1001 Acquisition Viewpoints

Efficient and Versatile View-Dependent Modeling of Real-World Scenes

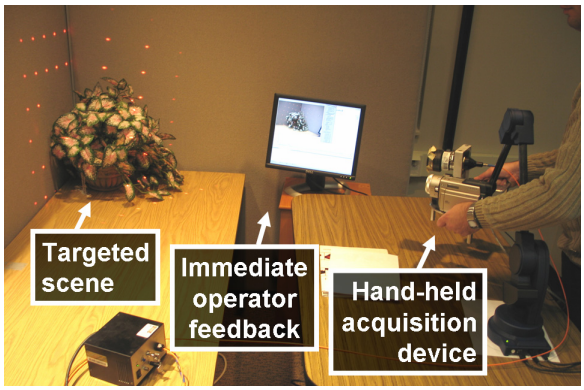


Figure 1: *Acquisition system*

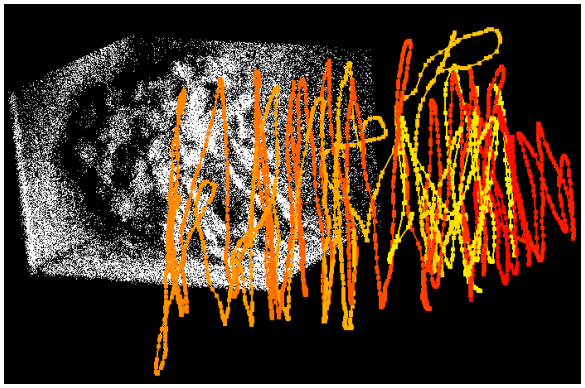
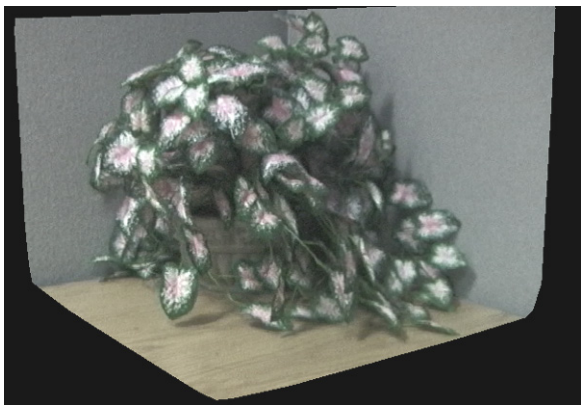


Figure 2: *Typical acquisition path with 3,684 viewpoints*



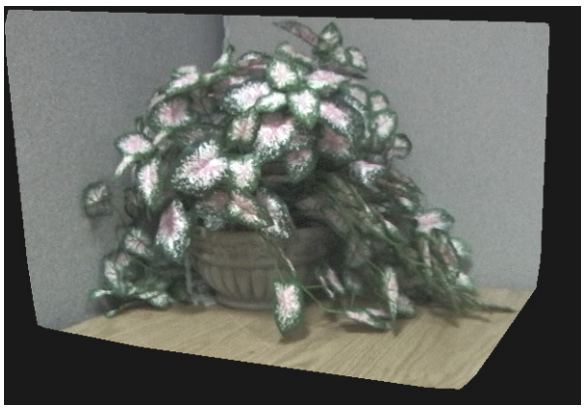


Figure 3*: Images rendered from plant model acquired in 10 minutes.*

Abstract

Three dimensional modeling is a severe bottleneck for computer graphics applications. Manual modeling is time consuming and, even so, the resulting models fail to capture the true complexity of real world scenes. Automated modeling based on acquiring color and depth data is a promising alternative. The conventional approach is to sample the scene densely from a sparse set of acquisition viewpoints. Unfortunately, even with careful view planning, a sparse set of acquisition viewpoints does not and cannot ensure adequate coverage for complex scenes. Moreover, the approach is inefficient due to considerable data redundancy between neighboring acquisition viewpoints: acquisition from an additional viewpoint has the same cost while it contributes fewer and fewer new samples.

We propose an automated modeling approach based on sampling the scene sparsely from a dense set of acquisition viewpoints. We show that the sparse data quickly accumulates to generate models with good scene coverage. The sparse depth is acquired efficiently and robustly, which enables an interactive, operator-in-the-loop acquisition pipeline. We describe a modeling system that implements this approach. The system acquires scenes with complex geometry and complex reflective properties from thousands of viewpoints in minutes. The resulting model has a compact memory footprint and it supports photorealistic rendering at interactive rates. The system is robust, yet it does not require displacing scene objects or altering scene lighting conditions.

Categories and Subject Descriptors (ACM CCS): I.3.3. [Computer Graphics]—Three-Dimensional Graphics and Realism.

1. Introduction

Constructing high-fidelity 3D models of real-world scenes is an important bottleneck for many computer graphics applications. The conventional approach of manual modeling using CAD or animation software requires artistic talent and a huge time investment. Automated modeling based on acquiring color and depth data is a promising alternative.

The typical automated modeling approach is to densely sample the color and geometry of the scene from several acquisition viewpoints, and then to merge the datasets to obtain the scene model. However, dense depth sampling of complex scenes remains a laborious process (e.g. sequential scanning in laser rangefinding, robust off-line correspondence computation in depth from stereo, or sequential light pattern projection in depth from structured light). This limits acquisition to a few viewpoints. Unfortunately, even with optimal acquisition viewpoint planning, complex scenes cannot be adequately sampled from only a sparse set of viewpoints.

Instead of dense depth sampling from a sparse set of acquisition viewpoints we propose an automated modeling approach based on *sparse depth sampling from a dense set of acquisition viewpoints*. We show that in sparse-depth/dense-viewpoint (SDDV) scanning depth data accumulates quickly to generate models with superior scene coverage. Moreover, sparse depth can be acquired robustly and efficiently for each viewpoint, which enables an interactive, operator-in-the-loop acquisition pipeline. The operator detects and addresses acquisition problems right away, and avoids over-scanning parts of the scene with simple geometry, maximizing the impact of the sparse depth. This ensures that a quality model is obtained in a single scanning session.

Our SDDV approach enables adding detail at linear cost: a quick scan can be refined at will by revisiting scanned regions in order to increase model fidelity. This is in sharp contrast with the dense-depth/sparse-viewpoint (DDSV) approach where, due to data redundancy, acquisition from each additional viewpoint comes at the same cost yet contributes fewer and fewer new samples.

Our approach of interactive acquisition from a dense set of viewpoints allows the operator to avoid grazing scanning angles, which leads to increased acquisition robustness for reflective surfaces. Such surfaces are challenging for passive methods because they complicate correspondence searching, and also for active methods because they limit the amount of emitted energy reflected back to the sensor.

We describe a modeling system that implements the SDDV modeling approach (Figure 1). The system acquires the scene from thousands of viewpoints in minutes (Figure 2). The color and depth data is combined into a view dependent model, which produces quality novel views of the scene at interactive rates (Figure 3, Figure 10, and accompanying video). The system is versatile—it supports complex geometry and surfaces with complex reflective properties—and robust—all the models shown here were acquired in a single take in one afternoon. We currently target scenes that fit in a 0.5m cube. However, the system has ample depth acquisition range, the resulting models are compact, and the system does *not* require altering scene lighting conditions or displacing scene objects, which are important pre-conditions for a future extension to SDDV modeling of large environments.

2. Prior work

We give a brief discussion of prior work organized according to the method used for depth acquisition.

No-depth methods

One automated modeling approach is to bypass depth acquisition altogether and to sample the scene color densely from a dense set of acquisition viewpoints. The resulting 4D ray database (light field [LH96], and lumigraph [GGS*96]) supports photorealistic rendering at interactive rates. Light fields and their extensions remain the only approach for acquiring extremely complex scenes (e.g. feathers, fur, translucent gaze [MPN*02]). However, the approach suffers from the disadvantages of large model sizes. Panoramas [Che95] and their extensions are compact 2D ray databases, but the reduction in size comes at the price of restricting the desired viewpoint to the center of the panorama. All ray database methods preclude quantitative applications that require explicit geometry.

User-specified-depth methods

Another approach is to rely on the user to specify a coarse geometric model of the scene using an interactive software tool that leverages model and image space geometric constraints [DTM96, HH02, QTZ*06]. The approach takes advantage of the user's understanding of the scene, who specifies the most relevant geometric entities. However, the method scales poorly with the geometric complexity of the scene, when manual geometry specification becomes prohibitively expensive.

Dense-depth methods

One of the oldest and most appealing ways of acquiring scene geometry is depth from stereo, where the disparity of corresponding points in overlapping photographs is translated into depth [PG02, DRR03, LHY*05,PHT06]. A simple camera suffices to capture scenes on small or large scale, indoors or outdoors. However, finding correspondences robustly and efficiently remains challenging.

In depth from structured light one camera of the stereo pair is replaced with a light source that casts a special light pattern in the field of view of the remaining camera. The approach simplifies the search for correspondences and is robust to surfaces that lack color texture. Using complex patterns of light enables recovering dense depth maps from a few frames [RHL02, KGG03], but the approach relies on projectors which require strict control of scene lighting.

The alternative is to use a simpler and brighter light pattern such as a plane of laser light. Such triangulation laser rangefinding produces high resolution depth maps

[Lev00, GFS05] but at a large time cost due to sequential scanning. Moreover, the laser stripe degrades quickly with the distance to the acquired surface, which requires that the acquisition device be kept close to the scanned surface or to use laser source power levels that are not eye safe. Complex geometry is also a challenge for the laser stripe approach due to laser scattering and secondary reflections.

A third approach for dense depth acquisition is time-of-flight laser rangefinding [MNP99, STY03, GFS05]. Quality depth maps are acquired in minutes [DLS, RGL]. The approach allows scanning from afar. Both laser rangefinding approaches do not acquire color directly.

One disadvantage common to all dense depth methods is the high redundancy between depth maps.

Interactive methods

We are not the first to notice the advantages of an interactive automated modeling pipeline. An early type of interactive modeling system allows the operator to move a light pattern inside the field of view of the fixed camera [Bor*98, BP99]. This enables the operator to avoid grazing angles between the scene surfaces and the pattern of light, but is fundamentally limited to a single viewpoint by the fixed camera. The most the operator can do is to assign depth to all the pixels of the camera image, which is equivalent to one dense depth map.

The single-viewpoint limitation is overcome by systems that employ a laser stripe triangulation rangefinder whose position and orientation is provided by an electromagnetic [FFG*96] or mechanical [Hi00] tracker. However, the laser stripe limitations discussed above remain. A different approach for achieving interactive acquisition from many viewpoints is to move the targeted scene in the field of view of a fixed scanner using a rotating platter [TOF05], or by relying on the operator to rotate the object [RHL02]. Such systems are essentially object scanners and scale poorly with the size of the scene.

The ModelCamera system [PSB03, BPM*05, BPM06] employs a structured light pattern that consists of a matrix of laser beams, which ensures robust acquisition of sparse depth data at interactive rates. Our acquisition device is similar to the ModelCamera and we describe the similarities and differences in Section 4.1. The ModelCamera scans simple scenes (e.g. large pieces of furniture) in a hand-held mode, without reliance on an external tracker [PSB03]. Complex scenes (e.g. plants, rooms, buildings) are acquired by panning and tilting the ModelCamera about a fixed acquisition viewpoint [BPM*05, BPM06]. The single viewpoint constraint limits the fidelity of the acquired models.

3. System overview

We developed a system that executes the sparse-depth/dense-viewpoint automated modeling approach. We employ a hand-held acquisition device that acquires dense color and sparse depth at interactive rates (Section 4.1). The operator sweeps the scene with the device and monitors the acquisition process through immediate visual feedback provided on a nearby monitor (Figure 1). Data acquisition proceeds in two steps: depth acquisition (Section 4.2), which achieves adequate scene coverage, followed by color acquisition (Section 4.3). The depth and color data is combined into a view-dependent model for rendering (Section 5).

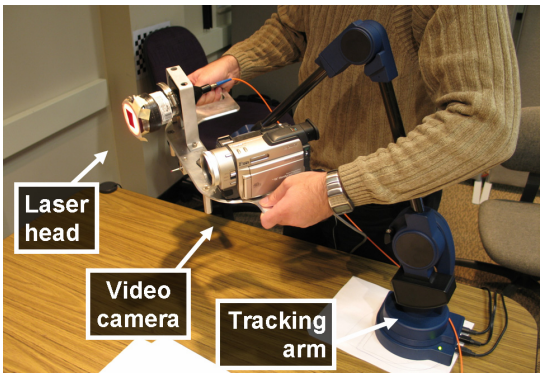


Figure 4: *Acquisition device.*

4. Acquisition

To support the SDDV approach the system should:

- acquire sparse depth efficiently and robustly
- acquire high-quality color
- allow the operator to freely position the acquisition device
- provide real-time feedback during acquisition.

4.1. Acquisition device

We employ an acquisition device similar to the ModelCamera [BPM05*, BPM06]. Like the ModelCamera, our acquisition device consists of a video camera with an attached laser system that casts a matrix of laser beams into the field of view of the camera (Figure 4). Each beam creates a bright dot in the frame which is detected and triangulated into a depth sample. By construction, the laser beams project to disjoint image plane epipolar segments which makes dot detection efficient and robust. Compared to a laser stripe, the dot pattern has the advantage of higher brightness and of fewer scattering and secondary reflection artifacts. The baseline of 20 cm provides a depth range of 50 to 400 cm and sufficient accuracy (a depth error less than 3mm at 100 cm range).

Whereas the ModelCamera was restricted to a single acquisition viewpoint using a parallax-free pan-tilt bracket, our acquisition device is attached to a mechanical tracking arm which allows 6-degree-of-freedom motion. The operator can position and orient the acquisition device freely within a sphere with a radius of 60 cm to acquire high quality color (i.e. progressive scan, 3 CCDs, 720x480 pixels) and 7x7 depth samples robustly at 15 fps.

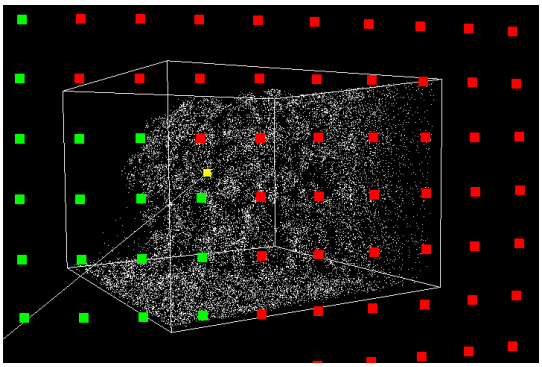


Figure 5: *Visual feedback during color acquisition.*

4.2. Depth acquisition

In a first step depth is acquired by continuously moving the acquisition device around the targeted scene. Dot detection parameters (e.g. minimum intensity, maximum size, and maximum asymmetry) are set conservatively in order to avoid false positives. Dot detection also fails when a dot is simply not visible from the camera viewpoint due to complex scene geometry with numerous depth discontinuities. Since the acquisition device can be freely moved in 6 degrees of freedom, the operator can roll the device to scan with a vertical baseline which removes the bias against vertical edges.

A typical acquisition path is shown in Figure 2. The point cloud comprises 78,345 3D points and was acquired along the path shown from 3,684 viewpoints (frames) in about 4 minutes. This yields approximately 21 dots per frame, or a dot detection rate of 43% on this challenging leafy plant model. The operator inspects the point cloud in real time and increases depth sampling resolution by revisiting poorly sampled areas. In Figure 2, the color used to visualize the acquisition viewpoints changes gradually from red to yellow, which shows that the operator scanned the plant once, producing a half-resolution model in half the time, and then backtracked to double the resolution.

4.3. Color acquisition

Since we primarily target computer graphics applications, the acquisition of quality color data is an important goal of our system. The video camera incorporated in our acquisition device acquires the equivalent of high quality 720x480 digital stills. Although our acquisition device acquires depth and color simultaneously, we acquire color in a second stage, with the laser turned off, to avoid the red dots in the reference color frames. In order to decide about the required density for color sampling locations, one has to consider the goals of color acquisition.

One goal is to provide color for diffuse surfaces. This goal is achieved if all surfaces are visible in at least one frame that does not sample them too obliquely. Simple diffuse surfaces are quickly acquired with a few frames. Complex geometry requires a denser set of frames to ensure visibility. Another goal is to provide color for non-diffuse surfaces. Color sampling location density has to increase with the specularity of the surfaces.

A third goal is to compensate for geometry undersampling. View dependent texture mapping [DYB98] and light field research has shown that increasing color sampling location density can reduce or even eliminate the need for geometry. The exact trade off between the amount of geometry and amount of color data depends on the targeted scene and on the application.

Our system allows the operator to easily control the density of color acquisition viewpoints using a sphere circumscribing the bounding box of the acquired depth samples. The operator specifies the desired angular spacing of the color acquisition viewpoints and then simply matches the current position of the acquisition device with the nodes of the regular grid of desired viewpoints, relying on visual feedback.

In Figure 5 the red dots are viewpoints from which color is yet to be acquired, green dots show viewpoints that already have data, and the yellow dot is the current position of the acquisition device. When the yellow dot is sufficiently close to a red dot, a frame is automatically captured. Note that the yellow dot is completely determined by the center of projection of the video camera, and does not dependent on the camera orientation. The operator adjusts only two degrees of freedom (the location of the acquisition device on the sphere), which enables efficient color acquisition. In a typical acquisition session, color is acquired in 5 minutes from ~100 locations covering a 50^{o}x30^{o} sphere sector with a 4^{o} angular resolution.

5. Rendering

The acquired point and color data could, in principle, be rendered with a variety of existing algorithms such as splatting (e.g. q-splats [RL00], or surfels [PZB*00]) and 3D triangulation [Hop*92, TL94, SH04]. These reconstruction approaches would produce good results for simple scenes where the geometry is adequately sampled. However, for complex scenes such as the leafy plant, 3D triangulation fails, and splatting produces low quality images since one cannot accurately estimate the desired image footprint of the projected samples.

We have developed a reconstruction algorithm that performs well even in such challenging cases. The algorithm uses the depth samples to automatically define a morph between the reference color frames. This provides the flexibility of compensating for low geometric resolution by increasing the resolution of color acquisition viewpoints.

A second defining characteristic of our reconstruction algorithm is that it is view dependent: a reconstruction is computed for every desired view. In light of the increased sophistication and raw processing power of GPUs, view-dependent modeling has become an appealing alternative to conventional, view-independent geometric modeling for graphics applications. Modeling is considerably simpler if the modeling algorithm caters to only one specific view.

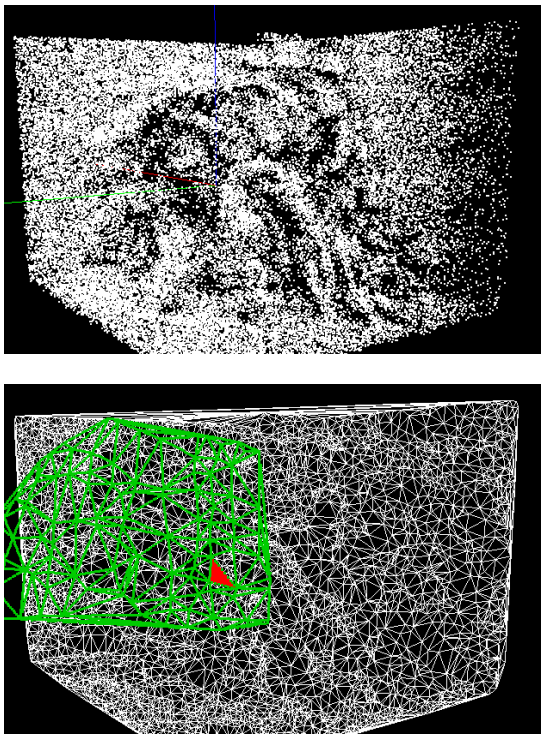


Figure 6: *Visualization of the stages of the run-time rendering algorithm.*

The run time algorithm performs the following steps for each frame:

A. *Visibility*. The points visible in the frame are computed by splatting all points and by collecting those whose splats remain present through at least one pixel. We use square splats with an image-space size derived from a model-space size which is estimated during preprocessing.

B. *2D triangulation*. The projections of the visible points are Delaunay 2D triangulated in the frame and the connectivity information so inferred is applied to the visible 3D points to derive a view-dependent 3D triangle mesh.

C. *3D triangle mesh coloring*. The 3D mesh provides depth for each frame pixel. A GPU shader projects each fragment into 3 nearby reference color frames. Visibility in a color frame is assessed using a z-buffer pre-computed according to steps A and B. The final fragment is computed by blending the colors looked up in the reference frames. In order to find the 3 neighboring color frames efficiently, the color acquisition viewpoints are 2D triangulated onto the color acquisition sphere as a preprocess. The 3 color frames are determined by the triangle that contains the projection of the desired viewpoint onto the sphere.

Figure 6 shows the entire set of 3D points (top), the points visible in the current frame (middle), and the bottom image shows the view dependent triangle mesh in white, the mesh of color reference viewpoints in green, and the triangle used for coloring the current frame, in red.

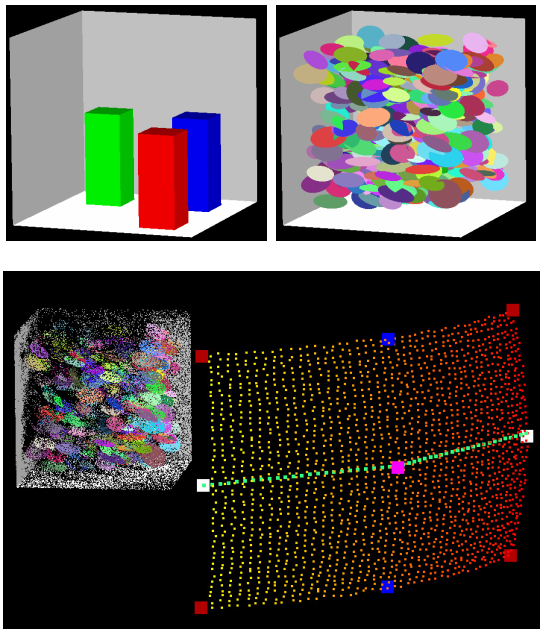


Figure 7: *Scenes (top) and sampling locations (bottom) used for scene coverage analysis.*

6. Results

This section first provides a comparison between our SDDV method and conventional DDSV on synthetic scenes, and then provides results of scanning real-world scenes with our system.

6.1. SDDV vs. DDSV

We compared the scene coverage performance of SDDV to that of DDSV quantitatively on two synthetic scenes. The geometry of the synthetic scenes is known which provides ground truth for the simulations. One scene is simple, consisting of 3 blocks, and one has complex geometry, consisting of a set of 400 disks with various radii and orientations (Figure 7).

The SDDV approach was simulated using a virtual acquisition device that gathers 7x7 depth samples per frame. The acquisition path samples a 50^{o}x30^{o} solid angle and has 1,979 viewpoints, which are typical values for our system. The bottom image in Figure 7 shows the acquired point cloud and the acquisition path with small dots whose color changes from red to yellow.

The DDSV approach was simulated in 5 scenarios with various numbers of acquisition locations: 1 (large pink dot in bottom image of Figure 7), 2 (blue dots), 2 (white dots), 4 (red dots), and 6 (red and blue dots). Scene coverage was measured along a horizontal rendering path shown with green dots.

Although it is possible to measure scene coverage in an absolute, view independent way, such a figure is less

Scene		Sampling Approach					
		SDDV	Dense depth				
			1L	2VL	2HL	4L	6L
Disks	avg.	0.83	11	11	3.6	4.3	3.1
	max	1.6	23	15	5.6	5.4	4.5
Blocks	avg.	0.11	6.2	3.9	0.38	0.24	0.15
	max	0.34	13	7.7	1.9	1.3	0.80

Table 1: *Percentage of missing samples.*

relevant since it includes surfaces that are never visible in any rendering path of practical interest (i.e. deeply hidden disks, or back faces of blocks). Instead, we investigated scene coverage by measuring the fraction of missing samples for each frame along the rendering path. For each frame, the scene is rendered from geometry to obtain ground truth. For SDDV, a ground truth image sample is missing if it is not present in a reconstruction of the scene built from the acquired samples. Since all surfaces in the scene are planar (i.e. disks, walls, block faces), reconstruction is computed accurately by 2D triangulation on each planar surface. For DDSV, a sample is missing if it is not present in any of the dense depth maps. Leveraging the fact that all scene surfaces are planar, this can be computed accurately, without using thresholds.

The results are given in Table 1. The dense depth simulations are labeled according to the number of acquisition locations. The simulations employing 2 vertical/horizontal locations are labeled 2VL/2HL. The 2HL simulation is favored by the rendering path chosen which is horizontal. The SDDV approach substantially outperforms the 1L simulation in all cases. For the disks scene, SDDV achieves superior scene coverage even compared to 6L (also see Figure 8). For the blocks scene, SDDV has an average performance comparable to 6L. Notice that 2VL is substantially worse than 2HL, which indicates a strong dependence on the rendering path for DDSV. In other words, even for a simple scene such as the blocks scene, two DDSV locations are not sufficient for general rendering paths.

Our 15fps acquisition device would capture the 1,979 simulation frames in under 2 ½ minutes, or 5 minutes when factoring in the actual dot detection rate. Even with the latest laser rangefinders, scanning a scene from 6 locations in 5 minutes remains challenging, since, in addition to the actual acquisition time one also needs to factor in the time

Scene		Sampling Approach					
		SDDV	Dense depth				
			1L	2VL	2HL	4L	6L
Disks	avg.	0.52	0.06	0.88	0.54	1.86	3.30
	max	0.57	0.2	1.09	0.60	2.09	3.64
Blocks	avg.	0.56	0.11	0.94	0.75	2.26	4.02
	max	0.65	0.30	1.23	0.90	2.69	4.63

Table 2: *Number of redundant samples per pixel.*

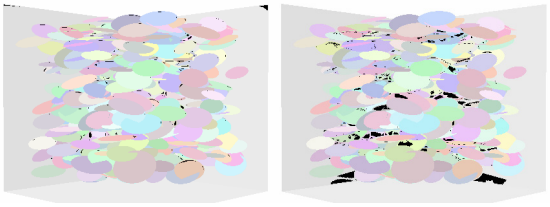


Figure 8: *Visualization of missing samples for SDDV (left, 0.57%) and DDSV 6L (right, 2.70%).*

needed to post-process the data.

We have also investigated sampling redundancy along our rendering path. Sample redundancy for a frame is computed as the average pixel sample redundancy. The pixel sample redundancy is the number of visible samples that project to the pixel, minus one. Since our surfaces are planar, there is no possibility of surface self occlusion. Therefore, visibility can be determined robustly by testing whether a sample belongs to the same surface as its ground truth image pixel. The rendered image and DDSV depth maps have the same resolution (720x480).

The average and maximum redundancy values over the path frames are given in Table 2. Note that redundancy is not 0 for the 1L case due to the sampling rate perturbation induced by scene geometry. The 6L case implies massive redundancy. For DDSV, redundancy is larger for the blocks scene, whereas redundancy changes little with the scene in the SDDV approach.

Moreover, in these SDDV acquisition simulations, the virtual device blindly sampled the scene along a preset path. In a real scanning scenario the SDDV approach benefits from operator guidance, which further improves scene coverage and further reduces redundancy.

6.2. Scanning results

We tested our system with good results on a variety of scenes with simple and complex geometry, with simple and complex reflective properties (Figure 3, Figure 10, and accompanying video). The system is efficient and robust: each model was acquired in a single take, in approximately 10 minutes. The models comprise fewer than 66,000 points and fewer than 154 reference color frames (Table 3), which is substantially more compact than a light field that would achieve a similar rendering quality.

Our rendering algorithm takes advantage of the depth samples which provide a smooth morph between the reference color frames. Figure 9 compares light field reconstructions that rely solely on the color data to the reconstruction achieved with our algorithm. Clearly there are not enough reference color frames for a quality light field reconstruction. Even if the light field is focused at the centroid of the scene, considerable blurriness persists. The 3D points in the model guide the transition between neighboring color frames, in a sense achieving a high-quality compression of the scene light field.

	Books	Box	Clothes	Flower2	Carafe	Pillow	Toys	Vase
Points [x 1,000]	51	32	56	50	28	16	47	66
Color frames	127	99	80	118	109	123	154	122
Acq. time [min]	10	7	10	10	8	7	11	12

Table 3*:* Statistics for models shown in Figure 10.

The rendering algorithm runs at 15fps for all the models. Rendering performance was measured on a Pentium 4 Xeon 3.4GHz 2GB PC with an NVIDIA Quadro FX1400 graphics card. Most of the frame rendering time goes to the Delaunay triangulation of the visible points, which is executed on the CPU. The preprocessing time is dominated by the model-space splat size evaluation, which takes about 3 minutes.

7. Discussion

We have demonstrated the importance of a large number of acquisition viewpoints. We conducted simulations that show that the sparse-depth/dense-viewpoint approach is fast, has ample modeling power, low data redundancy, and good scene coverage. We implemented the SDDV approach in an efficient and effective automated modeling system that handles a variety of scenes robustly. We developed a view-dependent modeling algorithm that computes per-frame triangle meshes for a quality reconstruction at interactive rates.

Our system has the following limitations. Although it acquires glossy surfaces (e.g. shoes or vase in Figure 10), it cannot detect dots on very specular surfaces such as perfect mirrors. The giraffe in the toys scene also proved to be challenging for our system. Because of the high contrast between the dark brown and the white regions, the video camera adjusts the white balance assigning very high intensities to the white region pixels (>250). Therefore the laser dots do not stand out sufficiently as required by the conservative settings used in dot detection.

A second limitation comes from an occasional loss in tracking precision during color acquisition, which translates in occasional frame instability during rendering. This issue appears to be caused by certain positions of the arm when tension in the joints causes less than accurate pose readings.

A third limitation of the current system comes from the fact that the reference color frames are used as acquired, without intensity balancing. For bright scenes this causes occasional intensity variation during rendering.

We will continue to extend our modeling system and we will address these issues. Another short term goal is to enable simultaneous acquisition of color and depth, which will halve the typical acquisition time. This requires replacing the laser dot regions in a frame with color from nearby frames.

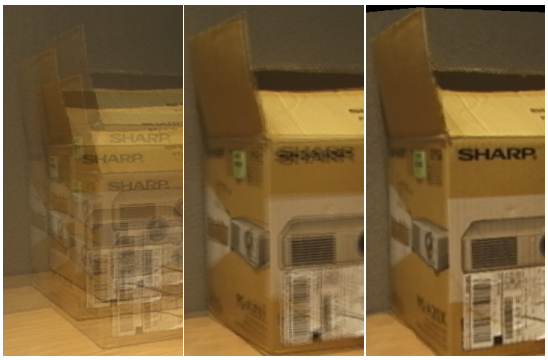


Figure 9: *Images rendered from light field focused at infinity, from light field focused at the scene centroid, and from acquired model.*

A longer term goal is to extend the SDDV approach to large indoor scenes. Many of the requirements are already met. The depth acquisition has sufficient depth range and operates robustly under unaltered scene lighting conditions. The tracking arm can reach virtually anywhere within a sphere with a 60cm radius. This should provide sufficient acquisition viewpoint parallax for good coverage of room-sized environments without the need to move the arm base. SDDV promises to contribute substantially towards eliminating the 3D computer graphics modeling bottleneck.

8. Acknowledgments

<Withheld for double-blind review.>

9. References

[BPM*05] Bahmutov G., Popescu V., Mudure M., Sacks E.: Depth Enhanced Panoramas. *Proceedings of Video Vision and Graphics (VVG)*, 2005

[BPM06] BAHMUTOV G., POPESCU V., AND MUDURE M. Efficient Acquisition of Large-Scale Building Interiors In *Proceedings of EUROGRAPHICS* 2006.

[Bor*98] BORGHESE, N.A., et al. 1998. Autoscan: A Flexible and Portable 3D Scanner. *IEEE Computer Graphics and Applications,* Vol.18, No.3, 1998, pp. 38-41.

[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] CHEN E., Quicktime VR - An Image-Based Approach to Virtual Environment Navigation, *Proc. SIGGRAPH 95*, 29-38 (1995).

[DLS] http://www.deltasphere.com/

[DRR03] DAVIS, J., RAMAMOORTHI, R., AND RUSINKIEWICZ, S. Spacetime stereo: a unifying framework for depth from triangulation. In Proceedings of *Conference on Computer Vision and Pattern Recognition*, (2003).

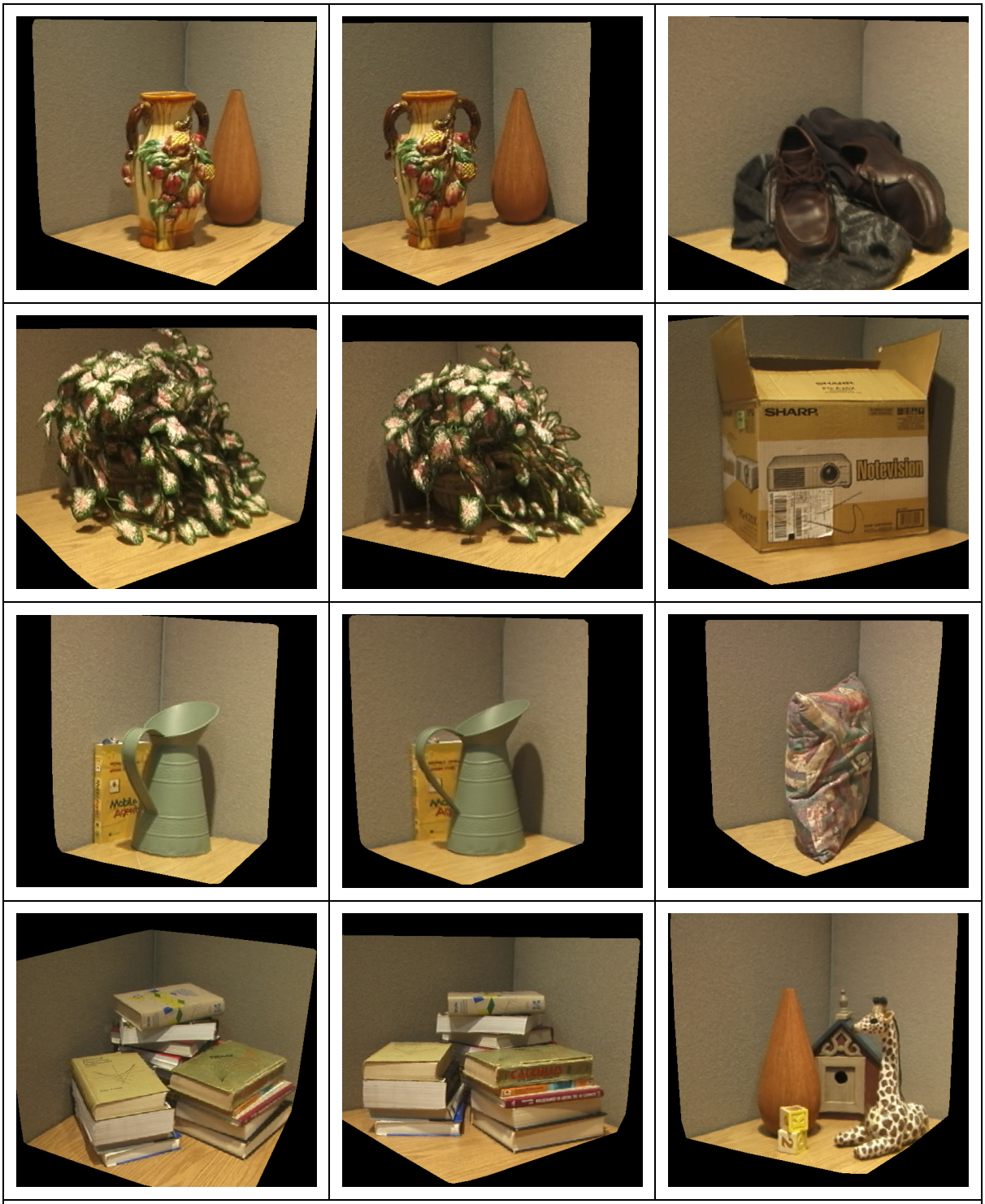


Figure 10: *Images rendered from the vase, clothes, flower2, box, carafe, pillow, books, and toys models acquired with our system.*

[DTM96] DEBEVEC P., TAYLOR, C., AND MALIK J. Modeling and Rendering Architecture from Photographs: A Hybrid Geometry and Image Based Approach. Proc. *SIGGRAPH* '96, 11-20 (1996).

[FFG*96] FISHER, R., FITZGIBBON, A., GIONIS, A., WRIGHT, M., AND EGGERT, D. "A Hand-Held Optical Surface Scanner for Environment Modeling and Virtual Reality" In Proc. *Virtual Reality World*, pages 13-15 1996.

[GFS05] GUIDI G.,FRISCHER B., SPINETTI A.: 3D digitization of a large model of imperial Rome. Proceedings of the *Fifth International Conference on 3-D Digital Imaging and Modeling* (3DIM'05).

[GGS*96] GORTLER S., GRZESZCZUK R., SZELISKI R., COHEN M.: The Lumigraph. *In SIGGRAPH '96*.

[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.

[HI00] HILTON, A. AND ILLINGWORTH. "Geometric Fusion for a Hand-Held 3D Sensor" 12(1)::44-51m, *Machine Vision Applications*, 2000.

[Hop*92] HOPPE H., et al. Surface reconstruction from unorganized points, *Computer Graphics*, (26:2), (1992).

[KGG03] T. P. KONINCKX, A. GRIESSER, AND L. VAN GOOL, Real-Time Range Scanning of Deformable Surfaces by Adaptively Coded Structured Light. Proceedings of *Fourth International Conference on 3D Digital Imaging and Modeling 2003*, pp. 293-301.

[Lev*00] M. LEVOY ET AL. The Digital Michelangelo Project: 3D Scanning of Large Statues, Proc. *ACM SIGGRAPH*, 2000.

[LHY*05] LIM J.,HO J.,YANG M.,KRIEGMAN D. : Passive Photometric Stereo from Motion . Proceedings of the Proceedings of the Tenth *IEEE International Conference on Computer Vision* (ICCV'05).

[LH96] LEVOY M., HANRAHAN P.: Light Field Rendering. *Proc. of SIGGRAPH 96* (1996), 31-42. UUU

[MPN*02] MATUSIK W., PFISTER H., NGAN A., BEARDSLEY P., ZIEGLER R., MCMILLAN L. Image-Based 3D Photography using Opacity Hulls. *In SIGGRAPH '02*.

[MNP*99] D. MCALLISTER, L. NYLAND, V. POPESCU, A. LASTRA, C. MCCUE. Real-Time Rendering of Real-World Environments. Proceedings of the *Eurographics Workshop on Rendering*, June 21-23, 1999.

[PG02] M. POLLEFEYS AND L. VAN GOOL. From Images to 3D Models, *Communications of the ACM*, July 2002/Vol. 45, No. 7, pp.50-55.

[PHT06] http://labs.live.com/photosynth/

[PZB*00] H. PFISTER, M. ZWICKER, J. VAN BAAR, AND M. GROSS. Surfels: Surface Elements as Rendering Primitives. Proc. of *SIGGRAPH '98*, 335-342 (2000).

[PSB03] POPESCU, V., SACKS, E., AND BAHMUTOV, G. 2003. The ModelCamera: A Hand-Held Device for Interactive Modeling. In *Proceedings of Fourth International Conference on Digital Imaging and Modeling*, Banff, Canada, 2003.

[QTZ*06] LONG O.,PING T.,GANG Z., YUAN L., WANG J., KANG S.: Image-based Plant Modeling. Proceedings of ACM Siggraph 2005, *ACM Transactions on Graphics (TOG)* Volume 25 , Issue 3 (July 2005) p. 599-604

[RGL] http://www.riegl.com/

[RHL02] S. RUSINKIEWICZ, O. HALL-HOLT, AND M. LEVOY. Real-Time 3D Model Acquisition. Proc. *SIGGRAPH* 2002.

[RL00] S. RUSINKIEWICZ, M. LEVOY. QSplat: A Multiresolution Point Rendering System for Large Meshes. Proc. *SIGGRAPH* 2000.

[SH04] SOLEM, H.E. AND HEYDEN, A. Reconstructing open surfaces from unorganized data points. In Proc. *Conference on Computer Vision and Pattern Recognition*, 2004.

[STY03] STUMPFEL J., TCHOU C., YUN N., MARTINEZ P., HAWKINS T., JONES A., EMERSON B., DEBEVEC P.. Digital Reunification of the Parthenon and its Sculptures, *4th International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage*, Brighton, UK, 2003.

[TOF05] TERAUCHI T., OUE Y., FUJIMURA K. : A Flexible 3D Modeling System Based on Combining Shape-from-Silhouette with Light-Sectioning Algorithm. Proceedings of the *Fifth International Conference on 3-D Digital Imaging and Modeling* (3DIM'05).

[TL94] TURK G. AND LEVOY M., Zippered Polygon Meshes from Range Images, *SIGGRAPH* 94, 311-318, (1994).