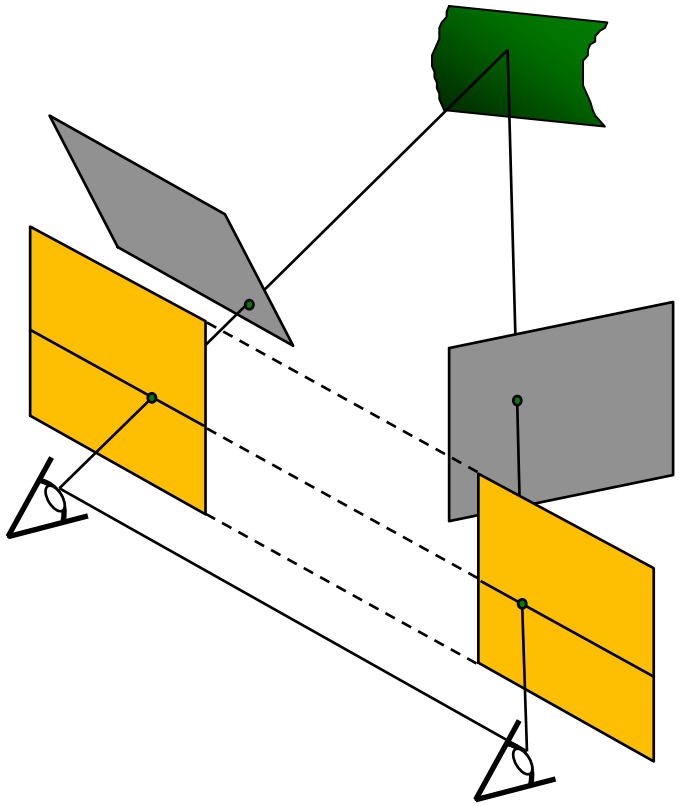
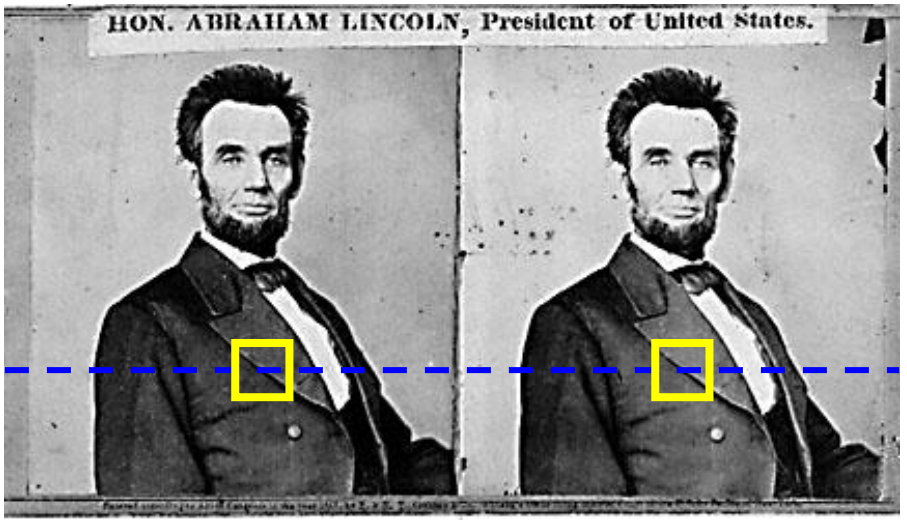
Topics to Review

- image processing
  - filtering & convolution
  - edge detection
  - resampling
  - seam carving
- cameras
  - projective geometry
  - single view modeling
  - epipolar geometry
- alignment
  - structure from motion
  - mosaics
- features
  - corner detection
  - SIFT
  - matching & RANSAC
- dense reconstruction
  - photometric stereo
  - two-view stereo
  - multiview stereo
- other
  - segmentation & recognition
  - color & lighting
  - tracking & optical flow

Last Time: 2-view stereo

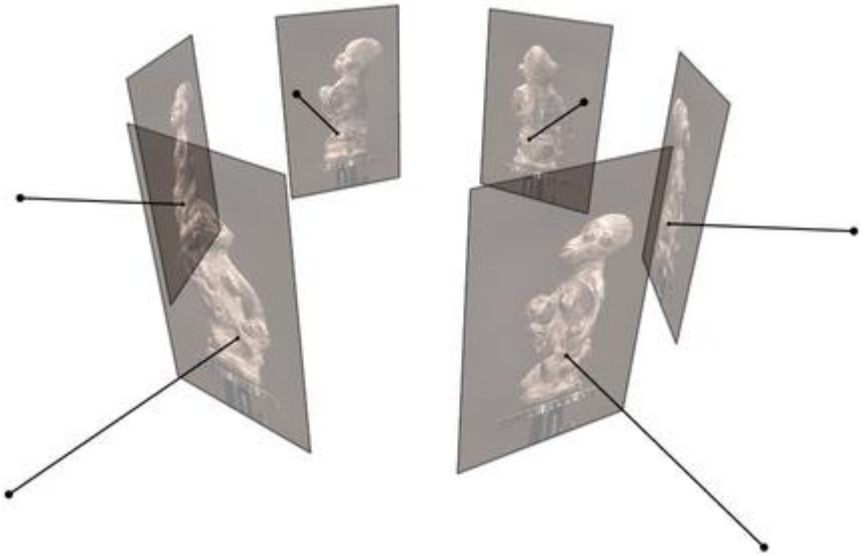




Multiview Stereo

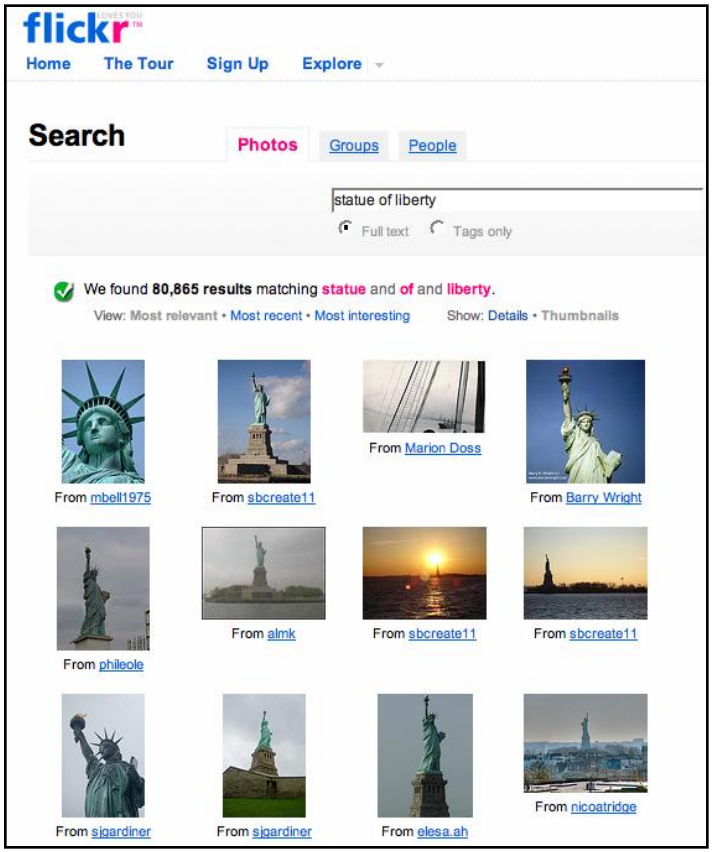
Input: calibrated images from several viewpoints

Output: 3D object model

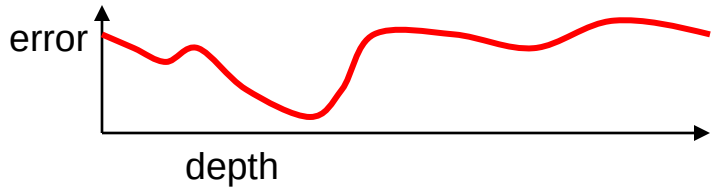


Figures by Carlos Hernandez

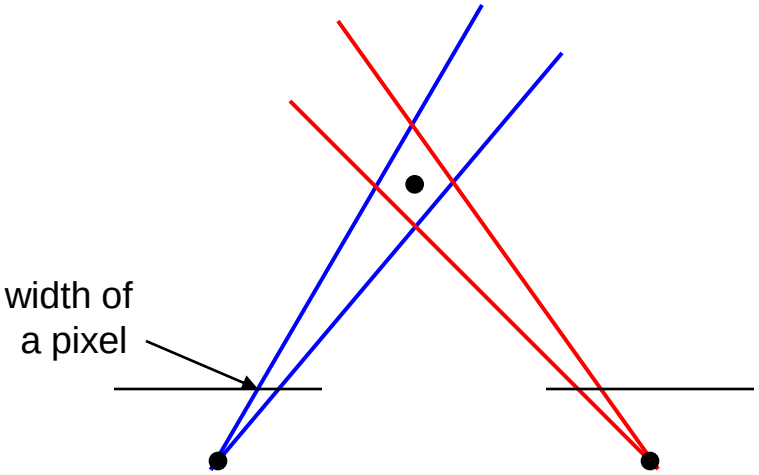
Multiview Stereo



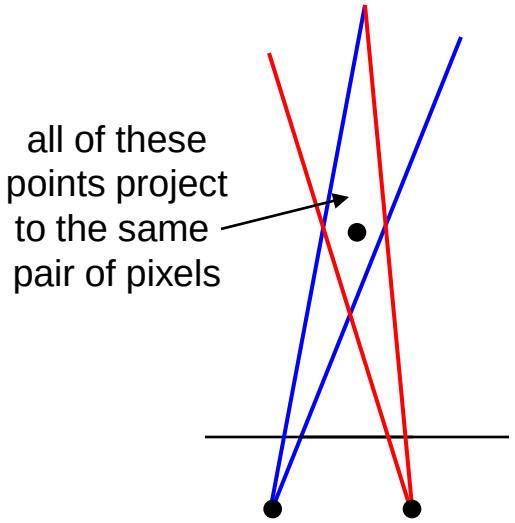
Stereo: basic idea



Choosing the stereo baseline



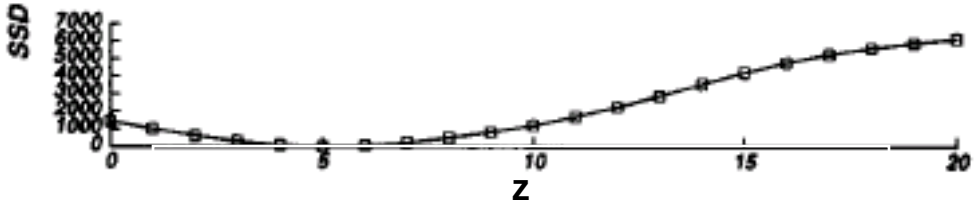
Large Baseline

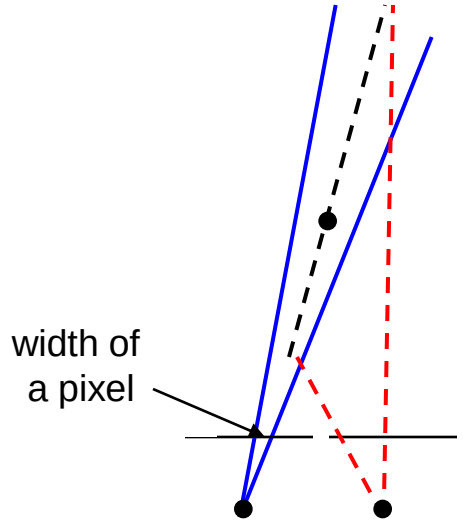


Small Baseline

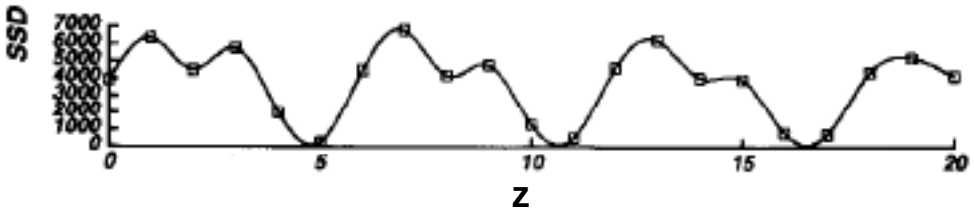
What's the optimal baseline?

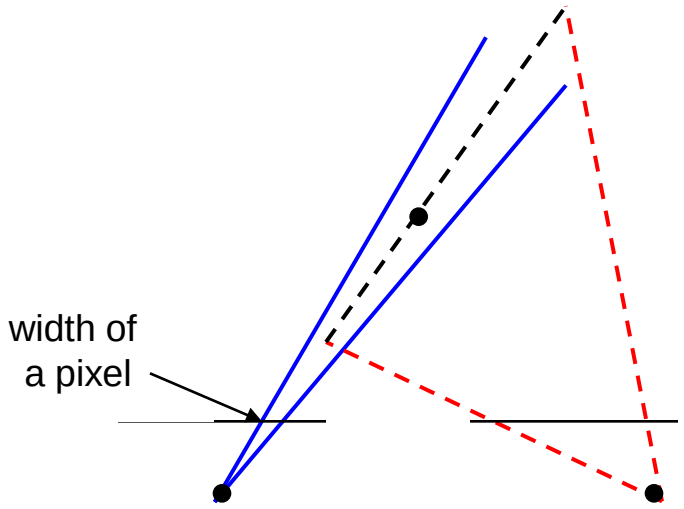
- Too small: large depth error
- Too large: difficult search problem





pixel matching score





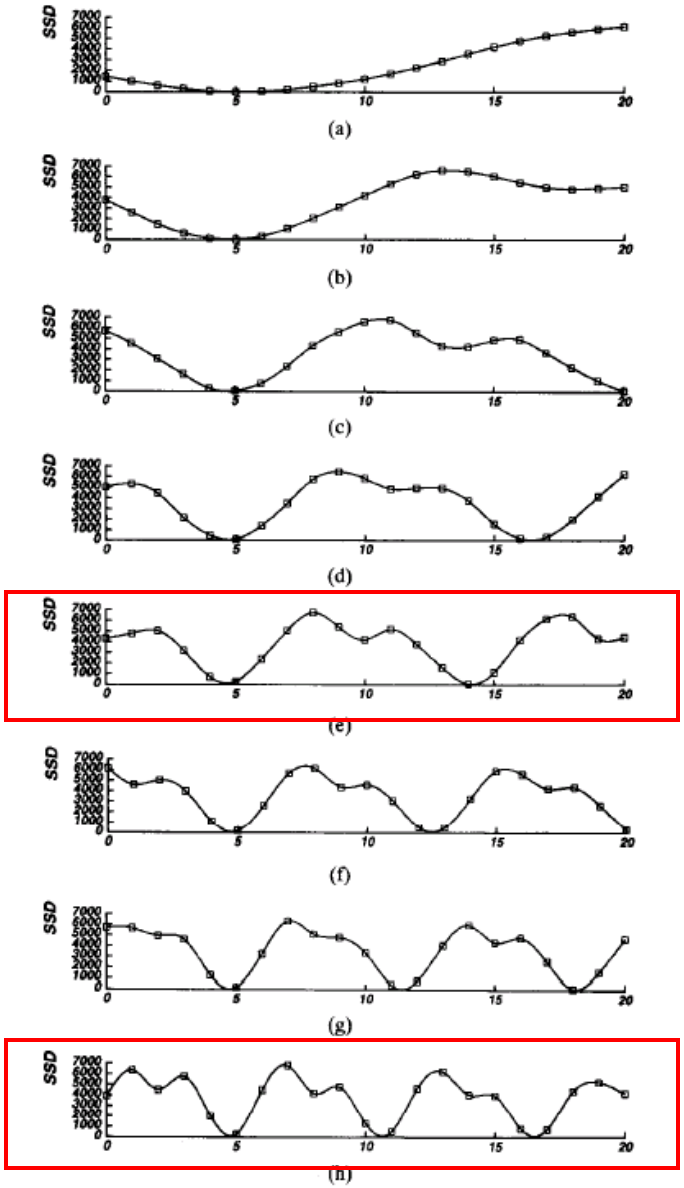


Fig. 5. SSD values versus inverse distance: (a) $B = b$; (b) $B = 2b$; (c) $B = 3b$; (d) $B = 4b$; (e) $B = 5b$; (f) $B = 6b$; (g) $B = 7b$; (h) $B = 8b$. The horizontal axis is normalized such that $8bF = 1$.

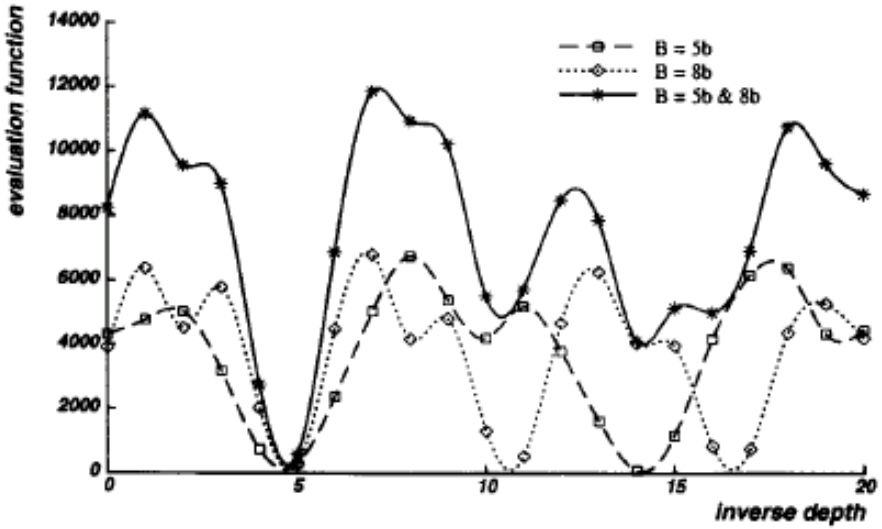


Fig. 6. Combining two stereo pairs with different baselines.

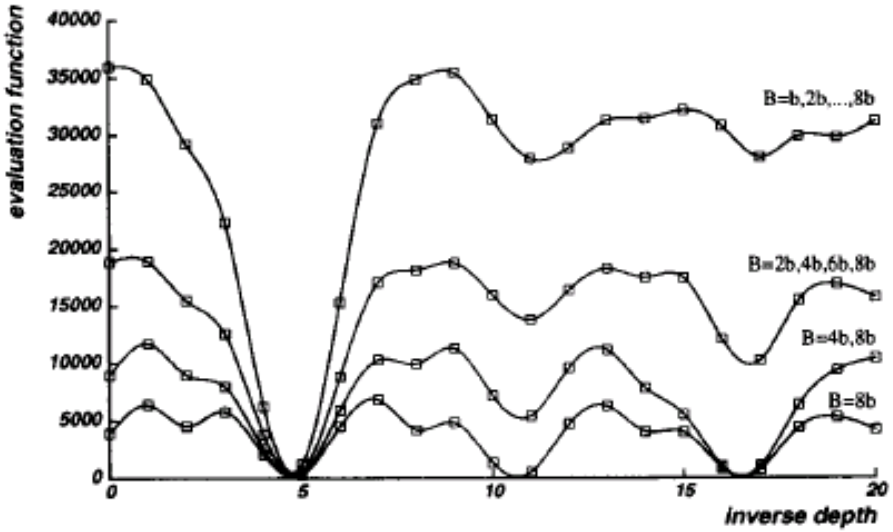


Fig. 7. Combining multiple baseline stereo pairs.

Multibaseline Stereo

Basic Approach

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

Limitations

- Only gives a depth map (not an “object model”)
- Won’t work for widely distributed views:

Problem: *visibility*

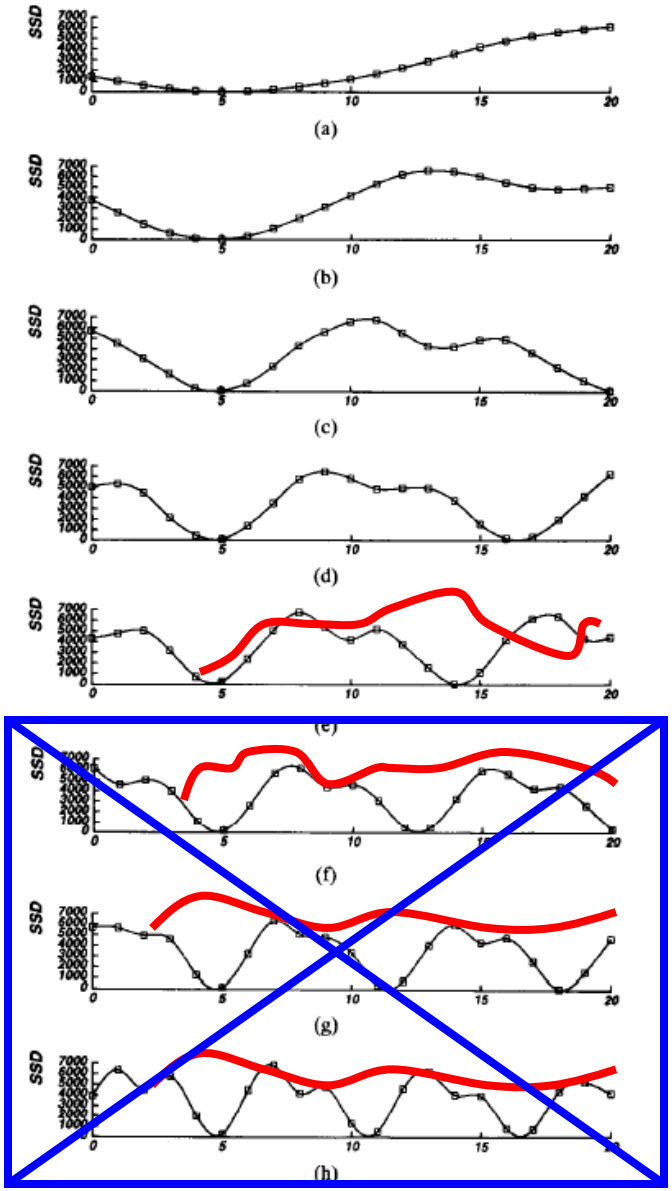


Fig. 5. SSD values versus inverse distance: (a) $B = b$; (b) $B = 2b$; (c) $B = 3b$; (d) $B = 4b$; (e) $B = 5b$; (f) $B = 6b$; (g) $B = 7b$; (h) $B = 8b$. The horizontal axis is normalized such that $8bF = 1$.

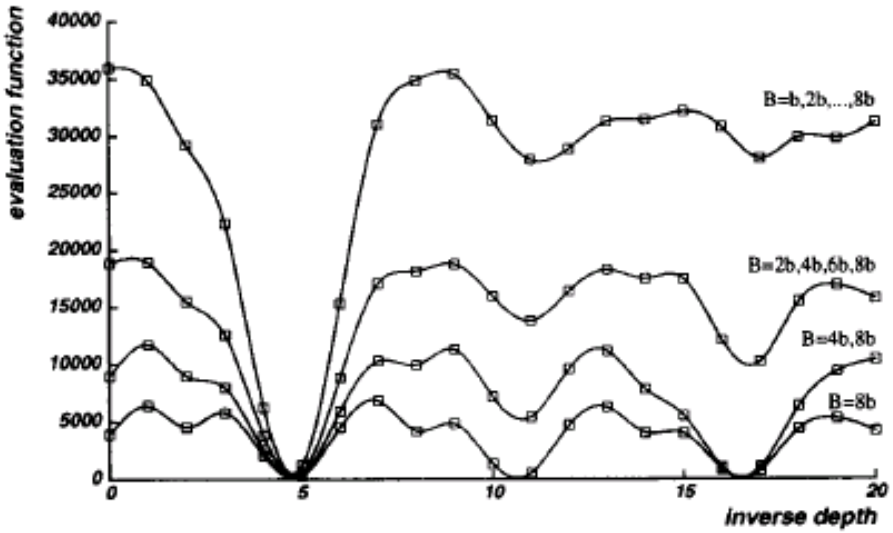


Fig. 7. Combining multiple baseline stereo pairs.

Some Solutions

- Match only nearby photos [Narayanan 98]
- Use NCC instead of SSD,
  Ignore NCC values < threshold
  [Hernandez & Schmitt 03]

Popular matching scores

- SSD (Sum Squared Distance)

$$\sum_{x,y} |W_1(x,y) - W_2(x,y)|^2$$

- NCC (Normalized Cross Correlation)

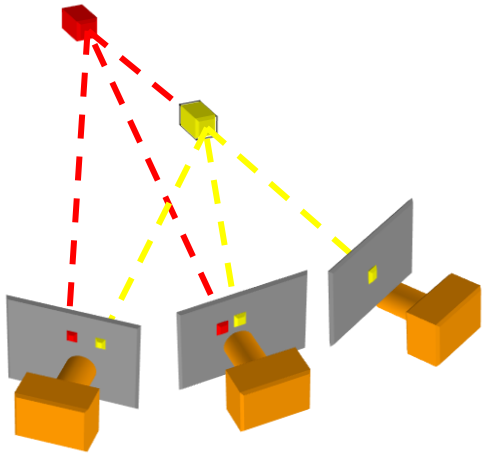
$$\frac{\sum_{x,y}(W_1(x,y) - \overline{W_1})(W_2(x,y) - \overline{W_2})}{\sigma_{W_1}\sigma_{W_2}}$$

  - where $\overline{W_i} = \frac{1}{n}\sum_{x,y} W_i \qquad \sigma_{W_i} = \sqrt{\frac{1}{n}\sum_{x,y}(W_i - \overline{W_i})^2}$
  - what advantages might NCC have?

Handling Visiblity: Two Approaches

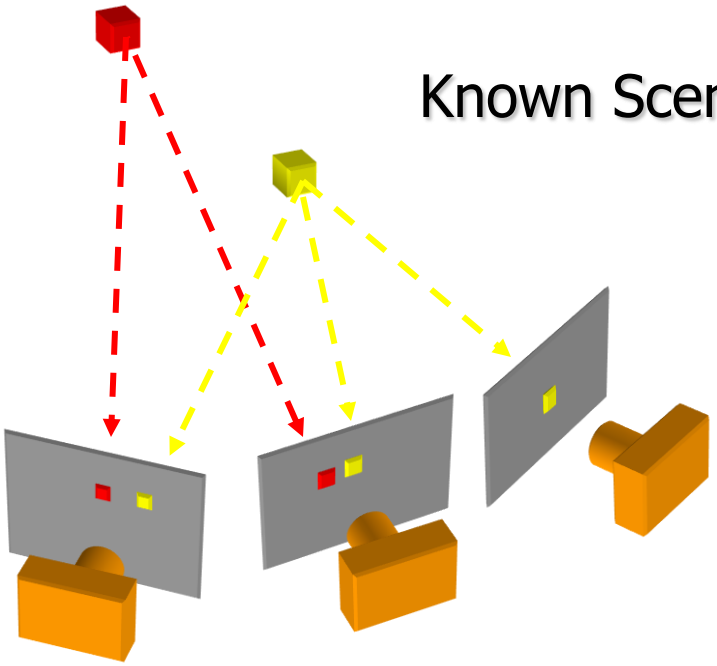
- Treat occlusions as outliers

- **Model occlusions geometrically**

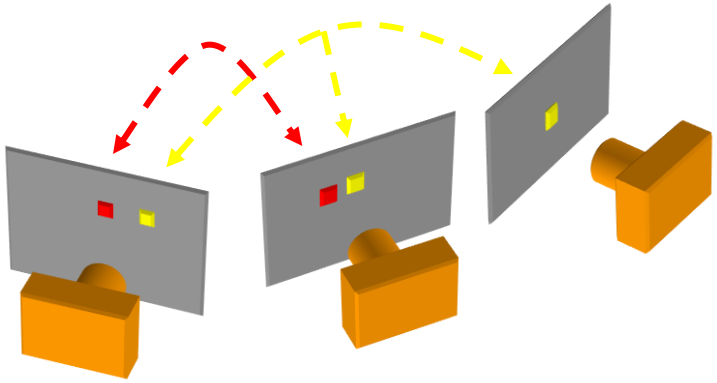


The visibility problem

Which points are visible in which images?



Known Scene

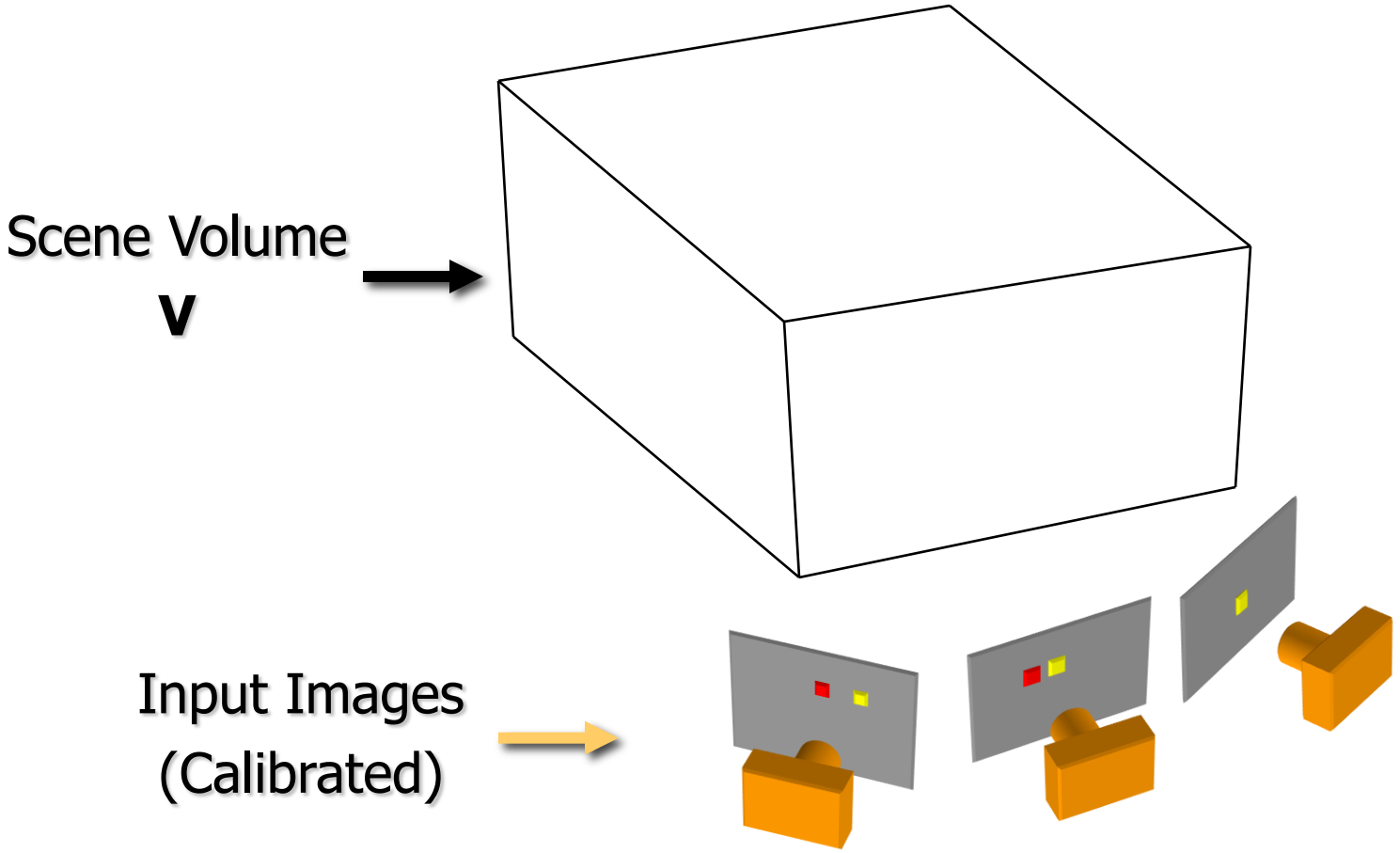


Unknown Scene

Forward Visibility

Inverse Visibility

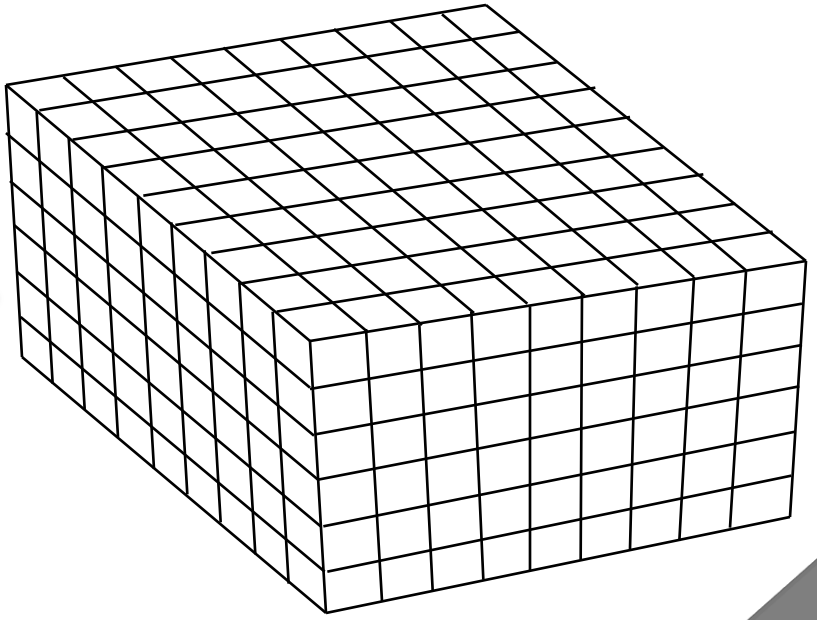
Volumetric stereo



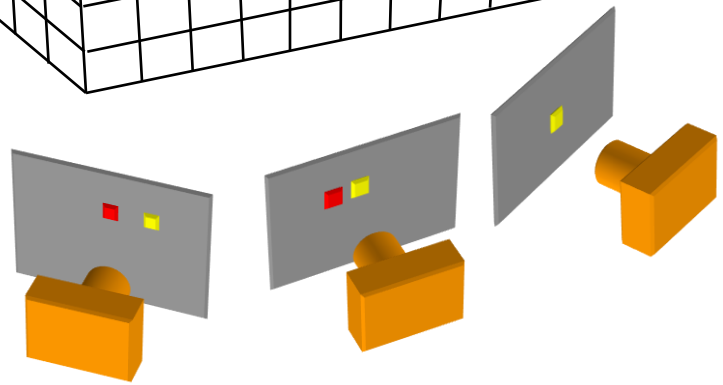
Goal: **Determine occupancy, "color" of points in V**

Discrete formulation: Voxel Coloring

Discretized
Scene Volume



Input Images
(Calibrated)



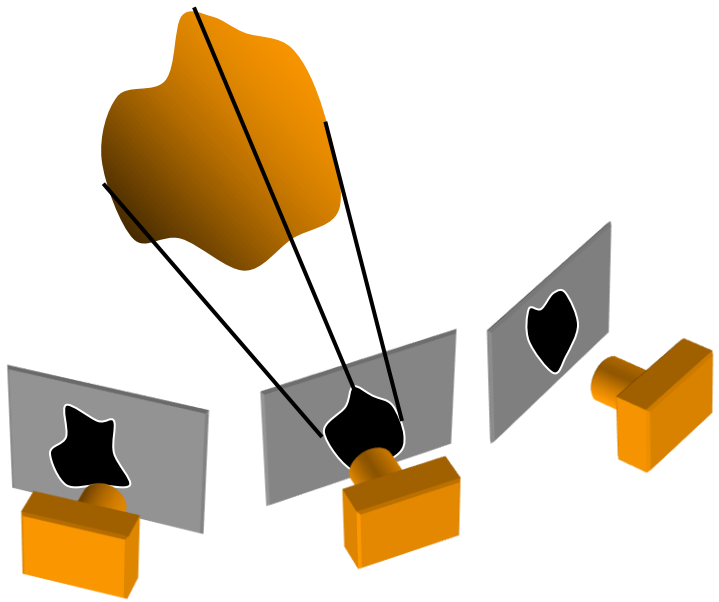
Goal: **Assign RGB (or empty) values to voxels in V**
photo-consistent **with images**

Voxel coloring solutions

1. Two colors (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) or
     > W. Matusik, C. Buehler, R. Raskar, L. McMillan, and S. J. Gortler, *Image-Based Visual Hulls*, SIGGRAPH 2000 (pdf 1.6 MB)
2. Many colors, viewpoint constraints
   - Voxel coloring algorithm [Seitz & Dyer 97]
3. General Case
   - Space carving [Kutulakos & Seitz 98]

Reconstruction from Silhouettes

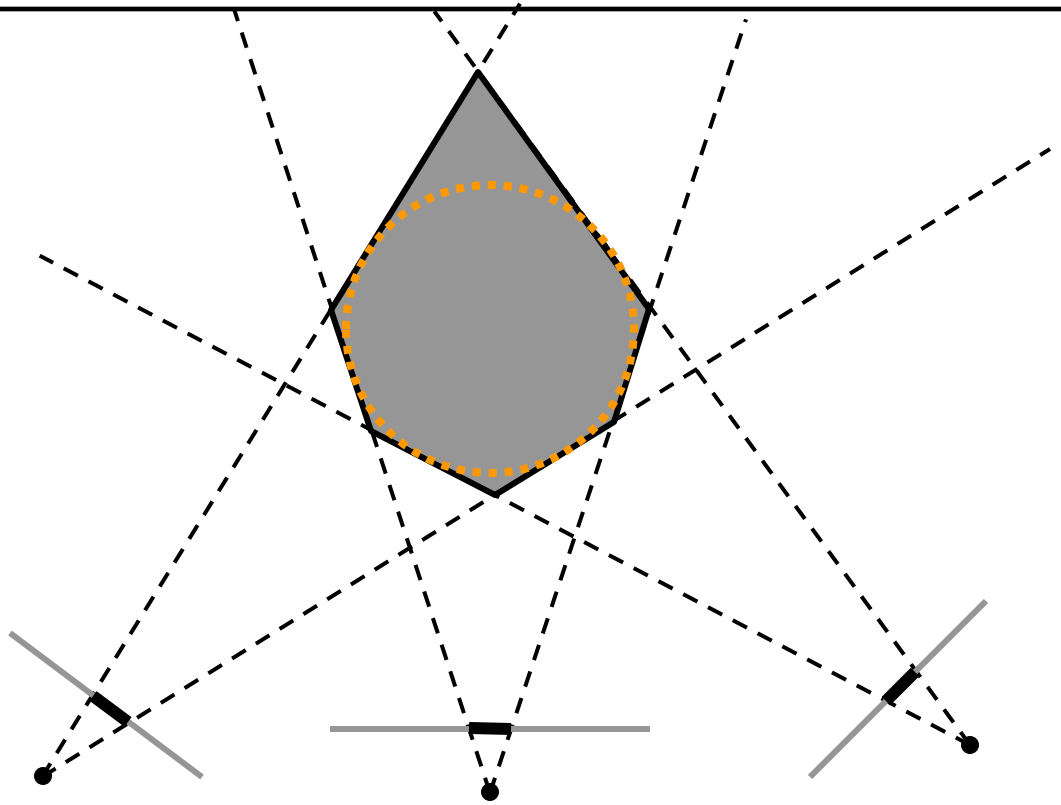
Binary Images →



Approach:

- *Backproject* each silhouette
- Intersect backprojected volumes

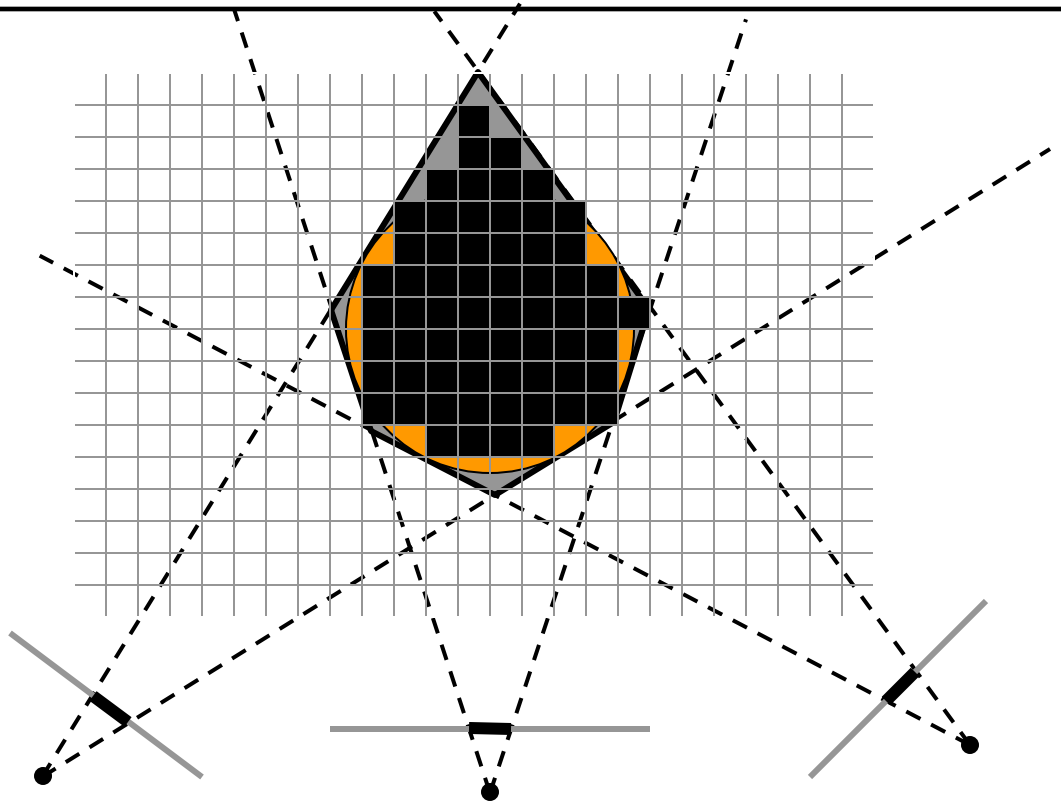
Volume intersection



Reconstruction Contains the True Scene

- But is generally not the same
- 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 in silhouette for every image

- O( ), 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] or interval techniques [Matusik 2000]

Cons

- No concavities
- Reconstruction is not photo-consistent

  (though it does agree with the silhouettes)
- Requires identification of silhouettes

Voxel Coloring Solutions

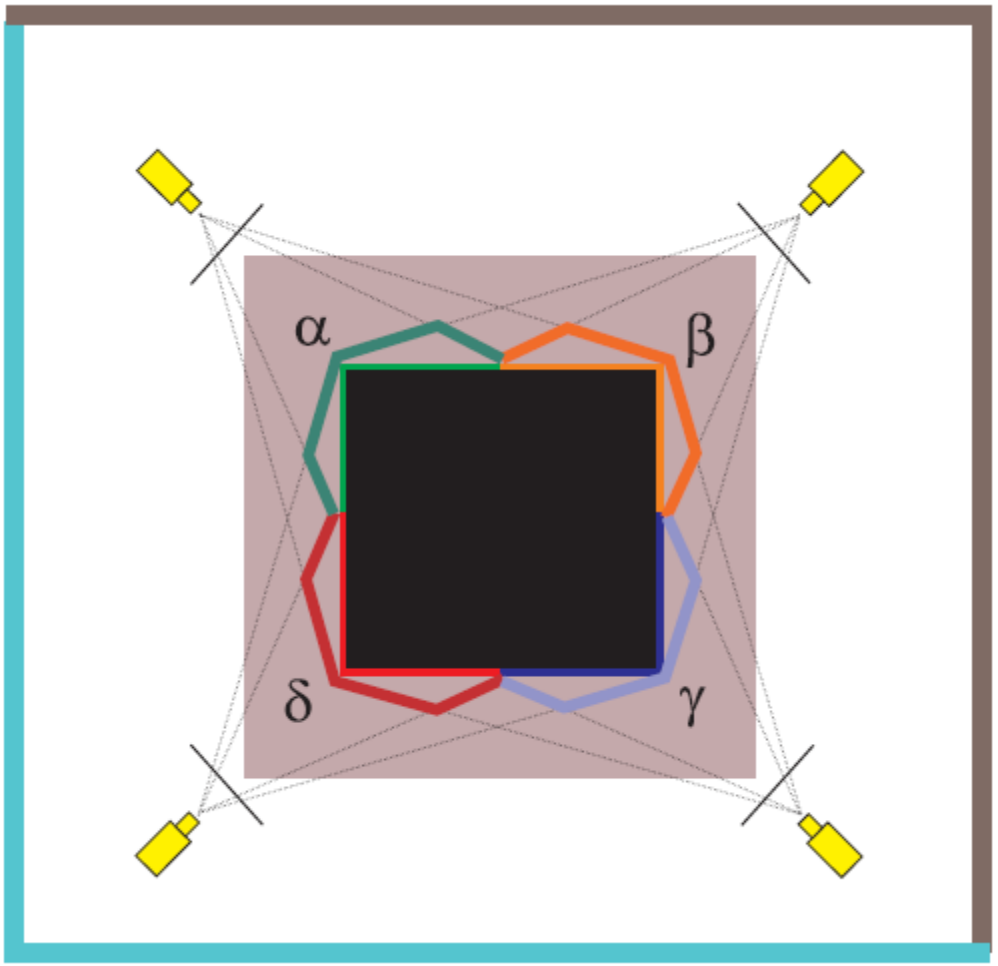
1. Two colors (silhouettes)
   - Volume intersection [Baumgart 1974]

2. Many colors, 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]

Problem: non-uniqueness

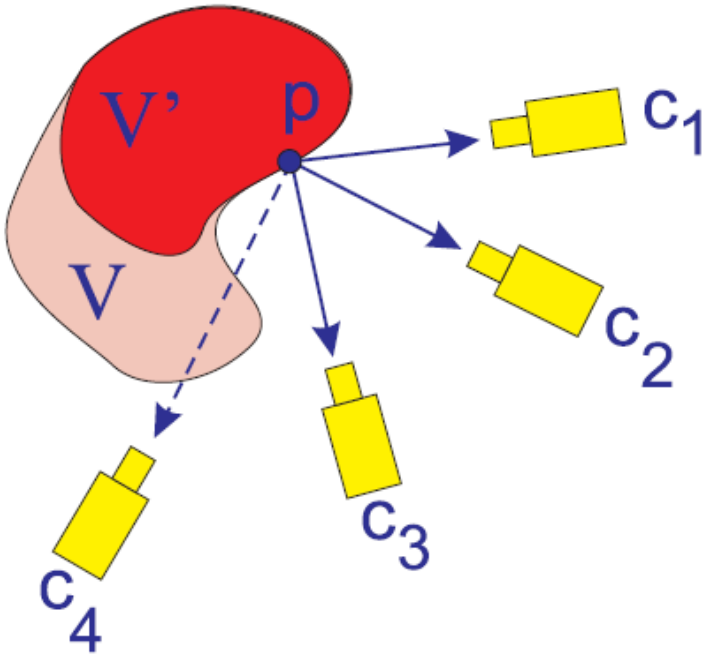
- Many scenes could give rise to the same images.



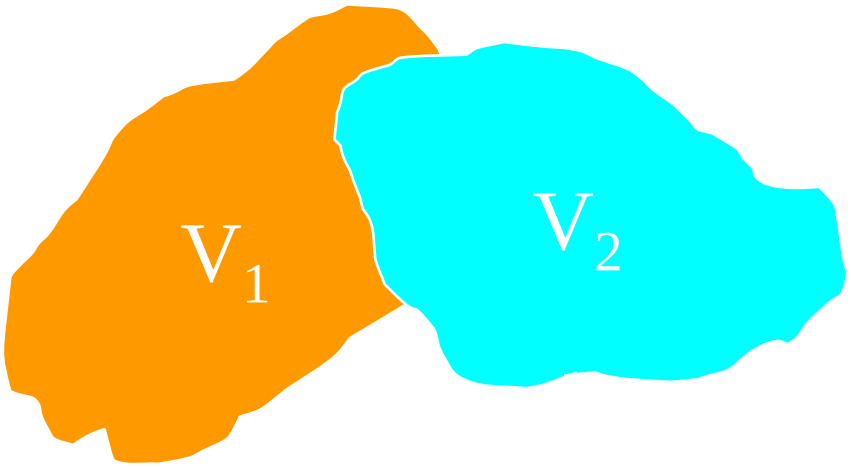
Photoconsistency and Visibility

A point p is *photoconsistent* with respect to volume V if its projection is the same color in all cameras in which it is visible.

Claim: if p is not photoconsistent with respect to V, and V' is a subset of V, then p is not photoconsistent with respect to V'.



Question

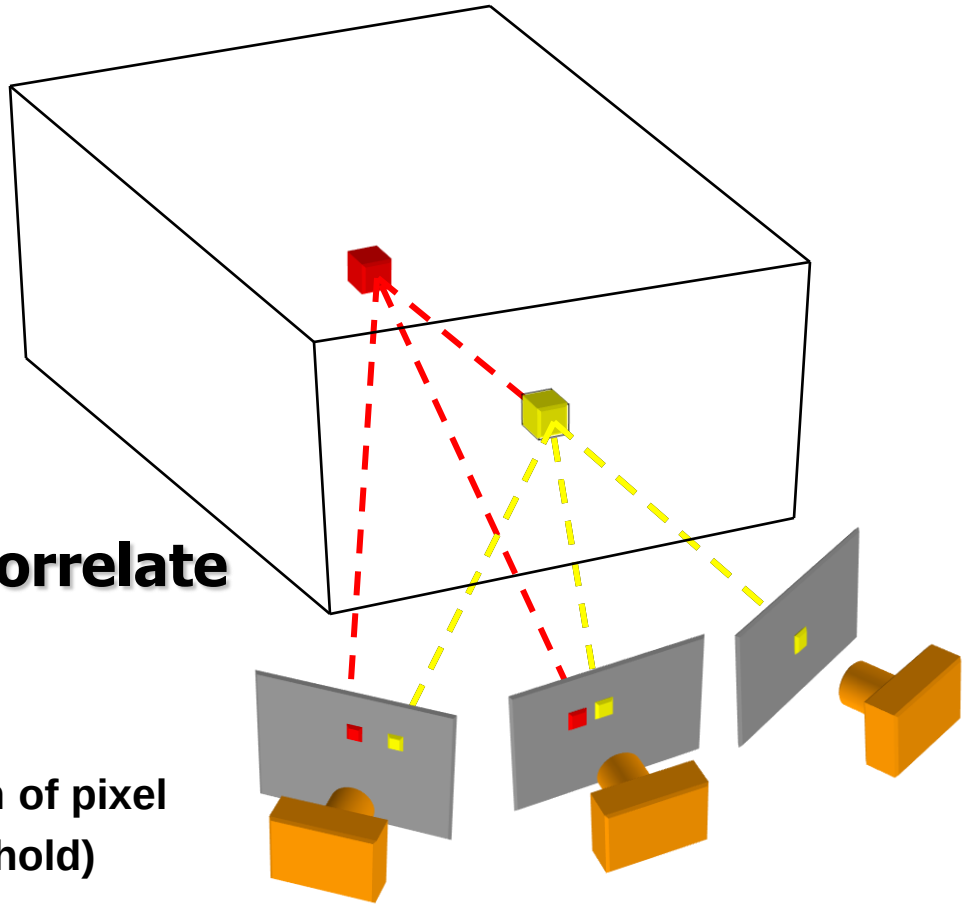


If all points on V_1 are photoconsistent,

and all points on V_2 are photoconsistent,

then…

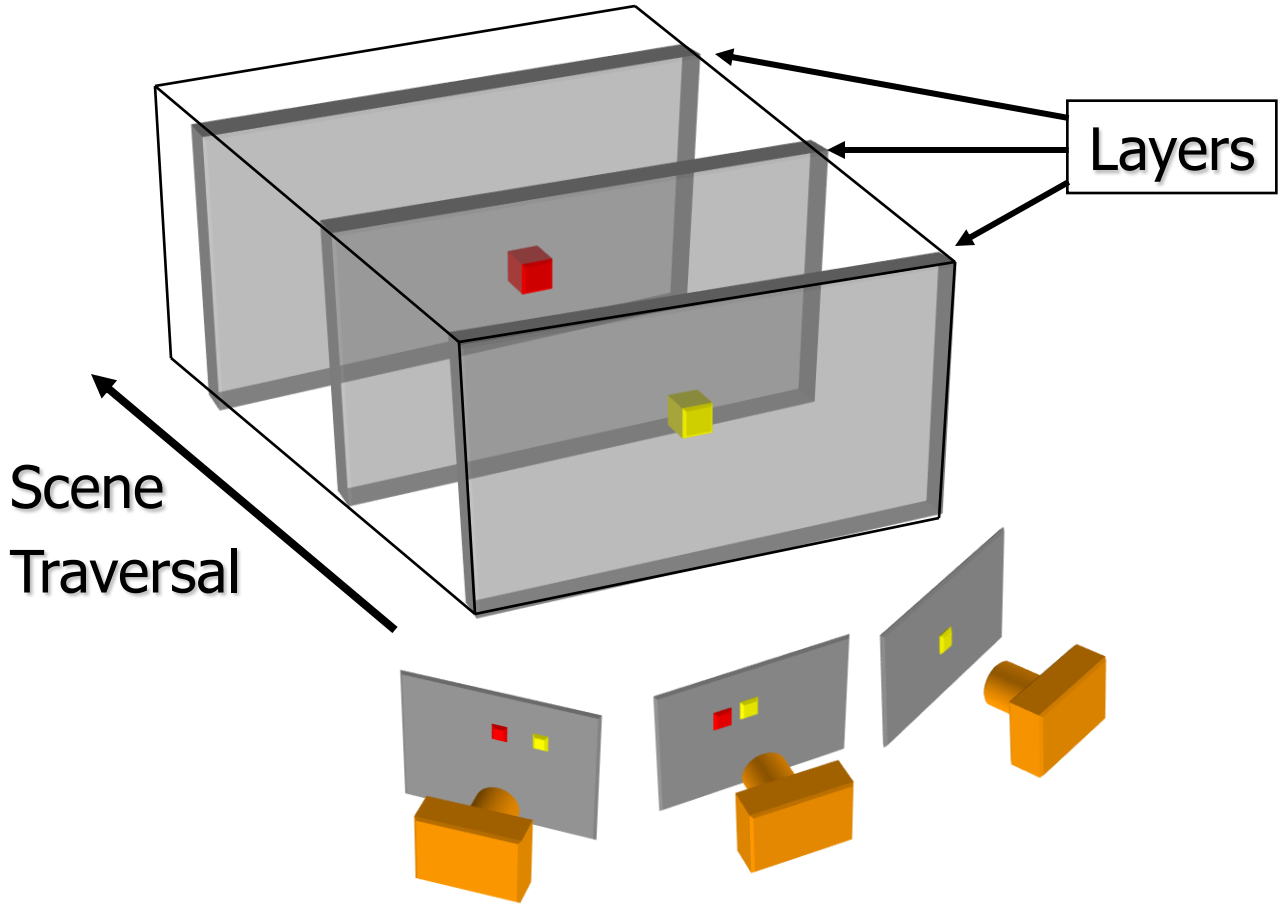
Space Carving Approach



1. **Choose voxel**
2. **Project and correlate**
3. **Remove if inconsistent**
   (standard deviation of pixel colors above threshold)

Problem: in what order should we look at the voxels?

Depth Ordering: visit occluders first!



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

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

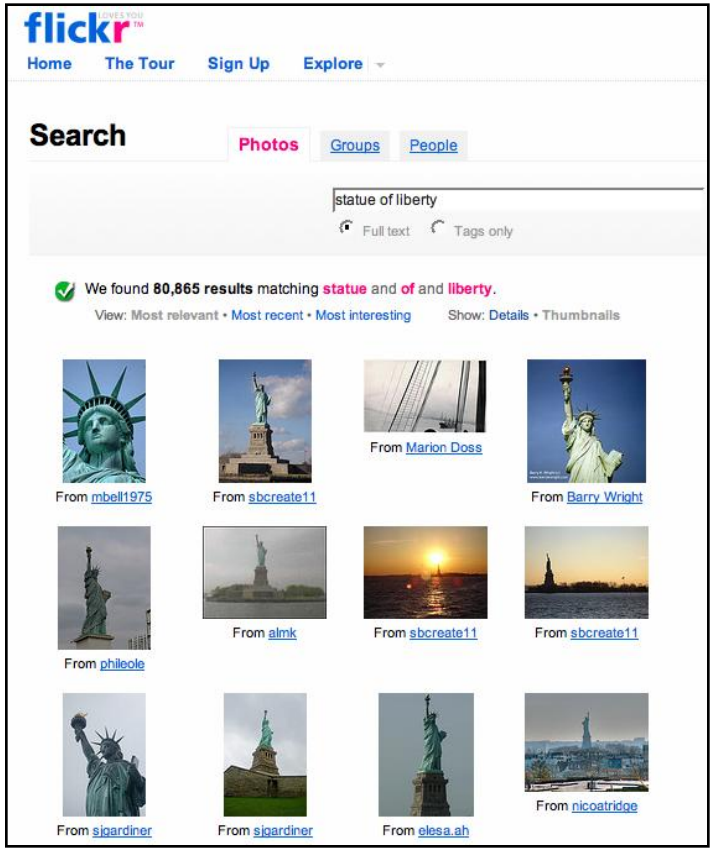
Handling Visiblity: Two Approaches

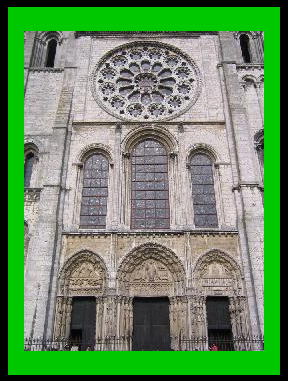
- **Treat occlusions as outliers**

- Model occlusions geometrically

Multi-view stereo from Internet Collections

[Goesele, Snavely, Curless, Hoppe, Seitz, ICCV 2007]



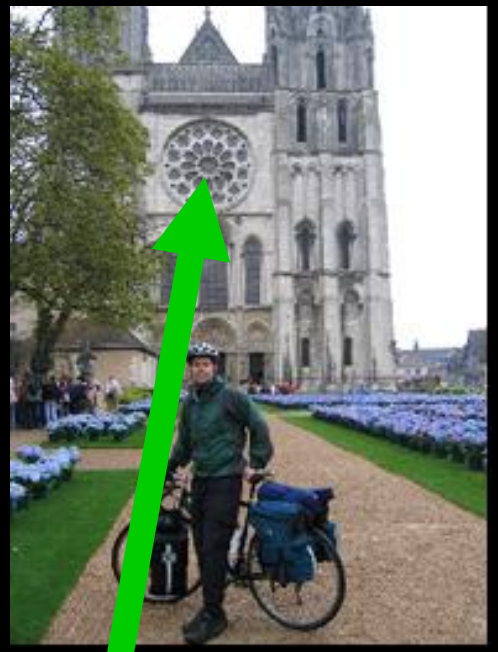





4 best neighboring views

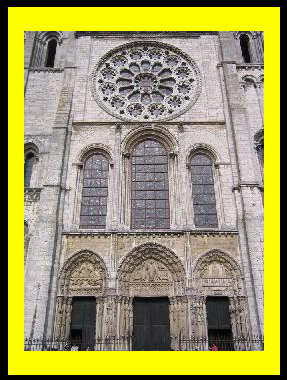





reference view

Local view selection

- Automatically select neighboring views for each point in the image
- Desiderata: good matches AND good baselines


4 best neighboring views

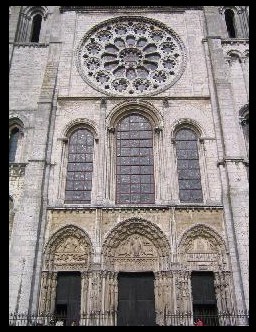




reference view

Local view selection

- Automatically select neighboring views for each point in the image
- Desiderata: good matches AND good baselines



4 best neighboring views





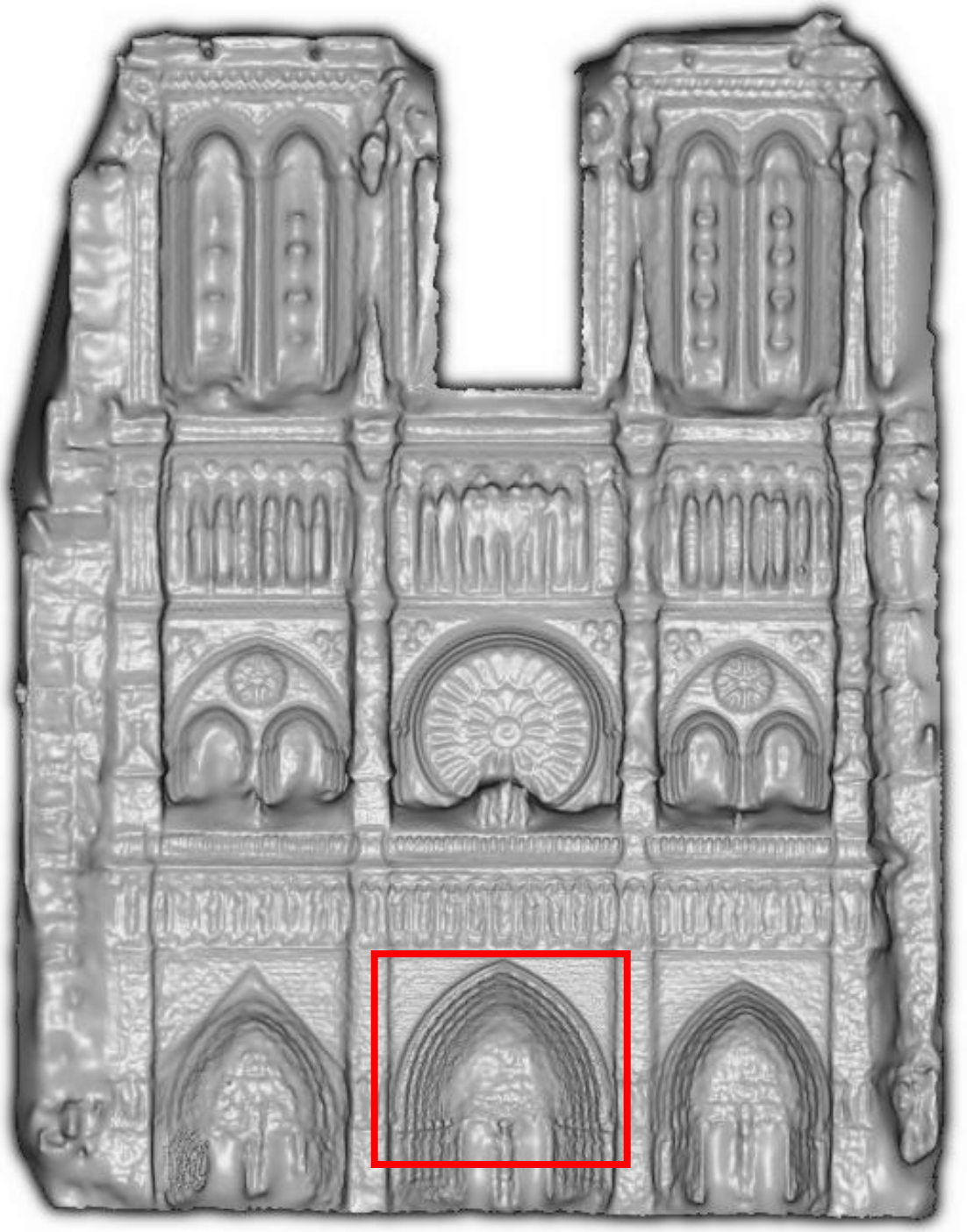

reference view

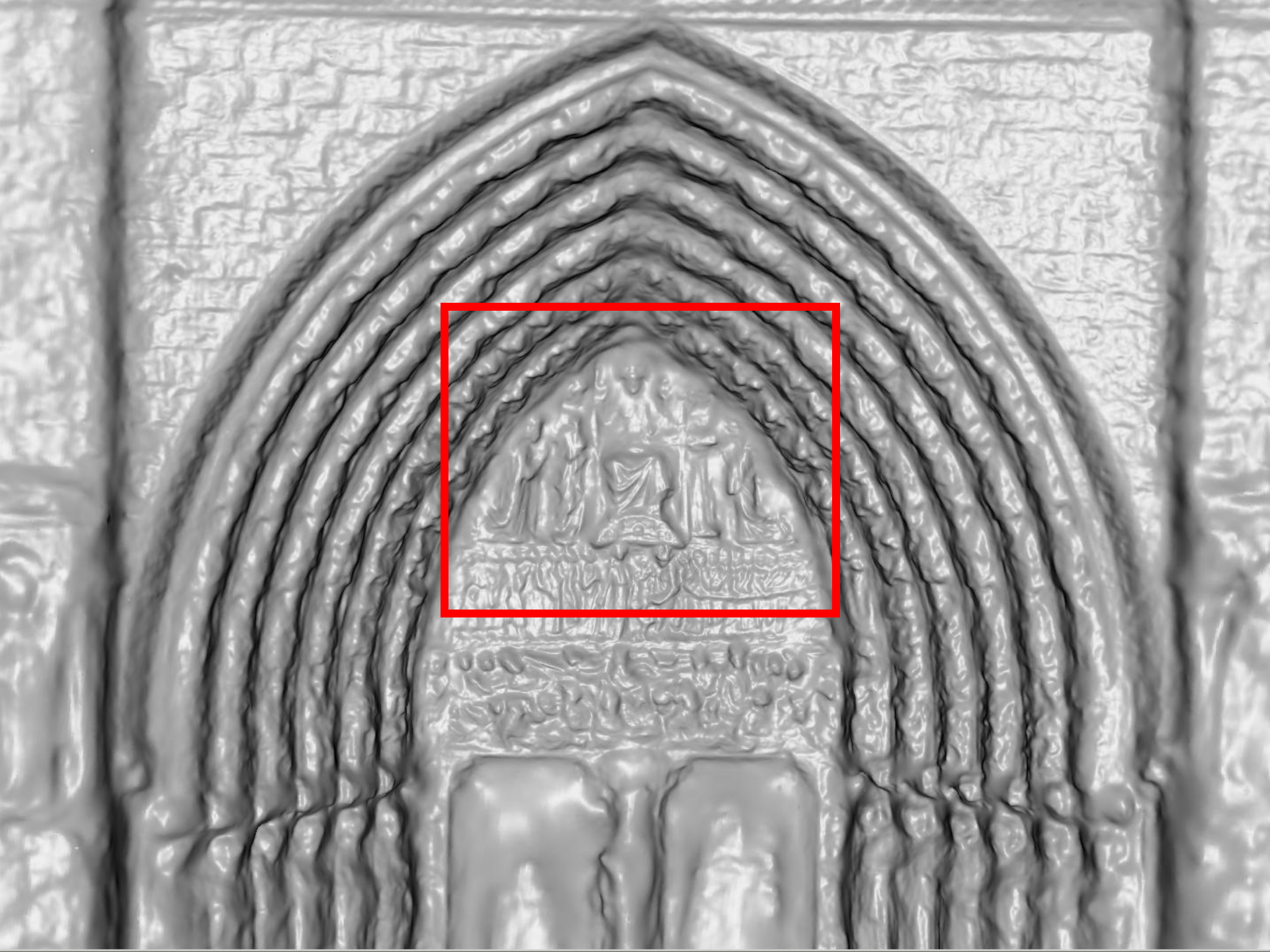
Local view selection

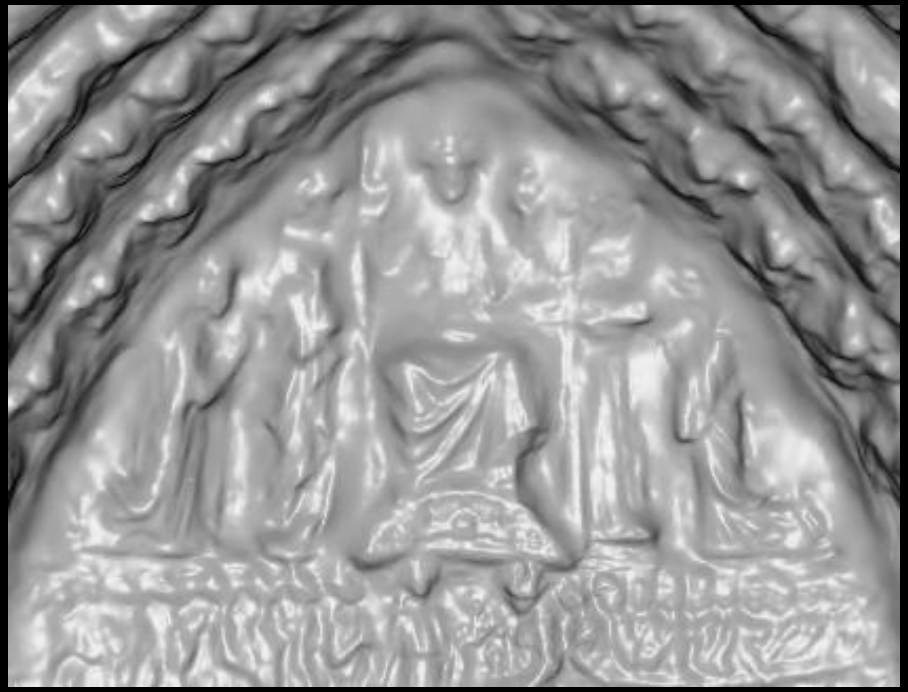
- Automatically select neighboring views for each point in the image
- Desiderata: good matches AND good baselines

Notre Dame de Paris

653 images
313 photographers



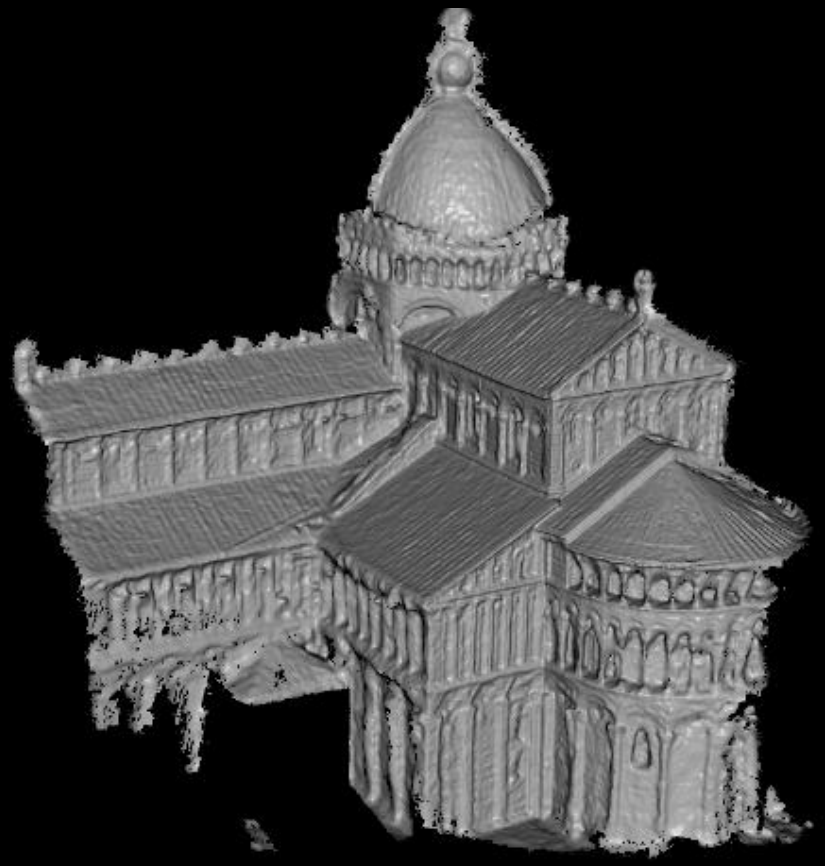




129 *Flickr* images taken by 98 photographers

merged model of Venus de Milo

56 *Flickr* images taken by 8 photographers



merged model of Pisa Cathedral

Some other work…

Unconstrained camera viewpoints

- Space carving [Kutulakos & Seitz 98]

Evolving a surface

- Level sets [Faugeras & Keriven 98]
- More recent work by Pons et al.

Global optimization

- Graph cut approaches
  - > [Kolmogoriv & Zabih, ECCV 2002]
  - > [Vogiatzis et al., PAMI 2007]

Modeling shiny (and other reflective) surfaces

- e.g., Zickler et al., *Helmholtz Stereopsis*

Combining photoconsistency and silhouettes

- [Kolev & Cremers 2008]