Understanding Touch and Motion Gestures for Blind People on Mobile Devices

Marco Romano, Andrea Bellucci, Ignacio Aedo

▶ **To cite this version:**

HAL Id: hal-01599659
https://inria.hal.science/hal-01599659

Submitted on 2 Oct 2017

Understanding touch and motion gestures for blind people on mobile devices

Marco Romano, Andrea Bellucci and Ignacio Aedo

Universidad Carlos III de Madrid, Leganés, Madrid, Spain

{ mromano, abellucc }@inf.uc3m.es, aedo@ia.uc3m.es

ABSTRACT - Considering users preferences and behaviour is a necessity to develop accessible interaction for blind people. Mainstream mobile devices are widely used by people with disabilities but, despite the growing interest of the research community around accessibility issues of touch interfaces, there is still much to understand about how best to design the interaction of blind people with mobile technologies. To this end, we conducted a preliminary elicitation study (8 participants) to understand how blind people perform touch and motion gestures for common tasks on a mobile phone. We found that blind people do not use motion gestures. We provide a discussion of our results according to the type of gestures performed.

Keywords. User-defined, gestures, blind, accessibility, touch screens.

1   INTRODUCTION

Mainstream mobile devices, like smartphones, are widely used by blind people [7]: they offer opportunities to increase their level of independence by making it possible to keep in touch with family and caregivers [18], orientate themselves in the street [3] and have portable access to assistive technologies like screen-readers [11]. Apart from screen-reading software that are now included in major mobile platforms, the interaction with mobile devices presents important accessibility barriers, since touch-enabled displays are designed for sighted interactions —users need to locate objects on the screen— and do not provide tactile feedback, conversely from hardware buttons.

Likewise previous research [6][8], we are interested in understanding how to improve the accessibility of modern mobile interaction. Off-the-shelf mobile phones allow to interact both by touch and motion gestures to perform specific computing tasks: our goal is to explore the design space and provide insights to understand how best to design gestural interaction with mobile devices for a blind person.

Here we present results of a preliminary elicitation study [19] in which 8 blind users are asked to provide touch or motion gestures for common tasks on a mobile device. Our insights provide "prima facie" implications for the design of accessible mobile gestures sets. Compared with the study of Kane et al. [8], in which researchers used a 10.1-inches tablet pc to elicit touch gestures from 10 blind people, our results

suggest that users have different gestures preferences for a mobile form factor. Moreover, we found that blind people are not particularly keen to motion gestures.

2 BACKGROUND

Our research is informed by previous work on touch and motion gestures, accessible mobile touch interface for blinds and user-defined gestures for surface computing.

Touch and motion gestures.
Touch gestures (or surface gestures) are hand movements performed on a two-dimensional surface, like the touch-screen of a mobile phone [13]. Over the last two decades, surface gestures have been studied under different perspectives, from enabling hardware [15], to software frameworks and architectures [4], to design principles and taxonomies for touch interaction and evaluations of gestures vocabularies [13]. Modern mobile devices allow users to perform also *motion gestures* that are gestures performed in three dimensions by translating and rotating the device thanks to embedded sensors [17]. One of the first examples of a motion gesture was provided by Rekimoto [16], who exploited tilting data to interact with virtual objects. Recently, motion gestures have been explored for a variety of input tasks, such as text entry [5] and maps navigation [2]. *One of the limitations of gestural interaction is the fact that, without any user participation, the gestures set is defined by designers. Although such gestures can be legitimate to test technical aspects of gesture systems —since they are easy to recognize— they do not consider users preferences and performances.*

User-defined gestures for surface computing.
Previous research exhibits that users preference patterns match better with gestures created by groups of potential users rather than expert designers [10]. To this end, researchers have begun to focus on participatory design techniques [19] to elicit touch gestures from users, aiming at generating better interactions that are directly informed by users behaviour. In their influential study, Wobbrock et al. [20] created a *user-defined* touch gestures set for surfaces by showing the outcome of an action to participants and asking them to provide a gesture that would produce that action. In the recent past a plethora of other studies have exploited the same method to elicit user preferences for a variety of conditions, e.g.: motion gestures for mobile devices [15], touch gestures for coupling mobile devices with interactive surfaces and multi display environments [9] or for discerning cultural differences and similarities in touch behaviours [12]. Significant to our research is the work of Kane et al. [8]: they applied the method of Wobbrock et al. with the goal of devising what kind of gestures are the most intuitive and easy to perform for blind people. In their study they asked participants to provide touch gestures for standard tasks on a touch-enabled 10.1–inches tablet. Insights from the study provide implications for the design of accessible touch interfaces for blind user, such as avoiding symbols used in print writing, use physical edges and corners as landmarks and reduce the need of spatial accuracy.

Our study follows the method proposed by Wobbrock et al. [20] and differs from the study of Kane et al. [8] in the following aspects: (1) we explore gesture preferences for a different form factor —a 3.5-inches smartphone— since the size of a device can be an important factor for the definition of a gestures set, and (2) we include motion gestures, which are a relevant input modality supported by modern smartphones that, nonetheless, is underutilized [16]. Moreover, there is a lack of knowledge if or not motion gestures can be a valid input method for blind users.

Mobile touch interfaces for blind people.
The widespread diffusion of mobile devices requires to assess accessibility issues of touch interfaces in order to provide efficient interaction techniques for *visual impaired users* [1]. Speech, audio and haptic has been widely used as output channels. For instance, screen-readers have been developed for different platforms, such as VoiceOver[1] for Apple OS or Eyes-Free Shell[2] for Android. The Talking Tactile Tablet [10] uses a special hardware with an overlay that enables tactile feedback. With respect to the input, various touch techniques have been explored for eyes-free interactions. Slide Rule [6], for example, offers adapted gestural interaction for navigating lists of items. Lastly, the role of touch location has been studied in order to provide different triggers for gestures that depend on the screen region [17].

Research on touch interaction techniques for blind people shows that we have not yet developed a clear understanding of users preferences, which can be achieved by involving the end users in the design process [19]. The results are promising, since it has been shown that blind people have significant spatial and tactile abilities and therefore are capable of using touch- and motion-based interfaces [8].

3   USER STUDY

We are interested in understanding what kind of gestures blind users consider intuitive for mobile devices and whether or not motion gestures can be a valid input modality together with touch gestures. We conducted a preliminary study with a group of 8 blind people in which participants were asked to perform touch or motion gestures for common tasks on a mobile phone.

We recruited 8 users, 4 male and 4 female with average age 61.1 (SD=11.29). Participants were selected considering their level of blindness and their familiarity with touch technologies. Participants have congenital blindness (2 of 8) or are early blind (6 of 8), a condition that requires the use of a screen reader to access digital information. As one of the participants stated *"blind people are not an homogeneous population"* and there are differences, for instance, with respect to spatial perception between early and late blind [14]. Since it would not have been possible, at this stage, to isolate this condition, we recruited the subjects in order to have a sample as uniform as possible. The participants belong to a Spanish organization for blind people

---

[1] https://www.apple.com/es/accessibility/ios/voiceover/
[2] https://code.google.com/p/eyes-free/

(ONCE). The organization provides general IT education and technological support to their members. In particular, it offers a course that focuses on interaction with iPhone, in which attendees learn how to interact with the VoiceOver interface. All of our participants attended the course and are familiar with IT technologies. They use a desktop pc daily for working purposes and use touch-screen devices mainly as personal assistants both when they are at home (e.g. reading emails) or outside (e.g. getting the current location and routes). Moreover, all participants consider themselves as technologically proactive and eager to improve the access of their peers to mobile technologies. We conducted the study in a laboratory setting: we chose the ONCE central office to make participants confortable, since the location was well known by all.

The study was conducted on a 3.5 inches iPhone 4s running iOS 7.1.2. We developed an HTML5 and Javascript application that captures multi-touch and motion gestures. As in [8], we used an operating system-agnostic application aiming to provide a neutral setup for the experiment. The ubiquity of HTML allows to repeat the experiment with heterogeneous hardware without implementing an ad-hoc version of the application. The application stores gestures data in a XML-formatted file and provides the researchers with a visualization tool for the later analysis. Figure 1 (a) shows a participant performing a touch gesture (on the left) and the resulting data captured by the application (on the right). All the sessions were video-recorded with a stationary camera pointing to the device to capture the interactions and protect subjects' privacy. No other objects were present and the users sat to side of a table during the whole session. Just a ONCE helper was present in case of necessity.
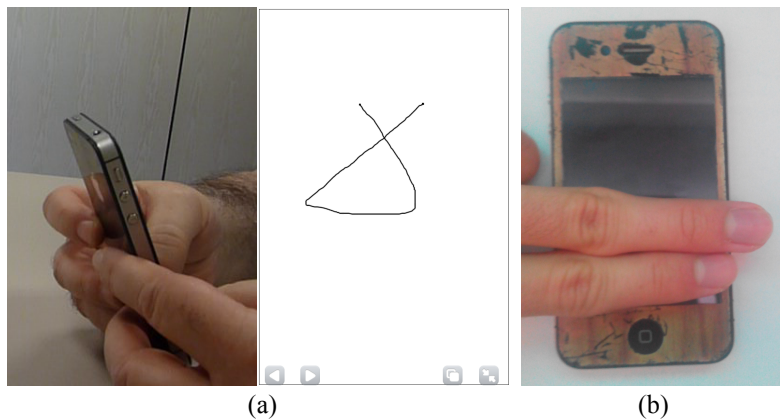

Fig. 1. (a) Example of a touch gesture and its visualization. (b) A no traditional gesture.

Previous to the experiment, we interviewed one of the ONCE IT specialists in order to fine-tune the setup. From the interview, we decided to give participants 10 minutes to get familiar with the device, since the IT specialist reported that "*it is paramount for visually impaired users to understand the dimensions of the screen in order to create a mental map of the device they are going to use*". We also estimated each session to last less than 40 minutes, which is the time limit the specialist suggested to avoid excessively stressing the users. Each session began with an interview on the kind of

disability, technological skills, and the participant experience with mobile technologies. Participants were then introduced to the tasks. They had to invent two gestures (a preferred one and a second choice), motion or touch, which can be used to trigger a specific command. We used the command list from Kane et al. [8], which we reviewed and modified in collaboration with ONCE IT specialist in order to match the context of mobile devices and to assure that all the participants could know and understand it. Selected commands were: *Context menu, Help, Undo, Switch application, Next, Previous, Ok, Reject, Move object, Open, Close, Copy, Cut, Paste, Quit, Select, Change input field, Answer up, Hang up*. Participants were asked to think aloud while performing the gestures and motivate their preferences. Before performing a gesture for a command, the experimenter read the name of the command and a description of the expected outcome.

Results.
The 8 participants were asked to perform 2 gestures for 19 commands for a total of 304 samples. Some of them were not able to provide two different gestures for the same command, therefore we have 152 gestures as first choice plus 126 for the second choice (total of 278). Only 12 of the 278 (4,32%) were motion gestures while multi-finger were 43 (15,46%). Figure 2 illustrates the gesture rationale for each category. We classified motion and touch gestures using the nature dimension of the taxonomy from Wobbrock et al. [20], in the same way Kane et al. [8] did for touch gestures and Ruiz et al. [17] for motion gestures.
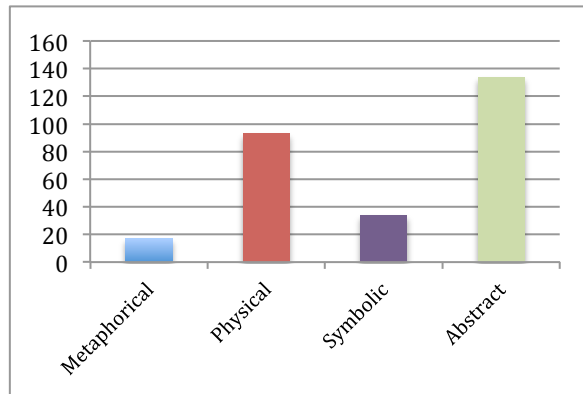

Fig. 2. Gestures rationale according to the nature dimension.

The nature dimension groups together *symbolic, metaphorical, physical and abstract gestures*. Gestures are defined as metaphorical when they represent the metaphor to act on an object different from a mobile phone (e.g., emulating an eraser on the touchscreen to remove an object). They are physical when they act directly on an object (e.g., drag and drop), symbolic when gestures indicate a symbol (e.g., typographic symbols like '?'). Finally gestures are abstract when their mapping is arbitrary (e.g., double tap). We grouped our samples taking into account users explanations and di-

rect observations. The *abstract* gestures are the largest category, with 134 samples (48.02%), followed by *physical* gestures (93 samples), *symbolic* (34) and *metaphorical* (17). The *abstract* category presents the highest prevalence of multi-finger gestures (24 gestures, 15 physical and 4 metaphor). Multi-finger gestures have been employed 18 times as first choice, 25 as second. Finally, 37% of the provided gestures stem from previous experiences with desktop and touch interfaces as explicitly commented by the testers.

4   DISCUSSION

When compared with the study of Kane et al [8], in which metaphorical and abstract gestures were dominant for a tablet form factor, our results show that physical and abstract are prevailing for a mobile device. Those results suggest that the physicality of the device, which determine the way users interact with it (e.g., making gestures with the same hand that is holding the device), the tasks selected for the mobile context (e.g., answering the phone) and the user background (e.g., technological knowledge) influence the kind of gestures made by blind users. In particular, the physical nature of the gestures is prevalent for the following commands: *previous* (60%), *next* (47%), *move object* (67%), *cut* (36%), *reject* (47 %). Some users commented to be inspired by the interaction mechanisms of real world that make them to feel the interaction as more "natural". In general abstract is the main gesture category and get the best percentages in commands such as *ok* (75%), *copy* (64%) and *select* (77%). Many of the abstract gestures stem from participants experience with the other user interfaces. For example, regarding the *ok* command, 7 users performed a tap or a double tap because it is what they normally do in their mobile or desktop devices.

Moreover most of the elicited gestures (203) are one-finger and multi-finger variations of tap or flick (an unidirectional movement of the finger). This is probably due to the fact that the VoiceOver interface, which participants have learnt to operate in the IT courses, uses mostly tap and flick gestures. Flicks vary on the number of fingers, from one to four. Taps vary on the number of fingers, one or two, and repetitions. In VoiceOver multi-finger gestures are mainly used for system control or are personalized gestures. This suggests the reason why testers demonstrated to be familiar with multi-finger gestures but to prefer mainly one-finger strokes.

Our study shows that blind people tent to use gestures that are different from common touch interfaces for sighted users but that are familiar to them. For instance, the 2-finger double-tap gesture is used frequently in the VoiceOver interface: its uses include answering and ending phone calls, starting and stopping music and video playback, etc. A preliminary advice for the designers of mobile accessible interaction is to consider the technological literacy of the users and exploit their familiarity with touch gestures.

Another outcome of the study is related to the reduced use of motion gestures. Causes are due to users lack of familiarity with motion gestures and because they were afraid to hit some objects moving the device. Conversely from other studies that focus on sighted users [17] and show that motion gestures are a significant for mobile

devices, we found that motion gestures do not seem to be suitable for blind people: there is no general agreement in using a common gesture for a given command.

Together with touch gestures that can be easily recognized, participants provided other kinds of no traditional hand postures, consisting in covering portions of the screen, that are not detected by current mobile phones hardware and software. Such gestures were performed by three different users for four different commands. For instance, covering the bottom part of the screen was used to confirm a choice, as it recalls the paper's position in which we usually sign an agreement, as depicted in Figure 1(b). Another participant used the palm to cover the whole screen with the intent to trigger the quit command.

A considerable number of gestures was performed in specific screen locations that participants associated to a semantic interpretation. Performing gestures in the left side was interpreted as "go back" or "recover from an action". For instance gestures on the left side were associated to commands like undo, previous, reject, close, quit and hang up. On the other hand, the right side was associated to opposite actions, such as next, accept, open, answer a call.

With respect to spatial awareness, Kane et al. [8] suggest to favor edges and corners of the screen because they can be used as physical landmarks that help blind people to locate spots in the surface. Our study reveals that this behavior largely depends on the experience of the users with mobile technology. We observed that users with less experience start by searching for the edges of the touch-screen and, when they are ready, they perform the gesture nearby it. Users with more years of experience with touch-screens generally start their gestures from the center of the device.

5   CONCLUSIONS AND FUTURE WORK

We presented a preliminary elicitation study to understand preferences of blind people with respect to touch and motion gestures on mobile devices. First insights from our study can be summarized as follows: 1) our results highlight that gestures are influenced by the form factor and the users background, 2) the gesture rationale is prevalently abstract and physical, 3) users tend to assign a behavior to screen locations, 4) users prefer touch gestures rather than motion gestures. Our work provides new insights on how blind people use mobile technologies. However, given the open-ended nature of the study, further research is needed. Since the lack of prior knowledge and experience affects the users' preferences, we consider that an in-depth study could reveal new interesting insights. We plan to modify the experiment requiring users to provide and then evaluate both a touch and a motion gesture for each command. This is to confirm whether motion gestures can be an effective input modality. The final aim is to reach an agreement on the mapping between gestures and commands and thus to define a robust and accessible gestures set for mobile phones.

6   REFERENCES

1. Abascal, J. and Civit, A. Mobile communication for people with disabilities and older people: New opportunities for autonomous life. In Proc. ERCIM Workshop 2000, 255-268 (2000)
2. Bellucci, A., Malizia, A., Diaz, P., and Aedo, I. Don't touch me: multi-user annotations on a map in large display environments. In Proc. AVI 2010, ACM, 391–392 (2010)
3. Di Chiara, G.; Paolino, L.; Romano, M.; Sebillo, M.; Tortora, G.; Vitiello, G.; Ginige, A., "The Framy user interface for visually-impaired users," *Digital Information Management (ICDIM), 2011 Sixth International Conference on* , vol., no., pp.36,41, 26-28 (2011)
4. Echtler, F. and Klinker, G. A multitouch software architecture. In Proc. NordiCHI 2008, ACM Press, 463 (2008)
5. Jones, E., Alexander, J., and Andreou, A. GesText: accelerometer-based gestural text-entry systems. In Proc. CHI 2010, 2173–2182 (2010)
6. Kane, S., Bigham, J., and Wobbrock, J. Slide rule: making mobile touch screens accessible to blind people using multi-touch interaction techniques. In *Proc. SIGACCESS 2008*, ACM Press, 73–80 (2008)
7. Kane, S.K., Jayant, C., Wobbrock, J.O., and Ladner, R.E. Freedom to Roam : A Study of Mobile Device Adoption and Accessibility for People with Visual and Motor Disabilities. In Proc. SIGACCESS 2009, ACM Press, 115–122 (2009)
8. Kane, S., Wobbrock, J., and Ladner, R. Usable gestures for blind people: understanding preference and performance. In Proc. CHI 2011, ACM Press, 413–422 (2011)
9. Kristensson, P.O., Clawson, J., Dunlop, M., et al. Designing and evaluating text entry methods. In *CHI EA 2012*, ACM Press, 2747 (2012)
10. Landau, S. and Wells, L. Merging Tactile Sensory Input and Audio Data by Means of The Talking Tactile Tablet. In Proc. EuroHaptics 2003, 60, 291–297 (2003)
11. Manduchi, R. and Coughlan, J. Portable and Mobile Systems in Assistive Technology. In Computers helping people with special needs, Springer Berlin Heidelberg, 1078–1080 (2008)
12. Mauney, D., Howarth, J., Wirtanen, A., and Capra, M. Cultural Similarities and Differences in User-Defined Gestures for Touchscreen User Interfaces. In CHI EA 2011, ACM Press, 4015–4020 (2010)
13. Morris, M., Wobbrock, J., and Wilson, A. Understanding users' preferences for surface gestures. In Proc. GI 2010, Canadian Information Processing Society, 261–268 (2010)
14. Postma, A., Zuidhoek, S., Noordzij, M. L., and Kappers, A. M. Differences between early-blind, late-blind, and blindfolded-sighted people in haptic spatial-configuration learning and resulting memory traces. Perception, 36(8), 1253–1265 (2007)
15. Rekimoto, J. Smartskin: an infrastructure for freehand manipulation on interactive surfaces. In Proc. CHI 2002, ACM Press, 113–120. (2002)
16. Rekimoto, J. Tilting Operations for Small Screen Interfaces. In Proc. UIST 2016, ACM Press, 2–3 (1996)
17. Ruiz, J., Li, Y., and Lank, E. User-defined motion gestures for mobile interaction. In Proc. CHI 2011, ACM Press, 197–206 (2011)
18. Sánchez, J. and Aguayo, F. Mobile messenger for the blind. In Universal Access in Ambient Intelligence Environments, Springer Berlin Heidelberg (2007), 369– 385.
19. Schuler, D., & Namioka, A., Participatory design: Principles and practices. L. Erlbaum Assoc. Inc., Hillsdale, NJ, USA. (1993)
20. Wobbrock, J., Morris, M., and Wilson, A. User-defined gestures for surface computing. In Proc. CHI 2009, ACM Press, 1083–1092 (2009)