RC22273 (W0112-016) December 4, 2001
Computer Science

IBM Research Report

Transforming Surfaces into Touch-Screens

Claudio S. Pinhanez, Frederik C. Kjeldsen, Anthony Levas, Gopal S. Pingali,
Jacob Hartman, Mark E. Podlaseck, Vivek Kwatra,
Paul B. Chou
IBM Research Division
Thomas J. Watson Research Center
P.O. Box 218
Yorktown Heights, NY 10598

Preliminary version, submitted to CHI'02.

Transforming Surfaces Into Touch-Screens

Claudio Pinhanez, Rick Kjeldsen, Anthony Levas, Gopal Pingali,
Jacob Hartman, Vivek Kwatra, Mark Podlaseck, Paul Chou
IBM T.J. Watson Research Center
P.O. Box 218, Yorktown Heights, NY 10598 USA, +1 914 945 3251
pinhanez | fcmk | levas | gpingali | jhartman | vkwatra | podlasec | pchou@us.ibm.com

ABSTRACT

Can ordinary surfaces such as desks and walls behave like touch-screens, providing us information that we can interact with? We present a new device called the *Everywhere Displays projector* (*ED-projector*) that combines projection and vision technologies to bestow interactivity on any surface. The ED-projector acts as a pervasive display by directing light from an LCD projector onto different surfaces using a steerable mirror while correcting for distortion caused by oblique projection. It also acts as an input device by recognizing gestures on the projected surface with a pan/tilt camera. Based on an experiment involving hundreds of novice users in an augmented reality assembly task, we find that users readily relate to the idea of projected touch-screens. The experiment also reveals several interesting phenomena in the way users interact with projected displays and suggests directions for future research.

Keywords

Interaction technology, display devices, ubiquitous computing, gesture interfaces, augmented reality, computer vision, computer graphics, projection systems.

INTRODUCTION

The use of projection systems to augment reality has been demonstrated by researchers in many different situations [7, 8, 11, 13, 15, 16]. However, these systems were constrained to a fixed projector that could only project information on a limited area of an environment.

This paper introduces a new display device that overcomes this limitation by employing a rotating mirror to steer the light of the projector onto different surfaces. Unlike the above-mentioned systems, the *Everywhere Displays projector*, or *ED-projector*, can be easily reconfigured to create displays on any surface in a room.

Projected images can become interactive by using a vision system to detect the users' hand gestures [2], moving objects [13], or the users' body position [6]. The ED-projector employs a computer vision system with a pan/tilt video camera. The vision system is programmed to track hands and to recognize gestures similar to those made when people press or touch a button. This results in interaction capabilities similar to those found in touch-screen displays: a graphical output device coupled with low-resolution pointing input.

The key advantage of the ED-projector is that it can simultaneously steer the mirror and the camera around the environment. Using this capability, it is possible to bring computer access to where the user is located; to augment reality without requiring users to wear any device; and to enrich everyday objects with information without wiring them.

This paper starts with a basic technical description of the projection and vision components of the ED-projector. Following, we describe an experiment at SIGGRAPH'01 where users were confronted for the first time with the concept of surfaces transformed into touch-screens. Observations and results from this experiment are then presented and discussed, and their influence on our future research on Everywhere Displays is examined. We also briefly elaborate on possible applications of ED-projector-like devices.

THE EVERYWHERE DISPLAYS PROJECTOR

The *Everywhere Displays projector* is composed of an LCD projector, a computer-controlled pan/tilt mirror, and a pan/tilt/zoom camera. The projector is connected to the display output of a host computer that also controls the

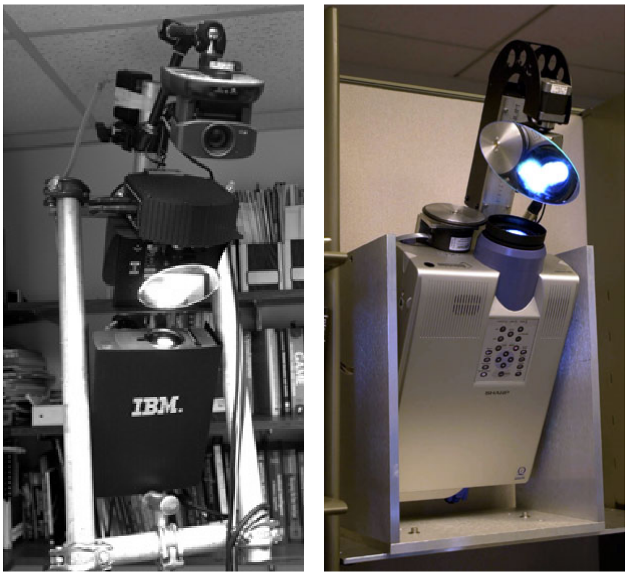

Fig. 1. Two prototypes of the Everywhere Displays projector.

mirror and the camera. Figure 1 shows two prototypes of ED-projectors built with off-the-shelf components — rotating mirrors used in theatrical/discotheque lighting, steerable cameras and LCD projectors.

In both prototypes, the projector's light can be directed anywhere within a cone defined by approximately 60 degrees in the vertical axis and 230 degrees in the horizontal axis. Due to hardware constraints in the mirror and the projector controls, changing surfaces currently takes 2 to 9 seconds.

Brightness and Contrast

It is possible to project images even in a brightly-lit room because the human vision system perceives brightness and contrast locally. Consider a white wall in an environment with normal lighting: if no image were projected, subjects would describe the brightness of the wall as "white." However, if a white and black pattern with sufficient brightness is projected on the same wall, viewers perceive the white area of the projected pattern as "white" and any neighboring black area of the pattern, receiving only the ambient light, as "black" [4].

Our first prototype employs a 1200 lumens LCD projector. Since in typical home and office lighting conditions a white pattern projected by our prototype is approximately 10 times brighter than its surroundings and, the projector is able create the illusion of contrast. The second prototype we built employs a 3000 lumens LCD projector, enabling even sharper contrast and better image quality.

Correcting for Oblique Distortion

When projection is not orthogonal to the projected surface, the projected image is distorted. In fact, most times the ED-projector is used to create displays on precisely such oblique surfaces. To correct the distortions caused by oblique projection and by the shape of the projected surface (if not flat), the image to be projected must be inversely distorted prior to projection.

Given a surface in the real world and an image to be projected onto it, the goal of the distortion correction process is to find a warping of the image that makes its projection appear without distortions to an observer orthogonal to the real-world surface (see Fig. 2). In general, this process is non-linear and involves the selective compression and expansion of different areas of the original image.

We have developed two simple methods to speed up this process that use computer graphics hardware present now in virtually every computer. The first method relies on the fact that, geometrically speaking, cameras and projectors with the same focal length are identical (as observed in [8, 10, 12]). Therefore, to project an image obliquely without distortions, it is sufficient to simulate the inverse process —i.e., viewing the projected image with a camera— in a virtual 3D computer graphics world that models the real world geometry (see [9] for details). This "viewed image" is inversely transformed and when projected onto the real world surface appears free of distortion.

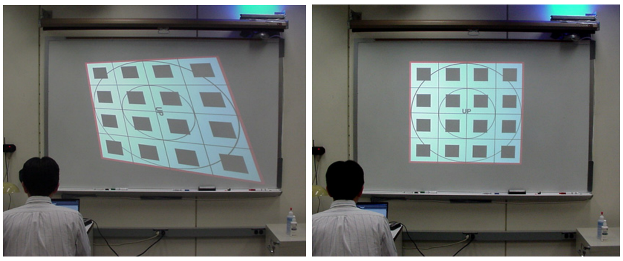
Fig. 2. Correction of oblique distortions: (left) the typical result of the oblique projection of a pattern on a surface; (right) the projection of a properly warped pattern creates a projected image free of distortion.

In the case of planar surfaces, we can also employ a second method for correction of distortion based on a 2D *homography* [3] — a 3x3 homogeneous projective matrix that represents the relationship between the projector and viewer's points of view (see [9]).

For both methods, it is necessary to individually calibrate each projected surface. When using the first distortion-correction method, the calibration parameters are determined manually by simply projecting a calibration pattern and interactively adjusting the scale, rotation, and position of a virtual surface in the 3D world, and the "lens angle" of the 3D virtual camera. This process typically takes between 10 to 20 minutes.

In the case of homography-based correction of distortion, calibration is obtained simply by finding the appropriate position of four points on the projected image (using similar algorithms to [8, 12]). This process normally takes 1 to 2 minutes for each surface. We are also investigating completely automated calibration procedures using techniques similar to [18].

Display Resolution

One problem with these distortion-correcting techniques is that they project displays with resolution that are smaller than the projector's resolution. The distortion-correction process normally fits an irregular quadrangular into the 4:3 viewing area of typical displays. A considerable amount of display area, and thus resolution, is lost in the process.

We have employed standard 1024x768 XVGA projectors in our prototypes. However, due to the loss of display area created by the distortion correction process, the final display resolution corresponds approximately to VGA, i.e., 640x480 pixels.

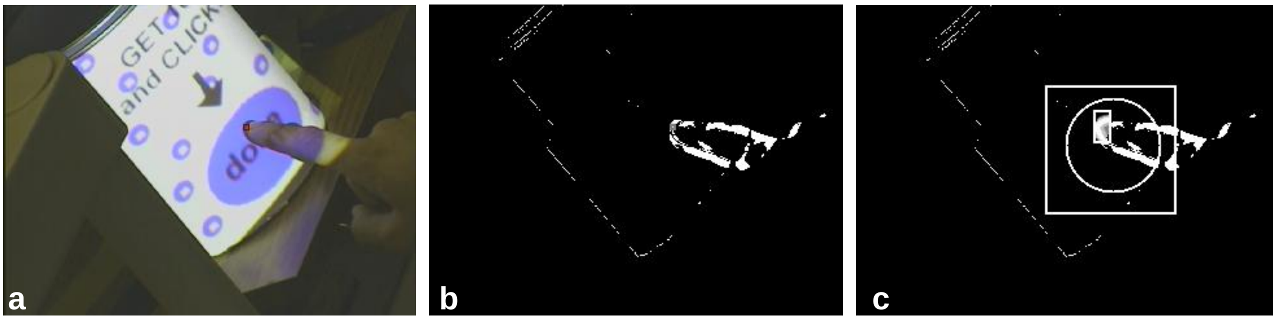

Fig. 3. a) Camera view of an interaction with a bucket; b) image difference data; c) overlay of search region (square), button active area (circle), and the fingertip template shown at the pointing location.

DETECTION OF USER INPUT

To detect user interaction with the projected display we use a single pan/tilt/zoom camera that is steered to follow the projected image. A vision system examines the video stream from this camera for user actions and generates events for the application software. The camera is located adjacent to the mirror assembly so that the camera view of the user's hand interacting with the projected image is occluded only if the user also occludes the projected image, a situation that is apparent to the users and which they can quickly correct.

There are two functional goals for the vision system: to detect when the user touches a projected "button"; and to track the location where the user is pointing on a projected "screen". Detecting these types of activities in the context of an ED-projector presents several challenges, from a computer vision standpoint. For one the appearance of the user's hand changes drastically as it moves through the projection. This makes techniques based on color or appearance unusable. Similarly, techniques based on background subtraction often give unreliable results as the projected image can completely overwhelm the inherent surface color of the background. Hence, the gesture recognition techniques used here are different from those not involving projected displays [17].

The vision system performs three basic steps: detecting when the user is pointing; tracking where the user is pointing; and detecting salient events such as a button touch from the pointing trajectory.

Pointing Detection

Even though the appearance of an object will change as it moves across a projected image, it will create a region of changed pixels that retains the basic shape of the moving object (see Figs. 3a and 3b). To find pointing fingertips we subtract each video frame from the frame before it, removing noise with simple computational morphology, and then convolving a fingertip template (Fig. 4) over the difference image using a custom matching function. If the template does not match well anywhere in the image, we assume the user is not pointing.

This approach supports a tracking rate between 5 and 30 frames/second on a 500 MHz workstation for a 320x240 image, depending on the size of the search region and the size of the fingertip template (as determined by the expected size of the user's hand).

Point Location

The fingertip template is deliberately kept short so that it will match fingertips that extend only slightly beyond their neighbors and will match fingertips within a wider range of angles. As a result the template often matches well at several points in the image. We resolve between these hypotheses by using the "fingertip" furthest from the user.

This approach requires knowledge of where the user is located. Rather than trying to infer this information from the shape and motion of the user's hand, we obtain it during calibration, assuming that the user will approach the interaction area from a consistent location each time.

Button Touch Detection

A button touch event is defined to occur when a fingertip is in the vicinity of a button, travels away from the user to a point within the button and then returns toward the user. Button touches are detected by examining the hand trajectory for several specific patterns that indicate this type of motion. If more than one button appears at a given moment, as with the "selection" screen described later, and more than one seem to be touched, only the button furthest from the user generates a "touch" event.

This algorithm works very well on interactions where the user is asked to touch one button at a time. Importantly, it also resists generating an event when the user's hand moves through a button on the way to or from another location. The algorithm can fail when the user touches several buttons without retracting their hand, or "flies" their finger around in the image before touching a button.

Notice that although our vision system was built to recognize the touching of a button, the image of a button tends to elicit from the user a slightly different gesture, the pressing of the button. We will address the consequences of this difference later.

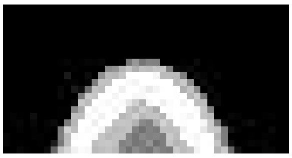

Fig. 4. Fingertip template.

Fig. 5. A picture made of M&Ms.

Fig. 6. Space used in the augmented assembly demo at SIGGRAPH'01.

Calibration

This algorithm requires prior information for each surface the user will interact with about the expected location of the user and size of their hand, the location of buttons in the image, and if tracking is to be used, the corners of the projected screen image. We supply this in a few calibration steps by sizing and rotating a hand icon to match the image of the user's hand, clicking on the corners of the projected "screen" and drawing shapes over the buttons. A search region is also identified for each surface in order to help the system ignore extraneous movement in the image and speed response (see Fig. 3c).

AUGMENTING AN ASSEMBLY TASK

Among the various applications envisioned for the ED-projector, we are particularly interested in augmented reality (AR) scenarios. Unlike traditional AR, the use of ED-projectors enable user visualization of information without encumbering goggles and interaction without contact with input devices. A typical AR application is the augmentation of an assembly task. In this case, assembly instructions are superimposed on the objects to be assembled. Each individual user contributes to the assembly of the whole system by executing specific parts of the assembly process, many times without knowledge of the complete structure of the assembled object.

Based on these premises, we conceived and implemented an augmented reality assembly environment to be experienced by the visitors of a computer graphics conference, SIGGRAPH'01. The goals of this experiment were twofold. First, we wanted to create an experience where people could be introduced to the ED-projector concept and technology. Second, and most important, we wanted to observe users interacting for the first time with surfaces transformed into touch screens.

Given the characteristics and interests of our audience, we chose the object to be assembled to be a picture made of M&M®[1]'s (multi-colored sugar-coated chocolates) where each M&M is regarded as a "pixel" of the picture. Figure 5, shows an example of an M&M picture. Our experiment built pictures of 60x50 M&M "pixels" based on portraits painted by Van Gogh.

The Space

Figure 6 shows the demo space constructed at SIGGRAPH'01. In the front of the space is a table; on the table lies a flat transparent Plexiglas board coated with double-sided transparent stick tape. This board is used as a mounting "canvas" for the M&M's and after assembly, as the hanging support for the pictures. Under the Plexiglas board, we placed a white sheet of paper where the positions (but not the colors) of the M&M's are printed. Such marks are visible through the transparent Plexiglas and the double-sided tape and are used to aid the M&M placement process.

On the back walls of the space we shelved unlabeled buckets, each containing one of the 18 different colors of M&M's used. Each bucket is covered with paper of different color and texture. Two other surfaces are also used for projection: a painter's palette mounted on an easel (see right of Fig. 6) and a wood board first covered with 1 inch-thick foam and then topped with white fabric (shown in the extreme left of Fig. 6). On the surrounding walls are completed M&M pictures, giving users a notion of the final goal of the assembly task. The ED-projector, not visible in Fig. 6, is positioned directly above the Plexiglas board.

The Experience

As the visitor arrives in the space, she sees the image of an M&M projected on the painter's palette with the instruction "Touch to begin", as shown in Fig. 7a (see also the video figure). The goal of this image is simply to familiarize the user with the idea of touching a surface in order to interact with the system.

---

[1] "M&M" is a registered trademark of Mars, Inc..

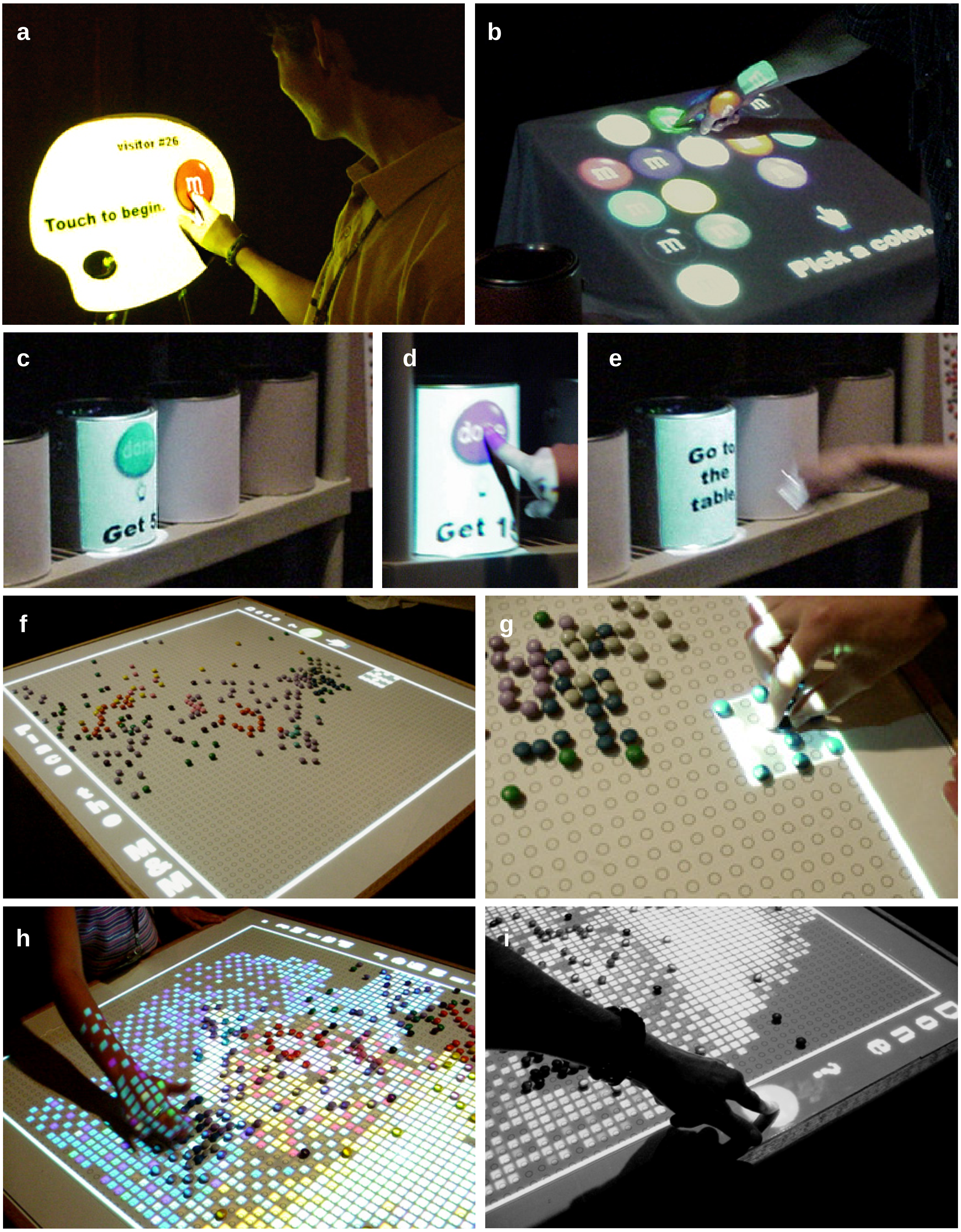

Fig. 7. Steps followed by each user when placing the M&M's in the SIGGRAPH'01 demo: a) the entrance panel is touched to start the process; b) a color is selected; c) the bucket containing the selected color is projected upon; d) the "done" button bucket is "clicked" after the M&M's are picked up; e) an instruction to go to the table is projected on the bucket; f) the positions where the M&M's are to be placed is shown; g) the user places the M&M's on the sticky surface; h) finger-painting to reveal the full image; and i) touching a "yes" button on the board after all interaction is done.

When the visitor touches the M&M "button," the system reacts by projecting an arrow onto the palette and a message inviting the visitor to "Come in." The ED-projector is then steered towards the foam-covered-board where images of different colored M&Ms are projected along with the invitation "Pick a color" (see fig. 7b). Here, the user can select which color to place in this iteration by simply making a "clicking" gesture on the M&M of the desired color.

As soon as the system detects the user selecting the color she wants to place, it projects a message on the foam covered-board, asking the user to get a certain number of M&M's from a highlighted can. The ED-projector then rotates the mirror so it highlights exactly the bucket that contains the M&Ms of that color. It displays the message "Get *n*", (where *n* is the number of M&Ms to be picked up from the bucket) on the bucket's frontal face. In addition, a button with the word "done" on it, *i.e.*, an image of an M&M underlined by a pointing finger (see Fig. 7c), is also displayed on the can. After retrieving the appropriate number of M&M's from inside the bucket, the visitor has to touch this "done" button to communicate that the picking has been completed (Fig. 7d). Immediately afterward an instruction to "Go to the table" is projected on the bucket (Fig. 7e) and the ED-projector is redirected to the Plexiglas board.

The ED-projector then indicates the precise location where the M&M's should be placed (Figs. 7f and 7g) by projecting an image of the target locations on the canvas. It also displays the instructions "Place the M&M's", and "Done ?" and an M&M button labeled "yes". After the user places the M&M's, she communicates the end of her action by touching this button (similar to Fig. 7i). At this point, the user is invited to "Finger paint", that is, to wave her hand over the board to reveal the completed picture. The camera system tracks the position of the user's hand and fills in circles of the appropriate color, interactively completing the picture (Fig. 7h). At any time the user can stop this activity by touching the "yes" button to the right of the image (Fig. 7i). The system times out after 40 seconds and shows the complete image to the user.

RESULTS FROM THE SIGGRAPH'01 EXPERIENCE

During the 6 days of the exhibition more than 650 people went through our demo, 4,000 were estimated to have watched, and 4 complete M&M pictures were composed. The imagery of the ED-projector camera was recorded into tape and the users' selections and the system's transitions to different surfaces were logged. However, we did not test competing ideas of configurations or interface designs. Also, since most of our users were confronting a novel interaction paradigm, we were not shy of giving them instruction as needed.

The objective of the SIGGRAPH experiment was to use the video and logged data, coupled with our own observations, to determine what are the main HCI issues associated with the usage of ED-projectors. In other words, the goal of this study is to use a large number of subjects to find the interesting questions for us to investigate in the future in more controlled studies.

Table 1. Performance of the vision system (130 users).

	success	false positives	false negatives
entrance	88%	9%	3%
select board	94%	2%	4%
buckets	55%	33%	12%
board (place)	91%	8%	2%
board (paint)	80%	2%	18%
total w/o buckets	89%	6%	5%
total	**81%**	**12%**	**7%**

A problem that we detected early on was the length of time required to move projections between surfaces. As mentioned before, the process of moving the mirror and/or adjusting focus and zoom of the projector takes between 2 and 9 seconds. Even the fastest transitions seem to be perceived as too long by the users of the system. The problem seems to be exacerbated by the lack of feedback on what to do next, and not knowing on which surface of the environment the display would appear next.

To minimize the perceived delay, the interaction was designed so that the user always received some instruction conveying what to do next before the time-consuming mirror movement. For instance, after the user touched the "done" button on a bucket, the instruction "Go to the table" was immediately projected (see Fig. 7c-e). In fact, we did not have this kind of messages on the first day of the exhibition and we observed greater anxiety, confusion and frustration among users. Even after the change, however, many users would not read the instructions. After having pressed a button, they would simply ask us "What do I do next?" People do not read instructions, no matter on which surface they are projected.

Although the vision system and the projection system were integrated for the first time only in the weeks preceding the exhibition, the combined system worked remarkably well. A sample with 130 users of the data collected from the camera videotapes with 621 button touch events (touching gestures or false detections) yielded correct detection of touching gestures up to 81%, with 12% of false negatives and 7% of false positives, as shown in table 1. If the buckets are excluded from the count, the performance exceeds 89%.

The buckets yielded a high number of errors for several reasons. The biggest problem with buckets was that some times the users partially or completely occluded the displayed button while picking M&M's from the bucket. At these times, the display was on the user's body, typically on their head and on their back. The user's motion in such situations often triggered false positives for the "done" button on the display.

Some of false negative errors were due to the fact that the system was tuned for a straight, forward, out-and-back button touch motion. In some cases, the error happened because the user assumed some that he was interacting with “clickable” buttons and not “touchable” triggers. For example, when picking a color, the user would fly his hand over different buttons before selecting one. As discussed before, this kind of user behavior sometimes triggers an incorrect touch event. Interestingly, while most users interacted using regular button presses. some also tried to wave their hands through the buttons or made similar non-touching actions. Finally, we credit the high number of false negatives during finger painting to the reduced frame rate when simultaneously tracking over a large area and detecting touch events. Occasionally the brief interval when the finger was inside the button would fall between video frames.

Either when trying to click a button on the hard surface of the Plexiglas board or the soft foam board, our users almost always applied pressure to the “buttons.” Also, whenever a button touch was not detected by the vision system, the most common user’s reaction was to try to press again with more pressure and, if still not successful, using the whole hand instead of a finger. Since the users seem to apply excessive pressure on the first button (mounted on a fragile easel), we decided to change the wording of the instruction from “Press to begin” to “Touch to begin” (Fig. 7a). Although we got gentler touching on this first button, it seems that the overall amount of pressure remained high for all surfaces and especially for the foam-covered selection board. Perhaps people expect buttons, even virtual buttons, to react to pressure and not to touch, and to fail if enough pressure is not applied on them.

Our users also looked less afraid of interacting with the painter’s palette in the entrance and the Plexiglas board on the table than on the fabric-covered soft board and, especially, on the buckets. In general we observed no hesitation to interact with the entrance board (Fig. 7a), followed by a little more of concern about what to do to select a color on the selection board (Fig. 7b). However, when they went to a bucket to pick up M&M’s, many of them did not realize that they had to click the big “done” button on it. We expected that, since this was the third time in the demo that this kind of interaction was happening, the users’ would naturally press the button on the bucket. Most of the time, this was not the case.

To help the users in that situation we initially tried to give verbal instructions such as “Click the ‘done’ button.” After some time, we found that the most effective way was to tell them to “Click the bucket.” This instruction seems to make clearer that there is, in fact, a button on the bucket. Interestingly, as in the case of the entrance board, users did not have any problem with the idea of clicking the buttons projected on the Plexiglas board, both when placing the M&M’s and when finger painting the picture.

Is clicking a bucket a more difficult concept to grasp then clicking a wood board? Are buckets expected to be less interactive than tables? Of course the design of our system and the simplicity of our experiment allow us neither to state that the phenomenon does happen nor to draw definitive conclusions about its causes. However, our experience at SIGGRAPH suggests an interesting hypothesis: people seem to attribute different interactive capabilities to surfaces transformed into touch-screens according to the nature of the surfaces themselves.

We also observed that our users quite naturally accepted the idea of augmenting reality by directly projecting information onto objects. As shown in Figs. 7f and 7g, the placement of the M&M’s was guided by projecting a circle of light in the appropriate position. Very few users had trouble in understanding this instruction and in accomplishing the task. Moreover, we saw that they naturally changed their body’s position to minimize the shadow of their arms and body over the board. People seem not only to accept augmentation naturally but also to compensate for its limitations.

For practical reasons, the demo was designed so we could have not only users but also observers, i.e., people watching the demo. A major advantage of projection over goggles as a means for augmented reality is precisely this “social” nature of projection. While goggles restrict the visualization to the people wearing them, with ED-projectors it is easy to create a shared experience. We observed the value of sharing in different situations of our demo, but especially when users learned the interaction process by watching other users; and when observers helped users, suggesting what to do next.

FUTURE APPLICATIONS OF ED-PROJECTORS

We regard the ED-projector as a generic input/output device that can replace, in many situations, current displays and interactive devices. Besides, this technology also enables a new set of applications with its ability to steer information and interfaces around an environment.

A first class of applications relates to the creation of computer desktop-like interactive displays on non-tethered surfaces. For example, in an office, a desktop application can be projected directly on any surface, as previously suggested by Wellner [16]. Unlike the interactive whiteboard described in [2], the projected application can be easily moved around the room; for instance, from a whiteboard on the wall to the top of a desk for more detailed reading. Similarly, the ED-projector can be used to bring information to the physical location where it is used or needed. For example, a database application managing reports can be projected on top of the file cabinet with hard copies of the reports.

A second class of applications deals with enabling a computer to act on the physical world, almost like a robotic arm made of light. These applications can use the ED-projector to point to physical objects, show connections among them, and project patterns to indicate movement or change in the real world. The SIGGRAPH experience described here is a typical example of this class of applications.

The ED-projector can also be used to provide computer and information access in spaces where traditional displays can be broken or stolen, or create hazardous conditions, such as in public spaces and areas subject to harsh environmental conditions. The device also permits an interactive display to be brought to the proximity of a user without requiring the user to move. In particular, the ED-projector can facilitate the access and use of computers by people with locomotive disabilities. For instance, it can project an interactive display on a hospital bed sheet without patient contact with any device.

We also see the Everywhere Displays projector as a potential enabler of a new generation of games that happen not in the virtual world but in the physical, everyday world where we live. Unlike games based on phones and portable computers [1], the use of the ED-projector provides high-resolution displays where characters and fantastic objects can move to different surfaces, creating a game that surrounds the players. Using video input, various user actions can be detected depending on the game needs, for example, hand gestures, body movements, foot action, and/or facial expressions. Above all, a single ED-projector installed on the ceiling of a living room (or arcade environment), just using software commands, can completely reconfigure the space to create different games according to the interests, age, and motor skills of the players.

FINAL REMARKS

The ED-projector introduces a new concept: transforming ordinary surfaces into touch-screens using steerable projection and vision. The viability of this concept was clearly demonstrated in the sizeable SIGGRAPH'01 experiment, which was our first attempt to combine projection and vision in the ED-projector.

As mentioned before, the SIGGRAPH experiment is too simple to enable us to draw definitive conclusions about how people react to a system that transforms everyday surfaces into touch-screens. However, it has alerted us about some limitations of the technology and has suggested guidelines for the design of applications.

It is now evident that the switching time between surfaces has to be either shortened or filled with some sort of user feedback. The analysis of our errors in detecting user input points out many ways we can improve the performance of the vision system. It has become clear that, like in many other computer vision systems, domain knowledge is essential to obtaining good performance. This might include knowledge of where the user is located during the interaction, what types of activities she is likely to perform outside the context of the interaction, and what types of errors are more or less disruptive to the interaction. All these are issues that have to be addressed by our future research.

We believe that some of the most interesting research is, in fact, connected to how people will interact with a projected surface. Given our observations about the way users interacted, is the "button" the right metaphor for interaction? As we demonstrated, we can detect button-pressing actions with reasonable accuracy, even with a single camera, but is this the right paradigm from the user's point of view for a surface that detects mainly the hand position? How can more flexible and expressive hand gestures play a role in these interfaces?

Similarly, the difficulty experienced by our users to realize that a bucket can be interactive also poses interesting questions. In particular, we are interested in investigating the extent to which the object where the interface is projected on determines the design of the interaction occurring on it. Beyond ED-projectors, this is a question that can be relevant for the whole research on tangible interfaces [5] and ubiquitous computing [14].

REFERENCES

1. Björk, S., et al. Pirates! - Using the Physical World as a Game Board. In Proc. of Interact'01. 2001. Tokyo, Japan.
2. Crowley, J.L., et al., Things that See. Communications of the ACM, 2000. 43(3): p. 54-64.
3. Faugeras, O., Three-Dimensional Computer Vision: A Geometric Viewpoint. 1993, Cambridge, Massachusetts: The MIT Press.
4. Hoffman, D., Visual Intelligence: How We Create What We See. 1998: W. W. Norton.
5. Ishii, H. and B. Ullmer. Tangible Bits: Towards Seamless Interfaces between People, Bits, and Atoms. In Proc. of CHI'97. 1997. Atlanta, Georgia.
6. Keays, B. and R. Macneil. metaField Maze. In Proc. of SIGGRAPH'99. 1999. Los Angeles, California.
7. Krueger, M.W., Artificial Reality II. 1990: Addison-Wesley.
8. Morishima, S., et al. HyperMask: Talking Head Projected Onto Real Objects. In Proc. of Multimedia Modeling (MMM'00). 2000: World Scientific.
9. Pinhanez, C. The Everywhere Displays Projector: A Device to Create Ubiquitous Graphical Interfaces. In Proc. of Ubicomp'01. 2001. Atlanta, Georgia.
10. Raskar, R. Oblique Projector Rendering on Planar Surfaces for a Tracked User. In Proc. of SIGGRAPH'99. 1999. Los Angeles, California.
11. Rekimoto, J. A Multiple Device Approach for Supporting Whiteboard-based Interactions. In Proc. of CHI'98. 1998. Los Angeles, CA.
12. Sukthankar, R., et al. Smarter Presentations: Exploiting Homography in Camera-Projector Systems. In Proc. of ICCV'01. 2001. Vancouver, Canada.
13. Underkoffler, J., et al. Emancipated Pixels: Real-World Graphics in the Luminous Room. In Proc. of SIGGRAPH'99. 1999. Los Angeles, CA.
14. Weiser, M., The Computer for the Twenty-First Century. Scientific American, 1991: p. 94-100.
15. Welch, G., et al., Projected Imagery in Your "Office of the Future". IEEE Computer Graphics and Applications, 2000(July/August): p. 62-67.
16. Wellner, P., Interacting with Paper on the DigitalDesk. Communications of the ACM, 1993. 36(7).
17. Wu, Y. and T. Huang, Vision-Based Gesture Recognition: A Review. Lecture Notes in Artificial Intelligence, 1999. 1739.
18. Yang, R. and G. Welch. Automatic and Continuous Projector Display Surface Calibration Using Every-Day Imagery. In Proc. of 9th International Conf. in Central Europe in Computer Graphics, Visualization, and Computer Vision. 2001. Plzen, Czech Republic.