Reconstruction

CSE 576

Ali Farhadi

Several slides from Steve Seitz, Larry Zitnick, Lana Lazebnik, Carlos Hernández, George Vogiatzis,Yasutaka Furukaw

3d model

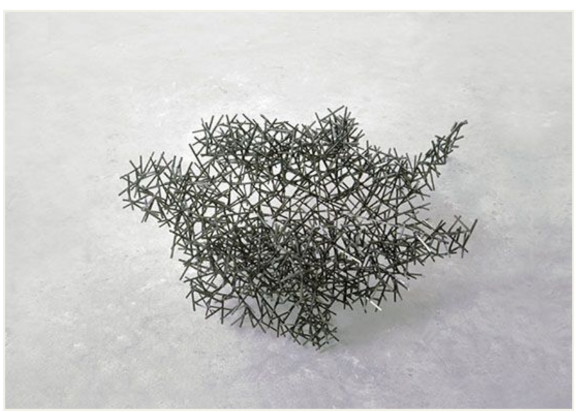
- “Digital copy” of real object
- Allows us to
  - Inspect details of object
  - Measure properties
  - Reproduce in different material
- Many applications
  - Cultural heritage preservation
  - Computer games and movies
  - City modelling
  - E-commerce

Applications: cultural heritage

SCULPTEUR European project

Applications: art

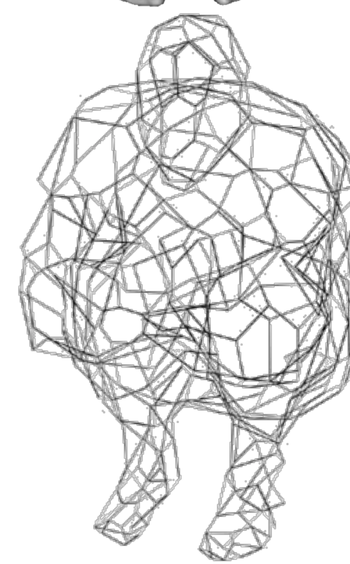
Block Works Precipitate III 2004
Mild steel blocks 80 x 46 x 66 cm



Domain Series Domain VIII Crouching
1999 *Mild steel bar* 81 x 59 x 63 cm

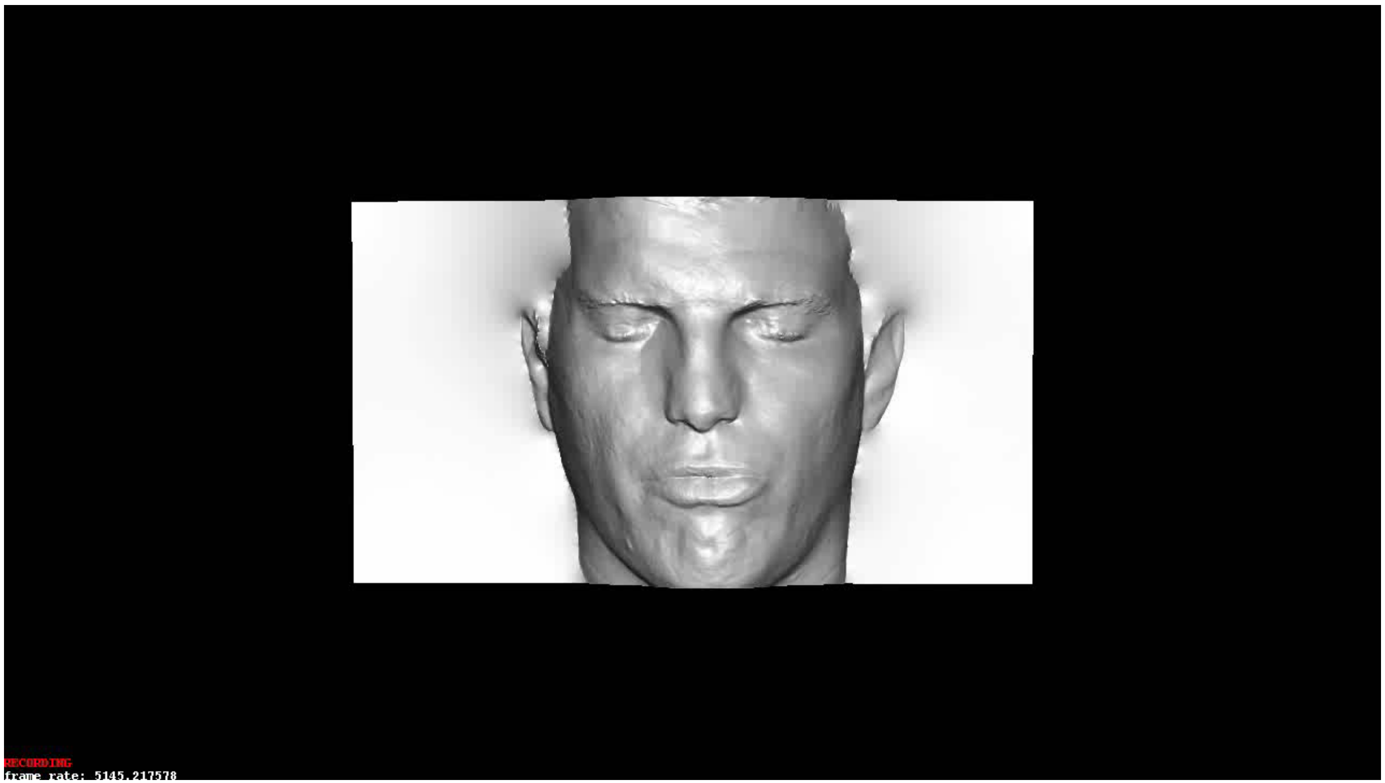
Applications: structure engineering

BODY / SPACE / FRAME, Antony Gormley, Lelystad, Holland

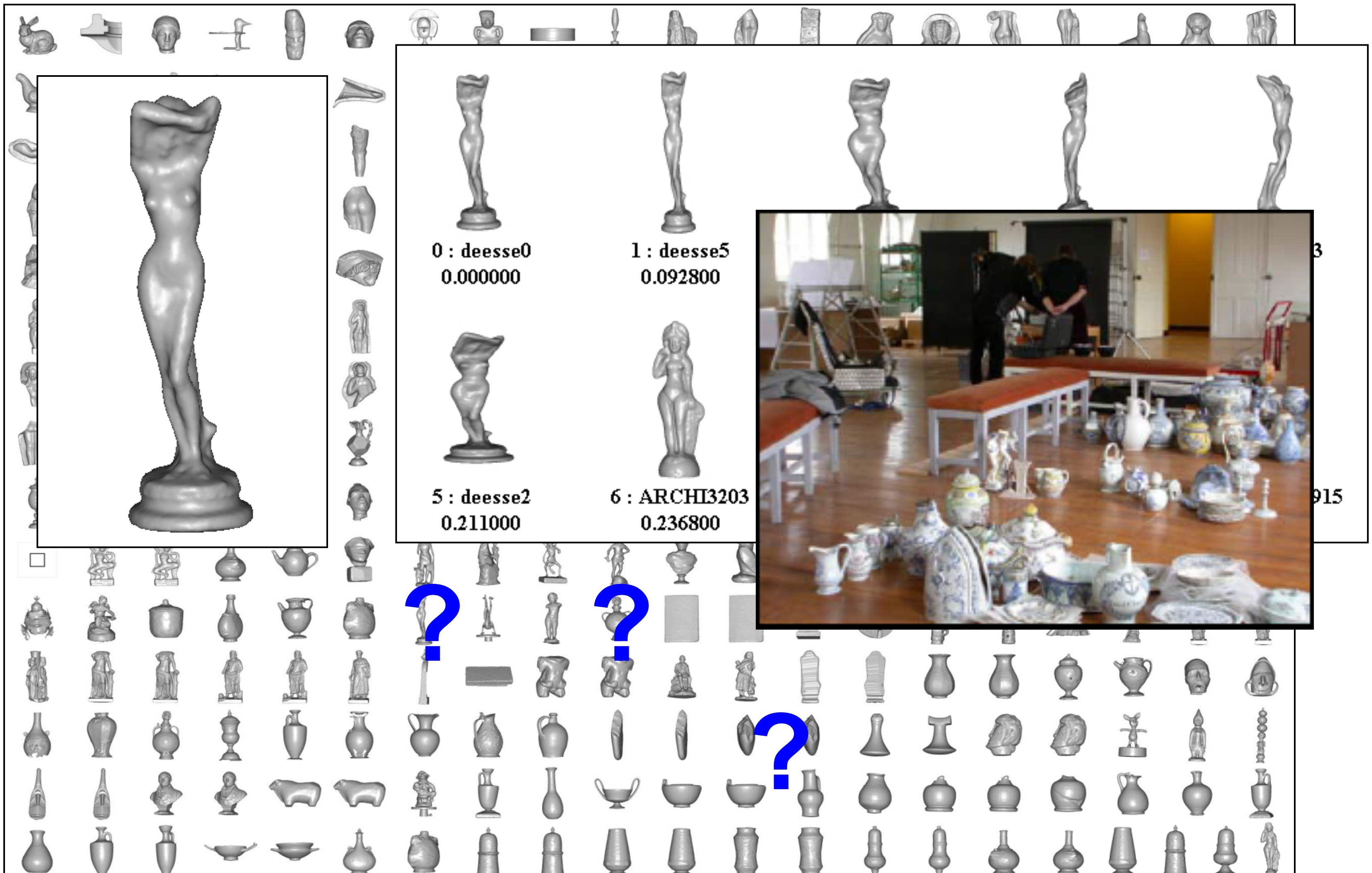


Applications: art

Applications: computer games



Applications: 3D indexation

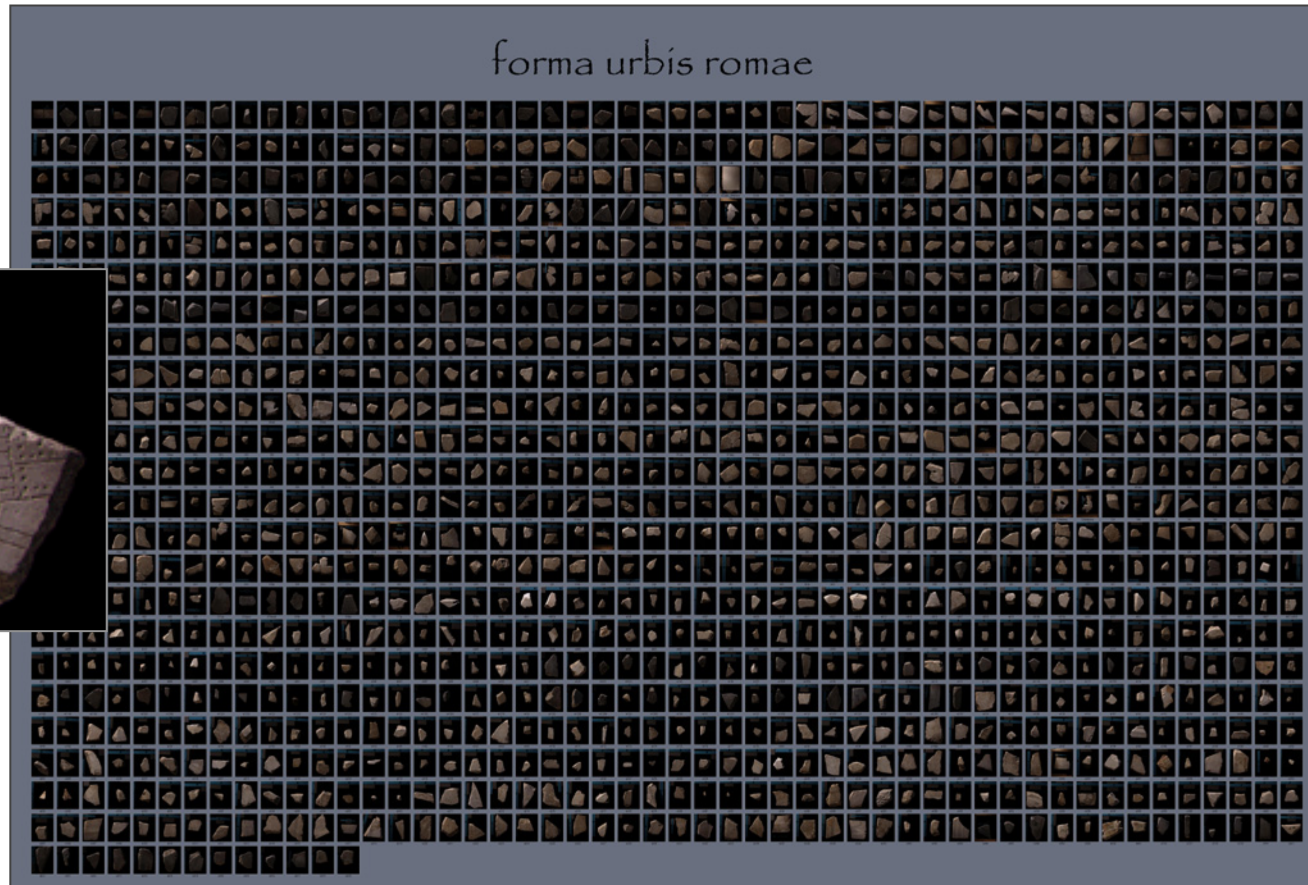


Applications: archaeology

- "forma urbis romae" project

Fragments of the City: Stanford's Digital Forma Urbis Romae Project
David Koller, Jennifer Trimble, Tina Najbjerg, Natasha Gelfand, Marc Levoy
Proc. Third Williams Symposium on Classical Architecture, Journal of Roman Archaeology supplement, 2006.

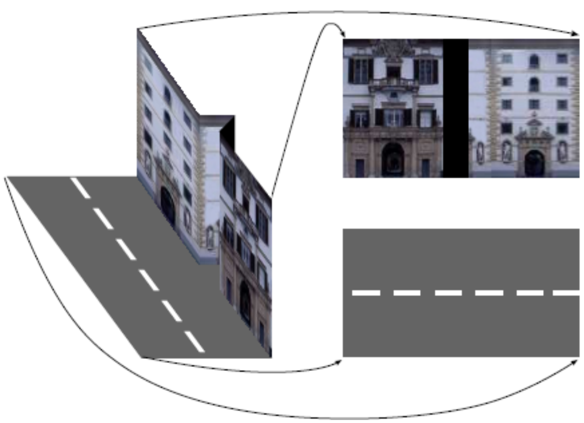
1186 fragments



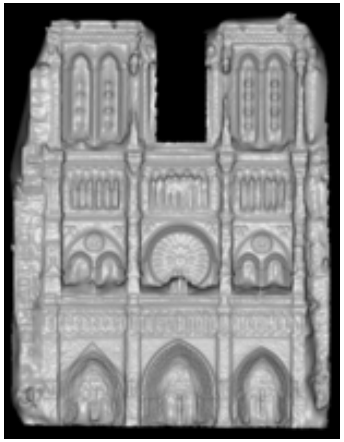
Applications: large scale modelling

[Furukawa10]

[Pollefeys08]



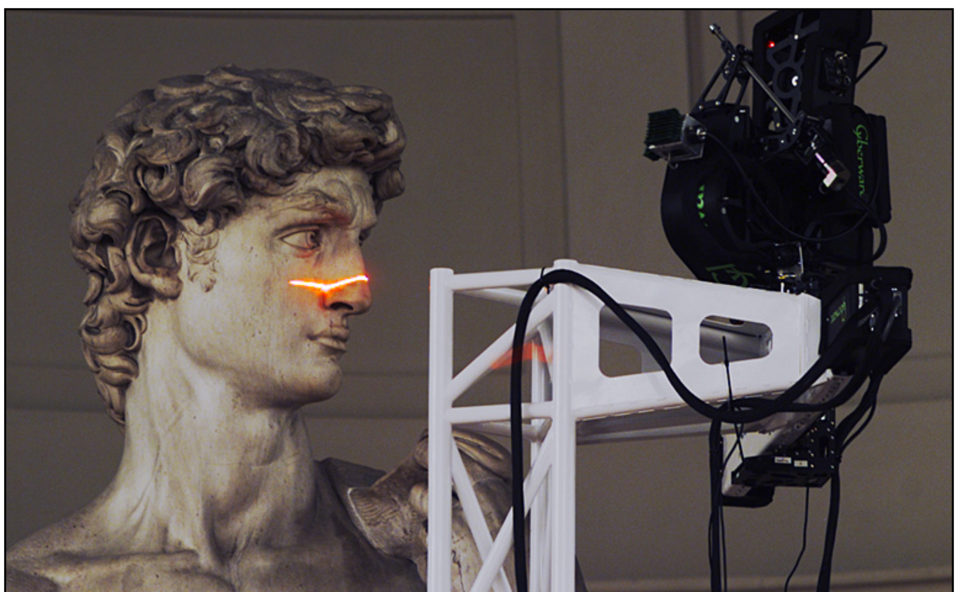
[Cornelis08]



[Goesele07]

Scanning technologies

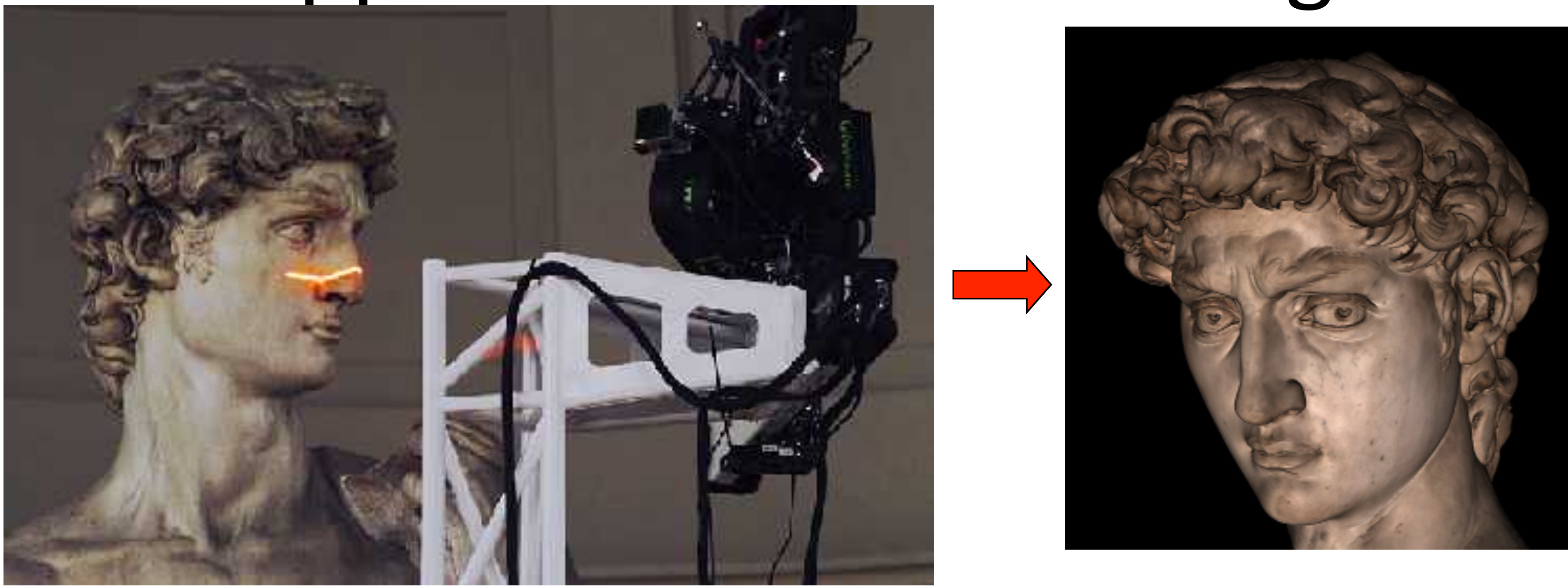
- Laser scanner, coordinate measuring machine
  - Very accurate
  - Very Expensive
  - Complicated to use



Minolta

Contura CMM

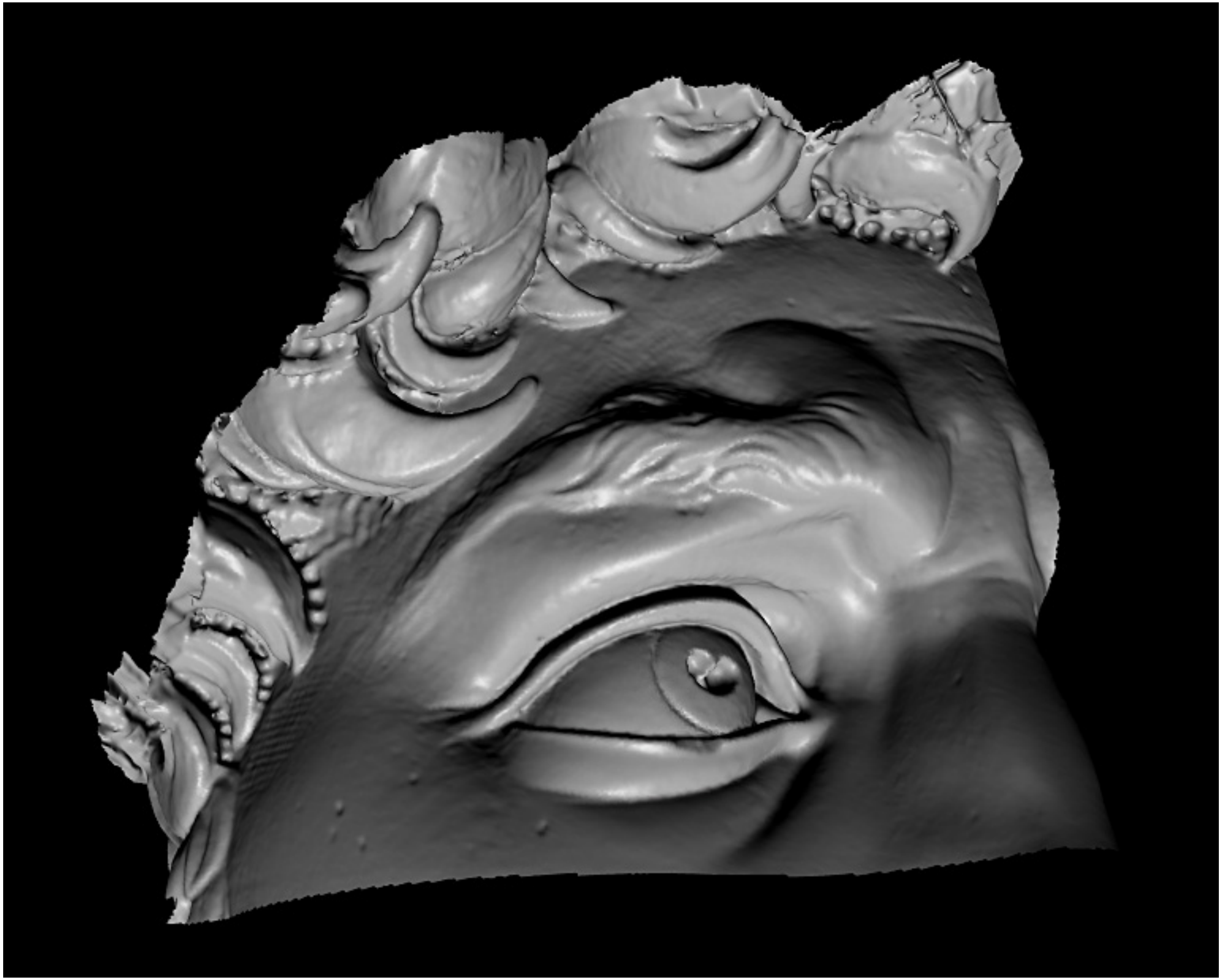
Applications: 3D Scanning

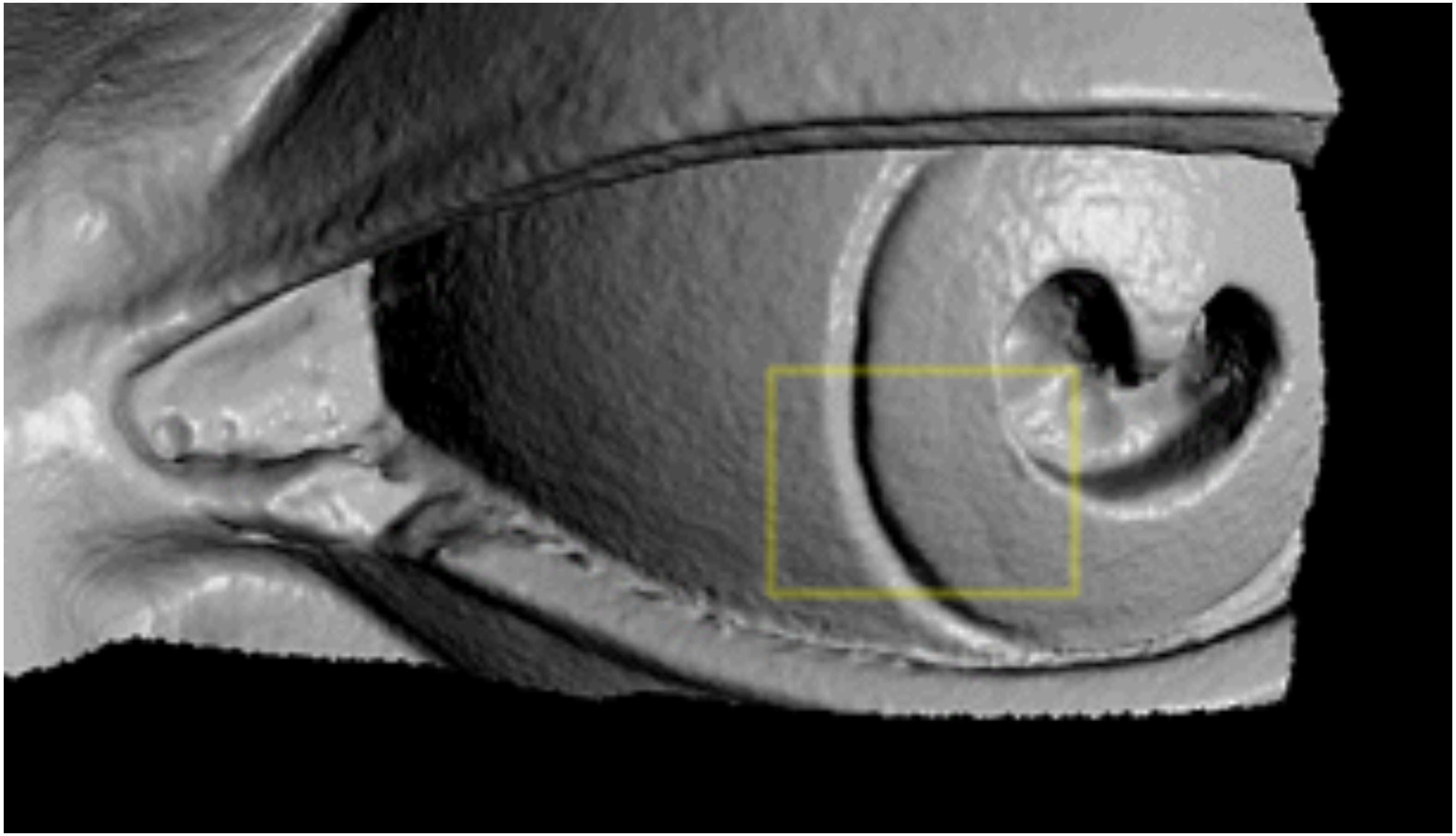


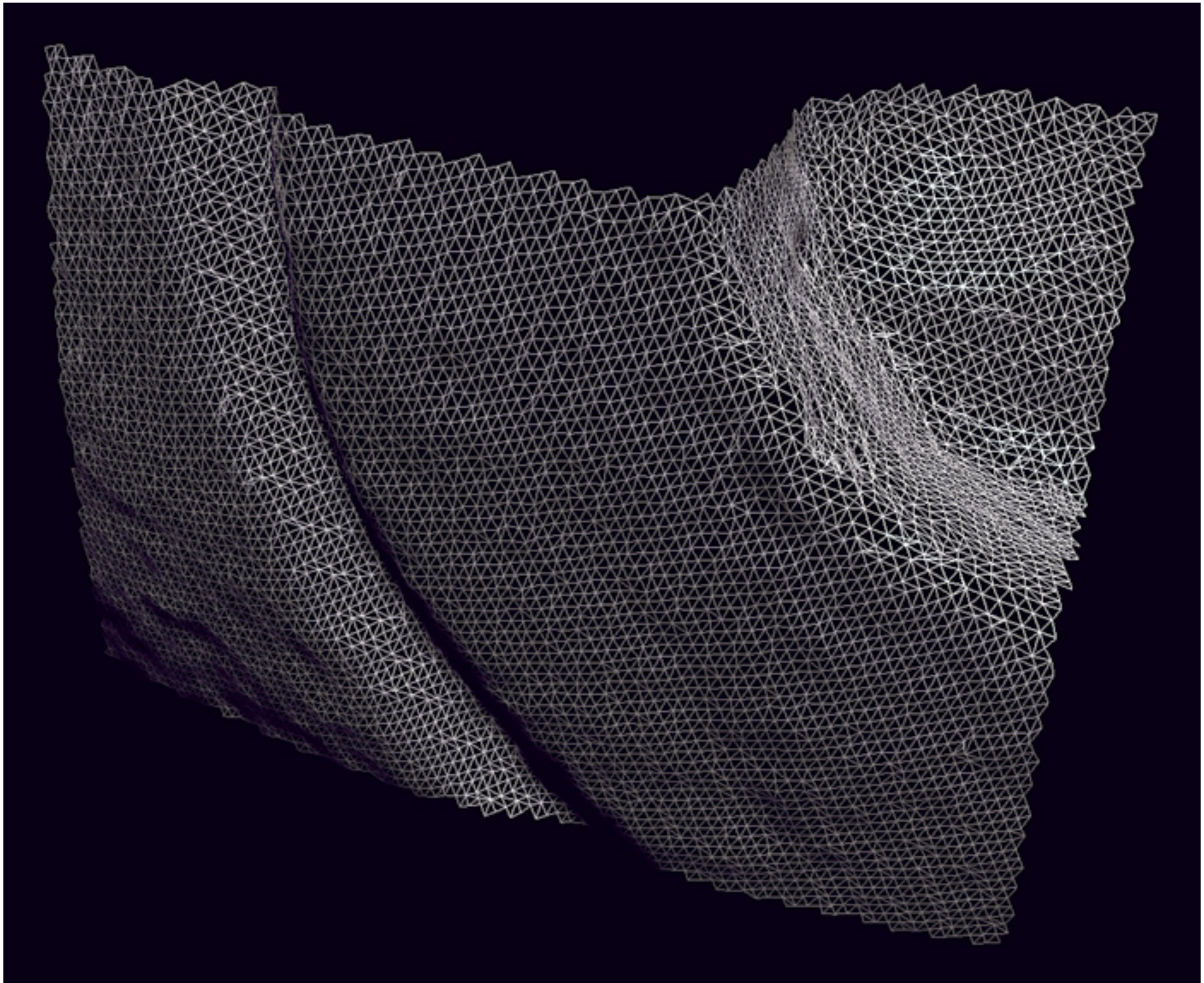
Scanning Michelangelo's *"The David"*

- [The Digital Michelangelo Project](http://graphics.stanford.edu/projects/mich/)
    - http://graphics.stanford.edu/projects/mich/
- 2 BILLION polygons, accuracy to .29mm

The Digital Michelangelo Project, Levoy et al.

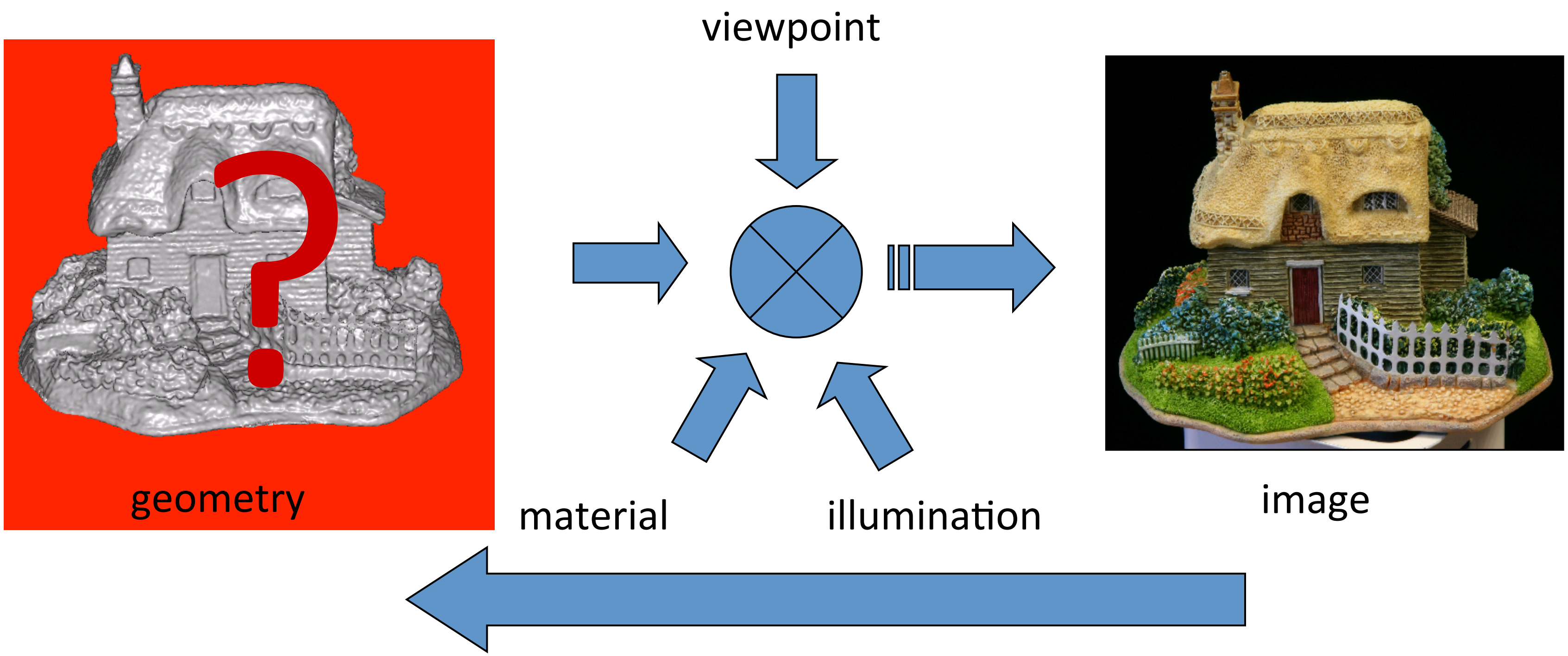






3d shape from photographs

"Estimate a 3d shape that would generate the input photographs given the same material, viewpoints and illumination"



3D shape from photographs

"Estimate a 3d shape that would generate the input photographs given the same material, viewpoints and illumination"

3d shape from photographs

Appearance strongly depends on the material and lighting

textured

rigid

deforming

textureless

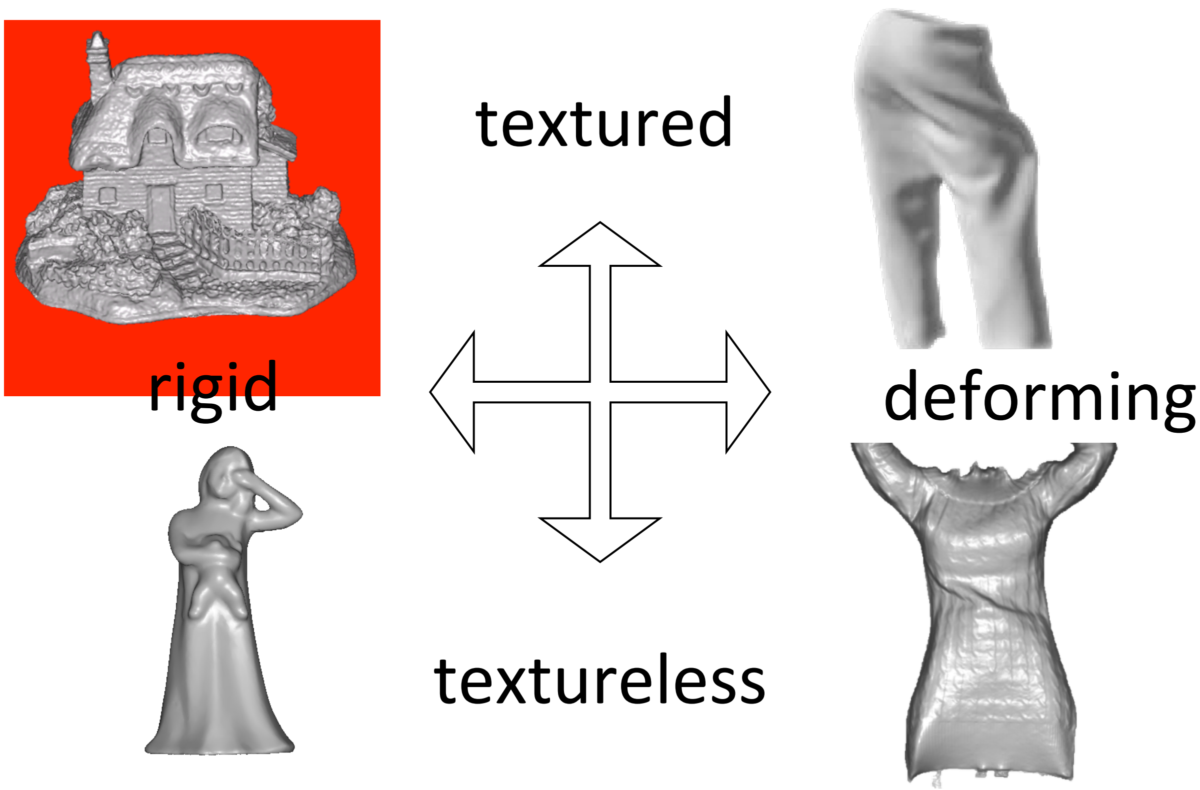
[redacted]

3d shape from photographs

Appearance strongly depends on the material and lighting

$\Downarrow$

No single algorithm exists dealing with **any** type of scene



3d shape from photographs

Photograph based 3d reconstruction is:

- ✓ practical
- ✓ fast
- ✓ non-intrusive
- ✓ low cost
- ✓ Easily deployable outdoors
- ✗ “low” accuracy
- ✗ Results depend on material properties

Reconstruction

- Generic problem formulation: given several images of the same object or scene, compute a representation of its 3D shape



Reconstruction

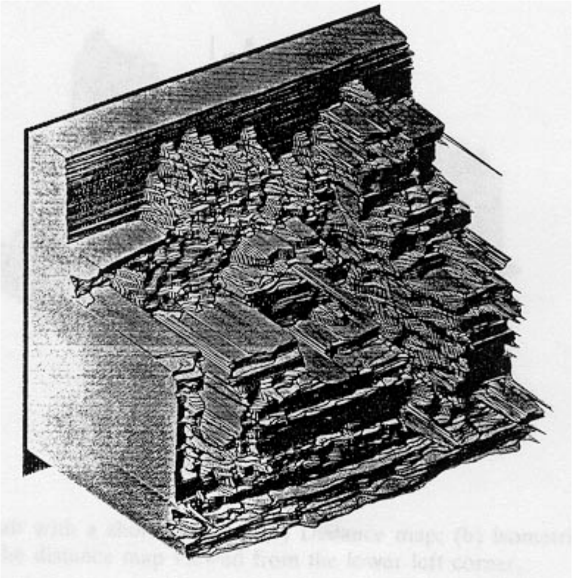
- Generic problem formulation: given several images of the same object or scene, compute a representation of its 3D shape
- “Images of the same object or scene”
  - Arbitrary number of images (from two to thousands)
  - Arbitrary camera positions (camera network or video sequence)
  - Calibration may be initially unknown
- “Representation of 3D shape”
  - Depth maps
  - Meshes
  - Point clouds
  - Patch clouds
  - Volumetric models
  - Layered models

Multiple-baseline stereo

I1

I2

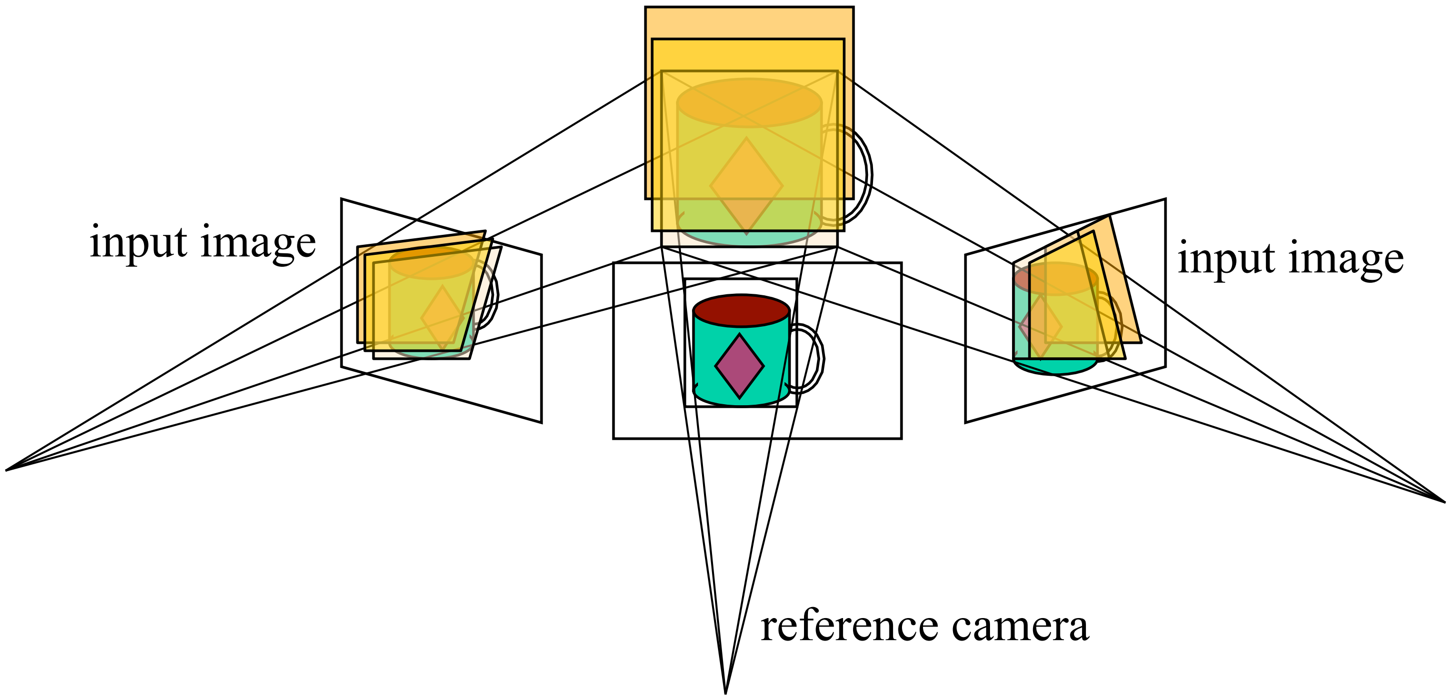
I10



M. Okutomi and T. Kanade, "A Multiple-Baseline Stereo System," IEEE Trans. on Pattern Analysis and Machine Intelligence, 15(4):353-363 (1993).

Plane Sweep Stereo

- Choose a reference view
- Sweep family of planes at different depths with respect to the reference camera

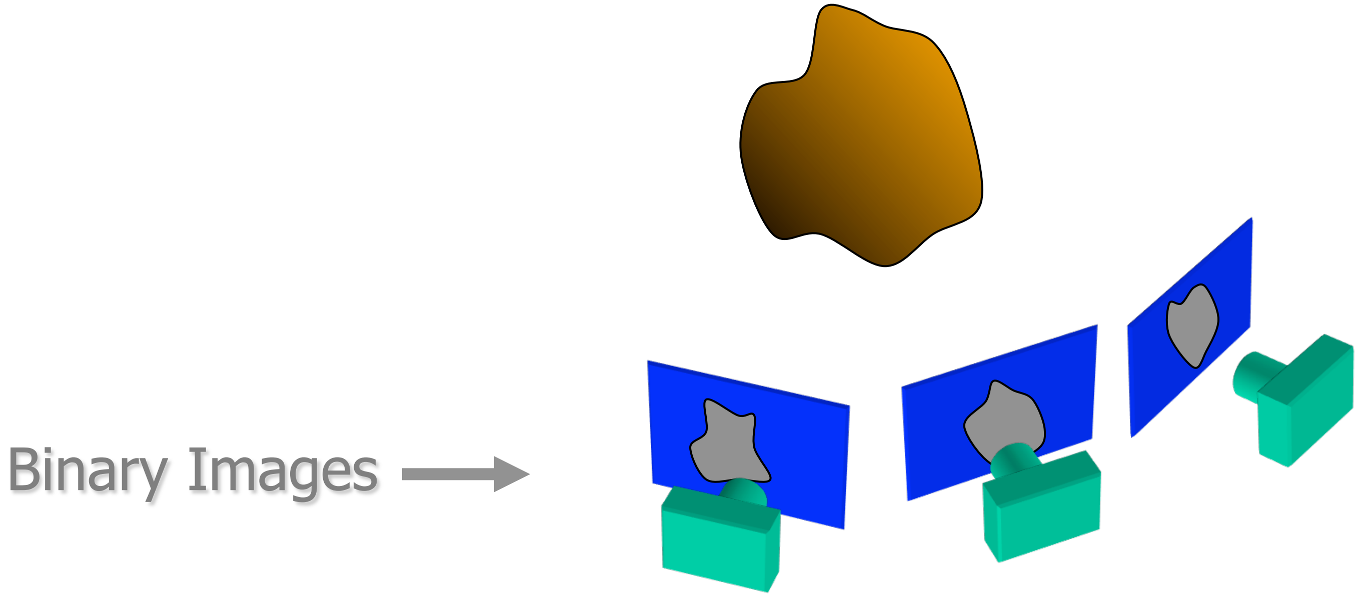


Each plane defines a homography warping each input image into the reference view

R. Collins. [A space-sweep approach to true multi-image matching.](#) CVPR 1996.

Reconstruction from Silhouettes

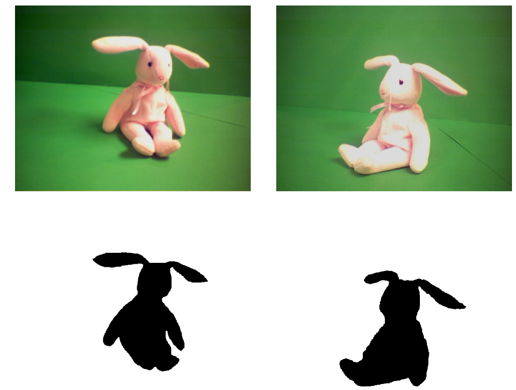
- The case of binary images: a voxel is photo-consistent if it lies inside the object's silhouette in all views

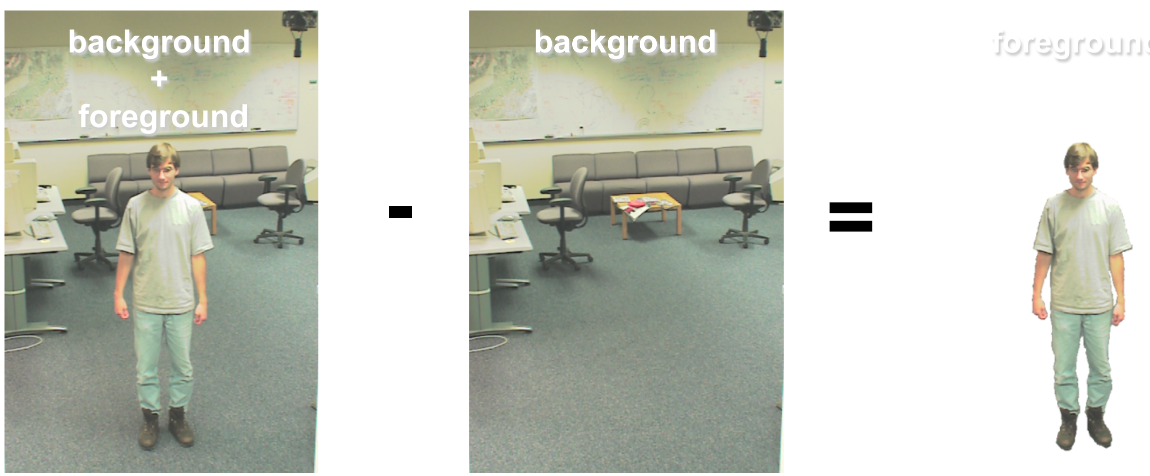


Use silhouettes

Can be computed robustly

Can be computed efficiently

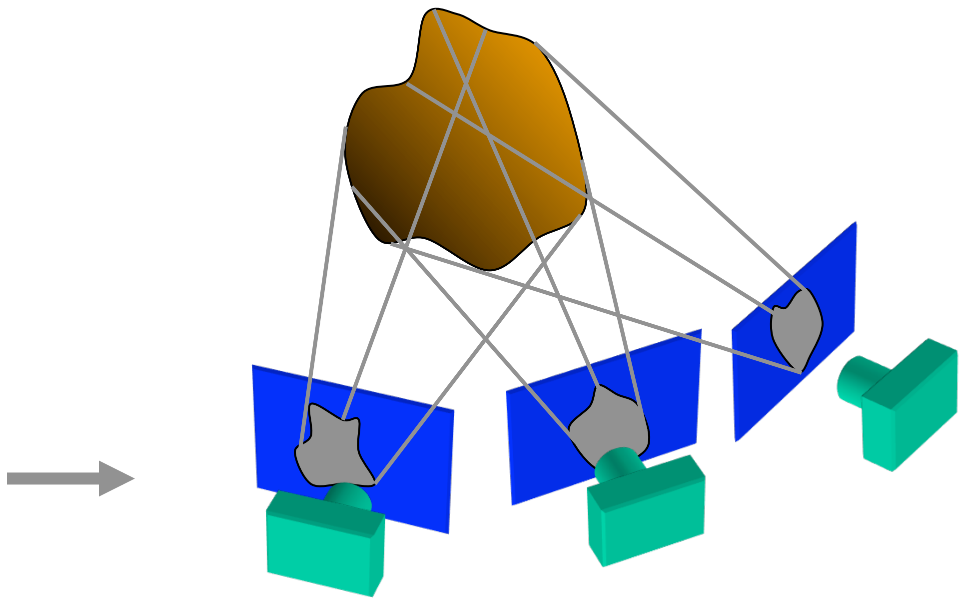




Reconstruction from Silhouettes

- The case of binary images: a voxel is photo-consistent if it lies inside the object's silhouette in all views

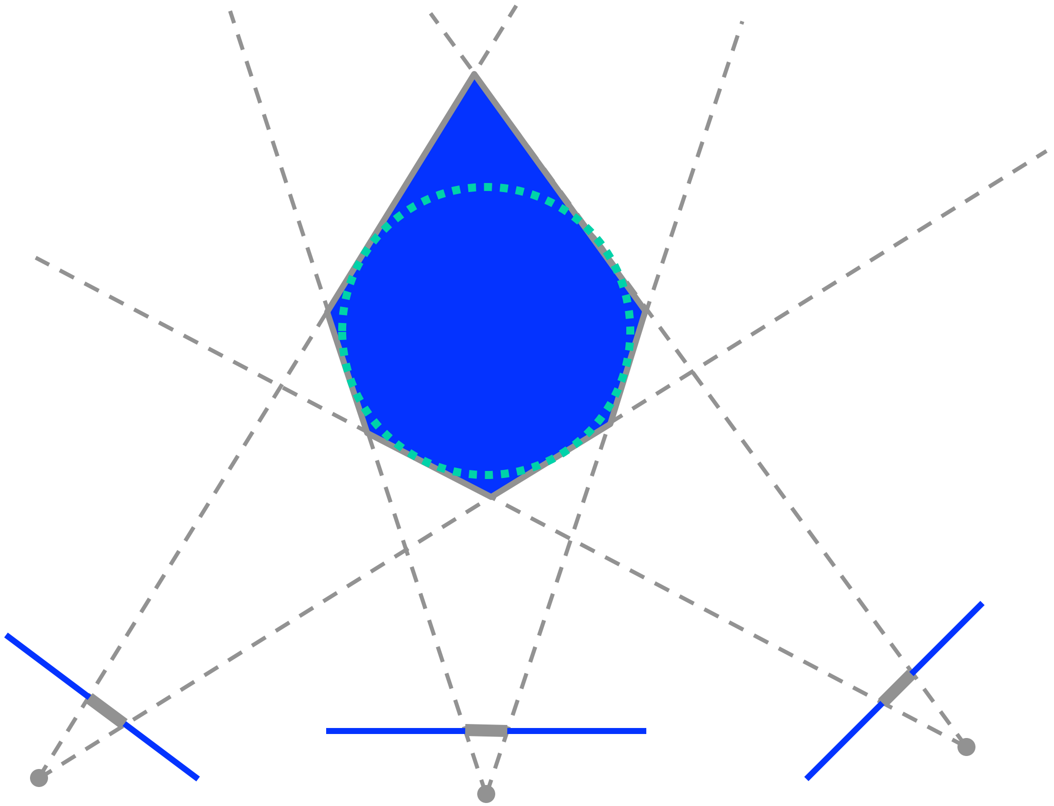
Binary Images



Finding the silhouette-consistent shape (*visual hull*):

- *Backproject* each silhouette
- Intersect backprojected volumes

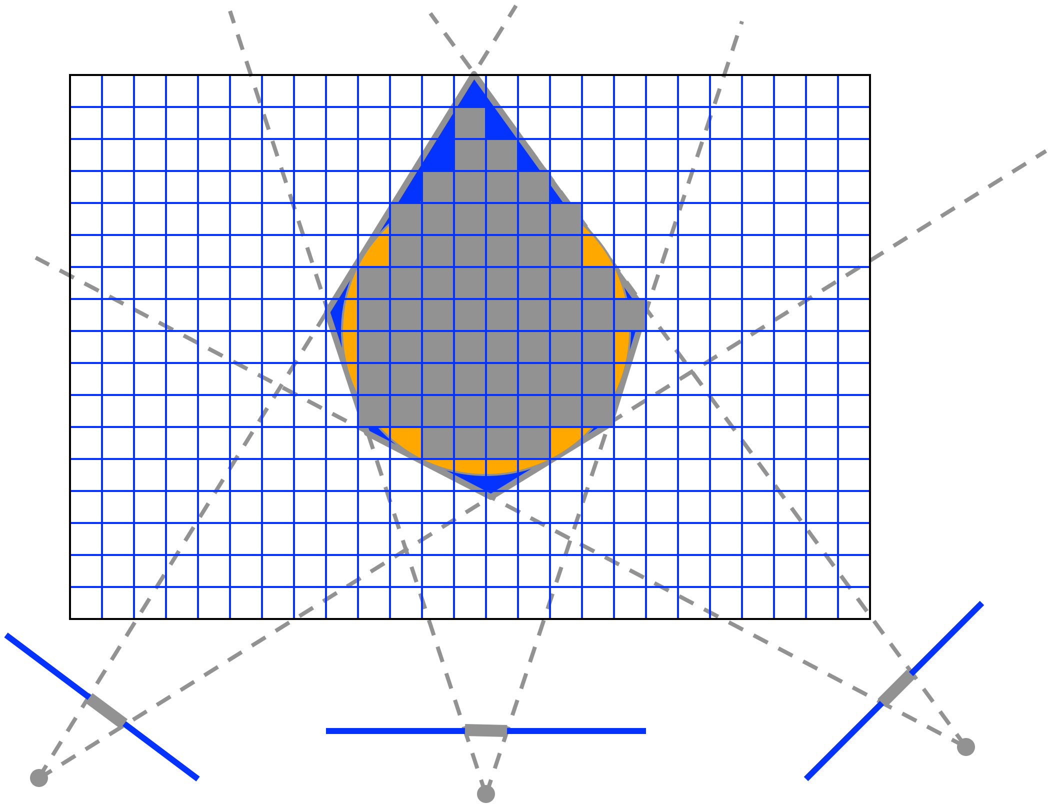
Volume intersection



Reconstruction Contains the True Scene

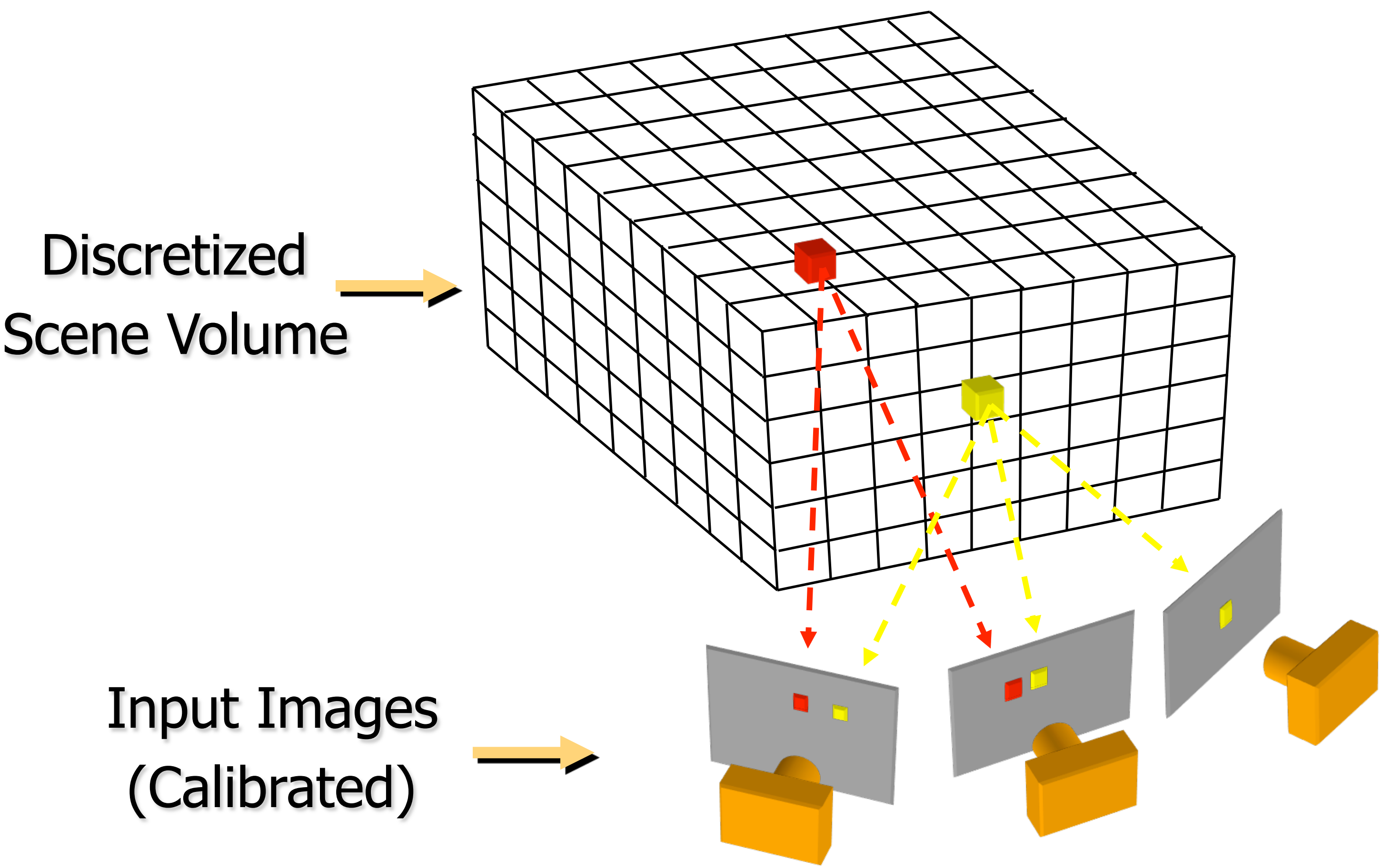
- But is generally not the same
- In the limit (all views) get *visual hull*

Voxel algorithm for volume intersection



Color voxel black if on silhouette in every image

Volumetric Stereo / Voxel Coloring



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

Photo-consistency of a 3d point

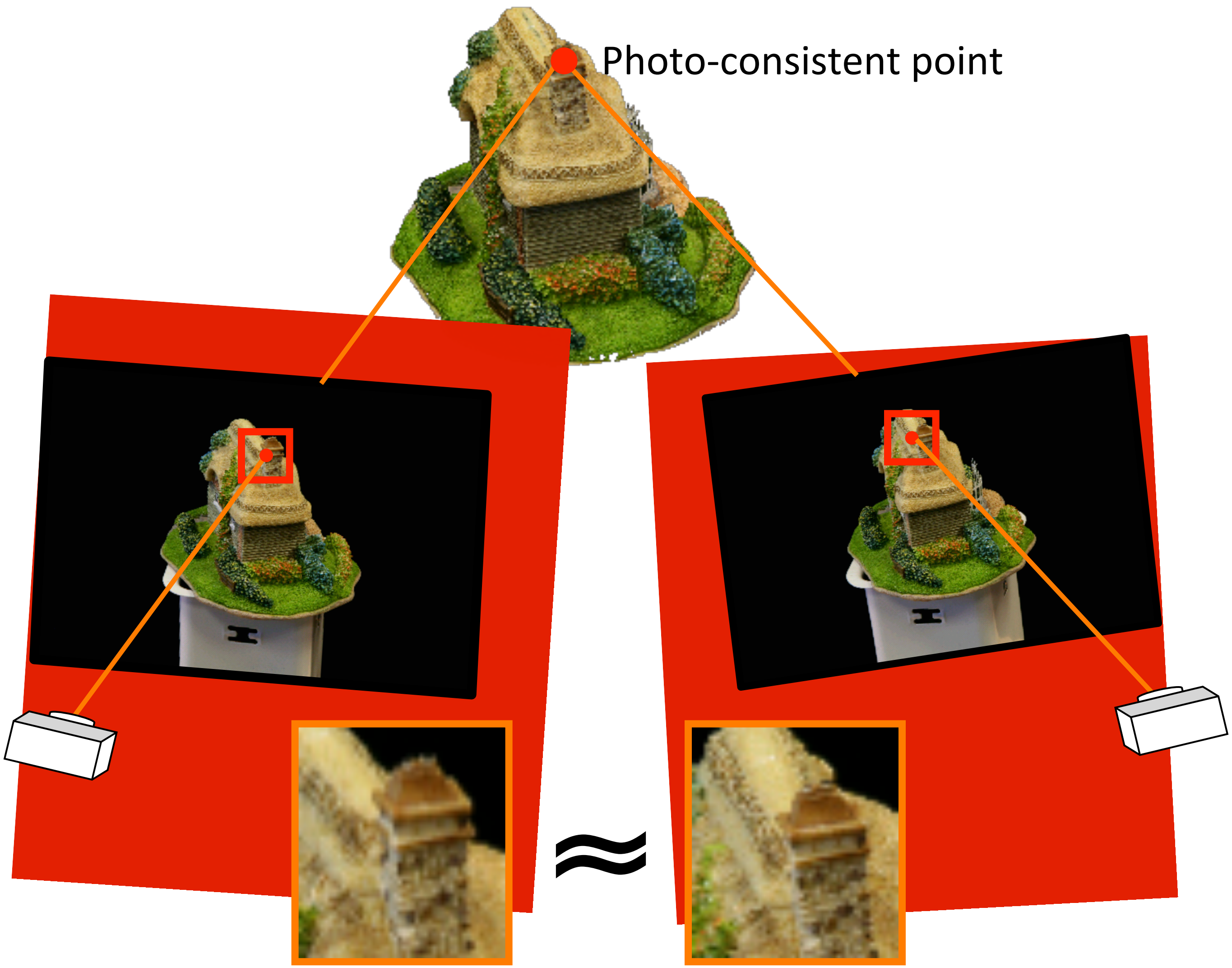


Photo-consistency of a 3d point

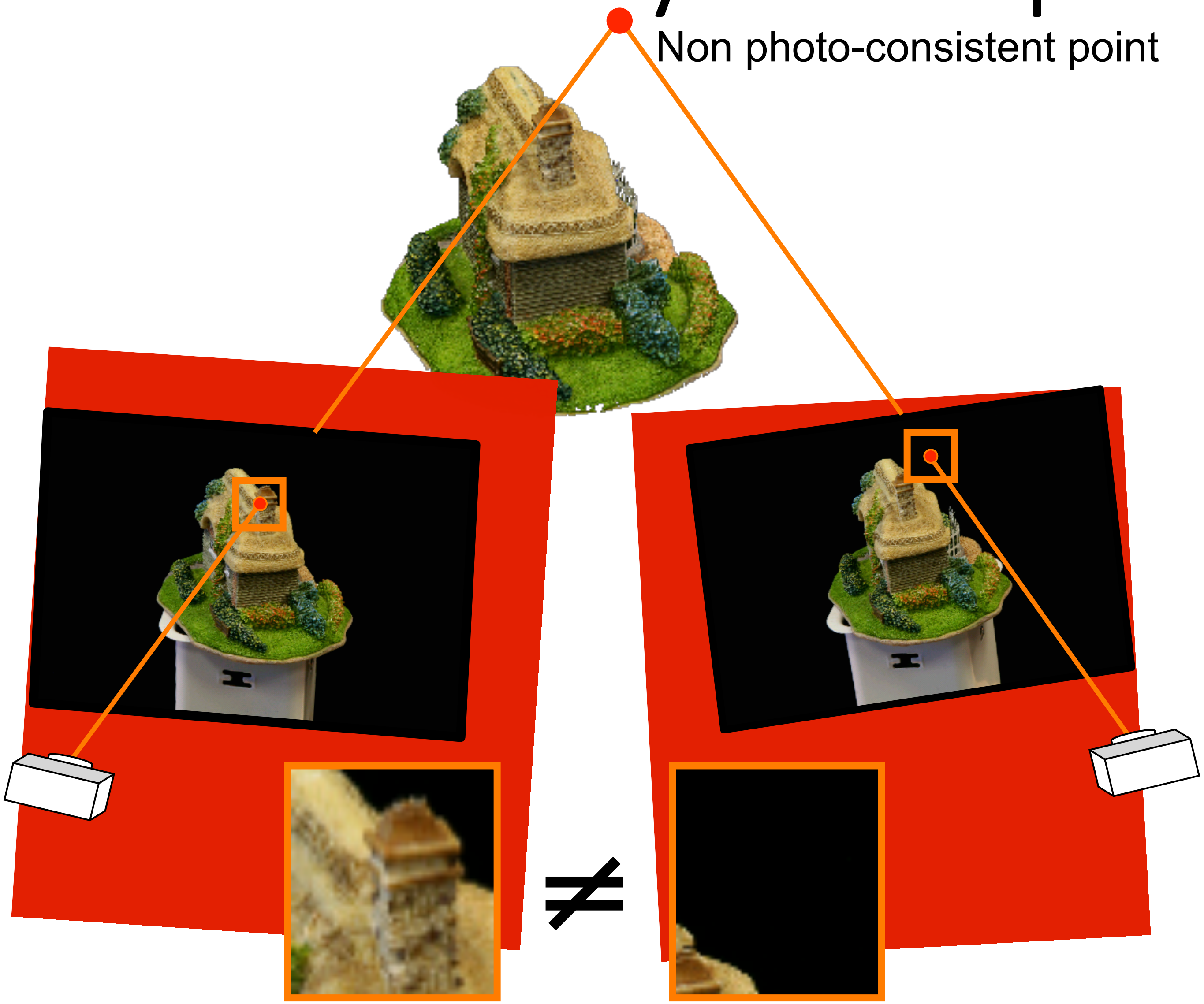
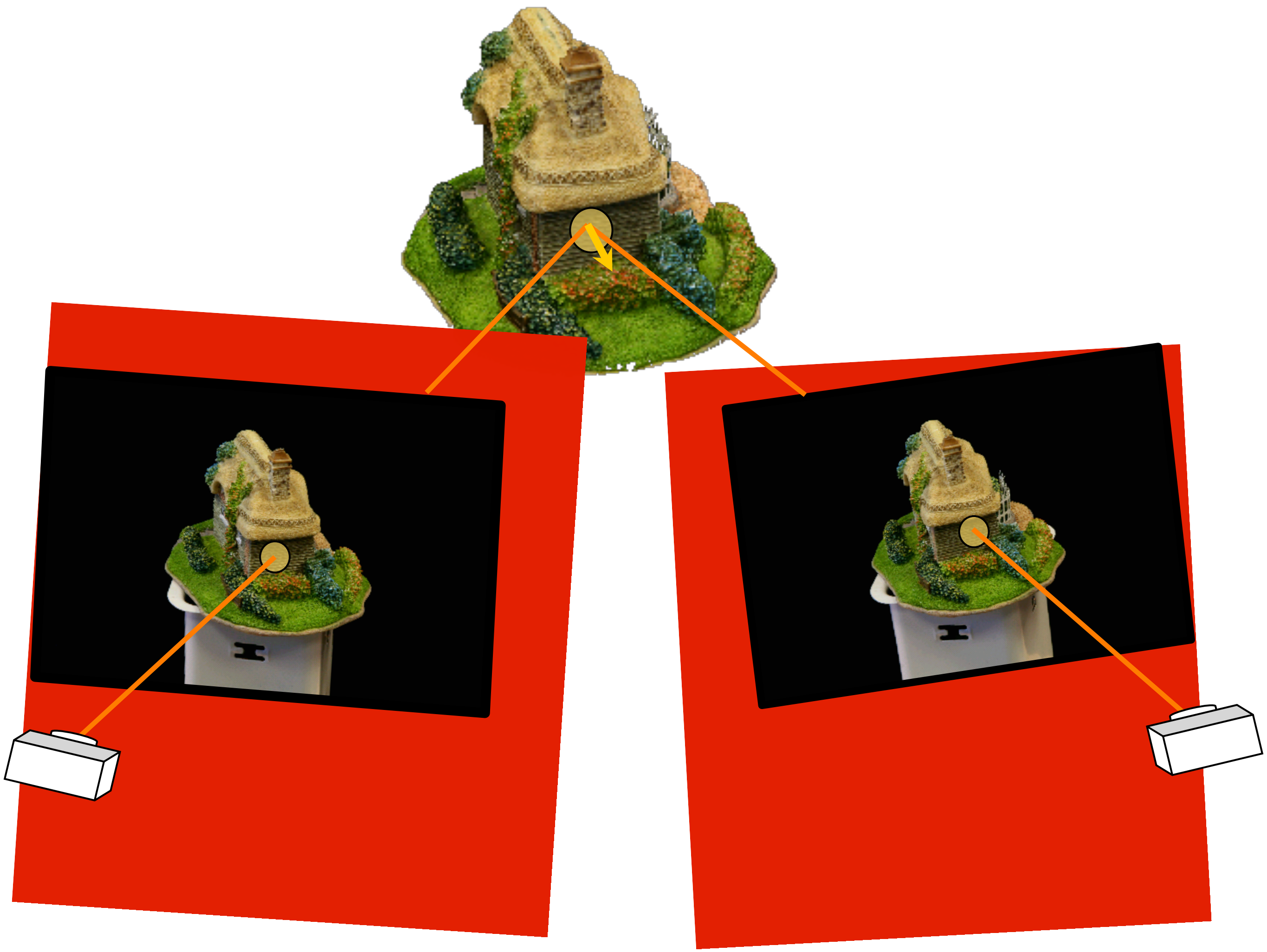


Photo-consistency of a 3d patch



Measuring photo-consistency

- Equivalent statements
  - voxel v is photo-consistent
  - image content is (nearly) identical for all projections of v
  - any pair of projections of v matches well

- Examples:

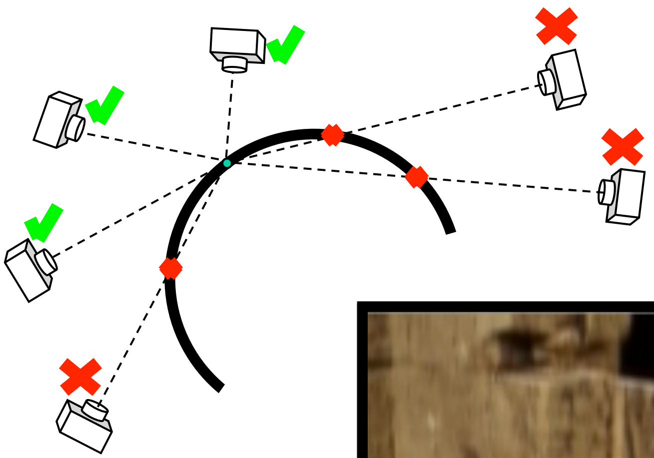
$$\Phi(v) = f\left(\frac{1}{K}\sum_{j=1}^{K}\left(c_j - c_{mean})^2\right)\right)$$

variance of average colour c_j over all K visible images [Seitz&Kutulakos]

$$\Phi(v) = \frac{2}{K(K+1)}\sum_{i=1}^{K}\sum_{j=i+1}^{K}\mathrm{NCC}(p_i, p_j)$$

average normalised cross-correlation over all pairs of visible images [Vogiatzis et al.]

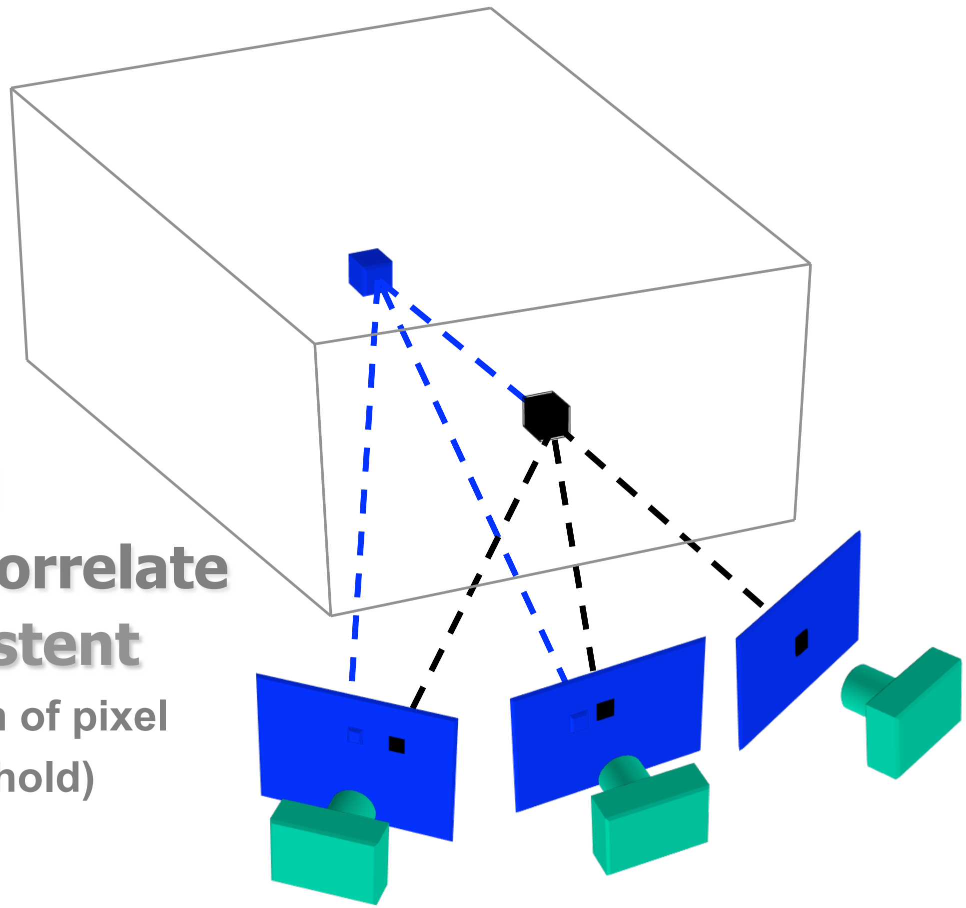
Challenges of photo-consistency



- Camera visibility

- Failure of comparison metric
  - repeated texture
  - lack of texture
  - specularities

Voxel Coloring Approach

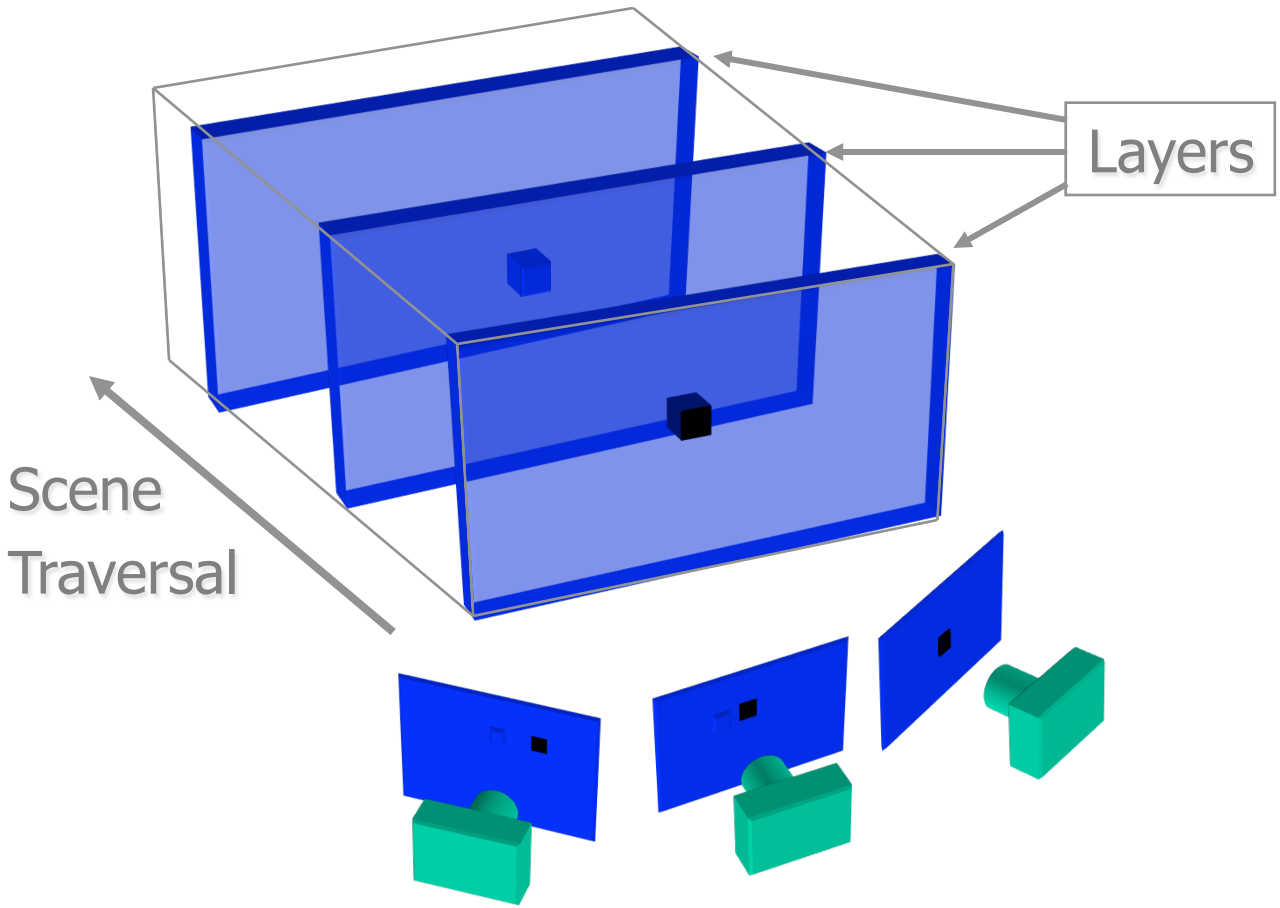


1. Choose voxel
2. Project and correlate
3. Color if consistent

   (standard deviation of pixel colors below threshold)

Visibility Problem: in which images is each voxel visible?

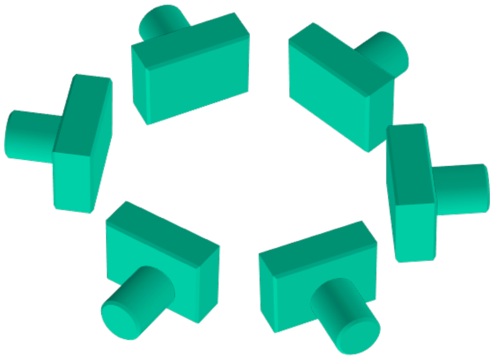
Depth Ordering: visit occluders first!



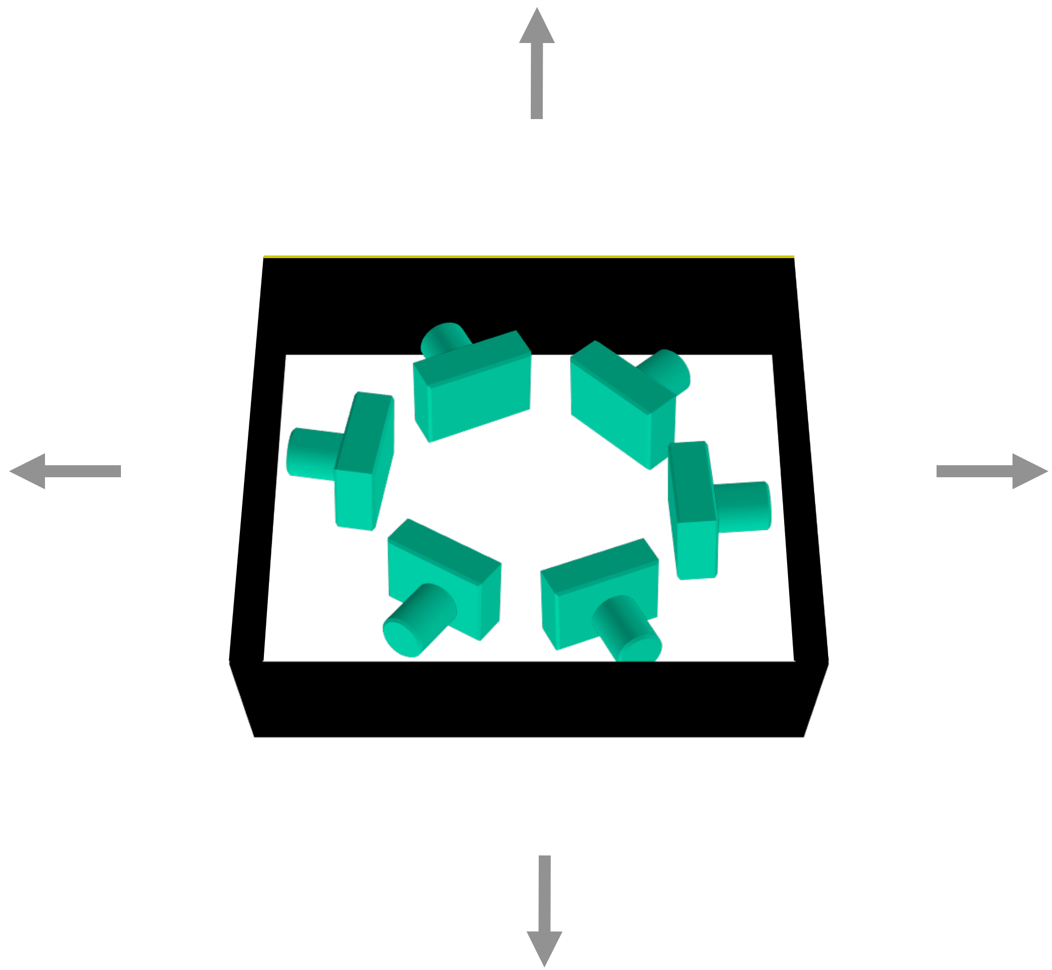
Condition: **depth order is the *same for all input views***

Panoramic Depth Ordering

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

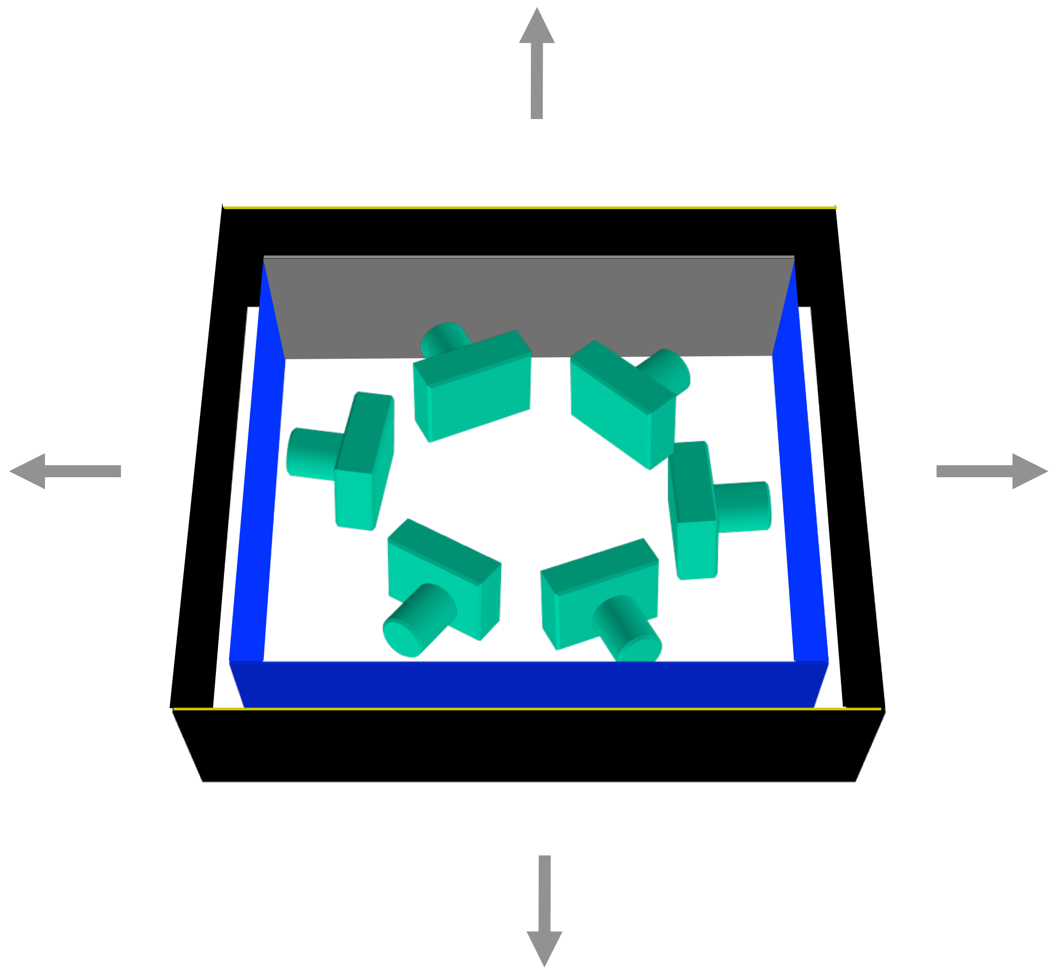


Panoramic Depth Ordering



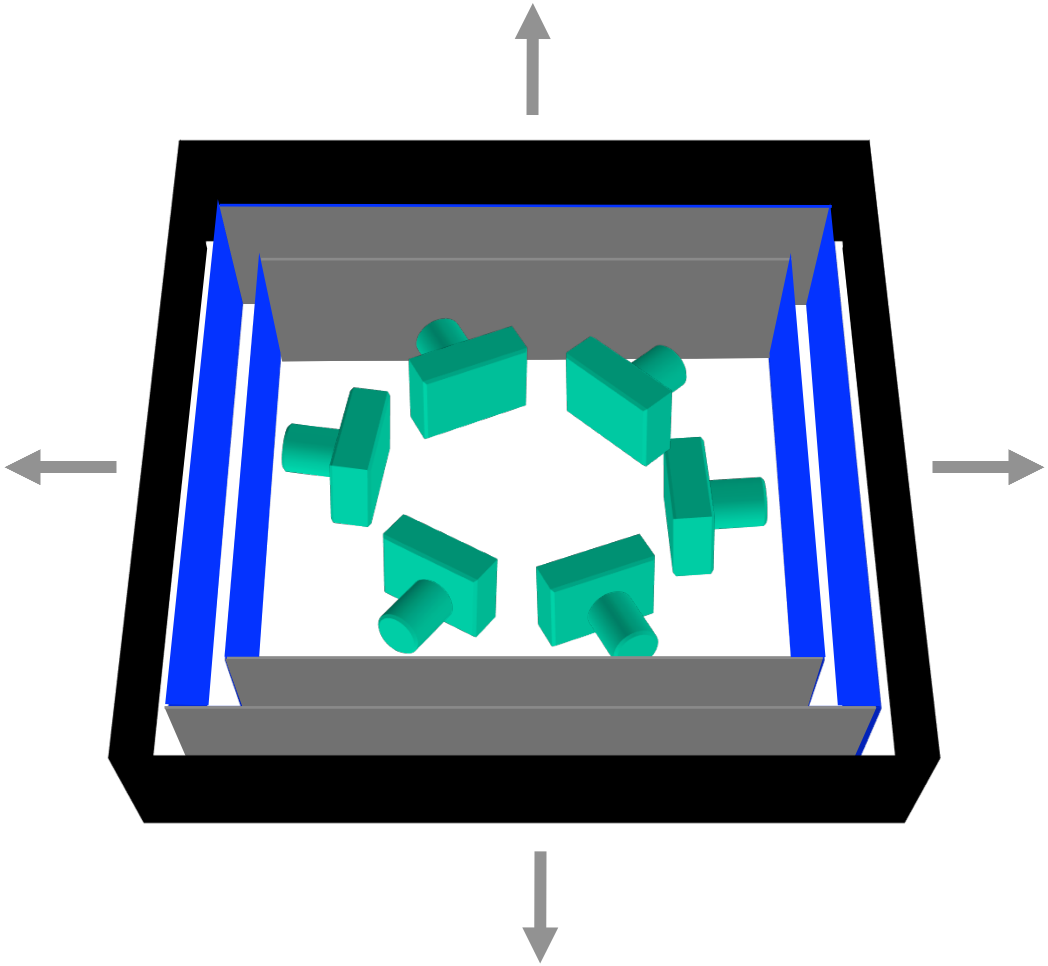
Layers radiate outwards from cameras

Panoramic Layering



Layers radiate outwards from cameras

Panoramic Layering

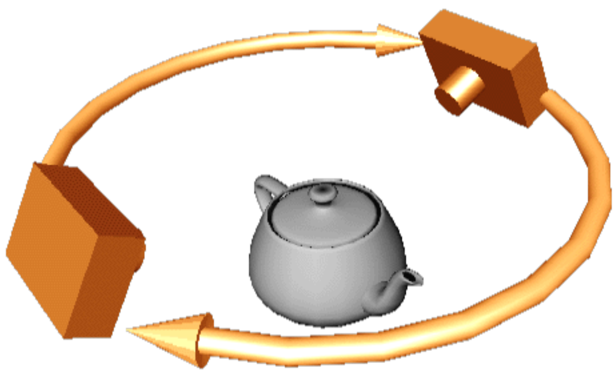


Layers radiate outwards from cameras

Compatible Camera Configurations

Depth-Order Constraint

- Scene outside convex hull of camera centers



Inward-Looking

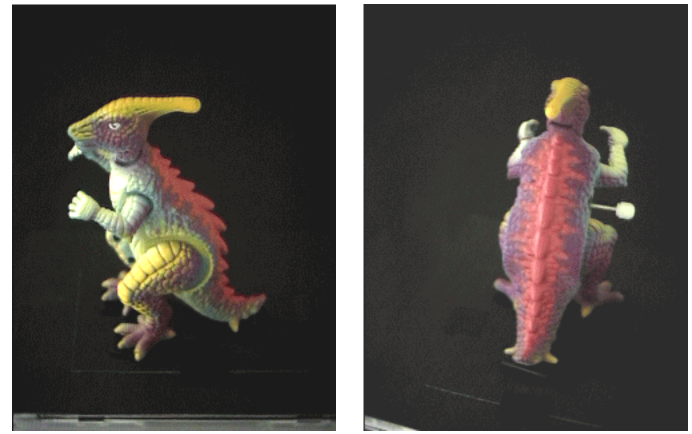
cameras above scene

Outward-Looking

cameras inside scene

Calibrated Image Acquisition

Calibrated Turntable



Selected Dinosaur Images

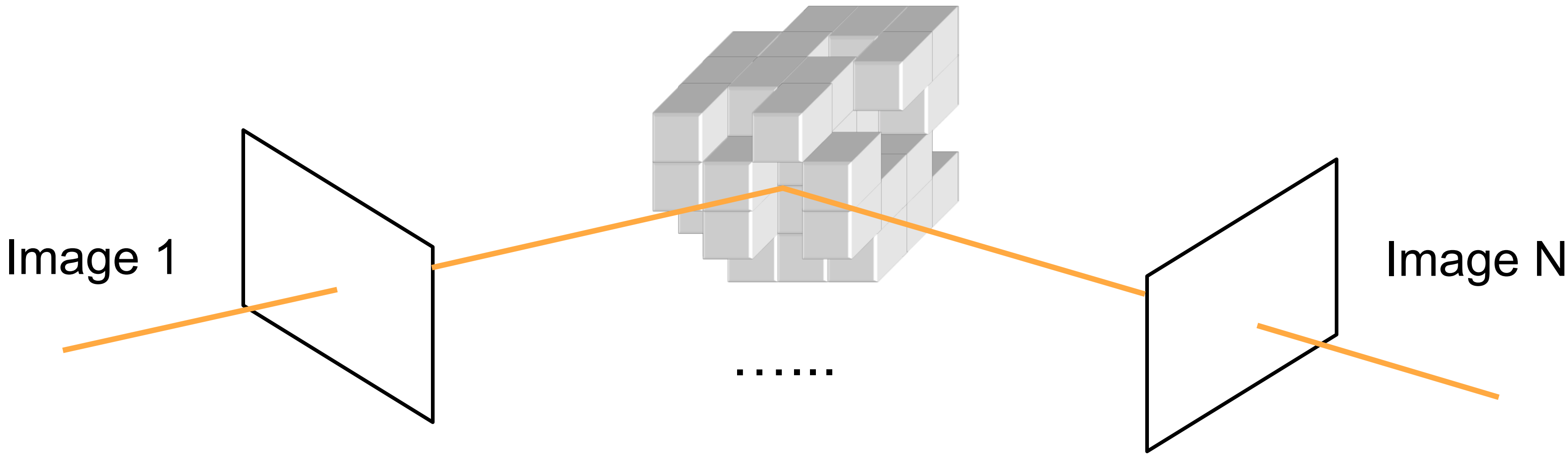
Selected Flower Images

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

Space Carving



Space Carving Algorithm

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

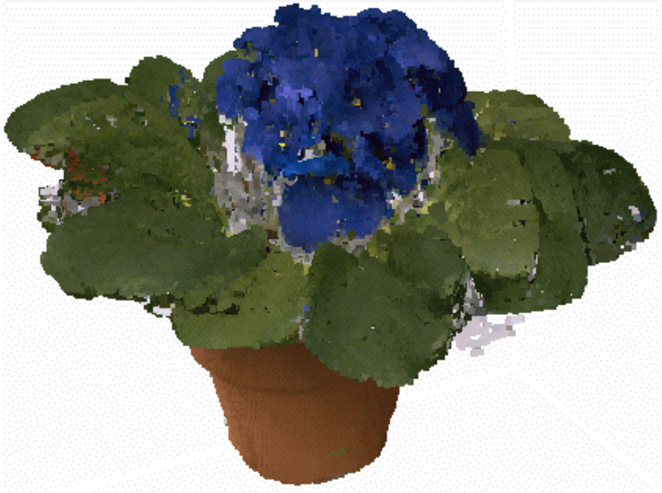
K. N. Kutulakos and S. M. Seitz, **A Theory of Shape by Space Carving**, *ICCV* 1999

Space Carving Results: African Violet

Input Image (1 of 45)

Reconstruction

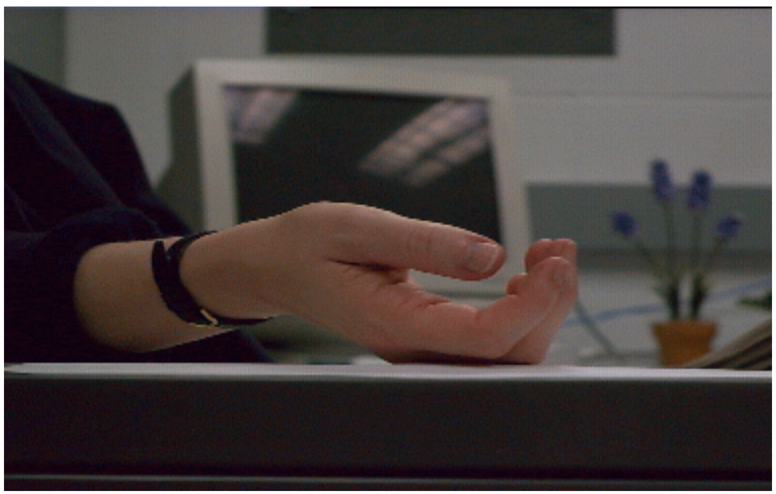
Reconstruction



Reconstruction

Source: S. Seitz

Space Carving Results: Hand



Input Image (1 of 100)

Views of Reconstruction

Carved visual hulls

- The visual hull is a good starting point for optimizing photo-consistency
  - Easy to compute
  - Tight outer boundary of the object
  - Parts of the visual hull (rims) already lie on the surface and are already photo-consistent

Yasutaka Furukawa and Jean Ponce,
Carved Visual Hulls for Image-Based Modeling, ECCV 2006.

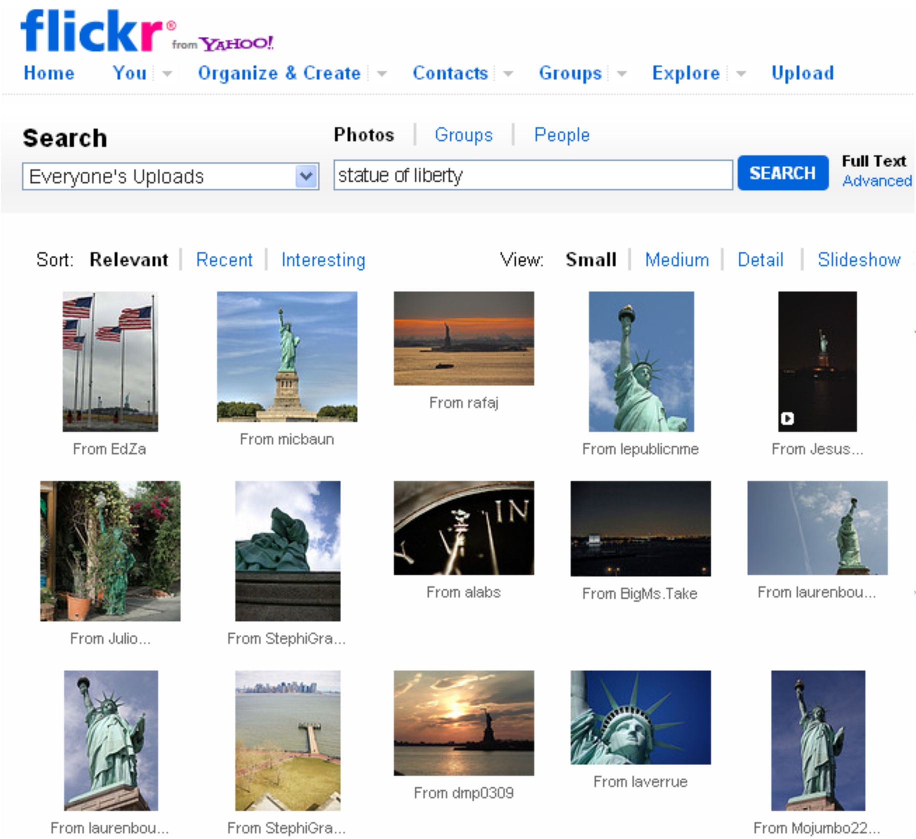
Carved visual hulls

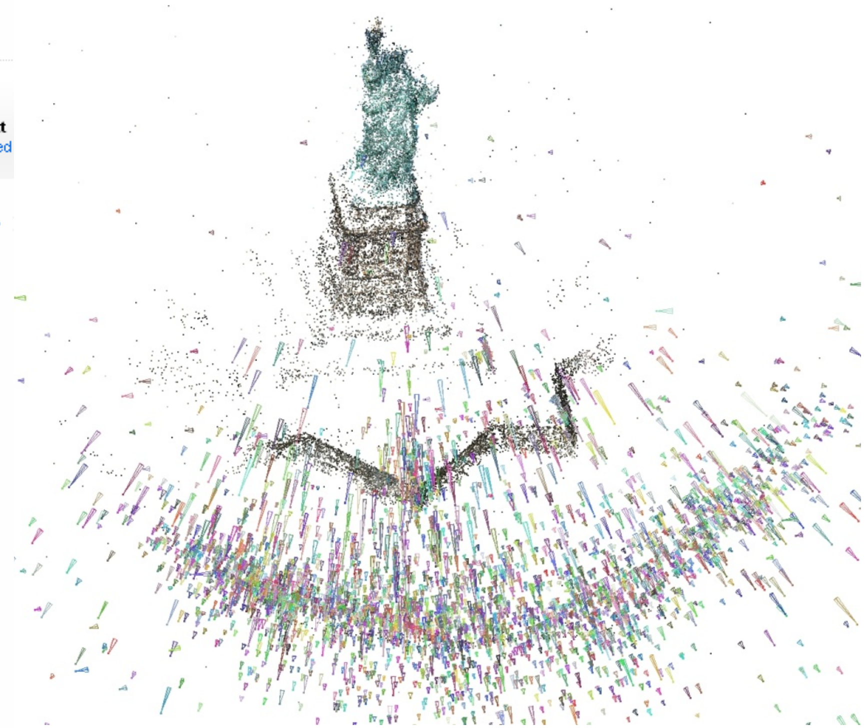
1. Compute visual hull
2. Use dynamic programming to find rims (photo-consistent parts of visual hull)
3. Carve the visual hull to optimize photo-consistency keeping the rims fixed

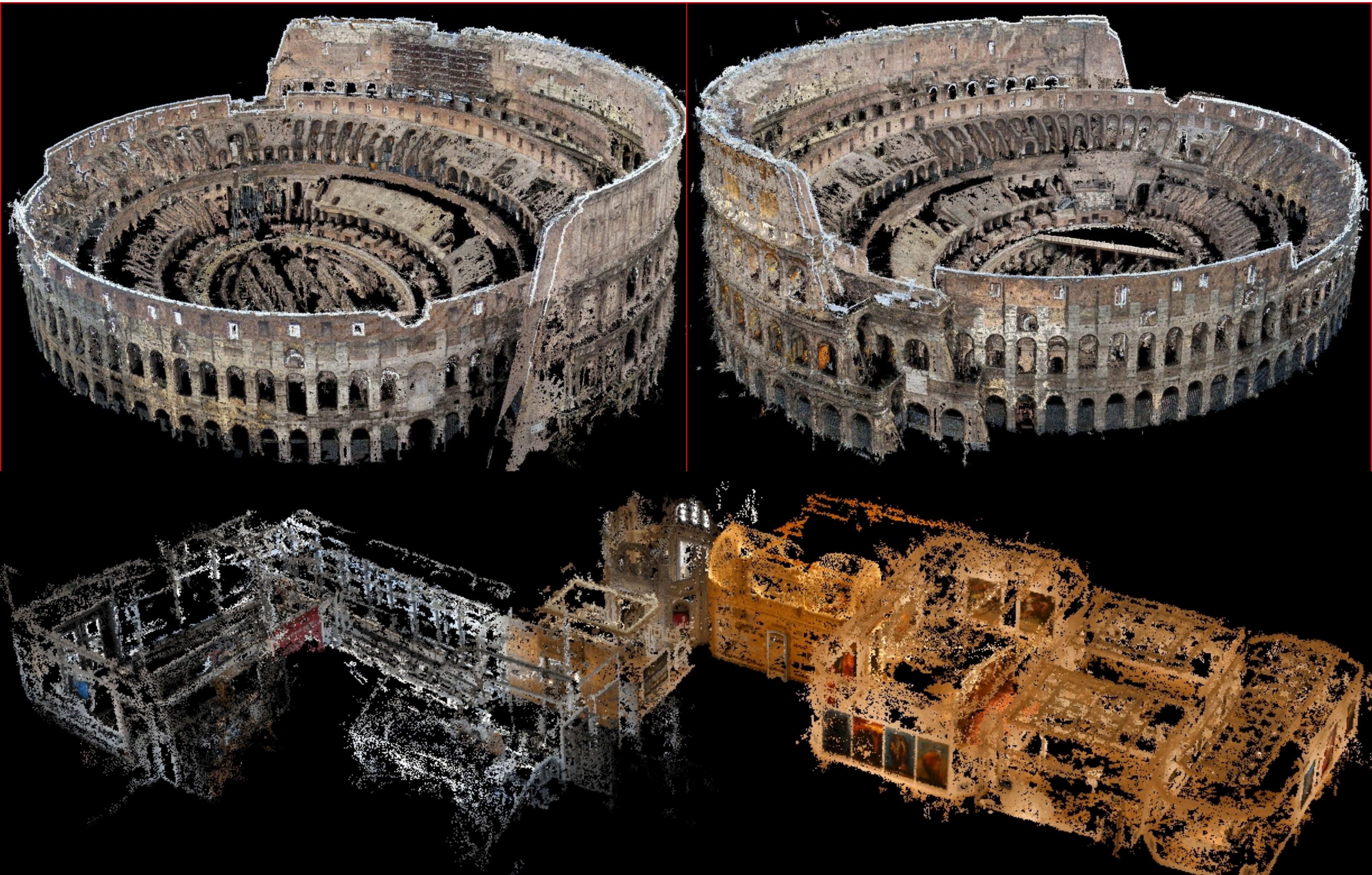
Yasutaka Furukawa and Jean Ponce,
[Carved Visual Hulls for Image-Based Modeling](), ECCV 2006.

Stereo from community photo collections

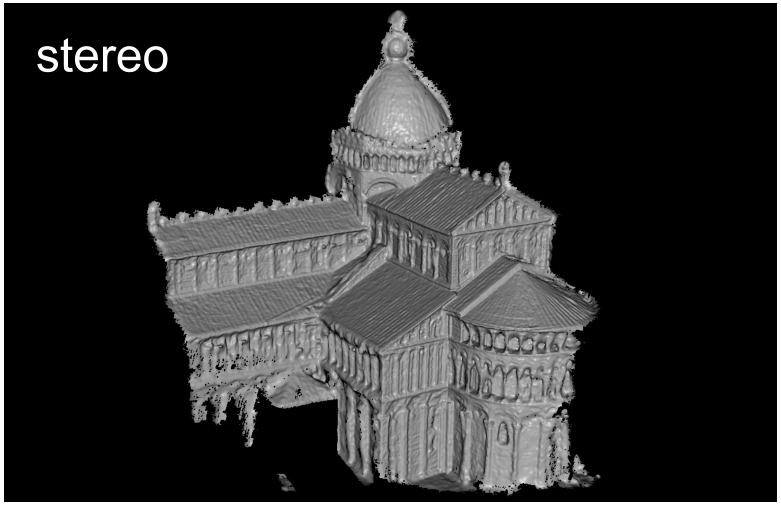
- Up to now, we've always assumed that camera calibration is known
- For photos taken from the Internet, we need *structure from motion* techniques to reconstruct both camera positions and 3D points

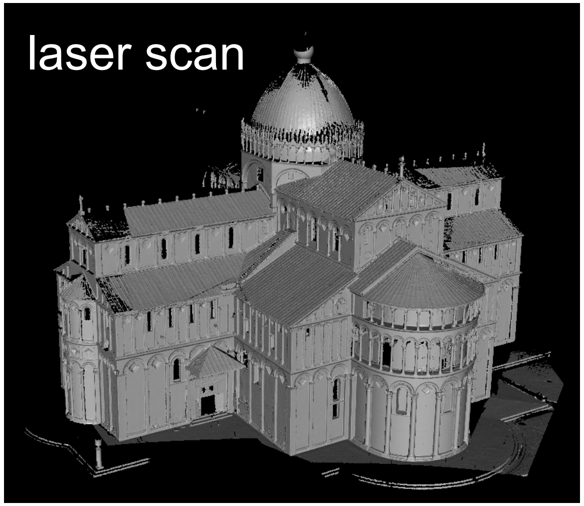


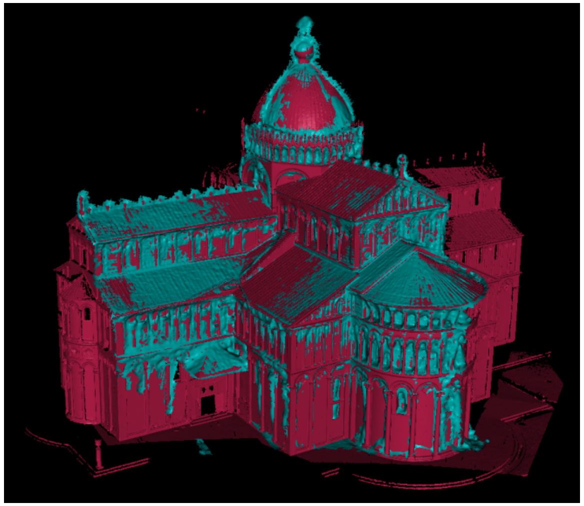




Stereo from community photo collections





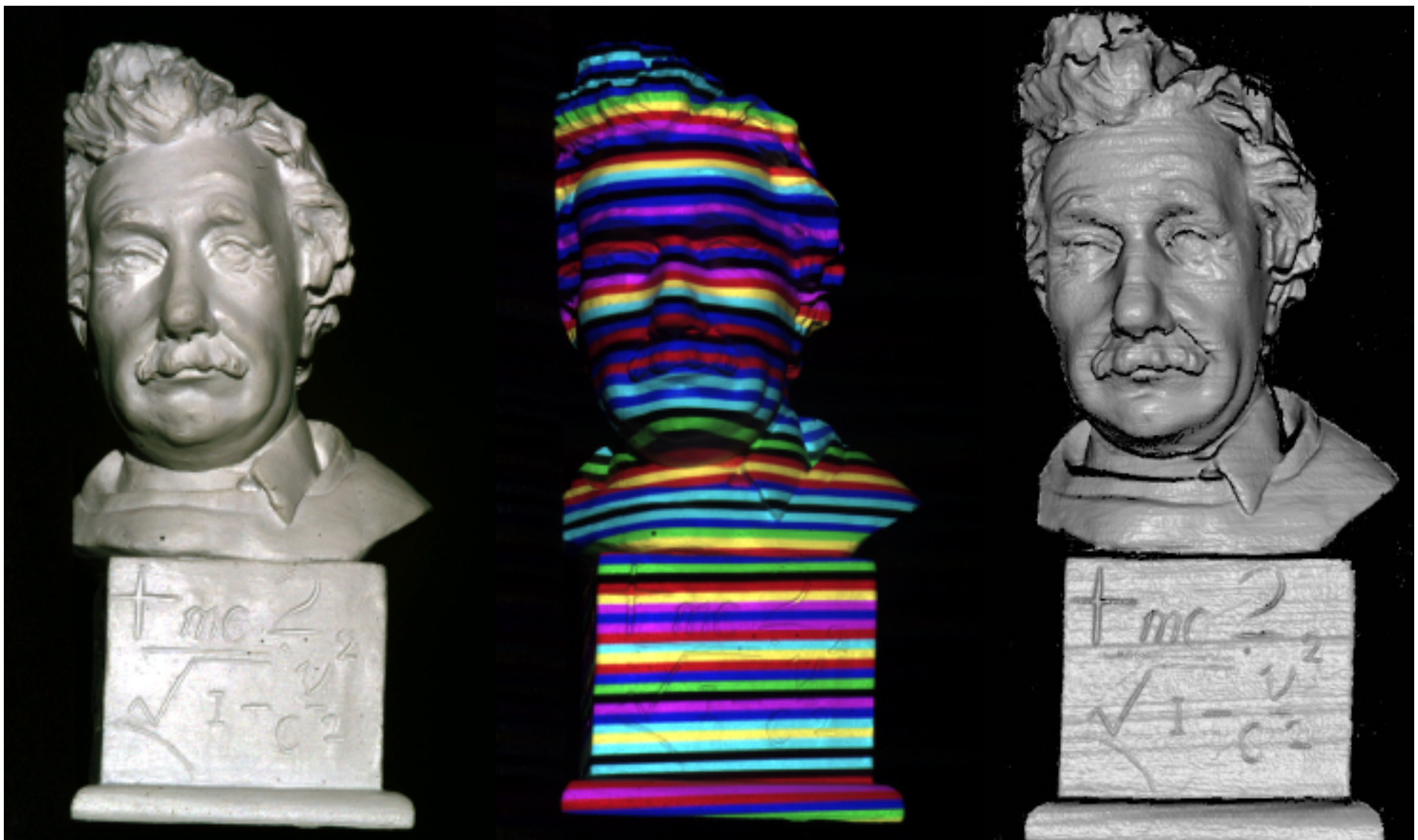


Comparison: 90% of points within 0.128 m of laser scan (building height 51m)

M. Goesele, N. Snavely, B. Curless, H. Hoppe, S. Seitz, [Multi-View Stereo for Community Photo Collections](), ICCV 2007

Scanning technologies

Structured light



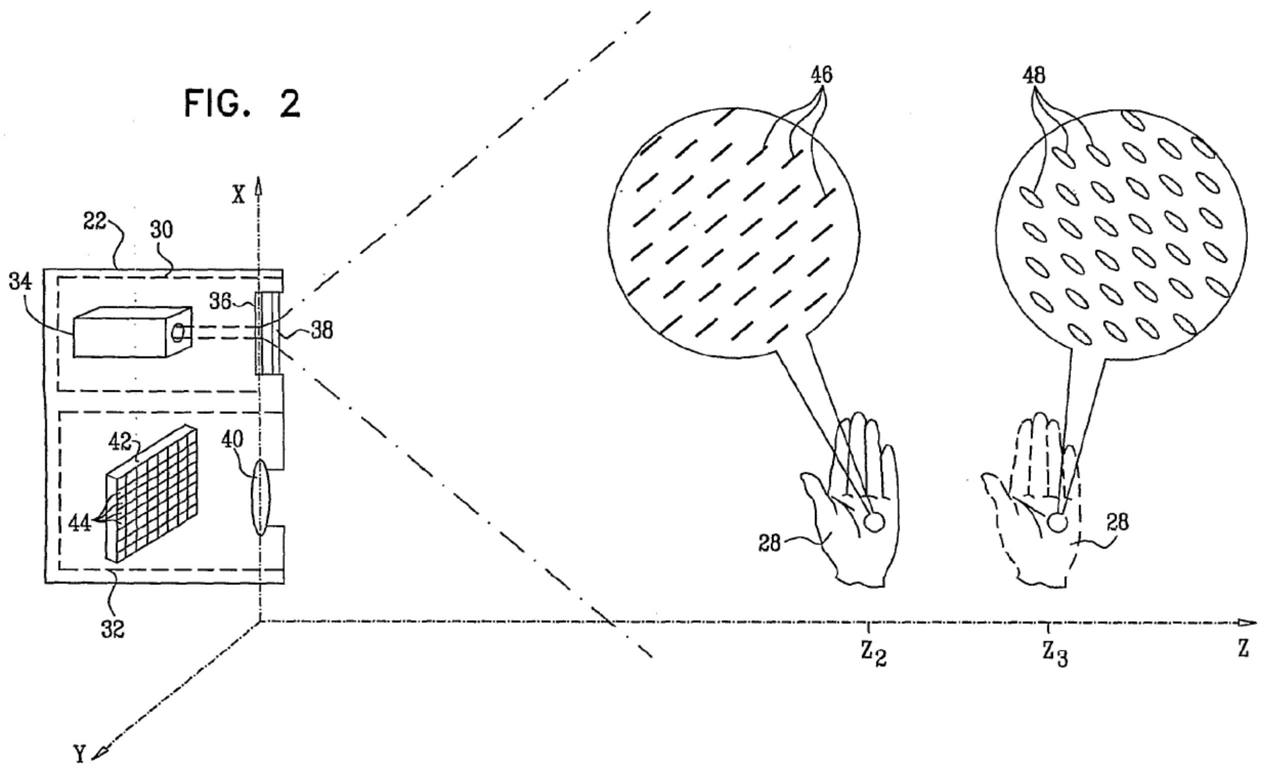
[Zhang02]

Kinect: Structured infrared light

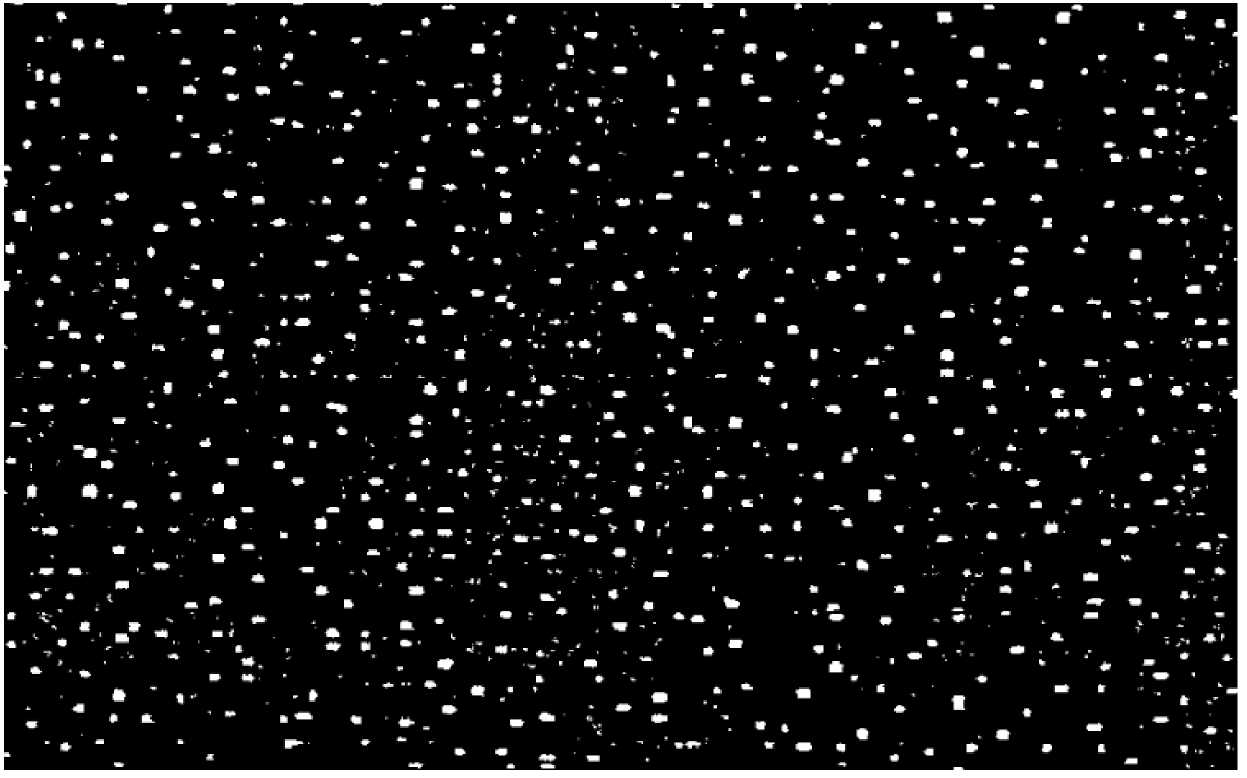
http://bbzippo.wordpress.com/2010/11/28/kinect-in-infrared/

PrimeSense Patents

FIG. 2



The Kinect uses infrared laser light, with a speckle pattern



Shpunt et al, PrimeSense patent application
US 2008/0106746

KinectFusion: Real-time 3D Reconstruction and Interaction Using a Moving Depth Camera*