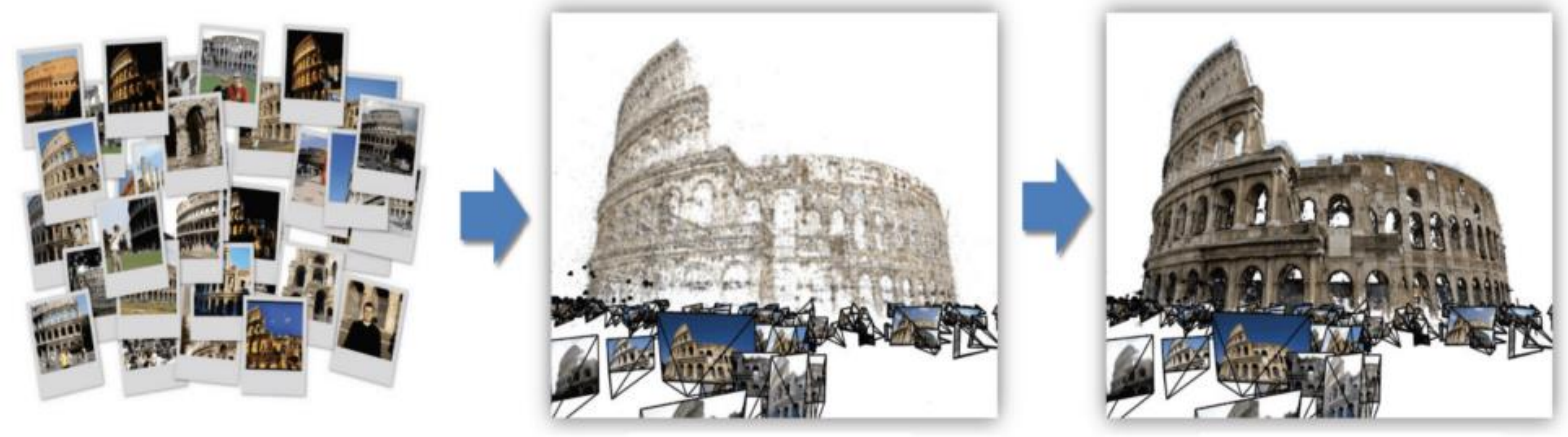
Multi-View Geometry

Jiayuan Gu

gujy1@shanghaitech.edu.cn

Many slides modified from Prof. Derek Hoiem CS543

Building Rome in a Day

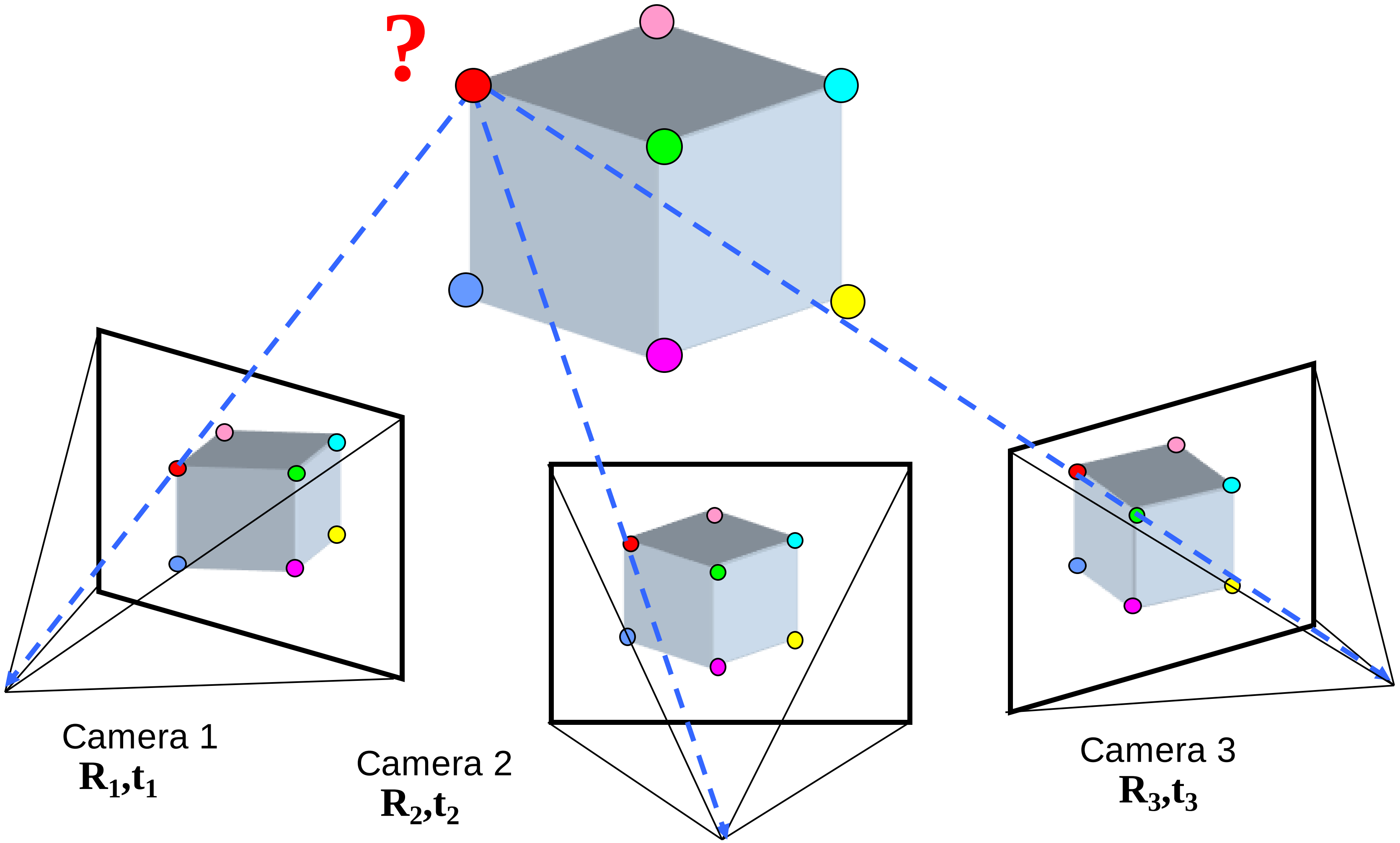


https://grail.cs.washington.edu/rome/

Figure from blog

Multi-view Geometry Problems

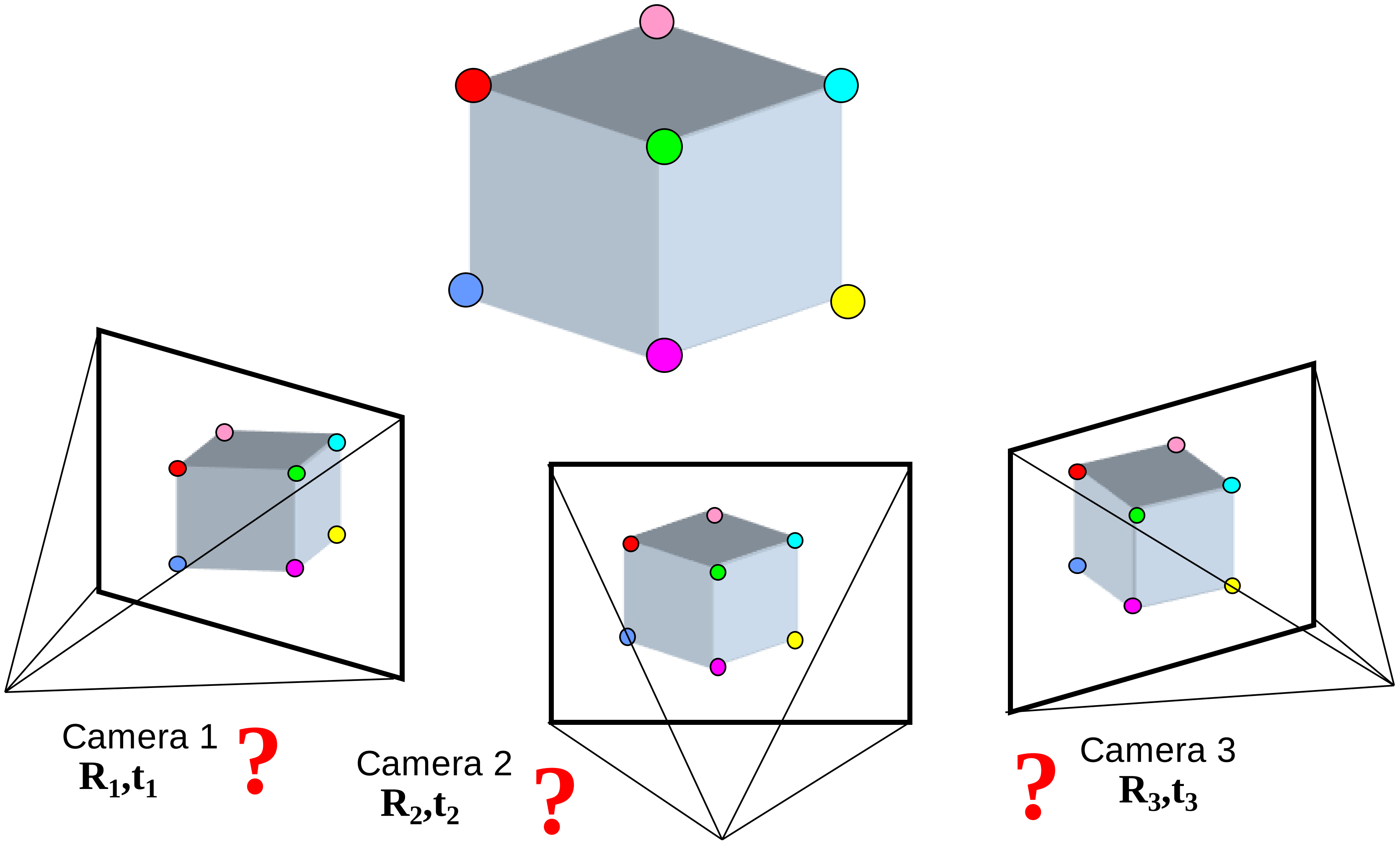
- **Structure:** Given projections of the same 3D point in two or more images, compute the 3D coordinates of that point



Slide credit: Noah Snavely

Multi-view Geometry Problems

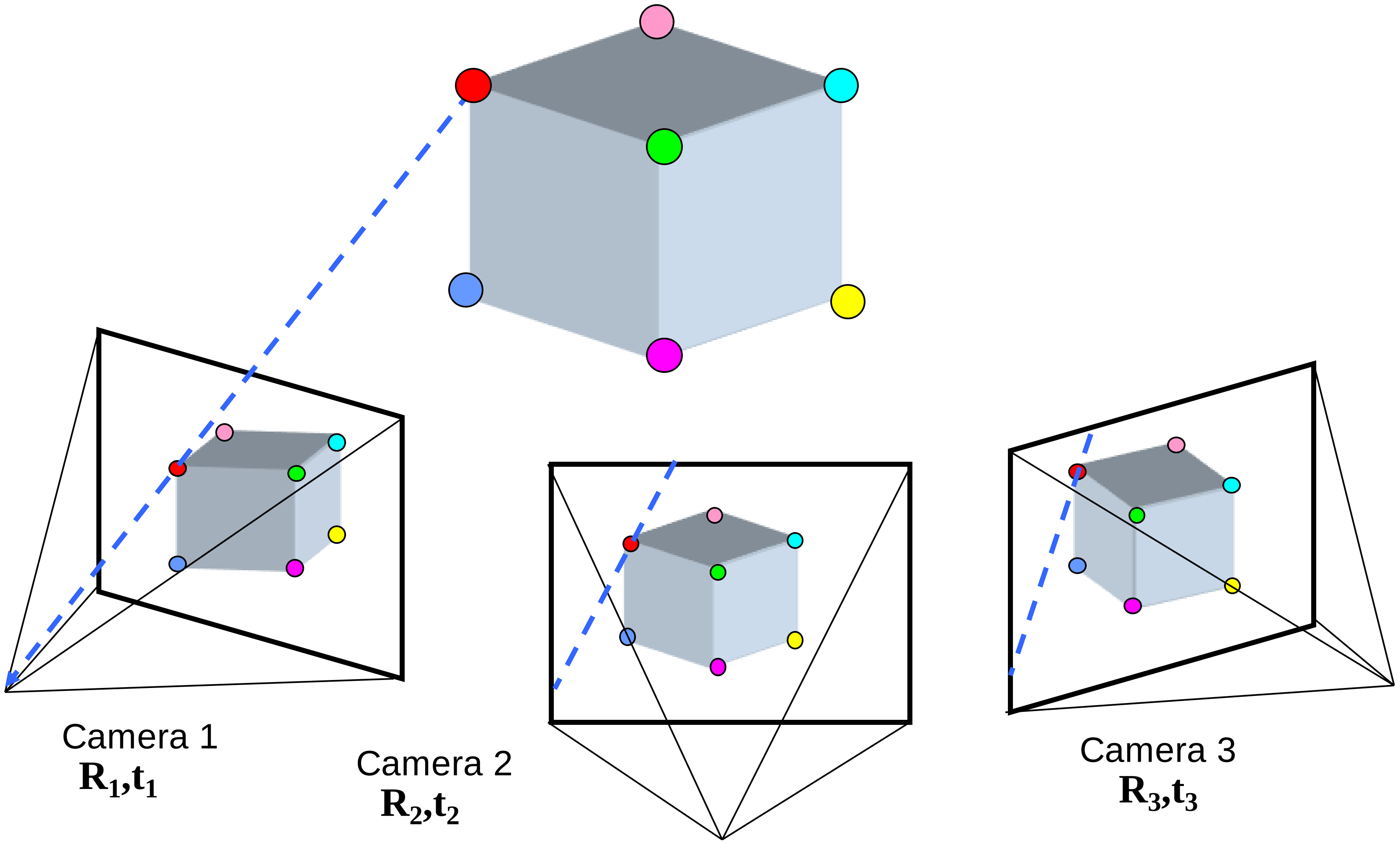
- **Motion:** Given a set of corresponding points in two or more images, compute the camera parameters



Slide credit: Noah Snavely

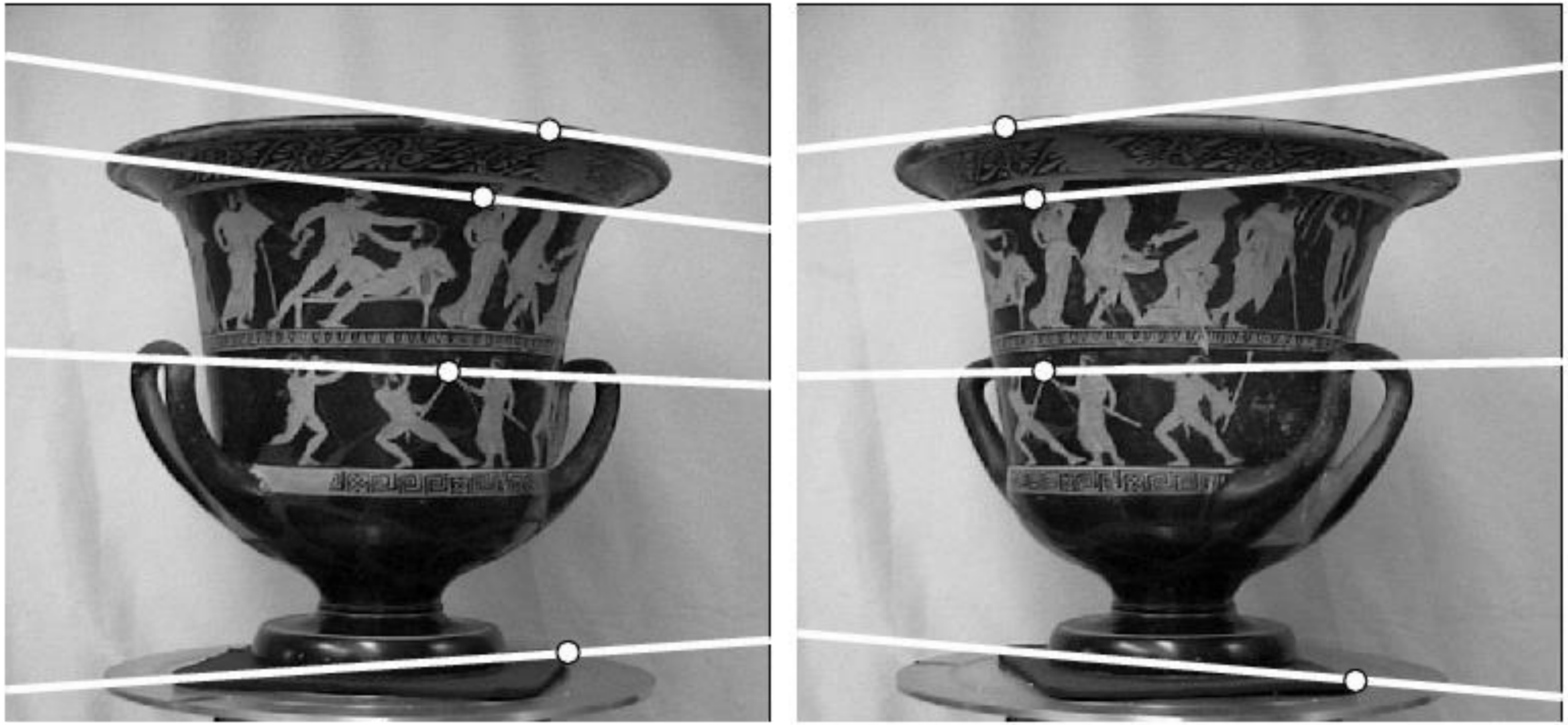
Multi-view Geometry Problems

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



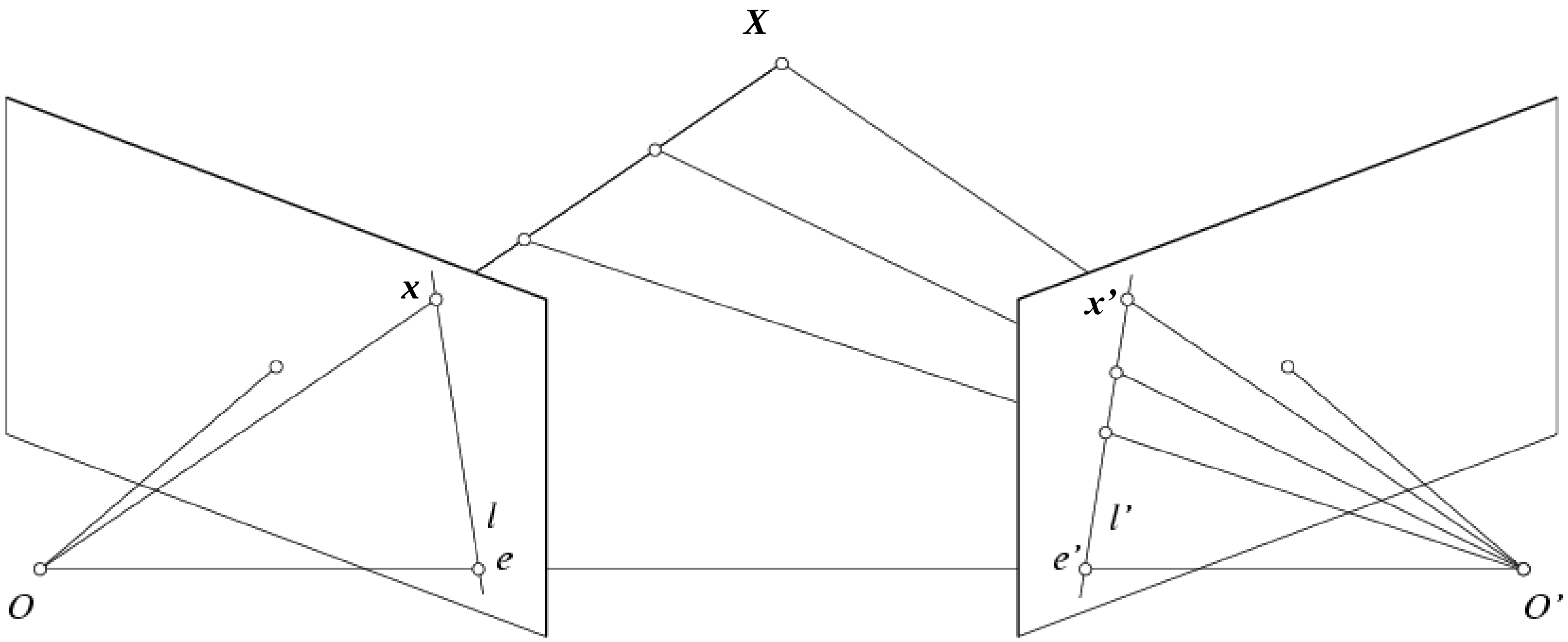
Slide credit: Noah Snavely

Two-View Geometry



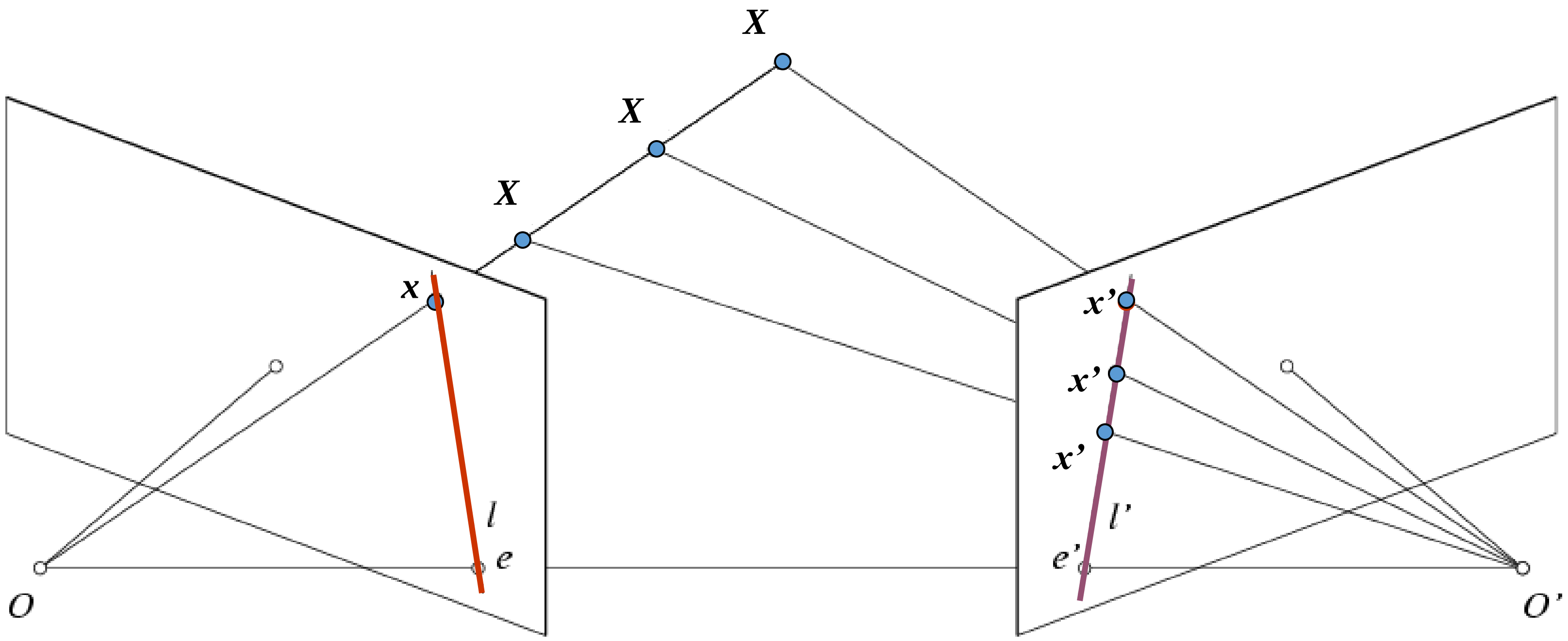
Example of stereo correspondence

Epipolar Constraint



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

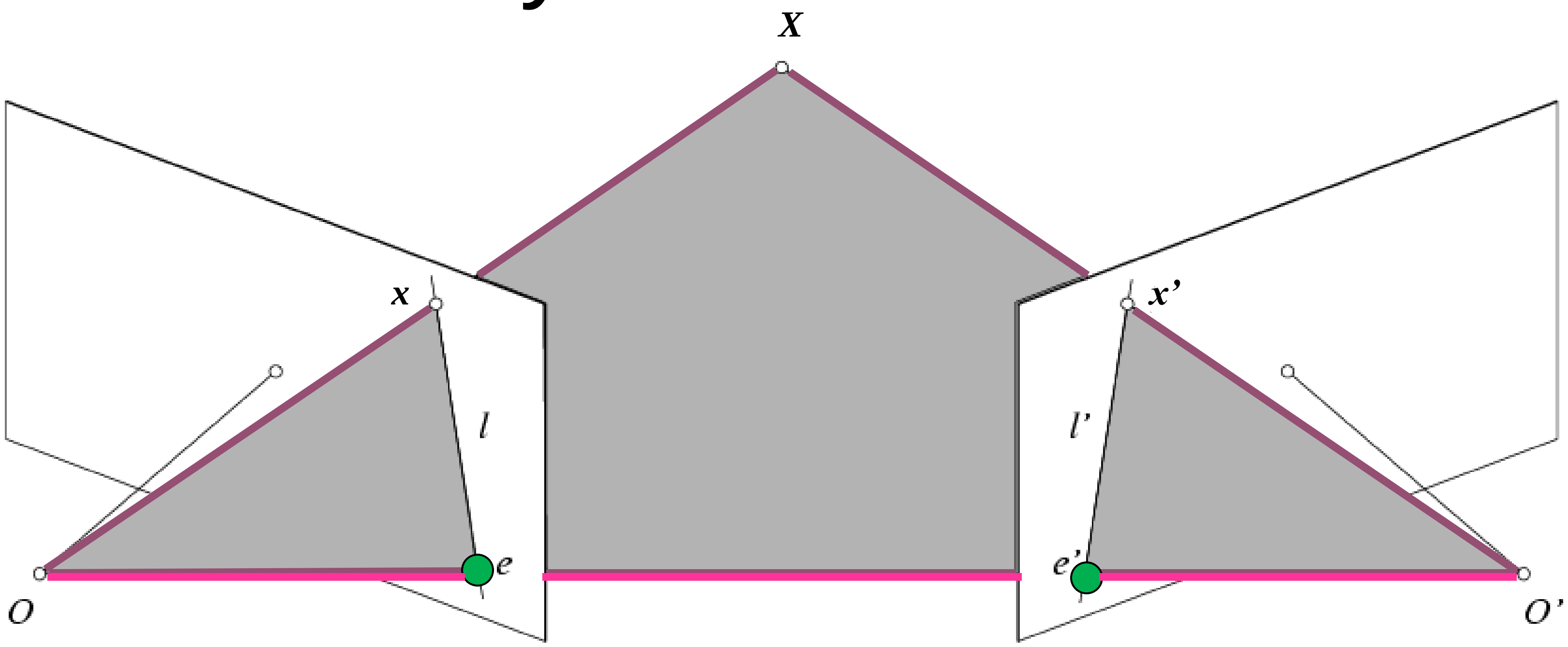
Epipolar Constraint



Potential matches for ***x*** must lie on the corresponding line ***l'***.

Potential matches for ***x'*** must lie on the corresponding line ***l***.

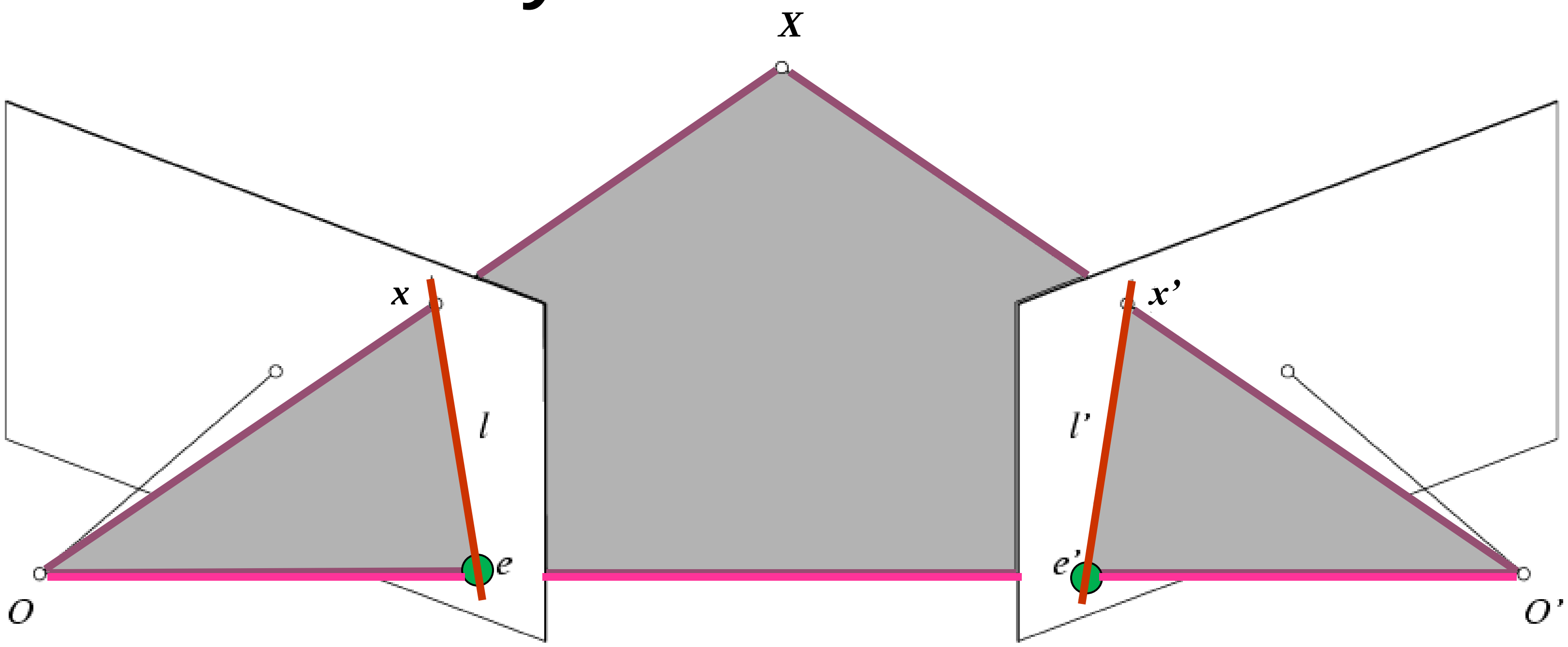
Epipolar Geometry



- **Baseline** – line connecting the two camera centers
- **Epipolar Plane** – plane containing baseline
- **Epipoles**

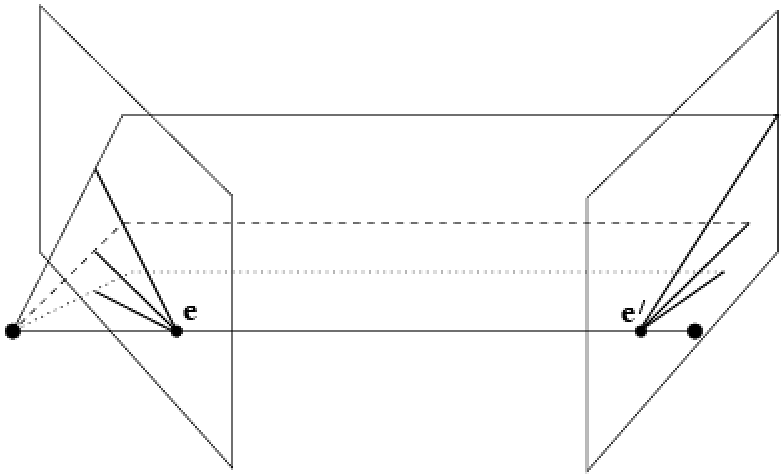
= intersections of baseline with image planes
= projections of the other camera center

Epipolar Geometry

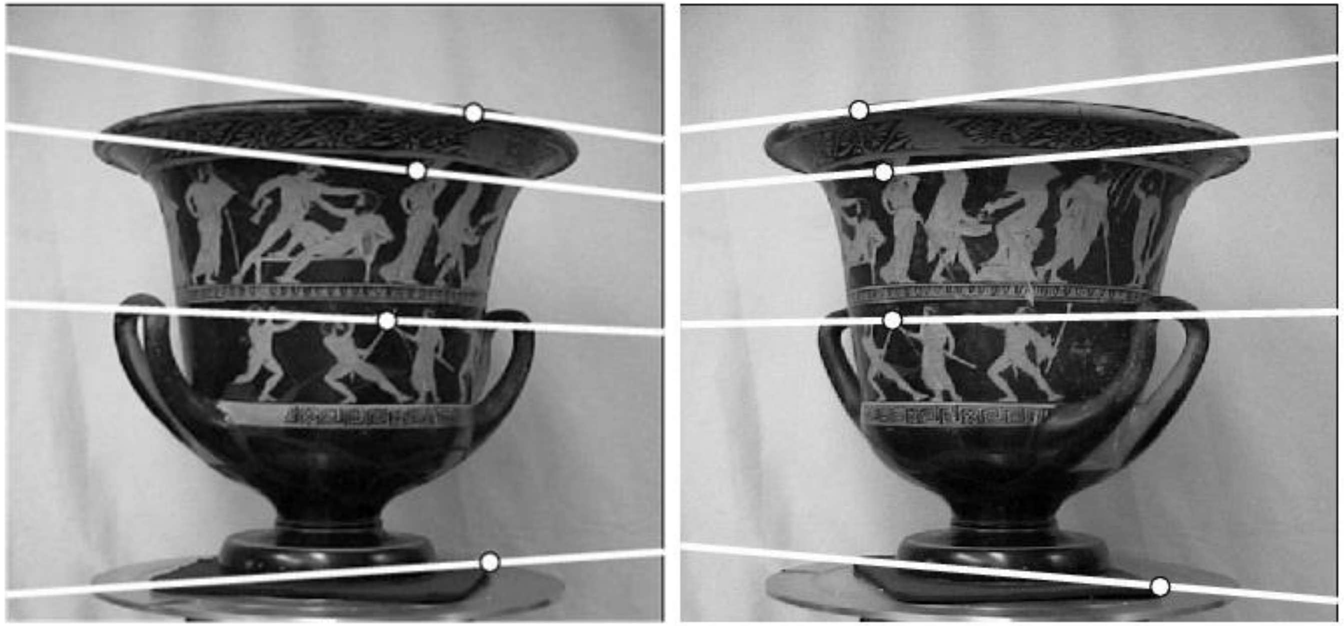


- **Baseline** – line connecting the two camera centers
- **Epipolar Plane** – plane containing baseline
- **Epipoles**
  = intersections of baseline with image planes
  = projections of the other camera center
- **Epipolar Lines** - intersections of epipolar plane with image planes (always come in corresponding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

Example Configuration of Two Image Planes

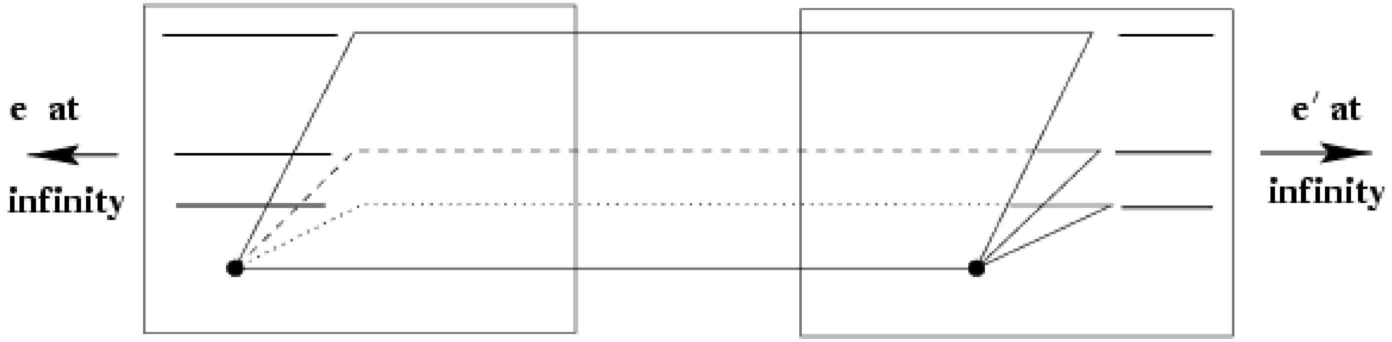


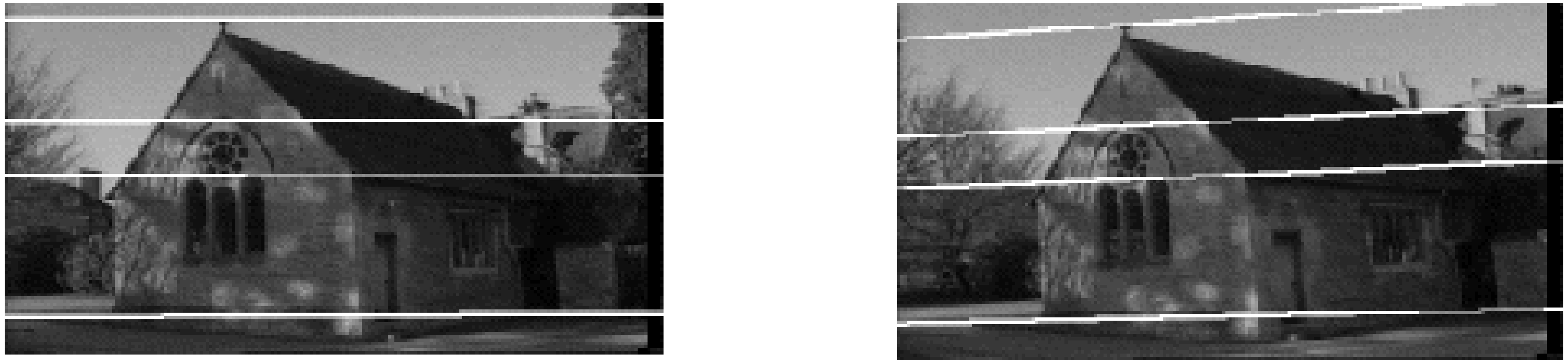
Converged cameras: *The two cameras are set apart by the inter-axial distance then converged, or angled towards each other (sometimes called a "toe-in").*



Epipolar lines must pass the epipole (finite point for converged cameras)

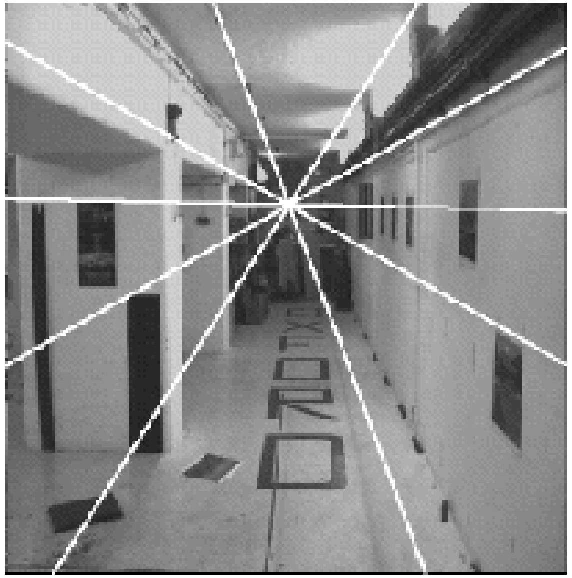
Example Configuration of Two Image Planes

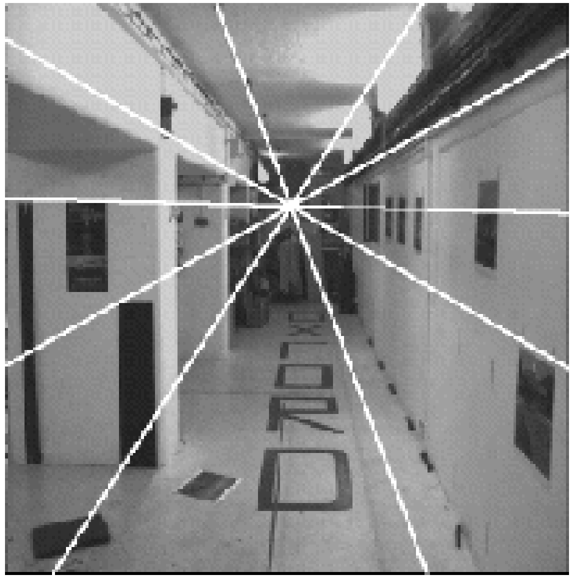


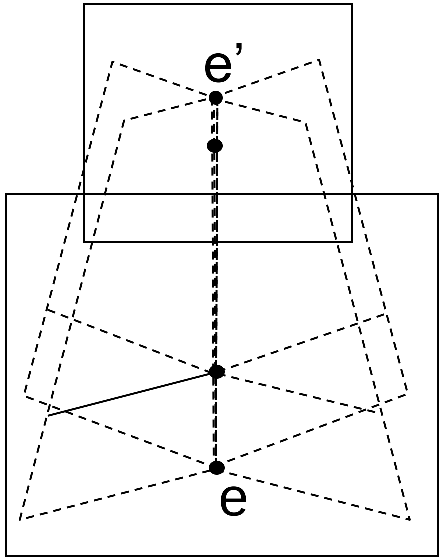


Parallel camera (epipole at infinity)

Example Configuration of Two Image Planes





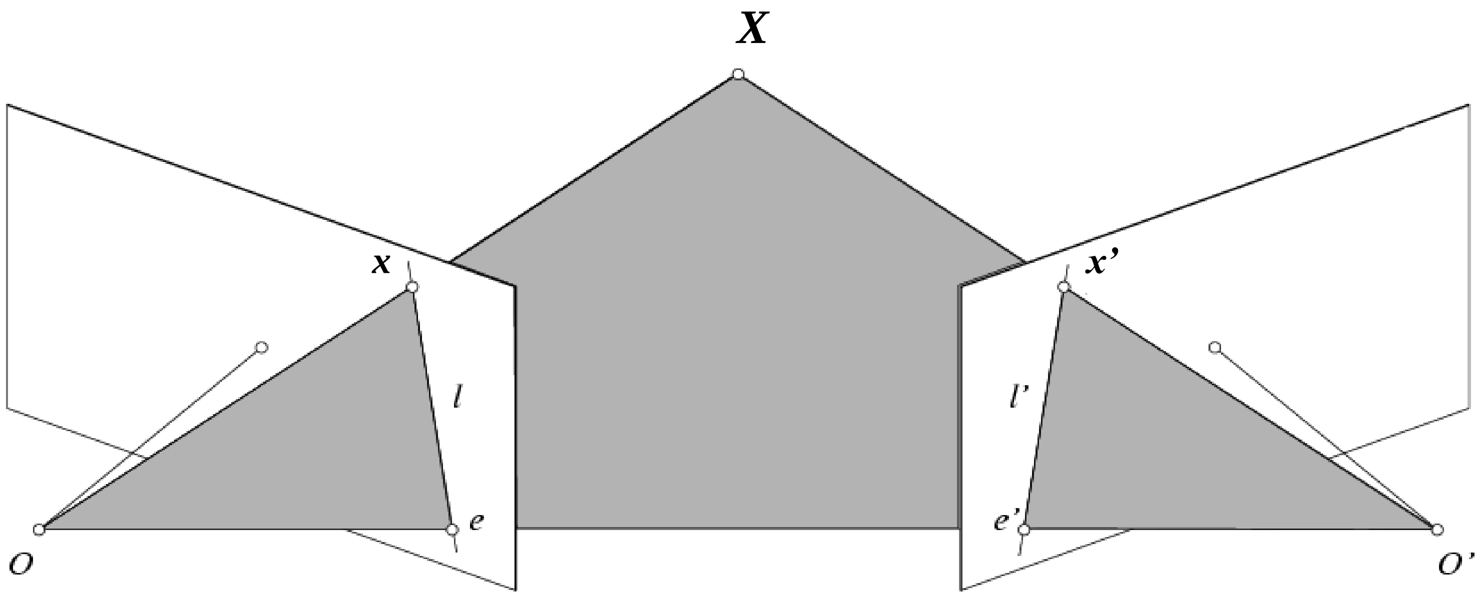


The second camera is moving along the principal axis of the first one

Epipole has same coordinates in both images.
Points move along lines radiating from e: “Focus of expansion”

Next, we will mathematically describe the epipolar constraint

Epipolar Constraint



Given the intrinsic parameters of the cameras:

- Convert to **normalized coordinates** by pre-multiplying all points with the inverse of the intrinsic matrix; set the first camera's coordinate system to world coordinates

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

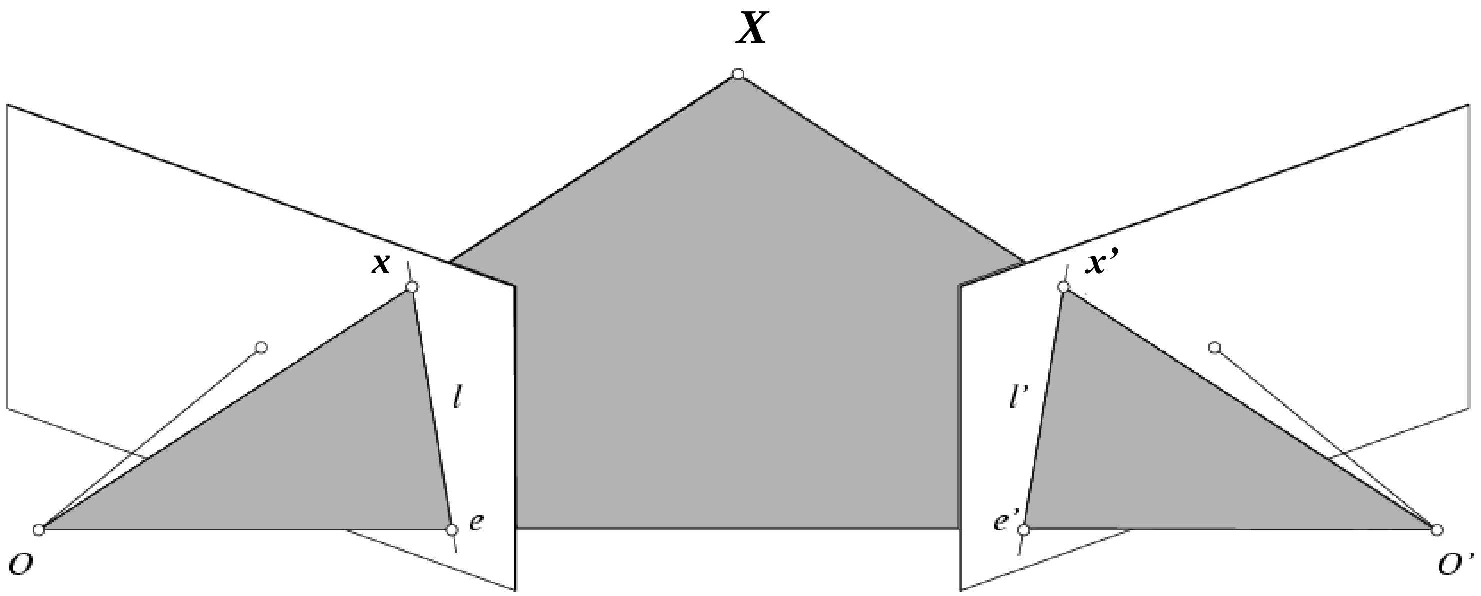
Homogeneous 2d point
(3D ray towards X)

2D pixel coordinate
(homogeneous)

3D scene point

3D scene point in 2nd
camera's 3D coordinates

Epipolar Constraint



The relation of two cameras are described by their extrinsic matrices.
Note that we set the first camera to align with the world frame

2nd camera rotation

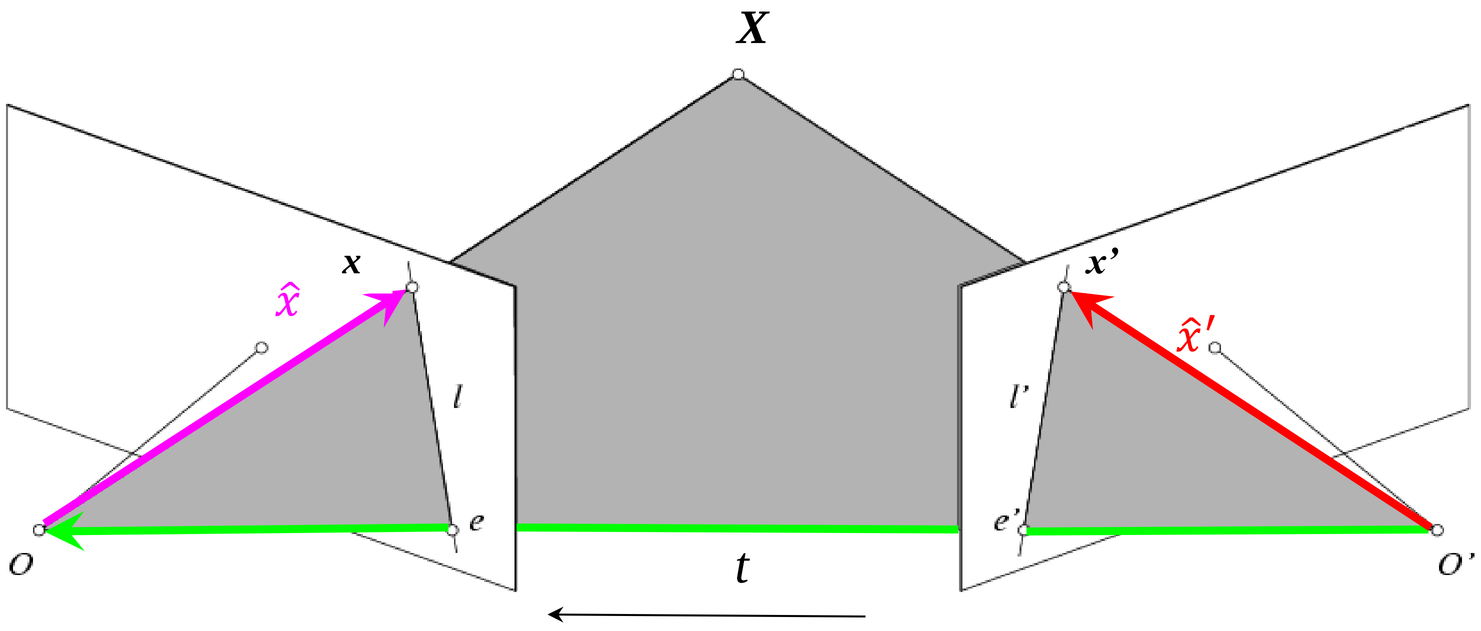
2nd camera translation

$$X' = RX + t$$

3D scene point in 2nd camera's 3D coordinates

3D scene point in 1st camera's 3D coordinates (also world coordinates)

Epipolar Constraint: Co-Planar



$$X' = RX + t$$

$$\lambda_2 \hat{x}' = R\lambda_1 \hat{x} + t$$

$$\lambda_2 \hat{x}' \cdot [t \times (R\lambda_1 \hat{x})] = 0$$

$$\hat{x}' \cdot [t \times (R\hat{x})] = 0$$

(because $\hat{x}, R\hat{x}'$, and t are co-planar)

The Algebraic View of Co-Planar

$X' = RX + t$

$RX \times t$ is the normal of the plane

$X' \cdot (RX \times t) = 0$

X' is a linear combination of RX and t

$\Downarrow$

X' is on the plane formed by RX and t

$\Updownarrow$

X' is orthogonal to the plane normal

Matrix Form of Cross Product

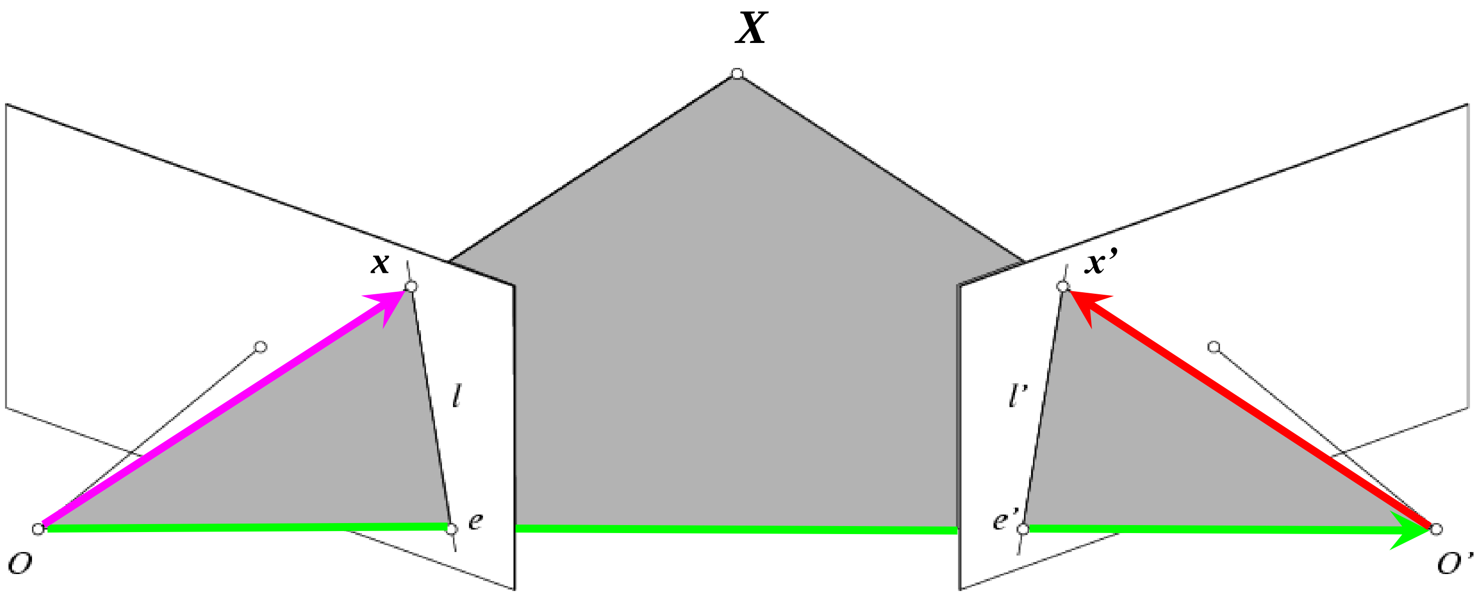
$$\vec{a} \times \vec{b} = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \vec{c}$$

Can be expressed as a matrix multiplication

$$[a_x] = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}$$

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

Epipolar Constraint

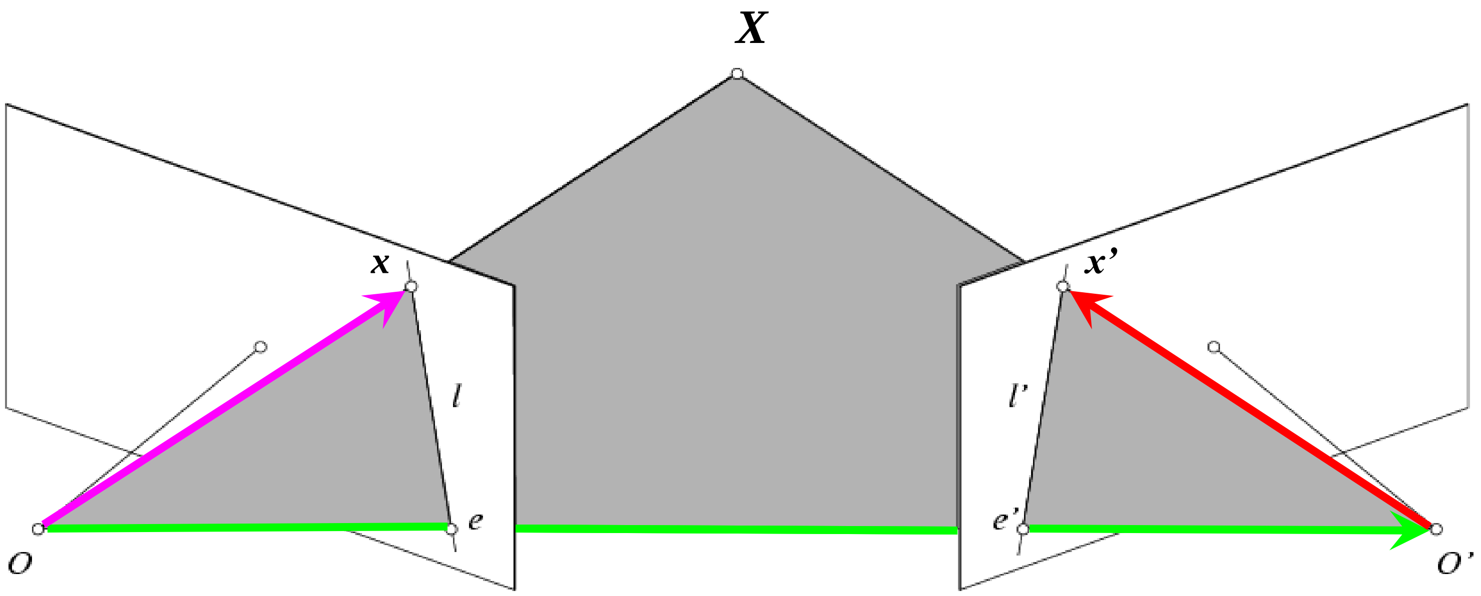


$$\hat{\boldsymbol{x}}' \cdot [\boldsymbol{t} \times (\boldsymbol{R}\hat{\boldsymbol{x}})] = \mathbf{0} \quad \Rightarrow \quad \hat{\boldsymbol{x}}'^{T}[\boldsymbol{t}_{\times}]\boldsymbol{R}\hat{\boldsymbol{x}} = \mathbf{0}$$

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

Essential Matrix

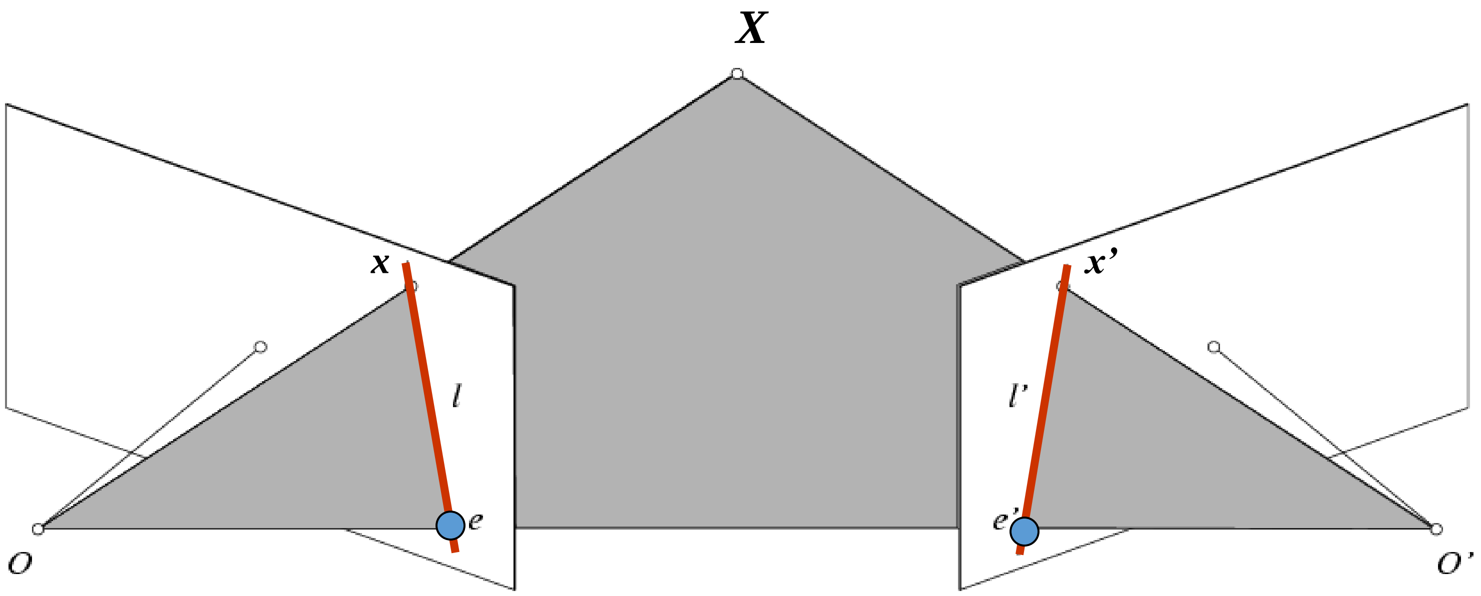


$$\hat{x}' \cdot [t \times (R\hat{x})] = 0 \Rightarrow \hat{x}'^T [t_\times] R\hat{x} = 0 \Rightarrow \hat{x}'^T E\hat{x} = 0$$

Essential Matrix
(Longuet-Higgins, 1981)

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

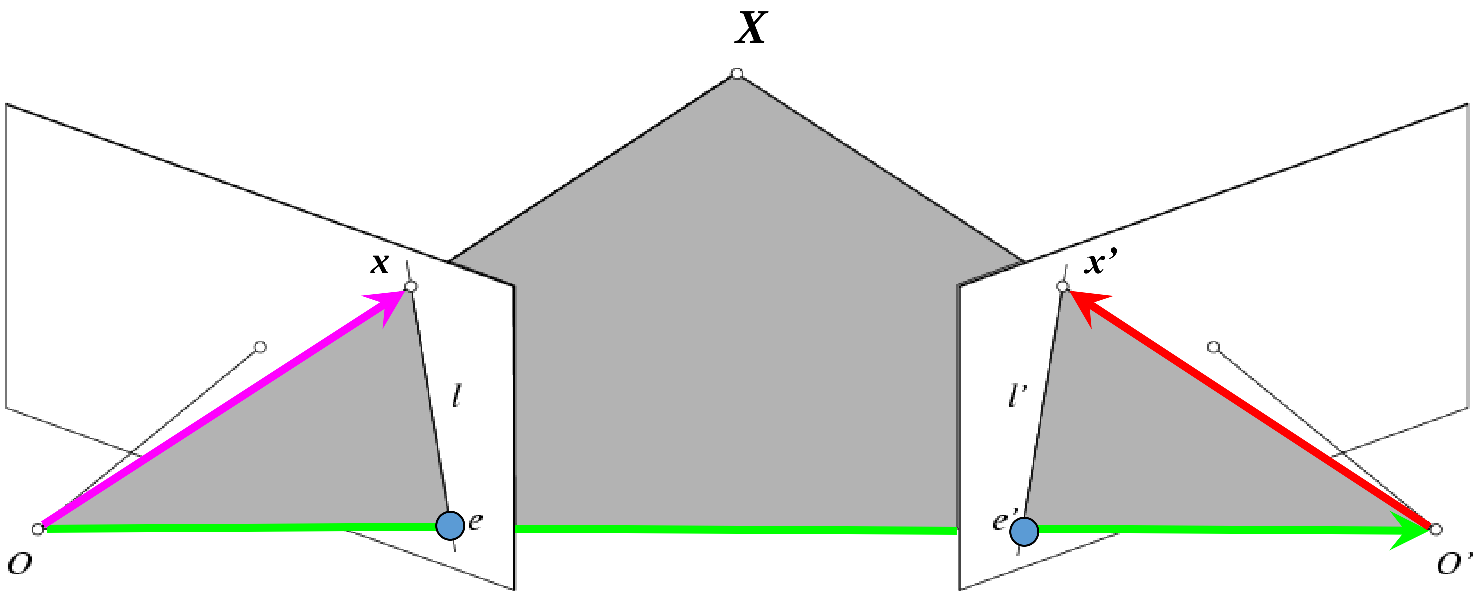
Properties of Essential Matrix



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

- $E\hat{x}$ is the epipolar line associated with $\hat{x}$ ($l' = E\hat{x}$)
- $E^T\hat{x}'$ is the epipolar line associated with $\hat{x}'$ ($l = E^T\hat{x}'$)
- $Ee = 0, E^T e' = 0$
- E is singular with rank 2 ($[t_\times]$ is rank 2)
- E has five degrees of freedom

Epipolar Constraint: Uncalibrated Case



If we do not know K and K', we can write the epipolar constraint in terms of unknown normalized coordinates

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

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

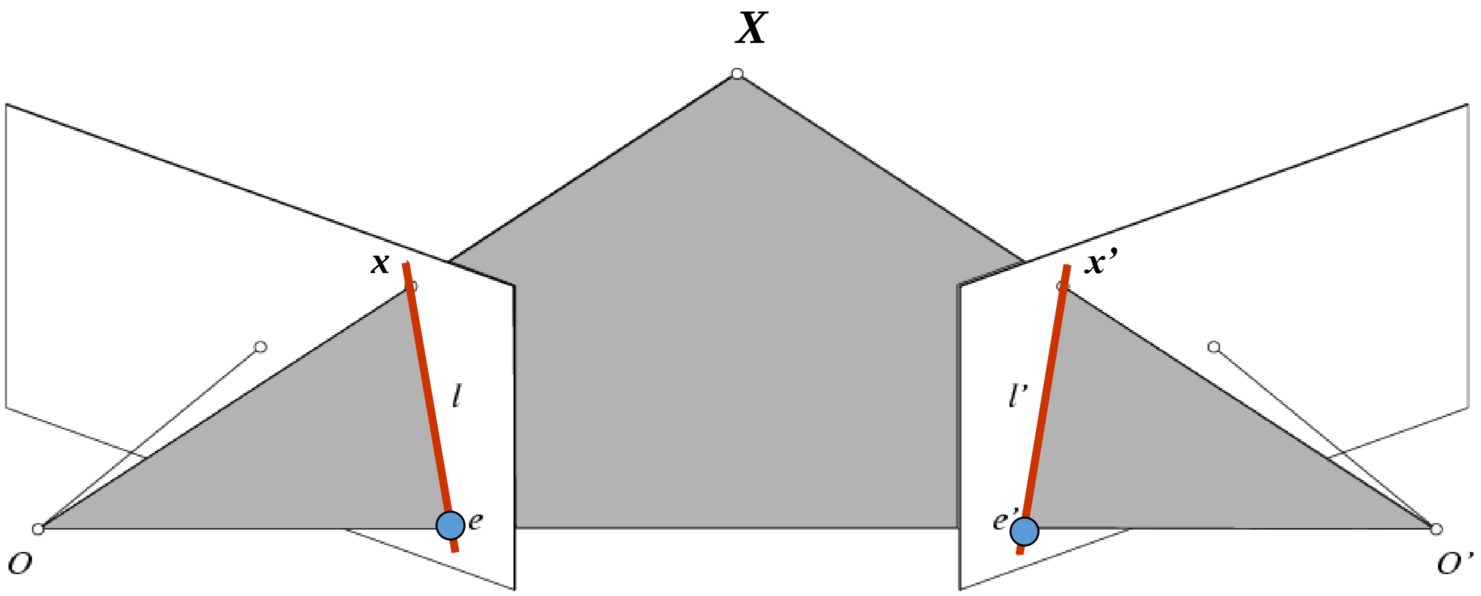
Fundamental Matrix

$$x'^{T} K'^{-T} E K^{-1} x \stackrel{\text{def}}{=} x'^{T} F x = 0$$

$$F = K'^{-T} E K^{-1}$$

Fundamental Matrix
(Faugeras and Luong, 1992)

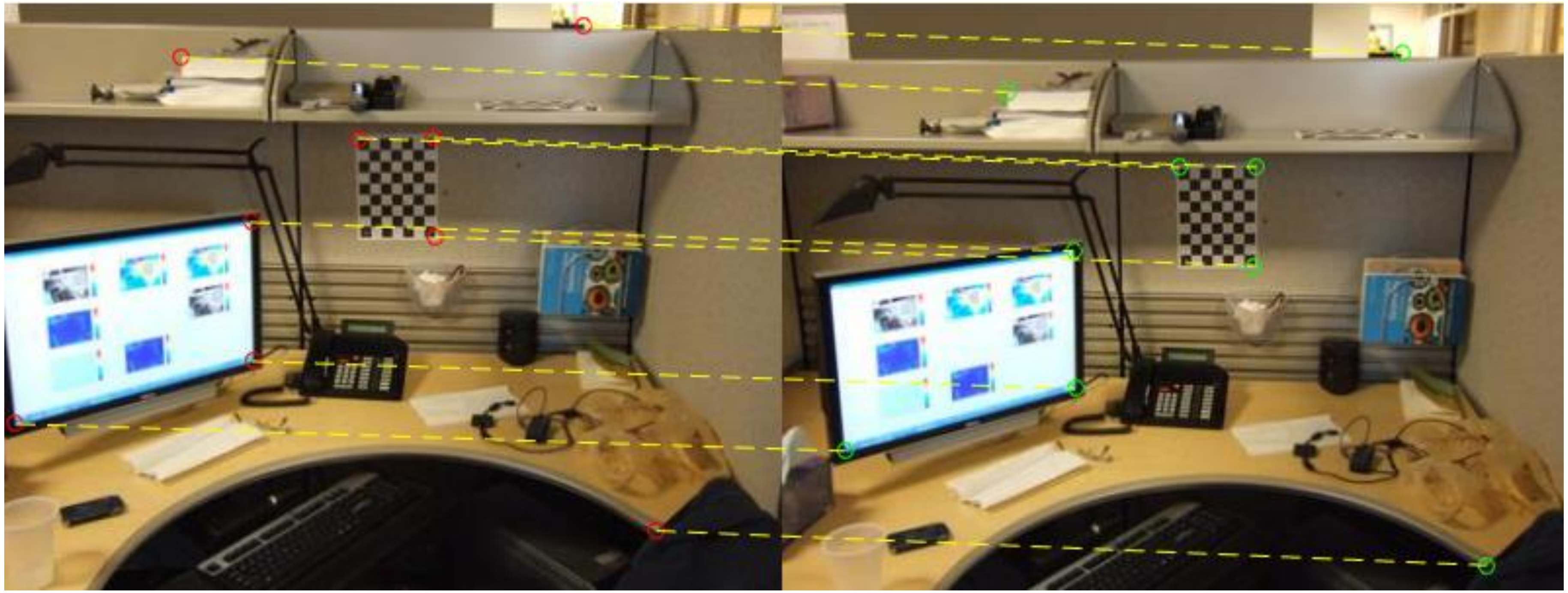
Properties of Fundamental Matrix



$$\boxed{\boldsymbol{x'^T F x = 0}}$$

- Fx is the epipolar line associated with x ($l' = Fx$)
- $F^T x'$ is the epipolar line associated with x' ($l = F^T x'$)
- $Fe = 0, F^T e' = 0$
- F is singular with rank 2 (E is rank 2)
- F has 7 degrees of freedom (up to scale and rank 2)

Estimating the Fundamental Matrix



Given correspondences in two images, estimate the fundamental matrix

Linear System of Fundamental Matrix

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

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

$$x'^T F x = 0$$

Definition of fundamental matrix

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

$$a^T f = 0$$

Homogeneous linear equation

8-Point Algorithm: Step 1

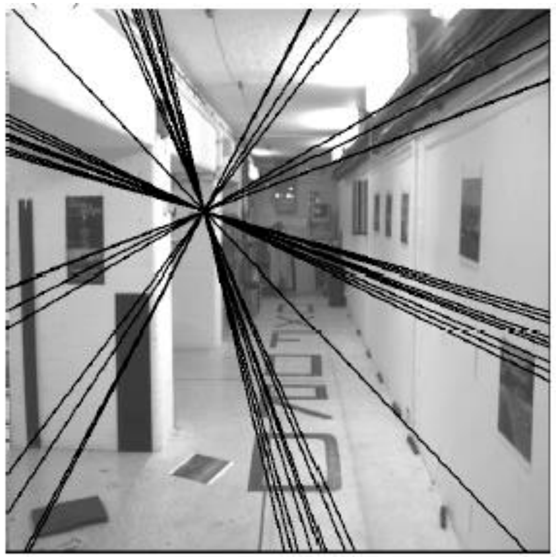
$$\mathrm{A}\boldsymbol{f} = \begin{bmatrix} u_1{}'u_1 & u_1{}'v_1 & u_1{}' & v_1{}'u_1 & v_1{}'v_1 & v_1{}' & u_1 & v_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ u_n'u_v & u_n{}'v_n & u_n{}' & v_n{}'u_n & v_n{}'v_n & v_n{}' & u_n & v_n & 1 \end{bmatrix} \begin{bmatrix} f_{11} \\ f_{12} \\ f_{13} \\ f_{21} \\ \vdots \\ f_{33} \end{bmatrix} = \mathbf{0}$$

Given 8 pairs, A is an 8x9 matrix with 1D null space

Question: How to solve Af=0?

Answer: Least square solution

8-Point Algorithm: Step 2



Left: uncorrected F
Epipolar lines are not conincident

Right: corrected F

Enforce the estimated fundamental matrix F to be rank-2

8-Point Algorithm: Step 2

Then, enforce rank 2 constraint (simple and convenient method using SVD)

$$\mathtt{F} = \mathtt{U}\Sigma\mathtt{V}^\top, \text{ where } \Sigma = \text{diag}(\sigma_1, \sigma_2, \sigma_3)$$
$$\mathtt{F} = \mathtt{U}\Sigma'\mathtt{V}^\top, \text{ where } \Sigma' = \text{diag}(\sigma_1, \sigma_2, 0)$$

Eigen values are in a descending order

Question: Why do we discard the smallest eigen value?

Source: Ben Ocha CSE252B-WI24

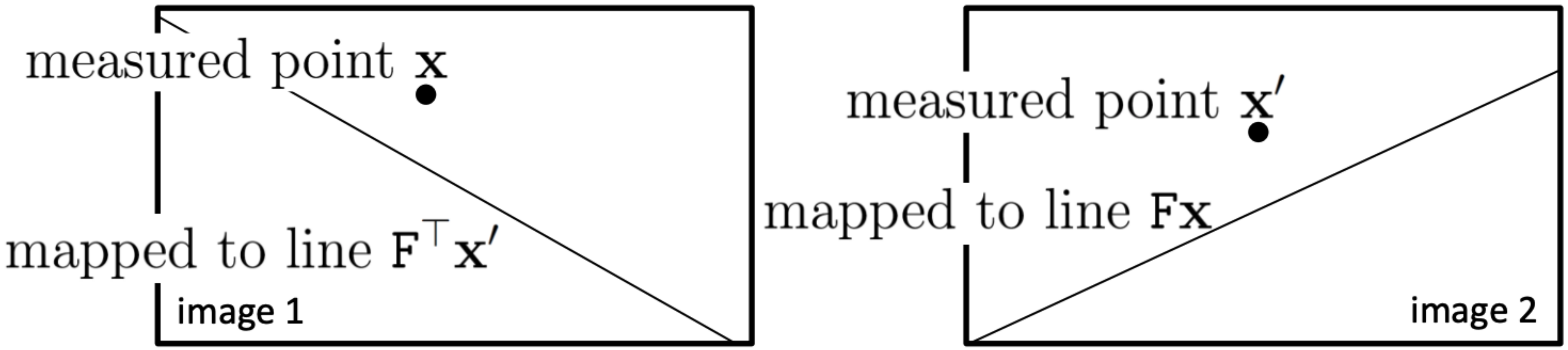
Other Algorithms

- 7-point algorithm
  - Similarly, solve the null space of A (7x9 matrix) with 7 pairs of correspondences
  - Solve linear combination of null space vectors that satisfy det(F)=0
  - More complicated but with fewer points and thus faster

- Non-linear optimization
  - Minimize reprojection error (how to define?)

Reprojection Error

$$\mathbf{x}'^{\top}\mathbf{F}\mathbf{x} = 0 \qquad \text{Epipolar constraint}$$



Measured points should be close to the mapped epipolar line

Problem with 8-Point Algorithm

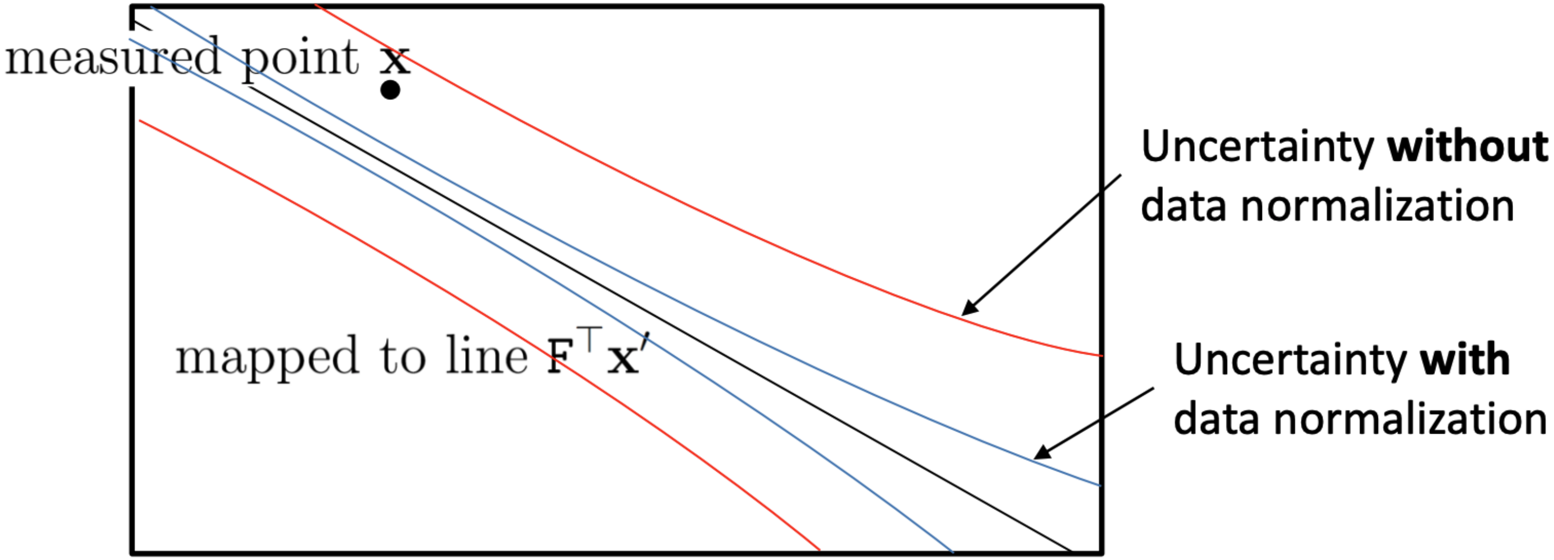
250906.36	183269.57	921.81	200931.10	146766.13	738.21	272.19	198.81
2692.28	131633.03	176.27	6196.73	302975.59	405.71	15.27	746.79
416374.23	871684.30	935.47	408110.89	854384.92	916.90	445.10	931.81
191183.60	171759.40	410.27	416435.62	374125.90	893.65	465.99	418.65
48988.86	30401.76	57.89	298604.57	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

An example of A in Af=0 (the first 8 columns) without data normalization

- Poor numerical conditioning
- Can be fixed by rescaling the data

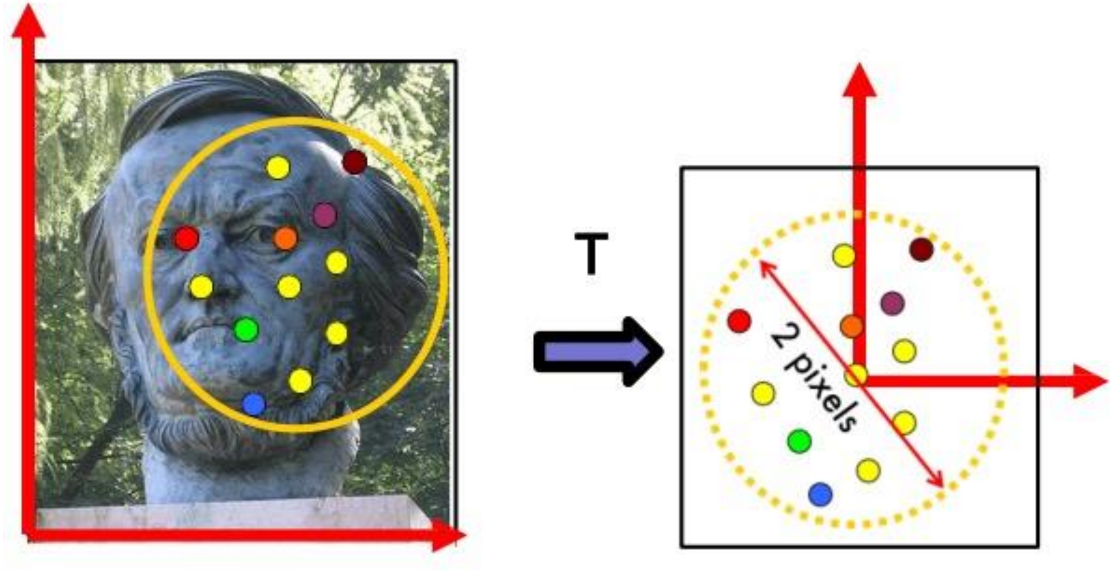
See the Sec 4.4.4 of *Multi View Geometry*

Normalization for Better Conditioning Number



Data normalization is required to reduce the propagated uncertainty (related to the conditioning number) under the nonlinear projection

Data Normalization



Determine the similarity transformation T such that the mean (i.e., centroid) of the transformed points is at the origin and their standard deviation from the origin is $\sqrt{2}$.

$$\mathtt{T} = \begin{bmatrix} s & 0 & -s\mu_{\tilde{x}} \\ 0 & s & -s\mu_{\tilde{y}} \\ 0 & 0 & 1 \end{bmatrix}, \text{ where } s = \sqrt{\frac{2}{\sigma_{\tilde{x}}^2 + \sigma_{\tilde{y}}^2}}$$

where $\mu_{\tilde{x}}$ and $\sigma_{\tilde{x}}^2$, and $\mu_{\tilde{y}}$ and $\sigma_{\tilde{y}}^2$ are the mean and variance of the $\tilde{x}$ and $\tilde{y}$ coordinates, respectively.

Normalized 8-Point Algorithm

- 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 ***F*** from the normalized points

- Enforce the rank-2 constraint (for example, take SVD of ***F*** and throw out the smallest singular value)

- Transform fundamental matrix back to original units: if ***T*** and ***T'*** are the normalizing transformations in the two images, then the fundamental matrix in original coordinates is $T'^{T} F T$

(Hartley, 1995)

Comparison of Different 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 intrinsic (calibration) matrices of two cameras, we can estimate the essential matrix: $E = K'^T F K$
- The essential matrix gives us the relative rotation and translation between the cameras, or their extrinsic parameters
- However, there exists an algorithm to solve the essential matrix if camera intrinsics are known, which generates more accurate results than converting the essential matrix from the fundamental matrix