Elderly User Evaluation of Mobile Touchscreen Interactions

Masatomo Kobayashi[1], Atsushi Hiyama[2], Takahiro Miura[3],
Chieko Asakawa[1], Michitaka Hirose[2], and Tohru Ifukube[4]

[1] IBM Research – Tokyo, 1623-14 Shimo-tsuruma, Yamato-shi, Kanagawa 242-8502, Japan
{mstm,chie}@jp.ibm.com
[2] Graduate School of Information Science and Technology, The University of Tokyo,
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan
{atsushi,hirose}@cyber.t.u-tokyo.ac.jp
[3] Research Center for Advanced Science and Technology, The University of Tokyo,
4-6-1 Komaba, Meguro-ku, Tokyo, 153-8904, Japan
[4] Institute of Gerontology, The University of Tokyo,
7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-8656, Japan
{miura,ifukube}@human.rcast.u-tokyo.ac.jp

Abstract. Smartphones with touchscreen-based interfaces are increasingly used
by non-technical groups including the elderly. However, application developers
have little understanding of how senior users interact with their products and of
how to design senior-friendly interfaces. As an initial study to assess standard
mobile touchscreen interfaces for the elderly, we conducted performance meas-
urements and observational evaluations of 20 elderly participants. The tasks in-
cluded performing basic gestures such as taps, drags, and pinching motions and
using basic interactive components such as software keyboards and photo view-
ers. We found that mobile touchscreens were generally easy for the elderly to
use and a week's experience generally improved their proficiency. However,
careful observations identified several typical problems that should be ad-
dressed in future interfaces. We discuss the implications of our experiments,
seeking to provide informal guidelines for application developers to design bet-
ter interfaces for elderly people.

Keywords: Mobile, Smartphones, Touchscreens, Gestures, Aging, Elderly,
Senior Citizens, User Evaluation.

1 Introduction

Mobile phones are filling essential roles in today's societies both for the elderly and
younger groups. The Ministry of Internal Affairs and Communications in Japan re-
ported that approximately 75% of people aged six or older had their own mobile
phones in 2009 [1]. At the same time, more than 70% of people aged 60-69 and 40%
of people aged 70-79 used mobile phones.

It is known that senior citizens tend to use their mobile phones for relatively lim-
ited purposes (e.g., [2]). The limiting factors include hard-to-see displays, hard-to-

press buttons, and hard-to-learn procedures. Special mobile phones for the elderly (such as the *Raku-Raku Phone* from NTT DoCoMo [3]) have been developed and are widely used, and they provide such features as simplified interfaces with larger buttons. However, their limited functionality may increase the digital divide between the older and younger generations.

Touchscreen-based smartphones have the potential to address these problems. First, the finger-based intuitive interactions can be useful regardless of the age of the user. Second, touchscreen-based interfaces allow us to offer senior-friendly interfaces through software-level adaptations. Senior citizens can use the same hardware devices as younger people and also access online services and communities. This increases the social inclusion of the elderly, improving the quality of their lives and making a more sustainable society by benefiting from the power of senior citizens.

The current penetration rate of touchscreen-based smartphones for the elderly is estimated to be low. However, in the next few years they are expected to become the majority of mobile phones, since many of the new mobile phones are smartphones and the major smartphone operating systems all support touchscreens. The market of smartphone applications for senior citizens is expected to grow due to the rapidly increasing numbers of elderly mobile phone users in developed countries. Our informal preliminary survey supports this expectation, indicating that at least some senior citizens, even those who usually feel hesitation about new technologies, were interested in the latest touchscreen-based smartphones and enjoyed using gestures on their screens. Given the explosive growth of mobile phones including smartphones (even in emerging countries [4]), senior-friendly smartphone applications will be strongly desired throughout the world.

A major problem is that application developers currently have little understanding of how to design better touchscreen interfaces for elderly users since the de facto standards of basic operations on touchscreen-based smartphones, which consist of tapping, dragging, and pinching, have only been adopted in the last few years. We need frameworks and guidelines supported by empirical evidence to help develop senior-friendly interfaces. As an initial step, we need to know how the elderly interact with current touchscreen interfaces as a basis for the development of frameworks and guidelines.

We observed and measured the actions of elderly people using touchscreen smartphones. Based on previous research on senior citizens' use of traditional button-based mobile interfaces (e.g., [5][6]) and gesture-based commands on touchscreens (e.g., [7][8]), we focused on typical finger-based operations on standard devices to assess the practical use of mobile touchscreen interfaces. The goal of this study is to determine the trends and problems of mobile touchscreen interfaces that elderly users actually encounter. The tasks included: (1) controlling basic operations with gestures such as taps, drags, and pinching motions, for which we measured task completion times, analyzed their behaviors while making the motions, and asked about the users' preferences; and (2) using interactive components operated with basic gestures such as software keyboards and photo viewers, for which we simply observed their behaviors and asked for user comments. Based on the quantitative and qualitative results, we try to provide informal guidelines for application developers to design senior-friendly interfaces. Our initial experiments also suggested future research directions and new experiments.

The remainder of this paper is organized as follows. First, in Section 2 we summarize related work. Next we describe the methods of our experiments. Then we present the results of the performance measurements and review the observational analyses and the subjective feedback. We also discuss the implications of the experiments and possible future research. Finally, we summarize this work.

2   Related Work

Much research has been done to evaluate the usability for elderly people of desktop or laptop user interfaces [9][10]. Guidelines for designing accessible Web user interfaces for elderly people have been proposed [11]. In contrast, developing design guidelines for the elderly using such new interfaces as mobile terminals with touchscreen interfaces is ongoing work. Design guidelines for general touchscreen-based mobile interfaces have only recently been adopted and guidelines that consider elderly users are yet to be investigated [12].

There have been a number of user evaluation studies of touchscreen interfaces. Leonardi et al. designed a tabletop touch panel interface and found that the direct interaction metaphor was easy to understand and had a pleasing effect that attracted and motivated elderly participants in their study [13]. Lepicard et al. found two-handed touchscreen input was difficult for elderly users [14]. Stößel et al. compared old users to young users in 42 different gesture inputs for touch surfaces and measured their speed and accuracy. They found that older users are a little slower but there was no significant difference in accuracy and suggested that older adults favor accuracy over speed [7].

Many studies of elderly users and mobile terminals have also been conducted. Seik et al. compared the performance using PDA applications between younger adults and older adults. Their results showed that both older and younger participants performed at the same level [5]. Darroch et al. examined the preferred font sizes on PDA screens, comparing older and younger adults. There were no significant differences in reading performance and accuracy between the older and younger adults, but the preferred size of the font was slightly larger for the older participants [15]. Kurniawan investigated the problems that older people face when using mobile phones and assessed some characteristics of a mobile phone for the elderly [16]. Older people are relatively passive adopters with fears of the consequences of using unfamiliar technologies, such as reduced face to face communications, or of accidents caused by careless use, such as by talking while driving. They like functions that support their declining functional capabilities.

Usability studies for new smartphones with touchscreen interfaces have just begun. Stone proposed a special free text input method for elderly people [17]. There have been some accessibility-related evaluations of mobile touchscreens for visually-impaired or motor-impaired users [18][19]. Stößel et al. compared older users and younger users in how they interacted with touchscreens including multi-touch systems. They focused on symbolic gestures and direct manipulations. Their results showed that older users prefer direct manipulations. There were no age-related differences in direct manipulations, but considerable age differences in the use of symbolic gestures. Symbolic gestures are relatively more accepted by older users. Also, they

found no age-related differences in single-finger gestures, but younger users were more likely to use double-fingers gestures [8].

These studies give us valuable information regarding various perspectives of touchscreen interfaces for elderly people. However, few studies have focused on the up-to-date, de facto standard set of touchscreen operations, which consists of simple gestures such as tapping for selecting items, dragging for scrolling, and pinching for zooming. To derive general guidelines for designing senior-friendly interfaces on a common device, we assessed the elder's trends in performance and behaviors in the basic set of touchscreen operations.

3 Methodology

3.1 Participants

Twenty elderly Japanese in their 60s and 70s (14 females and 6 males) were paid to participate in the experiments. Their profiles are summarized in Table 1.

All of the participants had at least two years of prior experience with personal computers, except for one participant whose experience was not recorded (P15). Twelve of the participants had at least two years of prior experience with mobile phones, one was not a mobile phone user (P9), and two did not report their mobile phone experience (P1 and P15). Four of the participants had up to two years of experience with touchscreen-based mobile devices (P1, P4, P5, and P14). Given their relatively longer experience with traditional information technologies, the results of our experiments are expected to reflect the problems specific to mobile touchscreen interfaces rather than those due to the lack of general IT skills and knowledge.

Only one participant was left-handed (P18), but she interacted with the touchscreen primarily with her right hand, as did all of the right-handed participants. Some stated that they had difficulties in their daily lives due to age-related problems with their eyes (P1, P4, and P9) or age-related hearing loss (P5).

3.2 Apparatus

We used an iPad (the "large device" with a 9.7-inch, 147.8×197.1 mm multi-touch screen with 768×1024 pixel resolution, weighing 680 g) and an iPod touch (the "small device" with a 3.5-inch, 49.3×74.0 mm multi-touch screen with 640×960 pixel resolution, weighing 101 g) running iOS 4.2 in the experiments. The large device was enclosed in a case while the small device was used without a case. The experimental software described below was implemented as a native application for the small device running in the low resolution mode (320×460 pixels without a status bar at the top of the screen) at 163 ppi. On the large device, it was run in the small device emulation mode (double size) at 66 ppi. We tested the small device as a representative of mobile devices that have similar size touchscreens. We included the large device in the experiments believing that the larger touchscreen would be especially helpful for the elderly participants, since they may have visual and motor limitations that would be eased by the larger screen.

The participants sat on a chair in front of a table during the experiments. We asked them to hold the devices in a comfortable manner. As a result, the participants who

used a large device generally put it on the table while the other participants who used a small device generally held it with their left hands.

3.3 Procedure

The participants performed the series of experimental procedures twice, separated by one week, except for one participant (P11) who could not return for the second session. Each session consisted of (a) performance measurements for gesture operations and (b) realistic uses of interface components. Each session took about two hours, including a post-experiment interview. During the one-week period, each participant was asked to practice the gesture operations using the experimental software at least once per day. They were also allowed to freely use the pre-installed applications.

Performance measurements. Each participant was asked to perform four standard touchscreen operations: *tapping*, *dragging*, *pinching without panning*, and *pinching with panning*. Fig. 1 shows images of each task.

Each tapping trial started when a target, a white rounded-square button, appeared at a random location on the screen. It ended when the target was successfully tapped, meaning the finger was briefly placed on and quickly removed from the target. For each trial, the target was randomly selected from three sizes: 30, 50, and 70 pixels. The target became blue as the finger was touching the target. Once a target had been dismissed, the next target appeared until the participant had completed 30 tapping trials. Before the timed trials, the participant was allowed to have some practice trials.

We used 30, 50, and 70 pixels as the typical sizes of the keys in a software keyboard, general buttons, and icons on the home screen, respectively. We controlled the logical size in pixels, instead of the physical size, aiming to identify the problems that elder users may face in the actual use of standard mobile interfaces.

Each dragging trial started when an image (200×200 pixels) appeared at a random location on the screen. It ended when the image was moved into the target using dragging motions. The target was a green square (also 200×200 pixels) with 20-pixel-wide borders, which was always displayed at the center of the screen. The width of the borders of the target rectangle represented the tolerance, so errors up to 10 pixels in any direction were permitted for the final image location when the trial ended. The participants had some practice trials and 30 timed trials, just as with the tapping trials.

Each pinching-without-panning trial started when a variable-size square image appeared at the center of the screen. It ended when the image was adjusted to match the target by using two-finger pinching or spreading motions. The initial image was 50, 100, 300, or 400 pixels. The target was the same 200-pixel target used in the dragging trials. The final image size could range from 190 to 210 pixels. The participants again had some practice trials but only 20 timed trials.

Each pinching-with-panning trial started when a square image appeared at a random location on the screen. It ended when the image's size and location had been adjusted to match the same target rectangle. This required two-finger pinching and spreading motions and one- or two-finger dragging. The initial image sizes and the target size were the same as in the "without panning" condition. The participants performed some practice trials and 20 timed trials.
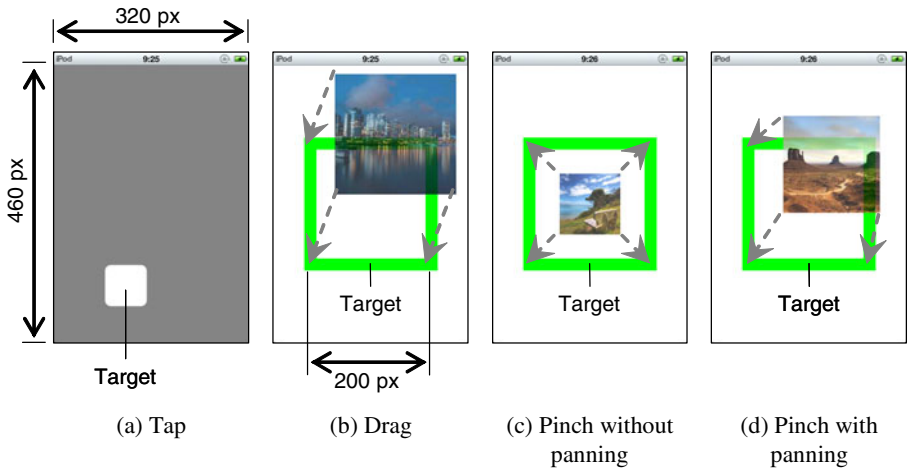
Fig. 1. Our experimental software was designed to practice and assess four standard touch-screen operations: (a) tapping, (b) dragging, (c) pinching without panning, and (d) pinching with panning. The text, lines, and arrows are annotations, not displayed in the actual experiments.

Realistic use. The participants were asked to use two standard interactive components: *photo viewer* and *software keyboard*. We chose these two components as representative of fundamental tasks in online communications: browsing and entering information. Also, they require the use of all of the tested gestures. First each participant was briefly shown how to use each component. Then they were asked to use the component for about 20 minutes with informal tasks such as finding photos that contain more than ten people and copying some printed reference sentences. We also briefly introduced them to some frequently used applications such as cameras, music players, e-book readers (for magazines and newspapers), and games.

3.4 Design

For the performance measurements, the *device* (large or small device) was a between-participant factor, where P0-P9 used the large device and P10-P19 used the small device. The *week* (1 or 2) was a within-participant factor. Other within-participant factors included the *target size* for tapping tasks and *initial size* for pinching tasks. We measured the *task completion time* as the primary dependent variable. We also investigated the participants' behaviors in depth and asked about their preferences.

For the realistic use experiment, we observed the participants' behaviors and then asked them for comments.

3.5 Hypotheses

Here are the main hypotheses we sought to test in this study:

1. Basic touchscreen operations are easy for the elderly to perform without training.
2. A week's practice will improve the performance.

3. The simplest gesture, tapping, is the easiest to perform.
4. Pinching with panning is the most difficult to perform because of its complexity.
5. The larger screen is preferred and more efficient for tapping, because the targets are larger, but the smaller screen is better for dragging and pinching operations, requiring smaller finger movements.
6. For the interactive components, the participants face problems similar to those of traditional user interfaces, such as unclear instructions and unclear indicators of the current state.

Table 1. Twenty senior citizens participated in the experiments. P0-P9 tested a large screen device while P10-P19 tested a small screen device

| # | Age | Gender | Dominant hand | Prior experiences | | | Difficulties in vision/hearing |
				PC	Mobile	Touch-screen	
P0	70s	Female	Right	10 years	6 years	none	-
P1	70s	Male	Right	10 years	-	1 year	Vision
P2	70s	Female	Right	10 years	6 years	none	-
P3	70s	Male	Right	2 years	5 years	none	-
P4	70s	Female	Right	10 years	5 years	once	Vision
P5	70s	Female	Right	8 years	2 years	2 years	Hearing
P6	70s	Male	Right	15 years	3 years	none	-
P7	70s	Male	Right	23 years	15 years	none	-
P8	70s	Female	Right	6 years	3 years	none	-
P9	60s	Female	Right	6 years	none	none	Vision
P10	60s	Female	Right	7 years	8 years	none	-
P11	60s	Female	Right	10 years	10 years	none	-
P12	70s	Female	Right	8 years	3 years	none	-
P13	70s	Male	Right	10 years	5 years	none	-
P14	70s	Male	Right	15 years	8 years	once	-
P15	60s	Female	Right	-	-	none	-
P16	60s	Female	Right	6 years	6 years	none	-
P17	60s	Female	Right	15 years	10 years	none	-
P18	70s	Female	Left	10 years	11 years	none	-
P19	60s	Female	Right	10 years	5 years	none	-

4 Basic Gesture Performances

We found no significant effects on the task completion times due to the age, gender, dominant hand, prior experience, or difficulties in vision/hearing. This does not prove that these factors cannot affect the performance because we did not control for them in this experiment. Their precise effects or irrelevance should be assessed in future research.

4.1 Taps

Fig. 2 shows the average task completion times for tapping operations for each target size. For the first week, the average times were 0.77 seconds and 1.02 seconds for

large and small devices, respectively. For the second week, they were 0.72 seconds and 0.94 seconds. A week's practice reduced the task completion times by 7% and 9% for large and small devices, respectively.

Analysis of variance showed significant main effects on the task completion times due to the device ($F_{1,1158} = 14.43$, $p < .001$) and the target size ($F_{2,1158} = 28.81$, $p < .001$). It also showed a significant interaction effect between the device and the target size ($F_{2,1158} = 8.96$, $p < .001$). There were no significant main effects of the week's practice ($F_{1,1158} = 1.12$, $p > .05$). A *post hoc* analysis found that tapping 30-pixel targets on the small device was significantly more difficult than the other conditions. It took about twice as long to tap the 30-pixel target compared to the 50- or 70-pixel targets.
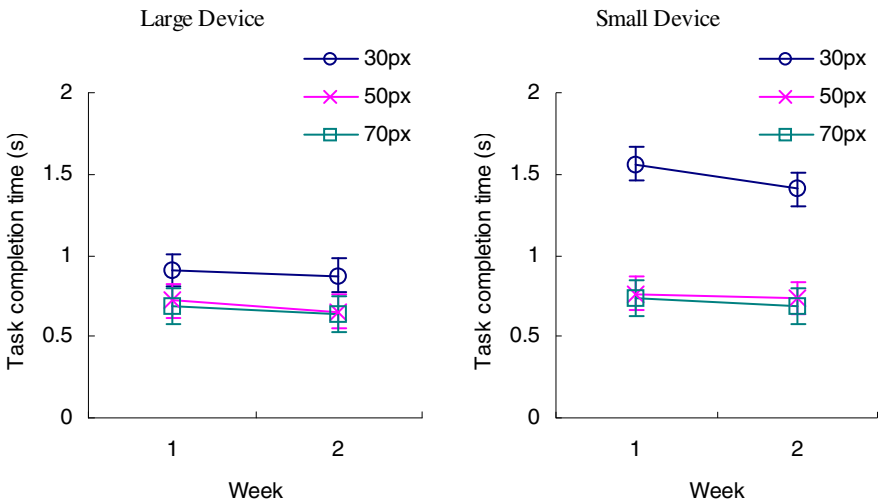

Fig. 2. Tapping task completion times for each target size with standard errors before and after a week's practice

4.2 Drag

Fig. 3 shows the average task completion times for dragging operations. For the first week, the average times were 2.09 seconds and 2.17 seconds for the large and small devices, respectively. For the second week, they were 1.59 seconds and 1.77 seconds. A week's practice reduced the task completion times by 24% and 18% for the large and small devices, respectively.

Analysis of variance showed significant main effects from the week's practice ($F_{1,1138} = 50.83$, $p < .001$) and for the device ($F_{1,1138} = 4.52$, $p < .05$) on the task completion times. There were no significant interaction effects.
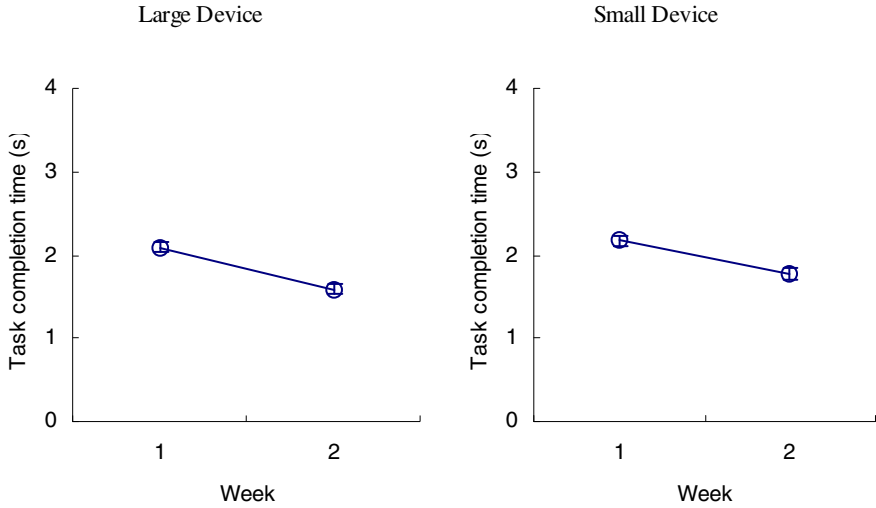
Fig. 3. Task completion times for dragging with standard errors before and after a week's practice

4.3 Pinching and Spreading without Panning

Fig. 4 shows the average task completion times for zooming in and out operations for each initial image size. In the first week, the average times were 2.69 seconds and 2.75 seconds for the large and small devices, respectively. For the second week, they were 1.92 seconds and 1.99 seconds. A week's practice reduced the task completion times by 29% and 28% for the large and small devices, respectively.

Analysis of variance showed significant main effects on the task completion time of the week's practice ($F_{1,753} = 100.15$, $p < .001$) and from the initial size ($F_{3,753} = 20.87$, $p < .001$). It also showed a significant interaction effect between the device and the initial size ($F_{3,753} = 6.51$, $p < .001$). There were no significant main effects of the device ($F_{1,753} = 0.63$, $p > .05$). A *post hoc* analysis found that spreading motions were more difficult than pinching motions, especially on the small device. It took 3.13 seconds to zoom in to resize the image from 50 pixels to 200 pixels while it took only 2.19 seconds to zoom out from 400 pixels to 200 pixels, even though the required amount of finger movement on the screen was larger for zooming out. Some participants commented that spreading (zooming-in) motions were more difficult than pinching (zooming-out) motions, reinforcing these results.

4.4 Pinching and Spreading with Panning

Fig. 5 shows the average task completion times for zooming with panning operations for each initial image size. In the first week, the average times were 4.57 seconds and 5.01 seconds for the large and small devices, respectively. In the second week, they were 3.20 seconds and 3.60 seconds. A week's practice reduced the task completion times by 30% and 29% for large and small devices, respectively. The times were
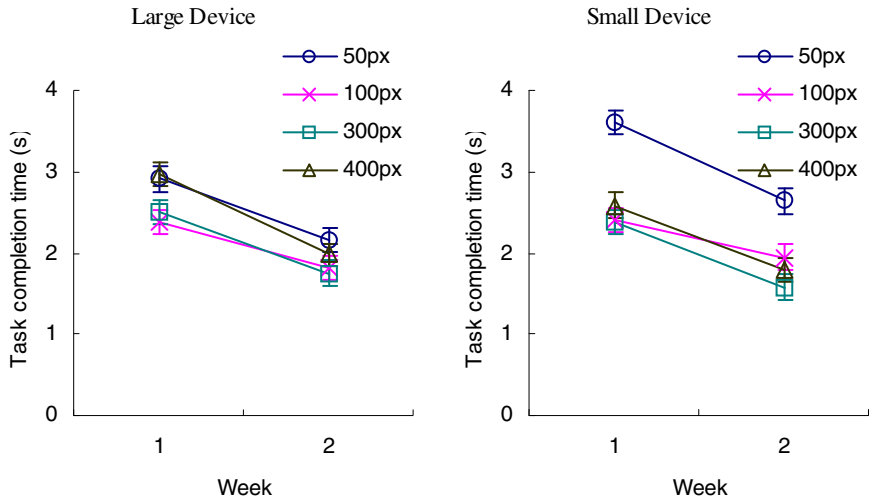
Large Device

Small Device

Fig. 4. Task completion times of pinching without panning with standard errors before and after a week's practice
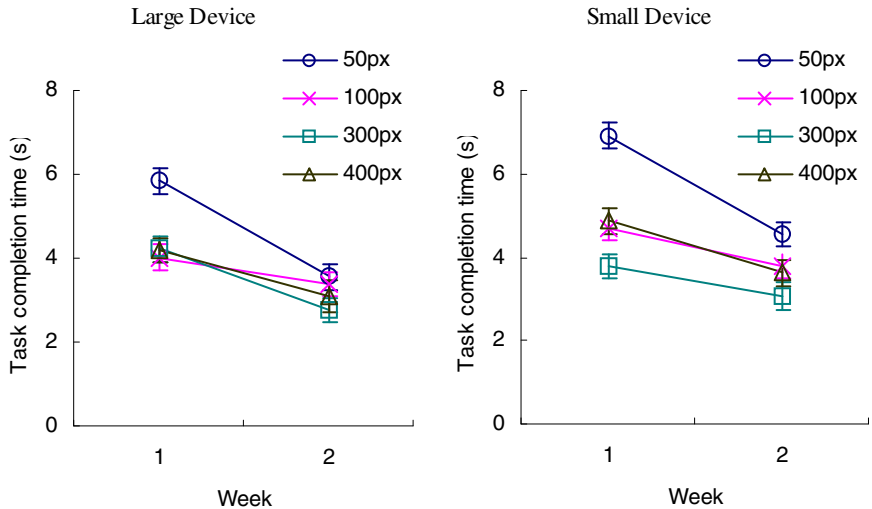
Large Device

Small Device

Fig. 5. Task completion times for pinching with panning with standard errors before and after a week's practice

approximately the total of the times for separately dragging and pinching without panning. This seems to confirm that the participants often used a 2-phase approach, where they first dragged the image to the center of the screen and then zoomed in or out to fit the target rectangle.

Analysis of variance showed significant main effects on the task completion times of the week ($F_{1,738} = 76.71$, $p < .001$), the device ($F_{1,738} = 11.92$, $p < .005$), and the initial size ($F_{3,738} = 18.85$, $p < .001$). It also showed a significant interaction effect between the week and the initial size ($F_{3,738} = 5.09$, $p < .005$). A *post hoc* analysis found that a week's practice greatly improved the performance, especially for zooming in from 50 pixels to 200 pixels (6.38 seconds and 4.06 seconds for weeks 1 and 2, respectively).

5   Observational Findings

5.1   Why Is Tapping a Small Target So Difficult?

To help explain why tapping 30-pixel targets showed much worse performance on the small device, Fig. 6 shows the locations where the participants actually touched in relative coordinates with respect to the target location. The average x-locations were off center by +12.9 ($SD = 10.0$) and +8.28 ($SD = 24.2$) pixels for Weeks 1 and 2, respectively. In contrast, the corresponding values for the y-axis were only +2.0 ($SD = 9.7$) and +1.8 ($SD = 31.2$) pixels. The participants clearly tended to touch the right side of the target. The size and location of the target did not affect this tendency. The gap between intended and actual touch locations could be explained by the fact that all of the participants used their right hands to perform the operations. The gap in the (x, y) range was similar for the 50- and 70-pixel targets on the small device while it was smaller for the large device. The touch locations mostly ranged from (−25, −25) to (+25, +25). As a result, the 30-pixel targets frequently caused errors and increased the task completion times, while the larger targets had few errors. For 30-pixel targets on the small device, the error rate was 39%, where an error means the participant tapped the screen outside of the target at least once during the trial. This error rate was much worse than the acceptable 4% value used in most human-computer interaction (HCI) literature. The error rates were 6.5% and 5.6% for 50- and 70-pixel targets, respectively, on the small device. The respective values were 13.6%, 1.4%, and 1.7% for 30-, 50-, and 70-pixel targets on the large device.

The distribution of touch locations also indicates why the performance was little improved after a week's practice. As shown in Fig. 6, the variance of the locations in Week 2 was actually larger than in Week 1. Our observations and the participants' comments showed that this was caused by a change in the participants' error correction strategy. Due to the gap between the intended and actual touch locations, the participants often tapped outside the target and saw that nothing happened. In that case, in Week 1 they tended to tap the same place until they received the expected feedback. In Week 2, they seemed to change their strategy and tapped at different places, now being aware of the gap between the intended and actual touch locations and trying to correct for it. However, this strategy often resulted in an error by overshooting the target, especially for small 30-pixel targets on the small device (Fig. 7-a). In addition, we found another type of strategy change. In Week 2, some participants tried to vary the speed and pressure by slowing down or speeding up their finger movements if they failed to tap the target on the first attempt. The average contact times with the screen when tapping were 105 ms ($SD = 43$ ms) and 114 ms
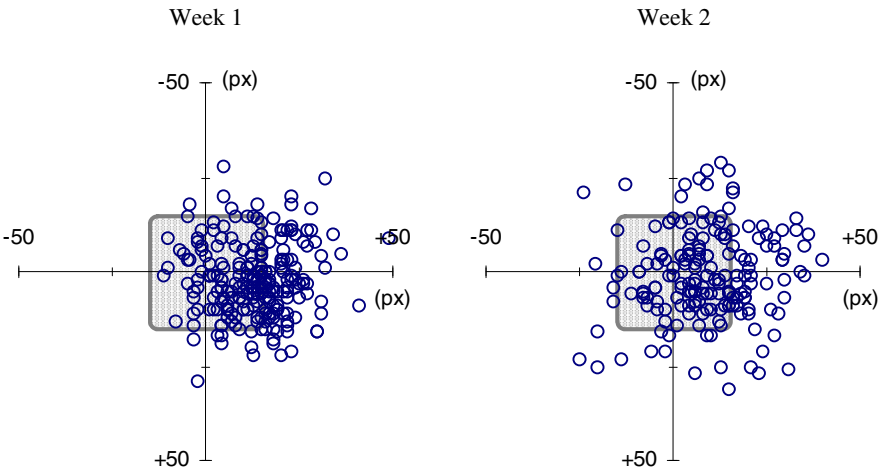
Week 1   Week 2

Fig. 6. The distribution of touch locations relative to the location of the 30-pixel target on the small device before and after a week's practice. The target area is shown as a rounded square. The participants tended to tap the right side of the target. The variation of touch locations was larger after the week's practice.
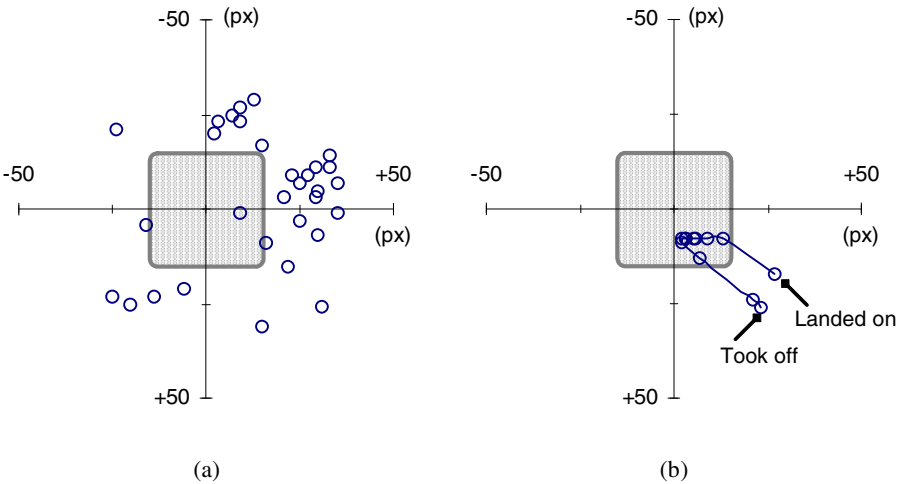

(a)   (b)

Fig. 7. (a) Due to repetitive overshooting, a participant touched the screen 32 times to tap a small target (P12, Week 2). (b) Adding extra pressure (i.e., "pressing") on the screen while tapping can cause an irregular trajectory (P13, Week 2).

(SD = 76 ms) for Weeks 1 and 2, respectively. This increased-pressure strategy often caused other problems. First, an overly slow tap can be recognized as a "hold" operation by many applications. Second, exerting more pressure on the screen during a slow motion can cause the finger to draw a random path, which is then recognized as

a drag (Fig. 7-b). The low visibility of the visual feedback seemed to make these problems even worse, since a dynamic change of the target color was often obscured by the finger when tapping a small (30-pixel) target. Due to these factors, five of the participants (P7 who used the large device and P12, P13, P15, and P16 who used the small device) actually became slower after a week's practice.

5.2 Problems in Applications

The problems with tapping small targets were frequently observed during the use of the practical interactive components. For example, the participants failed to tap the proper key on the software keyboard or failed in moving the cursor to a specific position in a text input area. In addition, we found that the participants faced general user interface problems, which seem to be common with traditional user interfaces.

First, the participants were often confused due to unclear instructions. A typical gesture-based interface on a touchscreen lacks textual labels and menus to show "what is possible with this application". It also provides few explicit instructions for each gesture. Even though an interface is carefully designed so that it is natural and intuitive for users to control many of its functions, the participants tended to stick to the few functions that the experimenters had explained. They were sometimes confused when they invoked an unexpected function with an unintentional gesture. A participant commented that he wanted standard menus as with a personal computer (PC). Though the participants were experienced PC users and familiar with standard QWERTY keyboards, they often could not find the backspace and shift keys on the software keyboard because those keys did not have the textual labels "Backspace" and "Shift".

The participants were also confused when an application has multiple modes without an indication of the current mode. For example, the photo viewer provides multiple thumbnail views such as by folder, by event, and by person, but it was difficult to tell which view was active. The software keyboard was more problematic. Since Japanese needs to input thousands of Chinese-related characters as well as characters using three alphabetic sets, Arabic digits, and symbols, the interface requires frequent mode-switching between predictive input [20], 5-touch input [21], and a QWERTY keyboard. The participants frequently lost track of their current mode, especially when another mode had a similar look-and-feel. One participant explicitly complained that there was no textual indication of the current mode.

6 Subjective Feedback

In contrast to our expectations, the participants commented that dragging and pinching were preferred and easier to do than tapping. In particular, most of the participants stated that dragging was the easiest operation to perform. They found that pinching was also easy, though some reported that pinching with panning was somewhat difficult.

As regards the device, the tradeoff between the size and weight was mentioned. Many participants said that the large device was too heavy to carry while the small device was too small to read. Some female participants wanted a middle-sized, 7-inch device that would fit into their handbags.

When it came to applications, the participants wanted to use various types of applications, such as cameras and photo viewers, newspapers, e-books, and games. Many of them stated that they especially wanted smartphones when they had a few extra minutes, such as when riding the train.

The participants frequently commented that using touchscreen interfaces was "enjoyable". For example:

It is fun to flip through photos [in the photo viewer]. (P5)
I want to do nothing but use this [a touchscreen]. (P14)

These comments confirm the findings of previous research [22]. The enjoyability is one of the most important factors in encouraging the elderly to continue participating in society via information technologies. We need more investigation of the factors influencing why they enjoyed using these devices. Perhaps it is part of the nature of touchscreen-based interfaces or the participants were excited by using new devices.

7   Discussion

The experiments partly confirmed and partly rejected our hypotheses. The first hypothesis was generally confirmed. The participants, even those who had never used touchscreens, performed the gesture-based operations reasonably well, except for tapping on small targets. The second hypothesis was partly confirmed. A week's practice significantly improved the performance for dragging and pinching operations, though there were no significant effects for tapping. The third and fourth hypotheses were rejected. The participants stated that dragging and pinching were easier and more comfortable than tapping. Though some of them found that pinching with panning was a little difficult, their performance was not seriously hindered by the difficulty. The fifth hypothesis was partly confirmed. It was surprising that even for dragging and pinching, as well as for tapping, the larger screen outperformed the smaller screen, even though it required more than twice the amount of finger movement on the screen. However it was reported that the device with the large screen was too heavy to carry and use outdoors. The last hypothesis was also confirmed.

7.1   Design Implications

Based on the experimental findings, we suggest these informal guidelines for application developers seeking to design better interfaces for the elderly.

Use larger targets (8 mm or larger in size). The elderly tend to make errors when tapping a small target, such as a 30-pixel button on a small screen. In the experiment, the touch locations were mostly distributed within 8 mm on the physical screen regardless of the device and the target size. Thus, interactive objects such as buttons, icons, and clickable text should at a minimum be larger than 8 mm. Note that some common components in the latest smartphone interfaces, such as keys in a software keyboard and a "back" button in the upper-left corner of a small device, conflict with this guideline. With targets located close to each other without spacing (e.g., software

keyboard), the size should be much larger to avoid invoking adjacent targets. This guideline confirms the result of previous work [23], which found that the speed and error rates were much worse when tapping buttons smaller than approximately 8 mm.

Address the gap between intended and actual touch locations. The elderly tend to miss their intended targets due to parallax and the large contact area of each finger. Providing a calibration mechanism might be a solution. However, practical calibration is difficult because of the nature of handheld devices. Users frequently tilt and rotate their devices depending on their situations. Another solution might include providing appropriate visual feedback to indicate where the users touched the screen, even when they have missed all of the interactive targets.

Consider using drag and pinch gestures rather than taps. The elderly tend to prefer dragging and pinching operations over tapping. This is a new finding that was not observed in an earlier experiment on multi-touch interactions for elder users [8]. It is also the opposite of the result for motor-impaired users [18]. Though applications need not avoid using drag and pinch operations, the application should provide instructions and clear visual clues to show which gesture invokes which function. Since typical functions invoked by dragging and pinching provide no visual clues, such as flipping and zooming photos, most participants in our experiment could not find those functions without guidance.

Explicitly display the current mode. Some elderly users are less likely to notice changes of the modes and can become confused. Applications should avoid multi-mode interfaces as much as possible. If a multi-mode design is needed, then there should be explicit feedback about mode-changes and a display of the current mode. The indicators of mode changes should be persistent and large, since the elderly users may fail to notice short alerts or small changes in the look-and-feel. Also, the feedback should be readable as they tend to have difficulty in interpreting the meanings of symbolic representations.

7.2   Future Work

A limitation of this study is that we did not have younger participants as a reference group. We focused on investigating how the elderly actually interact with mobile touchscreens and how practice improves their proficiency, as a first step towards guidelines and frameworks to develop senior-friendly interfaces. Hence in-depth analyses of aging effects were beyond our scope. An informal experiment we conducted with two young participants (2 females, 24 and 31 years old), in which they performed touchscreen operations 3%-49% faster than the senior participants, indicated that younger users interact more rapidly with touchscreen interfaces. However, the general trends were similar. For example, considering the result of previous research [23], the minimal target size that a user can tap with good speed and accuracy seems to be around 8 mm for both younger and older people. Additional investigation of the similarity and differences between younger and older users in their performances, behaviors, and preferences would help in designing more universal interfaces that satisfy users of all ages.

In addition to the comparison with younger users' performance, more investigation is needed to understand the trends and problems we encountered. Are they caused by physical and cognitive disabilities such as attention and memory loss related to aging? How do the experiences with other technologies affect performance? A study with more control of the variables such as age, disability, and experience with technologies may discover clues towards more practical solutions for specific user groups.

8   Conclusion

We conducted performance measurements and observational evaluations to assess standard mobile touchscreen interfaces when used by the elderly. Our participants were 20 Japanese in their 60s and 70s. The tasks included (1) controlling basic operations with gestures such as taps, drags, and pinching motions and (2) using basic interactive components such as software keyboards and photo viewers. The results show that touchscreen mobile interfaces are preferred and not too difficult to use, even by the elderly. A week's practice significantly improved the performance in dragging and pinching, but did not significantly affect tapping. We identified several typical problems, such as mismatches between the user's visual target and the touched position as detected by the sensor. We discussed the implications of the experiments on the design of better interfaces for the elderly and considered future research directions. This and future studies will provide the basis for the development of senior-friendly user interfaces.

References

1. Ministry of Internal Affairs and Communications in Japan: Communications Usage Trend Survey (2010)
2. Office for National Statistics in the U.K.: Use of ICT at Home (2007)
3. Universal Design Initiatives, http://www.nttdocomo.com/about/csr/universal_design/
4. ICT Statistics, http://www.itu.int/ITU-D/ict/statistics/ict/
5. Siek, K.A., Rogers, Y., Connelly, K.H.: Fat Finger Worries: How Older and Younger Users Physically Interact with PDAs. In: Costabile, M.F., Paternó, F. (eds.) INTERACT 2005. LNCS, vol. 3585, pp. 267–280. Springer, Heidelberg (2005)
6. Ueda, K., Okochi, N., Chikawa, A., Ito, A., Hiramatsu, Y., Ifukube, T.: Usability Evaluation of Mobile Phone for the Elderly: The Effect of Properties of the Button. In: Human Interface 2009, pp. 673–676 (2009) (in Japanese)
7. Stößel, C., Wandke, H., Blessing, L.: Gestural Interfaces for Elderly Users: Help or Hindrance? In: Kopp, S., Wachsmuth, I. (eds.) GW 2009. LNCS, vol. 5934, pp. 269–280. Springer, Heidelberg (2010)

8. Stößel, C., Blessing, L.: Mobile Device Interaction Gestures for Older Users. In: Proceedings of the 6th Nordic Conference on Human-Computer Interaction (NordiCHI 2010): Extending Boundaries, pp. 793–796 (2010)

9. Arias-Torres, D.: The Design and Evaluation of a Pen-Based Computer Interface for Novice Older Users. In: Proceedings of the 7th Mexican International Conference on Computer Science (ENC 2006), pp. 142–150 (2006)

10. Mahmud, M., Kurniawan, H.: Involving Psychometric Tests for Input Device Evaluation with Older People. In: Proceedings of the 19th Conference of the Computer-Human Interaction Special Interest Group (CHISIG) of Australia on Computer-Human Interaction: Citizens Online: Considerations for Today and the Future (November 21-25, 2005)

11. Web Accessibility and Older People: Meeting the Needs of Ageing Web Users, http://www.w3.org/WAI/older-users/

12. iOS Human Interface Guidelines, http://developer.apple.com/library/ios/documentation/userexp erience/conceptual/mobilehig/

13. Leonardi, C., Albertini, A., Pianesi, F., Zancanaro, M.: An Exploratory Study of a Touch-Based Gestural Interface for Elderly. In: Proceedings of the 6th Nordic Conference on Human-Computer Interaction (NordiCHI 2010): Extending Boundaries, pp. 845–850 (2010)

14. Lepicard, G., Vigouroux, N.: Touch Screen User Interfaces for Older Subjects: Effect of the Targets Number and the Two Hands Use. In: Miesenberger, K., Klaus, J., Zagler, W., Karshmer, A. (eds.) ICCHP 2010. LNCS, vol. 6180, pp. 592–599. Springer, Heidelberg (2010)

15. Darroch, I., Goodman, J., Brewster, S., Gray, P.: The Effect of Age and Font Size on Reading Text on Handheld Computers. In: Costabile, M.F., Paternó, F. (eds.) INTERACT 2005. LNCS, vol. 3585, pp. 253–266. Springer, Heidelberg (2005)

16. Kurniawan, S.: Older People and Mobile Phones: A Multi-Method Investigation. International Journal of Human-Computer Studies 66(12), 889–901 (2008)

17. Stone, R.G.: Mobile Touch Interfaces for the Elderly. In: Proceedings of ICT, Society and Human Beings (July 22-24, 2008)

18. Guerreiro, T., Nicolau, H., Jorge, J., Goncalves, D.: Towards Accessible Touch Interfaces. In: Proceedings of the 12th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS 2010), pp. 19–26 (2010)

19. McGookin, D., Brewster S., Jiang, W.: Investigating Touchscreen Accessibility for People with Visual Impairments. In: Proceedings of the 5th Nordic Conference on Human-Computer Interaction (NordiCHI 2008): Building Bridges, pp. 298–307 (2008)

20. Masui, T.: POBox: An efficient text input method for handheld and ubiquitous computers. In: Gellersen, H.-W. (ed.) HUC 1999. LNCS, vol. 1707, pp. 289–300. Springer, Heidelberg (1999)

21. Macdonald, A.S.: The UD Phenomenon in Japan: Product Innovation through Universal Design. In: Stephanidis, C. (ed.) HCI 2007. LNCS, vol. 4554, pp. 224–233. Springer, Heidelberg (2007)

22. Chikawa, A., Ueda, K., Okochi, N., Ito, A., Hiramatsu, Y., Ifukube, T.: Usability Evaluation of Touch Screen Mobile Phone for the Elderly. In: Human Interface 2009, pp. 677–680 (2009) (in Japanese)

23. Lee, S., Zhai, S.: The Performance of Touch Screen Soft Buttons. In: Proceedings of the 27th International Conference on Human Factors in Computing Systems (CHI 2009), pp. 309–318 (2009)