MULTI-TOUCH: A NEW TACTILE 2-D GESTURE INTERFACE FOR HUMAN-COMPUTER INTERACTION

Wayne Westerman, John G. Elias
Department of Electrical and Computer Engineering, University of Delaware
140 Evans Hall, Newark, DE 19716-3130.
and
Alan Hedge
Department of Design and Environmental Analysis
Cornell University, College of Human Ecology, MVR Hall, Ithaca, NY 14853-4401.

The naturalness and variety of a touch-based hand gesture interface offers new opportunities for human-computer interaction. Using a new type of capacitive sensor array, a Multi-Touch Surface (MTS) can be created that is not limited in size, that can be presented in many configurations, that is robust under a variety of environmental operating conditions, and that is very thin. Typing and gesture recognition built into the Multi-Touch Surface allow users to type and perform bilateral gestures on the same surface area and in a smaller footprint than is required by current keyboard and mouse technologies. The present approach interprets asynchronous touches on the surface as conventional single-finger typing, while motions initiated by chords are interpreted as pointing, clicking, gesture commands, or hand resting. This approach requires learning only a few new chords for graphical manipulation, rather than a vocabulary of new chords for typing the whole alphabet. Graphical manipulation seems a better use of chords in today's computing environment.

INTRODUCTION

Hand gestures play an integral role in human communication. Gestures are a physical expression of mental concepts (Huang et al., 1996). As Pavlovic et al. (1997) emphasize, when compared with current interfaces, the naturalness and variety of hand gestures offers unique opportunities for new forms of human computer interaction (HCI).

Most personal computers require the operation of at least two devices: a keyboard for bilateral text and numeric input; and a pointing device such as a mouse, trackball, trackpoint, touchpad, pen, joystick, etc. In terms of human performance, the typical desktop computer hardware interface arrangement is inefficient for several reasons:

a) The operation of the keyboard and pointing device requires non-productive movement times to switch a hand between typing and pointing activities.
b) The use of separate devices requires different skill sets: typing movements for keyboards, mostly point-and-click operations for other input devices.
c) With the exception of the keyboard, most input devices can only be operated with one hand.
d) The physical arrangement of a keyboard and a mouse or other input device can be a risk factor in the development of musculoskeletal disorders (e.g. National Academy of Sciences, 2000).
e) The physical area required for a keyboard and a mouse or other input device is an inefficient use of workspace.

Consequently, several kinds of gestural interface have been developed. This paper will describe the development and operation of a unique interface for multi-finger gestures performed on a surface.

Taxonomy of Hand Gestures

As used in HCI, the term 'gesture' is defined as a trajectory in a parameter space over a suitably defined interval, and a taxonomy of hand gestures has been developed (Pavlovic et al., 1997) as shown in Figure 1 (figure 1 is photo, not taxonomy!). In this taxonomy, hand movements are initially separated into unintentional movements and gestures (intentional movements are meant to communicate the user's intent). Gestures then are divided into manipulative or communicative forms. Examples of manipulative gestures are actions such as the movement or rotation of an object. Communicative gestures include the hand movements that are part of the nonverbal components of speech. Communicative gestures are divided into symbols and acts. Finally, symbols can be referential actions, such as those movements that are symbolic representations (e.g. the circular movement of fingers to symbolize opening an object) or modalizing actions, such as imitating a wave movement up and down. Acts can be either mimetic (i.e. movements that imitate an action) or deictic (acts of pointing). Temporal components and patterns of gestures are also important in the flow of communication. However, most current input devices restrict users to mainly deictic gestures and limit temporal components to sequences of keystrokes or clicks.

Research on Gestural Interfaces (GI)

Free-Space 3-D GIs

Most of the research on gestural interfaces has attempted to develop efficient and economical systems for capturing three-dimensional hand movements in free-space. Approaches have either used instrumented gloves or vision-based systems (Huang and Pavlovic, 1996). Glove systems tend to be restrictive since the gloves are wired to the computer, but these systems have been used in applications such as controlling 3-D molecular modeling software in molecular biology work (e.g. Pavlovic et al., 1996). Vision based approaches typically require the use of at least two video cameras and considerable computational power to process the resulting signals (e.g. Martin and Crowley., 1997). To date, the expense and complexity of these approaches has limited their usefulness. Distinguishing intentional gestures from unintentional postural adjustments is very difficult for free-space gesture recognition systems because surface contact and pressure, the simplest intention cues, are missing. Also, because of the lack of tactile feedback and a surface to support the hands, these approaches are less practical for typing input than a conventional keyboard.

Touch-Based 2-D GIs

Perhaps the best example of a widespread 2-D GI is the touchpad. This surface allows for deictic gestures and it has become an integral component in laptop computers and some keyboards, as well as being a standalone device. Some types of touchpads allow the capture of basic symbolic representations, such as capturing a signature, but typically these require an additional component, such as a stylus or special pen. Touchpads currently on the market use row and column electrodes that span the entire sensing area, which can be likened to projection scanning from the surface edges. Such projective scanning systems cannot distinguish all arrangements of multiple fingers, and are generally limited to tracking the centroid of all surface contacts (Lee 1984). Furthermore, capacitance-based projective systems do not scale up well because signal-to-noise ratios fall as the electrodes that must span the surface increase in length. Optical-based projective systems (Newcom, 1999) do not suffer this limitation.

Multi-Touch: a new approach to a 2-D GI

Multi-touch surface (MTS)

To overcome the sensing limitations of current touch surfaces and the typing limitations of free-space interfaces, a new technology, the Multi-Touch Surface (MTS) has been developed (Westerman, 1999; Westerman and Elias, 1999). Using scalable arrays of a new capacitive proximity sensor, the MTS is not limited in size and can be presented in many configurations. The use of different sensor densities allows the creation of surfaces with different resolutions. The Multi-Touch technology is robust under a variety of environmental operating conditions, and compared with a conventional keyboard, the surface depth can be very thin. A keyboard that uses the MTS allows users to type and perform bilateral gestures on the same surface area, and in a smaller footprint than is required by current keyboard and mouse technologies. Compared to free-space gesture systems, the MTS offers a mirror-like correspondence between surface contacts felt by the user and those detected by the proximity sensors. This correspondence inherently ignores many postural adjustments and ensures that the user is always aware of any touches (intentional or not) that the system may attempt to interpret.

In 1984, Lee at the University of Toronto first implemented a capacitive sensor array capable of tracking multiple touches simultaneously (Lee et al. 1985), but this early prototype suffered from interference between sensors. Also, relatively little computational power for interpreting gestures was available at the time, and graphical user interfaces were not yet mainstream in the PC world. The Toronto research group envisioned breaking surfaces up into multiple virtual windows wherein gestures within each window would have different effects (Brown et al. 1990; Buxton et al. 1985). In a different approach, McAvinney's SensorFrame (Rubine and McAvinney, 1990) recognized gestures involving up to three fingers that pass through a grid of light beams. This approach has recently been extended by Newcom, Inc. of Japan (Newcom, 1999), whose projective optical sensors mounted at the corners of a touchscreen can distinguish infrared reflective pen tips from fingertips, which absorb and block the infrared beams.

Previous attempts to implement both typing and pointing on the same region of a touch surface have assumed the user would press a button to switch modes (Boie et al. 1995). The present MTS implementation creates a seamless, unified interface in which mode or tool selection is implicit in hand configuration, not absolute placement of the hand in certain regions of the surface nor performance of an explicit mode-switching signal such as a button press. Unifying typing, manipulative gestures, command gestures, and handwriting requires a very carefully designed gesture set. Rather than apply the chord detection capability afforded by multi-touch sensing to a chord keying scheme, chords are reserved for hand resting, manipulative gestures, and command (symbolic) gestures. Traditional touch or hunt-and-peck typing can then easily be distinguished from chord gestures as long as finger touchdowns during typing are asynchronous, while finger touchdowns at the beginning of a chord are roughly simultaneous. The *initial* combination of touches in a manipulation chord selects the graphical manipulation 'channel' (i.e. point,drag, orscroll). This way users are free to relax their remaining fingers onto the surface while continuing the same manipulation. Hand configurations other than basic fingertip combinations can also be recognized, e.g. fists and pengrips.

Thus, with the present approach, asynchronous touches on a MTS are interpreted as conventional single-finger typing, while motions initiated by chords are interpreted as pointing, clicking, and other gesture commands. This approach is preferable because it requires learning only a few new chords for graphical manipulation rather than a vocabulary of new chords for typing the whole alphabet. Graphical manipulation seems a better use of chords in today's graphics intensive computing environment.

Vocabulary of 2-D gestures

By itself the MTS merely allows for the detection of finger and hand contacts on the surface. For this to be an effective GI, software has been developed that recognizes a number of 2-D gestures. For example, all current keyboard and 2-button and 3-button mouse operations are included in this vocabulary of 2-D gestures. The gestures that represent specific actions have been developed based upon both the naturalness of the finger movements, their conceptual associations, and results from user trials of the technology (Figure 1). In addition to the existing vocabulary, configurable gesture sets can be created, which allows users to create an interface that is naturally compatible with user expectations. The software is able to distinguish unintentional hand movements, such as resting the hand on the surface, from gestures.

Prototypes of the multi-touch GI have been tested with 5 subjects over a 3 months period in their normal work situations. Users report a fast learning curve for the new technology and they quickly master the gesture vocabulary. The design of the enclosure for a MTS can also improve the position of the hands during typing and the performance of input gestures. When users type on the MTS mere contact with the surface, not explicit finger pressure, is sufficient to activate keys. The lightness of the touch should benefit users experiencing tendonitis (one of the test subjects was experiencing such pain prior to using the multi-touch surface and this has been alleviated).

MTS Research

To date, two studies of a MTS have been conducted. The first study explored the use of a MTS keyboard with children doing fingerpainting and its use in a kindergarten classroom (Browne et al., 2000). They began by designing a finger painting program based on the input from six children between the ages of 6 and 11, and who came to the laboratory twice a week after school and work with adults to design and test the MTS technology. The final software design assigned colors based on the temporal order in which the 10 digits touched the multi-touch surface (Figure 2). Initially, the 10 colors, always assigned in the same order, but children could modify this by using menus to change color order for the 10 colors, brush type (5 choices), as well as clearing the screen. Changes in contact pressure changed the painting action, a light pressure painted the cursor mark being painted on the screen for a brief period of time before disappearing; a stronger pressure painted as usual. Informal testing was conducted with 7 children, 4 of whom had participated in pre-design questioning. The MTS was used along with the painting program by children in groups of 2 or 3 for about 10 minutes for each group. The groups received minimal instruction, mostly discovering how the MTS and software worked by exploration. Two adults observed the groups working. The contextual inquiry method was used to record time-synchronized notes from one observer recording the children's activities and the other recording their verbal utterances (Druin, 1999). Results showed that all of the children immediately liked and understood how to use the MTS. Most groups wanted to continue when their allotted time was up and four children specifically remarked that it was "cool". Although it was used by small groups, the MTS was shared remarkably well by the children, dividing the surface up equally according to sitting position and collaborating on the artwork. Working in pairs seemed to work best, whereas threesomes seemed cramped. Most of the children just scribbled with their fingers and enjoyed seeing the different colors fill the screen, though some children said that they didn't like how quickly the screen filled up. After this the children wrote and drew in journals about the experience. Many of the children drew both a touch-screen and a MTS in their illustrations. Interestingly, the MTS can be integrated with display-screen technology.

At the end of the study children worked on designing a "Classroom of the Future" for kindergarteners. Based on this, in August 2000 a Finger Painting Table, about 5 feet (153 cm) in diameter and covered it in 2-inch thick foam and incorporating MTS, was built . The table was positioned beneath a ceiling-mounted projector that projected the finger-painting screen image onto a white foam area of the table surface about 2 x 1 ½ feet (61x46 cm). The rest of the table was decorated with colorful cloth and fabric (Figure 3). Testing of this table by the children found that it provided a much softer, friendlier, and inviting painting environment, and it was more conducive to collaborative creations because the children stood and easily moved around the table. A combination of physical and virtual design tools using different colored foams were cut into shapes such as animals, people, houses, and clouds and placed on the projection surface so that children could paint around them using the MTS was immediately successful. This work determined that the MTS can readily be integrated into alternative uses for computer technology.

In a second study at the same HCI laboratory, the usability of an MTS keyboard (FingerBoard, see Figure 4) was compared to that of a standard keyboard and mouse setup using two identical tests involving basic typing and mousing activities. Fourteen college students (10 men, 4 women) were tested but the results were inconclusive in part because the learning curve for using the MTS exceeded the short test-time (30 minutes) available for become proficient at using the MTS. However, in this study keyboard layout of the MTS was confounded with the surface characteristics of the MTS, and the results obtained were similar to those for other ergonomic split-keyboards (e.g. kinesis), and this design change has been shown to decrease typing speed in the short-term (Chen et al., 1994).

CONCLUSIONS

Development of the MTS technology and supporting software allows the creation of an efficient touch-based GI for computers systems. A basic vocabulary of gestures has been created that can be expanded to include a greater range of operations than is currently available with input devices. The multi-touch 2-D GI may overcome limitations with existing 2-D interfaces, by harnessing many of the benefits of a 3-D interface through the use of more intuitive 2-D gestures performed on a surface. This approach offers several advantages over other attempts at creating effective GIs:

a) The use of a multi-touch surface affords some simple tactile and visual feedback to users
b) The use of a multi-touch surface allows users to rest their hands to relieve shoulder loads
c) The use of a multi-touch surface allows all keying and mousing activities to be performed using a single, space efficient unit
d) The use of a multi-touch surface provides users with a more intuitive interface that allows bilateral gestural controls and that speeds human performance
e) The use of a multi-touch surface is a very cost-effective approach to the creation of a new GI and it does not require the instrumentation of users hands or the use of twin video cameras and substantial computational operations.

In conclusion, the MTS represents a new direction in thinking about HCI and it holds the potential to create a new generation of more usable, more effective HCI products.

REFERENCES

Boie, R.A., Ruedisueli, L.W. and E.R.Wagner (1995) Computer Mouse or Keyboard Input Device Utilizing Capacitive Sensors. U.S. Patent 5463388, October 31.

Brown, E., Buxton, W.A.S. and K Murtagh (1990) Windows on Tablets as a Means of Achieving Virtual Input Devices, In D. Diaper et al. (Eds.) Human-Computer Interaction- INTERACT '90, Elsevier Science Publishers B.V., North Holland, 675-681.

Browne, H., Bederson, B., Druin, A., and L. Sherman (2000) Designing a Collaborative Finger Painting Application for Children, Human-Computer Interaction Laboratory, Institute for Advanced Computer Studies, University of Maryland, College Park, MD 20742 (unpublished report).

Buxton W., Hill, R. and P. Rowley (1985) Issues and Techniques in Touch-Sensitive Tablet Input (Extended Summary), Graphics Interface '85, Computer Systems Research Institute, Toronto, Ontario, Canada, 147-149.

Cakir, A.E., Cakir, G., Muller, T. and P. Unema (1995) The Trackpad - A Study on User Comfort and Performance, Proceedings of CHI '95, ACM.

Chen, C., Burastero, S., Tittiranonda, P., Hollerbach, K., Shih, M. and R. Denhoy (1994), Quantitative evaluation of 4 computer keyboards: wrist posture and typing performance, Proceedings of the Human Factors and Ergonomics Society 38th Annual Meeting, 2, HFES, Santa Monica, CA, 1094-1098.

Detwiler, B., Hun An, C. and P.Choi (2000)

The Battle of Input Devices: Examining the FingerWorks FingerBoard, Human-Computer Interaction Laboratory, Institute for Advanced Computer Studies, University of Maryland, College Park, MD 20742 (unpublished report)

Druin, A. (1999) Cooperative Inquiry: Developing New Technologies for Children with Children. Proceedings of CHI 99, 592-599.

Lee, S. A (1984) Fast Multiple-Touch-Sensitive Input Device, University of Toronto.

Lee, S.K., Buxton, W. and K.C. Smith, A (1985) Multi-Touch Three Dimensional Touch-Sensitive Tablet, Proceedings of CHI '85, ACM.

Martin, J. and J. L. Crowley (1997) An appearance-based approach to gesture-recognition. In Proceedings Int'l Conf. on Image Analysis and Processing, Florence, Italy, September.

National Academy of Sciences (2001) Musculoskeletal disorders and the workplace: low back and upper extremities. National Academy Press, Washington, DC.

Newcom, Inc. of Japan. http://newcom-uc.co.jp/ , 1999.

Pavlovic, V.I., Sharma, R. and T.S. Huang (1996) Gestural Interface to a Visual Computing Environment for Molecular Biologists, in Proceedings, International Conference on Automatic Face and Gesture Recognition, pp. 30-35.

Pavlovic, V.I. R. Sharma, T. S. Huang (1997) Visual interpretation of Hand Gestures for Human-Computer Interaction: A Review. IEEE PAMI, 19, (7),. 677-695.

Rubine, D. and McAvinney, P. (1990) Programmable Finger-tracking Instrument Controllers, Computer Music Journal, 14 (1), 26-41.

Westerman, W.C. (1999) Hand Tracking, Finger Identification, and Chordic Manipulation on a Multi-Touch Surface, University of Delaware PhD Dissertation.

Westerman, W.C. and J.G. Elias (1999) "A method and apparatus for integrated manual input," U.S. Patent Application Serial No. 09/236,513, January 25.

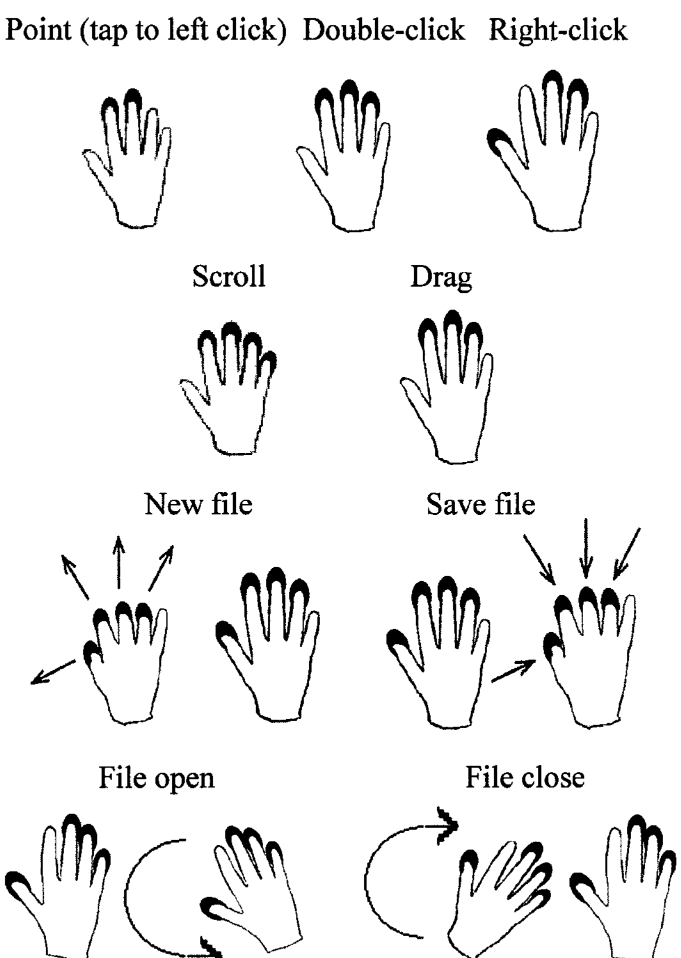

FIGURE 1: Examples of the 2-D gestures

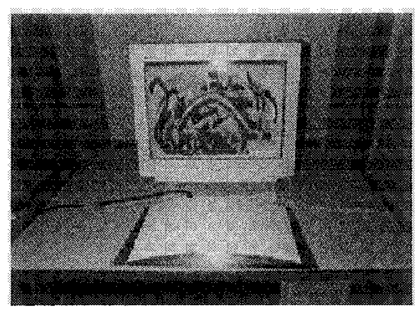

FIGURE 2: Finger-painting arrangement for children.

FIGURE 3: Finger painting table.

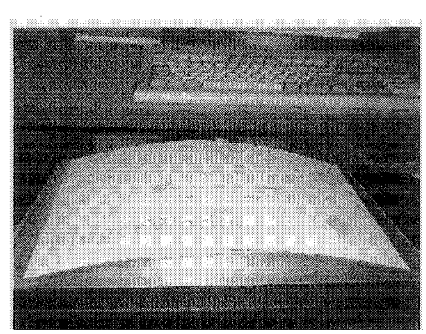

FIGURE 4: Prototype MTS keyboard.