Gesture Interfaces for Mobile Devices – Minimalist Approach for Haptic Interaction

Jukka Linjama
Nokia Technology Platforms
Itämerenkatu 11-13, 00180 Helsinki, Finland
jukka.linjama@nokia.com

Jonna Häkkilä, Sami Ronkainen
Nokia Multimedia
Yrttipellontie 6, 90230 Oulu, Finland
jonna.hakkila@nokia.com,
sami.ronkainen@nokia.com

ABSTRACT
Due to the limited input and output functionalities and varying usage situations, haptics offers a potential approach for interacting with mobile handheld devices. In this paper, we discuss the specific features related to haptic interaction methods for mobile devices and argue for a minimalistic gesture control. In addition, the paper presents a haptic mobile phone application utilizing this design principle, and implementing sensor based movement detection and haptic feedback.

Author Keywords
Mobile devices, user interface, haptic interaction, tactile feedback, gesture input, acceleration sensors, multi-modal interfaces

ACM Classification Keywords
H5.2. User interfaces: Haptic I/O, Input devices and strategies, Interaction styles.

INTRODUCTION
Mobile phones are typically used in very different kinds of situations. Often this includes situations, where user's full attention cannot be focused on the device or visual contact is hard, e.g. while riding a bicycle. The input methods used when interacting with the phone are limited, as they conventionally include small keypad or a pen-input. Thus, the actions usually require several button presses or pen strokes which makes the task completing slow, and the phone is hard to use for instance with gloves on. As an optional input technique, several mobile phones employ voice commands as an alternative input modality. However, speech input is often not fully functional option: it may not be socially suitable in a quiet environment, and the speech recognition may fail in noisy surroundings. In addition, it is often perceived as awkward to use in social situations. Addressing to these problems, haptic user interfaces offer an alternative way for interacting with a mobile phone.

Haptic interfaces in mobile devices have many limitations in comparison to desktop usage, where e.g. precise position input is possible, and enough energy is available for force feedback. However, there are opportunities for haptic interfaces in mobile device domain as well. Despite the limitations in the actual fidelity in haptic sensing and feedback, the use context of mobile devices raises many needs for the end user that haptic interfaces can help to satisfy. For instance, there is a need for a mobile device user to be able to

- control the device on the move, without access to keys
- know the status of the device or application with very limited access to the display

Haptic input methods have so far been used to modify the visual output on the screen. This has been demonstrated e.g. by [1], who use tilting the device for scroll content on a PDA display, and [4]. who additionally orientate the portrait and landscape view according to the device orientation. such as zooming or scroll the display. Context recognition has employed sensors to recognize different movements of the device, which information has then be used as an information source to determine e.g. if the user is walking or running [6], [3].

This paper addresses some of the opportunities with a demo [5] of minimalist gesture interface, which utilises the combination of inertial sensing and vibration feedback.

GESTURE INPUT
From the perspective of mobile devices, the technology used for gesture recognition becomes important. Traditionally, many gesture recognition systems have been based on visual recognition of gestures from a video stream, as e.g. in system described by O'Hagan et al [9]. Even though e.g. mobile phones equipped with an inbuilt camera are becoming more popular, the possibility of utilizing the technology for gesture recognition is limited. Robust

recognition of gestures from a video stream originated from a shaking handheld camera is a challenge for the recognition algorithms. Furthermore, this kind of gesture recognition requires that the user has prepared the device so that its camera is pointed at the user. This often requires several steps, limiting its usage possibilities. For instance, utilizing a gesture for silencing a phone when it is ringing, is pointless if the first steps in the interaction are to dig the phone up from a handbag and point its camera at the user.

Another approach for gesture recognition is to embed position and/or acceleration sensors into the device itself, so that the user makes the gestures using the hand holding the device. This kind of approach has been described e.g. in the XWand concept [11].

Interestingly, when talking about mobile devices, the word "handheld" is often used. However, mobile devices spend most of their time (when not stored off the user altogether as e.g. during charging) not in the user's hand but in a carrying location such as a pocket or a handbag. A device only becomes handheld when the user has dug up the device from its carrying location.

Hence, an interesting field for gesture recognition in relation to mobile devices is that of minimizing user interaction by allowing gesture input when the device is still stored in its storage location. In this paper, we introduce one such way of interaction, namely the tap input.

USER NEEDS IN THE MOBILE USE CONTEXT

When mobile devices become capable of providing more and more services for a user, the likelihood of the user wanting to interact with the device while engaged in some other activity (such as walking) becomes higher. This kind of ultra-mobile computing has been described by Schmidt et al [10] as computing devices that are operational and operated while on the move, as opposed to devices where the device is intended to be used out of the office but still reserving time and space to concentrate on the device usage.

The ultra-mobile computing brings new requirements for interaction, both in the field of traditional user interface design when applications may have to be used on the move, and in the field of new applications emerging from the mobile usage context. Examples of such new requirements are:

- Need to configure the device on the move, e.g. silent mode
- Need to be able to continue/terminate application when the usage context changes, e.g. from outdoors to office
- Need to share your data instantly, e.g. select Bluetooth pair and send data to it
- Need to easily check/know what the device state or mode is, e.g. that the device is secured (key lock)

Some of the required actions do not necessarily require lots of interaction (as they are simply one-shot activities such as changing a device profile or silencing an incoming call alert). On the other hand, the need to do the actions may occur unexpectedly (e.g. the incoming call or message). In those cases, minimal interaction (including avoidance of having to dig up the device from its carrying location, or to unlock the keys of the device) would be desirable. Haptic interaction could produce solutions to fulfil this, as with it the user generally does not have to find or feel buttons or look at a display. This kind of usage scenario occurs often in multitasking situations, where only limited attention can be focused on the mobile device.

As providing feedback is one of the key elements of successful UI design [7], a response for haptic input is needed to correspond the recognition of the user-initiated action. Following the haptic input, using the sense of touch is also an intuitive output channel. With mobile handheld devices, haptic feedback can be implemented by using a simple vibrator component.

HAPTIC SENSORS – TOUCH AND MOTION

Many of these challenging needs can be satisfied, if new input devices are introduced. Especially mechanical location, motion or force sensors have this kind of potential. These can be called also haptic sensors (from Greek “haptikos”, to touch or grasp), using an analogy to human sensing and action. Human action is inherently motoric by nature, and the haptic sensing (sense of touch and kinaesthetic sense) is the necessary feedback channel for regulating our motor activity.

Inertial (acceleration) sensors have been used in many interface concepts. They can indicate the motion of the device in 1 to 3 dimensions. User can thus move the device and indicate some controls by gestures. A 3-D sensor can indicate also the tilt angle of the device (using gravity as reference).

This has been used e.g. in MyDevice concept device from F-Origin Ltd. [2], which uses the tilt angle measurement to change the screen layout from vertical to landscape depending on the holding orientation. Also tilt movements, together with key presses, are used to control the scrolling or zooming the web pages. Tilt movements with associated haptic feedback in scrolling have been studied e g in [8].

There are challenges for gesture input to be usable. Large motions are not always preferable due to social restrictions. Also the indication of the gesture command start and end is not easy to solve in practice. One solution is to use only minimal set of gestures, which can be detected continuously.

Simply tapping the device in any direction may act as a robust and clear control gesture. Demo of this minimalist “Tap-it” gesture interaction was presented in NordiCHI [5], and discussed in this paper further.

TRUST YOUR DEVICE: NEED FOR FEEDBACK

Crucial for any novel input technologies is to understand the possible lack natural feedback. Touch screens do not provide the click feeling (and sound) that is always present in mechanical keys. Motion gesture sensors can recognise the gestures, but the user may not be sure when the recognition started, or did it start at all.

Feedback can be done with all available modes (graphics, light, sound, haptic vibration or force). The challenge is in the content design:

- Feedback must be immediate enough (no latency in SW)
- All output modes are in sync and support each other (feel and sound "match")
- Feedback must not be excessive (not too long, too strong) but still perceivable in wide range of usage conditions

Tactile feedback in phones

The traditional vibrating component in mobile phones is a tiny rotating motor having an eccentric weight on top of the shaft (see Figure 1).

Figure 1 Vibrating motor component: DC motor with counterweight, rotating typically in 150 – 200 Hz range.

The feeling of the vibration produced by a rotating motor is relatively soft. However, many applications are possible with this inexpencive and existing technology, using more advanced electronic control than pure on/off drive with the nominal voltage.

TOWARDS MULTIMODAL INTERFACES

Innovations lie in combinations of input and output technologies. Traditionally, the term "multimodal" has often been used for interfaces, where the speech modality is added to the graphical UI. However, in mobile context, haptic sensing and feedback offers a new dimension to multimodal interactions.

Our demo is "just" an entertaining application but it illustrates some important elements of haptics in a multimodal experience

- Minimalist input – touch by tapping/drumming
- Minimalist haptic feedback – vibration pulse
- These haptic elements in sync with graphical content and metaphor of moving, tangible elements with natural dynamics (rubber ball)

Figure 2. Handheld demonstration device, equipped with 2-D acceleration sensor, display and vibration motor.

DEMO DESCRIPTION

Interaction concept

The demonstration concept in this paper has focus in the "low-fi" extreme of basic interactions that might be useful and simple to implement in mobile devices [5]. Motivation is to

- Use the most simple and robust gesture input – finger "tap" detected by an acceleration sensor
- Associated, naturally matching vibrotactile feedback – "bump" kick produced by the standard vibration motor
- Find potential applications for the combinations of these two, together with or without graphics

By carefully integrating these specific haptic I/O events with graphical content enables multi-modal, visuo-haptic experiences to be created.

Input. Specific gesture, drumming the device with fingers, is used for controlling the device. Motion sensor (2-dimensional accelerometer, ADXL202E) is tuned to detect the tiny motion of the device caused by finger taps on the surface of the device. Basically, drumming in any direction and position on the surface of the device is detected. Simple trigger from acceleration value exceeding approx +/-2 g is interpreted as the control input event in X or Y directions.

Output. Output from the device is provided by graphics and tactile means. Vibrating alert actuator is used to give a "kick" pulse (duration 20 ms) that briefly shakes the device.

Application: bouncing ball game

Example application running in the demo is the bouncing ball game. Position of the ball is calculated and displayed on the display. By tapping the device in X or Y direction user can put the ball moving in corresponding direction on the screen, see Figure 2. It bounces from the walls, and these events are supported by short vibration bump feedback. Simulation model includes a centering spring and damping effects, so that finally the ball returns to the center of the screen. Synchrony of graphical and tactile events (tap input and kick feedback) creates a kind of a kinesthetic illusion of a soft ball being tapped and bouncing inside the device. In informal evaluations most users rated this illusion very natural, impressive, and enjoyable.

CONCLUSION: POSITION STATEMENT

In mobile device usage and technology context, minimalist gesture control and related haptic feedback have potential. Haptic content fidelity can be rather low, but if the interaction and content design is in balance with technology, many innovations for mobile device control and applications used through the handheld devices become possible. Haptic interfaces must be designed to be multi-modal – visual and haptic content are in sync, and support each other. The simple bouncing ball game demo is intended to facilitate discussion and spark new thinking for mobile device interfaces.

ACKNOWLEDGMENTS

We thank Jukka Jalkanen for the SW implementation of the demo, and many others in NRC Helsinki for the actual realization of the PDP demo platform, on which this demonstration has been created.

REFERENCES

1. Eslambolchilar, P. and Murray-Smith R. Tilt-Based Automatic Zooming and Scaling in Mobile Devices – A State-Space Implementation. In *Proc. MobileHCI 2004*, (2004), 120-131.
2. F-Origin Ltd. http://www.f-origin.com/references/mydevice.htm (checked 10.1.2005).
3. Gellersen, H.W., Schmidt, A., and Beigl, M. Multi-Sensor Context-Awareness in Mobile Devices and Smart Artifacts. *Mobile Networks and Applications 7*, Kluwer Academic Publisher (2002), 341-351.
4. Hinckley, K., Pierce, J., Sinclair, M., and Horvitz, E. Sensing Techniques for Mobile Interaction. *CHI Letters 2*, 2, ACM Press (2000), 91-100.
5. Linjama, J. and Kaaresoja, T. Novel, minimalist haptic gesture interaction for mobile devices. In *Proc. NordiCHI2004*, (2004), 457-458.
6. Mäntyjärvi, J., Himberg, J. and Huuskonen, P. (2003). Collaborative Context Recognition for Handheld Devices. In *Proc. of PerCom’03*, (2003), 161-168.
7. Nielsen, J. and Mack, R.L. *Usability Inspection Methods.* John Wiley & Sons, Inc., Canada, 1994.
8. Oakley, I., Ängeslevä, J., Hughes, S and O’Modhrain, S. Tilt and feel: scrolling with vibrotactile display. In *Proc. of Eurohaptics 2004*, (2004), 316-323.
9. O'Hagan, R. and Zelinsky, A. Finger Track - A Robust and Real-Time Gesture Interface. In *Proc. of Australian Joint Conference on Artificial Intelligence*, (1997).
10. Schmidt, A., Beigl, M. and Gellersen, H.-W. There is more to Context than Location. *Computers and Graphics 23*, 6, (1998), 893-901.
11. Wilson, A. and Shafer, S. XWand: UI for intelligent spaces. In *Proc. of ACM CHI Conference on Human Factors in Computing Systems*, ACM Press (2003), 545-552.