SAIR
Spatial AI & Robotics Lab

CSE 473/573-A
L15: EPIPOLAR GEOMETRY

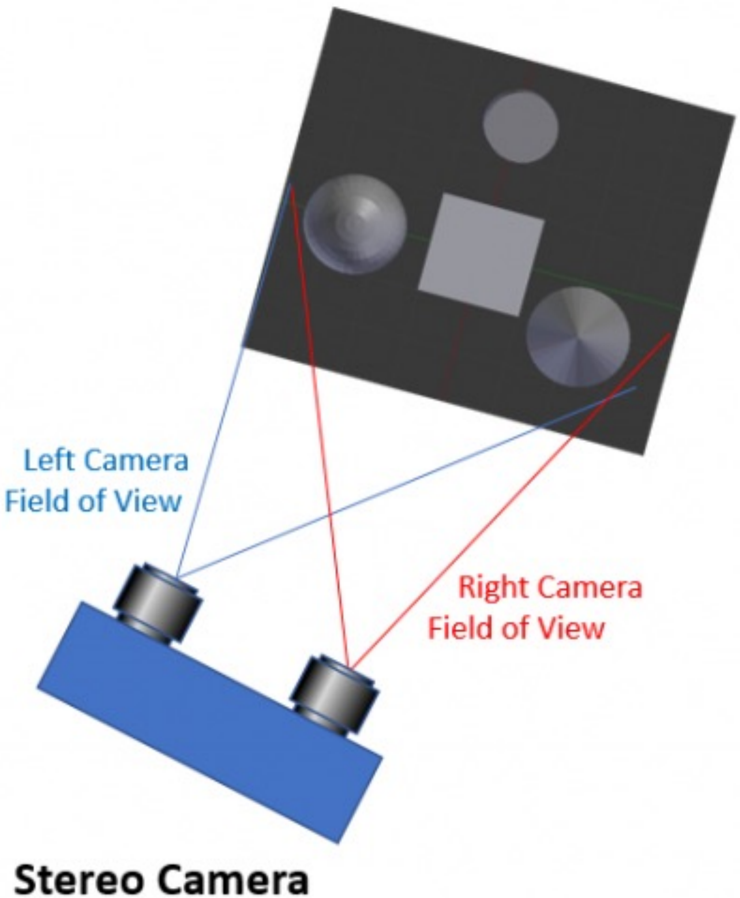
Chen Wang
Spatial AI & Robotics Lab
Department of Computer Science and Engineering

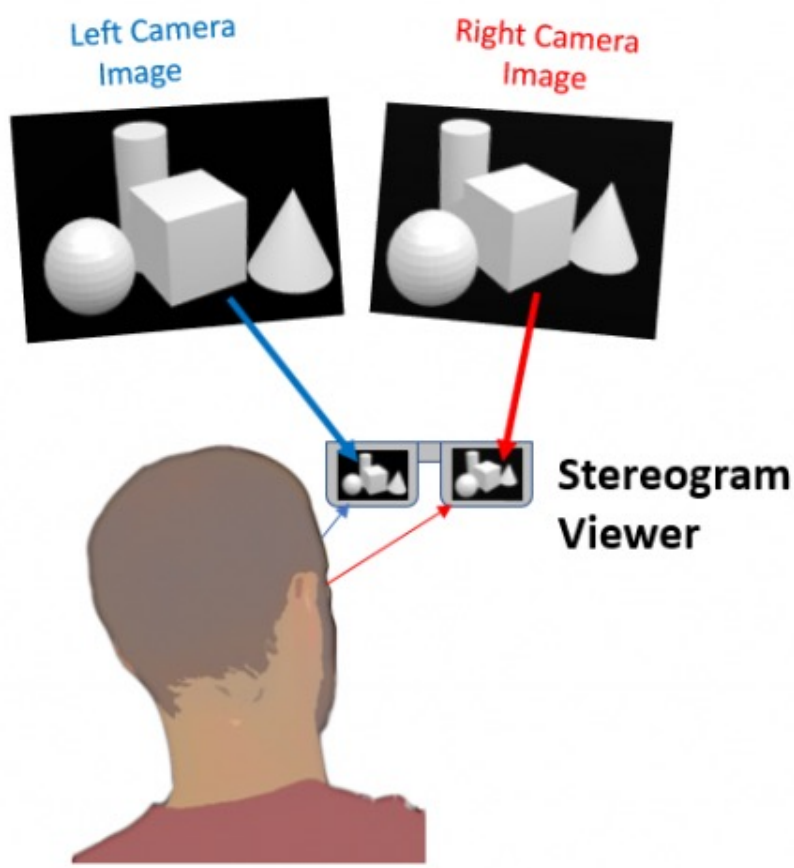
University at Buffalo The State University of New York

Many Slides from Lana Lazebnik

Stereo Vision (binocular camera)

Stereo Vision



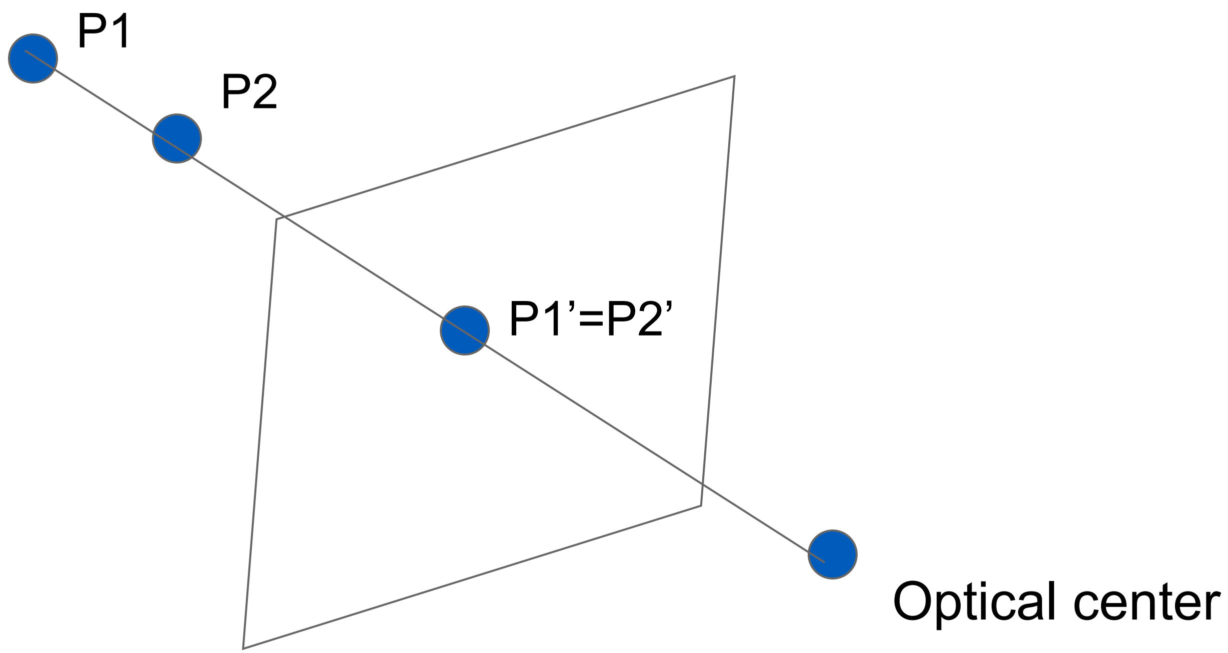


Why multiple views?

- Structure and depth are inherently ambiguous from single views.

Why multiple views?

- Structure and depth are inherently ambiguous from single views.

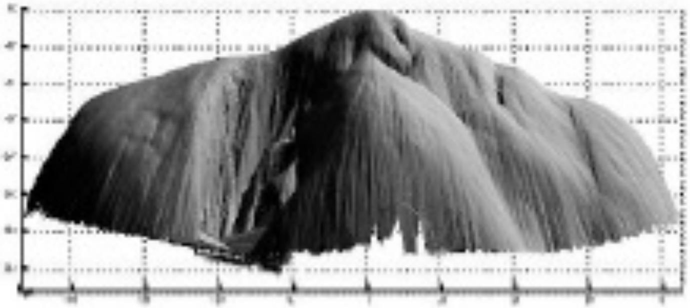


Stereo Vision

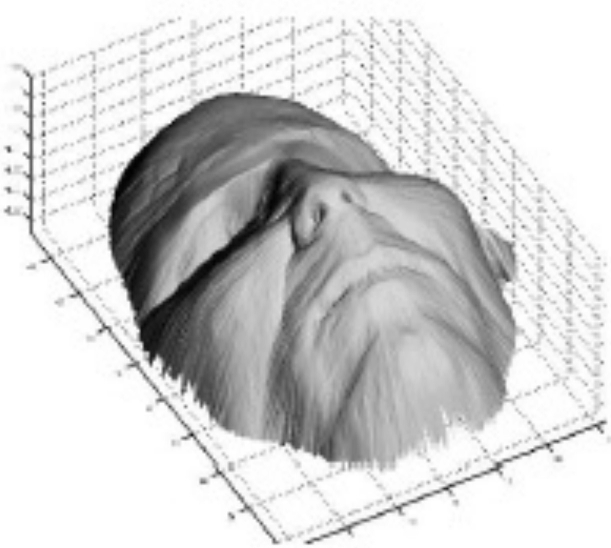
- What cues help perceive 3D shape and depth?
  - Shading
  - Focus/Defocus
  - Texture
  - Perspective
  - Motion
  - Occlusion

Shading

a)



b)



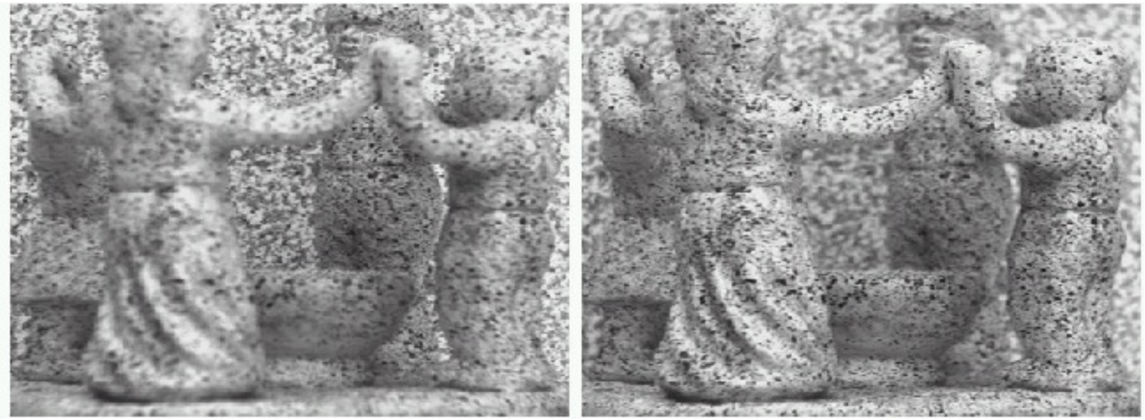
c)

Prados & Faugeras 2006

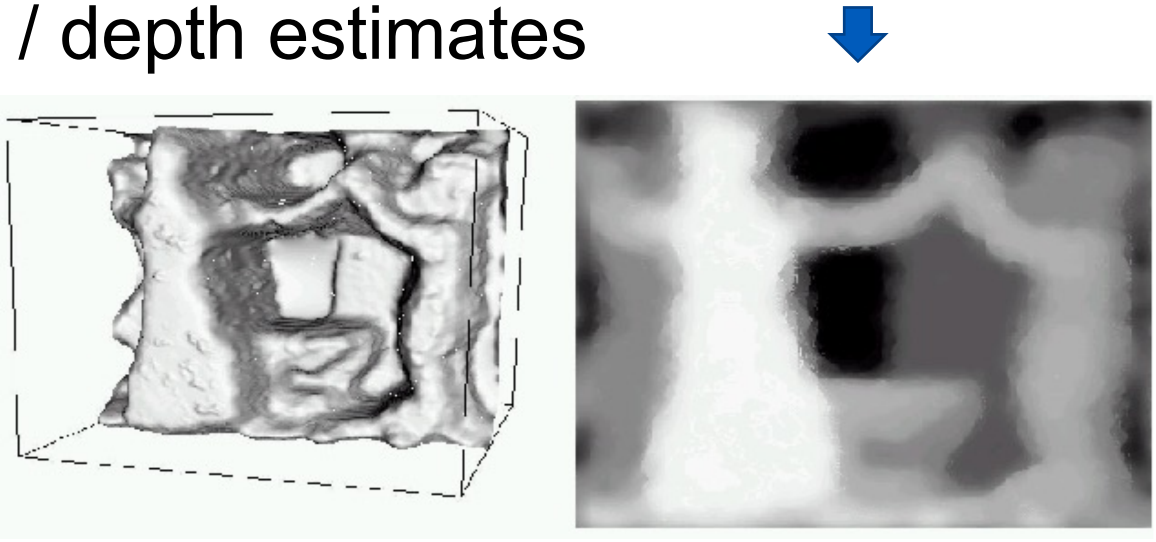
Spatial AI & Robotics Lab

Focus/defocus

- Same point of view, different camera parameters

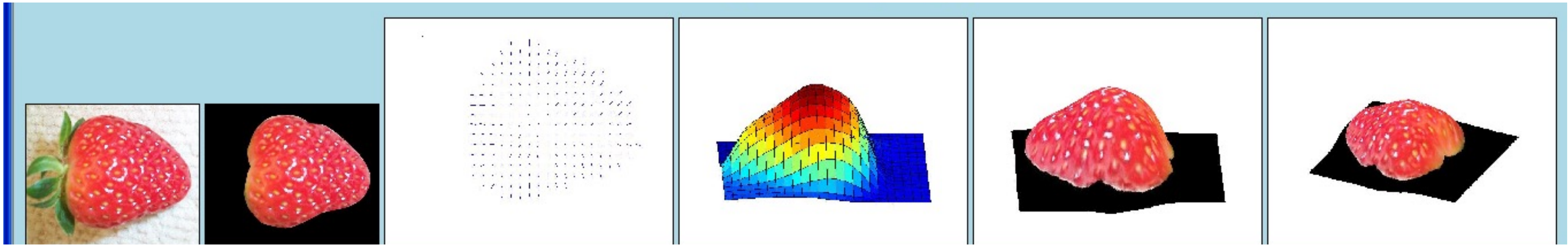


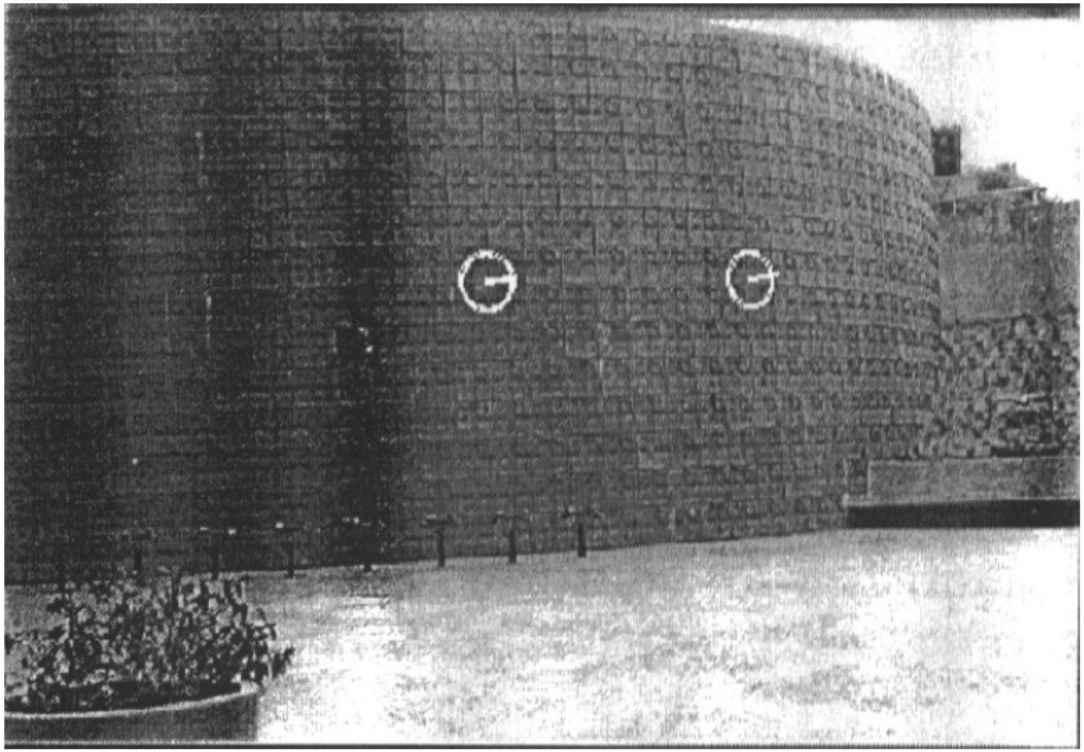
- 3D shape / depth estimates



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

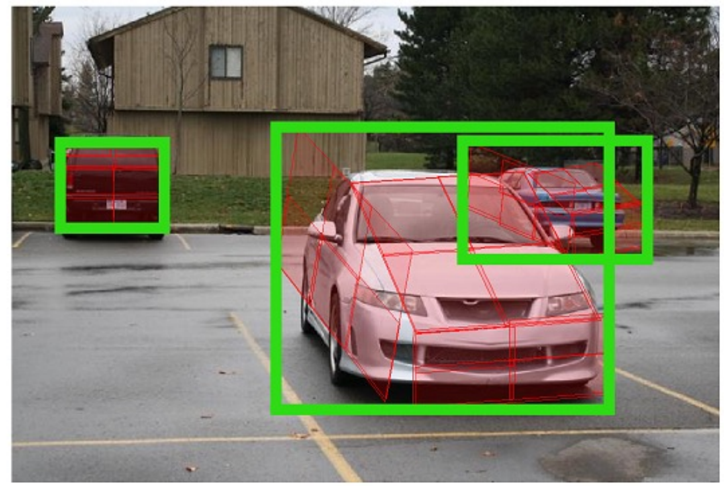
Image credit: S. Seitz

Motion

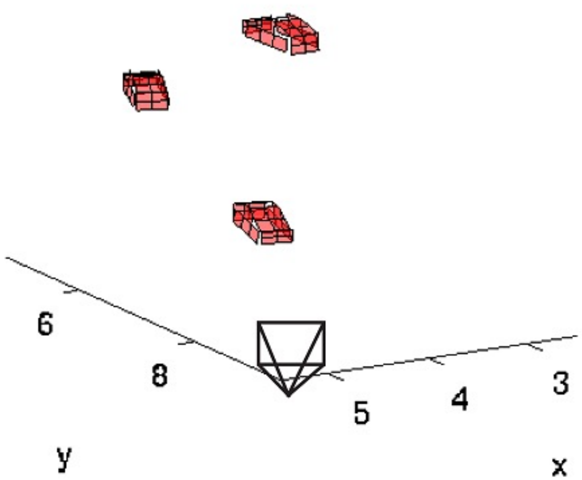
Figures from L. Zhang

Occlusion

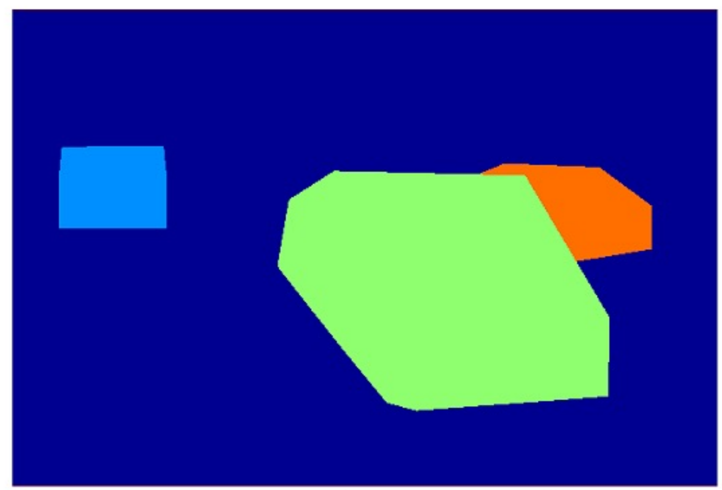
(a) input image



(b) 2D detection



(c) 3D spatial layout



(d) 2D object mask

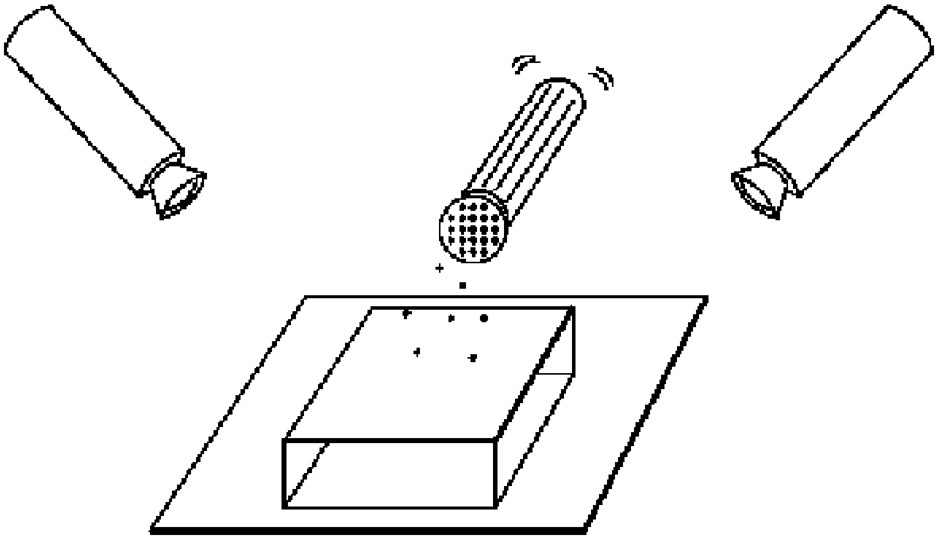
Yu Xiang and Silvio Savarese. Object Detection by 3D Aspectlets and Occlusion Reasoning. In the 4th International IEEE Workshop on 3D Representation and Recognition (3dRR), 2013

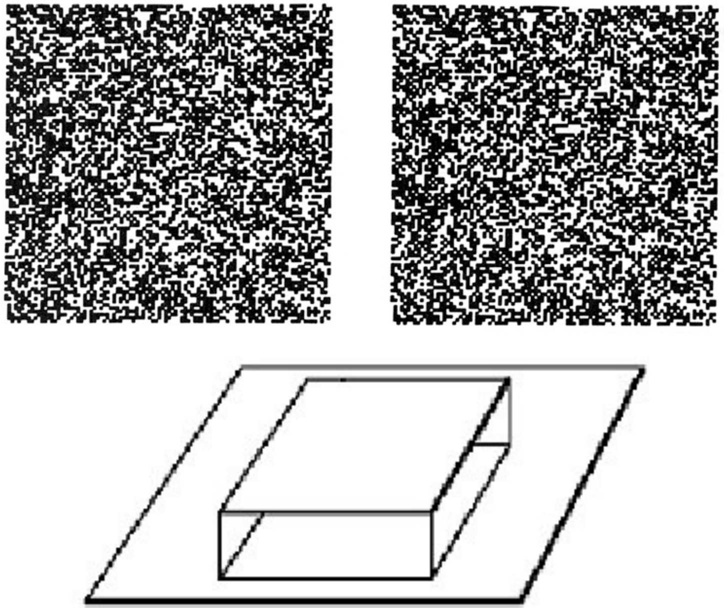
Animal Binocular Systems

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

Random dot stereograms

- Julesz 1960: Do we identify local brightness patterns before fusion (monocular) or after (binocular)?
- To test: pair of synthetic images obtained by randomly spraying black dots on white objects





From Palmer, “Vision Science”, MIT Press

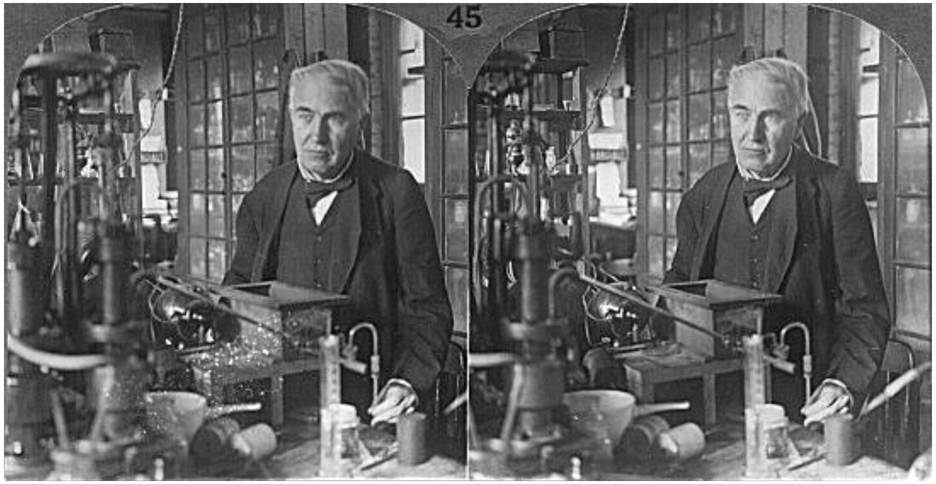
Random dot stereograms

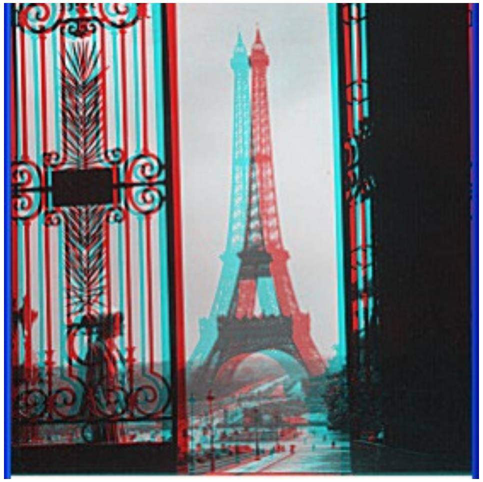
- When viewed monocularly, they appear random; when viewed stereoscopically, see 3D structure.
- Human binocular fusion not directly associated with the physical retinas; must involve the central nervous system.
- High level scene understanding not required for Stereo
- High level scene understanding is arguably *better* than stereo

Stereo photography and stereo viewers

- Invented by Sir Charles Wheatstone, 1838
- Take two pictures of the same subject from two slightly different viewpoints and display so that each eye sees only one of the images

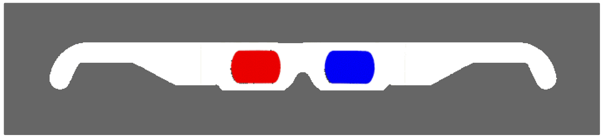
Stereo photography and stereo viewers




http://www.johnsonshawmuseum.org

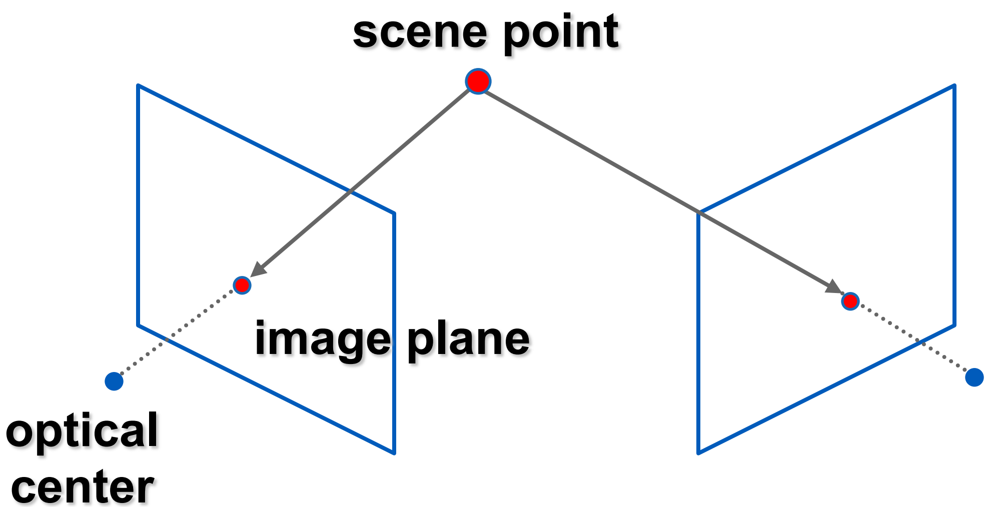
Stereo photography and stereo viewers

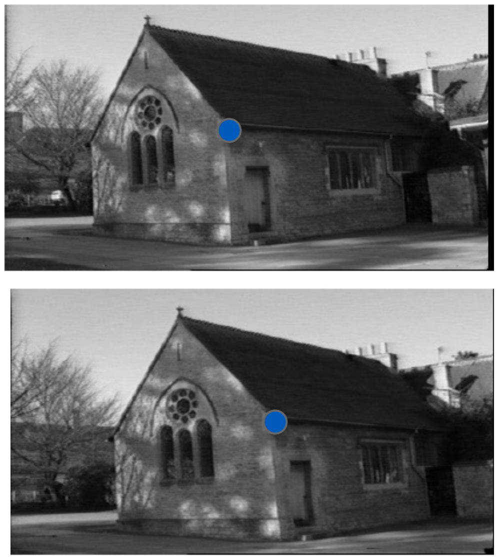


Public Library, Stereoscopic Looking Room, Chicago, by Phillips, 1923

Estimating depth with stereo

- We'll need to consider:
  - Info on camera pose ("calibration")
  - Image point correspondences (feature detection/matching)





Stereo Vision

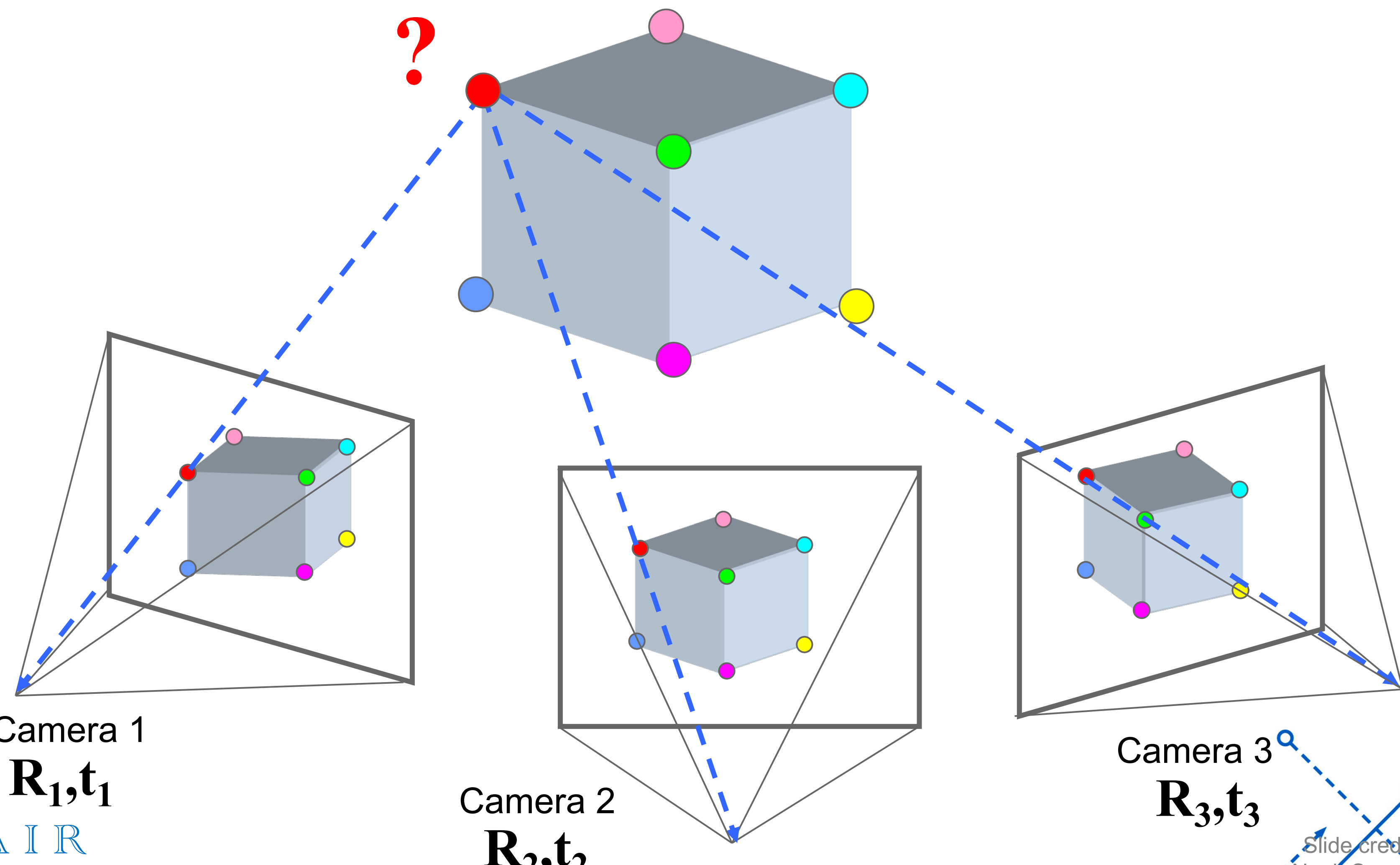
- Structure from motion
  - Shape from “motion” between two views
  - Infer 3D shape of scene from two (multiple) images from different viewpoints

Two cameras, simultaneous views

Single moving camera and static scene

Structure from motion (SFM)

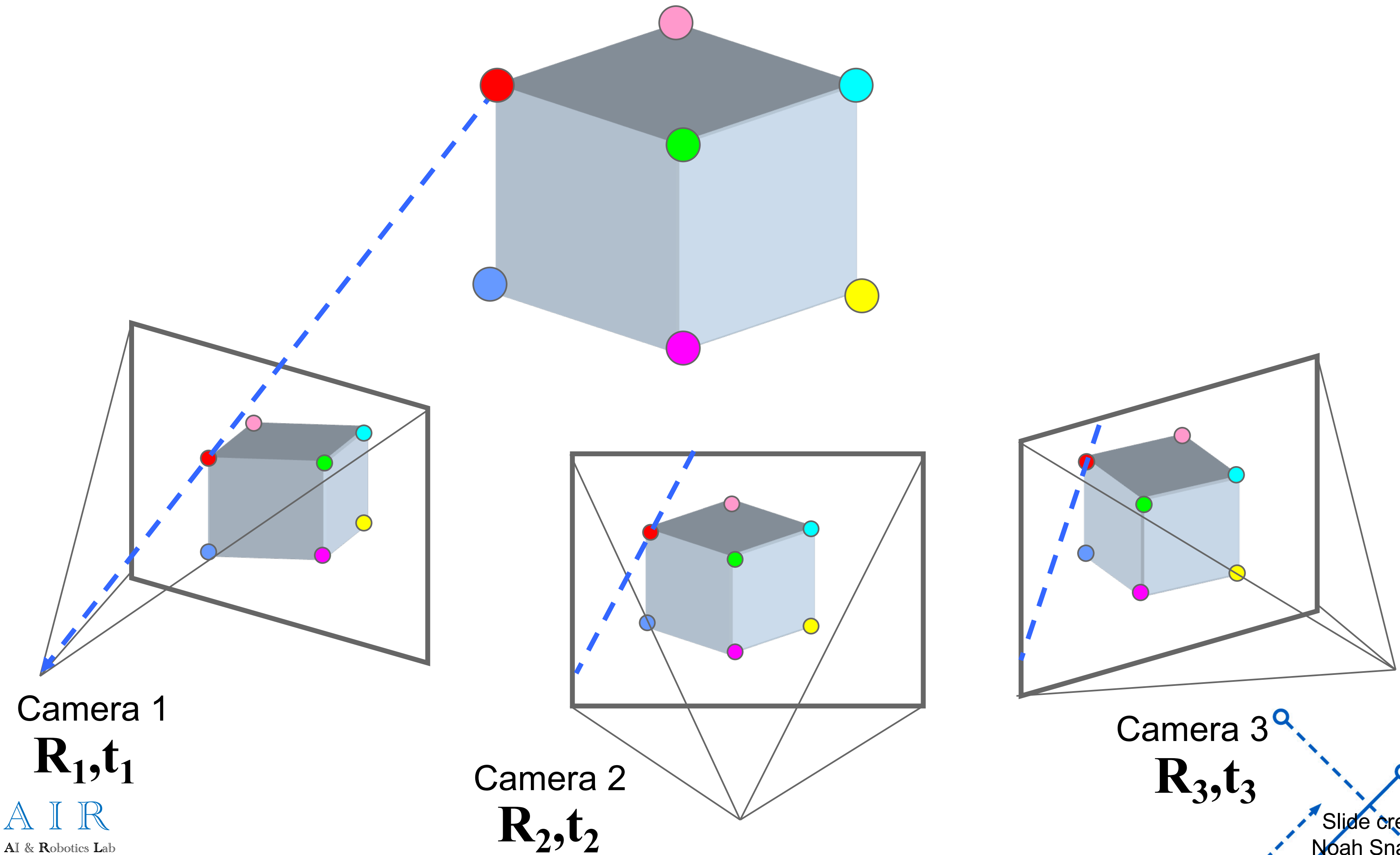
- **Structure:** Given projections of the same 3D point in two or more images, compute the 3D coordinates of that point.



Slide credit: Noah Snavely

Structure from motion (SFM)

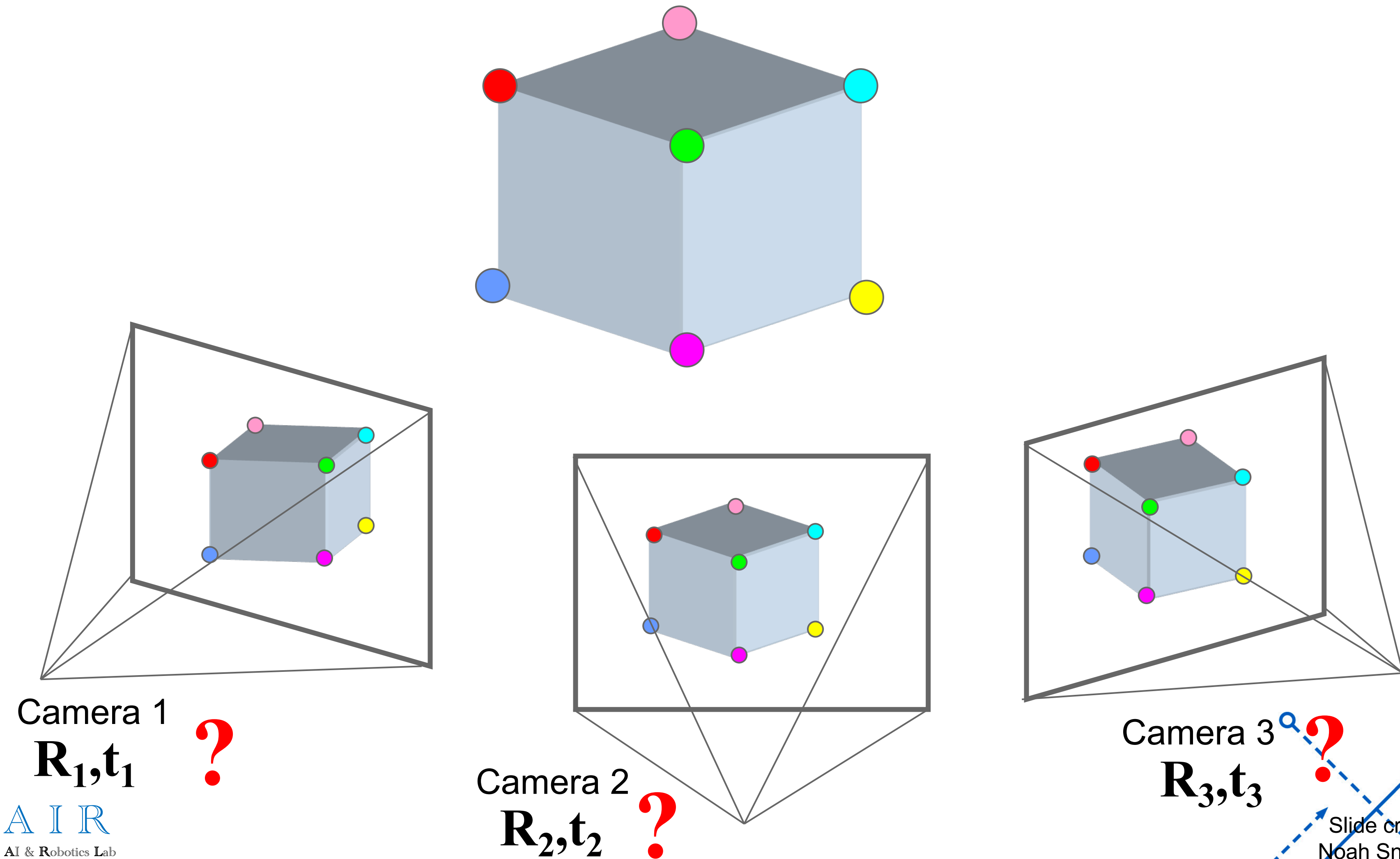
- **Stereo correspondence:** Given a point in one of the images, where could its corresponding points be in the other images?



Slide credit: Noah Snavely

Structure from motion (SFM)

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



SAIR
Spatial AI & Robotics Lab

Slide credit: Noah Snavely

Human eye

- Rough analogy with human visual system:
  - Pupil/Iris: control light passing through lens
  - Retina: contains cells, where image is formed
  - Fovea: highest concentration of cones

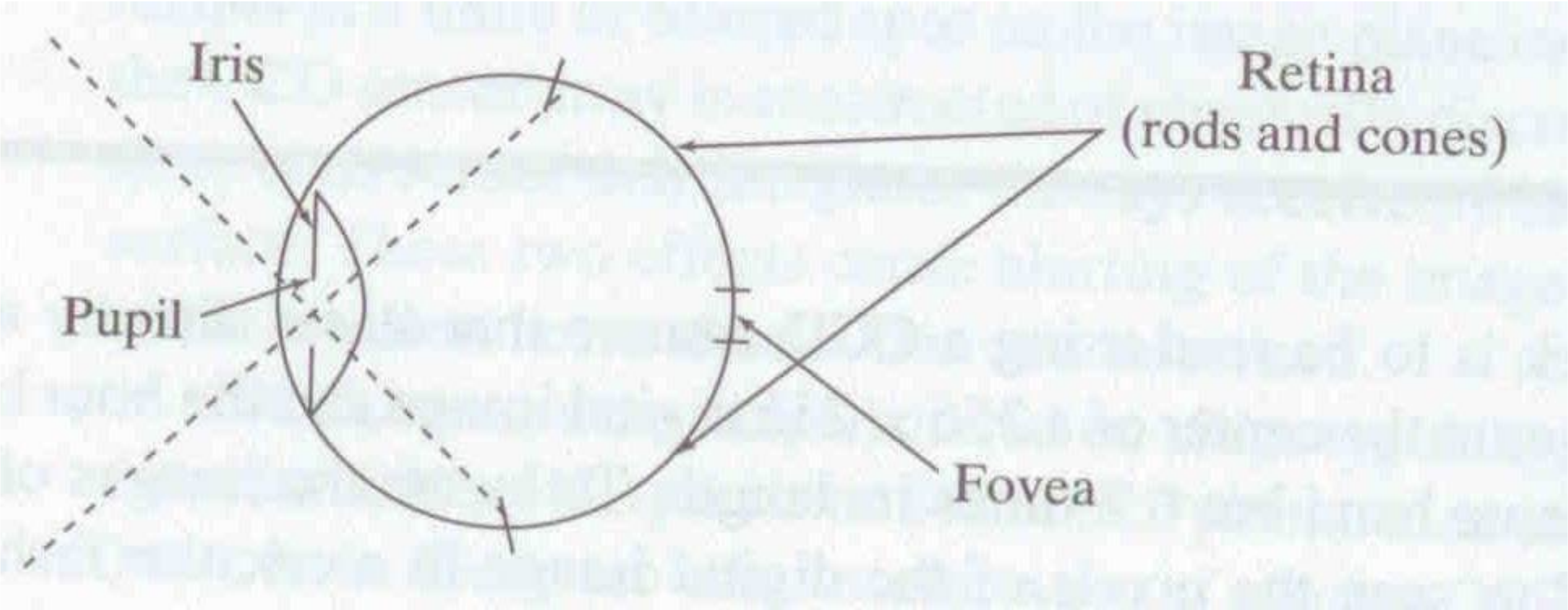
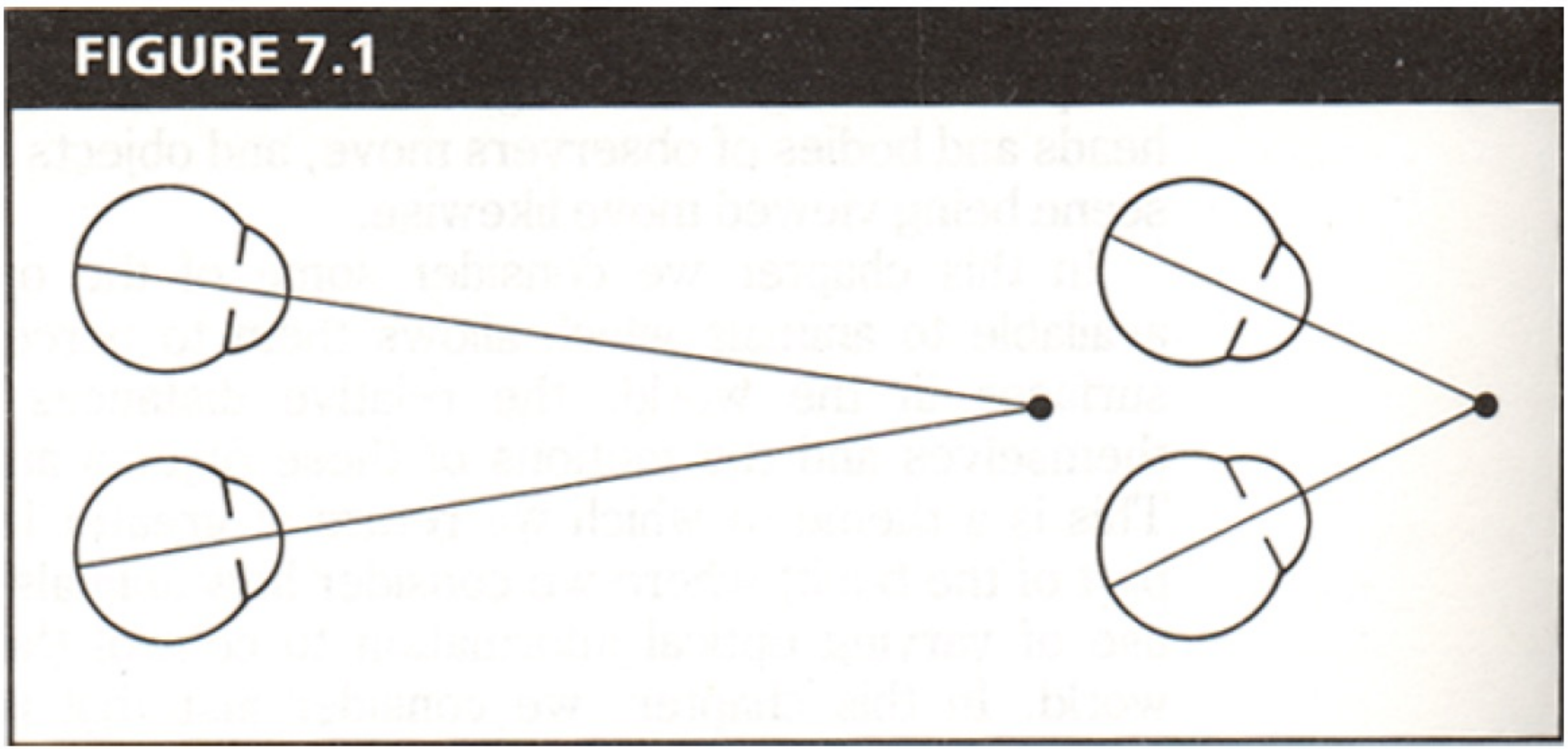


Fig from Shapiro and Stockman

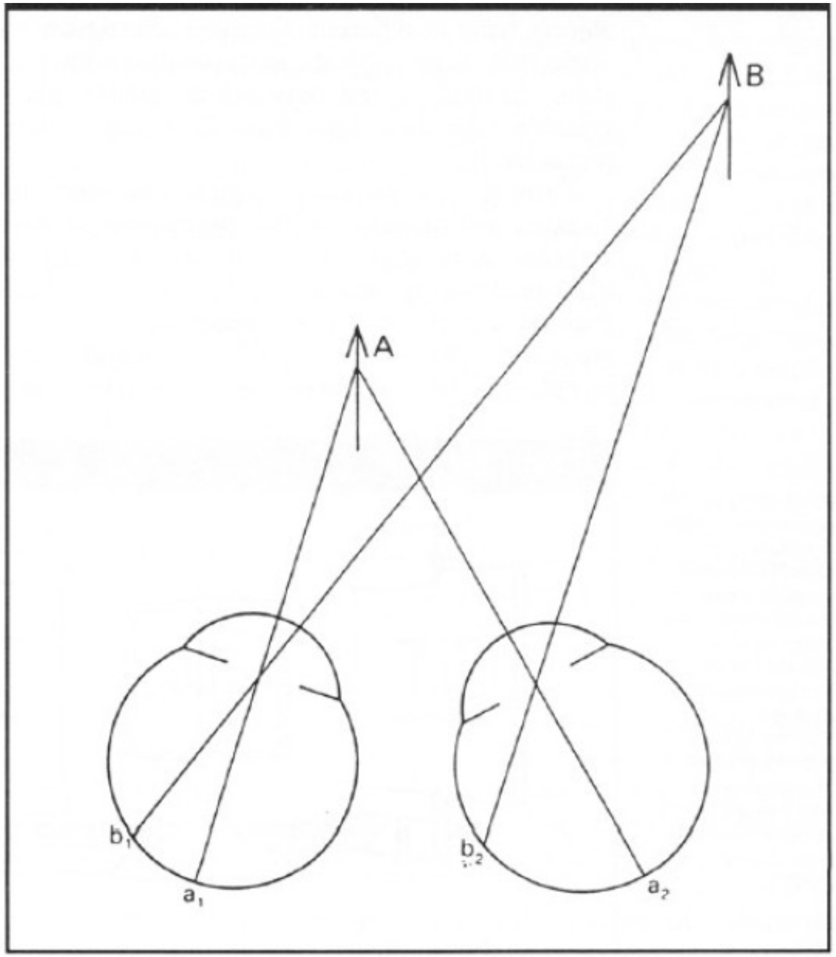
Human stereopsis: disparity

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



Disparity

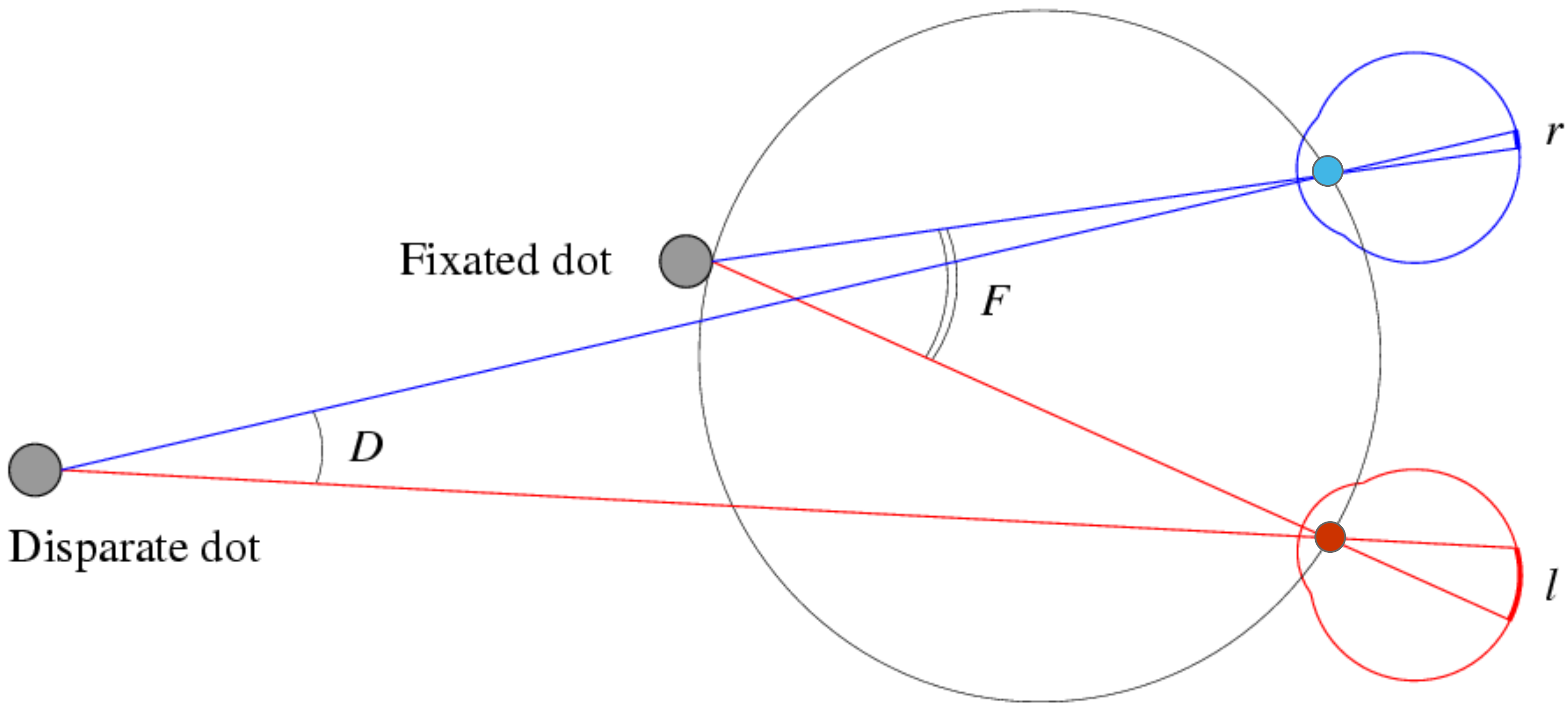
- **Disparity** occurs when eyes fixate on one object; others appear at different visual angles



Disparity

- Disparity:

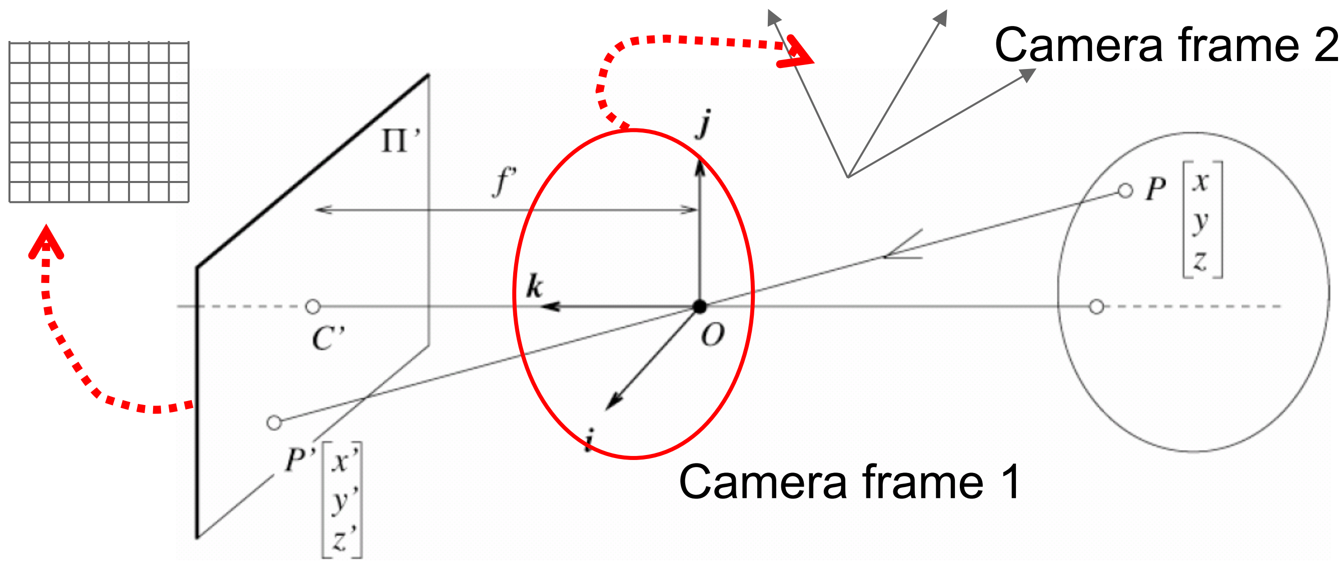
$$d = r - l = D - F.$$



Forsyth & Ponce

Camera parameters (Recap)

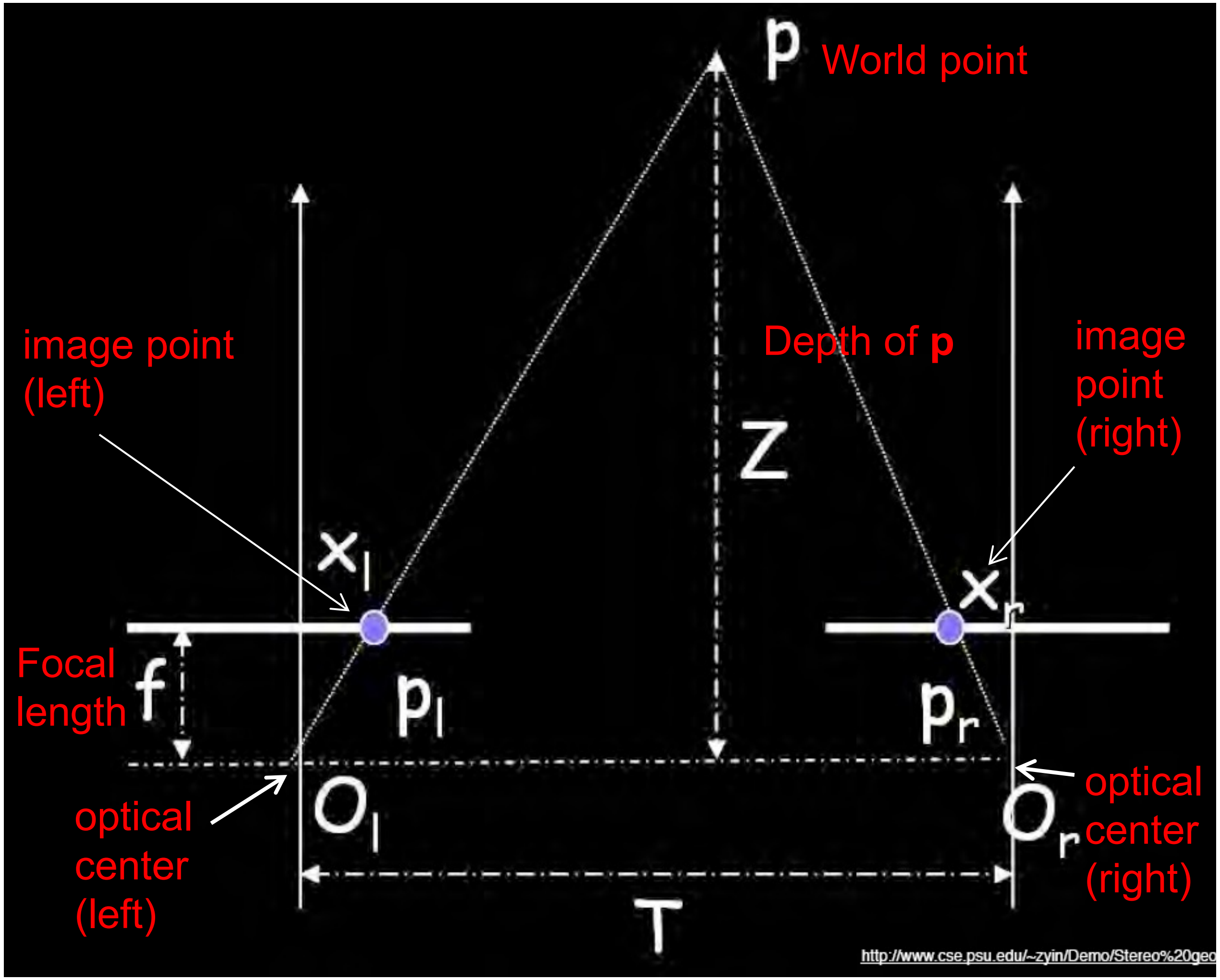
- *Intrinsic* params:
  - Focal length, image center, radial distortion parameters
  - Coordinates relative to camera ←→ Pixel coordinates
- *Extrinsic* params:
  - Rotation matrix and translation vector
  - Camera frame 1 ←→ Camera frame 2



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

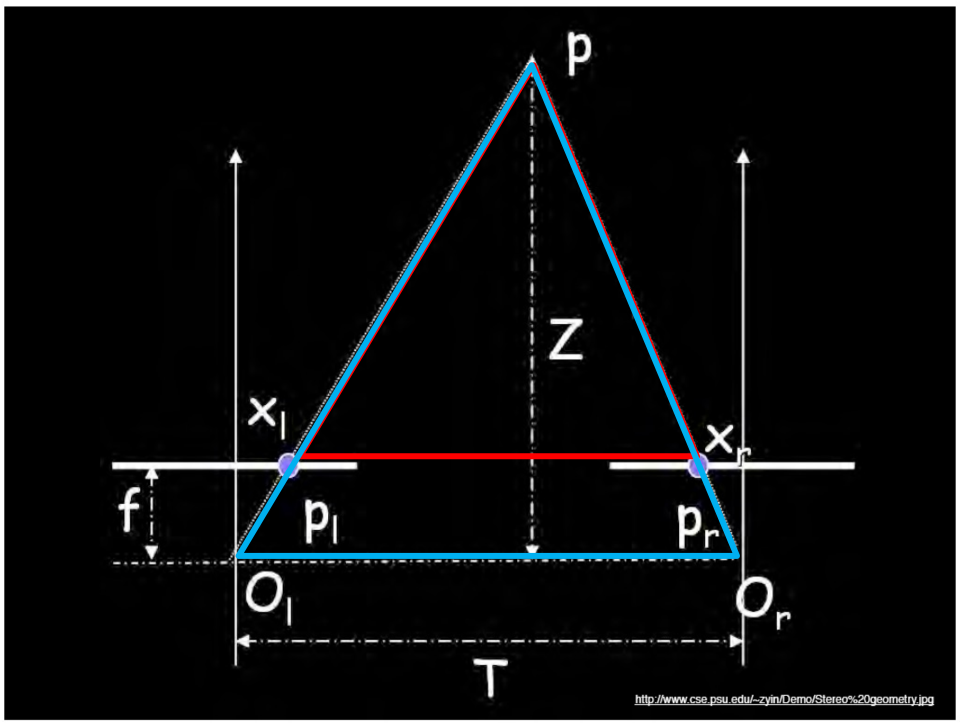
Geometry for a simple stereo system

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



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):



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

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

Disparity

Depth from disparity

- If we could find the **corresponding points** in two images, we could **estimate relative depth.**

Image $I(x, y)$

Disparity map $D(x, y)$

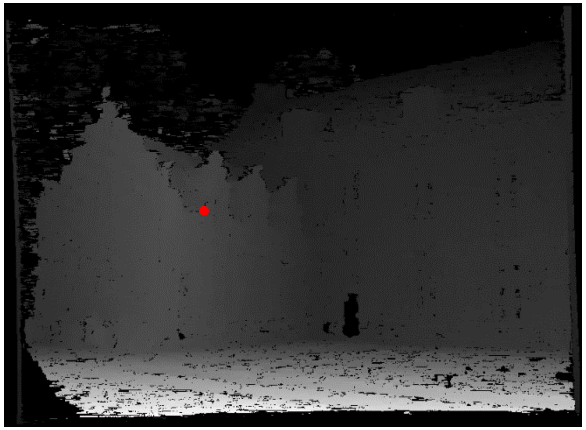
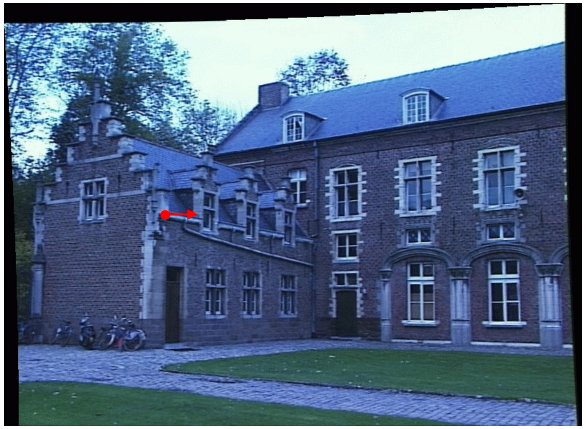


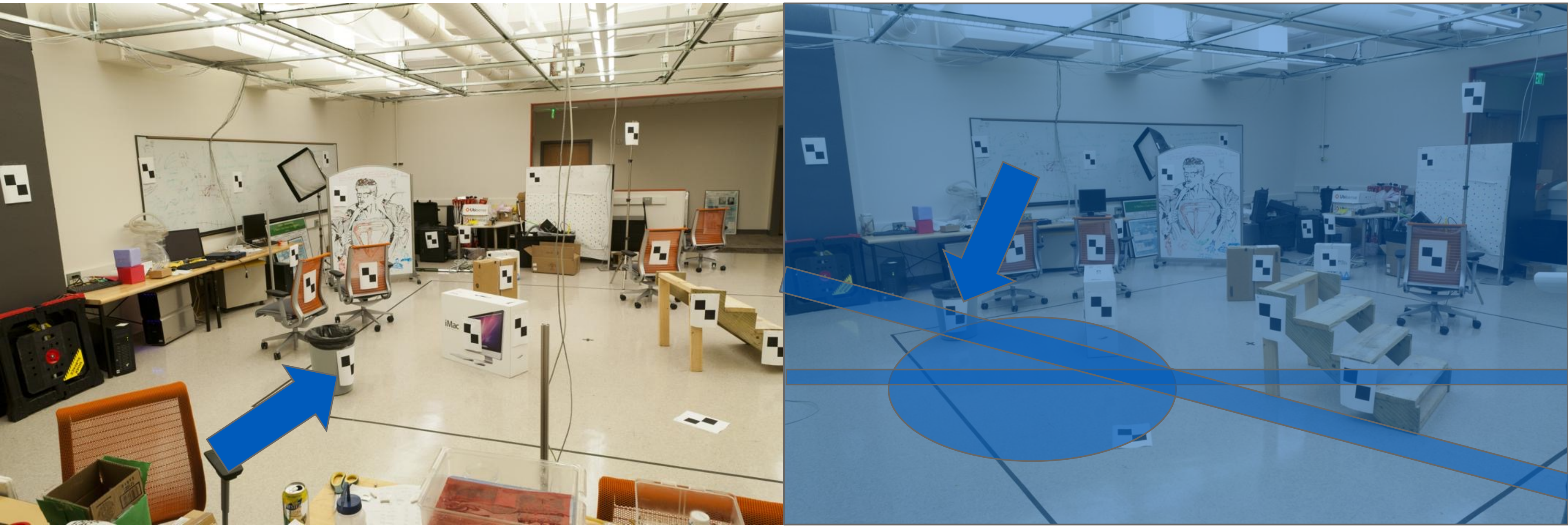
Image $I'(x', y')$



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

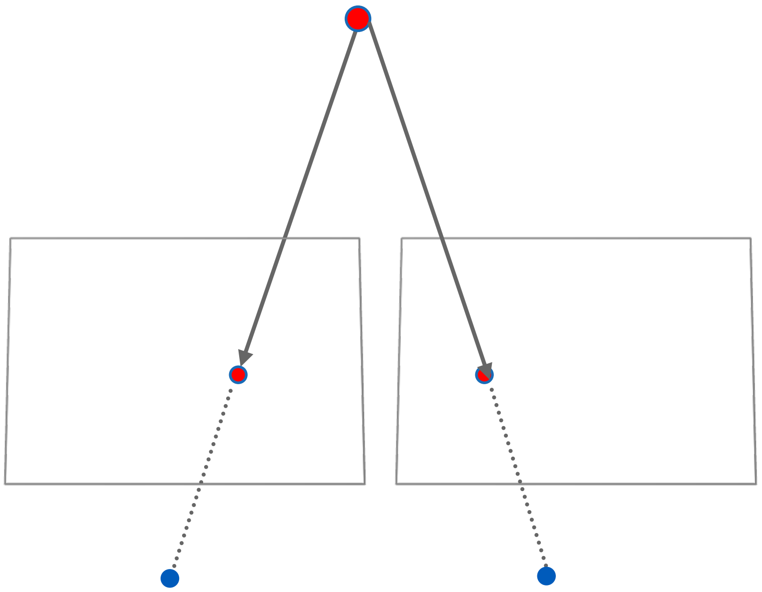
Depth from disparity

- Where do we need to search?

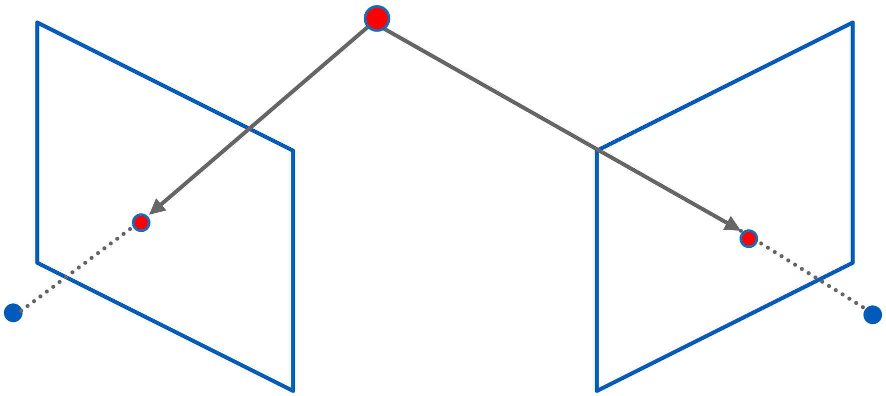


General case (calibrated cameras)

- The two cameras need not have parallel optical axes.



Vs.

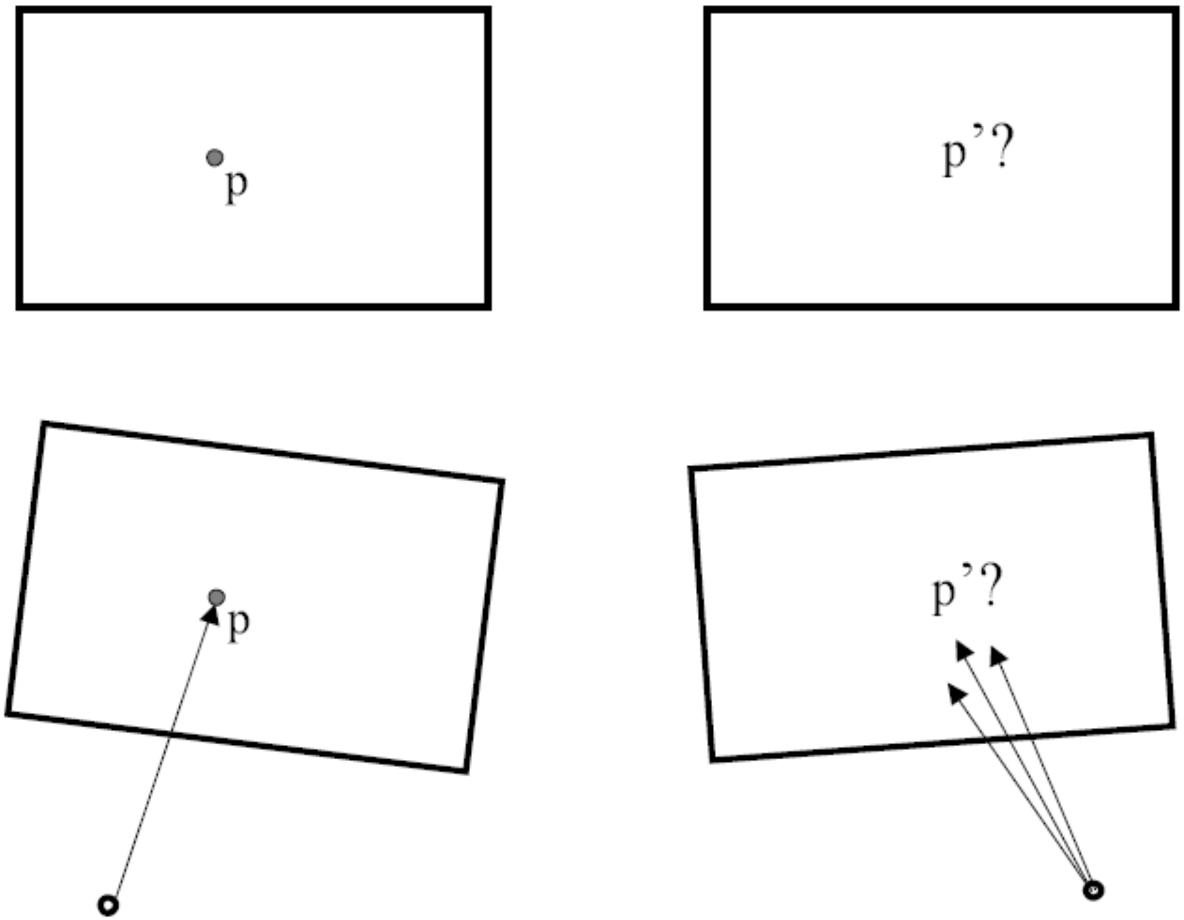


Grauman

Spatial AI & Robotics Lab

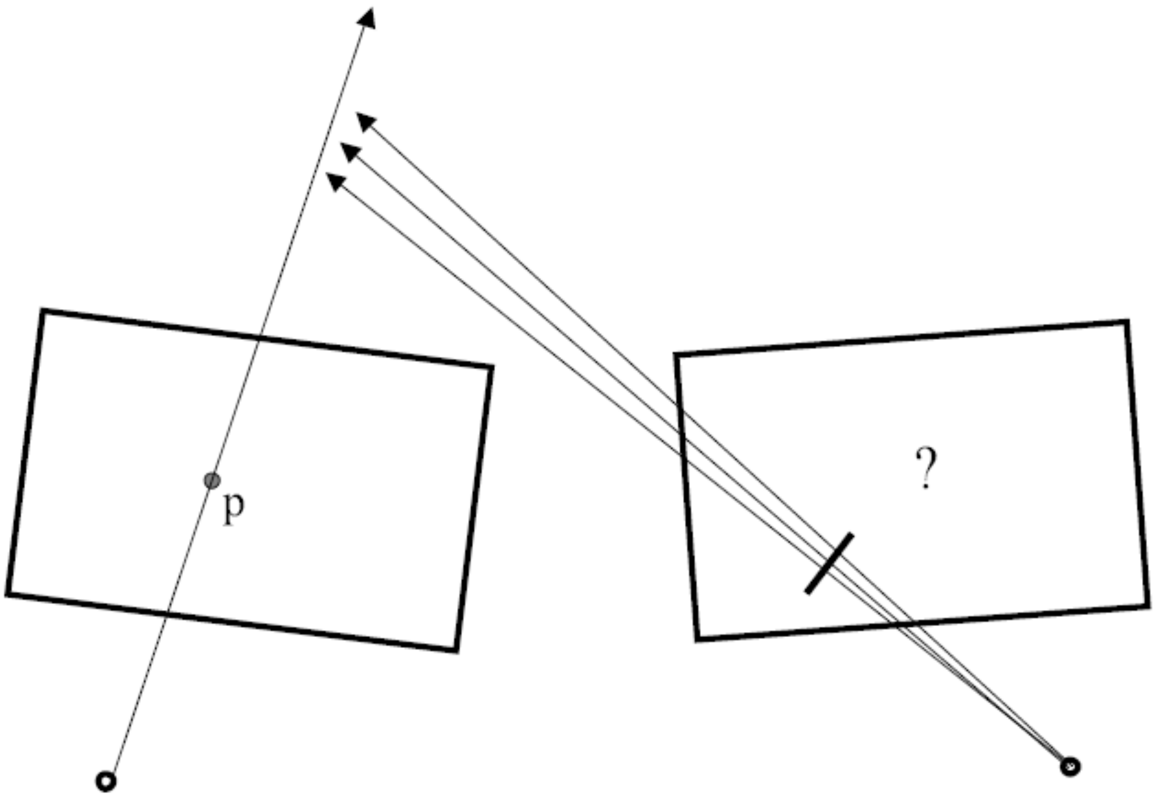
Stereo correspondence constraints

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



Stereo correspondence constraints

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



Correspondence problem

- Multiple match hypotheses satisfy epipolar constraint, but which is correct?

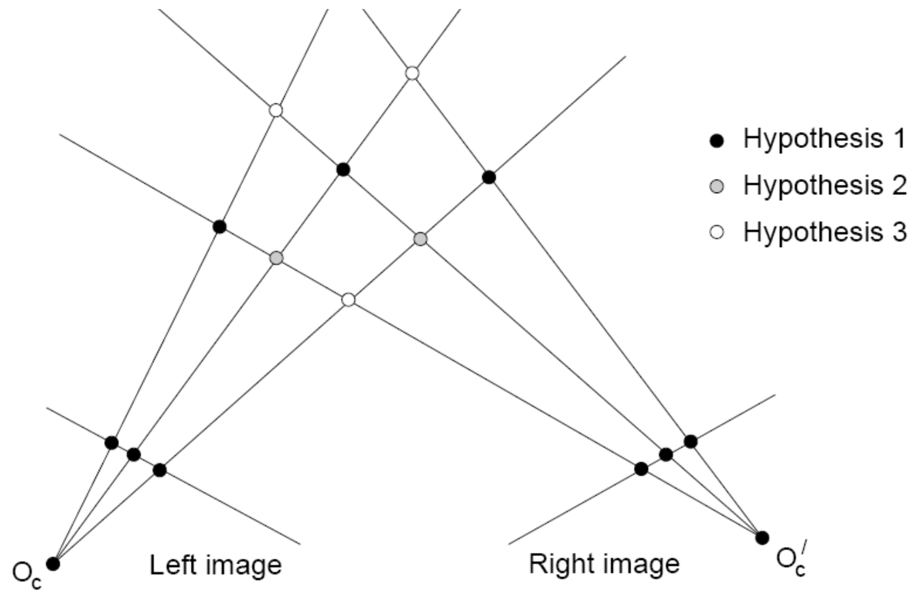
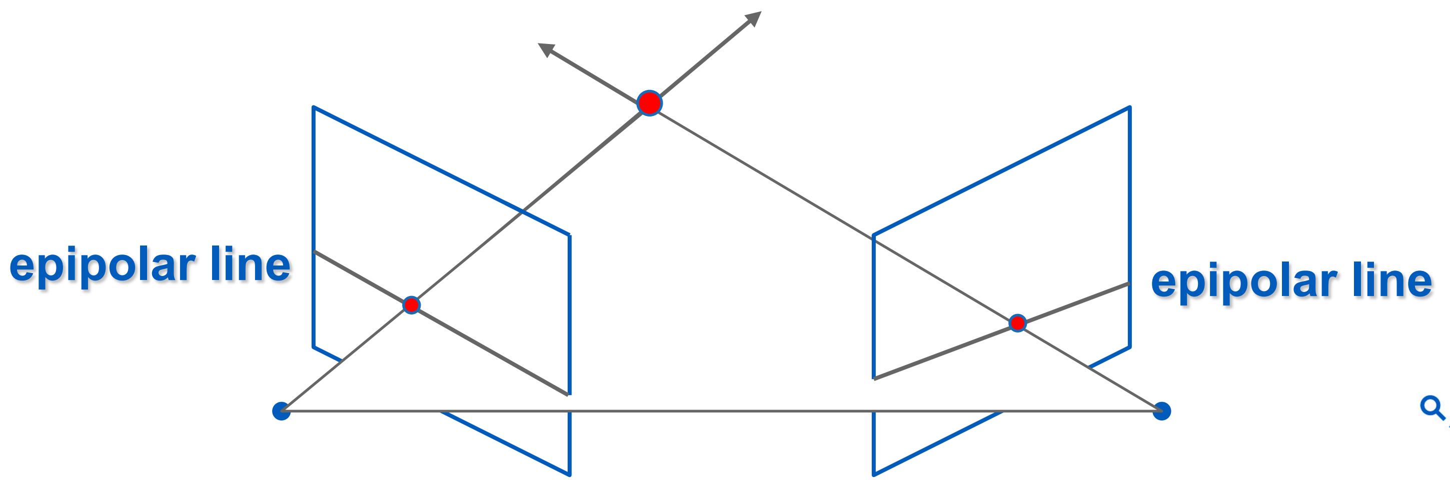


Figure from Gee & Cipolla 1999

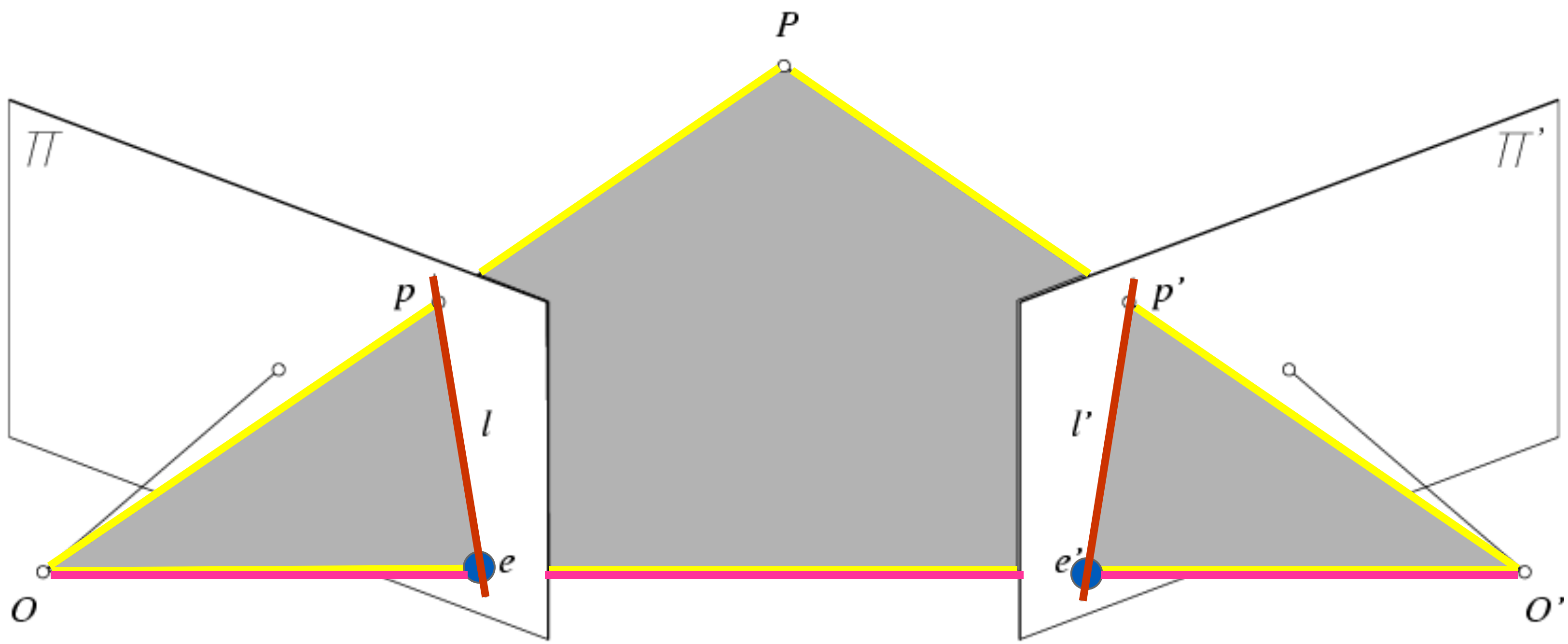
Stereo correspondence constraints

- Geometry of two views allows us to constrain where the corresponding pixel for image points in the first view must occur in the second view.
- Epipolar constraint
  - Reduces correspondence problem to 1D search along conjugate epipolar lines



Adapted from Steve Seitz

Epipolar geometry



- Epipolar Plane
- Baseline
- Epipoles
- Epipolar Lines

M. Pollefeys, UNC

Spatial AI & Robotics Lab

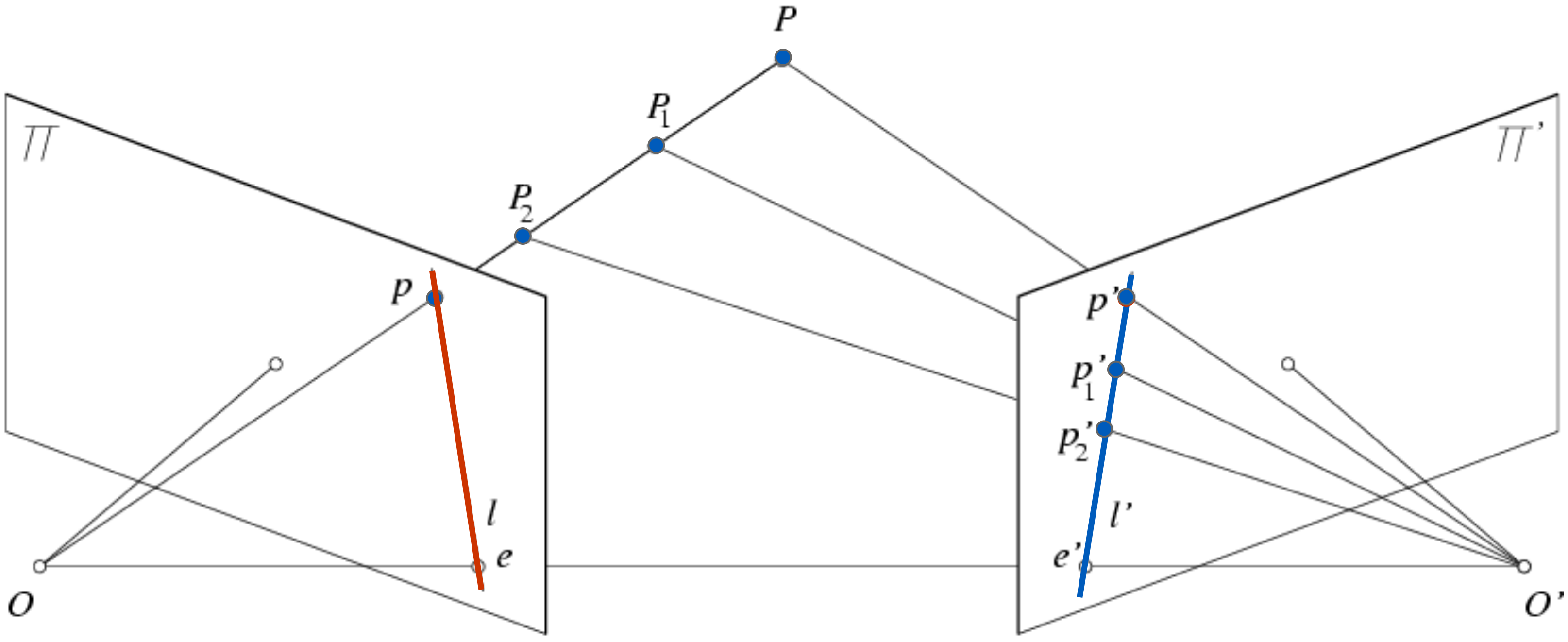
Epipolar geometry: terms

- **Baseline**: line joining the camera centers
- **Epipole**: point of intersection of baseline with the image plane
- **Epipolar plane**: plane containing baseline and world point
- **Epipolar line**: intersection of epipolar plane with image plane
- All epipolar lines intersect at the epipole
- An epipolar plane intersects the left and right image planes in epipolar lines

Grauman

Epipolar constraint

- Potential matches for p must lie on epipolar line l'.
- Potential matches for p' must lie on epipolar line l.

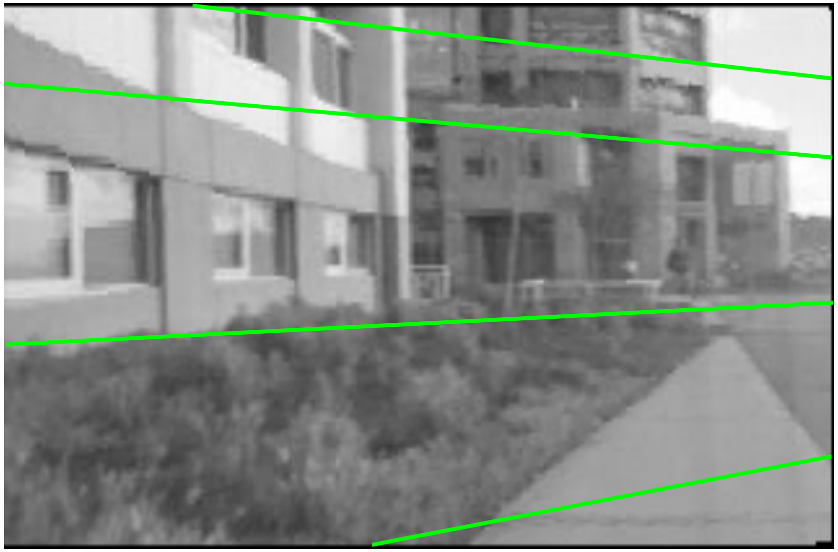


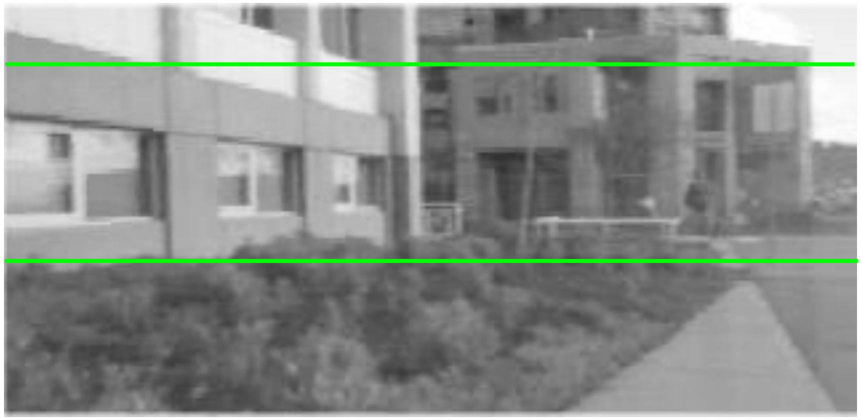
Source: M. Pollefeys

S A I R
Spatial AI & Robotics Lab

Rectification

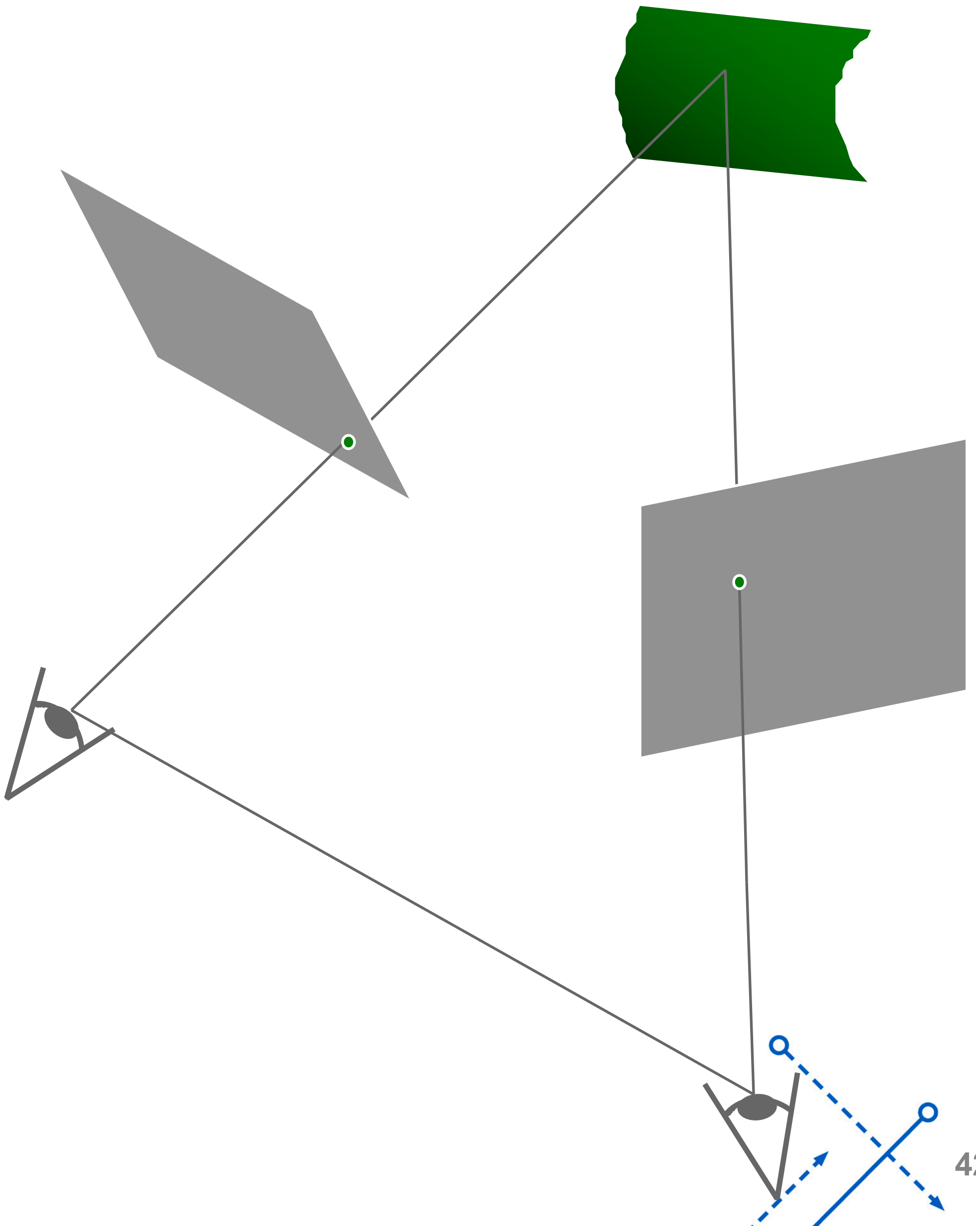
- Searching along epipolar lines at arbitrary orientation is intuitively expensive.
- We prefer searching along the image row.
  - Epipolar lines parallel to the rows of the image.
- This transformation is called *rectification*.





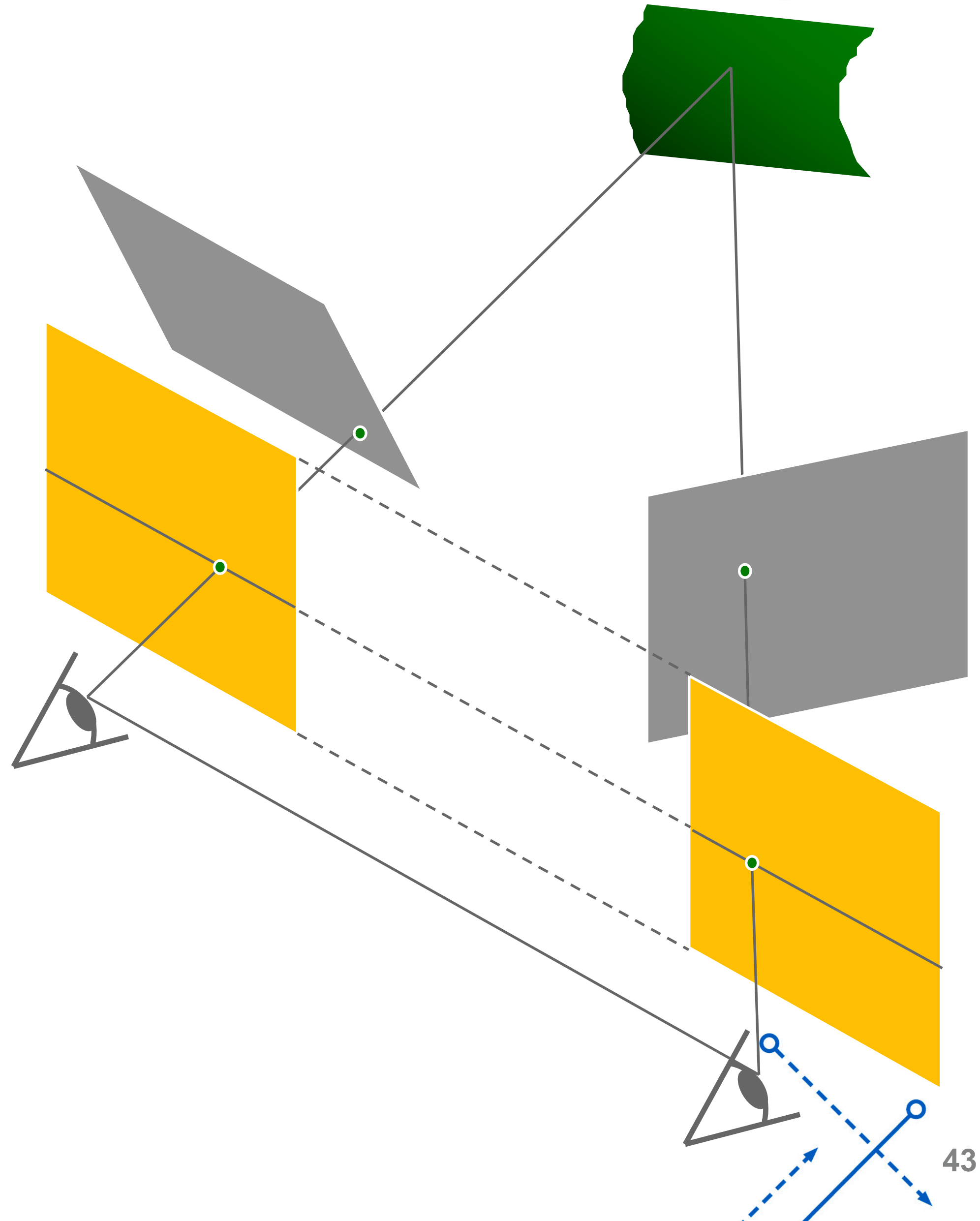
Stereo Image Rectification

- Reproject image planes onto a common plane parallel to the line between optical centers



Stereo Image Rectification

- Reproject image planes onto a common plane parallel to the line between optical centers
- pixel motion is horizontal after this transformation



Stereo Image Rectification

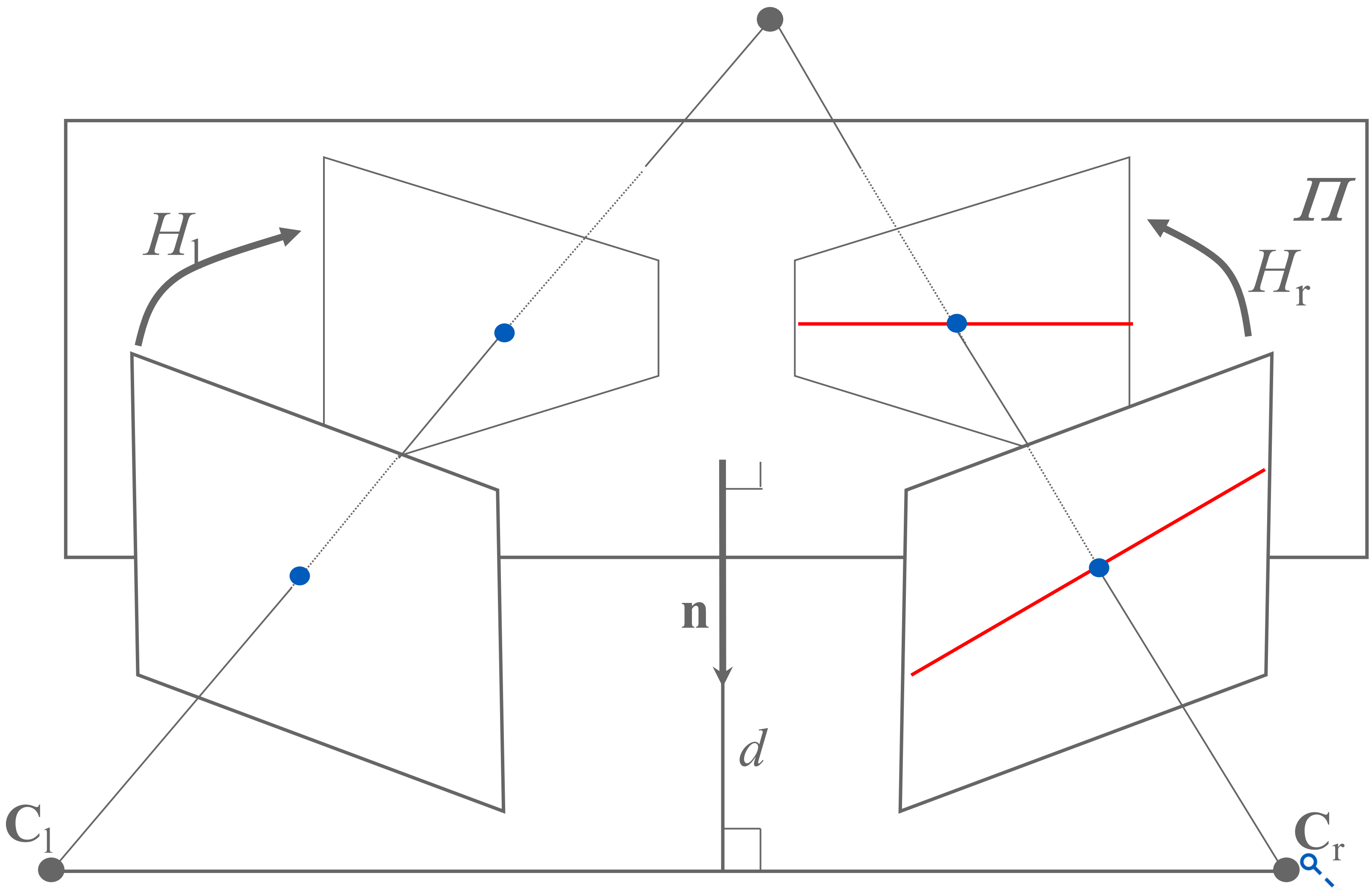
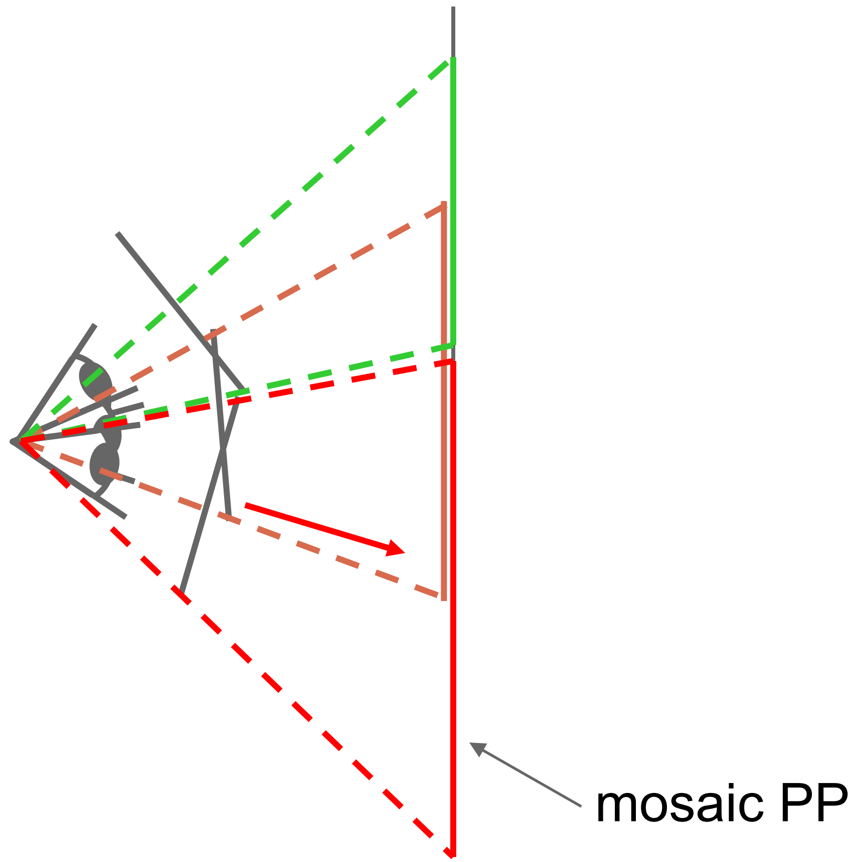


Image reprojection (Recap)

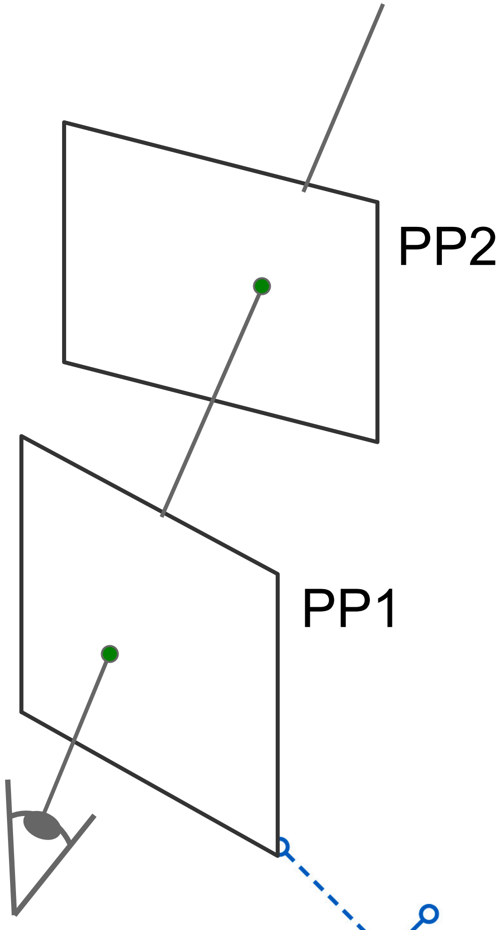


- The mosaic has a natural interpretation in 3D
  - The images are reprojected onto a common plane
  - The mosaic is formed on this plane
  - Mosaic is a *synthetic wide-angle camera*

Source: Steve Seitz

Homography (Recap)

- How to relate two images from the same camera center?
    - how to map a pixel from PP1 to PP2?
- Take as a 2D **image warp** using projective transform.
- A **projective transform** is a mapping between any two PPs with the same center of projection
    - rectangle should map to arbitrary quadrilateral
    - parallel lines aren't preserved.
    - but straight lines are preserved.
- Called **Homography**

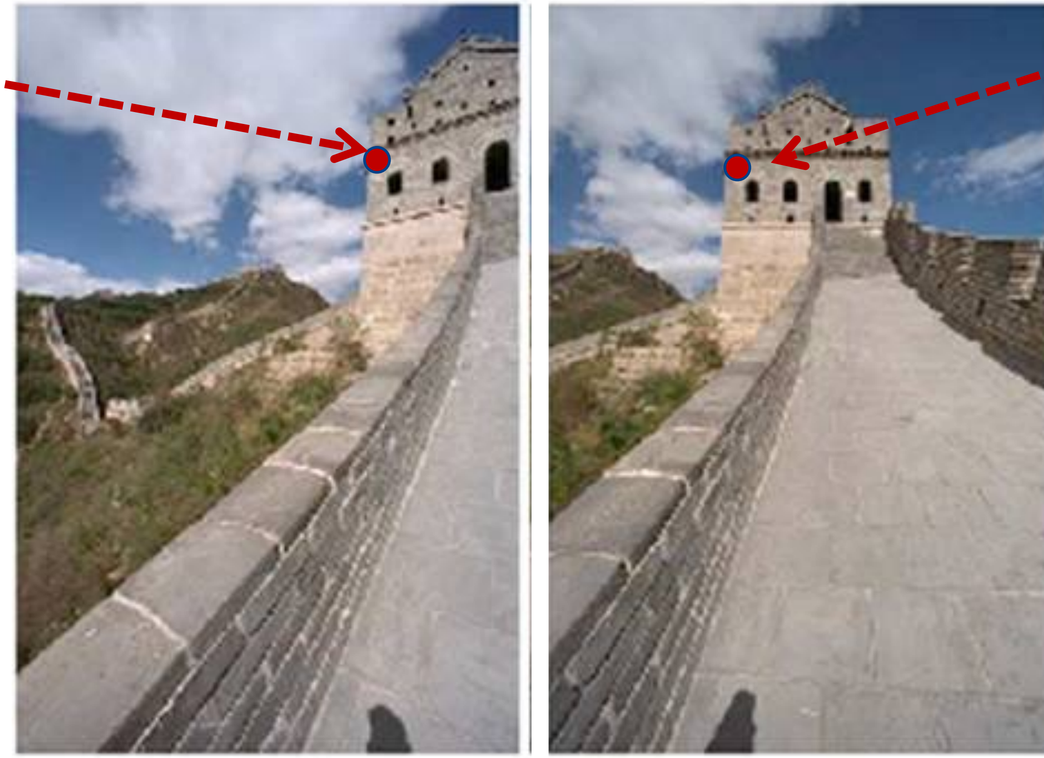


$$\underbrace{\begin{bmatrix} wx' \\ wy' \\ w \end{bmatrix}}_{\mathbf{p'}} = \underbrace{\begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \end{bmatrix}}_{\mathbf{H}} \underbrace{\begin{bmatrix} x \\ y \\ 1 \end{bmatrix}}_{\mathbf{p}}$$

Source: Alyosha Efros

Solving for homographies (Recap)

$$(x, y)$$



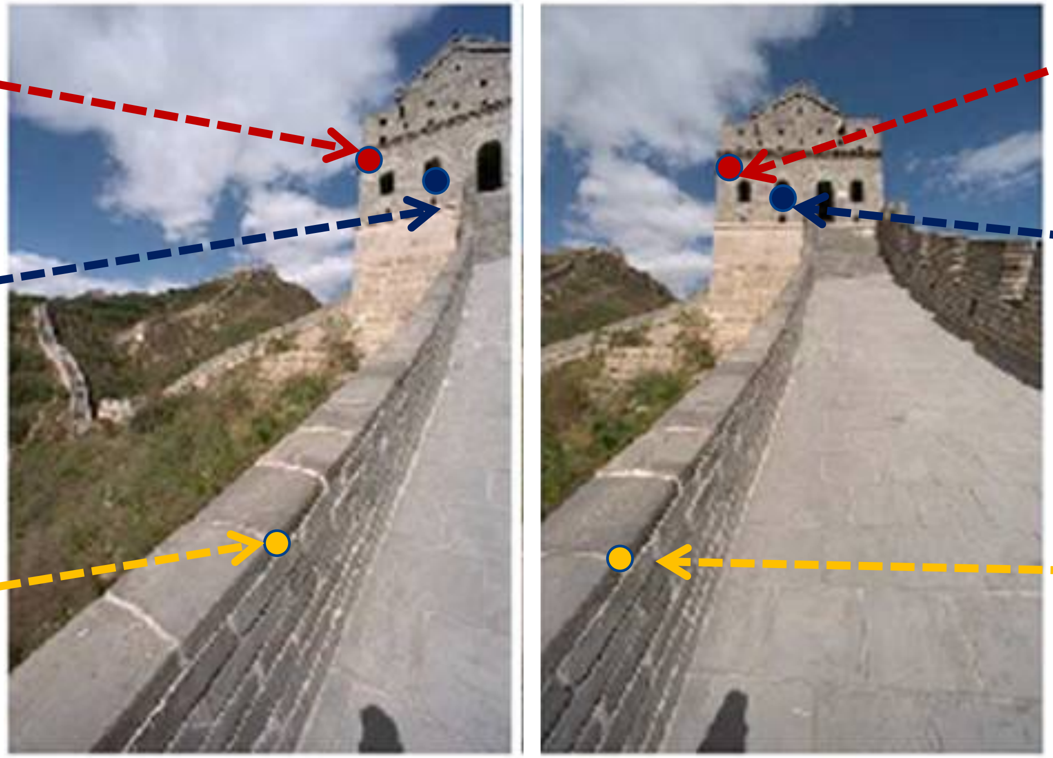
$$\left(\frac{wx'}{w}, \frac{wy'}{w} \right) = (x', y')$$

To **apply** a given homography **H**

- Compute **p' = Hp** (regular matrix multiply)
- Convert **p'** from homogeneous to image coordinates

$$\underbrace{\begin{bmatrix} wx' \\ wy' \\ w \end{bmatrix}}_{\mathbf{p'}} = \underbrace{\begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \end{bmatrix}}_{\mathbf{H}} \underbrace{\begin{bmatrix} x \\ y \\ 1 \end{bmatrix}}_{\mathbf{p}}$$

Solving for homographies (Recap)



To **compute** the homography given pairs of corresponding points, we need to set up an equation where the parameters of **H** are the unknowns…

Solving for homographies (Recap)

p' = Hp

$$\begin{bmatrix} wx' \\ wy' \\ w \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

- Can set scale factor *i*=1 or ||H|| = 1. So, there are 8 unknowns.
- Set up a system of linear equations:
  - **Ah = b**

  where vector of unknowns h = $[a, b, c, d, e, f, g, h]^T$
- Need at least 8 equations (4 points), but the more the better…
- Solve for H. If over-constrained, solve using least-squares:

$$\min \|Ah - b\|^2$$

$$h = (A^T A)^{-1} A^T b$$

Grauman

Proof of least squares (Recap)

- $F(h) = ||Ah - b||^2 = (Ah - b)^T \, (Ah - b)$
- $F(h) = h^T A^T Ah - h^T A^T b - b^T Ah + b^T b$
- $\frac{\partial}{\partial h} F(h) = 2A^T Ah - A^T b - (b^T A)\text{^}T$
- Setting derivative to 0: $\frac{\partial}{\partial h} F(h)$=0
- $A^T Ah = A^T b$
- $h = (A^T A)^{-1} A^T b$

Contents

- Epipolar Geometry
  - Disparity
  - Epipolar line
  - Epipolar constraints
  - Rectification