THE UNIVERSITY OF TEXAS AT DALLAS

Epipolar Geometry and Stereo

CS 4391 Introduction to Computer Vision

Professor Yapeng Tian
Department of Computer Science

Slides borrowed from Professor Yu Xiang

Depth Perception

Metric

- The car is 10 meters away

Ordinary

- The tree is behind the car

Depth Cues

Information for sensory stimulation that is relevant to depth perception

Monocular cues: single eye

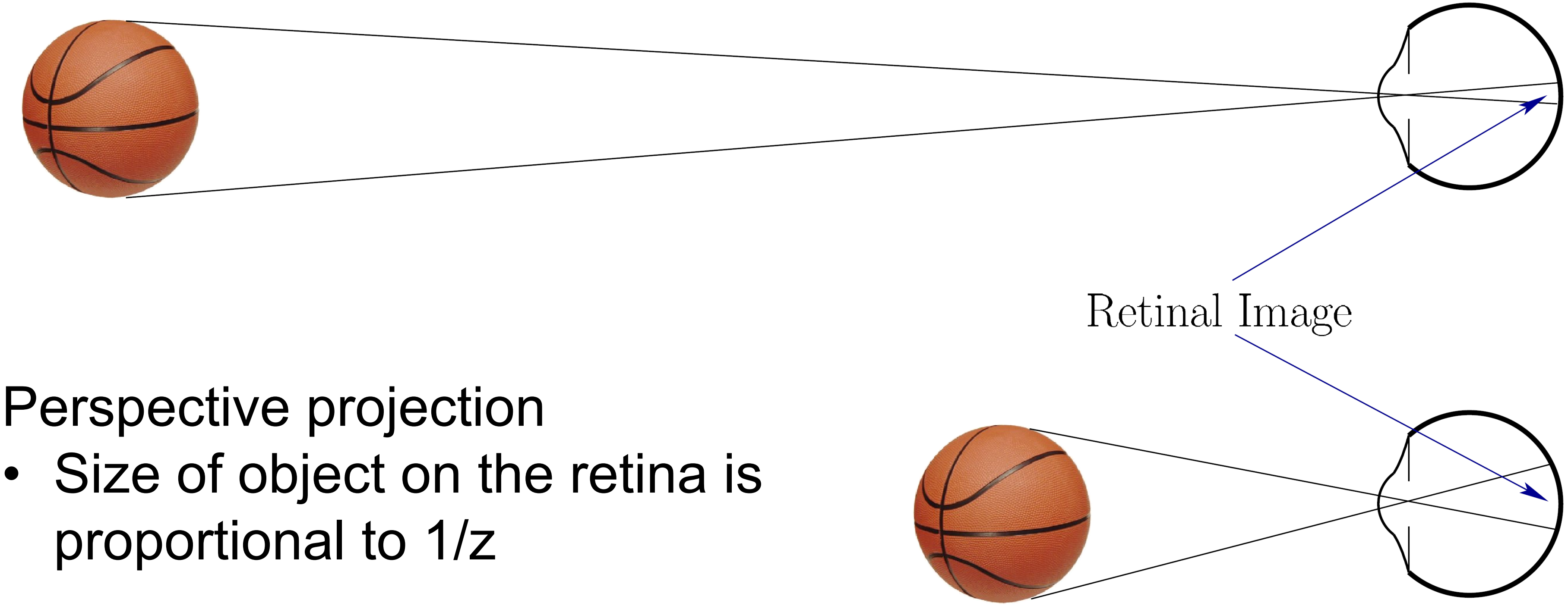
Stereo cues: both eyes

"Paris Street, Rainy Day," Gustave Caillebotte, 1877. Art Institute of Chicago

- Texture of the bricks
- Perspective projection
- Etc.

Monocular Depth Cues

Retinal image size



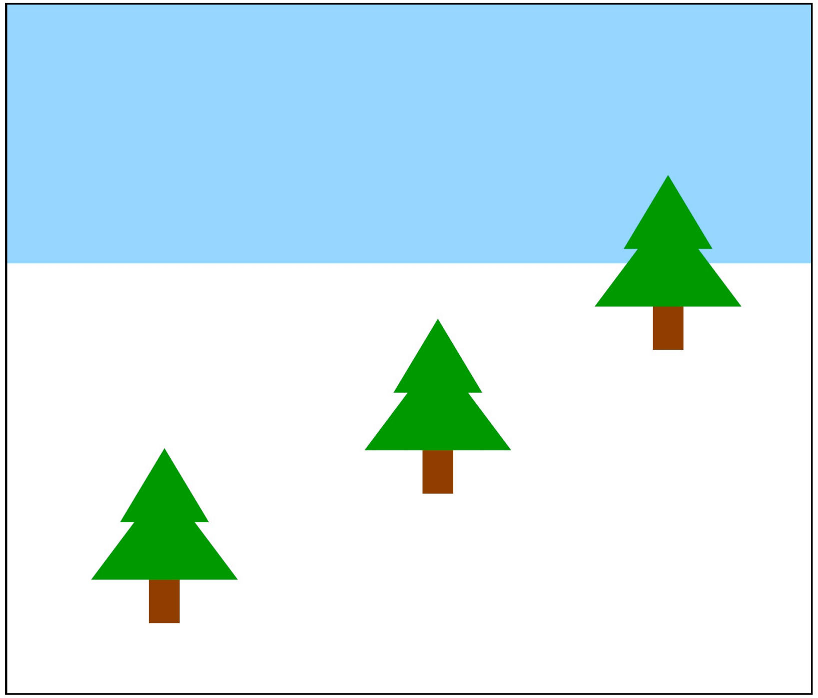
Perspective projection

- Size of object on the retina is proportional to $1/z$

Monocular Depth Cues

Height in visual field

- The closer to the horizon, the further the perceived distance

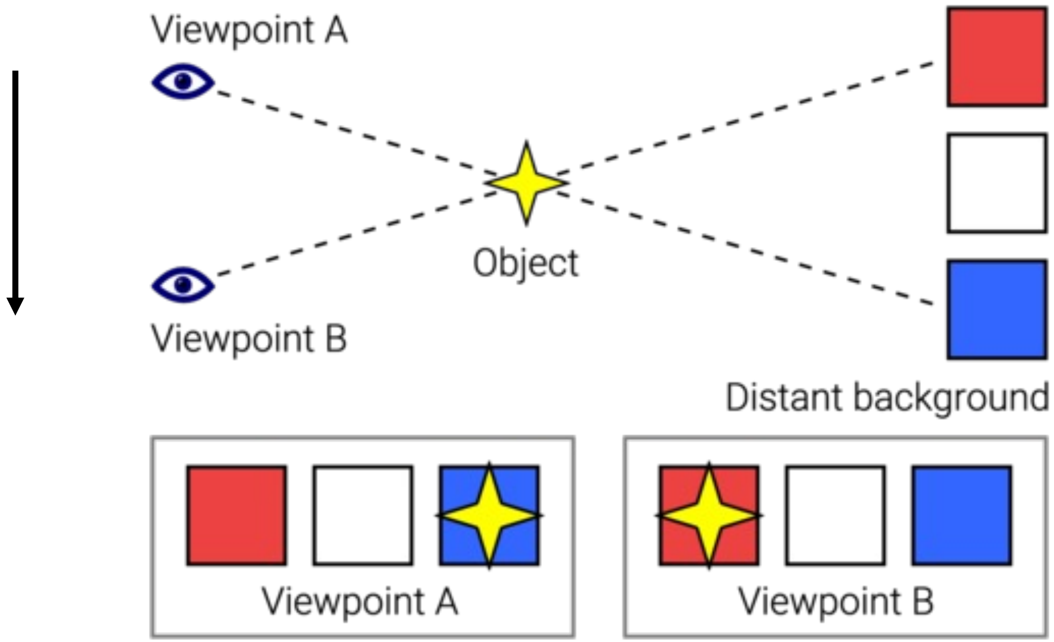


size constancy scaling

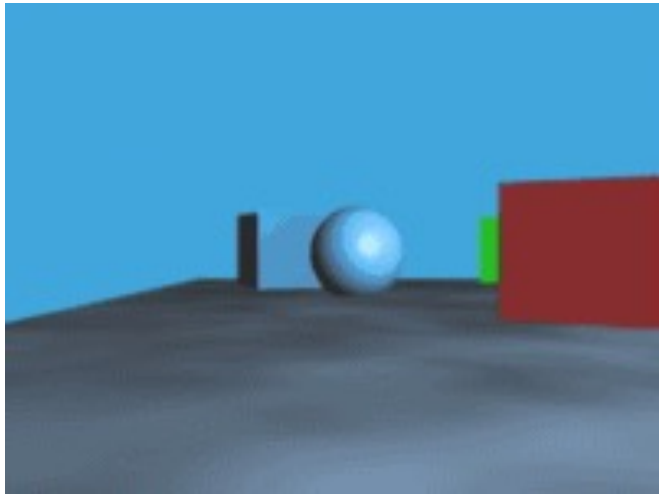
Monocular Depth Cues

Motion parallax

- Parallax: relative difference in speed

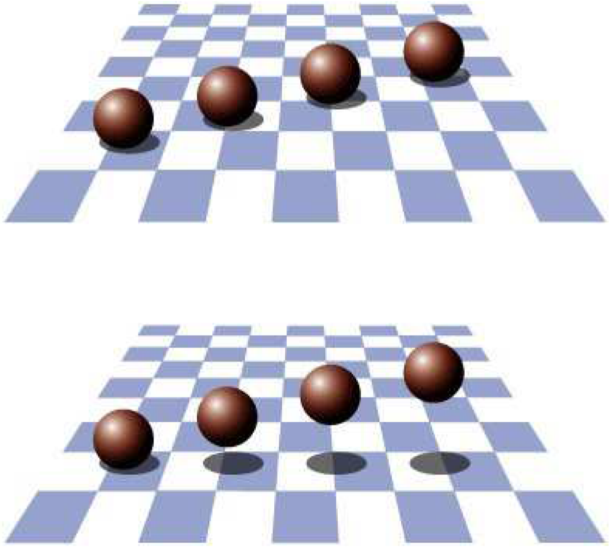


Closer objects have larger image displacements than further objects

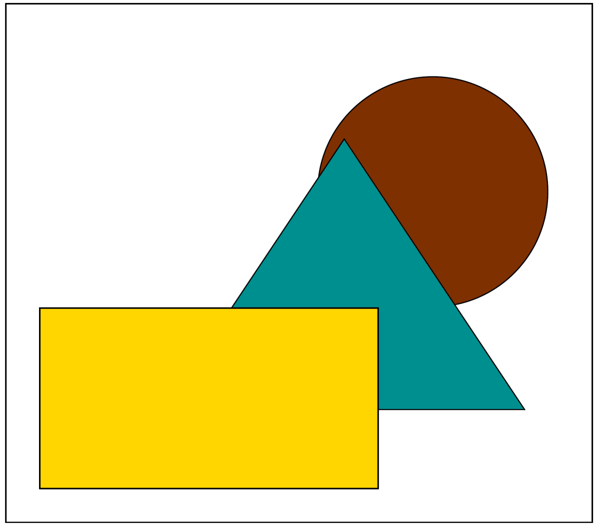


Further objects move slower

Monocular Depth Cues



Shadow



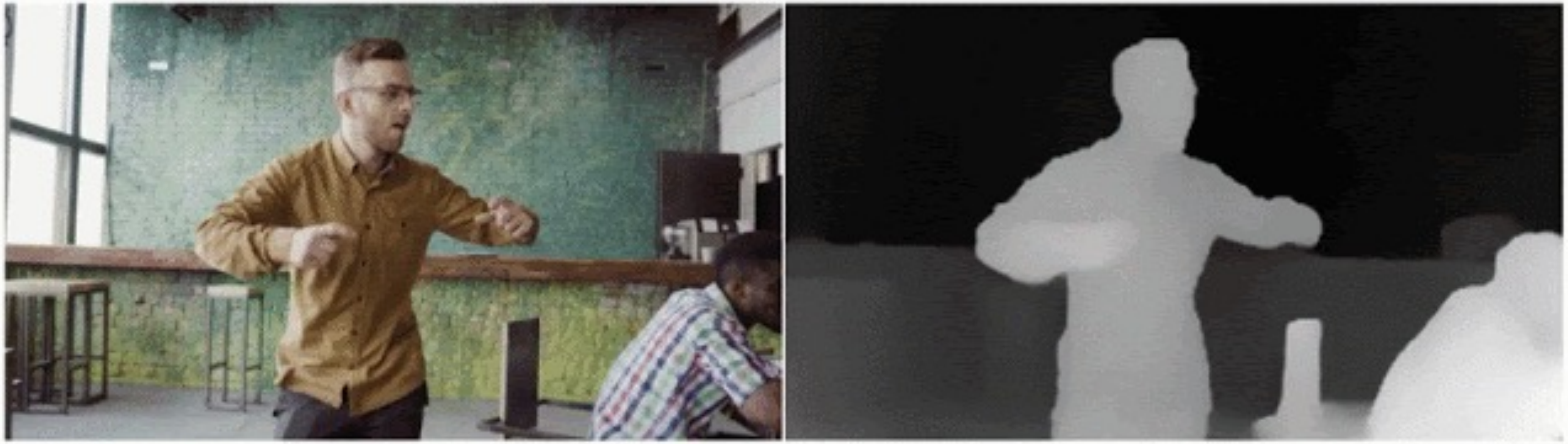
Occlusions

Image blur

Atmospheric cue

further away because it has lower contrast

Monocular Depth Estimation



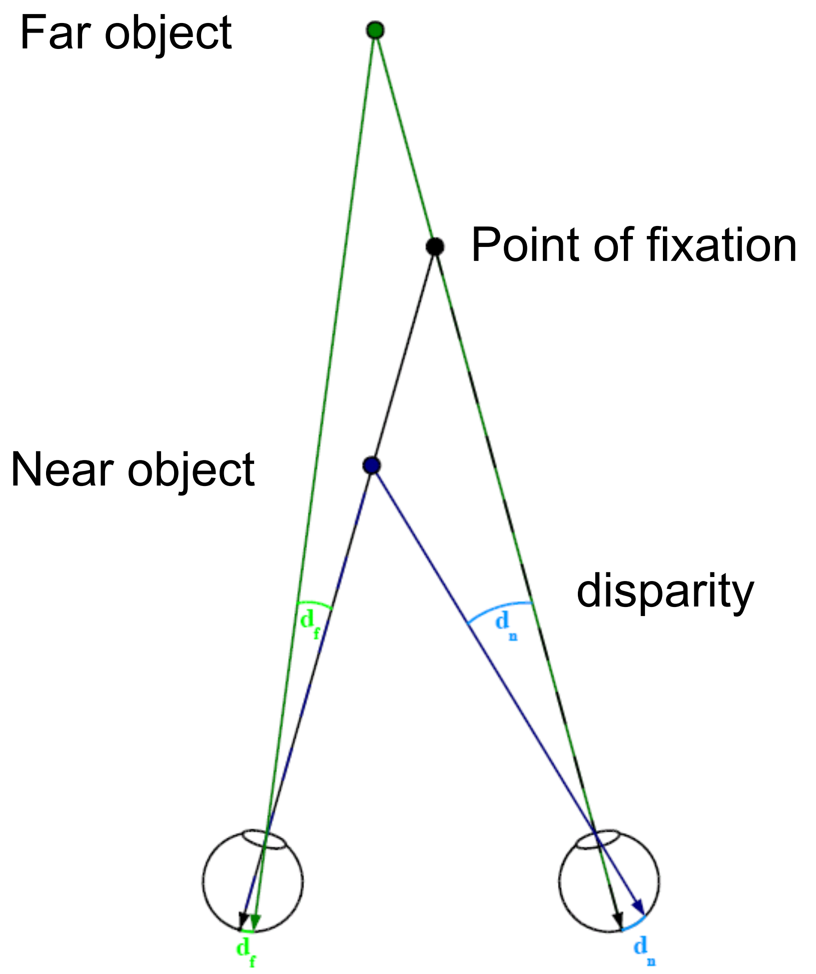
Input video | Our depth predictions

https://heartbeat.fritz.ai/research-guide-for-depth-estimation-with-deep-learning-1a02a439b834

Stereo Depth Cues

Binocular disparity

- Each eye provides a different viewpoint, which results in different images on the retina

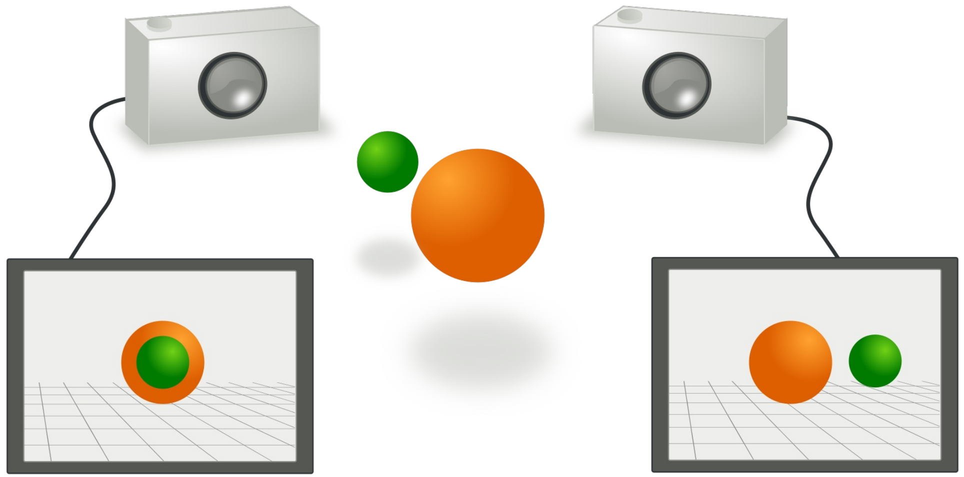


From Wikipedia

Epipolar Geometry

The geometry of stereo vision

- Given 2D images of two views
- What is the relationship between pixels of the images?
- Can we recover the 3D structure of the world from the 2D images?



Wikipedia

Geometry of Stereo Vision

Basics: points and lines

Homogeneous representation of lines

A line in a 2D plane $ax + by + c = 0$ $(a, b, c)^T$

$k(a, b, c)^T$ represents the same line for nonzero k

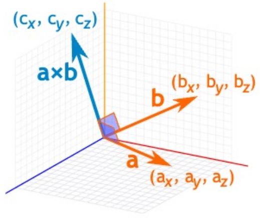
A point lies on the line $\mathbf{x}^T\mathbf{l} = 0$ $\mathbf{x} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$ $\mathbf{l} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$

Points and Lines

https://www.mathsisfun.com/algebra/vectors-cross-product.html

When **a** and **b** start at the origin point (0,0,0), the Cross Product will end at:

- $c_x = a_y b_z - a_z b_y$
- $c_y = a_z b_x - a_x b_z$
- $c_z = a_x b_y - a_y b_x$



Example: The cross product of **a** = (2,3,4) and **b** = (5,6,7)

- $c_x = a_y b_z - a_z b_y = 3×7 - 4×6 = -3$
- $c_y = a_z b_x - a_x b_z = 4×5 - 2×7 = 6$
- $c_z = a_x b_y - a_y b_x = 2×6 - 3×5 = -3$

Answer: **a × b** = (−3,6,−3)

cross product example

Intersection of lines

$$\mathbf{l} = (a, b, c)^T \qquad \mathbf{l}' = (a', b', c')^T$$

The intersection is $\mathbf{x} = \mathbf{l} \times \mathbf{l}'$ (vector cross product)

$$\mathbf{l} \cdot (\mathbf{l} \times \mathbf{l}') = \mathbf{l}' \cdot (\mathbf{l} \times \mathbf{l}') = 0$$

$$\mathbf{l}^T \mathbf{x} = \mathbf{l}'^T \mathbf{x} = 0$$

Points and Lines

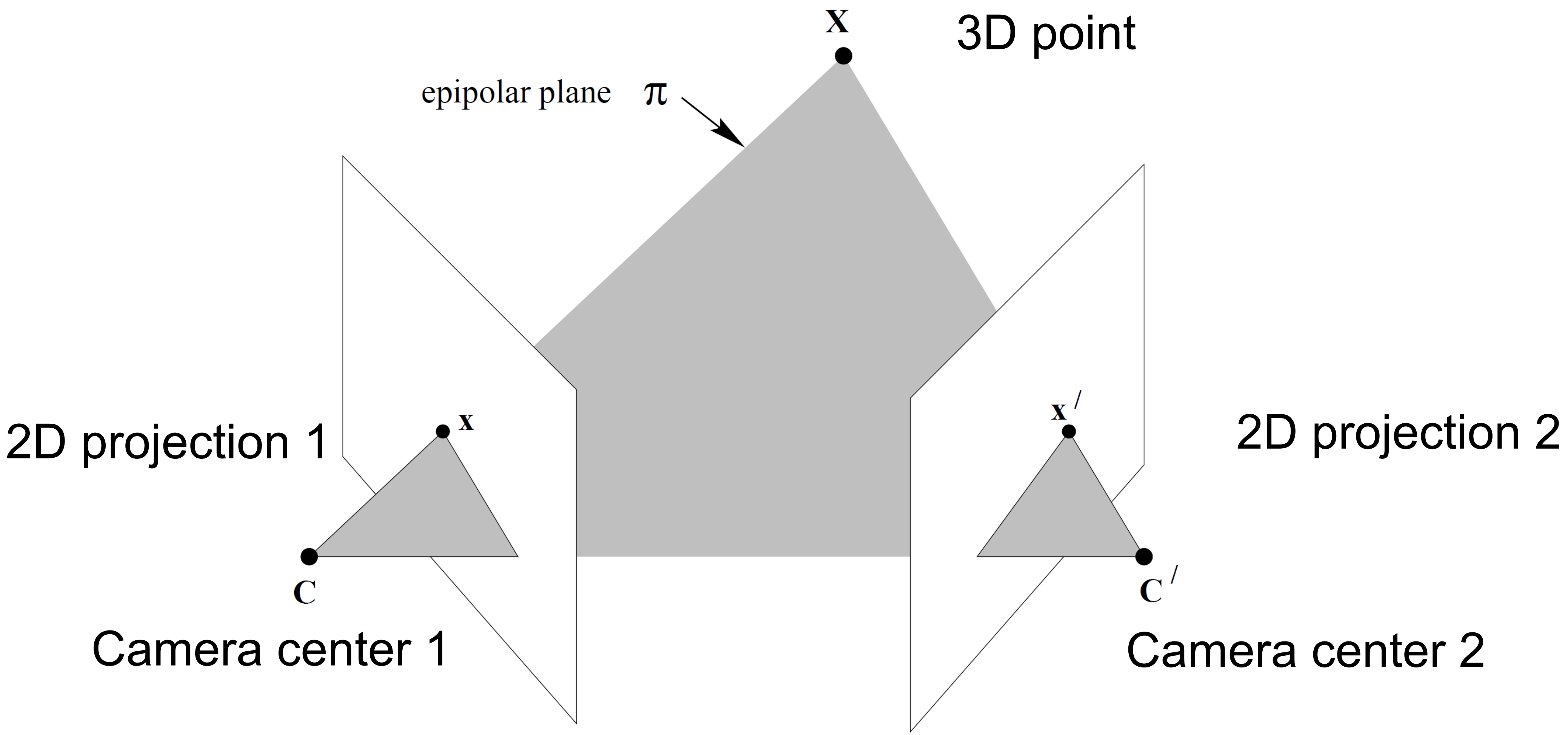
Line joining points

$$\mathbf{l} = \mathbf{x} \times \mathbf{x}'$$

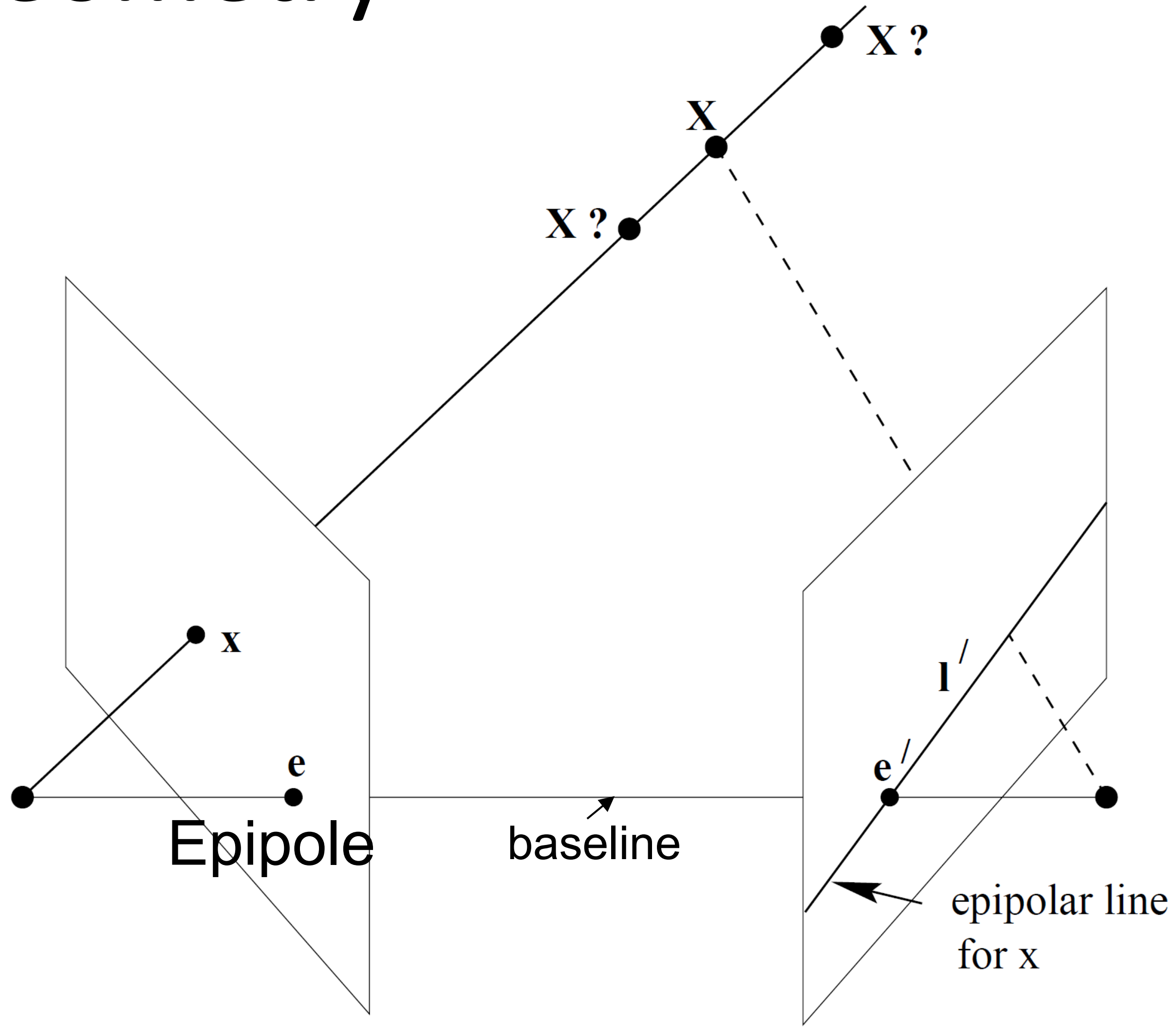
$$\mathbf{x} \cdot (\mathbf{x} \times \mathbf{x}') = \mathbf{x}' \cdot (\mathbf{x} \times \mathbf{x}') = 0$$

$$\mathbf{x}^T \mathbf{l} = \mathbf{x}'^T \mathbf{l} = 0$$

Epipolar Geometry



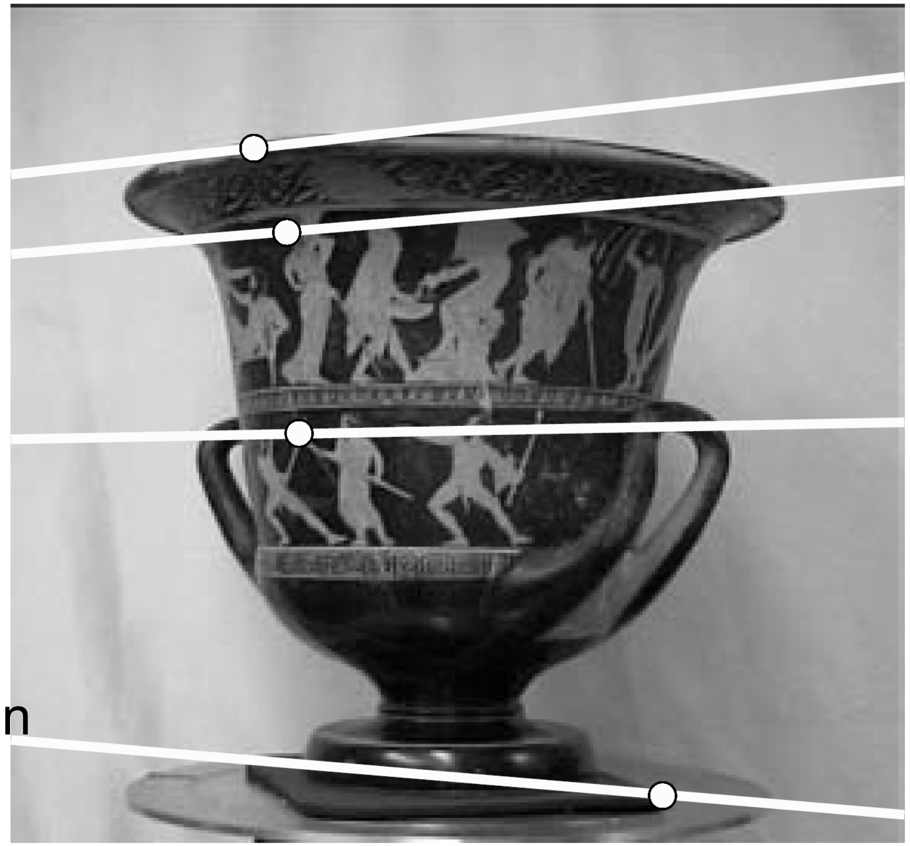
Epipolar Geometry



Epipolar Geometry

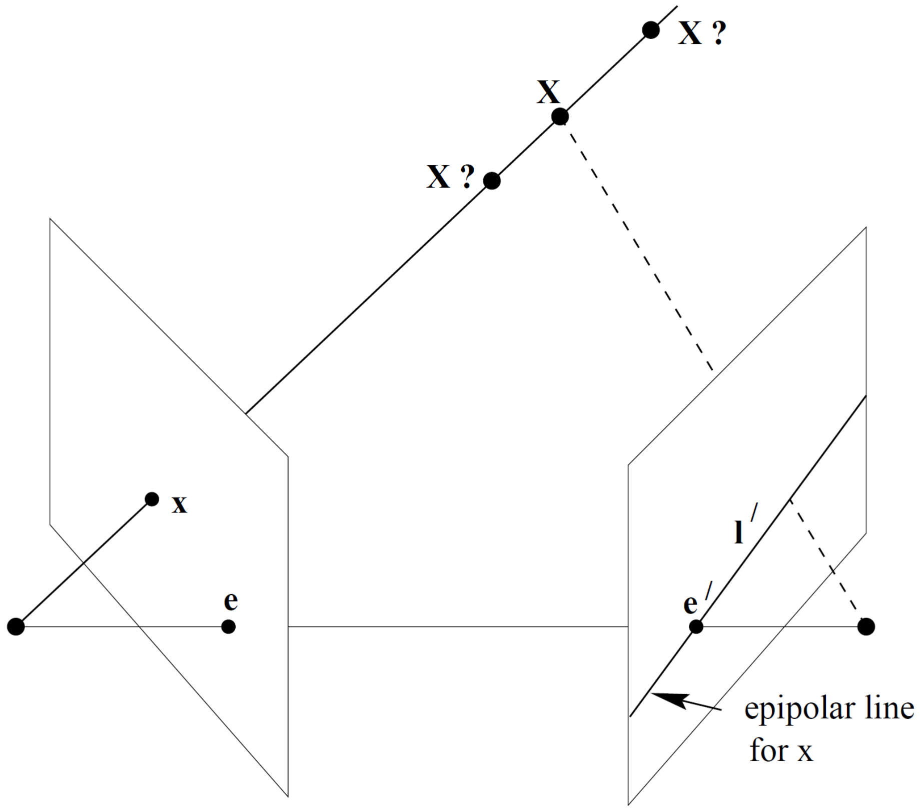
Epipolar lines

Rotation and Translation between two views



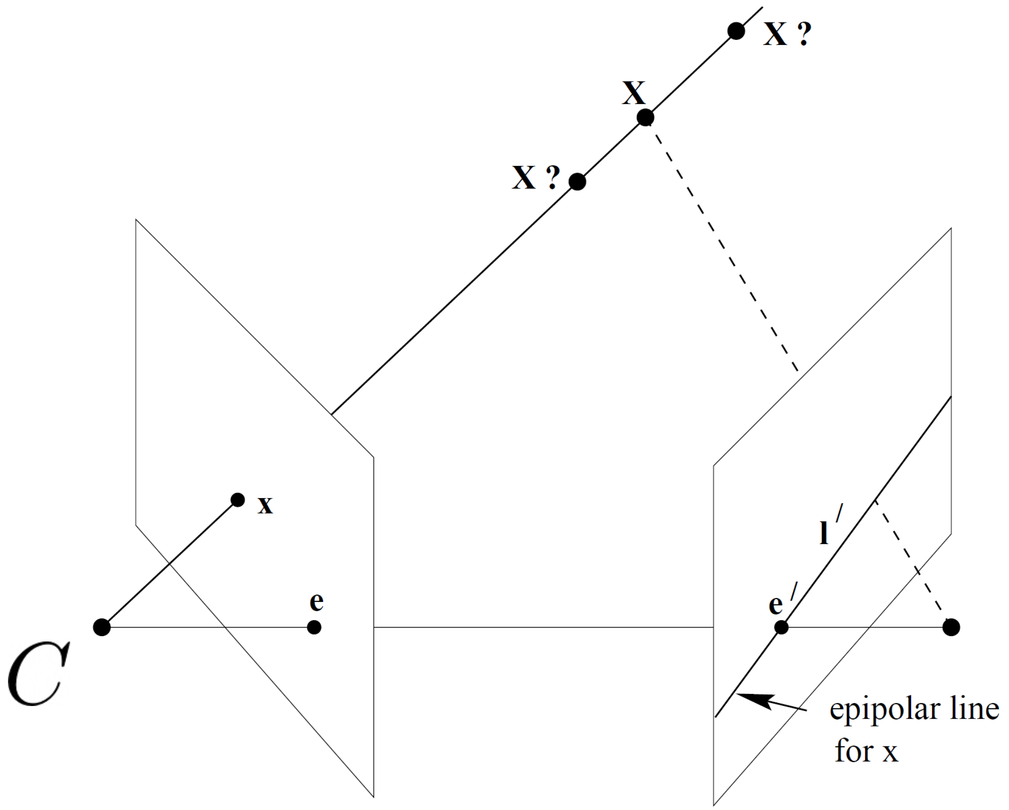
Epipolar Geometry

What is the mapping for a point in one image to its epipolar line?



$$\mathbf{x} \mapsto \mathbf{l}'$$

Fundamental Matrix



- Recall camera projection

$$P = K[R|\mathbf{t}]$$

$$\mathbf{x} = P\mathbf{X}$$ Homogeneous coordinates

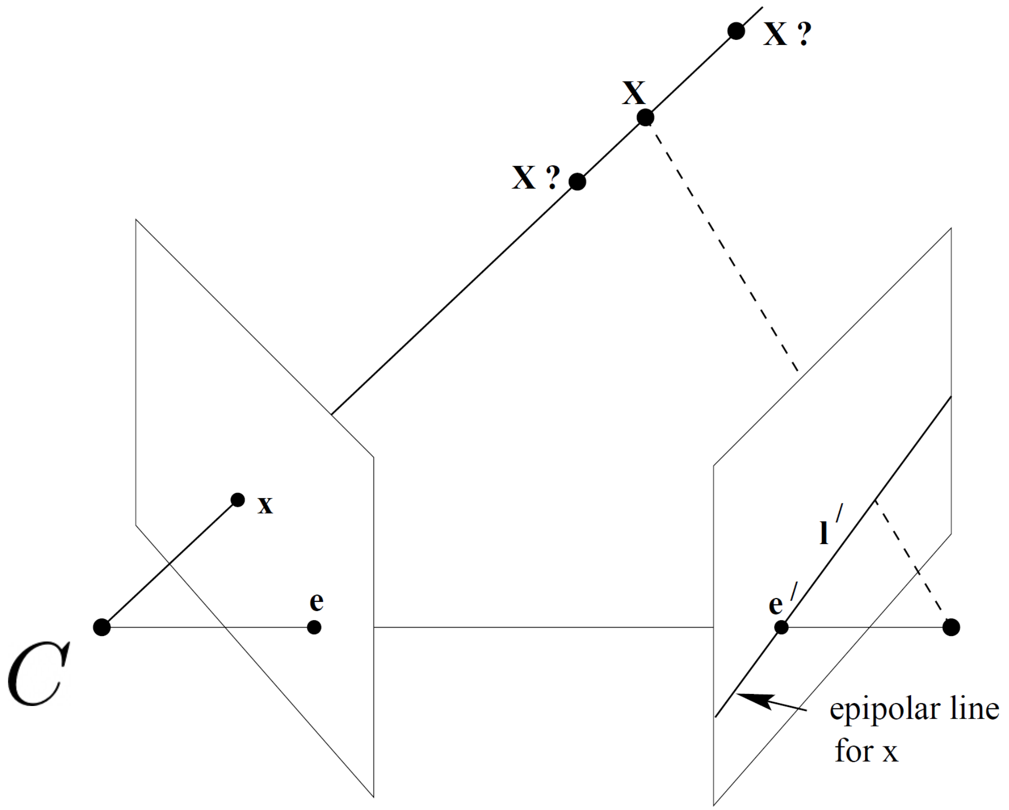
- Backprojection

$$\mathbf{X}(\lambda) = \mathrm{P}^{+}\mathbf{x} + \lambda\mathbf{C}$$

P^{+} is the pseudo-inverse of P, $PP^{+} = I$

$P^{+}\mathbf{x}$ and C are two points on the ray

Fundamental Matrix



- Project to the other image

$P^+\mathbf{x}$ and C are two points on the ray

$P'P^+\mathbf{x}$ and $P'C$

- Epipolar line

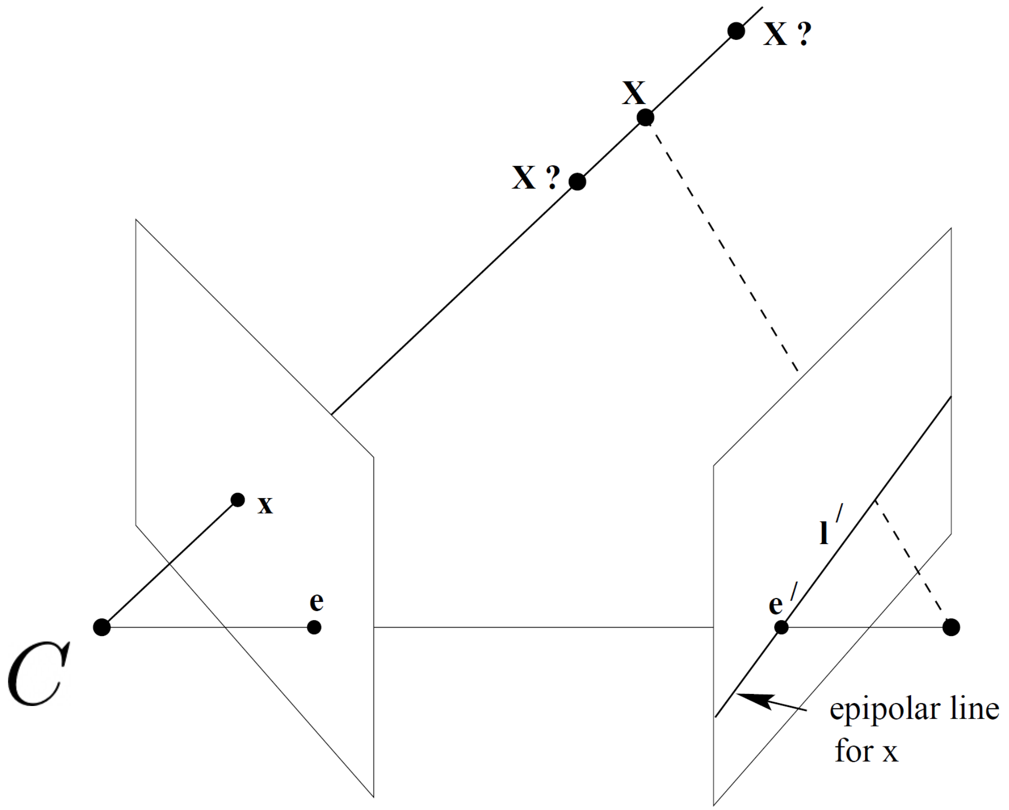
$$\mathbf{l}' = (P'C) \times (P'P^+\mathbf{x})$$

Epipole $\mathbf{e}' = (P'C)$

$$\mathbf{l}' = [\mathbf{e}']_\times (P'P^+\mathbf{x})$$

Cross product matrix

Fundamental Matrix



- Epipolar line

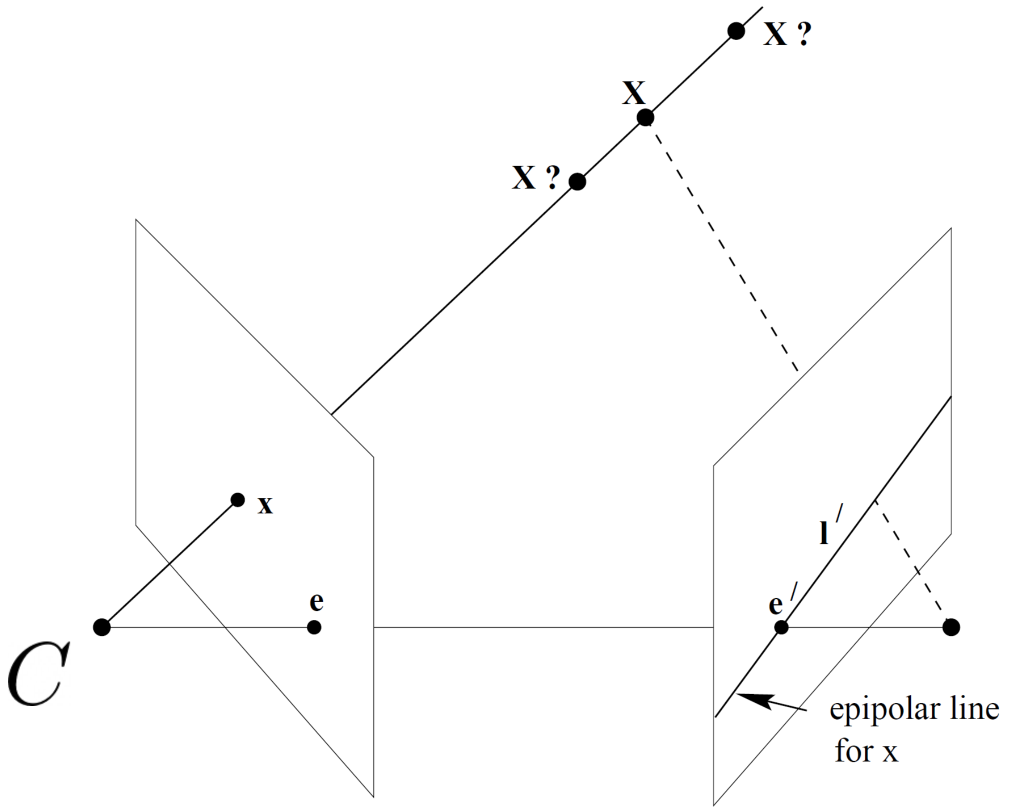
$$\mathbf{l}' = [\mathbf{e}']_\times (P' P^+ \mathbf{x}) = F\mathbf{x}$$

- Fundamental matrix

$$F = [\mathbf{e}']_\times P' P^+$$

3x3

Properties of Fundamental Matrix



$\mathbf{x}'$ is on the epiploar line $\mathbf{l}' = F\mathbf{x}$

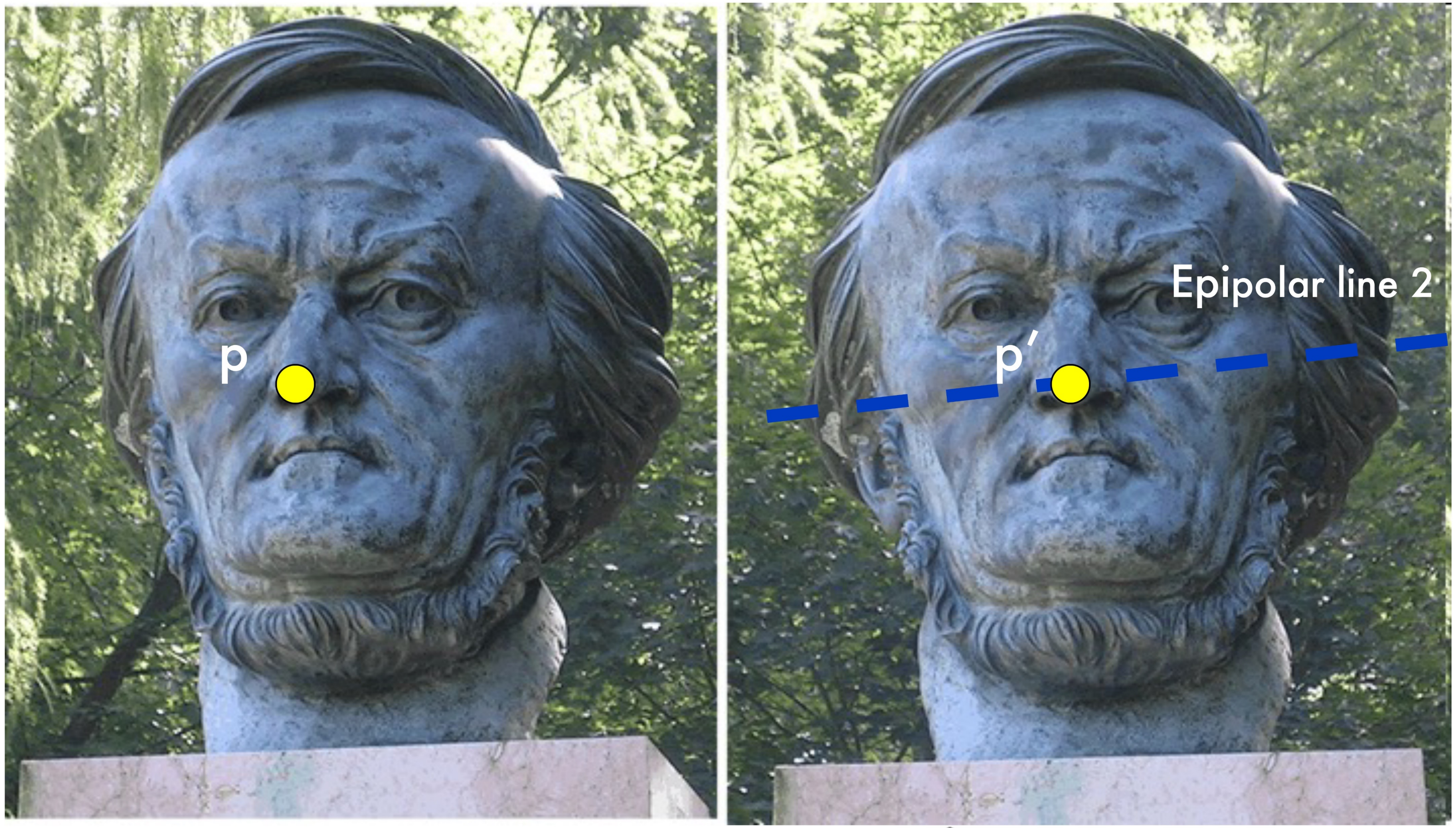
$$\mathbf{x}'^T F \mathbf{x} = 0$$

- Transpose: if F is the fundamental matrix of (P, P'), then F^T is the fundamental matrix of (P', P)
- Epipolar line: $\mathbf{l}' = F\mathbf{x}$ $\quad \mathbf{l} = F^T\mathbf{x}'$
- Epipole: $\mathbf{e}'^{\mathsf{T}}\mathbf{F} = 0 \quad \mathbf{F}\mathbf{e} = \mathbf{0}$

$$\mathbf{e}'^{\mathsf{T}}(\mathbf{F}\mathbf{x}) = (\mathbf{e}'^{\mathsf{T}}\mathbf{F})\mathbf{x} = 0 \text{ for all } \mathbf{x}$$

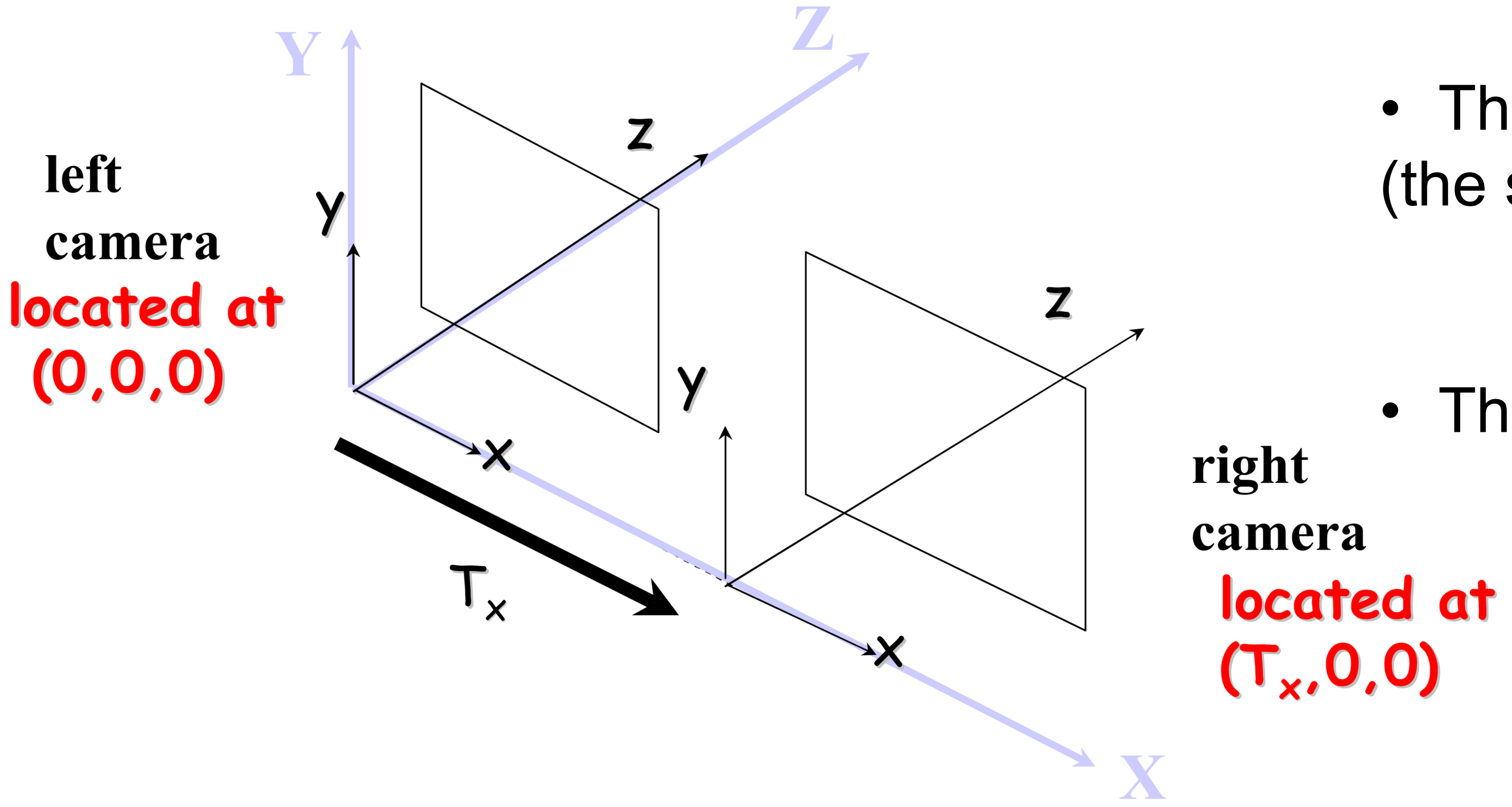
- 7 degrees of freedom $\det \mathbf{F} = 0$

Why the Fundamental Matrix is Useful?



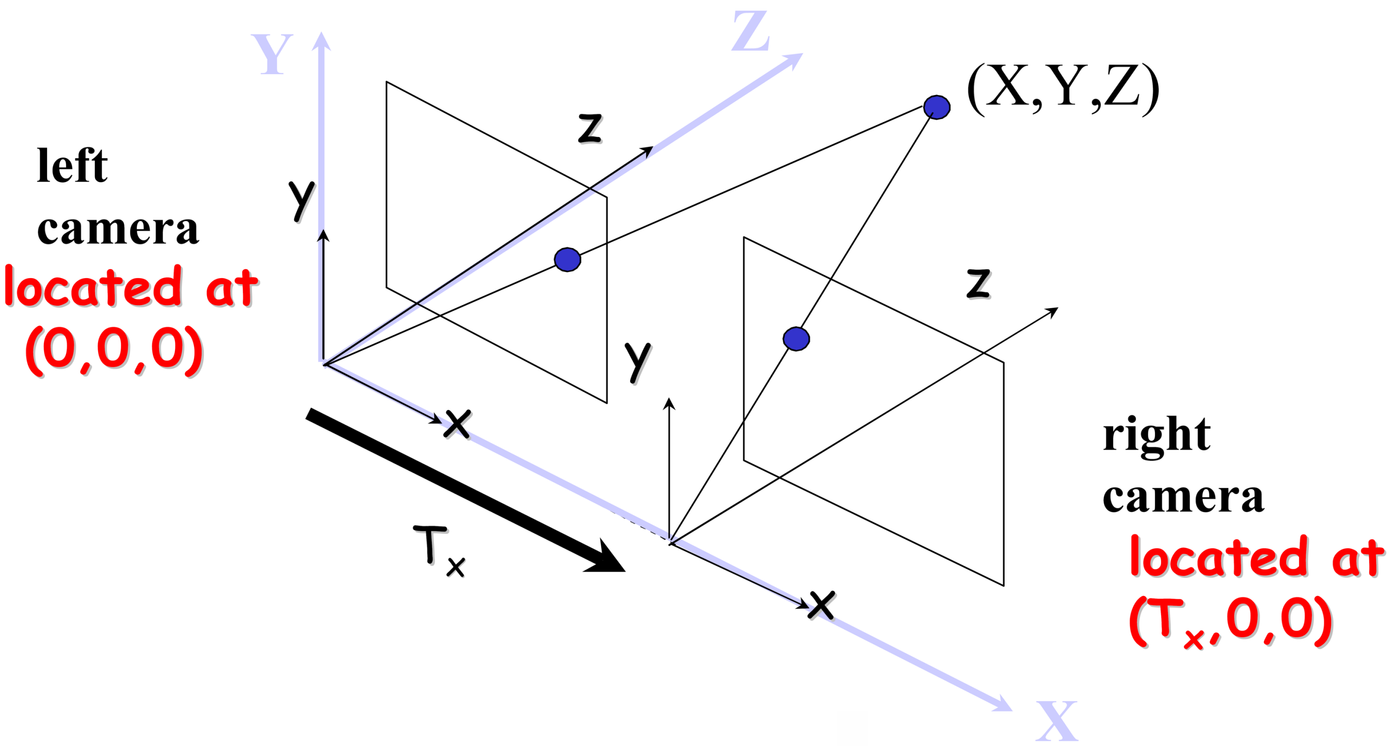
$$\mathbf{l}' = F\mathbf{p}$$

Special Case: A Stereo System



- The right camera is shifted by T_x (the stereo baseline)
- The camera intrinsics are the same

Special Case: A Stereo System



- Left camera

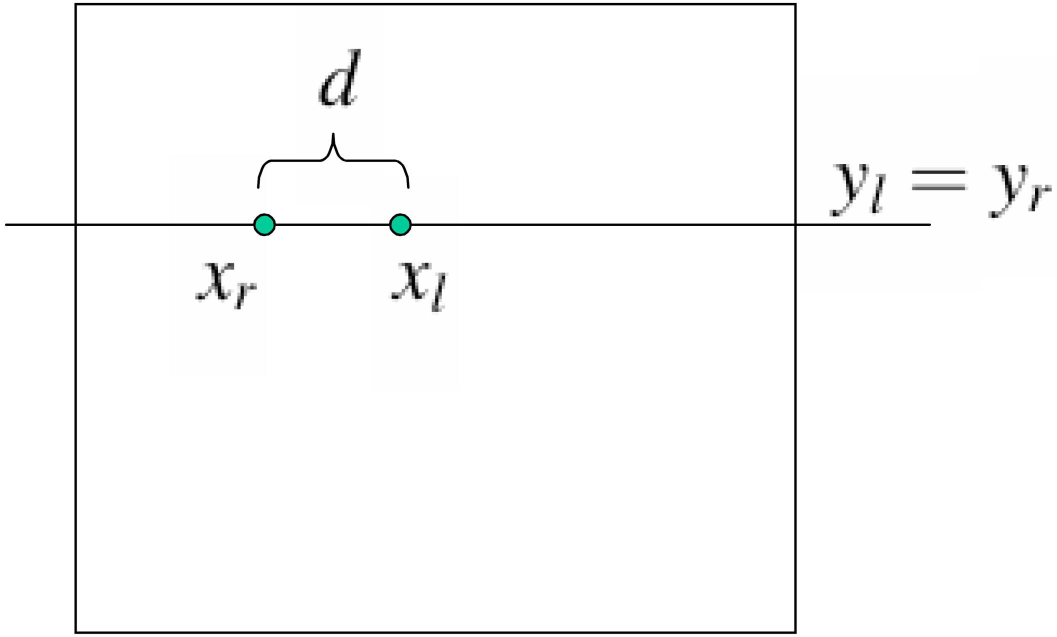
$$x_l = f\frac{X}{Z} + p_x \qquad y_l = f\frac{Y}{Z} + p_y$$

- Right camera

$$x_r = f\frac{X - T_x}{Z} + p_x$$

$$y_r = f\frac{Y}{Z} + p_y$$

Stereo Disparity



- Disparity

$$\begin{aligned} d &= x_l - x_r \\ &= \left(f\frac{X}{Z} + p_x\right) - \left(f\frac{X - T_x}{Z} + p_x\right) \\ &= f\frac{T_x}{Z} \end{aligned}$$

- Depth

$$Z = f\frac{T_x}{d}$$

Baseline (T_x)

Disparity (d)

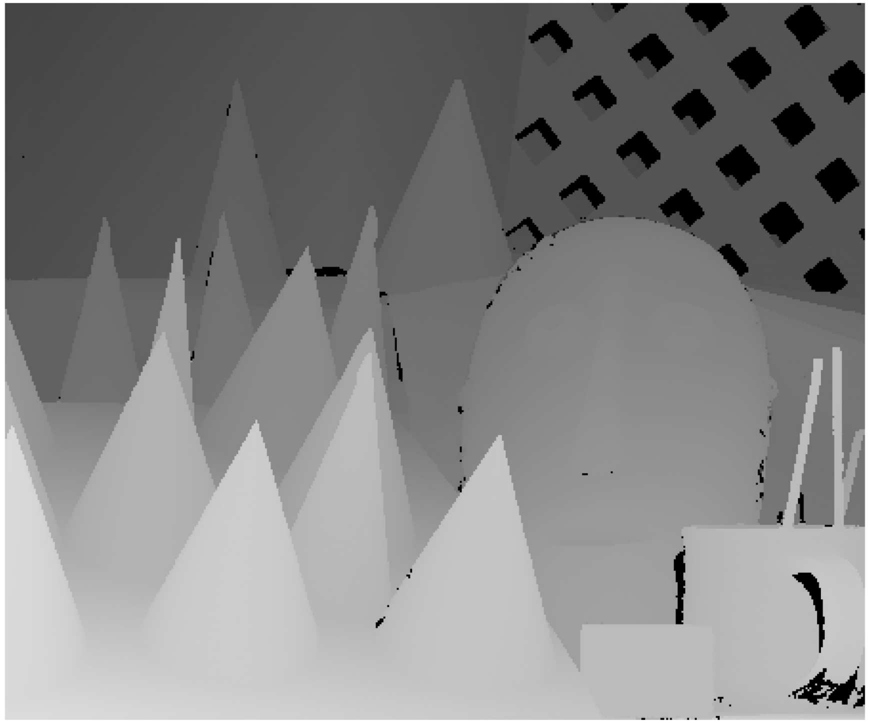
Recall motion parallax: near objects move faster (large disparity)

Stereo Example


Left image


Right image

Disparity values (0-64)



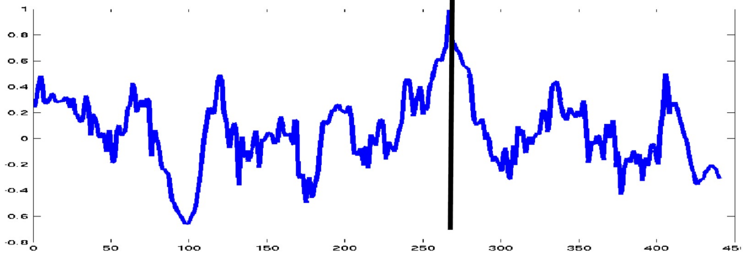
Note how disparity is larger (brighter) for closer surfaces.

$$d = f\frac{T_x}{Z}$$

Computing Disparity

Left Image

Right Image



Match Score Values

For a patch in left image

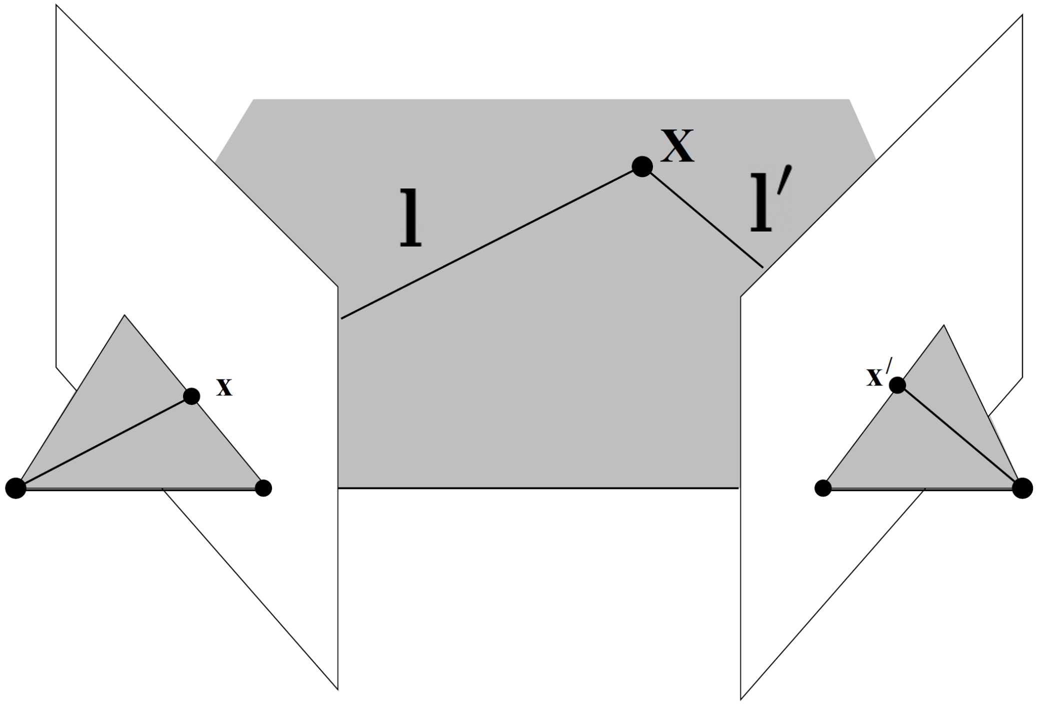
Compare with patches along same row in right image

- Eipipolar lines are horizontal lines in stereo
- For general cases, we can find correspondences on eipipolar lines
- Depth from disparity

$$Z = f\frac{T_x}{d}$$

Triangulation

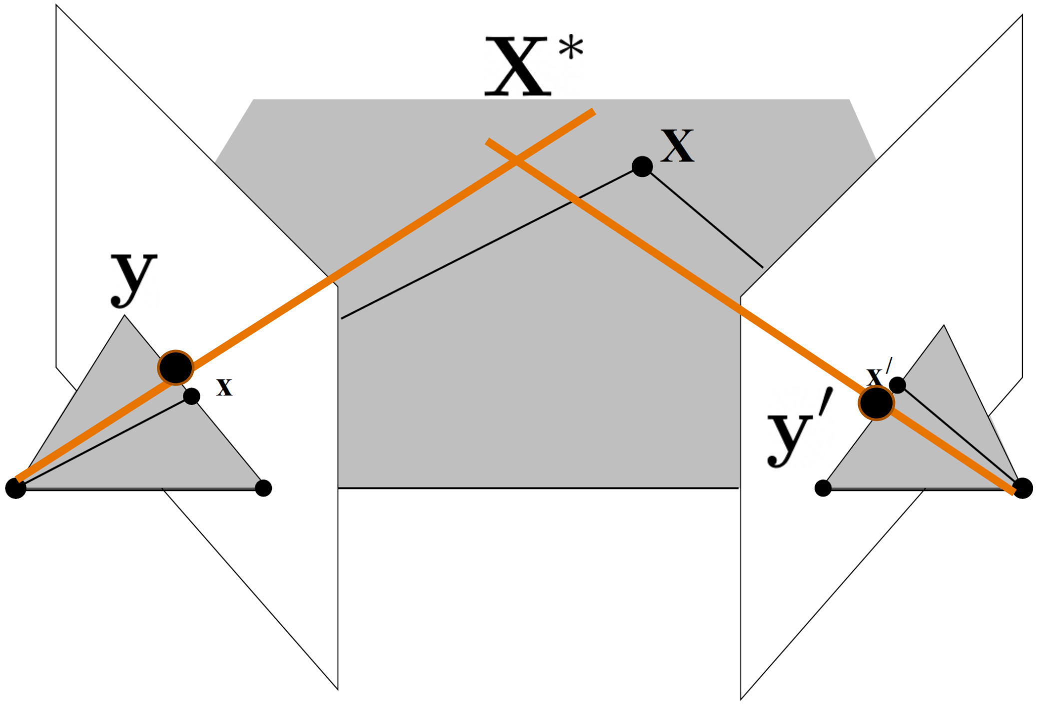
Compute the 3D point given image correspondences



Intersection of two backprojected lines

$$\mathbf{X} = \mathbf{l} \times \mathbf{l}'$$

Triangulation



- In practice, we find the correspondences

  $\mathbf{y} \ \ \mathbf{y}'$

- The backprojected lines may not intersect
- Find X* that minimizes

$$d(\mathbf{y}, P\mathbf{X}^*) + d(\mathbf{y}', P'\mathbf{X}^*)$$

Projection matrix

Summary

Depth perception

- Monocular cues
- Stereo cues

Computational models for stereo vision

- Epipolar geometry
- Stereo Systems
- Triangulation

Further Reading

Multiview Geometry in Computer Vision, Richard Hartley and Andrew Zisserman, Chapter 9, Epipolar Geometry and Fundamental Matrix

Stanford CS231A: Computer Vision, From 3D Reconstruction to Recognition, Lecture 5
https://web.stanford.edu/class/cs231a/syllabus.html