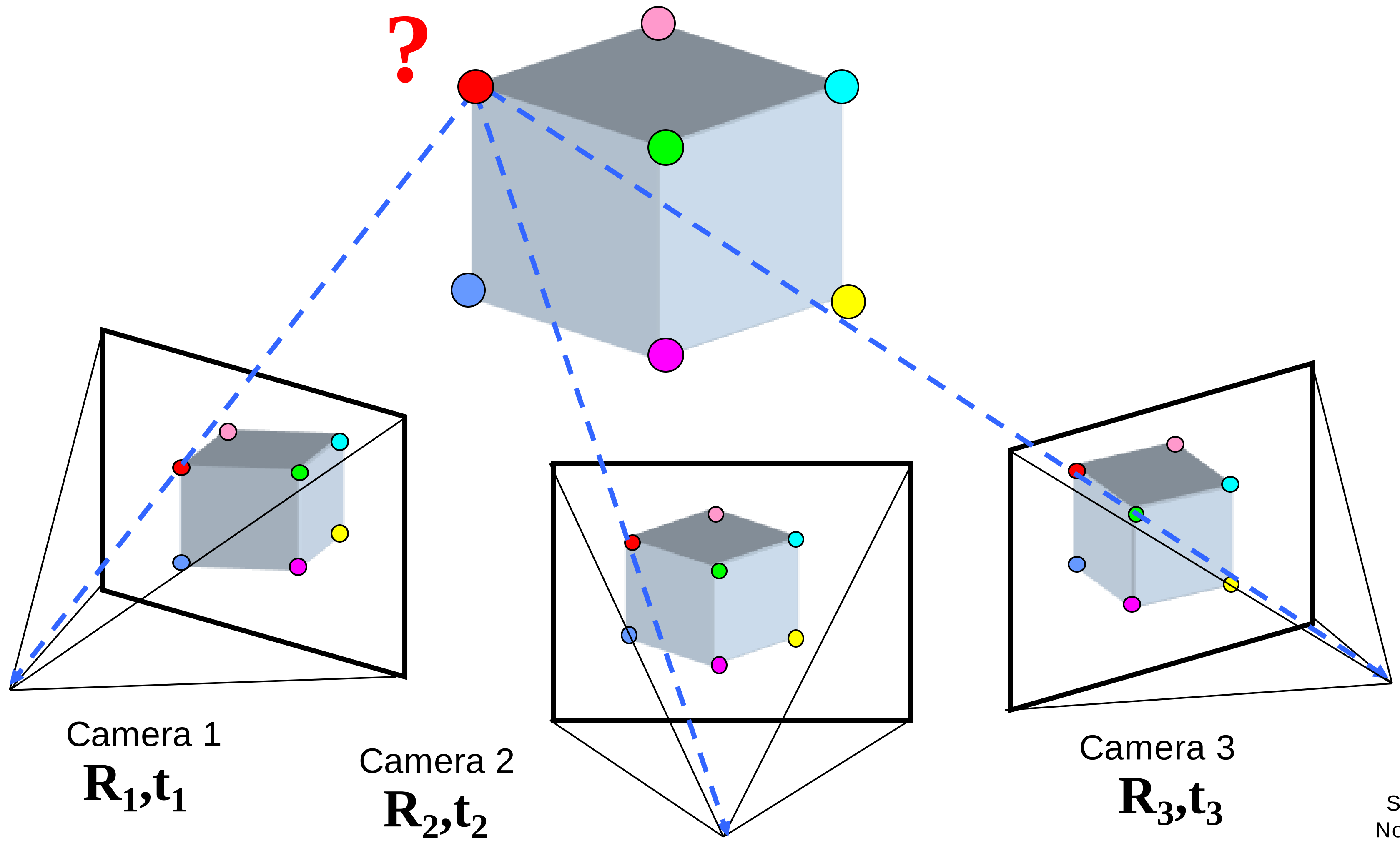
Multi-view geometry

Multi-view geometry problems

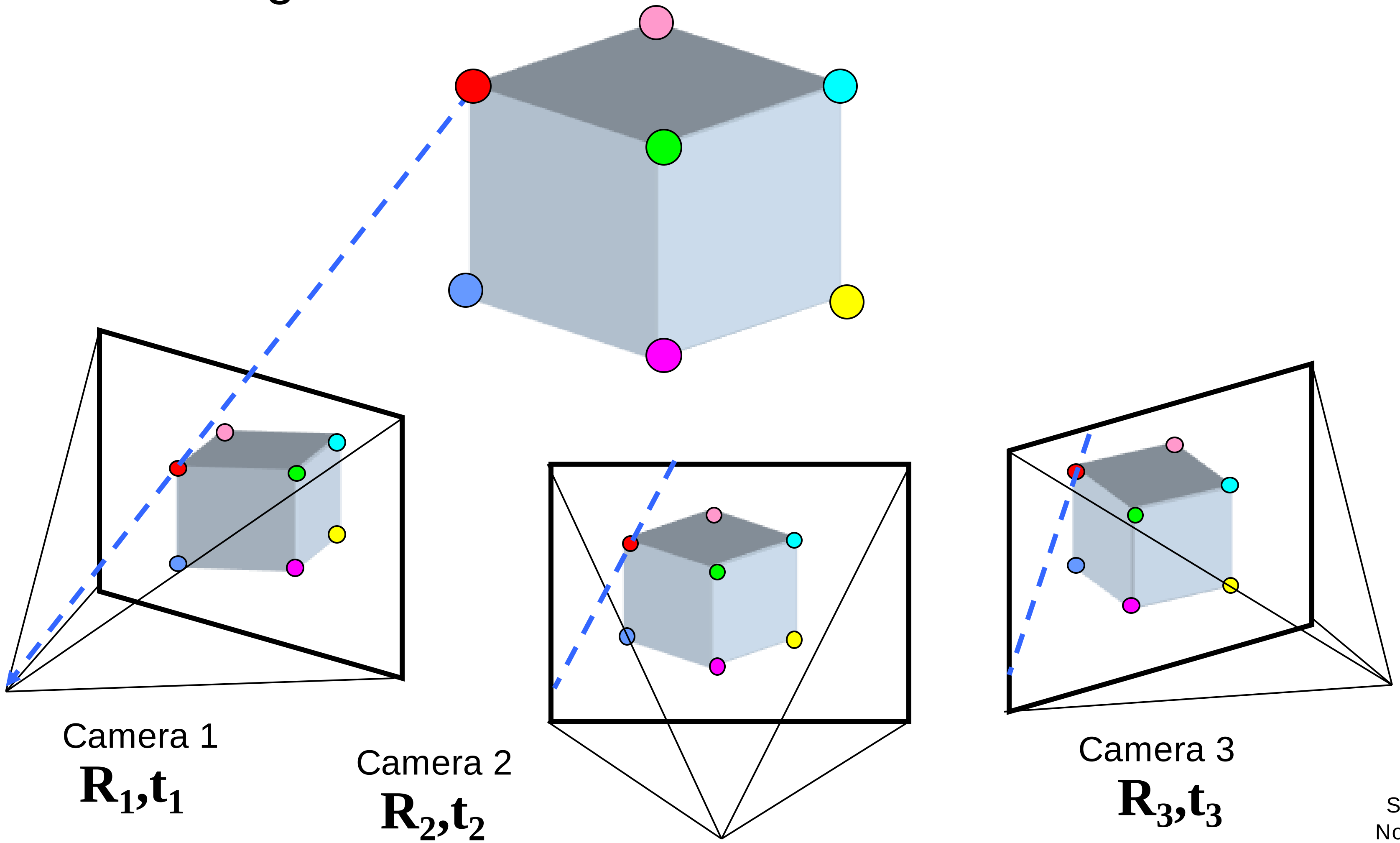
- **Structure:** Given projections of the same 3D point in two or more images, compute the 3D coordinates of that point



Slide credit: Noah Snavely

Multi-view geometry problems

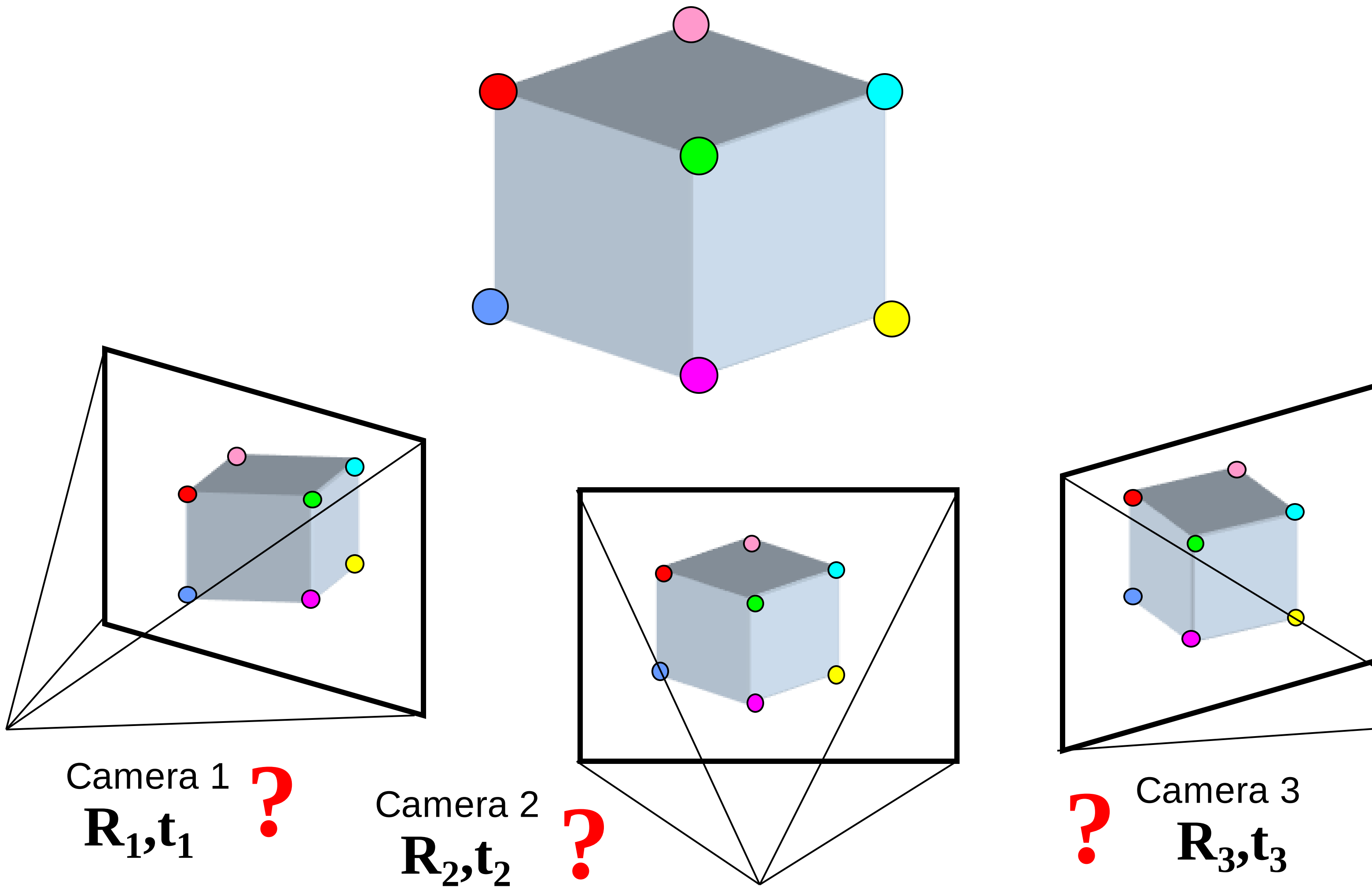
- **Stereo correspondence:** Given a point in one of the images, where could its corresponding points be in the other images?



Slide credit: Noah Snavely

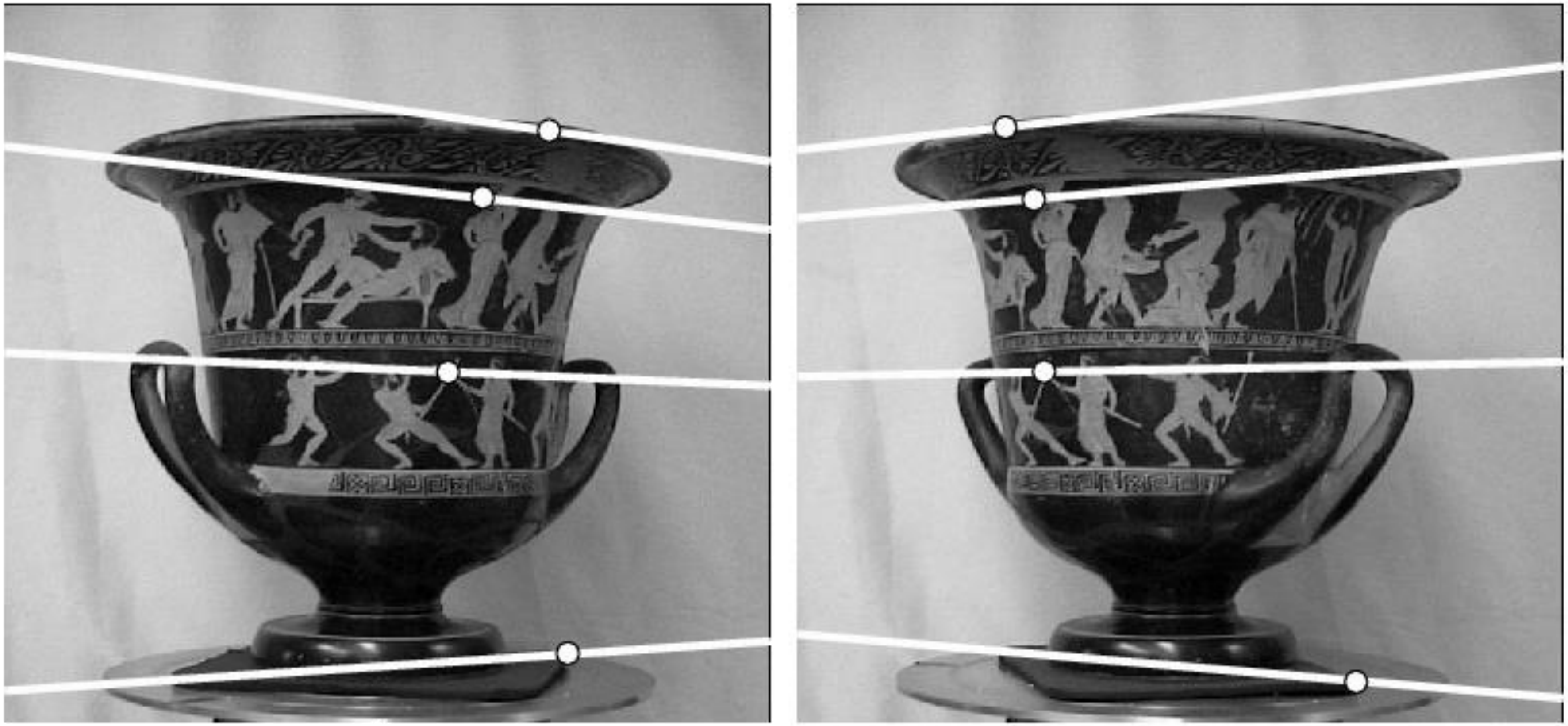
Multi-view geometry problems

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

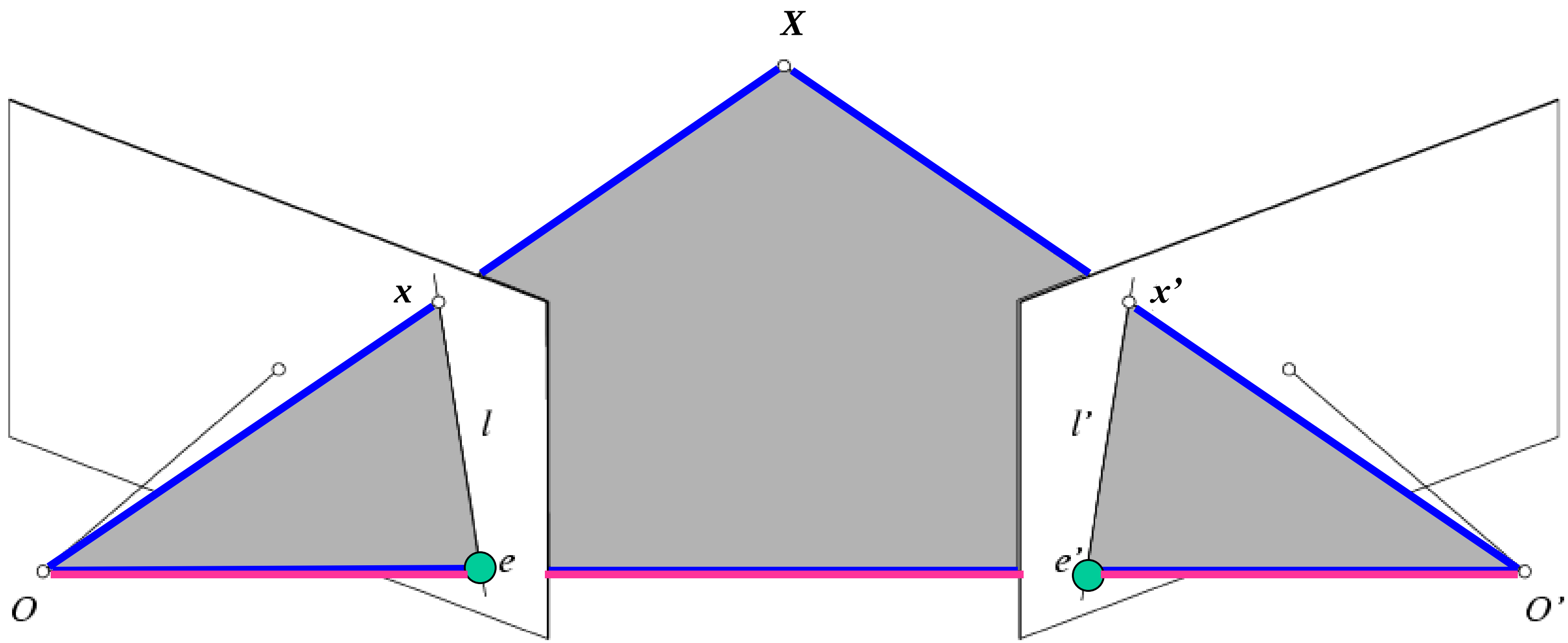


Slide credit: Noah Snavely

Two-view geometry



Epipolar geometry

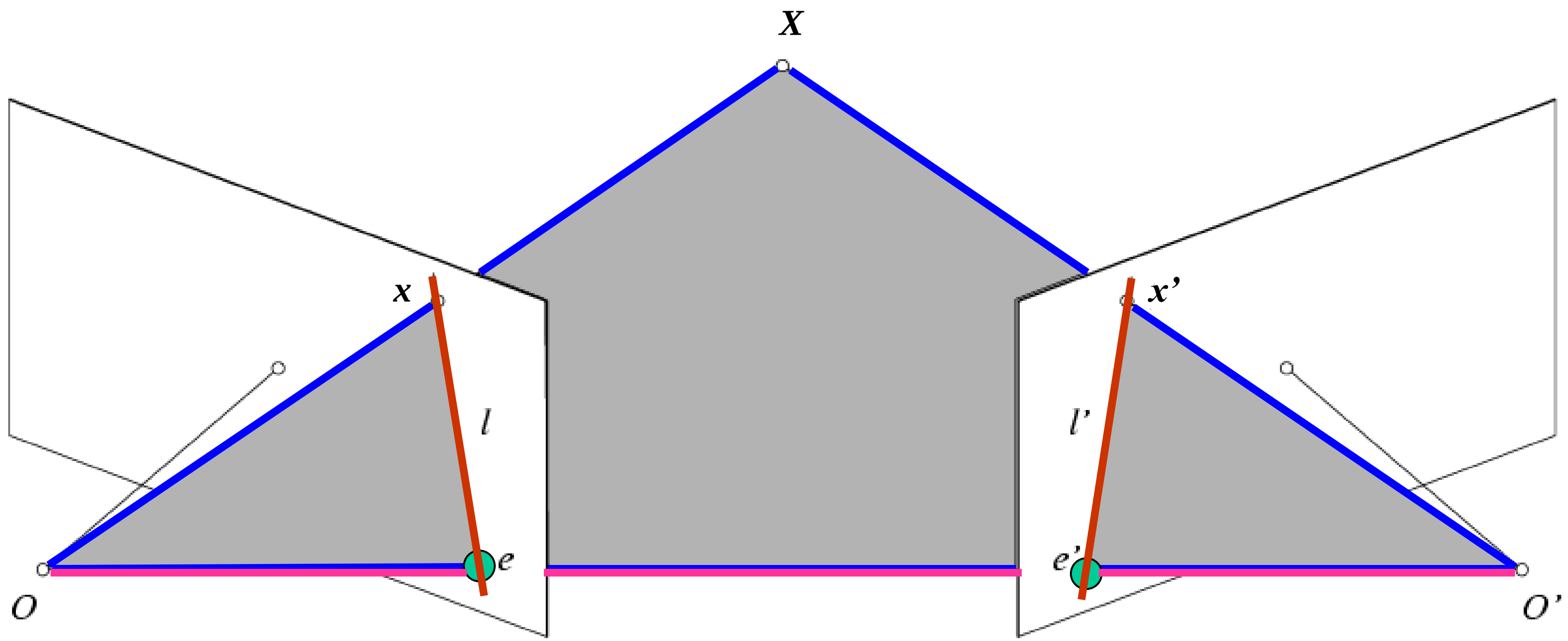


- **Baseline** – line connecting the two camera centers
- **Epipolar Plane** – plane containing baseline (1D family)
- **Epipoles**

= intersections of baseline with image planes

= projections of the other camera center

Epipolar geometry

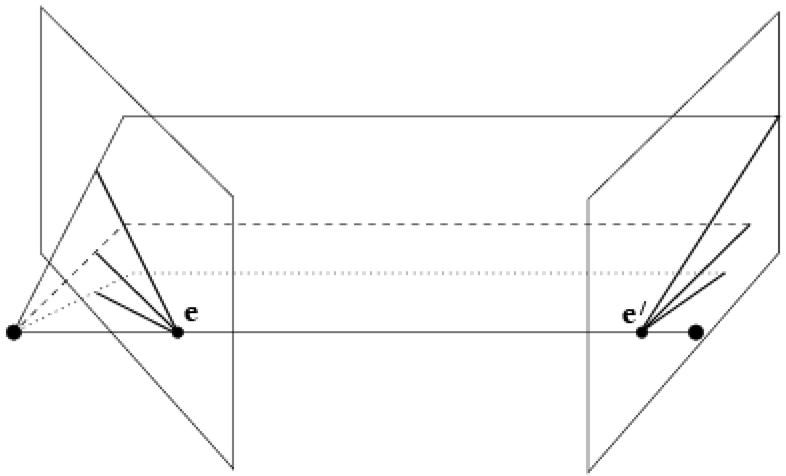


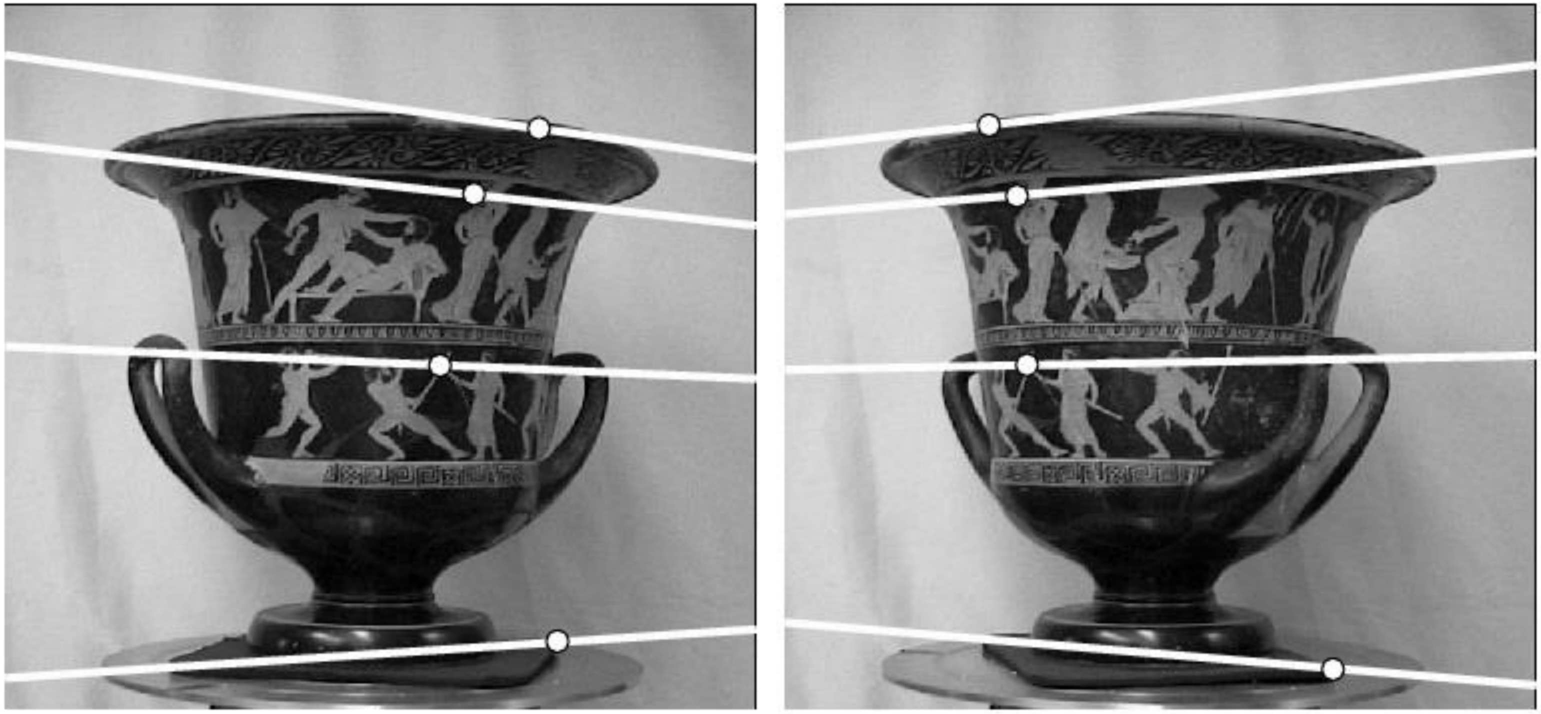
- **Baseline** – line connecting the two camera centers
- **Epipolar Plane** – plane containing baseline (1D family)
- **Epipoles**

= intersections of baseline with image planes
= projections of the other camera center

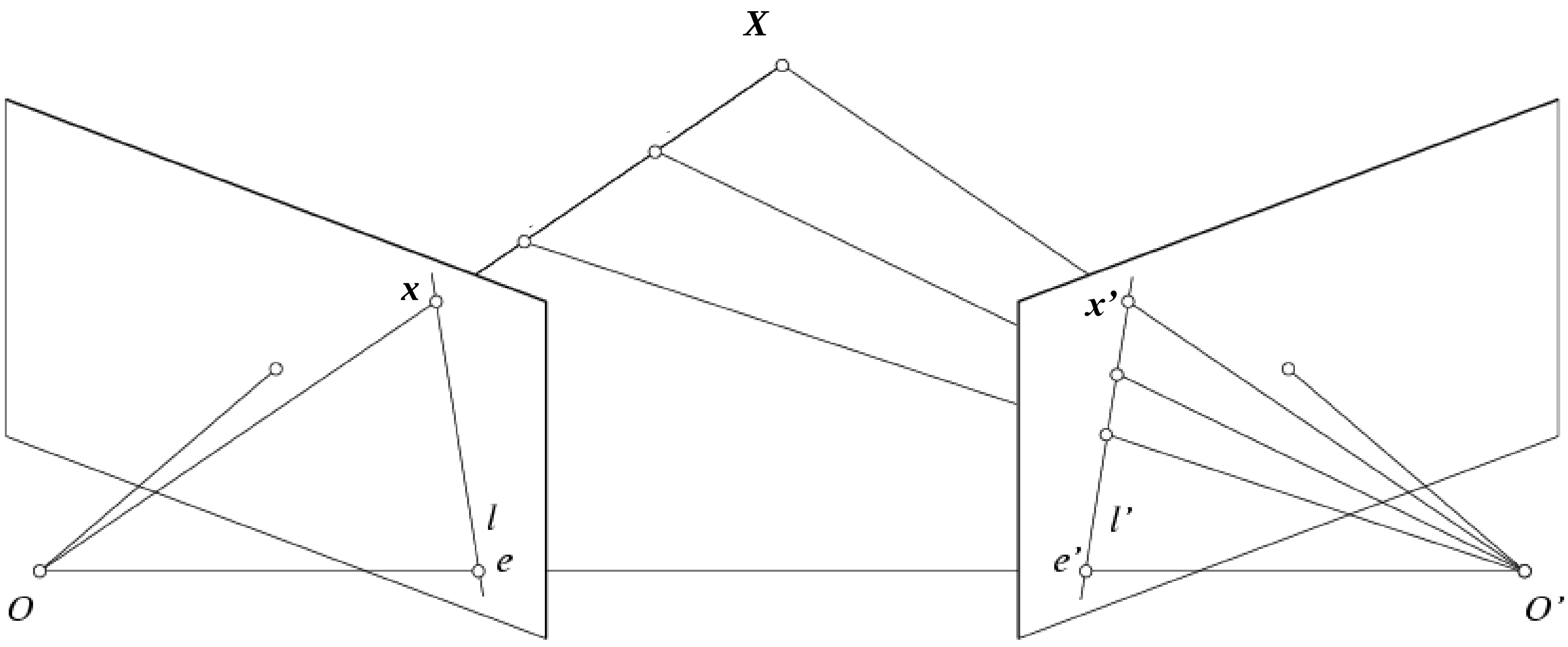
- **Epipolar Lines** - intersections of epipolar plane with image planes (always come in corresponding pairs)

Example: Converging cameras



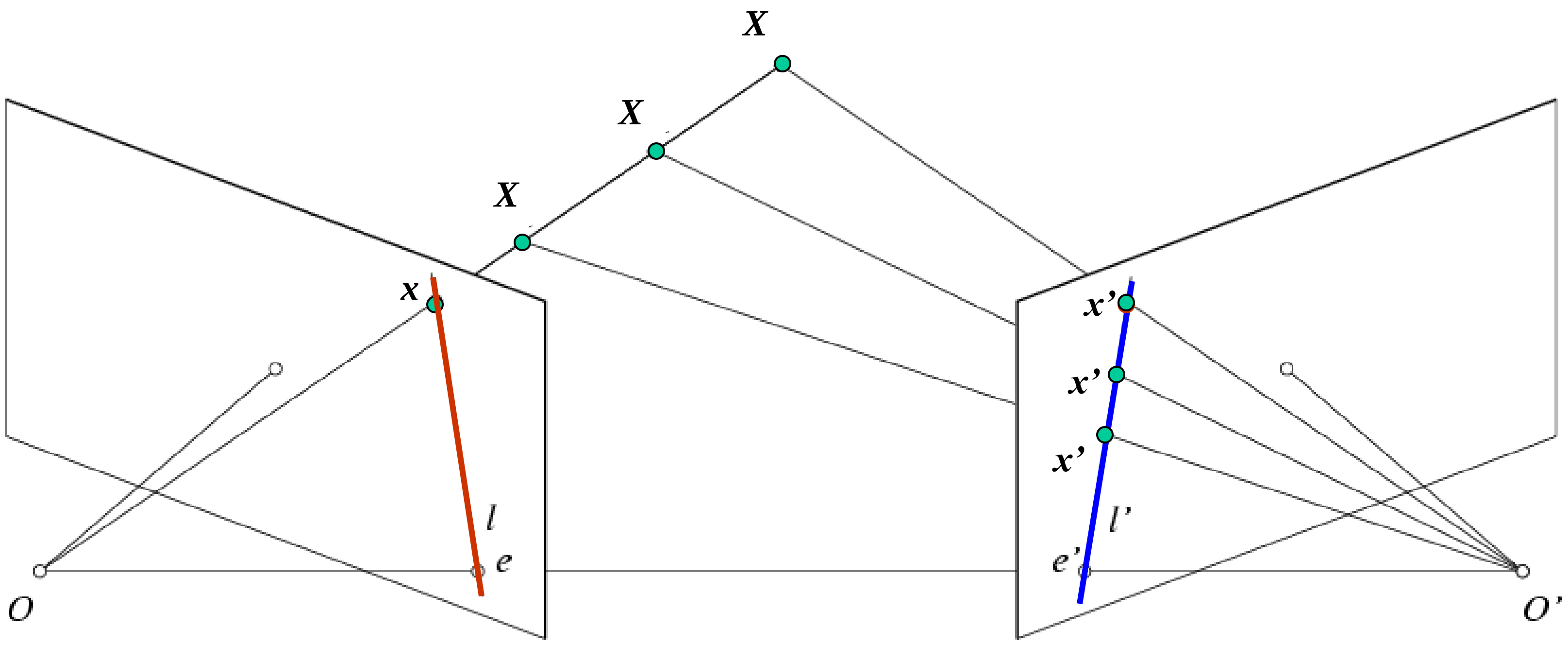


Epipolar constraint



- If we observe a point ***x*** in one image, where can the corresponding point ***x'*** be in the other image?

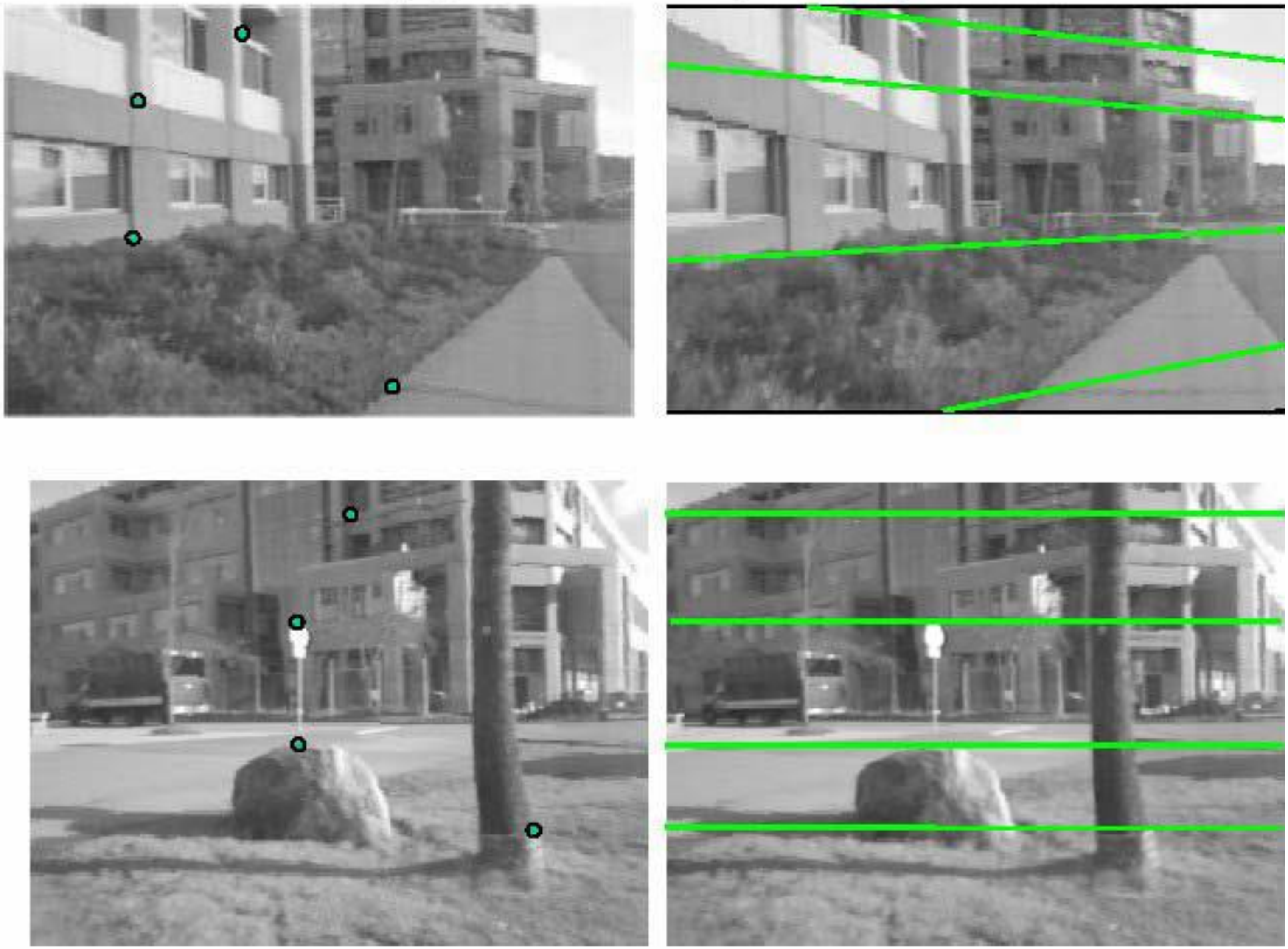
Epipolar constraint



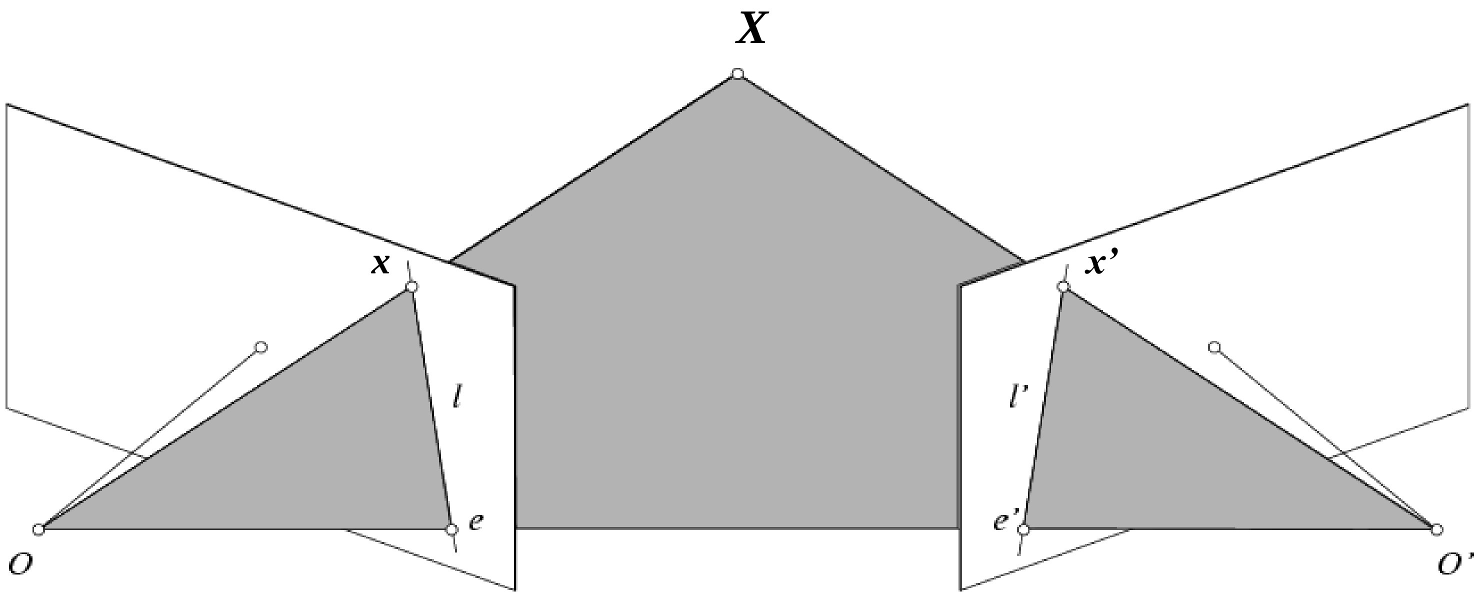
- Potential matches for ***x*** have to lie on the corresponding epipolar line ***l'***.

- Potential matches for ***x'*** have to lie on the corresponding epipolar line ***l***.

Epipolar constraint example



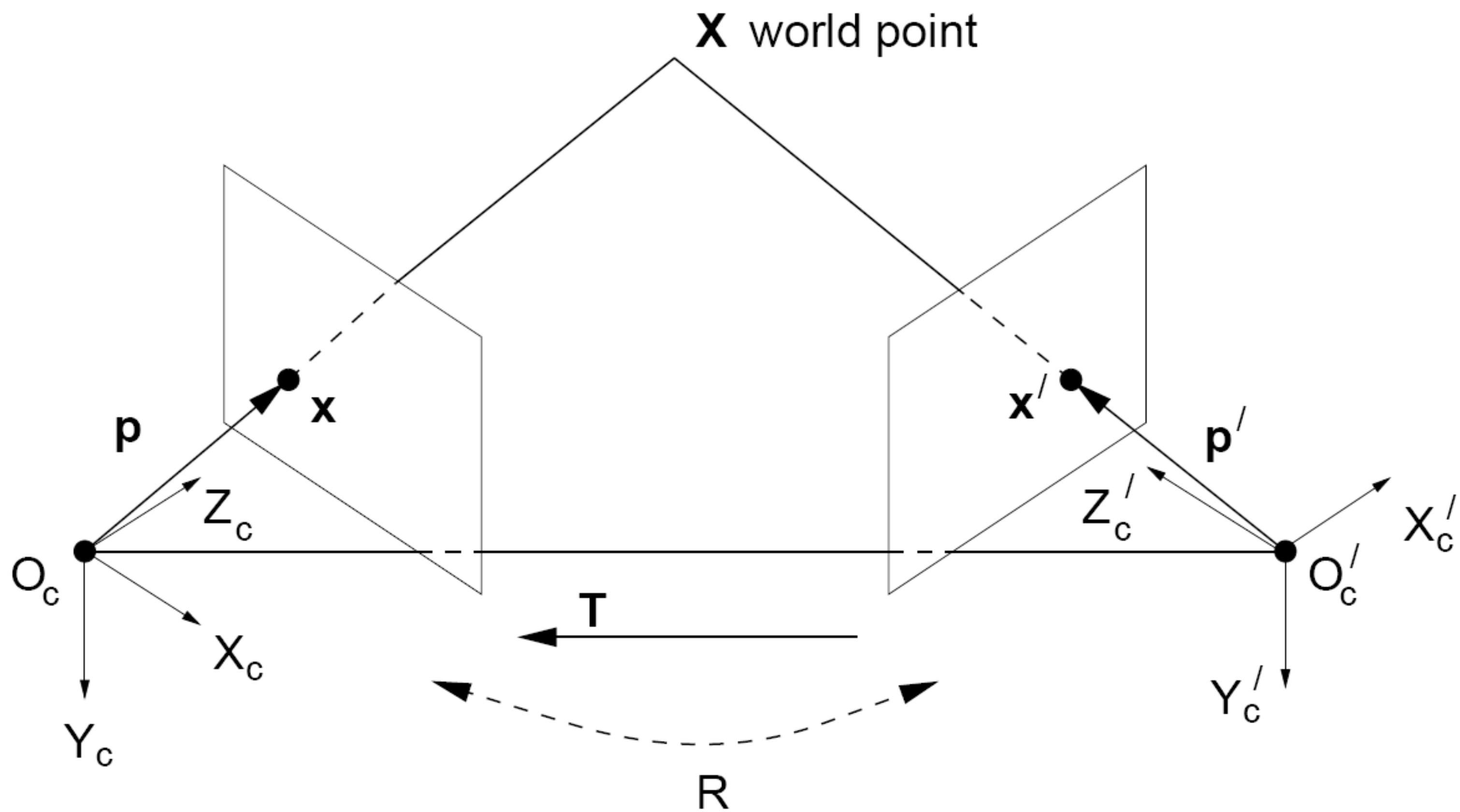
Epipolar constraint: Calibrated case



- 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 $\boldsymbol{K}[\boldsymbol{I} \mid \mathbf{0}]$ and $\boldsymbol{K}'[\boldsymbol{R} \mid \boldsymbol{t}]$
- We can multiply the projection matrices (and the image points) by the inverse of the calibration matrices to get *normalized* image coordinates:

$$\boldsymbol{x}_{\text{norm}} = \boldsymbol{K}^{-1}\boldsymbol{x}_{\text{pixel}} = [\boldsymbol{I} \;\; 0]\,\boldsymbol{X}, \qquad \boldsymbol{x}'_{\text{norm}} = \boldsymbol{K}'^{-1}\boldsymbol{x}'_{\text{pixel}} = [\boldsymbol{R} \;\; \boldsymbol{t}]\,\boldsymbol{X}$$

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

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

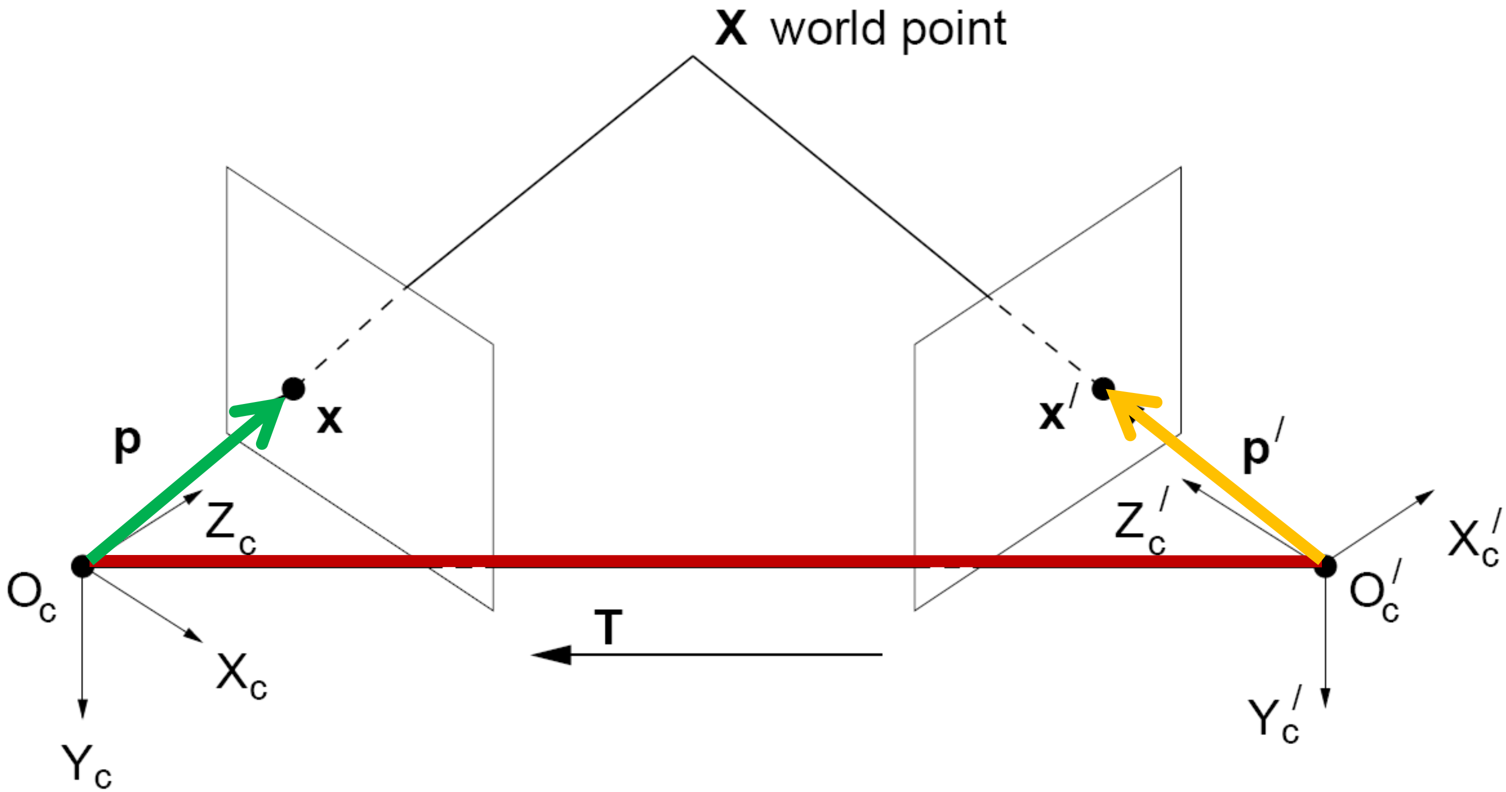
$$\vec{a} \cdot \vec{c} = 0$$

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

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



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

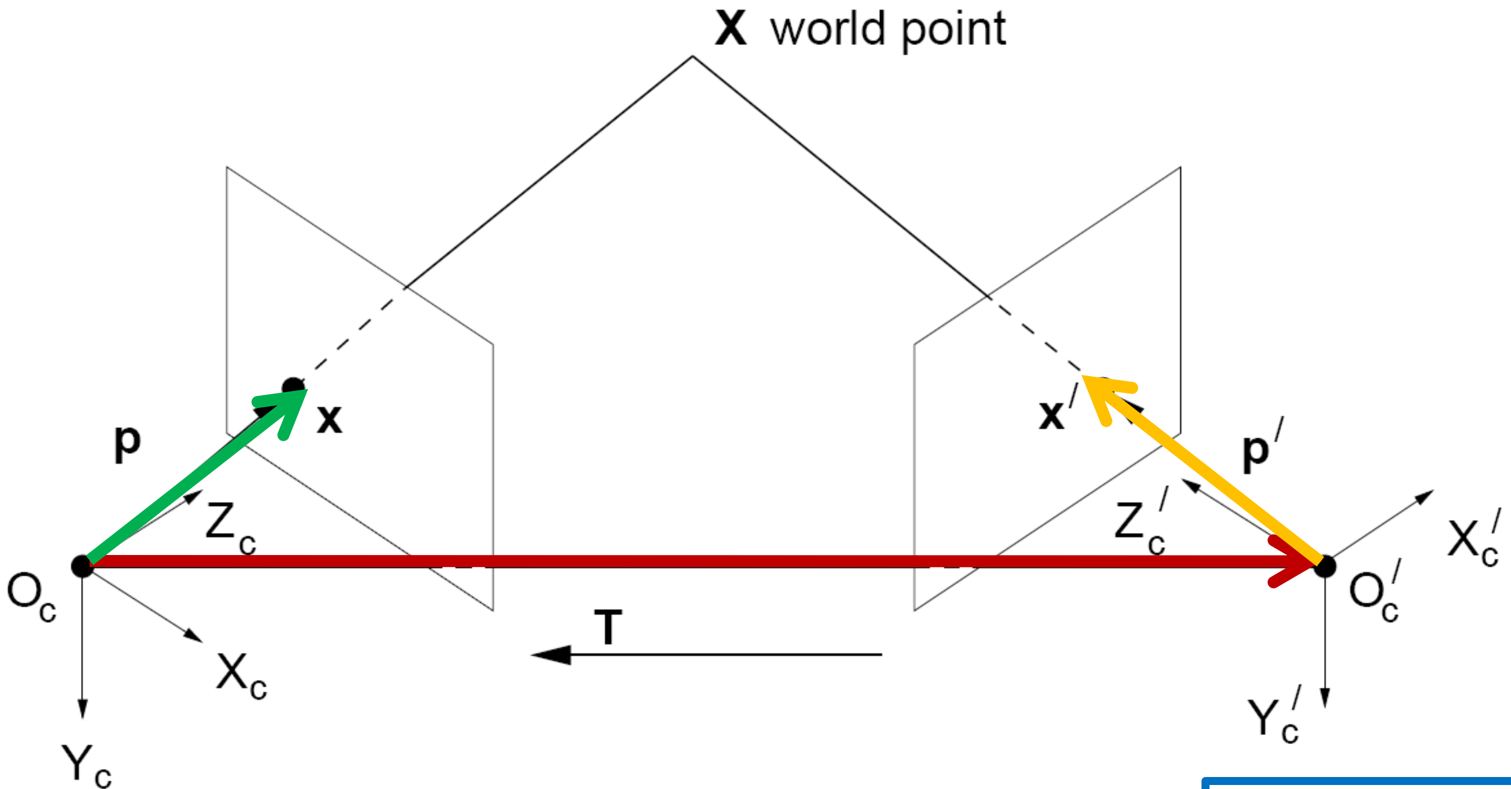
$$\underbrace{\mathbf{T} \times \mathbf{X'}}_{\text{Normal to the plane}} =$$

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

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

$$= 0$$

From geometry to algebra



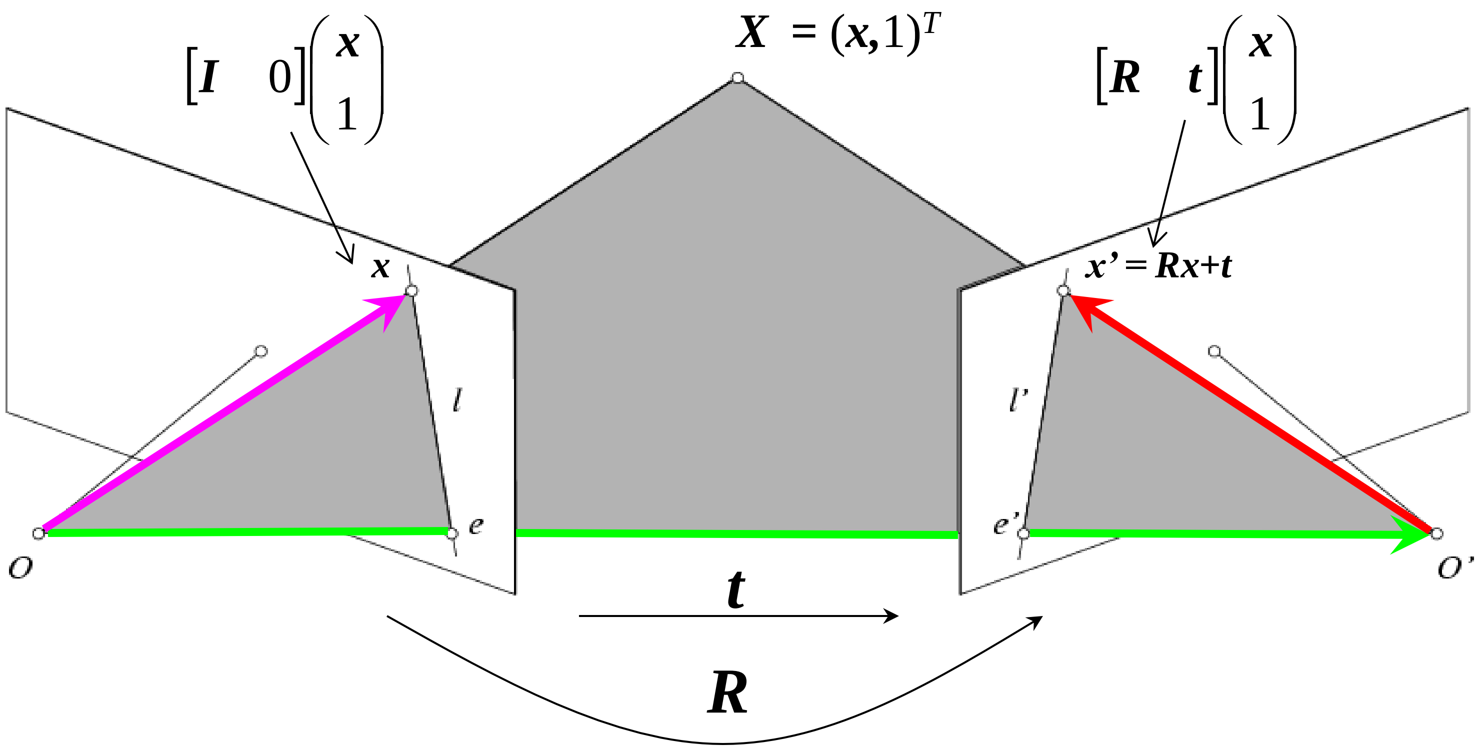
$$\mathbf{X'} = \mathbf{RX} + \mathbf{T}$$

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

$$= \mathbf{T} \times \mathbf{RX}$$

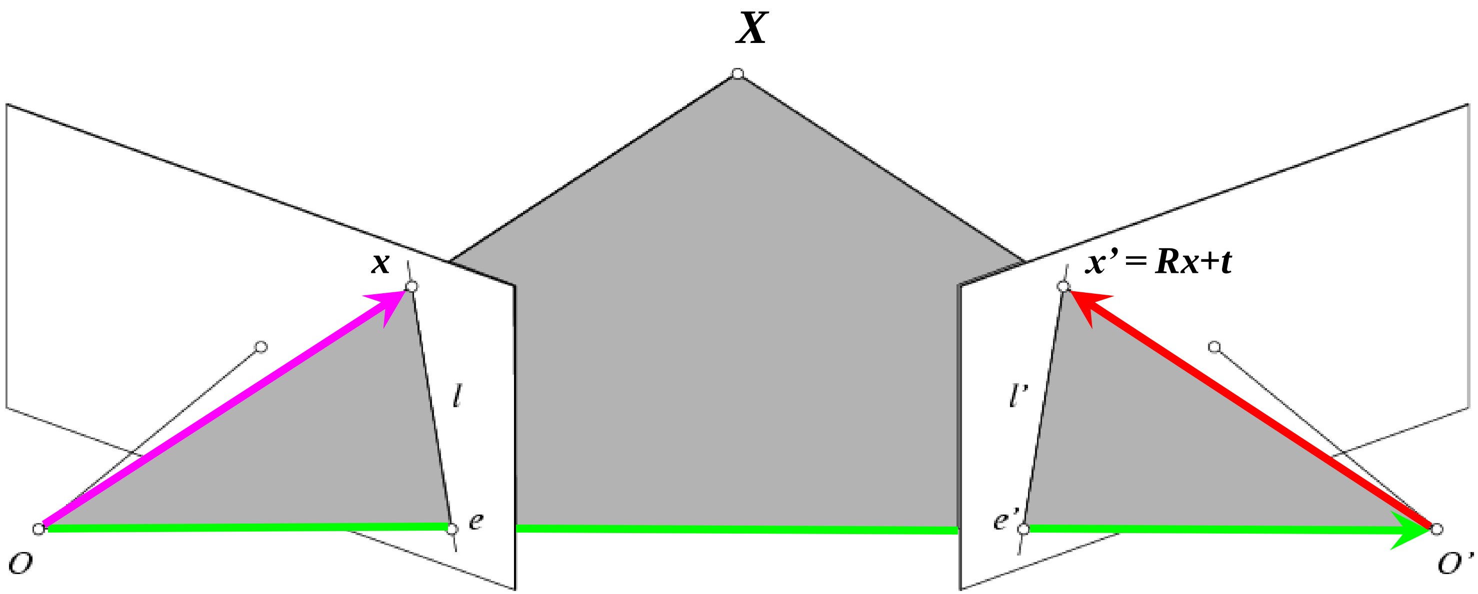
$$\mathbf{X'} \cdot (\mathbf{T} \times \mathbf{X'}) = \boxed{\mathbf{X'} \cdot (\mathbf{T} \times \mathbf{RX}) = 0}$$

Epipolar constraint: Calibrated case



The vectors $\color{magenta}{Rx}$, $\color{lime}{t}$, and $\color{red}{x'}$ are coplanar

Epipolar constraint: Calibrated case

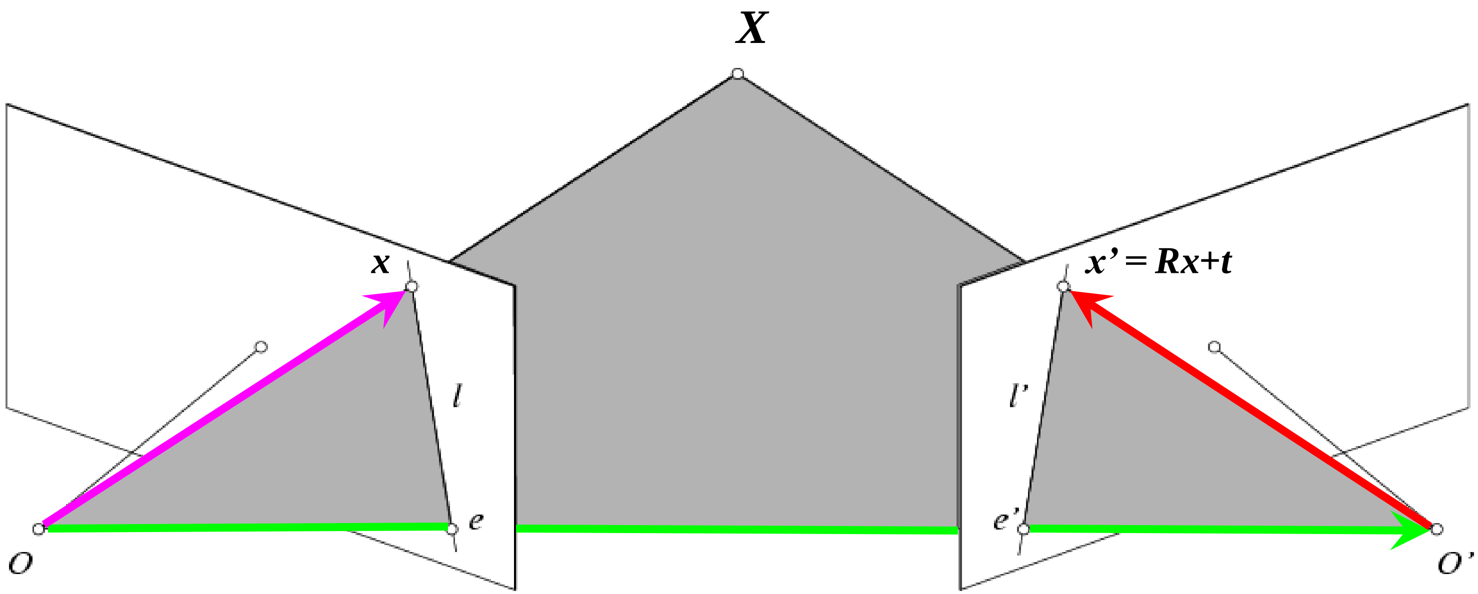


$$\boldsymbol{x}' \cdot [\boldsymbol{t} \times (\boldsymbol{R}\boldsymbol{x})] = 0 \quad \Rightarrow \quad \boldsymbol{x}'^T [\boldsymbol{t}_\times] \boldsymbol{R}\boldsymbol{x} = 0$$

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 ***Rx***, ***t***, and ***x'*** are coplanar

Epipolar constraint: Calibrated case

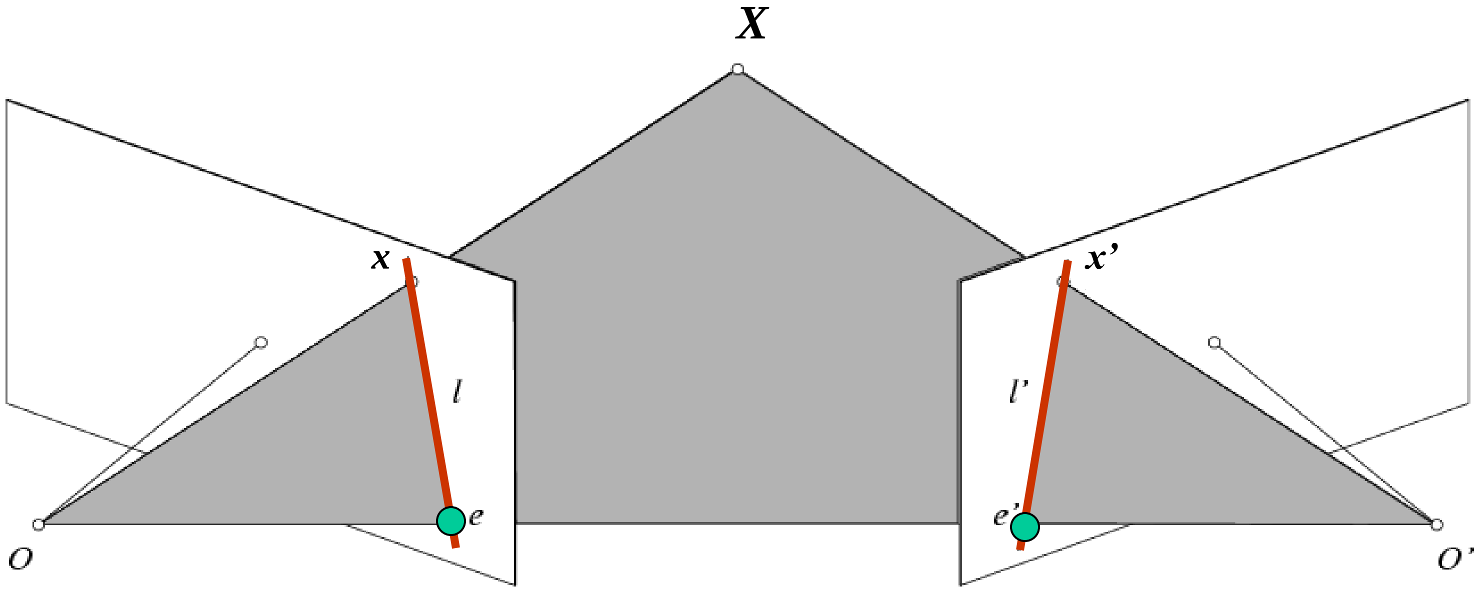


$$\boldsymbol{x}' \cdot [\boldsymbol{t} \times (\boldsymbol{R}\boldsymbol{x})] = 0 \qquad \boldsymbol{x}'^T [\boldsymbol{t}_\times] \boldsymbol{R}\boldsymbol{x} = 0 \qquad \boldsymbol{x}'^T E \boldsymbol{x} = 0$$

Essential Matrix
(Longuet-Higgins, 1981)

The vectors ***Rx***, ***t***, and ***x’*** are coplanar

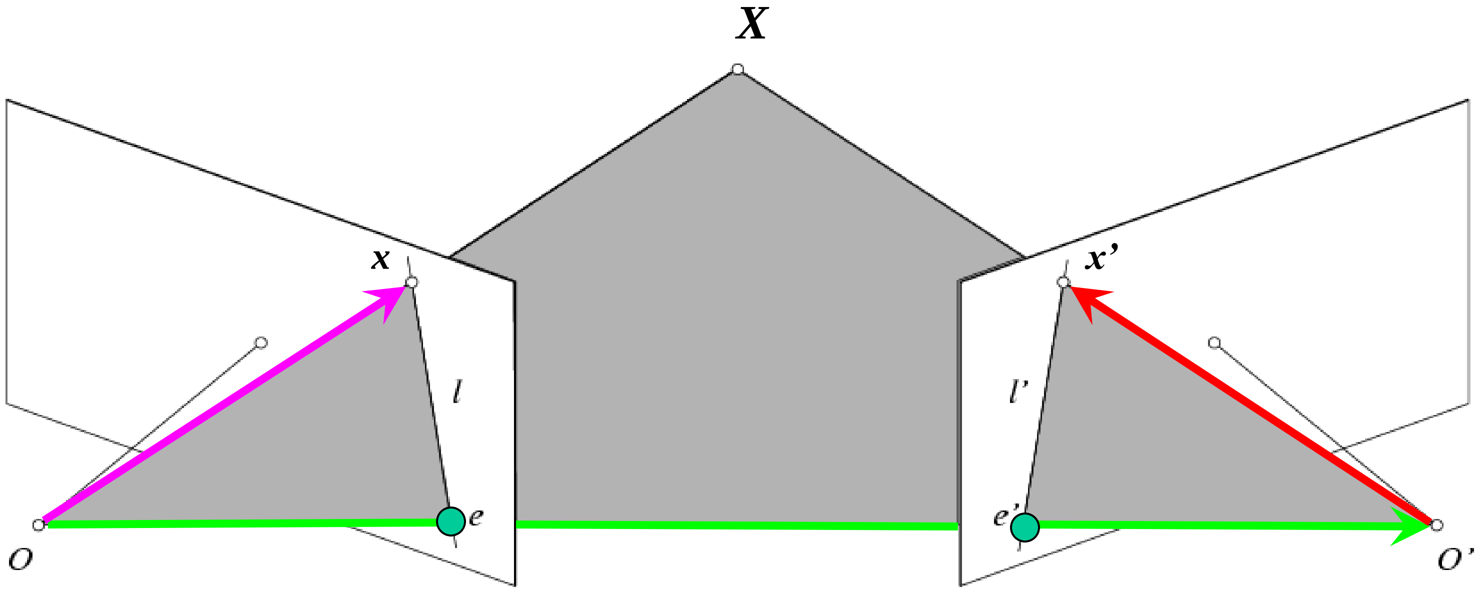
Epipolar constraint: Calibrated case



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

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

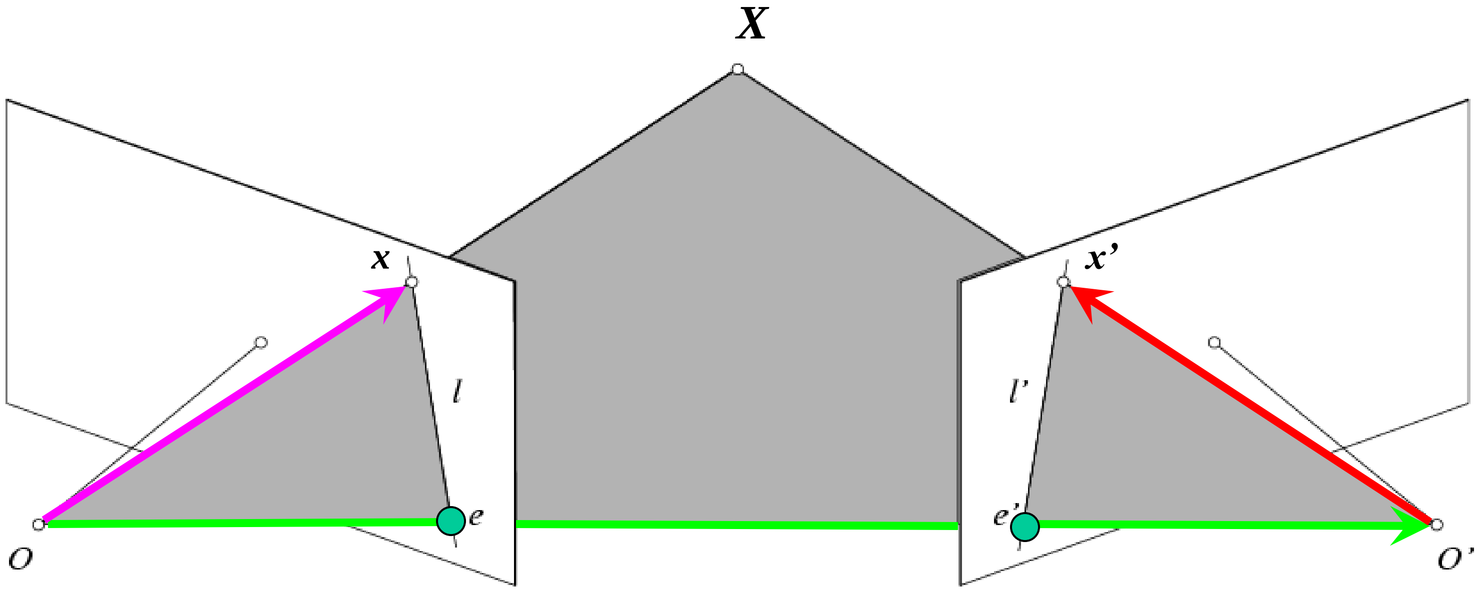
Epipolar constraint: Uncalibrated case



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

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

Epipolar constraint: Uncalibrated case



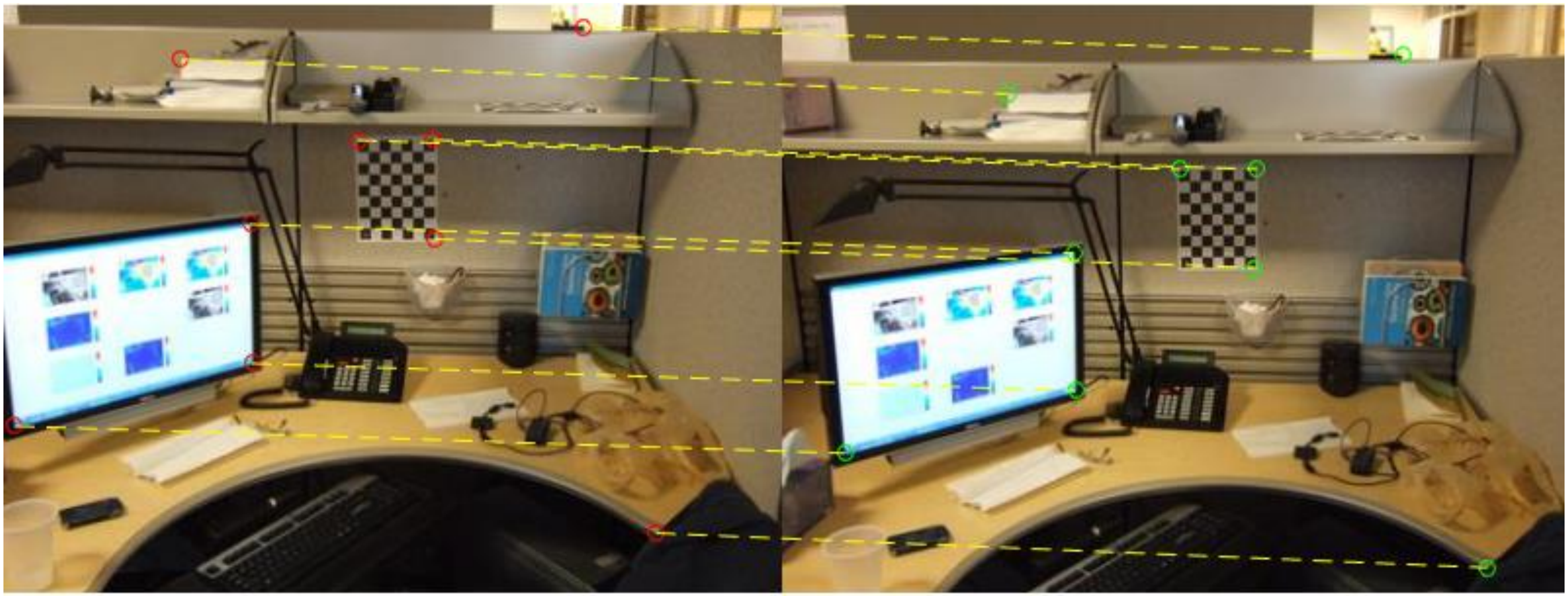
$$\hat{\mathbf{x}}'^T \boldsymbol{E}\, \hat{\mathbf{x}} = 0 \quad \Rightarrow \quad \boldsymbol{x}'^T \boldsymbol{F}\, \boldsymbol{x} = 0 \quad \text{with} \quad \boldsymbol{F} = \boldsymbol{K}'^{-T} \boldsymbol{E}\, \boldsymbol{K}^{-1}$$

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

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

Fundamental Matrix
(Faugeras and Luong, 1992)

Estimating the fundamental matrix



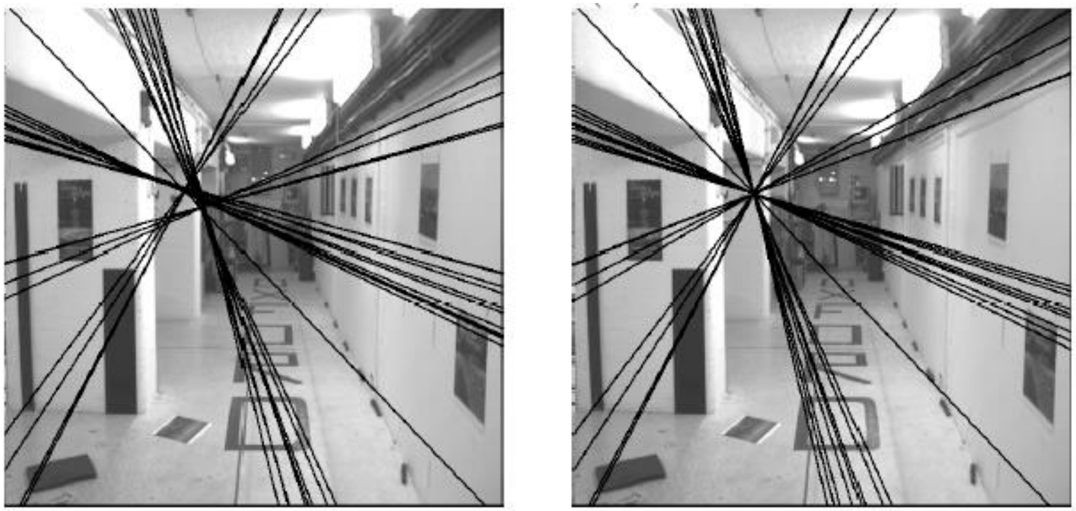
The eight-point algorithm

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

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

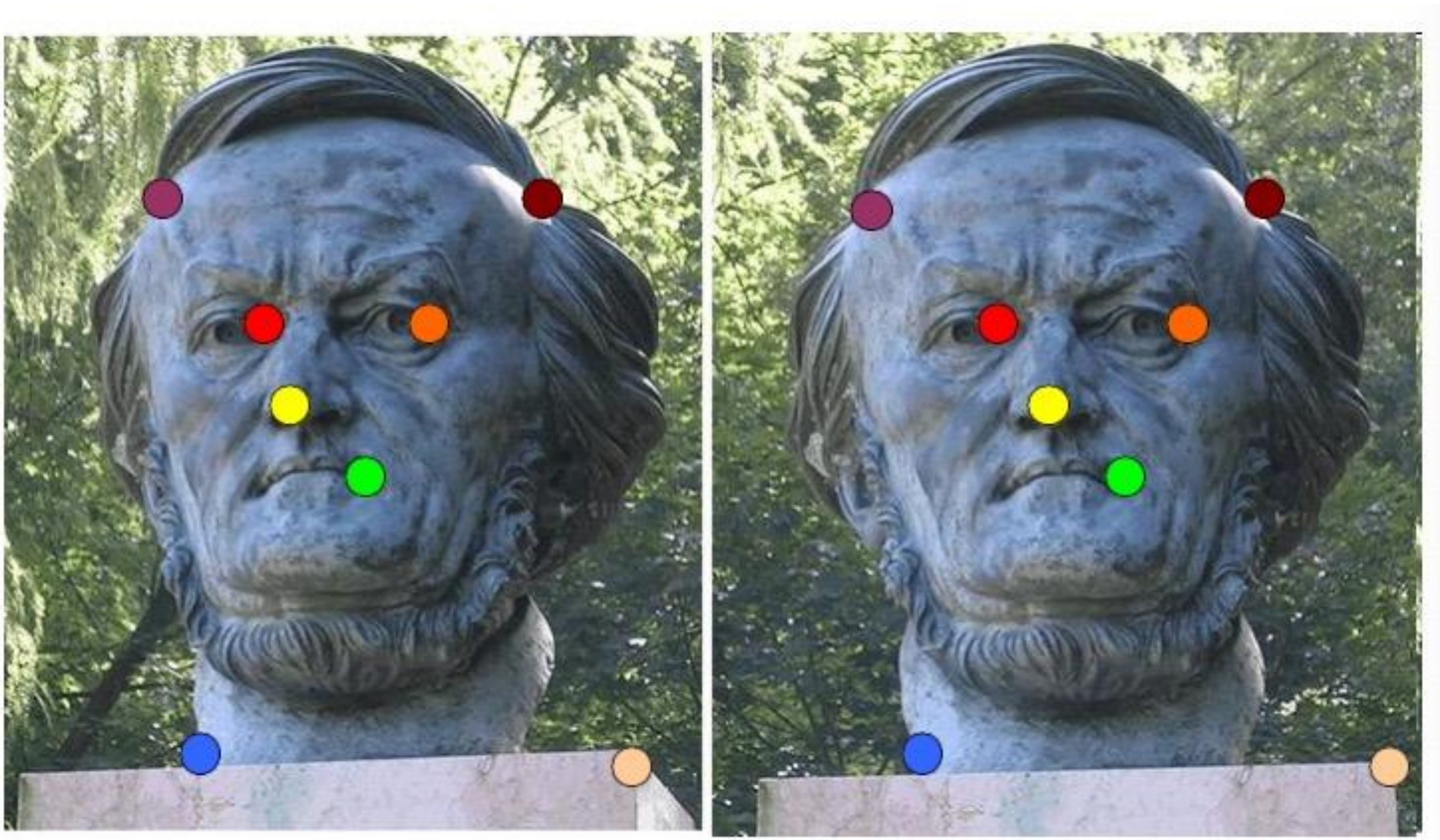
Solve homogeneous linear system using eight or more matches

Enforce rank-2 constraint (take SVD of ***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

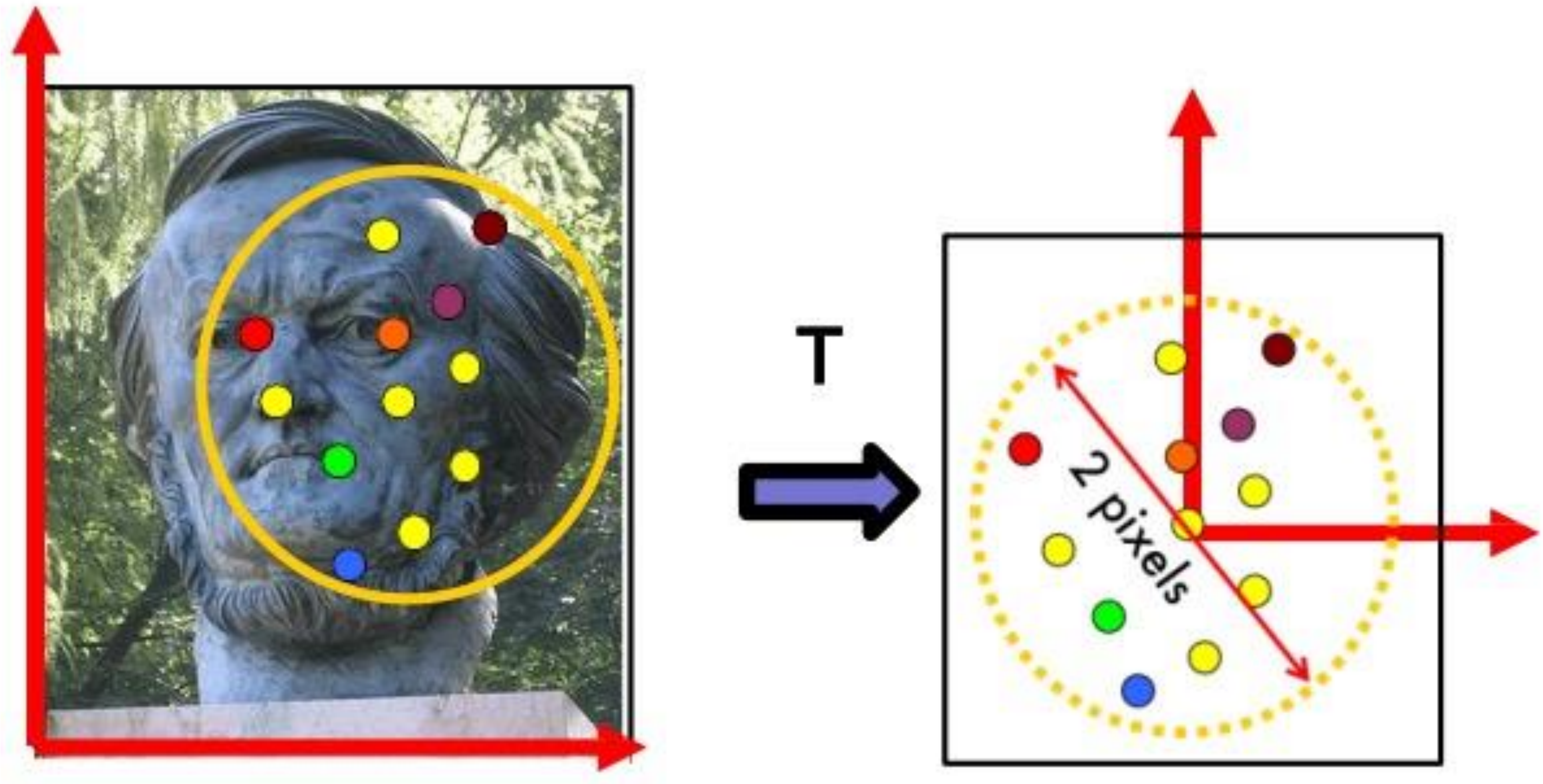
The normalized eight-point algorithm

(Hartley, 1995)

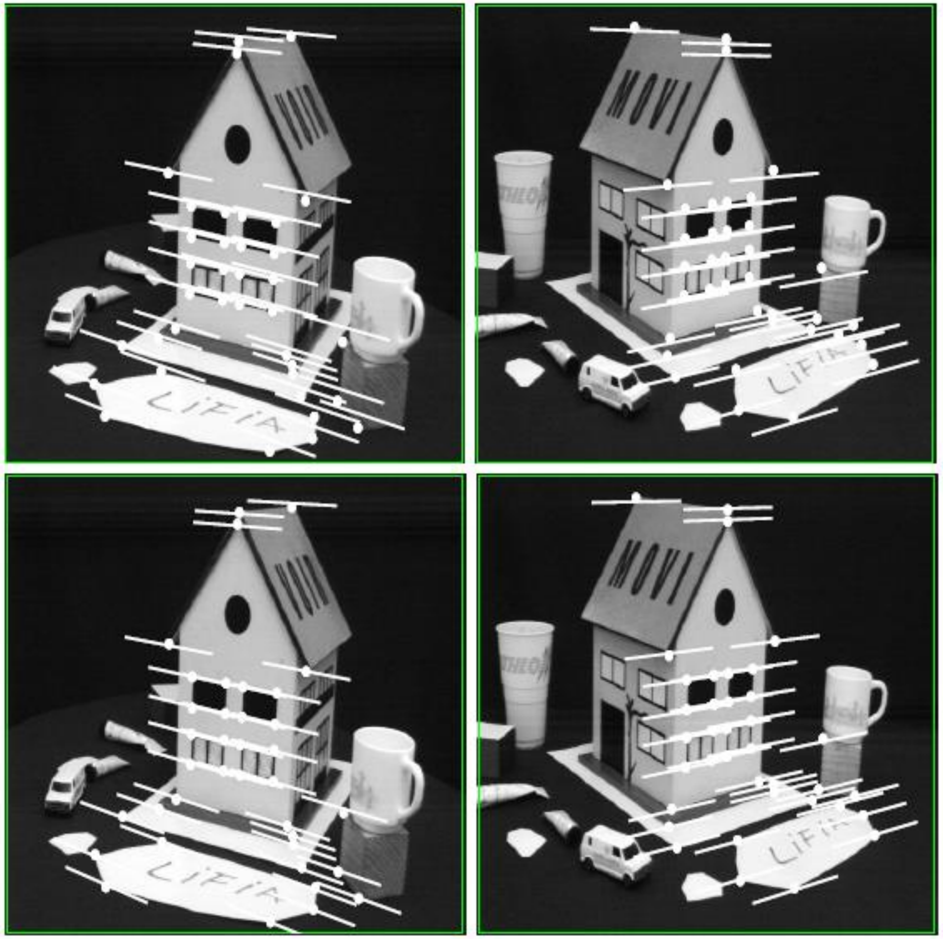
- Center the image data at the origin, and scale it so the mean squared distance between the origin and the data points is 2 pixels
- Use the eight-point algorithm to compute $\boldsymbol{F}$ from the normalized points
- Enforce the rank-2 constraint (for example, take SVD of $\boldsymbol{F}$ and throw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

The normalized eight-point algorithm

(Hartley, 1995)



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

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: $\boldsymbol{E} = \boldsymbol{K}'^{T}\boldsymbol{F}\boldsymbol{K}$
- The essential matrix gives us the relative rotation and translation between the cameras, or their extrinsic parameters