6.801/866

Multi-view geometry

T. Darrell

Multi-view geometry and 3-D

We have 2 eyes, yet we see 3-D!

Using multiple views allows inference of hidden dimension.

3-D: The hidden dimension…

Multiple views to the rescue!

How to see in 3-D

(Using geometry…)

- Find features
- Triangulate & reconstruct depth

Correspondence

Given a point in one image, find the point in a second image of the same 3-D location.

One of the hardest vision problems!

Next lecture: Algorithms for (quickly) estimating best correspondences.

Now: Where do we search? What are the constraints between images of 3-D points in multiple views?

Outline

- Multi-view geometry
- Epipolar constraint
- Essential matrix
- Fundamental matrix
- Trifocal tensor

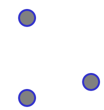
Multi-view geometry

Relate

Multi-view geometry

Relate

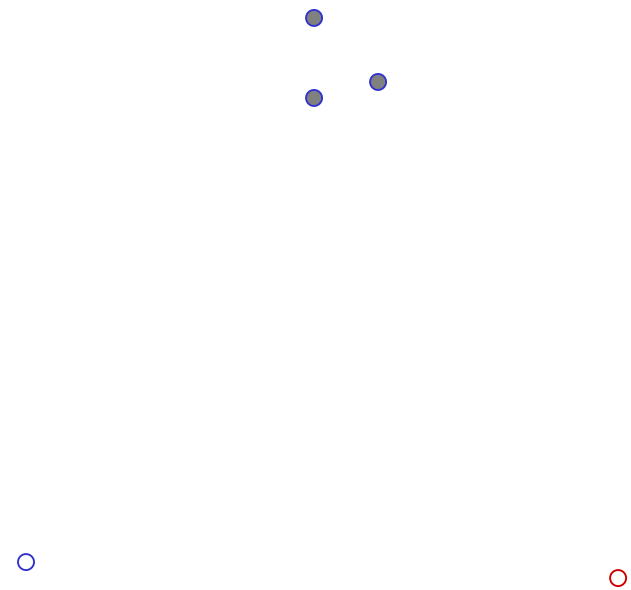
- 3-D points



Multi-view geometry

Relate

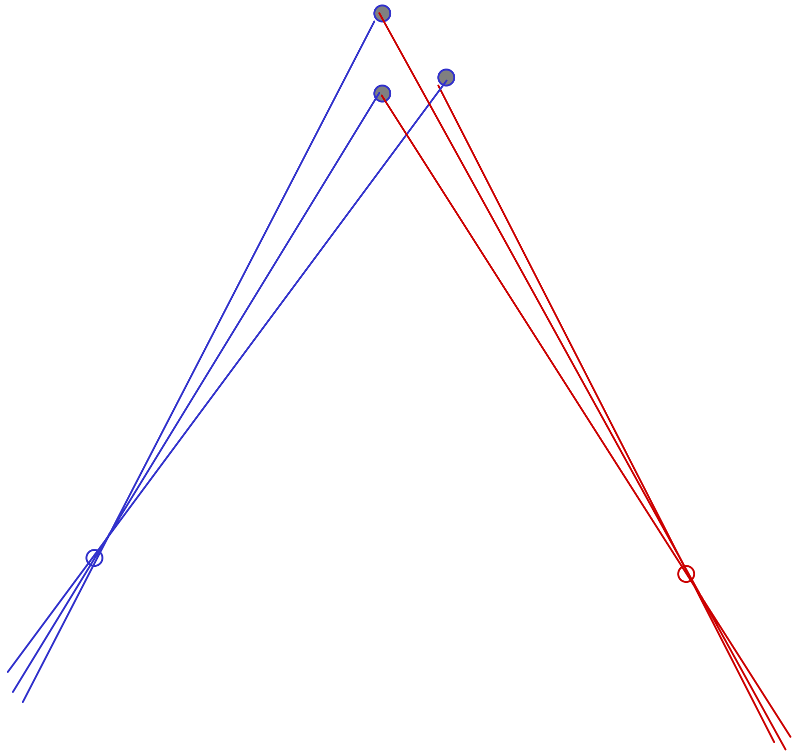
- 3-D points
- Camera centers



Multi-view geometry

Relate

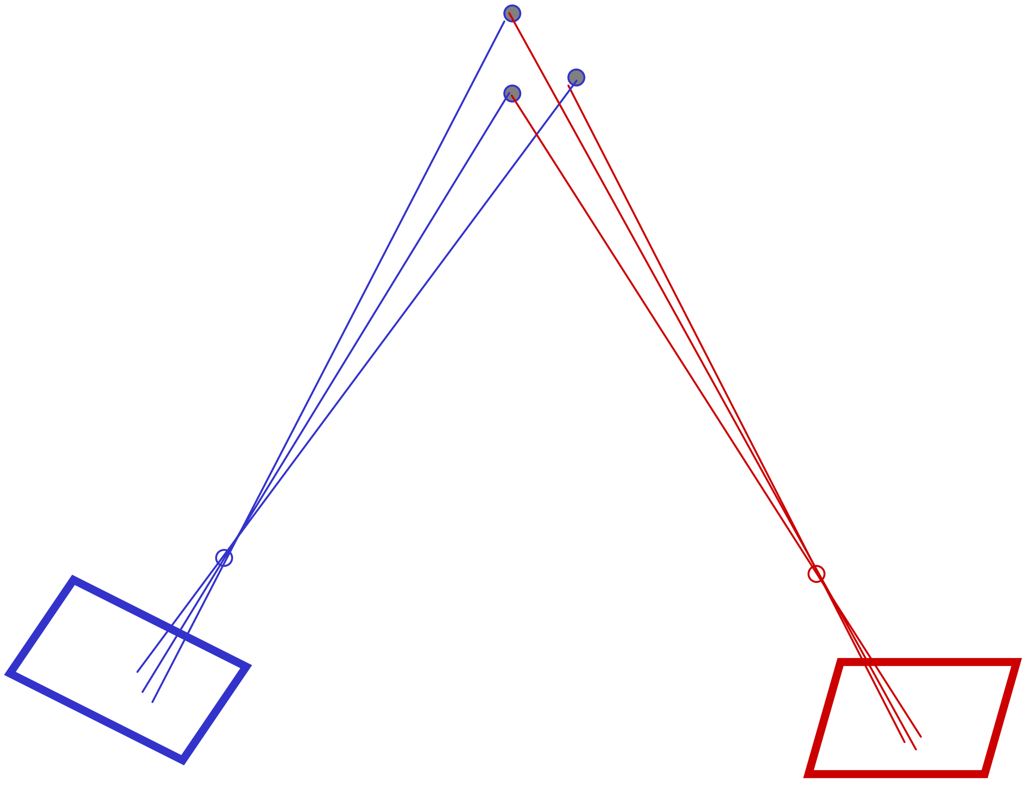
- 3-D points
- Camera centers
- Camera orientation



Multi-view geometry

Relate

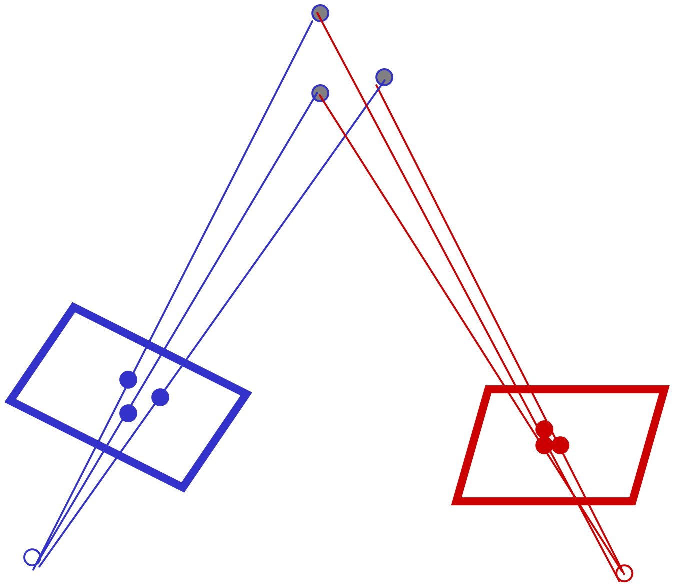
- 3-D points
- Camera centers
- Camera orientation
- Camera intrinsics



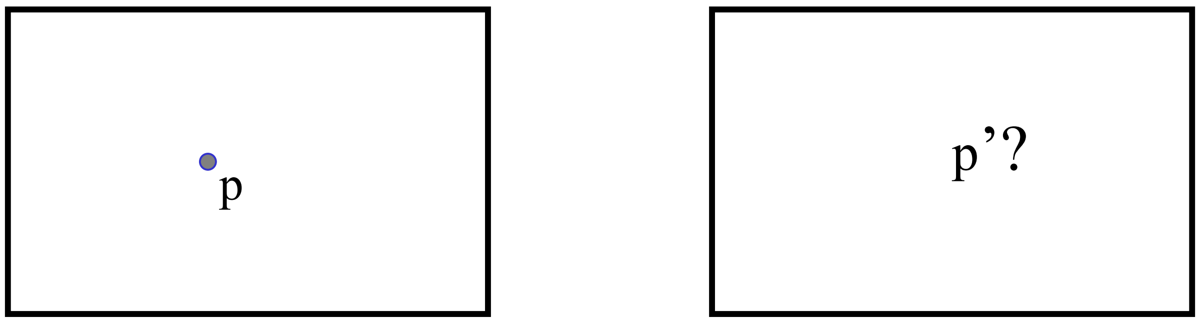
Multi-view geometry

Relate

- 3-D points
- Camera centers
- Camera orientation
- Camera intrinsics



Stereo constraints



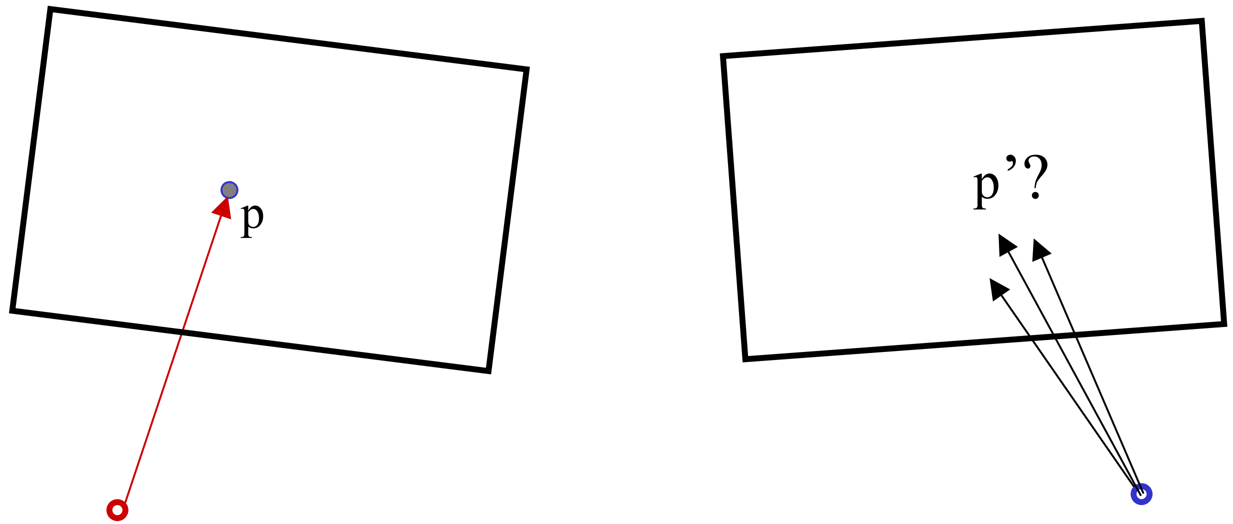
Given p in left image, where can corresponding point p' be?

Could be anywhere! Might not be same scene!

… Assume pair of pinhole views of static scene:

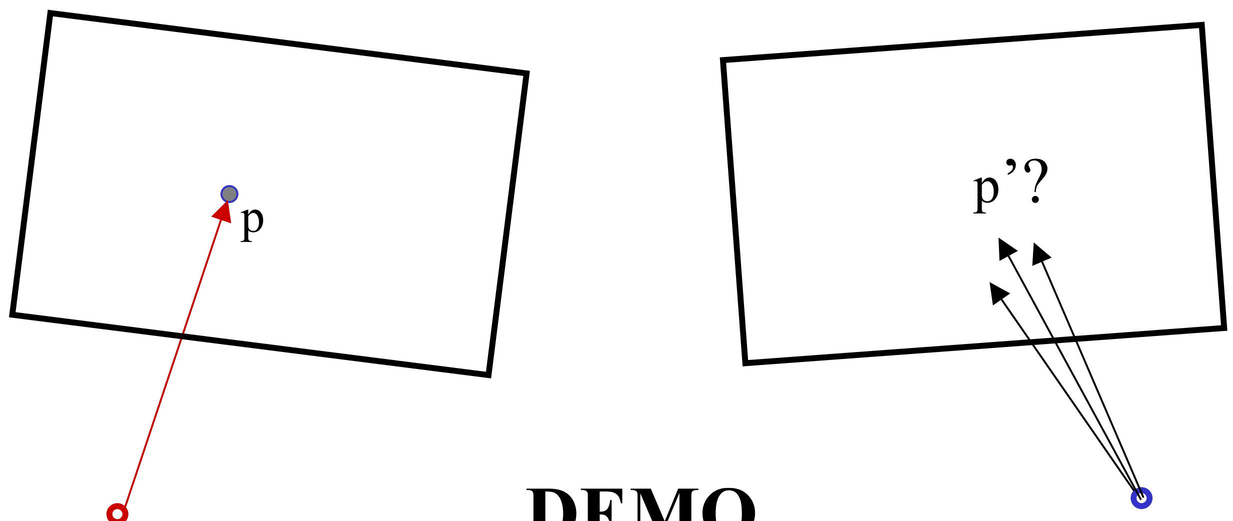
Stereo constraints

Given p in left image, where can p' be?



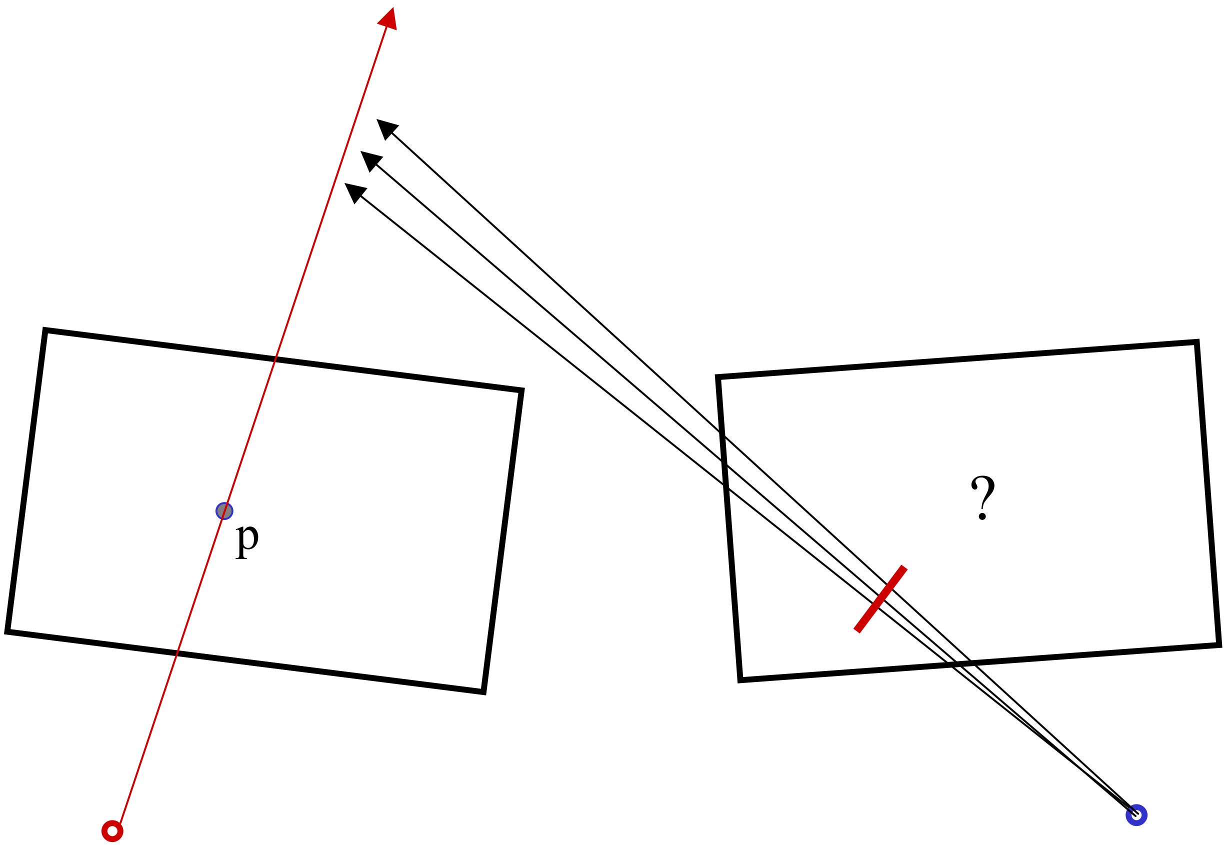
Stereo constraints

Given p in left image, where can p’ be?



DEMO

Epipolar line



Epipolar constraint

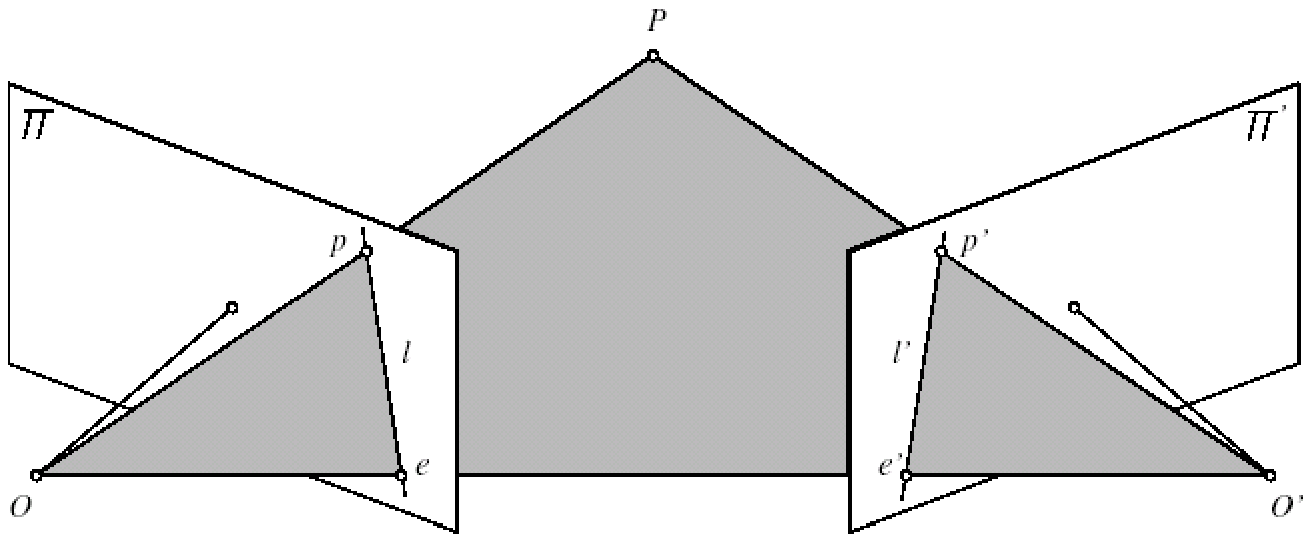


FIGURE 11.1: Epipolar geometry: the point P, the optical centers O and O' of the two cameras, and the two images p and p' of P all lie in the same plane.

All epipolar lines contain epipole, the image of other camera center.

From geometry to algebra…

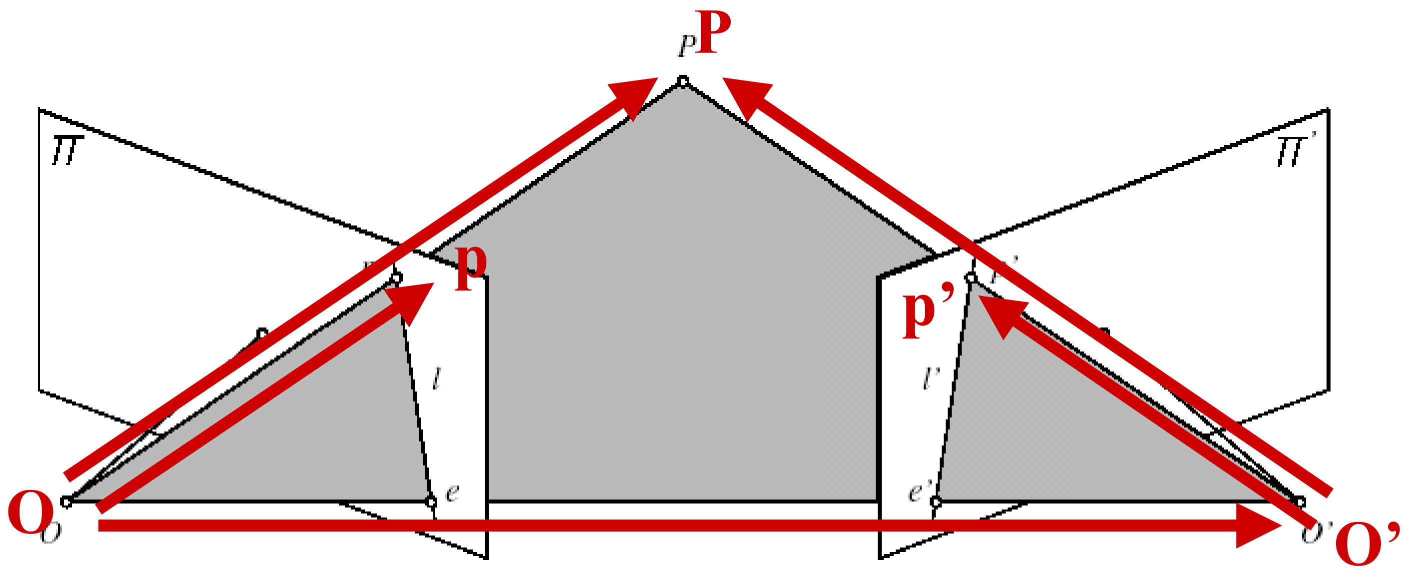
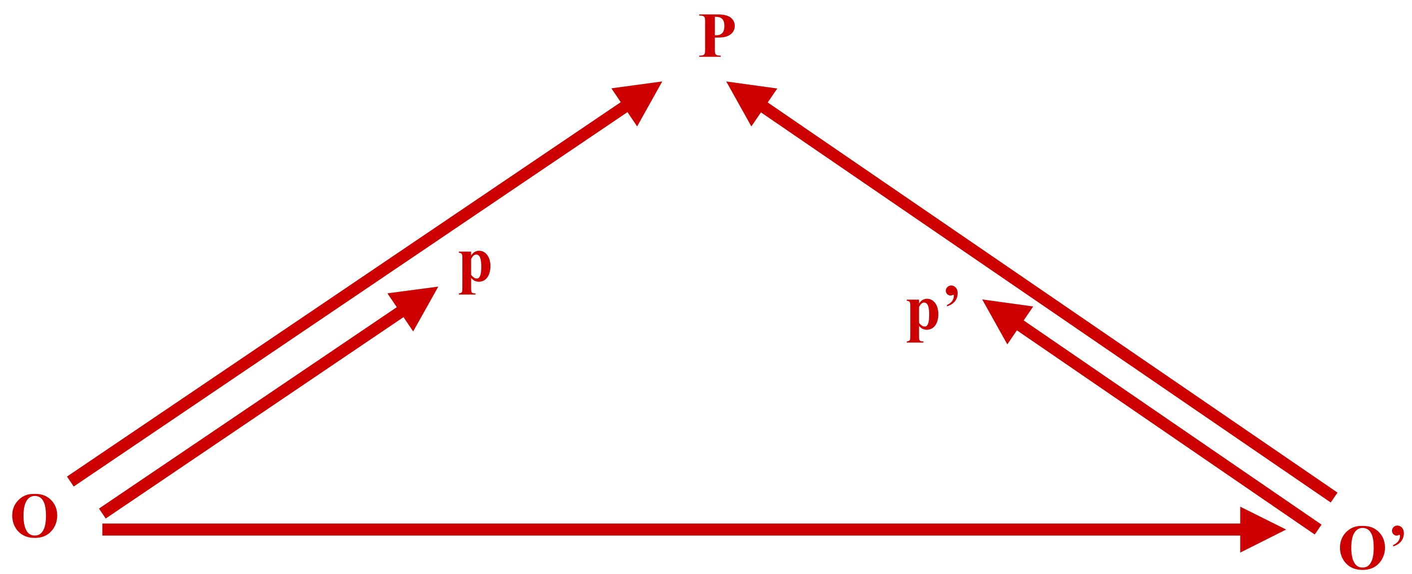


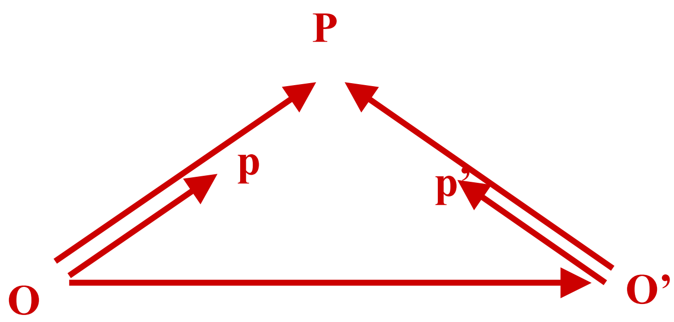
FIGURE 11.1: Epipolar geometry: the point P, the optical centers O and O' of the two cameras, and the two images p and p' of P all lie in the same plane.

From geometry to algebra…

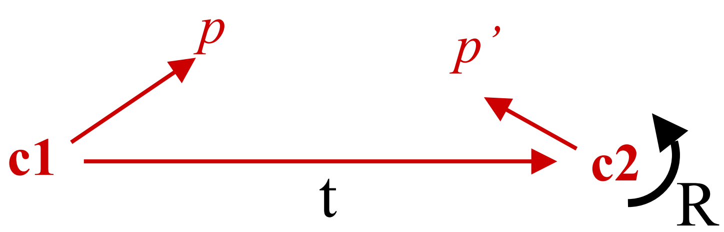


The epipolar constraint: these vectors are coplanar:

$$\overrightarrow{Op} \cdot [\overrightarrow{OO'} \times \overrightarrow{O'p'}] = 0$$



$$\overrightarrow{Op} \cdot [\overrightarrow{OO'} \times \overrightarrow{O'p'}] = 0$$



p,p' are image coordinates of P in c1 and c2…

c2 is related to c1 by rotation R and translation t

$$\boldsymbol{p} \cdot [\boldsymbol{t} \times (\mathcal{R}\boldsymbol{p}')] = 0$$

Matrix form

$$\boldsymbol{p} \cdot [\boldsymbol{t} \times (\mathcal{R}\boldsymbol{p}')] = 0$$

Linear constraint, should be able to express as matrix equation…

Review: matrix form of cross-product

The vector cross product also acts on two vectors and returns a third vector. Geometrically, this new vector is constructed such that its projection onto either of the two input vectors is zero.

$$\vec{a} \times \vec{b} = \begin{bmatrix} a_y b_z - a_z b_y \\ a_z b_x - a_x b_z \\ a_x b_y - a_y b_x \end{bmatrix}$$

$$\vec{a} \times \vec{b} = \begin{bmatrix} 0 & -a_z & a_y \\ a_z & 0 & -a_x \\ -a_y & a_x & 0 \end{bmatrix} \begin{bmatrix} b_x \\ b_y \\ b_z \end{bmatrix} = \vec{c} \qquad \begin{aligned} \vec{a} \cdot \vec{c} &= 0 \\ \vec{b} \cdot \vec{c} &= 0 \end{aligned}$$

Review: matrix form of cross-product

$$\vec{a} \times \vec{b} = \begin{bmatrix} 0 & -a_z & a_y \\ a_z & 0 & -a_x \\ -a_y & a_x & 0 \end{bmatrix} \begin{bmatrix} b_x \\ b_y \\ b_z \end{bmatrix} = \vec{c}$$

$$\vec{a} \cdot \vec{c} = 0$$

$$\vec{b} \cdot \vec{c} = 0$$

$$[a_x] = \begin{bmatrix} 0 & -a_z & a_y \\ a_z & 0 & -a_x \\ -a_y & a_x & 0 \end{bmatrix}$$

$$\boxed{\vec{a} \times \vec{b} = [a_x]\vec{b}}$$

Matrix form

$$\boldsymbol{p} \cdot [\boldsymbol{t} \times (\mathcal{R}\boldsymbol{p}')] = 0$$

$$\vec{a} \times \vec{b} = [a_x]\vec{b}$$

$$p^T [t_x] \Re p' = 0$$

$$\varepsilon = [t_x] \Re$$

$$\boxed{\boldsymbol{p}^T \mathcal{E} \boldsymbol{p}' = 0}$$

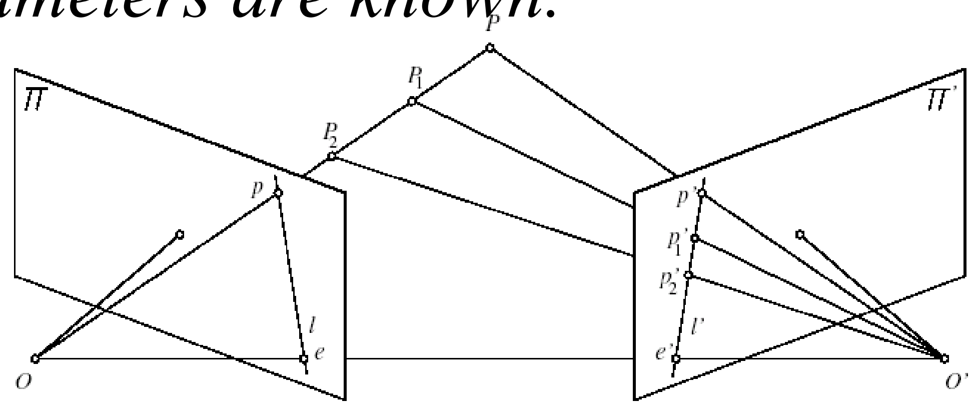
The Essential Matrix

Matrix that relates image of point in one camera to a second camera, given translation and rotation.

5 independent parameters (up to scale)

Assumes intrinsic parameters are known.

$$\varepsilon = [t_x]\Re$$



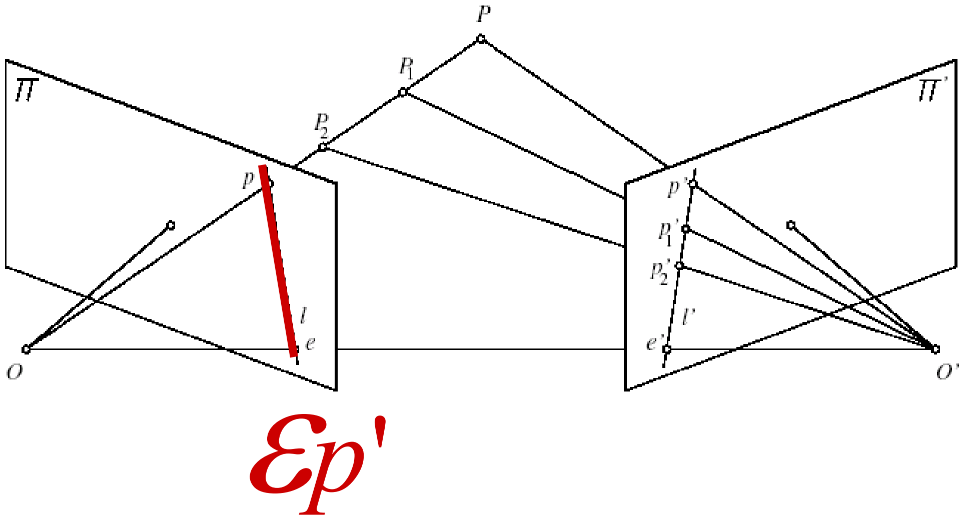
$$\boldsymbol{p}^T \mathcal{E} \boldsymbol{p}' = 0$$

$$\vec{a} \times \vec{b} = [a_x]\vec{b}$$

The Essential Matrix

$\mathcal{E}p'$ is the epipolar line corresponding to p' in the left camera.

$$au + bv + c = 0$$



$$p = (u, v, 1)^T$$

$$l = (a, b, c)^T$$

$$l \cdot p = 0$$

$$\mathcal{E}p' \cdot p = 0$$

$$\boldsymbol{p}^T \mathcal{E} \boldsymbol{p}' = 0$$

Essential matrix in instantanous case

For small motion given translation and rotation velocity:

$$\boldsymbol{t} = \delta t\, \boldsymbol{v},$$
$$\mathcal{R} = \mathrm{Id} + \delta t\, [\boldsymbol{\omega}_\times]$$
$$\boldsymbol{p}' = \boldsymbol{p} + \delta t\, \dot{\boldsymbol{p}}.$$

$$\boldsymbol{p}^T [\boldsymbol{v}_\times](\mathrm{Id} + \delta t\, [\boldsymbol{\omega}_\times])(\boldsymbol{p} + \delta t\, \dot{\boldsymbol{p}}) = 0$$

FOE for translating camera

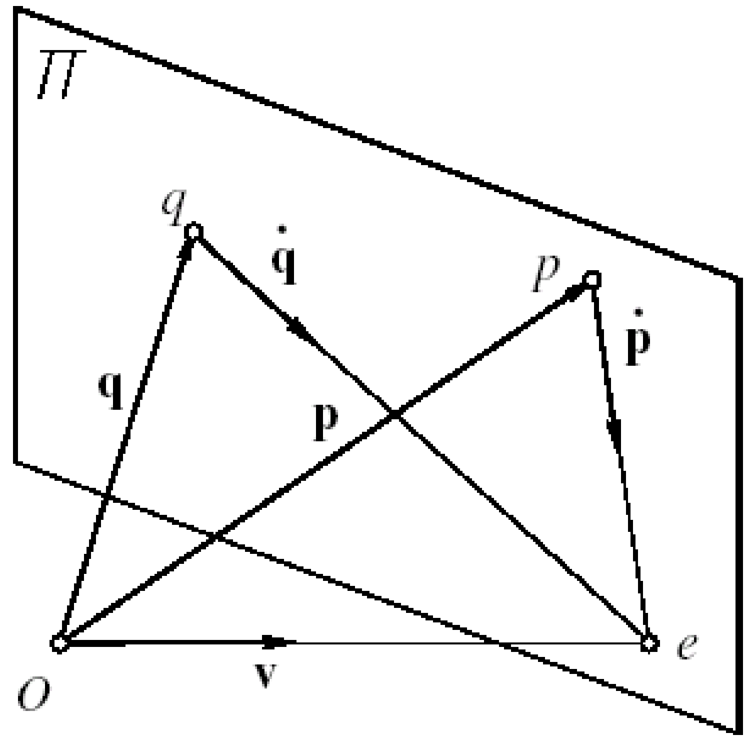
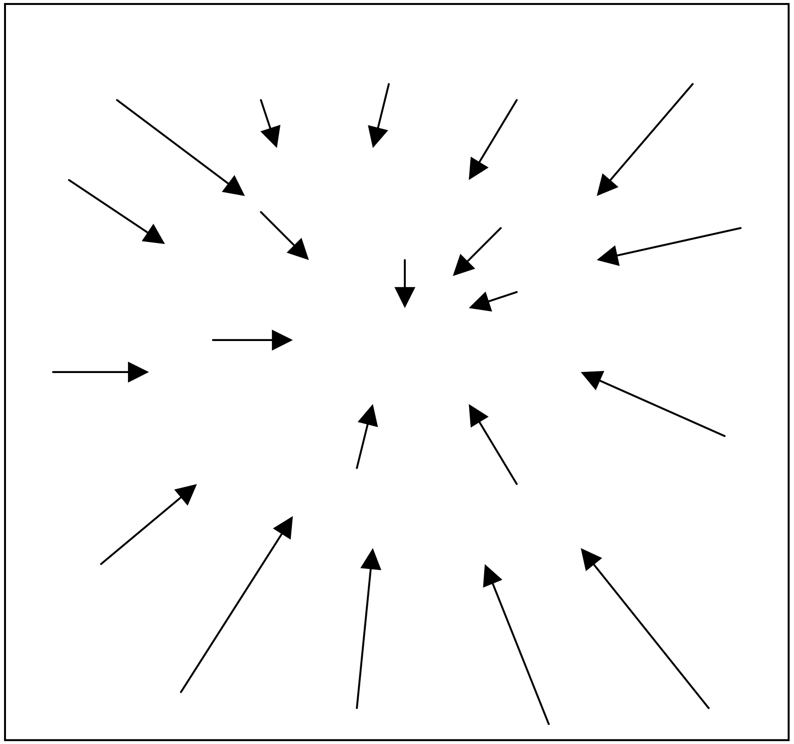


FIGURE 11.3: Focus of expansion: under pure translation, the motion field at every point in the image points toward the focus of expansion.

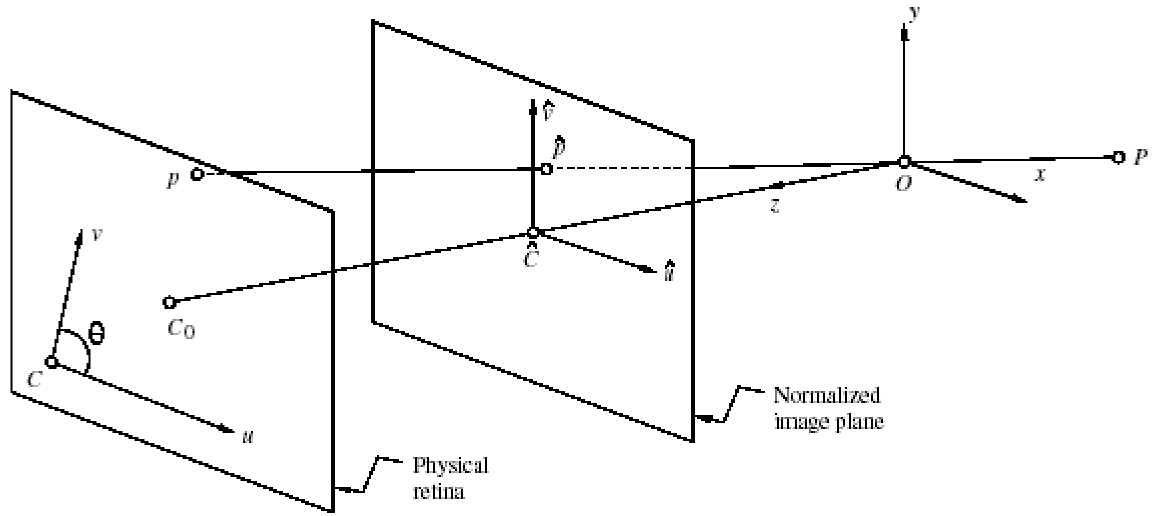
FOE for translating camera



What if calibration is unknown?

Recall calibration eqn:

$$\boldsymbol{p} = \mathcal{K}\hat{\boldsymbol{p}}, \quad \text{where} \quad \boldsymbol{p} = \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} \quad \text{and} \quad \mathcal{K} \stackrel{\text{def}}{=} \begin{pmatrix} \alpha & -\alpha\cot\theta & u_0 \\ 0 & \dfrac{\beta}{\sin\theta} & v_0 \\ 0 & 0 & 1 \end{pmatrix}.$$



Fundamental matrix

Essential matrix for points on normalized image plane,

$$\hat{p}^T \mathcal{E} \hat{p}' = 0$$

assume unknown calibration matrix:

$$p = K\hat{p}$$

yields:

$$\boxed{\boldsymbol{p}^T \mathcal{F} \boldsymbol{p}' = 0} \qquad \mathcal{F} = \mathcal{K}^{-T} \mathcal{E} \mathcal{K}'^{-1}$$

Estimating the Fundamental Matrix

$$\boldsymbol{p}^T \mathcal{F} \boldsymbol{p}' = 0$$

Each point correspondence can be expressed as a single linear equation

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

Estimating the Fundamental Matrix

$$\boldsymbol{p}^T \mathcal{F} \boldsymbol{p}' = 0$$

Each point correspondence can be expressed as a single linear equation

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

Estimating the Fundamental Matrix

How many correspondences are needed to estimate $\mathcal{F}$?

$\mathcal{E}$ has 5 independent parameters up to scale.

In principle $\mathcal{F}$ has 7 independent parameters up to scale, and can be estimated from 7 point correspondences.

Direct, simpler method uses 8 correspondences….

The 8 point algorithm

8 corresponding points, 8 equations.

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

Invert and solve for $\mathcal{F}$.

(Use more points if available; find least-squares solution to minimize $\sum_{i=1}^{n} (\boldsymbol{p}_i^T \mathcal{F} \boldsymbol{p}_i')^2$)

The 8 point algorithm

$$\boldsymbol{p}^T \mathcal{F} \boldsymbol{p}' = 0$$

is $\mathcal{F}$ (or $\mathcal{E}$) full rank?

No…singular with rank=2.
Has zero eigenvalue corresponding to epipole.

$$\mathcal{F}^T e = 0$$

(Note that $\mathcal{E}$ has two equal singlar values [Huang and Faugeras 1989])

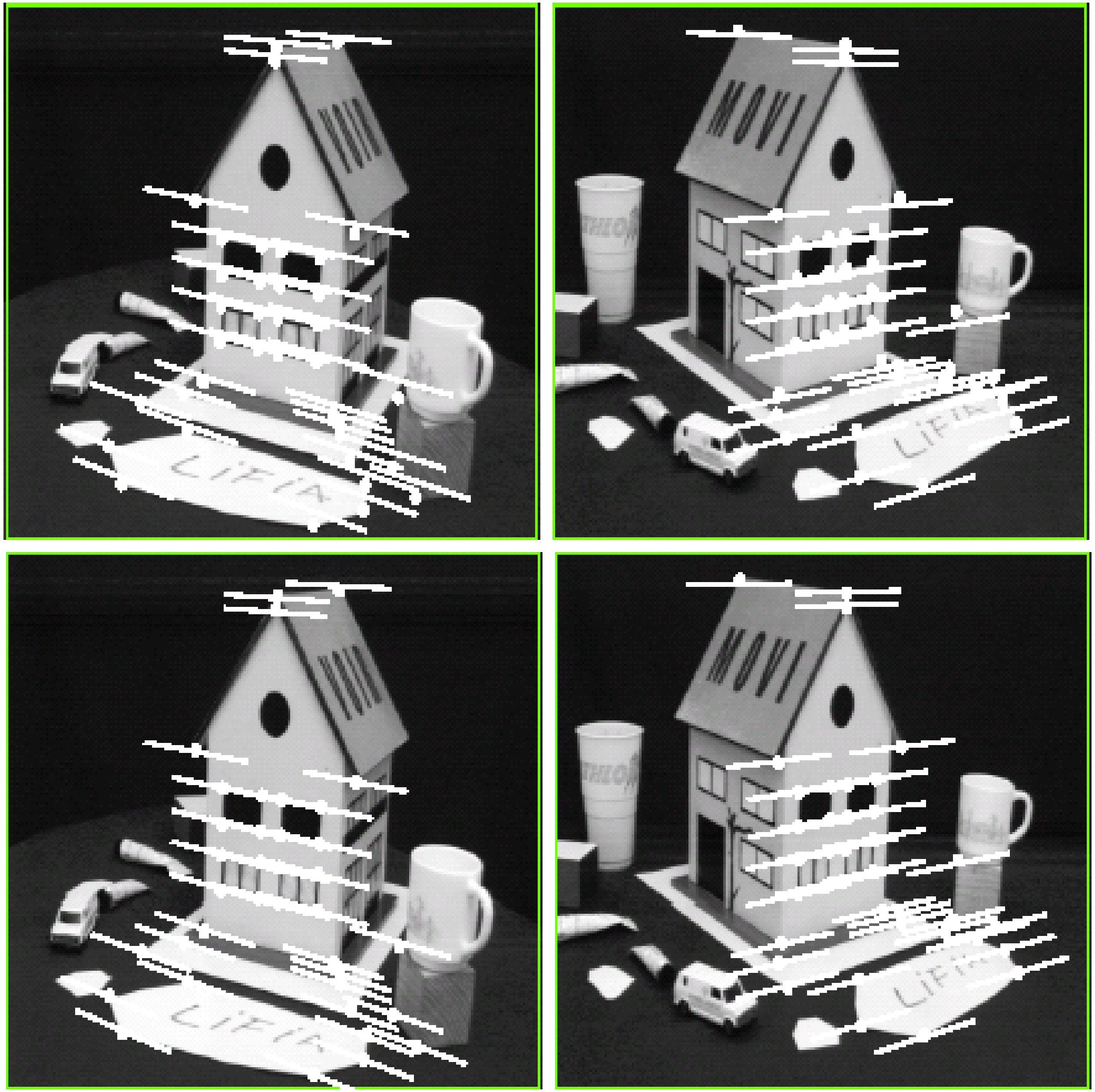
Improved 8 point algorithm

Enforce rank 2 constraint!

(Also pay attention to numerical conditioning…)

Hartley 1995: use SVD.

1. Transform to centered and scaled coordinates
2. Form least-squares estimate of F
3. Set smallest singular value to zero.

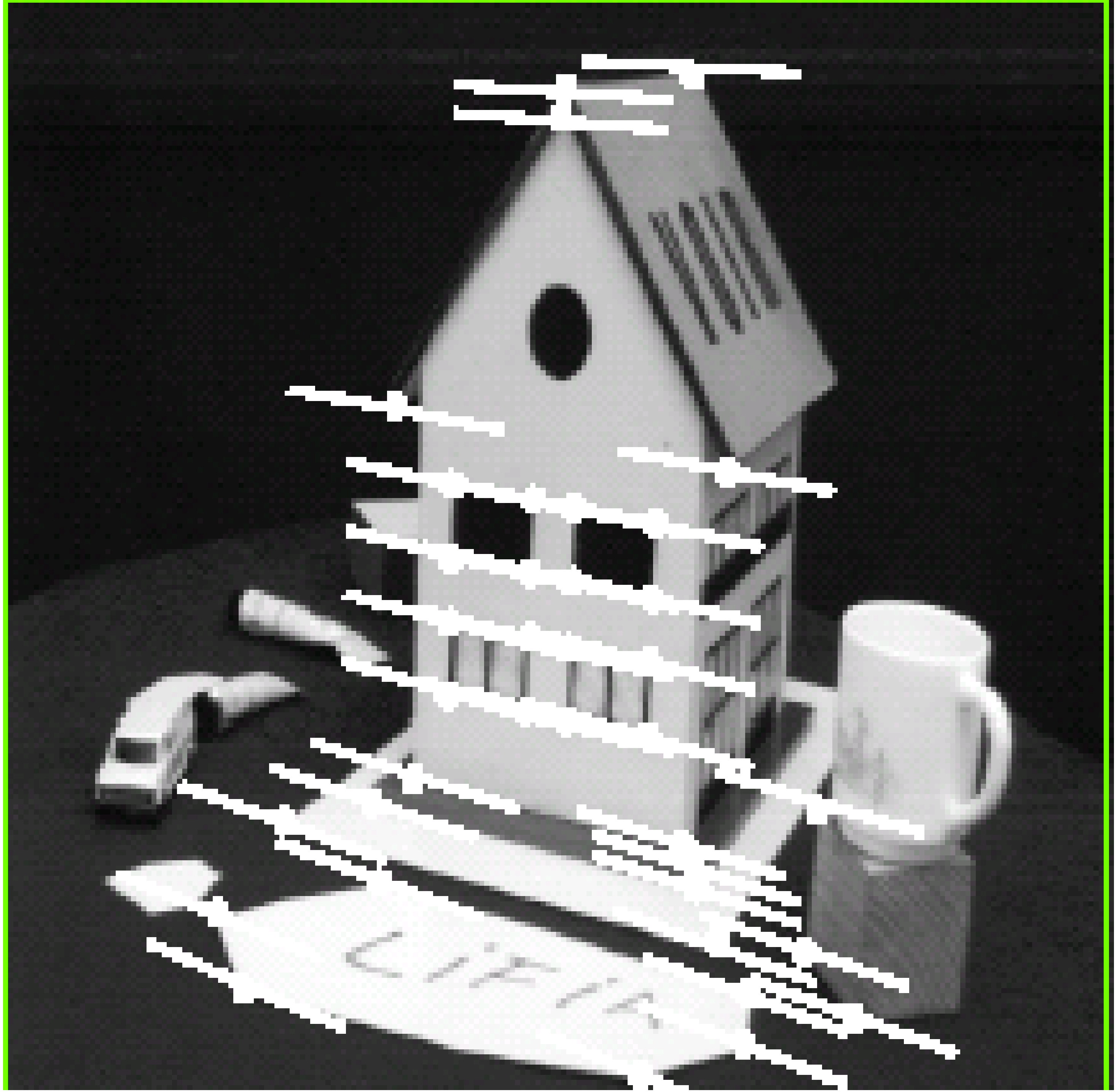


(a)

(b)

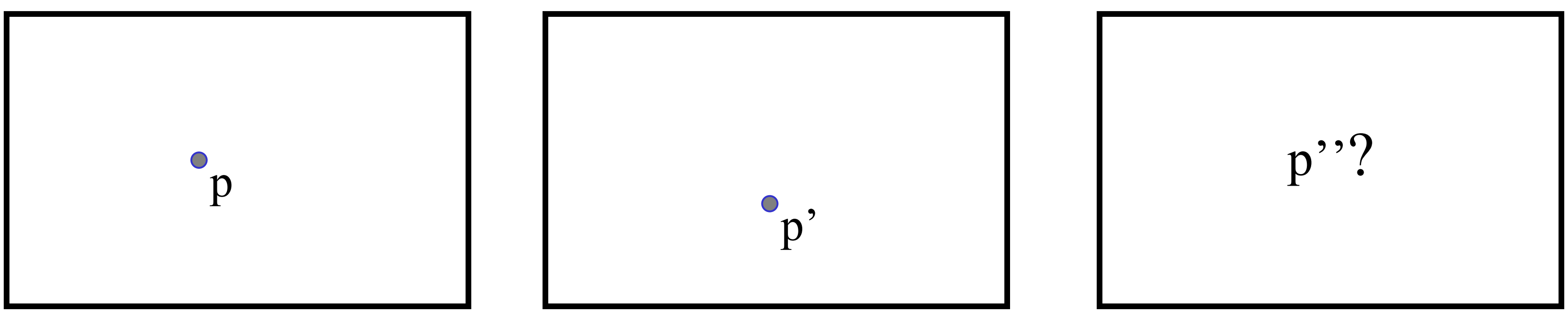
	Linear Least Squares	[Hartley, 1995]	[Luong *et al.*, 1993]
Av. Dist. 1	2.33 pixels	0.92 pixel	0.86 pixel
Av. Dist. 2	2.18 pixels	0.85 pixel	0.80 pixel


LIFIA



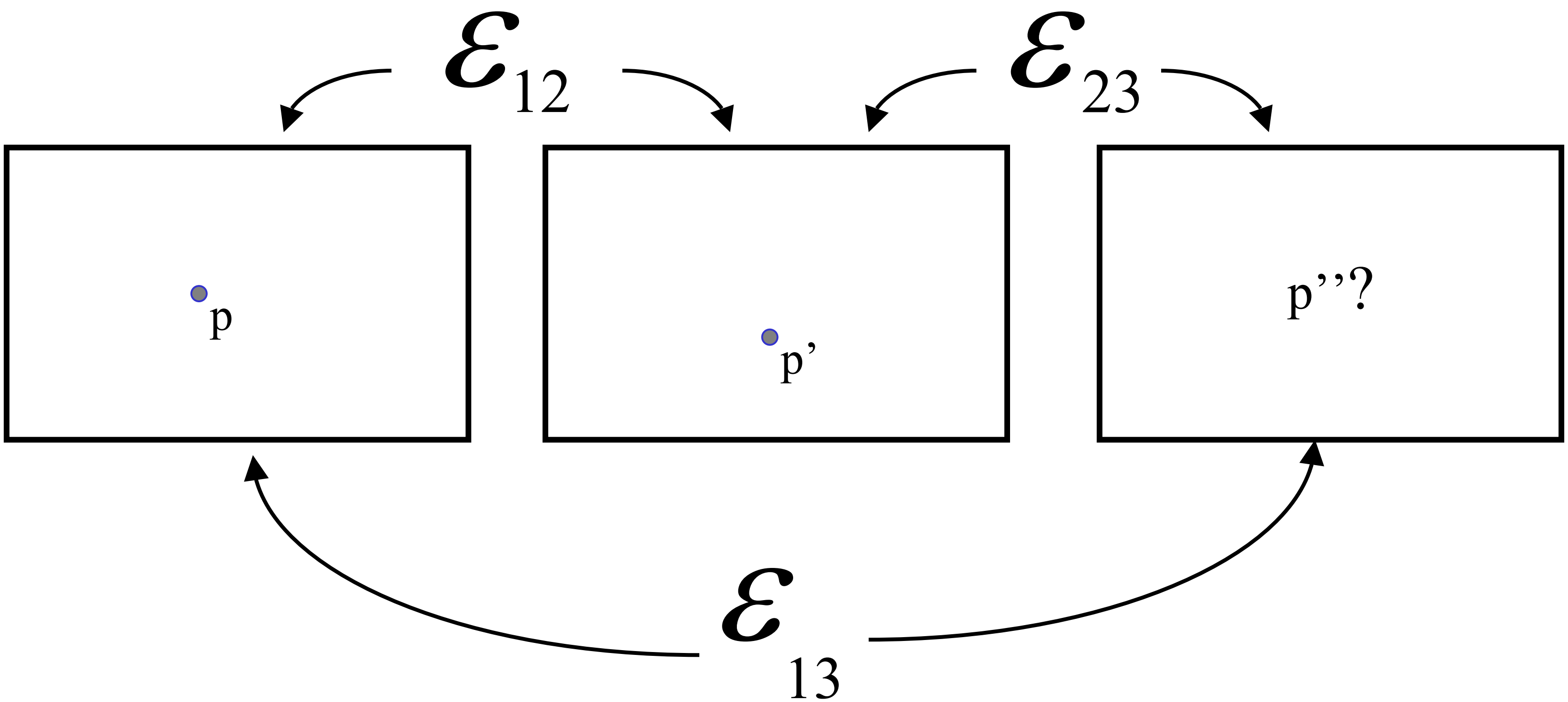
Stereo constraints

Given p’,p’’ in left and middle image, where is p’’ in a third view?



Three essential matrices

Essential matrices relate each pair:
(calibrated case)



Three essential matrices

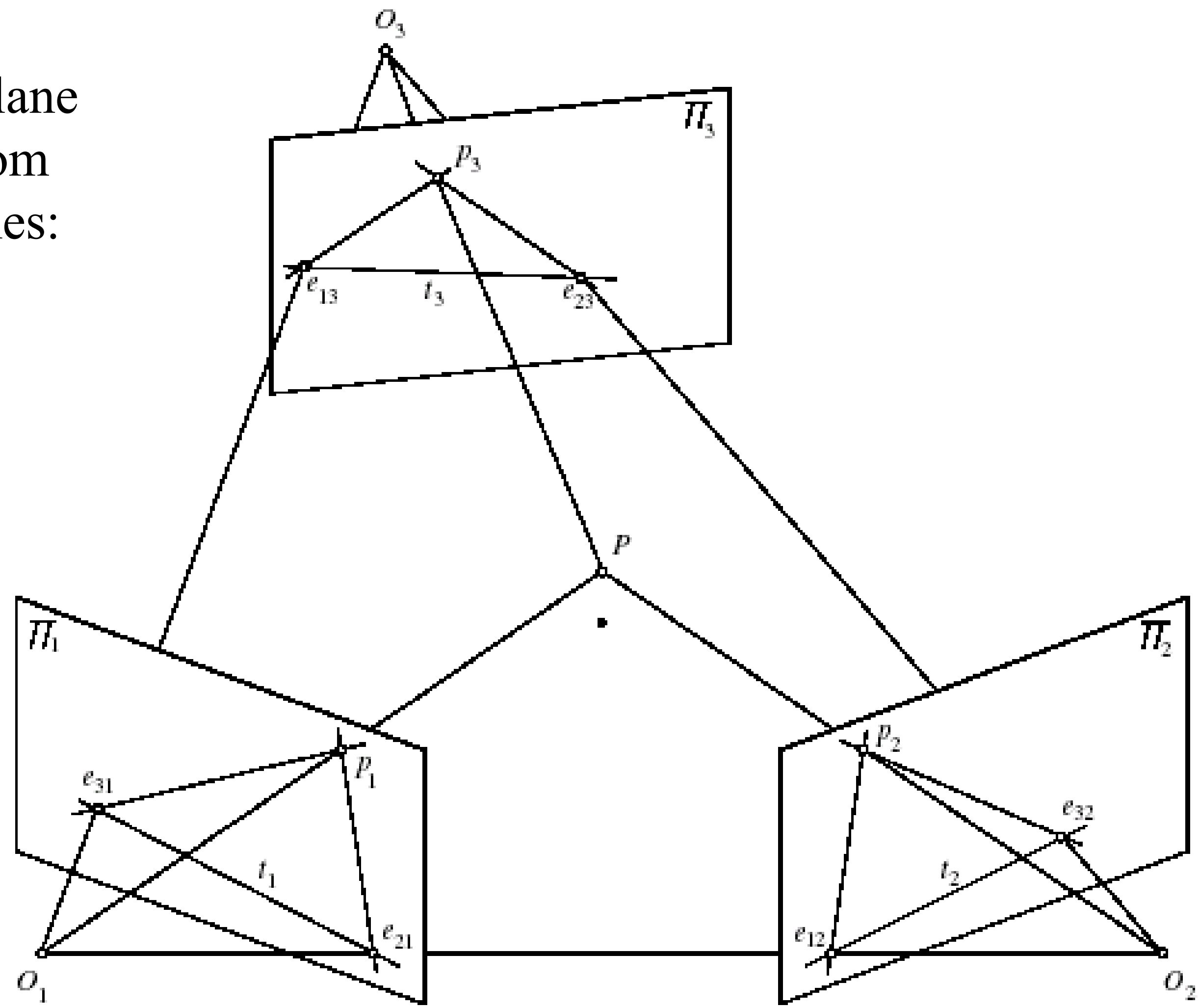
$$\begin{cases} \boldsymbol{p}_1^T \mathcal{E}_{12} \boldsymbol{p}_2 = 0, \\ \boldsymbol{p}_2^T \mathcal{E}_{23} \boldsymbol{p}_3 = 0, \\ \boldsymbol{p}_3^T \mathcal{E}_{31} \boldsymbol{p}_1 = 0, \end{cases}$$

any two are independent!

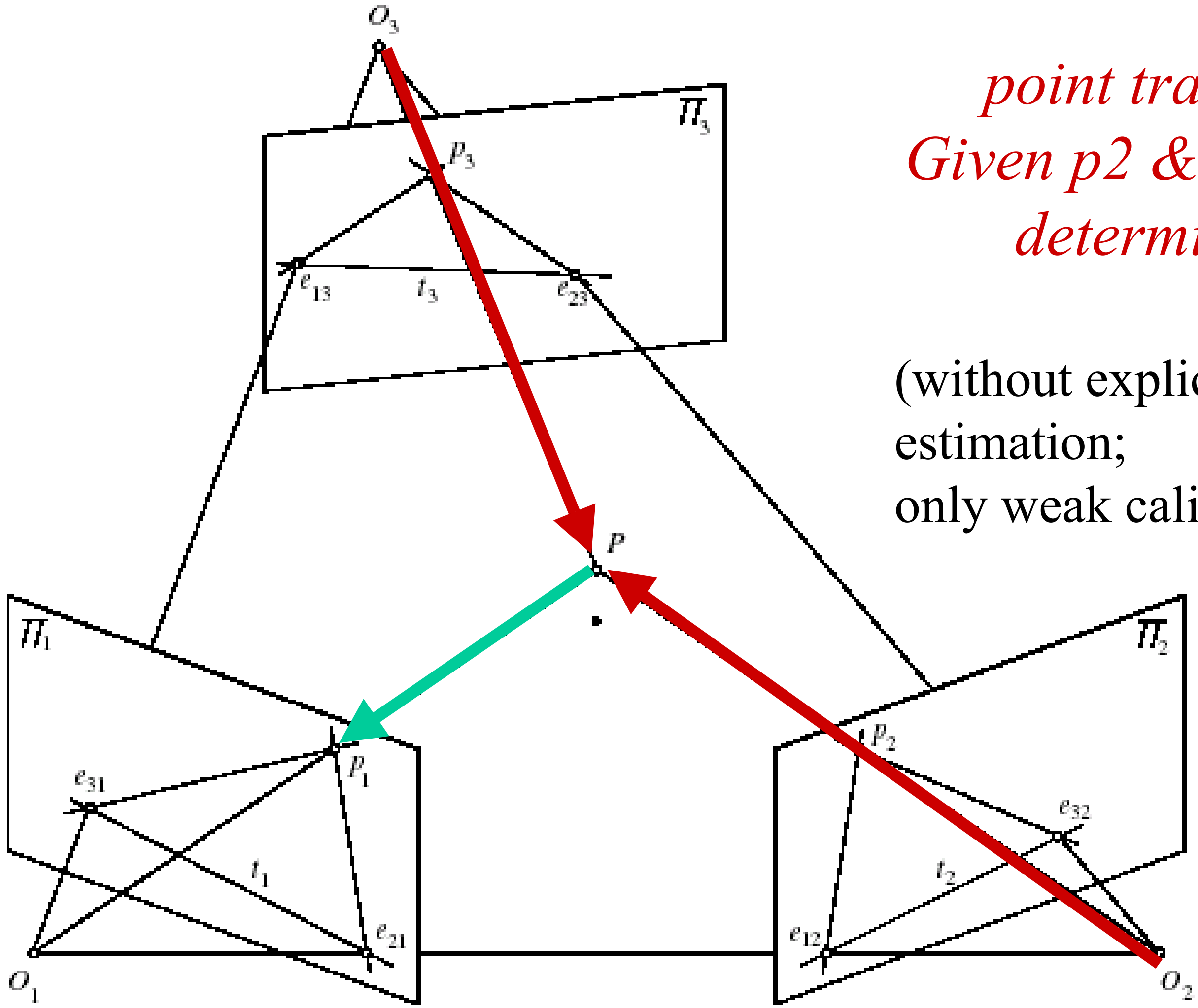
can predict third point from two others.

Trinocular epipolar geometry

Trifocal plane formed from trifocal lines:



Trinocular epipolar geometry

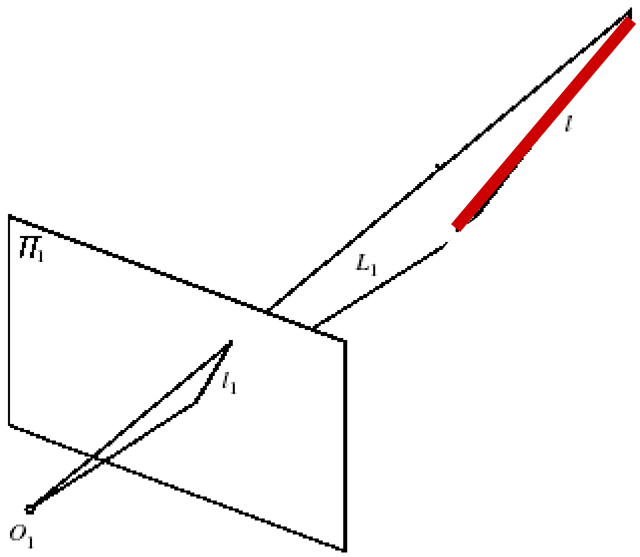


point transfer:
Given p2 & p3, p1 is determined!

(without explicit depth estimation;
only weak calibration)

Trifocal line constraint

Form the plane containing a line l and optical center of one camera:



$$\boldsymbol{l}^T \mathcal{M} \boldsymbol{P} = 0,$$

Trifocal line constraint

3 cameras, 3 plane equations:

L=M^Tl

$$\begin{pmatrix} \boldsymbol{L}_1^T \\ \boldsymbol{L}_2^T \\ \boldsymbol{L}_3^T \end{pmatrix} \boldsymbol{P} = \boldsymbol{0}$$

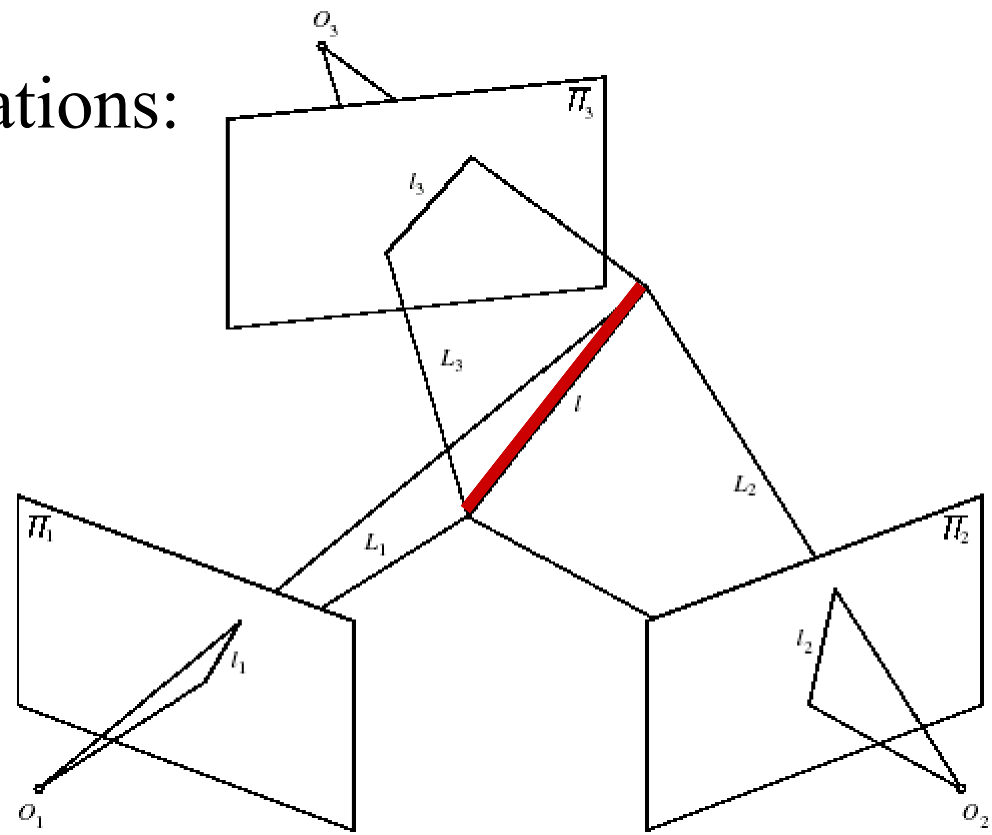


Figure 12.6. Three images of a line define it as the intersection of three planes in t
same pencil.

Trifocal line constraint

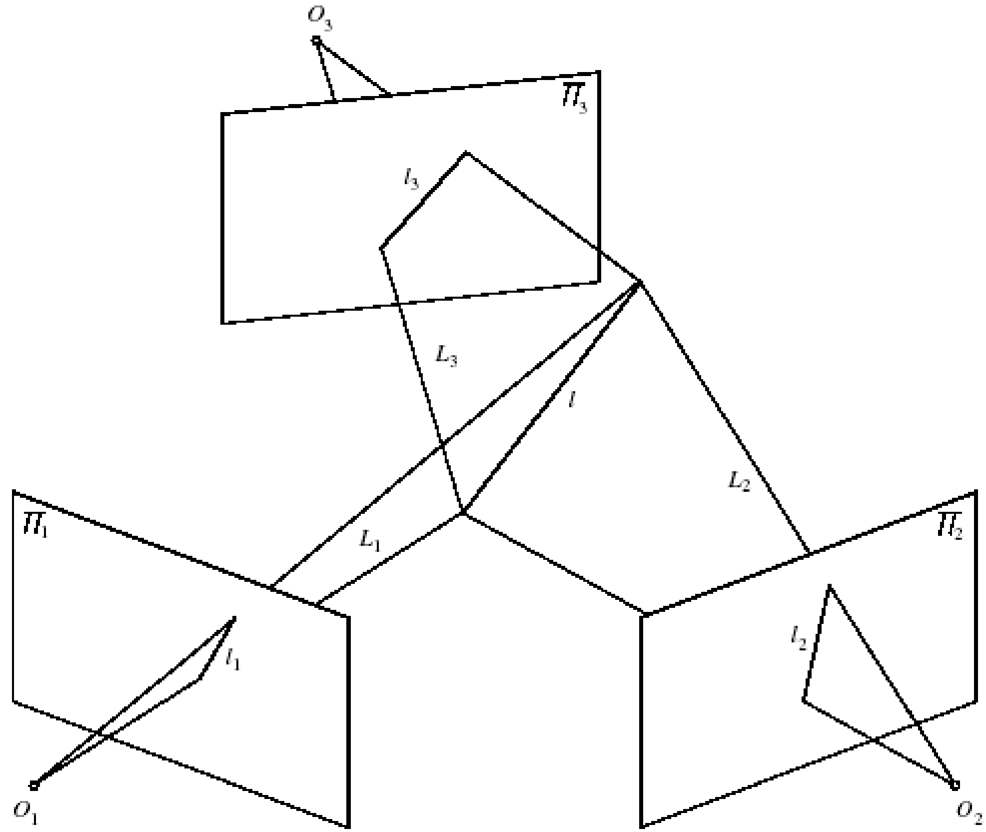
$$\boldsymbol{l}^T \mathcal{M} \boldsymbol{P} = 0,$$

$$\boldsymbol{L} = \boldsymbol{M}^T \boldsymbol{l}$$

$$\begin{pmatrix} \boldsymbol{L}_1^T \\ \boldsymbol{L}_2^T \\ \boldsymbol{L}_3^T \end{pmatrix} \boldsymbol{P} = \boldsymbol{0}$$

$$\mathcal{L} \stackrel{\text{def}}{=} \begin{pmatrix} \boldsymbol{l}_1^T \mathcal{M}_1 \\ \boldsymbol{l}_2^T \mathcal{M}_2 \\ \boldsymbol{l}_3^T \mathcal{M}_3 \end{pmatrix}$$

If 3 lines intersect in more than one point (a line) this system is degenerate and is rank 2.



Rank of $\mathcal{L} \stackrel{\text{def}}{=} \begin{pmatrix} \boldsymbol{l}_1^T \mathcal{M}_1 \\ \boldsymbol{l}_2^T \mathcal{M}_2 \\ \boldsymbol{l}_3^T \mathcal{M}_3 \end{pmatrix} = 2$

Trifocal line constraint

Assume calibrated camera coordinates

$$\mathcal{M}_1 = \begin{pmatrix} \mathrm{Id} & \mathbf{0} \end{pmatrix}$$

$$\mathcal{M}_2 = \begin{pmatrix} \mathcal{R}_2 & \boldsymbol{t}_2 \end{pmatrix}$$

$$\mathcal{M}_3 = \begin{pmatrix} \mathcal{R}_3 & \boldsymbol{t}_3 \end{pmatrix}$$

then

$$\mathcal{L} = \begin{pmatrix} \boldsymbol{l}_1^T & 0 \\ \boldsymbol{l}_2^T \mathcal{R}_2 & \boldsymbol{l}_2^T \boldsymbol{t}_2 \\ \boldsymbol{l}_3^T \mathcal{R}_3 & \boldsymbol{l}_3^T \boldsymbol{t}_3 \end{pmatrix}$$

$$\mathcal{L} = \begin{pmatrix} \boldsymbol{l}_1^T & 0 \\ \boldsymbol{l}_2^T \mathcal{R}_2 & \boldsymbol{l}_2^T \boldsymbol{t}_2 \\ \boldsymbol{l}_3^T \mathcal{R}_3 & \boldsymbol{l}_3^T \boldsymbol{t}_3 \end{pmatrix}$$

Rank $\mathcal{L} = 2$ means det. of 3x3 minors are zero, and can be expressed as:

$$\boldsymbol{l}_1 \times \begin{pmatrix} \boldsymbol{l}_2^T \mathcal{G}_1^1 \boldsymbol{l}_3 \\ \boldsymbol{l}_2^T \mathcal{G}_1^2 \boldsymbol{l}_3 \\ \boldsymbol{l}_2^T \mathcal{G}_1^3 \boldsymbol{l}_3 \end{pmatrix} = \mathbf{0},$$

with

$$\mathcal{G}_1^i = \boldsymbol{t}_2 \boldsymbol{R}_3^{iT} - \boldsymbol{R}_2^i \boldsymbol{t}_3^T$$

The trifocal tensor

These 3 3x3 matrices are called the trifocal tensor.

$$\mathcal{G}_1^i = \boldsymbol{t}_2 \boldsymbol{R}_3^{iT} - \boldsymbol{R}_2^i \boldsymbol{t}_3^T$$

the constraint

$$\boldsymbol{l}_1 \times \begin{pmatrix} \boldsymbol{l}_2^T \mathcal{G}_1^1 \boldsymbol{l}_3 \\ \boldsymbol{l}_2^T \mathcal{G}_1^2 \boldsymbol{l}_3 \\ \boldsymbol{l}_2^T \mathcal{G}_1^3 \boldsymbol{l}_3 \end{pmatrix} = \mathbf{0},$$

can be used for point or line transfer.

Trifocal line constraint

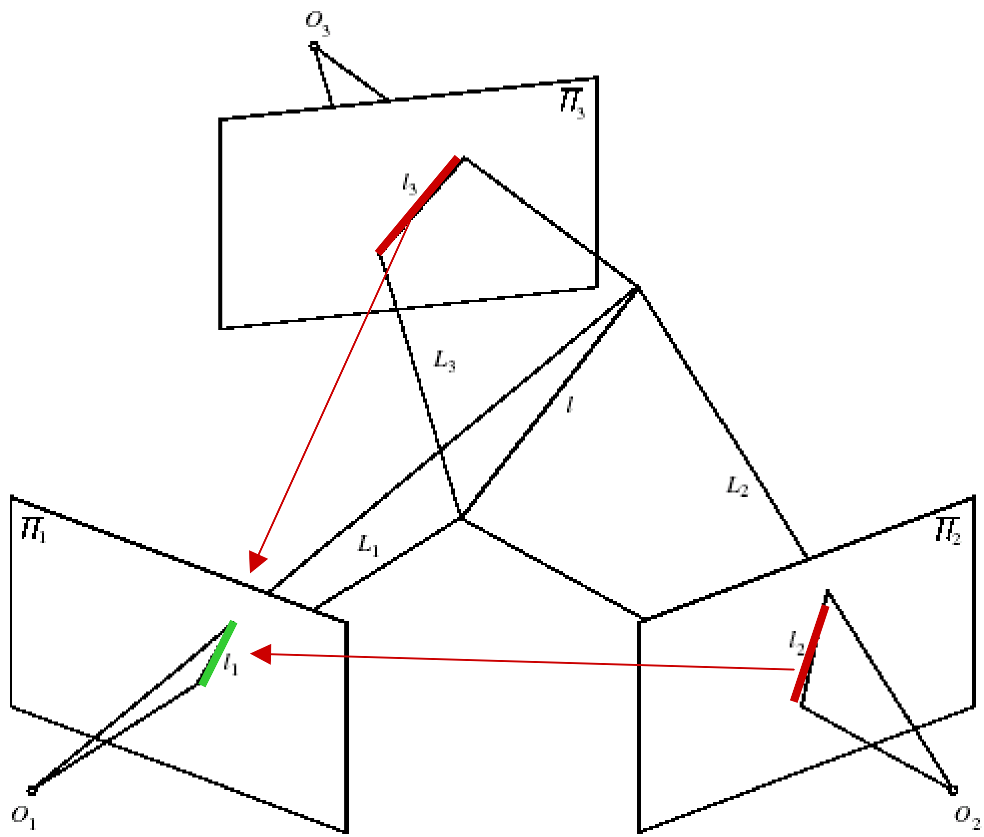
line transfer:

$$\boldsymbol{l}_1 \approx \begin{pmatrix} \boldsymbol{l}_2^T \mathcal{G}_1^1 \boldsymbol{l}_3 \\ \boldsymbol{l}_2^T \mathcal{G}_1^2 \boldsymbol{l}_3 \\ \boldsymbol{l}_2^T \mathcal{G}_1^3 \boldsymbol{l}_3 \end{pmatrix}$$

point transfer via lines: form independent pairs of lines through p2,p3, solve for p1.

Line transfer

$$\boldsymbol{l}_1 \approx \begin{pmatrix} \boldsymbol{l}_2^T \mathcal{G}_1^1 \boldsymbol{l}_3 \\ \boldsymbol{l}_2^T \mathcal{G}_1^2 \boldsymbol{l}_3 \\ \boldsymbol{l}_2^T \mathcal{G}_1^3 \boldsymbol{l}_3 \end{pmatrix}$$



Uncalibrated case

$$\mathcal{L} = \begin{pmatrix} \boldsymbol{l}_1^T \mathcal{K}_1 & 0 \\ \boldsymbol{l}_2^T \mathcal{K}_2 \mathcal{R}_2 & \boldsymbol{l}_2^T \mathcal{K}_2 \boldsymbol{t}_2 \\ \boldsymbol{l}_3^T \mathcal{K}_3 \mathcal{R}_3 & \boldsymbol{l}_3^T \mathcal{K}_3 \boldsymbol{t}_3 \end{pmatrix}$$

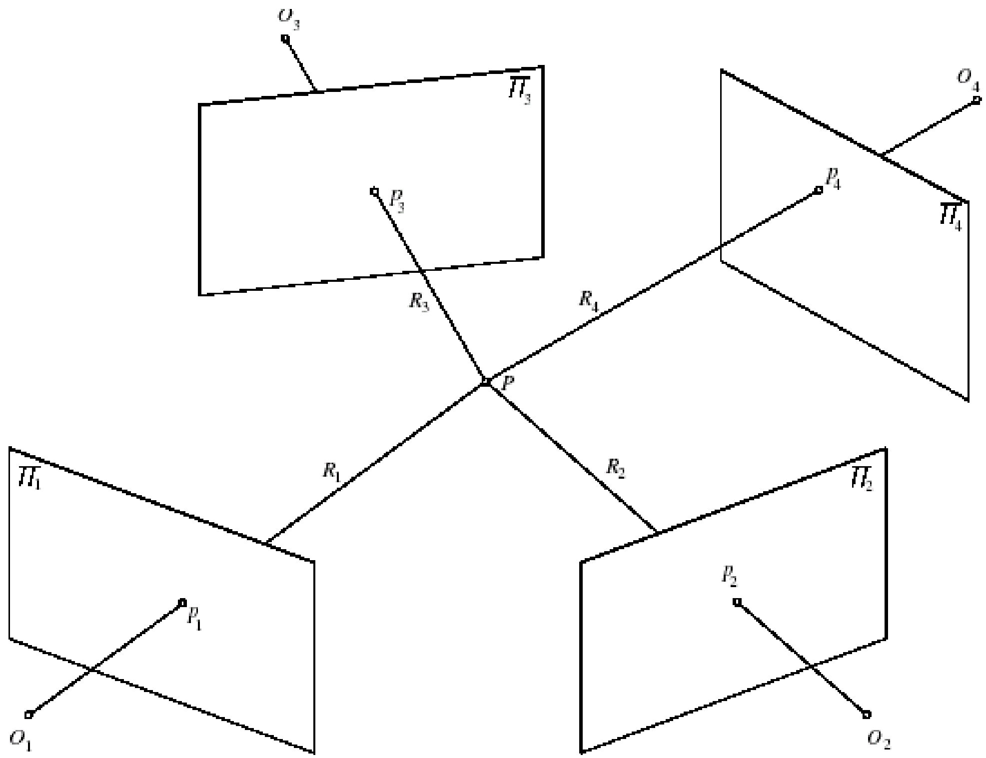
$$\mathcal{A}_i \stackrel{\text{def}}{=} \mathcal{K}_i \mathcal{R}_i \mathcal{K}_1^{-1} \qquad\qquad \boldsymbol{a}_i \stackrel{\text{def}}{=} \mathcal{K}_i \boldsymbol{t}_i$$

$$\mathcal{M}_1 = (\, \mathcal{K}_1 \quad \mathbf{0} \,),\ \mathcal{M}_2 = (\, \mathcal{A}_2 \mathcal{K}_1 \quad \boldsymbol{a}_2 \,),$$

$$\mathcal{M}_3 = (\, \mathcal{A}_3 \mathcal{K}_1 \quad \boldsymbol{a}_3 \,)$$

$$\operatorname{Rank}(\mathcal{L}) = 2 \Longleftrightarrow \operatorname{Rank}(\mathcal{L} \begin{pmatrix} \mathcal{K}_1^{-1} & 0 \\ 0 & 1 \end{pmatrix}) = \operatorname{Rank} \begin{pmatrix} \boldsymbol{l}_1^T & 0 \\ \boldsymbol{l}_2^T \mathcal{A}_2 & \boldsymbol{l}_2^T \boldsymbol{a}_2 \\ \boldsymbol{l}_3^T \mathcal{A}_3 & \boldsymbol{l}_3^T \boldsymbol{a}_3 \end{pmatrix} = 2$$

Quadrifocal geometry



Can form a "quadrifocal tensor"

Faugeras and Mourrain (1995) have shown that it is algebraically dependent on associated essential/fundamental matricies and trifocal tensor: no new constraints added.

No additional independent constraints from more than 3 views.

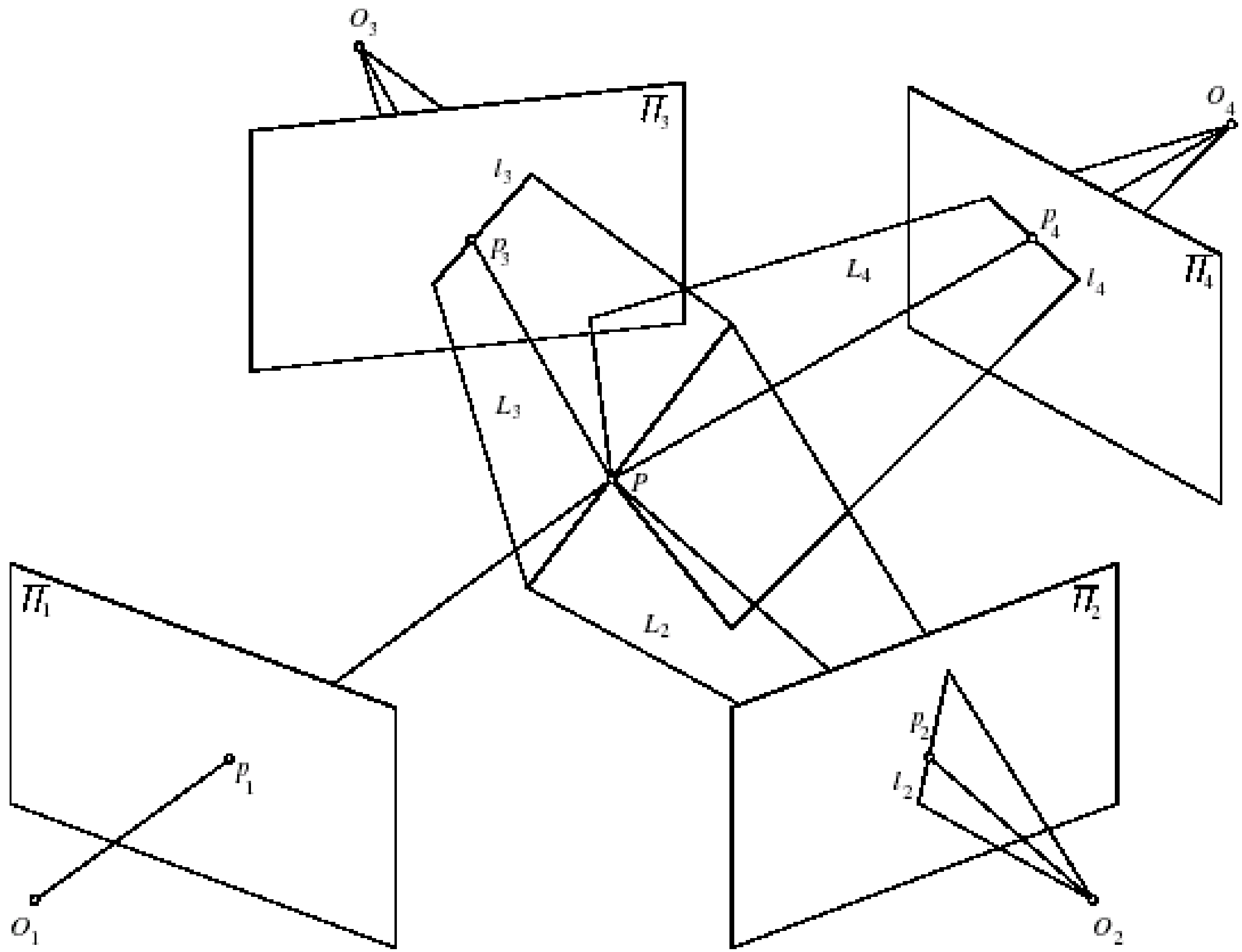


Figure 12.10. Given four images p_1, p_2, p_3 and p_4 of some point P and three arbitrary image lines l_2, l_3 and l_4 passing through the points p_2, p_3 and p_4, the ray passing through O_1 and p_1 must also pass through the point where the three planes L_2, L_3 and L_4 formed by the preimages of these lines intersect.

Trifocal constraint with noise

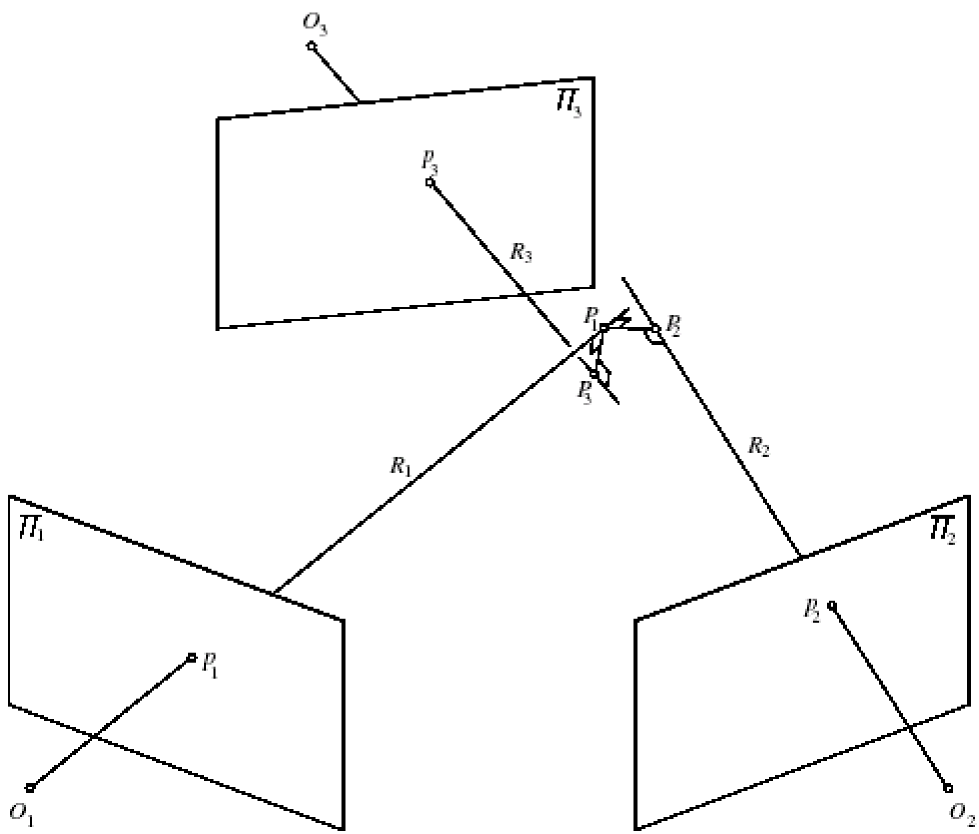


Figure 12.11. Trinocular constraints in the presence of calibration or measurement errors: the rays R_1, R_2 and R_3 may not intersect.

Multi-view geometry and 3-D

We have 2 eyes, yet we see 3-D!

Using multiple views allows inference of hidden dimension.

Geometric and algebraic constraints

…can you see 3-D without multiple views?

Outline

- Multi-view geometry
- Epipolar constraint
- Essential matrix
- Fundamental matrix
- Trifocal tensor

Next: how do we find those correspondences?

[Most figures adapted from Forsythe and Ponce]