Gestures and Widgets: Performance in Text Editing on Multi-Touch Capable Mobile Devices

Vittorio Fuccella **Poika Isokoski** **Benoît Martin**

ABSTRACT

We describe the design and evaluation of a *gestural* text editing technique for touchscreen devices. The gestures are drawn on top of the soft keyboard and interpreted as commands for moving the caret, performing selections, and controlling the clipboard. Our implementation is an Android service that can be used in any text editing task on Android-based devices. We conducted an experiment to compare the gestural editing technique against the widget-based technique available on a smartphone (Samsung Galaxy II with Android 2.3.5). The results show a performance benefit of 13-24% for the gestural technique depending on the font size. Subjective feedback from the participants was also positive. Because the two editing techniques use different input areas, they can co-exist on a device. This means that the gestural editing can be added on any soft keyboard without interfering with user experience for those users that choose not to use it.

Author Keywords

Gestures; Text editing; Caret movement; Clipboard; Android

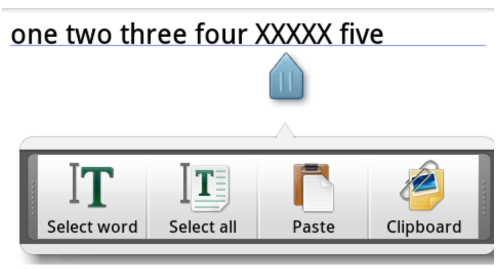


Figure 2. The editing widget and menu in Samsung Galaxy S II (Android 2.3).

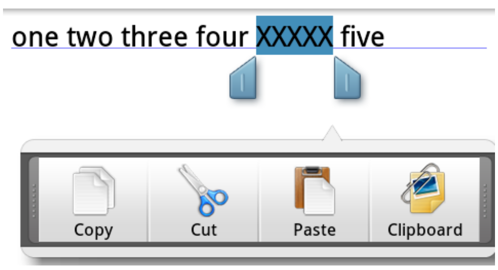


Figure 3. Text selection, the adjustment handles, and the associated menu on Samsung Galaxy S II (Android 2.3).

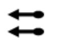
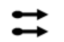
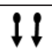
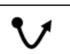
Name	Category	Actions	Shape
Left	CP/Sel	Moves the caret/selection endpoint one char left	
Right	CP/Sel	Moves the caret/selection endpoint one char right	
Up	CP/Sel	Moves the caret/selection endpoint one row up	
Down	CP/Sel	Moves the caret/selection endpoint one row down	
2Left	Sel	Moves selection endpoint one word left	
2Right	Sel	Moves selection endpoint one word right	
2Up	Sel	Moves selection endpoint to the beginning of the text	
2Down	Sel	Moves selection endpoint to the end of the text	
Copy	CC	Copies the selected text to the clipboard	
Cut	CC	Cuts the selected text and copies it to the clipboard	
Paste	CC	Pastes the text from the clipboard	

Table 1. The gesture-operation mapping used in our design. Each gesture was associated with action(s) from three categories: Caret Positioning (CP); Selection Manipulation (Sel); Clipboard Control (CC).

EXPERIMENT

As it is apparent in the related work section above, we were not alone in believing that gestures may make editing more pleasant and perhaps more efficient on touchscreen devices. However, we have not seen empirical evidence of this. This is why we ran an experiment to record users' subjective impressions and performance in editing tasks.

Design

The experiment was a two-factor within-subjects design. The factors were the font size (1.75, 3.25, and 4.75 mm) and the editing technique (gestural and the widget based technique that comes with Android 2.3 on Samsung Galaxy II phones). The font sizes were selected by picking a comfortable font size for the phone model in question and then a smaller and bigger at 1.5 mm intervals.

The primary dependent variable was the task completion time. In addition it was possible to compute a variety of other measures that may correlate with perceived ease of use such as the number of gestures, text selection events, key presses, etc. Some of these are reported in the results section.

To record subjective impressions we used the System Usability Scale (SUS) [6] that was filled in for both editing techniques. The questionnaire is composed of 10 statements to which participants assign a score indicating their strength of agreement in a 5-point scale. The final SUS score ranges from 0 to 100. Higher scores indicate better perceived usability. In addition we made notes of participants' comments during and after the experiment. We also collected background information including whether the participants owned and operated touchscreen devices and whether they edited text on their devices.

Participants

We recruited a total of 12 participants among the staff and students of our universities in Finland and France. Five of the participants were female and seven were male. The ages ranged from 25 to 52 (M=34.7, SD=7.8). All participants had some experience with touchscreen devices. Eight participants had mobile phones with touchscreens (four iPhones and four other devices). In addition one participant owned an iPad, but reported that the only editing task she used was entering URLs to a web browser. Interestingly, three participants did not yet own or operate a touchscreen device although all were mobile phone owners.

When asked about text editing habits with the touchscreen devices, the participants who used touchscreen devices typically mentioned using backspace to erase text and tapping on the text to move the caret if the error was far away from the current position. Use of copy, cut, and paste commands was rare and limited to copying URLs to the web browser and sometimes copying the whole message in order to construct the reply around the quoted material.

All participants reported their skill in the English language to be at least good (on the scale of poor, good, near native). Note that while the tasks were presented in the English language, only the last task required language skills beyond knowing the order of small numerals.

Apparatus

The experimental software had two parts. The first part was an Android text entry service that implemented the QWERTY soft keyboard with the ability to interpret the gestural commands listed in Table 1. This soft keyboard was constructed based on the text entry method example delivered with the Android Software Development Kit. An Android *toast* (a message shown for a short time in a small window) appeared superimposed on the keyboard after each clipboard operation. It was used to give a positive or negative feedback on the success of the operation.

The second part of the experimental software was a text editor whose initial screen presented a task list. The participants started a task by selecting it from the list. The text editor underlined the erroneous passages that needed editing and displayed a message when the editing task had been completed. Besides logging input events, the editor also automatically measured the duration of a task from the time when the task was shown to the time when it was completed. The software determined the completion status by comparing the edited text to the correct solution.

The phone model was chosen because of the Android operating system, sufficient processing and storage capabilities and the widespread availability[14]. The display had a 480x800 pixel resolution and measured approximately 110 mm in diagonal. The multi-touch sensing was capacitive.

The editing tasks used in our experiment are listed in Table 3. Earlier work did not always report the tasks used in the experiments. We would prefer to have a standard task set to use in editing experiments. Table 3 is included in this paper in full to support the development of such a set.

The choice of tasks should ideally be based on statistics on the frequency of editing actions that occur in real-life situations. Lacking such statistics, we based our design on the inclusion of the widest possible range of editing situations without making the experimental sessions too time consuming. We chose the editing operations on the basis of a previous research on the use of gestures in text editing [25]. There was also a significant overlap with the basic tasks by Roberts[20] as almost all operations were included in our tasks. Furthermore, character insertion and deletion operations were considered both individually and when included in a series. We also designed the tasks to let the user exploit all of the gestures provided by the gestural editing technique.

Although all tasks contained more than one type of interaction, it is useful to divide the task set into subsets according to the dominant type of interaction to aid in the analysis of the performance results. This made it possible to relate possible performance differences to the types of interaction. In our analysis of the tasks we ended up with three sets. The first set was dominated by keyboard use (tasks 2, 3, 7, and 8), the second set was dominated by caret movement (tasks 1, 4, 5, 6, 15), and the third set is dominated by text selection and clipboard use (tasks 9, 10, 11, 12, 13, and 14).

[14]Internet sources seem to agree that about 20 million phones were sold in ten first ten months of availability

Figure 5. A screen shot of the Samsung Galaxy II display during an editing task.

Procedure

Each participant participated in one session lasting about 1.5 hours. In the beginning of the session each participant was given a sheet of paper. One side of the paper contained instructions in the use of the editing techniques including a table of the available gestures. The instruction sheet remained visible during the experiment. The other side of the sheet contained an informed consent form that the participants signed before the experimental tasks began.

The experiment consisted of eight blocks. In each block, the participants had to complete all of the the previously described fifteen tasks. The participants were allowed to rest as long as they wanted between tasks. Longer breaks were taken between blocks while the experimenter was setting up the next block. Four of the blocks were completed using the widget-based technique and the remaining four blocks were completed using the gesture-based technique. Half of the participants did the four gesture blocks first and then the four blocks with the widgets, while the rest reversed the order of the techniques. The first block of each editing technique was completed using the medium (3.25mm) font size. It was aimed at training the use of the editing technique and at familiarizing the participant with the tasks. The three blocks after the initial training block varied in the font size. This way, we completely counterbalanced the order for the two factors, technique (widget, gesture) and font size (small, medium, large), obtaining 12 different permutations for 12 users, as summarized in Table 2.

Participant	Order of Font Sizes		Order of Techniques
	Training	Blocks	
1	m	s-m-l	gesture-widget
2	m	s-l-m	gesture-widget
3	m	l-s-m	gesture-widget
4	m	l-m-s	gesture-widget
5	m	m-s-l	gesture-widget
6	m	m-l-s	gesture-widget
7	m	s-m-l	widget-gesture
8	m	s-l-m	widget-gesture
9	m	l-s-m	widget-gesture
10	m	l-m-s	widget-gesture
11	m	m-s-l	widget-gesture
12	m	m-l-s	widget-gesture

Table 2. The counterbalancing scheme used in the experiment. The font size is reported abbreviated (s=small, m=medium, l=large).

After completing all blocks with one editing technique, the participants responded to the System Usability Scale concerning that technique. At the end of the experiment we asked which technique they would prefer if they had to make the choice. There was room for free-form feedback at the end of the SUS form and we kept notes on the comments the participants made during the experiment and during the final debriefing.

Task	Title	Instruction	Presented form	Correct form
1	Delete Character	Delete the X character in the sentence	one two thrXee four five	one two three four five
2	Delete Word	Delete the X characters in the sentence	one two three four XXXXX five	one two three four five
3	Delete Phrase	Delete the incorrect phrase in the sentence	one XXXXX XXX two three four five	one two three four five
4	Delete Characters	Delete the X characters in the sentence	oneX XXXXX two thrXee fouXXXr Xfive	one two three four five
5	Insert Character	Insert a space in the sentence	one two threefour five	one two three four five
6	Insert Characters	Insert spaces in the sentence	onetwothreefourfive	one two three four five
7	Insert Word	Insert the correct word in the sentence	one three four five	one two three four five
8	Insert Phrase	Insert the correct words in the sentence	one four five	one two three four five
9	Move Word	Move a word to restore the correct order	one three two four five	one two three four five
10	Move Word 2	Move a word to restore the correct order	one three two four five	one two three four five
11	Move Phrase	Move the words to restore the correct order	one four five two three	one two three four five
12	Move Line	Move a line to restore the correct order	one one one one three three three three two two two two four four four four five five five five	one one one one two two two two three three three three four four four four five five five five
13	Move Lines	Move the lines to restore the correct order	one one one one four four four four five five five five two two two two three three three three	one one one one two two two two three three three three four four four four five five five five
14	Complete Text	Fill in the missing text	one one one one two three three three three four five five five five	one one one one two two two two three three three three four four four four five five five five
15	Correct Errors	Correct the misspelled words	Twenty years form now you will be more disappointed by the thXings you didn't do than by the ones you dd. So throw off the bowlines, Sail away from the safe harborX. Catch the trade winds in oyur sails. Explore. Drem.	Twenty years form now you will be more disappointed by the things you didn't do than by the ones you did. So throw off the bowlines, Sail away from the safe harbor. Catch the trade winds in your sails. Explore. Dream.

Table 3. The editing tasks. Spotting the errors was easier in the experiment than it seems here because the spell checker underlined words that did not match the corrected form.

Vulture: A Mid-Air Word-Gesture Keyboard
Markussen et al., CHI 2014

"Word-gesture keyboards enable fast text entry by letting users draw the shape of a word on the input surface. Such keyboards have been used extensively for touch devices, but not in mid-air, even though their fluent gestural input seems well suited for this modality. We present Vulture, a word-gesture keyboard for mid-air operation. Vulture adapts touch based word- gesture algorithms to work in mid-air, projects users' movement onto the display, and uses pinch as a word delimiter. A first 10-session study suggests text-entry rates of 20.6 Words Per Minute (WPM) and finds hand-movement speed to be the primary predictor of WPM. A second study shows that with training on a few phrases, participants do 28.1 WPM, 59% of the text-entry rate of direct touch input. Participants' recall of trained gestures in mid-air was low, suggesting that visual feedback is important but also limits performance. Based on data from the studies, we discuss improvements to Vulture and some alternative designs for mid-air text entry."