Image Formation

Jana Kosecka
kosecka@gmu.edu

- pinhole camera
- perspective projection
- coordinate frames (world, camera, sensor)
- thin lens model, field of view, depth of field
- radial distortion

Image Formation

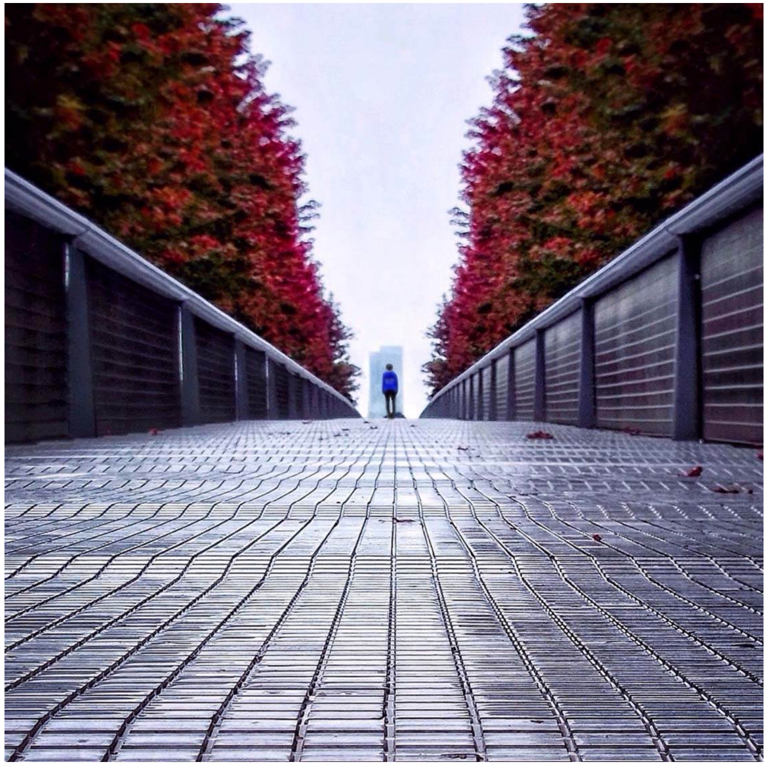
- If the object is our lens the refracted light causes the images

- How to integrate the information from all the rays being reflected from the single point on the surface ?

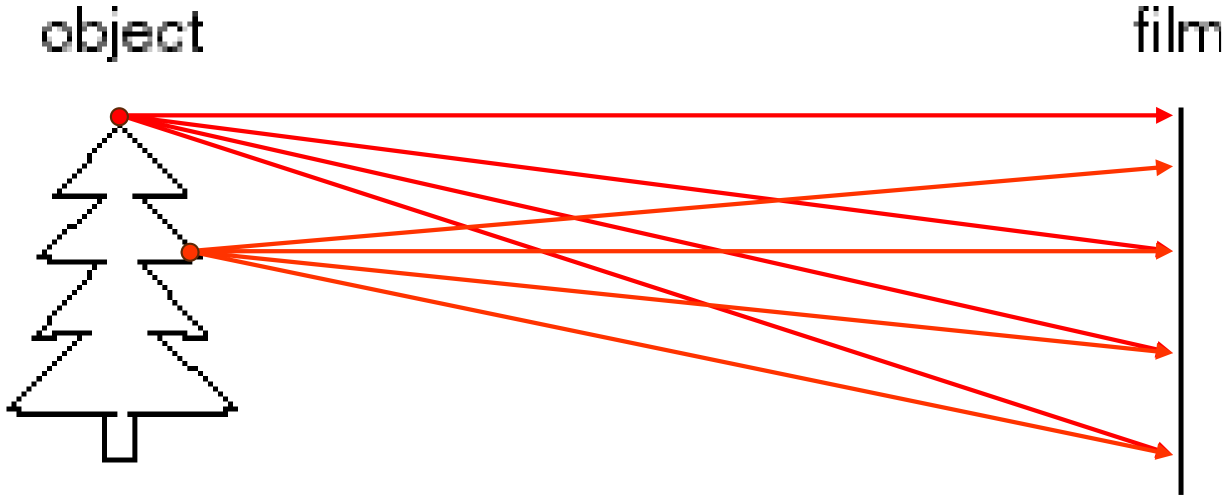
- Depending in their angle of incidence, some are more refracted then others – refracted rays all meet at the point – basic principles of lenses

- Also light from different surface points may hit the same lens point but they are refracted differently - Kepler's retinal theory


photopills.com

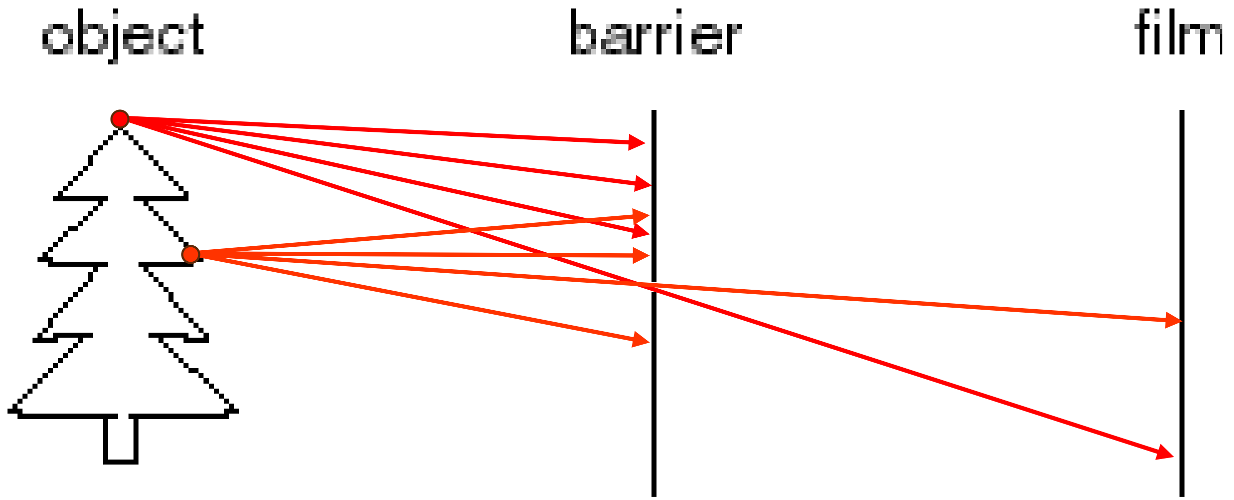


Let's design a camera



- Idea 1: put a piece of film in front of an object
- Do we get a reasonable image?

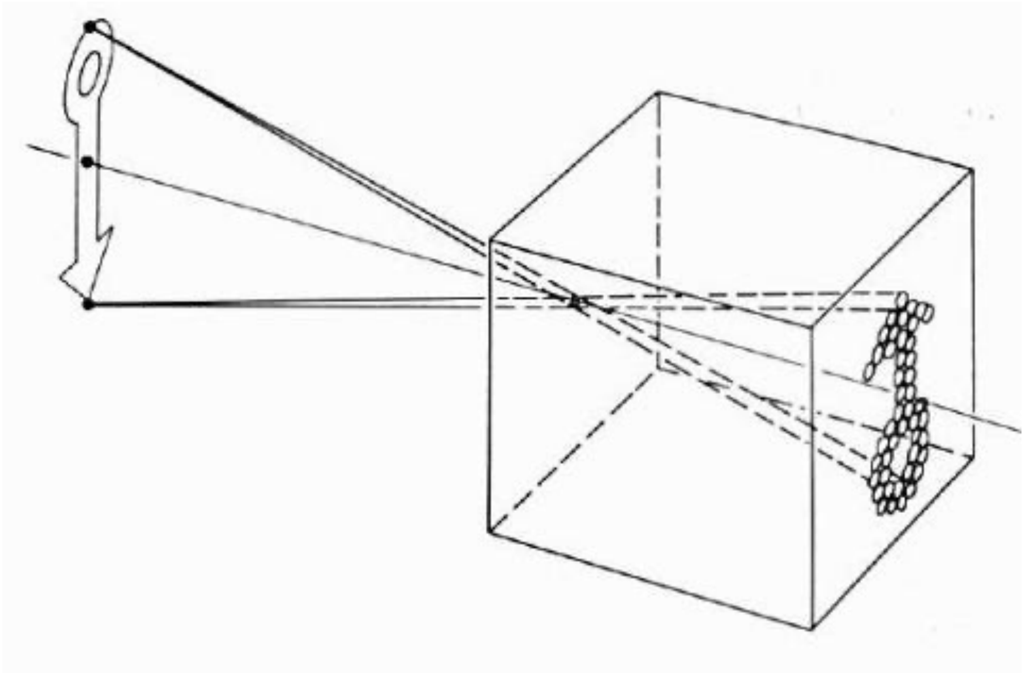
Pinhole camera



- Add a barrier to block off most of the rays
  - This reduces blurring
  - The opening is known as the **aperture**

Slide by Steve Seitz

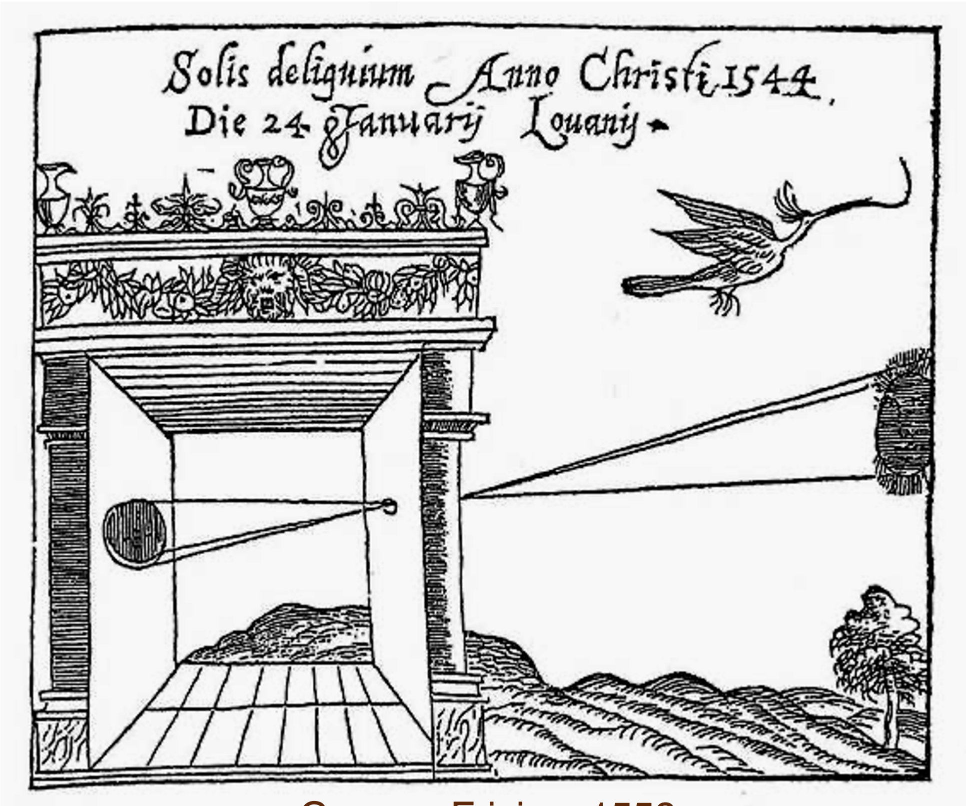
Pinhole camera model



- Pinhole model:
  - Captures **pencil of rays** – all rays through a single point
  - The point is called **Center of Projection (focal point)**
  - The image is formed on the **Image Plane**

Slide by Steve Seitz

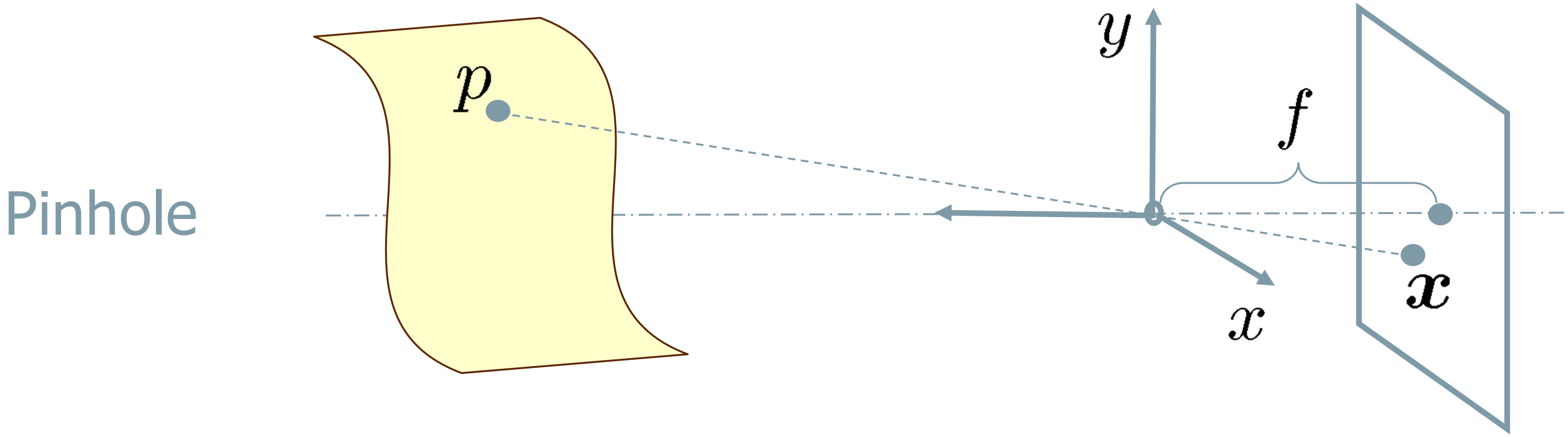
Camera Obscura



Gemma Frisius, 1558

- Basic principle known to Mozi (470-390 BCE), Aristotle (384-322 BCE)
- Drawing aid for artists: described by Leonardo da Vinci (1452-1519)

Pinhole Camera Model

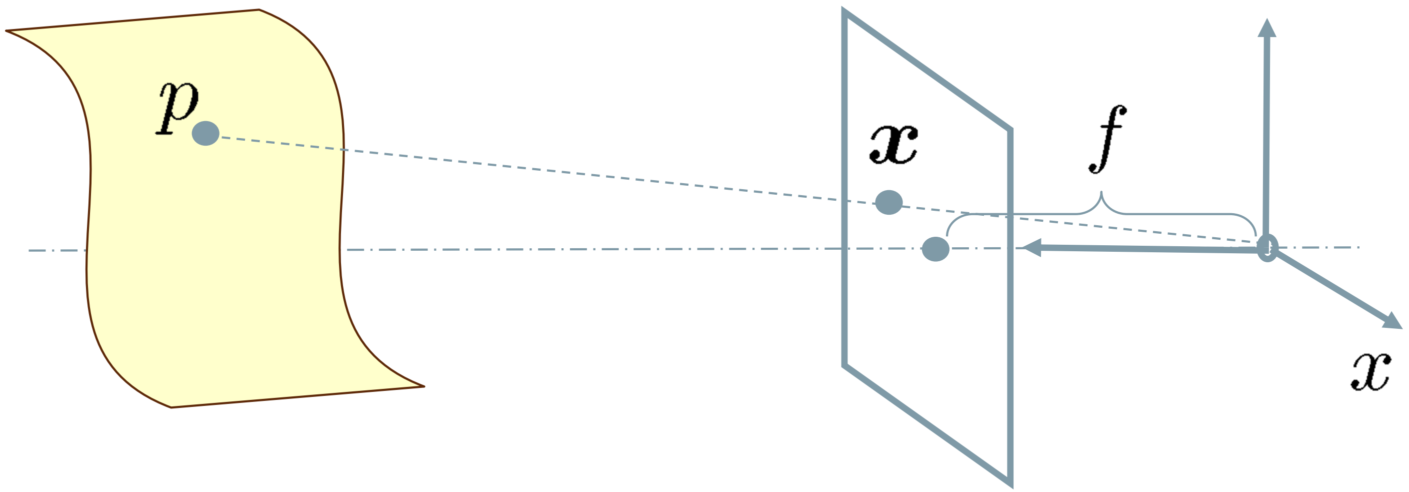


$$\boldsymbol{X} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \rightarrow \boldsymbol{x} = \begin{bmatrix} x \\ y \end{bmatrix} = \frac{f}{Z} \begin{bmatrix} X \\ Y \end{bmatrix}$$

Pinhole Camera Model

Note for the image plane moved in front of the center of projection (aperture) – the image will not be flipped

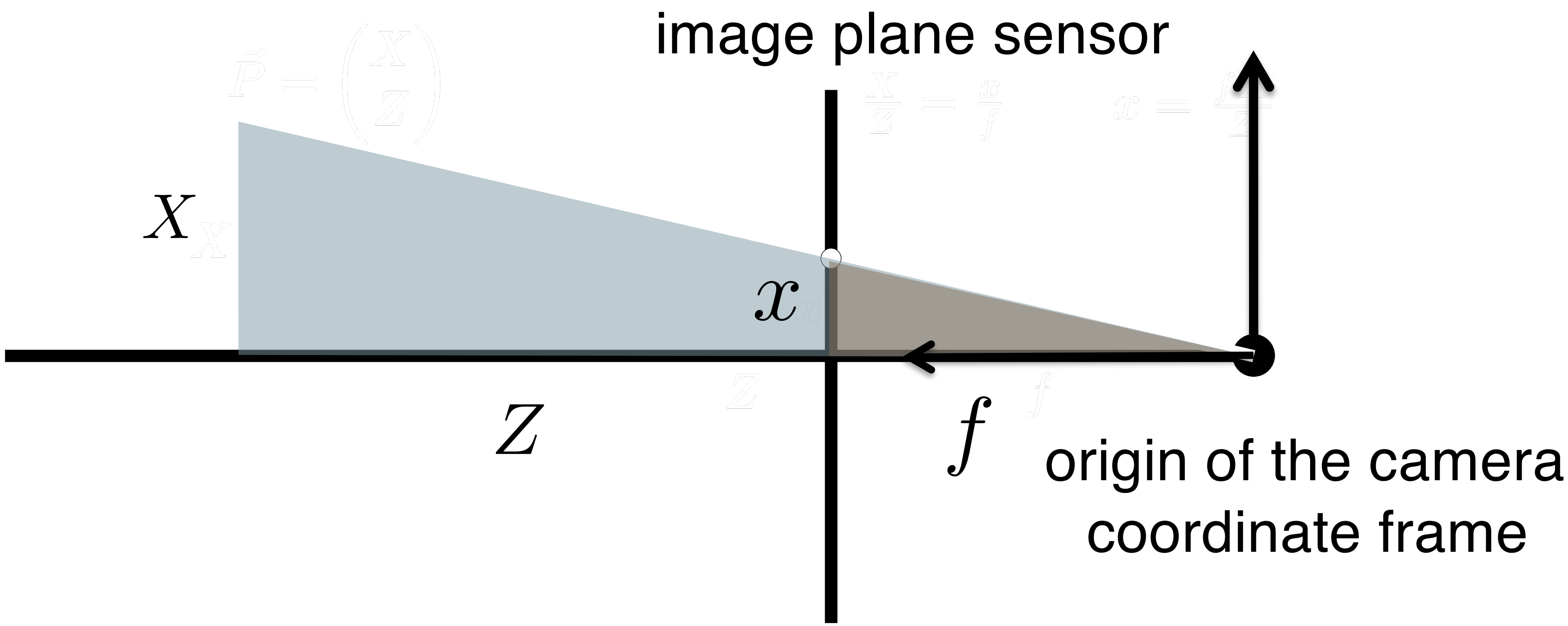
Frontal pinhole



Pinhole Camera Model

By similarity of triangles

$$P = \begin{bmatrix} X \\ Z \end{bmatrix} \qquad \frac{X}{Z} = \frac{x}{f} \qquad \text{solve for x} \qquad x = \frac{fX}{Z}$$



Here the relationship between x and X is non-linear

Pinhole Camera Model
Homogeneous coordinates

Image coordinates are nonlinear function of coordinates
In the camera coordinate frame

$$x = \frac{fX}{Z} \quad f = 1 \quad x = \frac{X}{Z} \quad \text{and} \quad y = \frac{Y}{Z}$$

Homogenous coordinates in 2D – add dimension

$$\begin{bmatrix} x \\ y \end{bmatrix} \rightarrow \begin{bmatrix} u \\ v \\ w \end{bmatrix} \rightarrow \begin{bmatrix} u/w \\ v/w \\ 1 \end{bmatrix}$$

Physical points

Divide by 3rd coordinate
What if *w = 0* ?

Points and lines

Homogeneous coordinates of a point are equivalent
If they are proportional

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \lambda \begin{bmatrix} u' \\ v' \\ w' \end{bmatrix}$$

Advantage of homogenous coordinates – line equation

General equation of a line

$$\boldsymbol{l}^T \boldsymbol{x} = ax + by + c = 0 \qquad \mathbf{l} = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \qquad \mathbf{x} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Points and lines

Homogeneous coordinates of a point are equivalent
If they are proportional

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \lambda \begin{bmatrix} u' \\ v' \\ w' \end{bmatrix}$$

Advantage of homogenous coordinates – line equation

General equation of a line in 2D

$$\boldsymbol{l}^T \boldsymbol{x} = ax + by + c = 0$$

$$\mathbf{l} = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \qquad \mathbf{x} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

set of all points x that satisfy
the equation

Image of a line – homogeneous representation

Homogeneous representation of and 2D image of a line

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

cro

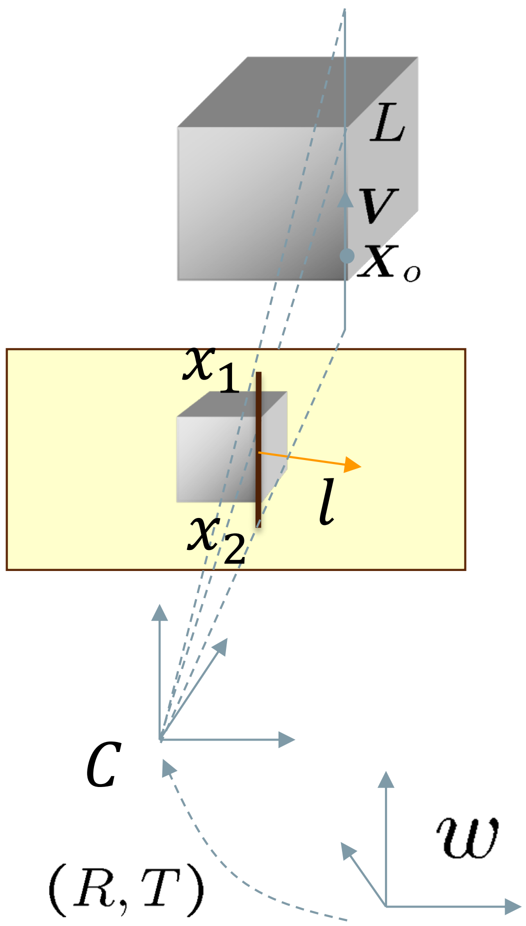
Is plane normal spanned by $C, \boldsymbol{x}_2, \boldsymbol{x}_1$

Equation of a line that goes through two points $\boldsymbol{x}_2, \boldsymbol{x}_1$

$$\boldsymbol{l} = \boldsymbol{x_1} \times \boldsymbol{x_2}$$

Intersection of two lines $\boldsymbol{l}, \boldsymbol{m}$ $is\ a\ point$

$$\boldsymbol{p} = \boldsymbol{l} \times \boldsymbol{m}$$

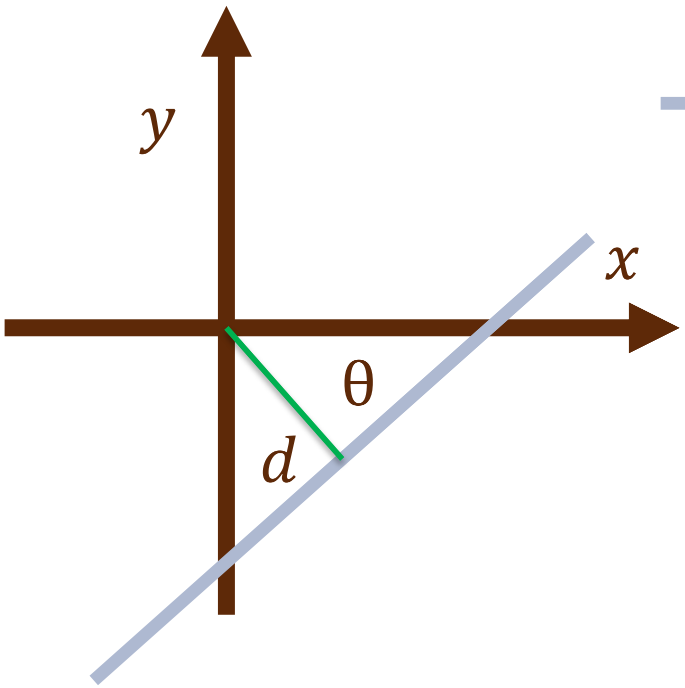


2D lines – homogeneous representation

Polar line representation

$$x \cos\theta + y \sin\theta = d$$

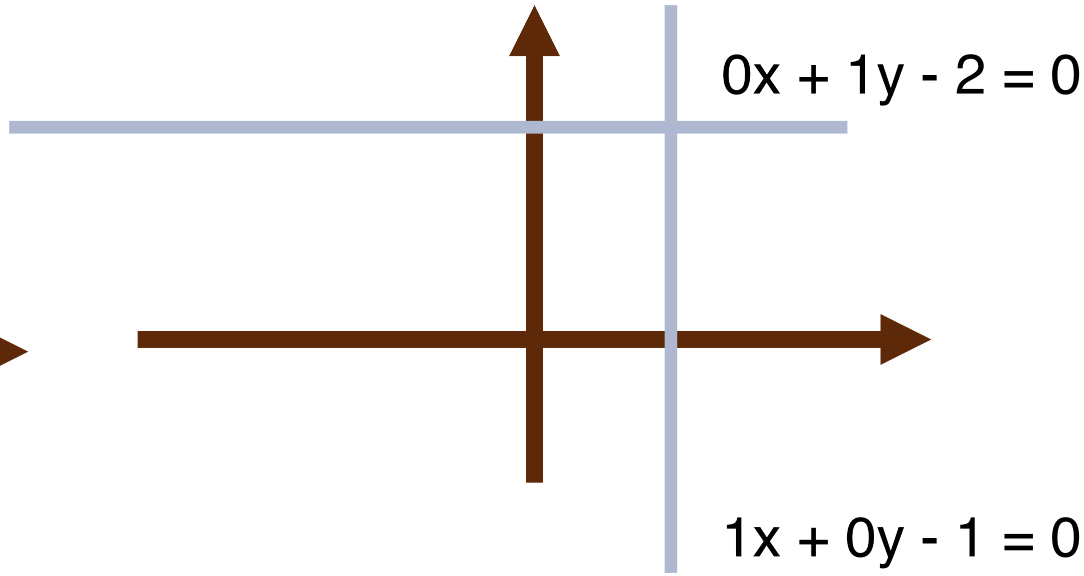
$$[a, b, c][x, y, 1]^t = 0$$



Slope and intercept representation

$$y = mx + b$$

What's the intersection?



[0,1,-2] x [1,0,-1] = [-1,-2,-1]
Converting back (divide by -1)
(1,2)

Recall Vector (cross) Product

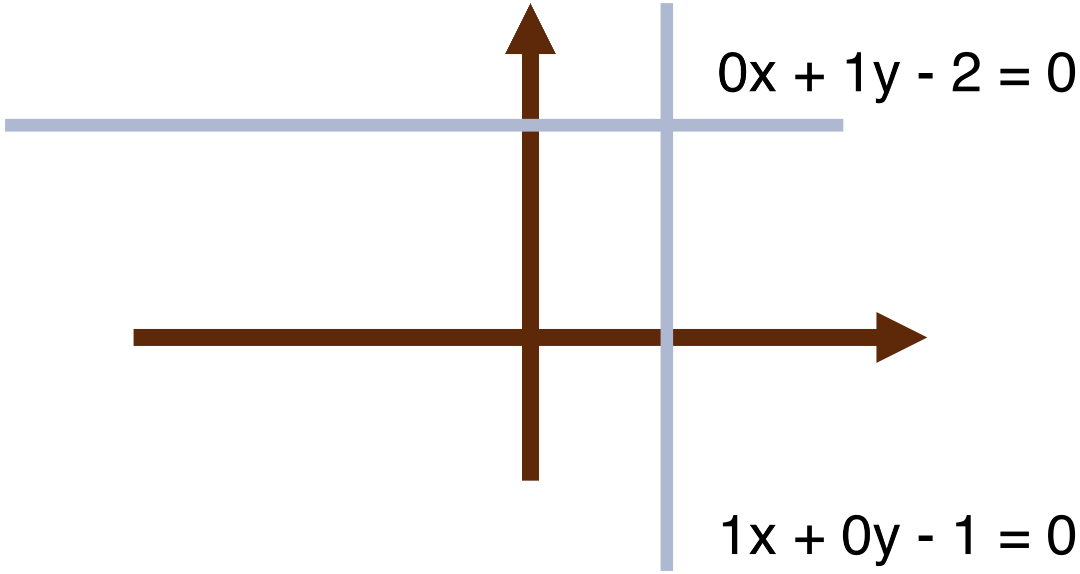
- Cross product between two vectors in $c = a \times b$

where

$$c = a \times b = \hat{a}b = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

$$\hat{a} = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix} \quad c = \begin{bmatrix} -a_3 b_2 + a_2 b_3 \\ a_3 b_1 - a_1 b_3 \\ -a_2 b_1 + a_1 b_2 \end{bmatrix}$$

2D lines – homogeneous representation

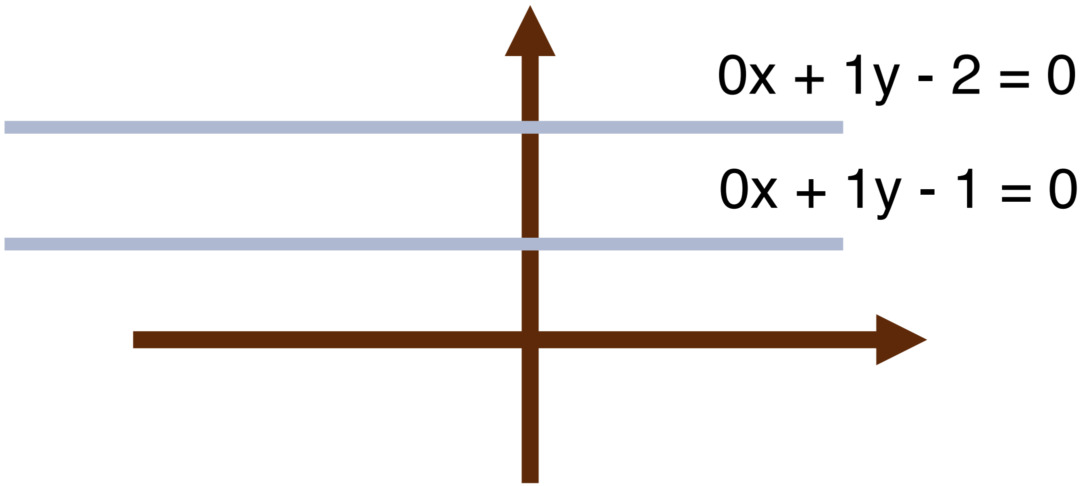


[0,1,-2] x [1,0,-1] = [-1,-2,-1]
Converting back (divide by -1)
(1,2)

Cross product above:

$$c = [0, 1, -2] \times [1, 0, -1] = \hat{a}b = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 2 \\ -1 & -2 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ -2 \\ -1 \end{bmatrix}$$

2D lines – homogeneous representation



[0,1,-2] x [0,1,-1] = [1,0,0]
Third coordinate is zero – point at infinity

Pinhole Camera Model

- Image coordinates are nonlinear function of world coordinates
- Relationship between coordinates in the camera frame and sensor plane

2-D coordinates

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

Perspective projection in homogeneous coordinates

divide by 3rd coordinate

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

same as

$$x = \frac{fX}{Z}$$

$$y = \frac{fY}{Z}$$

Pinhole Camera Model

- Image coordinates are nonlinear function of world coordinates
- Relationship between coordinates in the camera frame and sensor plane

2-D coordinates

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

In the matrix form: each matrix represents stage of the projection

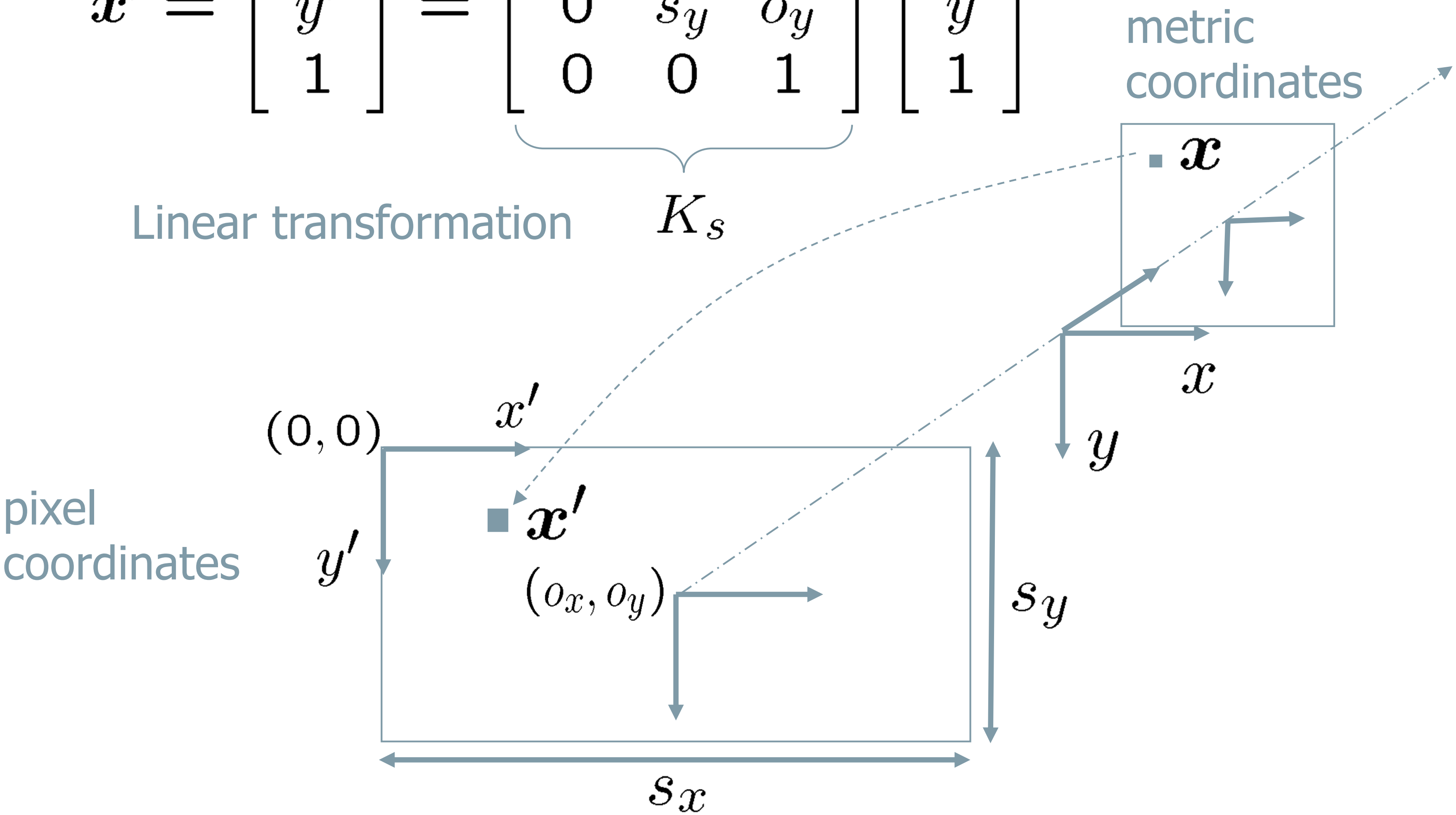
$$Z \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{K_f} \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{\Pi_0} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Image Coordinates

- Relationship between coordinates in the sensor plane and image

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

Linear transformation



Calibration Matrix and Camera Model

- Relationship between coordinates in the camera frame and image

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)

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

Projection matrix

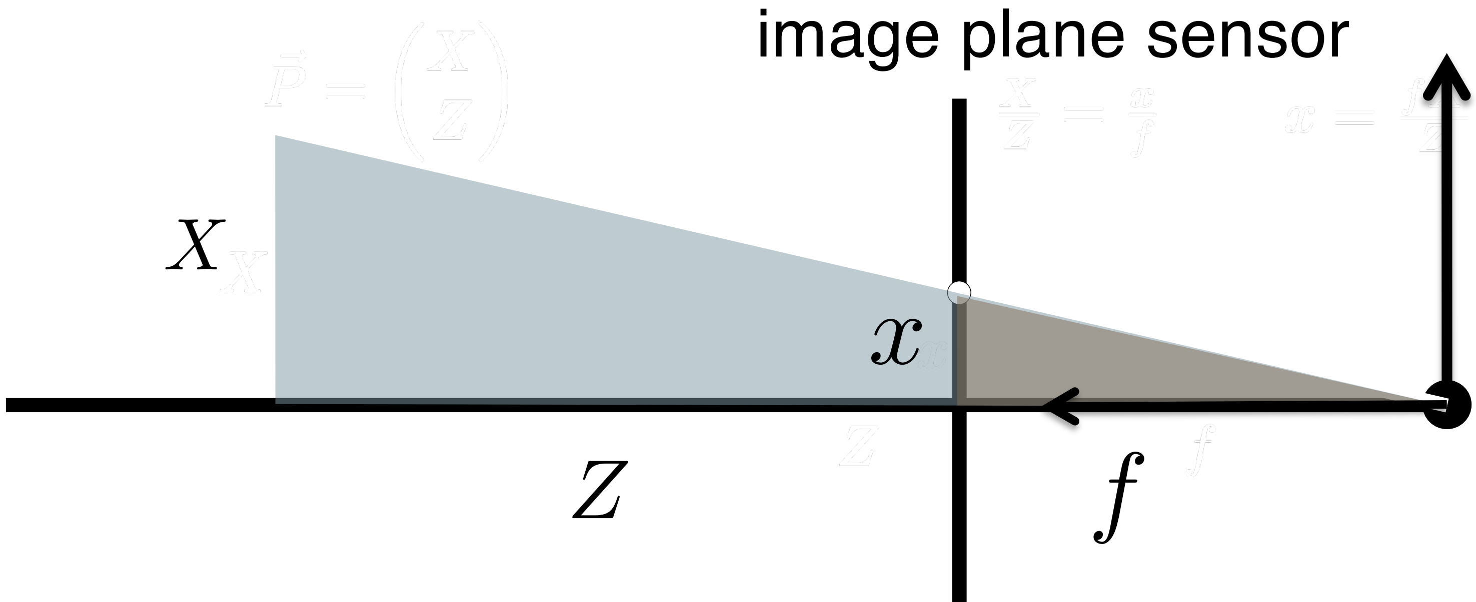
$$\Pi = [K, 0] \in \mathbb{R}^{3\times 4}$$

Camera model

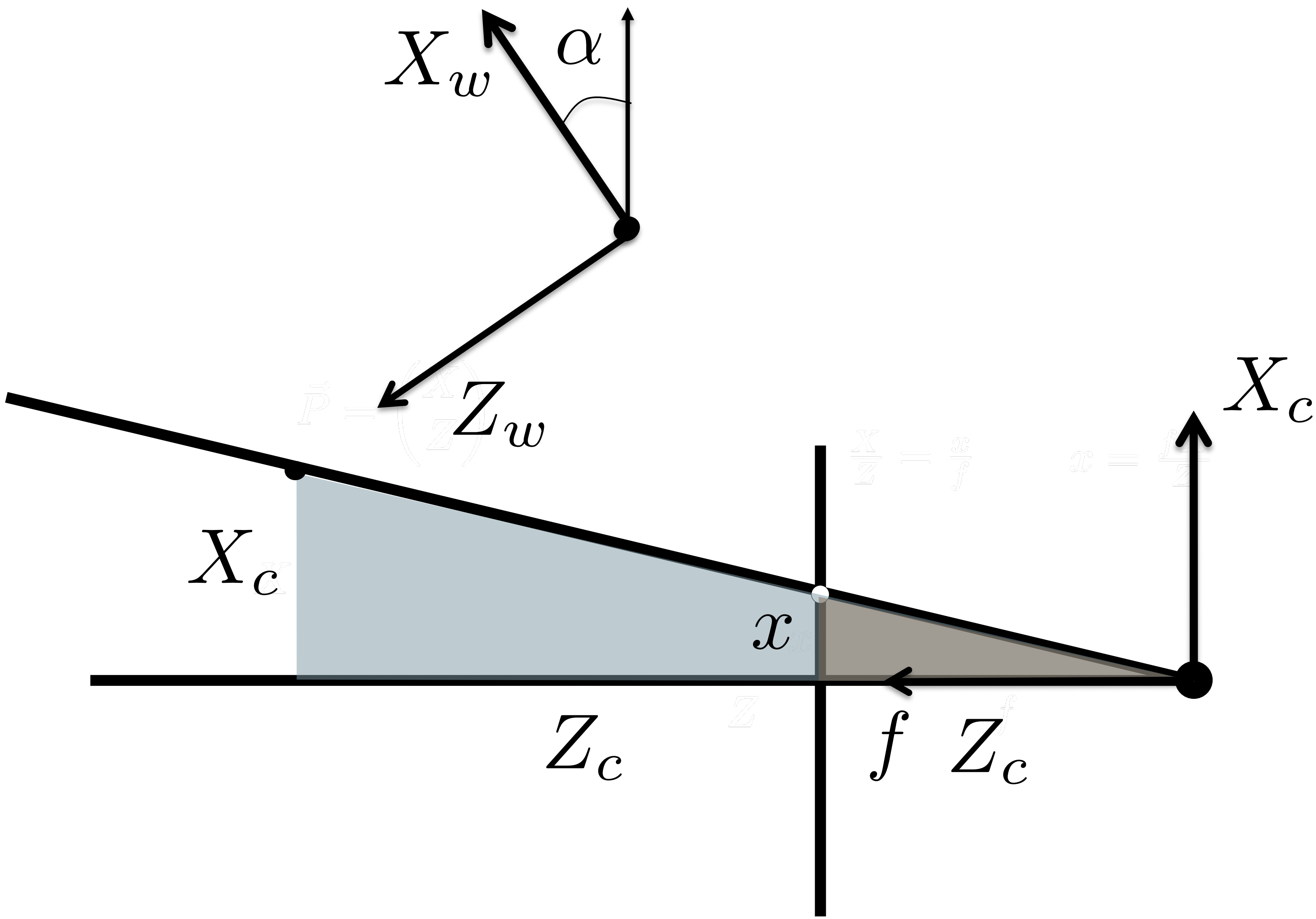
$$\lambda \boldsymbol{x}' = K \Pi_0 \boldsymbol{X} = \Pi \boldsymbol{X}$$

Pinhole Camera Model

$$P = \begin{bmatrix} X \\ Z \end{bmatrix} \qquad \frac{X}{Z} = \frac{x}{f} \qquad x = \frac{fX}{Z}$$

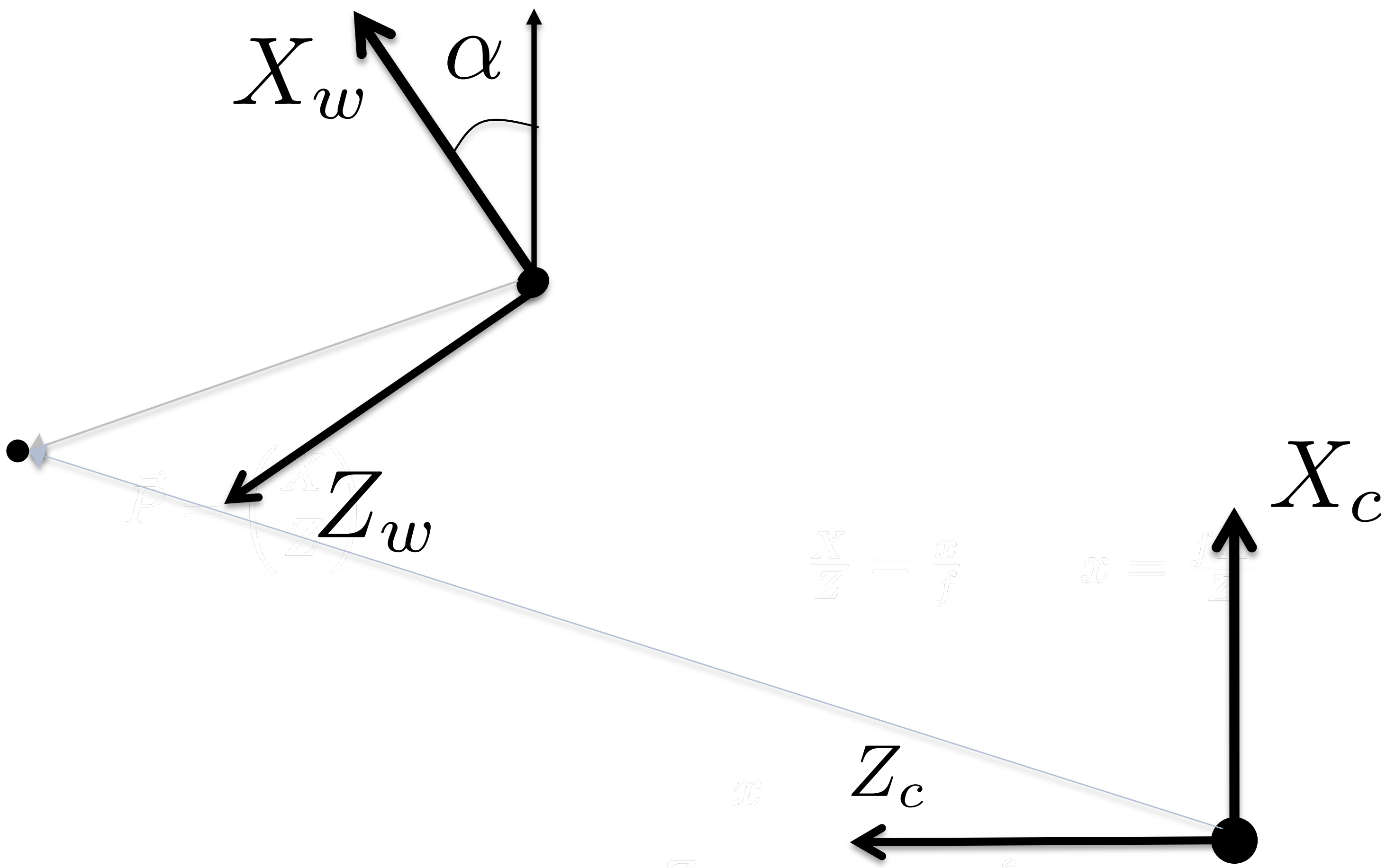


$$P = \lambda \begin{bmatrix} X \\ Z \end{bmatrix} \qquad x = \frac{fX}{Z} = \frac{f\lambda X}{\lambda Z}$$



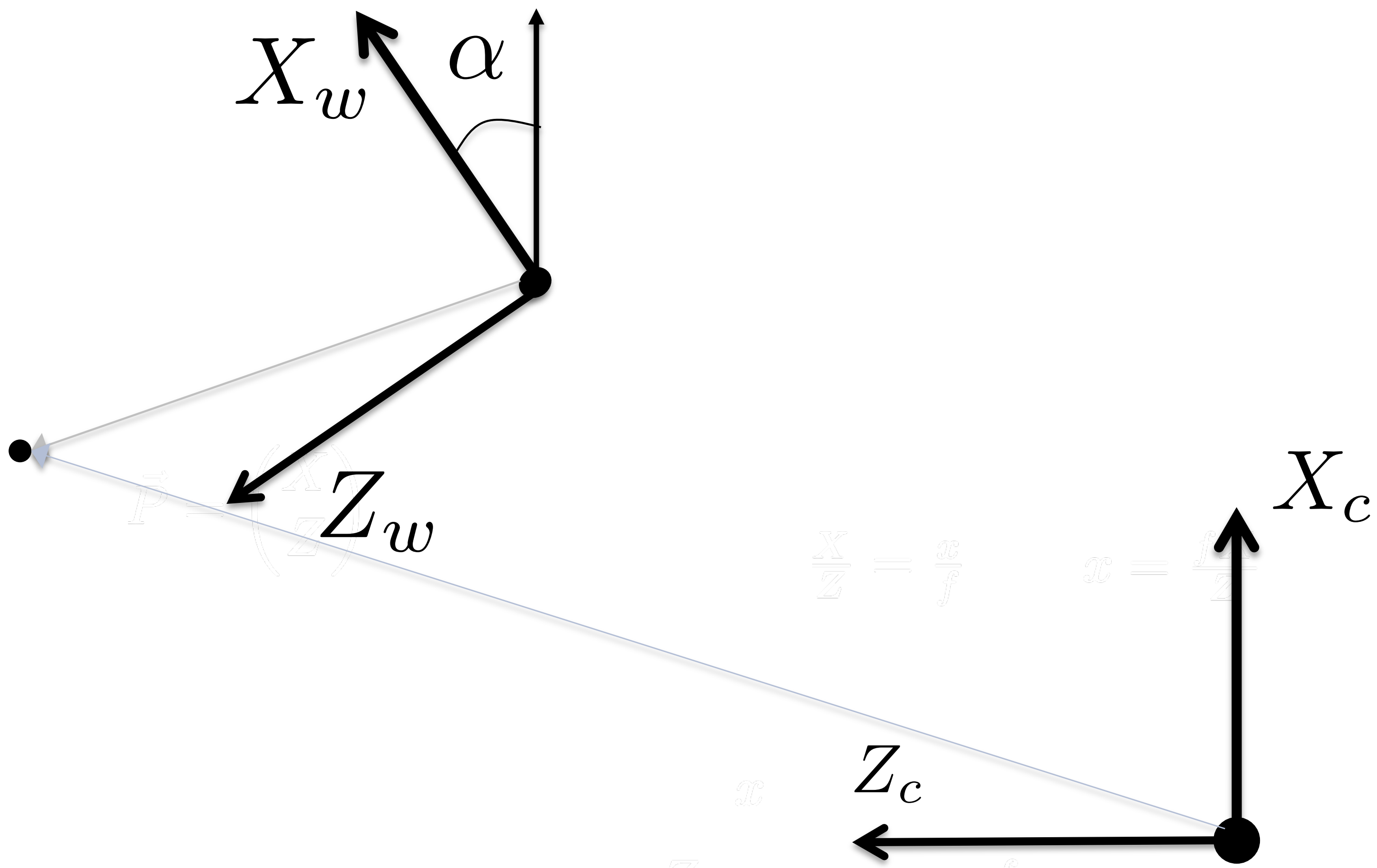
X_w
α
Z_w
X_c
X_c
x
Z_c
f
Z_c

2D Rigid Body Motion



$$\begin{bmatrix} X_c \\ Z_c \end{bmatrix} = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} X_w \\ Z_w \end{bmatrix} + \begin{bmatrix} t_x \\ t_z \end{bmatrix}$$

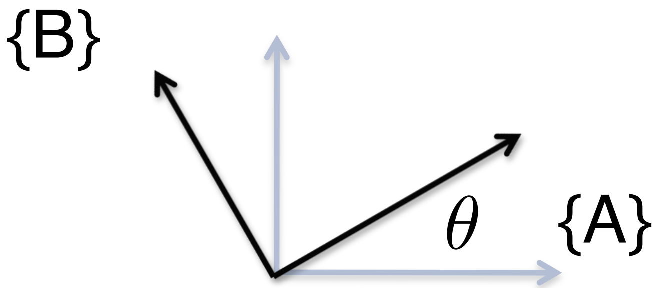
2D Rigid Body Motion in homogeneous coordinates



$$\begin{bmatrix} X_c \\ Z_c \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\alpha & -\sin\alpha & t_x \\ \sin\alpha & \cos\alpha & t_z \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Z_w \\ 1 \end{bmatrix}$$

2D Rotation Matrix

Interpretations of the rotation matrix R_{AB}



$$R_{AB} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

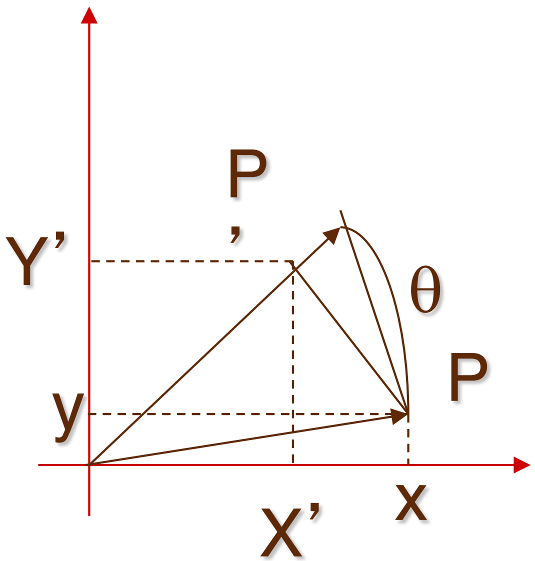
Columns of R_{AB} are the unit vectors of the axes of frame B expressed in coordinate frame A. Such rotation matrix transforms coordinates of points in frame B to points in frame A

Use of the rotation matrix as transformation R_{AB}

$$\mathbf{X}_A = R_{AB}\mathbf{X}_B$$

2D Rotation Matrix

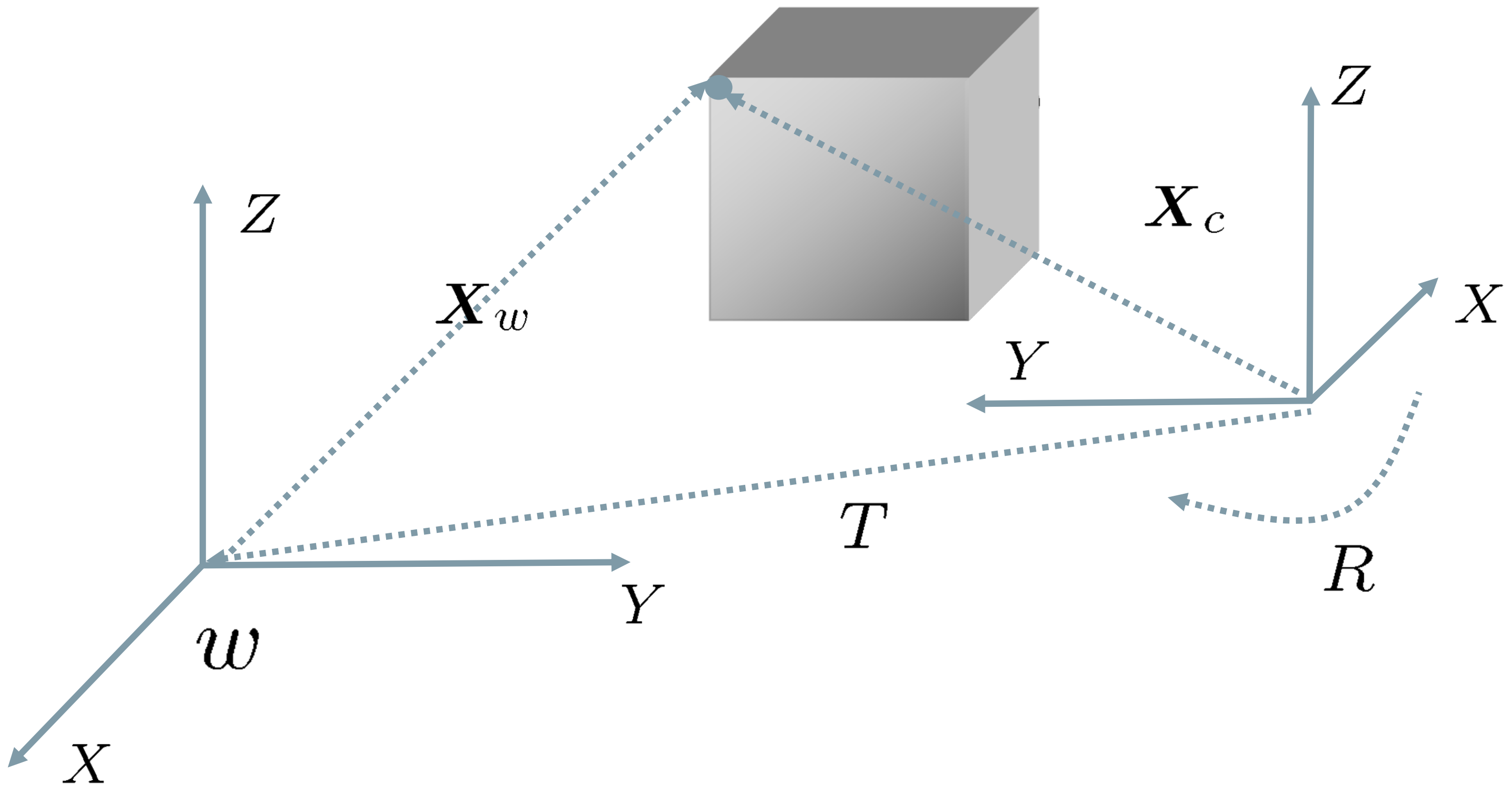
Counter-clockwise rotation of a point by an angle θ



$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Counter-clockwise rotation of a coordinate frame attached to a rigid body by an angle θ

3D Rigid Body Transformation

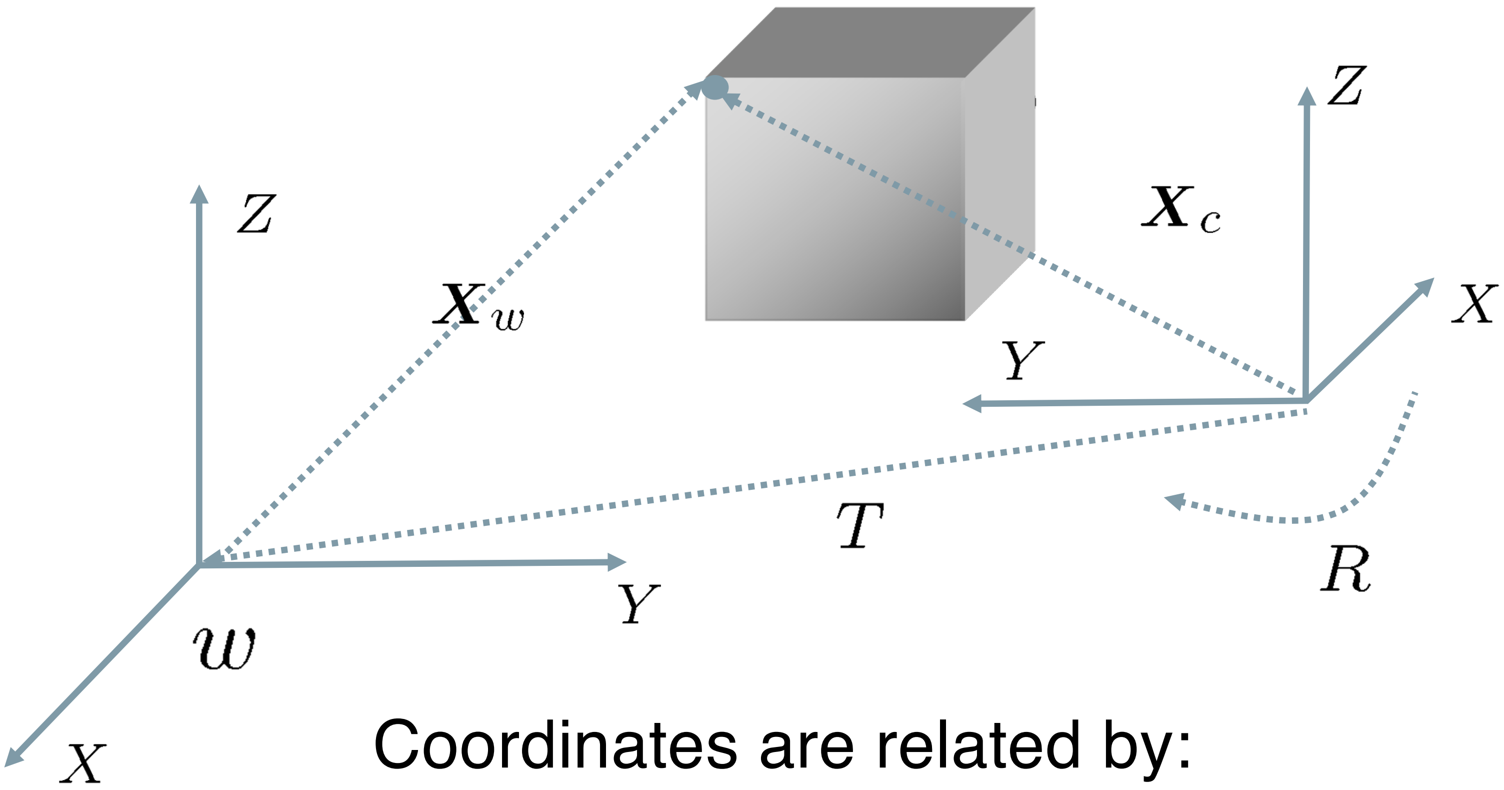


What is camera moves ?

{w} world coordinate frame

{c} camera coordinate frame

3D Rigid Body Transformation



Coordinates are related by:

$$\boldsymbol{X}_c = R\boldsymbol{X}_w + T,$$

Camera pose is specified by: $g = (R, T) \in SE(3)$

3D Rigid Body Motion in Homogeneous Coordinates

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 are related by:

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

3D Rotation of Points – Euler angles

Where rotation matrix is:

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}$$

$$R = R_z(\gamma) R_y(\beta) R_x(\alpha)$$

Rotation around the coordinate axes, counter-clockwise:

$$\begin{bmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{bmatrix}$$

$R_z(\theta)$ $R_x(\alpha)$ $R_x(\alpha)$

Properties of Rigid Body Motions

Rigid body motion composition

$$\bar{g}_1\bar{g}_2 = \begin{bmatrix} R_1 & T_1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} R_2 & T_2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} R_1R_2 & R_1T_2 + T_1 \\ 0 & 1 \end{bmatrix} \quad \in SE(3)$$

Rigid body motion inverse

$$\bar{g}^{-1} = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} R^T & -R^T T \\ 0 & 1 \end{bmatrix} \quad \in SE(3).$$

Rigid body motion acting on vectors

Vectors are only affected by rotation – 4^{th} homogeneous coordinate is zero

Calibration Matrix and Camera Model

- Relationship between coordinates in the world frame and image

Pinhole camera

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

Pixel coordinates

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

$$\lambda \boldsymbol{x}' = \begin{bmatrix} f s_x & f s_\theta & o_x \\ 0 & f s_y & o_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

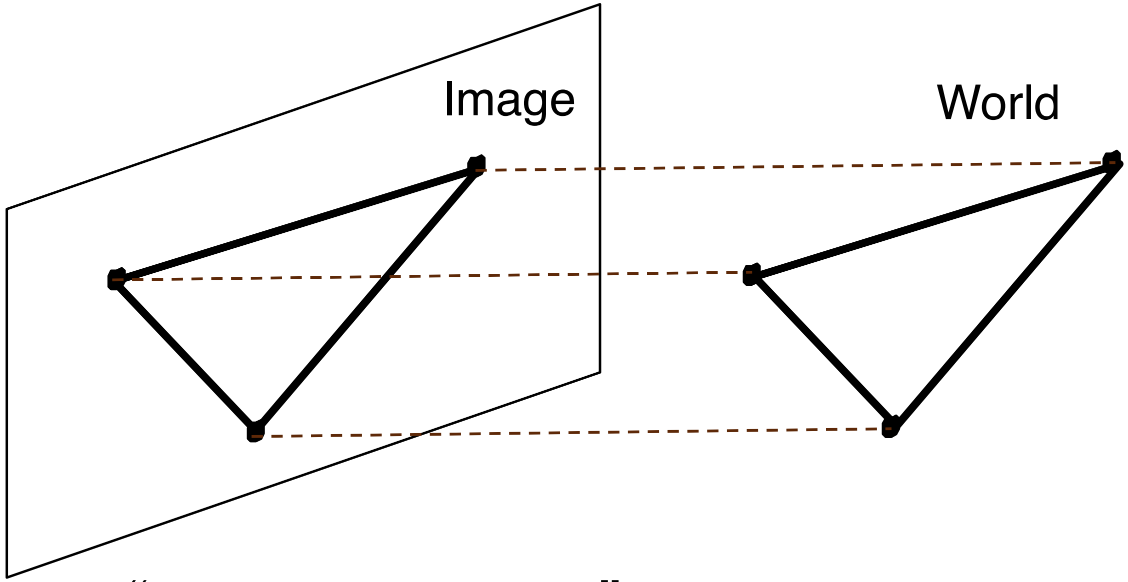
$$\lambda \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} f s_x & f_{s_\theta} & o_y \\ 0 & f s_y & o_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

More compactly $\lambda \boldsymbol{x} = K_f \Pi_0 g \boldsymbol{X} = \Pi \boldsymbol{X}$

Transformation between camera coordinate systems and world coordinate system

Orthographic Projection

- Special case of perspective projection
  - Distance from center of projection to image plane is infinite



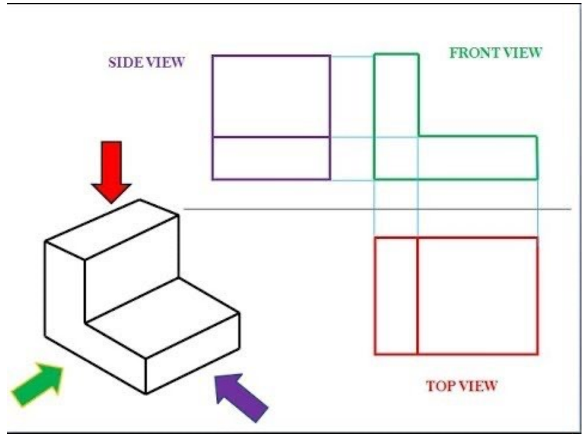
  - Also called “parallel projection”
  - What’s the projection matrix (in homogenous coordinates)?

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \Rightarrow (x, y)$$

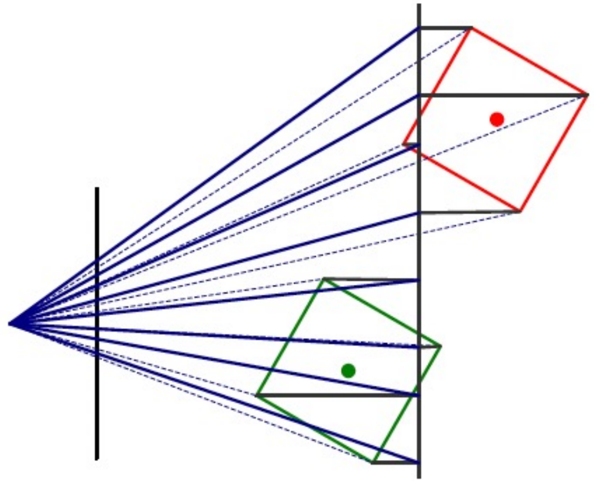
Slide by Steve Seitz

Orthographic Projection

- Special case of perspective projection
  - All objects appear at the same scale
  - Parallel lines remail parallel
  - Often used in engineering drawings



- Scaled orthographic projection
  - project to fronto-parallel plane
  - scale the image using regular projection



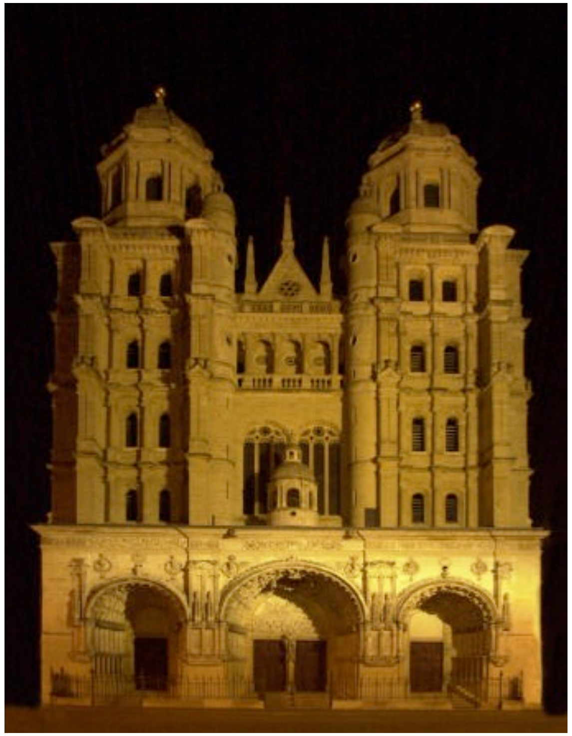
In non-homogeneous coordinates

$$\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} s & 0 & 0 \\ 0 & s & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

Adopted from R. Szeliski

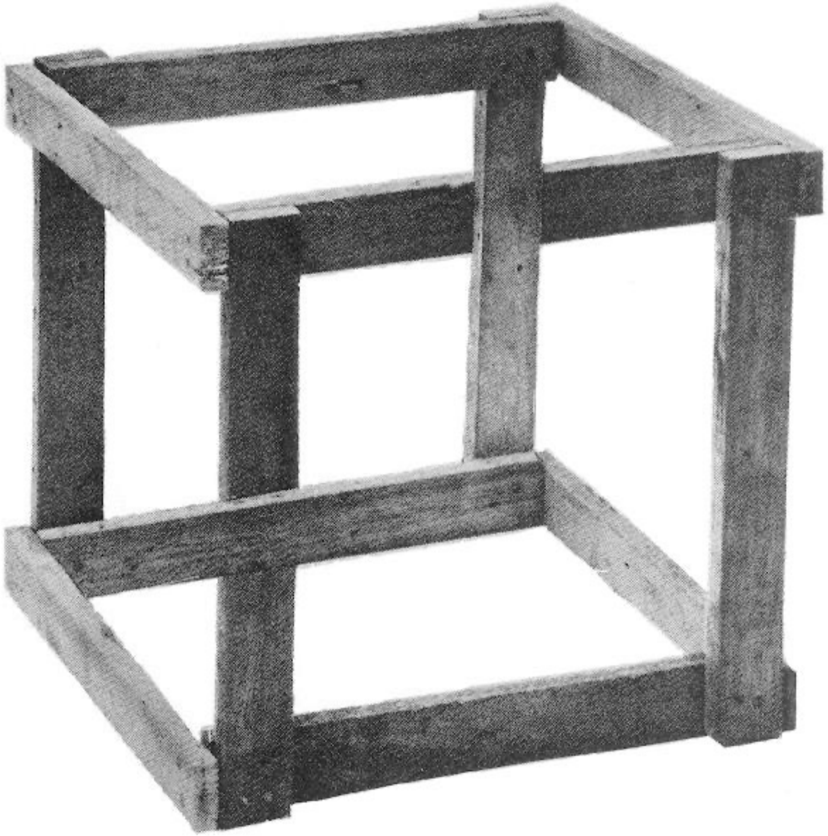
Perspective distortion

- Problem for architectural photography: converging verticals
- Parallel lines in the world are not parallel in the image
- Correction; design special cameras to correct the distortion
- Useful for architectural photography
- http://en.wikipedia.org/wiki/Perspective_correction_lens



Source: F. Durand

Visual Illusions, Wrong Perspective

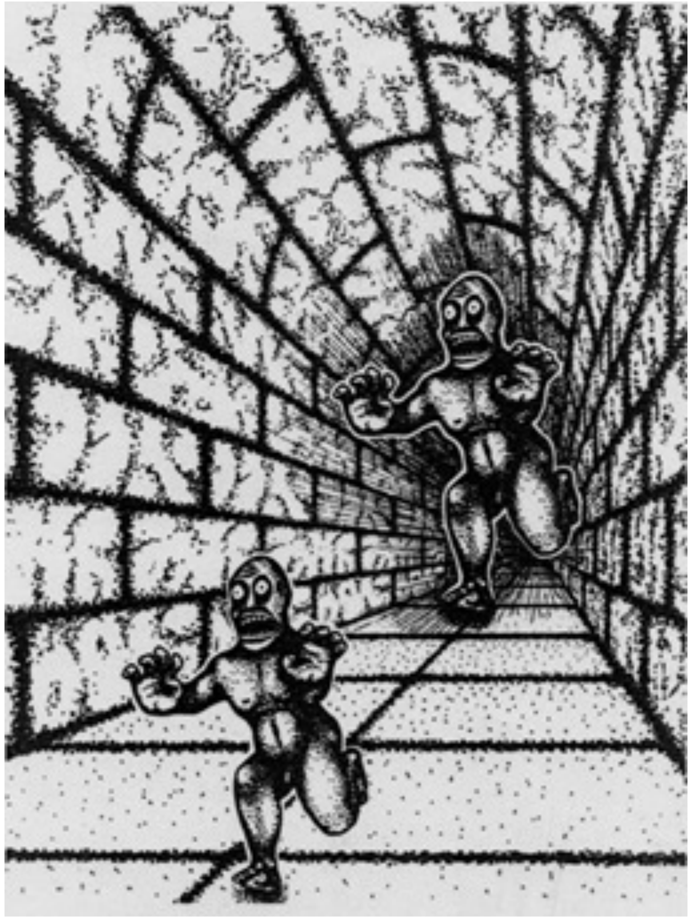


Vanishing points

Different sets of parallel lines in a plane intersect
at vanishing points, vanishing points form a horizon line

Ames Room Illusions

More Illusions



Which of the two monsters is bigger ?