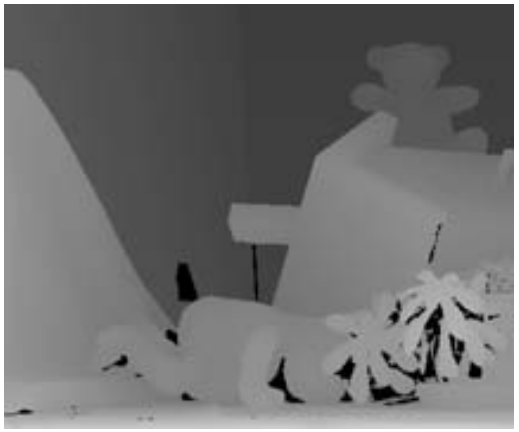
Stereo Vision

Source: S. Lazebnik

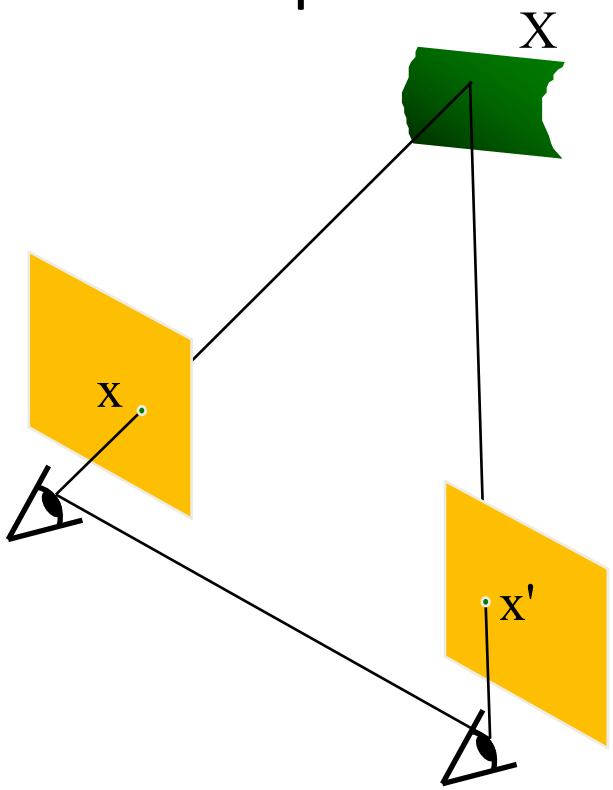
Amount of horizontal movement is …

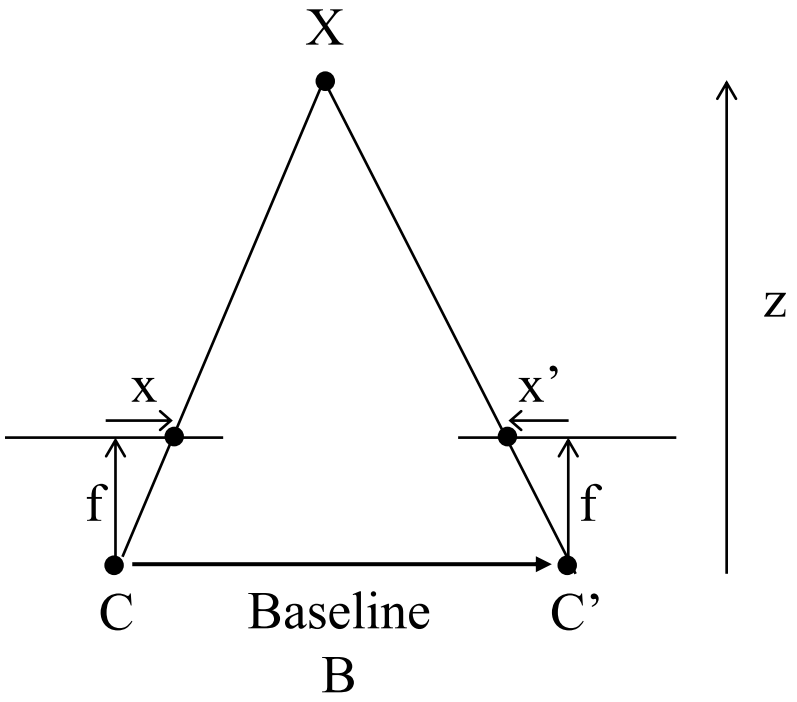
…inversely proportional to the distance from the camera



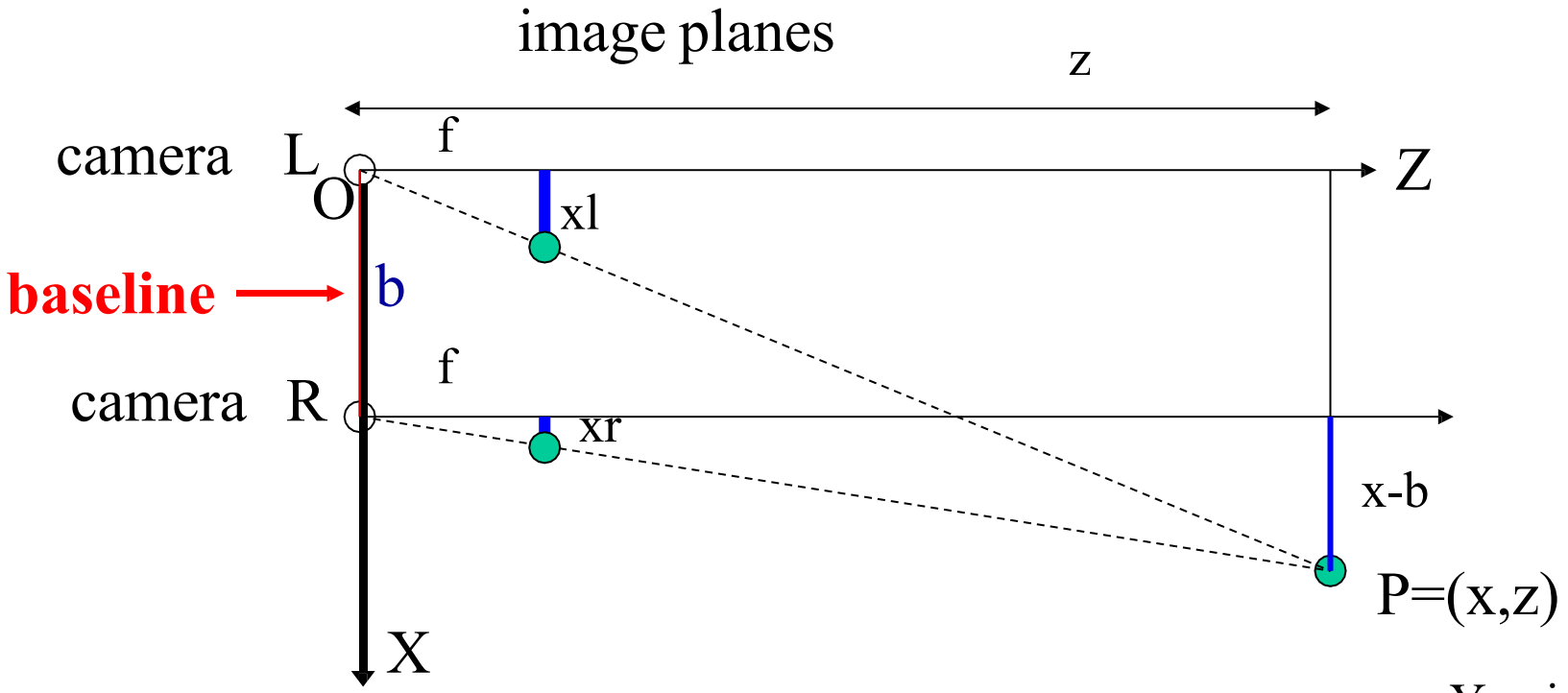
Depth from Stereo

- Goal: recover depth by finding image coordinate x’ that corresponds to x





Optic axes of 2 cameras are parallel



$$\frac{z}{f} = \frac{x}{xl}$$

$$\frac{z}{f} = \frac{x-b}{xr}$$

$$\frac{z}{f} = \frac{y}{yl} = \frac{y}{yr}$$

Y-axis is perpendicular to the page.

(from similar triangles)

3D from Stereo Images

For stereo cameras with parallel optical axes, focal length f, baseline b, corresponding image points (xl,yl) and (xr,yr), the location of the 3D point can be derived from previous slide's equations:

Depth z = f*b / (xl - xr) = f*b/d

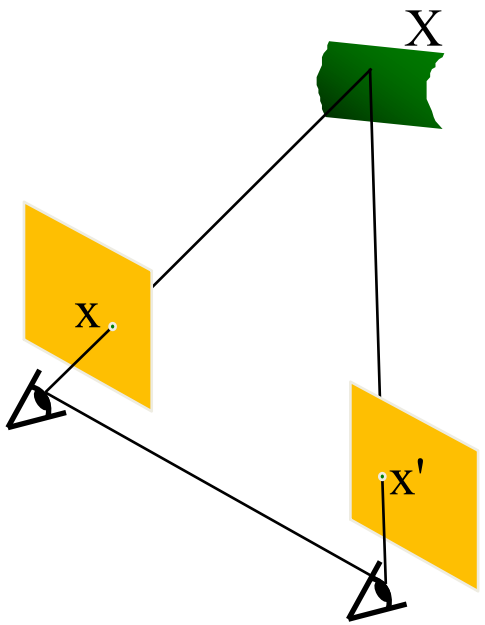
$x = xl*z/f$ or $b + xr*z/f$

$y = yl*z/f$ or $yr*z/f$

Note that **depth is inversely proportional to disparity**

Depth from Stereo

- Goal: recover depth by finding image coordinate x’ that corresponds to x
- Sub-Problems
  1. Calibration: How do we recover the relation of the cameras (if not already known)?
  2. Correspondence: How do we search for the matching point x’?

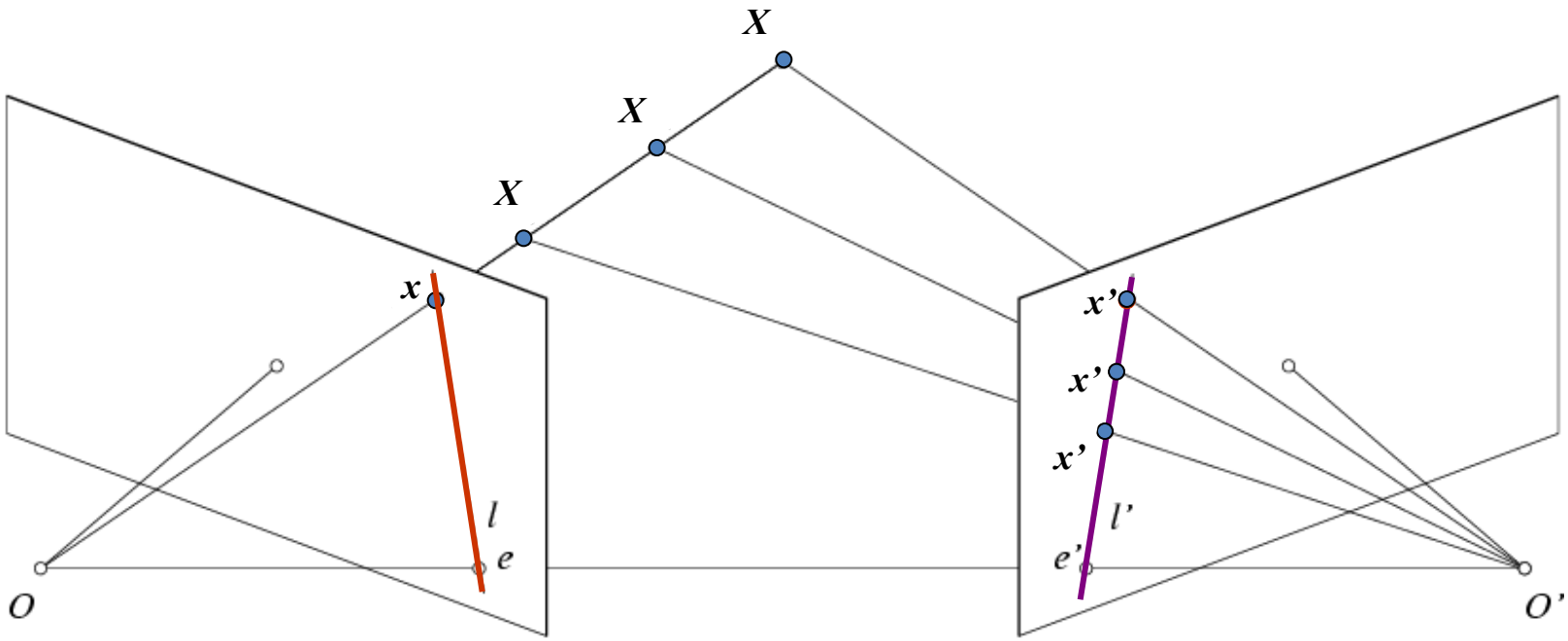


Correspondence Problem

- We have two images taken from cameras with different intrinsic and extrinsic parameters

- How do we match a point in the first image to a point in the second? How can we constrain our search?

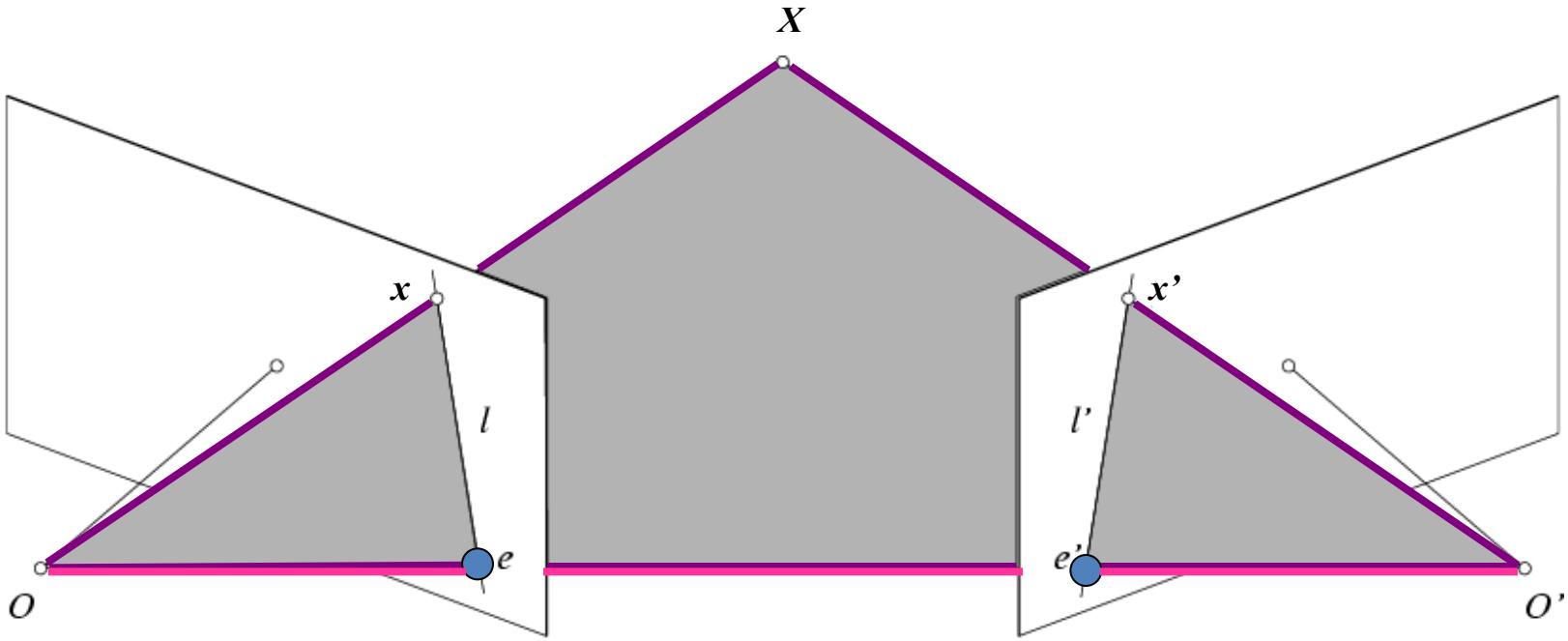
Key idea: Epipolar constraint



Potential matches for x have to lie on the corresponding line l'.

Potential matches for x' have to lie on the corresponding line l.

Epipolar geometry: notation



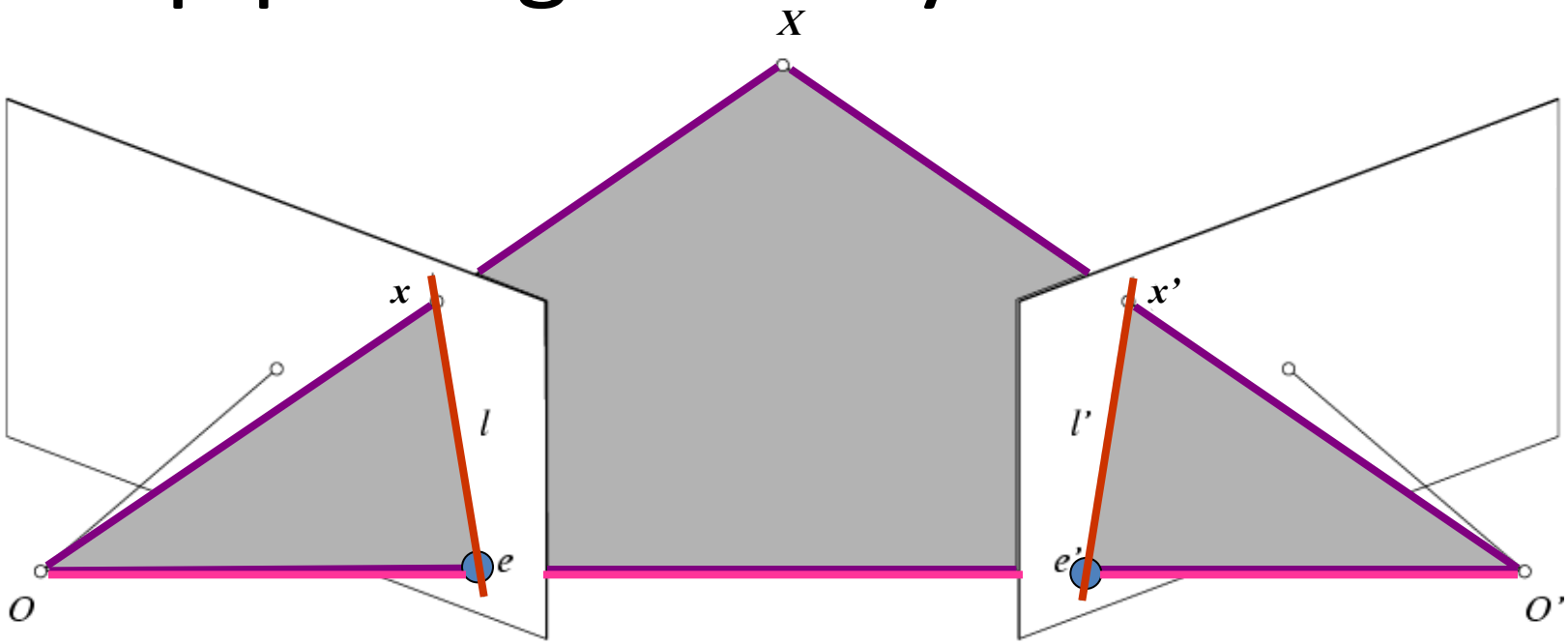
- **Baseline** – line connecting the two camera centers
- **Epipoles**

= intersections of baseline with image planes

= projections of the other camera center

- **Epipolar Plane** – plane containing baseline (1D family)

Epipolar geometry: notation



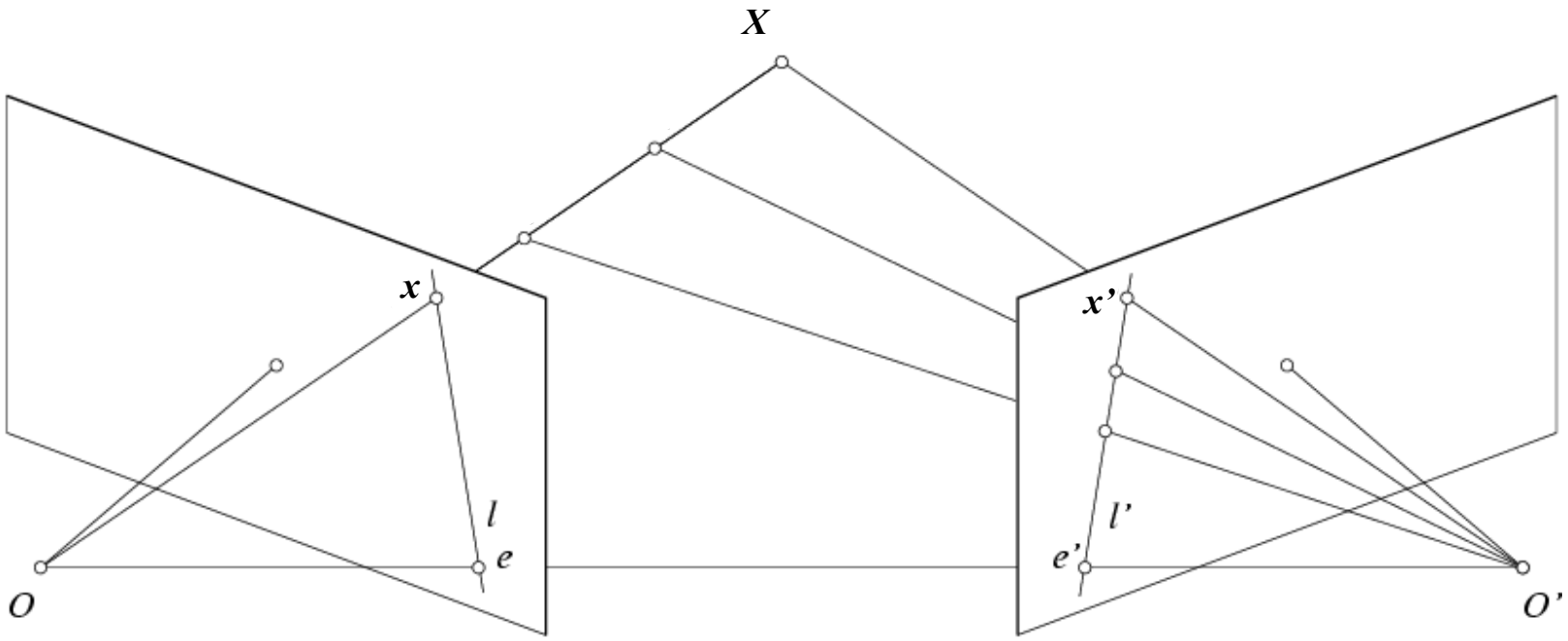
- **Baseline** – line connecting the two camera centers
- **Epipoles**

= intersections of baseline with image planes

= projections of the other camera center

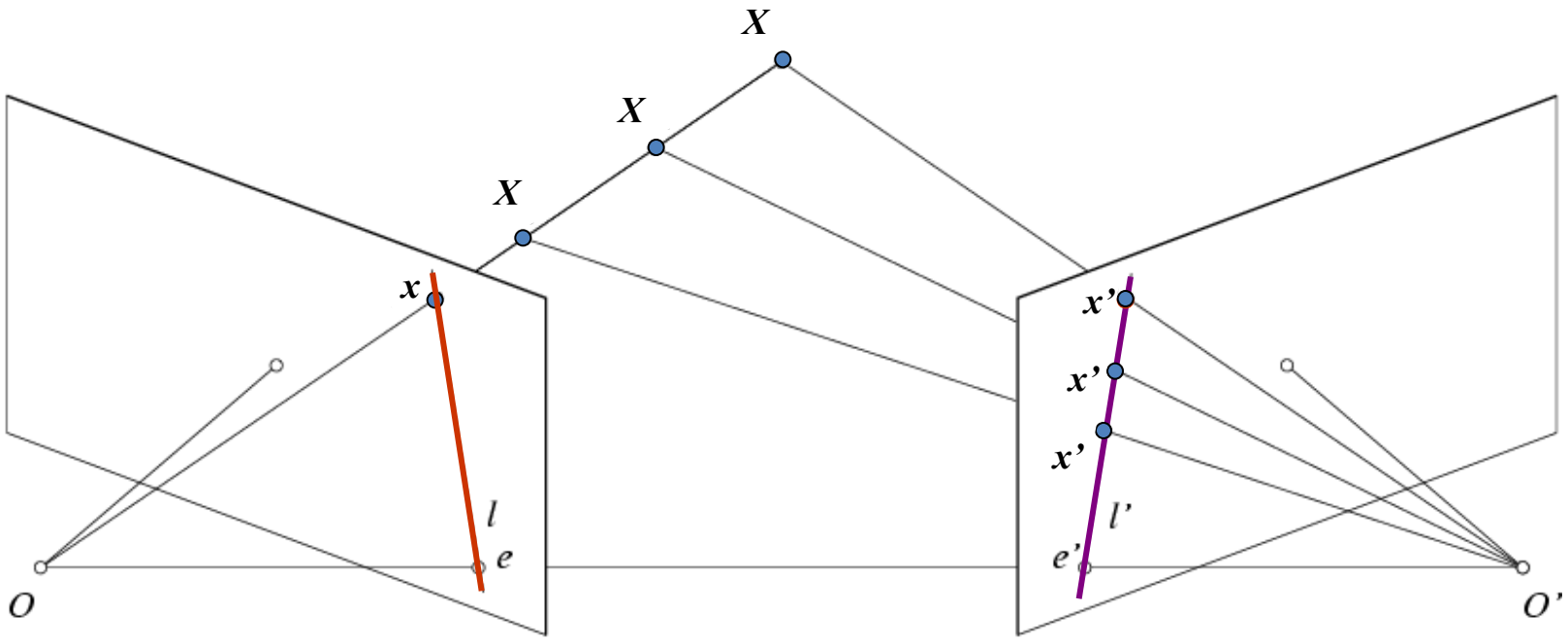
- **Epipolar Plane** – plane containing baseline (1D family)
- **Epipolar Lines** - intersections of epipolar plane with image planes (always come in corresponding pairs)

Epipolar constraint



- If we observe a point ***x*** in one image, where can the corresponding point ***x'*** be in the other image?

Epipolar constraint



- Potential matches for ***x*** have to lie on the corresponding epipolar line ***l'***.

- Potential matches for ***x'*** have to lie on the corresponding epipolar line ***l***.

Moving on to stereo…

Fuse a calibrated binocular stereo pair to produce a depth image

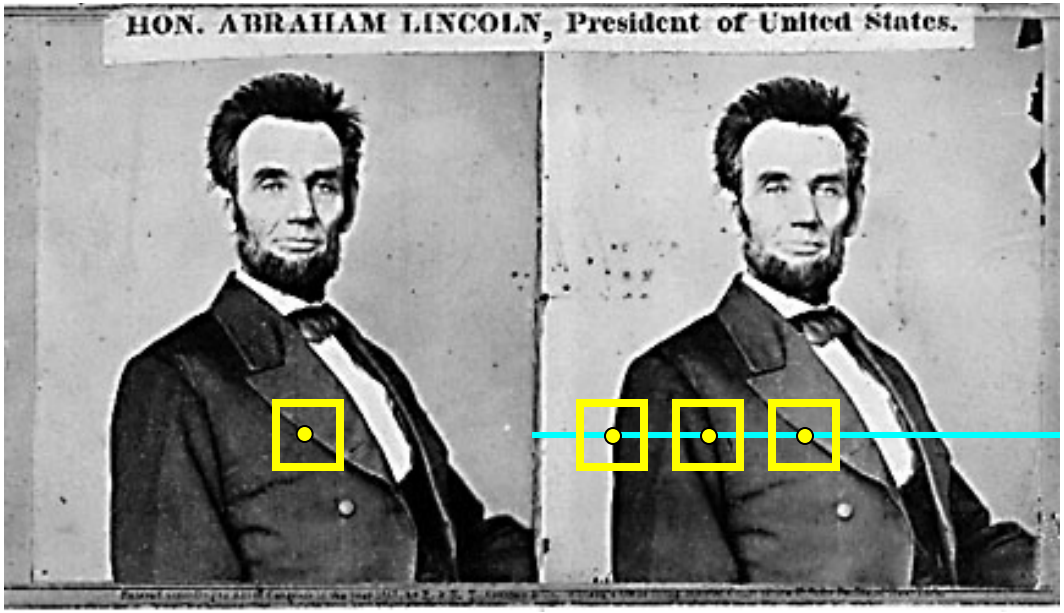
image 1

image 2

Dense depth map

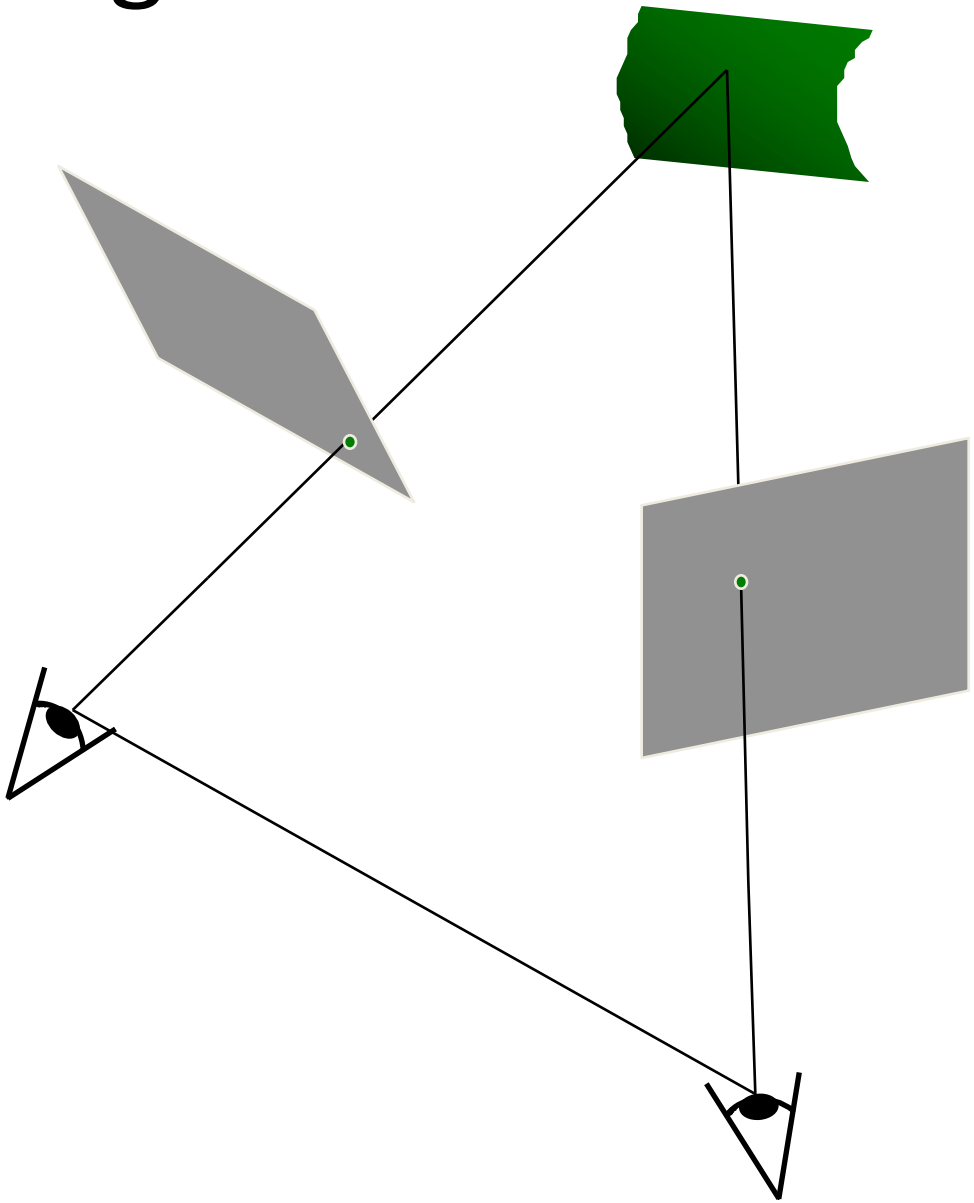
Many of these slides adapted from Steve Seitz and Lana Lazebnik

Basic stereo matching algorithm

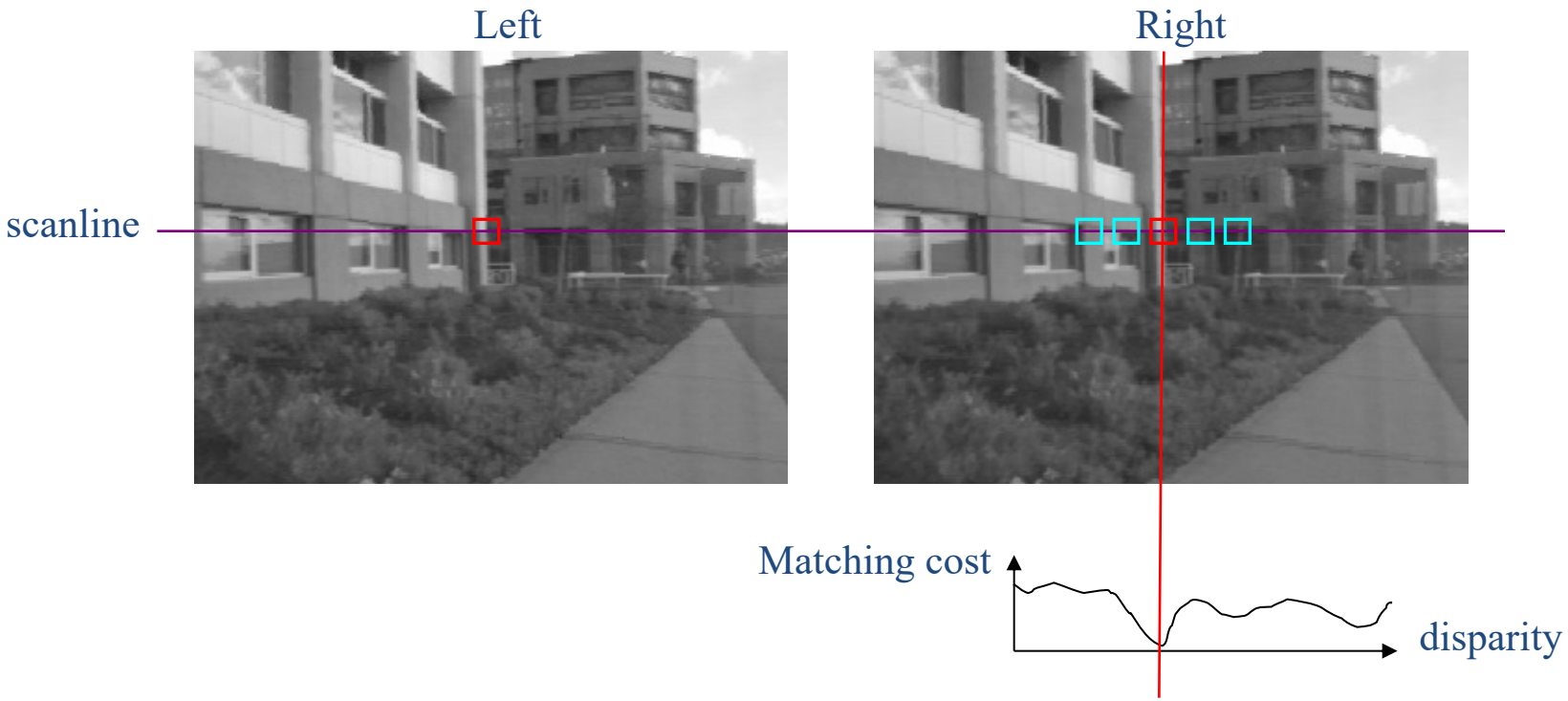


- If necessary, rectify the two stereo images to transform epipolar lines into scanlines
- For each pixel x in the first image
  - Find corresponding epipolar scanline in the right image
  - Search the scanline and pick the best match x'
  - Compute disparity x-x' and set depth(x) = fB/(x-x')

Stereo image rectification



Example



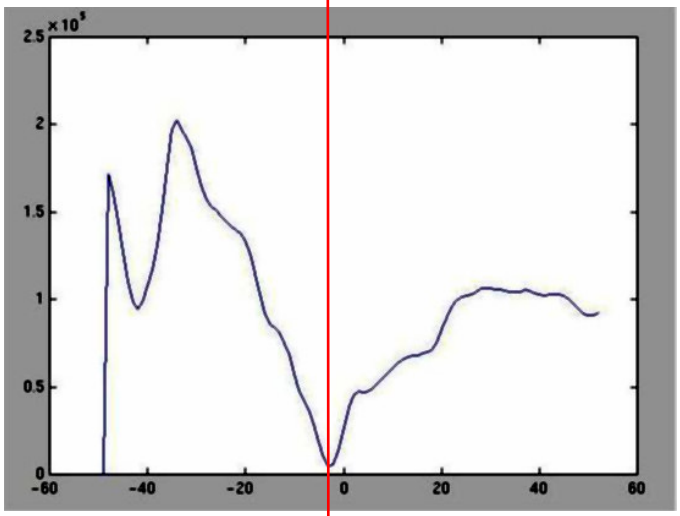
- Slide a window along the right scanline and compare contents of that window with the reference window in the left image
- Matching cost: SSD, SAD, or normalized cross correlation

Correspondence search

Left

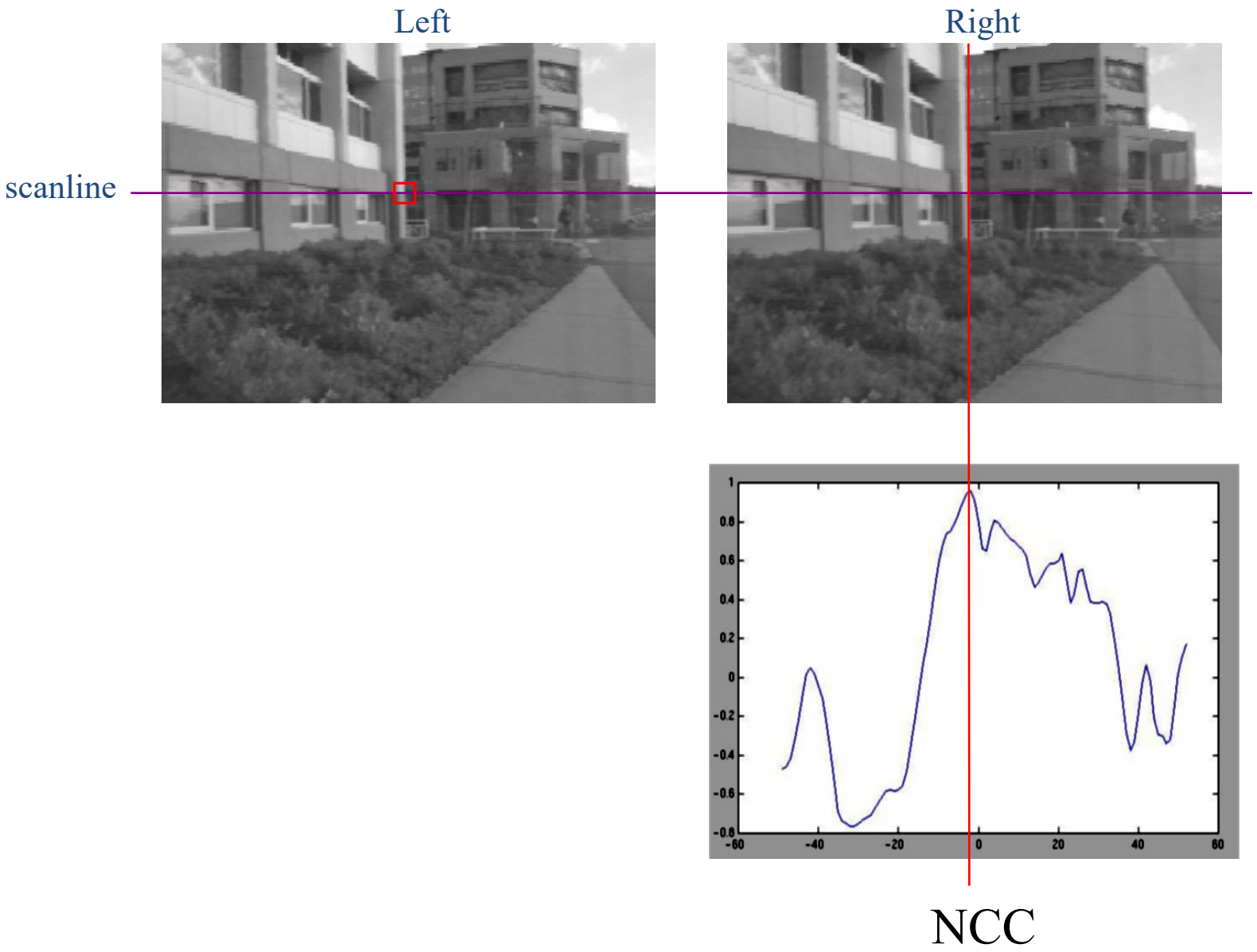
Right

scanline



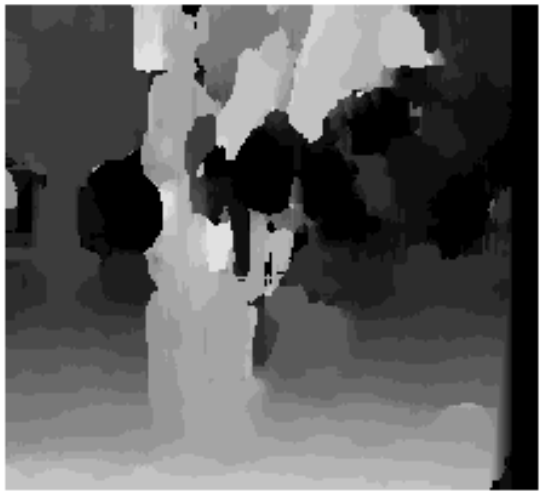
SSD

Correspondence search



Effect of window size

W = 3



W = 20

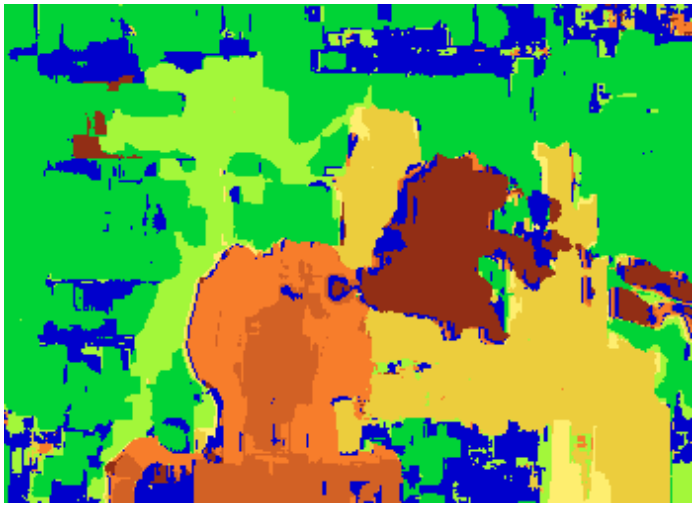
- Smaller window
  - \+ More detail
  - – More noise

- Larger window
  - \+ Smoother disparity maps
  - – Less detail
  - – Fails near boundaries

Results with window search

Data

Window-based matching



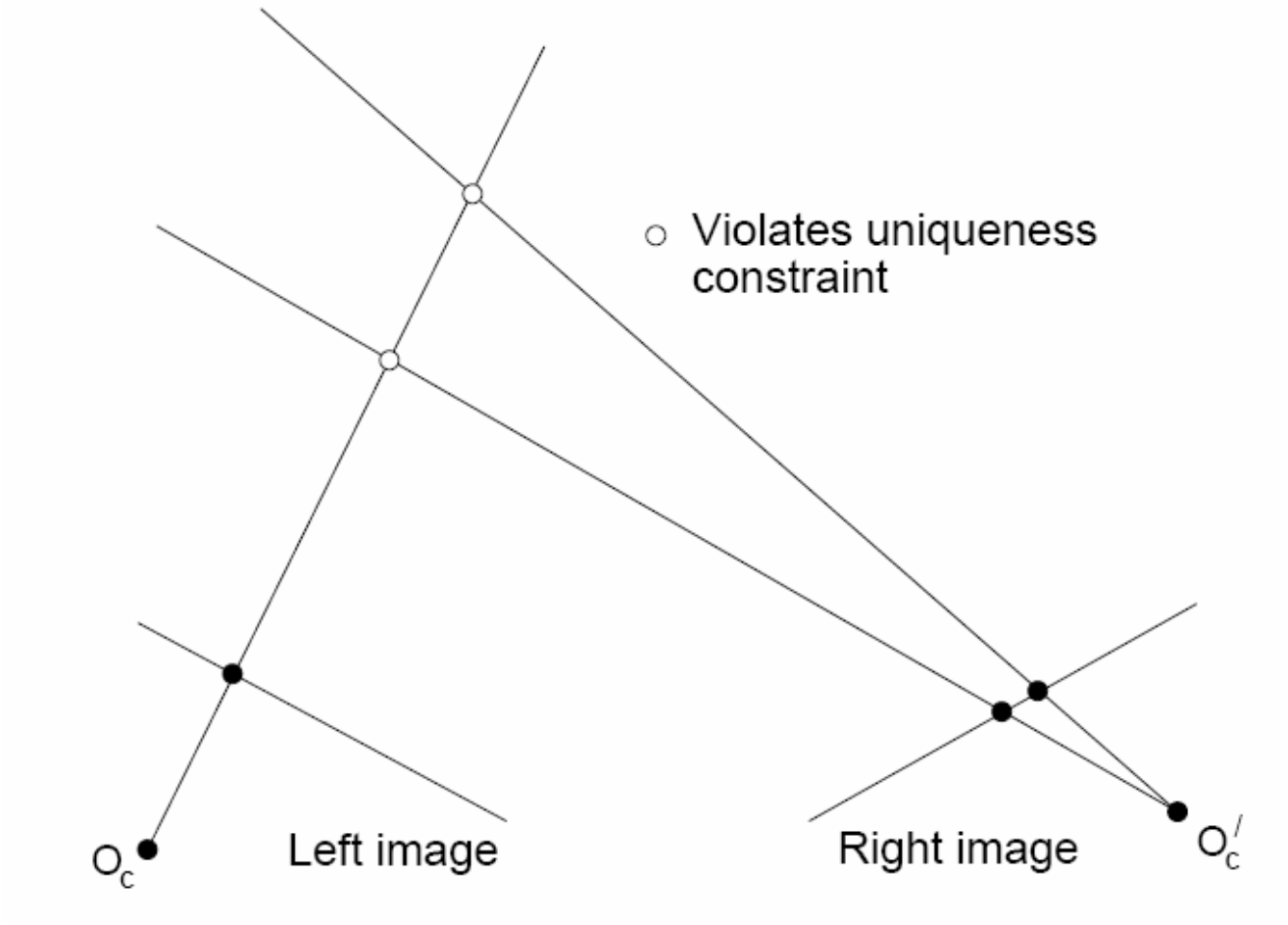
Ground truth

How can we improve window-based matching?

- So far, matches are independent for each point
- What constraints or priors can we add?

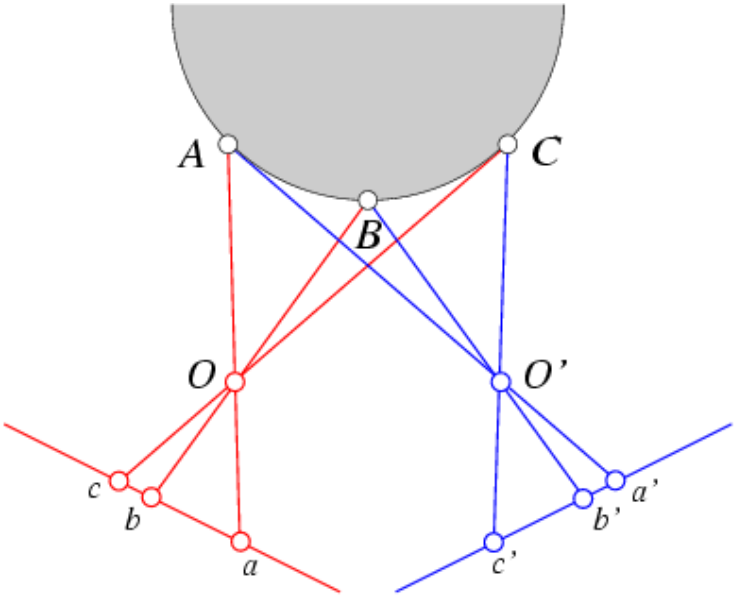
Stereo constraints/priors

- Uniqueness
  - For any point in one image, there should be at most one matching point in the other image



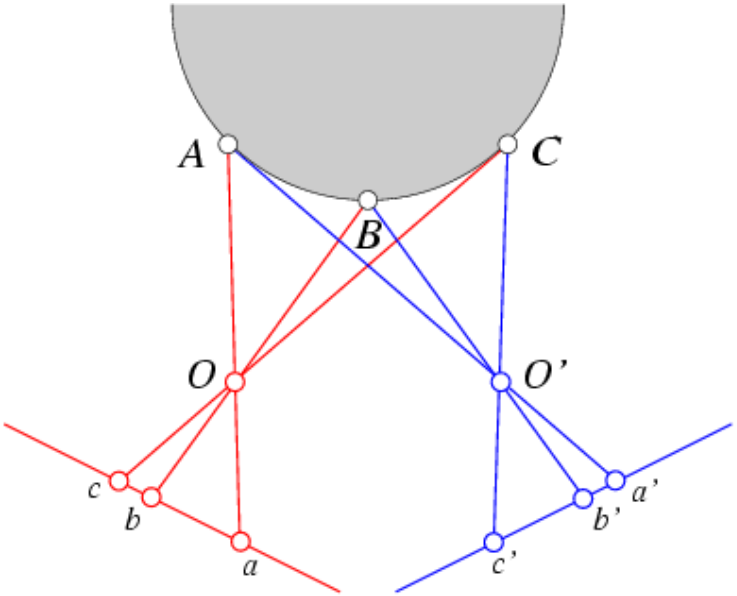
Stereo constraints/priors

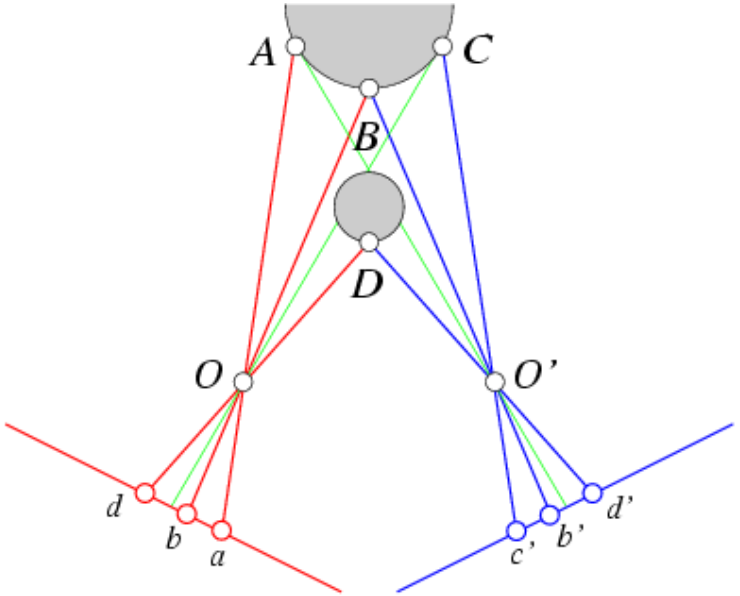
- Uniqueness
  - For any point in one image, there should be at most one matching point in the other image
- Ordering
  - Corresponding points should be in the same order in both views



Stereo constraints/priors

- Uniqueness
  - For any point in one image, there should be at most one matching point in the other image
- Ordering
  - Corresponding points should be in the same order in both views



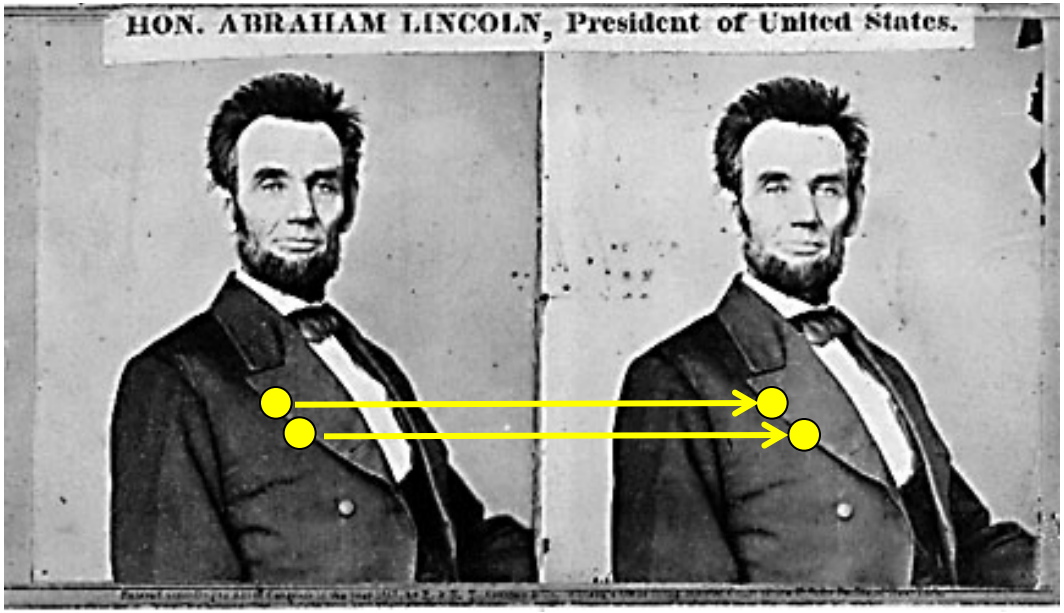


Ordering constraint doesn't hold

Priors and constraints

- Uniqueness
  - For any point in one image, there should be at most one matching point in the other image
- Ordering
  - Corresponding points should be in the same order in both views
- Smoothness
  - We expect disparity values to change slowly (for the most part)

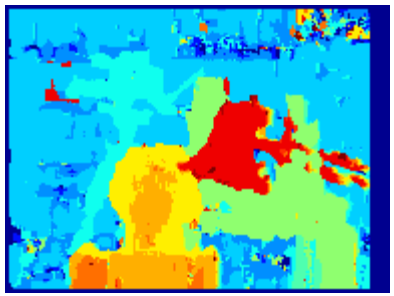
Stereo as energy minimization



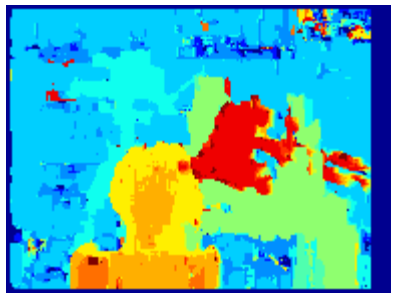
- What defines a good stereo correspondence?
  1. Match quality
     - Want each pixel to find a good match in the other image
  2. Smoothness
     - If two pixels are adjacent, they should (usually) move about the same amount

Matching windows:

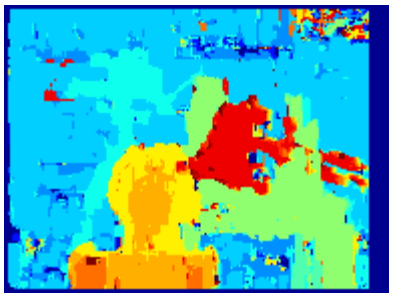
Similarity Measure	Formula
Sum of Absolute Differences (SAD)	$\sum_{(i,j)\in W} \lvert I_1(i,j) - I_2(x+i, y+j)\rvert$
Sum of Squared Differences (SSD)	$\sum_{(i,j)\in W} \left(I_1(i,j) - I_2(x+i, y+j)\right)^2$
Zero-mean SAD	$\sum_{(i,j)\in W} \lvert I_1(i,j) - \bar{I}_1(i,j) - I_2(x+i, y+j) + \bar{I}_2(x+i, y+j)\rvert$
Normalized Cross Correlation (NCC)	$\frac{\sum_{(i,j)\in W} I_1(i,j).I_2(x+i, y+j)}{\sqrt{\sum_{(i,j)\in W} I_1^2(i,j).\sum_{(i,j)\in W} I_2^2(x+i, y+j)}}$



SAD



SSD



NCC

Ground truth

http://siddhantahuja.wordpress.com/category/stereo-vision/

Stereo reconstruction pipeline

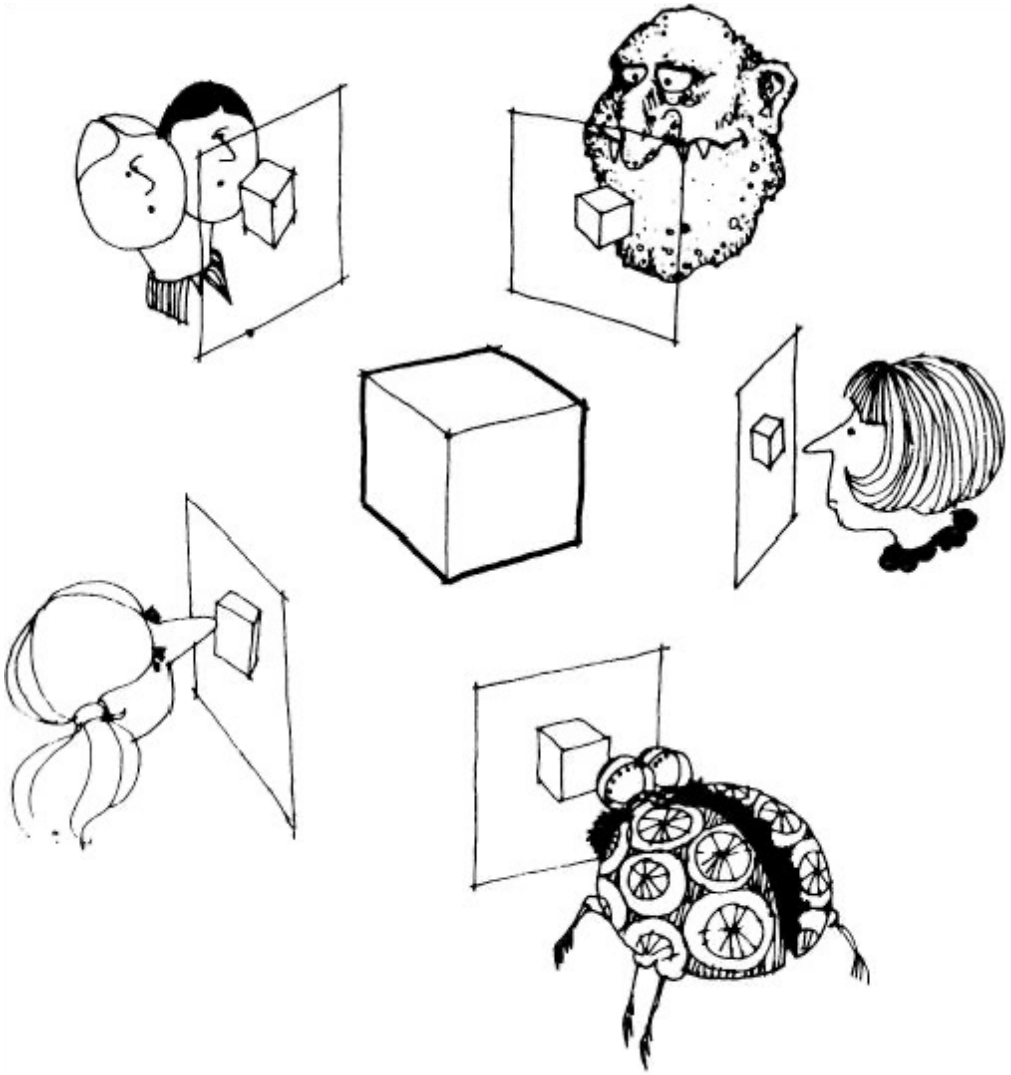
Steps

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

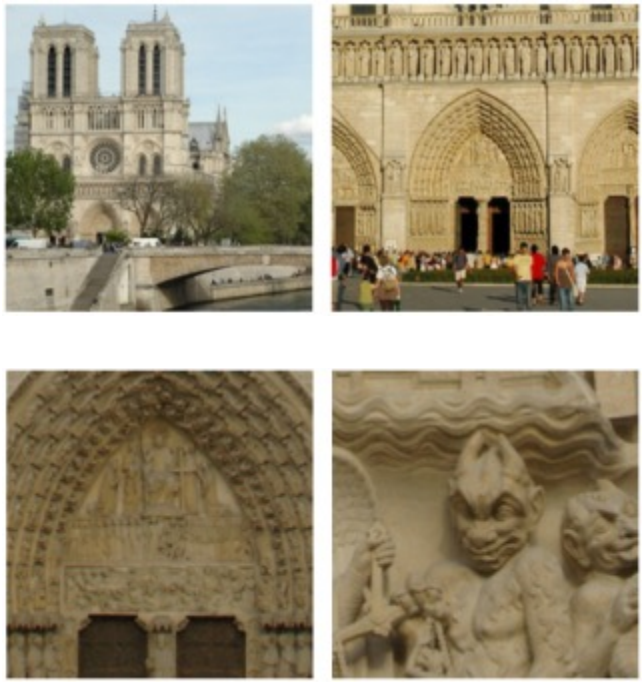
What will cause errors?

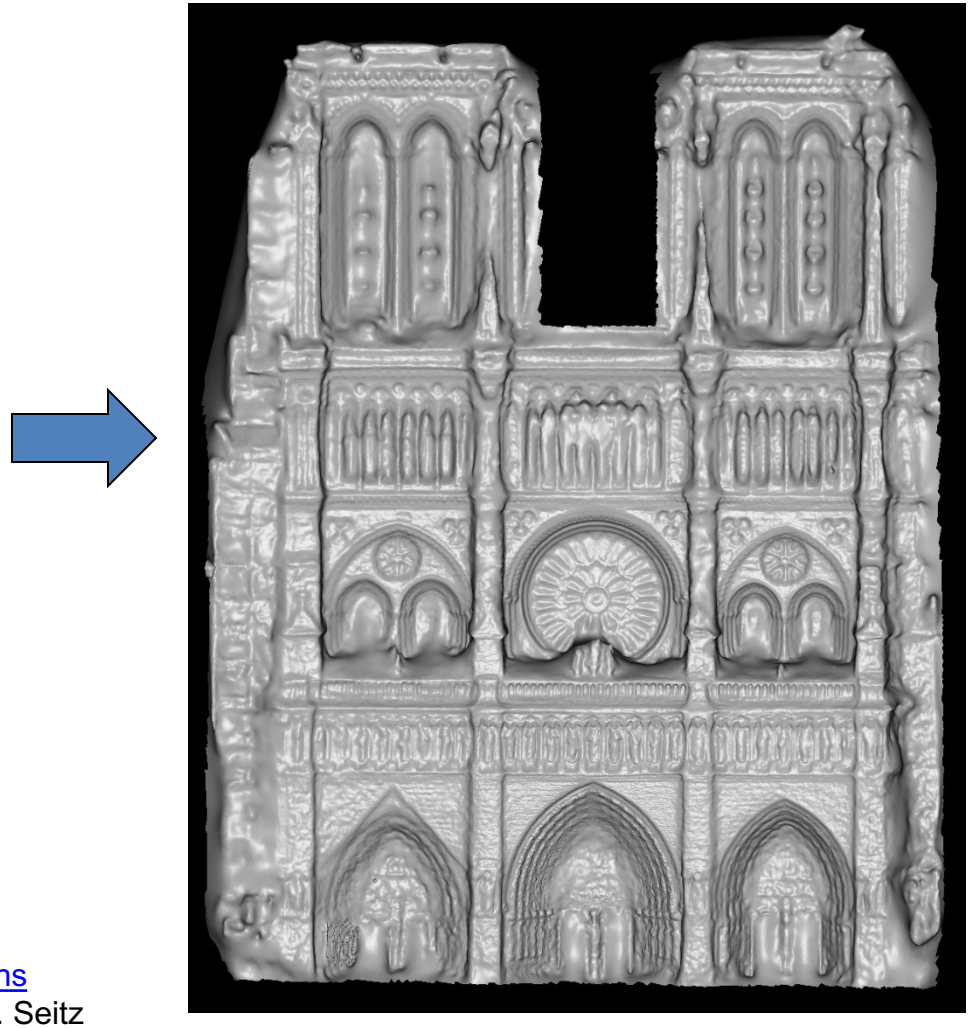
- Camera calibration errors
- Poor image resolution
- Occlusions
- Violations of brightness constancy
- Low-contrast image regions

Multi-view stereo ?



Using more than two images





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