Development of a Mixed Reality Device for Interactive On-Site Geo-visualization

Volker Paelke, Claus Brenner
Leibniz Universität Hannover

Abstract

This paper reports on the development of a novel mixed reality I/O device tailored to the requirements of interaction with geo-spatial data in the immediate environment of the user.
The I/O device is suitable for expert and casual users, integrates with existing applications using spatial data and can be used for a variety of applications that require geo-visualization including urban planning, public participation, large scale simulation, tourism, training and entertainment.

1 Motivation

Many applications deal with data that is based on geographic information. Sometimes the spatial nature of the data is obvious to the user as in car navigation systems, often it is an implicit aspect of it as in urban planning or large-scale simulations. Users require adequate tools in order to interpret, analyze and manipulate data. Visualization and interactivity are central features of most applications since visual representations support the effective interpretation of data and interactive manipulation can help users comprehend the possible effects of decisions. A key aspect of spatial (or geo-referenced) data is the requirement that it must be viewed, interpreted and manipulated in relation to a real physical environment. Typically, this spatial context is provided by a map.
While the use of abstracted maps is adequate for many tasks, alternative forms of presentation can help to make spatial data more accessible and simplify the interpretation. The success of Google Earth and similar 3D "Earth Viewers" hints at the potential of more concrete presentation formats. However, the completely virtual representation of the spatial environment remains very different from the corresponding physical surroundings. The user interface paradigm of "Mixed Reality" (MK94), in which virtual information objects are integrated seamlessly into a real environment may seem like the ideal solution but the use of mixed reality in real applications has been very limited so far due to technological limitations. Mixed reality displays with adequate quality and resolution are one central problem in this regard, spatial positioning techniques with sufficient resolution and precision are another. Most existing mixed reality systems are only demonstrators of new technologies. For the work reported here we have chosen to analyze the specific requirements of a central category of geo-visualization applications for which the benefit of a mixed reality interface seem especially relevant: Applications in which users view and interact with geo-spatial data in their immediate surroundings. Based on these requirements a novel mixed reality I/O device was developed.

2 Interaction with spatial data

2.1 Spatial interaction tasks

The selection of relevant data, the configuration of adequate data presentation techniques and the input or manipulation of data are central tasks in an interactive system. In conventional WIMP (windows, icons, menus, pointer) interfaces the tasks "position", "select", "quantify" and "text" are usually regarded as basic interaction tasks from which complex interactions can be composed. In systems that use geo-spatial data these tasks can be significantly more complex than their 2D counterparts, e.g. 3D positioning can require the control of up to 6 degrees of freedom and while 2D selection is easily implemented by clicking with the mouse, spatial selection can involve the specification of areas, volumes or individual objects in 3D space. As a first step in our design process we aimed to establish basic interaction tasks with geo-spatial data. The following categories were derived from interviews and questionnaires conducted among 18 GIS professionals and 11 students with advanced GIS knowledge. The following spatial interaction tasks were identified:

- „Identification": The identification of objects or areas
- „Information": The presentation of position specific information, e.g. the annotation of objects
- „Localization": The localization of objects or the user himself („Self-Localization")
- „Guidance": Finding locations or objects in space
- „Navigation": Selfpositioning of users in real or virtual space
- „Spatial Selection": Spatial selection of objects or areas
- „Spatial Positioning" and „Spatial Orientation": Spatial manipulation of objects up to six degrees of freedom
- „Data Collection": Acquisition of locations or objects with spatial information.

In addition to the interaction tasks themselves the spatial source and extend of the data is relevant in the selection or design of appropriate interaction techniques as discussed in the following section.

2.2 Scopes of interaction

Interaction tasks with real or virtual objects that correspond to spatial data can be distinguished according to where the objects are located relative to the user. For our study we established the following scopes of interaction:

- Local objects: Objects within the immediate range of the user that can be directly manipulated, e.g. real objects on a table or virtual objects on a local display.
- Local groups: Groups of object in the immediate range of the user that can be manipulated simultaneously.

- Workspace: The space surrounding the user that can be viewed from the user's current position, e.g. a real physical room, a panorama view or the current (virtual) workspace on a display.
- Environment: The environment surrounding the user that can be accessed by previous actions of the user, e.g. the building surrounding the user or the (virtual) workspace on a display that can be accessed by scrolling or self-positioning.
- World: The space surrounding the user that cannot be practically accessed by self-positioning.

The development of appropriate interaction techniques depends on the scope of interaction. For example, the concepts of so called tangible user interfaces (IU97), in which physical place-holders are manipulated to control virtual data appear to be well suited for handling "local objects" (and possibly "local groups"). They would be less suitable for interactions in "workspace" or "environment" scope since an indirect mapping from placeholder to manipulated object would be required that could break the underlying paradigm. However, mixed reality techniques appear well suited for interaction at "workspace" and "environment" scope because they can exploit the real environment to place the data in the correct spatial context. For "local" scopes mixed reality presentation techniques can be combined with tangible user interface concepts. The application of mixed reality or tangible user interface concepts to spatial data at "world" scope would be problematic since the data is neither in the immediate grasp nor view of the user. Here a completely virtual presentation, e.g. a 3D world-viewer concept like Google Earth, would seem most appropriate to handle presentation and interaction with aggregated "world" scope data at various levels of detail. In our development we have focused on the "workspace" and "environment" scope as most geo-spatial data is acquired and analyzed at these scopes.

2.3 Mixed reality interaction concepts for spatial tasks

Following the decision to focus on "workspace" and "environment" scope the next step was to identify how mixed reality could be used to create interaction techniques for the basic spatial interaction tasks at these scopes. Initial ideas were created at a brainstorming session with GIS experts and promising ideas were later followed up to various stages of prototype implementation as part of student projects. The following list provides a summary of promising uses of mixed reality to address these interaction tasks:

- „Identification“: Identification of given objects can be supported by visual highlighting and/or annotation. Spatially selected objects in the user's view can be identified by annotations.
- „Information“: Information on objects or locations can be provided by various means of annotation.
- „Localization“: Object localization can be supported by visual hints in the user's view (e.g. markers). Self-localization can be supported by direct output of text, coordinate or map information.
- „Guidance“: Users can be supported to find locations or objects in the surrounding space by visual guides (e.g. pointers).

- „Navigation“: User motion in the real environment can be used to control self-positioning in a virtual information space. Physical guidance information can be used to guide uses in self-positioning tasks in real space.
- „Spatial Selection“: Spatial selection of objects or areas can be simplified significantly by using the current position and viewing direction of the user. Based on the specific requirements the explicit interaction task could either be replaced completely by the intuitive self-positioning of the user or be used as a pre-selection technique so that detailed selection becomes a more simple selection task from a reduced set or volume.
- „Spatial Positioning“ and „Spatial Orientation“: Based on the techniques for spatial selection, spatial manipulation of objects can be simplified by using the user’s position and view to reduce the available degrees of freedom or directly control some of them.
- „Data Collection“: User position and orientation as well as spatial context data can be embedded in new data sets at acquisition time. For simple (position) acquisition tasks the interaction task can become completely transparent; for more complex acquisition the workload can be reduced.

Central prerequisites for the implementation of mixed reality techniques are appropriate base technologies for determining and tracking of the user’s view of the environment, the acquisition of spatial context information as well as means for displaying data spatially correct within the environment and means of interaction. The current state of the art of mixed reality technologies and the requirements are discussed in the following section.

3 Base technology

3.1 State of the art

The central elements required for a mixed reality system are displays suitable for the combination of virtual and real content and sensors that allow to register virtual information spatially within the physical environment in real time (Azu97).

The development of suitable technologies can be traced back to Ivan Sutherland who proposed the concept of the "ultimate display" (in which virtual content should be indistinguishable from real objects) in 1965 and implemented a first hardware prototype of an optical-see-through head-mounted-display (HMD) with mechanical tracking in 1968 (Sut68). While HMDs have seen mayor improvements in recent years commercially available examples still fall short of user expectations in display resolution, and field-of-view. Also HMDs face a huge acceptance problem in many applications and are unsuitable for fast switching between users. Therefore, researchers in mixed reality have experimented increasingly with handheld displays like PDAs and Tablet PCs in recent years. Use of PDAs in practical application is the limited by the small size and low resolution of the displays. Tablet PCs on the other hand have larger displays and higher resolution but are cumbersome to carry around.

The development of positioning technologies that are suitable for mixed reality is an ongoing challenge. Welch and Foxlin (WF02) provide an overview of available techniques. 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.

3.2 Requirements

To make geo-visualizations at “workspace” or “environment” scope accessible to large groups of non expert users a device which allows for easy switching between users is required. It should also enable the implementation of interaction techniques for spatial interaction tasks that are easy to use, intuitive and that do not require additional devices or huge amounts of training. Since visual presentation and analysis are a key objective in the targeted geo-visualization applications data display must be high quality and high resolution. Furthermore the device should provide a stable platform using existing technologies that can serve both as an evaluation platform for mixed reality presentation and interaction techniques as well as a possible starting point for future developments.
The key requirements for the system hardware can be summarized as follows:

- High resolution display, good colour reproduction, similar to desktop displays. Since the intended applications operate with objects at medium to large distance stereoscopic display is not necessary.
- High precision and low-latency tracking as a prerequisite for a correct overlay of the augmented information.
- Support for interaction beyond selection by positioning. Interaction should be easy to understand for occasional users without training.
- No dependency on external infrastructure.
- Size and weight suitable for easy transport.
- Based on reliable, robust, inexpensive technologies.
- Suitable for public applications, e.g. fairs, exhibitions, tourism or public participation.

4 The GeoScope I/O device

4.1 Concept

The GeoScope is a device that can be installed on a standard geodetic tripod at arbitrary locations. Its main components are a high resolution display with touch-screen facing the user and a camera with corresponding resolution that is mounted on the back of the display, looking into the environment. By augmenting the video-stream with computer generated graphics a video-see-through mixed reality setup can be realized. Similar to a telescope the GeoScope can be turned in two degrees of freedom (pitch and yaw). The rotation angles are captured with high resolution and precision by mechanical sensors. In combination with the position of the tripod (which can be determined by geodetic means since the same mount is used by the GeoScope) all position parameters of the GeoScope can be determined with high precision, allowing for spatially correct augmentation of the video image. Using geodetic calibration techniques for the sensors and the camera pixel-precise positioning of virtual overlays becomes possible. In addition to the high precision further advantages of the GeoScope in mixed reality applications are the lag and latency free sensor-data, the intuitive means of interaction, as well as the high-resolution display that can be viewed by several users simultaneously.

4.2 Hardware design

Figure 1 shows the GeoScope hardware prototype. It consists of a standard LCD panel with touchscreen, corresponding driver hardware and an AVT Guppy camera with IEEE 1394 interface that employs standard C-mount lenses. Special consideration has been devoted to the measurement of the horizontal and vertical orientation angles that can be easily changed by rotating the device with the handles on both sides. To enable pixel-precise positioning of overlays adequate angular resolution for the selected focal length of the lens must be provided. In a "telescope" with a long focal range a higher resolution is required, compared to normal views or wide-angle views. For pixel-precise positioning with a 300mm small format camera (35mm) equivalent lens approximately 16 bit of resolution for the angular measurements are required. Our current implementation with industrial type potentiometers and A/D conversion circuitry provides approximately 20 bits, so that the limiting factors become the sensor linearity and the (long term- and temperature-) drift that must be addressed by calibration. Angular measurements are transmitted by a serial connection to the connected computer, thus providing up-to-date orientation information for every rendered frame.

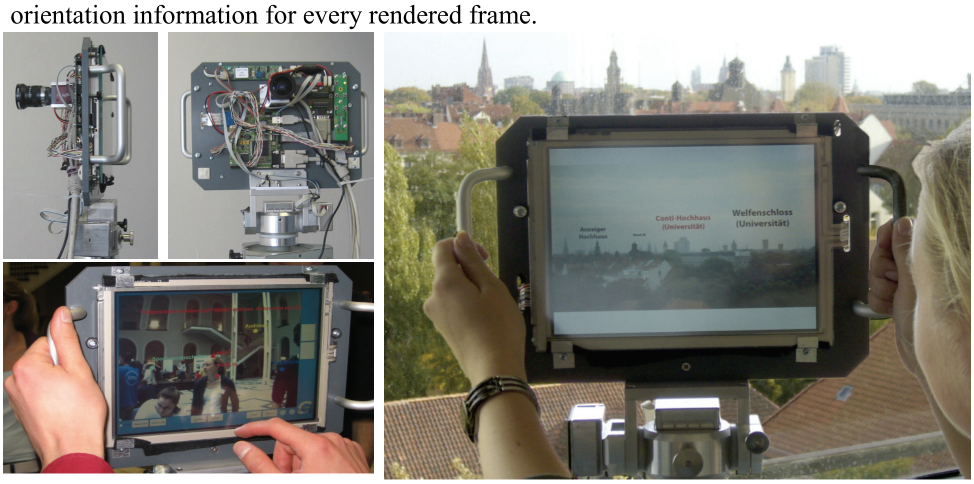

Figure 1: The GeoScope Prototype

4.3 Software architecture

Content management in mixed-reality application is often handled in an ad-hoc fashion, either by providing the augmentation information directly or by using the data structures of the graphics software (e.g. scene-graphs) to store the information. While such an approach is suitable for demonstrators and small-scale applications it does not scale to large scale applications or the visualization of existing (spatial) datasets.

In the GeoScope system augmentation content is stored and managed in a Geographic Information System (GIS), a spatial database that allows to capture, analyze and visualize geographic information. The GeoScope software is based on ArcObjects from ESRI (Esri06), the most widespread professional GIS application (Figure 2, left). Using spatial queries the spatial features within arbitrary "views" (positions and orientations) of the GeoScope as well as their attributes and relationships can be accessed from the underlying database. The GeoScope software can visualize the current view-frustrum of the GeoScope on a map to indicate which areas are currently queried. A key advantage of the GIS integration is the ability to manage large amounts of augmentation content with established tools and that potential to access and visualize large amounts of (existing) geo-spatial data that is accessible through the GIS system. By using a widespread GIS system like ESRI's ArcGIS existing data from city-planning, urban simulation, spatial measurement and similar use cases can be directly used with the GeoScope.

One special consideration in mixed reality applications is the correct blending of virtual information with real objects. In many applications (e.g. city planning, urban simulation) correct occlusion between real and virtual objects is paramount. Even if 3D environment data (e.g. building geometry) is available in a given database it is usually not of sufficient accuracy and detail to match the high accuracy of the GeoScope. In our system we provide several means to acquire masking-geometry for the surroundings based on close range photogrammetry or terrestrial laser scanning. Close range photogrammetry uses the triangulation principle to recover 3D information from at least two images from separate standpoints. Laser-scanning is the preferred mode of environment acquisition with the GeoScope. A Laser scanner can be operated on the same geodetic tripod as the GeoScope. Since only the visibility of objects from the standpoint of the GeoScope is of interest, a single terrestrial laser scan is sufficient to establish a correct depth map, eliminating the need for the acquisition and registration of multiple scans, which is usually the most time-consuming part during data acquisition. The acquired range data can be used directly as a depth map in the rendering of correctly overlayed images. In applications that involve several standpoints the depth-masks can also be stored in the GIS database for later access. For city planning and urban simulation applications we have acquired depth data with a Riegl LMS Z360I scanner, which has a measurement rate of 8,000 points per second, a range of 200 m and a single point accuracy of 12 mm (Rie06). Especially useful in the context of the GeoScope is the scanner's ability to cover a full 360 x 90 degrees field of view (Figure 2, right).

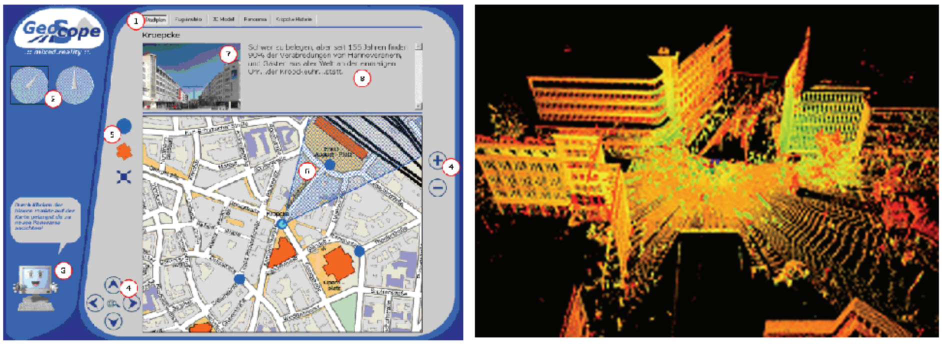

Figure 2: GIS Interface of the Geoscope (left); Laser range scan data (right)

5 Supported Interaction Techniques

A wide range of spatial interaction tasks can be realized on the GeoScope using the position and orientation information of the GeoScope in combination with the 2D pointing data and discrete interaction events generated on the touchscreen. For non spatial tasks arbitrary interaction techniques can be implemented similar to mouse controlled applications by using graphical interaction elements (widgets/controls) on the touchscreen. The combination of spatial pointing with the GeoScope with the touchscreen events allows to adapt many “action at a distance” interaction techniques that were developed for immersive virtual reality applications to mixed reality applications. The following list provides an overview of possible implementations of the spatial interaction techniques identified earlier:

- „Identification“: Pointing of the device; visual highlighting and/or annotation of objects in view.
- „Information“: Annotation of real objects in current view.
- „Localization“: Visual markers in the user’s view.
- „Guidance“: Visual guides/pointers.
- „Navigation“: Pointing of the device (restricted to view volume). Extension possible by use of 2D/3D widgets on screen.
- „Spatial Selection“: Pointing of the device combined with either discrete selection or area selection using the touchscreen.
- „Spatial Positioning“: Spatial “drag and drop” using pointing of the device and touchscreen for selection/deselection. Distance control requires additional 2D widget on screen.
- „Spatial Orientation“: Spatial “drag and drop” with touchscreen as clutching-mechanism; alternative use of a 3D rotation widget on screen.
- „Data Collection“: Use of position and orientation in combination with touchscreen. Distance control requires additional 2D widget on screen.

6 Calibration and Validation

The GeoScope consists of a number of components that can cause systematic errors in the mixed reality displays produced (positioning, lens distortions, angular measurement). In a calibration study these were examined systematically to establish their influence and derive possible measures for compensation. For the study we selected a city planning scenario as the baseline in which the GeoScope is used with a pentax 4.2 mm wide-angle lens. The camera has a resolution of 1024×768 pixels with a sensor size of 4,76 mm × 3,57 mm (corresponding to a pixel spacing of 4,65 µm). Typical distances of real objects in this scenario range from 10m upwards. To achieve pixel-precise overlay the position of the GeoScope must be known with a precision of 10cm, which is easily achieved using geodetic position measurements but very difficult using the sensors commonly employed in MR systems (e.g. low-cost GPS). Calibration of the camera/lens system established significant radial distortions (equivalent to up to 35 pixels offset). Using the calibration data these distortions can be removed from the video images by correcting distortions. Figure 3 (left) shows uncompensated radial distortions (magnified by factor 100), figure 3 (right) shows the remaining (random) measurement errors after application of distortion correction (magnified by factor 1000 for better visibility).

The rotation sensors of the GeoScope were calibrated using a field of geodetic measurement fiducials in a test room. The fiducials were captured using the (calibrated) GeoScope camera and used to determine accuracy and linearity of the rotation measures. The evaluation resulted in a maximum angular error of 0,168° which is insignificant for the city planning scenario but could be corrected if required for high precision applications. In outdoor use temperature differences and sensor aging could introduce additional errors that may require additional calibration.

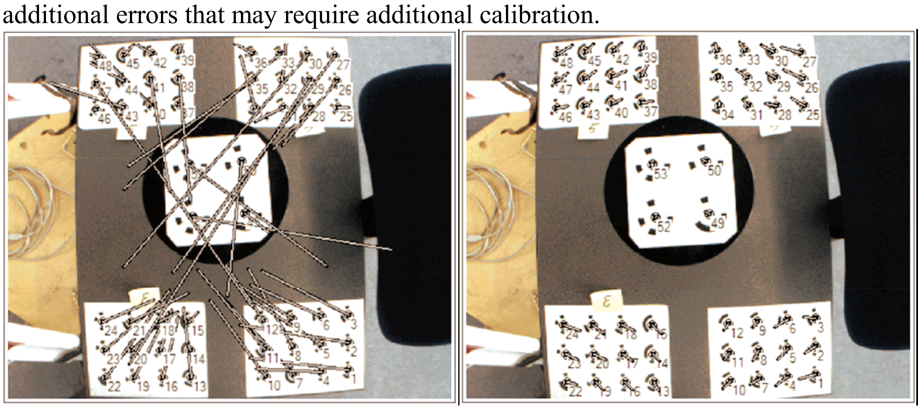

Figure 3: radial distortions
(left: uncorrected, magnified by factor 100; right: corrected, magnified by factor 1000)

The GeoScope was validated in a city planning application for a central location in downtown Hannover. Augmentation information was provided for various planning scenarios as well as historic views of the same area. 3D models of the historic views were acquired by close range photogrammetry from existing dioramas. 3D models for future plans were derived from CAD models and also modelled from scratch to generate planning alternatives for interaction. Laser range scan data for the location was acquired for correct occlusion. The GeoScope was publicly presented to large audiences at two events using this application and various other application scenarios that were developed as part of a year-long student project. Casual users rated the interaction possibilities provided as easy/intuitive and stated a high interest in using such a system in an actual public participation processes.

Figure 4: Planning application: depth map for occlusion (left); image overlay (right)

6 Discussion and Conclusions

The GeoScope supports many different applications involving public audiences, including city and landscape visualization, public participation or entertainment. By combining the touch-screen with the angular measurements from the GeoScope direct manipulation of distant (virtual) objects at “workspace” and “environment” scope becomes easy, supporting a large variety of spatial interaction tasks.

A number of mixed reality platforms have recently been presented that share some features with the GeoScope. The augurscope (SKF02) is a mobile tripod-mounted mixed reality display that can be rotated and tilted. In contrast to the GeoScope it is mobile in use but therefore cannot provide geodetic precision in positioning. Ydreams "Virtual Sightseeing" (YD06) is a tripod mounted LCD designed to operate as a coin operated device in outdoor settings. Interaction is limited to point&click. The XC-01 Augmented Reality Teleskop (XC06) is similar in functionality and limited interaction to the Ydreams system but uses a micro-display instead of an LCD. It therefore is not suitable for applications in which multiple users share the same view. This restricts the use in public participation activities but could be commercially more interesting in tourism applications for the same reason.

The precision of the positioning and orientation measures in the GeoScope in combination with careful calibration allows to augment images with high accuracy, especially when compared to existing mixed reality systems. Together with the high quality display and the ease of switching users this enables the use in applications like urban planning and makes the GeoScope a suitable evaluation platform for the study of mixed reality visualization techniques with larger groups of test users. The integration with a commercial GIS system supports the creation and management of mixed reality contents for large scale applications as well as the (re-)use of existing spatial datasets, making the use of mixed reality interfaces a viable solution for application domains in which spatial data surrounding the user must be visualized and mobility is not a key requirement. Future work will focus on the study of novel visualization and interaction techniques afforded by the GeoScope as well as the exploration of the GeoScope in other application domains, e.g. in visualization tasks with expert users. A number of promising application scenarios were developed in a student project group and future developments can build on these.

Acknowledgements

We would like to thank our colleagues Nora Ripperda and Jan Haunert for their work on the GeoScope project. Also 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 GeoScope topics in the course of a one year seminar.

Literature

[Azu97] Azuma, R: A Survey of Augmented Reality. In: Presence: Teleoperators and Virtual Environments 6, 4, pp. 355 – 385, 1997.

[Esri06] ESRI Website: http://www.esri.com/; accessed 1.11.2006.

[IU97] Ishii, H. and Ullmer, B.: Tangible bits: towards seamless interfaces between people, bits and atoms. In Proc. CHI '97. ACM Press, NY, 234-241, 1997.

[MK94] Milgram, P. and Kishino, F., A Taxonomy of Mixed Reality Visual Displays, in: IEICE Transactions on Information Systems, Vol E77-D (12), Dec. 1994.

[Rie06] Riegl Website: http://www.riegl.co.at/terrestrial_scanners/lms-z360i_/360i_all.htm; accessed 1.11.2006.

[SKF02] Schnädelbach, H., Koleva, B., Flintham, M., et al..: "The augurscope: a mixed reality interface for outdoors". In: Proc. CHI '02. ACM Press, New York, 9-16, 2002.

[Sut68] Sutherland, I.: A head-mounted three-dimensional display. 1968 Fall Joint Computer Conference, AFIPS Conference Proceedings, 33, pp. 757-764, 1968.

[WF02] Welch, G., Foxlin, E.: Motion Tracking: No Silver Bullet, but a Respectable Arsenal. In: IEEE Computer Graphics and Applications, Vol. 22., No. 6, pp. 24- 38, 2002.

[XC06] XC01 Website: http://www.xc-01.de, accessed 1.11.2006.

[YD06] Ydreams Website: http://www.ydreams.com/virtualsightseeing/eng/tour.html, accessed 1.11.2006.