CEE598 - Visual Sensing for Civil Infrastructure Eng. & Mgmt.

Session12 – Epipolar Geometry

Mani Golparvar-Fard
Department of Civil and Environmental Engineering
3129D, Newmark Civil Engineering Lab
e-mail: mgolpar@illinois.edu

Reminders

- An updated Syllabus will be posted.
- Term project mid-term presentations
  - Mar 28
  - Each group presenting for 5minutes

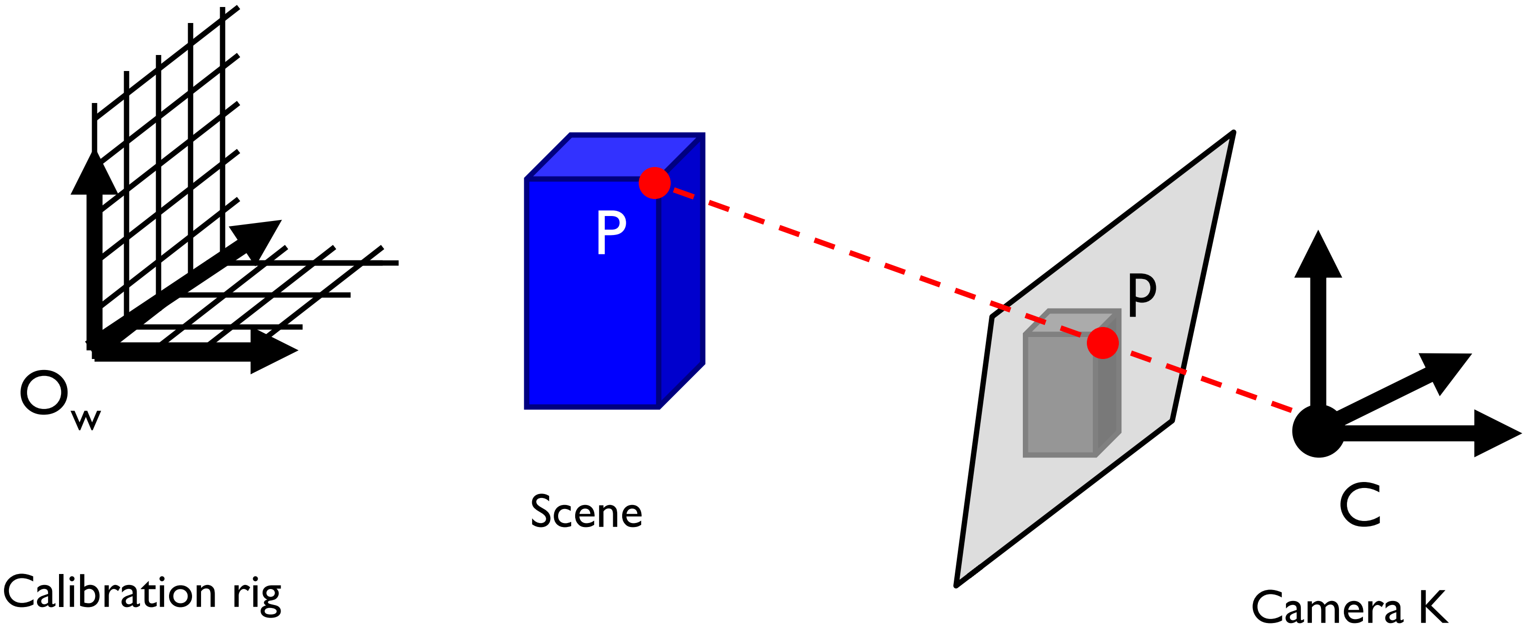
Slide 1- Introduction to the Engineering problem your work is addressing;
Slide 2- Review of previous works, or the works you're following for your own implementation;
Slide 3- A summary of your technical solution;
Slide 4- Presenting experimental results and discuss validation approach (if you do not have experimental results by then, that's fine..though please prepare a detailed plan as to how you will implement and validate your algorithm; and
Slide 5- Plan and schedule of activities for the remainder of the project.

Outline

- Why is stereo useful?
- Epipolar constraints
- Essential and fundamental matrix
- Estimating F
- Examples

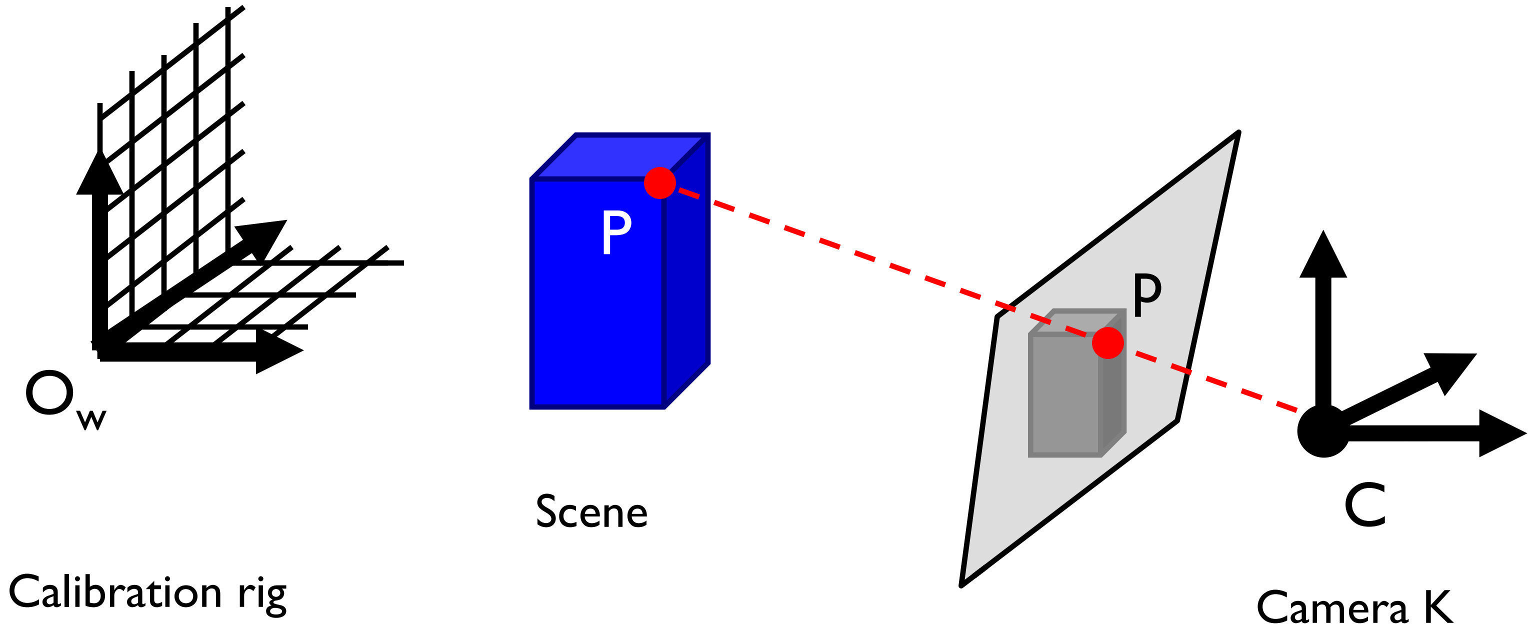
Reading: [HZ] Chapters: 4, 9, 11
[FP] Chapters: 10

Recovering structure from a single view



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

Recovering structure from a single view



Why is it so difficult?

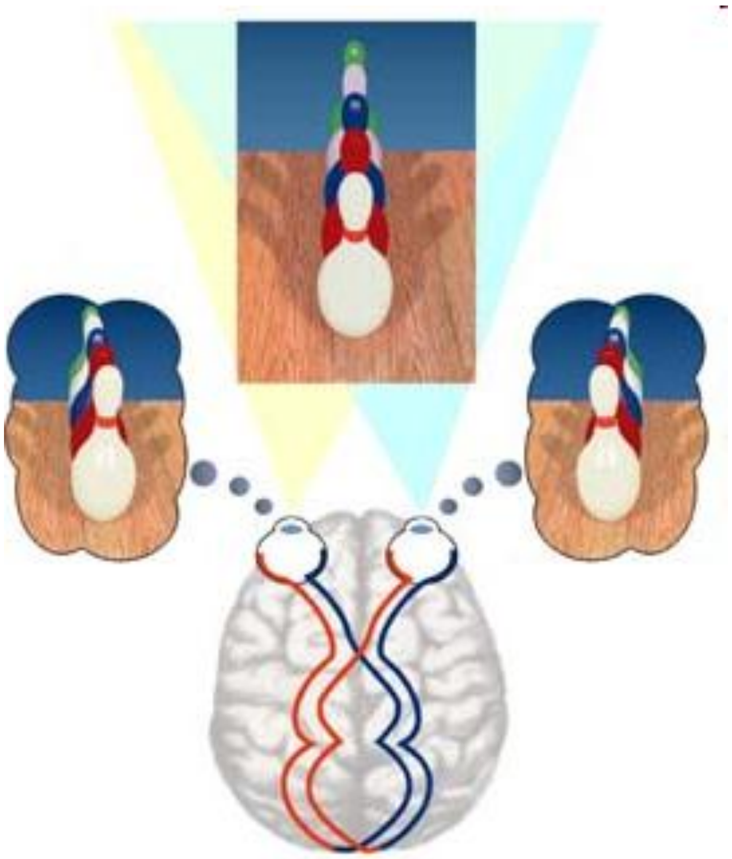
Intrinsic ambiguity of the mapping from 3D to image (2D)

Recovering structure from a single view

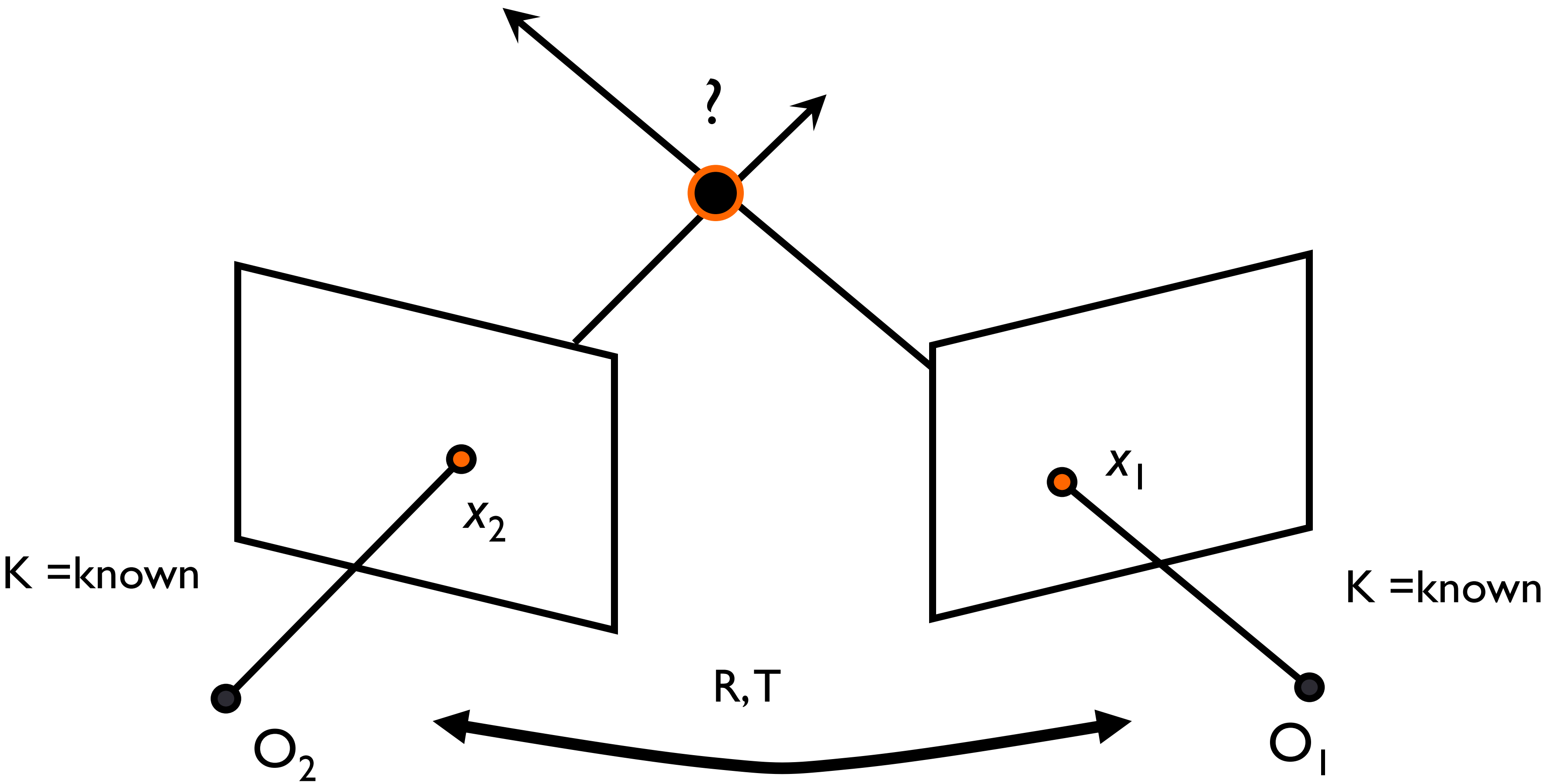
- Intrinsic ambiguity of the mapping from 3D to image (2D)

Courtesy slide S. Lazebnik

Two eyes help!



Two eyes help!

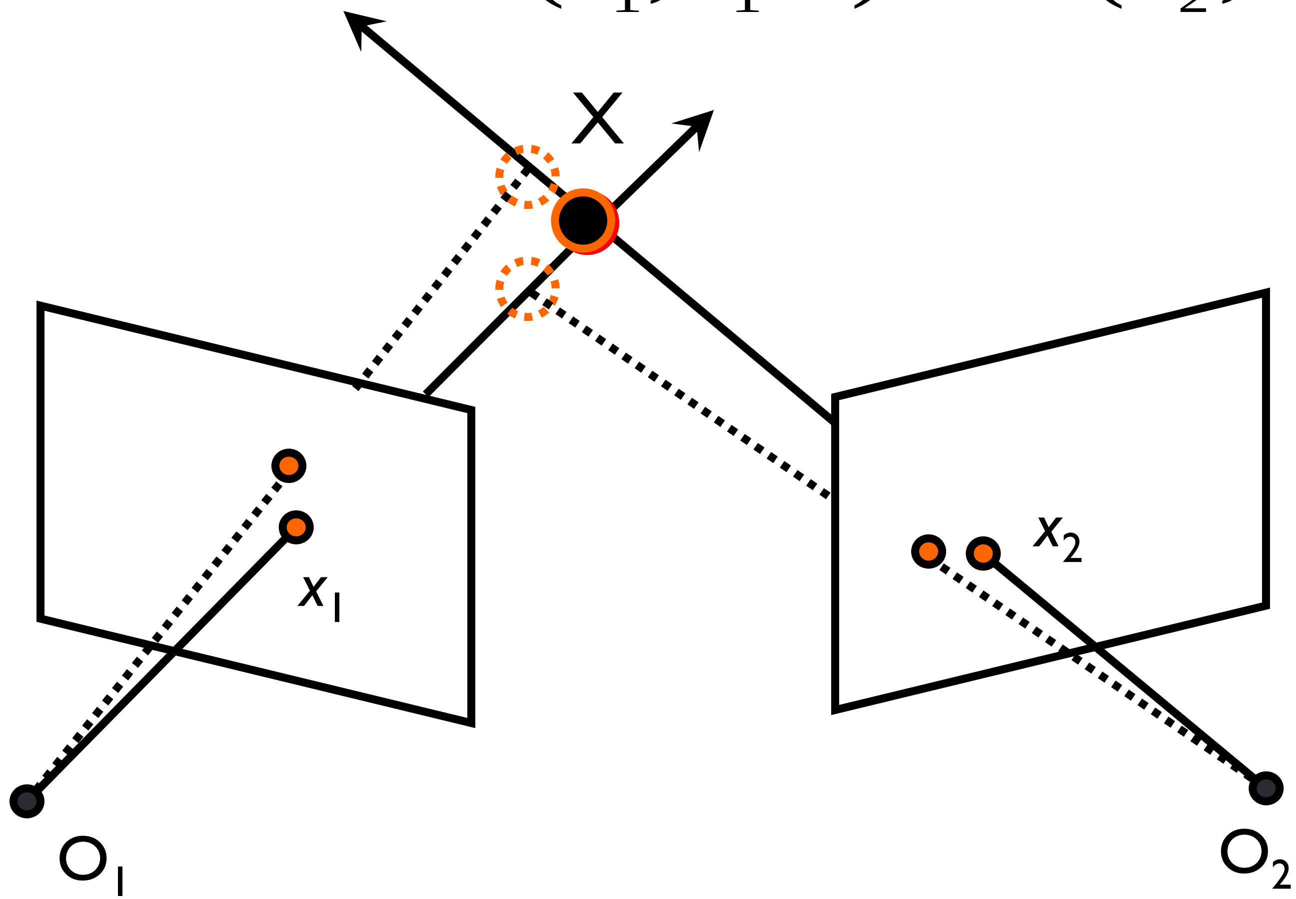


This is called triangulation

Triangulation

- Find X that minimizes

$$d^2(x_1, P_1X) + d^2(x_2, P_2X)$$



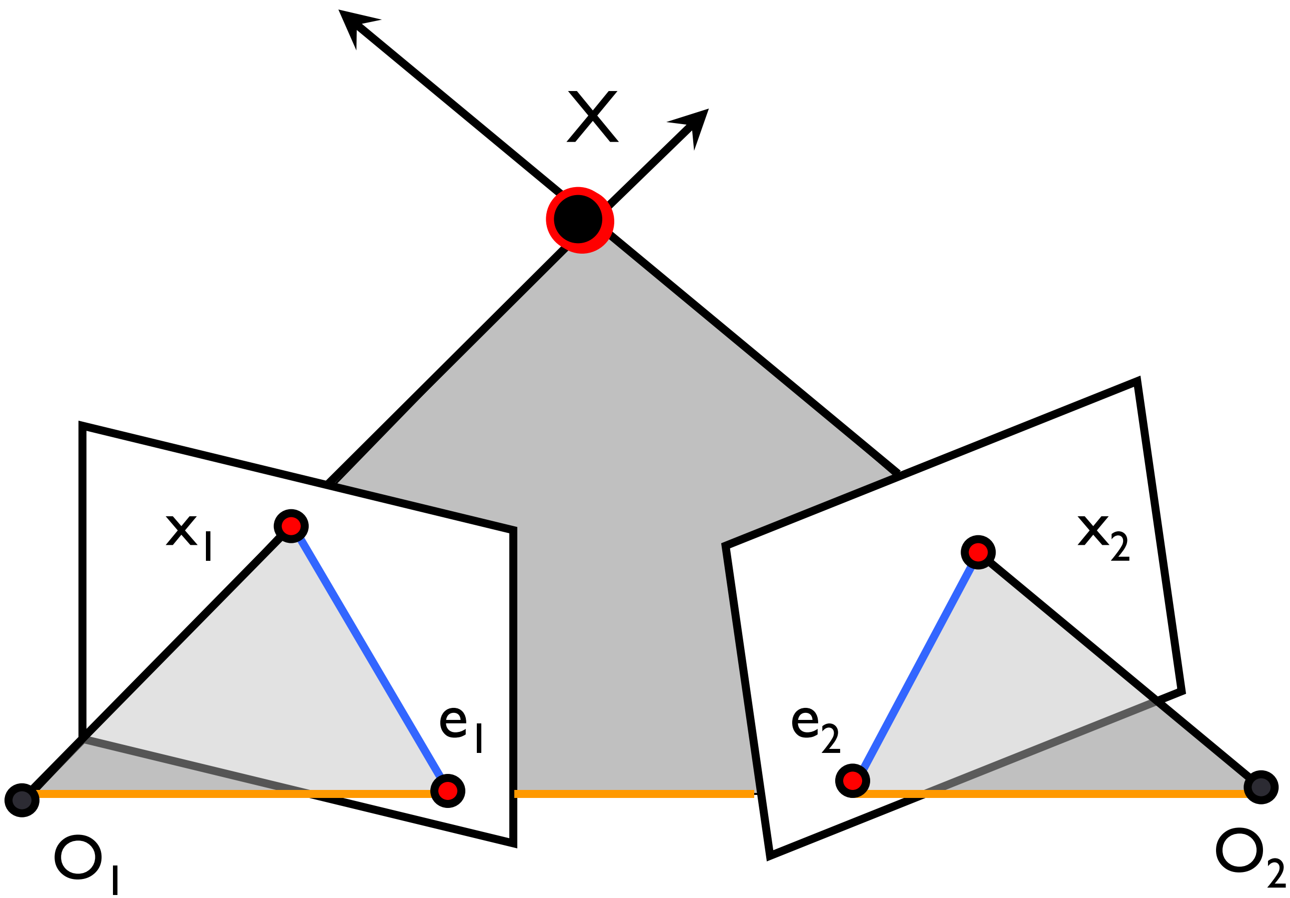
Stereo-view geometry

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

- **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.

Epipolar geometry

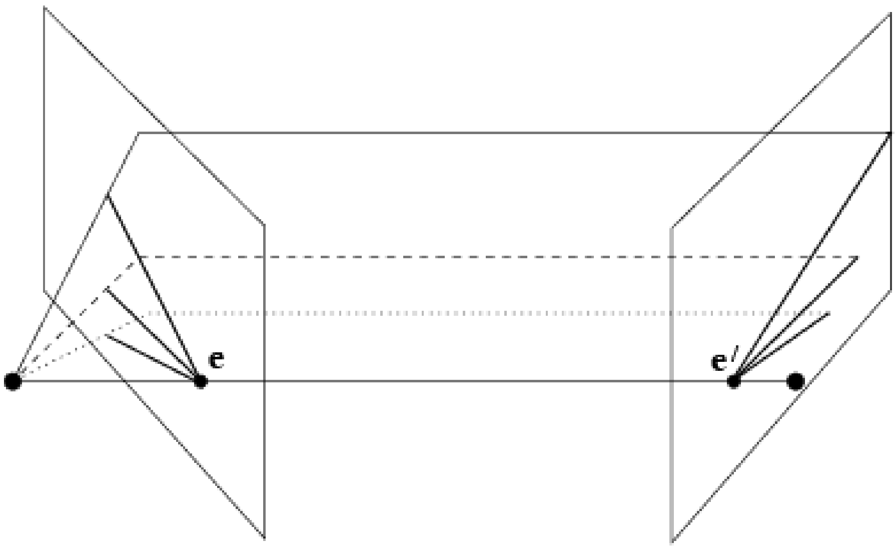


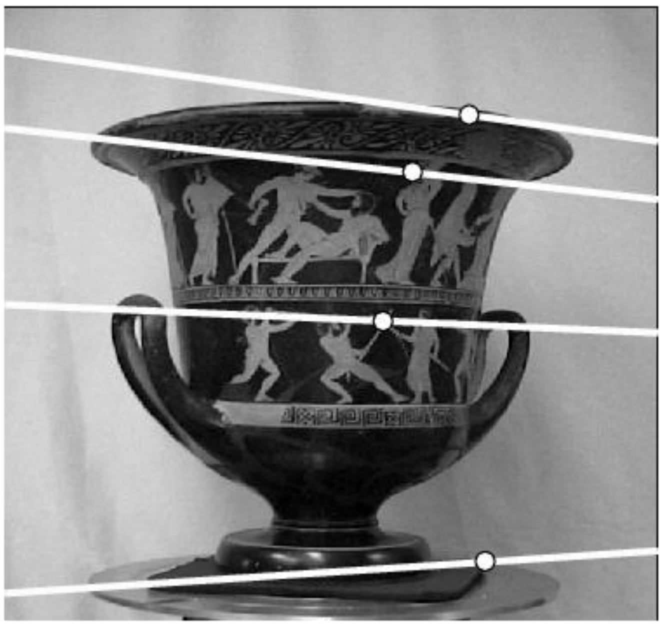
- Epipolar Plane
- Baseline
- Epipolar Lines

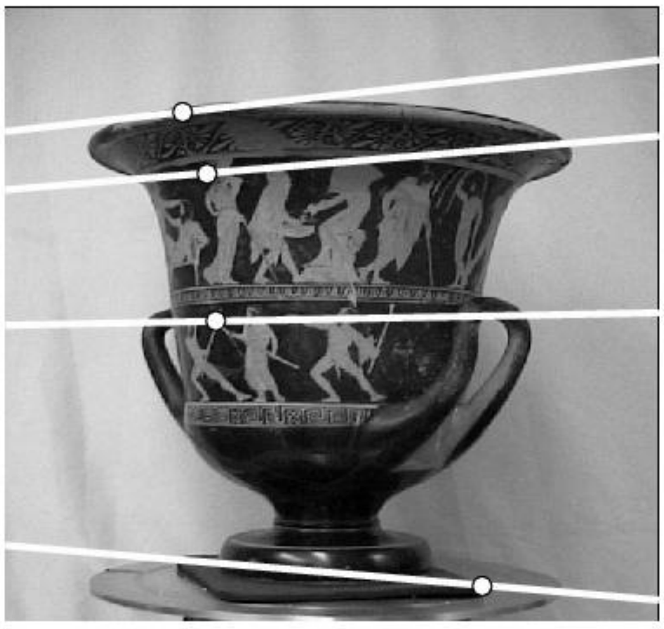
- Epipoles e_1, e_2

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

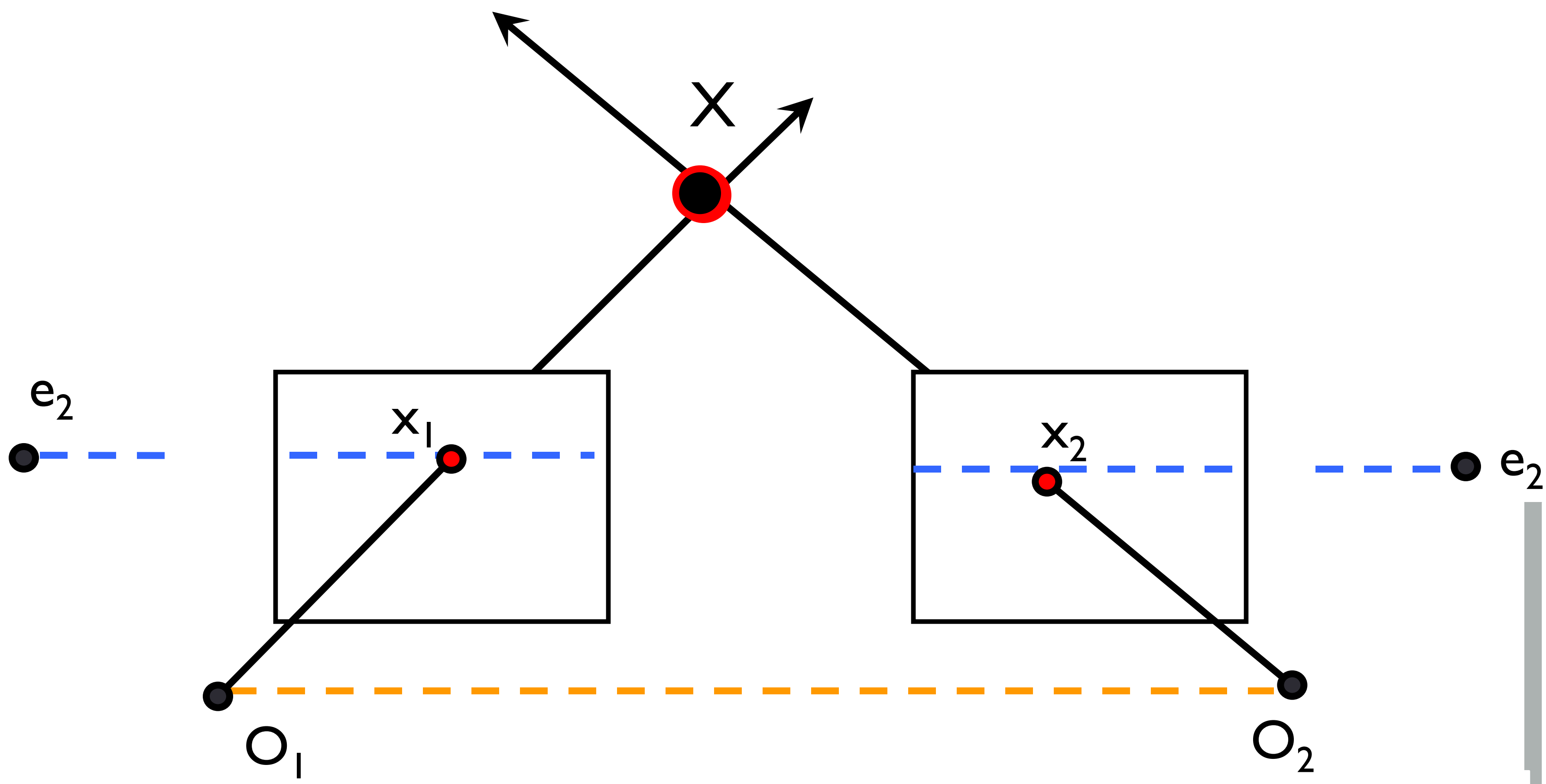
Example: Converging image planes





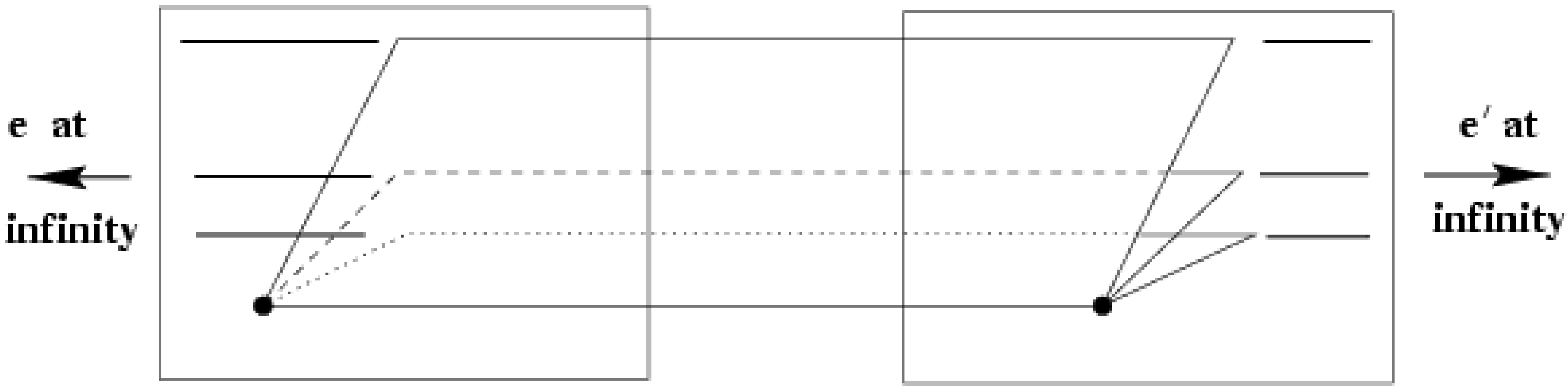


Example: Parallel image planes

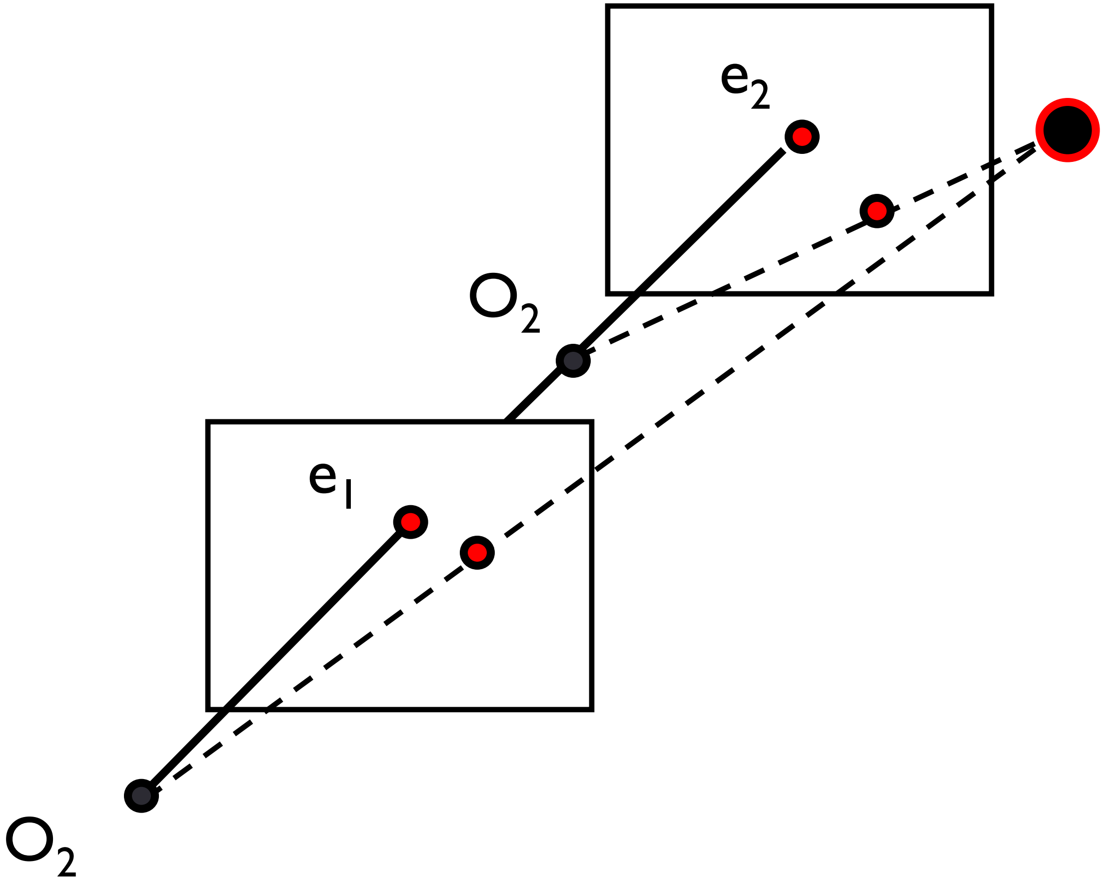


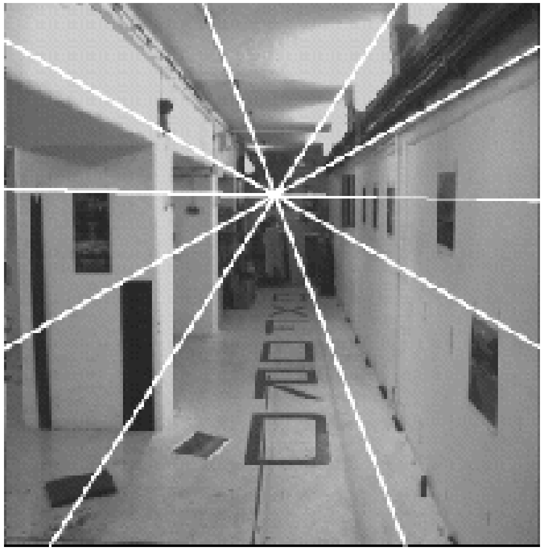
- Baseline intersects the image plane at infinity
- Epipoles are at infinity
- Epipolar lines are parallel to x axis

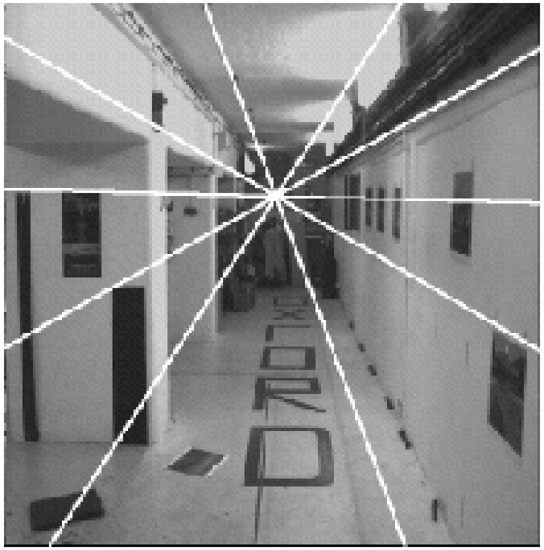
Example: Parallel image planes



Example: Forward translation

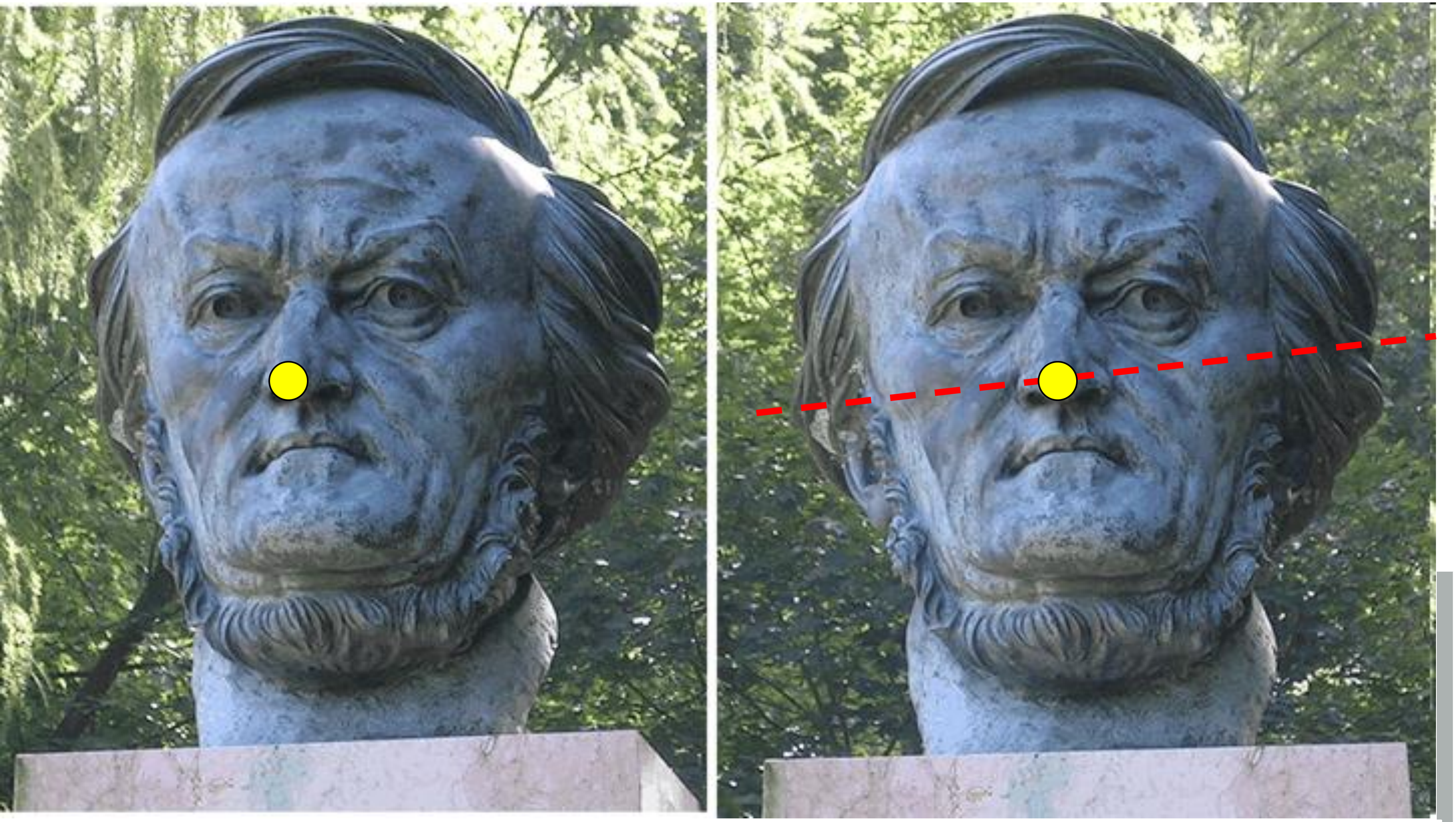






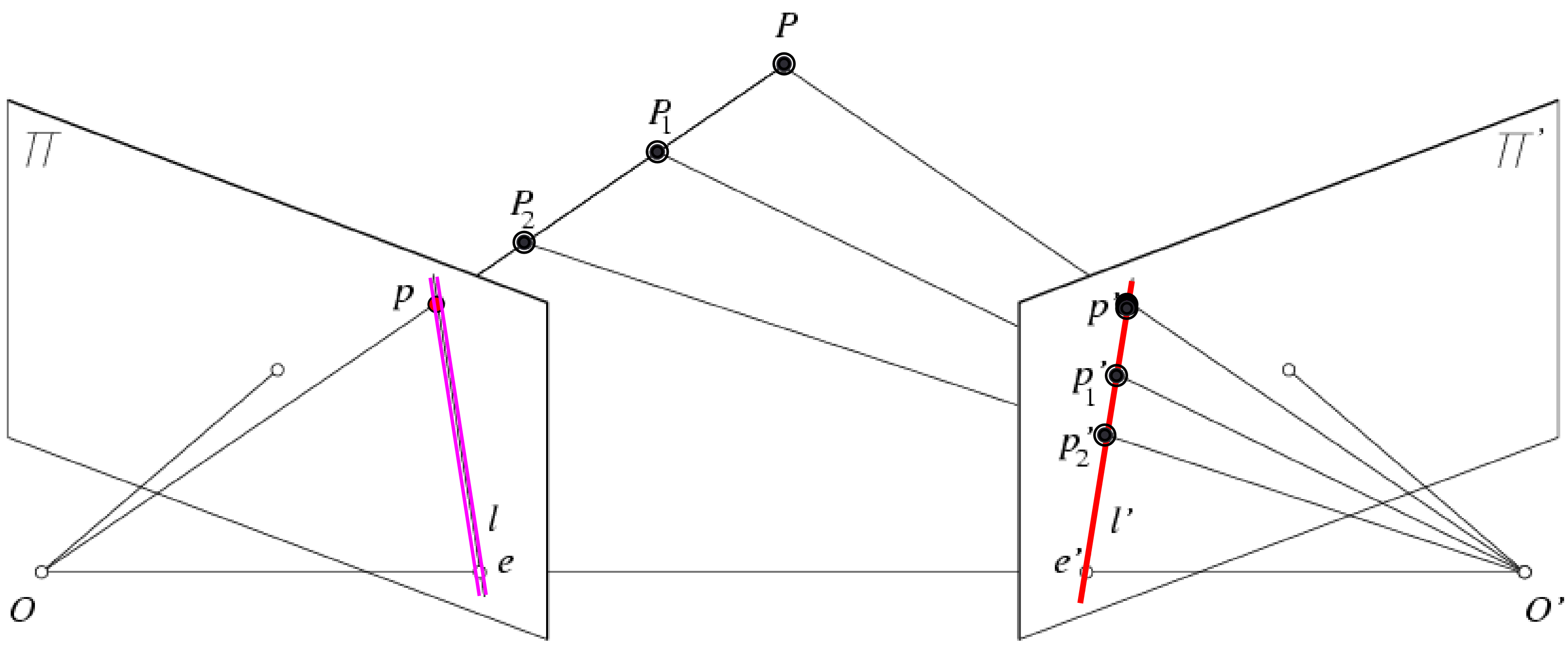
- The epipoles have same position in both images
- Epipole called FOE (focus of expansion)

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?

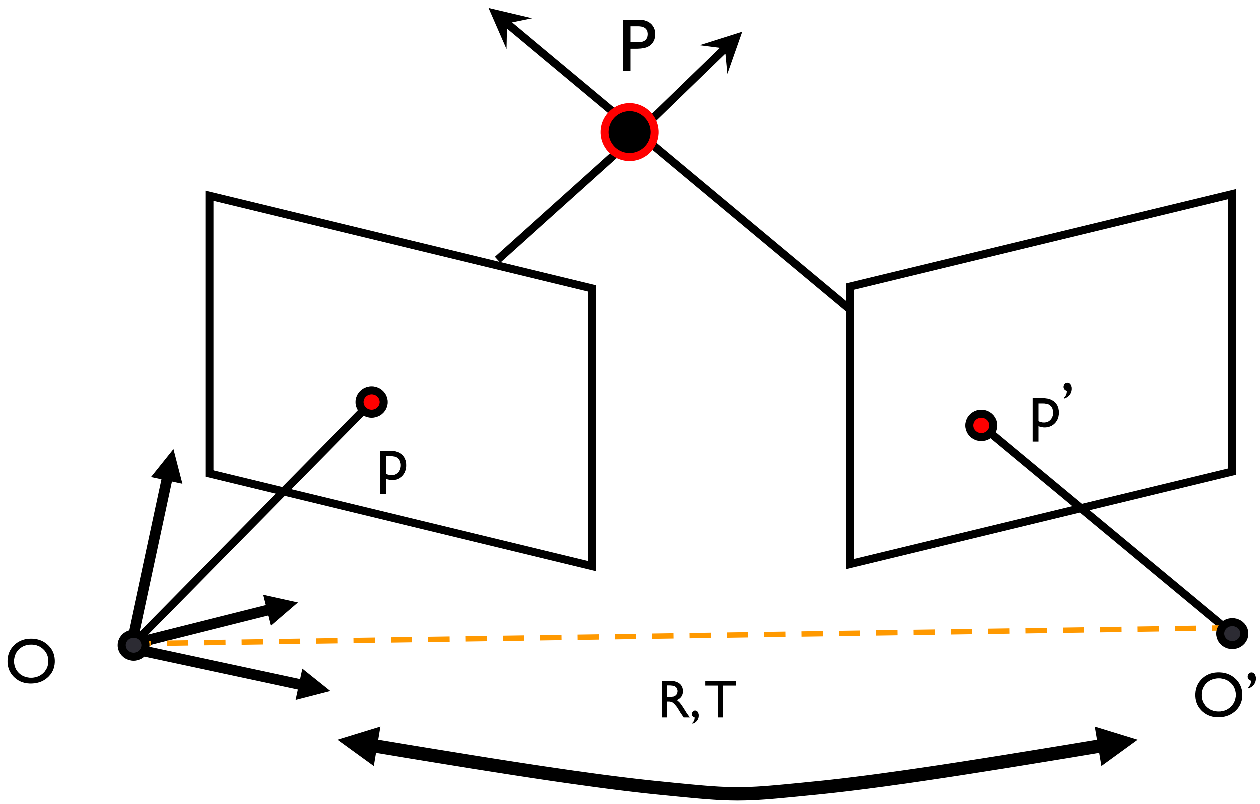
Epipolar Constraint



- 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*.

Epipolar Constraint



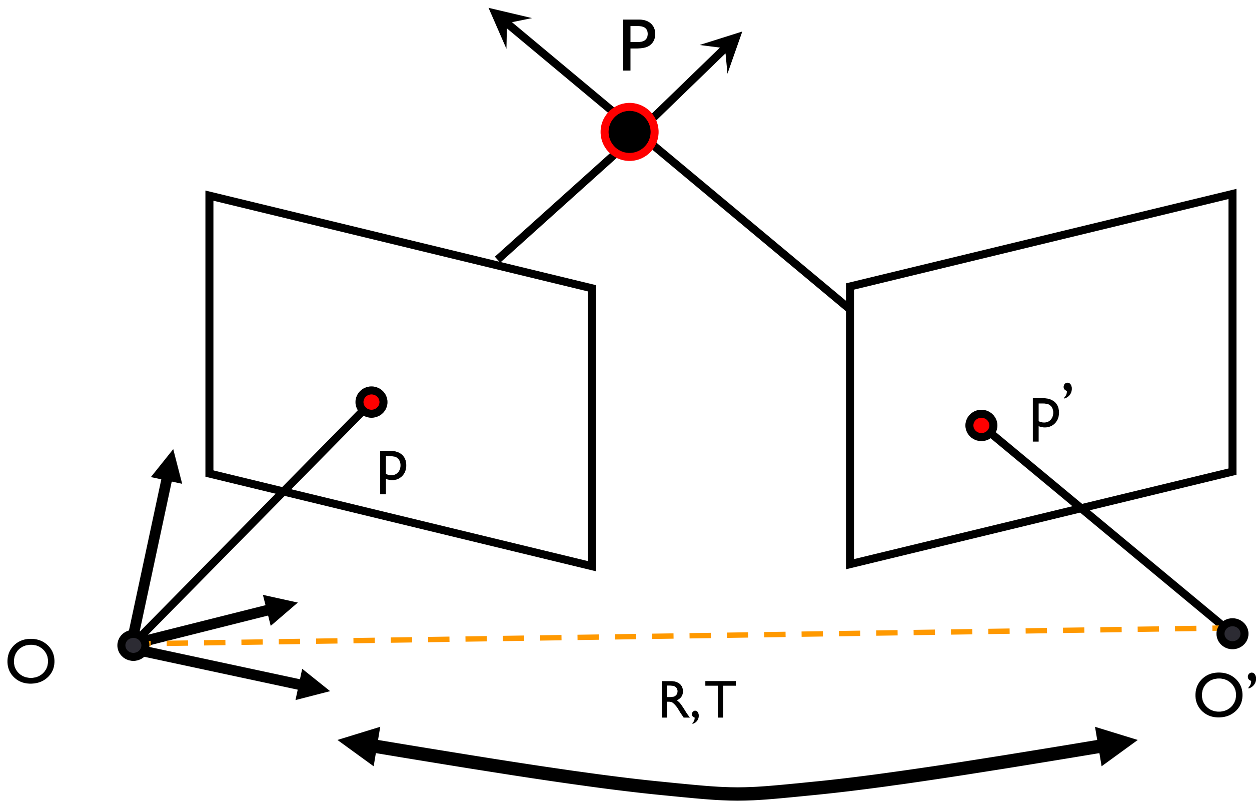
$$M = K\begin{bmatrix} I & 0 \end{bmatrix}$$

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

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

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

Epipolar Constraint



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

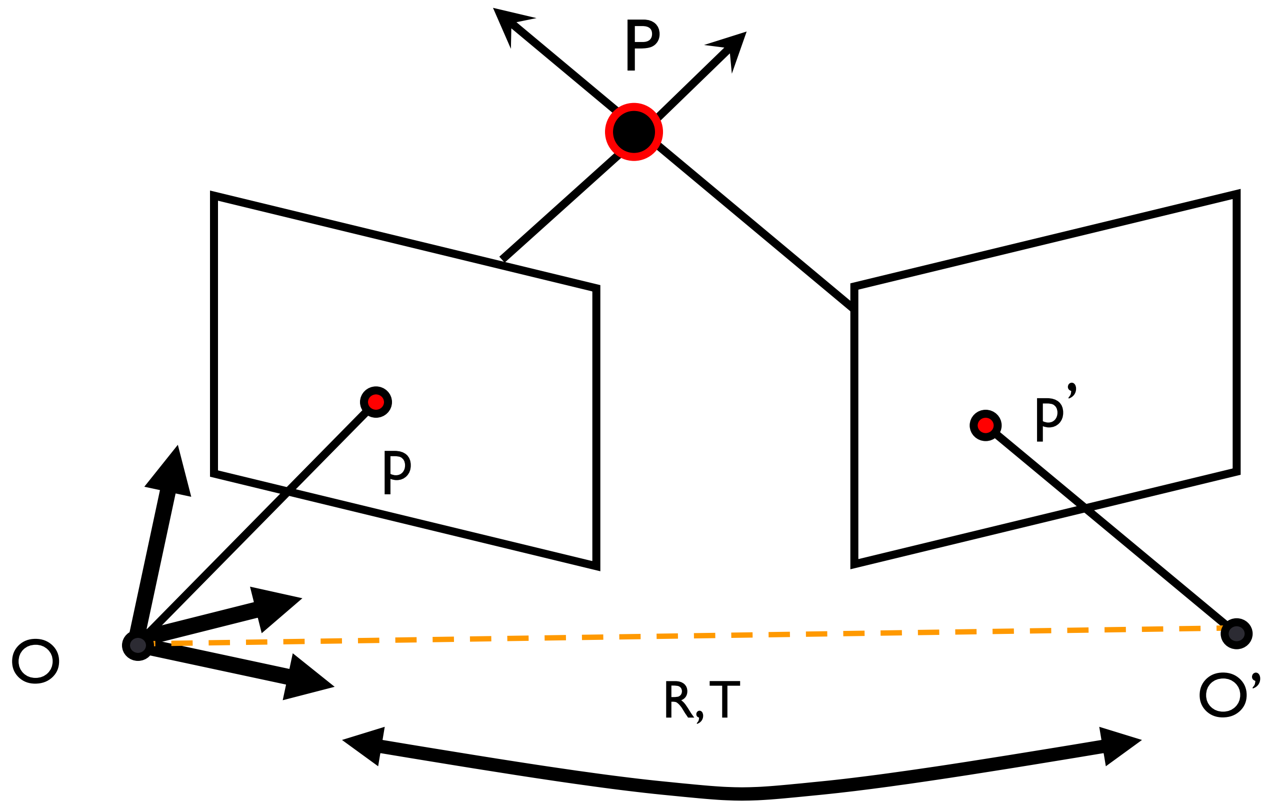
Perpendicular to epipolar plane

K_1 and K_2 are known (calibrated cameras)

Cross product as matrix multiplication

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

Epipolar Constraint

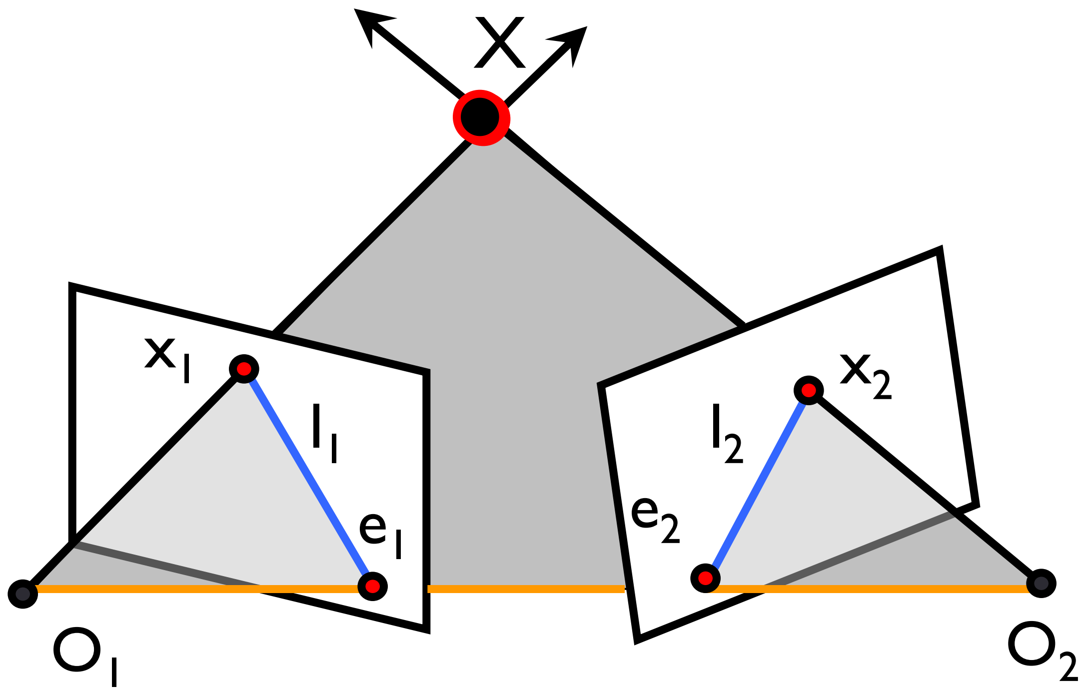


$$p^T \cdot [T \times (R\, p')] = 0 \rightarrow p^T \cdot [T_\times] \cdot R\, p' = 0$$

E = essential matrix

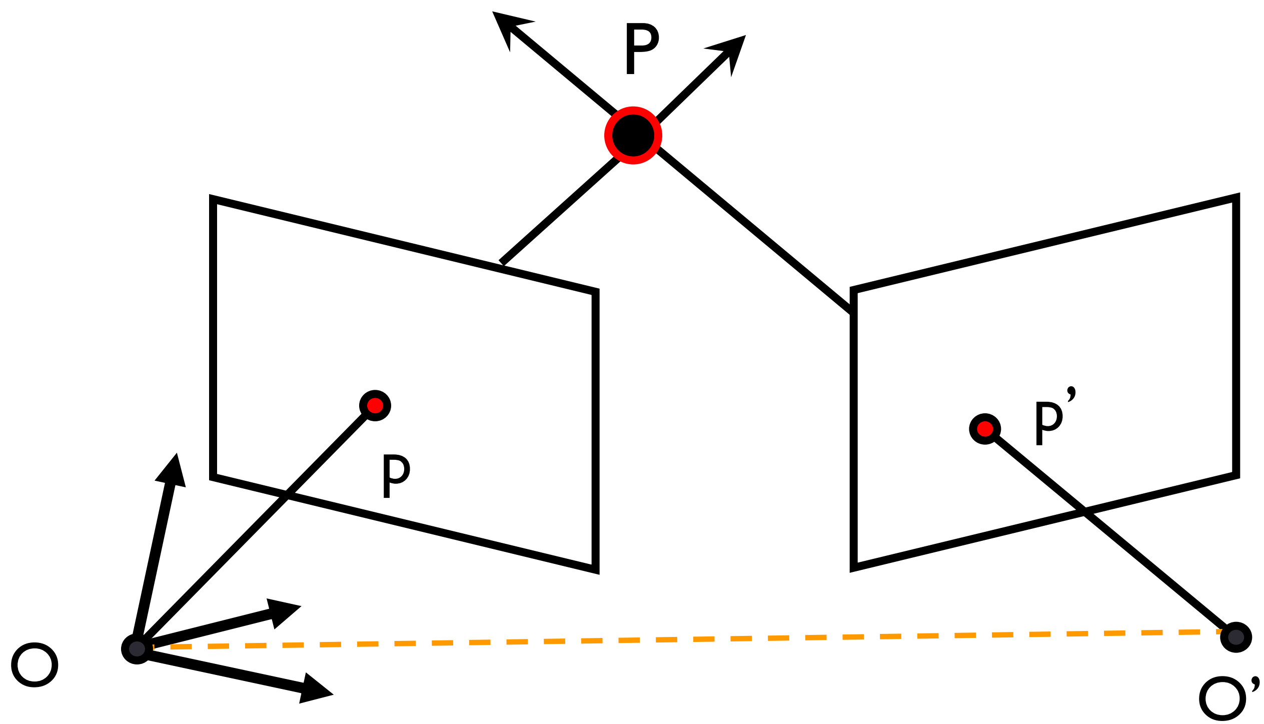
(Longuet-Higgins, 1981)

Epipolar Constraint



- $E\,x_2$ is the epipolar line associated with x_2 ($l_1 = E\,x_2$)
- $E^T x_1$ is the epipolar line associated with x_1 ($l_2 = E^T x_1$)
- E is singular (rank two)
- $E\,e_2 = 0$ and $E^T e_1 = 0$
- E is 3x3 matrix; 5 DOF

Epipolar Constraint



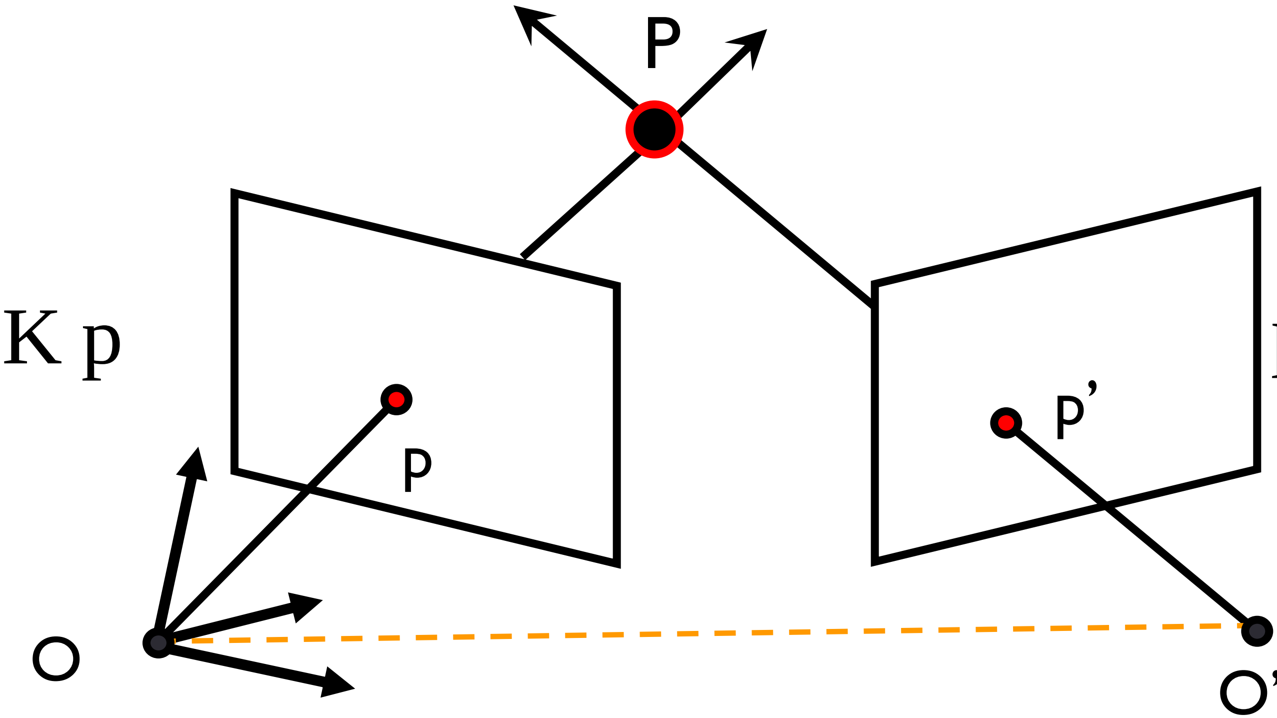
$$P \rightarrow M\,P \quad \rightarrow p = \begin{bmatrix} u \\ v \end{bmatrix}$$

$$M = K[I \quad 0]$$

unknown

Epipolar Constraint

$p \rightarrow K\,p$

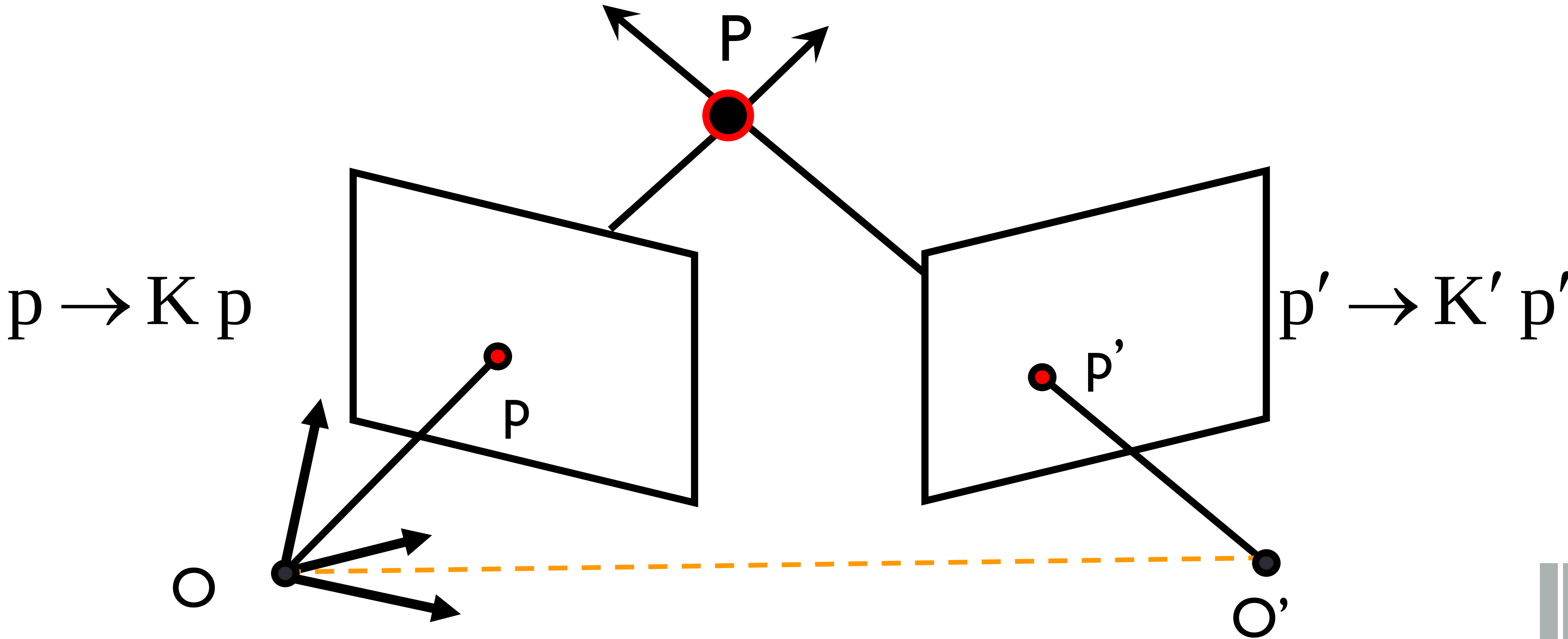


$p' \rightarrow K'\,p'$

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

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

Epipolar Constraint

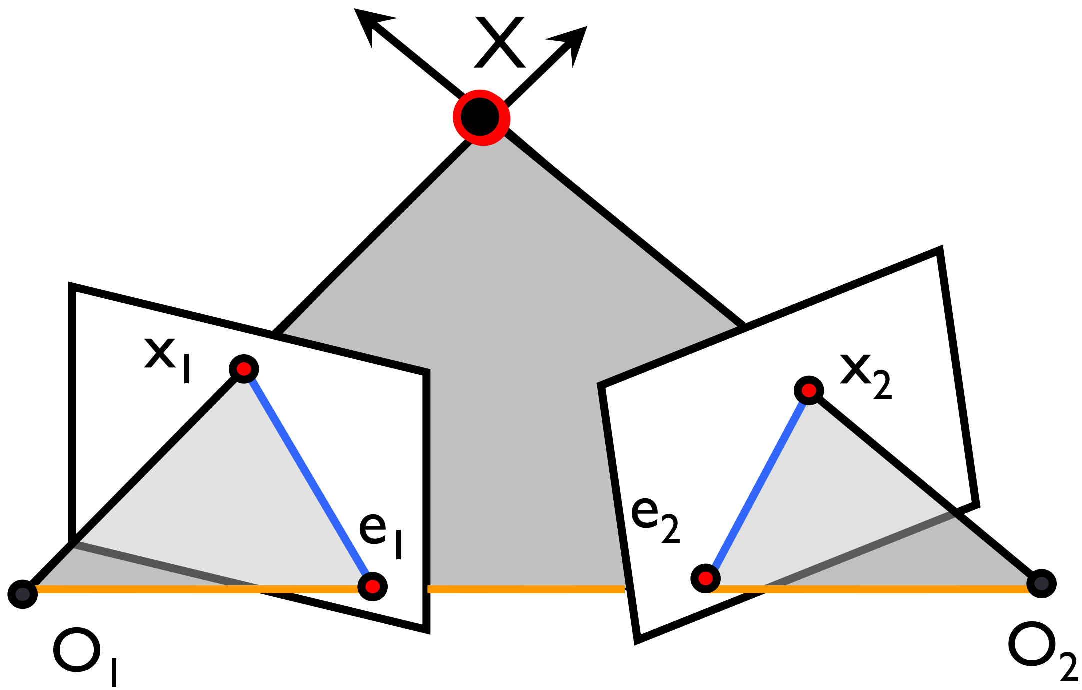


$$p^T F\, p' = 0$$

F = Fundamental Matrix

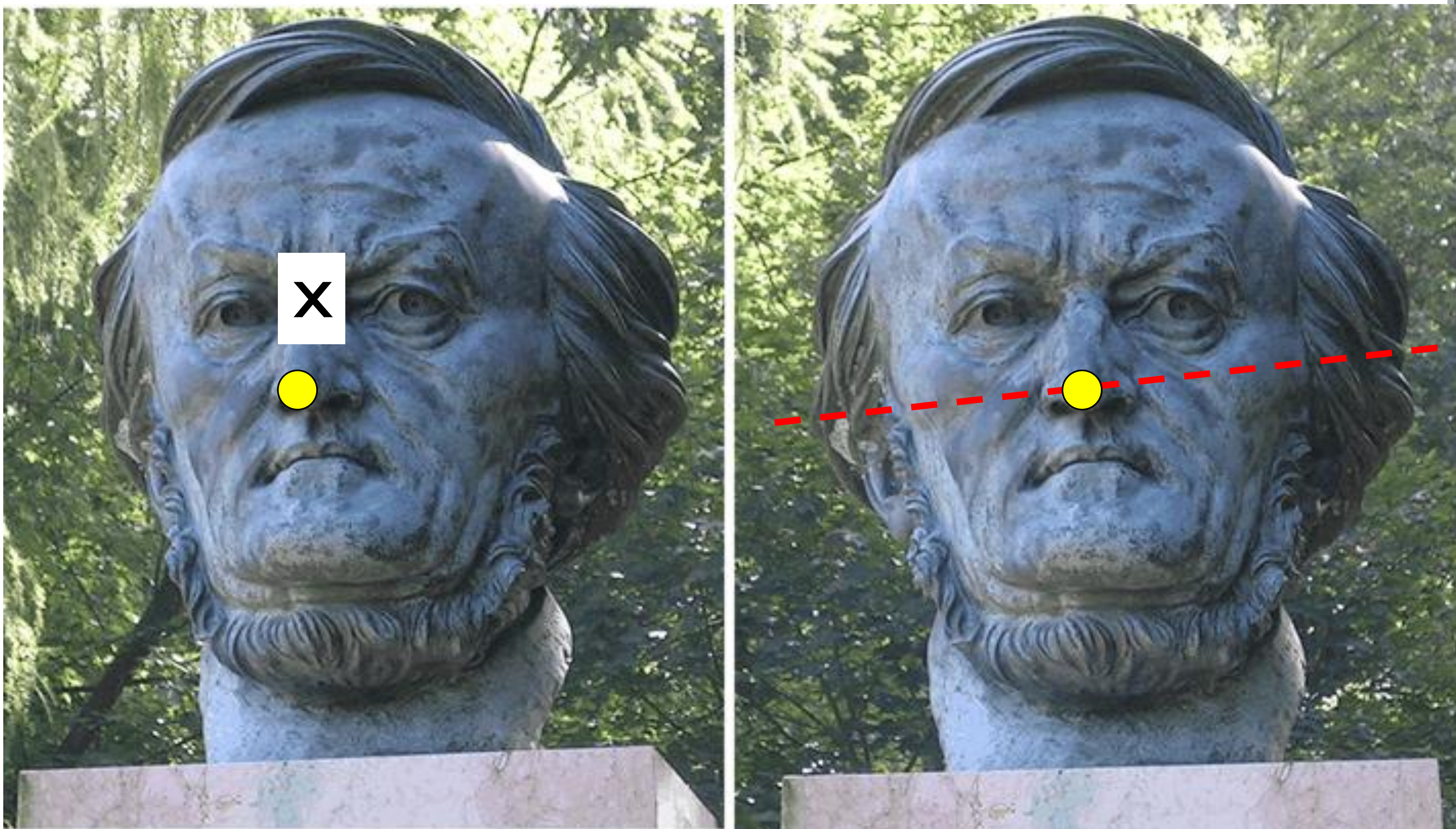
(Faugeras and Luong, 1992)

Epipolar Constraint



- $F x_2$ is the epipolar line associated with x_2 ($l_1 = F x_2$)
- $F^T x_1$ is the epipolar line associated with x_1 ($l_2 = F^T x_1$)
- F is singular (rank two)
- $F e_2 = 0$ and $F^T e_1 = 0$
- F is 3x3 matrix; 7 DOF

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?

Why F is useful?

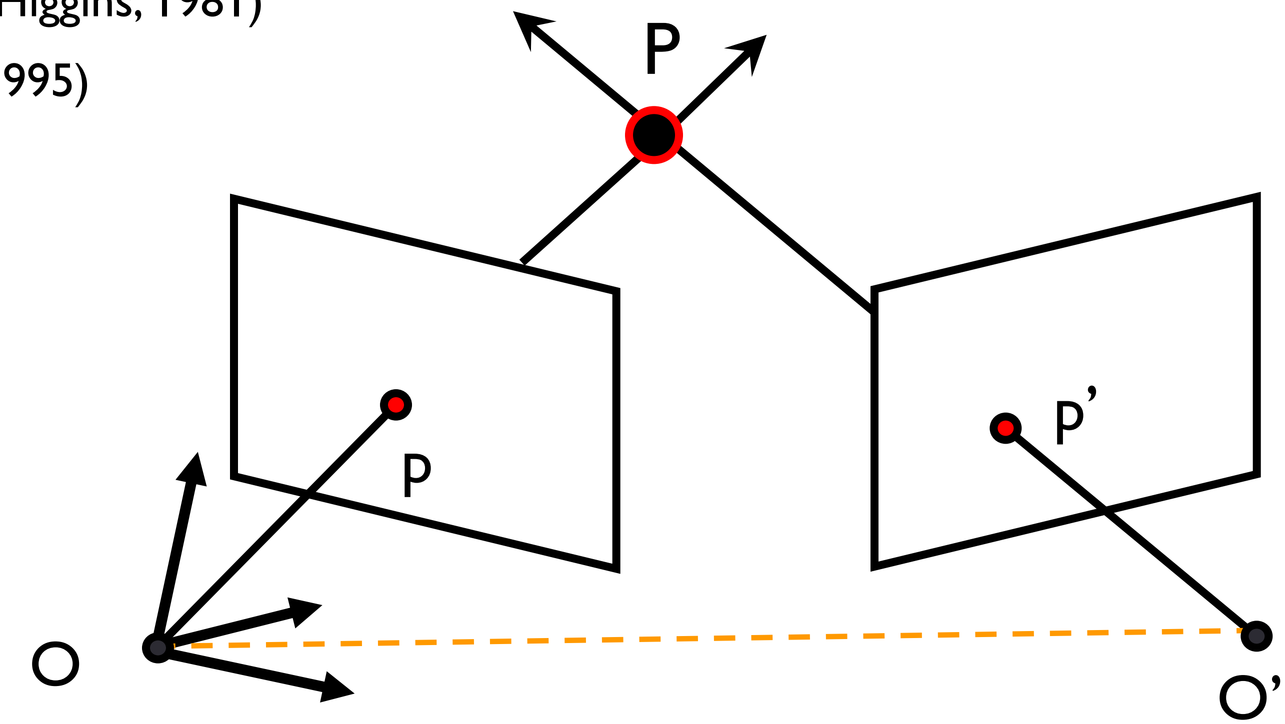
- F captures information about the epipolar geometry of 2 views + camera parameters
- MORE IMPORTANTLY: F gives constraints on how the scene changes under view point transformation (without reconstructing the scene!)
- Powerful tool in:
  - 3D reconstruction
  - Multi-view object/scene matching

Estimating F

The Eight-Point Algorithm

(Longuet-Higgins, 1981)

(Hartley, 1995)



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

Estimating F

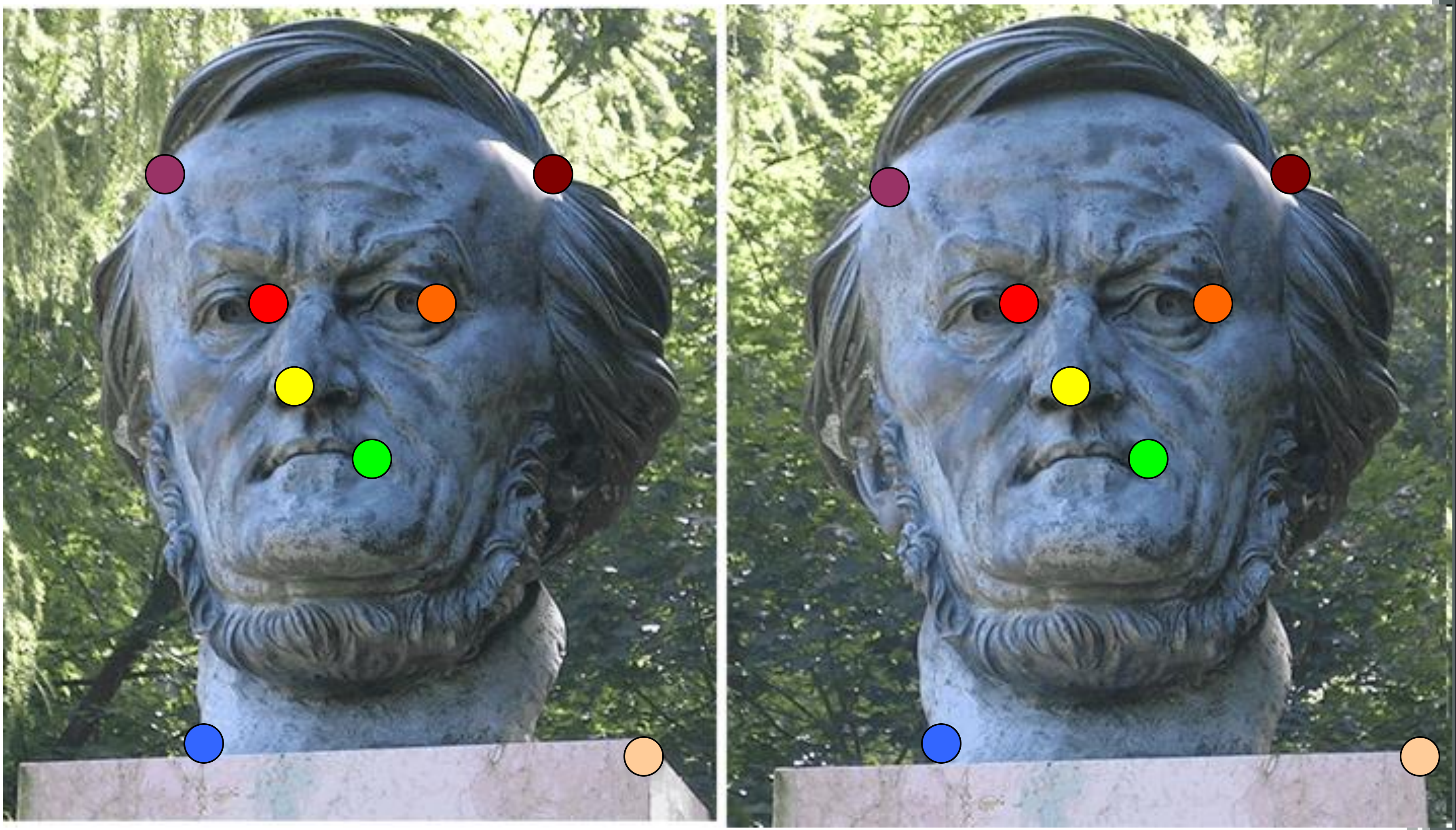
$$p^T \; F \, p' = 0 \quad \Rightarrow$$

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

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

Let's take 8 corresponding points

Estimating F



Estimating F

$$\mathbf{W}\quad \begin{pmatrix} u_1u'_1 & u_1v'_1 & u_1 & v_1u'_1 & v_1v'_1 & v_1 & u'_1 & v'_1 & 1 \\ u_2u'_2 & u_2v'_2 & u_2 & v_2u'_2 & v_2v'_2 & v_2 & u'_2 & v'_2 & 1 \\ u_3u'_3 & u_3v'_3 & u_3 & v_3u'_3 & v_3v'_3 & v_3 & u'_3 & v'_3 & 1 \\ u_4u'_4 & u_4v'_4 & u_4 & v_4u'_4 & v_4v'_4 & v_4 & u'_4 & v'_4 & 1 \\ u_5u'_5 & u_5v'_5 & u_5 & v_5u'_5 & v_5v'_5 & v_5 & u'_5 & v'_5 & 1 \\ u_6u'_6 & u_6v'_6 & u_6 & v_6u'_6 & v_6v'_6 & v_6 & u'_6 & v'_6 & 1 \\ u_7u'_7 & u_7v'_7 & u_7 & v_7u'_7 & v_7v'_7 & v_7 & u'_7 & v'_7 & 1 \\ u_8u'_8 & u_8v'_8 & u_8 & v_8u'_8 & v_8v'_8 & v_8 & u'_8 & v'_8 & 1 \end{pmatrix} \begin{pmatrix} F_{11} \\ F_{12} \\ F_{13} \\ F_{21} \\ F_{22} \\ F_{23} \\ F_{31} \\ F_{32} \\ F_{33} \end{pmatrix} = 0 \quad \mathbf{f}$$

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

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

Rank-2 constraint

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

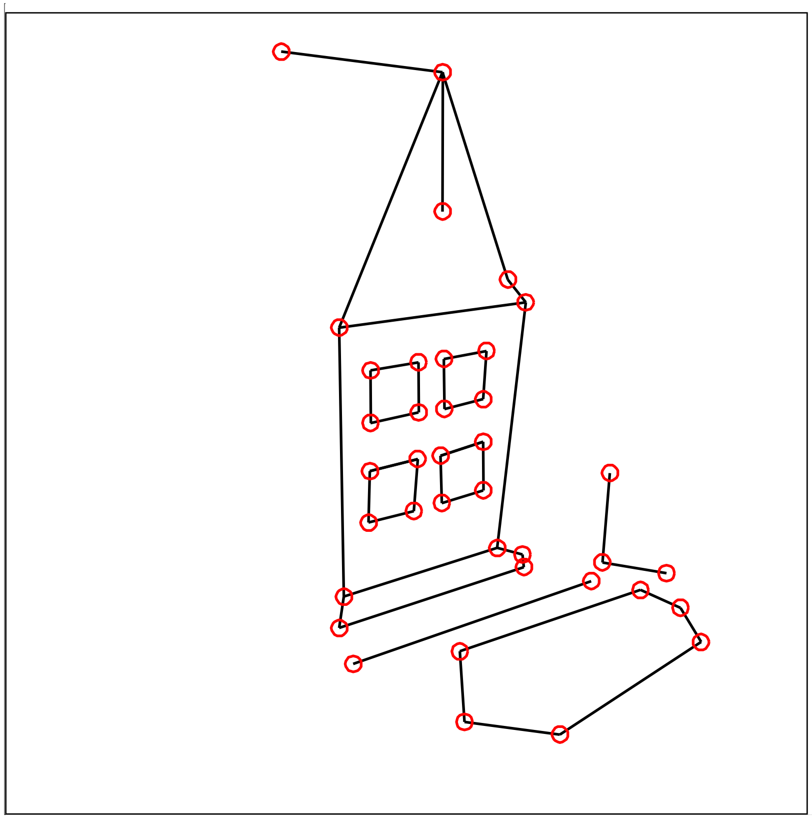
The estimated $\hat{F}$ may have full rank ($\det(\hat{F}) \neq 0$)
(F should have rank=2 instead)

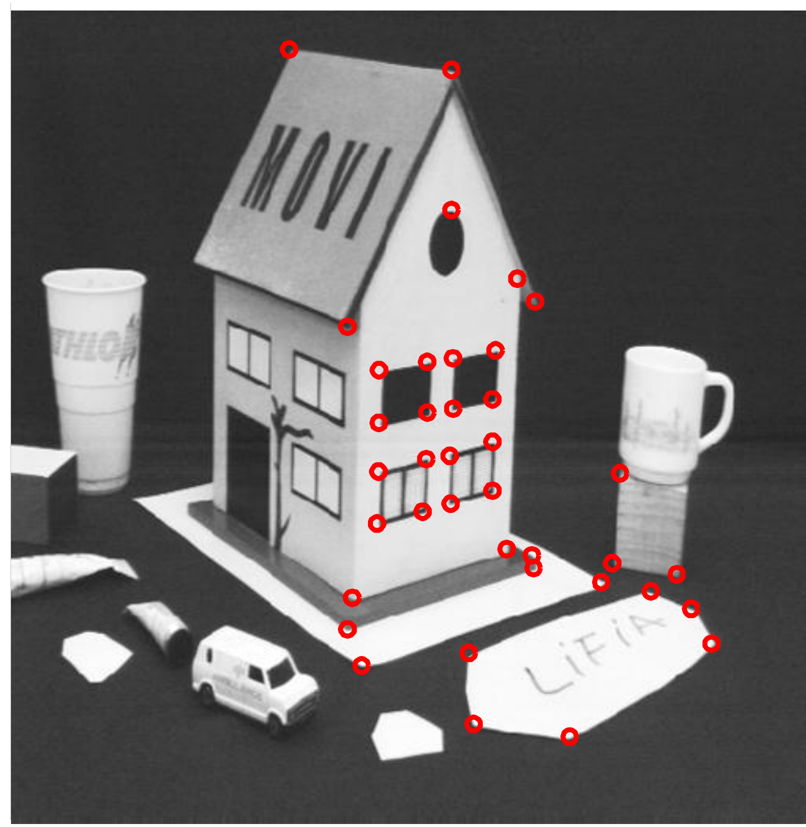
Find F that minimizes $\left\| F - \hat{F} \right\| = 0$

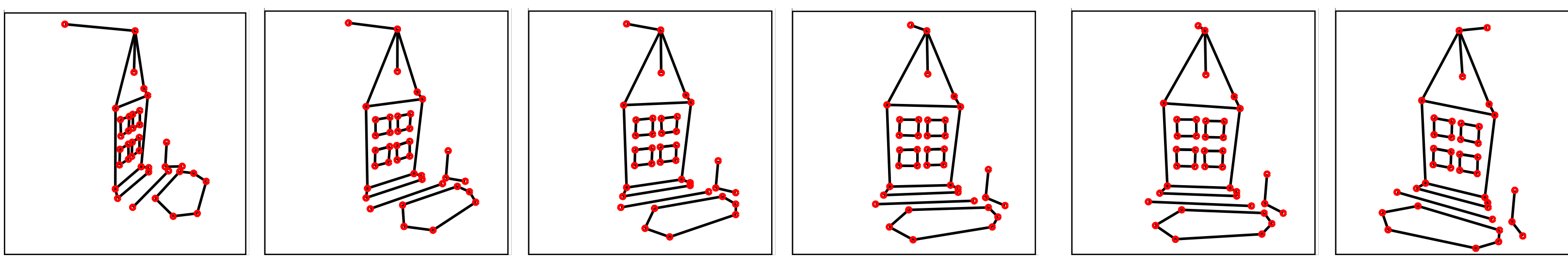
Frobenius norm

Subject to det(F)=0

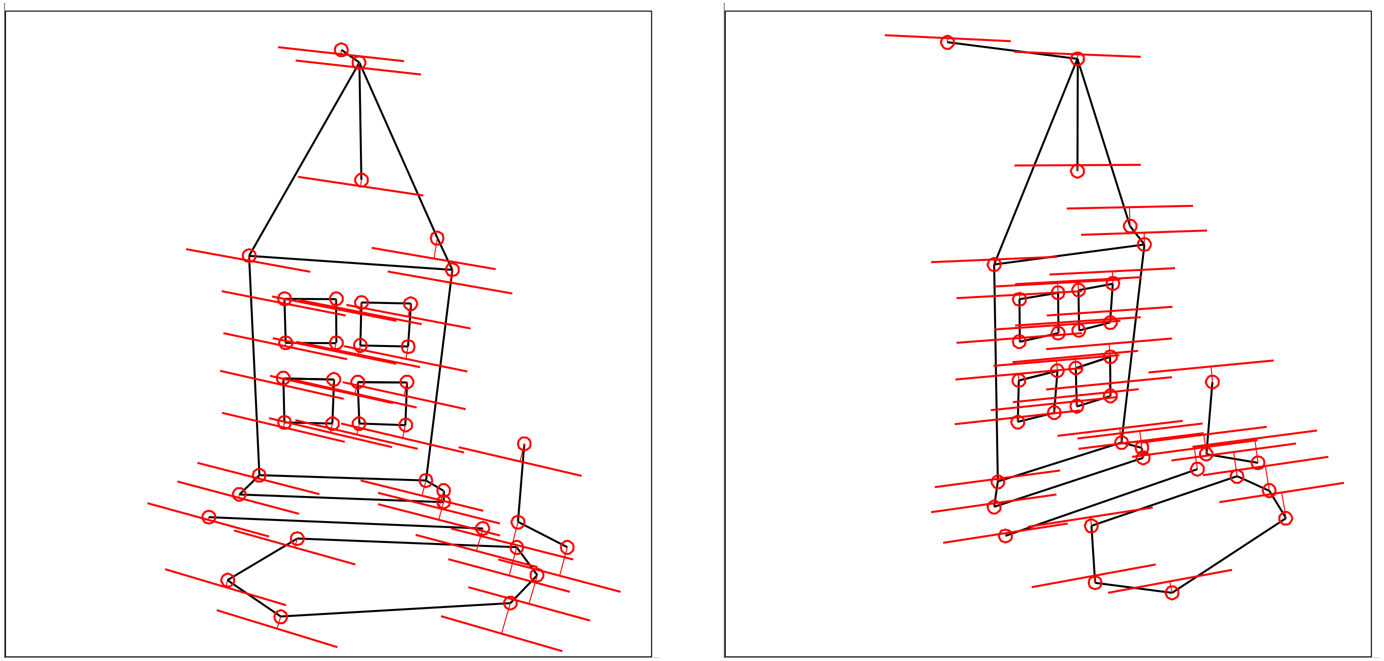
SVD (again!) can be used to solve this problem





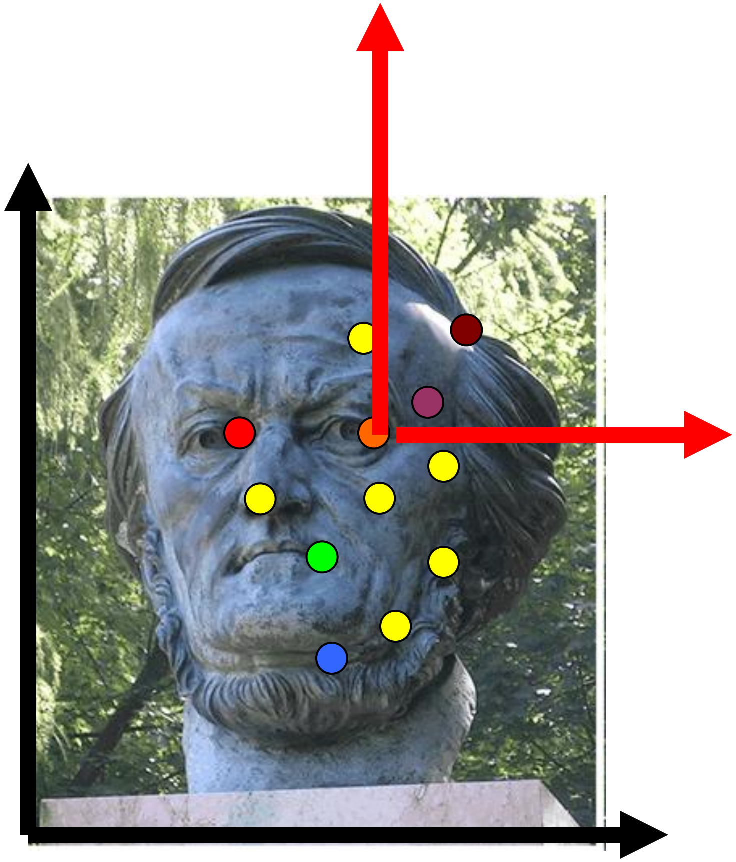


Data courtesy of R. Mohr and B. Boufama.



Mean errors:
10.0pixel
9.1pixel

Normalization



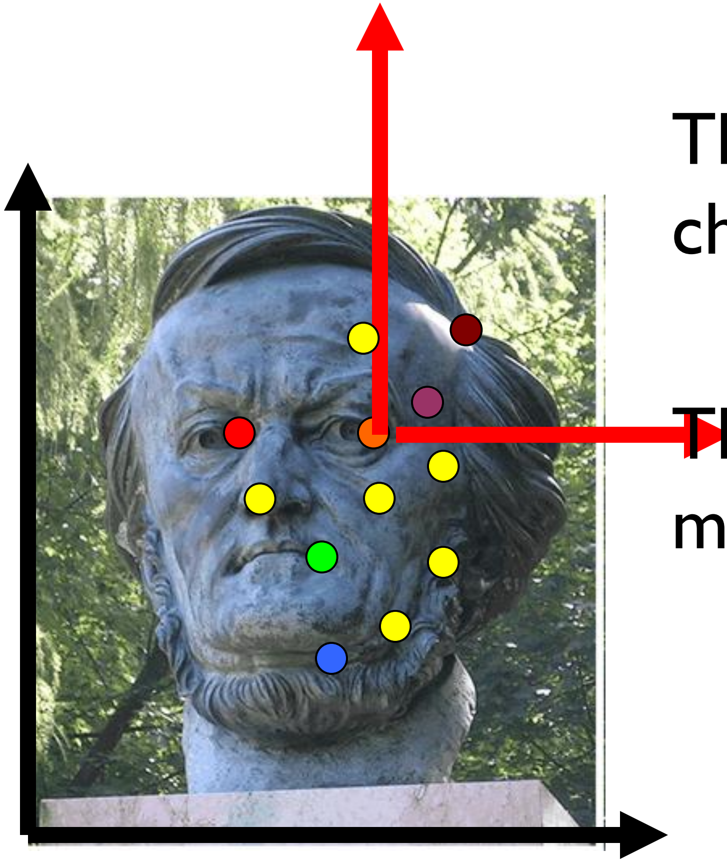
Is the accuracy in estimating F function of the ref. system in the image plane?

E.g. under similarity transformation (T = scale + translation):

$$q_i = T_i \; p_i \qquad q'_i = T'_i \; p'_i$$

Does the accuracy in estimating F change if a transformation T is applied?

Normalization



The accuracy in estimating F does change if a transformation T is applied

There exists a T for which accuracy is maximized

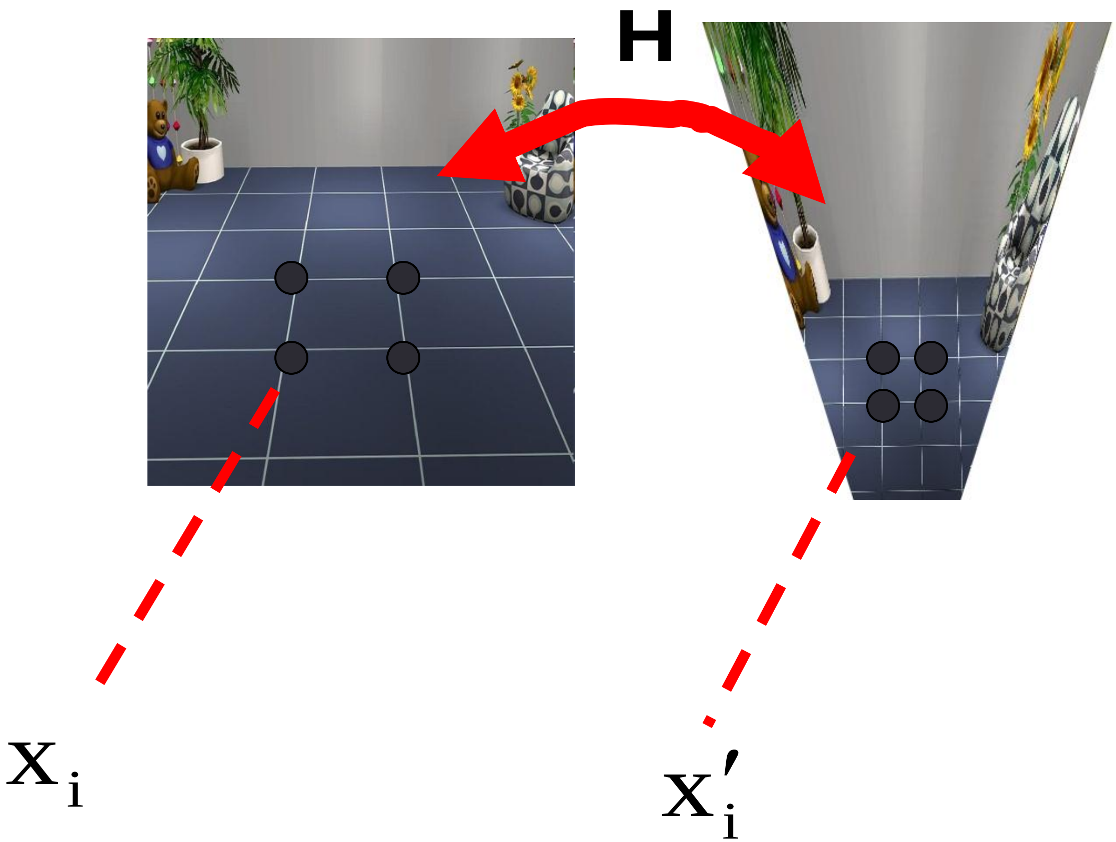
$$
\mathbf{W} = \begin{pmatrix}
u_1u_1' & u_1v_1' & u_1 & v_1u_1' & v_1v_1' & v_1 & u_1' & v_1' & 1 \\
u_2u_2' & u_2v_2' & u_2 & v_2u_2' & v_2v_2' & v_2 & u_2' & v_2' & 1 \\
u_3u_3' & u_3v_3' & u_3 & v_3u_3' & v_3v_3' & v_3 & u_3' & v_3' & 1 \\
u_4u_4' & u_4v_4' & u_4 & v_4u_4' & v_4v_4' & v_4 & u_4'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

- Homogeneous system $\mathbf{W}\,\mathbf{f} = 0$
- If N>8 → Lsq. solution by SVD! → $\hat{\mathrm{F}}$

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

- uv's ~ 100000
- Last column of ones -> numerical values are unbalanced;
- Small errors in measurements (u, v) are amplified

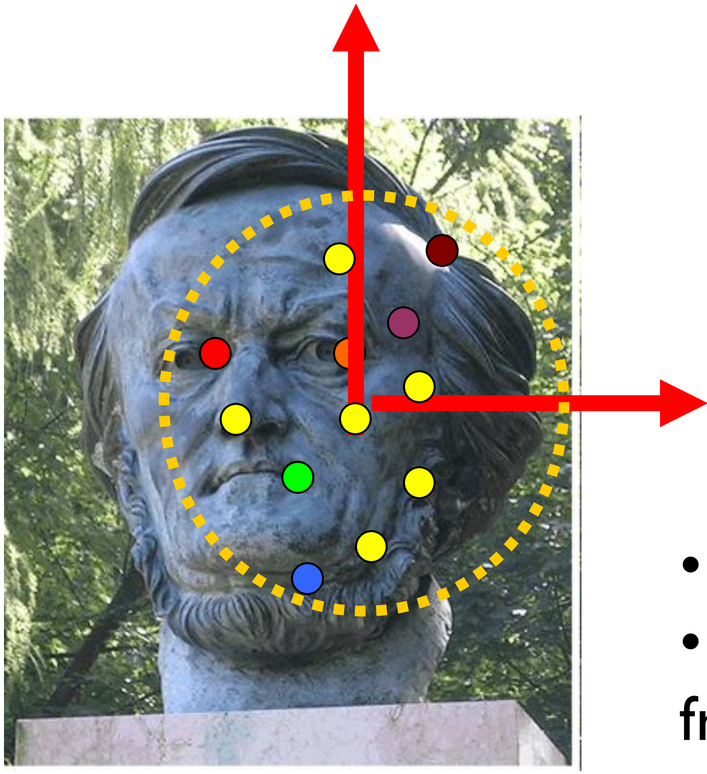
Same issue for the DLT algorithm



$$x'_i = H x_i$$

[Section 4.4 in AZ]

Normalization



Transform image coordinate system (T = translation+scaling) such that:

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

$$q_i = T_i \, p_i \qquad q'_i = T'_i \, p'_i \qquad \text{(normalization)}$$

The Normalized Eight-Point Algorithm

0. Compute T_i and T_i'

1. Normalize coordinates:

$$q_i = T_i \; p_i \qquad\qquad q'_i = T'_i \, p'_i$$

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

1. Enforce the rank-2 constraint.

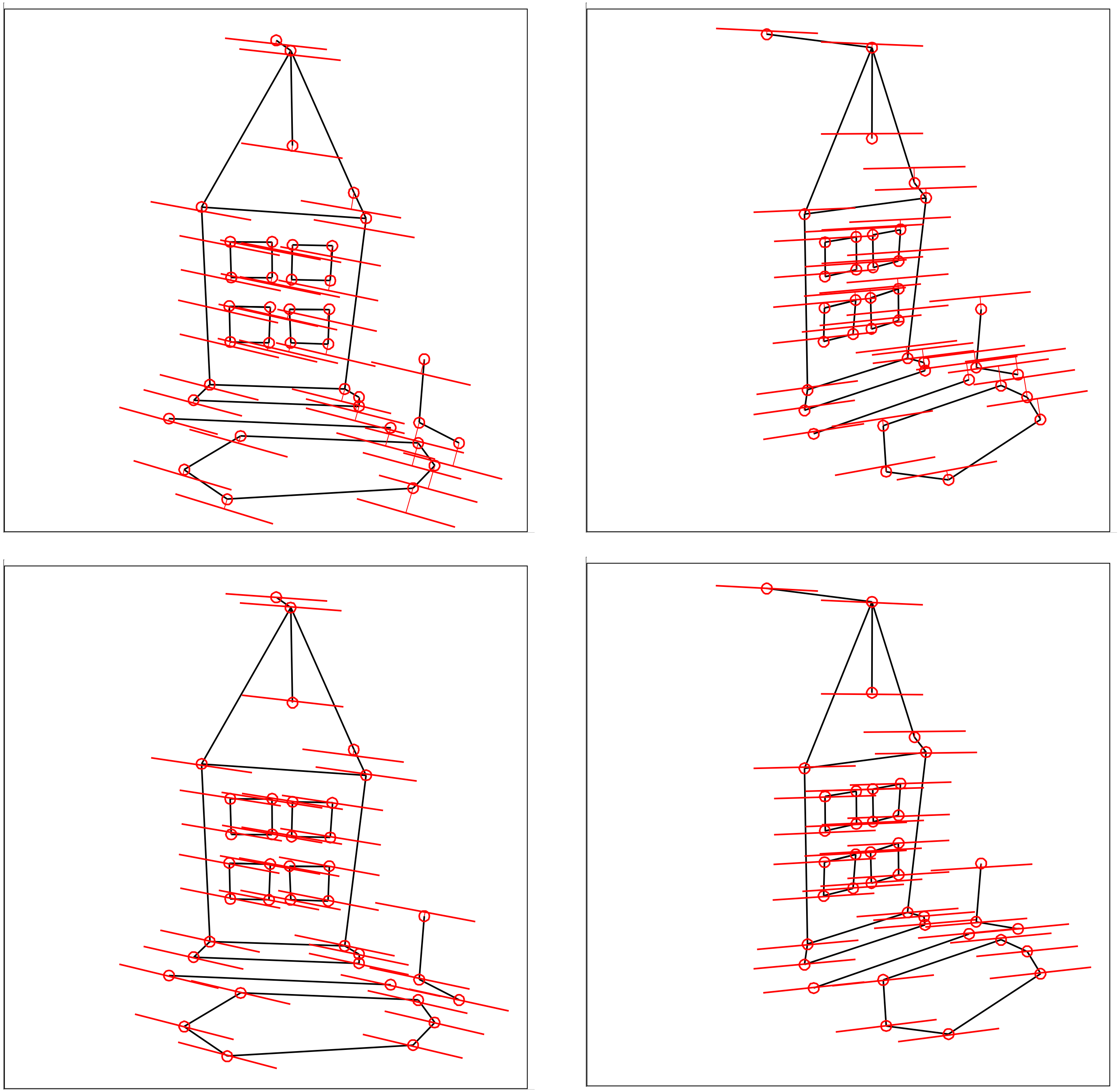
$$\rightarrow F_q \quad \begin{cases} q^T \; F_q \; q' = 0 \\ \det(F_q) = 0 \end{cases}$$

2. De-normalize F_q:

$$F = T'^T F_q T$$

Without transformation

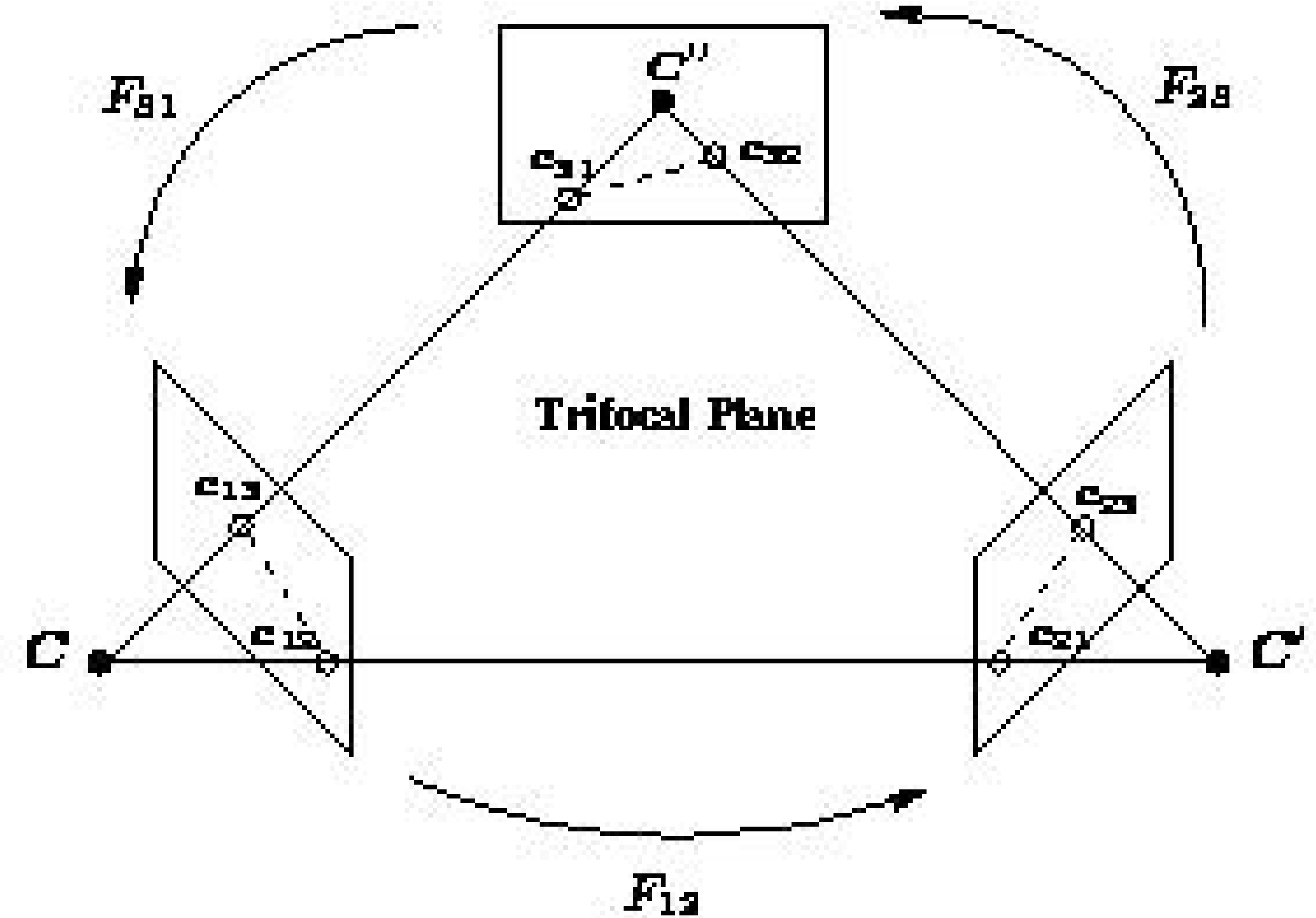
With transformation



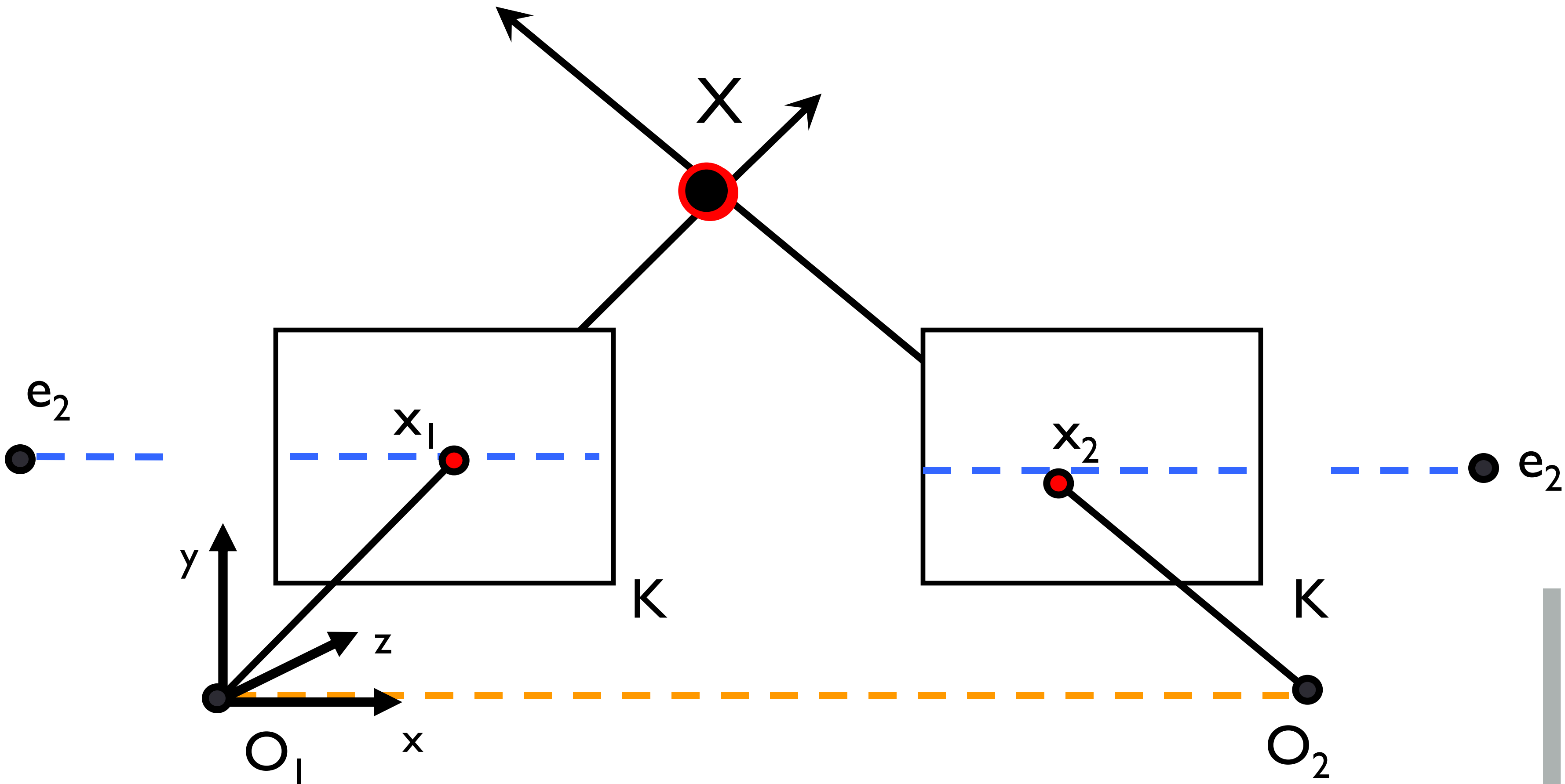
Mean errors:
10.0pixel
9.1pixel

Mean errors:
1.0pixel
0.9pixel

Trifocal tensor T



Example: Parallel image planes



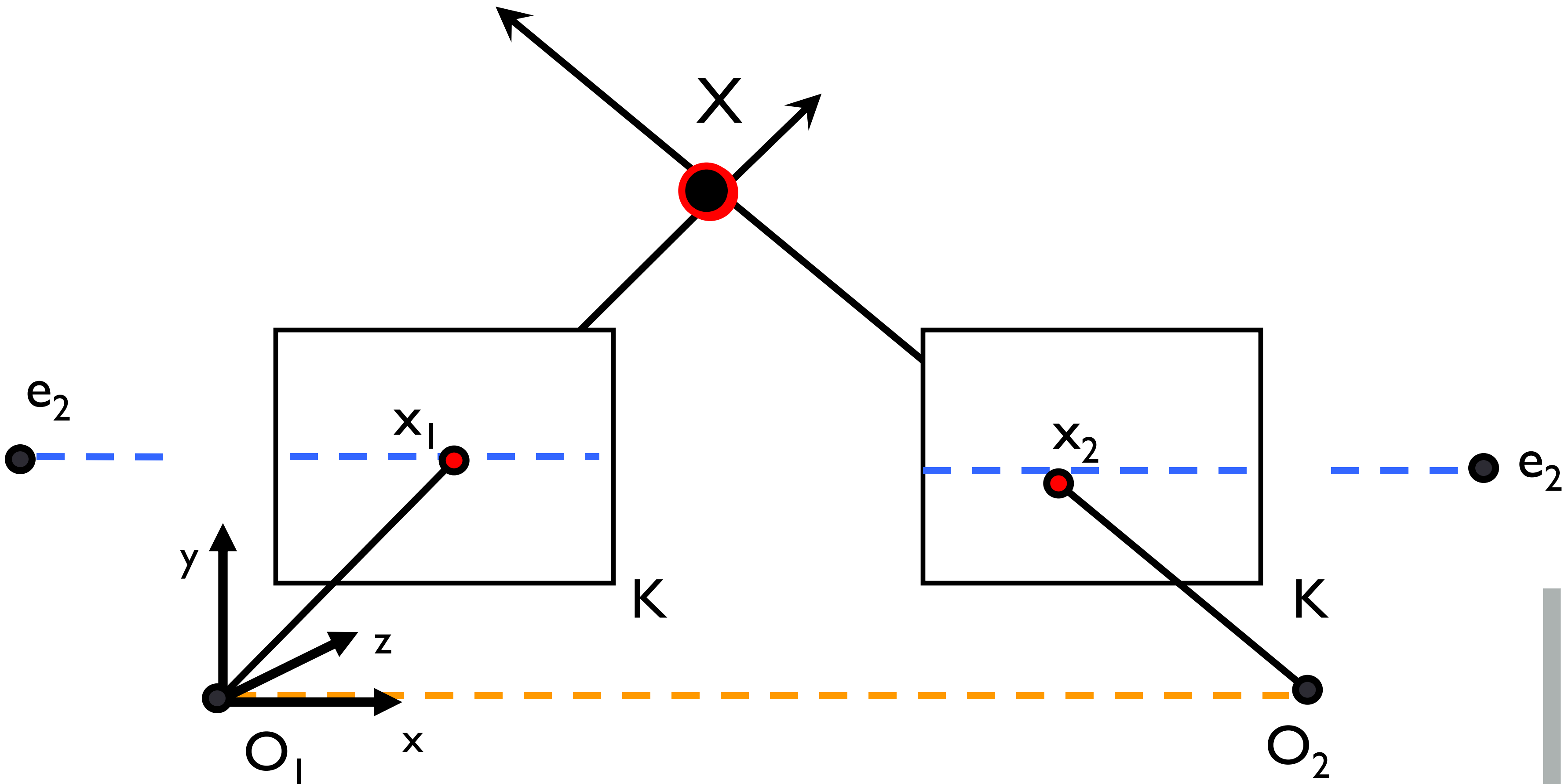
$K_1 = K_2$ = known

x parallel to O_1O_2

E=?

$$E = [t_\times]R = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -T \\ 0 & T & 0 \end{bmatrix}$$

Example: Parallel image planes

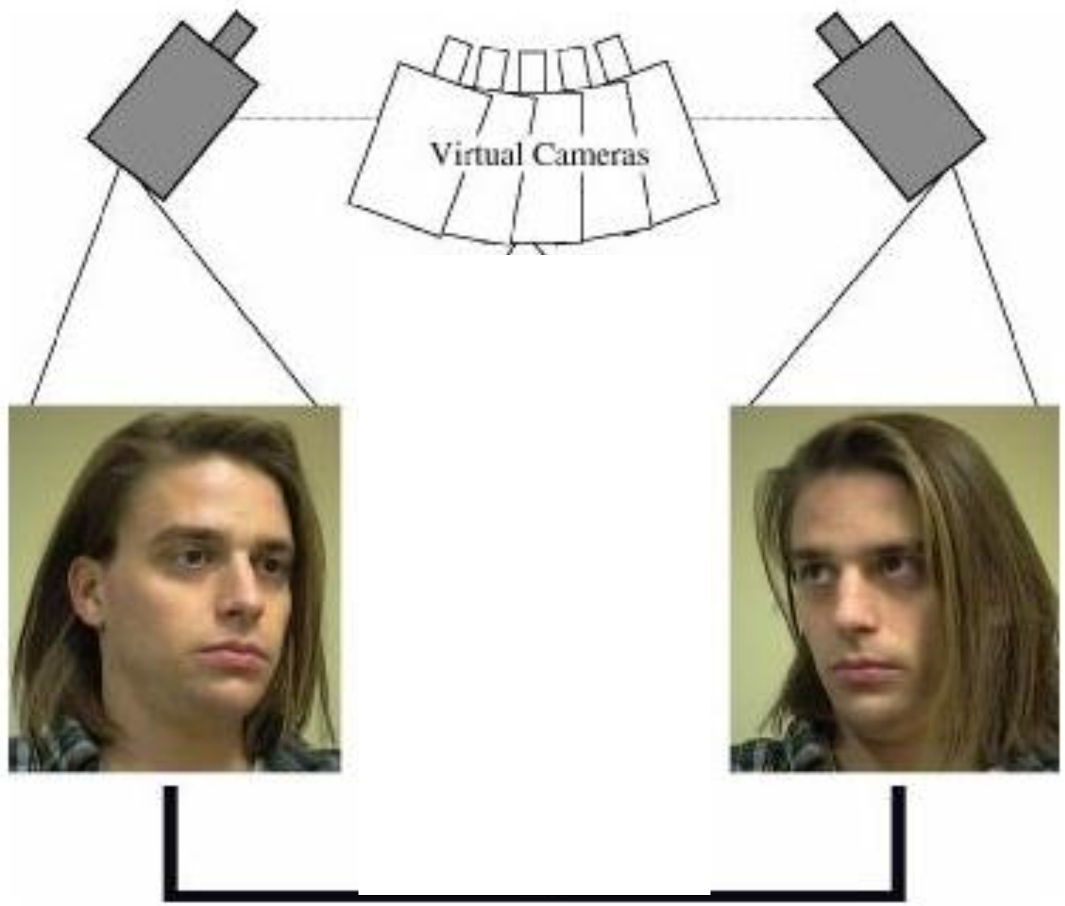


Rectification: making two images "parallel"

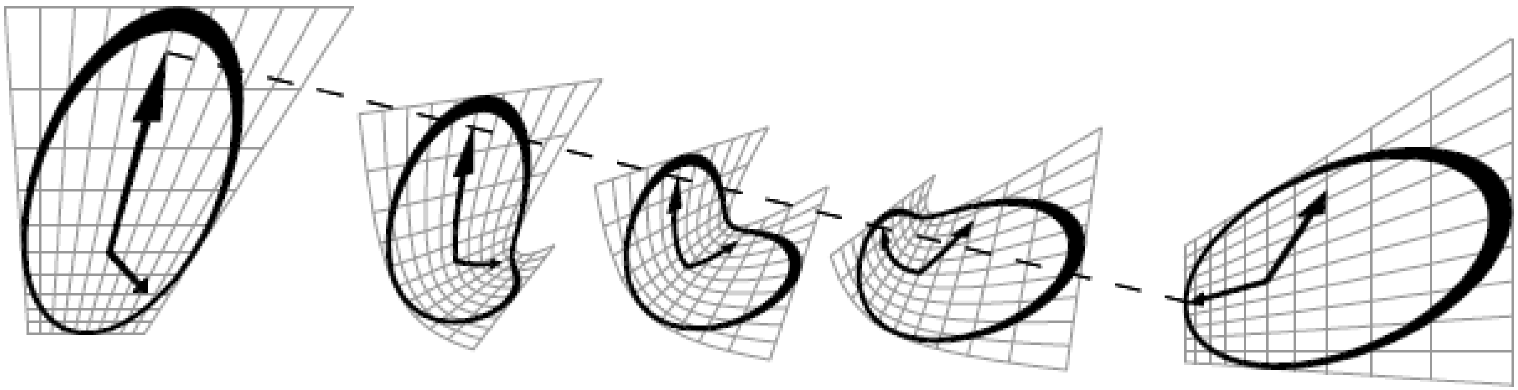
Why it is useful? Epipolar constraint → y = y'

Application: view morphing

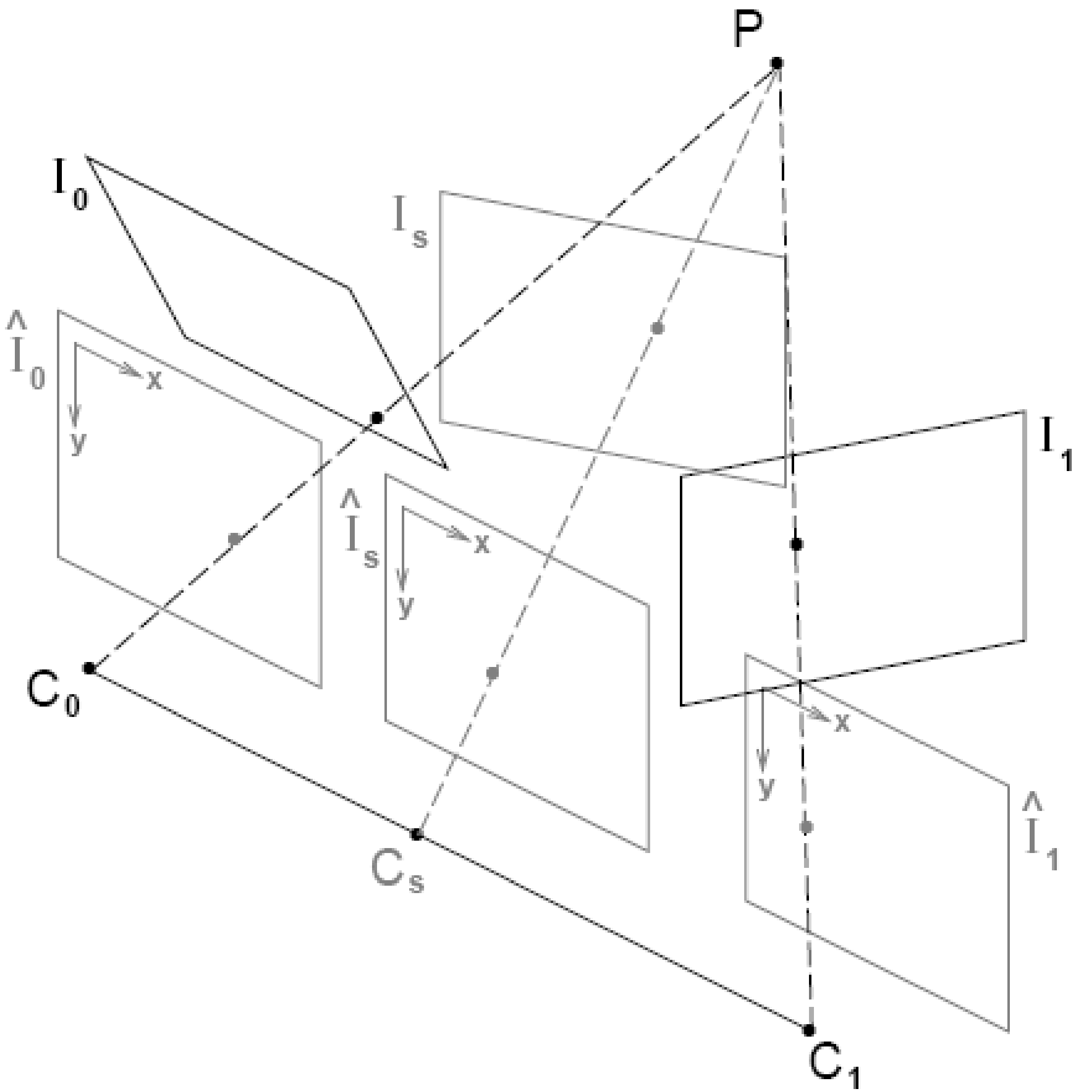
S. M. Seitz and C. R. Dyer, *Proc. SIGGRAPH 96*, 1996, 21-30

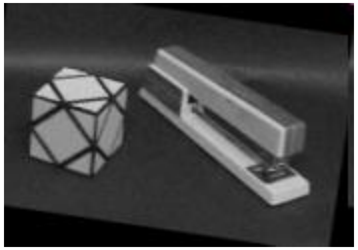


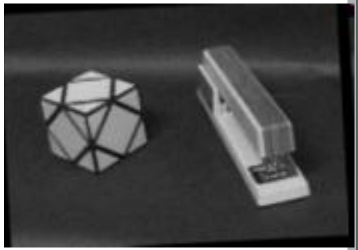
Morphing without using geometry

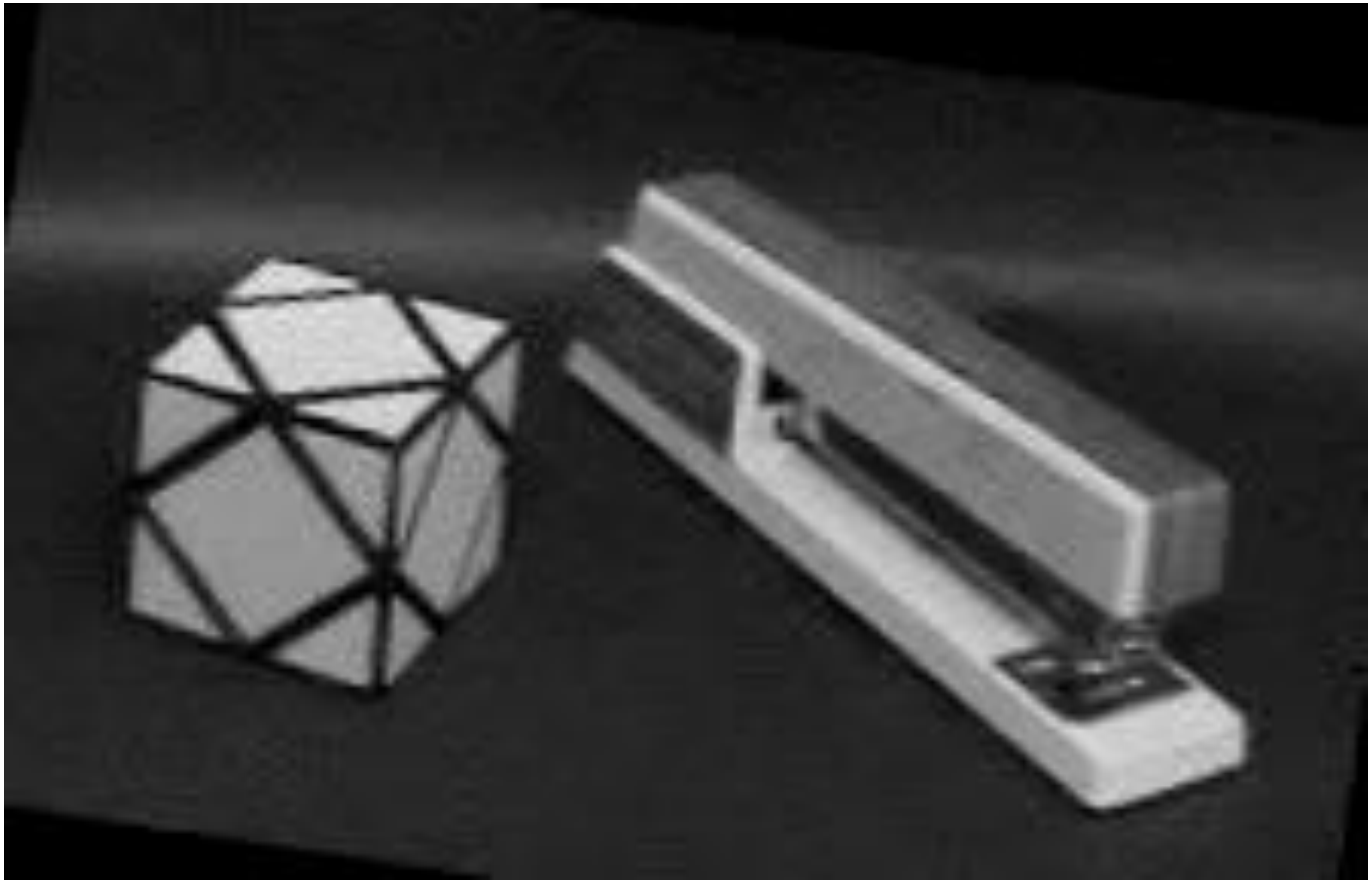


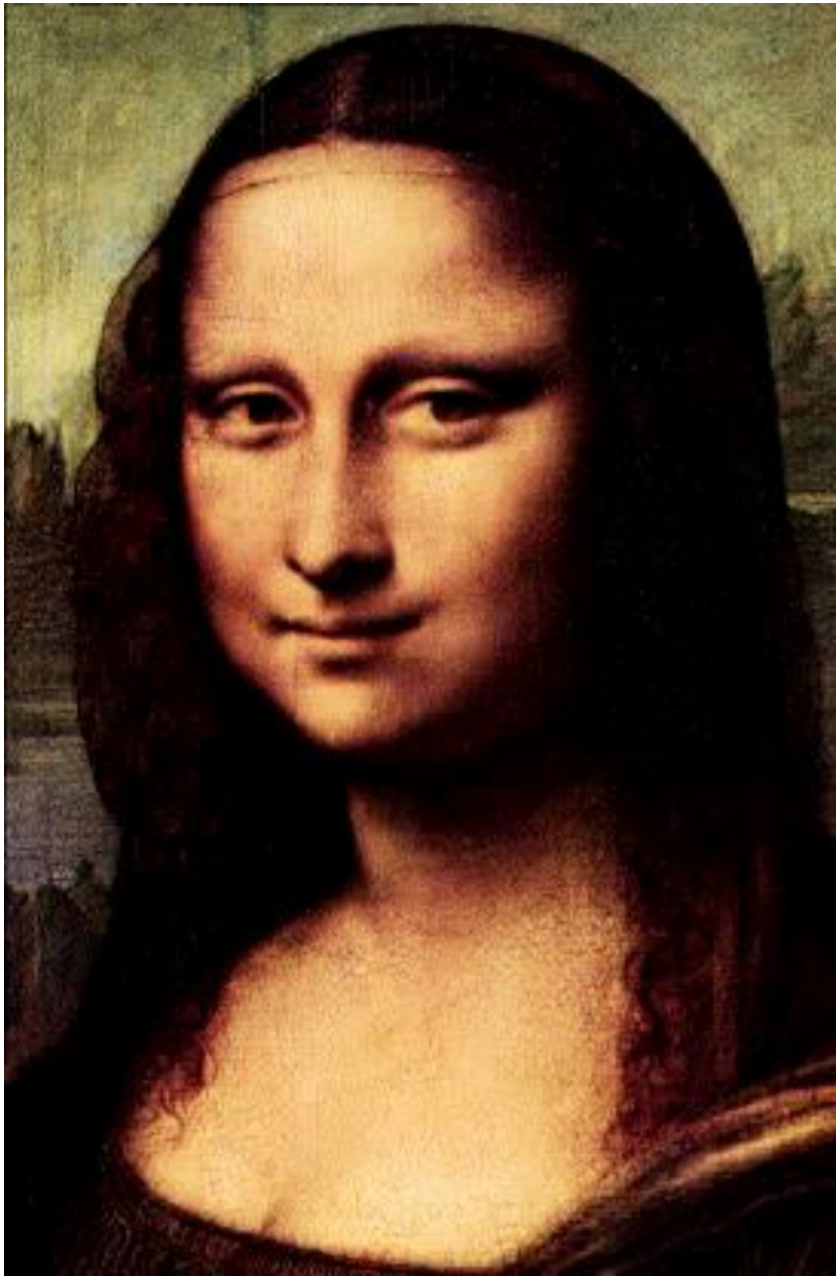
Rectification











From its reflection!

The Fundamental Matrix Song

http://danielwedge.com/fmatrix/

Next Lecture

- Reconstruction using stereo systems