Lecture 2

Rigid-Body Motion
and
Imaging Geometry

OUTLINE

3-D EUCLIDEAN SPACE & RIGID-BODY MOTION
- Coordinates and coordinate frames
- Rigid-body motion and homogeneous coordinates

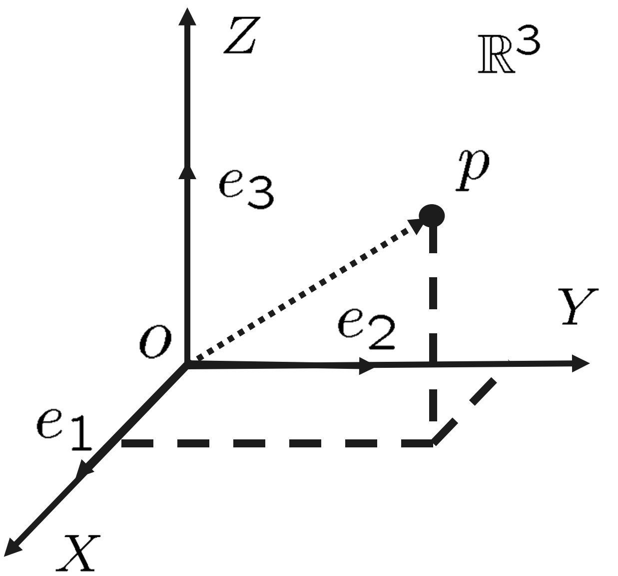
GEOMETRIC MODELS OF IMAGE FORMATION
- Pinhole camera model

CAMERA INTRINSIC PARAMETERS
- From metric to pixel coordinates

SUMMARY OF NOTATION

Standard base vectors:

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad e_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

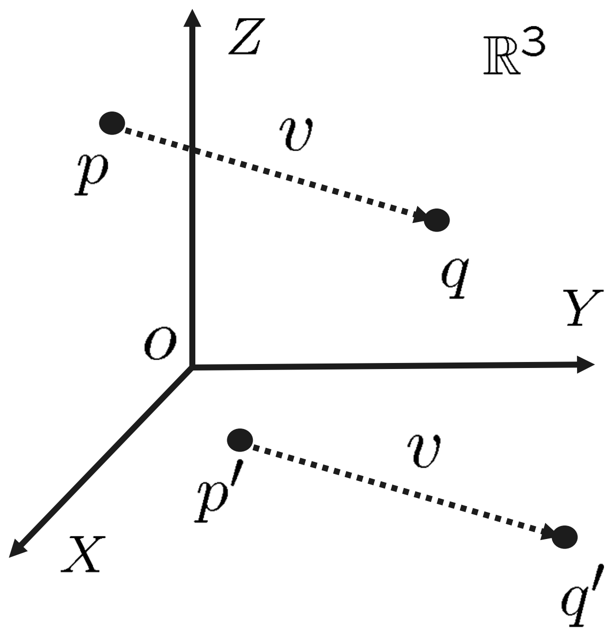


Coordinates of a point p in space:

$$\boldsymbol{X} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \in \mathbb{R}^3$$

A "free" vector is defined by a pair of points (p, q) :

$$\boldsymbol{X}_p = \begin{bmatrix} X_1 \\ Y_1 \\ Z_1 \end{bmatrix} \in \mathbb{R}^3, \ \boldsymbol{X}_q = \begin{bmatrix} X_2 \\ Y_2 \\ Z_2 \end{bmatrix} \in \mathbb{R}^3,$$



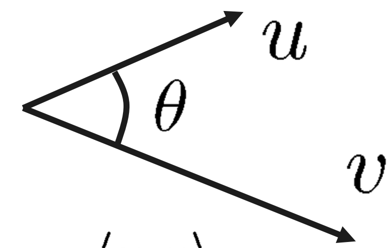
Coordinates of the vector v :

$$v = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} X_2 - X_1 \\ Y_2 - Y_1 \\ Z_2 - Z_1 \end{bmatrix} \in \mathbb{R}^3$$

3-D EUCLIDEAN SPACE - **Inner Product and Cross Product**

Inner product between two vectors:

$$u = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \quad v = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$$



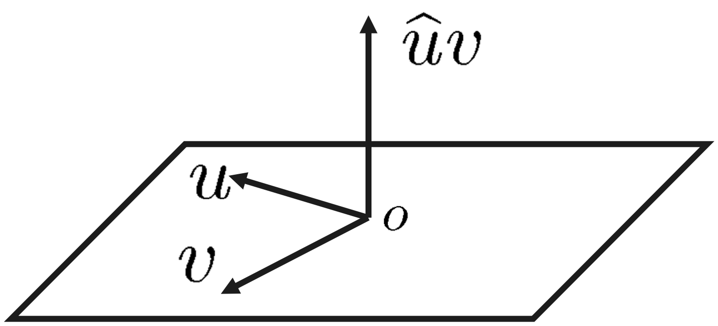
$$\langle u, v \rangle \doteq u^T v = u_1 v_1 + u_2 v_2 + u_3 v_3$$

$$\cos(\theta) = \frac{\langle u, v \rangle}{\|u\| \|v\|}$$

$$\|u\| \doteq \sqrt{u^T u} = \sqrt{u_1^2 + u_2^2 + u_3^3}$$

Cross product between two vectors:

$$u \times v \doteq \widehat{u} v, \qquad u, v \in \mathbb{R}^3$$

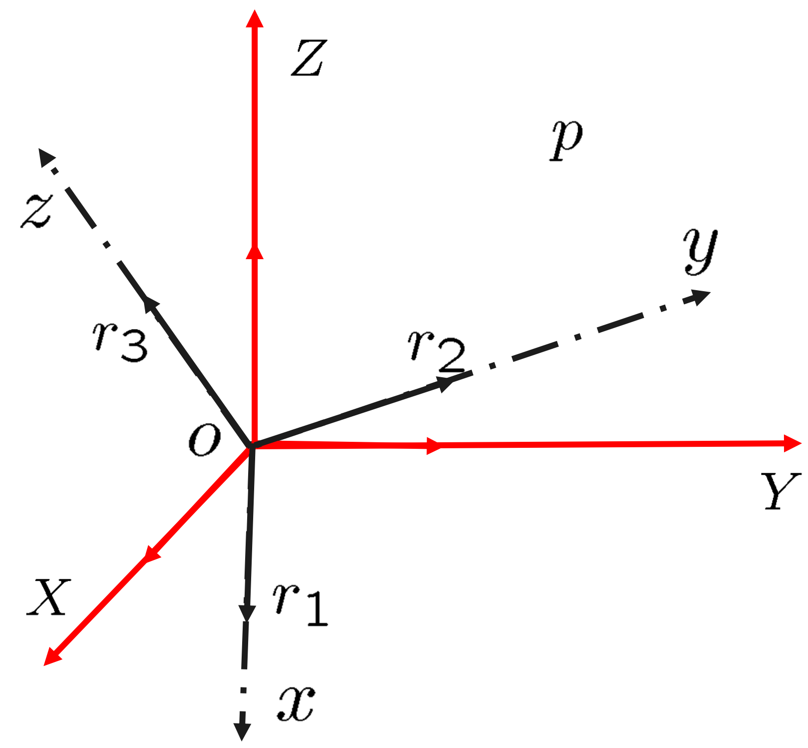


$$\widehat{u} = \begin{bmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{bmatrix} \in \mathbb{R}^{3 \times 3}$$

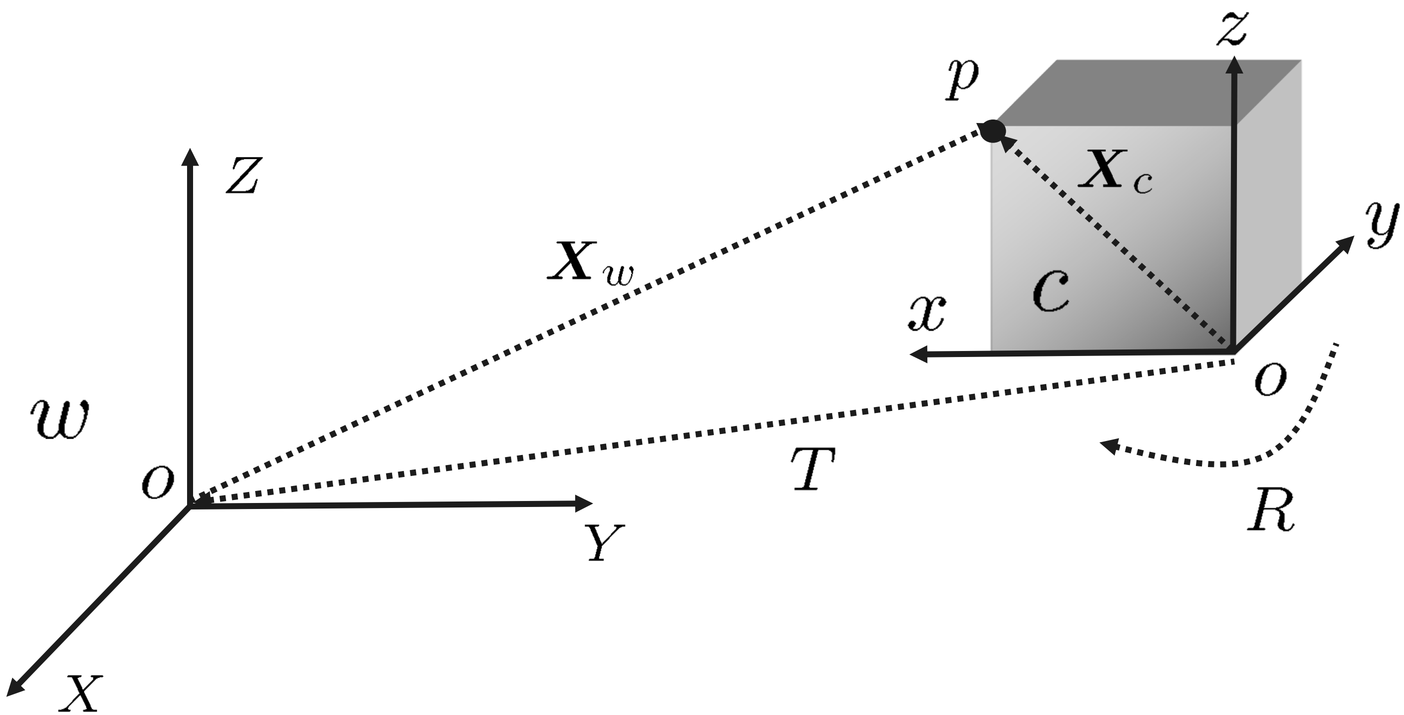
Rotation matrix:

$$R \doteq [r_1, r_2, r_3] \in \mathbb{R}^{3\times 3}$$

$$R^T R = I, \ \det(R) = +1$$



Coordinates are related by: $\boldsymbol{X}_c = R\boldsymbol{X}_w$



Coordinates are related by: $\boldsymbol{X}_c = R\boldsymbol{X}_w + T,$

Velocities are related by: $\dot{\boldsymbol{X}}_c = \widehat{\omega}\boldsymbol{X}_c + v.$

3-D coordinates are related by: $\boldsymbol{X}_c = R\boldsymbol{X}_w + T,$

Homogeneous coordinates:

$$\boldsymbol{X} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad \rightarrow \quad \boldsymbol{X} = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \in \mathbb{R}^4,$$

Homogeneous coordinates/velocities are related by:

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \qquad \begin{bmatrix} \dot{X}_c \\ \dot{Y}_c \\ \dot{Z}_c \\ 1 \end{bmatrix} = \begin{bmatrix} \widehat{\omega} & v \\ 0 & 0 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix}$$

"The Scholar of Athens," Raphael, 1518

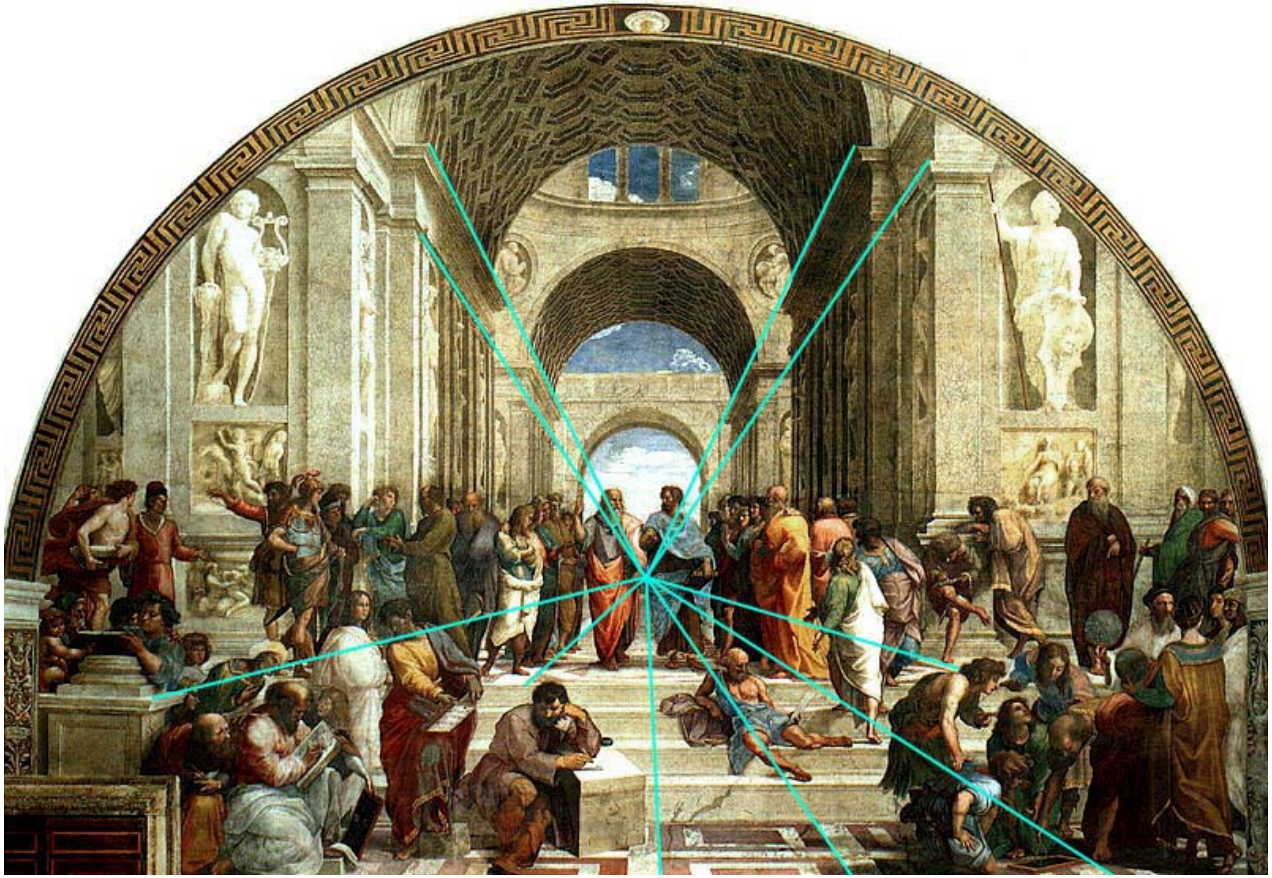
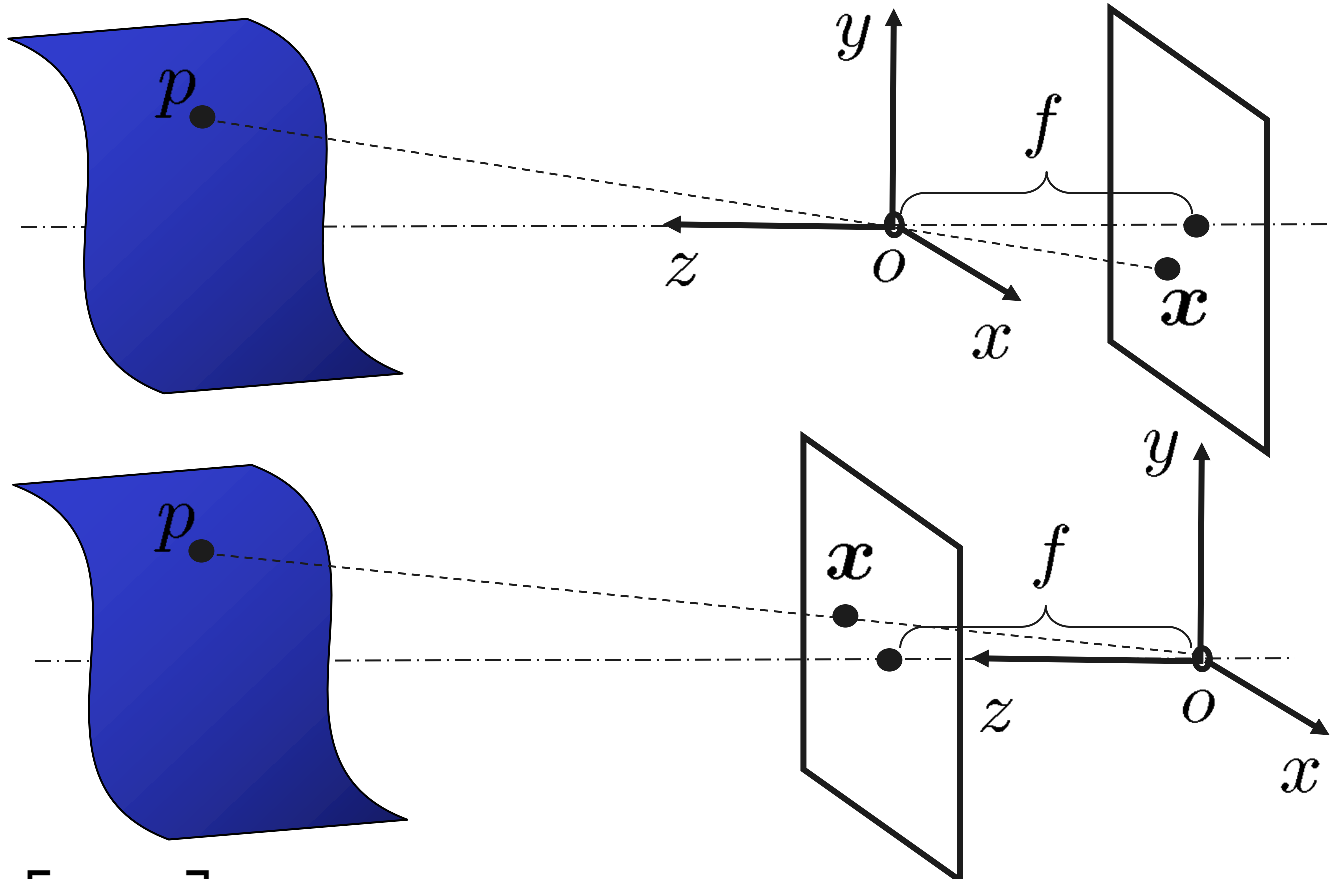


Image courtesy of C. Taylor

Pinhole

Frontal pinhole



$$\boldsymbol{X} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \rightarrow \boldsymbol{x} = \begin{bmatrix} x \\ y \end{bmatrix} = \frac{f}{Z} \begin{bmatrix} X \\ Y \end{bmatrix}$$

2-D coordinates $\boldsymbol{x} = \begin{bmatrix} x \\ y \end{bmatrix} = \frac{f}{Z} \begin{bmatrix} X \\ Y \end{bmatrix}$

Homogeneous coordinates

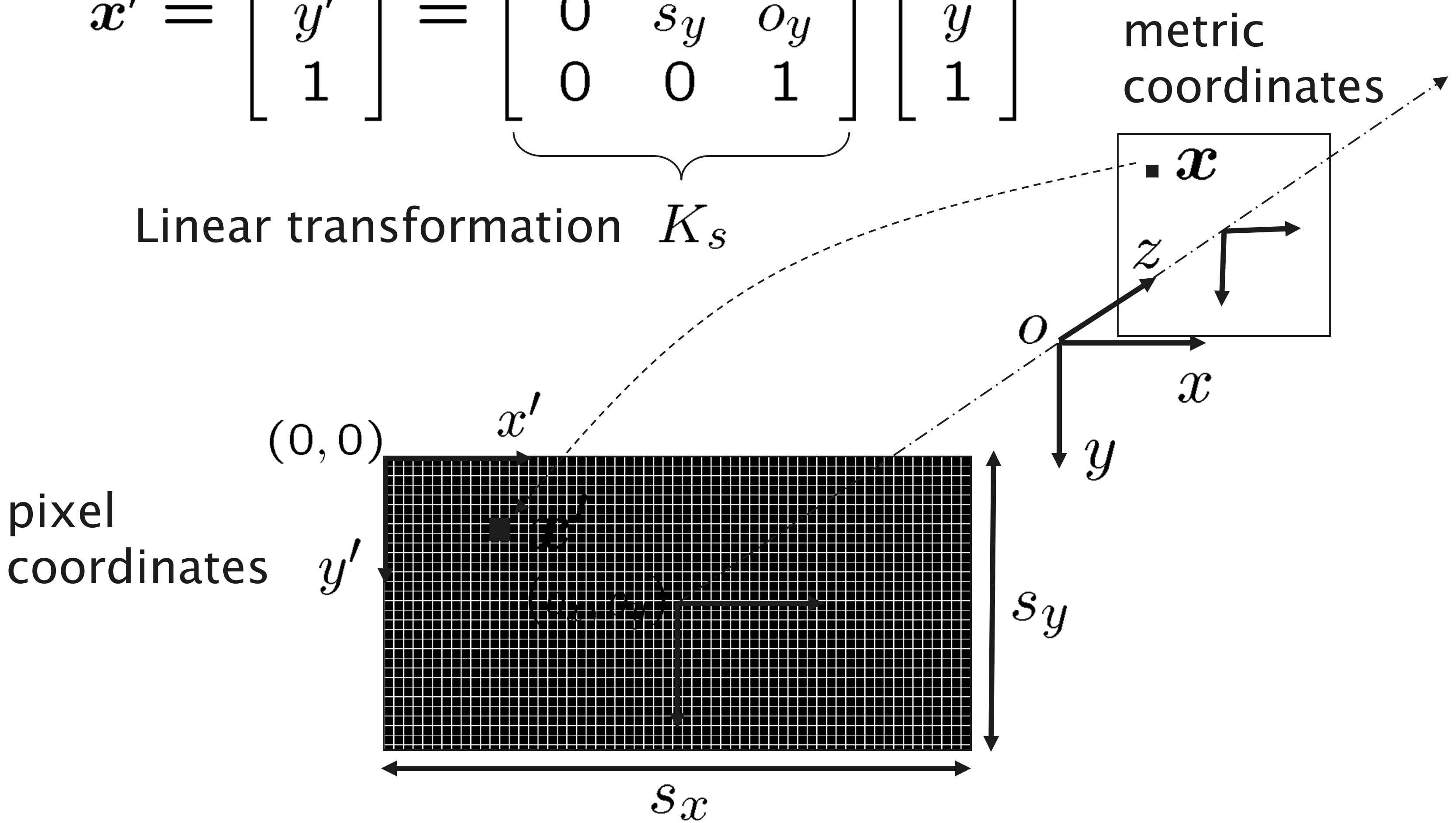
$$\boldsymbol{x} \rightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \frac{1}{Z} \begin{bmatrix} fX \\ fY \\ Z \end{bmatrix}, \qquad \boldsymbol{X} \rightarrow \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix},$$

$$Z \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{K_f} \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{\Pi_0} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

CAMERA PARAMETERS – **Pixel Coordinates**

$$\boldsymbol{x}' = \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} s_x & s_\theta & o_x \\ 0 & s_y & o_y \\ 0 & 0 & 1 \end{bmatrix}}_{} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Linear transformation K_s



Pinhole camera

$$\lambda \boldsymbol{x} = K_f \Pi_0 \boldsymbol{X}$$

Pixel coordinates

$$\boldsymbol{x}' = K_s \boldsymbol{x}$$

$$\lambda \boldsymbol{x}' = K_s K_f \Pi_0 \mathbf{X} = \underbrace{\begin{bmatrix} f s_x & f s_\theta & o_x \\ 0 & f s_y & o_y \\ 0 & 0 & 1 \end{bmatrix}}_{K = K_s K_f} \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{\Pi_0} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Calibration matrix
(intrinsic parameters)

$$\underbrace{K = K_s K_f \qquad \Pi_0}$$

Projection matrix

$$\Pi = [K, 0] \in \mathbb{R}^{3\times 4}$$

Camera model

$$\lambda \boldsymbol{x}' = K \Pi_0 \boldsymbol{X} = \Pi \boldsymbol{X}$$

IMAGE FORMATION – **Image of a Point**

Homogeneous coordinates of a 3-D point p

$$\boldsymbol{X} = [X, Y, Z, W]^T \in \mathbb{R}^4, \quad (W = 1)$$

Homogeneous coordinates of its 2-D image

$$\boldsymbol{x} = [x, y, z]^T \in \mathbb{R}^3, \quad (z = 1)$$

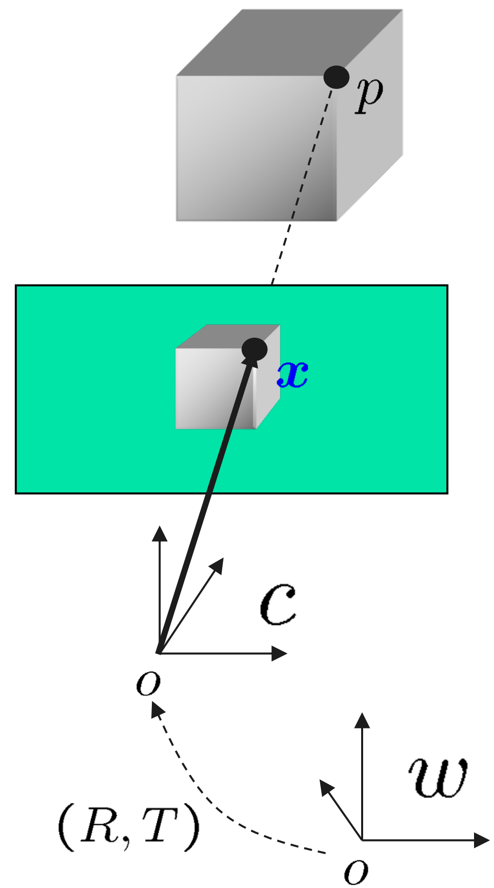
Projection of a 3-D point to an image plane

$$\boxed{\lambda \boldsymbol{x} = \Pi \boldsymbol{X}}$$

$$\lambda \in \mathbb{R}, \ \Pi = [R, T] \in \mathbb{R}^{3\times 4}$$

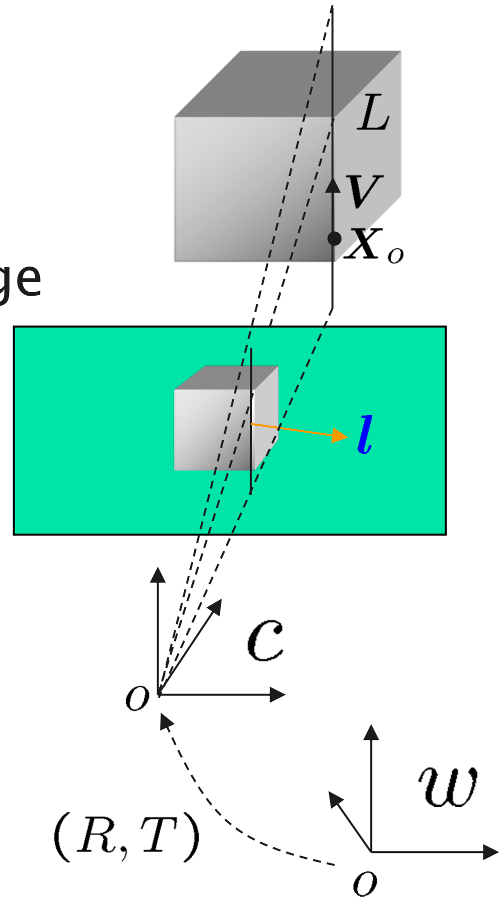
$$\boxed{\lambda \boldsymbol{x}' = \Pi \boldsymbol{X}}$$

$$\lambda \in \mathbb{R}, \ \Pi = [KR, KT] \in \mathbb{R}^{3\times 4}$$



Homogeneous representation of a 3-D line L

$$\boldsymbol{X} = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} X_o \\ Y_o \\ Z_o \\ 1 \end{bmatrix} + \mu \begin{bmatrix} V_1 \\ V_2 \\ V_3 \\ 0 \end{bmatrix}, \quad \mu \in \mathbb{R}$$



Homogeneous representation of its 2-D image

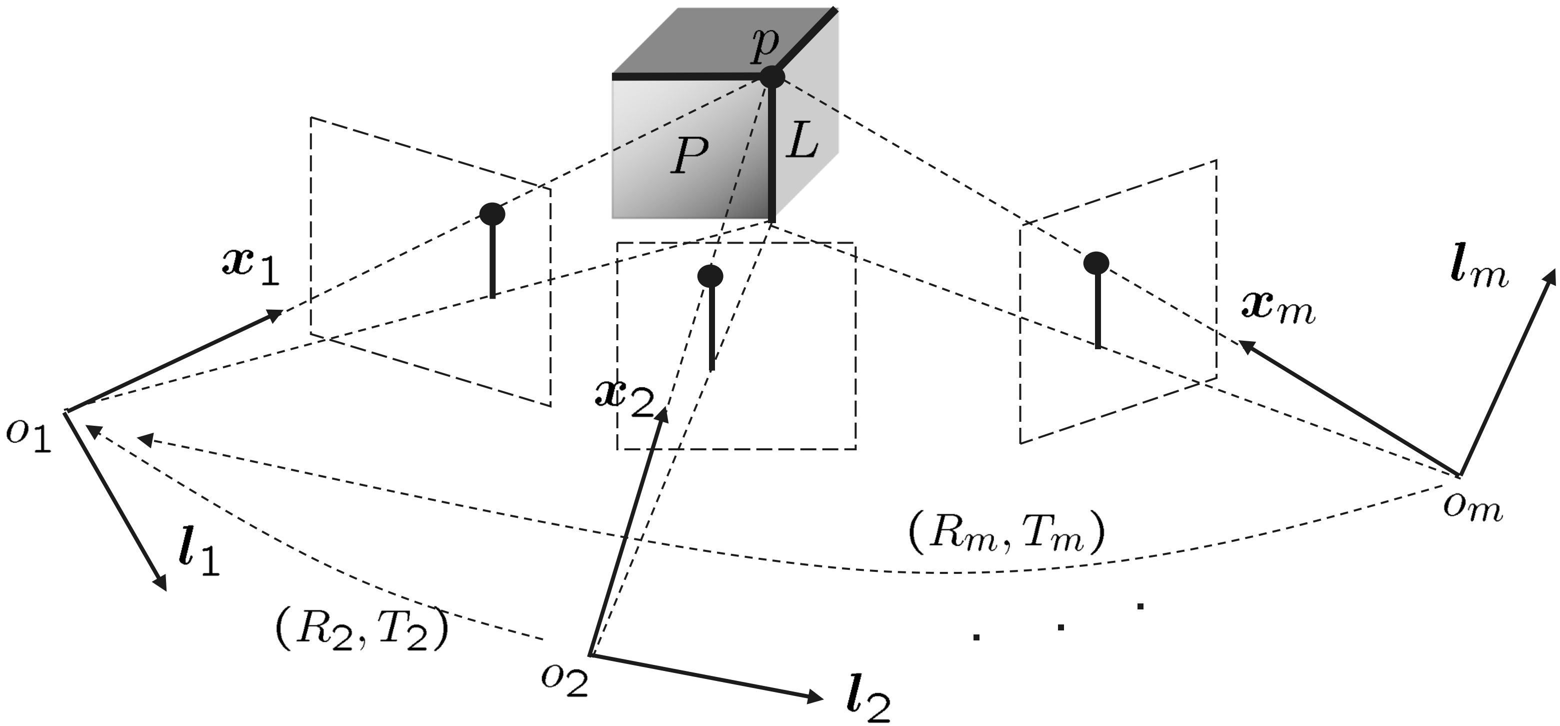
$$\boldsymbol{l} = [a, b, c]^T \in \mathbb{R}^3$$

Projection of a 3-D line to an image plane

$$\boxed{\boldsymbol{l}^T \boldsymbol{x} = \boldsymbol{l}^T \Pi \boldsymbol{X} = 0}$$

$$\Pi = [KR, KT] \in \mathbb{R}^{3 \times 4}$$

SUMMARY OF NOTATION – **Multiple Images**



1. Images are all "incident" at the corresponding features in space;
2. Features in space have many types of incidence relationships;
3. Features in space have many types of metric relationships.