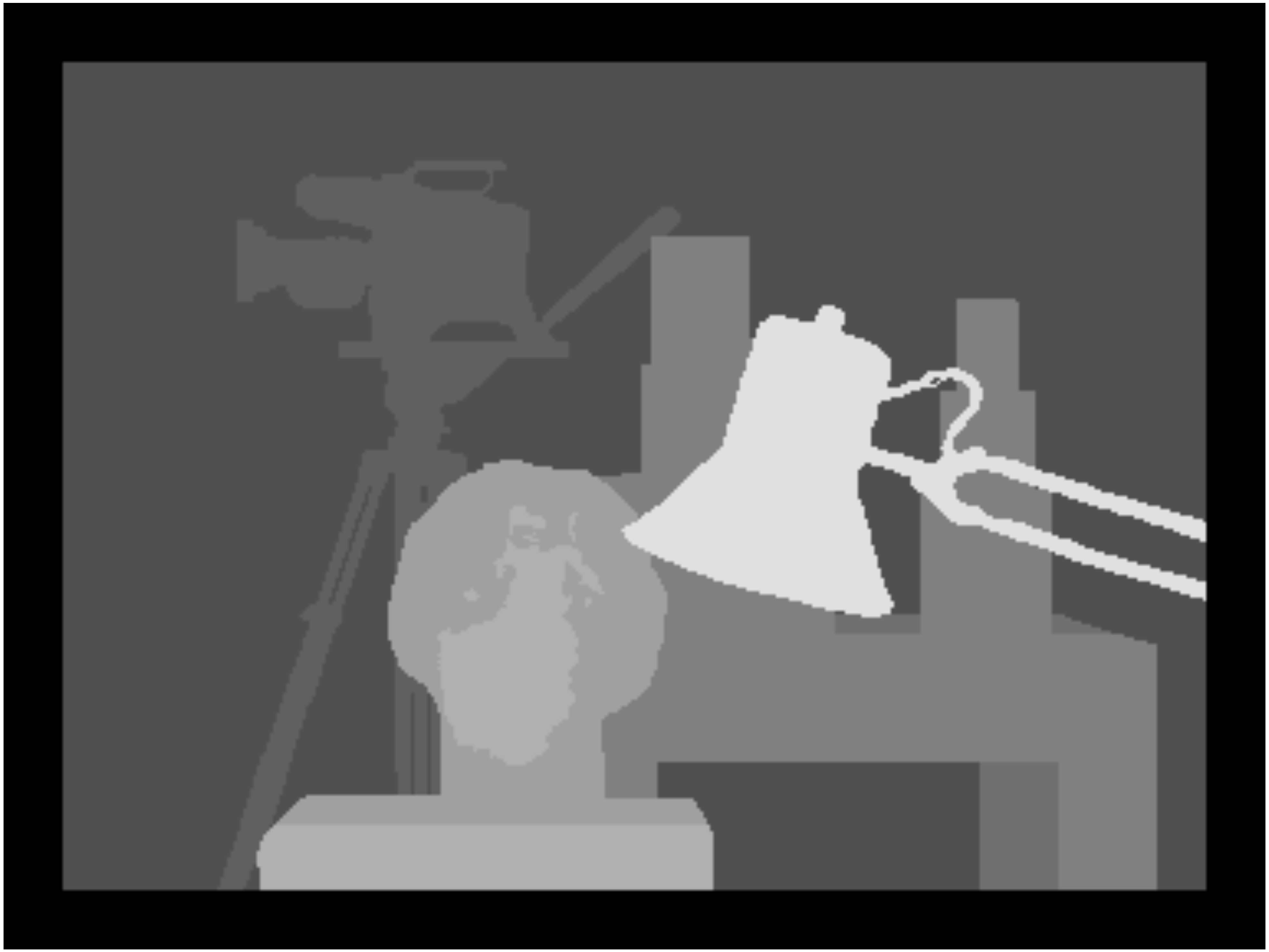
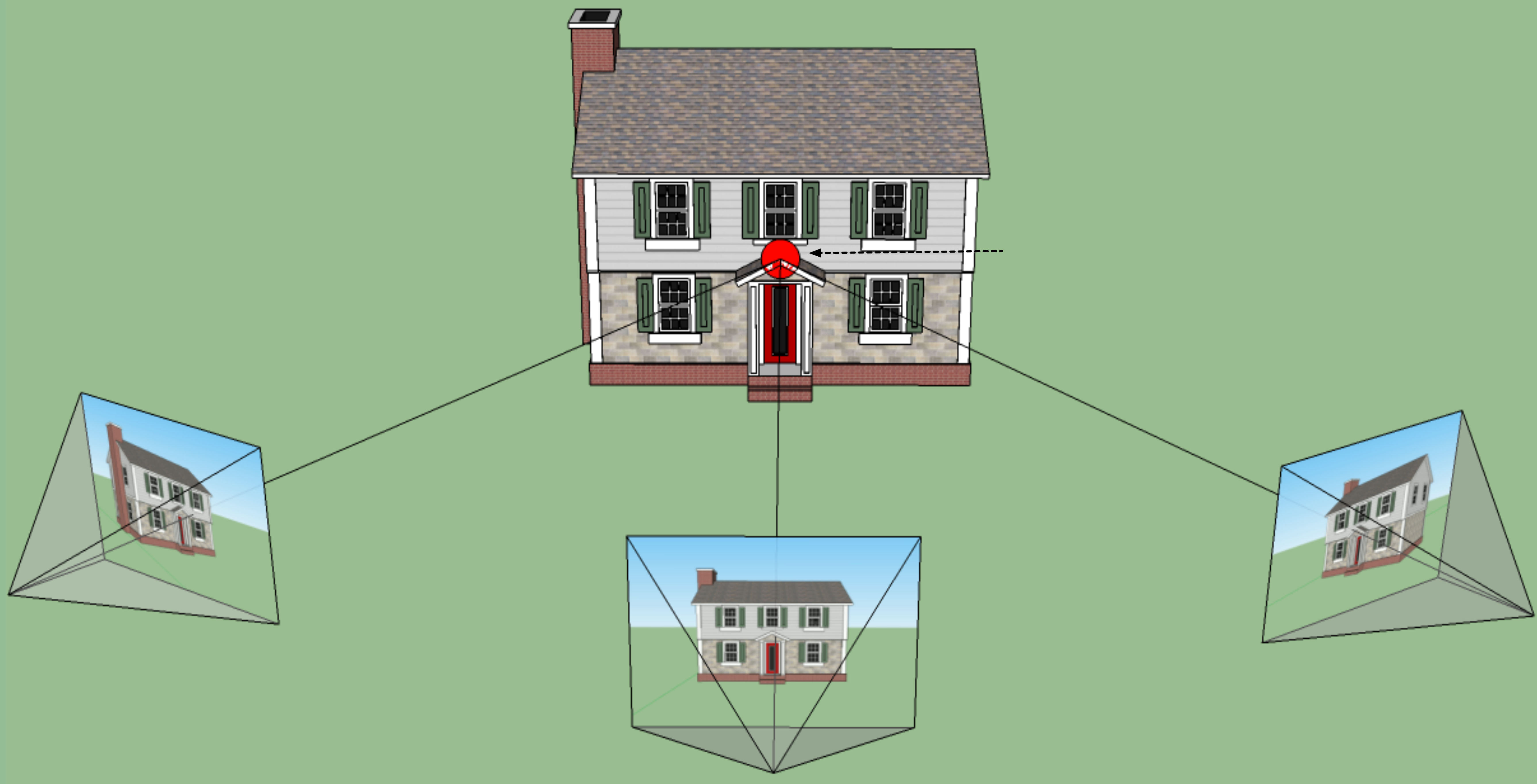
3D Reconstruction with Computer Vision
Meeting 8: Stereo and Epipolar Geometry



Slides by Kristen Grauman, Richard Szeliski and others
CS 378 Fall 2014, UT Austin, Bryan Klingner, 23 September

Idea: 3D from multiple images of a scene

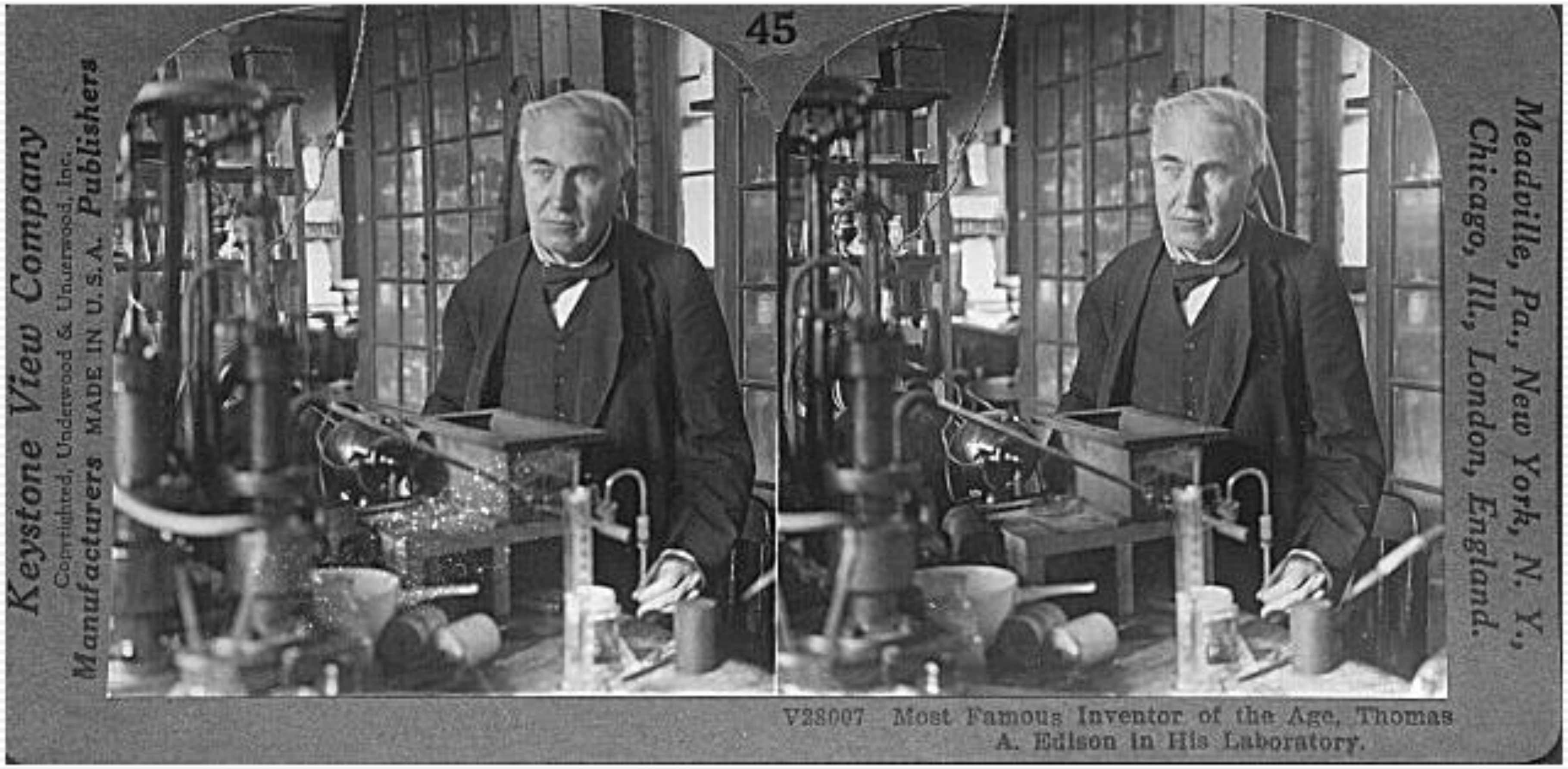


Stereo photography and stereo viewers

Take two pictures of the same subject from two slightly different viewpoints and display so that each eye sees only one of the images.

Invented by Sir Charles Wheatstone, 1838

Image from fisher-price.com



http://www.johnsonshawmuseum.org

http://www.johnsonshawmuseum.org

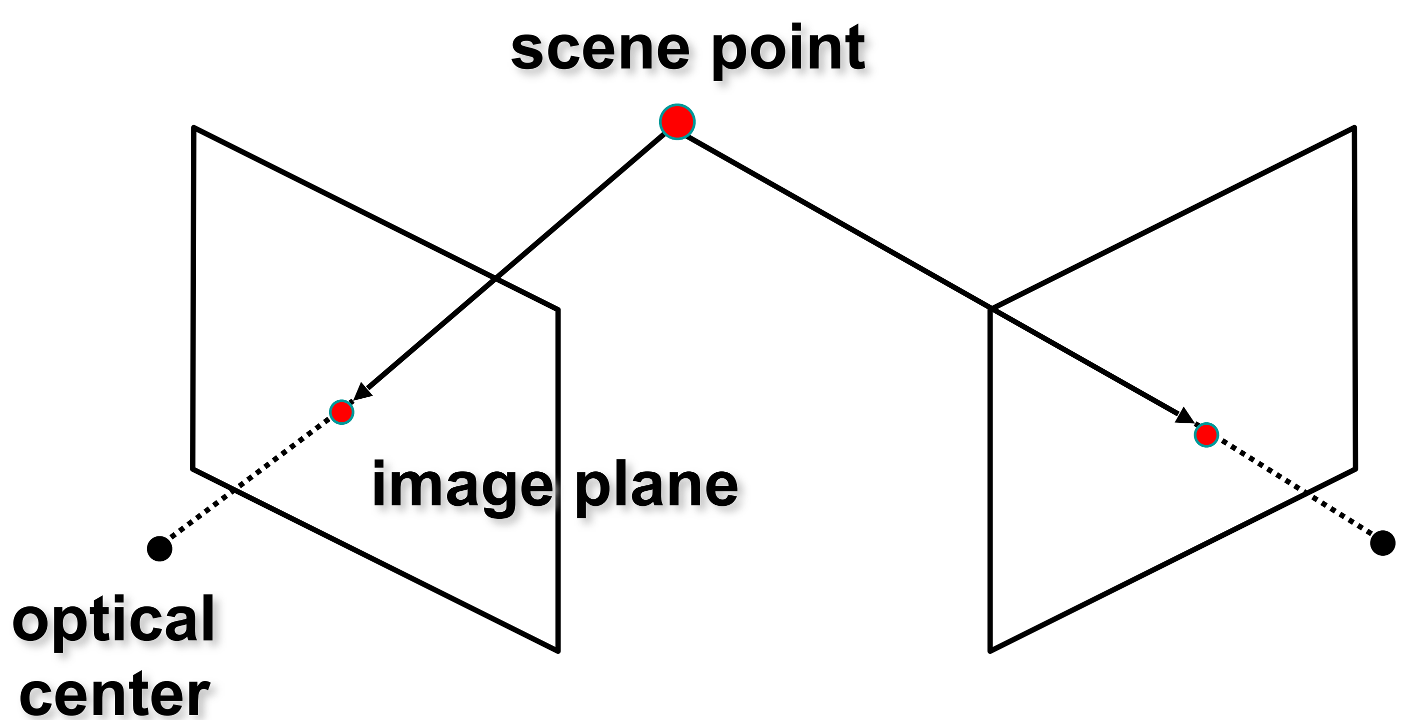
Stereo vision

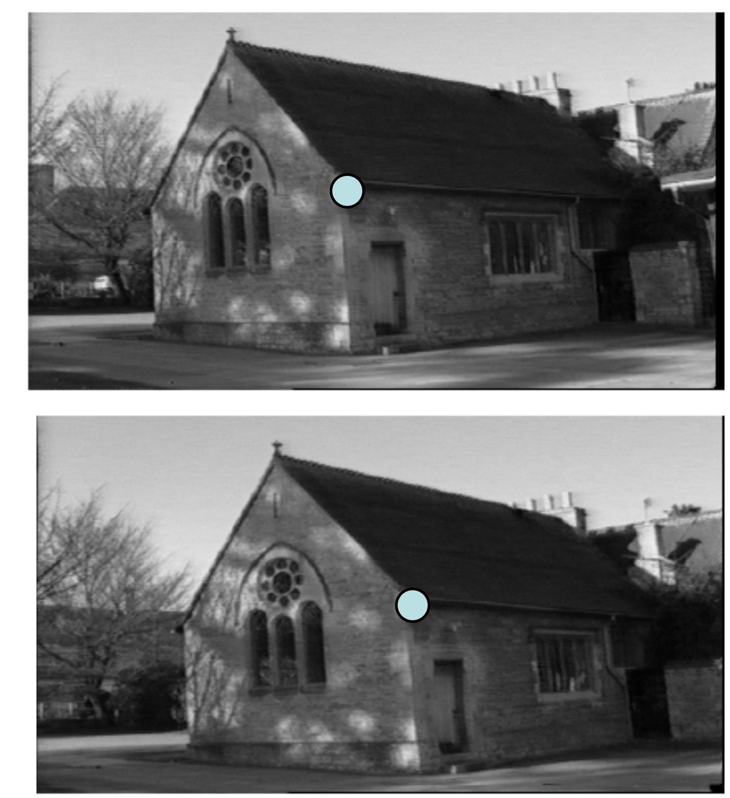
Two cameras, simultaneous views

Single moving camera and static scene

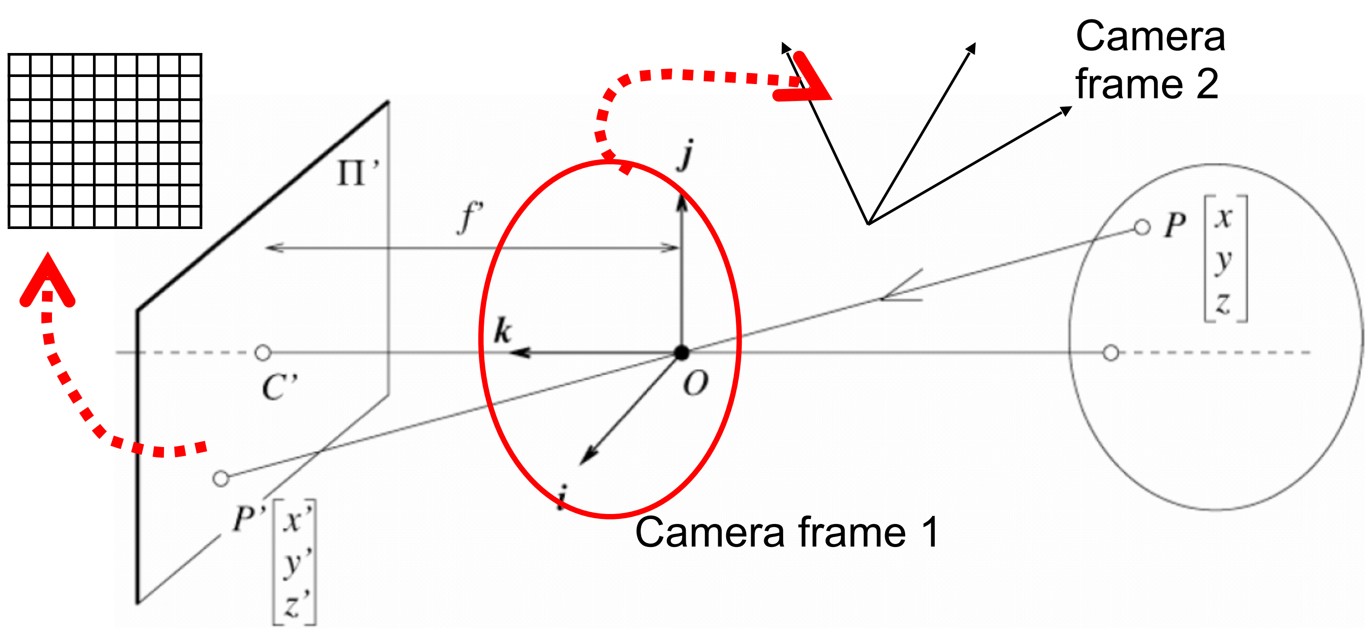
Estimating depth with stereo

- **Stereo**: shape from “motion” between two views
- We’ll need to consider:
  - Info on camera pose (“calibration”)
  - Image point correspondences





Camera parameters



Extrinsic parameters:
Camera frame 1 ←→ Camera frame 2

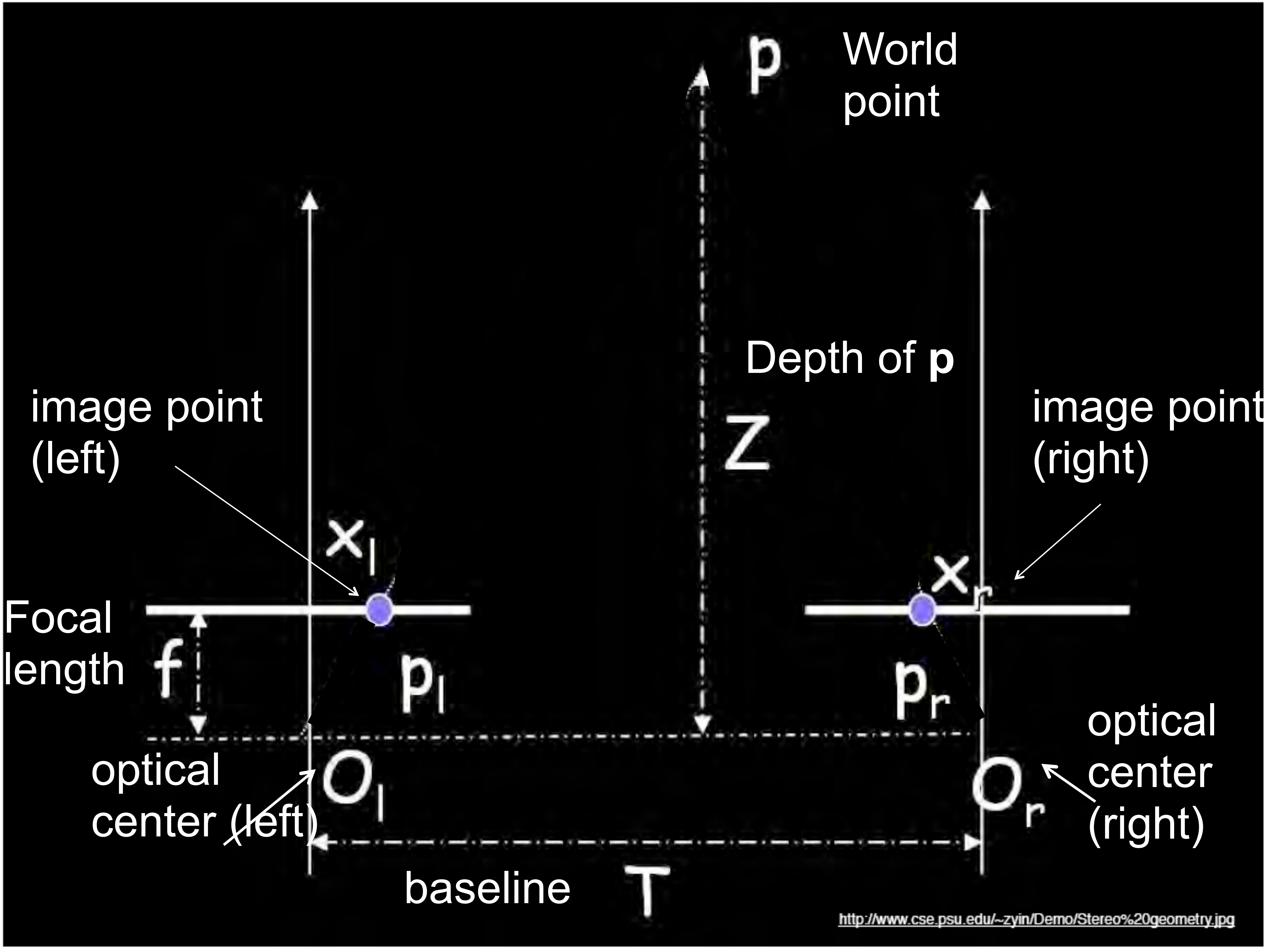
Intrinsic parameters:
Image coordinates relative to camera ←→ Pixel coordinates

- *Extrinsic* params: rotation matrix and translation vector
- *Intrinsic* params: focal length, pixel sizes (mm), image center point, radial distortion parameters

We'll assume for now that these parameters are given and fixed.

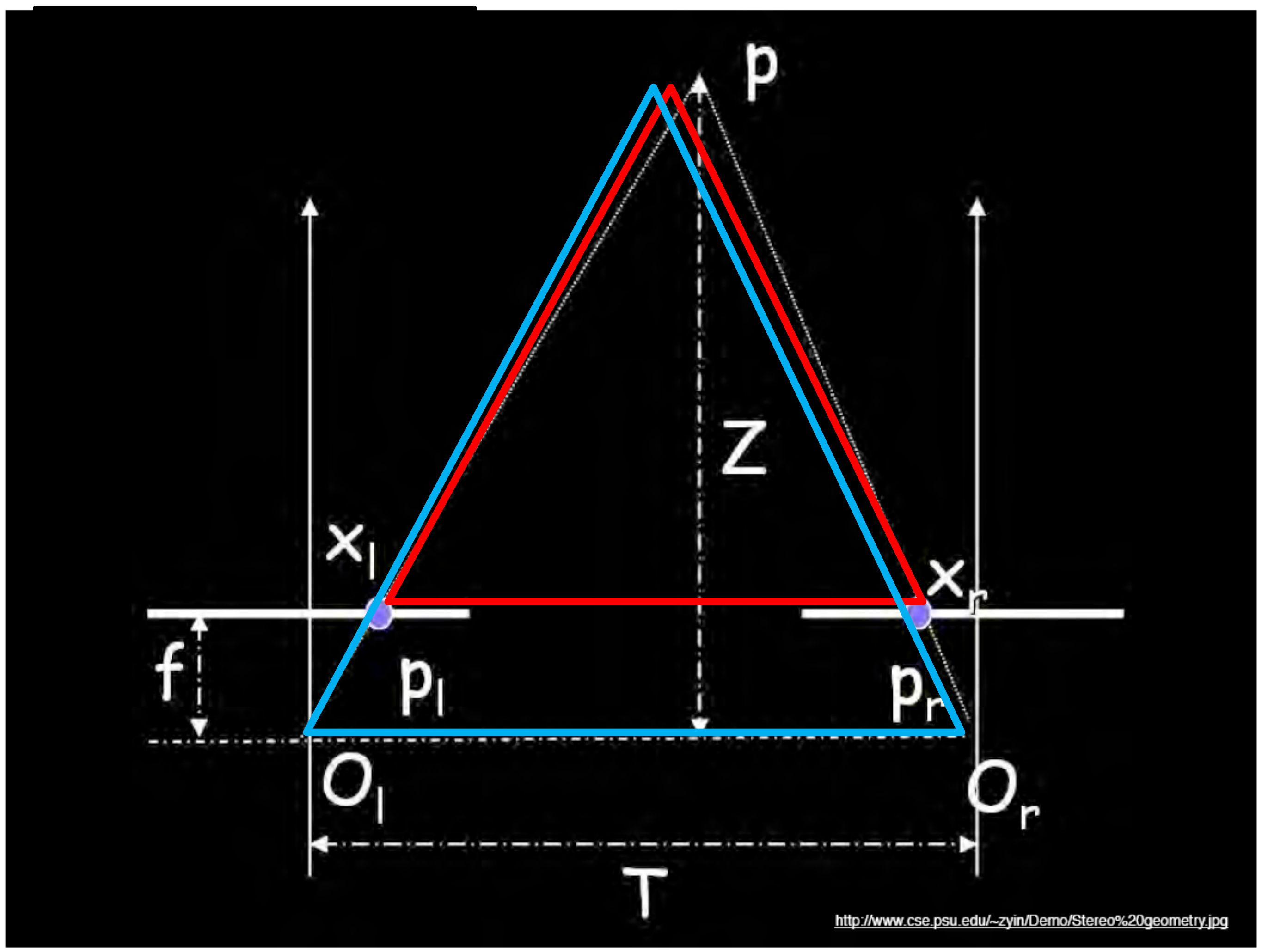
Geometry for a simple stereo system

- First, assuming parallel optical axes, known camera parameters (i.e., calibrated cameras):


p
World
point
Depth of **p**
Z
image point
(left)
image point
(right)
x_l
x_r
Focal
length
f
p_l
p_r
optical
center (left)
O_l
O_r
optical
center
(right)
baseline
T
http://www.cse.psu.edu/~zyin/Demo/Stereo%20geometry.jpg

Geometry for a simple stereo system

- Assume parallel optical axes, known camera parameters (i.e., calibrated cameras). **What is expression for Z?**



Similar triangles (p_l, P, p_r) and (O_l, P, O_r):

$$\frac{T + x_l - x_r}{Z - f} = \frac{T}{Z}$$

$$Z = f \frac{T}{x_r - x_l}$$

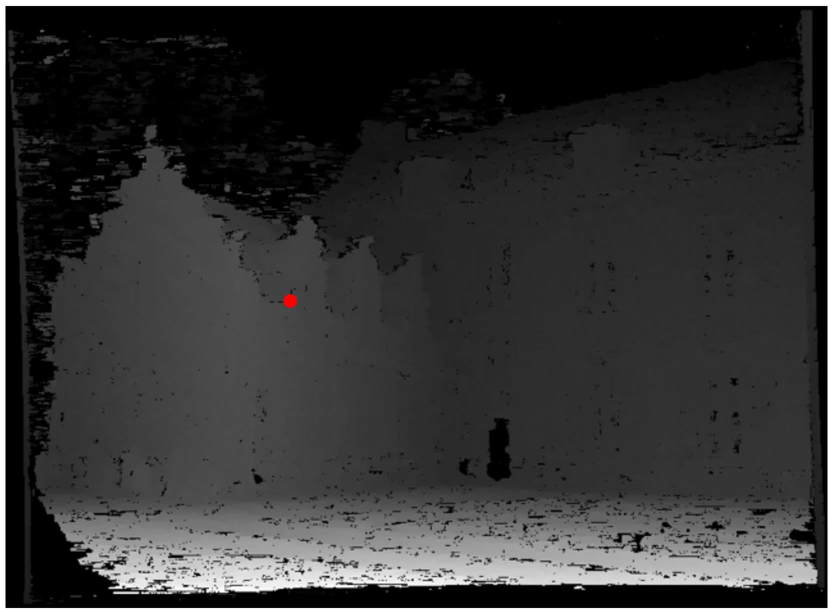
disparity

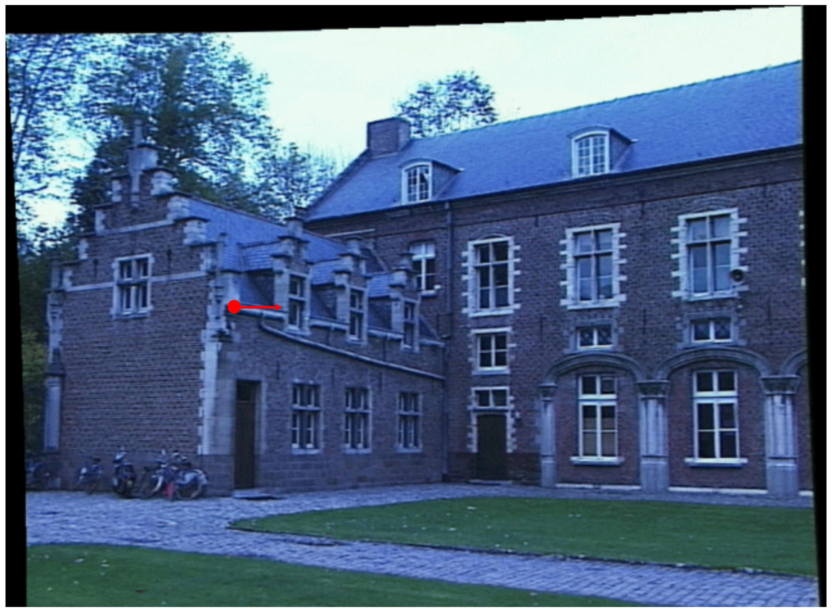
Depth from disparity

image I(x,y)

Disparity map D(x,y)

image I´(x´,y´)



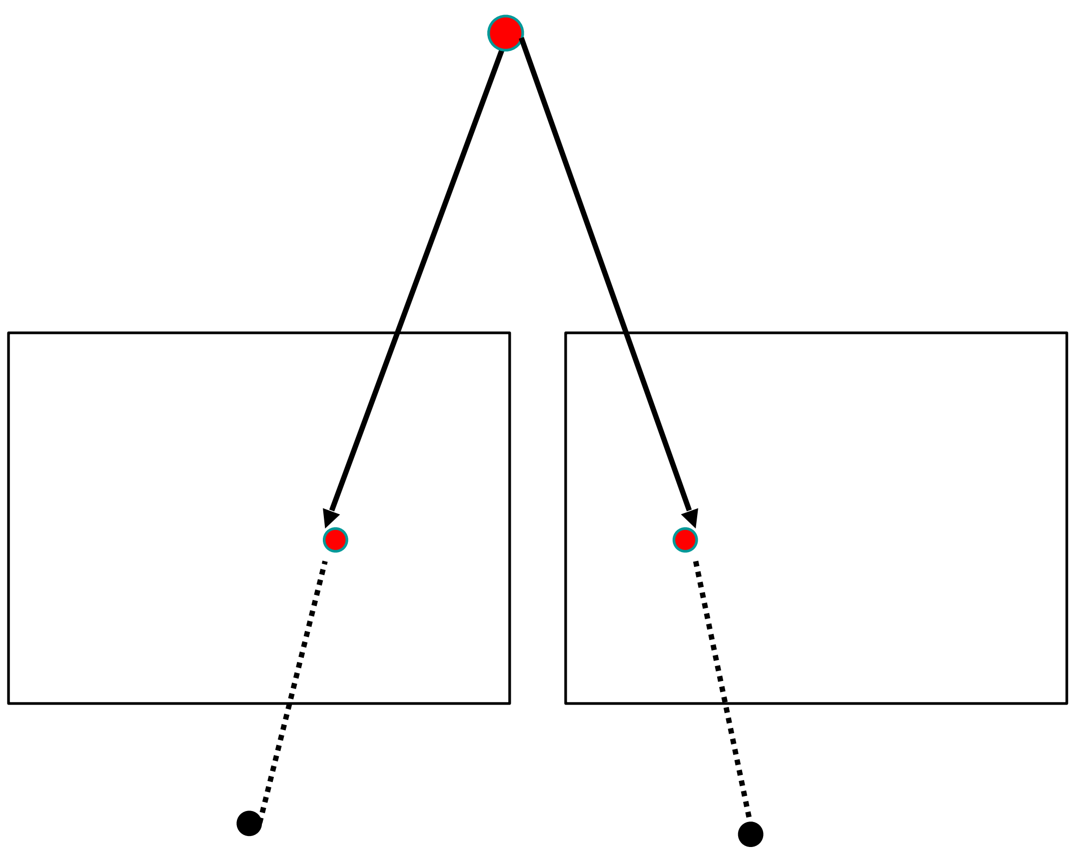


$$(x',y')=(x+D(x,y),\ y)$$

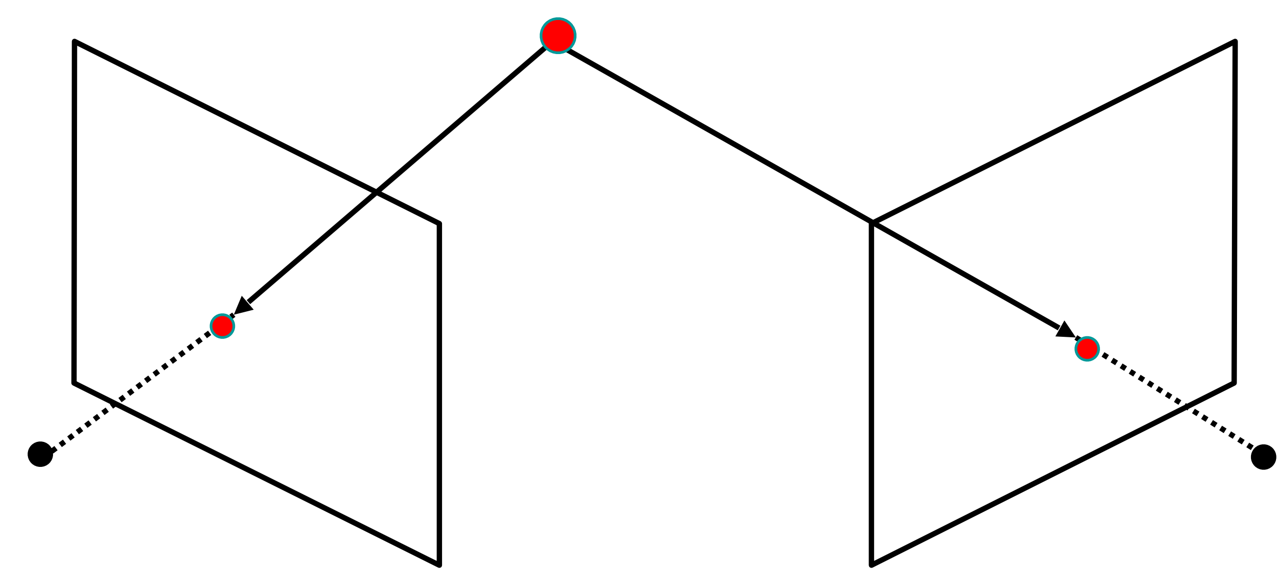
So if we could find the **corresponding points** in two images, we could **estimate relative depth**…

General case, with calibrated cameras

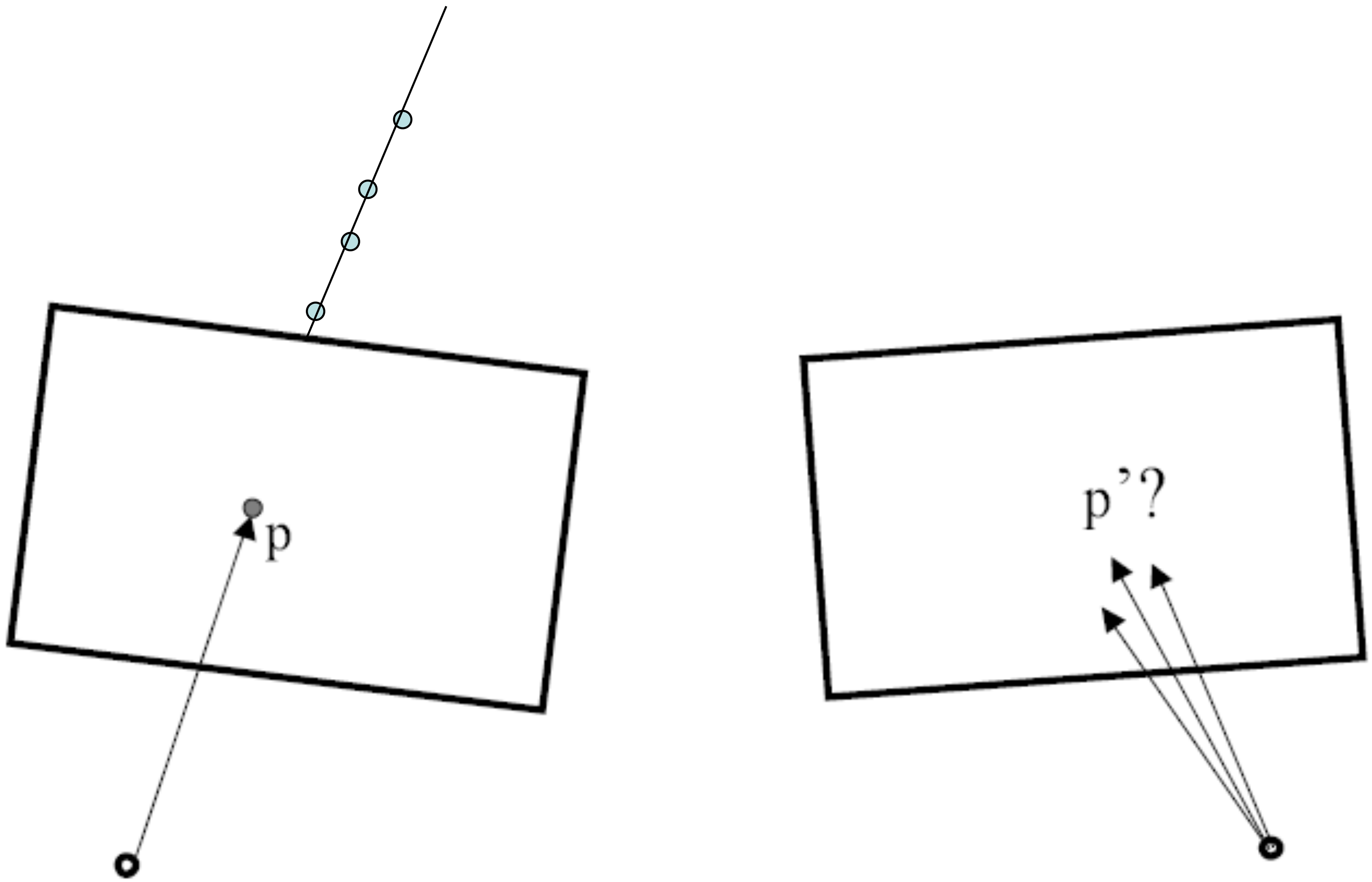
- The two cameras need not have parallel optical axes.



Vs.

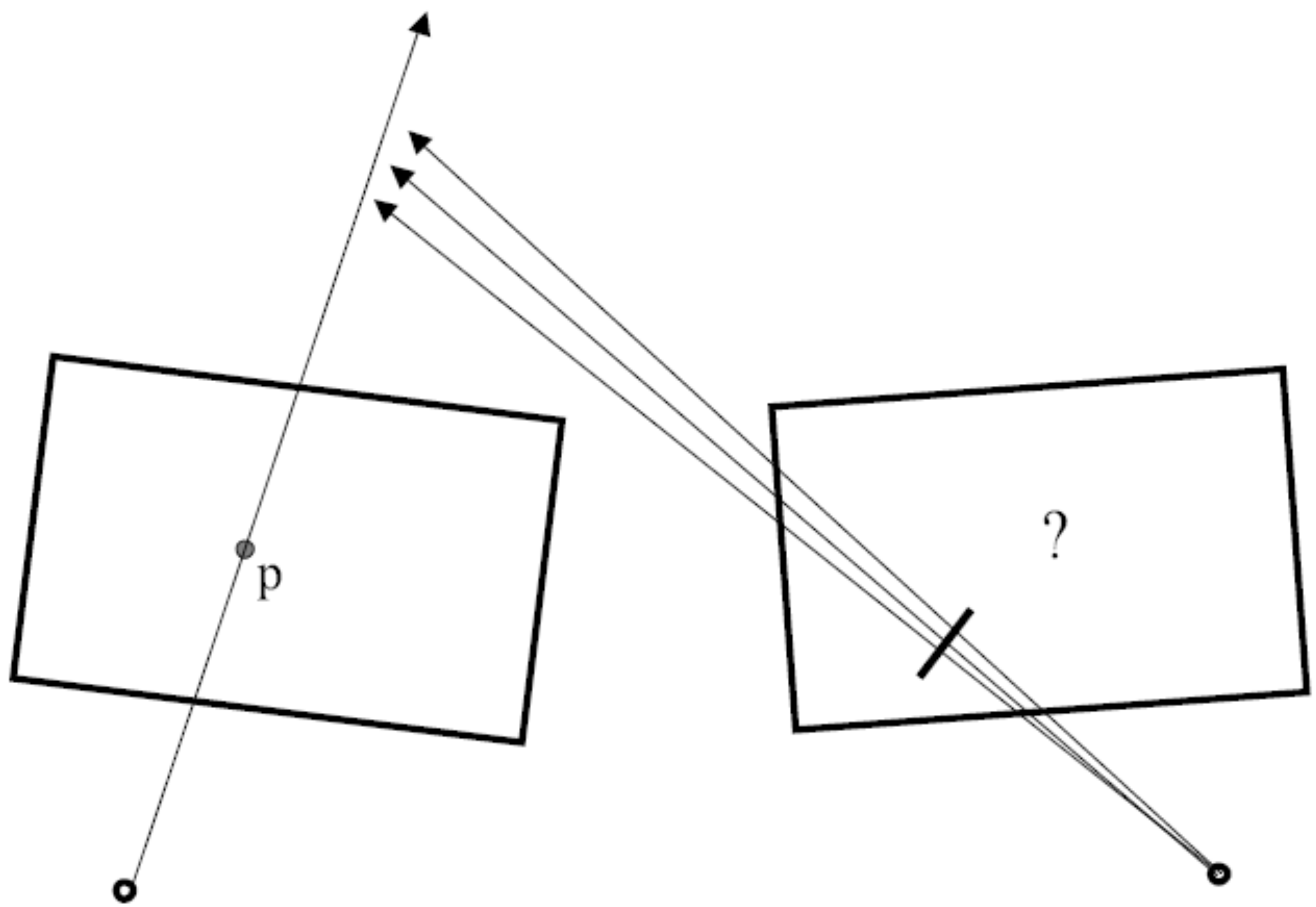


Stereo correspondence constraints

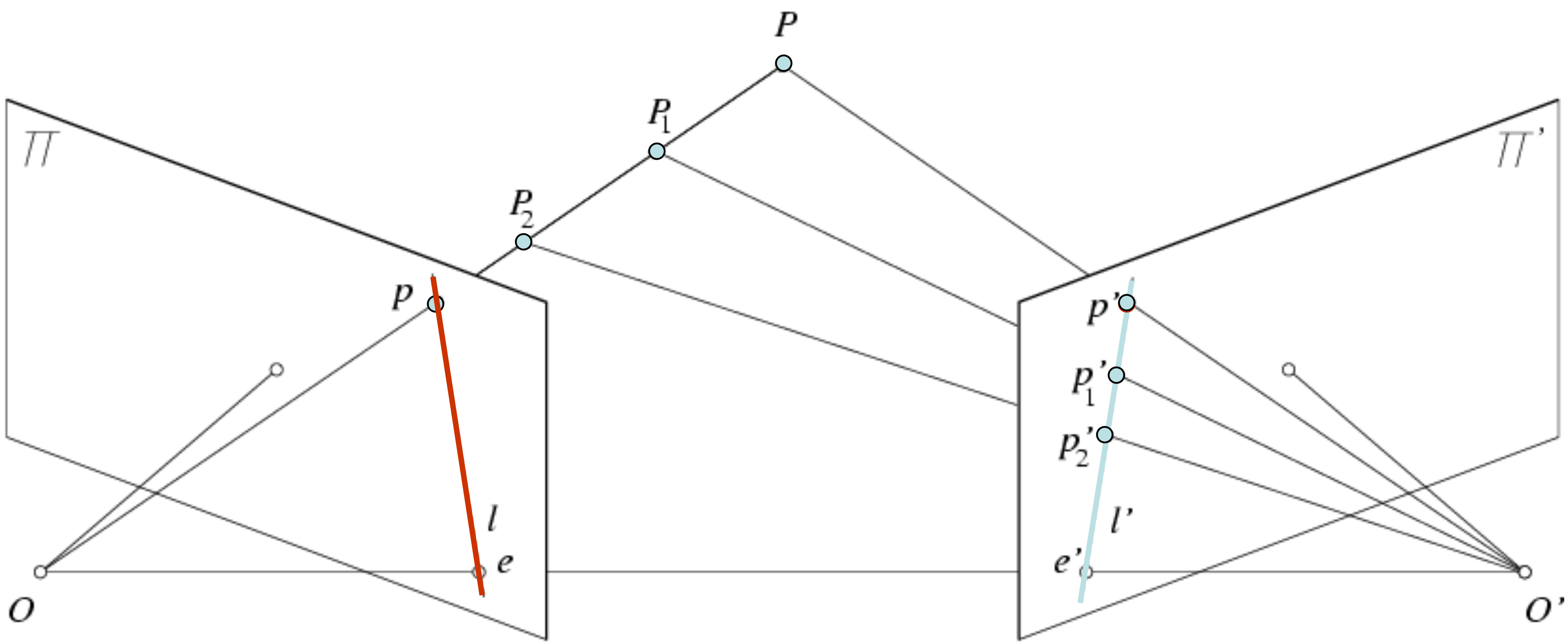


- Given p in left image, where can corresponding point p' be?

Stereo correspondence constraints



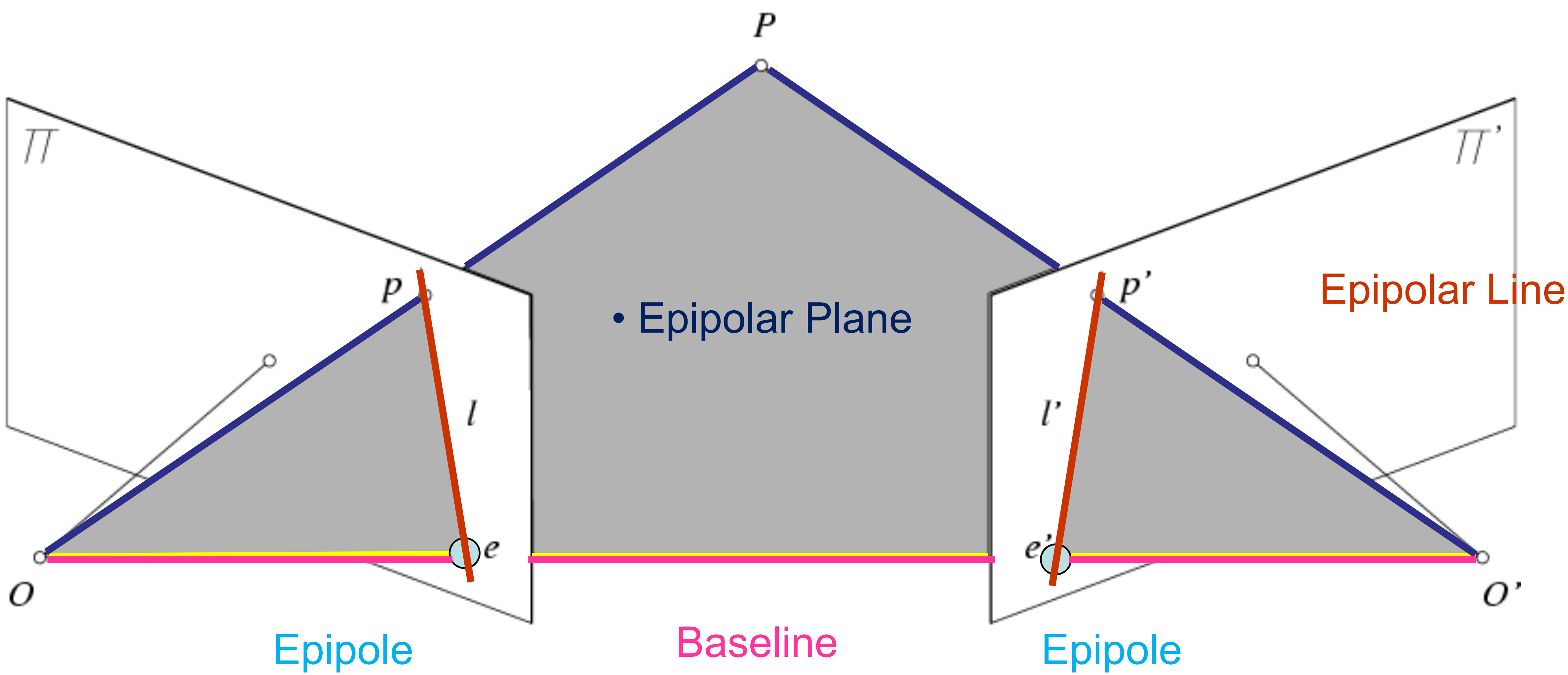
Epipolar constraint



Geometry of two views constrains where the corresponding pixel for some image point in the first view must occur in the second view.

- It must be on the line carved out by a plane connecting the world point and optical centers.

Epipolar geometry



http://www.ai.sri.com/~luong/research/Meta3DViewer/EpipolarGeo.html

Epipolar geometry: terms

- **Baseline**: line joining the camera centers
- **Epipole**: point of intersection of baseline with image plane
- **Epipolar plane**: plane containing baseline and world point
- **Epipolar line**: intersection of epipolar plane with the image plane

- All epipolar lines intersect at the epipole
- An epipolar plane intersects the left and right image planes in epipolar lines

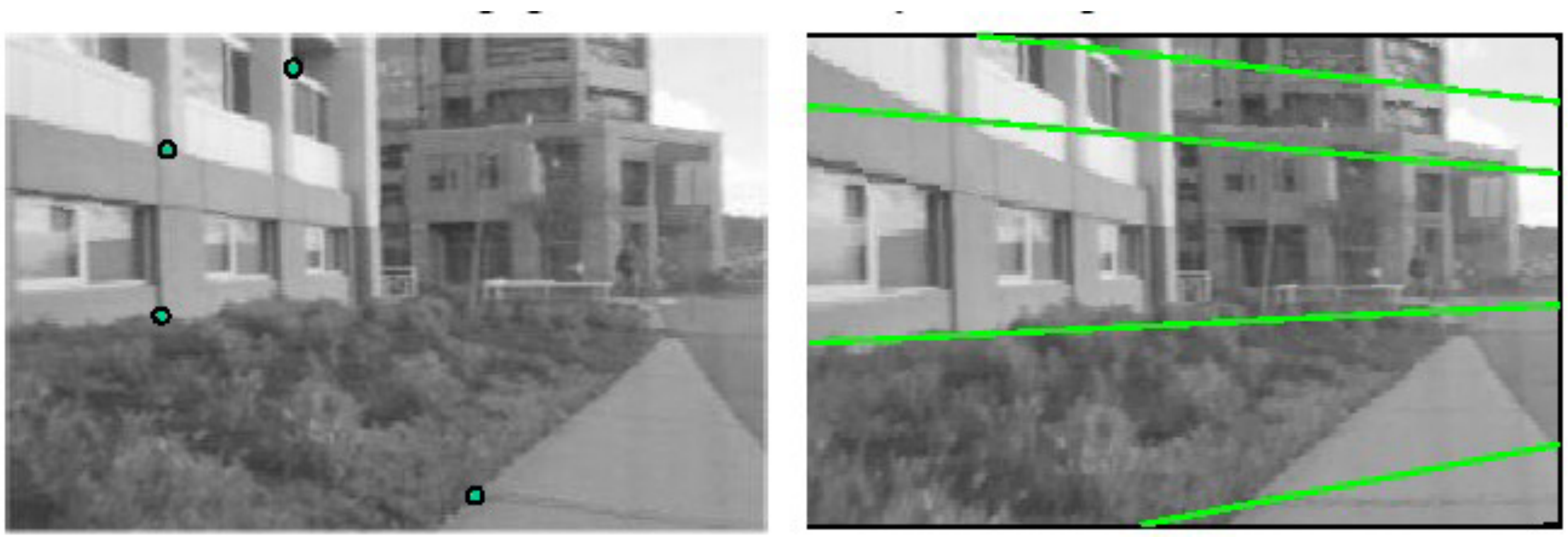
Why is the epipolar constraint useful?

Epipolar constraint

This is useful because it reduces the correspondence problem to a 1D search along an epipolar line.

Image from Andrew Zisserman

Example



Example: converging cameras

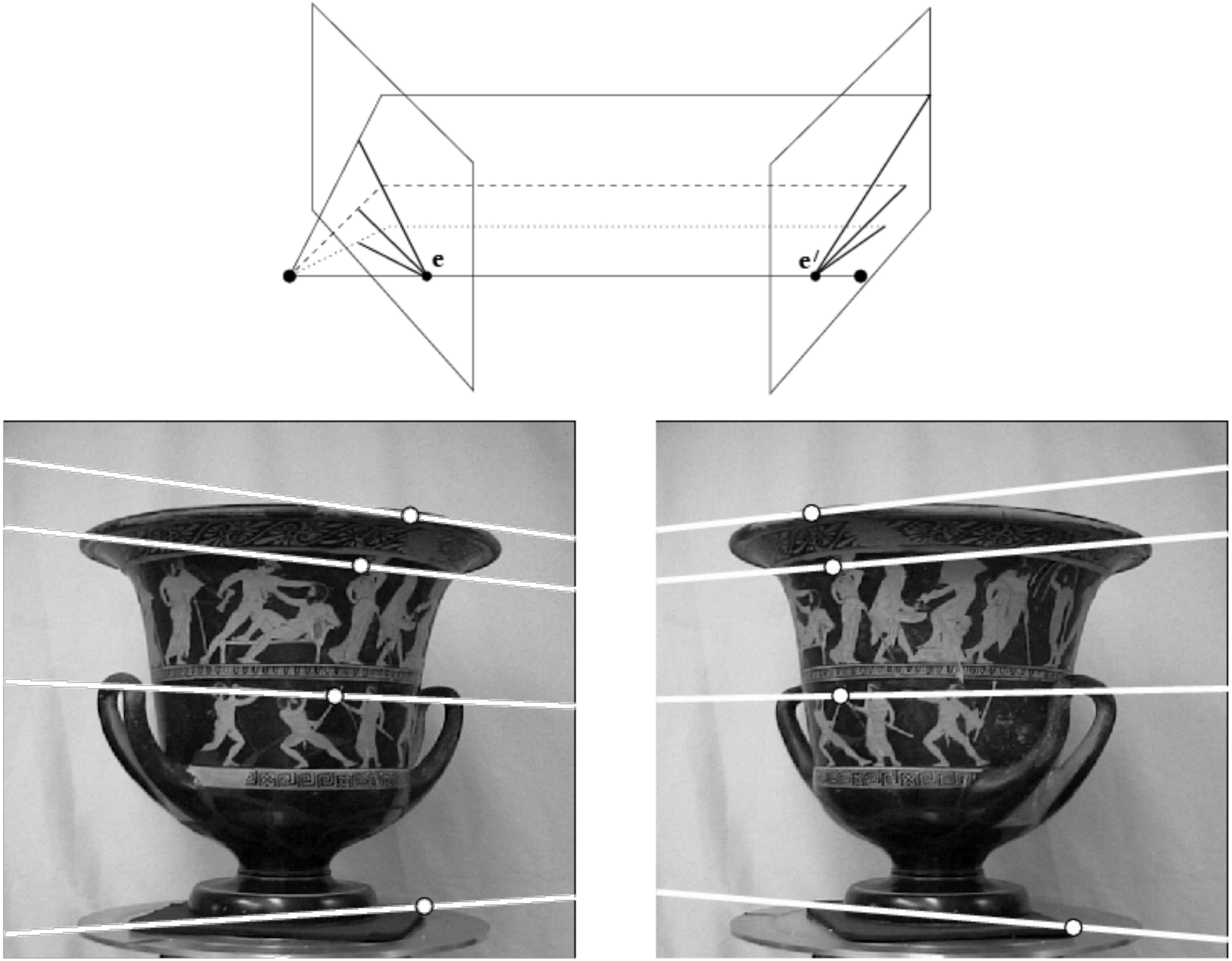
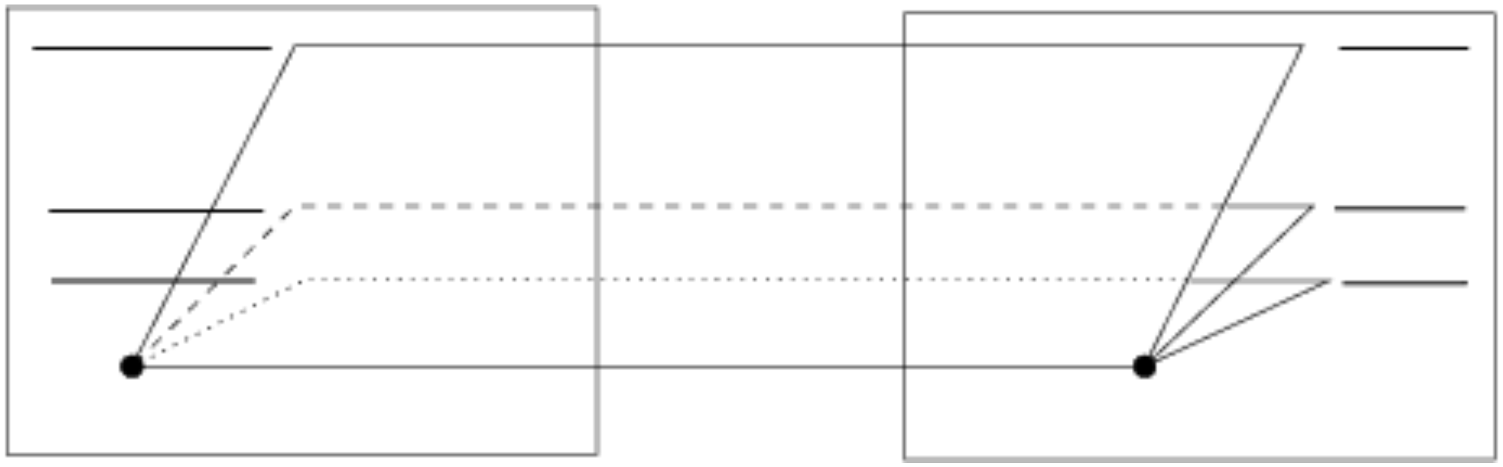


Figure from Hartley & Zisserman

Example: parallel cameras



Where are the epipoles?

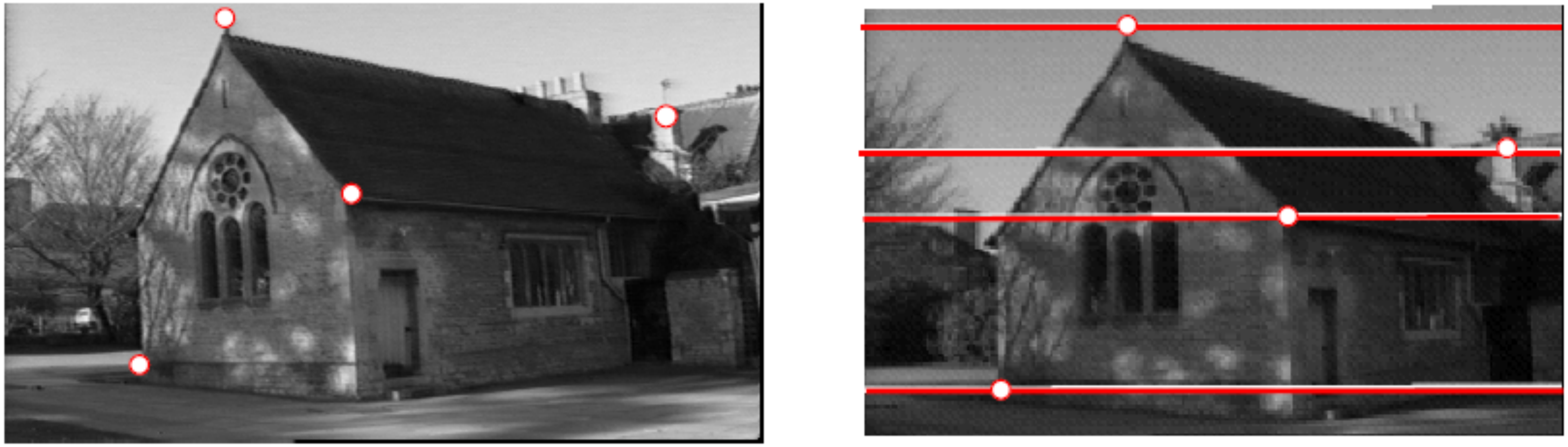
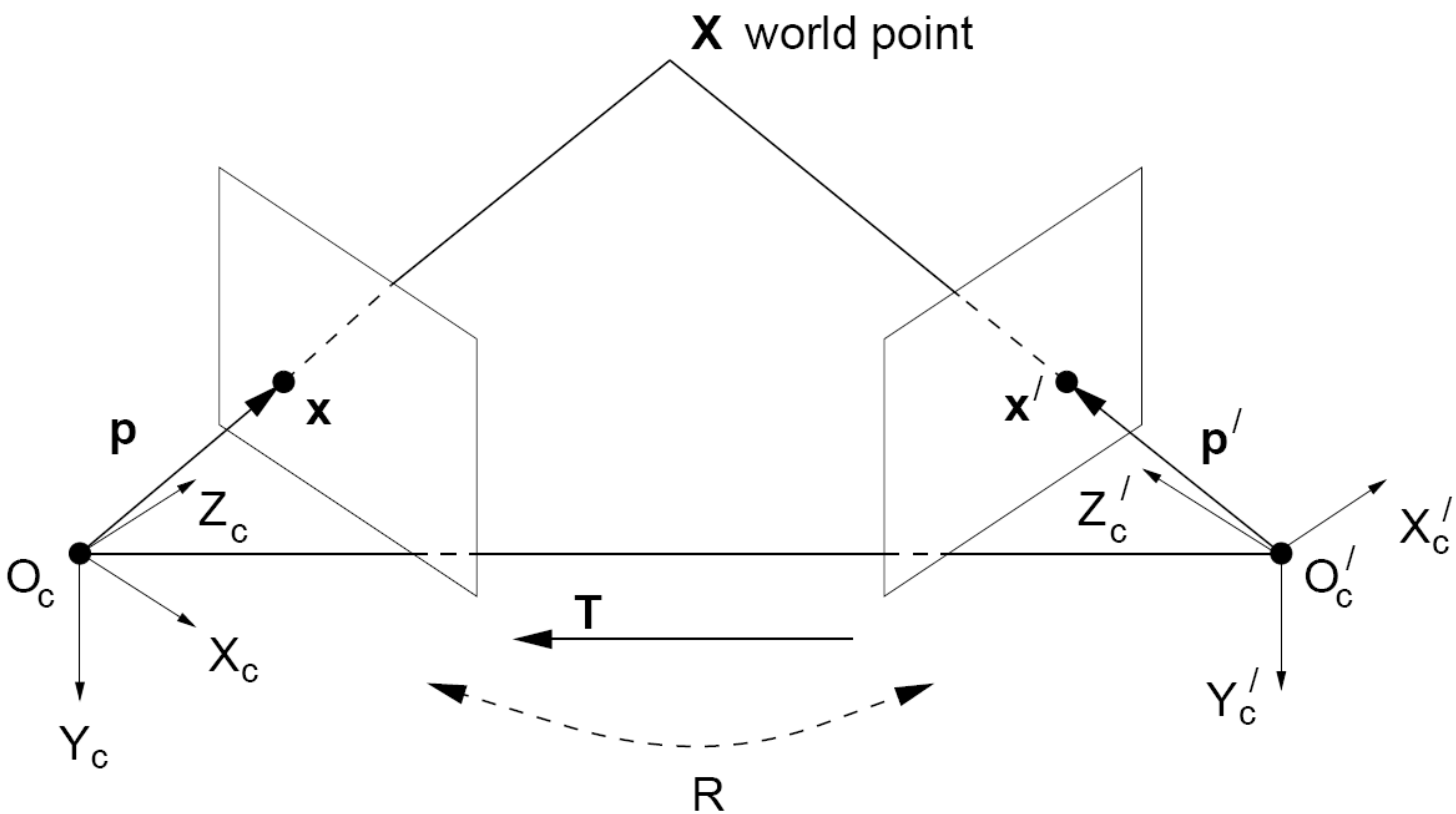


Figure from Hartley & Zisserman

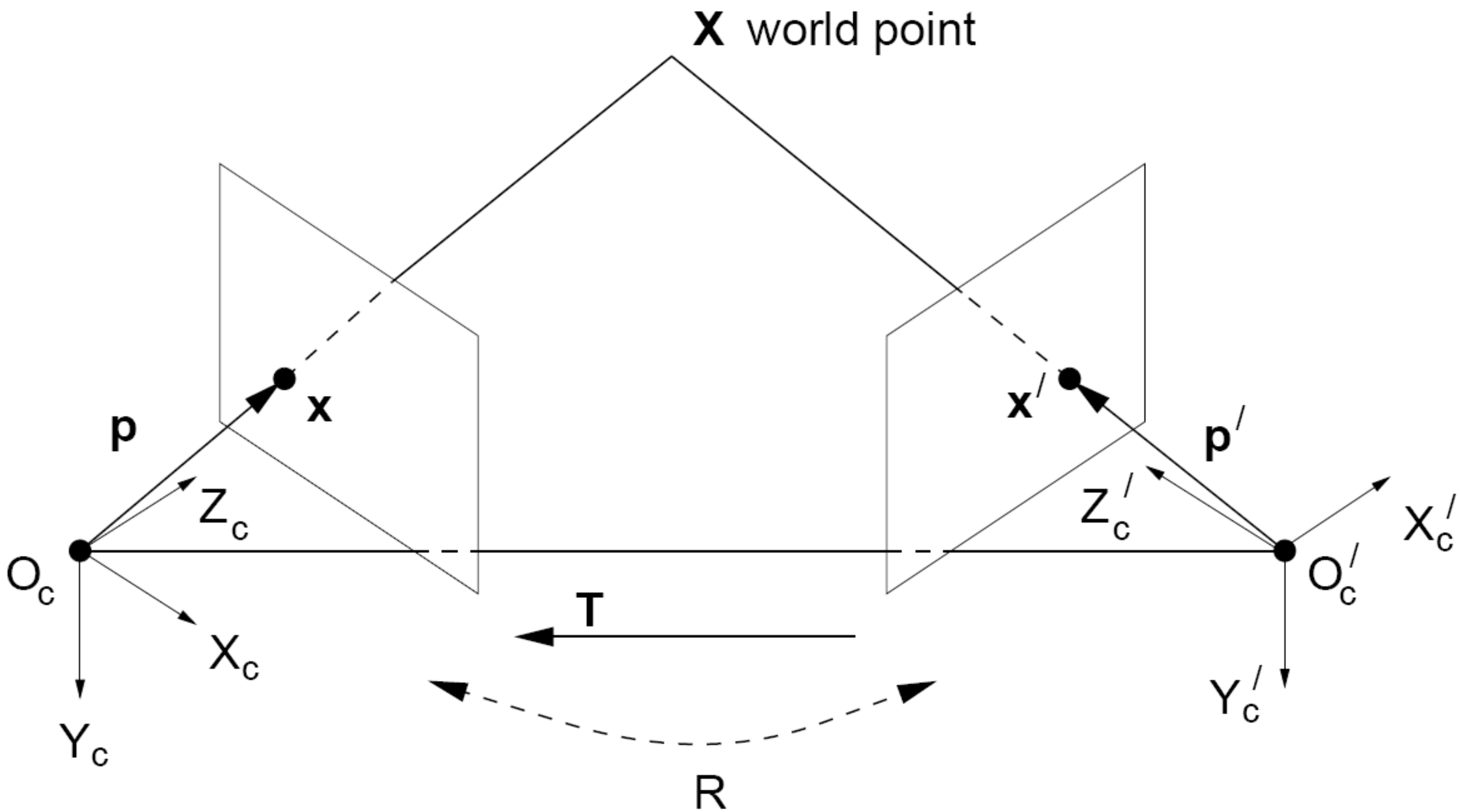
- So far, we have the explanation in terms of geometry.
- Now, how to express the epipolar constraints algebraically?

Stereo geometry, with calibrated cameras



Main idea

Stereo geometry, with calibrated cameras

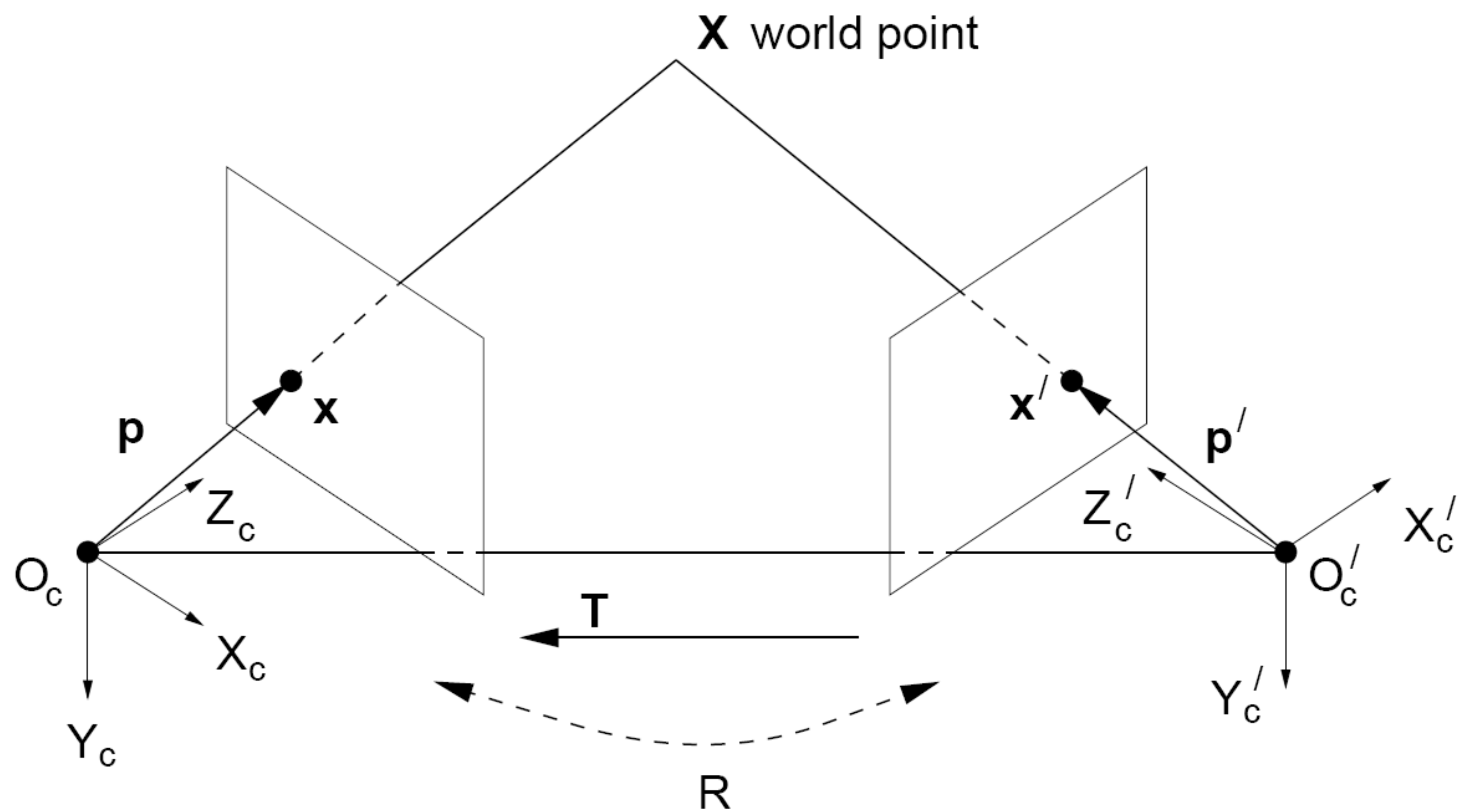


If the stereo rig is calibrated, we know :

how to **rotate** and **translate** camera reference frame 1 to get to camera reference frame 2.

Rotation: 3 x 3 matrix **R**; translation: 3 vector **T**.

Stereo geometry, with calibrated cameras



If the stereo rig is calibrated, we know :

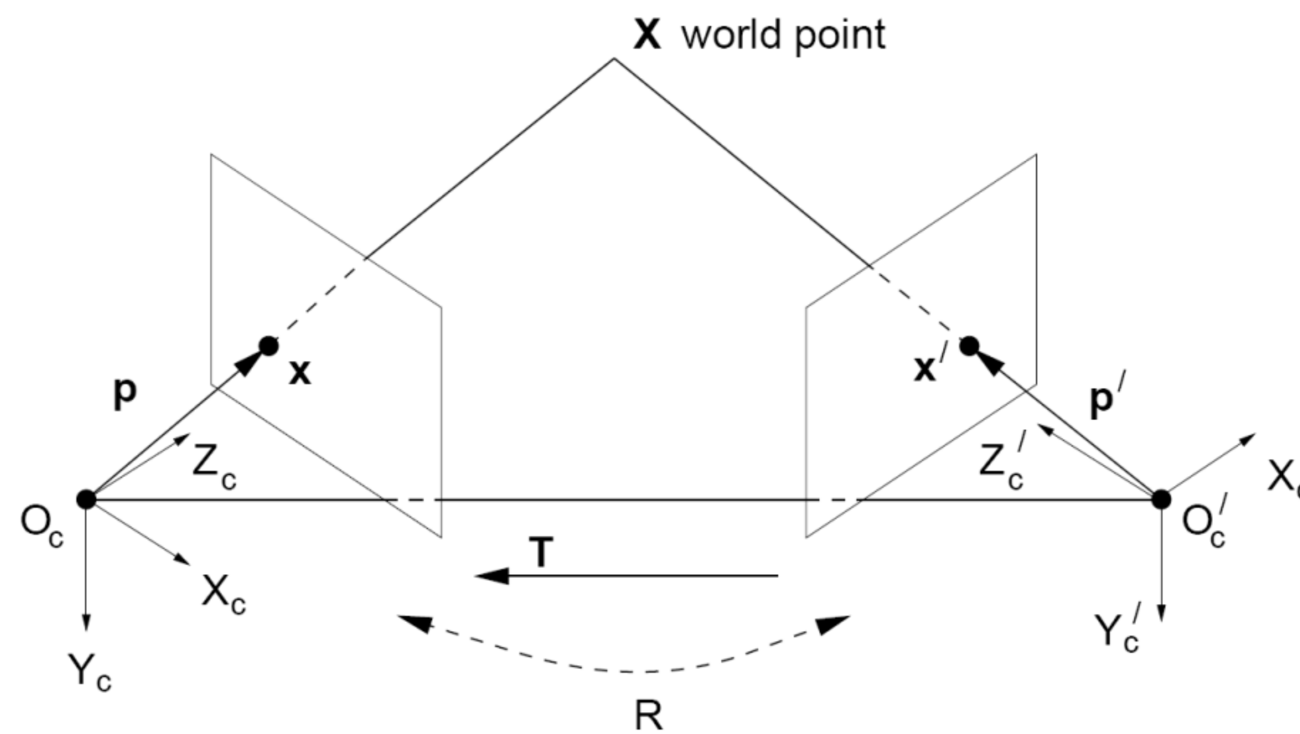
how to **rotate** and **translate** camera reference frame 1 to get to camera reference frame 2.

$$\mathbf{X}'_c = \mathbf{R}\mathbf{X}_c + \mathbf{T}$$

Essential matrix

$$\boxed{\mathbf{X}' \cdot \left(\mathbf{T} \times \mathbf{R}\mathbf{X}\right) = 0}$$

$$\mathbf{X}' \cdot \left([\mathrm{T}_x]\mathbf{R}\mathbf{X}\right) = 0$$



Let $\mathbf{E} = [\mathrm{T}_x]\mathbf{R}$

$$\mathbf{X}'^{T}\mathbf{E}\mathbf{X} = 0$$

E is called the **essential matrix**, and it relates corresponding image points between both cameras, given the rotation and translation.

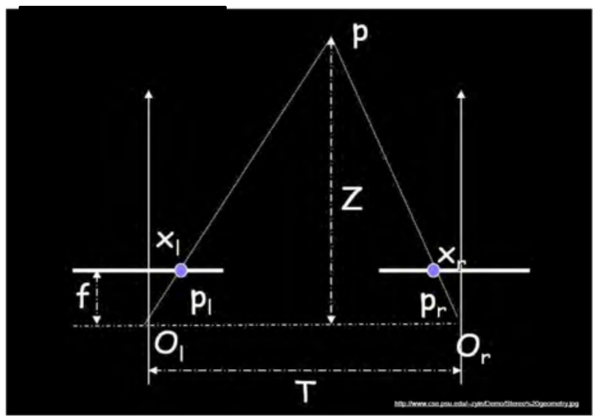
If we observe a point in one image, its position in other image is constrained to lie on line defined by above.

Note: these points are in **camera coordinate systems**.

Essential matrix **E**:
constrains points in one **camera frame** to lie on epipolar lines in another **camera frame**.

Fundamental matrix **F**:
constrains points in one **image** to lie on epipolar lines in another **image**. It’s just E, but measured in **pixels, not inches**!

Essential matrix example: parallel cameras



$\mathbf{R} =$

$\mathbf{T} =$

$\mathbf{E} = [\mathrm{T}_{\mathbf{x}}]\mathbf{R} =$

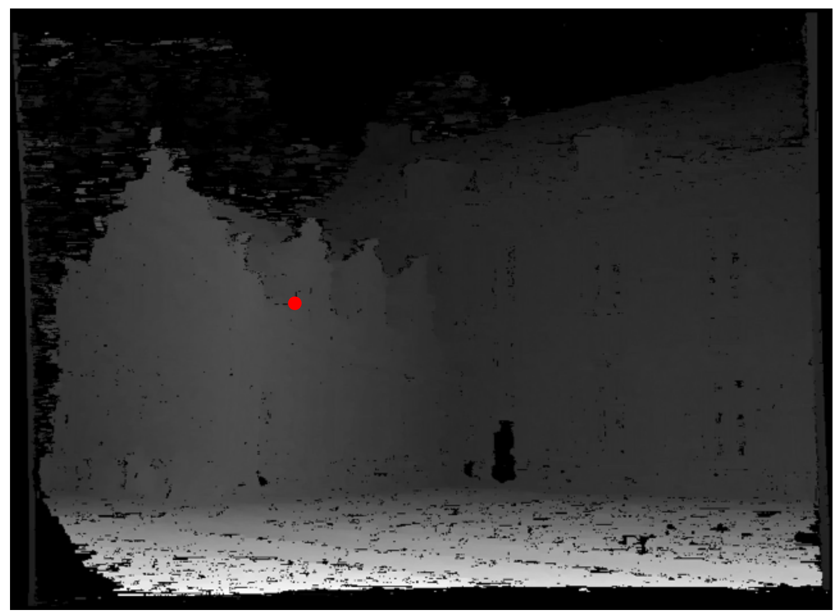
$\mathbf{p} = [x, y, f]$

$\mathbf{p}' = [x', y', f]$

$$\mathbf{p}'^{\mathrm{T}}\mathbf{E}\mathbf{p} = 0$$

For the parallel cameras, image of any point must lie on same horizontal line in each image plane.

image I(x,y) Disparity map D(x,y) image I´(x´,y´)

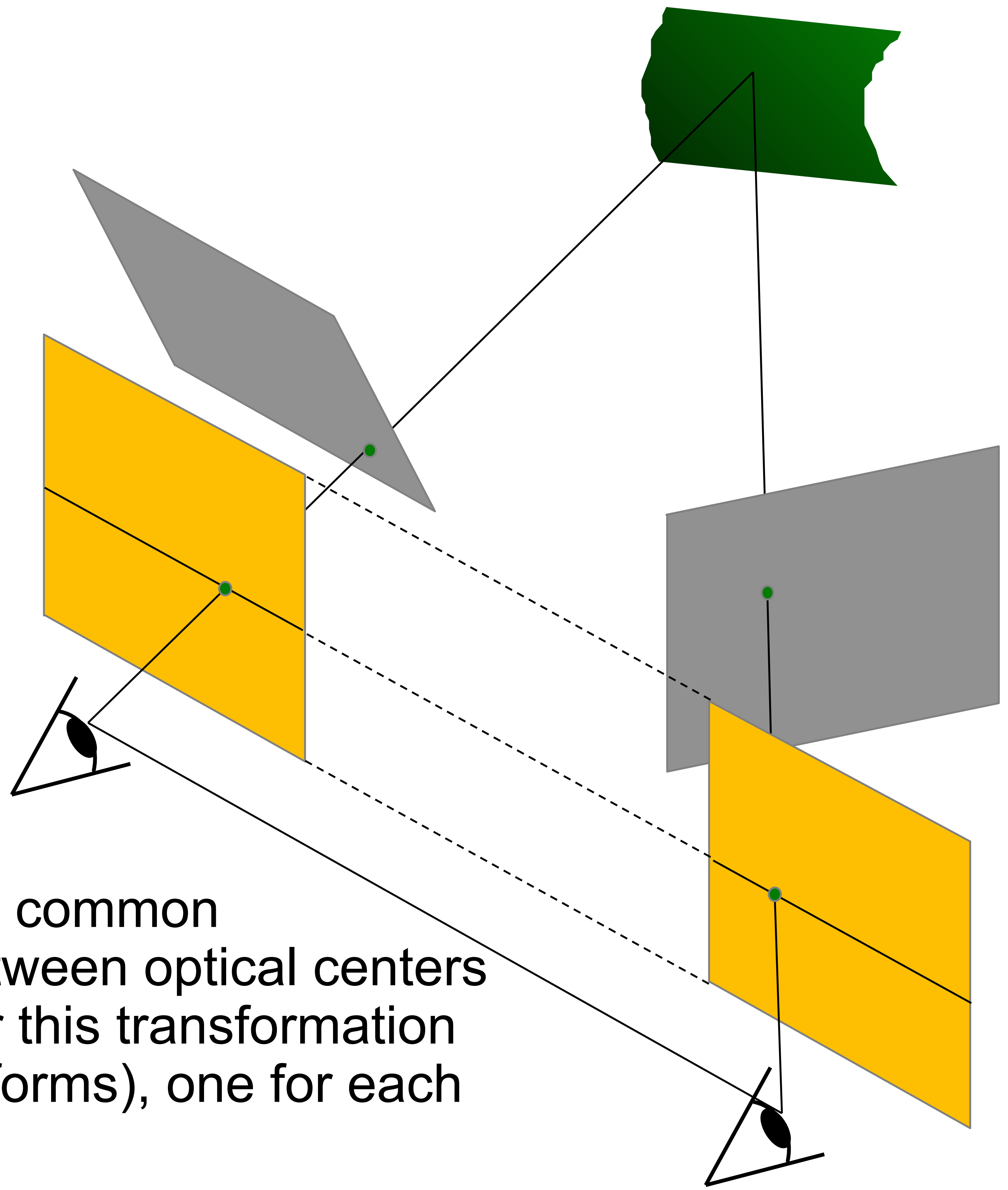


$$(x',y')=(x+D(x,y),y)$$

What about when cameras' optical axes are not parallel?

Stereo image rectification

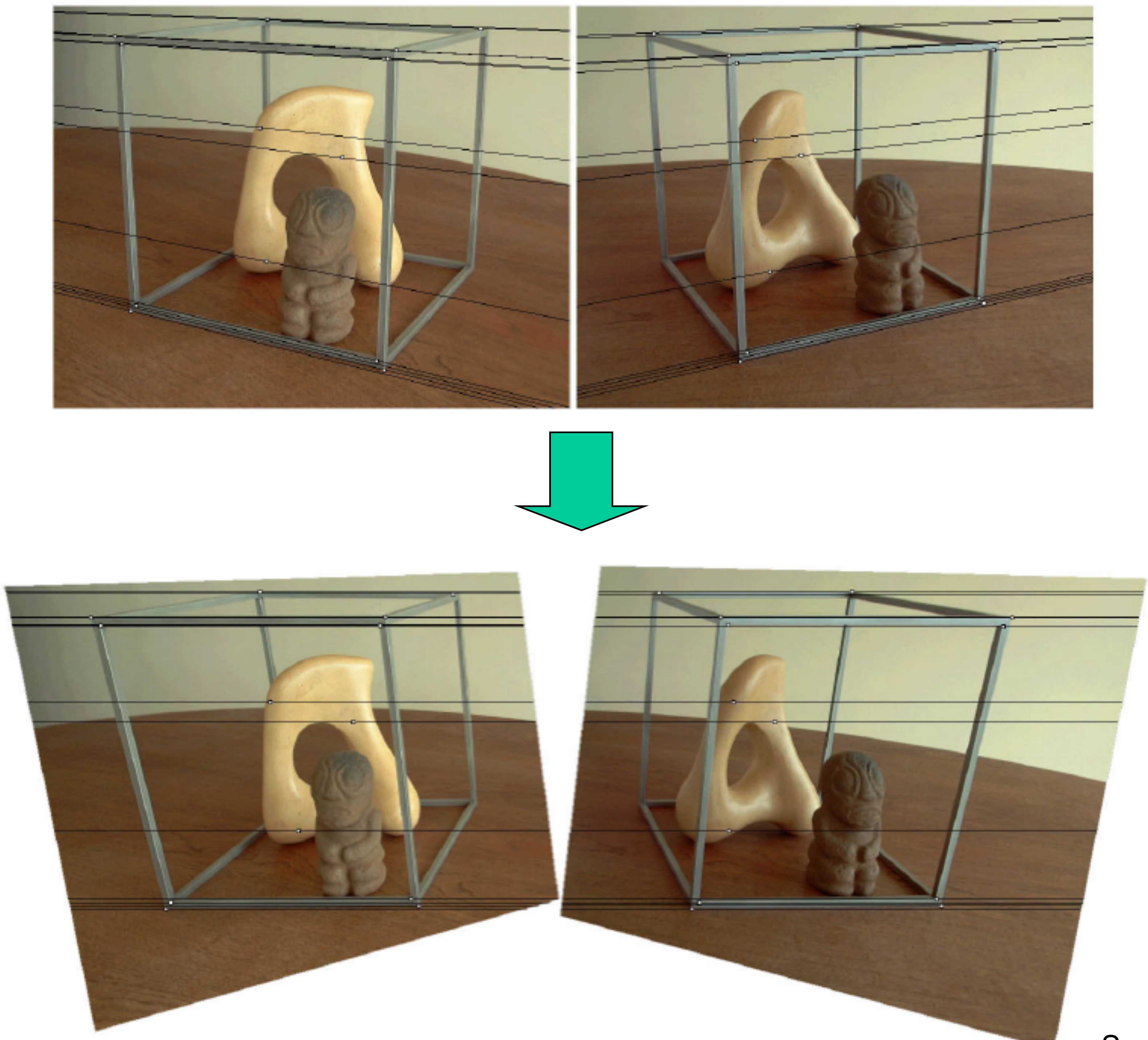
In practice, it is convenient if image scanlines (rows) are the epipolar lines.



- reproject image planes onto a common plane parallel to the line between optical centers
- pixel motion is horizontal after this transformation
- two homographies (3x3 transforms), one for each input image reprojection

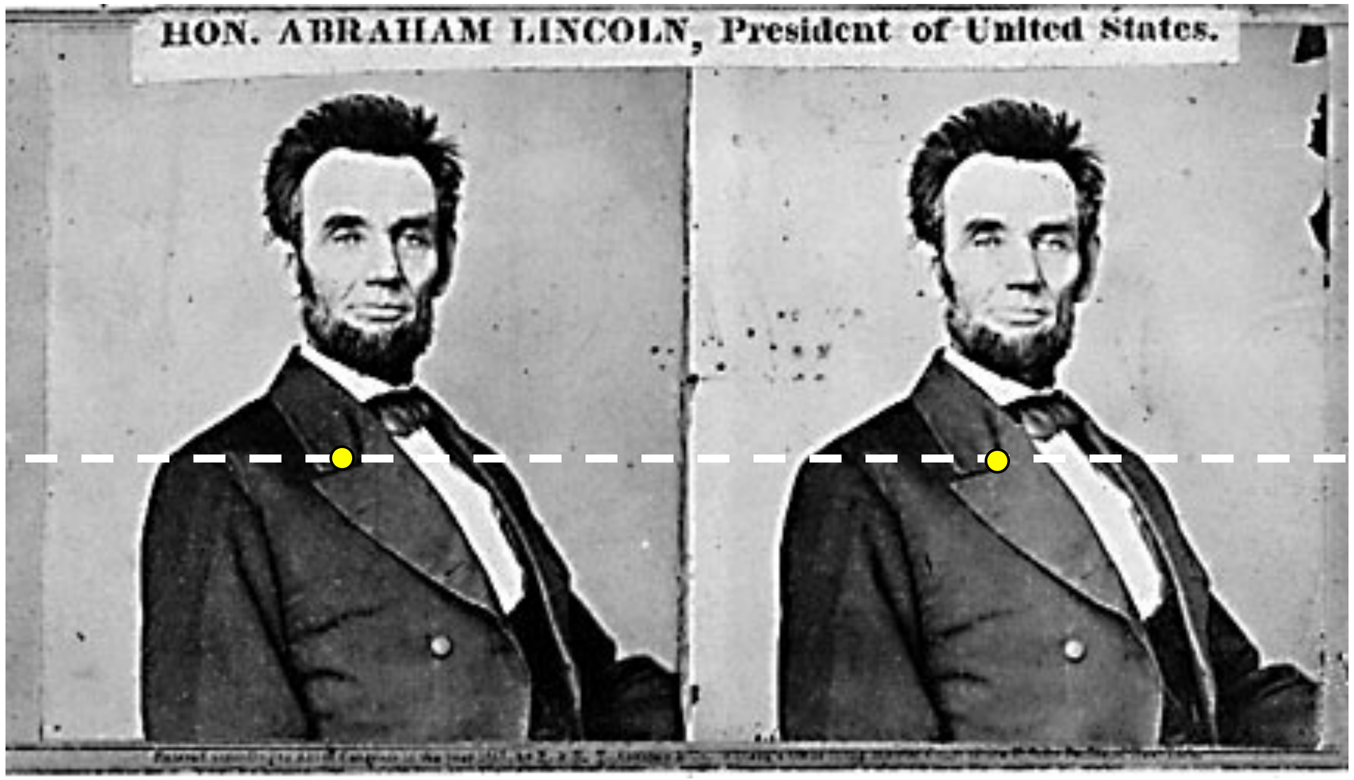
Slide credit: Li Zhang

Stereo image rectification: example



Source: Alyosha Efros

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

Worktime: Project 1