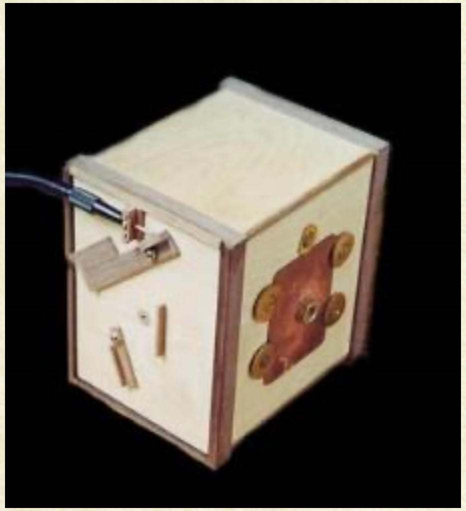
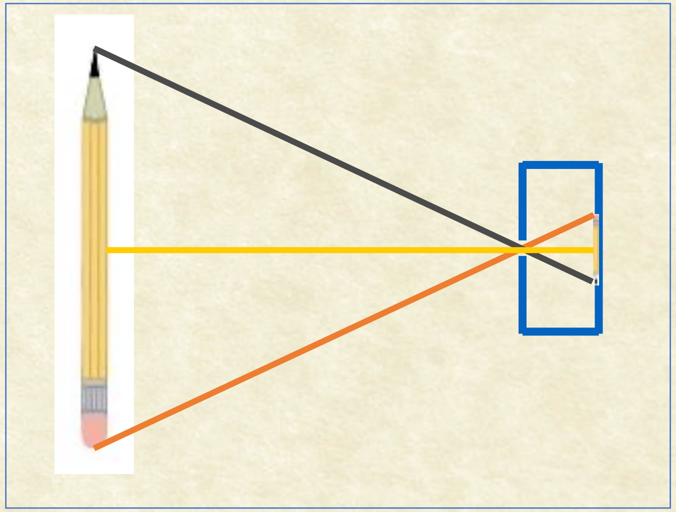
# CS7.505: Computer Vision

## Spring 2024: Pinhole Camera Model



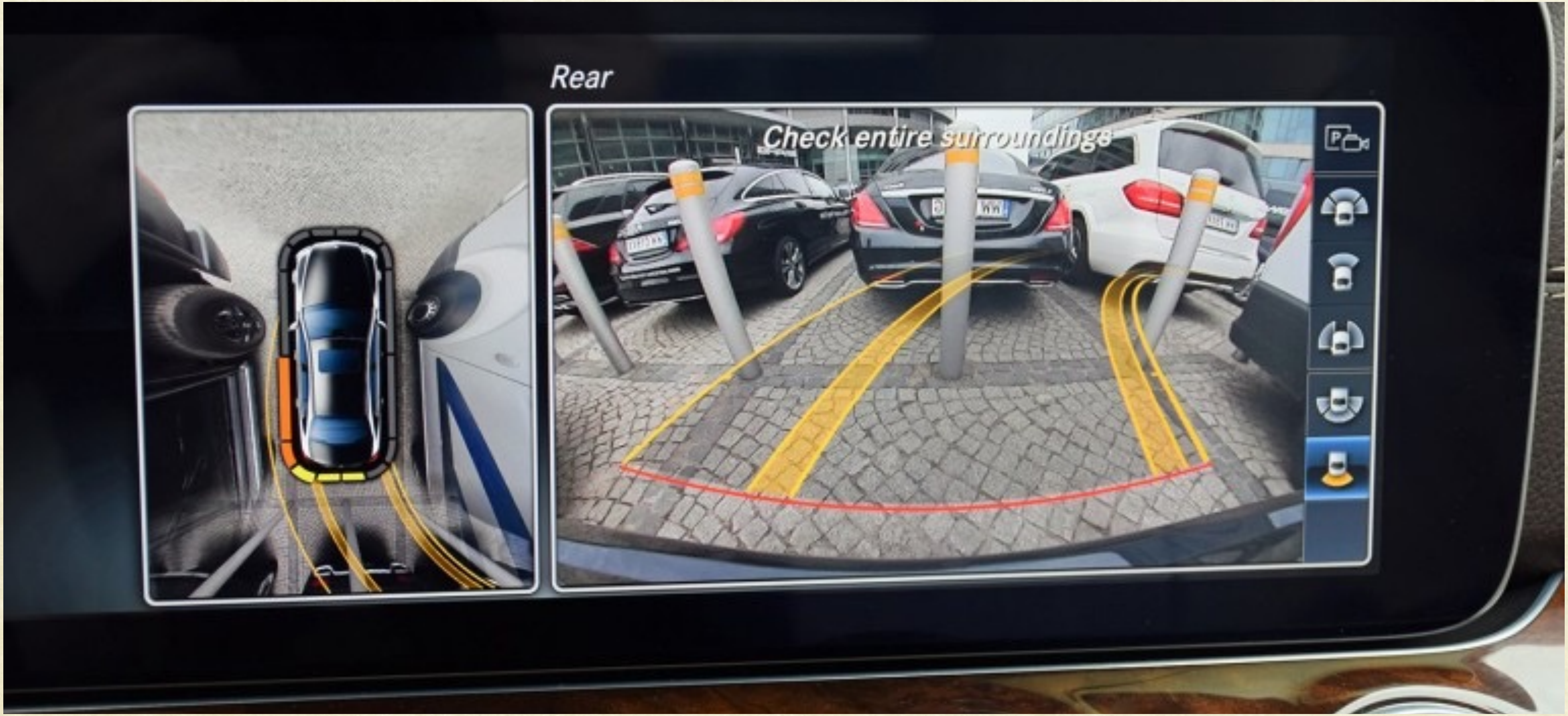



Makarand Tapaswi
Katha AI Lab, CVIT,
IIIT Hyderabad

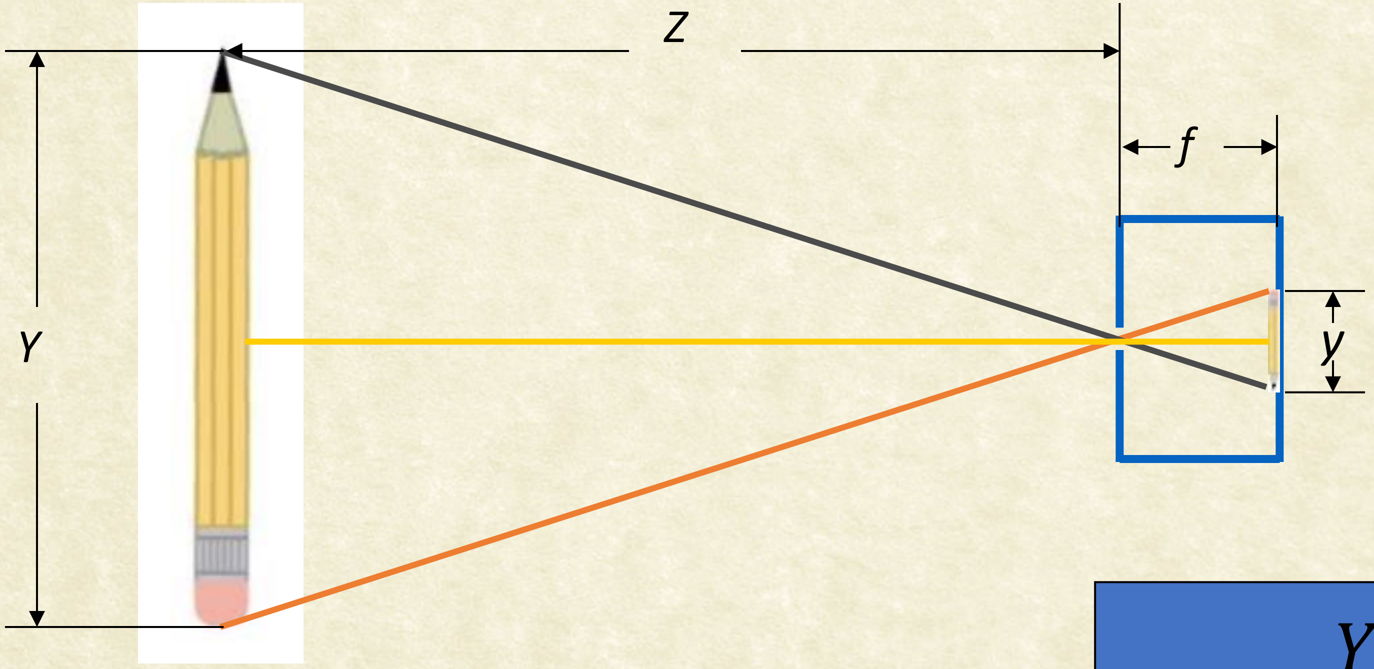
Anoop Namboodiri
Biometrics and Secure ID Lab, CVIT,
IIIT Hyderabad

# Projections and Homography

# Projections!



# The Pinhole Camera

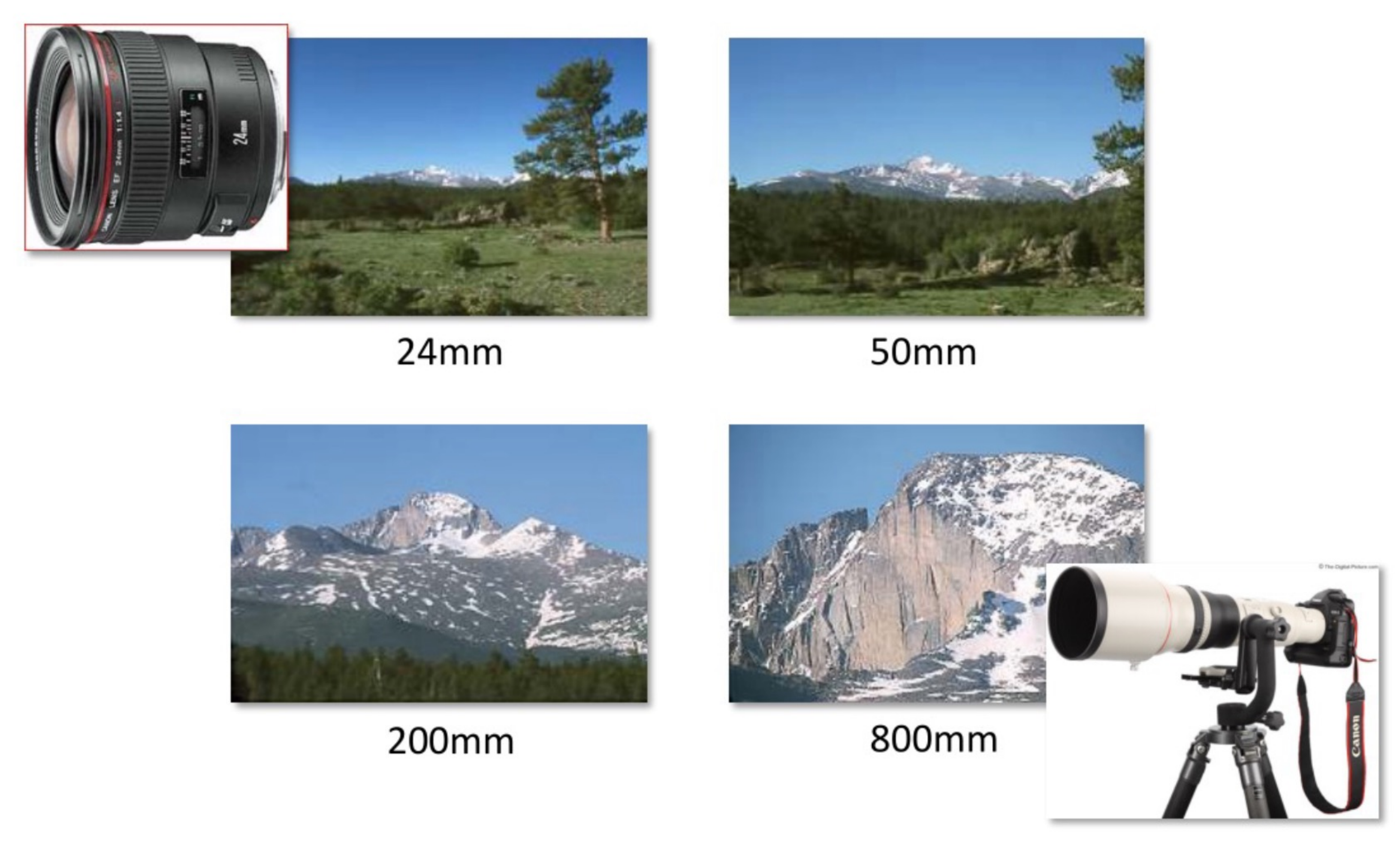



$$y = f\frac{Y}{Z}$$

# Pinhole camera recap

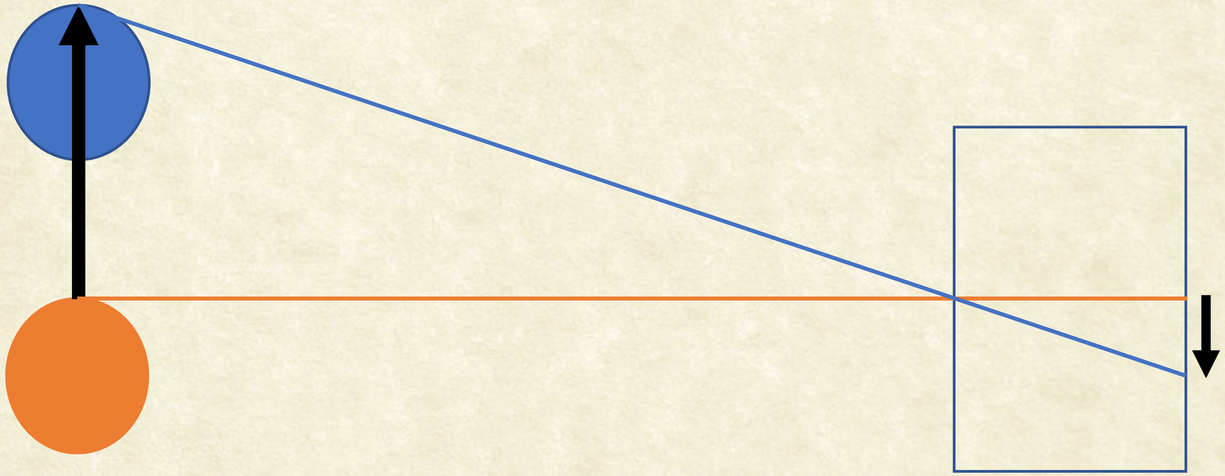
- http://www.cs.toronto.edu/~fidler/slides/2023Winter/CSC420/lecture10.pdf slides 7-37

# Field of view, relation to zoom

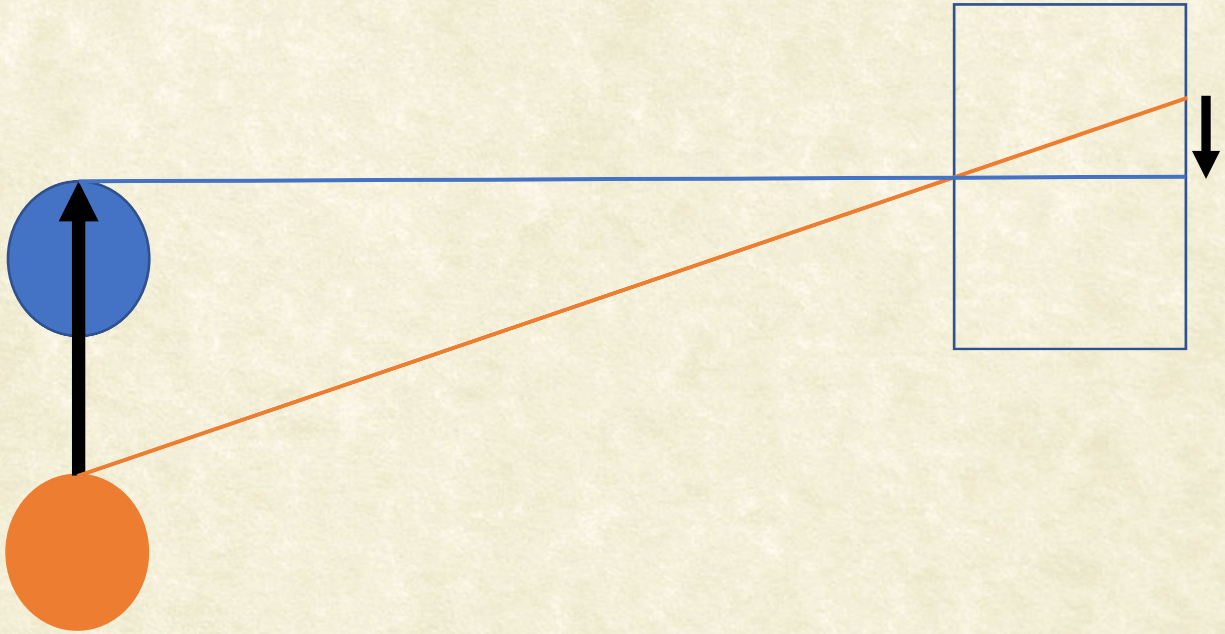



Slide credit: Noah Snavely

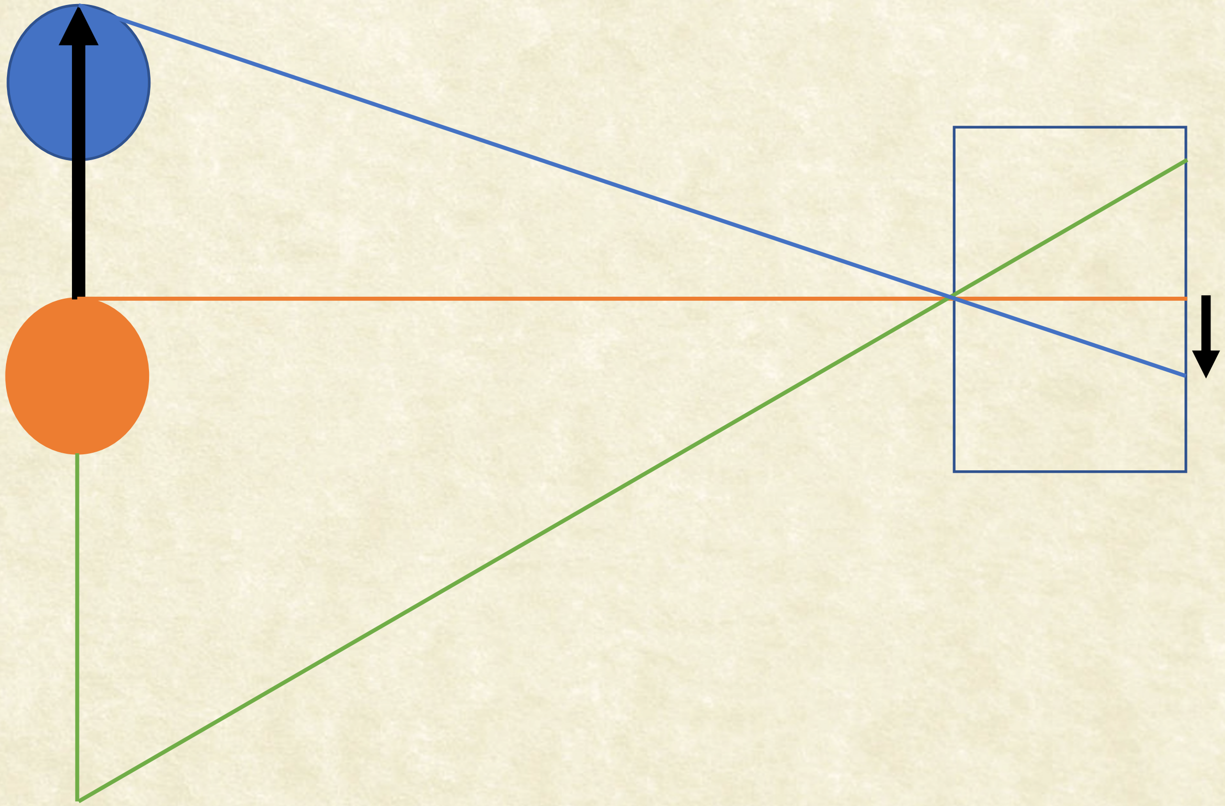
# Object Size vs Object Motion vs Camera Motion



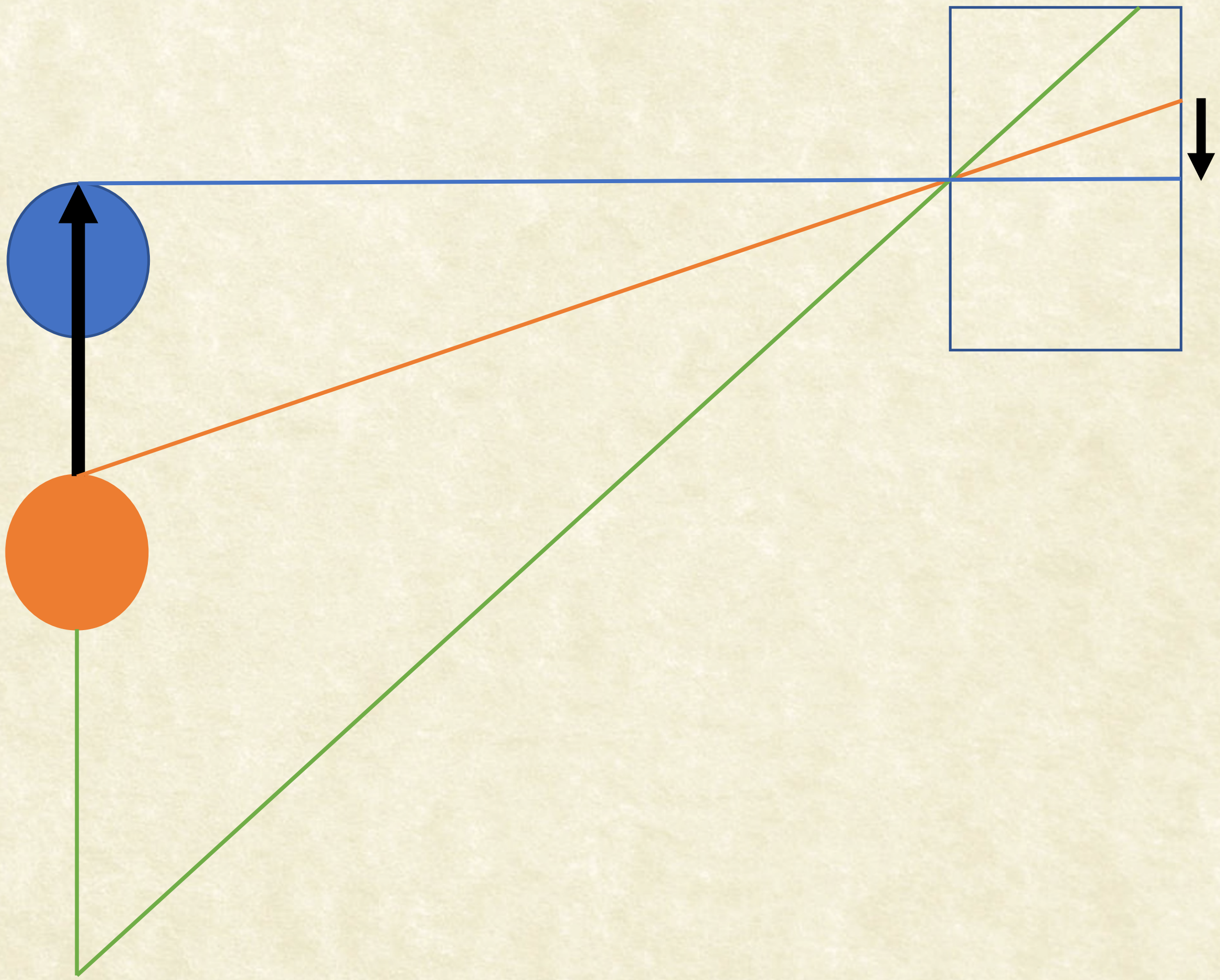
# Object Size vs Object Motion vs Camera Motion



# Object Size vs Object Motion vs Camera Motion



# Object Size vs Object Motion vs Camera Motion



# Problem

- You have a person who is **1.75m** tall standing at a distance of **7m** from a camera. The pinhole camera has a focal length of **20mm**. The sensor is 1cm tall and has a resolution of **4000x3000**.
  - Find the height of the person in pixels in the image.
  - If the camera is raised by **1m**, how much does the person move in the sensor (in pixels)?
  - How much does the Sun move in the above case
    Note: Sun is **150 million kms** away (in pixels)?

# Coordinate System

- $[X, Y, Z]$: World coordinate system
- $[x, y]$: Image coordinate system
- Camera Model: a function that maps world coordinates to image coordinates
- Perspective Projection: the mechanism of this many-to-one mapping

# Transformations: 3D Translation

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = \begin{bmatrix} X + d_X \\ Y + d_Y \\ Z + d_Z \end{bmatrix}$$

How do I write this as a linear transformation?

$$\boldsymbol{X}' = \boldsymbol{T}\boldsymbol{X}$$

# Transformations: 3D Translation

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = \begin{bmatrix} X + d_X \\ Y + d_Y \\ Z + d_Z \end{bmatrix}$$

$$T = \begin{bmatrix} 1 & 0 & 0 & d_X \\ 0 & 1 & 0 & d_Y \\ 0 & 0 & 1 & d_Z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} X' \\ Y' \\ Z' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & d_X \\ 0 & 1 & 0 & d_Y \\ 0 & 0 & 1 & d_Z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

$$T^{-1} = \begin{bmatrix} 1 & 0 & 0 & -d_X \\ 0 & 1 & 0 & -d_Y \\ 0 & 0 & 1 & -d_Z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

# Transformations: 3D Scaling

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = \begin{bmatrix} XS_X \\ YS_Y \\ ZS_Z \end{bmatrix}$$

$$T = \begin{bmatrix} S_X & 0 & 0 & 0 \\ 0 & S_Y & 0 & 0 \\ 0 & 0 & S_Z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} X' \\ Y' \\ Z' \\ 1 \end{bmatrix} = \begin{bmatrix} S_X & 0 & 0 & 0 \\ 0 & S_Y & 0 & 0 \\ 0 & 0 & S_Z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

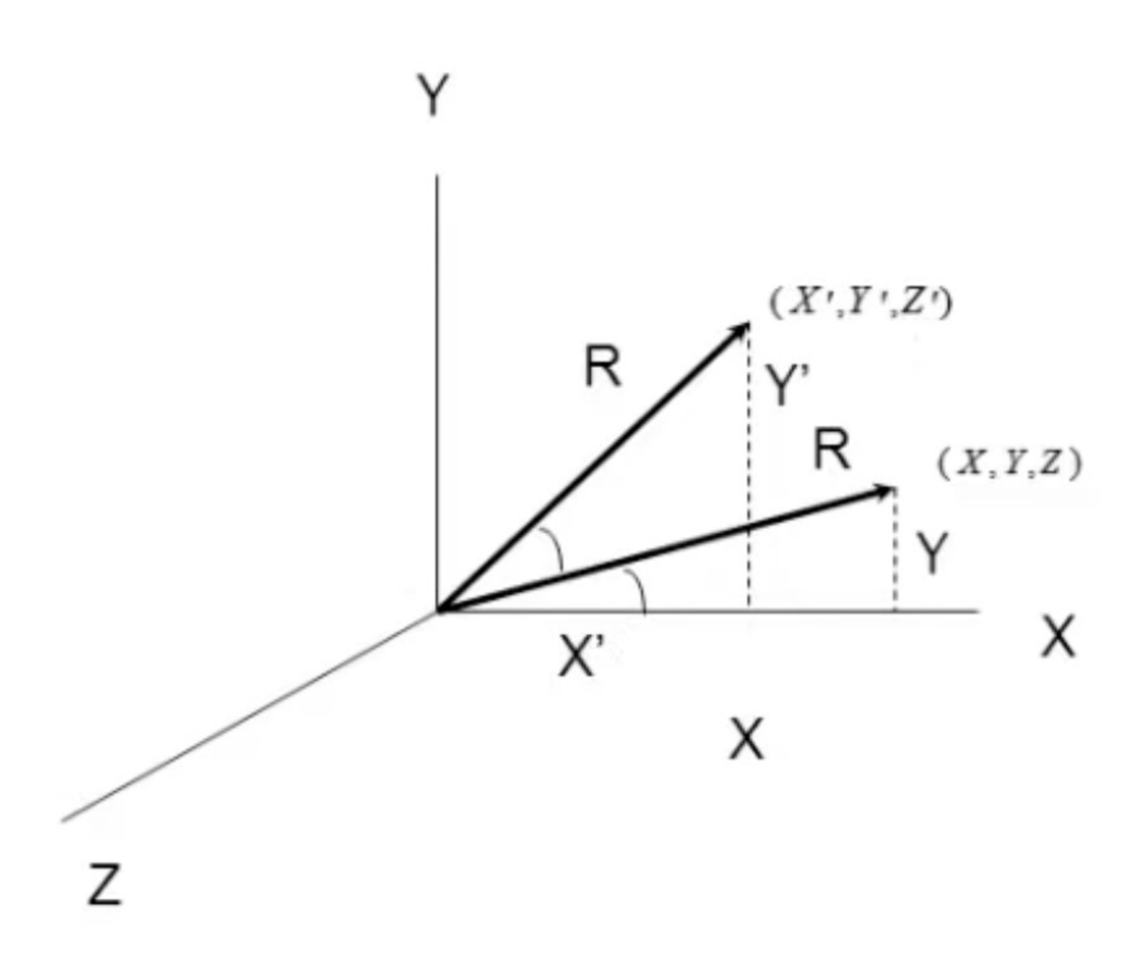
$$T^{-1} = \begin{bmatrix} 1/S_X & 0 & 0 & 0 \\ 0 & 1/S_Y & 0 & 0 \\ 0 & 0 & 1/S_Z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

# Transformations: 3D Rotation (around z-axis)

$$X = R\cos\emptyset$$
$$Y = R\ \sin\emptyset$$

$$X' = R\cos(\emptyset + \theta)$$
$$Y' = R\sin(\emptyset + \theta)$$

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$



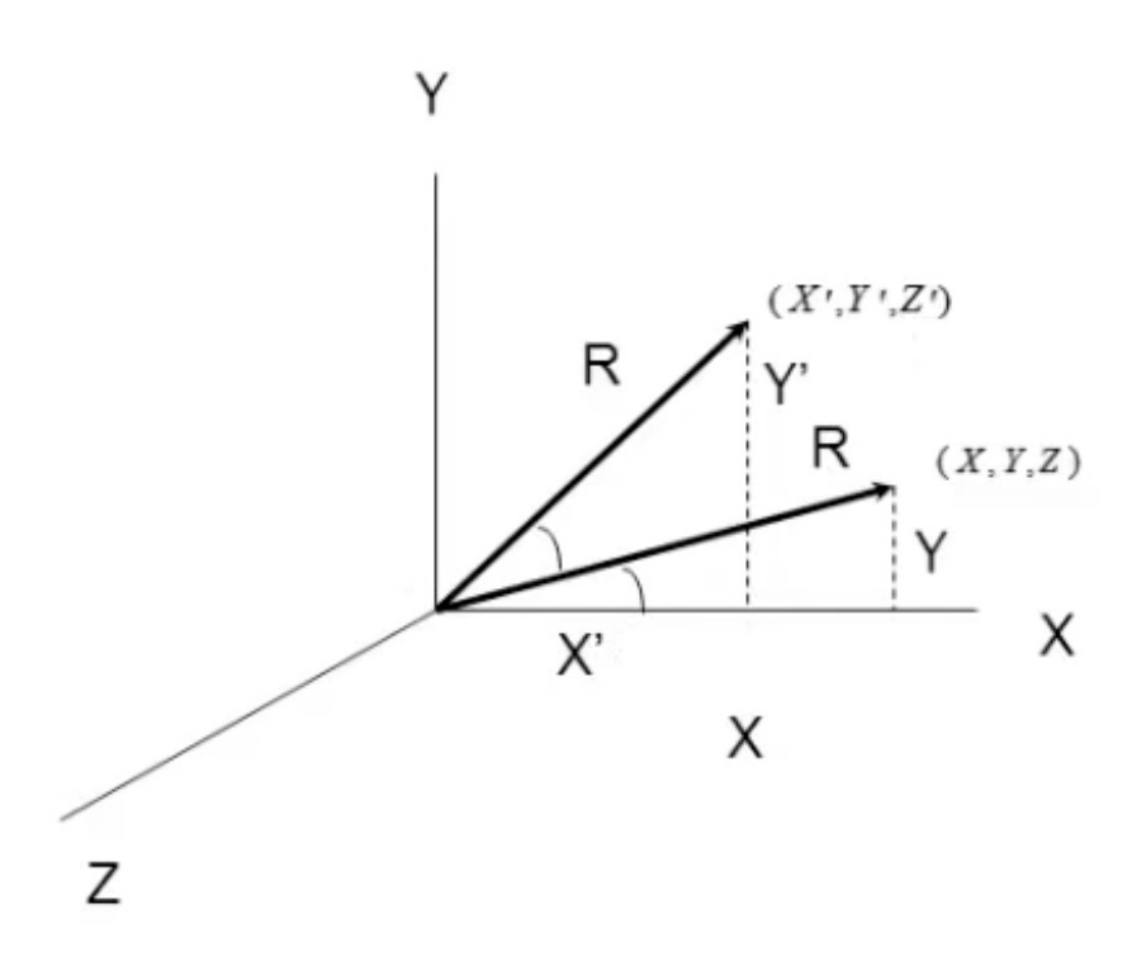
# Transformations: 3D Rotation (around z-axis)

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

$$R = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$R^{-1} = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$R^{-1} = \mathrm{R}^{\mathrm{T}}$$



# Transformations: 3D Rotation (Euler Angles)

Rotation around an arbitrary axis:

$$R = R_Z^{\alpha} R_Y^{\beta} R_X^{\gamma} = \begin{bmatrix} \cos\alpha\cos\beta & \cos\alpha\sin\beta\sin\gamma - \sin\alpha\cos\gamma & \cos\alpha\sin\beta\cos\gamma + \sin\alpha\sin\gamma \\ \sin\alpha\cos\beta & \sin\alpha\sin\beta\sin\gamma + \cos\alpha\cos\gamma & \sin\alpha\sin\beta\cos\gamma - \cos\alpha\sin\gamma \\ -\sin\beta & \cos\beta\sin\gamma & \cos\beta\cos\gamma \end{bmatrix}$$

Lecture 05: Spatial Transformations
Computer Graphics course by Keenan Crane
https://youtu.be/QmFBHSJS0Gw?list=PL9_jI1bdZmz2emSh0UQ5iOdT2xRHFHL7E&t=803

# Cartesian → Homogenous coordinates

- Add a new dimension:

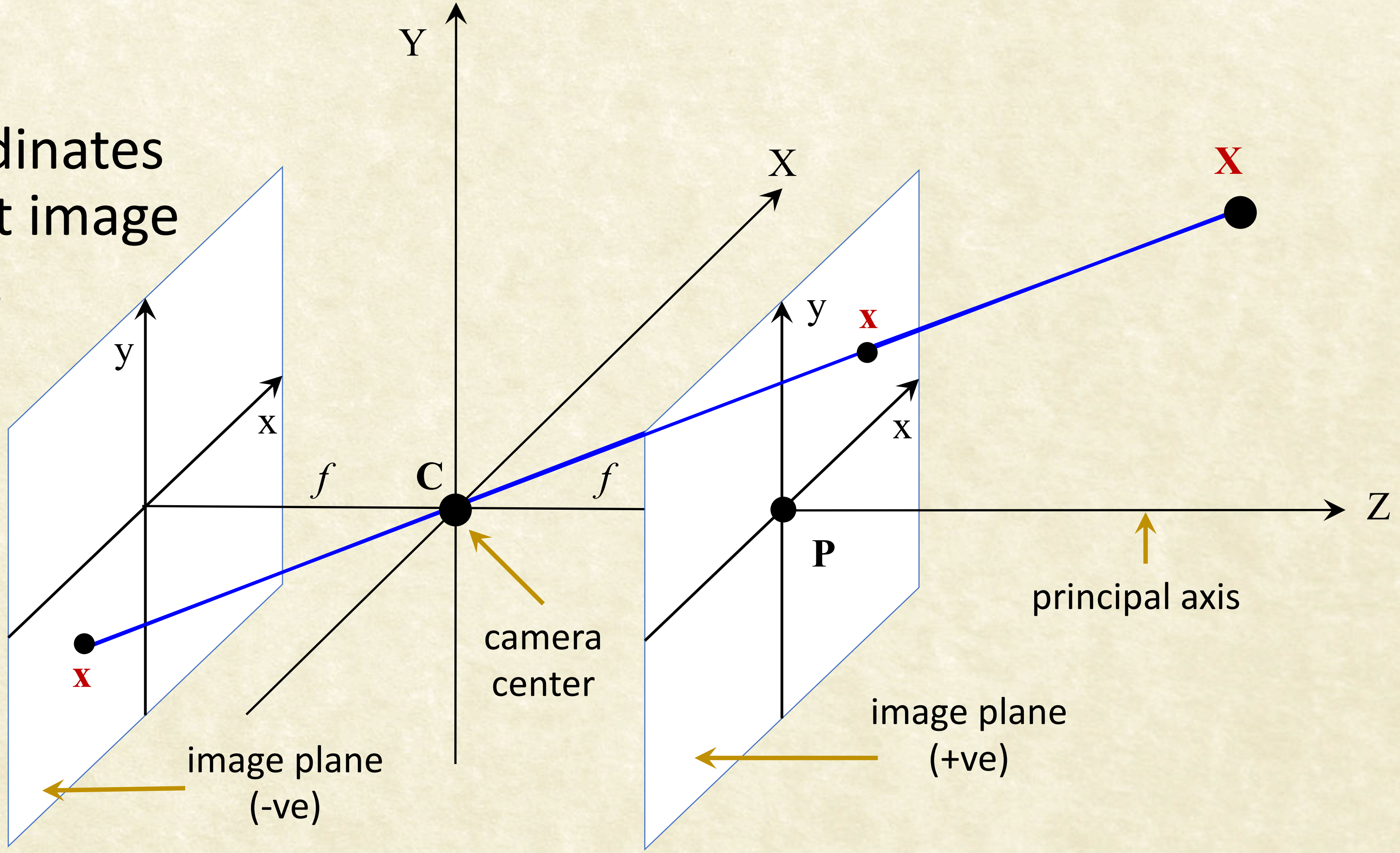
$$[x, y] \rightarrow \begin{bmatrix} kx \\ ky \\ k \end{bmatrix} \qquad [X, Y, Z] \rightarrow \begin{bmatrix} kX \\ kY \\ kZ \\ k \end{bmatrix}$$

- Convert back:

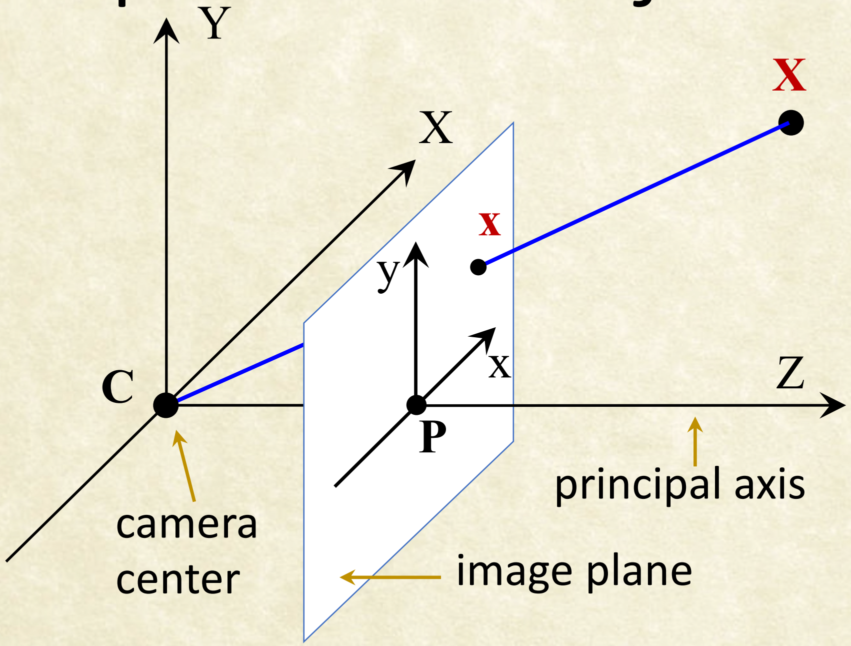
$$\begin{bmatrix} x \\ y \\ w \end{bmatrix} \rightarrow \left[\frac{x}{w}, \frac{y}{w}\right] \qquad \begin{bmatrix} X \\ Y \\ Z \\ W \end{bmatrix} \rightarrow \left[\frac{X}{W}, \frac{Y}{W}, \frac{Z}{W}\right]$$

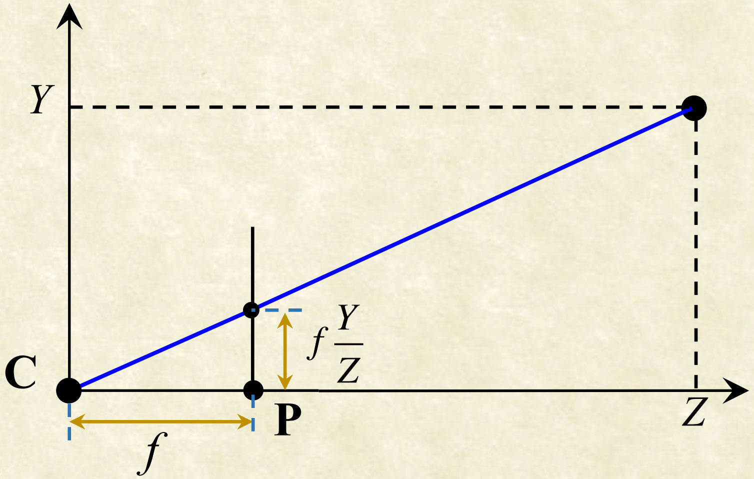
# Perspective Projection

- Inverted Image
- Multiply image coordinates with -1 to get upright image
- Equivalent to placing the image plane in front of C



# Perspective Projection




- Cartesian image coordinates: $x = f\frac{X}{Z}, \qquad y = f\frac{Y}{Z}$

- In matrix form (homogeneous):

$$\mathbf{x} = \begin{bmatrix} x \\ y \\ w \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \mathbf{PX}$$

# Basic Camera Equation

A pinhole camera projects a 3D point $\mathbf{X_c}$ in camera coords to an image point $\mathbf{x}$ via the 3x4 camera matrix $\mathbf{P}$ as:

$$\mathbf{x} = \mathbf{P}\mathbf{X}_c = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} [\mathbf{I}|\mathbf{0}]\mathbf{X}_c = \mathbf{K}[\mathbf{I}|\mathbf{0}]\mathbf{X_c},$$

where $\mathbf{K}$ is the internal camera calibration matrix.

Note that:

- The camera is at the origin
- Z is the Camera or Optical axis
- Principal Point: Center of the image
- Focal length in pixel units
- Orthogonal image axes with uniform scale

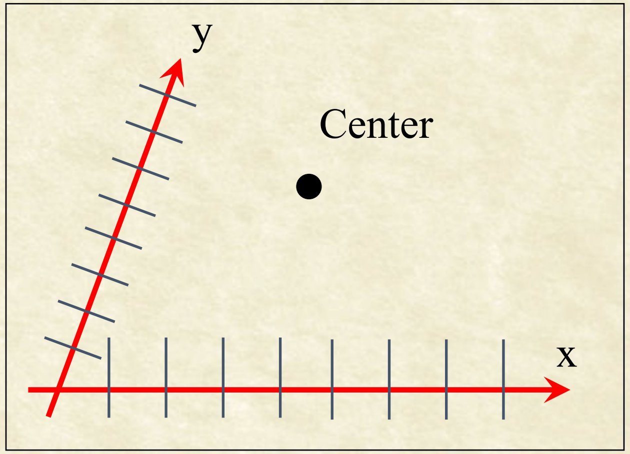
# A General Camera

Image center at $(x_0, y_0)$,

Non-orthogonal axes with skew $s$,

and different scales for axes with focal lengths, $\alpha_x$ and $\alpha_y$.

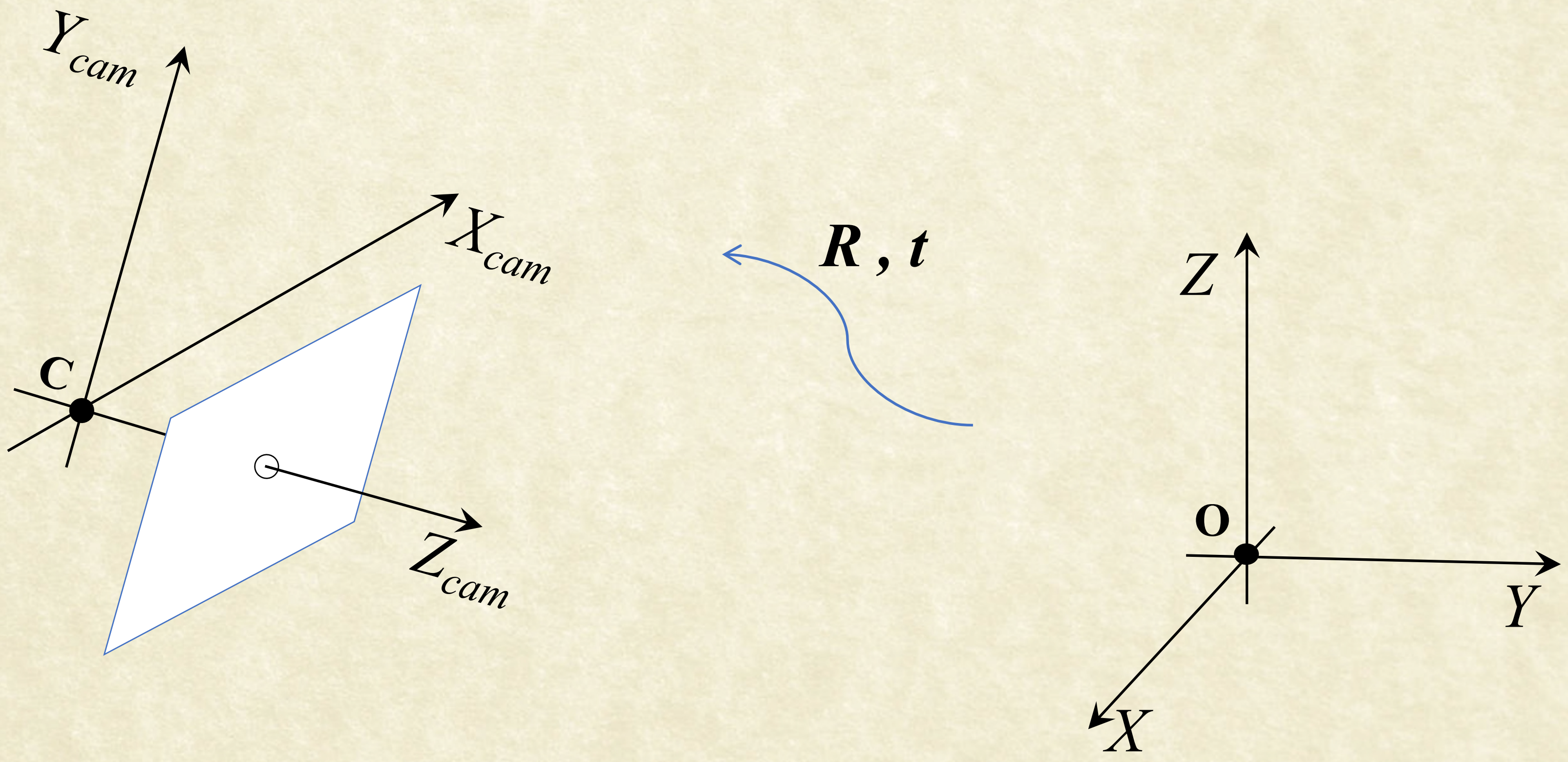
$$\mathbf{K} = \begin{bmatrix} \alpha_x & s & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{bmatrix}$$



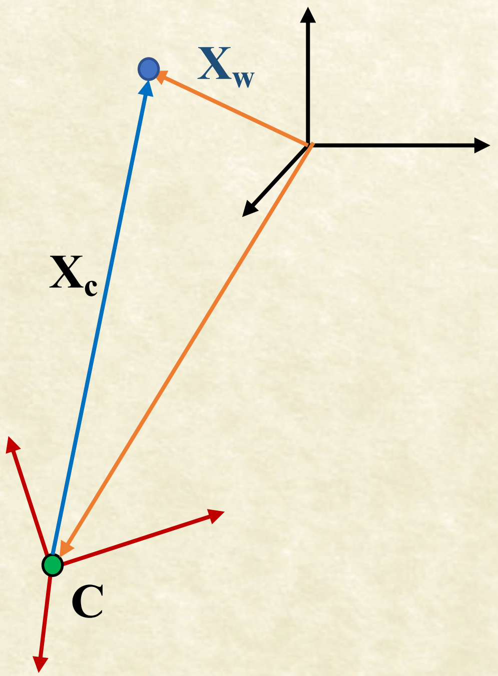

$\mathbf{K}$ an upper triangular matrix with 5 degrees of freedom.

# Moving the Camera from Origin

- General Setting: Camera is not at origin and Z is not the optical axis.
- Camera is at a point **C** in world coordinates. The world to camera rotation is **R** (rows of R are $X_{cam}$ $Y_{cam}$ $Z_{cam}$ in world coordinates).



# General Camera Equation




- In cartesian coordinates, $\mathbf{X_c = R\,(X_w - C) = RX_w - RC}$
- In homogeneous co-ordinates:

$$\mathbf{X_c} = \begin{bmatrix} \mathbf{R} & \mathbf{-RC} \\ \mathbf{0} & \mathbf{1} \end{bmatrix} \mathbf{X_w}$$

- 2D projection $\mathbf{x}$ of a 3D point $\mathbf{X_w}$ given by:
  - $\mathbf{x = K[\,I \mid 0\,]\,X_c = K\,[\,R \mid -RC\,]\,X_w}$
- $\mathbf{x = PX_w}$ ; camera matrix $\mathbf{P = [KR \mid -KRC] = [M \mid p_4]}$
- Common K: $\begin{bmatrix} f & 0 & x_0 \\ 0 & f & y_0 \\ 0 & 0 & 1 \end{bmatrix}$ General K: $\begin{bmatrix} \alpha_x & s & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{bmatrix}$

Questions?