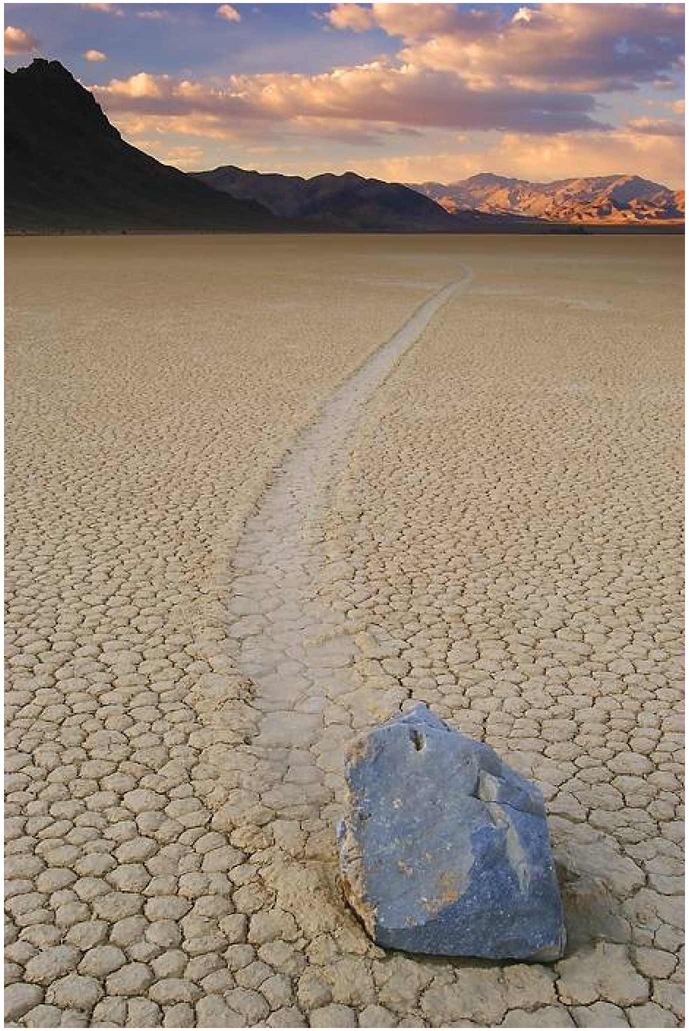
Multiple View Geometry

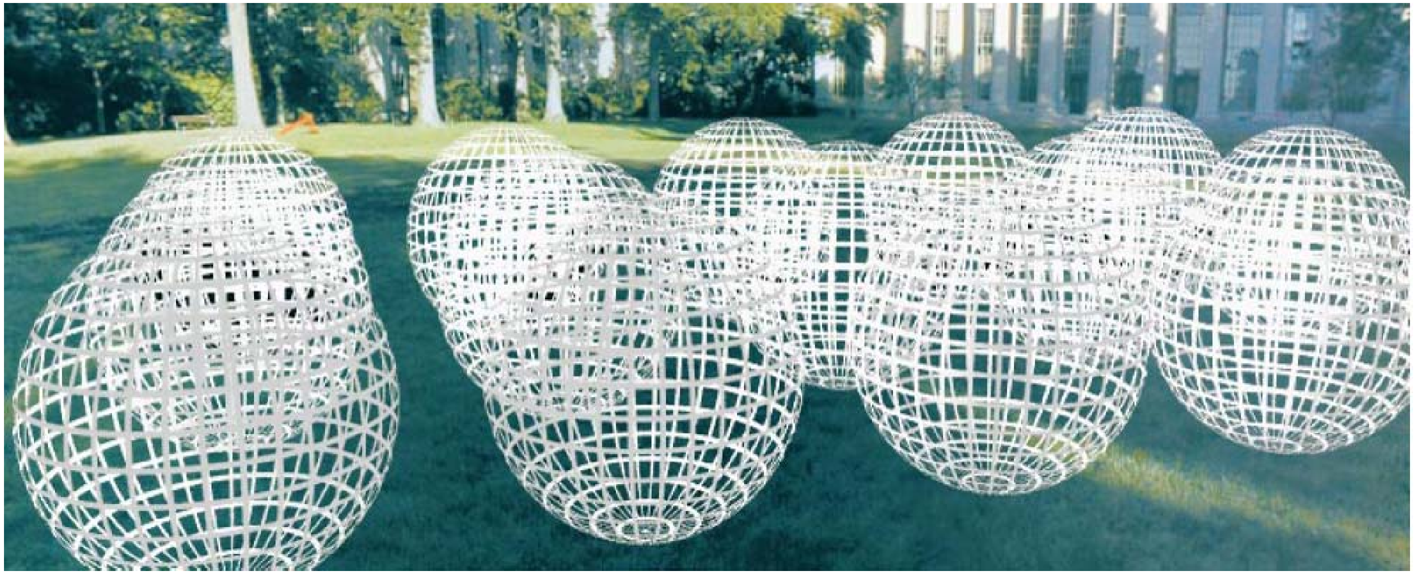


…with a lot of slides stolen from Steve Seitz and Jianbo Shi

15-463: Computational Photography
Alexei Efros, CMU, Fall 2005

Our Goal

The Plenoptic Function



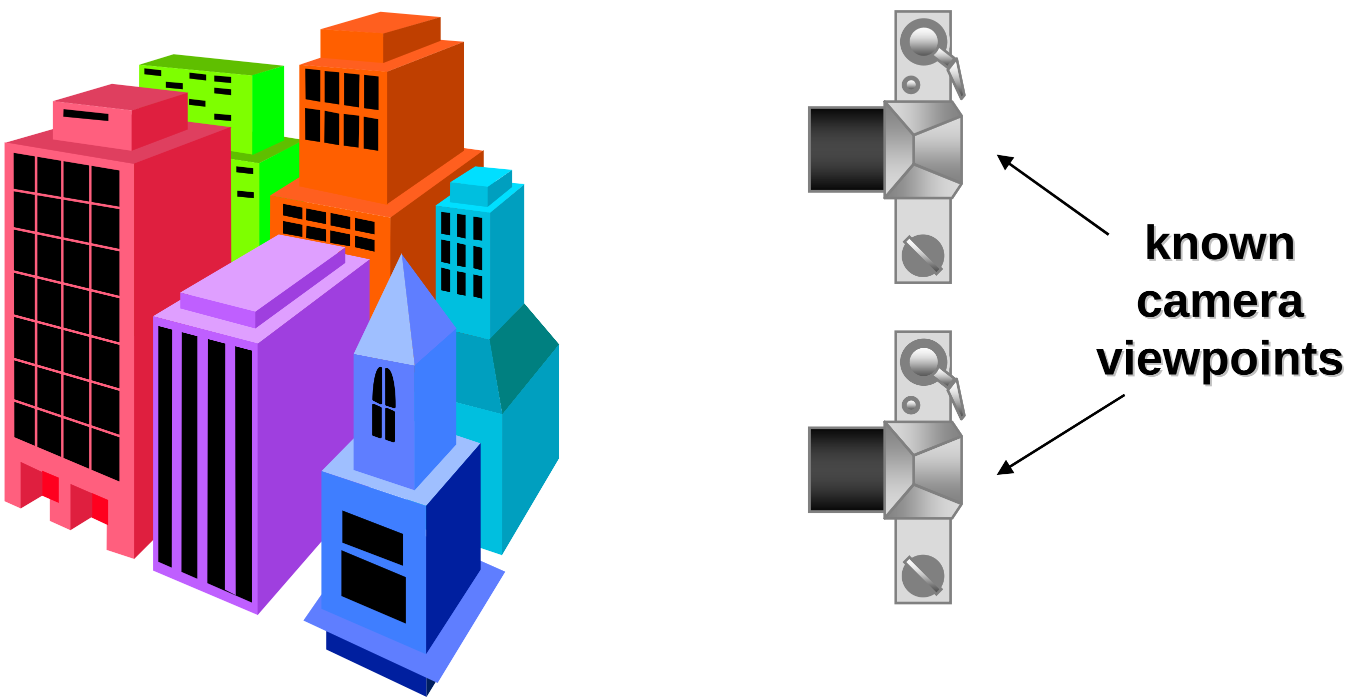
$$P(\theta,\phi,\lambda,t,V_X,V_Y,V_Z)$$

How can we compress this into something manageable?

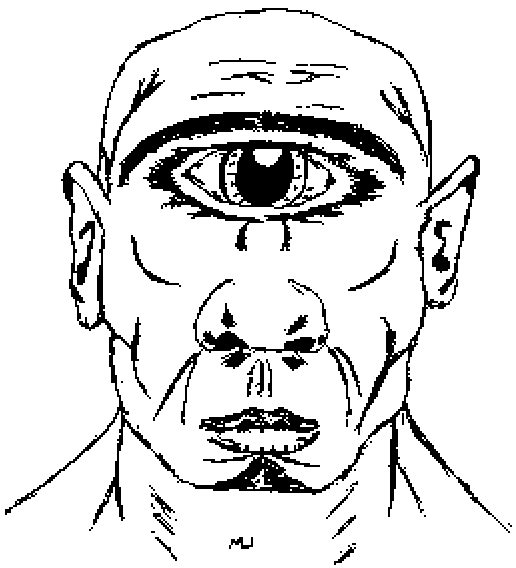
Stereo Reconstruction

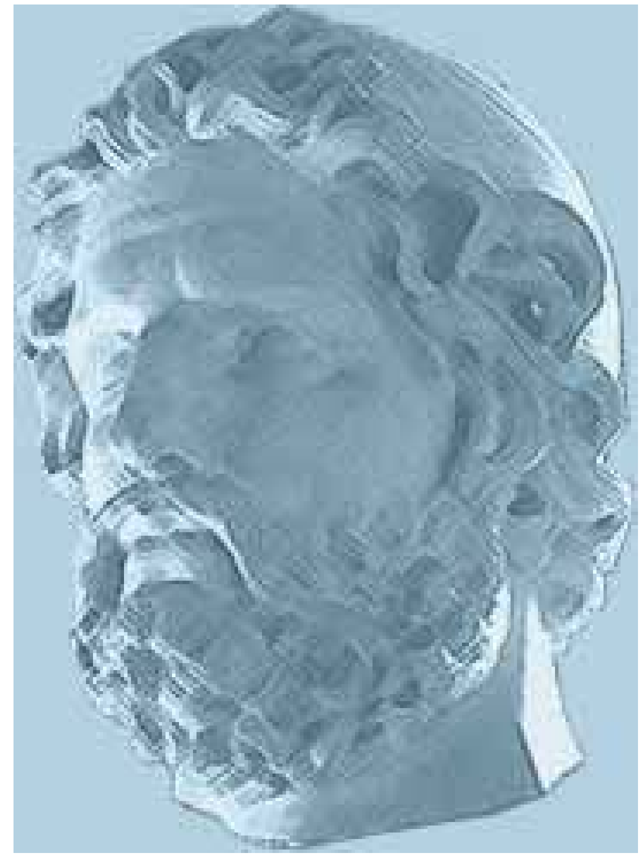
The Stereo Problem

- Shape from two (or more) images
- Biological motivation



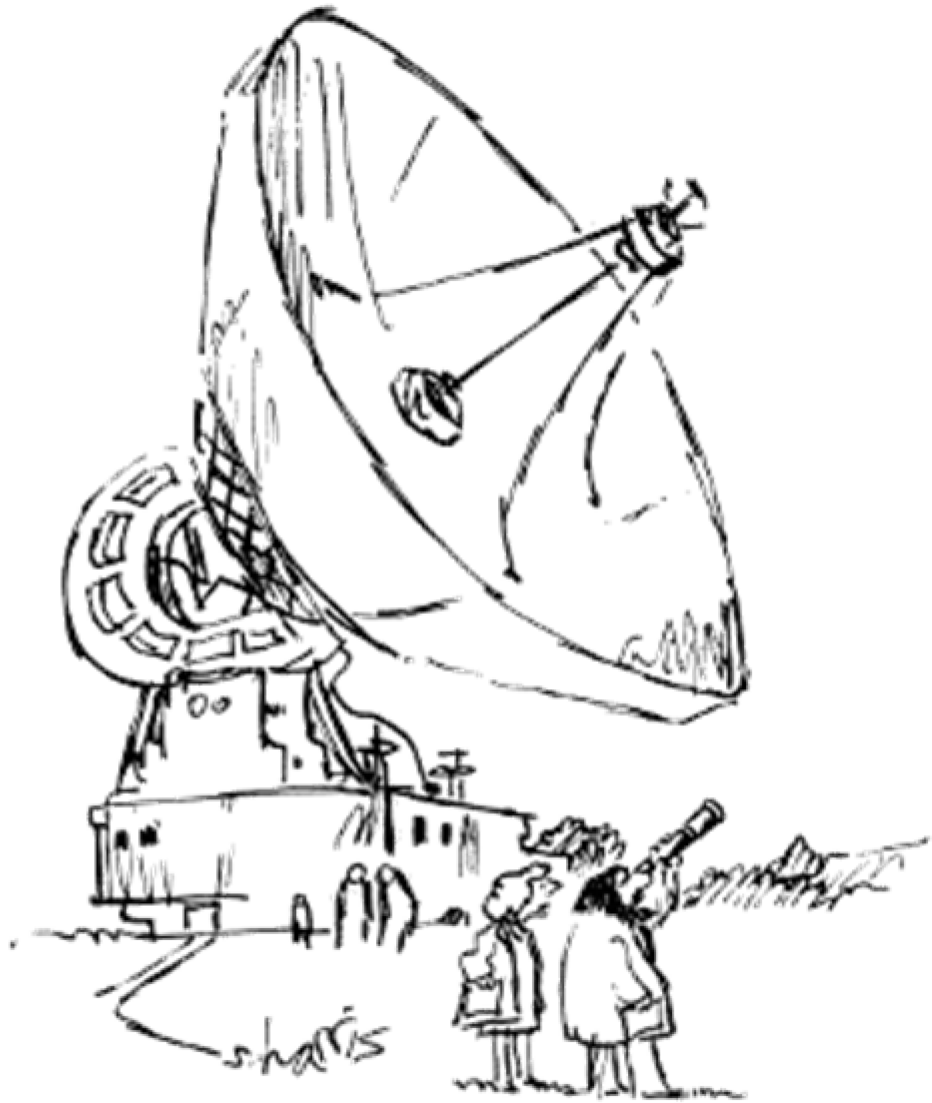
Why do we have two eyes?





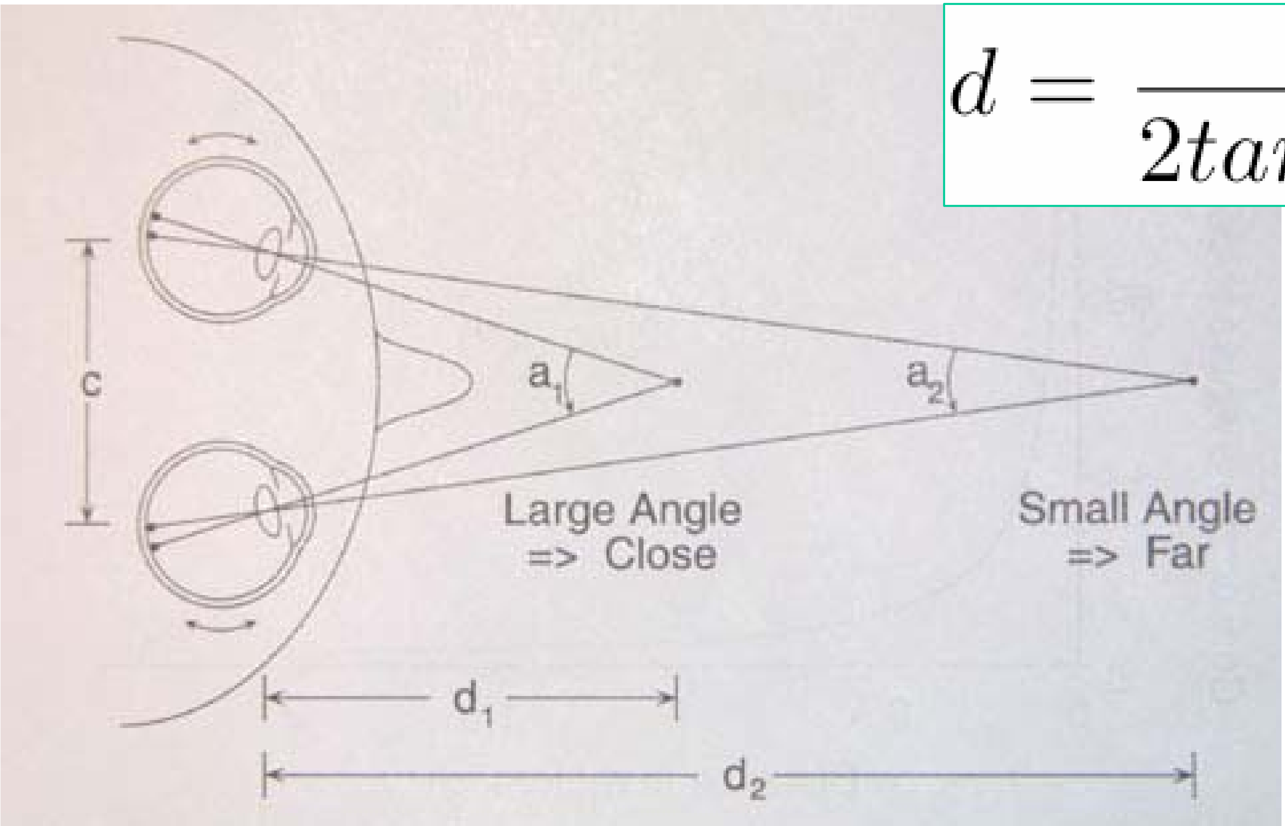
Cyclope **vs.** **Odysseus**

1. Two is better than one



"Just checking."

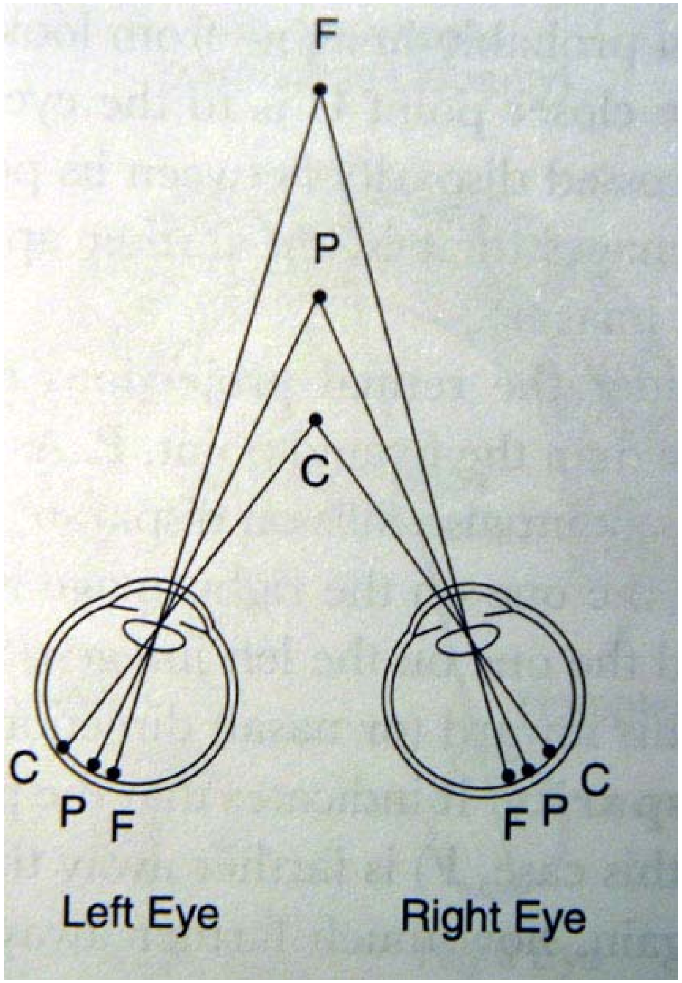
2. Depth from Convergence



$$d = \frac{c}{2tan(a/2)}$$

Human performance: up to 6-8 feet

3. Depth from binocular disparity

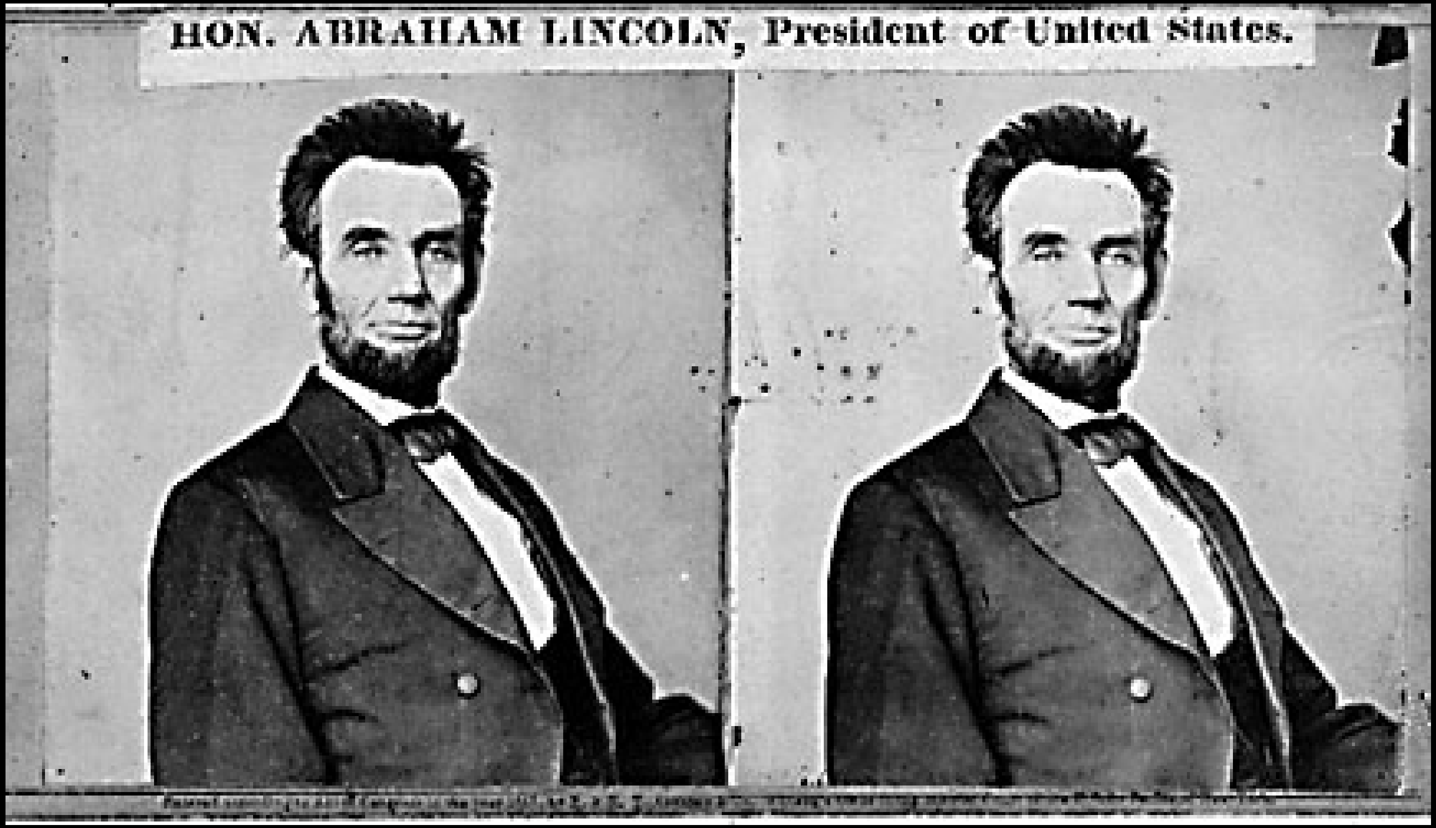


Sign and magnitude of disparity

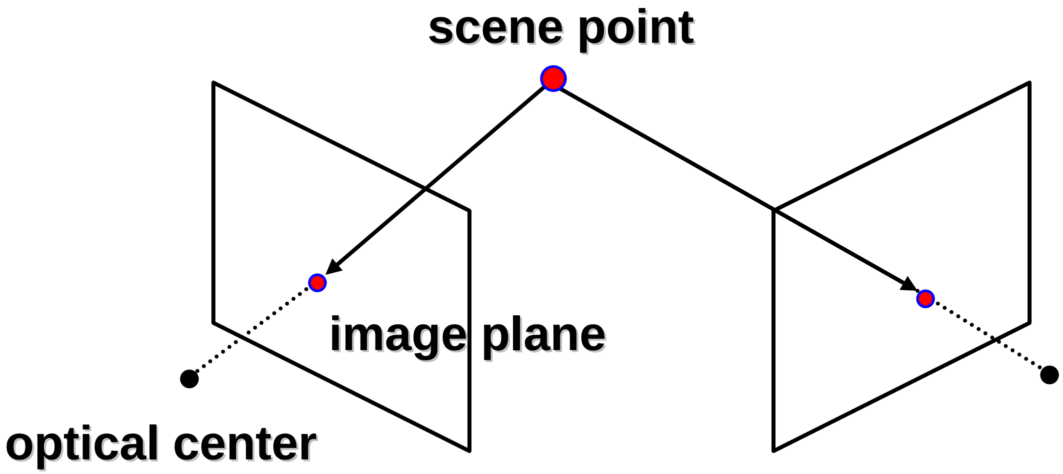
P: converging point

C: object nearer projects to the outside of the P, disparity = +

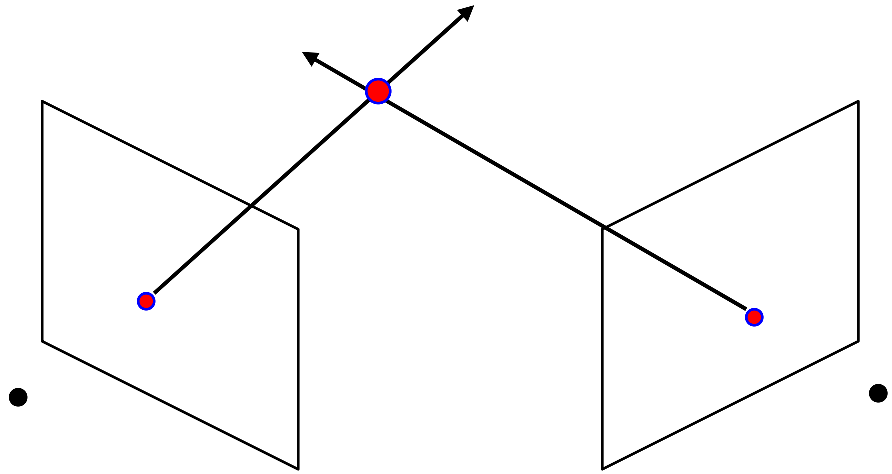
F: object farther projects to the inside of the P, disparity = -


HON. ABRAHAM LINCOLN, President of United States.

Stereo



Stereo



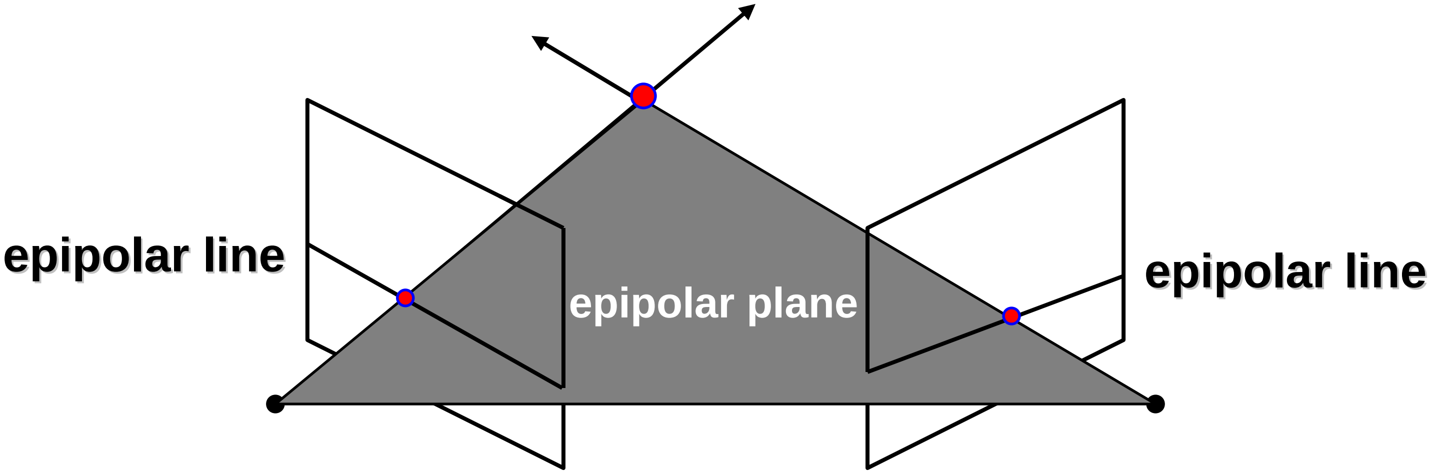
Basic Principle: Triangulation

- Gives reconstruction as intersection of two rays
- Requires
  - calibration
  - ***point correspondence***

Stereo correspondence

Determine Pixel Correspondence

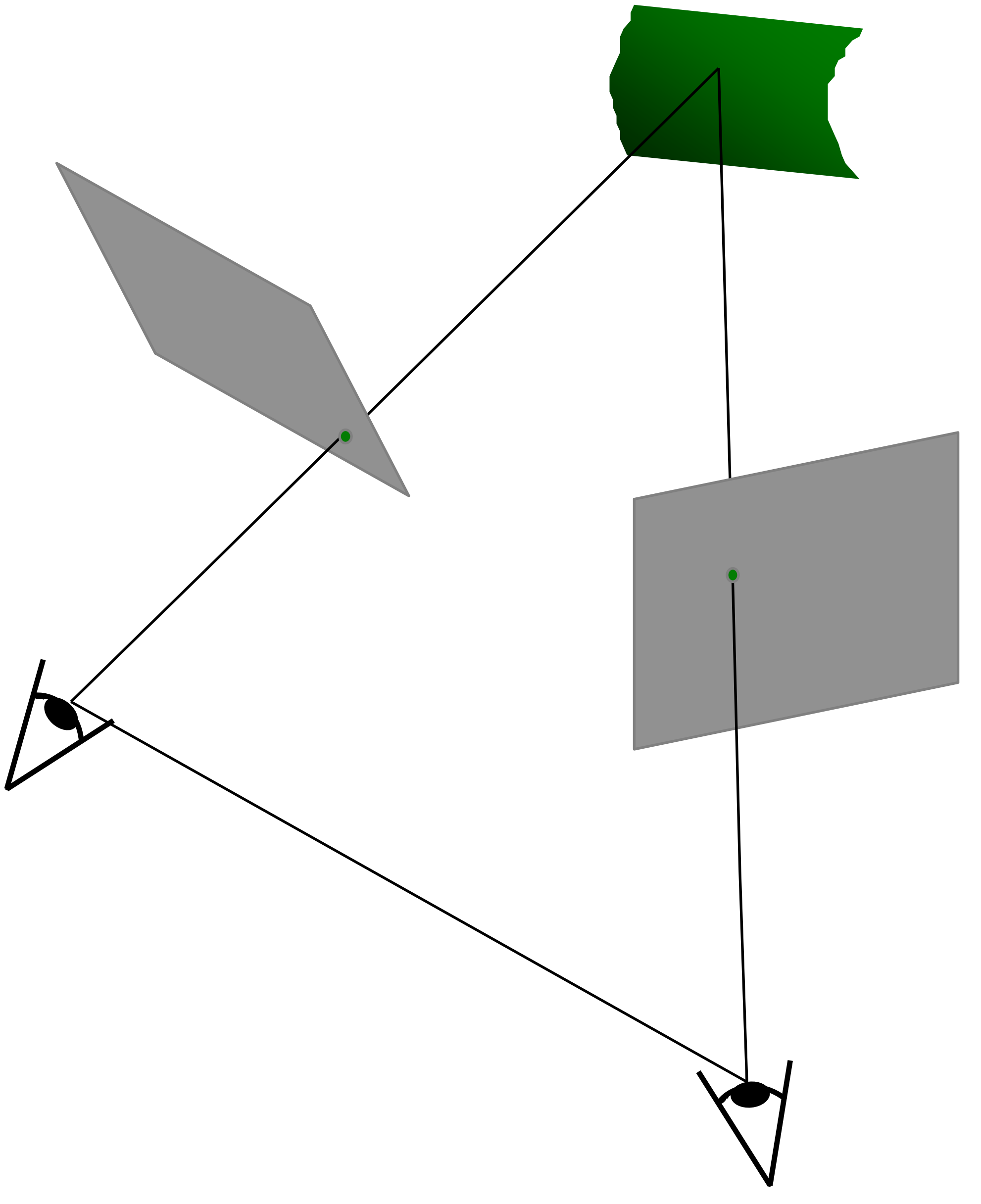
- Pairs of points that correspond to same scene point



Epipolar Constraint

- Reduces correspondence problem to 1D search along *conjugate epipolar lines*

Stereo image rectification



Stereo image rectification

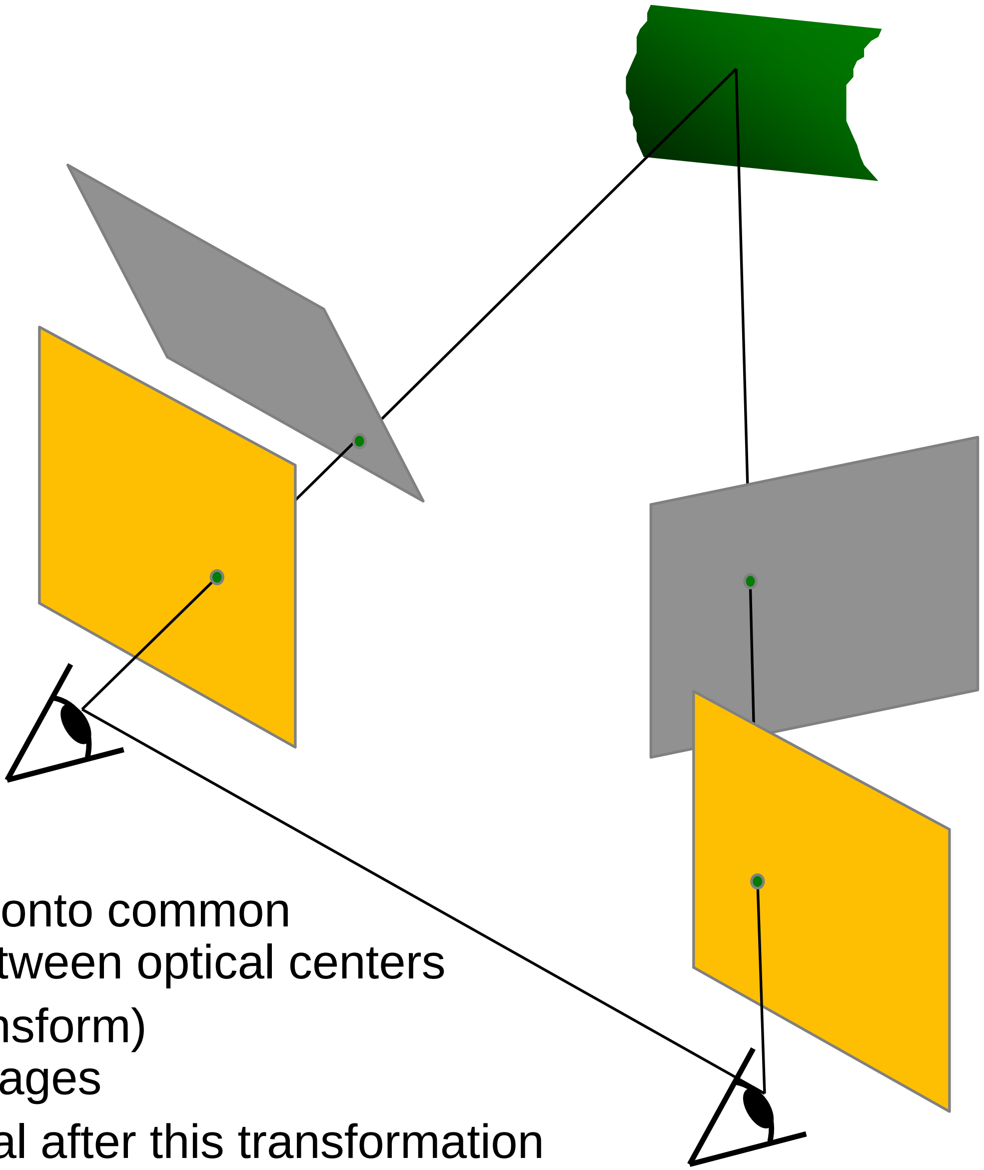
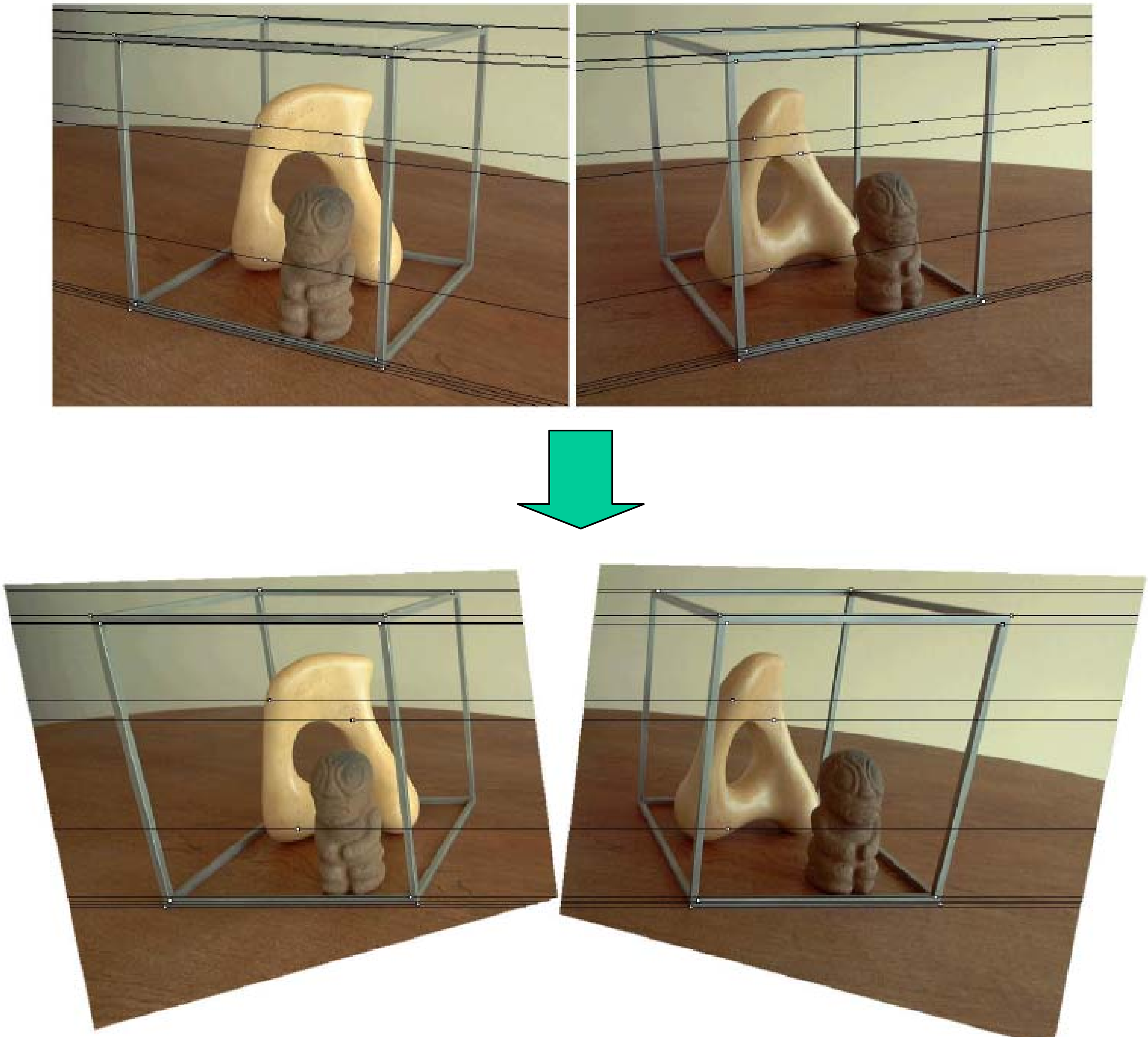


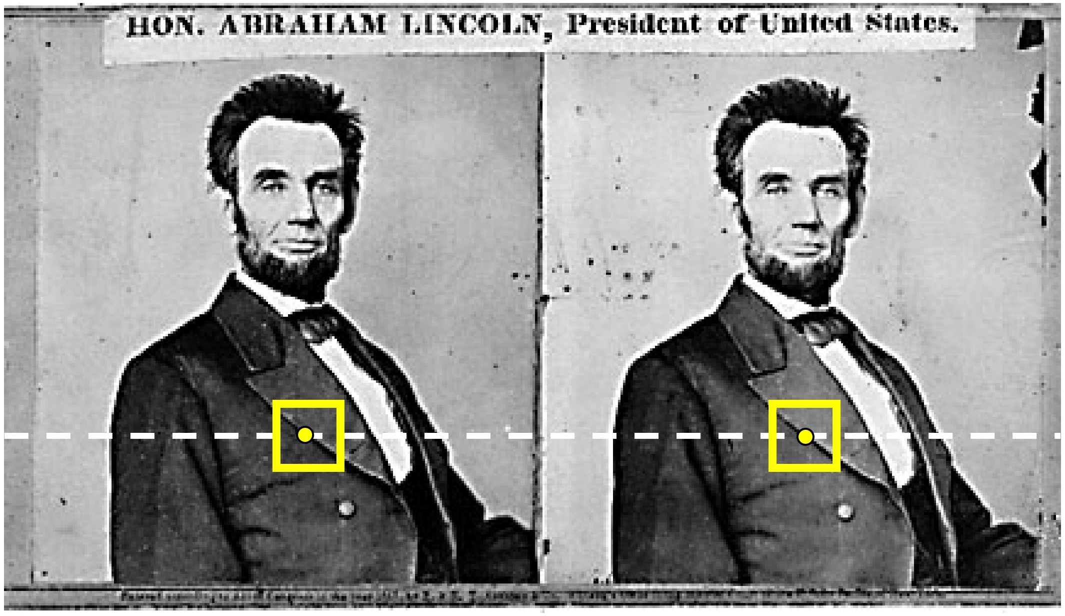
Image Reprojection

- reproject image planes onto common plane parallel to line between optical centers
- a homography (3x3 transform) applied to both input images
- pixel motion is horizontal after this transformation
- C. Loop and Z. Zhang. [Computing Rectifying Homographies for Stereo Vision]. IEEE Conf. Computer Vision and Pattern Recognition, 1999.

Stereo Rectification



Your basic stereo algorithm



For each epipolar line

For each pixel in the left image

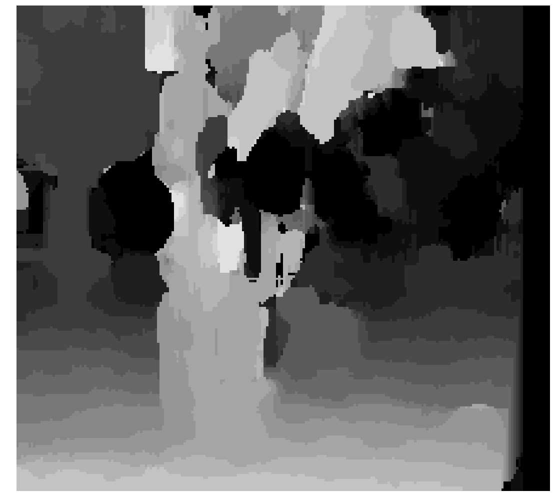
- compare with every pixel on same epipolar line in right image
- pick pixel with minimum match cost

Improvement: match ***windows***

- This should look familar...
- Can use Lukas-Kanade or discrete search (latter more common)

Window size

W = 3



W = 20

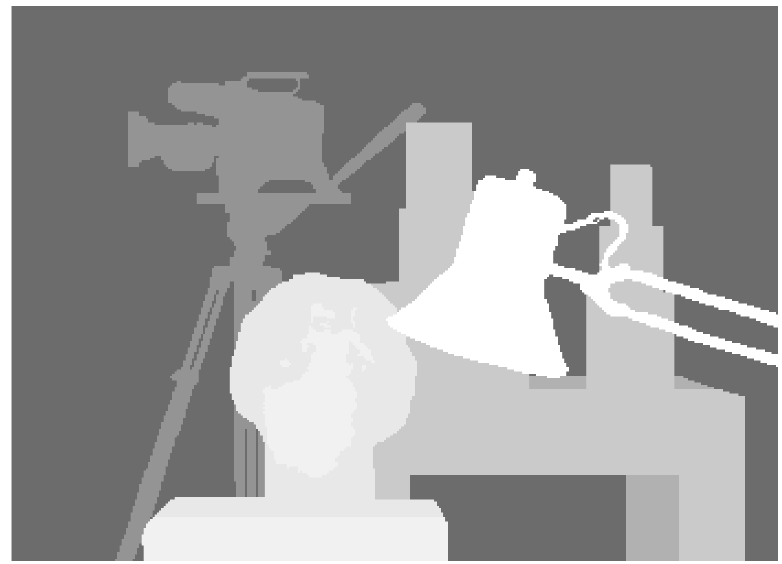
Effect of window size

- Smaller window
  - +
  - –
- Larger window
  - +
  - –

Stereo results

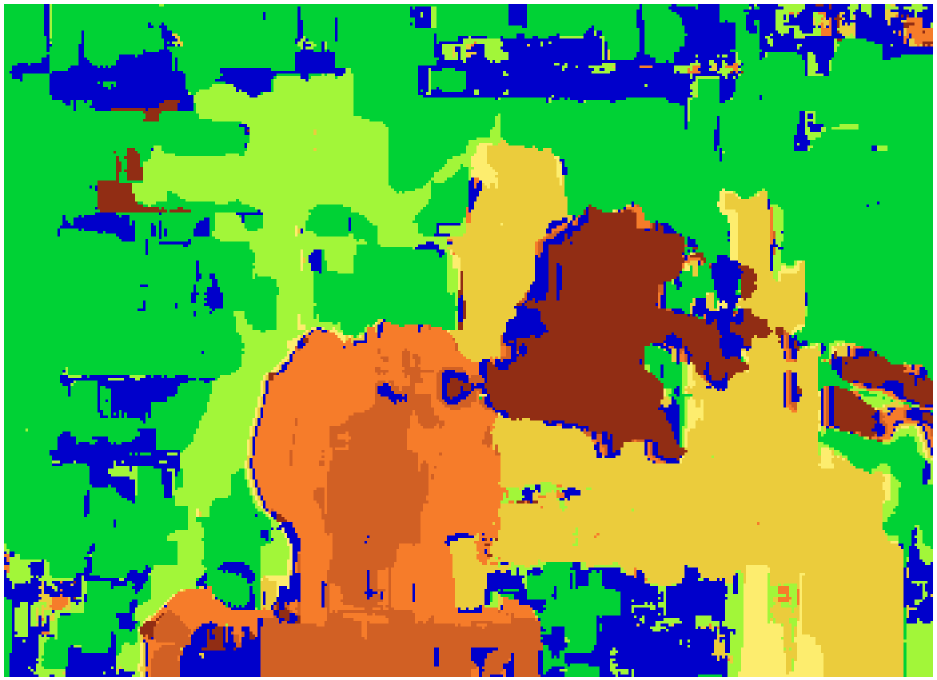
- Data from University of Tsukuba
- Similar results on other images without ground truth

Scene



Ground truth

Results with window search



Window-based matching
(best window size)

Ground truth

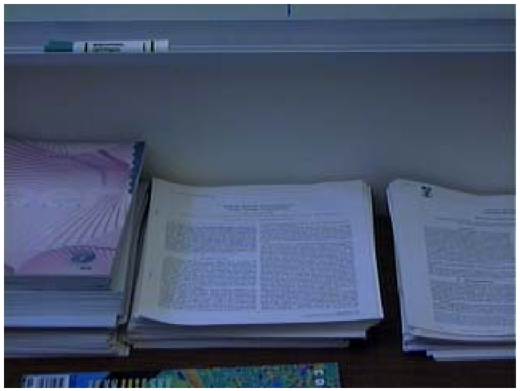
Better methods exist...

State of the art method

Ground truth

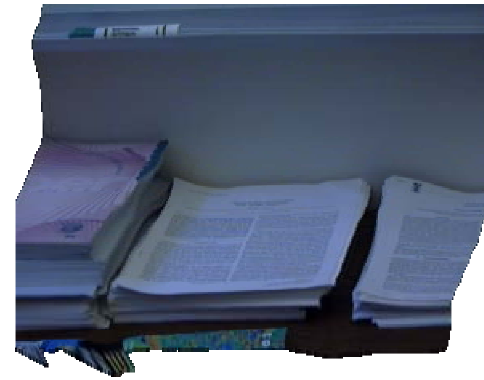
Boykov et al., [Fast Approximate Energy Minimization via Graph Cuts](), International Conference on Computer Vision, September 1999.

Depth from disparity

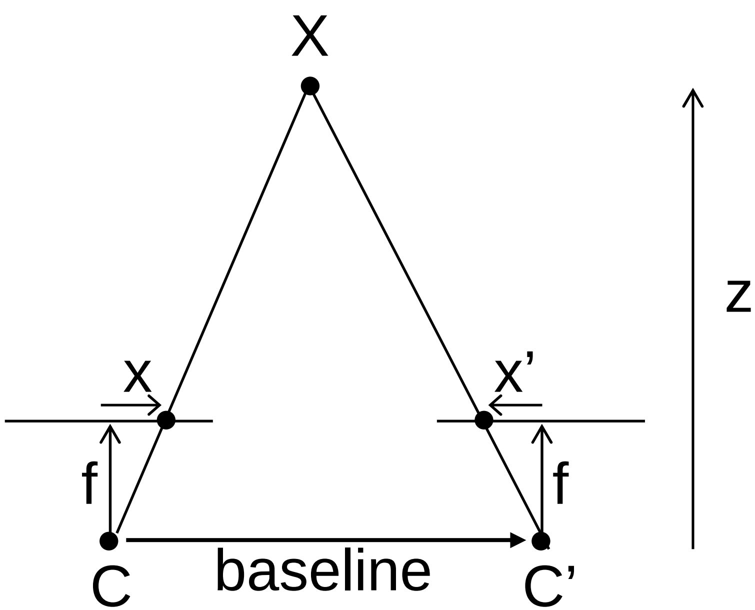


input image (1 of 2)

depth map
[Szeliski & Kang ‘95]



3D rendering



$$disparity = x - x' = \frac{baseline * f}{z}$$

Stereo reconstruction pipeline

Steps

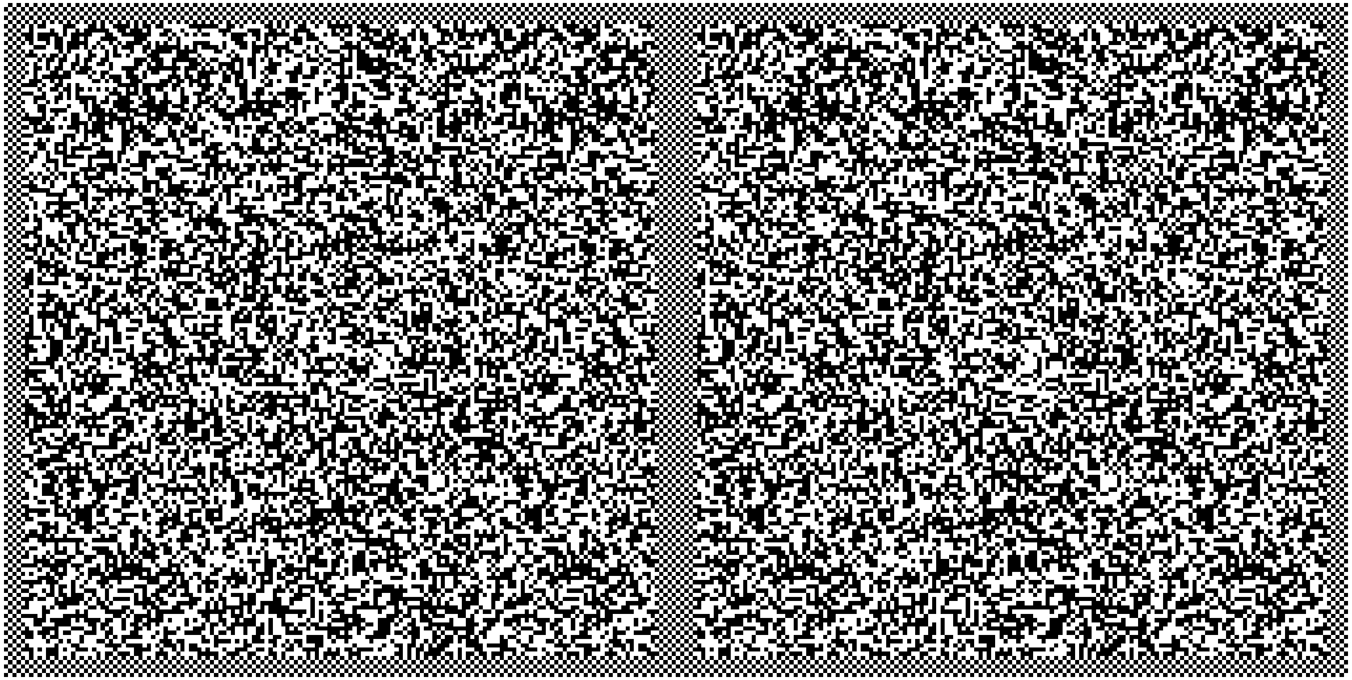
- Calibrate cameras
- Rectify images
- Compute disparity
- Estimate depth

What will cause errors?

- Camera calibration errors
- Poor image resolution
- Occlusions
- Violations of brightness constancy (specular reflections)
- Large motions
- Low-contrast image regions

Stereo matching

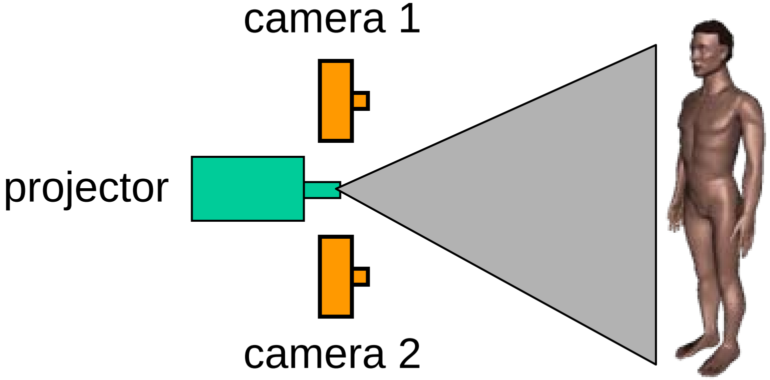
Need texture for matching

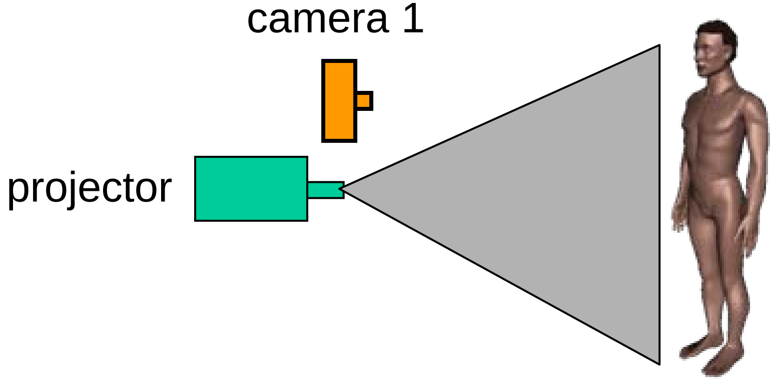


Julesz-style Random Dot Stereogram

Active stereo with structured light

Li Zhang's one-shot stereo

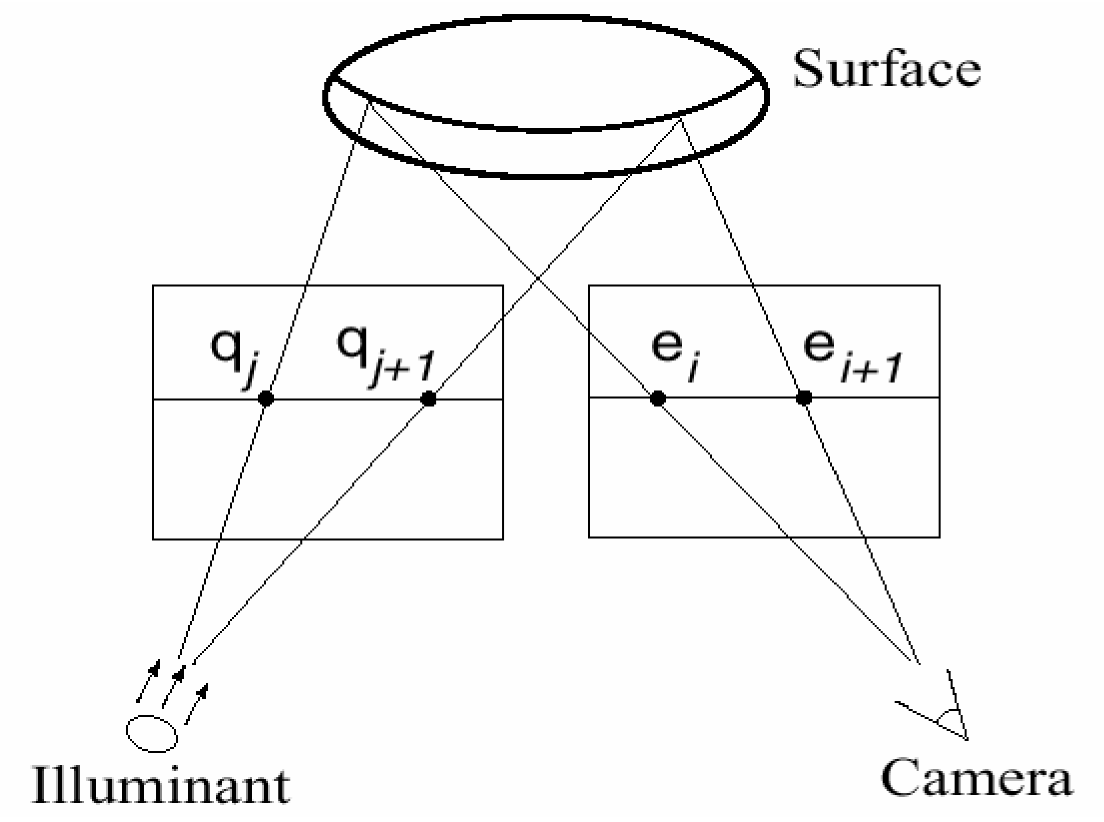




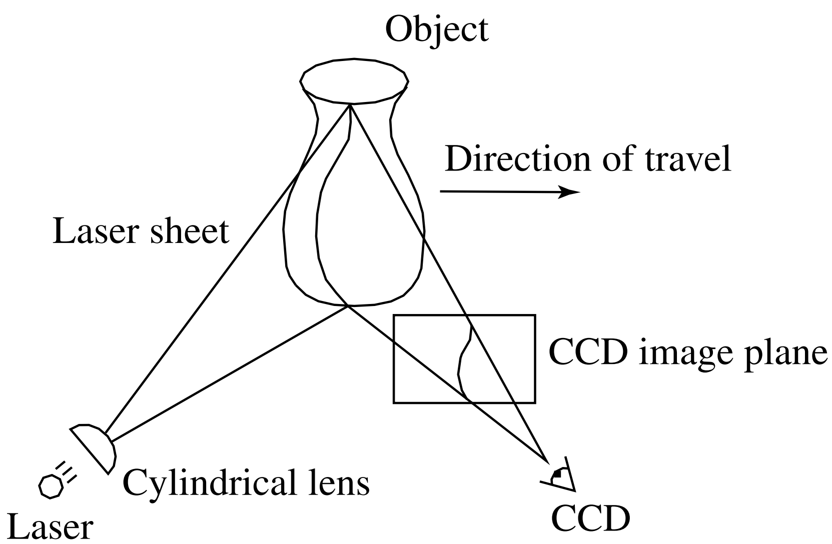
Project "structured" light patterns onto the object

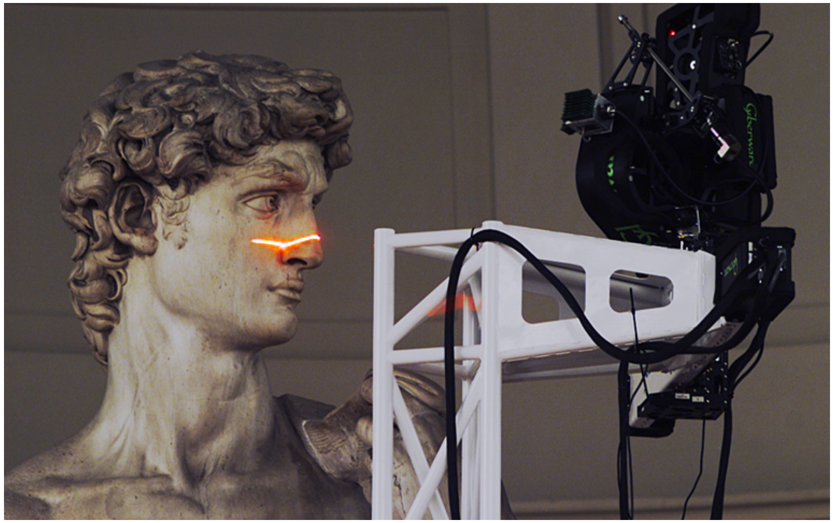
- simplifies the correspondence problem

Active stereo with structured light



Laser scanning



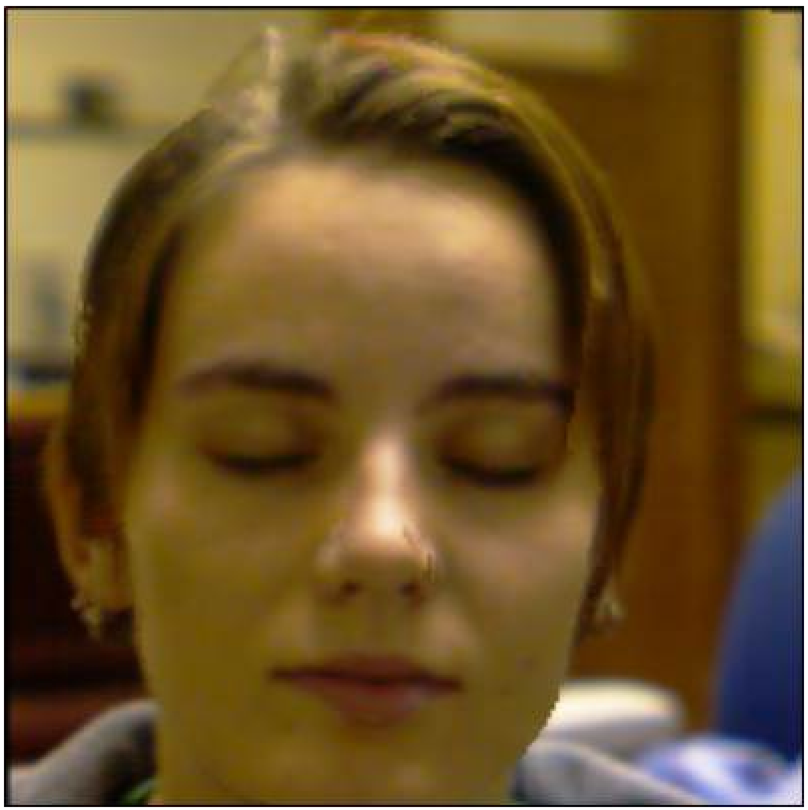


Digital Michelangelo Project
http://graphics.stanford.edu/projects/mich/

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

Portable 3D laser scanner (this one by Minolta)



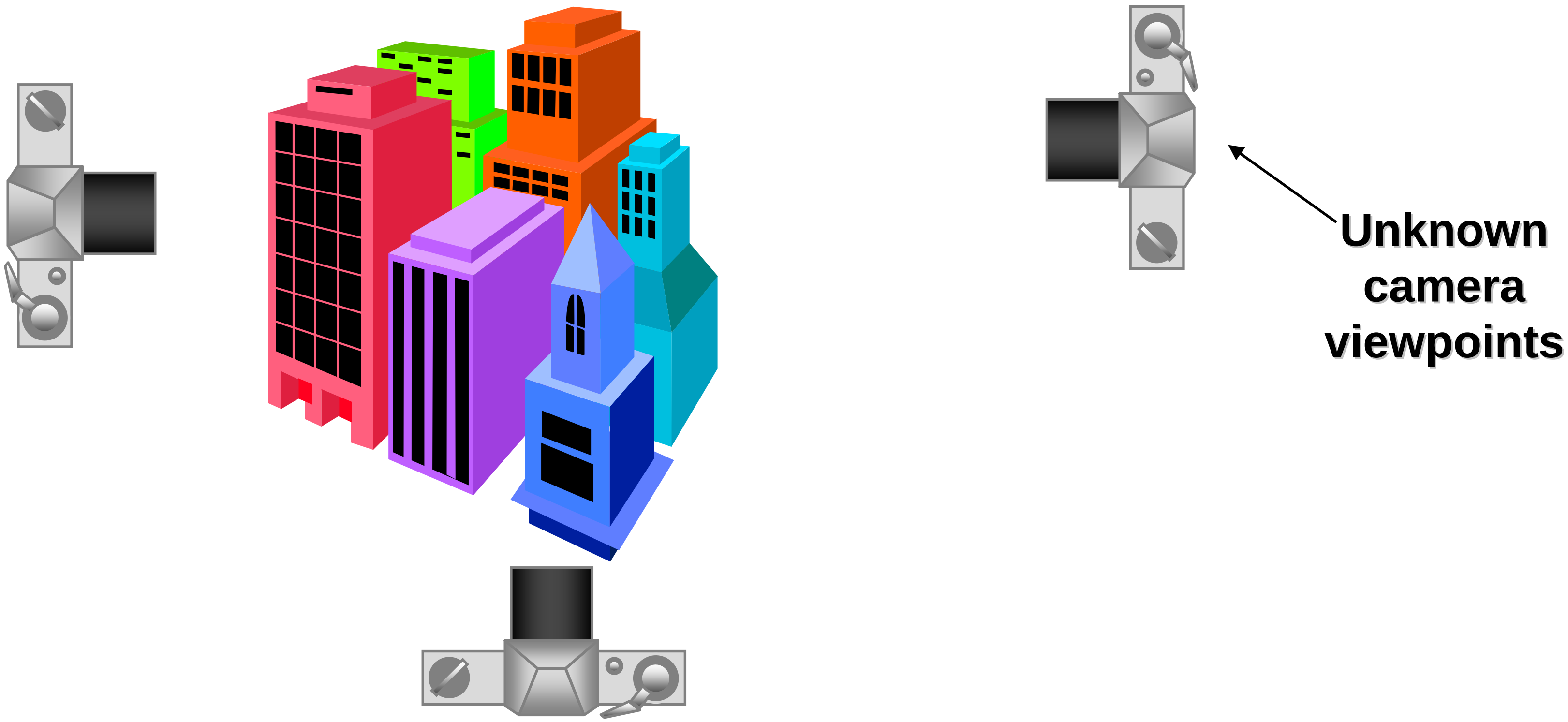
Real-time stereo

Nomad robot searches for meteorites in Antartica
http://www.frc.ri.cmu.edu/projects/meteorobot/index.html

Used for robot navigation (and other tasks)

- Several software-based real-time stereo techniques have been developed (most based on simple discrete search)

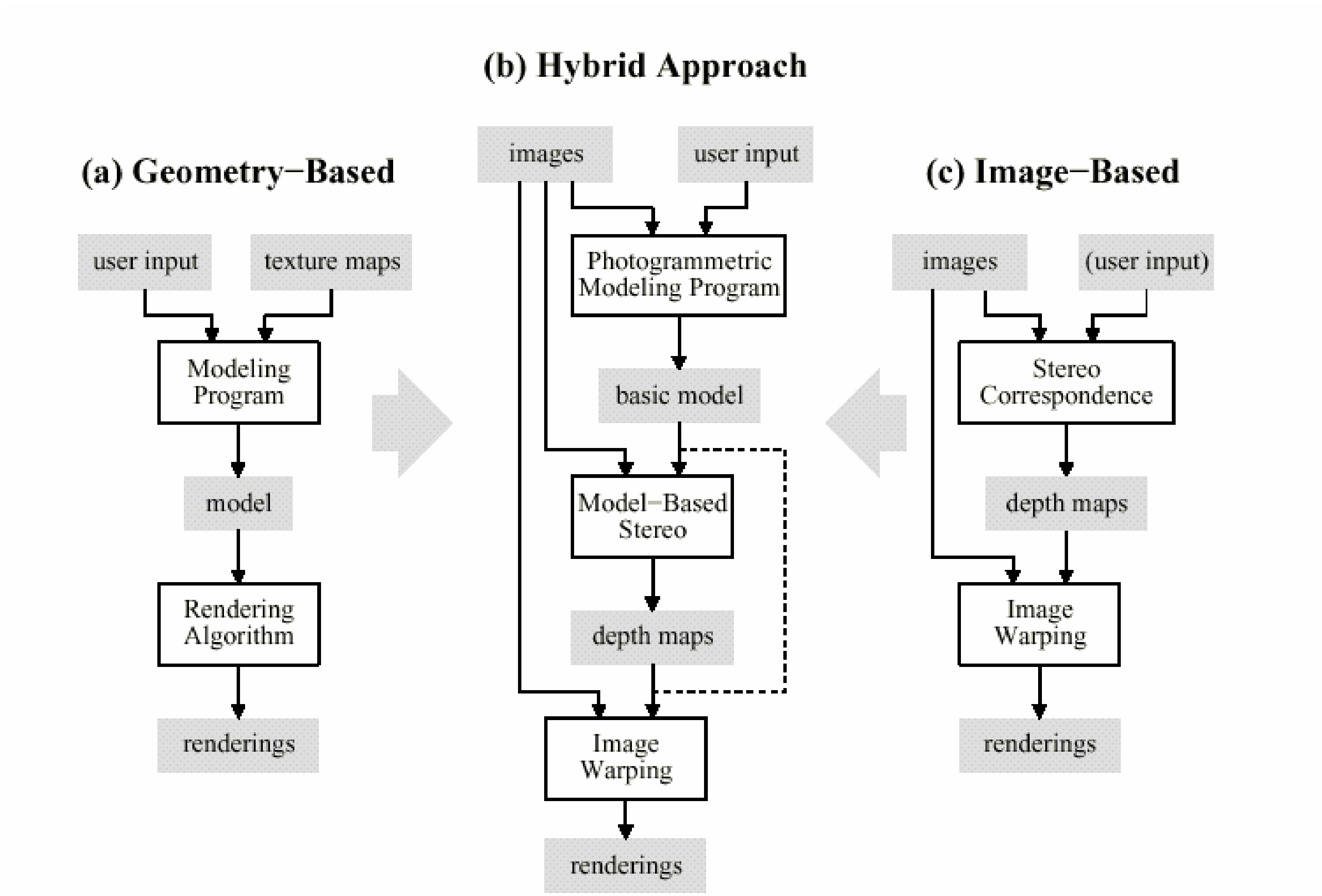
Structure from Motion



Reconstruct

- Scene geometry
- Camera motion

Three approaches

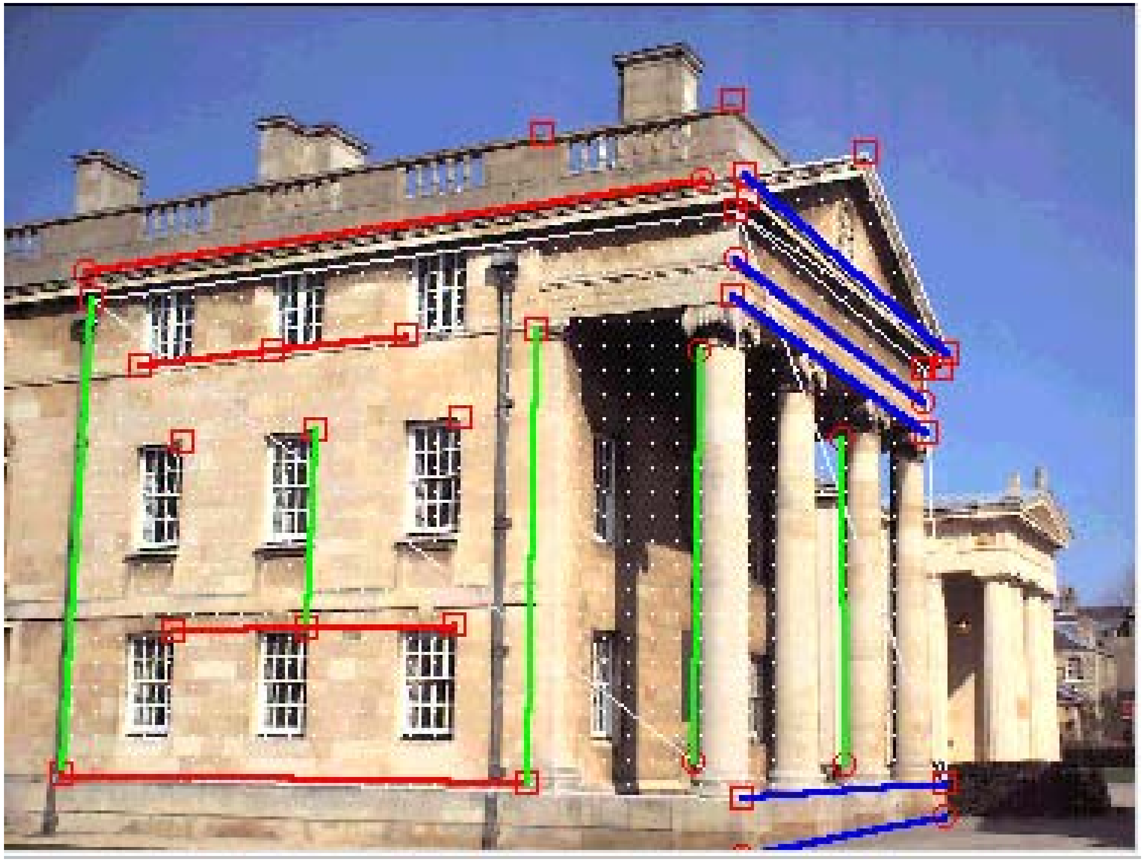


Outline of a simple algorithm (1)

- Based on constraints
- Input to the algorithm (1): two images

Outline of a simple algorithm (2)

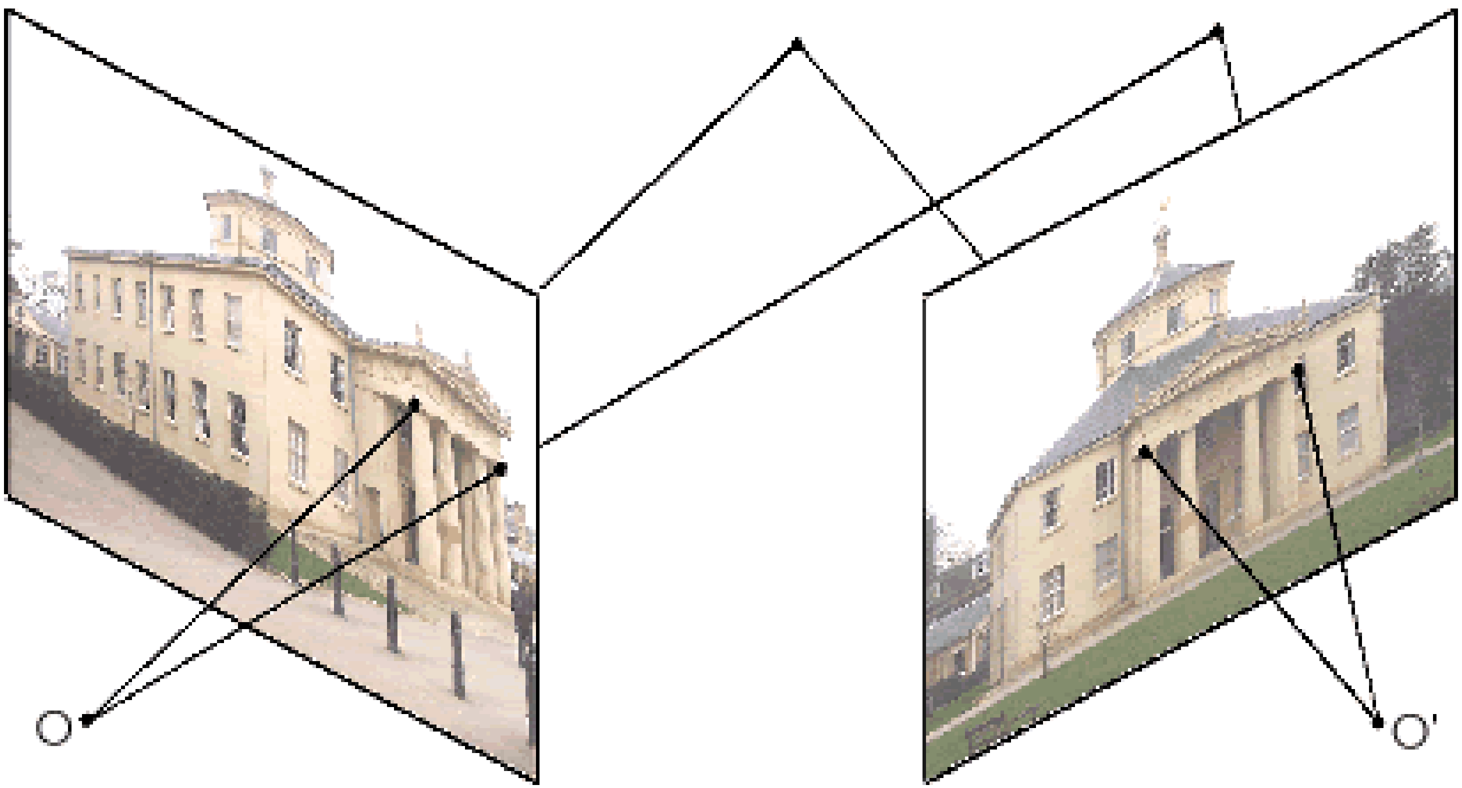
- Input to the algorithm (2):
  User select edges and corners



Outline of a simple algorithm (3)

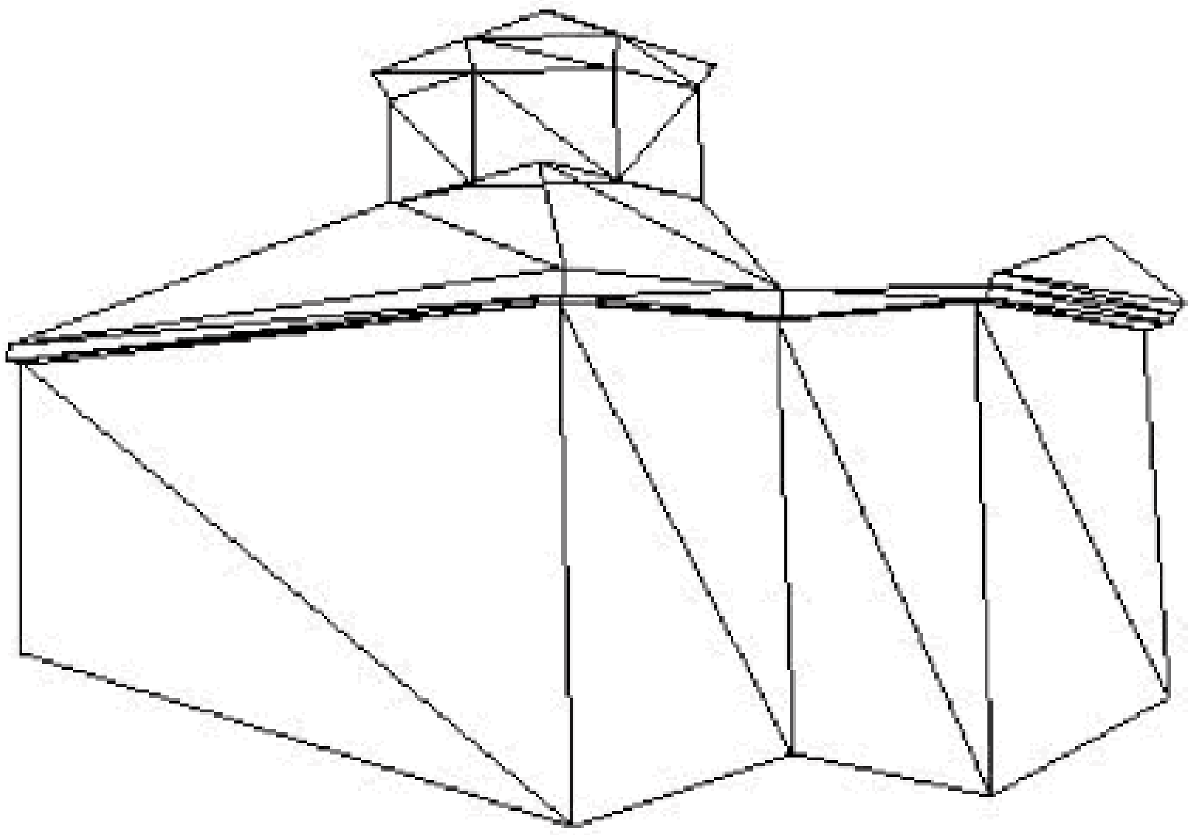
- Camera Position and Orientation

  Determine the position and orientation of camera



Outline of a simple algorithm (4)

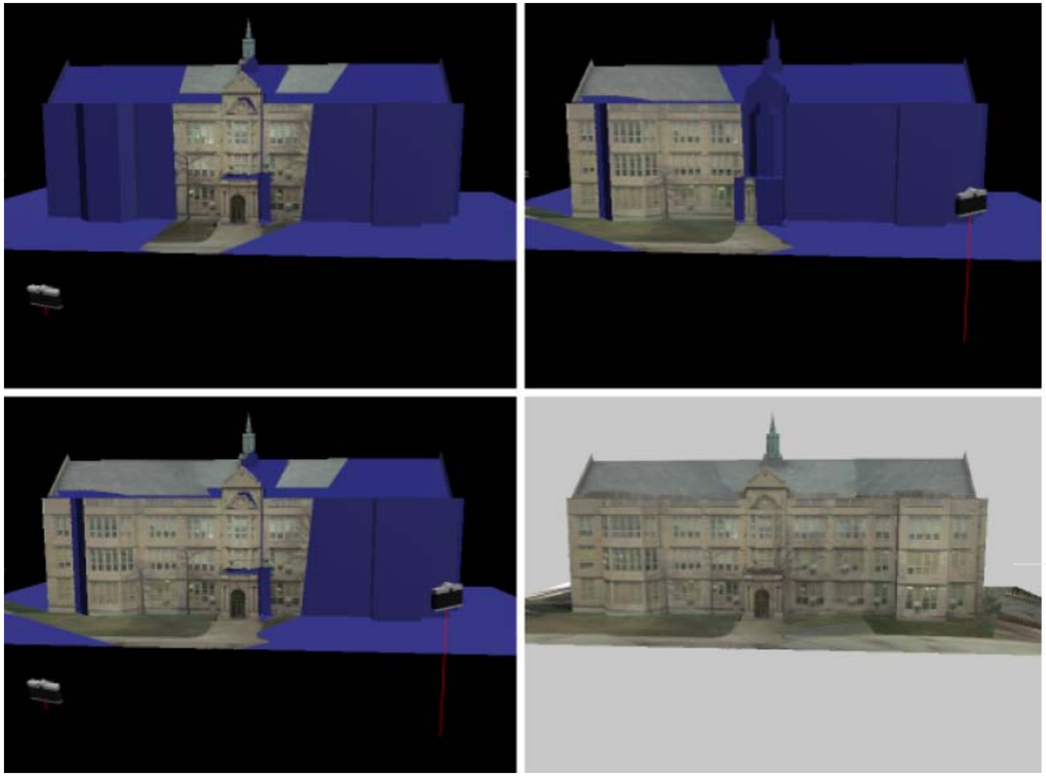
- Computing projection matrix and Reconstruction



Outline of a simple algorithm (5)

- Compute 3D textured triangles

View-Dependant Texture Mapping



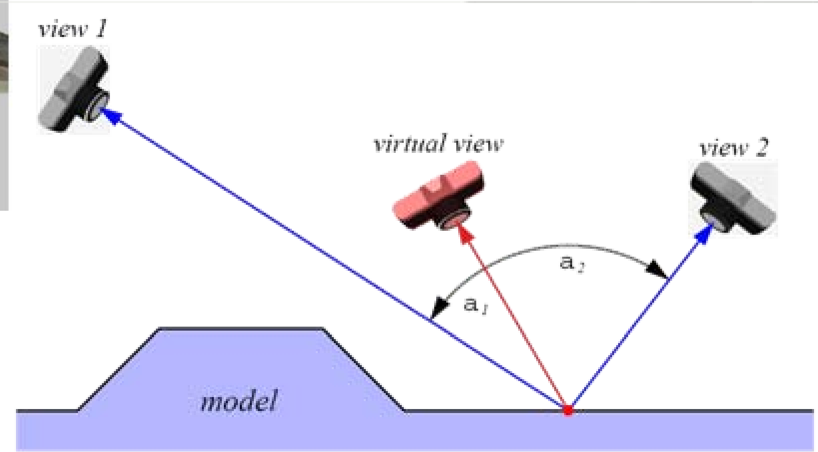
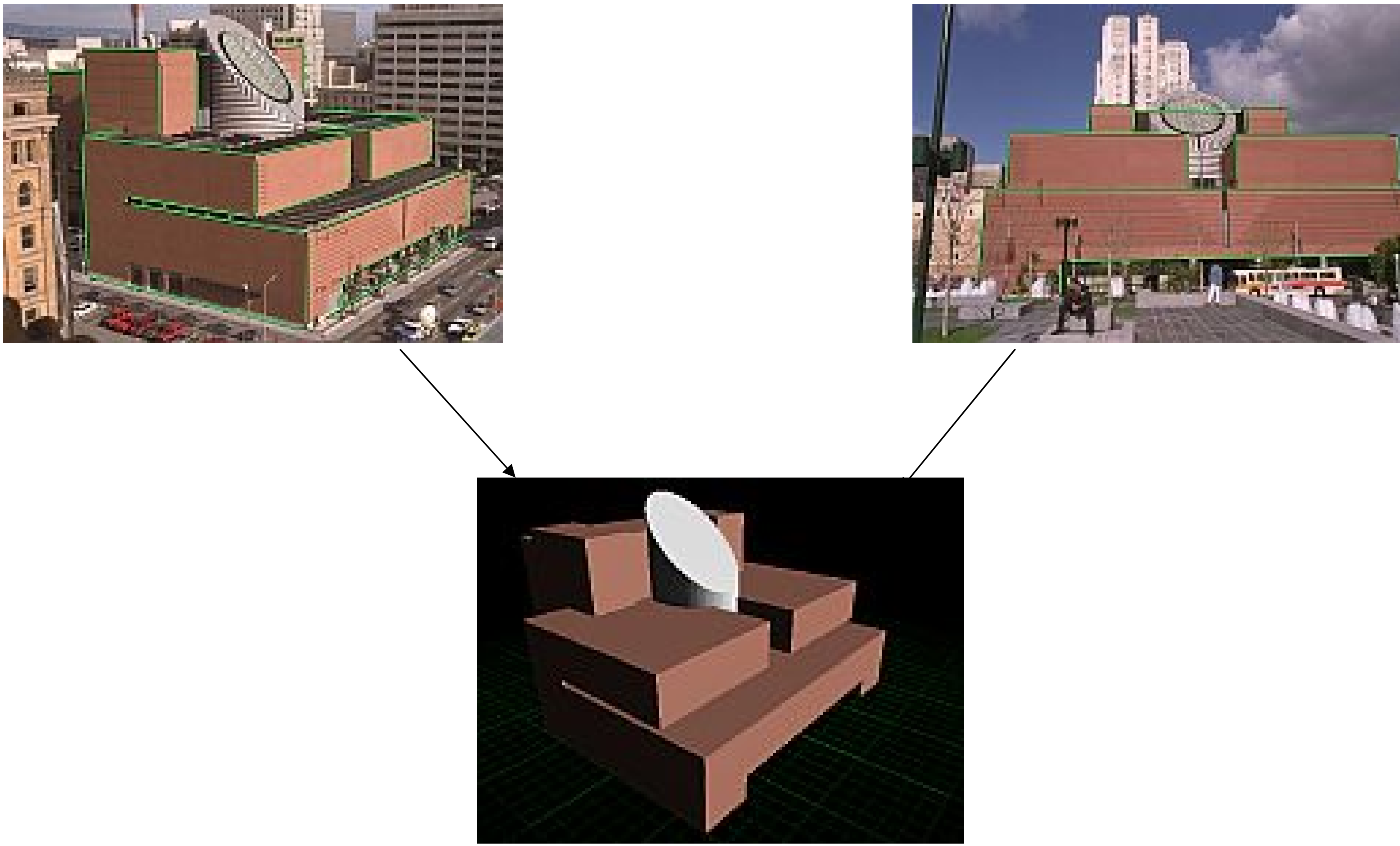


Figure 12: *The weighting function used in view-dependent texture mapping. The pixel in the virtual view corresponding to the point on the model is assigned a weighted average of the corresponding pixels in actual views 1 and 2. The weights w_1 and w_2 are inversely inversely proportional to the magnitude of angles a_1 and a_2. Alternately, more sophisticated weighting functions based on expected foreshortening and image resampling can be used.*

Facade



SFMOMA (San Francisco Museum of Modern Art) by Yizhou Yu,

Façade (Debevec et al) inputs

Façade (Debevec et al)