Lecture 9: Epipolar Geometry

Professor Fei-Fei Li

Stanford Vision Lab

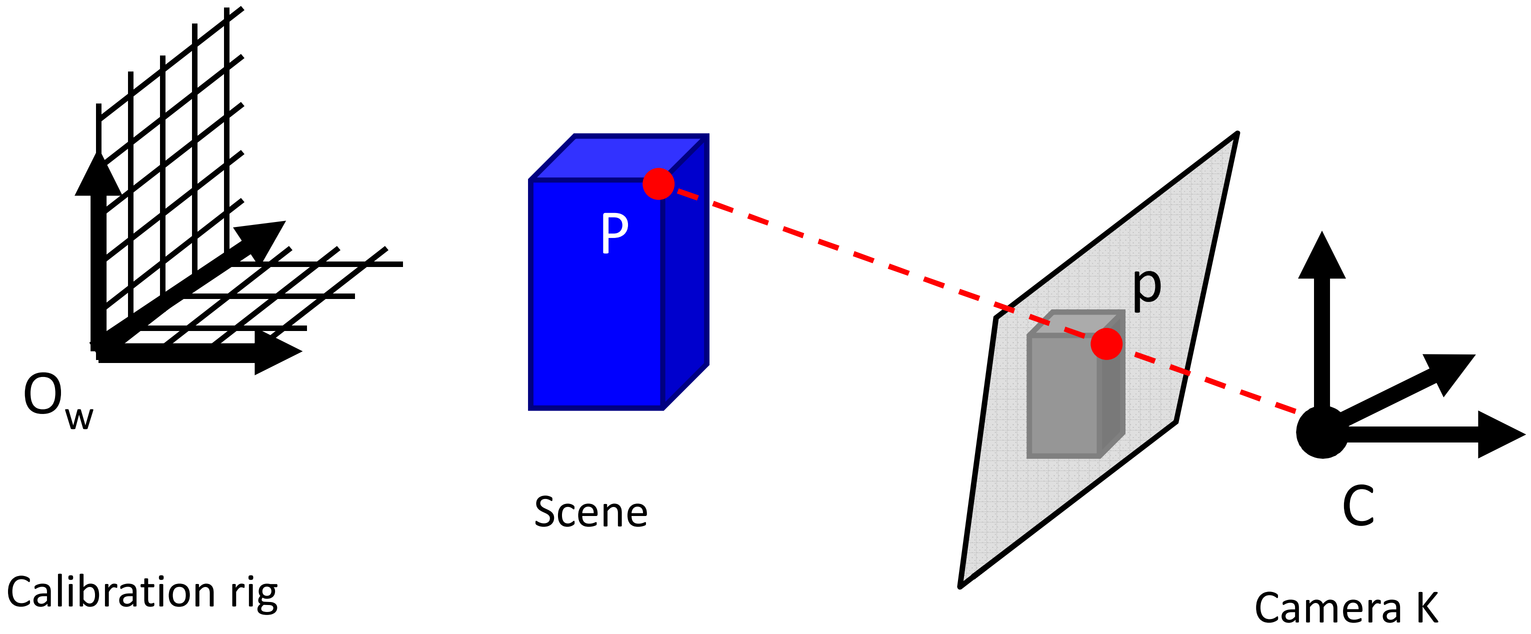
What we will learn today?

- Why is stereo useful?
- Epipolar constraints
- Essential and fundamental matrix
- Estimating F (Problem Set 2 (Q2))
- Rectification

10 : 20am

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

Recovering structure from a single view



Why is it so difficult?

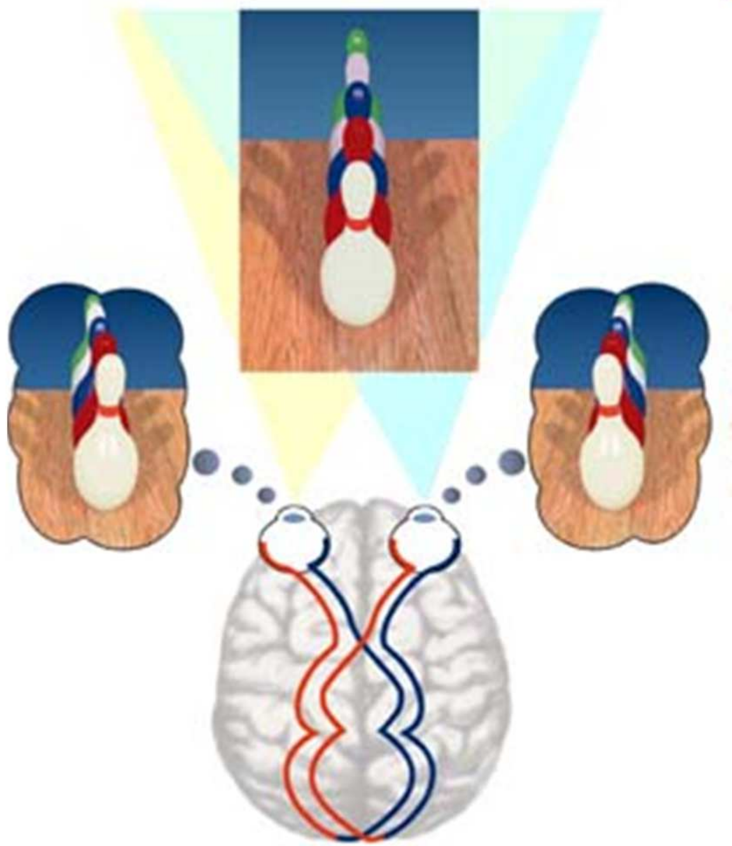
Intrinsic ambiguity of the mapping from 3D to image (2D)

Recovering structure from a single view

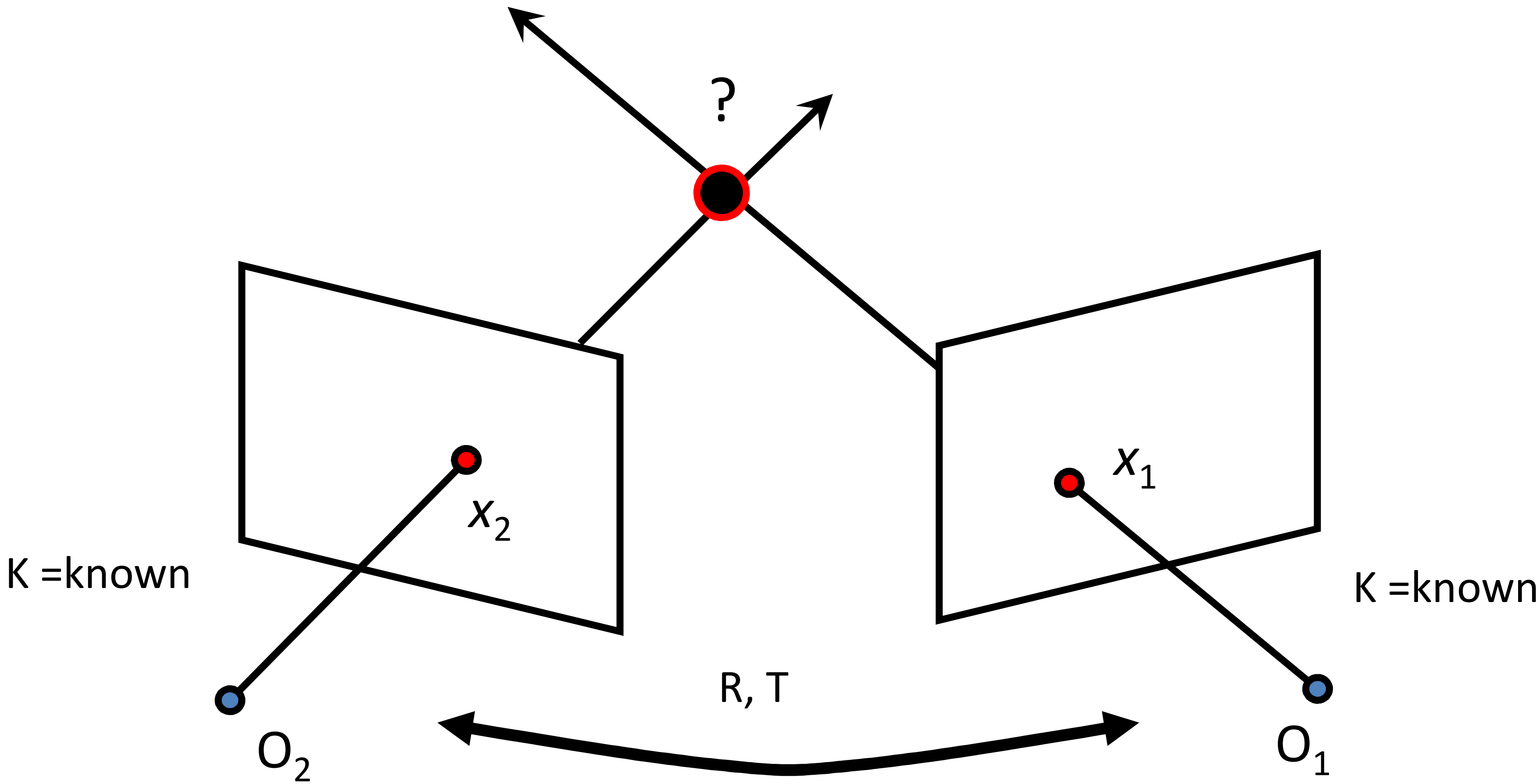
Courtesy slide S. Lazebnik

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

Two eyes help!



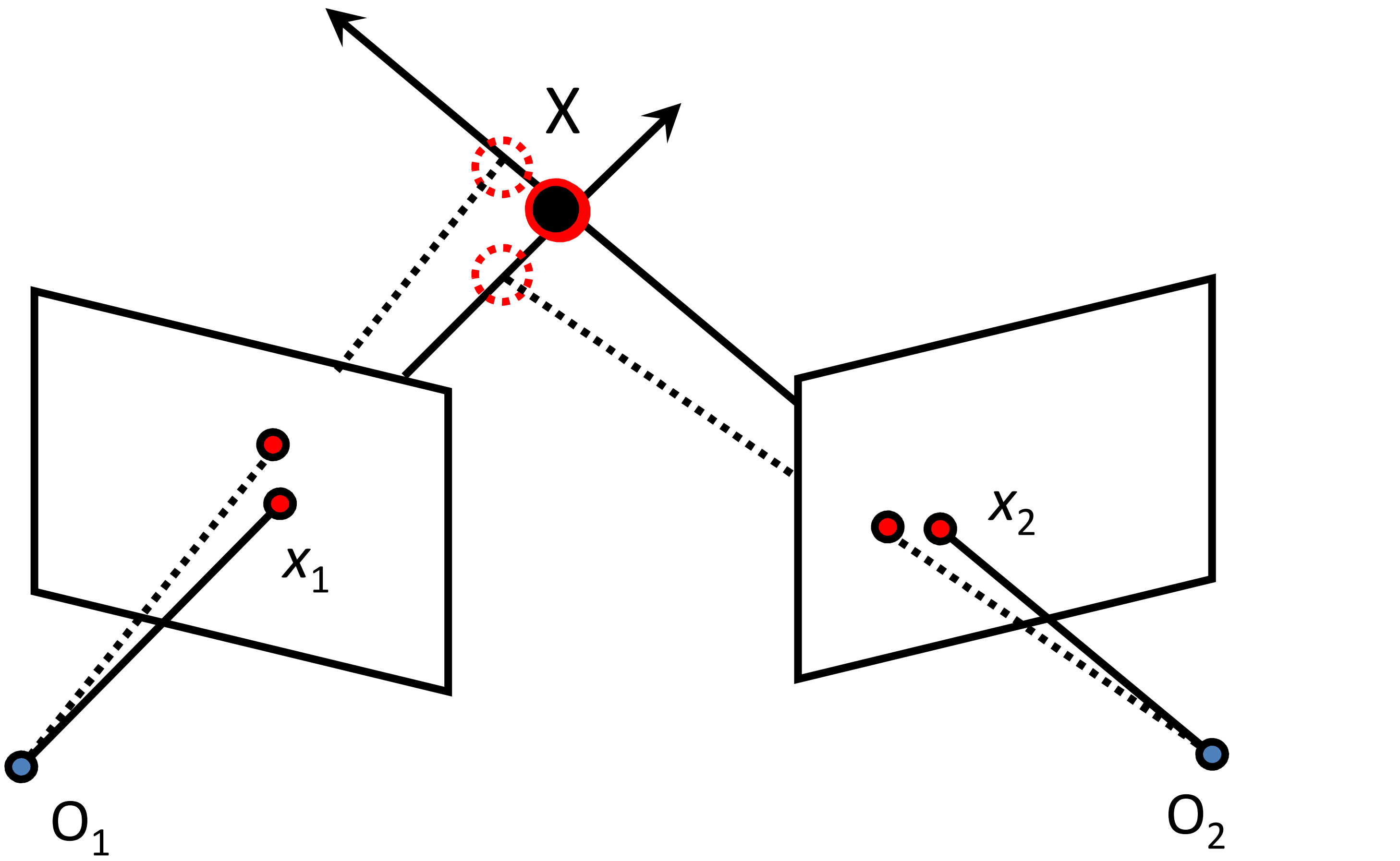
Two eyes help!



This is called triangulation

Triangulation

- Find X that minimizes $d^2(x_1, P_1 X) + d^2(x_2, P_2 X)$



Stereo-view geometry

- **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.

This lecture (#9)

Stereo-view geometry

Next lecture (#10)

- **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.

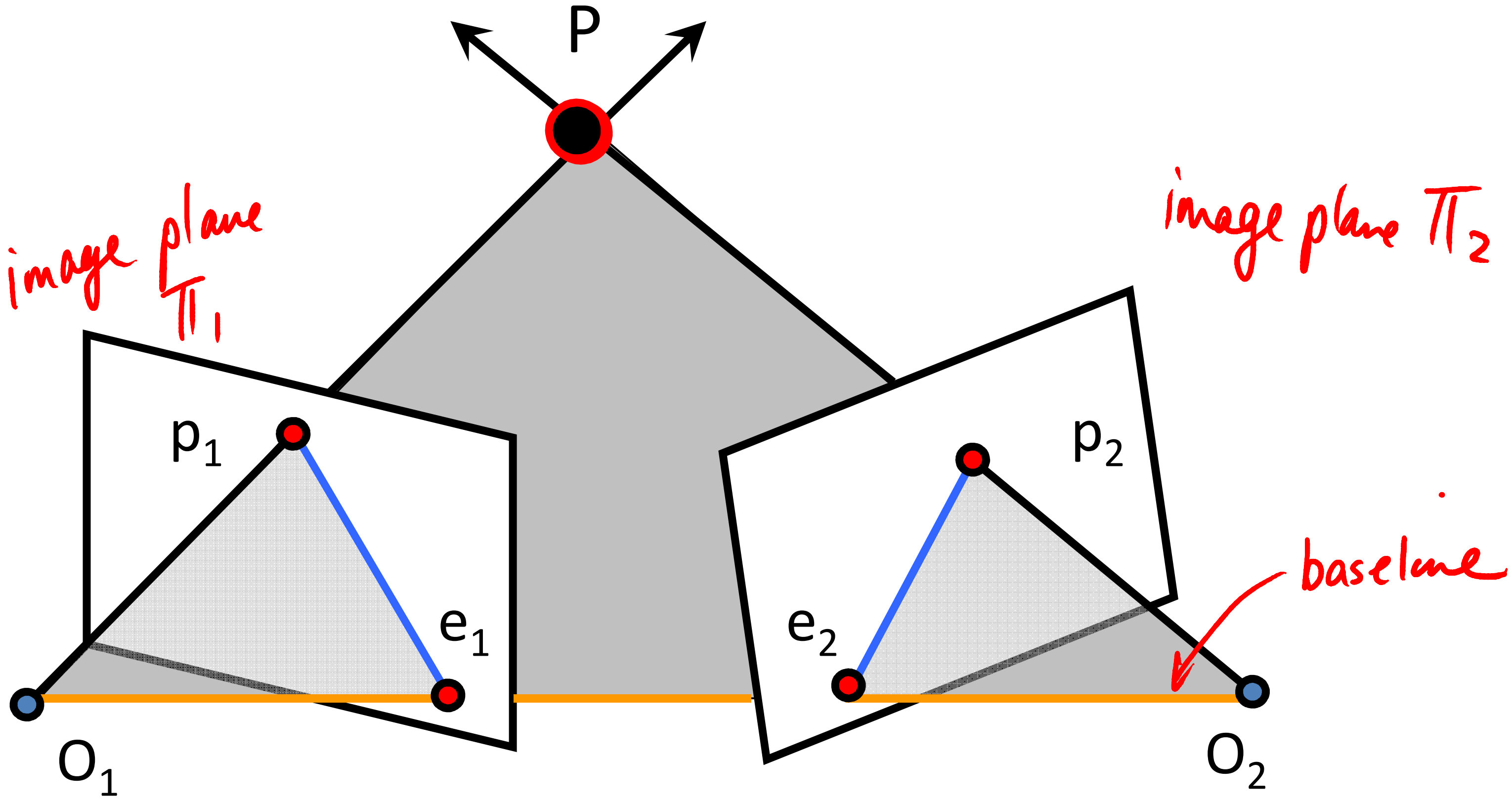
- **Scene geometry:** Find coordinates of 3D point from its projection into 2 or multiple images.

What we will learn today?

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

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

Epipolar geometry

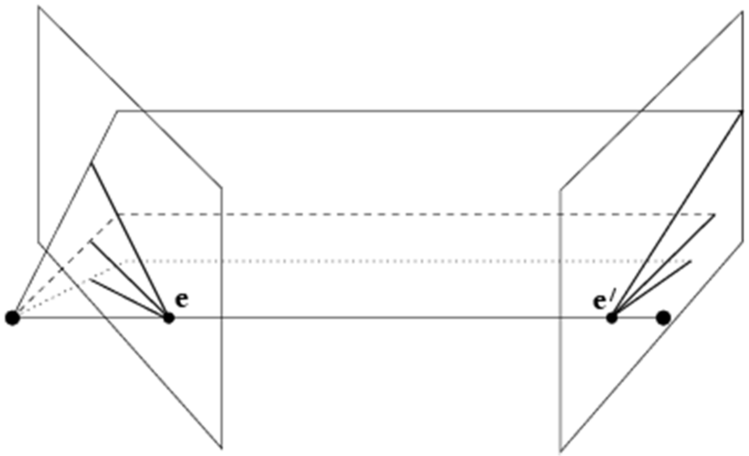


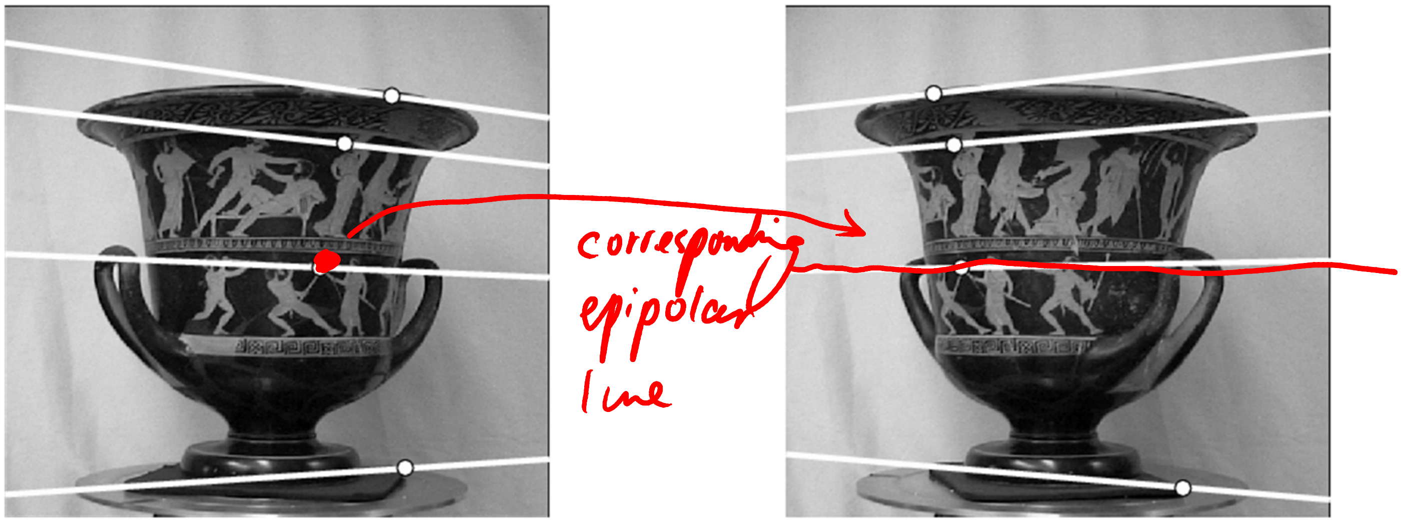
- Epipolar Plane
- Baseline
- Epipolar Lines

- Epipoles e_1, e_2

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

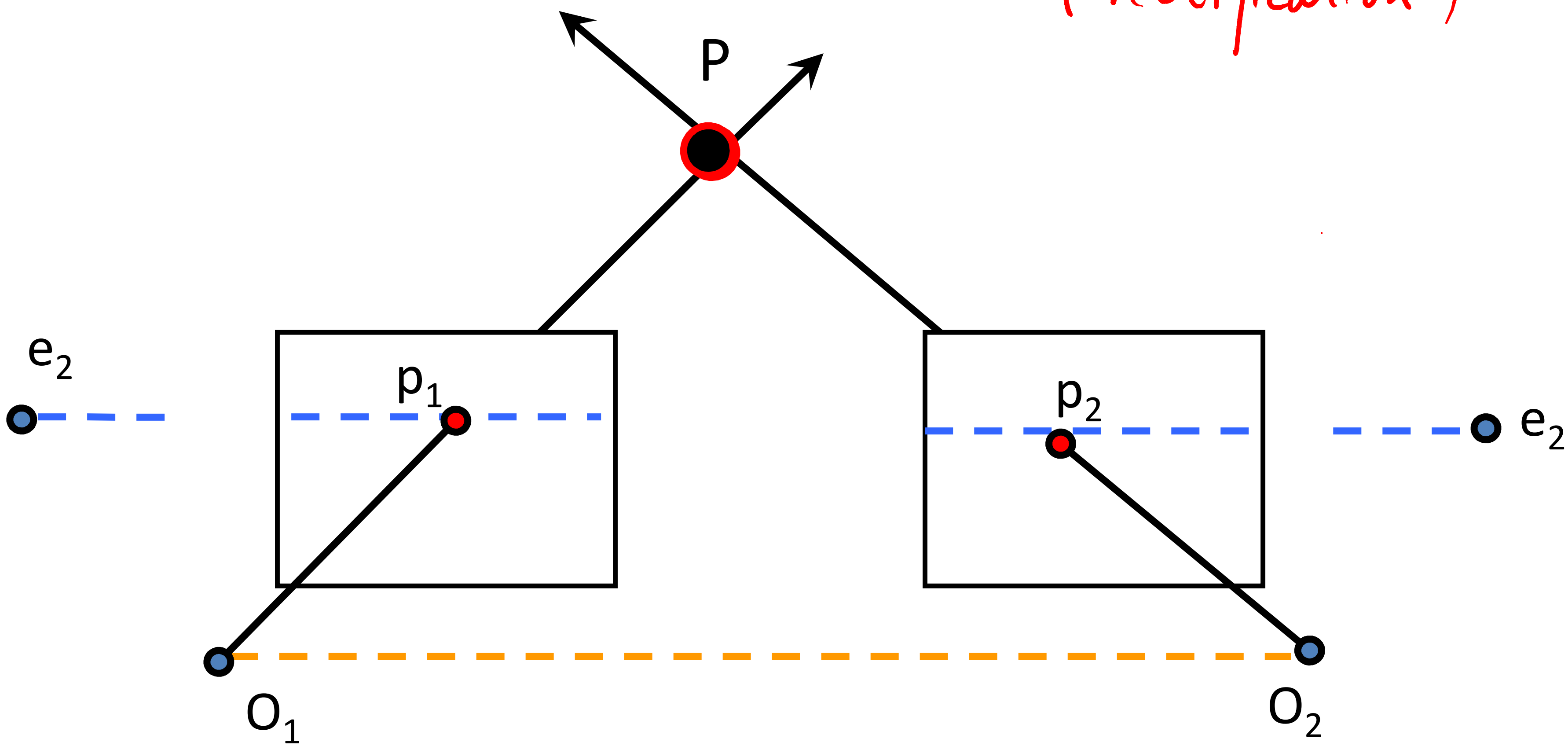
Example: Converging image planes





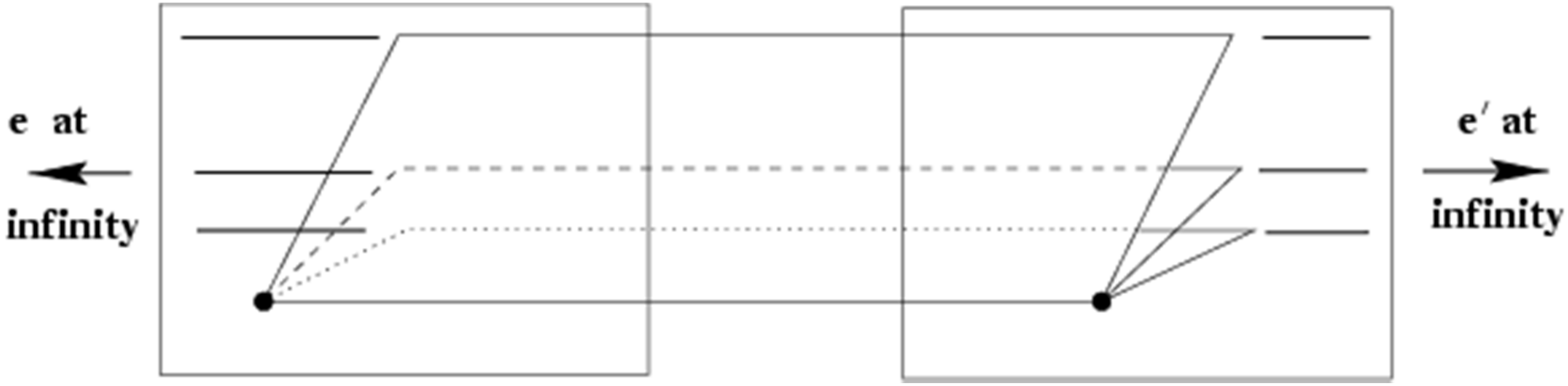
Example: Parallel image planes

("rectification")

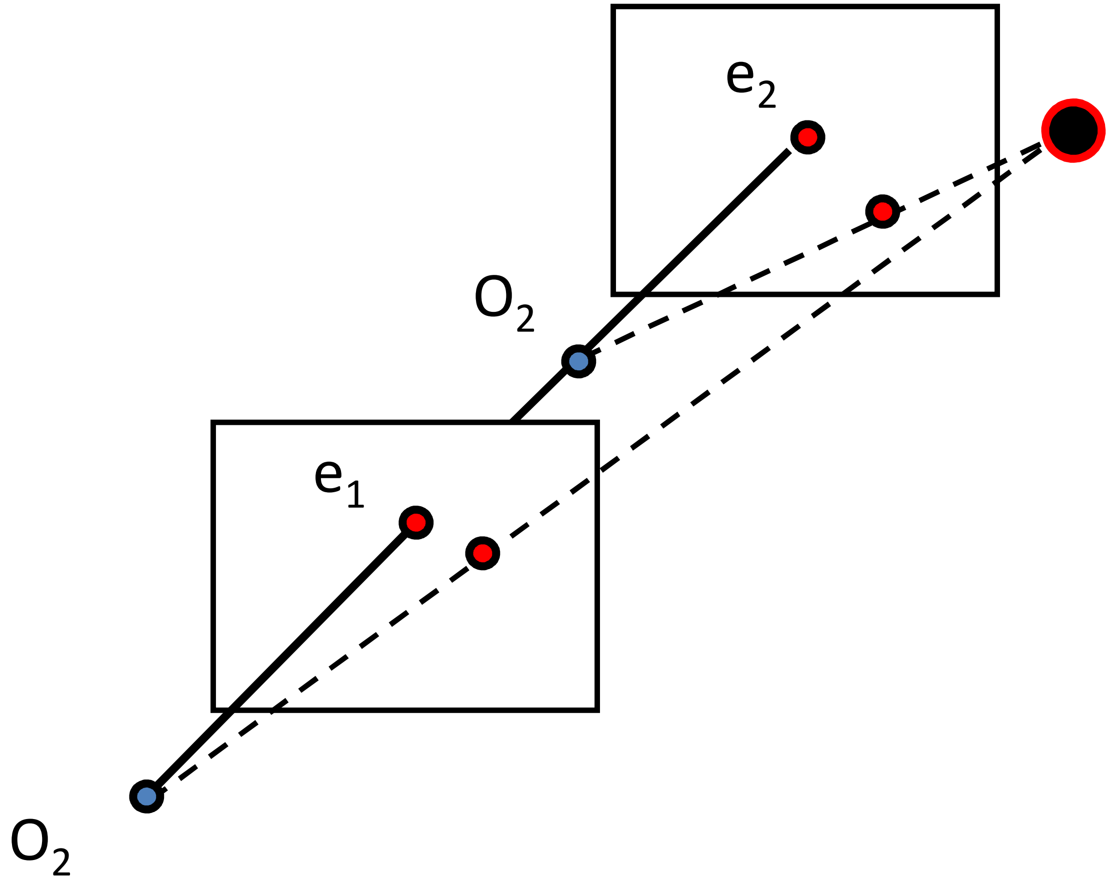


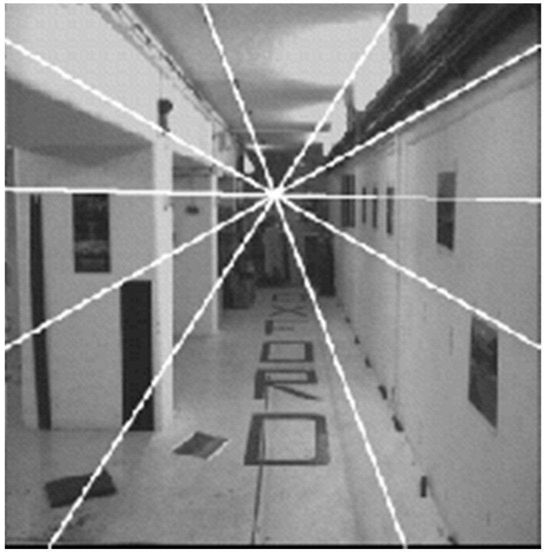
- Baseline intersects the image plane at infinity
- Epipoles are at infinity
- Epipolar lines are parallel to x axis

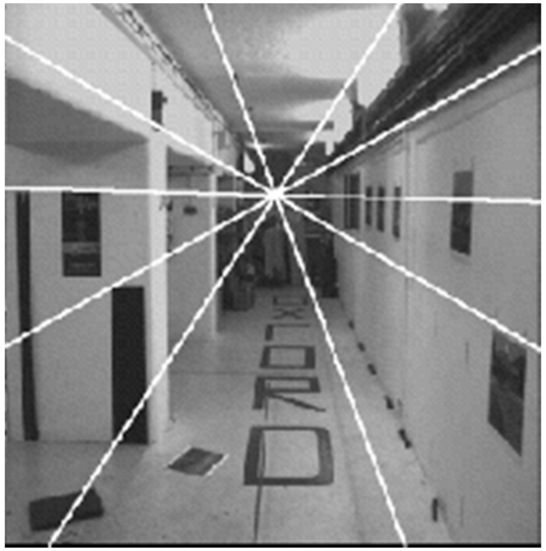
Example: Parallel image planes



Example: Forward translation

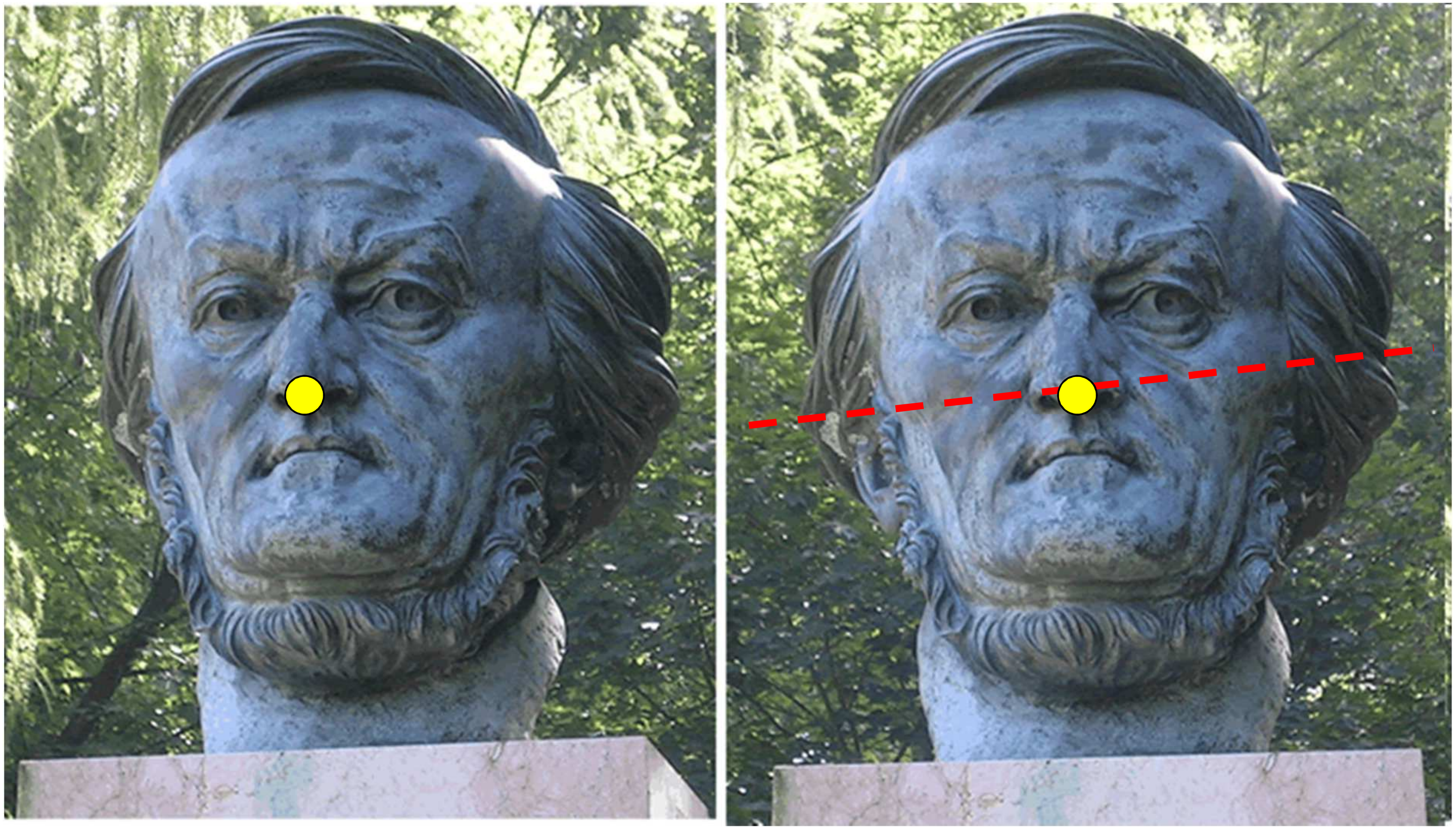






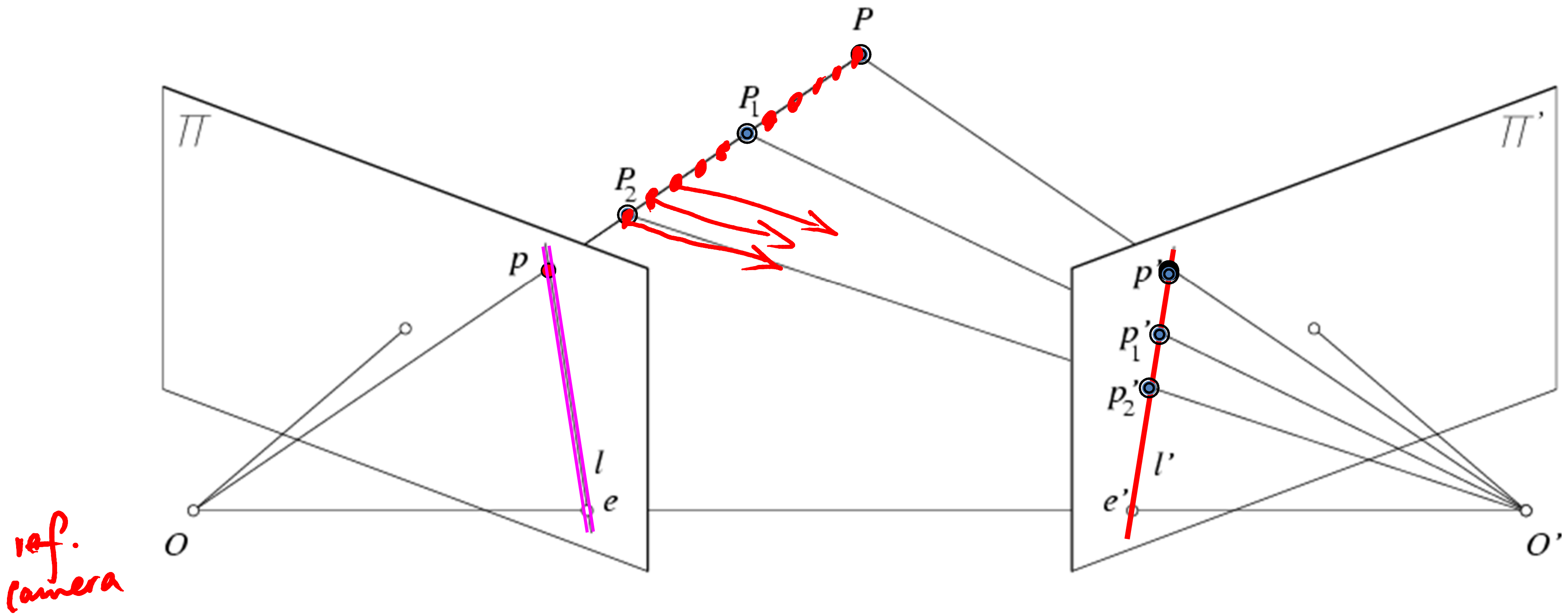
- The epipoles have same position in both images
- Epipole called FOE (focus of expansion)

Epipolar Constraint



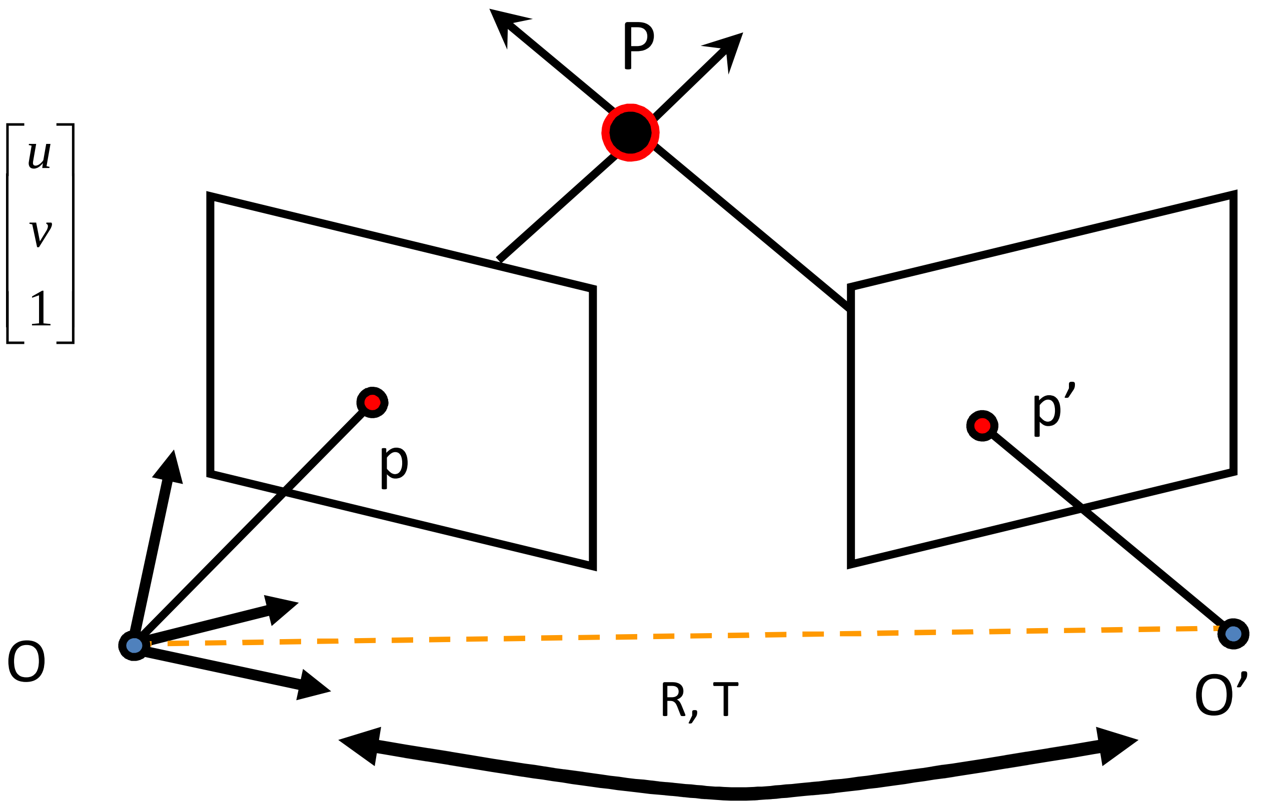
- Two views of the same object
- Suppose I know the camera positions and camera matrices
- Given a point on left image, how can I find the corresponding point on right image?

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



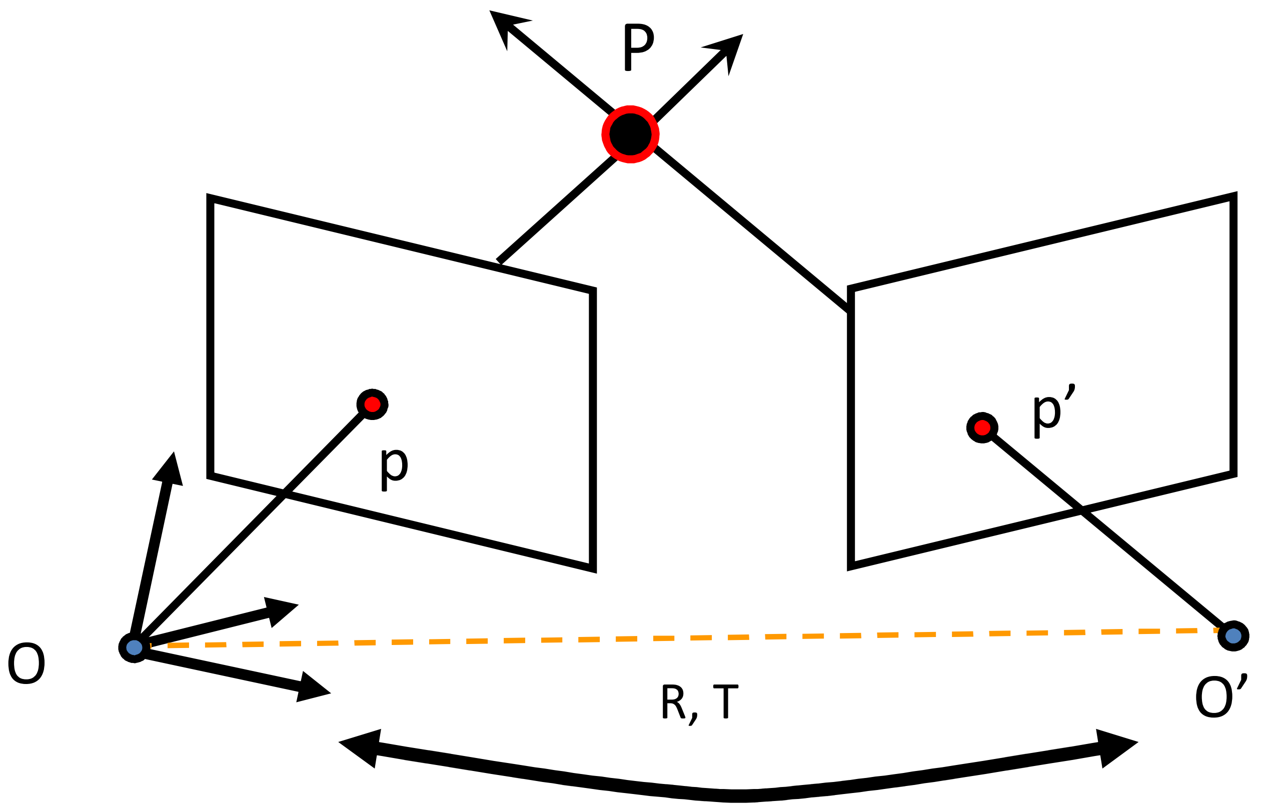
$$p \rightarrow M\ P = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}$$

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

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

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

Epipolar Constraint



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

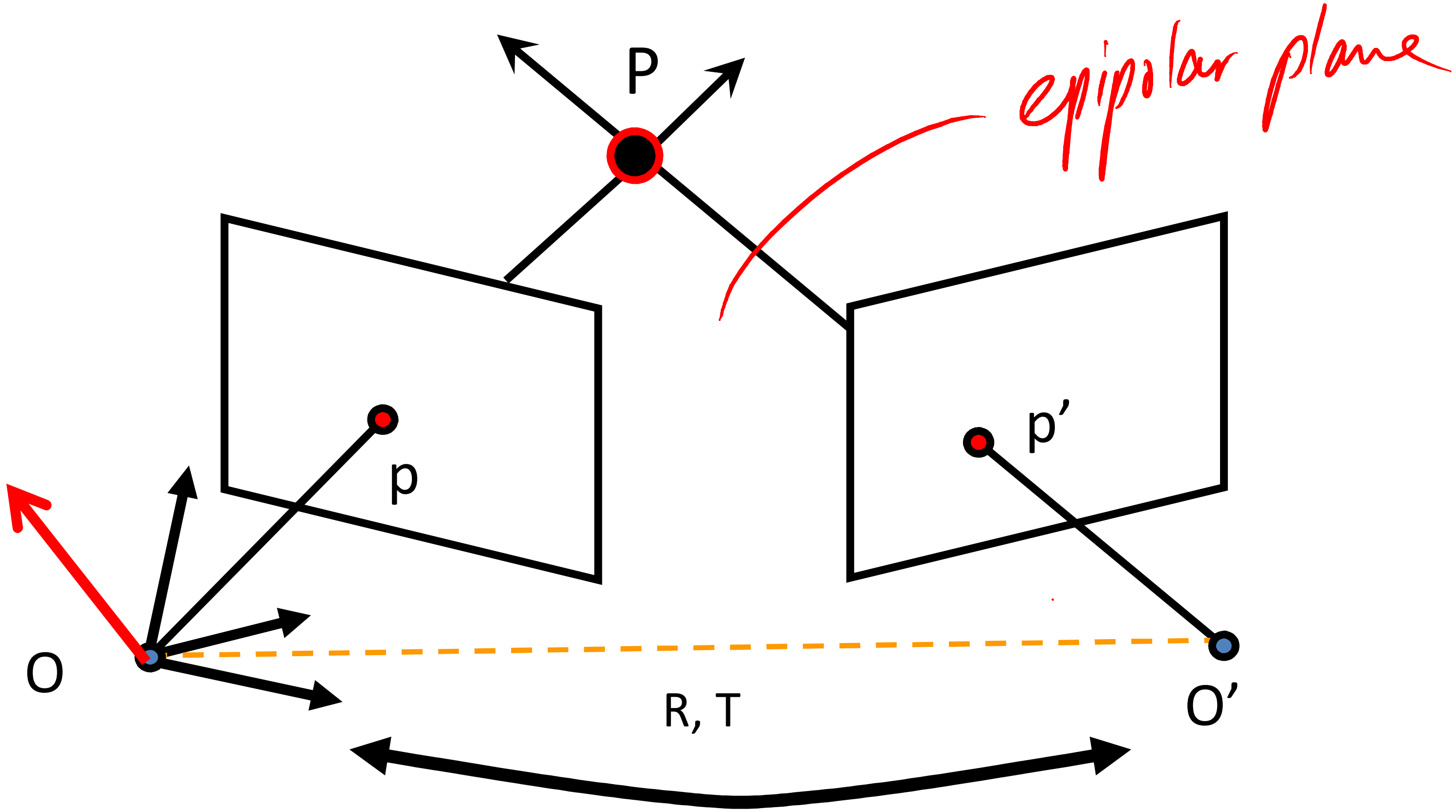
$$M = [I \quad 0]$$

K$_1$ and K$_2$ are known
(calibrated cameras)

$$M' = K[R \quad T]$$

$$M' = [R \quad T]$$

Epipolar Constraint



$$T \times (R\, p') \perp$$ epipolar plane

Perpendicular to epipolar plane

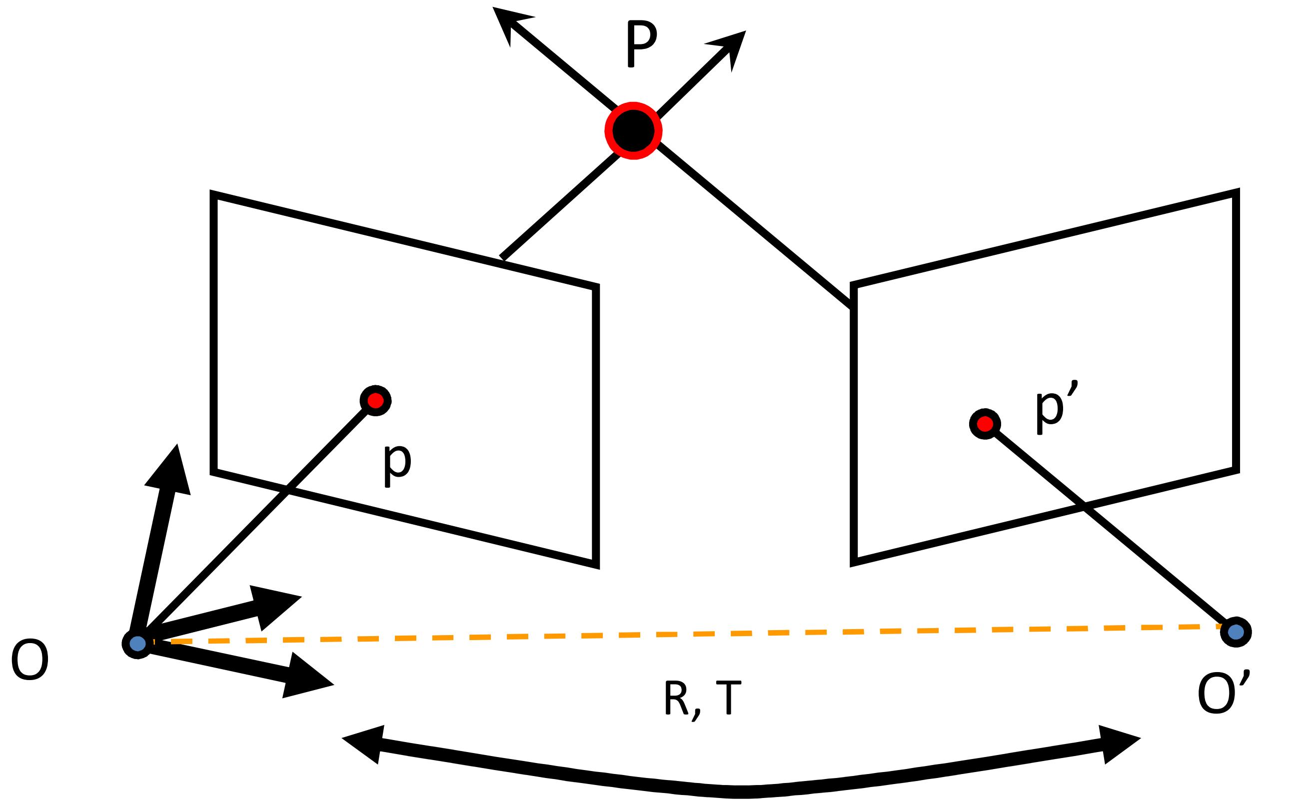
$$\mathrm{p}^{\mathrm{T}} \cdot [\mathrm{T} \times (\mathrm{R}\, \mathrm{p}')] = 0$$

Cross product as matrix multiplication

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

"skew symmetric matrix"

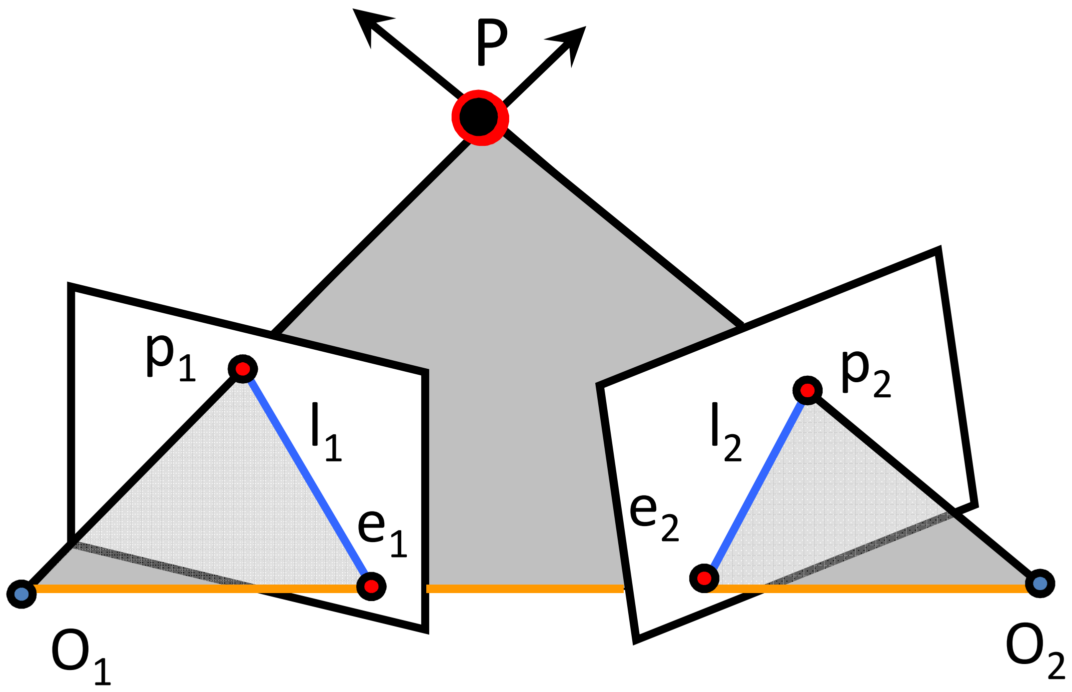
Triangulation



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

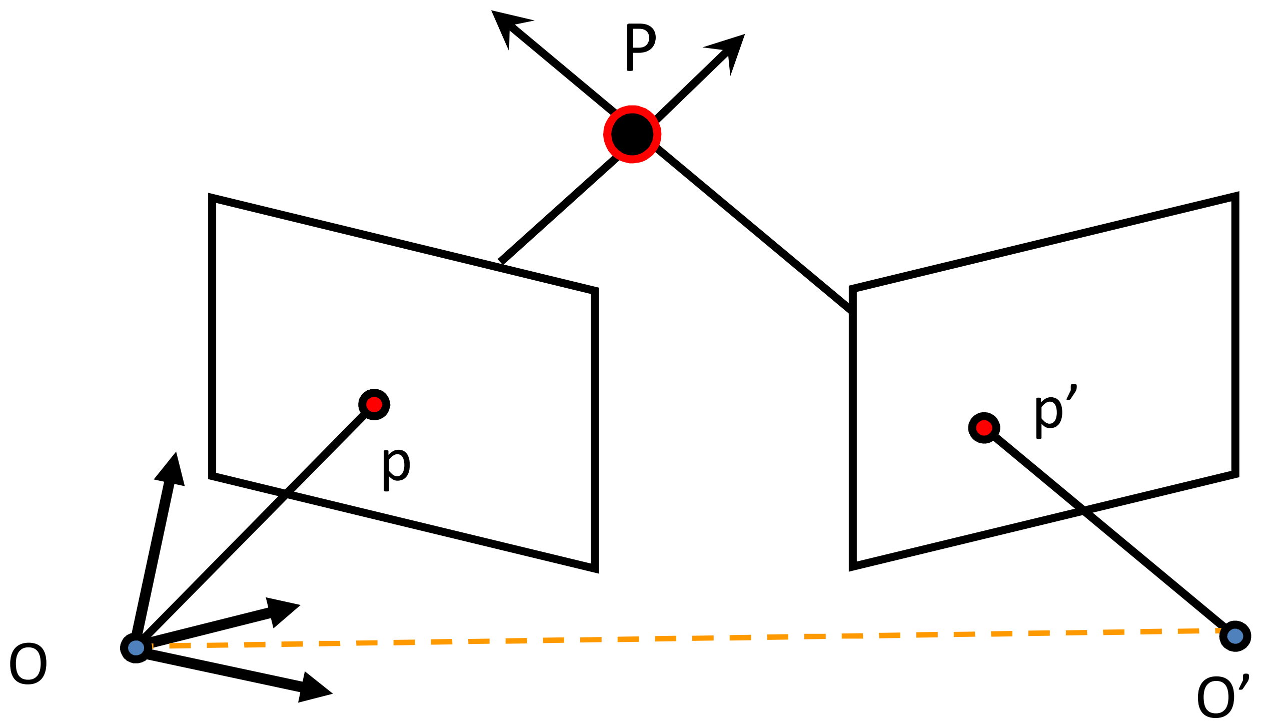
(Longuet-Higgins, 1981) E = essential matrix

Triangulation



- $E\ p_2$ 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 is singular (rank two)
- $E\ e_2 = 0$ and $E^T e_1 = 0$
- E is 3x3 matrix; 5 DOF

Triangulation

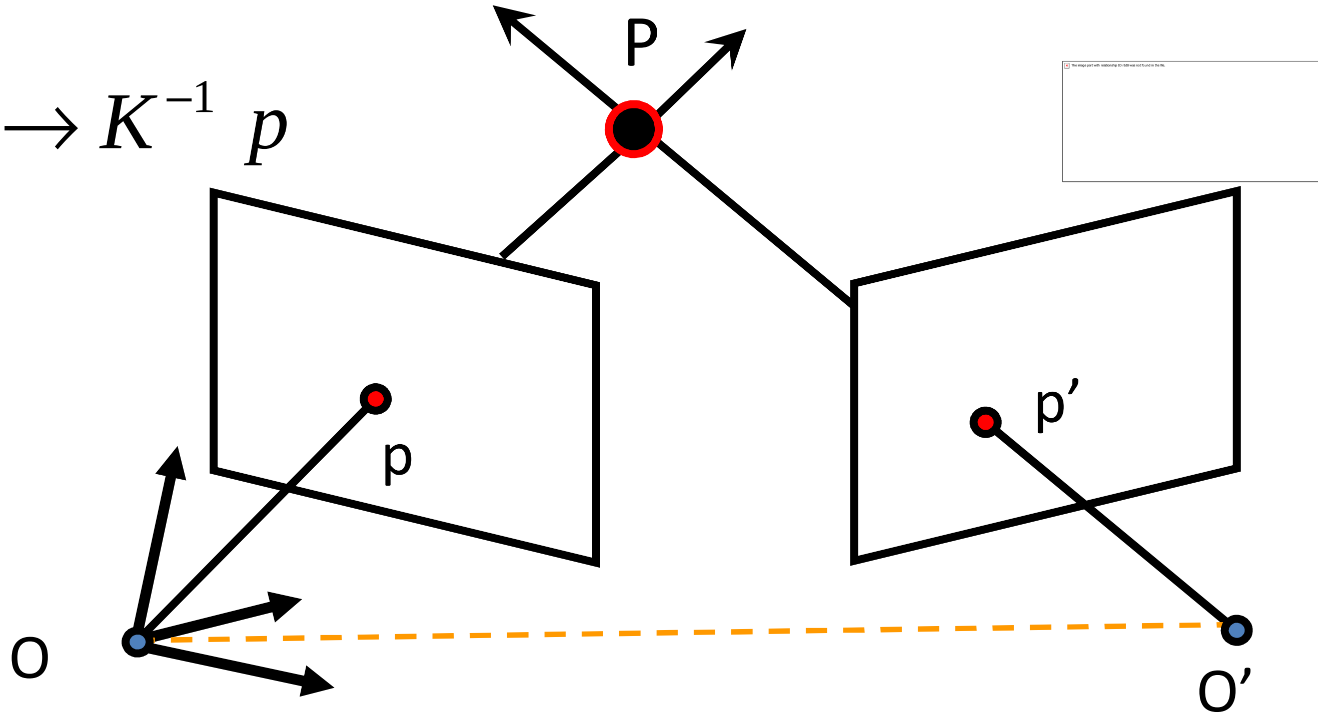


$$P \rightarrow M\,P \;\rightarrow p = \begin{bmatrix} u \\ v \end{bmatrix} \qquad\qquad M = \boxed{K}\begin{bmatrix} I & 0 \end{bmatrix}$$

unknown

Triangulation

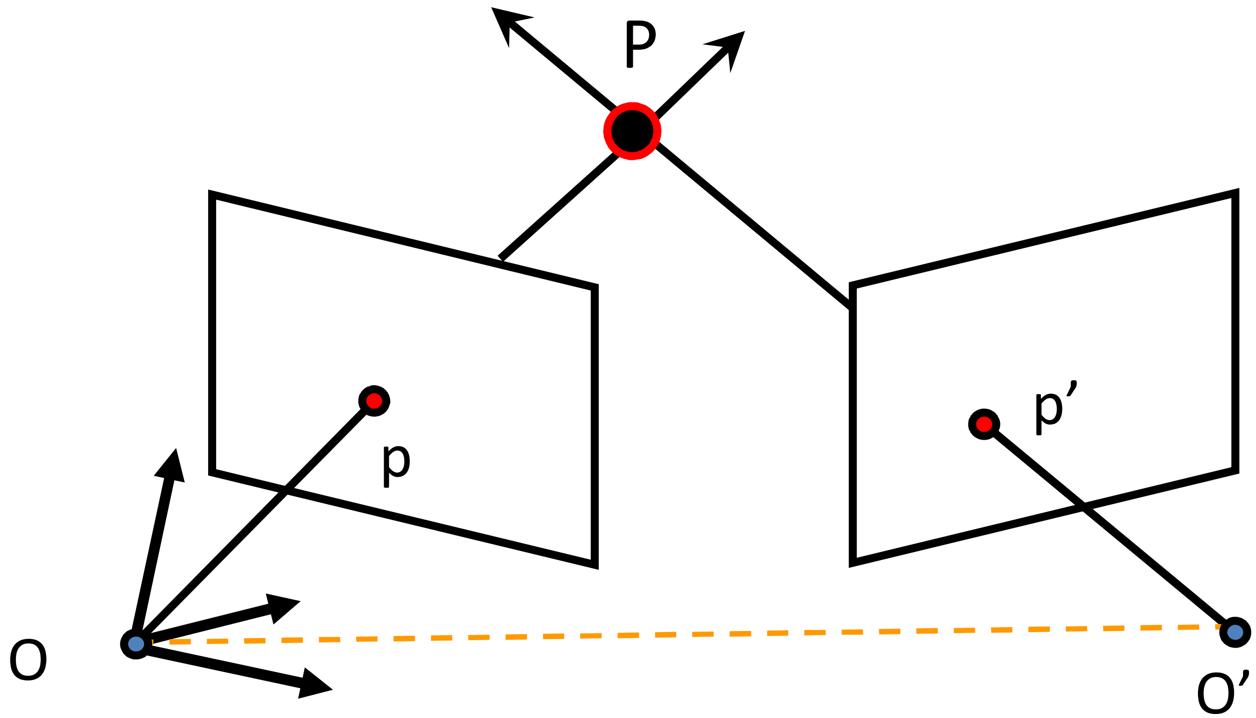
$$p \rightarrow K^{-1}\ p$$



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

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

Triangulation

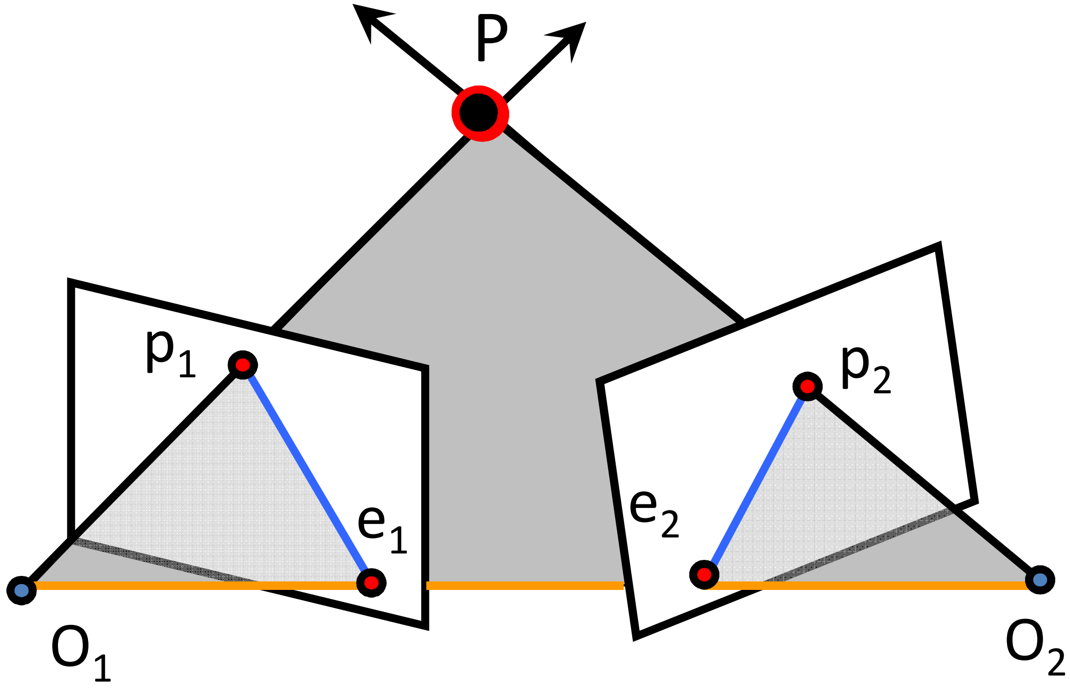


$$\boxed{p^{T}\, F\, p' = 0}$$

F = Fundamental Matrix

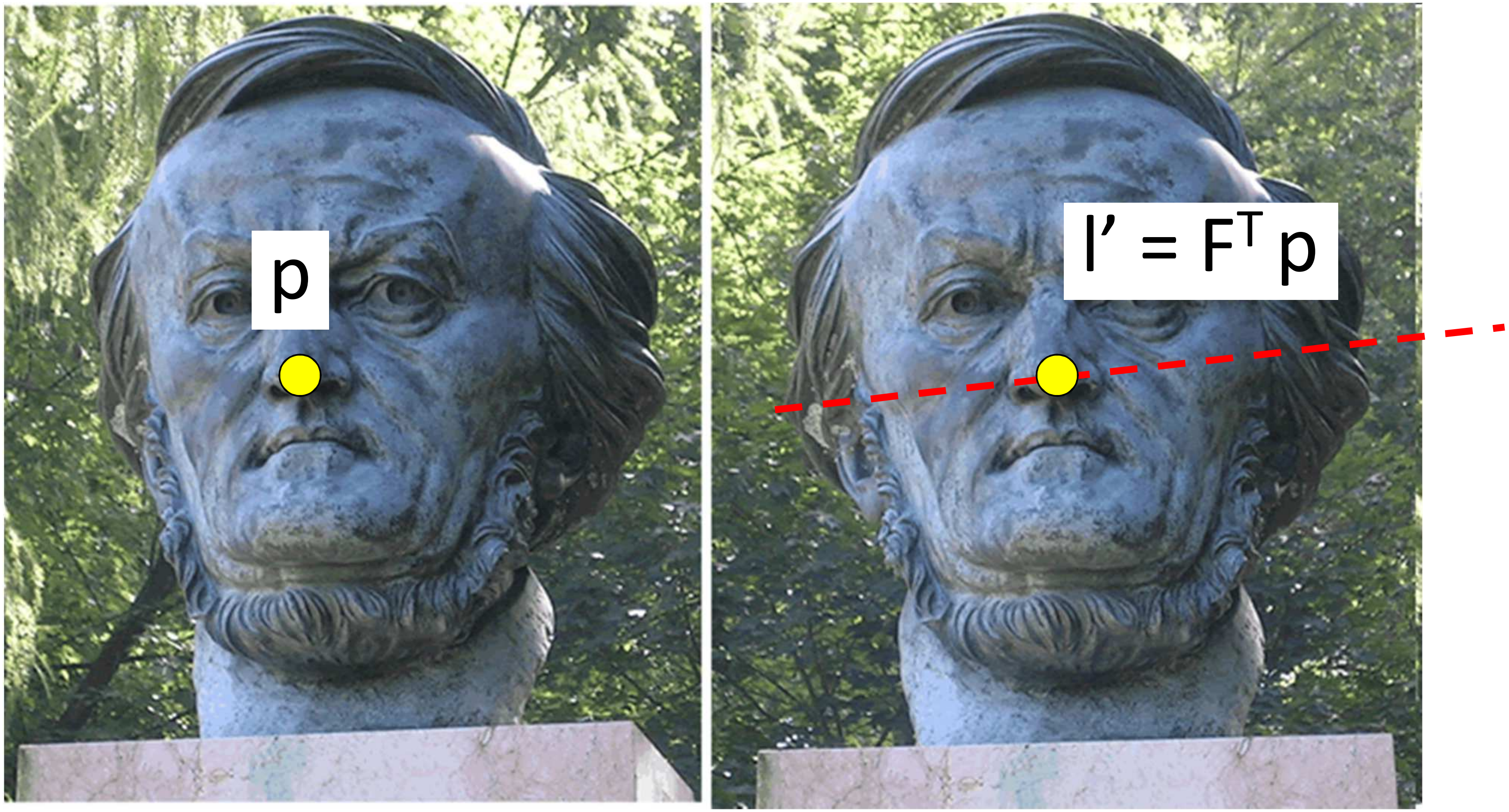
(Faugeras and Luong, 1992)

Triangulation



- $F\ p_2$ 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 is singular (rank two)
- $F\ e_2 = 0$ and $F^T e_1 = 0$
- F is 3x3 matrix; 7 DOF

Why is F 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 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

What we will learn today?

- Why is stereo useful?
- Epipolar constraints
- Essential and fundamental matrix
- Estimating F (Problem Set 2 (Q2))
- Rectification

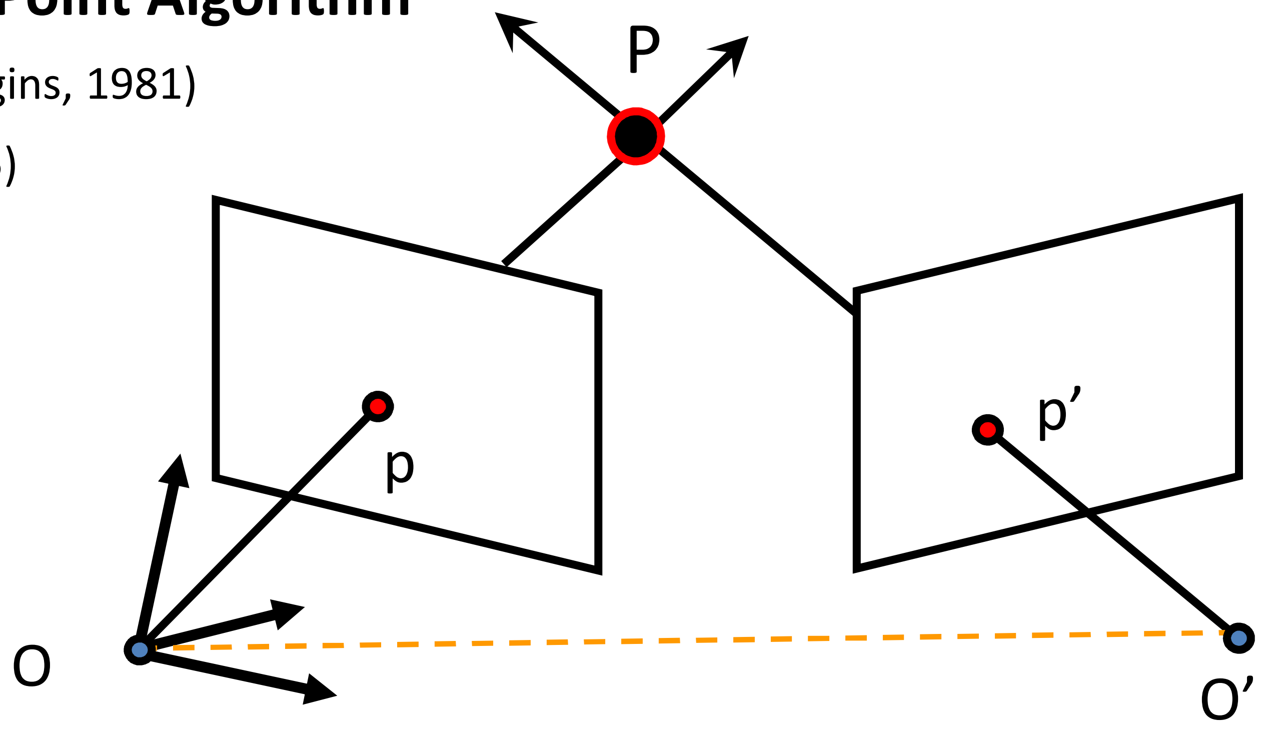
Reading:
[HZ] Chapters: 4, 9, 11
[FP] Chapters: 10

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

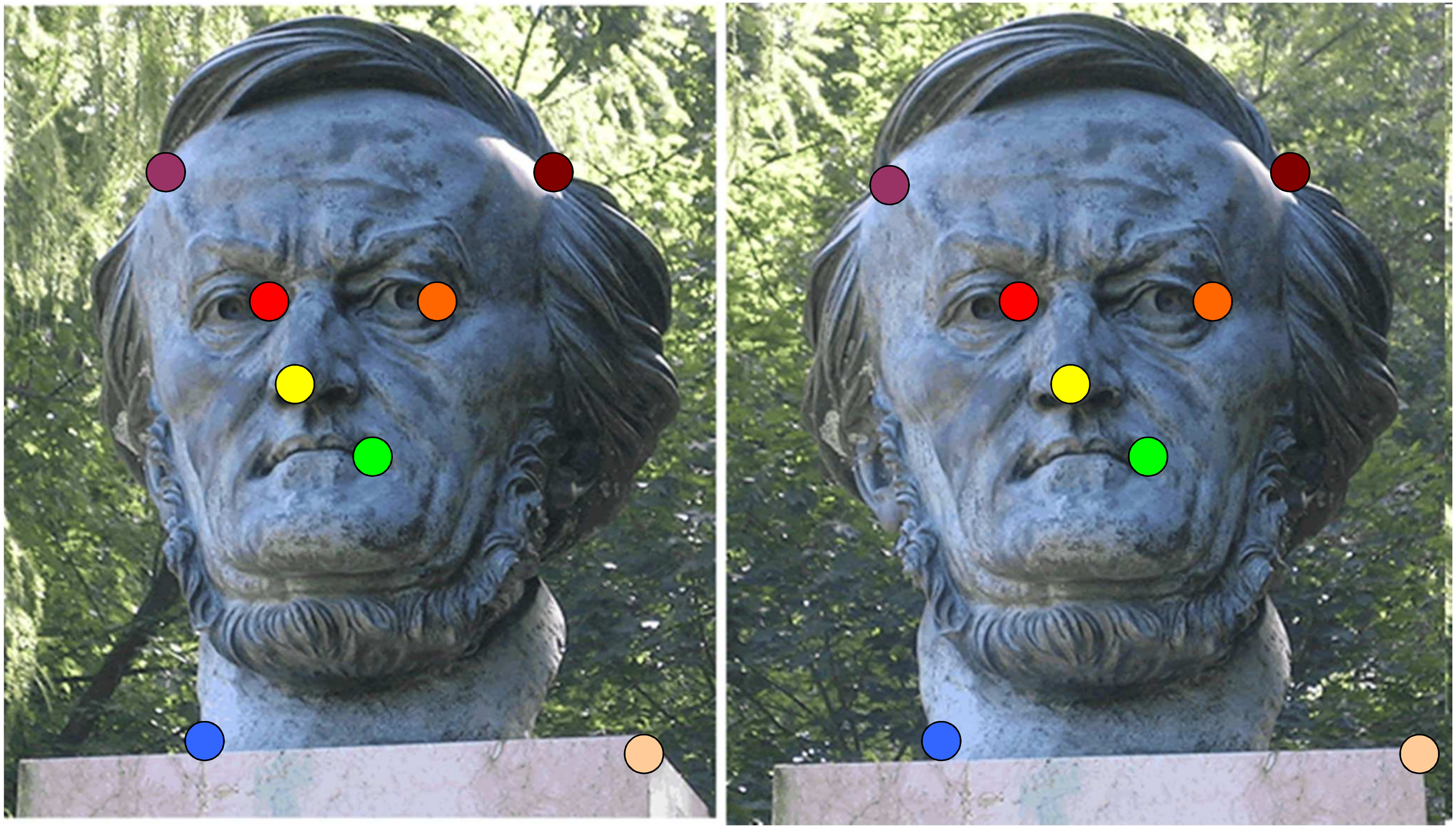
$$\mathrm{p}^{\mathrm{T}}\ \mathrm{F}\,\mathrm{p}' = 0 \quad \Rightarrow$$

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

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

Let's take 8 corresponding points

Estimating F



Estimating F

known observations   unknowns

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

- Homogeneous system $\mathbf{W}\,\mathbf{f} = 0$

constraint

- Rank 8 ⟶ A non-zero solution exists (unique)
- If N>8 ⟶ Lsq. solution by SVD! ⟶ $\hat{\mathrm{F}}$ $\|\mathbf{f}\| = 1$

Estimating F

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

The estimated F may have full rank (det(F) ≠0)

(F should have rank=2 instead)

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

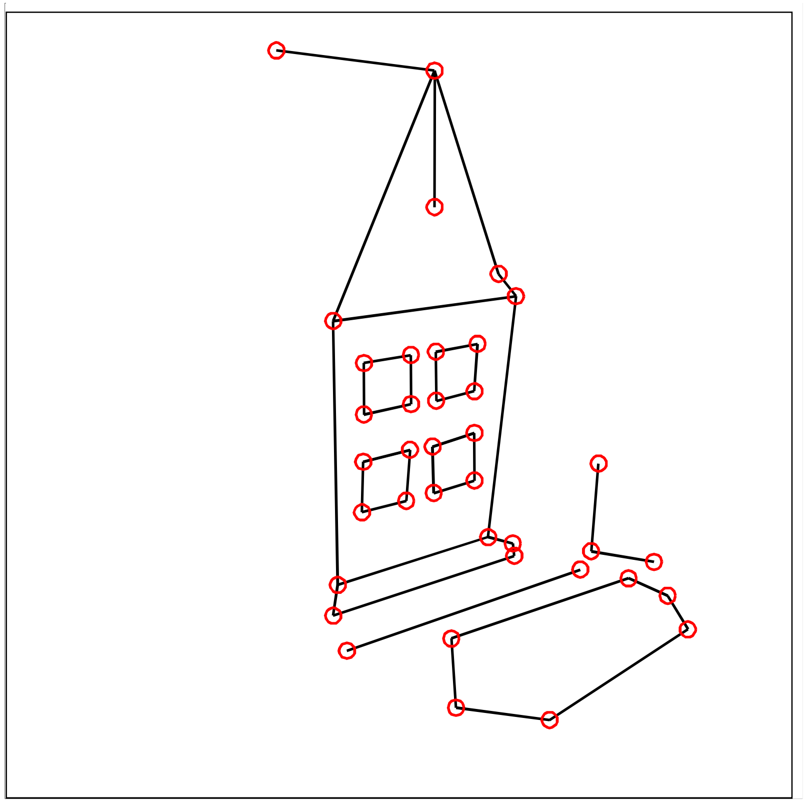
Frobenius norm (*)

Subject to det(F)=0

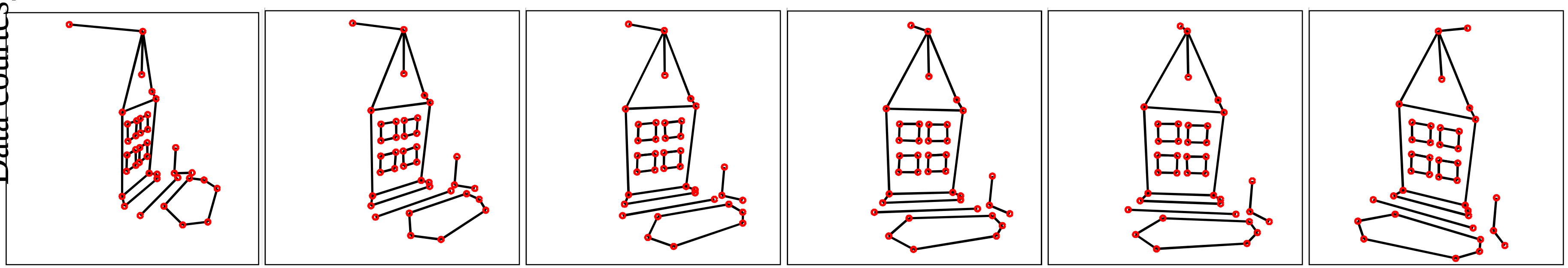
SVD (again!) can be used to solve this problem

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

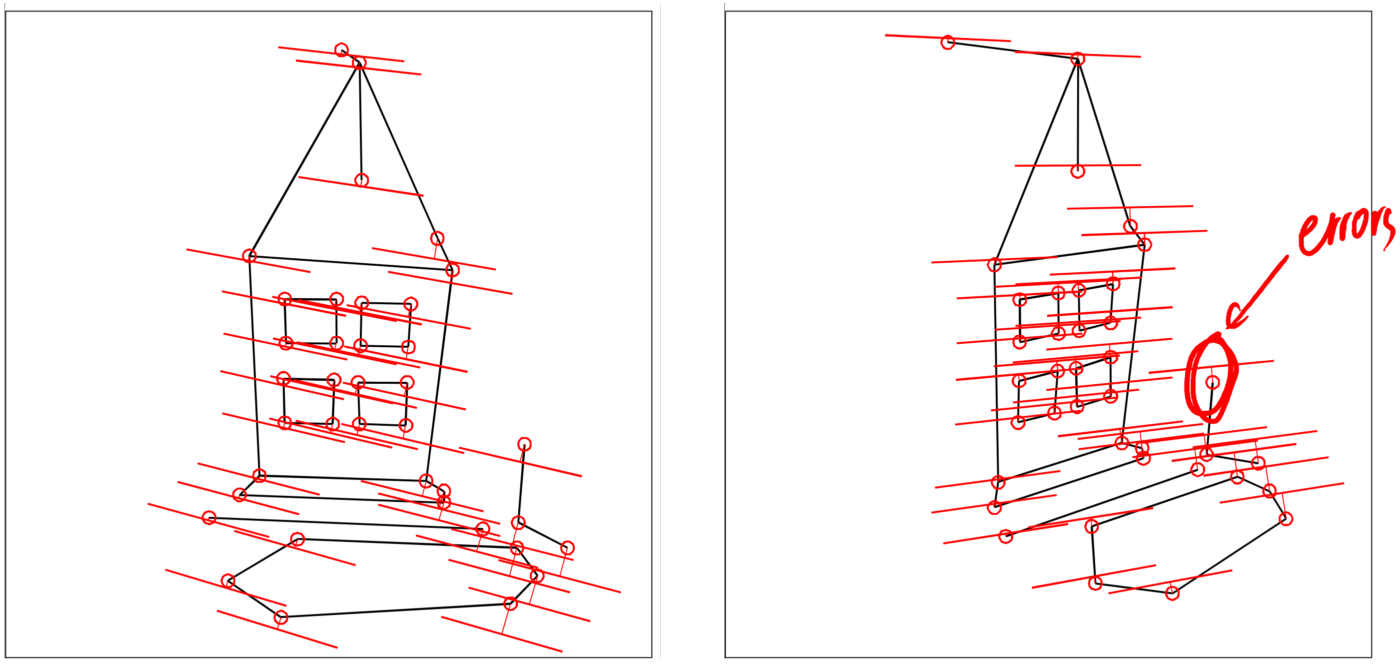
Example



Data courtesy of R. Mohr and B. Boufama.



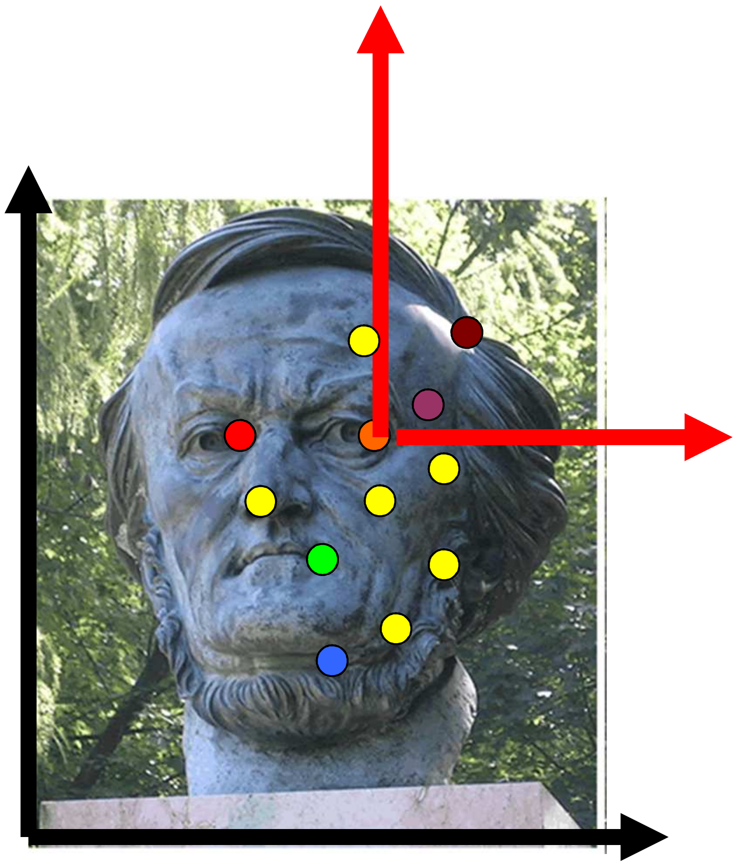
Example



errors

Mean errors:
10.0pixel
9.1pixel

Normalization



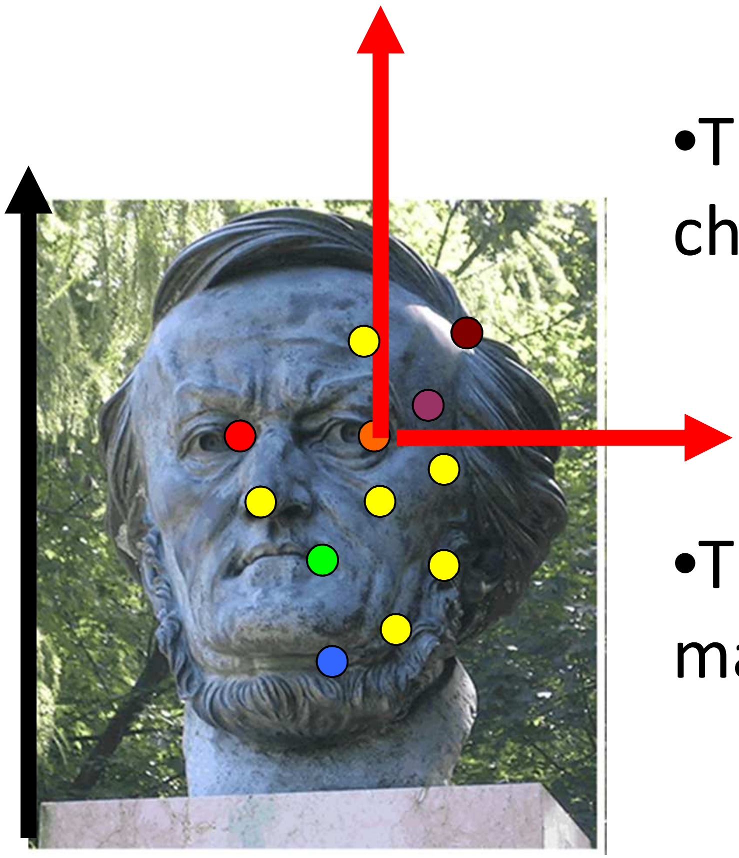
Is the accuracy in estimating F function of the ref. system in the image plane?

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

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

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

Normalization

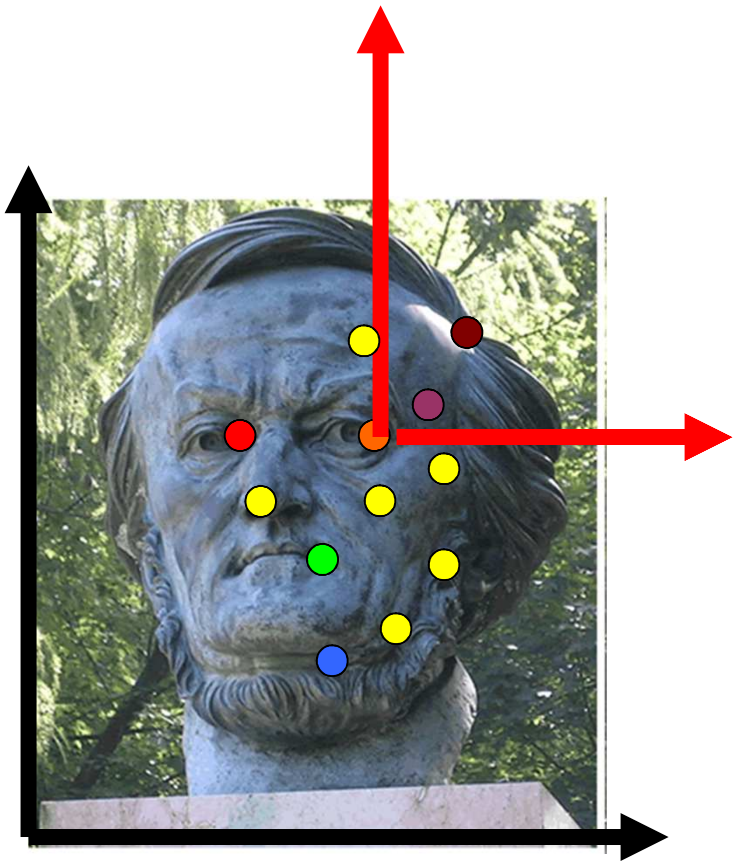


- The accuracy in estimating F does change if a transformation T is applied

- There exists a T for which accuracy is maximized

Why?

Normalization



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

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

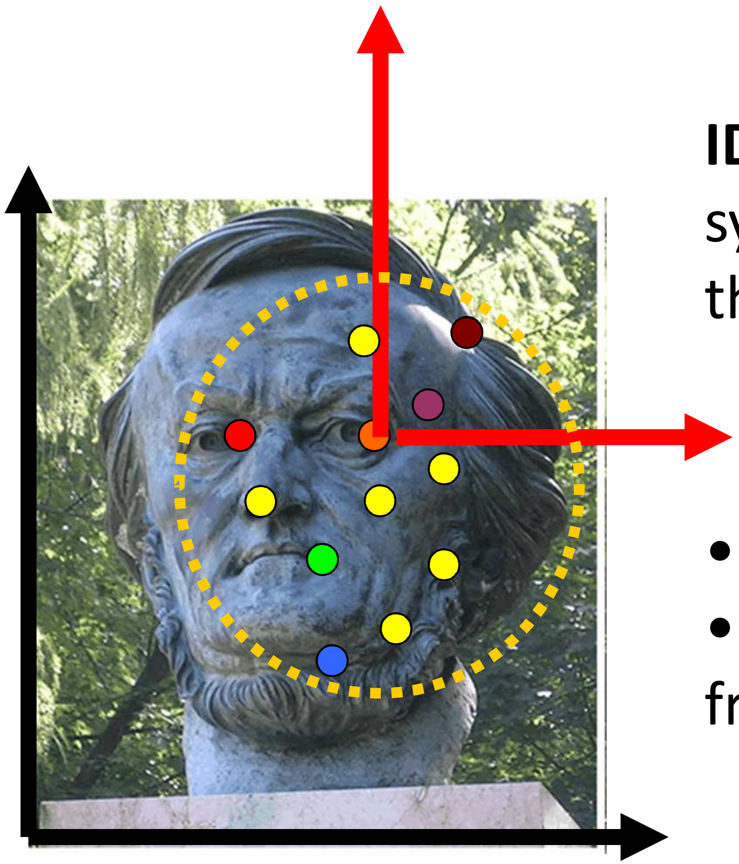
Lsq solution by SVD $\longrightarrow$ F

- SVD enforces Rank(W)=8
- Recall the structure of W:
  Highly un-balance
  (not well conditioned)

-Values of W must have similar magnitude

More details HZ pag 108

Normalization



IDEA: 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:

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

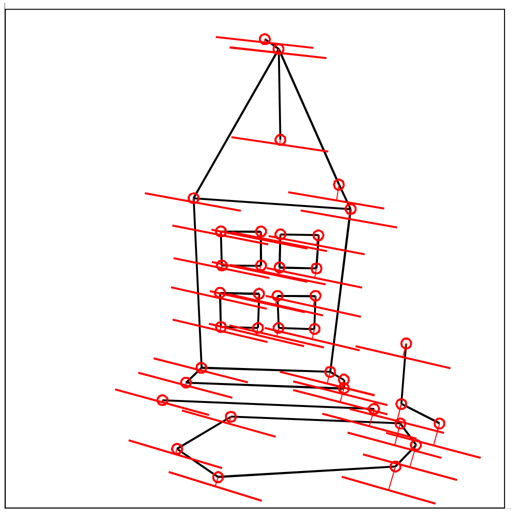
2. Use the eight-point algorithm to compute F'_q from the points q_i and q_i' .

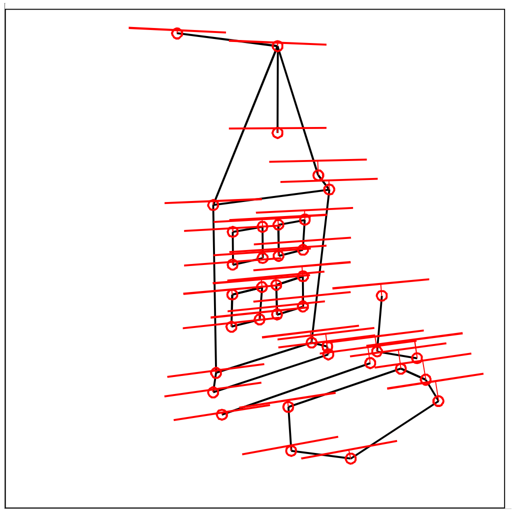
3. Enforce the rank-2 constraint. $\quad \rightarrow F_q \quad \begin{cases} q^T \; F_q \; q' = 0 \\ \det(\, F_q \,) = 0 \end{cases}$

4. De-normalize F_q: $\quad F = T'^T F_q T$

Example

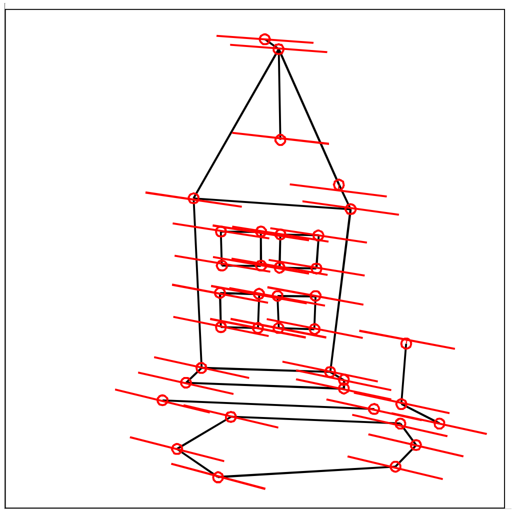
Without transformation

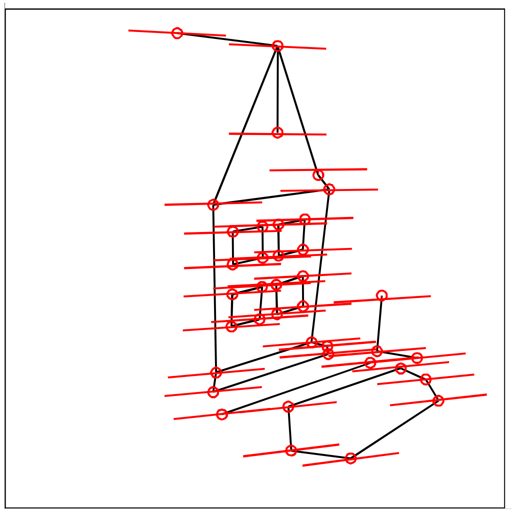




Mean errors:
10.0pixel
9.1pixel

With transformation





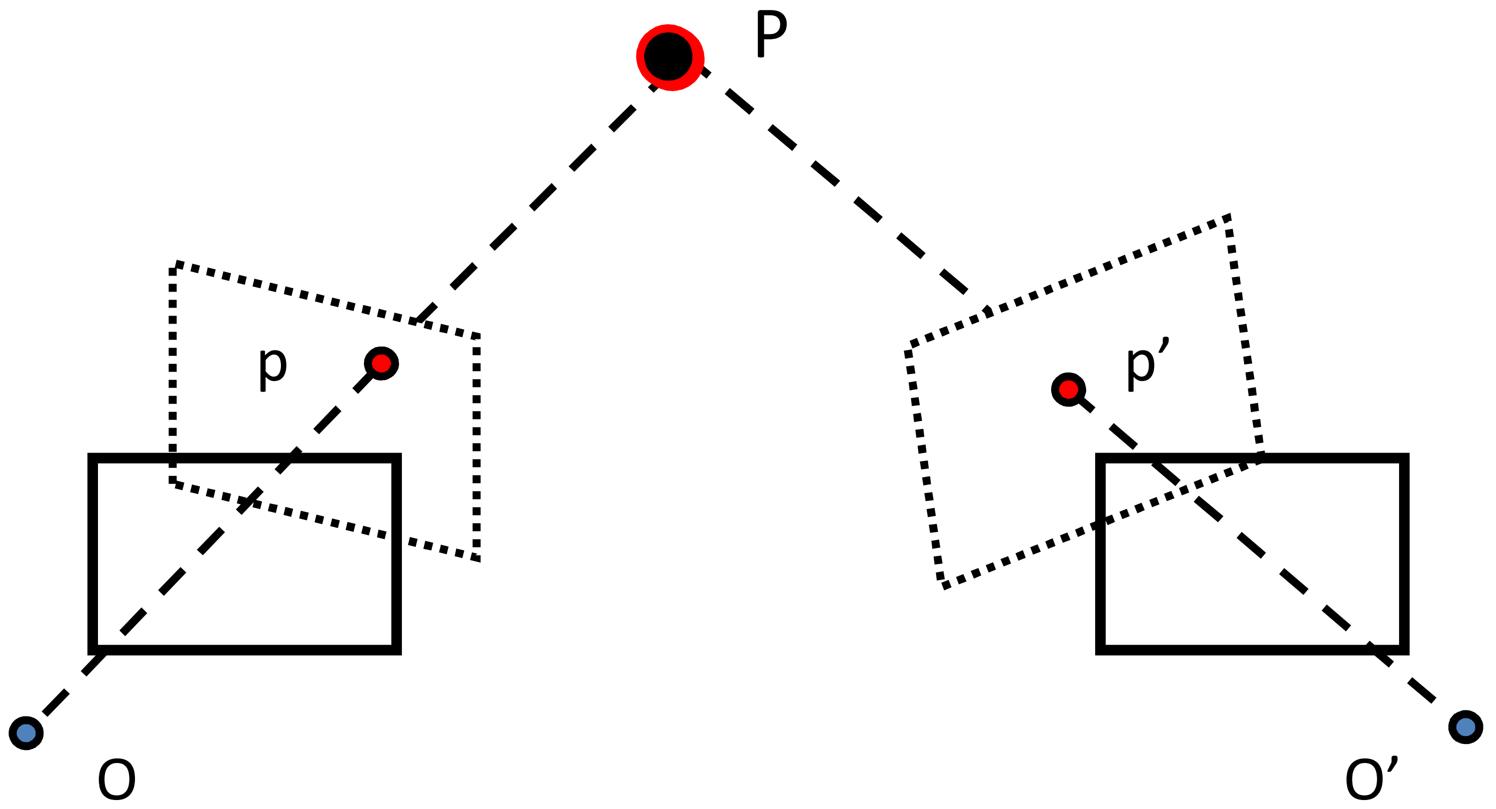
Mean errors:
1.0pixel
0.9pixel

What we will learn today?

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

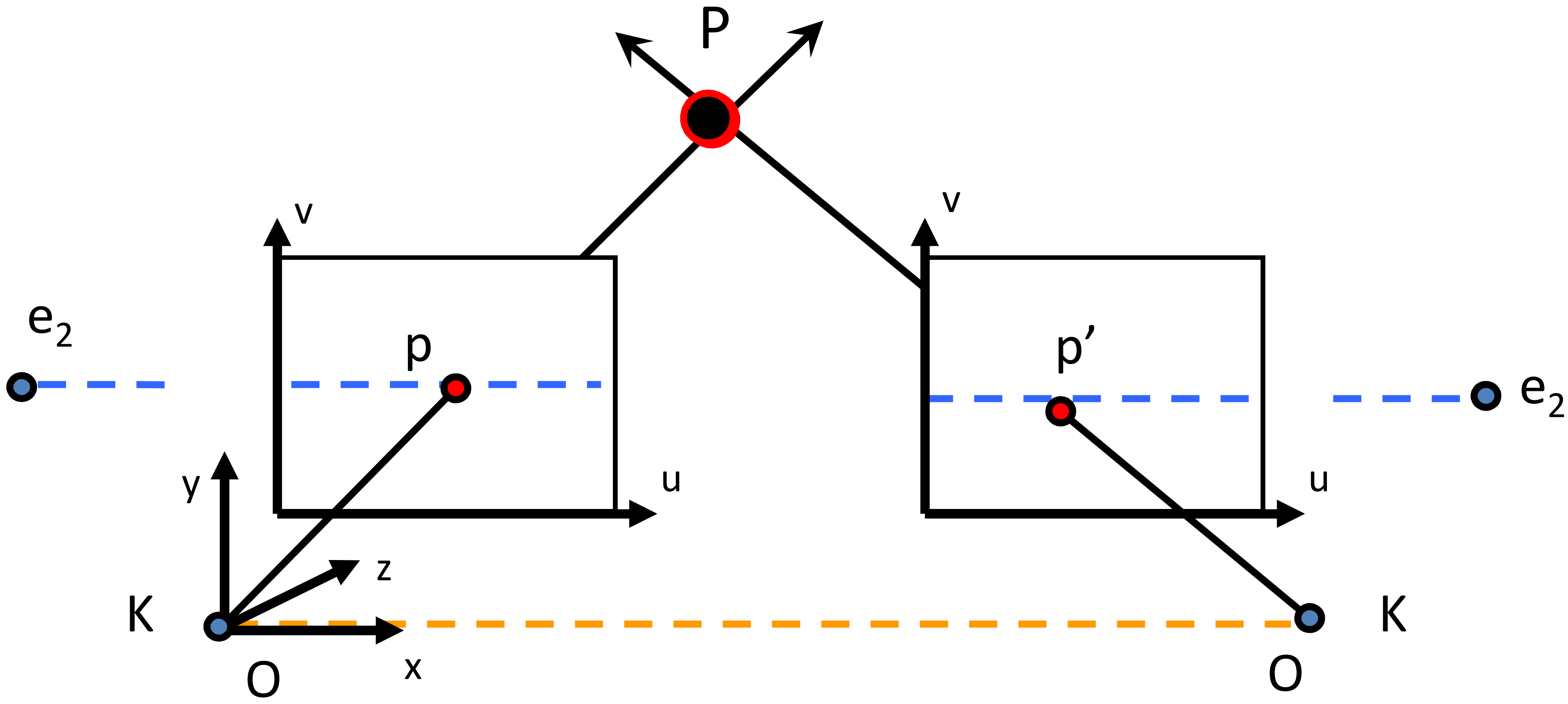
Reading:
[HZ] Chapters: 4, 9, 11
[FP] Chapters: 10

Rectification



- Make two camera images "parallel"
  - Correspondence problem becomes easier

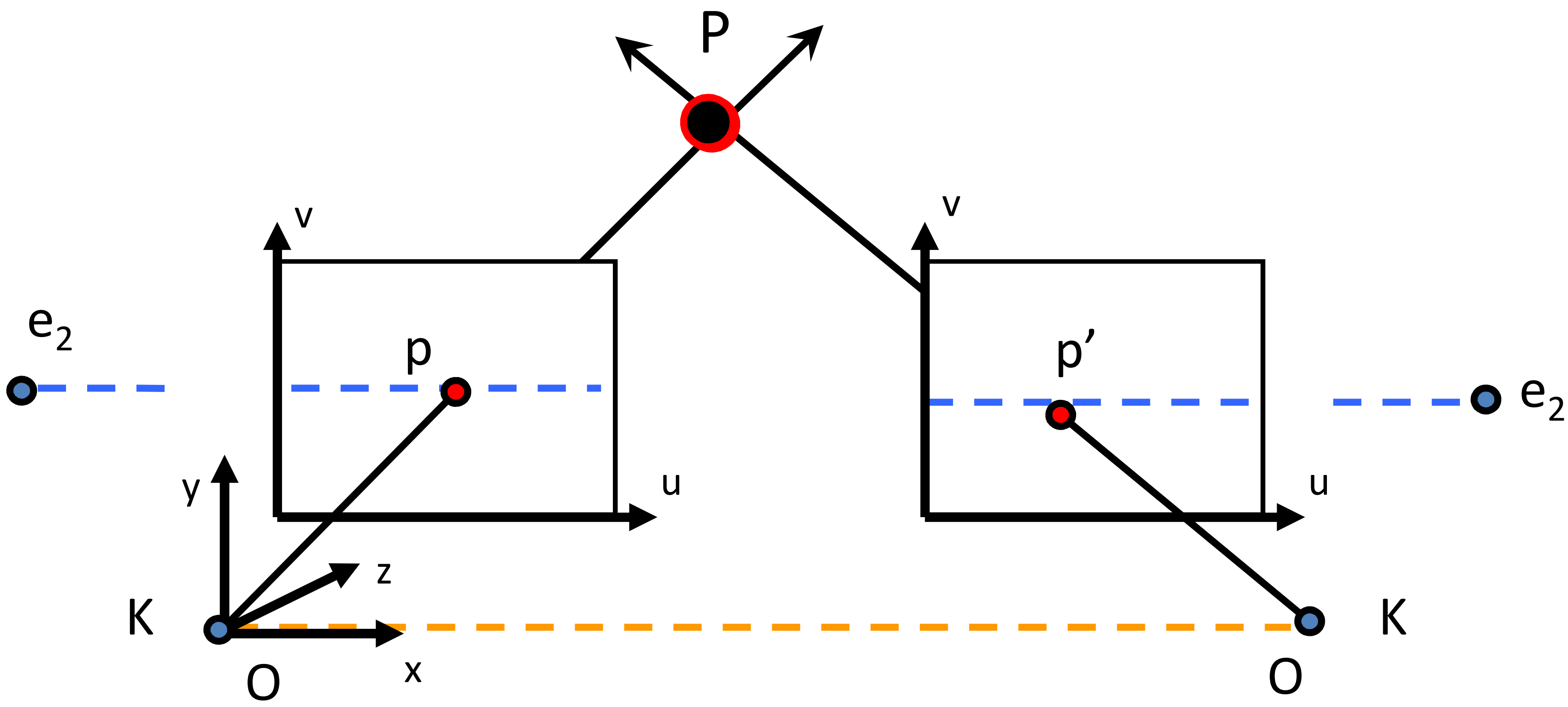
Rectification



- Parallel epipolar lines
- Epipoles at infinity
- v = v'

Let's see why….

Rectification



K_1=K_2 = known

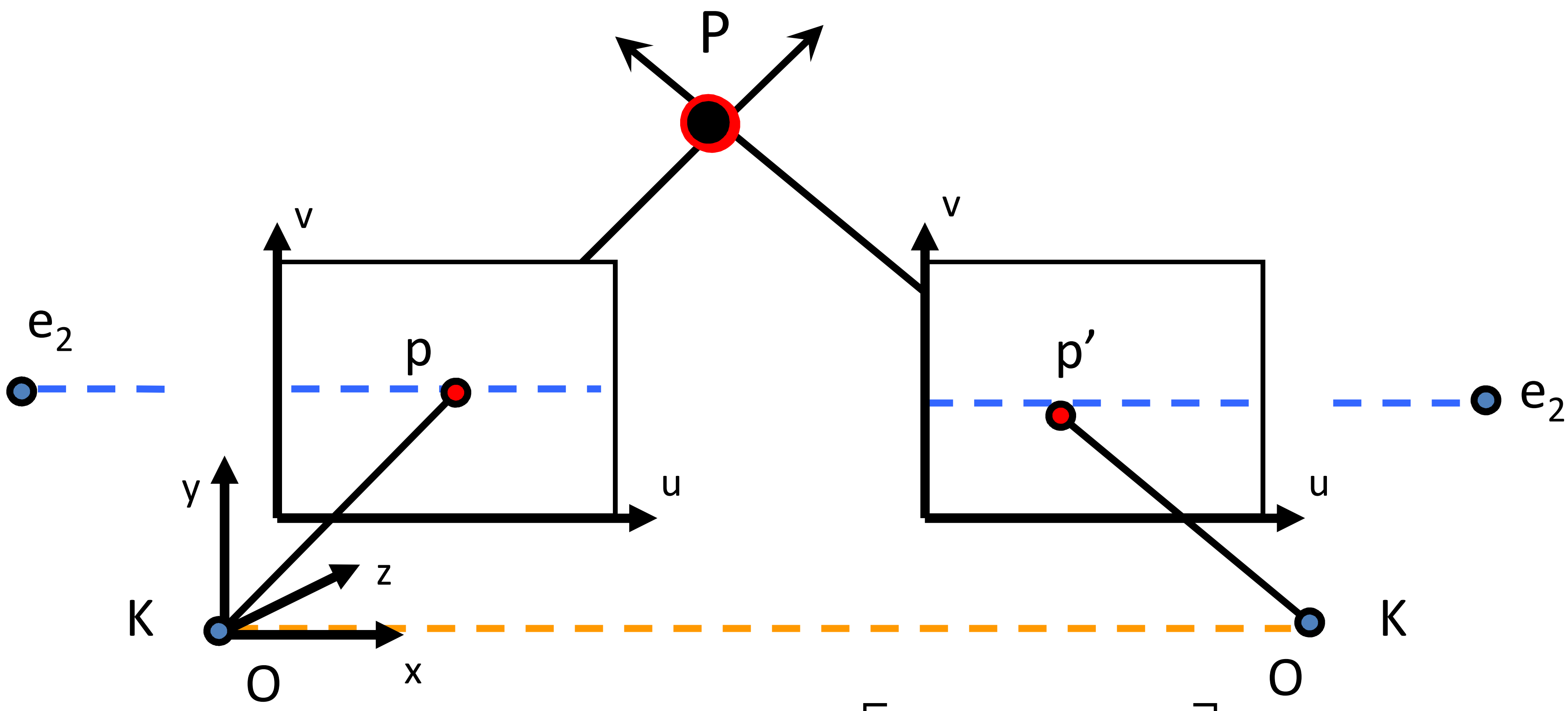
x parallel to O_1O_2

$$E = [t_\times]R$$

Cross product as matrix multiplication

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

Rectification



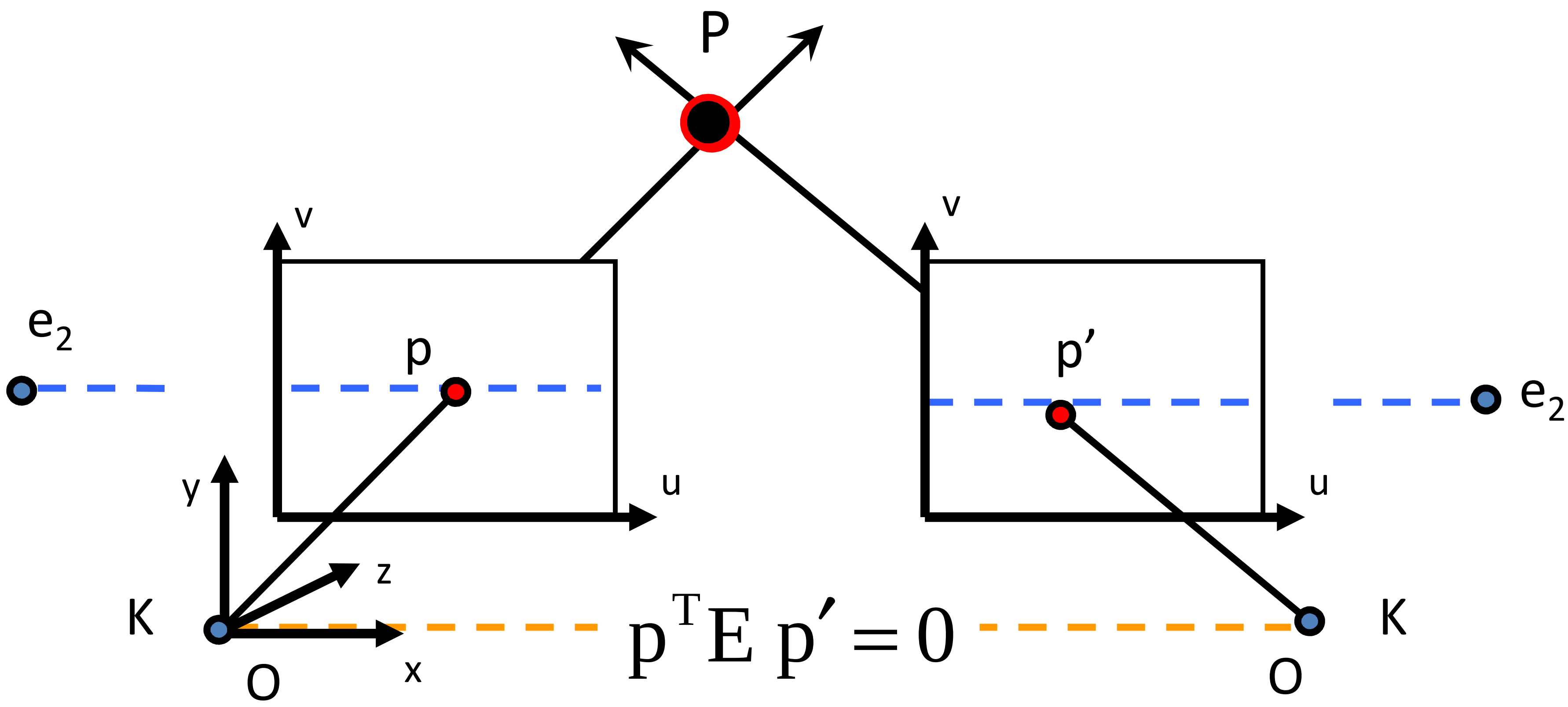
K_1=K_2 = known

x parallel to O_1O_2

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

$\rightarrow$ v = v'?

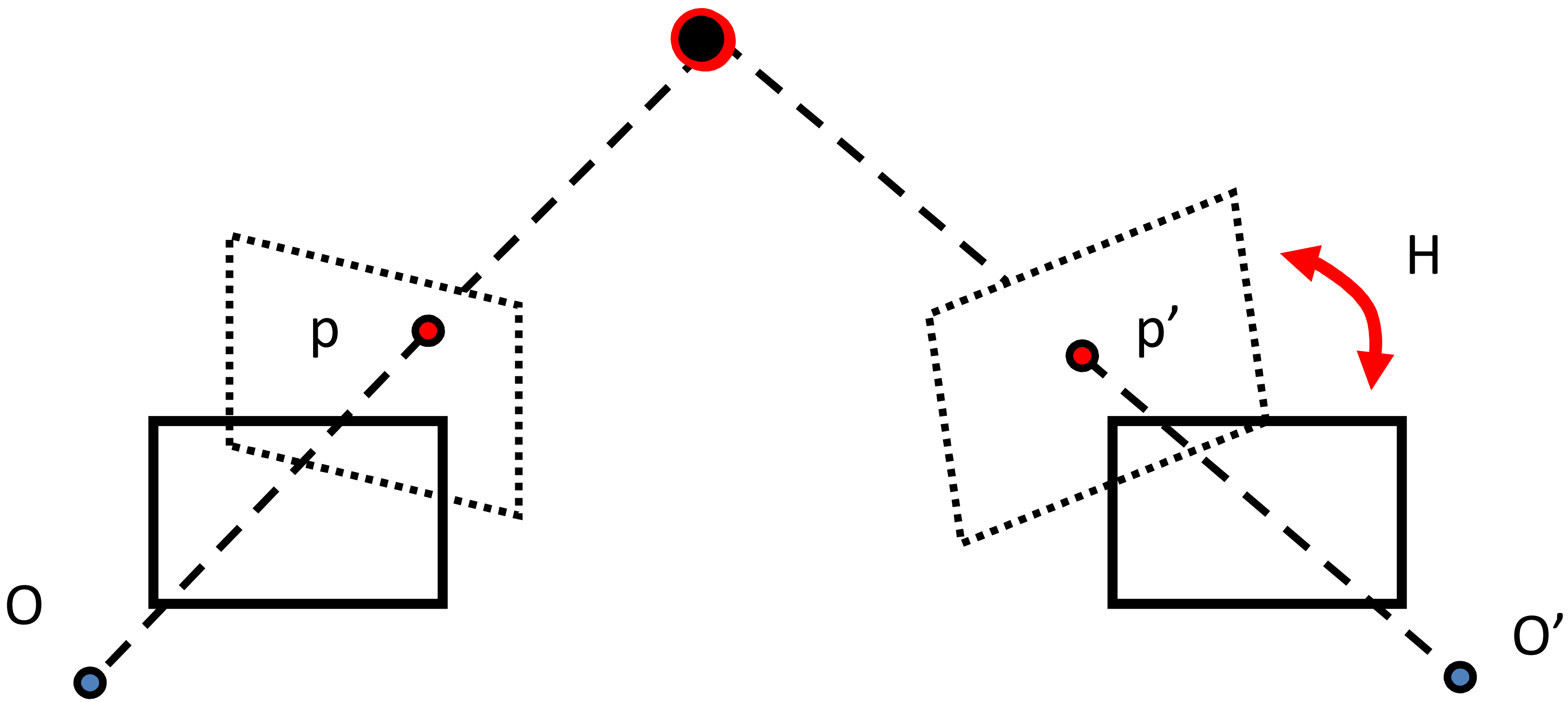
Rectification



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

$$\begin{pmatrix} u & v & 1 \end{pmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -T \\ 0 & T & 0 \end{bmatrix} \begin{pmatrix} u' \\ v' \\ 1 \end{pmatrix} = 0 \qquad \begin{pmatrix} u & v & 1 \end{pmatrix} \begin{pmatrix} 0 \\ -T \\ Tv' \end{pmatrix} = 0 \qquad \begin{aligned} Tv &= Tv' \\ \rightarrow v &= v' \end{aligned}$$

Rectification



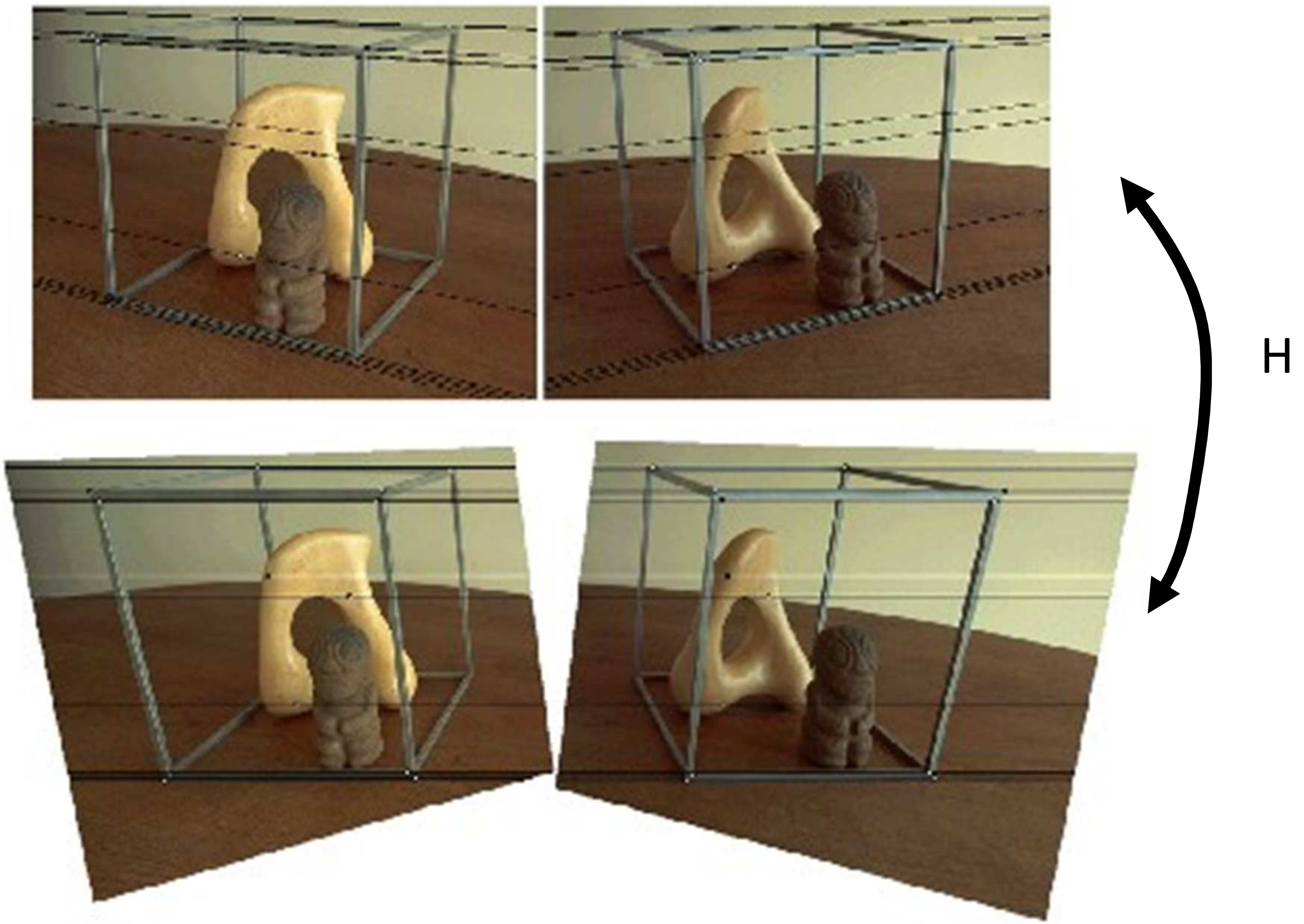
GOAL of rectification : Estimate a perspective transformation H that makes images parallel

Impose v’=v

- This leaves degrees of freedom for determining H
- If not appropriate H is chosen, severe projective distortions on image take place
- We impose a number of restriction while computing H

[HZ] Chapters: 11 (sec. 11.12)

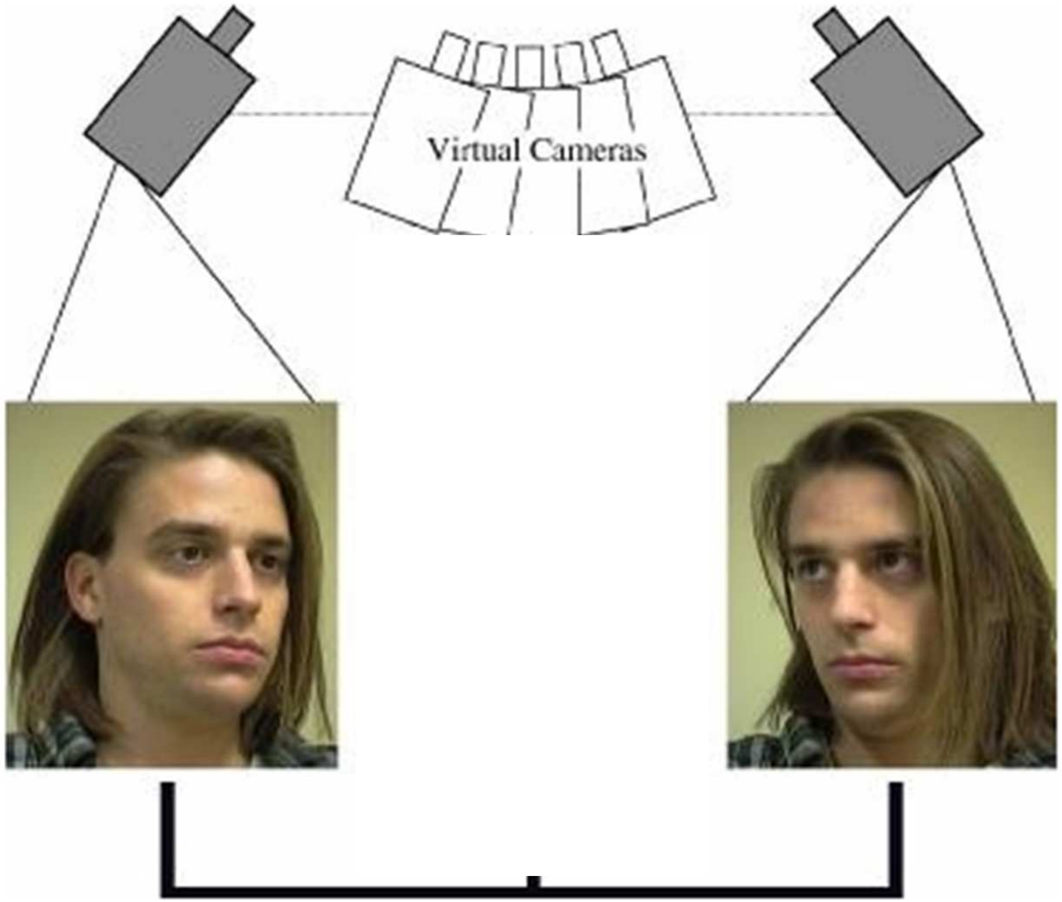
Rectification



Courtesy figure S. Lazebnik

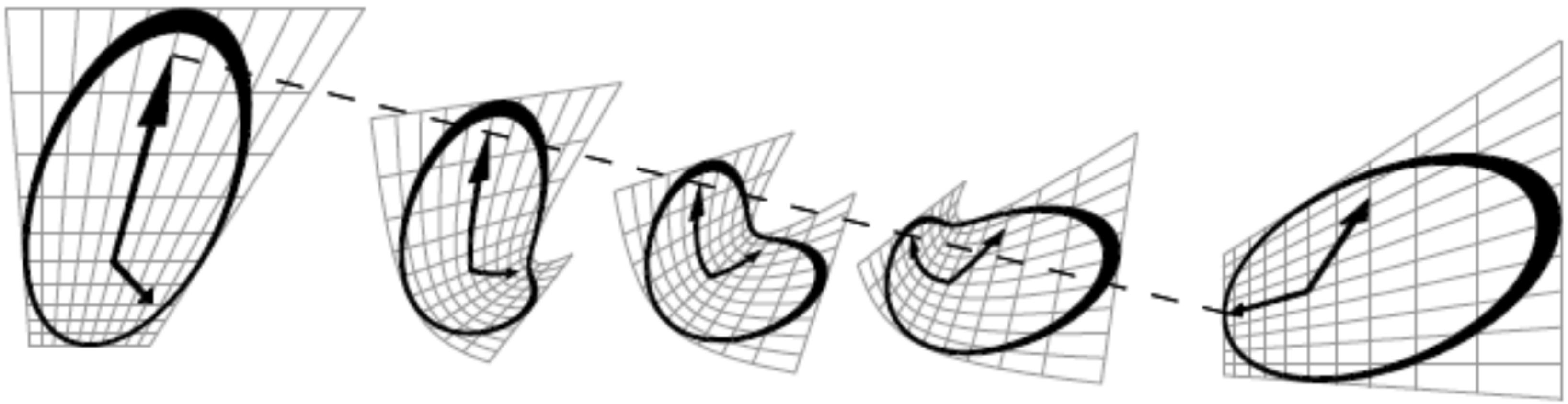
Application: view morphing

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

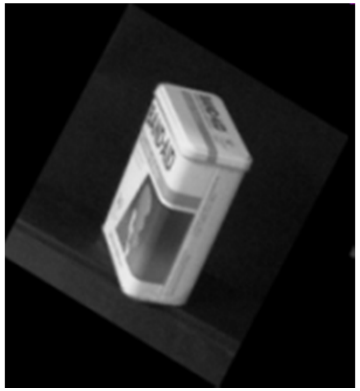


Application: view morphing

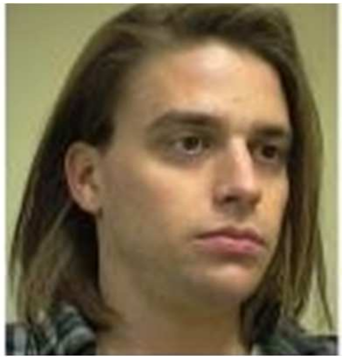
If rectification is not applied, the morphing procedure does not generate geometrically correct interpolations



Application: view morphing

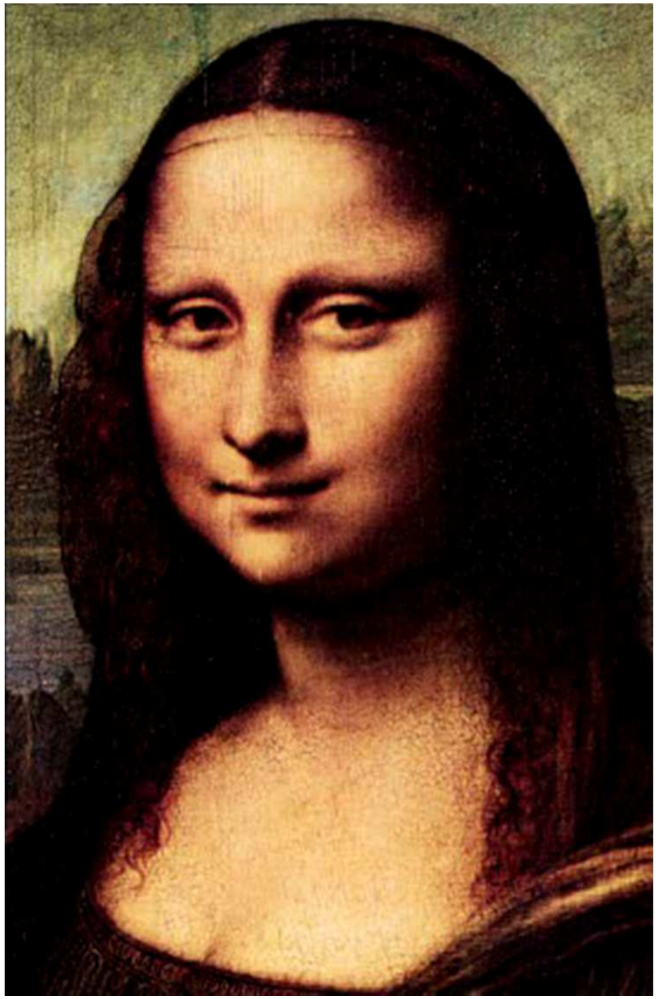


Application: view morphing



Application: view morphing

Application: view morphing



The Fundamental Matrix Song

http://danielwedge.com/fmatrix/

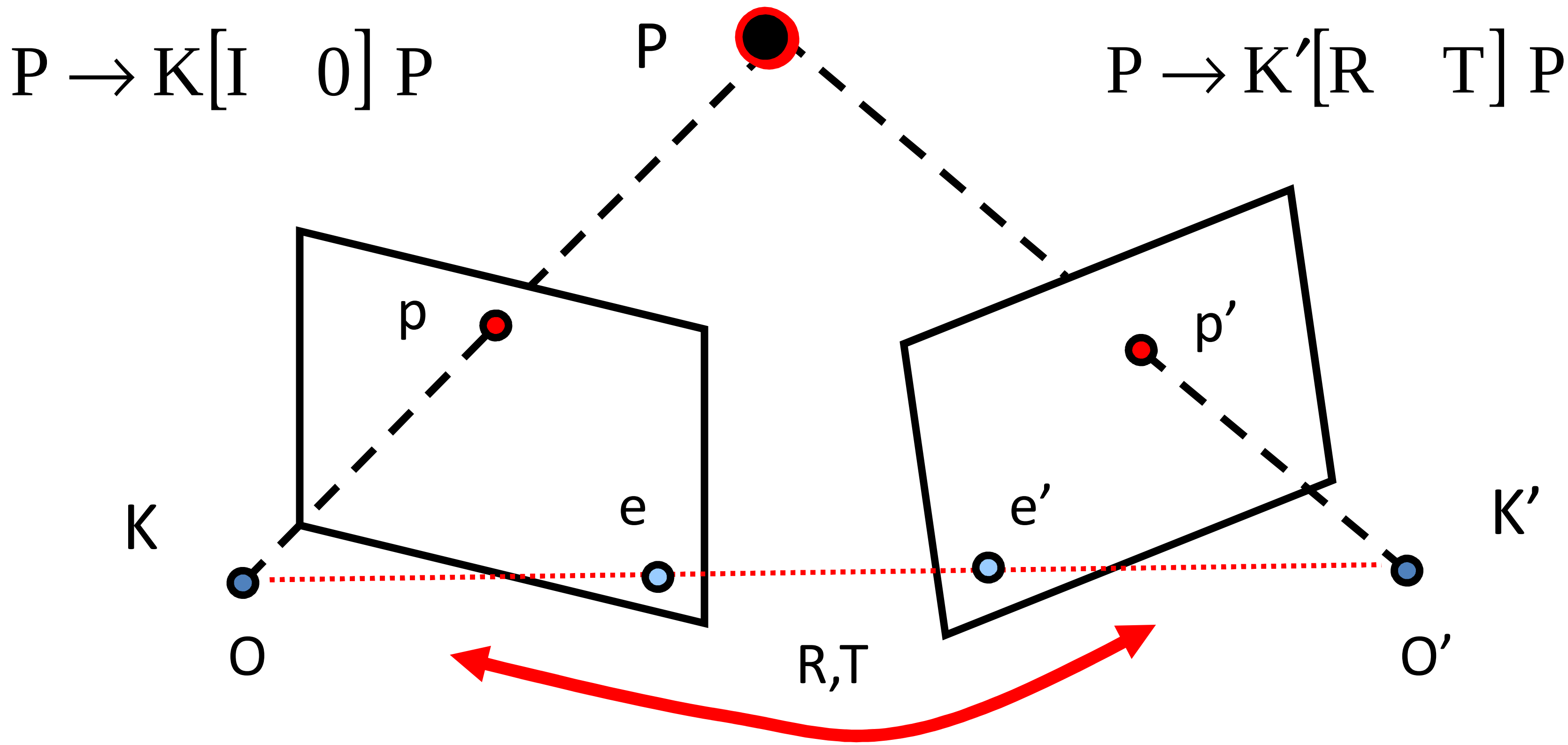
What we have learned today?

- Why is stereo useful?
- Epipolar constraints
- Essential and fundamental matrix
- Estimating F (Problem Set 2 (Q2))
- Rectification

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

Supplementary materials

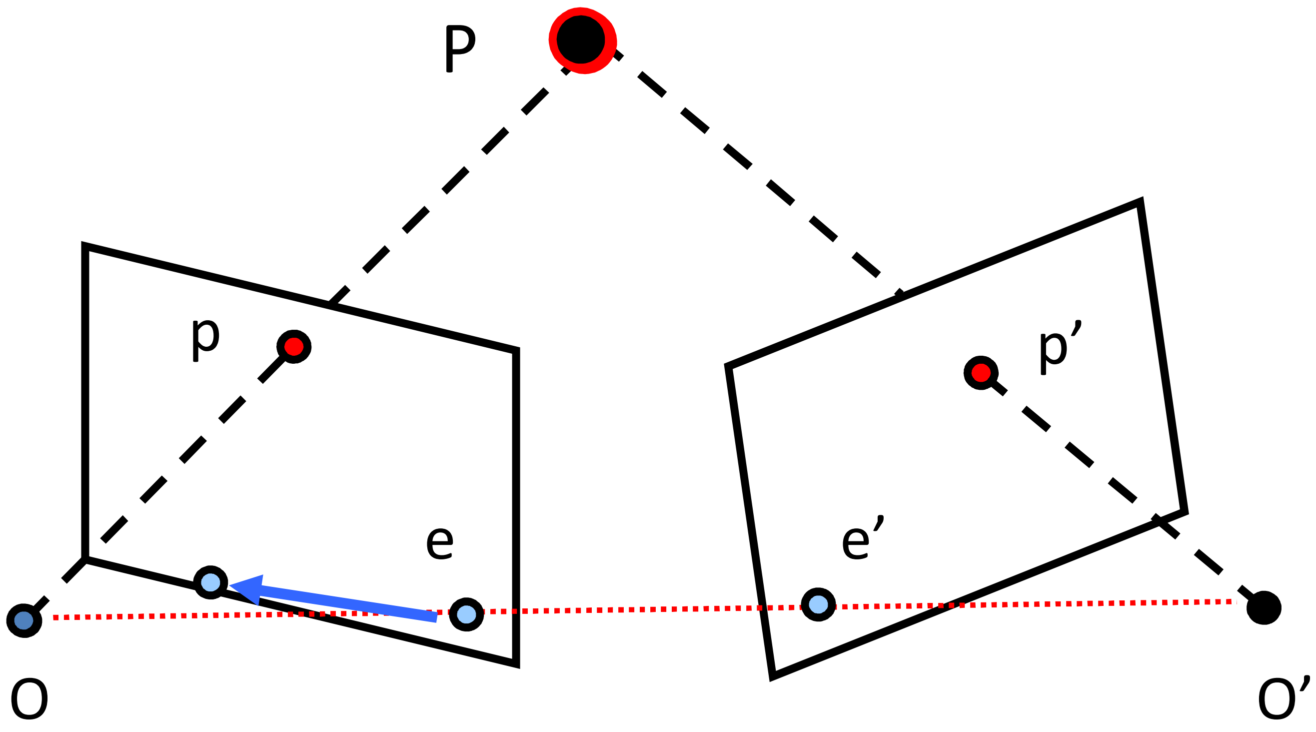
Making image planes parallel



0. Compute epipoles

$$e = K\, R^T\, T = [e_1 \quad e_2 \quad 1]^T \qquad e' = K'\, T$$

Making image planes parallel

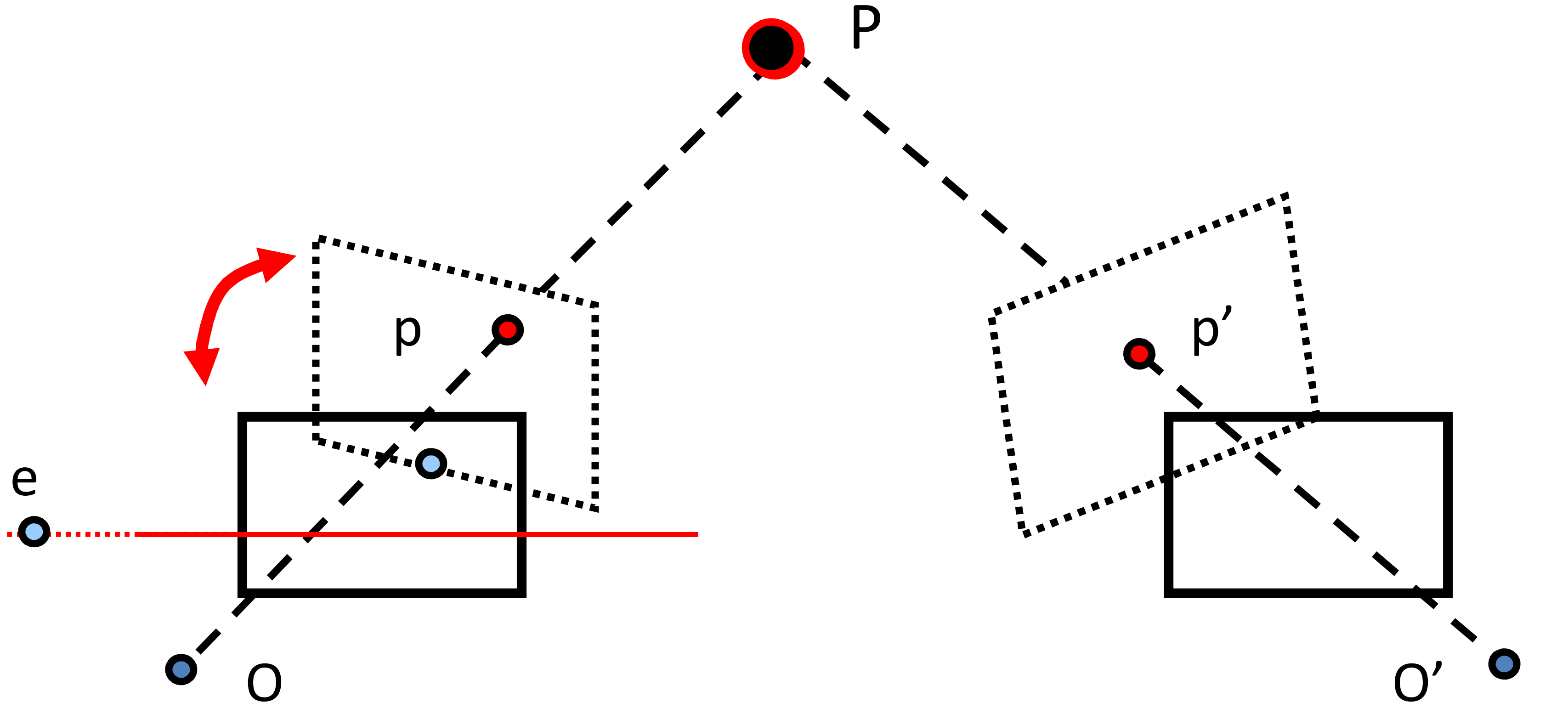


1. Map e to the x-axis at location $[1,0,1]^T$ (normalization)

$$e = [e_1 \quad e_2 \quad 1]^T \rightarrow [1 \quad 0 \quad 1]^T$$

$$\boxed{H_1 = R_H T_H}$$

Making image planes parallel



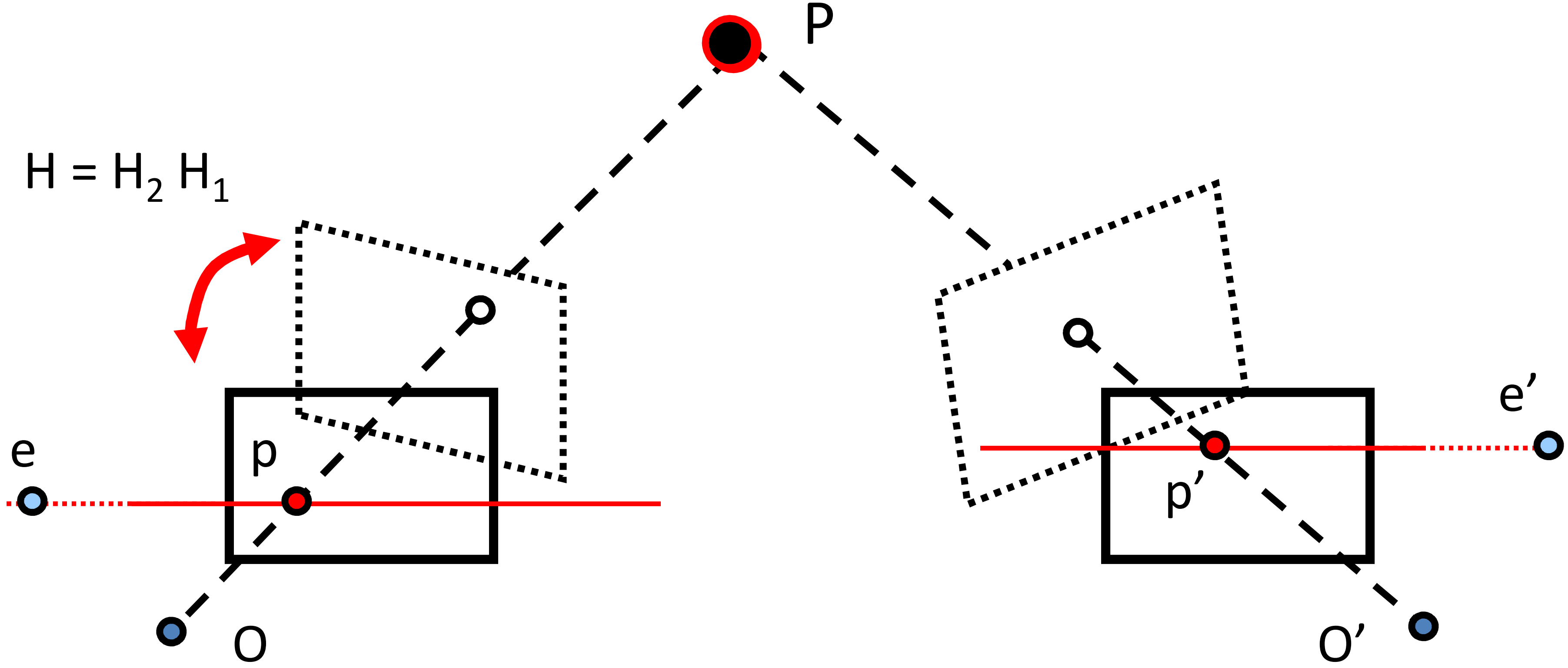
2. Send epipole to infinity:

$$e = \begin{bmatrix} 1 & 0 & 1 \end{bmatrix}^T \rightarrow \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^T$$

$$\boxed{H_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}}$$

Minimizes the distortion in a neighborhood (approximates id. mapping)

Making image planes parallel

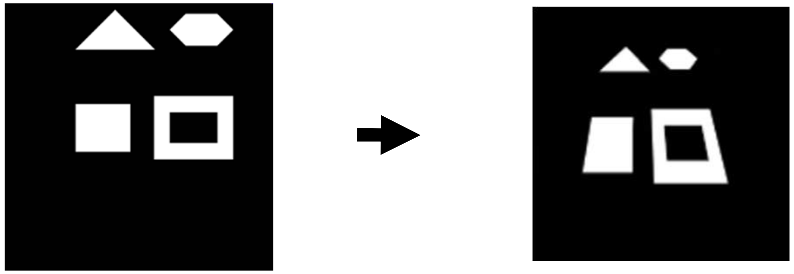


3. Define: $H = H_2 H_1$

4. Align epipolar lines

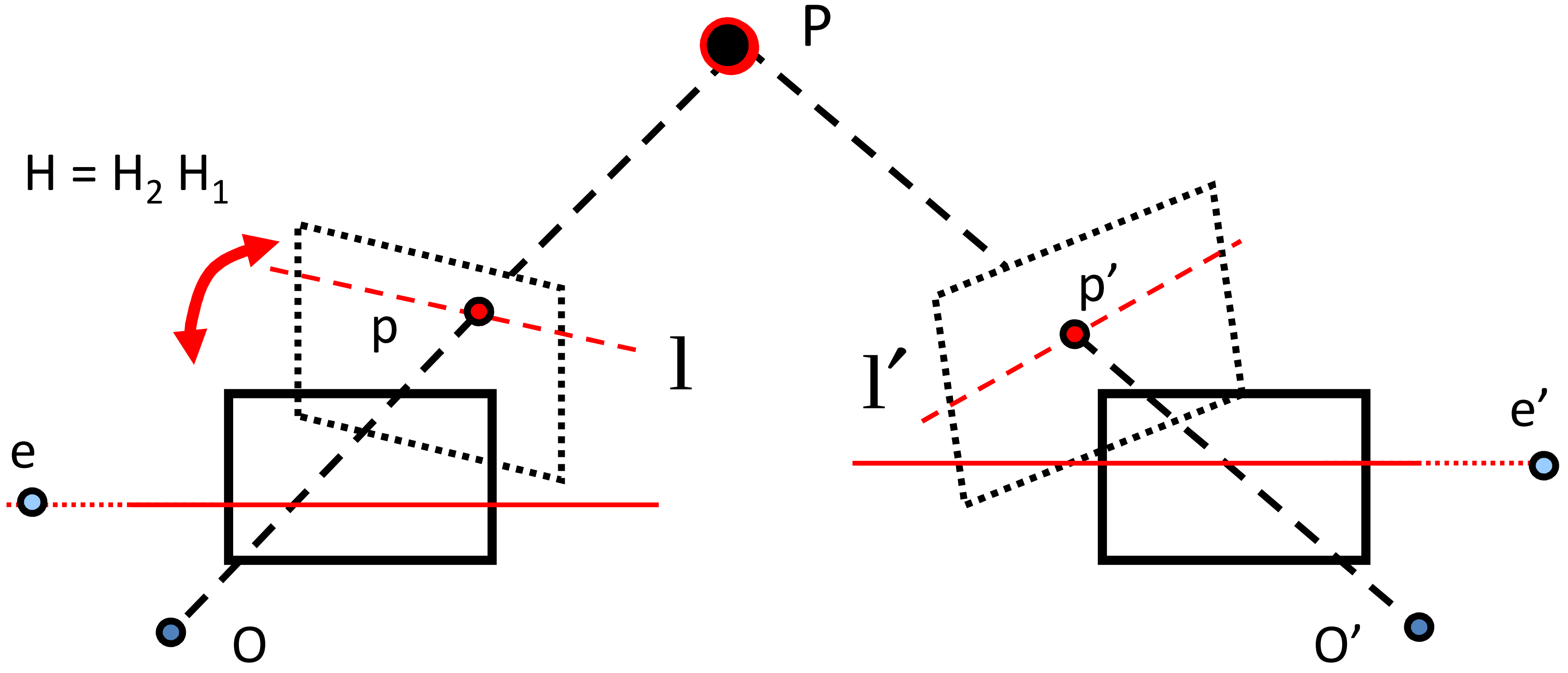
Projective transformation of a line (in 2D)

$$H = \begin{bmatrix} A & t \\ v & b \end{bmatrix}$$



$$l \rightarrow H^{-T}\, l$$

Making image planes parallel



3. Define: $H = H_2 H_1$

4. Align epipolar lines

[HZ] Chapters: 11 (sec. 11.12)

$$\overline{H'}^{-T} l' = \overline{H}^{-T} \, l$$

These are called matched pair of transformation