NEW APPROACHES FOR MIXED REALITY IN URBAN ENVIRONMENTS: THE CINESPACE PROJECT

Mans Shapshak
Project Manager. Tracasa S.A.
mshapshak@tracasa.es

Software Engineering and Information Technology Department. Tracasa S.A.
C/ Cabárceno, 6. 3162, Sarriguren, Spain
Teléfono + 34 948 289 000
Fax + 34 948 249 209

Key words: Mixed Reality, Human-computer interaction (HCI), Displays, Sensors

Abstract

The CINeSPACE (www.cinespace.eu) project allows tourists to access the rich cultural heritage of urban environments by literally morphing the user into the past through the use of multimedia archives. Tourists use the device which includes both a PDA type of device with a GIS interface displayed on a touch screen to help the user navigate and select multimedia content, and video binoculars to create the augmented reality effects. In addition to this mode of interaction, a survey of Mixed Reality user interaction paradigms will be presented. A key feature of Mixed Reality user interfaces is the object identification and annotation methods available to the user, of which a survey, including a review of the GeoConcepts ontology annotation methodology used in the CINeSPACE device, will be presented.

1. Introduction

The use of Mixed Reality (MR) in Urban Spaces has been a popular area of interest and many reviews of the state of the art can be found in the literature (Azuma, 1997; Costanza, Kunz, & Fjeld, 2009). A commonly used and very inclusive definition of MR is that of all applications between pure Virtual Reality and the real world (Milgram & Kishino, 1994). This definition includes applications where the real environment is mixed with a virtual world. A narrower domain might be considered to be visual output MR applications which are focused on mixing the real and the virtual world specifically in the visual output device (computer screen, near the eye displays, etc.) as opposed to the mixing of audio or tactile real and virtual worlds. Another commonly accepted but not necessary restraint is that the virtual world model be 3 dimensional in nature. For example common automotive navigation Geographic Information System (GIS) has a nominal model of a 2D virtual world projected at an oblique angle while projecting the user location.

Using the broad definition of MR it becomes clear that there is an enormous amount of modes possible for user interaction. In the last few years there has been a leap in the number of user interaction paradigms being used in the laboratory and in commercial applications. In addition, many of these new and novel MR applications are characterized by the blurring of input and output devices.

The blurring of input and output in the same interface is typical of the very fluid interaction which exists with real world items and but much less so in a traditional computer world. For example in the real world we push a cube of metal and it pushes back whereas in the virtual world we press a key on the keyboard and the cube moves on the computer screen. In a MR environment these concepts of input and output become more closely linked. In MR the user interaction can be achieved through haptic devices where the position of a finger is tracked in 3 dimensions and then a haptic feedback device applies the appropriate reverse force. In a simpler mode of interaction the motion of the finger can be traced by various cameras the results can be shown in the virtual world or a virtual world projection on the real world. Another mode of fluid interaction is characterized by computer output being incorporated into items themselves (Merrill, Kalanithi, & Maes, 2007). In summary we find that the basic interaction paradigm is the physical action of input (mouse, keyboard, hand motion or voice) with visual and/or audio output. The physical output of force and real motion are becoming more pervasive as computers become embedded in everyday objects.

2. The CINeSPACE project

The CINeSPACE project has developed a MR device which tourists visiting cities will be able to rent. The novel design, shown in the figure 1 below, includes both a PDA type of device with a GIS interface displayed on a touch screen to help the user navigate and select multimedia content, and video binoculars to create the augmented reality effects. This mode change is typical of MR applications and allows the user to move from a GIS navigation system with touch screen interaction to a near the eye display with 3 dimensional MR as the device orientation is changed.

Figure 1. Image of a user with the CINeSPACE device in San Sebastian

Source: CINeSPACE

The user position and orientation is calculated from a module that has been developed by the Fraunhofer IGD group in Darmstadt, Germany and uses four different technologies to allow maximum positioning accuracy. GPS is used to establish the position within approximately 10 meters. Then marker-less optical tracking is used to establish a more defined location and orientation. Marker-less tracking is based on standard computer vision algorithms (Lowe, 1999; Zhang, 1999). For quick motions inertial tracking is used by mathematically integrating the equations of motion to find the current position from acceleration data provided by integrated circuit accelerometers.

Once the user position and viewpoint orientation has been found the device is able to carry out MR functionality such as the replay of historical film clips superimposed on the real scene being viewed, including morphing effects, creating a sensation something like daydreaming.

The particular challenge presented in MR is the need to allow the user to access a wide variety of geo-referenced information in a visual near the eye display and annotate that information. To address these needs a novel semantic GIS technology with its MR interface has been developed called GeoConcepts ontology (Ruiz, Shapshak, Achaerandio, Villadangos, & Clerigué, 2008).

2.1 The Semantic GeoConcepts Ontology and its use for Annotation Functionality in MR

The GeoConcepts ontology provides a way to capture in a GIS system the spatial dimension of concepts, defining whether they are traditional geographic features (street, building or town) or vague concepts (events, daily activities, historical milestones or personalities) common to geospatial thinking. In order to initially populate this GeoConcepts ontology, data are collected from social websites such as Panoramia or Flickr, where user multimedia content is normally stored in a geo-encoded format.

Polygons are calculated that include clustered data points and reject data points which are outside of chosen statistical tolerance limits to eliminate erroneous or badly tagged photos. The resulting ontology is shown in the figure below for a set of GeoConcepts for the city of San Sebastian. Both normal GIS concepts like City Hall (Ayuntamiento in Spanish) and conceptual objects like fireworks (fuegos in Spanish) are shown. Also of interest are those concepts with an implied interrelationship to other polygons, such as "nice view of Santa Clara."

Figure 2. A Google Earth image of San Sebastian with superimposed GeoConcepts

Source: Ruiz, J. Google

With the GeoConcepts ontology loaded into the CINeSPACE device, the simple act of looking in a direction through the video binoculars provides the user with feedback as to the GeoConcepts at which he is looking by overlaying the selected GeoConcept 3D polygon representation in MR. This functionality is used for video and image annotation and for content selection in the CINeSPACE device.

3. A survey of MR user interaction paradigms relevant to the CINeSPACE device

New interaction paradigms or functionalities being developed for MR applications are closely linked to the developments of new technologies of input and output devices. In the sections below we survey some at the MR applications, use cases or functionalities and devices.

3.1 The MR Toolbox

We present a list of some of the most common technologies used in MR for user output and input in tables 1 and 2 we then continue to present the advances in the mixing of various devices leading to a fluid interaction between input and output in table 3.

Table 1. A list of commonly used displays in MR Technologies

Description	*Input*	*Output*	*CINeSPACE*
Microdisplay Near the Eye Optics and Display		X	X
Touch Screen,	X	X	X
Muti Touch Screen	X	X	
MicroProjector		X	
Hologram		X	
Immersive video environment		X	
Embeded Small Screens		X	

Table 2. A list of commonly used sensors in MR Technologies

Description	*CINeSPACE*
GPS	X
Accelerometer(s)	X
Digital compass	X
Marker-less optical tracking	X
Marker based optical tracking	
Eye tracking near the eye	
Eye tracking Up to 10m	
Gesture recognition	
Gesture recognition with 3D marker	
Subsonic beacons placed in environment	
High accuracy short distance < 1m 6 degree of freedom pulsed DC magnetic field sensor	
RFID or WiFi beacons place in environment	
RFID tags in real objects linked in wireless network	

Table 3. Advanced Combinations of MR Technique

Name	*Input*	*Output*	*CINeSPACE*
Gesture Recognition without 3d marker and personal micro projector display (Mistry, 2009).	X	X	
Embedded small screens on blocks with edge sensitive interactions (Merrill D., 2007).	X	X	
Physical orientation sensor and display screen orientation	X	X	X
Using device orientation and 3D scene as a item selector	X	X	X

3.2 Functionality with particular relevance to MR

Virtual Reality offers fantastic possibilities of navigating, altering and managing the virtual world completely breaking the laws of physics and reality. However, Virtual Reality worlds have in many instances been deemed to be devoid of the rich details and lacking of meaningful content. MR on the other hand offers the user the functionality of Virtual Reality but blends it with the detail of real world content. Some of the functionalities which have shown most interest are:

- Adding Virtual Reality content to a real world:
- Moving content and mapping it in three dimensions from another time and/or space to now and close by
- Morphing into mapped content
- Annotation of objects in three dimensions on a two dimensional interface
- Zoom and analysis of detail in the three dimensions
- Non destructive experimentation interaction with the environment

These abilities have shown to provide value in many domains. Over the last few years advances in new devices and technologies relevant to MR have further reduced the barriers to entry of MR into new fields. Of particular note are advances in software algorithms in computer vision and the development of low cost display and tactile devices.

One of the most common user interface paradigms in the above functionalities is the “magic lens” paradigm. In this mode a device with a camera and video display is held up in front of the user and images are superimposed in 3D onto the video image. Since the camera and the display are mounted in the same device, most typically a mobile phone or PDA, the rotation of the device automatically rotates the real world image on the display. The problem now becomes to position the computer generated output onto this scene. Marker-less or marker based optical tracking techniques are designed to work with rotational freedom and hence the MR content is easily overlaid. If marker-less or marker based optical tracking cannot be used other techniques must be used to provide the information of the device location and orientation in respect to the computer generated content. One solution is fixing the location to a particular point and measuring the device orientation. This is the method used with MR telescopes solutions. For a mobile device like CINeSPACE, extra sensors like GPS or RFID in combination with accelerometers and magnetic compass are used. The implementation of a "magic lens" has become straightforward to implement with the development of software libraries to do marker based optical tracking. Nonetheless since markers are not available in the urban environment the implementation is somewhat more complex. Currently the availability highly popularized open source accelerometer projects (Lee, 2007) need to be complemented with a simple and inexpensive electric compass. This development is to be expected in the next few years in consumer electronics which already have incorporated accelerometers making them suitable for "magic lens" applications without the need for markers in neither the environment nor the computational and reference image calibration of marker-less optical tracking.

Another paradigm used in the design and study of Urban Spaces with Mixed Reality is the "model camera". In this mode a 3D icon of a camera is moved around in a real model space as a sort of MR probe on a workbench (Kienzl (n.d.); Wang, 2007; Wang & Schnabel, 2008) The user positions the camera with the current orientation in respect to objects normally on a MR workbench. Ideally the MR workbench also incorporates images into displays and sensor networks into the workbench objects themselves so that they can be configured to provide a more fluid interaction with the user. The relevance to CINeSPACE is that once the "model camera" has been moved to the appropriate location the binocular device can then be used to look around. The two modes are interchanged as if in the in situ mode. The physical act of walking around is replaced by moving the "model camera". A benefit for urban planners and architects is that with a workbench is that the urban landscape can be reconfigured. It may be noted that in this interaction paradigm there is a reversal of the typical MR scenario. The real world in this case is a physical model of the urban landscape.

3.3 MR Application Domains of Particular Interest

The uses of the functionalities mentioned in the previous section have been put to use in a wide array of applications from many domains including games, military and law enforcement, emergency services, medical surgery, airport and port control towers, advertising, tourism and navigation. Below we discuss some of these applications. The use of MR in game consoles has gained much interest with the widespread popularity of game consoles with gesture based and accelerometer input devices. On the other hand use of MR in outdoor games have gained limited acceptance. Some interesting examples include A-rage (A-rage web site) and Can You see Me Now (Tandavantij, et al., 2006).

Augmented Reality technology has long been used in the military applications and to a lesser degree in law enforcement. But here the combination of binoculars with context information similar to the CINeSPACE form factor could provide a viable product offering for users who are typically very sensitive to changes in form factor because of limitations in their work pattern. As an example PDAs have had poor acceptance in police forces because it is difficult for officers to use PDAs in urban and dangerous environments. On the other hand car mounted computer terminals have had excellent acceptance. Binoculars, like the CineSpace unit, have an accepted work pattern in law enforcement and military use. Currently only simple information about distance and heading are typical in military binocular devices but the availability of better user interfaces may speed the development of more sophisticated MR solutions.

MR technology is used by Fire Departments in training and actual incident response. For training there are systems which let firemen practice using simulated fire and smoke. The use of head mounted displays (HMDs) is being studied for the use in urban environments (Wilson, et al., 2005). MR can also be very relevant in a rural setting where Fire Fighters can identify and get important information about the risks, flammability of geographical from a binocular system.

In architecture complete Virtual Reality has been used frequently for virtual visits of designs while MR applications are gaining interest as they offer the ability to visualize new buildings with the "magic lens" paradigm implemented in a telescope(Stricker) or mobile device in the archaeological context ancient buildings (Dahne & Karigiannis, 2002) at a site. Additionally the MR workbench with the "model camera" paradigm has been evaluated for its use in urban design and architecture. MR applications in television broadcasting, navigation, medicine have achieved mature market position with specific devices and interaction paradigms.

4. Conclusions and further work

As the field of MR is of active interest in the academic and commercial sectors many new tools and functionalities are being introduced. A wide variety of MR platforms are gathering attention from developers and many novel interaction paradigms are being developed with a growing number of MR compatible input devices. The MR applications are becoming easier to create and deploy on smaller devices such as mobile phones and PDAs. Previous limitations of not being able to integrated sophisticated sensors, lack of powerful CPUs or limited wireless access to distributed databases are quickly being removed accelerating the pace of development of MR applications.

Of particular interest to the user interface paradigm used in the CINeSPACE device are series of improvements including the ability to navigate to other positions than the actual user position for content viewing. Once in this new position the functions of zoom and pan would provide an excellent way to discover elements of the urban environment from this new vantage point. Semantic Ontologies provide a rich level of detail in relationships between objects but the challenge of creating a simple and intuitive semantic navigation system in a MR system so far has been elusive (Shapshak, 2008) and further work should be done. While the CINeSPACE project has placed a large emphasis in user evaluations the vast majority of research projects have not evaluated user reaction to MR user interaction paradigms. Work needs to continue in this area to evaluate the combinations of technologies, MR functionalities and application domains that are most useful to the user.

5. Acknowledgements

This work is partially funded by the European Commission within the Sixth Framework Programme, as part of the project CINeSPACE, Experiencing urban film and cultural heritage while on-the move.

Bibliography

A-rage web site. (n.d.). Retrieved 1 10, 2008, from http://www.a-rage.com/

Azuma, R. (1997). A survey of augmented reality. Presence: Teleoperators and Virtual , 355–385.

Cobos, ,. Y., Sarasua, ,. C., Linaza, ,. M., Jimenez, ,. I., & Garcia, ,. A. (2008). Retrieving Film Heritage Content Using an MPEG-7 Compliant Ontology. (pp. 63--68). IEEE Computer Society.

Costanza, E., Kunz, A., & Fjeld, M. (2009). Mixed Reality: A Survey. In Human Machine interaction: Research Results of the MMI Program. In E. D. Lalanne and J. Kohlas, Lecture Notes In Computer Science, vol. 5440 (pp. 47-68). Berlin, Heidelberg: Springer-Verlag.

Dahne, ,. P., & Karigiannis, ,. J. (2002). Archeoguide: System Architecture of a Mobile Outdoor Augmented Reality System. (pág. 263). IEEE Computer Society.

Fritz, F., Susperregui, A., & Linaza, M. T. (2005). Enhancing Cultural Tourism experiences with Augmented Reality Technologies. The 6th International Symposium on Virtual Reality, Archaeology and Cultural Heritage VAST.

Kienzl, T. (n.d.). The MRI (Mixed Reality Interface). Retrieved 5 20, 2009, from http://www.kommerz.at/en/Produkte/MRI.html

Lee, J. C. (2007). Johnny Chung Lee Wii Project. Retrieved 5 15, 2009, from http://johnnylee.net/projects/wii/

Lowe, D. G. (1999). Object recognition from local scale-invariant features. Proceedings of the International Conference on Computer Vision 2, (pp. 1150–1157).

Merrill, D., Kalanithi, J., & Maes, P. (2007). Siftables: Towards Sensor Network User Interfaces. In the Proceedings of the First International Conference on Tangible and Embedded Interaction (TEI'07). Baton Rouge, Louisiana, USA.

Milgram, P., & Kishino, F. (1994). A Taxonomy of Mixed Reality Visual Displays. IEICE. IEICE Transactions on Information Systems 77 , 1321–1329.

Mistry, P., Maes, P., & Chang, L. (2009). A Wearable Gestural Interface. WUW - Wear Ur World - In the CHI '09 extended abstracts on Human factors in computing systems . Boston,, USA.

Ruiz, J., Shapshak, M., Achaerandio, J., Villadangos, J., & Clerigué, R. (2008). A Semantic Framework for Dealing with Human Geospatial Reasoning on Geo-Reference Data. 5th Int. Conference on Geographic Systems, ICGIS. Istanbul.

Shapshak, M. (2008, 6 3). GeoSpace. Retrieved 5 1, 2009, from http://manskj.blogspot.com/2008/06/wireframes-semantic-interfaces.html

Stricker, D. (n.d.). AR Telescope XC-01. Retrieved 5 10, 2009, from http://www.xc-01.de/english/index.html

Tandavantij, N., Flintham, M., Crabtree, A., Adams, M., Row-Farr, J., Iddon, J., et al. (2006). Can You See Me Now? ACM Transactions on Computer-Human Interaction (TOCHI) , 100 - 133.

Wang, X. (2007). Specifying Augmented Virtuality Systems for Creative Architectural Design. Proceedings of the 11th International Conference on Information Visualization.

Wang, X., & Schnabel, M. A. (2008). Mixed Reality in Architecture, Design, and Construction. Springer.

Wilson, J., Steingart, D., Romero, R., Reynolds, J., Mellers, E., Redfern, A., et al. (2005). Design of monocular head-mounted displays for increased indoor firefighting safety and efficiency. Helmet- and Head-Mounted Displays X: Technologies and Applications. 5800, pp. 103--114. SPIE.