Bezel Swipe: Conflict-Free Scrolling and Multiple Selection on Mobile Touch Screen Devices

Volker Roth
FX Palo Alto Laboratory
3400 Hillview Ave, Bldg. 4
Palo Alto, CA 94304
vroth@fxpal.com

Thea Turner
FX Palo Alto Laboratory
3400 Hillview Ave, Bldg. 4
Palo Alto, CA 94304
turner@fxpal.com

ABSTRACT
Zooming user interfaces are increasingly popular on mobile devices with touch screens. Swiping and pinching finger gestures anywhere on the screen manipulate the displayed portion of a page, and taps open objects within the page. This makes navigation easy but limits other manipulations of objects that would be supported naturally by the same gestures, notably cut and paste, multiple selection, and drag and drop. A popular device that suffers from this limitation is Apple's iPhone. In this paper, we present Bezel Swipe, an interaction technique that supports multiple selection, cut, copy, paste and other operations without interfering with zooming, panning, tapping and other pre-defined gestures. Participants of our user study found Bezel Swipe to be a viable alternative to direct touch selection.

Author Keywords
Bezel Swipe, touch interaction, multiple selection, mobile device.

ACM Classification Keywords
H.5.2 [Information Interfaces and Presentation]: User Interfaces – Input devices and strategies, Interaction styles.

INTRODUCTION
Mobile touch screen devices are gaining popularity. Apple's iPhone [2] is a prominent example. Even though it targets a premium price segment, iPhone sales are increasing in lower-income levels [6] and a quarter of iPhone users say that it displaces a notebook computer [10]. The iPhone is a phenomenon that is changing the mobile phone landscape, leading other vendors to strive to match its design. Therefore, the iPhone is an important subject for study as an indicator of future trends.

The screens of mobile devices are necessarily small, which limits the size of displayed pages and the space that can be devoted to user interface widgets such as scroll bars. But smaller widgets are harder to hit.

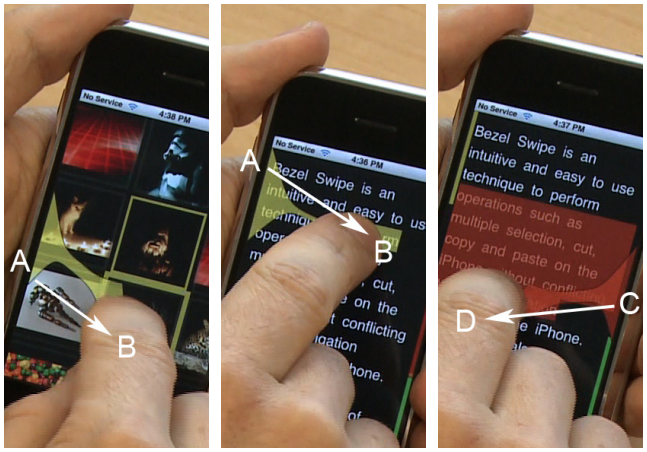

Figure 1. The Bezel Swipe prototype in action. Left panel: A swipe from A to B selects an image. Middle panel: A swipe from A to B marks the beginning of a text selection. Right panel: A swipe from C to D marks the end of the text selection.

Apple worked around these limitations by adopting a zooming interface that is controlled by swiping (for panning) and pinching (for zooming) finger gestures anywhere on the display. Users tap on objects within a page to open them. For example, tapping on a hyperlink loads the referenced page, and tapping on an image thumbnail shows a large view of the image. This design eases navigation tasks at the expense of other operations that would be supported naturally by the same gestures, for example: dragging and dropping of objects within a page, and the selection of objects and page areas for the purpose of cutting and pasting. These operations now require a mode change if we want to identify the object first and then the action (known as *noun-verb* [9] interaction). However, input modes are a significant source of errors and complexity in user interfaces [9].

It is worth noting that at the time of writing the iPhone does not support cut and paste operations. This limitation has been debated in Internet forums, accompanied by speculation on how and when Apple intends to remedy the situation. The developers of the MagicPad text editor [8] for the iPhone used timeouts to enter a text marking mode. However this adds an undesirable delay to user interactions [4] which

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
CHI 2009, April 4 - 9, 2009, Boston, Massachusetts, USA.
Copyright 2009 ACM 978-1-60558-246-7/09/04...$5.00.

Samsung Exhibit 1012, Page 1 of 4

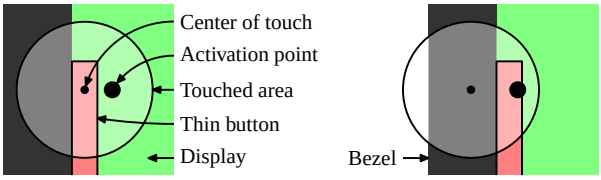

Figure 2. Touching narrow targets close to the bezel: a touch (large circle) centered on the target (left) may not activate it whereas a touch significantly overlapping the bezel (right) does.

may not be tolerable when mode changes are frequent. Alternatively, more complex (and less intuitive) multi-finger gestures could be used but the larger number of fingers also exacerbates the existing display occlusion problems.

To address the disadvantages of mode changes we propose *Bezel Swipe*. Bezel Swipe supports multiple object selection, cutting, copying, pasting and other operations on mobile touch screen devices without conflict with the current panning and zooming gestures.

BEZEL SWIPE

The user starts a Bezel Swipe gesture on the bezel, which is the physical touch *insensitive* frame that *surrounds* the display. The finger is then swiped through a part of the display edge into the display, which enters selection mode. Next, the user moves the finger onto an object or display position and selects it by lifting the finger off the display, which ends selection mode. Different edge portions represent different actions and are distinctively marked. In our prototype, we use a thin colored bar. Single objects, such as image thumbnails, can be selected with one swipe. Text regions can be marked by first selecting the start with one swipe. A second swipe selects the end and performs the action. This two-step process is comparable to the *mark* and *yank* operations in the Emacs text editor, and allows users to select regions that are larger than the display without the need for automatic scrolling (see Fig. 1 for illustrations).

Bezel Swipe combines *verb-noun* [9] interaction (the gesture starts at the action and ends at the object) with a crossing gesture to affect a temporal mode change. The mode change ends when the gesture ends and therefore users cannot "forget" to leave the Bezel Swipe mode and accidentally perform Bezel Swipe operations outside their locus of attention. While a Bezel Swipe is in progress, we keep it at the user's locus of attention by means of graphical feedback. In our prototype, we used a colored line that stretches from the activation bar to the user's fingertip on the display. These measures avoid disadvantages that verb-noun interactions may exhibit over noun-verb interactions [9].

Swipes are recognized by detecting touches inside a thin rectangular area positioned flush against the display edges. Bezel Swipe does not typically reduce the accessible display area because the swiping gesture minimizes the area of initial contact with the touch sensitive display. Figure 2 illustrates this. A touch sensitive display translates the contact area of the finger with the display into an *activation point.* The activation point represents the area center. The left side of Fig. 2 illustrates the contact area for a touch that is centered on a narrow target (the red bar) positioned flush against the display edge. Since the bezel is touch insensitive the activation point is not the center of the touch (small dot) but is instead offset to the right (large dot) and does not actually fall into the narrow bounds of the target. The Bezel Swipe gesture, on the other hand, starts on the bezel and moves into the display. Therefore the contact area is gradually increased until it exceeds the contact threshold of the display. The activation point for the initial contact is therefore farther to the left than the activation point of a tap that is "right on" the target. Bezel Swipe therefore leverages display pixels that are not typically touched during regular interactions.

RELATED WORK

The Bezel Swipe gesture is similar to crossing-based interfaces [1]. However, in contrast to crossing interfaces, Bezel Swipe requires a non-zero activation area and is therefore more comparable to a button. In fact, no crossing is detected. Instead, the crossing of the touch-insensitive bezel in conjunction with the properties of touch sensitive displays enables the reliable activation of a button that is so narrow that it would be difficult to activate by a direct touch.

Bezel Swipe differs from crossing interfaces, and most other interaction techniques, by following an verb-noun interaction design instead of the more common noun-verb design. This role reversal allows users to begin their gestures at relatively unoccupied display areas and move towards the display interior where objects are displayed and navigation gestures originate. Shizuki et al. [11] describe a related technique that allows a user to control a slide presentation or a simple drawing tool with a laser pointer. For instance, a swipe from the right outside of the slide to the inside selects the next slide. The movement of the laser pointer with respect with the slide projection is tracked using a camera.

Bezel Swipe could be regarded as the opposite of *Barrier Pointing* as suggested by Froehlich et al. [5]. Whereas Bezel Swipe relies on flat bezels and gestures from the bezel into the display, Barrier Pointing relies on elevated bezels that stop a stylus motion from the inside of the display area towards the edges. The *SimPress* [3] technique may be suitable to click on multiple objects while in Bezel Swipe mode. Outside of Bezel Swipe mode, the rocking finger motion that triggers SimPress, even if it is small, does conflict with the navigation gestures.

Bezel Swipe is susceptible to finger occlusion, particularly when used to select text. In order to support the acquisition of small targets, Bezel Swipe must be combined with occlusion compensation techniques such as *Shift* [12].

EXPERIMENTAL EVALUATION

We evaluated Bezel Swipe against one single mode condition, Direct, and two mode-switching alternatives, Single and Multi. In *Direct*, users selected an image by tapping on it and navigated by swiping gestures. In a real application, a tap on the image would likely show a full screen ver-

Samsung Exhibit 1012, Page 2 of 4

sion of the image instead, which is why we cannot assume that this selection method is generally available. In *Single*, users tapped on a button to enter selection mode. Then they tapped on an image to select it. This also canceled selection mode. In *Multiple*, users tapped on the button to enter and to leave selection mode. While in selection mode, users could select as many images as they wished but they could not navigate. Instead, moving the finger across the display highlighted the image under the finger with a thick border and selection occurred on take-off. The selection highlighter also appeared in Direct as long as the finger touched an image, and vanished when finger movement indicated a swipe. For all methods, selected images were framed with a thin border. Users could deselect images by selecting them again.

We assume that Direct, arguably the simplest and most "ideal" method, is unavailable because of conflicts with other gestures. Consequently, our hypothesis was that Bezel Swipe would not perform as well as Direct (H1). However, we expected that Bezel Swipe would be easier to use than the two moded alternatives (H2). We realize that H1 is a somewhat unusual hypothesis from an evaluation standpoint but we are interested here in evaluating how close Bezel Swipe would be to some notion of an ideal.

We chose image thumbnail selection over text selection as our experimental task because we wanted to start with a simple, straightforward task to examine the differences between Bezel Swipe and other methods. Scanning images for insects is a simpler task than scanning text for a target word. A simple image classification (insect or face) can be accomplished more quickly than a word recognition. We also wanted selection targets for which finger occlusion has limited influence. For a text selection task, we would have to add a compensation technique to attenuate the effects of finger occlusion, further complicating the task. Text selection and occlusion compensation techniques using Bezel Swipe are interesting subjects for future study.

We kept parameters as consistent across conditions as possible to make the conditions comparable. The presentation of the buttons and the Bezel Swipe activation bar were identical. We used a green 8×160 pixels bar centered vertically against the left display edge, with an active area that was 12 pixels wide. When in selection mode, the bar turned yellow.

We presented a grid of 10×20 images of which 40 showed an insect and all other images showed faces. About six full images and six partial images could be seen on the display. We scaled all images proportionally to fit into a 128×128 pixels bitmap, which amounts to about 1.5×1.5 in^2 on the iPhone display. This puts the image dimensions into a safe target size range [7] and limits the influence of target acquisition problems [12] on the image selection performance.

Study Design

We used a repeated measurements within-subject design for our experiments. Participants first answered a questionnaire that asked for basic demographics, hand dominance, device usage and native writing system. Next, participants received written and oral instructions on how to perform all four selection methods accompanied by a brief trial of each on a separate trial image set. For the tasks, we instructed participants to select all insect images in the grid as swiftly as they could. After each task, participants filled out a questionnaire that captured subjective ratings of difficulty and satisfaction on a nine-point Likert scale. Participants had to agree (ratings 1 to 4) or disagree (ratings -1 to -4) with statements such as "The mechanism requires a lot of accuracy." A rating of 0 represented neutrality.

Methods and data sets were presented in counterbalanced order using Latin Squares. To avoid image learning effects between methods, each data set used different images arranged in one of four randomized grids.

Participants included eleven males and five females aged 19 to 50 years old (median 39). Seven were recruited from a small research company and nine from cafes in San Francisco. Seventy-five percent had used the iPhone previously. No differences in performance were found between those who had previously used the iPhone and those who had not.

Results

Participants achieved similar accuracy for each of the four techniques. On completion, they selected an average of 94.5% of the insect images ($F_{(3,45)} = 0.457$, $p > 0.7$) while selecting an average of only 0.43 face images in error ($F_{(3,45)} = 0.244$, $p > 0.9$).

Although task performance was comparable across methods, participants exhibited some difficulty during the task. They selected faces in error, then generally deselected them immediately. People made slightly more errors in Direct than in Single or Bezel Swipe, but these errors were far more likely when they used Multiple ($F_{(3,45)} = 19.518$, $p > 0.001$). We believe this is due to attempts to navigate when they were still in selection mode when using Multiple. In the Bezel Swipe and Single methods, this type of error is not an issue, since selection is disabled after selecting an image. This error also cannot occur in Direct, since images are not selected if the finger is in motion across the display area.

Participants also deselected insect images they had selected, and then had to reselect them. This occurred less often because of the relative infrequency of insects to faces. As in the case of incorrectly selected faces, more insects were accidently deselected when using Multiple, followed by Direct, with Single having the least ($F_{(3,45)} = 4.667$, $p > 0.01$).

Participants were quickest to complete the task when using Direct (101.34 sec.) and the slowest with Multiple (170.01 sec.), with Bezel Swipe (138.15 sec.) and Single (140.66 sec.) in between ($F_{(3,45)} = 34.292$, $p < 0.001$). This result cannot be accounted for by looking at the errors alone, since Direct resulted in more errors during the task than Single or Bezel Swipe. We believe the shorter time is related to the relative simplicity of Direct, which only requires a single tap on an image. Bezel Swipe and Single each require contact be made with the bezel edge or the bar, respectively,

Samsung Exhibit 1012, Page 3 of 4

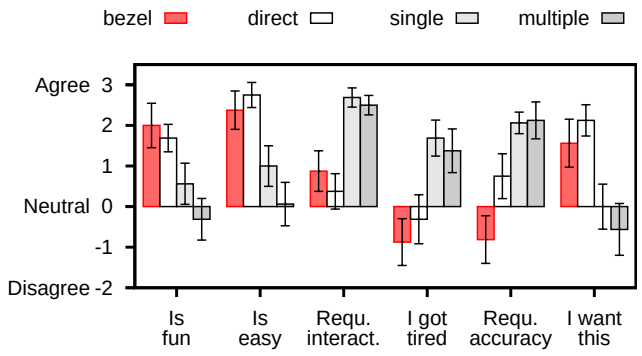

Figure 3. Responses to post-task questionnaire on a scale from -4 (Disagree) to +4 (Agree).

before moving to the image to be selected. Although users of Multiple had a simpler task in terms of actions, the moded interaction proved to be difficult to use.

Results from the questionnaire (see Fig. 3) show that people viewed Bezel Swipe very positively. When compared to Direct, it was rated the same or nearly the same for the following: "I had fun using the mechanism," "The mechanism is easy to use," "The mechanism requires a lot of interaction," "I was getting tired using the mechanism," and "I would like to have the mechanism on my mobile device." In addition, Bezel Swipe was viewed more positively for these statements than Single and Multiple, which did not differ from each other. The only statement rated more positively for Direct than Bezel Swipe was "Using the mechanism requires high accuracy." Ratings for "I am sure I marked everything correctly" and "I had difficulties seeing all necessary UI elements" were not significantly different for the four methods. This suggests that for all methods people tried to be equally accurate and that they were able to see the UI elements.

In summary, performance measures for four methods show similar accuracy levels for all conditions. As expected, the simpler method of Direct was faster than other methods. Participants took more time when using Bezel Swipe and Single. Multiple took significantly longer and had more errors than the other three methods. In subjective ratings, Bezel Swipe was rated as favorably as Direct on most measures, while the other methods were rated significantly worse.

CONCLUSIONS

Since its introduction, the iPhone has gained significant popularity and its user interface design has had an impact on the design of newer phones. While it supports navigation tasks very well, it has shortcomings with respect to tasks such as cutting, copying and pasting multiple object selections.

We proposed Bezel Swipe, a simple interaction technique that enables many of these operations without limiting the ease of navigation. Bezel Swipe takes advantage of the edge of a touch display, enabling users to easily access functionality by activating a thin button. The finger movement continues to the target, without requiring the user to lift it and reposition it over the target. In addition, because the finger is continuously in contact with the display, Bezel Swipe can provide feedback to the user, enabling them to fine tune their selections before committing to them. We evaluated Bezel Swipe's usefulness in an image selection task. Participants in our usability study liked Bezel Swipe and found it to be a reasonable alternative to direct touch selection.

In the future, we will explore the use of Bezel Swipe for text editing tasks, such as highlighting, copying or deleting selected text. We will also examine its scalability and limitations. Possible extensions may include its use on other types of devices, such as multiple or large display environments.

Acknowledgments

We thank the reviewers, Gene Golovchinsky, Scott Carter and Jake Biehl or their helpful comments. We also thank all participants of our study for their time.

REFERENCES

1. J. Accot and S. Zhai. More than dotting the i's — foundations for crossing-based interfaces. In *CHI 2002*, pages 73–80, 2002. ACM.
2. Apple Inc. iPhone. http://www.apple.com/iphone, Nov. 2008.
3. H. Benko, A. D. Wilson, and P. Baudisch. Precise selection techniques for multi-touch screens. In *CHI 2006*, pages 1263–1272, 2006. ACM.
4. W. Buxton. A three-state model of graphical input. In *INTERACT 1990*, pages 449–456, 1990. North-Holland Publishing Co.
5. J. Froehlich, J. O. Wobbrock, and S. K. Kane. Barrier pointing: using physical edges to assist target acquisition on mobile device touch screens. In *ASSETS 2007*, pages 19–26, 2007. ACM.
6. S. Lawson. Low-income users latch on to iPhone. http://www.itworld.com/print/57039, Oct. 2008.
7. P. Parhi, A. K. Karlson, and B. B. Bederson. Target size study for one-handed thumb use on small touchscreen devices. In *MobileHCI 2006*, pages 203–210, 2006. ACM.
8. Proximi. MagicPad – The first rich text editor for the iPhone. http://magicpad.proximi.com/, Nov. 2008.
9. J. Raskin. *The Humane Interface: New Directions for Designing Interactive Systems*. Addison-Wesley, 2000.
10. Rubicon Consulting Inc. The Apple iPhone: Successes and challenges for the mobile industry. White paper available at http://rubiconconsulting.com, Mar. 2008.
11. B. Shizuki, T. Hisamatsu, S. Takahashi, and J. Tanaka. Laser pointer interaction techniques using peripheral areas of screens. In *AVI 2006*, pages 95–98, 2006. ACM.
12. D. Vogel and P. Baudisch. Shift: a technique for operating pen-based interfaces using touch. In *CHI 2007*, pages 657–666, 2007. ACM.

Samsung Exhibit 1012, Page 4 of 4