Usability evaluation of single-stroke vs. multi-stroke gesture recognition for rapid text input on small screen mobile devices.

Saad Chaudhry
Department of Computer Science and Engineering
University of Oulu
mchaudhr@ee.oulu.fi

STATEMENT OF THE PROBLEM
Recent years have seen tremendous growth in handheld mobile devices, most of which interact with the user through a rather small touch screen interface. With their increased capabilities, mobile devices are increasingly being used for more text-intensive tasks like SMS, e-mail, Internet browsing and even for creating and editing documents.

Despite the obvious disparity between desktop and mobile systems, most mobile devices rely on traditional desktop-environment-tested interface paradigms [1].

Without physical keyboard, users have to turn to key-based or stylus-based text entry device for solutions. Stylus-based handwriting interaction seems more natural. However, problems of segmentation and recognition come out [3]. (Note that since 2007’s announcement of the iPhone, and widespread adaptability of capacitive touchscreens, stylus has gradually been replaced by finger-touch user interfaces.). Performances of recognizer are closely correlated to speed and accuracy and text entry has been widely studied in recent years [4].

This research shall focus on two types of handwriting-based text recognition algorithms; particularly studying the usability of single-stroke and multi-stroke gesture recognition algorithms for text input, rather than their efficiency. The target is to develop a relation between the accuracy of the two techniques for different types of applications.

Unlike soft keyboard single-stroke text entry can afford eye-free entry. Strokes are made anywhere on the digitizing surface. Successive strokes are spatially independent and can occur, for example, directly on top of preceding strikes [5].

On the other hand, in case of multi-stroke gestures a timeout between strokes is used to segment adjacent characters [4]. However, they tend to be more accurate than single-stroke algorithms.

RESEARCH QUESTIONS AND/OR HYPOTHESES
The aim of the study is to evaluate the usability of single-stroke gestures against multi-stroke gesture when used in a mobile environment for two different application-specific tasks. The two applications are categorized with respect to the reader that will consume the entered texts (machine-reader / human-reader). The evaluation shall be performed to benchmark their usability rather than accuracy and performance in either case, and shall then be weighed against the time required to enter the text.

In short, the study shall answer the question of possibility of using single-stroke (faster but more erroneous) recognition for entering human-targeted and machine-targeted texts.

METHODS AND PROCEDURES
The study shall include testing of different available algorithms for both single and multi-stroke gestures in a time constrained environment.

The two gesture types shall be scrutinized in context of two different types of applications i.e. SMS for human-reader and Online Search for machine-reader.

In the end, user acceptability and preference in using the two techniques will be taken into account in terms of time-tolerance, and accuracy of results.

Variables
The two recognition techniques (single-stroke and multi-stroke), and the two application scenarios (human-reader and machine-reader) will be cross tested using the same pool of sample texts and by same testing volunteers.

All the testing will be performed on identical devices to eliminate any discrimination that may arise because of differences in underlying hardware or platform. It must however be noted that all the testers will test all possible technique-application combinations; one at a time.

The sample texts shall be same for all tests, but will be picked at random from a large pool, to prevent any subconscious learning by volunteers.
The volunteers shall be chosen at random from different fields of life having different technological backgrounds. The number of volunteers from each category will be controlled to create a uniform population.

Lastly, the tests shall be performed with the volunteer while being stationary and while walking.

**Sampling**
As per the basic aim of the study, the samples shall be collected separately (under identical conditions) for the two situations i.e. single-stroke gesture and multi-stroke gestures.

The testing volunteers shall be grouped into categories according to their previous experience on working with similar technologies. Once categorized by their technological background, equal number of samples shall be taken from each group and shall be merged together in the end to mimic general population.

The result samples will be collected as success or failure of entered text in performing the required search, or to convey the required information to the SMS receiver.

The available recognition algorithms for each of the two types will be tested randomly.

**Instrumentation**
Specially developed applications shall be used for testing purposes. At least two applications will be designed (one for single-stroke recognition, and one for multi-stroke recognition) to be integrated on an Android smartphone.

Both the applications shall be able to display the recognized text on screen, or to forward it to default Internet browser as keywords for performing online search. The online search, as mentioned before, will test the machine-reader test case.

For testing the human-reader situation, the volunteers shall be grouped in pairs, who will then read each other’s entered texts. The accuracy in this case will be calculated as a function of the percentage of information successfully conveyed by the (crippled/misrecognized) text.

**Data collection**
Data will be collected in separate testing sessions, over a period of time, every time with a new group of test volunteers.

**Data analysis**
The collected data will be analyzed in terms of accuracy of reading for human-reader, and percentage of times the intended search results are produced by the search engine for machine-reader.

The data collected from the different available algorithms for each of the two gesture types will be analyzed separately as well as collectively as an average. The algorithms available for single-stroke will be analyzed for determining the best single-stroke algorithm, and same shall be done for multi-stroke algorithms. The separate algorithm data for each type will also be averaged together and compared with similar data compiled from the other gesture type for comparing the gesture recognition techniques as a whole.

**LIMITATIONS AND DELIMITATIONS**

The following points list some of the limitations that could possibly be faced during the period of research:

- Limited number of volunteers who do not have any previous experience of working with gestures. Because of the nature of this research, it is important to conduct testing with people who have never used gestures before.
- Unavailability of volunteers from all age groups, for replicating the general population.
- Unavailability of different mobile devices with different screen sizes.
- Because of the limitations in processing power of available mobile devices, all the strokes have to be sized down to low resolution before processing. This in turn, lowers the efficiency of recognition algorithms.

The scope of study, especially data collection and sampling is narrowed down by applying the following delimitations.

- It is assumed that testing the applications on only one of the leading mobile platform shall be enough.
- The touch screen sensitivity is not being taken into account. Moreover, the availability of screen resolution beyond a required minimum is disregarded since the algorithms will be fed scaled down images in case of higher resolution devices.
- No new algorithms will be developed for the sake of this study. Instead, the research shall focus on scrutinizing the already available recognition APIs.
- All the samples will be collected for only two applications – Online Search and SMS. It can be argued that even different machine-readers like Online search and local search may behave differently (for example, phonebook search can yield results by using only first three characters), the study will consider only the broader perspective.
- Any special features like auto-correct will be turned off. Since there is a possibility that such features may favor one of the recognition type more than the other, auto-corrections shall not be taken into account to reach an unbiased conclusion.

**SIGNIFICANCE**
This research shall be able to present a study of tradeoff between inaccuracies when using single-stroke gestures against the time-delays required by multi-stroke gestures for telling separate characters apart. A final conclusion can be drawn for selecting either of the two methods as a method of choice for gesture recognition for text entry.
REFERENCES


