

Chapter 6

Urban Terrain Modeling For Augmented Reality Applications

Simon Julier, Yohan Baillot, Marco Lanzagorta, Lawrence Rosenblum and Dennis Brown

Baillot, Brown and Julier are with ITT AES/NRL. Rosenblum is with the Naval Research Laboratory, Washington D. C.. Lanzagorta is with Scientific Solutions

Key words: Augmented reality, virtual reality, urban modeling, terrain reconstruction, computer vision

Abstract: Augmented reality (AR) systems have arguably some of the most stringent requirements of any kind of three-dimensional synthetic graphic systems. AR systems register computer graphics (such as annotations, diagrams and models) directly with objects in the real-world. Most of the AR applications require the graphics to be precisely aligned with the environment. For example, if the AR system shows wire frame versions of actual buildings, we cannot afford to see them far apart from the position of the real buildings. To this end, an accurate tracking system and a detailed model of the environment are required. Constructing these models is an extremely challenging task as even a small error in the model (order of tens of centimeters or larger) can lead to significant errors, undermining the effectiveness of an AR system. Also, models of urban structures contain a very large number of different objects (buildings, doors and windows just to name a few). This chapter discusses the problem of developing a detailed synthetic model of an urban environment for a mobile augmented reality system. We review, describe and compare the effectiveness of a number of different modeling paradigms against traditional manual techniques. These techniques include photogrammetry methods (using automatic, semi-automatic and manual segmentation) and 3 dimensional scanning methods (such as aircraft-mounted LIDAR) and conventional manual techniques.

Report Documentation Page				Form Approved OMB No. 0704-0188	
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.					
1. REPORT DATE 2001		2. REPORT TYPE		3. DATES COVERED 00-00-2001 to 00-00-2001	
4. TITLE AND SUBTITLE Urban Terrain Modeling for Augmented Reality Applications				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Research Laboratory, Washington, DC				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited					
13. SUPPLEMENTARY NOTES 3D Synthetic Environments Reconstruction (pages 119-136). Dordrecht, The Netherlands: Kluwer Academic Publishers, 2001.					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT Same as Report (SAR)	18. NUMBER OF PAGES 20	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

1. INTRODUCTION

Augmented Reality (AR) has the potential to literally revolutionize the way in which information is disseminated to mobile users. The basic principle of augmented reality is illustrated in Figure 1—a user wears a see-through head mounted display and his position and orientation is tracked. Using a model of the user’s environment, computer graphics are generated; through the head-mounted display, they appear to be aligned directly with the objects in the user’s environment. Experimental AR prototypes have been demonstrated in task domains ranging from aircraft manufacturing (Caudell, 1992; Caudell, 1994) to image-guided surgery (Fuchs, 1998), and from maintenance and repair (Feiner, 1993; Hoff, 1996) to building construction (Webster, 1996).

Recent developments in wearable computers have begun to make *mobile* augmented reality systems a reality (Feiner, 1997; Piekarski, 1999, Julier, 2000). Systems such as that shown in Figure 1 can now be constructed using commercially available hardware and software. With this freedom comes a new domain—outside of a laboratory and into the “real world”—and many new possible applications.

One of the most potentially most important benefits of AR is for providing situation awareness to military personnel in urban environments. Urban environments are complicated, dynamic, and inherently three-dimensional, and military personnel need to receive data to ensure safe operation and coordination with other team members. AR can provide information such as virtual signposts (name labels that appear to be attached to the side of a building), routes (perhaps as a trail of breadcrumbs which need to be followed), or even various types of infrastructure (such as the location of pipes). This information can be presented in a hands-off manner; it can be integrated directly into the environment, and does not block the user’s view of the “real world.” An actual output from the mobile AR system of Figure 1 is shown in Figure 2. This image shows various types of computer graphics including the outline of buildings and windows.

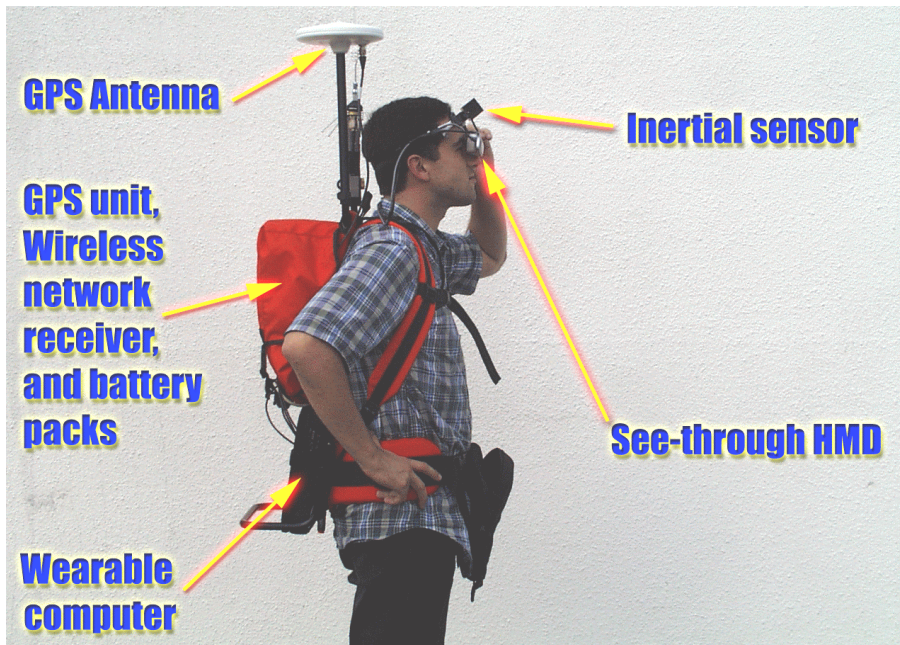


Figure 1. A wearable augmented reality system. The large size of the system is the result of the fact that it is developed from purely using COTS hardware



Figure 2. Actual output captured from the headmounted display of the hardware system shown in Figure 1.

However, an AR system is only effective if the computer graphics it generates are aligned with the objects in the environment. If the graphics are incorrectly aligned, the result can be a system that is annoying or possibly even misleading. There are several factors that contribute to the accuracy of the registration. These include:

- *Accuracy of the tracking system.* How well is the user's position and orientation known?
- *Accuracy of the calibration of the head mounted display.* How well is the mapping from the 2D graphics display to the view of the user's eye known?
- *Accuracy of the underlying models.* How well is the underlying environment known?

The first two issues have been extensively examined and reported upon in the literature. Azuma (Azuma, 1994), for example, studied the effect of tracking errors (including prediction lag) when a user looks at a scene whose properties are extremely well-known. Holloway (Holloway, 1995) developed detailed error models that examined how the unknown optical characteristics of the display affected registration errors. These studies have shown that tracking errors are much more significant than calibration errors and, for most applications, calibration errors (apart from the static offset of how a user puts the display on their head) can be ignored.

However, the third issue—model acquisition—has received relatively little attention in the mobile AR literature. This is despite the fact that the importance of model accuracy is well recognized for AR systems. Indeed, it could be that AR systems apply some of the most stringent requirements of any kind of three-dimensional synthetic graphic systems. The reason is that unlike virtual or visualized display systems, where a user looks at a purely synthetic environment, an AR system locates the graphics directly with the real world. Even though a model might be qualitatively correct, quantitative¹ modeling errors are readily apparent. However, outside the computer vision community, it appears that little research has been done into the third problem of model acquisition for mobile AR. The prevailing assumptions appear to be that either the system is working in an environment where accurate models can be constructed (for example, in a laboratory or an operating theatre) or the modeling errors are secondary to the other types of errors that were listed above.

This chapter discusses the problem of developing a detailed synthetic model of an urban environment for a mobile augmented reality system. We review, describe, and compare the effectiveness of a number of different modeling paradigms against traditional manual techniques. The structure of this chapter is as follows. Section 2 describes the role and function of a mobile AR system in more detail and presents an analysis of the model requirements that provide a lower bound on the required model accuracy. In Section 3 we survey a number of different modeling techniques and assess their advantages and disadvantages in a typical urban scenario. A summary and conclusions are given in Section 4.

¹ By qualitatively we mean that the model, when viewed on its own, *appears* to be correct. For example, the model might contain the correct number of buildings with the correct relative locations with respect to one another.

2. MODELING REQUIREMENTS FOR A MOBILE AUGMENTED REALITY SYSTEM

The requirements of a model depend on the purpose to which that model will be used. In this section we identify a set of requirements that will be used to assess the appropriateness of different modeling approaches.

Our specific application is the Battlefield Augmented Reality System (BARS), a visualization tool which can be used to provide situation awareness to Marines operating in urban environments. BARS is motivated by the fact that with the proliferation of urbanization throughout the world, it is expected that many future military operations (such as peace keeping or hostage rescue) will occur in urban environments (CFMOUT, 1997). These environments present many challenges. First, urban environments are extremely complicated and inherently three-dimensional. Above street level, the infrastructure of buildings may serve many different purposes (such as hospitals or communication stations) and can harbor many types of risks (such as snipers or instability due to structural damage). These features are often distributed and interleaved over several floors of a multi-floor building. Below street level, there may be a complex network of sewers, tunnels and utility systems. Cities can be confusing (especially if street signs are damaged or missing) and coordinating multiple team members can be difficult. To ensure the safety of both civilian and military personnel, it has long been argued that environmental information must be delivered to the individual user *in situ*. Some of the types of environmental information that must be shown include:

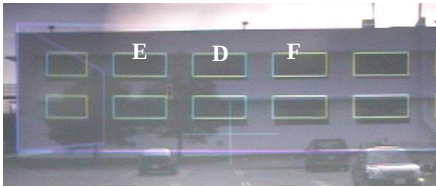
1. **Information local to the user.** Information which is localized and is a function of the user's current position and orientation. This type of information will be overlaid on relatively large-scale features in the environment. Examples include:
 - Building data (e.g., name of building, known function and floor plans).
 - Routing information (e.g., path that has to be followed to reach a particular destination).
 - Signpost information (e.g., translations of road signs).
2. **Highly localized information.** Unlike the local information described in the previous type, this type of information must be accurately registered to specific features in the environment.
 - Warnings (e.g., the alert that a particular window in a particular building contains a sniper).
 - Infrastructure and utility information, such as the location of power lines, service tunnels and water supplies, including 3D representations

of otherwise hidden features that can be viewed as if seen with “X-ray vision”.

This problem statement introduces two sets of requirements: what components should be in the model, and how accurately must these components be known?

The components of the model are defined by the need to be able to access and display individual “fine-grained” features such as windows and doors. Therefore, the model cannot simply be a “polygon soup” which consists of 3D representations of buildings that are covered with textures. Rather, the model must be composed of many hundreds (or thousands) of individually identified features possibly with their own textures.

The acceptable level of accuracy is highly context and domain dependent. We assess the minimum accuracy requirements by considering the motivating scenario shown in Figure 3.



The AR system for the person (A) needs to be able to register graphics with the center of a window (D) on a wall of a building (B). The target window is surrounded by two other similar windows (centers at E and F respectively). The spacing between each window is uniform and is of length m . The user looks along the Y-axis. In general, the user does not look directly at the side of the building. Rather, the angle subtended between the user's viewing direction and the side of the building is α . The *augmentation error* is the difference between where an object appears on the head mounted display and where the computer rendered augmentation for that object appears. Because the optical characteristics of the head mounted display are assumed to be known, this error is equivalent to the angular error between the ray that

points to the object and the ray that is formed by projecting the location of the graphics (drawn on the head-mounted display) out into the world.

Since the purpose of the system is to unambiguously show the user the correct window, we limit the augmentation error so that the computer generated augmentation of D lies less than half way between D and the adjacent features (E or F)². Therefore, the computer-generated graphics should lie within the sector with interior angle θ .

The main factors influencing whether or not the augmentation is displayed correctly are the modeling and tracking errors. Modeling and tracking errors modify the size of θ as a function of the position and orientation of the building with respect to the user. Model errors are considered to be errors in position only because the models are generally constructed from the measurement of the location of their corners. Therefore, for a given modeling error the augmentation error will decrease as the distance between the user and the building increases. Tracking errors affect both position and orientation. Position errors can be treated to be the equivalent of modeling errors. Orientation errors lead to augmentation errors that are constant irrespective of the distance between the user and the building. For this reason, estimating the orientation of a moving user is one of the most difficult challenges in mobile augmented reality (Azuma, 94).

Modeling errors have their greatest impact on the augmentation error when the building is orthogonal to the user's point of view. Consider the case when the user looks directly at the face of the building ($\alpha=90^\circ$). In this case, the horizontal error between the actual position of the target feature (D) and the neighboring feature (E) is

$$m = e_{\text{position}} + e_{\text{modeling}} + 2 * Y_f * \tan (e_{\text{orientation}} / 2)$$

The effects of this function are illustrated in Figure 4, which plots the maximum permissible modeling error for different viewing distances and tracker orientation errors. It is assumed that the windows are 2m apart and the errors in position are 0.1m. As an example, if the user looks at the building at a distance of 35m with an angular error of 2 degrees, the maximum modeling error should be less than 0.5m.

To illustrate this function with a concrete example, we can consider the case of location of window on a building outdoors. We consider the center of the windows to be separated of 2 meters (m). A realistic position error using a kinematics-differential GPS is 0.1 m. An orientation error using state-of-the-art inertial platform (gyroscope, accelerometers and compass) is within 1

² The horizontal spacing between windows on the same floor is usually much less than the vertical spacing between windows on adjacent floors. Therefore, our analysis only considers the first case.

degree. Figure 4 is a chart representing, for this specific case, what should be the maximum modeling error as a function of the viewing distance so that a specific window can be highlighted without confusing it with a neighboring one. As an example, if the user is looking at a building at a distance of 35m with an angular error of 2° , the maximum permissible error in the model is less than 0.5m.

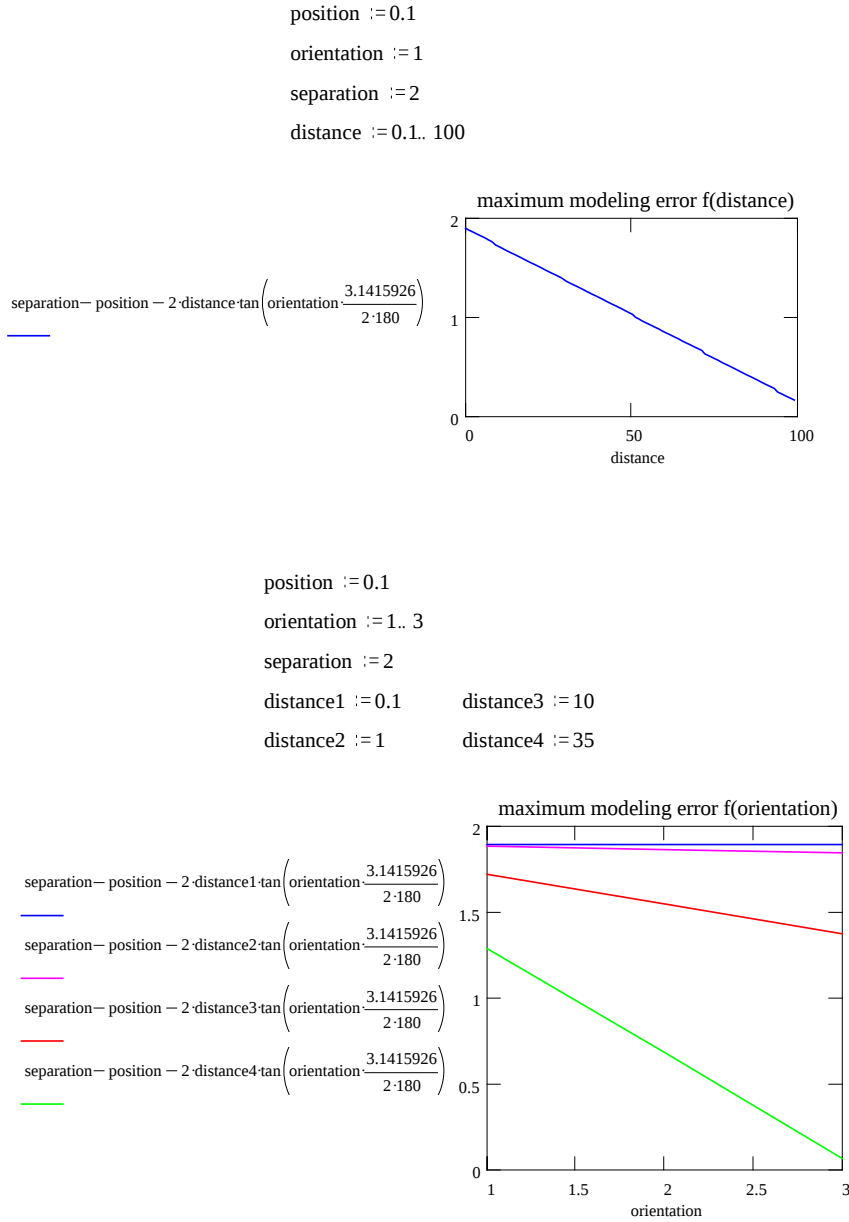


Figure 4. The results of the analysis

In summary, this section has shown that our mobile AR application requires the following:

- The model must be composed of building as well as “fine-grained” building features. These features include windows and doors. Each

object must be identified individually – it is not sufficient to build a model that is a “polygon soup” of building shapes and textures.

- The maximum permissible error in estimating any feature must be less than 0.5m.

We now consider a number of different modeling approaches that are available.

3. MODELING METHODS

1. Surveying Methods

Probably the oldest (and simplest) approach to constructing a model is to use conventional surveying techniques. It includes equipment such as tape measures, theodolites, laser range finders, and kinematic GPS receivers. This type of approach is relevant because it can be used as the “ground truth” against which other methods can be compared. State-of-the-art surveying tools can give errors, when surveying a large site, on the order of centimeters.

However, manual methods have two obvious drawbacks. First, they do not scale well. Because the model must be constructed using many measurements, data acquisition and model building can take on the order of days. Second, certain types of building features (such as windows on a high story) are difficult to survey using these methods.

2. Topological LIDAR

A common type of system uses LIght Detection And Ranging (LIDAR). This scanning method use the same principle as RADAR, and it can be thought of as a laser radar. The LIDAR instrument transmits light out to a target. The transmitted light interacts with and is changed by the target. Some of this light is reflected or scattered back to the instrument where it is analyzed. The change in the properties of the light enables some property of the target to be determined. The time for the light to travel out to the target and back to the LIDAR is used to determine the range of the target. LIDAR operates in the ultraviolet, visible, and infrared region of the electromagnetic spectrum. One of the most important practical advantages is that topographical LIDAR methods utilize an airborne ranging sensor to measure highly accurate distances to objects and surfaces. Distances from the airborne sensor are calculated through thousands of laser pulses within a

scanned width beneath the aircraft. As a result, it is possible to acquire models of large environments extremely rapidly. Several commercial services, such as 3Di's EagleScan, provide commercial data sets of urban environments for municipal and government customers.

The use of LIDAR methods for topographical reconstruction can be traced back to NASA's application of LIDAR technology for oceanographic applications back in the 1970s. Although the US Geological Survey and the Jet Propulsion Laboratory experimented with these technologies during the 1980s, no successful low cost, high resolution results were obtained until the 1990s. Common LIDAR resolution ranges between 1 and 3 meters (X and Y) with a 1 meter horizontal accuracy, and delivering elevation accuracy (Z) of 30 centimeters or better. The ground coverage or 'swath' of the LIDAR sensor is a direct function of the altitude of the aircraft together with the scan angle (about 18 degrees to each side) of the laser itself. A general rule of thumb result is that the ground swath width to be one-half of the altitude height above ground level. So, multiple flight lines are required to cover wide areas.

LIDAR offer several advantages for topographical applications. First of all, it allows for the rapid generation of large scale Digital Terrain Models (DTM). Second, it is daylight and relatively weather independent. Third, it is extremely fast and precise in comparison to other topographic methods—historically, elevation data acquisition for the production of digital terrain data and DTMs was very costly and time consuming, and was usually done by acquiring and analyzing many stereo pairs of aerial photographs. Finally, LIDAR data can be fused directly with images to provide 3D textured models of an environment.

However, LIDAR methods are not sufficient, on their own, to fill the needs of Augmented Reality Systems. There are several difficulties with their use. First, the typical spatial errors recorded by a LIDAR model are not sufficient to meet the needs of mobile AR identified in the first section. Second, LIDAR does not, in itself, identify fine-grained building features. Rather, the best one can do is to use the LIDAR data and combine it with other data (such as images). However, as explained later, there can be significant difficulties unless the image data is extremely high quality. Finally, it is not clear that such approaches are capable of picking up crucial features such as the geometry in narrow alleyways. Together, these difficulties imply that LIDAR is not sufficient to meet the needs of mobile AR systems.

3. Photogrammetric and Computer Vision-Based Techniques

A popular alternative to explicit range-based modeling algorithms are those that attempt to extract model parameters directly from photographs and video images. Given a sufficient number of pictures of an environment and sufficient camera calibration information (such as focal length and radial distortion), it is possible to construct a model of the scene at which a camera has been pointing (Maybank-92). Almost all such systems are designed to extract the geometry of buildings and to texture these to provide models that can be used for flythrough and other applications.

UMass's ASCENDER system (Jaynes-96), for example, provides a suite of software that allows the construction of textured models of an environment from aerial photographs. A calibrated camera is mounted to the bottom of an aircraft and a series of images are taken. Using template matching, the system uses an edge detector to determine the footprints of buildings, which are registered between multiple images. From this information, the geometric structure of the buildings can be determined. Textures are extracted in several steps. From those faces that are clearly visible, the texture is warped to offset the fact that it was taken from a non-oblique angle. For those faces that are obscured, the system has the capability to "fill in" and correct for the textures. Given information about the location of the sun, the system calculates the shadows cast from one building onto the surface of another one so that the color histogram of the shadowed region can be made to match that of the unshadowed region. Occluded textures can be extrapolated from visible building features.

Although these systems provide displays sufficient to meet the needs of many applications including cartography, land-use surveying, and urban planning, these models do not appear to be appropriate for our application. Many of these problems stem from the same limitations as airborne LIDAR sensors: the errors in the models can be fairly large and difficulties such as occlusion and the angle at which walls are viewed (near vertical) makes it difficult to recover the types of features which we need to include in the model.

Many of these difficulties can be overcome by using imagery that is collected directly from within the urban environment itself, for example, by a user walking through the environment. A number of software systems and packages, already marketed for computer graphics, are available for this purpose. One such system is Canoma, a commercial system that was inspired by the FAÇADE system (Debevec-96). Canoma uses a human operator to help identify correspondences between several pictures. The system is given a set of photographs that have been taken of the object to be modeled. The

user identifies the same features (such as edges of buildings) between different pictures. The system then attempts to find a model that is consistent with the images that have been taken. However, we have encountered two difficulties with Canoma, both of which are illustrated in Figure 5, which shows a model constructed using the Canoma software. The first difficulty is that the software does not attempt to predict the accuracy of the model that it is constructing. As a result, it is only possible to assess the errors in the model by directly measuring them against ground truth. The second difficulty is that the texture is significantly distorted. Using this system, it is not possible to construct a model of the environment and subsequently use the texture data in any meaningful way.

A more sophisticated system for model reconstruction is PhotoModeler, developed by EOS Systems. PhotoModeler adopts broadly the same user interface principles as Canoma. The system uses a set of photographs taken from a calibrated (or approximately calibrated) camera. A user identifies the same features in multiple photographs and a model is constructed. Unlike Canoma, which only uses geometric primitives, PhotoModeler can be used to register point or line features. In Figure 6 we show a set of input images to PhotoModeler. These consist of the outline of the building as well as certain critical features such as windows or doors. The generated model is shown in Figure 7.

Wasilewski (Wasilewski-96) has developed a toolkit for urban terrain construction that combines elements of both aerial photogrammetry with the precise reconstruction from the PhotoModeler system. The model is constructed in several stages. First, aerial images are used to identify the footprints of buildings. Height is also entered (if it is already known) or is estimated from the shadows cast by the buildings. Finer-scale structures are reconstructed using PhotoModeler. A reconstruction of Atlanta using their system is shown in Figure 8.

However, the greatest difficulty with manual systems such as those described here is the problem of scale. Because a manual operator must analyze each photograph and identify the correspondences between successive images, constructing a model can be an extremely difficult process. Therefore, a number of authors are attempting to develop systems that minimize the role that must be played by a user. These systems usually attempt to estimate structure (what a camera looks at) and motion (how the camera moves through the scene). Unlike the manual approaches described above, these systems attempt to track image primitives (or tokens) between multiple frames (Beardsley-95, Ayache-87, Zhang-92, Faugeras-98). Furthermore, these systems attempt to estimate the parameters of the camera directly as well, obviating the need for a calibrated camera. Although progress in this research seems extremely encouraging, most systems and

results only consider the problem of developing a relatively small number of models (e.g., for a single building).

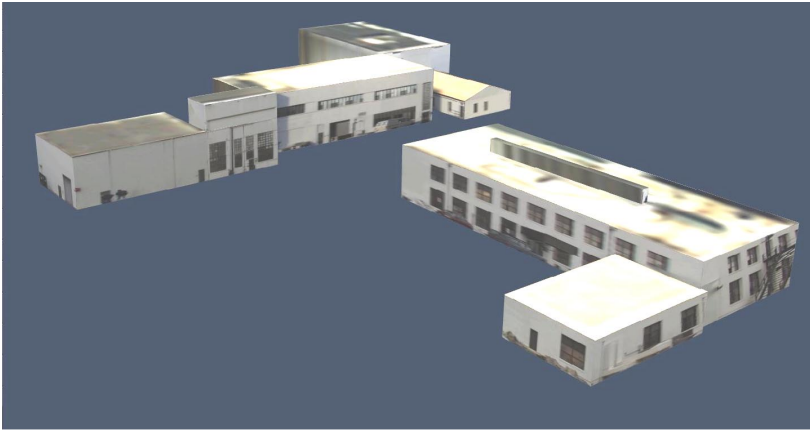


Figure 5. Model constructed using MetaCreation's Canoma software package. Note that although the broad geometric relationship between the buildings is correct, the textured building features (important for a mobile AR application) show significant distortion



Figure 6. Input images required to build the model shown below. In this (and similar manual systems) the user takes a series of photographs using a calibrated camera. The user then manually identifies common features between groups of photographs. In this case, the user identifies edges of the building as well as significant features (windows and a partially open door on the top floor of the main building).

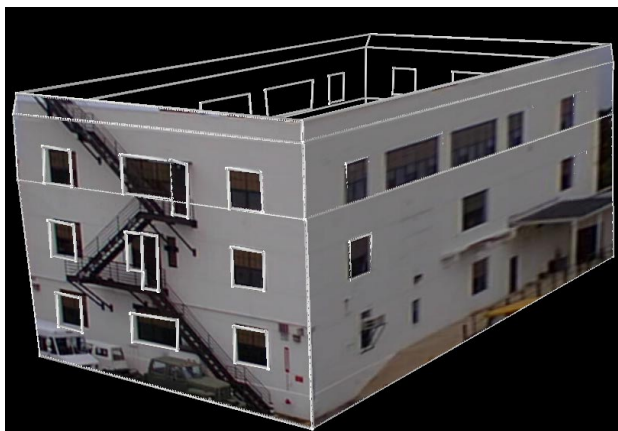


Figure 7. Model of test building constructed using EOS System's PhotoModeler system. The user has to manually register the location of the individual features. The software assesses its accuracy using dimensionless units.

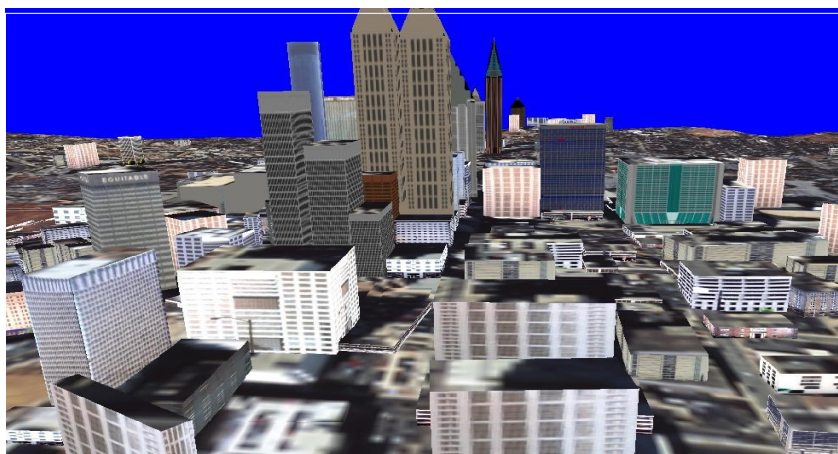


Figure 8. Reconstruction of Atlanta performed by Ribarsky and Faust at Georgia Tech. A combination of aerial photographs and more refined photometrics leads to more accurate building models. However, note that small-scale building features are provided by photographs. Many buildings, in fact, possess "default" textures which do not necessarily reflect the actual physical appearance of the building.

Recently, MIT has embarked on the MIT City Scanning Project (Teller-98). The purpose of this project is to make a fully automated system for building an end-to-end system that “scans” an urban environment and constructs a 3D model that is suitable for use within a CAD package. A mobile robot is driven along a prearranged path. Every 10-15 meters the vehicle stops and, using a high resolution camera which is mounted on a pan or tilt head, the system records a mosaic of 47 or 71 images. These images are combined to form high resolution panoramic images at each location. The collection of images, known as a pose image dataset, is processed using a collection of algorithms to identify buildings and building structures. Although the scope and scalability of this algorithm is ideal for our application, there do not appear to be any detailed results published yet as to the actual accuracy achieved with the system. Columbia University is also developing a mobile robot that incorporates range and vision data in its urban model reconstruction efforts (Reed-99, Gueorguiev-00). Although this system does not appear to be as mature as the MIT system, it has the potential to automatically construct accurate urban models of sufficient accuracy, detail and resolution that they can be used with a mobile augmented reality system.

4. CONCLUSIONS

In this chapter we have considered the problem of constructing the model of an urban environment for mobile augmented reality applications. Unlike fly through, walk through, and other types of virtual reality applications, augmented reality applies two strict conditions. First, the models must be extremely accurate. A preliminary analysis suggests that errors cannot exceed 0.5m. Second, because the system must highlight individual building features, it is not sufficient to extract the geometry of the buildings and simply apply a texture to them.

We have considered a number of systems, both commercially available and currently under academic research, which aspire to construct urban models. However, although many of these systems yield models that are qualitatively correct, most do not meet our conditions identified earlier. Either the systems are not able to yield models of sufficient accuracy (for example, errors in LIDAR measurements are twice our acceptable levels) or the systems are not capable of identifying individual features.

Of the methods we have surveyed, we believe that two types of systems are likely to be most applicable. The first are the largely manual methods and, in particular, precision photogrammetric systems such as PhotoModeler. These systems are established products and have been available for many

years. However, the problem with these systems is that they can be highly labor intensive and, as a result, constructing a model of a large-scale urban environment can be an extremely difficult prospect. Second, the fully autonomous systems currently under development appear extremely promising both in terms of the potential accuracy and detail of the models that they construct.

- Ayache, N., and Lustman, F. (1987). Fast and Reliable Passive Trinocular Stereovision, *Proceedings of the International Conference on Computer Vision*.
- Azuma, R., and Bishop, G. (1994). Improving Static and Dynamic Registration in an Optical See-Through HMD, *Proceedings of SIGGRAPH 94*, Orlando, Florida.
- Beardsley, P., Torr, P., and Zisserman, A. (1995). 3D Model Acquisition from Extended Image Sequences, Oxford University Technical Report No. OUEL 2089/96.
- Brooks, F. (1999). What's Real About Virtual Reality?, *Computer Graphics and Applications*, November/December 1999, pp. 16–27.
- Caudell, T.P., and Mizell, D.W. (1992). Augmented Reality: An Application of Heads-Up Display Technology to Manual Manufacturing Processes, *Proceedings of 1992 IEEE Hawaii International Conference on Systems Sciences*, January.
- Caudell, T.P. (1994). Introduction to Augmented Reality, *Proceedings of the Conference on Telem manipulator and Telepresence Technologies*, SPIE, Vol. 2351.
- Concept Division, Marine Corps Combat Development Command "CMOUT". (1997). A Concept for Future Military Operations on Urbanized Terrain.
- Debevec, P.E., Camillo J. T., and Malik, J. (1996). Modeling and Rendering Architecture from Photographs. In *SIGGRAPH '96*.
- Faugeras, O., Rober, L., Laveau, S., Csurka, G., Zeller, C., Gauclin, C., and Zoghalmi, I. (1997). 3-D Reconstruction of Urban Scenes from Image Sequences, *CVGIP-IU*.
- Feiner, S., MacIntyre, B., Seligmann, D. (1993). Knowledge-based Augmented Reality, *Communications of the ACM*, Vol.37, No.6, pp. 52-62
- Feiner, S., MacIntyre, B., Höllerer, T., and Webster, T. (1997). A Touring Machine: Prototyping 3D mobile augmented reality systems for exploring the urban environment, *Proceedings of the International Symposium on Wearable Computers*, Cambridge MA.
- Fuchs, H., Livingston, M., Raskar, R., Colucci, D., Keller, K., State, A., Crawford, J., Rademacher, P., Drake, S., and Meyer, A. (1998). Augmented Reality Visualization for Laparoscopic Surgery, *Proceedings of the First International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI '98)*.
- Gueorguiev, A., Allen, P., Gold, E., and Blaer, P. (2000). Design, Architecture, and Control of a Mobile Site Modeling Robot, *IEEE Int. Conf. on Robotics and Automation*, San Francisco.
- Hoff, W.A., Lyon, T., and Nguyen, K. (1996). Computer Vision-Based Registration Techniques for Augmented Reality, *The Proceedings of Intelligent Robots and Control Systems XV, Intelligent Control Systems and Advanced Manufacturing*, SPIE, Vol. 2904, pp. 538-548.
- Holloway, R. (1995). Registration Errors in Augmented Reality Systems, PhD dissertation, University of North Carolina at Chapel Hill. TR95-016.
- Jaynes, C., Collins, R., Cheng, Y.Q., Wang, X.Q., Stolle, F., Schultz, F., Hanson, A., and Riseman, E. (1996). Automatic Construction of Three-Dimensional Models of Buildings, in *ARPA IU RADIUS*, Edited by: Yiannis Aloimonos.

- Julier, S., Baillot, Y., Lanzagorta, M., Brown, D., and Rosenblum, L. (2000). BARS: Battlefield Augmented Reality System, Presented at the 2000 International Information Systems Technology Panel (IST) Symposium New Information Processing Techniques for Military Systems.
- Maybank, S. and Faugeras, O. (1992). A Theory of Self-Calibration of a Moving Camera, *International Journal of Computer Vision*, 8(2):123-151.
- Piekarski, W., Gunther, B., and Thomas, B. (1999). Integrating Virtual and Augmented Realities in an Outdoor Application, *Proceedings of the Second International Workshop on Augmented Reality*, pp. 45 – 54, October 20-21.
- Reed, M., and Allen, P. (1999). 3-D Modeling from Range Imagery: An Incremental Method with a Planning Component, *Image and Vision Computing*, V. 17, pp. 99-111, 1999.
- Teller, S. (1998). Automated Urban Model Acquisition: Project Rationale and Status, *Proceedings of the 1998 Image Understanding Workshop*, pp. 455-462.
- Wasilewski, T., Faust, N. and Ribarsky, W. (1996). Semi-Automated and Interactive Construction of 3D Urban Terrain Models.
- Webster, A., Feiner, S., MacIntyre, B., Massie, W., and Krueger, T. (1996). Augmented reality applications in architectural construction”, *Designing Digital Space: An Architect's Guide to Virtual Reality*, Ed. D. Bertol, Pub. John Wiley & Sons, pp. 193–200.
- [Zhang-92] Zhang, Z., and Faugeras, O. (1992). 3D Dynamic Scene Analysis, Springer-Verlag.