Stereo

Many slides adapted from Steve Seitz

Binocular stereo

- Given a calibrated binocular stereo pair, fuse it to produce a depth image

image 1

image 2

Dense depth map

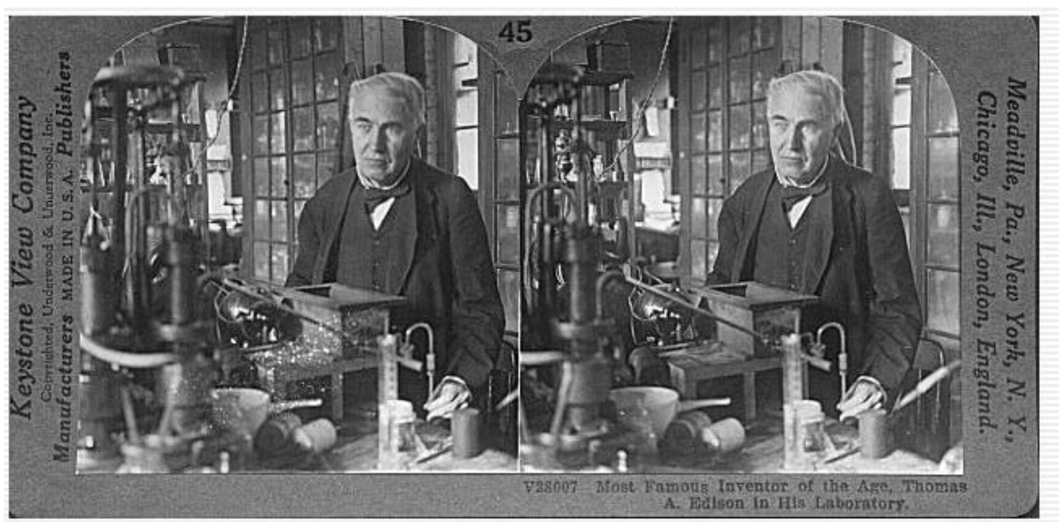
Binocular stereo

- Given a calibrated binocular stereo pair, fuse it to produce a depth image

Where does the depth information come from?

Binocular stereo

- Given a calibrated binocular stereo pair, fuse it to produce a depth image
  - Humans can do it



Stereograms: Invented by Sir Charles Wheatstone, 1838

Binocular stereo

- Given a calibrated binocular stereo pair, fuse it to produce a depth image
  - Humans can do it

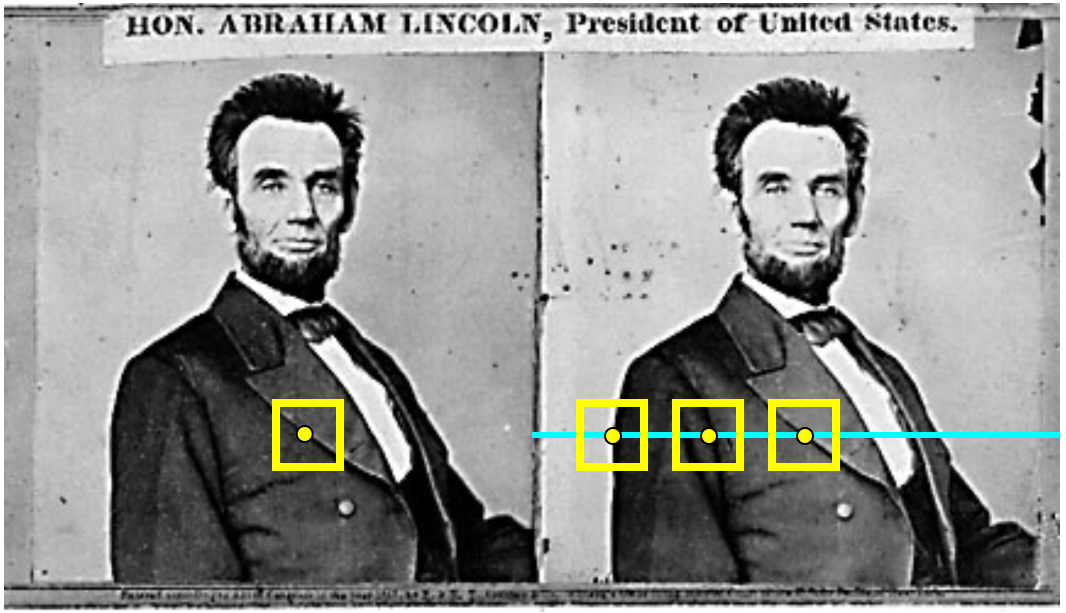
Autostereograms: www.magiceye.com

Binocular stereo

- Given a calibrated binocular stereo pair, fuse it to produce a depth image
  - Humans can do it

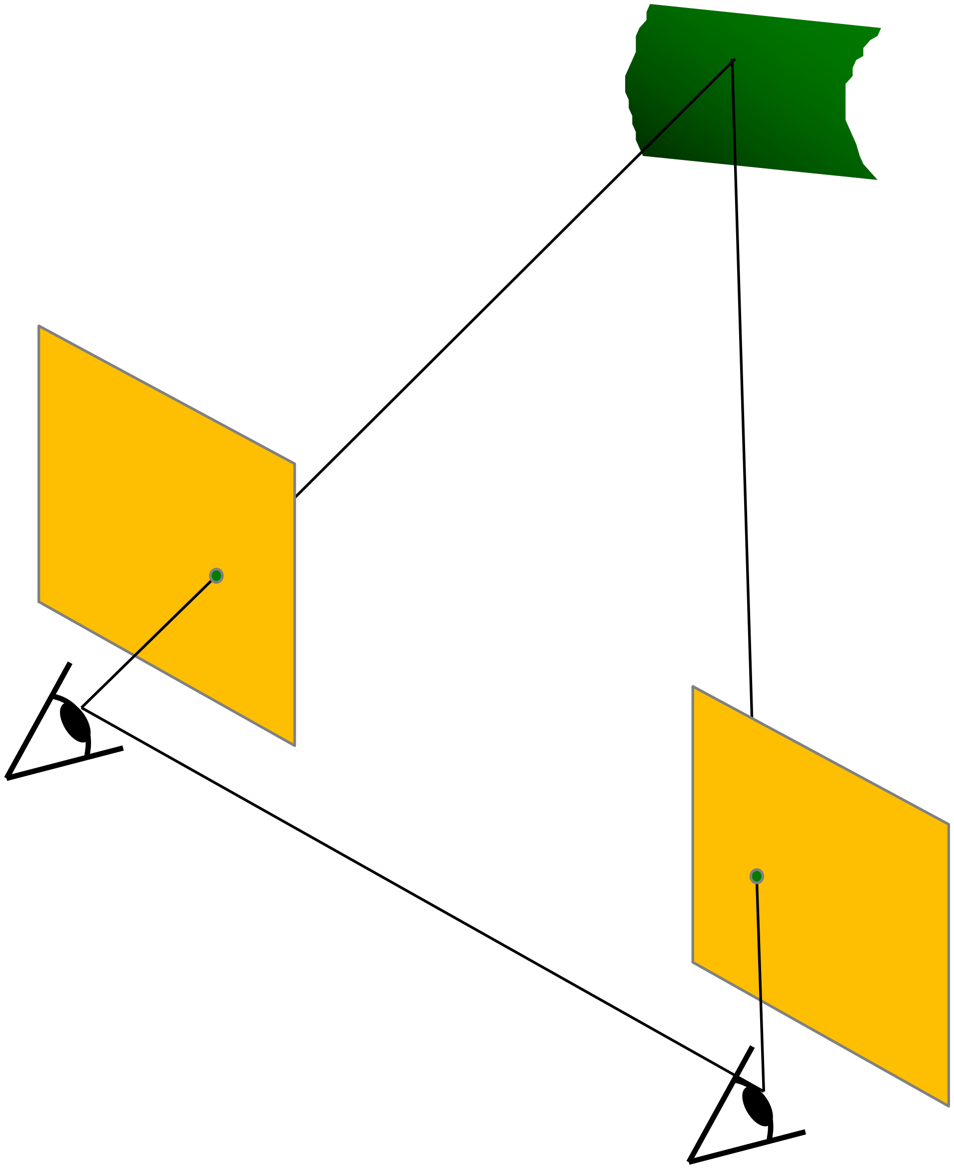
Autostereograms: www.magiceye.com

Basic stereo matching algorithm



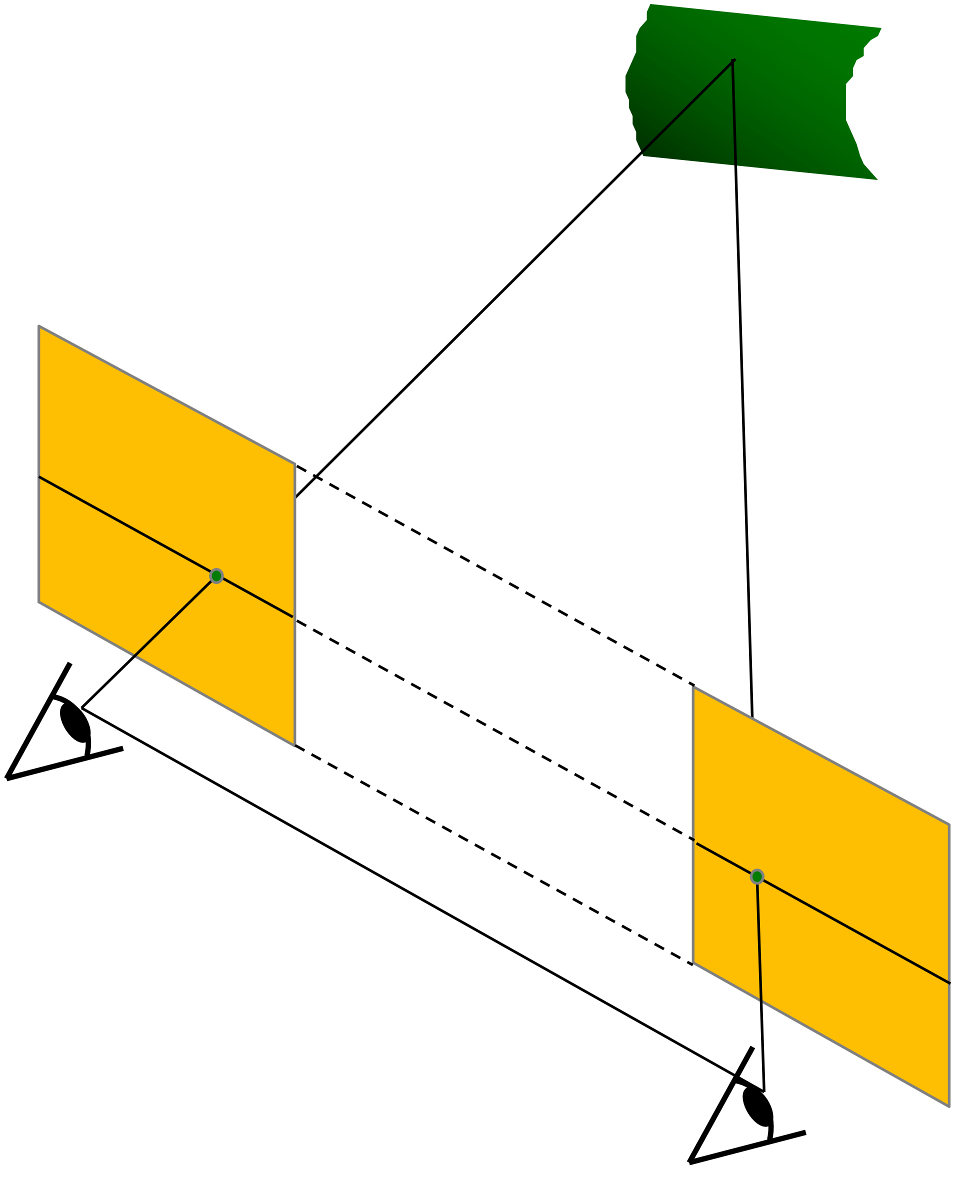
- For each pixel in the first image
  - Find corresponding epipolar line in the right image
  - Examine all pixels on the epipolar line and pick the best match
  - Triangulate the matches to get depth information
- Simplest case: epipolar lines are corresponding scanlines
  - When does this happen?

Simplest Case: Parallel images



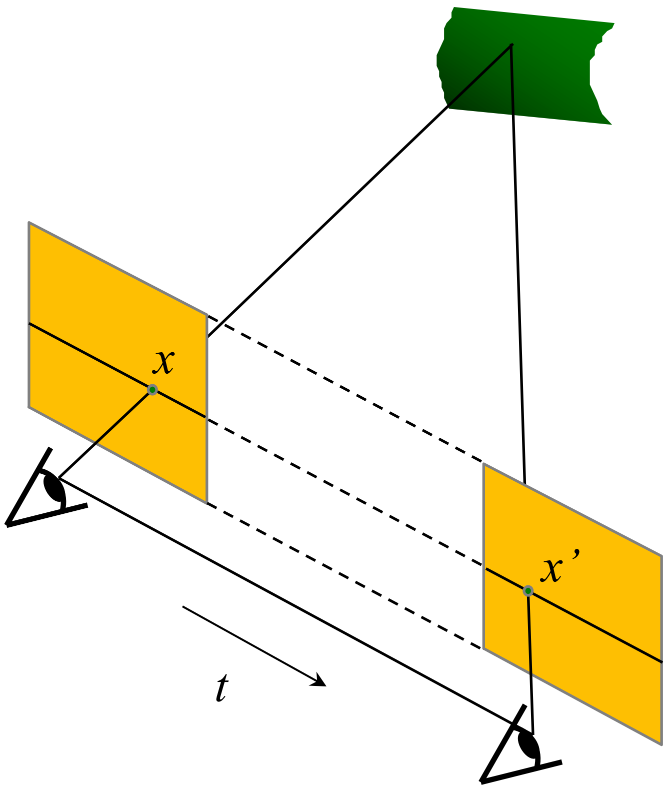
- Image planes of cameras are parallel to each other and to the baseline
- Camera centers are at same height
- Focal lengths are the same

Simplest Case: Parallel images



- Image planes of cameras are parallel to each other and to the baseline
- Camera centers are at same height
- Focal lengths are the same
- Then epipolar lines fall along the horizontal scan lines of the images

Essential matrix for parallel images



Epipolar constraint:

$$\boldsymbol{x}'^T \boldsymbol{E} \boldsymbol{x} = 0, \quad \boldsymbol{E} = [\boldsymbol{t}_\times]\boldsymbol{R}$$

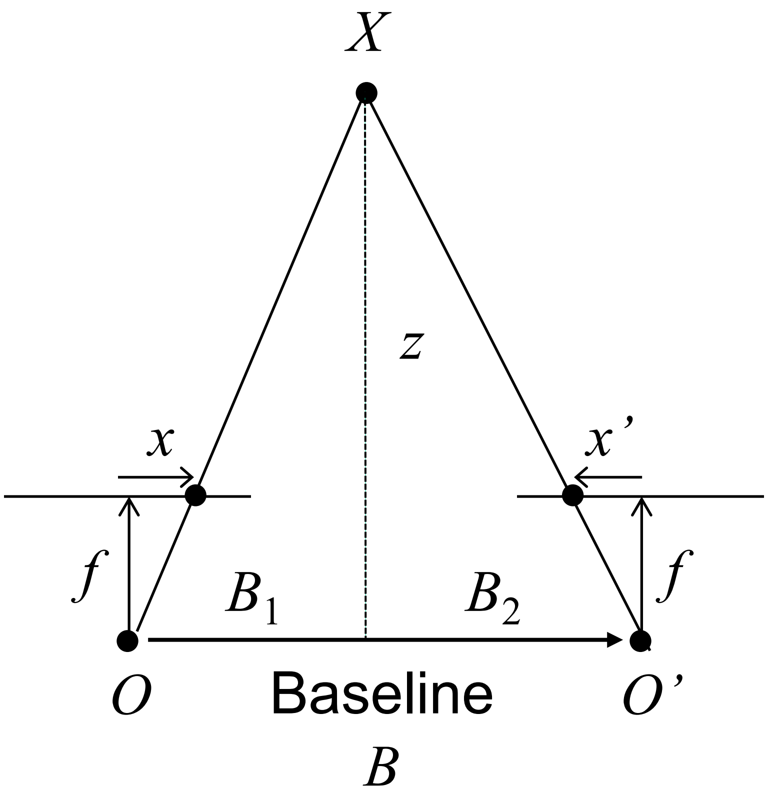
$$\boldsymbol{R} = \boldsymbol{I} \qquad \boldsymbol{t} = (T, 0, 0)$$

$$\boldsymbol{E} = [\boldsymbol{t}_\times]\boldsymbol{R} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -T \\ 0 & T & 0 \end{bmatrix}$$

$$\begin{pmatrix} u' & v' & 1 \end{pmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -T \\ 0 & T & 0 \end{bmatrix} \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = 0 \qquad \begin{pmatrix} u' & v' & 1 \end{pmatrix} \begin{pmatrix} 0 \\ -T \\ Tv \end{pmatrix} = 0 \qquad Tv' = Tv$$

The y-coordinates of corresponding points are the same!

Depth from disparity



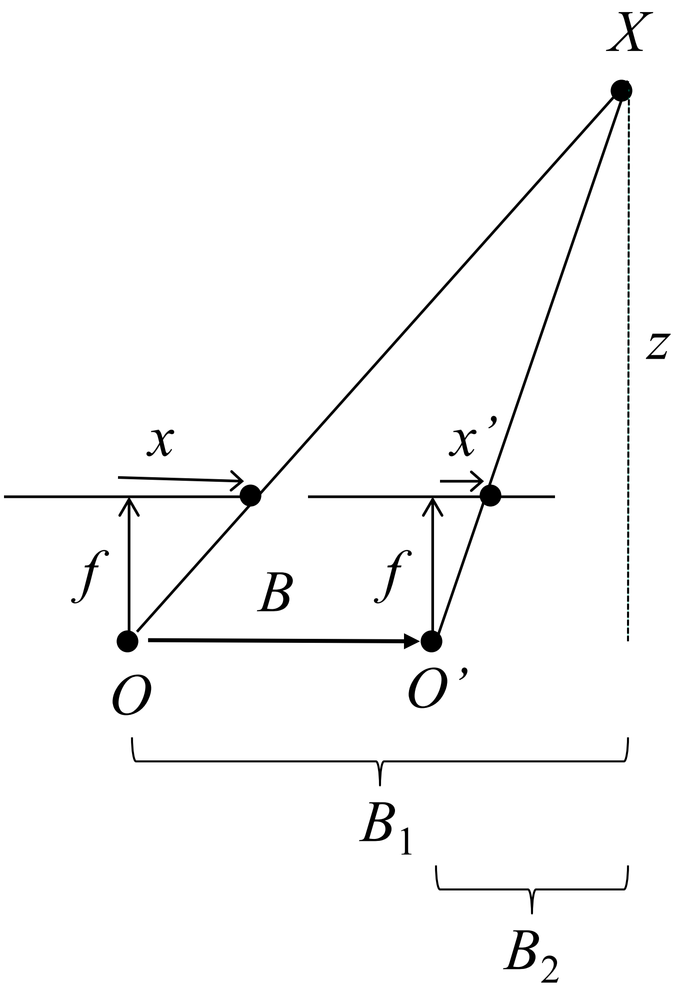
$$\frac{x}{f} = \frac{B_1}{z} \qquad \frac{-x'}{f} = \frac{B_2}{z}$$

$$\frac{x - x'}{f} = \frac{B_1 + B_2}{z}$$

$$\boxed{disparity = x - x' = \frac{B \cdot f}{z}}$$

Disparity is inversely proportional to depth!

Depth from disparity

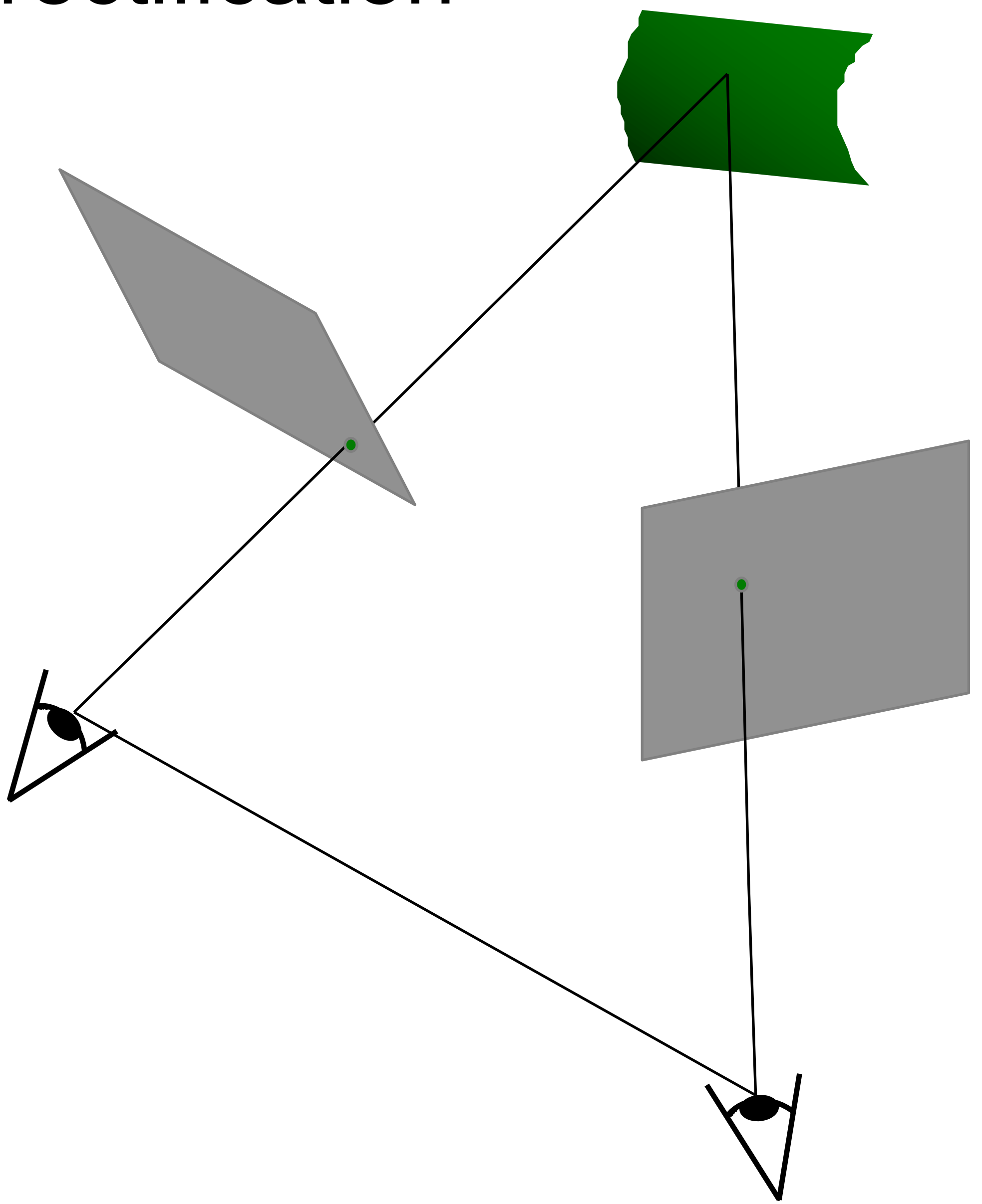


$$\frac{x}{f} = \frac{B_1}{z} \qquad \frac{x'}{f} = \frac{B_2}{z}$$

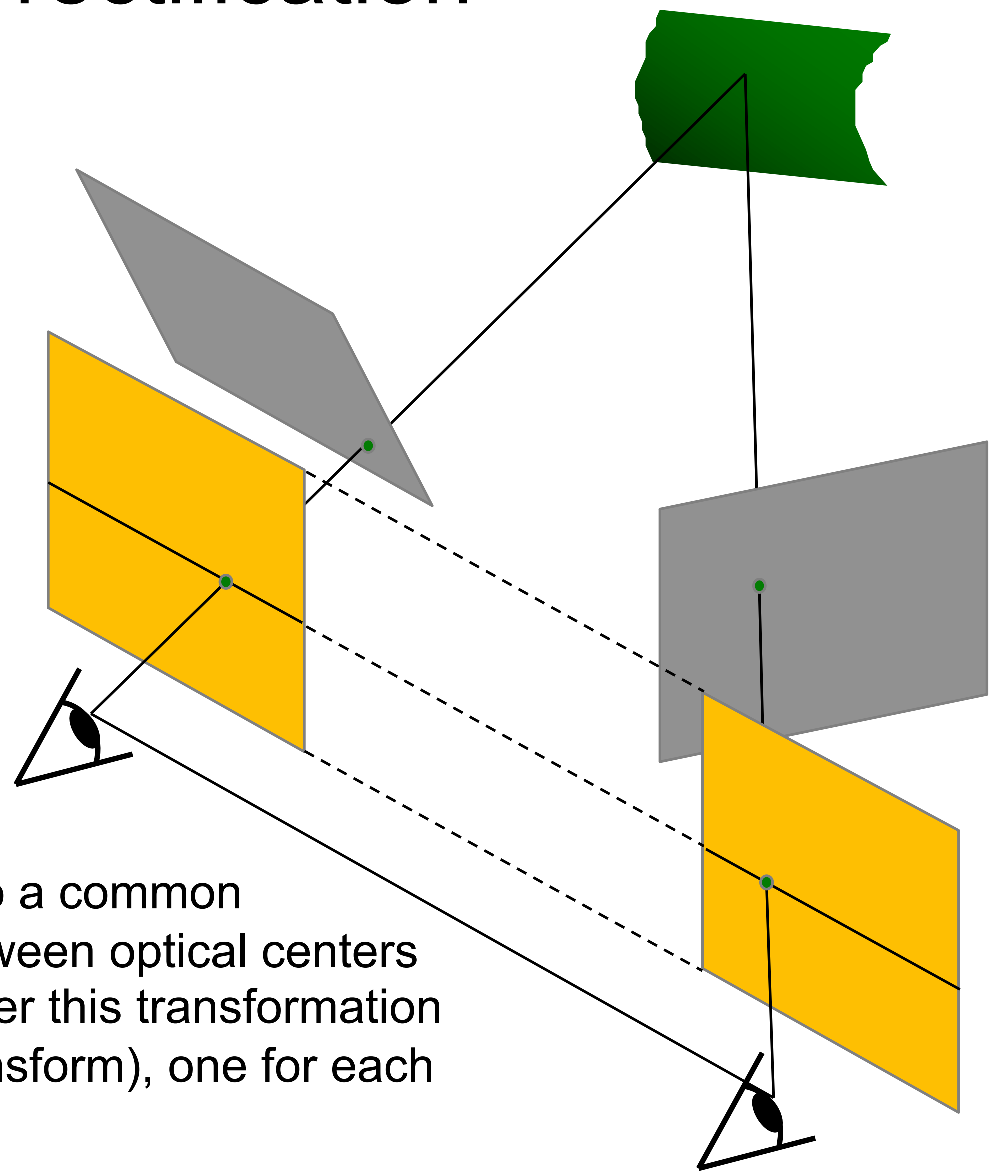
$$\frac{x - x'}{f} = \frac{B_1 - B_2}{z}$$

$$\boxed{disparity = x - x' = \frac{B \cdot f}{z}}$$

Stereo image rectification



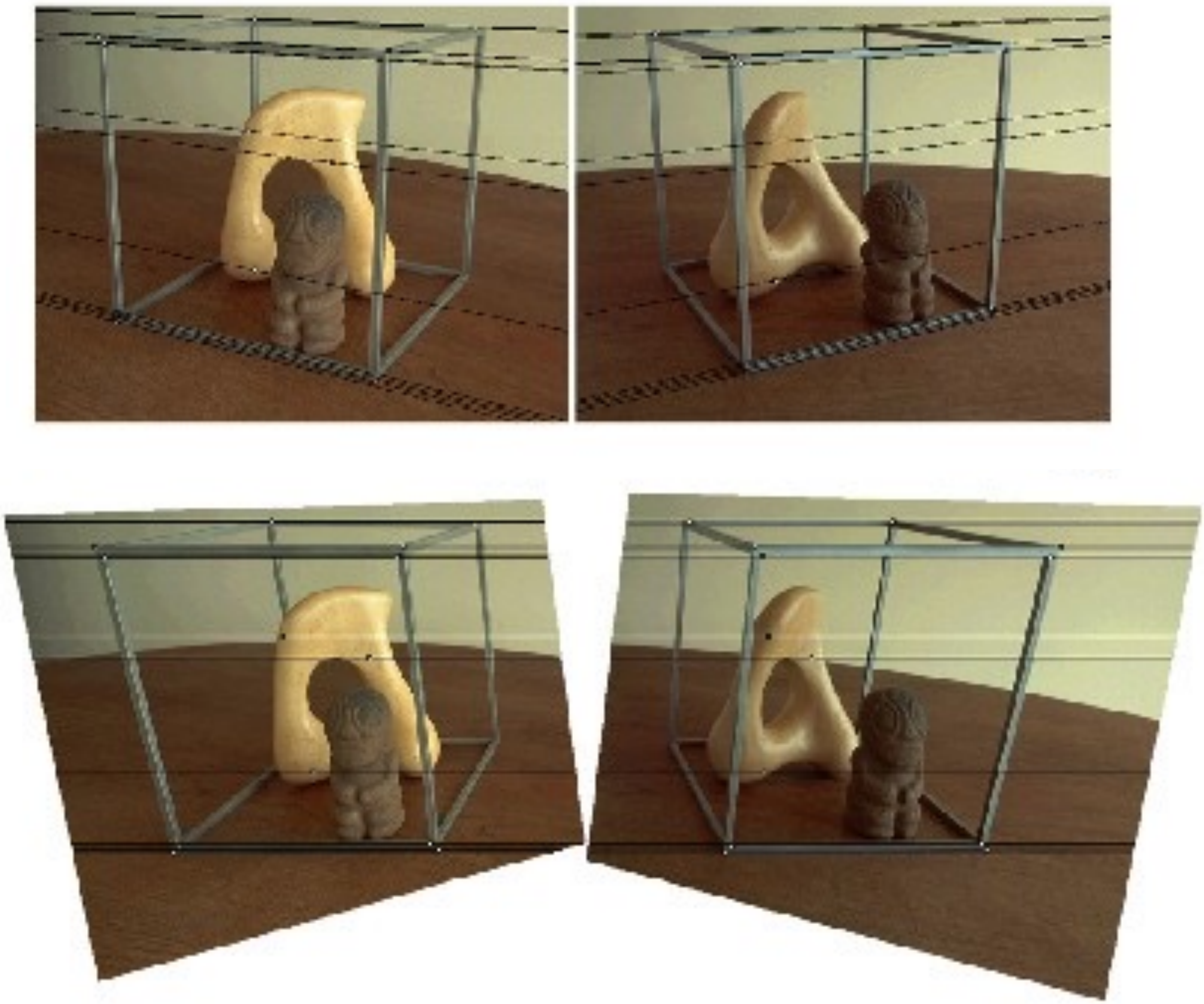
Stereo image rectification



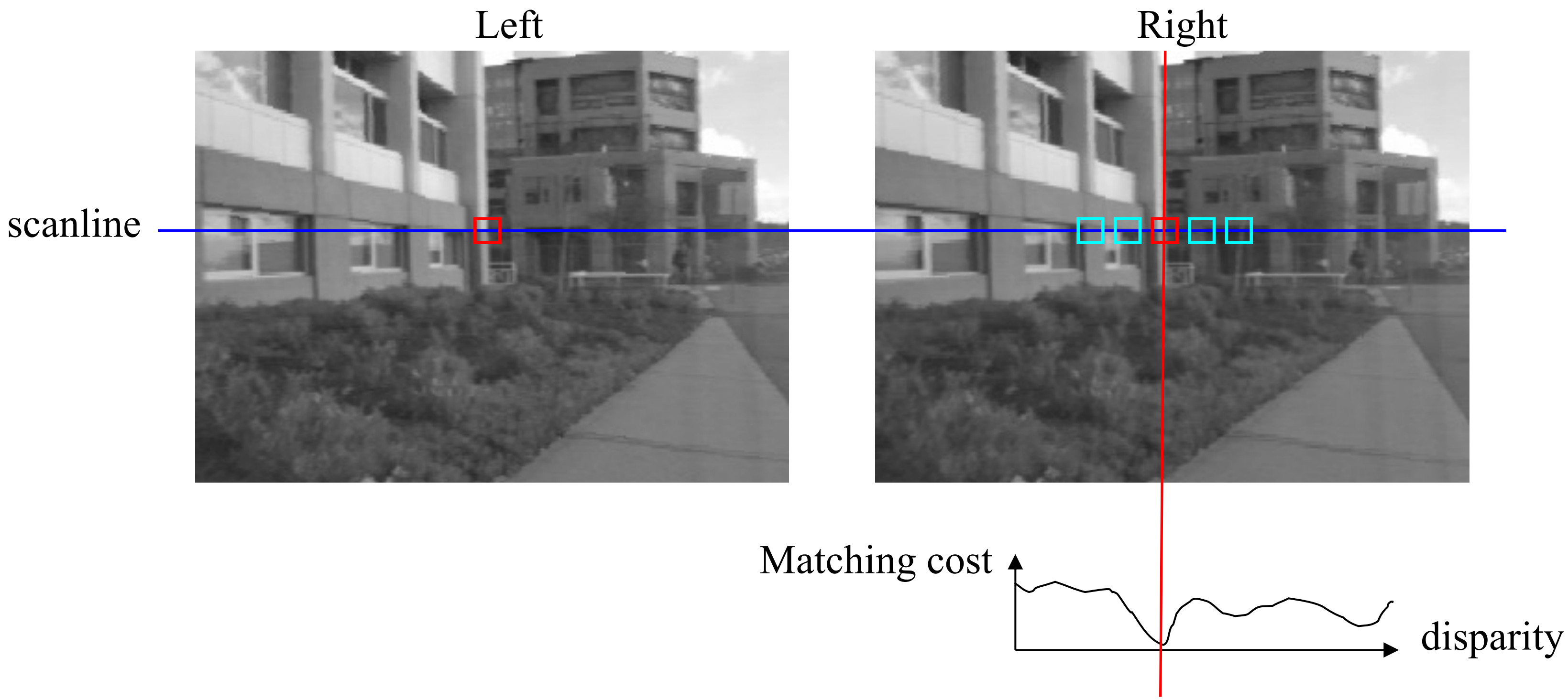
- Reproject image planes onto a common plane parallel to the line between optical centers
- Pixel motion is horizontal after this transformation
- Two homographies (3x3 transform), one for each input image reprojection

C. Loop and Z. Zhang. [Computing Rectifying Homographies for Stereo Vision](). IEEE Conf. Computer Vision and Pattern Recognition, 1999.

Rectification example

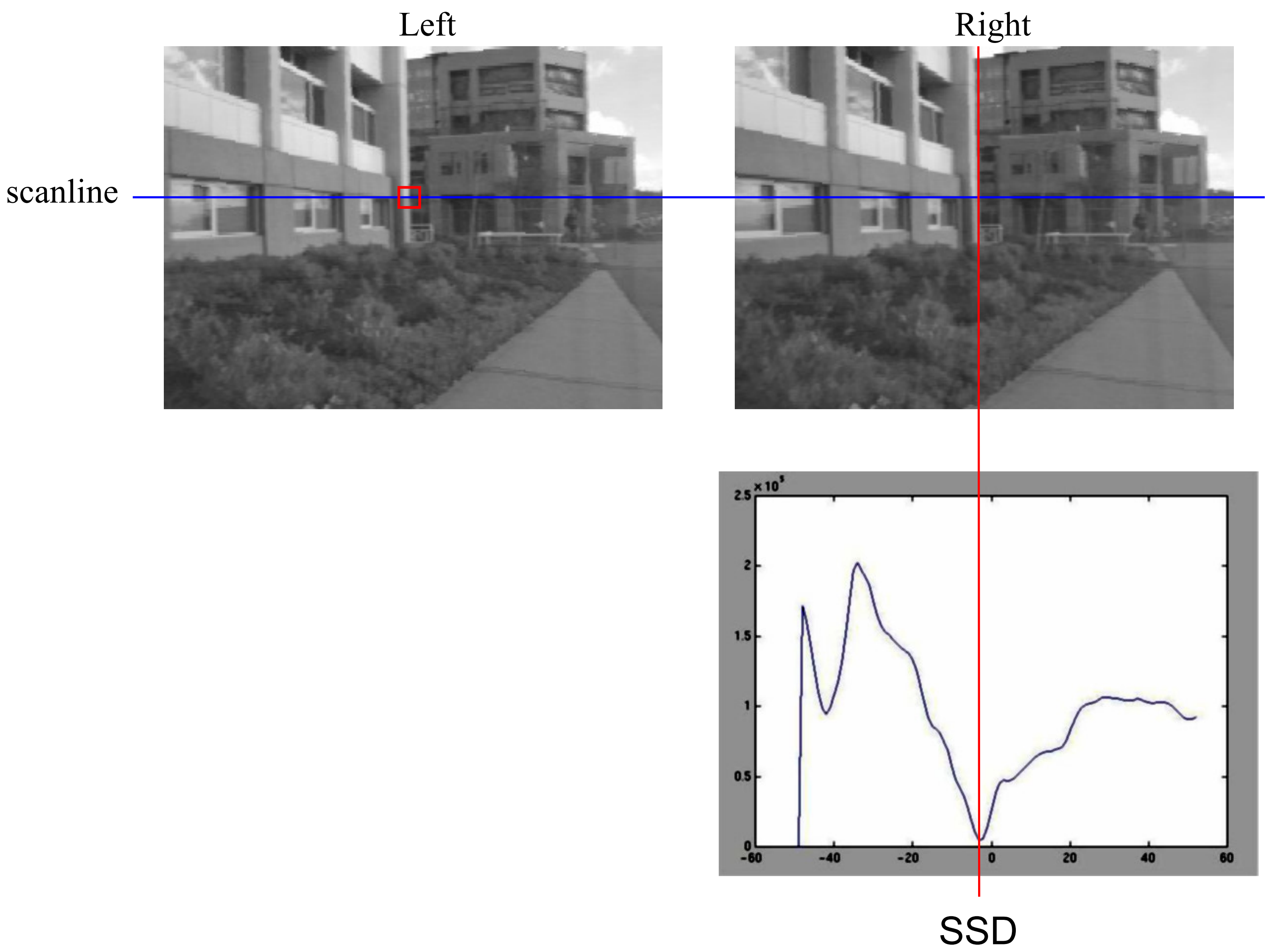


Correspondence search

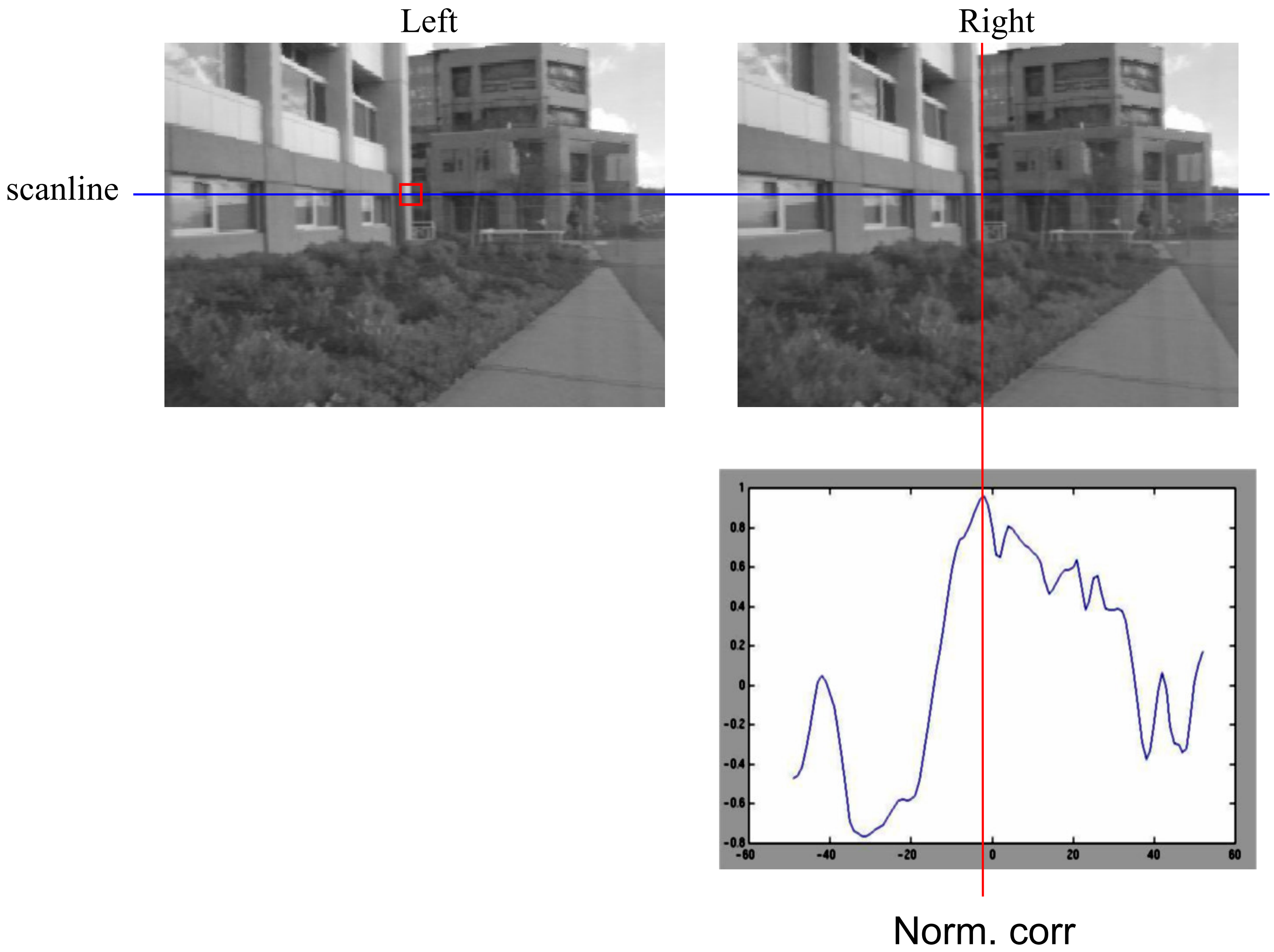


- Slide a window along the right scanline and compare contents of that window with the reference window in the left image
- Matching cost: SSD or normalized correlation

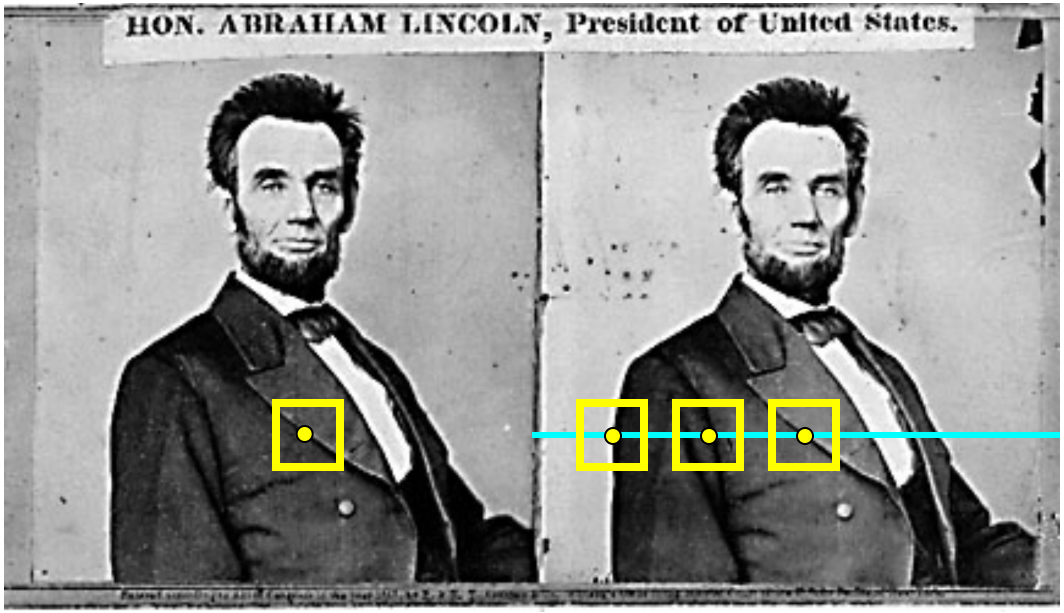
Correspondence search



Correspondence search

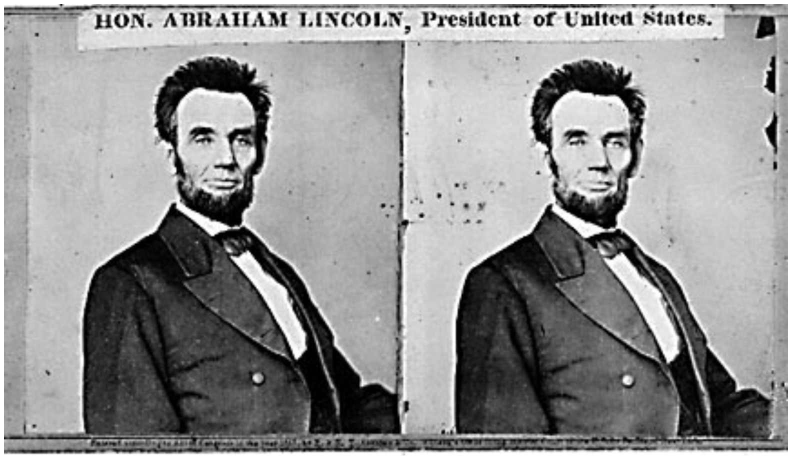


Basic stereo matching algorithm



- 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
  - Examine all pixels on the scanline and pick the best match *x'*
  - Compute disparity *x–x'* and set depth(*x*) = *B***f*/(*x–x'*)

Failures of correspondence search



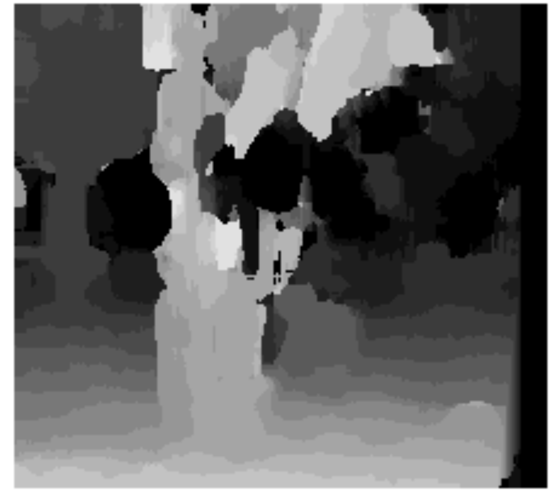
Textureless surfaces

Occlusions, repetition

Non-Lambertian surfaces, specularities

Effect of window size

W = 3



W = 20

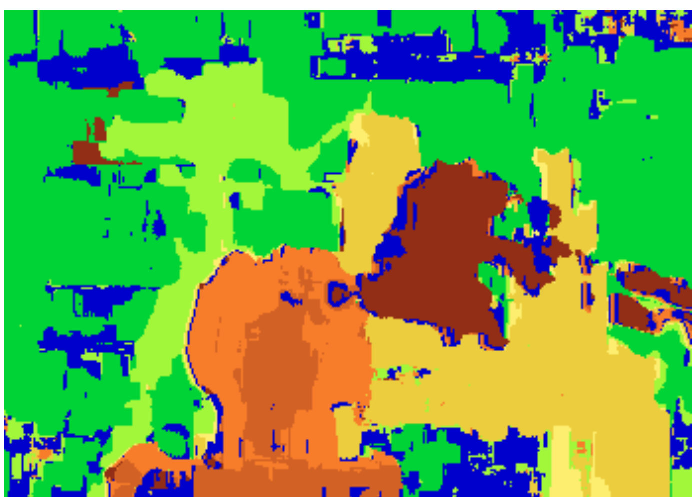
- Smaller window
  - \+ More detail
  - – More noise

- Larger window
  - \+ Smoother disparity maps
  - – Less detail

Results with window search

Data

Window-based matching



Ground truth

Better methods exist...

Graph cuts

Ground truth

Y. Boykov, O. Veksler, and R. Zabih, [Fast Approximate Energy Minimization via Graph Cuts](), PAMI 2001

For the latest and greatest: http://www.middlebury.edu/stereo/