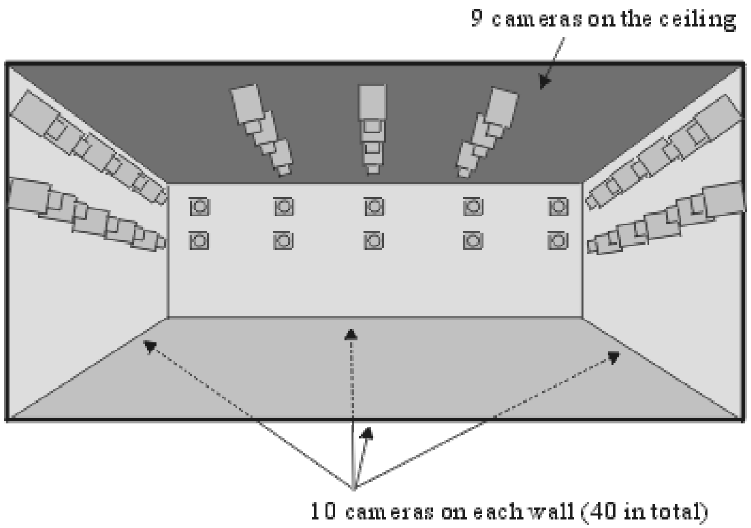
Multiview stereo

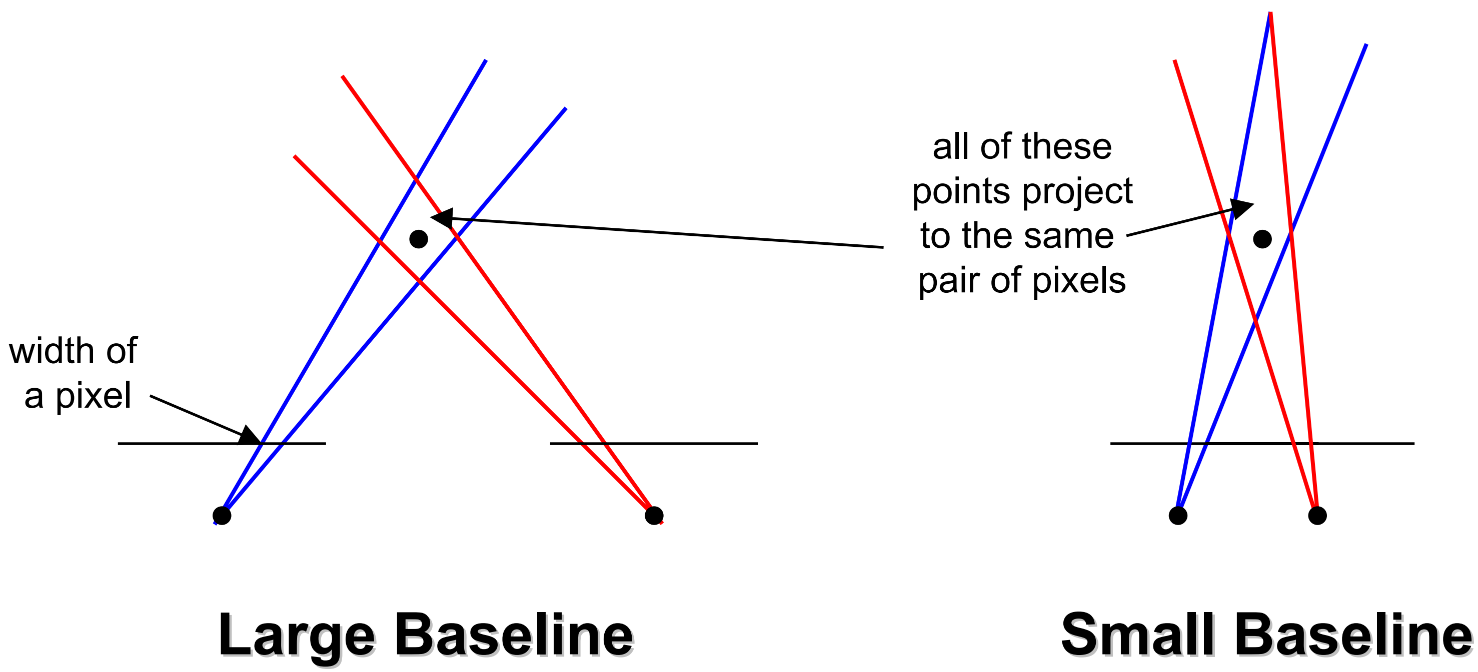


CMU's 3D Room

Readings **(Optional)**

- S. M. Seitz and C. R. Dyer, Photorealistic Scene Reconstruction by Voxel Coloring, *International Journal of Computer Vision*, 35(2), 1999, pp. 151-173.

Choosing the Baseline

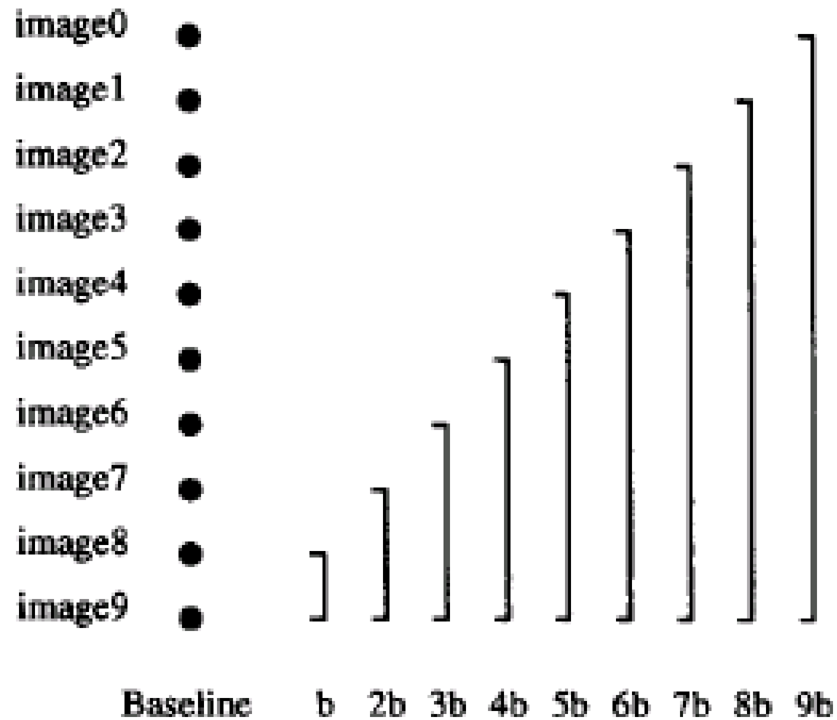


What's the optimal baseline?

- Too small: large depth error
- Too large: difficult search problem

The Effect of Baseline on Depth Estimation

Figure 2: An example scene. The grid pattern in the background has ambiguity of matching.



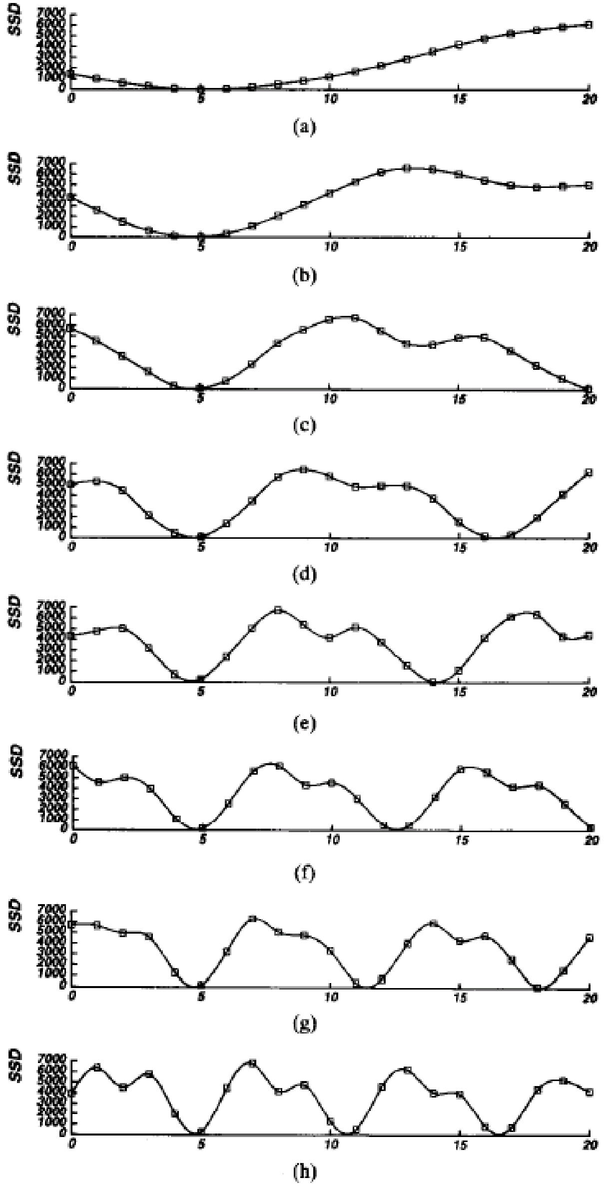


Fig. 5. SSD values versus inverse distance: (a) $B = b$; (b) $B = 2b$; (c) $B = 3b$; (d) $B = 4b$; (e) $B = 5b$; (f) $B = 6b$; (g) $B = 7b$; (h) $B = 8b$. The horizontal axis is normalized such that $8bF = 1$.

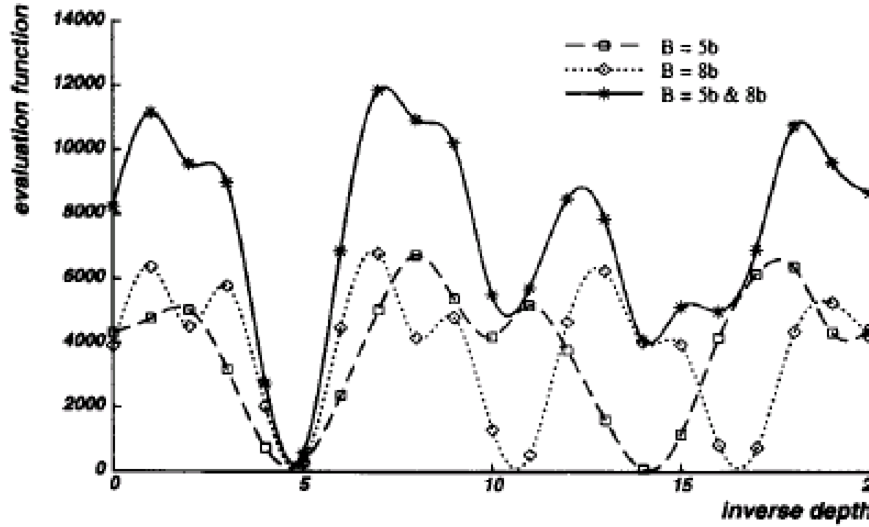


Fig. 6. Combining two stereo pairs with different baselines.

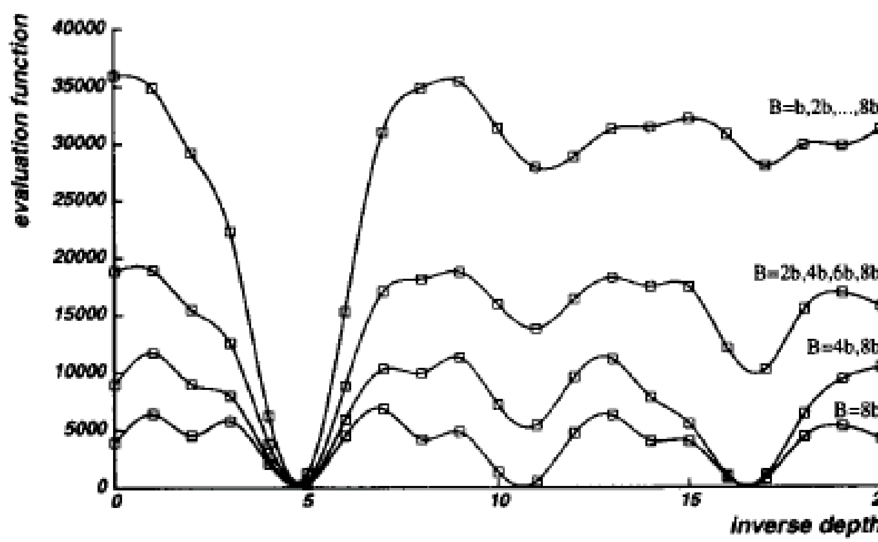


Fig. 7. Combining multiple baseline stereo pairs.

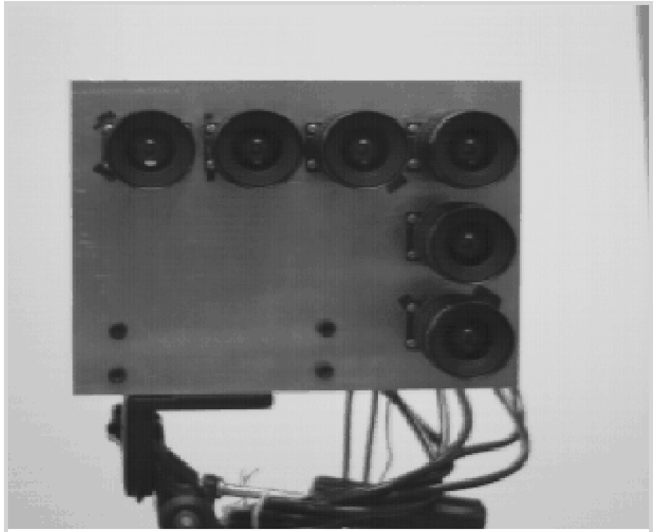
Multibaseline Stereo

Basic Approach

- Choose a reference view
- Use your favorite stereo algorithm BUT
  - > replace two-view SSD with SSD over all baselines

Limitations

- Must choose a reference view (bad)
- Visibility!



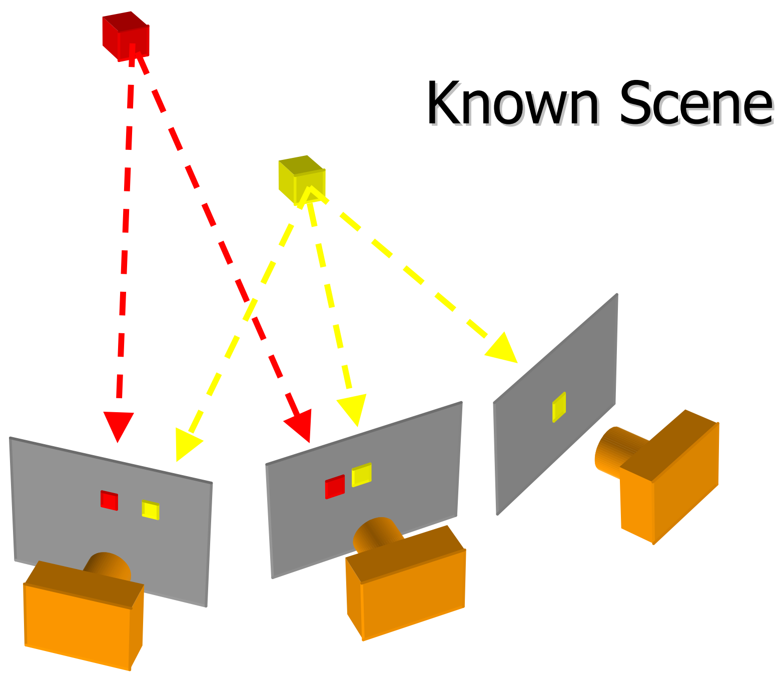
(a)

(b)

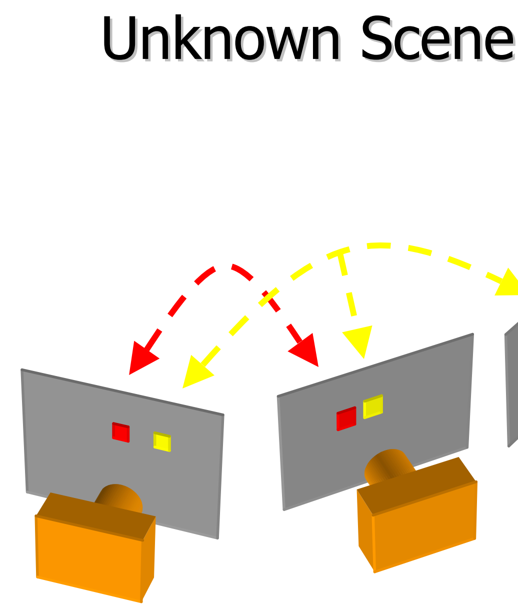
Figure 7: The CMU Video-Rate Stereo Machine Prototype System: (a) camera head; (b) processor boards

The global visibility problem

Which points are visible in which images?



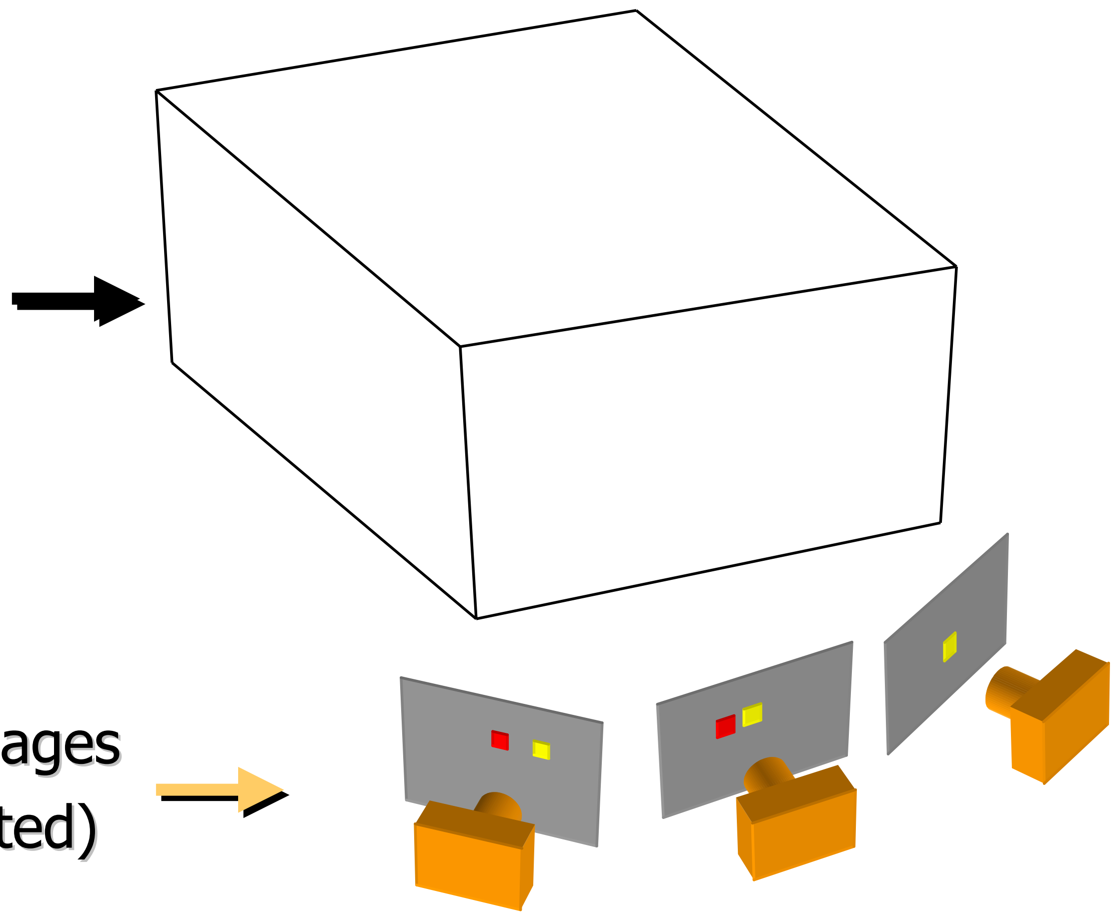
Forward Visibility



Inverse Visibility

Volumetric stereo

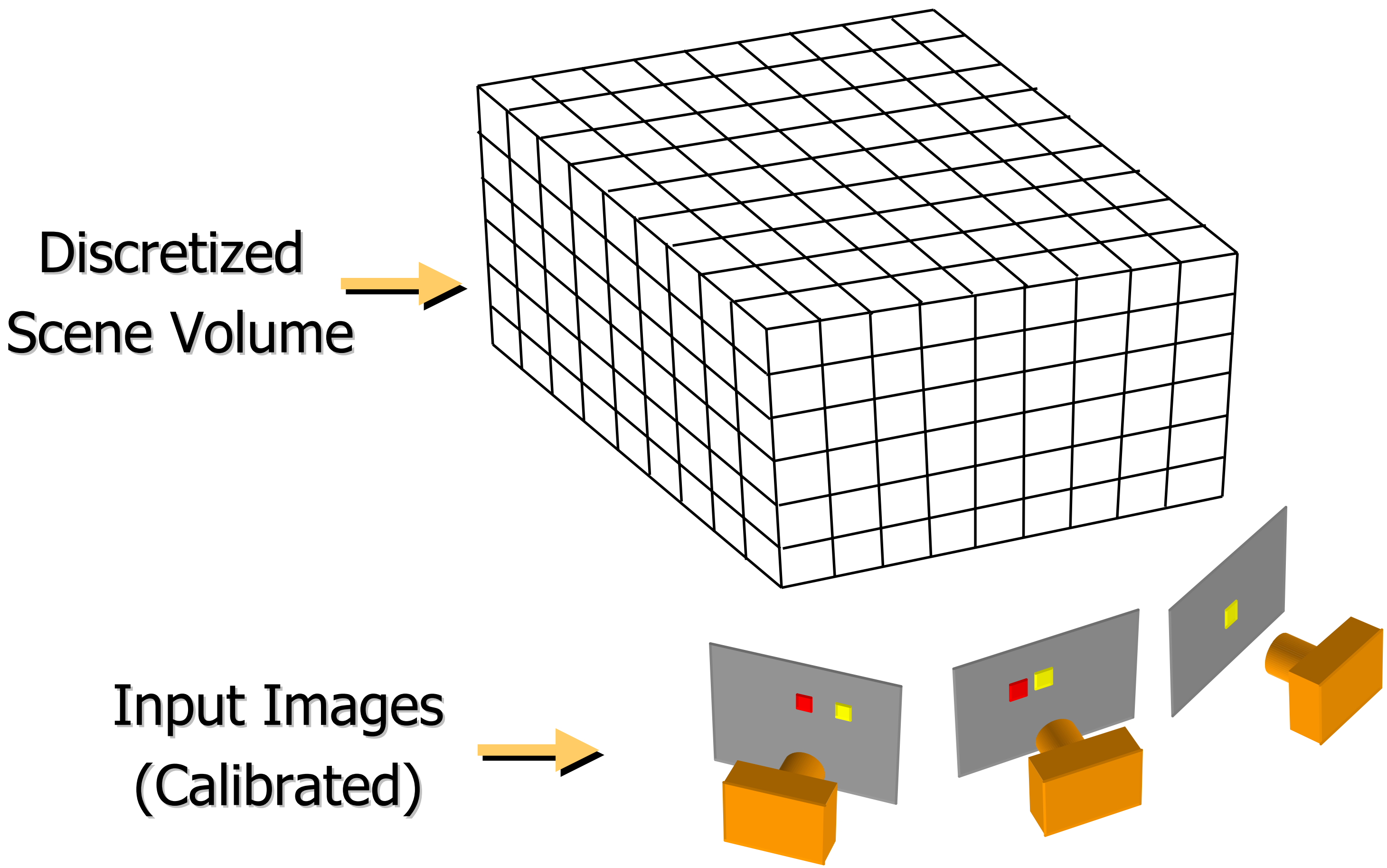
Scene Volume
V



Input Images
(Calibrated)

Goal: Determine occupancy, "color" of points in V

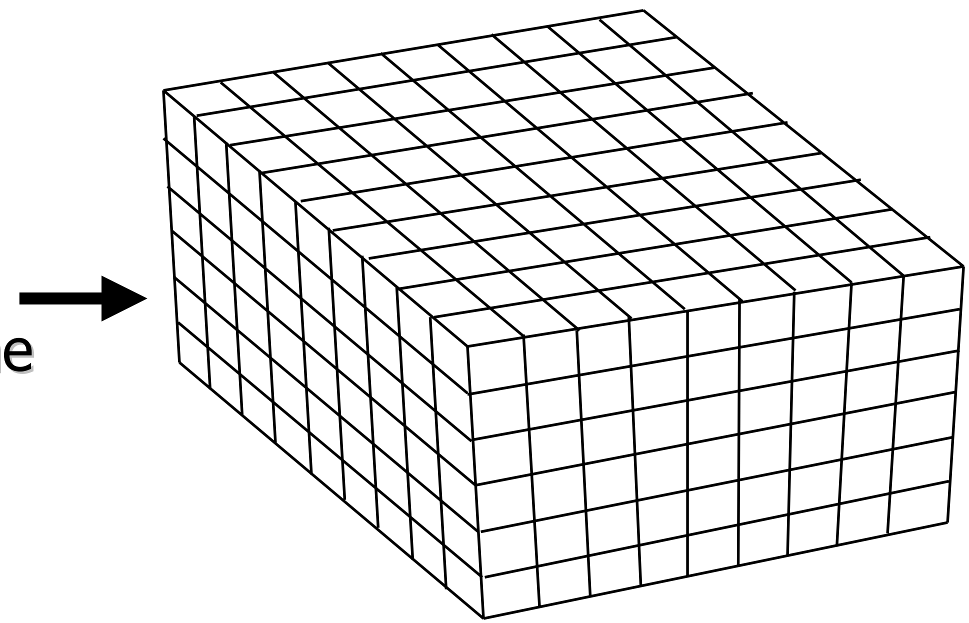
Discrete formulation: Voxel Coloring

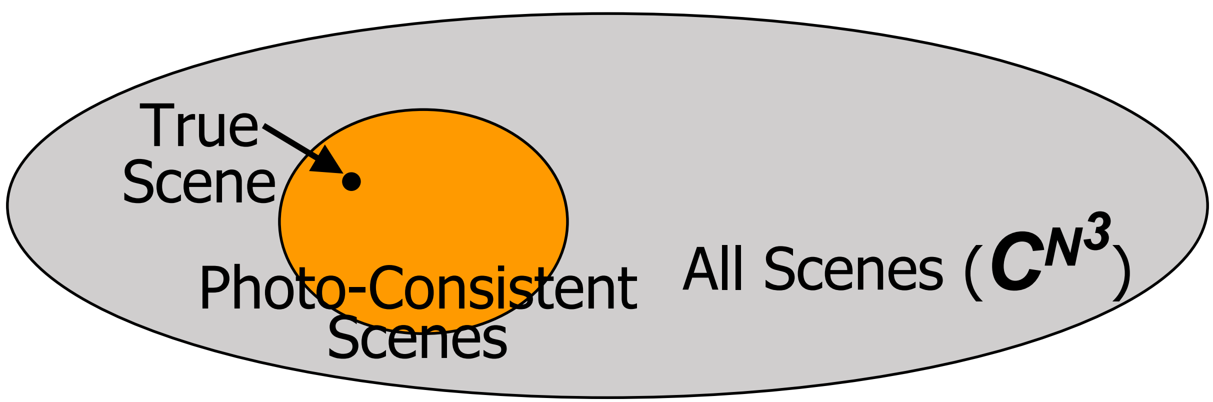


Goal: **Assign RGBA values to voxels in V**
photo-consistent **with images**

Complexity and computability

Discretized
Scene Volume
N^3 voxels
C colors





Issues

Theoretical Questions

- Identify class of *all* photo-consistent scenes

Practical Questions

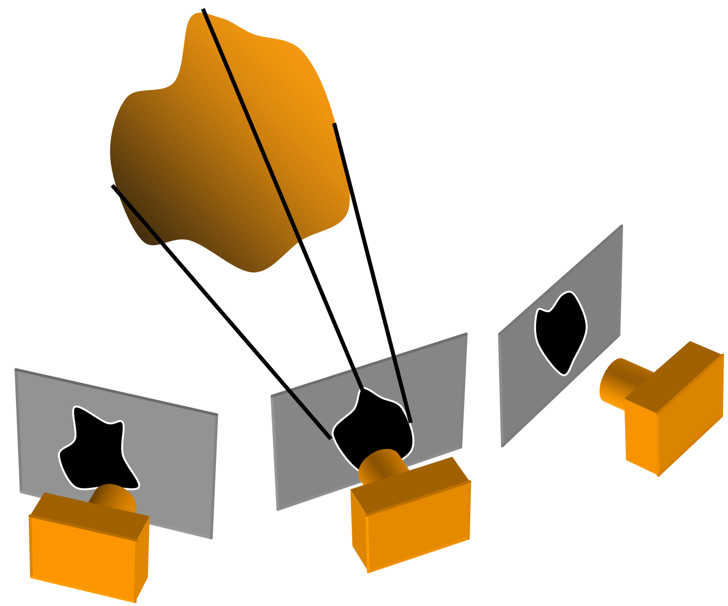
- How do we compute photo-consistent models?

Voxel coloring solutions

1. C=2 (shape from silhouettes)
   - Volume intersection [Baumgart 1974]
     - > For more info: *Rapid octree construction from image sequences.* R. Szeliski, CVGIP: Image Understanding, 58(1):23-32, July 1993. (this paper is apparently not available online)
2. C unconstrained, viewpoint constraints
   - Voxel coloring algorithm [Seitz & Dyer 97]
3. General Case
   - Space carving [Kutulakos & Seitz 98]

Reconstruction from Silhouettes (C = 2)

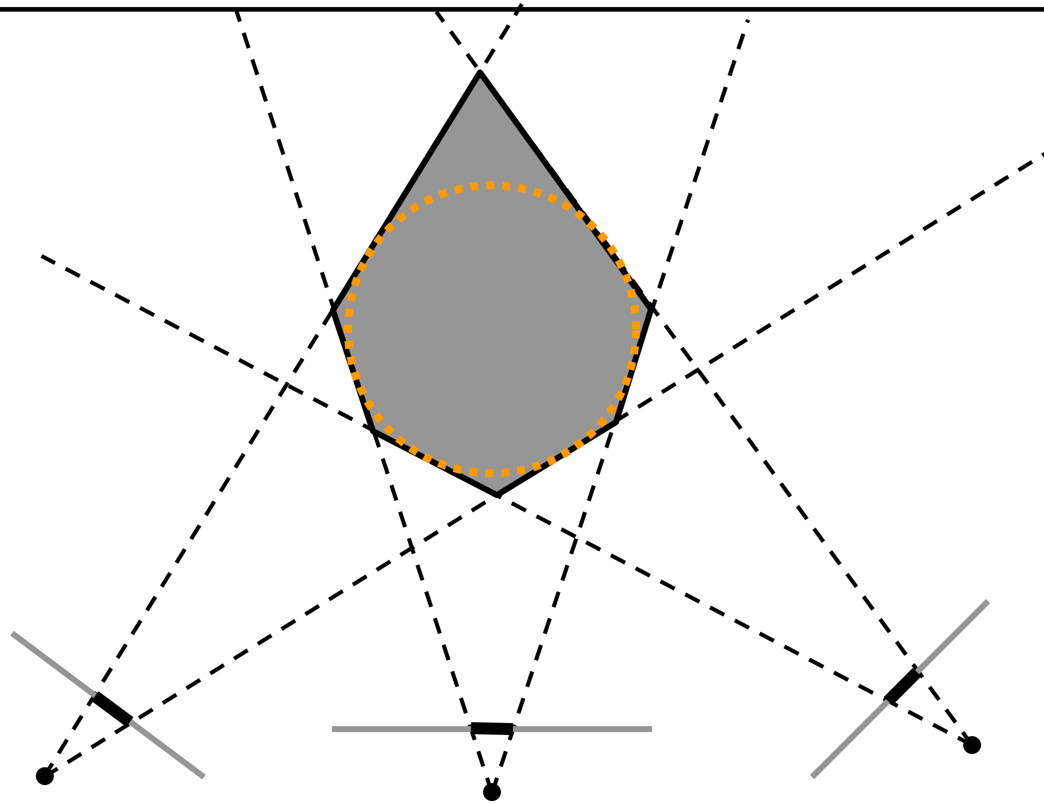
Binary Images



Approach:

- *Project* each silhouette
- Intersect projected volumes

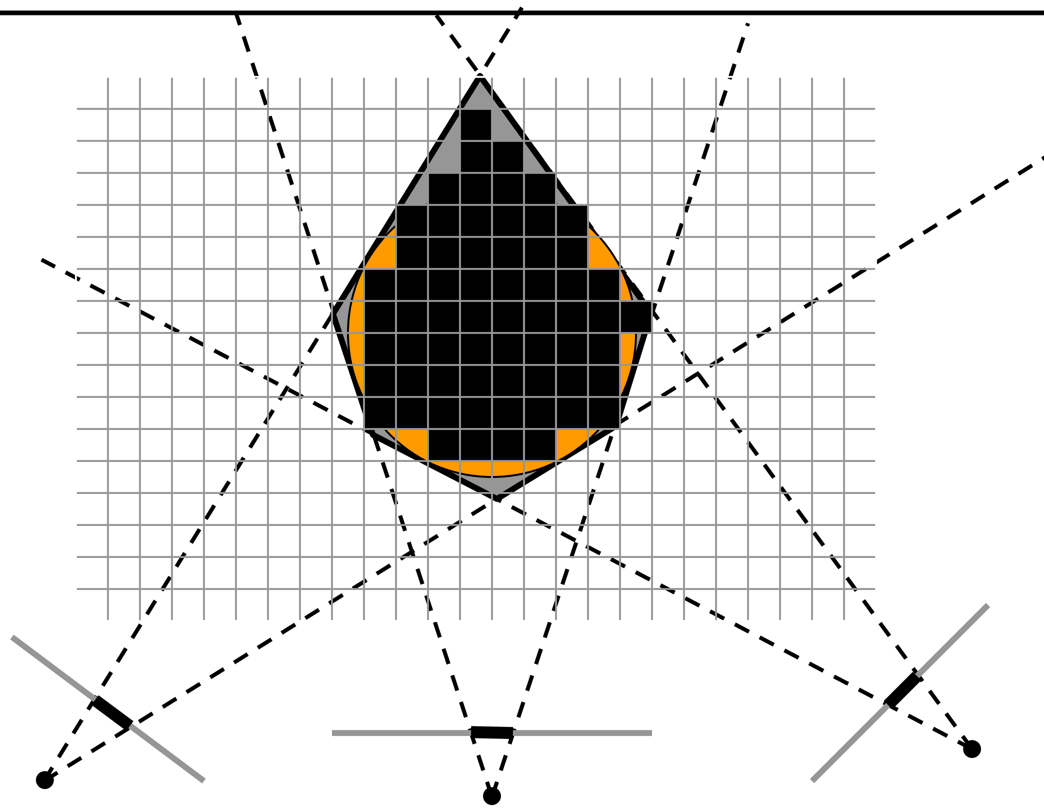
Volume intersection



Reconstruction Contains the True Scene

- In the limit (all views) get *visual hull*
  - > Complement of all lines that don’t intersect S

Voxel algorithm for volume intersection



Color voxel black if on silhouette in every image

- $O(MN^3)$, for M images, N^3 voxels
- Don't have to search 2^{N^3} possible scenes!

Properties of Volume Intersection

Pros

- Easy to implement, fast
- Accelerated via octrees [Szeliski 1993]

Cons

- No concavities
- Reconstruction is not photo-consistent
- Requires identification of silhouettes

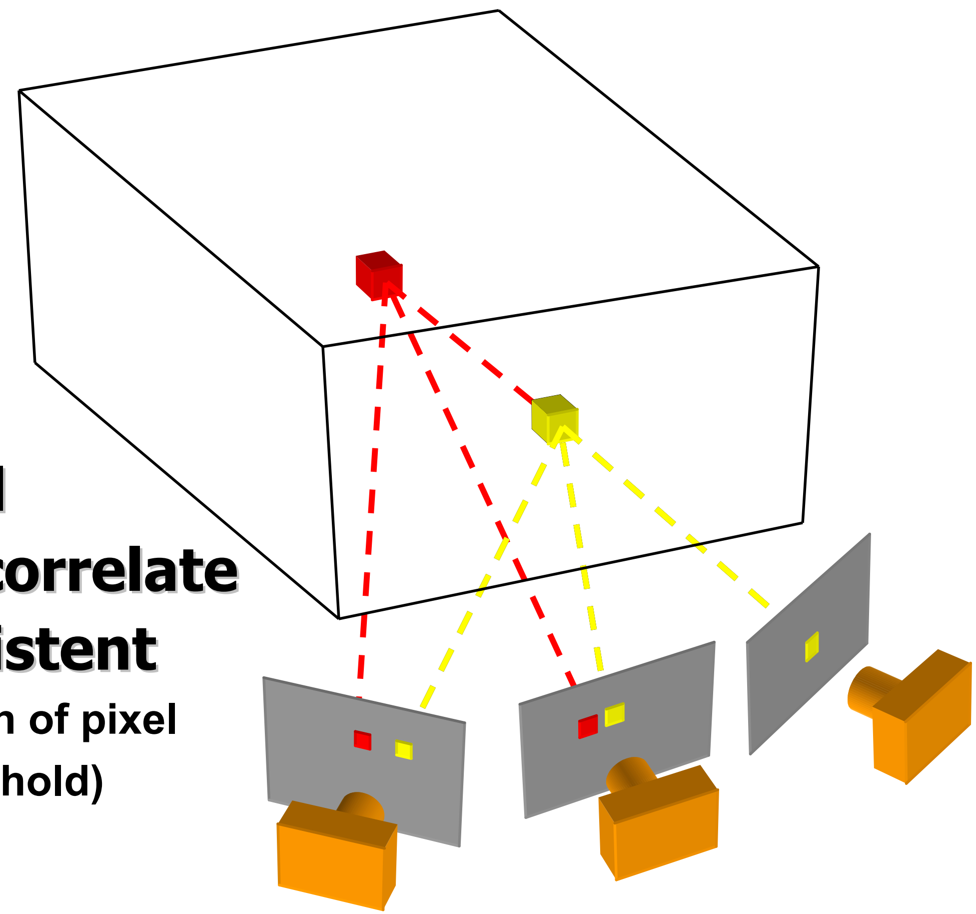
Voxel Coloring Solutions

1. C=2 (silhouettes)
   - Volume intersection [Baumgart 1974]

2. C unconstrained, viewpoint constraints
   - Voxel coloring algorithm [Seitz & Dyer 97]
     - For more info: http://www.cs.washington.edu/homes/seitz/papers/ijcv99.pdf

3. General Case
   - Space carving [Kutulakos & Seitz 98]

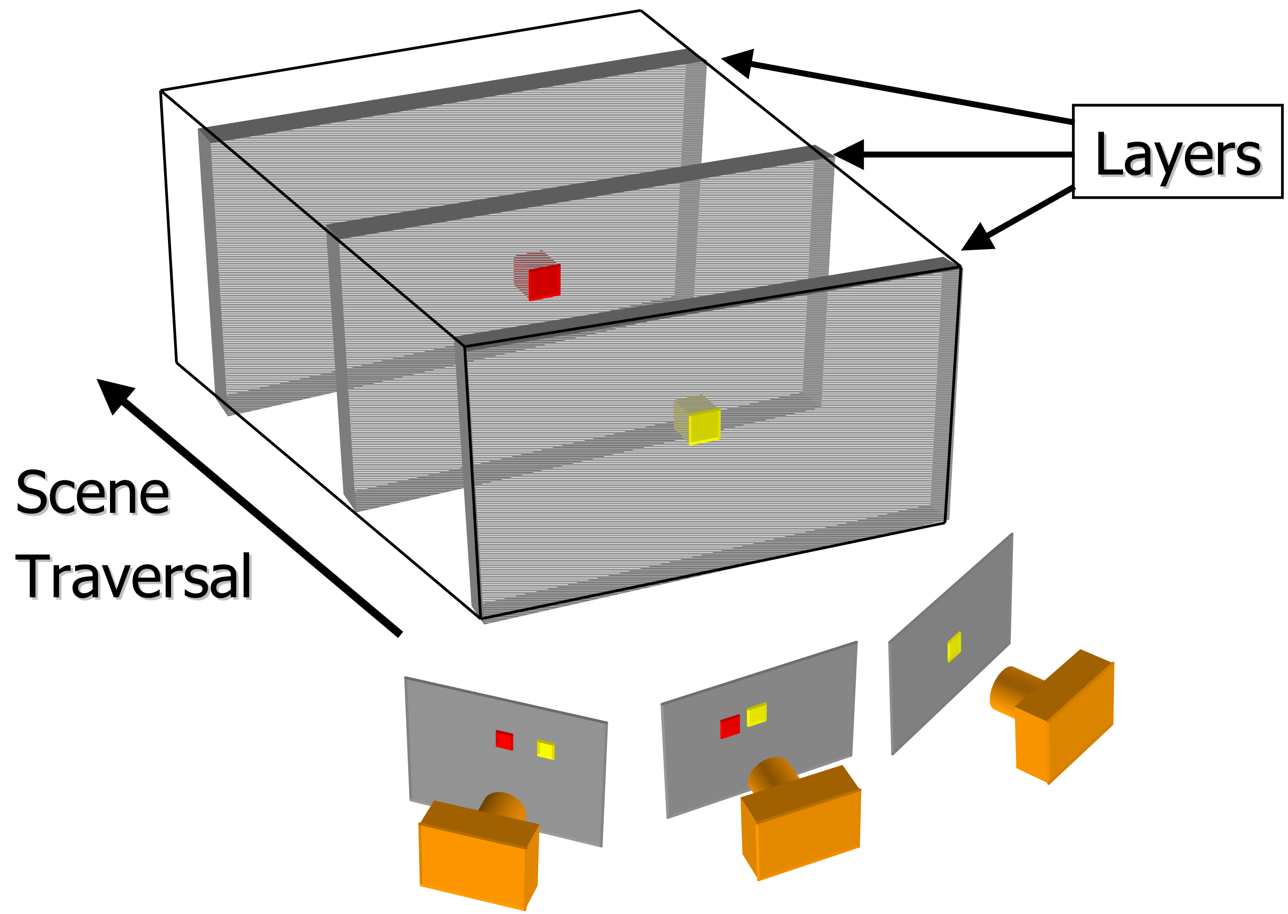
Voxel Coloring Approach



1. **Choose voxel**
2. **Project and correlate**
3. **Color if consistent**
   (standard deviation of pixel colors below threshold)

Visibility Problem: in which images is each voxel visible?

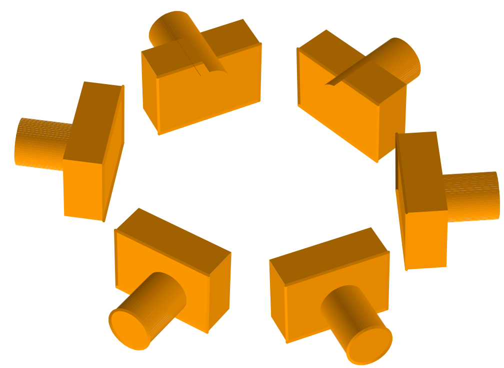
Depth Ordering: visit occluders first!



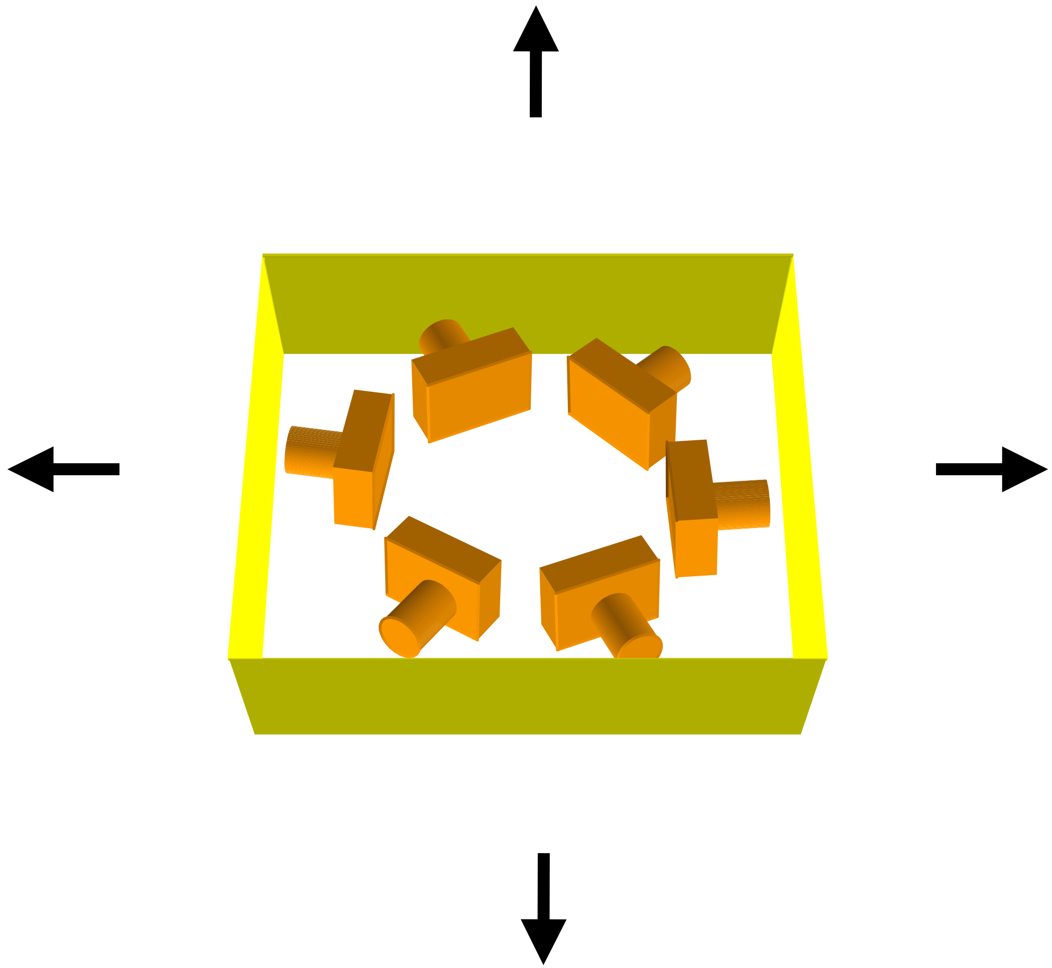
Condition: **depth order is the *same for all input views***

Panoramic Depth Ordering

- Cameras oriented in many different directions
- Planar depth ordering does not apply

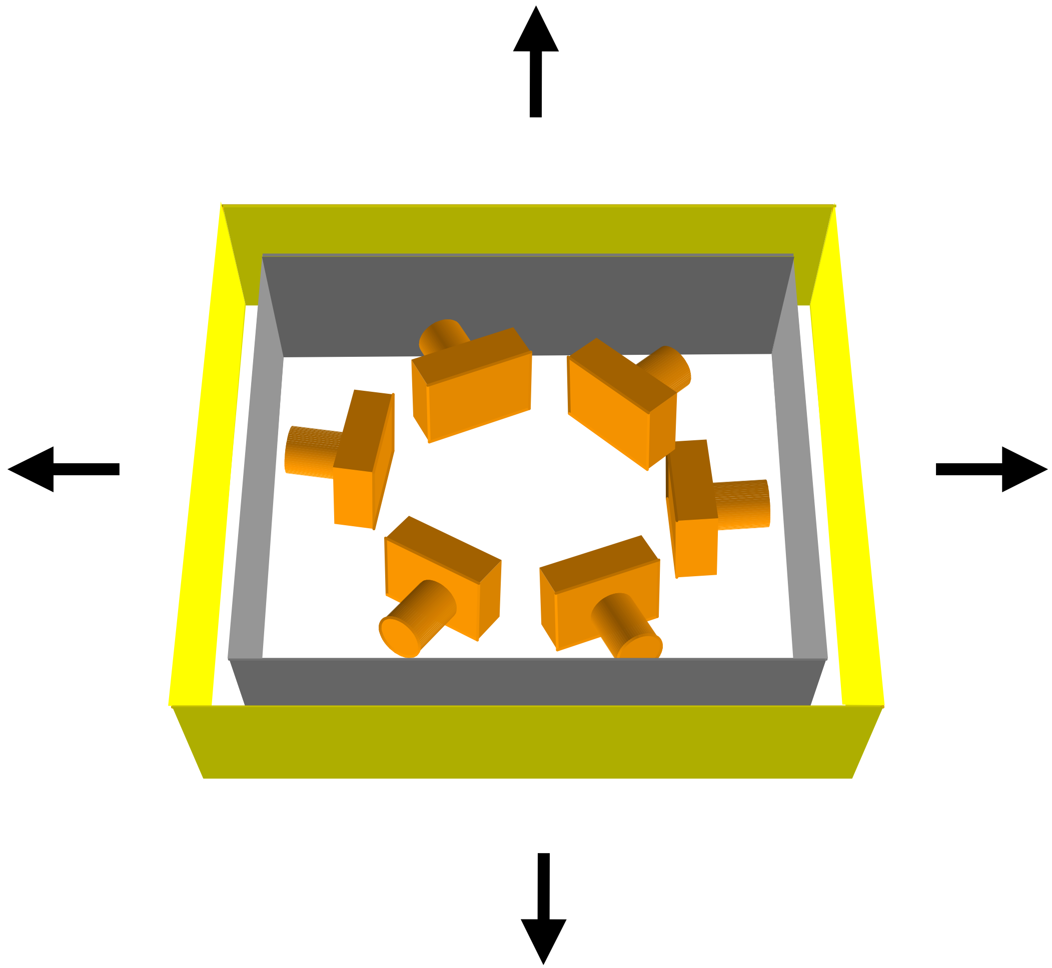


Panoramic Depth Ordering



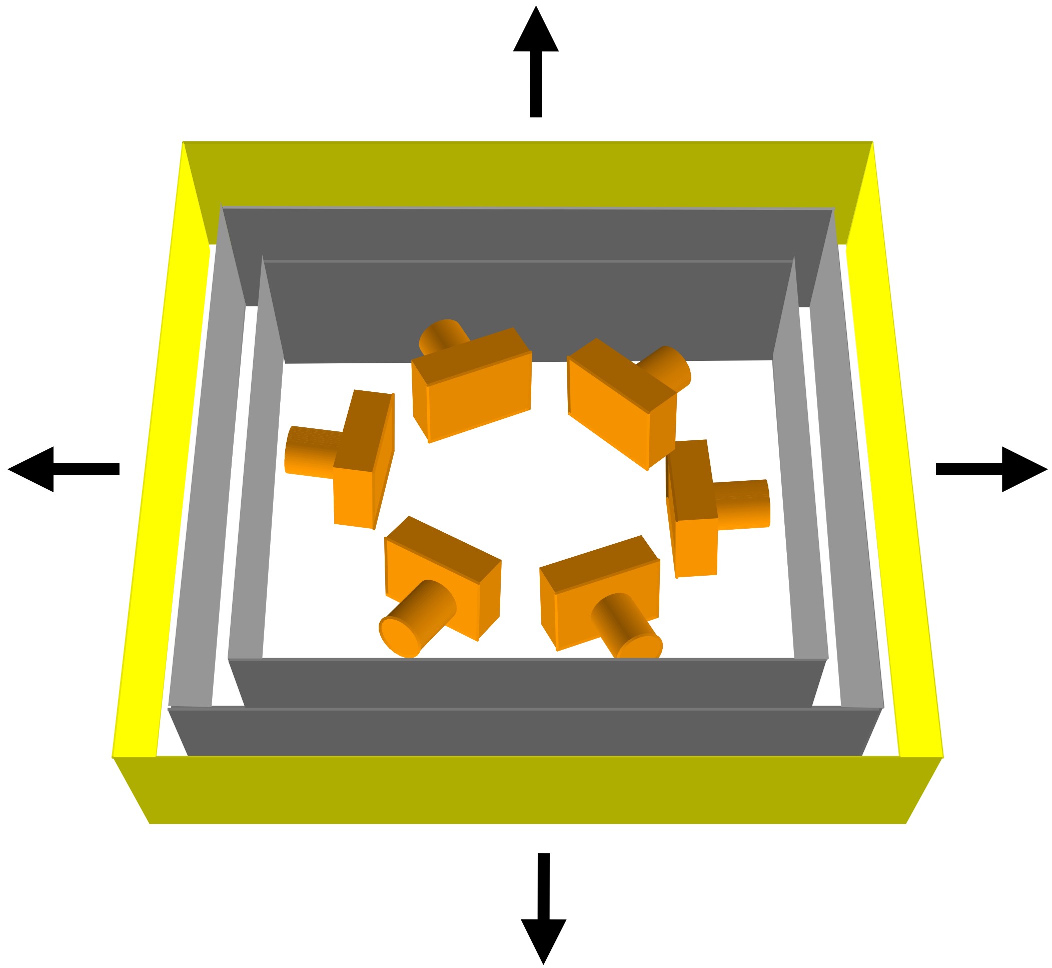
Layers radiate outwards from cameras

Panoramic Layering



Layers radiate outwards from cameras

Panoramic Layering

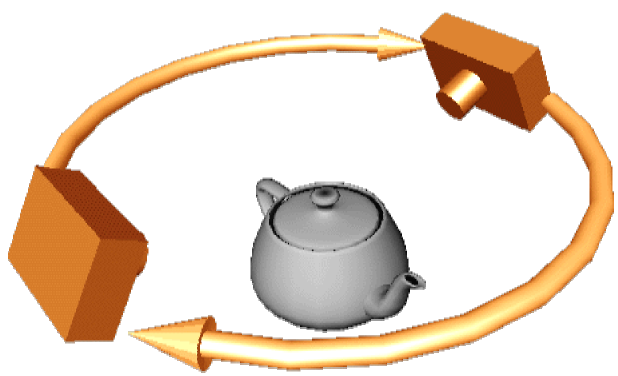


Layers radiate outwards from cameras

Compatible Camera Configurations

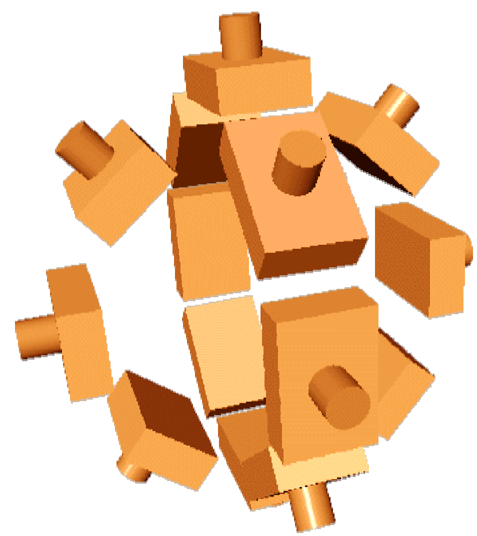
Depth-Order Constraint

- Scene outside convex hull of camera centers



Inward-Looking

Camera above scene

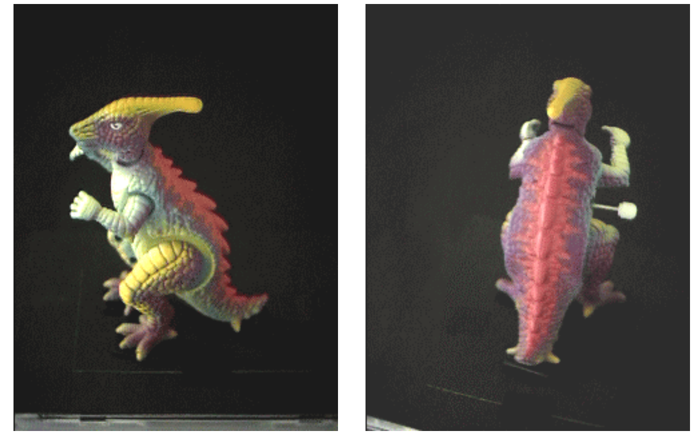


Outward-Looking

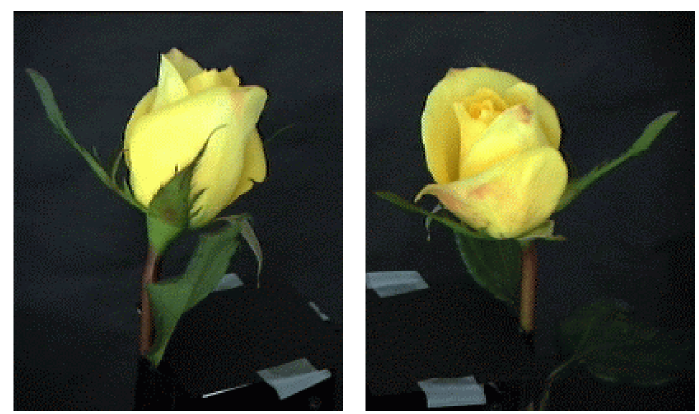
Camera inside scene

Calibrated Image Acquisition

Calibrated Turntable



Selected Dinosaur Images



Selected Flower Images

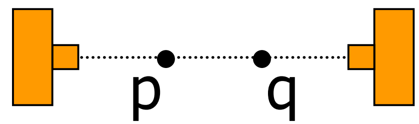
Voxel Coloring Results (Video)

Dinosaur Reconstruction
72 K voxels colored
7.6 M voxels tested
7 min. to compute
on a 250MHz SGI

Flower Reconstruction
70 K voxels colored
7.6 M voxels tested
7 min. to compute
on a 250MHz SGI

Limitations of Depth Ordering

A view-independent depth order may not exist



Need more powerful general-case algorithms

- Unconstrained camera positions
- Unconstrained scene geometry/topology

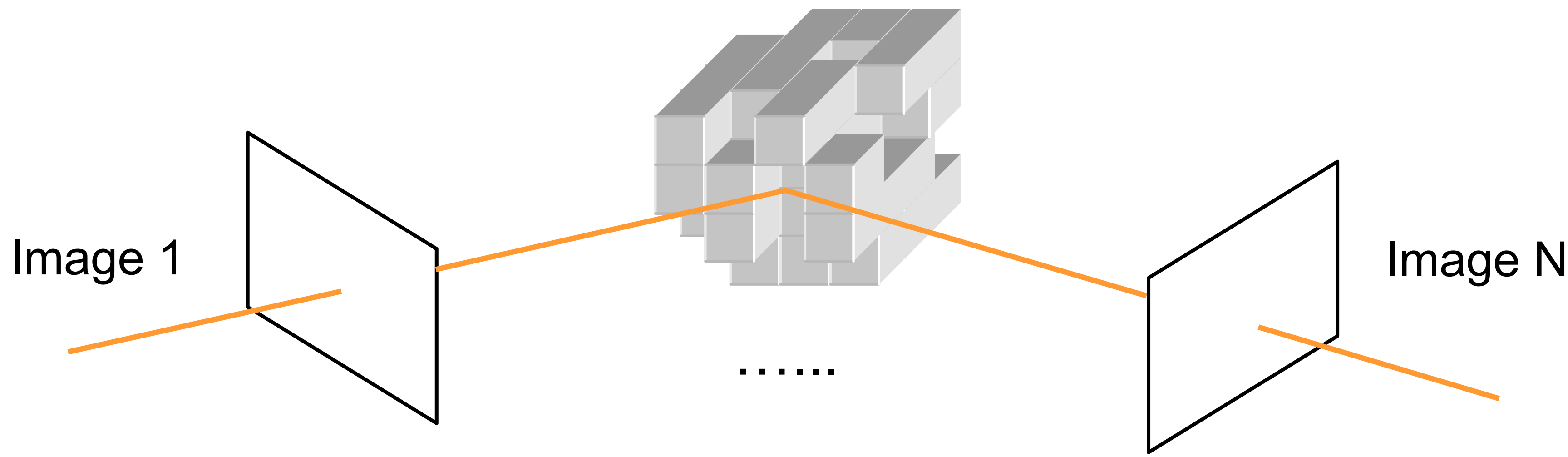
Voxel Coloring Solutions

1. C=2 (silhouettes)
   - Volume intersection [Baumgart 1974]

2. C unconstrained, viewpoint constraints
   - Voxel coloring algorithm [Seitz & Dyer 97]

3. General Case
   - Space carving [Kutulakos & Seitz 98]
     - For more info: http://www.cs.washington.edu/homes/seitz/papers/kutu-ijcv00.pdf

Space Carving Algorithm



Space Carving Algorithm

- Initialize to a volume V containing the true scene
- Choose a voxel on the current surface
- Project to visible input images
- Carve if not photo-consistent
- Repeat until convergence

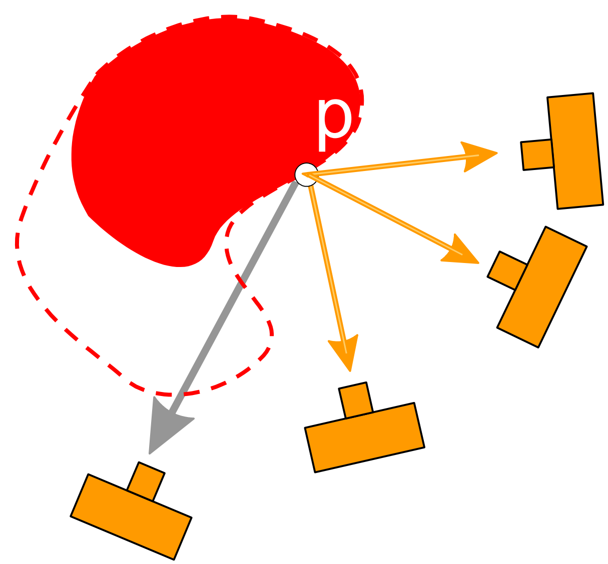
Convergence

Consistency Property

- The resulting shape is photo-consistent
  - > all inconsistent points are removed

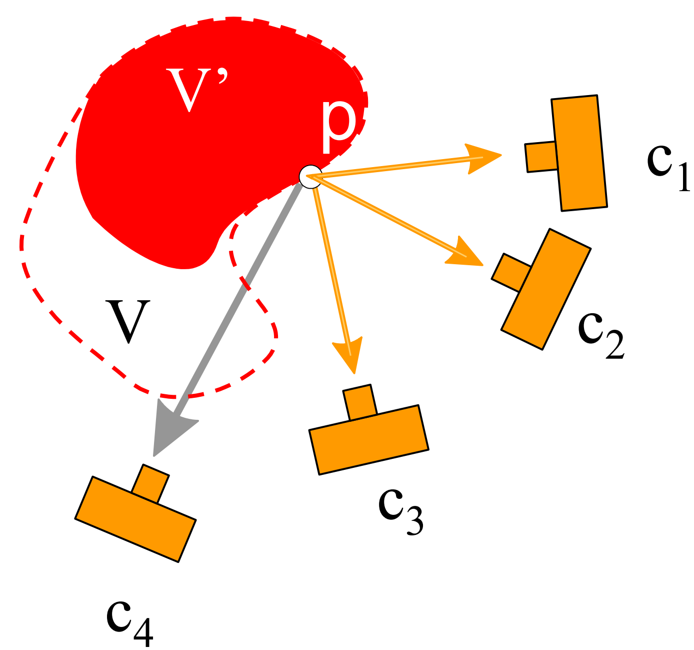
Convergence Property

- Carving converges to a non-empty shape
  - > a point on the true scene is *never* removed



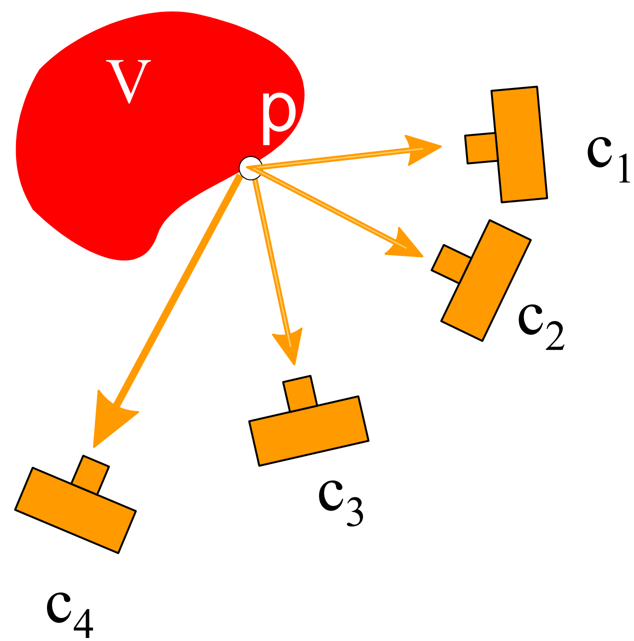
Visibility lemma

Let p be a point on V's surface, Surf(V), and let $Vis_v(p)$ be the collection of input images in which V does not occlude p. If V', a subset of V, is a shape that also has p on its surface, $Vis_v(p)$ is a subset of $Vis_{v'}(p)$.

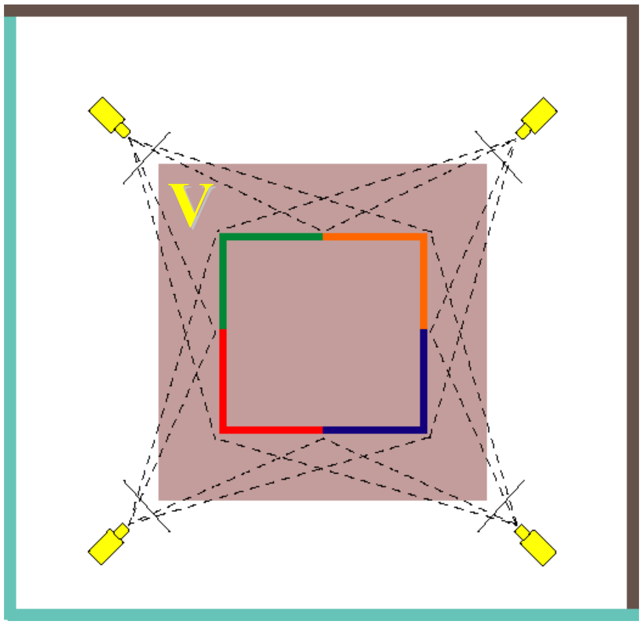


Non-photo-consistency lemma

Let p, which is in Surf(V), is not photo-consistent with a subset of $Vis_v(p)$, it is not photo-consistent with the entire $Vis_v(p)$.



Which shape do you get?



True Scene

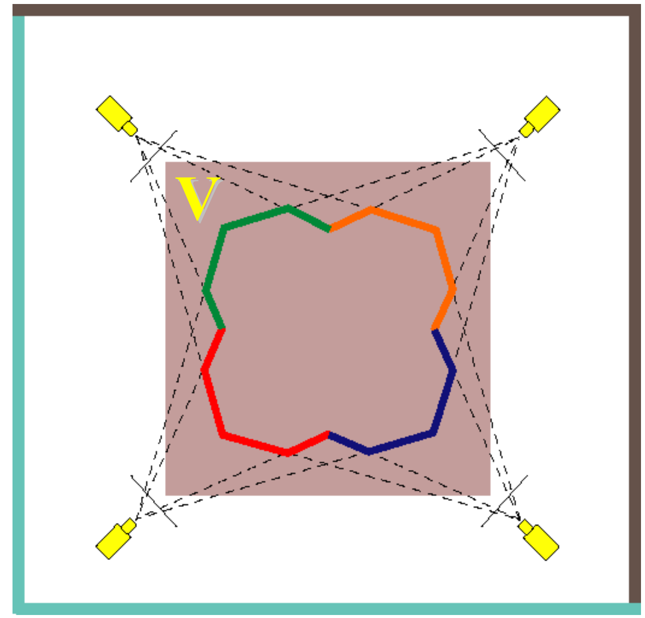


Photo Hull

The Photo Hull *is the UNION of all photo-consistent scenes in V*

- It is a photo-consistent scene reconstruction
- Tightest possible bound on the true scene

Space Carving Algorithm

The Basic Algorithm is Unwieldy

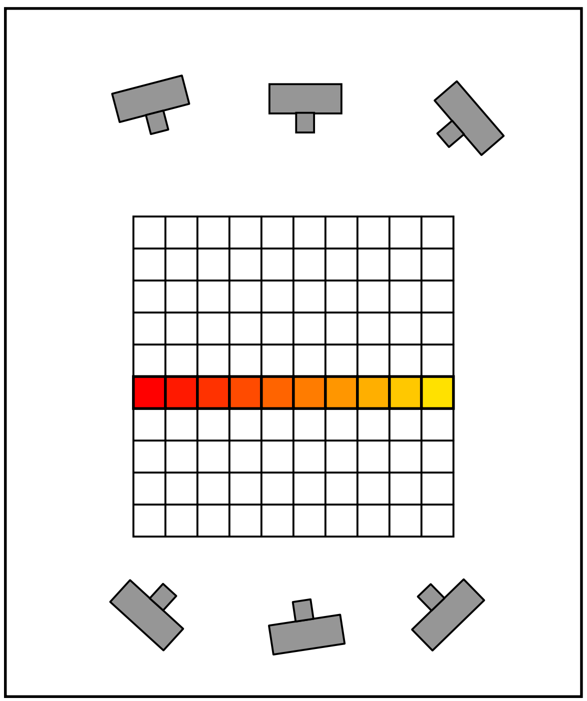
- Complex update procedure

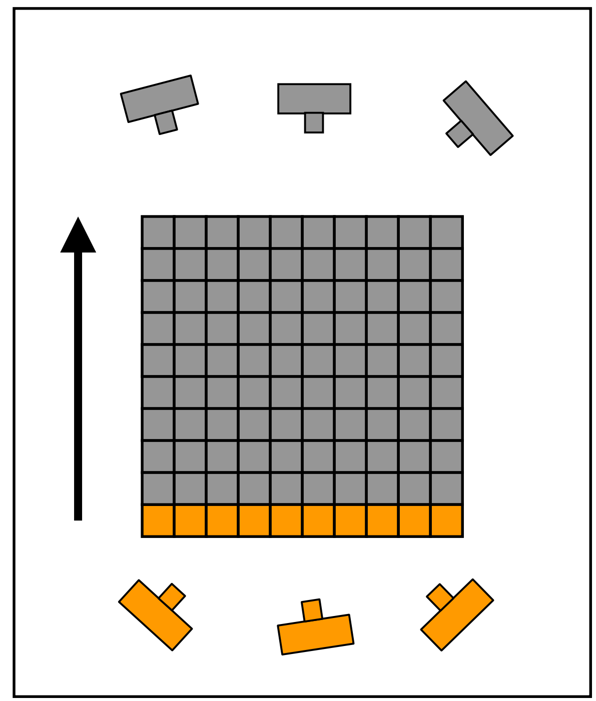
Alternative: Multi-Pass Plane Sweep

- Efficient, can use texture-mapping hardware
- Converges quickly in practice
- Easy to implement

Multi-Pass Plane Sweep

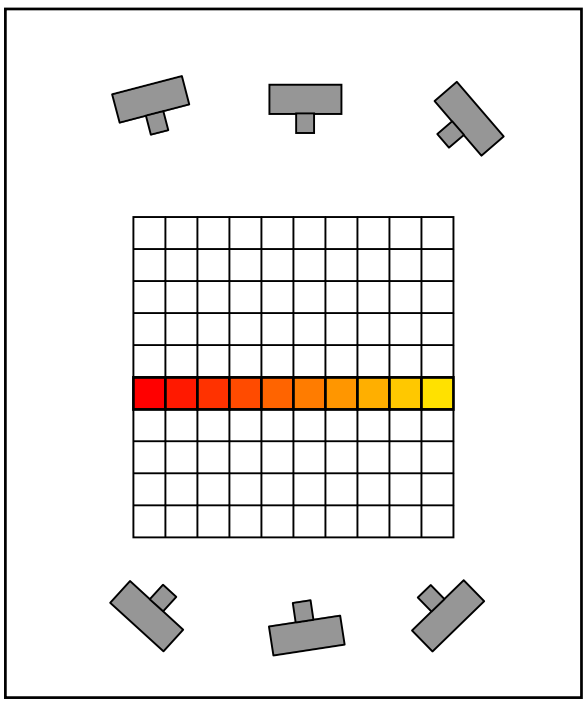
- Sweep plane in each of 6 principle directions
- Consider cameras on only one side of plane
- Repeat until convergence

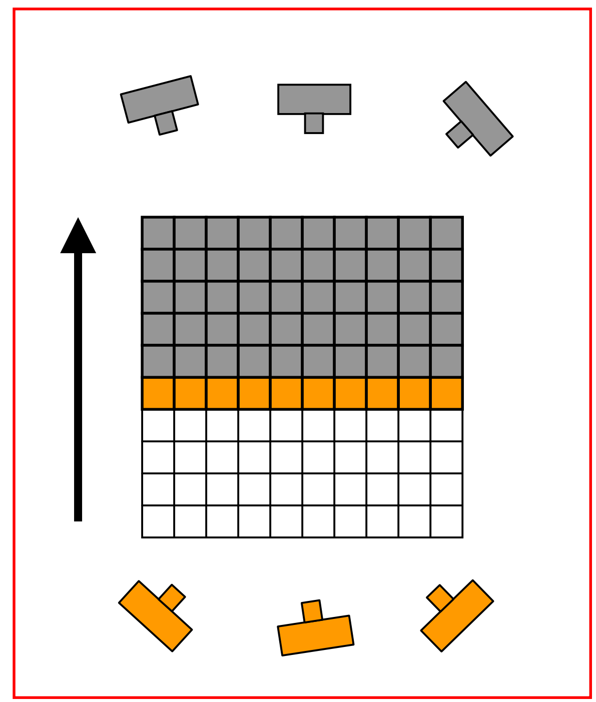




Multi-Pass Plane Sweep

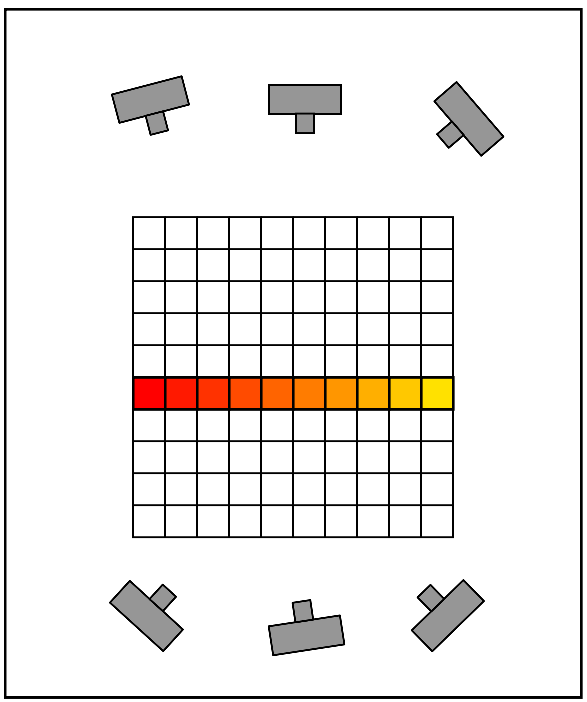
- Sweep plane in each of 6 principle directions
- Consider cameras on only one side of plane
- Repeat until convergence

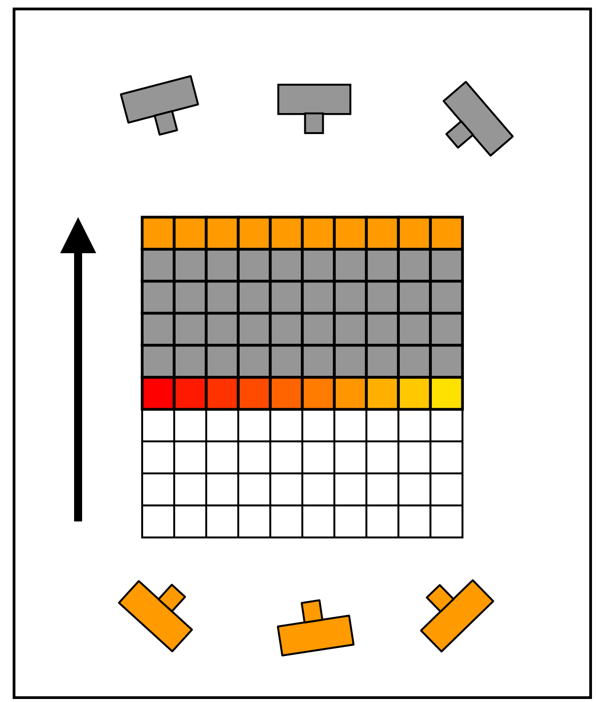




Multi-Pass Plane Sweep

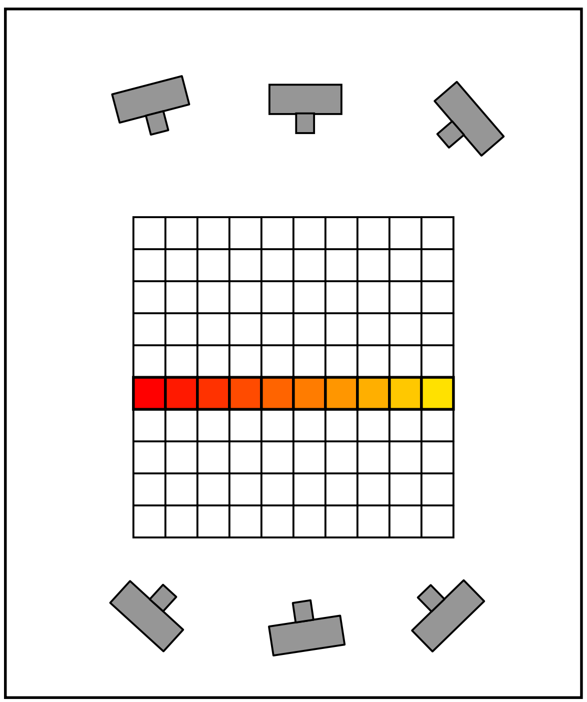
- Sweep plane in each of 6 principle directions
- Consider cameras on only one side of plane
- Repeat until convergence

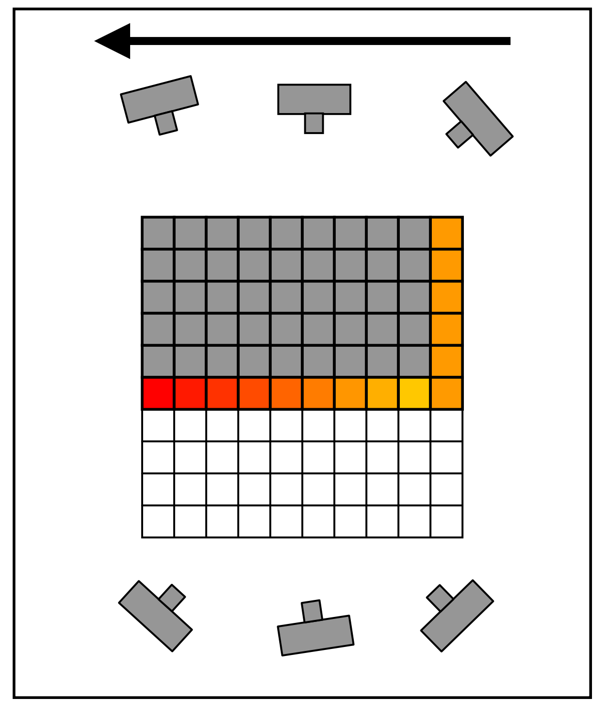




Multi-Pass Plane Sweep

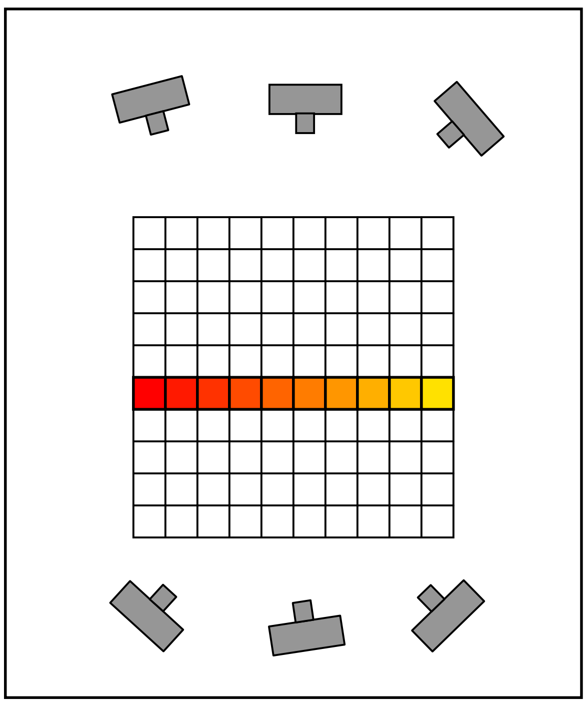
- Sweep plane in each of 6 principle directions
- Consider cameras on only one side of plane
- Repeat until convergence

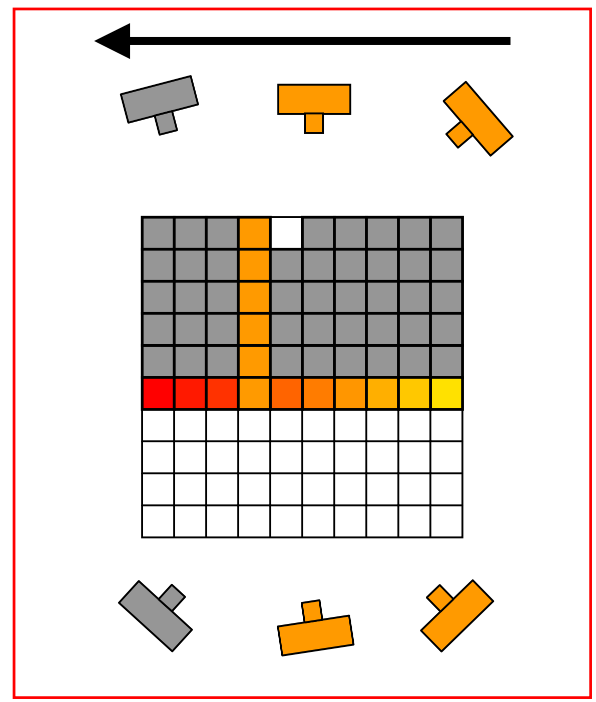




Multi-Pass Plane Sweep

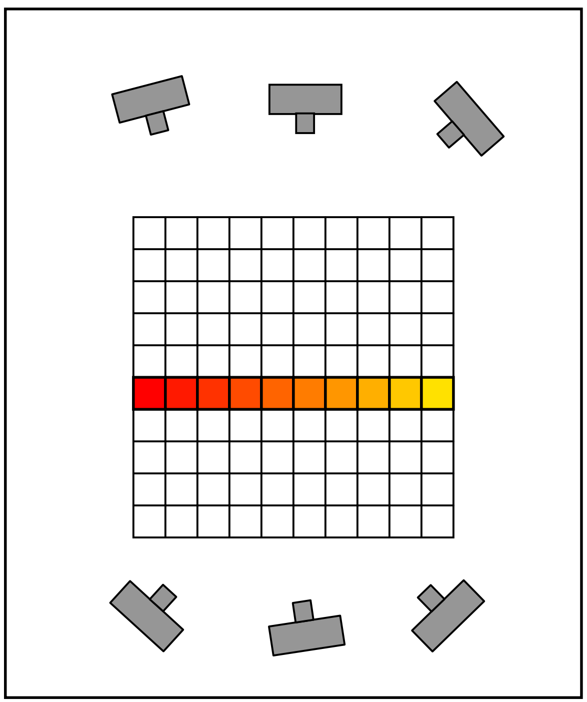
- Sweep plane in each of 6 principle directions
- Consider cameras on only one side of plane
- Repeat until convergence

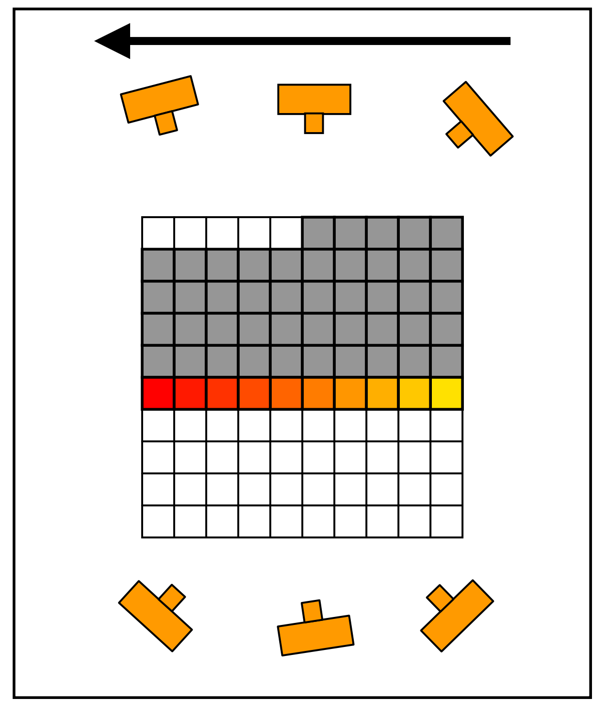




Multi-Pass Plane Sweep

- Sweep plane in each of 6 principle directions
- Consider cameras on only one side of plane
- Repeat until convergence



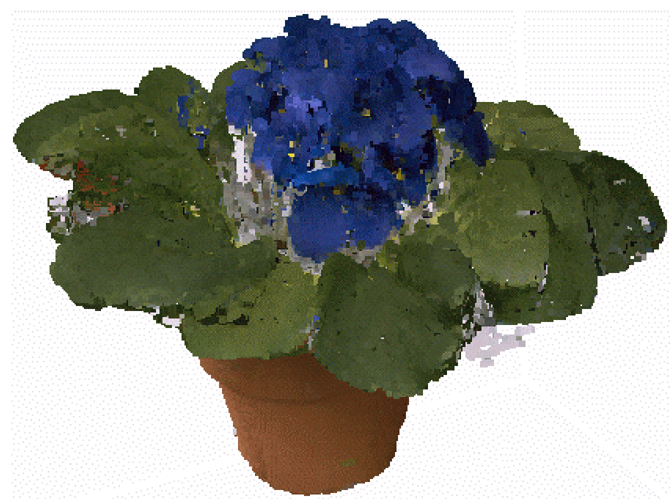


Space Carving Results: African Violet

Input Image (1 of 45)

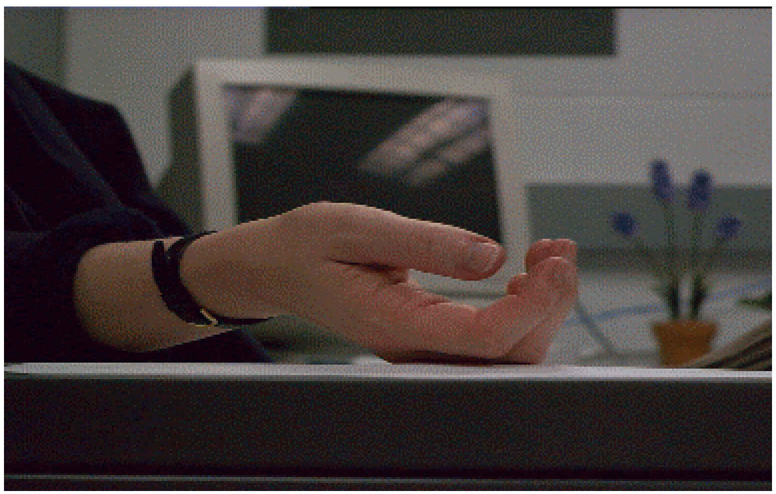
Reconstruction

Reconstruction



Reconstruction

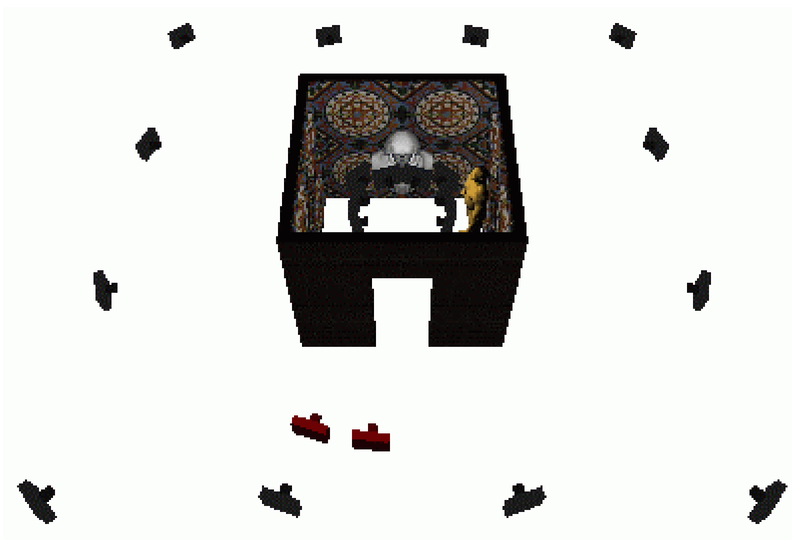
Space Carving Results: Hand

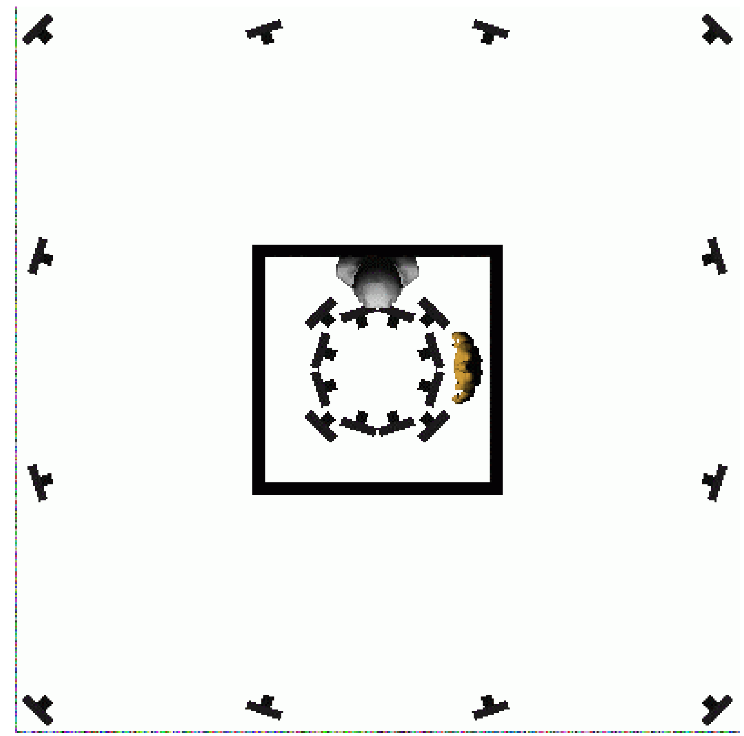


Input Image
(1 of 100)

Views of Reconstruction

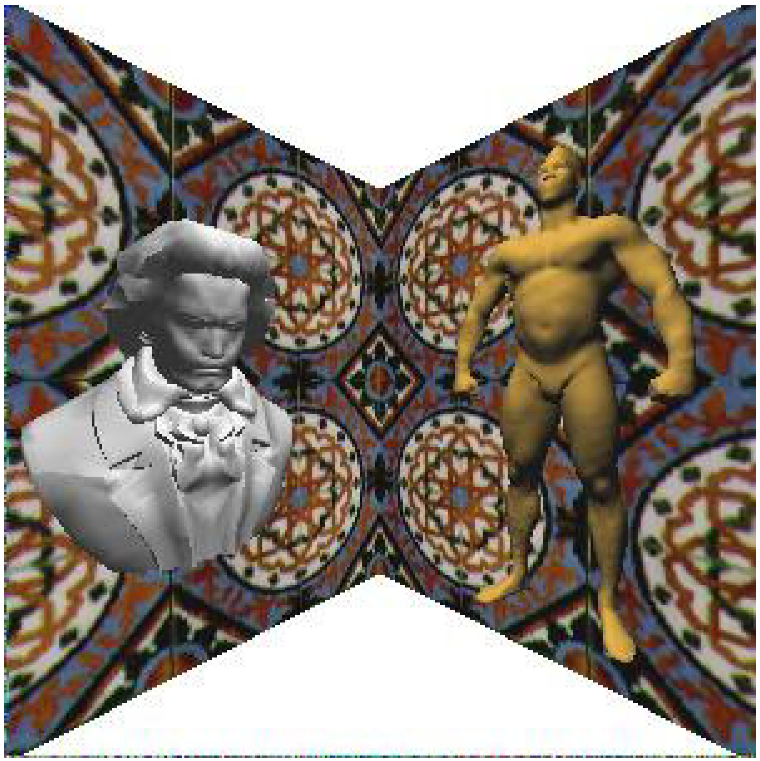
House Walkthrough





24 rendered input views from inside *and* outside

Space Carving Results: House



Input Image
(true scene)

Reconstruction
370,000 voxels

Space Carving Results: House

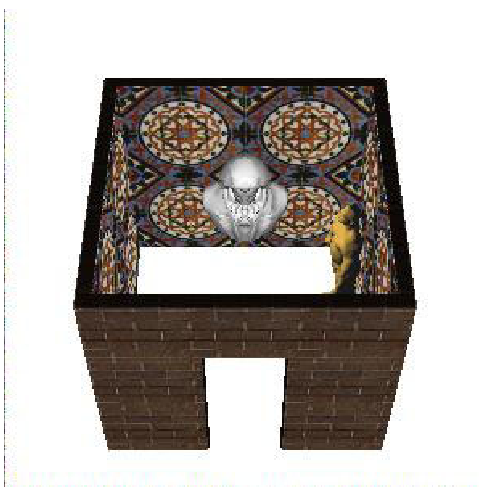
Input Image
(true scene)

Reconstruction
370,000 voxels

Space Carving Results: House

New View (true scene)

Reconstruction



New View
(true scene)

Reconstruction

Reconstruction
(with new input view)

Other Features

Coarse-to-fine Reconstruction

- Represent scene as octree
- Reconstruct low-res model first, then refine

Hardware-Acceleration

- Use texture-mapping to compute voxel projections
- Process voxels an entire plane at a time

Limitations

- Need to acquire calibrated images
- Restriction to simple radiance models
- Bias toward maximal (fat) reconstructions
- Transparency not supported

Other Approaches

Level-Set Methods ***[Faugeras & Keriven 1998]***

- Evolve implicit function by solving PDE's

Probabilistic Voxel Reconstruction ***[DeBonet & Viola 1999], [Broadhurst et al. 2001]***

- Solve for voxel uncertainty (also transparency)

Transparency and Matting ***[Szeliski & Golland 1998]***

- Compute voxels with alpha-channel

Max Flow/Min Cut ***[Roy & Cox 1998]***

- Graph theoretic formulation

Mesh-Based Stereo ***[Fua & Leclerc 1995], [Zhang & Seitz 2001]***

- Mesh-based but similar consistency formulation

Virtualized Reality ***[Narayan, Rander, Kanade 1998]***

- Perform stereo 3 images at a time, merge results

Bibliography

Volume Intersection

- Martin & Aggarwal, “Volumetric description of objects from multiple views”, Trans. Pattern Analysis and Machine Intelligence, 5(2), 1991, pp. 150-158.
- Szeliski, “Rapid Octree Construction from Image Sequences”, Computer Vision, Graphics, and Image Processing: Image Understanding, 58(1), 1993, pp. 23-32.

Voxel Coloring and Space Carving

- Seitz & Dyer, “Photorealistic Scene Reconstruction by Voxel Coloring”, Proc. Computer Vision and Pattern Recognition (CVPR), 1997, pp. 1067-1073.
- Seitz & Kutulakos, “Plenoptic Image Editing”, Proc. Int. Conf. on Computer Vision (ICCV), 1998, pp. 17-24.
- Kutulakos & Seitz, “A Theory of Shape by Space Carving”, Proc. ICCV, 1998, pp. 307-314.

Bibliography

Related References

- Bolles, Baker, and Marimont, “Epipolar-Plane Image Analysis: An Approach to Determining Structure from Motion”, International Journal of Computer Vision, vol 1, no 1, 1987, pp. 7-55.
- DeBonet & Viola, “Poxels: Probabilistic Voxelized Volume Reconstruction”, Proc. Int. Conf. on Computer Vision (ICCV) 1999.
- Broadhurst, Drummond, and Cipolla, "A Probabilistic Framework for Space Carving“, International Conference of Computer Vision (ICCV), 2001, pp. 388-393.
- Faugeras & Keriven, “Variational principles, surface evolution, PDE's, level set methods and the stereo problem", IEEE Trans. on Image Processing, 7(3), 1998, pp. 336-344.
- Szeliski & Golland, “Stereo Matching with Transparency and Matting”, Proc. Int. Conf. on Computer Vision (ICCV), 1998, 517-524.
- Roy & Cox, “A Maximum-Flow Formulation of the N-camera Stereo Correspondence Problem”, Proc. ICCV, 1998, pp. 492-499.
- Fua & Leclerc, “Object-centered surface reconstruction: Combining multi-image stereo and shading", International Journal of Computer Vision, 16, 1995, pp. 35-56.
- Narayanan, Rander, & Kanade, “Constructing Virtual Worlds Using Dense Stereo”, Proc. ICCV, 1998, pp. 3-10.

Summary

Things to take away from this lecture

- Baseline tradeoff
- Multibaseline stereo approach
- Voxel coloring problem
- Volume intersection algorithm
- Voxel coloring algorithm
- Space carving algorithm