I, ROBOT
ISAAC ASIMOV

1950

Future Vision

EYE ROBOT
CSCI 1430
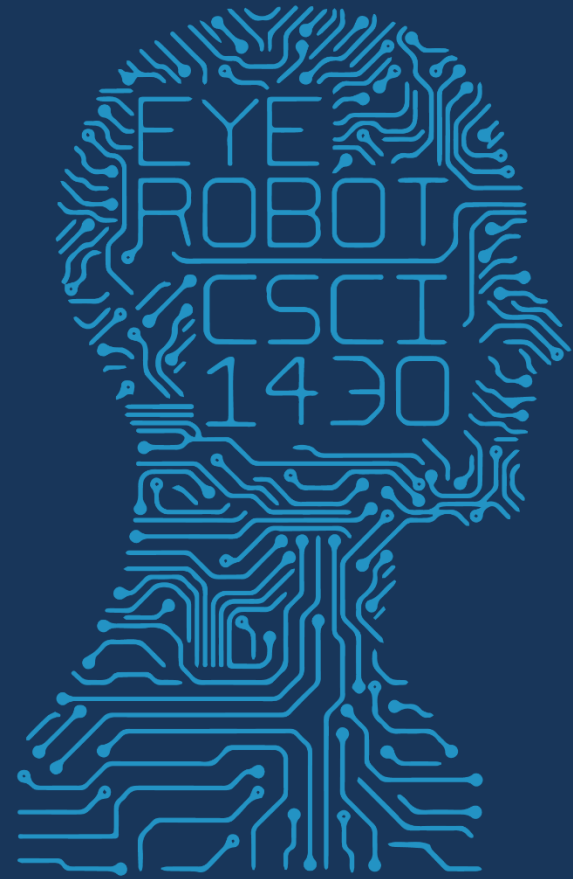
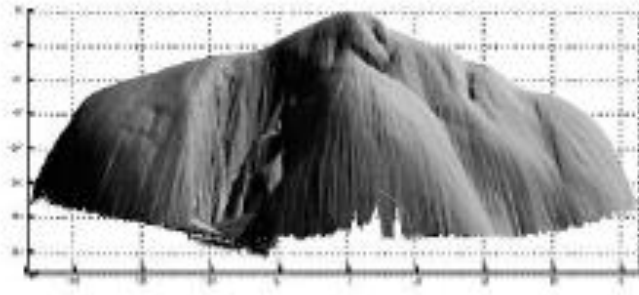2017 MWF 1PM 368

Computer Vision

Think-Pair-Share

What visual or physiological cues help us to perceive 3D shape and depth?

Shading

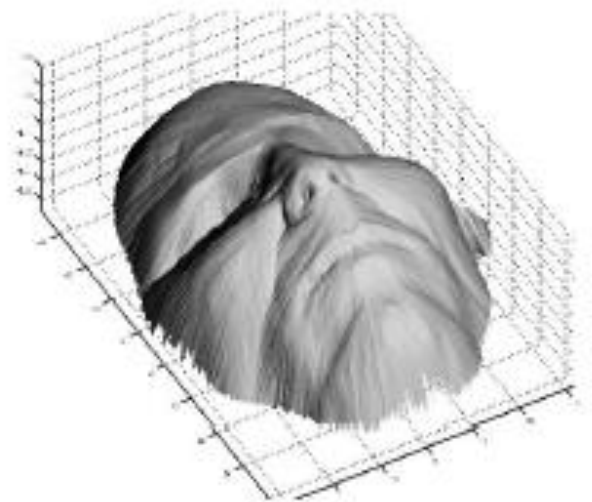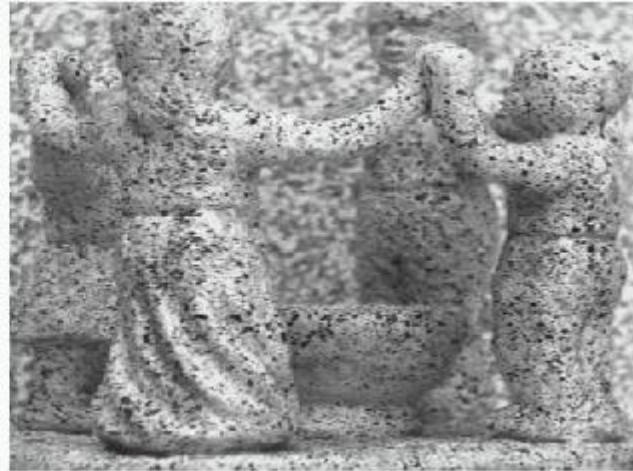
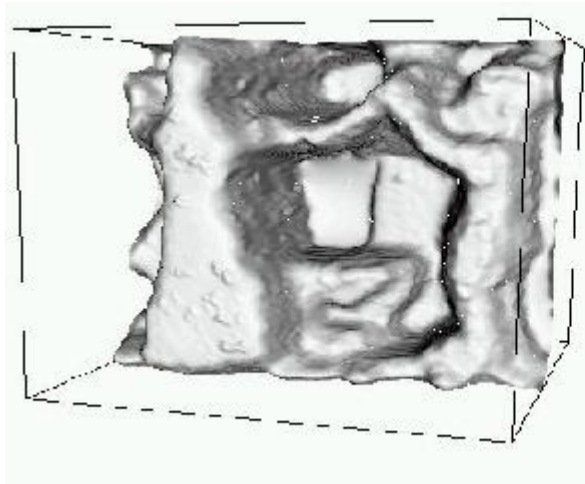a)   b)   c)
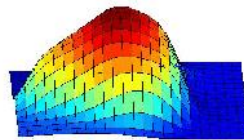
[Figure from Prados & Faugeras 2006]

Focus/defocus

Images from same point of view, different camera parameters

3d shape / depth estimates

[figs from H. Jin and P. Favaro, 2002]
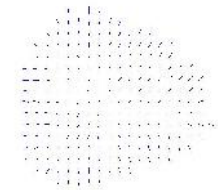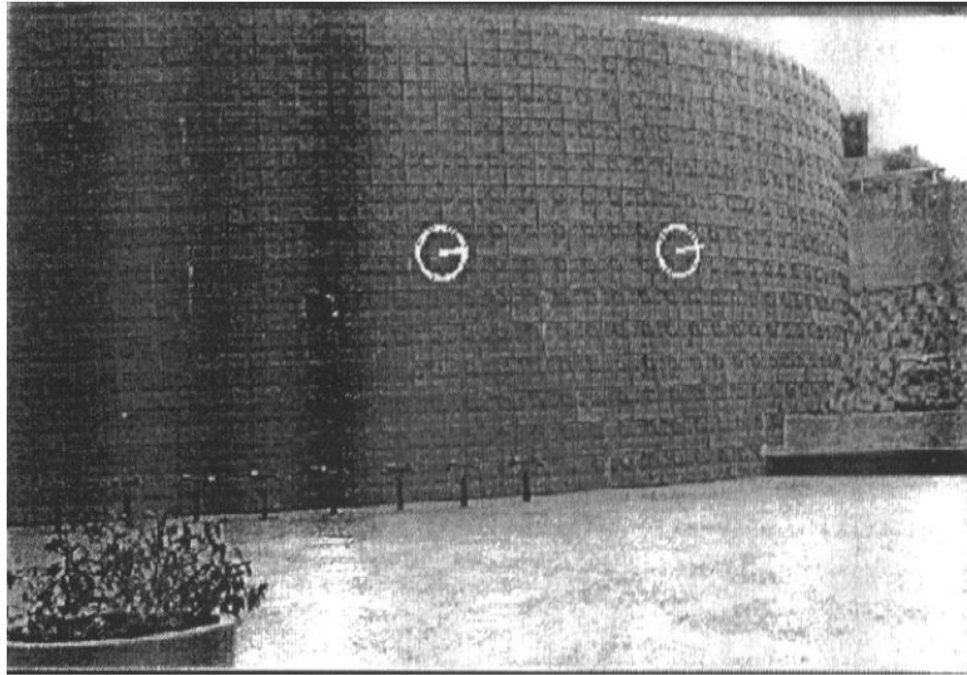
Texture

[From A.M. Loh. The recovery of 3-D structure using visual texture patterns. PhD thesis]
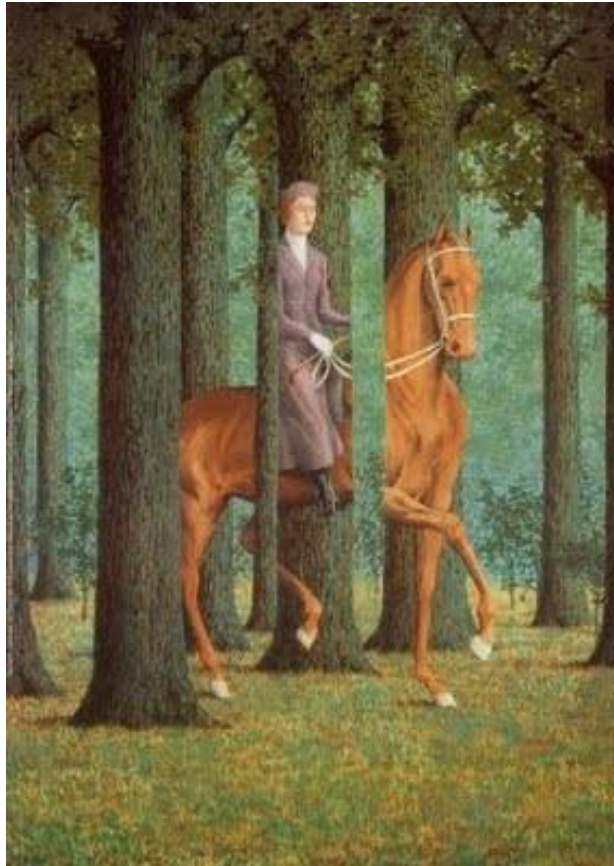
Perspective effects

Image credit: S. Seitz

Motion
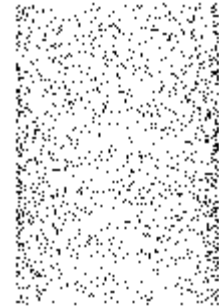
Occlusion

Rene Magritt'e famous painting *Le Blanc-Seing* (literal translation: "The Blank Signature") roughly translates as "free hand" or "free rein".

Stereo
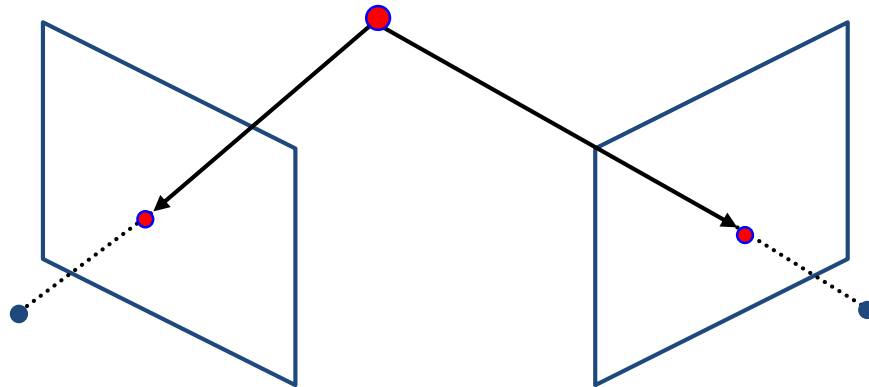
If stereo were critical for depth perception, navigation, recognition, etc., then rabbits would never have evolved.

Human stereopsis

FIGURE 7.1

From Bruce and Green, Visual Perception,
Physiology, Psychology and Ecology

Human eyes **fixate** on point in space – rotate so that
corresponding images form in centers of fovea.

Human stereopsis: disparity

FIGURE 7.3

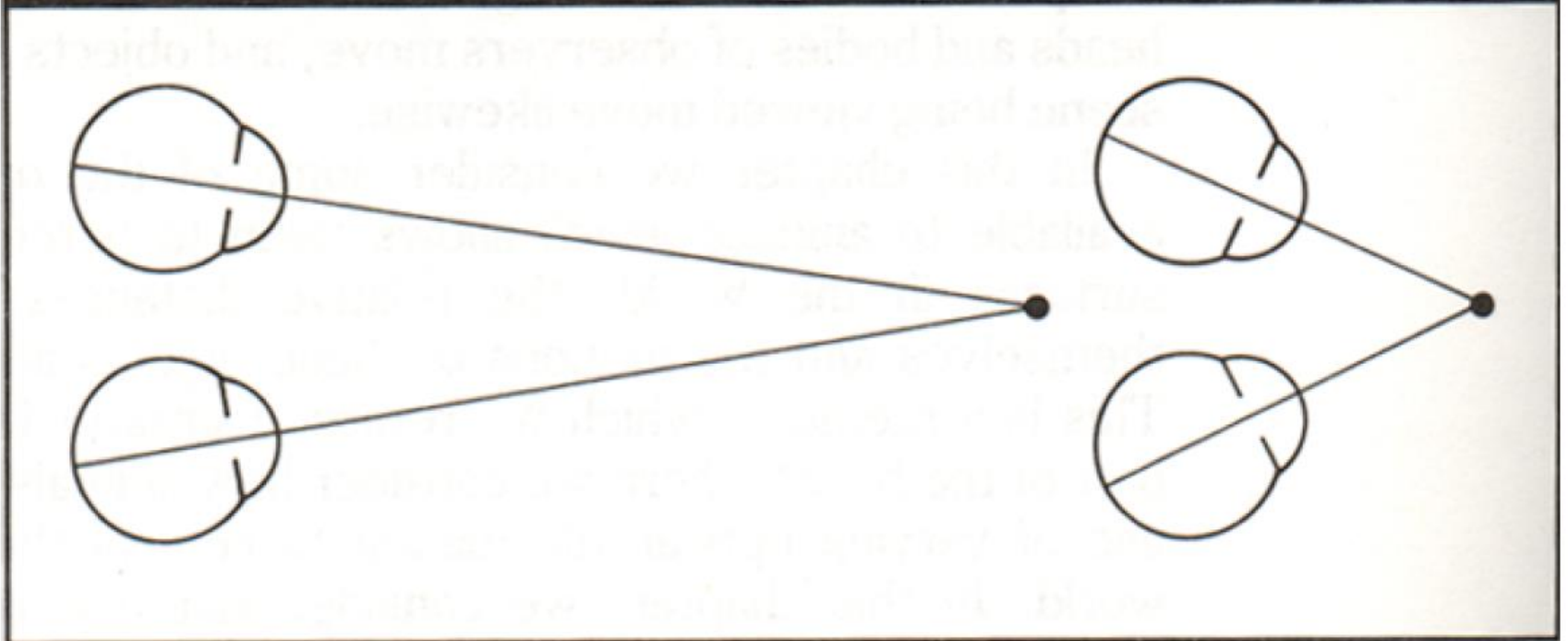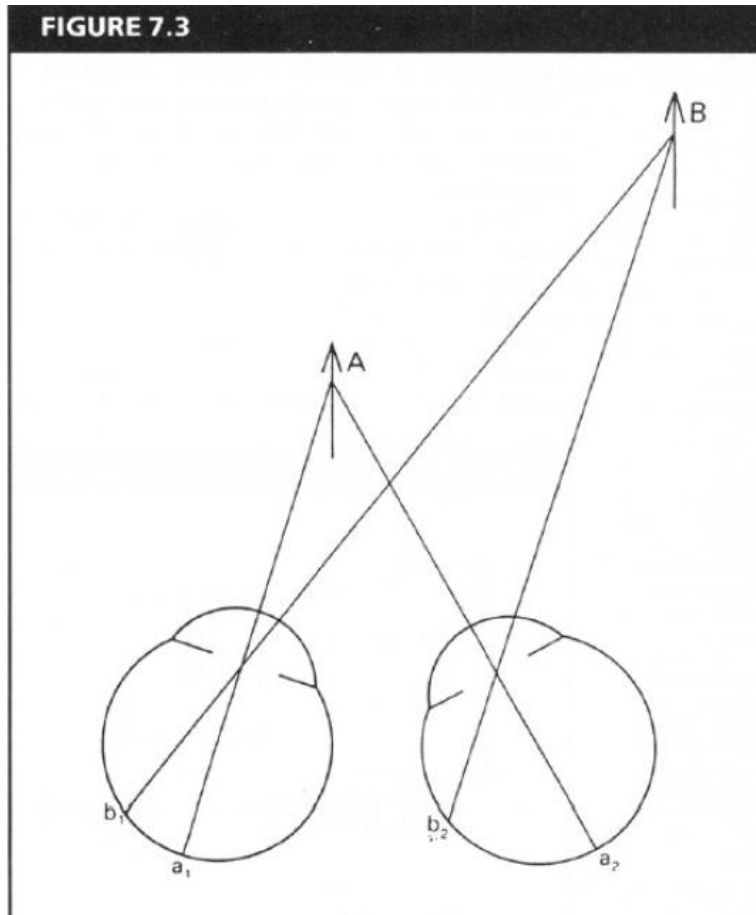From Bruce and Green, Visual Perception,
Physiology, Psychology and Ecology

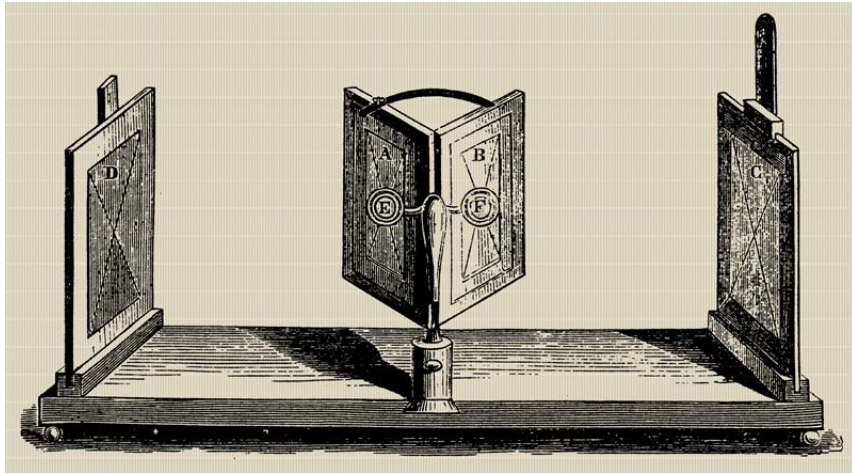**Disparity** occurs when eyes fixate on one object; others appear at different visual angles.

Disparity is distance from b1 to b2 along retina.

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

Most Famous Inventor of the Age, Thomas A. Edison In His Laboratory.

Keystone View Company

Copyrighted, Underwood & Underwood, Inc.

Manufacturers MADE IN U.S.A. Publishers

Meadville, Pa., New York, N. Y., Chicago, Ill., London, England.

45
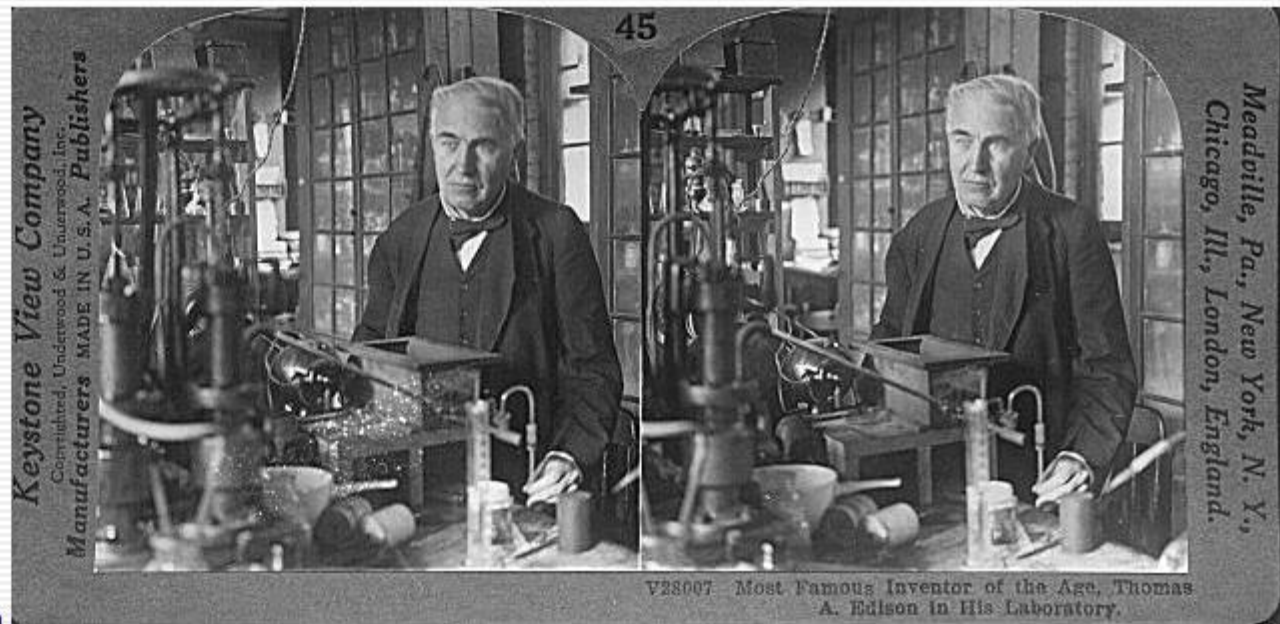
© Copyright 2001 Johnson-Shaw Stereoscopic Museum

http://www.johnsonshawmuseum.org

http://www.johnsonshawmuseum.org
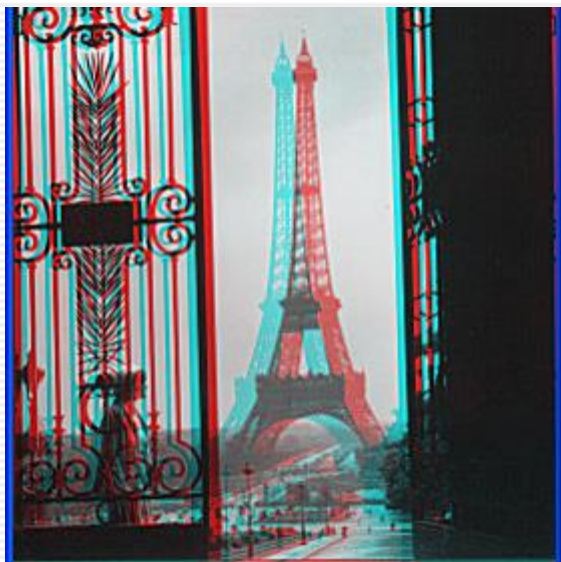
Wiggle images

http://www.well.com/~jimg/stereo/stereo_list.html

Stereo vision

Two cameras, simultaneous views

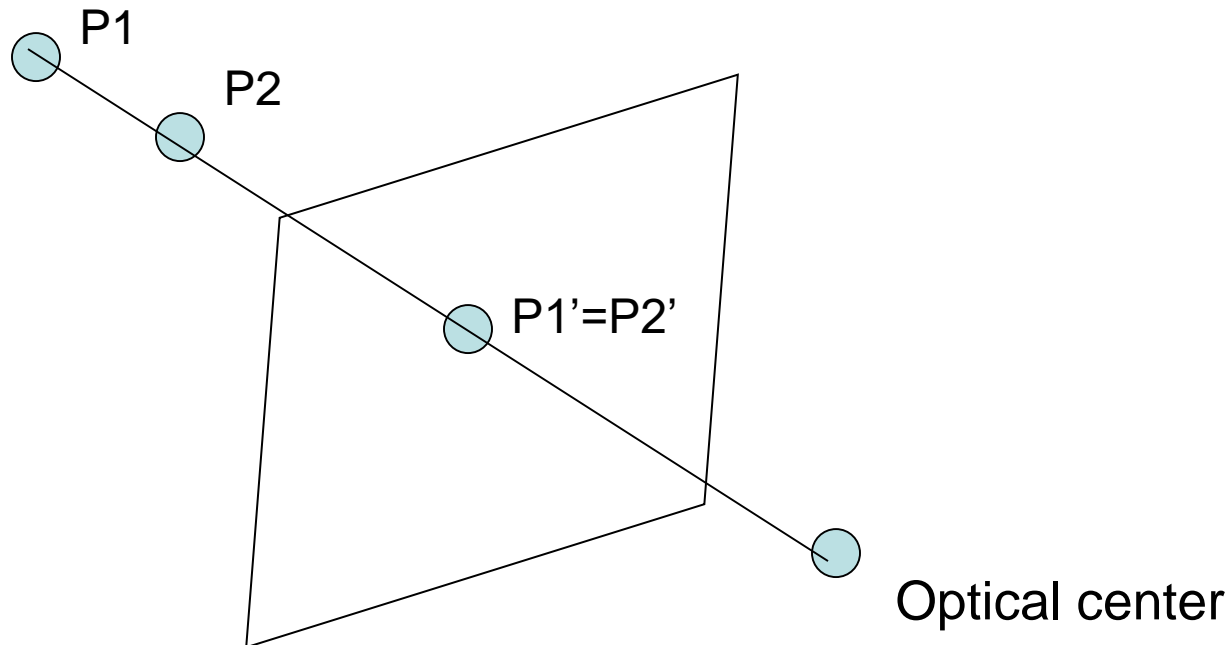Single moving camera and static scene

Why multiple views?

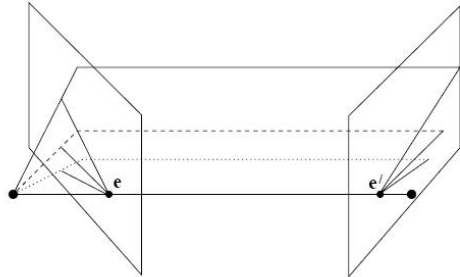Structure and depth can be ambiguous from single views...

Why multiple views?

Points at different depths along a line project to a single point

P1

P2

P1'=P2'

Optical center

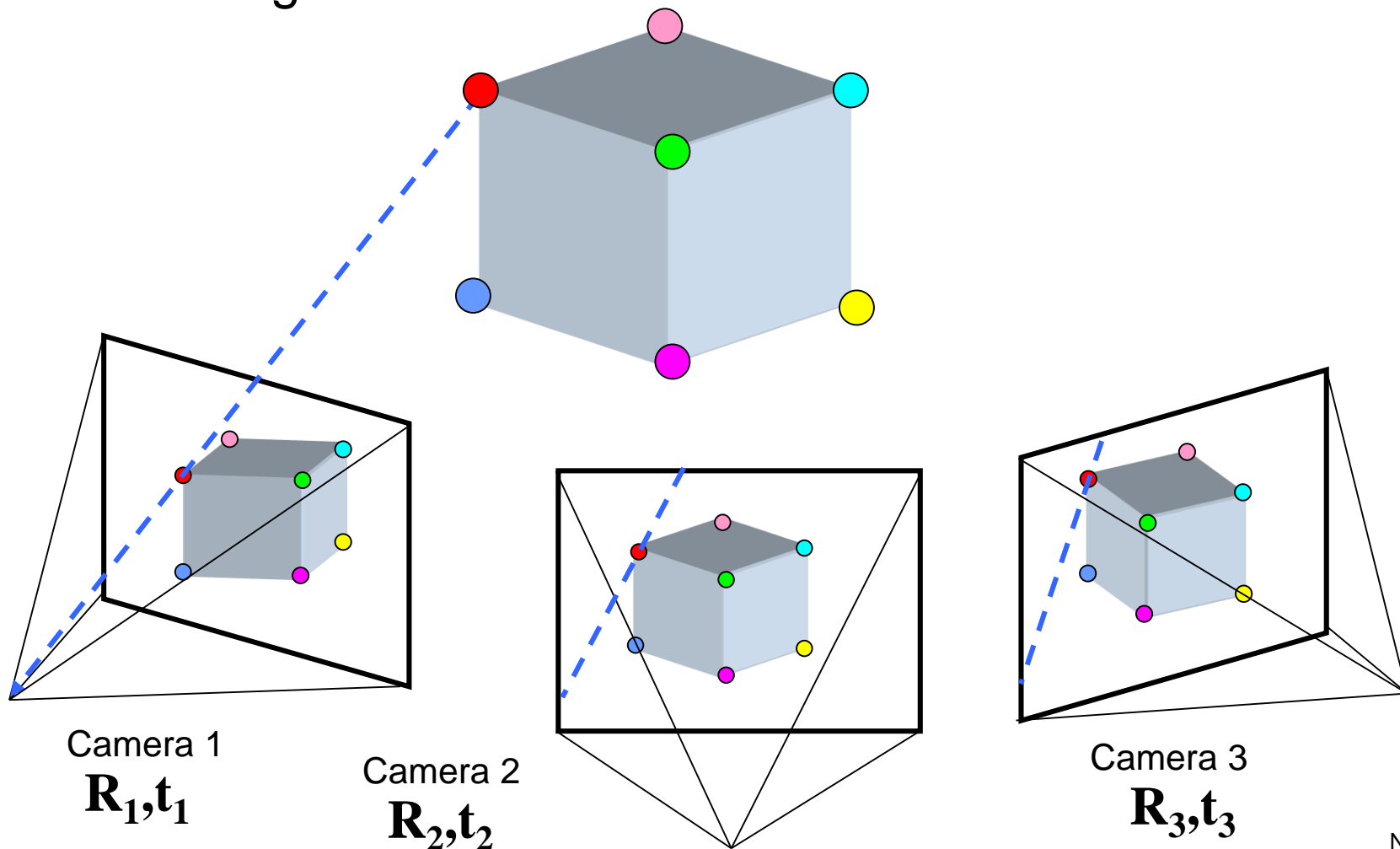# Multiple views
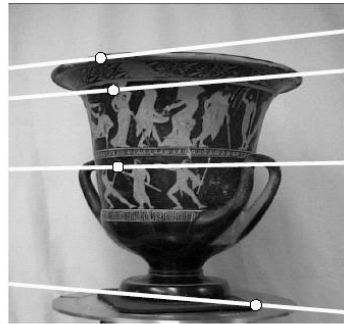
Hartley and Zisserman

Stereo vision
Structure from motion
Optical flow

Multi-view geometry problems

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

Camera 1
R_1, t_1

Camera 2
R_2, t_2

Camera 3
R_3, t_3

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
$\mathbf{R_1, t_1}$

Camera 2
$\mathbf{R_2, t_2}$
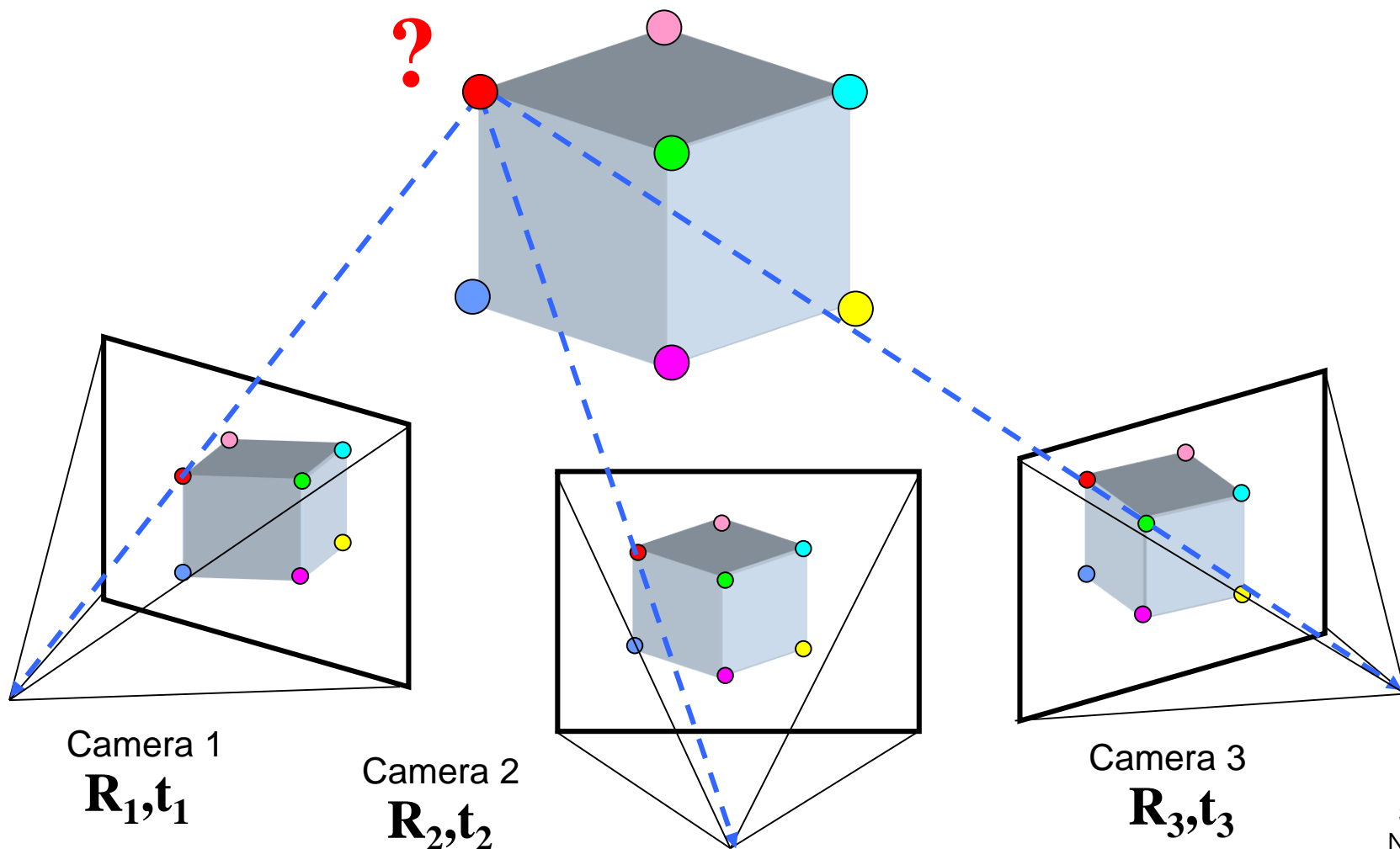
Camera 3
$\mathbf{R_3, t_3}$

Multi-view geometry problems

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

Camera 1
R_1, t_1 ?

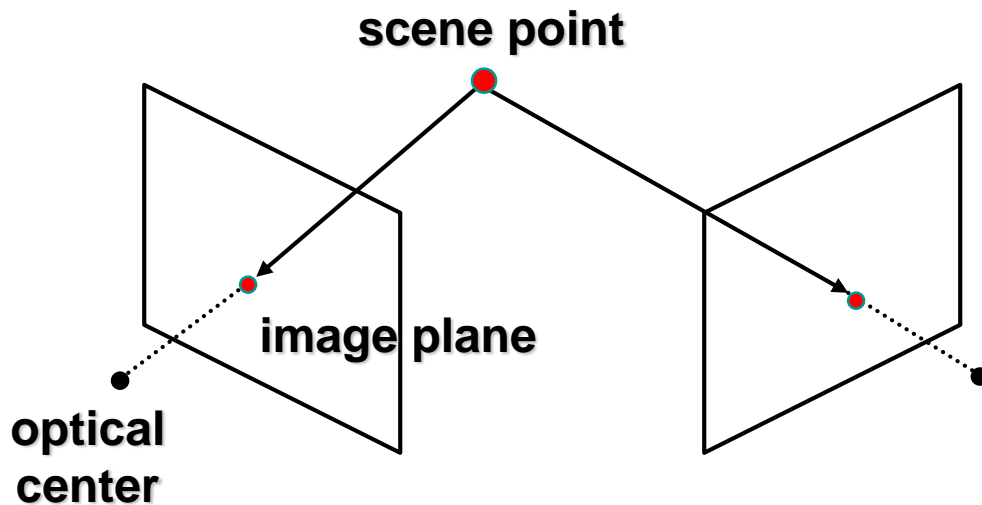Camera 2
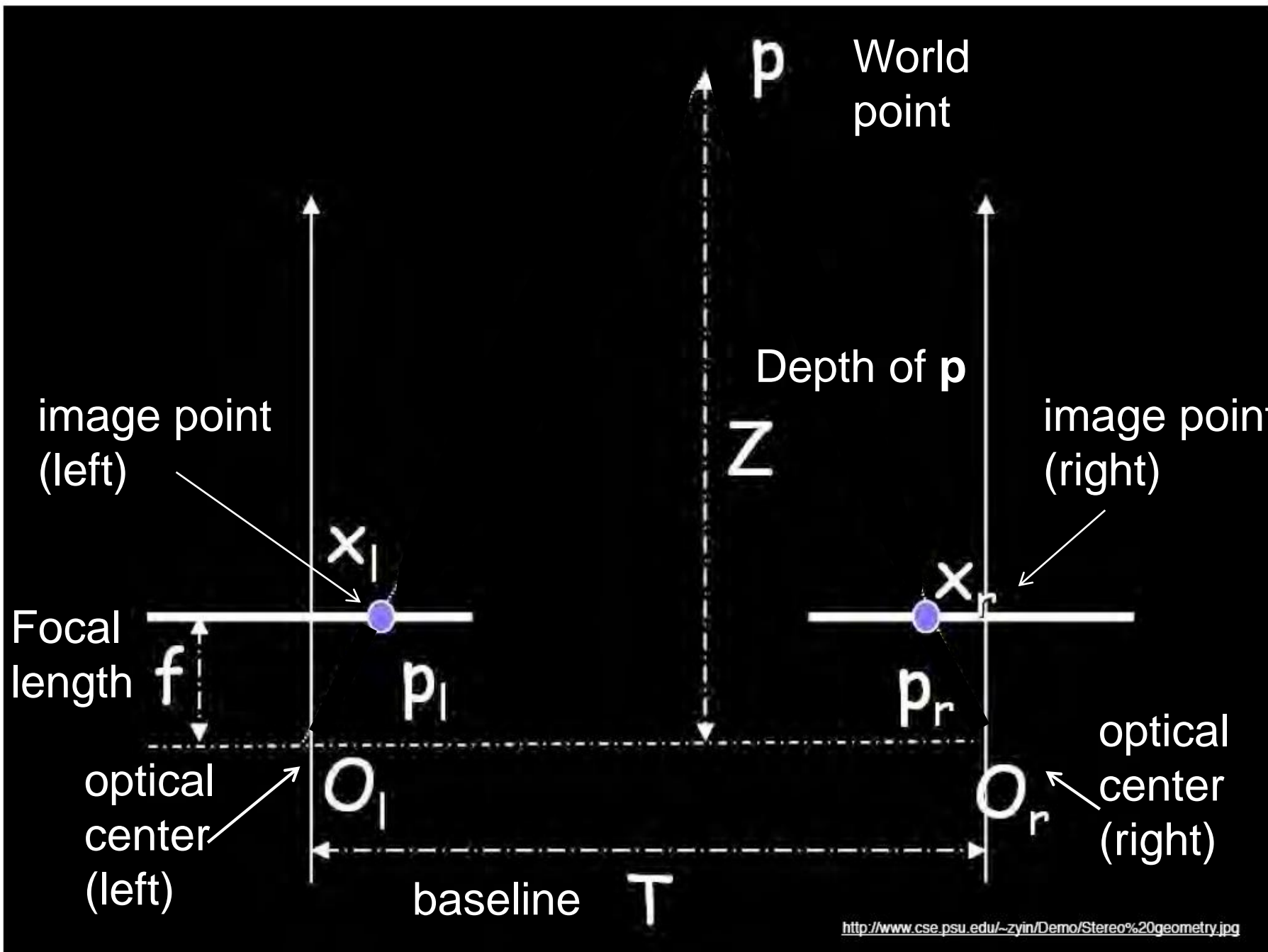R_2, t_2 ?
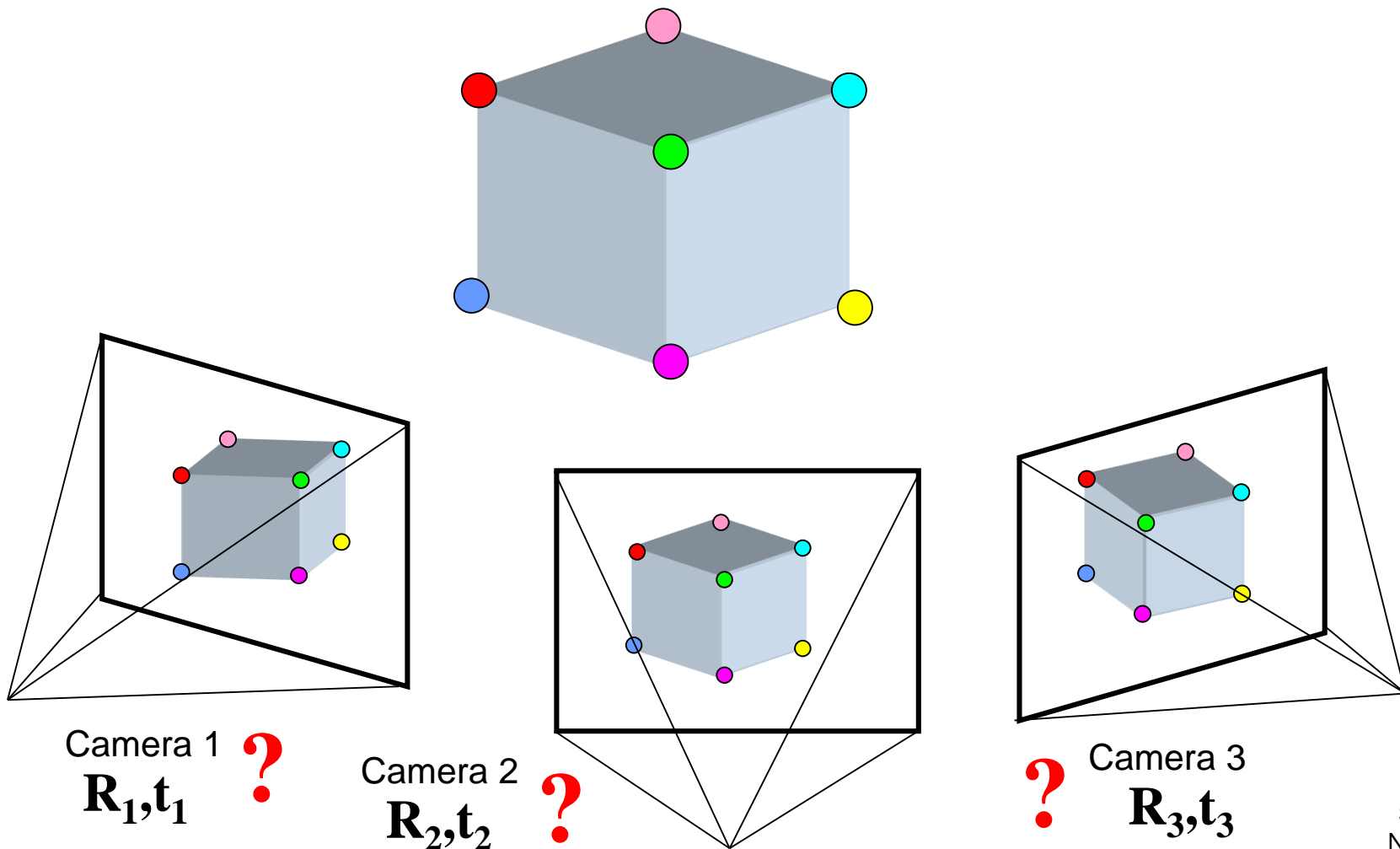
? Camera 3
R_3, t_3

Multi-view geometry problems

- **Optical flow:** Given two images, find the location of a world point in a second close-by image with no camera info.
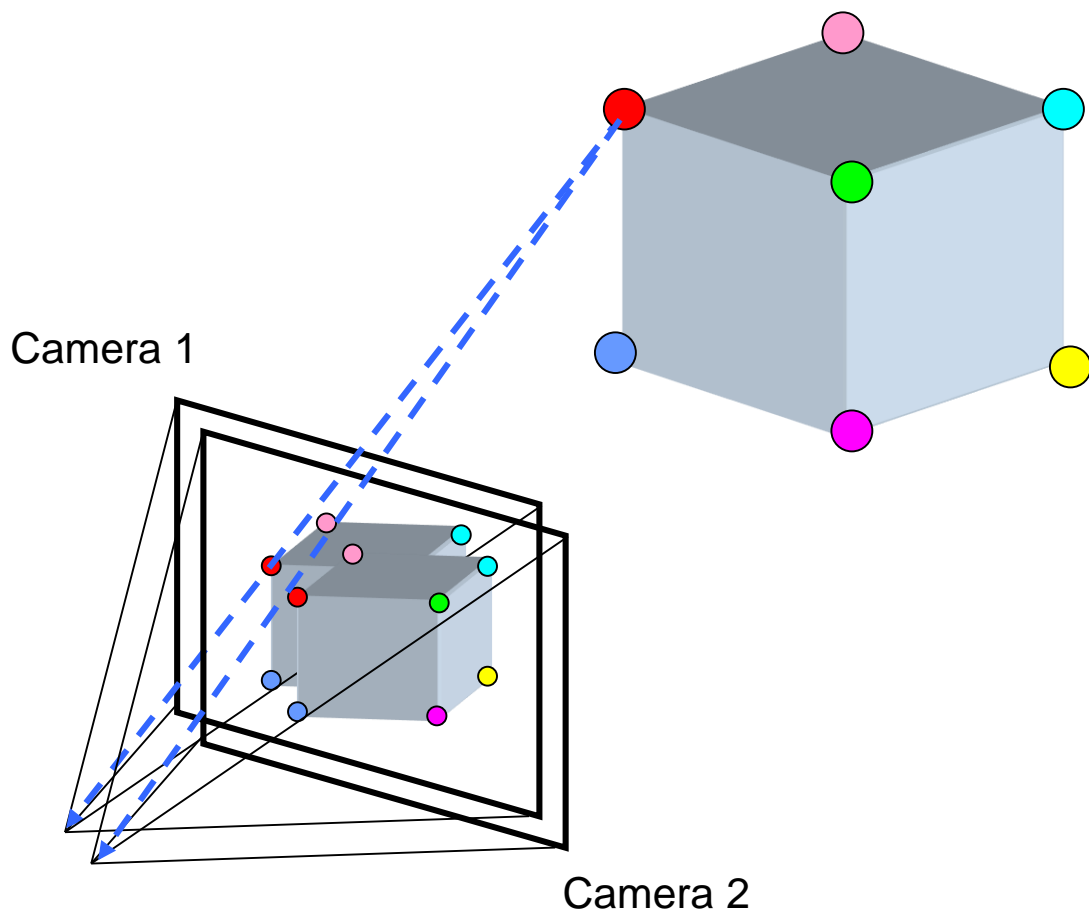
Camera 1

Camera 2

Estimating depth with stereo

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

p World point

Depth of **p**

Z

image point (left)

x_l

image point (right)

x_r

Focal length f

p_l

p_r

optical center (left)

O_l

O_r

optical center (right)

baseline T

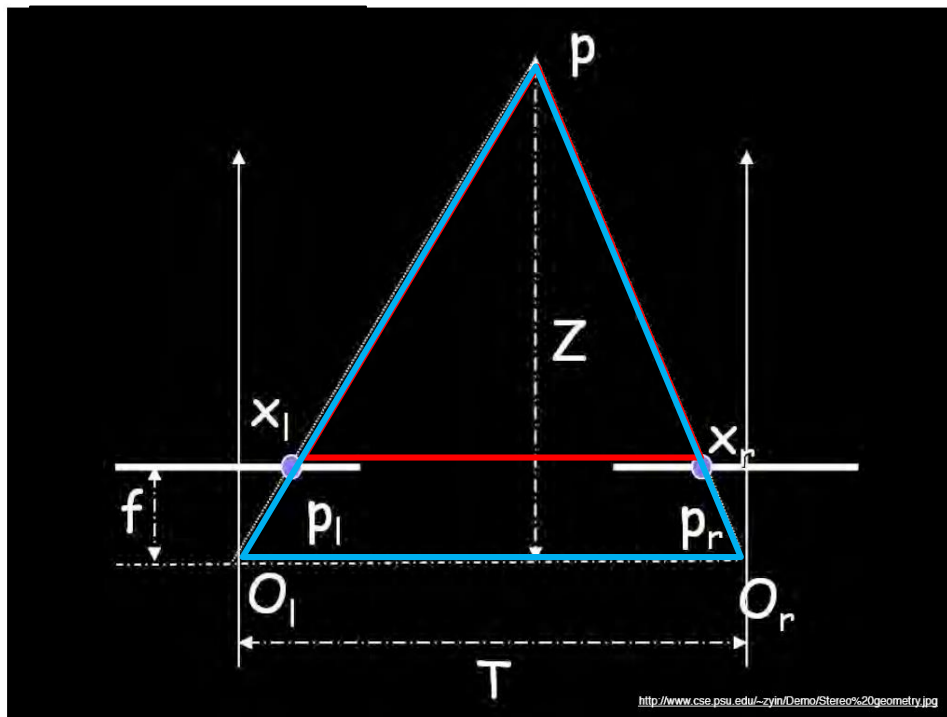http://www.cse.psu.edu/~zyin/Demo/Stereo%20geometry.jpg

Geometry for a simple stereo system

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

http://www.cse.psu.edu/~zyin/Demo/Stereo%20geometry.jpg

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´)
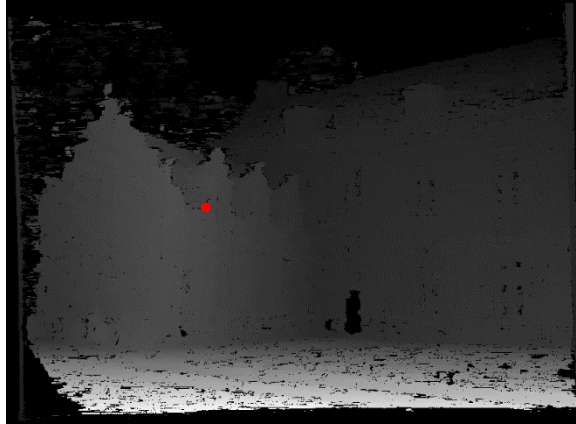

$$(x´,y´)=(x+D(x,y), y)$$

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