Heads-up Engagement with the Real World: Multimodal Techniques for Bridging the Physical-Digital Divide

Simon Robinson
Future Interaction Technology Lab
Swansea University
SA2 8PP, UK
s.n.w.robinson@swansea.ac.uk

CHI 2010, April 10–15, 2010, Atlanta, Georgia, USA.
ACM 978-1-60558-930-5/10/04.

Abstract
The vast and ever-increasing collection of geo-tagged digital content about the physical world around us has prompted the development of interaction methods for various different scenarios. However, the map-based views common on desktop computers are not always appropriate when considering mobile usage. The aim of this research is to provide suitable methods that can encourage user interaction with geo-located digital content, avoiding unnecessary interference with the user's immersion in the physical world around them. This extended abstract outlines the work published to date, suggests future areas of research, and highlights the key contributions brought to the HCI community.

Keywords
Mobile, multimodal, heads-up, engaging, interaction, location-aware, geo-web, haptics, sensors.

ACM Classification Keywords
H.5.2 User Interfaces: Interaction Styles, Prototyping; H.5.1 Multimedia Information Systems: Artificial, augmented and virtual realities.

General Terms
Design, Experimentation, Human Factors.

Introduction

Situated, geo-tagged content about the places we visit has become increasingly common in recent years with the proliferation of online mapping and low-cost location-based systems. Spontaneous photos can be automatically location tagged, then uploaded to websites such as *Flickr* or *Panoramio* to view amongst people's memories of different experiences, while others can add their thoughts, or perhaps just a record of what they're doing at that moment. Building upon these technologies, an array of services has emerged to promote and facilitate the sharing of all types of digital content associated with real-world locations.

Interactions with geolocated content from services like these work well on the large screens of conventional computers, but lack elements of usability in mobile scenarios. For mobile devices it is often tempting to simply shrink the interface and provide the same interaction on a smaller screen. Crucially, though, implementations of this type can require the user to transfer attention from their physical surroundings to a digital representation of the location, and concentrate on interpreting the links between the two. The work presented here is motivated by a desire to remove the process of grappling with the digital world; instead, our aim is to provide engaging, 'heads-up' ways for users to discover digital content associated with the real-world places they visit.

As this extended abstract will demonstrate, novel ways of engaging the user in their real-world surroundings are recognised as an important focus for current HCI research. Our innovative approaches move away from traditional augmented reality work, helping people remain immersed in their environment. Evidence of the importance of this research has been published in a journal and in the proceedings of several conferences, the most recent of which was awarded *best full paper*.

Heads-up interaction

In this work, we consider heads-up to mean a fluid interaction between the user, their mobile device and the world around. There are clear links between this and previous research into mobile or wearable eyes-free interfaces (e.g. [1]). However, rather than completely remove the need for a visual interface, our aim is to provide an exploratory, negotiated interaction flow between the user and their environment. Not all stages of the interaction need be eyes-free, but the natural immersion of the user in their environment should be supported rather than interrupted.

We consider pointing-based interaction to be an ideal example of this – the user can maintain their link with the environment around them, but use their mobile device to probe, discover and take advantage of interesting content in their digitally-enriched surroundings. Previous work has considered ways of enabling a more natural interaction with the physical world via mobile pointing-based interfaces (e.g. [9]), providing 'point-to-select' interaction with real objects. However, these methods have often required a detailed three-dimensional model of the user's environment, or a visual interface for confirmation of the user's desired target. Our approach uses currently-available geolocated content, complementing the user's view of the physical world.

The goal of this project is to study different modalities and multimodal interaction techniques to provide and support these engaging interaction styles. The ever-

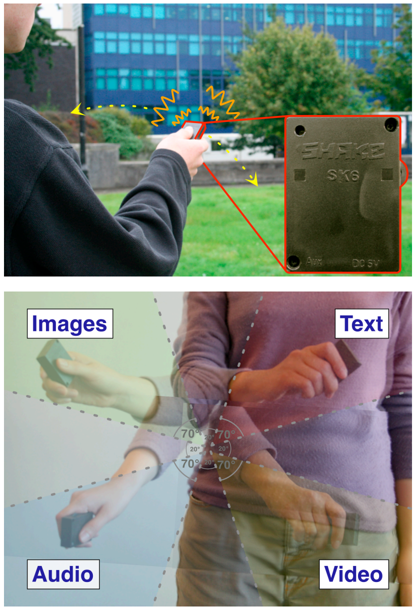

Figure 1: *Top:* As the user scans their environment, the mobile device vibrates when they point toward a location with geotagged content. *Bottom:* When a potential point of interest is found, pointing to each of the clusters triggers feedback if there is content available.

decreasing cost of the mobile sensors used for this work has made possible the creation of realistic interfaces and interaction concepts that are viable for today's devices, bringing key contributions to the mobile interaction domain. Evaluation and publication of our prototypes has, and will continue to demonstrate the development of innovative mobile interaction techniques, and this will be brought together into the PhD thesis at the end of this research.

Approach

Each part of this project will explore methods for supporting engaged interaction with physical environments. Work to date has considered possibilities for providing feedback, both real-time and after the event; these are detailed in the following sections.

Non-visual, mobile information seeking

Our initial work focused on *geoblogging* while on the move, supporting the creation of a rich record of people's journeys taken and places encountered [6]. In a similar approach to that of *LifeTag* [5], participants' positions were logged at all times while on the move. However, our mobile prototype also allowed selective tagging of particularly interesting places. Participants used the device to point at anything in their environment, tilting to refine the distance of their selection and record the gesture.

We designed the device to offer a delayed interaction interface [4]; at the time of the interaction users were given no feedback as to what they had selected, or what sort of information was available. Later the device was connected to a desktop computer and users were able to retrace their journey, browsing automatically-retrieved information about the places they had marked. Comparing this to a visual interface in a field study, we found that casual pointing and selecting was seen as an attractive way to browse and select real-world locations, and offered advantages over the visual alternative. However, some difficulties arose due to the lack of feedback, leading to further work augmenting point-to-select interaction with haptics.

Haptic browsing and filtering

Our more recent work moved to real-time information discovery using vibrotactile feedback [8]. Rather than pointing and tilting, the user scanned their environment with a mobile device to feel for the location of possible points of interest (see Fig. 1, top); as the user moved their focus the device vibrated when interesting content was targeted. Similar to work on audio or haptic-based navigation [3, 10], our novel prototype provided handheld vibrotactile feedback. Building on our initial work, the device allowed the user to interact with digital information in their nearby environment without the need to concentrate on a screen during the discovery process. Extending this browsing interaction, we allowed users to filter retrieved content by type with simple directional gestures (see Fig. 1, bottom). Results from a user study were encouraging, and showed that, despite the low-resolution of the haptic feedback, users were able to successfully discover and filter information from points of interest in their surroundings.

Ongoing research

Our work to-date has considered heads-up interaction with sensor-based, non-visual techniques. Further work is possible to refine these early prototypes, such as improving gesture recognition or clarifying the haptic feedback provided. In addition, the effect of the system on aspects of the user's behaviour can also be quantified (we have studied this in [7]).

Further research could investigate the modalities that best support heads-up interaction. Audio feedback is one area in particular that could provide an engaging user experience with low impact on users' natural behaviour. In addition, previous research has found that combining audio with haptic feedback significantly improves the user experience [2], a clear extension of our existing work.

Until now we have studied visual systems as an example of devices that do not support heads-up interaction, but the predicted ubiquity of handheld projectors promises to change this view. Previous work highlights the huge range of potential interactions with augmented reality displays; mobile projectors could offer these views in the world around us.

Conclusions

The aim of this project is to provide methods for supporting engaging heads-up interaction with the physical world. We have considered the use of delayed interaction and haptic feedback to support place browsing, finding benefits to users in both cases. Future work will investigate different heads-up modalities and the use of multimodal interfaces to provide engaging user experiences. In summary, this work will contribute in the following key areas:

1. Development of interaction styles to support users' natural engagement with the physical world.
2. Study of feedback modalities for discovering and browsing geotagged place information.

Acknowledgements

This work is funded by EPSRC project EP/E042171/1, undertaken in collaboration with colleagues at Glasgow University. Thanks especially to Matt Jones and Parisa Eslambolchilar for their support and feedback.

References

[1] Brewster, S., Lumsden, J., Bell, M., Hall, M., Tasker, S. Multimodal 'eyes-free' interaction techniques for wearable devices. In *Proc. CHI 2003*, 473–480.

[2] Chang, A., O'Sullivan, C. Audio-haptic feedback in mobile phones. In *Proc. CHI '05*, 1264–1267.

[3] Holland, S., Morse, D. R., Gedenryd, H. AudioGPS: spatial audio navigation with a minimal attention interface. *PUC*, 6(4): 253–259, 2002.

[4] Jones, M., Buchanan, G., Cheng, T-C., Jain, P. Changing the pace of search: supporting "background" information seeking. *JASIST*, 57(6): 838–842, 2006.

[5] Rekimoto, J., Miyaki, T., Ishizawa, T. LifeTag: WiFi-based continuous location logging for life pattern analysis, In *Proc. LOCA '07*, 35–49.

[6] Robinson, S., Eslambolchilar, P., Jones, M. Point-to-GeoBlog: gestures and sensors to support user generated content creation. In *Proc. MobileHCI '08*, 197–206.

[7] Robinson, S., Eslambolchilar, P., Jones, M. Evaluating haptics for information discovery while walking. In *Proc. HCI '09*, 93–102.

[8] Robinson, S., Eslambolchilar, P., Jones, M. Sweep-Shake: finding digital resources in physical environments. In *Proc. MobileHCI '09*, 85–94.

[9] Simon, R., Kunczier, H., Anegg, H. Towards orientation-aware location based mobile services. In *Proc. LBS '05*, 279–290.

[10] Strachan, S., Williamson, J., Murray-Smith, R. Show me the way to Monte Carlo: density-based trajectory navigation. In *Proc. CHI '07*, 1245–1248.