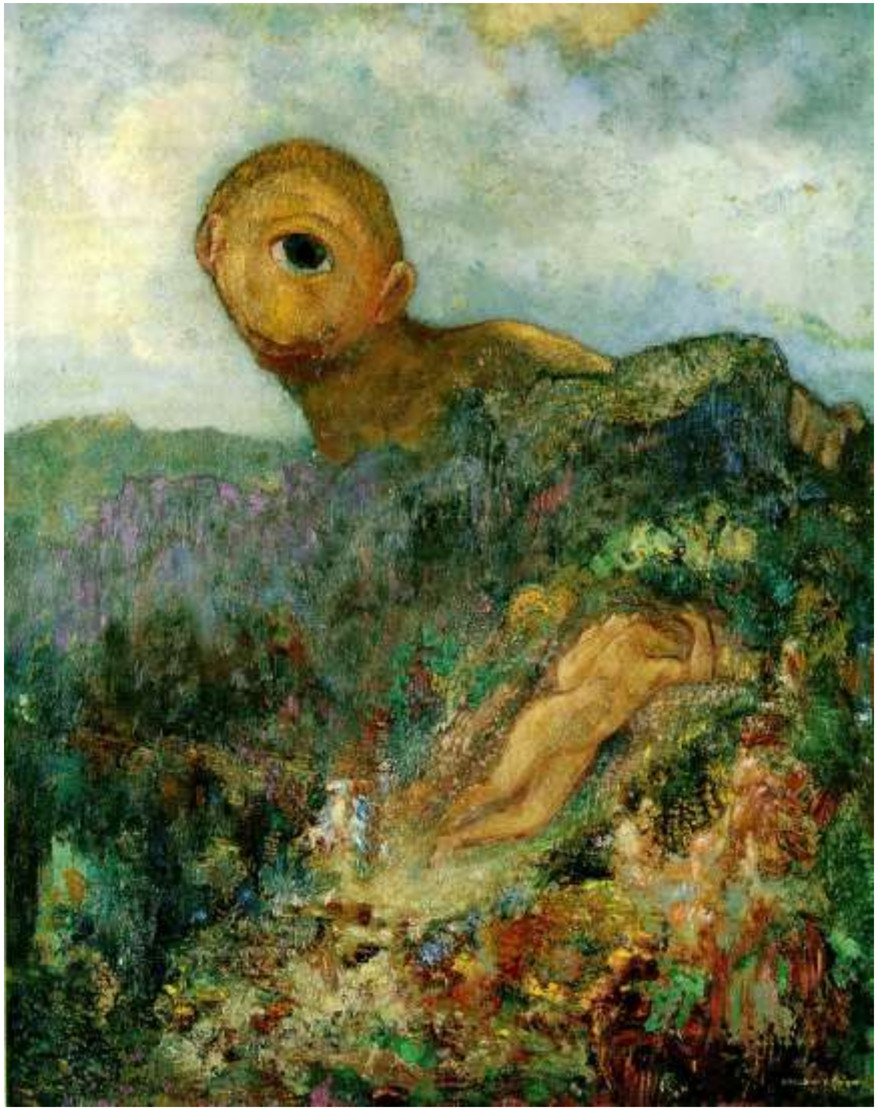
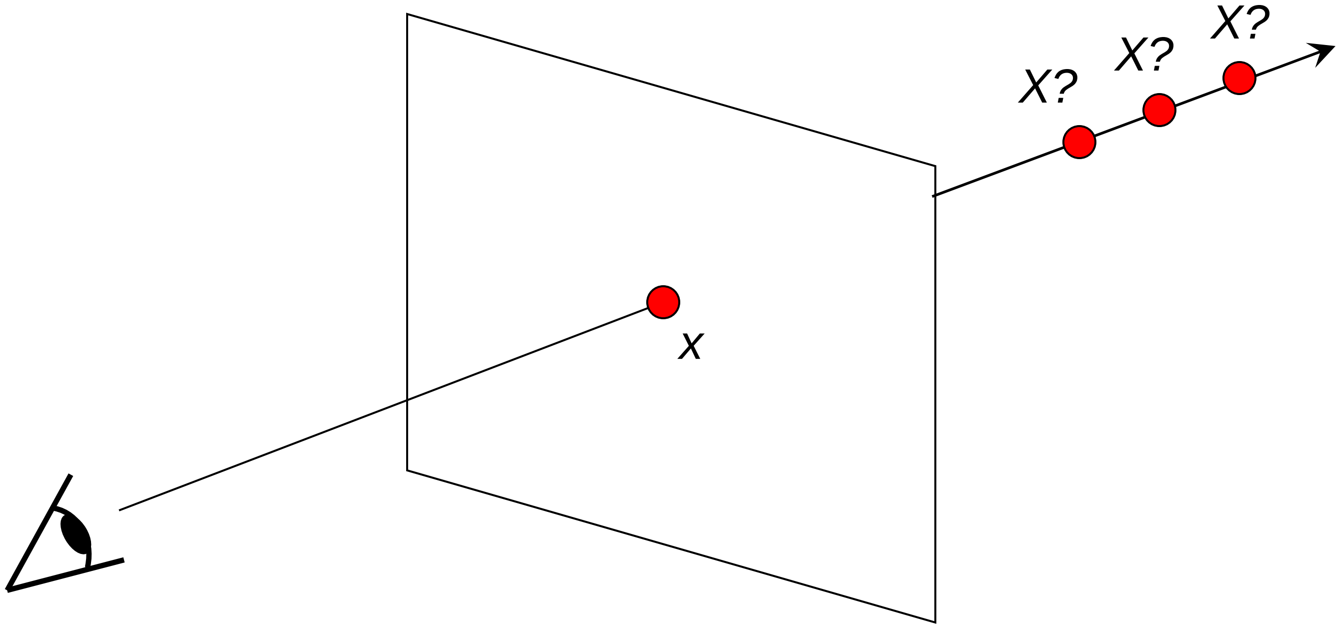
Calibrating a single camera



Odilon Redon, *Cyclops*, 1914

Our goal: Recovery of 3D structure

- Recovery of structure from one image is inherently ambiguous

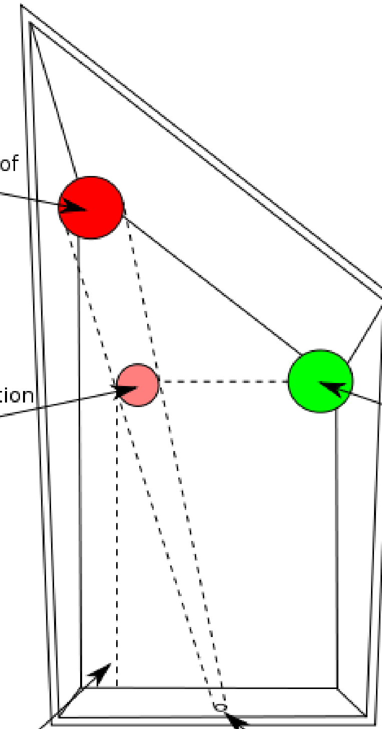


Single-view ambiguity

Single-view ambiguity

[Rashad Alakbarov shadow sculptures]

Single-view ambiguity

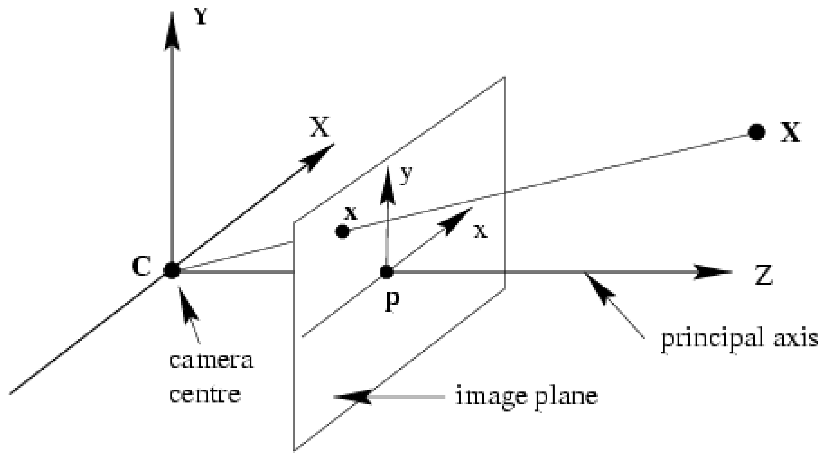


[Ames room]

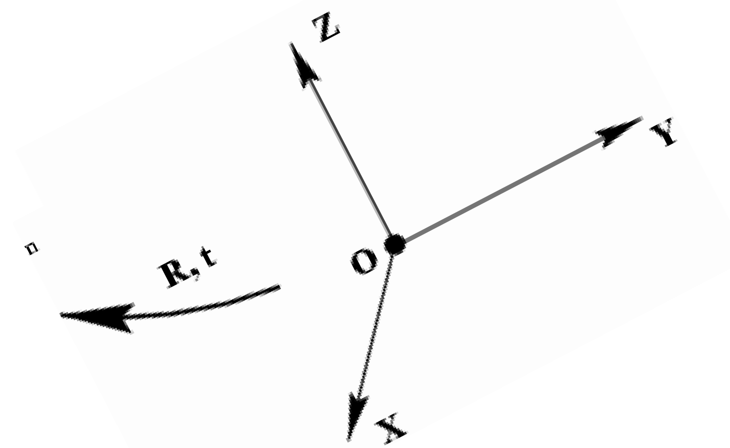
Our goal: Recovery of 3D structure

- We will need *multi-view geometry*

Review: Pinhole camera model



world coordinate system



- **Normalized (camera) coordinate system:** camera center is at the origin, the ***principal axis*** is the z-axis, x and y axes of the image plane are parallel to x and y axes of the camera

- Goal of camera calibration: go from *world* coordinate system to *image* coordinate system

Perspective Projection (pinhole projection)

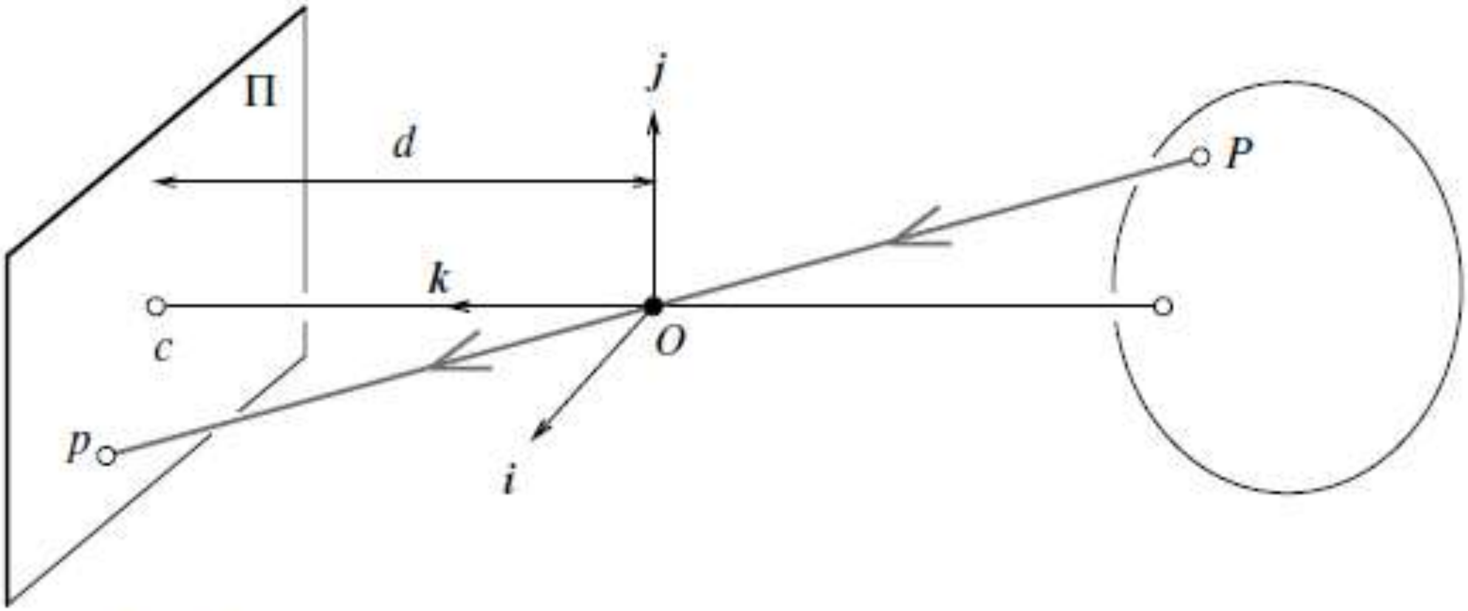


FIGURE 1.4: The perspective projection equations are derived in this section from the collinearity of the point P, its image p, and the pinhole O.

$$\begin{cases} x = \lambda X \\ y = \lambda Y \\ d = \lambda Z \end{cases} \iff \lambda = \frac{x}{X} = \frac{y}{Y} = \frac{d}{Z},$$

$$\begin{cases} x = d\dfrac{X}{Z}, \\ y = d\dfrac{Y}{Z}. \end{cases}$$

Projection Equation in Homogenous Coordinates

For a point **P** in some fixed world coordinate

P=(X, Y, Z, 1)^T, and its image **p** in the camera's reference frame (normalized image plane) *p*= (x,y,1)^T, the projection equation is represented as:

$$\boldsymbol{p} = \frac{1}{Z}\mathcal{M}\boldsymbol{P}.$$

Intrinsic Parameters

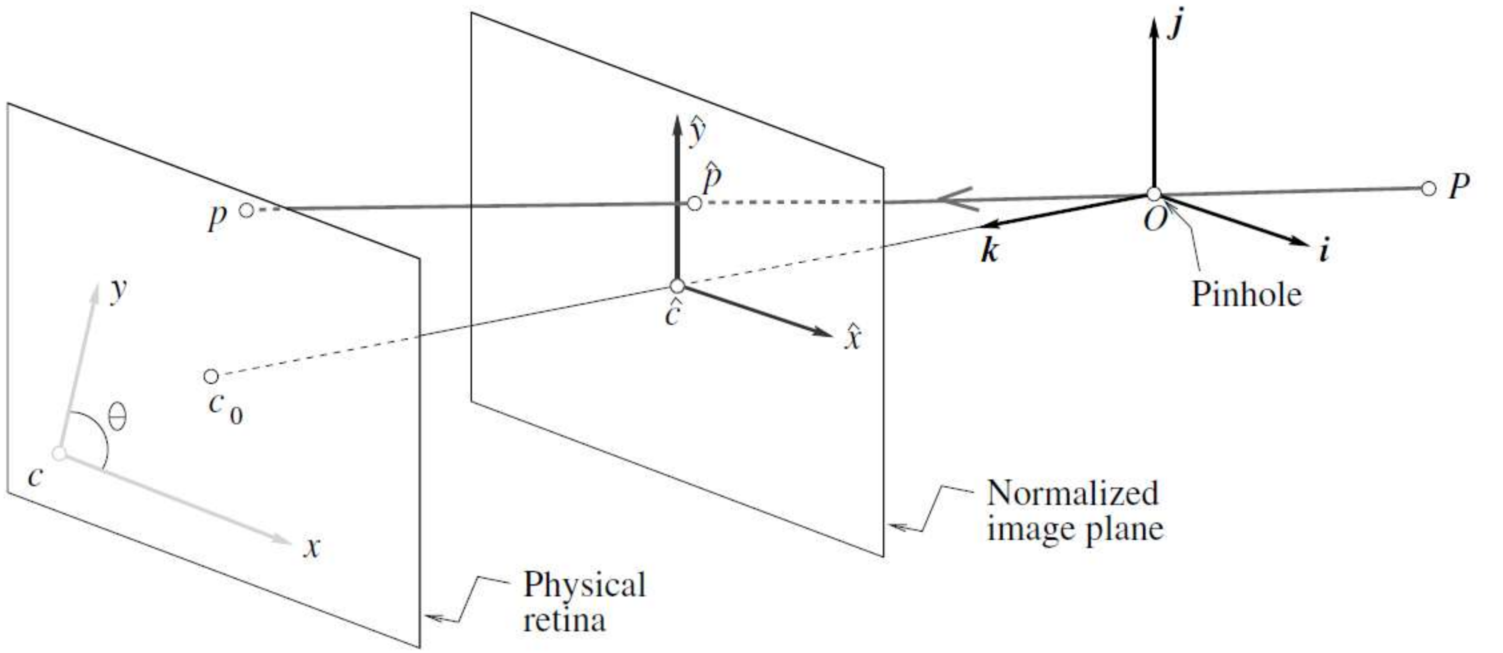


FIGURE 1.14: Physical and normalized image coordinate systems.

A point at normalized image plane

$$\begin{cases} \hat{x} = \dfrac{X}{Z} \\ \hat{y} = \dfrac{Y}{Z} \end{cases} \Longleftrightarrow \hat{\boldsymbol{p}} = \frac{1}{Z}\begin{pmatrix} \mathrm{Id} & \mathbf{0} \end{pmatrix} \boldsymbol{P}$$

Intrinsic Parameters

- The coordinates (x, y) of the image point p are expressed in pixel units (not meters).
- Pixels may be rectangular instead of square(skewed).

$$\begin{cases} x = kf\dfrac{X}{Z} = kf\hat{x}, \\ y = lf\dfrac{Y}{Z} = lf\hat{y}. \end{cases} \qquad \alpha = kf \text{ and } \beta = lf$$

- The center of the CCD matrix usually does not coincide with the image center c_0

$$\begin{cases} x = \alpha\hat{x} + x_0, \\ y = \beta\hat{y} + y_0. \end{cases}$$

- Due to manufacturing error, the angle between two image axes is not 90 degrees.

$$\begin{cases} x = \alpha\hat{x} - \alpha\cot\theta\hat{y} + x_0, \\ y = \dfrac{\beta}{\sin\theta}\hat{y} + y_0. \end{cases}$$

Intrinsic Parameters

Putting all equations together, we get

$$\boldsymbol{p} = \mathcal{K}\hat{\boldsymbol{p}}, \quad \text{where} \quad \boldsymbol{p} = \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad \text{and} \quad \mathcal{K} \stackrel{\text{def}}{=} \begin{pmatrix} \alpha & -\alpha\cot\theta & x_0 \\ 0 & \dfrac{\beta}{\sin\theta} & y_0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Here $\mathcal{K}$ is called (Internal) calibration matrix of the camera.

$$\boldsymbol{p} = \frac{1}{Z}\mathcal{K}(\text{Id} \quad \mathbf{0})\boldsymbol{P} = \frac{1}{Z}\mathcal{M}\boldsymbol{P}, \quad \text{where} \quad \mathcal{M} \stackrel{\text{def}}{=} (\mathcal{K} \quad \mathbf{0}).$$

Intrinsic parameters: $\alpha,\ \beta,\ \theta,\ x_0,$ and y_0

Non-homogenous Coordinates

A point P in some coordinate frame (F) = (O, ***i, j, k***) is represented as: $\overrightarrow{OP} = X\boldsymbol{i} + Y\boldsymbol{j} + Z\boldsymbol{k}.$

The same point P in different coordinate systems (A) and (B): ${}^{A}\boldsymbol{P} = \mathcal{R}\,{}^{B}\boldsymbol{P} + \boldsymbol{t},$

Here ***R*** is a rotation matrix, ***t*** is a translation vector.

$$\mathcal{R} \stackrel{\text{def}}{=} \left({}^{A}\boldsymbol{i}_B, {}^{A}\boldsymbol{j}_B, {}^{A}\boldsymbol{k}_B\right) = \begin{pmatrix} \boldsymbol{i}_A \cdot \boldsymbol{i}_B & \boldsymbol{j}_A \cdot \boldsymbol{i}_B & \boldsymbol{k}_A \cdot \boldsymbol{i}_B \\ \boldsymbol{i}_A \cdot \boldsymbol{j}_B & \boldsymbol{j}_A \cdot \boldsymbol{j}_B & \boldsymbol{k}_A \cdot \boldsymbol{j}_B \\ \boldsymbol{i}_A \cdot \boldsymbol{k}_B & \boldsymbol{j}_A \cdot \boldsymbol{k}_B & \boldsymbol{k}_A \cdot \boldsymbol{k}_B \end{pmatrix}$$

By introducing homogenous coordinates, we have

$${}^{A}\boldsymbol{P} = \mathcal{T}\,{}^{B}\boldsymbol{P}, \quad \text{where} \quad \mathcal{T} = \begin{pmatrix} \mathcal{R} & \boldsymbol{t} \\ \boldsymbol{0}^T & 1 \end{pmatrix},$$

Extrinsic Parameters:

Camera coordinate frame:C

$$\boldsymbol{p} = \frac{1}{Z}\mathcal{K}(\mathrm{Id} \quad \mathbf{0})\boldsymbol{P} = \frac{1}{Z}\mathcal{M}\boldsymbol{P}, \quad \text{where} \quad \mathcal{M} \stackrel{\text{def}}{=} (\mathcal{K} \quad \mathbf{0}). \qquad \boldsymbol{p} = \frac{1}{Z}\mathcal{M}^{C}\boldsymbol{P}$$

World coordinate frame:W

$${}^{C}\boldsymbol{P} = \begin{pmatrix} \mathcal{R} & \boldsymbol{t} \\ \mathbf{0}^T & 1 \end{pmatrix} {}^{W}\boldsymbol{P},$$

Taking P = ^{W}P

$$\boldsymbol{p} = \frac{1}{Z}\mathcal{M}\boldsymbol{P}, \quad \text{where} \quad \mathcal{M} = \mathcal{K}(\mathcal{R} \quad \boldsymbol{t}).$$

Extrinsic Parameters

Extrinsic Parameters: 3 independent parameters in rotation matrix ***R*** and 3 parameters in translation vector ***t.***

Camera calibration

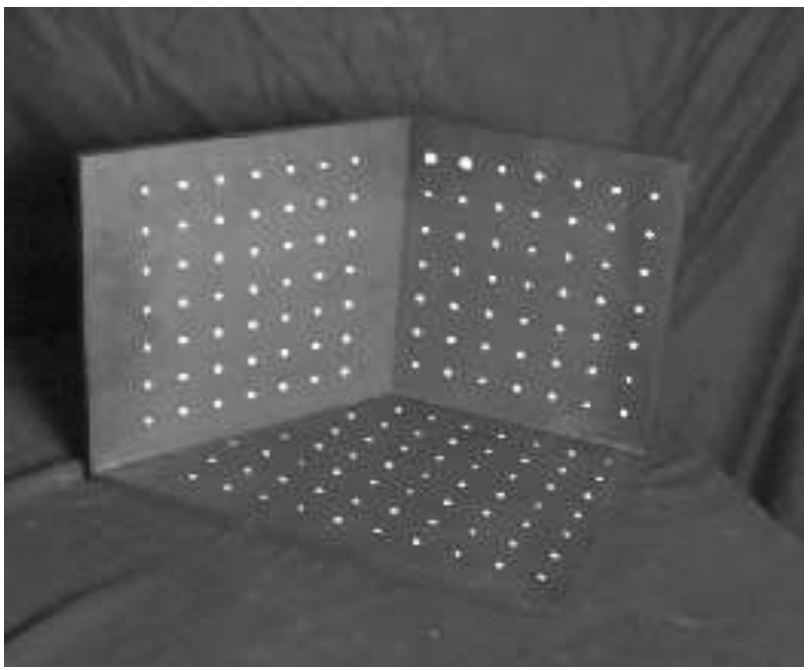
$$\mathbf{x} = \mathbf{K}\begin{bmatrix}\mathbf{R} & \mathbf{t}\end{bmatrix}\mathbf{X}$$

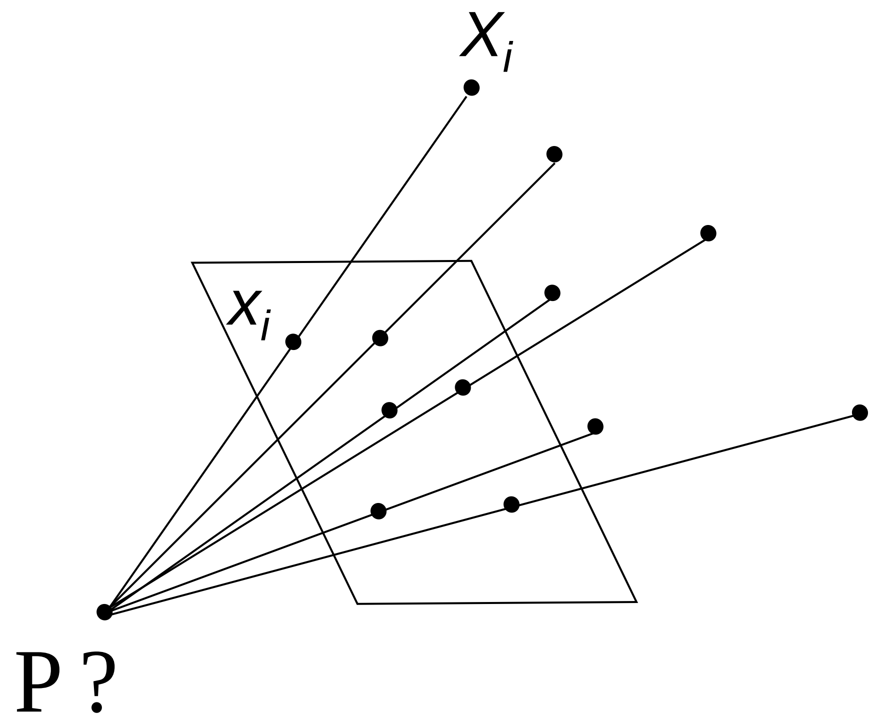
$$\begin{bmatrix}\lambda x \\ \lambda y \\ \lambda\end{bmatrix} = \begin{bmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{bmatrix}\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Source: D. Hoiem

Camera calibration

- Given *n* points with known 3D coordinates $\boldsymbol{X}_i$ and known image projections $\boldsymbol{x}_i$, estimate the camera parameters





Camera calibration: Linear method

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

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

Two linearly independent equations

Camera calibration: Linear method

$$\begin{bmatrix} 0^T & \mathbf{X}_1^T & -y_1\mathbf{X}_1^T \\ \mathbf{X}_1^T & 0^T & -x_1\mathbf{X}_1^T \\ \cdots & \cdots & \cdots \\ 0^T & \mathbf{X}_n^T & -y_n\mathbf{X}_n^T \\ \mathbf{X}_n^T & 0^T & -x_n\mathbf{X}_n^T \end{bmatrix} \begin{pmatrix} \mathbf{P}_1 \\ \mathbf{P}_2 \\ \mathbf{P}_3 \end{pmatrix} = 0 \qquad \mathbf{Ap} = \mathbf{0}$$

- P has 11 degrees of freedom
- One 2D/3D correspondence gives us two linearly independent equations
  - 6 correspondences needed for a minimal solution
- Homogeneous least squares: find $\mathbf{p}$ minimizing $||\mathbf{Ap}||^2$
  - Solution given by eigenvector of $\mathbf{A}^\mathbf{T}\mathbf{A}$ with smallest eigenvalue

Camera calibration: Linear method

$$\begin{bmatrix} 0^T & \mathbf{X}_1^T & -y_1\mathbf{X}_1^T \\ \mathbf{X}_1^T & 0^T & -x_1\mathbf{X}_1^T \\ \cdots & \cdots & \cdots \\ 0^T & \mathbf{X}_n^T & -y_n\mathbf{X}_n^T \\ \mathbf{X}_n^T & 0^T & -x_n\mathbf{X}_n^T \end{bmatrix} \begin{pmatrix} \mathbf{P}_1 \\ \mathbf{P}_2 \\ \mathbf{P}_3 \end{pmatrix} = 0 \qquad \mathbf{Ap} = \mathbf{0}$$

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

Camera calibration: Linear method

- The linear method only estimates the entries of the projection matrix:

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

- What we ultimately want is a decomposition of this matrix into the intrinsic and extrinsic parameters:

$$\mathbf{x} = \mathbf{K}\begin{bmatrix} \mathbf{R} & \mathbf{t} \end{bmatrix}\mathbf{X}$$

- State-of-the-art methods use nonlinear optimization to solve for the parameter values directly

Camera calibration: Linear method

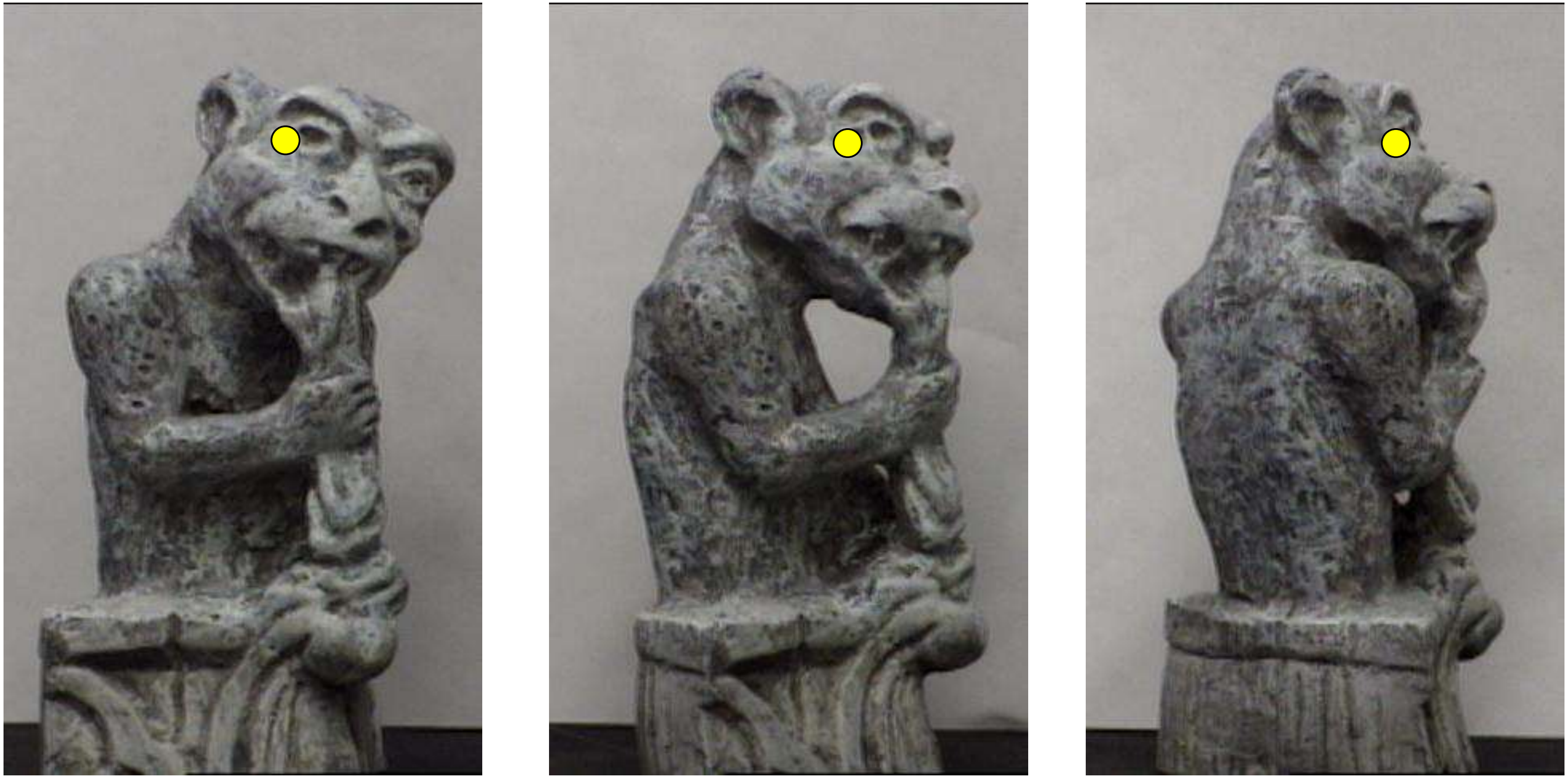
- Advantages: easy to formulate and solve
- Disadvantages
  - Doesn’t directly tell you camera parameters
  - Doesn’t model radial distortion
  - Can’t impose constraints, such as known focal length and orthogonality

- Non-linear methods are preferred
  - Define error as sum of squared distances between measured 2D points and estimated projections of 3D points
  - Minimize error using Newton’s method or other non-linear optimization

Source: D. Hoiem

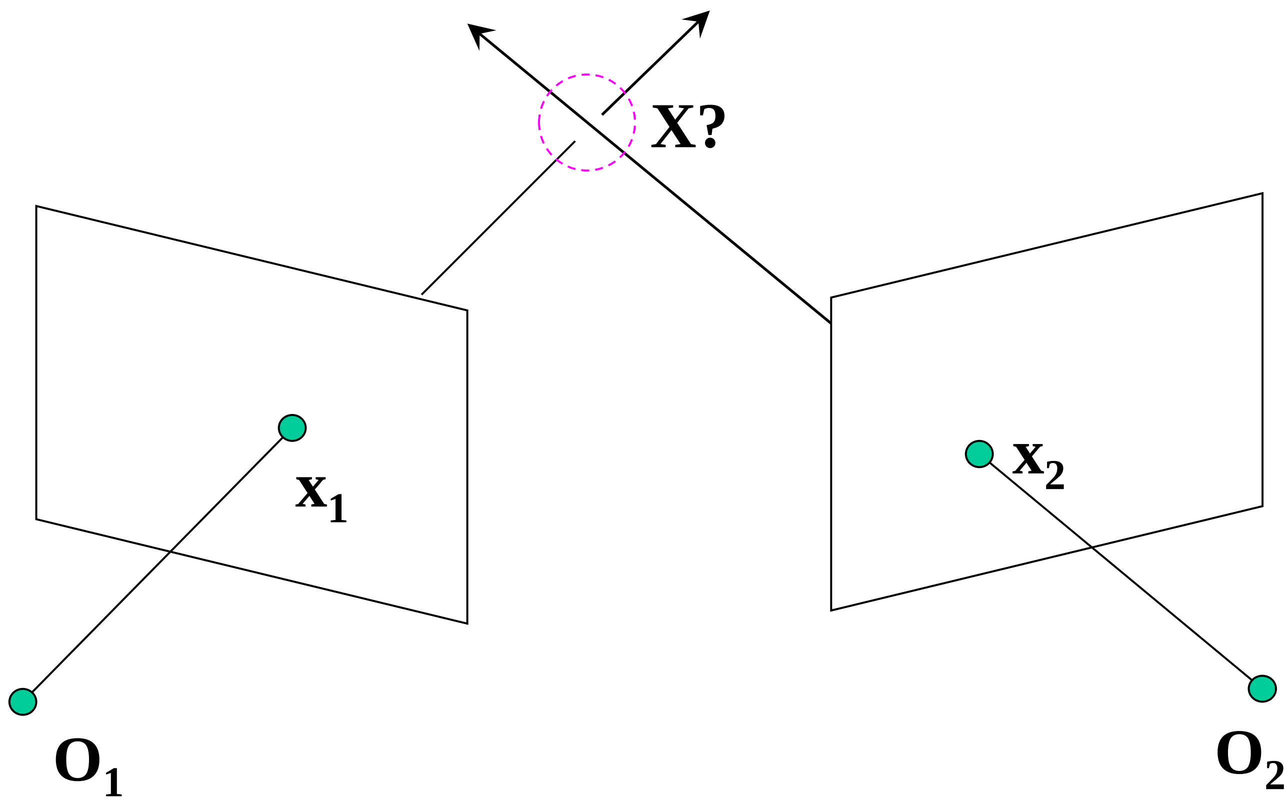
A taste of multi-view geometry: Triangulation

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



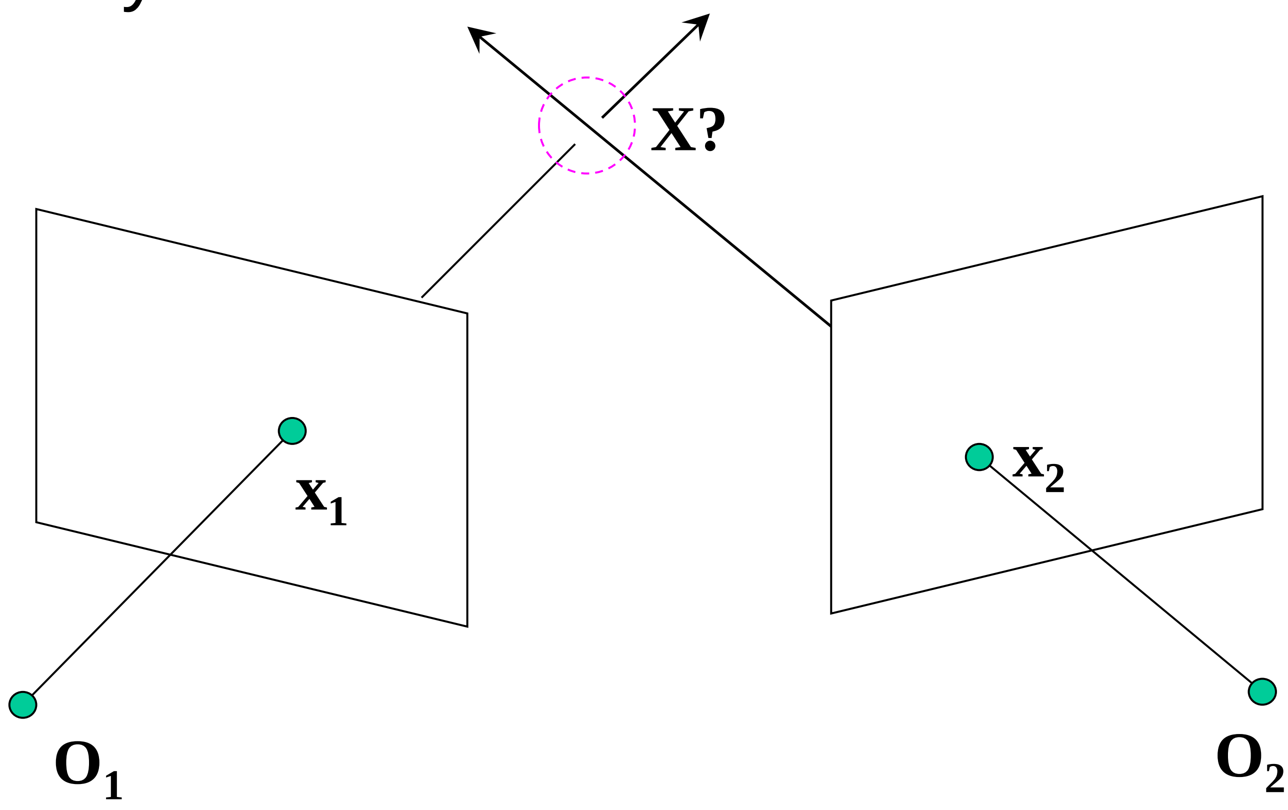
Triangulation

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



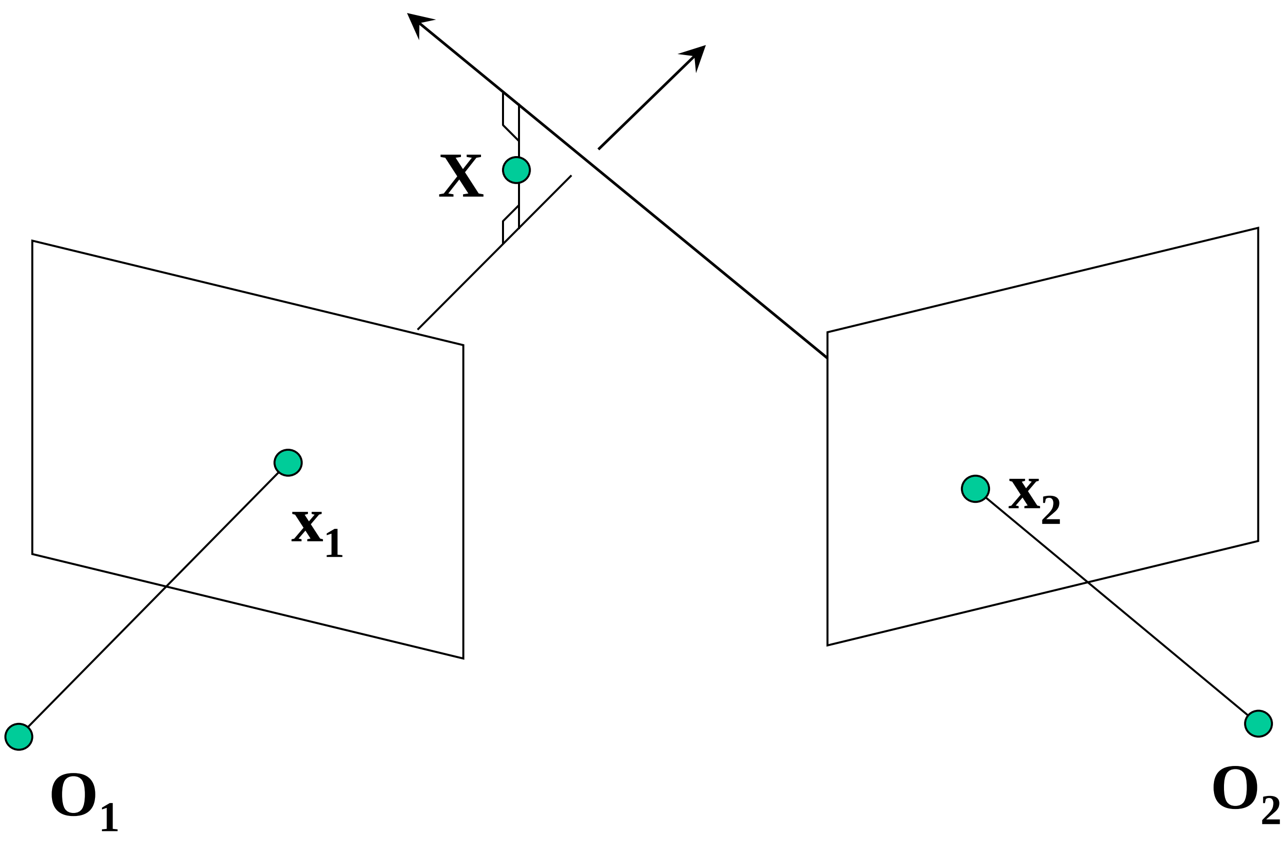
Triangulation

- We want to intersect the two visual rays corresponding to $\mathbf{x}_1$ and $\mathbf{x}_2$, but because of noise and numerical errors, they don't meet exactly



Triangulation: Geometric approach

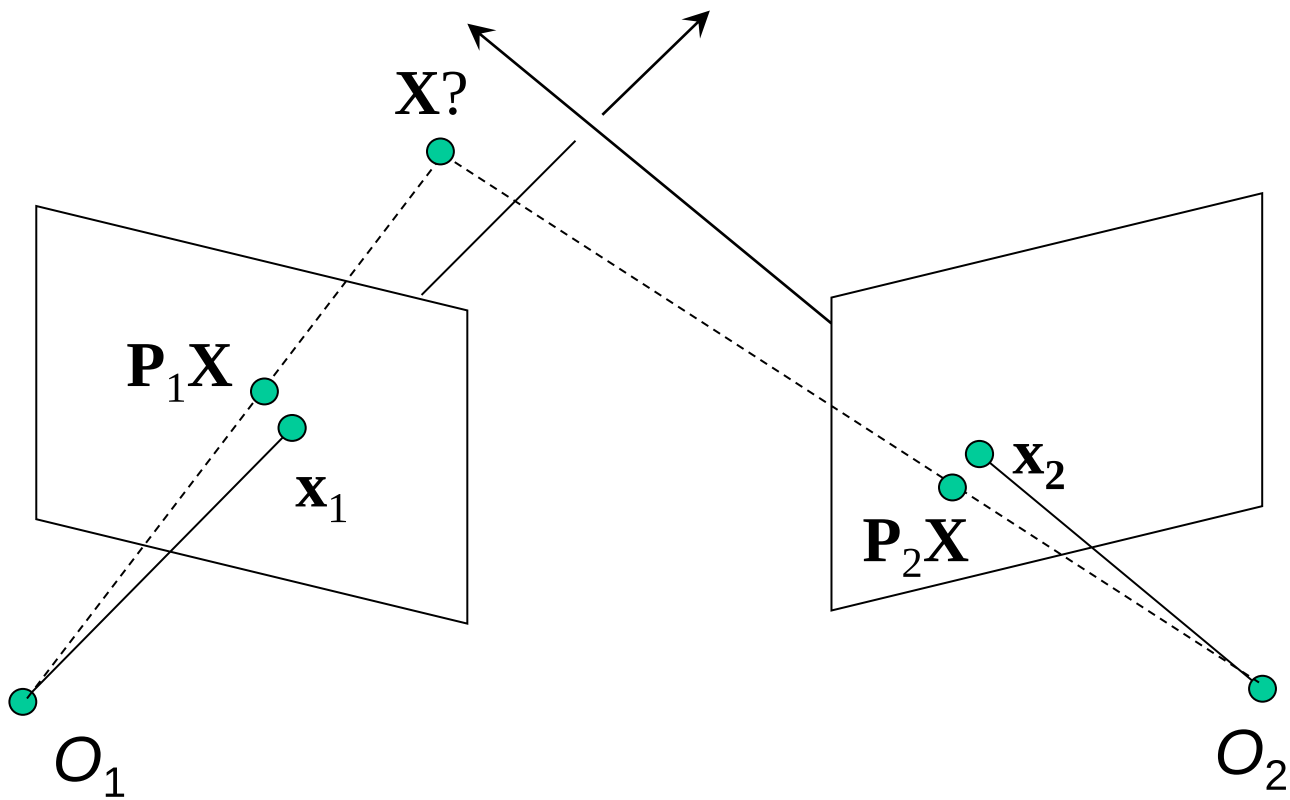
- Find shortest segment connecting the two viewing rays and let **X** be the midpoint of that segment.



Triangulation: Nonlinear approach

Find X that minimizes

$$d^2(\mathbf{x_1}, \mathbf{P_1 X}) + d^2(\mathbf{x_2}, \mathbf{P_2 X})$$



Triangulation: Linear approach

$$\lambda_1 \mathbf{x_1} = \mathbf{P_1 X} \qquad \mathbf{x_1} \times \mathbf{P_1 X} = \mathbf{0} \qquad [\mathbf{x_{1\times}}]\mathbf{P_1 X} = \mathbf{0}$$

$$\lambda_2 \mathbf{x_2} = \mathbf{P_2 X} \qquad \mathbf{x_2} \times \mathbf{P_2 X} = \mathbf{0} \qquad [\mathbf{x_{2\times}}]\mathbf{P_2 X} = \mathbf{0}$$

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

$$\lambda_1 \mathbf{x_1} = \mathbf{P_1 X} \qquad \mathbf{x_1} \times \mathbf{P_1 X} = \mathbf{0} \qquad [\mathbf{x_{1\times}}]\mathbf{P_1 X} = \mathbf{0}$$

$$\lambda_2 \mathbf{x_2} = \mathbf{P_2 X} \qquad \mathbf{x_2} \times \mathbf{P_2 X} = \mathbf{0} \qquad [\mathbf{x_{2\times}}]\mathbf{P_2 X} = \mathbf{0}$$

Two independent equations each in terms of three unknown entries of **X**