Association for Information Systems
AIS Electronic Library (AISeL)

SIGHCI 2005 Proceedings

Special Interest Group on Human-Computer Interaction

2005

Investigating the Usability of the Stylus Pen on Handheld Devices

Xiangshi Ren
Kochi University of Technology, ren.xiangshi@kochi-tech.ac.jp

Sachi Mizobuchi
Keio University, sachi.mizobuchi@nokia.com

Investigating the Usability of the Stylus Pen on Handheld Devices

Xiangshi Ren
Kochi University of Technology
ren.xiangshi@kochi-tech.ac.jp

Sachi Mizobuchi
Nokia Research Center/ Keio University
sachi.mizobuchi@nokia.com

ABSTRACT

Many handheld devices with stylus pens are available on the market, however, there have been few studies which examine the effects of the size of the stylus pen on user performance and subjective preferences for hand-held device interfaces. Two experiments were conducted to determine the most suitable dimensions (pen-length, pen-tip width and pen-width) for a stylus pen. In Experiment 1, five pen-lengths (7, 9, 11, 13, 15 cm) were evaluated. In Experiment 2, six combinations of three pen-tip widths (0.5, 1.0 and 1.5mm) and the two pen widths (4 and 7mm) were compared. In both experiments, subjects conducted pointing, steering and writing tasks on a PDA. The results were assessed in terms of user performance and subjective evaluations for all three pointing, steering and writing tasks. We determined that the most suitable pen dimensions were 11 cm for length, 0.5 mm for tip width, and 7mm for pen width.

Keywords

Mobile computing, pen-based devices, pointing task, steering task, handwriting task.

INTRODUCTION

With handheld information devices such as personal digital assistants (PDAs), users are often required to interact with a number of objects on a small display. Pen input, which has already been adopted in most current PDAs, is considered to be an efficient input method which permits direct and intuitive manipulation. Pen-based input is well suited to jotting down text and accessing information in mobile computing situations.

There have been studies which compared the pen with other input methods for PDAs [Mizobuchi et al. 2002], or which investigated target selection strategies [Ren and Moriya 2000] and human performance with a pen in different tasks [Brewster 1999]. These studies worked on the design or usability of PDA interfaces. However, only a few studies have been reported on the physical aspects of the input device, such as the size of the stylus pen. One exception is the study [Baird et al. 2002] which investigated the effects of probe length on tests which are based on Fitts' law. However, the range of probe length they examined (100 mm to 400 mm) was relevant to heavy mechanical tasks and was thus outside the range that would suit handheld devices and stylus usability.

Many kinds of papers and pens are available for traditional handwriting. Each individual is able to find a combination of materials and tools uniquely suited to his or her needs. However, in the digital world, for instance, in designing handheld devices, industrial designers are often confronted with a conflict between limited hardware space and the usability of the input device. Smaller device design permits only a small space for storing a pen, but a too small pen tends to impair the usability of the device. Currently the length, width of a pen provided with a device seems to be limited to the size of the device itself with little or no consideration being given to the usability of the pen, e.g. various sizes of pen are available on the market, such as Palm Source，Revo，Palm, IPAQ，G-FORT.

To maximize the usability of these devices serious consideration must be given to the physical aspects of the pen, such as its length, width and tip-width. This study looks at how the length, tip-width and width of the pen affect human performance and user preferences when using handheld devices. We performed two experiments in which we evaluated the effect of pen size on PDA usage. In these experiments, we used three basic and common PDA tasks, a pointing task, a steering task and a handwriting task because human activities on PDAs can be classified broadly into these three categories.

We consider that the most suitable size of the pen will include the following characteristics: high performance (e.g. minimum movement time, minimum error rate, and the larger index of performance in pointing and steering tasks, high character recognition rate, a minimum number of error corrections, minimal number of protruding strokes, minimum handwriting character input time and movement time between input boxes in handwriting tasks), and high subjective ratings (e.g. ease of writing and minimum degree of fatigue).

User interface designers often have to conduct empirical comparisons among many candidate devices. In order to measure the user performance during completion of simple tasks, Fitts [Fitts 1954] carried out an experiment to establish a model for movement time in pointing type tasks, known as Fitts' law, commonly expressed in the following equation [MacKenzie 1992]:

$$MT = a + b\log_2(\frac{A}{W} + 1) \quad (1)$$

where *MT* is the acquisition time of a pointing task, *A* is the distance or amplitude from the starting position to the target, *W* is the width of the target, and *a* and *b* are empirically determined constants; the reciprocal of *b*, called the index of performance (*IP*), is often used as a measure of input device efficiency. The log term of the equation is defined as the index of difficulty for a pointing task (ID_P).

$$ID_P = \log_2(\frac{A}{W} + 1) \quad (2)$$

Modern computer interactions, however, are more than just pointing tasks. One limitation of Fitts' law paradigm is that it can be applied for pointing tasks and cannot be applied for other tasks such as drawing, writing and navigating through a menu and its nested menus. Accot and Zhai [Accot and Zhai 1997] provided the first quantitative tool for predicting the difficulty of HCI steering tasks. A steering task requires one to move the input device (or pointer of the device) a certain distance through a tunnel. A daily example of the steering task is driving an automobile without crossing the road boundaries [Accot and Zhai 1999]. Examples of the steering task performed with input devices include steering through a menu, moving the scroll bar of a window. Accot and Zhai called the following model a "steering law" which models the relationship between completion time *MT* and task parameters:

$$MT_S = a + bID_S \quad (3)$$

where *a* and *b* are empirically determined constants, *1/b* is called the index of performance (*IP*) in a steering law. *IDs* is the index of difficulty of the steering task which, for linear movement, is defined as:

$$ID_S = \frac{A}{W_S} \quad (4)$$

where *A* represents tunnel length, *Ws* represents tunnel width. For circular tunnel movement, *IDs* can be defined as:

$$ID_S = \frac{2\pi\ r}{W_S} \quad (5)$$

where *r* is the radius of the circle and *Ws* is the width of the tunnel.

To date the steering law has been well verified in manual movement tasks under various conditions [Accot and Zhai 1999][Dennerlein et al. 2000][Accot and Zhai 2001]. We can use the steering model as a quantitative measure for determining the performance of candidate devices in our experiments.

The third task encountered within PDAs is the handwriting task. There is no measurement model for this task, however, we looked at some performance factors (e.g. character recognition rates) and subjective preferences (e.g. ease of writing) while writing characters into the text entry boxes because there are many applications that provide text entry boxes on the screen for handwriting input.

We hypothesized that the significance (or insignificance) of the differences between the candidate devices may be adequately observed in each of the three tasks (pointing, steering and handwriting).

METHOD

Subjects

Twelve subjects (9 male and 3 female) participated the experiments. Subject ages ranged between 21 and 23 years. The average age was 21.25. All subjects were right-handed. None of the subjects had previous experience using PDAs.

Each subject held the PDA with their non-dominant hand and the input device with their dominant hand while in a sitting posture. All subjects were instructed not to rest their hands (elbow or arms) on the table or any other objects during the experiment. This ensured that the environment was common to each user and that it was also a typical and universally available PDA environment.

Apparatus

In the experiments, the device used was an "iPAQ Pocket PC" by Compaq Co., running Windows CE 3.0. The size was 84mm (W) x 16mm (D) x 134mm (H). It weighed 184g. It had a 240 x 320 pixel display with a 0.24 mm pixel pitch. The display was accurately calibrated before the experiments. The experimental programs were developed using Sun Microsystems JAVA (for the pointing task and the steering task) and Microsoft embedded Visual C++ (for the handwriting task).

Pen-length, pen-tip width and pen-width

We tested the effects of pen-length using five lengths (7, 9, 11, 13, 15cm) all with 1.0 mm pen-tip width and 5 mm pen-width in Experiment 1. In Experiment 1, the independent variable range of the pen-length was set at 7cm, 9cm, 11cm, 13cm, and 15cm. 95% of the 12 subjects tested in a pilot study revealed that any pen length less than 7 cm was too difficult to handle. Regarding the longest pen tested, 15cm is approximately the same length as an ordinary pencil or a ball point pen with a cap.

In Experiment 2, we set two pen-widths (4mm and 7mm) and three pen-tip widths (0.5mm, 1.0mm and 1.5mm) as the independent variables range. The reason we chose these settings was that the pen-width attached to existing PDAs was close to 4mm or 7mm and they are close to the pen-width of a real pen and pencil. We combined the three pen-tip widths (0.5, 1.0 and 1.5mm) and the two pen-widths (4 and 7mm) giving us six pen-tip widths/pen-widths combinations, i.e. 0.5mm/4mm, 1.0mm/4mm, 1.5mm/4mm, 0.5mm/7mm, 1.0mm/7mm, 1.5mm/7mm pens (pen-tip width/pen-width).

Each of the two experiments included the three tasks described in detail in the following sections.

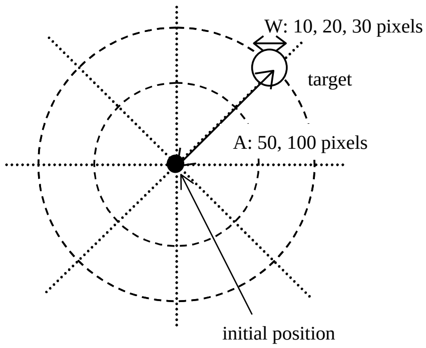

Figure 1: In factorial design for pointing task.

Pointing task

The steps for pointing at a target were as follows:

(1) *The center circle was displayed*: In the beginning of this task, the white circle (the center circle) was displayed in the center of the PDA display. Subjects tapped the center circle. When the center circle was tapped, the start time was recorded.

(2) *The target was displayed*: The center circle turned black when the subject tapped it, and the other white circle (target) was displayed randomly on the PDA display. The subject was asked to tap the target as quickly and accurately as possible. When the target was tapped, the end time was recorded, and if the target had not been tapped accurately, an error was recorded.

(3) *Repeat*: The white center circle was displayed again immediately after the subject tapped the target. The subject repeated (1) and (2).

After they finished testing each pen, the subjects were asked to answer a questionnaire which included: ease of pointing, degree of fatigue, and a overall evaluation on a scale from 1 (worst) to 7 (best).

The factorial design and levels were as follows (see Figure 1):

- The target appearance positions were located in 8-directions around the center circle.
- The distances between the center of the center circle and the center of the target were 50 and 100 pixels.
- The diameters of the targets were 10, 20 and 30 pixels.

After the procedure for the task was explained to the subject, a practice session with the accessory pen was performed. After this, the subjects were told to do the real trials. For each of the eleven pens (five pen-lengths + six pen-tip widths/pen-widths), each subject had a total of 48 test trials (8 directions x 2 distances x 3 target sizes). Each subject completed 240 test trials (Experiment 1) and 288 test trials (Experiment 2). In each pen, 576 test trials (12 subjects x 48 test trials) were completed. The order for all pens was different for each of the subjects.

Steering task

Two types of steering task, a straight tunnel and a circular tunnel, were used in the experiment (see Figure 2). The steps for the steering tasks were as follows:

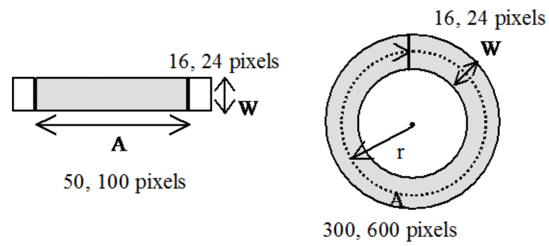

Figure 2: Two steering tasks. Straight tunnel steering and circle tunnel steering.

(1) *The tunnel was displayed*: In the beginning of this task, the tunnel (either linear or circular) was displayed on the PDA display. This is the path that subjects had to drag the pen-tip along.

(2) *Start segment and goal segment*: In each tunnel, there were the start segment and goal segment. In the linear tunnel, the frame of the start segment was black, and the frame of the goal segment was red. In the circular tunnel, the start segment and goal segment were located in the same position. When the subjects put the pen on the start segment, the label "START" was displayed in the upper left of the display as a signal to begin the task, and the start time was recorded. When the pen entered the goal segment, all objects on the display disappeared as a signal to end the trial, and the end time was recorded. After that, the next trial was displayed. Releasing the stylus pen from the display after leaving the start segment and before entering the goal segment, or crossing the borders of the path, resulted in an "error" and the trial would be recorded as the invalid trial. The subjects were asked to continue to attempt the task until they succeeded in each trial.

The factorial design and levels were as follows:

- The task types were linear and circular.
- Distances were 50 and 100 pixels in the linear task, and 300 and 600 pixels in circular task.
- Tunnel widths were 16, 24 pixels.
- Directions were left to right and right to left.

After the procedure of the task was explained to the subject, a practice session with the accessory pen was performed. The subjects were then told to do the real trials. For each of the eleven pens (five pen-lengths + six pen-tip widths/pen-widths), each subject had a total of 16 test trials (2 task types x 2 distances x 2 widths x 2 directions). Each subject completed 80 test trials (Experiment 1) and 96 test trials (Experiment 2). For each pen, 192 test trials (12 subjects x 16 test trials) were completed. The order for the pens was different for each of the subjects.

After the steering task was completed, we asked the subjects to rate ease of dragging, degree of fatigue, and to give an overall evaluation on a scale from 1 (worst) to 7 (best).

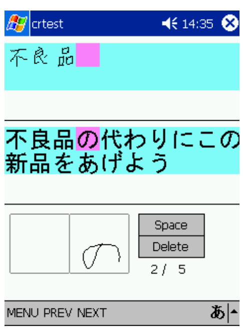

Figure 3: Handwriting task

Handwriting task

The design of the task was based on other studies which determined the optimal size of handwriting character input boxes [Kato et al. 2003]. Figure 3 shows the interface for the experiment. The two boxes on the display use the optimal size (1.44 x 1.44 cm) determined by previous experiments for Japanese handwriting and for alphanumeric input as a baseline. The procedure for writing both alphanumeric and Japanese characters was as follows:

(1) *The target character was displayed*: The target character, which the subject was to input using a pen, was displayed and highlighted in pink in the middle rows. The two character input boxes were displayed on the lower part of the display. The characters actually input into the boxes by the subjects were displayed in the upper section of the display. When the character was successfully input, the next target character would be highlighted in pink.

(2) *Character input*: The subject identified the target character and input the character in the boxes with the pen. The character which had been input was then displayed without recognition in the upper section of the display. A space was inserted between the characters whenever the subject touched the "Space" icon in the lower right of the display. Touching the "Delete" icon had the effect of a backspace key on a keyboard. Subjects used the "Delete" icon to remove any character they wanted to rewrite or correct, e.g. if the character which was written by the subject was an incorrect character. Character recognition was not carried out during the experiment. The recognition rate was derived from the data after the experiment. This procedure was followed so that the subject would not develop stress caused by having to rewrite a character when the wrong character recognition result was displayed.

(3) After the input of all characters was completed, we asked the subjects to rate ease of writing, degree of fatigue, and to give an overall evaluation on a scale from 1 (worst) to 7 (best).

After the procedure of the task was explained to the subject, a practice session with the accessory pen was performed. The subjects were then told to do the real trials. For each of the eleven pens (five pen-lengths + six pen-tip widths/pen-widths), each subject had a total of 141 test trials (72 alphanumeric and 69 Japanese). Each subject completed 705 test trials (Experiment 1) and 846 test trials (Experiment 2). For each pen, 1692 test trials (12 subjects x 141 test trials) were completed. The order for the pens was different for each of the subjects.

RESULTS & DISCUSSIONS

According to the results of Experiment 1, there was no significant difference in movement time or error rates in any of the three tasks, however, the *IP* values showed that the 9 cm and 11cm pens were better than the other pens in the pointing task; the 11 cm pen was better than the other pens in the steering task. Moreover, the subjective evaluations showed that the pens more than 11cm in length received high scores and the subjects preferred the pen-length to be 11cm or more.

Taking these results and considerations together, a length of 11cm pen can be regarded as the best choice to use a PDA.

Thus, we used the 11 cm long pen to test pen-tip width combined with pen-width in Experiment 2. According to the results of Experiment 2, there was no significant difference in movement time in each of the three tasks, however, ANOVA results showed that the 0.5mm/7mm and 1.0mm/7mm pens were better than 1.0mm/7mm in error rate in the pointing task. Moreover, the *IP* values showed that 0.5mm/7mm pens were better than the others in the pointing task, and 1.0mm/7mm and 0.5mm/4mm pens were better than the others in the steering task.

Furthermore, the subjective evaluation and reactions showed that they preferred the 0.5mm/7mm pen in the three tasks.

Taking these results and considerations together, a pen-tip width of 0.5 mm and a pen-width of 7 mm can be regarded as the best choice for use with a PDA.

CONCLUSION

This study investigated the effects of pen size on user performance through two experiments based on the pointing, steering, and handwriting tasks that PDA users perform in their daily lives.

The results of the experiments show the dimensions of the pen affect user performance a little but they affect user preferences quite significantly. There were no significant differences in performance between the sizes of the pens in most comparisons. This is due to the fact that the various dimensions of the pens designed by us only had a small range, but the *IP* values and the subjective evaluations provided useful information which user interface designers can refer to for PDA pen design.

Taking Experiments 1 and 2 together, we determined that the most suitable dimensions are as follows: pen length 11cm, pen-tip width 0.5mm, and pen width 7mm. We believe the findings of this study provide a base point for further research in this field with a view to the development of more useful pens. They will further enable handheld designers to design stylus pens for handheld devices which will offer users greater comfort and greater efficiency.

The data that were collected in mobile situations also support the relationship between performance and task difficulty as proposed by [Accot and Zhai 1999].

We are still working on this issue and are only in the very early stage. The limitations of the conclusions are quite normal for most laboratory-based studies. Many essential issues shall be considered such as rigidity, shape of stylus pen, center of gravity – the weighting and balance of pens, surface friction, finger and hand size, subject posture, age issues and so on. There are many challenges for future study to make this kind of research complete. Experiments under other conditions should be conducted. We would like to emphasize that the conclusions on the sizes (length, width, tip-width) of the pen were based upon the pens which we designed. We expect different conclusions once the designers test different variables of the pen. These will have a valuable impact on studies regarding the physical aspects of the pen.

REFERENCES

1. J. Accot and S. Zhai, Beyond Fitts' Law: Models for Trajectory-Based HCI Tasks, *Proc. CHI'97*,1997, 295-302.
2. J. Accot and S. Zhai, Performance Evaluation of Input Devices in Trajectory-based Tasks: An Application of Steering Law, *Proc. CHI'99*, 1999, 466-472.
3. J. Accot and S. Zhai, Scale effects in steering law tasks, *Proc. CHI 2001*, 2001, 1-8.
4. K.M. Baird, E.R. Hoffmann, and C.G. Drury, The effects of probe length on Fitts' law. *Applied Ergonomics*, 33, 2002, 9-24.
5. S.A. Brewster, Sound in the interface to a mobile computer, *Proc. HCI International'99*, Lawrence Erlbaum Associates, NJ, 1999, 43-47.
6. J.T. Dennerlein, D.B.Martin and C.Hasser, Force-feedback improves performance for steering and combined steering-targeting tasks, *Proc. CHI2000*, 2000, 423 - 429.
7. P.M. Fitts, The information Capacity of Human Motor Systems in Controlling the Amplitude of a Movement, *Journal of Experimental Psychology*, 47(6), 1954, 381-391.
8. T. Kato, X. Ren, N. Sakai, and Y. Machi, The optimal sizes of input squares for the pen-input characters on PDAs, *Human-Computer Interaction - Theory and Practice, 2,* 2003, 686-690.
9. I.S. MacKenzie, Fitts' law as a research and design tool in human-computer interaction. *Human-Computer Interaction*, *7*, 1992, 91-139.
10. I.S. MacKenzie, A. Sellen, and W. Buxton, A Comparison of input devices in elemental pointing and dragging tasks. *Proc. CHI'91*, 1991, 161-166.
11. S. Mizobuchi, K. Mori, X. Ren, Y. Yasumura, An empirical study of the minimum required size and the minimum number of targets for pen input on the small display, *Proc. MobileHCI 2002*, 2002, 184-194.
12. X. Ren, and S. Moriya, Improving Selection Performance on Pen-Based System: A Study of Pen-Based Interaction for Selection Tasks, *ACM Transactions on Computer-Human Interaction (ToCHI)*, 7(3), 2000, 384-416.
13. S. Zhai and R. Woltjer, Human Movement Performance in Relation to Path Constraint - The Law of Steering in Locomotion, *Proc. of IEEE Virtual Reality 2003*, 2003, 149-156.