EVALUATING DIFFERENT TOUCH-BASED INTERACTION TECHNIQUES IN A PUBLIC INFORMATION KIOSK

Roope Raisamo
Department of Computer Science
University of Tampere
P.O. Box 607 (Pinninkatu 53B)
FIN-33101 Tampere, Finland
email: rr@cs.uta.fi

ABSTRACT

Public information kiosks are becoming more and more common. However, their user interfaces are still based on simple button interfaces, which may be implemented with physical buttons or with virtual buttons drawn on a touchscreen. We have implemented a multimodal kiosk prototype that is based on touch and speech input. This paper describes an evaluation in which 23 users compared five area selection techniques for touchscreens. These techniques are based on selection time, touch pressure and direct manipulation. The results show that pressure-sensitive techniques offer a possible alternative to the other selection techniques, but require careful designing. Our time-based technique was the most intuitive and the direct manipulation technique was understood well after an initial learning phase.

KEYWORDS

Public information kiosks, touchscreens, interaction techniques, pressure-sensitive techniques, empirical tests

1 INTRODUCTION

Many public information kiosks are based on conventional desktop computers equipped with a keyboard and a mouse, or on a computer that communicates with the user using a large touchscreen that is used for both input and output. Still, with these modern kiosks human-computer interaction is limited. Current kiosks may have good multimedia presentation capabilities, but the user can only use buttons or on-screen keyboards to give input for the system.

We present new area selection techniques that are based on touch pressure. Area selection involves the selection of a circular area on the display. These techniques are compared in an empirical evaluation to find out how intuitive they are and how the users react on them. In addition to public information kiosks similar techniques could be applied in desktop systems that are equipped with a touchscreen, and with other touch-based input devices such as touchpads and touch tablets.

Developing new ways to use touch sensing as an input modality has gained popularity lately. MacKenzie and Oniszczak (1998) have developed techniques for touchpads to better simulate real button presses. Harrison *et al* (1998) and Hinckley and Sinclair (1999) have experimented with using touch in other ways than to detect position information.

This paper is organized as follows. First, our kiosk user interface is described using a restaurant information system as an example. Next, we have a description of our touchscreen selection techniques. Then we describe our experimental design and the results. Finally, we discuss our observations and highlight potential areas for future research.

2 THE KIOSK USER INTERFACE

Our kiosk user interface framework, Touch’n’Speak, was introduced in an earlier paper (Raisamo, 1998). It is presented in this section as it appears to the user. We have developed a demonstration application (Figure 1) that lets the user select restaurants in Cambridge, Massachusetts. The user has two available input modalities: touch and speech. The modalities can be freely mixed. For example, the user can start the interaction by pressing some touchscreen buttons, but can select items by voice. When Figure 1 was captured, a user was selecting a circular area on the map with her finger. Speech commands were given to carry out the search operation.

Touch is used in several ways. First, conventional large touch buttons are presented on-screen when needed. For example, in Figure 1 the large buttons at the bottom let the user select the criteria that he or she wants to refine. Pressing a button shows a specific input screen, such as the one for choosing a location shown in Figure 1. The search can be started any time by pressing the "Search" button. The second way to use touch is to select options in the "Your options" list at the left by touching them. Finally, the touchscreen is used in picture area selection.

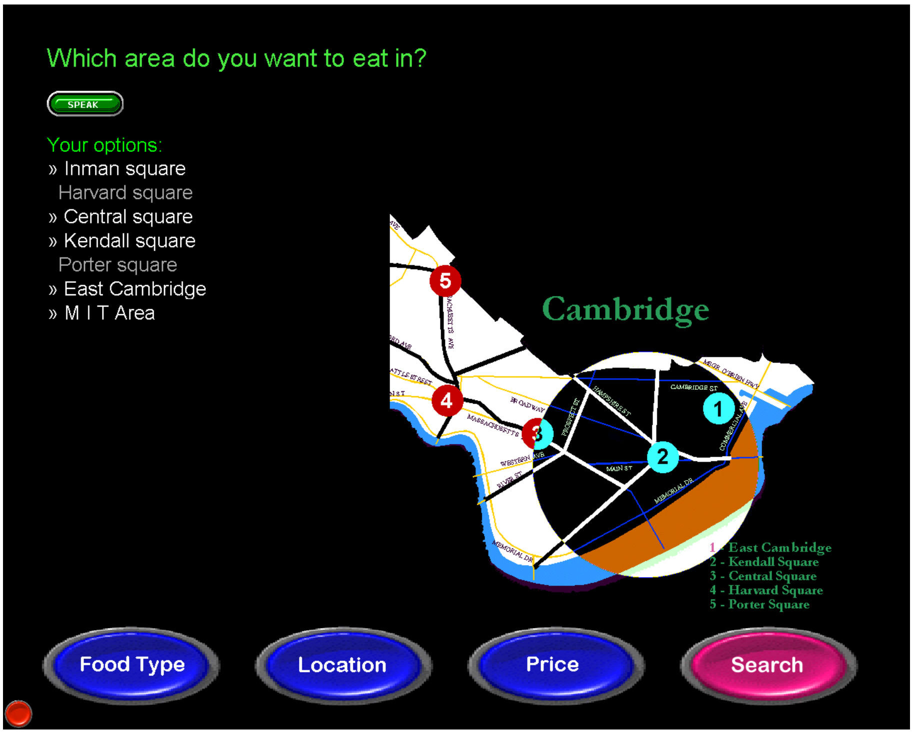

Figure 1. An area is being selected in the restaurant system. The user selects a circular area on the map of Cambridge, and the options are automatically selected based on this area.

In addition to touching the options, the user has an option to speak the option words. In the example system, the user can navigate from Figure 1 to different input screens by saying "Food type", "Location" or "Price". The options in the "Your options" list can be selected and deselected by voice, and the user can also use some additional voice commands in the selection process, including "Help", "All", "None" and "Undo".

When the user has defined the criteria and initiates the search, the system generates results in HTML format and presents them in an integrated Web browser (Raisamo, 1998). The resulting pages can also be browsed using speech in addition to pressing buttons.

3 TOUCHSCREEN TECHNIQUES

We have developed different kinds of selection techniques that can be used to select a part of a picture. All of the techniques rely on a touchscreen. We used a 20" Elo IntelliTouch surface wave touchscreen (Elo Touchsystems, 1999) with screen resolution 1024 * 768. The screen can detect touch pressure in 256 levels based on the area that the finger occupies on screen. The way the radius of the circle changes depends on the selection technique. Default values for different parameters are inside parentheses.

3.1 Time-based selection

This selection technique is based on a timer. When the user touches the picture, the radius of the selection starts growing and grows as long as the user keeps the finger on screen. The selection area can be simultaneously moved while it is growing by sliding the finger on the screen. This technique has two parameters: timer interval (100 ms) and increment (15 pixels). The default values seemed to work well with our 20" touch monitor, but some users thought that the speed was too fast. The selection grows as long as the finger is pressed on the screen and stops growing when the finger is lifted off the screen.

3.2 Incremental pressure-based selection

Incremental pressure-based selection technique works as follows: When the pressure level is greater than a threshold value, the selection radius is increased by a pre-defined number of pixels per time unit. When the level is smaller than the same threshold value, the selection radius is decreased by the same number of pixels per time unit. This selection technique has three parameters: the threshold value (128), the increment or decrement (10 pixels), and the time unit (100 ms). Optimal values depend on touchscreen capabilities, i.e., how the pressure is defined, and on the speed in which the change is being added or subtracted from the current value.

3.3 Nonlinear pressure-based selection 1

In our preliminary experiments we noticed that direct use of the pressure level does not result in an appropriate selection area. The selection area changed too rapidly, and since the pressure level decreases instantly when the user lifts the finger off the screen, the final selection area was not what the user thought it would be, but much smaller. The users noticed that they need to lift the finger off the screen fast to prevent the area from shrinking.

To cure the hastiness of direct pressure mapping we decided to introduce a coefficient function to transform the pressure value. Obviously, since direct z-value mapping did not work well, all linearly behaving coefficients would give the same results, just in different scales. We therefore decided to experiment with nonlinear functions. A solution that seems to behave quite well is to multiply the pressure value by *sine of (pressure value/3)*, in which pressure/3 represents an angle in degrees. Another similar coefficient function could have been the square root of the pressure value multiplied by a factor.

3.4 Nonlinear pressure-based selection 2

We have also experimented with a slower-changing coefficient, *sine of (pressure value/5)*. With the Elo touchscreens, direct pressure mapping was too hasty in small pressure levels, and our coefficient functions make it steadier as all pressure values are multiplied by a value that is smaller than 1. In addition, both of our nonlinear functions speed up growing the area at the highest pressure levels, which is good since pressing the screen hard is not as comfortable as pressing it normally. Especially moving the finger while pressing it hard on screen is inconvenient. The two coefficient functions and similar square root functions are presented in Figure 2.

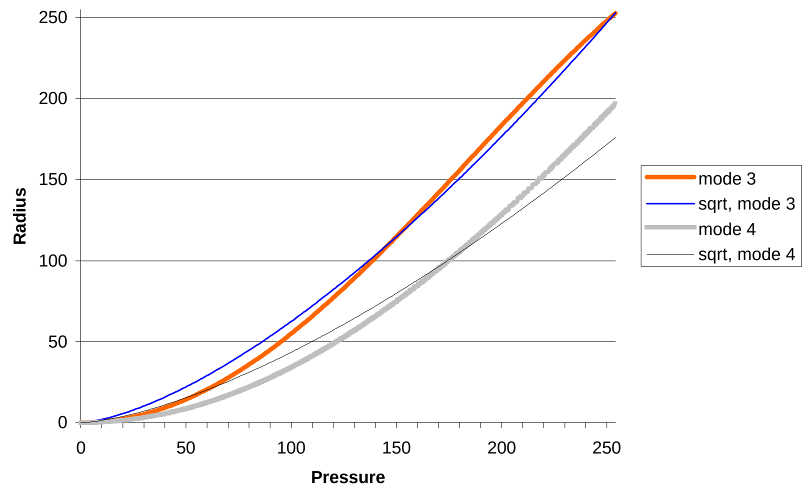

Figure 2. The pressure level multiplied by different coefficient functions.

3.5 Direct manipulation

Our direct manipulation selection technique resembles closely the mouse usage in drawing programs. In this technique, the user draws the radius of a circle with his or her finger pressed on screen. The circle is drawn at the same time. This technique is based on the principles of conventional direct manipulation (Shneiderman, 1982). This technique was implemented to allow the comparison of a basic direct manipulation technique with our new time- and pressure-based techniques. Direct manipulation is intuitive when the user knows what to do. However, problems in understanding this technique were expected since nothing happens if the user does not move the finger.

4 EXPERIMENT

We were comparing five different area selection techniques. We also wanted to see how the users react on a multimodal interface that uses touch and speech. In this section we describe our experimental procedure and the results. We focused on comparing the selection techniques and did not try to determine the accuracy of speech recognition since we were using an English speech recognizer and the users were non-native English speakers.

4.1 Experimental procedure and setup

The experiment was carried out in our usability laboratory. Each trial lasted about 30 minutes with the kiosk and 20-30 minutes in a post-trial interview. The use of the kiosk was videotaped using a screen converter to capture the screen and a video camera to record how the user worked with his or her hands. The interviews were recorded. In addition, the prototype was set up to record log information each time that the user performed an action. This provided us with objective information on the use of different modalities and on the way the users tried different selection techniques.

We purposefully chose not to train our users before the trial. This was done because we wanted to determine the intuitiveness of different selection techniques and speech commands that we had selected. Naturally, an end-user system would have a help system and more hints for the user, but the prototype that we tested did not have them. There was an assistant with the test user all the time the kiosk was used. His task was to give the user different tasks to do following a similar script for each user and to ensure that the user understood what to do.

The intuitiveness was measured in the following way: the assistant explained that the map works in a new way. The user was asked to try it and deduce how it works. If the user did not figure out a selection technique within three minutes, the assistant showed it to the user and asked him or her to try it. In these cases the technique surely was not intuitive, or at least not in accordance with prior computer or kiosk usage. However, the users needed to know how each interaction technique worked in order to compare them later in the experiment.
The test was divided in five phases:

1. experimenting with and becoming familiar with the system,
2. using only speech input to find information,
3. trying five different touch-based area selection modes and finding out how they respond to touch,
4. using all the possible ways to find information, and
5. filling in an evaluation sheet and answering questions in an interview.

The touch-based selection modes that we tested were listed in Section 3: 1. the time-based selection, 2. the incremental pressure-based selection, 3. the nonlinear pressure-based selection with coefficient *sine of (pressure value/3)*, 4. the nonlinear pressure-based selection with coefficient *sine of (pressure value/5)*, and 5. the direct manipulation technique. They are referred to as modes 1-5 later in this paper.

In total, there were 23 users in the final experiment, made up of 11 female and 12 male users. The users' computing skills varied much, which was planned since information kiosks should be usable by everyone. 15 users had taken an introductory course in HCI, and 9 of them considered themselves as technically oriented persons. The other 8 users were typical end-users that were not interested in technology, but knew how to use a mouse.

4.2 Qualitative results

Figure 2 shows the average rankings of different area selection techniques. The users ranked the modes by giving the best mode rank 1 and the worst mode rank 5. They could not give the same rank to two or more modes. These rankings show that the users liked the time-based selection technique best, and ranked the direct manipulation technique second. Two pressure-sensitive techniques, modes 2 and 3 were ranked the worst. The more controllable and slower mode 4 received the best rankings of the pressure-sensitive techniques. There were users who said that controlling the pressure level is very hard. It is also clear that delicate control may be impossible for elder people and for some people with disabilities. The differences in rankings were statistically significant with $F_{4,110}=7.25$ ($p<0.00005$).

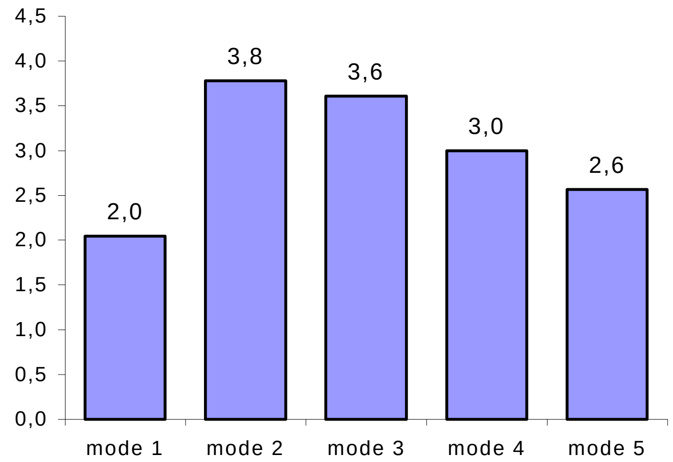

Figure 2. Averages of the rankings of different selection techniques (the smaller the better).

We also asked the users to rank the modalities. They were ranked using scale 1-3. Combined touch and speech received best average ranking (1.4). The second best average ranking was received by touch alone (1.7) and the worst ranking by speech alone (2.9). The differences between the rankings were statistically significant in one-way analysis of variance with $F_{2,66}=57.92$ ($p<0.00005$).

Even if the users liked to use speech, they would not have trusted in it as the only input modality due to recognition errors. Without exception the users thought that speech recognition was useful, but it should work well. Two users could not use speech input at all, but all the 21 others could control the kiosk using speech commands with a few recognition errors. However, in the fourth phase as many as 6 users did not use touch input at all but just speech. They were observed to be the users for whom speech recognition worked best.

4.3 Quantitative results

Average times for determining how the techniques work are summarized in Table 1 with other statistics of the trials. This data shows that direct manipulation (mode 5) required more time to understand. The average time to try each area selection technique was 1 minute 40 seconds. Mode 2 was not understood well and it took time to understand mode 3 even though it was completely understood by half of our subjects. Mode 4 seems to be a possible rival for the direct manipulation technique since it was understood by almost as many users as mode 5 and this took less time. The highest average trial time for mode 5 is in line with our observation that most users did not understand it at all in the beginning of the trial. They believed that the technique could only be used to select small areas, because they did not move their fingers. After many repetitive touches they noticed its behavior. The number of fast learners and fast quitters contain those users that either learned or quitted faster than in the average time (1 minute 40 seconds). The differences between modes in discovering their use were statistically significant in one-way analysis of variance with $F_{4,110}=3.34$ ($p<0.05$).

TABLE 1

A SUMMARY OF DATA ON INTUITIVENESS (n=23)

	Mode 1	Mode 2	Mode 3	Mode 4	Mode 5
Average time	1 min 38 s	1 min 39 s	1 min 33 s	1 min 36 s	1 min 52 s
Understood by	17 users	6 users	11 users	13 users	15 users
Fast learners	10 users	3 users	4 users	8 users	8 users
Fast quitters	5 users	9 users	10 users	7 users	2 users

We measured the pressure level in all five selection techniques. The following data is based on all longer-lasting selection sessions. A selection session means many repetitive selection operations without other operations between them. The average pressure level within all sessions was 145. The maximum pressure level was 245 in average. The levels and the standard deviations are summarized in Table 2.

TABLE 2

A SUMMARY OF DATA ON PRESSURE LEVELS (level ranging from 0 to 255)

	Mode 1	Mode 2	Mode 3	Mode 4	Mode 5
Average pressure level	137	157	134	149	150
Average max. pressure level	246	249	237	243	251
Standard deviation	63	63	56	60	58

As can be seen in Table 2, there were no big differences between the modes in the pressure level used. The differences were not statistically significant.

5 DISCUSSION

It is not at all sure that if a technique works well with a mouse it works as well with a touchscreen. This was noticed in our test for intuitiveness. All the users (23) had used a drawing program with a mouse, but only 15 of them noticed a similar behavior in the direct manipulation technique. It also took more time to notice how it worked than with the other selection techniques. When the technique was explained to them, the users liked the accuracy and sense of control of direct manipulation. But eight users said that they would never have guessed how this technique works. This may be due to prior conditioning with mouse use, which offers an opportunity for further research.

It seems that the direct manipulation selection that we used would benefit from a small hint of its function: “Move your finger”. In our experiment the assistant told this hint to the users if they did not manage to use the technique. After this advice the technique was clear to all the users. All the other techniques started working instantly when the user touched the screen, and the problems in these techniques were related to finding out why the area is changing and how to control it. They would also benefit from a small hint text if the user does not understand them. The problem with the direct manipulation technique was the same as these techniques have in drawing programs: the user has to know that moving the mouse with a mouse button pressed draws a circle or a rectangle. If only a button is pressed, nothing happens. But once these operations are learned it is easy to draw the shapes.

It was interesting to see that the users thought that the incremental pressure-based technique behaved just like the time-based technique, but sometimes they had problems with understanding why the selection shrank. Ten pixels seemed to be a suitable increment value in this technique, but the threshold value should be higher than what we used in this experiment in order to make this technique usable. This technique would also benefit from using two threshold values as MacKenzie and Oniszczak (1998) used to simulate button clicks with a touchpad. They used threshold values so that button activation requires larger pressure value than button deactivation. This would give the user a safe margin within which the pressure can vary without affecting the way the selection is changing.

Two-handed interaction techniques have proved their usefulness in many interactive systems dating back to Buxton and Myers (1986). When trying to find out how the area selection techniques worked, four users tried to use their both hands: "what if I press it with two fingers…" We would have liked to provide them with a two-handed technique, but we were constrained by touchscreen technique and driver limitations. Our interviews showed that touchscreens should allow many simultaneous touch points, which is not the case with the commercial screens (Pickering, 1986).

Speech was a highly preferred input modality. The users thought that it is beneficial and they actually used it in almost all given search tasks. There were also many users that used speech exclusively in a simple search task. Our findings suggest that speech should be considered as an input modality for public kiosks, but the users expect it to work. If speech does not work well, the users reject it fast. Keyword-spotting speech recognition engines would make speech interaction more fluent and would be a better alternative than the command-based speech recognition engine that was used in this experiment.

6 CONCLUSION

This paper described an evaluation on five different touch-based selection techniques in a multimodal public information kiosk. The simplest technique was time-based and received the best rankings from the users. It was also the most intuitive technique: the technique that was understood best without any instructions. The second best technique was a basic direct manipulation technique that resembled drawing a circle in a drawing program. Its main problem was that it required the longest time to discover and come to a working understanding of the technique. Two of the pressure-based techniques were ranked the worst, but the slower non-linear technique that directly mapped the pressure level to the size of the selection area was almost as good as the direct manipulation technique. Many users found pressure-based techniques interesting, which supports further studies in which touch pressure is used as one way to control the interface. There was clearly a need for experimental touchscreens that can detect many simultaneous touch points.

ACKNOWLEDGMENTS

This work was partially supported by the Academy of Finland (Project 38118, Grant 52936). We wish to thank Toni Pakkanen for working as an assistant in this experiment. In addition, we wish to thank OzCHI'99 reviewers for their insightful comments.

REFERENCES

Buxton, W., Myers, B. A. (1986). A study in two-handed input. *Human Factors in Computer Systems, CHI '86 Conference Proceedings*, ACM Press, 321-326.

Elo TouchSystems (1999). *IntelliTouch surface-wave products* [WWW]. http://www.elotouch.com/.
link checked 10/20/99

Harrison, B. L., Fishkin, K. P., Gujar, A., Mochon, C., and Want, R. (1998). Squeeze me, hold me, tilt me! An exploration of manipulative user interfaces. *Human Factors in Computing Systems, CHI '98 Conference Proceedings*, ACM Press, 17-24.

Hinckley, K. and Sinclair, M. (1999). Touch-sensing input devices. *Human Factors in Computing Systems, CHI '99 Conference Proceedings*, ACM Press, 223-230.

MacKenzie, I. S. and Oniszczak, A. (1998). A comparison of three selection techniques for touchpads. *Human Factors in Computing Systems, CHI '98 Conference Proceedings*, ACM Press, 336-343.

Pickering, J. A. (1986) Touch-sensitive screens: the technologies and their application. *International Journal on Man-Machine Studies* **25:3**, 249-269.

Raisamo, R. (1998). A multimodal user interface for public information kiosks. *Proceedings of 1998 Workshop on Perceptual User Interfaces*, Matthew Turk (ed.), San Francisco, 7-12. Available online at:
http://research.microsoft.com/PUIWorkshop/Papers/Raisamo.pdf.
link checked 10/20/99

Shneiderman, B. (1982). The future of interactive systems and the emergence of direct manipulation. *Behaviour and Information Technology* **1:3**, 237-256.