Computers in Human Behavior 53 (2015) 71–81

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

Research Report

Touch screen gestures for web browsing tasks

Sabrina S. Billinghurst [*,1], Kim-Phuong L. Vu [1]

California State University Long, United States

ARTICLE INFO

Article history:
Received 9 June 2014
Revised 8 June 2015
Accepted 9 June 2015
Available online 6 July 2015

Keywords:
Touch screen
Gestures
Web browsing
Mobile computing

ABSTRACT

Touch screens have many advantages over traditional desktop devices. In addition to their direct interaction and portability, they allow for a novel interaction method, hand gestures. Most research on touch screen gestures focuses on the technical constraints of the input and comparison of the input method to other techniques. Fewer studies have defined gestures for a specific task, which is the purpose of the present study. We defined natural gestures for common web browsing tasks, and added to the growing research in this area by examining differences between postures, screen sizes, and users' educational background. It was found that most gestures produced by users are made with the dominant hand, the index finger, and a single motion. There was no effect of device tilt or posture since most gestures were performed with one hand. Five fingered gestures were used more often on the tablet than on the phone. There were little differences between the psychology and engineering participants. Design implications and limitations are discussed.

1. Touch screen gestures

Touch screens have been increasing used as the preferred input method for smartphones, tablets, and other mobile devices. This trend is not surprising given that previous research has shown that touch screens have advantages over traditional input devices. Touch screens have been described as the fastest pointing device and have easier hand-eye coordination than mice or keyboards (Shneiderman, 1991). Bragdon, Nelson, Li, and Hinckley (2011) found that, when gestures were used on touch screen devices, participants only had to look at the phone 3.5% of the time compared to 98.8% of the time when soft buttons were used. Gestures can also provide a sense of immersion (Srinivasan & Basdogan, 1997) because the user is directly manipulating objects on a screen with their fingers rather than being disconnected between the control and display. Touch screens are also commonly and effectively used for simple applications without prior experience and have been considered the most natural of all input devices (Holzinger, 2003). Touch screens have also been shown to be effective input devices for children (Rust, Malu, Anthony, & Findlater, 2014) and older adults (Gao & Sun, 2015). Furthermore, touch screens save workspace since they do not require extra accessories such as a keyboard and mouse (Holzinger, 2003; Shneiderman, 1991). These advantages have most likely contributed to its popularity over the recent years.

On the other hand, touch screens have some disadvantages. Users' hands may obscure certain parts of the screen when touching it (Shneiderman, 1991). They are also not ideal for pointing to an exact location on the screen (Holzinger, 2003). In addition, typing on a touch screen keyboard takes longer than a traditional keyboard, which may make large data entry cumbersome (Holzinger, 2003; Shneiderman, 1991). However, new interaction techniques may overcome some of the disadvantages of touch screens. One method of interacting with touch screens is by executing touch screen gestures. Gestures are a combination of movements or multiple touches by the fingers that command a response from the system. Touch screen gestures have been shown to have advantages over more traditional ways of interaction. As noted earlier, Bragdon et al. (2011) found that gestures do not have to be looked at when compared to tapping on-screen buttons. In addition, Findlater, Lee, and Wobbrock (2012) found that gestures were preferred over buttons and were perceived to require less visual attention. This direct interaction technique has engaged research interest in human–computer interaction. However, only a few studies have examined factors that contribute to the selection of gestures for information input (e.g., Rust et al., 2014; Wobbrock, Morris, & Wilson, 2009); therefore, the present study will examine gesture creation for one type of task, that of web-browsing.

1.1. User-defined gesture sets

Recent studies examining gesture sets created by users (e.g., Rust et al., 2014) used procedures similar to those established by

* Corresponding author.
E-mail addresses: sabrinabillinghurst@gmail.com (S.S. Billinghurst), kim.vu@csulb.edu (K.-P.L. Vu).
[1] Department of Psychology, CSULB, 1250 N Bellflower Blvd, Long Beach, CA 90840, United States.

http://dx.doi.org/10.1016/j.chb.2015.06.012
0747-5632/

Wobbrock et al. (2009). As such, we will describe Wobbrock et al.'s study in some detail in this section. Wobbrock et al. compared a gesture set created by expert designers to one created by 20 non-technical participants for general commands on a table-top touch screen. The commands were general computer commands such as zoom, move, rotate, and next. Participants were asked to create a gesture that commanded a change in the display with one and two hands while thinking aloud. Then, they were asked to indicate their agreement with two questions on a 7-point Likert scale: *The gesture I picked is a good match for its intended purpose* or *The gesture I picked is easy to perform*. Responses to these questions were intended to provide a measurement on gesture goodness and ease. Wobbrock et al. rated each of 27 commands on conceptual complexity before participants created gestures.

Wobbrock et al. (2009) found that gestures that were created most frequently by participants for the same purpose had significantly higher goodness ratings. However, conceptually complex gestures were rated as having lower goodness. For gesture articulation time, they found that gestures that took longer to perform were rated as easier to perform, possibly because they were less hasty. The authors believed that gestures with longer articulation times were smoother or less hasty. This hypothesis was evaluated by a number of touch events (down, move, and up) that were recorded. Gestures that took longer to perform had more touch events, and therefore were rated easier than gestures with fewer touch events. In addition, the author's complexity ratings correlated with planning time. It took participants longer to create gestures for conceptually complex commands such as undo, task switch, help, menu access, accept and reject. Also, planning time affected goodness ratings. As planning time increased, goodness ratings decreased. This result could indicate that gestures that come naturally are those for simple commands, and that complex commands may require gestures that need to be taught to users.

Wobbrock et al. (2009) also manually classified each of the 1080 gestures defined by participants into four dimensions: form (type of hand pose), nature (symbolic, physical, metaphorical or abstract), binding (e.g., object centric), and flow (discrete or continuous). When Wobbrock et al. compared conceptual complexity and gesture categories, they found that simple commands resulted in physical gestures and more complex commands resulted in metaphorical or symbolic gestures. Therefore, as more research is completed on taxonomy of touch screen gestures, we can begin to see trends on how people define gestures. Wobbrock et al. also compared the gesture set created by the authors, or experts, to the gesture set defined by the "novice" participants, who did not have a technical background. They found that experts predicted only 60% of user-defined gestures and 20% of the expert's gesture set was not used by one novice participant. This finding indicates that use of subject matter experts alone cannot lead to fully predictable scope of gestures defined by users. It has also been suggested that experts consider technological constraints of gesture recognition when creating gestures for different commands and functions (Olafsdottir & Appert, 2014; Urakami, 2012). Therefore, in order to create a comprehensive library of gestures, end users should be involved.

Similar to Wobbrock et al. (2009), previous research by Karam (2005) indicate that gestures for touch screens has focused mainly on what gestures can be interpreted by the system, and not for a specific task. Thus, many gestures being implemented based on technological limitations are "stressing, impossible for some people to perform, and [use] illogical imposed functionality" (Nielsen, Störring, Moeslund, & Granum, 2004). Instead, the goal of gestural interfaces should be to consider an efficient human–computer interaction for specific applications in addition to the abilities of the system. Thus, the present study will examine gesture sets conducted for particular applications. This approach can be effective as demonstrated next by studies on gestures for punctuation symbols (Findlater et al., 2012) and navigational tasks (Urakami, 2012).

Findlater et al. (2012) performed two studies to create and evaluate user-defined single touch and multi touch gestures for punctuation symbols. Their goal was to find punctuation gestures that can be used instead of, or in addition to, buttons on a keyboard. In the first study, Findlater et al. asked 20 participants from both technical and non-technical backgrounds to create gestures for punctuation symbols on an Apple iPad. All participants had experience using a touch screen device. Twenty-six symbols and commands were randomly presented, and participants were asked to complete three trials. The first trial asked participants to create a gesture while displaying the name, and the second trial asked participants to create a gesture while displaying the symbol itself (commands shows the name in the second trial). Then participants were asked to rate whether the gesture was, "a good match for its intended purpose" on a 7-point Likert scale. Lastly, participants were asked to change how they drew the gesture by using multi-touch instead of single touch or vice versa. Authors independently grouped similar gestures by the shape of the stroke, and an agreement score was calculated from 0 to 1.

Findlater et al. (2012) found that half of the symbols and commands had agreement scores of .90 or higher. Moreover, commands tended to have a lower agreement scores than the symbols. In response to the rating of good match, participants rated symbols significantly higher compared with commands, which reflects the results from agreement scores. In a second experiment, Findlater et al. showed that gestures can be used as efficiently as buttons with little practice. In the second experiment, the touch screen, or moded-keyboard had a special button that reveal a keyboard that had punctuation symbols. To enter a punctuation symbol, the participant tapped on the punctuation button, then press the punctuation symbol. After practice trials writing every symbol with each technique, participants were asked to type pairs of words including punctuation symbols. Results showed that participants typed, on average, 19.7 words per minute (WPM; $SD = 5.16$) with gestures and 19.9 WPM ($SD = 4.4$) with the moded-keyboard. They also found that the typing times were not significantly different between the gestures and buttons, though the gesture set was responded to more positively by participants when compared to the button presses. Moreover, participants commented that the gesture set was more efficient and required less visual attention when compared with the touch screen buttons. Thus, this study demonstrated that gestures can be used as efficiently as buttons, and users were excited to use gestures.

In another study investigating gestures for a single task, Urakami (2012) compared technical and non-technical users in their gesture choice for a map application. Urakami video-taped 20 participants creating gestures for map navigation tasks. Then, each gesture was coded into three main categories: form, nature, and binding based on a combination of taxonomies developed by Freeman, Benko, Morris, and Wigdor (2009) and Wobbrock et al. (2009). The form category described physical form of the hand(s) when making a gesture. The nature of the gesture described the meaning of the gesture; whether it was symbolic, physical, metaphorical, or abstract. Lastly, the binding category described whether the gesture related or did not relate to an object on the screen. Urakami utilized these categories to define and compare gestures that were created by the novice and expert participants.

In her first experiment, Urakami (2012) compared the type of gesture created by novices and experts based on this taxonomy. She found that experts used gestures that were more abstract, while novices used metaphorical gestures. Thus, technical expertise seemed to affect the type of gesture that was created. In a second experiment, Urakami tested the memorability of the gesture