Barehands:
Implement-Free Interaction with a Wall-Mounted Display

Meredith Ringel
Computer Science Department
Brown University
Providence, RI 02912
mringel@cs.brown.edu

Henry Berg, Yuhui Jin, Terry Winograd
Computer Science Department
Stanford University
Stanford, CA 94305-9035
{hgberg, yhjin, winograd}@cs.stanford.edu

ABSTRACT
We describe Barehands, a free-handed interaction technique, in which the user can control the invocation of system commands and tools on a touch screen by touching it with distinct hand postures. Using behind-screen infrared (IR) illumination and a video camera with an IR filter, we enable a back-projected SMARTBoard (a commercially available, 61 3/8'' x 47'' touch-sensing display) to identify and respond to several distinct hand postures. Barehands provides a natural, quick, implement-free method of interacting with large, wall-mounted interactive surfaces.

Keywords
Interaction technique, user interface, hand posture, infrared, image processing, touch interaction.

INTRODUCTION
As part of our project to develop a pervasive computing environment, we have created an *interactive workspace* which integrates a variety of devices, including laptops, PDAs, and large displays, both vertical (wall-mounted) and horizontal (tabletop). A key component of our space is a wall with three adjacent back-projected SMARTBoards [2]: commercially available, wall-mounted, touch-sensitive displays. We normally display a standard Windows 2000 desktop on each of these boards or one combined desktop covering all three.

In our experience over a year of using the boards, we recognized the strong appeal of direct hands-on manipulation without implements. Although the SMARTBoard provides a set of whiteboard tools, users gravitated toward performing simple interactions by touching a finger to the board. In fact, they often tried fruitlessly to do the same on the adjacent front-projected wall, which had no touch sensors. We therefore decided to explore the possibilities for bare-hands mechanisms that could provide appropriate interaction, and that did not require the user to hold separate implements.

Any touch on the SMARTBoard touch screen is interpreted by the standard software as a left mouse click. Simultaneous touch at multiple points is interpreted as a single mouse click with coordinates corresponding roughly to the center of the contact points. Right-button click is implemented by pressing a physical button on a tray at the bottom, right-hand edge of the SMARTBoard. The next touch is then interpreted as a right mouse click. There is no way to indicate a position without a button click (as is used for affordances such as rollover tool-tips).

The SMARTBoard and its accompanying software also provide electronic whiteboard capabilities using separate physical devices modeled after conventional whiteboard marker pens and erasers. There are a number of problems with these tools, and we wanted to see what could be done using a tool-free hands-only mechanism.

Our research goal is to explore the space of hand postures and gestures to identify which ones are best suited to the different aspects of overface and interface.

BAREHANDS IMPLEMENTATION
In order to obtain images of a user's hand as it touches the display, we use a technique similar to that of the HoloWall [1]. We illuminate the rear of the screen with a pair of IR LED arrays, approximately 1.5m behind the board, each containing twelve rows of eight 16mw/str SLI-0308CP LEDs, pulsed at 200mA. Centered between the light sources approx 3m from the board, is a Marshall V-1070 video camera (resolution 811 x 510 pixels) with an IR filter, and an 8-80 mm motorized lens, specially adjusted to focus with IR light. A standard video digitizing card is used to capture the image from the camera.

IR light from the LED arrays is reflected off the rear of the screen and picked up by the camera. When a hand touches the front of the display, it reflects additional IR light, and is perceived by the camera as a region of increased intensity.

When the user touches the display, we first identify which pixels in the video image correspond to his or her hand by examining the pixels' grayscale intensity, and filtering out random noise.

We take advantage of the touch sensing of the SMARTBoard to reduce image-processing overhead. Although we receive a continual input from the video camera, we only analyze the frames of the video that occur immediately after we receive a "contact down" event from the SMARTBoard driver. This contact event is intercepted before it can be sent to Windows as a left mouse click event (the SMARTBoard software's default response to touch).

We are able to avoid processing the entire image by using the x- and y-coordinates associated with the contact-down event as a seed for region growing. Pixels within a fixed radius of the seed are examined and those with a grayscale intensity similar to that of the seed are included in the region. This process is repeated by exploring a fixed radius around each new pixel found to be in or near the region of contact. This allows for the detection of contact regions with or without discontinuities (a two-fingered contact, with the fingers spaced slightly apart, would result in a discontinuous region). Isolated pixels are eliminated, even if their intensity is above threshold.

Region characteristics such as height-to-width ratio, ratio of a region's area to the area of its bounding box, ratio of perimeter to area, and the presence or absence of gaps in the region are used to classify the image as either one finger, two fingers, a vertical edge, a horizontal edge, a palm, or unknown.

Once the activated region has been classified as a particular hand posture, an appropriate action is produced, based on the mapping currently associated with that posture.

A simple interface is provided to allow users to set up the mapping of any of the hand postures onto any of a number of events, including mouse events (with various buttons), SMARTBoard whiteboard tools (such as highlight and erase) and standard Windows shortcuts (such as cut and paste). If the posture is mapped to a whiteboard tool, we send messages to the SMARTBoard driver to change the current tool. If the posture is mapped to a Windows command, we send an appropriate combination of keyboard and/or mouse events.

EVALUATION

Informal initial observations of Barehands use affirm that our system provides a means of interacting with a large display that avoids the inconvenience of walking to a different area of the device to push a button or of grasping a physical interaction tool. By mapping gestures to commands such as copy and paste the user saves time and motions traditionally required for selecting those commands from menus.

System response time is conducive to real-time interaction: image-processing time averaged 13.37 ms for a mixed-posture set of 120 touches. Average posture-classification accuracy rates are above 90%; we expect that more uniform IR illumination in a remodeled configuration will allow for improved accuracy (see *Future Work*). Since there is immediate visual feedback on the posture that was recognized and opportunity for correction, the system should be useable without achieving 100% accuracy, but it will take further experimentation to determine the best tradeoff between accuracy and the number of different postures recognized.

DISCUSSION

The key assumption underlying Barehands is that for certain classes of users and interactions, bare-handed interaction is better than using physical tools to determine the mapping from touch to action; One of our long-term research questions is understanding the criteria that make tool-based and bare-hand interaction appropriate for different kinds of situations and actions.

Remodeling of our workspace (which took place from August through November) necessitated dismantling of the Barehands setup shortly after we completed developing the system. Now that remodeling is complete, however, we are planning user tests that will compare our technique with alternative methods of working with a large screen.

Future Work

Lighting arrangements have been the greatest technical challenge in the implementation of Barehands. Employing two arrays of IR LEDs provides non-uniform lighting; the uneven light results in lower recognition accuracy for hand postures outside of the display's central area.

We plan to experiment with new arrangements of IR light sources to achieve more uniform illumination. Since each source is composed of many LEDs, they can be arranged into other geometries, such as linear strings rather than concentrated arrays. We will also experiment with using alternating lighting directions to improve recognition. We are planning to experiment with lighting schemes where successive frames of video are lit in different ways and from different directions, enabling more sophisticated image analysis.

More uniform lighting will facilitate more accurate image processing, which will enable us to expand our current set of five recognized postures. Although the space of recognizable postures is large, many of the theoretically possible ones will be awkward or uncomfortable. We plan to conduct user experiments, in which we will look for postures that are ergonomically appropriate and that have natural mappings to actions.

REFERENCES

1. Nobuyuki Matsushita and Jun Rekimoto, "HoloWall: Designing a Finger, Hand, Body, and Object Sensitive Wall", Proceedings of UIST '97, 1997.

2. See http://www.smarttech.com.