Human-Computer Interaction with Mobile Devices

Introduction

Much has been written about mobile computers and the changes that have occurred in human-computer interactions as a result of their popularity. A review of the literature on this subject reveals that the many of the writings are referring to notebook or laptop computers when they use the term “mobile computer”. These form factors certainly opened up new possibilities for working while traveling, but they have failed to deliver on the promise of enabling one to work “anytime, anyplace”. (Perry et al. 2001) Even the term “mobile” may not be appropriate for laptop computers, as their size, weight, and relatively long boot times limit their use to instances when one can remain stationary for the duration of the task at hand with an available surface on which to set the computer. Contrast this to computing devices such as wirelessly-enabled PDAs and converged PDA phone devices, which truly can be used while one is moving about and thus can be characterized accurately as “mobile computers”. These mobile computers will be the focus of this review.

Differentiating factors for interface design

Antii Aaltonen and Juha Lehikoinen point out that the tasks for which mobile computers were designed are fundamentally different from those for which the desktop computer has been used. Originally meant for personal information management, mobile devices have only recently begun taking on additional functions such as communication, entertainment, and information retrieval. Unlike desktop computers, which are generally the focal point when in use, the mobile device is very likely to be competing with other tasks for the user's attention. (Aaltonen and Lehikoinen 2005) Mobile computers break assumptions that have been implicit in the design of fixed-location computer applications, and thus create new design challenges for human-computer interaction. (Dix et al. 2000) Chao Li and Katharine Willis describe the interaction between individuals and mobile devices as "a new kind of human-computer interaction" where the surrounding environment is being brought into consideration. (Li and Willis 2006) Stripping the interaction with a computing device to its two most basic functions, input and output, gives us a good starting point for both reviewing the literature and suggesting topics for further research. Though this order may seem reversed, this discussion will begin with output considerations, as this is the aspect that one is likely to notice first.

Output considerations

The physical size of mobile computers is their most obvious distinguishing characteristic from desktop computers, so it is natural that HCI researchers have focused on the physical characteristics that impose constraints in designing user interfaces. Paelke et al listed these constraints on visual presentation designs for small screens:

- Limited resolution: Typical resolutions range from 100*80 pixels for mobile phones to 240*320 for PDAs, as compared to mega-pixel desktop displays. This constraint must be addressed by interface designers.
- Small display size, with limited screen real estate for information display and interaction.

- Limited processing power: This limits the use of some animation and presentation techniques that are commonly used on desktop computers. (Paelke, Reimann, and Rosenbach 2003)

An example of the effects of these constraints that the average mobile device owner encounters frequently is in presentation of Web content. Lari Karkkainen and Jari Laarni contend that despite the increasing availability and capabilities of wirelessly-connected devices, there is still a lack of Web design guidelines specifically for these devices. They state that though there are many resources for guidelines for Web publishing, these guidelines must be modified when publishing Web pages for mobile computers. They divide existing guidelines for Web design into three categories: (1) guidelines related to technical aspects of computers and communications; (2) guidelines related to content; and (3) guidelines related to layout and aesthetics. (Karkkainen and Laarni 2002)

In the first category of Web design guidelines, Karkainen and Laarni include recommendations regarding connection speeds, browser capabilities, and processing capacity that a Web designer should assume in creating pages. When adapted for mobile devices, the recommended parameters must be scaled down to match the hardware configuration of the mobile device. According to research by Virpi Roto and Antti Oulasvirta, even with network connections used for mobile browsing now becoming faster with 3G and other networks, Web browsing on a mobile device may require page download times of more than 5 seconds. They observe that network connection speed is a bottleneck, and the processing power of the device contributes to the problem by making page rendering much slower than on a PC. They argue that the most important response is when the first visible part of the page appears, not after images and other components have loaded. They define system response time as the time between “New page request” and “New page partially visible”. At this point, the user can start reading, scrolling, and following links to other pages, so the remainder of the information yet to be loaded is irrelevant in many cases. In their user studies, Roto and Oulasvirta noticed that the user’s attention shifts away from the screen before the page arrives, requiring that he constantly glance back at the screen to determine when the page loading has reached the desired stage. They advocate capitalizing on mobile device capabilities by using tactile feedback such as vibration to signal the user that the page has loaded.(Roto and Oulasvirta 2005) Tactile feedback is thus introduced an entirely new output mechanism, specific to mobile devices.

Kaarkainen and Laarni’s second category in Web design, guidelines related to content, also requires revision. Laarni referred to the small screen as a “keyhole though which only a limited amount of information can be seen at a time.” (Laarni 2002) Thus it is necessary to rearrange or reformat the content in order to eliminate the requirement for excessive scrolling by the user. Laarni advocates using dynamic presentation of text, and he compared empirically such methods as vertical continuous scrolling, horizontal continuous scrolling, teletype mode, and rapid serial presentation of one word at a time in the center of the screen. (Laarni 2002) I would argue that the user of a mobile device is rarely interested in reading large amounts of text. Instead, designers should use editing and summarizing judiciously to present just that content that would likely be of interest to a mobile user. Content management systems that automate the process of reformatting

Web content for multiple devices and screen sizes must do more than simply revise the layout and leave the content unchanged. To avoid this, content providers must be able to indicate to the content management system which information is to be displayed and which is to be discarded in the reformatting process. An alternate approach is offered by Jesse Steinberg and Joseph Pasquale, who contend that users of mobile devices should be able to customize their view of the web by removing data they are not interested in, filtering images, and displaying web pages in formats that are easier to surf. They have developed middleware "Customizers" to enable these changes. (Steinberg and Pasquale 2002)

The third of Karkkainen and Laarni's categories in Web design guidelines, layout, is of utmost importance in the usability of a mobile Web site. Laarni showed that small screens slow down reading speed by disrupting normal eye movement patterns, and that much time is spent manually scrolling or paging. (Laarni 2002) Designers should consider the site's purpose and design it based on task analysis rather than the design used for conventional computers. (Karkkainen and Laarni 2002) Though some of these techniques are by now well-known, the number of web sites that do not include adjustments for mobile devices is still unacceptably high.

Many researchers have noted that users are often engaged in other activities while using the mobile device. Oulasvirta et al, in extensive studies of page-loading on the mobile device while the user is performing other tasks, found that attention to the mobile device broke down to bursts of just 4 to 8 seconds during the page loading tasks. As a result, users often slow down, postpone, or stop interaction with a mobile device because of the cognitive demands of other tasks. (Oulasvirta et al. 2005) One way to address this attention issue is to consider another new form of output that many mobile devices are capable of providing—voice. Perry et al pointed out that voice output, with its low demand for visual attention, can be particularly useful when the user is engaged in an activity such as walking or driving. (Perry et al. 2001) Software to provide voice output is not yet widely implemented, and is usually confined to specific voice-enabled applications rather than the general interaction with the mobile device. Technology advances will solve this problem over time, however, and HCI specialists can then concentrate on implementation. Voice output is appropriate for some social situations and not others, requiring that designers enable control over the presentation of output as text or voice.

Input considerations

Input mechanisms, as well as output mechanisms, on mobile devices present HCI challenges. As they did with output considerations, Paelke et al focused on the physical characteristics of mobile computers in identifying the following constraints:

- Lack of full alphanumeric keyboards:
- No mouse
- Low-resolution touch-screens

Keyboard input is implemented in a variety of ways, all of which pose challenges. (Paelke, Reimann, and Rosenbach 2003) Virtual keyboard images on the screen require use of a stylus to tap the keys; despite the fact that touch screens enable tapping with a finger, the key images are too small to make this feasible. Handwriting recognition, again requiring use of a stylus, has improved significantly in accuracy, but still poses challenges because writing on such a small screen surface is difficult. Some mobile devices have hardware navigation buttons that enable moving about the screen without use of the stylus or keyboard, but these buttons fall short of the functionality of the mouse on a desktop computer. Designing interfaces with "thumbable" buttons that can be selected without using the stylus improves usability.

An input mechanism that is rarely present on a desktop computer but is maturing as a viable technology on mobile computers is speech input. Voice recognition, like handwriting recognition, is improving rapidly in accuracy. Steve Love defines two categories of speech input: speaker-dependent and speaker-independent. Speaker dependent systems require the user to "train" the system by speaking samples of words to provide the system a template for her voice. Speaker independent systems have vocabularies created by taking speech samples from a cross-section of the population, and do not need training. (Love 2005) Currently speech capability requires special application software that consumes valuable storage and memory resources on the mobile device. Moreover, each voice application is limited in scope and independent of others on the device, utilizing different software, activation methods, and recognized vocabularies. This lack of standardization is a prevailing theme in mobile device mechanisms and applications, a fact that makes skills transfer among devices difficult and prohibits the systematic use of these mechanisms in user interfaces. (Paelke, Reimann, and Rosenbach

2003) As is the case with output technologies, these problems will be solved in due time, requiring designers to begin addressing the specific usability issues they generate.

Improving the user experience

Thus far, only physical characteristics of mobile devices have been considered. It is essential that we also consider the human factors, looking more closely at the user's experience. Alistair Sutcliffe warns, "If HCI fails to employ knowledge about the very people it is designing for then it is left with technology and creative inspiration." (Sutcliffe 2000) In particular, the effects of limited input mechanisms on the user experience precipitate a search for ways to reduce the amount of input required for various tasks. One way to address this issue is to bring the environment into consideration, as context information. According to Anind Dey and Gregory Abowd, one goal of context-aware computing should be to facilitate interacting with computers. (Dey and Abowd 1999)

Context has been defined in many ways, but for purposes of this discussion, a broad and intuitive composite definition is sufficient. Context is the set of physical, social, and environmental conditions in which the mobile device is being used, such as location, time of day, orientation, and social situation (eg, in a meeting). These contexts can be used as input to information-seeking processes in place of explicit input by the user. This use of context is seen in many tour guide systems, where the tourist's location is sensed by the mobile device and used as input to the information retrieval process. Thus the tourist is supplied information about her current location without being required to enter location data. This method could be applied in many more IR situations than is currently the case. For example, requesting the weather, movie listings, nearby restaurants, and so forth typically require the user to enter the city or the postal code rather than sensing the location and using that sensed information to perform the retrieval. Similarly, the mobile device could use sensed time of day as automatic input to display just the upcoming times for movies, bus arrivals, and other schedule-driven events rather than including times that have already passed. Of course there must be allowance in the design for the user to override the assumption that it is the *current* location or *current* time that should be used as input.

A more complex issue was raised by Cheverst et al in their research on whether the device should push information to the user automatically when her context changes, or wait until she requests it. Perhaps not surprisingly, their study generated mixed results that indicate that a combination of push and pull combined at the correct level should be further investigated. (Cheverst, Mitchell, and Davies 2002)

Some context-sensing mechanisms, such as those for sensing location indoors, are still immature and thus can generate ambiguous context. This causes faulty input to the information retrieval process, and resulting faulty output for the user. Anind Dey and Jennifer Mankoff point out that design must include provisions for dealing with incorrect or ambiguous context before we can expect realistic and deployable applications. In particular, the degree of and techniques for user mediation must be considered. Their design recommendations include the following:

- Interpretations of ambiguous context should have *carefully chosen defaults* to minimize user mediation, particularly when users are not directly interacting with a system;

- *Ambiguity should be retained* until mediation is necessary for an application to proceed.(Dey and Mankoff 2005)

Location-sensing is well-developed enough for user studies, however. Ilkka Arminen argues that to understand the dynamic nature of location, we have to study the user practices in which location gains its value. These kinds of user-centered studies have design implications as well, enabling the researcher to pinpoint users' requirements exactly at the moment when they emerge. (Arminen 2006)

Conclusion

Hardware characteristics of mobile devices require new approaches to human-computer interface issues as compared to those required for traditional computers. In particular, input and output limitations of mobile devices affect their usability and compel researchers to seek new mechanisms for interaction between humans and their mobile computers. Two new forms of output, voice and tactile, have been suggested. Two new forms of input have also been suggested: voice and context. Further research is needed into using each of these new mechanisms. As voice input and output technologies mature to the point that they can be used with all functions of a mobile device rather than within specific applications, situational usability studies will be needed. Context sensing is still almost as much art as science, and thus further research is needed on dealing with ambiguity and the appropriate measure of user mediation to enable.

Aaltonen, Antti, and Juha Lehikoinen. 2005. Refining visualization reference model for context information. *Personal Ubiquitous Comput.* 9 (6):381-394.

Arminen, Ilkka. 2006. Social functions of location in mobile telephony. *Personal Ubiquitous Comput.* 10 (5):319-323.

Cheverst, Keith, Keith Mitchell, and Nigel Davies. 2002. Exploring Context-aware Information Push. *Personal Ubiquitous Comput.* 6 (4):276-281.

Dey, Anind K., and Gergory D. Abowd. 1999. Towards a Better Understanding of Context and Context-Awareness. Atlanta, Ga: Georgia Institute of Technology.

Dey, Anind K., and Jennifer Mankoff. 2005. Designing mediation for context-aware applications. *ACM Transactions on ComputerHuman Interaction.* 12 (1):53-80.

Dix, Alan, Tom Rodden, Nigel Davies, Jonathan Trevor, Adrian Friday, and Kevin Palfreyman. 2000. Exploiting space and location as a design framework for interactive mobile systems. *ACM Trans. Comput.-Hum. Interact.* 7 (3):285-321.

Karkkainen, Lari, and Jari Laarni. 2002. Designing for small display screens. Paper read at Proceedings of the second Nordic conference on Human-computer interaction, at Aarhus, Denmark.

Laarni, Jari. 2002. Searching for optimal methods of presenting dynamic text on different types of screens. Paper read at Proceedings of the second Nordic conference on Human-computer interaction, at Aarhus, Denmark.

Li, Chao, and Katharine Willis. 2006. *Modeling context aware interaction for wayfinding using mobile devices, Proceedings of the 8th conference on Human-computer interaction with mobile devices and services*. Helsinki, Finland: ACM Press.

Love, Steve. 2005. *Understanding Mobile Human-Computer Interaction*. Edited by D. Avison, *Elsevier Butterworth-Heinemann Information Systems Series*. Oxford, England: Architectural Press.

Oulasvirta, Antti, Sakari Tamminen, Virpi Roto, and Jaana Kuorelahti. 2005. *Interaction in 4-second bursts: the fragmented nature of attentional resources in mobile HCI*, *Proceedings of the SIGCHI conference on Human factors in computing systems*. Portland, Oregon, USA: ACM Press.

Paelke, Volker, Christian Reimann, and Waldemar Rosenbach. 2003. *A visualization design repository for mobile devices*, *Proceedings of the 2nd international conference on Computer graphics, virtual Reality, visualisation and interaction in Africa*. Cape Town, South Africa: ACM Press.

Perry, Mark, Kenton O'Hara, Abigail Sellen, Barry Brown, and Richard Harper. 2001. Dealing with mobility: understanding access anytime, anywhere. *ACM Trans. Comput.-Hum. Interact.* 8 (4):323-347.

Roto, Virpi, and Antti Oulasvirta. 2005. *Need for non-visual feedback with long response times in mobile HCI*, *Special interest tracks and posters of the 14th international conference on World Wide Web*. Chiba, Japan: ACM Press.

Steinberg, Jesse, and Joseph Pasquale. 2002. A web middleware architecture for dynamic customization of content for wireless clients. Paper read at Proceedings of the 11th international conference on World Wide Web, at Honolulu, Hawaii, USA.

Sutcliffe, Alistair. 2000. On the effective use and reuse of HCI knowledge. *ACM Trans. Comput.-Hum. Interact.* 7 (2):197-221.

Anne Bishop
INLS 881

December 11, 2006