

Exploring Urban Environments Using Virtual and Augmented Reality

Fotis Liarokapis, Vesna Brujic-Okretic, Stelios Papakonstantinou

City University

giCentre, Department of Information Science, School of Informatics

London EC1V 0HB

email: fotisl, vesna, stelios@soi.city.ac.uk

www: www.soi.city.ac.uk/organisation/is/research/giCentre/

Abstract

In this paper, we propose the use of specific system architecture, based on mobile device, for navigation in urban environments. The aim of this work is to assess how virtual and augmented reality interface paradigms can provide enhanced location based services using real-time techniques in the context of these two different technologies. The virtual reality interface is based on faithful graphical representation of the localities of interest, coupled with sensory information on the location and orientation of the user, while the augmented reality interface uses computer vision techniques to capture patterns from the real environment and overlay additional way-finding information, aligned with real imagery, in real-time. The knowledge obtained from the evaluation of the virtual reality navigational experience has been used to inform the design of the augmented reality interface. Initial results of the user testing of the experimental augmented reality system for navigation are presented.

Keywords: Mobile Interfaces, Augmented and Virtual Environments, Virtual Tours, Humancomputer interaction.

1 Introduction

Navigating in urban environments is one of the most compelling challenges of wearable and ubiquitous computing. The term navigation which can be defined as the process of moving in an environment can be extended to include the process of wayfinding [DS93]. Wayfinding refers to the process of determining one or more routes (also known as paths). Mobile computing has brought the infrastructure for providing navigational and wayfinding assistance to users, anywhere and anytime. Moreover, recent advances in positioning technologies - as well as virtual reality (VR), augmented reality (AR) and user interfaces (UIs) - pose new challenges to researchers to create effective wearable navigation environments. Although a number of prototypes have been developed in the past few years there is no system that can provide a robust solution for unprepared urban navigation. There has been significant research in position and orientation navigation in urban environments. Experimental systems that have been designed range from simple location-based services to more complicated VR and AR interfaces.

An account of the user's cognitive environment is required to ensure that representations are not just delivered on technical but also usability criteria. A key concept for all mobile applications based upon location is the 'cognitive map' of the environment held in mental image form by the user. Studies have shown that cognitive maps have asymmetries (distances between points are different in different directions), that they are resolution-dependent (the greater the den-

Digital Peer Publishing Licence

Any party may pass on this Work by electronic means and make it available for download under the terms and conditions of the current version of the Digital Peer Publishing Licence (DPPL). The text of the licence may be accessed and retrieved via Internet at <http://www.dipp.nrw.de/>.

First presented at the International Conference on Computer Graphics Theory and Applications (GRAPP) 2006, extended and revised for JVRB

sity of information the greater the distance between two points) and that they are alignment-dependent (distances are influenced by geographical orientation) [Tve81]. Thus, calibration of application space concepts against the cognitive frame(s) of reference is vital to usability. Reference frames can be divided into the egocentric (from the perspective of the perceiver) and the allocentric (from the perspective of some external framework) [Kla98].

End-users can have multiple egocentric and allocentric frames of reference and can transform between them without information loss [MA01]. Scale by contrast is a framing control that selects and makes salient entities and relationships at a level of information content that the perceiver can cognitively manipulate. Whereas an observer establishes a 'viewing scale' dynamically, digital geographic representations must be drawn from a set of preconceived map scales. Inevitably, the cognitive fit with the current activity may not always be acceptable [Rap00].

Alongside the user's cognitive abilities, understanding the spatio-temporal knowledge users have is vital for developing applications. This knowledge may be acquired through landmark recognition, path integration or scene recall, but will generally progress from declarative (landmark lists), to procedural (rules to integrate landmarks) to configurational knowledge (landmarks and their inter-relations) [SW75]. There are quite significant differences between these modes of knowledge, requiring distinct approaches to application support on a mobile device. Hence, research has been carried out on landmark saliency [MD01] and on the process of self-localisation [Sho01] in the context of navigation applications.

This work demonstrates that the cognitive value of landmarks is in preparation for the unfamiliar and that self-localisation proceeds by the establishment of rotations and translations of body coordinates with landmarks. Research has also been carried out on spatial language for direction-giving, showing, for example, those paths prepositions such as along and past is distance-dependent [KBZ⁺01]. These findings suggest that mobile applications need to help users add to their knowledge and use it in real navigation activities. Holl et al, [HLSM03] illustrate the achievability of this aim by demonstrating that users who pre-trained for a new routing task in a VR environment made fewer errors than those who did not. This finding encourages us to develop navigational wayfinding and commentary support on mobile devices accessible to the cus-

tomers.

The objectives of this research include a number of urban navigation issues ranging from mobile VR to mobile AR. The rest of the paper is structured as follows. In section 2, we present background work while in section 4 we describe the architecture of our mobile solution and explain briefly the major components. Sections 5 and 6 present the most significant design issues faced when building the VR interface, together with the evaluation of some initial results. In section 8, we present the initial results of the development towards a mobile AR interface that can be used as a tool to provide location and orientation-based services to the user. Finally, we conclude and present our future plans.

2 Background Work

There are a few location-based systems that have proposed how to navigate through urban environments. Campus Aware [BGKF02] demonstrated a location-sensitive college campus tour guide, which allows users to annotate physical spaces with text notes. However, user-studies showed that navigation was not well supported. The ActiveCampus project [GSB⁺04] tests whether wearable technology can be used to enhance the classroom and campus experience for a college student. The project also illustrates ActiveCampus Explorer, which provides location aware applications that could be used for navigation. The latest application is EZ NaviWalk, a pedestrian navigation service launched in Japan in October 2003 by KDDI [oTI04] but in terms of visualisation it offers only the 'standard' 2D map.

From the other hand, many VR prototypes have been designed for navigation and exploration purposes. A good overview of the potential and challenges for geographic visualisation has been previously provided [MEH99]. One example is LAMP3D - a system for the location-aware presentation of VRML content on mobile devices, applied in tourist mobile guides [BC05]. Although the system provides tourists with a 3D visualization of the environment they are exploring, synchronized with the physical world through the use of GPS data, there is no orientation information available. Darken and Sibert [DS96] examined whether real world wayfinding and environmental design principles can be effective in designing large virtual environments that support skilled wayfinding behaviour.

Another example is the mobile multi-modal interaction platform [WSK03] which supports both indoor and outdoor pedestrian navigation by combining 3D graphics with synthesised speed generation. Indoor tracking is achieved through infra-red beacon communication while outdoor via GPS. However, the system does not use georeferenced or accurate virtual representations of the real environment, neither report on any evaluation studies. For the route guidance applications, 3D City models have been demonstrated as useful for mobile navigation [KK02], but studies pointed out the need for detailed modelling of the environment and additional route information. To enhance the visualisation, to aid navigation, a combination of 3D scene representation and a digital map were previously used in a single interface [RV01], [LGS03].

In terms of AR navigation, a few experimental systems have been reported on, until present. One of the first wearable navigation systems is MARS (Mobile Augmented Reality Systems) [FMHW97], which aimed at exploring the synergy of two promising fields of user interface research: AR and mobile computing. Thomas et al, [TD⁺98] proposed the use of a wearable AR system with a GPS and a digital compass as a new way of navigating into the environment. Moreover, the ANTS project [RCD04] proposes an AR technological infrastructure that can be used to explore physical and natural structures, mainly for environmental management purposes. Finally, Reitmayr, et al., [RS04] demonstrated the use of mobile AR for collaborative navigation and browsing tasks in an urban environment.

Although the experimental systems listed above focus on some of the issues involved in navigation, they cannot deliver a functional system capable of combining all accessible interfaces, consumer devices and web metaphors. The motivation for the research reported on in this paper is to address those issues, namely an integration of a variety of hardware and software components to provide effective and flexible navigational and wayfinding tool for urban environments. In addition, we compare potential solutions for detecting the user location and orientation in order to provide appropriate urban navigation applications and services.

To realise this we have designed a mobile platform based on both VR and AR interfaces. To understand in depth all the issues that relate to location and orientation-based services, first a VR interface was designed and tested on a personal digital assistant (PDA)

as a navigation tool. Then, we have incorporated the user feedback into the design of an experimental AR interface. Both prototypes require the precise calculation of the user position and orientation, for the registration purpose. The VR interface is coupled with the GPS and digital compass output to correlate the model with the location and orientation of the user, while the AR interface is only dependent on detecting features belonging to the environment.

3 Urban Modelling



Figure 1: Accurate modelling of urban environment (a) high resolution aerial image (b) 3D building extruding

The objectives of this research include issues, such as modelling the urban environment and using visualisation concepts and techniques on a mobile device to help navigation. Currently, the scene surrounding the user is modelled in 3D, and the output is used as a base for both VR and AR navigation scenarios. A partner on the project, GeoInformation Group (GIG), Cambridge, provided a unique and comprehensive data

set, containing the building height/type and footprint data, for the entire City of London. We are using 3D modelling techniques, ranging from manual to semi-automated methods, to create virtual representation of the users immediate environment. The first step of the process involves the extrusion of a geo-referenced 3D mesh using aerial photographs as well as building footprints and heights (Figure 1).

The data set is enhanced by texture information, obtained from the manually captured photographs of the building sides, using a standard, higher resolution digital camera. The steps in the semiautomated technique for preparing and texturing the 3D meshes include: *detaching the objects in the scene; un-flipping the mesh normals; unifying the mesh normals; collapsing mesh faces into polygons and texturing the faces*. An example screenshot of the textured model is shown in Figure 3. All 3D content is held in the GIG City heights database for the test sites in London. The geo-referenced models acquire both the orientation information and the location through a client API on the mobile device, and the application is currently fully functional on a local device. In the final version, the models will be sent to the server in the packet-based message transmitted over the used network. The server will build and render the scene graph associated with the location selected and return it to the client for portrayal.

4 Mobile Platform and Functionality

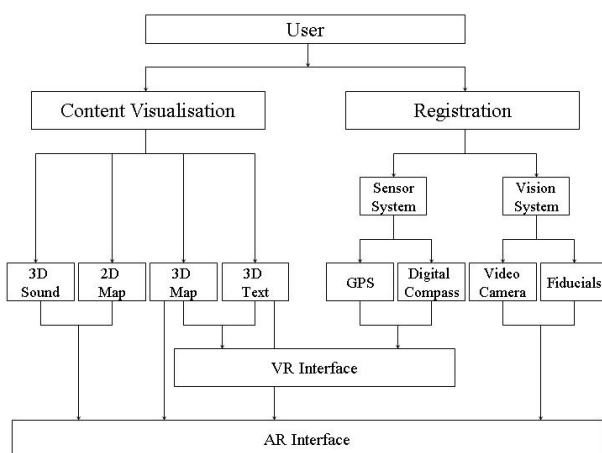


Figure 2: Architecture of our mobile interfaces

Based on these geo-referenced models as building blocks, a generic mobile platform architecture has been designed and implemented for urban navigation

and wayfinding applications and services (Figure 2).

4.1 System Configuration

Figure 2 illustrates the system architecture aimed at optimising navigation by using intelligent data retrieval inside an urban area and providing types of digital appropriately visualised information, suitable to be offered as a core of an enhanced location based service. The hardware configuration consists of two distinct sub-systems: i) the remote server equipment and ii) the client device (e.g. a PDA) enhanced with a selection of sensors and peripherals to facilitate the information acquisition, in real time. Both sides feed into the interface on a mobile device, in the form adequate for the chosen mode of operation.

4.2 System Functionality

Software applications are custom made and include the information retrieval application, clientserver communication software and a cluster of applications on the client side, which process sensory information, in real-time, and ensure seamless integration of the outputs into a unique interface. The calibration and registration algorithms are at the core of the client side applications ensuring all information is geo-referenced and aligned with the real scene. Registration, in this context, is achieved using two different methods: i) a sensor based solution, taking and processing the readings off the sensors directly, and ii) the image analysis techniques coupled with the information on user's location and orientation obtained from the sensors. The sensor system delivers position and orientation data, in real-time, while a vision system is used to identify fiducial points in the scene. All this information is used as input to the VR and AR interfaces. The VR interface uses GPS and digital compass information for locating and orientating the user.

4.3 Interface modalities

Information visualisation techniques used vary according to the nature of the digital content, and/or the navigational task in hand, throughout the navigation. In terms of the content to be visualised, the VR interface can present only 3D maps and textual information. On the other hand, the AR interface uses the calculated user's position and orientation coordinates from the image analysis to superimpose 2D and 3D

maps as well as text and auditory information on the 'spatially aware' framework.

4.4 Notes on Hardware Components

Initially, the mobile software prototype was tested on a portable hardware prototype consisting of a standard laptop computer (equipped with 2.0 GHz M-processor, 1GB RAM and a GeForce FXGo5200 graphics card), a Honeywell HMR 3300 digital compass, a Holux GPS component and a Logitech web-camera (with 1.3 mega-pixel resolution). Then, the prototype system has been ported to a mobile platform based on a Personal Digital Assistant (PDA) and is currently being tested with users.

4.5 Software infrastructure

In terms of the software infrastructure used in this project, both interfaces are implemented based on Microsoft Visual C++ and Microsoft Foundation Classes (MFC). The graphics libraries used are based on OpenGL, Direct3D and VRML. Video operations are supported by the DirectX SDK (DirectShow libraries).

5 Virtual Reality Navigation

Navigation within our virtual environment (the spatial 3D map) can take place in two modes: *automatic* and *manual*. In the automatic mode, GPS automatically feeds and updates the spatial 3D map with respect to the users position in the real space. This mode is designed for intuitive navigation. In the manual mode, the control is fully with the user, and it was designed to provide alternative ways of navigating into areas where we cannot obtain a GPS signal. Users might also want to stop and observe parts of the environment in which case control is left in their hands.

During navigation, there are minor modifications obtained continuously from the GPS to improve the accuracy, which results in minor adjustments in the camera position information. This creates a feeling of instability in user, which can be avoided by simply restricting minor positional adjustments. The immersion provided by GPS navigation is considered as pseudo-egocentric because fundamentally the camera is positioned at a height which does not represent a realistic scenario. If, however, the user switches to manual navigation, any perspective can be obtained, which is very helpful for decision-making purposes.

While in a manual mode, any model can be explored and analysed, therefore additional enhancements of the graphical representation are of vital importance.

One of the problems that quickly surfaced during the system evaluation is the viewing angle during navigation which can make it difficult to position the user. This can make it difficult to understand at which point the user is positioned. After informal observation of users during the development process, an altitude of fifty meters over the surface was finally adopted as adequate. In this way, the user can visualise a broader area plus the tops of the buildings, and acquire richer knowledge about their location, in the VR environment. The height information is hard-coded when the navigation is in the automatic mode because user testing (section 7) showed that it can be extremely useful in cases where a user tries to navigate between tall buildings, having low visibility.



Figure 3: FOV differences (a) low angle (b) high angle

Figure 3, illustrates to what extent the FOV is influenced by that angle and how much more information can be included from the same field-of-view, if the angle is favourable. In both Figure 3 (a) and Figure 3 (b), the camera is placed at exactly the same position

and orientation in the horizontal plane, with the only difference in the pitch angle. In Figure 3 (a), the pitch angle is very low and in the Figure 3 (b) it is set to maximum (90°). This feature was considered important to implement after initial testing. The obvious advantage is that, once in a position, no additional rotations are required from the user to understand the exact position of the camera. Taking into consideration the fact that the normal human viewing angle is about 60° and the application supports angles in the range from 0° to 90°, wide angles (including more objects of the landscape) can be interactively obtained. This can be extremely useful in cases where a user tries to navigate between tall buildings, having low visibility.

We are currently implementing two different technologies for presenting 3D maps on PDA interfaces, involving VRML and Managed Direct3D Mobile (MD3DM). The first solution operates as a stand-alone mobile application and uses VRML technology combined with GPS for determining the position and a digital compass for calculating orientation.

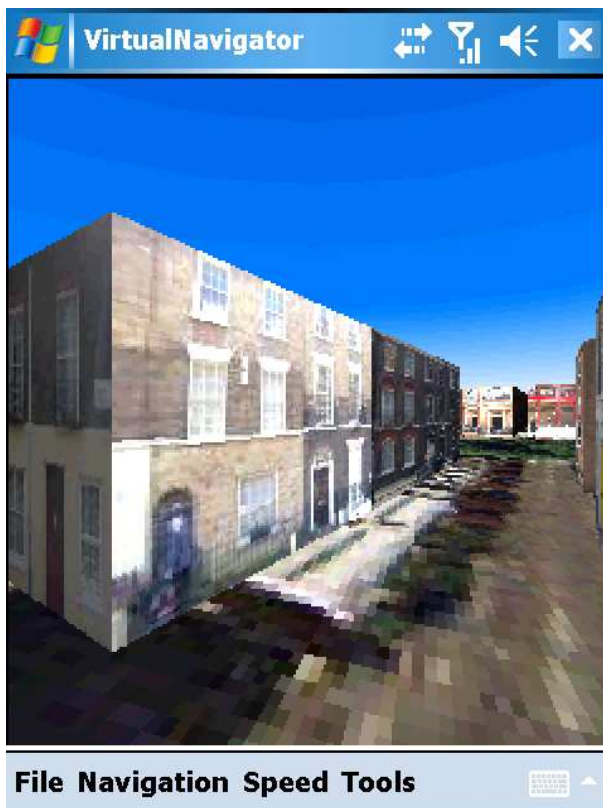


Figure 4: VR navigation in City University's campus

Figure 4 illustrates how the PDA-based navigation inside a virtual environment can be performed. Specifically, stylus interactions can be used to navigate in-

side a realistic virtual representation of City University's campus. Alternatively, menu interactions can be used as another medium for performing navigation and wayfinding tasks. In terms of performance, the frame-rate per second (FPS) achieved varies depending on the device capabilities. For example, using an HTC Universal device the efficiency ranges between 3 to 5 FPS while in a Dell Axim X51v PDA (with a dedicated 16 MB graphics accelerator) the efficiency ranges between 12 to 15 FPS.

The second interface is based on MD3DM that operates as a separate mode, with the aim of handling the output from the GPS/compass automatically providing sufficient functionality to generate mobile VR applications. Compared to the VRML interface, the major advantage of MD3DM is that it takes full advantage of graphics hardware support and enables the development of highperformance three-dimensional rendering [LRBO06]. On the other hand, the major disadvantage of MD3DM is that the Application Programming Interface (API) is low level and thus a lot of functionality which is standard in VRML has to be re-implemented.

6 Preliminary Evaluation

The aims of the evaluation of the VR prototype included assessment of the user experience with particular focus on interaction via movement, identification of specific usability issues with this type of interaction, and to stimulate suggestions regarding future directions for research and development. A 'thinking aloud' evaluation strategy was employed [DFAB04]; this form of observation involves participants talking through the actions they are performing, and what they believe to be happening, whilst interacting with the system. This qualitative form of evaluation is highly appropriate for small numbers of participants testing prototype software: Dix et al, [DFAB04] suggested that the majority of usability problems can be discovered from testing in this way. In addition, Tory and Möller [TM05] argued that formal laboratory user studies can effectively evaluate visualisation when a small sample of expert users is used.

The method used for the evaluation of our VR prototype was based on the Black Box technique which offers the advantage that it does not require the user to hold any low-level information about the design and implementation of the system. The user testing took place at City University campus which includes building structures similar to the surrounding area with

eight subjects in total (testing each one individually). All subjects had a technical background and some were familiar with PDAs. Their age varied between 25 and 55. For each test, each subject followed a pre-determined path represented by a highlighted line. Before the start of the walk, the GPS receiver was turned on and flow of data was guaranteed between it and the 'Registration' entity of the system. The navigational attributes that were qualitatively measured include the: *user perspective, movement with device and decision points.*

6.1 User Perspective

The main point of investigation was to test whether the user can understand where they are located in the VR scene, in correspondence to the real world position. An examination of the initial orientation and level of immersion was also evaluated after minimum interaction with the application and understanding of the available options. The information that was obtained by the users was concerning mainly four topics including: *level-of-detail (LOD), user-perspective, orientation and field-of-view (FOV).*

Most of the participants agreed that the LOD is not sufficiently high for a prototype navigational application. Some concluded that texture based models would be a lot more appropriate but others expressed the opinion that more abstract, succinct annotations would help, at a different level (i.e. A to Z abstract representations). Both groups of answers can fit in the same context, if all interactions could be visualised from more than one perspective. A suggested improvement was to add *geo-bookmarks* (also known as hotspots) that would embed information about the nature of the structures or even the real world functionality.

As far as the 'user-perspective' attribute is concerned, each user expressed a different optimal solution. Some concluded that more than one perspective is required to fully comprehend their position and orientation. Both perspectives, the egocentric and the allocentric, are useful during navigation for different reasons [LGM⁺05] and under different circumstances. During the initial registration, it would be more appropriate to view the model from an allocentric point of view (which would cover a larger area) and by minimising the LOD just to include annotations over buildings and roads. This proved easier to increase the level of immersion with the system but not being directly exposed to particular information such as

the structure of the buildings. In contrast, an egocentric perspective is considered productive only when the user was in constant movement. When in movement, the VR interface retrieves many updates and the number of decision points is increased. Further studies should be made on how the system would assist an everyday user, but a variation on the user perspective is considered useful in most cases.

The orientation mechanism provided by the VR application consists of two parts. The first maintains the user's previous orientation whilst the second restores the camera to the predefined orientation (which is parallel to the ground). Some users noted that when angle direction points towards the ground gives better appreciation of the virtual navigation. Another subject that the users agree in is the occurrence of fast updates. This can make it difficult to navigate, because the user needs to align the camera on three axes and not two. Based on our experiments we noticed that the used orientation mechanisms are inadequate for navigational purpose and it is imperative that the scene should be aligned in the same direction as the device in the real world.

Furthermore, all participants appreciated the user-maintained FOV. They agreed that it should be wide enough to include as much information, on the screen, as possible. They added that in the primary viewing angle, there should be included recognisable landmarks that would aid the user comprehend the initial positioning. One mentioned that the orientation should stay constant between consecutive decision points, and hence should not be gesturebased. Most users agreed that the functionality of the VR interface provides a wide enough viewing angle able to recognise some of the surroundings even when positioned between groups of buildings with low detail level.

6.2 Movement with the Device

The purpose of this stage was to explore how respondents interpreted their interaction with the device, whilst moving. The main characteristics include the *large number of updates* as well as the *change of direction* followed by the user. The elements, which are going to be discussed, are mainly considered with the issues of making the navigation easier, the use of the most appropriate perspective, and the accuracy of the underlying system as well as the performance issues that drive the application. One important issue is to consider the inheritance of a specific perspective for

use throughout the navigation process. Some participants mentioned the lack of accurate direction way-points that would assist route tracking. A potential solution is to consider the adoption of a user-focused FOV during navigation using a simple line on the surface of the model. However, this was considered partially inadequate because the user expects more guidance when reaching a decision point. Some participants suggested to use arrows on top of the route line which would be either visible for the whole duration of the movement or when a decision point was reached.

In addition, it was positively suggested that the route line should be more distinct, minimising the probability of missing it while moving. Some expressed the opinion that the addition of recognisable landmarks would provide a clearer cognitive link between the VR environment and the real world scene. However, the outcomes of this method are useful only for registering the users in the scene and not for navigation purposes. A couple of participants included in their answers that the performance of the system was very satisfactory. This is an important factor to consider, because in the change of the camera position occurs when new data is being retrieved from the external sensor. The characterisation, of the position transition, as smooth reflects that the main objective of any actor is to obtain new information about his position, at the time it is available. The latency that the system supports is equal to the latency the H/W receiver obtains meaning that the performance of the application is solely dependent on the quality of operating hardware. The adaptation to a mobile operating system (i.e. Windows Mobile 5.0) would significantly increase the latency of the system, since devices are not powerful enough to handle heavy operations.

Moreover, opinions, about the accuracy of the system, differ. One of respondents was convinced that the accuracy, provided by the GPS receiver, was inside the acceptable boundaries, which reflected the GPS specifications supporting that the level of accuracy between urban canyons was reflecting the correspondence to reality, in a good manner. A second test subject revealed that the occlusion problem was in effect due to GPS inaccuracy reasons underlining that when the GPS position was not accurate enough, the possibility to miss the route line or any developed direction system increased. Both opinions are equally respected and highlighted the need for additional feedback.

6.3 Decision Points

The last stage is concerned with the decision points and the ability of the user to continue the interaction with the system when it reaches them. A brief analysis of the users' answers is provided to identify ways forward with the design, but full analysis will be published in a separate publication. As described previously, the user has the feeling of full freedom to move at any direction, without being restricted by any visualisation limitations of the computergenerated environment. Nonetheless, participants may feel overwhelmed by the numerous options they may have available and be confused about what action to take next. We take into consideration that large proportion of users is not sufficiently experienced in 3D navigational systems and the appropriate time is given to them to familiarise with the system.

Preliminary feedback suggests that some users would prefer the application to be capable of manipulating the users perspective automatically, when a decision point (or, an area close to it) is reached. This should help absorb more information about the current position as well as supporting the future decision making process. Another interesting point relates to the provision of choice to the user in the future to accommodate sudden, external factors that may allow them to detour from a default path. Partially, some of these requirements would be met if the user could manually add geo-bookmarks in the VR environment representing points in space with supplementary personal context. The detailed analysis of the responses will be taken into account in further developments of the system, which is underway.

7 Augmented Reality Navigation

The AR interface is the alternative way of navigating in the urban environment using mobile systems. Unlike the VR interface, which uses the hardware sensor solution (a GPS component and a digital compass), the AR interface uses a webcam (or video camera) and computer vision techniques to calculate position and orientation. Based on the findings of the previous section and a previously developed prototypes [Lia05], [LGM⁺05], a high-level AR interface has been designed for outdoor use. The major difference with other existing AR interfaces, such as the ones described in [FMHW97], [TD⁺98], [RS04] and [RCD04], is that our approach allows for the combi-

nation of four different types of navigational information: 3D maps, 2D maps, text and spatial sound. In addition, two different modes of registration have been designed and experimented upon, based upon *fiducial points* and *feature* recognition. The purpose for the exercise was to understand some of the issues involved in two of the key aspects of urban navigation: *wayfinding* and *commentary*.

In the fiducial points recognition mode, the outdoor environment needs to be populated with fiducials prior to the navigational experience. Fiducials are placed in points-of-interest of the environment, such as corners of the buildings, ends of streets etc, and play a significant role in the decision making process. In our current implementation we have adopted ARToolKit's template matching algorithm [KB99] for detecting marker cards and we try to extend it for natural feature detection. Features that we currently detect can come in different shapes, such as square, rectangular, parallelogram, trapezium and rhomb [Lia05] similar to shapes that exist in the environment. In addition, it is not convenient, sometimes it is even impossible, to populate large urban areas with fiducials. Therefore, we have experimentally used road signs as fiducials to compute the users pose [LRBO06]. Road signs are most of the time printed in black colour on a white background. Also, they are usually placed at the decision points, such as beginning and ending of streets, corners and junctions. As a result, if a highresolution camera is used to capture the object, it is relatively easy to detect the road signs, as illustrated in Figure 5.

One of the known limitations of this technique is that, sometimes, road signs are not in a good condition, which makes it more difficult to recognise a pattern. Also, the size of the road signs is usually fixed (depending on the urban area) limiting severely the number of operations that can be done on it. An example screenshot of how road signs can be used in practice as fiducial points (instead of using pre-determined markers) during urban navigation is illustrated in 5.

Alternatively, in the feature recognition, the user is 'searching' to detect natural features of the real environment to serve as 'fiducial points and pointsof-interest, respectively. Distinctive natural features like door entrances, windows etc, have been experimentally tested to see whether they can be used as 'natural markers'. Figure 6 shows the display presented to a user navigating in City University's campus, to acquire location and orientation information using 'natural markers'.



Figure 5: Pattern recognition of road signs: (a) original image; (b) detected image

As soon as the user turns the camera (on a mobile device) towards these predefined natural markers, audio-visual information (3D arrows, textual and/or auditory information) can be superimposed on the real-scene imagery (Figure 7), thus satisfying some of the requirements identified in section 6.1. Userstudies for tour guide systems showed that visual information could sometimes distract the user [BGKF02], while audio information could be used to decrease the distraction [WAHS01]. With this in mind, we have introduced a spatially referenced sound into the interface, to be used simultaneously with the visual information. In our preliminary test case scenario, a pre-recorded sound file is assigned to the corresponding fiducial point, for each pointof-interest. As the user approaches a fiducial point, commentary information can be spatially identified; the closer the user to the object the louder the volume of the commentary audio information. Depending on the end-user's preferences,

or needs, the system allows for a different type of digital information to be selected and superimposed. For example, for visually impaired users audio information may be preferred to use over visual, or a combination of the two may be found optimal [Lia05]. A coarse comparison between the use of fiducial points and the feature recognition mode is shown in Table 1. Further testing is underway and the detailed analysis will be published in a separate publication.

Recognition Mode	Range	Error	Robustness
Fiducial	0.5~ 2m	Low	High
Feature	2~ 10m	High	Low

Table 1: Fiducial vs feature recognition mode

In the feature recognition mode, the advantage is that the range within which it may operate is much greater because it does not require preparation of the environment. Thus, it can be applied when wayfinding is the focus of the navigation. However, the natural feature tracking algorithm, which is used in this scenario, does require improved accuracy of the position and orientation information, which is currently limited. In contrast, the fiducial points recognition mode offers the advantage of a very low error during the tracking process (i.e. detecting fiducial points). However, the limited space of operation due to the need to populate the area with tags, makes it more appropriate for confined areas and commentary navigation modes. The research suggests, however, that the combination of fiducial and feature recognition modes allows the user to pursue both wayfinding and commentary based navigation into urban environments within a single application.



Figure 6: Road sign pedestrian navigation

8 Discussion

After completing the development of a portable prototype application (based on a laptop computer based) specific requirements to enhance the user interface and interaction mechanisms on a mobile device (PDA) were identified. Through this research, it was found obligatory to retrieve and visualise spatio-temporal content from a remote server in order to support real-time operation and meet the information needs of a user. This was accomplished by transmitting geographic coordinates (i.e. GPS input) to the server-side and automatically retrieving geo-referenced information in the form of VRML 3D maps. The 3D content was designed to cover an area encompassing the current position of the user and the position of one or more actors/points-of-interest in their proximity. The quality and accuracy of these models are proved good while the techniques used are customdeveloped and based on a semi-automated routine, developed in a specialised software development environment.

9 Conclusions

This paper addresses how virtual and augmented reality interface paradigms can provide enhanced location based services for urban navigation and wayfinding. The VR interface operates on a PDA and presents a realistic and geo-referenced graphical representation of the localities of interest, coupled with sensory information on the location and orientation of the user. The knowledge obtained from the evaluation of the VR navigational experience has been used to inform

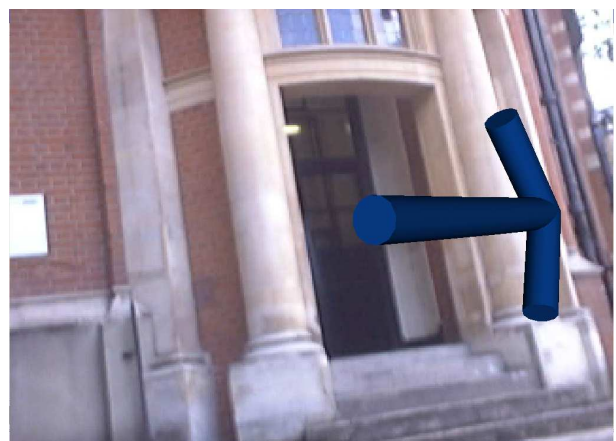


Figure 7: Detecting door entrances

the design of the AR interface which operates on a portable computer and overlays additional way-finding information onto the captured patterns from the real environment.

Both systems calculate the user's position and orientation, but using a different methodology. The VR interface relies on a combination of GPS and digital compass data whereas the AR interface is only dependent on detecting features of the immediate environment. In terms of information visualisation, the VR interface can only present 3D maps and textual information while the AR interface can, in addition, handle other relative geographical information, such as digitised maps and spatial auditory information.

Work on both modes and interfaces is in progress and we also consider a hybrid approach, which aims to find a balance between the use of hardware sensors (GPS and digital compass) and software techniques (computer vision) to achieve the best registration results. In parallel, we are designing a spatial database to store our geo-referenced urban data, which will feed the client-side interfaces as well as routing algorithms, which we are developing to provide more services to mobile users. The next step in the project is a thorough evaluation process, using both qualitative and quantitative methods. The results will be published in due course.

10 Acknowledgments

The work presented in this paper is conducted within the LOCUS project, funded by EPSRC, through the Location and Timing (KTN) Network. We would also like to thank our partner on the project, GeoInformation Group, Cambridge, for making the entire database of the City of London buildings available to the project. The invaluable input from David Mountain on resolving the sensor fusion issues and from Christos Gatzidis for generating components of the 3D content are greatly acknowledged.

References

- [BC05] Stefano Burigat and Luca Chittaro, *Location-aware visualization of VRML models in GPS-based mobile guides*, Proceedings of the 10th International Conference on 3D Web Technology, ACM Press, 2005, ISBN 1-59593-012-4, pp. 57–64.
- [BGKF02] Jenna Burrell, Geri K. Guy, Kiyo Kubo, and Nick Farina, *Context-aware computing: test case*, Proceedings of UbiComp, Lecture Notes in Computer Science Vol. 2498, Springer, 2002, ISBN 3-540-44267-7, pp. 1–15.
- [DFAB04] Alain J. Dix, Janet E. Finlay, Gregory D. Abowd, and Russel Beale, *Human-Computer Interaction*, 3rd edition ed., Prentice Hall, Harlow, 2004, ISBN 0-13-046109-1.
- [DS93] Rudy P. Darken and John L. Sibert, *A toolset for navigation in virtual environments*, Proceedings of the 6th annual ACM symposium on User interface software and technology (New York, NY, USA), ACM Press, 1993, ISBN 0-89791-628-X, pp. 157–165.
- [DS96] Rudy P. Darken and John L. Sibert, *Navigating Large Virtual Spaces*, International Journal of Human-Computer Interaction **8** (1996), no. 1, 49–72, ISSN 1044-7318.
- [FMHW97] Steven Feiner, Blair MacIntyre, Tobias Höllerer, and Antony Webster, *A touring machine: Prototyping 3D mobile augmented reality systems for exploring the urban environment*, Proceedings of the 1st IEEE International Symposium on Wearable Computers, IEEE Computer Society, 1997, ISBN 0-8186-8192-6, pp. 74–81.
- [GSB⁺04] William G. Griswold, Patricia Shanahan, Steven W. Brown, Robert S. Boyer, Matt Ratta. R. Benjamin Shapiro, and Tan Minh Truong, *ActiveCampus: Experiments in Community-Oriented Ubiquitous Computing*, Computer **37** (2004), no. 10, 73–81, ISSN 0018-9162.
- [HLSM03] Doris Höll, Bernd Lepow, Robby Schöpfung, and Maximilian Mehdorn, *Spatial cognition III*, Lecture Notes in Computer Science Vol. 2685, ch. Is it possible to learn and transfer spatial information from virtual to real worlds?, pp. 143–156, Springer, Berlin, 2003, ISBN 3-540-40430-9.

- [KB99] Hirokazu Kato and Mark Billinghurst, *Marker Tracking and HMD Calibration for a Video-Based Augmented Reality Conferencing System*, Proceedings of the 2nd IEEE and ACM International Workshop on Augmented Reality, IEEE Computer Society, 1999, ISBN 0-7695-0359-4, pp. 85–94.
- [KBZ⁺01] Christian Kray, Jörg Baus, Hubert D. Zimmer, Harry R. Speiser, and Antonio Krüger, *Two path prepositions: along and past*, Proceedings of the International Conference on Spatial Information Theory: Foundations fo Geographic Information Science (London) (D. Montello, ed.), Lecture Notes in Computer Science Vol. 2205, Springer, 2001, ISBN 3-540-42613-2, pp. 263–277.
- [KK02] M. Kulju and E. Kaasinen, *Guidance Using a 3D City Model on a Mobile Device*, Workshop on Mobile Tourism Support Mobile HCI 2002 Symposium, Pisa, Italy, Sept. 17th, 2002.
- [Kla98] R. L. Klatzky, *Spatial cognition - An interdisciplinary approach to representation and processing of spatial knowledge*, ch. Allocentric and egocentric spatial representations: Definitions, distinctions, and interconnections, pp. 1–18, Springer, Berlin, 1998, ISBN 3-540-64603-5.
- [LGM⁺05] Fotis Liarokapis, Ian Greatbatch, David Mountain, Anil Gunesh, Vesna Brujic-Okretic, and Johnathan Raper, *Mobile Augmented Reality Techniques for Geo-Visualisation*, Proceedings of the 9th International Conference on Information Visualisation IV'05, IEEE Computer Society, 2005, ISBN 0-7695-2397-8, pp. 745–751.
- [LGS03] K. Laakso, O. Gjesdal, and J.R. Sulebak, *Tourist information and navigation support by using 3D maps displayed on mobile devices*, Workshop on Mobile Guides, Mobile HCI 2003 Symposium, Udine, Italy, 2003.
- [Lia05] Fotis Liarokapis, *Augmented Reality Interfaces Architectures for Visualising and Interacting with Virtual Information*, Ph.D. thesis, School of Science and Technology, University of Sussex, Department of Informatics, 2005, Sussex theses S 5931, ISBN/ISSN/CNM0426866US.
- [LRBO06] Fotis Liarokapis, Johnathan Raper, and Vesna Brujic-Okretic, *Navigating within the urban environment using Location and Orientation-based Services*, European Navigation Conference, 710 May, Manchester, UK, 2006.
- [MA01] Christy R. Miller and Gary L. Allen, *Spatial frames of reference used in identifying directions of movement: an unexpected turn*, Proceedings of the International Conference on Spatial Information Theory: Foundations of Geographic Information Science (London) (D. Montello, ed.), Lecture Notes in Computer Science Vol 2205, Springer, 2001, ISBN 3-540-42613-2, pp. 206–216.
- [MD01] Pierre-Emmanuel Michon and Michel Denis, *Proceedings of the international conference on spatial information theory: Foundations of geographic information science*, Lecture Notes in Computer Science Vol 2205, ch. When and why are visual landmarks used in giving directions, pp. 292–305, London, 2001, ISBN 3-540-42613-2.
- [MEH99] A.M. MacEachren, R. Edsall, and D. Hauq, *Virtual environments for Geographic Visualization: Potential and Challenges*, Proceedings of the ACM Workshop on New Paradigms in Information Visualization and Manipulation, ACM Press, 1999, ISBN 1-58113-254-9, pp. 35–40.
- [oTI04] DTI Department of Trade and Industry, *Location-based services: understanding the japanese experience, global watch mission report*, http://www.oti.globalwatchonline.com/online_pdfs/36246MR.pdf, 2004, visited: 10/02/2006.

- [Rap00] Johnathan F. Raper, *Multidimensional geographic information science*, Taylor and Francis, London, 2000, ISBN 0-7484-0506-2.
- [RCD04] T. Romão, N. Correia, and E. Dias, *ANTS-Augmented Environments*, Computers & Graphics **28** (2004), no. 5, 625–633, ISSN 0097-8493.
- [RS04] Gerhard Reitmayr and Dieter Schmalstieg, *Collaborative Augmented Reality for Outdoor Navigation and Information Browsing*, Proceedings of the Symposium of Location Based Services and TeleCartography, Vienna, Austria, January 2004, 2004, pp. 31–41.
- [RV01] I. Rakkolainen and T. Vainio, *A 3D City Info for Mobile Users*, Computer & Graphics **25** (2001), no. 4, 619–625, ISSN 0097-8493.
- [Sho01] M. Jeanne Sholl, *The role of self reference system in spatial navigation*, Proceedings of the International Conference on Spatial Information Theory: Foundations of Geographic Information Science (London) (D. Montello, ed.), Lecture Notes in Computer Science Vol. 2205, Springer, 2001, ISBN 3-540-42613-2, pp. 217–232.
- [SW75] A.W. Siegel and S.H. White, *The development of spatial representation of large scale environments*, Advances in child development and Behaviour **10** (1975), 9–55, ISSN 0065-2407.
- [TD⁺98] B.H. Thomas, , V. Demczuk, W. Piekarski, D. Hepworth, and B. Gunther, *A Wearable Computer System with Augmented Reality to Support Terrestrial Navigation*, Proceedings of the 2nd International Symposium on Wearable Computers, IEEE and ACM, 1998, ISBN 0-8186-9074-7, pp. 168–171.
- [TM05] M. Tory and T. Möller, *Evaluating Visualizations: Do Expert Reviews Work?*, Computer Graphics and Applications **25** (2005), no. 5, 8–11, ISSN 0272-1716.
- [Tve81] B. Tversky, *Distortions in memory for maps*, Cognitive Psychology **13** (1981), 407–433, ISBN 0010-0285.
- [WAHS01] Allison Woodruff, Paul M. Aoki, Amy Hurst, and Margaret H. Szymanski, *Electronic Guidebooks and Visitor Attention*, Proceedings of 6th International Cultural Heritage Informatics Meeting ICHIM'01, Milan, Italy, Sep. 2001, 2001, ISBN 1-885626-24-X, pp. 437–454.
- [WSK03] Rainer Wasinger, Christoph Stahl, and Antonia Krüger, *Mobile Multi-Modal Pedestrian Navigation*, Second International Workshop on Interactive Graphical Communication IGC 2003, London, 2003.

Citation
Fotis Liarokapis, Vesna Brujic-Okretic, Stelios Papakonstantinou, <i>Exploring Urban Environments Using Virtual and Augmented</i> , Journal of Virtual Reality and Broadcasting, 3(2006), no. 5, December 2006, urn:nbn:de:0009-6-7720, ISSN 1860-2037.