

# Multi-view geometry

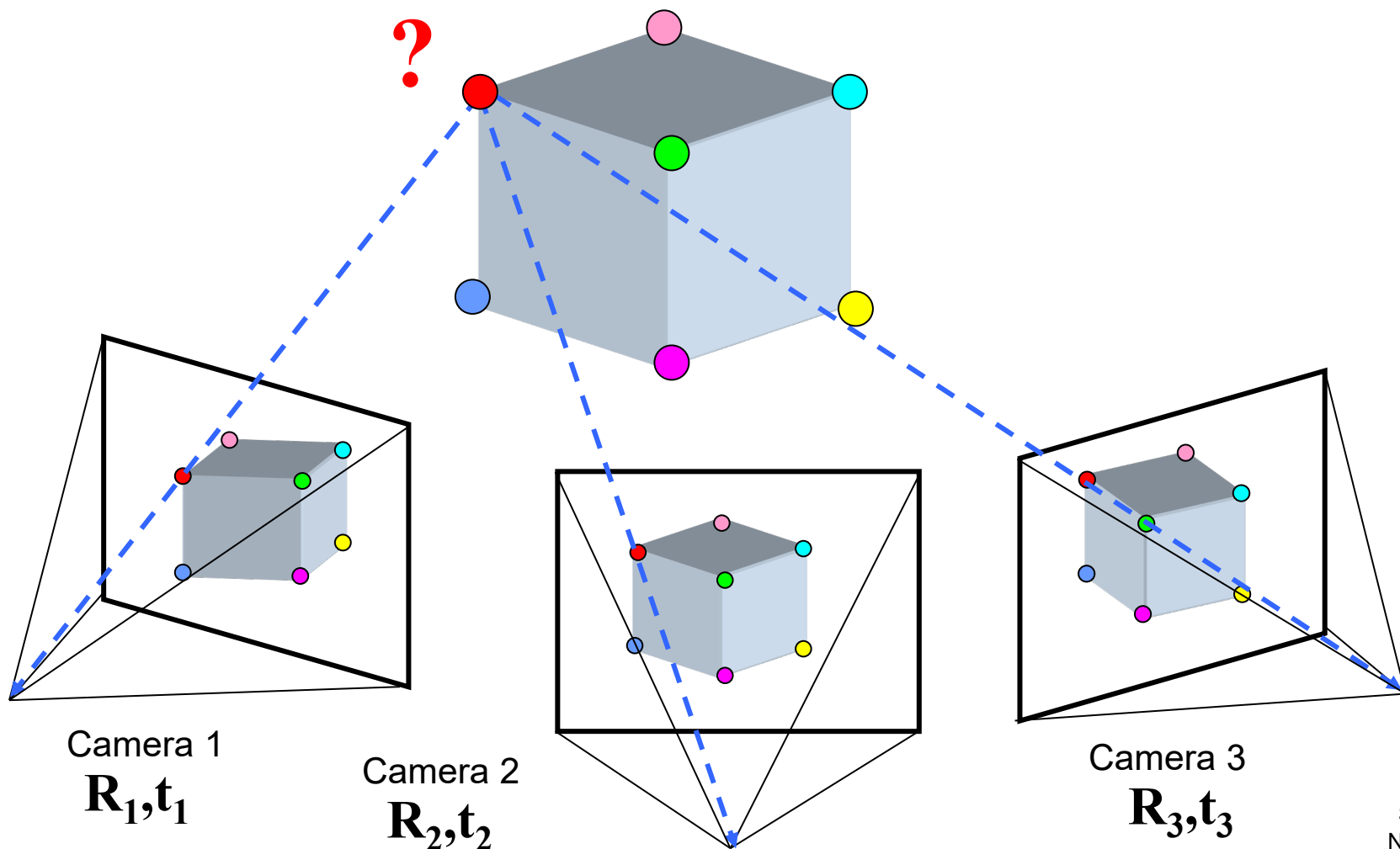
橫看成嶺側成峰  
遠近高低各不同  
不識廬山真面目  
只緣身在此山中

蘇軾詩 蘇軾書於津



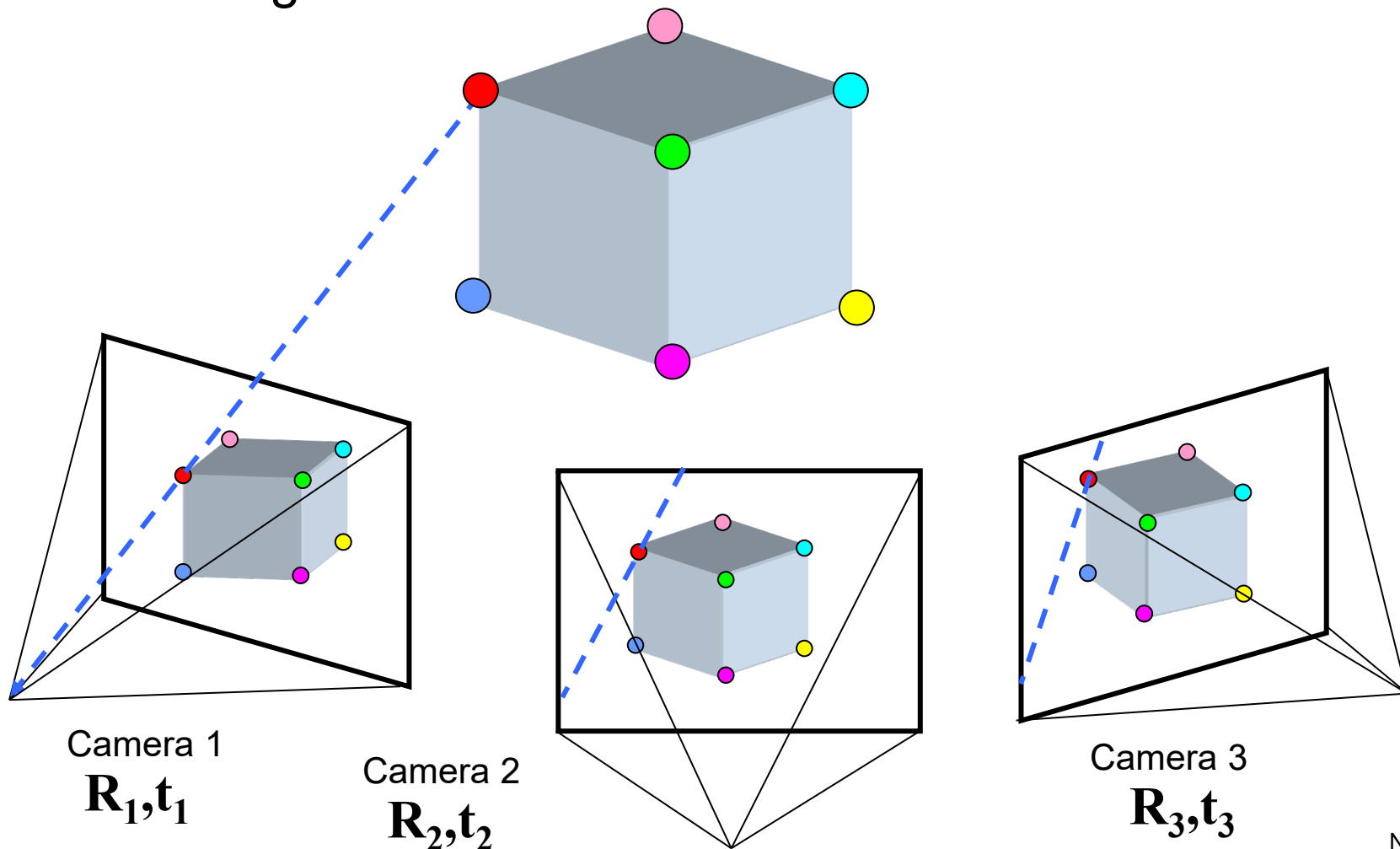
# Multi-view geometry problems

- **Structure:** Given projections of the same 3D point in two or more images, compute the 3D coordinates of that point



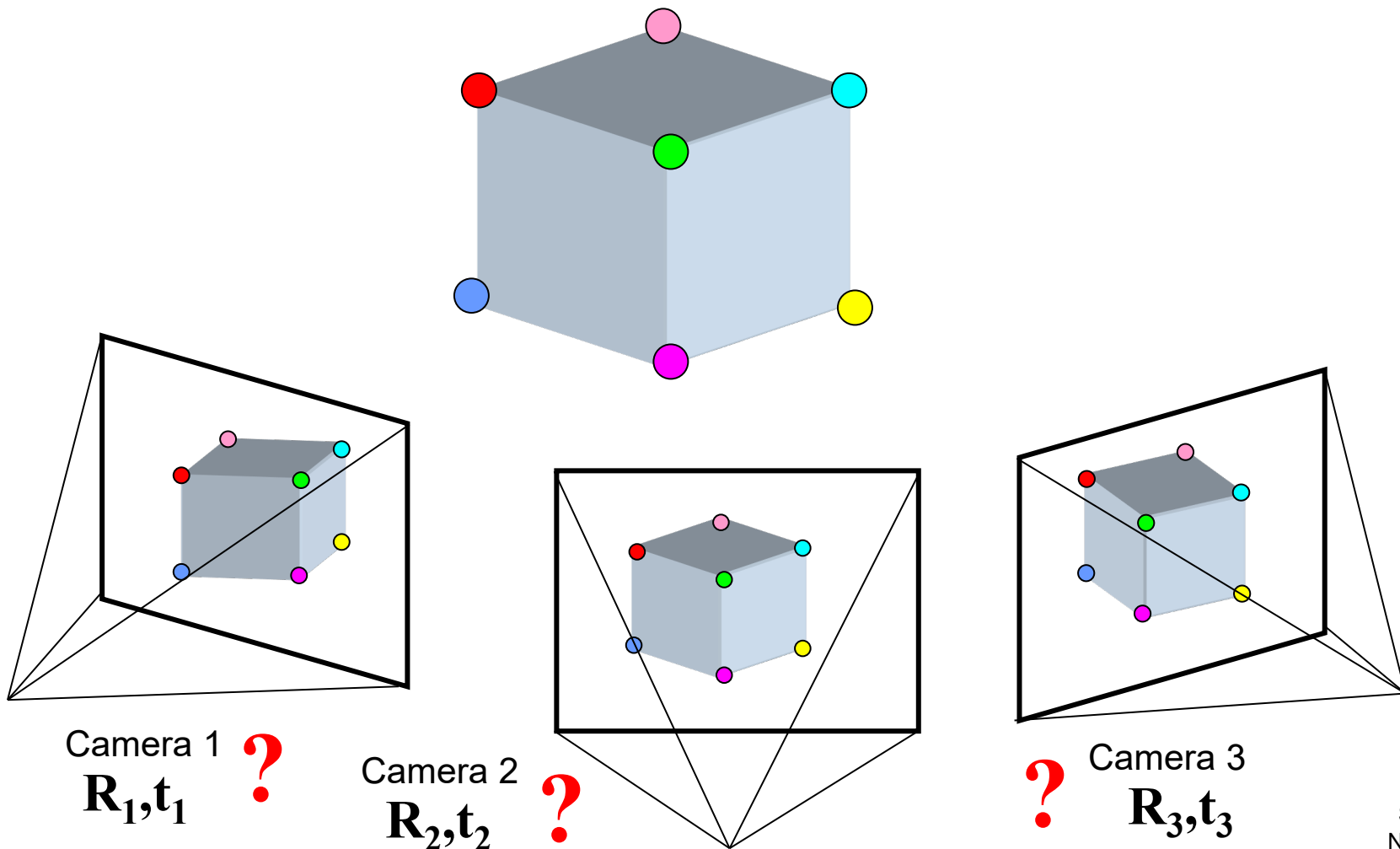
# Multi-view geometry problems

- **Stereo correspondence:** Given a point in one of the images, where could its corresponding points be in the other images?



# Multi-view geometry problems

- **Motion:** Given a set of corresponding points in two or more images, compute the camera parameters

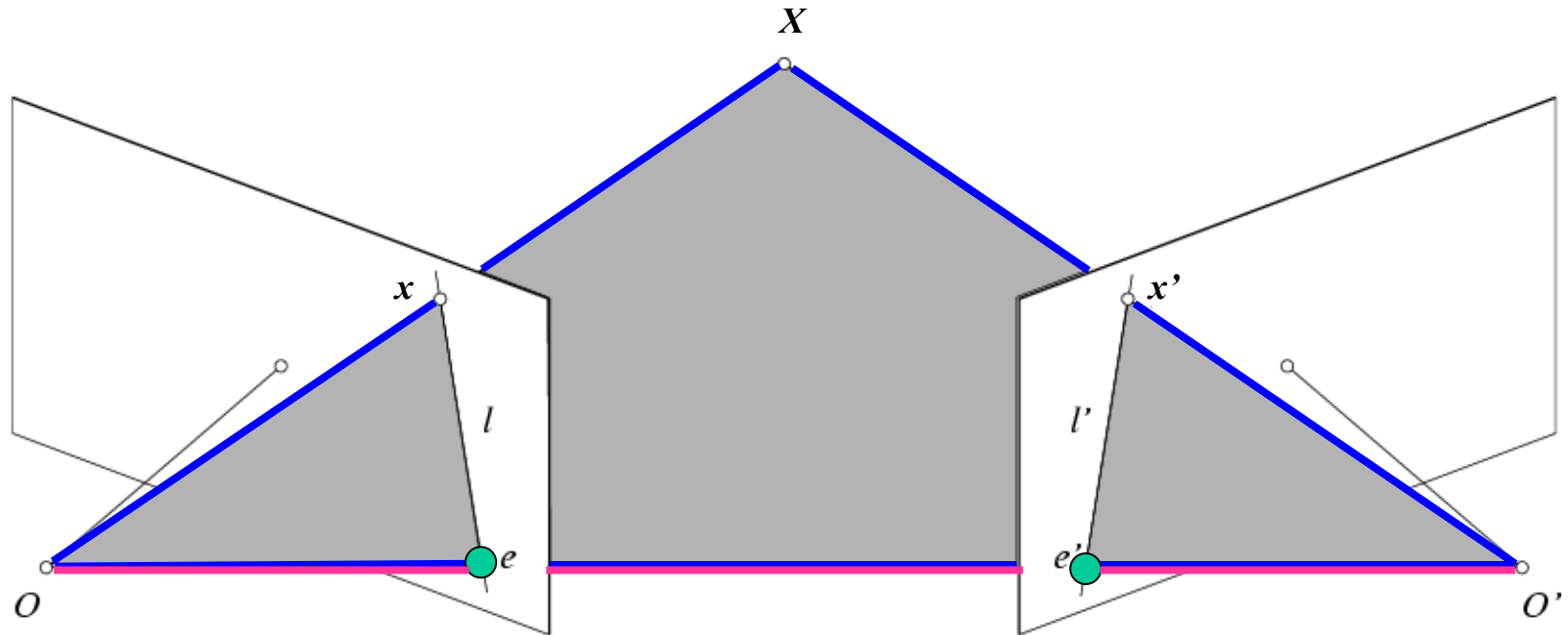


# Two-view geometry

---



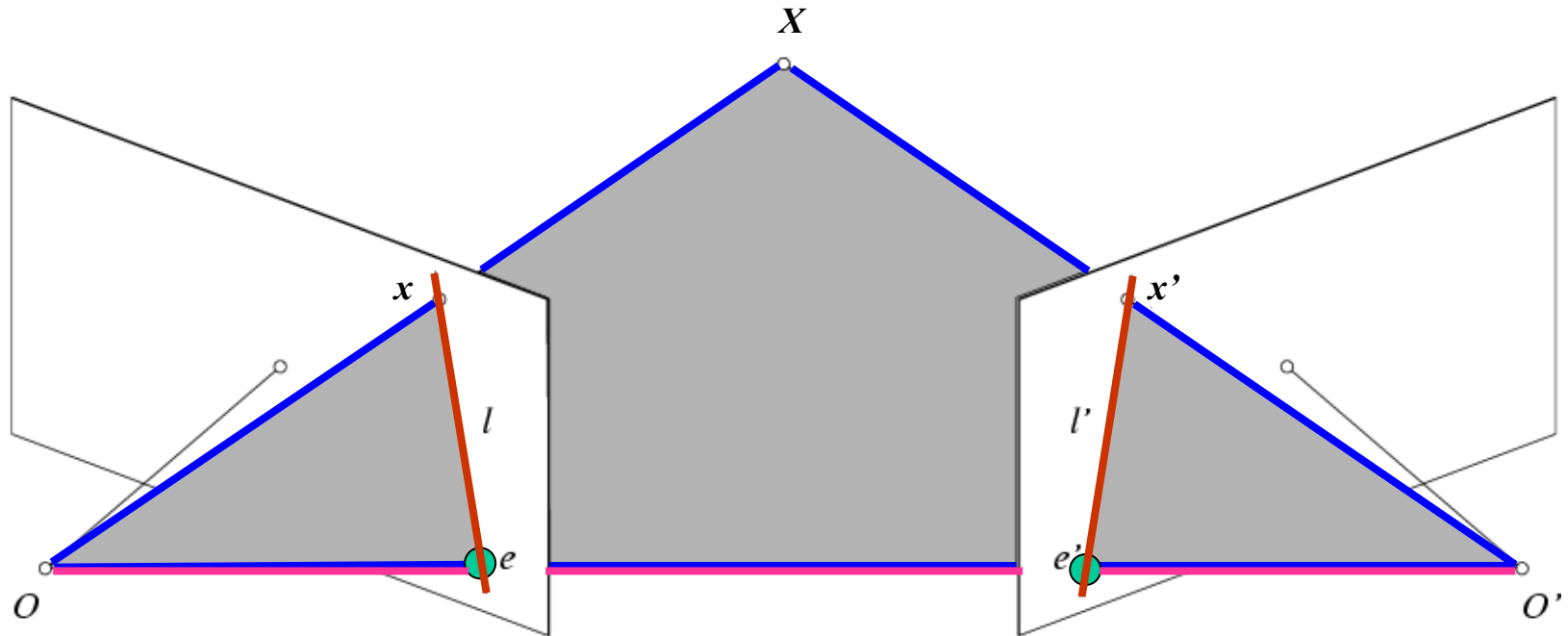
# Epipolar geometry 对极几何



- **Baseline** – line connecting the two camera centers
- **Epipolar Plane** – plane containing baseline (1D family)
- **Epipoles**
  - = intersections of baseline with image planes
  - = projections of the other camera center

# Epipolar geometry

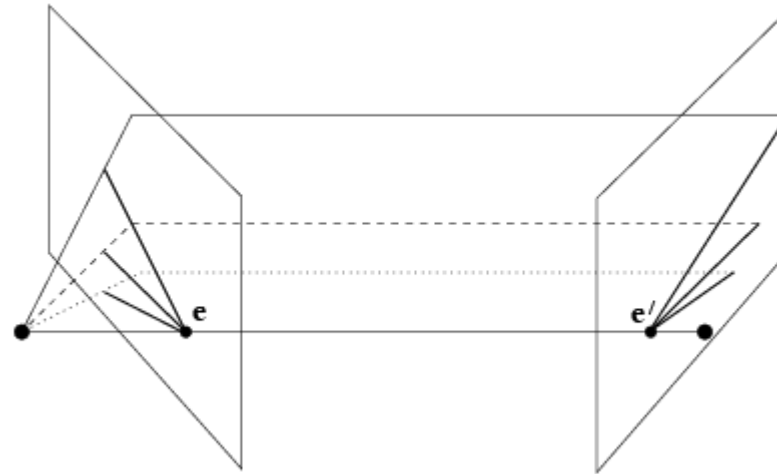
---



- **Baseline** – line connecting the two camera centers
- **Epipolar Plane** – plane containing baseline (1D family)
- **Epipoles**  
= intersections of baseline with image planes  
= projections of the other camera center
- **Epipolar Lines** - intersections of epipolar plane with image planes (always come in corresponding pairs)

# Example: Converging cameras

---





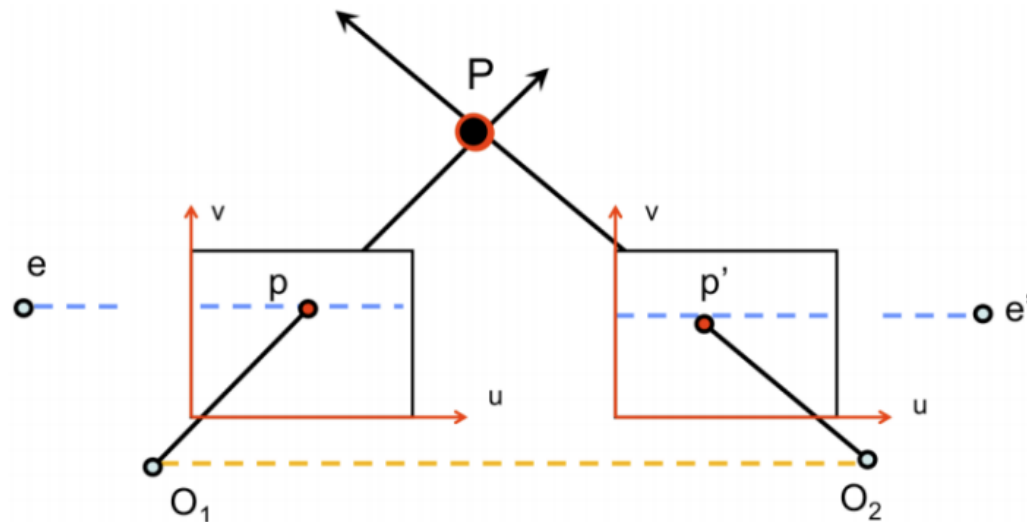


Figure 4: When the two image planes are parallel, then the epipoles  $e$  and  $e'$  are located at infinity. Notice that the epipolar lines are parallel to the  $u$  axis of each image plane.

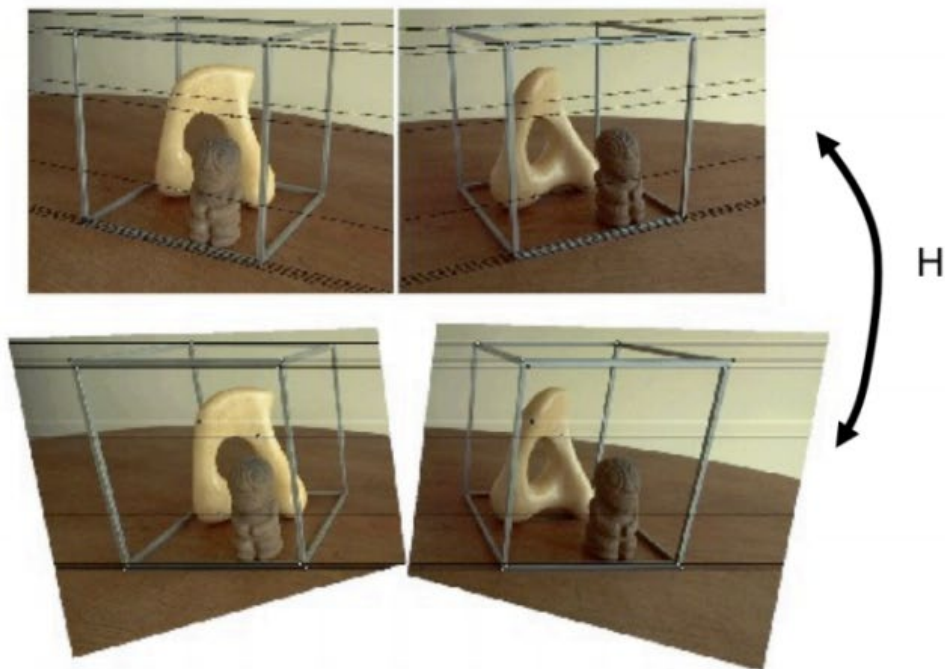
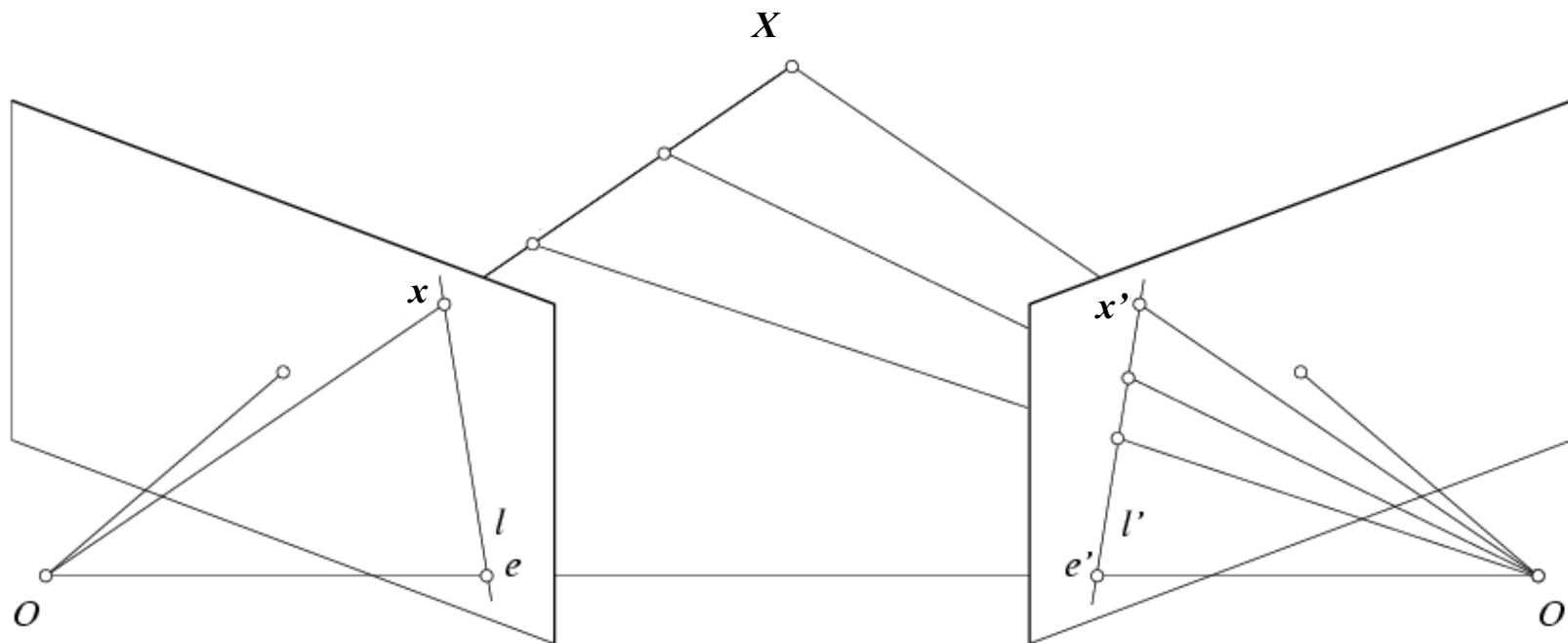


Figure 7: The process of image rectification involves computing two homographies that we can apply to a pair of images to make them parallel.

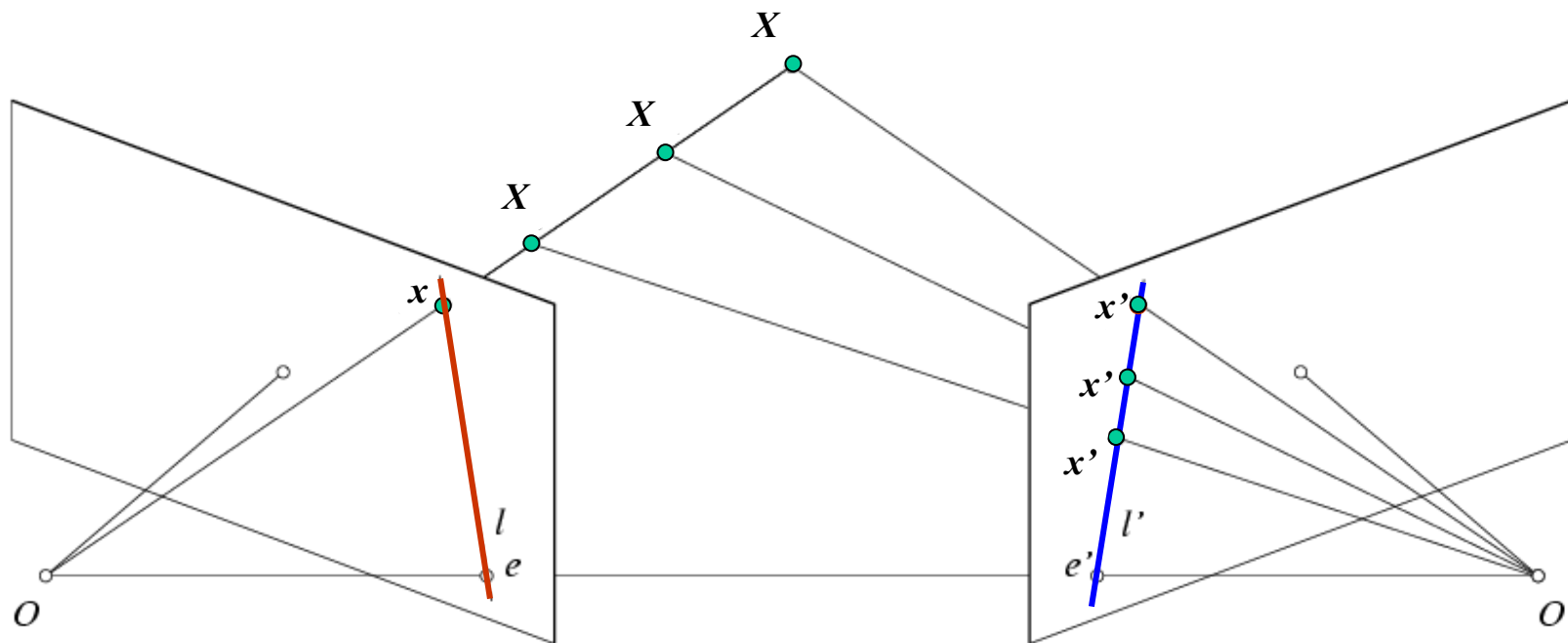
# Epipolar constraint

---



- If we observe a point  $\mathbf{x}$  in one image, where can the corresponding point  $\mathbf{x}'$  be in the other image?

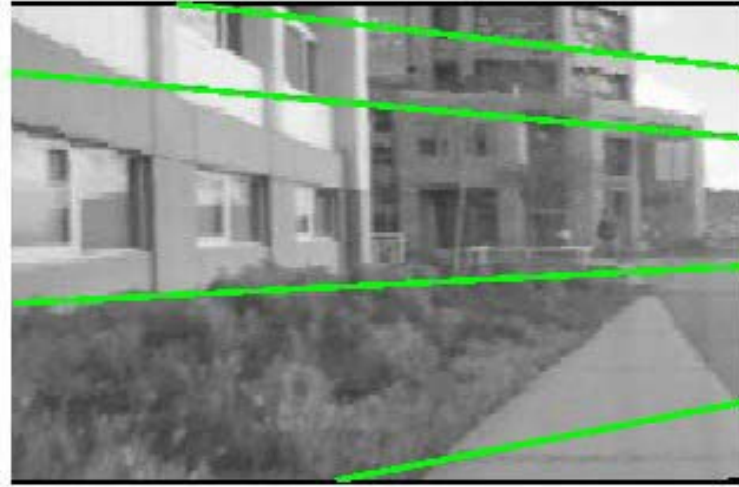
# Epipolar constraint



- Potential matches for  $\mathbf{x}$  have to lie on the corresponding epipolar line  $l'$ .
- Potential matches for  $\mathbf{x}'$  have to lie on the corresponding epipolar line  $l$ .

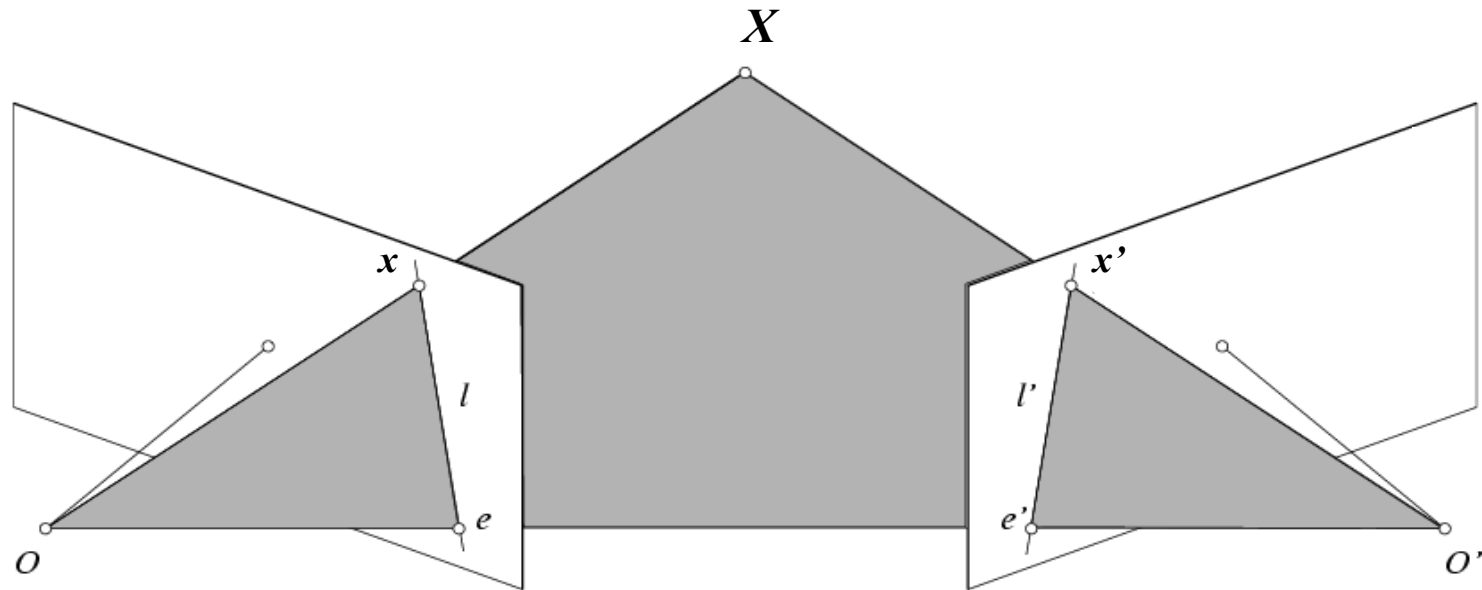
# Epipolar constraint example

---



# Epipolar constraint: Calibrated case

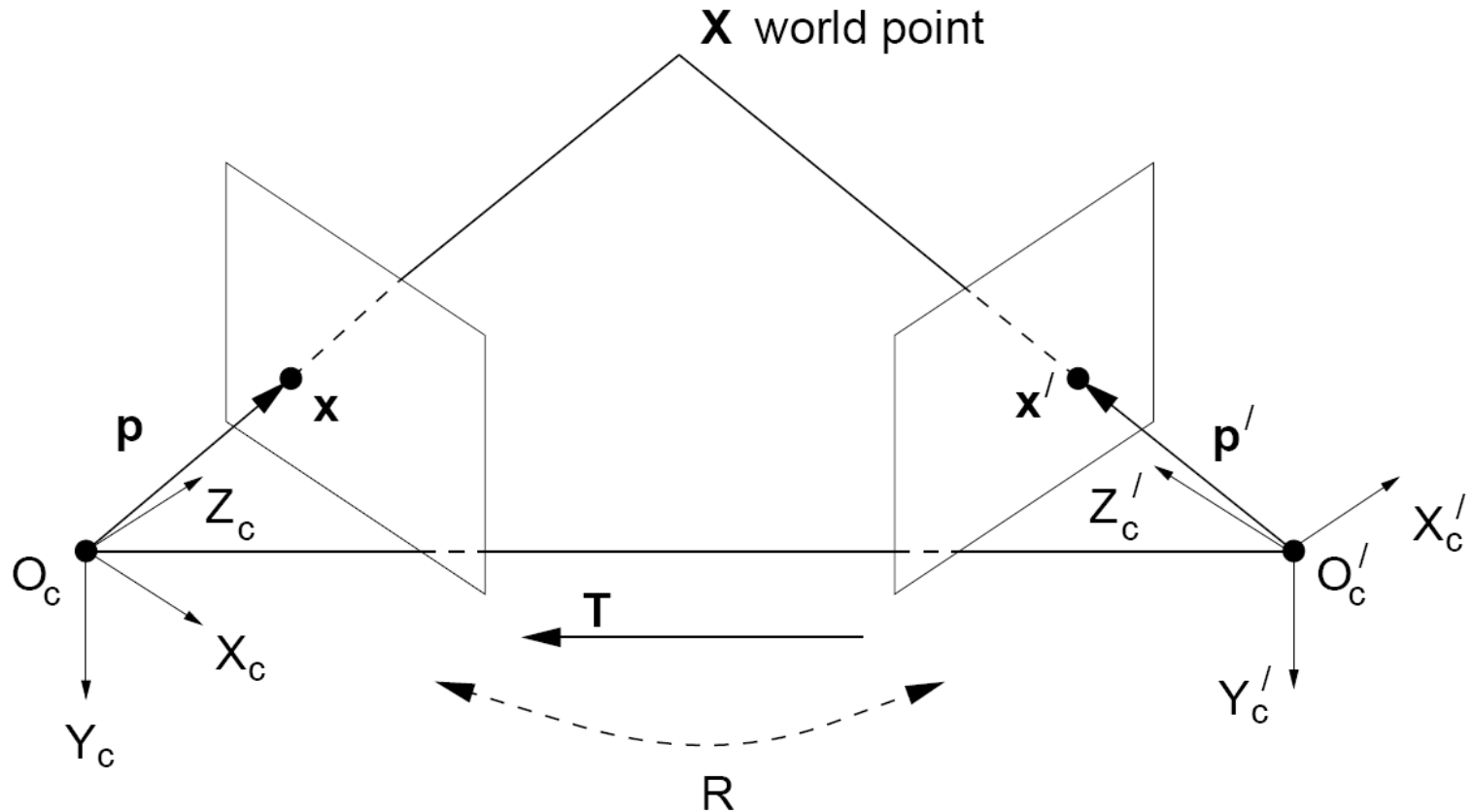
---



- Intrinsic and extrinsic parameters of the cameras are known, **world coordinate system is set to that of the first camera**
- Then the projection matrices are given by  $K[I \mid 0]$  and  $K'[R \mid t]$
- We can multiply the projection matrices (and the image points) by the inverse of the calibration matrices to get *normalized* image coordinates:

$$\mathbf{x}_{\text{norm}} = \mathbf{K}^{-1} \mathbf{x}_{\text{pixel}} = [\mathbf{I} \mid 0] \mathbf{X}, \quad \mathbf{x}'_{\text{norm}} = \mathbf{K}'^{-1} \mathbf{x}'_{\text{pixel}} = [\mathbf{R} \mid \mathbf{t}] \mathbf{X}$$

# Stereo geometry, with calibrated cameras



If the stereo rig is calibrated, we know :

how to **rotate** and **translate** camera reference frame 1 to get to camera reference frame 2.

$$\mathbf{X}'_c = \mathbf{R}\mathbf{X}_c + \mathbf{T}$$

# An aside: cross product

$$\vec{a} \times \vec{b} = \vec{c}$$

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

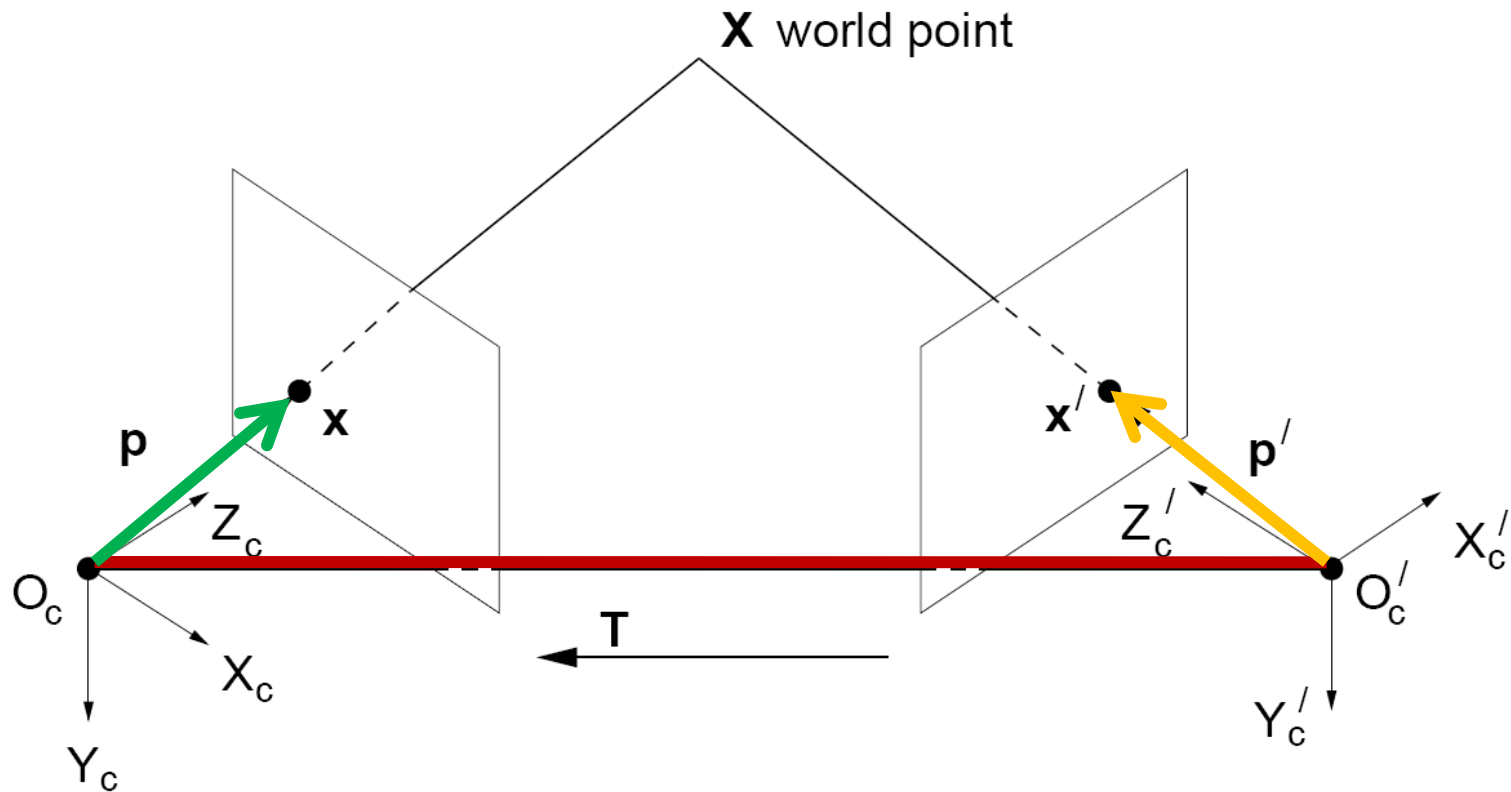
$$\vec{b} \cdot \vec{c} = 0$$

Vector cross product takes two vectors and returns a third vector that's perpendicular to both inputs.

So here, c is perpendicular to both a and b, which means the dot product = 0.



# From geometry to algebra



$$\boxed{X'} = \boxed{R} \boxed{X} + \boxed{T}$$

$$\underbrace{T \times X'}_{\text{Normal to the plane}} = T \times R X$$

$$X' \cdot (T \times X') = X' \cdot (T \times R X) = 0$$

# Another aside:

## Matrix form of cross product

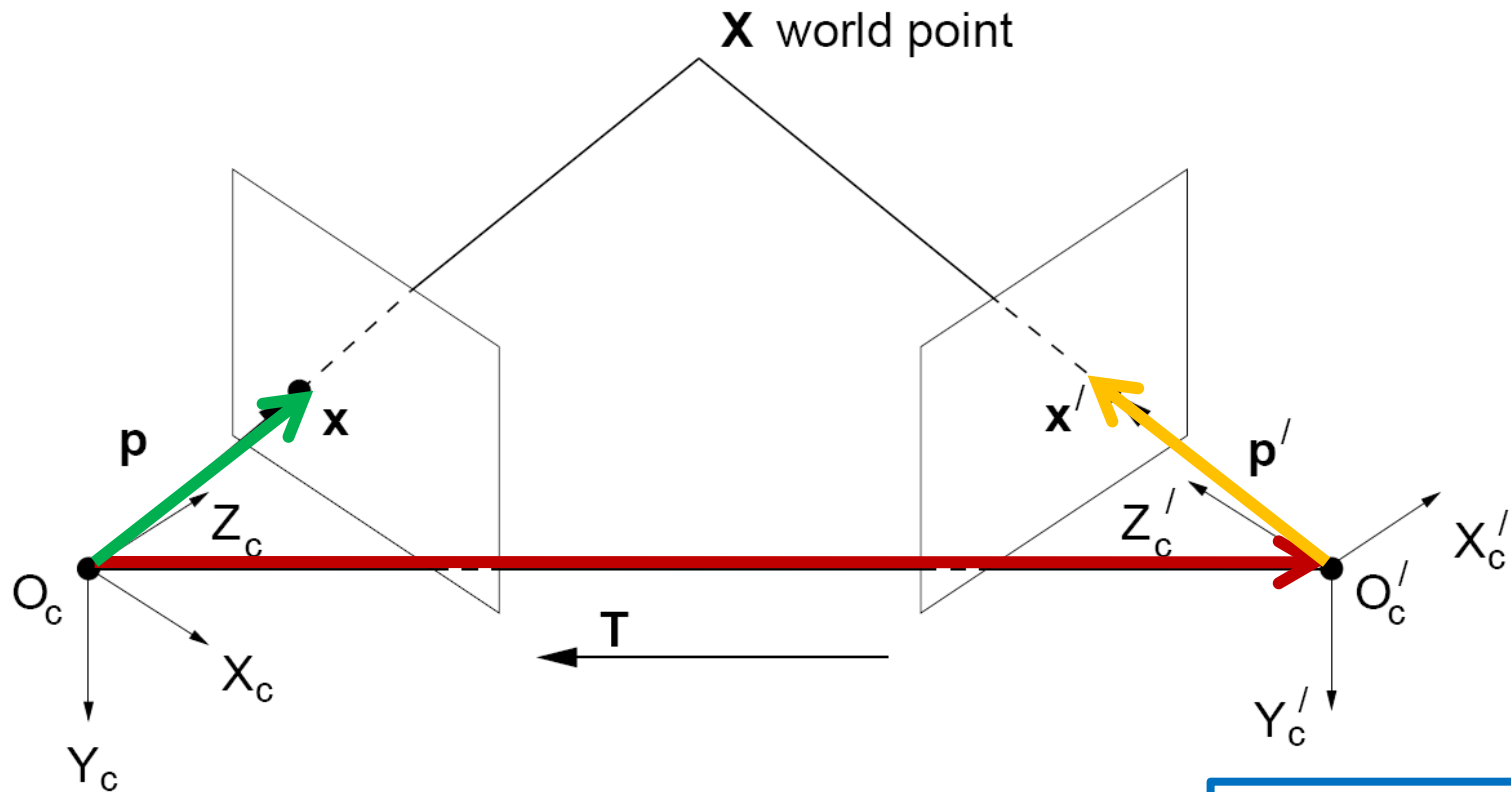
---

$$\vec{a} \times \vec{b} = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \vec{c} \quad \begin{array}{l} \vec{a} \cdot \vec{c} = 0 \\ \vec{b} \cdot \vec{c} = 0 \end{array}$$

Can be expressed as a matrix multiplication.

$$[a_x] = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix} \quad \boxed{\vec{a} \times \vec{b} = [a_x] \vec{b}}$$

# From geometry to algebra



$$\mathbf{X}' = \mathbf{R}\mathbf{X} + \mathbf{T}$$

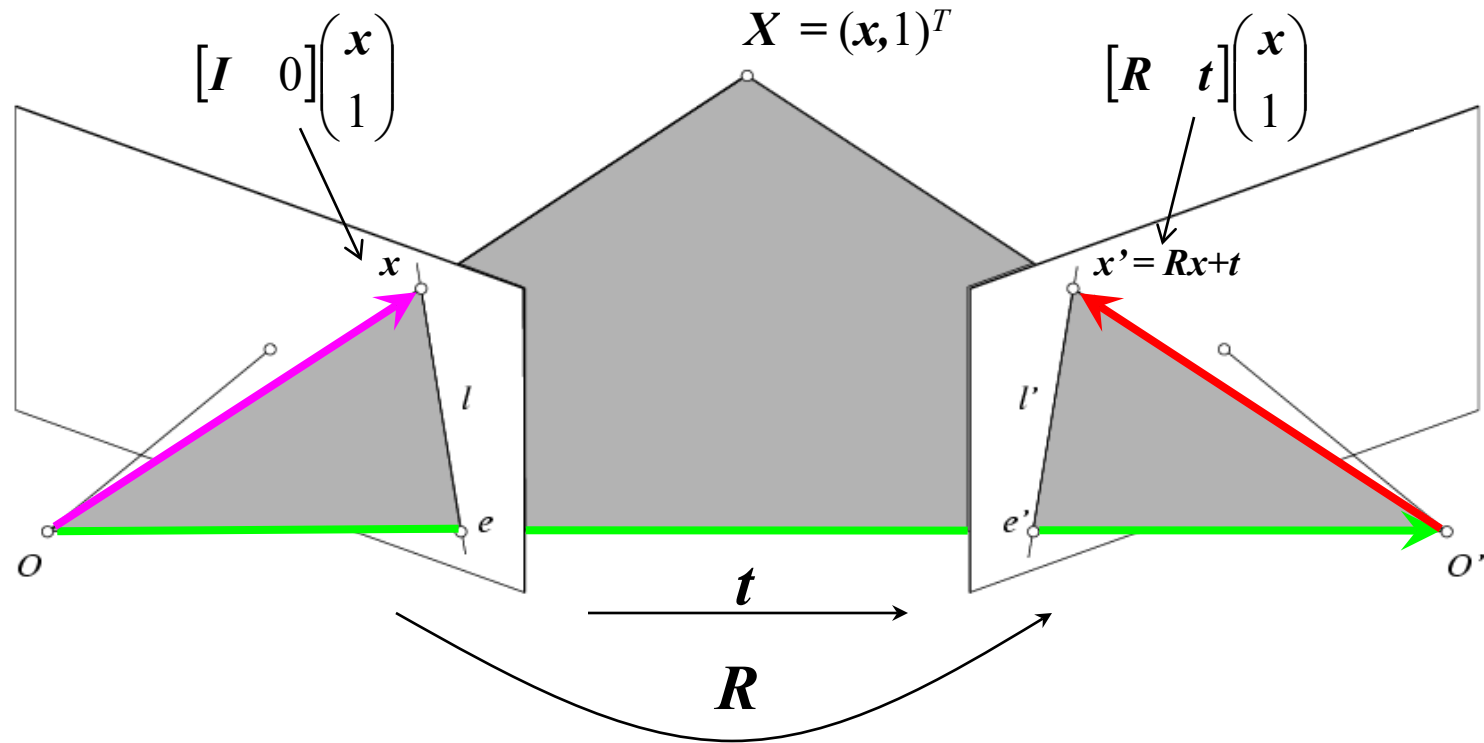
$$\underbrace{\mathbf{T} \times \mathbf{X}'}_{\text{Normal to the plane}} = \mathbf{T} \times \mathbf{R}\mathbf{X} + \mathbf{T} \times \mathbf{T}$$

$$= \mathbf{T} \times \mathbf{R}\mathbf{X}$$

$$\mathbf{X}' \cdot (\mathbf{T} \times \mathbf{X}') = \mathbf{X}' \cdot (\mathbf{T} \times \mathbf{R}\mathbf{X})$$

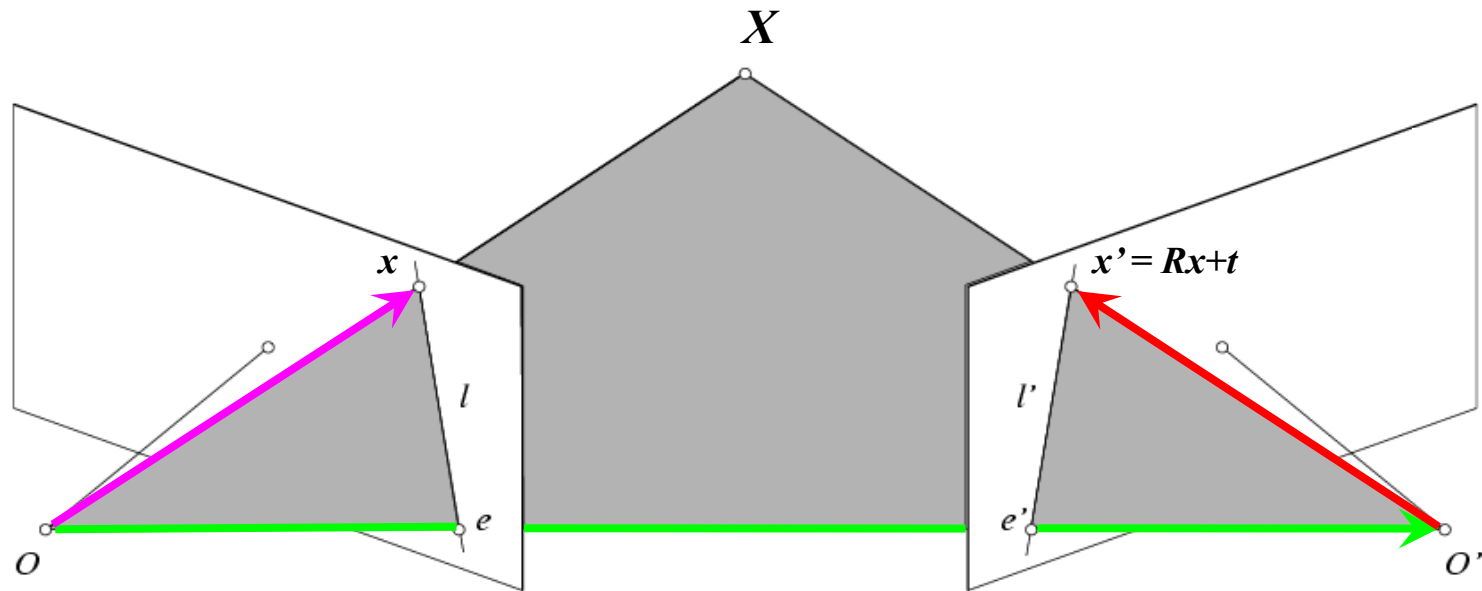
$$= 0$$

# Epipolar constraint: Calibrated case



The vectors  $Rx$ ,  $t$ , and  $x'$  are coplanar

# Epipolar constraint: Calibrated case

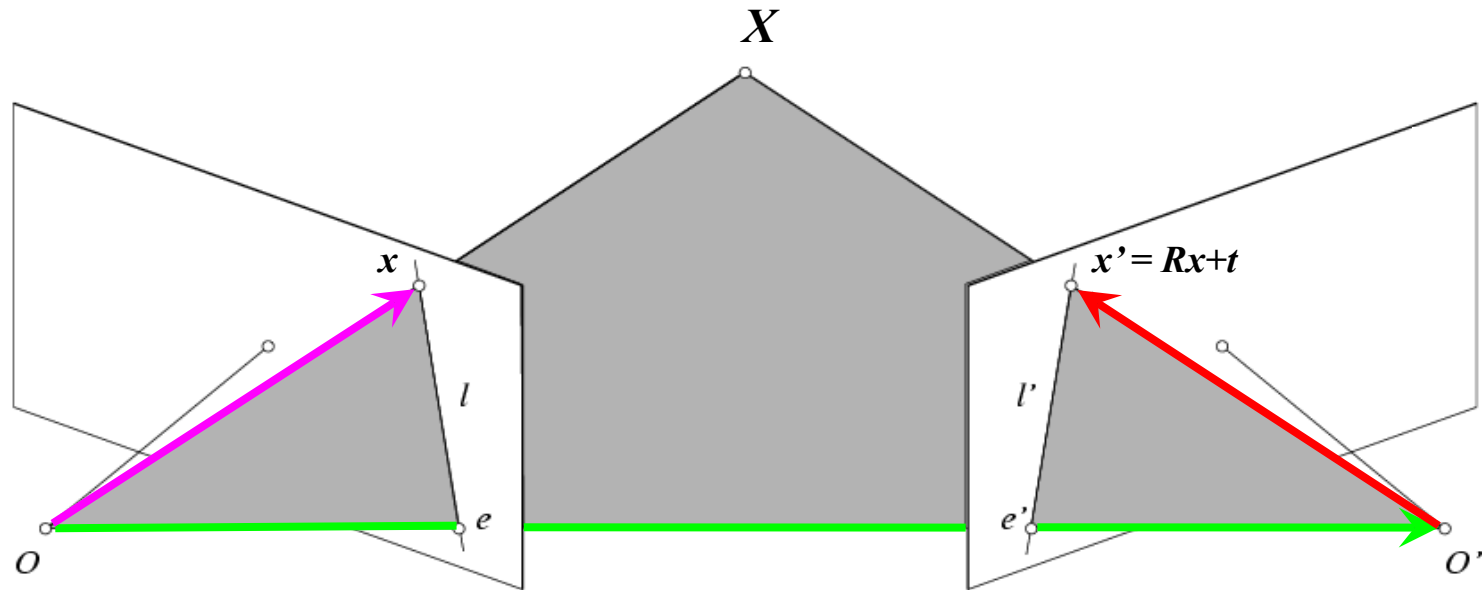


$$\mathbf{x}' \cdot [\mathbf{t} \times (R\mathbf{x})] = 0 \quad \Rightarrow \quad \mathbf{x}'^T [\mathbf{t}_\times] R\mathbf{x} = 0$$

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

The vectors  $R\mathbf{x}$ ,  $\mathbf{t}$ , and  $\mathbf{x}'$  are coplanar

# Epipolar constraint: Calibrated case



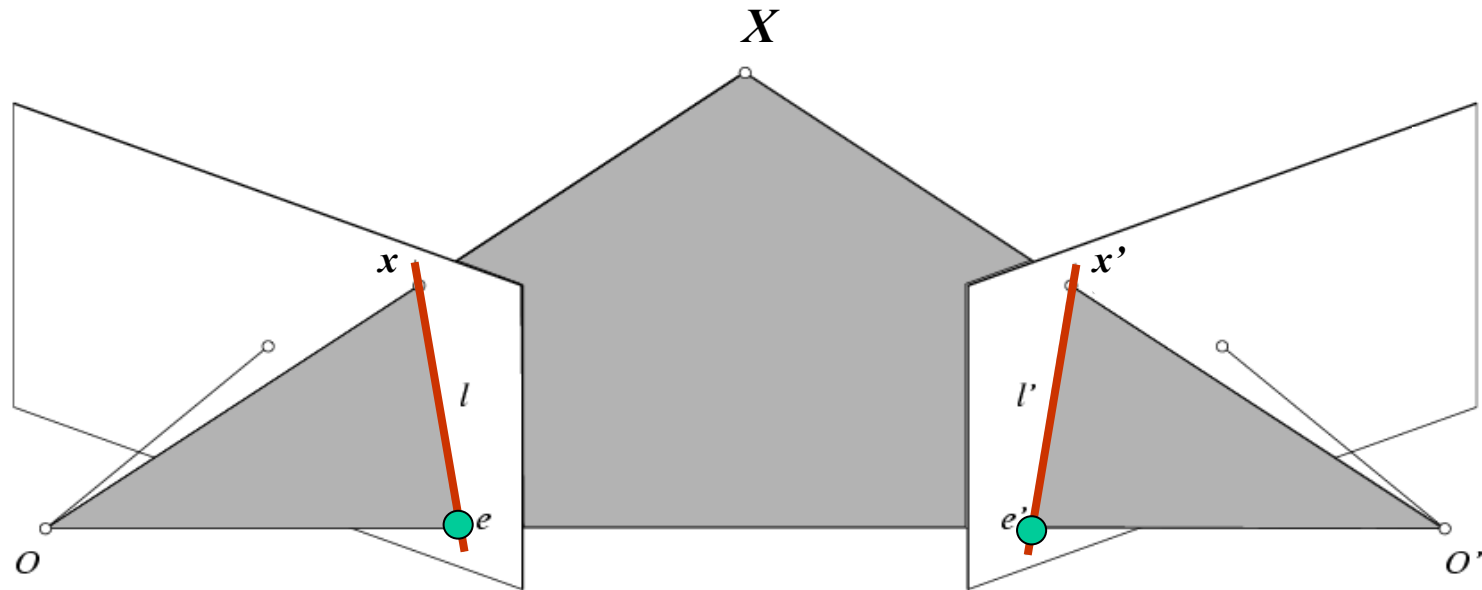
$$\mathbf{x}' \cdot [\mathbf{t} \times (R\mathbf{x})] = 0 \quad \Rightarrow \quad \mathbf{x}'^T [\mathbf{t}_\times] R \mathbf{x} = 0 \quad \Rightarrow \quad \mathbf{x}'^T E \mathbf{x} = 0$$

**Essential Matrix**  
(Longuet-Higgins, 1981)

The vectors  $R\mathbf{x}$ ,  $\mathbf{t}$ , and  $\mathbf{x}'$  are coplanar

# Epipolar constraint: Calibrated case

---

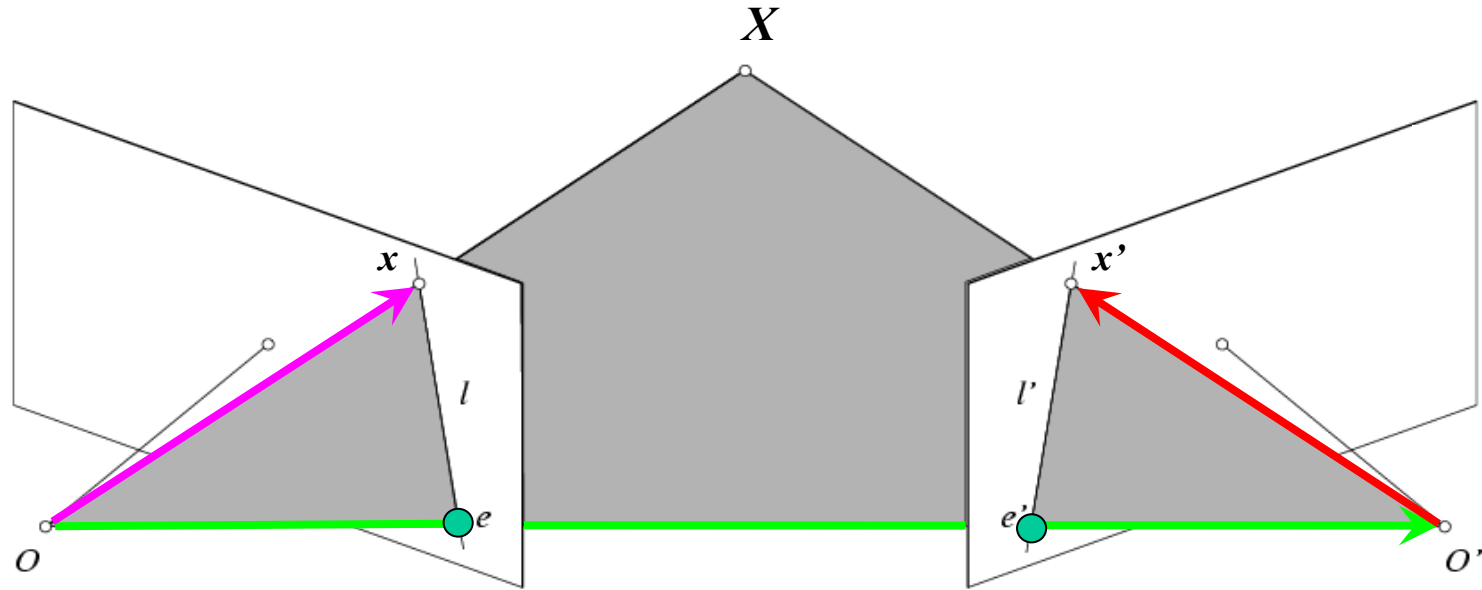


$$\mathbf{x}'^T \mathbf{E} \mathbf{x} = 0$$

- $\mathbf{E} \mathbf{x}$  is the epipolar line associated with  $\mathbf{x}$  ( $l' = \mathbf{E} \mathbf{x}$ )
- $\mathbf{E}^T \mathbf{x}'$  is the epipolar line associated with  $\mathbf{x}'$  ( $l = \mathbf{E}^T \mathbf{x}'$ )
- $\mathbf{E} \mathbf{e} = 0$  and  $\mathbf{E}^T \mathbf{e}' = 0$
- $\mathbf{E}$  is singular (rank two)
- $\mathbf{E}$  has five degrees of freedom

# Epipolar constraint: Uncalibrated case

---

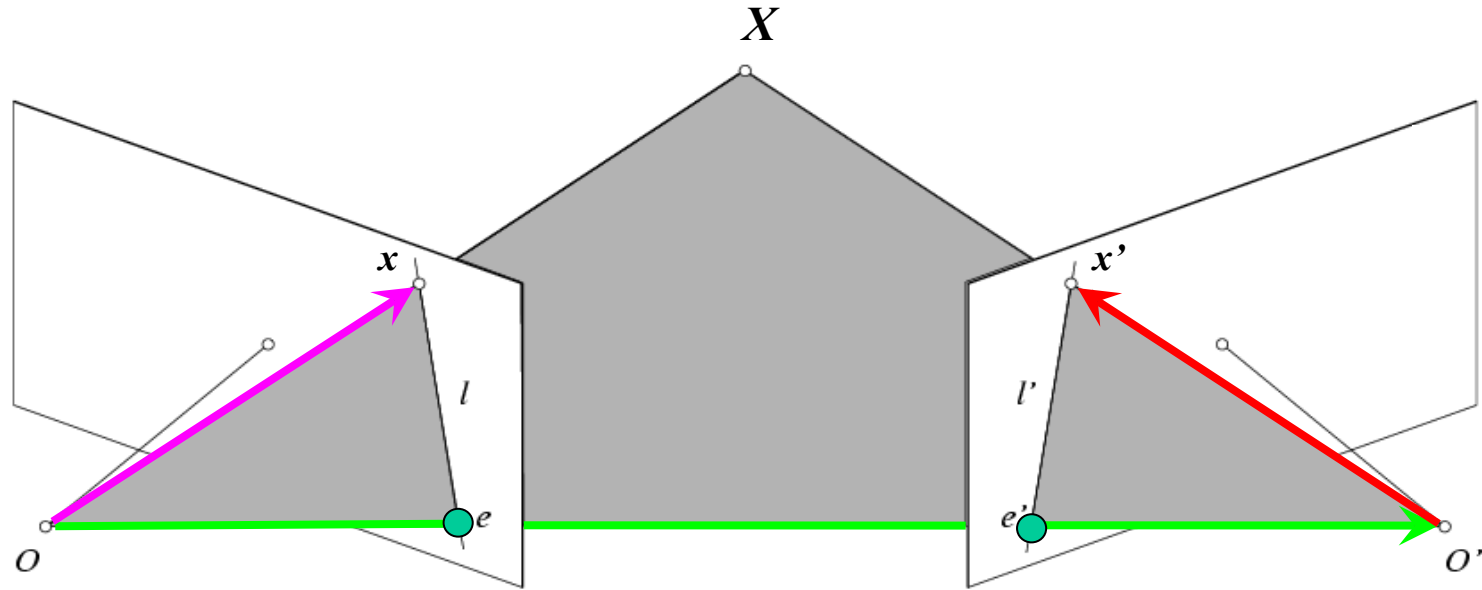


- The calibration matrices  $\mathbf{K}$  and  $\mathbf{K}'$  of the two cameras are unknown
- We can write the epipolar constraint in terms of *unknown* normalized coordinates:

$$\hat{\mathbf{x}}'^T \mathbf{E} \hat{\mathbf{x}} = 0 \qquad \hat{\mathbf{x}} = \mathbf{K}^{-1} \mathbf{x}, \quad \hat{\mathbf{x}}' = \mathbf{K}'^{-1} \mathbf{x}'$$



# Epipolar constraint: Uncalibrated case



$$\hat{x}'^T E \hat{x} = 0 \quad \Rightarrow \quad x'^T F x = 0 \quad \text{with} \quad F = K'^{-T} E K^{-1}$$

$$\hat{x} = K^{-1} x$$

$$\hat{x}' = K'^{-1} x'$$

**Fundamental Matrix**  
(Faugeras and Luong, 1992)

# Estimating the fundamental matrix

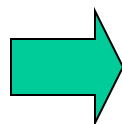
---



# The eight-point algorithm

$$\mathbf{x} = (u, v, 1)^T, \quad \mathbf{x}' = (u', v', 1)$$

$$\begin{bmatrix} u' & v' & 1 \end{bmatrix} \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = 0$$



$$\begin{bmatrix} u'u & u'v & u' & v'u & v'v & v' & u & v & 1 \end{bmatrix} \begin{bmatrix} f_{11} \\ f_{12} \\ f_{13} \\ f_{21} \\ f_{22} \\ f_{23} \\ f_{31} \\ f_{32} \\ f_{33} \end{bmatrix} = 0$$

Solve homogeneous linear system using eight or more matches



Enforce rank-2 constraint (take SVD of  $\mathbf{F}$  and throw out the smallest singular value)



# Problem with eight-point algorithm

---

$$\begin{bmatrix} u'u & u'v & u' & v'u & v'v & v' & u & v \end{bmatrix} \begin{bmatrix} f_{11} \\ f_{12} \\ f_{13} \\ f_{21} \\ f_{22} \\ f_{23} \\ f_{31} \\ f_{32} \end{bmatrix} = -1$$

# Problem with eight-point algorithm

---

250906.36	183269.57	921.81	200931.10	146766.13	738.21	272.19	198.81
2692.28	131633.03	176.27	6196.73	302975.59	405.71	15.27	746.79
416374.23	871684.30	935.47	408110.89	854384.92	916.90	445.10	931.81
191183.60	171759.40	410.27	416435.62	374125.90	893.65	465.99	418.65
48988.86	30401.76	57.89	298604.57	185309.58	352.87	846.22	525.15
164786.04	546559.67	813.17	1998.37	6628.15	9.86	202.65	672.14
116407.01	2727.75	138.89	169941.27	3982.21	202.77	838.12	19.64
135384.58	75411.13	198.72	411350.03	229127.78	603.79	681.28	379.48

$$\begin{bmatrix} f_{11} \\ f_{12} \\ f_{13} \\ f_{21} \\ f_{22} \\ f_{23} \\ f_{31} \\ f_{32} \end{bmatrix} = -1$$

Poor numerical conditioning

Can be fixed by rescaling the data

# The normalized eight-point algorithm

---

(Hartley, 1995)

- Center the image data at the origin, and scale it so the mean squared distance between the origin and the data points is 2 pixels
- Use the eight-point algorithm to compute  $\mathbf{F}$  from the normalized points
- Enforce the rank-2 constraint (for example, take SVD of  $\mathbf{F}$  and throw out the smallest singular value)
- Transform fundamental matrix back to original units: if  $\mathbf{T}$  and  $\mathbf{T}'$  are the normalizing transformations in the two images, then the fundamental matrix in original coordinates is  $\mathbf{T}'^T \mathbf{F} \mathbf{T}$

# Nonlinear estimation

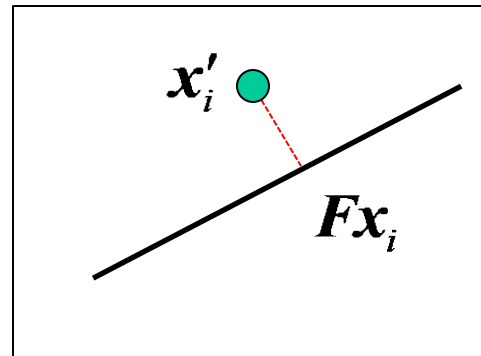
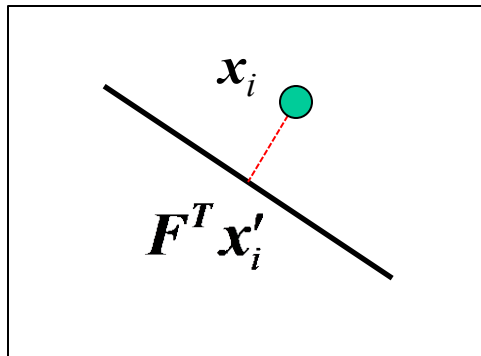
---

- Linear estimation minimizes the sum of squared *algebraic* distances between points  $\mathbf{x}'_i$  and epipolar lines  $\mathbf{F} \mathbf{x}_i$  (or points  $\mathbf{x}_i$  and epipolar lines  $\mathbf{F}^T \mathbf{x}'_i$ ):

$$\sum_{i=1}^N (\mathbf{x}'_i{}^T \mathbf{F} \mathbf{x}_i)^2$$

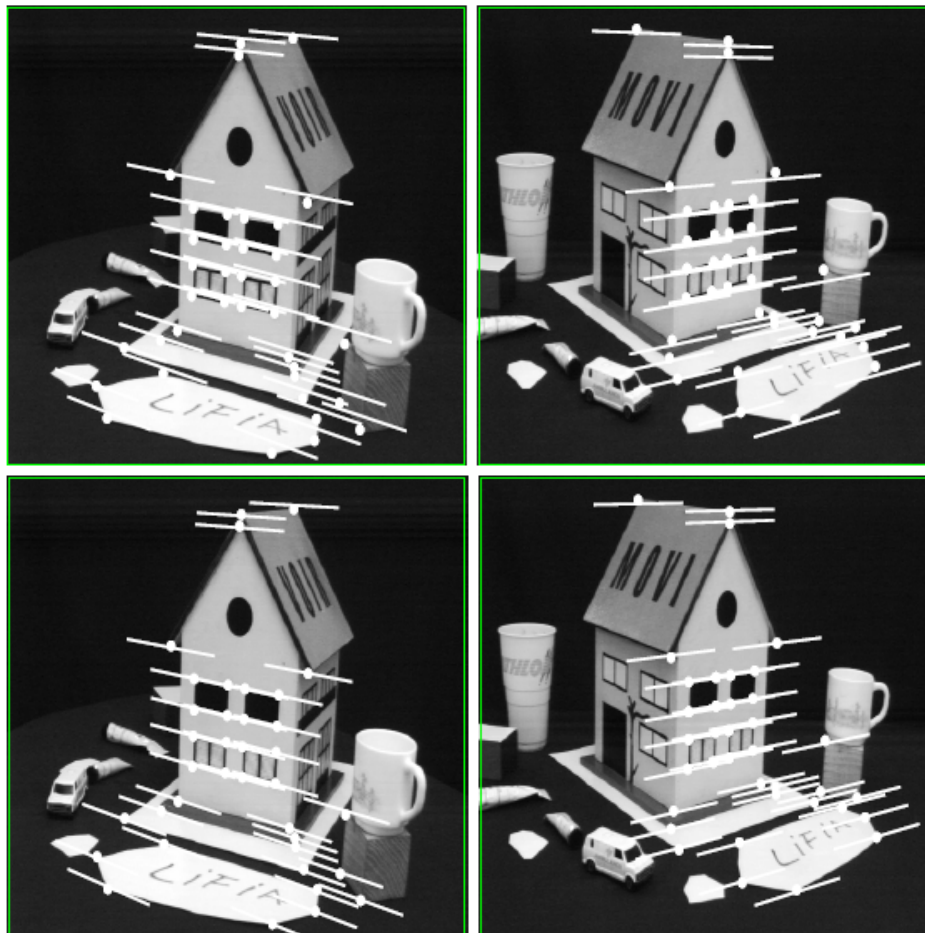
- Nonlinear approach: minimize sum of squared *geometric* distances

$$\sum_{i=1}^N \left[ d^2(\mathbf{x}'_i, \mathbf{F} \mathbf{x}_i) + d^2(\mathbf{x}_i, \mathbf{F}^T \mathbf{x}'_i) \right]$$



# Comparison of estimation algorithms

---

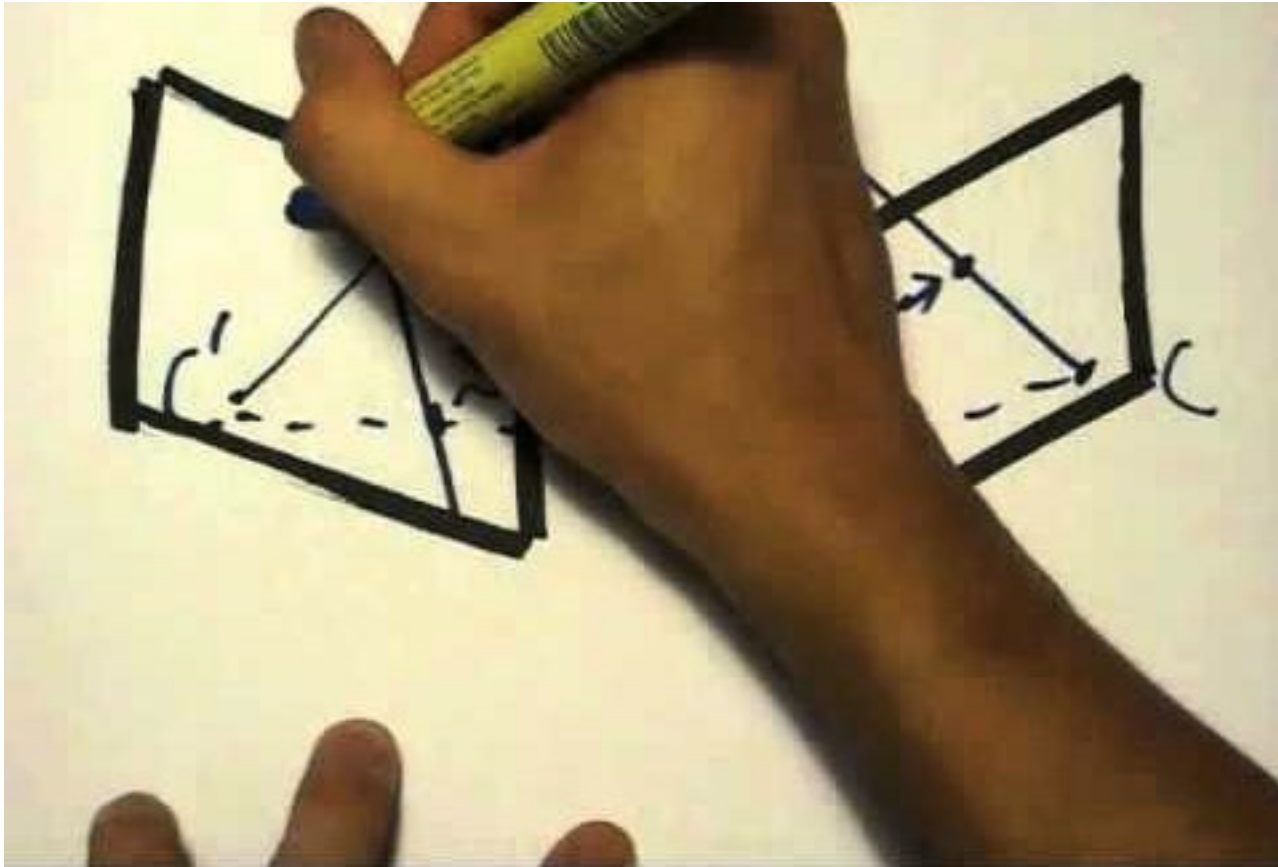


	8-point	Normalized 8-point	Nonlinear least squares
Av. Dist. 1	2.33 pixels	0.92 pixel	0.86 pixel
Av. Dist. 2	2.18 pixels	0.85 pixel	0.80 pixel



# The Fundamental Matrix Song

---



<http://danielwedge.com/fmatrix/>

# The Fundamental Matrix Song.

---

The [fundamental matrix](#)

Used in [stereo geometry](#)

A matrix with nine entries

It's square with size 3 by 3

Has seven [degrees of freedom](#)

It has a [rank deficiency](#)

It's only of rank two

Call the matrix  $F$  and you'll see...

[Two points that correspond](#)

Column vectors called  $x$  and  $x'$

$x'^T F x$  equals [zero](#) every time

The [epipolar constraint](#)

Involves epipolar lines

Postmultiplying  $F$  by  $x$

Results in vector  $l'$

It's the epipolar line

In the other view passing through  $x'$

A three component vector

Of [homogeneous design](#)

The left and right [nullspaces](#) of  $F$

Are the [epipoles](#)  $e$  and  $e'$

All of the epipolar lines

Should pass through these

Here's a linear estimation example:

Take a set of 8 point samples

Construct a matrix, take the [SVD](#)

And the elements of  $F$  are in the [last column of  \$V\$](#)

If you try to estimate

$F$  with a coplanar set of points

Your sample set will be degenerate

And will not bring you joy

When doing the estimation

If you don't perform [rank deprivation](#)

Your epipolar lines

And the epipoles will not coincide

But if your scene has three views

The [trifocal tensor](#) is what you'd use

Constraints from the third view act like glue

That can't be determined from just two views

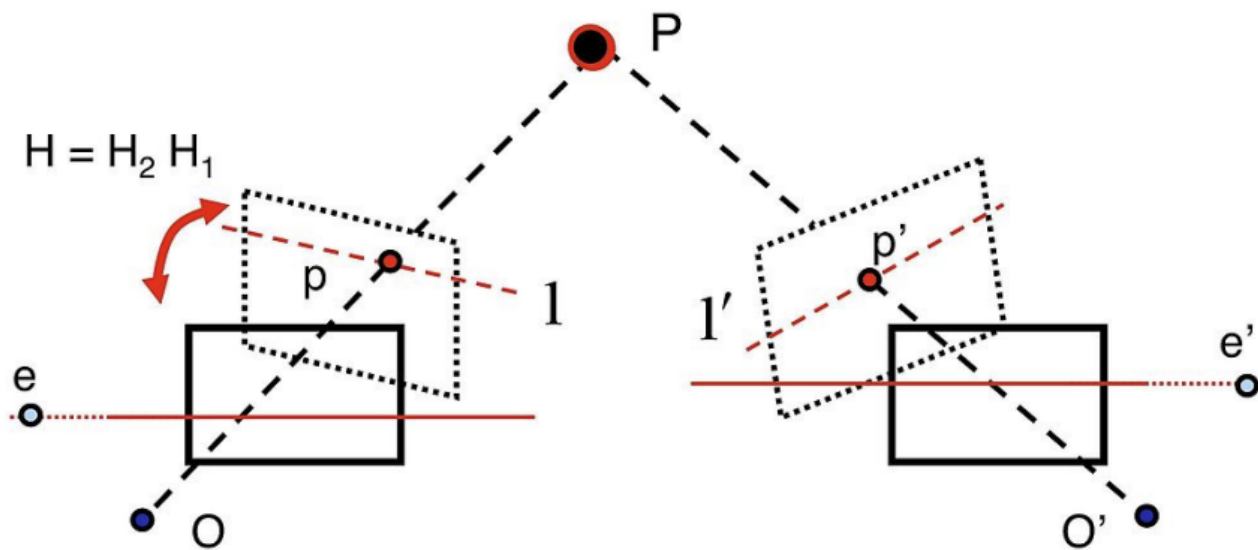


Figure 8: The rectification problem setup: we compute two homographies that we can apply to the image planes to make the resulting planes parallel.

# From epipolar geometry to camera calibration

---

- Estimating the fundamental matrix is known as “weak calibration”
- If we know the calibration matrices of the two cameras, we can estimate the essential matrix:  $\mathbf{E} = \mathbf{K}'^T \mathbf{F} \mathbf{K}$
- The essential matrix gives us the relative rotation and translation between the cameras, or their extrinsic parameters.