Agenda

Let's Go Out: Workshop on Outdoor Mixed and Augmented Reality
ISMAR 2009, Orlando, FL
Monday, October 19, 2009

Organizers: Christian Sandor (University of South Australia), Itaru Kitahara (University of Tsukuba), Gerhard Reitmayr (Graz University of Technology), Steven Feiner (Columbia University), Yuichi Ohta (University of Tsukuba)

Timeline of the workshop's 3 segments:

9:15-11:00	11:00 -12:00	13:00-17:00
Position Statements Fast Forward	**Discussion about overarching research challenges**	**Focused discussions**

Detailed workshop agenda

Time	Agenda item	Speakers
9:00-9:15	Welcome, Workshop Introduction	Sandor, Kitahara, Reitmayr, Feiner
9:15 - 10:30	**Position Statements**	11 research presentations, 5 minutes each (+20 minutes buffer). Full list of position papers and speakers at the end of this agenda.
10:30-11:00	Coffee Break	
11:00 – 12:00	**Discussion about overarching research challenges**	Joint discussion: what the major high-level problems for outside AR? This sessions serves as preparation for the afternoon sessions.
12:00 - 13:00	Lunch	
13:00-15:00	**Focused discussions**	4 * 45 minute theme sessions (Session Moderators in brackets) • Modeling (Kitahara) • Tracking and Interaction (Reitmayr) • Visualization (Feiner) • Applications (Sandor)
15:00-15:30	Coffee Break	
15:30-16:30	**Focused discussions continued**	
16:30-17:00	Concluding remarks	Sandor, Kitahara, Reitmayr, Feiner

List of position papers

Talk Nr.	Area	Title	Authors
1	Modeling	Mobile Lidar Mapping: Building the Next Generation of Outdoor Environment Models for AR	Rob HARRAP (Queen's University, GeoEngineering Center), Sylvie DANIEL (Laval University, Department of Geomatics)
2		Implications of Having 3D Models for Outdoor Augmented Reality	Zhuming Ai (Naval Research Laboratory), Mark A. Livingston (Naval Research Laboratory)
3		Precise Manipulation at a Distance in Wearable Outdoor Augmented Reality	Thuong N. Hoang (Wearable Computer Lab – University of South Australia), Bruce H. Thomas (Wearable Computer Lab – University of South Australia)
4	Tracking and Interaction	Pedestrian Dead Reckoning and its applications	Masakatsu Kourogi (Center for Service Research, National Institute of Advanced Industrial Science and Technology), Tomoya Ishikawa (Center for Service Research, National Institute of Advanced Industrial Science and Technology) ,Yoshinari Kameda (University of Tsukuba, Tsukuba), Jun Ishikawa (University of Shizuoka), Takeshi Kurata (Center for Service Research, National Institute of Advanced Industrial Science and Technology)
5		3D Pointing Interface by using Virtual Diorama for Attention Sharing	Masayuki Hayashi (University of Tsukuba), Itaru Kitahara (University of Tsukuba), Yoshinari Kameda (University of Tsukuba), Yuichi Ohta (University of Tsukuba)
6	Visualization	Adaptive Visualization in Outdoor AR Displays	Denis Kalkofen (TU Graz), Stefanie Zollman (TU Graz), Gerhard Schall (TU Graz), Gerhard Reitmayr (TU Graz), Dieter Schmalstieg (TU Graz)
7		Prototyping an Outdoor Mobile Augmented Reality Street View Application	Yoshitaka Tokusho (Software Division, Hitachi Ltd), Steven Feiner (Department of Computer Science, Columbia University)
8		Interaction with the Environment: Sensor Data Visualization in Outdoor Augmented Reality	Sean White (Department of Computer Science, Columbia University Department of Botany, Smithsonian Institution)
9		AR-based GIS: Towards closing the loop between acquisition and visualization of geographical data	Guillaume Moreau (Ecole Centrale de Nantes – CERMA), Myriam Servieres (Ecole Centrale de Nantes – CERMA), Hideo Saito (Keio University)
10	Application	2-Way Wireless Spherical Panoramic Video Telepresence: "The Killer App for VR/AR"	Kurtis J. Ritchey (U.S. Army, National Simulation Center)
11		Augmented Reality for Bulding and Construction – Software Architecture	Charles Woodward (VTT Technical Research Centre of Finland), Mika Hakkarainen, Kari Rainio (VTT Technical Research Centre of Finland)

Mobile Lidar Mapping: Building the Next Generation of Outdoor Environment Models for AR

Rob HARRAP[1]
Queen's University
GeoEngineering Center
Kinsgton (ON), Canada

Sylvie DANIEL[2]
Laval University
Department of Geomatics
Quebec City (QC) – Canada

ABSTRACT

There are broad synergies between the 'world' requirements for outdoor augmented reality applications, especially augmented games situated in realistic worlds, and the core capabilities of geomatics tools and techniques. AR applications are fundamentally rooted in space; they require accurae 3D models of the world to be able to provide highly immersive and interactive experiences. New high resolution modeling techniques using LiDAR scanning support the creation of centimeter-accurate models of real world settings for use in augmented reality and situated games; the cost of such models can be leveraged by the many other applications of such models.

KEYWORDS: Spatial Databases and GIS, Modeling and Recovery of Physical Attributes, Computer-Aided Design (CAD).

INDEX TERMS: H.2.8 [Database Management]: Databases Applications; I.2.10 [Artificial Intelligence]: Vision and Scenes Understanding; J.6 [Computer-Aided Engineering]

1 CONTEXT OF OUTDOOR MOBILE AUGMENTED REALITY

We have seen recently an emergence in the public domain of outdoor mobile augmented reality (AR) applications on a wide range of hardware platforms. Solutions like Layars (http://Layar.eu) or Wikitude (http://www.wkitude.org/) provide new platforms allowing users to browse the world. These applications put an information overlay on top of the camera view of a smartphone, bringing digital data of various sorts into play whenever the user is looking at or for something in the real world. The world augmentation in these applications is rather light, and interests to improve and enhance the user experience have already been expressed. Current developments are focusing on geo-tagging and authoring platforms. If buildings and open places are the general targets, interactions with items like doors, light poles, signage, and cultural assets that exist in a typical urban area, may be a real challenge.

Augmentation of reality pushes the problems of having a physical and photorealistic model of a domain to the limit as the overlay process relies on accurate alignment between that model and the world. This augmented world model, or set of models, must exist at a variety of scales – corresponding to the scales at which the user navigates a region (blocks) down to the scales of fine-grained interactions (centimeters). Features must also have rich annotations that support a variety of interaction styles, search, discovery, and community annotation.

2 HIGH RESOLUTION WORLD ACQUISITION

In the proposed position paper, we will lay out a general view of the relevance of high resolution mobile LiDAR mapping to AR development. Our perspective is that, although expensive to acquire and difficult to process, LiDAR data represents a key framework for future augmented reality applications that require centimeter level feature representation. This is especially true in the case of augmented reality games.

2.1 Mobile high resolution LiDAR

Our concern in constructing outdoor mobile augmented reality has been the difficulty in collecting data at the scale of buildings and especially building components. Support for situated activity [1] as well as high resolution urban mapping demands models where features down to 'doorknob scale' are represented. Thus, the overwhelming problem to tackle is that of data acquisition at this level of detail. This exceeds the difficulty in fields such as game world building and computer animation where models must be precise (detailed and photorealistic) but not accurate (they don't match any real world setting precisely).

As part of an ongoing research program on augmented reality, we are investigating the use of static and mobile LiDAR scanners to build geospatial data at unprecedented levels of accuracy. LiDAR scanners [2] rasterize a ranging laser over a target, generating a cloud of 3d locations at a density that can approach one point per square millimeter close to the scanner. LiDAR scanners rely on time of flight or phase change of the beam to determine ranging, and can deliver clouds that are accurate to a few centimeters out to hundreds of meters from the scanner. In the simplest case, the scanner is mounted on a tripod at a known location ('static' setup) and generates a cloud out to line-of-sight or maximum range, whichever is less.

Scanning large areas with a tripod-mounted LiDAR is cumbersome. We have begun to investigate the use of mobile scanners to model outdoor environments. Mobile scanners, like the Terrapoint TITAN system (http://www.ambercore.com) we are using, generally consist of multiple LiDAR units integrated with GPS and inertial movement unit sensors as a single package that can be mounted on a truck body. They can scan at flow-of-traffic speeds, and with one pass can collect about 100 points per square meter, sufficient for building and building component modeling. The GPS and inertial movement unit work together to provide precise positioning during scanning; An example is shown in Figure 1. Note that this scene is perfectly recognizable as buildings and street infrastructure to a human viewer, but is in fact a very large collection of points in 3d. Objects are captured in their precise locations - including those not shown on national

[1] Harrap@geol.queensu.ca
[2] sylvie.daniel@scg.ulaval.ca

mapping such as street furniture – effectively mapping the truly immersive 3D built environment.

For areas where even higher detail is needed we have successfully merged data from static scanners to TITAN data resulting in localized models accurate to millimeters.

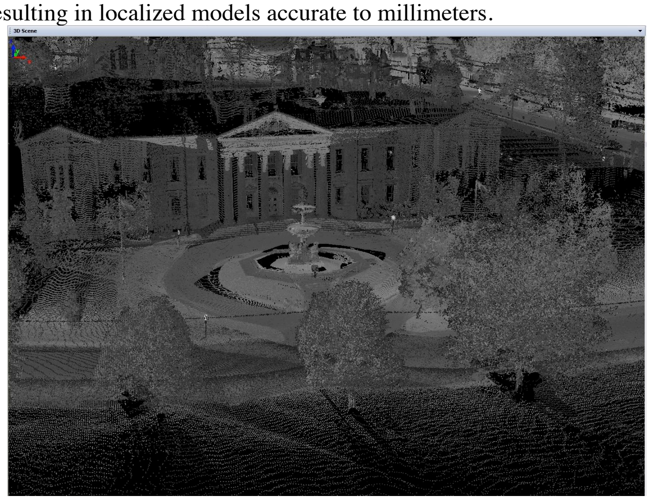


Figure 1. TITAN scan of a portion of downtown Kingston, Ontario, Canada. This data is in real world coordinates

2.2 Discussion

LiDAR scanners generate very large clouds of point locations that can constrain geometric models of 3d features in dense environments such as streets, the interior of buildings, and furthermore can provide detailed reference models for 3d shapes of items such as large vehicles. However, there are a number of challenges facing widespread use of this approach, including completeness of the scan, processing of the cloud, and markup of the resulting model.

First, the LiDAR and photograph collection process covers buildings based on orientation of the sensors and the presence of trees and other obstructions as can be clearly seen in Figure 1. The result is quite good when trees and other obstacles can be avoided, but may have gaps where these are close to building faces. There is abundant research on model gap-filling using procedural modeling, where the regularity of building features are used to estimate the content for gaps and so in cases where a procedural proxy is suitable and features are regular the effect of obscuring vegetation can be mitigated. Similar but typically more challenging issues arise when obscurance is due to buildings, where street access is impossible, or where law or policy prohibits the use of the scanner. Although there are gaps in the data as noted, geometric distortion is typically small since the mobile systems use sophisticated GPS and INS hardware to maintain position and pose accuracy. As a result, scanned objects show little distortion and even at significant driving speeds reproducible accuracy is on the order of 5cm, more than sufficient for highly detailed interactive feature modeling.

There are also significant challenges in processing the resulting point clouds into CAD or GIS surface representations for use in AR. The primary difficulty with processing data of this type is the time taken to convert point clouds into features. This is at the moment labour intensive; however, compared to any other available method (e.g. photogrammetry, surveying, CAD recovery) known to produce similar detailed models of large areas it is highly efficient, accurate, and repeatable. Although the exact ratio is highly context-dependent, we have found that about thirty times as much on-computer processing time is needed as field time: behind every lidar field survey is an army of technicians turning points into features and then into 3d world models with semantic annotations. Research work, relying on computer vision methods, is currently undertaken in order to help reducing the processing time by assisting with commonplace and repeating items such as light poles, curbs, and street furniture. In the long run, we may have systems that will be able to build a preliminary model during the scan acquisition.

Note that processing of points to CAD or GIS representations includes two significant components: the geometric feature and the semantics of that feature, the 'what is it' issue. Even given a fixed ontology of objects in a real world or game world domain, classification into a single consistent scheme of things, especially things that have interactive properties, is non-trivial. For example, a door scanned in LiDAR is a geometric feature, but has to be recoded as an object that can undergo constrained motion, with specific axes, and only after interaction with a subpart, the doorknob. And only then if the door is unlocked. Having the door as a physical representation is thus a very small step towards building a realistic, pervasive, and accurate world, and yet it is an essential first step. Building simpler models capturing only geometry and identity is somewhat simpler, and may suffice for many AR applications.

Although the cost to collect and process LiDAR data is significant, we have successfully argued in existing projects that such framework models are of interest to urban planners, first responders, urban business councils, and urban utility maintenance staff, and have partnered with these organizations to share the cost and distribute the workload of building accurate GIS models of urban areas. Augmented reality developers in general, especially game developers, can benefit from such partnerships; some of these partners are quite interested in games as a means to training, simulation, and outreach, and all are interested in high quality GIS data. The ability to capture all of this information at unrivalled resolution and coverage - all from a moving vehicle – makes this a very rapid and cost effective method of surveying. The ability to then do it again and carry out change detection allows unprecedented support for urban growth, vegetation, and natural hazards studies.

3 Conclusion

Even with the significant challenges in efficient model building, outdoor world acquisition using mobile terrestrial scanners is far more efficient than any other known approach for building centimeter-scale models of large domains. Mobile terrestrial scanners offer great advantages for building the 'where' in situated applications. Together with photography these models can provide 'sets' for mobile AR applications at a reasonable cost. Driven by new extraction techniques and cheaper sensors, LiDAR data will become as commonplace in AR and game development as motion tracking already is. As underlined by Large [3], "by using laser technology, we are embarking on the second revolution in the surveying and monitoring of our physical environment".

References

[1] R.M. Harrap. Situated interactions: humans, computers, and rich media in space and time. In REGARD Workshop on Mobile Augmented Reality, August 2008, to be published 2009 by Springer.

[2] R.A. Fowler, A. Samberg, M.J. Flood, T.J. Greaves. Modeling mobile terrestrial LiDAR to vector based models. In D.F. Maune, editor, Digital Elevation Model Technologies And Applications: The DEM Users Manual 2nd Edition, 2007, pp.199-248.

[3] A.R.G. Large, G.L. Heritage, M.E. Charlton. Laser scanning: the future. In George L. Heritage and Andrew R.G. Large, editors, Laser Scanning for the Environmental Sciences, 2009, pp.262-271

Implications of Having 3D Models for Outdoor Augmented Reality

Zhuming Ai* Mark A. Livingston†

Naval Research Laboratory

1 Introduction

One major issue for outdoor augmented reality (AR) systems is how much of the world must be modeled in order to operate the system. This paper discusses several issues as they relate to requirements for a 3D model of the operating environment. In many cases, a world model has been used in order to be able to make assumptions that simplify building other components of the system. Creating a model is not a trivial task in most cases, though automated methods are improving in performance (accuracy, degree of automation, and time required). We review techniques and reasons for building world models for outdoor AR systems and speculate on ways to reduce the requirements in various components of an AR system.

2 Model Creation

Many outdoor mixed and augmented reality projects rely on an accurate and detailed model of the environment. The models must often be extremely accurate. One analysis, using nominal position and orientation accuracy (at the time) of 0.1m and 2°, suggested that modeling errors could not exceed 0.5m [5] if the application is to highlight repeating urban features such as windows separated (at their centers) by 2m. A more recently proposed tracking system cited an application need for a maximum of 1m of error when sighting a location 100m away [2], approximately 0.57°. Our current application focus has adopted this performance standard for the registration error of the AR system.

Because many applications must highlight individual building features, it is not sufficient to extract the geometry of the buildings and simply apply a texture to them. There is no single method that can reliably and accurately create all the models. Techniques as simple as manual surveying or as complex as reconstruction from Light Detection and Ranging (LIDAR) sensing may be used to generate such a model.

For small areas for research and testing purposes, a combination of extruding from a blueprint, augmented with survey tools, is feasible. We built a very detailed and accurate model of our testing area using this method [5].

The generation of 3D urban maps has been an area of rapid development. There are some LIDAR systems that can rapidly map a city within a day or even hours [8]. Prototypical unmanned aerial vehicle (UAV) based LIDAR systems have already been demonstrated, which makes it possible for future small military units to generate this data for tactical missions. However, errors in current LIDAR measurements are still high, and it may not possible to create models as detailed as many MR/AR applications need. One fundamental problem is that the AR users will usually be on the ground; LIDAR is acquired from the air, leaving much building detail near the ground subject to errors and occlusions.

Many studies have been done to create urban models based on image sequences [6]. It is a non-trivial task to obtain these attributes in the general case of an arbitrary location in the world. Automated systems [7] are an active research topic, and semi-automated methods have been demonstrated at both large and small scales [5].

While the availability of the models of big fixed objects like buildings is a tractable issue, moving objects such as cars or persons are more difficult to handle. Research has been done to detect and track moving objects based on video images [3]. By integrating multiple 2D video streams into a 3D environment, it is possible to identify and track objects in 3D. This could be done by first detecting and tracking moving objects in video streams using existing algorithms. Then this information could be combined in 3D with the knowledge of the 3D relationships.

Google Earth has 3D buildings in many areas; this information may be available for Google Earth users and thus could be used for outdoor AR applications. There are also many community-created models available for Google Earth, including some models of whole cities. Many of these model are very detailed. However the accuracy of these models varies from site to site, and whether they are of sufficient quality for AR needs to be tested. A more accurate model may be needed to get the desired results.

3 Features Affected

Many features on an outdoor AR system may be affected by the fidelity of the model of the real environment. One theme that runs through these issues, as for AR systems in general, is the registration accuracy. There are numerous aspects to consider. Clearly, if something is to be registered to a real object, then the coordinates of that real object must be known to some degree in the AR database. One motivation for modeling the real world in our applications was that displaying the real-world model would give us a sense of how (in)accurate the registration was during operation of the system. Rough estimates may be derived without actually running the system (based on offline measurements and mathematical models) [4], but online feedback is perhaps the most reliable. Even for a simple case, such as a virtual object portrayed as sitting on the real ground, a model of the terrain may be necessary. A perfectly flat plane is one extreme example, but even gently rolling terrain or a single building introduces modeling requirements for achieving registration against the real environment.

3.1 Tracking and Calibration

Closely related to the issue of registration is the strategy by which tracking and calibration of the tracking components are achieved. Many technologies require some sort of base station, which must generally be at a known position. This is in effect a world model, albeit sometimes of extremely low fidelity. A dense arrangement of fiducial markers or acoustic beacons may constitute a more dense world model, although many times such systems may require only a plane of beacons – but a precise model (within that plane) for best accuracy. Computer vision systems that use features from the real environment have provided an active topic of research in tracking for AR, but a system that works outdoors, with high accuracy, and without a world model eludes the field. Hybrid systems may be used to boot-strap a vision component, but the problem still requires what the robotics community calls simultaneous localization and mapping (SLAM), a very challenging problem. How errors propagate forward from slightly inaccurately triangulated features is a major difficulty in the numerically-sensitive computations. The

*e-mail: ai@itd.nrl.navy.mil

†e-mail: mark.livingston@nrl.navy.mil

localization portion of the problem becomes much easier if a map in the form of a set of accurately known 3D points may be used as beacons, but then this becomes mathematically equivalent to building a low-fidelity world model. Thus outdoor tracking right now benefits greatly from having some form of modeled world, although the burden has been reduced by improved sensors and integration algorithms.

Calibration of an AR system presents another, smaller challenge. Inertial sensors are a popular component of outdoor tracking systems. A magnetic compass in particular has been a thorny issue in our outdoor tracking attempts, since in most urban environments, the amount of metal present in the environment distorts the compass data. (Many inexpensive compasses do not have the accuracy even in ideal operating environments.) We have typically solved this through a calibration procedure that directs the user to stand in a known location and align some virtual objects with known – i.e. modeled – real-world objects. While we re-use some of the world model acquired for demonstrating registration, this is a potential burden on the system. One method discussed (but not implemented and tested for accuracy to our knowledge) is to have two mobile users, both tracked in position through GPS, calibrate against each other's position. This would potentially avoid the need for a world model against which to calibrate.

3.2 Occlusion

One potentially powerful feature of AR is the capability to "see" objects that are occluded from view through graphical analogues. This "X-ray vision" concept was among the earliest stated benefits of AR applications. While in many applications, assumptions can be made that simplify the implementation, proper depiction of complex occlusion relationships between real and graphical objects also requires some fidelity in the world model. For example, if a graphical object is to be occluded by a real one, then the depth of the real object must be known so that the data may be considered in the rendering process. If the graphics should always occlude the real world, then one may not require any world model for occlusion. However, the registration issues discussed above may still inhibit the perceptual integration of the virtual overlay with the real world. The integration of multiple layers of graphical objects against multiple known real-world layers becomes a very complex question, requiring graphical depictions of occlusion relationships. The application requirements must be carefully reviewed.

3.3 Video overlays and other sensor data

Projecting real time images on top of 3D models has been widely practiced, and there are some attempts at augmenting live video streams for remote participation and remote videoconferencing. We have developed a prototypical implementation of a mixed reality (MR) based system that integrates georegistered information on a virtual globe (Google Earth) [1]. The application can be used for a command and control center to monitor the field operation where multiple AR users are engaging in a collaborative mission. The system integrates georegistered icons, live video streams from field operators or surveillance cameras, 3D models, and satellite or aerial photos into one MR environment.

The prerequisite of projecting the images on the wall or other 3D objects is that we need to have a database of the models of all the objects, so that the projection planes can be determined. The projection errors on the building are pretty obvious. There are several sources of errors involved. One is the accuracy of the models of the buildings. Other problems come from camera tracking, calibration, and lens distortion. In this study, we used the model we created to calculate the video overlay. Our model is very accurate, but the buildings shown in the final result are from Google Earth because they cover a much larger area. These two models are not matching very well and cause some confusion.

Moving objects such as cars or persons will cause blocked parts that can not be removed if we do not generate their models on the fly. Research has been done to detect moving objects based on video images [3]. While in theory it is possible to project the video image on these moving objects, it is difficult to recreate 3D models in real time with few images. In our applications it is not really necessary, and we are projecting the images on known models instead.

3.4 Interaction

To interact with real objects, whether for collaborating or for editing the geometric database, one must be able to select objects. One popular method for doing so is pointing with a hand-held pointer, leading to ray-casting methods and volumetric extensions. This also presumes the existence of a world model, at whatever level of detail the user wishes to select objects. Our system includes a hierarchical model of buildings containing walls, which in turn contain doors and windows. Such a model has many uses for interactive specification of targets or destinations in a collaborative mobile system. To avoid having a such model prior to running the system would require a way to measure depth from a user's position (e.g. a laser range finder or computer vision techniques). One could then specify not an object per se, but a location that is known (by measurement) to correspond to a real-world object. In doing so, one essentially acquires a (low-fidelity) world model online. The cost-benefit analysis of having such a system capable on online measurements versus acquisition of a priori model would depend on the application and perhaps even upon the task or role which a user plays.

4 Conclusions

Creating a model of the environment is an major task for outdoor AR applications. There are cases for which it is not possible to get an accurate model for the task, yet no solutions for the task have been documented without requiring at least some modeling. The implications of the lack of a (detailed or accurate) model are the concerns raised in this paper. One of the major impacts is on tracking and registration; others include occlusion, overlay, and interaction. Although there are ways to overcome these difficulties through improved sensing or display algorithms, not to mention simply ignoring the errors created, at the end, some kind of model has to be created either prior to the mission or on the fly. Model creation should thus receive greater focus in research for outdoor AR.

References

[1] Z. Ai and M. A. Livingston. Integration of georegistered information on a virtual globe. In *ISMAR '09*, October 2009.

[2] R. Azuma, H. Neely, M. Daily, and J. Leonard. Performance analysis of an outdoor augmented reality tracking system that relies upon a few mobile beacons. In *ISMAR'06*, pages 101–104, Santa Barbara, 2006.

[3] E. J. Carmona, J. M. Cantos, and J. Mira. A new video segmentation method of moving objects based on blob-level knowledge. *Pattern Recogn. Lett.*, 29(3):272–285, 2008.

[4] R. L. Holloway. Registration error analysis for augmented reality. *Presence: Teleoperators and Virtual Environments*, 6(4):413–432, 1997.

[5] S. Julier, Y. Baillot, M. Lanzagorta, L. Rosenblum, and D. Brown. Urban terrain modeling for augmented reality applications. In M. Abdelguerfi, editor, *3D Synthetic Environments Reconstruction*, pages 119–136. Kluwer Academic Publishers, Dordrecht, 2001.

[6] A. Jurisch and D. Mountain. Evaluating the viability of pictometry imagery for creating models of the built environment. In O. Gervasi, B. Murgante, A. Laganà, D. Taniar, Y. Mun, and M. L. Gavrilova, editors, *Lecture Notes in Computer Science*, volume 5072, pages 663–677. Springer, 2008.

[7] M. Pollefeys. *3-D Modeling from Images*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2005.

[8] C. Poullis and S. You. Automatic creation of massive virtual cities. In *VR '09*, pages 199–202, Lafayette, LA, March 2009. IEEE.

Precise Manipulation at a Distance in Wearable Outdoor Augmented Reality

Thuong N. Hoang[1] and Bruce H. Thomas[2]

Wearable Computer Lab – University of South Australia

Abstract

In this paper we present our research directions on the issue of precise manipulation and modeling interactions for outdoor augmented reality. We also discuss the related issues of action at a distance techniques and the relationship between virtual and physical objects. Techniques for action at a distance are limited by the accuracy of the sensors, as well as the efficiency of tracking solution. We examine possible solutions for precise interactions including symbolic entry of data, grid snapping, and alignment operations. Grid snapping and relative snapping are also discussed as ways to improve precision.

Keywords: User Interaction, Direct Manipulation, Outdoor AR.

Index Terms: H.5.2 [Information Interfaces and Presentation]: Graphical User Interfaces—Input Devices and Strategies; I.3.6 [Computer Graphics]: Methodology and Techniques—Interaction Techniques

1 Introduction

Our investigation into augmented reality has focused on outdoor AR on wearable computer. Tinmith wearable computer system is the main platform for our research, employing video see-through head-mounted display with belt-mounted computer.

We are motivated to find solutions to the question of "*How to perform precise interactions in outdoor AR?*". Successful solutions are beneficial to AR modeling in the creation and modification of high-fidelity and accurate models using AR, without the need for desktop-based CAD applications. Precise interactions also help to fine-tune the virtual part of AR via translation, scale, and rotation operations. For example, the user could scale a virtual building model to an exact size as specified in a building plan. We have recently been successful at precise manipulation for outdoor model refinements, using a handheld 3DOF orientation tracked single-point laser rangefinder [1].

Other issues arising from the main problem of precise manipulation include techniques for interaction at a distance in outdoor AR and relationship between the virtual and the real world. In an outdoor environment, it is common that manipulation tasks are generally out of arms reach, due to the inherently large scope of the environment and the size of outdoor models, such as buildings or street furniture. Research efforts in action at a distance interactions have been mainly limited in virtual reality, which does not directly apply to augmented reality, considering that physical world imposes more constraints on AR than VR.

Unlike desktop-based CAD applications, precise operations, such as affine transformation or modeling tasks, are not fully supported in outdoor AR, due to various restrictions on input devices and sensors noise. Further challenge is posed by the lack of suitable and robust coordinate systems for the combination of the virtual and physical environment. Such a coordinate system would assist in correctly register the virtual world against the real one.

1.1 Precise Manipulation

Precise manipulation allows a user to transform a graphical object at a much higher level of accuracy to that of free hand manipulation. Without robust hand tracking, free hand manipulation is prone to errors and imprecision.

1.1.1 Symbolic entry of the data

Desktop-based CAD applications use menus, dialogs, or toolbars to allow user to input exact measurement data via keyboard. Text entry in wearable computer could either be a forearm mounted keyboard [2], gesture-based hand input [3], or speech recognition [4]. Even though forearm keyboard is preferred over virtual keyboard [2], it requires the use of both hands. Gesture-based approach and speech recognition while walking suffer a higher error rate than pen-based solution [3, 4].

1.1.2 Snapping to a grid

Snapping the objects to an axis-aligned grid eliminates errors due to trackers noise or free hand operations. Grid snapping is a common feature for most 2D desktop-based CAD modeling system. Precision is also achievable for 3D grid snapping [5]. However, the operations can only be as precise as the grid is small. This leads to another problem of how to precisely resize the grid resolution.

1.1.3 Alignment operations

Instead of attaching to vertices of the grid in grid snapping, virtual objects could be aligned with one another or with other physical objects, via vertices, edges, and facets. Such alignment requires correct registration of virtual object to the real world. Depending on the task at hand, the virtual object could be either registered using a world coordinate system, or attached relatively to the physical objects.

1.2 Action-at-a-distance

It is well known action-at-a-distance suffers from tracker noise. A number of solutions have been examined for indoor VR systems, such as worlds in miniature [6], voodoo dolls [7], scaled manipulation [8] and scale world zoom [9]. Although a number of worlds in miniature interfaces have been developed for indoor and outdoor AR [10], these have focused on the relationship within the virtual domain only.

Scale world zoom is not appropriate for AR because the AR information is registered to the physical world. The use of scale world zoom would break the one-to-one relationship between the virtual and physical worlds. Because people can not physical scale up, the virtual information would have virtual scale down.

Outdoor augmented reality has inherent problems with action-at-a-distance due to the quality of the sensors for 6DOF tracking.

1) ngocthuong@gmail.com

2) Bruce.Thomas@unisa.edu.au

While indoor tracking can obtain millimeter tracking accuracy, outdoor tracking is left to GPS or vision based sensing systems. At the very best this would be 2-3 centimeter accuracy. While orientation sensors themselves are quite accurate, the act of physical walking adversely affects the accuracy of these sensors. A good solution for the relative tracking of a user's hands for wearable outdoor augmented reality systems is still an open research question. The tracking of a user's hand is a key technology for many of the action-at-a-distance VR solutions.

2 Relationship Between Virtual and Physical Objects

In order to perform precise manipulation, a coordinate system or reference framework for virtual models is required. While outside, virtual objects can be relative to some world coordination systems, such as GPS, or these virtual objects can be relative to a physical object, such as a building.

2.1 World Coordinate Relationships

The use of placing objects with some world coordinate system by itself places the virtual object in some absolute position and orientation. This form of a coordinate system makes for a quite straightforward implementation and a unified coordinate system. Employed by itself, it can only support an open loop tracking framework to specify the relationship between virtual and physical objects.

GPS is commonly used as the world coordinate system for virtual objects in AR. Physical objects inherently have a set of associated GPS coordinates. By embedding GPS coordinates into virtual objects, outdoor AR system is able to correctly position the virtual object into the physical world. Thus the relationship between virtual and physical objects is specified, within the unified Global Position System. GPS tracking of user position is required in combination with this world coordinate system.

2.2 Relative Relationships

The position and orientation of virtual objects can be specified via a relationship between each of them. Using the world coordinate of the physical and a transformation from that object, the position and orientation of a virtual object can be specified. For example a particular window can be augmented with a set of highlighting surrounding it. The location of the augmentation is specified by its relationship to the building. This relationship can be an offset from the physical structure. This enables the user to view the virtual object from different view points.

3 Snapping and Constrained manipulation in Outdoor AR

Snapping is a common approach to precise manipulation, and 3D grid snapping has been shown to improve precision [5]. However, is grid snapping applicable in an outdoor AR environment? The challenge of suitable input devices still remains, as well as interaction techniques for snapping operations at a distance. The appropriate visualization for snapping operation also requires further consideration. Rendering of a 3D grid provide obvious indication to the user that grid snapping is effective, but might be obtrusive for immersive AR. Mechanism for determining suitable grid size requires further research. A solution could be suggested to snap objects along its own coordinate axes, instead of a global grid. The pixels representing the objects could be used as snapping points.

Snapping in outdoor AR could leverage the physical world, which is a unique feature of AR systems, as opposed to VR. Instead of being snapped to a grid structure, virtual objects could also be relatively attached or constrained to physical objects, by vertices, edges or facets. An example of such relative constraints include a virtual door being attached to a wall of a physical building, or a virtual fence running parallel at a fixed distance to the walls of a physical building. Relative snapping could also be used between virtual objects. For example, a virtual gate could be fixed onto the virtual fence in the previous example. This technique could also be referred to as constrained manipulation.

The techniques of specifying such relative relationships between virtual and physical objects are open for further research. Baillot et al. [11] presents the techniques of creating points in space as an intersection between a virtual ray and another ray, or a plane/surface. Using a laser rangefinder, Wither et al. [12] proposes techniques to affix virtual annotations to physical buildings, with correct orientation. A hybrid solution could be proposed, in which a virtual replica of the physical object is registered at the exact location of its real world counterpart, using world coordinate system, so that other virtual objects could be positioned relative to the virtual replica.

Snapping in outdoor AR, whether grid snapping or relative snapping with physical objects, assists in precise manipulation of virtual objects. By applying snapping constraints on virtual objects, manipulation errors caused by tracker devices or caused by free hand movement could be eliminated.

Reference

[1] T. N. Hoang and B. Thomas, "In-situ Refinement Techniques for Outdoor Geo-Referenced Models Using Mobile AR," in IEEE/ACM International Symposium on Mixed and Augmented Reality, 2009. ISMAR 2009., Orlando, Florida USA, 2009.

[2] B. Thomas, S. Tyerman, and K. Grimmer, "Evaluation of text input mechanisms for wearable computers," Virtual Reality, vol. 3, pp. 187-199, 1998.

[3] Y. Liu, X. Liu, and Y. Jia, "Hand-Gesture Based Text Input for Wearable Computers," in Computer Vision Systems, 2006 ICVS '06. IEEE International Conference on, 2006, pp. 8-8.

[4] K. J. Price, M. Lin, J. Feng, R. Goldman, A. Sears, and J. A. Jacko, "Data Entry on the Move: An Examination of Nomadic Speech-Based Text Entry," in User-Centered Interaction Paradigms for Universal Access in the Information Society, 2004, pp. 460-471.

[5] E. A. Bier, "Snap-dragging in three dimensions," SIGGRAPH Comput. Graph., vol. 24, pp. 193-204, 1990.

[6] R. Stoakley, M. Conway, and R. Pausch, "Virtual reality on a WIM: interactive worlds in miniature," in Conference on Human Factors in Computing Systems, Denver, Colorado, 1995, pp. 265-272.

[7] J. S. Pierce, B. C. Steams, and R. Pausch, "Voodoo Dolls: Seamless Interaction at Multiple Scales in Virtual Environments," in 1999 Symposium on Interactive 3D Graphics, 1999, pp. 141-145.

[8] S. Frees and G. D. Kessler, "Precise and rapid interaction through scaled manipulation in immersive virtual environments," in Virtual Reality, 2005. IEEE, 2005, pp. 99-106.

[9] M. Mine, F. P. Brooks, and C. H. Sequin, "Moving Objects In Space: Exploiting Proprioception In Virtual-Environment Interaction," in ACM SIGGRAPH 1997, Los Angeles, Ca, 1997, pp. 19-26.

[10] T. Höllerer, S. Feiner, D. Hallaway, B. Bell, M. Lanzagorta, D. Brown, S. Julier, Y. Baillot, and L. Rosenblum, "User interface management techniques for collaborative mobile augmented reality," Computers & Graphics, vol. 25, pp. 799-810, 2001.

[11] Y. Baillot, D. Brown, and S. Julier, "Authoring of physical models using mobile computers," in The 5th International Symposium on Wearable Computers ISWC 2001, ETH Zurich, 2001, pp. 39-46.

[12] J. Wither, C. Coffin, J. Ventura, and T. Hollerer, "Fast annotation and modeling with a single-point laser range finder," in 7th IEEE/ACM International Symposium on Mixed and Augmented Reality, ISMAR 2008., 2008, pp. 65-68.

Pedestrian Dead Reckoning and its applications

Masakatsu Kourogi*, Tomoya Ishikawa*,Yoshinari Kameda**, Jun Ishikawa***, Kyota Aoki****, Takeshi Kurata*

* Center for Service Research, National Institute of Advanced Industrial Science and Technology, AIST Tsukuba, Japan
** University of Tsukuba, Tsukuba, Japan
*** University of Shizuoka, Shizuoka, Japan
**** Utsunomiya University, Utsunomiya, Japan

ABSTRACT

First, we briefly introduce the technologies of PDR (Pedestrian Dead Reckoning) that can be used to track the location and orientation of a pedestrian. Secondly, we show some enhancement of the PDR to improve its performance on accuracy. Finally, we describe the applications of the PDR.

KEYWORDS: Pedestrian Dead Reckoning, Pedestrian navigation system

INDEX TERMS: K.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems□Artificial, augmented, and virtual realities; I.4.8 [Image Processing and Computer Vision]: Scene Analysis□Sensor fusion.

1 INTRODUCTION

Recently, the PDR technologies draw strong attention from many fields of industries since by using the self-contained sensors (ex. accelerometers, gyrosensors, magnetometers and barometers) the PDR can be utilized to track location and orientation of pedestrians indoors where the GPS signal is blocked and becomes unavailable. Compared to the INS (inertial navigation system), the PDR has several advantages on cost, weight and energy consumption of the system thanks to the recent rapid advancement of the MEMS (microelectromechanical system) sensing components. The PDR is classified into two categories by the equipped position of the self-contained sensors on the pedestrian: (a) waist-mounted and (b) shoe-mounted [3]. We took the former approach on PDR since the equipped position allows the system to gather information about the user's action such as walking, going-up/down stairs, standing-up and sitting-down on a chair.

2 BASICS OF PEDESTRIAN DEAD RECKONING

The PDR is composed of the three key technologies: (a) tracking of the sensor's attitude, (b) detection of walking locomotion and (c) estimation of the walking velocity. In this section, we briefly describe the key technologies.

2.1 Tracking of the sensor's attitude

The attitude consists of the gravitational and the horizontal reference (north) components. They are estimated and updated by the Kalman filtering framework [1][2]. The gravitation can be directly observed by accelerometers and the north direction can be observed by the magnetometers. As the magnetic field measured by the sensors is often distorted by local structures of the environment, we use a selective mechanism to filter out the disturbance for the observation on the north direction [1]. The state vector of the Kalman filter is predicted by the angular velocity vector from the gyrosensors.

2.2 Detection of walking locomotion

Since human walking locomotion is highly coordinated, the pattern of motion repeatedly appears in acceleration and angular velocity. If the sensors are placed on the waist which is near the center of gravity of human body, the pattern can be easily recognized by detecting the pair of down-peak and up-peak of the components.

2.3 Estimation of the walking velocity

The walking velocity of pedestrians is empirically known to have strong linear correlation with the acceleration amplitude in the vertical direction within a walking cycle. By using linear regression, the equation to estimate the velocity from the accelerometers can be obtained. However, the linear regression parameters differ in person and thus a calibration procedure is required beforehand for more precise estimation of walking velocity.

3 ENHANCEMENT OF THE PDR

We introduce two enhancements of the PDR to improve its performance, by map matching and dynamic estimation of walking parameters from surveillance cameras.

Since the PDR is an incremental method of estimating location, its error is accumulated to significant amount over time. Therefore the accumulated error needs to be somehow reduced and corrected. We introduce the constraints of the environment (namely 2.5D map) to the PDR and mechanism to fuse the PDR results and the constraints by using particle filter.

As described in Section 2.3, the walking parameters are slightly different from person to person. Then, we use surveillance cameras to measure walking velocity in order to recalibrate the parameters. The camera can also be used to correct the estimated position by the PDR.

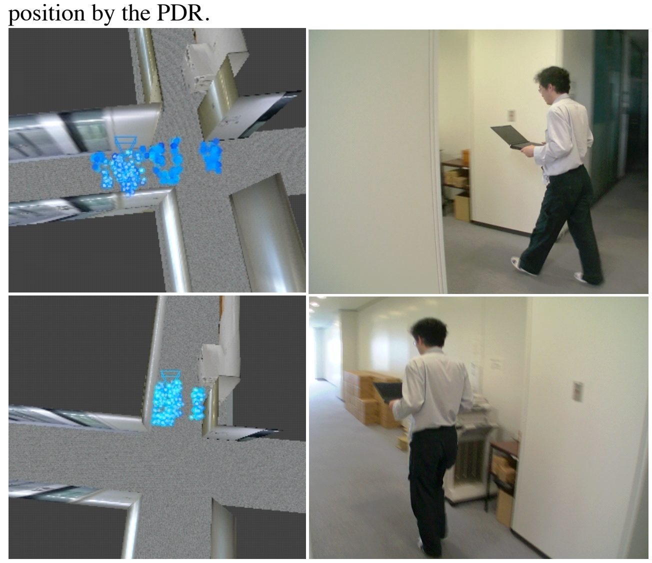


Figure 1. Map matching using a map converted from 3D environment model (Left-column presents probability distributions of locations of the user shown in right-column. Bottom-row presents the state one second before top-row).

3.1 Map matching and particle filter

The user's location and orientation are updated by fusing the measurements from the PDR estimation and 2.5D maps [4] generated from 3D environment models [6] and we use the particle filter framework for probabilistically fusing the data [5]. The particle filter which is kind of Bayesian filter efficiently estimates state of a system under the Markov assumption and Monte-Carlo approximation of probability distribution.

The state space of the particle filter is represented by the 4D vector whose components are 2D position, a polygon identification number of the 2.5D map, and an absolute orientation. In this state space, the probability distribution of the user's location is predicted from the estimated position, orientation, and its uncertainties. Note that, in the prediction process of the current probability distribution based on measurement from the PDR, a sample in those representing the probability distribution is eliminated when the displacement vector of the sample intersects with lines of walls or outgoes to outside of floors in the 2.5D map. This map matching improves the performance by utilizing knowledge of the environments. And the implementation of the map matching on a typical PC runs in real-time. The error of the PDR with map-matching is less than 5 percent of walking distance.

3.2 Dynamic estimation of walking parameters from surveillance video camera

We aim at improving localization performance by utilizing surveillance cameras used for surveillance service as existing infrastructure. The surveillance camera must be clearly placed in important areas of human traffic, so the feature is absolutely suitable for localization. The surveillance cameras are used for realizing the following two functions.

- Correction of localization error of the PDR
- Dynamic estimation of walking parameter

Camera has high-spatial resolution, and so its measurements are effective for estimating precise location of the pedestrian. In addition, the cameras can also be used to estimate the walking velocity of the pedestrian from the time-series measurements.

In order to realize above functions, our system needs to recognize the user wearing the self-contained sensor module from persons on the surveillance videos. Our system recognizes the user by matching and identifying two kinds of 2.5 D trajectories that are fusion-based trajectory and video-based trajectory. Here, the camera parameters to be used for converting 2D video-based trajectories into 2.5D video-based trajectories can be estimated by 3D environment models as contents and the 3D modeler described above from a photo without any special devices.

When the system successfully matches and identifies trajectories, the identified video-based trajectory is sent to the fusion unit for correction of localization errors. Moreover, the system estimates the walking velocity from the video-based trajectory and sends it to the part of walking parameter estimation in PDR. From the velocity and the acceleration amplitude, the system can estimate waking parameters by the method in [1][2].

4 Applications of the PDR

We have developed two types of applications of the PDR technologies mentioned above. These applications are described in this section.

4.1 The PDR Evaluation Kit

We have commercialized the PDR technologies into an evaluation kit, which allows developers to test the accuracy and availability of localization. The kit includes the PDR software and a self-contained sensor module. It can be used to develop indoor and outdoor localization system such as a pedestrian navigation system and locating (and tracking) system of factory personnel. The evaluation kit can be purchased from SHIBUYA KOGYO at around 5,000 USD.

4.2 Application toward visually impaired people navigation

We are currently working on a national project that aims to develop an advanced mobile device which supports visually impaired people to work out streets safely and comfortably. Figure 2 shows a schematic outlook of the device.

One of the problems of visually impaired people navigation is the inaccuracy and unavailability of the positioning of the GPS-based navigation system. Our new device will navigate the visually impaired people more accurately by the help of new sensors including the PDR system described in this paper. The device exploits the GPS for rough position estimation and its bootstrapping, and then the PDR system estimates the position in a map. Additional vision system also gives position correction information by checking landmarks and pre-recorded images in a scene. It also has a voice navigation interface and a refreshable braille display so that we can inform navigation instruction flexibly. The device is planned to be released next spring.

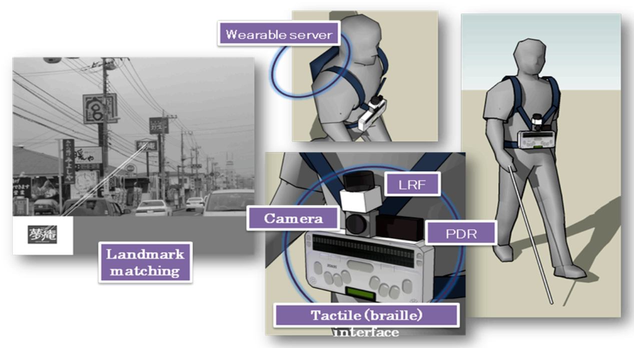


Figure 2. The navigation system for visually impaired people.

References

[1] M. Kourogi, N. Sakata, T. Okuma, and T. Kurata: "Indoor/Outdoor Pedestrian Navigation with an Embedded GPS/RFID/Self-contained Sensor System", In Proc. 16th International Conference on Artificial Reality and Telexistence (ICAT2006), pp.1310-1321 (2006).

[2] M. Kourogi, and T. Kurata, "Personal Positioning Based on Walking Locomotion Analysis with Self-Contained Sensors and a Wearable Camera", In Proc. 2nd IEEE and ACM Int. Symp. on Mixed and Augmented Reality (ISMAR2003), pp. 103-112, 2003.

[3] E. Foxlin: "Pedestrian Tracking with Shoe-Mounted Inertial Sensors", IEEE Computer Graphics and Applications, No. 6, Vol. 25, pp. 38-46, 2005.

[4] O. Woodman and R. Harle, "Pedestrian Localisation for Indoor Environments", In Proc. 10th Int. Conf. on Ubiquitous Computing (UbiComp2008), pp. 114-123, 2008.

[5] T. Ishikawa, M. Kourogi, T. Okuma, and T. Kurata, "Economic and Synergistic Pedestrian Tracking System for Indoor Environments", International Conference on Soft Computing and Pattern Recognition (SoCPaR2009), Dec., 2009 (To appear).

[6] T. Ishikawa, T. Kalaivani, M. Kourogi, A. P. Gee, W. Mayol, K. Jung, T. Kurata: "In-Situ 3D Indoor Modeler with a Camera and Self-Contained Sensors", In Proc. 13th International Conference on Human-Computer Interaction (HCII2009) in San Diego, CA, USA, LNCS 5622, pp. 454-464 (2009)

3D Pointing Interface by using Virtual Diorama for Attention Sharing

Masayuki Hayashi* — University of Tsukuba
Itaru Kitahara† — University of Tsukuba
Yoshinari Kameda‡ — University of Tsukuba
Yuichi Ohta§ — University of Tsukuba

ABSTRACT

We propose a new graphical user interface based on a Mixed-Reality (MR) technique by using a miniature CG model of surrounding environment of user. We call the CG model "Virtual Diorama (VD)." The interface easily specifies a 3D position on the surface of buildings outdoors so that others can easily understand where the point is. In cooperative task, it is important but difficult to indicate a specific point (e.g., a fourth floor window on a building) when the point is invisible from the indicator and co-workers have different viewpoints. Our system provides a user's view that is free from these problems. Our interface enables the indicator to specify any attention points including an invisible place at his actual viewpoint because he can move the VD virtually on the interface. The point is then sent and shown at the VD co-worker side. The VD overlapped on the real image of the co-worker viewpoint is aligned either to the world or to the view of the co-worker. As both alignment methods have pros and cons, we have had a user study for evaluating them.

CR Categories H.4.3 [Information Systems Applications]: Communications applications – Attention sharing; H.5.2 [Information Interfaces and Presentation]: User interfaces – Graphical user interfaces (GUI); I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism – Virtual reality

KEYWORDS Outdoor Mixed-Reality, 3D Pointing, Attention Sharing, Cooperative Task

1 INTRODUCTION

When we do some cooperative tasks such as setup of event site, we often need to draw a co-worker's attention to the place we are discussing. Ordinarily the target point is indicated by a finger or such equipment as a laser pointer. However, pointing to invisible places for users and indicating the exact location of distant objects by these ways is difficult. As a result, our communication ability is limited especially in such large-scale spaces as outdoor environments. Using maps addresses this limitation, but ordinary map has only 2D information and does not represent height information. Using a 3D map generated by CG technology is one of the solutions to support cooperative tasks in outdoor environment. In order to make users feel comfortable with browsing 3D maps, it is important to consider about the methodology to display 3D maps. Since, most users are not familiar with browsing 3D maps, even though it gives much geometrical information.

A Head Mounted Display (HMD) is a common display device for MR. In outdoor environments, however, HMD has a few problems. To conduct outdoor activities, users have to pay attention to obstacles and hazards, a task is complicated by the narrow field of view that is limited by the size of HMD's display screen. Power supply and video transmission cables also restrict user movements. Therefore, we use a mobile hand-held device with a video camera such as a Personal Digital Assistant (PDA) or a cellular phone [1] for an outdoor MR system. Users can watch the surrounding environment with their own eyes without being disturbed by cables. It is important to know that a user lose institutive understanding for spatial relationship between the real world and VD, instead of keeping their safety and comfortableness.

We propose a new attention point sharing system for outdoor environment that uses mobile hand-held device and 3D map. The system makes it possible to indicate invisible or distant places. As illustrated in Fig. 1, the system shows miniature CG models of the surrounding environment. We call the CG models "Virtual Diorama (VD)." Users can easily grasp the geometrical relationship of the virtual and real worlds by observing the VD in a MR fashion that displays the models in front of users. The scenario is as follows: When user indicates an attention point to some place on the VD by using a 2D input device such as a touch screen, the mobile device sends 3D position of the point to a server via a wireless network. The server sends the position to other users. They can see a visual icon on the position of their VD shown in their mobile device.

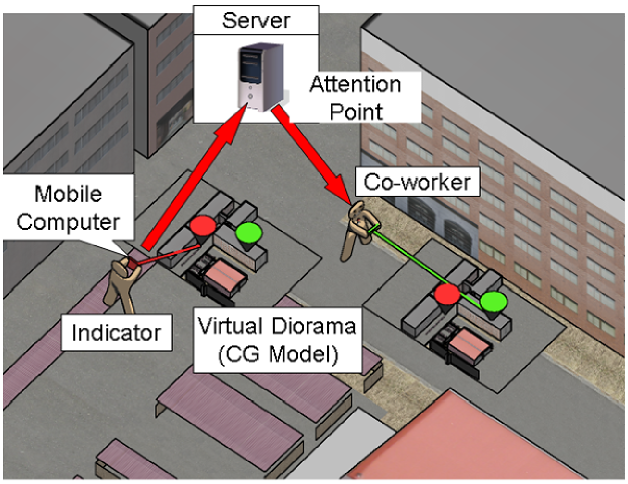


Fig. 1 Concept of Virtual Diorama. Users can observe a miniature CG model of real world around themselves through mobile hand-held device's monitor.

2 RELATED WORK

2.1 Mixed Reality Navigation System

Many mixed or augmented reality systems [1,2] have been developed that realize visual navigation in outdoor environments. For example, Kanbara et al. [2] developed a human and car navigation system for outdoor MR with a real time kinematic global positioning system and an inertial navigation system to

* e-mail: mhayashi@image.iit.tsukuba.ac.jp
† e-mail: kitahara@iit.tsukuba.ac.jp
‡ e-mail: kameda@iit.tsukuba.ac.jp
§ e-mail: ohta@acm.org

measure a user's position and orientation in real time. The system shows users such virtual objects as an arrow for navigation. Many such mixed or augmented reality systems, especially those using HMD, need to realize geometrical consistency to overlay virtual objects onto the real world. When position and orientation estimation (camera registration) is incorrect, the virtual objects are displayed at the wrong position, and navigation fails.

There are two ways to solve this problem. One is to improve camera registration accuracy by installing high-end devices [3]. The other is devising a display method that is insensitive to the camera registration error. Our proposed method is the latter case. It displays a visual icon that indicates the 3D attention points on the VD that are not in the real world. Thus, user input is always accurately allocated at the intended position, even though geometrical inconsistency is not well estimated. VD does not require accurate 3D position of mobile hand-held devices, because it works with only rough location information such as names of streets or blocks, and users set the 3D gazing position on VD by themselves using such input devices as keyboards. Therefore, our system has a feature that simplifies the camera registration problem as 3D orientation estimation of mobile hand-held devices.

2.2 Worlds in Miniature

The concept of Virtual Diorama resembles Worlds in Miniature (WIM), which was developed by Stoakley and Paush [4]. Originally, WIM was developed as a visual assistant technique for virtual reality environments. Thus, most WIM systems use HMD, and the target space is indoor environments, although there are WIM applications for outdoor environments. Höllerer et al. [5] developed a pedestrian navigation system that resembles our system. They use WIM in an MR fashion in outdoor environments by mobile hand-held devices and HMD. Our system's target space is also outdoor environments, but it uses a mobile hand-held device to show miniature CG models in an MR fashion. Moreover, different from ordinary systems, our proposed system can indicate and share 3D points with users in outdoor spaces.

3 Virtual Diorama

We propose a system that indicates and shares 3D attention points by displaying a Virtual Diorama of the real world on a mobile hand-held device, as if a real diorama exists in front of the user. Users can simultaneously observe both the real and virtual worlds. As shown in Fig. 2(a), when the orientations of the two worlds are different, users may lose their sense of orientation. To avoid this, our system aligns a VD to maintain the heading direction of both worlds coincident (Fig. 2(b)). Therefore, the system requires pose information of the mobile hand-held device. Users can intuitively control the virtual viewpoint to render the VD by moving a mobile hand-held device.

Figure 3 shows how to input an attention point using VD. When a user touches a point on the mobile hand-held device's display, the system sets a visual icon at the intersection point of a miniature 3D model and a line connecting the virtual viewpoint (e.g., point A) with the touching point. For example, if the user wants to indicate an invisible point occluded by buildings, he/she can move the virtual viewpoint so that the target point is visible, like virtual viewpoint B in Fig. 3.

The attention points given by other users are displayed as visual icons with different colors. Pose information of the mobile hand-held device is required to realize the system. Our system does not track the actual viewpoint of users to simplify the system. Thus, the virtual viewpoint is fixed in front of the mobile hand-held device's monitor.

(a) (b)

Fig. 2 Orientation issue of a miniature CG model: (a) orientation of real and virtual worlds are different. User may lose their sense of orientation: (b) orientation of two worlds are aligned to keep user's sense of orientation.

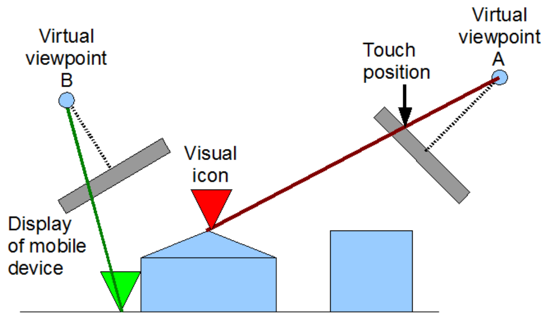


Fig. 3 Input method of attention point by Virtual Diorama. Users can input an attention point on a miniature 3D CG model. Our system shows visual icons with different colors for each user at the attention point.

4 Conclusion

We proposed a graphical user interface using a Mixed-Reality (MR) technique called "Virtual Diorama" and a system with which users can easily point and share 3D positions in outdoor spaces. Our system solves the problem of ordinary pointing methods in outdoor environment by introducing a user interface technique that displays a miniature CG model of the real world around the user on mobile hand-held devices, just as if the modeled environment exists in front of the user.

References

[1] Yoshinari Kameda, Taisuke Takemasa, and Yuichi Ohta, "Outdoor See-Through Vision Utilizing Surveillance Cameras," *Proceedings of the Third IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'04)*, 2004.

[2] Masayuki Kanbara and Naokazu Yokoya, "Outdoor Augmented Reality System Using RTK-GPS and Inertial Navigation System," *Technical Report of IEICE*, 2005.

[3] Kiyohide Satoh, Mahoro Anabuki, Hiroyuki Yamamoto, and Hideyuki Tamura, "A Hybrid Registration Method for Outdoor Augmented Reality," *Proceedings of the IEEE and ACM International Symposium on Augmented Reality (ISAR'01)*, 2001.

[4] Richard Stoakley, Matthew J. Conway, and Randy Pausch, "Virtual Reality on a WIM: Interactive Worlds in Miniature," *Proceedings of the ACM CHI'95*, 1995.

[5] Tobias Höllerer et al., "User interface management techniques for collaborative mobile augmented reality," *COMPUTERS AND GRAPHICS*, 2001.

Adaptive Visualization in Outdoor AR Displays

Denis Kalkofen* Stefanie Zollman† Gerhard Schall‡ Gerhard Reitmayr§ Dieter Schmalstieg¶

TU Graz
Graz, Austria

1 Introduction

Visualizations in an Augmented Reality environment consist of a visual interaction between the presentations of virtual and real word structure. This interaction is a key element towards the generation of comprehensible visualizations, and the AR presentation has to ensure the generation of visually harmonizing elements. Thus, both virtual and real world information have to be carefully analyzed and their visual parameters have to be adapted to fit to each other. Moreover, the dynamic and complex nature of outdoor AR environments does not allow for manual adjustments of visualization parameters. Consequently, any visualization technique which can be considered suitable for outdoor AR environments has to be able to control the interplay of real and virtual information online. In this work, we discuss visualization parameters which have to be addressed during the adaptation in an outdoor AR scenario and outline possible data sources and their potential usage to control the adaptation itself.

Current approaches [1, 2] use fixed visual designs comprising of different choices of colour, transparency or line drawing styles. which work well in indoor and static AR scenarios where colour schemes are known and environmental influences such as lighting and shadows can be controlled. To address the more complex issues of outdoor environments we envision a dynamic visualization pipeline that takes all the current available data into account. Through an on-the-fly analysis of the incoming data, we propose to derive appropriate visual composition methods to provide convincing visualizations which consist of visually harmonizing elements.

2 Adjustments

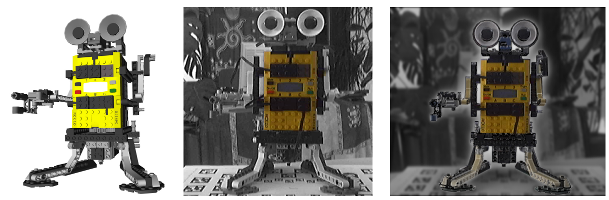
Three components of an AR visualization can be identified that influence the successful integration of virtual and real content. Firstly, these are parameters which directly control the appearance of an object (such as their material properties). Secondly, a classification of virtual and real world information controls where to apply which stylization parameter. Finally, parameters specific to the visualization technique have to be set to ensure the effectiveness of the resulting AR visualization. Notice, the order of the following subsections does not necessarily has to align with the processing pipeline in an adaptive rendering system.

2.1 Material Properties

Careless generation of visualization in AR environments may lead to misleading interactions of colors and shades representing the real and the virtual objects. Consequently, to be able to generate understandable visualizations in AR environments, the appearance of the additional 3D computer graphics has to fit into the real world environment. Both color and shading of the virtual and the real world objects have to be analyzed and adapted to correspond to each other.

Seligman et.al. [5] already addressed adaptive presentations in VR environments. In her dissertation work on intention based illustrations (IBIS) she demonstrated e.g. the rendering of an automatic

*e-mail:kalkofen@icg.tugraz.at
†e-mail:zollmann@icg.tugraz.at
‡e-mail:schall@icg.tugraz.at
§e-mail:reitmayr@icg.tugraz.at
¶e-mail:schmalstieg@icg.tugraz.at



(a) (b)

Figure 1: Different Classifications (a) Only two different classes have been selected for further stylization. While in VR environments effective visualizations are possible, the same stylization fails to communicate the intention of the visualization (b) More complex classifications preserve the information in a complex environment

highlighting of certain parts in an illustration. However, her approach did not take complex environments into account. Especially in outdoor AR it is not obvious which parameter to alter to perceive easily comprehensive results. Moreover, since visual information in an outdoor AR changes by their dynamic nature, the adaptation has to ensure visually stable results over time.

2.2 Classification

In a complex AR environment we also have to identify the right amount of characteristic structures which are considered to apply the modifications of material parameter to. For example, Fig. 1(a) shows a classification of the presented content in only two separate groups. A subsequent stylization using this classification generates an effective visualization in a VR environment but the same stylization fails in a more complex AR environment.

In comparison, Fig. 1(b) uses a more complex classification which allows to preserve characteristic structures in the close proximity of the object of interest. However this visualization rather heavily alters the appearance of the scenario which we only want to apply if less modifications fail to communicate the visualizations intent.

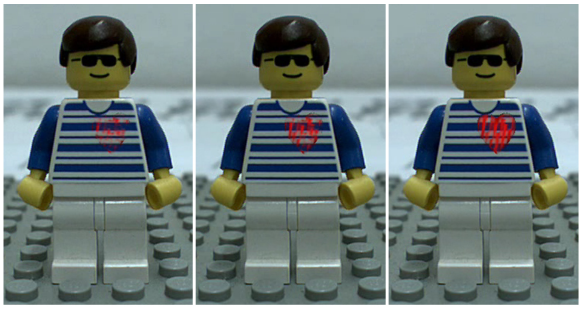
2.3 Visualization Parameters

Dynamic scenarios furthermore may demand an online adaptation of certain parameter of the visualization technique itself. For example, Fig. 2(b) shows a hatching texture which is used to control the amount of occlusion cues in an x-ray visualization. If this visualization technique is applied to a low frequent real world object the strokes in the hatching communicate effectively shape information of occluding structure. The occluding structure can mentally be reconstructed because the bending of the strokes implicitly encodes shape information. However, if the same visualization technique is applied using the same parameter on higher frequent real world object the strokes may interfere with real world texture information resulting in a cluttered presentation (Fig. 2(b)).

3 Data Analysis

Comprehensible visualization for AR was demonstrated using a registered 3D model of the environment to detect important features. However, since the necessary stylization to generate comprehensible visualizations are highly dependent on a very precise registration, these model-based approaches are more appropriate in a

(a)



(b)

Figure 2: Hatch Textures controlling the preservation of occluding structures. (a) While artificial stroke textures which follow curvature information are able to effectively preserve shape information on low frequent textures (b), the same technique fails to comprehensibly present high frequent textures which are independent of the preserved shape.

controlled environment where it is feasible to obtain a full environmental model. Outdoor scenarios neither afford sufficient control over the appearance of the environment nor are comprehensive and detailed models available. Thus these simple and static approaches are not sufficient for complex outdoor environments. Therefore we propose to directly analyse the video background to extract salient features and 2-1/2D scene representations.

3.1 Video Data

In Video-see-through-based setups scene information is readily available in the video background image. Pure image-based analysis of the scene can provide some hints for an adaptive visualization.

A simple example is the extraction of visible edges. Thereby the most important visual cues of the scene are detected and, through appropriate masking of the virtual content, are kept visible in the augmented rendering to provide a better understanding of the outdoor scene and the virtual objects spatial relationship to the scene. To avoid visual clutter the robust extraction of exclusively important cues is essential. Similar techniques were used successfully in label placement for AR applications [4].

Similarly extracting the dominant color scheme in the combined area can contribute to the adaption. While rendering a red object on green background may produce good contrast, moving to a red background will obliterate the object no matter what depth cues are provided. Here a dynamic adaption of the virtual object's color or rendering style is needed and it should take into account the actual appearance of the background.

Empty areas in the video background do not provide information about the scene depth. Adding hatchings (see Fig. 2) in these areas can also provide a better understanding of the scene, since non-textured parts of the scene are extended with a stylization.

The creation of explosion views is an example for the modification of the original video image. Explosion views transform occluders in such a way that they are still visible in the video image, but do not occlude the virtual scene. Challenges for this technique is the adaptive determination of adequate transformations of the occluder, that still provides information about its shape and avoid the occlusion of other important objects.

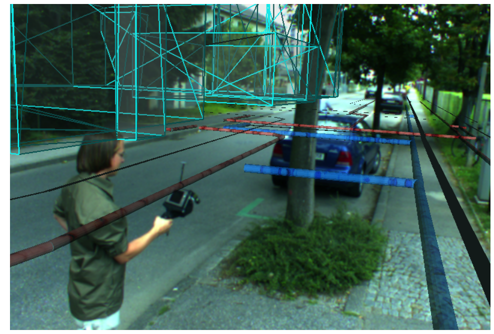


Figure 3: Visualization of subsurface infrastructure in VIDENTE, demonstrating common problems in outdoor X-ray visualization.

3.2 Adaptation using 3D models

Pure image-based methods are a first step towards adapting the visualization to the background, but having 3D model information of the real scene allows more extensive manipulations. This applies beyond the basic requirements for correct occlusions.

Explosion views in the video image plane look only correct for small camera motions. Under larger motion, both parallax as well as rotation show that the virtual plane on which the exploded parts are moved does not correspond to the depth of the real scene. Knowledge about the 3D model allows to create virtual geometry that fits the scene and move this placeholder appropriately.

Capturing the full 3D scene in real-time is still an open challenge for outdoor AR systems. Current developments include laser range finders to build a coarse model[6], or visual SLAM to create a basic model consisting of a few planes [3].

4 Application

We intend to apply various adaptive visualization techniques in an outdoor application[1] demonstrated in Figure 3. Our project aims at the visualization of underground infrastructure using mobile AR and assists field workers of utilities in common tasks such as maintenance, locating of assets and planning. In this application subsurface assets (e.g. gas pipes, water mains, telecommunication lines or sewers) are visualized below the ground level and occluded by real surfaces. To help correct depth perception we are investigating dynamic techniques that adapt to the unknown surface texture and appearance to provide occlusion cues.

5 Conclusion

To ensure comprehensible visualization in outdoor AR environments, the visualization system has to be able to analyze the current background followed by a selection of an appropriate visualization technique and parameter values. Moreover, the visualization technique should not only stylize the additional virtual data but also existing real world information. This work was funded by the Austrian Research Promotion Agency (FFG) under contract no. BRIDGE 811000, and FIT-IT 820922.

References

[1] D. Kalkofen, E. Mendez, and D. Schmalstieg. Comprehensible visualization for augmented reality. *IEEE Trans. Vis.*, 15(2):193–204, 2009.

[2] C. Bichlmeier, F. Wimme, S. Heining, and N. Navab. Contextual anatomic mimesis hybrid in-situ visualization method for improving multi-sensory depth perception in medical augmented reality. In *Proc. ISMAR 2007*, pages 129–138, Nov. 2007.

[3] A. P. Gee, D. Chekhlov, W. Mayol, and A. Calway. Discovering planes and collapsing the state space in visual slam. In *Proc. BMVC 2007*, Sept 2007.

[1] www.vidente.at

[4] A. Leykin and M. Tuceryan. Automatic determination of text readability over textured backgrounds for augmented reality systems. In *Proc. ISMAR 2004*, pages 224–230, Arlington, VA, USA, Nov 2004.

[5] D. D. Seligmann and S. Feiner. Automated generation of intent-based 3d illustrations. In *Proc. SIGGRAPH '91*, pages 123–132, New York, NY, USA, 1991.

[6] J. Wither, C. Coffin, J. Ventura, and T. Höllerer. Fast annotation and modeling with a single-point laser range finder. In *Proc. ISMAR 2008*, Cambridge, UK, Sept 15–18 2008.

Prototyping an Outdoor Mobile Augmented Reality Street View Application

Yoshitaka Tokusho[1]

Software Division, Hitachi Ltd

Steven Feiner[2]

Department of Computer Science, Columbia University

Abstract

We describe the early development of an experimental outdoor mobile augmented reality (AR) application, "AR street view," running on a Google Android Dev Phone 1. This project addresses the design of interaction and visualization techniques for AR user interfaces on mobile phones. Our experience suggests that usability is compromised by a variety of issues, including tracking and registration errors caused by the relatively inaccurate GPS and orientation sensor, the need to hold the phone away from the body in potentially awkward poses, and the relatively small field of view covered by the display and camera.

Keywords: Augmented reality, mobile user interfaces, information visualization.

Index Terms: H.5.1 [**Multimedia Information Systems**]: Artificial, augmented, and virtual realities—*Mobile augmented reality*; H.5.2 [**User Interfaces**]: Graphical user interfaces (GUI)—*3-D user interface*

1 Introduction

In this position paper, we present our early work on "AR street view" (Figure 1), an experimental outdoor mobile augmented reality (AR) application running on a Google Android Dev Phone 1, and discuss some of the issues we faced during its development. Our overarching observation is that the technical limitations imposed by current mobile phone hardware severely hinder the potential benefits of outdoor mobile-phone–based AR. To address this, it will be necessary to understand the weaknesses of current mobile devices, and how they might be overcome through improved hardware and software, as well as to develop effective design principles for outdoor mobile AR.

AR street view provides a kind of live video equivalent of Google Maps Street View. (In turn, one can think Google Maps Street View as an example of simulated AR, overlaying interactive graphical annotations on prerecorded panoramic stills [3], instead of on a live view of the environment.) AR street view aims to provide users with an intuitive sense of the relevant geographic information surrounding them, especially in urban environments. Targets for visualization and interaction include streets; addresses; building facades; placemarks for key locations and events, such as local businesses, public and government facilities, and festivals; representations of physical objects, such as historic structures, and virtual objects, such as routes; and all manner of site-specific documentation, such as bus schedules, restaurant reviews, and historical facts. Usage scenarios include local exploration and navigation. For example, a salesperson visiting a customer for the first time might be able to find a path to and identify the customer's building more quickly through information overlaid on a live view, instead of on older outdated images.

Thus far, we have implemented a basic egocentric AR view of streets and labels on a position- and orientation-tracked Android phone, and constructed an experimental street database server. Our project focuses on research in interaction and visualization for outdoor mobile phone AR user interfaces (UIs), rather than on outdoor registration and tracking for these devices. However, to develop UIs that will be appropriate for future AR systems, we are experimenting with better external sensors (e.g., the orientation tracker of Figure 2), to approximate the performance of future built-in technology.

Figure 1. AR street view running on Google Android Dev Phone 1. Representations of streets and their names are shown near a street corner in midtown Manhattan.

Figure 2. Application system components. Top to bottom: UMPC with Bluetooth adapter, Android Dev Phone 1, InertiaCube3 external orientation sensor.

Figure 3. AR street view 3D components. (a) Street objects are registered with the real world and annotated with labels. (b) Compass reticle shows heading. (c) Information display shows the quantitative values of GPS, orientation, and other sensor observations.

[1] yoshitaka.tokusho.zp@hitachi.com
[2] feiner@cs.columbia.edu

2 RELATED WORK

Our work builds in part on over 13 years of mobile AR research at Columbia, beginning with the development of the first bulky backpack-based mobile AR system with a head-worn display [1]. The current project is our first attempt to extend this research to smartphones, from the backpacks and large hand-held devices that we used before. It is interesting to note that while current smartphones have, as we might expect, far better processing power and wireless communications than our backpack systems, the position and orientation trackers that we were using over a decade ago are significantly *better* than what is in current smartphones.

A wide variety of research from other labs has addressed mobile AR on platforms ranging from backpack computers to smartphones (e.g., [2,5,6]). Over the past year, commercial mobile AR applications have appeared for smartphones (e.g., Wikitude and Layar).

3 USER INTERFACE DESIGN

AR street view provides a 3D egocentric view of the user's surroundings. The three primary 3D visual components overlaid on the camera view in our preliminary work include simplified street geometry, a compass reticle, and information labels (Figure 3).

The street geometry is made up of simplified linear representations of streets, annotated with street name labels. A compass reticle at the bottom of the display shows the direction in which the phone is pointed. Information labels indicate the quantitative values of GPS and sensor observations for diagnostic purposes, and are fixed to the 2D display coordinate system.

4 IMPLEMENTATION

Our prototype system consists of four components: AR street view running on the Google Android Dev Phone 1, the street database server, an external sensor, and a UMPC (or laptop PC) equipped with Bluetooth adapter (Figure 2).

AR street view is built on Google Android platform 1.5, supporting a mobile device equipped with a camera, Bluetooth adapter, GPS receiver, orientation sensor and cellular or Wi-Fi network adapter. 3D graphics are drawn by OpenGL ES 1.0. Registration and tracking are implemented using either the built-in location and orientation sensors, or external sensors.

The *street database server* manages the street coordinates (and attributes data for the US, decoded from the US Census 2006 TIGER/Line Files. The street database is queried by the MySQL 5.1 spatial extension, and accessed via HTTP from application.

We use the InterSense InertiaCube3 (IC3) orientation sensor device as the additional external sensor. Data generated by the IC3 is transferred via Bluetooth from IC3-connected UMPC to AR street view application.

5 USABILITY ISSUES

5.1 Position Inaccuracy

The position tracking errors introduced by the relatively inexpensive location sensors used in smartphones compromise the usability of applications that rely on them. For example, in our experience, the A-GPS in the Android Dev Phone 1 can sometimes report more than a 40 m offset from its actual position. To address this in our testbed, we are taking advantage of external sensors that we have used in our earlier backpack and hand-held AR research: an external RTK GNSS position tracking system, and a dead-reckoning module. When the user's position cannot be accurately obtained from available satellites, the dead-reckoning module uses waist-worn accelerometers and gyroscopes to implement a digital pedometer, counting the user's steps and the directions in which each is taken, to estimate the user's changing position.

5.2 Orientation Inaccuracy

Similarly, the phone's built-in orientation sensor can report significant static and dynamic orientation errors, off by tens of degrees. We have addressed this by using an InterSense IC3 orientation tracker rigidly fastened to the phone, but offset to avoid electrical interference.

An even more appealing approach to improving both position and orientation tracking will be to use camera-based feature tracking [4 ,6], with or without models of the environment.

5.3 Small Field of View

The optics used in smartphone cameras typically produce an image with a wider field of view than would be seen looking directly through a rectangular frame the size of the display. However, this is still far smaller than the field of view of even a single human eye. Thus, only a relatively small portion of the user's field of view can be visibly augmented at any point in time.

5.4 Awkward Pose

The incorporation of a camera on the opposite side from the display encourages a "magic lens" metaphor that is a defining characteristic of most smartphone AR systems. However, this design requires the user to assume a rather awkward pose when viewing relatively distant real world objects—the phone held in an outstretched hand, sufficiently far from the eye to focus on the display, and often angled uncomfortably upward (e.g., to overlay a tall building). This is the stereotypical smartphone AR equivalent of the texting stance of the Blackberry user. Partial solutions include using an angled camera mount or "freezing" the camera image to allow the user to see the desired augmented view while holding the phone in a comfortable pose. An alternative will be made possible by lightweight eyewear.

6 CONCLUSIONS

We have described our early work on an experimental outdoor smartphone AR application, and surveyed some of the issues caused by current smartphones. We are continuing to develop our system, and, in order to better understand usability issues, we have designed a user study that we will be conducting shortly.

ACKNOWLEDGMENTS

We thank Mengu Sukan for assistance in developing the software. This work was supported in part by a Google Research Award.

REFERENCES

[1] Feiner, S., MacIntyre, B., Höllerer, T., and Webster, A. A touring machine: Prototyping 3D mobile augmented reality systems for exploring the urban environment. *Proc. ISWC '97,* 74–81.

[2] Fröhlich, P., Simon, R. Baillie, L. and Anegg, H. Comparing conceptual designs for mobile access to geo-spatial information. *Proc. 8th Conf. on Human-Computer Interaction with Mobile Devices and Services*, 109–112. ACM, New York, NY, , 2006.

[3] Güven, S. and Feiner, S. Authoring 3D hypermedia for wearable augmented and virtual reality. Proc. ISWC 2003, 118-126.

[4] Klein, G. and Murray, D. Parallel tracking and mapping for small AR workspaces. *Proc. ISMAR 2007*, 1–10. 2007.

[5] Reitmayr, G. and Schmalstieg, D. Collaborative augmented reality for outdoor navigation and information browsing. *Proc. Symp. on Location Based Services and TeleCartography*, 31–41. 2004.

[6] Wagner, D., Langlotz, T. and Schmalstieg, D. Robust and unobtrusive marker tracking on mobile phones. *Proc. ISMAR 2008*, 121–124. 2008.

Interaction with the Environment: Sensor Data Visualization in Outdoor Augmented Reality

Sean White

Department of Computer Science, Columbia University
Department of Botany, Smithsonian Institution

ABSTRACT

Proliferating sensor systems provide a rich source of data for understanding the dynamic characteristics of the environment. Visualization of, and interaction with, such data in outdoor augmented reality poses several challenges for infrastructure, representation, interactive data collection, and information discovery and filtering. We discuss these challenges in the context of our experiences with an example system.

KEYWORDS: outdoor augmented reality, situated visualization, sensors

INDEX TERMS: H.5.1 [**Information Interfaces and Presentation**]: Multi-media Information Systems—*Artificial, augmented, and virtual realities*; H.5.2 [**Information Interfaces and Presentation**]: User Interfaces—*GUI*

1 INTRODUCTION

In our vision of a future outdoor augmented reality (AR), the world is filled with a media ecology of many co-existing experiences. These experiences include canonical example applications such as annotated buildings or points of interest in urban settings, first person driving or walking guidance (wayfinding), interactive games, helpful avatars and agents, historical or informational overlays on modern sites, and visualizations about the surrounding environment.

Each of these different experiences provides new challenges for the AR community and requires new technical advances. Many of the existing outdoor AR challenges are familiar. How do we make optical or video see-through displays that work well in sunlit outdoor conditions? How can we achieve millimeter tracking and registration in urban canyons and difficult settings? In this position paper, we focus on a distinct aspect of the outdoor AR experience: acquiring and visualizing sensor data as a means of enhancing our perception of the physical world.

Consider the situated visualization example of carbon monoxide (CO) in our SiteLens system (Figure 1) [8]. From our use and evaluation of the system, several new challenges arise when investigating the integration of sensors, not for tracking, but rather for enhancing perception of the world.

- What architecture and infrastructure best supports sensor data visualization?
- How are the sensors placed in the physical world? Are they mobile or stationary? Can they be used in tandem with visualization to direct gathering new data?
- What are the best representations and layout for individual visualizations and collections for related and unrelated visualizations?

swhite@cs.columbia.edu

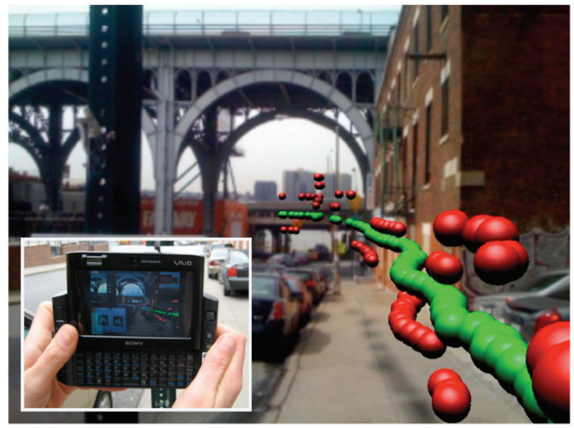


Figure 1. Example situated visualization of carbon monoxide (CO) data. Altitude of spheres represents CO ppm level. Position of spheres represent location of sensor during data acquisition.

- Given a proliferation of sensor systems, how is a given set of data discovered and how are vast collections filtered?

2 ARCHITECTURE AND INFRASTRUCTURE

Our current architecture maintains a collection of visualizations under a visualization manager and data input from a data importer and context services (Figure 2) [7]. While the architecture supports multiple visualizations in the scene, it does not support the dynamic discovery or creation of data as part of any larger infrastructure.

Reitmayr and Schmalsteig developed an infrastructure for outdoor information browsing that was meant to address large scale data sets[5]. However, the data in their system was static, not dynamic as one would find in sensor data.

Existing iinfrastructure for maintaining scalable representations

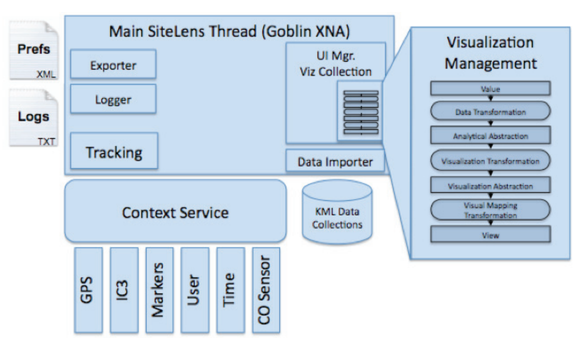


Figure 2. SiteLens architecture diagram

of such dynamic data sources does not meet the needs of AR visualization. Chang et al. have developed the SensorBase system [2] which associates sensor data blogs with geocodes, but groups of sensors are not related and characteristics that might be useful for AR visualization, such as orientation or preferred visual representation, are not present.

3 Scene Aware Layout and Representations

The specific visual representations and spatial layout of sensor data visualizations are based on the dynamics of the scene and the data itself. Moreover, sensor data visualization is distinct from annotation or display of models in that data may often be interpolated across multiple sensors or extrapolated into simulations.

Spatial layout of labels and annotations in AR typically takes into account location and orientation of other objects in the visual field based on existing models of the scene [1]. However, we believe a combination of cues that include the underlying content and texture of the scene together with spatial layout of existing objects will provide a better guide for layout of situated visualization. To achieve this, algorithms for presentation, layout, and even rendering should be guided by perceptual and cognitive principles for combining the physical with the virtual.

While Schall et al [6] have developed techniques for visualizing representations of static models, the challenges of dynamic representations may need to take the scene into account. Kalkofen [4] approaches this aspect of representation by looking at the context of the focus of attention in complex models, but remains within the bounds of static scenes.

4 Data Painting and Directed Sensing

One interesting observation regarding the SiteLens system was that the data was considered “stale” because it had been collected a month prior to the study. As the following quote indicates, participants in our study expressed a desire to have live or dynamic sensing coupled with existing data:

> It could still add more information to the analysis in different ways, ways that turn passive observation into overt surveillance […] to generate more qualitative information about a place, like how people perceive the environment, or how people sense pollution without really knowing if it’s there. These perceptions could help with the phase of design in which interviews and site surveys are done, but only with a handful of people. (anonymous participant).

Based on initial usage of the SiteLens prototype, we are interested in increasing the dynamics and symmetry of sensing and visualization by extending the system to live sensor data. In doing this, we want to close the gap between the act of sensing and the act of visualization. One approach to this involves the use of mobile sensors, which can be used for “painting” data as it is sensed, in real time, on the scene. Feedback from our users suggests that this would provide a way for them to further explore unknown regions by guiding them towards areas of interest through their own actions. We have prototyped an example mobile sensor which can be used together with a fiducial marker or similar tracker (Figure 3).

5 Discovery and Filtering of Visualizations

In our vision of situated visualization, a user will eventually have easy access to a wealth of information for any given object or scene in the world. This, combined with large data sets, will certainly lead to information overload. To address this, we see a

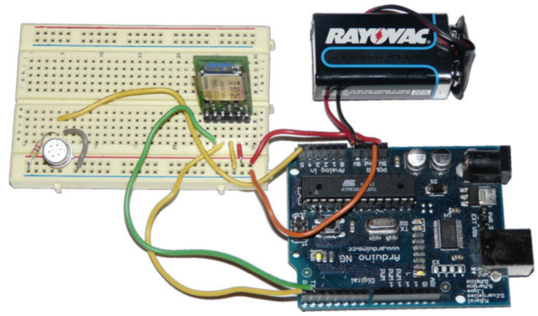


Figure 3. Wireless Bluetooth CO sensor prototype.

need for algorithms and infrastructure for discovering the presence of hidden information and filtering visualizations so that only the salient and useful information is left. Julier et al. [3] have provided some initial work in this area, but we see a need for a comprehensive solution to the problem both in terms of the algorithms and user interface. Our initial interface uses proximity for discovery and simple selection for filtering. However, we must look beyond interfaces that simply provide hierarchical lists for selection and filtering.

6 Conclusions

In this position paper, we have discussed several challenges for sensor data visualization in outdoor augmented reality. In particular, we have focused on architecture and infrastructure, data layout and representation, dynamic and directed sensing, and discovery and filtering issues.

While our initial example here deals with CO sensor data, we plan to explore this research in the context of a wider variety of sensor types and applications including geocoded botanical species identification.

Acknowledgements

This research was funded in part by NSF Grant IIS-03-25867 and a grant from the Boston Society of Architects. Steven Feiner, Sarah Williams, Petia Morozov, Ohan Oda, and Steve Henderson have all informed this work through invaluable discussions.

References

[1] B. Bell, S. Feiner, and T. Höllerer, "View Management for Virtual and Augmented Reality," *Proc. ACM UIST*, 2001, pp. 101–110.

[2] K. Chang, N. Yau, M. Hansen, and D. Estrin, "Sensorbase.org-a centralized repository to slog sensor network data," Center for Embedded Network Sensing 2006.

[3] S. Julier, M. Lanzagorta, Y. Baillot, L. Rosenblum, S. Feiner, T. Höllerer, and S. Sestito, "Information filtering for mobile augmented reality," in *Proc. ACM and IEEE ISAR*, 2000, pp. 3-11.

[4] D. Kalkofen, E. Mendez, and D. Schmalstieg, "Interactive Focus and Context Visualization for Augmented Reality," *Proc. IEEE and ACM ISMAR*, Nara, Japan, 2007, pp. 191-200.

[5] G. Reitmayr and D. Schmalstieg, "Collaborative Augmented Reality for Outdoor Navigation and Information Browsing," *Proc. Symp. Location Based Services and TeleCartography*, pp. 31-41, 2004.

[6] G. Schall, E. Mendez, E. Kruijff, E. Veas, S. Junghanns, B. Reitinger, and D. Schmalstieg, "Handheld Augmented Reality for Underground Infrastructure Visualization," *Journal of Personal and Ubiquitous Computing*, vol. 13, pp. 281-291, May 2009.

[7] S. White, "Interaction and Presentation Techniques for Situated Visualizations," PhD Diss., New York: Columbia University, 2009.

[8] S. White and S. Feiner, "SiteLens: situated visualization techniques for urban site visits," *Proc. ACM CHI*, Boston, MA, USA, 2009, pp. 1117-1120.

AR-based GIS: Towards closing the loop between acquisition and visualization of geographical data

Guillaume Moreau*
Ecole Centrale de Nantes - CERMA

Myriam Servières†
Ecole Centrale de Nantes - CERMA

Hideo Saito‡
Keio University

ABSTRACT

This paper presents the challenges of outdoor AR-based Geographical Information Systems and some ideas to address those issues. The main challenges of AR-based GIS are acquisition and visualization of information. In the first case, arises the problem of geo-referencing images to enrich the GIS. In the latter, the problem of pose computation in AR can be seen as the same problem.

1 INTRODUCTION

Sustainable development is raising new issues in urban data modeling, processing, acquisition and visualization: not only are data getting 3D and time-evolving but they are also coming from heterogeneous sources at different scales and qualities. Queries also evolve as demands become more complex: for instance, we have to face the construction of comfort indicators at the intersection of lighting, wind and acoustic phenomena.

In this paper, we focus on outdoor Augmented Reality (display of synthetic information onto the real world) for acquiring and visualizing urban data which are major challenges of 3D urban data. Without AR, in the first case, data is acquired on site, geo-referenced and then introduced in the urban GIS later. In the latter, the lack of landmarks and the difficulty of representing physical phenomena (3D+t semiology) in classical virtual environments are the most common issues. Therefore, we propose to use Augmented Reality to solve some of these issues. Both acquisition (including geo-referencing) and visualization of urban data can be performed on-site, thus overcoming previous difficulties. However, outdoor AR is a difficult challenge. It is not possible to use classical markers and registering with image databases can be too costly both in space and time. We believe that the link between images of the site, the 3D urban GIS can be seen as a closed loop: not only geo-referenced images are used to build 3D GIS but the GIS can also be used to geo-reference images [1]. We will show how we geo-reference images and compute AR-suitable poses from those images thanks to GIS information.

The rest of the paper is organized as follows: in section 2, we present our vision for AR-based 3D-GIS and the challenges that we have to address. In section 3, we will briefly present two different approaches of geolocalization from images. Finally, in section 4, we will discuss visualization issues based on the existing contradiction between current virtual environments and 2D maps and will give a short introduction on what we call 3D background maps.

2 OUR VISION

City authorities have built 2D GIS (including 3D information as attribute information). They now want to have on-site use of GIS in order to let their staff consult and interact with GIS data including

*e-mail: guillaume.moreau@ec-nantes.fr
†e-mail:myriam.servieres@ec-nantes.fr
‡e-mail:saito@ozawa.ics.keio.ac.jp

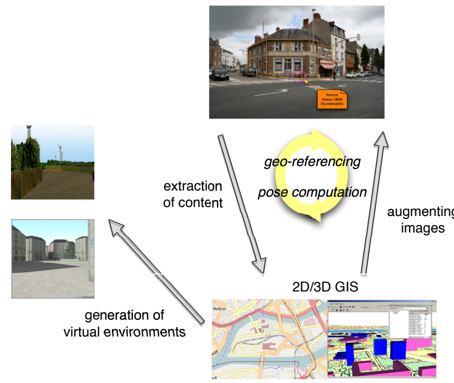


Figure 1: AR-based GIS.

3D parts such as annotation in 3D space or visualization of underground data. Given the fact, that nowadays there exists at least a partial 2D GIS of any city, we assume that starting from this 2D GIS base, 3D-GIS can either represented or enriched by digital images: on one side, images of the real world can be processed to extract information that will be inserted into the GIS (see figure 1). However, this information needs to be geo-referenced, i.e. we need to know its spatial position and extent. On the other side, from a 3D GIS, it is possible to generate classical virtual urban environments but also to augment images in order to bring GIS data on-site, i.e. in the street. To augment an image, it is required to compute the pose of the camera in world coordinates in order to superimpose synthetic information from the GIS. This process is generally separated into two different parts: a first initialization phase which is generally manual (GPS are not enough accurate and don't give orientation) and a tracking phase which has been extensively covered in the literature. In this paper, although we have also worked on tracking for outdoor AR [5], we will try to address the initialization problem and consider pose computation and geo-referencing of images as being the same problem that can be solved by cooperative use of computer vision and GIS.

3 GEO-LOCALIZATION AND POSE COMPUTATION

In this section, we will briefly describe two geo-localization techniques: one based on a GIS approach and one on computer-vision.

3.1 GIS-based approach

This approach relies on extracting some local geometric features (from the seen buildings) in the images and making spatial queries in the GIS that provide with candidate places and orientations where the picture might have been taken. Results are then filtered to reduce the number of possible locations using other techniques. For instance, in [1], we use angles between neighboring facades and length ratio between neighboring facades to query the GIS with a computed tolerance. As filters, we use 2D-visibility checking, road presence assessment. From this location, we can compute the pose of the camera. This pose can be used for initialize an AR system.

We have tested this method on various photos of the city of Nantes and getting interesting results. From now on, we are starting to use multiple images to restrict the number of poses in some cases. Yet, the main difficulty with this method is non up-to-date GIS.

3.2 Using an image database

This second method [6] is based on the fact that there exist geo-referenced image databases such a Google Street View. Google Street View consists in a huge set of spherical views. The idea is to start from a query image and to find this image in the set of panoramic images and then to compute the pose of the query image within the panoramic image. This is made possible by the matching of the query image and panoramics using SURF descriptors and then projecting the query in the panoramic image to compute the camera pose. While being quite tolerant to position differences with the panorama center, this method has to be enhanced for robustness to illumination changes, scene changes and weather conditions.

4 The display issue: 3D background maps

Assuming that we have a 3D+t urban GIS, we have to cope with display issues. Number of studies have pointed out the difficulty of self-orientation and localization in virtual environments. Moreover, another issue is dealing with information representation; while a 2D map is mainly an abstraction of the real word, virtual environments (and especially virtual cities) tend to be as most photo-realistic as possible which is a severe contradiction with 2D maps and results in too important information density which prevents for displaying additional useful information. This raises the issue of the *background map*: in 2D, to display cartographic information (such as a zone), it is common to use a background map to localize the data. In 3D, a background map can be either *natural*, i.e. additional information is displayed on a natural background. This is the principle of augmented reality and especially outdoor augmented reality. Or the background map can be *synthetic*, i.e. additional information is displayed in a synthetic environment (classical digital cities).

To solve the issues raised by self-orientation and information density, we plan to perform non only outdoor augmented reality but also to continue to use 2D maps in a simple desktop AR context.

4.1 Augmented maps: first steps

The main assumptions of this approach are that 2D maps can be usable tangible interface for people that are not used to virtual environments. The idea is augment a 2D map with 3D information, the user still being able to look at the map and to manipulate it. Augmented maps for mobile phones have already been proposed but they either use maps with AR-markers or are simply image based. Here [4], we compute the camera pose as another image geo-localization based problem: from the distribution of intersections seen on the query image and from the GIS, we make use of local geometric hashing to retrieve matchings between image and the GIS in order to compute the camera pose in the GIS frame. This allows us to display additional 3D information superimposed on the 2D map as seen in figure 2. This method is robust to occlusions but at this step, we have only tackled the tracking issue; it remains to find use cases and to make user studies to check whether tangible maps might be of use for displaying 3D information. We think of augmented maps as a good mean to represent evolving watersheds (somewhat 3D) on peri-urban areas whereas they are fast evolving due to roads and urbanization.

4.2 Displaying 3D+t information

As mentioned before, information density and 3D+t semiology are the key issues for displaying 3D+t GIS information. Moreover, the introduction of abstraction looks like a necessity to limit information density and improve understanding of urban virtual environments. To this end, we plan to develop *expressive* rendering, i.e. sketch-based rendering for urban data [3]. Based on the idea that sketching is a form of abstraction [2], we assume that this will provide a way to reduce information density with respect to classical photo-realistic digital cities thus providing a support to display new information. We also think that *expressive* rendering also has a potential to help finding the missing landmarks that reduce self-orientation capabilities while reducing rendering computation times. Next, we will study the semiology of 3D+t information, i.e. how information displayed in AR/MR is understood.

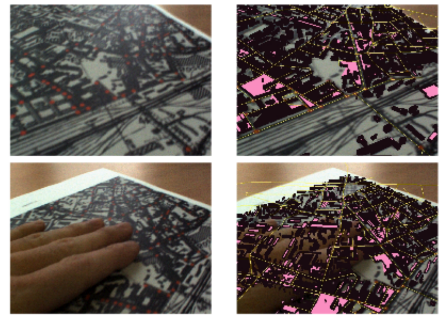

(a) Input (b) Output

Figure 2: Augmented maps.

5 Conclusion and future works

In this paper, we have briefly presented our works around the link existing between digital images and 3D+t Geographical Information Systems. We showed that AR-based GIS main challenges dealt with pose computation/geo-referencing images and information visualization and gave a few hints to solve those issues. Yet, we have just started working on information insertion and updating 3D GIS using AR and map constraints. Therefore, in the future, we will focus on both interaction with GIS data in AR context and on information visualization.

Acknowledgements

The authors wish to thank Hideaki Uchiyama and Nicolas Bioret for their research contributions to geo-localization.

References

[1] N. Bioret, G. Moreau, and M. Servières. Outdoor localization based on image/GIS correspondence using a simple 2d building layer. In T. Badard and S. Daniel, editors, *2nd International Workshop on Mobile Geospatial Augmented Reality*, Lecture notes in Geoinformation and Cartography, Québec, Canada, aug 2008. Springer-Verlag.

[2] B. Elias and V. Paelke. User-centered design of landmark visualization. In *Map-based Mobile Services - Design, Interaction and Usability*. Springer-Verlag, 2008.

[3] J.-C. Quillet, G. Thomas, X. Granier, P. Guitton, and J.-E. Marvie. Using expressive rendering for remote visualization of large city models. In *Web3D '06: Proceedings of the eleventh international conference on 3D web technology*, pages 27–35. ACM, apr 2006.

[4] H. Uchiyama, H. Saito, M. Servières, and G. Moreau. Ar gis on a physical map based on map image retrieval using llah tracking. In *IAPR Conference on Machine Vision Application*, Yokohama, Japan, may 2009.

[5] Y. Uematsu, H. Saito, M. Servières, and G. Moreau. AR mapping of GIS information by pattern-based tracking with particle filter. In *International Conference on Artificial Reality and Telexistence*, pages 305–308, Yokohama, Japan, dec 2008. VRSJ.

[6] N. Yazawa, H. Uchiyama, H. Saito, M. Servières, and G. Moreau. Image based view localization system retrieving from a panorama database by SURF. In *IAPR Conference on Machine Vision Application*, Yokohama, Japan, may 2009.

Personal Wireless Panoramic Video Tele-Augmented Reality Capable Communications Devices and Networks: "The Killer App for VR/AR"

Kurtis J. Ritchey

U.S. Army, National Simulation Center, kurtis.ritchey@us.army.mil, Phone 913-684-8147

ABSTRACT

Abstract: Video Telepresence, specifically embodiments of Immersive Interactive Spherical Panoramic Video Telecommunication's over a hand held wireless wearable cell-phone, PDA, or HMD, for 1-way, 2-way, and multi-point telecommunication is poised to become the "Killer Application" which the Virtual Reality industry has been awaiting for many years. This presentation will briefly discuss the historical developments in computers, video, 3G and 4G telecommunications technology, and societal changes that have contributed to this opportunity. The author's early pioneering patents 5,130,794 and 5,495,576 work on an associated Spherical RealityCam, VideoRoom, RealityRoom, and Head-Mounted Display will be mentioned. A demonstration of a new low cost spherical panoramic adapter lens, referred to as Virtual Video "RealityLens", that mounts onto a standard HDTV consumer camcorder for creating mixed media, virtual, and augmented realities will be demonstrated along with an example movie. Additionally, work the author has recently been involved with at the U.S. Army National Simulation Center at Fort Leavenworth, KS and it's implications for future combat operations and training will be presented. The presentation will conclude by discussing how the integration these efforts are leading to the killer application mentioned in the first sentence of this abstract; including the limitations and advantages of previous work, and opportunities for future research and development in video telepresence.

KEYWORDS: Virtual Reality, VR, Augmented Reality, AR, Telepresence, Wireless, Cell Phone, Personal Digital Assistant, PDA, head mounted display, HMD, wearable, hand held, Military, Army, Society, Immersive, geospatial, Interactive, Panoramic, Spherical, Video, Camera, Lens Computer, Telecommunications, 3G, 4G, Future, 3-D, history, pioneer, "Killer Application".

1 INTRODUCTION

"The Basis for Design" is mankind's continuing endeavor to make things that overcome their human physical, mental and emotional limitations. Incorporating Augmented Reality with Panoramic Video over wireless networks can overcome many current limitations and can greatly enhance situational awareness and social interaction between people located in different locations.

Early in the 1980's Jaron Lanier popularized the term "Virtual Reality" and was a pioneer in the field. Both Jaron and this author presented at the 1992 SPIE Conference in San Jose. While Jaron's focus was more on computer graphics, the authors focus was on using texture mapped video to create virtual environments.

Street Address and Electronic Mail Address

Other pioneers the author drew inspiration from Dr Furness in the 1980s who developed a large HMD system where a pilot could fly in a computer graphic environment that presented flight and threat information over terrain graphics. An important aspect of this work was that the graphics presented some things the pilot could not see with his innate senses like threat radar coverage. This same concept would later be a mainstay of Augmented Reality to increase situational awareness when orienting oneself using a PDA. Other pioneers included Dr. McGrevy of NASA texture mapping the Mars surface to create a virtual Martian landscape.

1.1 EARLY WORK

Figure 1. Left: Early Spherical Panoramic Camera, PC Laptop, and HMD Viewing System; Right: VideoRoom and RealityRoom.

COTS technology used by the author in a novel manner in the 1980s that was incorporated to create the first spherical FOV virtual video camera and simulated worlds included Quantel Ltd of England's television digital effects generators to texture map and distort live and recorded video on 3d objects to create realistic VR. And Phillip's of the Netherlands Video Wall controller technology that was instead used to distribute projectors and/or flat panel displays completely around participants instead of just in front of them. (See Fig 1)

Figure 2. Top Right: New Spherical Video Panoramic Adapter Lens for HDV Camcorder; Top Left: Raw Image from Spherical Adapter Lens; Bottom: Undistorted and Stitched Image from

Spherical Adapter ready for Augmented Reality graphic overlays, ROI sampling, and transmission to a remote viewer with an immersive PDA, cellphone, or HMD viewing capability.

1.2 RECENT WORK

Making spherical video accessible to the mass market.has been the goal of the author in more recent years. Powerful smaller computers, low cost high definition cameras, 3-D trackers and low cost see-thru head mounted displays have made this possible, as evidenced by the spherical camcorder in Figure 2a, 2b, and 2c. Even more recent work was done by the military in 2008 at the U.S. Army C4I BattleLab. The Command Post of the Future with Personal Assistant that Learns Experiment incorporated AI, Augmented Reality and interactive multi-media technology that demonstrated that machine augmentation of a Brigade staff conducting combat operations could increase situational awareness, understanding, and timeliness by several fold.

1.3 FUTURE DIRECTIONS

The above mentioned and below described VR/AR related technologies now enable a new paradigm for social interaction over personal wireless panoramic video tele-augmented capable devices and networks shown in Figures 4, 5, and 6. As shown in Figure 3 the new enabling technologies include RF soldier helmet & UAV cams like used in Iraq, new 3G and 4G cellular phone PDA, and HMD technology, HDTV ROI camera sensors, and geospatial location systems, and AR overlay capabilities on cell phone and PDA's which have been recently demonstrated.

Figure 3. New enabling technologies for 3-D panoramic video over a wireless network for social interaction over a personal wireless communication device.

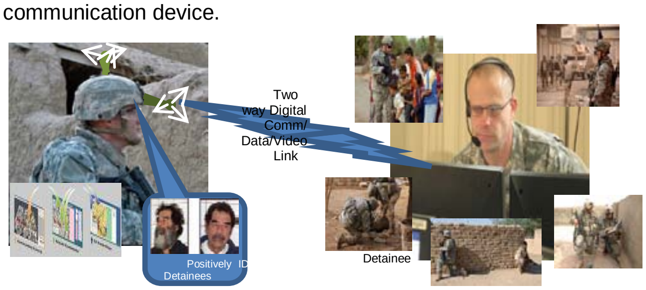


Figure 4. Illustrated here is how 2-way situational awareness and understanding are improved between Soldier and the Command Post of the Future when 3-D panoramic video and augmented reality inare incorporated in combat, peace keeping, and assistance operations. Two-way Spherical Panoramic Video will improve on the U.S. Army's initiative to make "Every Soldier a Sensor".

Virtual Enemy Displayed on Visor of actual soldier.

Actual Soldier Training

Figure 5. The same basic 2-way HMD augmented reality technology used for combat operations shown in Figure 5 can be incorporated for training. Here an actual soldier has a video simulating an attacking terrorist projected onto his HMD.

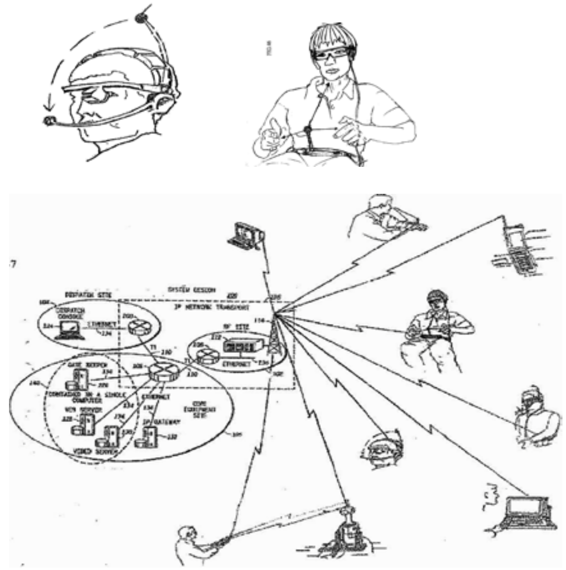


Figure 6. Upper Right: Spherical Panoramic Video Camera with HMD System for remote 2-way video telepresence and AR. Upper Left; Same used in a Gaming Application where the interface is the spherical camera which tracks the users body movements for an immersive HMD interactive gaming experience; Bottom: Diagram of the proposed wireless 3-D video and information network for social interaction on a 3-D multi-media capable personal wireless communication device.

2 DISCUSSION

There are still many research and development opportunities that need addressed in order to field the above described capabilities. Specifically, hardware and software solutions addressing image stabilization and hot-spot problems need to be developed. Also, high resolution sensors of at least 1920 x 1080 pixel resolution sensors need to be integrated into cell phones, PDA's and HMD's. A 2K x 4 K pixel array would is preferred. Of course, using high resolution sensors or multiple sensors brings up bandwidth problems transmitted across the network. A solution to this problem the author has discovered is to only sample out and transmit the Region-of-Interest (ROI) the user or users viewing the panoramic scene locally or at the other end want to see and hear. This dramatically reduces the bandwidth requirements for video telepresence. There are many HD sensors on the market that allow multiple ROI sampling for incorporation into a spherical FOV capable video camera PDA, cell phone, or HMD. And there are numerous head and eye tracking devices that could be integrated with the camera to output coordinates to tell the camera what ROI'simagery to sample out, stitch, and transmit downrange. Down range a server or remote user can add graphics or other video onto the existing incoming video and communicate back.

3 CONCLUSION

Being able to "put yourself in the other person's shoes" is what this virtual video reality and augmented reality technology allows people to do. This powerful technology has application in both the military and public sectors. Personal 3-D wireless VR/AR is the "killer application" the VR industry has long been awaiting.

Augmented Reality for Building and Construction – Software Architecture

Charles Woodward, Mika Hakkarainen, Kari Rainio

VTT Technical Research Centre of Finland

ABSTRACT

We describe a software architecture for providing mobile user at the construction site with two-way real-time augmented reality access to 4D CAD and BIM information. The system allows the user to compare scheduled building information models with the situation at the construction site in real time, as well as to attach position and time aligned visual and other feedback to the building models at the office.

1 INTRODUCTION

Building and construction is widely recognized as one of the most promising application fields for Augmented Reality [1, 2]. Recent developments in mobile processing units, camera quality, different sensors, wireless infrastructure and tracking technology enable AR applications to be implemented even in demanding mobile environments [3]. Building Information Models (BIM) are another main technology driver increasingly used for data sharing and communication purposes in real estate and construction sector [4]. Combined with Augmented Reality, 4D BIMs could facilitate comparisons of the actual situation at the construction site with the building's planned appearance at the given moment. Such mobile augmented information available at the construction site would have various applications for construction work planning, verification, training and safety, as well as for communication and marketing prior to construction work.

The related camera tracking technologies open up further application scenarios, enabling us to implement mobile location based visual feedback from the construction site to the CAD and BIM systems. We may think of adding elements of reality e.g. images, reports and other comments to the virtual 4D model, with full awareness of the user's location in time and space. Altogether, the tracking and interaction techniques can thus be made to serve the complete spectrum of Mixed Reality [5], forming a seamless interaction cycle between the real world (augmented with virtual 3D/4D model data) and digital building information (augmented with real world data): see Figure 1.

The AR4BC project (Augmented Reality for Building and Construction) between VTT and industrial partners from the Finnish B&C sector aims to providing the mobile user at the construction site with direct two-way access to 3D CAD and 4D BIM information. Our article describes the overall AR4BC software architecture, from content creation (4DStudio) and positioning tools (MapStudio) through onsite visualisation (OnSitePlayer) and wireless data sharing (OnSiteServer) up to mobile interaction and visualisation (OnSiteClient). Special emphasis is on authoring tools, i.e. managing different model formats, linking them to 4D information, placing the models in geo coordinates, as well as managing complex data intensive building model information on thin mobile clients. We also discuss model based tracking and other methods to match the camera view with the virtual model description of the building and its environment, visualisation and interaction tools, as well as the feedback mechanisms to be enabled.

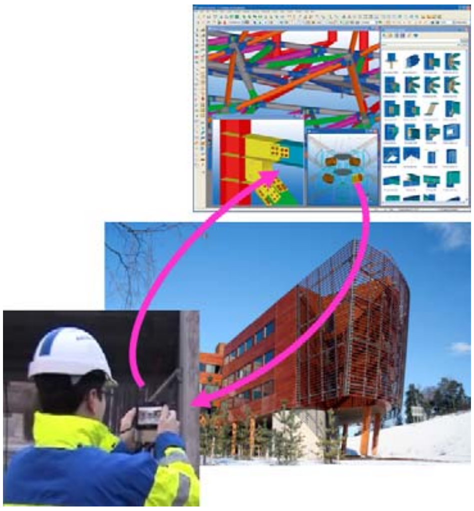


Figure 1. Mixed reality interaction cycle: mobile view of construction site augmented with BIM (arrow down), and feedback from mobile device back to BIM (arrow up).

2 SYSTEM OVERVIEW

2.1 Hardware

The onsite visualisation systems (OnSitePlayer and OnSiteClient extension) are developed with a lightweight UMPC (Ultra Mobile PC) such as the Sony Vaio UX in mind. However, stand-alone operation of the full functionality (OnSitePlayer) with complex BIMs typically requires a high-end laptop including good 3D display hardware. The GPS positioning information is received from the external GPS module via Bluetooth connection. Basically any GPS module with Bluetooth connection supporting virtual serial port communication can be used. For sensor based tracking we are currently using OS5000-US Kit 3 axis Digital Compass from OceanServer Technology Inc.

2.2 Software Modules

The prototype system is divided in three parts; 4DStudio, MapStudio and OnSitePlayer (see Figure 2). The Studio applications are in an authoring role of the system while the Player provides a rich augmented reality view and mobile feedback interface at the construction site. OnSitePlayer can be operated either as a stand-alone, if there is enough processing power and memory on the mobile device, or as a client-server solution distributing the heavy computation to the server, and interaction and display to the client.

Vuorimiehentie 3, Espoo, Finland
charles.woodward@vtt.fi

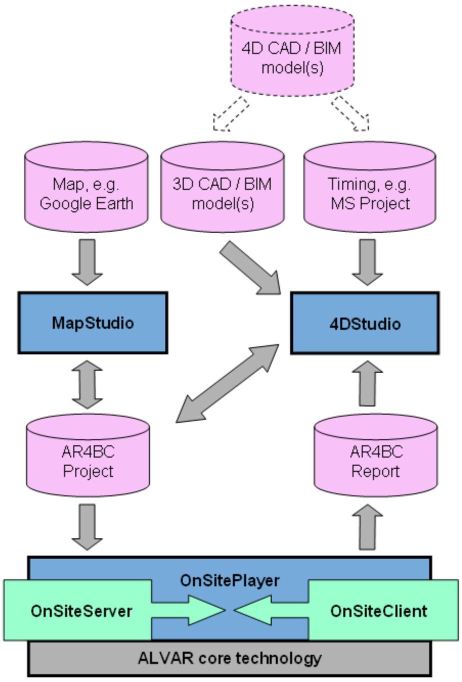


Figure 2. System overview

For rendering we use OpenSceneGraph 2.8.0 and the GUI is built using the wxWidgets 2.8.9.. The applications can handle all OSG supported file formats via OSG's plug-in interface (e.g., OSG's internal format, 3DS, VRML) as well as IFC, using the parser module developed by VTT.

The 4DStudio software is responsible for handling and modifying the 4th dimension i.e. the timing information of the BIM. It allows the user to define the construction timing related values part by part. The visualisation of the workflow in certain time range is another central feature.

The 4DStudio software also provides a user interface to browse and visualize incoming reports from the construction site, created by the user with the OnSitePlayer. The feedback reports are used to document various notes, verifications and problems detected on site. Typically, they contain text descriptions of the issue, perhaps photographs, time stamps, and position information. All these pieces of information are assembled into a single XML file, which is kept separate from the original model files, but can be viewed together with the model using 4Dstudio.

The MapStudio software is used to map the BIMs in geo coordinates (GPS+orientation). Now we don't insert the building model into Google Earth like we did in our previous implementation [6], instead we capture the Google Earth geo information into our application. Other map data bases could be used as well. The MapStudio can also be used to add some additional models around the construction site, so called block models. The block models can be used to mask the main construction model, or to add visual information of the surroundings.

OnSitePlayer is the main augmented reality visualisation, interaction and feedback software on the construction site. OnSitePlayer is able to visualize the models in right position (location, orientation and perspective) by utilizing the model's GPS coordinates in combination with the user's position. User positioning can be done automatically using GPS, or manually defined on the site.

Virtual models of the construction sites can be very complex and consequently mobile devices used for onsite visualisation may not be capable of smooth real time augmentation. To overcome this problem we employ the client-server architecture. The OnSiteClient software is used on the construction site for tracking and for visualizing a 2D image of the model, i.e. the 3D model projected to the client's viewing coordinates. The viewing projection is provided by the server software OnSiteServer and it needs to be updated only once in a while (not real time). Both the OnSiteClient and OnSiteServer modules are obtained as extensions of the stand-alone OnSitePlayer software with relatively small modifications.

Finally, shown at bottom of Figure 2, the tracking algorithms are implemented in our augmented reality subroutine library ALVAR (A Library for Virtual and Augmented Reality) [8]. It provides generic solutions for marker and markerless vision based tracking, as well as for hybrid solutions using electronic sensors.

3 Conclusions

At this point the system is on prototype level and we have focused mainly on authoring and interaction aspects. However, the current system already exceeds by far the functionality of our earlier implementation [6]: being able to deal with scheduled 4D BIMs; support of general OSG compatible file formats as well as IFC; new interactive tools for object placement in geo coordinates; integration of compass for orientation tracking; completely re-designed user interface with selectable 2D and 3D views; and mobile feedback to the office systems.

Further implementation details are provided in [7], including: integration into existing tools and data; communication between the software modules; linking the time schedules to the BIMs, interaction and visualization of with 4D BIMs; feedback mechanisms (reports); model placement in geo coordinates; mobile user interface and interaction; notes on tracking methods (combining model-based, feature based and sensor-based); client-server implementation; optimizations for mobility; and notes for rendering of augmented building models.

References

[1] G. Klinker, D. Stricker, D. Reiners. Augmented reality for exterior construction applications, in *Augmented Reality and Wearable Computers*, Barfield W. and Claudell T. (eds.), Lawrence Elrbaum Press 2001.

[2] C. Woodward, J. Lahti, J. Rönkkö, P. Honkamaa, M. Hakkarainen, J. Jäppinen., K. Rainio, S. Siltanen, J. Hyväkkä. Virtual and augmented reality in the Digitalo building project, *International Journal of Design Sciences and Technology*, Vol 14, No 1, Jan 2007, pp. 23-40.

[3] J. L. Izkara, J. Perez, X. Basogain, D. Borro. Mobile augmented reality, an advanced tool for the construction sector, *Proc. 24th CIB W78 Conference*, Maribor, Slovakia, June 2007, pp. 453-460.

[4] B. Koo, and M. Fischer. Feasibility study of 4D CAD in commercial construction, *Journal of Construction Engineering and Management*, Vol. 126, No. 4, July/August 2000, pp. 251-260.

[5] P. Milgram and F. Kishino. A Taxonomy of Mixed Reality Visual Displays, *IEICE Transactions on Information Systems*, Vol E77-D, No 12, December 1994.

[6] P. Honkamaa, S. Siltanen, J. Jäppinen, C. Woodward, O. Korkalo. Interactive outdoor mobile augmentation using markerless tracking and GPS, *Proc. Virtual Reality International Conference (VRIC)*, Laval, France, April 2007, pp. 285-288.

[7] M. Hakkarainen, C. Woodward, K. Rainio. Software Architecture for Mobile Mixed Reality and 4D BIM Interaction, *Proc. 25th CIB W78 Conference*, Istanbul, Turkey, Oct 2009, 8 pp.

[8] ALVAR home page www.vtt.fi/multimedia/alvar.html

SwissPeaks - Mobile augmented reality to identify mountains

Stephan Karpischek*
Information Management, ETH Zurich, Switzerland

Claudio Marforio†
ETH Zurich, Switzerland

Mike Godenzi‡
ETH Zurich, Switzerland

Stephan Heuel§
Ernst Basler + Partner, Zollikon, Switzerland

Florian Michahelles¶
Information Management, ETH Zurich, Switzerland

Abstract

The paper describes the implementation of SwissPeaks, a mobile augmented reality application for identifying mountains on Apple's iPhone 3GS. The presented prototype includes a novel approach for correcting inaccurate sensor data with manual user input and uses a web service to display only mountains which are actually visible from the user's point of view.

1 Introduction

It does feel great to get a panoramic 360-degree view of the horizon encompassing snow-topped mountains splattered with the myriad hues of a setting sun. Experiencing such a great moment in nature it does not take long until one asks "what's the name of this peak over there?" or "In which direction from here is Matterhorn?".

Today this question can be answered with mobile augmented reality using a mobile phone with camera, location, compass and accelerometer data. The combination of these features allows to superimpose information over objects displayed on the screen captured by the phone's the camera.

Several systems for mobile augmented reality exist on various platforms [1, 5, 2, 6, 7]. Just recently, first applications for identifying mountains have appeared on Apple's iPhone [3, 4]. However, existing mobile augmented reality applications still face two major challenges: First, how to find out which objects are actually visible from the user's point of view? And second the imprecision of the used sensor data. Especially digital compasses in mobile phones are prone to interference.

This paper describes our implementation of a light-weight augmented reality application for the iPhone 3GS, which provides users with information about mountains in sight, and our approach to the above mentioned challenges. The application tries to solve the visibility problem by accessing a geo-information service with pre-calculated view-sheds to display only the peaks actually visible from the observer's viewpoint. Our approach to correcting sensor errors is to use additional, manual user input. In the following we describe our implementation and first experiences with it.

2 Design Challenges

Accuracy of digital compass. First tests showed that accuracy of the phone's built-in digital compass is limited and prone to interference, e.g. when using it near metal surfaces or electric currents.

Accuracy of location information. Location information varies based on the line-of-sight to available GPS satellites. Additionally, the height information is even less accurate, typically by a factor of two to three compared to the horizontal location.

Field of view. Based on location and heading the actual view captured from the phones camera has to be calculated. Then the visible horizontal and vertical viewing angle can be mapped to the actual screen pixels of the mobile phone in order to superimpose the information over the image accordingly.

Selection of mountains to display. Only those visible mountains not occluded by higher mountains in the foreground should be selected for display. Furthermore, in order not to overload the screen and the CPU, the selection of labeled mountains has to be limited in a meaningful way. Tests showed that 60 peaks are reasonable.

3 Description of the Prototype

Our implementation uses the iPhone SDK 3.1 which provides an API for overlaying camera's video live-streams and access to the digital compass. The compass provides horizontal orientation of the phone, vertical orientation is derived from the accelerometer.

Sometimes an icon's location on the screen and the mountain's location in the camera image may not be correctly overlaid due to inaccurate GPS/compass data. The demo makes this error information visible with a semi-transparent rectangle around the information icons. Depending on the error range the color of the area goes from red to orange to green (see Figure 1).

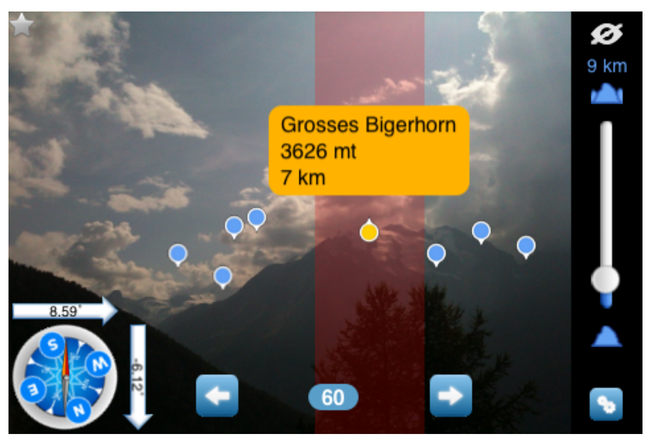


Figure 1: Superimposition of peaks.

The user can also correct this error manually: tapping on the compass in the bottom left corner of the screen enters the calibration mode (see Figure 2), dragging the overlay over the video stream adjusts the horizontal and vertical offset with simple finger movements. Thus, for a given location and environment the sensor error can be manually corrected. The correction values are shown in arrows around the compass.

*e-mail: skarpischek@ethz.ch
†e-mail:maclaudi@student.ethz.ch
‡e-mail:godenzim@student.ethz.ch
§e-mail:stephan.heuel@ebp.ch
¶e-mail: fmichahelles@ethz.ch

Figure 2: Calibration mode.

To find out, whether a peak is visible from a certain viewpoint, so-called view-sheds of 89 selected peaks are calculated: based on the terrain surface those locations from which a peak is visible are labeled (see Figure 3). Currently, a web-service provides access to view-sheds of 89 selected dominant peaks in Switzerland. Queries are returned in less than half a second. The number of peaks with visibility information will be increased in the future.

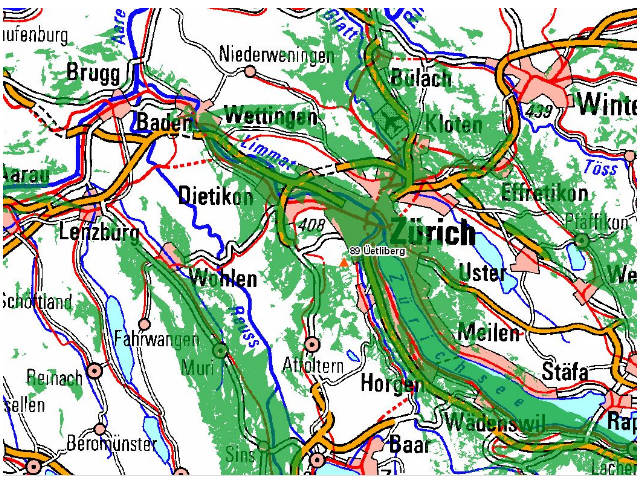


Figure 3: One out of 89 view-sheds. From the green areas Uetliberg is visible.

For all other peaks without calculated view-sheds the mode can be switched by the eye-icon in the upper right corner - a reverse geocoding service derives the current country the user is in in. Then mountains of the country can be downloaded from geonames.org and stored on the phone locally. Based on the phone's current view, the mountains to display are selected for the calculated overlay. The displayed peaks are limited to a number of 60 in order to preserve a satisfying refresh rate and user experience. A slider selects the maximum distance from the current location to the mountain to display (see Figure 1, right edge). If the returned result exceeds the threshold of 60 mountains, the user is asked if the set should be limited by either the closest or highest peaks. The user's decision is saved and can later be changed in the application preferences. Furthermore, favorite peaks can be bookmarked preserving the users preferences.

Each mountain within the field of view is displayed by a blue peak icon. The peak nearest to the center is pre-selected highlighted in yellow. A label displays name, height and distance of the selected peak (see Figure 1). The user can navigate through the selected mountain using the left/right arrows at the bottom of the screen. When the user taps on the label of mountain a browser window opens wikipedia for additional information using the mountain name as a query string. For Switzerland the prototype queries a user-generated tour description repository.

More details about the implementation can be found on our development website[1].

4 Discussion and Conclusion

We showed how information available today in geo-information systems can be made available to non-professional users providing value in a meaningful and intuitive way. The presented prototype works with mobile phone sensor data only without requiring more advanced AR techniques such as image processing and object recognition. It also allows users to correct inaccurate sensor data with manual input.

First informal evaluations with mountaineers testing the applications were positive. We see this as an indication to continue this work, conduct additional tests and user studies, and extend the application with more features in the future.

References

[1] Layar reality browser. http://layar.eu.
[2] Sekai cam. http://www.tonchidot.com/product-info.html.
[3] Peak.AR. http://peakar.salzburgresearch.at, September 2009.
[4] Peaks. http://peaks.augmented-outdoors.com, September 2009.
[5] Wikitude world browser. http://www.wikitude.org, March 2009.
[6] M. Kähäri and D. J. Murphy. MARA - sensor based augmented reality system for mobile imaging. Demo at ISMAR 2006.
[7] D. Wagner and D. Schmalstieg. Making augmented reality practical on mobile phones, part 1. *IEEE Computer Graphics and Applications*, 29(3):12–15, 2009.

[1] http://dev.im.ethz.ch/wiki/Peaks