

Introduction to Visual Computing

Prof. Feng Liu

Winter 2020

<http://www.cs.pdx.edu/~fliu/courses/cs410/>

02/11/2020

Last Time

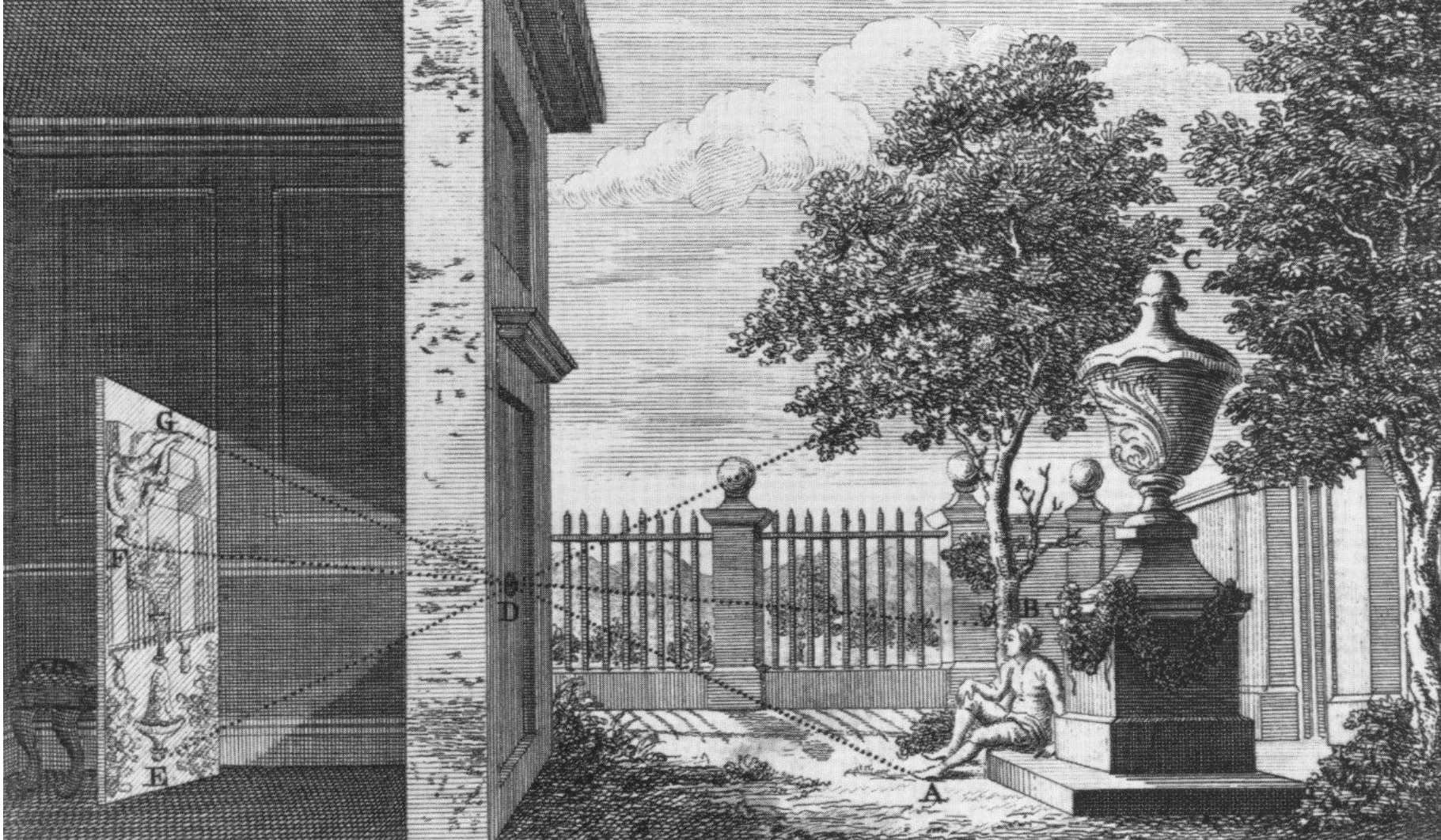
- Motion estimation
 - Parametric motion
 - image alignment
 - Optical flow
 - Feature tracking

Today

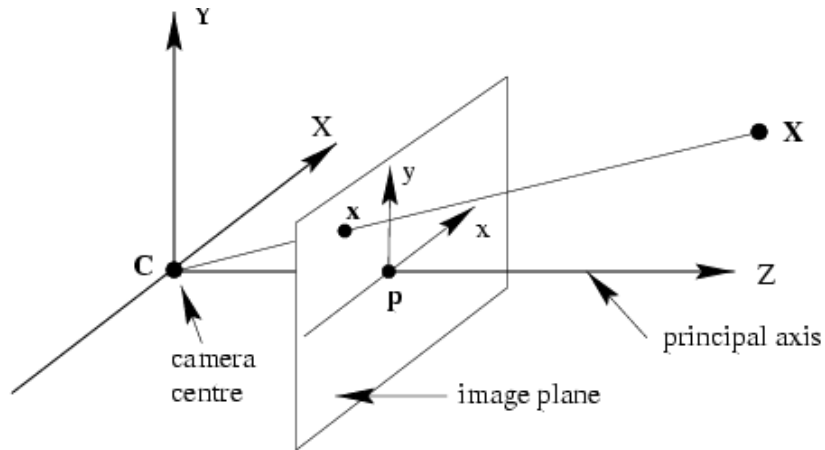
- Camera
- Calibration

The slides for this topic are used from Prof. S. Lazebnik.

Pinhole imaging

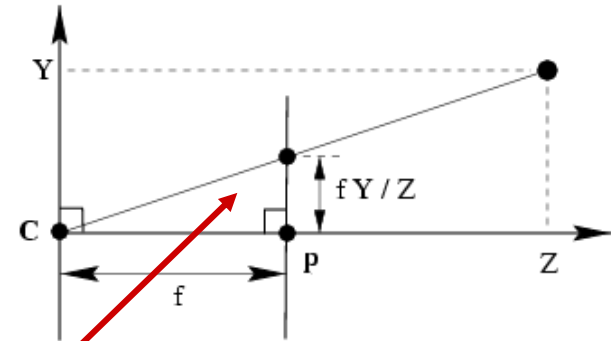
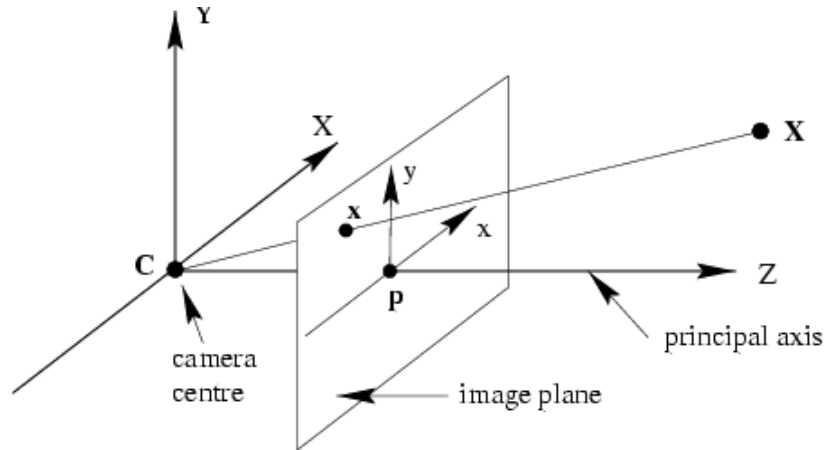


Pinhole camera model



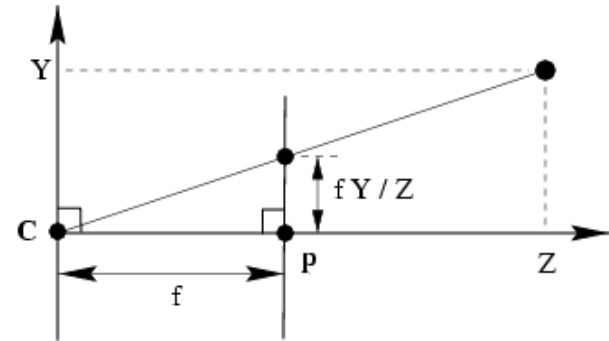
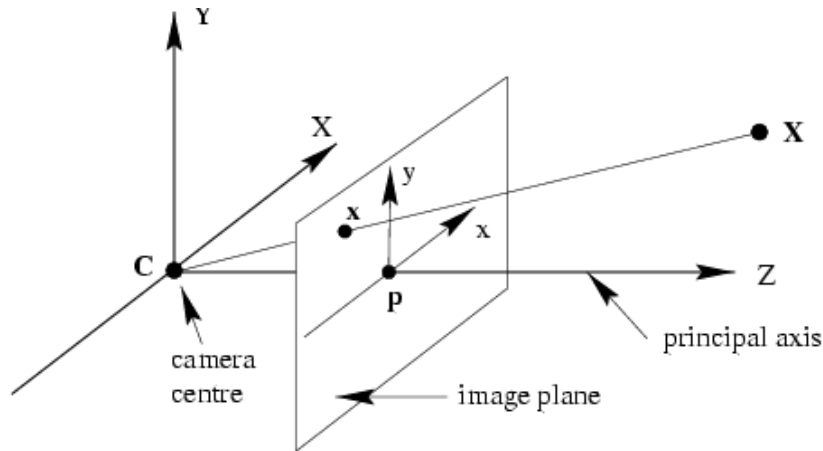
- **Principal axis:** line from the camera center perpendicular to the image plane
 - **Normalized (camera) coordinate system:** camera center is at the origin and the principal axis is the z -axis
-

Pinhole camera model



$$\frac{y}{f} = \frac{Y}{Z} \quad \mapsto \quad y = \frac{fY}{Z}$$
$$\frac{x}{f} = \frac{X}{Z} \quad \mapsto \quad x = \frac{fX}{Z}$$

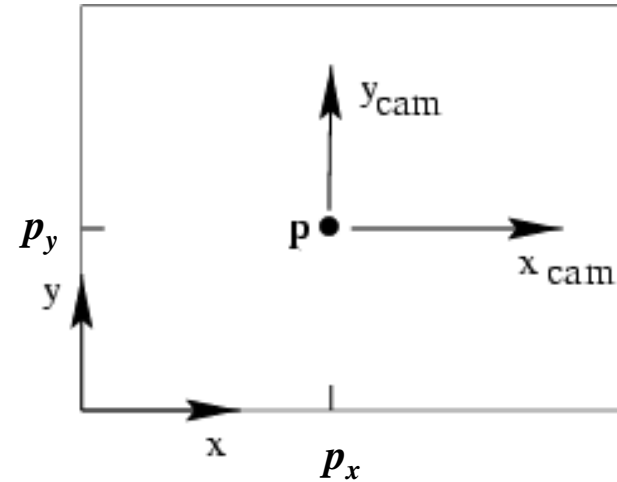
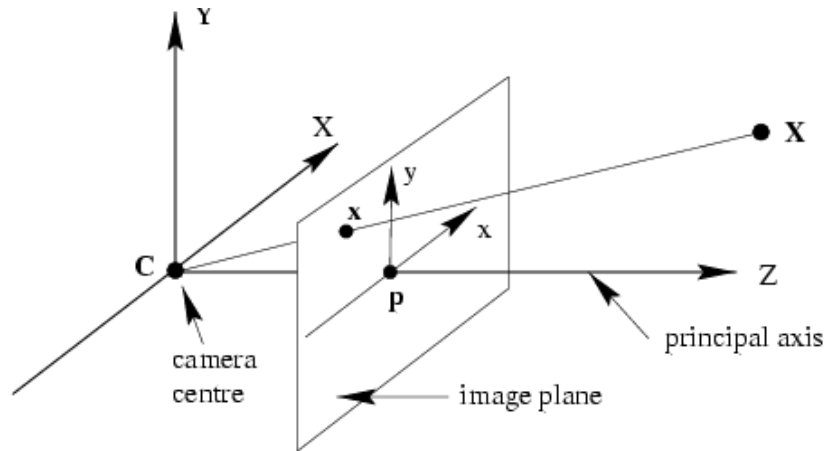
Pinhole camera model



$$(X, Y, Z) \mapsto (f X / Z, f Y / Z)$$

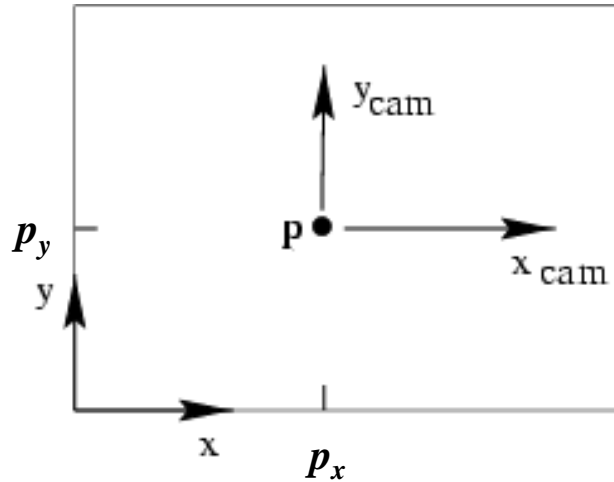
$$\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} f X \\ f Y \\ Z \end{pmatrix} = \begin{bmatrix} f & 0 & 0 \\ & f & 0 \\ & & 1 & 0 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \quad \mathbf{x} = \mathbf{P}\mathbf{X}$$

Principal point



- **Principal point (p):** point where principal axis intersects the image plane (origin of normalized coordinate system)
 - Normalized coordinate system: origin is at the principal point
 - Image coordinate system: origin is in the corner
 - How to go from normalized coordinate system to image coordinate system?
-

Principal point offset

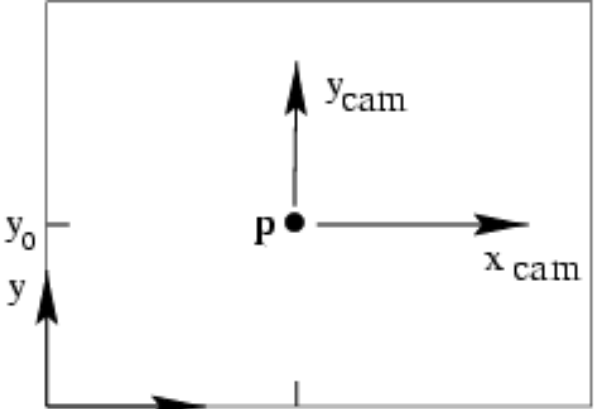


principal point: (p_x, p_y)

$$(X, Y, Z) \mapsto (fX/Z + p_x, fY/Z + p_y)$$

$$\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} fX + Zp_x \\ fY + Zp_y \\ Z \end{pmatrix} = \begin{bmatrix} f & p_x & 0 \\ & f & p_y & 0 \\ & & 1 & 0 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

Principal point offset

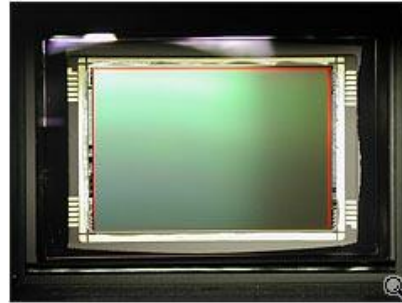


principal point: (p_x, p_y)

$$\begin{pmatrix} fX + Zp_x \\ fY + Zp_y \\ Z \end{pmatrix} = \begin{bmatrix} f & & & \\ & f & & \\ & & 1 & \\ & & & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ & & 1 \\ & & & 0 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

$$K = \begin{bmatrix} f & & p_x \\ & f & p_y \\ & & 1 \end{bmatrix} \text{ calibration matrix } \mathbf{P} = \mathbf{K}[\mathbf{I} \mid \mathbf{0}]$$

Pixel coordinates



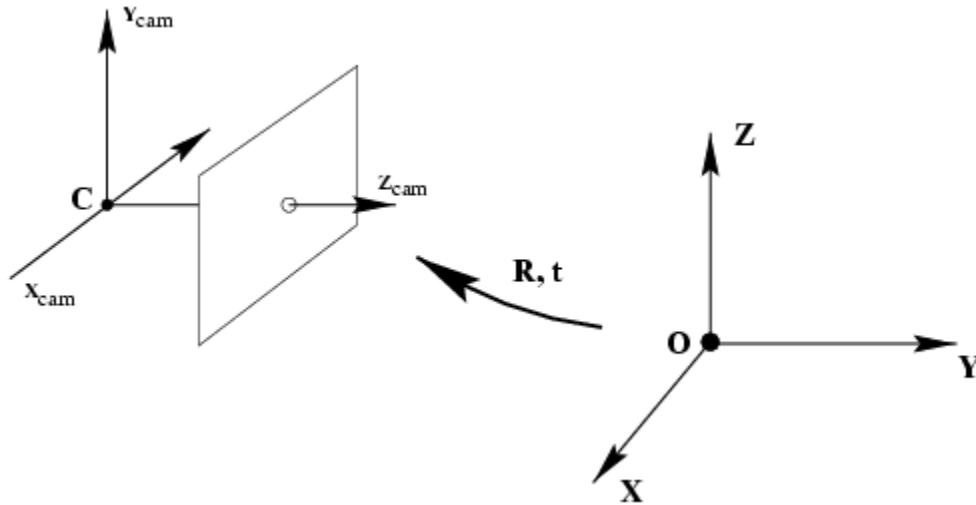
Pixel size: $\frac{1}{m_x} \times \frac{1}{m_y}$

- m_x pixels per meter in horizontal direction,
 m_y pixels per meter in vertical direction

$$K = \begin{bmatrix} m_x & & \\ & m_y & \\ & & 1 \end{bmatrix} \begin{bmatrix} f \\ f \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha_x & \beta_x \\ \alpha_y & \beta_y \\ & 1 \end{bmatrix}$$

pixels/m
m
pixels

Camera rotation and translation



- In general, the camera coordinate frame will be related to the world coordinate frame by a rotation and a translation

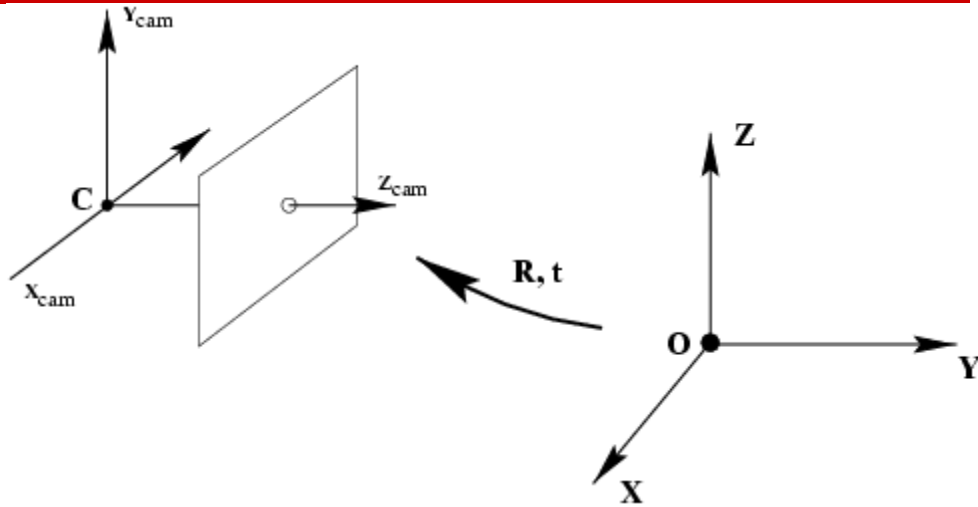
$$\tilde{X}_{cam} = R(\tilde{X} - \tilde{C})$$

coords. of point in camera frame

coords. of a point in world frame (nonhomogeneous)

coords. of camera center in world frame

Camera rotation and translation



In non-homogeneous coordinates:

$$\tilde{X}_{cam} = R(\tilde{X} - \tilde{C})$$

$$X_{cam} = \begin{bmatrix} R & -R\tilde{C} \\ 0 & 1 \end{bmatrix} \begin{pmatrix} \tilde{X} \\ 1 \end{pmatrix} = \begin{bmatrix} R & -R\tilde{C} \\ 0 & 1 \end{bmatrix} X$$

$$x = K[I | 0]X_{cam} = K[R | -R\tilde{C}]X \quad P = K[R | t], \quad t = -R\tilde{C}$$

Note: C is the null space of the camera projection matrix ($PC=0$)

Camera parameters

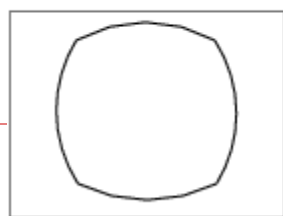
- Intrinsic parameters

- Principal point coordinates
- Focal length
- Pixel magnification factors
- *Skew (non-rectangular pixels)*
- *Radial distortion*

$$K = \begin{bmatrix} m_x & & \\ & m_y & \\ & & 1 \end{bmatrix} \begin{bmatrix} f & p_x \\ f & p_y \\ & 1 \end{bmatrix} = \begin{bmatrix} \alpha_x & \beta_x \\ & \alpha_y & \beta_y \\ & & 1 \end{bmatrix}$$



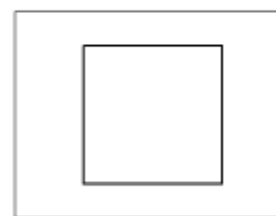
radial distortion



correction →



linear image



Camera parameters

- Intrinsic parameters
 - Principal point coordinates
 - Focal length
 - Pixel magnification factors
 - *Skew (non-rectangular pixels)*
 - *Radial distortion*
 - Extrinsic parameters
 - Rotation and translation relative to world coordinate system
-

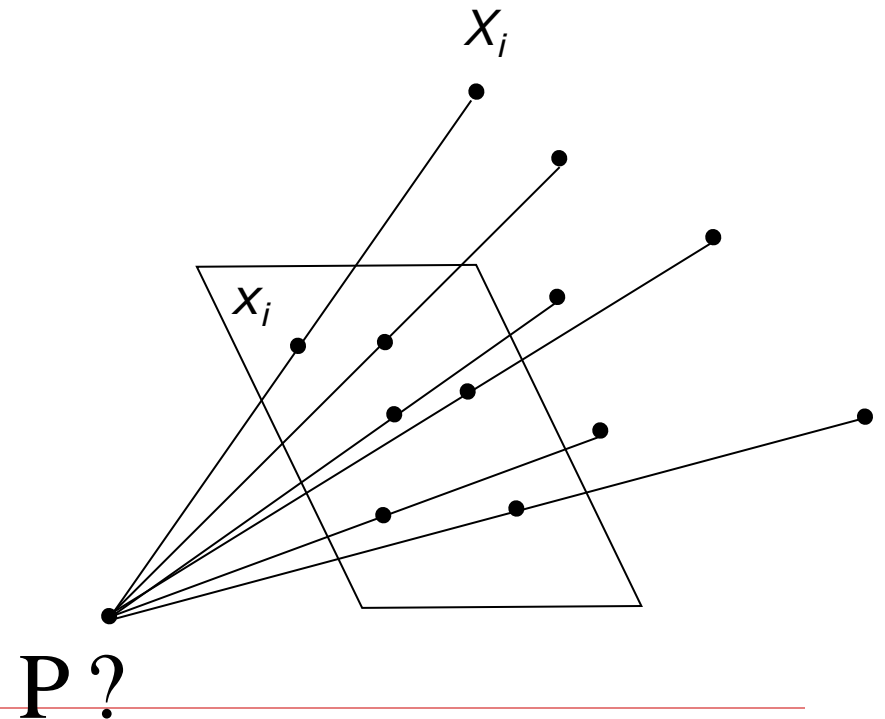
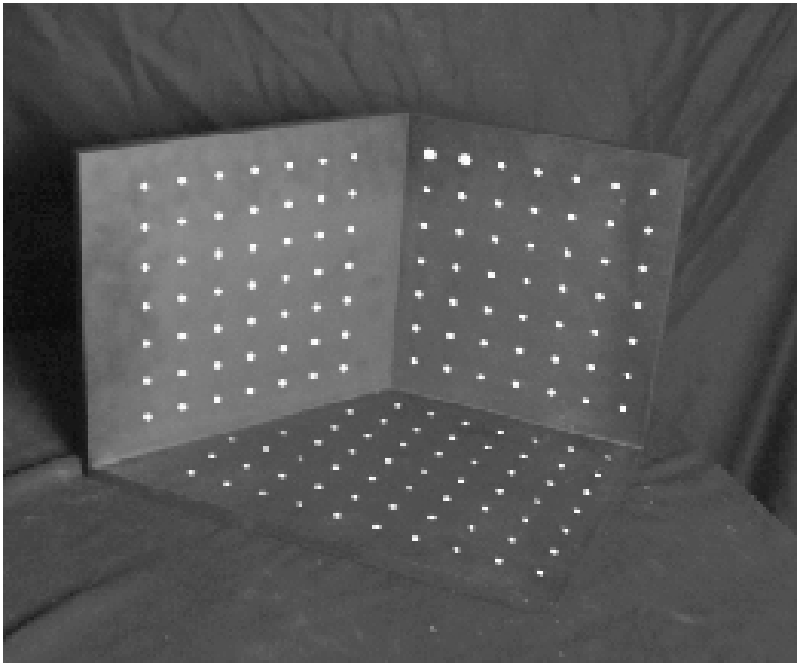
Camera calibration

$$\mathbf{x} = \mathbf{K}[\mathbf{R} \quad \mathbf{t}] \mathbf{X}$$

$$\begin{bmatrix} \lambda x \\ \lambda y \\ \lambda \end{bmatrix} = \begin{bmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Camera calibration

- Given n points with known 3D coordinates X_i and known image projections x_i , estimate the camera parameters



Camera calibration: Linear method

$$\lambda \mathbf{x}_i = \mathbf{P} \mathbf{X}_i$$

$$\mathbf{x}_i \times \mathbf{P} \mathbf{X}_i = \mathbf{0}$$

$$\begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} \times \begin{bmatrix} \mathbf{P}_1^T \mathbf{X}_i \\ \mathbf{P}_2^T \mathbf{X}_i \\ \mathbf{P}_3^T \mathbf{X}_i \end{bmatrix} = \mathbf{0}$$

$$\begin{bmatrix} 0 & -\mathbf{X}_i^T & y_i \mathbf{X}_i^T \\ \mathbf{X}_i^T & 0 & -x_i \mathbf{X}_i^T \\ -y_i \mathbf{X}_i^T & x_i \mathbf{X}_i^T & 0 \end{bmatrix} \begin{pmatrix} \mathbf{P}_1 \\ \mathbf{P}_2 \\ \mathbf{P}_3 \end{pmatrix} = \mathbf{0}$$

Two linearly independent equations

Camera calibration: Linear method

$$\begin{bmatrix} 0^T & \mathbf{X}_1^T & -y_1 \mathbf{X}_1^T \\ \mathbf{X}_1^T & 0^T & -x_1 \mathbf{X}_1^T \\ \dots & \dots & \dots \\ 0^T & \mathbf{X}_n^T & -y_n \mathbf{X}_n^T \\ \mathbf{X}_n^T & 0^T & -x_n \mathbf{X}_n^T \end{bmatrix} \begin{pmatrix} \mathbf{P}_1 \\ \mathbf{P}_2 \\ \mathbf{P}_3 \end{pmatrix} = 0 \quad \mathbf{A}\mathbf{p} = 0$$

- P has 11 degrees of freedom (12 parameters, but scale is arbitrary)
 - One 2D/3D correspondence gives us two linearly independent equations
 - Homogeneous least squares
 - 6 correspondences needed for a minimal solution
-

Camera calibration: Linear method

$$\begin{bmatrix} 0^T & \mathbf{X}_1^T & -y_1 \mathbf{X}_1^T \\ \mathbf{X}_1^T & 0^T & -x_1 \mathbf{X}_1^T \\ \dots & \dots & \dots \\ 0^T & \mathbf{X}_n^T & -y_n \mathbf{X}_n^T \\ \mathbf{X}_n^T & 0^T & -x_n \mathbf{X}_n^T \end{bmatrix} \begin{pmatrix} \mathbf{P}_1 \\ \mathbf{P}_2 \\ \mathbf{P}_3 \end{pmatrix} = 0 \quad \mathbf{A}\mathbf{p} = 0$$

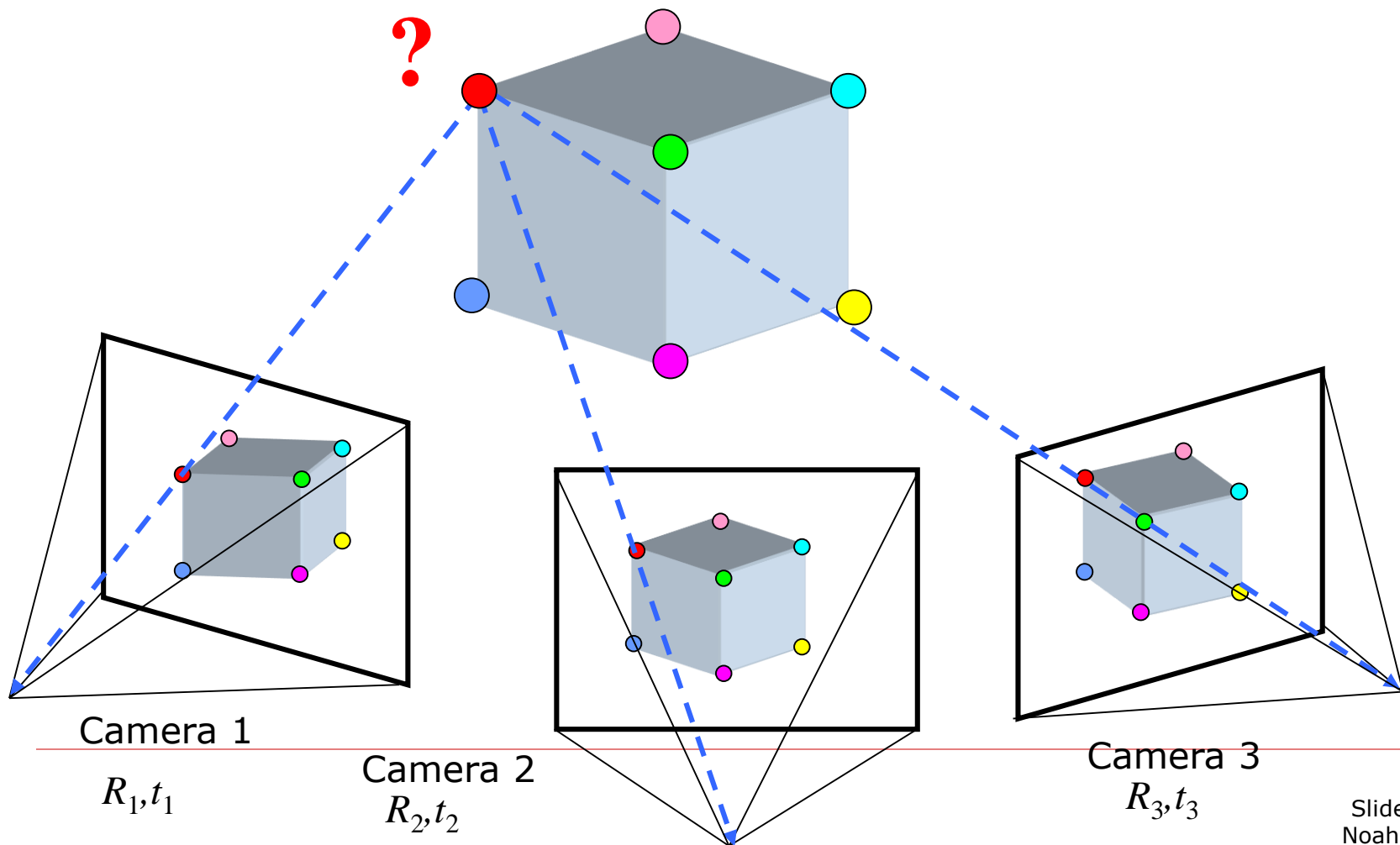
- Note: for coplanar points that satisfy $\Pi^T \mathbf{X} = 0$, we will get degenerate solutions $(\Pi, 0, 0)$, $(0, \Pi, 0)$, or $(0, 0, \Pi)$
-

Camera calibration: Linear method

- Advantages: easy to formulate and solve
- Disadvantages
 - Doesn't directly tell you camera parameters
 - Doesn't model radial distortion
 - Can't impose constraints, such as known focal length and orthogonality
- Non-linear methods are preferred
 - Define error as difference between projected points and measured points
 - Minimize error using Newton's method or other non-linear optimization

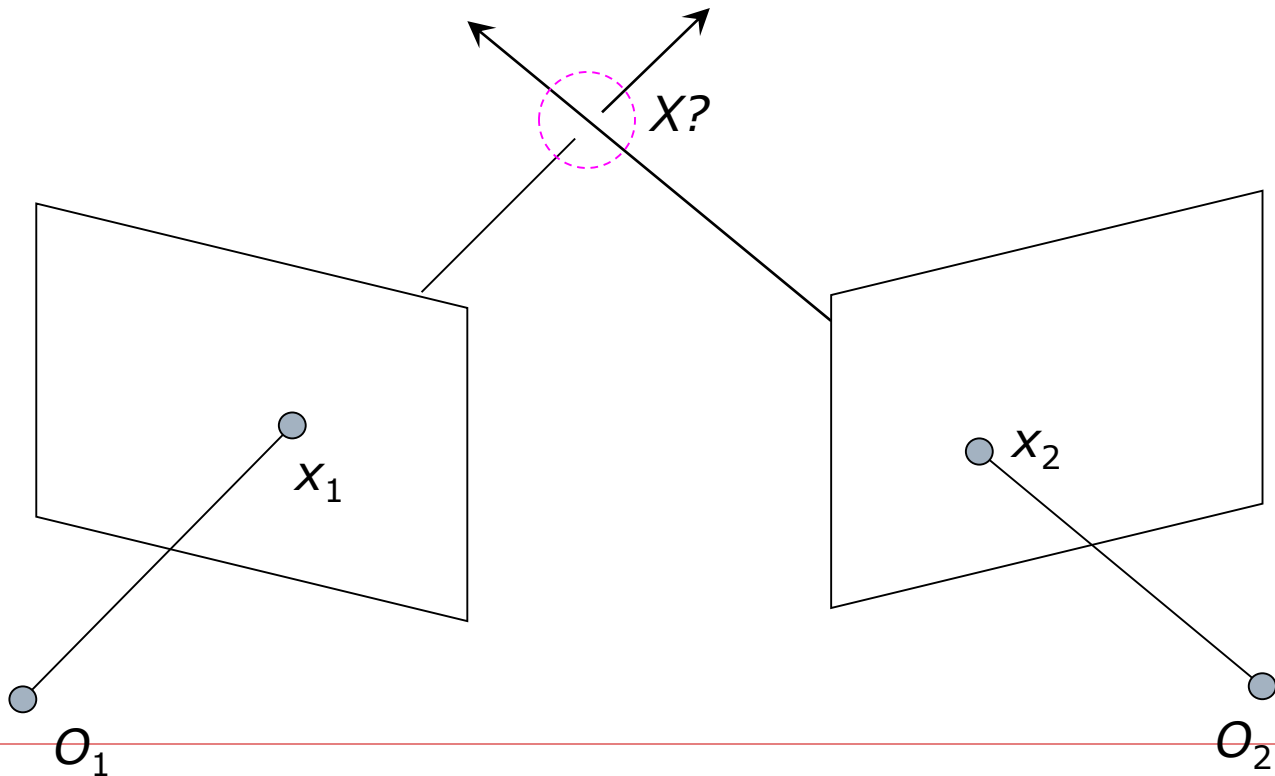
3D reconstruction

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



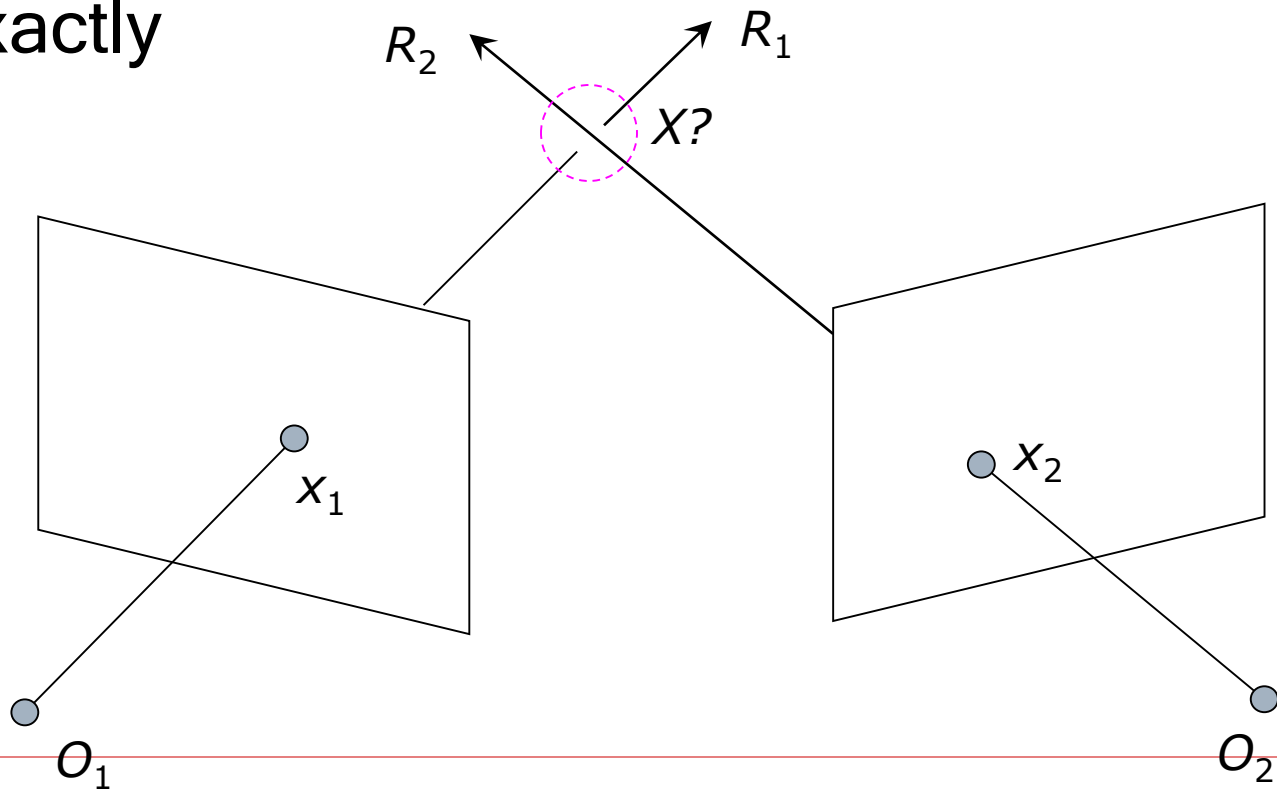
Triangulation

- Given projections of a 3D point in two or more images (with known camera matrices), find the coordinates of the point



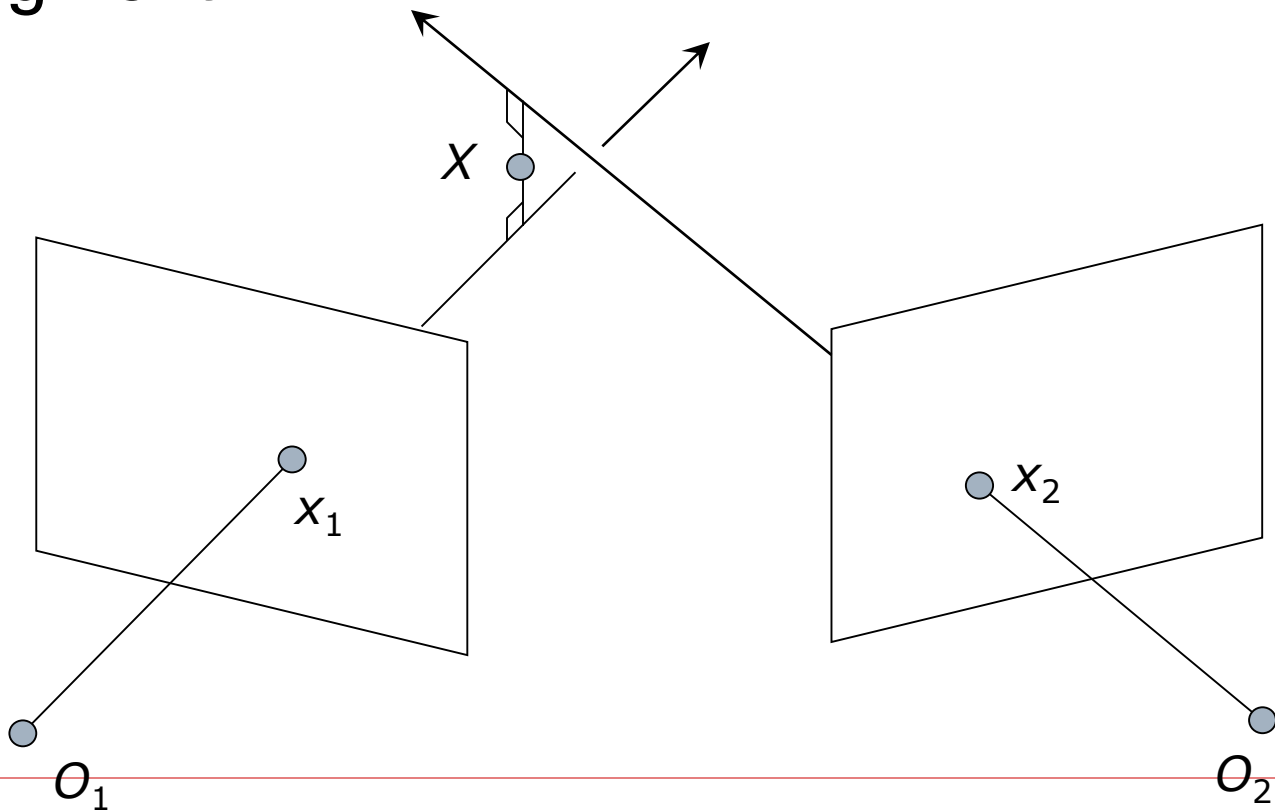
Triangulation

- We want to intersect the two visual rays corresponding to x_1 and x_2 , but because of noise and numerical errors, they don't meet exactly



Triangulation: Geometric approach

- Find shortest segment connecting the two viewing rays and let X be the midpoint of that segment



Triangulation: Linear approach

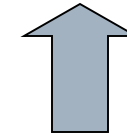
$$\begin{array}{lll} \lambda_1 \mathbf{x}_1 = \mathbf{P}_1 \mathbf{X} & \mathbf{x}_1 \times \mathbf{P}_1 \mathbf{X} = 0 & [\mathbf{x}_{1 \times}] \mathbf{P}_1 \mathbf{X} = 0 \\ \lambda_2 \mathbf{x}_2 = \mathbf{P}_2 \mathbf{X} & \mathbf{x}_2 \times \mathbf{P}_2 \mathbf{X} = 0 & [\mathbf{x}_{2 \times}] \mathbf{P}_2 \mathbf{X} = 0 \end{array}$$

Cross product as matrix multiplication:

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

Triangulation: Linear approach

$$\begin{array}{lll} \lambda_1 \mathbf{x}_1 = \mathbf{P}_1 \mathbf{X} & \mathbf{x}_1 \times \mathbf{P}_1 \mathbf{X} = 0 & [\mathbf{x}_{1 \times}] \mathbf{P}_1 \mathbf{X} = 0 \\ \lambda_2 \mathbf{x}_2 = \mathbf{P}_2 \mathbf{X} & \mathbf{x}_2 \times \mathbf{P}_2 \mathbf{X} = 0 & [\mathbf{x}_{2 \times}] \mathbf{P}_2 \mathbf{X} = 0 \end{array}$$

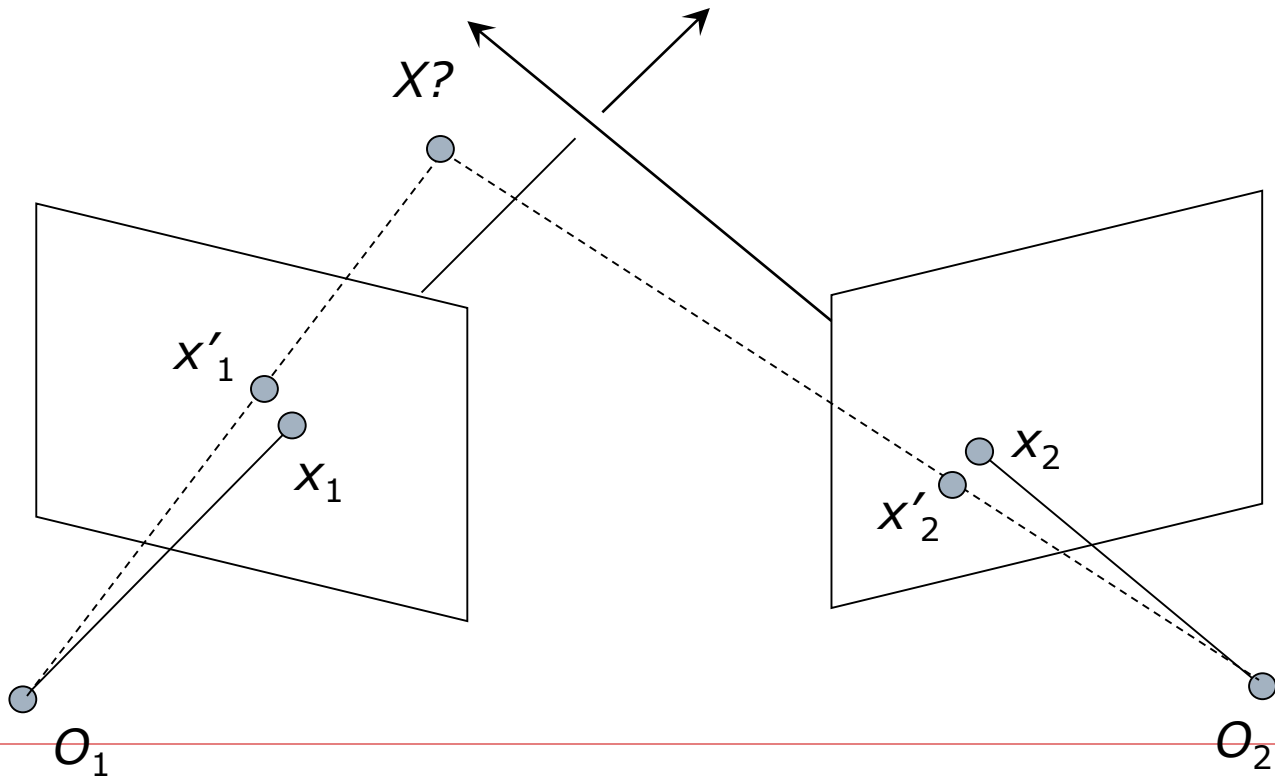


Two independent equations each in terms of three unknown entries of \mathbf{X}

Triangulation: Nonlinear approach

□ Find X that minimizes

$$d^2(x_1, P_1 X) + d^2(x_2, P_2 X)$$



Next Time

- Two-view geometry
- Structure from motion