Refashioning our Interaction with Space

Miten Sampat
Center for HCI and
Dept of Computer Science
Virginia Tech
Blacksburg, VA 24061 USA
msampat@vt.edu

Andrea Kavanaugh
Center for HCI and
Dept of Computer Science
Virginia Tech
Blacksburg, VA 24061 USA
kavan@cs.vt.edu

D. Scott McCrickard
Center for HCI and
Dept of Computer Science
Virginia Tech
Blacksburg, VA 24061 USA
mccricks@cs.vt.edu

Manuel A. Pérez-Quiñones
Center for HCI and
Dept of Computer Science
Virginia Tech
Blacksburg, VA 24061 USA
perez@cs.vt.edu

Joe Gabbard
Center for HCI and
Dept of Computer Science
Virginia Tech
Blacksburg, VA 24061 USA
jgabbard@cs.vt.edu

Jason Chong Lee
Center for HCI and
Dept of Computer Science
Virginia Tech
Blacksburg, VA 24061 USA
chonglee@cs.vt.edu

CHI 2007, April 28 – May 3, 2007, San Jose, USA
ACM 1-xxxxxxxxxxxxxxxxx.

Abstract
As spatial awareness technologies grow increasingly accurate and ubiquitous, the need for appropriate human-computer interfaces emerges. This paper summarizes our ongoing work in balancing the attention demanded from users with the utility provided through added reaction and comprehension. In so doing, we describe and discuss three techniques for mobile spatial interaction: "images first" presentation, mobile augmented reality, and audio notifications.

Keywords
Mobile interfaces, notifications, augmented reality

ACM Classification Keywords
H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

Introduction
Commercial off-the-shelf spatial awareness technologies, pervasive wireless networks, and cutting edge access tools are propelling us toward a new type of convergence. The convergence of information in cyberspace with information about a user's physical space increasingly attributes place of creation, access, manipulation, and relevance/scope as part of the metadata for our personal and public information. Take for example the camera that tags pictures with GPS data upon creation, allowing us to share experiences with other tourists, create mashups, and publish this

information in many novel ways. In-car navigation systems provide the user with location relevant information—in essence performing true local search—allowing us to find points of interest that are within a spatial scope of reference.

This convergence demands a refashioning of our understanding and use of the internet, where we consider abandoning the traditional link-to-link click-through paradigm of gaining access via tethered & semi-mobile portals to a model where information is made available in a context-appropriate form on pervasive computing devices: a geo-spatial web provides a layer of information, representations, and annotations prevalent over and above the physical reality we currently inhabit—naturalizing access to locally relevant information that exists in cyberspace, and creating frameworks for the better organization and utilization of these technological capabilities.

However, in creating such a future lie several challenges: identifying appropriate metaphors for interaction/manipulation, information modality, authoring of information relevant to this paradigm, and many more. An important consideration while designing mobile spatial interaction systems is the dual-task nature of these scenarios. Users will use these systems while performing varied primary tasks; for example returning to a somewhat (but not completely) familiar home town after an extended absence and exploring how things have changed. How do we design systems that appropriately manage the user's attention on the primary task while performing the secondary tasks of alerting and informing? How do we mix images and information about old and new elements of a place?

This paper describes our basic approach—rooted in the attention-utility theme as introduced in [5]—and explores three information presentation solutions to address these problems.

Affording Spatial Interaction

We live in a three-dimensional world, and with years of experience are well trained at interacting with physical space [6]. Spatial metaphors have been long used by interaction designers to familiarize users to virtual environments [3]. In refashioning our interactions with physical space to afford manipulating/visualizing the locally-prevalent cyberspace we are beginning to enter uncharted territory. Will map-based techniques lead the way to help us augment users to this paradigm? Systems like Google Earth and Google Maps have certainly taken the first steps toward this by using the metaphor of push-pins on a canvas style map. However, these are still primarily desktop-based two-dimensional interfaces that may not lend themselves to all the demand of on-the-go use—especially in situations where location is not the primary consideration.

We believe touch, proximity, and multi-modality are the primary affordances that systems targeted toward this paradigm need to enable. Touch allows the user to manipulate his environment by performing physical gestures: similar to the manner in which the mouse lets one manipulate their computer's desktop screen. In this manner users can interact with the virtual representations and annotations present at that place. Proximity on the other hand is the technological capability of determining the near and immediate virtual environment. It is a mix of hardware and software that is needed to present information that is

locally relevant and truly associated with that space. And last but equally important is the multi-modality of the interfaces. In the real world, a human being interacts with this physical environment through voice, visual, and olfactory senses. For the metaphors to be natural and easy to learn: voice and visual forms of interacting must be inherent.

Evaluating the effectiveness of these interfaces will be quite challenging. As described earlier, these systems will be used in dual-task situations, where the purpose of technology is to augment the primary task of the user in his surroundings. Attention-utility tradeoffs—specifically examined in our work are interruption, reaction, and comprehension [4,5]—need to be managed in creating an appropriate balance for affording such interaction. For example, an in-car navigation system might be best tailored for low interruption (to avoid unnecessary distraction of the driving task) and low comprehension (if the driver does not need to understand the geographic layout of the area) but high reaction (to aid in navigating).

Figure 1 User's perspective while using handheld AR to read location-aware augmented information overlaid to the left of a piece of artwork.

As shown, we were able to augment the artist name, the title of the piece, and the medium onto the wall, juxtaposed to the left of the piece of art.

Interaction Techniques

From our experiences working on mobile location-aware user interfaces, we know that traditional forms of user interaction are often too cumbersome or time consuming. Under the label of the SeeVT architecture—which provides location information based on wireless signal strength—we have built several mobile location-aware interfaces for guiding and informing users (e.g., [2,7,8]). This section provides an overview of three techniques we have encountered that seem to be particularly effective for mobile, multi-task situations.

One of our more interesting findings to date is that users often prefer an "images first" information technique over traditional map-based interfaces (as seen in [8]). In this case, users were often familiar with an area geographically, but they were not aware of the full history and events associated with the area. As such, users were not interested in interfaces that support wayfinding, but instead are interested in visual information that is tied to location and context. User goals were not to be interrupted with directional guidance (as in a wayfinding task), but rather to enhance comprehension of key elements in the environment. This is a case where the interruption and reaction aspects are relatively low, but the approach must afford fairly high levels of comprehension.

A second promising technique (building on the "images first" technique) is the use of mobile augmented reality: the presentation of an image close to the user's current location that is enhanced with information about the location. Recently we have been implementing a framework for handheld tracking that allows users to move freely within indoor spaces—un-cumbered by the wires and obtrusive head-mounted displays typical in augmented reality systems. Our ongoing research will enable users to interact with location-based content, receive geo-tagged notifications and interact with other user/friends based on location. Our first prototype of a handheld augmented reality system provides context-aware information about pieces of artwork, shown in Figure 1. Such interfaces convey very high levels of comprehension but also require significant attention to read information. Our ongoing work is seeking to find an appropriate balance for balancing the amount of information presented and the time required to read and understand it. A nested model of information presentation—in which

progressively more information is presented with a motion, click, or gesture—may allow users to appropriately manage their interruption levels.

Enhancing the presentation of image-based information, audio notifications can convey information directly, without the intermediate and potentially taxing stage of interpretation involved in non-auditory information [1]. Voice notification systems that focus on optimal allocation of attention to maximize system utility are of specific interest. Our prior work shows that audio notifications can be used in dual task situations to efficiently deliver information [2]. Notifications about points-of-interest, or items that users have subscribed to can be delivered through audio-on-the-go interfaces while achieving high comprehension and high reaction, though at the potential cost of higher interruption than a purely visual interface technique—both for the user and for surrounding people. For example, students could subscribe to podcasts about happenings at places of interest and have them delivered the next time they are near the place. Wayfinding interfaces would fall into this category as well.

Conclusions

The three presentation techniques described here—"images first", mobile augmented reality, and audio notifications—show promise in addressing the growing need to interact with information on-the-go. Each represents something of a departure from traditional desktop interfaces, and each shows promise in balancing the attention required to interact with the device with the utility gained from additional information.

References

[1] Alexanderson, P. (2004). Peripheral Awareness and Smooth Notification : the Use of Natural Sounds in Process Control Work. In *Proc. NordiCHI*, 281-284.

[2] Bhatia, S. and McCrickard, D. S. (2006). "Listening to your inner voices: Investigating means for voice notifications." In *Proc. CHI*, 1173-1176.

[3] Gaver, B. (1992). "*The Affordances of Media Spaces for Collaboration*", In *Proc. CSCW*, 17-24.

[4] McCrickard, D. S. and Chewar, C. M. (2003). "Attuning Notification Design to User Goals and Attention Costs." *Communications of the ACM, 46* (3), 67-72.

[5] McCrickard, D. S., Chewar, C. M., Somervell, J. P., and Ndiwalana, A. (2003). "A Model for Notification Systems Evaluation—Assessing User Goals for Multitasking Activity." *ACM TOCHI, 10*, 4, 312-338.

[6] Harrison, S. and Dourish, P. (1996). Re-place-ing space: the roles of place and space in collaborative systems. In *Proc.CSCW*, 67-76.

[7] Sampat, M., Kumar, A., Prakash, A. and McCrickard, D. S. (2005). "Increasing Understanding of a New Environment using Location-Based Notification Systems." Poster paper in *Proc HCII,* 5 pages (CD-ROM).

[8] Nair, S., Kumar, A., Sampat, M., Lee, J. C., and McCrickard, D. S. (2006). "Alumni Campus Tour: Capturing the Fourth Dimension in Location Based Notification Systems." In *Proc. ACMSE,* 500-505.