

Interacting with Virtual Reality models on mobile devices

David Mountain, Fotis Liarokapis, Jonathan Raper, Vesna Brujic-Okretic

GI Science Group, Department of Information Science
City University, London, EC1V 0HB, UK
+44 (0)20 7040 4044
{dmm, fotisl, raper, vesna}@soi.city.ac.uk

ABSTRACT

This paper discusses alternative approaches for interacting with virtual reality models on mobile devices, based upon work conducted as part of the locus project [6]. First, mobile virtual reality models are discussed within the established reality-virtuality continuum to demonstrate how this new mixed reality paradigm compares with augmented and virtual reality generally. Next, three prototypes are introduced that adopt different interaction paradigms for mobile virtual reality models: interaction can be via the screen only, movement and gestures within the real world environment, or a mixture of these two approaches. Finally this paper describes some potential applications of mobile virtual reality, including navigation, and the display of georeferenced information. The paper concludes by suggesting that interaction via movement and gestures within the may be a more intuitive approach for mobile virtual reality models.

Categories and Subject Descriptors

H.5.2 [User Interfaces]: Input devices and strategies; Interaction styles; Prototyping

General Terms

Design, Experimentation, Human Factors.

Keywords

Mobile computing; virtual reality; augmented reality; mixed reality; 3D display; positioning systems; scene registration; tracking; navigation; georeferenced information; information retrieval;

1. INTRODUCTION

The increased processing and screen display capabilities of modern handheld devices has made created a new mobile environment for applications, that could traditionally only be deployed on heavy, static hardware such as desktop machines or laptops. Mobile users tend to behave differently when using handheld devices outdoors [11] and new approaches for interacting with virtual reality models and georeferenced information are emerging in this context. The LOCUS project [6] is an EPSRC Pinpoint Faraday [14] research project aiming to enhancing Location Based Services (LBS) by extending the traditional map-based approach. LOCUS is currently investigating the use of both virtual reality (VR) and augmented reality (AR) interfaces for mobile devices such as Personal Digital Assistants (PDAs) and 3G phones. This paper discusses the research conducted into the rendering of VR models on mobile devices as

well as new interaction techniques for these models based upon both screen interaction and external sensors.

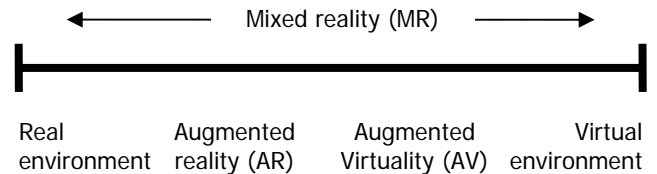


Figure 1. The Reality-Virtuality (RV) Continuum [7]

1.1 Reality-virtuality continuum

Within the field of mixed reality (MR) display systems, a taxonomy is used to distinguish different classes of display. This taxonomy takes the form of a continuum with real (physical) environments at one extreme and purely virtual environment on the other (figure 1). Between these two extremes, several classes of mixed reality displays systems have been defined. These classes were distinguished by three main factors [7];

Reality: the degree to which the world scene is 'real', or computer generated (virtual).

Immersion: whether the user perceives themselves to be immersed within the environment (egocentric), or observing from outside (exocentric).

Directness: whether the world objects are viewed directly (i.e. via optical see-through display) or electronically synthesised (i.e. via video see-through display).

The advent of mobile computing provides a new opportunity for MR applications. A VR model can be rendered on a handheld mobile device - such as a PDA or a 3G phone - and this model can be registered with the real environment through a combination of user interaction techniques and external sensors, discussed in section 2. By adopting this approach, movement in the real environment is emulated in the virtual environment hence the VR model provides an alternative view of the physical environment within which the user is moving. This model may be realistic, attempting to emulate the real world scene to a high degree of fidelity, or abstract, conveying some attribute information about surrounding objects, such as the primary use of a nearby building.

Such a mobile VR system does not map well to the established reality-virtuality taxonomy. In terms of the reality factor, the

mobile VR display system is both, since the user can see both the real world scene around them, and a computer generated scene on the device. There is a very high degree of immersion with this system since it is intended to be used by mobile users who are navigating within a real environment. This movement can be used to control observer location within the virtual environment providing an egocentric view within the VR model. There is some ambiguity about the directness factor. The user has a completely direct (unsynthesised) view of their surrounding real world environment. Unlike optical see-through devices however, virtual objects are not overlaid on this representation, but can be placed within the VR model that is displayed on the screen of their mobile device. This paper argues that this highly immersive VR system, that can offer realistic or abstract views of the scene surrounding a user, provides a new platform for navigation and the display of georeferenced information.

1.2 Gesture computing

Gesture recognition sensors have been suggested as an alternative form of input for handheld computers. These have taken many forms such as tilt, acceleration, pressure, conductivity and capacitance sensors for human-computer interaction [3]. Of particular interest to this study are sensors that can measure the location and orientation of a handheld mobile device. Such a gesture recognition system allows a user to interact with a computer through their physical behaviour in the real environment, for example walking and turning. As will be discussed in section 2, this behaviour can be used to control the viewpoint of an observer in a virtual environment.

1.3 Degrees of freedom

In mixed reality applications, *tracking* describes the process of calculating the scene coordinates of moving objects, in real time [17]. The viewpoint of a moving object is comprised of two components: *position* and *orientation*. Position can be described by three *degrees of freedom (DOF)*: two horizontal coordinates (x and y), and one vertical (z). Orientation can also be described by three DOF: heading, pitch and roll. These six DOF can be used to define the viewpoint of an observer in a virtual or augmented reality scene.



Figure 2: Registering the VR model with the real world scene via screen interaction on a mobile device

2. MOBILE VR INTERACTION

Three-dimensional VR models of real environments can be constructed from a variety of data sources. Digital elevation models provide an effective way of modelling the continuous variation in terrain associated with natural features at a coarse scale, such as ridges, channels, peaks and passes [19]. For modelling urban landscapes, building outlines and height data can better model the sharp variation in height associated with the edges of building. Realism can be added to the model through the use of textures added to building facades (figure 3a).

Interaction with such models on mobile devices can be via screen interaction, from external sensors that track the user's location and orientation, or a combination of the two. Various prototypes are being explored, which have dependencies on different technologies, operating systems, and software.

2.1 Screen interaction only

In this mode, all six DOFs controlling the observer location within the VR model are manipulated by user interaction with the display of the mobile device. Since there is no input from external sensors, the user must navigate within the virtual world to visualise and compare the virtual scene with the one seen in the real environment, in the same way that users interact with VR models on static (desktop) machines. Although screen interaction provides a poor form of tracking it allows the user to explore their surroundings in the virtual world in three dimensions, prior to navigating in the real world. This can be used for decision making for navigation prior to movement, for example to explore alternative routes to a destination, or search for a known location when lost. This system can be deployed easily by using a VRML client running on a Windows Pocket PC device [12].

2.2 Sensor control with screen interaction

In this prototype, user location is controlled from an external sensor - a Global Positioning System (GPS) receiver [13] - connected wirelessly to the handheld mobile device. The x and y coordinates from the GPS must be converted from the latitude and longitude coordinates, using the WGS84 datum, to the local coordinate system used by the VR model. GeoVRML [18] offers potential as a way of handling these transformations, although there is limited support for projected coordinate systems such as the British National Grid. GPS can also provide the elevation of the receiver above the WGS84 datum, which must first be converted to height above local sea level to be of use. Experience suggests that errors in the vertical (z) axis are more noticeable than those in the horizontal (x and y) axes, and could lead to the user's position being below or above the ground, or oscillating between the two: a more desirable solution may be to set the height of the observer location to a fixed value above ground level. This has the added advantage that the user can move from an egocentric view, where the observer location in the virtual world is set to eye level at their physical location in the real world scene, to an exocentric view that provides the user with a synopsis of the surrounding region [4]. By increasing the altitude of the observer location in the VR scene, through screen interaction, and altering the pitch to look directly down, the user gets a planar (bird's eye) view of their surroundings - analogous to a map view - that could be used for personal orientation and navigation (figure 3).

An orientation parameter - heading - can also be derived from GPS data, however this is calculated by comparing two or more

GPS positions and hence can only tell you the direction in which a user is moving, not the direction in which the mobile device is pointing. When the user is not moving, the heading information will not be accurate, reflecting the minor deviations in position associated with changing positional error.

2.3 Full sensor control

By integrating another sensor to measure orientation parameters, all six DOF can be controlled by the movement and gestures of the user in the real environment. Commercial three-axis digital compasses are available that can provide heading, pitch and roll measurements [15]. In this mode therefore, the orientation parameters of the observer in the virtual environment are controlled by the user interaction with the digital compass, which can be physically attached to the mobile device itself. By holding the device in front of you and turning around, the orientation of the observer location in the VR model will change correspondingly. Physically tilting the device up or down in the physical environment changes the observer view in the model to look up or down. Similarly by tilting the device from side to side, the roll within the VR model can be controlled. This full six DOF



a. egocentric perspective



b. exocentric perspective

Figure 3: Seamless navigation between two perspectives

control comes with a high processing overhead for a mobile device, which will introduce latency between movement and gestures in the virtual environment, and update of observer location in the VR model.

Beyond the performance issues associated with handling sensor input for all six DOFs, this fully automated sensor approach may also not be desirable for mobile VR models in terms of effective human-computer interaction. As mentioned in section 2.2, allowing the user to control altitude and pitch allows them to move from an egocentric to an exocentric perspective, which may offer advantages for navigation and self-localisation. The use of gestures to control pitch and roll is as yet untested. Roll is likely to be of limited use for ground based navigation, where users expect the world to be aligned vertically. Users of mobile devices also frequently change the pitch (tilt) of their mobile devices, depending upon whether they are sitting or standing, or to find an angle where the screen is more readable in strong sunlight, and this may prove to one degree of freedom too many for the usability of a mobile VR display system. Evaluation of these different sensor inputs aims to lead to the development of a system that allows control through movement and gestures, without introducing unwanted changes in the observer location in the VR model.

3. POTENTIAL APPLICATIONS

3.1 Navigation and exploration

The primary application for mobile VR display systems is likely to be for navigation and self-localisation tools. 3D models of buildings and landscape, when augmented with additional guidance information, such as route directions and street names, have been shown to be an effective way of providing people with route guidance on mobile devices [5].

The effectiveness of navigation tools is dependent to a large degree upon the selection of, and symbology used to represent, key decision points and landmarks along the route. Measures of the salience of 3D features within a geodatabase to act as landmarks for route guidance have been formalised, many of which can be extracted from the geometry of a building and its surroundings [10]. Effective symbology is also required for route guidance, to provide unambiguous guidance at decision points, and confirmation that the user is on route between decision points. As well as providing explicit routing guidance, mobile VR display systems offer the ability to explore unknown environments within the virtual environment via screen interaction, prior to physical movement in the real environment. This opportunity to plan ahead offers benefits in terms of providing familiarity of an unvisited location, which can save time when exploring a new area for the first time.

3.2 Displaying georeferenced information

Within the field of mobile computing, there are an increasing number of applications that can provide information about the facilities, services and features of interest that are within a specified spatial proximity of a mobile user (“where’s my nearest” services). Some of these services require the user to know and state their location using toponyms [20]. An alternative approach which makes fewer demands upon the cognitive load of the user is to use a positioning system that provides spatial filters such as the familiar “within x metres” of current location, or “search ahead x minutes”, to generate a probability surface of

where the user is likely to be based upon their previous spatial behaviour [9]. At present, these spatial search tools rely upon lists, often ranked by spatial proximity, or 2D planar maps for the presentation of retrieved results. The mobile VR display offers an egocentric approach for the display of these georeferenced search results, and may provide a more intuitive way of communicating the geographic component of relevance associated with individual search results.

Simultaneously, there is a great deal of research concerned with assigning spatial footprints to the wealth of existing Internet resources [2], and at least one commercial implementation of a search engine has been launched that can match spatial identifiers in a query to the locations of businesses and services in the real world, based upon their online presence alone [1]. For mobile users, a 3D display offers an alternative approach to displaying and selecting results than the traditional map- and list-based paradigms [16].

4. Conclusions

Mobile handheld devices provide a new computing environment where your location is a key factor in determining your information needs [8]. Research suggests that there are key differences in how mobile users interact with their devices compared to users of desktop systems, proposing that traditional interaction paradigms may not be appropriate in this new context. VR models on desktop machines exist in a co-ordinate space that is outside of the real environment coordinate system. AR attempts to bridge the real-virtual divide, however the hardware remains cumbersome and processing intensive. Mobile VR offers an environment for displaying the increasingly volume of geographic information available, such as maps, directions, and georeferenced Internet search results. This paper proposes that external sensors may allow a more intuitive approach for interaction with these mobile VR models. Development and evaluation conducted during the duration of the locus project aims to test this hypothesis.

5. ACKNOWLEDGMENTS

The authors would like to express their gratitude to PinPoint Faraday partnership for the funding of this project, to the GeoInformation Group for the use of building heights data, and to the Department of Information staff at City University for their inspiration and support.

6. REFERENCES

[1] Google (2005) Google local <http://local.google.co.uk/> (last accessed 28 Apr 2005)

[2] Jones, C. B., Purves, R., Ruas, A., Sanderson, M., Sester, M., van Kreveld, M. and Weibel, R. (2002) Spatial information retrieval and geographical ontologies. An overview of the SPIRIT project. SIGIR'02.

[3] Kallio, S., Kela, J. and Mantyjarvi, J. (2003) Online Gesture Recognition System for Mobile Interaction. Ieee International Conference on Systems Man and Cybernetics, 3, 2070-2076.

[4] Klatzky, R. L. (1998) Allocentric and egocentric spatial representations: Definitions, distinctions, and interconnections. Lecture Notes in Computer Science, 1-18.

[5] Kulju, M. and Kaasinen, E. (2002) Route guidance using a 3D city model on a mobile device. Mobile HCI02 Workshop: Mobile Tourism Support Systems.

[6] The LOCUS Project (2005), <http://www.locus.org.uk/>, (last accessed 03 May 2005).

[7] 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, 282-292.

[8] Mountain, D. and Raper, J. (2002) Location-based services in remote areas. AGI conference at GeoSolutions 2002, B05.3.

[9] Mountain, D. M., Goncalves, A. and Rodrigues, A. (2003) Geographic knowledge discovery and spatial agent interfaces for location-based services. GIS Research UK, 11, 206-209.

[10] Nothegger, C., Winter, S. and Raubal, M. (2004) Selection of Salient Features for Route Directions. Spatial Cognition and Computation, 4, 2, 113-136.

[11] Ostrem, J. (2002) Palm OS user interface guidelines http://www.palmos.com/dev/support/docs/ui/uiguide_front.html (last accessed 29 Jan 2003).

[12] Parallel graphics (2005) Pocket Cortona <http://www.parallelgraphics.com/products/cortonace/> (last accessed 28 April 2005)

[13] Pharos Science and Applications (2005) Pharos GPS <http://www.pharosgps.com/> (last accessed 27 April 2005)

[14] Pinpoint Faraday (2005) Time and Position solutions <http://www.pinpoint-faraday.org.uk/> (last accessed 27 April 2005)

[15] PNI Corporation (2005) 3 Axis Orientation Sensors <http://www.pnicorp.com/family?nodeId=c1a> (last accessed 27 April 2005)

[16] Reitmayr, G. and Schmalstieg, D. (2004) Collaborative Augmented Reality for Outdoor Navigation and Information Browsing. Symposium Location Based Services and TeleCartograph

[17] Ribo, M., Pinz, A. and Fuhrmann, A. L. (2001) A New Optical Tracking System for Virtual and Augmented Reality Applications. Instrumentation and measurement technology conference, 1932-1936.

[18] Web3D Consortium (2005) GeoVRML <http://www.geovrml.org/> (last accessed 27 April 2005)

[19] Wood, J. D. (1999) The Geomorphological Characterisation of Digital Elevation Models. PhD Thesis. Department of Geography. University of Leicester

[20] Yell (2005) The UK's local search engine <http://www.yell.com> (last accessed 27 April 2005)