Mixed reality (MR) interfaces for mobile information systems

Mountain, D. and Liarokapis, F.

Mountain, D., Liarokapis, F., Mixed reality (MR) interfaces for mobile information systems, Aslib Proceedings, Special issue: UK library & information schools, Emerald Press, 59(4/5): 422-436, (2007). ISSN: 0001-253X.

Mixed reality (MR) interfaces for mobile information systems

David Mountain
giCentre, Department of Information Science, City University London, London, UK, and
Fotis Liarokapis
Coventry University, Coventry, UK and Department of Information Science, City University London, London, UK

Abstract

Purpose - The motivation for this research is the emergence of mobile information systems where information is disseminated to mobile individuals via handheld devices. A key distinction between mobile and desktop computing is the significance of the relationship between the spatial location of an individual and the spatial location associated with information accessed by that individual. Given a set of spatially referenced documents retrieved from a mobile information system, this set can be presented using alternative interfaces of which two presently dominate: textual lists and graphical two-dimensional maps. The purpose of this paper is to explore how mixed reality interfaces can be used for the presentation of information on mobile devices.
Methodology/approach - A review of relevant literature is followed by a proposed classification of four alternative interfaces. Each interface is the result of a rapid prototyping approach to software development. Some brief evaluation is described, based upon thinking aloud and cognitive walk through techniques with expert users.
Findings - The most suitable interface for mobile information systems is likely to be user and task dependent, however, mixed reality interfaces offer promise in allowing mobile users to make associations between spatially referenced information and the physical world.
Research limitations/implications - Evaluation of these interfaces is limited to a small number of expert evaluators, and does not include a full scale evaluation with a large number of end users.
Originality/value of paper - The application of mixed reality interfaces to the task of displaying spatially referenced information for mobile individuals.
Keywords Mixed reality, Virtual reality, Augmented reality, Mobile computing, Mobile information systems, Geographic information.
Paper type: Research paper.

1. Introduction

Two of the most significant technological trends of the past 15 years have been the increased portability of computer hardware - such as laptop computers and personal digital assistants (PDAs) - and the increasing availability of wireless networks such as mobile telecommunications, and more recently wireless access points (Brimicombe and Li, 2006). The convergence of these technological drivers presents opportunities within the emerging field of mobile computing. Increasingly there is ubiquitous access to information stored via a variety of media (for example, text, audio, image and video) via mobile devices with wireless network connections. Advances in software development tools for mobile devices have resulted in the implementation of user-friendly interfaces that aim to appeal to a wide audience of end users. A key challenge for researchers of mobile information systems is to decide the type of interface to adopt when presenting this information on mobile devices. Additionally, developers should assess whether the most suitable interface is dependent upon the audience, the task-in-hand and geographic context in which the mobile information system is likely to be used (Jiang and Yao, 2006).

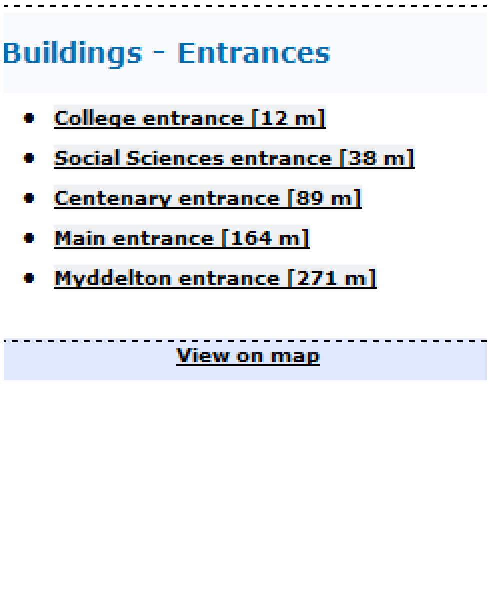
The LOCUS project (LOcation Context tools for UMTS Services) being conducted within the Department of Information Science at City University is addressing some of the research challenges described above (LOCUS, 2007). The main aim of the project is to enhance the effectiveness of location-based services (LBS) in urban environments by investigating how mixed reality interfaces compare with the current map- and text-based approaches used by the majority of location-based services for the tasks of navigation and wayfinding (Mountain and Liarokapis, 2005). To satisfy this aim, LOCUS is tackling a number of issues including the three-dimensional representation of urban environments, the presentation of spatially referenced information - such as the information retrieved as the result of a user query, and navigational information to specific locations - and advanced visualisation and interaction techniques (Liarokapis *et al.*, 2006).

The LOCUS system is built on top of the WebPark mobile client-server architecture (WebPark, 2006) which provides the basic functionality associated with LBS including the retrieval of information based upon spatial and semantic criteria, and the presentation of this information as a list or on a map (see Figures 1a and 1b). In common with the majority of LBS, the basic architecture provides no mechanism for the display of information in a three-dimensional environment, such as a mixed reality interface.

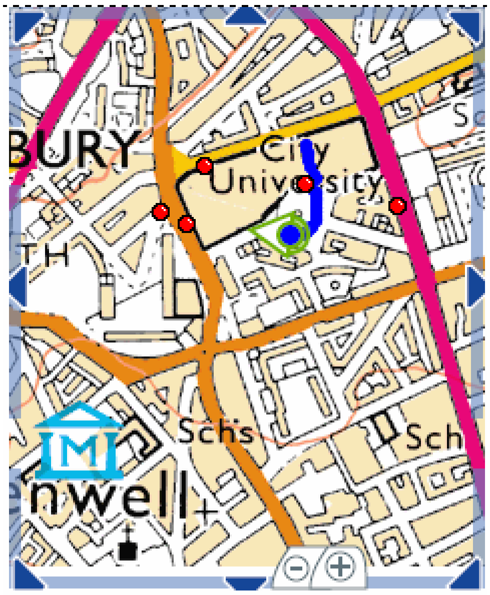
Mixed reality environments occupy a spectrum between entirely real environments at one extreme and entirely virtual environments on the other. This mixing of the real and the virtual domain offers great potential in terms of displaying information retrieved as a result of a location-based search, since this requires the presentation of digital information relative to your location in the physical world. This presentation may on the one hand be entirely synthetic, for example, placing virtual objects representing individual results within a virtual scene as a backdrop. Alternatively an augmented reality interface can superimpose this information over the real world scene in the appropriate spatial location from the mobile user’s perspective. Both interfaces can present the location of information within the scene as well as navigation tools that describe the routes to the spatial locations associated with retrieved information. The LOCUS project is extending the functionality of the WebPark architecture to allow the presentation of spatially referenced information via these mixed reality interfaces on mobile devices (see Figures 1c and 1d).

The rest of the paper is structured as follows. First, a review of relevant background literature in mobile computing and mixed reality is presented. Next, candidate interfaces for mobile information provision into mobile devices are suggested: these include the *list*, the *map*, and *virtual* and *augmented reality* interfaces. The paper closes with a discussion and conclusions.

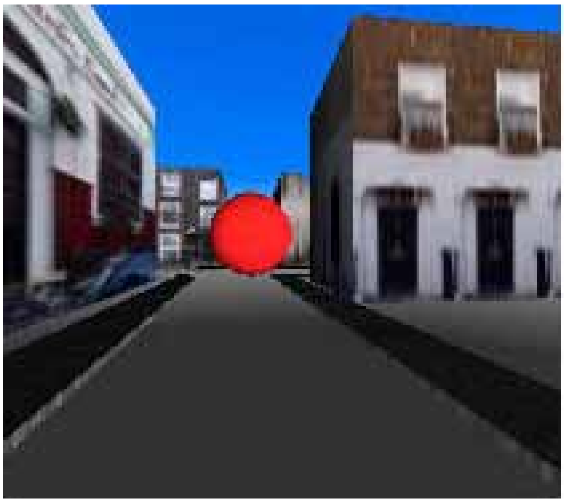
Aslib Proceedings, Special issue: UK library & information schools, Emerald Press, 59(4/5): 422-436, (2007). ISSN: 0001-253X



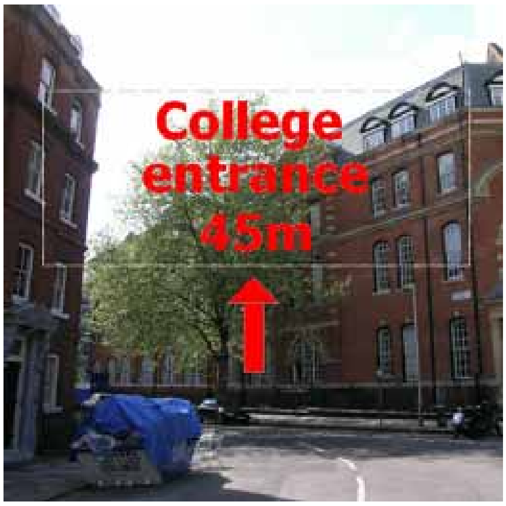
(a) List interface (WebPark platform)



(b) Map interface (WebPark platform)



(c) Mobile VR interface (LOCUS prototype application)



(d) Mobile AR interface (LOCUS prototype application)

Figure 1 Interfaces for presenting information retrieved from a mobile information system

2. Background

2.1 Mobile computing

Just as the evolution of the Internet has had a profound impact upon application development, forcing a change from a stand-alone desktop architecture to a more flexible, client-server architecture (Peng and Tsou, 2003), researchers in mobile computing are currently having a similar impact, forcing the development of web resources and applications that can be run on a wider range of devices than traditional desktop machines. According to Peng and Tsou (2003), mobile computing environments have three defining characteristics:

1. mobile clients that have limited processing and display capacity (e.g., PDAs and smart phones);
2. non-stationary users who may use their devices whilst on the move;
3. wireless connections that are often more volatile, and have more constrained bandwidth, compared to the "fixed" Internet.

These three characteristics suggest that mobile devices have both specific constraints and unique opportunities when compared to their desktop counterparts. First, screen real estate is limited; typically screens are small (usually less than 60mm by 80mm) with low resolution (typically 240 pixels width), and a relatively large proportion of this space may be taken up with *marginalia* such as scroll bars and menus, hence every pixel should be used wisely. Next, the outdoor environment is a more unpredictable and dynamic environment than the typically familiar indoor home and office environments in which desktop machines are used; hence user attention is more likely to be distracted in the mobile context. Mobile computer usage tends to be characterised by multiple short sessions per day, compared with desktop usage which tends to be for relatively few, longer durations (Ostrem, 2002). Given these constraints, there is a clear need for information to be communicated concisely and effectively for mobile users.

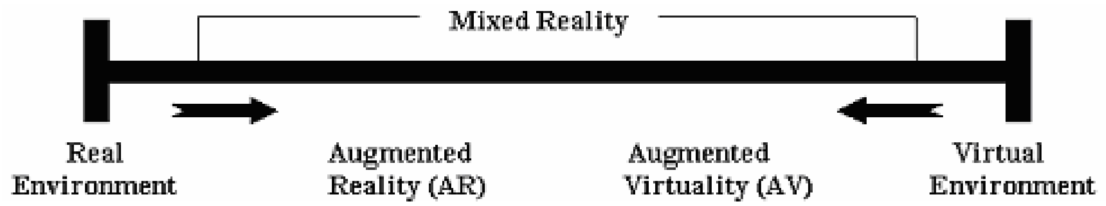
Despite constraints, the mobile computing environment offers a unique opportunity for the presentation of information, in particular taking advantage of location sensors to organise information relative to the device user's position, or their spatial behaviour (Mountain and MacFarlane, in press).

Whilst spatial proximity is perhaps the most intuitive and easily calculated measure of geographic relevance, it may not be the most appropriate in all situations and a variety of other measures of geographic relevance (Mountain and MacFarlane, in press; Raper, 2001) have been suggested. Individuals may be more interested in the relative accessibility of results, which can be quantified by travel time and can take account for natural and manmade boundaries (Golledge and Stimson, 1997) or the transportation network, to discount results which are relatively inaccessible despite being physically close (Mountain, 2005). Geographic relevance can also be quantified as the results are most likely to be visited in the future (Brimicombe and Li, 2006), or those which are most visible from the current location (Kray and Kortuem, 2004). However geographic relevance is quantified, there are opportunities to use this property to retrieve documents from document collections. Given a set of spatially referenced results that are deemed to be geographically relevant according to some criterion, there are a variety of different approaches to presenting this information.

Various mobile information systems have been developed. Kirste (1995) developed one of the first experimental mobile information systems based on wireless data communication. A few years later, Afonso *et al*. (1998) presented an adaptable framework for mobile computing information dissemination systems called UbiData. This model adopts a "push" model where relevant information is sent to the user, without them making a specific request, based upon their location. There are now a host of commercial and prototype mobile information systems that can present information dependent upon an individual's semantic and geographic criteria (Yell Group, 2006; WebPark, 2006), the majority of which present results either as a list or over a backdrop map.

2.2 Mixed reality

The mixed reality spectrum was proposed by Milgram and Kishino (1994) who depicted representations on a continuum with the tangible, physical ("real") world at one extreme and entirely synthetic Virtual Reality (VR) at the other. Two classes were identified between these extremes. Augmented Reality (AR) refers to virtual information placed within the context of the real world scene, for example, virtual chess pieces on a real chess board. The second case – augmented virtuality – refers to physical information being placed in a virtual scene, for example, real chess pieces on a virtual board. The resulting Reality-Virtuality continuum is shown in Figure 2.



Source: After Milgram *et al.* (1994)

Figure 2 The reality - virtuality continuum.

The first Virtual Reality (VR) system was introduced in the 1950s (Rheingold, 1991) and since then VR interfaces have taken two approaches: immersive *head-mounted displays* and *through the window* approaches. Head-mounted displays (HMDs) are very effective at blocking the signals from the real world and replacing this natural sensory information with digital information. Navigation within the scene can be controlled by mounting orientation sensors on top of the HMD, a form of gesture computing whereby the user physically turning their head results in a rotation of the viewpoint in the virtual scene. The ergonomic limitations of HMDs proved unpopular with users and this immersive interface has failed to be taken up on a wide scale (Ghadirian and Bishop, 2002). In contrast to HMDs, the through the window (Bodum, 2005) – or monitor-based VR/AR - approach exploits monitors on desktop machines to visualise the virtual scene, a far less immersive approach since the user is not physically cut-off from the physical world around them. This simplistic form of visualisation has the advantage that it is cost-effective (Azuma, 1997). Interaction is usually realised via standard input/output (I/O) devices such as the mouse or the keyboard but also more sophisticated devices (such as spacemouse, inertia cube, etc.) may be employed (Liarokapis, 2005).

Both HMDs and through the window approaches of VR aim to replace the physical world with the virtual. The distinction of *augmented* reality (AR) is that it aims to seamlessly combine real and virtual information (Tamura, and Katayama, 1999) by superimposing digital information directly into a user's sensory perception (Feiner, 2002) (see Figure 3). Whilst VR and AR can process and display similar information (for example three-dimensional buildings) the combination of the "real" and the "virtual" in the AR case is inherently more complex than the closed virtual worlds of VR systems.

Figure 3 Augmented reality representation: a computer vision sensor recognises the doorway outline, and augments the video stream with virtual information (the direction arrow). Developed as part of the LOCUS project

This combination of real and virtual requires accurate tracking of the location of the user (in three spatial dimensions: x, y and z) and the orientation of their view (around three axes of orientation: yaw, pitch and roll), in order to be able to superimpose digital information at the appropriate location with respect to the real world scene, a procedure known as *registration*. In the past few years, research has achieved great advances in tracking, display and interaction technologies, which can improve the effectiveness of AR systems (Liarokapis, 2005). The required accuracy of the AR tracking depends to a degree upon the scenario of use. In order to correctly superimpose an alternative building façade (for example, a historic or planned building façade) over an existing building, highly accurate tracking is required in terms of position and orientation, else the illusion will fail since the real and virtual facades will not align, or may drift apart as the user moves or turns their head (Hallaway *et al*., 2004). However, if simply augmenting the real world scene with annotations in the forms of text or symbols, for example, an arrow indicating the direction to turn at an upcoming junction, this tracking may not be required to be so accurate.

The two most common tracking techniques used in AR applications include *computer vision* and *external sensor systems*. The visual approach uses *fiducial* reference points where a specific number of locations act as links between the real and virtual scenes (Hallaway *et al*., 2004). These locations are usually marked with distinctive high contrast markers to assist identification, but alternatively can be distinctive landmarks within the real world scene. Computer vision algorithms first need to identify at least three reference points in real-time from a video camera input, then calculate the distance and orientation of the camera with respect to those reference points. Tracking using a computer vision-based system therefore establishes a *relative* spatial relationship between a finite number of locations in the real world scene and the observer, via a video camera carried or worn by that observer (Hallaway *et al*., 2004), which can allow very accurate registration between the real and virtual scenes in a well-lit indoor environment. This computer vision approach nevertheless has significant constraints. First, the system must be trained to identify these fiducial reference points, and may further require the real world scene to have markers placed within it. It requires both good lighting conditions (although infrared-cameras can be also used for night vision) and significant computing resources to perform real time tracking and therefore has usually been conducted in an indoor, desktop environment (Liarokapis and Brujic-Okretic, 2006).

An alternative to the vision based approach is to use external sensors to determine the position of the user and the orientation of their view. Positioning sensors such as the Global Positioning System (GPS) can determine position in three dimensions and digital compasses, gyroscopes and accelerometers can be employed to determine the orientation of the user's view. These sensor-based approaches have the advantage that they are not constrained to specific locations, unlike computer vision algorithms which must be trained to recognise specific reference points within a scene. Also the user's location is known with respect to an external spatial referencing system, rather than establishing relative relationships between the user and specific reference points. A major disadvantage is the accuracy of the positioning systems, which can produce errors measured in tens of metres and can produce poor results when attempting to augment the real world scene with virtual information. Whilst advances in GPS systems such as differential GPS and real-time kinematic GPS can bring down the accuracy to one meter and a few centimetres respectively, GPS receivers still struggle to attain a positional fix where there is no clear view of the sky, for example, in doors. Digital compasses also have limitations; the main flaw is that they are prone to environmental factors such as the magnetic fields.

Having identified a spatial relationship between the real world scene and the user location, virtual information needs to somehow be superimposed upon the real world scene. Traditionally there have been two approaches to achieving this: *video see-through displays* and *optical see-through displays*. Video see-through displays are comprised of a graphics system, a video camera, a monitor and a video combiner (Azuma, 1997). They operate by combining a HMD with a video camera. The video camera records the real environment and then sends the recorded video to the graphics system for processing. There the outputted video and the generated graphics images, by the graphics system, are blended together. Finally, the user perceives the augmented view in the closed-view display system. Using the alternative approach, optical see-through displays are usually comprised of a graphics system, a monitor and an optical combiner (Azuma, 1997). They work by simply placing the optical combiners in front of the user's view. The main characteristic of the optical combiners is that they are partially transmissive and reflective. That is because the combiners operate like half-silvered mirrors permitting only a portion of the light to penetrate. As a result, the intensity of the light which the user finally sees is reduced.

A novel approach to augmenting the real world scene with virtual information, emerging from within the field of mobile computing, is to use the screen of a handheld device to act as a *virtual window* on the physical world. Knowing the position and orientation of the device, the information displayed on screen can respond to movements and gestures of a mobile individual, for example, presenting the name of a building as text on the screen when a user points their mobile device at it, or updating navigational instructions via symbols or text as a user traverses a route.

MARS is one of the first outdoor AR systems and a characteristic example of a wireless mobile interface system for indoor and outdoor applications. MARS was developed to aid navigation and to deliver location-based information to tourists in a city (Höllerer *et al.*, 1999). The user stands in an outdoor environment wearing a prototype system consisting of a computer, a GPS system, a see-through head-worn display and a stylus-operated computer. Interaction is via a stylus and display is via a tracked see-through head-worn display. MARS like most current mobile AR systems has significant ergonomic restrictions which stretch the definitive of mobile and wearable computing beyond what is acceptable for most users (the system is driven by a computer contained in a backpack).

Tinmith-Hand AR/VR is a unified interface technology designed to support outdoor mobile AR applications and indoor VR applications (Piekarski, 2002). This system employs various techniques including 3D interaction techniques, modelling techniques, tracked input gloves and a menu control system, in order to build VR/AR applications that can be applied to construct complex models of objects in both indoor and outdoor environments. A location-based application that was designed for a mobile AR system is ARLib and aims to assist the user in typical tasks that are performed within a library environment (Umlauf *et al.*, 2002). The system follows a wide area tracking approach (Hedley *et al*., 2002) based on fiducial-based registration. Many distinct markers are attached to bookshelves and walls so that the book's positions are superimposed on the shelves as the user navigates inside the library. To provide extra support to the user, a simple interface and a search engine are integrated to provide maximum usability and speed during book searches.

3. Interfaces for mobile information systems

There are many candidate interfaces for the presentation of the results of an information retrieval query on mobile devices (Mannings and Pearson, 2003; Schofield and Kubin, 2002; Mountain and Liarokapis, 2005). This section describes how the interfaces described previously can be applied to the task of presenting information retrieved as the result of a mobile query.

As described in the introduction, the LOCUS project has developed alternative, mixed reality interfaces for existing mobile information system technology based upon the WebPark platform. The WebPark platform can assist users in formulating spatially-referenced, mobile queries. The retrieved set of spatially-referenced results can then be displayed using various alternative interfaces: a list, a map, virtual reality or augmented reality. Each interface is described in more detail in the rest of this section.

3.1 List interface

The most familiar interface for the presentation of the results of an information retrieval query is a list; this is the approach taken by the majority of Internet search engines where the most relevant result is placed at the top of the list, with relevance decreasing further down the list (see Figure 1a). In the domain of location-aware computing, results that are deemed to be particularly *geographically* relevant (Mountain and MacFarlane, in press; Raper, 2001) will be presented higher up the list (Google, 2006; WebPark, 2006). Whilst familiar, this approach of simply ordering the results does not convey their location relative to your current position.

3.2 Map interface

The current paradigm in the field of LBS is to present information relevant to an individual's query or task over a backdrop map (see Figure 1b). This information may include the individual's current position (and additionally some representation of the spatial accuracy), the locations of features of interest that were retrieved as the result of a user query (e.g., the results from a "find my nearest" search), or navigation information such as a route to be followed. This graphical approach has the advantage of displaying the direction and distance of results relative to the user's location (a vector value), as opposed to just an ordering results based on distance. The viewpoint is generally allocentric (Klatzky, 1998), adopting a bird's eye view looking straight down on a flat, two-dimensional scene (see Figure 1c). The backdrop contextual map used is usually an abstract representation and may choose to display terrain, points or regions of interest, transportation links, or other information, alternatively a degree of realism can be included by using aerial photography (WebPark, 2006; Google, 2006)

3.3 Virtual Reality (VR) interface

An alternative to the allocentric viewpoint of a two-dimensional, abstract scene is to choose an egocentric viewpoint within a three-dimensional scene (see Figure 4). Such a perspective is familiar from VR discussed in section 2.2. Whilst the concept of VR has existed for many decades, only during the past few years has it been used on handheld mobile devices. Traditionally, VR applications have been deployed on desktop devices and have attempted to create realistic looking models of environments to promote a feeling of immersion within a virtual scene. This has resulted in less opportunity for individuals to compare the virtual scene with its real-world counterpart. This separation of the real and the virtual is due in part to the static nature of desktop devices, and in addition that the appeal of many virtual scenes is that they allow the viewing of locations that cannot be visited easily, for example, virtual fly-throughs on other planets (NASA Jet Propulsion Laboratory, 2006) and imagined landscapes (Elf world, 2006).

(a) Egocentric perspective

(b) An oblique, allocentric perspective

Note: Scene rendered in Google Earth

Figure 4 A virtual representation of a London neighbourhood

In a location-aware, mobile computing context, the position of the user's viewpoint within a VR scene can be controlled from an external location-sensor such as GPS, and the orientation of the viewpoint can be controlled by sensing the direction of movement (from GPS heading), or an orientation sensor to gauge the direction an individual is facing (e.g., a digital compass). The VR scenes themselves can adopt different levels of detail and realism (Bodum, 2005). A particular building may be represented with an exact three-dimensional geometric representation, and graphics added as textures to the façades of the building to create as true a representation as possible – known as a verisimilar representation (Bodum, 2005). Alternatively, the building may be modelled with a generalised approximation of the geometry within specific tolerances. For texturing the building facades, generic images may be applied that are typical of that class of building. The building block can be left untextured, but more abstract information conveying using shading, icons, symbols or text (Bodum, 2005). The level of detail and realism required by different users for different tasks is an open question currently under investigation (Liarokapis and Brujic-Okretic, 2006).

Traditionally, for VR applications deployed in a static, desktop context, there has been greater emphasis placed upon scenes *looking* realistic than ensuring that the content of these scenes is spatially referenced. However, in a mobile context, accurate spatial referencing of VR scenes is required when setting the viewpoint within that scene (using position and orientation sensors) to ensure that the viewpoint in the virtual scene is registered accurately with the user's location in the real world scene. Realism is still important since this can help the user make associations between objects in the virtual scene and those in the real world.

For the applications developed as part of the LOCUS project, within this VR backdrop, additional, non-realistic visual information can be included to augment the scene. Such information can include nodes representing documents retrieved from a spatially referenced document collection (see Figure 1c), or navigational information and instructions (i.e., 3D textual directions). This approach has the advantage of promoting a feeling of immersion, and creating a stronger association between the physical world and relevant geo-referenced information, but is potentially less effective than a map in providing a quick synopsis of larger volumes of information relative to your location. There are opportunities to adopt multiple viewpoints within the VR scene that fall between the extremes of the allocentric-egocentric spectrum, for example, an oblique perspective several metres higher than the user's viewpoint (see Figure 4).

3.4 Augmented Reality (AR) interface

A fourth approach to the display of information in mobile computing is to use the device to merge the real world scene with relevant, spatially-referenced information by using an AR interface: the *virtual window* approach described in section 2.2. Just as for the mobile VR case described above, knowing the location and orientation of the device is an essential requirement for outdoor AR, in order to superimpose information in the correct location.

As described in the literature review, a GPS receiver and digital compass can provide sufficient accuracy for displaying points of interest in the approximate location relative to the user's position. At present, however, these sensor solutions lack the accuracy required for more advanced AR functionality, such as aligning an alternative façade on the front of a building in the real world scene. In the LOCUS system, the handheld mobile device presents text, symbols and annotations in response to the location and orientation of the device. There is no need for a HMD, since the screen on the device can be aligned with the real world scene. On the screen of the device, information can either be overlaid on imagery captured from the device's internal camera, or the screen can display just the virtual information with the user viewing the real world scene directly.

The information displayed is dependent upon the task in hand. When viewing a set of results, as the user pans the device around them, the name and distance of each result is displayed in turn as it coincides with the direction that the user is pointing the device, allowing the user to interrogate the real world scene by gesturing. By adopting an egocentric perspective to combine real and virtual information in this way, users of the system can base their decisions of which location to visit on more quantifiable criteria – such as the distance to a particular result, and the relevance on semantic criteria – but also the more subjective criteria that could never be quantified by an information system. For example, following a mobile search for places to eat conducted at a crossroads, by gesturing with a mobile device, users can see the distance and direction of candidate restaurants, and make an assessment based upon the ambience of the streets upon which different restaurants are located.

Having selected a particular result from the list of candidates, the AR interface can then provide navigational information, in the form of distance and direction annotations (see Figure 1d), to guide the user to the location associated with those results. Although most examples from location-based service suggest "where's my nearest" shop or service, there is no reason that this information could not be the location of breaking news stories from a news website, or spatially referenced html pages providing historical information associated with a particular era or event.

4. Discussion

An evaluation exercise was undertaken to assess appropriate levels of detail, realism and interaction for the mobile virtual reality interface. Whilst there has been extensive evaluation of these requirements in a static desktop context (Dollner, 2005), relatively little attention has been paid to the specific needs of mobile users. In order to gauge these specific requirements, an expert evaluation was conducted. Two common evaluation techniques were applied: *think aloud* and *cognitive walkthrough* (Dix *et al*., 2004). Think aloud is a form of observation that involves participants talking through the actions they are performing, and what they believe to be happening, whilst interacting with a system. The cognitive walkthrough technique was also used where a prototype of a mobile VR application and scenario of use were presented to expert users: evaluating in this way allows fast assessment of early mock-up, hence can influence subsequent development and the suitability of the final application. Both forms of evaluation are appropriate for small numbers of participants testing prototype software and it has been suggested that the majority of usability problems can be discovered from testing in this way (Dix *et al*., 2004).

The expert user testing took place at City University with a total of four users with varied backgrounds: one human-computer interaction expert, one information visualization expert, one information retrieval expert and one geographic information scientist. Each user spent approximately one hour, performing four tasks. The aims of the evaluation of the VR prototype included assessment of the expert user experience with particular focus on:

- the degree of realism required in the scene;
- the required spatial accuracy and level of detail of the building outlines;
- a comparison of 3D virtual scenes with 2D paper maps.

A virtual reality scene was created of the University campus and surrounding area, and viewpoints placed to describe trajectories of movement through the scene. The expert-evaluation process covered two tasks including *mobile search* and *navigation*. The first scenario was in relation to searching for, then locating, specific features. For example, a user searching on a mobile system for entrances to the City University campus from a nearby station. The second scenario was in relation to navigation from one point to another, for example, from the station to the University. Starting and target locations were marked in the 3D maps, and sequences of viewpoints were presented, to mimic movement through the scene.

There was a great deal of variation in terms of the level of photorealism required in the scene, and whether buildings should have image textures placed over the building faces, or whether the building outlines would be sufficient alone. Opinions varied between evaluators and according to the task in hand. Plain, untextured buildings are hard to distinguish from each other and, in contrast, buildings with realistic textures were considered easy to recognise in a micro-scale navigation context (for example, trying to find the entrance to a particular building). However, many evaluators thought that much of this realism would not be required or visible on a small screen device when an overview of the area was required, for example, when considering your present location in relation to information retrieved from a mobile search. Expert users also suggested various departures from the realism traditionally aspired to within the field of virtual reality. These included transparency, to allow users to see through buildings as an aid to navigation, since this will allow you to identify the location of a concealed destination point. Other suggestions included labelling of objects in the scene (for example, building and street names). The inclusion of symbology in the scene to represent points, and routes to those points, was considered to be beneficial to the task of navigation.

In terms of the level-of-detail and spatial accuracy, some users thought that it was not important to have very detailed models of building geometry. Building outlines that are roughly the right size and shape are sufficient, especially when considering an overview of an area, as often required in the mobile search task. For micro-navigation, a higher degree of accuracy may be required.

Virtual 3D scenes were found to have many advantages when compared to paper maps: the most positive feature was found to be the possibility to recognize the features in the surrounding environment, which provides a link between the real and virtual worlds. This removes the need to map-read, which is required when attempting to link your position in the real world with a 2D map, hence the VR interface offers an effective way to gauge your initial position and orientation. A more intangible response was the majority of the users enjoyed interacting with the VR interface more than a 2D map. However, the 3D interface also has significant drawbacks. Some users said that they are so used to using 2D maps that they do not really need a 3D map for navigating, however they thought this attitude may change with the next generation. The size, resolution and contrast of the device screen were also highlighted as potential problems for the VR interface.

5. Conclusions

This paper has presented some insights on how mixed reality interfaces can be used in conjunction with mobile information systems to enhance the user experience. We have explored how the LOCUS project has extended LBS through different interfaces to aid the tasks of urban navigation and wayfinding. In particular, we have described how virtual and augmented reality interfaces can be used in place of text and map based interfaces, which can provide an egocentric perspective to location-based information which is lacking from map and text based representations.

Expert user evaluation has proven to be a useful technique to aid development, and suggests that the most suitable interface is likely to vary according to the user and task in hand. Continued research, development and evaluation is required to provide increasingly intuitive interfaces for location-based services that can allow users to make associations between spatially referenced information retrieved from mobile information systems, and their location in the physical world.

Acknowledgements

The work presented in this paper is conducted within the LOCUS project, funded by EPSRC through the Pinpoint Faraday Partnership. We would also like to thank our partner on the project, GeoInformation Group, Cambridge, for contributing the building geometry and heights data used by the project. We are also grateful to Hulya Guzel for her assistance in the expert user evaluation.

References

Afonso, A.P., Regateiro, F.S. and Silva, M.J. (1998), “UbiData: an adaptable framework for information dissemination to mobile users”, *Object Oriented Technology, ECOOP'98 Workshop on Mobility and Replication*, 1543.

Azuma, R. (1997), “A survey of augmented reality”, *Teleoperators and Virtual Environments*, Vol. 6, No. 4, pp. 355-385.

Bodum, L. (2005), “Modelling virtual environments for geovisualization: a focus on representation”, in MacEachren, A.M., *Exploring Geovisualization*, Elsevier, London, pp. 389-402.

Brimicombe, A. and Li, Y. (2006), “Mobile space-time envelopes for location-based services”, *Transactions in Gis*, Vol. 10 No. 1, pp. 5-23.

Dix, A., Finlay, J.E., Abowd, G.D. and Beale, R. (2004), *Human-Computer Interaction*, Prentice Hall, Harlow.

Dollner, J. (2005), “Geovisualization and real-time compute graphics”, in Dykes, J.A., Kraak, M.J. and MacEachren, A.M. (Eds), *Exploring Geovisualization*, Elsevier, London, pp. 325-344.

Elf world (2006), *Elven Forest*, *3D* available at: http://www.allelves.ru/forest/ (accessed 12 December 2006).

Feiner, S.K. (2002), "Augmented reality: a new way of seeing", *Scientific American*, Vol. 4 No. 24, pp. 48-55.

Ghadirian, P. and Bishop, I.D. (2002), "Composition of augmented reality and GIS to visualise environmental changes", *Proceedings of the Joint AURISA and Institution of Surveyors Conference*, Adelaide, Australia, 25-30 Nov 2002.

Golledge, R.G. and Stimson, R.J. (1997), *Spatial Behaviour: A Geographic Perspective,* The Guildford Press, New York.

Google (2006), *Google local* http://local.google.co.uk/ (accessed 10 December 2006).

Hallaway, D., Hollerer, T. and Feiner, S. (2004), "Bridging the gaps: hybrid tracking for adaptive mobile augmented reality", *Applied Artificial Intelligence*, Vol. 18, No. 6, pp. 477-500.

Hedley, N.R., Billinghurst, M., Postner, L., May, R. and Kato, H. (2002), "Explorations in the use of augmented reality for geographic visualization", *Presence,* Vol. 11 No. 2, pp**.** 119-133.

Höllerer, T., Feiner, S.K., Terauchi, T., Rashid, G. and Hallaway, D. (1999), "Exploring MARS: developing indoor and outdoor user interfaces to a mobile augmented reality system", *Computers and Graphics*, Vol. 23 No. 6, pp**.** 779-785.

Jiang, B. and Yao, X. (2006), "Location-based services and GIS in perspective", *Computers Environment and Urban Systems*, Vol. 30 No. 6, pp. 712-725.

Kirste, T. (1995), "An infrastructure for mobile information systems based on a fragmented object model", *Distributed Systems Engineering Journal,* Vol. 2 No. 3, pp. 161-170.

Klatzky, R.L. (1998), "Allocentric and egocentric spatial representations: definitions, distinctions, and interconnections", in Freksa, C., Habel, C. and Wender, K.F. (Eds), *Spatial cognition - An interdisciplinary approach to representation and processing of spatial knowledge*, Springer-Verlag, Berlin, pp. 1-18.

Kray, C. and Kortuem, G. (2004), "Interactive positioning based on object visibility", in Brewster, S. and Dunlop, M. (Eds) *Mobile human-computer interaction,* Springer, Berlin / Heidelberg, pp. 276-287.

Liarokapis, F. (2005), *Augmented Reality Interfaces - Architectures for Visualising and Interacting with Virtual Information*, DPhil. Department of Informatics, School of Science and Technology, University of Sussex.

Liarokapis, F. and Brujic-Okretic, V. (2006), "Location-based mixed reality for mobile information services", *Advanced Imaging,* Vol. 21 No. 4, pp**.** 22-25.

Liarokapis, F., Mountain, D., Papakonstantinou, S., Brujic-Okretic, V. and Raper, J. (2006), "Mixed reality for exploring urban environments, *1st International Conference on Computer Graphics Theory and Applications***,** Setúbal, Portugal, 25-28 Feb 2006, pp. 208-215.

LOCUS (2007) *Homepage*, available at: http://www.locus.org.uk (accessed 22 Jan 2007*)*

Mannings, R. and Pearson, I. (2003), "`Virtual Air' A novel way to consider and exploit location-based services with augmented reality", *Journal-Communications Network,* Vol. 2 No. 1, pp**.** 29-33.

Milgram, P. and Kishino, F. (1994), "A taxonomy of mixed reality visual displays", *IEICE Transactions on Information Systems E Series D,* Vol. 77, No. 12, pp. 1321-1329.

Milgram P, Takemura H, Utsumi A and Kishino F (1994) “Augmented reality: a class of displays on the reality-virtuality continuum”. *Telemanipulator and telepresence technologies*, Vol. 2351, pp. 282-292.

Mountain, D.M. (2005), *Exploring Mobile Trajectories: An Investigation of Individual Spatial Behaviour and Geographic Filters for Information Retrieval*, PhD Thesis, Department of Information Science, City University London, London.

Mountain, D.M. and Liarokapis, F. (2005), “Interacting with Virtual Reality scenes on mobile devices”, *Human Computer Interaction with Mobile Devices and Services*. University of Salzburg, Austria, 19-22 Sep 2005.

Mountain, D.M. and MacFarlane, A. (in press) “Geographic information retrieval in a mobile environment: evaluating the needs of mobile individuals”, *Journal of Information Science*.

NASA Jet Propulsion Laboratory (2006), *Lander and Rover on Mars*, available at: http://mars.sgi.com/worlds/pathfinder/pathfinder.html (accessed 12 December 2006).

Ostrem, J (2002) *Palm OS user interface guidelines*, available at: http://www.palmos.com/dev/support/docs/ui/UIGuide_Front.html (accessed 10 April 2006).

Peng, Z-R and Tsou, M-H (2003), *Internet GIS: Distributed Geographic Information Services for the Internet and Wireless Networks*, John Wiley and Sons, New York.

Piekarski, WaT, H.B. (2002), *Unifying Augmented Reality and Virtual Reality User Interfaces*, University of South Australia, Adelaide, SA.

Raper, J.F. (in press), “Geographic Relevance”, *Journal of Documentation*.

Rheingold, H.R. (1991), *Virtual Reality*, Summit Books, New York, Tokyo.

Schofield, E. and Kubin, G. (2002), “On interfaces for mobile information retrieval”, *Lecture Notes in Computer Science*, 2411**,** pp. 383-387.

Tamura, H. and Katayama, A. (1999), “Steps toward seamless mixed reality”, in Ohta, Y. and Tamara, H. (Eds), *Mixed Reality Merging Real and Virtual Worlds*, Ohmsha Ltd., pp. 59-84.

Umlauf, E. Piringer, Reitmayr, G. Schmalstieg, D. (2002), ARLib: The Augmented Library. *Proceedings of the First IEEE International Augmented Reality ToolKit Workshop*. Darmstadt, Germany.

WebPark (2006), *Geographically Relevant Information for Mobile Users in Protected Area*, available at: http://www.webparkservices.info (accessed 12 December 2006*)*

Yell Group (2006), *The UK's Local Search Engine*, available at: http://www.yell.com (accessed 12 December 2006).

Received: 15 December 2006
Accepted: 12 June 2007