An Empirical Study of Foot Gestures for Hands-Occupied Mobile Interaction

Mingming Fan, Yizheng Ding, Fang Shen, Yuhui You, Zhi Yu
Department of Computer Science
University of Toronto
{mfan, yizheng, fionashen, huihuihui, zy1122}@cs.toronto.edu

ABSTRACT
User-defined foot-gesture is a promising approach to interacting with mobile devices when both hands are occupied. In this research, we first present a survey to identify how often interacting with mobile devices is needed when both hands are busy and how many tasks people commonly would like to do on the devices in such situations. We then present a study to compare a traditional approach with foot-gesture interaction in a simulated hands-occupied scenario. Results show that foot-gesture saved over 70% of the time compared with the traditional approach and was perceived more useful and satisfying.

Author Keywords
Foot gestures; hands-busy, user-defined gestures.

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION
Touch interaction on mobile devices can be inconvenient or unavailable when hands are wet, dirty or occupied. It can also be unfavorable when the weather is too cold to remove gloves or when the intended action is too trivial compared to the overhead of fetching and unlocking the device. Foot gesture is an alternative approach to interacting with mobile devices in such hands-unavailable scenarios. It has been used in games [5], office work [6], walk-up installations [7] and even dance [3] to enhance user experience. Previous research has grouped foot gestures [9] into four categories and provided comprehensive understandings of the physical capability of performing foot gestures, such as kicking [4], tapping [2], raising and shaking a toe [10], tilting and rotating [8]. Previous research has also explored users' preferences in mappings foot gestures to mobile interaction commands [1].

ISWC '17, September 11–15, 2017, Maui, HI, USA
© 2017 Association for Computing Machinery.
ACM ISBN 978-1-4503-5188-1/17/09…$15.00
https://doi.org/10.1145/3123021.3123043

Despite of the research in existing related work, there is little data to show how often people may encounter hands-occupied scenarios and how many interactions they may want to accomplish on the device in such scenarios. Furthermore, there is no empirical evaluation of the performance of using foot gestures in hands-occupied scenarios. We aim to explore these questions in this work.

We first conducted an online survey to understand:1) *how often people wanted to use their mobile phone when both hands were occupied* (rate on the scale of 1-5: 1: never; 5: very often); 2) *the number of tasks on the phone that they would like to use when both hands are occupied*. Among the responses from 52 participants, 49 reported that they had experienced both hands occupied situations when they wanted to use the phone; 21 participants had experienced such situations frequently (ratings of 4 and 5). The average number of tasks that they wanted to do when both hands were occupied is 2 (*SD:1*). The top three tasks mentioned are: answer a phone call (63.5%), pause a music (42.3%) and resume a music (42.3%). Since there are maximum five tasks that people often want to do when both hands are occupied, we design a foot-gesture recognition app, *FootSketch*, that allows a user to define 5 foot-gestures (see Figure 1) and use them to trigger commands on the phone. We conducted a study to compare it with a traditional mobile interaction (TMI) in a simulated hands-occupied scenario.

EXPERIMENT
FootSketch: We used a mobile phone (Samsung Note 3) inserted into a running band as a prototype and tied it to the side of an ankle to sense and recognize foot gestures (see the

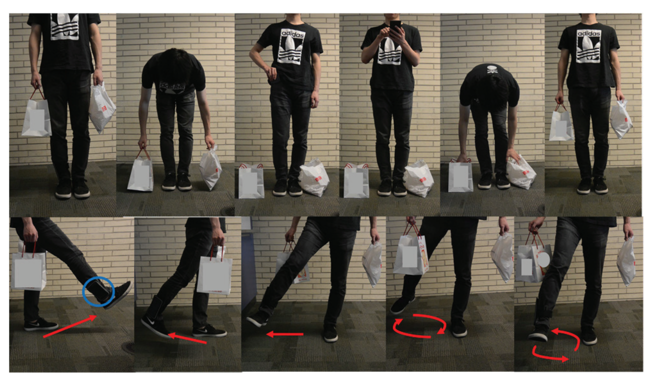

Figure 1 (Top) traditional mobile interaction (TMI): free hands, take phone out of a pocket, interact and resume two hands; (Bottom) FootSketch: moving forward, moving backward, moving to the side, moving clockwise; moving counter-clockwise.

blue circle in Figure 1). FootSketch constantly samples and quantizes accelerometer data and uses a dynamic time warping (DTW) algorithm to match a performed gesture with pre-recorded personalized gesture templates (source code: https://github.com/mingminf/InAirGestureRecognizer). To use footSketch, a user first records templates for each foot gesture and associates it with a command that she wants to trigger (*e.g.,* answering a phone call). This process only needs to be done once, and all templates and function mappings are stored internally with the app. During the real-time use, the user first activates the FootSketch by stomping the floor twice, and then sketches gestures with her foot. FootSketch finds the most matched template to the sketched gesture and triggers the command accordingly. A threshold is set to detect the stomping gesture and two consecutive stomping gestures are used to reduce false alarms.

Procedure: We recruited 12 volunteered participants between 20-30 (3 females). They are all right-footed and have healthy legs. We simulated a hands-occupied scenario (*e.g.,* after shopping) by asking them to hold a bag (~3 lb) in each hand (see Figure 1) and used within-subjects design to test traditional mobile interaction (TMI) and FootSketch. We asked each participant to draw 5 pre-defined gestures in a random order using TMI and FootSketch respectively. When using TMI, the participant must put down bags first, take the phone out of the pants' pocket, sketch the required gesture on the phone, put the phone back in the pocket, and pick up the bags. When using FootSketch, the participant first stomps the floor twice to activate the system, which vibrates for half a second as feedback. She then can start to draw the gesture using her foot. Upon completion, she pauses for two seconds and the system provides half a second vibration indicating that it is deactivated. Participants had a practice session (<3mins) before the study. They received real-time feedback of recognition results as they would do in realistic scenarios. We video recorded all sessions and identified the start and stop moments for both TMI and FootSketch in videos. For TMI, the recording starts when the participant began to bend her body to put down bags and stops when she resumed her normal posture after picking up the bags. For FootSketch, the recording starts when she stomped her foot and stops when she finished sketching the foot gesture. We asked them to rate the usability of two approaches on a scale of 1 (extremely negative) to 5 (extremely positive) in the end.

Results: Quantitative measures are shown in Table 1. Repeated measure ANOVA on the completion time shows

Table 1 Measures of two approaches for the hands-occupied scenario (bold items highlight significant differences)

	completion time (second)	ease of learning	Ease of rememb ering	usefulness	satisfaction
TMI	**22.5 (SD:4.2)**	4.3 (SD:0.7)	4.3 (SD:0.8)	**2.4 (SD:2.1)**	**2.4 (SD:1.2)**
FootSketch	**6.6 (SD:1.3)**	4.2 (SD:1.0)	4.4 (SD:0.7)	**3.8 (SD:1.0)**	**4.4 (SD:0.7)**

the difference is significant ($F(1,11)=175.2, p<.001, \eta_p^2 = .94$). FootSketch's average recognition accuracy of the five gestures for all participants is 85.8% (SD:9.5%). Participants performed the stomping gesture on average 1.3 times to successfully trigger the system. Subjective ratings show that FoodSketch was perceived as being more useful and satisfying than TMI. Wilcoxon Signed Rank test results show the difference was significant for usefulness ($Z=-2.14, p<.05$) and satisfaction ($Z=-2.90, p<.05$).

CONCLUSION & FUTURE WORK

Through the survey study, we identify that almost all people encountered situations when they wanted to use the mobile phone but both hands were occupied, and people would like to do 2 tasks on average with the device in these scenarios. We discovered that user-defined foot gesture interaction can save over 70% of the time compared to a traditional mobile interaction that requires users to release hands and take their phone out of pocket. This work is an initial exploration of using foot-gestures in realistic hands-occupied scenarios and can be improved in following ways. We used a smartphone attached to the ankle as a prototype but foot chain would be an ideal form. There are other ways to perform TMI, such as moving the bag in one hand to the other, which may cost less time than the TMI examined in our study. We required participants to stand still while sketching foot gesture. However, people may walk when their hands are occupied. Whether people can perform foot gestures while walking and how to design better triggers for such scenarios must be further examined. Social and cultural acceptability of performing foot gestures is worth further exploration too.

REFERENCES

[1] Alexander, J., *et al*. 2012. Putting Your Best Foot Forward: Investigating Real-World Mappings for Foot-based Gestures. *CHI'12*, 1229-1238.

[2] Crossan, A., Brewster, S. and Ng, A. 2010. Foot tapping for mobile interaction. *BCS HCI'10*, 418–422.

[3] Fujimoto, M., *et al.* 2009. A Motion Recognition Method for a Wearable Dancing Musical Instrument. *ISWC'09*, 11-18.

[4] Han, T., *et al*. 2011. Kick: investigating the use of kick gestures for mobile interactions. *MobileHCI'11,* 29-32.

[5] Lv, Z., *et al*. 2014. Multimodal Hand and Foot Gesture Interaction for Handheld Devices. *ACM Trans. Multimedia Comput. Commun. Appl. 11, 1s, Article 10* (Oct, 14), 19 pages.

[6] Saunders, W. and Vogel, D. 2016. Tap-Kick-Click: Foot Interaction for a Standing Desk. *DIS'16*, 323–333.

[7] Schmidt, D. *et al.* 2014. Kickables: tangibles for feet. *CHI'14*, 3143–3152.

[8] Scott, J., *et al*. 2010. Sensing foot gestures from the pocket. *UIST'10*, 199-208.

[9] Velloso, E., *et al.* 2015. The Feet in Human--Computer Interaction: A Survey of Foot-based Interaction. *ACM Comput. Surv.* 48, 2 Article 21 (Sep. 2015), 35 pages.

[10] Yamamoto, T., Terada, T. and Tsukamoto, M. 2011. Designing gestures for hands and feet in daily life. *MoMM'11*, 285-288.