An Investigation on Selection Mechanisms for Mobile Camera-based Cursor Manipulation

Liang Chen, Dongyi Chen, Shiji Xiahou and Jie Li

School of Automation Engineering,
University of Electronic Science and Technology of China,
Chengdu, 611731, P.R China
liangchen0513@gmail.com, dychen@uestc.edu.cn, xiahousj@uestc.edu.cn,
li_jie_world@sina.com

Abstract

To avoid the occlusion problem and the fat finger problem while manipulating mobile touch devices, researchers in HCI community have explored the area of mobile camera-based interaction which utilizes camera input instead of touch input. For mobile camera-based cursor manipulation, selection mechanisms may have an effect on users' pointing performance, and therefore worthy of study. In this research, we conducted a user study to explore how selection mechanisms influenced users' performance in mobile camera-based cursor manipulation and users' preferences. The quantitative results indicate that the users' performances of using two selection mechanisms (the button-based selection and tapping-based selection) are similar to each other from a statistical point of view. However, subjective feedback from the participants shows that the tapping-based selection is believed to be more comfortable and easy to use.

Keywords: Mobile Devices, Mobile Camera-based Interaction, Cursor Manipulation, Pointing Tasks, Selection Mechanism

1. Introduction

Nowadays mobile touch devices have already been ubiquitously used and become an indispensable part of people's everyday lives. In spite of many compelling merits of direct touch input, it has some conspicuous drawbacks which induce users' dissatisfaction when they interact with touch screens. The two most notable problems in direct touch manipulation are the occlusion problem and the fat finger problem [1].

To address these two problems, a number of novel techniques have been devised and implemented in the past decade. Vogel and Baudisch proposed Shift [2] which addressed these problems by displaying the occluded contents, which were actually beneath the fingertip, in a callout placed in a unblocked location on the screen. Other researchers investigated on the area of back-of-device interaction which allows users to manipulate mobile devices from the backside, thus resolving both occlusion and "fat fingers". NanoTouch [3], LucidTouch [4] and RearType [5] are typical techniques in this type of interaction.

Besides techniques mentioned in previous paragraph, mobile camera-based interaction can solve the abovementioned problems as well. Just as its name implies, mobile camera-based interaction makes use of video input captured by inbuilt or mounted cameras for manipulation. Since inbuilt cameras have already been standard components integrated in smart phones and tablet computers, it is more feasible to implement camera-based interfaces on mobile devices as a replacement or supplement for direct touch manipulation to combat occlusion and "fat fingers" in comparison to back-of-device interfaces.

The selection mechanism is an important component of cursor manipulation. Different selection mechanisms may have an important impact on users' pointing performance or preference. In this paper, we present both quantitative and subjective results of our user study in which we explored the selection mechanism's effect on users' pointing performance in conducting Fitts' reciprocal pointing tasks [6]. The results indicated that although there were no significant differences in users' pointing performance between the button-based selection and tapping-based selection, the latter selection mechanism was believed to be more comfortable and easy to use while conducting mobile camera-based cursor manipulation.

2. Related Work

Since our research concentrates on making use of the rear-facing camera for daily mobile device manipulation, various studies on augmented reality or utilizing the front-facing camera are not closely related to our current work.

Several researchers investigated on generating mobile camera-based manipulation by detecting the movement of the mobile device. In this interaction style, a user moves the device to manipulate virtual contents on the display. For example, Rohs [7] realized such a prototype which detected device movement by tracking 2-dimensional barcodes. Hansen *et al.* [8] presented a similar device movement detection technique by tracking a printed or hand drawn circle near the user and in front of the camera. Instead of tracking a specific object, Wang *et al.* [9] and Haro *et al.* [10] respectively proposed device movement detection method by continuously analyzing the differences in the incoming images captured by the camera.

Unlike the abovementioned mobile camera-based interfaces manipulated by moving the device in a user's hand, other researchers investigated on how to manipulate mobile devices by moving an object in front of the camera. Hachet and colleagues [11] presented an interface which tracked the movement of a rectangle target to interact with 2D or 3D virtual contents on a mobile device's display. Parlke *et al.* [12] made use of kicking actions detected through camera input to conduct mobile device menu manipulation. Gallo *et al.* [13] and Baldauf *et al.* [14] respectively presented mobile camera-based interfaces using fingertip detection techniques.

3. Selection Mechanisms for Mobile Camera-Based Cursor Manipulation

With mobile camera-based input, cursor manipulation, which is widely used in desktop environment, can be enabled on mobile devices. There are two major components in cursor manipulation: pointing and selection. In this paper we focus on investigating on selection mechanisms for mobile camera-based cursor manipulation.

There are several selection mechanisms which can be utilized to invoke selection events for mobile camera-based interfaces, e.g. depressing a button on the device or tapping on the front touch display. Camera-based input can also be utilized to generate selection events, for example making use of dwelling (keeping the cursor above the target for a short time, say 1 second), making clicking actions in the air, and so on.

Do selection mechanisms affect human performance in mobile camera-based cursor manipulation? Or among the many selection mechanisms which one is more preferred by users? In order to answer these questions, we conducted the following user study.

4. User Study

In this study we explored selection mechanisms for mobile camera-based cursor manipulation. We were especially interested in which selection mechanism was more efficient and/or accurate for mobile camera-based cursor input when performing Fitts'

reciprocal pointing tasks. We also wanted to know the participants' subjective preference between the two selection mechanisms.

In our study, the cursor was controlled by tracking a finger-worn color marker in absolute mode. There were two selection mechanisms that we intended to compare in this study - button-based selection and tapping-based selection. The method of dwelling was ruled out due to its apparent deficiency – too time-consuming for committing a selection event. Another camera-based method, clicking in the air, was also ruled out, since we observed in our pilot study that it caused more errors when selecting small target, which was partly due to the relatively low frame rate which our current technique (both hardware and software) could provide.

4.1. Apparatus

The whole experiment was carried out on a HTC One S smart phone running Android 4.0.3 Ice Cream Sandwich. The device possesses a 1.5GHz dual-core Qualcomm Krait CPU, an Adreno 225 GPU, and 1GB RAM. Its inbuilt screen, which is 4.3 inches in size, has a resolution of 960 x 540. The built-in 8-megapixel rear-facing camera was utilized for capturing video input. The software was implemented using Java and Android Software Development Kit.

4.2. Participants

Ten participants (5 men and 5 women) were recruited from within the university community, ranging in the age from 23 to 31 (AVG = 27, SD = 3.13). All participants claimed to be right-handed and had sufficient prior experiences with mobile touch devices, with half of them using their mobile devices more than 4 hours per day. They all volunteered to partake in the study, thus no compensation of any kind was paid to any of them.

4.3. Task and Procedure

We employed the Fitts' reciprocal pointing task since it was a classic task which had been widely used in previous HCI studies [15-17]. During the study, participants were presented with a variety of vertical rectangular targets in pairs, with different target widths and varying distances from each other. In the two displayed targets, the active one was highlighted in green while the inactive target was rendered in red. Participants were instructed to place the onscreen cursor (rendered as a blue circle) onto the active target and then commit a selection as quickly and accurately as they could. In our implementation, button-based selection was triggered by depressing a physical button located on the right side of the device while tapping-based selection was invoked by tapping on the front touch screen by the thumb of the holding hand. Once the active target was successfully selected, it immediately changed into inactive mode while simultaneously the other rectangular targets turned into an active one. If a selection was committed outside the active target, an error would be recorded by the system and an error sound was played to get the participant's attention.

Before the study, a pre-study questionnaire, which was aimed for collecting each participant's basic personal information such as age and handedness, was required to be filled out. After the paperwork, each participant was trained how to manipulate camera-based interfaces - both how to position the cursor and how to commit selections. After that, at least one block of trials was performed for practice. Once the participant claimed he or she had already mastered how to perform acquisition tasks with camera-based interfaces, three blocks of trials were given to them to complete. During the study, participants were allowed to take breaks after each block. Following completion of the

whole experiment, each participant was asked to fill out a post-study questionnaire to answer questions as well as share thoughts on both selection mechanisms.

4.4. Experiment Design

In the pointing tasks, four target amplitudes (A equals 200, 250, 300, 350 pixels respectively) and two target widths (W equals 50 and 100 pixels) were used. The widths were based on the widths of UI elements on smart phones. The values of ID (index of difficulty) of each A-W condition were 2.32, 2.59, 2.81, 3, 1.59, 1.81, 2, 2.17 bits, calculating by equation (1) which is the logarithm portion of the equation (2) for Fitts' law of Shannon formulation [15].

$$ID = \log_2\left(\frac{A}{W} + 1\right) \tag{1}$$

$$MT = a + b \log_2\left(\frac{A}{W} + 1\right) \tag{2}$$

Each participant performed pointing tasks for all eight A-W conditions using both selection mechanisms. The ordering of two selection mechanisms was counterbalanced and trials of each A-W condition in each block were presented in a random order. The design of the experiment could be summarized as:

10 participants   x

2 selection mechanisms (Button-based selection and Tapping-based selection) x

3 blocks   x

8 A-W conditions (ID values from 1.59 to 3 bits)   x

10 trials per A-W condition

= 4800 trials in total.

4.5. Results

3.5.1 Selection Time Analysis: Before any analysis was conducted, trials labeled as outliers and errors were deleted from the collected dataset. Error trials were the ones which participants selected outside the active targets while outliers referred to the ones whose selection time was longer than three standard deviation from the mean selection time of the same participant in one specific selection mechanism.

We analyzed the data using a repeated measures ANOVA. The results demonstrated that there was no significant difference between button-based selection and tapping-based selection on selection time ($F_{1, 9} = 2.389$, $p = 0.157$), with mean selection times of 1.803 seconds and 1.667 seconds respectively. As one would expect from Fitts' law, there was a significant effect for ID on selection time ($F_{7, 63} = 58.039$, $p < 0.001$). There was no significant interaction between selection mechanisms and A-W conditions. Figure 1 demonstrated the mean selection times of both selection mechanisms for each ID. As shown in Figure 1, we could see that although no significant effect for selection mechanisms on selection time was found through data analysis, for each ID the mean selection time of button-based selection was slightly higher than that of tapping-based selection.
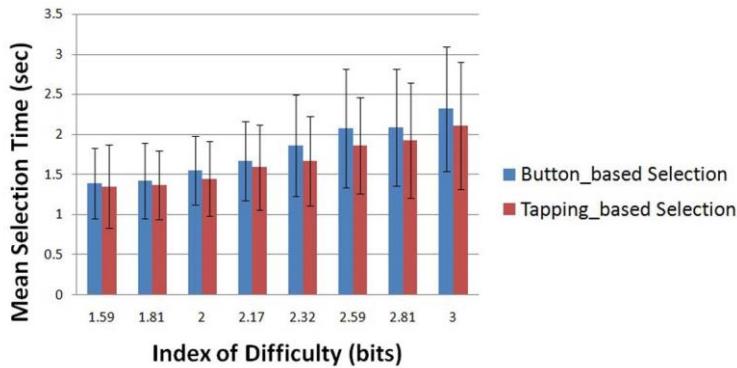
Figure 1. Mean Selection Time for Each ID

3.5.2. Selection Error Analysis: As mentioned above, an error selection was recorded whenever a selection was committed outside an active target. The result of a repeated measures ANOVA of the collected data showed that there was no significant effect existing in selection mechanisms on selection error ($F_{1, 9} = 0.837$, $p = 0.384$), with a mean error rate of 3.33% for button-based selection and 4.20% for tapping-based selection. There was a significant main effect for ID on selection error ($F_{7, 63} = 10.680$, $p < 0.001$). No significant interaction was found between selection mechanisms and A-W conditions. Figure 2 showed the error rates of both selection mechanisms for each ID. As demonstrated in Figure 2, we could see that within the same distance between active and inactive targets, the overall trend is the narrower the target width the higher the error rate.
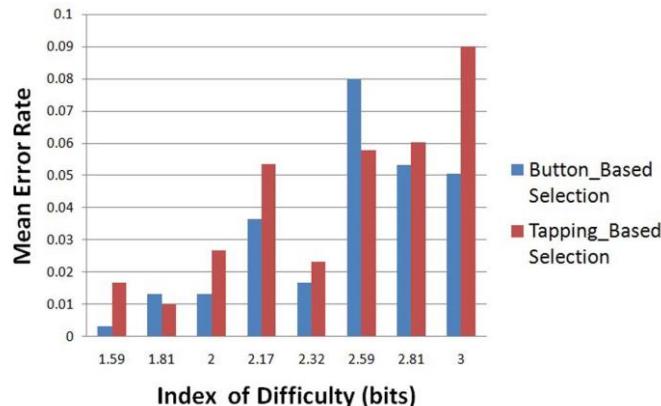
Figure 2. Mean Error Rate for each ID

3.5.3. User Feedback: The collected answers from post-study questionnaires were demonstrated in Figure 3. The corresponding questions asked in the post-study questionnaires were listed as follows:

(1) Which selection mechanism do you prefer?

(2) Which selection mechanism can provide better performance in terms of selection time in your opinion?

(3) Which selection mechanism is easier to use?

(4) Which selection mechanism is more comfortable to use?

(5) Which selection mechanism can get a higher error rate in selecting targets?
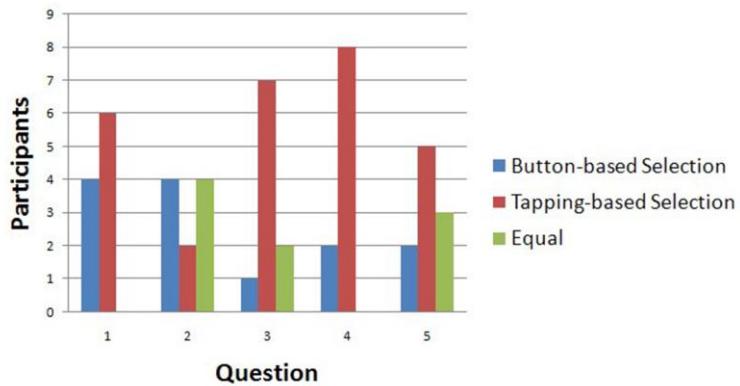
Figure 3. Answers to Questions of Post-study Questionnaires

As shown in Figure 3, the tapping-based selection mechanisms gained a slightly higher preference from our participants, for six people chose this selection mechanism as their preferred one. This result was somewhat in keeping with the quantitative results in which no significant effect was found for selection mechanisms in both selection time and error rate. Two more people chose tapping-based selection was probably because it was more comfortable and easier to use as we could find in the following.

In terms of selection time, two people thought tapping-based selection did a better performance; four participants believed button-based selection worked faster; the other four deemed there was no big difference between the two selection mechanisms. Participants' subjective feeling was almost equivalent to the quantitative results we got from the user study that we could see in Figure1 that for each ID the mean selection time of button-based selection was only slightly longer than that of tapping-based selection.

When asked which selection mechanism was easier to use, 7 participants voted for tapping-based selection while only one participant chose button-based selection. Similarly, 8 people thought tapping-based selection was more comfortable to use while the other two selected button-based selection. One participant supported his choice of tapping-based selection by remarking that "I can hold the phone anyway I like, not having to put my finger on the button". A woman, who also favored tapping-based selection, commented that "felt like my fingers were not long enough for the button". Several other participants mentioned that it took less effort to commit a selection using tapping-based method. So, we could conclude that it was the convenience of tapping anywhere on the front surface which made participants felt tapping-based selection to be easier and more comfortable to use.

When asked which selection mechanism might get a higher error rates, 5 people chose tapping-based selection while only two people selected button-based selection. This qualitative result matched corresponding quantitative results that error rate on tapping-based selection was a little higher than that of button-based selection. One participant mentioned that button-based selection had an advantage of offering the user a haptic feedback when the user pressed down the button. Another participant found that when using tapping-based selection, an error might be caused by accidentally double tap on the display.

4. Discussion and Future Work

Participants' feedback showed that most participants considered tapping-based selection was more comfortable and easier to use. However, we should not neglect the fact that some participants also complained about the position of the button responsible

for committing selection actions. Since we made use of a built-in button, which was designed for adjusting the volume not for invoking a selection, on a off-the-peg phone, perhaps it was the position of the button that affected users' satisfaction. A more ergonomic design of the position of the button on a mobile device may improve user experience. Therefore, a further research should be conducted to determine the most appropriate position for the button which is used for committing selections on a mobile device.

Furthermore, we mentioned in previous section that if future mobile devices possess high-speed cameras and sufficient computing power, vision-based selection mechanisms, *e.g.*, tapping the finger in the air, can be enabled. Since in this case both pointing and selection are controlled by camera input, users may feel more intuitive when interacting with mobile devices.

5. Conclusion

In this paper we presented our user study which was designed to explore selection mechanisms for mobile camera-based cursor manipulation. Our findings indicates that the performance of tapping-based selection is statistically similar to that of button-based selection in terms of both selection time as well as error rate. However, most participants deems that tapping-based selection is more comfortable to use and easier to conduct, a little more error-prone though.

Since quantitative findings indicates that no significant differences exist in efficiency and accuracy between the two selection mechanisms, user interface designers should concentrate more on the tasks themselves when designing a mobile camera-based interface. For example, if the task should avoid any finger occlusion while using, button-based selection should be adopted. If the task can allow a certain extent occlusion over corners of the display, tapping-based selection is the better choice since it is believed to be more comfortable and easy to use.

Acknowledgments

References

[1] K. A. Siek, Y. Rogers and K. H. Connelly, "Fat Finger Worries: How Older and Younger Users Physically Interact with PDAs", Proceedings of the 2005 IFIP TC13 international conference on Human-Computer Interaction, **(2005)**, pp. 267-280.

[2] D. Vogel and P. Baudisch, "Shift: a technique for operating pen-based interfaces using touch", Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, **(2007)**, pp. 657-666.

[3] P. Baudisch and G. Chu, "Back-of-device interaction allows creating very small touch devices", Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, **(2009)**, pp. 1923-1932.

[4] D. Wigdor, C. Forlines, P. Baudisch, J. Barnwell and C. Shen, "Lucid-Touch: A See-Through Mobile Device", Proceedings of the 20th annual ACM symposium on User interface software and technology, **(2007)**, pp. 269-278.

[5] J. Scott, S. Izadi, L. S. Rezai, D. Ruszkowski, X. Bi and R. Balakrishnan, "Reartype: text entry using keys on the back of a device", Proceedings of the 12th international conference on Human computer interaction with mobile devices and services, **(2010)**, pp. 171-180.

[6] P. M. Fitts, "The information capacity of the human motor system in controlling the amplitude of movement", Journal of Experimental Psychology, vol. 47, no. 6, **(1954)**, pp. 381-391.

[7] M. Rohs, "Real-World Interaction with Camera-Phones", Proceedings of the 2nd International Symposium on Ubiquitous Computing Systems, **(2004)**, pp. 74-89.

[8]   T. R. Hansen, E. Eriksson and A. Lykke-Olesen, "Mixed interaction space: designing for camera based interaction with mobile devices", CHI'05 extended abstracts on Human factors in computing systems, **(2005)**, pp. 1933-1936.

[9]   J. Wang, S. Zhai and J. Canny, "Camera phone based motion sensing: Interaction techniques, applications and performance study", Proceedings of the 19th annual ACM symposium on User interface software and technology, **(2006)**, pp. 101-110.

[10] A. Haro, K. Mori, T. Capin and S. Wilkinson, "Mobile camera-based user interaction", Proceedings of the International Conference on Computer Vision Workshop on Human Computer Interaction, **(2005)**, pp. 79-89.

[11] M. Hachet, J. Pouderoux and P. Guitton, "A camera-based interface for interaction with mobile handheld computers", Proceedings of the 2005 symposium on Interactive 3D graphics and games, **(2005)**, pp. 65-72.

[12] V. Paelke, C. Reimann and D. Stichling, "Kick-up menus", CHI '04 extended abstracts on Human factors in computing systems, **(2004)**, pp. 1552-1552.

[13] O. Gallo, S. M. Arteaga and J. E. Davis, "Camera-based pointing interface for mobile devices", Proceedings of the 15th IEEE International Conference on Image Processing, **(2008)**, pp. 1420 - 1423.

[14] M. Baldauf, S. Zambanini, P. Fröhlich and P. Reich, "Markerless visual fingertip detection for natural mobile device interaction", Proceedings of the 13th international conference on Human computer interaction with mobile devices and services, **(2011)**, pp. 539-544.

[15] I. S. MacKenzie, A. Sellen and W. Buxton, "A comparison of input devices in element pointing and dragging tasks", Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, **(1991)**, pp. 161-166.

[16] C. Forlines, R. Balakrishnan, "Evaluating tactile feedback and direct vs. indirect stylus input in pointing and crossing selection tasks", Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, **(2008)**, pp. 1563-1572.

[17] J. O. Wobbrock, B. A. Myers, H. H. Aung, "The performance of hand postures in front- and back-of-device interaction for mobile computing", Int. J. Human-Computer Studies, vol. 66, no.12, **(2008)**, pp. 857-875.

Authors

Liang Chen, He is now a doctoral student, majored in Computer Science, in School of Automation Engineering at University of Electronic Science and Technology of China (UESTC). His current research interest focuses on human-computer interaction, especially on prototyping and evaluating mobile or wearable devices and interfaces. He used to work at Department of Computer Science at University of Toronto as a visiting researcher from 2012 to 2014.

Dongyi Chen, He is a professor in School of Automation Engineering at UESTC where he directs Mobile Computing Center. He used to work at University of Toronto, Georgia Tech and some other foreign universities. His research interest includes ubiquitous computing, mobile computing, wearable computing, WSN, augmented reality, and so on.

Shiji Xiahou, He is an associate professor at University of Electronic and Science Technology of China, IEEE member, whose primary research involves wearable computing, sensor network, and human-computer interaction, with funding from the international science and technology cooperation (China-Canada and Sino-German) in the past five years.

Jie Li, He is a master student at UESTC. He mainly conducts research in mobile computing and wearable computing. His current project focuses on interactive techniques with head mounted displays.