Robot Vision:
Image formation

Prof. Friedrich Fraundorfer

SS 2024

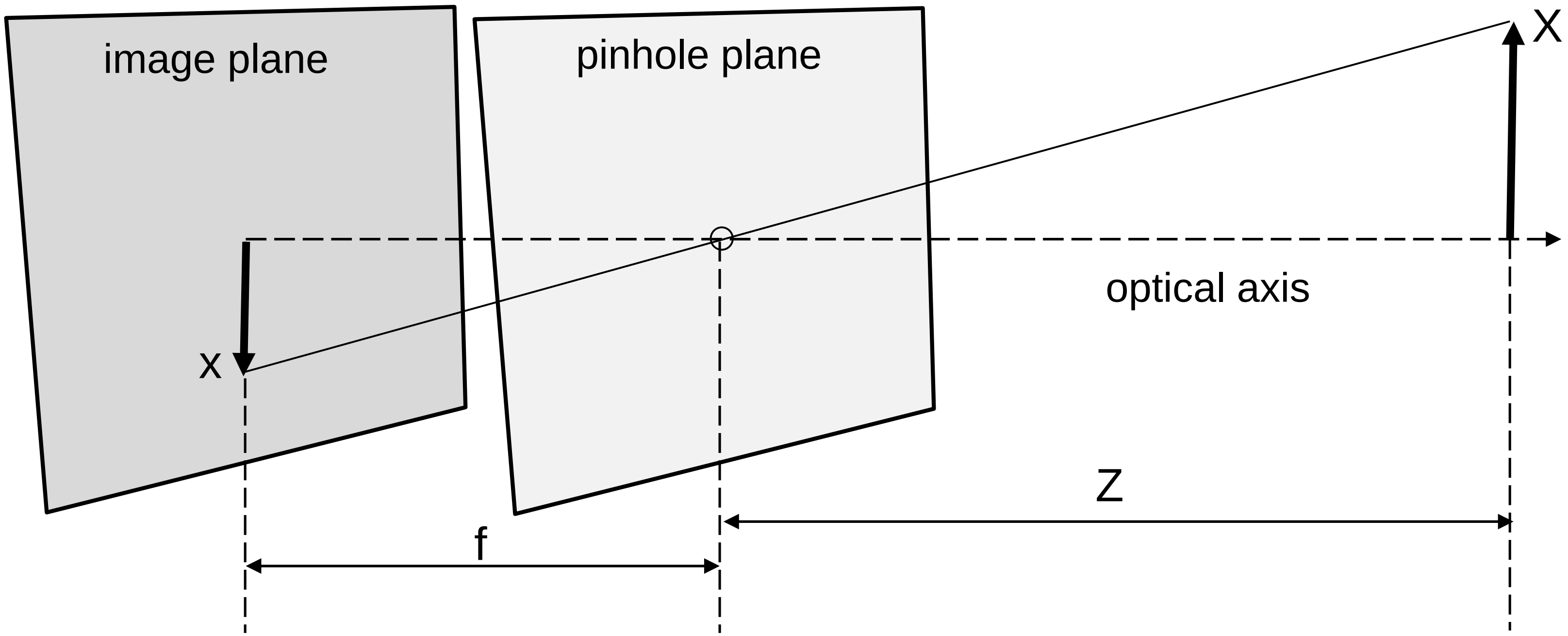
Outline

- Image formation
  - The pinhole camera model
  - The projection equation
  - Camera matrix estimation

Learning goals

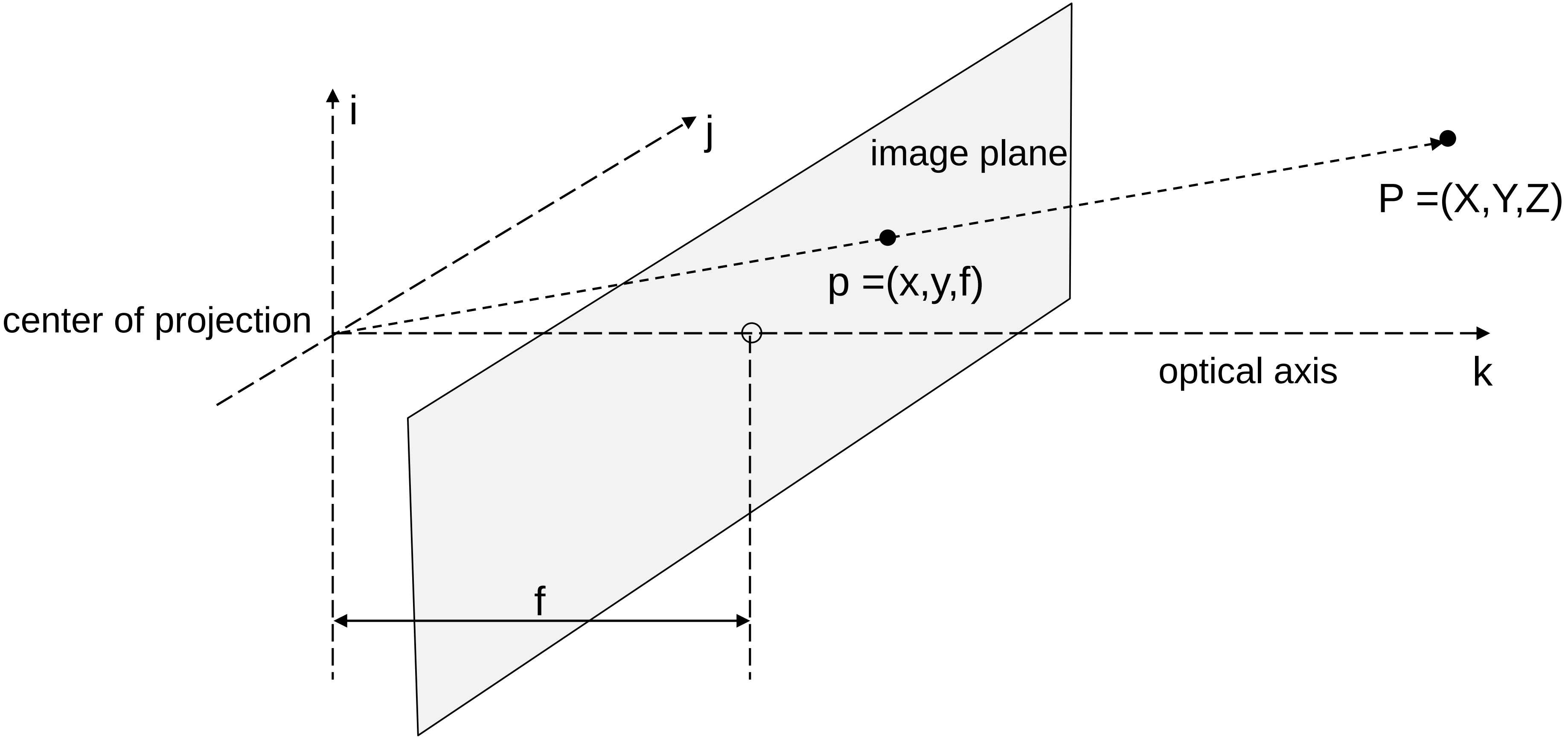
- To be able to explain the pinhole camera model
- To be able to explain the image projection process mathematically
- To be able to explain camera matrix estimation

The pinhole camera model



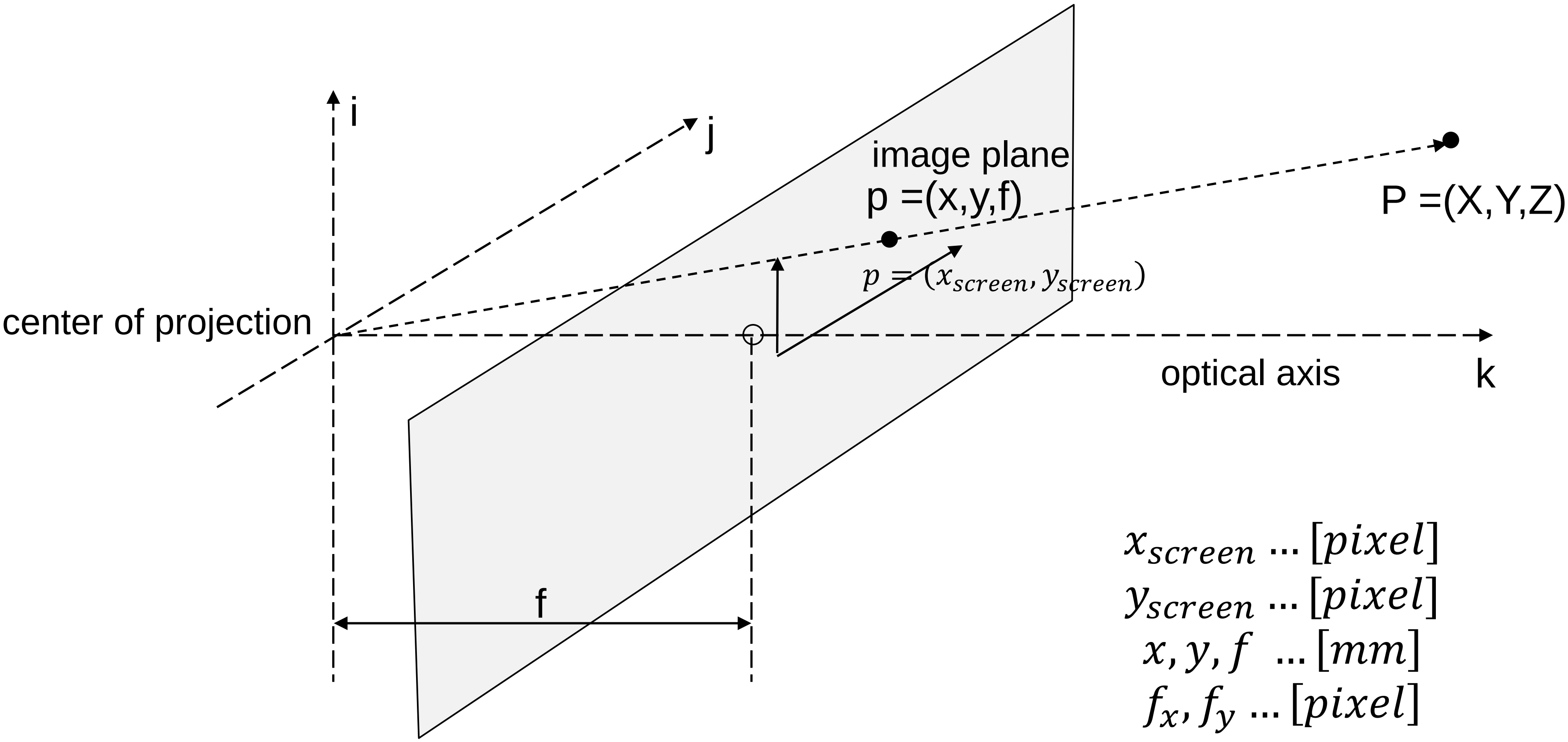
$$-x = f\,\frac{X}{Z}$$

The pinhole camera model



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

From mm's to pixels



$$x_{screen} \text{ ... } [pixel]$$
$$y_{screen} \text{ ... } [pixel]$$
$$x, y, f \text{ ... } [mm]$$
$$f_x, f_y \text{ ... } [pixel]$$
$$c_x, c_y \text{ ... } [pixel]$$
$$s_x, s_y \text{ ... } [\frac{pixel}{mm}]$$

$$x_{screen} = f_x\left(\frac{X}{Z}\right) + c_x$$

$$y_{screen} = f_y\left(\frac{Y}{Z}\right) + c_y$$

$$f_x = f s_x$$

$$f_y = f s_y$$

Matrix notation

$$x_{screen} = f_x \left(\frac{X}{Z}\right) + c_x$$

$$y_{screen} = f_y \left(\frac{Y}{Z}\right) + c_y$$

$$\begin{pmatrix} f_x X + c_x Z \\ f_y Y + c_y Z \\ Z \end{pmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \qquad K = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}$$

$$\begin{pmatrix} f_x \frac{X}{Z} + c_x \\ f_y \frac{Y}{Z} + c_y \\ 1 \end{pmatrix} \cong \begin{pmatrix} f_x X + c_x Z \\ f_y Y + c_y Z \\ Z \end{pmatrix}$$

- K-Matrix is often called “calibration matrix” or “interior orientation”

Normalization of image coordinates

- Convert 2D image coordinates into 2D projective coordinates with image plane distance = 1

$$K = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}$$

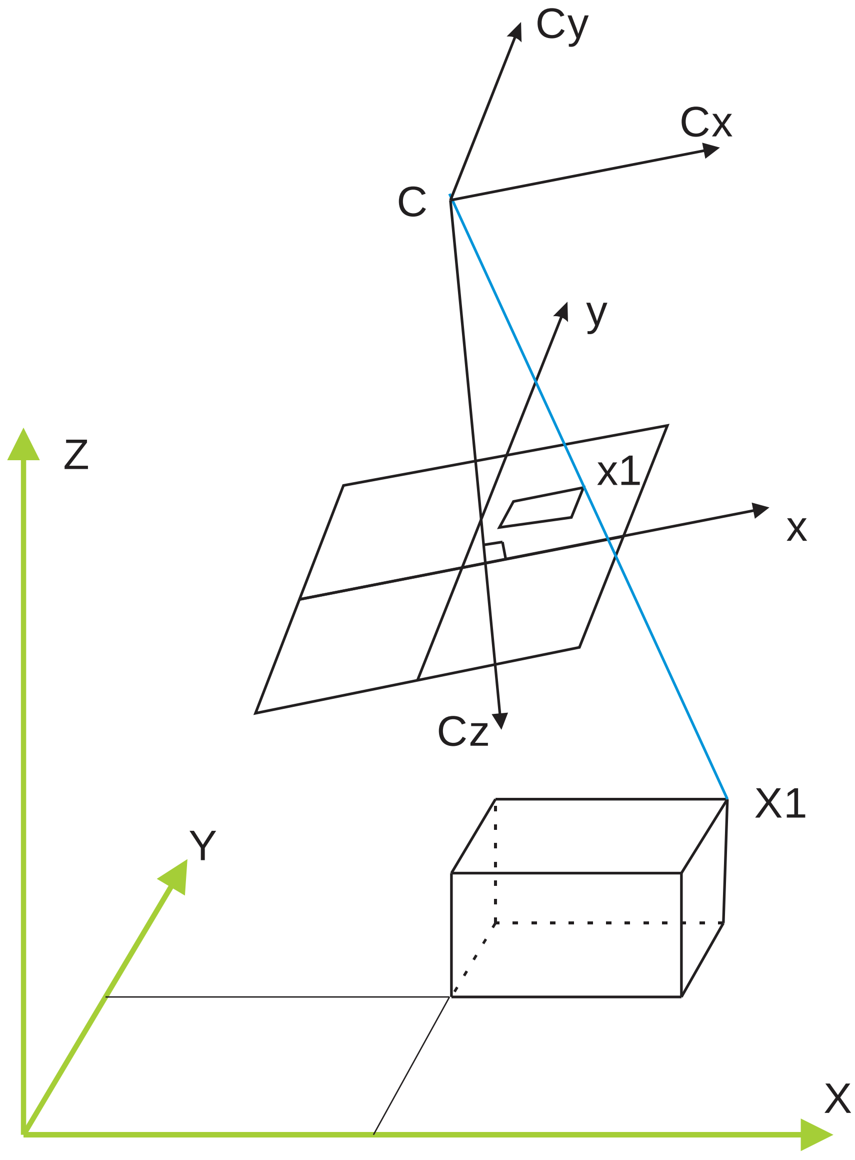
$$x_{screen} = f_x \left(\frac{X}{Z}\right) + c_x$$

$$\begin{pmatrix} x_n \\ y_n \\ 1 \end{pmatrix} = K^{-1} \begin{pmatrix} x_{screen} \\ y_{screen} \\ 1 \end{pmatrix}$$

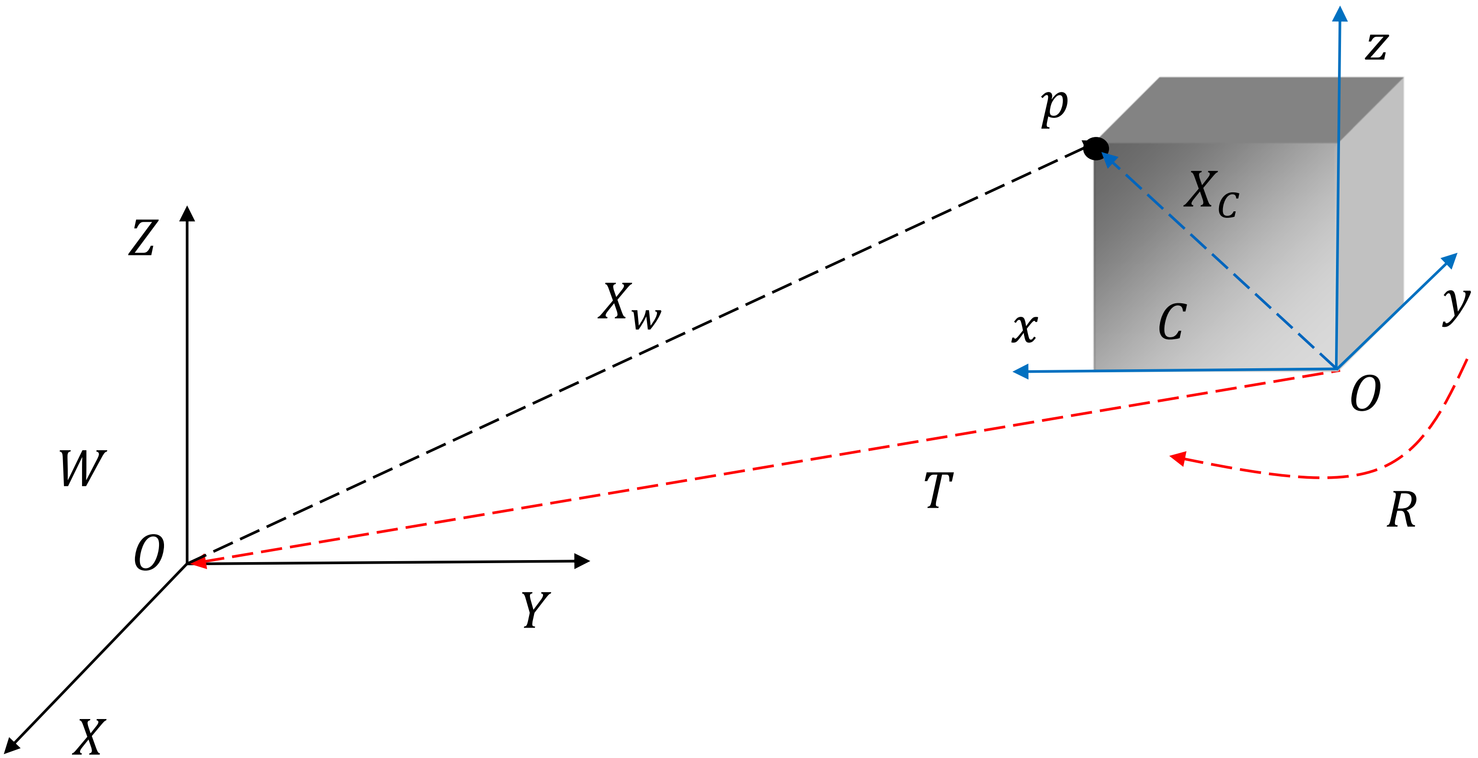
$$y_{screen} = f_y \left(\frac{Y}{Z}\right) + c_y$$

- x_n, y_n are unit-less 2D projective coordinates with a z-distance of 1

Exterior orientation



Rigid transformations



$x \in \mathbb{R}^n$

$\mathrm{T} \in \mathbb{R}^n$

$R \in \mathbb{R}^{n \times n}$

- Coordinates are related by: $X_c = RX_w + T$

$$\begin{bmatrix} X_c \\ 1 \end{bmatrix} = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ 1 \end{bmatrix}$$

The complete projection equation

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \quad \textbf{?} \quad \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

Perspective projection

- Mapping 3D projective space onto 2D projective space
- A projection onto a space of one lower dimension can be achieved by eliminating one of the coordinates
- General projective transformation in 3D is a 4x4 matrix

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} t_{11} & t_{12} & t_{13} & t_{14} \\ t_{21} & t_{22} & t_{23} & t_{24} \\ t_{31} & t_{32} & t_{33} & t_{34} \\ t_{41} & t_{42} & t_{43} & t_{44} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{bmatrix}$$

- Image projection from 3D to 2D

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \end{bmatrix} = \begin{bmatrix} t_{11} & t_{12} & t_{13} & t_{14} \\ t_{21} & t_{22} & t_{23} & t_{24} \\ t_{31} & t_{32} & t_{33} & t_{34} \\ & & & \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{bmatrix}$$

- The coordinate x4 is dropped

$$\begin{bmatrix} x \\ y \\ w \end{bmatrix} = P_{3\times 4} \begin{bmatrix} X \\ Y \\ Z \\ W \end{bmatrix}$$

The complete projection equation

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \qquad \textbf{?} \qquad \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

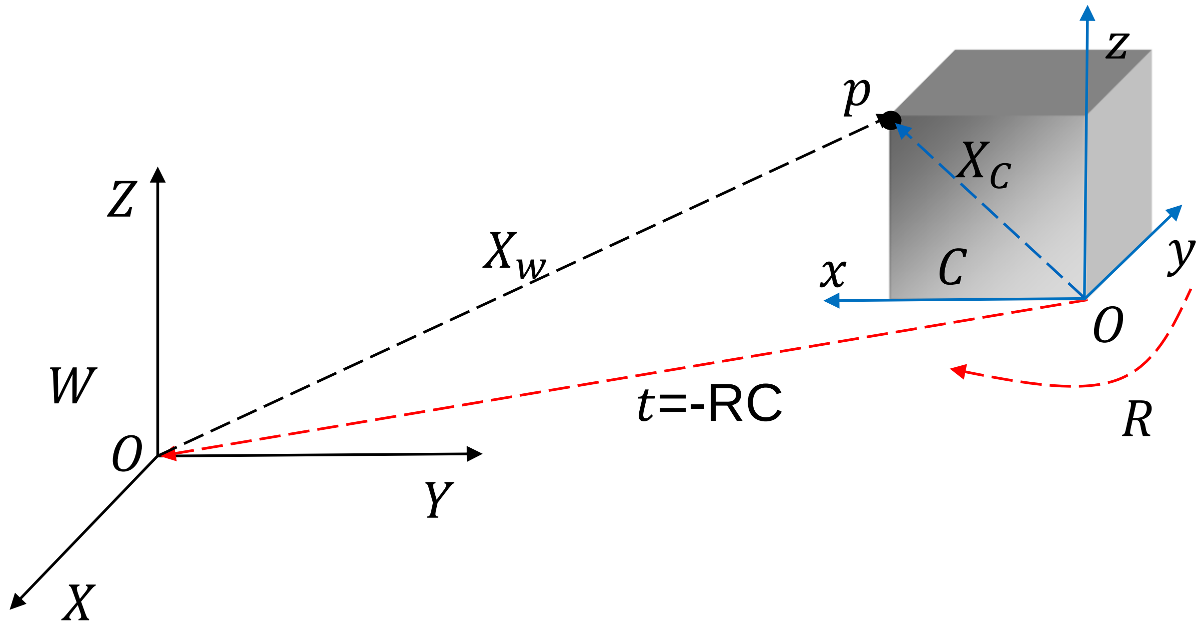
$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R & T \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = P_{3\times 4} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

Camera matrix

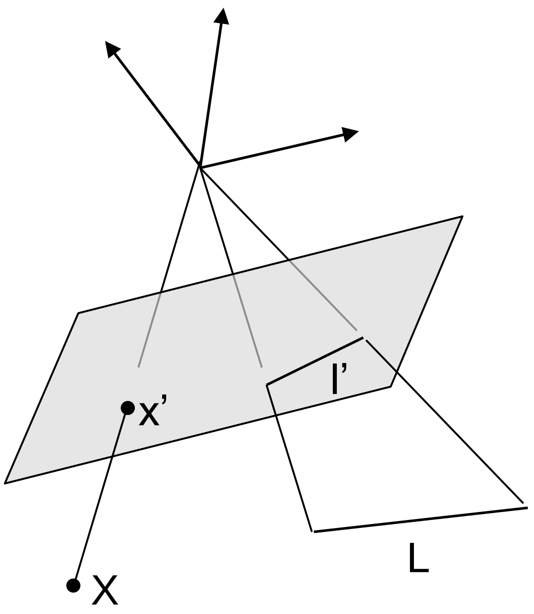
$$P = KR[I|-C] = K[R|-RC] = K[R|t]$$
$$t = -RC$$



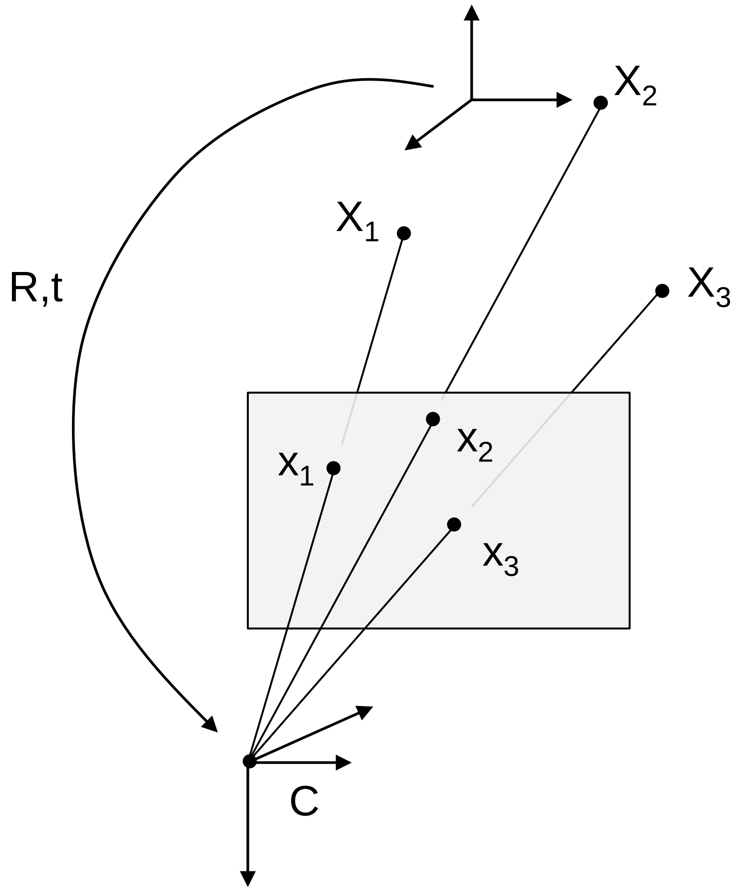
- Camera matrix P is a coordinate transformation and then a projection
- C … 3x1 coordinate of the camera center in world coordinate
- R … 3x3 rotation matrix representing the orientation of the camera coordinate frame
- K … 3x3 calibration matrix

Line projection

- Point projection x = PX
- Line projection is more involved (line l is a 4x4 matrix)
- Therefore indirect projection:
  $l' = x' \times y' = PX \times PY$
  $L = \overline{X\,Y}$

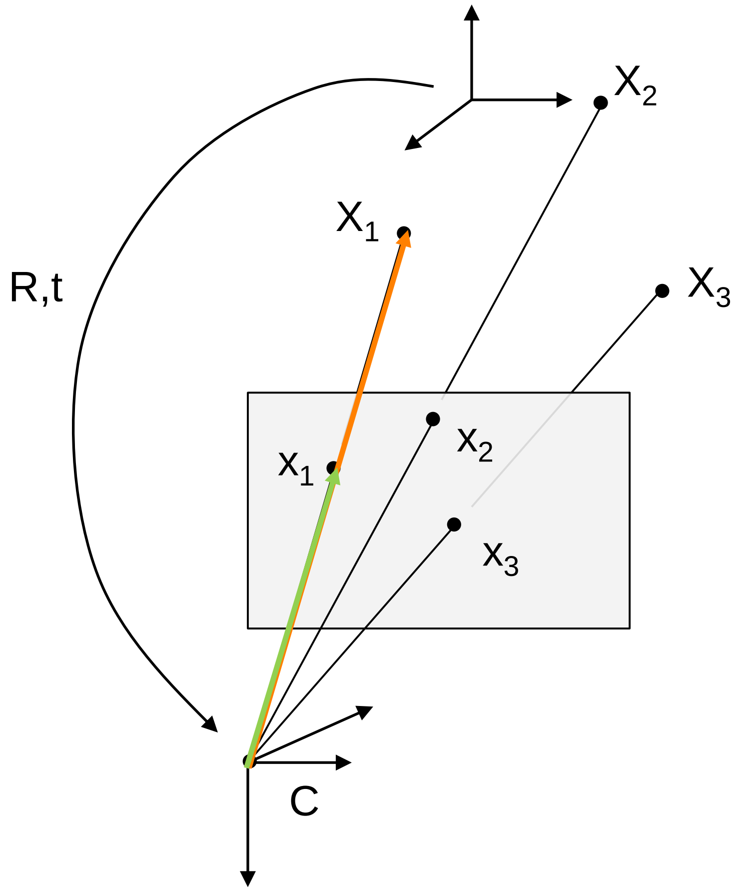


Camera matrix estimation



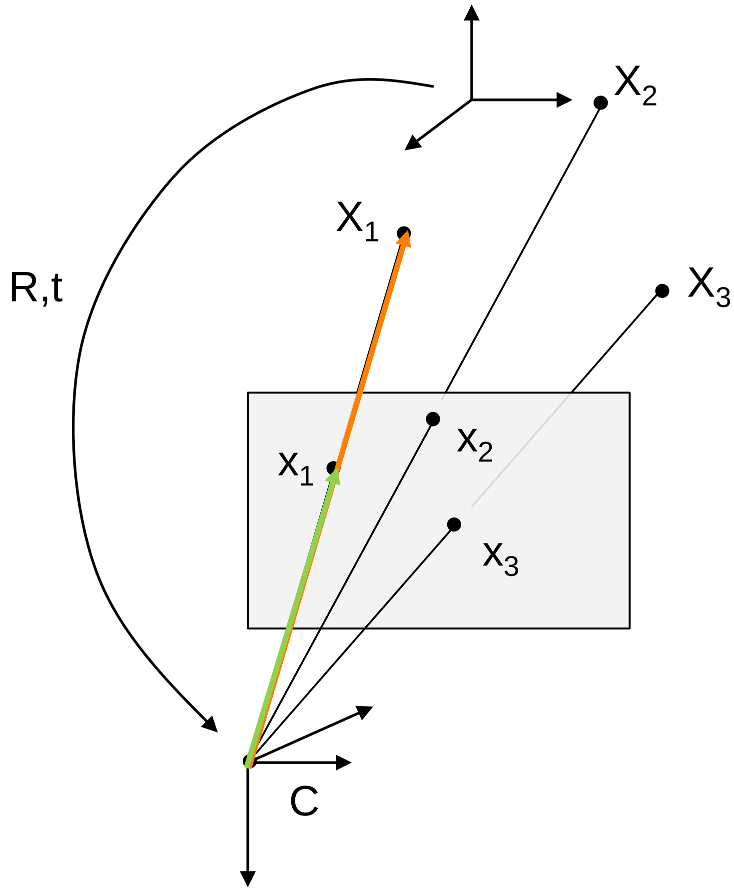
- perspective-n-point problem
- Goal is to estimate camera matrix P such that $x_1 = PX_1$
- $x_1, X_1, x_2, X_2, x_3, X_3$ are known
- Algebraic linear solution requires 6 3D-2D point correspondences, minimal nonlinear solution requires only 3

Camera matrix estimation



- Condition: Measurement vector x needs to have the same direction as projection of X (cross-product equals 0)

Camera matrix estimation



- Condition: Measurement vector x needs to have the same direction as projection of X (cross-product equals 0)

$$x \times (PX) = 0 \; for\; all\; pairs\; x \leftrightarrow X$$
$$y(P_3^T X) - w(P_2^T X) = 0$$
$$w(P_1^T X) - x(P_3^T X) = 0$$
$$x(P_2^T X) - y(P_1^T X) = 0$$

$$\begin{bmatrix} 0 & -wX^T & yX^T \\ wX^T & 0 & -xX^T \\ -yX^T & xX^T & 0 \end{bmatrix} \begin{bmatrix} P_1 \\ P_2 \\ P_3 \end{bmatrix} = 0$$

Recap - Learning goals

- To be able to explain the pinhole camera model
- To be able to explain the image projection process mathematically
- To be able to explain camera matrix estimation