Design Space for Finger Gestures with Hand-held Tablets

Katrin Wolf
TU Berlin
Telekom Innovation Laboratories
Ernst-Reuter-Platz 7, 10587 Berlin
Germany
katrin.wolf@acm.org

ABSTRACT
This paper presents research how a finger-gesture design space for interacting with hand-held tablets may be defined. The parameters that limit or extend this space, such as anatomy-dependent gesture feasibility, grasp requirement, gesture occlusion or complexity, are discussed based on initial explorative expert interviews and following user studies. The goal of this research is defining parameters that have to be taken into account for developing a finger-gesture UI model for hand-held tablets. Although this model has a strong user-centric design approach, rather than being technology driven, technical solutions for detecting finger gestures are also considered. A model design is presented and research questions for investigating this model in greater detail are outlined.

Categories and Subject Descriptors
H.5.2 [Information inter-faces and presentation]: User Interfaces - Ergonomics; Haptic I/O.

Keywords
Gestures; back-of-device; feedback; occlusion.

1. INTRODUCTION
Since human-computer interaction has become mobile and is no longer limited to mouse and keyboard, there are now many possibilities for interacting with computers, e.g. through free-hand gestures, touch, speech, device-gestures, gaze. The interactions with desktop computers are – compared to mobile devices – well investigated and paradigms for designing desktop interactions, such as WIMP, direct touch, and instrumental interactions [3] have been established. Current mobile interfaces mostly rely on the established paradigms mentioned above or on some modification of them with respect to a mobile scenario and different device form factors. For instance, the UIs on mobile phones usually make use of icons and menus but in a different appearance, such as icon groups on the iPhone. The greatest difference between mobile UIs and desktop computers is probably having a single device with far fewer physical buttons and which has a screen, which has replaced the mouse and touchpad by being able to detect finger-touches directly screen positions of icons. The thumb, with respect to mobility, is replacing mouse and cursor, but can also lead to problems, such as content occlusion and the "fat finger" problem.

In recent years researchers have developed interfaces that avoid occlusion and that allow access to the back of a hand-held device for input, such as pointing from the rear at content that is displayed at the user-facing screen [2,16]. Widgor *et al*. [16] describe how a hovering finger may be sensed with an external camera and Baudisch and Chu [2] use a back-of-device touch-pad for input. Moreover, the first game controller with a rear touch-pad was recently released by Sony. In a similar manner to the research projects mentioned above, this device uses back-of-device touch events for absolute pointing at graphical elements at the front screen. Devices that are larger than Apple's Nano Touch device [2] or the Sony Vita console, such as pad devices or tablets, are often held in such a way so that the fingers are already touching the rear surface. This can result in input errors through unintentionally released touch events. This phenomenon is known as the "Midas touch" problem [7], and is already known to cause problems with gaze input [8] through misinterpreted eye movements.

Back-of-device interactions address the problem of content occlusion caused by the fingers pointing directly at the displayed content. However, such an approach requires the ability to distinguish between intentional and unintentional touch-event releases. In addition, it might be the case that areas in the middle of a tablet screen may not be easily reachable from the device rear and therefore direct pointing may not be the best solution for back-of-device interactions. In order to address these problems as well as designing interaction techniques with respect to ergonomics, error resistance and mobility, this paper describes an approach for developing a mobile UI model that strongly focuses on capitalizing on the grasp and fingers that already hold the device.

2. RELATED WORK
2.1 Interacting while grasping
As Wimmer and Boring [17] suggest, detecting how users hold their devices could serve as a means for designing new implicit and explicit interactions as well as a way to enhance existing mobile interfaces. Researchers have already developed interfaces that rely on releasing grasping fingers or on re-adjusting the grasping hand. Kratz et al. [9] and Wigdor et al. [16] detect finger movements around grasped devices, and Baudisch [2] interpret back-of-device touch events as pointing commands. These projects mainly focus on specific interaction techniques that have become possible through novel interfaces. So far no research has been carried out that summarizes existing back- and around-device interaction techniques or that investigates UI requirements

with respect to a user-centric and ergonomic design. Moreover, the problem of content occlusion might be solved through back-of-device interactions; however, this may lead to the problem of gesture occlusion by the device itself. If visual information is not available, humans tend to rely more on other modalities. Usually in HCI, other system modalities, such as audio is provided if vision in not available. So far the potential for humans' internal feedback modalities, such as touch and kinesthetics have not been investigated. Therefore, in addition to ergonomics, internal modalities for gestural motor control are also relevant aspects that ought to be taken into account when modeling a gestural mobile UI for hand-held devices in general and for tablets specifically.

2.2 Ergonomic studies

The goal of ergonomic design approaches is to develop products, machines and working environments that are suited to human body size, and in such a manner that the user does not incur a physical or psychological penalty even over a long period of usage. Therefore handles are designed to fit in user's hand. For human-computer interaction design that relies on back-of-device finger movements, an ergonomic design could rely on hand sizes and finger lengths. Ergonomic studies for designing ergonomic physical products have been carried out. Lange [10] and Tilley [13] collected human's body sizes as well as degrees of freedom for flexing, bending and moving body parts, such as hands and arms. And Otl Aicher [1] described the ergonomics of grasping regarding the design of handles and hand-held tools. The ergonomics of human-computer interactions is a much younger discipline and therefore a lot of work remains to be carried out in order to understand how ergonomic hand-movement based interactions should be designed. Some lessons can be learned from research driven by prosthesis design. For instance, Feix *et al.* [5] merged several grasp taxonomies and presented an overview of the three main grasp types and their variations according to different grasp functions as well as form factors of the grasped objects.

2.3 Motor control modalities

In HCI research, feedback is a central issue regarding good interaction design. However usually system feedback of the computer is meant, and humans' internal feedback modalities have so far not been investigated with respect to their potential for supporting interaction design. The benefit in investigating human feedback modalities is to keep systems simple. Moreover in mobile scenarios there is a lot of audio and visual noise that may overlay or distract from system feedback. During the last ten years in cognitive science a lot work has been done in investigating internal feedback modalities for motor control: The Sensorimotor Adaption Model [14] is based on information of multiple sensory modalities that serve to monitor physical movement. For instance, vision and proprioception both provide information about hand movements. Vision had been thought to dominate this process [15], but studies, such as [14] have shown that the sensorimotor system weights modalities based on their information quality. If a modality loses information, such as when poor lighting reduces vision, proprioceptive information is given more weighting, becoming the dominant feedback control modality. Taking Sensorimotor Adaption Theory as inspiration, it can be suggested that proprioception can serve as an internal feedback modality when the gestures are occluded without decreasing back-of-device interaction performances, such as pointing or scrolling.

3. DESIGNING A GRASP-BASED UI

3.1 Research questions

The general question that underlies the development of a fingergesture based UI is: to what extent can users interact with hand-held tablets through tiny microgestures that are performed while grasping? As there are several interaction design aspects to be taken into account, three sub-questions are leading the research for the model design:

Q1: Which finger movements are feasible while grasping?

Q2: Which touch positions and finger movements should be excluded from the design space to avoid the Midas touch problem?

Q3: What are the appropriate microgetures to use for common tasks such as pointing, selecting, and sliding?

Q4: How should microgestures be supported through guidance and end-of-gesture feedback?

The methods for answering these questions are interviews with experts in anatomy and human physiology as well as evaluations of interactive prototypes.

3.2 Feasibility based design space

3.2.1 Gestural atoms

In expert interviews [18] we asked physiotherapists and sport scientists, which gestures were easily feasible while holding objects in three main grasp types [5]. For all grasp types, just touching or sliding on the grasped objects using the thumb, index and middle fingers, were identified by the experts as being easily feasible without decreasing the stability of the grasp.. Actually these two movements can be slightly modified or combined and then a lot of potential gestures can be designed using these gestural "atoms". Any further gesture design will basically be built out of active touches and slides. Just by modifying the intensity, speed or acceleration of touches and slide movements means that taps, press events, flip gestures, and slide movements of several shapes, such as lines or circles, are added to the design space.

3.2.2 Touchable surface regions

Depending on the type of device, the grasp, and the finger lengths of the user, some regions of the device surface are already touched when the device is held, while others would be still easy to reach through releasing and readjusting single fingers or one of the two grasping hands (while still maintaining a loose grip on device with this hand). However, some surface regions might be harder or impossible to reach without risking dropping the held device or it slipping towards the floor. In a user study (see Fig. 1) we

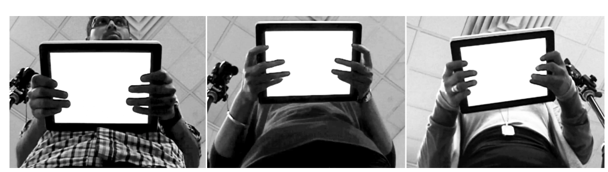
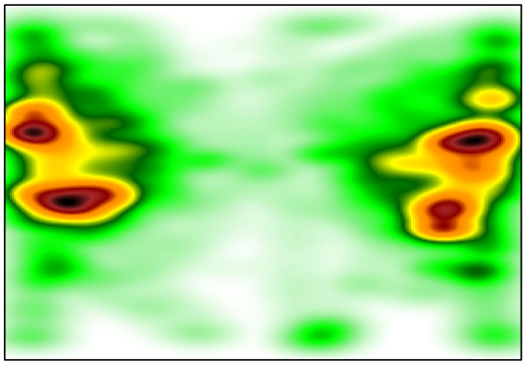


Figure 1. Participants were asked to touch with every finger and the thumb all possible device regions through drawing on the touch screen.



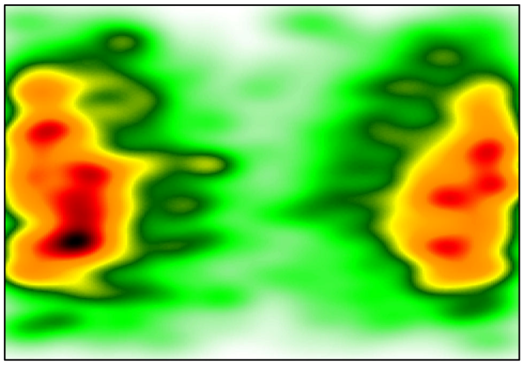


Figure 2. The touched areas at the hand-held tablet's front (top) and rear (bottom) are presented through heatmaps. The locations that are touched while grasping are displayed in dark orange, red, and brown.

investigated the touchable regions on the front and on the rear of held tablets (see Fig. 2). One fact that restricts UI design a lot is the unpredictability of the grasp and unintentional touch events; based on how different people hold devices. Therefore touchable regions can serve for placing touch-based interactions, but because the finger touches that are necessary for holding the device and those that indent interactions are located in the same area, these interactions have to be labeled for been differentiated from the natural grasp touches. Gesture combinations or touch releases could potentially provoke fewer errors. Also gestural trigger events that label gestures, similar to push-to-gesture or end-of-gesture events can help to design error resistant interactions.

3.2.3 Guidance feedback

Back-of-device interactions, such as finger movements and touches at the device's rear are occluded by the held device. The section on related work outlined how missing visual information is substituted with the hand behind the device [16] or fingers that are touching the device [2] are visually simulated on screens at the user facing device side. In many mobile scenarios a user's visual attention might be needed for other issues, rather than being directed to the display of a mobile device. Much work has already been done on designing non-visual displays [12], progress information [4] and end-of-gesture design [11]. Inspired by the Sensorimotor Adaption Model [14] we investigated whether widgets that are displayed on the screen of a mobile device can be pointed to from its rear without either system feedback or displayed touch positions. We discovered that users do not feel confident with absolute pointing from the back of the device, but that their performance without additional visual system guidance is as good as when it is present, provided that they point relatively from their current finger position [19].

3.3 Technology

Even though the presented work has a user-centric approach, many human-computer interfaces are rather technology-driven developed than designed focusing on human skills and needs With respect to mobility and changing environments any camera-based approach would cause too much noise as well as limiting mobility, (and therefore scalability to a large number of use cases). For any touch-based gestures on the device surface, we often used common touch screens. For gestures that might be hard to detect with that technology, we used inertia sensors that are worn on the fingers, which perform the gestures (see Fig. 3). Therefore it becomes possible to classify gestures that would not easily be distinguished from simple touch and slide gestures with common capacitive sensors, such as pitch or rapid releases that serve as end-of-gesture event. This aspect actually points to a big challenge in grasp-based interfaces: the Midas touch problem that describes unintentional movements that are interpreted by computers as commands.

3.4 Conceptual mismatches

Midas touch problems and many other input errors usually cause gesture classification errors because of certain differences between humans and computers that can be summarized as conceptual mismatches. These errors include instances where natural movements are misunderstood by the computer as gesture, as well as instances where intentional gestures performed by the user are either not recognized by the system or else wrongly classified. Even the simple touch gesture that is positioned by users at a certain point is not exactly measured by common capacitive screens at the same position [6]. However in this case the error is small enough that it does not cause completely wrong and unexpected system reactions.

The raw data of inertia sensors has a completely different structure than those which serve for humans recognize and control body movements. For instance, if a gesture is done without acceleration, in human communication this gesture might still be understood. Computers excuse far fewer modifications in gesture execution. The interactive prototype that is shown in Fig. 3 classifies swipe gestures through accelerations in certain directions.

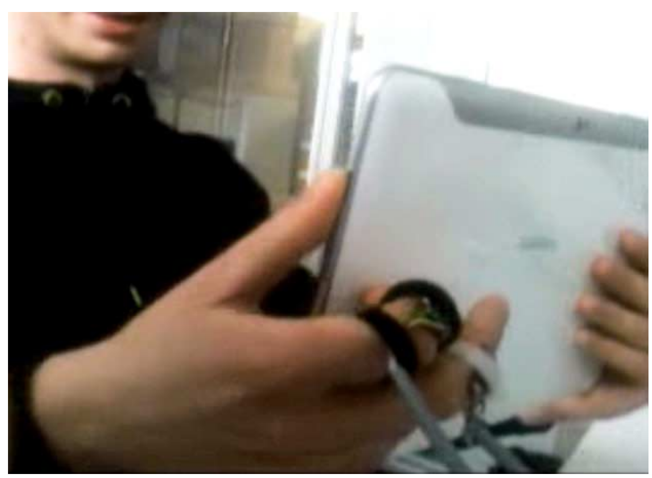


Figure 3. Inertia sensors (accelerometer and gyroscope) augment fingers for detecting gestures beyond touch and slide.

Because finger worn sensors are used, the gesture can be executed at the rear surface or in the air. To avoid tiny finger movements that release a command, a tap is labeling the swipe gesture. In the displayed interface, the index and the middle finger perform the gesture, and if the index finger is tapping, the middle finger might unintentional tap as well. Therefore having the same gesture for all fingers might be easy to learn but is also very error prone. Alternative gesture triggers could rely on other modalities or also on using two fingers from two hands instead of just one.

3.5 Future work

User studies on exploring the ability to perform gestures with more than one finger at the same time are currently underway, and studies on the performance of absolute and relative pointing as well as the requirements and needs for feedback will be carried out in the future. Moreover, a prediction model for Midas touch error for certain gestures and their trigger events will be developed, because successful handling of the Midas touch problem very much influences the success or failure of a gestural interface. All these findings will lead to the development of a post-WIMP paradigm for hand-held devices that focuses on mobility, ergonomic gesture design, and suitability for tasks that require visual attention or that cannot use audio due to environment noise, or when auditory input and output is not suitable for the scenario.

4. SUMMARY

This paper presents an approach for developing a UI paradigm for finger gestures with hand-held devices using the example of tablets. The work is still in progress and at this stage, parameters that ought to be taken into account are introduced based on the requirements of mobile scenarios and the focus on ergonomic and user-centric design. For designing a scalable model, the parameters are hard to define statically and ad-hoc layouts and dynamic interface component setups seems to be promising. Also non-visual and non-audio feedback modalities are promising with respect to providing interaction support and feedback. Finally, any gestural design has to be different from natural behavior so as not not provoke Midas touch events. This work will be continued during the next 1.5 years and finally be published as my Ph.D. thesis.

5. ACKNOWLEDGMENTS

I would like to thank Sebastian Möller, Michael Rohs, and Robert Schleicher for their valuable advice and input.

6. REFERENCES

[1] Aicher, O., Kuhn, R. Griffe und Greifen, Köln, 1995.

[2] Baudisch, P., Chu, G. 2009. Back-of-device interaction allows creating very small touch devices. In *Proc. CHI 2009*, 1923-1932.

[3] Beaudouin-Lafon, M. Instrumental interaction: An interaction model for designing post-WIMP user interfaces. In *Proc. CHI 2000*, 446–453.

[4] Brewster, S. and King, A.. 2005. An investigation into the use of tactons to present progress information. In *Proc. INTERACT* 2005, 6-17.

[5] Feix, T. et al. Grasp Taxonomy Comparison Sheet, DOI = http://web.student.tuwien.ac.at/~e0227312/ documents / taxonomy_comparison.pdf.

[6] Holz, C. and Baudisch, P. Understanding Touch, In Proc. *CHI 2011*, 2501-2510.

[7] Istance, H., Bates, R., Hyrskykari., Vickers, S. Snap clutch, a moded approach to solving the Midas touch problem. In *Proc. Symposium on Eye tracking research 2008*, 221-228.

[8] Jacob, R.J. 1990. What You Look at is What You Get: Eye Movement-Based Interaction Techniques, In *Proc. CHI 1990,* 11–18.

[9] Kratz, S., Rohs, M. 2009. Hoverflow: exploring around-device interaction with IR distance sensors. In *Proc. Mobile HCI 2009*, 4 pages.

[10] Lange, W. Kleine Ergonomische Datensammlung, Köln, 1991.

[11] McAdam, C. and Brewster, S.A. *Distal Tactile Feedback for Text Entry on Tabletop Computers*. In Proc. BCS HCI 2009, 504-511.

[12] Sawhney, N. and Schmandt, C.. 2000. Nomadic radio: speech and audio interaction for contextual messaging in nomadic environments. *ACM Trans. Comput.-Hum. Interact.* 7, 3 (September 2000), 353-383.

[13] Tilley, A. R. The Measure of Man and Woman: Human Factors in Design, New York, 1993.

[14] van Beers, R. J., Wolpert, D. M., Haggart, P. 2002. When Feeling Is More Important Than Seeing in Sensory Adaption, *Current Biology*, Vol. 12, May 14, 2002, 834–837.

[15] Welch, R. B., Warren, D. H. 1986. Intersensory interactions. In *Handbook of Perception and Human Performance*, Vol. 1, 1986, 25.1–25.36.

[16] Wigdor, D., Forlines, C., Baudisch, P., Barnwell, J., Shen, C. 2007. LucidTouch: A See-Through Mobile Device, In *Proc. UIST 2007,* 269-278.

[17] Wimmer, R., Boring, S. HandSense - Discriminating Different Ways of Grasping and Holding a Tangible User Interface, In *Proc. TEI 2009*, 359-362.

[18] Wolf, K., Naumann, A., Rohs, M., and Mueller, J. 2011. A Taxonomy of Microinteractions: Defining Microgestures based on Ergonomic and Scenario-dependent Requirements, In *Proc. INTERACT 2011*, Springer-Verlag, Berlin, Heidelberg, 559-575.

[19] Wolf, K., Müller-Tomfelde, C., Cheng, K., and Wechsung, I. 2012. Does proprioception guide back-of-device pointing as well as vision? In *Proc. CHI 2012*, 1739-1744.