High Precision Touchscreens: Design Strategies and Comparisons with a Mouse

Andrew Sears
Ben Shneiderman

Department of Computer Science
Human-Computer Interaction Laboratory
University of Maryland
College Park, MD 20742
January 23, 1989

Abstract

Three studies were conducted comparing speed of performance, error rates, and user preference ratings for three selection devices. The devices tested were a touchscreen, a touchscreen with stabilization (stabilization software filters and smooths raw data from hardware), and a mouse. The task was the selection of rectangular targets 1, 4, 16, and 32 pixels per side (0.4x0.6, 1.7x2.2, 6.9x9.0, 13.8x17.9 mm respectively). Touchscreen users were able to point at single pixel targets, thereby countering widespread expectations of poor touchscreen resolution. The results show no difference in performance between the mouse and touchscreen for targets ranging from 32 to 4 pixels per side. In addition, stabilization significantly reduced the error rates for the touchscreen when selecting small targets. These results imply that touchscreens, when properly used, have attractive advantages in selecting targets as small as 4 pixels per size (approximately one-quarter of the size of a single character). A variant of Fitts' Law is proposed to predict touchscreen pointing times. Ideas for future research are also presented.

Introduction

Overview

Many pointing devices are available for use with computers, but none are as natural to use as the touchscreen. Pointing at an item, or touching it, is one of the most natural ways to select it. Touchscreens allow the software designer to take advantage of this convenient selection method by having the users simply touch the item they are interested in.

Touchscreens are easy to learn to use, require no additional work space, have no moving parts, and are very durable (Pickering, 1986; Shneiderman, 1987; Stone, 1987; Muratore, 1987; Potter, Weldon, and Shneiderman, 1988a). Durability has made touchscreens popular in many applications, including kiosks at airports, shopping malls, amusement parks, and home automation. Even with these positive features, the touchscreens reputation for a lack of precision, high error rates, arm fatigue, and smudging the screen have resulted in limited use (Shneiderman, 1987; Pickering, 1986). Current touchscreen implementations do not include tasks requiring high resolution or tasks that are performed by frequent or experienced users. An adequate reduction in error rates, combined with the speed of the touchscreen may help expand this relatively limited use.

Previous Experiments

Many studies have compared touchscreens to other selection devices for various tasks. Our summary motivates our experiments. First, studies that compared the touchscreen to other selection devices are reviewed. Then several studies that explored the use of alternative selection strategies, an error reduction method we employ, will also be summarized.

Muratore (1987) did an extensive literature survey, reviewing fourteen studies that compared various cursor control devices. Her interpretation of these results implies that the touchscreen was the fastest but least accurate of the devices studied.

Hall, Cunningham, Roache, and Cox (1988) investigated the effects of various factors on touchscreen performance. The display was an IBM InfoWindow color terminal with a piezoelectric touchscreen using the land-on selection strategy forcing the selection at the location of the initial touch. Feedback was not provided about the accuracy of selections. They reported that accuracy varied from 66.7% for targets 10 mm per side, to 99.2% for targets 26 mm per side, and that accuracy was maximized once targets were approximately 26 mm per side.

Ostroff and Shneiderman (1988) compared a touchscreen, mouse, number keys, and arrow keys. The touchscreen was a Carroll Touch infrared touchscreen using the land-on strategy. The study involved selecting words from an interactive encyclopedia (Hyperties^{tm}). The results were similar to those of most other studies comparing the touchscreen and the mouse, indicating that the touchscreen was faster. They found no significant difference between error rates for the mouse and the touchscreen. This finding may be due in part to the relatively large size of the targets used and the rapid but awkward form of the jump mouse. (A jump mouse moves the cursor from one target to the next, skipping the space between them.)

Ahlstrom and Lenman (1987) compared a conductive touchscreen using the land-on strategy and mouse for the selection of a six character word from a list of words. This study indicated that the touchscreen was faster, but resulted in a much higher error rate.

Karat, McDonald, and Anderson (1986) compared a touchscreen, mouse, and keyboard for selection tasks. The touchscreen used was an Elographics analog membrane touchscreen using the land-on strategy. The task involved selecting items from a menu in a calendar program and a telephone directory. Some tasks also involved a typing subtask. The results indicated that the touchscreen was the preferred device for the task without the typing subtask, while the keyboard was preferred when the subtask was included. The touchscreen was the fastest for both tasks.

These studies have been limited to relatively large targets for selection tasks, but they do give some insight into the potential use of touchscreens. It is clear that a touchscreen can be used for rapidly selecting relatively large targets. Unfortunately, most of these studies also indicate that error rates were significantly higher for touchscreens. There are two explanations that may account for the majority of these errors, the inability of the touchscreens used in these studies to provide precise information about the location of a touch, and inadequate selection strategies for the tasks studied. While these studies were limited to relatively large targets, Garwin and Levine (1989) reported single pixel accuracy when using a laser-scanned touchscreen. Selection time and error rate data were not reported.

The inability of the touchscreen hardware to provide precise information may be due to a lack of resolution or the result of multiple pixel locations, possibly as many as twenty or more, being returned for a touch in a single location. While research by touchscreen manufacturers has dramatically increased the resolution of touchscreens, the problem of returning multiple pixel locations for a single touch remains. The extent to which this is a problem depends on both the touchscreen technology and manufacturer. Carroll Touch has published a Touch Handbook which provides a brief review of current touchscreen technologies including resolution, response time, and environmental resistance (Carroll Touch, 1989). Stabilization of the touchscreen will allow a single touch to result in the selection of a single pixel, possibly resulting in a significant reduction in errors, primarily for small targets. Ideally stabilization would be accomplished at the hardware level, but can also be done in software. Our studies will use software stabilization to filter and smooth raw data from the touchscreen hardware. Stabilization is an important idea that can be applied to many technologies including touchscreens, data gloves, and light pens, but has never been tested with touchscreens.

Many alternate selection strategies have been suggested to help reduce errors including take-off, first-contact, land-on, and others requiring a second touch. The land-on strategy uses the location of the initial touch for the selection. If the initial touch corresponds to a selectable region, that region is selected, otherwise no selection is made. The first-contact strategy results in the selection of the first selectable region the finger comes in contact with. This strategy the users to move their fingers on the screen until a selectable region is touched, this region is then selected and the appropriate process is initiated. Once again, all additional contact is ignored until the finger is removed from the screen. The take-off strategy allows users to place their fingers on the screen and move to the desired region on the screen before a selection is made. A cursor is placed slightly above the users fingers when they touch the screen indicating the exact location where a selection would be made. Users can then drag the cursor to the desired region, and lift their fingers from the screen to select it. A selection is made only if there is a selectable region under the cursor when users lifts their fingers. Others have also reported using variants of our lift-off strategy (Levine & Garwin, 1983; Murphy, 1987; Eller, 1987).

Several studies have been conducted to compare alternate selection strategies. The results indicate that some strategies may be promising for a wide range of task, and a significant reduction in error rates is possible (Murphy, 1987; Potter et al.,1988; Potter, Berman, and Shneiderman, 1989). Murphy (1987) compared seven selection strategies. He conducted an experiment that involved selecting targets that were .75 in^2 from a matrix of sixty targets. His results indicated few significant differences among the selection strategies, making it difficult to promote any single strategy as the best with respect to either selection time, or error rates for this target size.

Researchers at the University of Maryland Human-Computer Interaction Laboratory have performed two experiments comparing the land-on, first contact, and take-off strategies. The first experiment involved the selection of a two character state abbreviation from a 5x10 matrix. This study indicated that the first-contact strategy was the fastest, while the take-off strategy produced the fewest errors. The second experiment involved the traversal of a hypertext database by selecting highlighted words. There were no significant differences in the time needed to perform the task, while the first-contact and take-off strategies produced fewer errors than land-on (Potter et al., 1988; Potter et al., 1989).

These experiments indicate that first-contact may have been the fastest selection strategy, while the results pertaining to error rates did not consistently favor one strategy over the others. While these studies do provide a comparison of the selection strategies, they do not indicate how well a touchscreen using these strategies will perform compared to other selection devices.

Some researchers have claimed that the current touchscreen technology would not allow high-resolution selection, saying that selection of a single character with a touchscreen would be slow if it is even possible (Sherr, 1988; Greenstein & Arnaut, 1988). Others have blamed the size of the human finger for the lack of precision, claiming that the size of the user's finger limits the size of selectable regions (Beringer, 1985; Sherr, 1988; Greenstein & Arnaut, 1988). Previous studies have made no attempt at evaluating a touchscreen for high resolution tasks, restricting targets to relatively large sizes ranging from a square that is 0.25 inches per side, to targets that were approximately 1.0 x 1.6 inches. In addition, many of these studies have indicated that touchscreens result in significantly higher error rates than many other selection devices, including the mouse. Our experiments studied the selection of small targets with the touchscreen as compared with the mouse. We also study the effects that stabilization and the use of an alternative selection strategy have on these selections. Error rates and selection speed were measured. User preference data were also collected.

Experiment One: Stabilized Touchscreen, Non-stabilized Touchscreen, and Mouse

Introduction

The main purpose of the first experiment was to provide a comparison of a touchscreen, using an improved selection strategy, to a mouse for high resolution tasks. The secondary purpose was to investigate the effect stabilization has on speed of performance, error rates, and user preference for selection tasks when using a touchscreen.

Due to the difficulty involved in modifying hardware, stabilization was accomplished using software that filters and smooths raw data from the touchscreen hardware. These results should generalize to stabilization performed by either hardware or software.

The first step was to determine which selection strategy should be tested. To do this, we must understand the requirements of the task being evaluated. A typical high resolution task may be the selection of the start and stop points for a line in a graphics package, or possibly the selection of a character in a word processing program. Since it is difficult to accurately touch a single character, let alone a single pixel on the first attempt, the land-on strategy is not adequate. In addition, many high resolution tasks involve the selection of targets that are not defined before the selection is made, such as the starting point of a line, making the first-contact strategy inappropriate. On the other hand, the take-off strategy provides continuous feedback about cursor location, allowing the user to position the cursor before a selection is made by lifting the finger, making take-off the best candidate for many high resolution tasks.

Pilot Study Results

A pilot study helped determine that the original target sizes (16, 8, 4, and 2 pixels per side) were inappropriate. We decided that a larger range of target sizes would make the trends in selection time, and error rate data more apparent. The addition of a single pixel target allows predictions to be made about the selection of the smallest possible target we could represent.

Subjects

Thirty-six subjects volunteered from the Psychology Department subject pool of the University of Maryland. The amount of computer experience the subjects had was not controlled. Three subjects had used a touchscreen one time, while the remaining subjects had no experience. Experience using a mouse ranged from none to every day, with the majority of the subjects using the mouse very infrequently.

Equipment

All tasks were performed on an IBM PC-AT with an IBM Enhanced Color Display and a MicroTouch touchscreen. The monitor was placed on the desk in the normal monitor position with the keyboard placed in from of it. The monitor measured 27.6x19.5 cm and was used in EGA mode (640x350 pixels) resulting in pixels that were 0.4x0.6 mm. The MicroTouch touchscreen is a capacitive touchscreen that provides continuous information about the location of a touch on a 1024x1024 grid. It requires only a light touch to be activated and averages the location of all simultaneous touches and returns a centroid location. The touchscreen was cleaned once before the first subject began the experiment and was not cleaned again until the last subject had completed the experiment. Software was written to convert the touchscreen coordinates to pixel coordinates and to stabilize the resulting pixel coordinates. A MouseSystems Optical PC-Mouse with three buttons was used with a mouse pad that measured 22.9x19.7 cm. The mouse was calibrated to so a single pass horizontally on the pad resulted in the cursor moving the width of the screen, and a single pass vertically on the pad resulted in the cursor moving the height of the screen. Users were free to place the mouse pad anywhere they wanted.

After the experiment was completed it was discovered that the software provided with the mouse only allowed the cursor to be moved in two pixel increments horizontally. This did not impair the selection of targets, however, the resolution of the screen for the mouse tasks was essentially half (320x350) that of the screen for tasks using the touchscreen, possibly influencing the results in favor of the mouse. New mouse software was obtained for the second and third experiments to correct these problems.

Stabilization Software

Stabilization allowed a touch to result in a single pixel coordinate. The first attempt at stabilizing the touchscreen used running-means of the x and y coordinates. Although stability was improved dramatically, the selection of a single pixel was still not reliable, and the cursor lagged far behind the user's finger. Several additional steps were necessary to solve these problems.

First, a small region (0.9x1.7 mm) around the current cursor location was deactivated, requiring the user's finger to move beyond this region before the cursor moves (Figure 1, Region A). The second step was to define a larger region (8.6x16.8 mm) around the cursor that resulted in a movement that was only a fraction of the actual distance between the cursor and finger locations (Figure 1, Region B). For instance, if the user's finger was at point X (Figure 1), the cursor would only move to point Y. In this way, it was possible to perform very precise movements by dragging a finger on the screen.

Regions defined for stabilizing the touchscreen

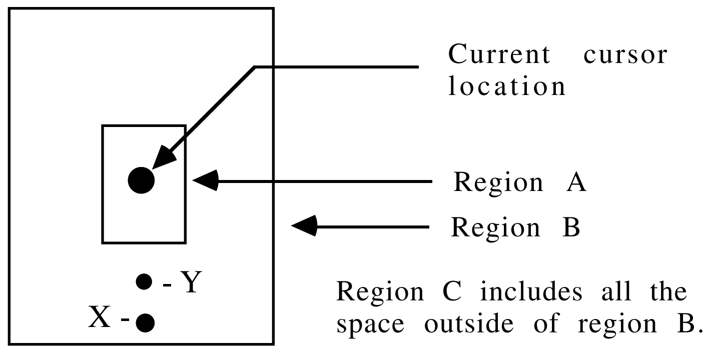

Figure 1 - If the current touch is within region A, the cursor does not move. If the touch is in region B, then the cursor moves a percentage of the distance between the current touch and the current cursor position. If the current touch is in region C, then the cursor moves to the location of the current touch.

These steps allowed the selection of a single pixel, but resulted in a significant delay between a movement of the finger and the movement of the cursor. One additional step was necessary to eliminate this delay. Whenever the location of the current touch was far enough from the current cursor location, the cursor moved directly to the location of the touch (Figure 1, Region C). In this way, the cursor could be dragged across the screen very rapidly without a significant distance between the cursor and the user's finger.

Although it may appear that stabilization will lead to a loss of directness between the movement of the finger and the movement of the cursor, careful manipulation of the size of Regions A and B allows stabilization without a loss of directness.

Design and Procedure

Selection device and target size were within subject variables. There were three selection devices, a mouse, a non-stabilized touchscreen, and a stabilized touchscreen. There were four rectangular targets: 1, 4, 16, and 32 pixels per side (0.4x0.6, 1.7x2.2, 6.9x9.0, 13.8x17.9 mm respectively). The four pixel target was approximately one-quarter of the size of a character which is 9x7 pixels. With this range of target sizes the results will be applicable to many practical tasks.

Each subject was tested with all selection devices and target sizes, resulting in three groups of four tasks for each subject. Each task required the selection of a series of six targets that were presented on the screen. Targets appeared in one of four positions, about 2.5 cm from each corner of the screen (Figure 2). Each subject had one practice trial for each task.

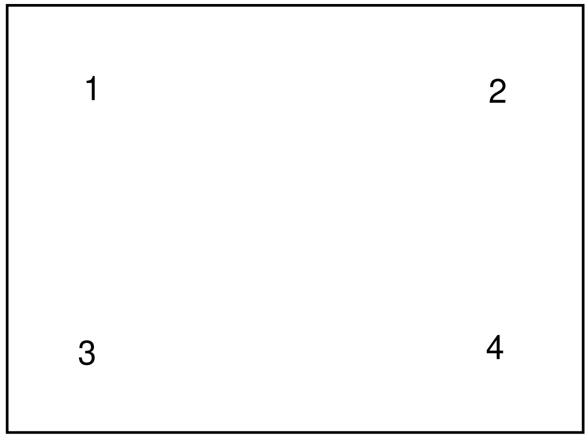

Figure 2 - Targets appeared in one of the four positions indicated on the diagram above.

Selection device was held constant in each group of tasks, and target size was held constant within a task. Within each group of tasks the target size decreased, in order from 32 pixels per side down to a single pixel. We chose to provide decreasing target sizes to facilitate the subjects' skill acquisition as they moved to smaller and more difficult targets. We recognized the disadvantages of non-random ordering, but we felt the additional experience was important. Each subject performed a set of selection tasks similar to the following list.

Task number	Device	Target Size (Pixels per side)
1	Non-stable Touchscreen	32
2		16
3		4
4		1
5	Mouse	32
6		16
7		4
8		1
9	Stable Touchscreen	32
10		16
11		4
12		1

The order that devices were used was randomized among subjects to prevent any possible bias. The order that the six targets within each task were presented was also varied to prevent subjects from anticipating the correct location for the next target.

Instructions were presented on the computer screen. Before each task a short message was presented telling the subject which device would be used. Subjects then pressed ENTER to begin the task. A target was presented and subjects had to select it with the appropriate device. When the target was successfully selected, or five errors were made on the current target, a tone sounded and the next target was presented. An error occurred each time subjects lifted their fingers without making a successful selection. A maximum of five errors were allowed per target to prevent subjects from getting stuck indefinitely on a target if they were not able to select it. Six targets were presented for each task. After the sixth target was selected, a message indicating the number of errors and time taken was presented. Subjects then pressed ENTER to continue to the next task.

When using the mouse, selections were made by moving the mouse until the center of the cursor was on the target and clicking any of the mouse buttons. Selection using both touchscreens involved touching the computer screen, dragging the cursor until the center of the cursor was on the target, and then lifting the finger from the screen. In all cases the cursor was a plus sign, '+', made by five pixels vertically, and five pixels horizontally and was presented 6.1 mm above the subject's finger to allow the subject to view both the cursor and target when selecting small targets. The cursor was blue and targets were red; when the cursor and target overlapped, the intersection became white making it easier to know when the cursor was correctly positioned.

The time to select each group of six targets and the number of errors per group were recorded for each task. In addition, subjects were asked to indicate their preference for each device on a scale from 1 to 9 (1 being strongly disliked, 9 being strongly liked). All data were recorded on the computer.

Results

Selection Times

The mean time from the initial presentation of a target until either successful selection or until five errors occurred appear with standard deviations in Table 1 and the means are plotted in Figure 3. An ANOVA with repeated measures for selection device and target size showed significant main effects for selection device, $F(2,70)=5.0$, $p < .01$, and target size, $F(3,105)=232.5$, $p < .001$. A significant interaction between selection device and target size, $F(6,210)=50.0$, $p < .001$, was also found. Tukey's post hoc HSD test showed that both touchscreens are faster than a mouse for targets 16 pixels per side ($p<.05$), and the mouse is faster than both touchscreens for a single pixel ($p<.05$). There were no other significant differences across the devices.

	Target Size (Pixels per side)			
	32	16	4	1
Mouse	3.13 (1.28)	3.47* (1.60)	4.97 (1.98)	6.08* (1.87)
Stabilized Touchscreen	1.83 (0.37)	1.98 (0.33)	4.27 (1.27)	11.78 (4.42)
Non-Stabilized Touchscreen	1.86 (0.45)	1.93 (0.47)	4.57 (1.65)	12.28 (4.95)

Table 1 - Mean selection time (in seconds) per target (standard deviation in parentheses). (* $p < .05$)

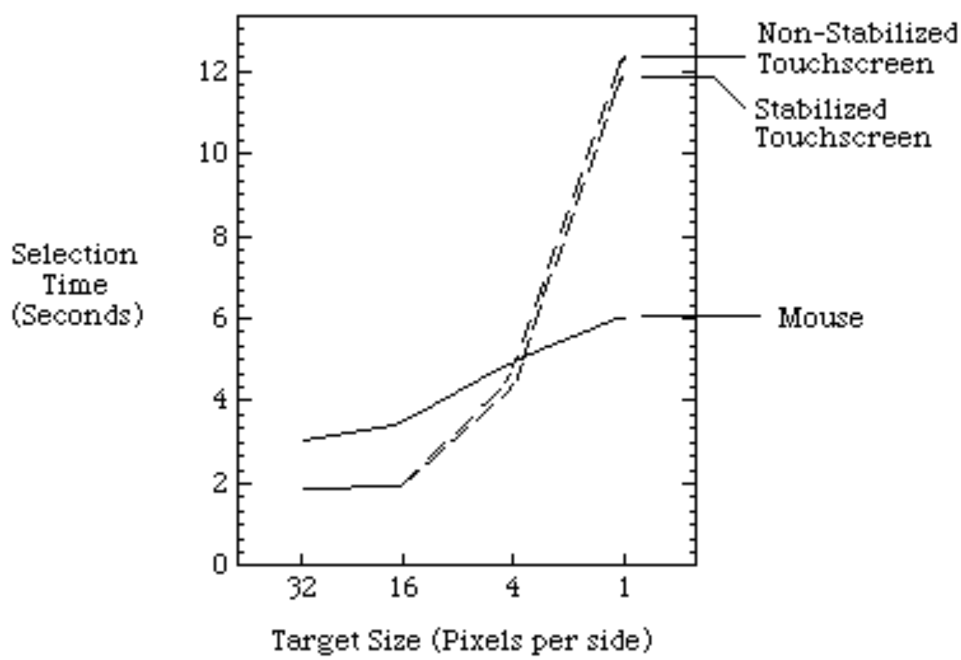

Figure 3 - Selection time for four target sizes and three selection devices.

<u>Error Rates</u>

The mean error rate per target and standard deviations appear in Table 2 and the means are plotted in Figure 4. An ANOVA with repeated measures for selection device and target size showed significant main effects for selection device, $F(2,70)=186.4$, $p < .001$, and target size, $F(3,105)=356.6$, $p < .001$. A significant interaction between selection device and target size, $F(6,210)=177.44$, $p < .001$, was also found. Tukey's post hoc HSD test showed that the non-stabilized touchscreen resulted in more errors than either of the other devices for the 4x4 pixel target ($p<.05$). For the single pixel target, the mouse resulted in fewer errors than either of the other devices, and the stabilized touchscreen resulted in fewer errors than the non-stabilized touchscreen ($p<.05$).

	Target Size (Pixels per side)			
	32	16	4	1
Mouse	0.08 (0.15)	0.06 (0.12)	0.08 (0.18)	0.50 (0.68)
Stabilized Touchscreen	0.03 (0.06)	0.05 (0.10)	0.35 (0.58)	1.53* (1.08)
Non-Stabilized Touchscreen	0.02 (0.06)	0.06 (0.15)	0.77* (0.60)	4.38* (0.62)

Table 2 - Mean number of errors per target (standard deviation in parentheses). (* $p < .05$)

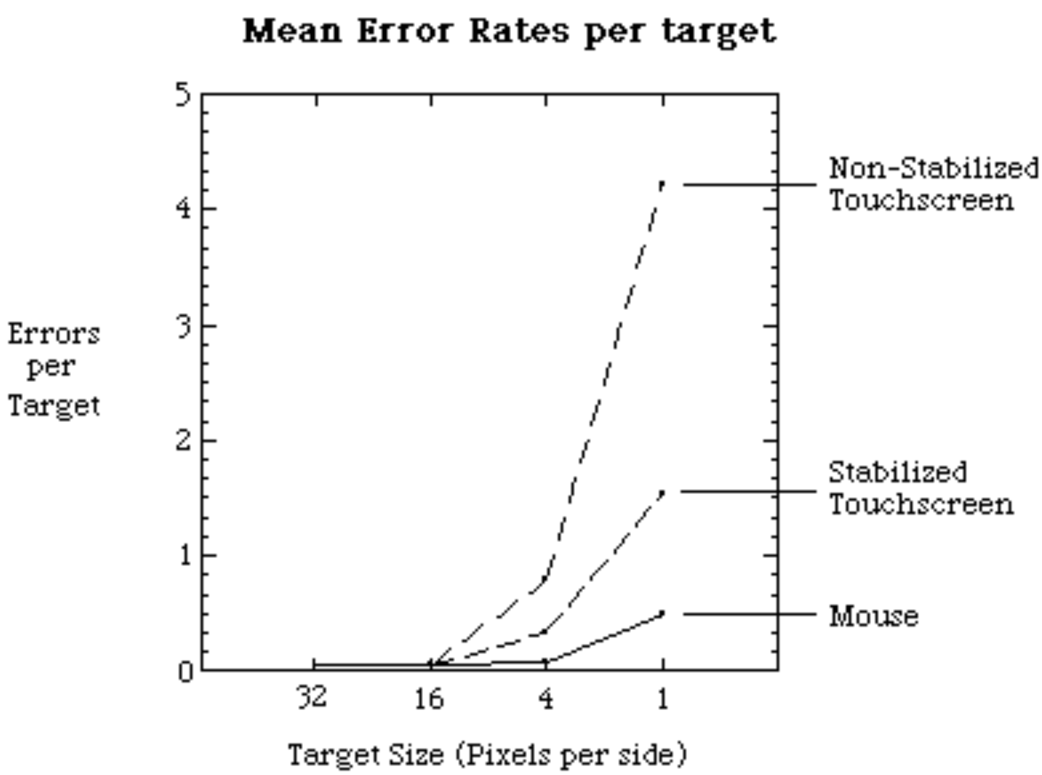

Figure 4 - Error rates for four target sizes and three selection devices.

<u>User Preference</u>

User preference means and standard deviations appear in Table 3. A one-way repeated measures ANOVA on selection device showed a effect for selection device, $F(2,70)=106.9$, $p<.001$. Tukey's post hoc HSD test showed that the non-stabilized touchscreen received lower preference ratings than either of the other devices ($p<.05$).

	Mouse	Stabilized Touchscreen	Non-Stabilized Touchscreen
Mean User Preference Rating	7.5 (1.7)	6.7 (1.9)	1.9* (1.5)

Table 3 - Mean user preference ratings for three selection devices (standard deviation in parentheses). (* $p < .05$)

Discussion

The stabilized touchscreen was as fast or faster than the mouse while making no more errors for targets as small as four pixels per side. This indicates that a touchscreen can be used for selection of single characters, which are 9 x 7 pixels, in many applications that currently use a mouse. There were no differences in selection times between the two touchscreen implementations, and stabilization resulted in a significant reduction in errors for the two smaller targets.

The results indicate that it was possible to select a single pixel with a touchscreen although the mouse resulted in faster, more accurate selections than either touchscreen. The significant increase in selection time and error rates from the four pixel targets to the single pixel indicates that none of the selection devices, as currently implemented, are appropriate for the selection of single pixel targets. Additional work must be done to improve the input devices if they are to be used for selecting single pixel targets without a zooming feature.

User preference ratings indicate that the stabilized touchscreen was preferred over the non-stabilized touchscreen. Since subjects are using their fingers to move the cursor on the screen, it seems reasonable to expect the cursor movements to correspond directly to movements of their fingers. When using the non-stabilized touchscreen, the jitter caused by the lack of stability violates this expectation, possibly resulting in lower preference ratings. When stabilization is added the cursor tracks their fingers accurately, resulting in both higher preference ratings and lower error rates.

The targets explored in this experiment allow predictions to be made about a wide range of practical target sizes. Considering that the majority of high resolution tasks are performed by experienced users, studies that include additional practice, or instructions for selecting small targets, may prove useful. Several subjects devised strategies for selecting the single pixel targets. Two subjects learned that they could position the cursor near the target and then simply roll their fingers up and down or left and right to make fine manipulations. Subjects that were observed using this strategy on the stabilized touchscreen had a mean error rate of only 0.25 when selecting six single pixel targets. When this mean is compared to the overall mean of 1.53, it becomes apparent that this strategy is very successful in reducing errors. If all subjects were exposed to this method of selecting small targets, the error rates might decrease. The second and third experiments incorporated this idea, presenting brief instructions to subjects before the experiment.

Although many people have claimed that smearing will be a significant problem when using touchscreens this is problem did not occur in the office like conditions of this experiment. The touchscreen surface is lightly ground, rather than polished, thereby reducing glare and the impact of fingerprints. The touchscreen used for this study was cleaned once before the experiment began and was not cleaned again. Small amounts of oil and dust accumulated on the screen but the accumulation was similar to that on standard monitors. Actually, less dust appears to collect on the touchscreen used in this experiment than on many standard monitors. No subjects complained that the accumulation affected their performance, and the experimenters did not notice a difference in performance between the early subjects and those at the end of the study.

Experiment Two: Stabilized versus Non-stabilized Touchscreen

Introduction

The major purpose of the second experiment was to eliminate potential problems with the first experiment. Comparisons were limited to the stabilized and non-stabilized touchscreens. In the first study the targets were presented in one of four locations with decreasing target sizes. In this experiment, the target sizes and locations were randomly determined. In addition, the maximum number of errors allowed per target was increased from five to ten to lessen the ceiling effect. Subjects were also given brief instructions on the selection of small targets with the touchscreen. Stabilization was accomplished using the software from the first experiment and the take-off selection strategy was used.

Pilot Study Results

A pilot study indicated that making the center pixel of the cursor a different color would aid the subjects in identifying when the single pixel targets were correctly selected. The cursor was also placed slightly higher above the finger.

Subjects

Twenty subjects volunteered from the Psychology Department subject pool of the University of Maryland. The amount of computer experience the subjects had was not controlled. Three subjects had used a touchscreen one time, while the remaining subjects had no experience at all. Experience using a mouse ranged from none to every day, with the majority of the subjects using the mouse very infrequently.

Equipment, Design, and Procedure

The stabilized and non-stabilized touchscreens from the first experiment were used. Once again, the touchscreen was cleaned once before the first subject began the experiment and was not cleaned again until the last subject had completed the experiment. Selection device and target size were within subject variables. There were two selection devices, a stabilized touchscreen and a non-stabilized touchscreen. There were four rectangular target sizes: 1, 4, 16, and 32 pixels per side.

Each subject was tested with both selection devices and all target sizes. The order that devices were used was randomized across subjects. The elimination of one selection device allowed the number of targets to be increased from 6 to 9 per target size providing additional data. Subjects selected thirty-six targets with each device. The target size used for each trial was chosen randomly, with a total of nine targets of each size being selected with each device. Targets were presented in random locations on the screen.

Subjects received brief instructions on the selection of small targets with each selection device. Subjects were informed that fine manipulations may be easier if they rolled their fingers once the cursor was near the desired target but they were not required to use this method.

Instructions were presented on the computer screen. Before each task a short message was presented informing subjects that the touchscreen would be used for this task. Subjects then touched the center of the screen to begin the task. A target was presented and subjects had to select the target using the touchscreen. When the target was successfully selected, or ten errors were made on the current target, a tone sounded and the next target was presented. An error occurred each time subjects lifted their fingers without making a successful selection. A maximum of ten errors were allowed per target to prevent subjects from getting stuck indefinitely on a target if they were not able to select it. After the last target was selected, a message was presented indicating the total number of selections made, counting both correct and errors. Subjects started the next task by touching the screen center.

Selection using the touchscreen involved touching the computer screen, dragging the cursor until the center of the cursor was on the target, and then lifting the finger from the screen. In all cases the cursor was a plus, '+', made by five pixels vertically, and five pixels horizontally and was presented 8.4 mm above the subject's finger. The cursor was blue, except the center pixel which was black, and the targets were yellow. When the cursor and target overlapped, the intersection turned red, except when the center pixel overlapped the intersection turned back to yellow. This allowed subjects to identify when the single pixel targets could be correctly selected.

The time to select and the number of errors were recorded for each target. In addition, subjects were asked to indicate their preference for using each device on a scale from 1 to 9 (1 being strongly disliked, 9 being strongly liked). All data were recorded on the computer.

Results

<u>Selection Times</u>

The mean time from the initial presentation of a target until either successful selection or until ten errors occurred appear with standard deviations in Table 4. An ANOVA with repeated measures for selection device and target size showed significant main effects for selection device, $F(1,179)=59.6$, $p < .001$, and target size, $F(3,537)=490.7$, $p < .001$. A significant interaction between selection device and target size, $F(3,537)=65.9$, $p < .001$, was also found. Tukey's post hoc HSD test showed that the stabilized touchscreen was faster than the non-stabilized touchscreen for the single pixel target ($p < .05$). There were no other significant differences across the devices.

	Target Size (Pixels per side)			
	32	16	4	1
Non-Stabilized Touchscreen	1.5 (0.3)	1.8 (0.5)	4.3 (3.1)	21.9* (13.9)
Stabilized Touchscreen	1.6 (0.3)	1.9 (0.5)	4.1 (2.5)	11.4 (9.6)

Table 4 - Mean selection time (in seconds) per target (standard deviation in parentheses). (* $p < .05$)

<u>Error Rates</u>

Error rate means and standard deviations appear in Table 5. Statistics experts were consulted regarding the lack of variance for the non-stabilized touchscreen when selecting the 32 pixel target. Following their suggestion, a value of 0.001 was inserted for one subject allowing an ANOVA to be performed. An ANOVA with repeated measures for selection device and target size showed significant main effects for selection device, $F(1,179)=292.3$, $p < .001$, and target size, $F(3,537)=370.8$, $p < .001$. A significant interaction between selection device and target size, $F(3,537)=267.9$, $p < .001$, was also found. Tukey's post hoc HSD test showed that the stabilized touchscreen resulted in significantly fewer errors than the non-stabilized touchscreen for the single pixel target ($p < .05$). There were no other significant differences across devices.

	Target Size (Pixels per side)			
	32	16	4	1
Non-Stabilized Touchscreen	0.00 (0.00)	0.06 (0.25)	0.61 (1.13)	5.95* (3.96)
Stabilized Touchscreen	0.02 (0.18)	0.01 (0.11)	0.15 (0.56)	0.73 (1.36)

Table 5 - Mean number of errors per target (standard deviation in parentheses). (* $p < .05$)

<u>User Preference</u>

A t-test was performed on the user preference data,T = 4.7, indicating that the stabilized touchscreen (mean = 6.5, S.D. = 2.1) was preferred over the non-stabilized touchscreen (mean=2.6, S.D. = 2.3) ($p<.001$).

Discussion

The stabilized touchscreen was as fast or faster than the non-stabilized touchscreen while making no more errors for targets as small as four pixels per side. Subjects made an average of 5.95 errors per target when selecting single pixels with the non-stabilized touchscreen, and made the maximum number of errors over 39% of the time. When selecting a single pixel with the stabilized touchscreen, the average number of errors was 0.73, and subjects made the maximum number of errors less than 0.01% of the time. This indicates that the selection of a single pixel is possible with few errors with the stabilized touchscreen, and that increasing the maximum number of errors allowed would increase the average number of errors for the non-stabilized touchscreen, while making no difference for the stabilized touchscreen. A comparison of the error rates for the single pixel targets from the first and second experiments seems to indicate that providing a simple set of instructions on the selection of small targets may result in a significant decrease in error rates, also indicating that experience may significantly improve performance when using the touchscreen. User preference ratings indicate that subjects preferred to work with a stabilized touchscreen. Once again, there was no problem with the screen becoming dirty.

Experiment Three: Stabilized Touchscreen versus Mouse

Introduction

With the exception of the selection devices being used the third experiment was identical to the second. For this experiment comparisons were limited to the stabilized touchscreen and the mouse to eliminate any potential negative effects using the non-stabilized touchscreen may have.

Subjects

Twenty subjects volunteered from the Psychology Department subject pool of the University of Maryland. Once again, the amount of computer experience the subjects had was not controlled. Eight subjects had used a touchscreen one time, while the remaining subjects had no experience at all. Experience using a mouse ranged from none to every day, with the majority of the subjects using the mouse very infrequently.

Equipment, Design, and Procedure

The equipment, design, and procedure for the third experiment were the same as for the second experiment with the following exceptions. The stabilized touchscreen and mouse were compared in this experiment. New mouse driver software was obtained that allowed single pixel movements both vertically and horizontally. When using the mouse subjects began a task by pressing any of the mouse buttons. Selections were made by positioning the center of the cursor on the target and pressing any of the mouse buttons. After completing the experiment subjects were interviewed about their impressions of the various devices for selecting various size targets.

Results

Selection Times

The mean time from the initial presentation of a target until either successful selection or until ten errors occurred appear with standard deviations in Table 6. An ANOVA with repeated measures for selection device and target size showed significant main effects for selection device, $F(1,179)=38.2$, $p < .001$, and target size, $F(3,537)=276.1$, $p < .001$. A significant interaction between selection device and target size, $F(3,537)=55.3$, $p < .001$, was also found. Tukey's post hoc HSD test showed that the mouse was faster than the stabilized touchscreen for the single pixel target ($p < .05$). There were no other significant differences across the devices.

	Target Size (Pixels per side)			
	32	16	4	1
Mouse	2.2 (0.9)	2.7 (2.2)	3.5 (1.5)	5.4 (3.6)
Stabilized Touchscreen	1.7 (0.5)	2.3 (0.8)	4.4 (2.2)	9.8* (6.9)

Table 6 - Mean selection time (in seconds) per target (standard deviation in parentheses). (* $p < .05$)

Error Rates

Error rate means and standard deviations appear in Table 7. An ANOVA with repeated measures for selection device and target size showed significant main effects for selection device, $F(1,179)=14.5$, $p < .001$, and target size, $F(3,537)=28.9$, $p < .001$. A significant interaction between selection device and target size, $F(3,537)=12.8$, $p < .001$, was also found. Tukey's post hoc HSD test showed that the mouse resulted in significantly fewer errors than the stabilized touchscreen for the single pixel target. There were no other significant differences across devices.

	Target Size (Pixels per side)			
	32	16	4	1
Mouse	0.03 (0.17)	0.02 (0.15)	0.04 (0.19)	0.20 (0.57)
Stabilized Touchscreen	0.01 (0.10)	0.04 (0.26)	0.10 (0.35)	0.64* (1.64)

Table 7 - Mean number of errors per target (standard deviation in parentheses). (* $p < .05$)

<u>User Preference</u>

A t-test was performed on the user preference data,$T = 3.9$, indicating that the mouse (mean = 7.9, S.D. = 1.6) was preferred over the stabilized touchscreen (mean = 5.6, S.D. = 2.1) ($p<.001$).

Discussion

The stabilized touchscreen was as fast or faster than the mouse while making no more errors for targets as small as four pixels per side. The results for the single pixel indicate that the mouse was faster and resulted in fewer errors than the stabilized touchscreen. The user preference ratings indicate that the mouse was preferred.

Informal interviews provided important information about how subjects rated the devices. Just as rolling their finger caused small movements when positioning the cursor, rolling their fingers when lifting them from the screen also cause small movements which often resulted in errors. Additional studies may remedy this problem by using a touch location slightly before the user actually lifts their finger (Beringer, 1989). All subjects felt that the touchscreen was fine for selecting the two largest targets (32x32 and 16x16 pixels), and the majority of subjects felt that the touchscreen was acceptable for the second smallest target (4x4 pixels). Most subjects indicated that their rating for the touchscreen was strongly influenced by their feelings about selecting the single pixel targets. Subjects felt that selecting a single pixel was significantly harder using the touchscreen. Since previous research has shown strong positive subjective satisfaction for touchscreens, it is recommended that future experiments collect preference data about using each device for each target size used, rather than about the devices in general. We believe that preference for touchscreen would be greater than for the mouse if the single pixel tasks were avoided.

Case study with an Experienced User

One subject (Sears) who was experienced with both the touchscreen and mouse performed the second and third experiments five times each. Average selection times and error rates for this subject are presented in Tables 8 and 9 respectively. Preference ratings were not collected due to the familiarity of the subject with the purpose of the experiment. These results are meant to provide some insight into the performance of experienced users for the selection of small targets. Statistical analyses were not performed, however these data should provide some useful information.

	Target Size (Pixels per side)			
	32	16	4	1
Non-Stabilized Touchscreen	0.95 (0.12)	1.16 (0.16)	2.36 (1.05)	8.32 (4.80)
Stabilized Touchscreen	0.93 (0.08)	1.15 (0.17)	1.94 (0.35)	3.86 (1.71)
Mouse	1.26 (0.24)	1.44 (0.23)	2.06 (0.33)	2.62 (0.54)
Stabilized Touchscreen	0.95 (0.10)	1.21 (0.19)	2.02 (0.44)	3.84 (1.42)

Table 8 - Mean selection times (in seconds) per target for an experienced user (standard deviation in parentheses).

	Target Size (Pixels per side)			
	32	16	4	1
Non-Stabilized Touchscreen	0.00 (0.00)	0.00 (0.00)	0.36 (0.68)	4.87 (3.91)
Stabilized Touchscreen	0.00 (0.00)	0.00 (0.00)	0.02 (0.15)	0.09 (0.29)
Mouse	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.04 (0.21)
Stabilized Touchscreen	0.02 (0.15)	0.00 (0.00)	0.00 (0.00)	0.07 (0.25)

Table 9 - Mean number of errors per target for an experienced user (standard deviations in parentheses).

Data for the stabilized and non-stabilized touchscreens appears to indicate that there is no difference in performance for the two larger targets. Data for the smaller targets strongly indicates that both selection time and error rates are decreased when stabilization is used with the touchscreen. The most dramatic differences are for the single pixel targets, where the selection time for the stabilized touchscreen is less than one half of the time for the non-stabilized touchscreen, and the error rates are reduced from 4.9 errors per target to less than 0.1.

Data for the stable touchscreen and the mouse is equally interesting. There appears to be no difference in error rates for any of the target sizes. In addition, the touchscreen appears to be faster than the mouse for the two larger size targets. There is no difference for selection time for the four pixel targets while the selection time for the single pixel target favored the mouse.

Comparing the data for experienced users (Tables 8 and 9) to that for novices (Tables 4, 5, 6, and 7) provides several interesting differences. Selection times for all devices appear to decrease for all target sizes, indicating that training may improve performance for all of these devices. The most dramatic differences were for the smaller targets where selection times decreased by as much as 60% for the stabilized touchscreen and 50% for the mouse. Error rates decreased for all devices for the two smaller targets, once again indicating that training may improve performance. The most dramatic differences were for the smaller targets where error rates decreased by as much as 80% for the stabilized touchscreen and the mouse.

Overall, these data appear to confirm the expectation that experience speeds selection and reduces errors for both the touchscreen and mouse. The data also indicates that the stabilized touchscreen may be appropriate in many situations where a mouse would have been used for the selection of targets as small as four pixels per side. These results indicate that training may be very important when using any of these devices for the selection of small targets.

Conclusions

Future Research

Additional experiments that trained subjects to use both the mouse and the touchscreen, and exposed them to strategies for selecting small targets may prove interesting. This would provide valuable information about the use of touchscreens by frequent or experienced users as well as information about how fast users learn to use each device.

"Gravity" refers to the idea of expanding the area that actually selects a target. To do this the designer simply makes a larger area around that target selectable or the cursor could be "pulled" to the target when it is close enough, and released when the finger is dragged far enough away. This would not aid in the selection of densely packed targets, but it would help for even the smallest targets if they are not too close together. Other possibilities, such as zooming, for selecting very small targets should also be explored.

One problem with touchscreens that continues to be mentioned is arm fatigue. Changing the screen mounting angle may substantially reduce this problem (Ahlstrom and Lenman, 1990). Some researchers, as well as optometrists (Grant, 1987), have suggested that lower screen placements may reduce eye strain.

If this is true, a simple change in the screen position may not only reduce arm fatigue, but also reduce eye strain. Studies frequent touchscreen users would provide important information about the effect of fatigue on performance.

The fear that the touchscreen would become dirty and would need frequent cleaning has not been supported by these experiments. The touchscreen used for these studies was cleaned once before each experiment was started and was not cleaned again until the last subject had completed the experiment. Small amounts of oil and dust accumulated on the screen but the accumulation was similar to that on standard monitors. No subjects complained that the accumulation affected their performance, and experimenters did not notice any difference in performance between early subjects and those at the end of the studies. It does not appear that dirt will be a problem in most environments although there are bound to be some environments where the accumulation of dirt will pose a problem.

<u>Revising Fitts' Law for selection with a touchscreen</u>

Fitts' Law, $T = a+b[\log(2d/w)]$, predicts that the time necessary to select a target (T) is proportional to the distance (d) to the target, the width of the target (w), and two device dependent parameters a and b. This law has been demonstrated in many situations and devices (Card, English & Burr, 1978). It is not clear that this law, in its original form, applies to all selection methods with all targets.

Cursor positioning with a touchscreen is inherently different than cursor positioning with most other devices. With most devices, there is always a cursor present and movement of the device moves the cursor. With touchscreens, there may not be a cursor present when the user's finger is not touching the screen. Once the user touches the screen, the cursor is placed near the finger and can then be dragged around the screen. In addition, when using most devices, there is some time dedicated to locating the device which moves the cursor.

With the touchscreen there is no time needed to locate the cursor position device, but the positioning time must include the time it takes the user to initially place their finger on the screen. Since, users typically first touch the screen close to the actual target, the distance they must move the cursor should be relatively small.

For these reasons, Fitts' Law must be modified to predict positioning times for a touchscreen. The modified version of Fitts' Law must have an additional factor that accounts for the time used to initially place a finger on the screen. This time will be relatively small and should be a function of the size of the target and the distance, measured in three dimensions, from the original location of the hand to the target. The second factor in the modified Fitts' Law will account for the time necessary to move the finger to the target once it is placed on the screen. This factor will be a function of the target size; distances will not effect this factor since the user will initially place their finger near the target. Using this information, the following can be offered as a modification of Fitts' Law for touchscreens: $T=a+b[\log(cD/W)]+d[\log(e/W)]$, where the first factor, b[log(cD/W)], measures the time to place the finger on the screen initially. The second factor, d[log(e/W)], measures the time to position the cursor once the finger has been placed on the screen. The distance, measured in three dimensions, from the original hand location to the location of first contact is D, and W = some measurement of target size. The constants a, b, c, d, and e are determined for each specific case. The distance, D, could be the distance from the previous selection if selections are made in rapid succession, but it may be difficult to measure in many circumstances.

Of course many other factors influence the time to select a target and therefore the values of the constants in both original and modified Fitts' Law. These include the target background and shape (a long narrow target may be harder to select than a square target of the same area).

Analyzing this revised Fitts' Law using the data from these experiments has proven inconclusive. The limited number of target sizes (four) and the lack of precise distance measurements (from the user's hand to the target) makes analysis difficult, allowing this data to fit many predictive equations including both the original and modified Fitts' Law. Future experiments to investigate this or any other revision to Fitts' Law will need to use additional target sizes and carefully measure the distance between the user's hand and the target.

Impact on practitioners

These results may have a significant impact on the number and type of applications that touchscreens are used for. Knowledge of the effect stabilization has on touchscreen accuracy and speed will allow designers to include touchscreens in applications requiring rapid performance, high precision, and low error rates. The performance achieved with the touchscreen in this experiment should be attainable by any designer, making it a desirable input device for many applications. The results obtained from these experiments indicate that although both the mouse and touchscreen could be used, neither is adequate for selection of single pixel targets. Additional research is necessary for selection tasks involving targets of this small size.

More applications could be designed using touchscreens, taking advantage of their simple operation, ease of learning, rapid performance, and potentially low error rates. Another direction would be to develop software to replace existing mouse drivers, allowing a touchscreen to be used with many applications that currently use a mouse. These results indicate that the touchscreen can easily be used for character selection in word processors, and possibly even for graphics design. Most graphics packages include a zoom option which allows the user to examine a magnified screen image which would make the touchscreen adequate for these applications.

Acknowledgements

We would like to thank Kent Norman, Linda Weldon, Yoram Kochavy, John Schnizlein, and NCR and all of the members of the Human-Computer Interaction Laboratory and the User Interface Design class (CMCS 828S, Fall 1988) for their useful suggestions. We appreciate the excellent comments of three anonymous reviewers who led us to make more precise descriptions of our experimental methods. We also thank the volunteers for their assistance.

References

Ahlstrom, B. and Lenman, S., (1990). Fatigue when using a touch screen, *Work in progress*.

Ahlstrom, B. and Lenman, S., (1987). Touch screen, cursor keys, and mouse interaction, *Work with Display Units*, Elsevier, North-Holland, 831-837.

Beringer, D., (1989). Touch panel sampling strategies and keypad performance comparisons., *Proceedings of the Annual Meeting of the Human Factors Society*.

Beringer, D. and Peterson, J., (1985). Underlying behavioral parameters of the operation of touch-input devices: biases, models, and feedback., *Human Factors*, 27(4), 445-458.

Card, S, English, W., and Burr, B., (1978). Evaluation of mouse, rate-controlled isometric joystick, step keys, and text keys for text selection on a CRT, *Ergonomics*, Volume 21, 8, 601-613.

Carroll Touch, (1989). *Touch Handbook*, Carroll Touch Inc. Round Rock, TX.

Eller, T.S., (1987). Personal communication.

Garwin, R. L. and Levine, J. L., (1989). Light-sensitive pen to substitute for finger in laser-scanned touch screen", *IBM Technical Disclosure Bulletin*, Volume 32, Number 3B, 435-436.

Grant, A., (1987). The computer user syndrome, *Journal of the American Optometric Association*, Volume 58, Number 11, 892-901.

Greenstein, J. and Arnaut, L., (1988). Input devices in: Helander (ed.), *Handbook of Human-Computer Interaction*, 495-519.

Hall, A., Cunningham, J., Roache, R., and Cox, J., (1988). Factors affecting performance using touch-entry systems: tactal recognition fields and system accuracy, *Journal of Applied Psychology*, Volume 73, No. 4, 711-720.

Karat, J., McDonald, J., and Anderson, M., (1986). A comparison of selection techniques: touch panel, mouse, keyboard, *International Journal of Man-Machine Studies*, Volume 1, 73-92

Levine, J. L. and Garwin, R. L., (1983). A laser-scanned touch input device. For Display '83 in Japan.

Muratore, D.A., (1987). Human performance aspects of cursor control devices, *Mitre Corporation working paper 6321*, Houston, TX.

Murphy, R., (1987). Evaluation of methods of touch screen implementation for interactive computer displays, Abstract presented at 2nd International Conference of Human-Computer Interaction, Honolulu, HI.

Ostroff, D. and Shneiderman, B., (1988). Selection devices for user of an electronic encyclopedia: an empirical comparison of four possibilities, *Information Processing & Management*, Volume 24, Number 6, 665-680.

Pickering, J.A., (1986). Touch-sensitive screens: the technologies and their applications, *International Journal of Man-Machine Studies*, 25, 249-269.

Potter, R., Weldon, L., and Shneiderman, B., (1988). Improving the accuracy of touch screens: an experimental evaluation of three strategies, *Proceedings of the Conference on Human Factors in Computing Systems*, ACM SIGCHI, New York, 27-32.

Potter, R., Berman, M., and Shneiderman, B., (1989). An experimental evaluation of three touch screen strategies within a Hyperties database, *International Journal of Human-Computer Interaction*, Volume 1, No. 1, 41-52.

Sherr, S., (1988). *Input Devices*, Academic Press, San Diego, CA.

Shneiderman, B., (1987). *Designing the User Interface: Strategies for Effective Human-Computer Interaction*, Reading, MA.

Stone, M. D., (1987). Touch-screens for intuitive input", *PC Magazine*, August, 183-192.