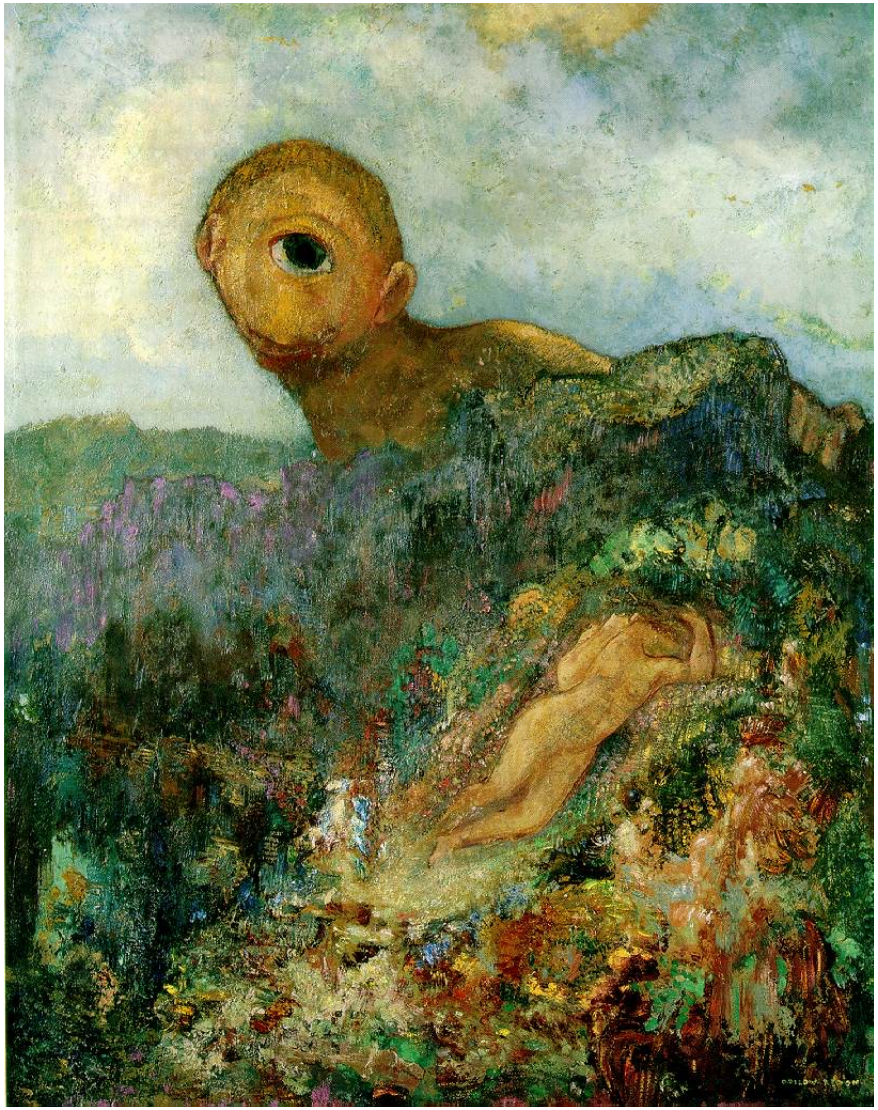
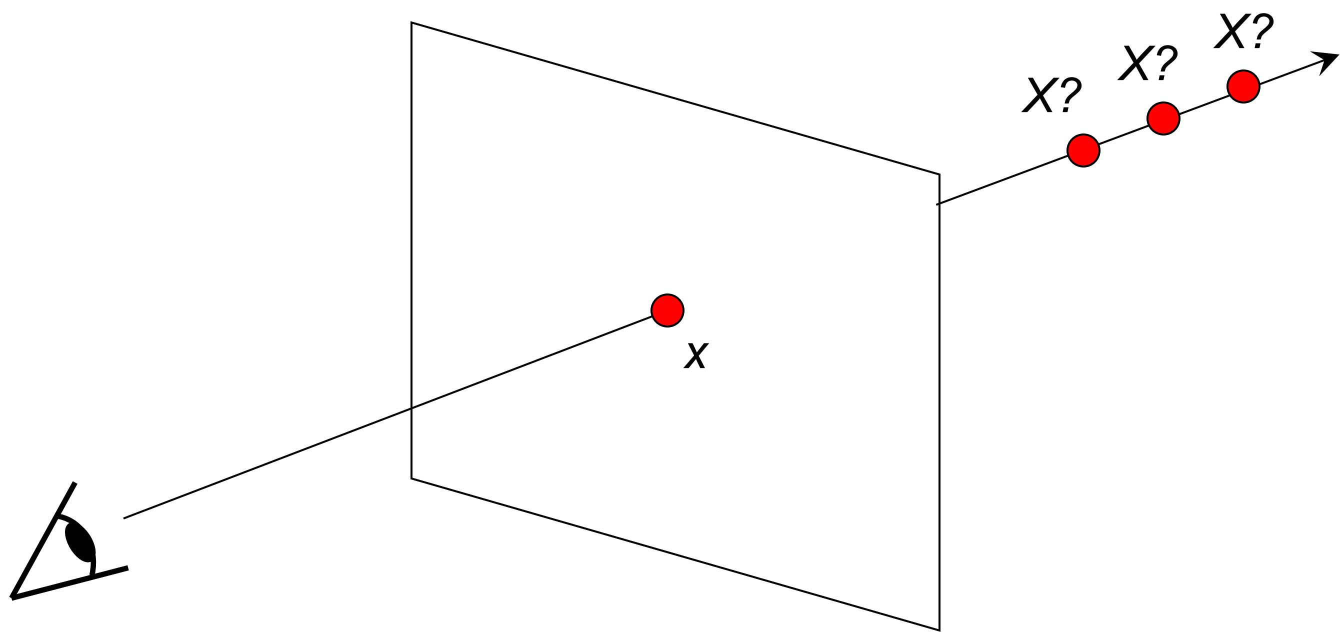
Calibrating a single camera



Odilon Redon, *Cyclops*, 1914

Our goal: Recovery of 3D structure

- Recovery of structure from one image is inherently ambiguous

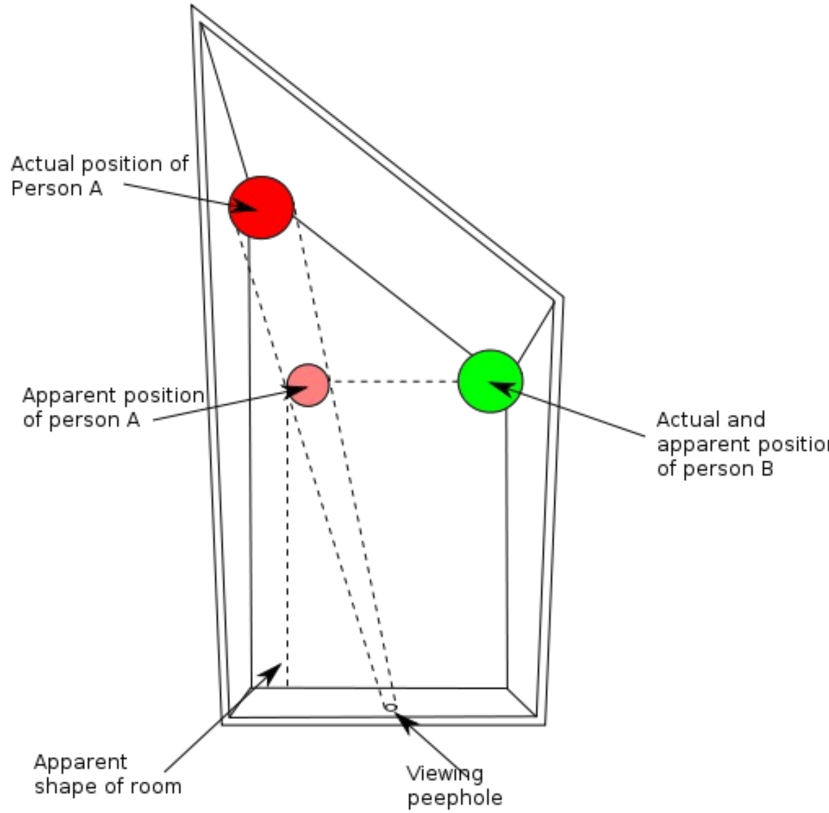


Single-view ambiguity

Single-view ambiguity

[Rashad Alakbarov shadow sculptures]

Single-view ambiguity

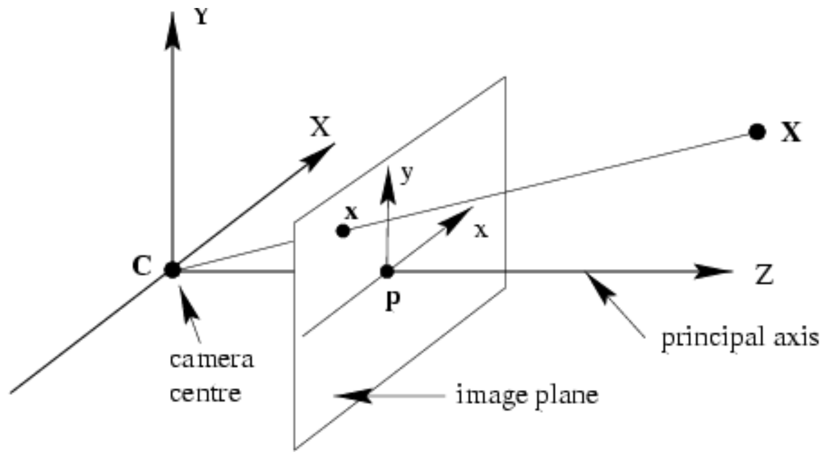


[Ames room]

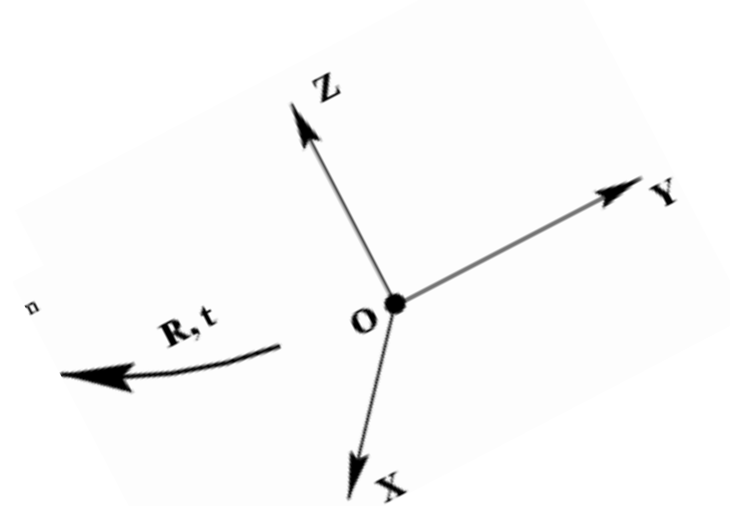
Our goal: Recovery of 3D structure

- 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)

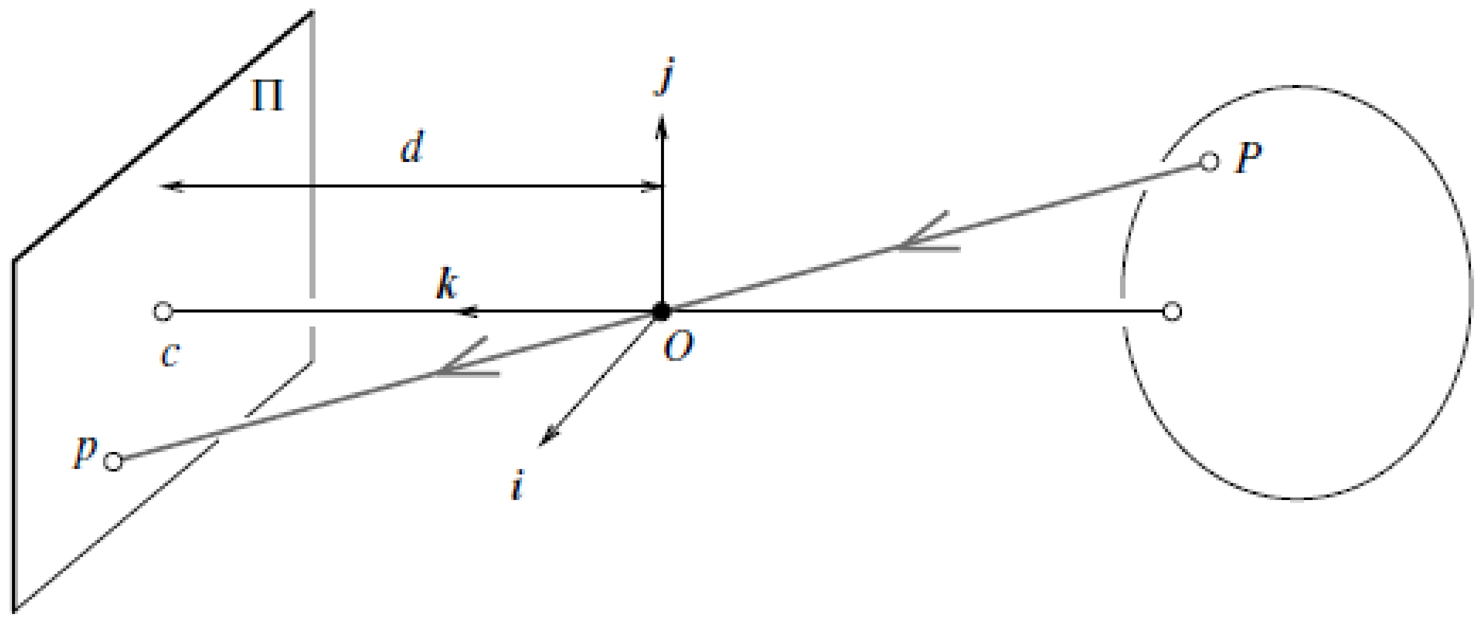


FIGURE 1.4: The perspective projection equations are derived in this section from the collinearity of the point P, its image p, and the pinhole O.

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

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

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:

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

Intrinsic Parameters

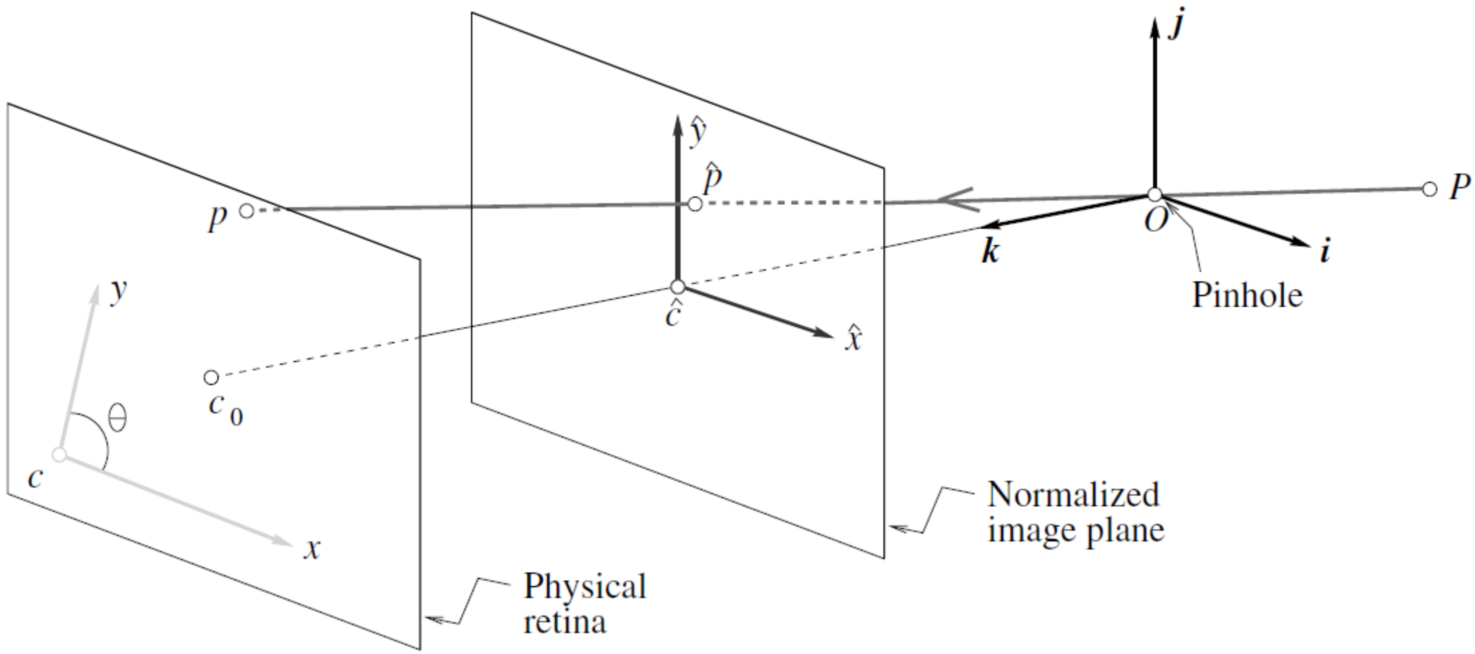


FIGURE 1.14: Physical and normalized image coordinate systems.

A point at normalized image plane

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

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.

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

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

Extrinsic Parameters:

Camera coordinate frame:C

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

$$p = \frac{1}{Z}\mathcal{M}^C P$$

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

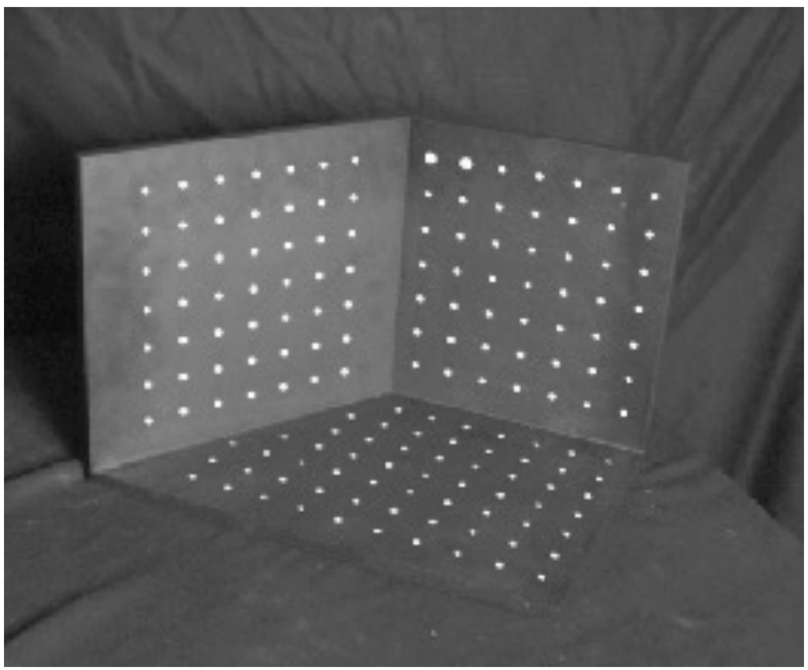
$$\mathbf{x} = \mathbf{K}\begin{bmatrix}\mathbf{R} & \mathbf{t}\end{bmatrix}\mathbf{X}$$

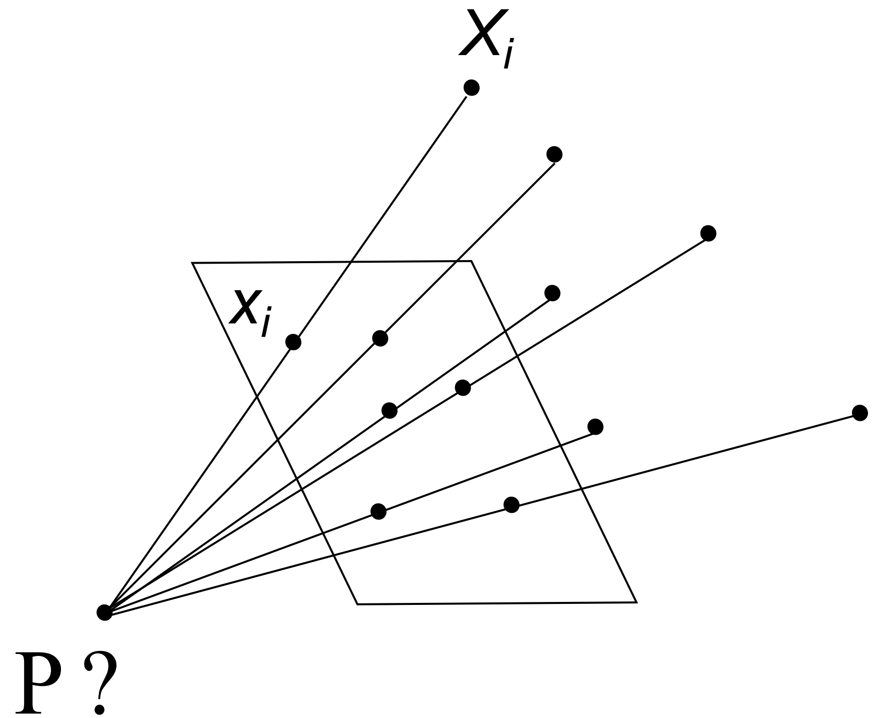
$$\begin{bmatrix}\lambda x \\ \lambda y \\ \lambda\end{bmatrix} = \begin{bmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{bmatrix}\begin{bmatrix} X \\ Y \\ Z \\ 1\end{bmatrix}$$

Source: D. Hoiem

Camera calibration

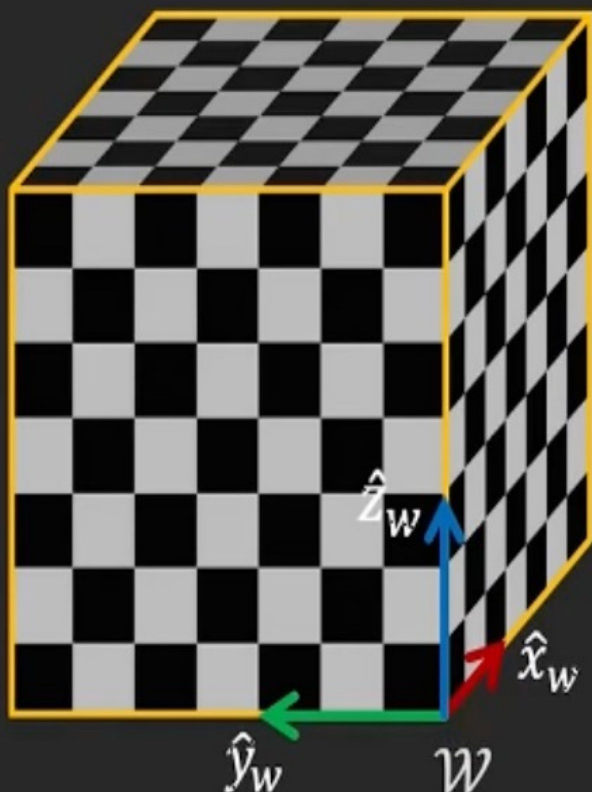
- Given n points with known 3D coordinates $\boldsymbol{X}_i$ and known image projections $\boldsymbol{x}_i$, estimate the camera parameters



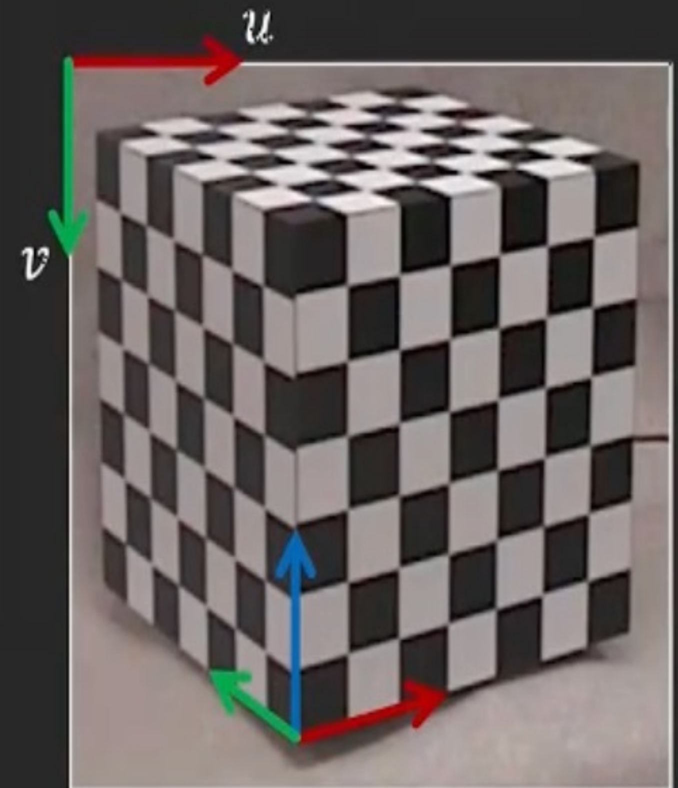


Camera Calibration Procedure

Step 1: Capture an image of an object with known geometry.



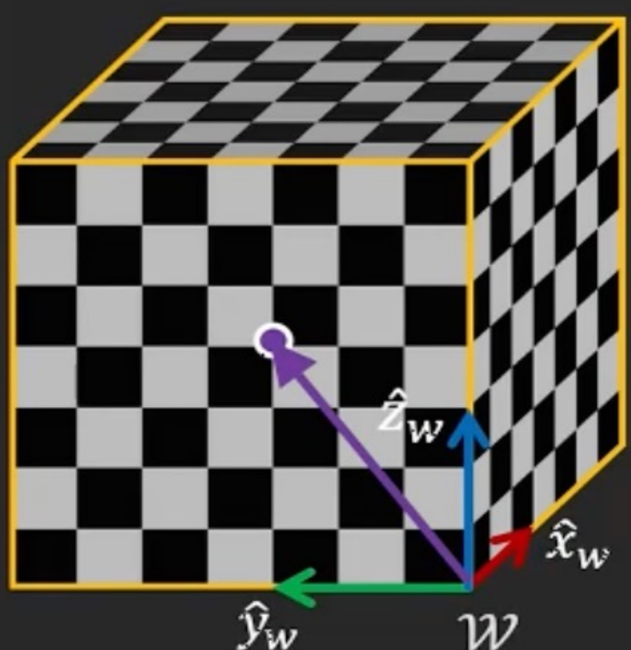
Object of Known Geometry



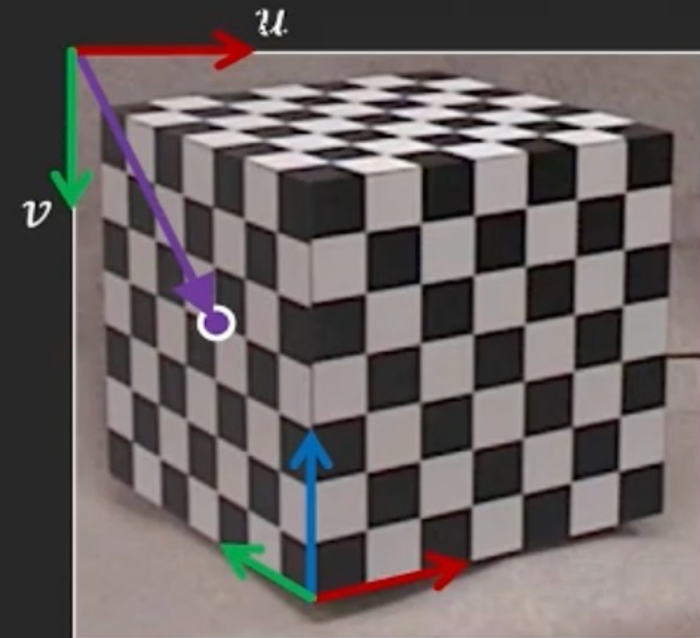
Captured Image

Camera Calibration Procedure

Step 2: Identify correspondences between 3D scene points image points.



Object of Known Geometry



Captured Image

$$\mathbf{x}_w = \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \\ 4 \end{bmatrix}$$

(inches)

$$\mathbf{u} = \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 56 \\ 115 \end{bmatrix}$$

(pixels)

Camera Calibration Procedure

Step 3: For each corresponding point i in scene and image

$$\underbrace{\begin{bmatrix} u^{(i)} \\ v^{(i)} \\ 1 \end{bmatrix}}_{\text{Known}} \equiv \underbrace{\begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix}}_{\text{Unknown}} \underbrace{\begin{bmatrix} x_w^{(i)} \\ y_w^{(i)} \\ z_w^{(i)} \\ 1 \end{bmatrix}}_{\text{Known}}$$

Expanding the matrix as linear equations:

$$u^{(i)} = \frac{p_{11}x_w^{(i)} + p_{12}y_w^{(i)} + p_{13}z_w^{(i)} + p_{14}}{p_{31}x_w^{(i)} + p_{32}y_w^{(i)} + p_{33}z_w^{(i)} + p_{34}}$$

$$v^{(i)} = \frac{p_{21}x_w^{(i)} + p_{22}y_w^{(i)} + p_{23}z_w^{(i)} + p_{24}}{p_{31}x_w^{(i)} + p_{32}y_w^{(i)} + p_{33}z_w^{(i)} + p_{34}}$$

Camera Calibration Procedure

Step 4: Rearranging the terms

$$\underbrace{\begin{bmatrix} x_w^{(1)} & y_w^{(1)} & z_w^{(1)} & 1 & 0 & 0 & 0 & 0 & -u_1 x_w^{(1)} & -u_1 y_w^{(1)} & -u_1 z_w^{(1)} & -u_1 \\ 0 & 0 & 0 & 0 & x_w^{(1)} & y_w^{(1)} & z_w^{(1)} & 1 & -v_1 x_w^{(1)} & -v_1 y_w^{(1)} & -v_1 z_w^{(1)} & -v_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_w^{(i)} & y_w^{(i)} & z_w^{(i)} & 1 & 0 & 0 & 0 & 0 & -u_i x_w^{(i)} & -u_i y_w^{(i)} & -u_i z_w^{(i)} & -u_i \\ 0 & 0 & 0 & 0 & x_w^{(i)} & y_w^{(i)} & z_w^{(i)} & 1 & -v_i x_w^{(i)} & -v_i y_w^{(i)} & -v_i z_w^{(i)} & -v_i \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_w^{(n)} & y_w^{(n)} & z_w^{(n)} & 1 & 0 & 0 & 0 & 0 & -u_n x_w^{(n)} & -u_n y_w^{(n)} & -u_n z_w^{(n)} & -u_n \\ 0 & 0 & 0 & 0 & x_w^{(n)} & y_w^{(n)} & z_w^{(n)} & 1 & -v_n x_w^{(n)} & -v_n y_w^{(n)} & -v_n z_w^{(n)} & -v_n \end{bmatrix}}_{\substack{A \\ \text{Known}}} \underbrace{\begin{bmatrix} p_{11} \\ p_{12} \\ p_{13} \\ p_{14} \\ p_{21} \\ p_{22} \\ p_{23} \\ p_{24} \\ p_{31} \\ p_{32} \\ p_{33} \\ p_{34} \end{bmatrix}}_{\substack{\mathbf{p} \\ \text{Unknown}}} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Step 5: Solve for $\mathbf{p}$

$$\boxed{A\,\mathbf{p} = \mathbf{0}}$$

Scale of Projection Matrix

Projection matrix acts on homogenous coordinates.

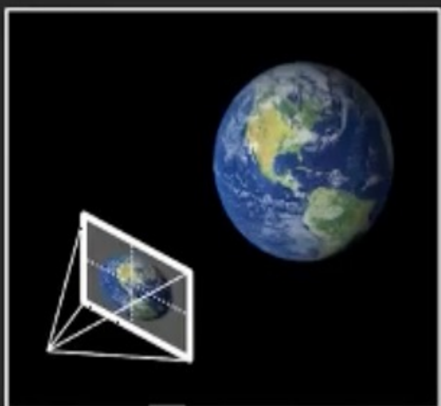
We know that: $\begin{bmatrix} \tilde{u} \\ \tilde{v} \\ \widetilde{w} \end{bmatrix} \equiv k \begin{bmatrix} \tilde{u} \\ \tilde{v} \\ \widetilde{w} \end{bmatrix}$ ($k \neq 0$ is any constant)

That is:

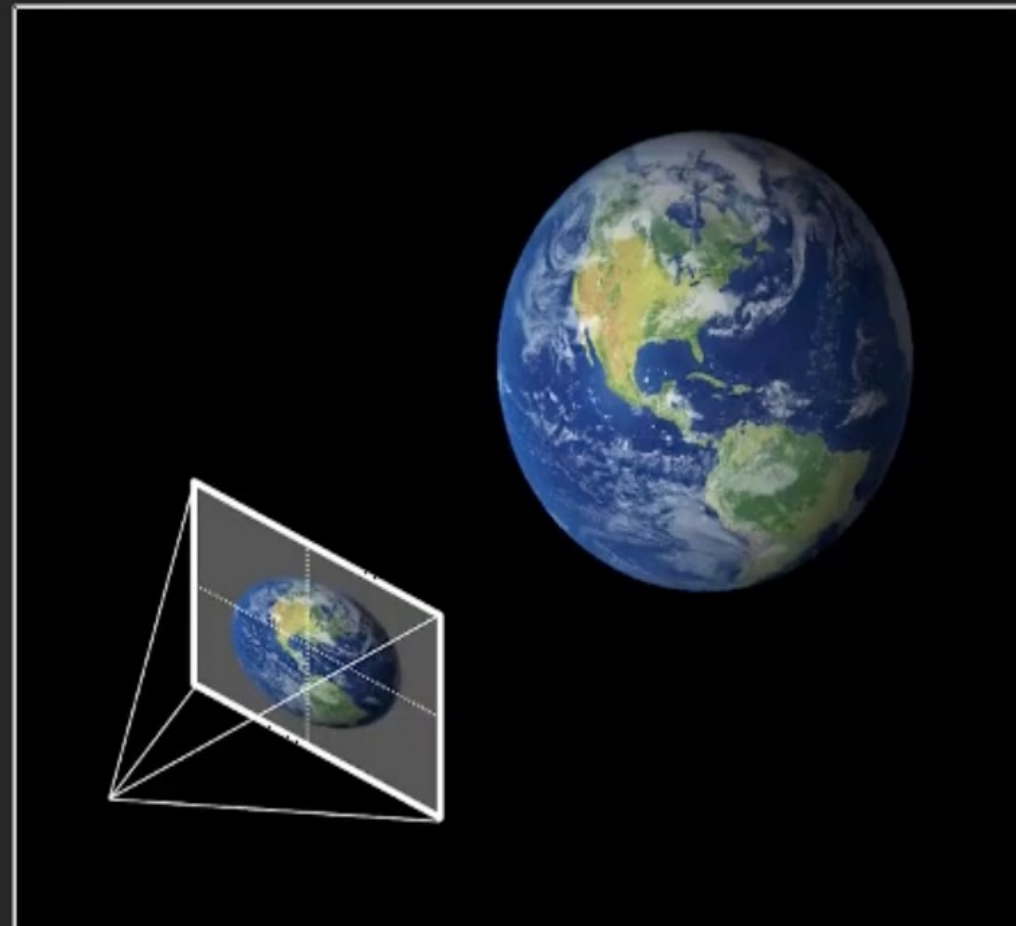
$$\begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \equiv k \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}$$

Therefore, Projection Matrices P and kP produce the same homogenous pixel coordinates.

Scale of Projection Matrix



Scale $= k_1$



Scale $= k_2$

Scaling projection matrix, implies simultaneously scaling the world and camera, which does not change the image.

Least Squares Solution for P

Option 1: Set scale so that: $p_{34} = 1$

Option 2: Set scale so that: $\boxed{\|\mathbf{p}\|^2 = 1}$

We want $A\mathbf{p}$ as close to 0 as possible and $\|\mathbf{p}\|^2 = 1$:

$$\min_{\mathbf{p}} \|A\mathbf{p}\|^2 \text{ such that } \|\mathbf{p}\|^2 = 1$$

$$\min_{\mathbf{p}} (\mathbf{p}^T A^T A \mathbf{p}) \text{ such that } \mathbf{p}^T \mathbf{p} = 1$$

Define Loss function $L(\mathbf{p}, \lambda)$:

$$L(\mathbf{p}, \lambda) = \mathbf{p}^T A^T A \mathbf{p} - \lambda(\mathbf{p}^T \mathbf{p} - 1)$$

(Similar to Solving Homography in Image Stitching)

Constrained Least Squares Solution

Taking derivatives of $L(\mathbf{p}, \lambda)$ w.r.t $\mathbf{p}$: $2A^T A\mathbf{p} - 2\lambda\mathbf{p} = \mathbf{0}$

$$\boxed{A^T A\mathbf{p} = \lambda\mathbf{p}}$$ Eigenvalue Problem

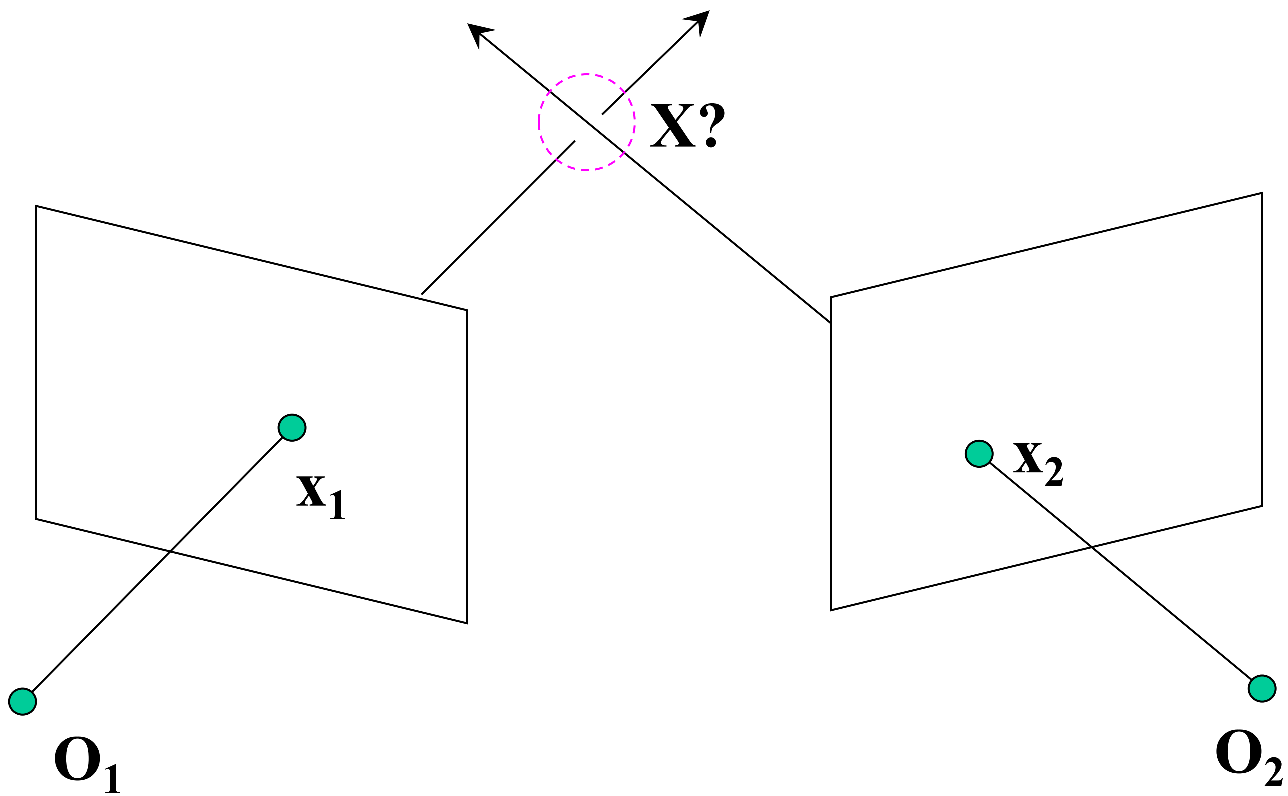
Eigenvector $\mathbf{p}$ with smallest eigenvalue λ of matrix $A^T A$ minimizes the loss function $L(\mathbf{p})$.

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

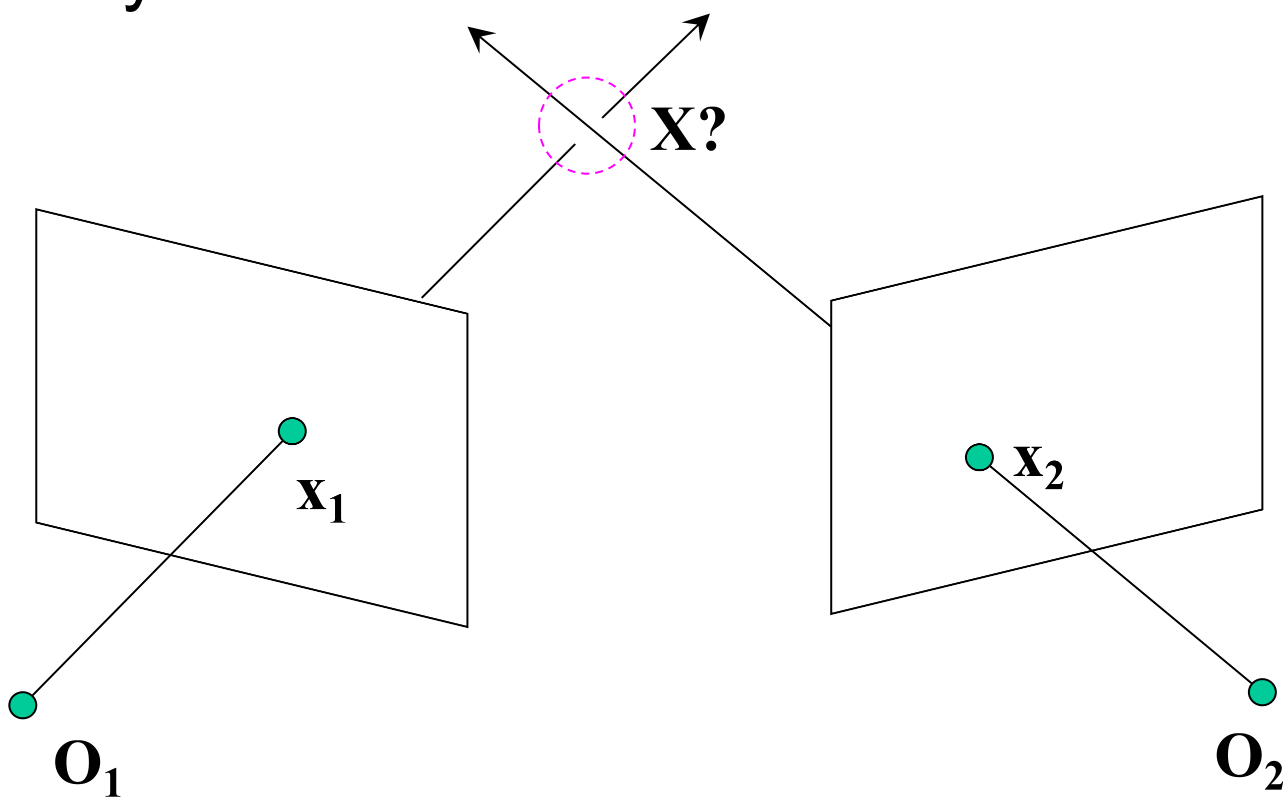
Triangulation

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



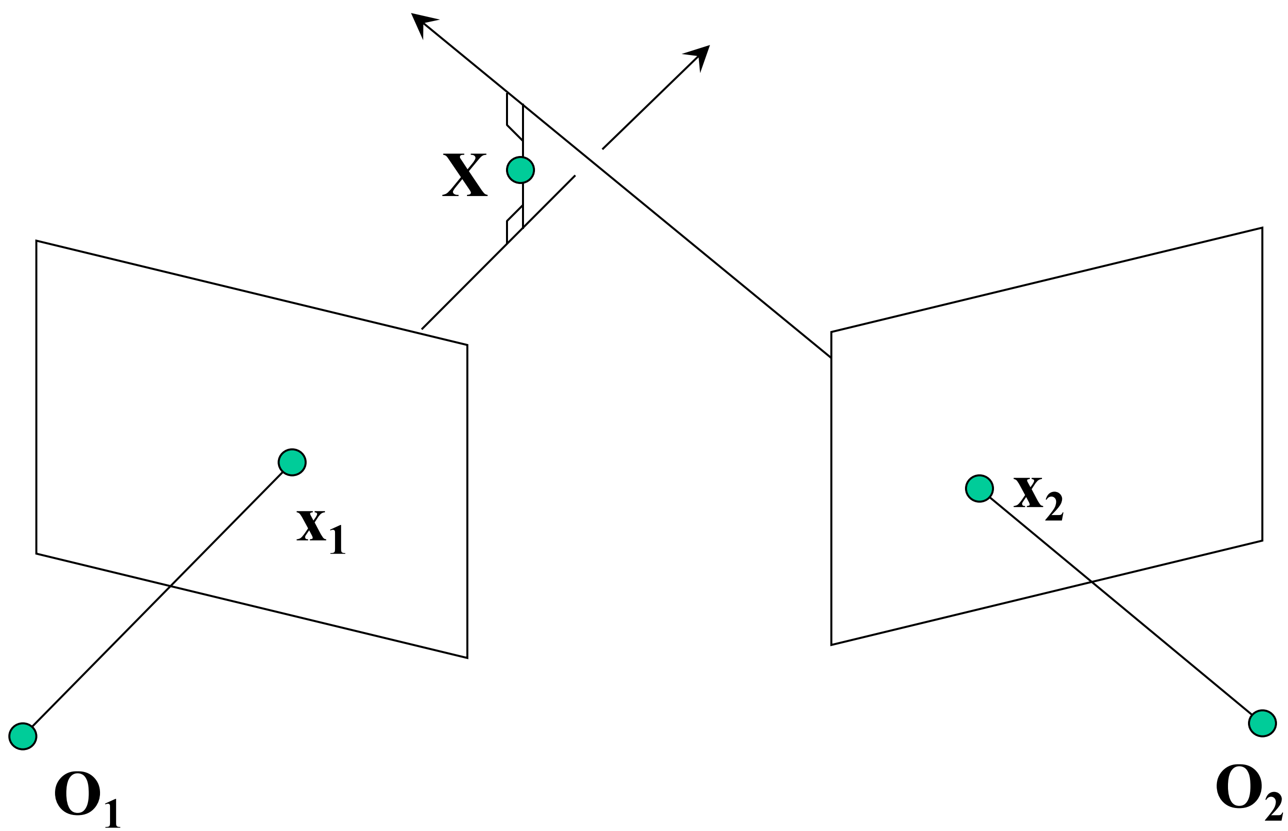
Triangulation

- 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

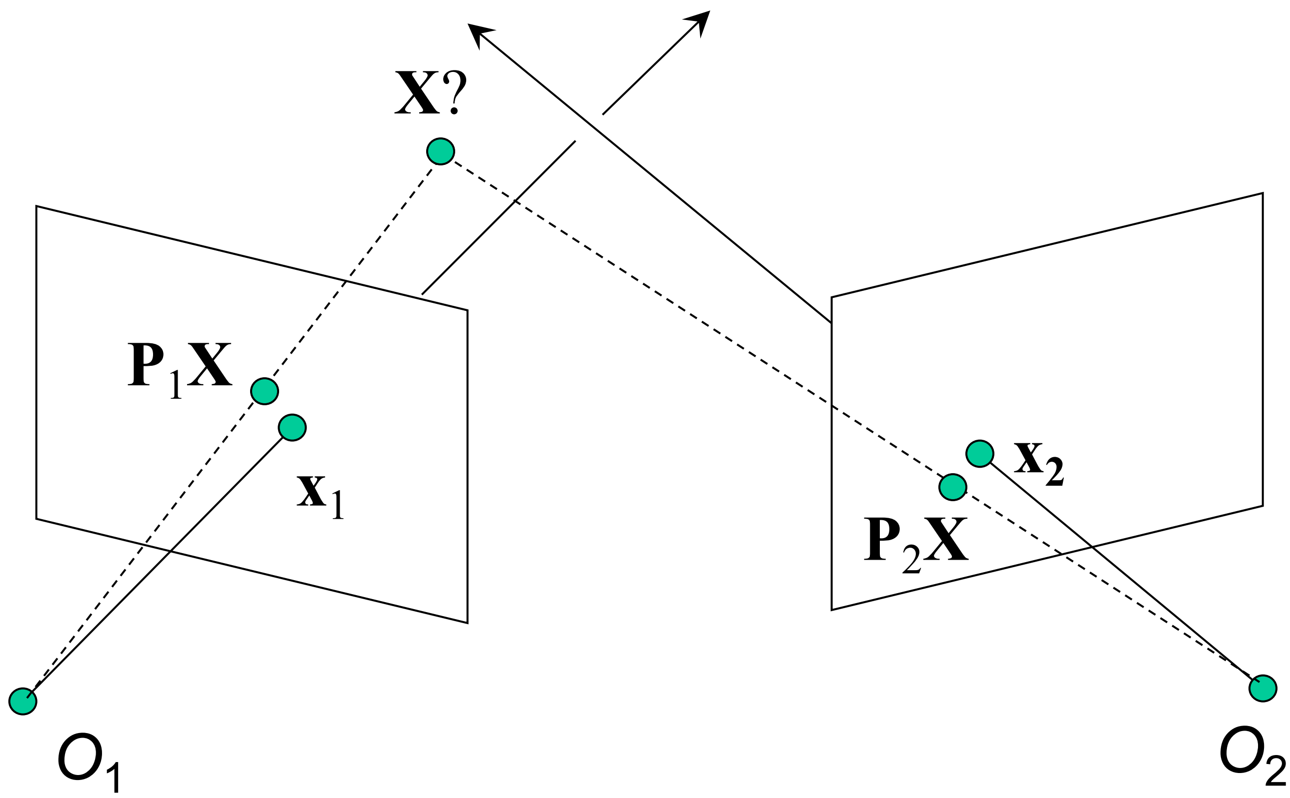
- Find shortest segment connecting the two viewing rays and let **X** be the midpoint of that segment



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**