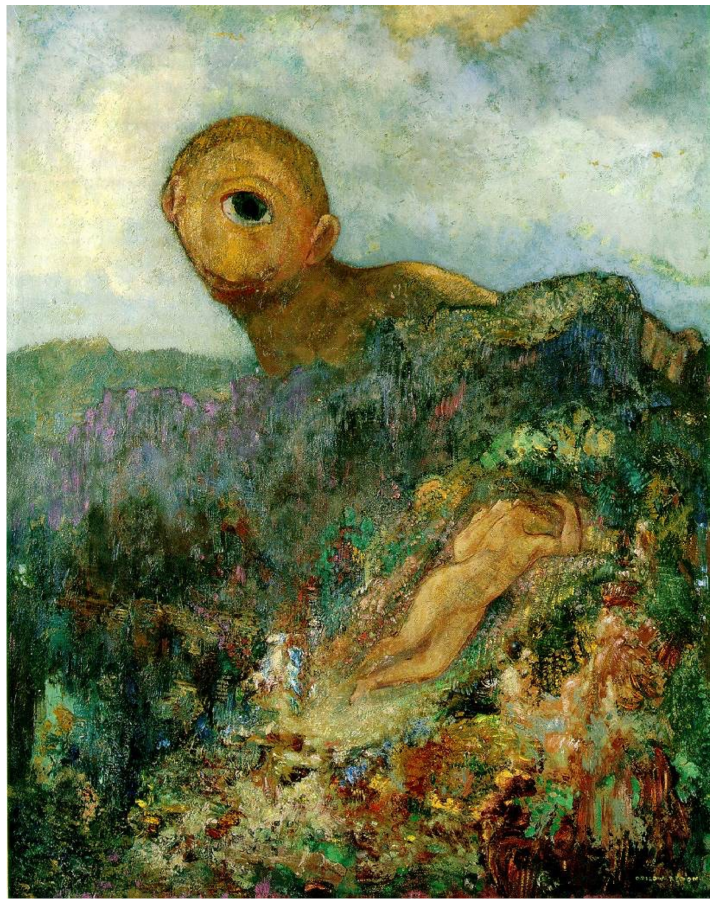
Camera calibration

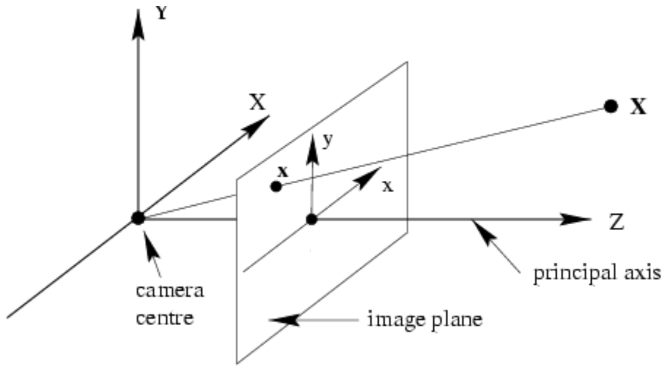


Odilon Redon, *Cyclops*, 1914

Overview

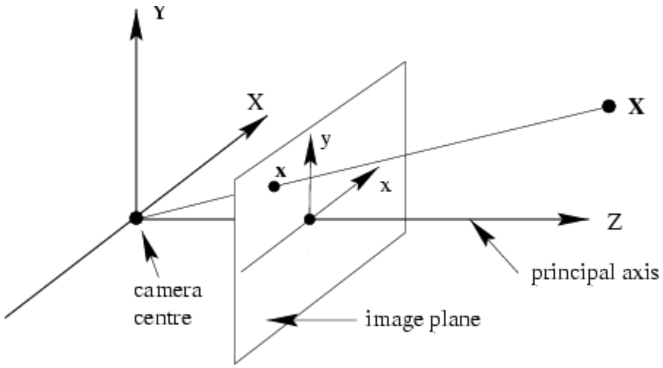
- Camera calibration
  - Intrinsic camera parameters
  - Extrinsic camera parameters
  - Estimation
- First taste of 3D reconstruction: triangulation

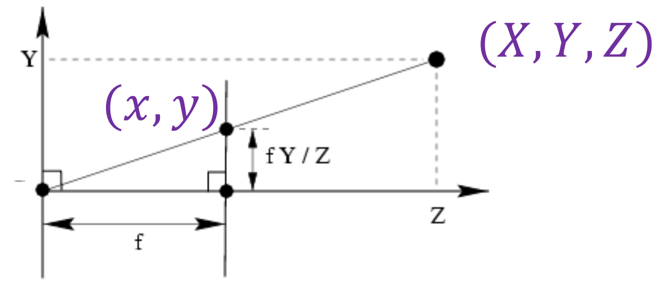
Perspective projection in normalized coordinates



- **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 world

Perspective projection in normalized coordinates





$$x = f\frac{X}{Z}, y = f\frac{Y}{Z}$$

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \cong \begin{pmatrix} fX \\ fY \\ Z \end{pmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

$$\boldsymbol{x} \cong \boldsymbol{P}\boldsymbol{X}$$

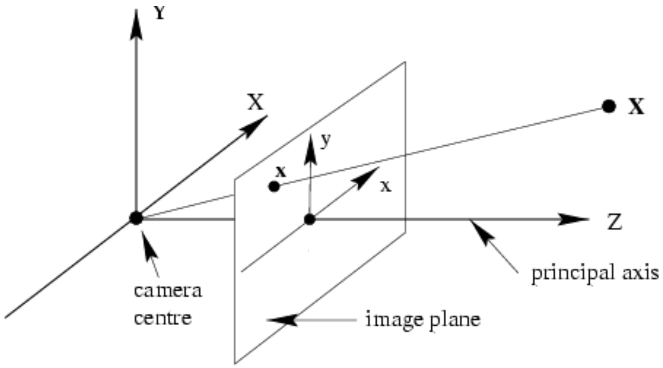
Equality up to scale

$\boldsymbol{x}$: Homogeneous coord. vec. of image point

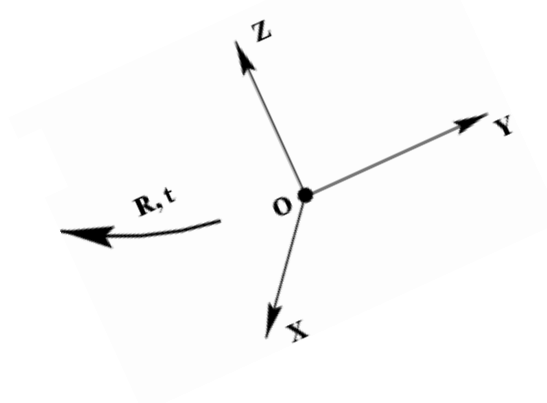
$\boldsymbol{P}$: Camera projection matrix

$\boldsymbol{X}$: Homogeneous coord. vec. of 3D point

Camera calibration: Overview



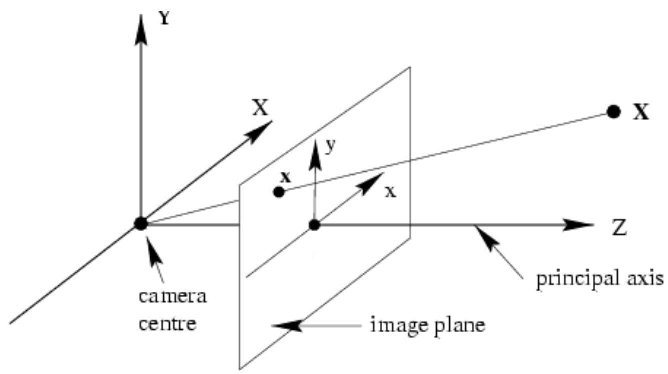
world coordinate system

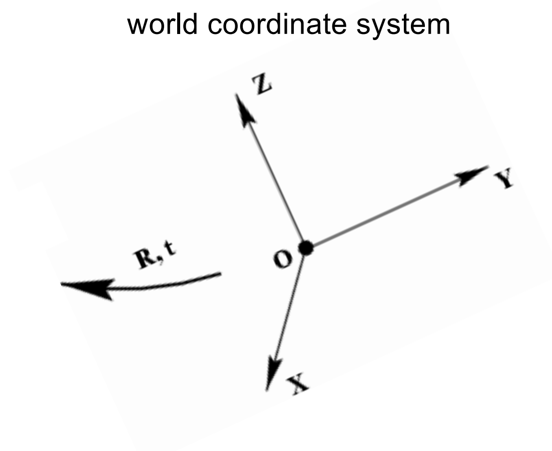


- **Camera calibration:** figuring out transformation from *world* coordinate system to *image* coordinate system

2D point (3x1)	$\cong$	Camera to pixel coord. trans. matrix $\boldsymbol{K}$ (3x3)	*Canonical* projection matrix $[\boldsymbol{I} \mid \boldsymbol{0}]$ (3x4)	World to camera coord. trans. matrix $\begin{bmatrix} \boldsymbol{R} & \boldsymbol{t} \\ \boldsymbol{0}^T & 1 \end{bmatrix}$ (4x4)	3D point (4x1)
$\boldsymbol{x}$		***Intrinsic* camera parameters**: principal point, scaling factors		***Extrinsic* camera parameters**: rotation, translation	$\boldsymbol{X}$

Camera calibration: Overview





- **Camera calibration:** figuring out transformation from *world* coordinate system to *image* coordinate system

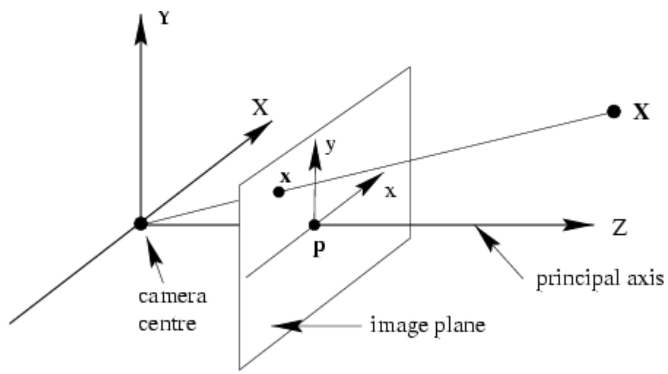
(2D point (3x1)) ≅ (Camera to pixel coord. trans. matrix $\boldsymbol{K}$ (3x3)) (*Canonical* projection matrix $[\boldsymbol{I} \mid \boldsymbol{0}]$ (3x4)) (World to camera coord. trans. matrix $\begin{bmatrix} \boldsymbol{R} & \boldsymbol{t} \\ \boldsymbol{0}^T & 1 \end{bmatrix}$ (4x4)) (3D point (4x1))

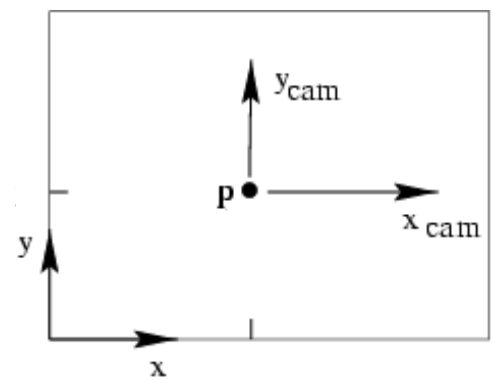
$\boldsymbol{x}$ — $\boldsymbol{X}$

$$\boldsymbol{P} = \boldsymbol{K}[\boldsymbol{R}|\boldsymbol{t}]$$

General camera projection matrix

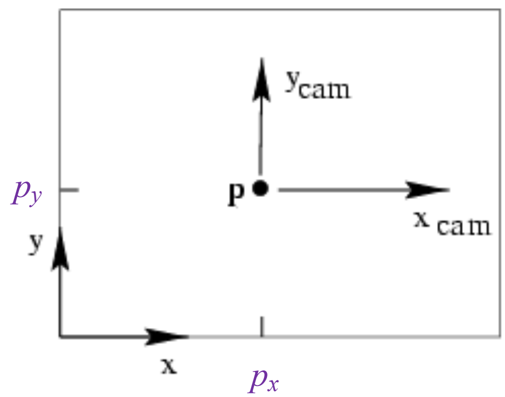
Intrinsic parameters: Principal point





- **Principal point (p):** point where principal axis intersects the image plane
- In the *normalized* coordinate system, the origin of the image is at the principal point
- In the *image* coordinate system: the origin is in the corner

Intrinsic parameters: Principal point

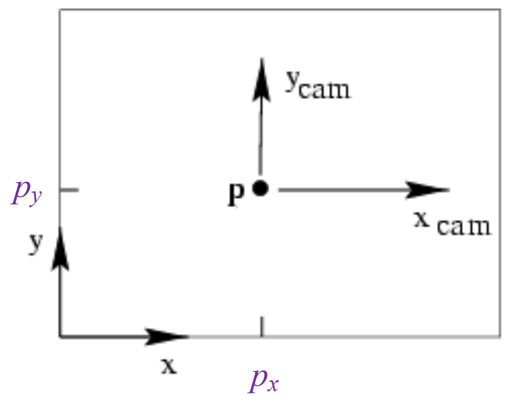


We want the principal point to map to (p_x, p_y) instead of $(0,0)$

$$x = f\frac{X}{Z} + p_x, \qquad y = f\frac{Y}{Z} + p_y$$

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \cong \begin{pmatrix} fX + Zp_x \\ fY + Zp_y \\ Z \end{pmatrix} = \left[\qquad \qquad \qquad \right] \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

Intrinsic parameters: Principal point



Principal point: (p_x, p_y)

$$\begin{bmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} f & 0 & p_x & 0 \\ 0 & f & p_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

calibration matrix $\boldsymbol{K}$

Canonical projection matrix $[\boldsymbol{I} \mid \boldsymbol{0}]$

$$\boldsymbol{P} = \boldsymbol{K}[\boldsymbol{I}|\boldsymbol{0}]$$

Intrinsic parameters: Principal point

- What are the units of the focal length f and principal point coordinates (p_x, p_y)?
  - Same as world units – presumably metric units
- What units do we want for measuring image coordinates?
  - Pixel units
- Thus, we need to introduce *scaling factors* for mapping from world to pixel units

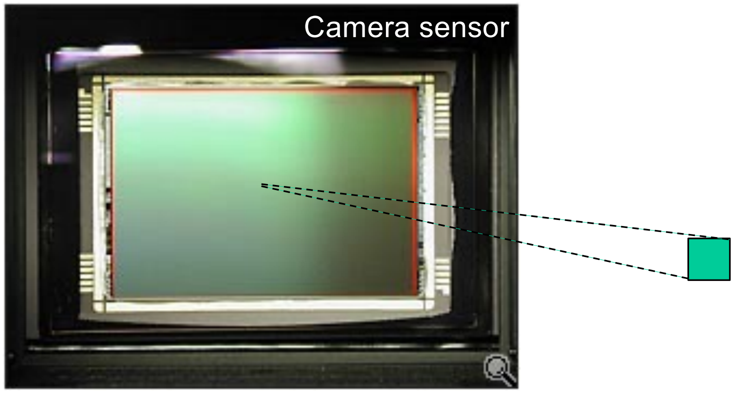
$$\begin{bmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} f & 0 & p_x & 0 \\ 0 & f & p_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

calibration matrix $\boldsymbol{K}$

Canonical projection matrix $[\boldsymbol{I} \mid \boldsymbol{0}]$

$\boldsymbol{P} = \boldsymbol{K}[\boldsymbol{I}|\boldsymbol{0}]$

Intrinsic parameters: Scaling factors



m_x pixels/m in horizontal direction,
m_y pixels/m in vertical direction

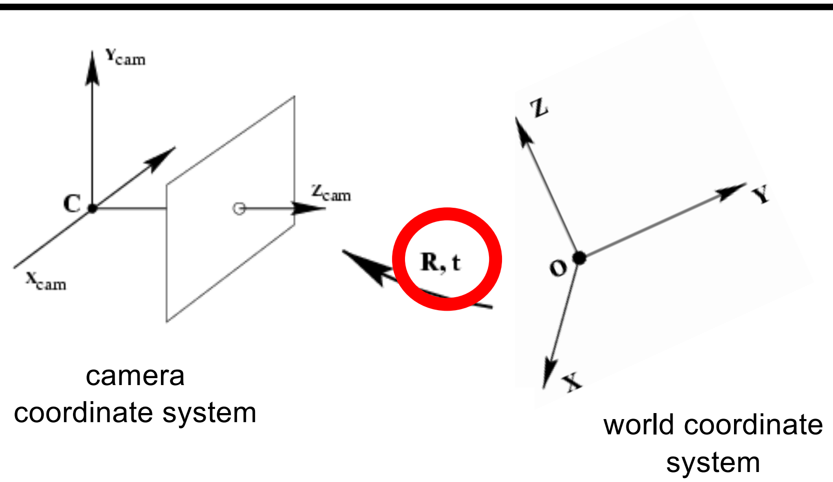
Pixel size (m): $\frac{1}{m_x} \times \frac{1}{m_y}$

Scaling factors · Calibration matrix $\boldsymbol{K}$ in metric units · Calibration matrix $\boldsymbol{K}$ in pixel units

$$\begin{bmatrix} m_x & 0 & 0 \\ 0 & m_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \alpha_x & 0 & \beta_x \\ 0 & \alpha_y & \beta_y \\ 0 & 0 & 1 \end{bmatrix}$$

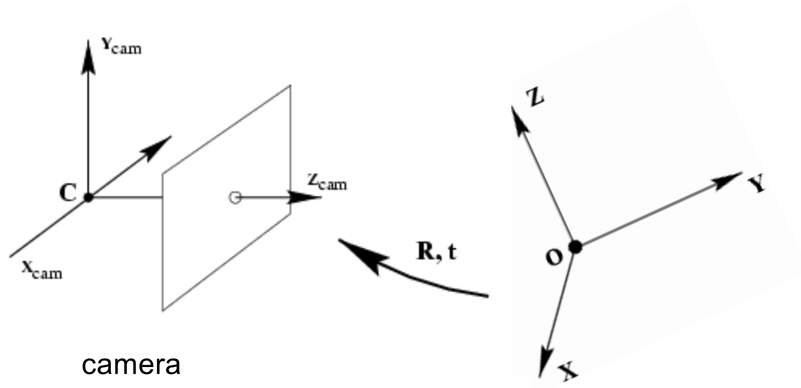
pixels/m · m · pixels

Extrinsic parameters: Rotation and translation



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

Extrinsic parameters: Rotation and translation



camera coordinate system

world coordinate system

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

- In non-homogeneous coordinates, the transformation from world to normalized camera coordinate system is given by:

$$\widetilde{\boldsymbol{X}}_{\text{cam}} = \boldsymbol{R}\left(\widetilde{\boldsymbol{X}} - \widetilde{\boldsymbol{C}}\right) = \boldsymbol{R}\widetilde{\boldsymbol{X}} + \boldsymbol{t}$$

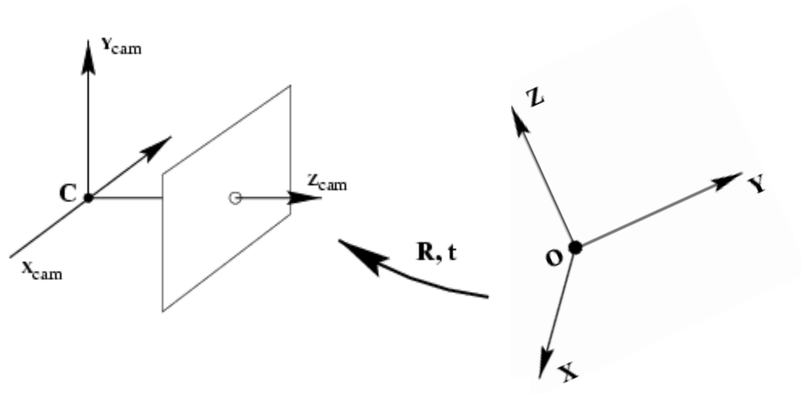
coords. of point in normalized camera frame

3x3 rotation matrix

coords. of a point in world frame

coords. of camera center in world frame

Extrinsic parameters: Rotation and translation



In *non-homogeneous* coordinates:

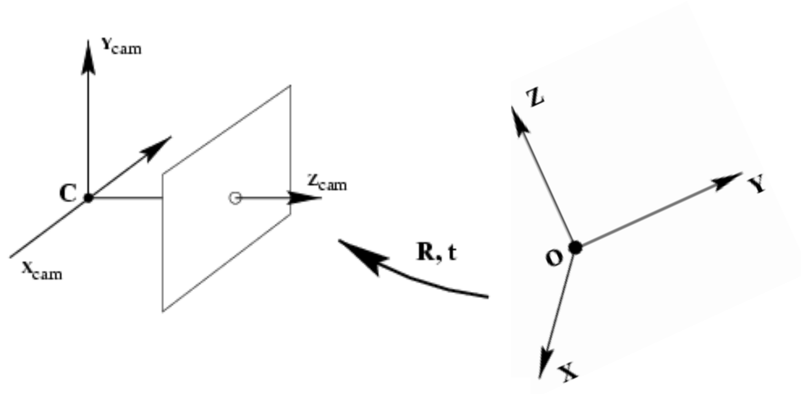
$$\widetilde{\boldsymbol{X}}_{\text{cam}} = \boldsymbol{R}\widetilde{\boldsymbol{X}} + \boldsymbol{t}$$

In *homogeneous* coordinates:

$$\begin{pmatrix} \widetilde{\boldsymbol{X}}_{\text{cam}} \\ 1 \end{pmatrix} = \begin{bmatrix} \boldsymbol{R} & \boldsymbol{t} \\ \boldsymbol{0}^{T} & 1 \end{bmatrix} \begin{pmatrix} \widetilde{\boldsymbol{X}} \\ 1 \end{pmatrix}$$

3D transformation matrix (4 x 4)

Extrinsic parameters: Rotation and translation



In *non-homogeneous* coordinates:

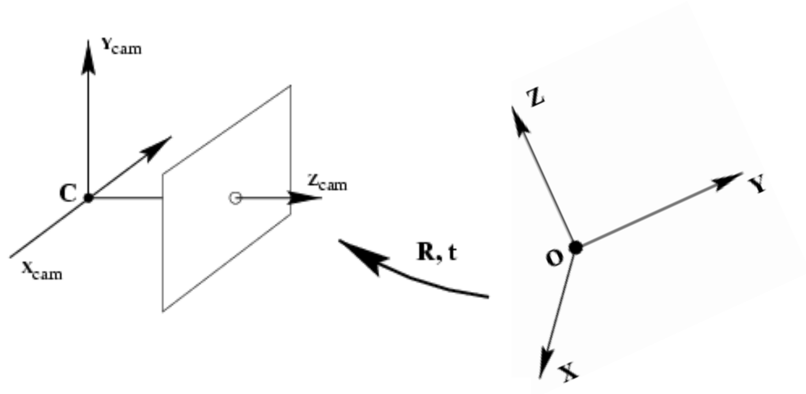
$$\widetilde{\boldsymbol{X}}_{\text{cam}} = \boldsymbol{R}\widetilde{\boldsymbol{X}} + \boldsymbol{t}$$

In *homogeneous* coordinates:

$$\boldsymbol{X}_{\text{cam}} = \begin{bmatrix} \boldsymbol{R} & \boldsymbol{t} \\ \boldsymbol{0}^T & 1 \end{bmatrix} \boldsymbol{X}$$

3D transformation matrix (4 x 4)

Extrinsic parameters: Rotation and translation



In *non-homogeneous* coordinates:

$$\widetilde{\boldsymbol{X}}_{\text{cam}} = \boldsymbol{R}\widetilde{\boldsymbol{X}} + \boldsymbol{t}$$

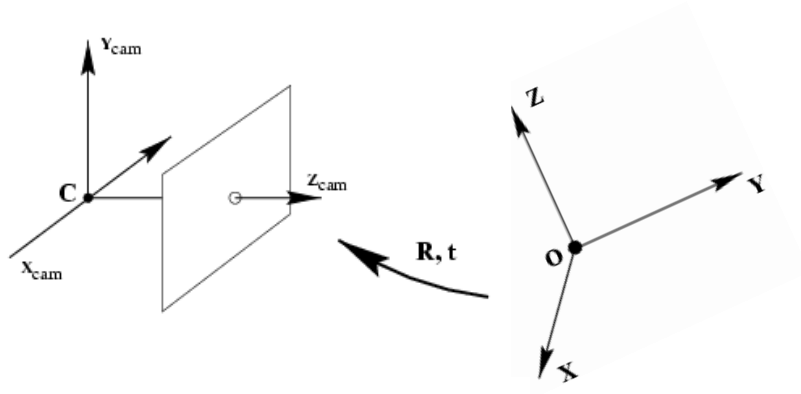
In *homogeneous* coordinates:

$$\boldsymbol{X}_{\text{cam}} = \begin{bmatrix} \boldsymbol{R} & \boldsymbol{t} \\ \boldsymbol{0}^T & 1 \end{bmatrix} \boldsymbol{X}$$

Transformation from normalized 3D coordinates to pixel image coordinates:

$$\boldsymbol{x} \cong \boldsymbol{K}[\boldsymbol{I}|\boldsymbol{0}]\boldsymbol{X}_{\text{cam}}$$

Extrinsic parameters: Rotation and translation



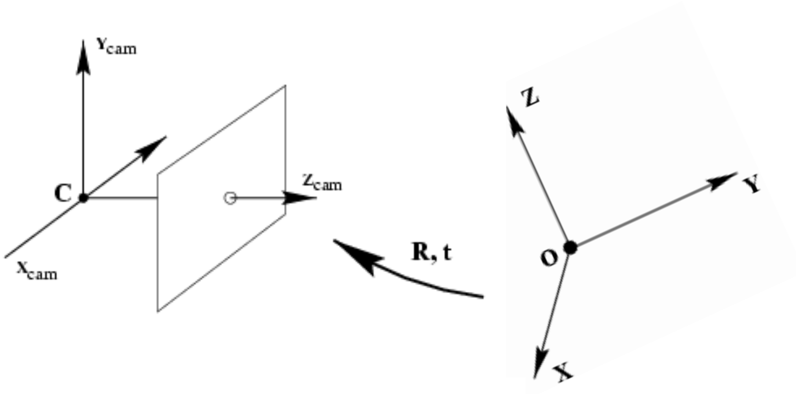
In *homogeneous* coordinates:

$$\boldsymbol{x} \cong \boldsymbol{K}[\boldsymbol{I}|\boldsymbol{0}] \begin{bmatrix} \boldsymbol{R} & \boldsymbol{t} \\ \boldsymbol{0}^T & 1 \end{bmatrix} \boldsymbol{X}$$

Finally:

$$\boldsymbol{x} \cong \boldsymbol{K}[\boldsymbol{R}|\boldsymbol{t}]\boldsymbol{X} \qquad \boldsymbol{t} = -\boldsymbol{R}\widetilde{\boldsymbol{C}}$$

Extrinsic parameters: Rotation and translation



$$x \cong K[R|t]X \qquad t = -R\tilde{C}$$

coords. of
camera center
in world frame

- What is the projection of the camera center in world coordinates?

$$PC = K[R \mid -R\tilde{C}]\begin{pmatrix}\tilde{C}\\1\end{pmatrix} = K(R\tilde{C} - R\tilde{C}) = \mathbf{0}$$

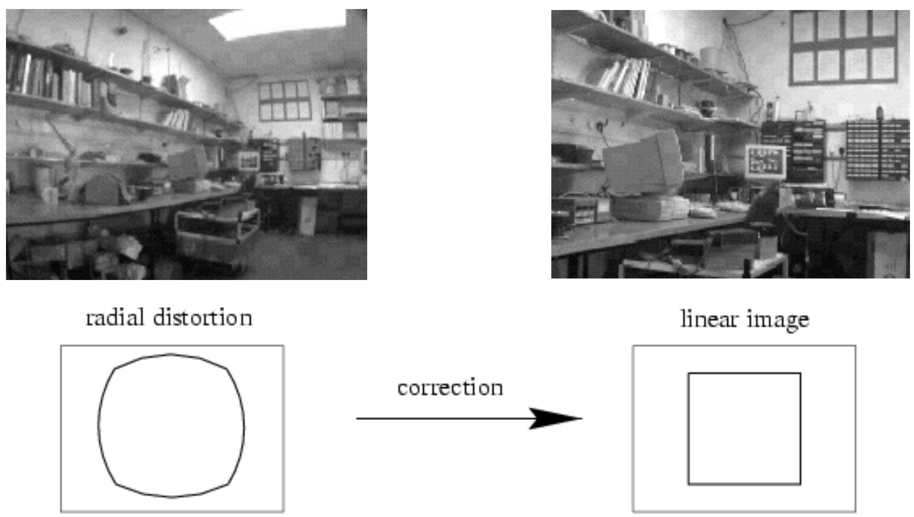
- The camera center is the null space of the projection matrix!

Camera parameters: Summary

$$\boldsymbol{P} = \boldsymbol{K}[\boldsymbol{R}|\boldsymbol{t}]$$

- Intrinsic parameters
  - Principal point coordinates
  - Focal length
  - Pixel magnification factors
  - *Skew (non-rectangular pixels) – not important in practice*
  - *Radial distortion – important in practice!*

$$\boldsymbol{K} = \begin{bmatrix} m_x & 0 & 0 \\ 0 & m_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \alpha_x & 0 & \beta_x \\ 0 & \alpha_y & \beta_y \\ 0 & 0 & 1 \end{bmatrix}$$



Camera extrinsics

$$\mathcal{T}_i \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \mathcal{T}_e$$

Intrinsics

Extrinsics

$$\begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix} = \begin{bmatrix} \text{Transformation} \\ \text{mapping image} \\ \text{plane coords to} \\ \text{pixel coords} \end{bmatrix} \mathcal{C}_p \begin{bmatrix} \text{Transformation} \\ \text{mapping world} \\ \text{coords to camera} \\ \text{coords} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{bmatrix}$$

Rotation matrix - orthornormal, det=1

$$\begin{bmatrix} \mathcal{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}$$

Camera intrinsics

$$\mathcal{T}_i \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \mathcal{T}_e$$

Intrinsics | Extrinsics

$$\begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix} = \begin{bmatrix} \text{Transformation} \\ \text{mapping image} \\ \text{plane coords to} \\ \text{pixel coords} \end{bmatrix} \mathcal{C}_p \begin{bmatrix} \text{Transformation} \\ \text{mapping world} \\ \text{coords to camera} \\ \text{coords} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{bmatrix}$$

↑

$$\begin{bmatrix} as & k & c_x \\ 0 & s & c_y \\ 0 & 0 & 1/f \end{bmatrix}$$

Overview

- Camera calibration
  - Intrinsic camera parameters
  - Extrinsic camera parameters
  - Estimation

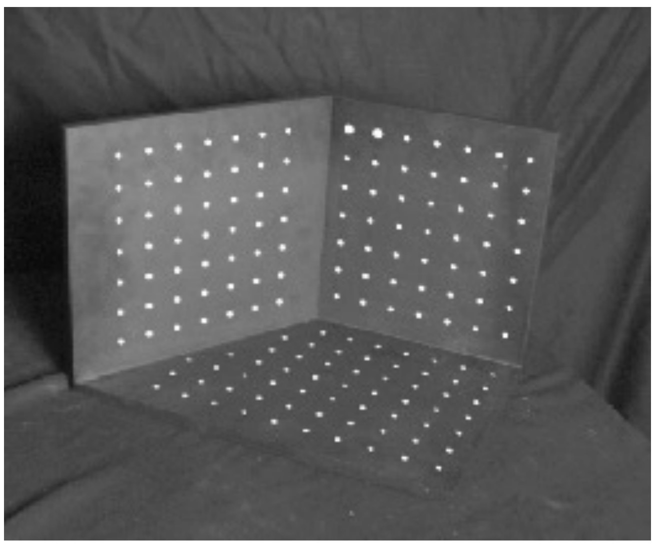
Camera calibration

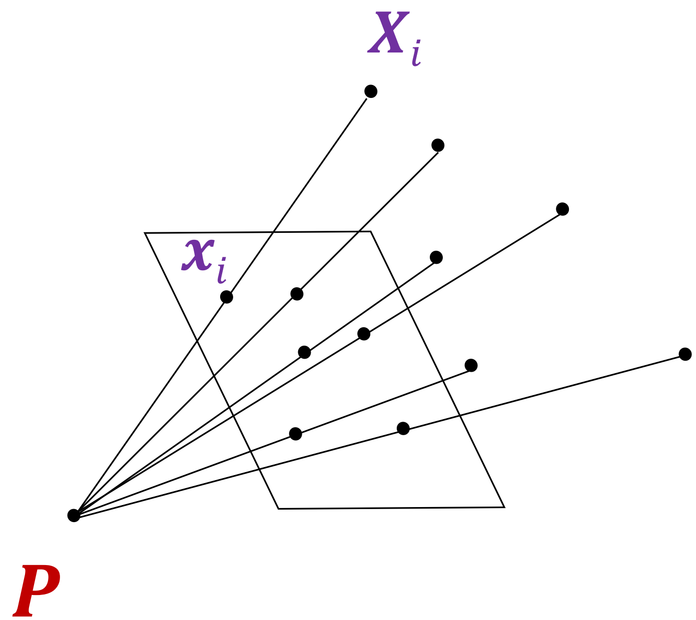
$$\boldsymbol{x} \cong \boldsymbol{K}[\boldsymbol{R} \quad \boldsymbol{t}]\boldsymbol{X}$$

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \cong \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

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

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

Known 2D image coords		Known 3D locations		
880	214	312.747	309.140	30.086
43	203	305.796	311.649	30.356
270	197	307.694	312.358	30.418
886	347	310.149	307.186	29.298
745	302	311.937	310.105	29.216
943	128	311.202	307.572	30.682
476	590	307.106	306.876	28.660
419	214	309.317	312.490	30.230
317	335	307.435	310.151	29.318
783	521	308.253	306.300	28.881
235	427	306.650	309.301	28.905
665	429	308.069	306.831	29.189
655	362	309.671	308.834	29.029
427	333	308.255	309.955	29.267
412	415	307.546	308.613	28.963
746	351	311.036	309.206	28.913
434	415	307.518	308.175	29.069
525	234	309.950	311.262	29.990
716	308	312.160	310.772	29.080
602	187	311.988	312.709	30.514

Image credit: J. Hays

Camera calibration: non-linear method

N reference points with known position in 3D $\mathbf{s}_i = [s_{x,i}, s_{y,i}, s_{z,i}]$

Predict the locations of the known points in camera using estimated camera parameters $\mathbf{t}_i$

Compare to observed locations $\hat{\mathbf{t}}_i = [\hat{t}_{x,i}, \hat{t}_{y,i}]$

$$\hat{\mathbf{t}}_i = \mathbf{t}_i + \xi_i$$

Minimize least-squares error

$$\sum_i \xi_i^T \xi_i.$$

Camera calibration: non-linear method

Two issues:

write out equations for optimization problem

good start point for optimization

Camera extrinsics

$$\mathcal{T}_i \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \mathcal{T}_e$$

Intrinsics

Extrinsics

$$\begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix} = \begin{bmatrix} \text{Transformation} \\ \text{mapping image} \\ \text{plane coords to} \\ \text{pixel coords} \end{bmatrix} \mathcal{C}_p \begin{bmatrix} \text{Transformation} \\ \text{mapping world} \\ \text{coords to camera} \\ \text{coords} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{bmatrix}$$

Rotation matrix - orthornormal, det=1

$$\begin{bmatrix} \mathcal{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}$$

Camera intrinsics

$$\mathcal{T}_i \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \mathcal{T}_e$$

Intrinsics | Extrinsics

$$\begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix} = \begin{bmatrix} \text{Transformation} \\ \text{mapping image} \\ \text{plane coords to} \\ \text{pixel coords} \end{bmatrix} \mathcal{C}_p \begin{bmatrix} \text{Transformation} \\ \text{mapping world} \\ \text{coords to camera} \\ \text{coords} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{bmatrix}$$

↑

$$\begin{bmatrix} as & k & c_x \\ 0 & s & c_y \\ 0 & 0 & 1/f \end{bmatrix}$$

Camera calibration: non-linear, equations

$$\sum_i \xi_i^T \xi_i = \sum_i (t_{x,i} - p_{x,i})^2 + (t_{y,i} - p_{y,i})^2$$

The i's are intrinsic parameters (remember, this is upper triangular)

$$\begin{aligned} p_{x,i} &= \frac{i_{11} g_{x,i} + i_{12} g_{y,i} + i_{13} g_{z,i}}{g_{z,3}} \\ p_{y,i} &= \frac{i_{22} g_{x,i} + i_{23} g_{z,i}}{g_{z,i}} \end{aligned}$$

The e's are extrinsic parameters (they are constrained)

$$\begin{aligned} g_{x,i} &= e_{11} s_{x,i} + e_{12} s_{y,i} + e_{13} s_{z,i} + e_{14} \\ g_{y,i} &= e_{21} s_{x,i} + e_{22} s_{y,i} + e_{23} s_{z,i} + e_{24} \\ g_{z,i} &= e_{31} s_{x,i} + e_{32} s_{y,i} + e_{33} s_{z,i} + e_{34} \end{aligned}$$

$$1 - \sum_v e_{j,1v}^2 = 0 \text{ and } 1 - \sum_v e_{j,2v}^2 = 0 \text{ and } 1 - \sum_v e_{j,3v}^2 = 0$$

$$\sum_v e_{j,1v} e_{j,2v} = 0 \text{ and } \sum_v e_{j,1v} e_{j,3v} = 0 \text{ and } 1 - \sum_v e_{j,2v} e_{j,3v} = 0 \quad .$$

Camera calibration: non-linear

Strategy:

chuck it into a constrained optimizer and run

this fails – you need a good starting point

Start point:

neat linear construction

Camera calibration: Linear method

Write $\mathbf{C}_j^T$ for the j'th row of the camera matrix, and $\mathbf{S}_i = [s_{x,i}, s_{y,i}, s_{z,i}, 1]^T$ for homogeneous coordinates representing the i'th point in 3D. Then, assuming no errors in measurement, we have

$$\hat{t}_{x,i} = \frac{\mathbf{C}_1^T \mathbf{S}_i}{\mathbf{C}_3^T \mathbf{S}_i} \text{ and } \hat{t}_{y,i} = \frac{\mathbf{C}_2^T \mathbf{S}_i}{\mathbf{C}_3^T \mathbf{S}_i}, \tag{24.3}$$

which we can rewrite as

$$\mathbf{C}_3^T \mathbf{S}_i \hat{t}_{x,i} - \mathbf{C}_1^T \mathbf{S}_i = 0 \text{ and } \mathbf{C}_3^T \mathbf{S}_i \hat{t}_{y,i} - \mathbf{C}_2^T \mathbf{S}_i = 0. \tag{24.4}$$

x component of location of i'th point in image

One match gives two linearly independent constraints on the camera matrix

Calibration: Linear method

N points gives 2N homogeneous equations

$$\begin{pmatrix} -\mathbf{S}_1^T & 0 & \mathbf{S}_1^T t_{x,1} \\ 0 & -\mathbf{S}_1^T & \mathbf{S}_1^T t_{y,1} \\ \dots & \dots & \dots \\ -\mathbf{S}_N^T & 0 & \mathbf{S}_N^T t_{x,N} \\ 0 & -\mathbf{S}_N^T & \mathbf{S}_N^T t_{y,N} \end{pmatrix} \begin{pmatrix} \mathbf{C}_1 \\ \mathbf{C}_2 \\ \mathbf{C}_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \dots \\ 0 \\ 0 \end{pmatrix}$$

The camera matrix is 3 x 4 but scale doesn't matter so there are 11 degrees of freedom – we can estimate it with 6 points

Camera calibration: Linear method

- Final linear system:

$$\begin{pmatrix} -\mathbf{S}_1^T & 0 & \mathbf{S}_1^T t_{x,1} \\ 0 & -\mathbf{S}_1^T & \mathbf{S}_1^T t_{y,1} \\ \cdots & \cdots & \cdots \\ -\mathbf{S}_N^T & 0 & \mathbf{S}_N^T t_{x,N} \\ 0 & -\mathbf{S}_N^T & \mathbf{S}_N^T t_{y,N} \end{pmatrix} \begin{pmatrix} \mathbf{C}_1 \\ \mathbf{C}_2 \\ \mathbf{C}_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \cdots \\ 0 \\ 0 \end{pmatrix}$$

- What if all the n 3D points are *coplanar*, i.e., there exists a set of line parameters $\mathbf{\Pi}^T = (a, b, c, d)^T$ such that $\mathbf{\Pi}^T \boldsymbol{X}_i = 0$ for all i?
  - Then 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 parameters from camera matrix

Remember this: *Given a* 3×4 *camera matrix* $\mathcal{P}$, *the homogeneous coordinates of the focal point of that camera are given by* $\mathbf{X}$, *where* $\mathcal{P}\mathbf{X} = [0, 0, 0]^T$

Camera matrix is:

$$\mathcal{T}_i \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \mathcal{T}_e$$

Camera parameters from camera matrix

Remember this: *Given a 3×4 camera matrix $\mathcal{P}$, the homogeneous coordinates of the focal point of that camera are given by* $\mathbf{X}$, *where* $\mathcal{P}\mathbf{X} = [0, 0, 0]^T$

Camera matrix is:

$$\mathcal{T}_i \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \mathcal{T}_e$$

Remember this: *Assume camera matrix* $\mathcal{P}$ *has null space* $\lambda\mathbf{u} = \lambda\left[\mathbf{f}^T, 1\right]^T$. *Then we must have* $\mathcal{T}_e\mathbf{u} = [0, 0, 0, 1]^T$, *so we must have*

$$\mathcal{T}_e = \begin{bmatrix} \mathcal{R} & -\mathcal{R}\mathbf{f} \\ \mathbf{0}^T & 1 \end{bmatrix} \tag{24.5}$$

Parameters from camera matrix

This means that, if we know $\mathcal{R}$, we can recover the translation from the focal point. We must now recover the intrinsic transformation and $\mathcal{R}$ from what we know.

$$\lambda \mathcal{P} = \mathcal{T}_i \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mathcal{R} & -\mathcal{R}\mathbf{f} \\ \mathbf{0}^T & 1 \end{bmatrix} = \begin{bmatrix} \mathcal{T}_i\mathcal{R} & -\mathcal{T}_i\mathcal{R}\mathbf{f} \end{bmatrix} \tag{24.6}$$

Parameters from camera matrix

We do not know λ, but we do know $\mathcal{P}$. Now write $\mathcal{P}_l$ for the left 3×3 block of $\mathcal{P}$, and recall that $\mathcal{T}_i$ is upper triangular and $\mathcal{R}$ orthonormal. The first question is the sign of λ. We expect $\mathsf{Det}(\mathcal{R}) = 1$, and $\mathsf{Det}(\mathcal{T}_i) > 0$, so $\mathsf{Det}(\mathcal{P}_l)$ should be positive. This yields the sign of λ – choose a sign $s \in \{-1, 1\}$ so that $\mathsf{Det}(s\mathcal{P}_l)$ is positive.

We can now factor $s\mathcal{P}_l$ into an upper triangular matrix $\mathcal{T}$ and an orthonormal matrix $\mathcal{Q}$. This is an RQ factorization (Section 35.2). Recall we could not distinguish between scaling caused by the focal length and scaling caused by pixel scale, so that

$$\mathcal{T}_i = \begin{bmatrix} as & k & c_x \\ 0 & s & c_y \\ 0 & 0 & 1 \end{bmatrix} \tag{24.7}$$

In turn, we have $\lambda = s(1/t_{33})$, $c_y = (t_{23}/t_{33})$, $s = (t_{22}/t_{33})$, $c_x = (t_{13}/t_{33})$, $k = (t_{12}/t_{33})$, and $a = (t_{11}/t_{22})$.

Camera calibration: nonlinear

- *non-linear* methods are preferred
  - Can include radial distortion and constraints such as known focal length, orthogonality, visibility of points

Multi-view geometry "Bible"