Computer Vision

CS-E4850, 5 study credits

Lecturer: Juho Kannala

Lecture 11: Two-view geometry & stereo vision

- **Two-view geometry** (a.k.a. epipolar geometry) describes the geometric constraints between two views

- **Stereo vision** is the principle of using two views to measure depths of scene points

Acknowledgement: many slides from Svetlana Lazebnik, Steve Seitz, Yuri Boykov, Noah Snavely, and others (detailed credits on individual slides)

Reading

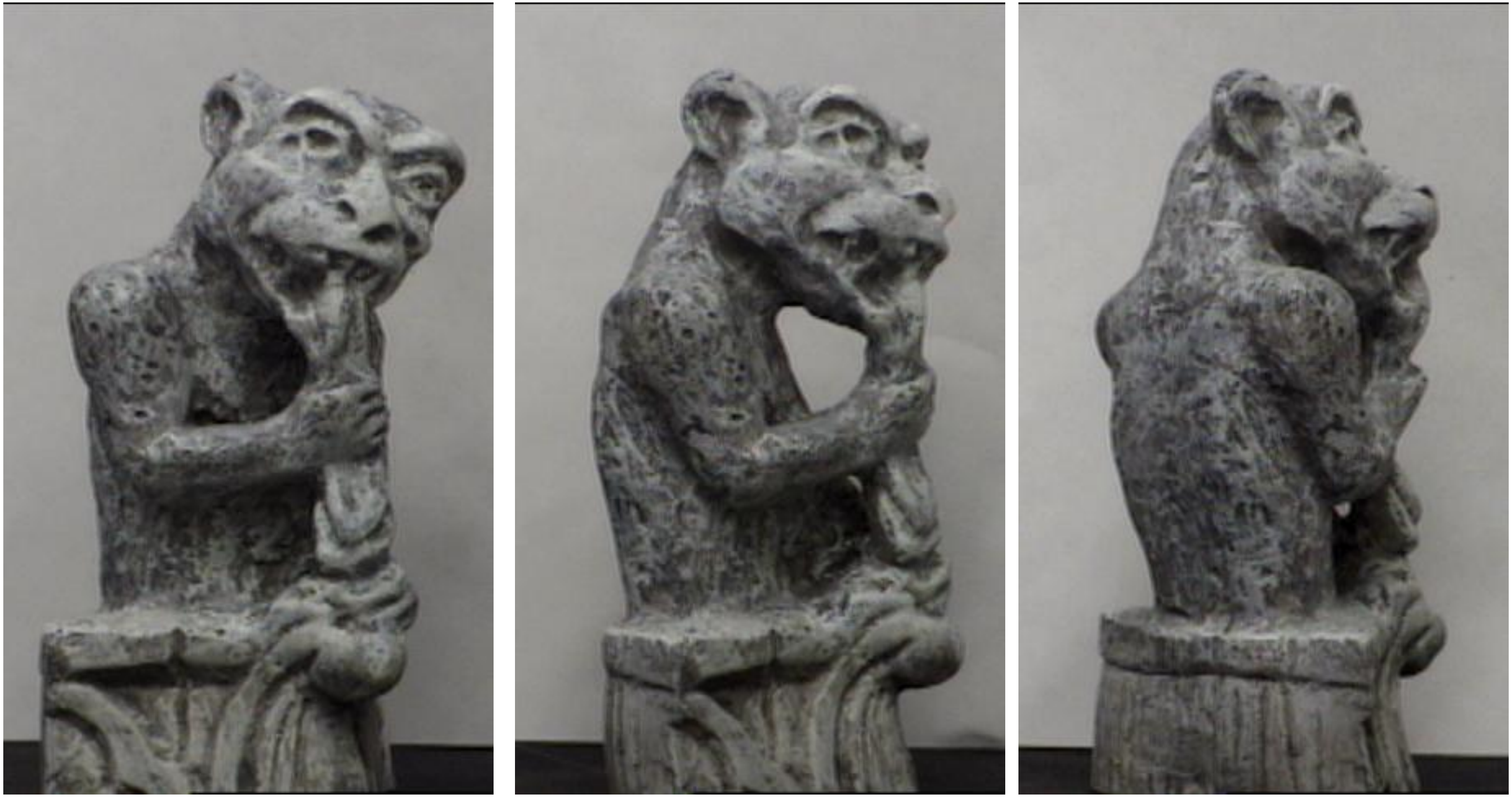
- Szeliski's book, Section 7.2 and Chapter 11 in 1st edition

and/or

- Hartley & Zisserman book, Chapters 9-12

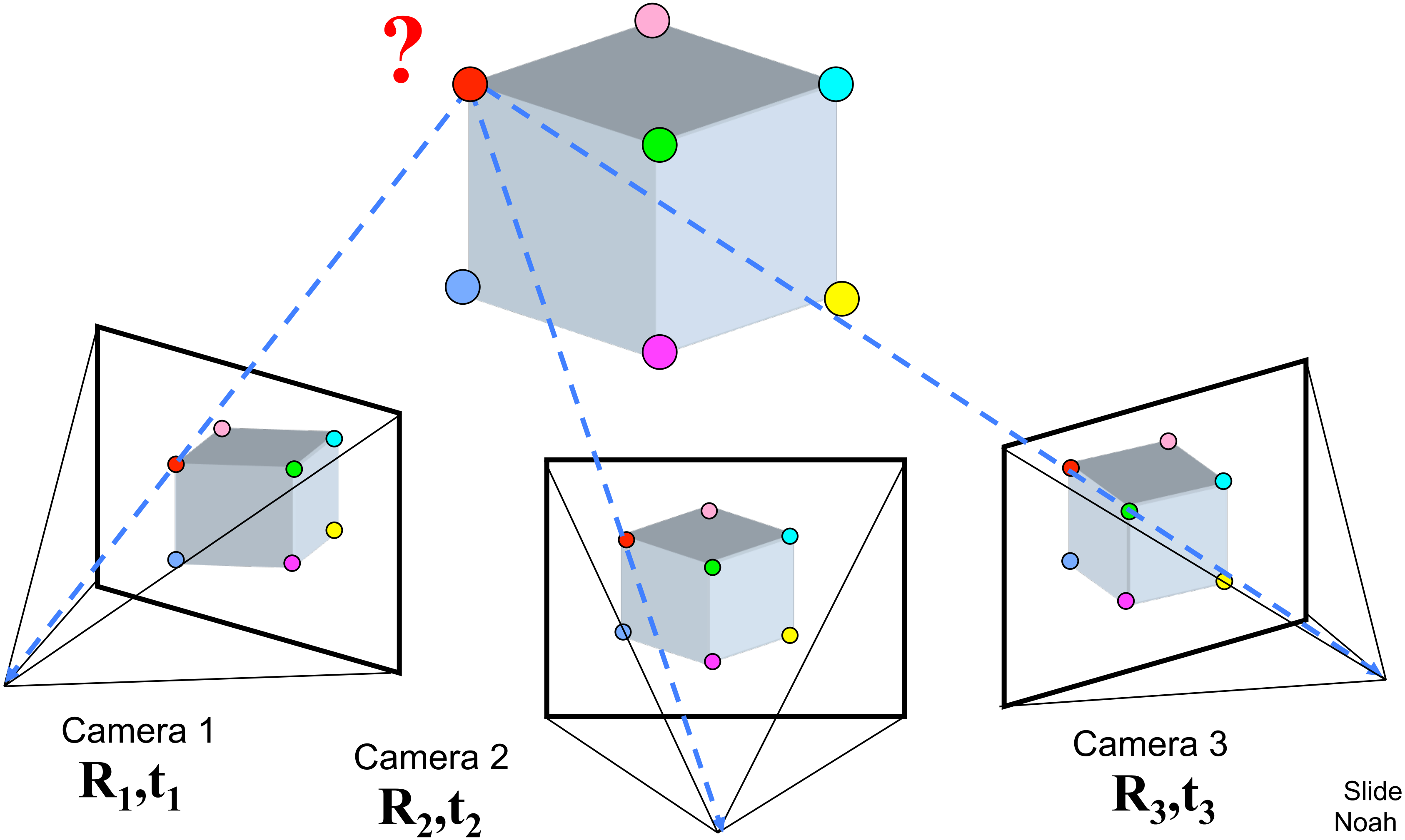
Source: S. Lazebnik

Multi-view geometry



Multi-view geometry problems

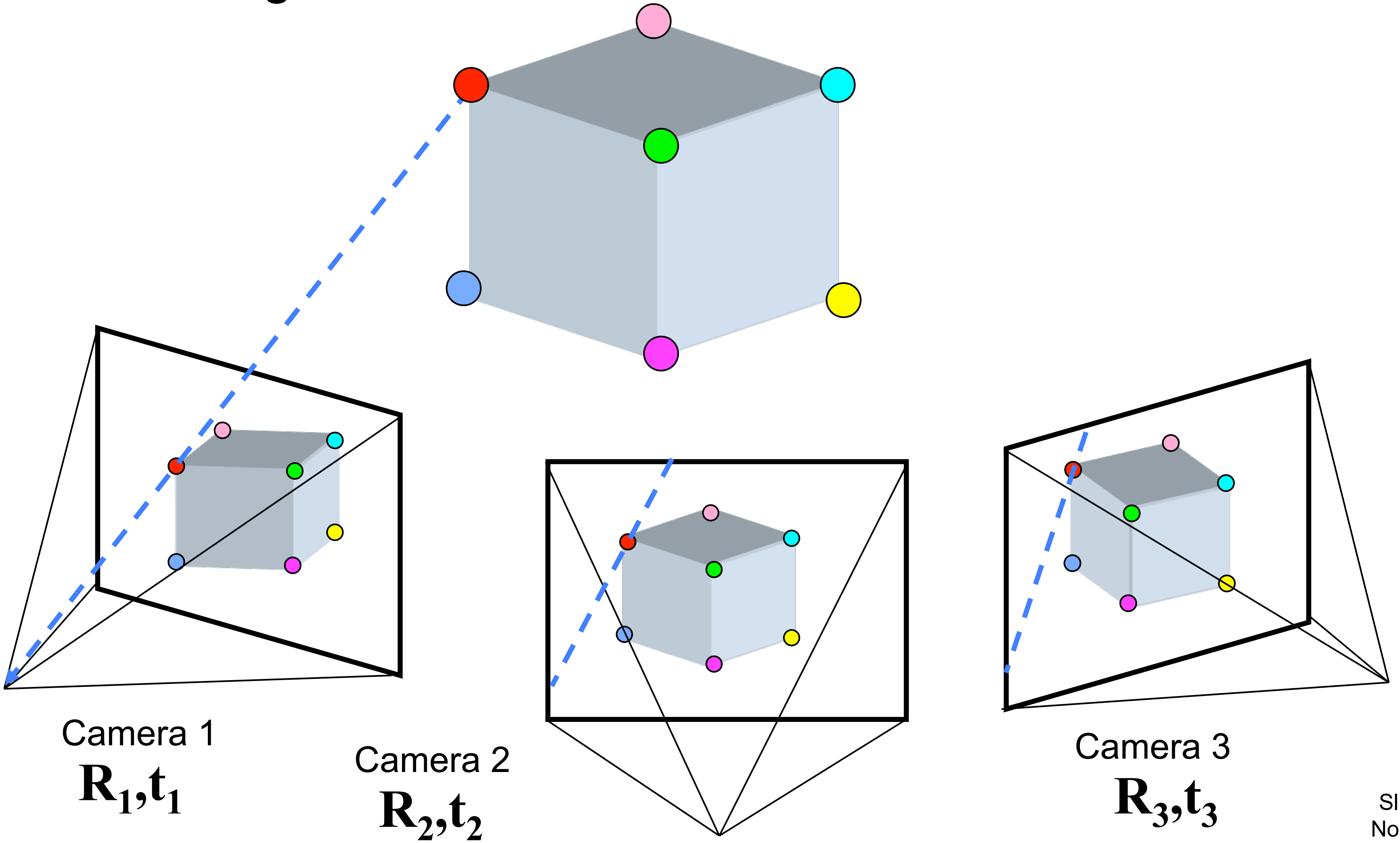
- **Structure:** Given projections of the same 3D point in two or more images, compute the 3D coordinates of that point


?
Camera 1
$\mathbf{R}_1,\mathbf{t}_1$
Camera 2
$\mathbf{R}_2,\mathbf{t}_2$
Camera 3
$\mathbf{R}_3,\mathbf{t}_3$

Slide credit: Noah Snavely

Multi-view geometry problems

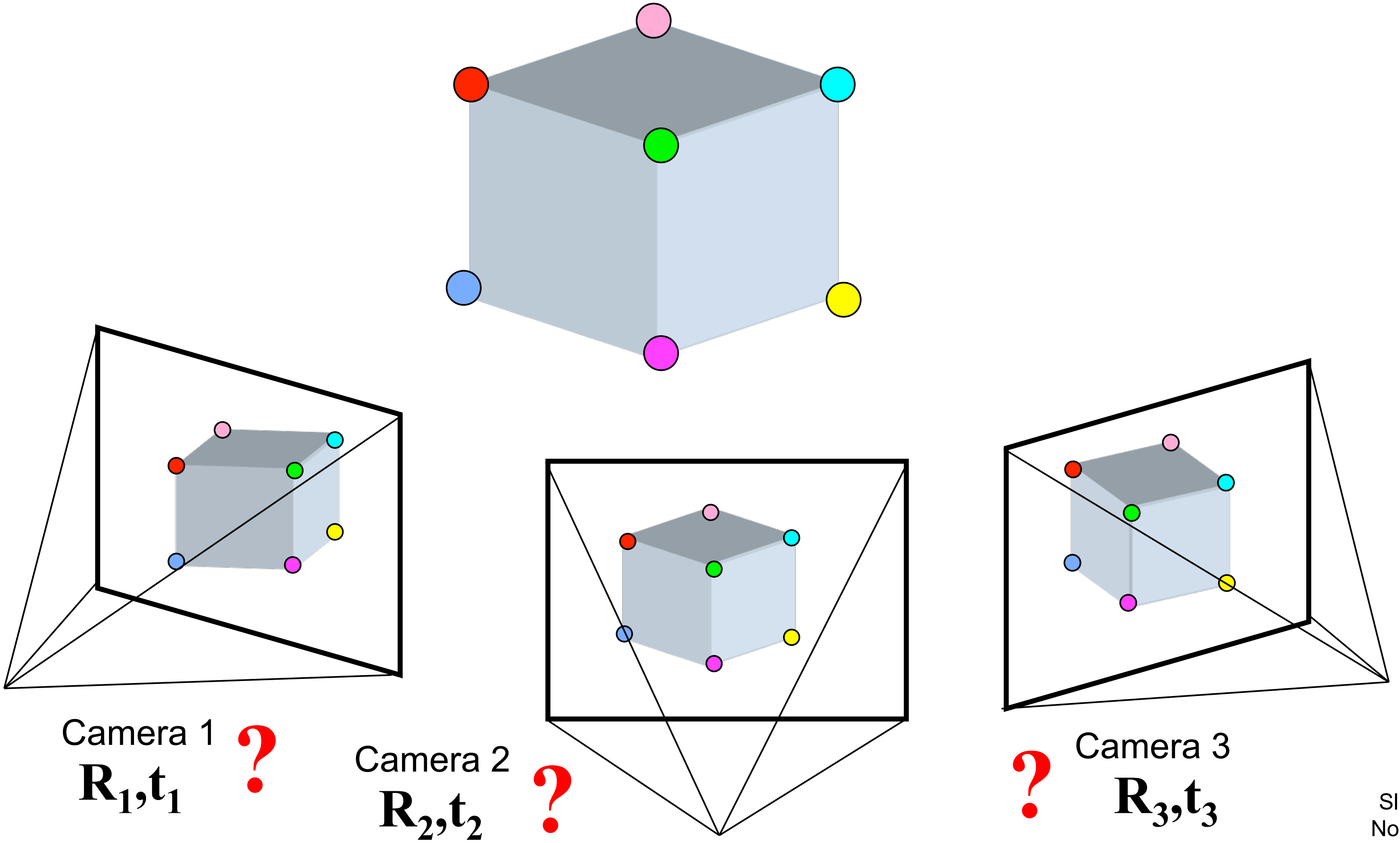
- **Stereo correspondence:** Given a point in one of the images, where could its corresponding points be in the other images?



Slide credit: Noah Snavely

Multi-view geometry problems

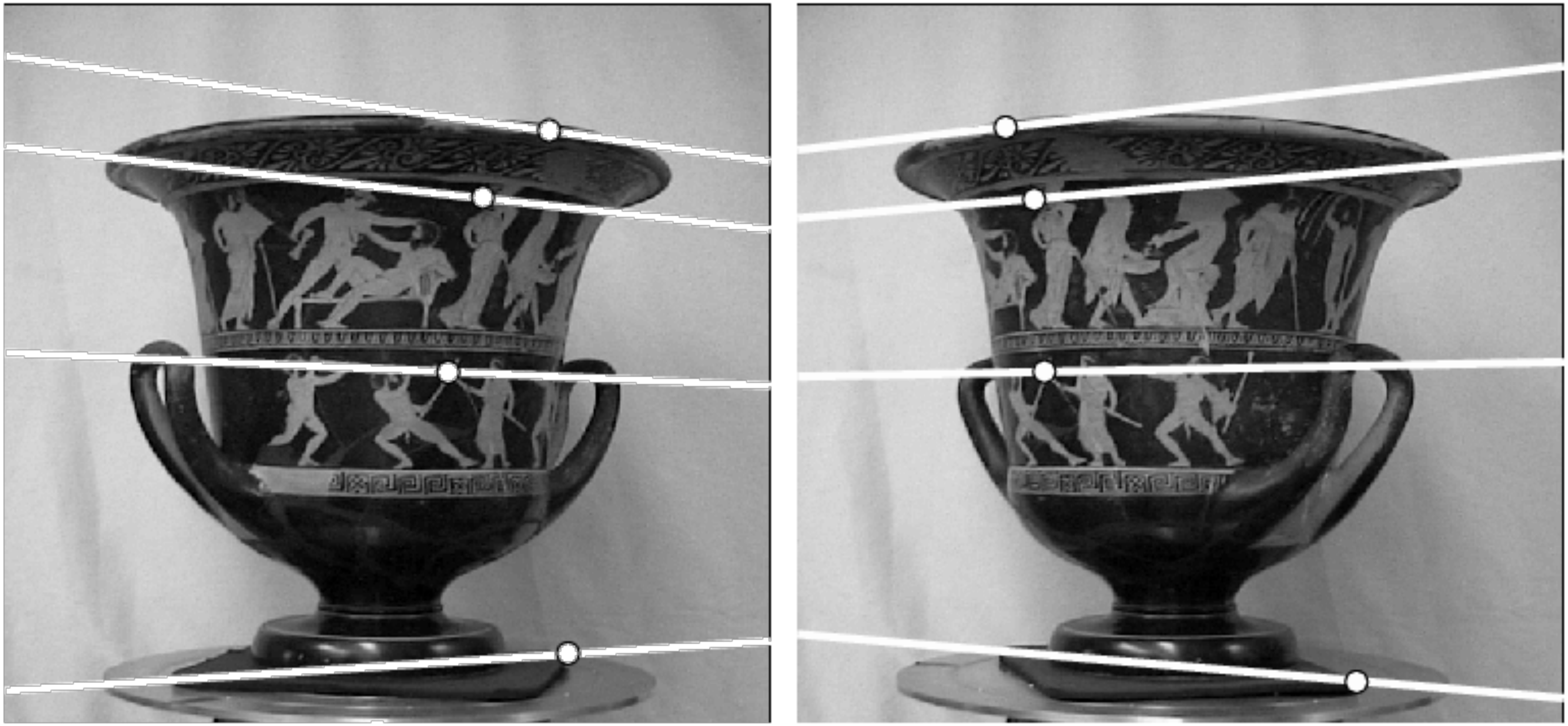
- **Motion:** Given a set of corresponding points in two or more images, compute the camera parameters



Slide credit: Noah Snavely

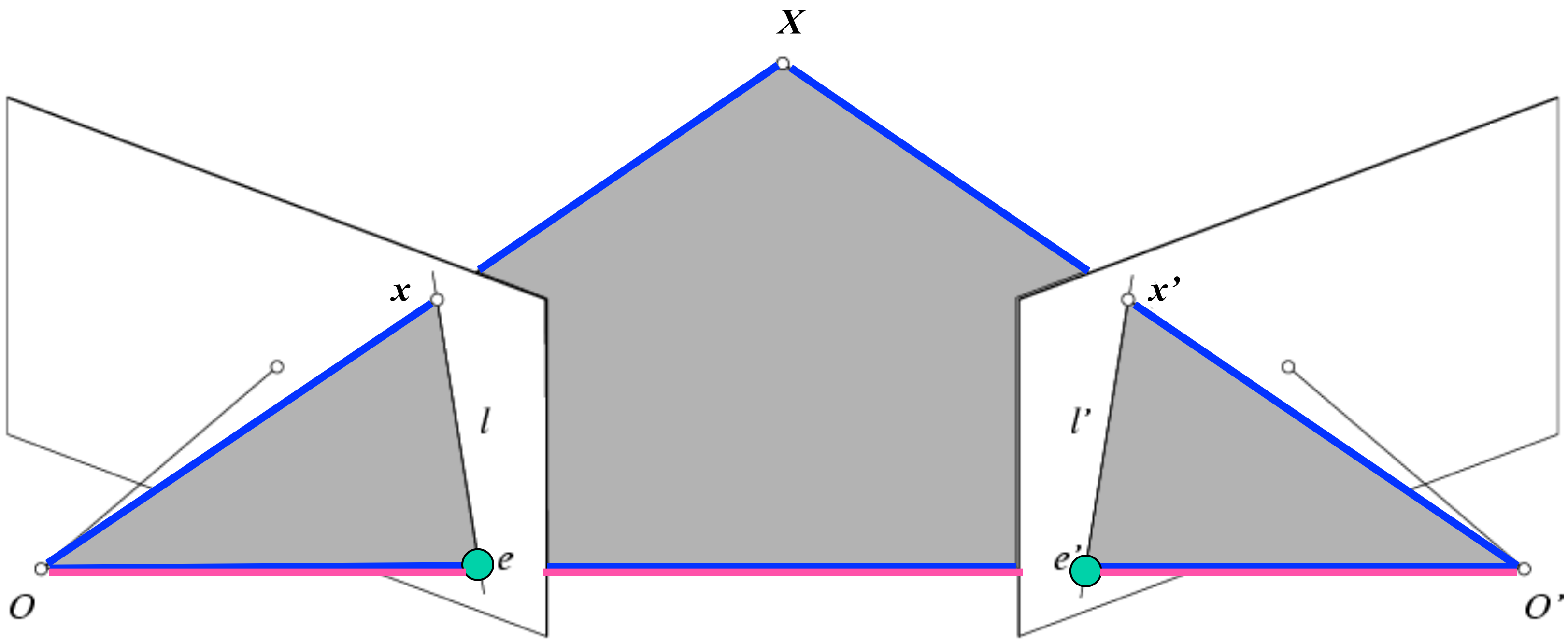
Source: S. Lazebnik

Two-view geometry



Source: S. Lazebnik

Epipolar geometry

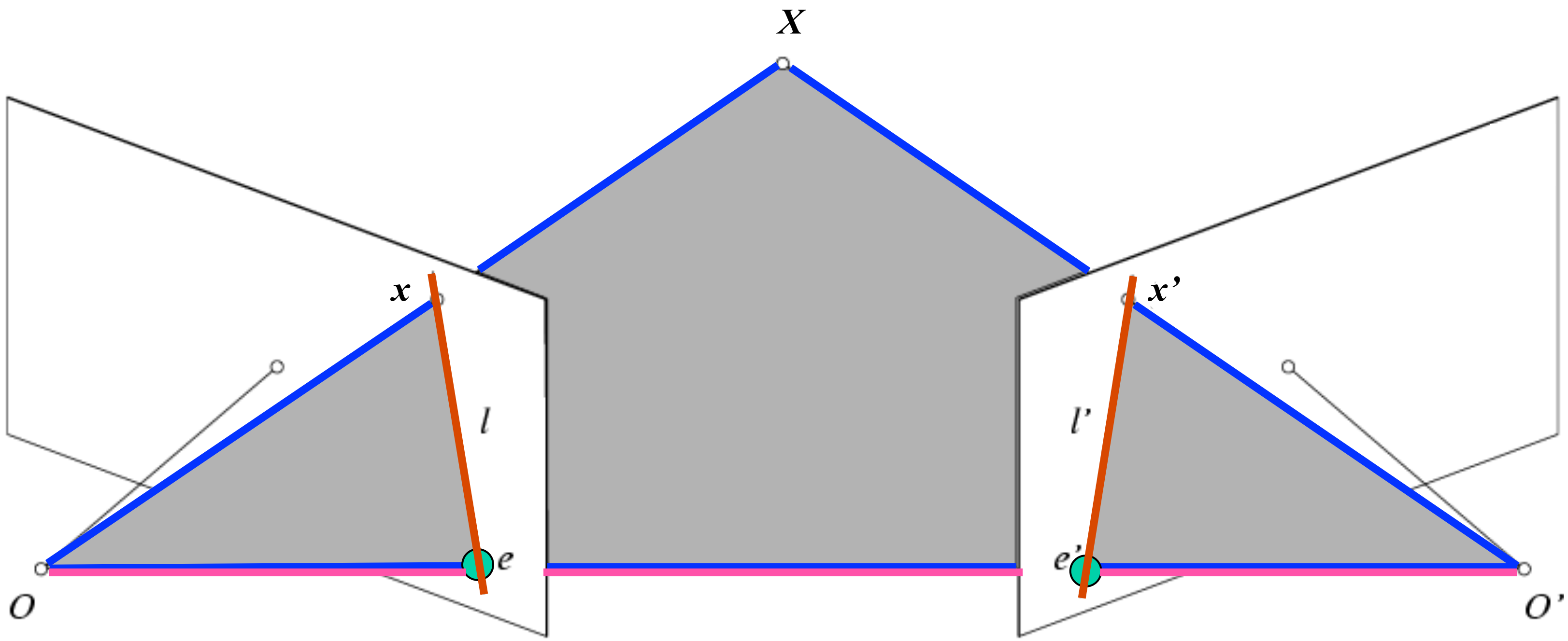


- **Baseline** – line connecting the two camera centers
- **Epipolar Plane** – plane containing baseline (1D family)
- **Epipoles**

= intersections of baseline with image planes
= projections of the other camera center
= vanishing points of the motion direction

Source: S. Lazebnik

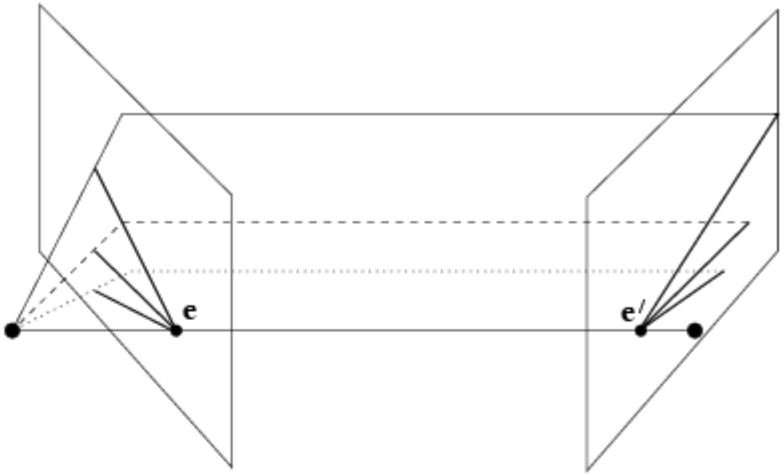
Epipolar geometry

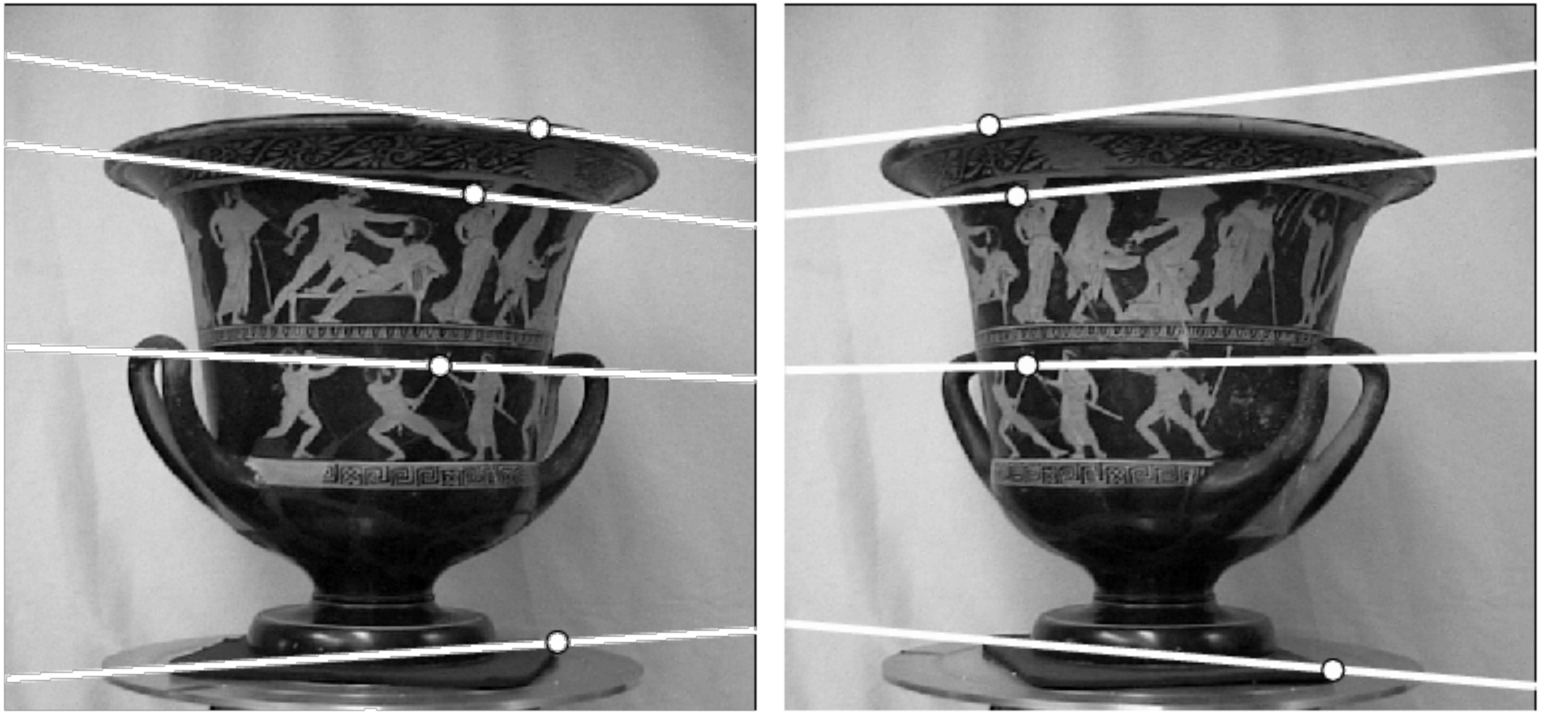


- **Baseline** – line connecting the two camera centers
- **Epipolar Plane** – plane containing baseline (1D family)
- **Epipoles**
= intersections of baseline with image planes
= projections of the other camera center
= vanishing points of the motion direction
- **Epipolar Lines** - intersections of epipolar plane with image planes (always come in corresponding pairs)

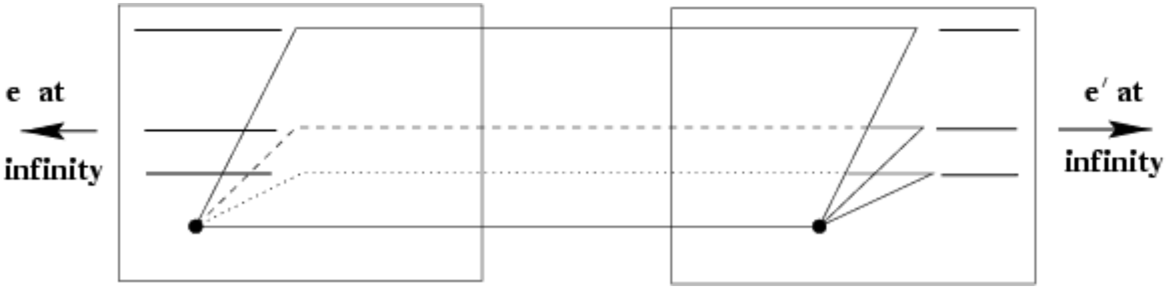
Source: S. Lazebnik

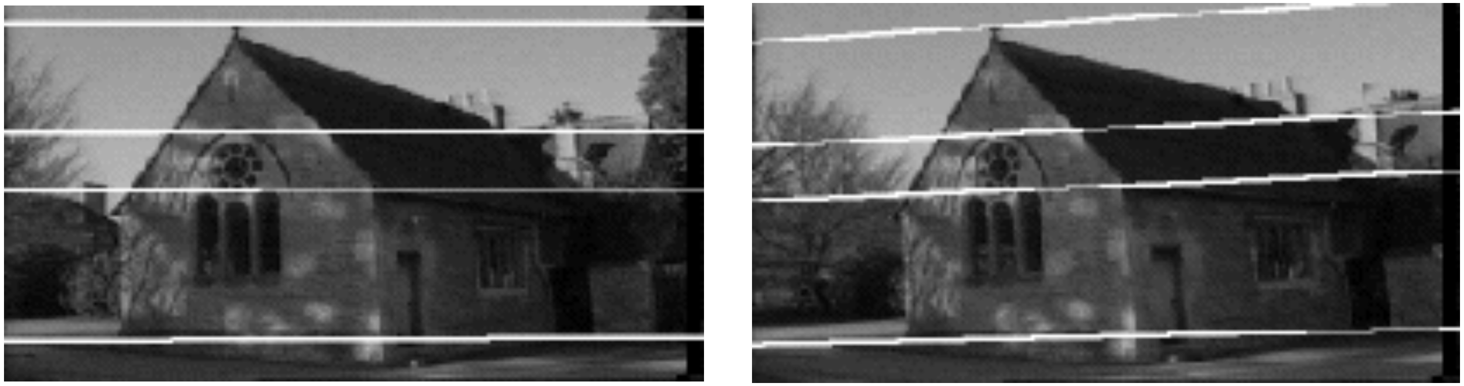
Example: Converging cameras





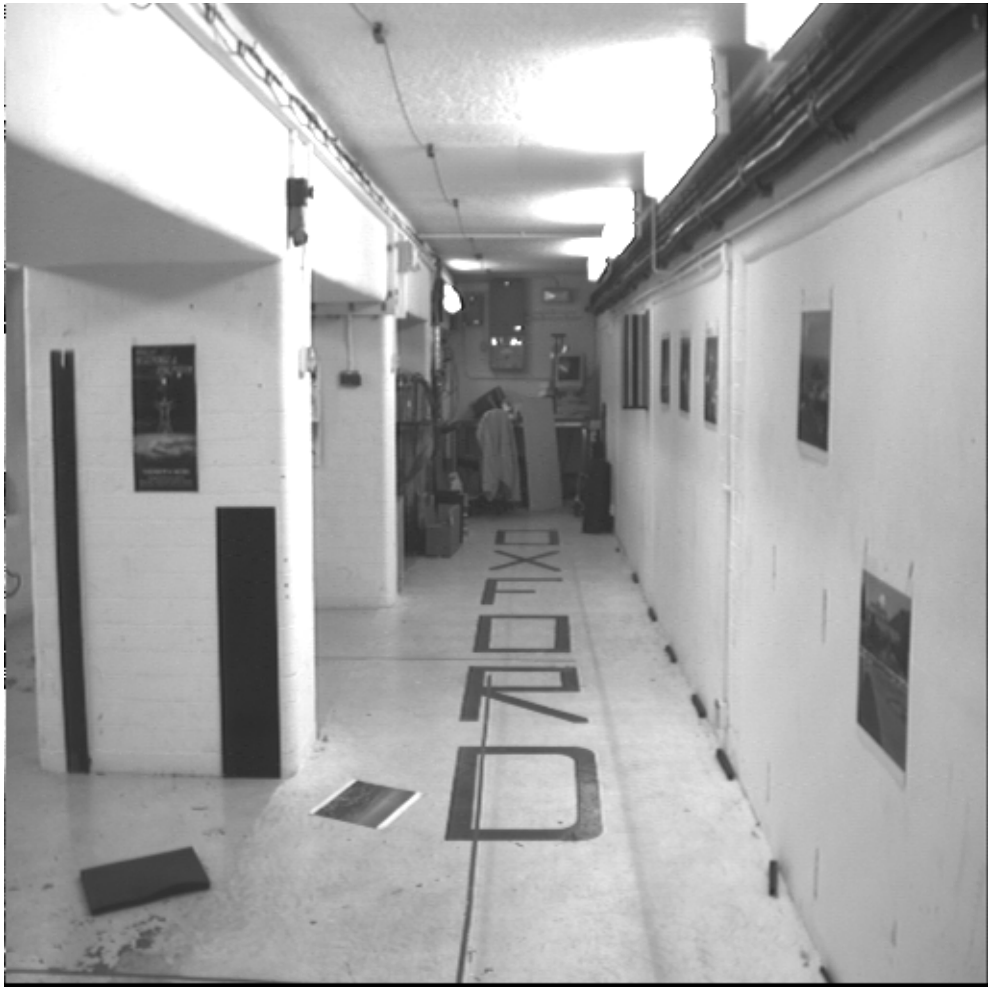
Example: Motion parallel to image plane





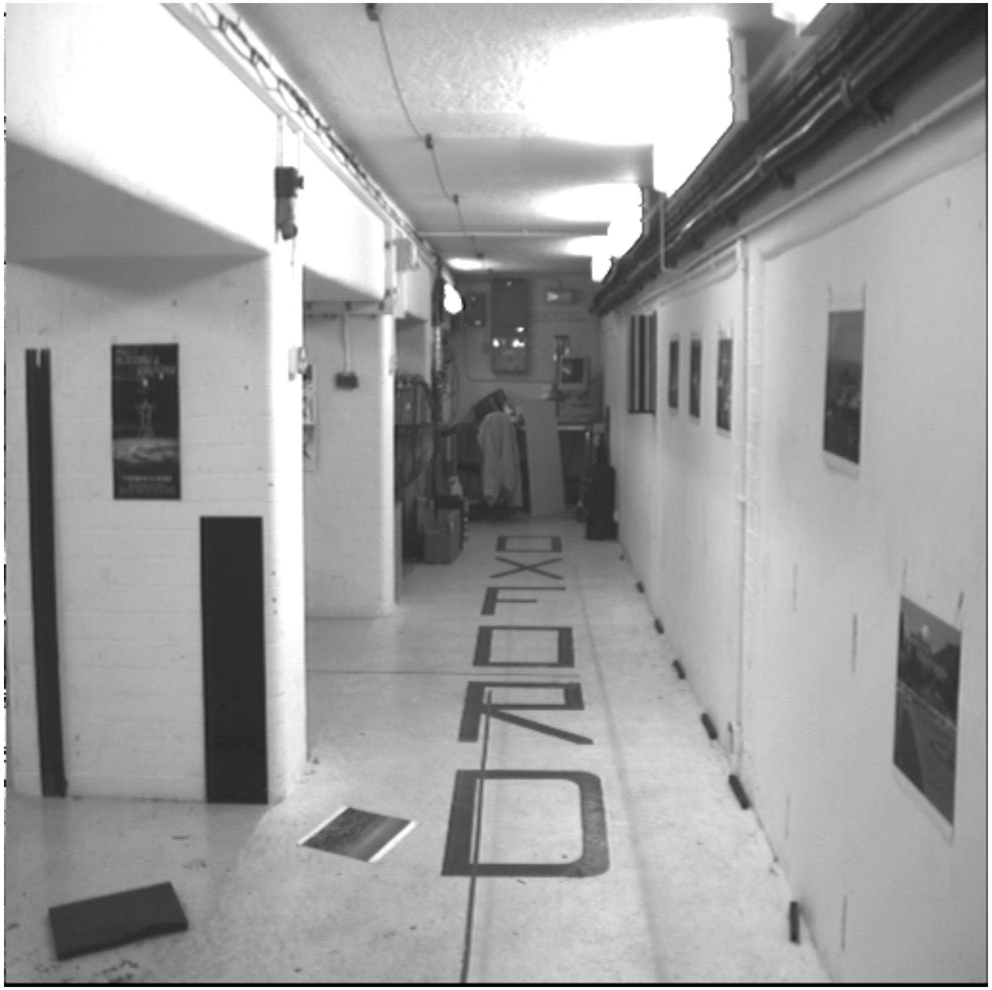
Source: S. Lazebnik

Example: Motion perpendicular to image plane



Source: S. Lazebnik

Example: Motion perpendicular to image plane

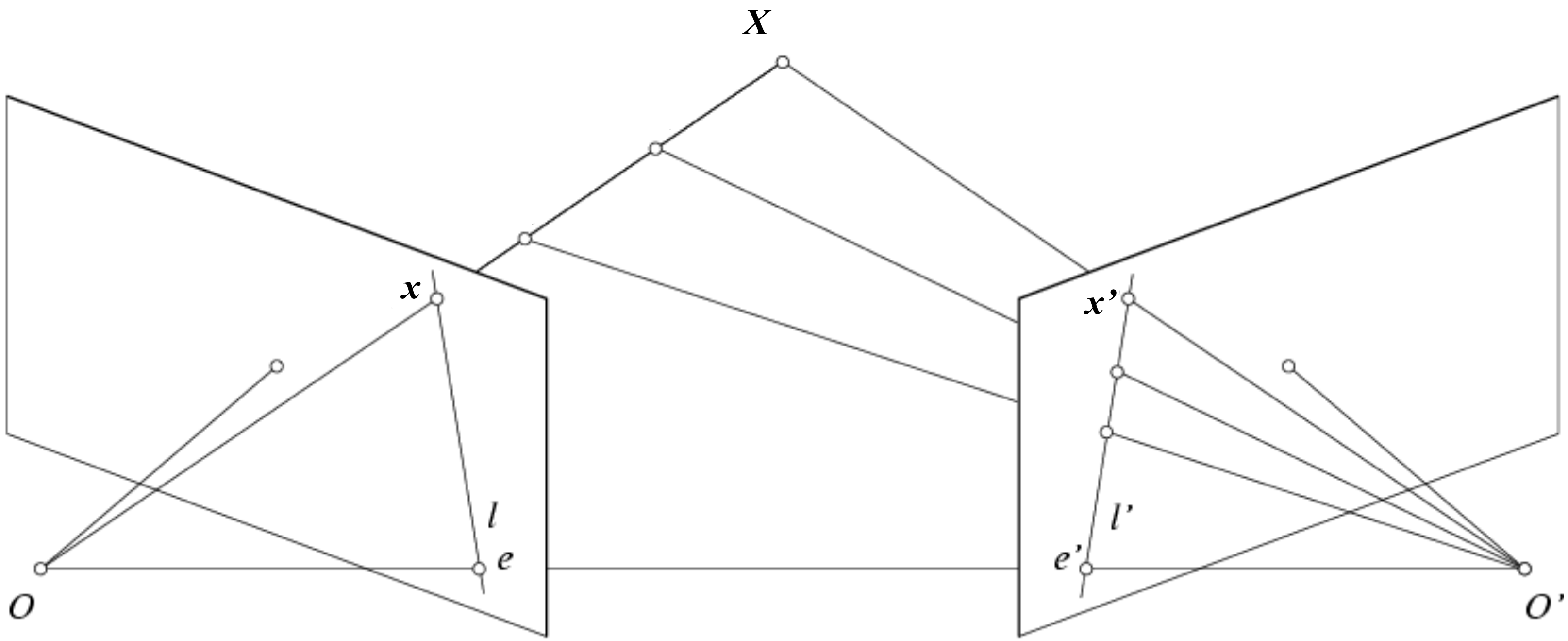


- Points move along lines radiating from the epipole: “focus of expansion”
- Epipole is the principal point

Source: S. Lazebnik

Source: S. Lazebnik

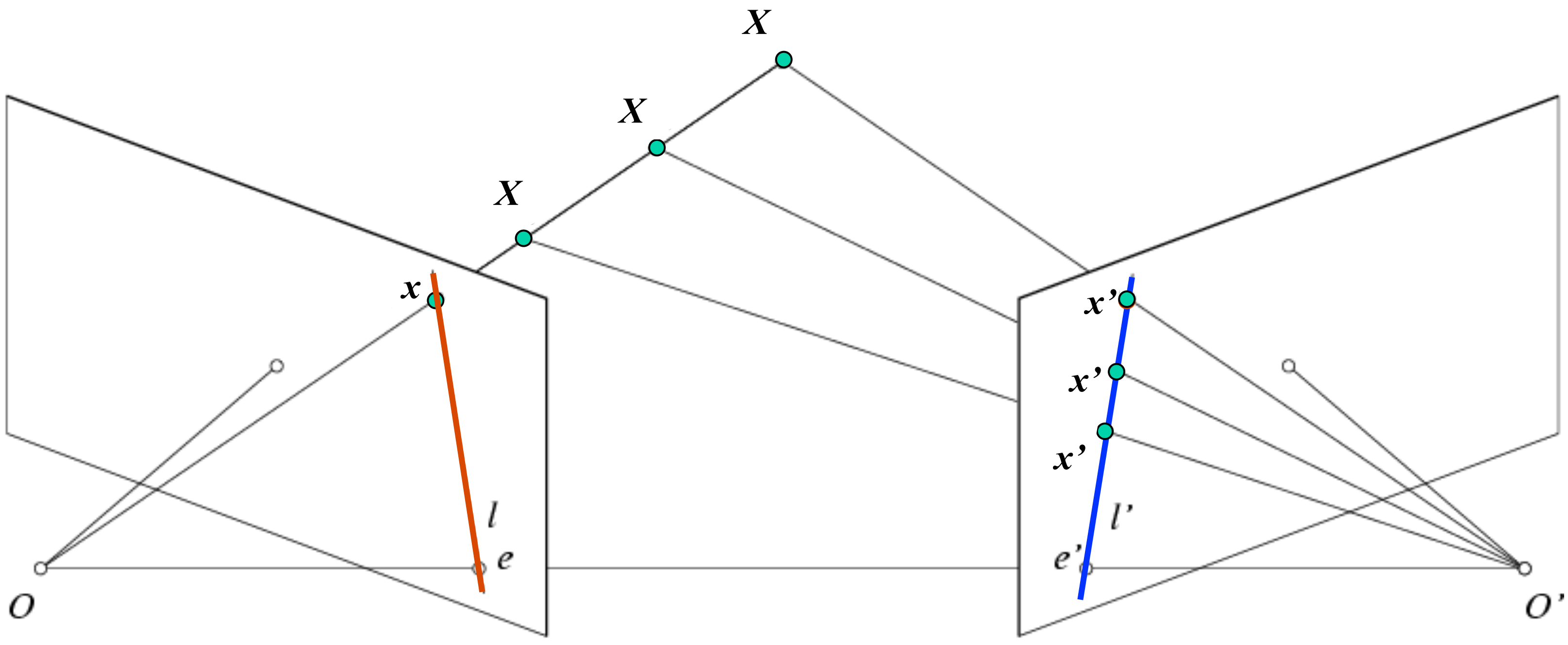
Epipolar constraint



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

Source: S. Lazebnik

Epipolar constraint

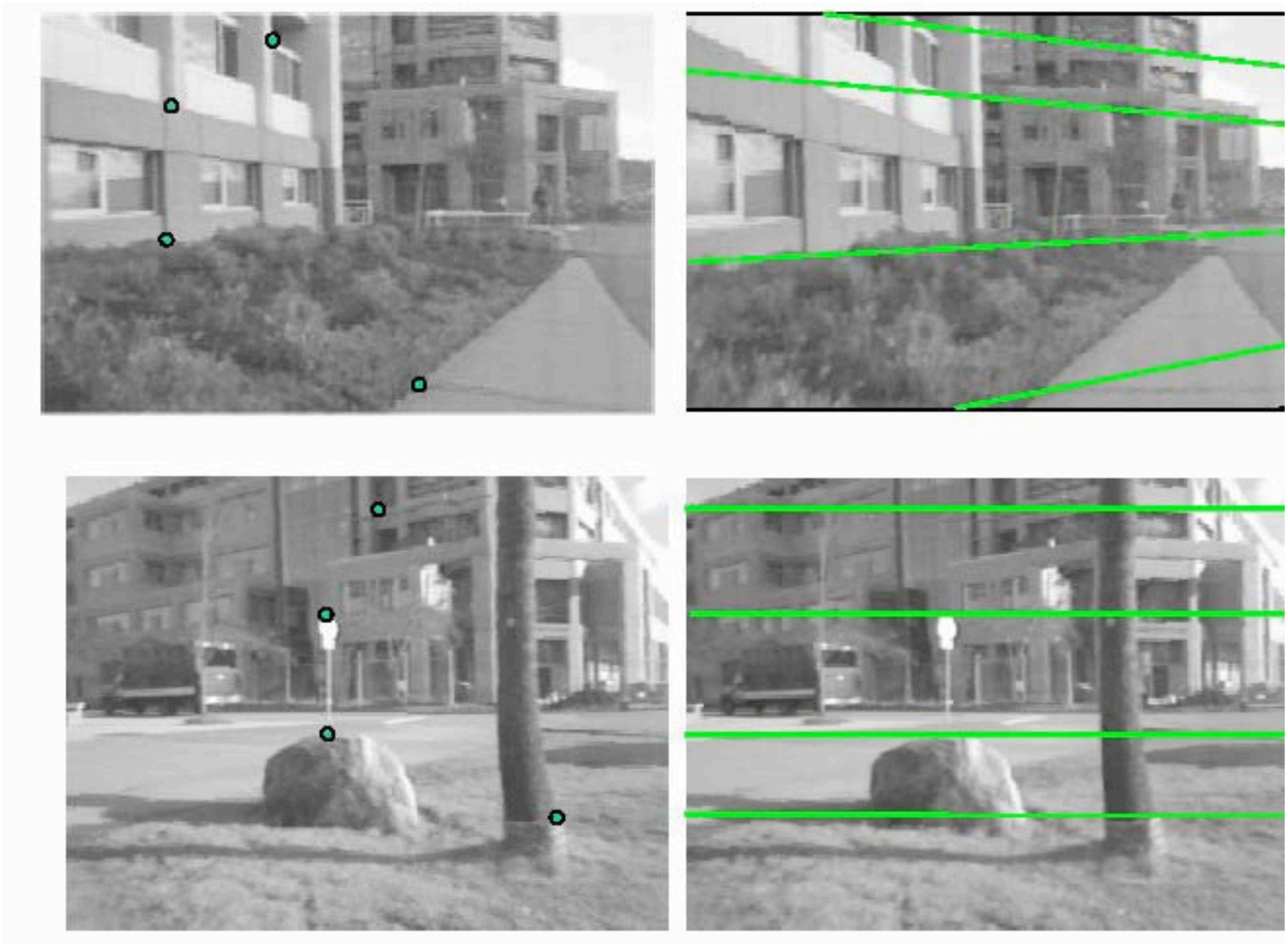


- Potential matches for ***x*** have to lie on the corresponding epipolar line ***l'***.

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

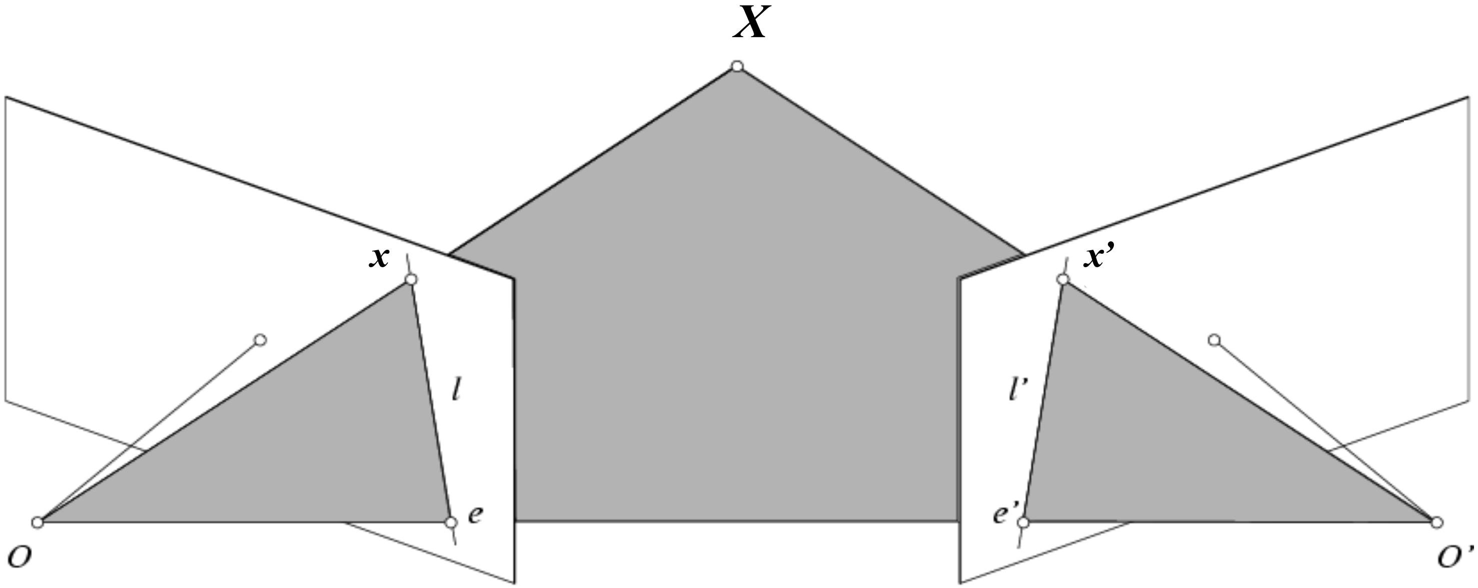
Source: S. Lazebnik

Epipolar constraint example



Source: S. Lazebnik

Epipolar constraint: Calibrated case

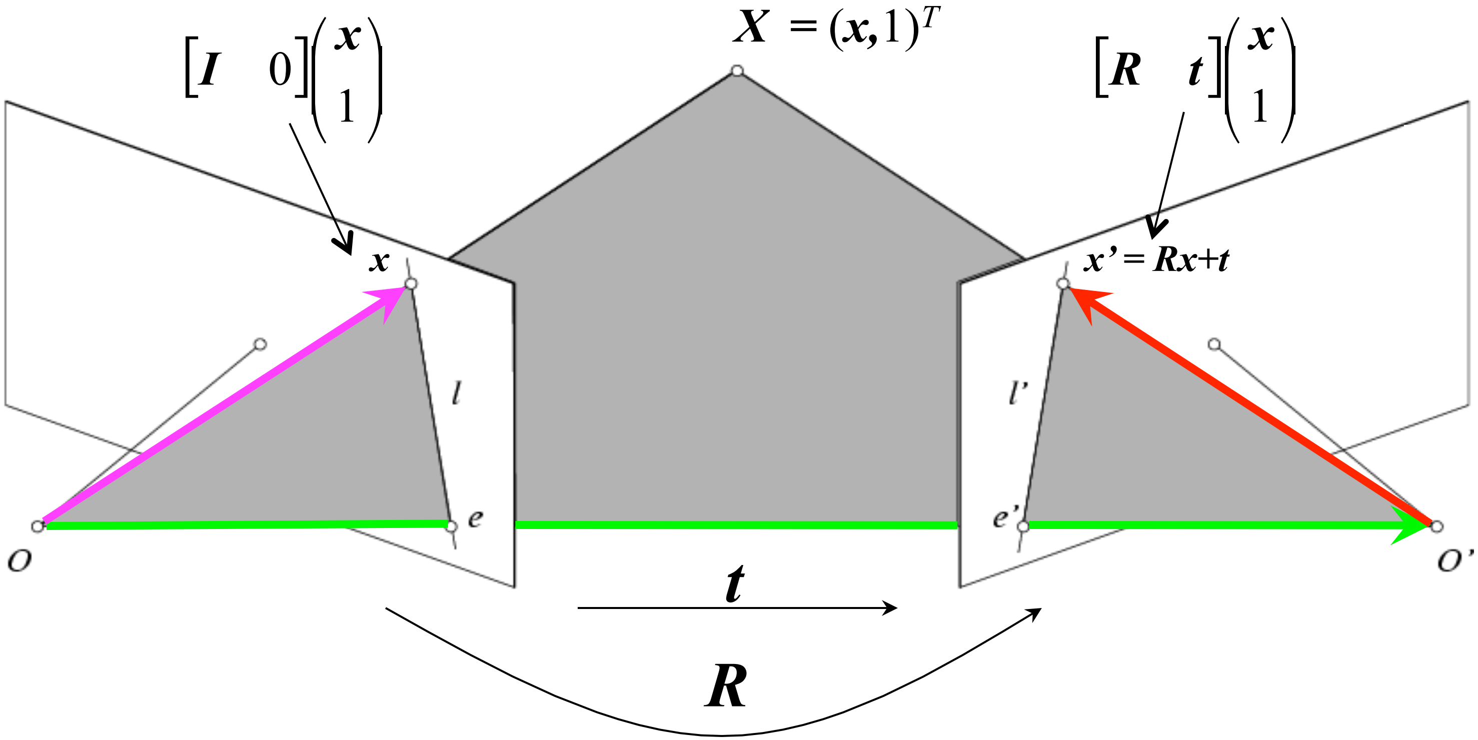


- Intrinsic and extrinsic parameters of the cameras are known, world coordinate system is set to that of the first camera
- Then the projection matrices are given by $\boldsymbol{K}[\boldsymbol{I} \mid \boldsymbol{0}]$ and $\boldsymbol{K}'[\boldsymbol{R} \mid \boldsymbol{t}]$
- We can multiply the projection matrices (and the image points) by the inverse of the calibration matrices to get *normalized* image coordinates:

$$\boldsymbol{x}_{\text{norm}} = \boldsymbol{K}^{-1}\boldsymbol{x}_{\text{pixel}} = [\boldsymbol{I}\ \ 0]\boldsymbol{X}, \qquad \boldsymbol{x}'_{\text{norm}} = \boldsymbol{K}'^{-1}\boldsymbol{x}'_{\text{pixel}} = [\boldsymbol{R}\ \ \boldsymbol{t}]\boldsymbol{X}$$

Source: S. Lazebnik

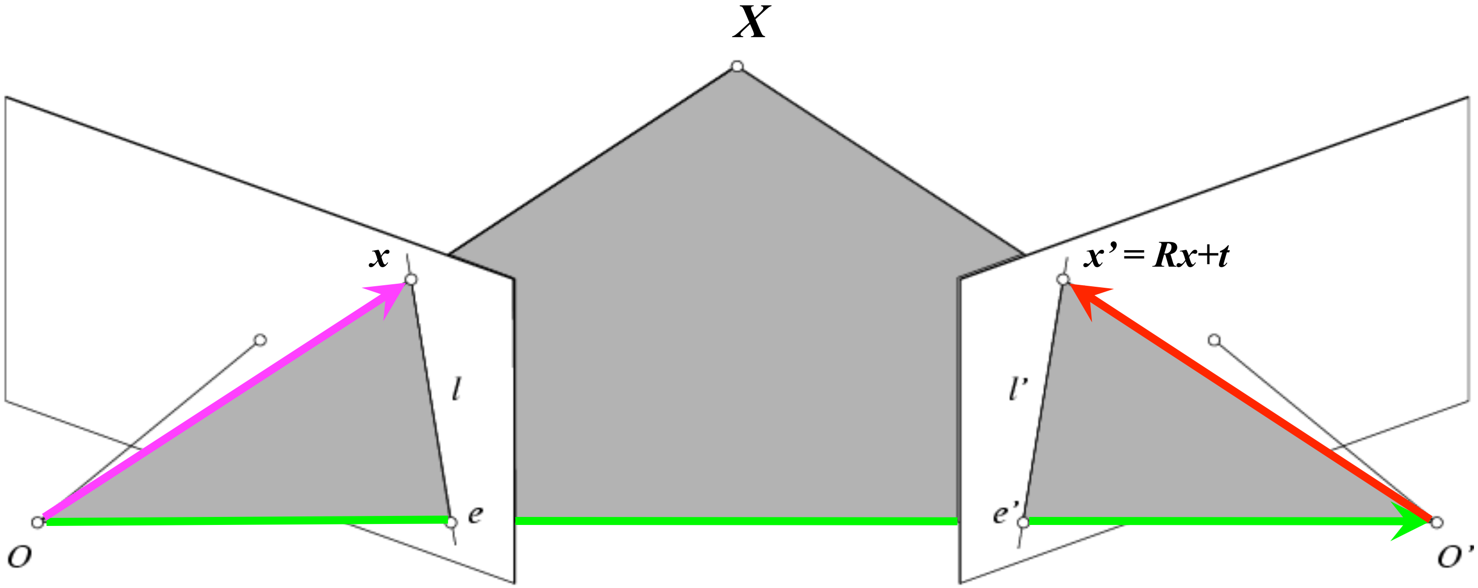
Epipolar constraint: Calibrated case


$\begin{bmatrix} I & 0 \end{bmatrix} \begin{pmatrix} \boldsymbol{x} \\ 1 \end{pmatrix}$
$\boldsymbol{X} = (\boldsymbol{x}, 1)^T$
$\begin{bmatrix} \boldsymbol{R} & \boldsymbol{t} \end{bmatrix} \begin{pmatrix} \boldsymbol{x} \\ 1 \end{pmatrix}$
x
x' = Rx+t
l
l'
e
e'
O
O'
t
R

The vectors *Rx*, *t*, and *x'* are coplanar

Source: S. Lazebnik

Epipolar constraint: Calibrated case



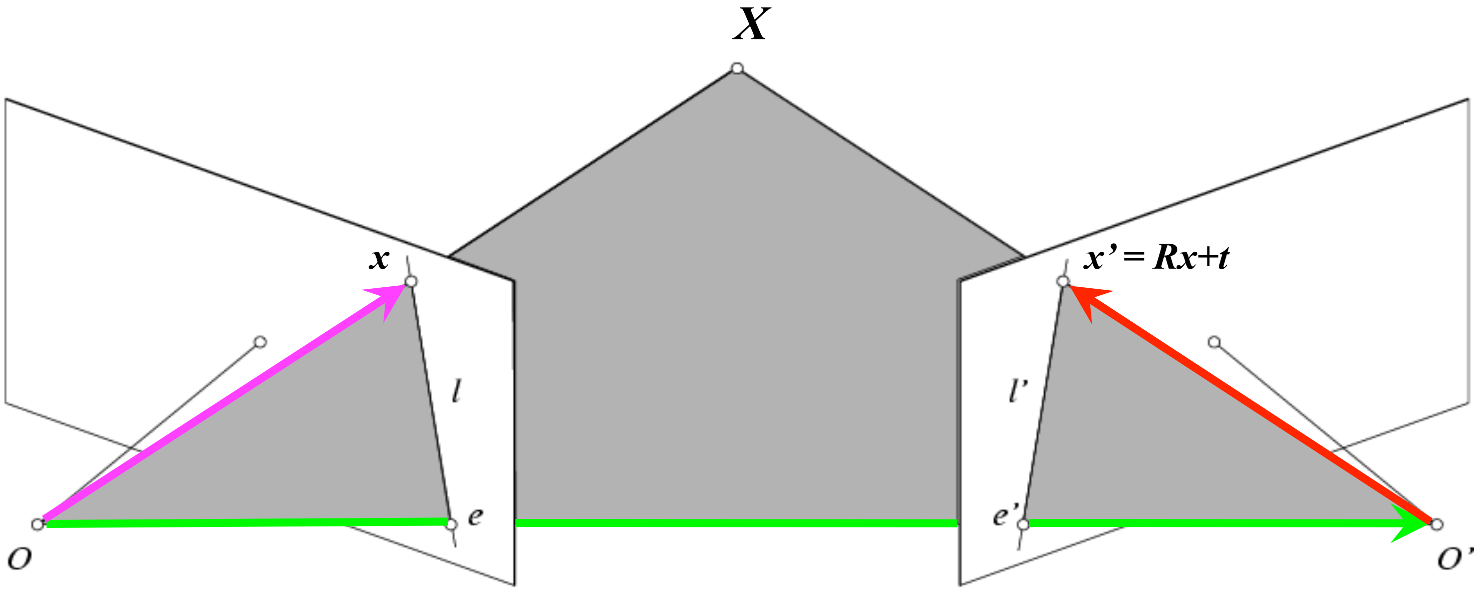
$$\boldsymbol{x}' \cdot [\boldsymbol{t} \times (\boldsymbol{R}\boldsymbol{x})] = 0 \Rightarrow \boldsymbol{x}'^{T}[\boldsymbol{t}_{\times}]\boldsymbol{R}\boldsymbol{x} = 0$$

Recall: $\mathbf{a} \times \mathbf{b} = \begin{bmatrix} 0 & -a_z & a_y \\ a_z & 0 & -a_x \\ -a_y & a_x & 0 \end{bmatrix} \begin{bmatrix} b_x \\ b_y \\ b_z \end{bmatrix} = [\mathbf{a}_{\times}]\mathbf{b}$

The vectors ***Rx***, ***t***, and ***x’*** are coplanar

Source: S. Lazebnik

Epipolar constraint: Calibrated case



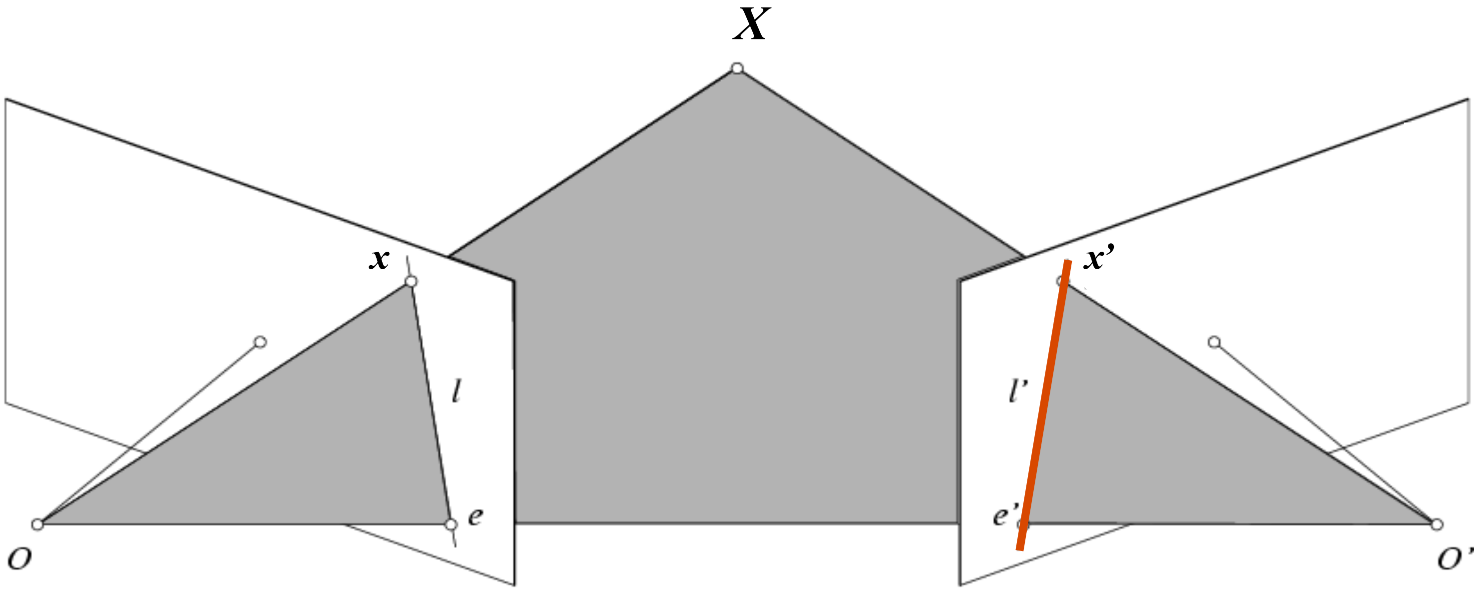
$$\boldsymbol{x}'\cdot[\boldsymbol{t}\times(\boldsymbol{R}\boldsymbol{x})]=0 \Rightarrow \boldsymbol{x}'^{T}[\boldsymbol{t}_{\times}]\boldsymbol{R}\boldsymbol{x}=0 \Rightarrow \boldsymbol{x}'^{T}E\boldsymbol{x}=0$$

Essential Matrix
(Longuet-Higgins, 1981)

The vectors ***Rx***, ***t***, and ***x'*** are coplanar

Source: S. Lazebnik

Epipolar constraint: Calibrated case



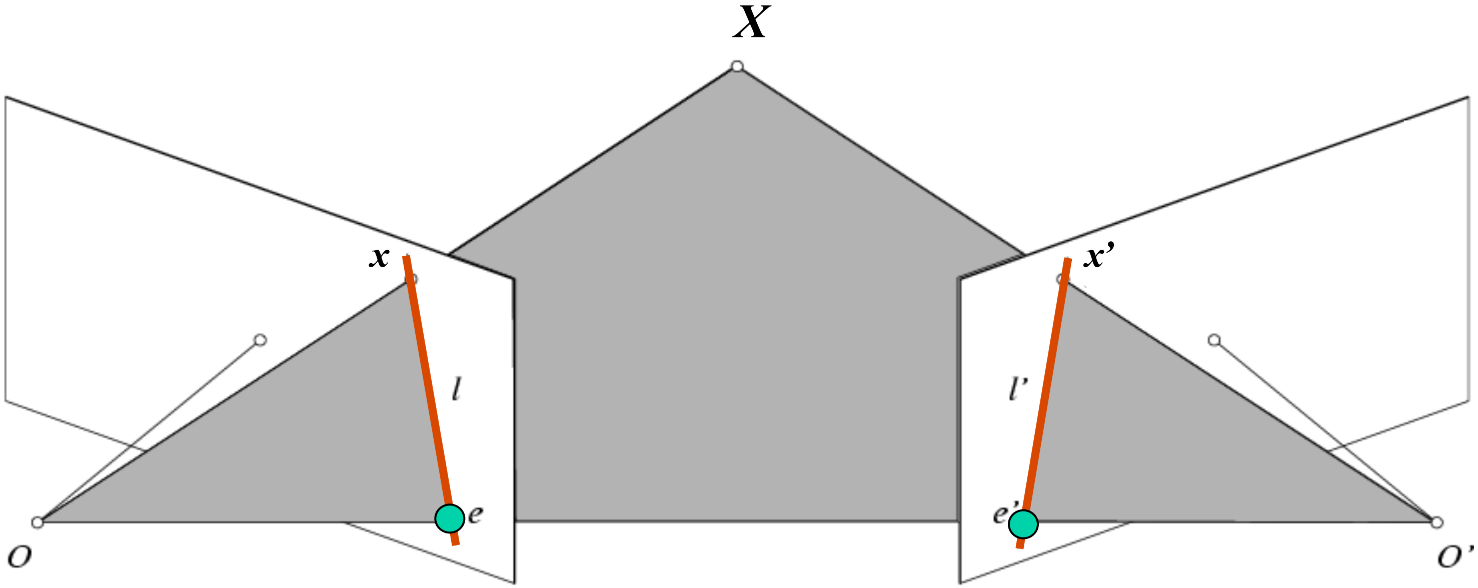
$$\boxed{x'^{T} E x = 0}$$

- ***E x*** is the epipolar line associated with ***x*** (***l'*** = ***E x***)
  - Recall: a line is given by $ax + by + c = 0$ or

$$\mathbf{l}^T \mathbf{x} = 0 \quad \text{where} \quad \mathbf{l} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Source: S. Lazebnik

Epipolar constraint: Calibrated case

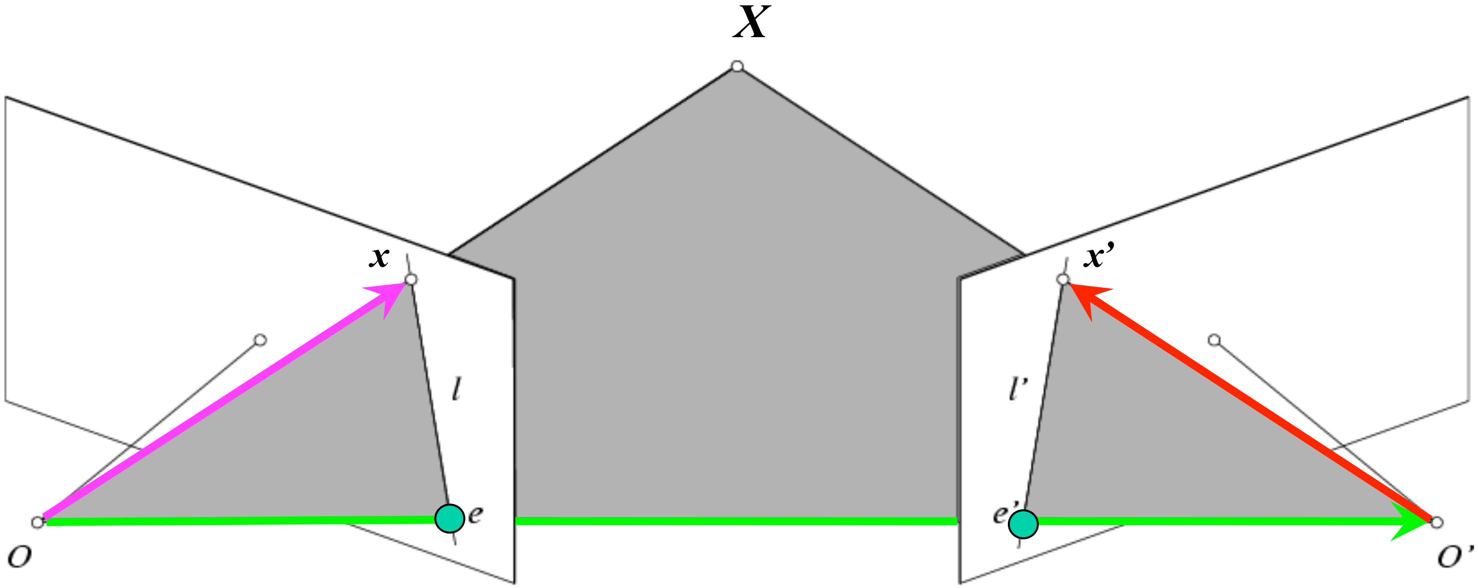


$$\boxed{x'^{T} E x = 0}$$

- ***E x*** is the epipolar line associated with ***x*** (***l'*** = ***E x***)
- ***E**^T**x'*** is the epipolar line associated with ***x'*** (***l*** = ***E**^T**x'***)
- ***E*** **e** = 0 and ***E**^T**e'** = 0*
- ***E*** is singular (rank two)
- ***E*** has five degrees of freedom

Source: S. Lazebnik

Epipolar constraint: Uncalibrated case

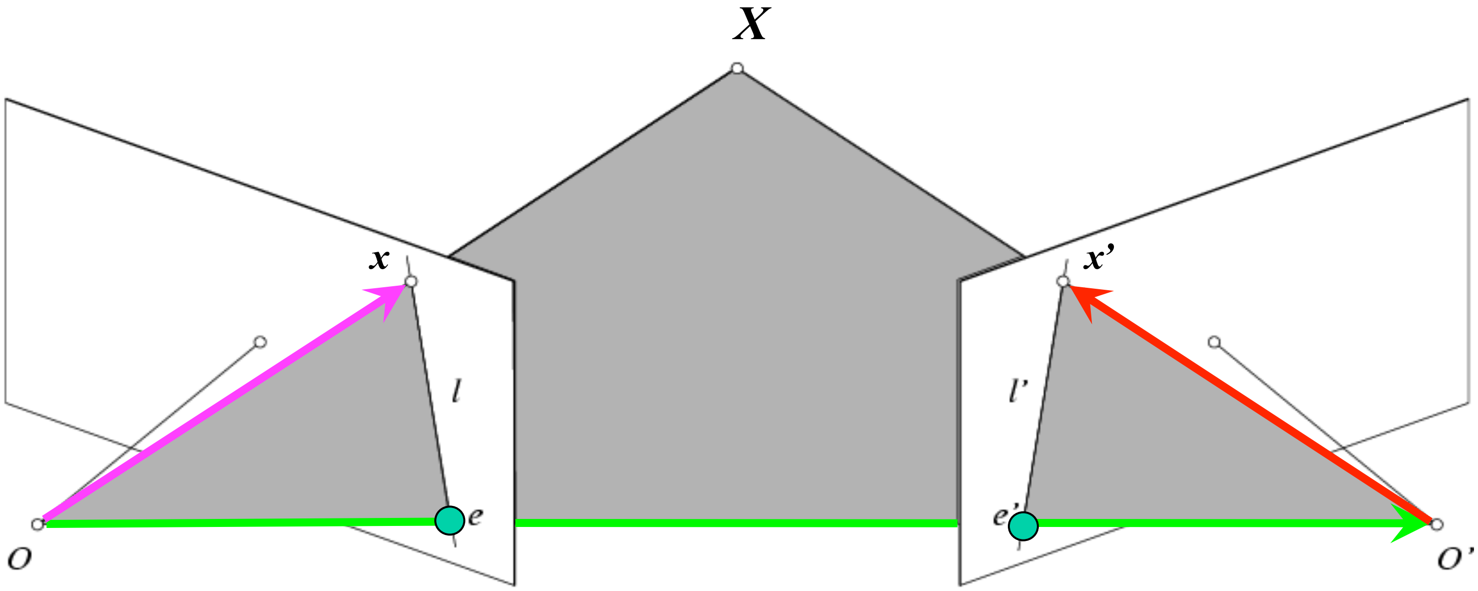


- The calibration matrices ***K*** and ***K'*** of the two cameras are unknown
- We can write the epipolar constraint in terms of *unknown* normalized coordinates:

$$\hat{\boldsymbol{x}}'^T \boldsymbol{E} \hat{\boldsymbol{x}} = 0 \qquad \hat{\boldsymbol{x}} = \boldsymbol{K}^{-1}\boldsymbol{x}, \quad \hat{\boldsymbol{x}}' = \boldsymbol{K}'^{-1}\boldsymbol{x}'$$

Source: S. Lazebnik

Epipolar constraint: Uncalibrated case



$$\hat{\boldsymbol{x}}'^{T} \boldsymbol{E} \hat{\boldsymbol{x}} = 0 \Rightarrow \boldsymbol{x}'^{T} \boldsymbol{F} \boldsymbol{x} = 0 \quad \text{with} \quad \boldsymbol{F} = \boldsymbol{K}'^{-T} \boldsymbol{E} \boldsymbol{K}^{-1}$$

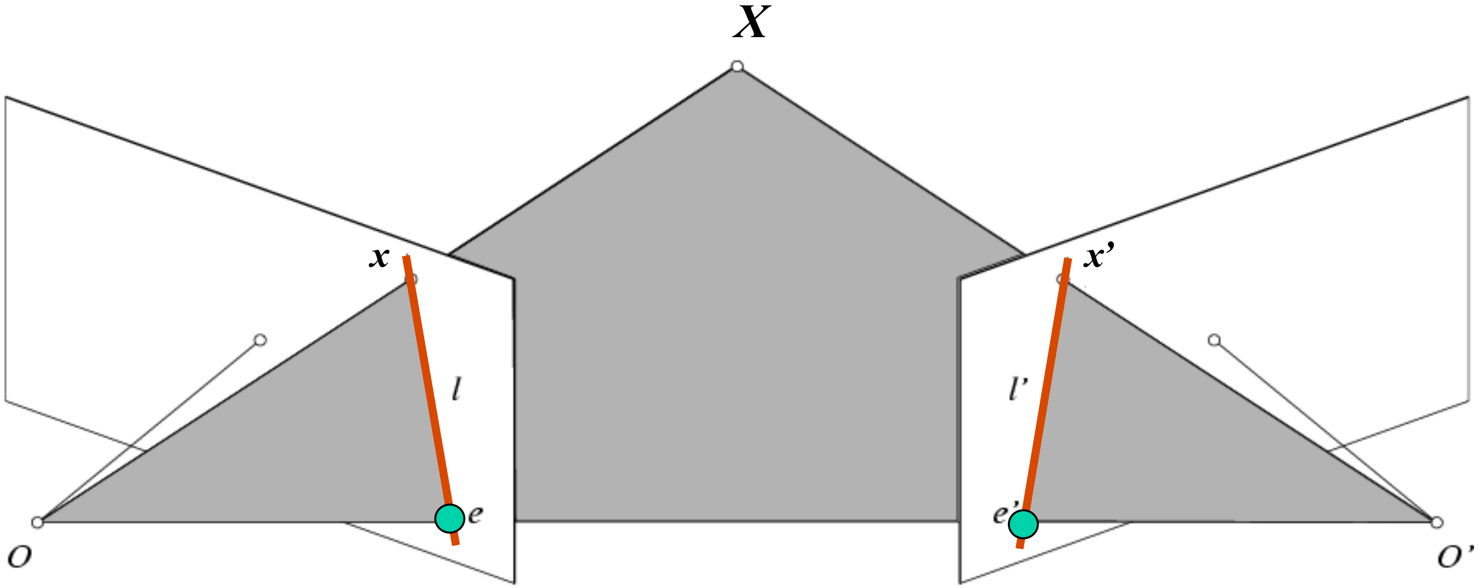
$$\hat{\boldsymbol{x}} = \boldsymbol{K}^{-1} \boldsymbol{x}$$

$$\hat{\boldsymbol{x}}' = \boldsymbol{K}'^{-1} \boldsymbol{x}'$$

Fundamental Matrix
(Faugeras and Luong, 1992)

Source: S. Lazebnik

Epipolar constraint: Uncalibrated case

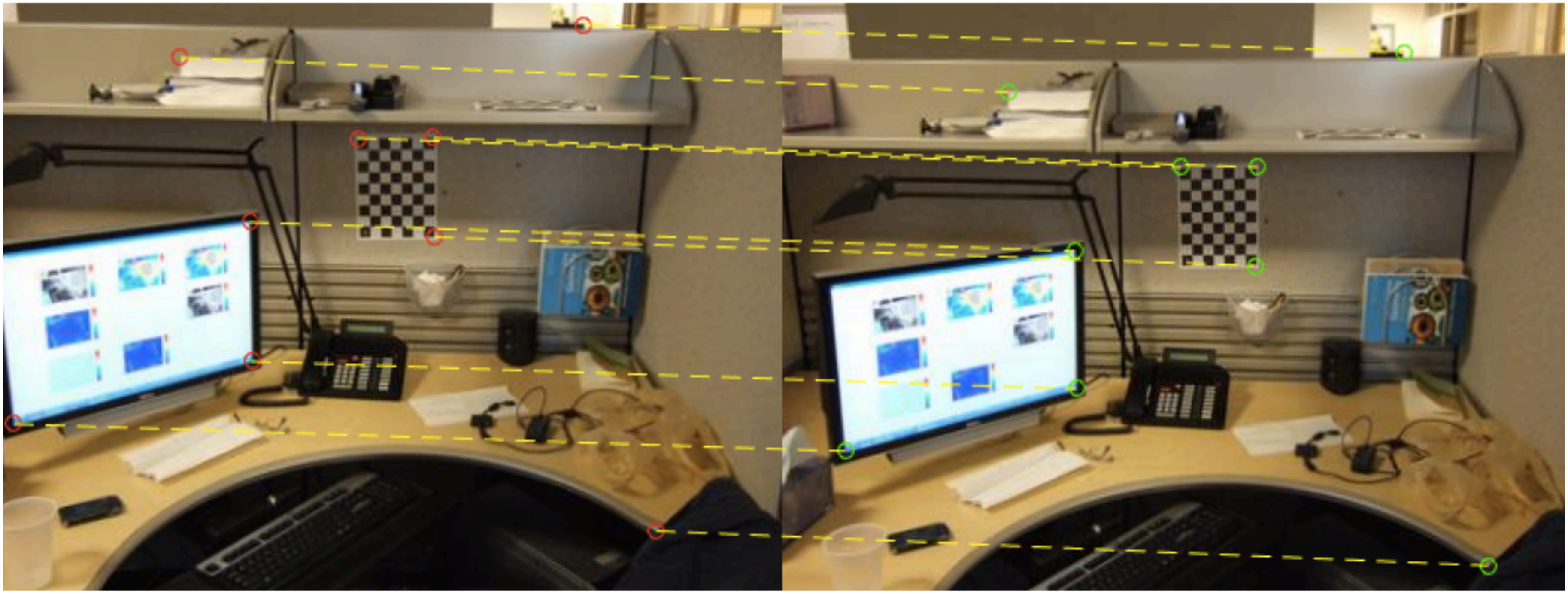


$$\hat{x}'^{T} E \hat{x} = 0 \quad \Rightarrow \quad x'^{T} F x = 0 \quad \text{with} \quad F = K'^{-T} E K^{-1}$$

- ***F x*** is the epipolar line associated with ***x*** (***l'*** = ***F x***)
- ***F***^{T}***x'*** is the epipolar line associated with ***x'*** (***l*** = ***F***^{T}***x'***)
- ***F e*** = 0 and ***F***^{T}***e'*** = 0
- ***F*** is singular (rank two)
- ***F*** has *seven* degrees of freedom

Source: S. Lazebnik

Estimating the fundamental matrix



Source: S. Lazebnik

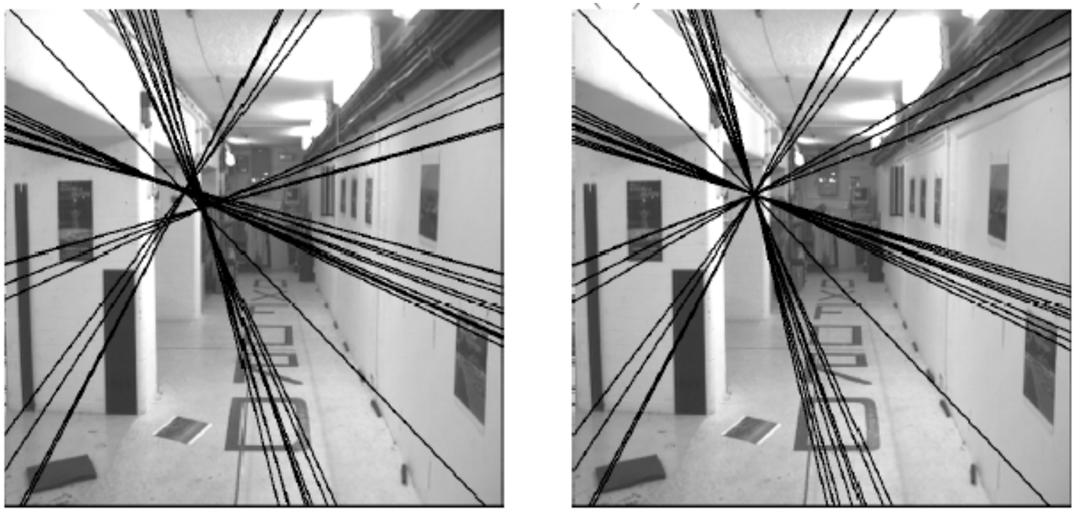
The eight-point algorithm

$$\boldsymbol{x} = (u, v, 1)^T, \quad \boldsymbol{x}' = (u', v', 1)$$

$$\begin{bmatrix} u' & v' & 1 \end{bmatrix} \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = 0 \quad \Rightarrow \quad \begin{bmatrix} u'u & u'v & u' & v'u & v'v & v' & u & v & 1 \end{bmatrix} \begin{bmatrix} f_{11} \\ f_{12} \\ f_{13} \\ f_{21} \\ f_{22} \\ f_{23} \\ f_{31} \\ f_{32} \\ f_{33} \end{bmatrix} = 0$$

Solve homogeneous linear system using eight or more matches

Enforce rank-2 constraint (take SVD of ***F*** and throw out the smallest singular value)



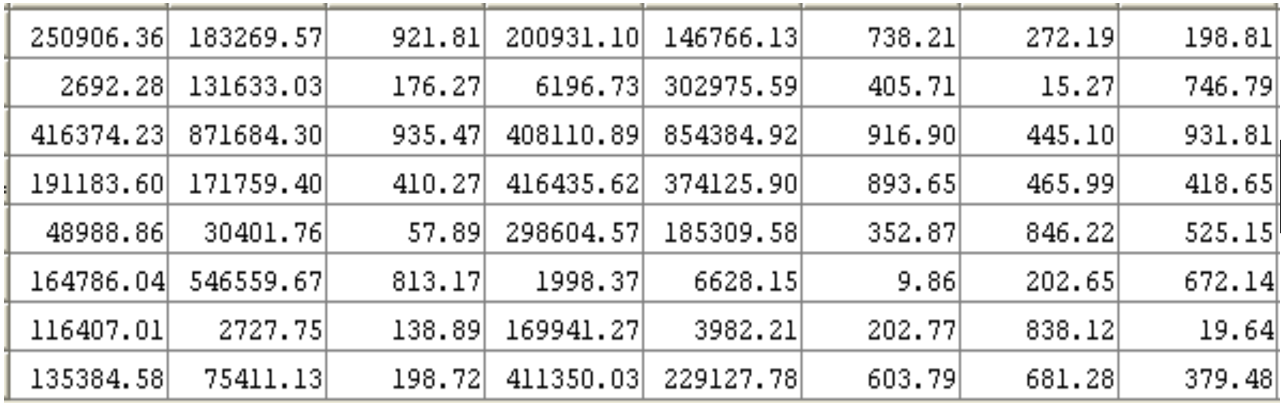
Source: S. Lazebnik

Problem with eight-point algorithm

$$\begin{bmatrix} u'u & u'v & u' & v'u & v'v & v' & u & v \end{bmatrix} \begin{bmatrix} f_{11} \\ f_{12} \\ f_{13} \\ f_{21} \\ f_{22} \\ f_{23} \\ f_{31} \\ f_{32} \end{bmatrix} = -1$$

Source: S. Lazebnik

Problem with eight-point algorithm



250906.36	183269.57	921.81	200931.10	146766.13	738.21	272.19	198.81
2692.28	131633.03	176.27	6196.73	302975.59	405.71	15.27	746.79
416374.23	871684.30	935.47	408110.89	854384.92	916.90	445.10	931.81
191183.60	171759.40	410.27	416435.62	374125.90	893.65	465.99	418.65
48988.86	30401.76	57.89	298604.57	185309.58	352.87	846.22	525.15
164786.04	546559.67	813.17	1998.37	6628.15	9.86	202.65	672.14
116407.01	2727.75	138.89	169941.27	3982.21	202.77	838.12	19.64
135384.58	75411.13	198.72	411350.03	229127.78	603.79	681.28	379.48

$$\begin{bmatrix} f_{11} \\ f_{12} \\ f_{13} \\ f_{21} \\ f_{22} \\ f_{23} \\ f_{31} \\ f_{32} \end{bmatrix} = -1$$

Poor numerical conditioning

Can be fixed by rescaling the data

Source: S. Lazebnik

The normalized eight-point algorithm

(Hartley, 1995)

- Center the image data at the origin, and scale it so the mean squared distance between the origin and the data points is 2 pixels
- Use the eight-point algorithm to compute ***F*** from the normalized points
- Enforce the rank-2 constraint (for example, take SVD of ***F*** and throw out the smallest singular value)
- Transform fundamental matrix back to original units: if ***T*** and ***T'*** are the normalizing transformations in the two images, than the fundamental matrix in original coordinates is $\boldsymbol{T'^T F T}$

Source: S. Lazebnik

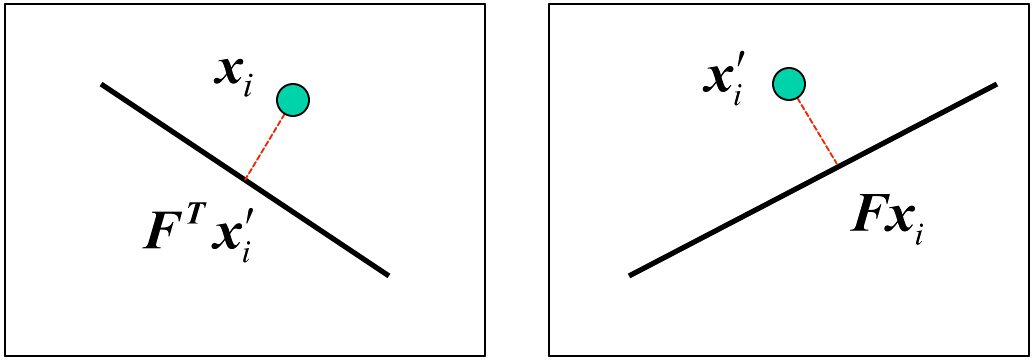
Nonlinear estimation

- Linear estimation minimizes the sum of squared *algebraic* distances between points ***x'*** $\boldsymbol{x}'_i$ and epipolar lines $\boldsymbol{F}\,\boldsymbol{x}_i$ (or points $\boldsymbol{x}_i$ and epipolar lines $\boldsymbol{F}^T\boldsymbol{x}'_i$):

$$\sum_{i=1}^{N} (\boldsymbol{x}_i'^T \boldsymbol{F} \boldsymbol{x}_i)^2$$

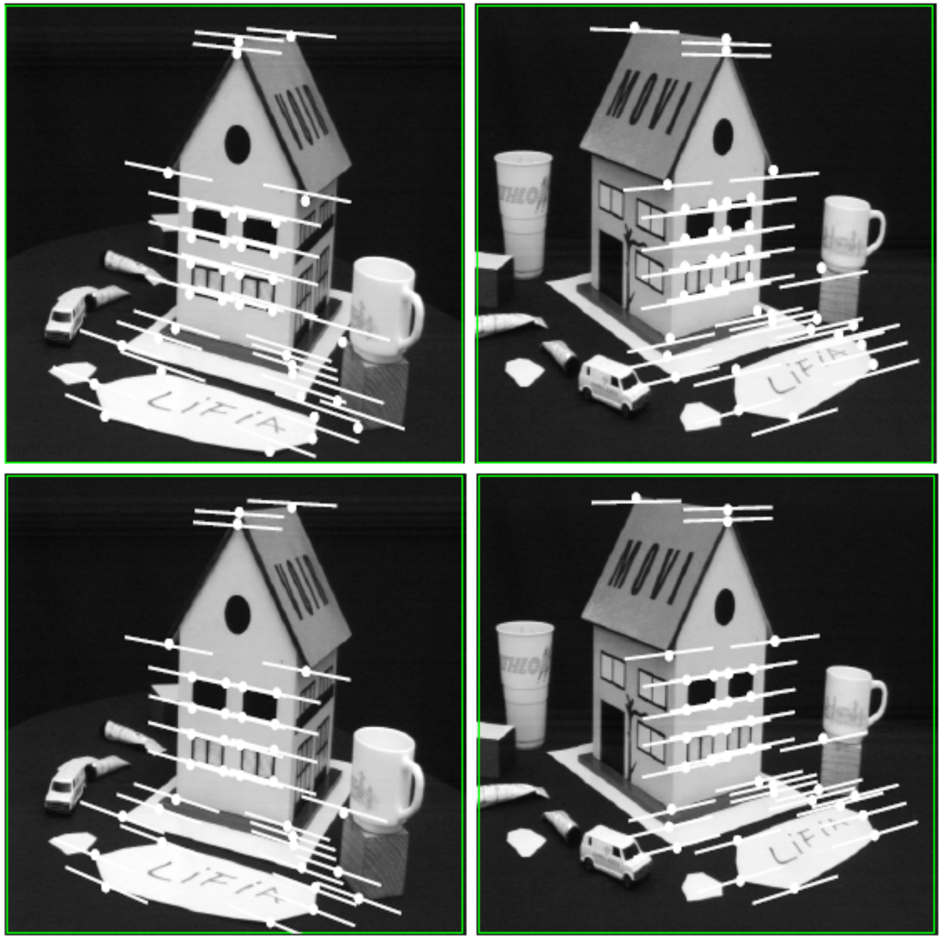
- Nonlinear approach: minimize sum of squared *geometric* distances

$$\sum_{i=1}^{N} \left[\mathrm{d}^2(\boldsymbol{x}'_i, \boldsymbol{F}\boldsymbol{x}_i) + \mathrm{d}^2(\boldsymbol{x}_i, \boldsymbol{F}^T\boldsymbol{x}'_i) \right]$$


$\boldsymbol{x}_i$
$\boldsymbol{F}^T\boldsymbol{x}'_i$
$\boldsymbol{x}'_i$
$\boldsymbol{F}\boldsymbol{x}_i$

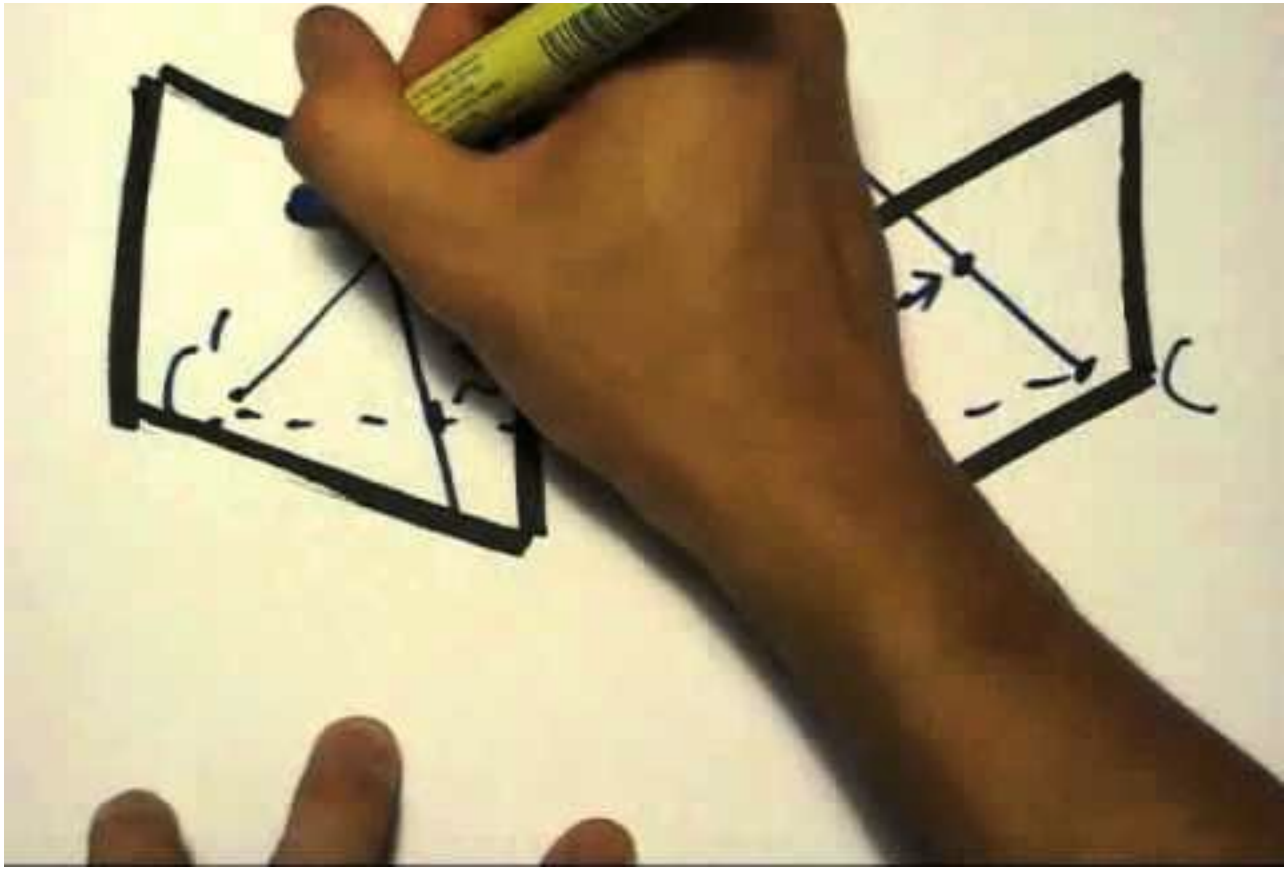
Source: S. Lazebnik

Comparison of estimation algorithms



	8-point	Normalized 8-point	Nonlinear least squares
Av. Dist. 1	2.33 pixels	0.92 pixel	0.86 pixel
Av. Dist. 2	2.18 pixels	0.85 pixel	0.80 pixel

The Fundamental Matrix Song



http://danielwedge.com/fmatrix/

Source: S. Lazebnik

From epipolar geometry to camera calibration

- Estimating the fundamental matrix is known as “weak calibration”
- If we know the calibration matrices of the two cameras, we can estimate the essential matrix: $\boldsymbol{E} = \boldsymbol{K}'^{T}\boldsymbol{F}\boldsymbol{K}$
- The essential matrix gives us the relative rotation and translation between the cameras, or their extrinsic parameters

Source: S. Lazebnik

Stereo

Many slides adapted from Steve Seitz

Source: S. Lazebnik

Binocular stereo

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

image 1

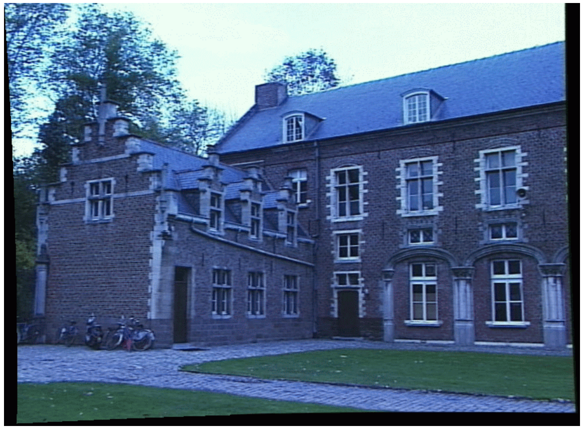
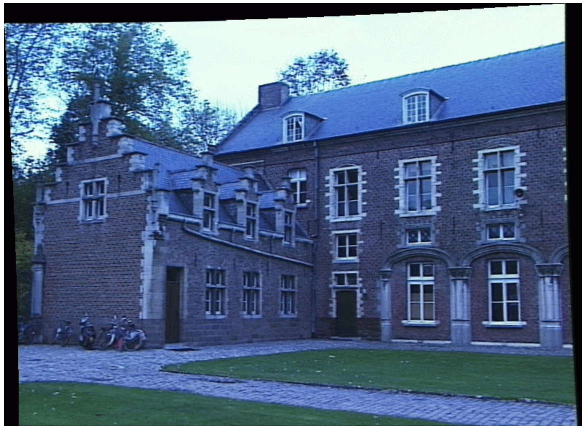
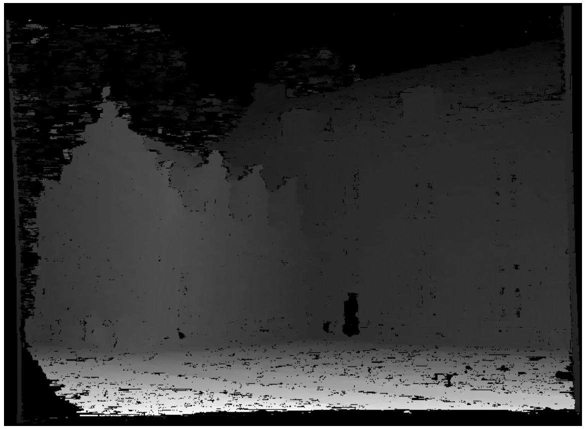


image 2



Dense depth map



Source: S. Lazebnik

Binocular stereo

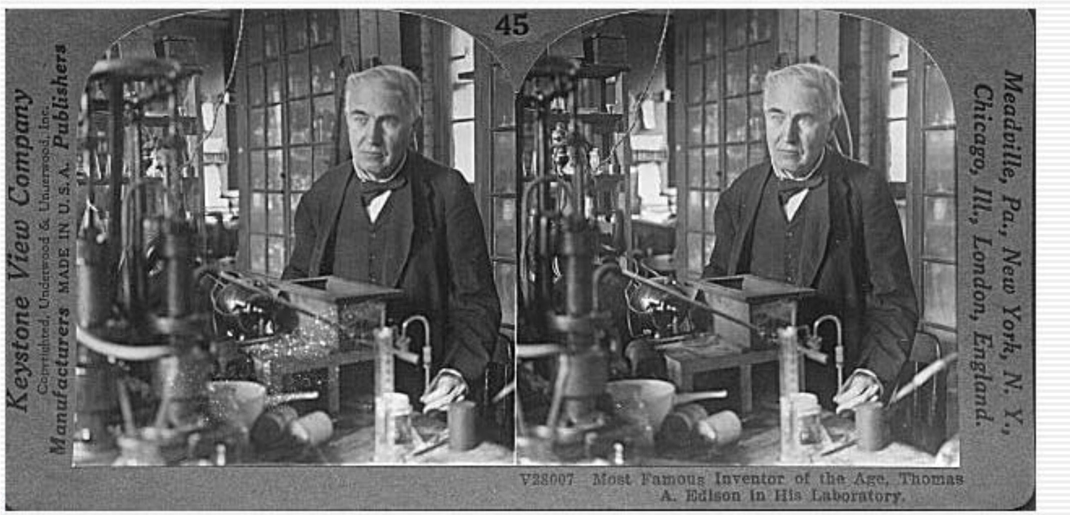
- Given a calibrated binocular stereo pair, fuse it to produce a depth image

Where does the depth information come from?

Source: S. Lazebnik

Binocular stereo

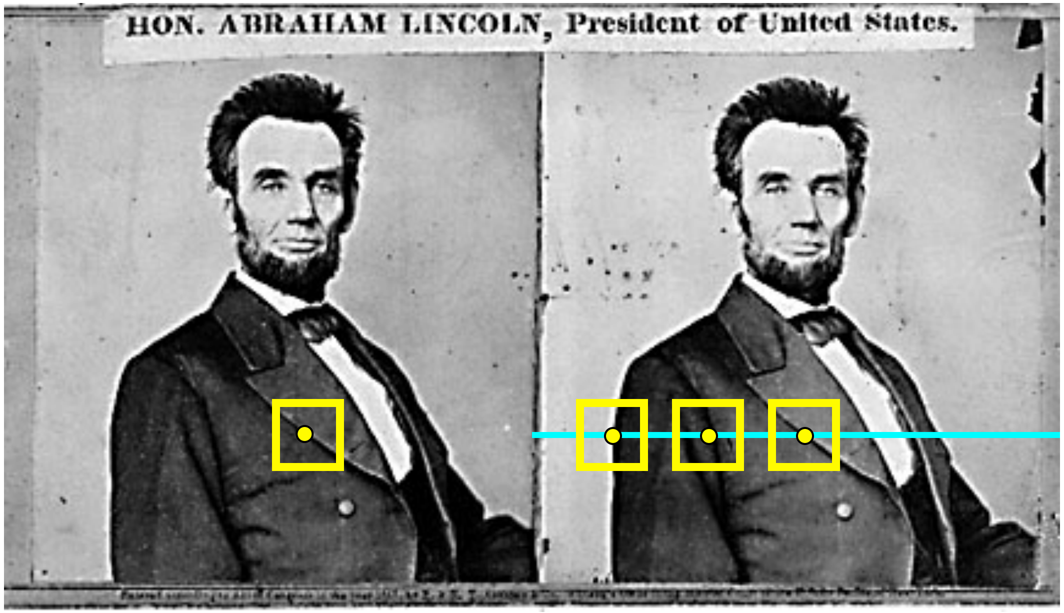
- Given a calibrated binocular stereo pair, fuse it to produce a depth image
  - Humans can do it


45
Keystone View Company
Copyrighted, Underwood & Underwood, Inc.
Manufacturers MADE IN U.S.A. Publishers
Meadville, Pa., New York, N. Y., Chicago, Ill., London, England.
V28007 Most Famous Inventor of the Age, Thomas A. Edison in His Laboratory.

Stereograms: Invented by Sir Charles Wheatstone, 1838

Source: S. Lazebnik

Basic stereo matching algorithm

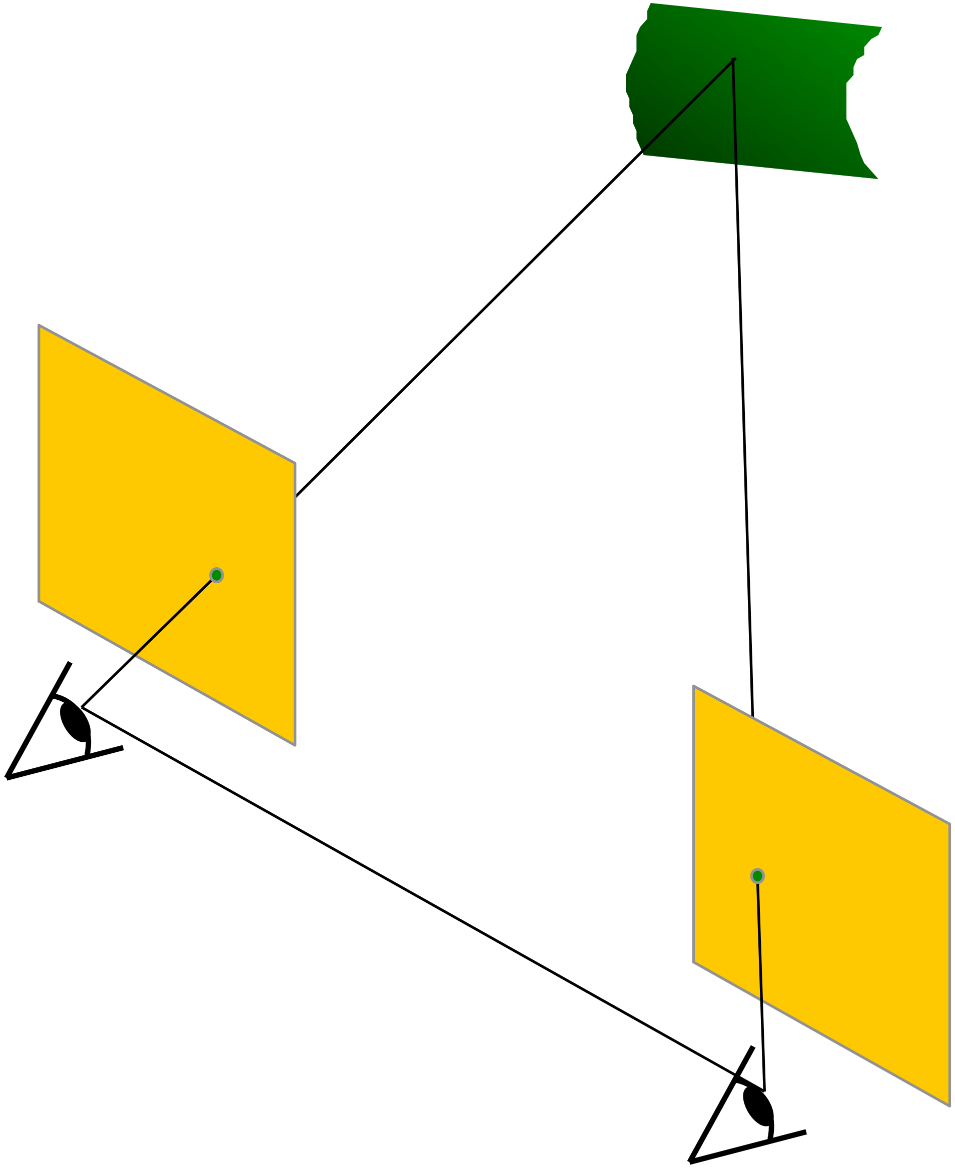


- For each pixel in the first image
  - Find corresponding epipolar line in the right image
  - Examine all pixels on the epipolar line and pick the best match
  - Triangulate the matches to get depth information

- Simplest case: epipolar lines are corresponding scanlines
  - When does this happen?

Source: S. Lazebnik

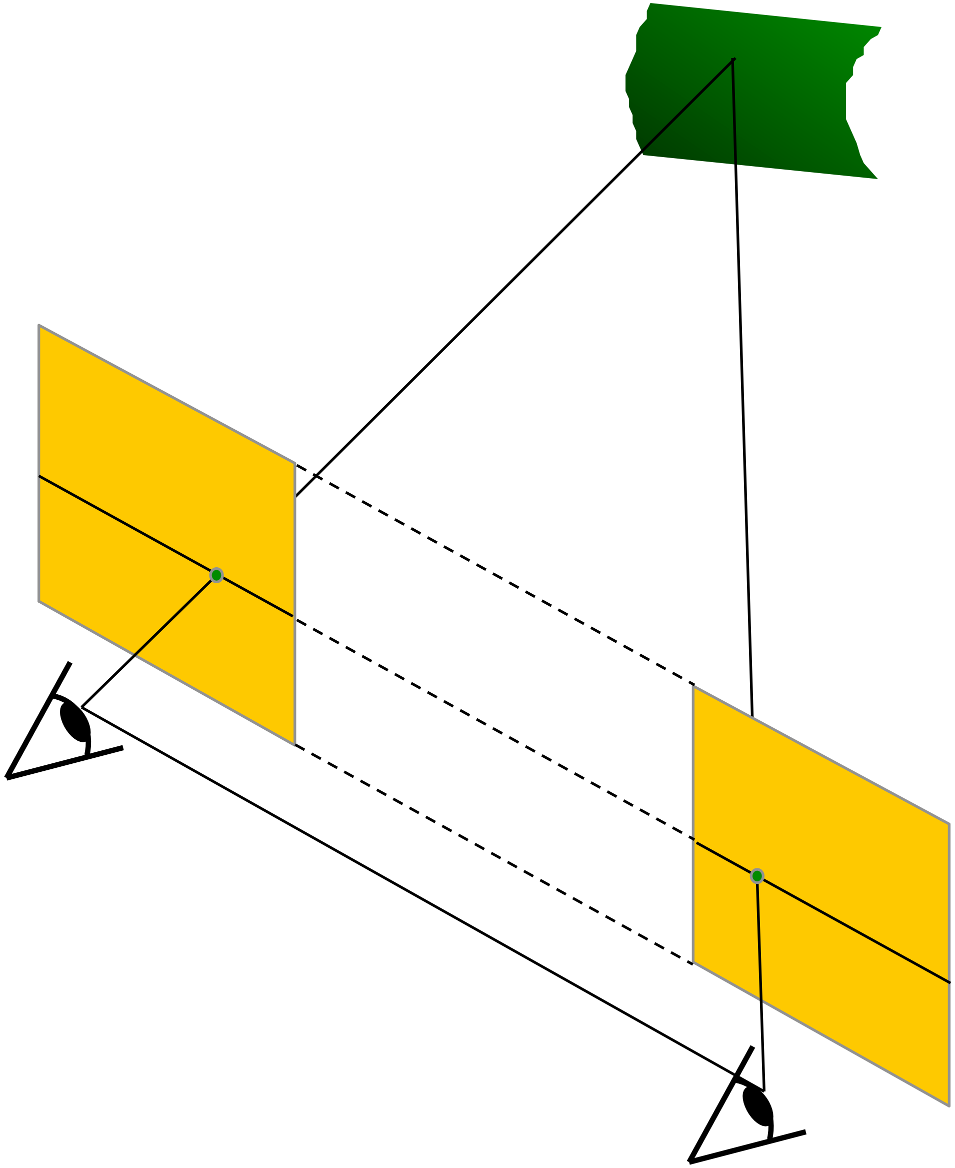
Simplest Case: Parallel images



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

Source: S. Lazebnik

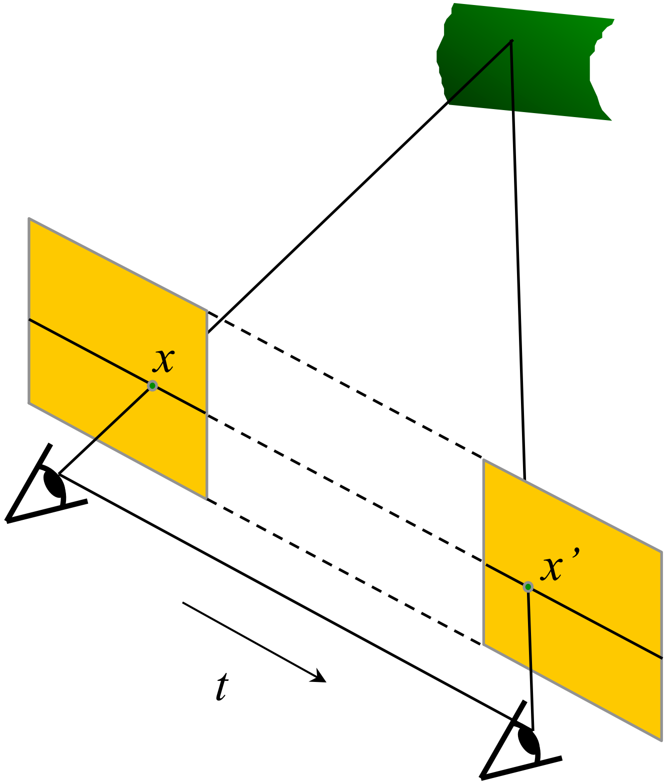
Simplest Case: Parallel images



- 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

Source: S. Lazebnik

Essential matrix for parallel images



Epipolar constraint:

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

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

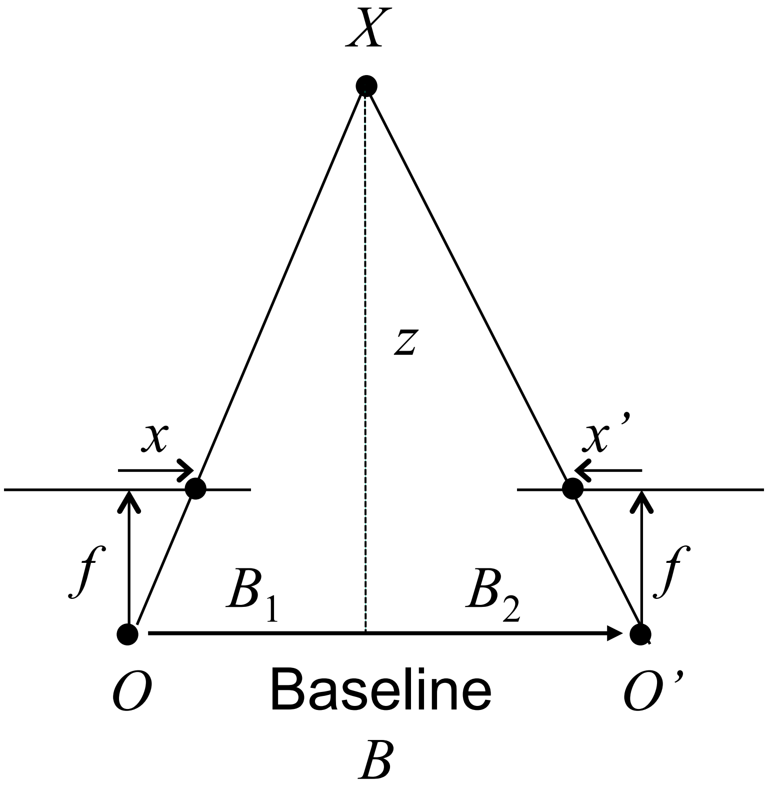
$$\boldsymbol{E} = [\boldsymbol{t}_{\times}]\boldsymbol{R} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -T \\ 0 & T & 0 \end{bmatrix}$$

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

The y-coordinates of corresponding points are the same!

Source: S. Lazebnik

Depth from disparity



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

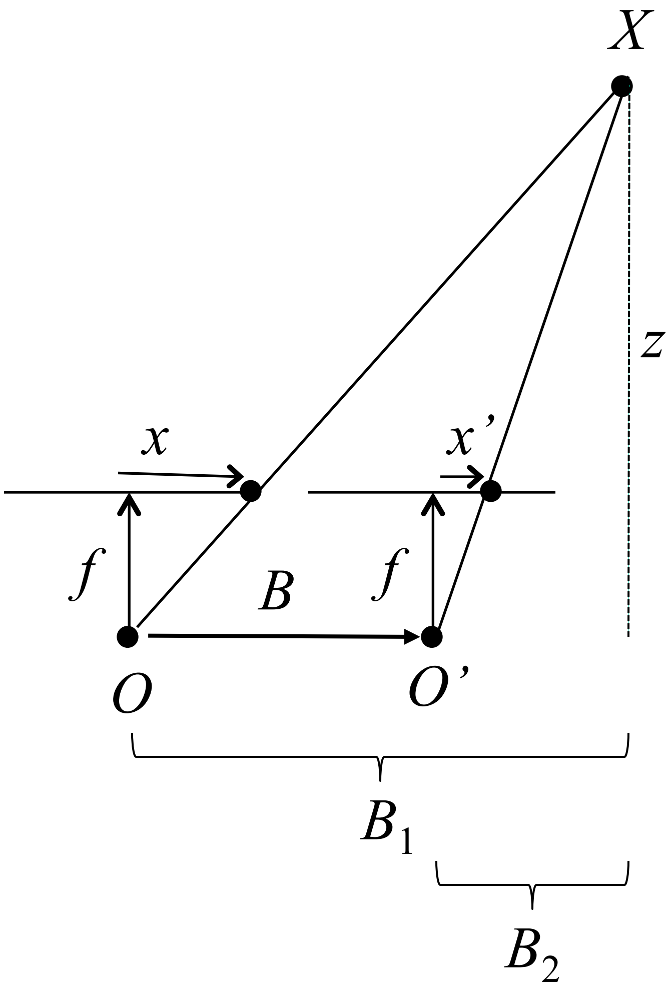
$$\frac{x - x'}{f} = \frac{B_1 + B_2}{z}$$

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

Disparity is inversely proportional to depth!

Source: S. Lazebnik

Depth from disparity



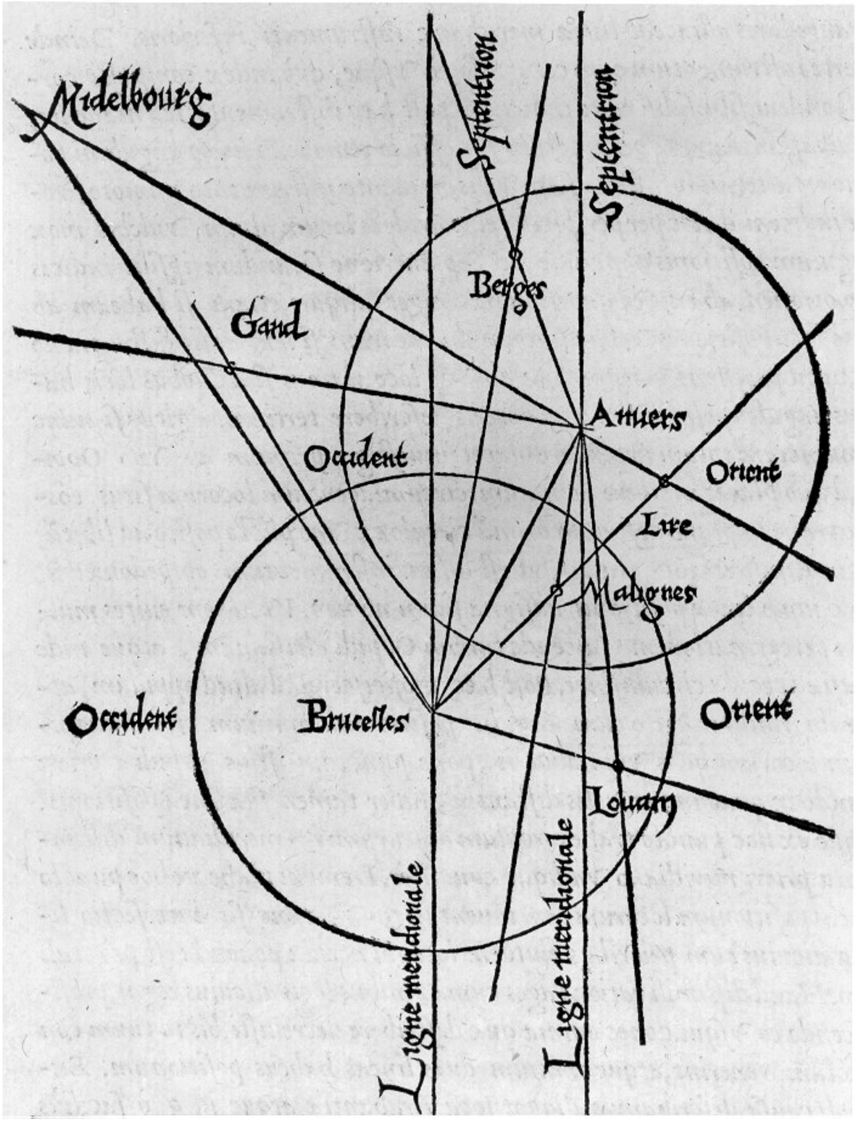
$$\frac{x}{f} = \frac{B_1}{z} \qquad \frac{x'}{f} = \frac{B_2}{z}$$

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

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

Source: S. Lazebnik

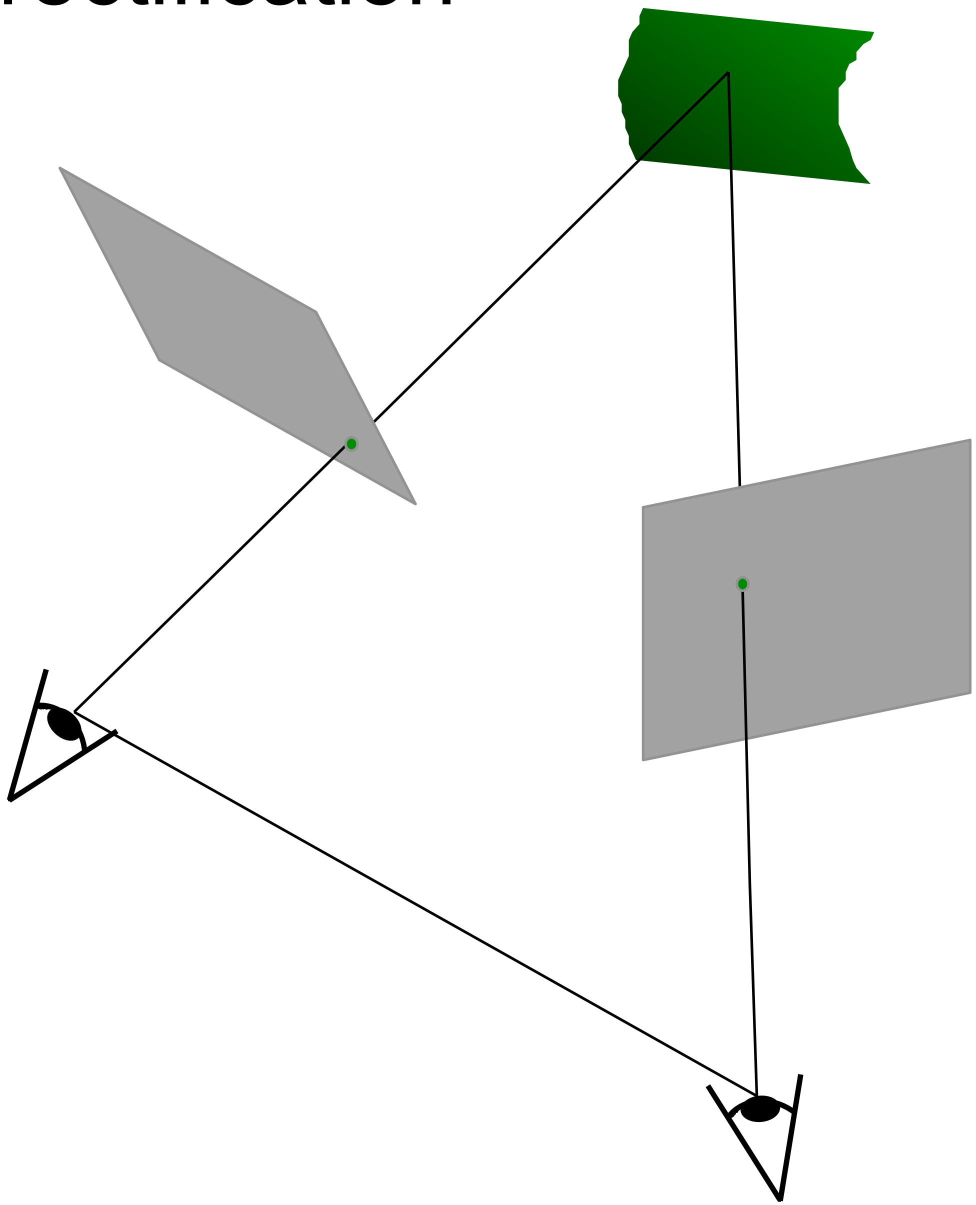
Triangulation: History



From Wikipedia: *Gemma Frisius's 1533 diagram introducing the idea of triangulation into the science of surveying. Having established a baseline, e.g. the cities of Brussels and Antwerp, the location of other cities, e.g. Middelburg, Ghent etc., can be found by taking a compass direction from each end of the baseline, and plotting where the two directions cross. This was only a theoretical presentation of the concept — due to topographical restrictions, it is impossible to see Middelburg from either Brussels or Antwerp. Nevertheless, the figure soon became well known all across Europe.*

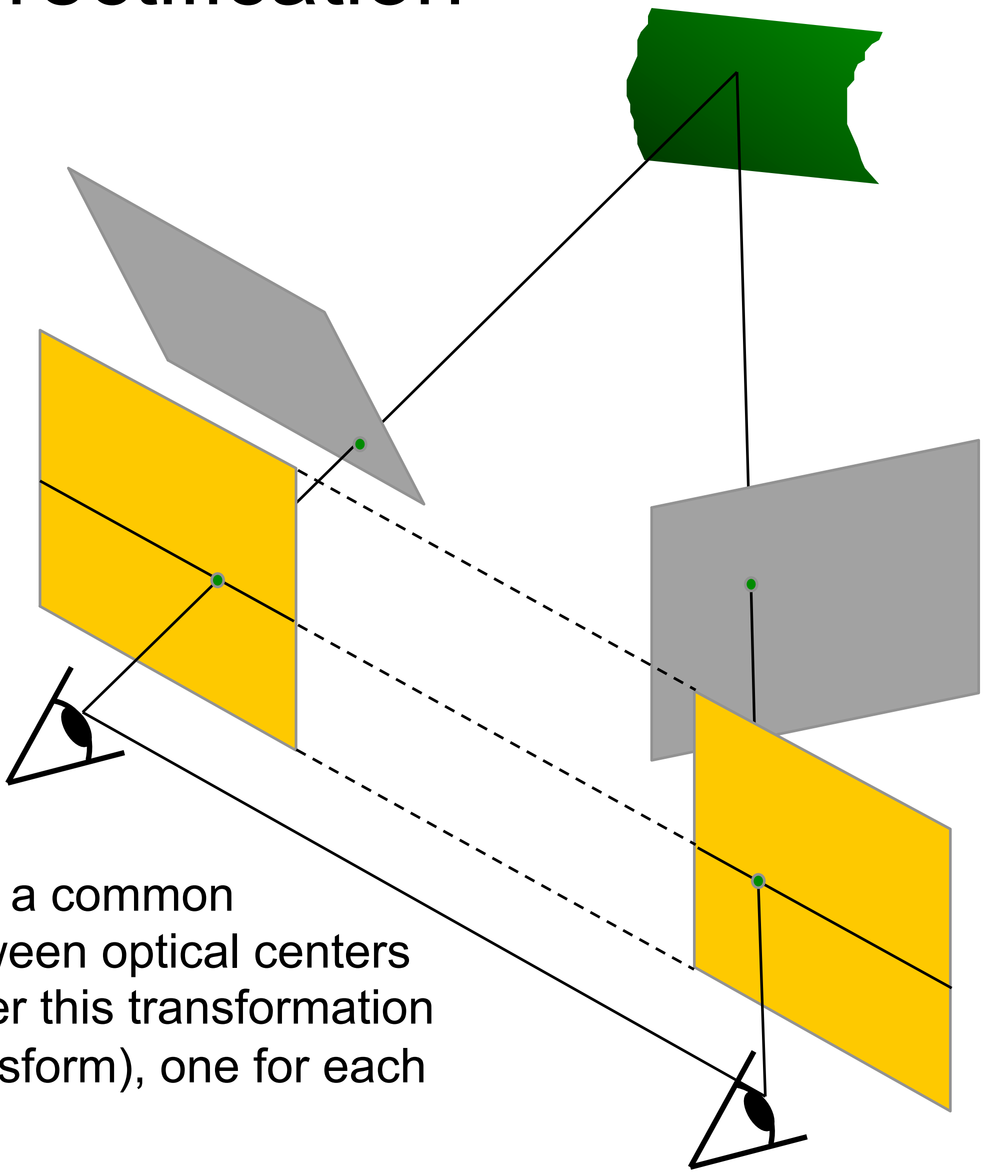
Source: S. Lazebnik

Stereo image rectification



Source: S. Lazebnik

Stereo image rectification

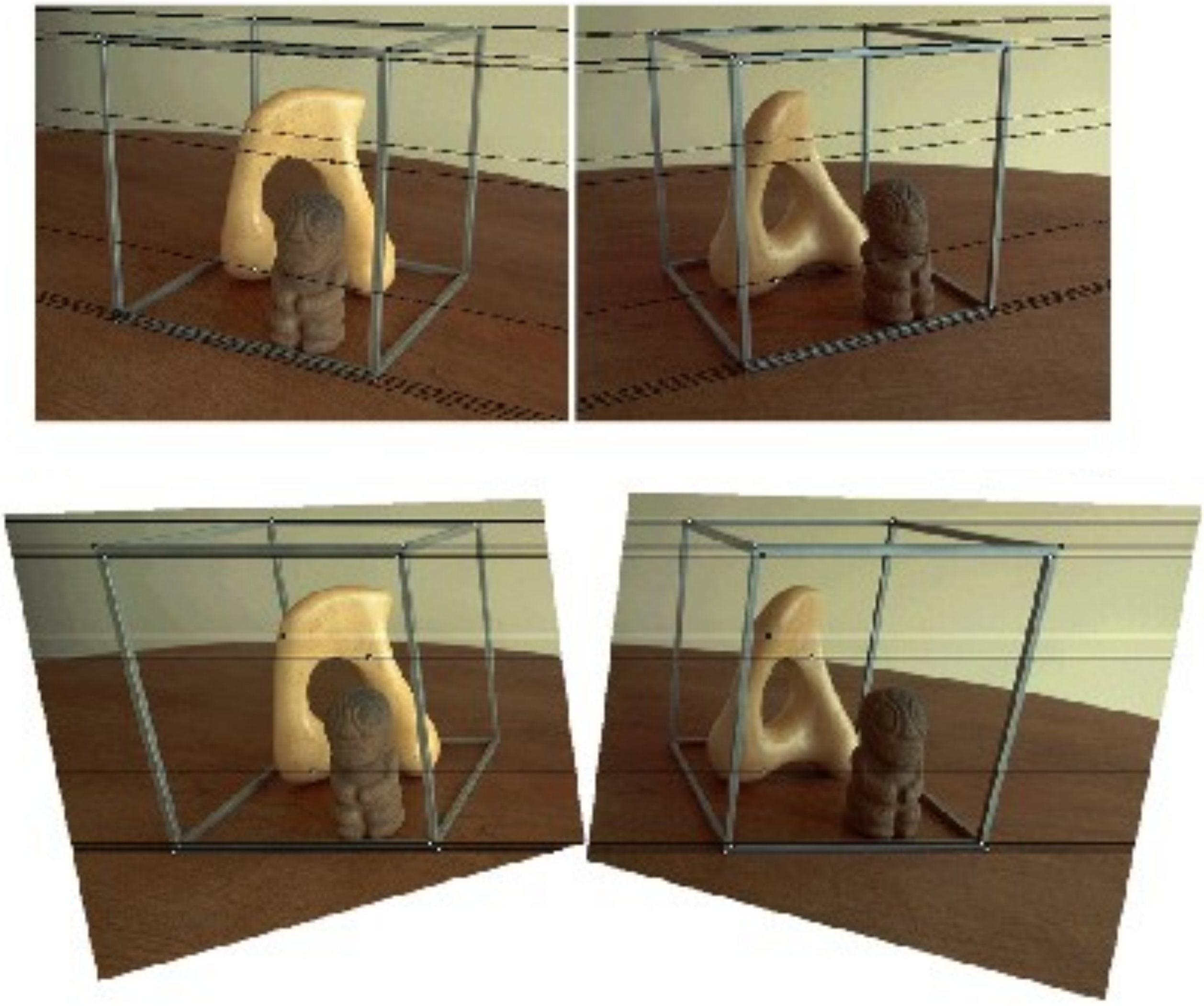


- Reproject image planes onto a common plane parallel to the line between optical centers
- Pixel motion is horizontal after this transformation
- Two homographies (3x3 transform), one for each input image reprojection

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

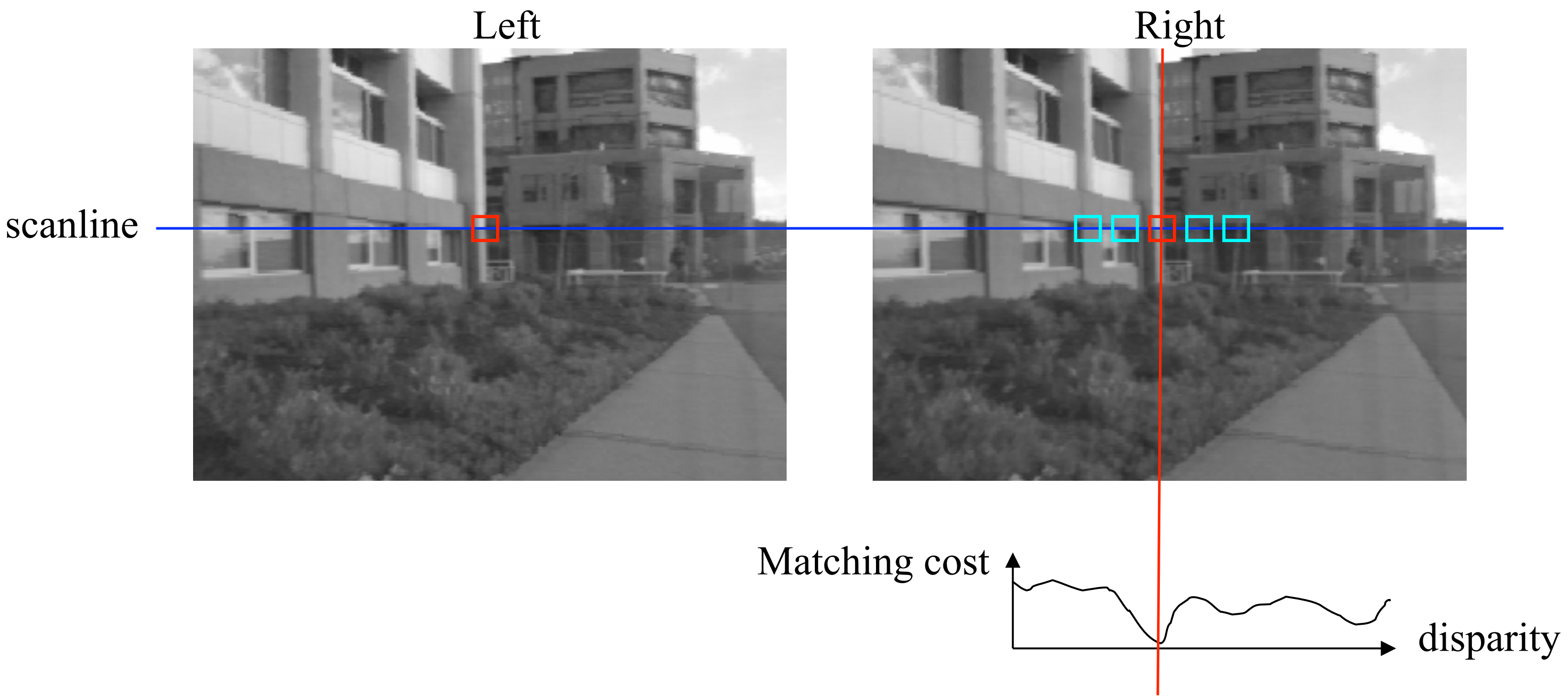
Source: S. Lazebnik

Rectification example



Source: S. Lazebnik

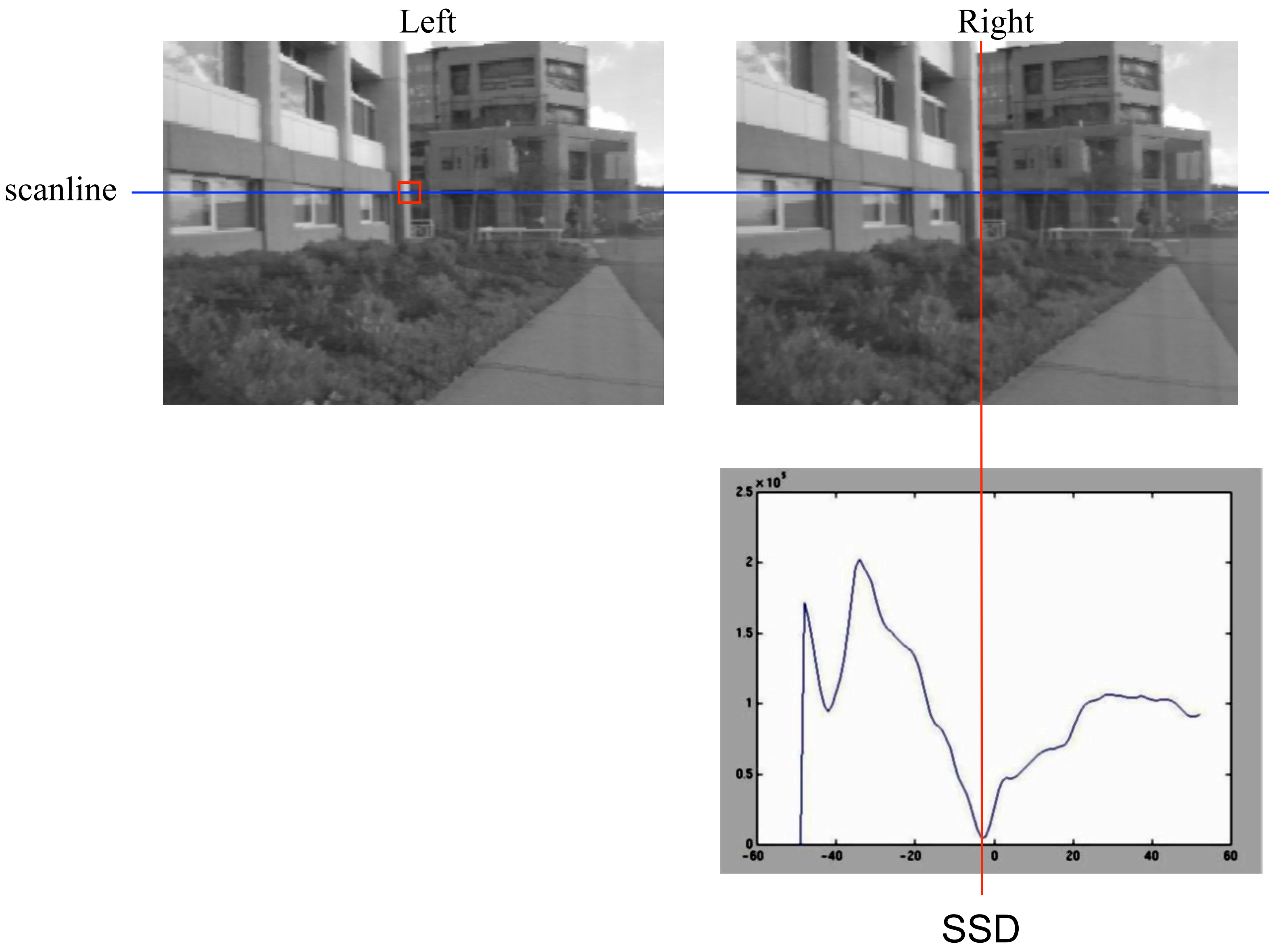
Correspondence search



- 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

Source: S. Lazebnik

Correspondence search



Left

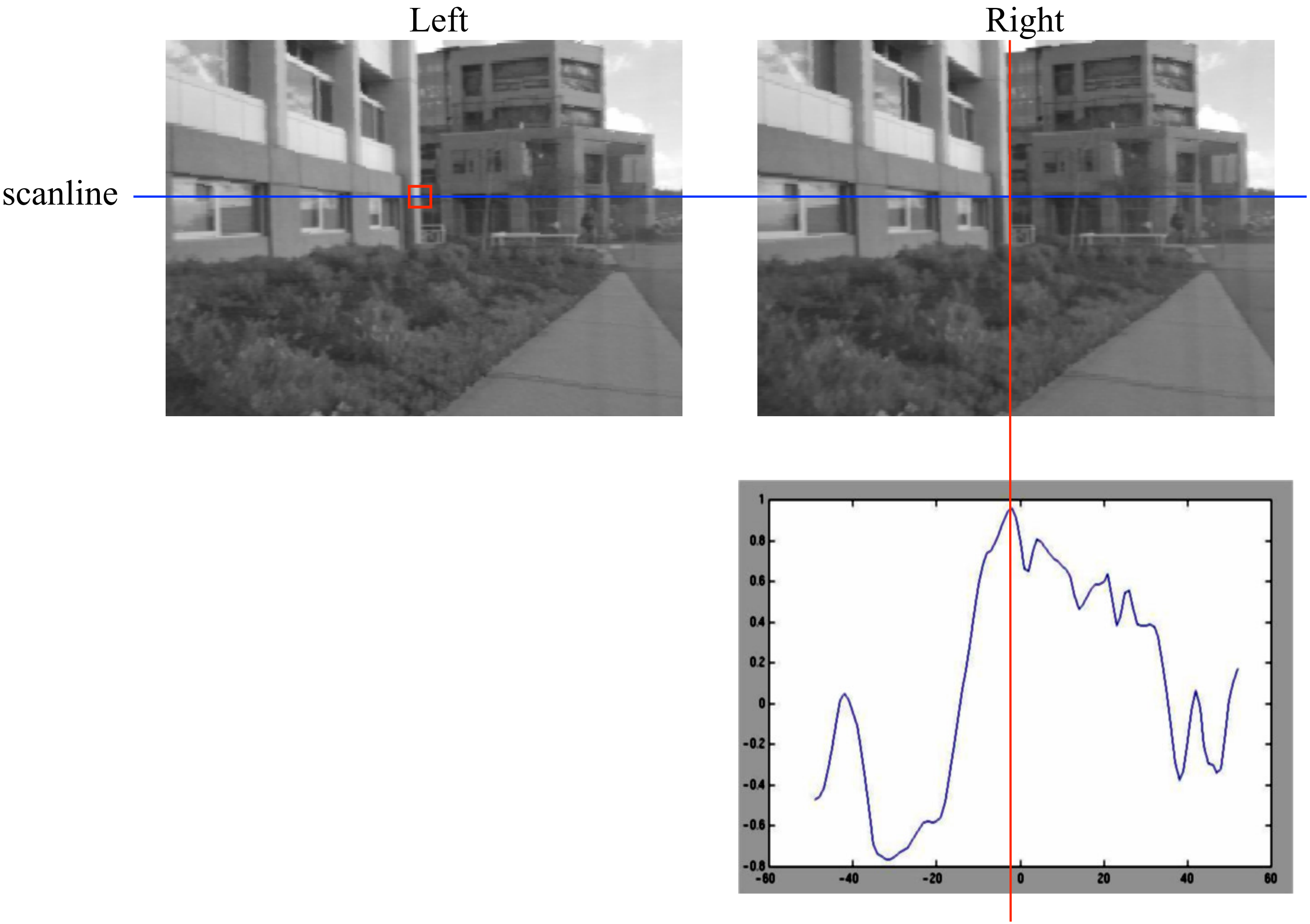
Right

scanline

SSD

Source: S. Lazebnik

Correspondence search



Left

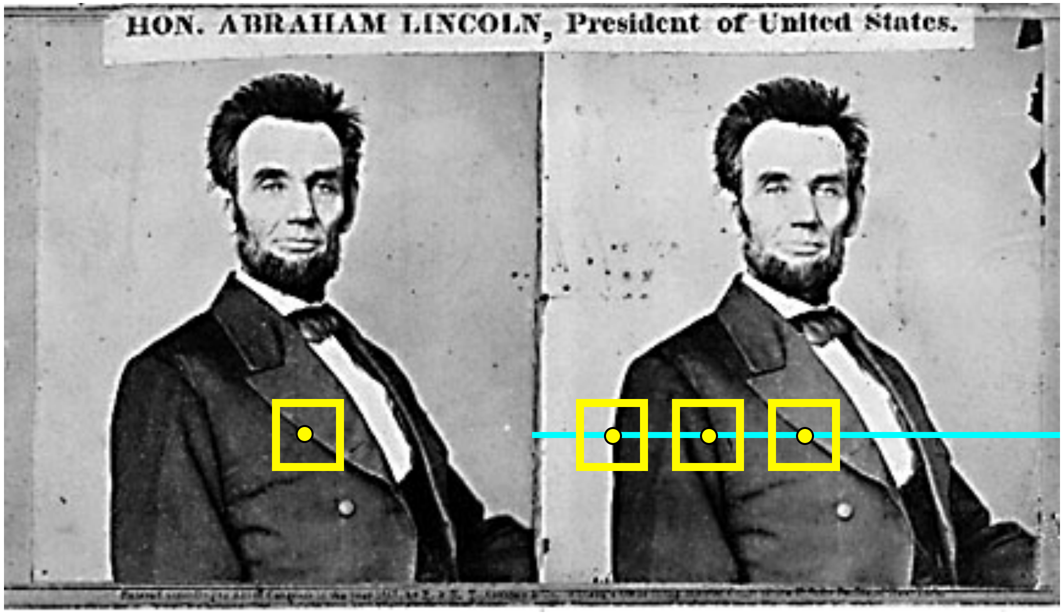
Right

scanline

Norm. corr

Source: S. Lazebnik

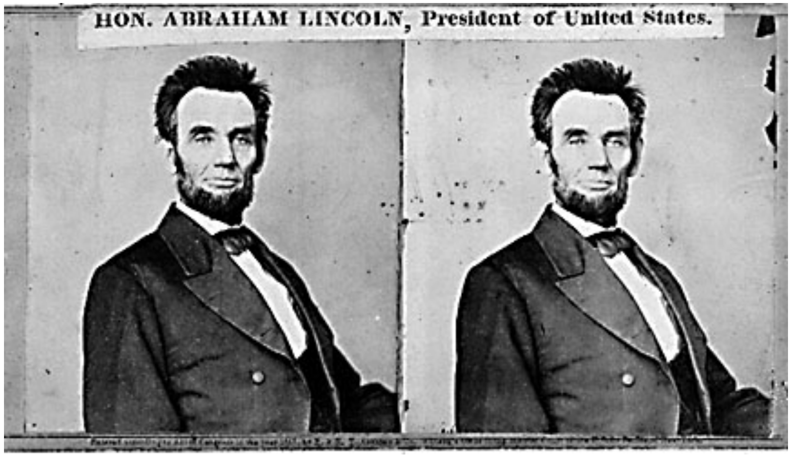
Basic stereo matching algorithm



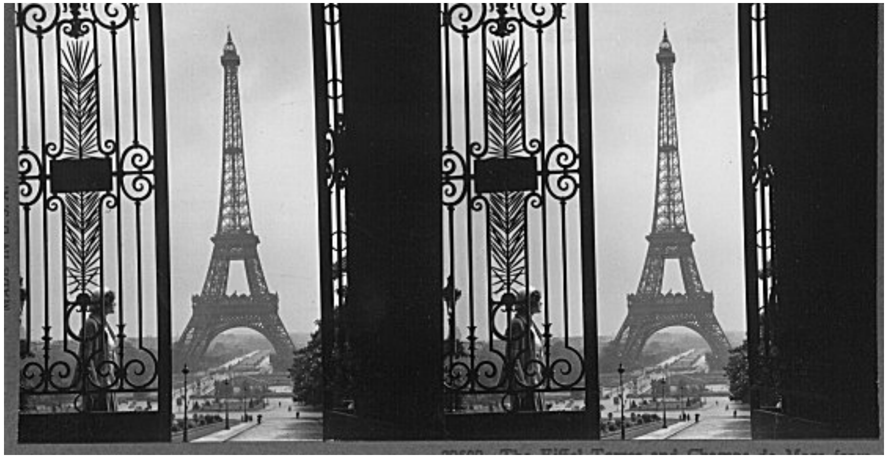
- If necessary, rectify the two stereo images to transform epipolar lines into scanlines
- For each pixel x in the first image
  - Find corresponding epipolar scanline in the right image
  - Examine all pixels on the scanline and pick the best match x'
  - Compute disparity $x–x'$ and set depth(x) = $B*f/(x–x')$

Source: S. Lazebnik

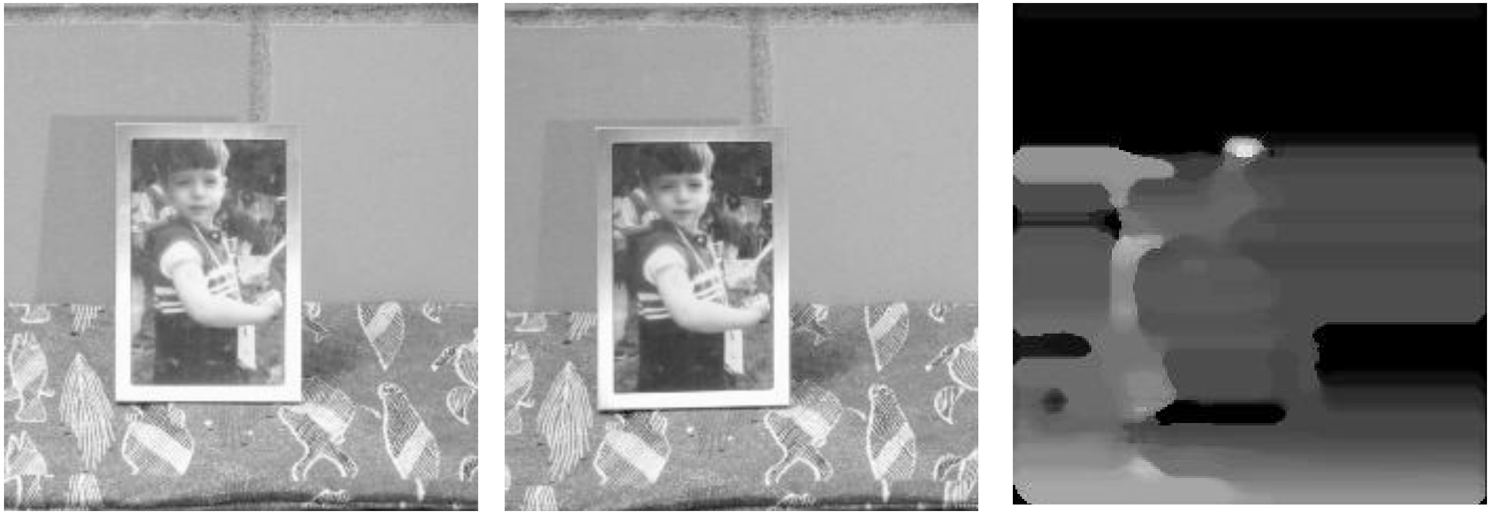
Failures of correspondence search



Textureless surfaces



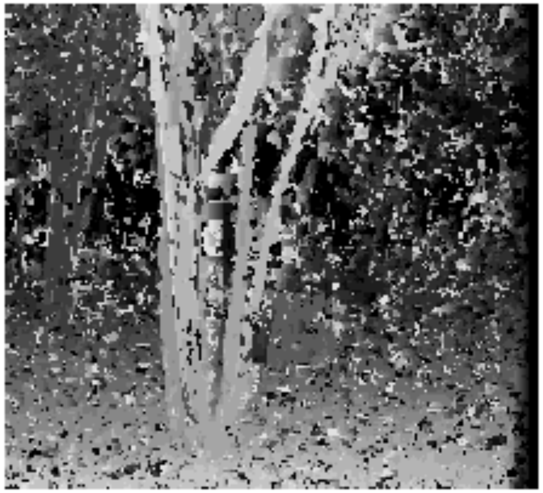
Occlusions, repetition



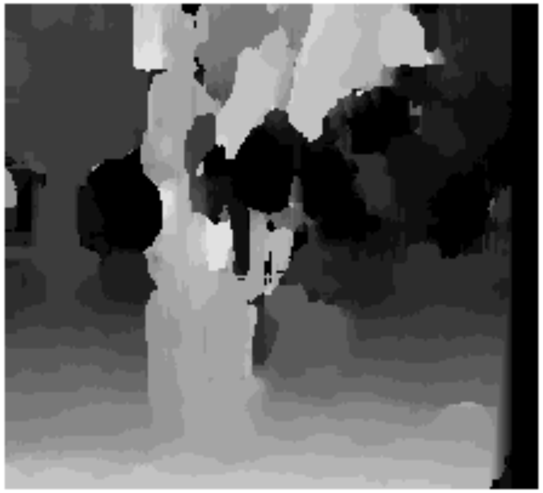
Non-Lambertian surfaces, specularities

Source: S. Lazebnik

Effect of window size



W = 3



W = 20

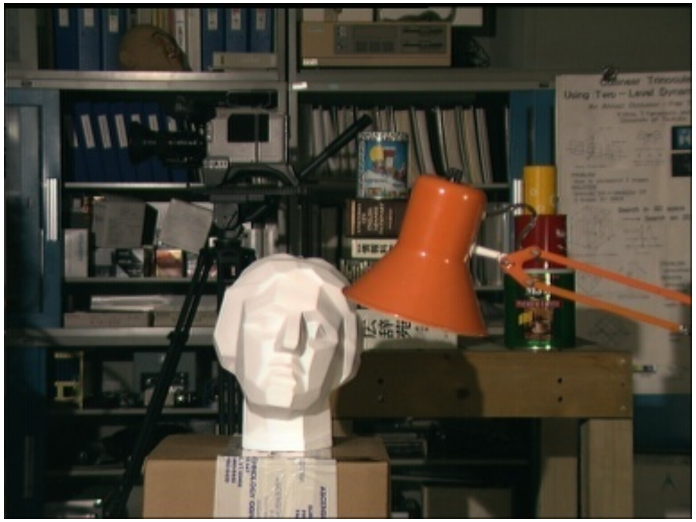
- Smaller window
  - + More detail
  - – More noise

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

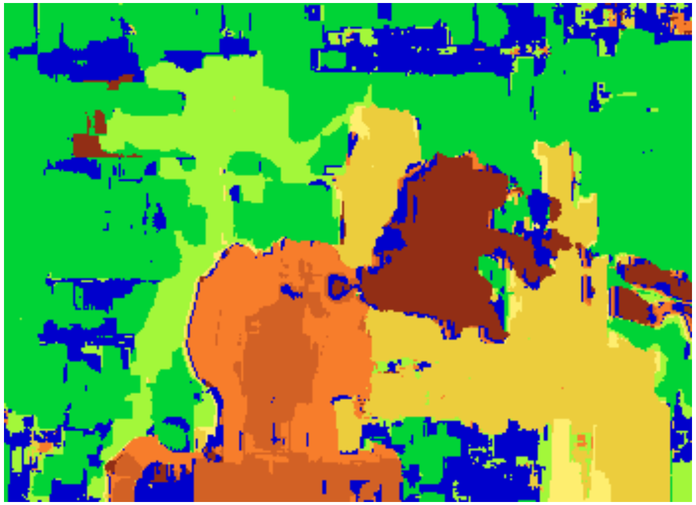
Source: S. Lazebnik

Results with window search

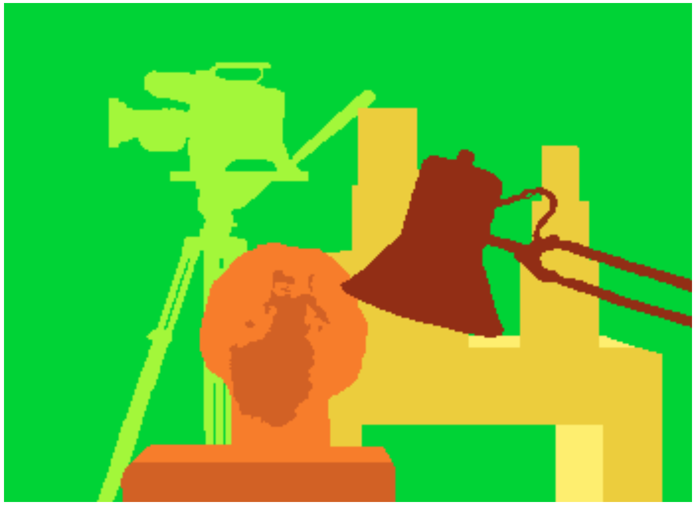
Data



Window-based matching

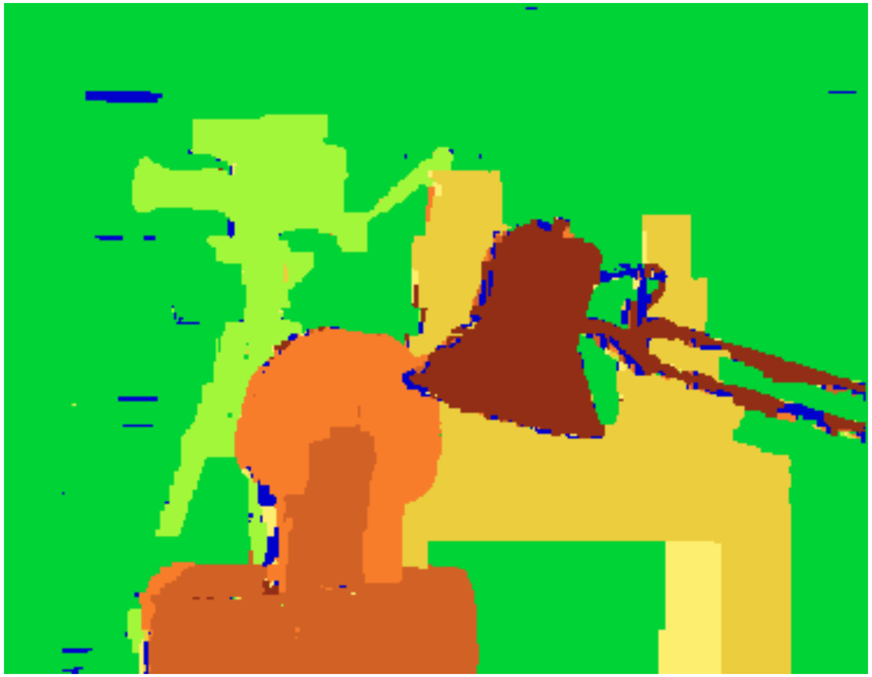


Ground truth

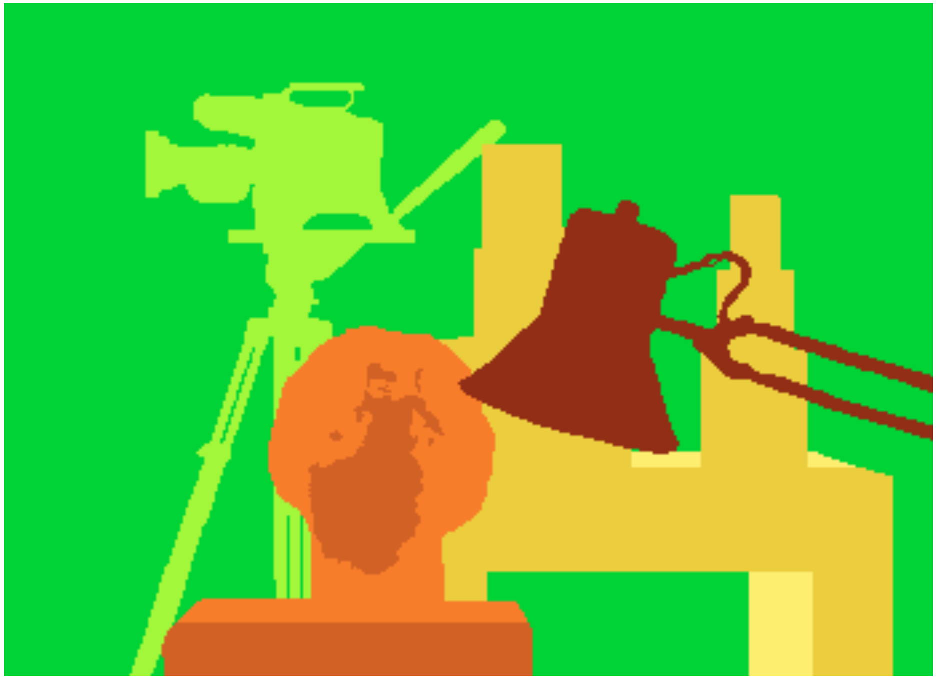


Source: S. Lazebnik

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/

Source: S. Lazebnik

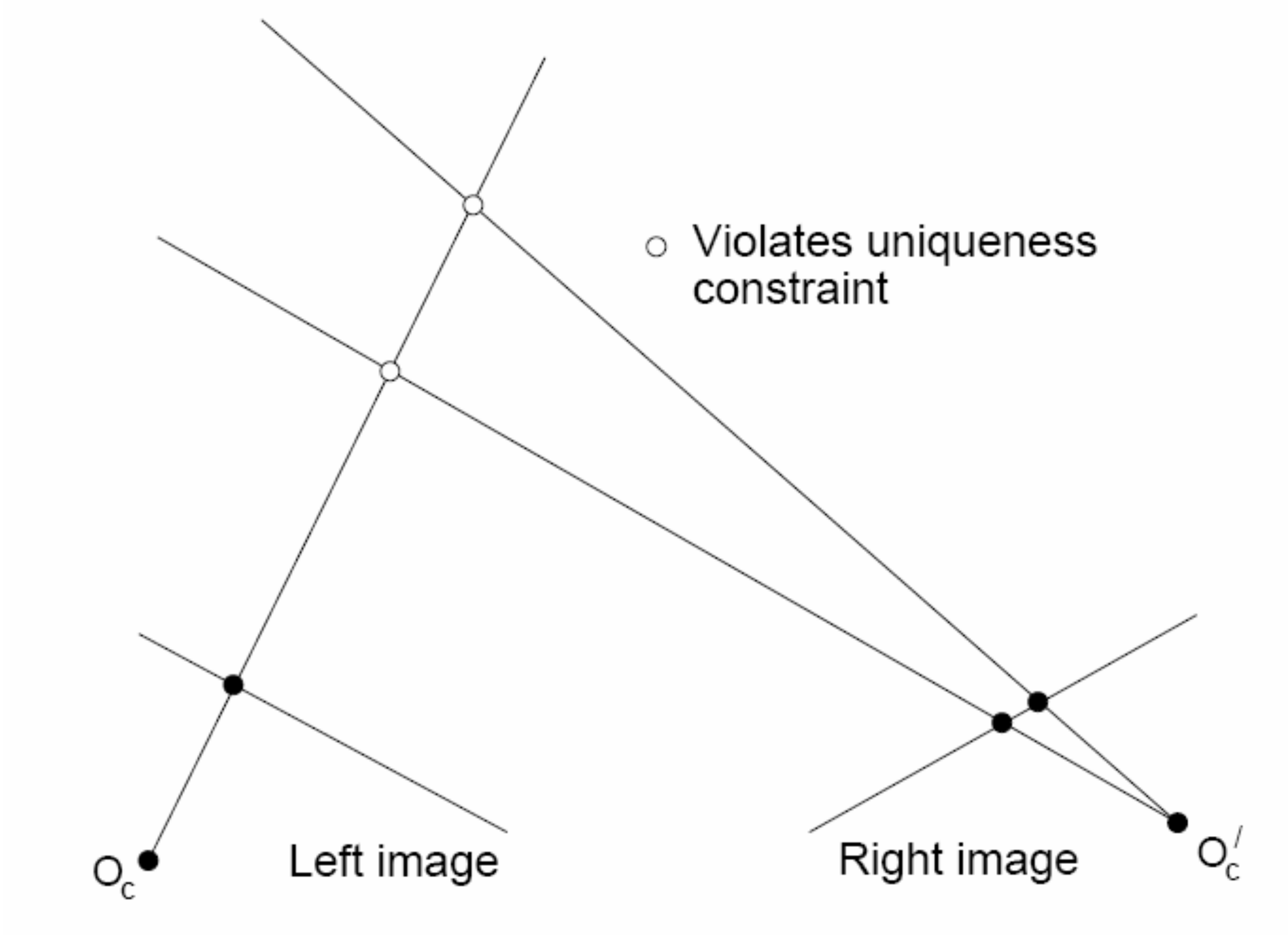
How can we improve window-based matching?

- The similarity constraint is **local** (each reference window is matched independently)
- Need to enforce **non-local** correspondence constraints

Source: S. Lazebnik

Non-local constraints

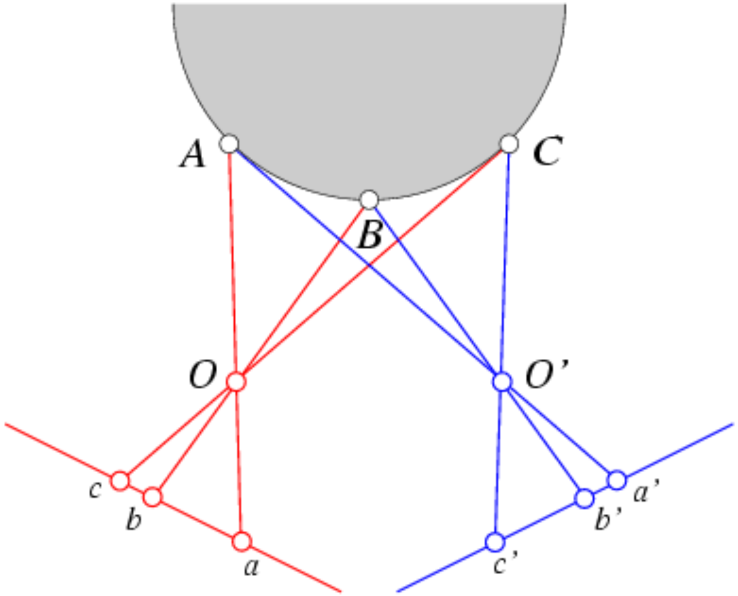
- Uniqueness
  - For any point in one image, there should be at most one matching point in the other image



Source: S. Lazebnik

Non-local constraints

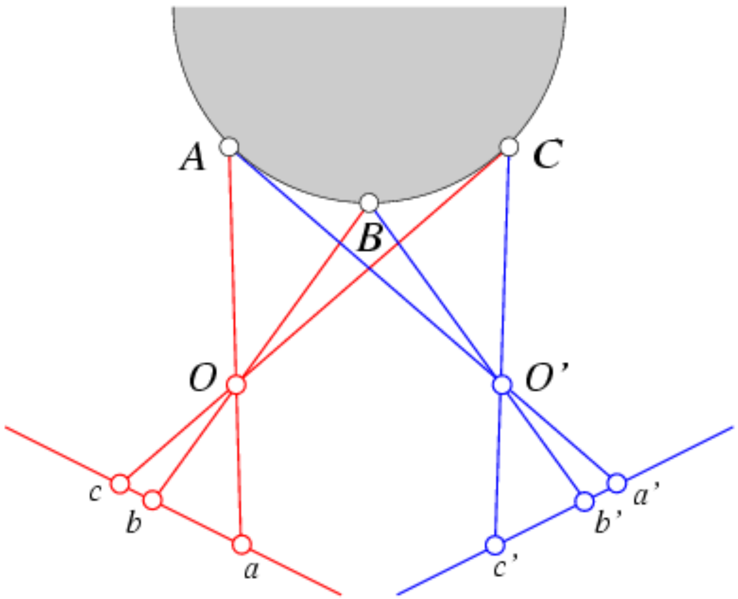
- Uniqueness
  - For any point in one image, there should be at most one matching point in the other image
- Ordering
  - Corresponding points should be in the same order in both views

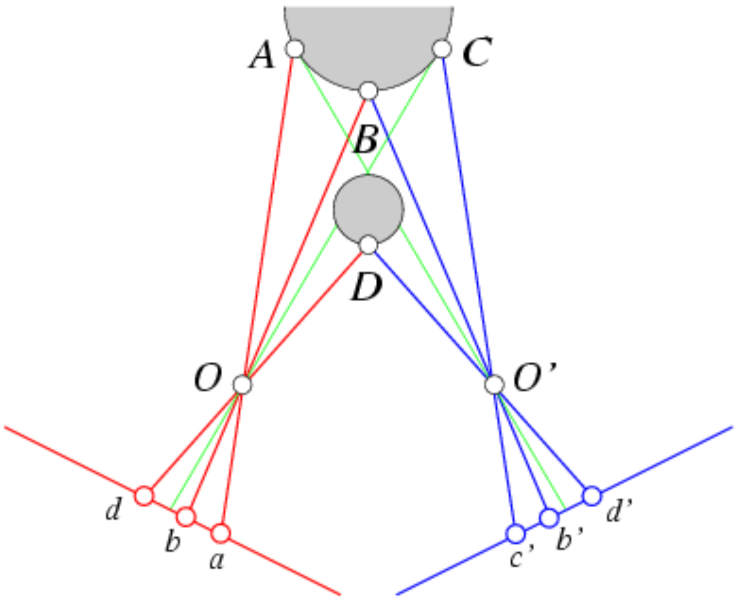


Source: S. Lazebnik

Non-local 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





Ordering constraint doesn't hold

Source: S. Lazebnik

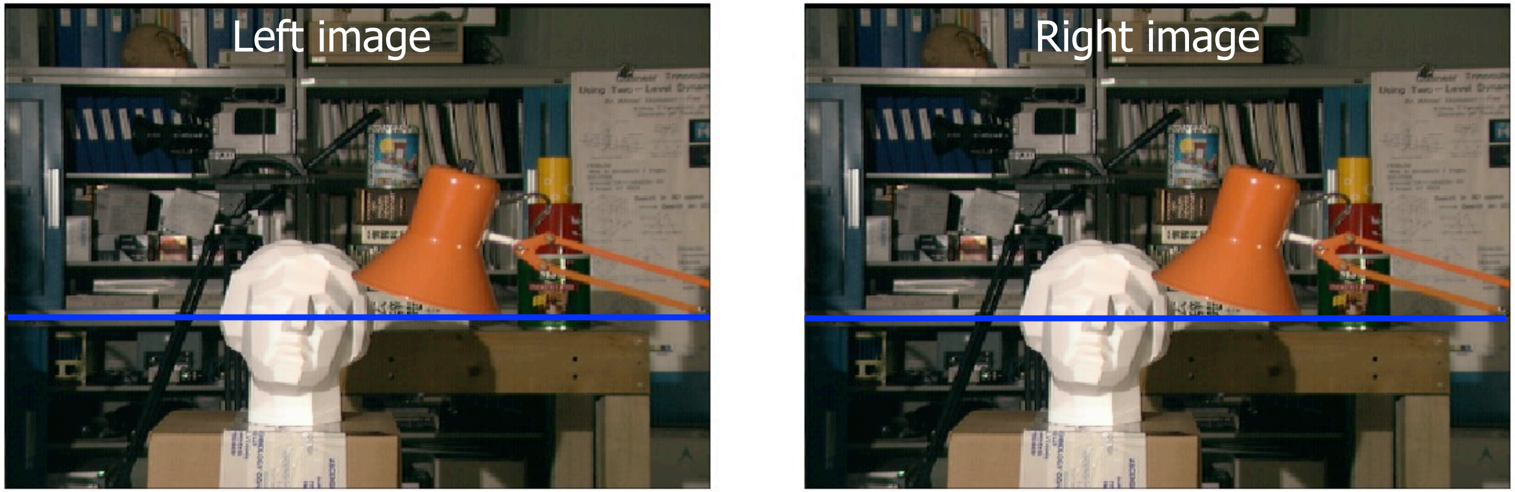
Non-local constraints

- 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)

Source: S. Lazebnik

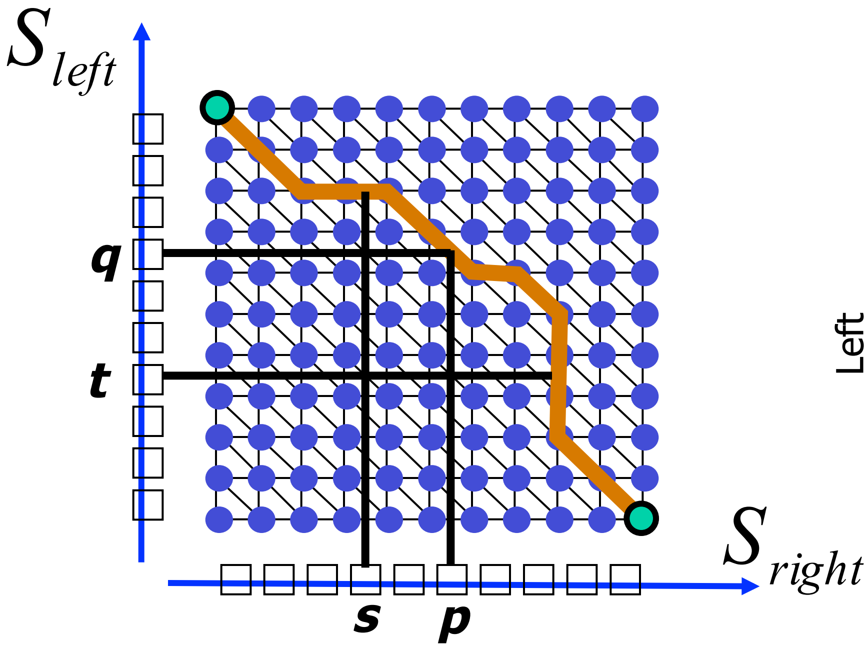
Scanline stereo

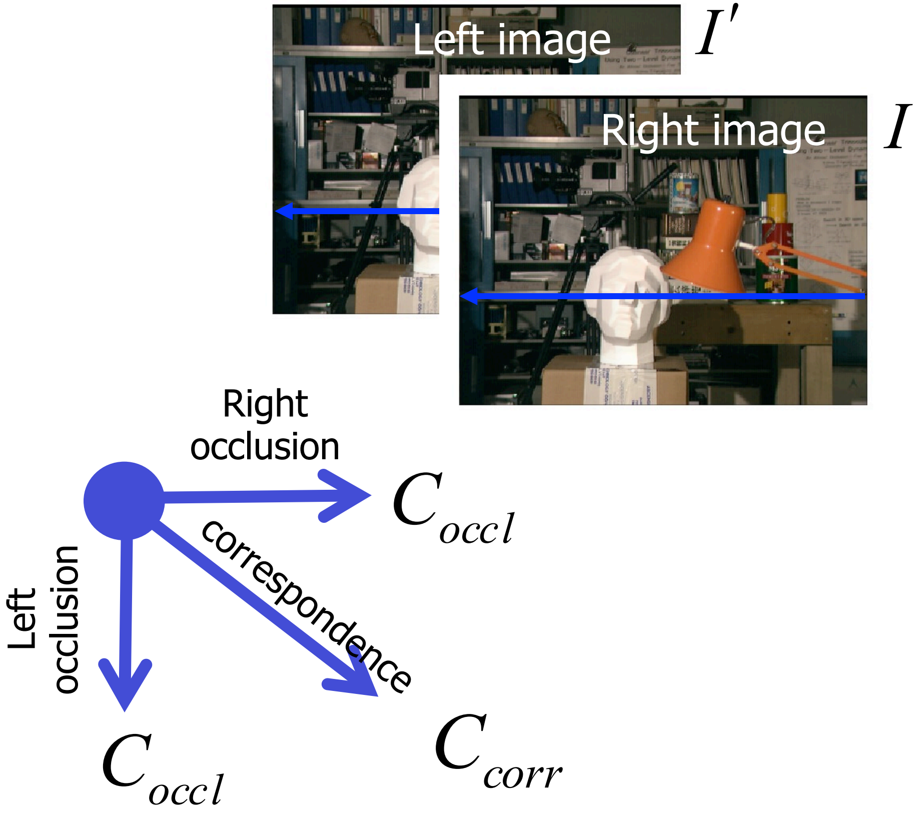
- Try to coherently match pixels on the entire scanline
- Different scanlines are still optimized independently



Source: S. Lazebnik

“Shortest paths” for scan-line stereo





Can be implemented with dynamic programming

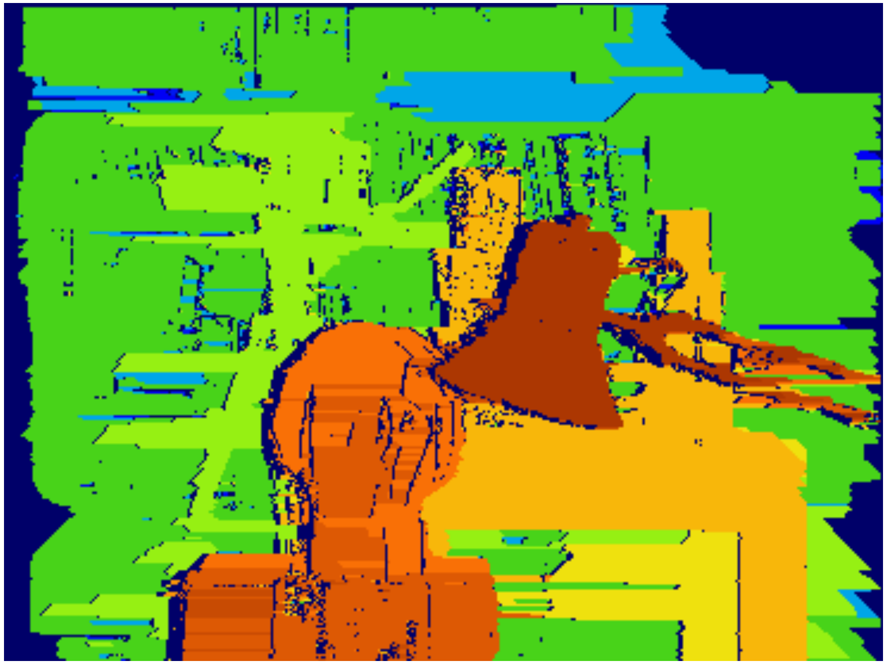
Ohta & Kanade ’85, Cox et al. ‘96

Slide credit: Y. Boykov

Source: S. Lazebnik

Coherent stereo on 2D grid

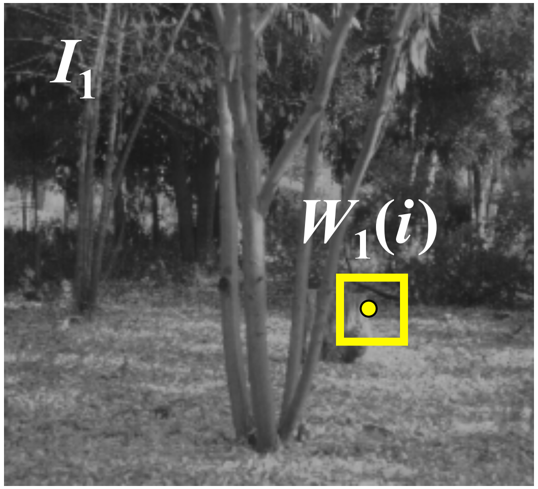
- Scanline stereo generates streaking artifacts

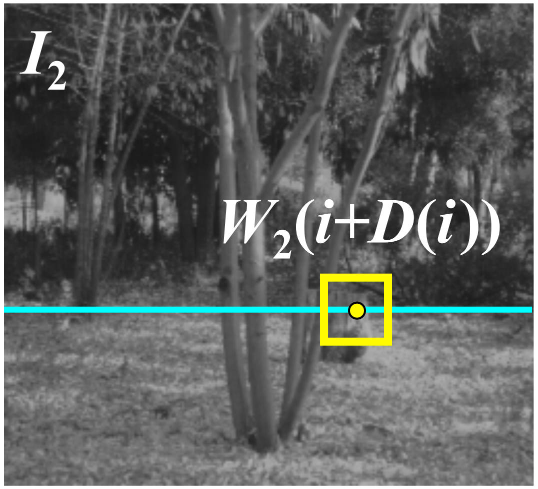


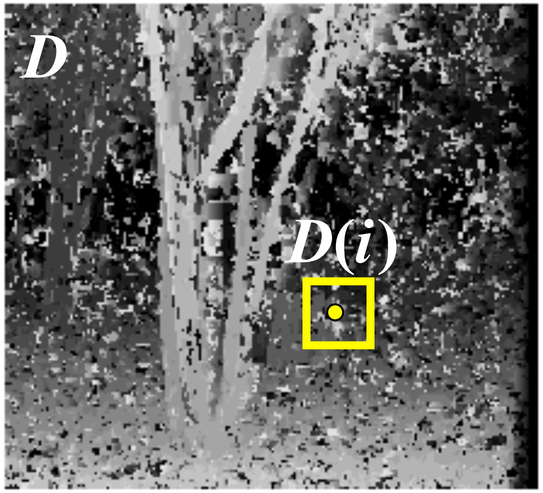
- Can’t use dynamic programming to find spatially coherent disparities/ correspondences on a 2D grid

Source: S. Lazebnik

Stereo matching as energy minimization







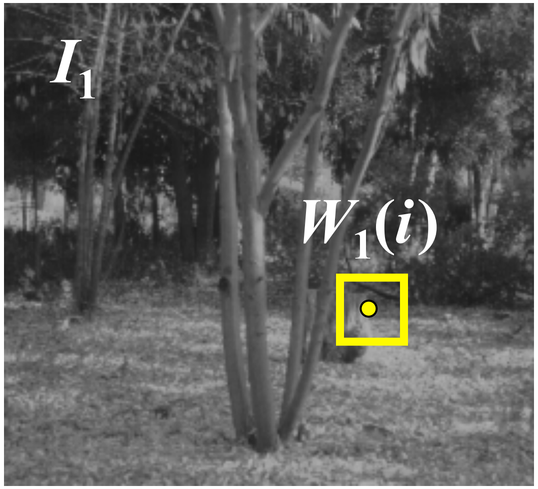
$$E(D) = \underbrace{\sum_i \left(W_1(i) - W_2(i + D(i))\right)^2}_{\text{data term}} + \lambda \underbrace{\sum_{\text{neighbors } i,j} \rho\left(D(i) - D(j)\right)}_{\text{smoothness term}}$$

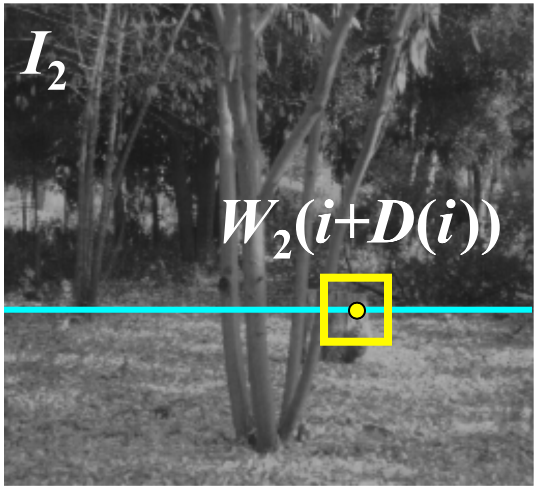
- Energy functions of this form can be minimized using *graph cuts*

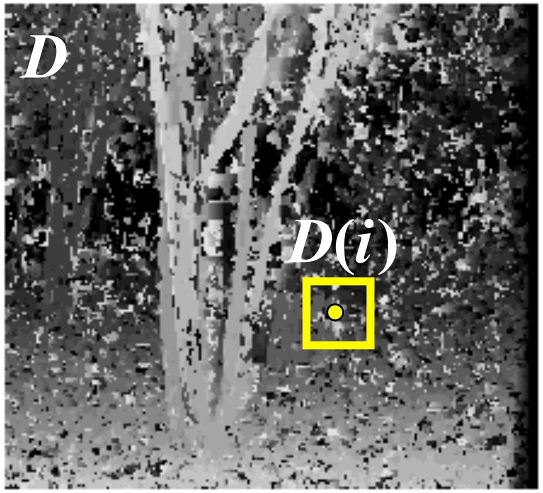
Y. Boykov, O. Veksler, and R. Zabih,
Fast Approximate Energy Minimization via Graph Cuts, PAMI 2001

Source: S. Lazebnik

Stereo matching as energy minimization







- Probabilistic interpretation: we want to find a Maximum A Posteriori (MAP) estimate of disparity image *D*:

$$P(D \mid I_1, I_2) \propto P(I_1, I_2 \mid D) P(D)$$

$$-\log P(D \mid I_1, I_2) \propto -\log P(I_1, I_2 \mid D) - \log P(D)$$

$$\boxed{E = E_{\text{data}}(I_1, I_2, D) + \lambda E_{\text{smooth}}(D)}$$

Source: S. Lazebnik

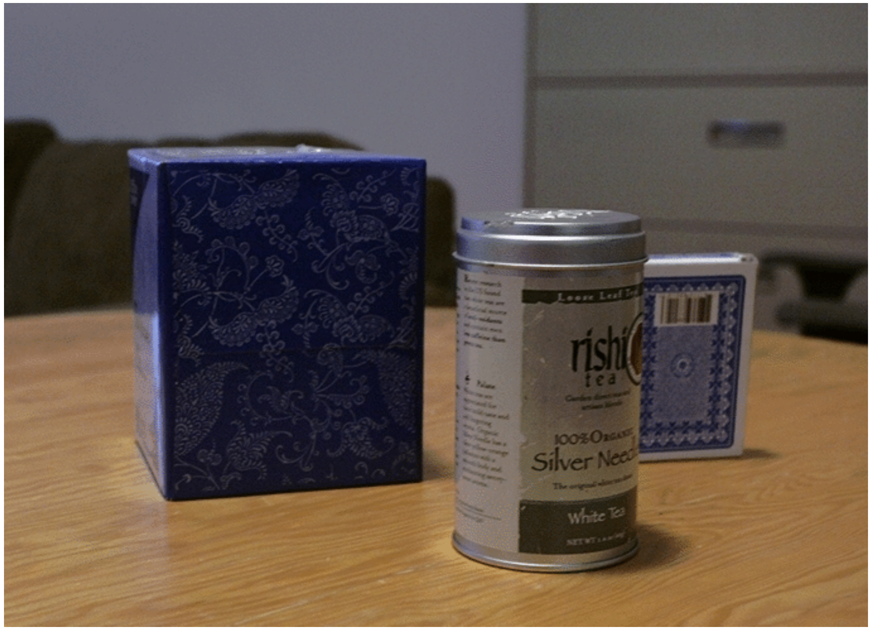
Stereo matching as energy minimization

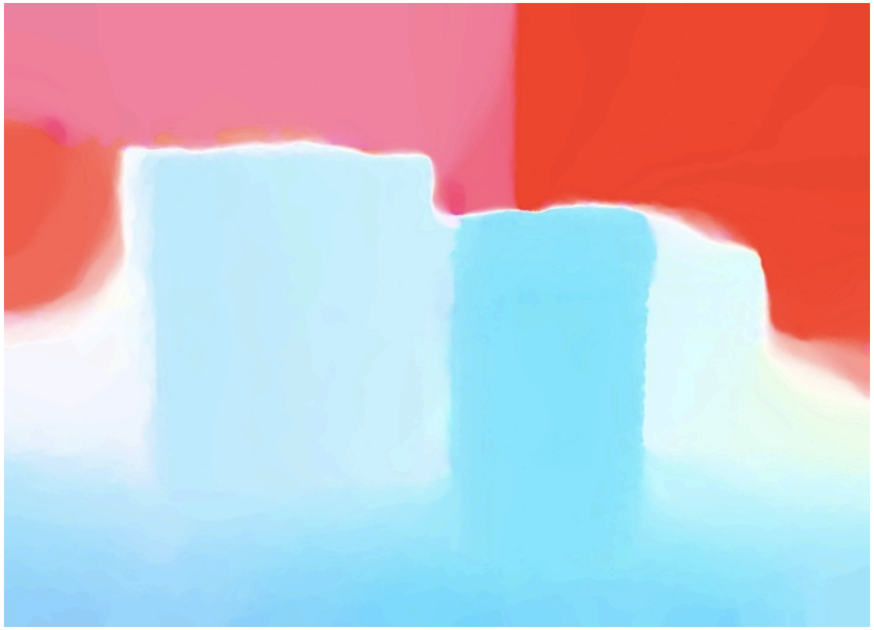
- Note: the above formulation does not treat the two images symmetrically, does not enforce uniqueness, and does not take occlusions into account
- It is possible to come up with an energy that does all these things, but it's a bit more complex
  - Defined over all possible sets of matches, not over all disparity maps with respect to the first image
  - Includes an *occlusion term*
  - The smoothness term looks different and more complicated

V. Kolmogorov and R. Zabih,
Computing Visual Correspondence with Occlusions using Graph Cuts, ICCV 2001

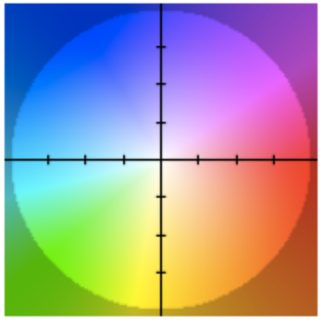
Source: S. Lazebnik

Optical flow estimation for stereo





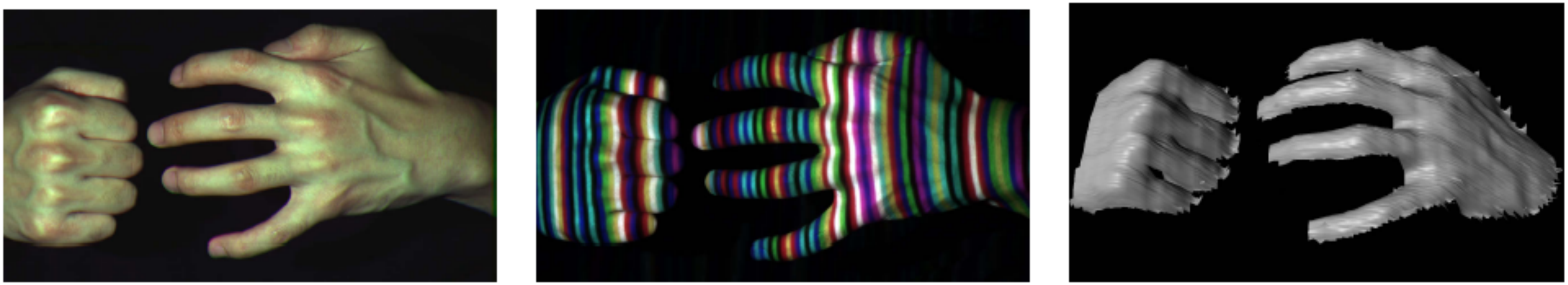
Source: http://people.csail.mit.edu/celiu/OpticalFlow/



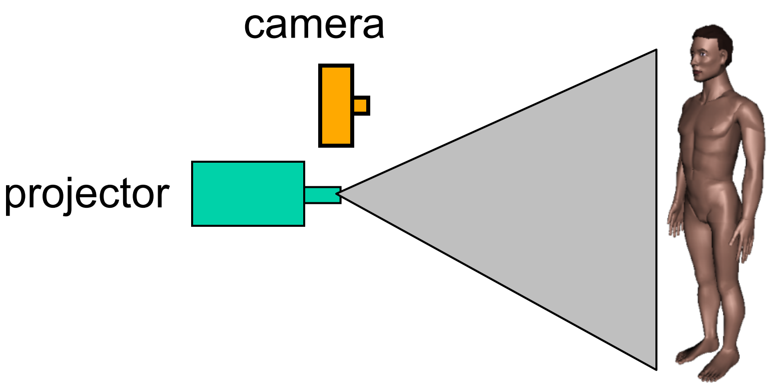
flow color coding

Source: S. Lazebnik

Active stereo with structured light



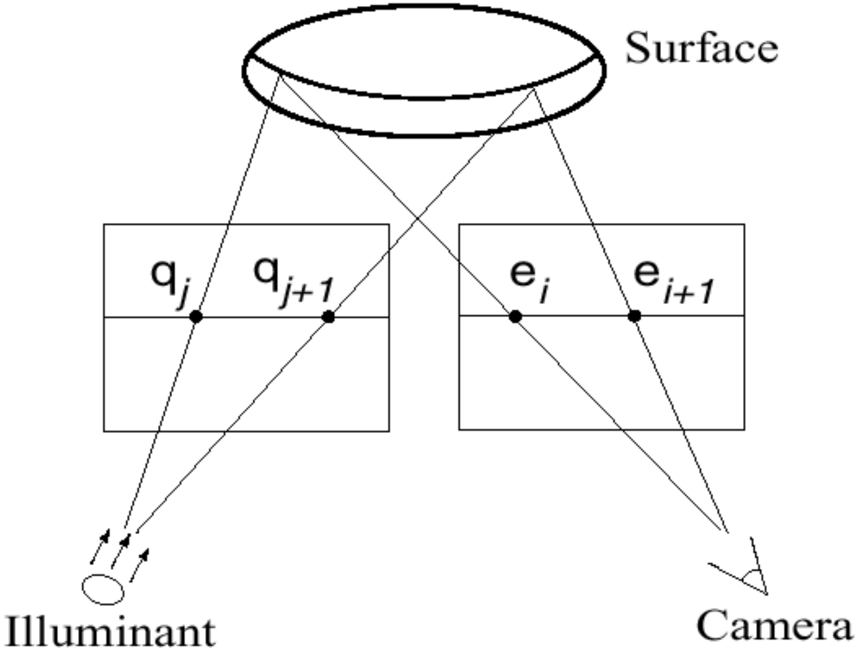
- Project “structured” light patterns onto the object
  - Simplifies the correspondence problem
  - Allows us to use only one camera

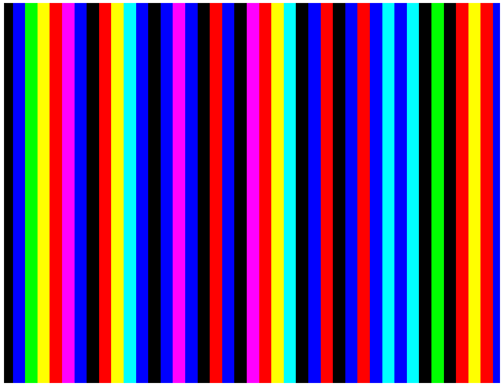


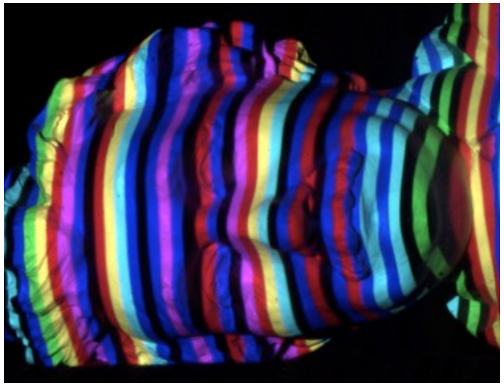
L. Zhang, B. Curless, and S. M. Seitz. [Rapid Shape Acquisition Using Color Structured Light and Multi-pass Dynamic Programming.]() *3DPVT* 2002

Source: S. Lazebnik

Active stereo with structured light



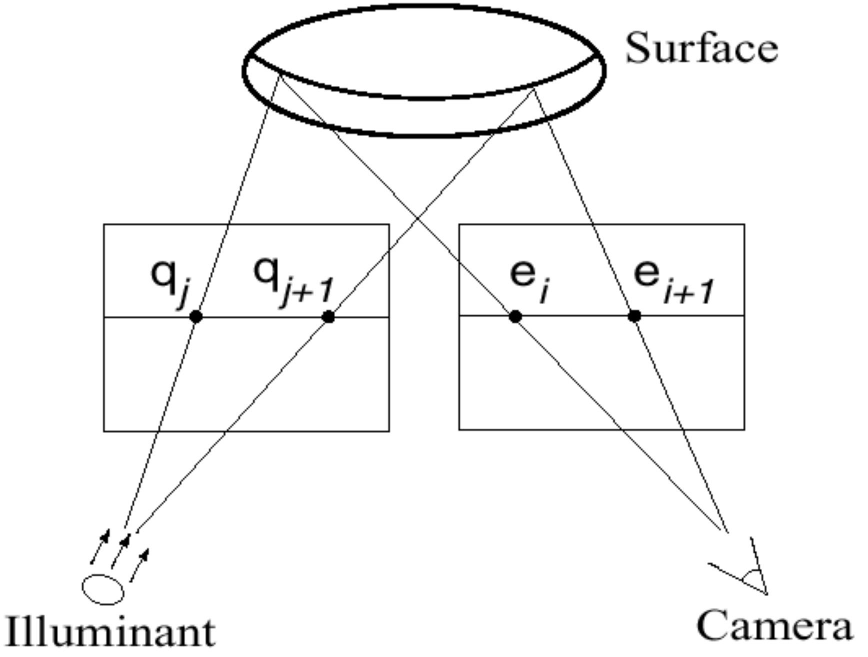


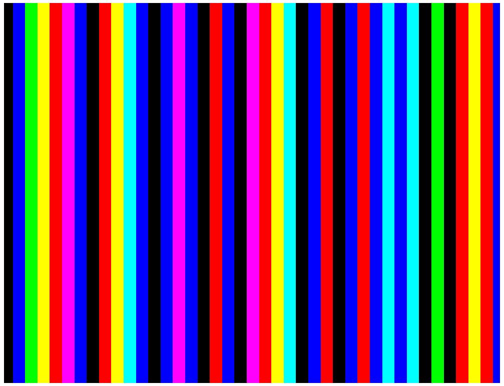


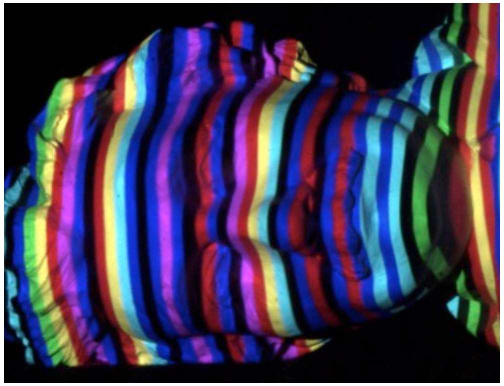
L. Zhang, B. Curless, and S. M. Seitz. Rapid Shape Acquisition Using Color Structured Light and Multi-pass Dynamic Programming. *3DPVT* 2002

Source: S. Lazebnik

Active stereo with structured light



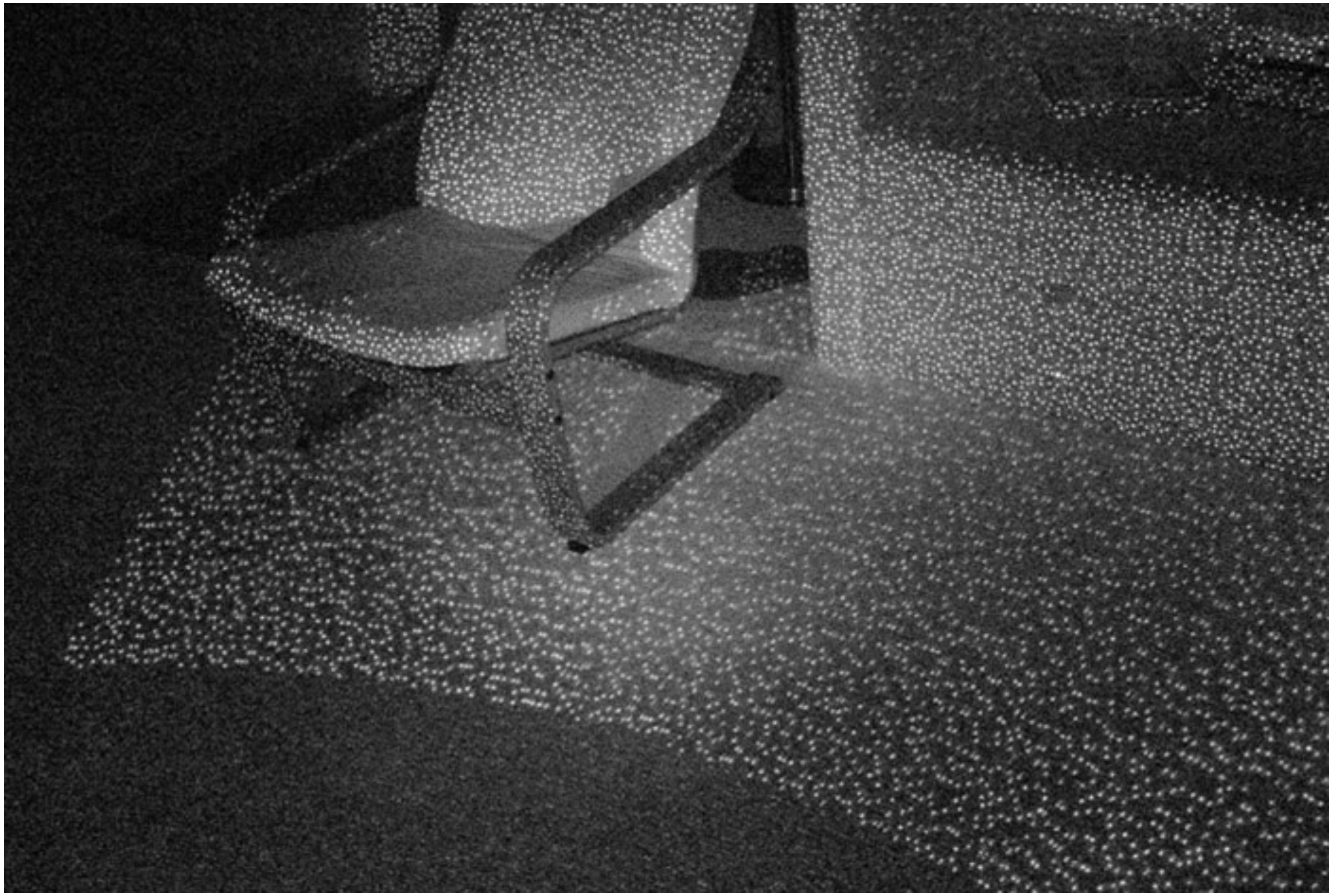




http://en.wikipedia.org/wiki/Structured-light_3D_scanner

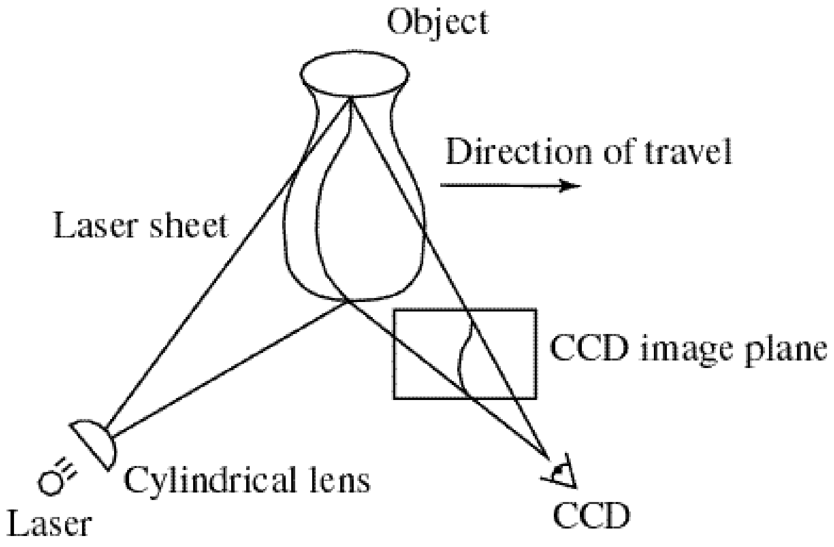
Source: S. Lazebnik

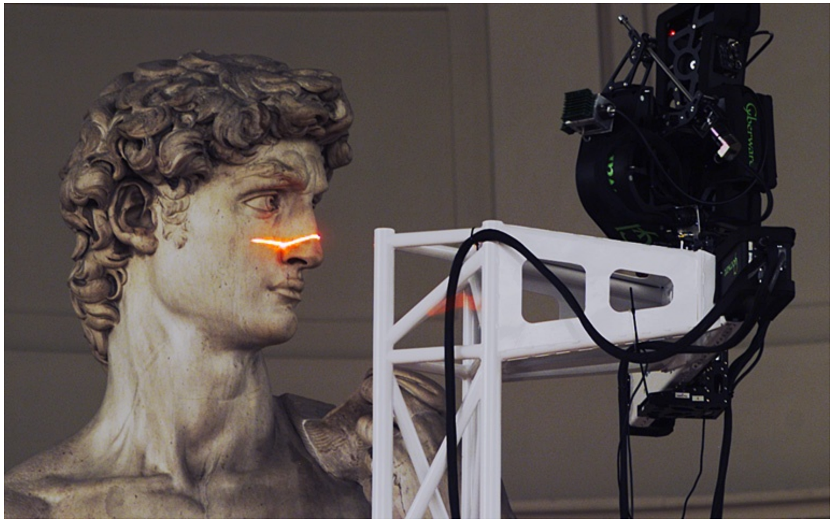
Kinect: Structured infrared light



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

Laser scanning





Digital Michelangelo Project
Levoy et al.
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

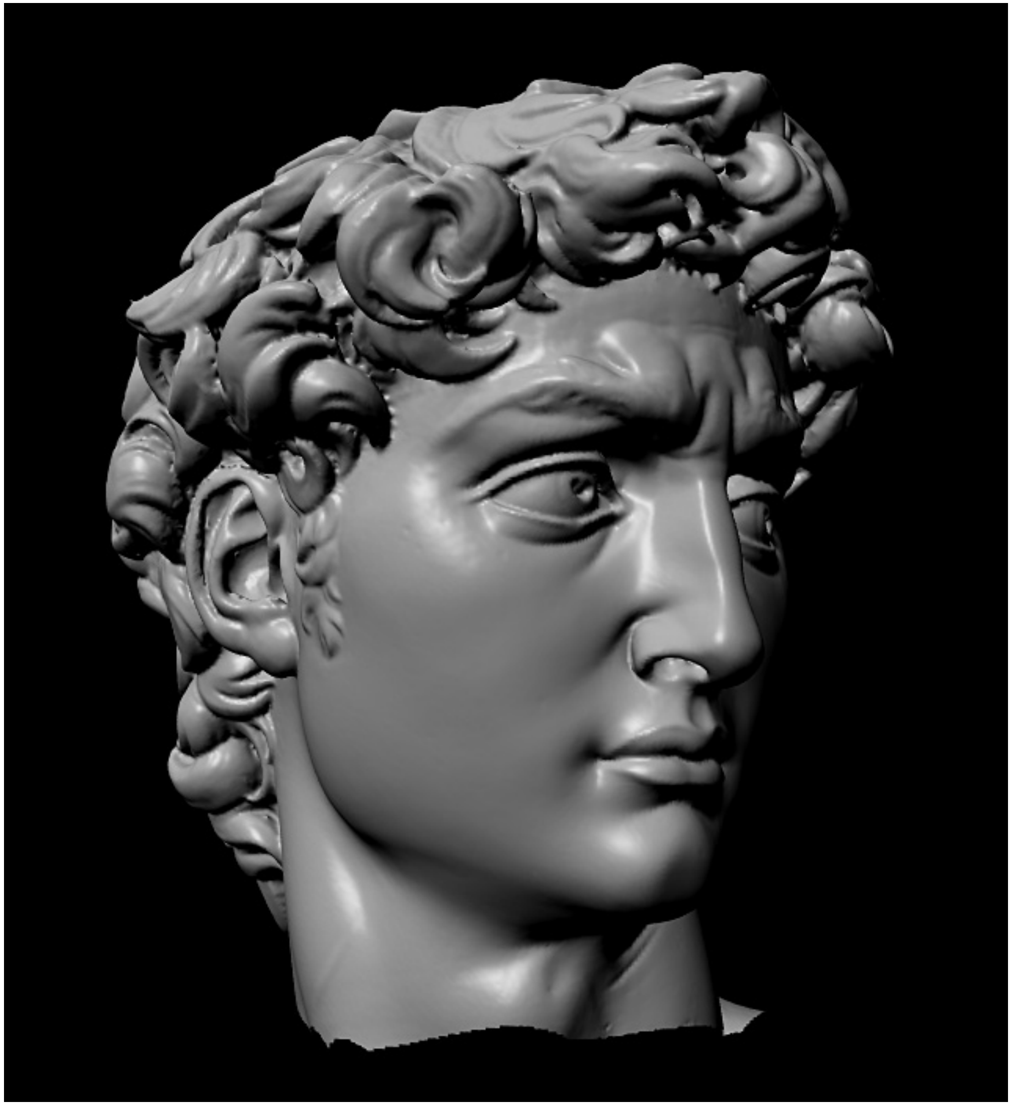
Source: S. Seitz

Laser scanned models

The Digital Michelangelo Project, Levoy et al.

Source: S. Seitz

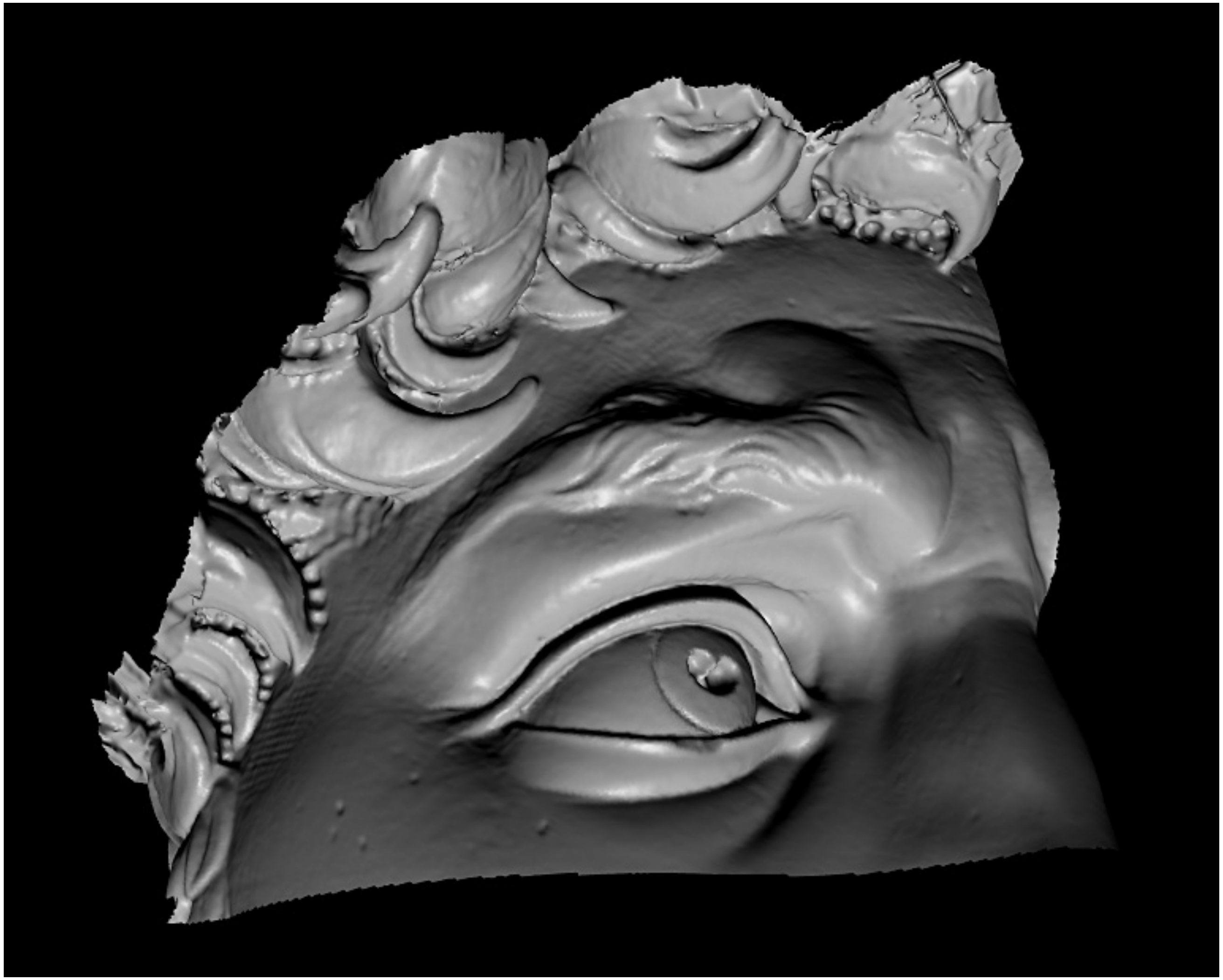
Laser scanned models



The Digital Michelangelo Project, Levoy et al.

Source: S. Seitz

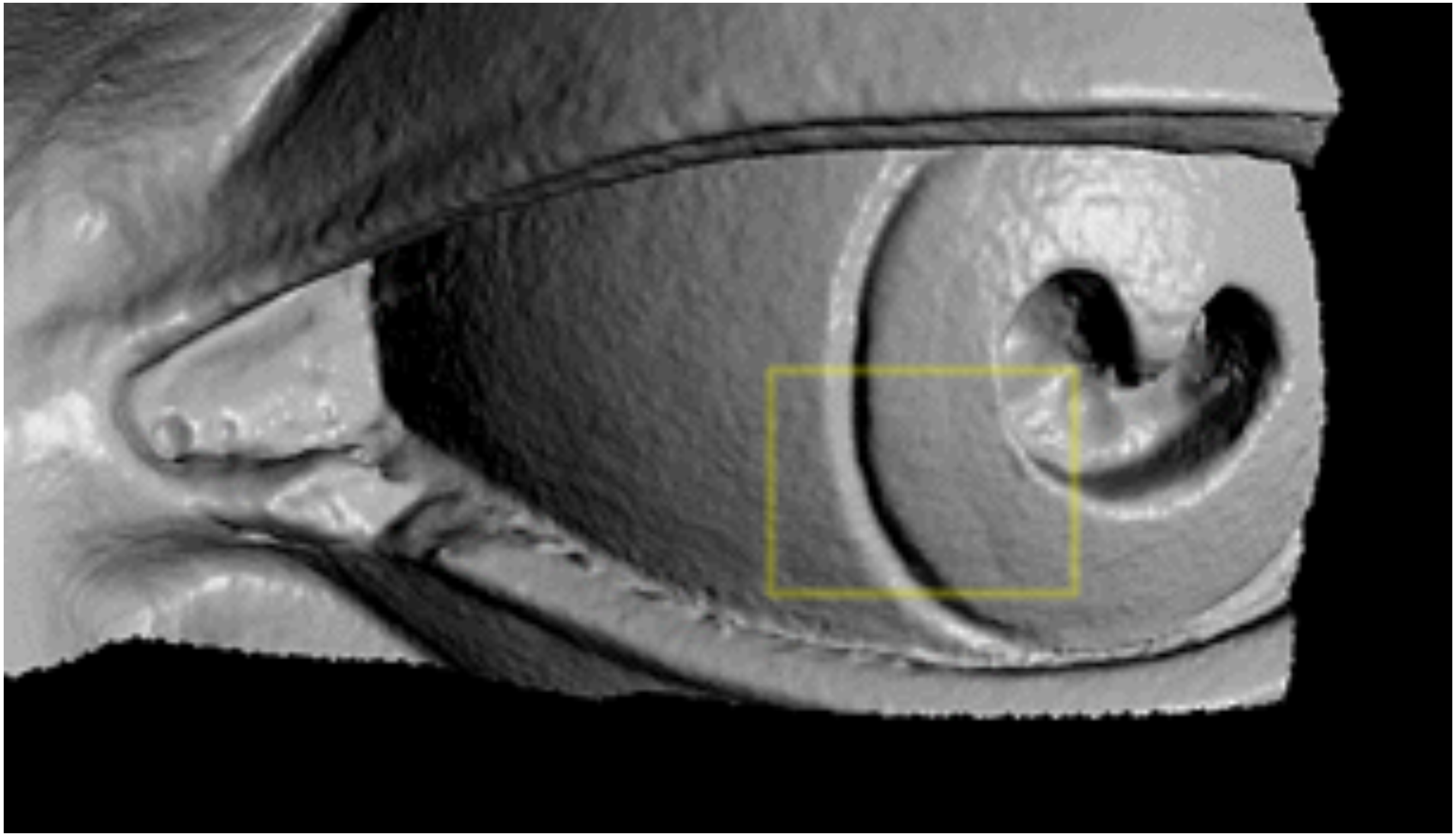
Laser scanned models



The Digital Michelangelo Project, Levoy et al.

Source: S. Seitz

Laser scanned models

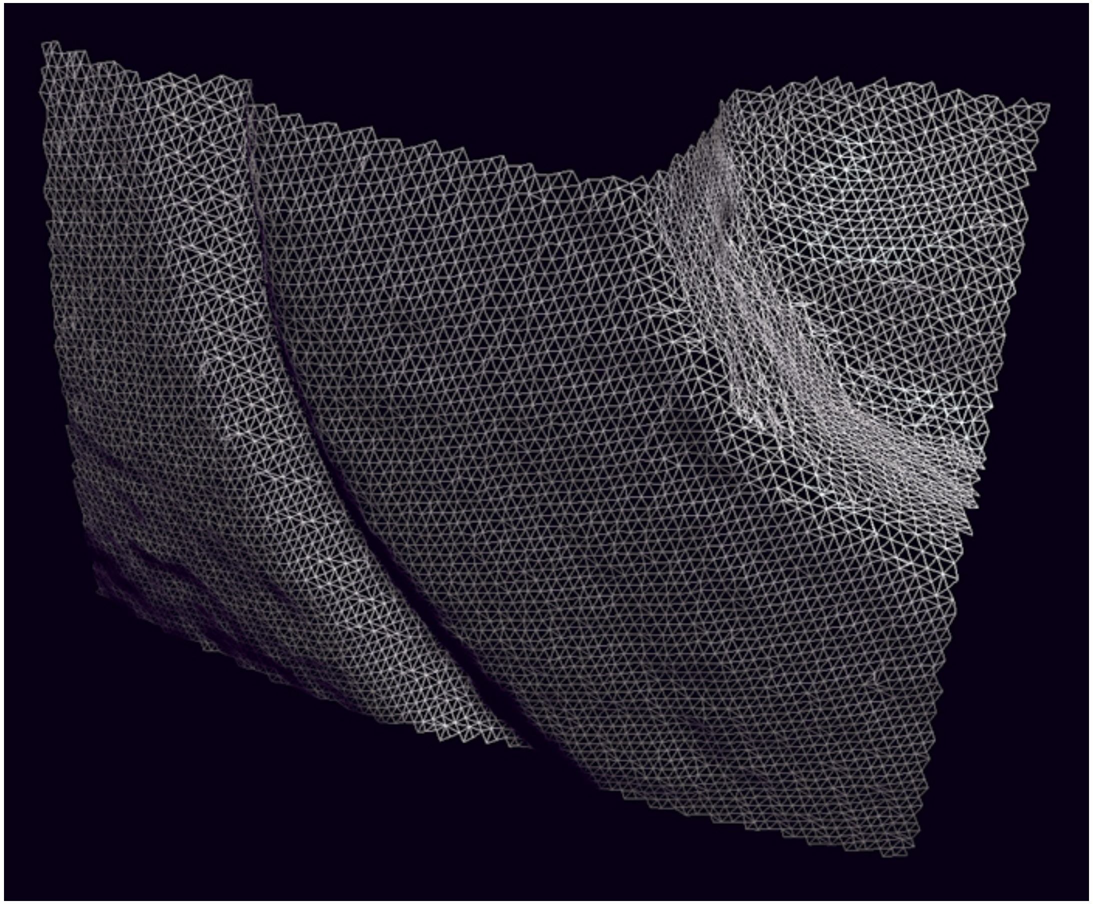


The Digital Michelangelo Project, Levoy et al.

Source: S. Seitz

Laser scanned models

1.0 mm resolution (56 million triangles)

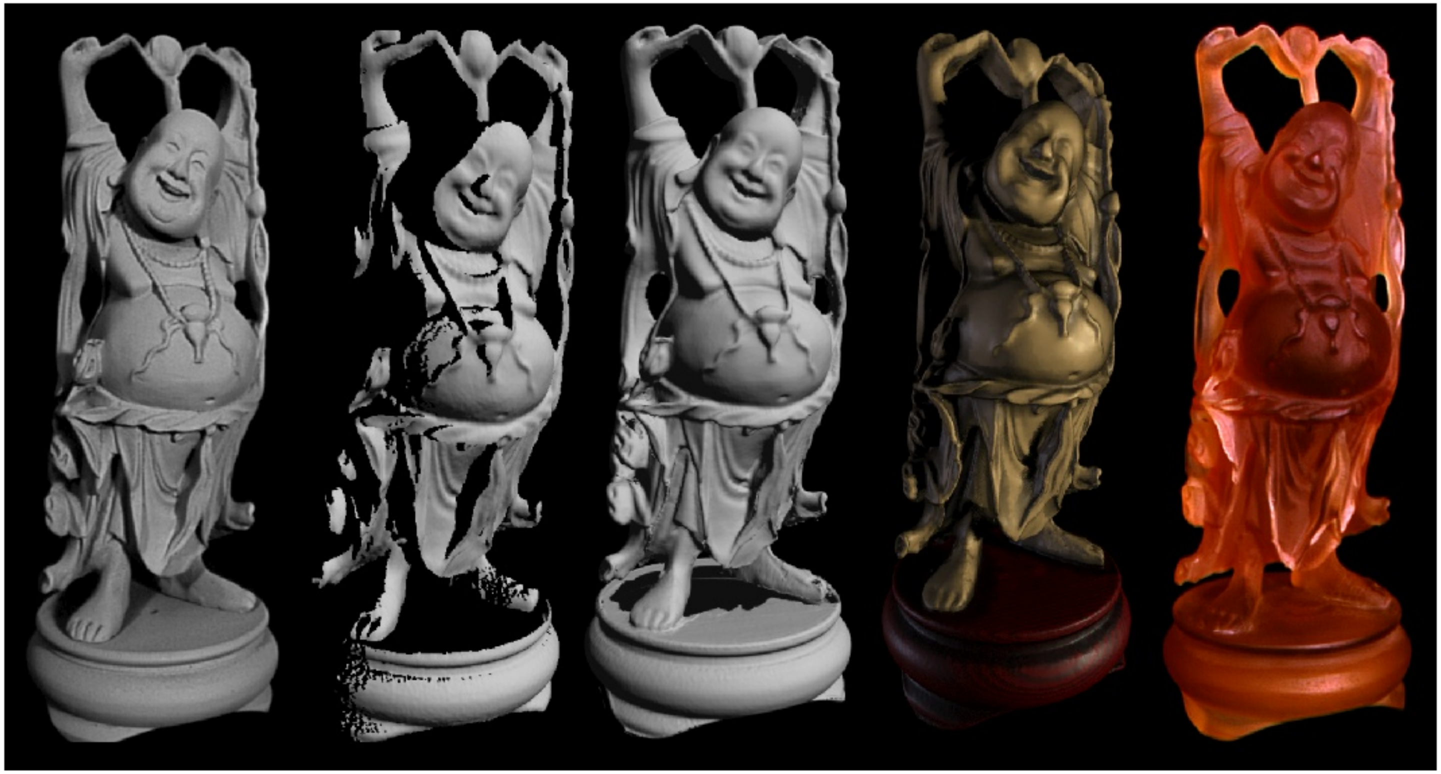


The Digital Michelangelo Project, Levoy et al.

Source: S. Seitz

Aligning range images

- A single range scan is not sufficient to describe a complex surface
- Need techniques to register multiple range images



B. Curless and M. Levoy,
A Volumetric Method for Building Complex Models from Range Images, SIGGRAPH 1996

Source: S. Lazebnik

Aligning range images

- A single range scan is not sufficient to describe a complex surface
- Need techniques to register multiple range images

… which brings us to *multi-view stereo*