# Project 4 Results

- Representation
  - SIFT and HoG are popular and successful.
- Data
  - Hugely varying results from hard mining.
- Learning
  - Non-linear classifier usually better.

Zachary, Hung-I, Paul, Emanuel

# Project 5

- Project 5 handout

# Final Project

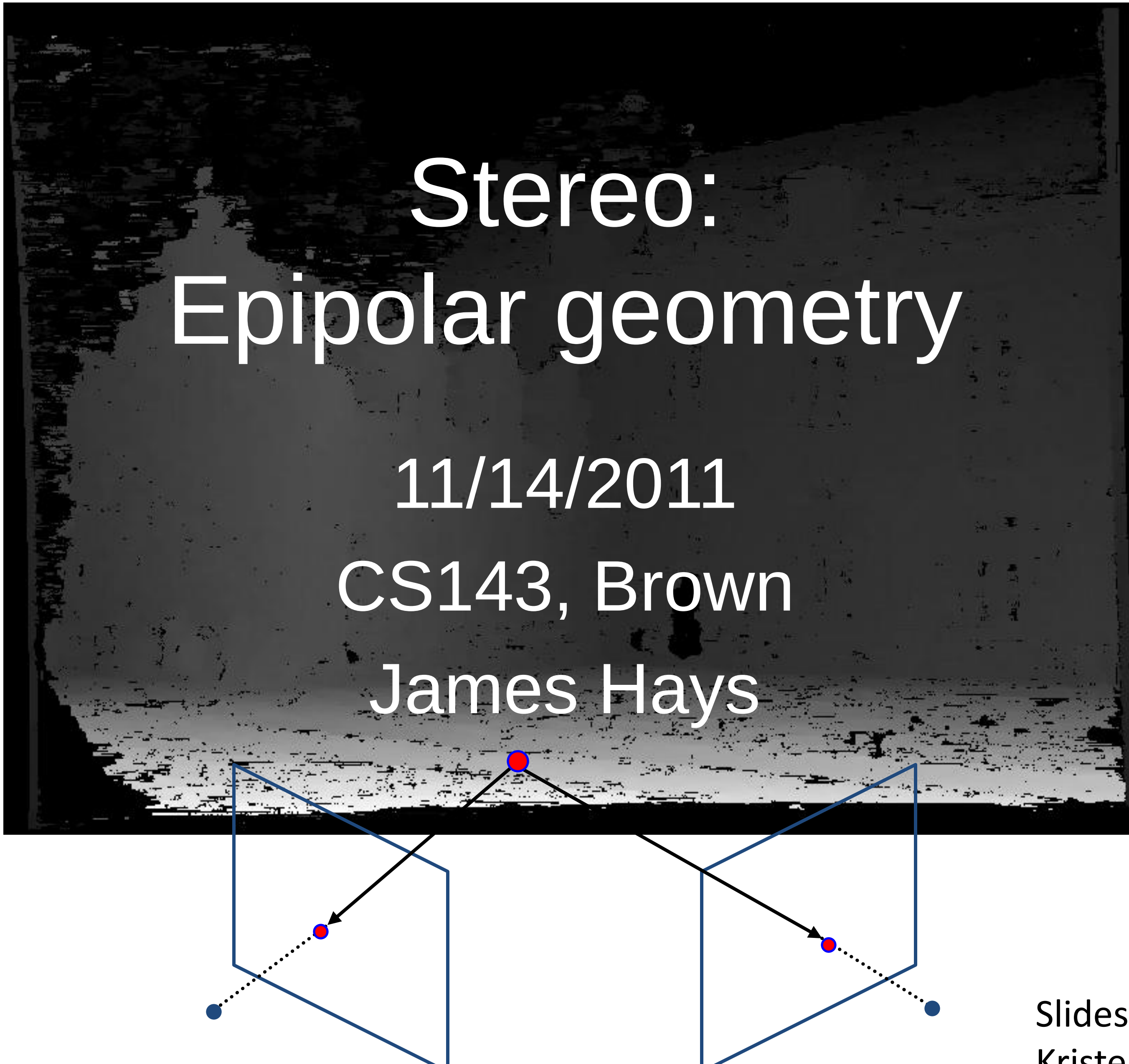
- Proposals due Friday.
  - One paragraph with one paper link.
- Can opt out of project 5.
- Higher expectations; must be able to present during final exam slot.

# Stereo: Epipolar geometry

11/14/2011

CS143, Brown

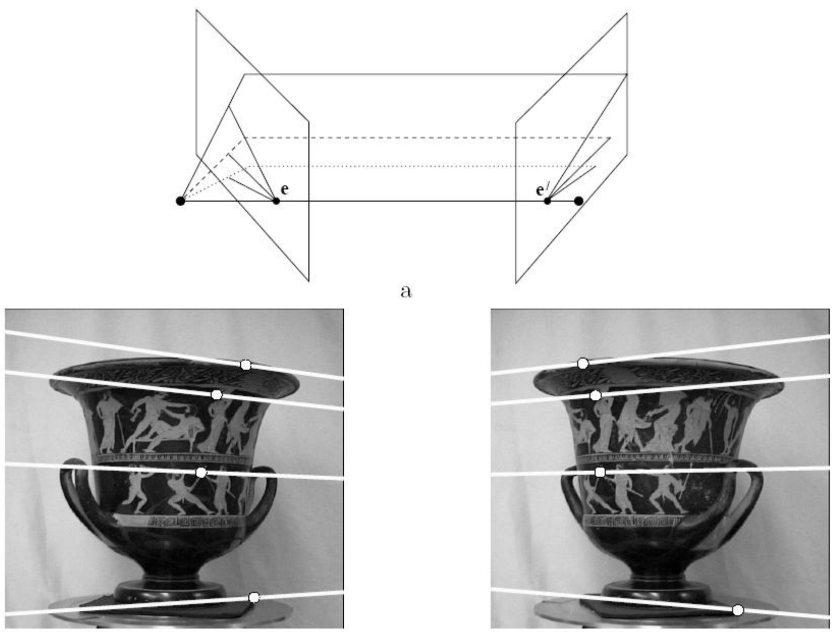
James Hays



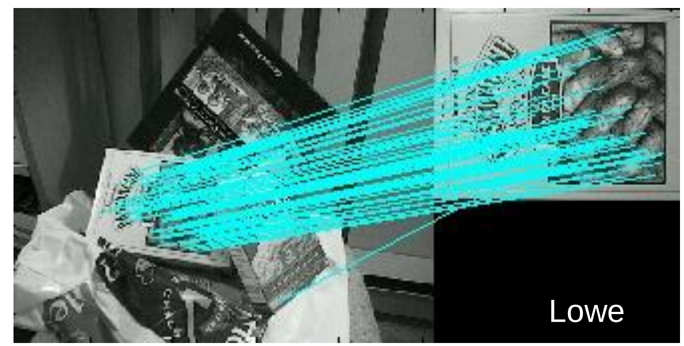
Slides by
Kristen Grauman

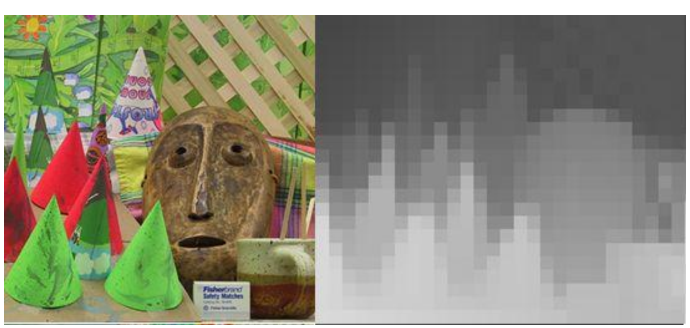
# Multiple views

Multi-view geometry, matching, invariant features, stereo vision




Hartley and Zisserman



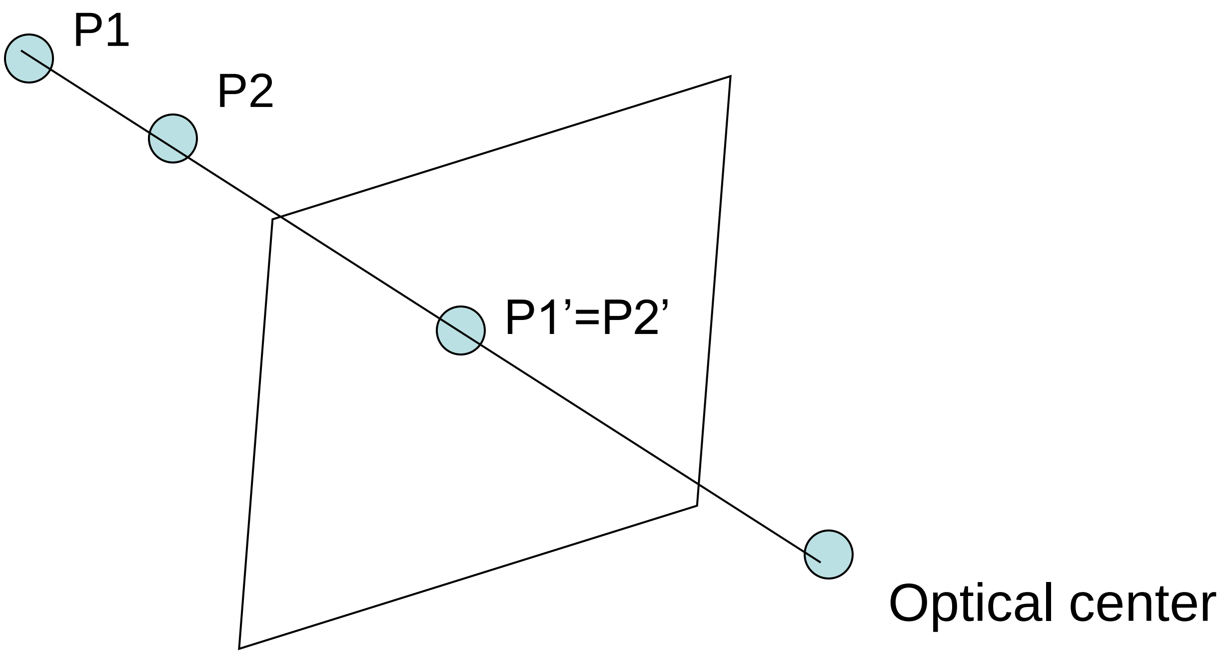
# Why multiple views?

- Structure and depth are inherently ambiguous from single views.

Images from Lana Lazebnik

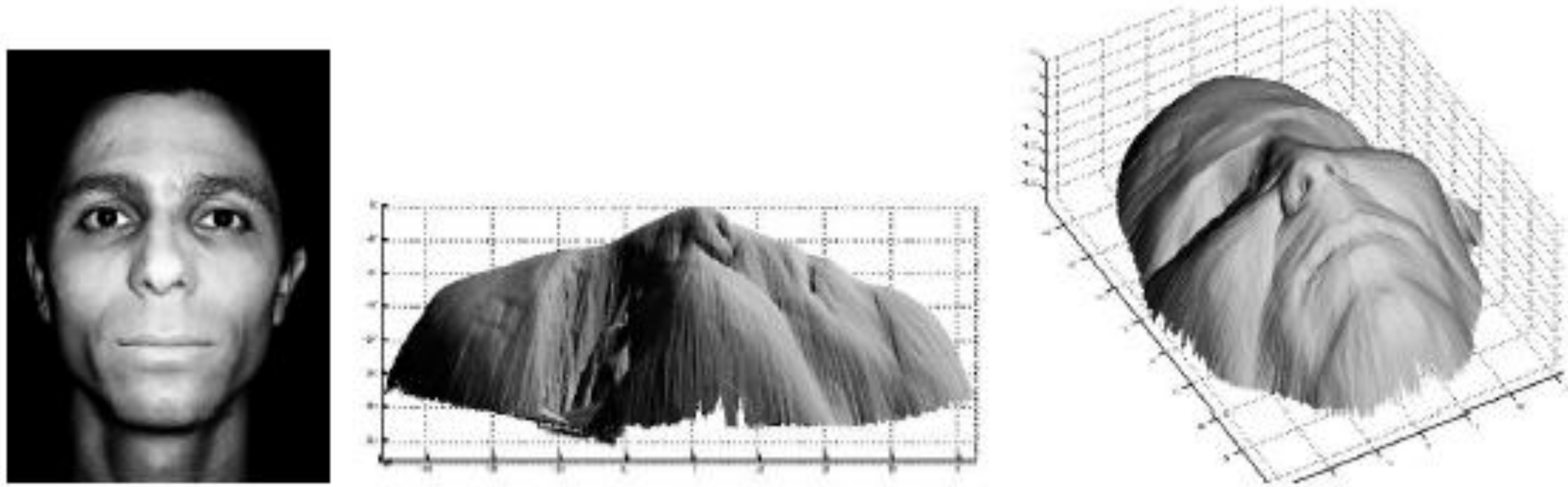
# Why multiple views?

- Structure and depth are inherently ambiguous from single views.



- What cues help us to perceive 3d shape and depth?

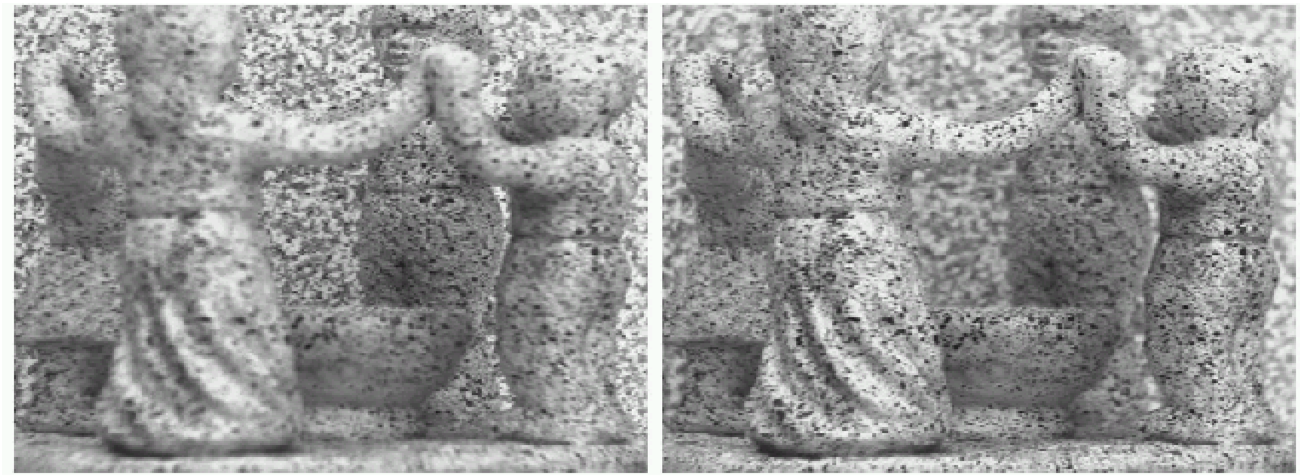
# Shading



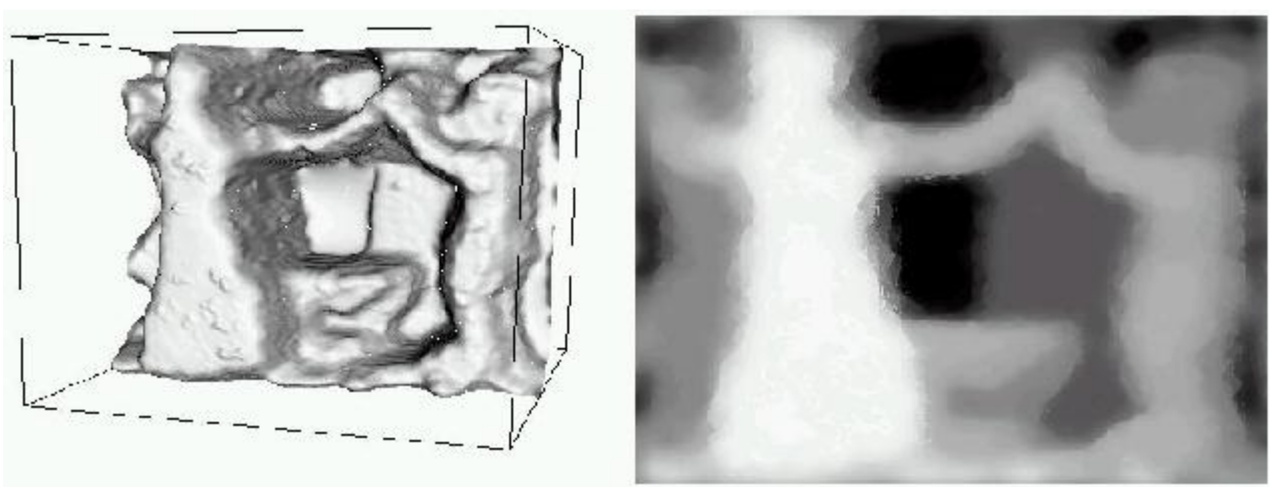
a) b) c)

[Figure from Prados & Faugeras 2006]

# Focus/defocus



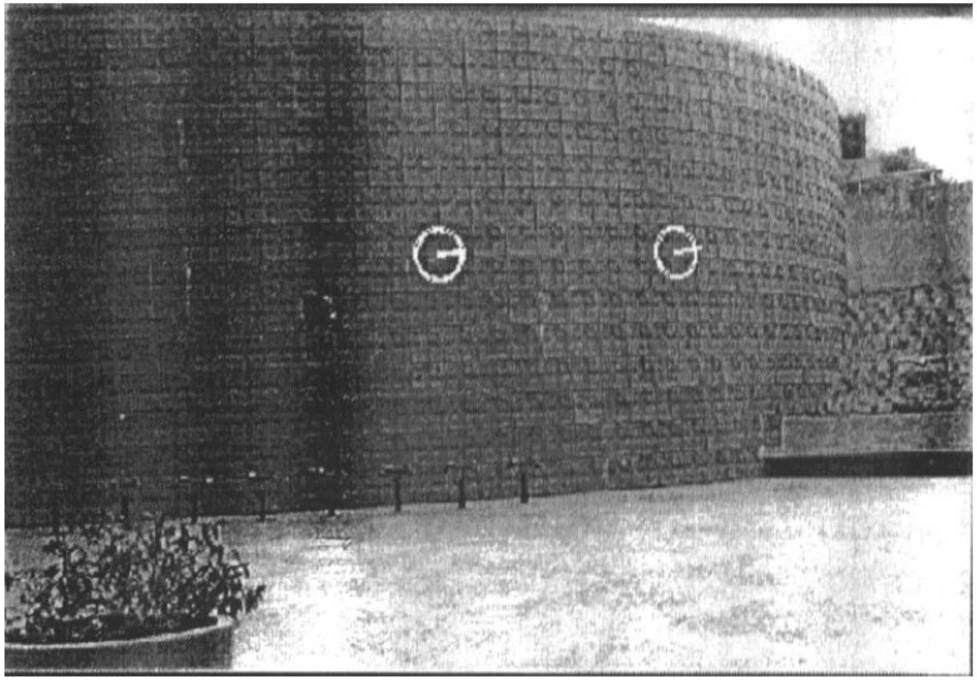
Images from same point of view, different camera parameters

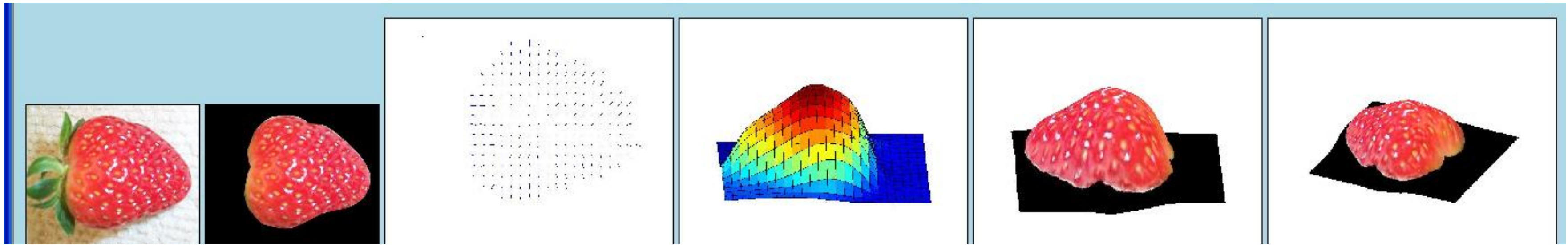


3d shape / depth estimates

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

# Texture



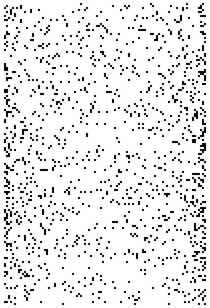


[From A.M. Loh. The recovery of 3-D structure using visual texture patterns. PhD thesis]

# Perspective effects


Image credit: S. Seitz

# Motion



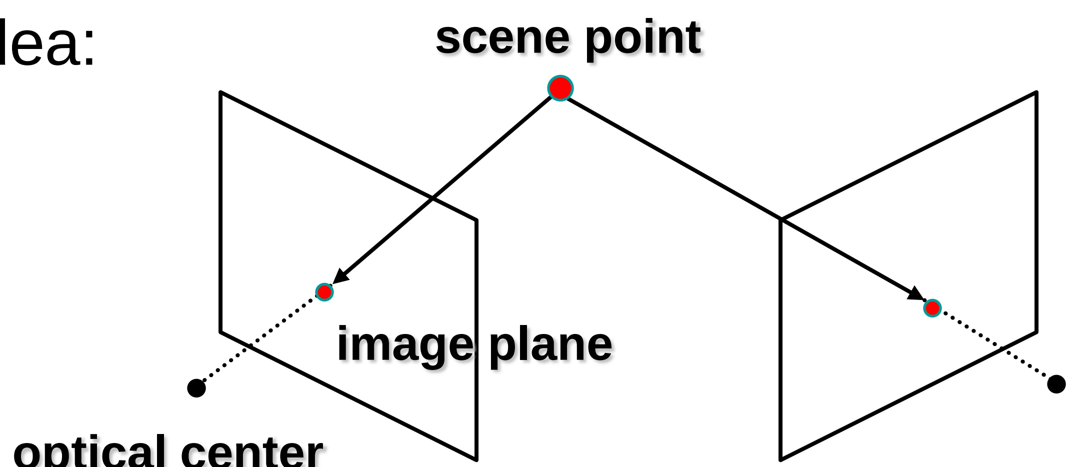
Figures from L. Zhang

http://www.brainconnection.com/teasers/?main=illusion/motion-shape

# Estimating scene shape

- “Shape from X”: Shading, Texture, Focus, Motion…
- **Stereo**:
  - shape from “motion” between two views
  - infer 3d shape of scene from two (multiple) images from different viewpoints

Main idea:

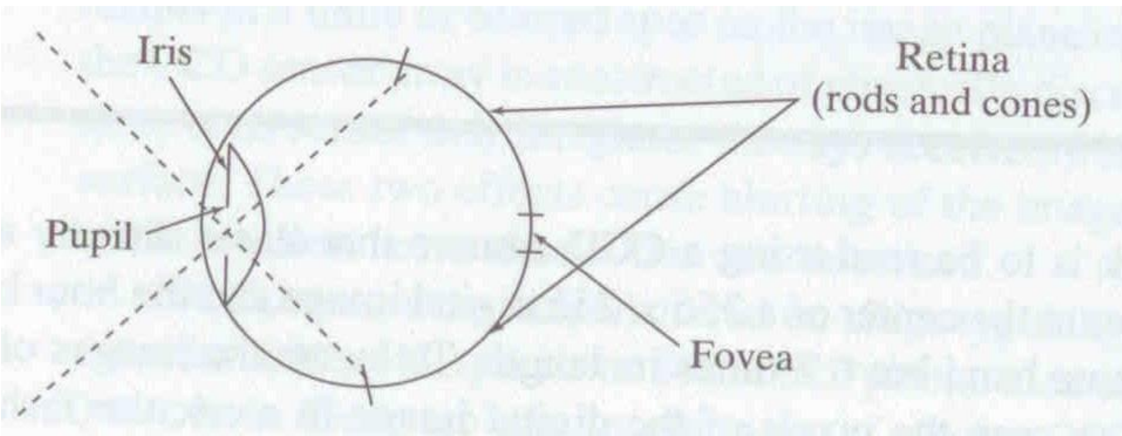


# Outline

- Human stereopsis
- Stereograms
- Epipolar geometry and the epipolar constraint
  - Case example with parallel optical axes
  - General case with calibrated cameras

# Human eye

Rough analogy with human visual system:



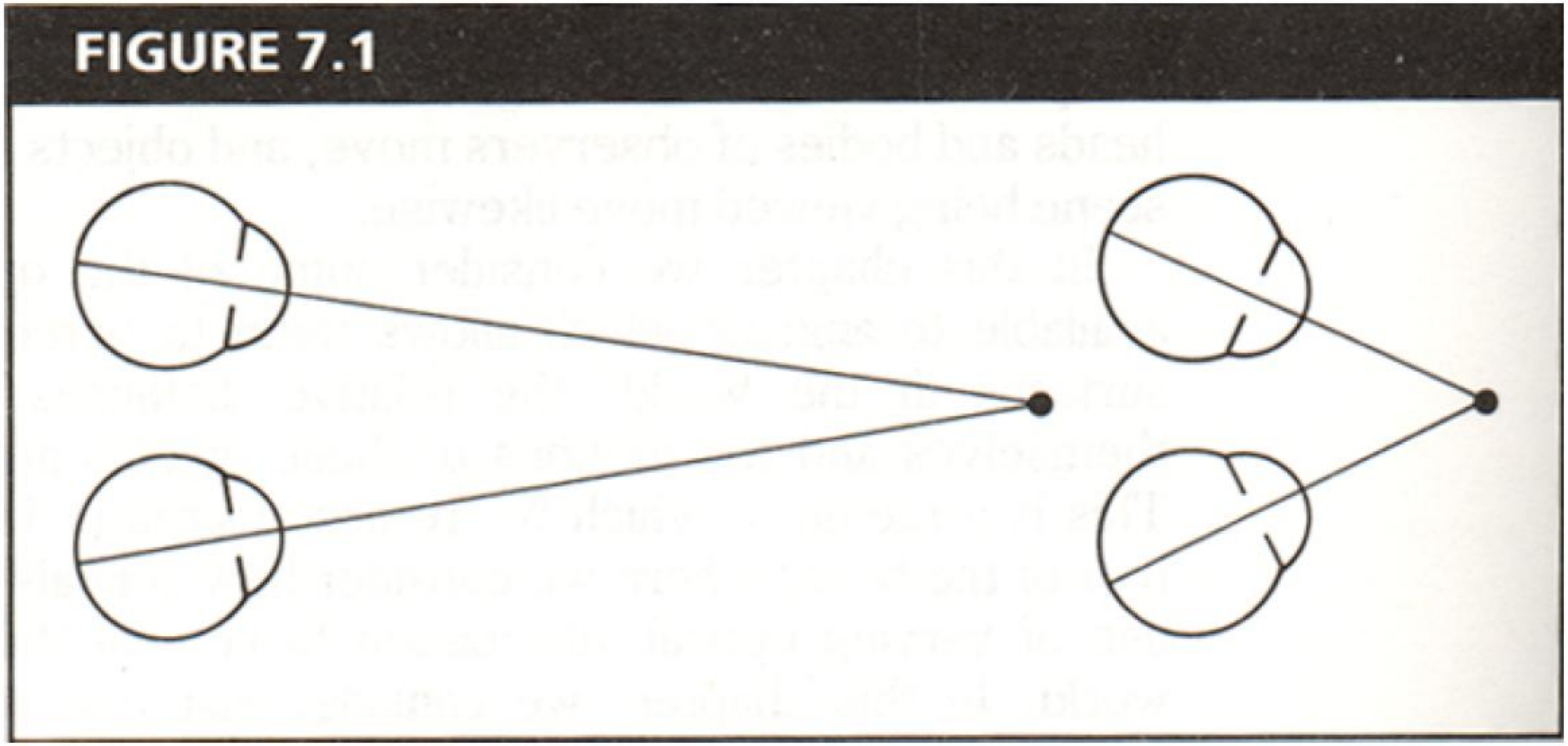

Pupil/Iris – control amount of light passing through lens

Retina - contains sensor cells, where image is formed

Fovea – highest concentration of cones

Fig from Shapiro and Stockman

# Human stereopsis: disparity

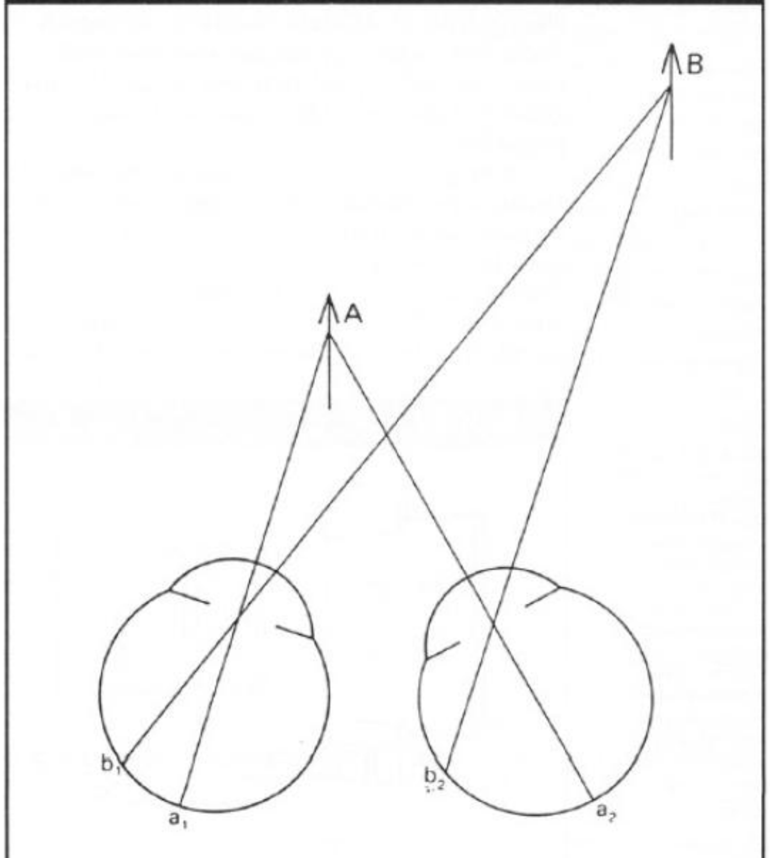



From Bruce and Green, Visual Perception, Physiology, Psychology and Ecology

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

# Human stereopsis: disparity

FIGURE 7.3

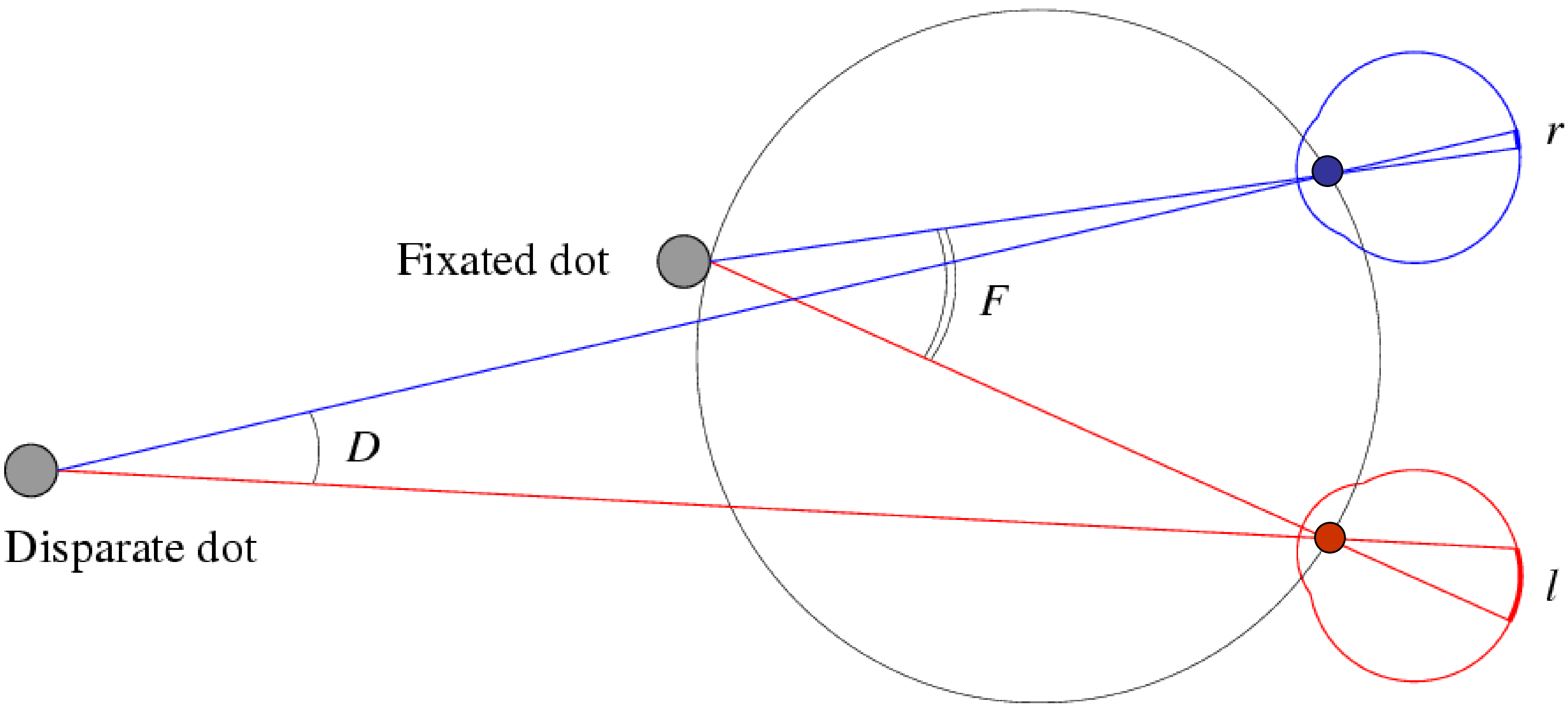



From Bruce and Green, Visual Perception, Physiology, Psychology and Ecology

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

Adapted from David Forsyth, UC Berkeley

# Human stereopsis: disparity




Disparity: $d = r - l = D - F.$

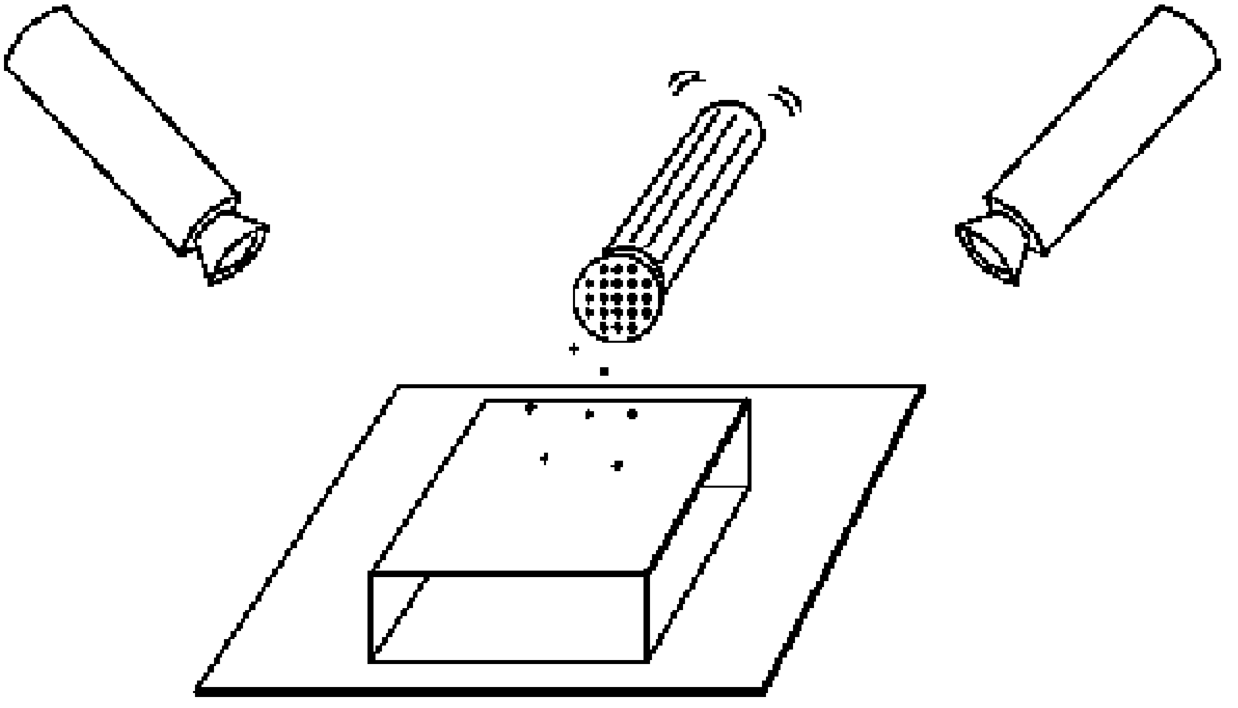
Forsyth & Ponce

# Random dot stereograms

- Julesz 1960: Do we identify local brightness patterns before fusion (monocular process) or after (binocular)?

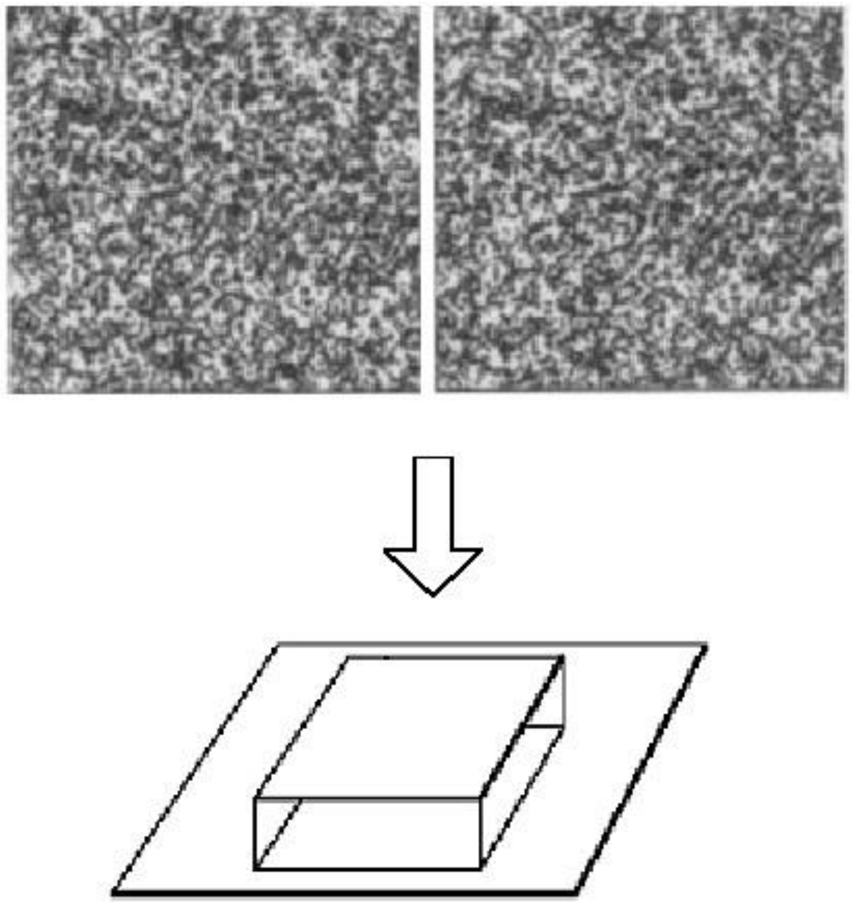
- To test: pair of synthetic images obtained by randomly spraying black dots on white objects

# Random dot stereograms



Forsyth & Ponce

# Random dot stereograms



# Random dot stereograms

- When viewed monocularly, they appear random; when viewed stereoscopically, see 3d structure.

- Conclusion: human binocular fusion not directly associated with the physical retinas; must involve the central nervous system
- Imaginary "*cyclopean retina*" that combines the left and right image stimuli as a single unit

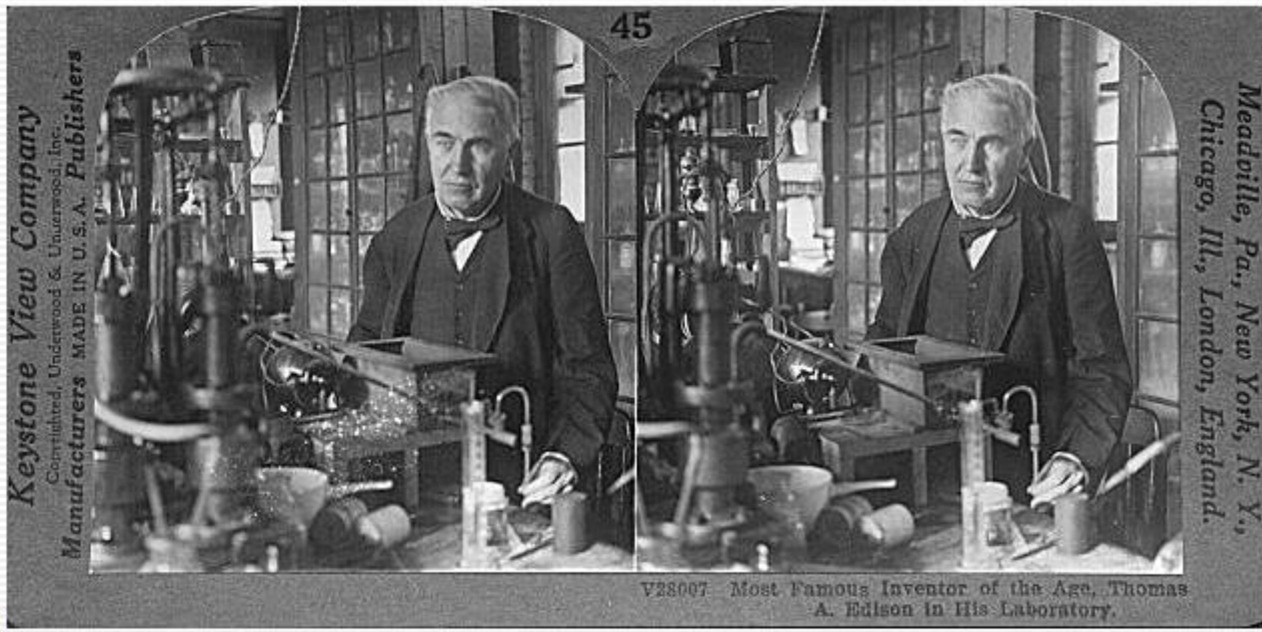
# Stereo photography and stereo viewers

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

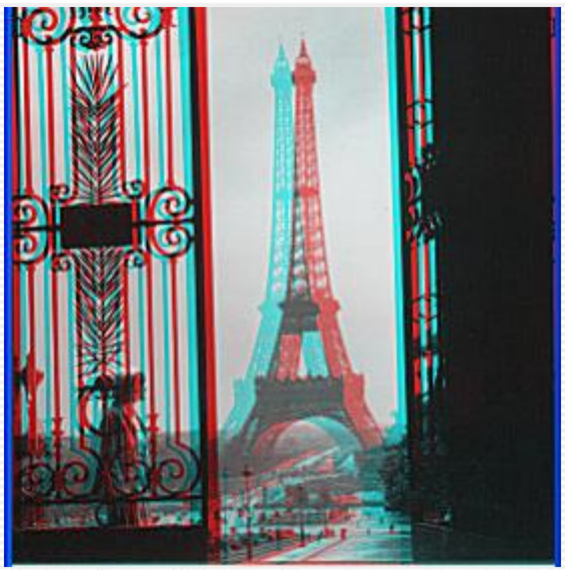
Invented by Sir Charles Wheatstone, 1838

Image from fisher-price.com






http://www.johnsonshawmuseum.org



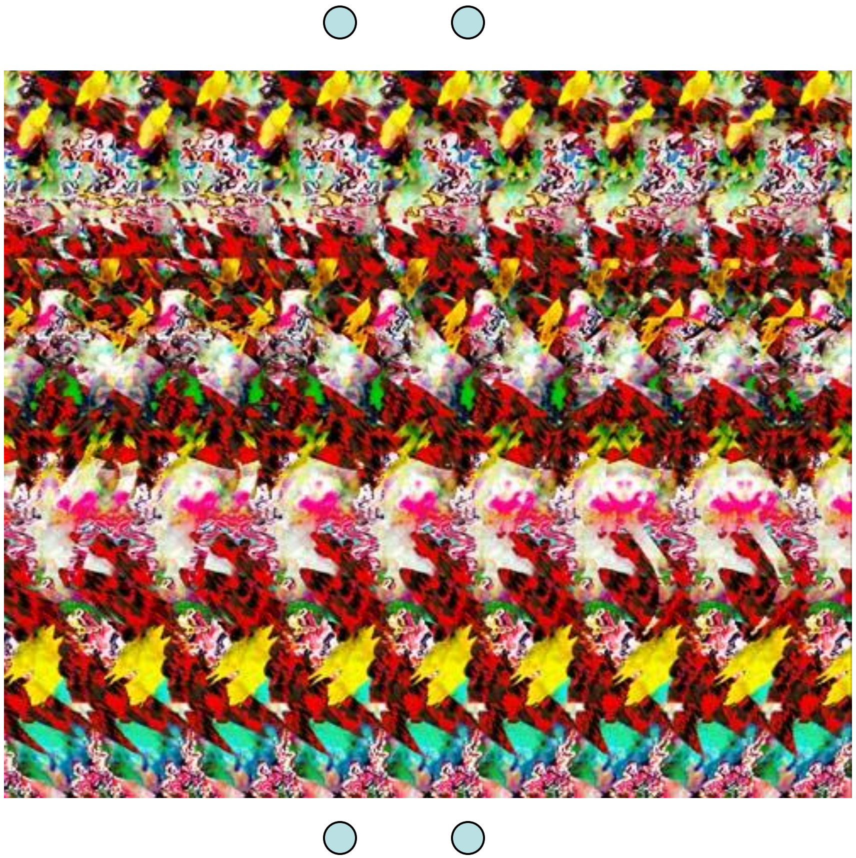



http://www.johnsonshawmuseum.org

**Public Library, Stereoscopic Looking Room, Chicago, by Phillips, 1923**

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

# Autostereograms



Exploit disparity as depth cue using single image.

(Single image random dot stereogram, Single image stereogram)

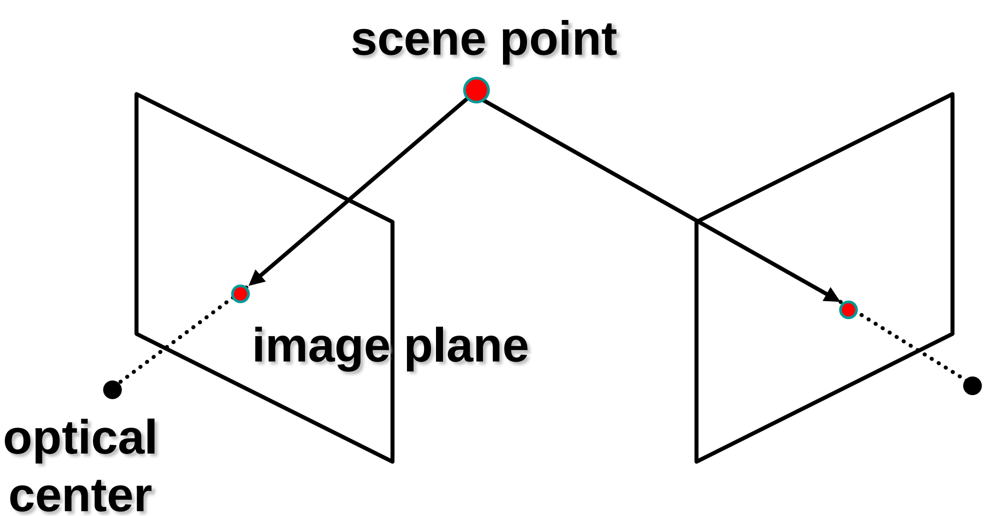
Images from magiceye.com

# Autostereograms

Images from magiceye.com

# Estimating depth with stereo

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

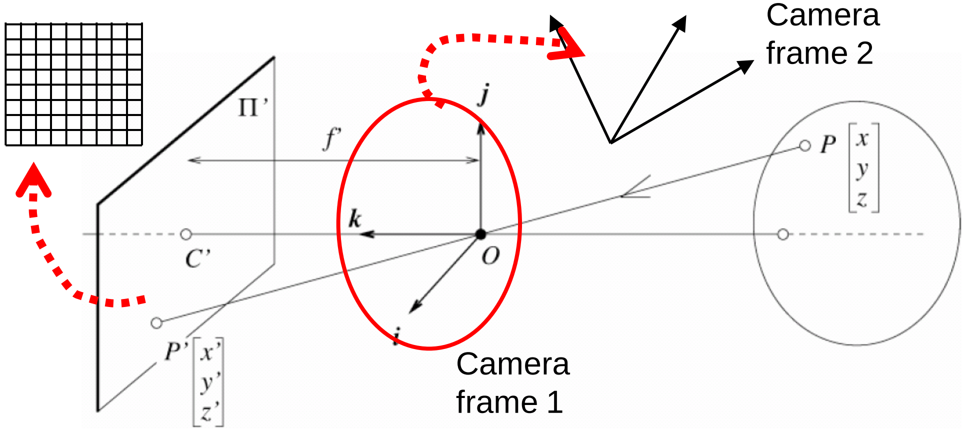



# Stereo vision


Two cameras, simultaneous views

Single moving camera and static scene

# 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

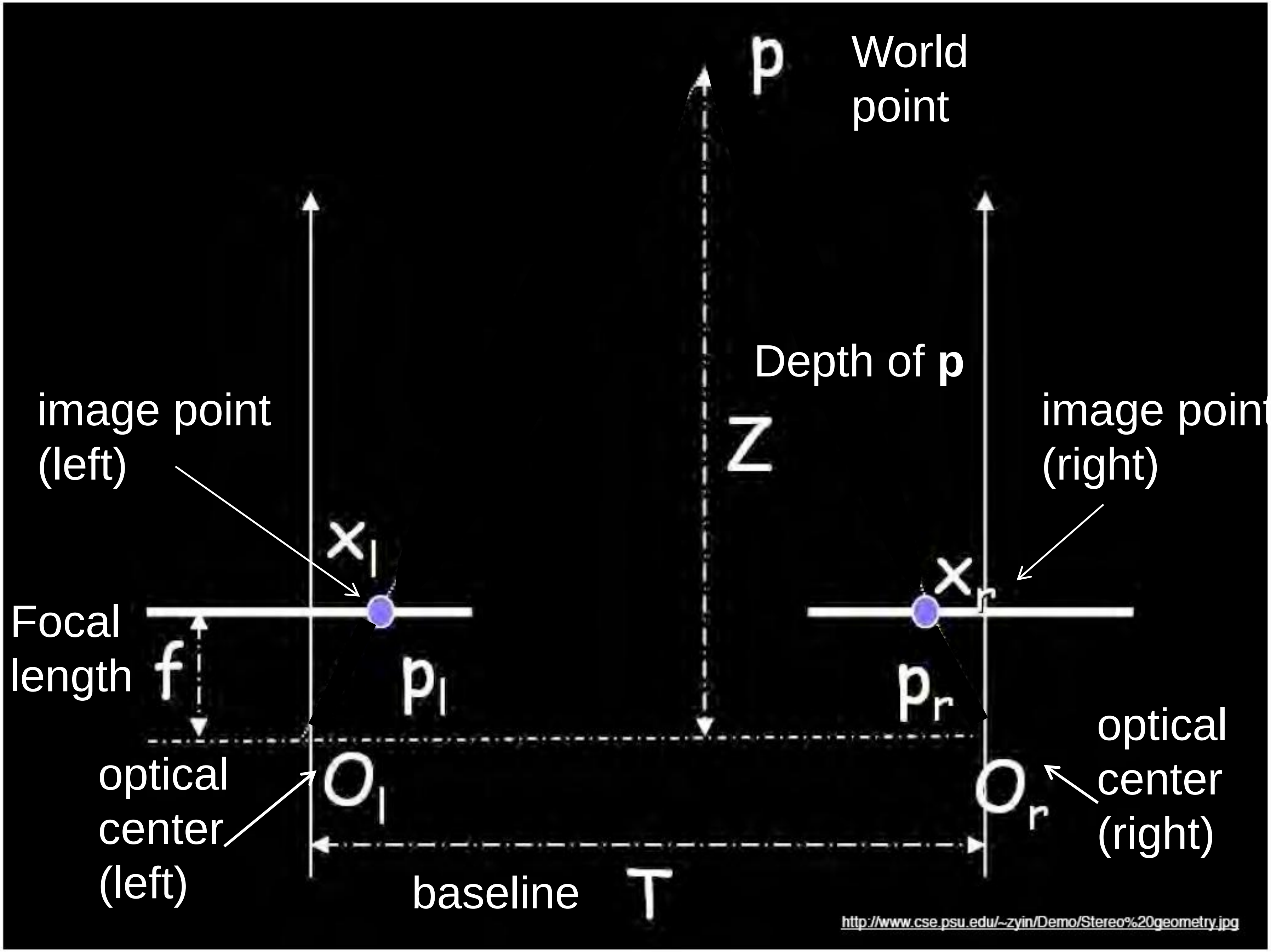
*We'll assume for now that these parameters are given and fixed.*

# Outline

- Human stereopsis
- Stereograms
- Epipolar geometry and the epipolar constraint
  - Case example with parallel optical axes
  - General case with calibrated cameras

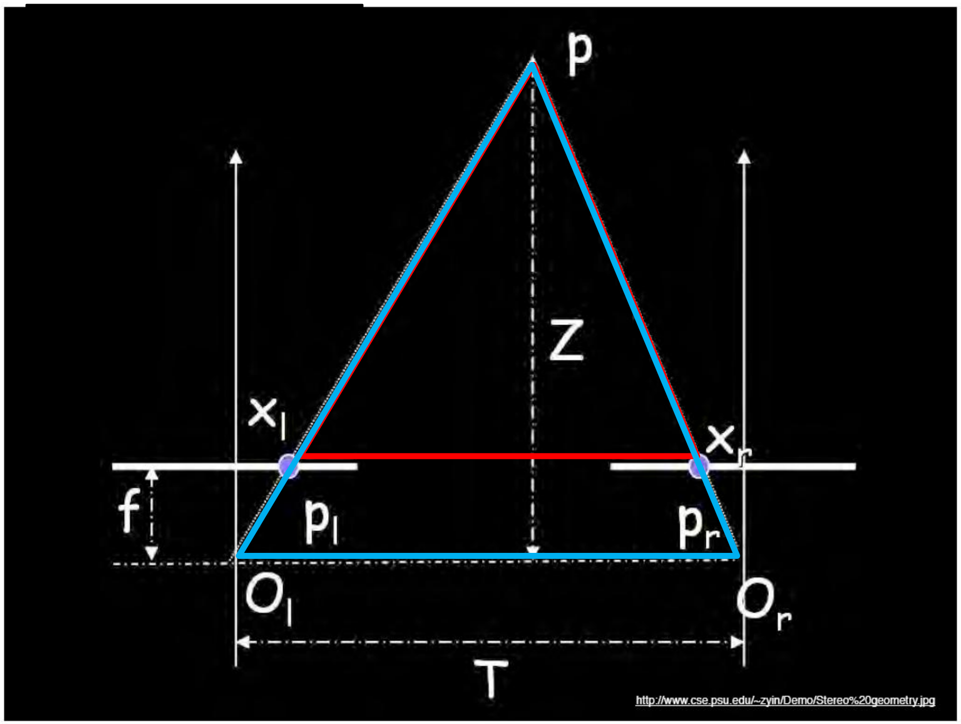
# Geometry for a simple stereo system

- 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, $O_r$):

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

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

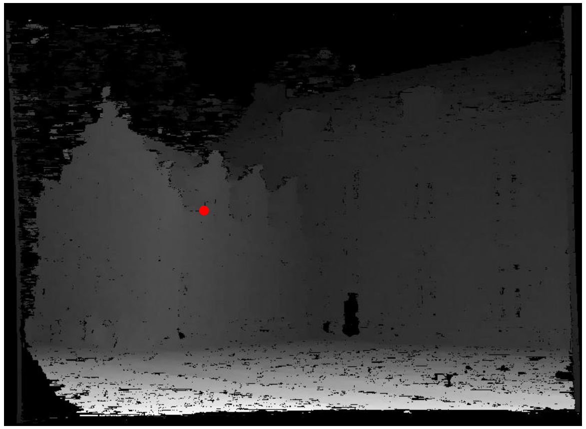
**disparity**

# 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**…

# Summary

- Depth from stereo: main idea is to triangulate from corresponding image points.
- Epipolar geometry defined by two cameras
  - We've assumed known extrinsic parameters relating their poses
- Epipolar constraint limits where points from one view will be imaged in the other
  - Makes search for correspondences quicker

- **Terms**: epipole, epipolar plane / lines, disparity, rectification, intrinsic/extrinsic parameters, essential matrix, baseline

# Coming up

– Computing correspondences
– Non-geometric stereo constraints