

View Planning for Site Modeling

Peter K. Allen Michael K. Reed Ioannis Stamos
Department of Computer Science, Columbia University
{allen,m-reed,istamos}@cs.columbia.edu *

Abstract

3-D models of complex environments, known as site models, are used in many different applications ranging from city planning, urban design, fire and police planning, military applications, virtual reality modeling and others. Site models are typically created by hand in a painstaking and error prone process. This paper focuses on two important problems in site modeling. The first is how to create a geometric and topologically correct 3-D solid from noisy data. The second problem is how to plan the next view to alleviate occlusions, reduce data set sizes, and provide full coverage of the scene. To acquire accurate CAD models of the scene we are using an incremental volumetric method based on set intersection that can recover multiple objects in a scene and merge models from different views of the scene. These models can serve as input to a planner that can reduce the number of views needed to fully acquire a scene. The planner can incorporate different constraints including visibility, field-of-view and sensor placement constraints to find correct view points that will reduce the model's uncertainty. Results are presented for acquiring a geometric model of a simulated city scene and planning viewpoints for targets in a cluttered urban scene.

1 Introduction

Realistic 3-D computer models are fast becoming a staple of our everyday life. These models are found on TV, in the movies, video games, architectural and design programs and a host of other areas. One

of the more challenging applications is in building geometrically accurate and photometrically correct 3-D models of complex outdoor urban environments. These environments are typified by large structures (i.e. buildings) that encompass a wide range of geometric shapes and a very large scope of photometric properties. 3-D models of such environments, known as site models, are used in many different applications ranging from city planning, urban design, fire and police planning, military applications, virtual reality modeling and others. This modeling is done primarily by hand, and owing to the complexity of these environments, is extremely painstaking. Researchers wanting to use these models have to either build their own limited, inaccurate models, or rely on expensive commercial databases that are themselves inaccurate and lacking in full feature functionality that high resolution modeling demands. For example, many of the urban models currently available are a mix of graphics and CAD primitives that visually may look correct, but upon further inspection are found to be geometrically and topologically lacking. Buildings may have unsupported structures, holes, dangling edges and faces, and other common problems associated with graphics vs. topologically correct CAD modeling. Further, photometric properties of the buildings are either missing entirely or are overlaid from a few aerial views that fail to see many surfaces and hence cannot add the appropriate texture and visual properties of the environment. Our goal is to have a mobile system that will autonomously move around a site and create an accurate and complete model of that environment with limited human interaction.

There are a number of fundamental scientific issues involved in automated site modeling. The first is

*This work was supported in part by an ONR/DARPA MURI award ONR N00014-95-1-0601, DARPA AASERT awards DAAH04-93-G-0245 and DAAH04-95-1-0492, and NSF grants CDA-96-25374 and IRI-93-11877.

how to create a geometric and topologically correct 3-D solid from noisy data. A key problem here is merging multiple views of the same scene from different viewpoints to create a consistent model. In addition, the models should be in a format that is CAD compatible for further upstream processing and interfacing to higher level applications. A second fundamental problem is how to plan the next view to alleviate occlusions and provide full coverage of the scene. Given the large data set sizes, reducing the number of views while providing full coverage of the scene is a major goal. If a mobile agent is used to acquire the views, then planning and navigation algorithms are needed to properly position the mobile agent. Third, the models need to integrate photometric properties of the scene with the underlying geometry of the model to produce a realistic effect. This requires developing methods that can fuse and integrate range and image data. Fourth, methods that reduce the complexity of the models while retaining fidelity are needed. This paper focuses on solving the first two problems, model acquisition and view planning.

Previous work in the model acquisition phase focuses on construction of models of 3-D objects from range data, typically small objects for reverse engineering or virtual reality applications. Examples of these efforts include the groups at Stanford [17, 4], CMU [18], UPENN [7], and Utah [16]. However, these methods have not been used on larger objects with multiple parts. Research specifically addressing the modeling of large outdoor environments includes the FACADE system developed at Berkeley [5]. This is an example of a system that merges geometric 3-D modeling with photometric properties of the scene to create realistic models of outdoor, urban environments. The system however, requires human interaction to create the underlying 3-D geometrical model and to make the initial associations between 2D imagery and the model. Teller et al. [15, 3] are developing a system to model outdoor urban scenes using 2-D imagery and large spherical mosaics. A number of other groups are also creating Image-Based panoramas of outdoor scenes including [11, 6].

Our approach to automatic site modeling is fundamentally different from other systems. First, we are

explicitly using range data to create the underlying geometric model of the scene. We have developed a robust and accurate method to acquire and merge range scans into topologically correct 3-D solids. This system has been tested on indoor models and we are extending it to outdoor scenes with multiple objects. Secondly, we are using our own sensor planning system to limit the number of views needed to create a complete model. This planner allows a partially reconstructed model to drive the sensing process, whereas most other approaches assume coverage of the scene is adequate or use human interaction to decide which viewing positions will be needed/used. Details on our approach are in the following sections.

The testbed we are using for this research consists of a mobile vehicle we are equipping with sensors and algorithms to accomplish this task. A picture of the vehicle is shown in figure 1. The equipment consists of an RWI ATRV mobile robot base, a range scanner (80 meter range spot scanner with 2-DOF scanning mirrors for acquiring a whole range image), centimeter accuracy onboard GPS, color cameras for obtaining photometry of the scene, and mobile wireless communications for transmission of data and high level control functions. Briefly, we will describe how a site model will be constructed. The mobile robot base will acquire a partial, incomplete 3-D model from a small number of viewpoints. This partial solid model will then be used to plan the next viewpoint, taking into account the sensing constraints of field of view and visibility for the sensors. The robot will be navigated to this new viewpoint and merge the next view with the partial model to update it. At each sensing position, both range and photometric imagery will be acquired and integrated into the model. By accurately calculating the position of the mobile base via the onboard GPS system, we can integrate the views from multiple scans and images to build an accurate and complete model. Both 3-D and 2-D data, indexed by the location of the scan, will be used to capture the full complexity of the scene.

2 Model Acquisition

We have developed a method which takes a small number of range images and builds a very accu-

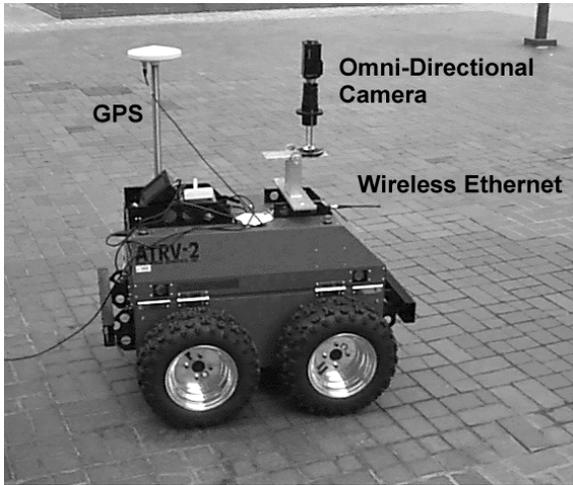


Figure 1: Mobile robot base and sensors (laser range finder not shown).

rate 3-D CAD model of an object [8, 10, 9, 2]. The method is an incremental one that interleaves a sensing operation that acquires and merges information into the model with a planning phase to determine the next sensor position or “view”. The model acquisition system provides facilities for range image acquisition, solid model construction, and model merging: both mesh surface and solid representations are used to build a model of the range data from each view, which is then merged with the model built from previous sensing operations. The planning system utilizes the resulting incomplete model to plan the next sensing operation by finding a sensor viewpoint that will improve the fidelity of the model and reduce the uncertainty caused by object occlusion (including self-occlusion).

We now describe how our system works. For each range scan, a mesh surface is formed and “swept” to create a solid volume model of both the imaged object surfaces and the occluded volume. This is done by applying an extrusion operator to each triangular mesh element, sweeping it along the vector of the rangefinder’s sensing axis, until it comes in contact with a far bounding plane. The result is a 5-sided triangular prism. A regularized set union operation is applied to the set of prisms, which produces a polyhedral solid consisting of three sets of surfaces: a mesh-like surface from the acquired range data, a number of lateral faces equal to the number of

vertices on the boundary of the mesh derived from the sweeping operation, and a bounding surface that caps one end. Each of these surfaces are tagged as “imaged” or “unimaged” for the sensor planning phase that follows.

Each successive sensing operation will result in new information that must be merged with the current model being built, called the composite model. The merging process itself starts by initializing the composite model to be the entire bounded space of our modeling system. The information determined by a newly acquired model from a single viewpoint is incorporated into the composite model by performing a regularized set intersection operation between the two. The intersection operation must be able to correctly propagate the surface-type tags from surfaces in the models through to the composite model. Retaining these tags after merging operations allows viewpoint planning for unimaged surfaces to proceed.

3 View Planning

The sensor planning phase plans the next sensor orientation so that each additional sensing operation recovers object surface that has not yet been modeled. Using this planning component makes it possible to reduce the number of sensing operations to recover a model: systems without planning tend to use human interaction or overly large data sets with significant overlap between them. This concept of reducing the number of scans is important for reducing the time and complexity of the model building process.

In cluttered and complex environments such as urban scenes, it can be very difficult to determine where a sensor should be placed to view multiple objects and regions of interest. It is important to note that this sensor placement problem has two intertwined components. The first is a purely geometric planner that can reason about occlusion and visibility in the scene. The second component is an understanding of the optical constraints imposed by the particular sensor (i.e. cameras and range scanners) that will affect the view from a particular chosen viewpoint. These include depth-of-field, resolution of the image, and field-of-view, which are controlled by aperture settings, lens size focal

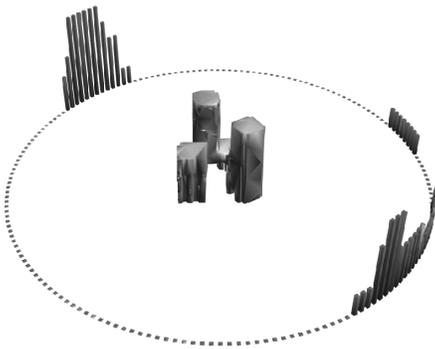
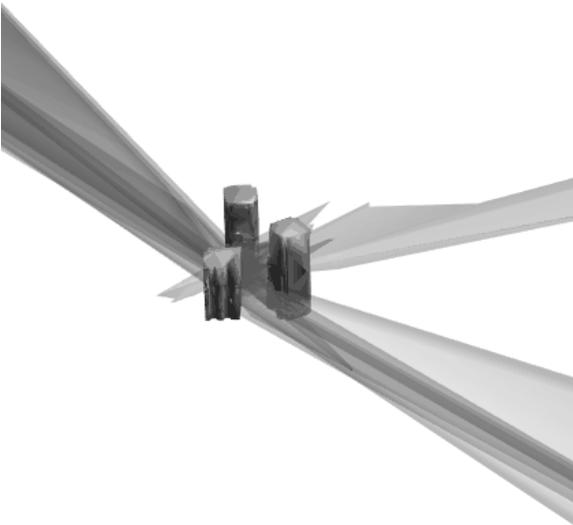
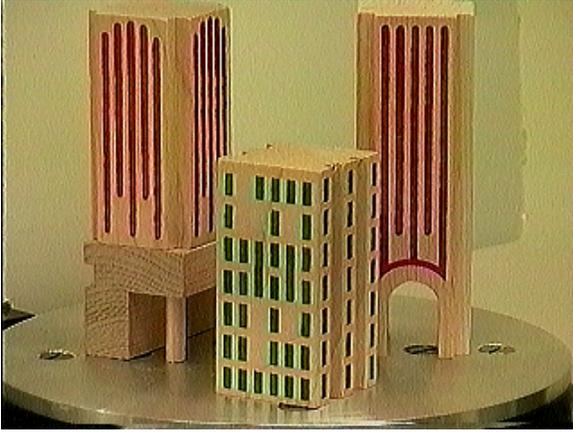


Figure 2: a) Simulated city environment on turntable. b) Visibility volume after 4 scans. c) Discretized sensor positions used to determine next view.

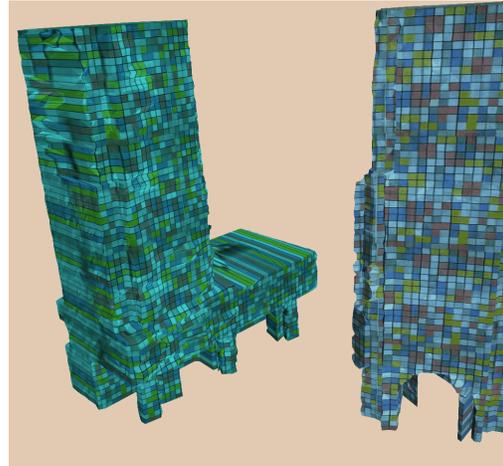
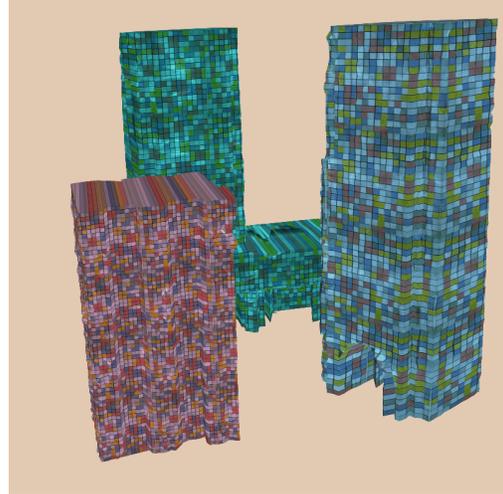
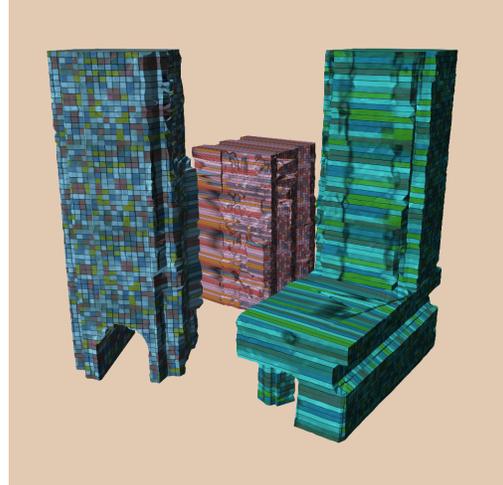


Figure 3: Recovered 3-D models - all 3 objects were recovered at once, using 12 scans with planning after the initial 4 scans. Visibility and occlusion volumes have been used to plan the correct next views for the scene to reduce the uncertainty in the model. Note recovered arches and supports

length for cameras and kinematic constraints in the case of a spot ranging sensor. To properly plan a correct view, all of these components must be considered.

The core of our system is a sensor planning module which performs the computation of the locus of *admissible* viewpoints in the 3-D space with respect to a 3-D model of objects and a set of target features to be viewed. This locus is called the *Visibility Volume*. At each point of the visibility volume a camera has an unoccluded view of all target features, albeit with a possibly infinite image plane. The finite image plane and focal length constraints will limit the field of view, and this imposes a second constraint which leads to the computation of *field of view cones* which limit the minimum distance between the sensor and the target for each camera orientation. The integration of visibility and optical constraints leads to a volume of *candidate* viewpoints. This volume can then be used as the goal region of the mobile robot navigation algorithm which will move the robot to a viewpoint within this volume.

The computation of the visibility volume involves the computation of the boundary of the *free space* (the part of the 3-D space which is not occupied by objects) and the boundary between the visibility volume and the *occluding volume*, which is the complement of the visibility with respect to the free space. In order to do that we decompose the boundary of the scene objects into convex polygons and compute the *partial* occluding volume between each convex boundary polygon and each of the targets which are assumed to be convex polygons. Multiple targets can be planned for, and the system can handle concave targets by decomposing them into convex regions. We discard those polygons which provide redundant information, thus increasing the efficiency of our method. The boundary of the intersection of all partial visibility volumes (see next section) is guaranteed to be the boundary between the visibility and the occluding volume. The boundary of the free space is simply the boundary of the scene objects.

We now describe how the planner computes visibility taking into account occlusion. The method is based on our previous work in automated visual in-

spection [13, 1]. Our model building method computes a solid model at each step. The faces of this model consist of correctly imaged faces and faces that are the result of the extrusion/sweeping operation. We can label these faces as “imaged” or “unimaged” and propagate/update these labels as new scans are integrated into the composite model. The faces labeled “unimaged” are then the focus of the sensor planning system which will try to position the sensor to allow these “unimaged” faces to be scanned.

Given an unimaged target face T on the partial model, the planner constructs a visibility volume V_{target} . This volume specifies the set of all sensor positions that have an unoccluded view of the target. This can be computed in four steps:

1. Compute $V_{unoccluded}$, the visibility volume for T assuming there were no occlusions - a half space on one side of T .
2. Compute M , the set of occluding model surfaces by including model surface F if $F \cap V_{unoccluded} \neq \emptyset$
3. Compute the set O of volumes containing the set of sensor positions occluded from T by each element of M .
4. Compute $V_{target} = V_{unoccluded} - \cup O, \forall O \in O$

The volume described by $V_{unoccluded}$ is a half-space whose defining plane is coincident with the target’s face, with the half-space’s interior being in the direction of the target’s surface normal. Each element of O is generated by the decomposition-based occlusion algorithm presented in [14], and describes the set of sensor positions that a single model surface occludes from the target. It is important to note that this algorithm for determining visibility does not use a sensor model, and in fact part of its attractiveness is that it is sensor-independent. However, for reasons of computational efficiency it makes sense to reduce the number of surfaces in M , and therefore the number of surfaces used to calculate O . This can be done by embodying sensor-specific constraints into the planner.

3.1 Example: City Scene

We now show a planning example of a complex scene using multiple targets. Figure 2a is a simulated city scene made up of three model buildings placed on a laser scanner turntable. This scene is composed of multiple objects and has high self occlusion. The modeling process was initiated by the acquisition of four range images, with 90 turntable rotations between them, to produce a preliminary model that contained many unimaged surfaces. Approximately 25% of the entire acquirable model surface is at this point composed of “occluded” surface (“acquirable model surface” in this context means those “occluded” surfaces that are not in a horizontal orientation, such as the roofs). After decimating the occluded surfaces, the 30 largest by area were chosen and a plan was generated for them. Figure 2b shows V_{target} for each of these 30 surfaces, with a decimated copy of the city scene at the center to allow the reader to observe the relative orientations. These visibility volumes are then intersected with $V_{placement}$, which is the volume representing the sensor placement constraints, to yield the sets of occlusion-free sensor positions for the targets, as shown in figure 2c. In this imaging setup of a turntable-laser, $V_{placement}$ is a cylindrical volume. A discrete solution is desired for the proper number of degrees to rotate the turntable for the next view. To accomplish this, the sensor space has been discretized every 2° , with the total target area acquired at each position found by testing the continuous-space plans for intersection with a vertical line at the appropriate position on the cylinder representing the sensor placement constraint. This is a planning histogram where the height of each bar represents the area of target surfaces visible from that sensor location, with higher bars denoting desirable sensor locations, lower ones less so. The angle of turntable rotation is found by selecting the peak in the planning histogram. After the next range image is taken, its model is merged with the existing composite model, and the planning process is repeated. After a total of 12 images have been automatically acquired, modeled, and integrated, the final model is shown in figure 3. The models have been texture mapped with a Mondrian painting *Checkerboard with Light Colors* to high-

light the geometric recovery.

3.2 Analysis: City Scene

Figure 4 shows some quantitative results from the model building phase. The entries in the table are:

- Vol - The total volume of the model.
- Surface Area - The total surface area of the model.
- Occ. Area - The total area of all occluded “unimaged” surfaces that have a significant component of their surface normals in the world x-y plane. This prevents the inclusion of “roof” features, which can not be acquired and should not be planned for, in this sum.
- Plan Area - The total surface area of the targets for which plans have been generated.
- Percent Planned - The surface area of planned-for targets, as a percentage of the total “occluded” surface area.

Each of these metrics was calculated algorithmically on the computer model. As shown in figure 4, the first 4 views were acquired without any planning (View 0 is just the entire volume before scanning). In the data describing the remaining views there are some features that seem intuitive. The total model volume decreases over time, as indeed it must for a system that uses set intersection for integration and has not duplicated any sensor viewpoints. Of particular import is the data in the final column. Because the plans are computed using a fixed number of surfaces at each iteration, it is interesting to see what percentage of the total available target area is being planned for. Clearly, if every target surface were considered, this would be 100% each time. Even though only 30 of the largest targets by area are planned for, the percent of the planned area never drops below 10% of the total area, and in most cases is over 20%. This shows that the considerable computational cost saved by selecting a subset of the targets to plan for is a viable strategy. The actual volume of the city scene has been calculated from measurements made by hand as $362cm^3$.

View No.	Vol.	Surface Area	Occ. Area	Plan Area	Percent Planned
0	4712	1571	1571		
1	1840	1317	942		
2	1052	1151	590		
3	506	733	200		
4	432	658	140		
5	416	656	121	61	50%
6	404	659	104	28	27%
7	391	657	90	12	13%
8	386	647	84	8	10%
9	382	644	75	15	20%
10	380	651	62	7	11%
11	374	622	53	16	30%
12	370	604	36	9	25%

Figure 4: Analysis of the planner’s ability to reduce uncertainty and create accurate models.

4 Integrating the Field of View Constraint for Cameras

To fully plan views, we need to take into consideration the constraints on the range scanner to scan unimaged surfaces as we did in the previous section. We also need to understand the constraints on cameras which will be used to acquire photometric properties of the scene. We now discuss the constraints related to 2-D imaging sensors. A viewpoint which lies in the visibility volume has an unoccluded view of all target features in the sense that all lines of sight do not intersect any object (other than the target) in the environment. This is a *geometric* constraint that has to be satisfied. Visual sensors however impose *optical* constraints having to do with the physics of the lens (Gaussian lens law for thin lens), the finite aperture, the finite extent of the image plane and the finite spatial resolution of the resulting image formed on the image plane, as well as lens distortions and aberrations.

An important constraint is the *field of view* constraint for a camera which is related to the finite size of the active sensor area on the image plane. Given a partial model with some target surfaces t_i we can plan a viewpoint to acquire a correct camera image of these surfaces. The targets t_i are imaged if their projection lies entirely on the active sensor area on the image plane. This active sensor area is

a 2-D planar region of finite extent. Thus the projection of the target features in their entirety on the image plane depends not only on the viewpoint P_f , but also on the orientation of the camera, the effective focal length and the size and shape of the active sensor area.

For a specific field of view angle a and a specific viewing direction \mathbf{v} we can compute the locus of viewpoints which satisfy the field of view constraint for the set of targets T . If we approximate the set of targets with a sphere S_f of radius R_f and of center \mathbf{r}_s containing them, then this locus is a circular cone $C_{fov}(\mathbf{v}, a, \mathbf{r}_s, R_f)$, called the field of view cone (see figure 5). The cone axis is parallel to \mathbf{v} and its angle is a . Viewpoints can translate inside this volume (the orientation is fixed) while the targets are imaged on the active sensor area. The locus of the apices of these cones for all viewing directions is a sphere whose center is \mathbf{r}_s and its radius is $R_f / \sin(a/2)$ (figure 5). For every viewpoint lying outside of this sphere there exists at least one camera orientation which satisfies the field of view constraint, since this region is defined as:

$$\bigcup_{\forall \mathbf{v}} C_{fov}(\mathbf{v}, a, \mathbf{r}_s, R_f)$$

For viewpoints inside the sphere there does not exist any orientation which could satisfy the field of view constraint (the camera is too close to the targets). The approximation of the targets by a sphere simplifies the field of view computation. It provides, however, a conservative solution to the field of view problem since we require the whole sphere to be imaged on the active sensor area.

The field of view angle for a circular sensor having a radius of I_{min} , is $a = 2 \tan^{-1}(I_{min}/2f)$, where f is the effective focal length of the camera. For rectangular sensors the sensor area is approximated by the enclosing circle. The field of view angle a does not depend on the viewpoint or the orientation of the camera. In our system, both the visibility volume and field of view cone are represented as solid CAD models. This allows us to use boolean set intersection on these regions of space to find an admissible viewing volume that encodes both constraints. Intuitively, this solid is the result of the intersection of the visibility volume for a target with

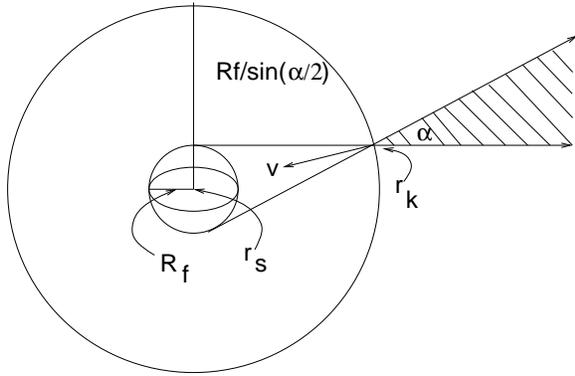


Figure 5: Field of view cone (shaded region) for viewing direction \mathbf{v} and field of view angle a . The targets are enclosed in the sphere of radius R_f

a field of view cone. For multiple targets, we simply intersect all the volumes and find the admissible region. Examples of these regions are given in the next section.

4.1 Example: Merging Visibility and Field-of-View Constraints

To test these algorithms, we have built an interactive graphical simulator where sensor planning experiments can be performed [12]. Figure 6 is an overview of this system which allows us to generate, load and manipulate different types of scenes and interactively select the target features that must be visible by a camera. The results of the sensor planning experiments are displayed as 3-D volumes of viewpoints that encode the constraints. Virtual cameras placed in those volumes provide a means of synthesizing views in real-time and evaluating viewpoints. The system can also be used to provide animated “fly throughs” of the scene.

We now describe an example of how this system works for a planning camera viewpoints in an urban environment. Figure 7 is a site model of Rosslyn, Virginia. Our input was an Open Inventor model of the city (given to us by GDE Systems Inc.), that is a set of polygons without topological information. While visually compelling, the model is not topologically correct. As stated earlier, this is not unusual - these models typically have dangling faces, unsupported structures and empty voids that can cause problems in upstream applications that expect a correct CAD model. Once we have created

a correct CAD model, we can then use the sensor planner to improve the model via navigation to regions in the scene that will allow visibility and correct field of view for imaging sensors. To do this, we have transformed this model into a CAD model using a set of interactive tools we have developed [12]. The CAD model consists of 488 buildings and we tested our sensor planning algorithms on a portion of this model whose boundary consisted of 566 planar faces (see figure 8).

In the first experiment (figure 9a) one target (black face) is placed inside the urban area of interest. The visibility volume is computed and displayed (transparent polyhedral volume). For a viewing direction of $\mathbf{v}_1 = (0^\circ, 22^\circ, 0^\circ)$ (Euler angles with respect to the global Cartesian coordinate system) and field of view angle of $a_1 = 44^\circ$, the field of view locus is the transparent cone on the left. The set of candidate viewpoints $I_1(\mathbf{v}_1, a_1)$ (intersection of visibility with field of view volume) is the partial cone on the left. For a different viewing direction $\mathbf{v}_2 = (0^\circ, 91^\circ, 0^\circ)$ the set of candidate viewpoints $I_1(\mathbf{v}_2, a_1)$ is the partial cone on the right.

In the second experiment (figure 9b) a second target is added so that two targets (black planar faces) need to be visible. The visibility volume, the field of view cone for the direction \mathbf{v}_1 and the candidate volumes $I_2(\mathbf{v}_1, a_1)$ (left) and $I_2(\mathbf{v}_3, a_1)$ (right) are displayed. The viewing orientation \mathbf{v}_3 is equal to $(0^\circ, 71^\circ, 0^\circ)$. The visibility volume and the candidate volume $I_2(\mathbf{v}_1, a_1)$ are subsets of the corresponding ones in the first experiment.

If we place a virtual camera inside the volume $I_1(\mathbf{v}_2, a_1)$ (point $(300.90, 56.18, 325.56)$), set the field view angle to a_1 and the orientation to \mathbf{v}_2 , then the synthesized view is displayed on figure 10a. The target is clearly visible. Placing a virtual camera outside of the visibility volume (point $(509.92, 41.70, 366.80)$) results in the synthesized view of figure 10b. Clearly the target is occluded by one object of the scene. The orientation of the camera is $(0^\circ, 84^\circ, 0^\circ)$ (for every viewpoint outside the visibility volume there does not exist any camera orientation that would result in an unoccluded view of the target). If we place a virtual camera on the boundary of the the candidate volume $I_1(\mathbf{v}_2, a_1)$ (point $(375.59, 52.36, 348.47)$), then in the result-

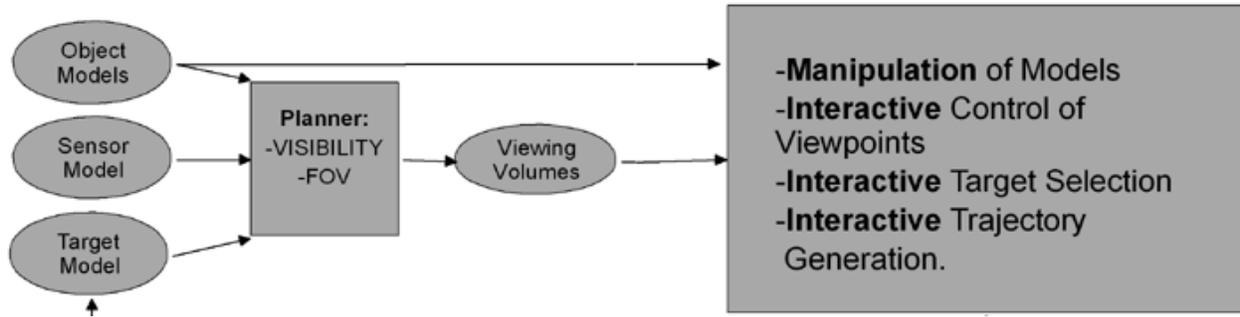


Figure 6: Interactive sensor planning system



Figure 7: VRML Graphics model of Rosslyn

ing synthesized view (figure 10c) we see that the image of the target is tangent to the image of one object of the scene. Again the camera orientation is \mathbf{v}_2 and the field of view angle a_1 .

In figure 10d we see a synthesized view when the camera is placed on the conical boundary of the candidate volume $I_2(\mathbf{v}_3, a_1)$. The camera's position is $(159.42, 30.24, 347.35)$. The transparent sphere is the sphere S_f used to enclose the targets. We see that S_f is tangent to the bottom edge of the image, because the viewpoint lies on the boundary of the field of view cone. Finally the figure 10e has been generated by a camera placed on the polyhedral boundary of the candidate volume $I_2(\mathbf{v}_3, a_1)$ (position $(254.78, 49.28, 350.45)$).

5 Conclusions

This paper describes a method for acquiring complex 3-D models from outdoor urban scenes that in-

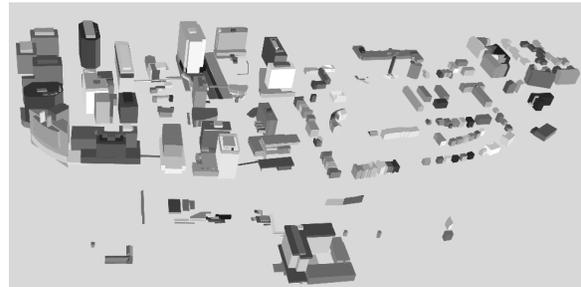


Figure 8: Solid CAD model computed from the graphics model.

cludes an integral planning component. The model acquisition process is based upon an incremental volumetric method that can merge multiple range data scans into a coherent, topologically correct 3-D model. The planner can incorporate visibility, field-of-view and sensor placement constraints in determining where to take the next view to reduce model uncertainty. Results have been presented for range data acquisition of a model of a simulated urban scene with high occlusion and for planning correct viewpoints for a camera in a model of Rosslyn, Virginia. using an interactive planning system. It can compute visibility and field of view volumes and their intersection which yields a locus of viewpoints which are guaranteed to be occlusion-free and places targets within the field of view. Object models and targets can be interactively manipulated and camera positions and parameters selected to generate synthesized images of the targets that encode the viewing constraints.

Given a partial site model of a scene, the system can be used to plan view positions for a variety of

tasks. We are currently extending this system to include resolution constraints and to create mobile robot navigation algorithms based upon the planner's output. Future work also includes fusing the range data models with the 2-D camera imagery to create even more realistic site models.

References

- [1] S. Abrams. *Sensor Planning in an active robot work-cell*. PhD thesis, Department of Computer Science, Columbia University, January 1997.
- [2] P. K. Allen and R. Yang. Registering, integrating, and building CAD models from range data. In *IEEE Int. Conf. on Robotics and Automation*, May 18-20 1998.
- [3] S. Coorg, M. Master, and S. Teller. Acquisition of large pose-mosaic dataset. In *Conference on Computer Vision and Pattern Recognition*, pages 872–878, June 23-25 1998.
- [4] B. Curless and M. Levoy. A volumetric method for building complex models from range images. In *Proceedings of SIGGRAPH*, 1996.
- [5] P. Debevec, C. J. Taylor, and J. Malik. Modeling and rendering architecture from photographs: A hybrid geometry and image-based approach. In *Proceedings of SIGGRAPH*, August 1996.
- [6] L. McMillen and G. Bishop. Plenoptic modeling: An image-based rendering system. In *SIGGRAPH '95*, pages 39–46, August 1995.
- [7] R. Pito and R. Bajcsy. A solution to the next best view problem for automated CAD model acquisition of free-form objects using range cameras. In *Proceedings SPIE Symposium on Intelligent Systems and Advanced Manufacturing, Phila. PA*, 1995.
- [8] M. Reed and P. K. Allen. 3-D modeling from range imagery: An incremental method with a planning component. *Image and Vision Computing*, (to appear) 1998.
- [9] M. Reed, P. K. Allen, and I. Stamos. Automated model acquisition from range images with view planning. In *Computer Vision and Pattern Recognition Conference*, pages 72–77, June 16-20 1997.
- [10] M. Reed, P. K. Allen, and I. Stamos. Solid model construction using meshes and volumes. In *Proc. DARPA Image Understanding Workshop*, pages 921–926, May 12-14 1997.
- [11] H. Shum and R. Szeliski. Panoramic image mosaics. Technical Report TR-97-23, Microsoft Research, 1997.
- [12] I. Stamos and P. K. Allen. Interactive sensor planning. In *Computer Vision and Pattern Recognition Conference (CVPR)*, pages 489–494, June 21-23 1998.
- [13] K. Tarabanis, P. K. Allen, and R. Tsai. The MVP sensor planning system for robotic vision tasks. *IEEE Transactions on Robotics and Automation*, 11(1):72–85, February 1995.
- [14] K. Tarabanis, R. Y. Tsai, and A. Kaul. Computing occlusion-free viewpoints. *IEEE Transactions Pattern Analysis and Machine Intelligence*, 18(3):279–292, March 1996.
- [15] S. Teller. Automatic acquisition of hierarchical, textured 3d geometric models of urban environments: Project plan, view planning for site modeling. In *Proc. DARPA Image Understanding Workshop*, pages 767–770, New Orleans, 1997.
- [16] W. Thompson, H. de St. Germain, T. Henderson, and J. Owen. Constructing high-precision geometric models from sensed position data. In *Proceedings 1996 ARPA Image Understanding Workshop*, pages 987–994, 1996.
- [17] G. Turk and M. Levoy. Zippered polygon meshes from range images. In *Proceedings of SIGGRAPH*, pages 311–318, 1994.
- [18] M. Wheeler. *Automatic Modeling and localization for object recognition*. PhD thesis, School of Computer Science, Carnegie Mellon University, October 1996.

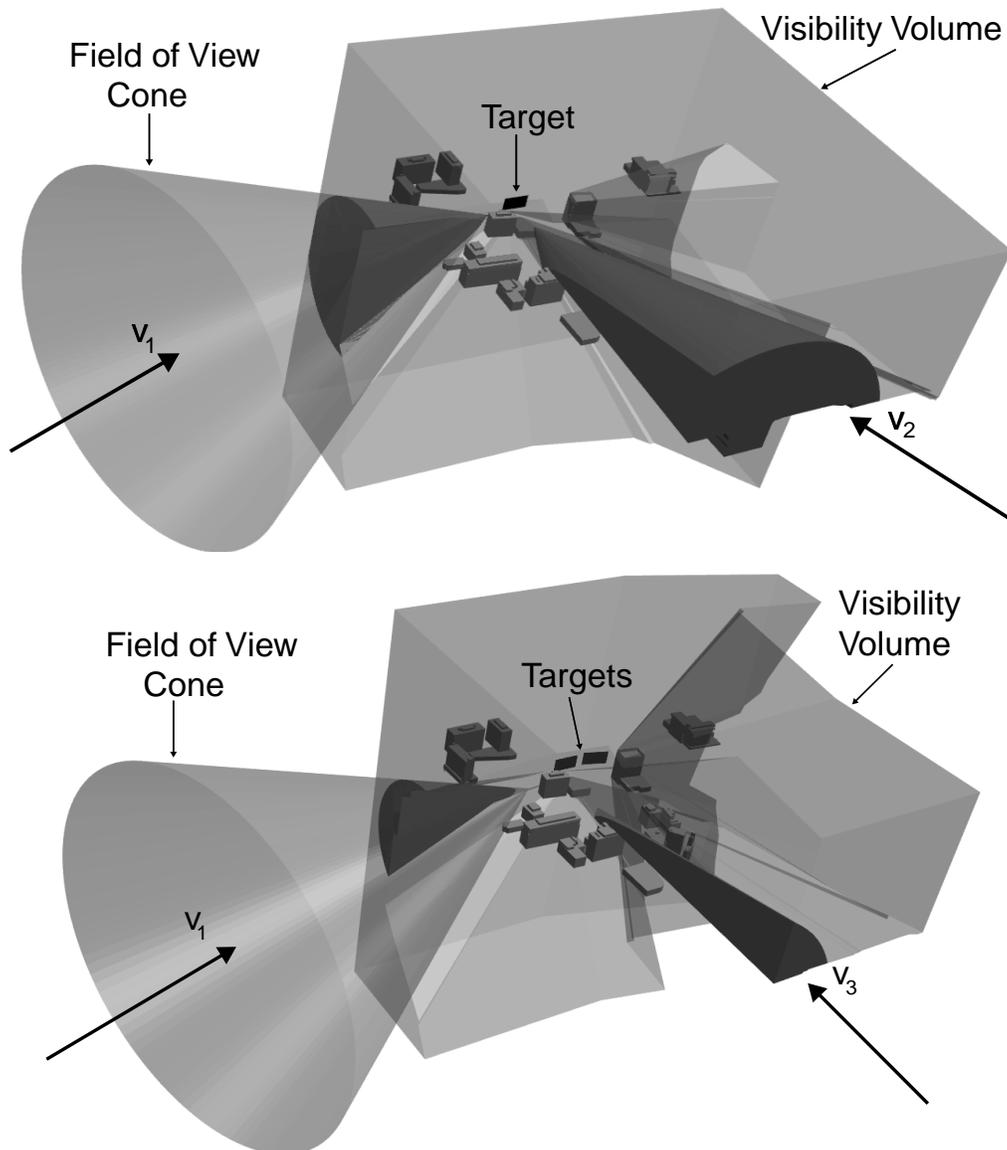


Figure 9: Two experiments. a) (top figure) One target and b) (bottom figure) two targets are placed in the urban area. The targets are planar faces. The Visibility Volumes (transparent polyhedral volumes), the Field of View Cones for the direction v_1 (transparent cones) and the Candidate Volumes (intersection of the visibility volumes with the field of view cones) for the viewing direction v_1 (left partial cones) and for the directions v_2 (right partial cone, top figure) and v_3 (right partial cone, bottom figure) are displayed. The Field of View Cones for the directions v_2 (top) and v_3 (bottom) are not shown.



Figure 10: Synthesized views. Single target (black face): the camera is placed a) (left image) inside the candidate volume, b) out of the visibility volume and c) on the boundary of the candidate volume. Two targets: the camera is placed on d) the conical boundary and e) the polyhedral boundary of the candidate volume.