Geometry of Multiple views

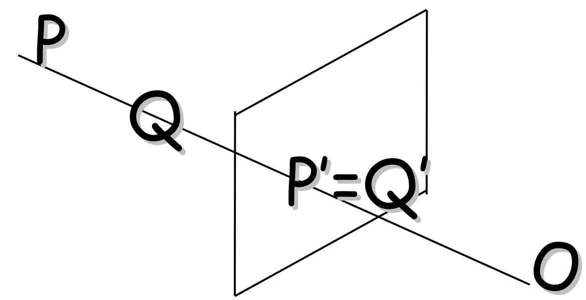
CS 554 – Computer Vision
Pinar Duygulu
Bilkent University

Multiple views

Despite the wealth of information contained in a a photograph, the **depth** of a scene point along the corresponding projection ray **is not directly accessible in a single image**

3D Points on the same viewing line have the same 2D image:

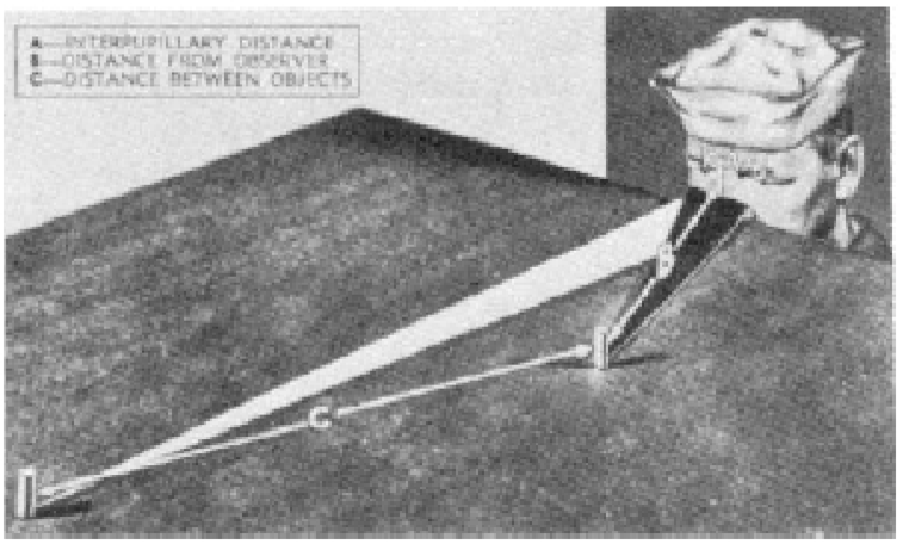
2D imaging results in depth information loss



With at least two pictures, depth can be measured by triangulation.

Human/Animal Visual system

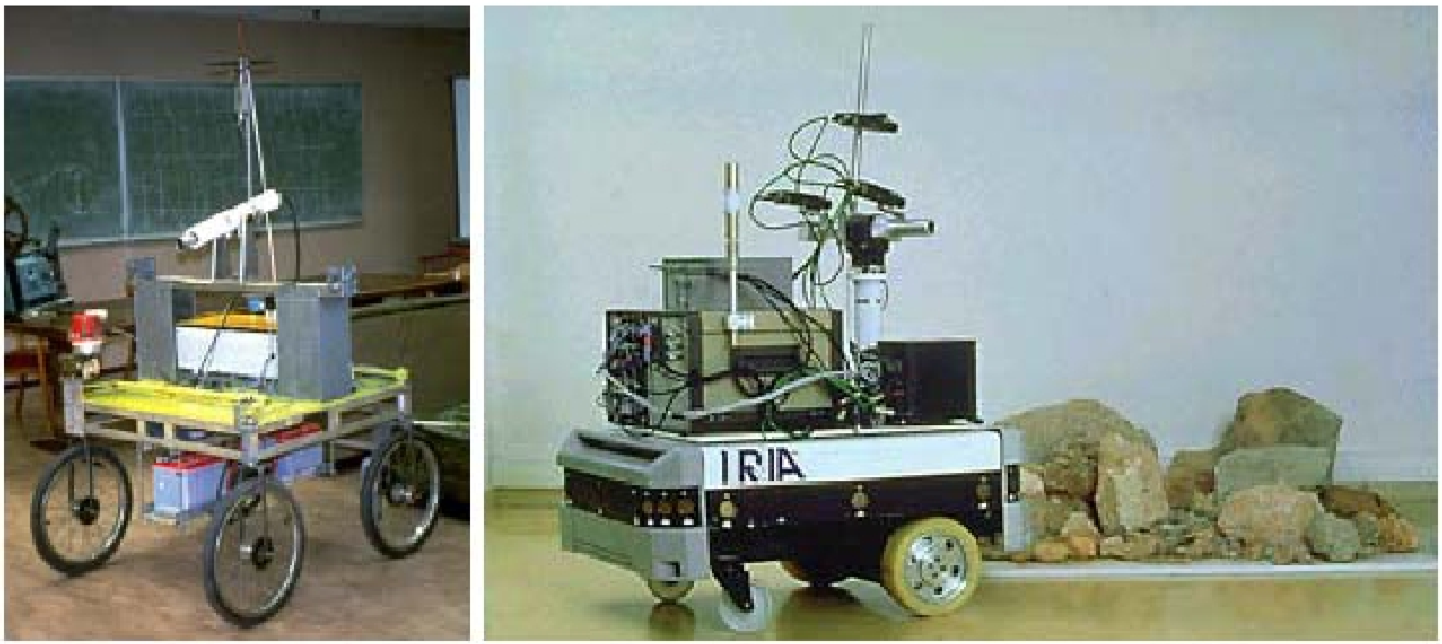
It is the reason that most animals have at least two eyes and/or move their head when looking around



Adapted from David Forsyth, UC Berkeley

Visual Robot Navigation

This is also the motivation for equipping an autonomous robot with a stereo or motion analysis system.

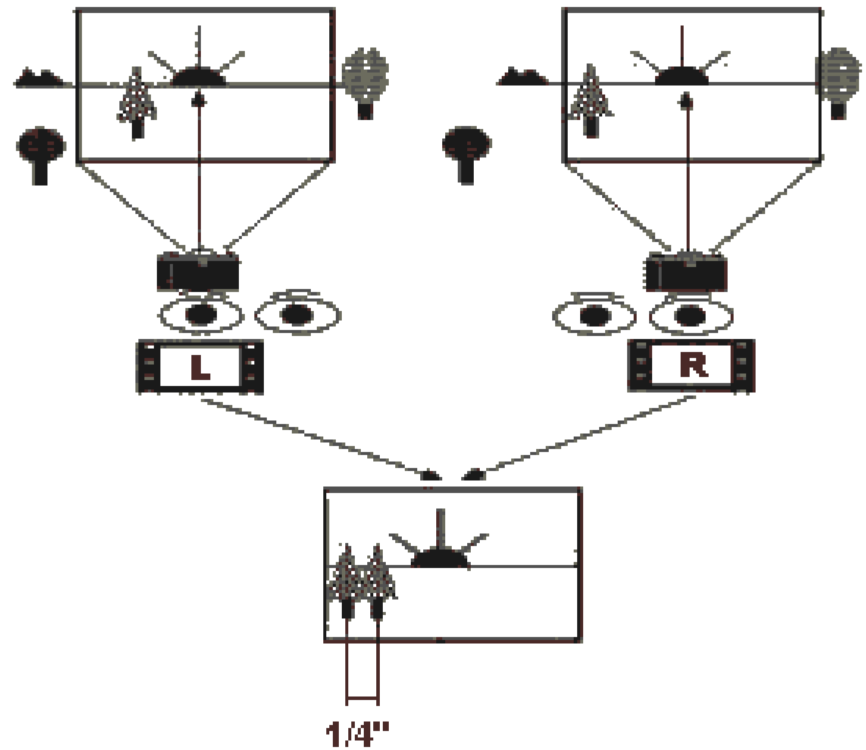


Left : The Stanford cart sports a single camera moving in discrete increments along a straight line and providing multiple snapshots of outdoor scenes
Right : The INRIA mobile robot uses three cameras to map its environment

Forsyth & Ponce

Human Vision

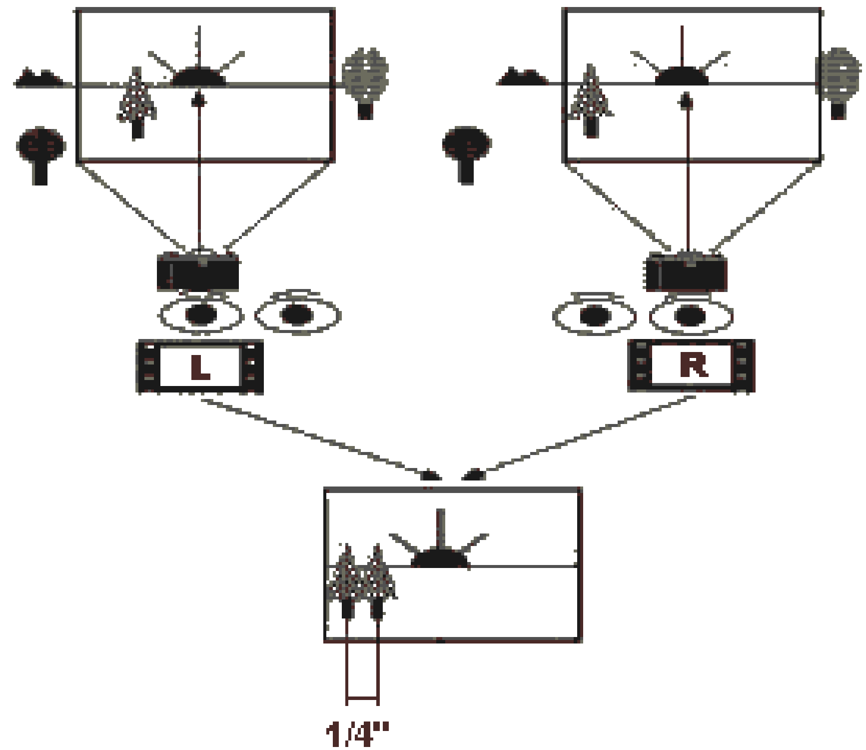
- Humans have two eyes, both forward facing but horizontally spaced by approximately 60mm.
- When looking at an object, each eye will produce a slightly different image, as it will be looking at a slightly different angle.
- The human brain combines both these images into one to give a perception of depth.
- This processing is so quick and seamless that the perception is that we are looking through one big eye rather than two.



http://www.photostuff.co.uk/stereo.htm

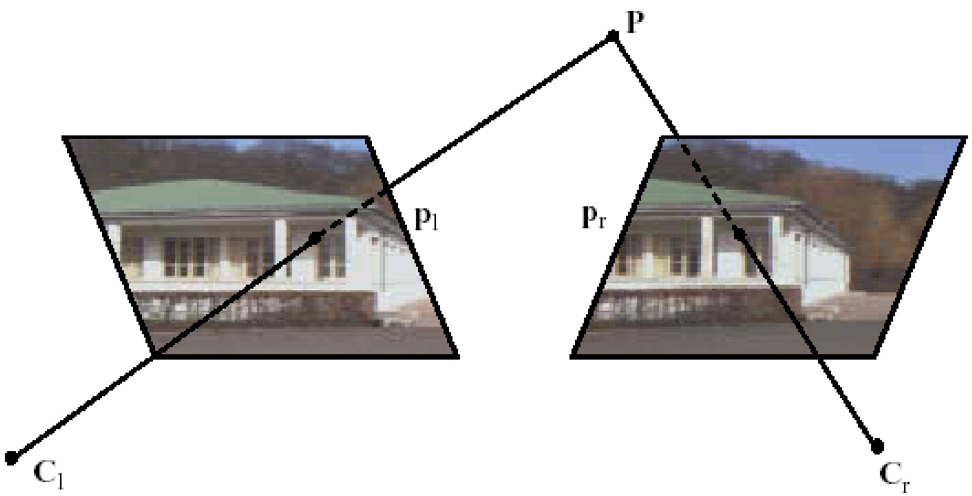
Human Vision

- The brain can also determine depth and how far objects are away from each other by the amount of difference between the two images that it receives.
- The further the subject is from the eye, the less will be the difference between the two images and conversely the nearer the subject, the greater the difference.
- The left and right eyes see the sun in the same place as it is in the distance. The tree being much closer is seen in slightly different places



http://www.photostuff.co.uk/stereo.htm

Stereo vision = correspondences + reconstruction



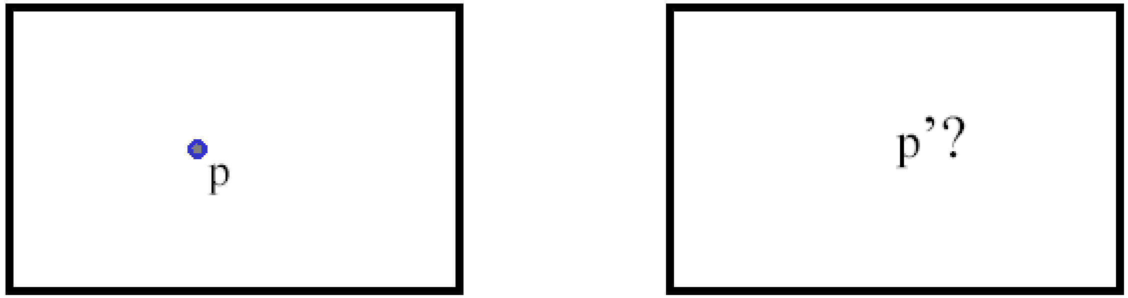
Stereovision involves two problems:

Correspondence : Given a point p_l in one image, find the corresponding point in the other image

Reconstruction: Given a correspondence (p_l, p_r) compute the 3D coordinates of the corresponding point in space, P

Adapted from Martial Hebert, CMU

Stereo constraints



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

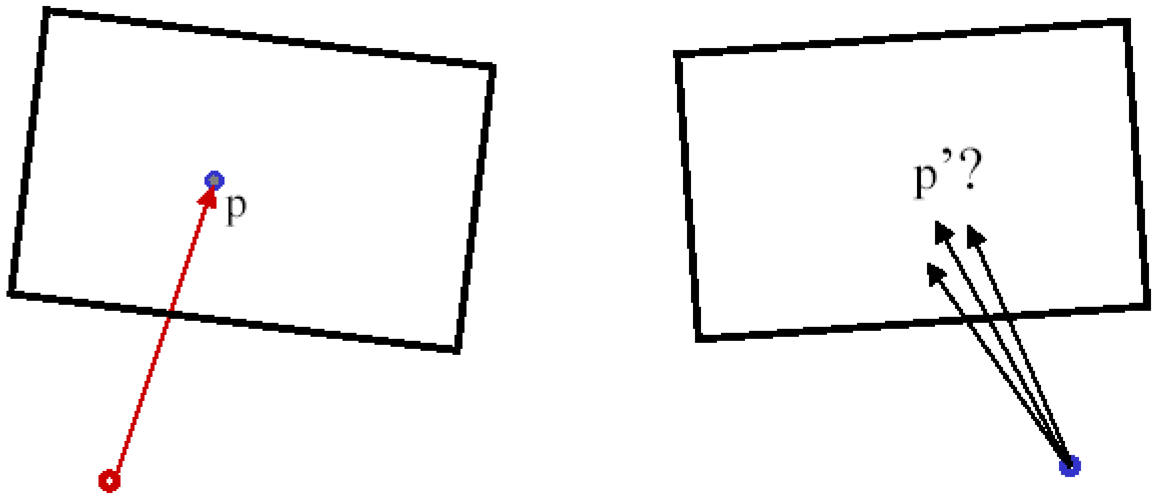
Could be anywhere! Might not be same scene!

… Assume pair of pinhole views of static scene:

Adapted from Trevor Darrell, MIT

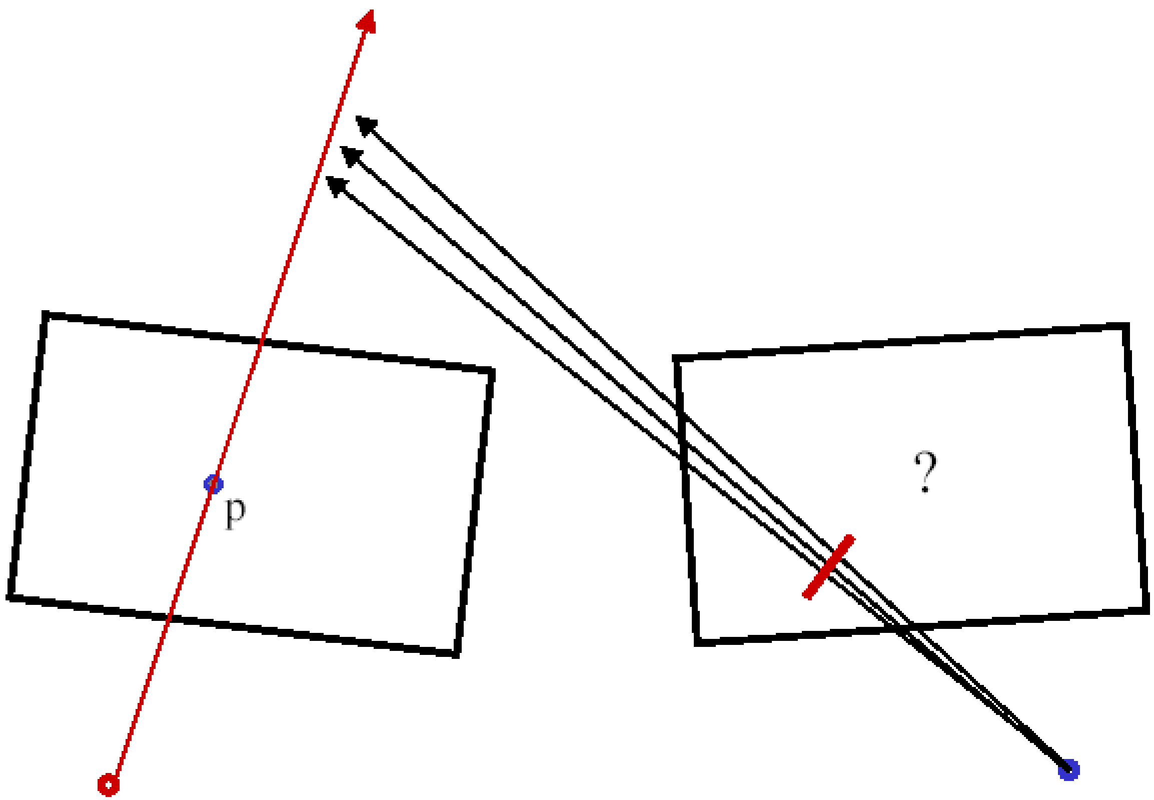
Stereo constraints

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



Adapted from Trevor Darrell, MIT

Epipolar line

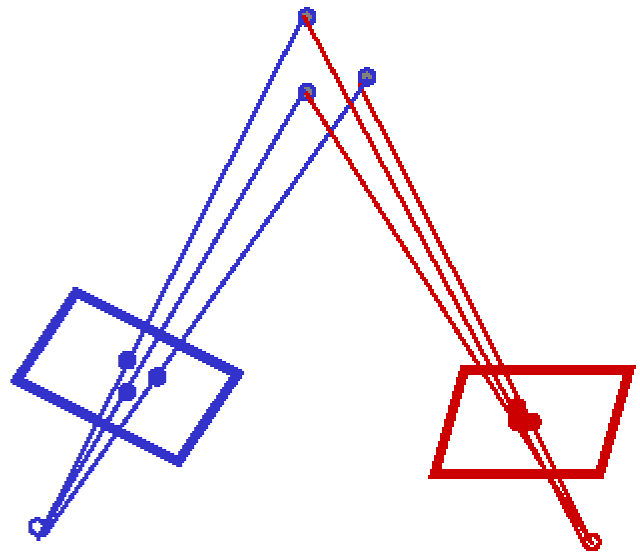


Adapted from Trevor Darrell, MIT

Multi-view Geometry

Relate

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



Adapted from Trevor Darrell, MIT

Epipolar constraint

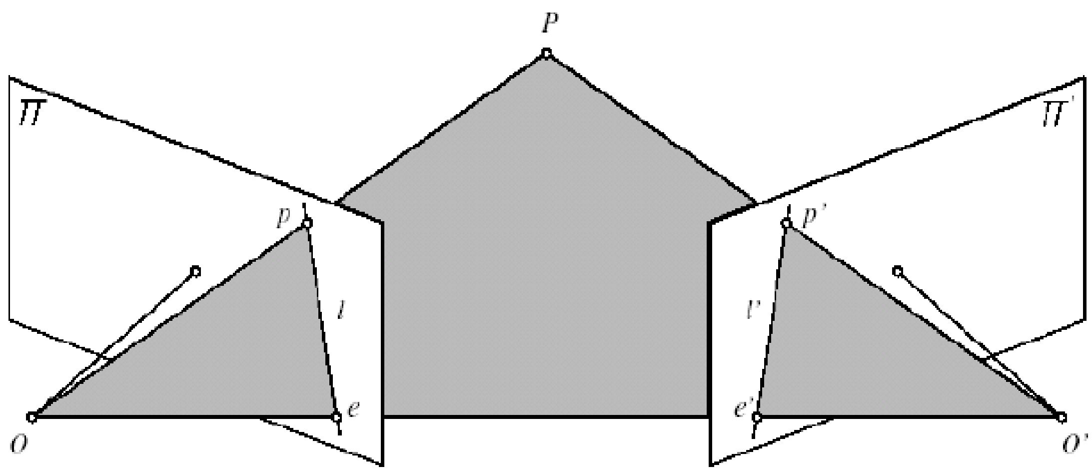


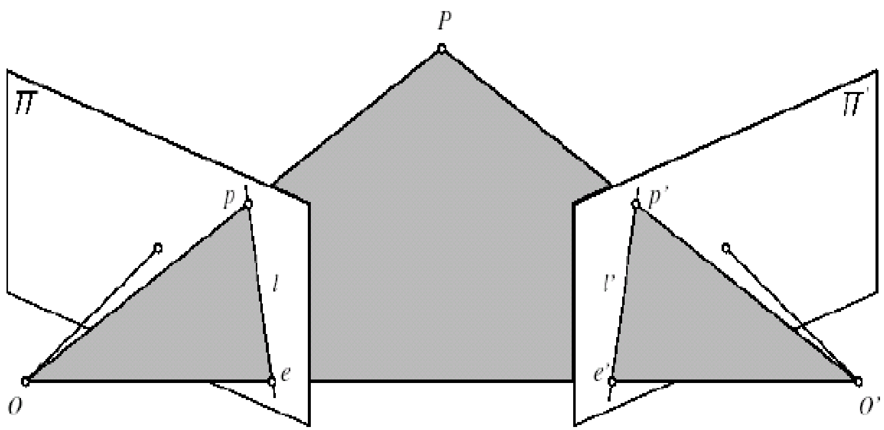
FIGURE 11.1: Epipolar geometry: the point P, the optical centers O and O' of the two cameras, and the two images p and p' of P all lie in the same plane.

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

O, O' : optical centers
p & p' are the images of P → These 5 points all belong to epipolar plane

Forsyth & Ponce

Epipolar constraint

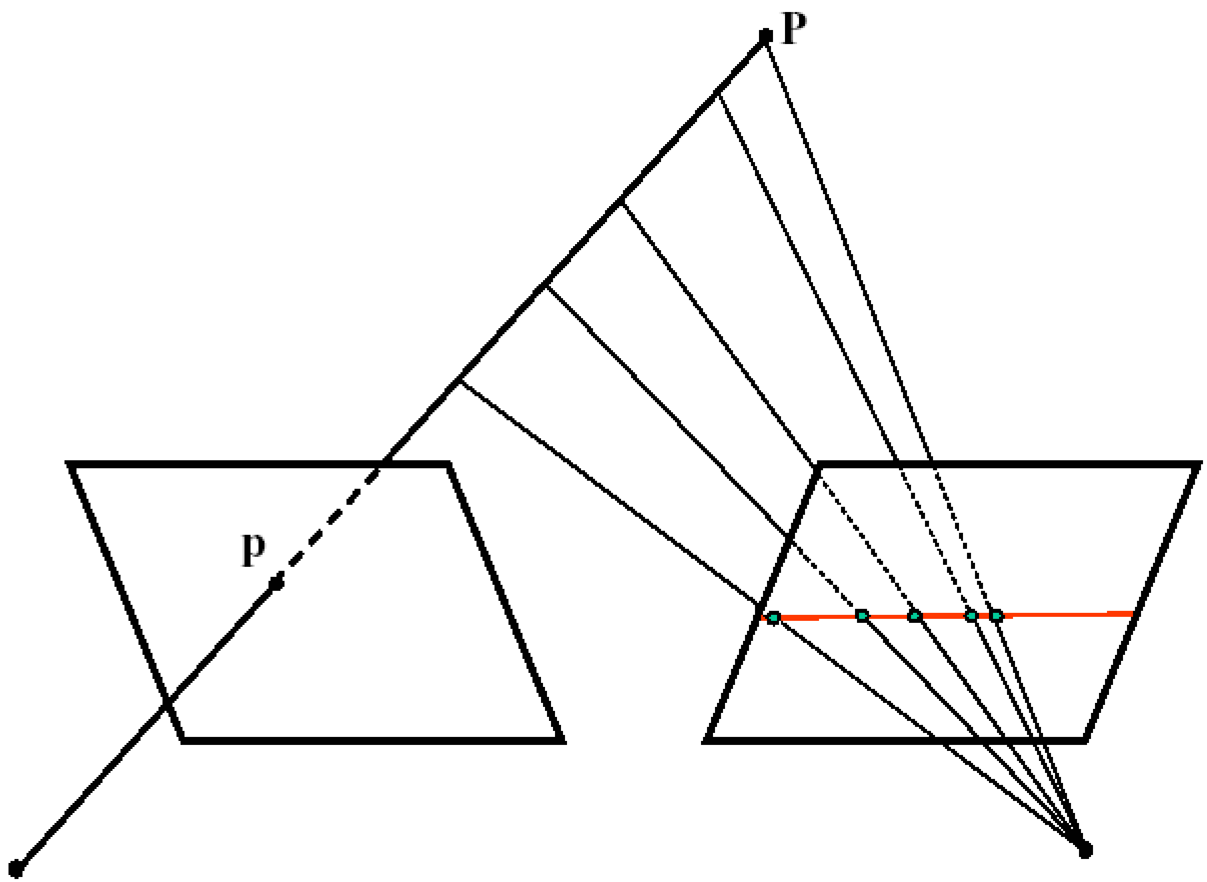


Point p' lies on the line l' where epipolar plane and the retina π' intersect.
The line l' is the epipolar line associated with the point p
It passes through the point e' where the baseline joining the optical centers o and O' intersects
The points e and e' are called the epipoles of the cameras

If p and p' are the images of the same point P, then p' must lie on the epipolar assoiciated with p → Epipolar constraint

Adapted from Trevor Darrell, MIT

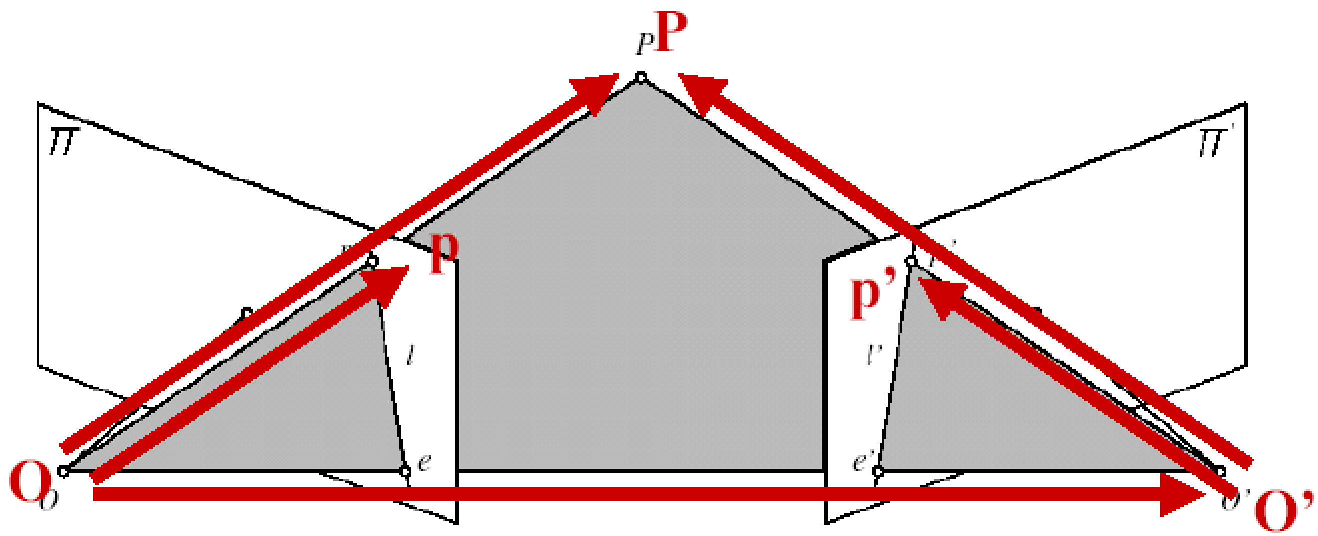
Epipolar constraint



Epipolar constraint greatly limits the search of corresponding points.

Adapted from Martial Hebert, CMU

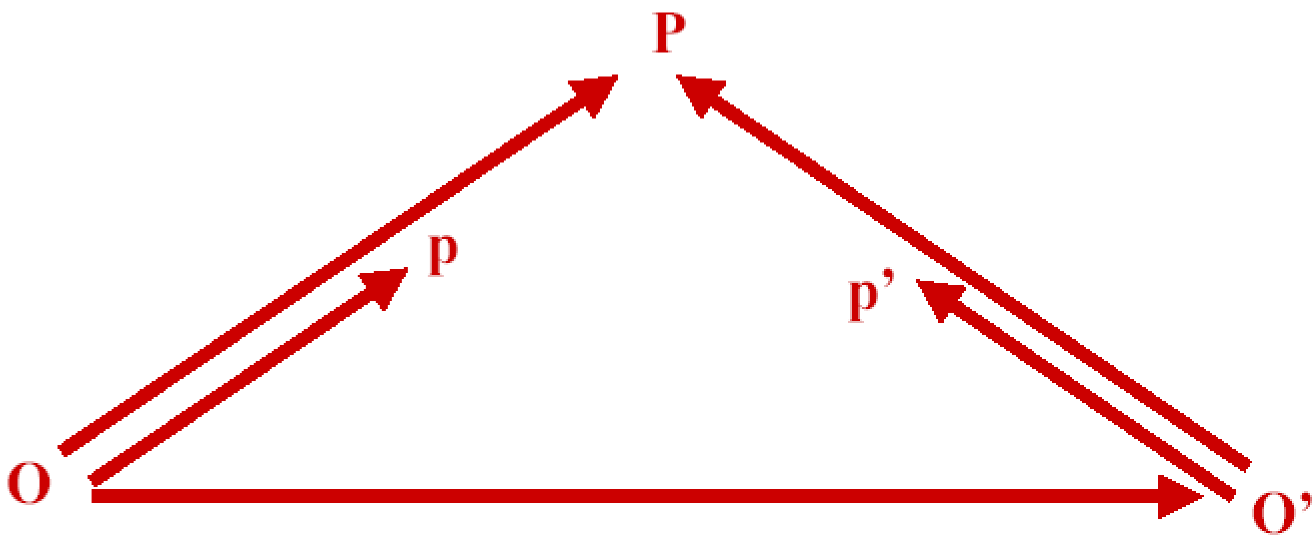
Epipolar constraint – Calibrated case



Assume that the intrinsic parameters of each camera are known

Adapted from Trevor Darrell, MIT

Epipolar constraint – Calibrated case

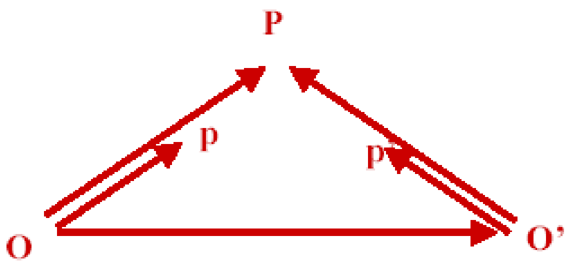


The epipolar constraint: these vectors are coplanar:

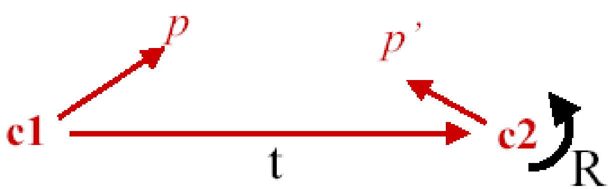
$$\overrightarrow{Op} \cdot [\overrightarrow{OO'} \times \overrightarrow{O'p'}] = 0$$

Adapted from Trevor Darrell, MIT

Epipolar constraint – Calibrated case



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



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

c2 is related to c1 by rotation R and translation t

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

$p = (u,v,1)^T$ $p' = (u',v',1)^T$

Adapted from Trevor Darrell, MIT

Epipolar constraint – Calibrated case

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.

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

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

Adapted from Trevor Darrell, MIT

Epipolar constraint – Calibrated case

Review: matrix form of cross-product

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

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

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

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

Adapted from Trevor Darrell, MIT

Epipolar constraint – Calibrated case

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

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

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

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

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

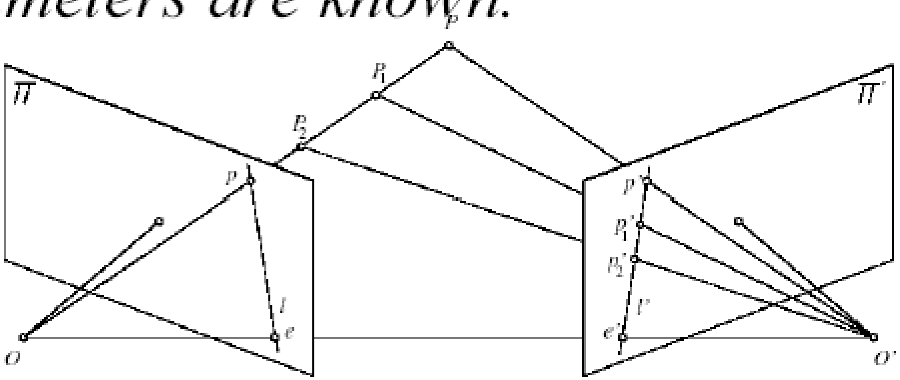
Adapted from Trevor Darrell, MIT

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.



$$\mathcal{E} = [t_x]\Re$$

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

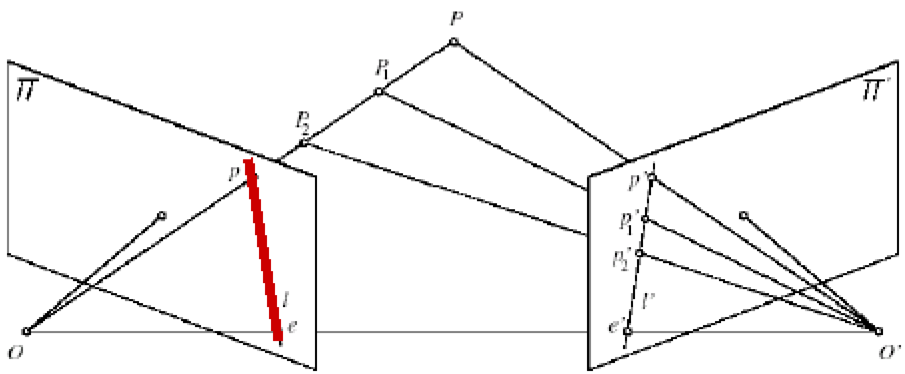
$$\vec{a} \times \vec{b} = [a_x]\vec{b}$$

Adapted from Trevor Darrell, MIT

The essential matrix

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

$$au + bv + c = 0$$



$\mathcal{E}p'$

$$p = (u, v, 1)^T$$
$$l = (a, b, c)^T$$
$$l \cdot p = 0$$

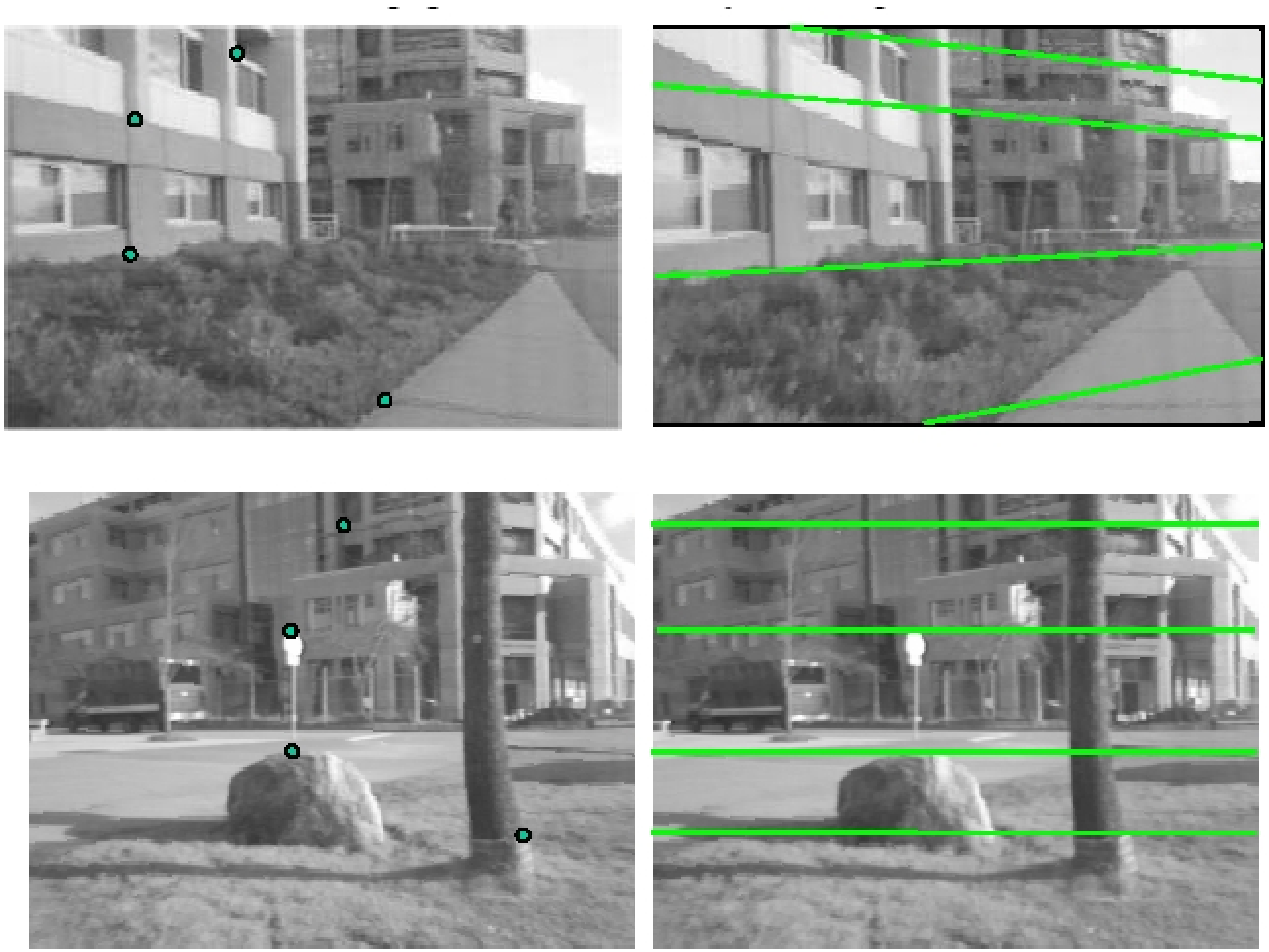
$$\mathcal{E}p' \cdot p = 0$$

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

p lies on the epipolar line associated with the point p'

Adapted from Trevor Darrell, MIT

Epipolar geometry example

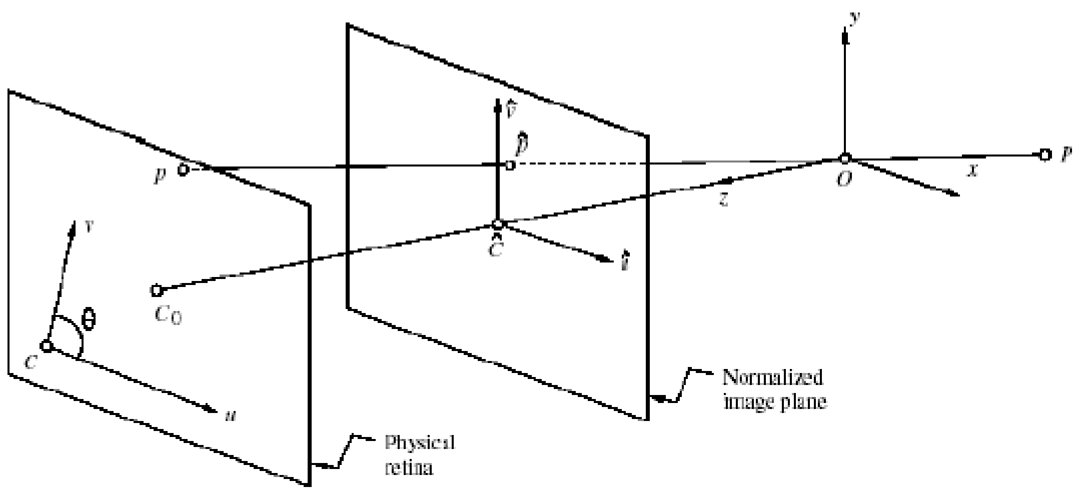


Adapted from Martial Hebert, CMU

What if calibration is unknown?

Recall calibration eqn:

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



Adapted from Trevor Darrell, MIT

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:

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

Adapted from Trevor Darrell, MIT

Estimating the Fundamental Matrix

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

Each point correspondence can be expressed as a single linear equation

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

Adapted from Trevor Darrell, MIT

Estimating the Fundamental Matrix

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

Each point correspondence can be expressed as a single linear equation

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

Adapted from Trevor Darrell, MIT

The 8 point algorithm (Longuet-Higgins, 1981)

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 $\sum_{i=1}^{n} (\boldsymbol{p}_i^T \mathcal{F} \boldsymbol{p}'_i)^2$)

Adapted from Trevor Darrell, MIT

Improved 8 point algorithm (Normalized)

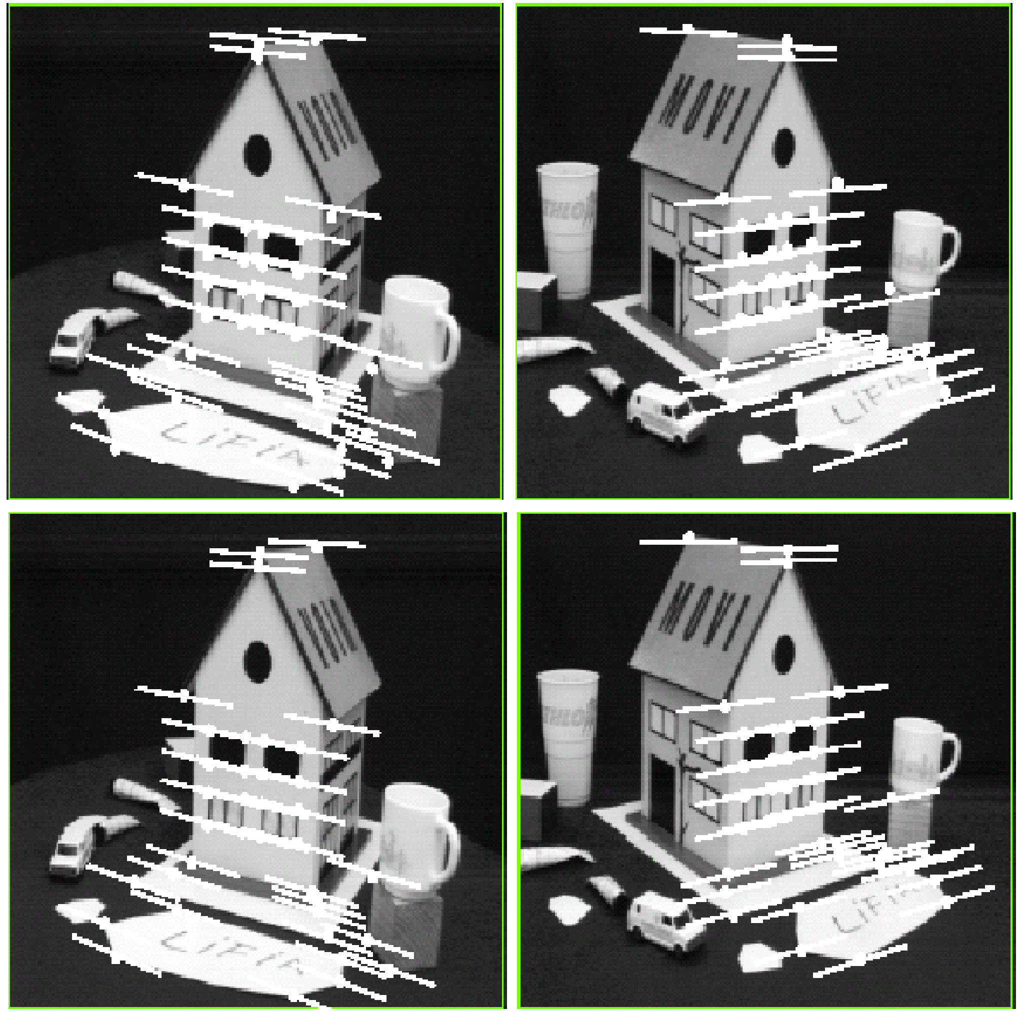
Hartley 1995: use SVD.

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

Adapted from Trevor Darrell, MIT

8 point algorithm

(a)

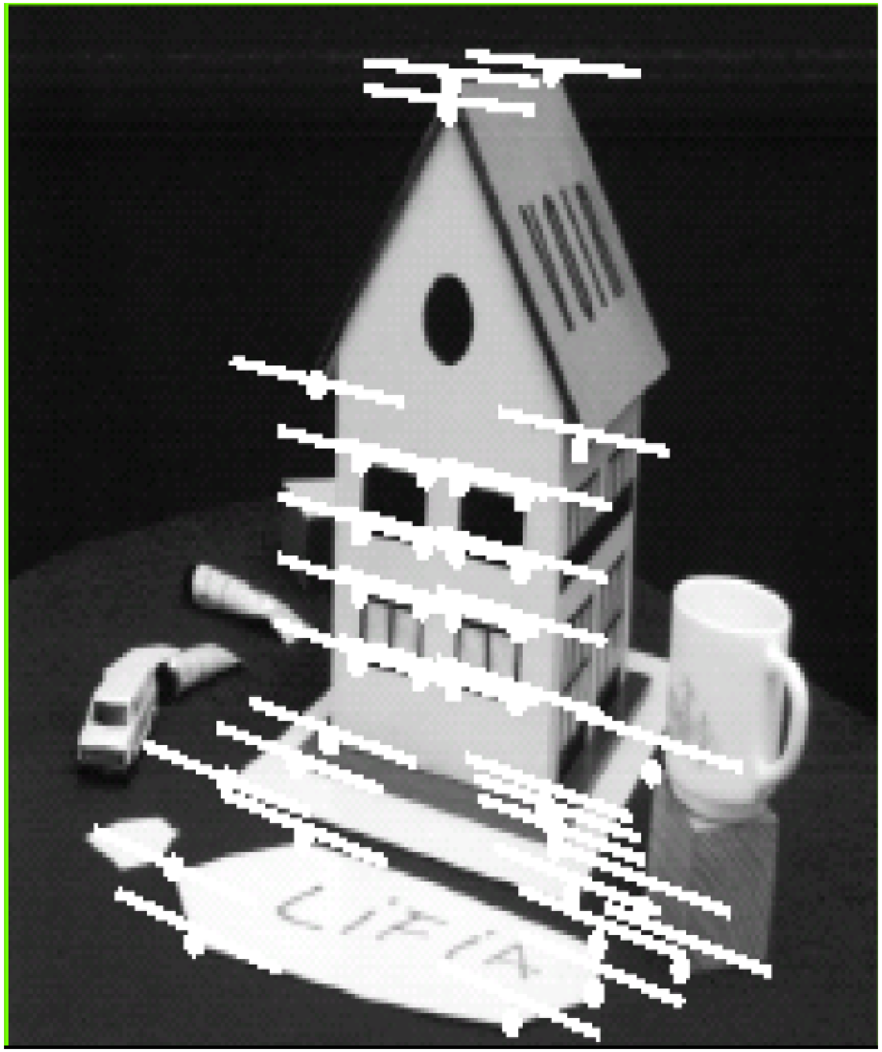


(b)

	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

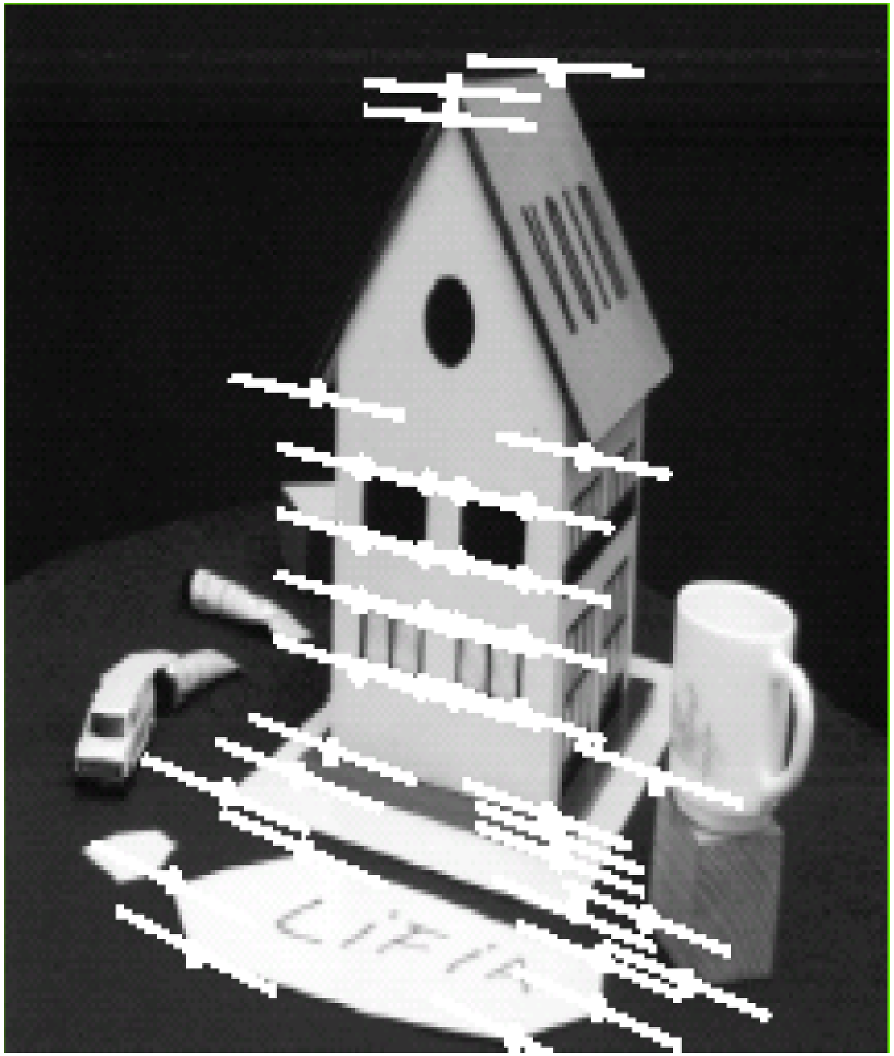
Adapted from Trevor Darrell, MIT

8 point algorithm



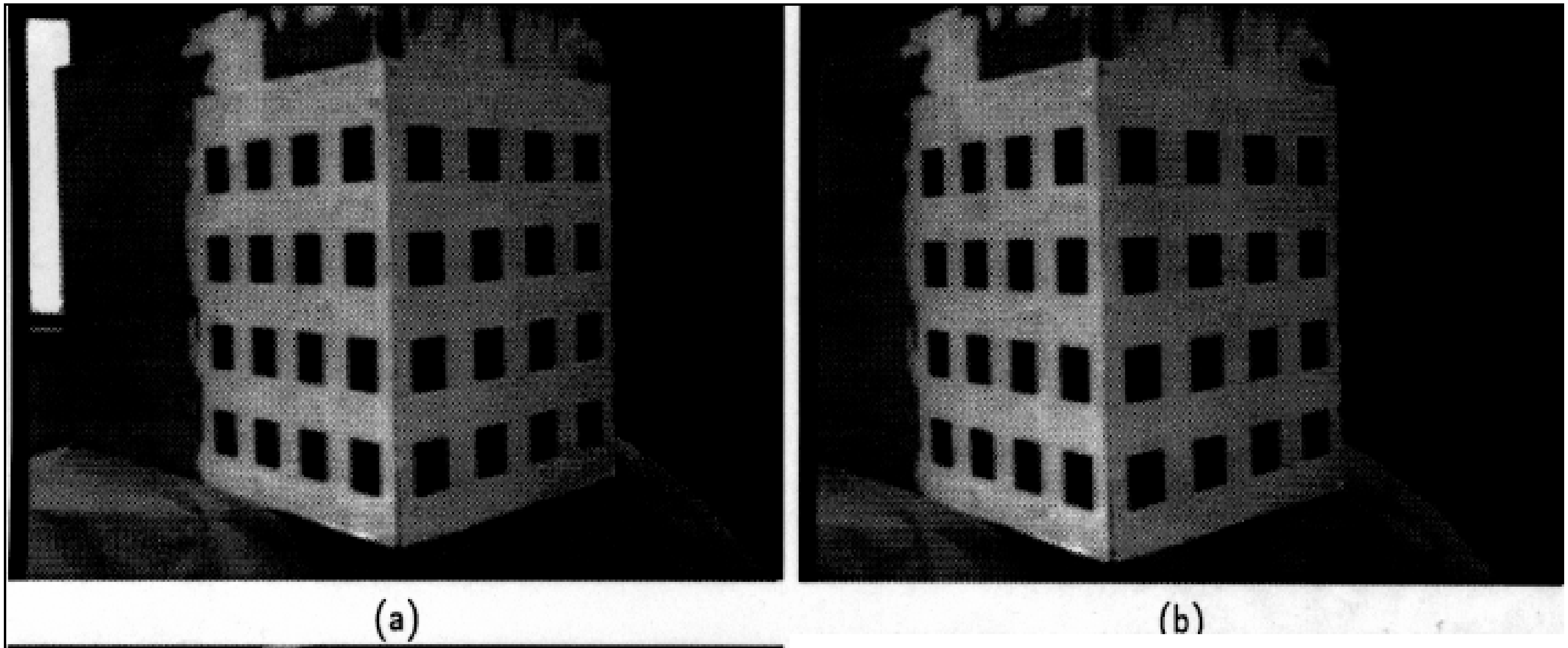
Adapted from Trevor Darrell, MIT

8 point algorithm



Adapted from Trevor Darrell, MIT

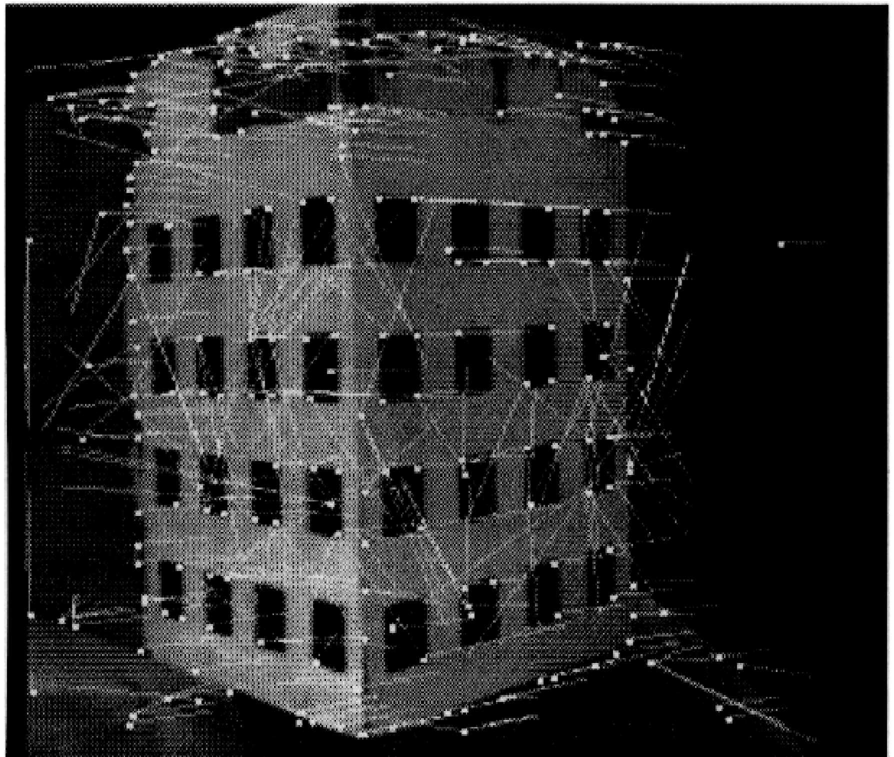
Example

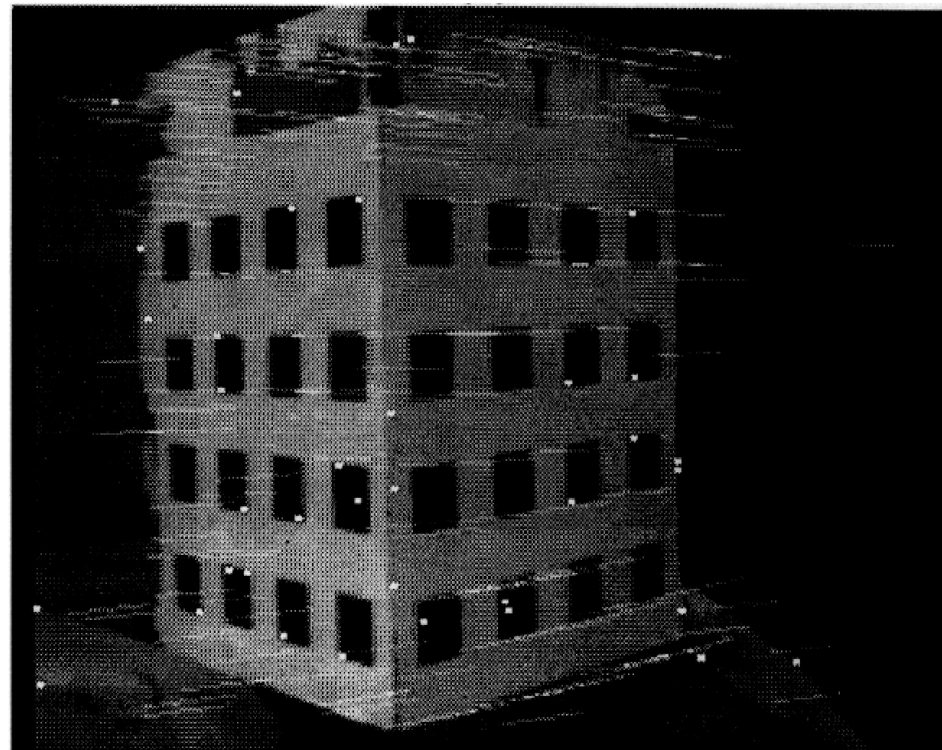


From Torr and Murray, "The development and comparison of robust methods for estimating the fundamental matrix"

Adapted from David Forsyth

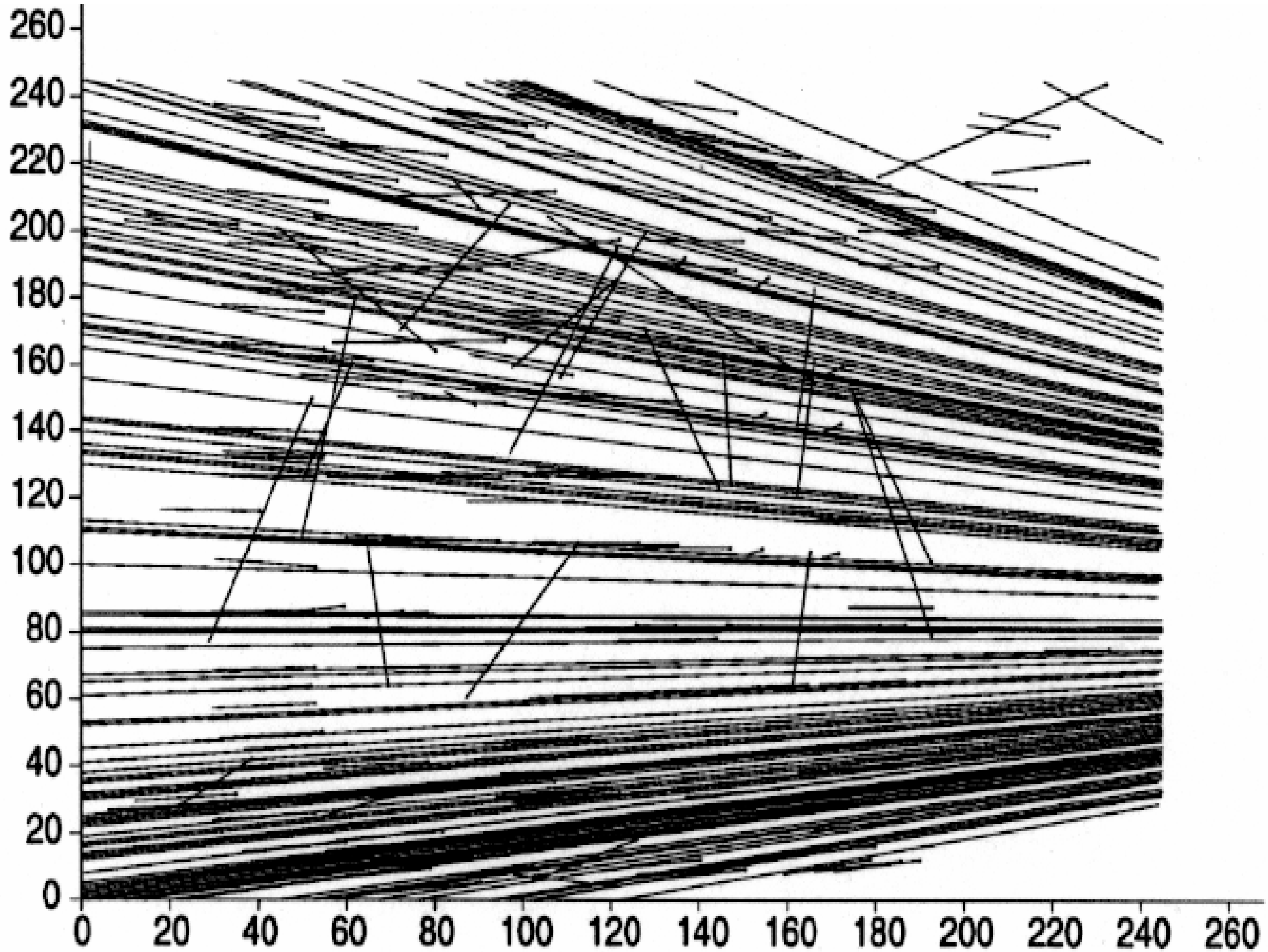
Example





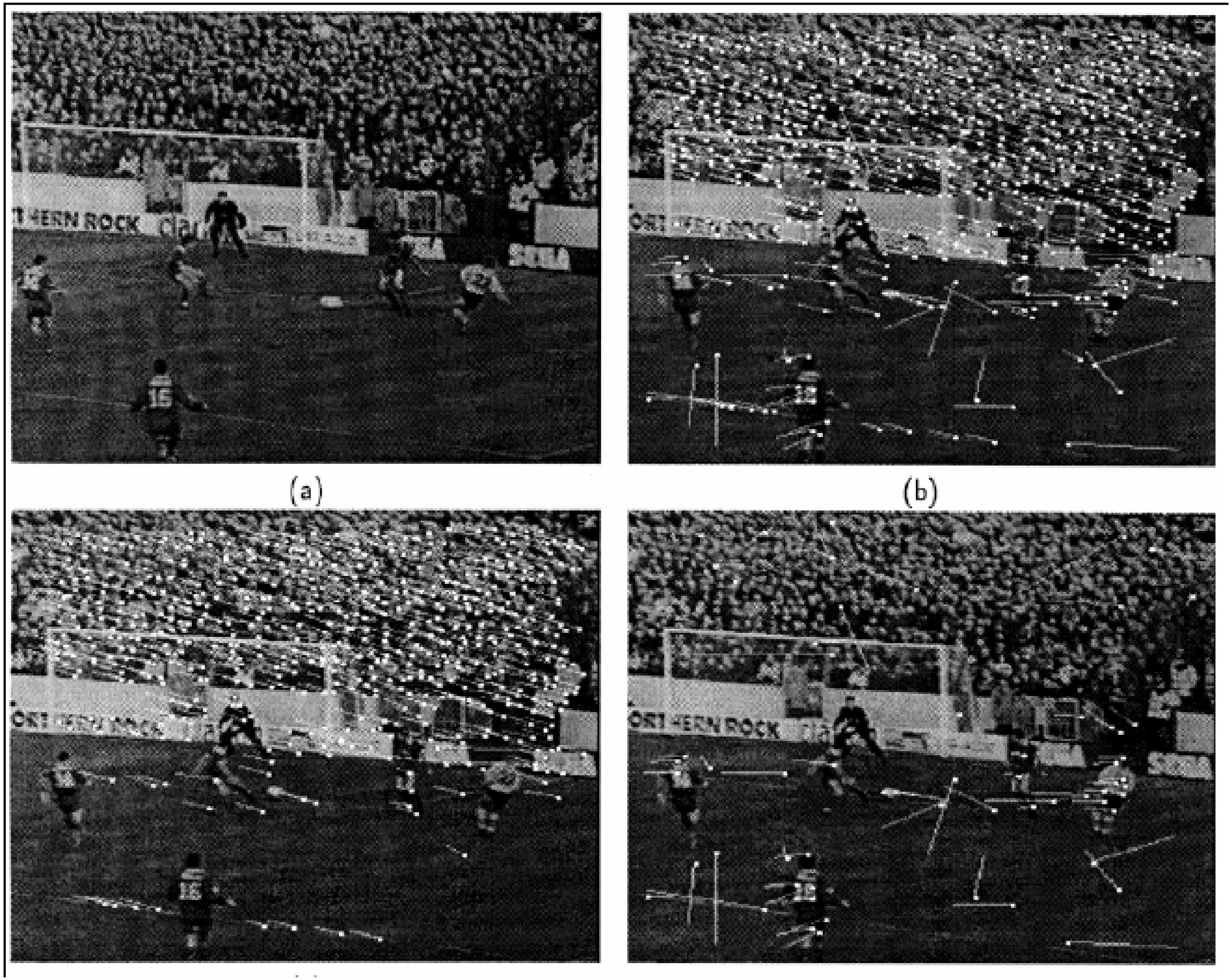
Adapted from David Forsyth

Example



Adapted from David Forsyth

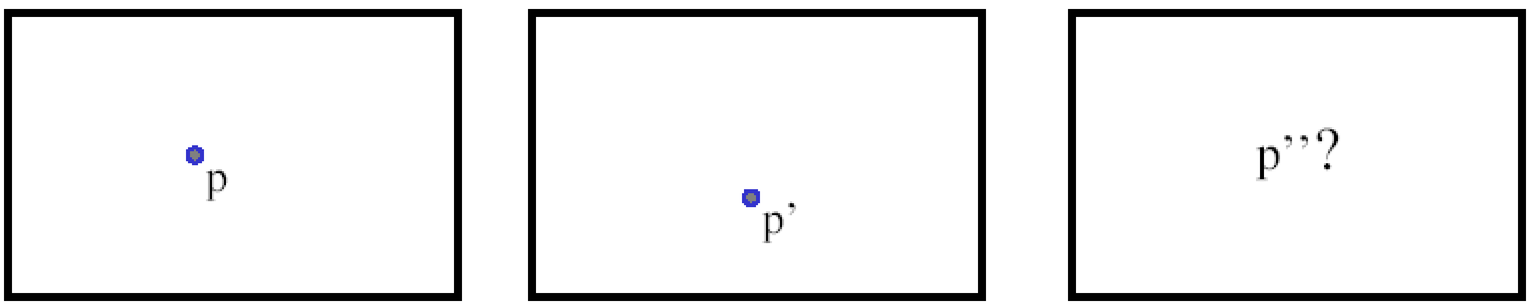
Example



Adapted from David Forsyth

Multiple Cameras

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

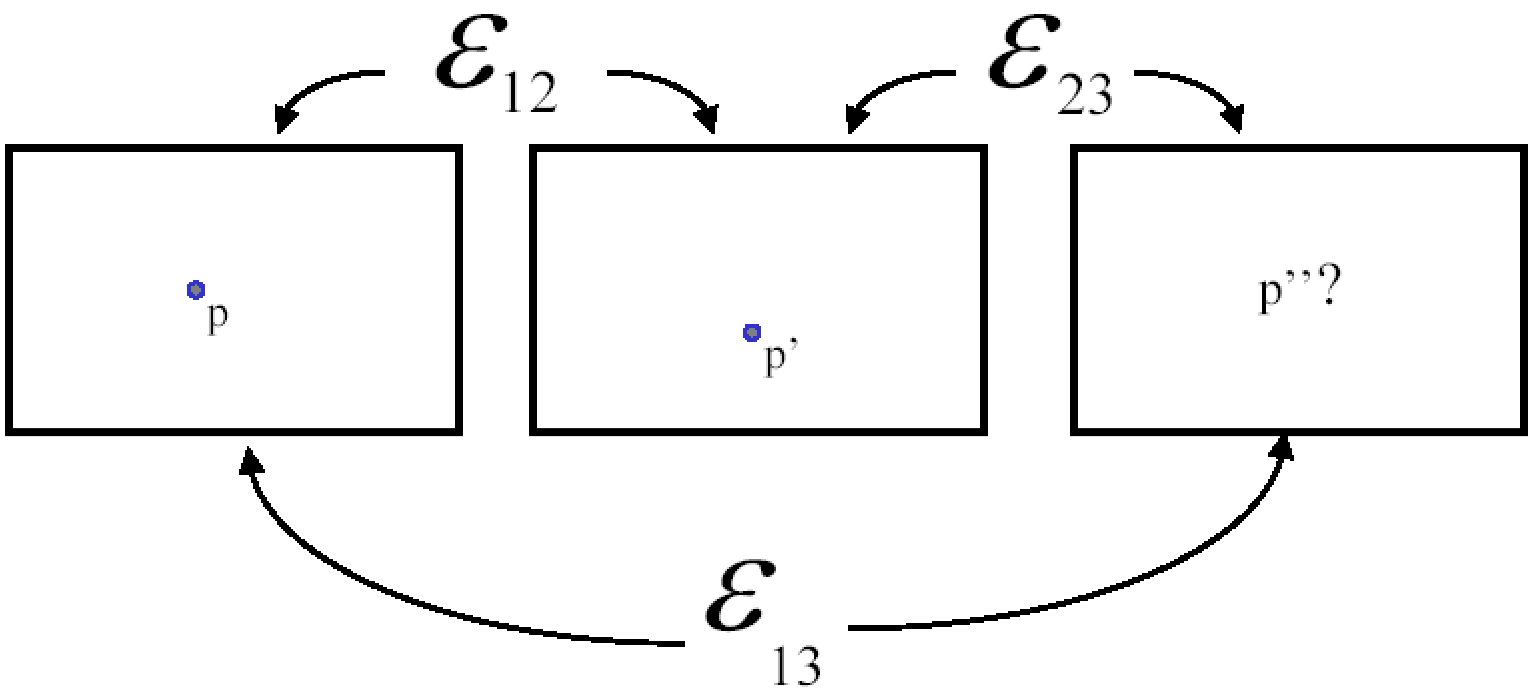


Adapted from Trevor Darrell, MIT

Three essential matrices

Essential matrices relate each pair:

(calibrated case)



Adapted from Trevor Darrell, MIT

Three essential matrices

$$\left\{\begin{array}{ll} \boldsymbol{p}_1^T \mathcal{E}_{12} \boldsymbol{p}_2 & = 0, \\ \boldsymbol{p}_2^T \mathcal{E}_{23} \boldsymbol{p}_3 & = 0, \\ \boldsymbol{p}_3^T \mathcal{E}_{31} \boldsymbol{p}_1 & = 0, \end{array}\right.$$

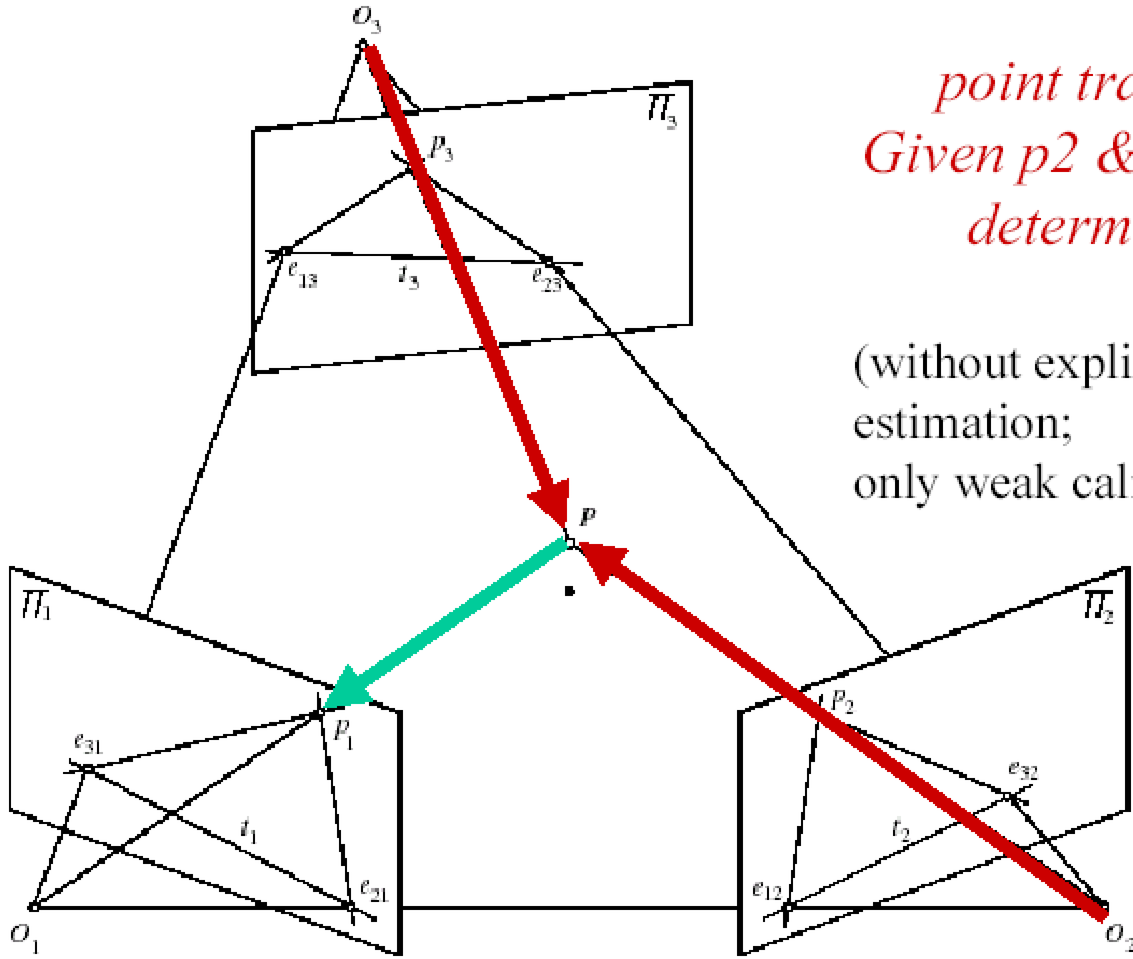
any two are independent!

can predict third point from two others.

Adapted from Trevor Darrell, MIT

Trinocular epipolar geometry

Trifocal plane formed from trifocal lines



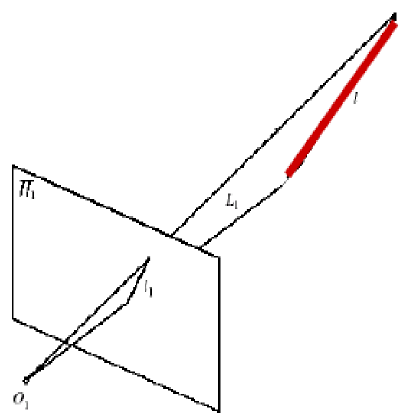
point transfer: Given p2 & p3, p1 is determined!

(without explicit depth estimation; only weak calibration)

Adapted from Trevor Darrell, MIT

Trifocal line constraint

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



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

Adapted from Trevor Darrell, MIT

Trifocal line constraint

3 cameras, 3 plane equations:

$$L = M^T l$$

$$\begin{pmatrix} L_1^T \\ L_2^T \\ L_3^T \end{pmatrix} P = 0$$

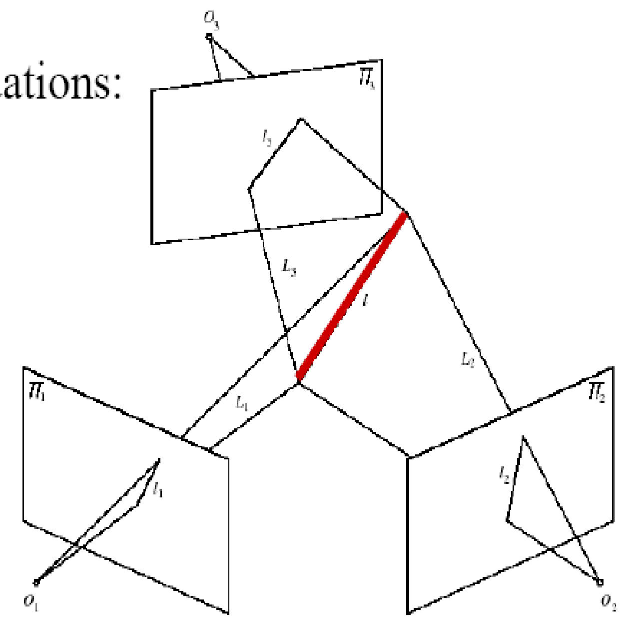


Figure 12.6. Three images of a line define it as the intersection of three planes in t same pencil.

Adapted from Trevor Darrell, MIT

Trifocal line constraint

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

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

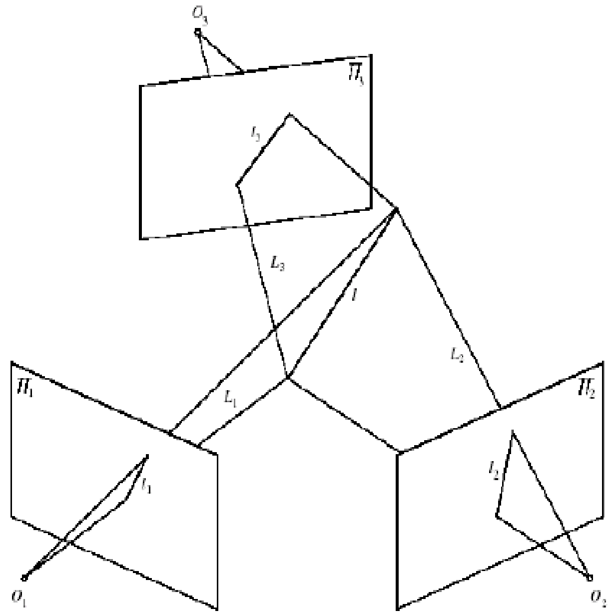
$$\begin{pmatrix} \boldsymbol{L}_1^T \\ \boldsymbol{L}_2^T \\ \boldsymbol{L}_3^T \end{pmatrix} \boldsymbol{P} = \boldsymbol{0}$$

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

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

Adapted from Trevor Darrell, MIT

Trifocal line constraint



Rank of $\mathcal{L} \stackrel{\text{def}}{=} \begin{pmatrix} l_1^T \mathcal{M}_1 \\ l_2^T \mathcal{M}_2 \\ l_3^T \mathcal{M}_3 \end{pmatrix} = 2$

Adapted from Trevor Darrell, MIT

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

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

Adapted from Trevor Darrell, MIT

Trifocal line constraint

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

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

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

with

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

Adapted from Trevor Darrell, MIT

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

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

can be used for point or line transfer.

Adapted from Trevor Darrell, MIT

Trifocal line constraint

line transfer:

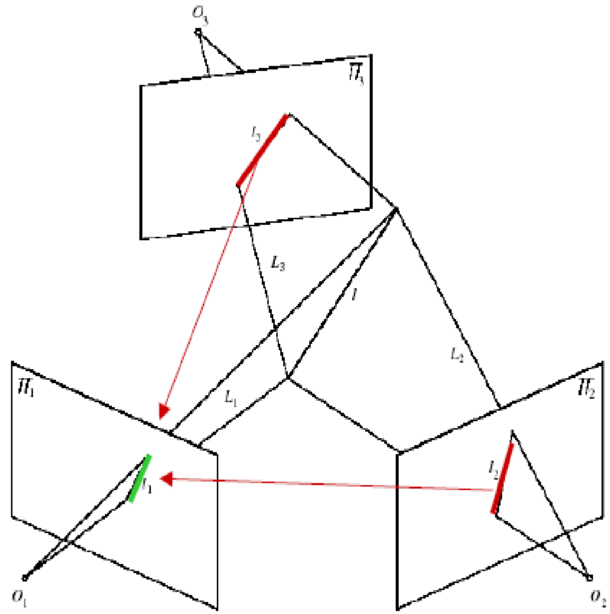
$$\boldsymbol{l}_1 \approx \begin{pmatrix} \boldsymbol{l}_2^T \mathcal{G}_1^1 \boldsymbol{l}_3 \\ \boldsymbol{l}_2^T \mathcal{G}_1^2 \boldsymbol{l}_3 \\ \boldsymbol{l}_2^T \mathcal{G}_1^3 \boldsymbol{l}_3 \end{pmatrix}$$

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

Adapted from Trevor Darrell, MIT

Line transfer

$$l_1 \approx \begin{pmatrix} l_2^T \mathcal{G}_1^1 l_3 \\ l_2^T \mathcal{G}_1^2 l_3 \\ l_2^T \mathcal{G}_1^3 l_3 \end{pmatrix}$$



Adapted from Trevor Darrell, MIT

Uncalibrated case

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

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

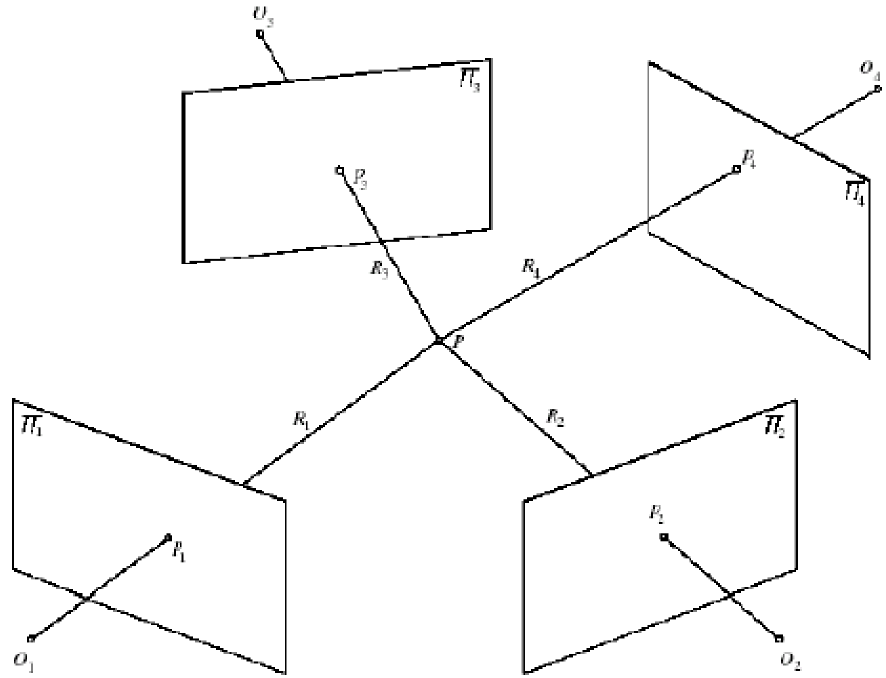
$$\mathcal{M}_1 = \begin{pmatrix} \mathcal{K}_1 & \boldsymbol{0} \end{pmatrix}, \ \mathcal{M}_2 = \begin{pmatrix} \mathcal{A}_2 \mathcal{K}_1 & \boldsymbol{a}_2 \end{pmatrix},$$

$$\mathcal{M}_3 = \begin{pmatrix} \mathcal{A}_3 \mathcal{K}_1 & \boldsymbol{a}_3 \end{pmatrix}$$

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

Adapted from Trevor Darrell, MIT

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.

Adapted from Trevor Darrell, MIT

Quadrifocal Geometry

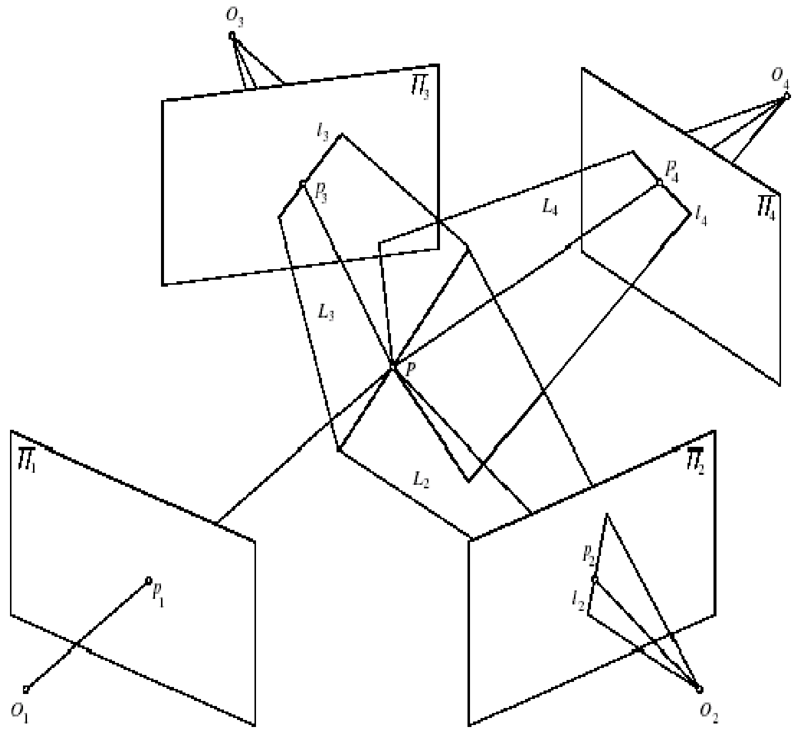


Figure 12.10. Given four images p_1, p_2, p_3 and p_4 of some point P and three arbitrary image lines l_2, l_3 and l_4 passing through the points p_2, p_3 and p_4, the ray passing through O_1 and p_1 must also pass through the point where the three planes L_2, L_3 and L_4 formed by the preimages of these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

Adapted from Trevor Darrell, MIT

Trifocal constraint with noise

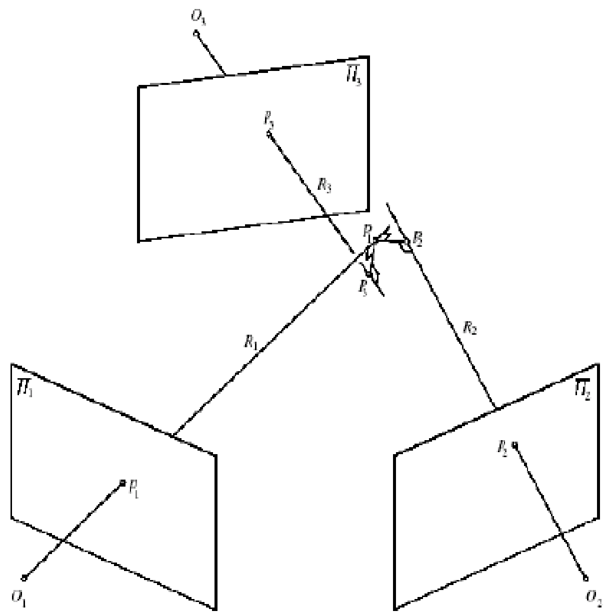


Figure 12.11. Trinocular constraints in the presence of calibration or measurement errors: the rays R_1, R_2 and R_3 may not intersect.

Adapted from Trevor Darrell, MIT