Lecture 8
Active stereo & Volumetric stereo

- Active stereo
  - Structured lighting
  - Depth sensing
- Volumetric stereo:
  - Space carving $O(N^3)$
  - Shadow carving $O(2N^3)$
  - Voxel coloring $O(LN^3)$

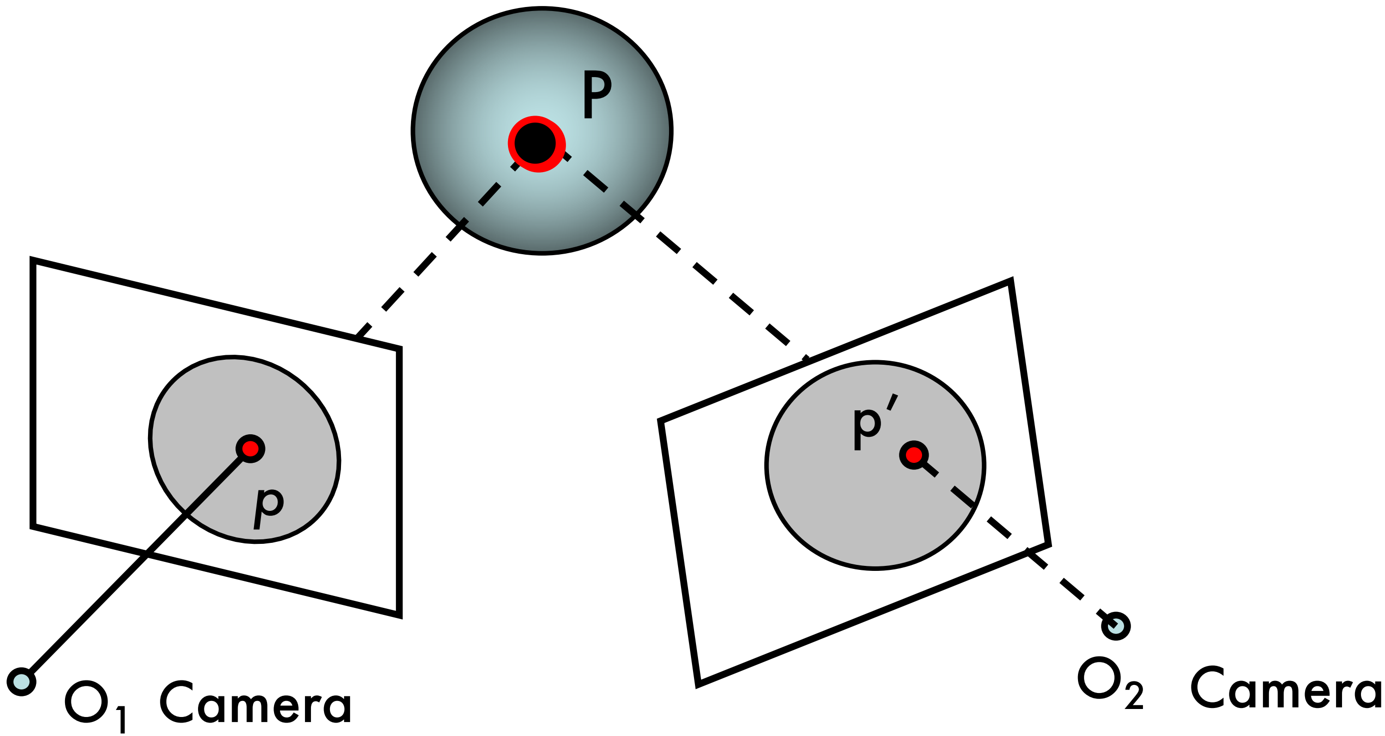
Reading:

[Szelisky] Chapter 11 "Multi-view stereo"

S. Savarese, M. Andreetto, H. Rushmeier, F. Bernardini and P. Perona, 3D Reconstruction by Shadow Carving: Theory and Practical Evaluation, International Journal of Computer Vision (IJCV) , 71(3), 305-336, 2006

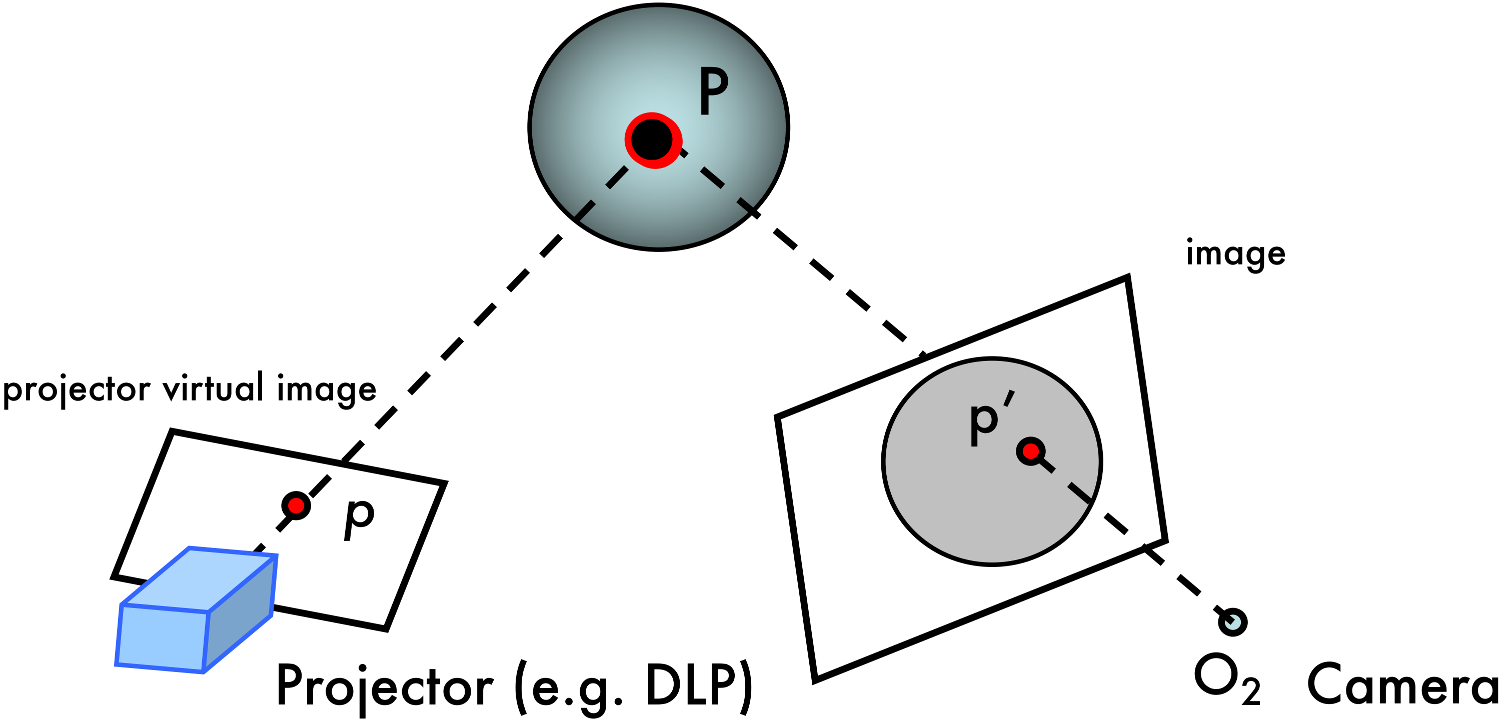
Seitz, S. M., & Dyer, C. R. (1999). Photorealistic scene reconstruction by voxel coloring. *International Journal of Computer Vision*, *35*(2), 151-173.

Traditional stereo



What's the main problem in traditional stereo?
We need to find correspondences!

Active stereo (point)

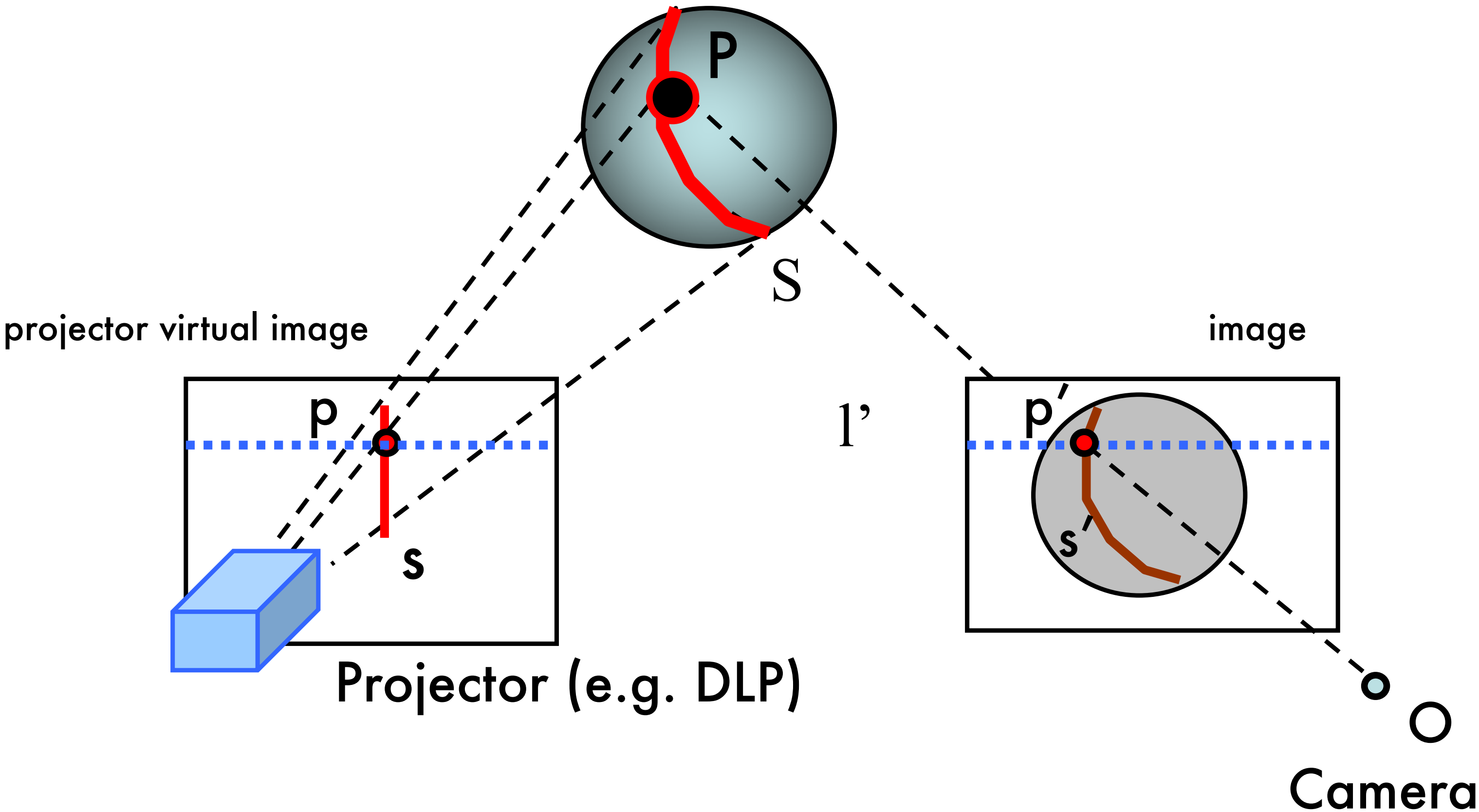


Replace one of the two cameras by a projector

- Projector geometry calibrated
- What's the advantage of having the projector? Correspondence problem solved!

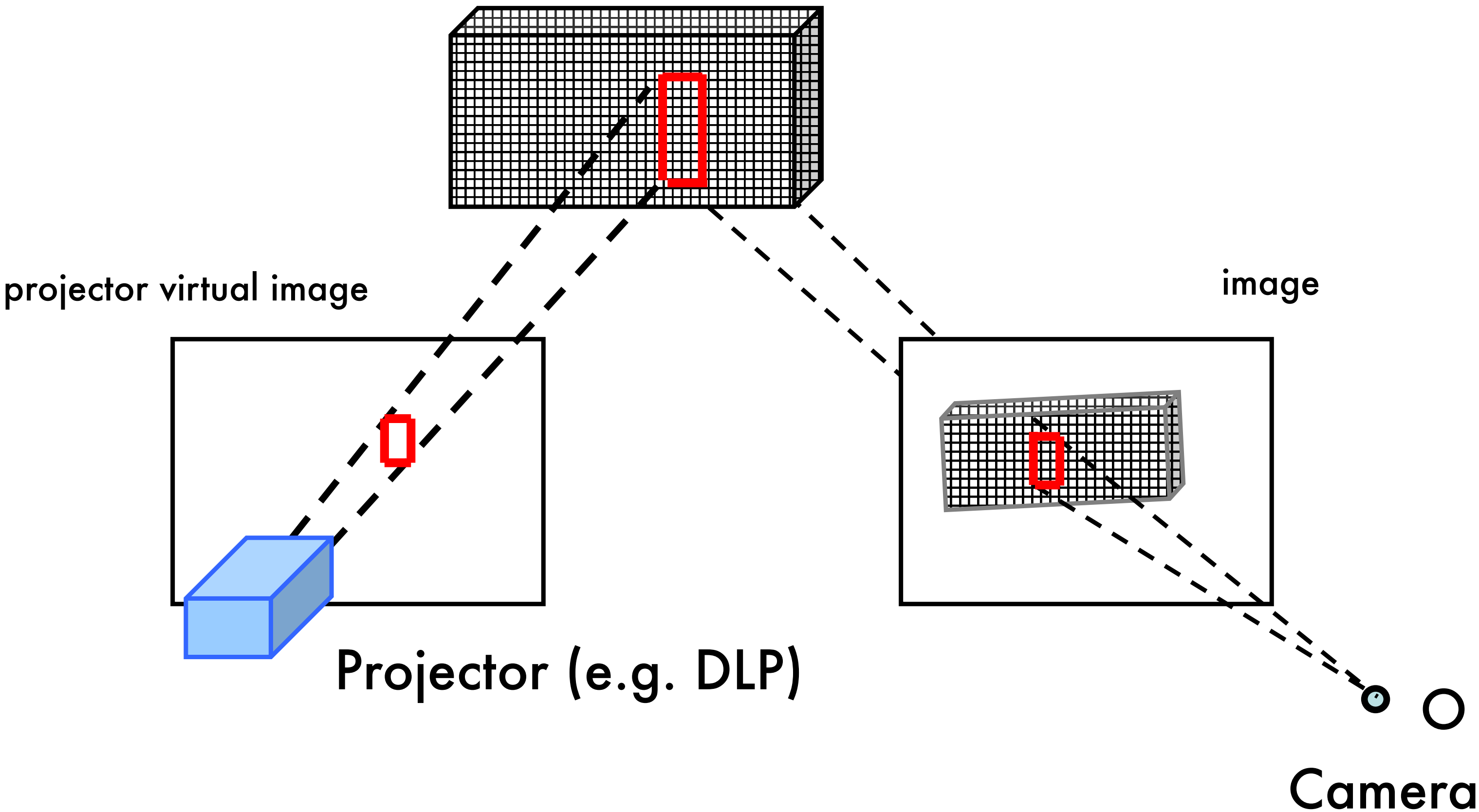
Any limitation??

Active stereo (stripe)



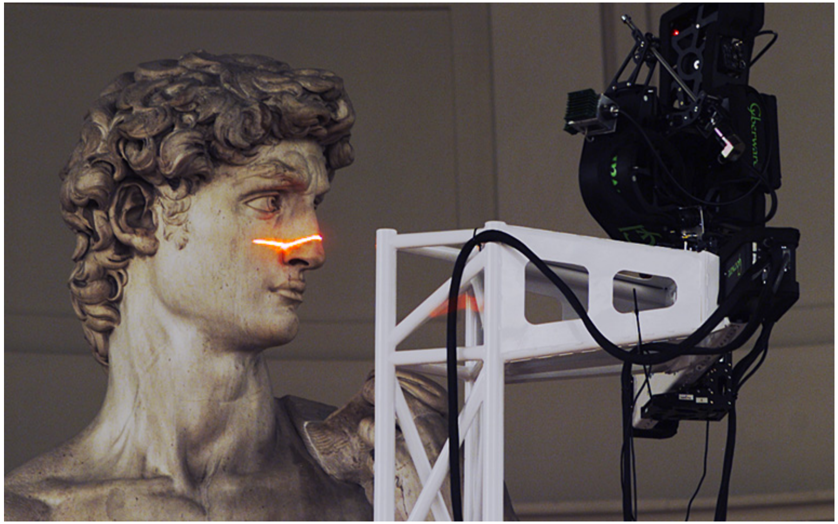
- Projector and camera are parallel
- Correspondence problem is solved!

Calibrating the system



- Use calibration rig to calibrate camera and localize rig in 3D
- Project patterns on rig and calibrate projector

Laser scanning



Digital Michelangelo Project (1990)
http://graphics.stanford.edu/projects/mich/

- Optical triangulation
  - Project a single stripe of laser light
  - Scan it across the surface of the object
  - This is a very precise version of structured light scanning

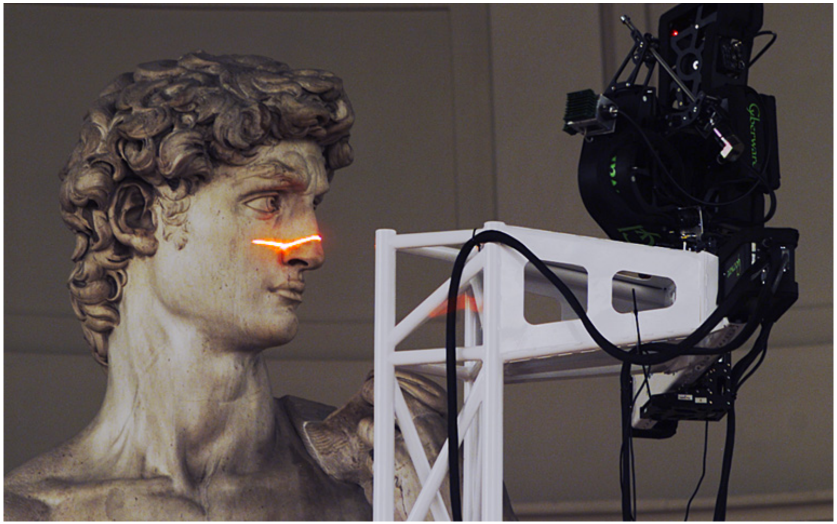
Source: S. Seitz

Laser scanning

The Digital Michelangelo Project, Levoy et al.

Source: S. Seitz

Limitations of Laser scanning

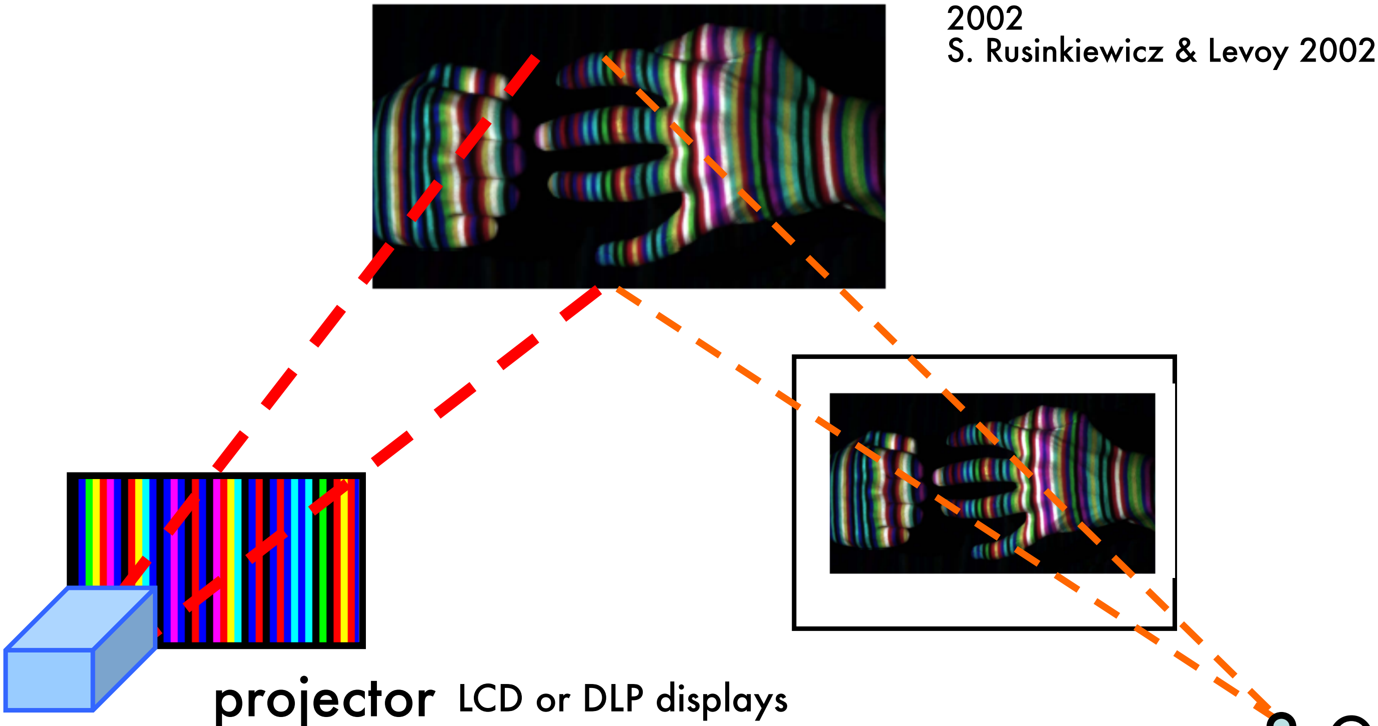


- Slow
- Cannot capture deformations in time

Source: S. Seitz

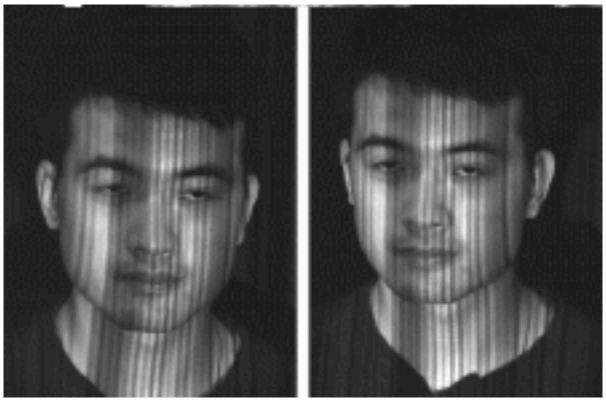
Active stereo (color-coded stripes)

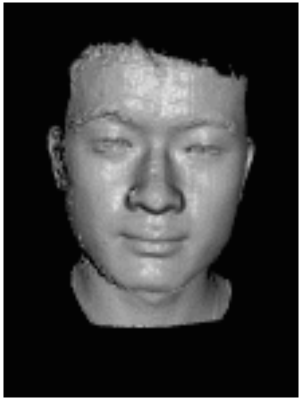
L. Zhang, B. Curless, and S. M. Seitz 2002
S. Rusinkiewicz & Levoy 2002



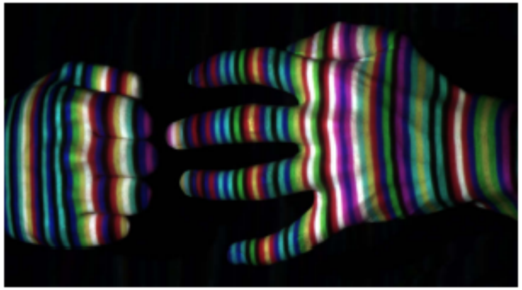
- Dense reconstruction
- Correspondence problem again
- Get around it by using color codes

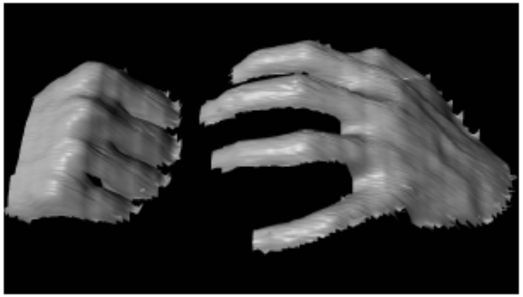
Active stereo (color-coded stripes)





Rapid shape acquisition: Projector + stereo cameras



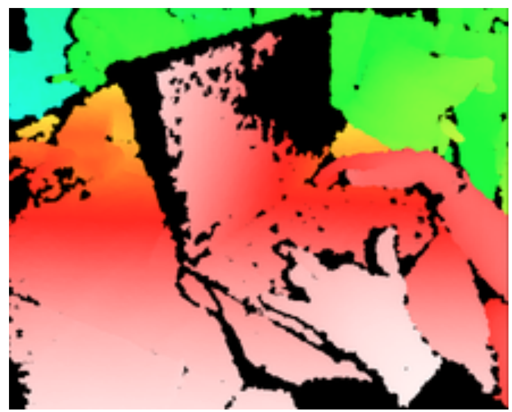


L. Zhang, B. Curless, and S. M. Seitz. Rapid Shape Acquisition Using Color Structured Light and Multi-pass Dynamic Programming. *3DPVT* 2002

Active stereo – the kinect sensor

- Infrared laser projector combined with a CMOS sensor
- Captures video data in 3D under any ambient light conditions.

Pattern of projected infrared points to generate a dense 3D image



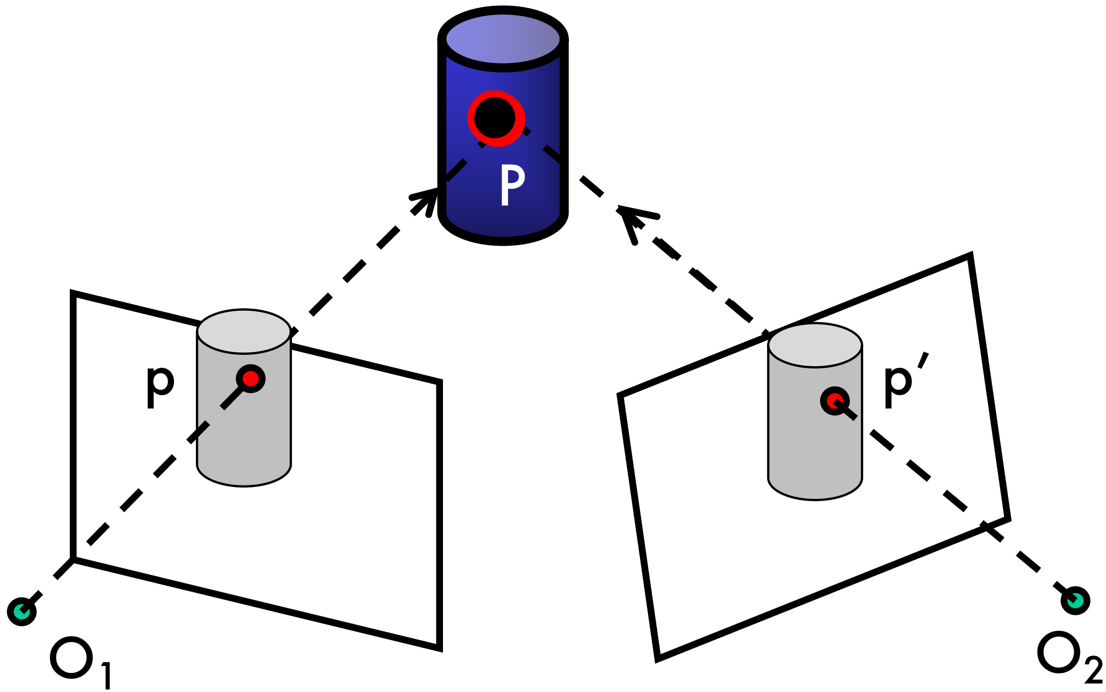
Depth map

Source: wikipedia

Lecture 8
Active stereo & Volumetric stereo

- Active stereo
  - Structured lighting
  - Depth sensing
- Volumetric stereo:
  - Space carving
  - Shadow carving
  - Voxel coloring

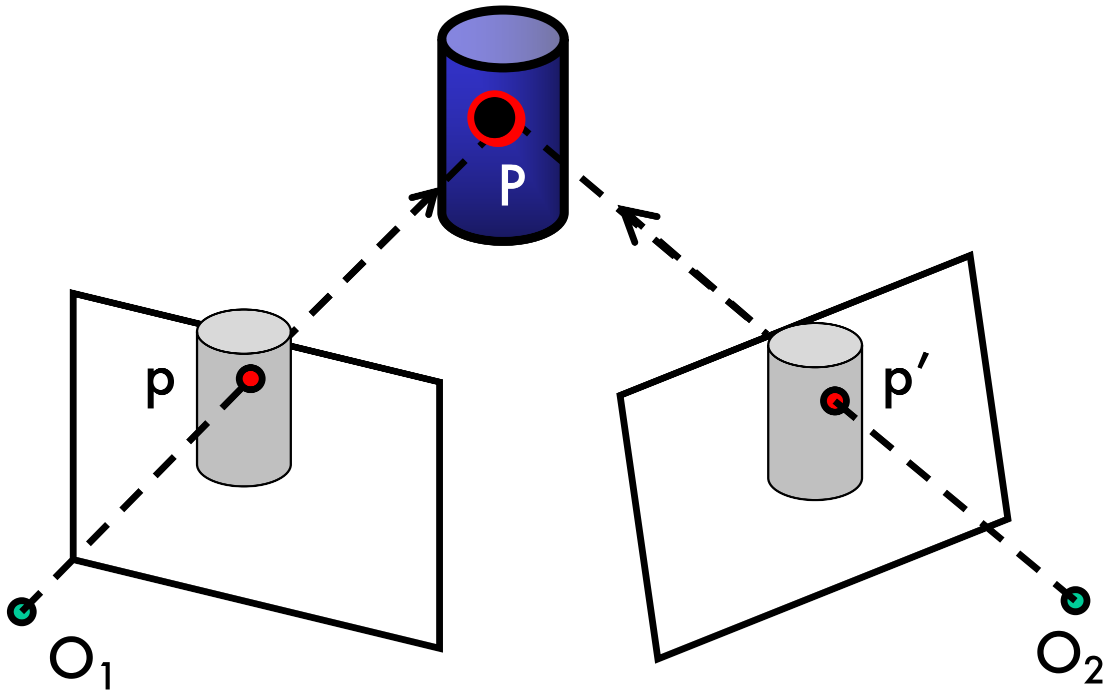
"Traditional" Stereo



Goal: estimate the position of P given the observation of P from two view points

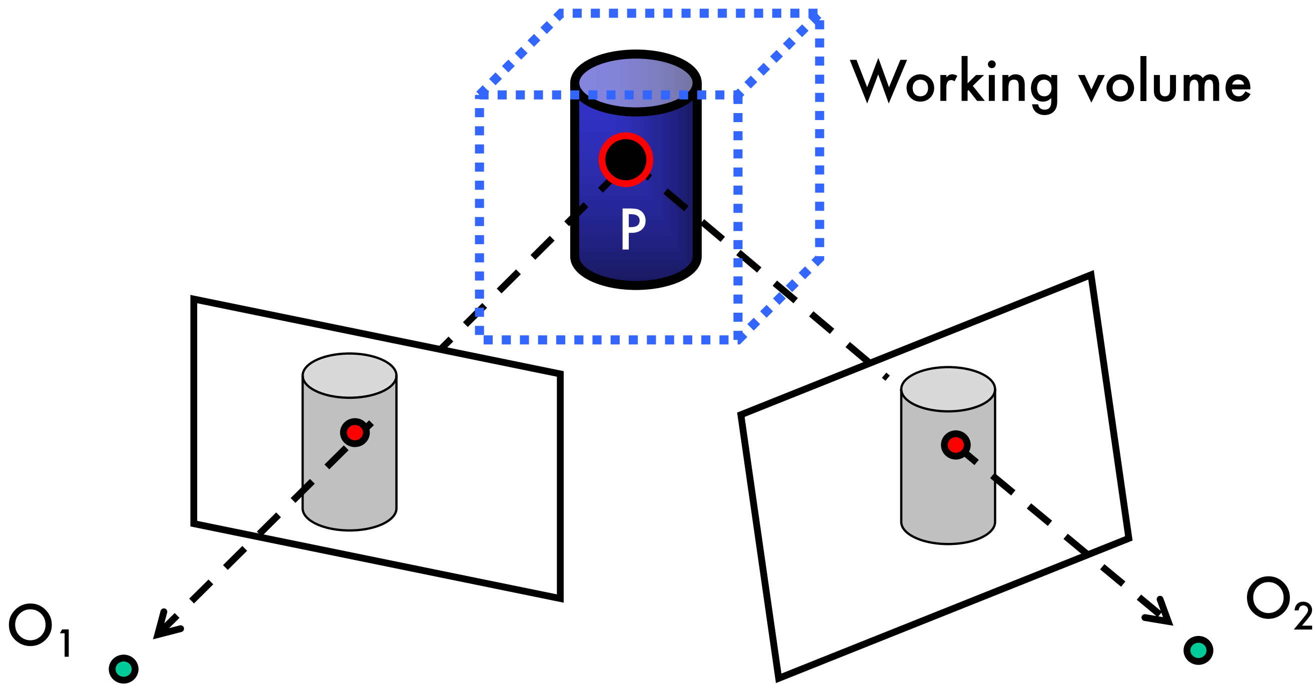
Assumptions: known camera parameters and position (K, R, T)

"Traditional" Stereo



Subgoals:
1. Solve the correspondence problem
2. Use corresponding observations to triangulate

Volumetric stereo



1. Hypothesis: pick up a point within the volume
2. Project this point into 2 (or more) images
3. Validation: are the observations consistent?

Assumptions: known camera parameters and position (K, R, T)

Consistency based on cues such as:

- Contours/silhouettes → Space carving
- Shadows → Shadow carving
- Colors → Voxel coloring

Lecture 8
Active stereo & Volumetric stereo

- Active stereo
  - Structured lighting
  - Depth sensing
- Volumetric stereo:
  - Space carving
  - Shadow carving
  - Voxel coloring

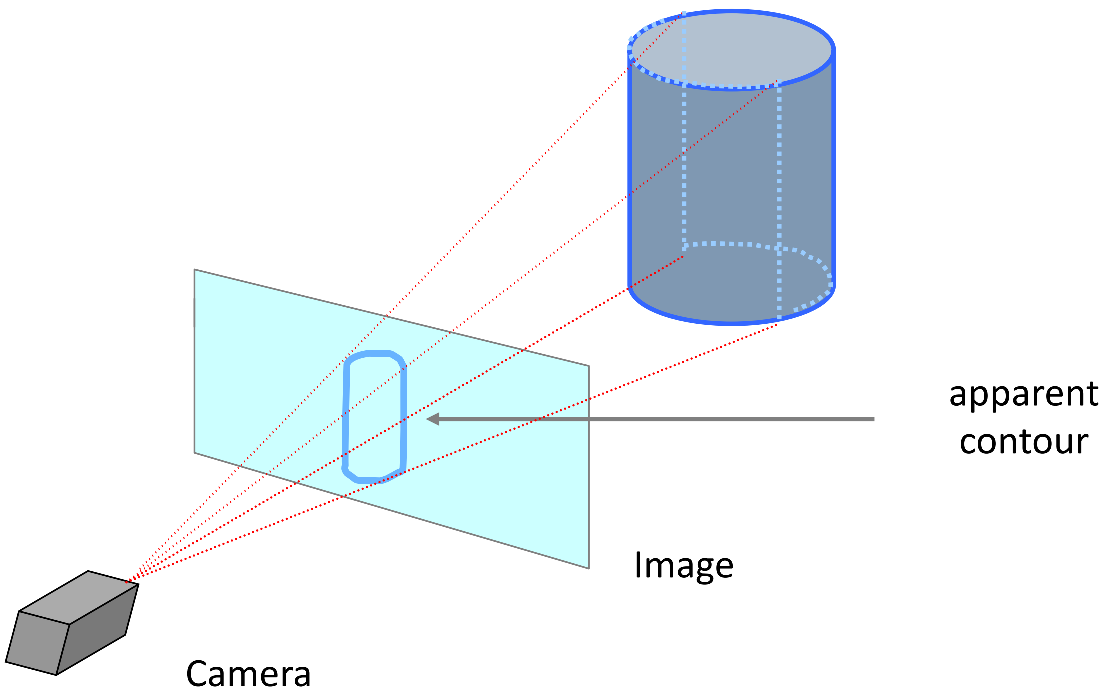
Reading:
[Szelisky] Chapter 11 “Multi-view stereo”

Contours/silhouettes

- Contours are a rich source of geometric information

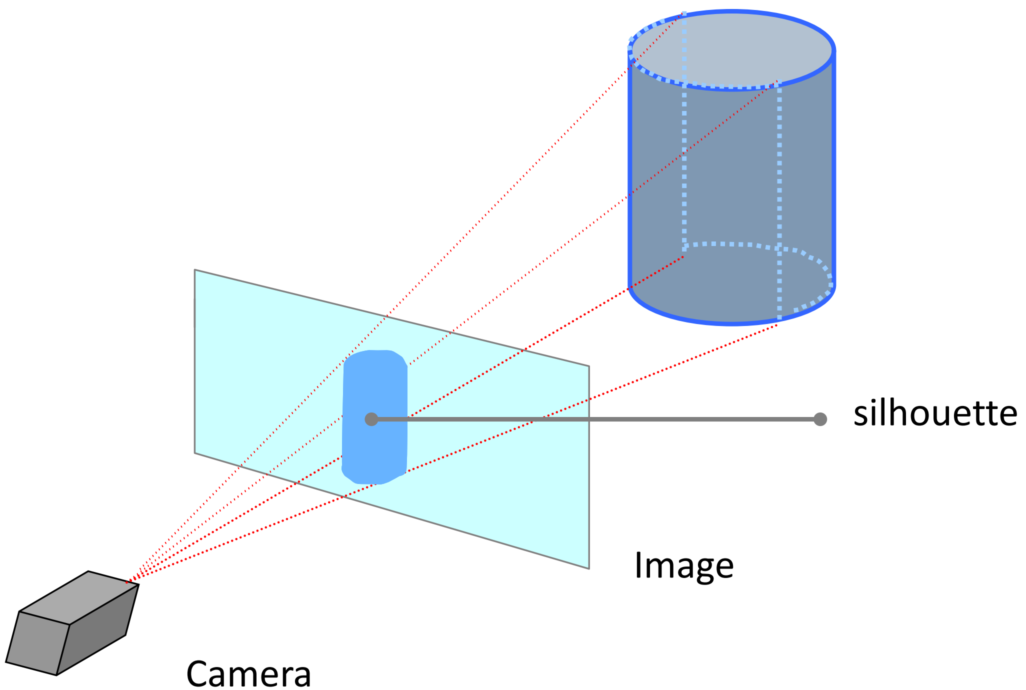
Apparent Contours

- Projection of the locus of points on the object surface which separate the visible and occluded parts on the surface [sato & cipolla]

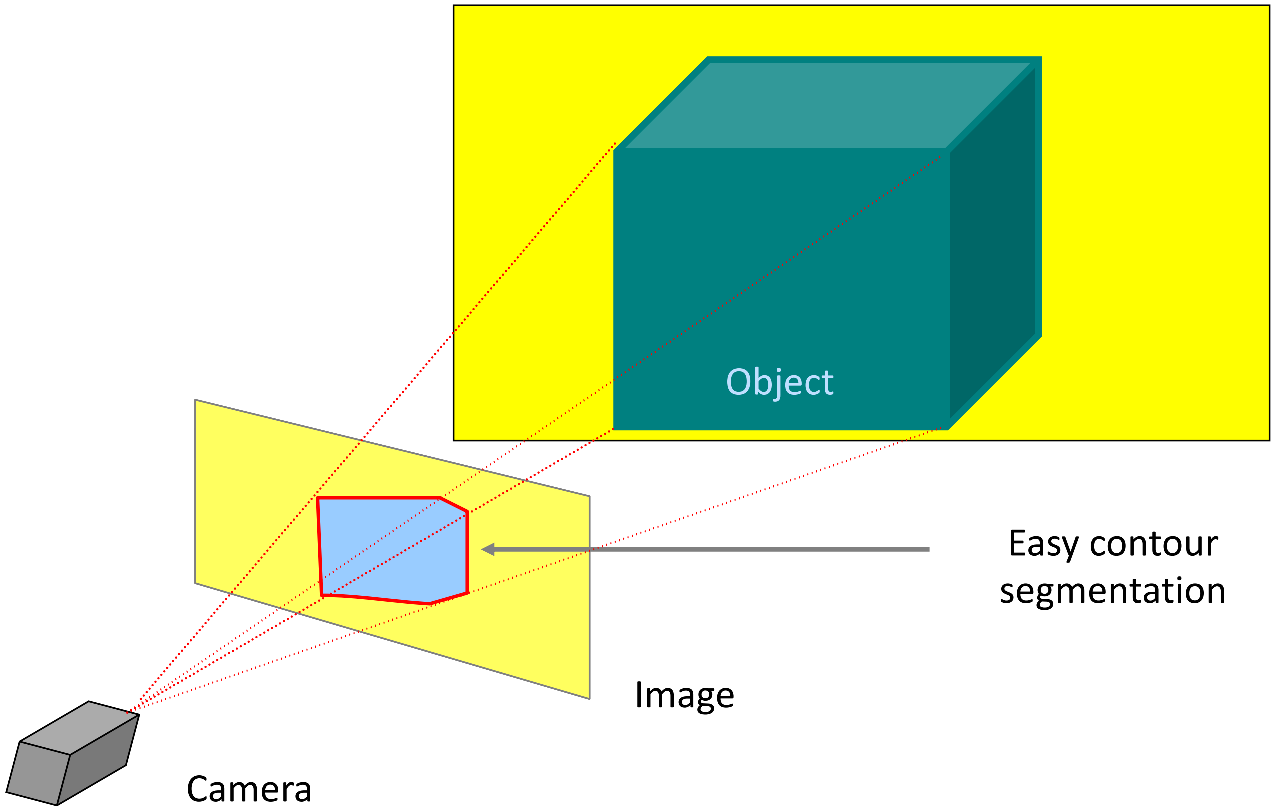


Silhouettes

A silhouette is defined as the area enclosed by the apparent contours

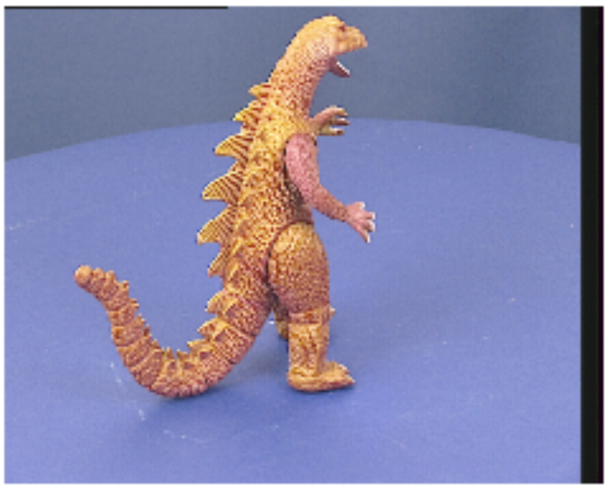


Detecting silhouettes



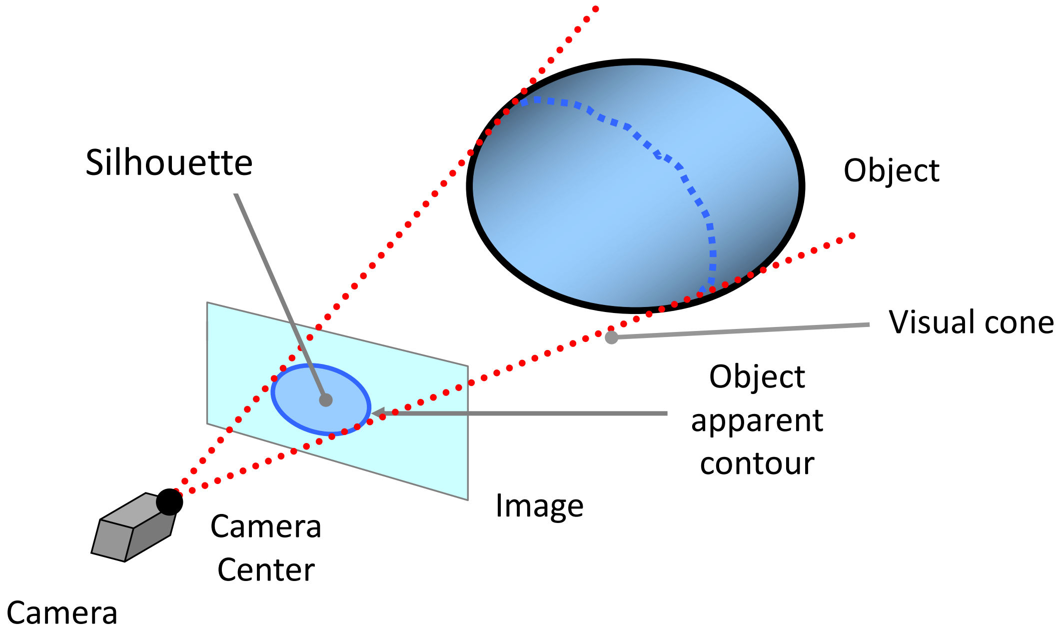
Detecting silhouettes

Original Image

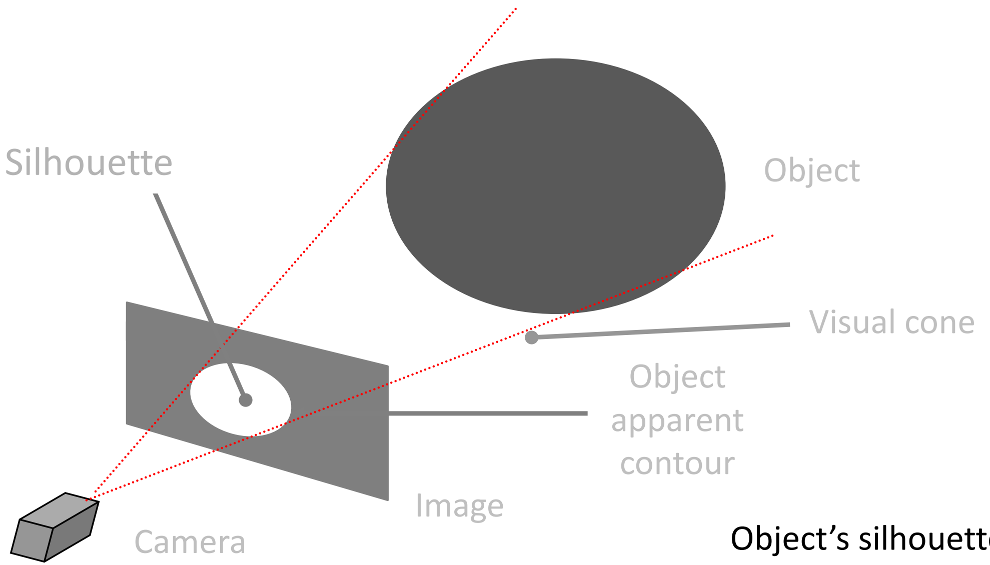


Silhouette

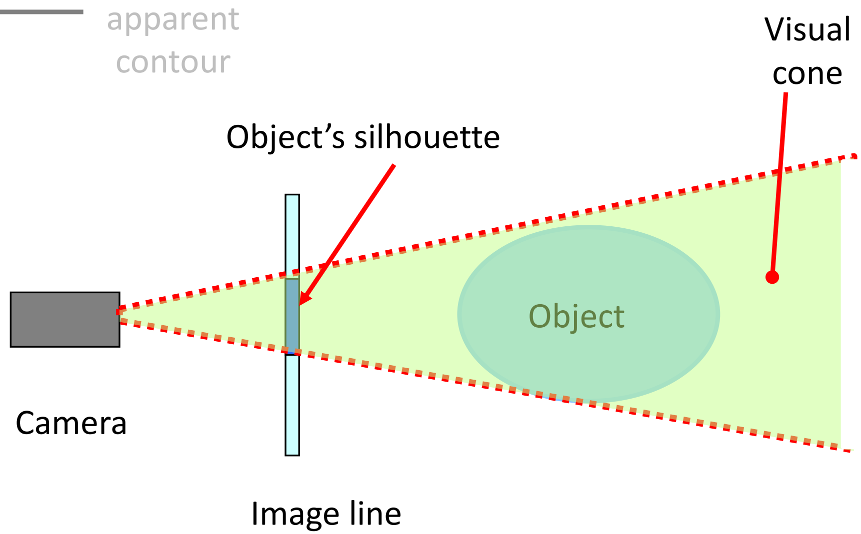
How can we use contours?



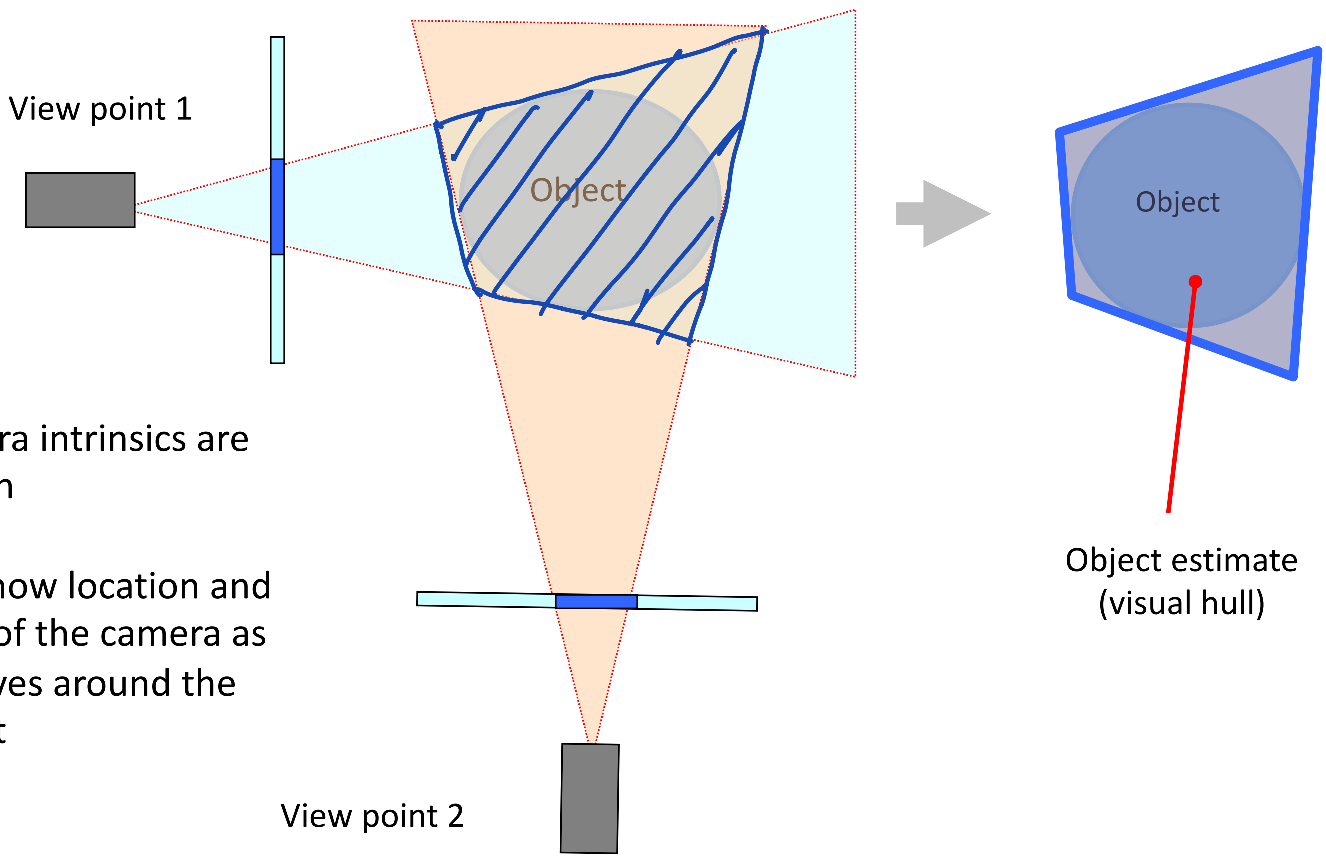
How can we use contours?



2D slice:



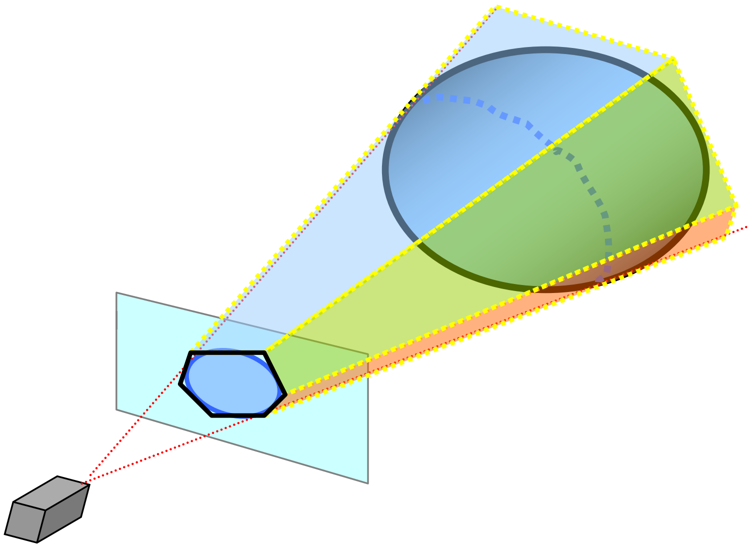
How can we use contours?



- Camera intrinsics are known
- We know location and pose of the camera as it moves around the object

How to perform visual cones intersection?

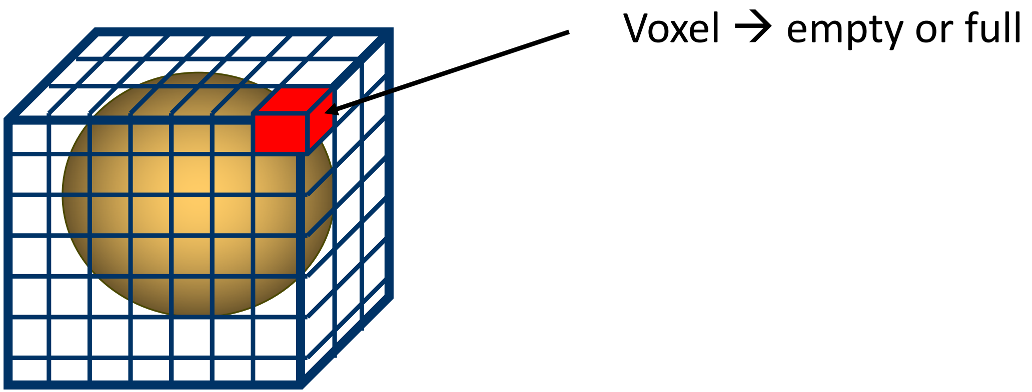
- Decompose visual cone in polygonal surfaces
  (among others: Reed and Allen '99)



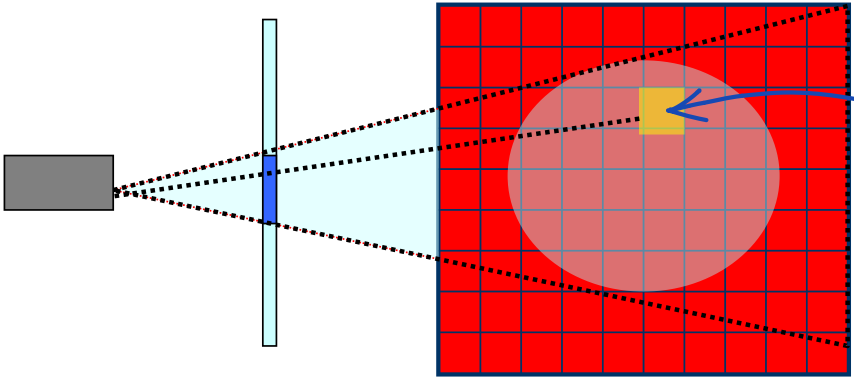
Space carving

[Martin and Aggarwal (1983)]

- Using contours/silhouettes in volumetric stereo, also called **space carving**



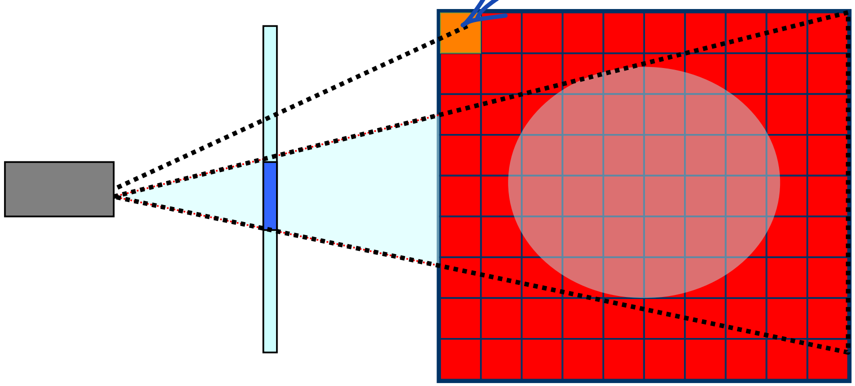
Computing Visual Hull in 2D



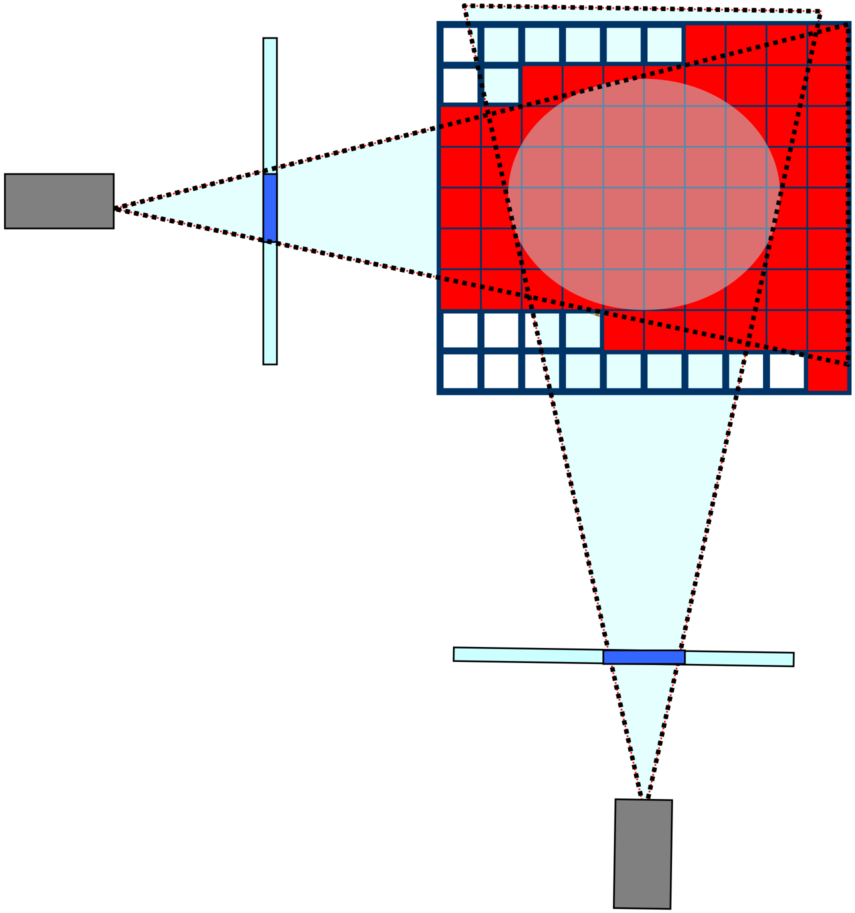
Belongs to the object

Computing Visual Hull in 2D

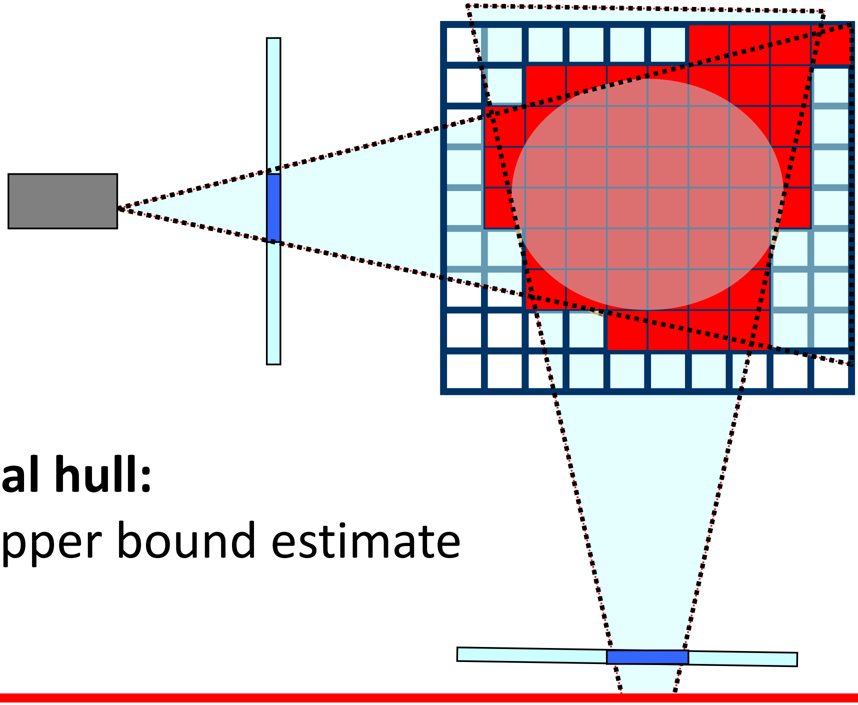
Doesn't belong to the object



Computing Visual Hull in 2D



Computing Visual Hull in 2D

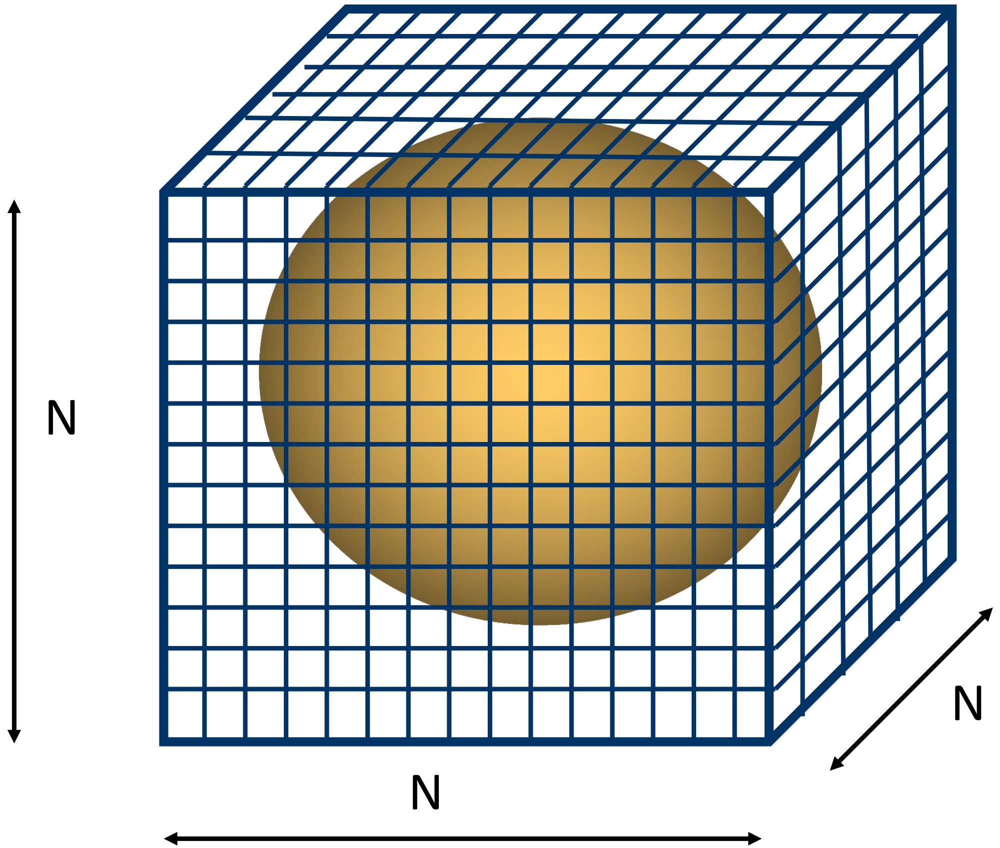


Visual hull:
an upper bound estimate

Consistency:

A voxel must be projected into a silhouette in each image

Space carving has complexity …



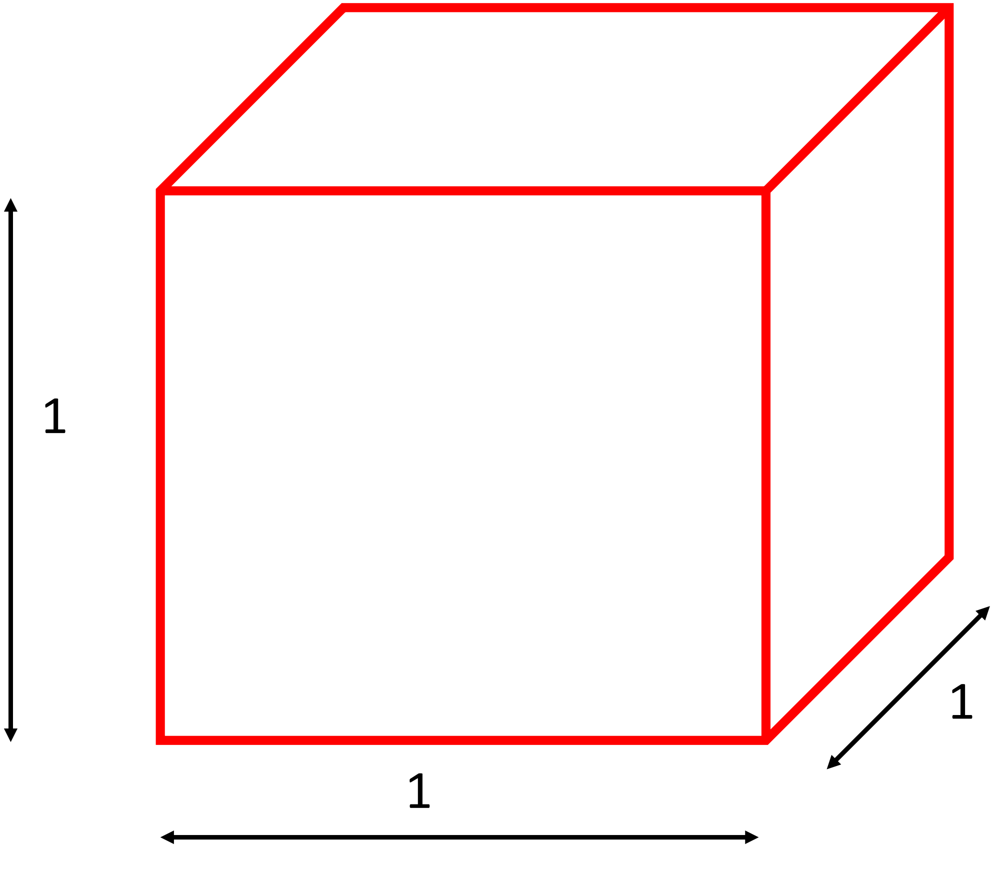
$O(N^3)$

Any suggestion for speeding this process up?

Octrees!

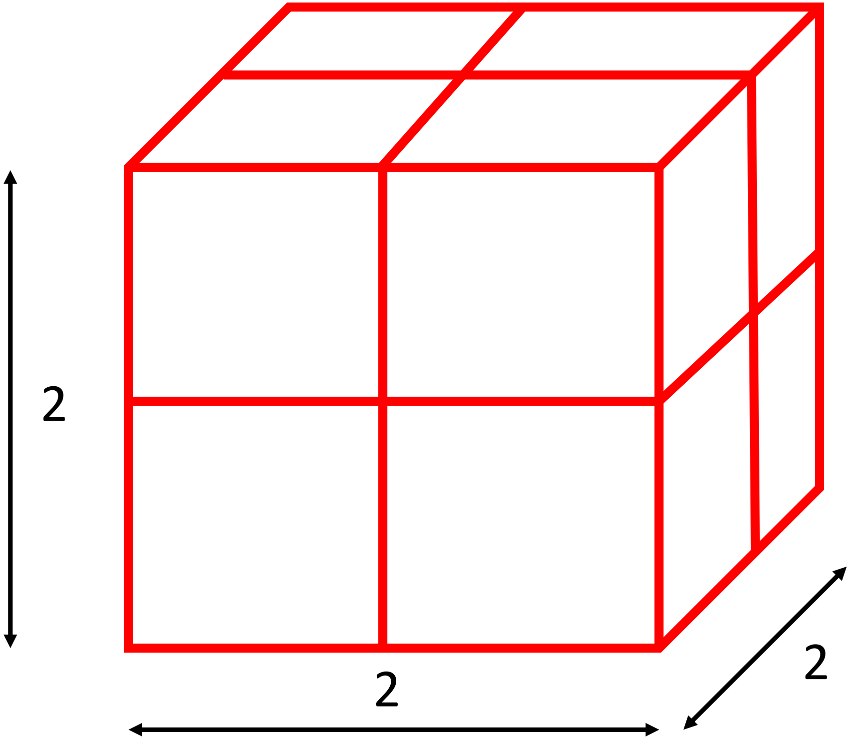
(Szeliski ’93)

Complexity Reduction: Octrees

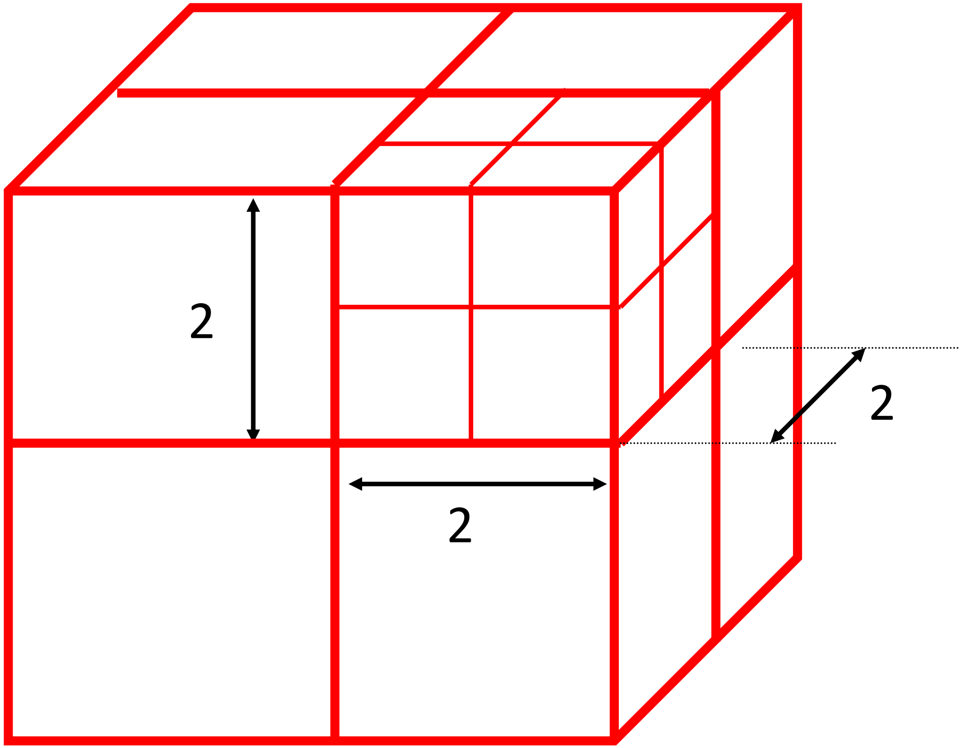


Complexity Reduction: Octrees

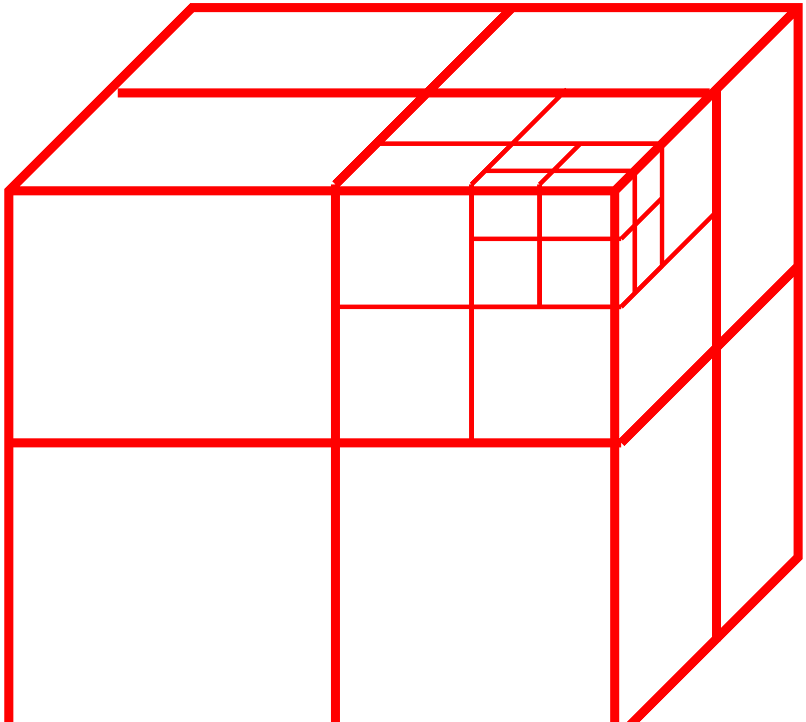
- Subdiving volume in sub-volumes of progressive smaller size



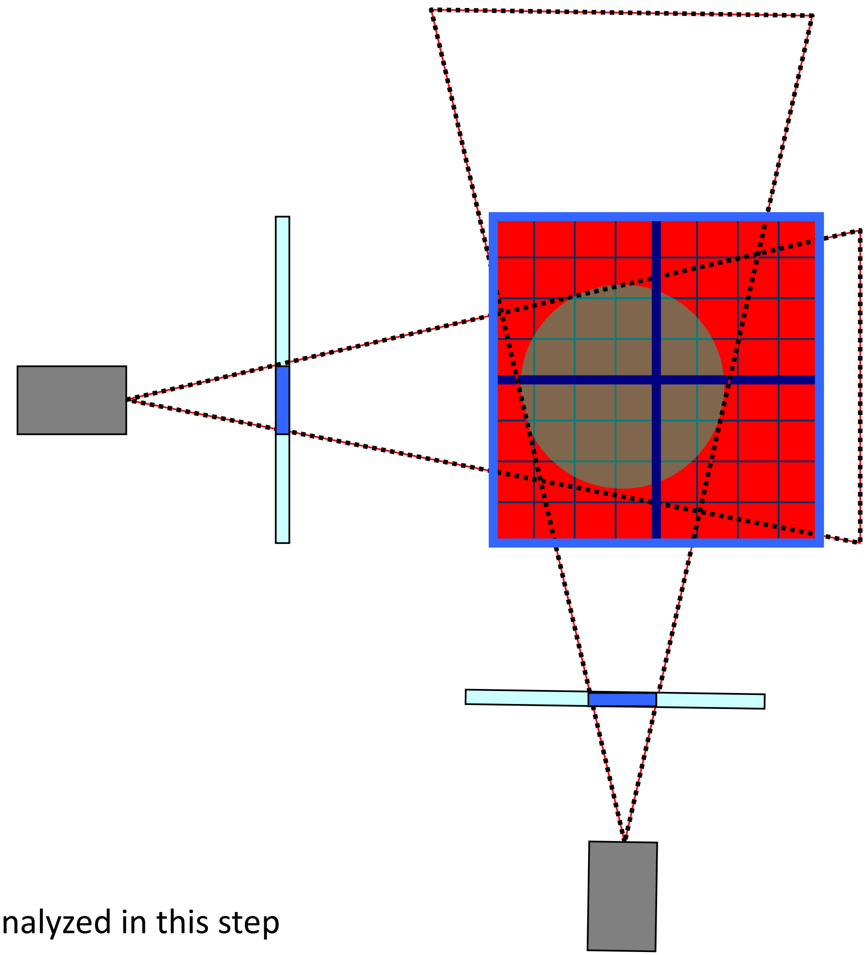
Complexity Reduction: Octrees



Complexity Reduction: Octrees

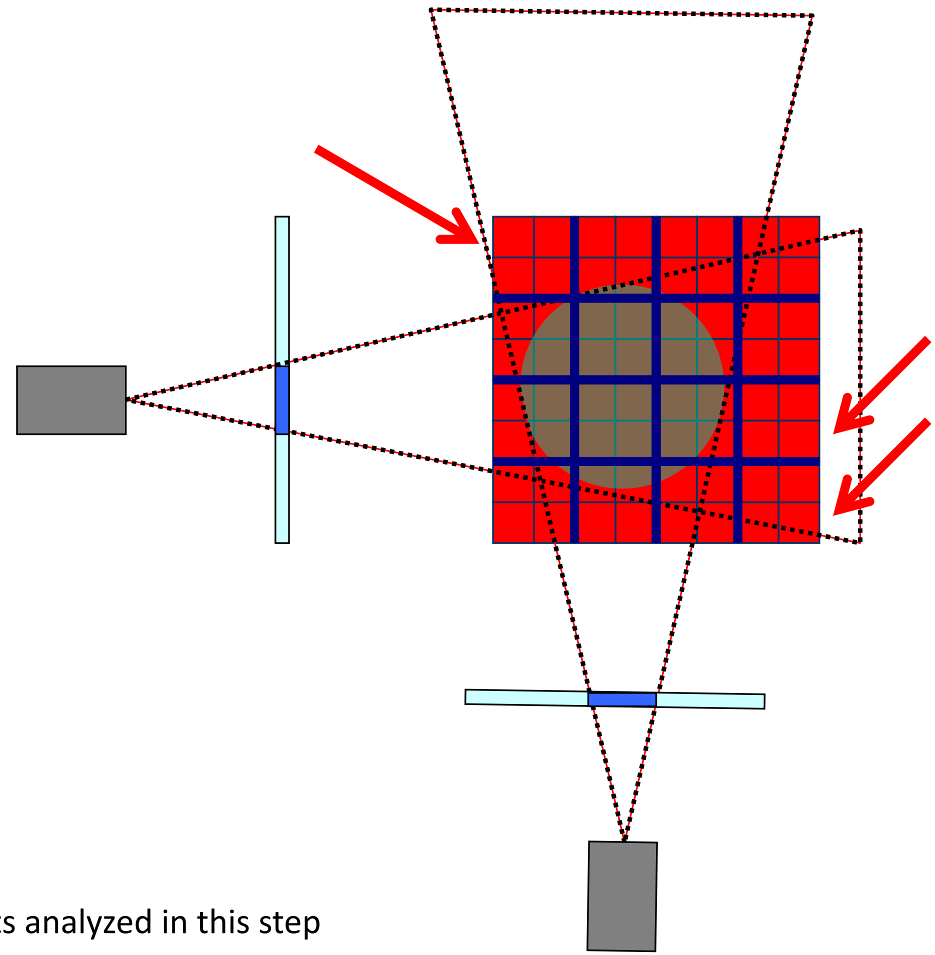


Complexity reduction: 2D example



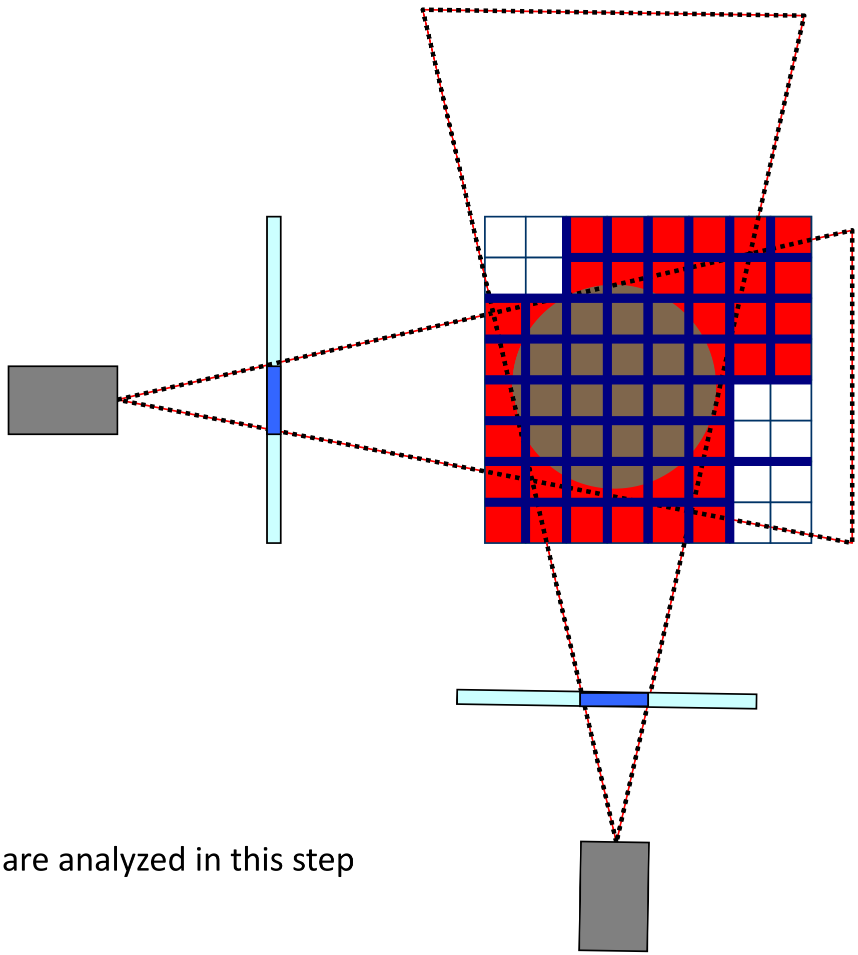
4 elements analyzed in this step

Complexity reduction: 2D example



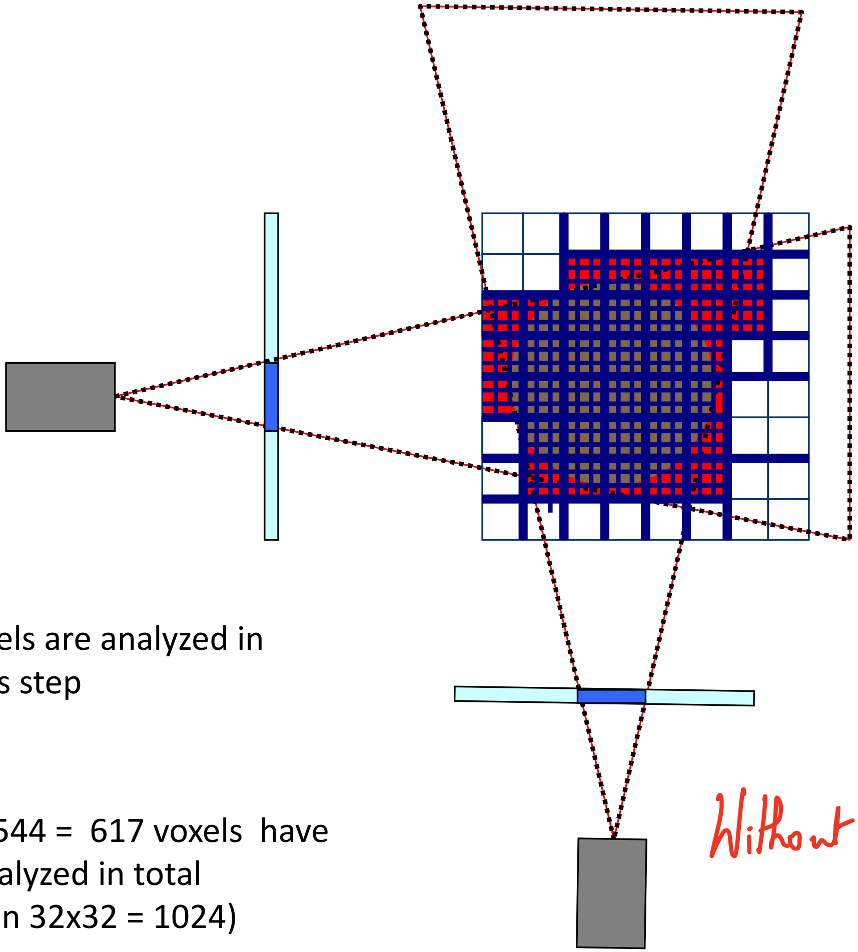
16 elements analyzed in this step

Complexity reduction: 2D example



52 elements are analyzed in this step

Complexity reduction: 2D example



16x34=544 voxels are analyzed in this step

1+ 4 + 16 + 52 + 544 = 617 voxels have been analyzed in total
(rather than 32x32 = 1024)

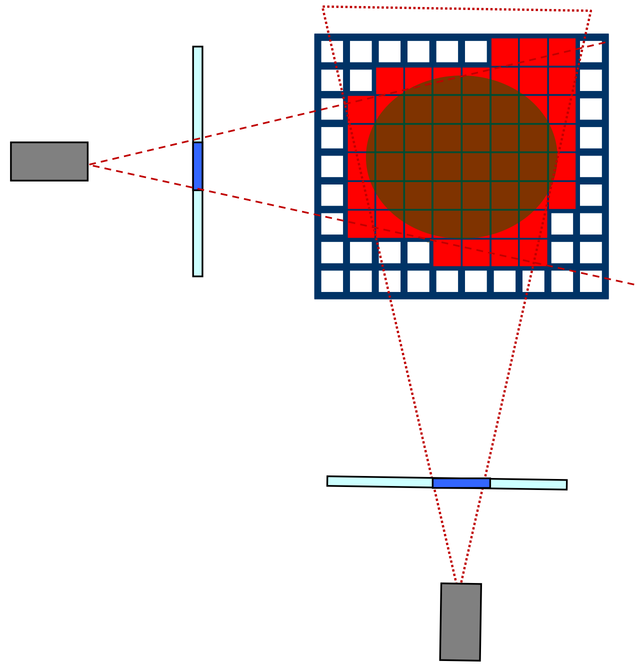
Without Octree $\rightarrow 8^3 = 1024$

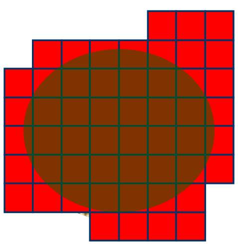
Advantages of Space carving

- Robust and simple
- No need to solve for correspondences

Limitations of Space carving

- Accuracy function of number of views



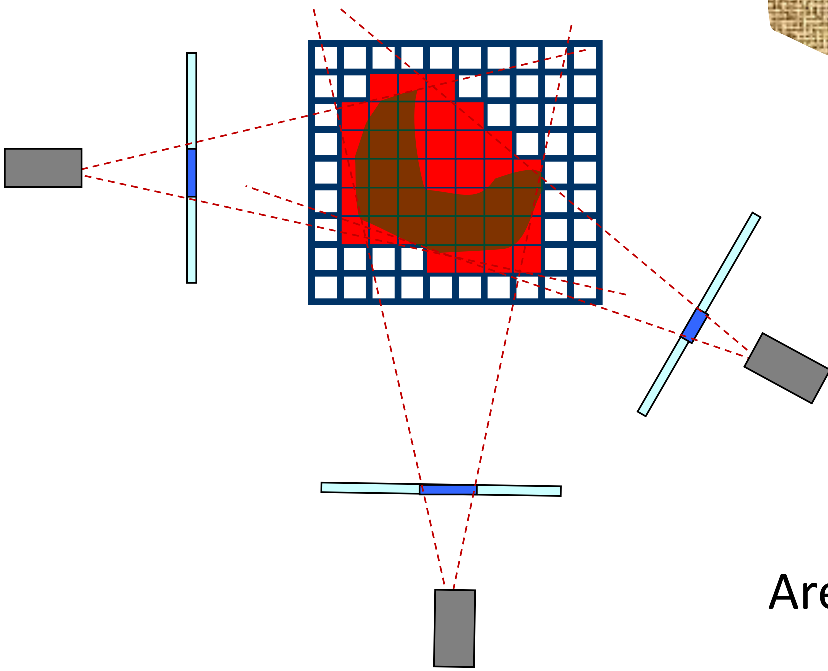


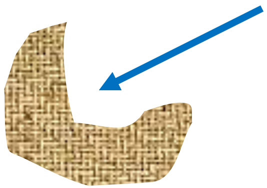
Not a good estimate

What else?

Limitations of Space carving

- Concavities are not modeled



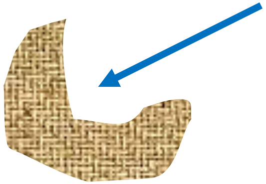


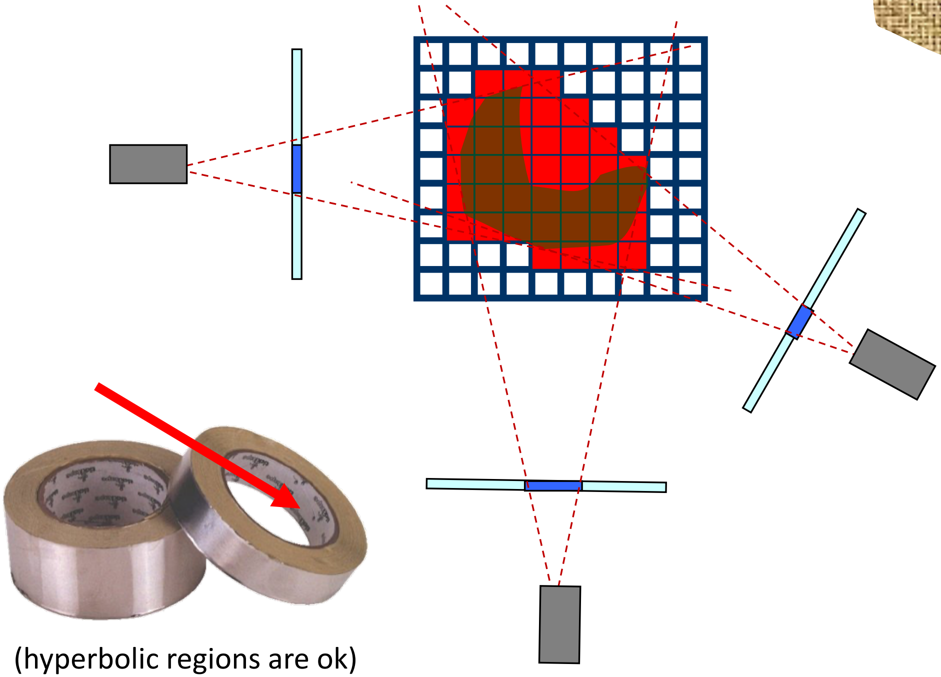
For 3D objects:
Are all types of concavities problematic?

Limitations of Space carving

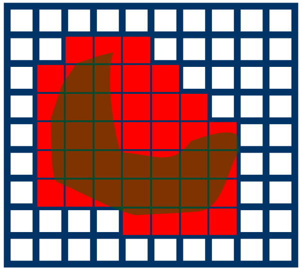
- Concavities are not modeled

Concavity





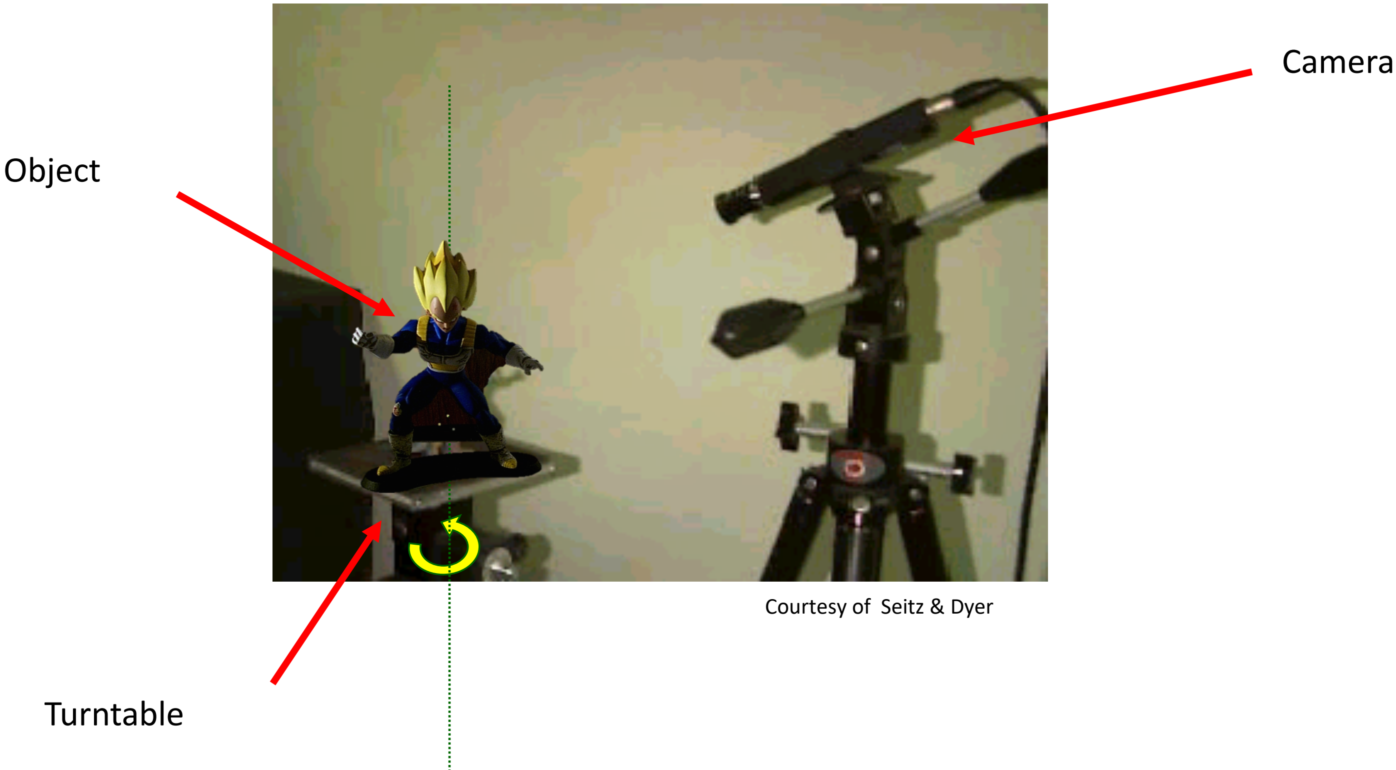
Visual hull



(hyperbolic regions are ok)

-- see Laurentini (1995)

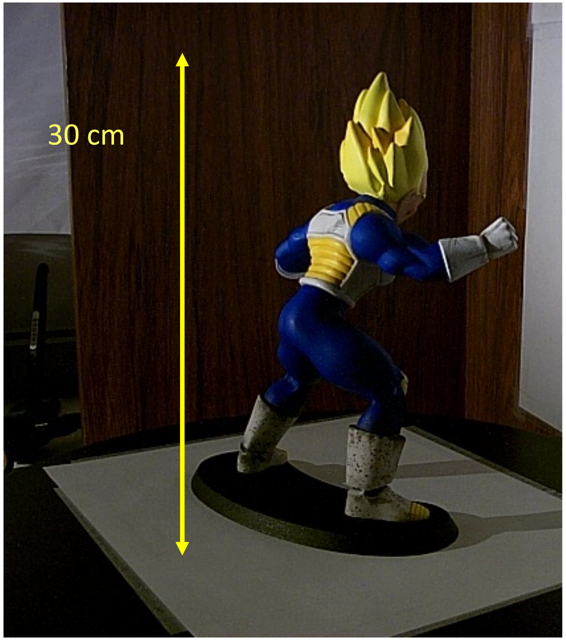
Space carving: A Classic Setup

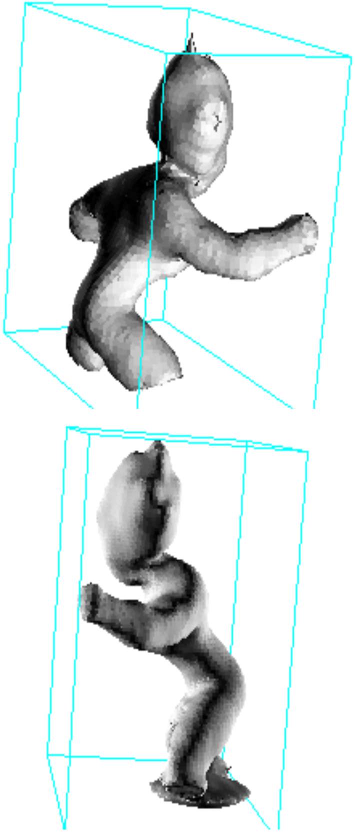


Courtesy of Seitz & Dyer

Space carving: A Classic Setup

Space carving: Experiments

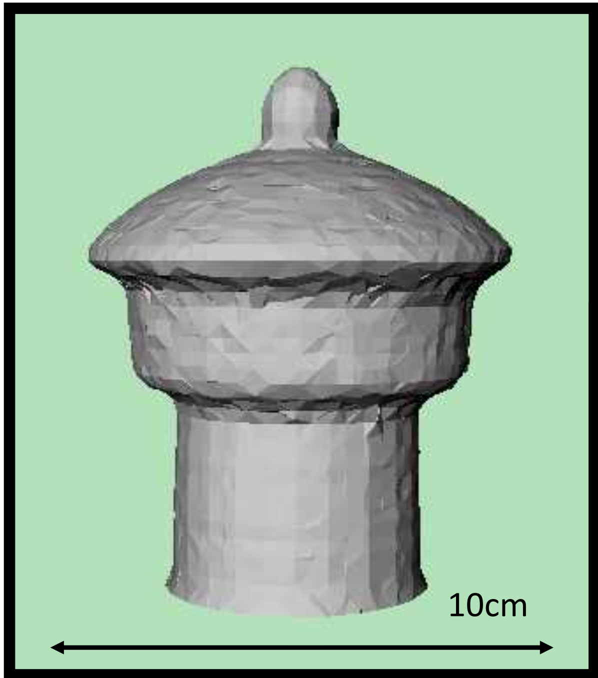




24 poses (15^{O})

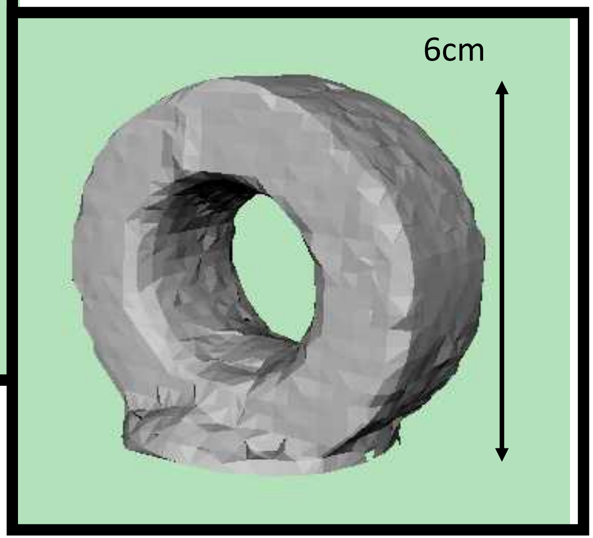
voxel size = 1mm

Space carving: Experiments



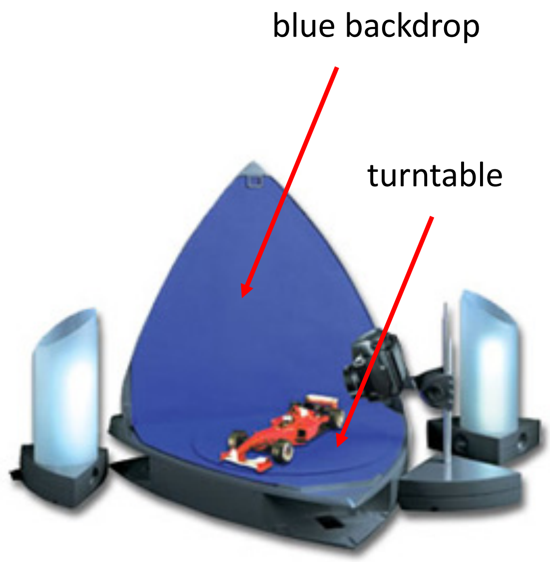
24 poses (15^{O})

voxel size = 2mm



Space carving: Conclusions

- Robust
- Produce conservative estimates
- Concavities can be a problem
- Low-end commercial 3D scanners



Lecture 8

Active stereo & Volumetric stereo

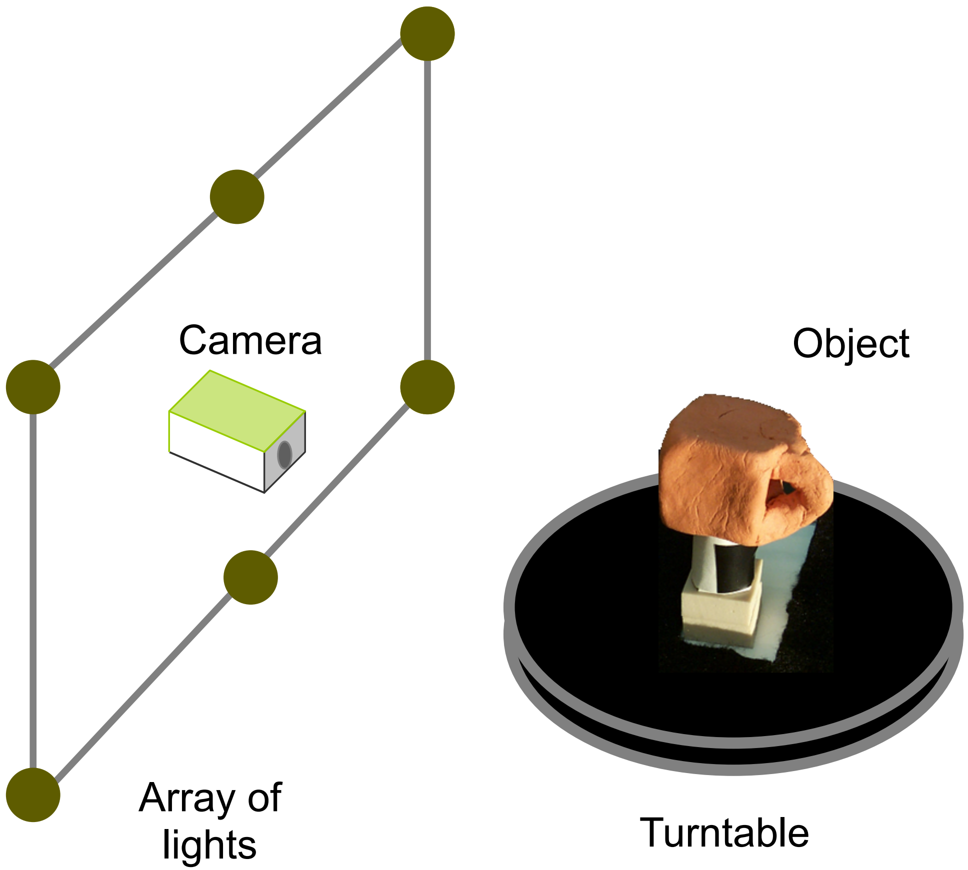
- Active stereo
  - Structured lighting
  - Depth sensing
- Volumetric stereo:
  - Space carving
  - Shadow carving
  - Voxel coloring

Shape from Shadows

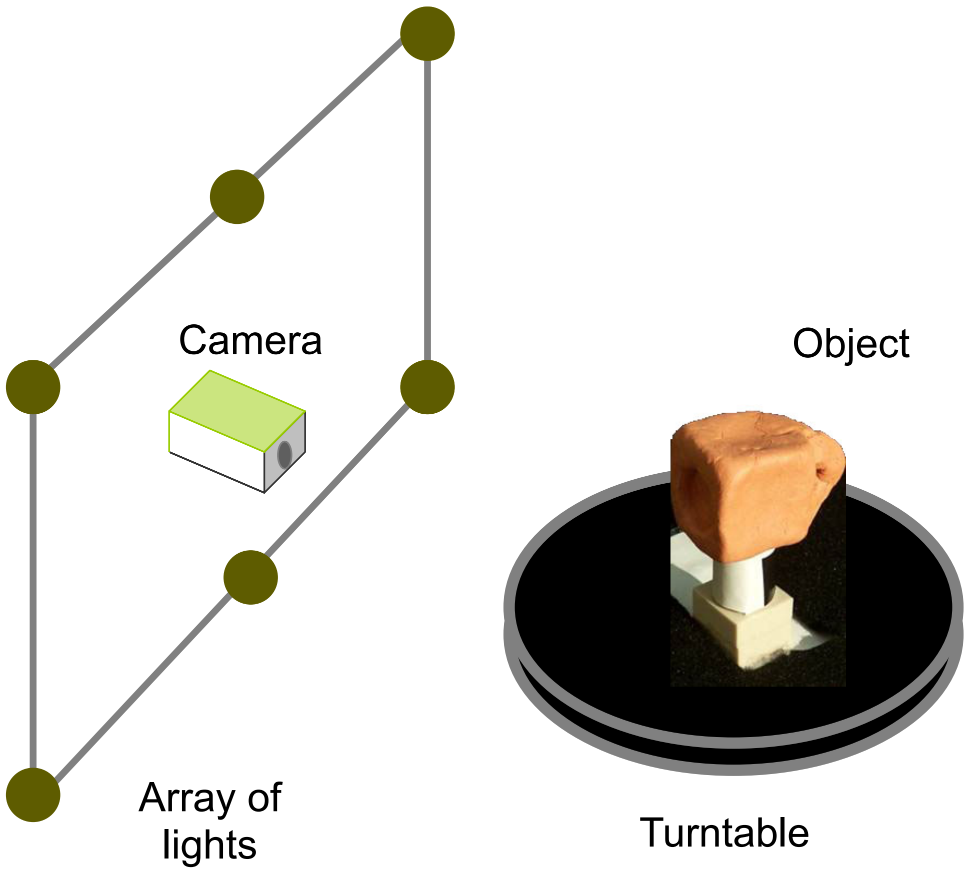
- Self-shadows are visual cues for shape recovery

Self-shadows indicate concavities
(no modeled by contours)

Shadow carving: The Setup

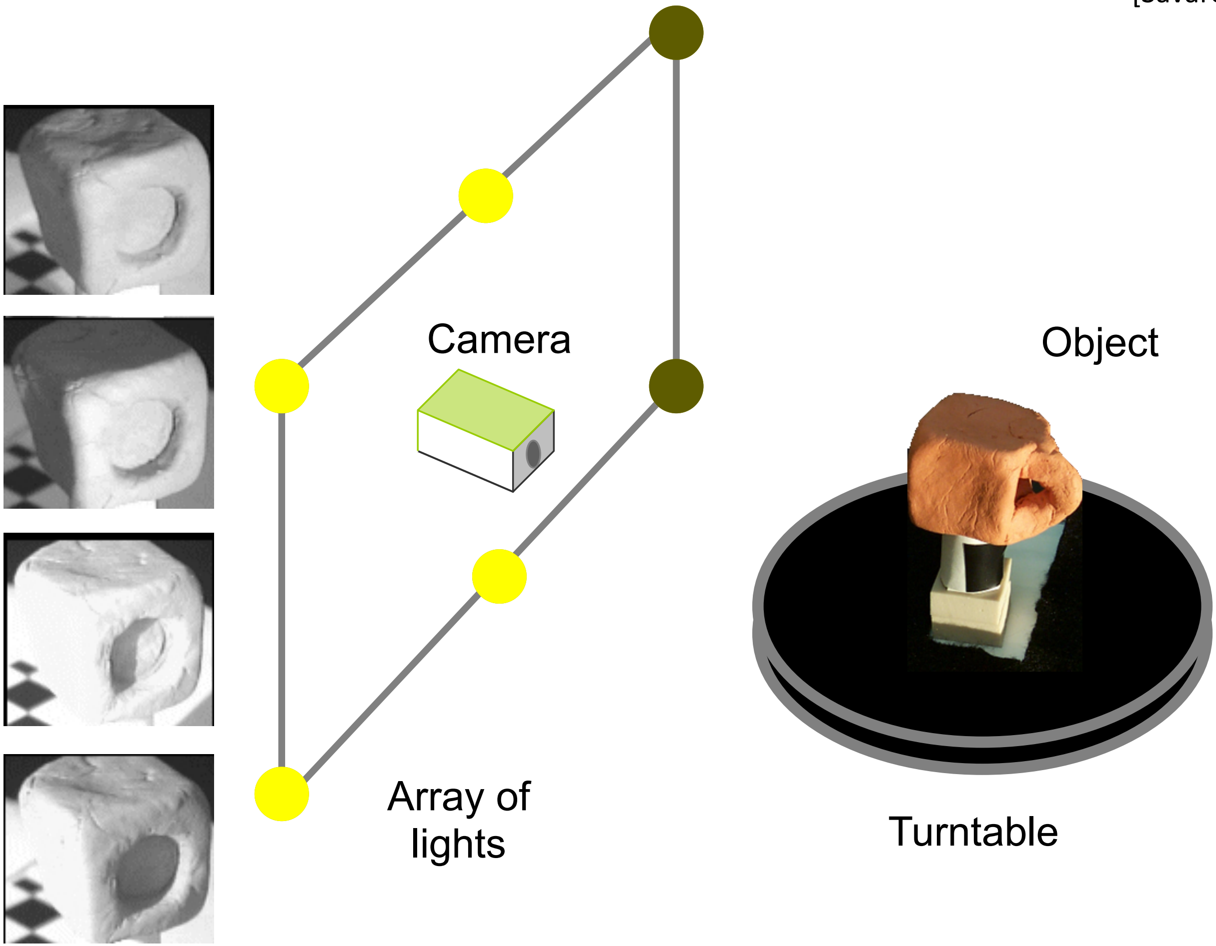


Shadow carving: The Setup



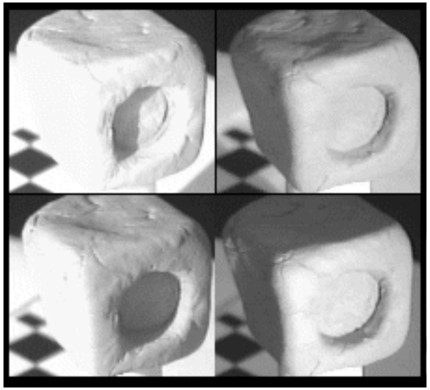
Shadow carving: The Setup

[Savarese et al '01]



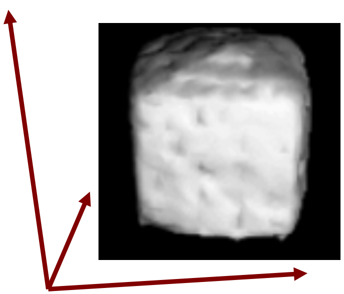
Shadow carving

[Savarese et al. 2001]

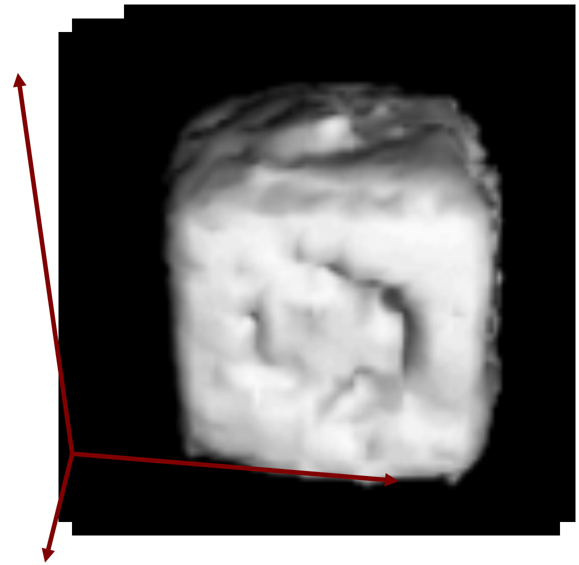


Self-shadows

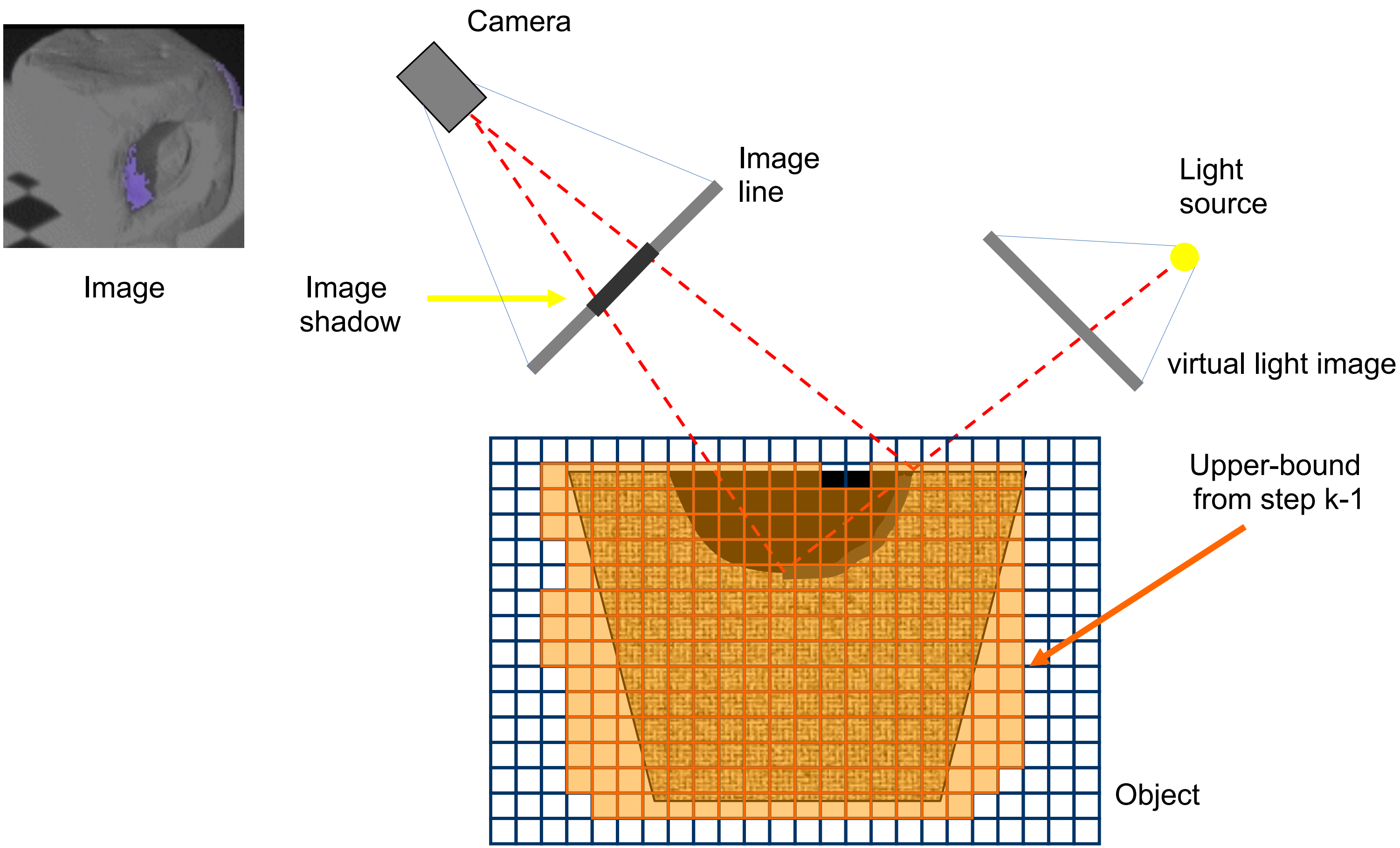
+



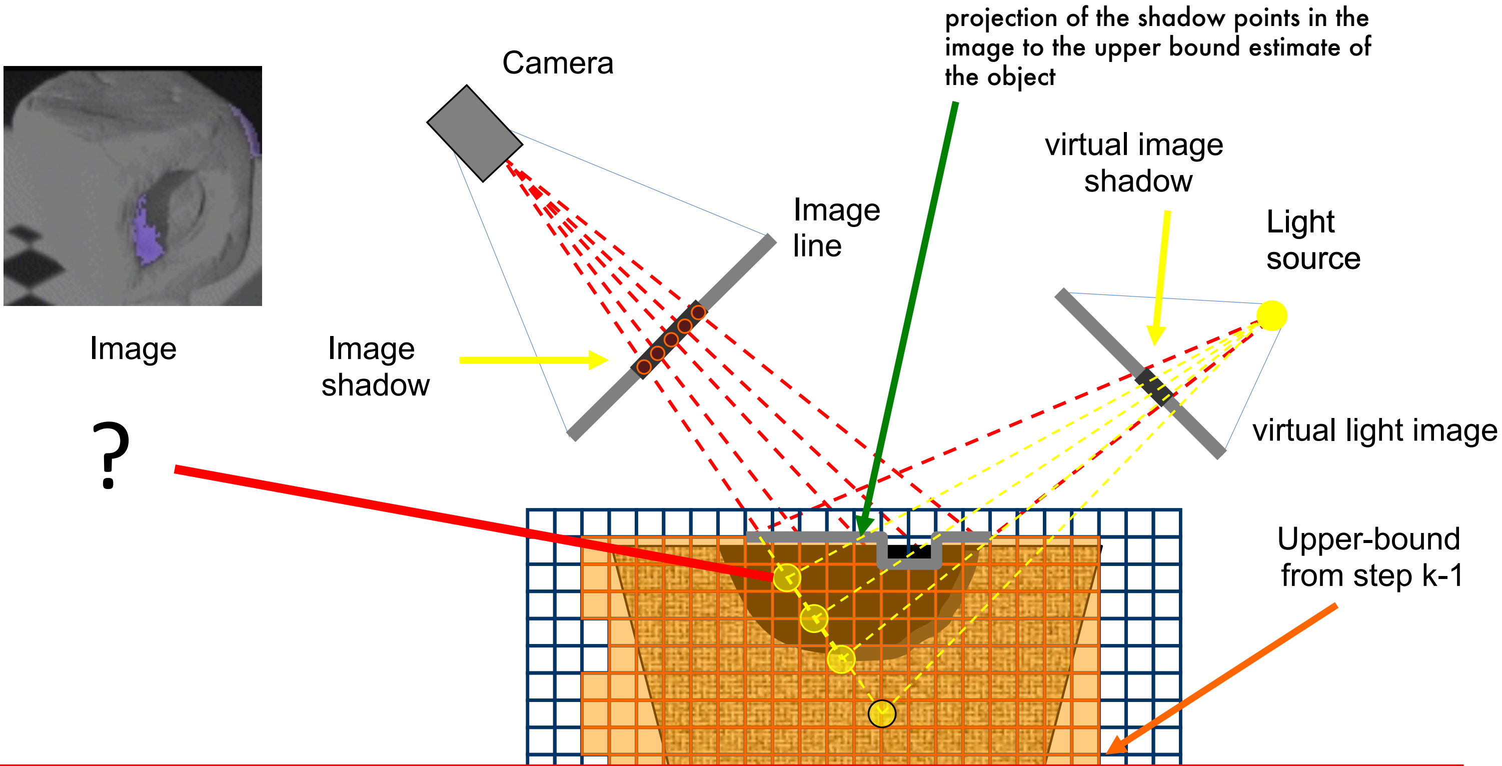
Object’s upper bound



Algorithm: Step k



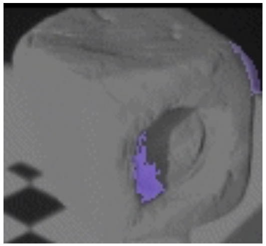
Algorithm: Step k



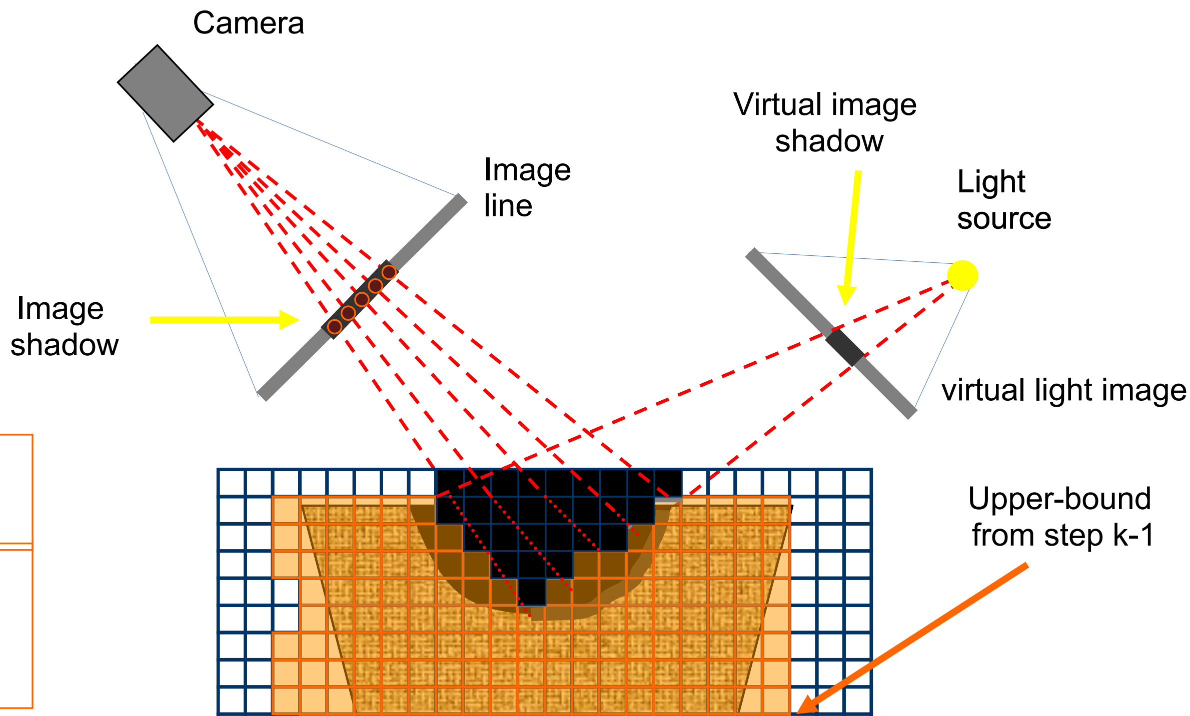
Theorem:

A voxel that projects into an image shadow AND an virtual image shadow cannot belong to the object.

Algorithm: Step k



Image



Properties:

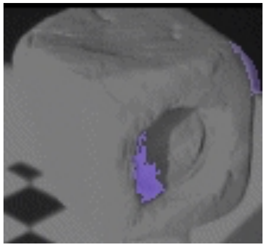
No further volume can be removed

Carving process always conservative

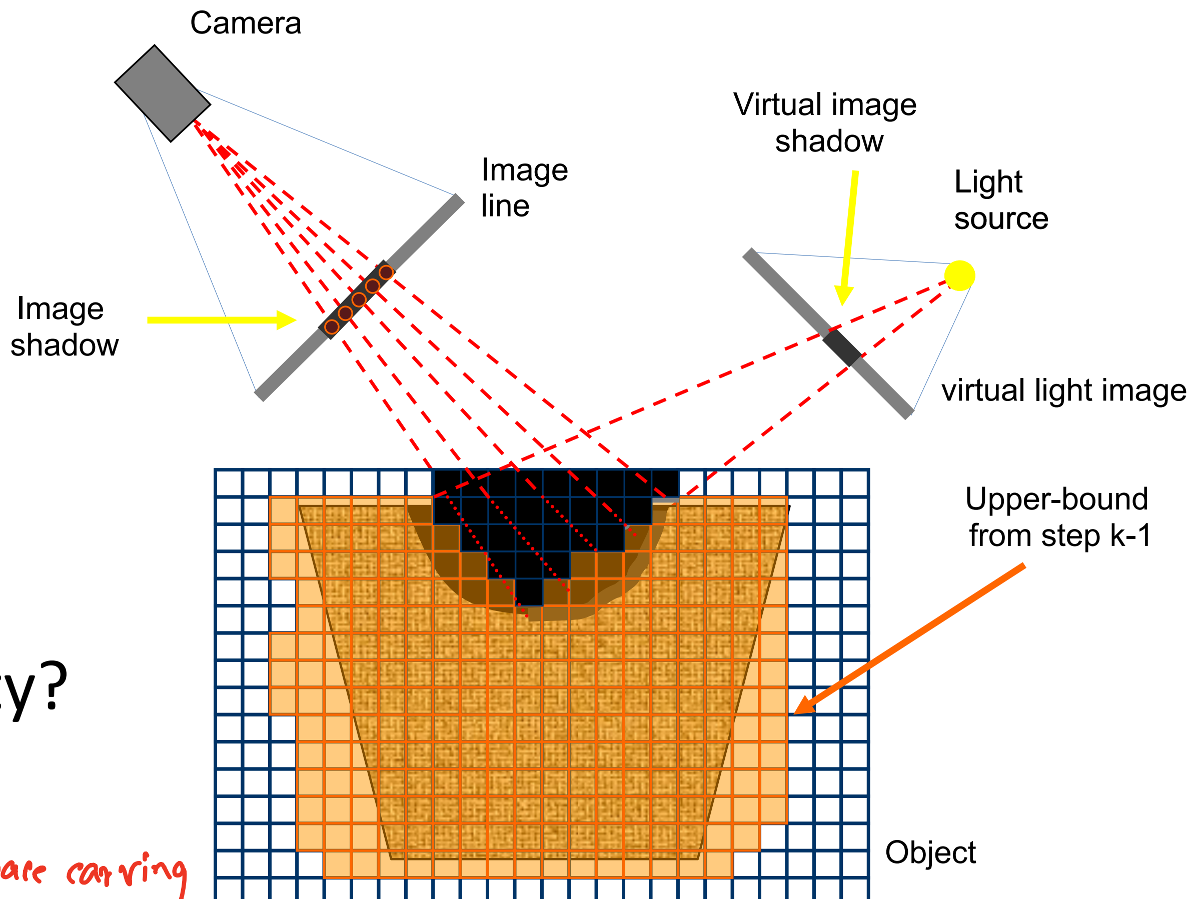
Consistency:

In order for a voxel to be removed it must project into both image shadow and virtual image shadow

Algorithm: Step k



Image

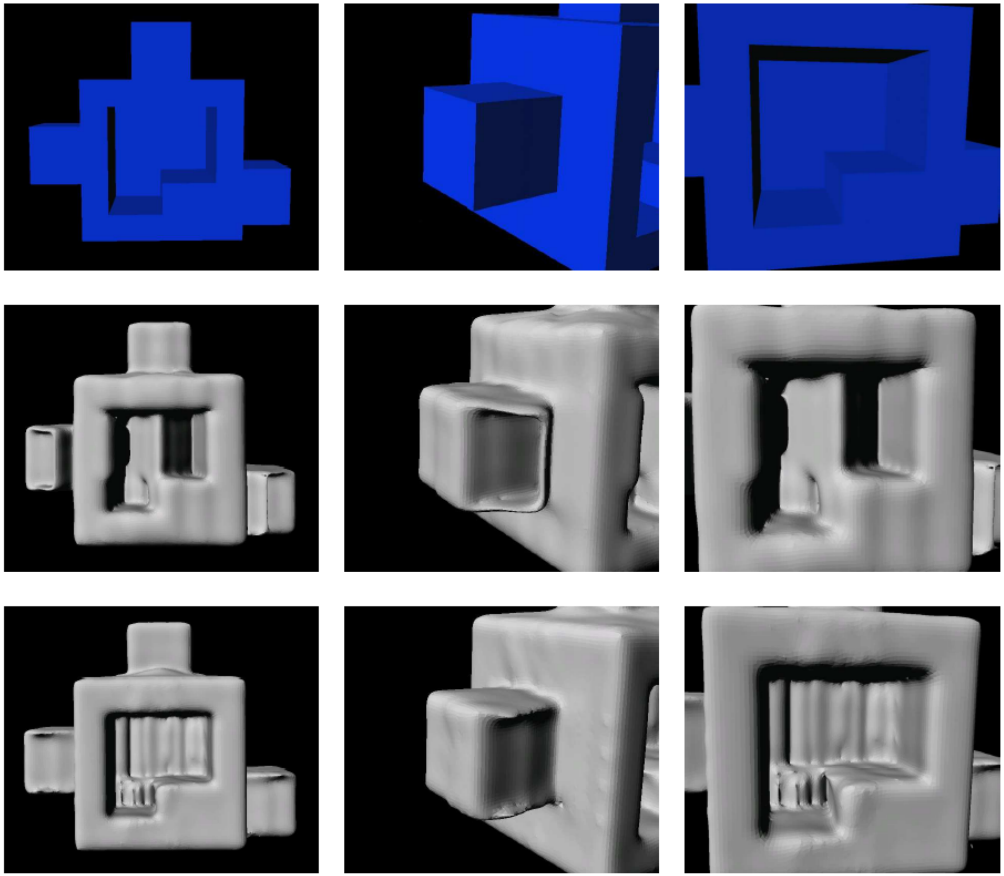


Complexity?

$O(2N^3)$

Same as Space carving
but for two image planes.

Simulating the System

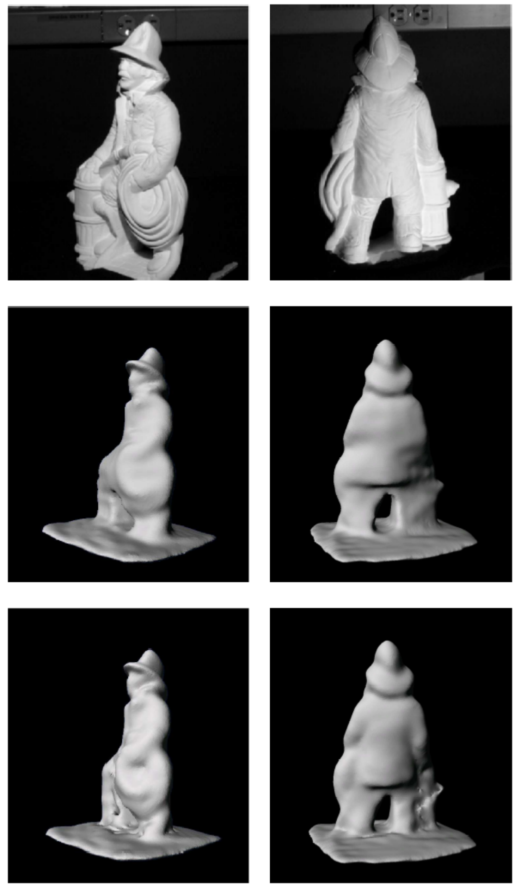


- 24 positions
- 4 lights

- 72 positions
- 8 lights

Results

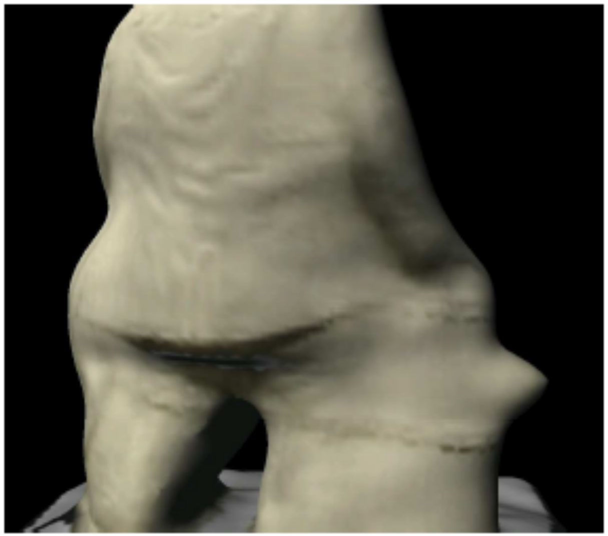
- 16 positions
- 4 lights



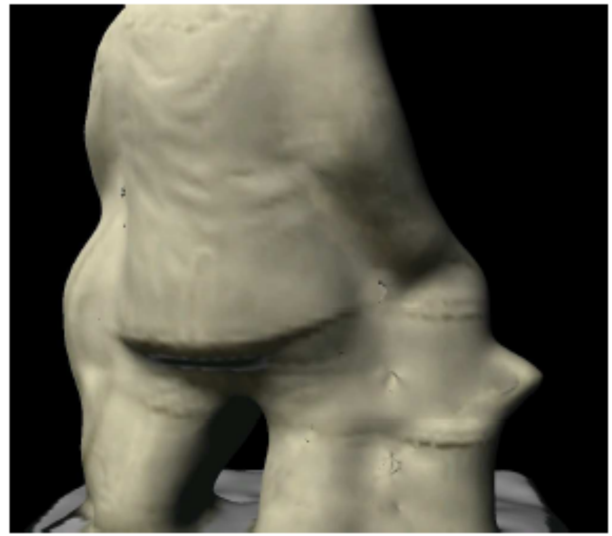
Space carving

Shadow carving

Results



Space carving



Shadow carving

Shadow carving: Summary

- Produces a conservative volume estimate
- Accuracy depending on view point and light source number
- Limitations with reflective & low albedo regions

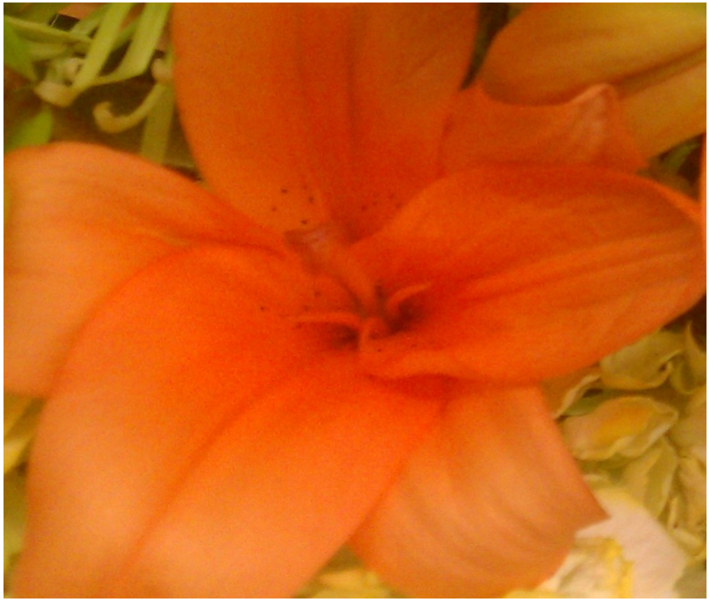
Lecture 8

Active stereo & Volumetric stereo

- Active stereo
  - Structured lighting
  - Depth sensing
- Volumetric stereo:
  - Space carving
  - Shadow carving
  - Voxel coloring

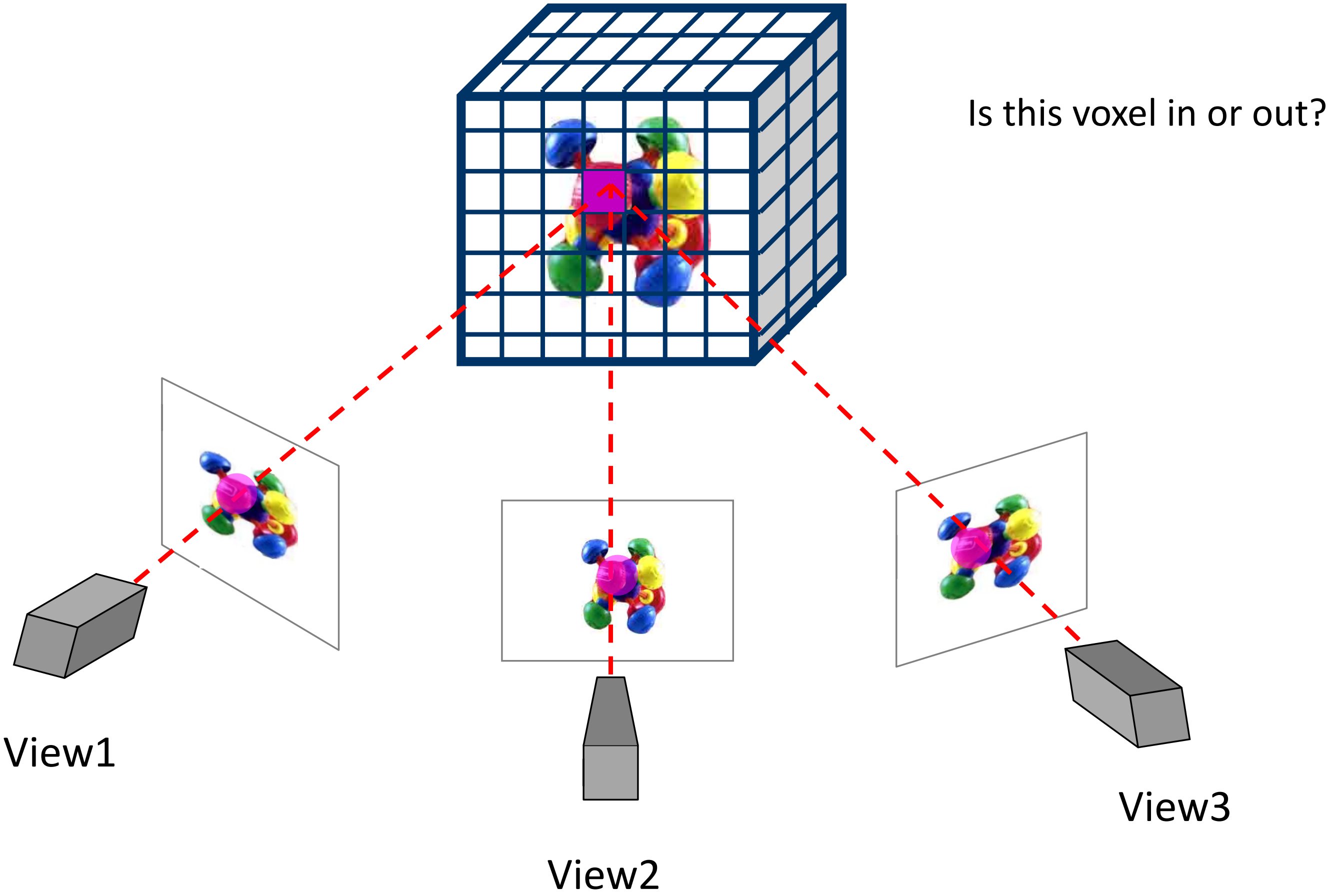
Voxel Coloring

[Seitz & Dyer (‘97)]
[R. Collins (Space Sweep, ’96)]

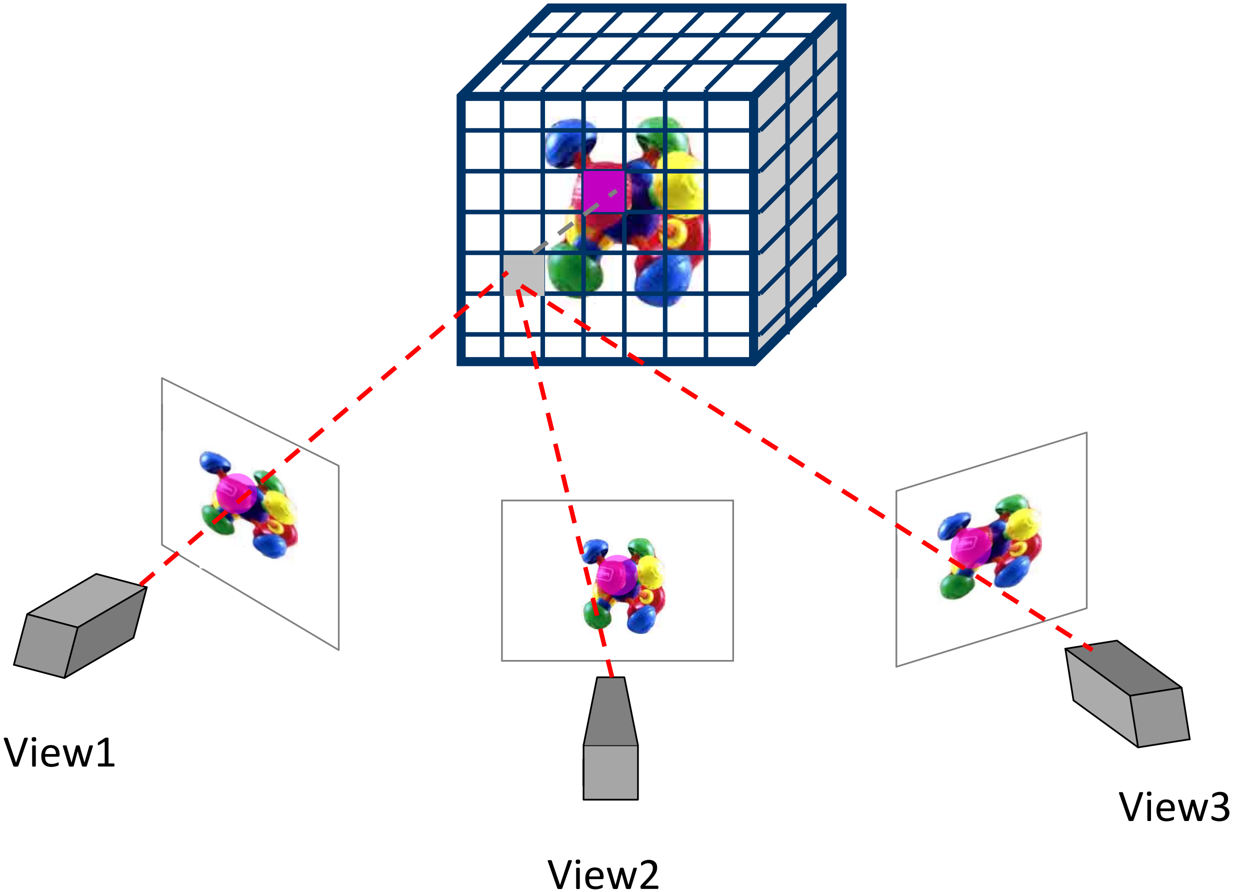


- Color/photo-consistency
- Jointly model structure and appearance

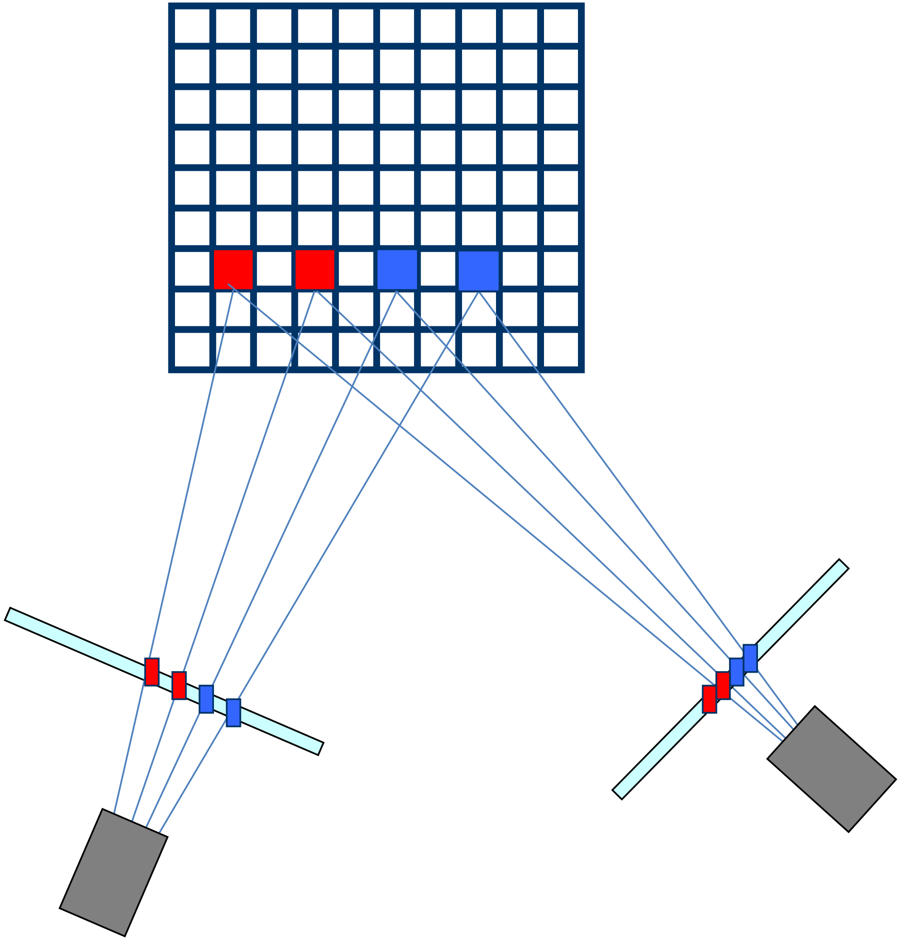
Basic Idea



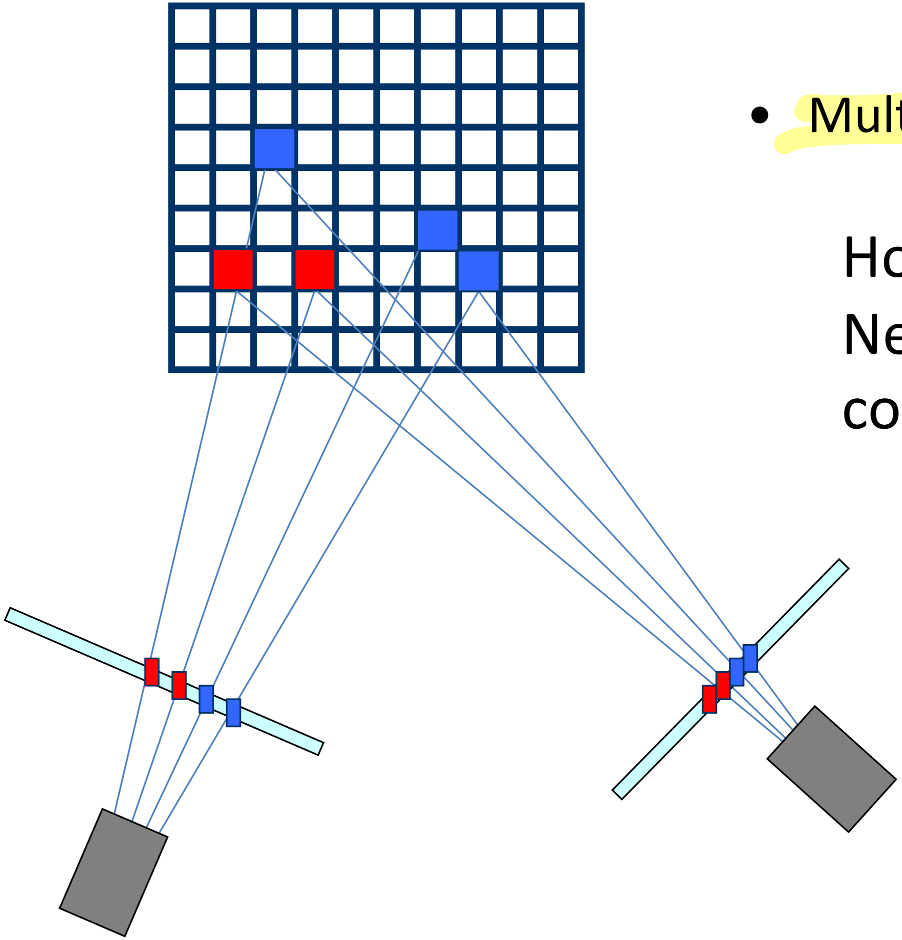
Basic Idea



Uniqueness



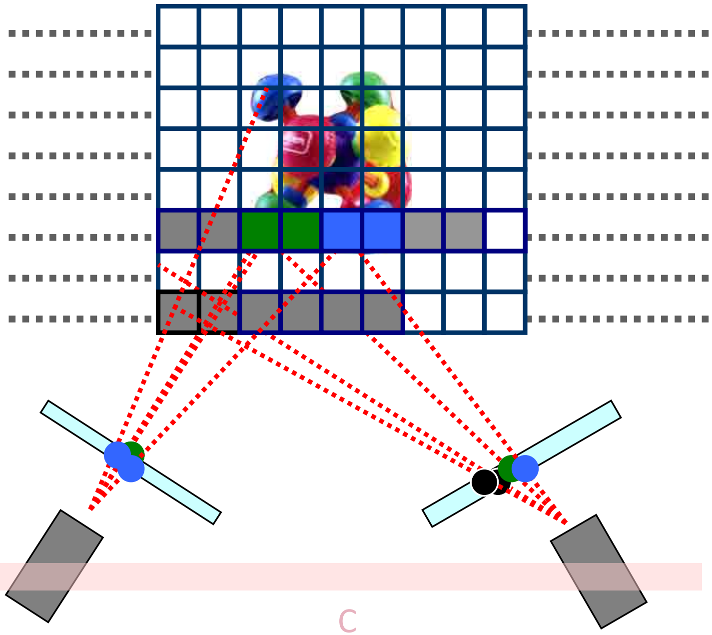
Uniqueness



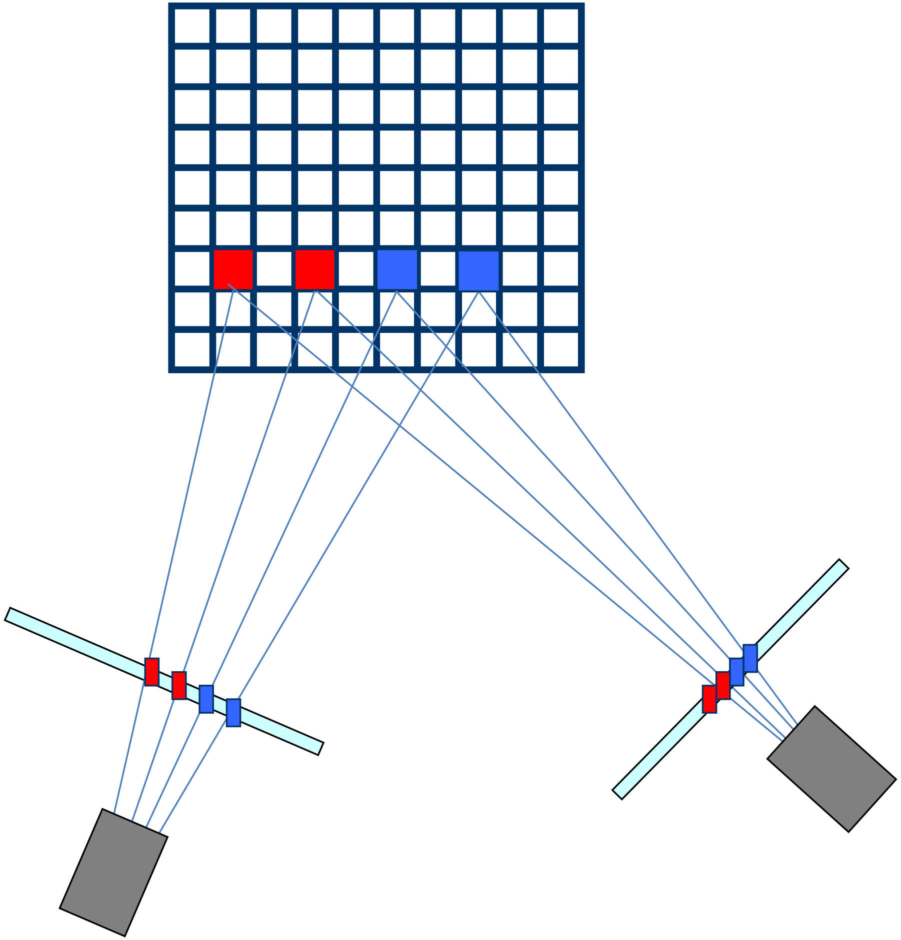
- Multiple consistent scenes

How to fix this?
Need to use a visibility constraint

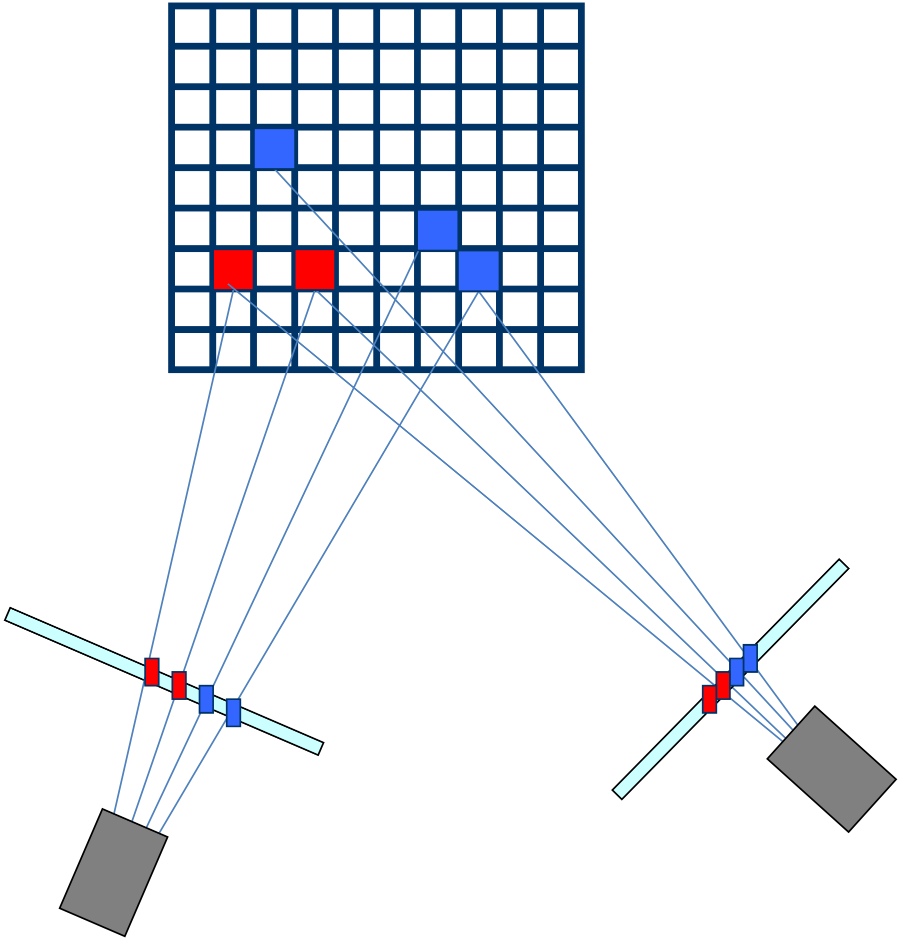
Algorithm for enforcing visibility constraints



The previous ambiguity is removed!



The previous ambiguity is removed!



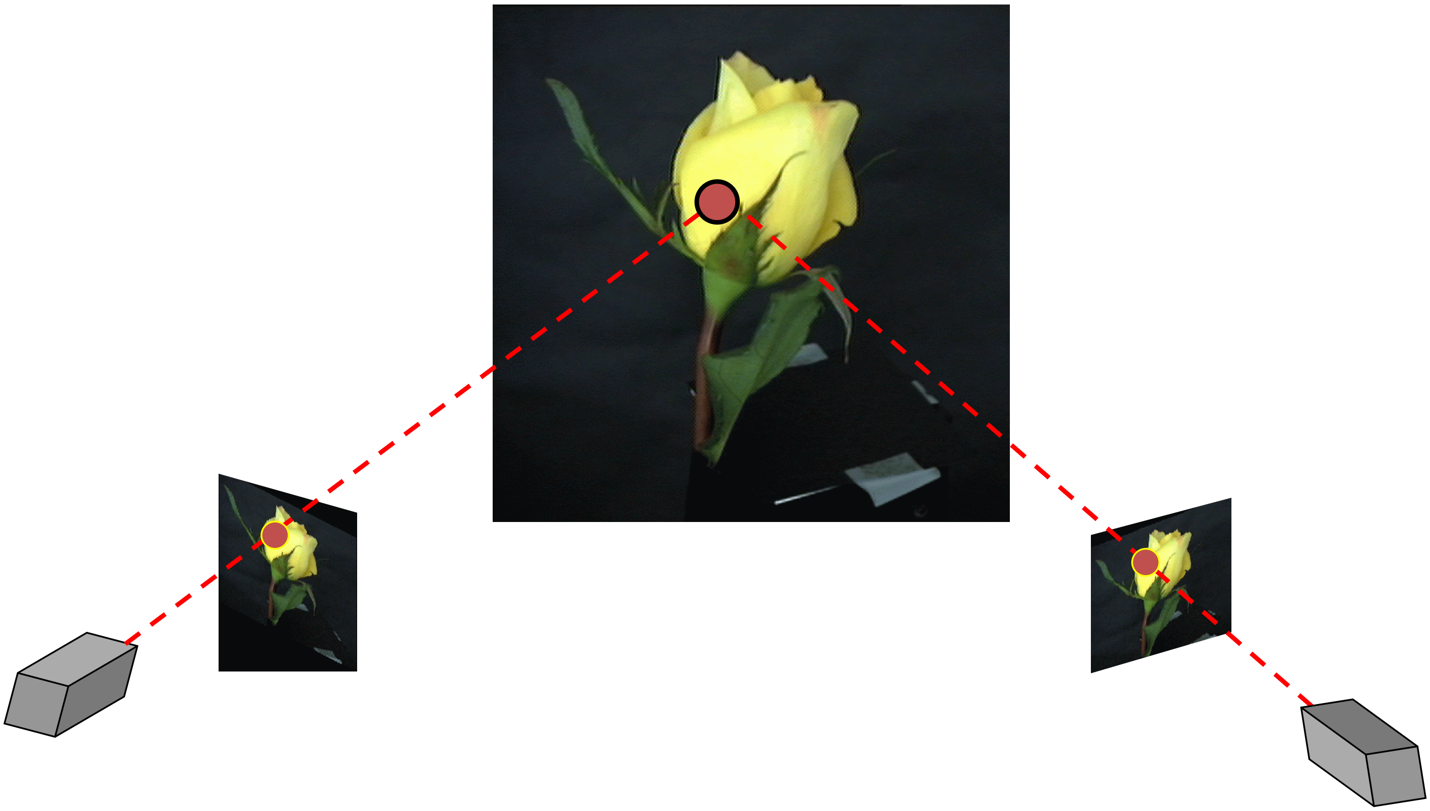
Algorithm Complexity

- Voxel coloring visits each N^3 voxels only once
- Project each voxel into L images

$$\rightarrow O(L\,N^3)$$

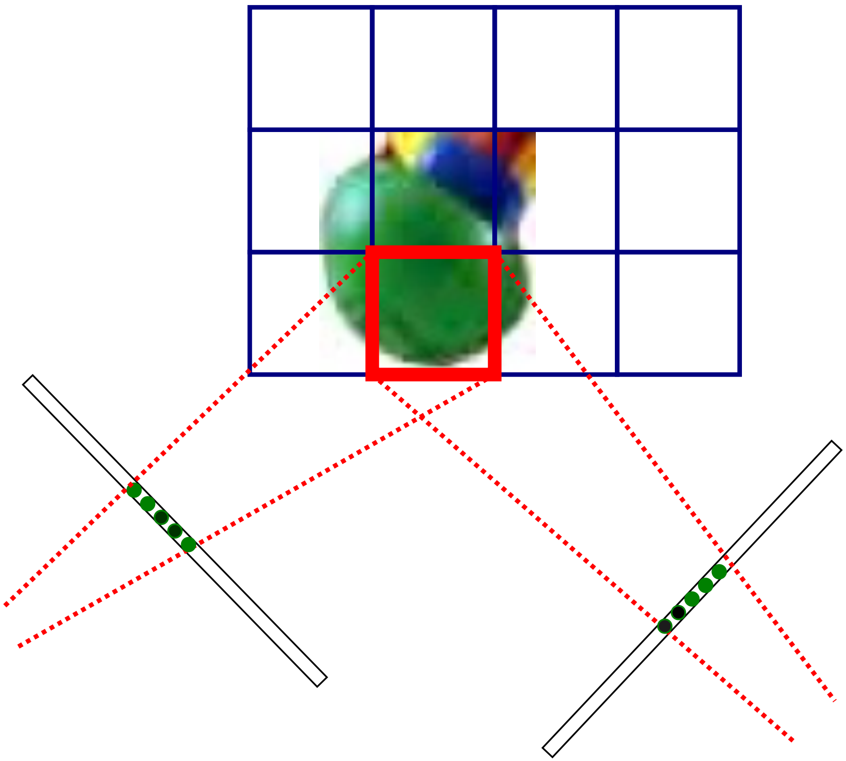
NOTE: not function of the number of colors

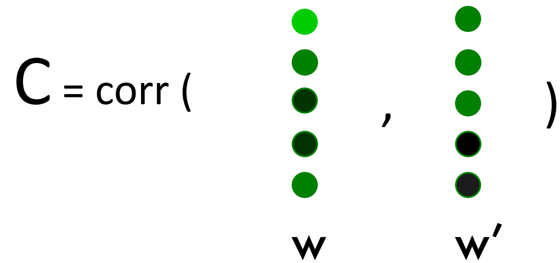
A Critical Assumption: Lambertian Surfaces



Non Lambertian Surfaces

Photoconsistency Test





$$C = \frac{(w-\bar{w})(w'-\bar{w}')}{\|(w-\bar{w})\|\,\|(w'-\bar{w}')\|}$$

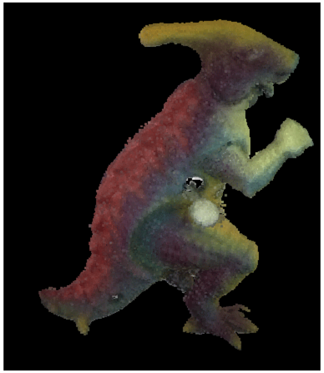
$$\bar{w} = mean(w)$$

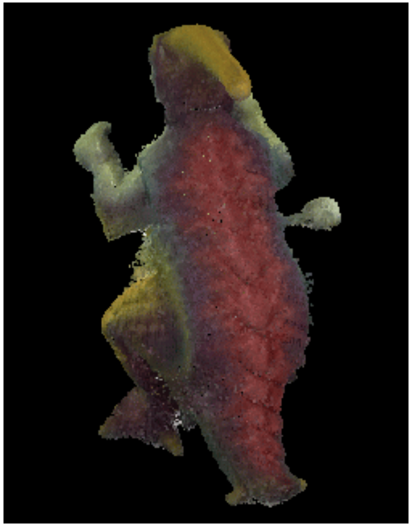
$$\bar{w}' = mean(w')$$

If C $>$ λ= threshold ➔ voxel consistent

Experimental Results

Dinosaur



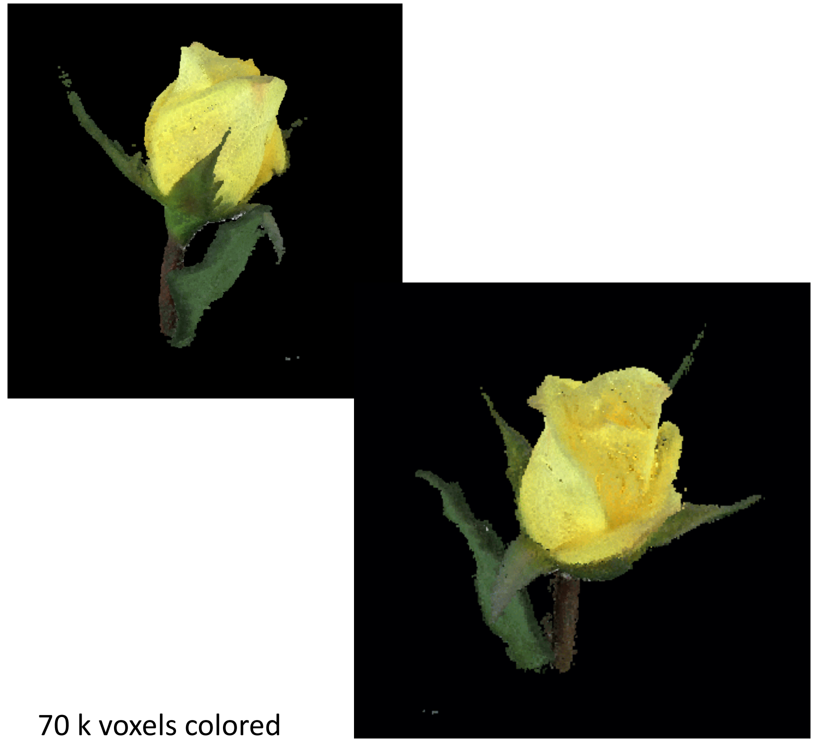


72 k voxels colored
7.6 M voxels tested
7 min to compute on a 250MHz

Image source: http://www.cs.cmu.edu/~seitz/vcolor.html

Experimental Results

Flower

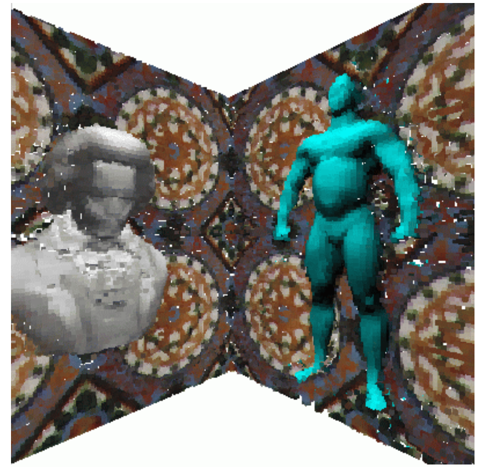


70 k voxels colored
7.6 M voxels tested
7 min to compute on a 250MHz

Image source: http://www.cs.cmu.edu/~seitz/vcolor.html

Experimental Results

Room + weird people



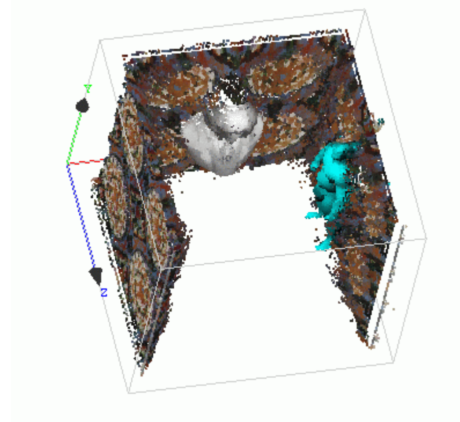


Image source: http://www.cs.cmu.edu/~seitz/vcolor.html

Voxel Coloring: Conclusions

- Good things
  - Model intrinsic scene colors and texture
  - No assumptions on scene topology

- Limitations:
  - Constrained camera positions
  - Lambertian assumption

Further Contributions

- A Theory of Space carving [Kutulakos & Seitz ’99]
  - Voxel coloring in more general framework
  - No restrictions on camera position

- Probabilistic Space carving

  [Broadhurst & Cipolla, ICCV 2001]
  [Bhotika, Kutulakos et. al, ECCV 2002]

Next lecture…

Fitting and Matching