Purdue University

Purdue e-Pubs

Department of Computer Science Technical Reports

Department of Computer Science

2006

Efficient Large Scale Acquisition of Building Interiors

Gleb Bahmutov

Voicu Popescu
Purdue University, popescu@cs.purdue.edu

Mihai Mudure

Report Number:

06-007

EFFICIENT LARGE SCALE ACQUISITION OF BUILDING INTERIORS

Gleb Bahmutov
Voicu Popescu
Mihai Mudure

Department of Computer Sciences
Purdue University
West Lafayette, IN 47907

CSD TR #06-007
March 21, 2006

Efficient Large Scale Acquisition of Building Interiors

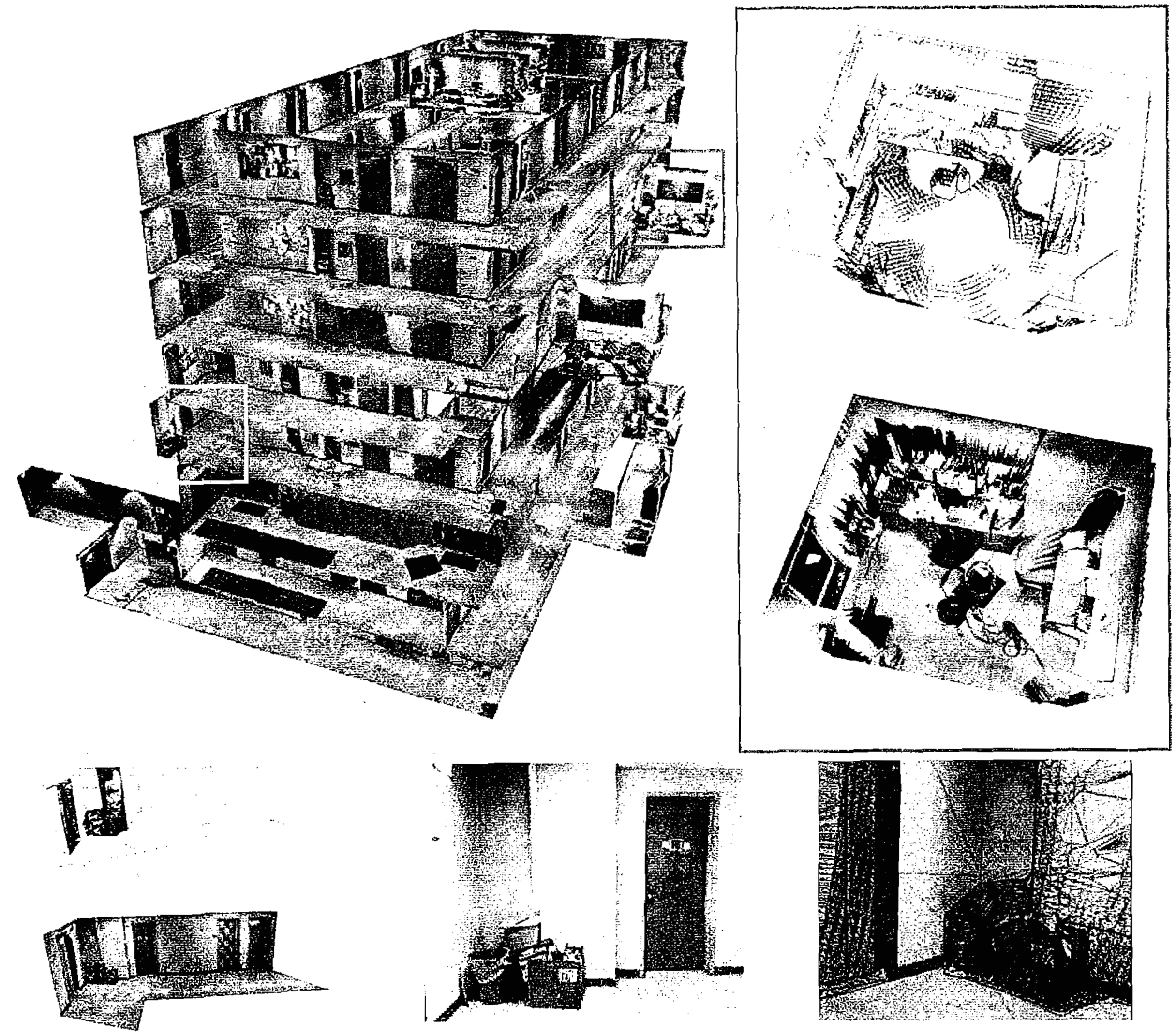

Figure 1: *This model was acquired in 40 hours by a two person team using a single acquisition device. The model spans 6 floors with 20 attached rooms. Individual rooms (see blue) and corridors (see orange) are modeled with fitted proxy geometry enhanced with embedded detail.*

Abstract

We describe a system for the rapid acquisition of building interiors. In 40 hours, a two member team with a single acquisition device captured a model of the corridors and 20 individual rooms spanning 6 floors of a large building. Our custom acquisition device operates at interactive rates. The system provides immediate feedback to the operator. The operator guides the acquisition device in real time and trivially avoids over sampling the planar parts of the scene such as floors, ceilings, walls, or doors. Most of the acquisition time budget is spent on the parts of the scene with complex geometry. A corridor section is modeled by acquiring a depth enhanced panorama (DEP) at each one of its two ends and by fitting proxy geometry to the two DEPs. A room is acquired with a single DEP and proxy geometry is fitted to the planar parts. A room or a corridor section is acquired in less than 15 minutes. The acquisition device acquires high quality color intrinsically registered with the depth data. The resulting model is a texture-mapped triangle mesh that supports photorealistic interactive rendering and is suitable for applications such as virtual training and simulation.

Categories and Subject Descriptors (ACM CCS): I.3.5 [Computer Graphics]: Computational Geometry and Object Modeling.

1 Introduction

Geometry and color models of large scale indoor environments are invaluable to numerous applications in science, engineering, defense, art, and entertainment. Simulations of fire propagation in models of existing buildings lead to safer constructions. A model that captures the individual furniture configuration of hundreds of rooms enables a high-fidelity simulation of the propagation of an airborne contaminant. Virtual training of emergency response personnel is more effective in environments that capture the true complexity of real-world scenes. Virtual tourists will benefit from exploring the buildings of an entire historic district. Real estate designers, constructors and marketers will benefit from photorealistic interactive architectural walkthroughs.

The state of the art in automated modeling offers good solutions for *outside-looking-in* modeling, when a relatively small scene is acquired using a few viewpoints located outside the scene. No complete solution exists for modeling in the *inside-looking-out* case, when the acquisition device and the operator are immersed in the scene to be acquired. The sheer size of the scene with complex occlusions and large range of depths make inside-looking-out acquisition particularly challenging.

Methods that acquire dense depth, such as depth from stereo, depth from structure light, or time-of-flight laser range finding scale poorly with the size of the scene. The only automated modeling technique that allows capturing large scale scenes at an accessible time and equipment cost are color panoramas [Che95]. Panoramas are easy to acquire, support photorealistic interactive visualization, but the user is confined to a single viewpoint. While an object can be easily shown from all angles with a collection of images, relying solely on images in the inside-looking-out case limits precludes many of the important applications listed above.

We describe a system for modeling and visualizing large scale building interiors. The system is efficient and the resulting model supports high-quality visualization at interactive rates. We remove the single viewpoint limitation of color panoramas without sacrificing their efficiency and low cost. We use a custom structured light device consisting of a camera and an attached laser that casts a matrix of 11x11 laser beams in its field of view. The operator sweeps the scene to collect color and depth data, which is visualized in real time. The operator uses the immediate feedback to avoid over-sampling flat regions and to assess the quality of the model. Parts of the scene with complex geometry are refined until the desired level of detail is reached. Data acquisition problems are detected and addressed right away. Operator input is used again during modeling to guide the fitting of proxy geometry and to assemble model sections.

The accompanying video and Figure 1 show a model efficiently acquired with our system, which covers a significant part of the interior of a large building. The model is assembled from individual room and corridor sections. A section is modeled as a texture-mapped mesh obtained by fitting proxy geometry to the depth data and keeping embedded detailed geometry where needed. The model captures the large indoor space convincingly by preserving the complexity of the cluttered rooms, yet supports interactive visualization by exploiting the geometric simplicity of corridors, walls, and ceilings.

In the next section we discuss prior systems for indoor environment acquisition. Section 3 gives an overview of our system. Section 4 describes our acquisition device. Section 5 describes the acquisition and modeling of individual sections. In Section 6 covers assembling the model from individual sections. We conclude with results and discussion.

2 Prior work

Modeling large scale indoor scenes requires capturing color and depth data from multiple viewpoints. The wide availability of photo and video cameras makes acquisition of high quality color an easy task. Sampling the scene geometry is far more challenging. We structure the discussion of prior work according to the method employed for depth acquisition.

Dense depth acquisition techniques have been used to model complex scenes such as Jefferson's Monticello [Wil03] or the Parthenon [Stu03]. Acquired accurate depth maps can be processed into high-quality models. Systems that rely on dense depth suffer from long per-view acquisition times, which limits the number of viewpoints, and from the fact that the operator has little or no control over the acquisition process. Due to the delay between the scanning phase and the time the model is available for inspection and validation, addressing problems with calibration, depth acquisition, or scene coverage is usually impractical due to the high cost of returning to the scanning site. In the case of indoor building environments, a time-of-flight laser rangefinder over-samples planar regions, which hurts the modeling pipeline at every stage: time is wasted acquiring, transferring, and simplifying a large number of coplanar points.

Other techniques rely on the user to manually enter the geometric data, either directly using a modeling package (AutoCad, 3dMax) or indirectly by specifying geometric constraints (such as line, plane and object relations) [DTM96], [HH02], [ZD*01]. This approach leverages the domain knowledge of the user, who maximizes the expressivity of the model while minimizing the complexity of the geometry. The main disadvantage of these methods is over-simplifying the parts of the scene with complex geometry. Adding geometric detail by hand in every one of tens of rooms is prohibitively slow. Therefore the resulting model fails to capture the complexity of individual rooms, and has an artificially clean appearance. This adversely affects, for example, virtual training applications where the simulated conditions have to be as realistic as possible.

Another set of techniques avoid depth acquisition altogether and rely exclusively on color. This reduces acquisition times and equipment costs, making them appealing for large scale modeling. Color panoramas [Che95] are two-dimensional ray databases built from multiple images sharing the same center of projection.

Color panoramas produce high-quality images of the scene in any view direction, but the user cannot translate away from the acquisition point. Several attempts have been made to alleviate this problem. Shum [SSB*98] extends color panoramas by inferring simple scene geometry from user specified geometric constraints. A similar approach is taken by the ICARUS system [GCH*02], where the user places geometric primitives guided by calibrated photographs. The resulting model has the advantage of photorealistic color originating from the photographs used as texture, but over-simplifies the geometry of complex rooms.

The work presented here builds upon depth enhanced panoramas (DEP) [Bah05]. DEPs are built from sequences of dense color and sparse depth frames that share the same center of projection. Leveraging the single perspective constraint, frames are registered automatically by minimizing color differences at overlap regions. A cube map color panorama is built which is enhanced with the depth samples contributed by the individual frames. The depth samples are triangulated in 2D on the faces of the cube map and the connectivity data so inferred is applied to the corresponding 3D points to obtain a texture-mapped 3D triangle mesh.

The contributions of this paper are:

- Development of an acquisition device optimized for large scale indoor modeling. Color panoramas and sparse depth are acquired robustly, and the acquisition direction is controlled by the operator in real time.
- A software system for the efficient processing of the acquired color and depth data into a compact photorealistic model of the indoor environment.
- The acquisition of a large-scale model to demonstrate the capabilities of our system.

3 System overview

Out modeling system targets buildings where each floor has a corridor with attached rooms. Examples include office buildings (the model shown in Figure 1 is a building that houses a department on our university campus), hospitals, hotels, and apartment buildings. The corridors are assumed to have a rectangular cross-section. We handle corridor turns, loops, and junctions, as well as occasional objects or corridor sections with high geometric complexity. There are no restrictions on the geometric complexity of the rooms. A single acquisition viewpoint is used for each room, so a geometric model of higher fidelity is obtained when a viewpoint can be found from where most room surfaces are visible. Our system does not have the depth acquisition range needed for large indoor spaces such as theaters or warehouses.

The acquisition device is mounted on a tripod. The operator acquires a DEP in each room, and several DEPs along the corridor. Once the DEP of a room is acquired a room section is built by fitting a box, and by removing the unnecessary points lying on the walls, ceiling or floor.

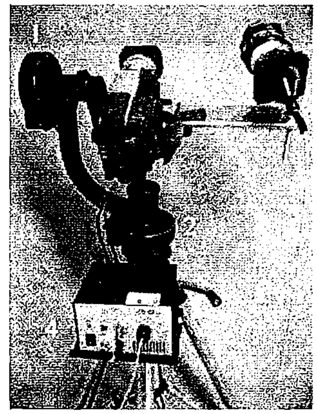

Figure 2: Acquisition device on a tripod. Shaft encoders (1, 2) report current tilt and pan angles. Laser diode (3) is powered by its own power converter (4).

Once a corridor DEP is acquired it is used to generate an I, L, or T corridor section. The geometric detail is preserved where needed. Corridor sections are connected into corridor loops using minimal operator input, by leveraging same-plane constraints. Color is blended over the transition region to alleviate exposure differences between DEPs. Room sections are attached to corridor loops the same way, to generate the building model.

4 Acquisition device

We have developed a structured light acquisition device (Figure 2) based on the ModelCamera [Bah05]. The matrix of laser beams is generated with a single laser source whose beam is split using a diffraction grating. Since the laser is fixed with respect to the camera, each beam projects in the frame to a constant epipolar segment. We optimized the diffraction grating to make use of the entire vertical field of view of the video camera. This allowed increasing the number of laser dots from 7x7 to 11x11, while maintaining the same distance between neighboring dots. In order to be able to register frames robustly even in the absence of color, we enhanced the parallax free bracket with shaft encoders, which report the current pan and tilt angles.

There are five steps to the calibration of the acquisition device. In a first step, intrinsic optical properties of the camera are calibrated using a calibration grid [BP99]. In a second step, the epipolar line for each laser is found from laser dot snapshots. In a third step, the corresponding 3D laser ray is computed. These steps are similar to those described in detail in [PSB04].

In a fourth step the pan and tilt axes of the bracket are computed in the coordinate system of the camera. As the camera is rotated around one axis, several overlapping frames are registered by minimizing color error over three rotational degrees of freedom. The found angles are converted to a single rotation which gives the axis. To reuse the tilt axis we start each acquisition system from the same tilt angle marked on the shaft encoder drum. These

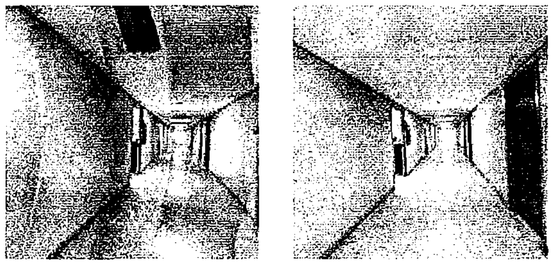

Figure 3: *Panorama face without blending (left) and panorama face blended in real time (right).*

four steps take approximately 20 minutes, and their result is reused in many acquisition sequences.

In a fifth step, which is repeated for each acquisition sequence, we measure the offset between the PC clock used to poll the shaft encoders and the camera clock used to time stamp the video frames. This allows synchronizing the angle readings with the frames without using an explicit (hardware) sync between the computer and the camera.

The offset is determined by taking advantage of the fact that the shaft encoders can be polled very frequently (10 times each millisecond), which for our application is equivalent to instantaneous angle reads. The second fact used in the calibration of the clocks offset is that the acquisition times of the video frames are evenly spaced in time. The operator pans the camera over a part of the scene with high color detail. The shaft encoder angles are read in frequently (every millisecond) and stored in a buffer together with the PC time when they were acquired. Using the known pan axis, the frames are registered in 1D using color. Once the pan angle of a frame with camera timestamp c is known, the angle is looked up in the buffer of shaft encoder angles. The corresponding PC clock timestamp p is used to compute the delay as p-c. The precision of the calibration increases with panning speed, since this shortens the time interval where the angle values stay constant. By panning 30 degrees in 4-5 seconds, we typically obtain 6 – 10 delay values agreeing within 1 millisecond. The calibrated delay is used during acquisition to look up the pan and tilt angles for each frame in a buffer indexed this time by the video frame timestamp.

5 Modeling of building sections

5.1 Data acquisition

We first acquire a color cube map using blending on per tile basis to minimize frame to frame camera exposure differences. The operator pans and tilts the video camera to cover all directions, using immediate feedback. A complete cube map is acquired in 4-5 minutes. Registration is achieved robustly using the shaft encoders. After color acquisition the operator turns on the laser. Bright red dots are found in the frame using a 1D search along epipolar lines, and the corresponding 3D points are computed by triangulation. Using the immediate feedback the operator sweeps the dots over complex geometry until the desired geometry resolution is attained.

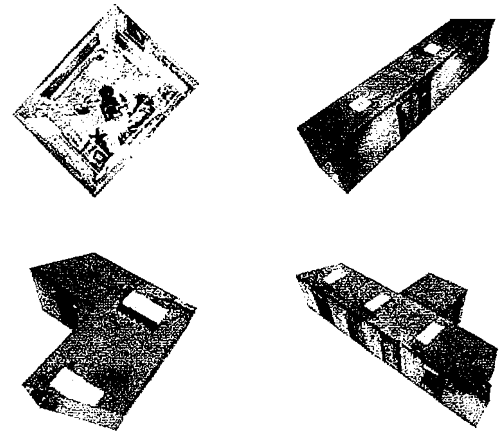

Figure 4: *Four section types are fitted to the DEP data. Room section is shown with the point cloud in the top left corner.*

Modern consumer video cameras capture very high quality color, but often overlapping frames have drastically different brightness, as the camera automatically adjusts exposure. Our real-time blending method minimizes these differences, creating smooth color transitions between parts of the color panorama taken from different frames (see Figure 3). We divide each face of the cubic panorama into square tiles. A tile is filled with the color from a given video frame if it is empty, or if the frame's brightness is higher than the tile's current brightness. This approach works well indoors: detail is captured in the darker parts of the scene, while only saturating the fluorescent lights on the ceiling. Each tile is larger than its contribution to the panorama to allow efficient blending of tiles with its neighbors. We found that 32 by 32 pixel tiles, with an additional border of 16 pixels is a good compromise between processing speed and quality of the resulting color.

The bracket does not allow capturing color right above the camera. As a temporary solution we fill this gap with color from the surrounding regions. The tripod interferes with color acquisition directly beneath the camera. The hole in the floor can be easily filled in for corridors due to the repetitive nature of the texture, but filling in color for complex rooms is more challenging.

We acquire between 60 and 200 thousand depth samples in 3-4 minutes. In order to reduce the number of false positive laser dot detections, candidate dots have to pass two tests. Firstly, a new dot has to be within epsilon of the location where it was found in the previous k frames. A legitimate jump from one surface to another is validated after k frames. Secondly, dots cannot be located at the same location in the cube map as the camera is rotated. This indicates confusing the dot with a scene feature. The number of false positives that pass these tests is negligible (less than 100 per DEP). Moreover, the false positives generate points that are clustered in front or behind the scene, which makes selecting and deleting them straightforward.

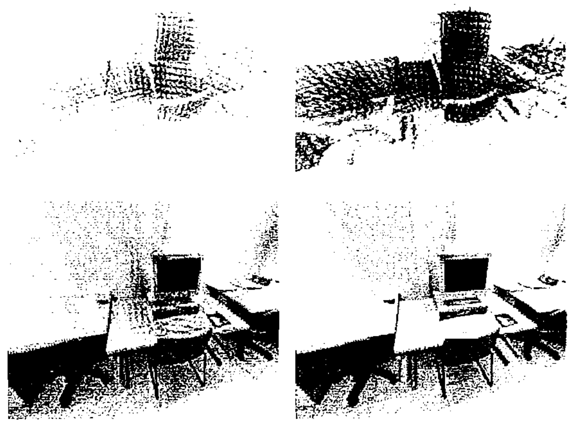

Figure 5: *Points lying within threshold distance to the fitted box are marked as invalid (red). Triangles connecting all invalid points are discarded, leaving only triangulation of the fragmented geometry. A complete room model is a union of fragmented geometry and fitted section triangles (bottom right).*

5.2 Proxy fitting

We fit proxy geometry to the corridor or DEP rooms. A corridor DEP has depth points for only a 2m band of the corridor tube. Outside of the band the DEP stores only color. We assume that the ceiling and floor are parallel, and that opposite walls are parallel to each other and perpendicular to the ceiling and floor. We fit four types of proxy geometry (see Figure 4): a rectangular box for rooms, an I section for a simple straight corridor piece, an L section for a corridor corner, and a T section for a corridor junction.

The fitting process starts by fitting floor and ceiling planes through points selected by the operator. Then the operator specifies the lines in the color panorama where the walls intersect the ceiling plane. A downhill simplex search finds wall planes perpendicular to the ceiling plane closest to the lines specified by the operator. The walls are intersected with floor and the ceiling to complete the proxy geometry. Orthographic texture is computed for each triangles from the cube map.

In some cases the fitted section obstructs part of the scene: shelves behind the fitted wall of the room section, open door in the corridor section leading into a room, etc. The operator can specify a region of the section to cut out. The remaining region is retriangulated automatically. The resulting triangles reuse the texture from the previous triangulation saving color reprojection costs.

5.3 Fragmented geometry

Building interiors are more complicated than the simple planar boxes of the proxies. Our system allows modeling the complex geometry inside rooms and enhancing the corridors with occasional geometric detail.

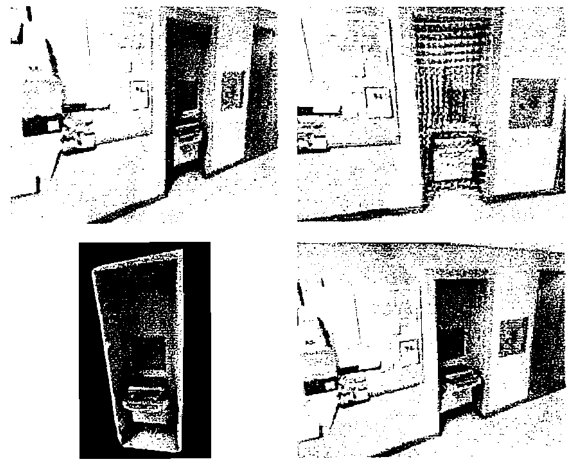

Figure 6: *Some objects appear flat in the corridor section. Operator can select the region, all depth samples are triangulated into a texture mapped mesh. The result embedded into the original corridor section provides motion parallax effect.*

Recall that a DEP is a texture-mapped triangle mesh acquired from the acquisition point. The DEP is combined with a room box by first eliminating the DEP points that are close to the box. The threshold used in practice is 7cm. To make geometry recessed behind wall planes visible (e.g. water fountain in Figure 6), the operator cuts an opening in the wall in 2D, using a view from the center of the DEP. The points mapping to the opening are excluded from the planarity test. The second step is to sieve the connectivity data eliminating triangles for which all three vertices were discarded. This effectively flattens the part of the floor, ceiling, and walls visible in the DEP. The resulting section shows the overall room with flat walls, ceiling and floor, and with furniture and other geometric detail "sticking out" from these planes (Figure 5).

Many objects in the corridors can be truthfully modeled with texture data alone: doors, posters, ceiling lights. Occasionally, there are objects in the building corridors whose lack of geometry cannot be hidden with texture (e.g. benches, trash cans, fire extinguishers, and water fountains). They appear noticeably flat on the triangles of the corridor sections. Moreover, many of these are important for applications since they hinder access, or are useful in emergency response During the acquisition the operator can sample the geometry of these objects. After the section is fitted through the data, the operator can select region of the panorama with additional objects. We construct a plane cutting through the points in the selected region. The points then are triangulated in 2D by projecting onto the plane, color from the panorama is used to texture map the resulting 3D mesh connecting the points. The mesh is stored with the section, and when rendered the additional geometry provides correct visual clues in the novel views of the objects (see Figure 6).

The triangulation of the sparse depth samples does not preserve correct depth discontinuities, resulting in visible artifacts in the rendered model: broken edges of the tables, shelves, monitors. The operator can improve the quality of

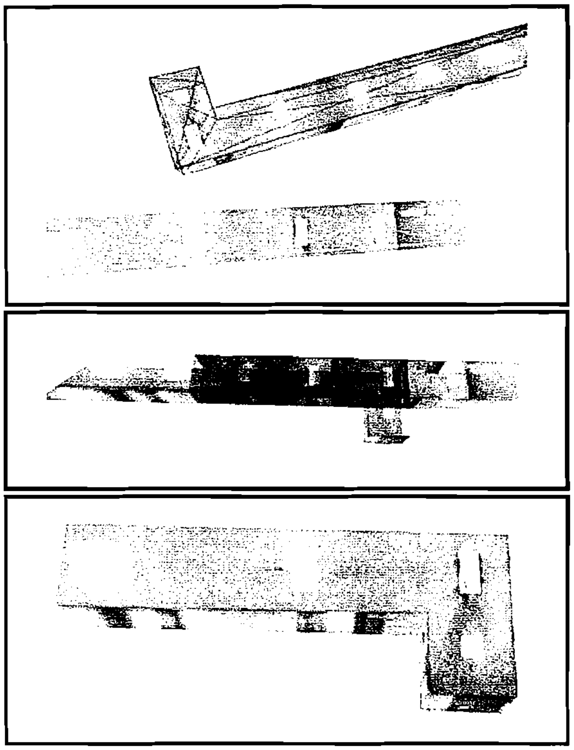

Figure 7: *Section registration: (top) original sections, (middle) after applying geometrical constraints operator slides one section along the remaining free translation to match colors in the shared region, (bottom) recomputed triangles show the final result.*

the model, by introducing additional depth samples into the model. Two primitives can be fitted to regions specified by the operator: lines and planes. If a operator specifies a 2D segment, the corresponding 3D segment passing through the depth samples lying nearby in 2D is used to introduce new 3D samples computed at regular intervals (1 cm works well). Similar process interpolates points on the boundary of the plane fitted through points in the region specified by the operator. Line interpolation helps to improve the straight edges of the shelves, and plane interpolation is useful with tables, monitors and other planar patches. Typically, the operator spends 10-15 minutes improving the visual appearance of the scanned room.

5.4 Combining sections

After acquiring depth enhanced panorama and fitting corridor section, the operator registers new section with a previous one. There are six free registration parameters to position the new section using a solid motion. The same floor and one of the wall planes seen in both sections determine five parameters. The remaining translation along the corridor is found using input from the operator, who slides the section until the colors in the overlapping region match (see Figure 7).

The triangulated model of the building corridors is recomputed every time a new corridor piece is registered. Because new section can be linked to sections besides its

Figure 8: *This floor consists of 12 corridor sections and 4 rooms. Different sections and corresponding point clouds are shown with different colors.*

registration partner, for each corridor section the adjacent pieces are found. Then for pairs of neighboring sections, shared triangles connecting vertices of the cross sections are computed. The texture for shared triangles is a blend of color from two panoramas to smoothen the color transition from one acquisition point to another.

The room sections are registered the same way as corridor sections, but the room section is positioned on the outside of the corridor. The model of the building does not have to be recomputed, since a solid wall divides new room section from the corridor sections. Using these four section types we can reconstruct fairly complex floor plans, such as the floor shown in Figure 8. Note that the fitted sections allowed registering non overlapping point clouds.

6 Results

We have placed the computer on a cart with wheels, together with the monitor to display the results in real time and an uninterruptible power supply to allow switching from one power outlet to another without shutting down the system. The acquisition device was placed on a sturdy tripod. Two people were needed to move this setup: one to push the cart, and another to carry the tripod with the camera. Only one person was needed during the actual scanning. The second person was identifying the next room to be scanned. The itinerary was finalized on the fly since we did not want to impose pre-established scanning times on those that had offices in the building. The 20 rooms were acquired over two days, with a single pass on each floor.

Corridors were captured by acquiring DEPs every 7 – 9 meters apart. The longest corridor section measures 36m and it was acquired with 5 DEPs from end to end. To minimize disruption to the normal activity in the building we did not cordon off the scene during scanning. The interruptions due to people moving through the corridors had a negligible impact due to the interactive nature of our acquisition pipeline. During the acquisition of a corridor

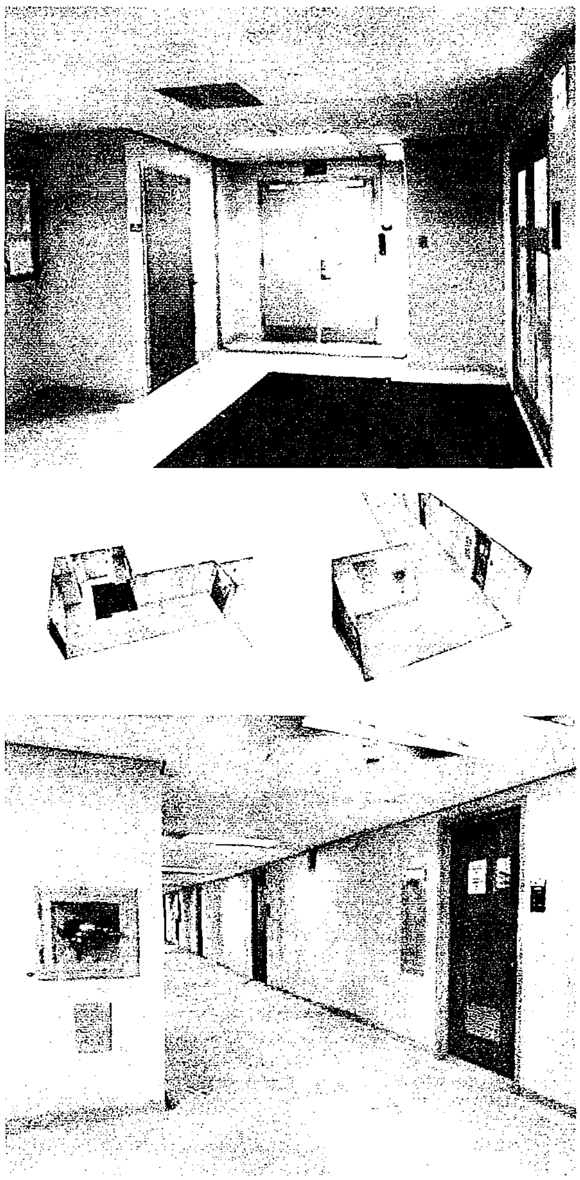

Figure 9: *Corridors rendered from novel viewpoints.*

DEP the cart was moved to remain outside of the field of view of the camera and therefore outside the panorama.

For room acquisition, the room had to be vacated for a total of 10 minutes. The device was positioned in the center of the room, and the cart was in the door frame. The door frame was cut out in the fitted box with operator input. On average, we spent 7/9 minutes acquiring depth and color for a corridor/room DEP. The longer acquisition times for the DEP are necessary to capture the more complex geometry.

The building model shown in Figure 1 contains 56 corridor sections and 20 individual rooms, spanning 6 floors. The corridor sections cover about 1,130 square meters of floor space. The room models cover 320 square meters of floor space. The original data for each room contains on average 110K depth samples. After discarding samples lying within 7 cm from the fitted section planes, 60K samples remain in each room, on average. For corridors sections we have acquired on average 38K samples, from which 5K samples were kept.

Section fitting takes on average 3 minutes, including computation of orthographic textures for the triangles of the section. It took about 2 minutes to register a pair of sections, and to recompute the shared textures, for a total per section time of less than 15 minutes. The proxies used in the model total less than 1,000 triangles. The fragmented geometry inside the room sections is modeled with 97K triangles per room, on average.

We have measured the dimensions of the longest corridor span on each floor. The average error in our model was 2.5%, although in one case the length of the corridor was off by 4.5%.

Our model is a set of texture mapped triangles saved in the VRML format. The model can be rendered with standard graphics APIs implemented in hardware (see Figure 10 and Figure 9). The full resolution model contains ~2 million triangles and over 2GB of textures. When the application desires to render the entire model a version with down sampled textures (4x4) and decimated geometry (90%) is used to enable interactive rates.

7 Conclusions and future work

We have described a system for the large scale acquisition of building interiors. The system relies on a custom acquisition device that captures color panoramas and sparse depth reliably, in real time. The short acquisition time enables an interactive automated modeling pipeline which is substantially more efficient than pipelines based on lengthy acquisition of dense depth maps. Once the operator is effectively integrated in the modeling loop, modeling greatly benefits from the operator's understanding of the scene. The operator monitors data acquisition and naturally aims the acquisition device towards scene with complex geometry, maximizing scanning efficiency. We validated the system by acquiring a significant fraction of a large building. With only a minor interference with the normal activity in the building, a team of two operators built what is, to the best of our knowledge, the largest inside looking out model.

We will continue to perfect the system. Low level development will improve the usability of the various software tools, as well as making the hardware more maneuverable. A tripod with wheels and battery power will allow acquiring data a lot more efficiently, by a single operator. Another small improvement that with great benefit is modeling some of the materials frequently repeated throughout a large building. The first candidate is the linoleum on the corridor and room floors. The specularity is not negligible and accounting for it will increase the realism of the walkthroughs by replacing the presently frozen highlights with correct, dynamic highlights. This also requires solving the problem of locating the light sources. Again, we plan to exploit the model regularity. Selecting two fluorescent lighting groups should allow for the automatic instantiation of the remaining lights.

Adding wireless connectivity will allow the second operator to fit the proxies and connect the sections remotely, from a model integration station. We do not foresee any major difficulty since the incremental updates to the color cube map and to the set of 3D points have a compact memory footprint. Also as future work we will investigate scanning with several acquisition devices in parallel, which could all be served by the same model integration station.

8 Acknowledgments <withheld>

References

[Bah05] Bahmutov, G., et al 2005. Depth Enhanced Panoramas. Second International Conference on Video, Vision, and Graphics. Edinburgh, July 2005.

[BP99] Bouguet, J.-Y. and Perona, P. 1999. 3D Photography using Shadows in Dual-Space Geometry. *International Journal of Computer Vision,* Vol. 35, No. 2, 1999, pp. 129-149.

[Che95] S. Chen, Quicktime VR - An Image-Based Approach to Virtual Environment Navigation, Proc. SIGGRAPH 95, 29-38 (1995).

[DTM96] Debevec, P., Taylor, C., and J. Malik. Modeling and Rendering Architecture from Photographs: A Hybrid Geometry and Image Based Approach. Proc. SIGGRAPH '96, 11-20 (1996).

[GCH*02] Gibson, S., Cook, J., Howard, T., and Hubbold, R.: ICARUS: Interactive Reconstruction from Uncalibrated Image Sequences, In Proc of SIGGRAPH 2002 courses, (2002).

[HH02] Hidalgo, E. and R. J. Hubbold. Hybrid geometric-image-based-rendering. Proceedings of Eurographics 2002, Computer Graphics Forum, 21(3):471-482, September 2002.

[PSB04] Popescu, V., Sacks, E., and G. Bahmutov, 2004. Interactive Modeling from Dense Color and Sparse Depth. In *Proceedings of Second International Symposium on 3D Data Processing, Visualization, and Transmission.* Thessaloniki, Greece, September 2004.

[SSB*98] Shum, H., Szeliski, R., Baker, S., Han, M., and P.Anandan., Interactive 3D Modeling from Multiple Images Using Scene Regularities. Proceedings of the European Workshop on 3D Structure from Multiple Images of Large-scale Environments (at ECCV '98), 1998.

[Stu03] Stumpfel, J. *et. al.* Digital Reunification of the Parthenon and its Sculptures, 4th International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage, Brighton, UK, 2003.

[ZD*01] Zhang, L., Dugas-Phocion, G., Samson, J., Seitz, S.. "Single View Modeling of Free-Form Scenes," *CVPR*, p. 990, 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'01) - Volume 1, 2001.

[Wil03] Williams, N. *et. al.* Monticello Through the Window. Proceedings of the 4th International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage (VAST 2003), Brighton, UK (November 2003).

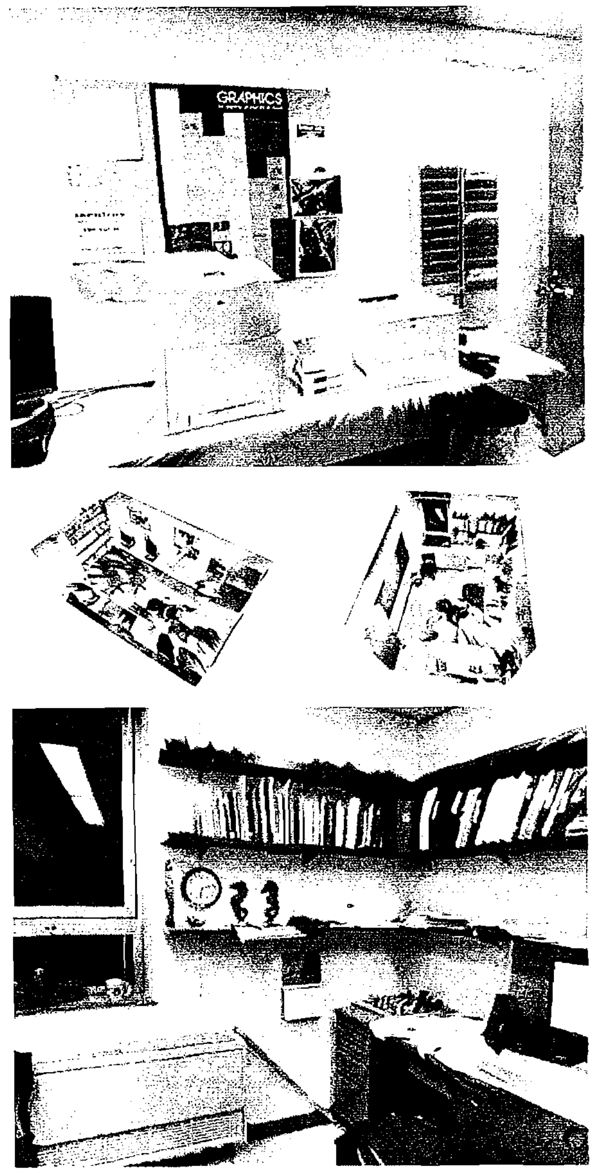

Figure 10: *Rooms rendered from novel viewpoints.*