COLLEGE OF ENGINEERING
ELECTRICAL ENGINEERING & COMPUTER SCIENCE
UNIVERSITY OF MICHIGAN

Basic Stereo & Epipolar Geometry

EECS 598-08 Fall 2014
Foundations of Computer Vision

Instructor: Jason Corso (jjcorso)
web.eecs.umich.edu/~jjcorso/t/598F14

Readings: FP 7; SZ 11.1; TV 7
Date: 10/22/14

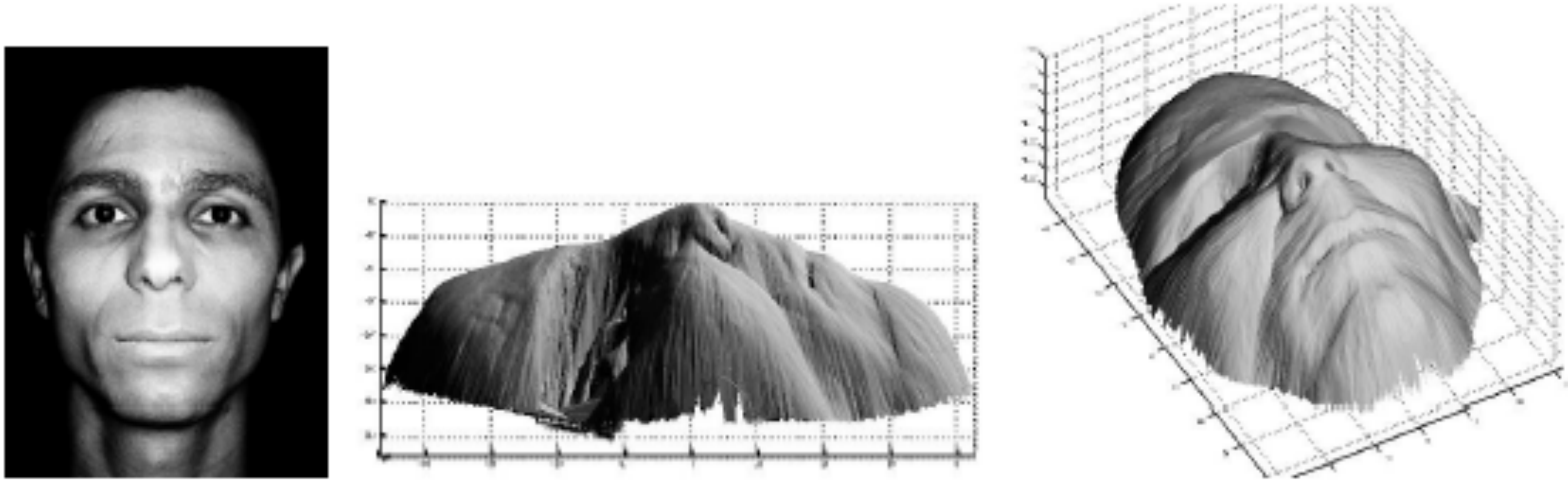
Plan

- Basic Stereo
  - Why is more than one view useful?
  - Triangulation
- Epipolar Geometry

On to Shape

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

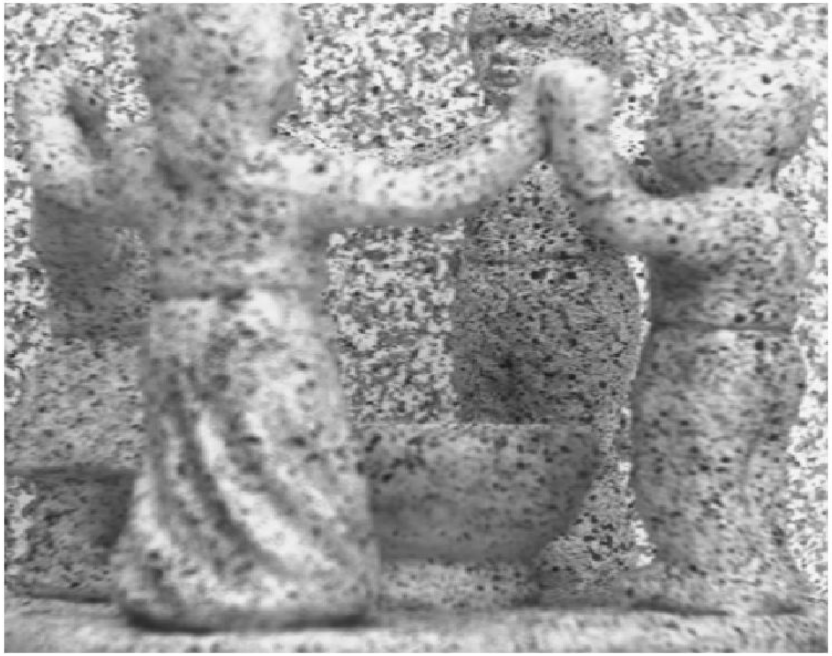
Shading

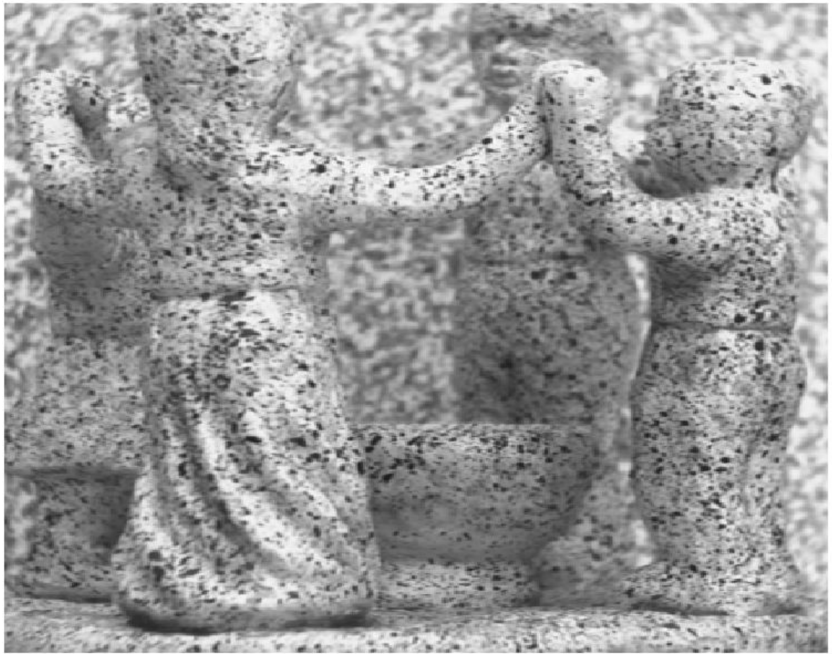


a) b) c)

[Figure from Prados & Faugeras 2006]

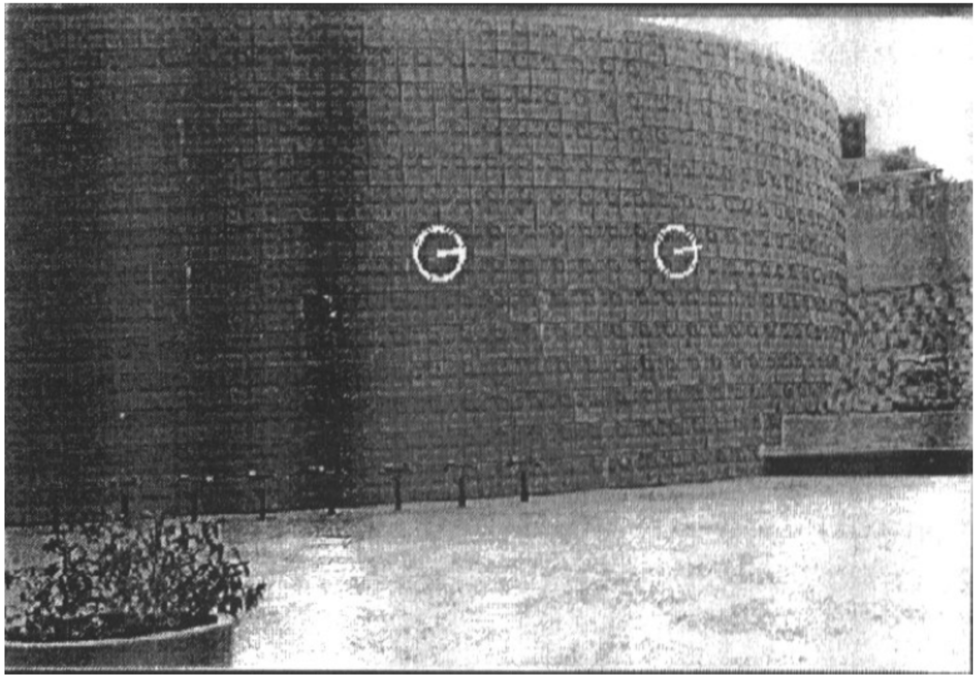
Focus/Defocus

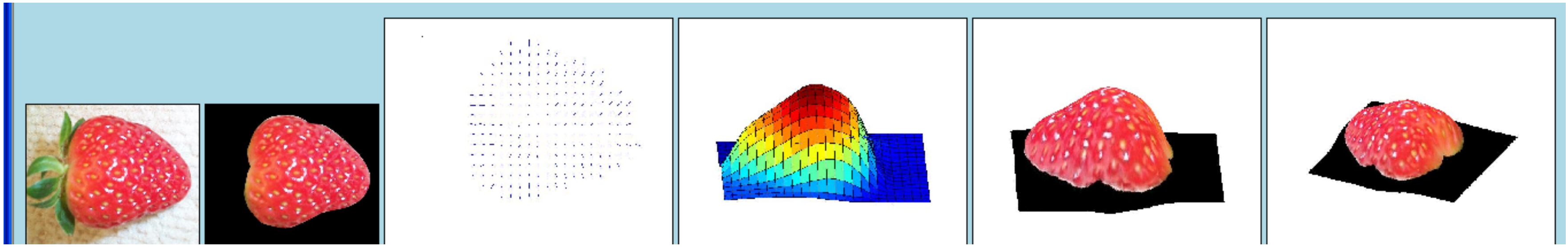




[Figure from H. Jin and P. Favaro, 2002]

Texture



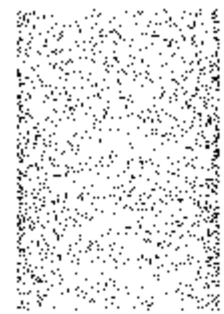


[From A.M. Loh. The recovery of 3-D structure using visual texture patterns. PhD thesis]

Perspective effects

Image credit: S. Seitz

Motion



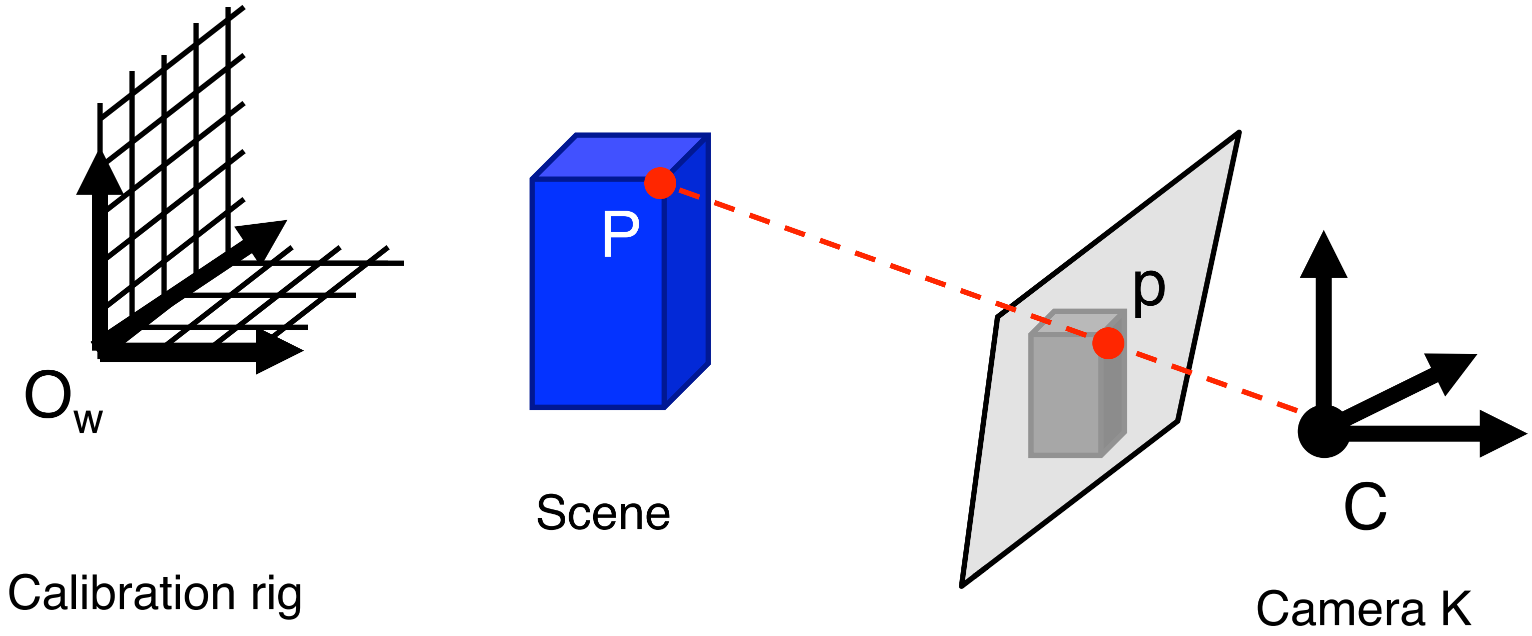
Figures from L. Zhang

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

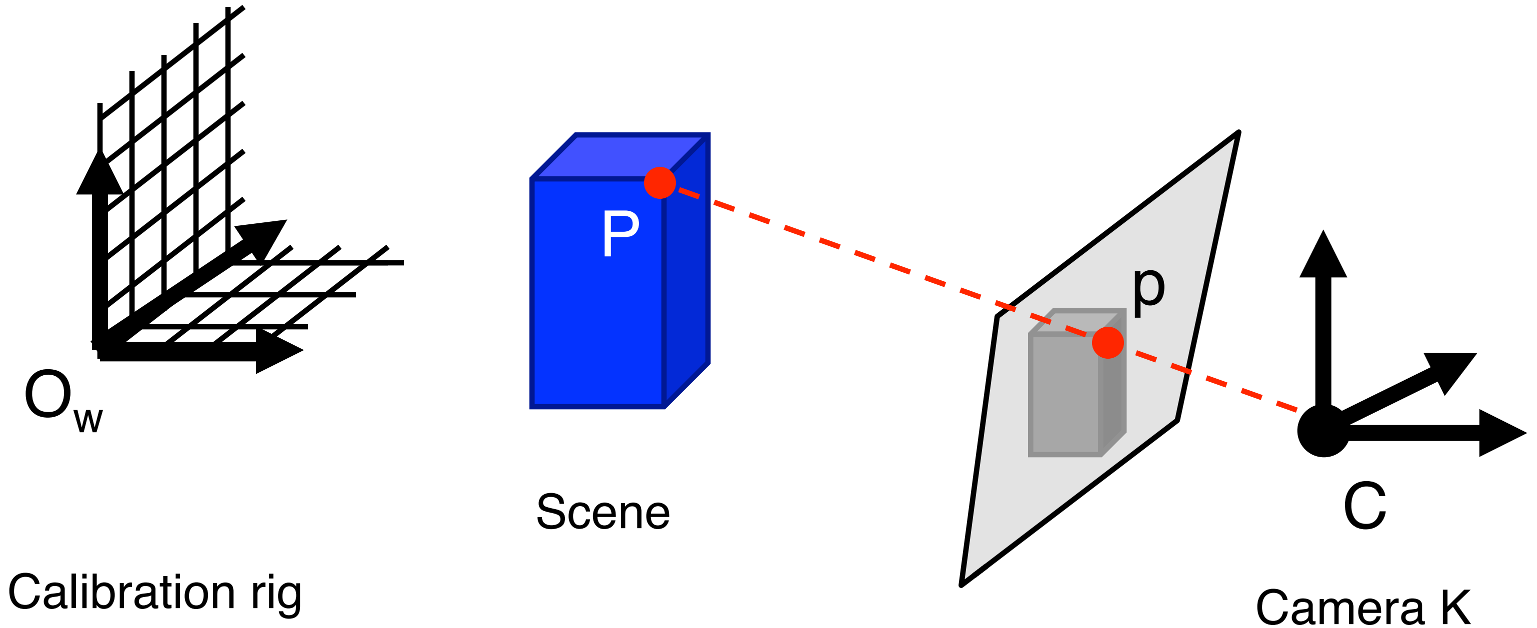
Can Structure Be Recovered from a Single View?



From calibration rig	→ location/pose of the rig, K
From points and lines at infinity + orthogonal lines and planes	→ structure of the scene, K
Knowledge about scene (point correspondences, geometry of lines & planes, etc…	

Slide source: S. Savarese.

Can Structure Be Recovered from a Single View?



Why is it so difficult?

Intrinsic ambiguity of the mapping from 3D to image (2D)

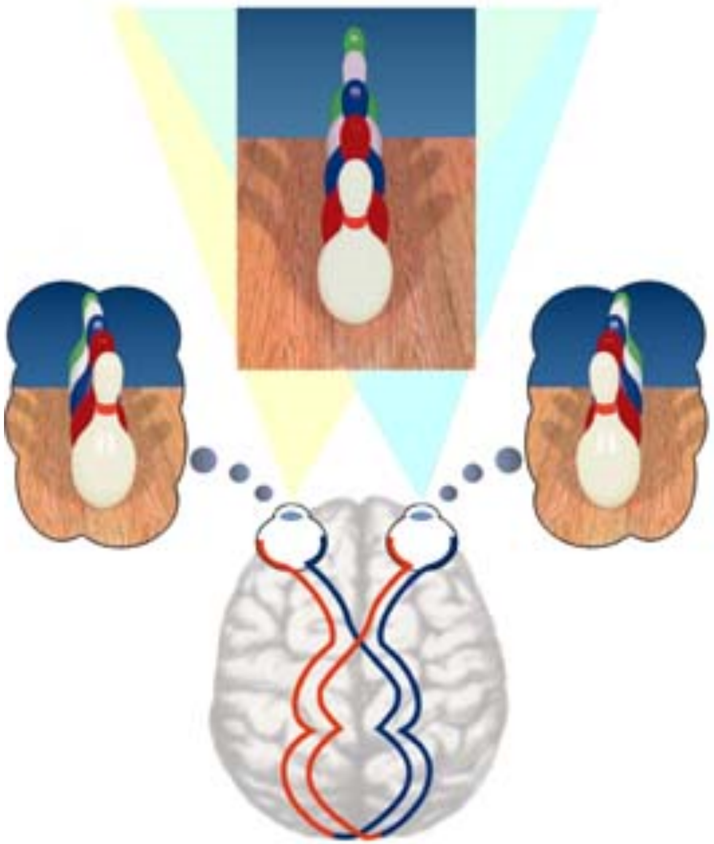
Slide source: S. Savarese.

Can Structure Be Recovered from a Single View?

Intrinsic ambiguity of the mapping from 3D to image (2D)

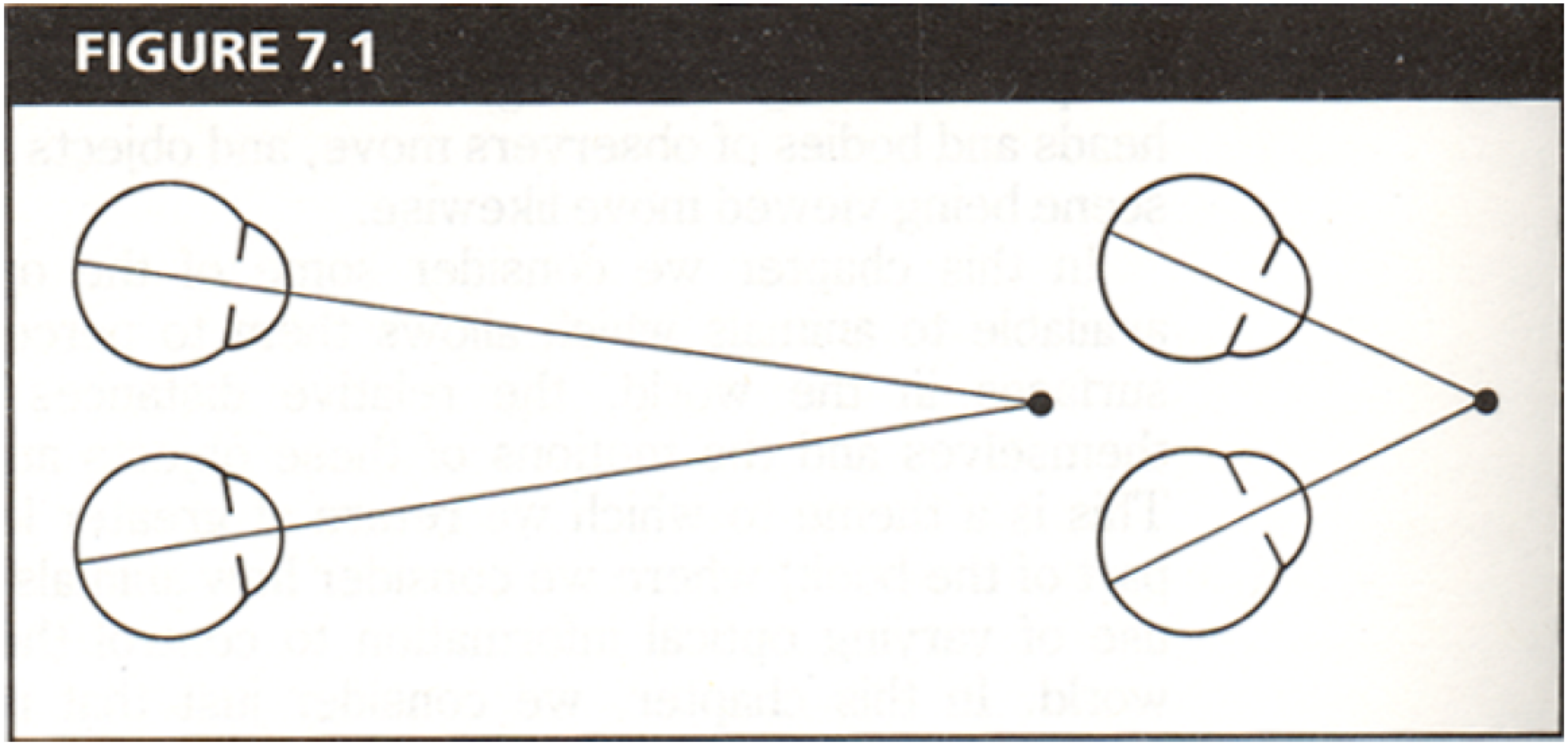
Courtesy slide S. Lazebnik

Two Eyes Help!



Slide source: S. Savarese.

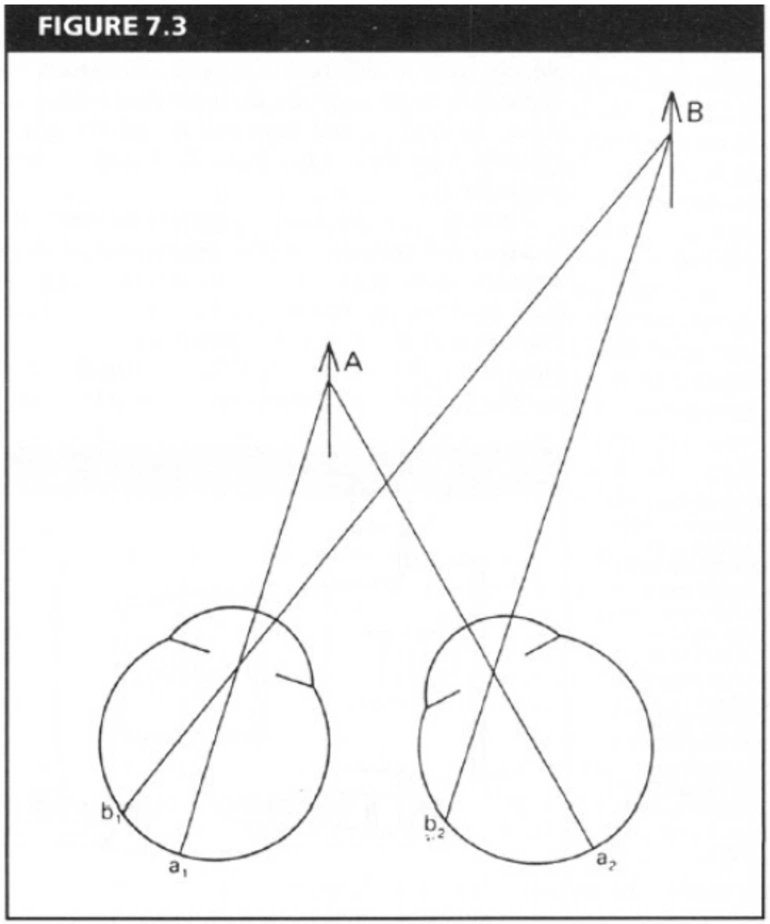
Fixation, convergence



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

Slide source: K. Grauman.

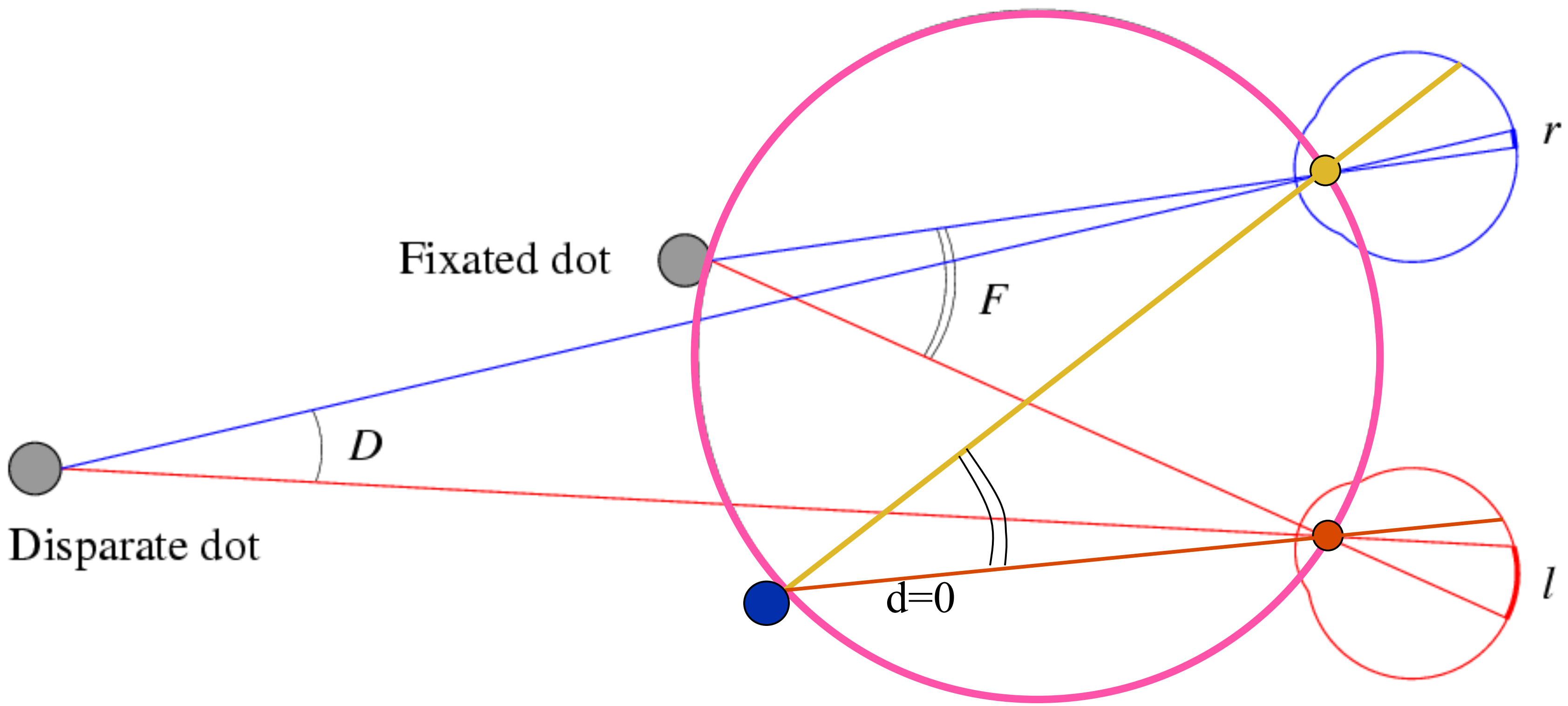
Human stereopsis: disparity



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

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

Human Stereopsis; Disparity



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

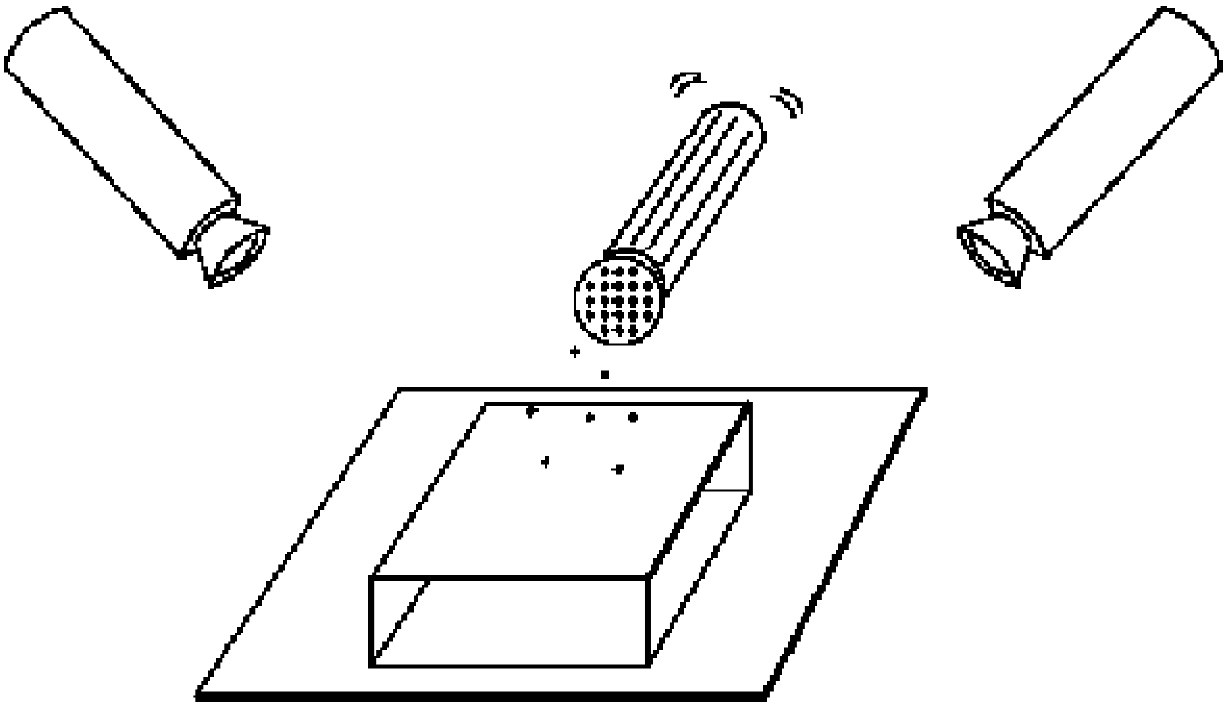
Adapted from M. Pollefeys

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

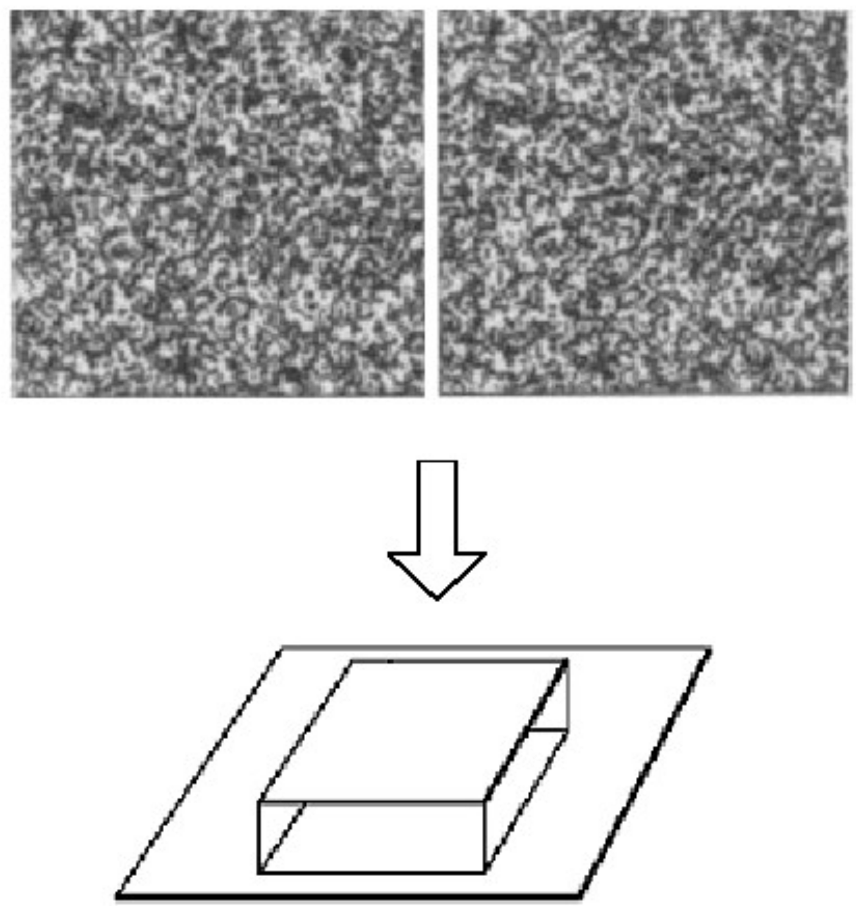
Slide source: K. Grauman.

Random dot stereograms



Forsyth & Ponce

Random dot stereograms



Slide source: K. Grauman.

Random dot stereograms

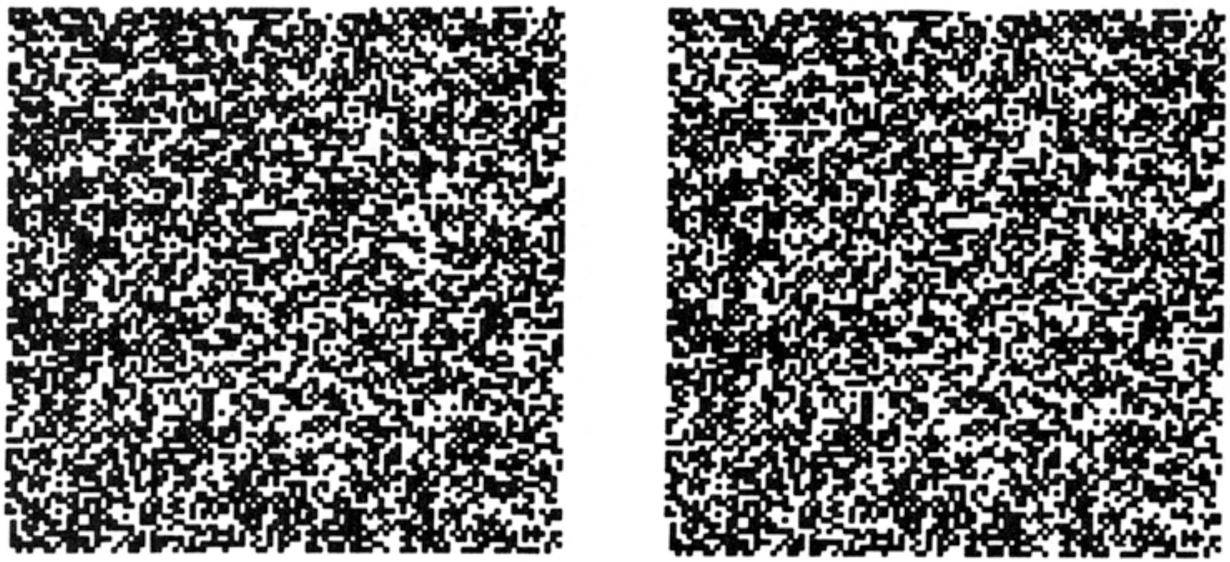


Figure 5.3.8 A random dot stereogram. These two images are derived from a single array of randomly placed squares by laterally displacing a region of them as described in the text. When they are viewed with crossed disparity (by crossing the eyes) so that the right eye's view of the left image is combined with the left eye's view of the right image, a square will be perceived to float above the page. (See pages 210–211 for instructions on fusing stereograms.)

From Palmer, "Vision Science", MIT Press

Slide source: K. Grauman.

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

Slide source: K. Grauman.

Autostereograms

Exploit disparity as depth cue using single image

(Single image random dot stereogram, Single image stereogram)

Images from magiceye.com

Slide source: K. Grauman.

Autostereograms

Images from magiceye.com

Slide source: K. Grauman.

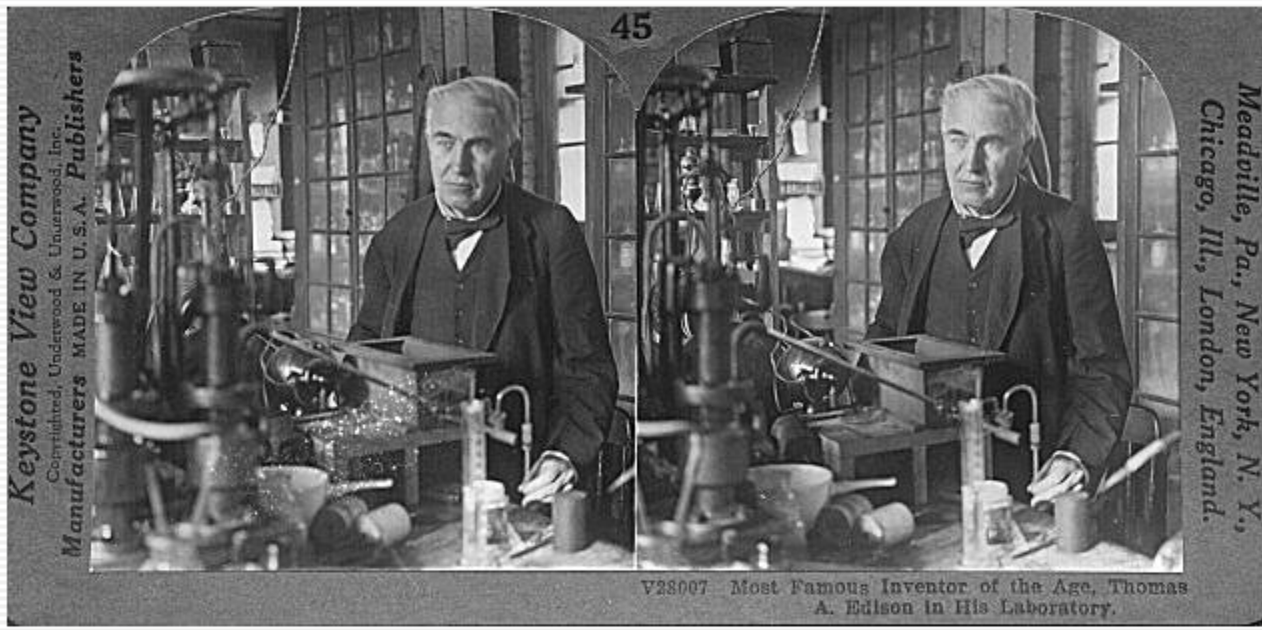
Stereo photography and stereo viewers

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

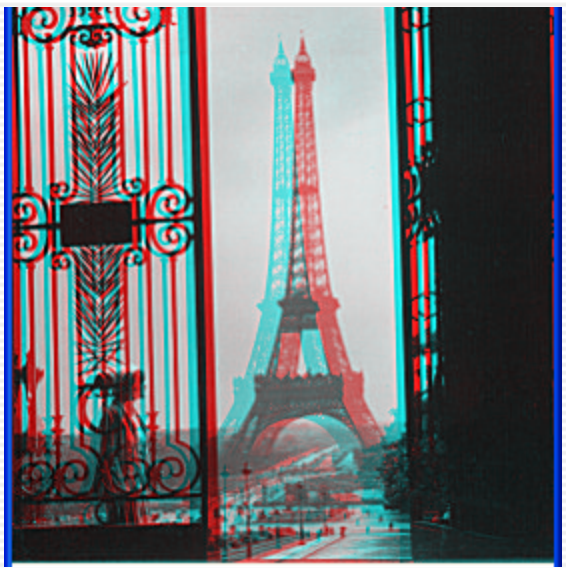
Image courtesy of fisher-price.com

Slide source: K. Grauman.



http://www.johnsonshawmuseum.org

Slide source: K. Grauman.



© Copyright 2001 Johnson-Shaw Stereoscopic Museum

http://www.johnsonshawmuseum.org

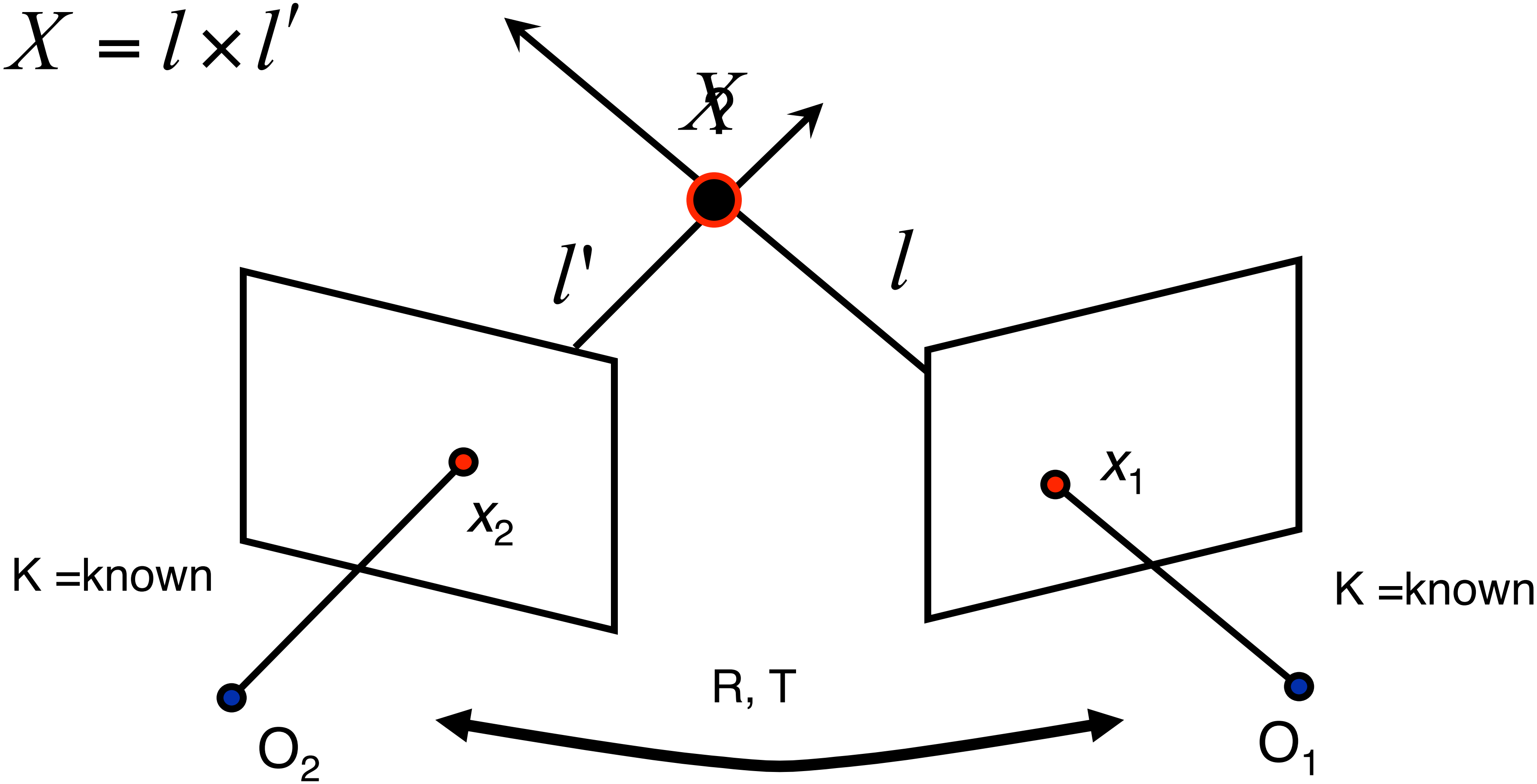
Slide source: K. Grauman.

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

Slide source: K. Grauman.

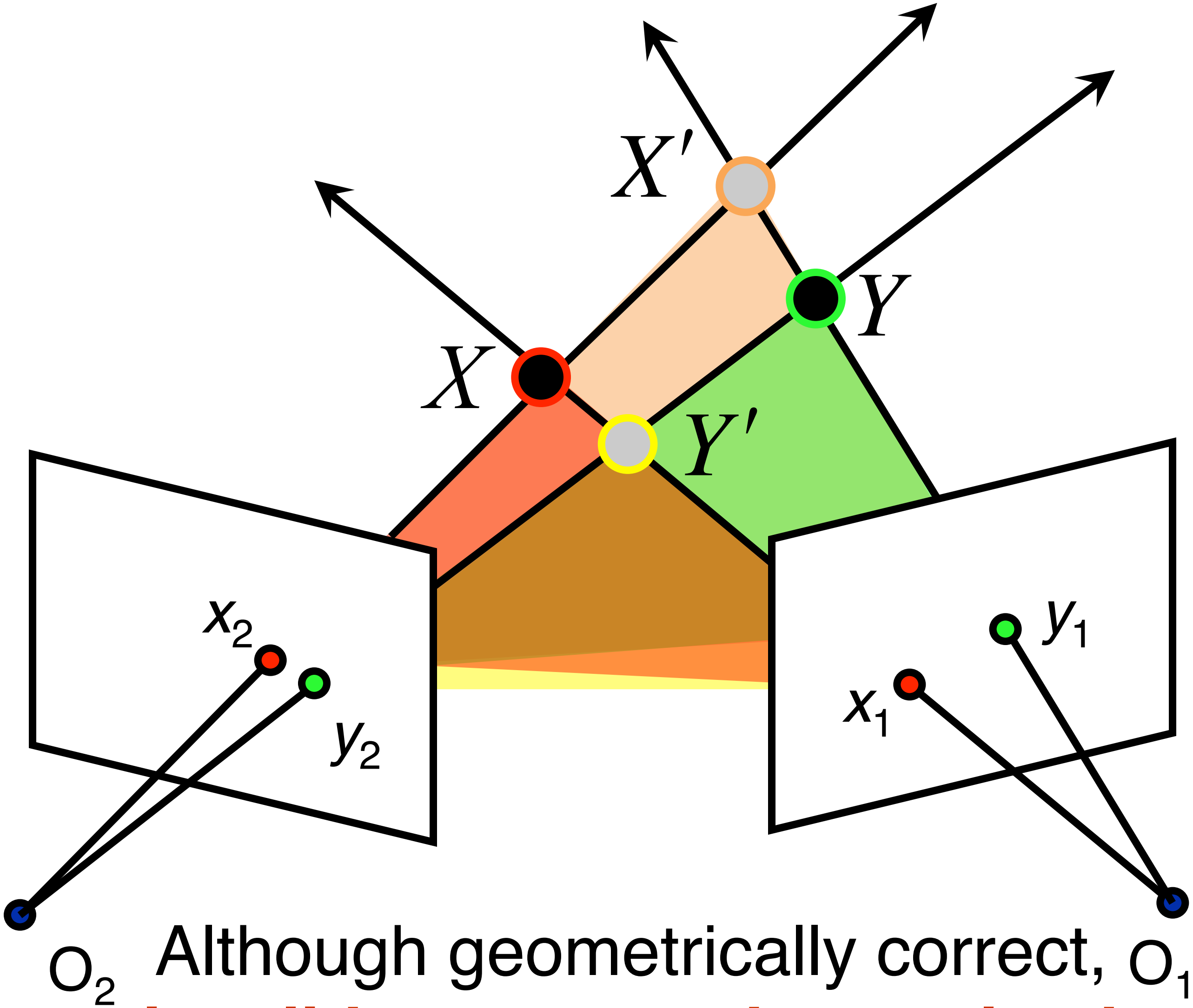
Two Eyes Help!

$$X = l \times l'$$



This is called **triangulation**

Triangulation



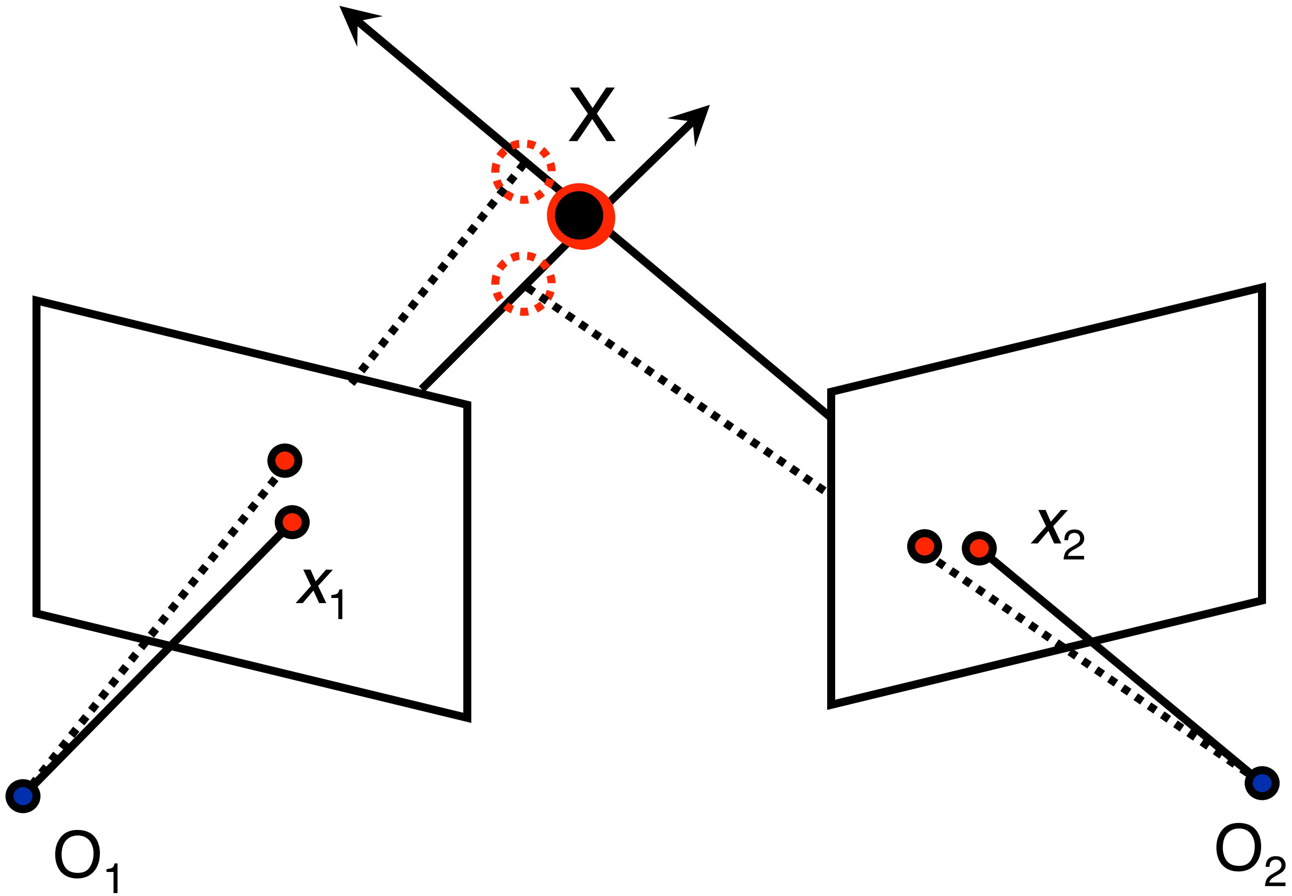
Although geometrically correct,

invalid correspondences lead to incorrect triangulation

Triangulation

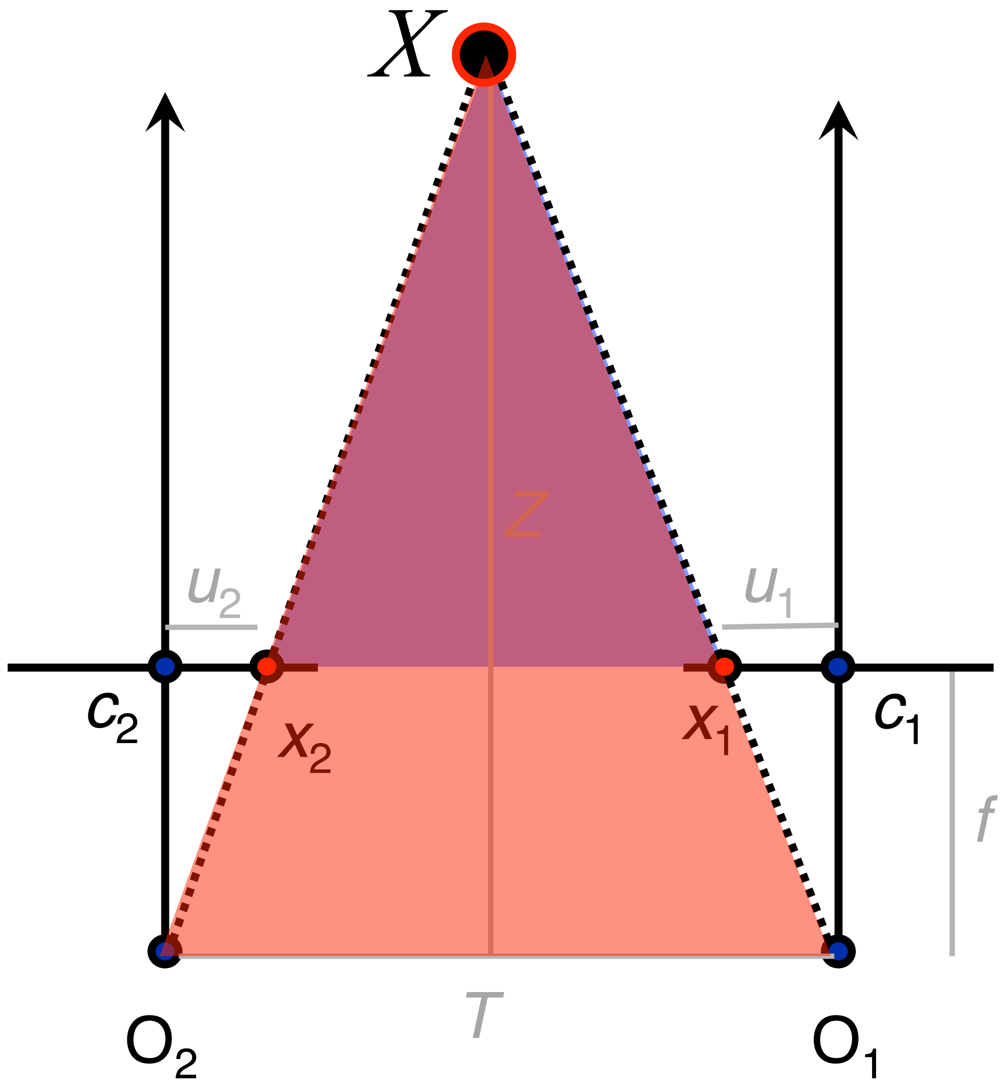
- Find X that minimizes

$$d^2(x_1, M_1X) + d^2(x_2, M_2X)$$



Slide source: S. Savarese.

Triangulation Geometry in Simple Stereo



$$\frac{T + u_2 - u_1}{Z - f} = \frac{T}{Z}$$

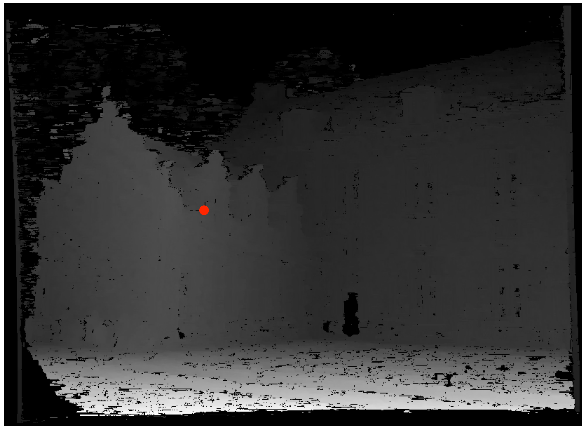
$$Z = f\frac{T}{u_1 - u_2}$$

Depth from disparity

image I(x,y)

Disparity map D(x,y)

image I´(x´,y´)



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

Slide source: K. Grauman.

Core Problems in Stereo

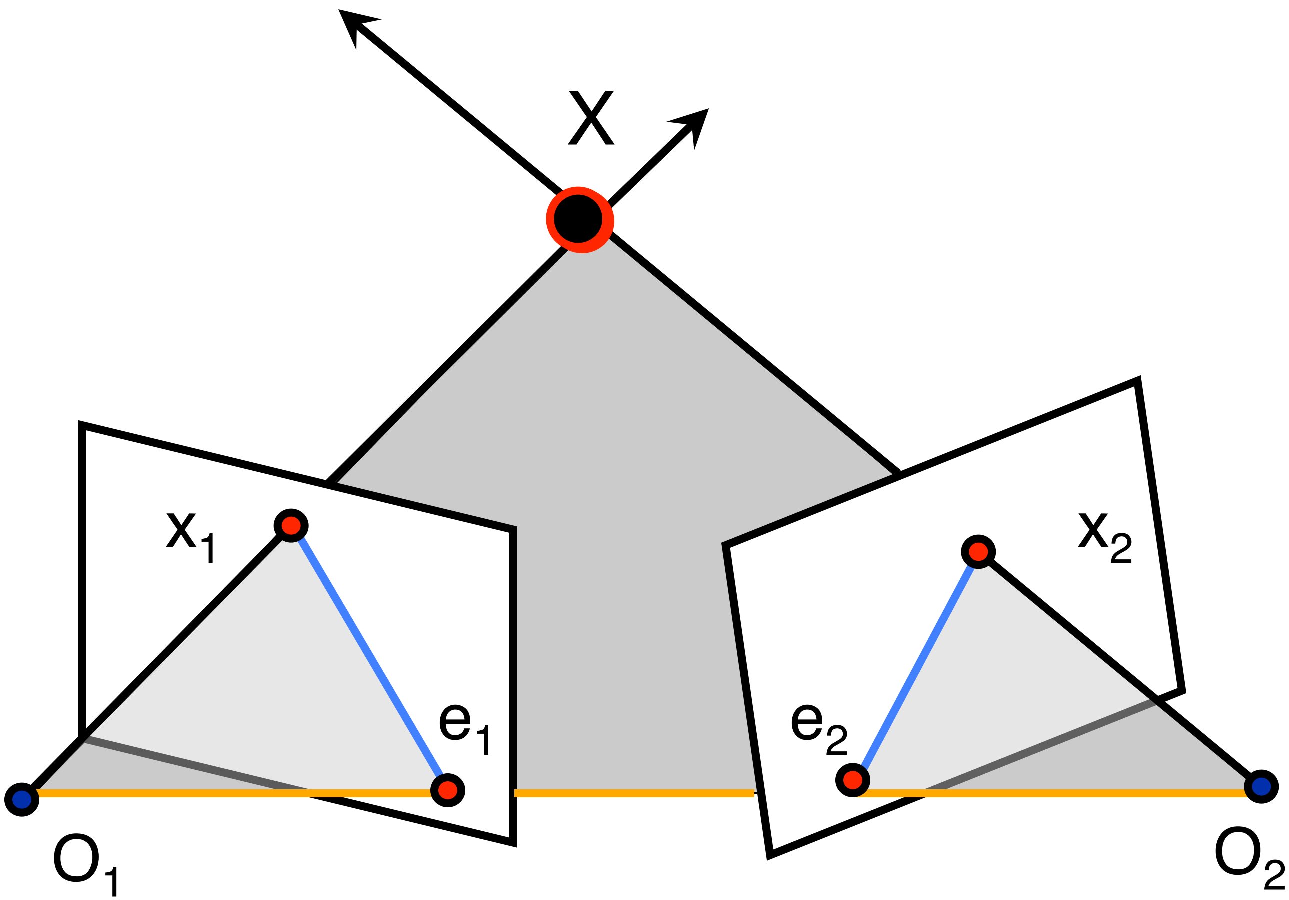
- **Correspondence:** Given a point in one image, how can I find the corresponding point x' in another one ?

- **Camera geometry:** Given corresponding points in two images, find camera matrices, position and pose.

- **Scene geometry:** Find coordinates of 3D point from its projection into 2 or multiple images.

Slide source: S. Savarese.

Epipolar Geometry



- Epipolar Plane
- Baseline
- Epipolar Lines

- Epipoles e_1, e_2

  = intersections of baseline with image planes
  = projections of the other camera center
  = vanishing points of camera motion direction

Slide source: S. Savarese.

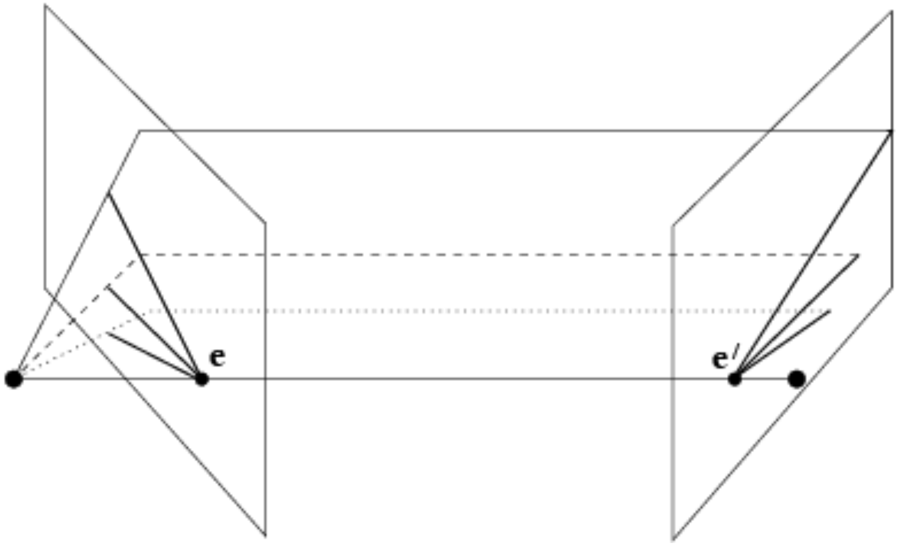
Epipolar Geometry Terms

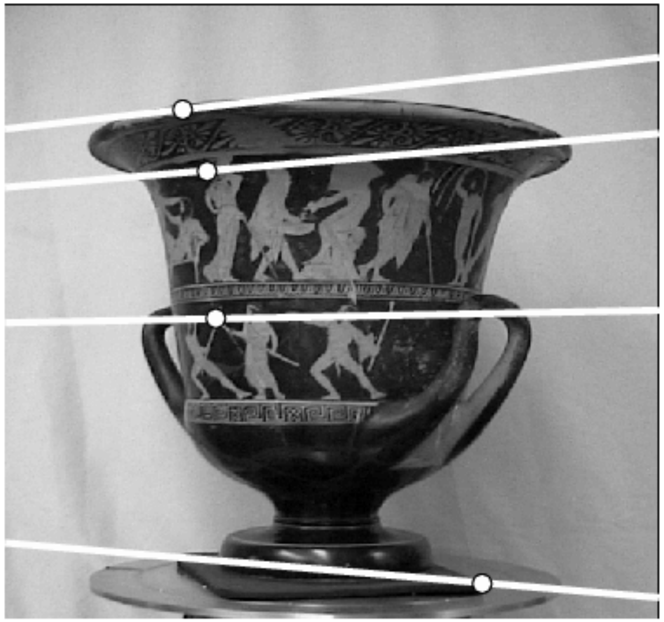
- **Baseline**: line joining the camera centers
- **Epipole**: point of intersection of baseline with the image plane
- **Epipolar plane**: plane containing baseline and world point
- **Epipolar line**: intersection of epipolar plane with the image plane

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

Slide source: K. Grauman.

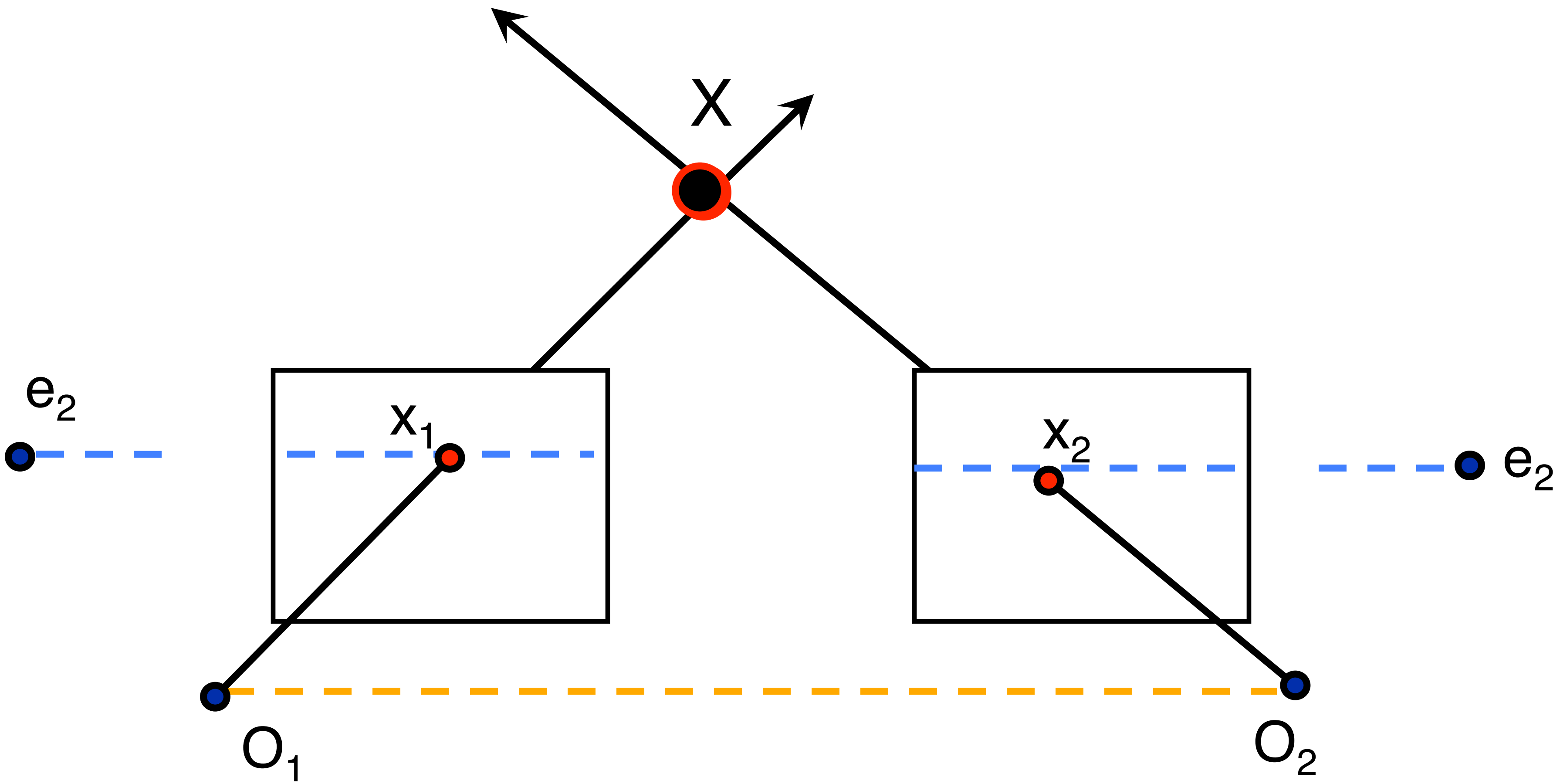
Example: Converging image planes





Slide source: S. Savarese, K. Grauman; Figure from Hartley and Zisserman.

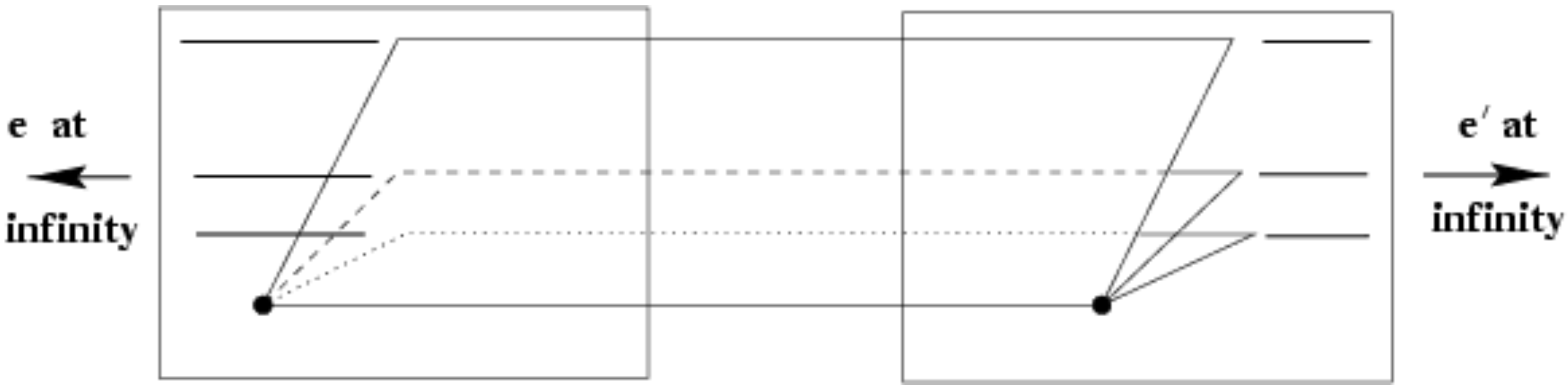
Example: Parallel Image Planes



- Baseline intersects the image plane at infinity
- Epipoles are at infinity
- Epipolar lines are parallel to x axis

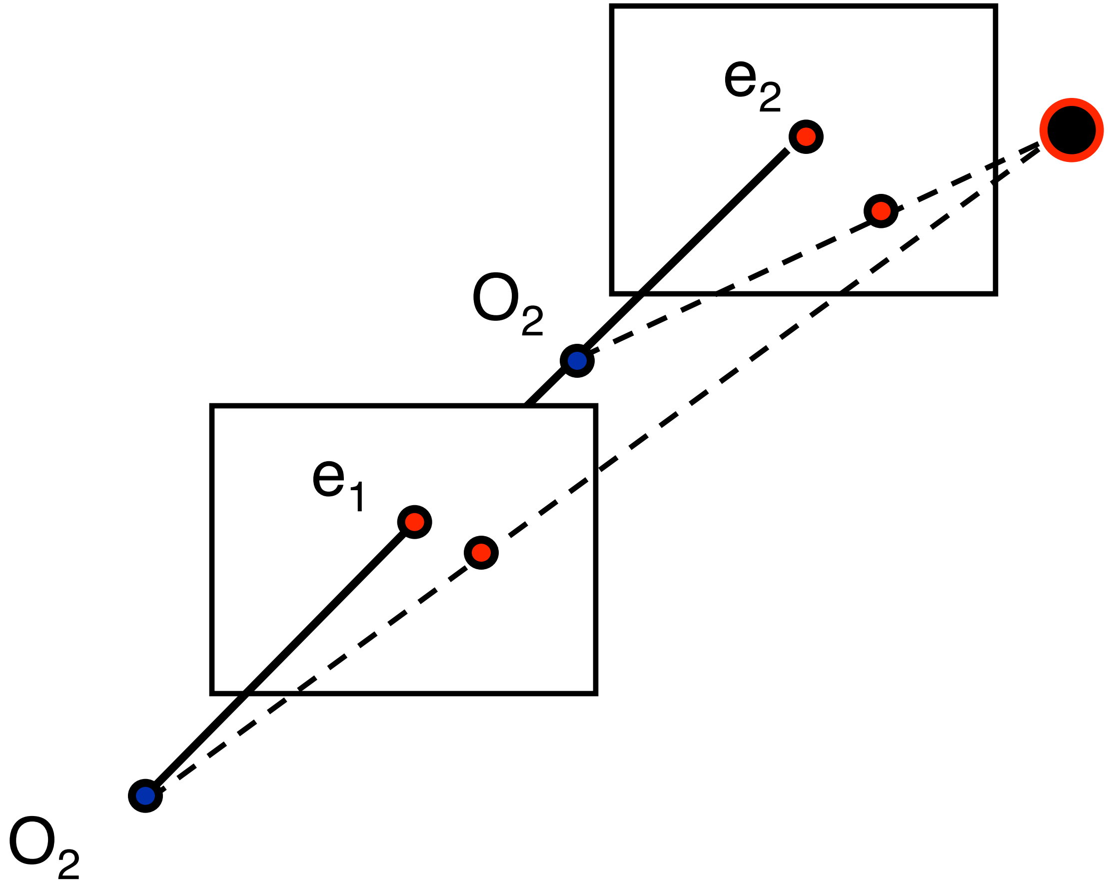
Slide source: S. Savarese.

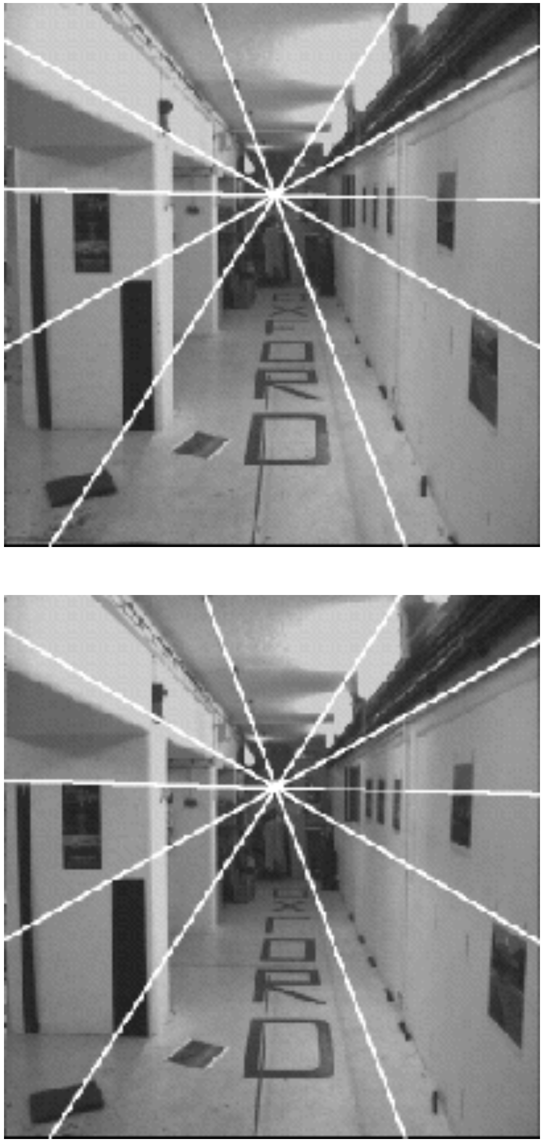
Example: Parallel Image Planes



Slide source: S. Savarese, K Grauman; Figure from Hartley and Zisserman.

Example: Forward translation

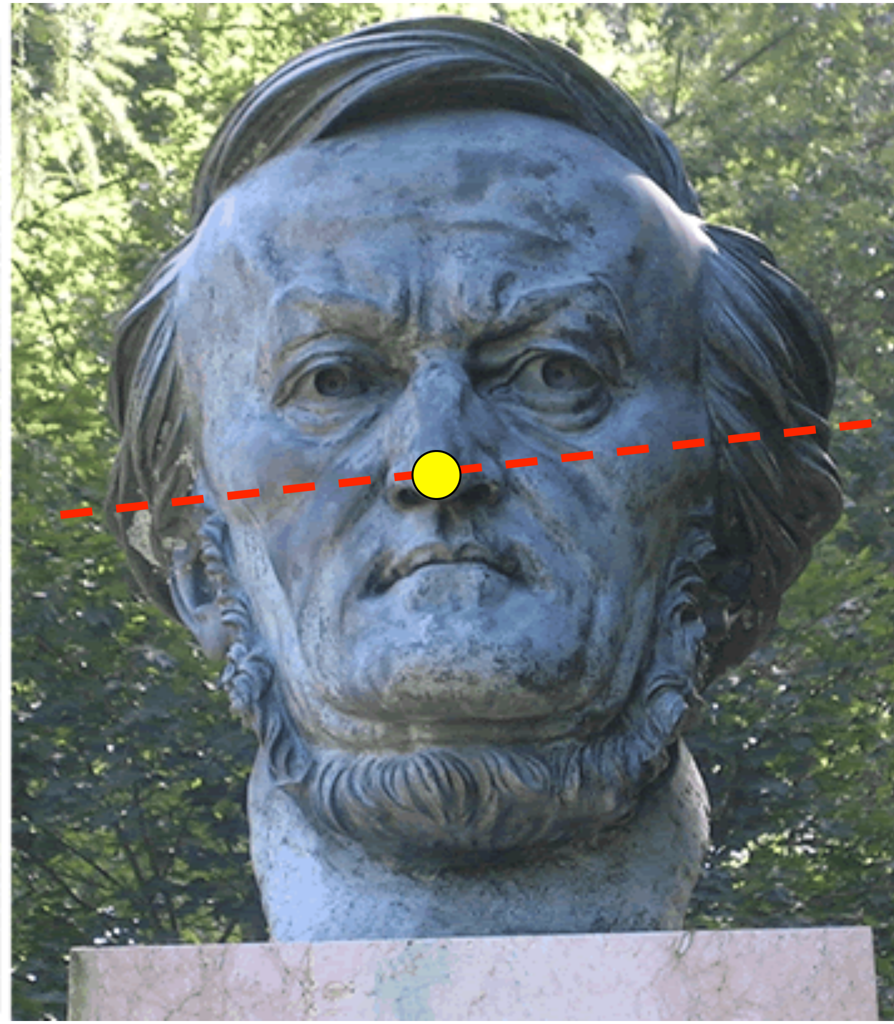




- The epipoles have same position in both images
- Epipole called FOE (focus of expansion)

Slide source: S. Savarese, K Grauman; Figure from Hartley and Zisserman.

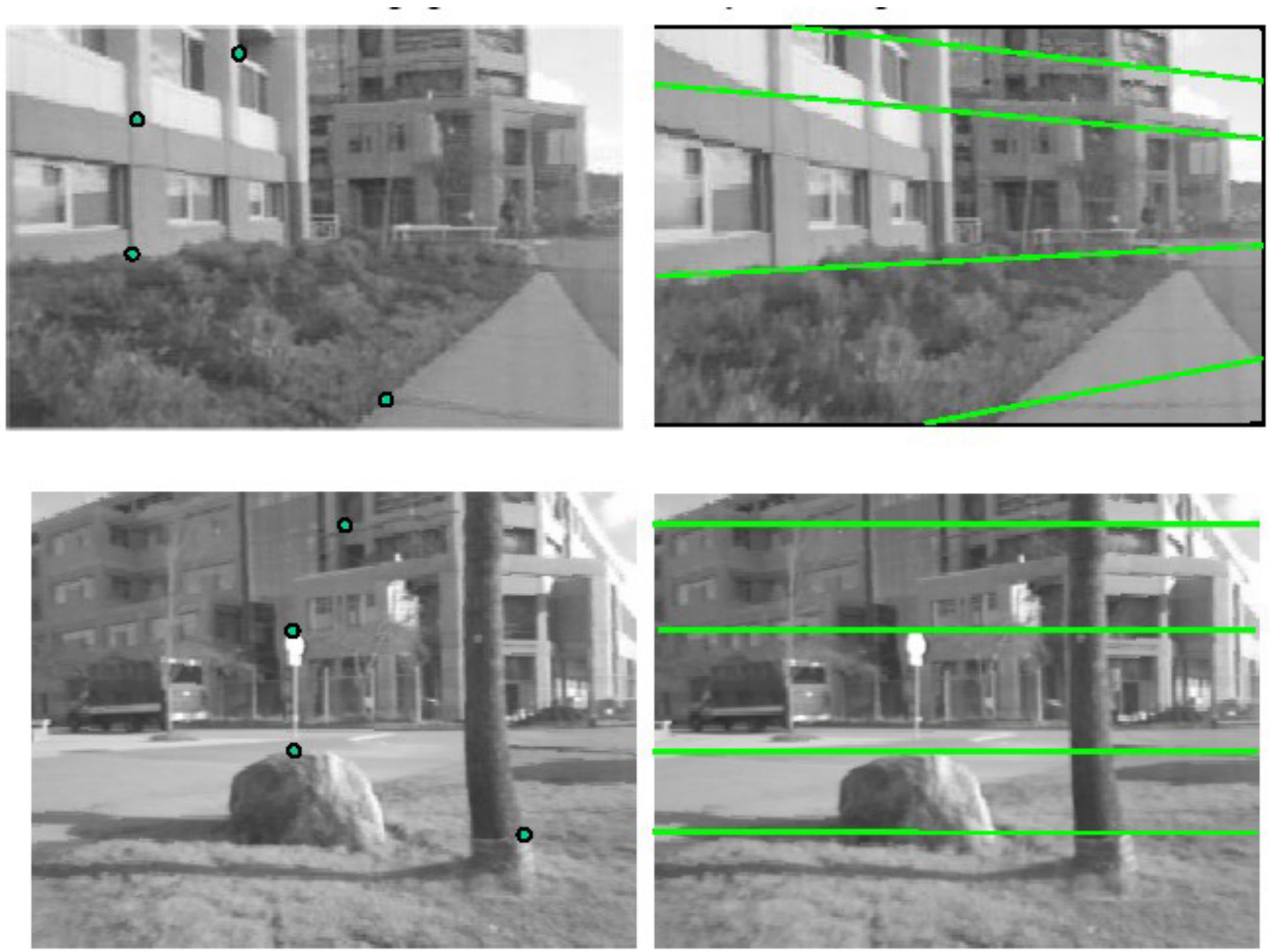
The Epipolar Constraint



- Two views of the same object
- Suppose I know the camera positions and camera matrices
- Given a point on left image, how can I find the corresponding point on right image?

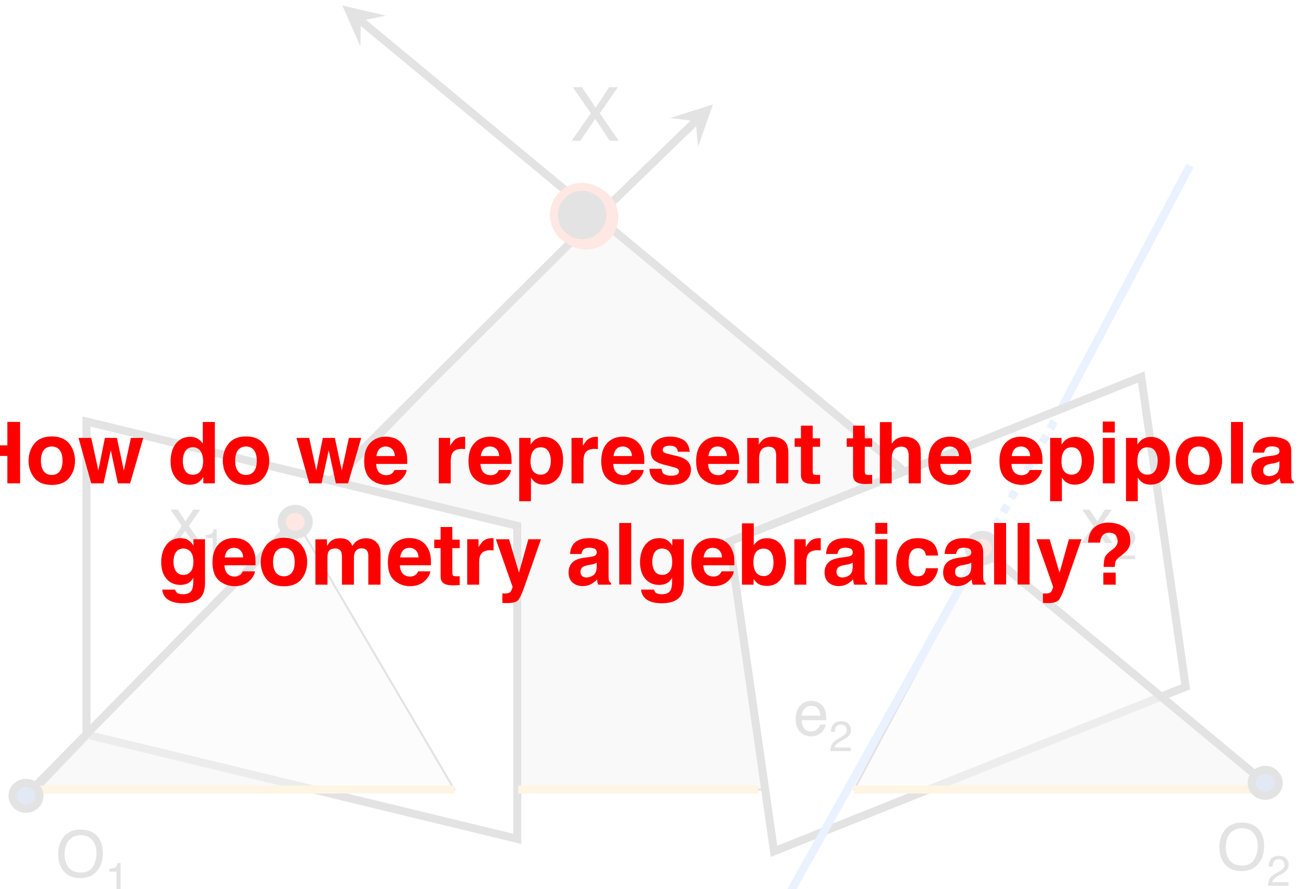
Slide source: S. Savarese.

Image Examples of the Epipolar Constraint



Slide source: K. Grauman.

Epipolar Geometry



How do we represent the epipolar geometry algebraically?

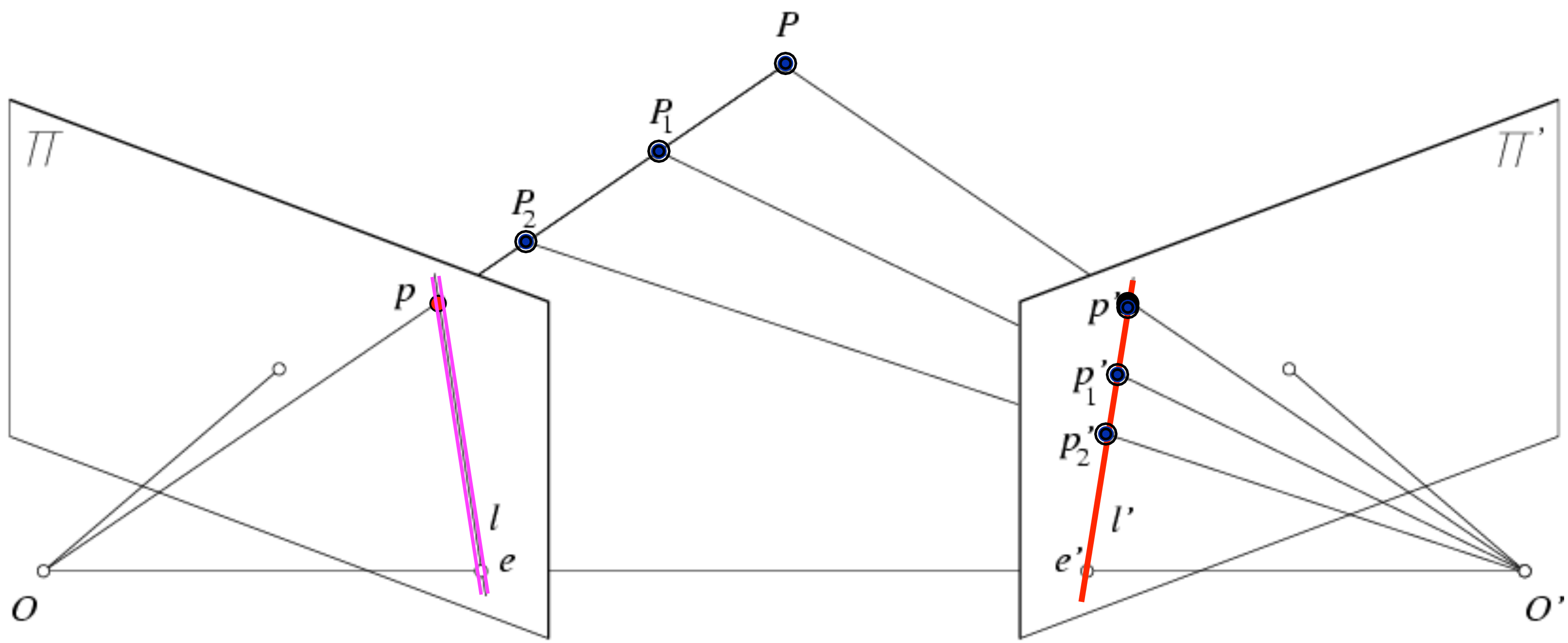
- Epipolar Plane
- Baseline
- Epipolar Lines

- Epipoles e_1, e_2

  = intersections of baseline with image planes
  = projections of the other camera center
  = vanishing points of camera motion direction

Slide source: S. Savarese.

Epipolar Constraint

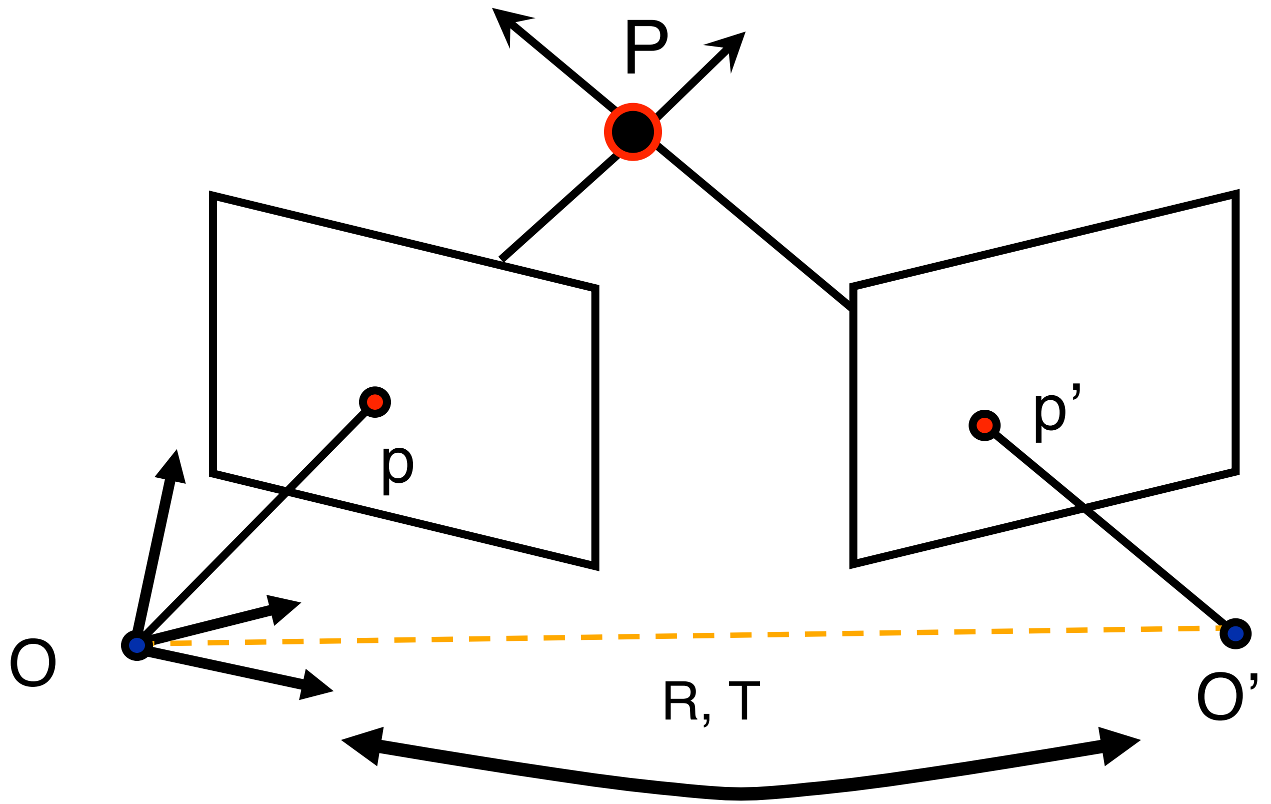


- Potential matches for *p* have to lie on the corresponding epipolar line *l'*.

- Potential matches for *p'* have to lie on the corresponding epipolar line *l*.

Slide source: S. Savarese.

The Epipolar Constraint

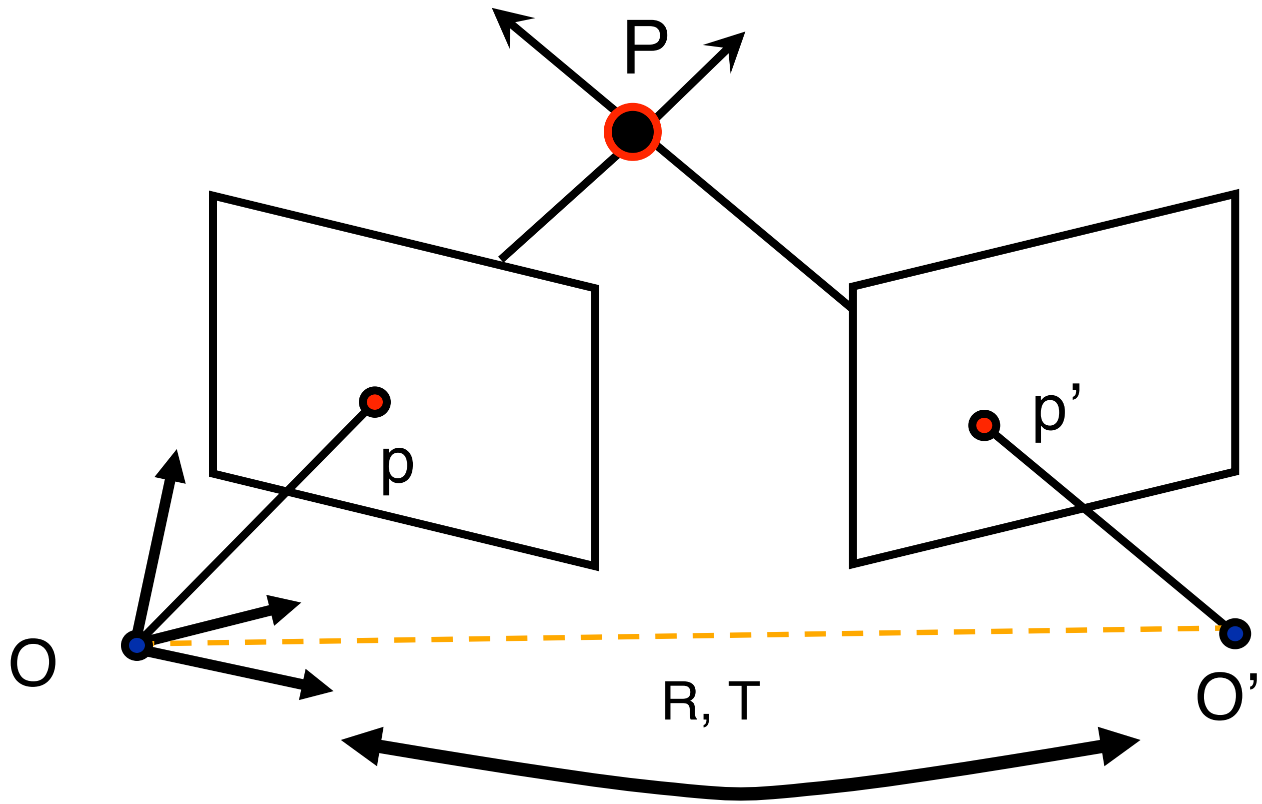


$$M = K\begin{bmatrix} I & 0 \end{bmatrix} \qquad M' = K\begin{bmatrix} R & T \end{bmatrix}$$

$$P \rightarrow M\,P = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \qquad P \rightarrow M'\,P = \begin{bmatrix} u' \\ v' \\ 1 \end{bmatrix}$$

Slide source: S. Savarese.

The Epipolar Constraint



$$\mathrm{M} = \mathrm{K}\begin{bmatrix} \mathrm{I} & 0 \end{bmatrix}$$

$$M = \begin{bmatrix} I & 0 \end{bmatrix}$$

K is known
(canonical cameras)

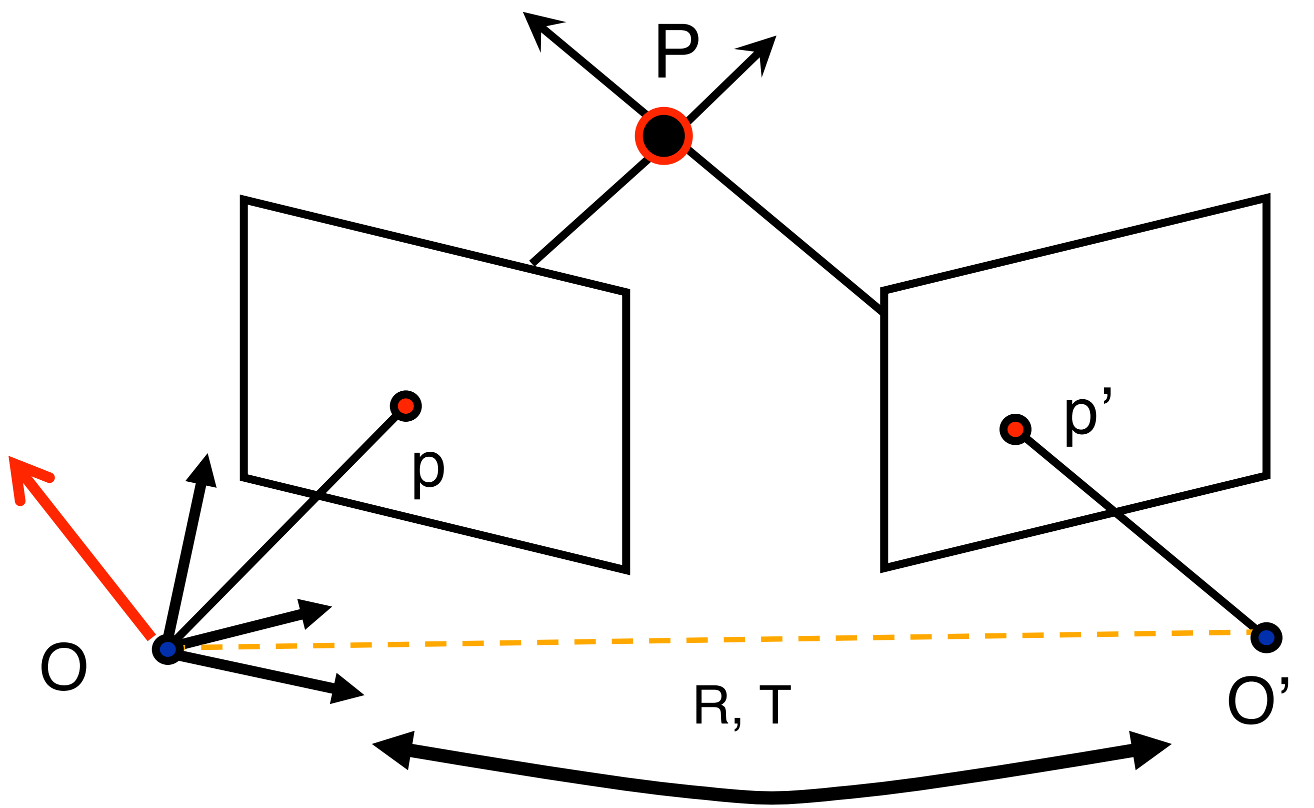
$$K = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\mathrm{M}' = \mathrm{K}\begin{bmatrix} \mathrm{R} & \mathrm{T} \end{bmatrix}$$

$$M' = \begin{bmatrix} R & T \end{bmatrix}$$

Slide source: S. Savarese.

The Epipolar Constraint



p' in first camera reference system is $= R\,p'$

$T \times (R\,p')$ is perpendicular to epipolar plane

$$\rightarrow p^T \cdot \left[T \times (R\,p') \right] = 0$$

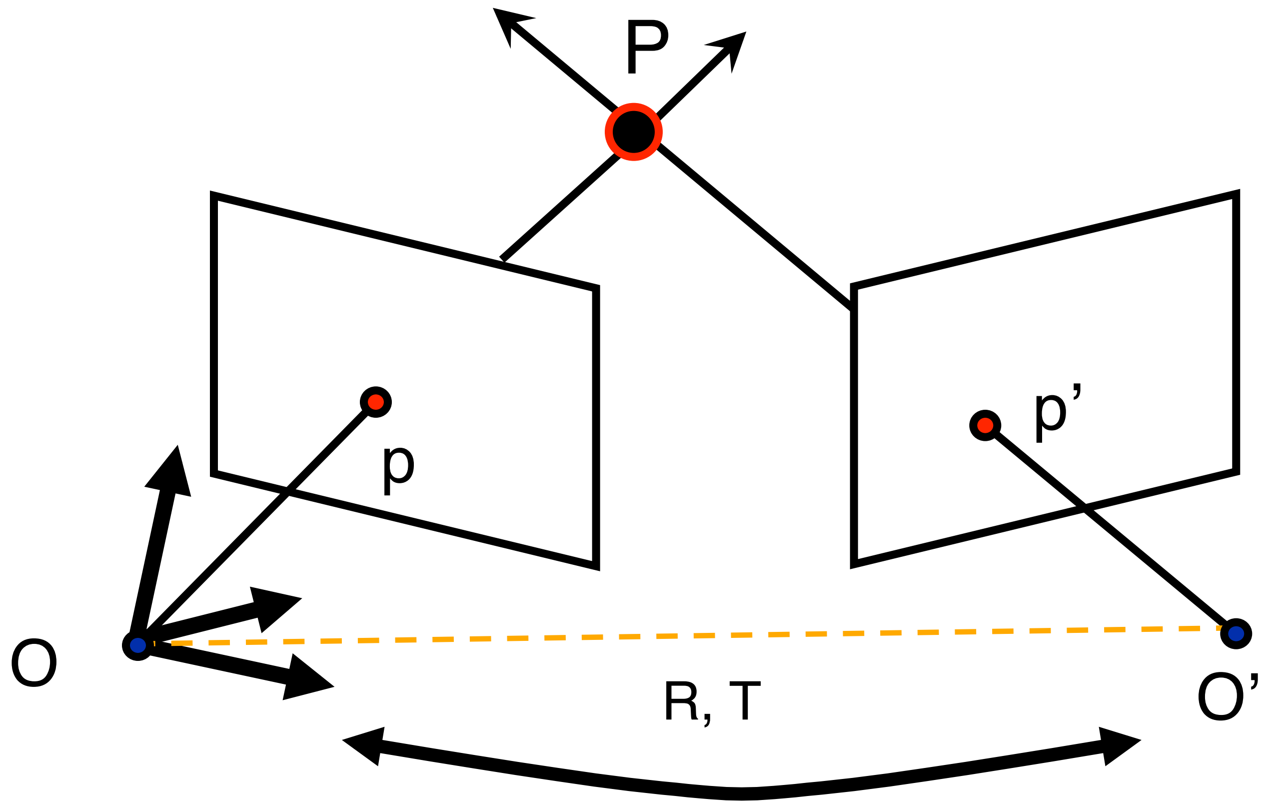
Slide source: S. Savarese.

Cross product as matrix multiplication

$$\mathbf{a} \times \mathbf{b} = \begin{bmatrix} 0 & -a_z & a_y \\ a_z & 0 & -a_x \\ -a_y & a_x & 0 \end{bmatrix} \begin{bmatrix} b_x \\ b_y \\ b_z \end{bmatrix} = [\mathbf{a}_\times]\mathbf{b}$$

Slide source: S. Savarese.

The Epipolar Constraint



$$p^T \cdot [T \times (R\,p')] = 0 \rightarrow p^T \cdot [T_\times] \cdot R\,p' = 0$$

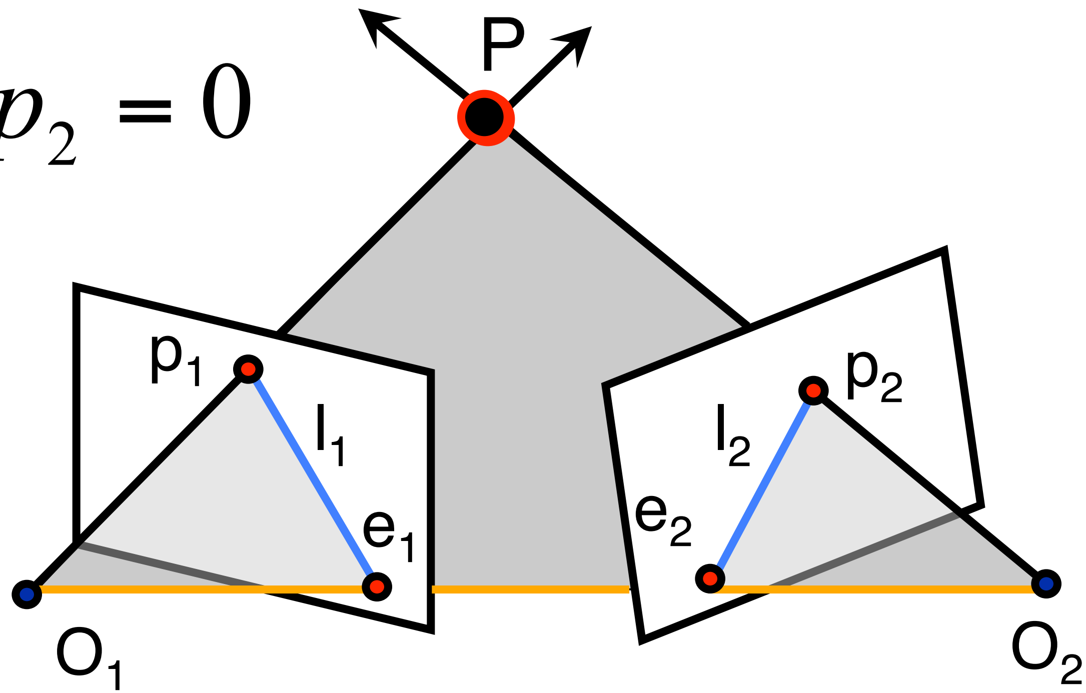
E = essential matrix

(Longuet-Higgins, 1981)

Slide source: S. Savarese.

The Epipolar Constraint

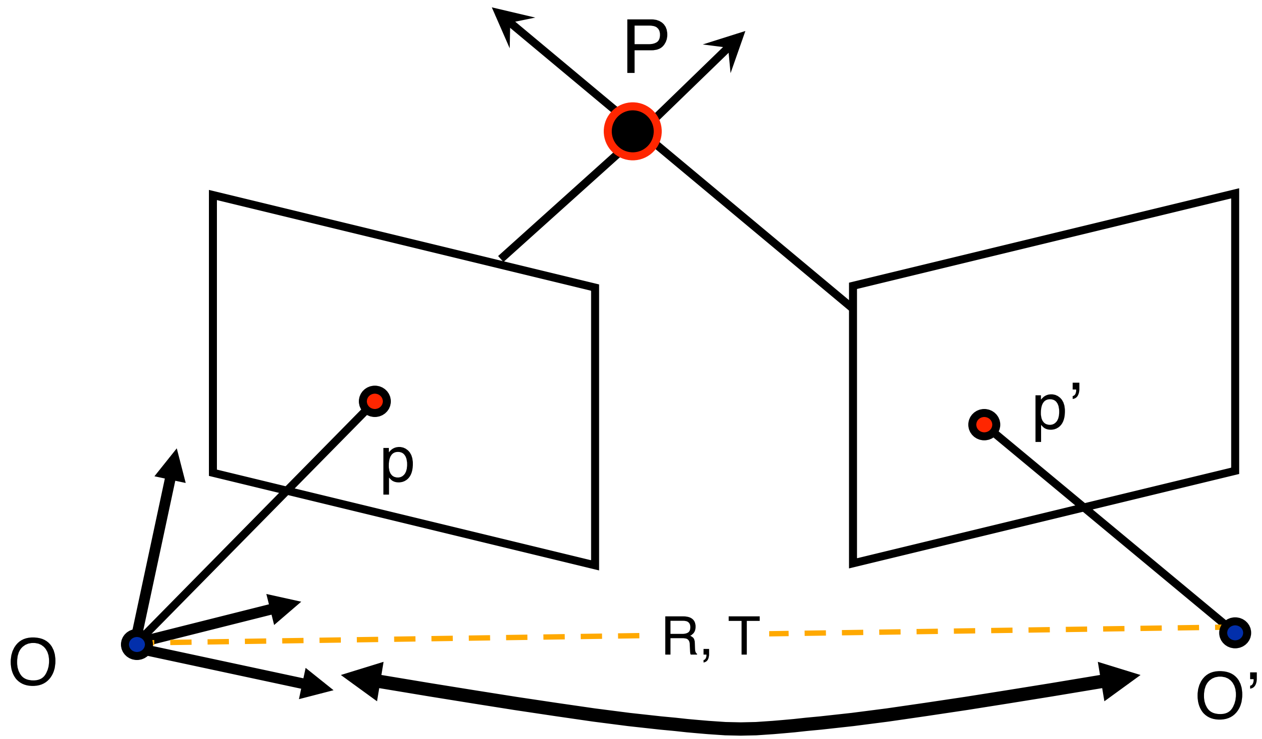
$$p_1^T \cdot E\, p_2 = 0$$



- $E\, p_2$ is the epipolar line associated with p_2 ($l_1 = E\, p_2$)
- $E^T p_1$ is the epipolar line associated with p_1 ($l_2 = E^T p_1$)
- $E\, e_2 = 0$ and $E^T e_1 = 0$
- E is 3x3 matrix; 5 DOF
- E is singular (rank two)

Slide source: S. Savarese.

The Epipolar Constraint



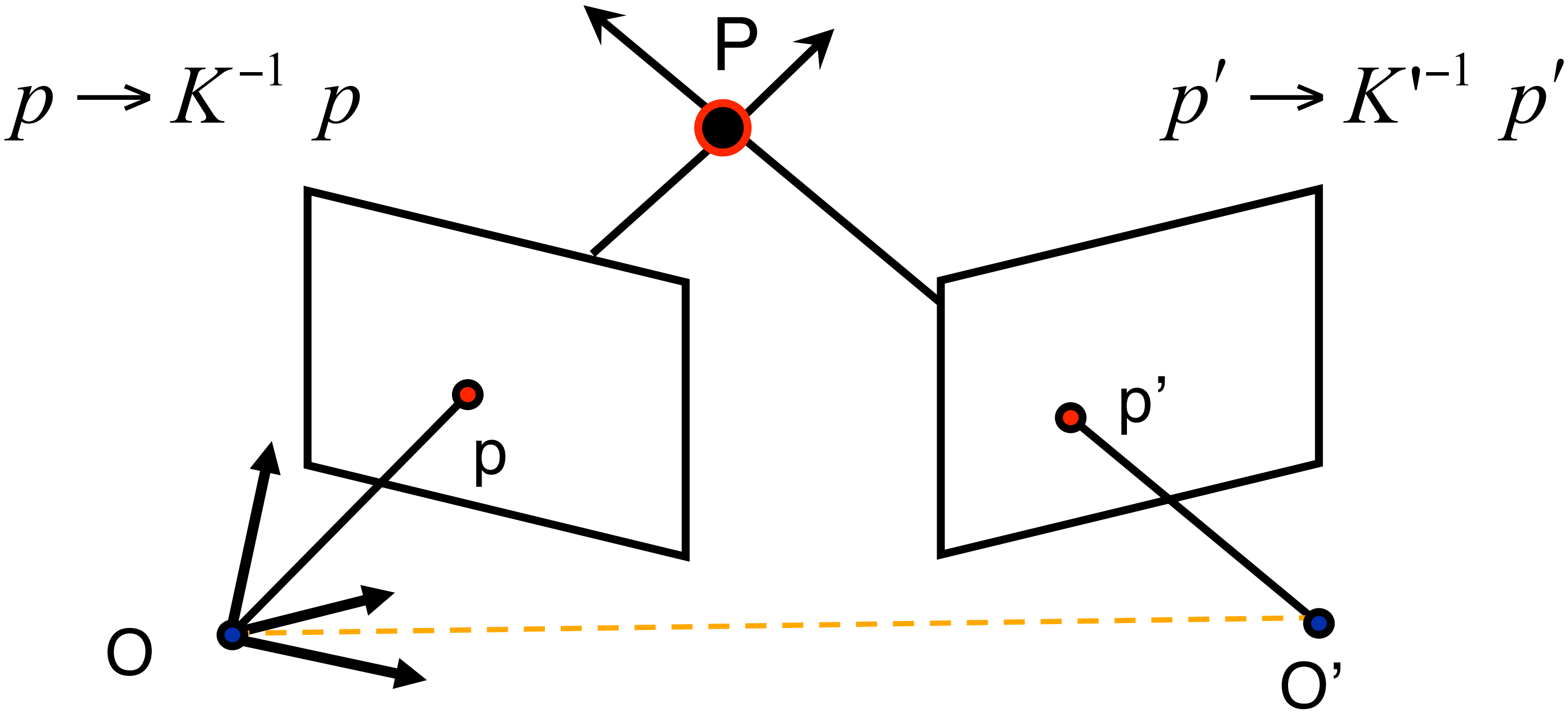
$$M = K\begin{bmatrix} I & 0 \end{bmatrix}$$

K is unknown

$$M' = K\begin{bmatrix} R & T \end{bmatrix}$$

Slide source: S. Savarese.

The Epipolar Constraint

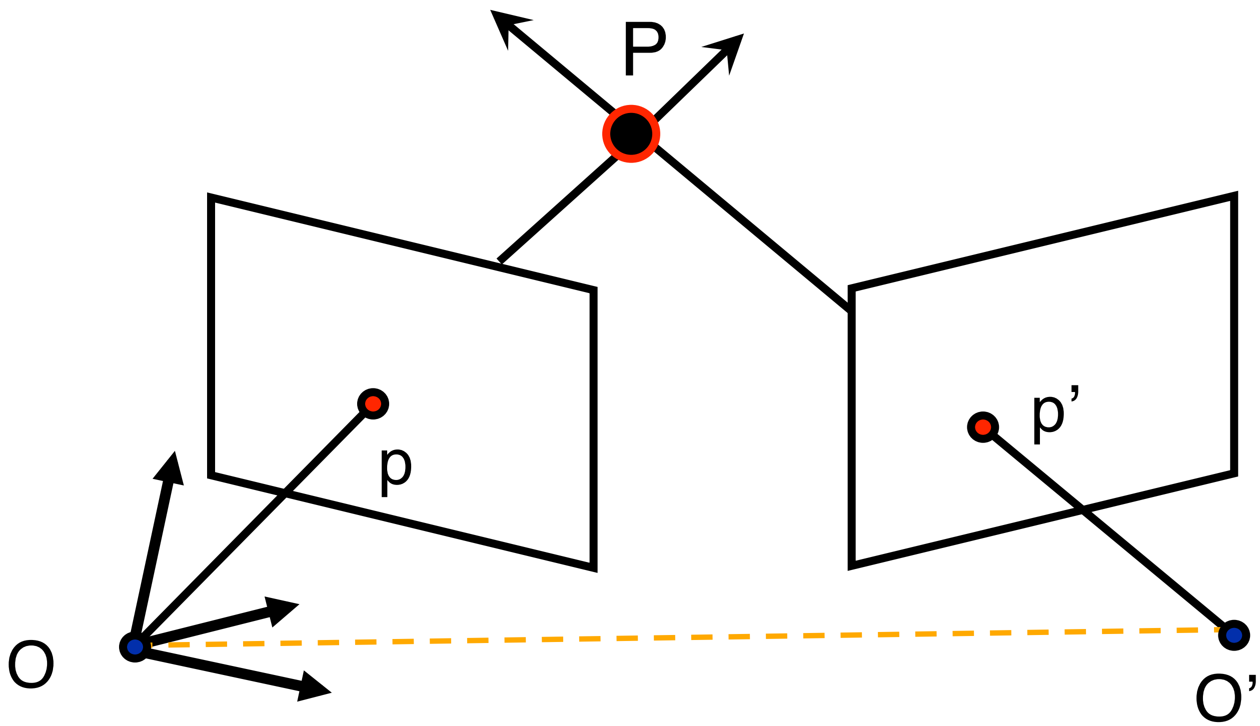


$$p^T \cdot [T_\times] \cdot R\, p' = 0 \rightarrow (K^{-1}\, p)^T \cdot [T_\times] \cdot R\, K'^{-1}\, p' = 0$$

$$p^T\, K^{-T} \cdot [T_\times] \cdot R\, K'^{-1}\, p' = 0 \rightarrow p^T\, F\, p' = 0$$

Slide source: S. Savarese.

The Epipolar Constraint



$$p^T F p' = 0$$

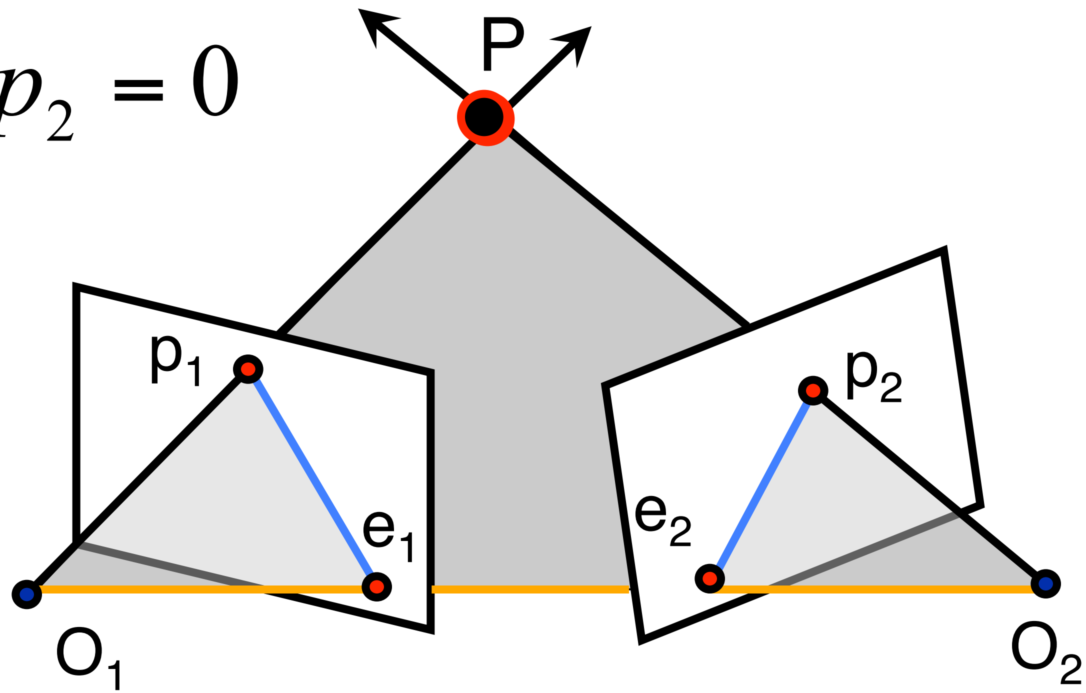
$$F = K^{-T} \cdot [T_\times] \cdot R K'^{-1}$$

F = Fundamental Matrix

(Faugeras and Luong, 1992)

The Epipolar Constraint

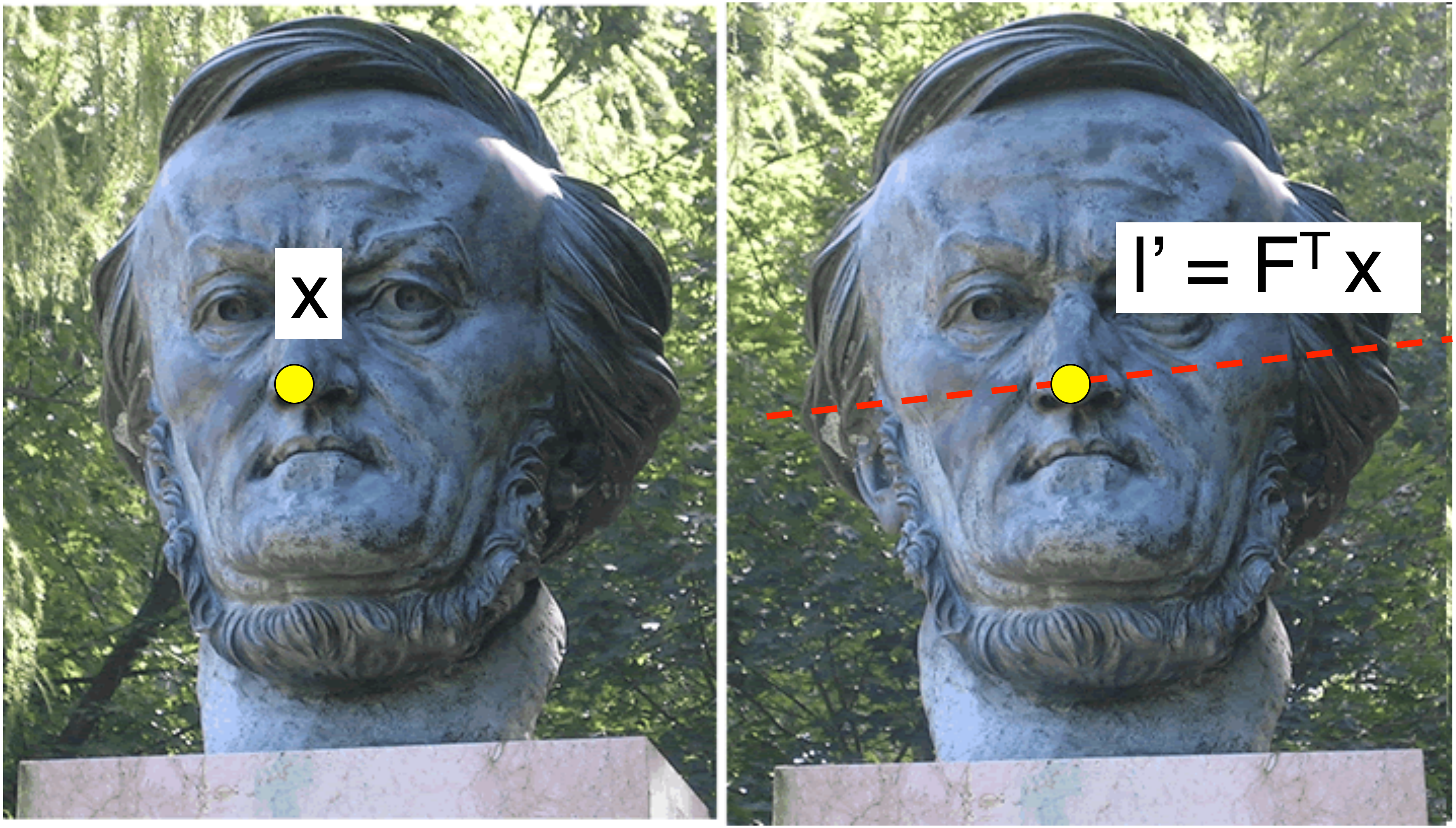
$$p_1^T \cdot F\ p_2 = 0$$



- $F\ p_2$ is the epipolar line associated with p_2 ($l_1 = F\ p_2$)
- $F^T p_1$ is the epipolar line associated with p_1 ($l_2 = F^T p_1$)
- $F\ e_2 = 0$ and $F^T e_1 = 0$
- F is 3x3 matrix; 7 DOF
- F is singular (rank two)

Slide source: S. Savarese.

Why F is useful?



- Suppose F is known
- No additional information about the scene and camera is given
- Given a point on left image, how can I find the corresponding point on right image?

Slide source: S. Savarese.

Why is F Useful?

- F captures information about the epipolar geometry of 2 views + camera parameters
- MORE IMPORTANTLY: F gives constraints on how the scene changes under view point transformation (without reconstructing the scene!)
- Powerful tool in:
    - 3D reconstruction
    - Multi-view object/scene matching

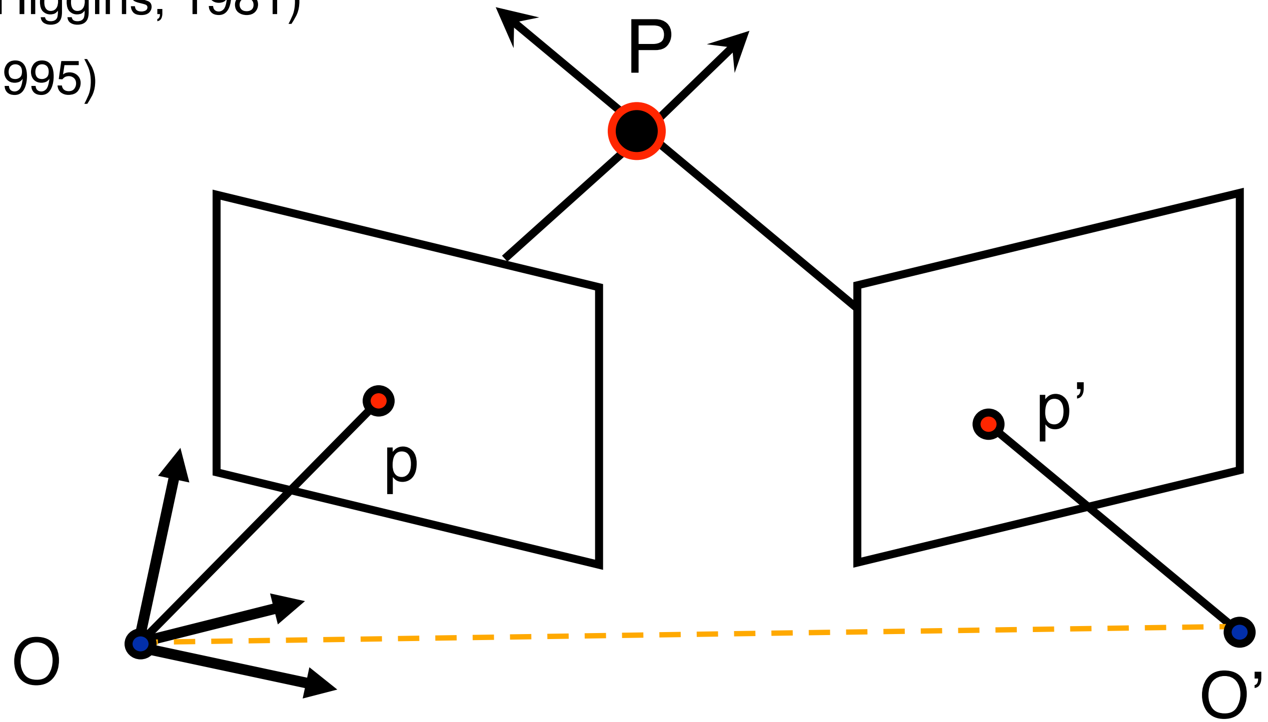
Slide source: S. Savarese.

Estimating F

The Eight-Point Algorithm

(Longuet-Higgins, 1981)

(Hartley, 1995)



$$P \rightarrow p = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \qquad P \rightarrow p' = \begin{bmatrix} u' \\ v' \\ 1 \end{bmatrix} \qquad p^{T}\, F\, p' = 0$$

Slide source: S. Savarese.

Estimating F

$$\mathrm{p}^{\mathrm{T}}\,\mathrm{F}\,\mathrm{p}' = 0 \quad \Rightarrow$$

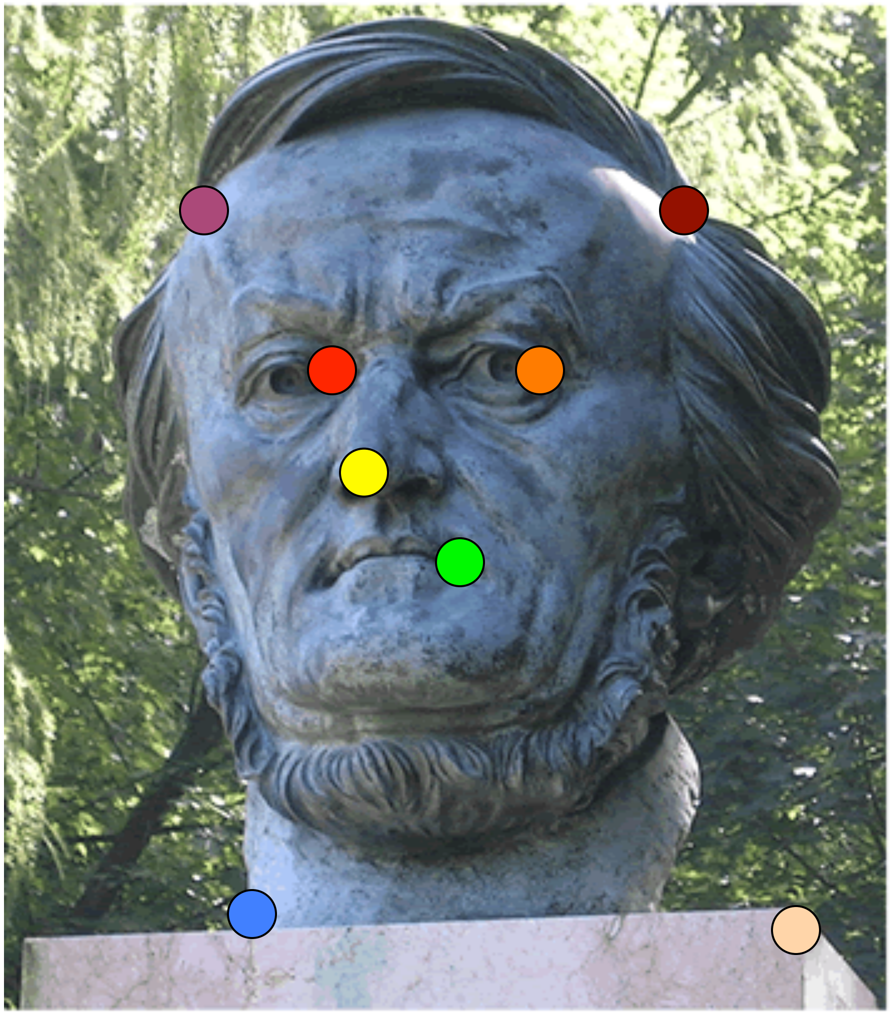
$$(u, v, 1)\begin{pmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{pmatrix}\begin{pmatrix} u' \\ v' \\ 1 \end{pmatrix} = 0$$

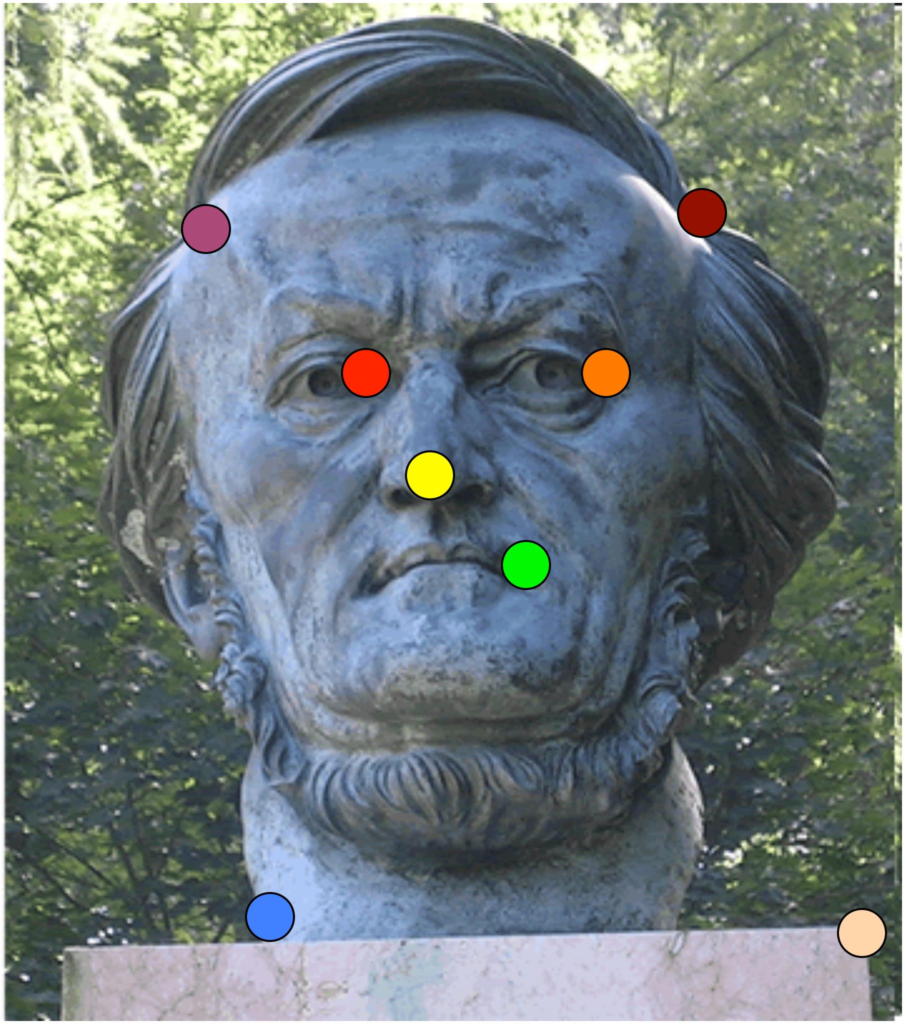
$$\Rightarrow \quad (uu', uv', u, vu', vv', v, u', v', 1)\begin{pmatrix} F_{11} \\ F_{12} \\ F_{13} \\ F_{21} \\ F_{22} \\ F_{23} \\ F_{31} \\ F_{32} \\ F_{33} \end{pmatrix} = 0$$

Let’s take 8 corresponding points

Slide source: S. Savarese.

Estimating F





Slide source: S. Savarese.

Estimating F

$$\underbrace{\begin{pmatrix} u_1u_1' & u_1v_1' & u_1 & v_1u_1' & v_1v_1' & v_1 & u_1' & v_1' & 1 \\ u_2u_2' & u_2v_2' & u_2 & v_2u_2' & v_2v_2' & v_2 & u_2' & v_2' & 1 \\ u_3u_3' & u_3v_3' & u_3 & v_3u_3' & v_3v_3' & v_3 & u_3' & v_3' & 1 \\ u_4u_4' & u_4v_4' & u_4 & v_4u_4' & v_4v_4' & v_4 & u_4' & v_4' & 1 \\ u_5u_5' & u_5v_5' & u_5 & v_5u_5' & v_5v_5' & v_5 & u_5' & v_5' & 1 \\ u_6u_6' & u_6v_6' & u_6 & v_6u_6' & v_6v_6' & v_6 & u_6' & v_6' & 1 \\ u_7u_7' & u_7v_7' & u_7 & v_7u_7' & v_7v_7' & v_7 & u_7' & v_7' & 1 \\ u_8u_8' & u_8v_8' & u_8 & v_8u_8' & v_8v_8' & v_8 & u_8' & v_8' & 1 \end{pmatrix}}_{\mathbf{W}} \underbrace{\begin{pmatrix} F_{11} \\ F_{12} \\ F_{13} \\ F_{21} \\ F_{22} \\ F_{23} \\ F_{31} \\ F_{32} \\ F_{33} \end{pmatrix}}_{\mathbf{f}} = 0$$

- Homogeneous system $\mathbf{W}\,\mathbf{f} = 0$
- Rank 8 → A non-zero solution exists (unique)
- If N>8 → Lsq. solution by SVD! → $\hat{\mathrm{F}}$

$$\|\mathbf{f}\| = 1$$

Slide source: S. Savarese.

$\hat{F}$ satisfies: $p^T \hat{F} p' = 0$

and estimated $\hat{F}$ may have full rank ($\det(\hat{F}) \neq 0$)

But remember: fundamental matrix is Rank 2

Find F that minimizes $\|F - \hat{F}\| = 0$

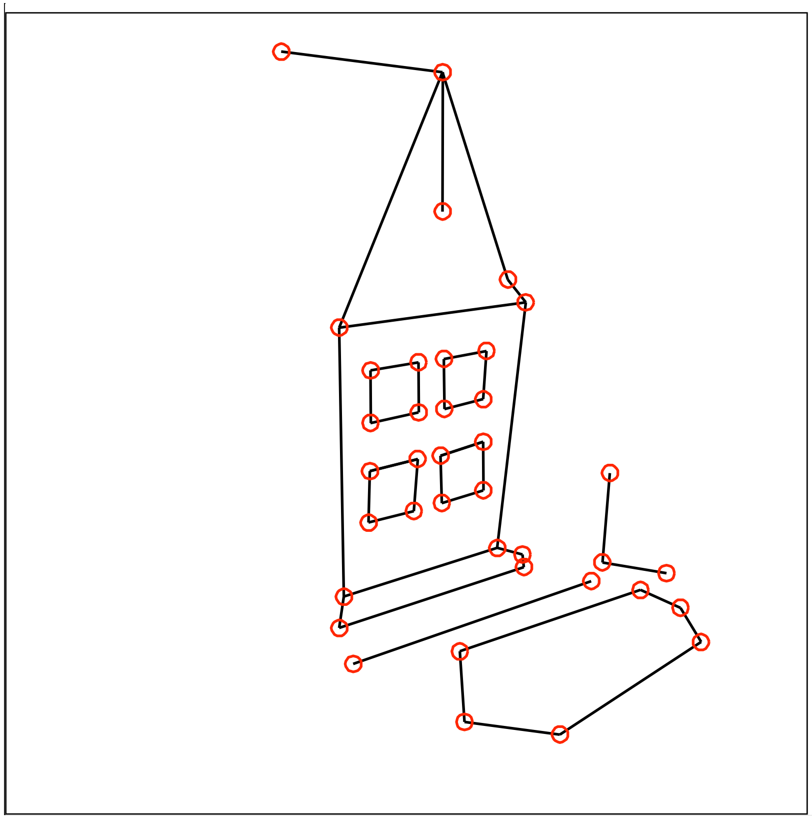
Frobenius norm (*)

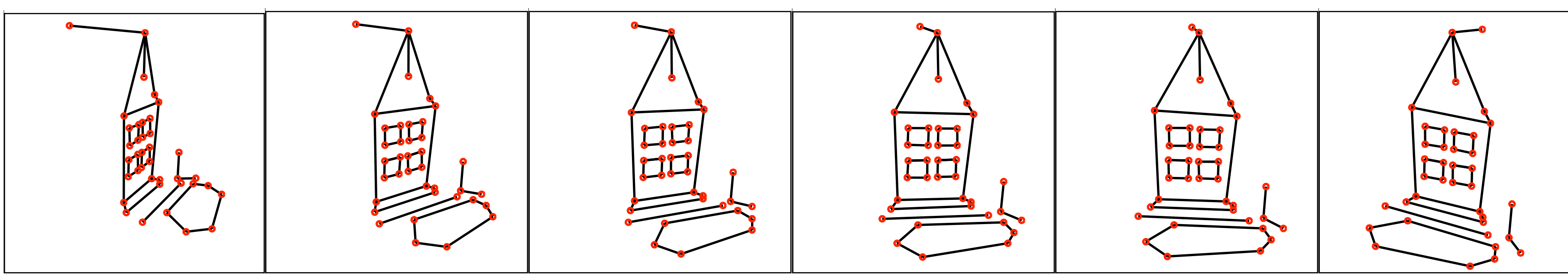
Subject to $\det(F)=0$

SVD (again!) can be used to solve this problem

(*) Sqrt root of the sum of squares of all entries

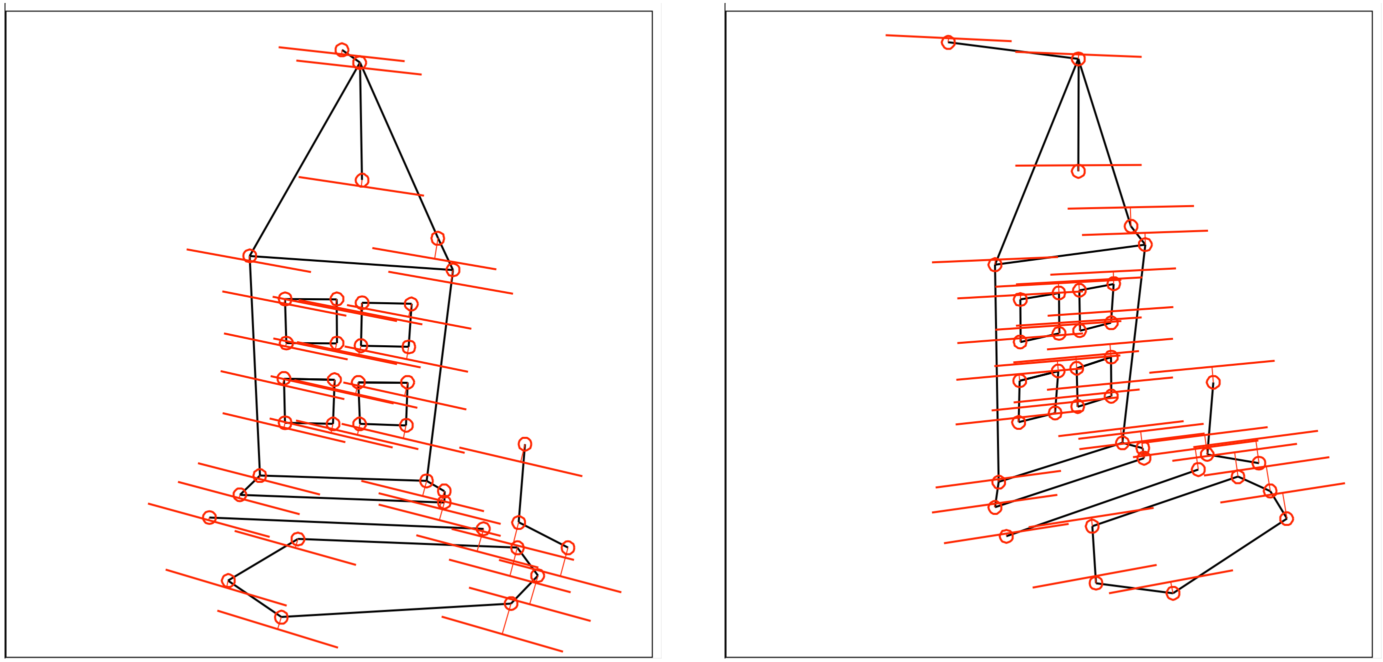
Slide source: S. Savarese.





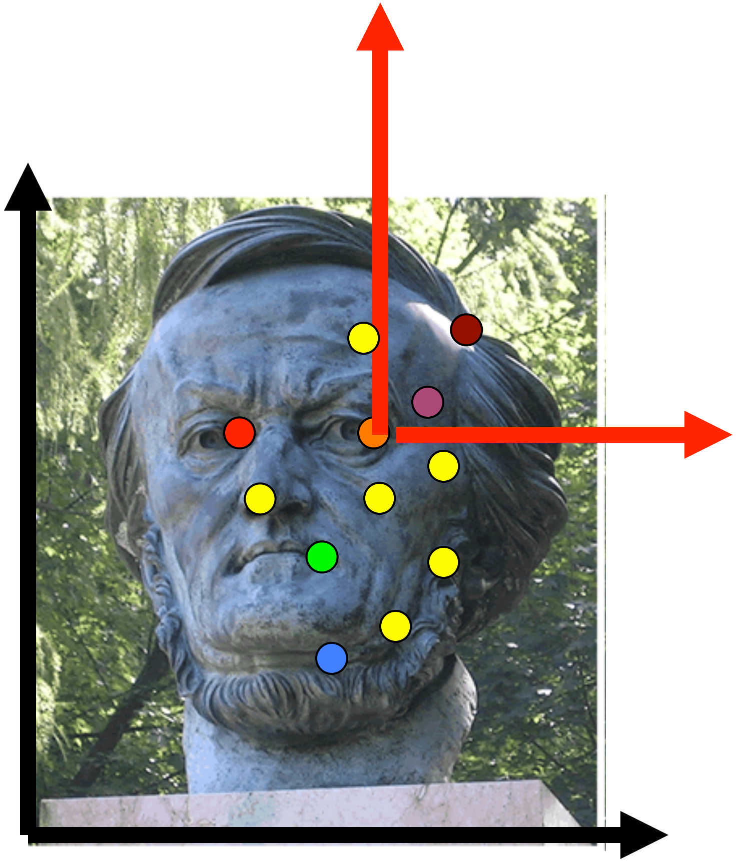
Data courtesy of R. Mohr and B. Boufama.

Slide source: S. Savarese.



Mean errors:
10.0pixel
9.1pixel

Problems with the 8-point Algorithm



$$\mathbf{W}\,\mathbf{f} = 0, \quad \|\mathbf{f}\| = 1 \xrightarrow{\text{Lsq solution by SVD}} \mathrm{F}$$

- Recall the structure of W:
  - do we see any potential (numerical) issue?

Slide source: S. Savarese.

Problems with the 8-point Algorithm

$$\mathbf{W}\,\mathbf{f} = 0$$

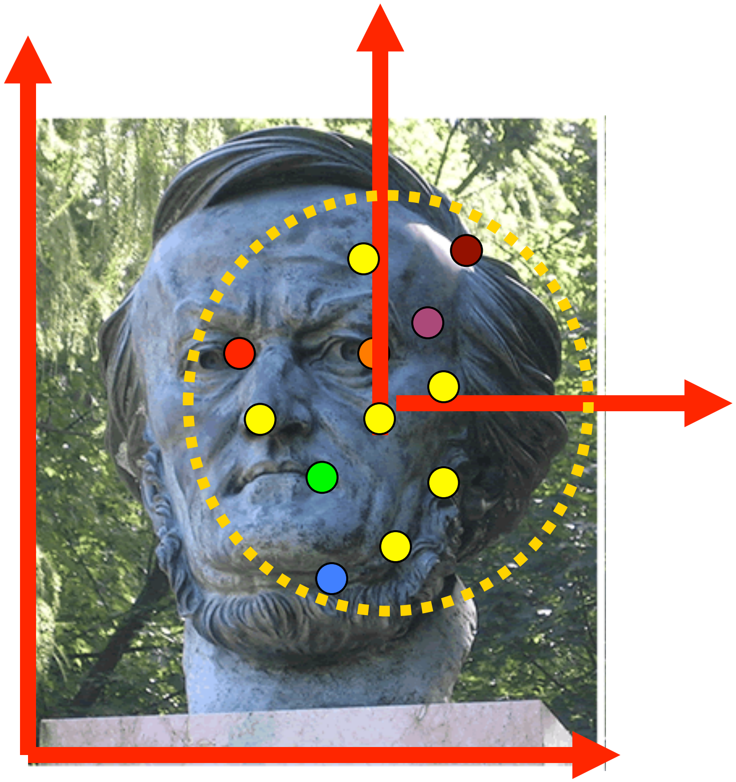
$$\mathbf{W}\; \begin{pmatrix} u_1u_1' & u_1v_1' & u_1 & v_1u_1' & v_1v_1' & v_1 & u_1' & v_1' & 1 \\ u_2u_2' & u_2v_2' & u_2 & v_2u_2' & v_2v_2' & v_2 & u_2' & v_2' & 1 \\ u_3u_3' & u_3v_3' & u_3 & v_3u_3' & v_3v_3' & v_3 & u_3' & v_3' & 1 \\ u_4u_4' & u_4v_4' & u_4 & v_4u_4' & v_4v_4' & v_4 & u_4' & v_4' & 1 \\ u_5u_5' & u_5v_5' & u_5 & v_5u_5' & v_5v_5' & v_5 & u_5' & v_5' & 1 \\ u_6u_6' & u_6v_6' & u_6 & v_6u_6' & v_6v_6' & v_6 & u_6' & v_6' & 1 \\ u_7u_7' & u_7v_7' & u_7 & v_7u_7' & v_7v_7' & v_7 & u_7' & v_7' & 1 \\ u_8u_8' & u_8v_8' & u_8 & v_8u_8' & v_8v_8' & v_8 & u_8' & v_8' & 1 \end{pmatrix} \begin{pmatrix} F_{11} \\ F_{12} \\ F_{13} \\ F_{21} \\ F_{22} \\ F_{23} \\ F_{31} \\ F_{32} \\ F_{33} \end{pmatrix} \;\mathbf{f} = 0$$

- Highly un-balanced (not well conditioned)
- Values of W must have similar magnitude
- This creates problems during the SVD decomposition

HZ pag 108

Slide source: S. Savarese.

Normalization



IDEA: Transform image coordinate such that the matrix **W** become better conditioned

Apply following transformation T: (translation and scaling)

- Origin = centroid of image points
- Mean square distance of the data points from origin is 2 pixels

$$q_i = T_i \, p_i \qquad q_i' = T_i' \, p_i' \qquad \text{(normalization)}$$

Slide source: S. Savarese.

The Normalized 8-point Algorithm

0. Compute T_i and T_i'

1. Normalize coordinates:

$$q_i = T_i \, p_i \qquad q_i' = T_i' \, p_i'$$

2. Use the eight-point algorithm to compute F'_q from the points q_i and q'_i

1. Enforce the rank-2 constraint. $\rightarrow F_q$

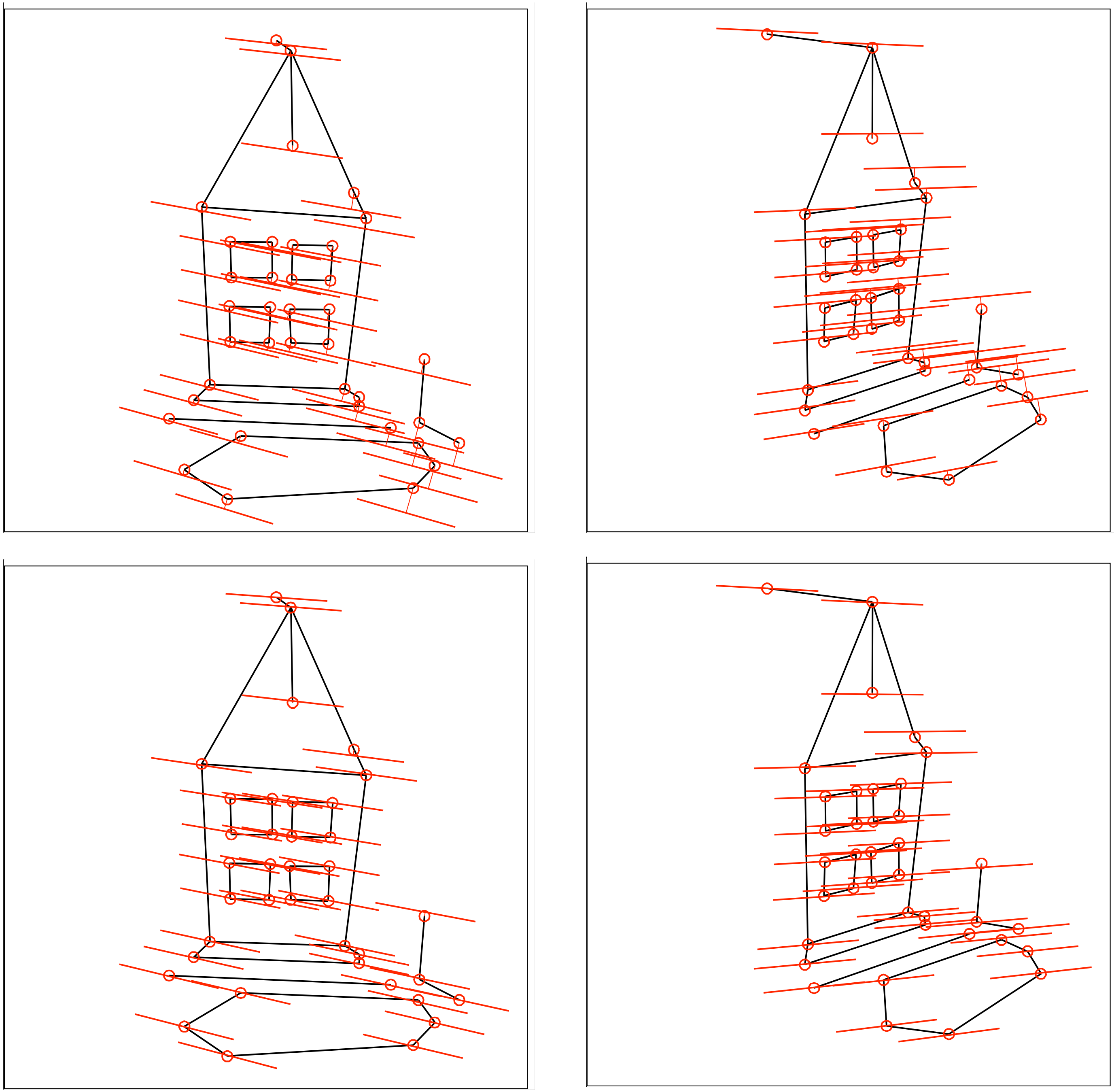
$$\begin{cases} q^T F_q \, q' = 0 \\ \det(F_q) = 0 \end{cases}$$

2. De-normalize F_q: $F = T'^T F_q T$

Slide source: S. Savarese.

Without transformation

With transformation



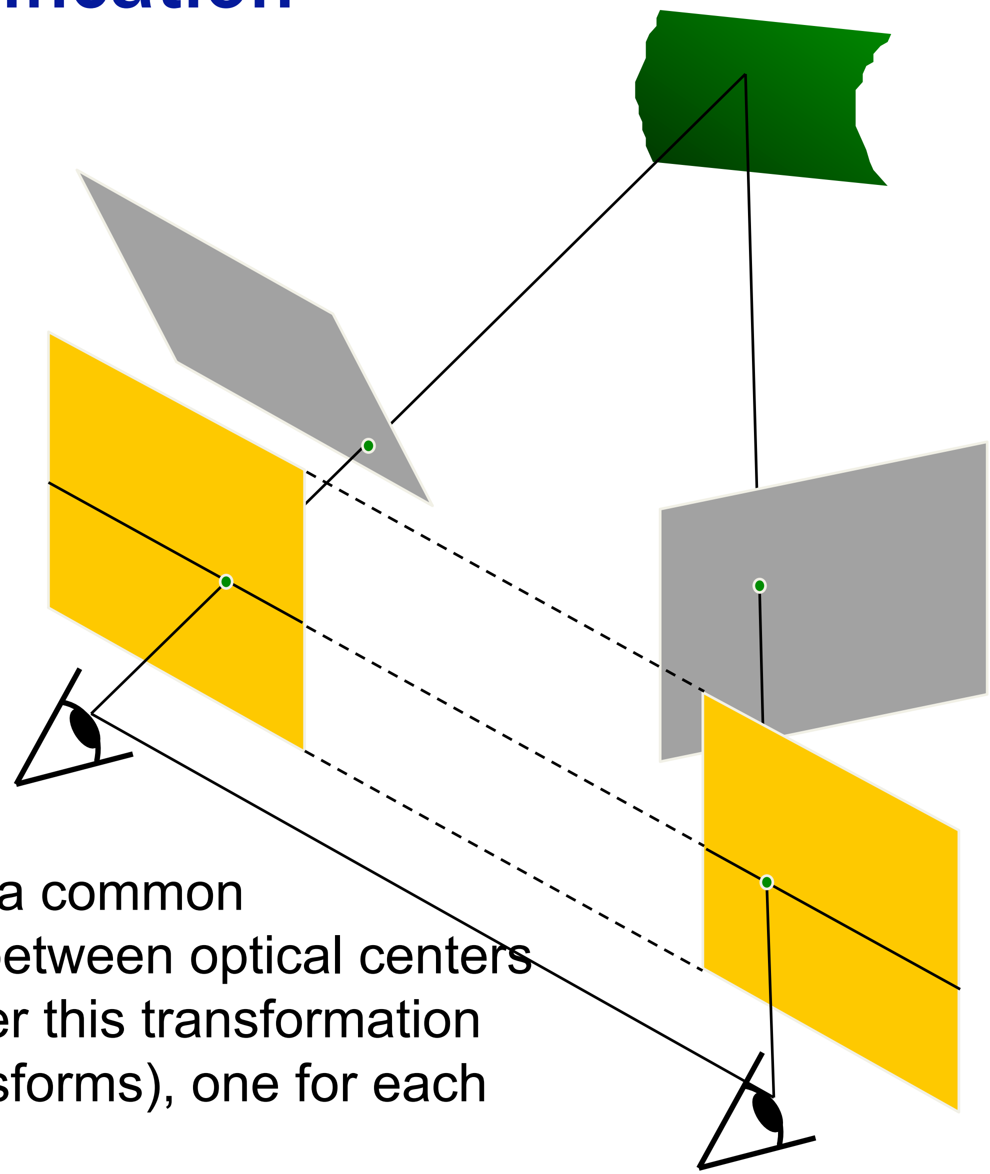
Mean errors:
10.0pixel
9.1pixel

Mean errors:
1.0pixel
0.9pixel

Slide source: S. Savarese.

Stereo image rectification

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

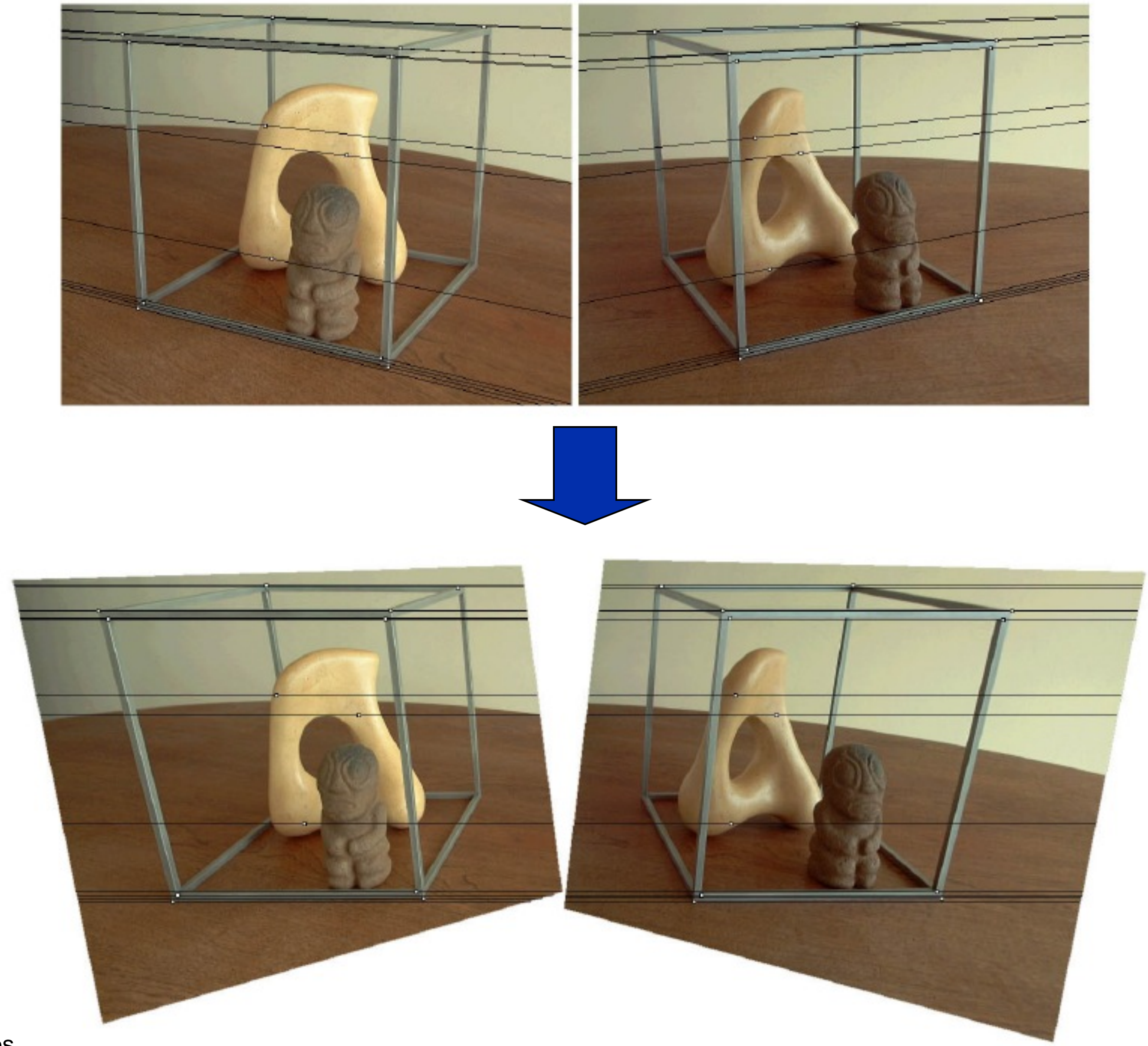


- reproject image planes onto a common plane parallel to the line between optical centers
- pixel motion is horizontal after this transformation
- two homographies (3x3 transforms), one for each input image reprojection

Adapted from Li Zhang

C. Loop and Z. Zhang. Computing Rectifying Homographies for Stereo Vision. CVPR 1999.

Stereo image rectification: example



Source: Alyosha Efros

Next Lecture: Stereo Vision

- Readings: FP 7; SZ 11; TV 7