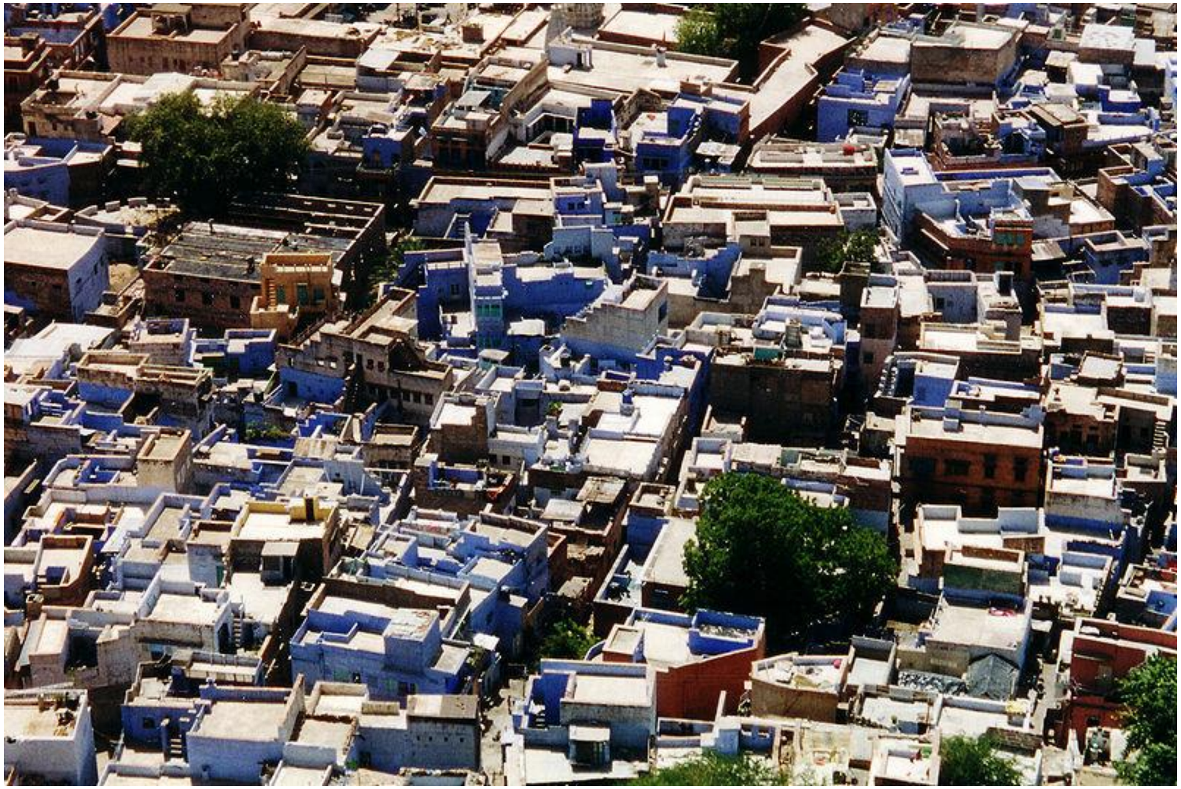
Miniature faking

In close-up photo, the depth of field is limited.

http://en.wikipedia.org/wiki/File:Jodhpur_tilt_shift.jpg

Miniature faking



Miniature faking

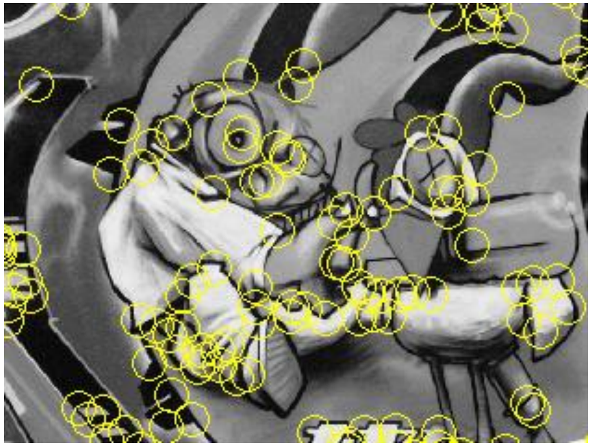
http://en.wikipedia.org/wiki/File:Oregon_State_Beavers_Tilt-Shift_Miniature_Greg_Keene.jpg

Review

- Previous section:
  - Feature detection and matching
  - Model fitting and outlier rejection

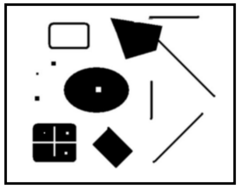
Review: Interest points

- Keypoint detection: repeatable and distinctive
  - Corners, blobs, stable regions
  - Harris, DoG, MSER



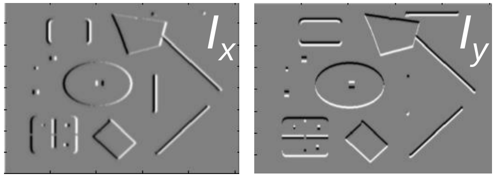
Harris Detector [Harris88]

- Second moment matrix



$$\mu(\sigma_I,\sigma_D) = g(\sigma_I) * \begin{bmatrix} I_x^2(\sigma_D) & I_xI_y(\sigma_D) \\ I_xI_y(\sigma_D) & I_y^2(\sigma_D) \end{bmatrix}$$

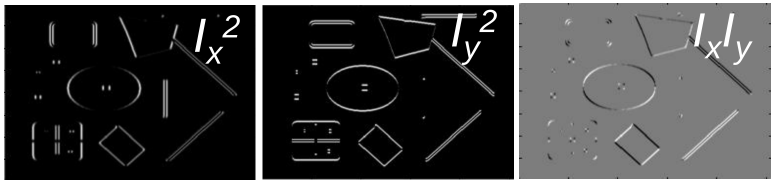
1. Image derivatives (optionally, blur first)



$$\det M = \lambda_1 \lambda_2$$

$$\operatorname{trace} M = \lambda_1 + \lambda_2$$

2. Square of derivatives



3. Gaussian filter $g(\sigma_I)$

4. Cornerness function – both eigenvalues are strong

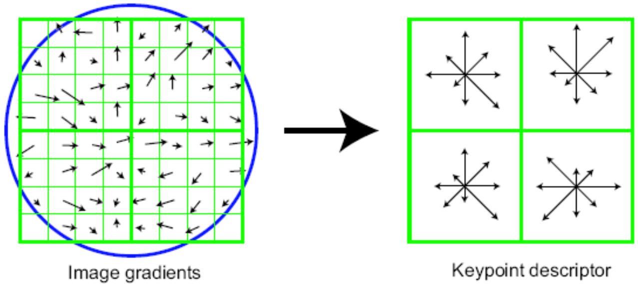
$$har = \det[\mu(\sigma_I,\sigma_D)] - \alpha[\operatorname{trace}(\mu(\sigma_I,\sigma_D))^2] =$$

$$g(I_x^2)g(I_y^2) - [g(I_xI_y)]^2 - \alpha[g(I_x^2)+g(I_y^2)]^2$$

5. Non-maxima suppression

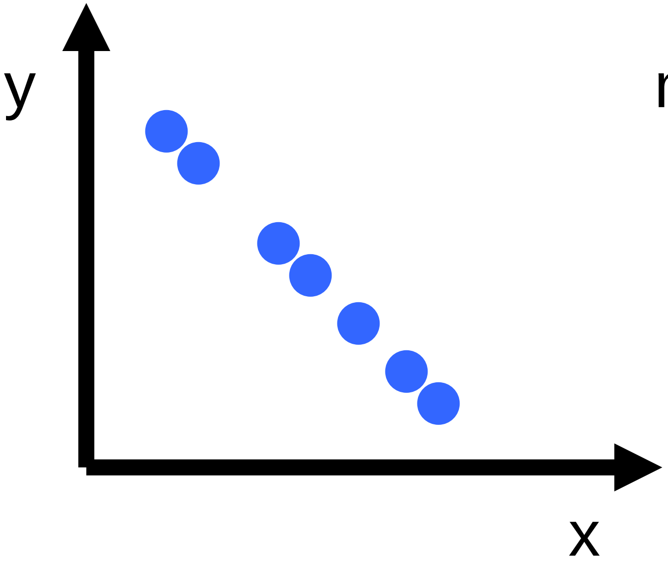
Review: Local Descriptors

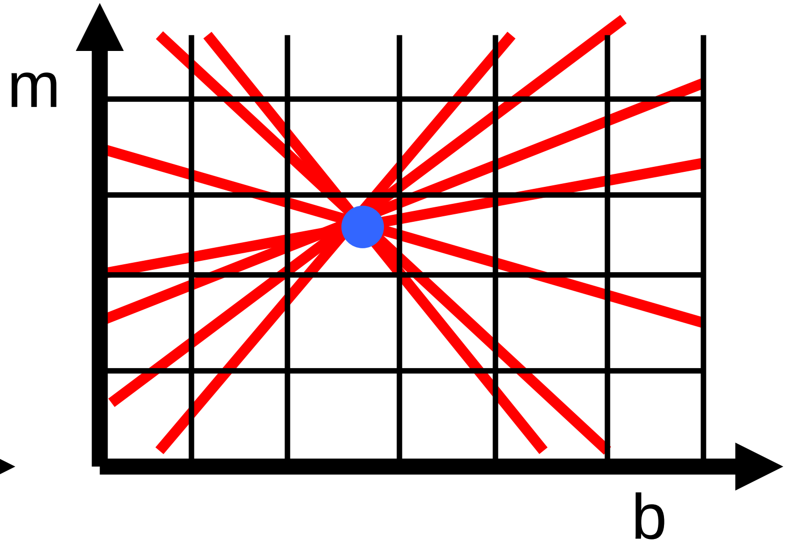
- Most features can be thought of as templates, histograms (counts), or combinations
- Most available descriptors focus on edge/gradient information
    - Capture texture information
    - Color rarely used

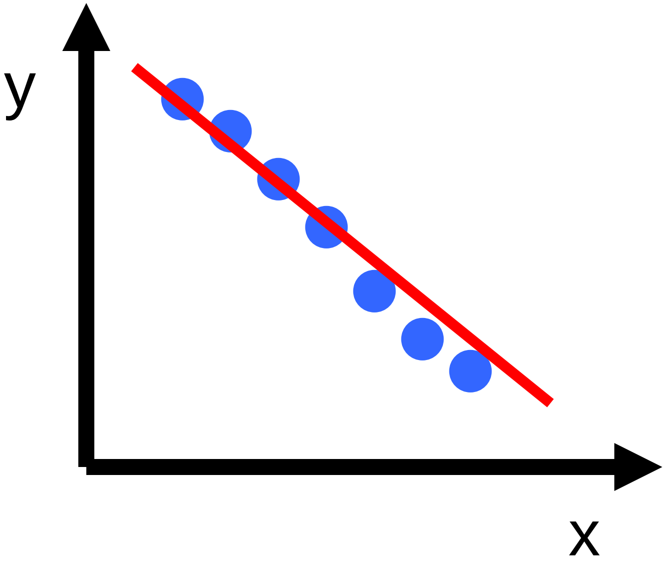


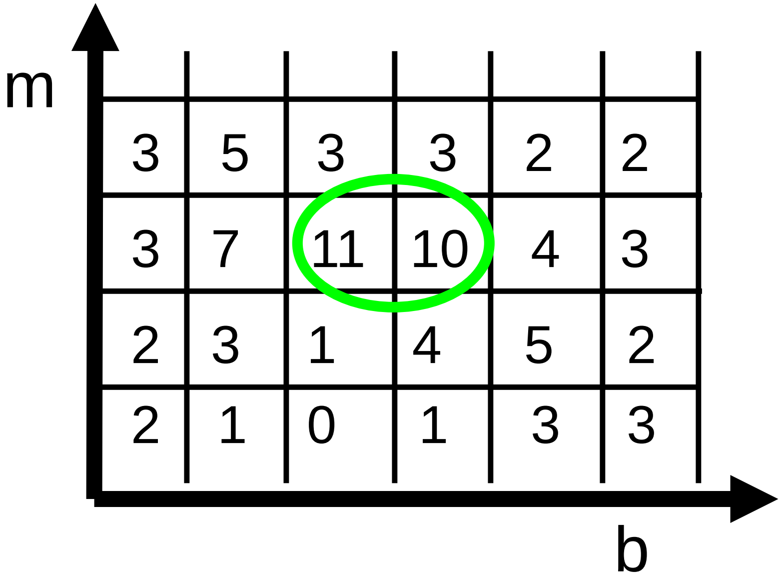
K. Grauman, B. Leibe

Review: Hough transform





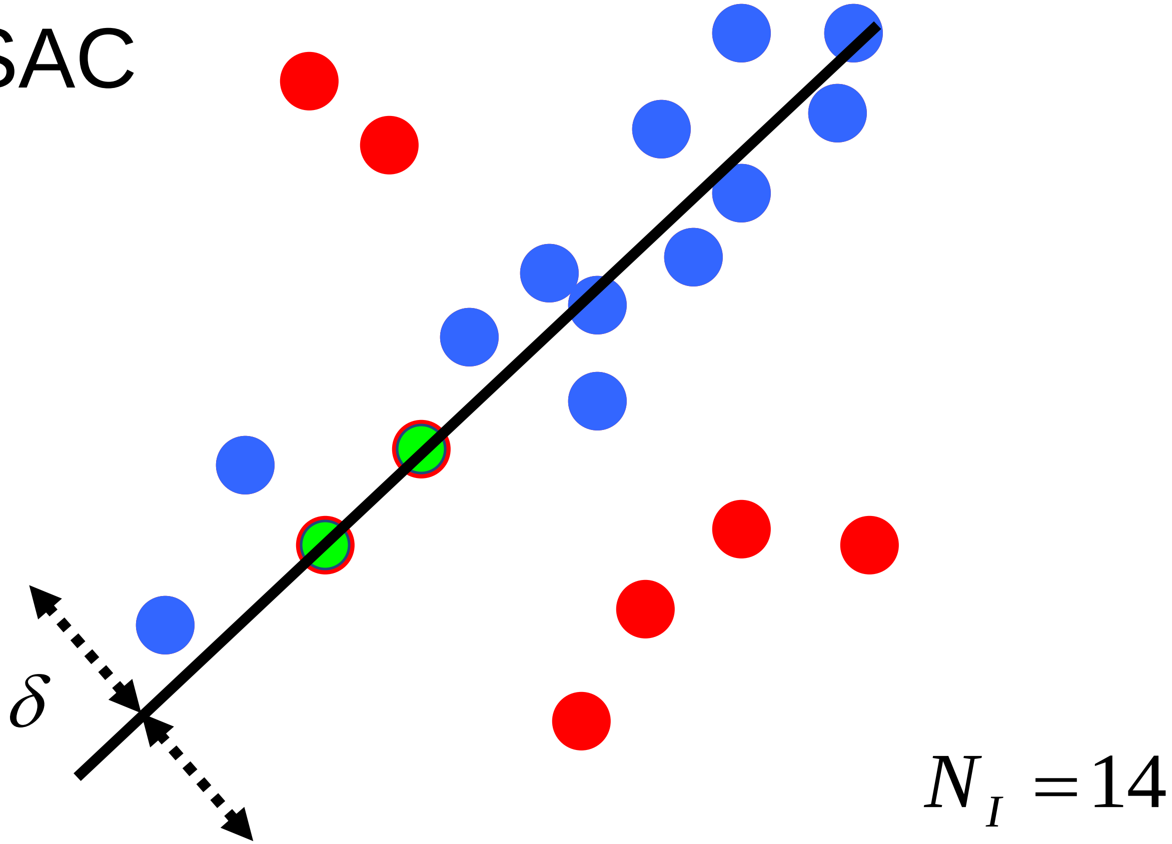




3	5	3	3	2	2
3	7	11	10	4	3
2	3	1	4	5	2
2	1	0	1	3	3

Slide from S. Savarese

Review: RANSAC

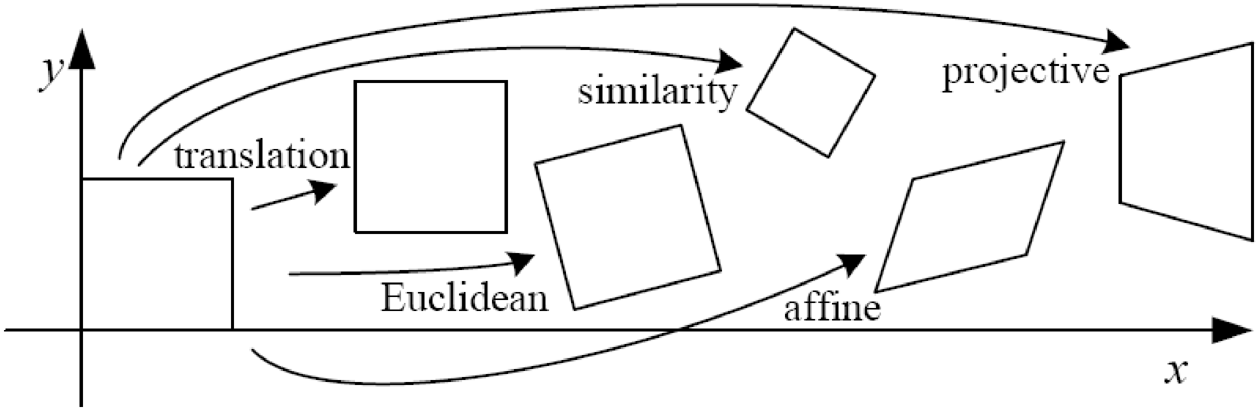


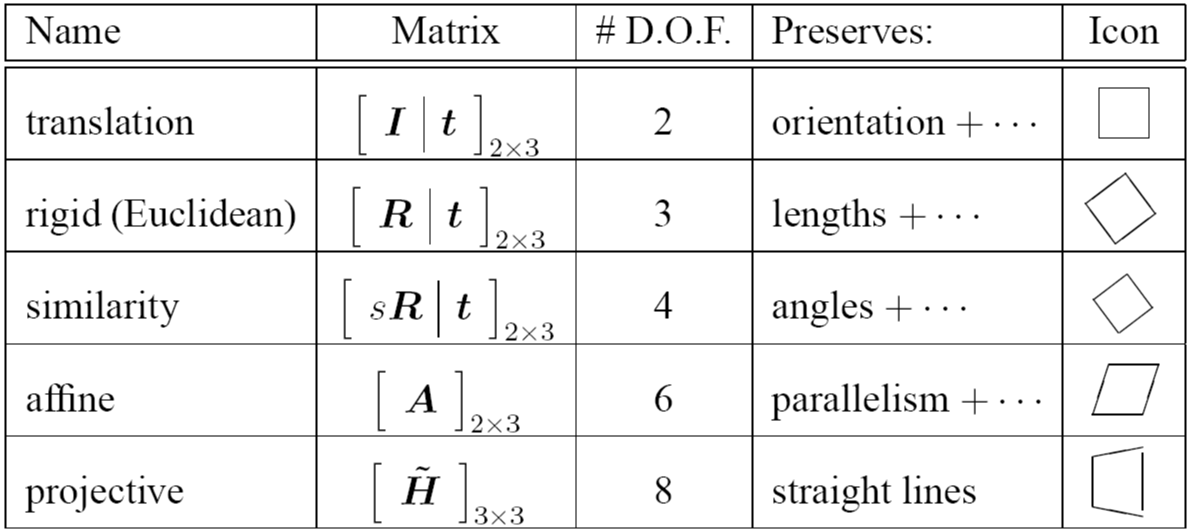
Algorithm:

1. **Sample** (randomly) the number of points required to fit the model (#=2)
2. **Solve** for model parameters using samples
3. **Score** by the fraction of inliers within a preset threshold of the model

Repeat 1-3 until the best model is found with high confidence

Review: 2D image transformations





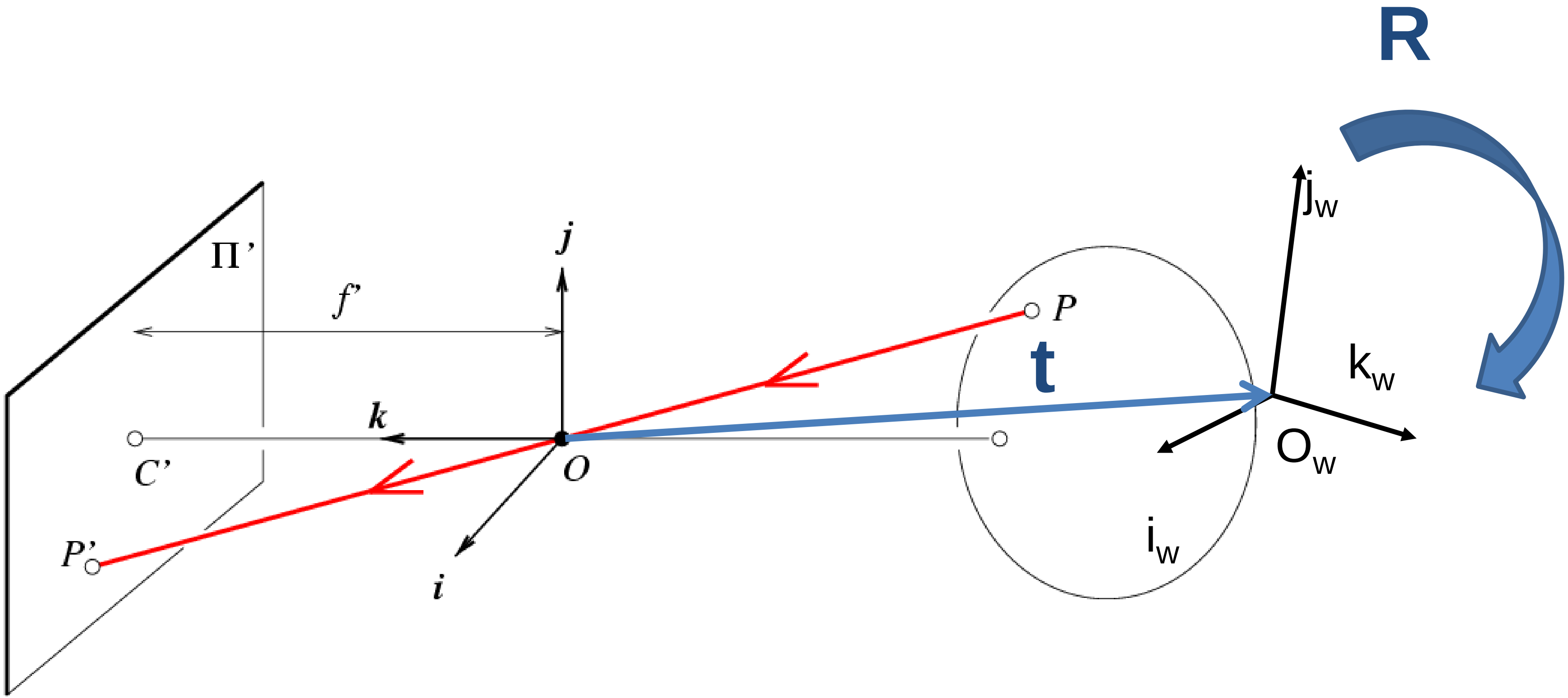
Name	Matrix	# D.O.F.	Preserves:	Icon
translation	$\left[\, \boldsymbol{I} \,\middle|\, \boldsymbol{t} \,\right]_{2\times 3}$	2	orientation $+ \cdots$	□
rigid (Euclidean)	$\left[\, \boldsymbol{R} \,\middle|\, \boldsymbol{t} \,\right]_{2\times 3}$	3	lengths $+ \cdots$	◇
similarity	$\left[\, s\boldsymbol{R} \,\middle|\, \boldsymbol{t} \,\right]_{2\times 3}$	4	angles $+ \cdots$	◇
affine	$\left[\, \boldsymbol{A} \,\right]_{2\times 3}$	6	parallelism $+ \cdots$	▱
projective	$\left[\, \tilde{\boldsymbol{H}} \,\right]_{3\times 3}$	8	straight lines	⏢

Szeliski 2.1

This section – multiple views

- Today – Intro to multiple views and Stereo
- Wednesday – Camera calibration
- Friday – Fundamental Matrix
- Monday – Optical Flow
- Wednesday – Multiview wrapup

Oriented and Translated Camera



Degrees of freedom

$$\mathbf{x}=\mathbf{K}\begin{bmatrix}\mathbf{R} & \mathbf{t}\end{bmatrix}\mathbf{X}$$

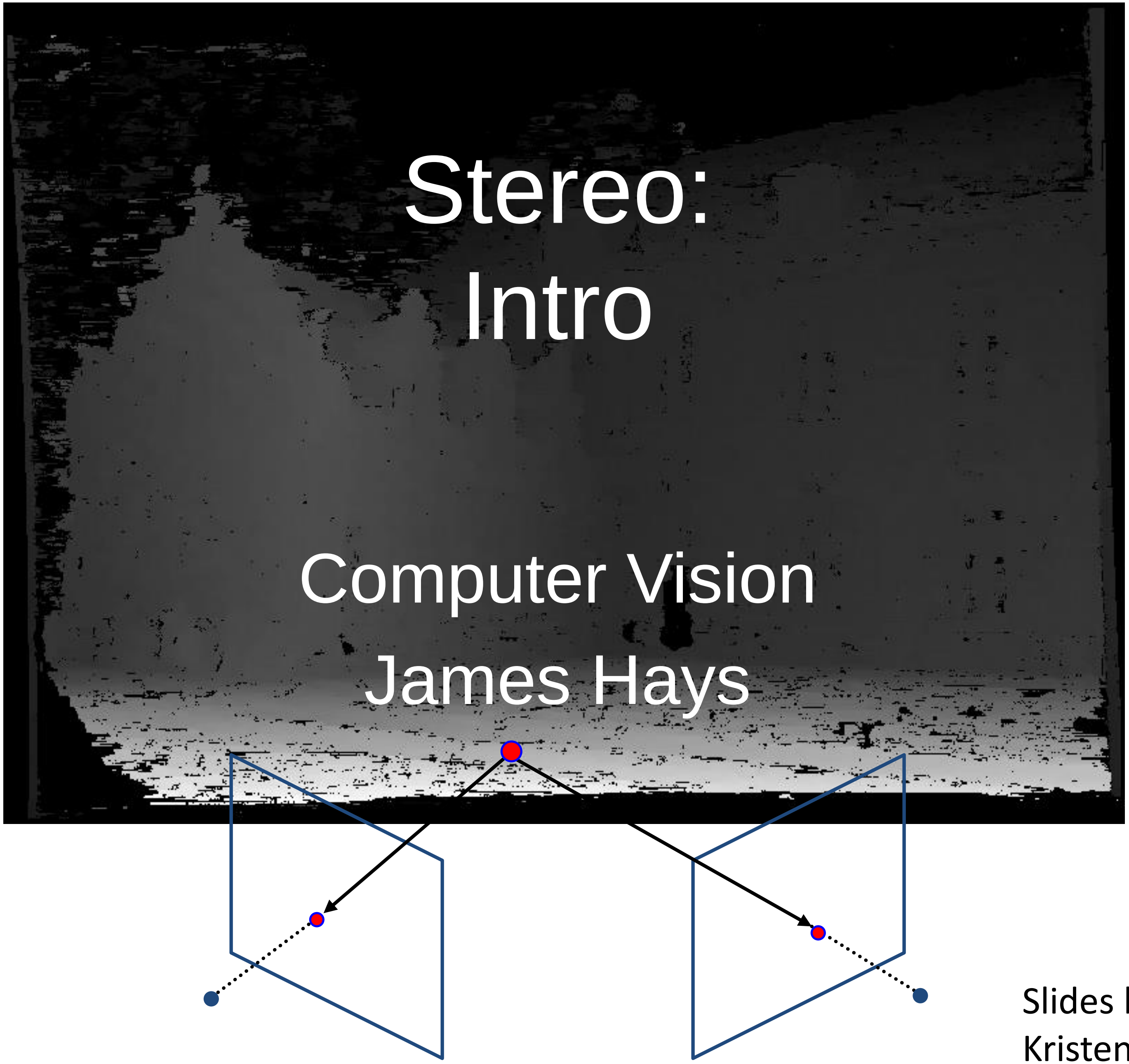
⇩

$$w\begin{bmatrix}u\\ v\\ 1\end{bmatrix}=\overset{5}{\begin{bmatrix}\alpha & s & u_0\\ 0 & \beta & v_0\\ 0 & 0 & 1\end{bmatrix}}\overset{6}{\begin{bmatrix}r_{11} & r_{12} & r_{13} & t_x\\ r_{21} & r_{22} & r_{23} & t_y\\ r_{31} & r_{32} & r_{33} & t_z\end{bmatrix}}\begin{bmatrix}x\\ y\\ z\\ 1\end{bmatrix}$$

How to calibrate the camera?

$$\mathbf{x} = \mathbf{K}\begin{bmatrix}\mathbf{R} & \mathbf{t}\end{bmatrix}\mathbf{X}$$

$$\begin{bmatrix} su \\ sv \\ s \end{bmatrix} = \begin{bmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$



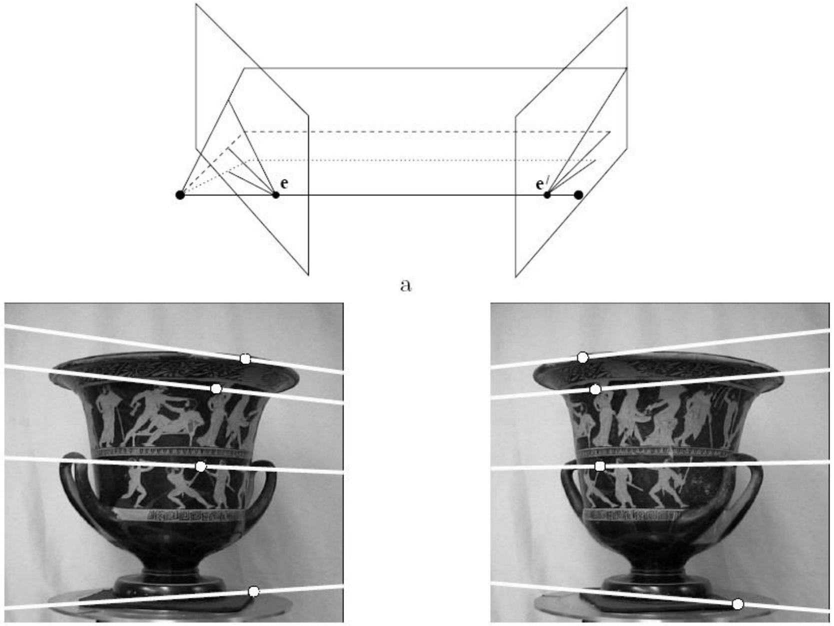
Stereo:
Intro

Computer Vision

James Hays

Slides by
Kristen Grauman

Multiple views



Hartley and Zisserman

stereo vision
structure from motion
optical flow

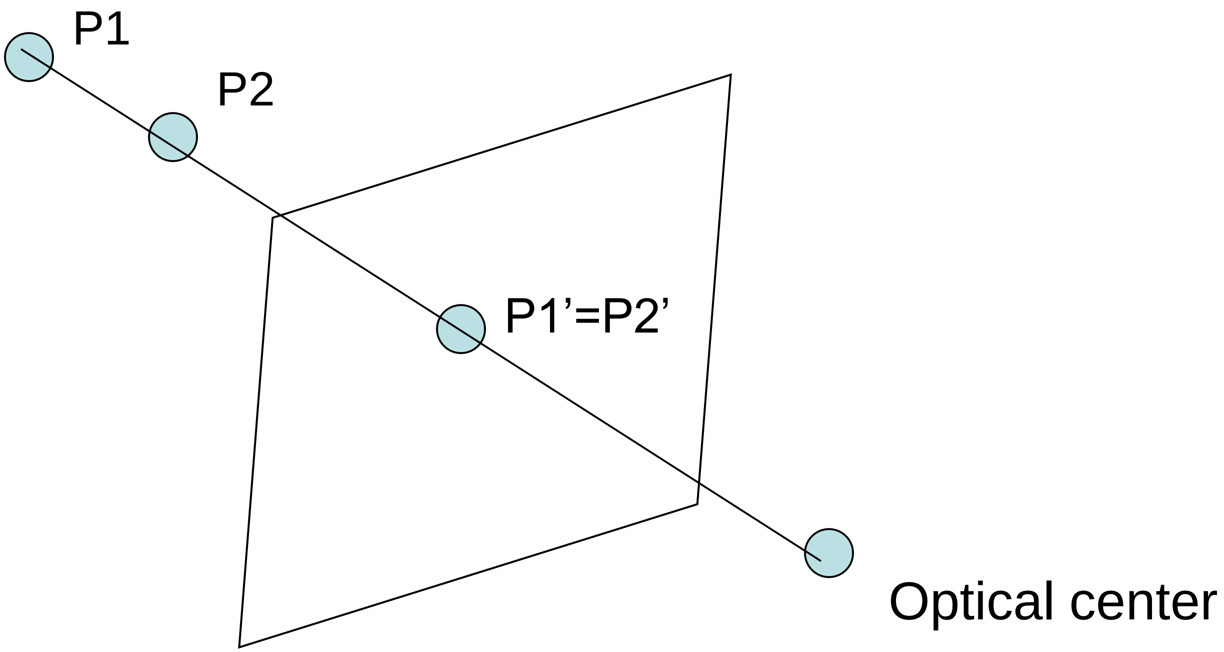
Why multiple views?

- Structure and depth are inherently ambiguous from single views.

Images from Lana Lazebnik

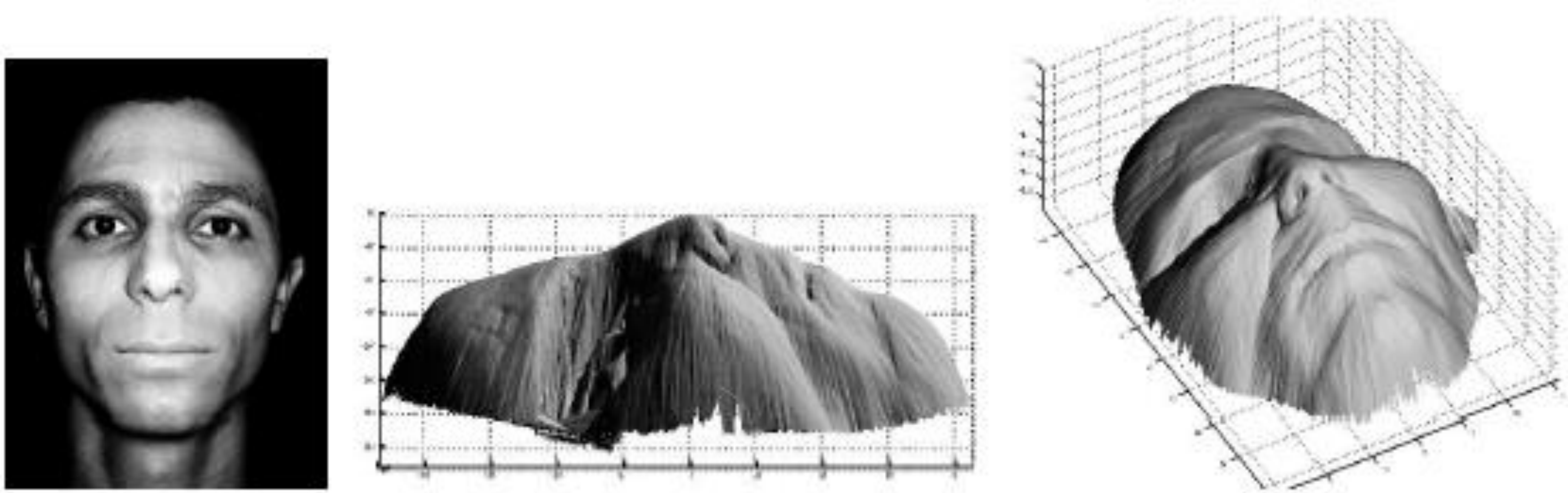
Why multiple views?

- Structure and depth are inherently ambiguous from single views.



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

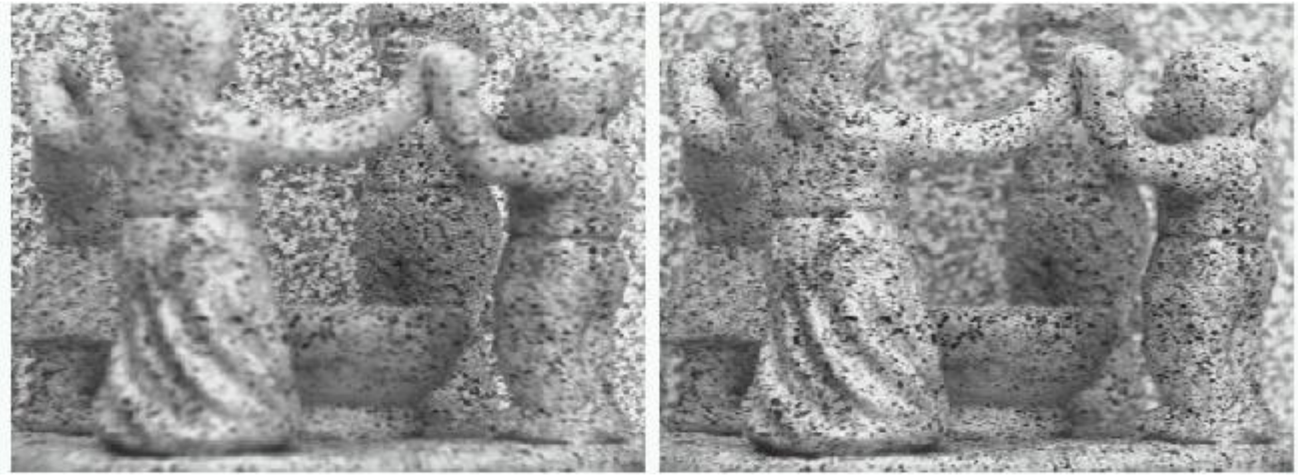
Shading



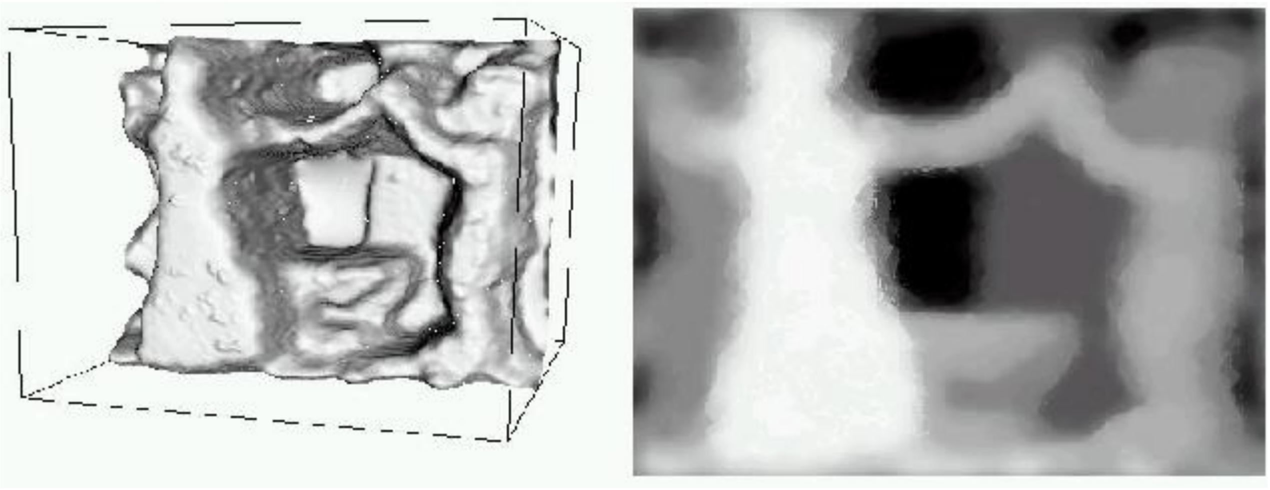
a) b) c)

[Figure from Prados & Faugeras 2006]

Focus/defocus



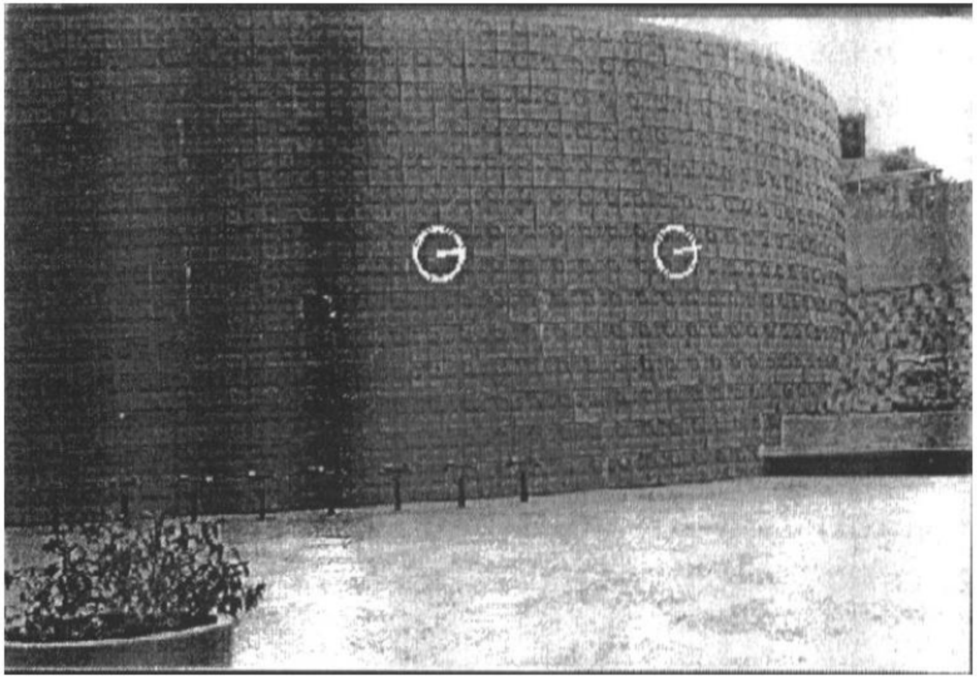
Images from same point of view, different camera parameters

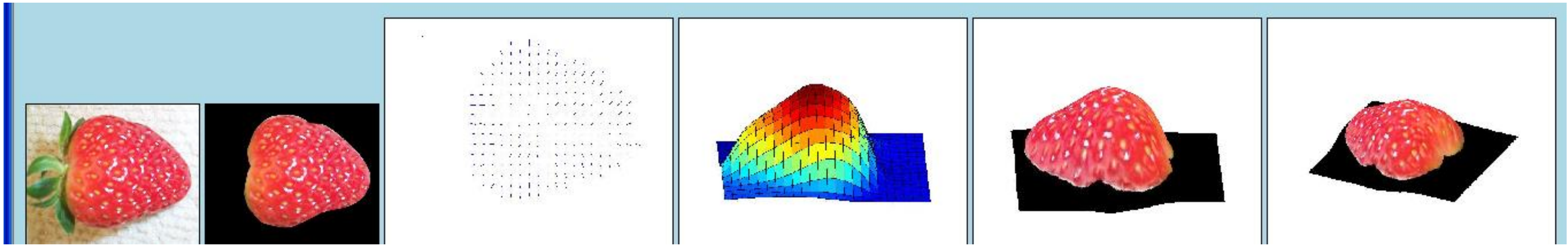


3d shape / depth estimates

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

Texture



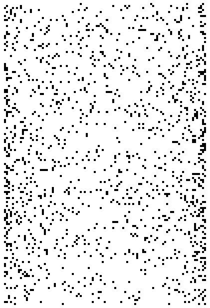


[From A.M. Loh. The recovery of 3-D structure using visual texture patterns. PhD thesis]

Perspective effects

Image credit: S. Seitz

Motion



Figures from L. Zhang

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

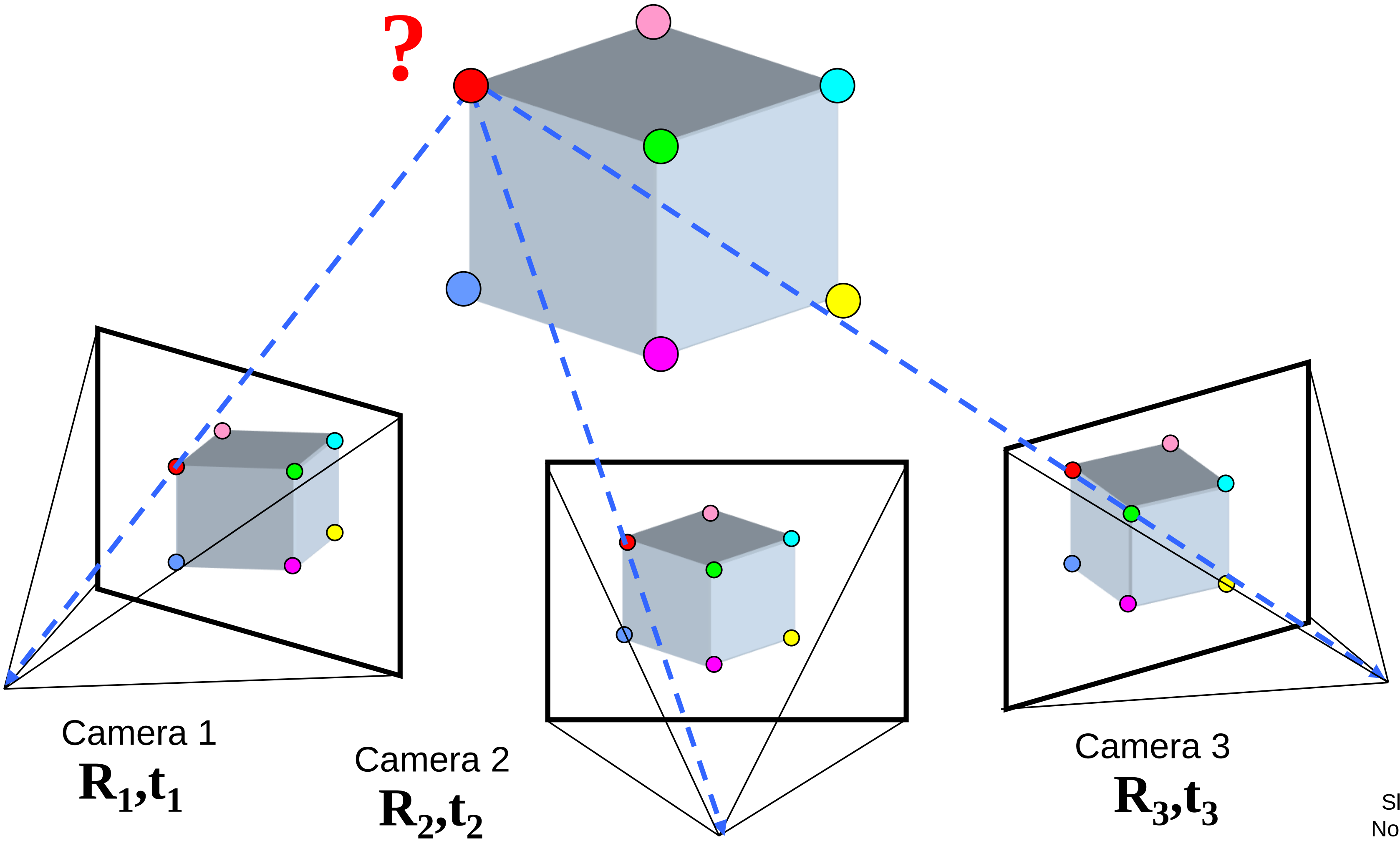
Occlusion

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

If stereo were critical for depth perception, navigation, recognition, etc., then this would be a problem

Multi-view geometry problems

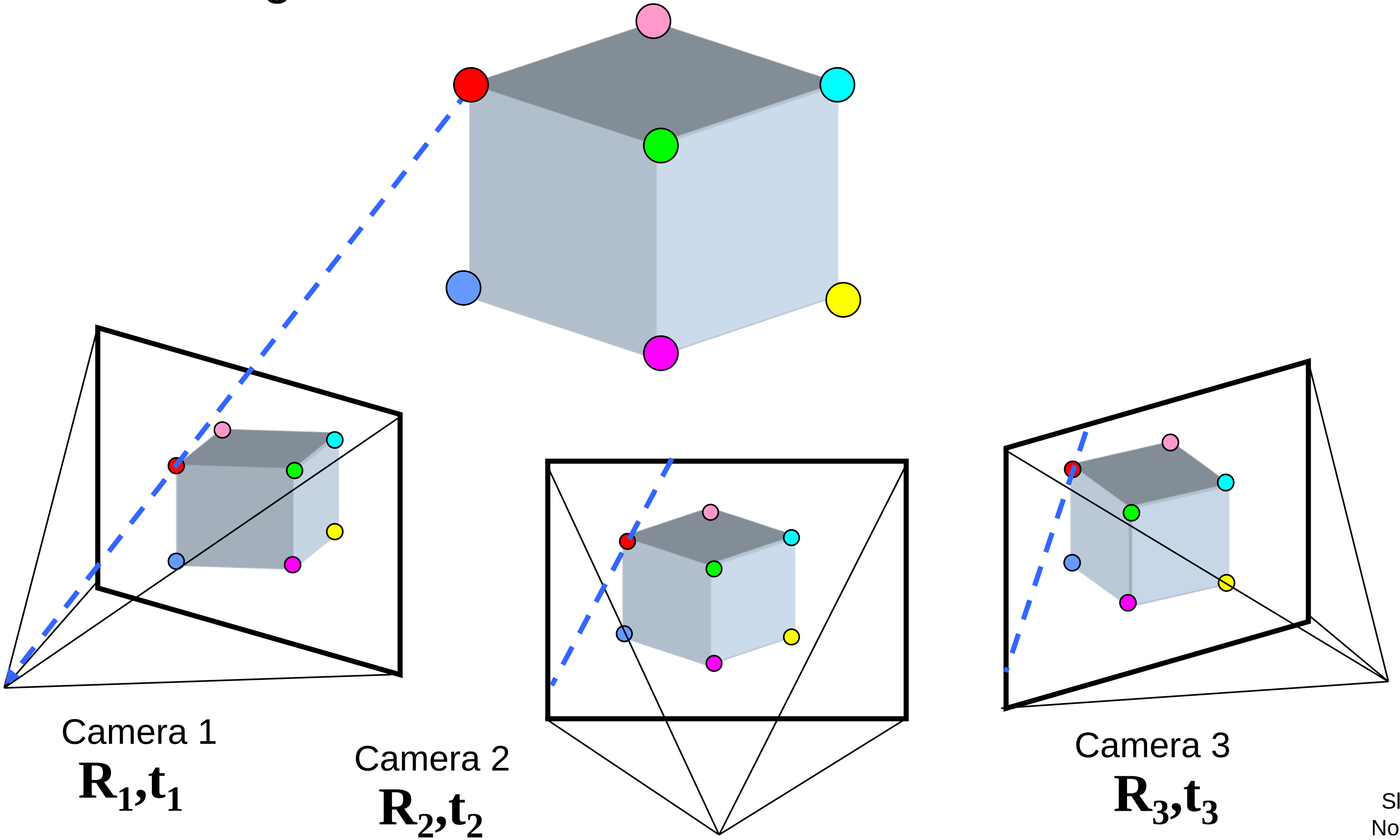
- **Structure:** Given projections of the same 3D point in two or more images, compute the 3D coordinates of that point



Slide credit: Noah Snavely

Multi-view geometry problems

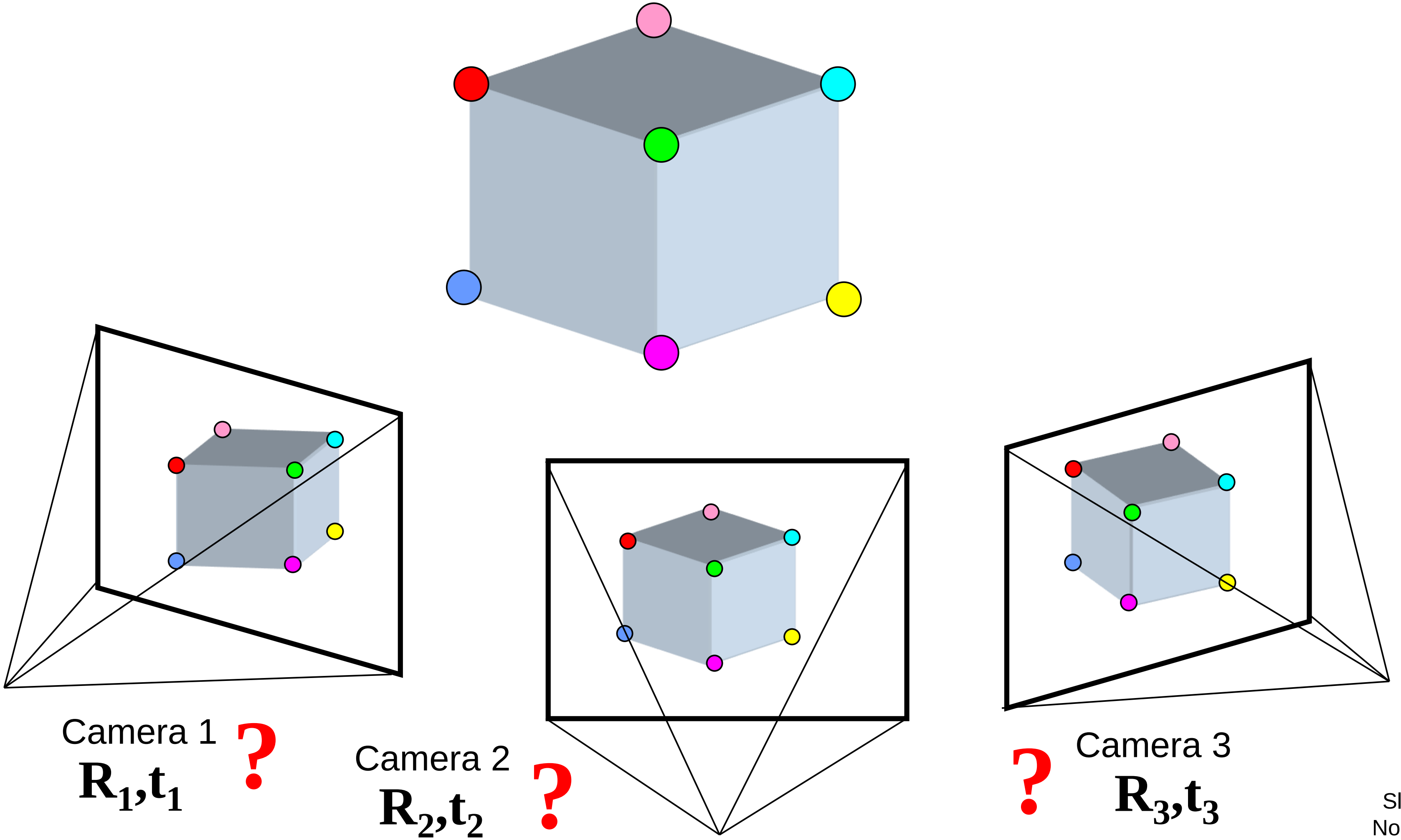
- **Stereo correspondence:** Given a point in one of the images, where could its corresponding points be in the other images?



Slide credit: Noah Snavely

Multi-view geometry problems

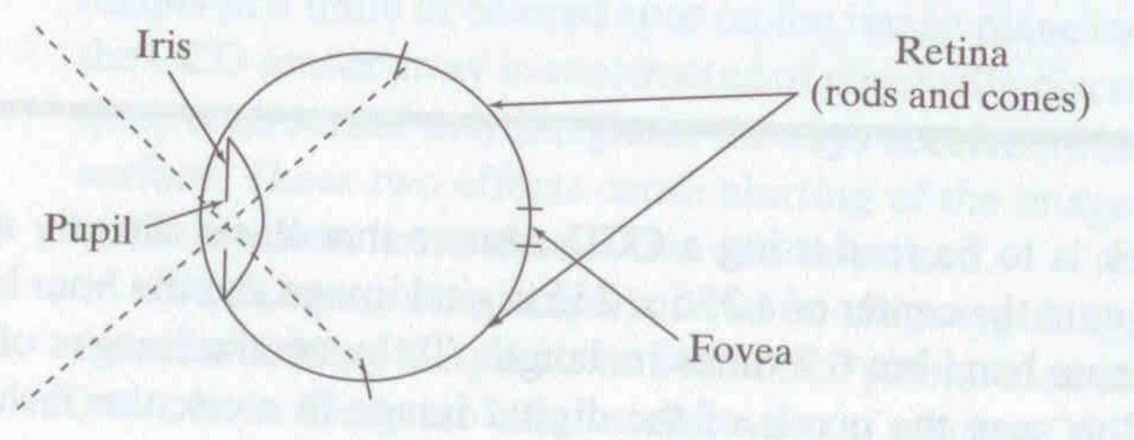
- **Motion:** Given a set of corresponding points in two or more images, compute the camera parameters



Slide credit: Noah Snavely

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

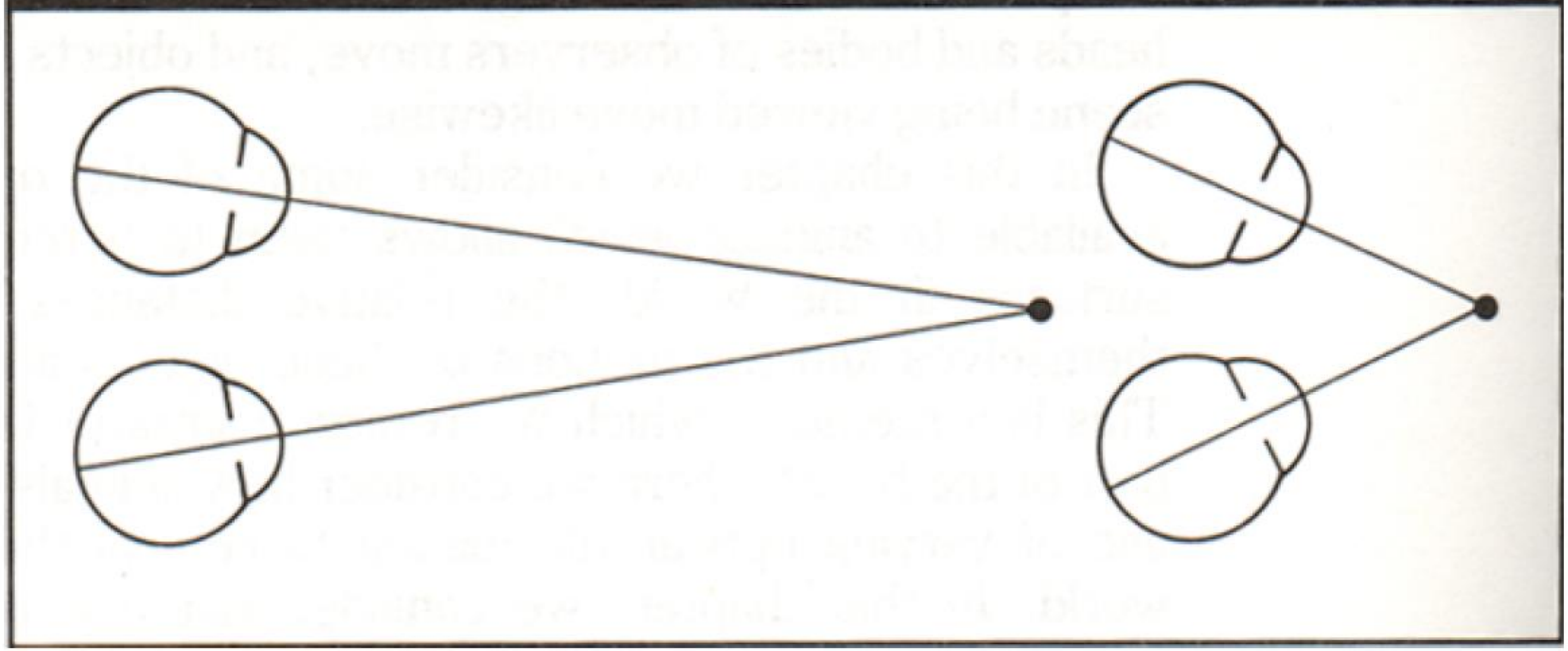


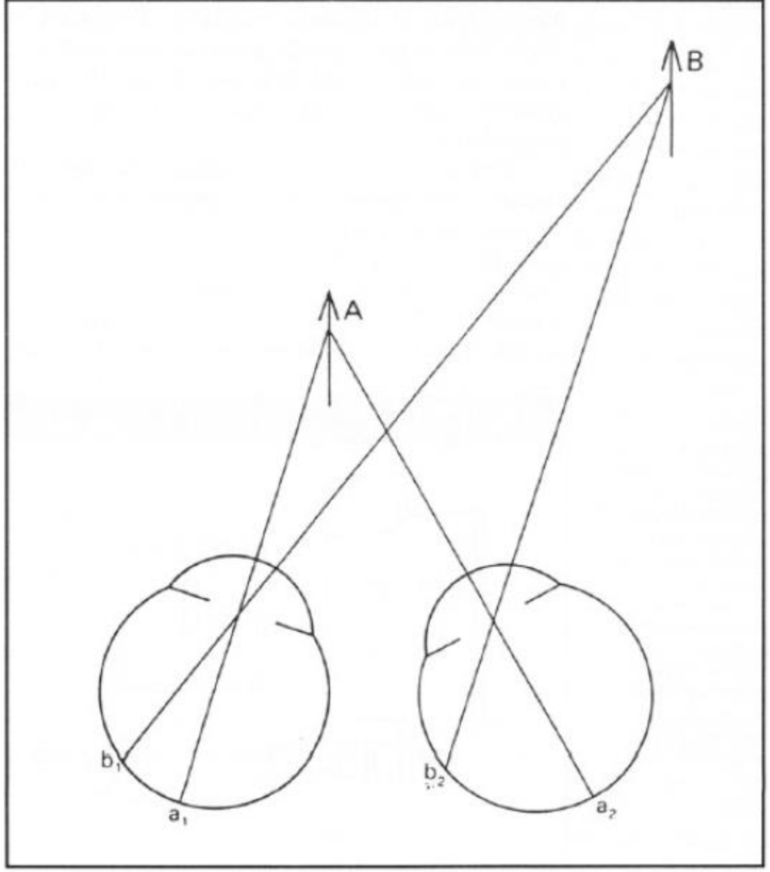
FIGURE 7.1

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



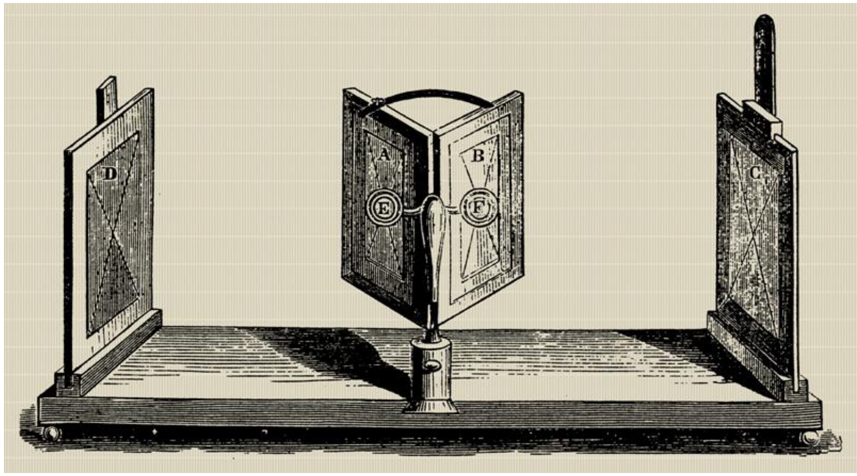
From Bruce and Green, Visual Perception, Physiology, Psychology and Ecology

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

Adapted from David Forsyth, UC Berkeley

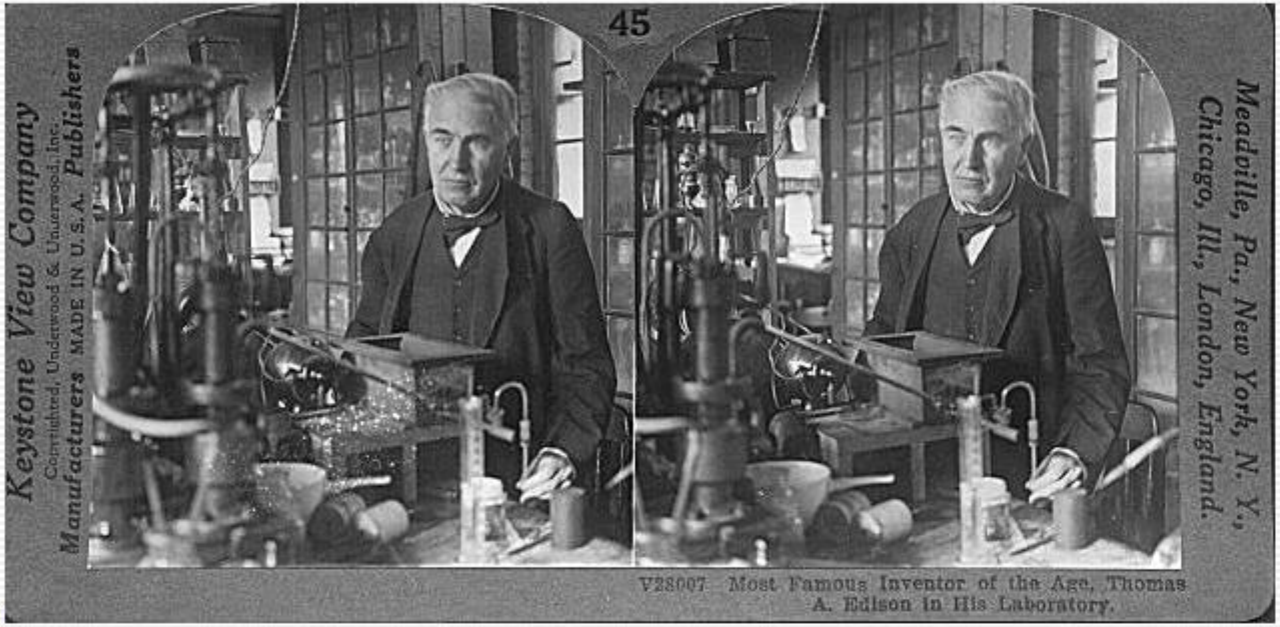
Stereo photography and stereo viewers

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



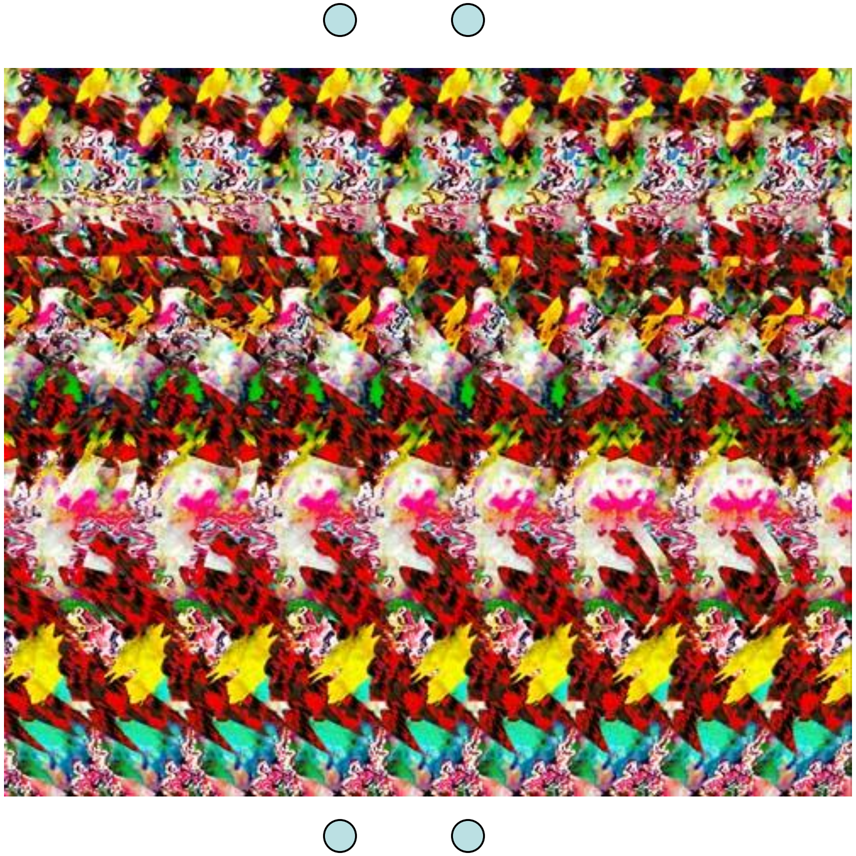
http://www.johnsonshawmuseum.org

http://www.johnsonshawmuseum.org

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

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

Autostereograms

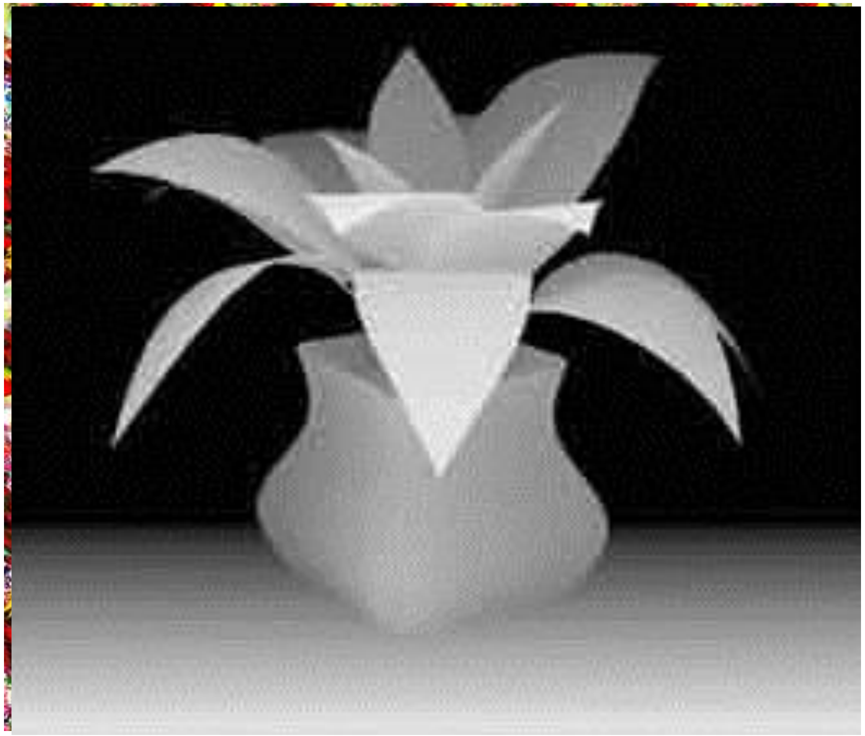


Exploit disparity as depth cue using single image.

(Single image random dot stereogram, Single image stereogram)

Images from magiceye.com

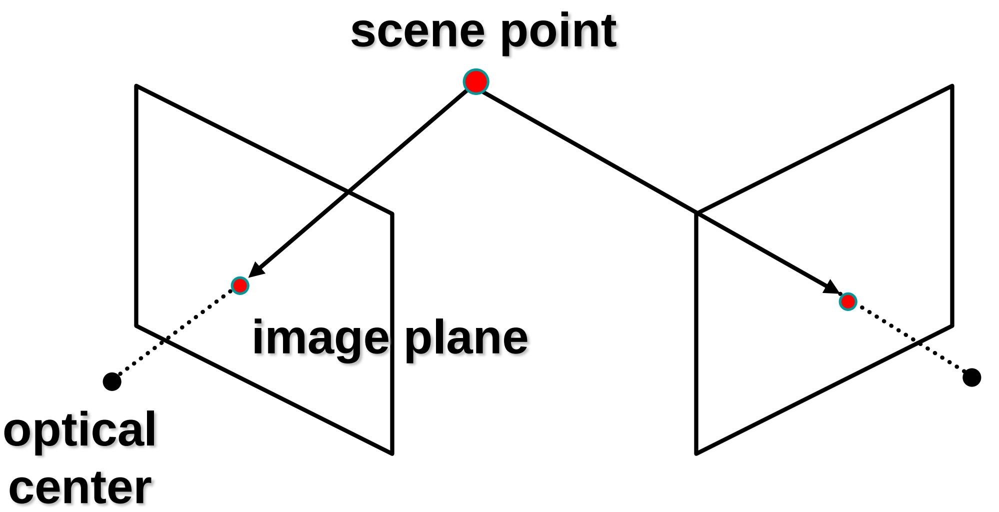
Autostereograms

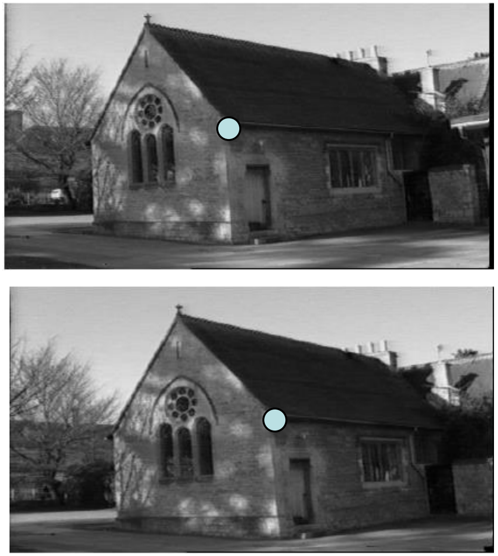


Images from magiceye.com

Estimating depth with stereo

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



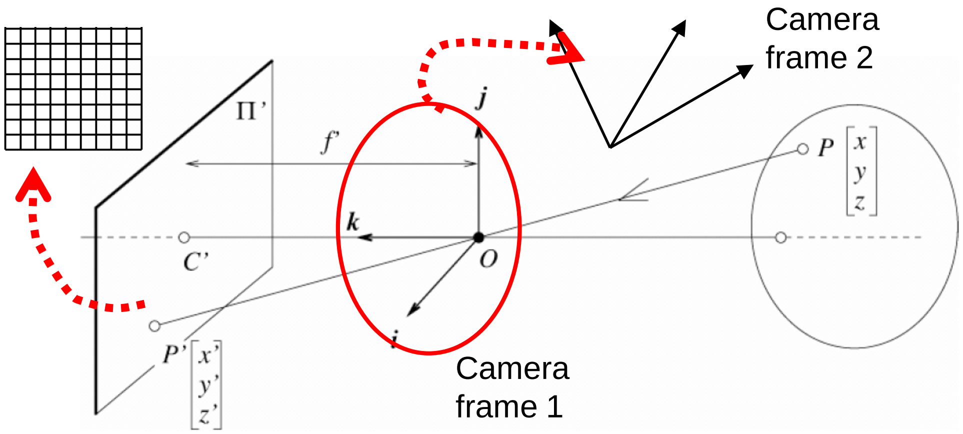


Stereo vision

Two cameras, simultaneous views

Single moving camera and static scene

Camera parameters



Extrinsic parameters:
Camera frame 1 ←→ Camera frame 2

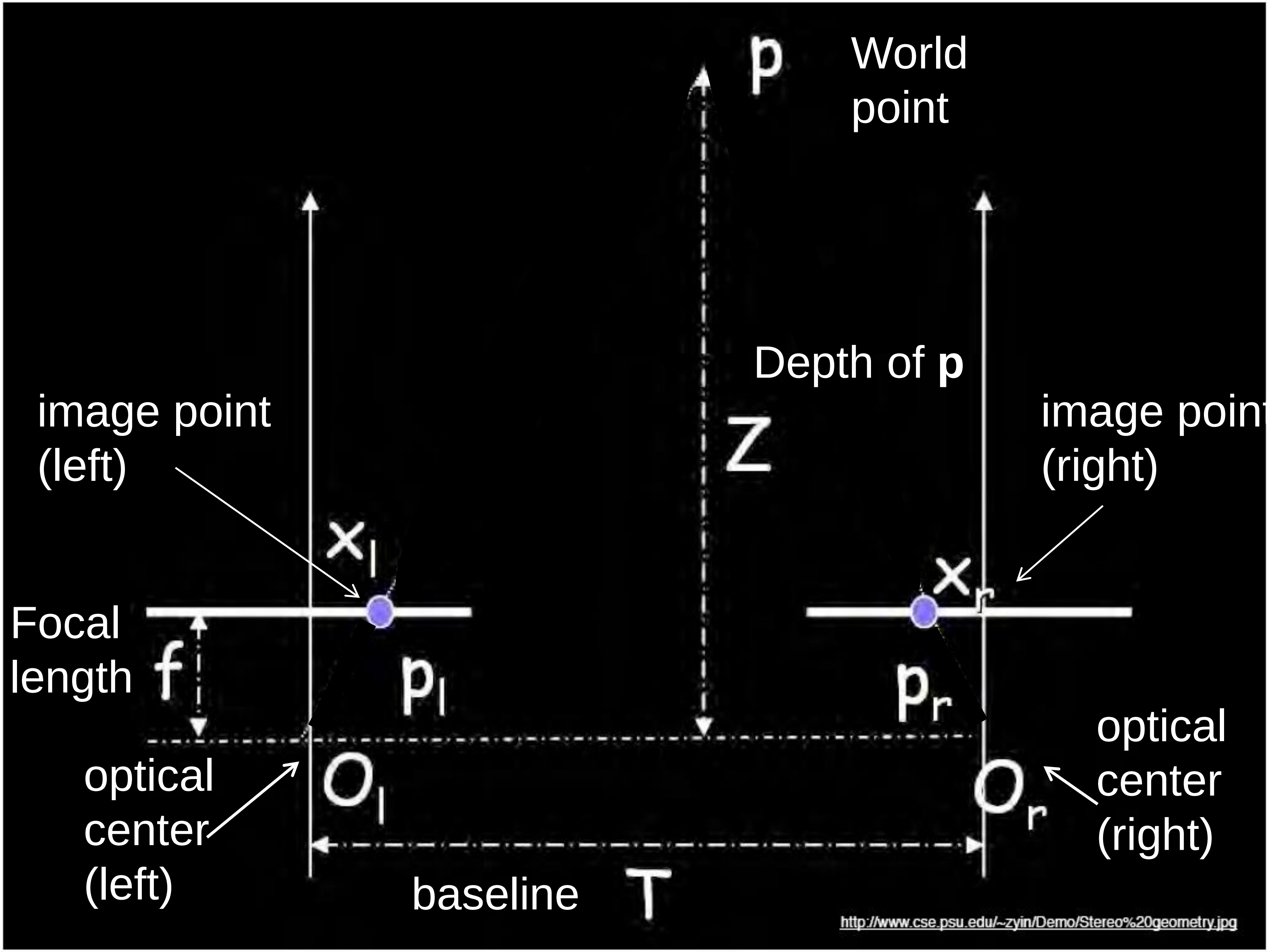
Intrinsic parameters:
Image coordinates relative to camera ←→ Pixel coordinates

- *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.

Geometry for a simple stereo system

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



World point

Depth of **p**

image point (left)

image point (right)

Focal length

optical center (left)

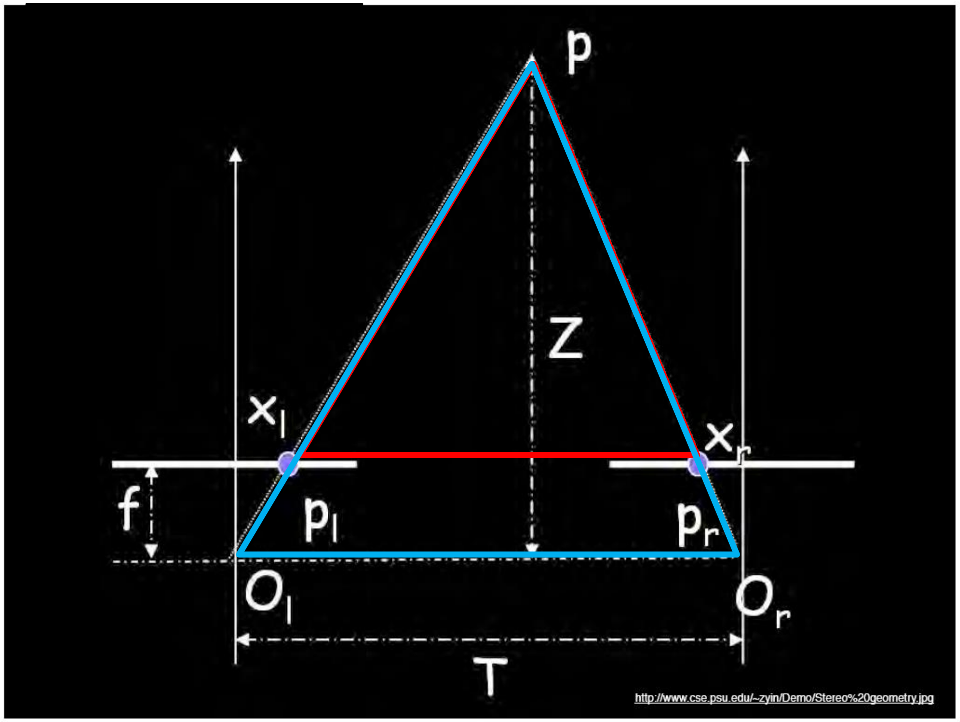
optical center (right)

baseline

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):

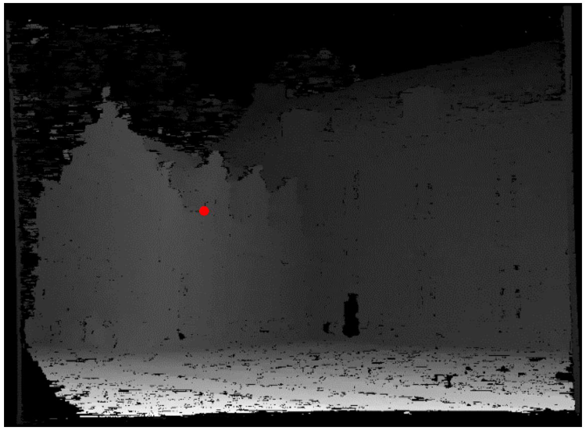
$$\frac{T + x_l - x_r}{Z - f} = \frac{T}{Z}$$

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

disparity

Depth from disparity

image $I(x,y)$



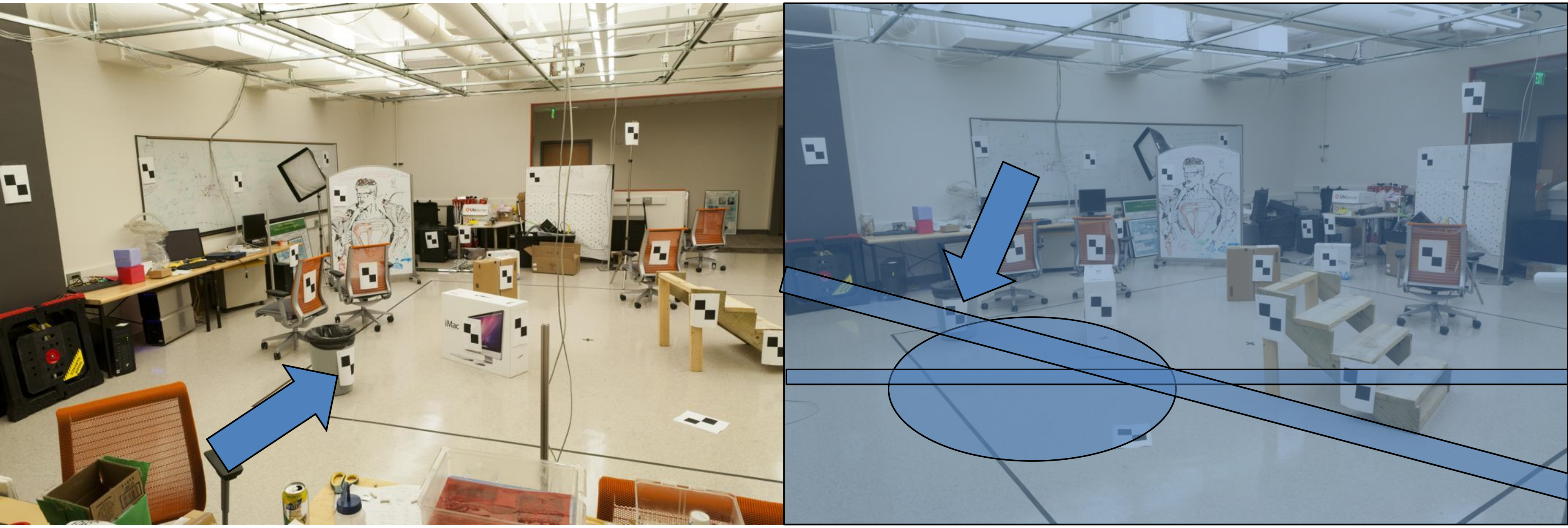
Disparity map $D(x,y)$

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**…

Where do we need to search?



To be continued…