THE GEOSCOPE – A MIXED- REALITY SYSTEM FOR PLANNING AND PUBLIC PARTICIPATION

CLAUS BRENNER, Volker Paelke, Jan Haunert, Nora Ripperda

Institute of Cartography and Geoinformatics, University of Hannover
Germany

ABSTRACT

The augmentation of real environments with additional computer generated information (augmented reality, AR), and the addition of virtual contents (mixed reality, MR) are promising interaction paradigms for the support of applications outside the classical office scenario. While the potential of AR for applications such as geo-visualization has been shown using a number of demonstration systems, there are virtually no systems in daily use up to now. This is mainly due to technical reasons. Besides suitable displays, there is a lack of reliable and accurate tracking systems, which are however required for an exact registration of the virtual contents and the real scene. Additionally, there is a lack of experience and tools for the creation of mixed reality applications and often, of intuitive interaction techniques. Especially for applications involving public audience (e.g. exhibitions, museums, public participation), additional requirements apply, such as reliability, robustness, easy adaptation to different users, and the necessity for relatively low operating expenses.

In this paper, we introduce the GeoScope, a mixed-reality input/output device, which addresses those problems, especially for public applications. The GeoScope is mounted at a fixed position and consists of a display, oriented towards the user, and a camera, which points at the surroundings. Just as a telescope, the GeoScope can be turned around two axes, the two angles being captured by measuring devices. Together with the known position, a fast and highly precise tracking of the current view direction is possible, allowing the superposition of the real scene, as delivered by the camera, and virtually generated information.

INTRODUCTION

According to *Azuma (1997)* an augmented reality system is characterized by the following features:

- Virtual and real content are mixed within the real environment of the user.
- The system is interactive and operates in real-time.
- The virtual content elements are spatially registered in three dimensions.

This definition is not limited to graphical augmentation – even though this is currently the main output for augmented reality applications – but also covers other channels of perception (e.g. audio and haptics). The idea of removing parts of the real environment (often referred to as mediated or diminished reality) is also covered by this definition.

To implement a system that conforms to this criteria a number of hardware and software components are required. Figure 1 illustrates the basic components of such an augmented reality system. A central requirement is an appropriate model of the real environment that is to be augmented in order to associate virtual augmentation information with real world objects and locations. Within an AR system this augmentation is typically implemented through a spatially correct superimposition of computer generated visuals corresponding to the augmentation information over the real environment. To achieve such a spatial superimposition requires suitable sensors for fast and accurate positioning to ensure the correct registration of the virtual content with the real environment as well as suitable displays that support the visual combination of real environment and virtual content. Within the AR application the virtual content must be provided in a suitable format that can be rendered on demand, e.g. in the form of 3D graphics or textual labels. For interactive applications additional sensors and interaction techniques are required that enable users to influence and manipulate the augmented virtual content.

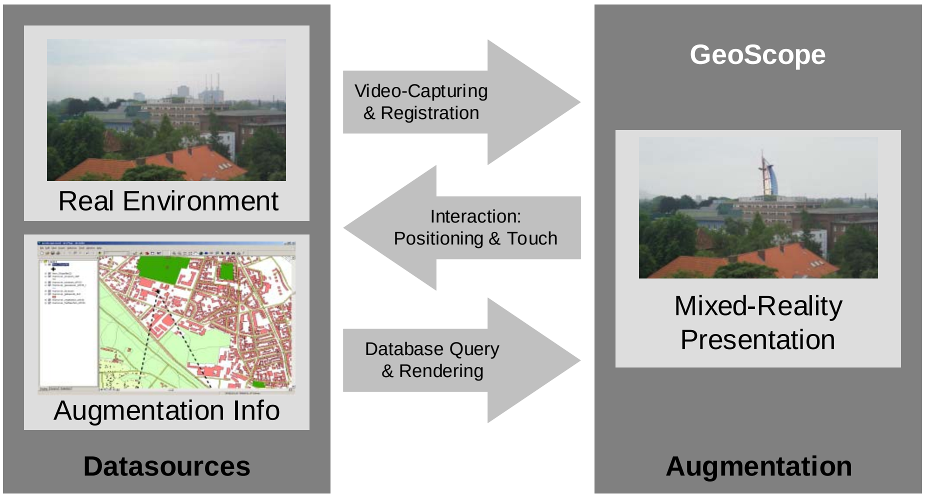

Figure 1: Basic components of an AR System as realized in the GeoScope.

The selection of a display suitable for the combination of virtual and real content requires a careful trade-off between costs, usability and the acceptance with the intended user group. One key criteria is the principle used to combine virtual and real content. Display devices can be distinguished according to whether they combine the information by (various) visual means or through video processing. Visual combination in the eye of the user is used in retinal scanning displays. These create crisp output and maintain the full field of view of the user but suffers from a lack of acceptance and the limited resolution of the available displays. A combination within the real environment (e.g. using projectors) can achieve high visual quality in a suitable environment and can be built with commodity hardware. However, such an approach is unsuitable for many outdoor and mobile applications. Optical augmentation, e.g. using semitransparent displays or mirrors as in see-through head-mounted-displays (HMD), maintain the optical impression of the real environment and are suitable for mobile use. However, only very few displays of this type are currently available and the requirements for low latency real-time and high precision registration are difficult to meet with current sensor technology in unprepared environments. Due to these limitations the combination of real and virtual information through video processing – video-see-through augmentation – in which a real-time video-stream of the environment is modified with virtual content is currently the most popular visual combination technology. The main benefits of this approach are the ease of practical implementation, the relatively low cost using commodity hardware components, the possible use in mobile settings and the reduced registration requirements compared to optical-see-through solutions. Key limitations of video-see-through solutions are the reduced optical resolution of the "real" environment (limited to the video resolution) and the restricted field of view.

Similar trade-offs apply to the technologies available for registering virtual information spatially within the real environment. Reliability, precision, resolution, update rate, latency and the requirements for a supporting infrastructure are key criteria that need to be considered. Positioning technologies based on the time-of-flight principle cover a wide range of possible usage scenarios (e.g. ultra-sonic trackers for indoor use or (d)GPS for outdoor applications) but require an elaborate infrastructure, are constrained to limited usage environments (ultrasound) or do not provide the required update frequency and precision for correct augmentation (GPS). Optical tracking approaches are used in a wide range of AR applications but are also subject to a number of critical constraints. Mechanical positioning techniques support fast and precise positioning but are not applicable in classical AR systems due to the need for mechanical linkages.

PREVIOUS WORK

The task of developing suitable displays and positioning technogies for augmented and virtual reality applications was first addressed by Ivan Sutherland in 1965 *(Sutherland, 1965)*. Sutherland proposed the concept of the "ultimate display" in which virtual content should be indistinguishable from real objects. In 1968 this concept was implemented in a first hardware prototype of an optical-see-through head-mounted-display (HMD) with mechanical tracking *(Sutherland, 1968)* because other techniques proved unsuited to address the requirements.

Because even today the available tracking and positioning techniques are still unsuitable for a wide variety of augmented reality applications mechanical tracking continues to be used in a number of display devices. Well known examples include the Boom (Binocular Omni-Orientation Monitor) device from Fakespace *(Fakespace, 2005)* and the Window VR system from Virtual Research *(Virtual research, 2005)* that are both aimed at virtual reality applications without integration into a real environment. The question of how augmented reality applications could be made accessible to the general public was first addressed in the AR-PDA project *(AR-PDA, 2005)* that explored the use of PDAs and Smartphones as augmented reality platforms and spawned a wide variety of current developments in hand-held AR systems. While the use of the PDA touchscreen as an interaction modality proved very useful in the AR PDA project, the size and resolution of the available devices as well as the limited processing power restrict the use of mobile devices to relatively simple AR applications. An approach that is closely related to the GeoScope presented in this paper is the Trivisio AR-Telescope XC-01, that also employs mechanical tracking in two degrees of freedom in combination with video-see-through AR *(XC01, 2005)*. Key differences to our approach are the use of a microdisplay in the XC-01 that restricts it to single user operation and excludes the use of advanced interaction techniques as in the GeoScope. The XC-01 is optimized for robustness in order to operate as a replacement for coin operated binoculars, where the restriction to a single user is desirable for economic reasons and vandalism prevention is a key priority.

DESIGN OF THE GEOSCOPE

The goal behind the development of the GeoScope was to provide a system platform for MR applications for the public such as for exhibitions, museums, vantage points, and public participation. Besides the basic requirements for display and tracking, the device has to be reliable, robust, inexpensive and it should be easy to switch between users. Additionally, easy interaction possibilities were required, since for MR applications there is so far no broad design experience, no standardised interaction techniques or guidelines exist and no user experience can be assumed with the intended audience. The technical requirements were:

- The display should have a high resolution and a good colour reproduction, similar to desktop displays. Since the intended applications require larger distances, no stereoscopic display is necessary.
- A highly precise and low-latency tracking is required as a prerequisite for a correct overlay of the augmented information.
- Besides content selection by positioning, additional possibilities for the interaction with augmented contents is desirable. Interaction should be easy to understand for occasional users without training.
- Dependency on external infrastructure should be low. Size and weight should allow for easy transport.

Figure 2: Snapshot of the GeoScope prototype, front side (left), back side (upper right), compact two axes head with angular measurement units (lower right).

Figure 2 shows the GeoScope prototype which has been developed on the basis of those requirements. The GeoScope is installed at a fixed location. It has a display headed towards the user and a camera headed towards the environment. Like a traditional telescope, it can be turned around two axes. Since the turn angles are captured by angular measurement units, together with the fixed position of the GeoScope, all six orientation parameters are known which allows to augment the video image with virtual contents.

Special consideration has been devoted to the measurement of the horizontal and vertical angles. The required resolution depends on the application scenario: a typical telescope application has a large focal length and thus angular measurement resolution must be higher as compared to a wide angle close range application (as in a museum, for example). In order to derive the necessary angular measurement resolution, it is assumed that one pixel of the camera corresponds to one pixel of the display. Thus, one can compute the angle which corresponds to one screen pixel. Figure 3 shows the aperture angle, angle corresponding to one screen pixel, and required digitisation resolution of the angular measurement units, depending on the focal length. For simplicity, the focal length is given relative to a standard small format camera. As one can see from the table, a focal length of 300 mm requires a measurement resolution of nearly 16 bits.

Focal length [mm]	Aperture angle [grad]	Angle corr. to one pixel [mgrad]	Required resolution [bits]
24	81,9	93,3	12,1
50	44,0	44,8	13,1
200	11,4	11,2	15,1
300	7,6	7,5	15,7

Figure 3: Relationship between focal length, aperture angle and required angle measurement resolution.

Absolute shaft encoders can be used to measure the horizontal and vertical angle. Compared to incremental encoders, the advantage is that measurement commences immediately after power-up, with no need to turn the device initially to a predefined null position. Also, they are very robust and long term stable. However, being precision devices, they are relatively expensive (17 bits resolution in the order of 600 €). Thus, we have explored the use of potentiometers in the GeoScope prototype which are digitised by A/D converters. For the potentiometers, we used industrial types with a conductive plastic resistance element, protection class IP65, a temperature range of –40 to +150 degrees Celsius, and 50 million moves, making them suitable for usage in outside conditions. The effective resolution of our A/D conversion circuitry is around 20 bits, so that the limiting factors are the linearity and the (long term- and temperature-) drift. The results of the A/D conversion are sent with a rate of 25 Hertz over USB. Altogether, all input and output signals of the GeoScope can be directly interfaced to a standard PC or laptop.

POSSIBILITIES FOR USER INTERACTION

The GeoScope supports interaction techniques for the standard tasks selection, manipulation and navigation. The augmentation information can be navigated intuitively by physically pointing the GeoScope in the desired direction. The two corresponding degrees of freedom are captured with high precision and minimal latency so that a fast and precise movement becomes possible. This mode of interaction is so intuitive that it is typically not noted by users as an explicit interaction. For manipulation and selection the touchscreen enables direct access to the objects on screen. More complex interaction techniques can be implemented by means of standard controls/widgets like scroll-bars, sliders and control boxes that are familiar to PC users and can be operated just like in desktop or web applications. The benefit in comparison to the custom interaction techniques currently employed in most AR systems are the recourse to operating knowledge of the users (they don't have to learn new techniques) as well as the availability of a large selection of well designed interaction techniques for application developers and the ease of scaling interface designs. Restrictions are due to the limitation to a two dimensional input space. As with mouse based interaction techniques this is not a major restriction for the intended application domain of public applications where complex spatial interaction tasks are unlikely to form a large part of the interaction.

USAGE SCENARIOS FOR THE GEOSCOPE

One can envision many display and interaction techniques for the GeoScope. Simple examples are text labels, 2D icons, or 3D models which are superimposed in real time over the video image. In addition to MR contents, purely virtual data can be shown such as panoramic views, 3D scenes, 3D virtual flights, or 2D top views such as satellite or aerial images, as well as topographic or thematic maps. Instead of controlling directly pan and tilt, the two input angles of the GeoScope can also control pitch and yaw of a virtual flight or pan a 2D top view (see Figure 4 and Figure 5).

Opera house

The opera house was built in 1845-52 based on a plan drawn by Laves. Originally it served as the royal theatre, as the king considered the theatre in the Leineschloß too small. The new opera house is a classical style building with two large wings and a balcony with statues of famous poets and composers. The balcony used to be open so that visitors could drive straight to the entrance in their carriages. Nowadays they can use the stairs to the underground car park. Hannover's opera house was badly damaged in the Second World War and rebuilt in 1948.

Figure 4: Examples for simple display and interaction techniques: panoramic view with overlaid virtual labels (left), textual description of a landmark appearing after a label has been selected (middle), purely virtual flight controlled by the GeoScope's pan and tilt angle (right).

In our student group, we have identified a large number of possible usage scenarios, such as time travel with historic panoramic images, video sequences, or virtual animated 3D models built according to historic photographs; panoramic mountain views with labelled mountain names and superimposed hiking trails or ski slopes, as well as purely virtual flights over large mountain areas; sea shore applications with low-tide and high-tide visualizations, explanatory material, virtual flights to nearby islands, and ferry timetables; and mounting the GeoScope on cruise ships, using the ship's navigation, possibly with a modified GeoScope construction to compensate for the ship's roll.

The main application to be discussed here is, however, planning and public participation. Two main aspects are important in this case. Planning usually needs a solid geometric base, as can be provided by Geographic Information Systems. And, usually, a 3D model of the existing surroundings is needed. Those two aspects are discussed in the next two sections.

Figure 5: Example for an overlay of a real image and virtual contents. Left: original image, centre: masking, right: real image with overlaid virtual content.

MAP VIEW AND DATABASE QUERIES BASED ON GEOGRAPHIC INFORMATION SYSTEMS

In general, a Geographic Information System (GIS) is a computer-based system which enables the capture, administration, analysis and visualization of geographic information. A fundamental component of a GIS is a database management system, which provides the access to a spatial database. The GIS allows to query spatial features, their attributes and their relationships which are modelled in this database. Common GIS software products offer map views in which the cartographic data can be explored. Maps are built of different layers, which can be switched on or off to focus on certain topics *(Longley et al., 2001)*.

Offering such functionality is advantageous, especially for planning purposes. For example, navigating through a scene from a fixed standpoint does not suffice in all cases. Users probably want to get an idea about their surroundings, including objects which are not in sight. For this a classical 2D map will be much more helpful than having only the egocentric view which is offered by the GeoScope. On the other hand, offering only a map does not take full advantage of the GeoScope's functionalities. Therefore, our approach is to combine both techniques. In the simplest way, this can be done by showing a map of the surrounding in a separate frame which can be enlarged or hidden. This can be supplemented with standard GIS functionalities for the map exploration such as pan and zoom as well as selection of features and query of their attributes. Initially, the map will be centred on the standpoint and starting from this, the scene can be explored.

However, without interaction between the map view and the augmented camera view no benefit compared to a standalone map application would be gained. Thus, different features have been implemented that allow a better association between the contents of both views. At first, the map can be rotated by turning the GeoScope and thus it is oriented always in the viewing direction. Secondly, the current viewing frustum of the camera is always displayed on the map. For the further development it is aimed to link labels, that are shown in the camera view, with buildings or other features in the map. By this, the user can query information about features either by pointing on the label in the camera view or by selecting the feature in the map. The queried information such as additional images, opening times of buildings or texts will be retrieved from the database and displayed in a way which is adapted to the potential user.

For the implementation of the GIS functionalities we have chosen ArcObjects from ESRI, Inc. This software offers ample libraries for the management and display of spatial data. The provided interfaces are based on COM technology and thus the integration to custom applications is rather simple. ArcObjects can be used with licences for ArcGIS Desktop or ArcGIS Server, which are large and expensive GIS software products. However, for a custom application as needed for the GeoScope a cost-effective deployment is offered with the ArcEngine Developer Kit and the ArcEngine Runtime *(ESRI, 2006)*.

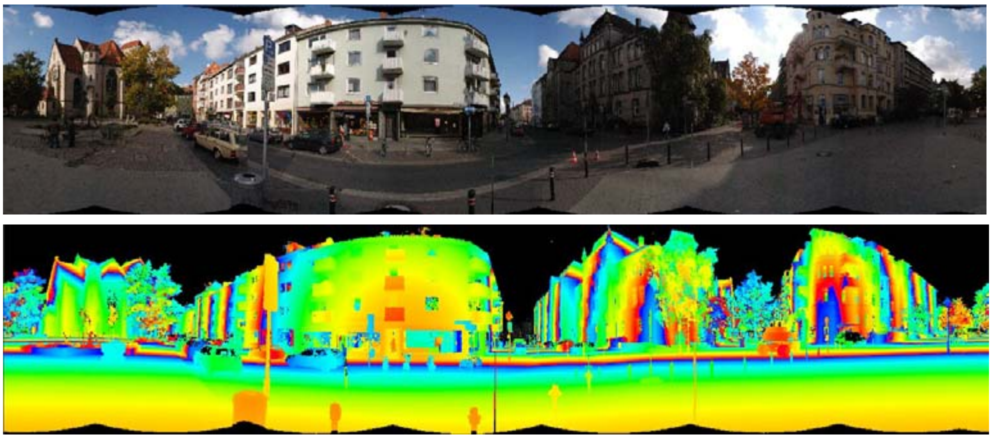

Figure 6: Panoramic and depth image, acquired with a laser scanner Riegl LMS Z360I. Measurements can be made in the 2D panoramic image but are readily available in full 3D due to the pixel synchronous depth image.

ACQUISITION AND MODELLING OF THREE DIMENSIONAL SURROUNDINGS

For modelling the 3D environment, existing data sources like aerial images, digital terrain models, two dimensional maps and 3D city models can be used. The acquisition of 3D city models has been a research topic for many years *(Brenner, 2005)*. Today, large 3D city models can be created using standard software tools *(Brenner & v. Gösseln, 2004)* although usually still a relatively large amount of manual interaction is required.

Especially for close range scenes, measurement of the surroundings is usually necessary since object shape in the required detail is most often not available. Small areas can be acquired using close range photogrammetry or terrestrial laser scanning. Both techniques usually require the use of several standpoints, although in the case of laser scanning this is only to obtain larger scenes or to remove occlusions, whereas close range photogrammetry, working according to the triangulation principle, needs at least two standpoints to recover 3D information. Since the GeoScope operates from a single standpoint, only the visibility of objects from this point is of interest so that a single terrestrial laser scan is sufficient. This eliminates the need for the acquisition and registration of multiple scans, which is usually the most time-consuming part during data acquisition. Also, single standpoint modelling does not need elaborated point cloud modelling but can be based directly on image and scan rasters as acquired by the terrestrial laser scanner.

Figure 6 shows a scan raster of a public place which contains pixel synchronous colour and depth information. It was acquired with a terrestrial scanner Riegl LMS Z360I, which has a measurement rate of 8,000 points per second, a range of 200 m and a single point accuracy of 12 mm *(Riegl, 2006)*. Especially useful in the context of the GeoScope is the scanner's ability to cover a full 360 x 90 degrees field of view and to directly map colour information from a mounted camera.

CONCLUSIONS AND OUTLOOK

The GeoScope supports many different applications involving public audience, including city-, landscape-, and architectural visualization, public participation and entertainment. Its main advantages are simple installation, high quality visualization, precise tracking, intuitive usage, support for advanced interaction using the touch screen and a relatively low price. By using a PC plattform, standard products can be used both for development and in the final end product. Some future developments are currently in the focus of our student group, including further evaluation of different presentation and interaction techniques and tools for the creation and management of mixed reality contents.

ACKNOWLEDGEMENTS

We thank our students (in alphabetical order) Xing Fang, Marion Horn, Julia Köpke, Juliane Mondzech, Michael Nierychlo, Matthias Roland, Marc Schmitz, and Dun Wu, who worked on the GeoScope topic in the course of a one year student project. Claus Brenner and Nora Ripperda are funded by the VolkswagenStiftung, Germany.

REFERENCES

AR-PDA (2005): http://www.ar-pda.de/, last access 31.01.2006.

Azuma, R (1997): A Survey of Augmented Reality. In: Presence: Teleoperators and Virtual Environments 6, 4, pp. 355 – 385.

Brenner, C. (2005): Building reconstruction from images and laser scanning, Int. Journal of Applied Earth Observation and Geoinformation, Theme Issue on "Data Quality in Earth Observation Techniques", **6**(3-4), March 2005, Elsevier, 187-198.

Brenner, C. and Gösseln, G. v. (2004): Tools and methods for the rapid acquisition of 3D city models, in: \`Proceedings of 24th Urban Data Management Symposium´, Chioggia (Venice), Italy, 2004.

ESRI (2006): ArcGIS for Developers Brochure, http://www.esri.com/library/brochures/pdfs/arcgis-for-developers.pdf, last access 31.01.2006.

Fakespace (2005): http://www.fakespace.com/, last access 31.01.2006.

Longley, P. A., Goodchild, M. F., Maguire, D. J., Rhind, D. W. (2001). Geographic Information Systems and Science. John Wiley & Sons, Chichester, England.

Riegl (2006): http://www.riegl.co.at/terrestrial_scanners/lms-z360i_/360i_all.htm, last access 31.01.2006.

Sutherland, I. (1965): The ultimate display. In Proc. of the IFIP Congress 2, pp. 506-508.

Sutherland, I. (1968): A head-mounted three-dimensional display. 1968 Fall Joint Computer Conference, AFIPS Conference Proceedings, 33, pp. 757-764.

Virtual Research (2005): http://www.virtualresearch.com/products.html, last access 31.01.2006.

XC01 (2005): http://www.xc-01.de, last access 31.01.2006.

CVS OF THE AUTHORS

Claus Brenner has obtained a doctoral degree of engineering from the University of Stuttgart, Germany, in 2000, his thesis being on “Three-Dimensional Building Reconstruction from Digital Surface Models and Ground Plans”. After having spent two years with Robert Bosch corporate research in the area of telematics, he is since 2002 with the Institute of Cartography and Geoinformatics, University of Hannover, where he heads a junior research group on data fusion in the area of geoinformation.

Volker Paelke is Juniorprofessor for 3D Geovisualization and Augmented Reality at the Institute of Cartography and Geoinformatics. Before joining the University of Hannover he worked as a research scientist at C-LAB, the joint Cooperative Computing & Communication Laboratory of the University of Paderborn and Siemens where he obtained a PhD in Computer Science with a thesis on the "Design of interactive 3D illustrations" in 2002.

Jan-Henrik Haunert was born in 1978 and studied Surveying at the University of Hannover. He graduated in 2003 and obtained the Master’s degree (Dipl.-Ing.). Since January 2004 he is a scientific assistant at the Institute of Cartography and Geoinformatics at the University of Hannover. In a project funded by the German Research Foundation he develops methods for the automatic updating process of topographic databases with multiple scales.

Nora Ripperda studied mathematics with specialisation in computer science at the University of Hanover and graduated in 2004 (Dipl.-Math.). Since October 2004 she works as a research assistant at the Institute of Cartography and Geoinformatics. She is member of the junior research group “Automatic methods for the fusion, reduction and consistent combination of complex heterogeneous geoinformation”, where she works on the automatic extraction of objects from laser scanner data.

CO-ORDINATES

Dr.-Ing. Claus Brenner, Prof. Volker Paelke, Dipl.-Ing. Jan Haunert, Dipl.-Math. Nora Ripperda
Institution : Institute of Cartography and Geoinformatics, University of Hannover
Address : Appelstr. 9a, Hannover
Postal Code : 30167
Country : Germany
Telephone number : +49 511 762 5076
Fax number : +49 511 762 2780
E-mail address : {claus.brenner,volker.paelke,jan.haunert,nora.ripperda}
@ikg.uni-hannover.de
Website : www.ikg.uni-hannover.de