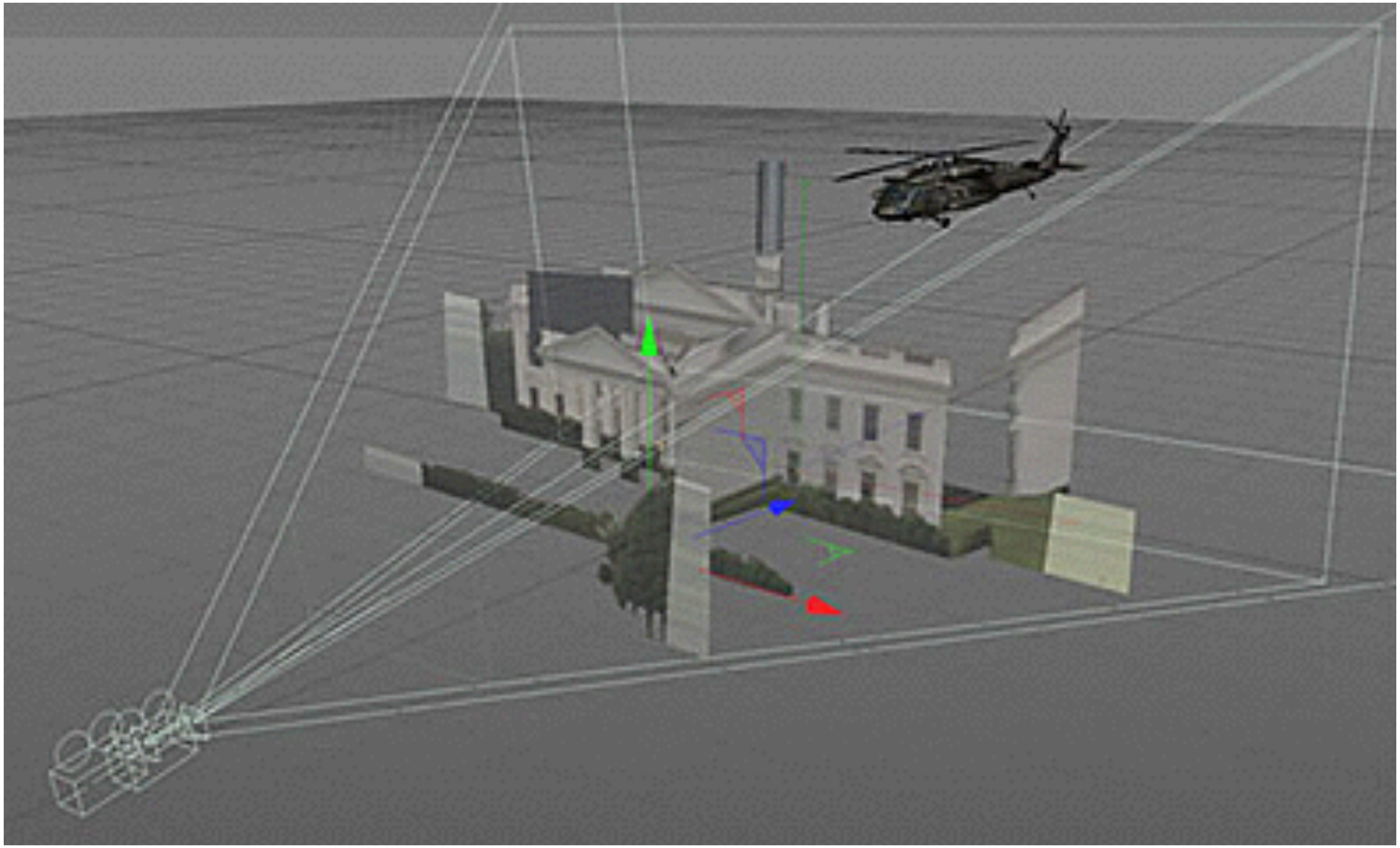
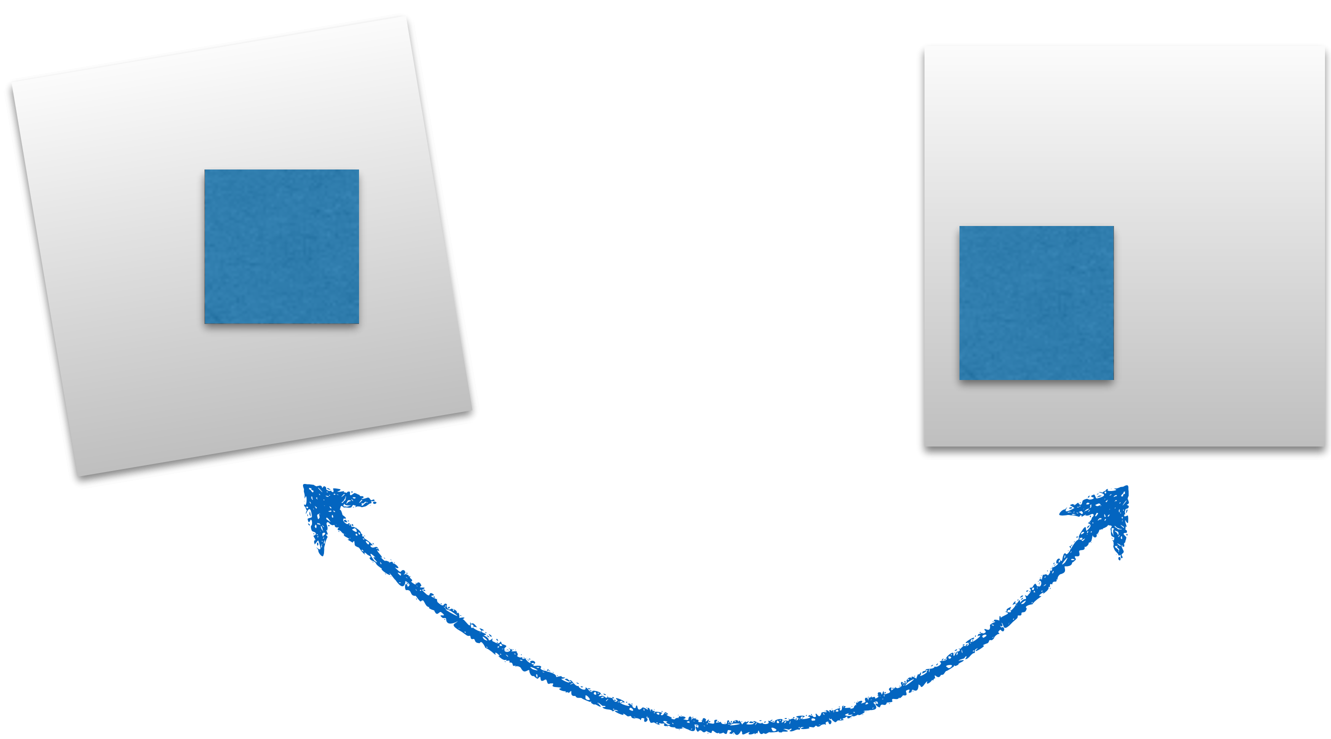


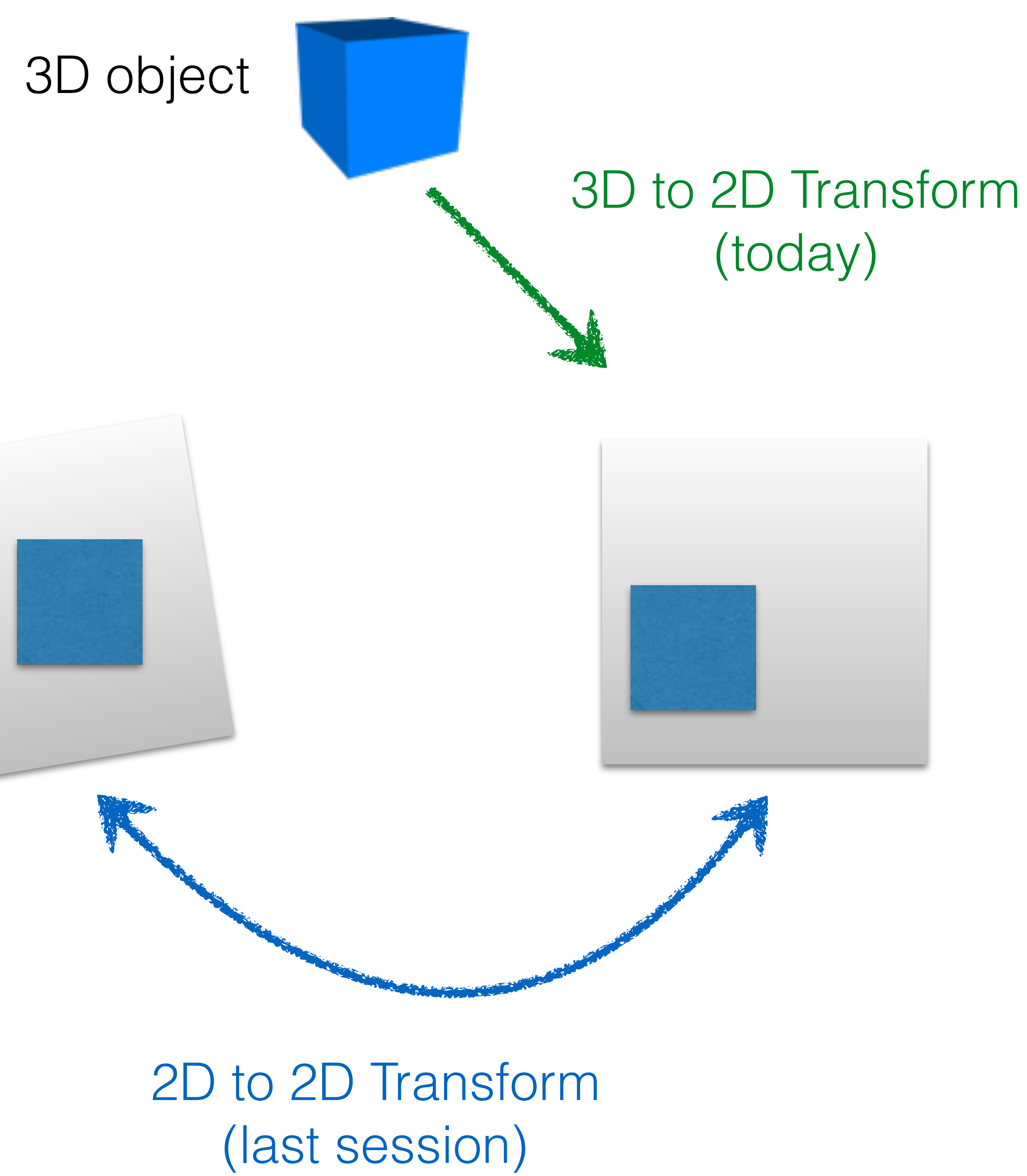
# Camera Matrix

16-385 Computer Vision (Kris Kitani)

**Carnegie Mellon University**



2D to 2D Transform
(last session)


3D object
3D to 2D Transform
(today)
2D to 2D Transform
(last session)

A camera is a mapping between

the **3D world**

and

a **2D image**

A camera is a mapping between
the 3D world and a 2D image

$$\boldsymbol{x} = \mathbf{P}\mathbf{X}$$

2D image point

camera matrix

3D world point

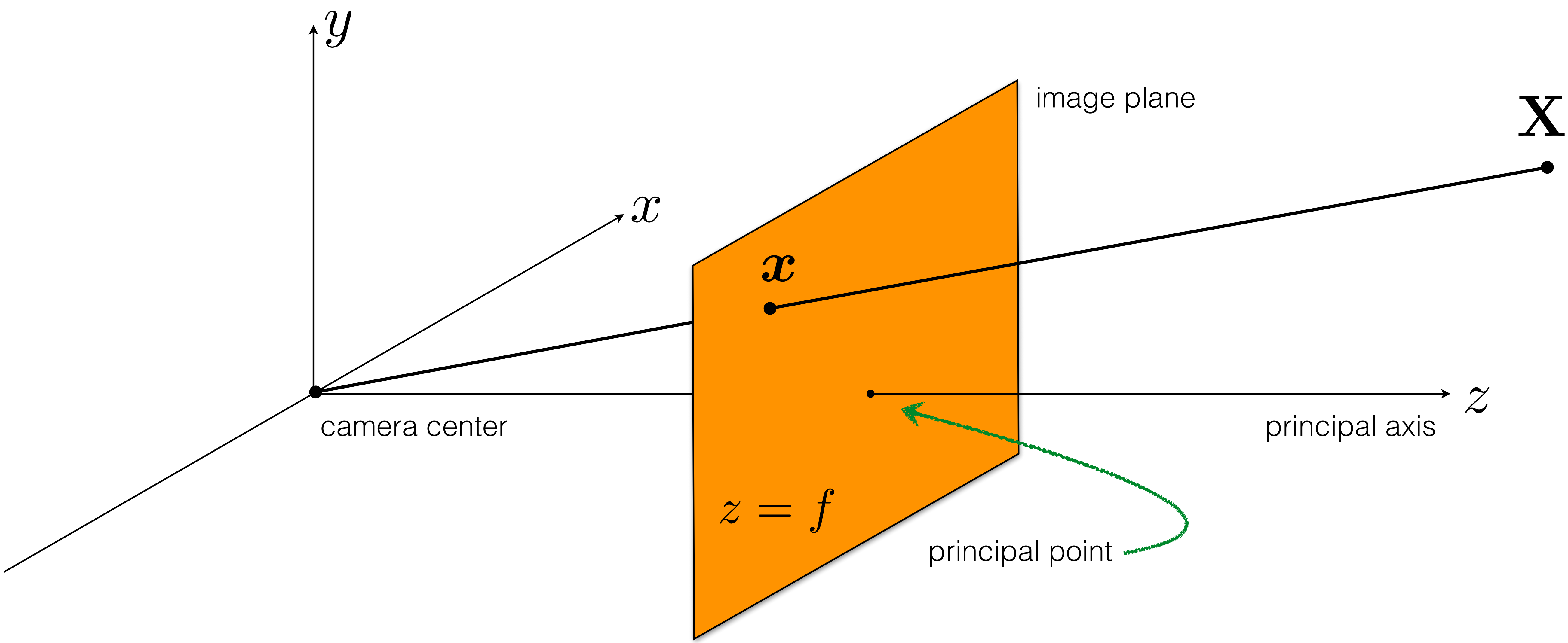
*What do you think the dimensions are?*

$$\boldsymbol{x} = \mathbf{PX}$$

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} p_1 & p_2 & p_3 & p_4 \\ p_5 & p_6 & p_7 & p_8 \\ p_9 & p_{10} & p_{11} & p_{12} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

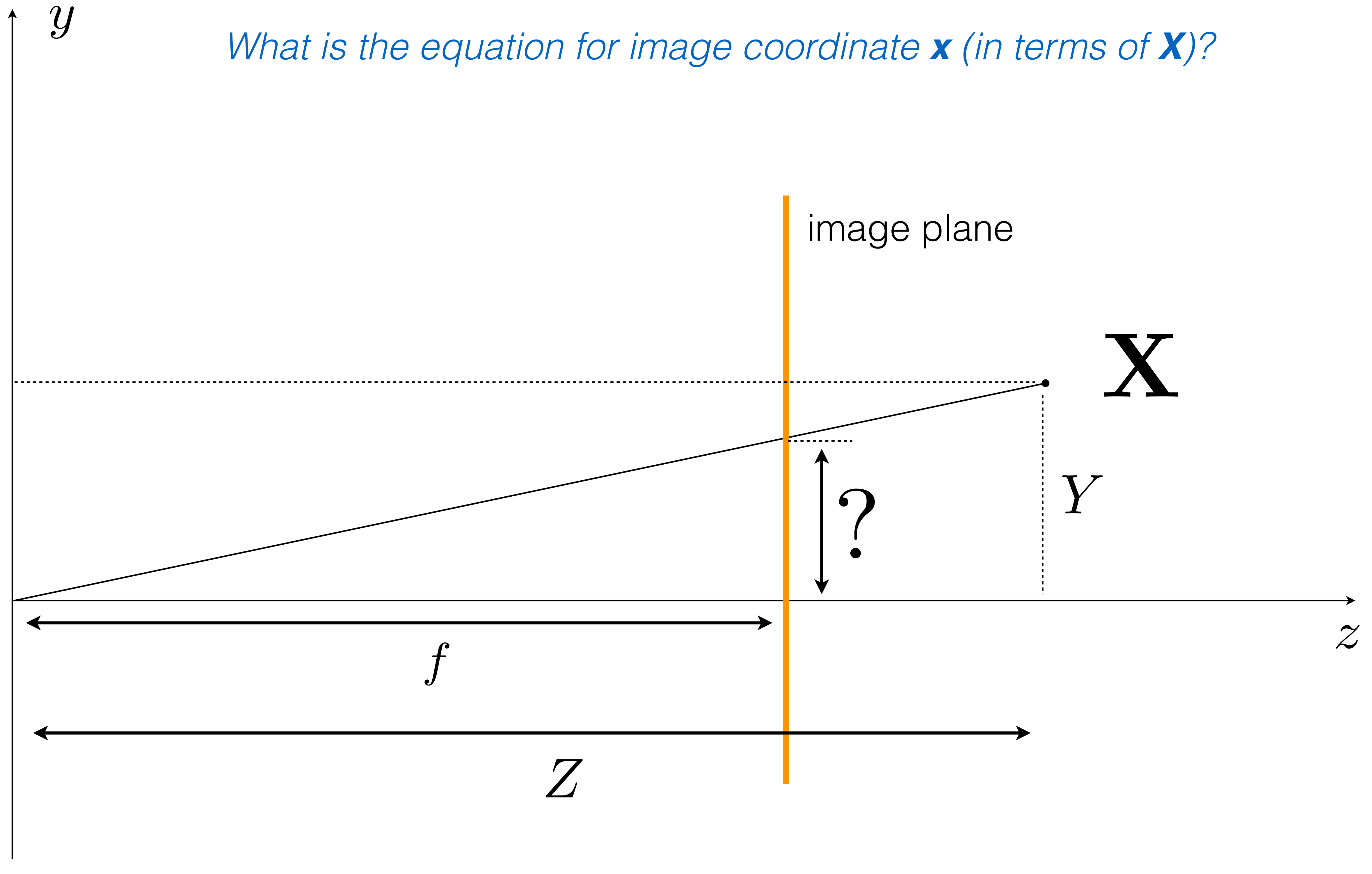
| homogeneous image 3 x 1 | Camera matrix 3 x 4 | homogeneous world point 4 x 1 |
|---|---|---|

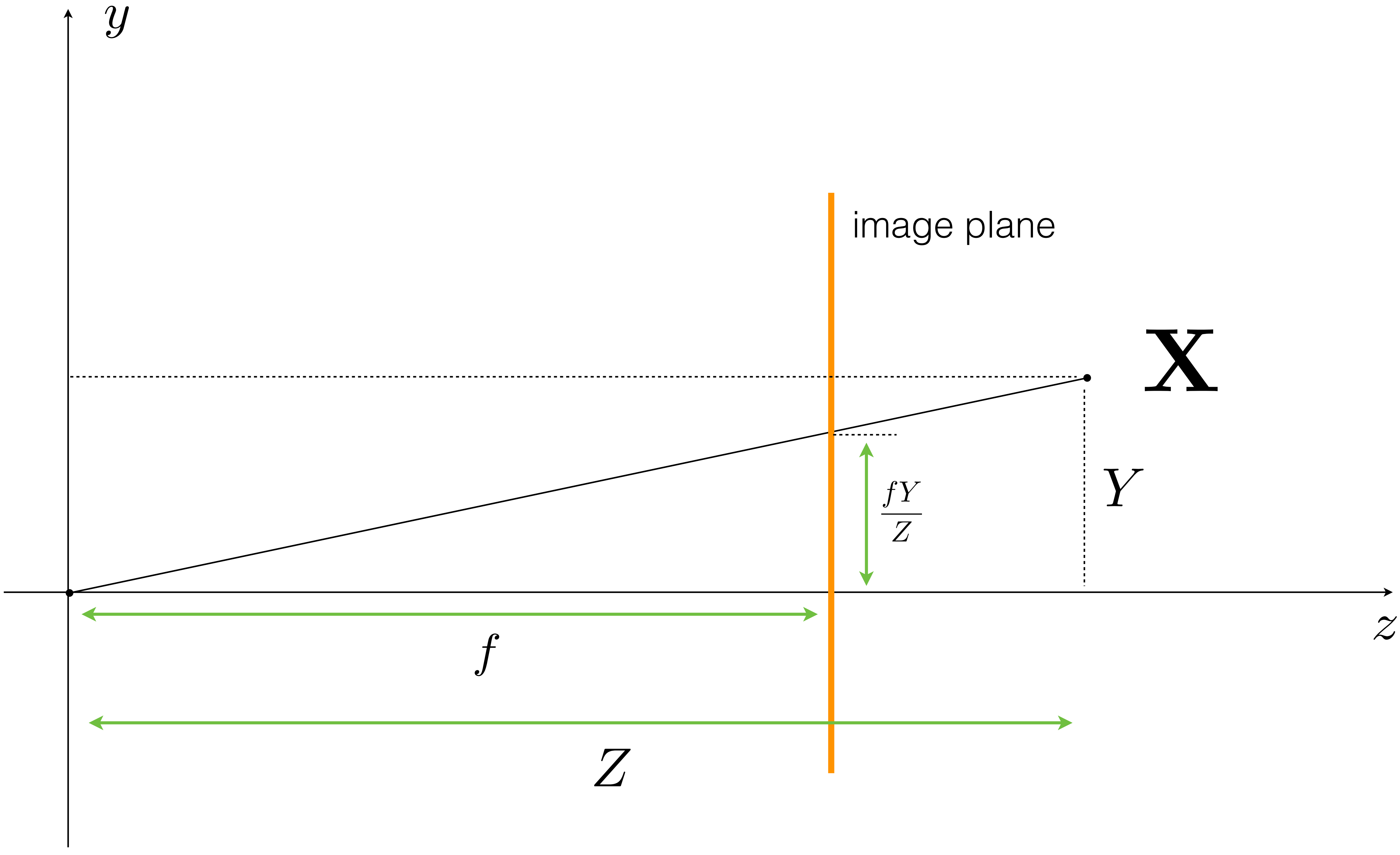
# The pinhole camera




*What is the equation for image coordinate **x** (in terms of **X**)?*

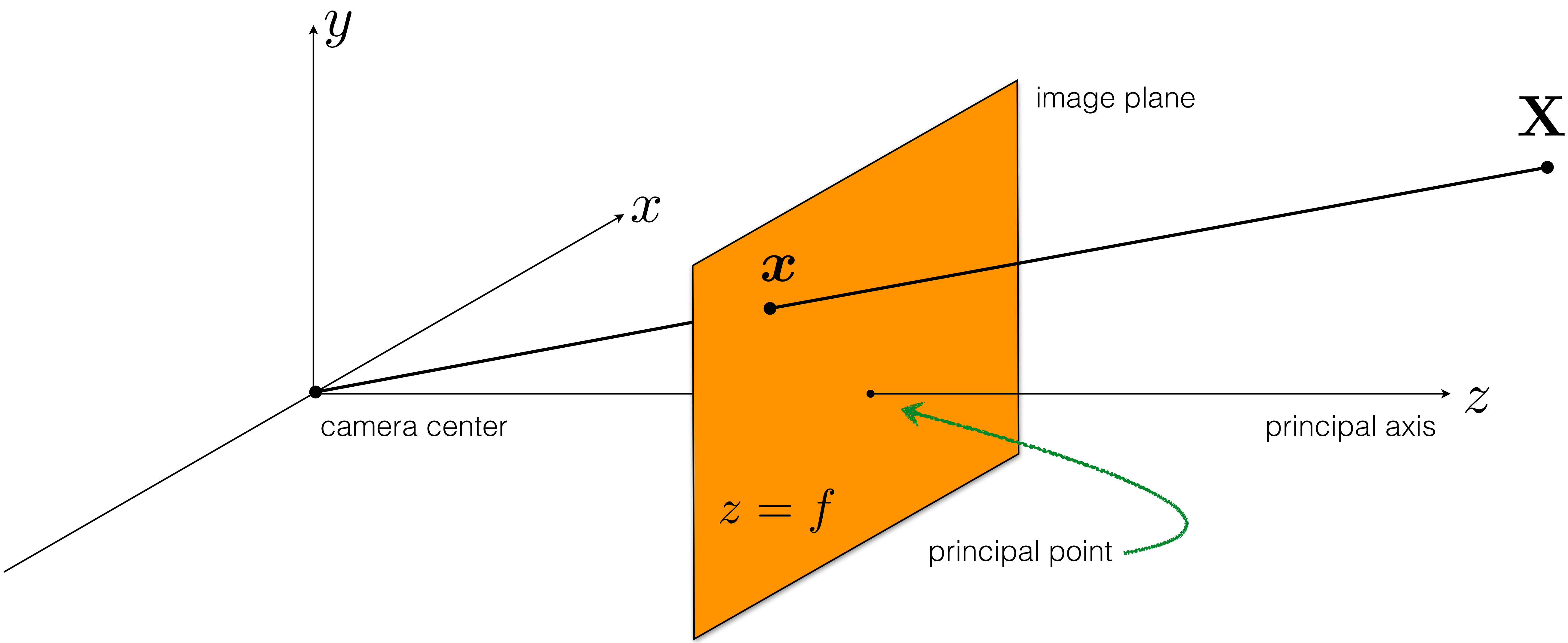
*What is the equation for image coordinate **x** (in terms of **X**)?*






$$[X \quad Y \quad Z]^\top \mapsto [fX/Z \quad fY/Z]^\top$$

# Pinhole camera geometry




*What is the camera matrix* **P** *for a pinhole camera model?*

$$\boldsymbol{x} = \mathbf{P}\mathbf{X}$$

Relationship from similar triangles…

$$[X \quad Y \quad Z]^\top \mapsto [fX/Z \quad fY/Z]^\top$$

generic camera model

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} p_1 & p_2 & p_3 & p_4 \\ p_5 & p_6 & p_7 & p_8 \\ p_9 & p_{10} & p_{11} & p_{12} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

*What does the pinhole camera model look like?*

$$\mathbf{P} = \begin{bmatrix} ? & ? & ? & ? \\ ? & ? & ? & ? \\ ? & ? & ? & ? \end{bmatrix}$$

Relationship from similar triangles…

$$[X \quad Y \quad Z]^\top \mapsto [fX/Z \quad fY/Z]^\top$$

generic camera model

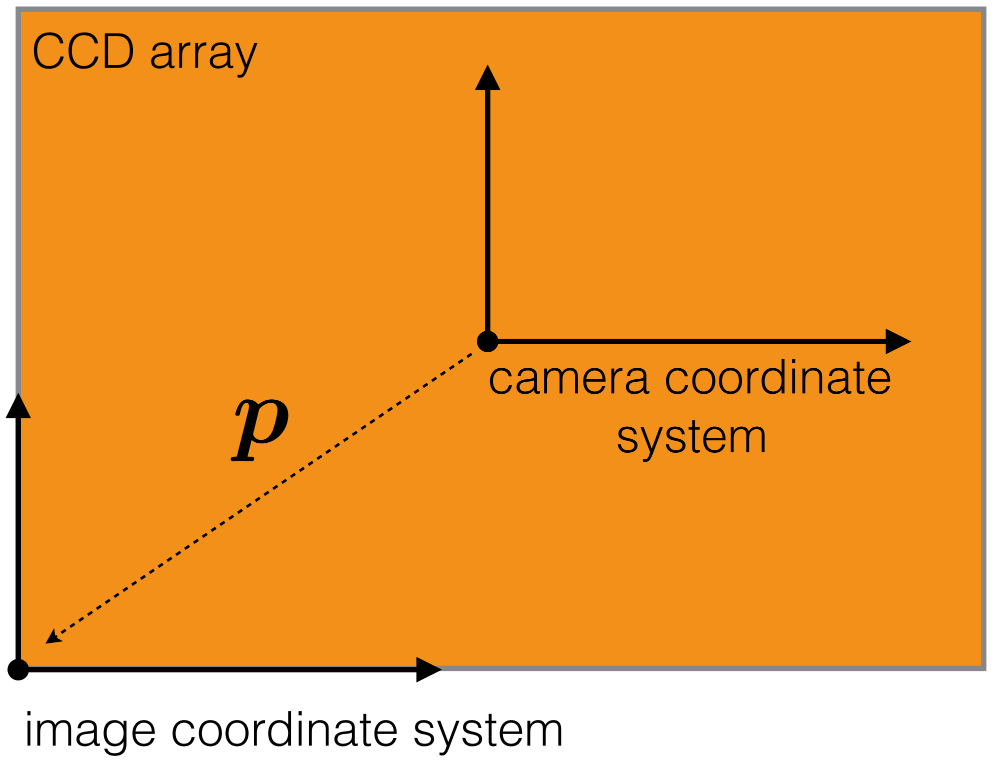
$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} p_1 & p_2 & p_3 & p_4 \\ p_5 & p_6 & p_7 & p_8 \\ p_9 & p_{10} & p_{11} & p_{12} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

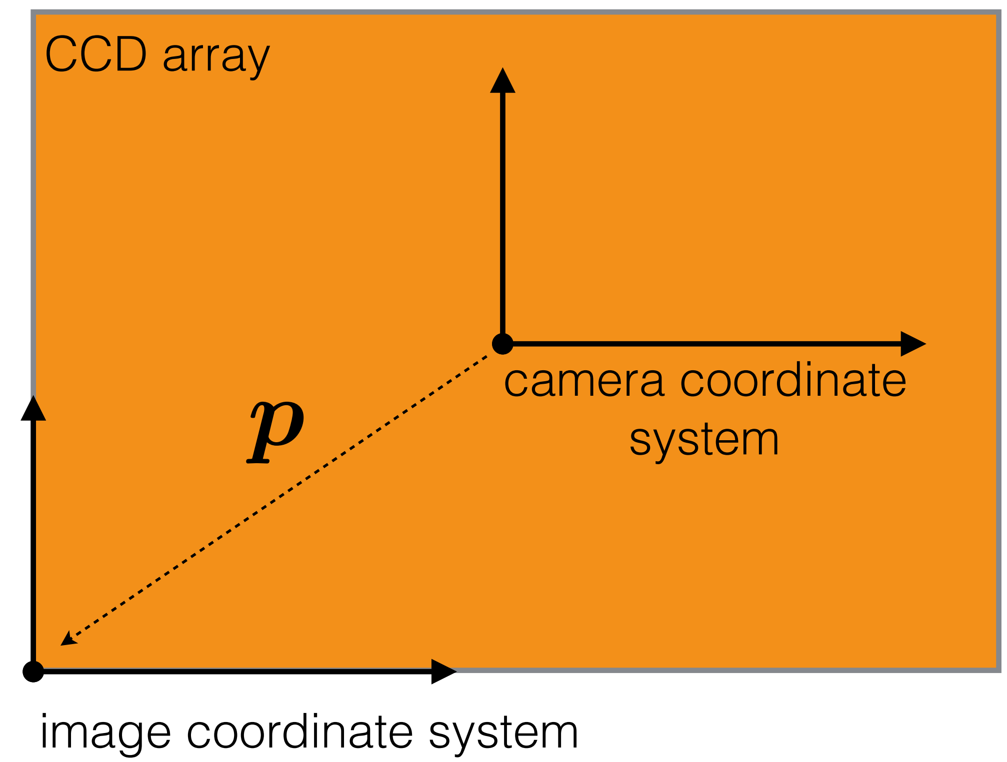
*What does the pinhole camera model look like?*

$$\mathbf{P} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$$\mathbf{P} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Camera origin and image origin might be different

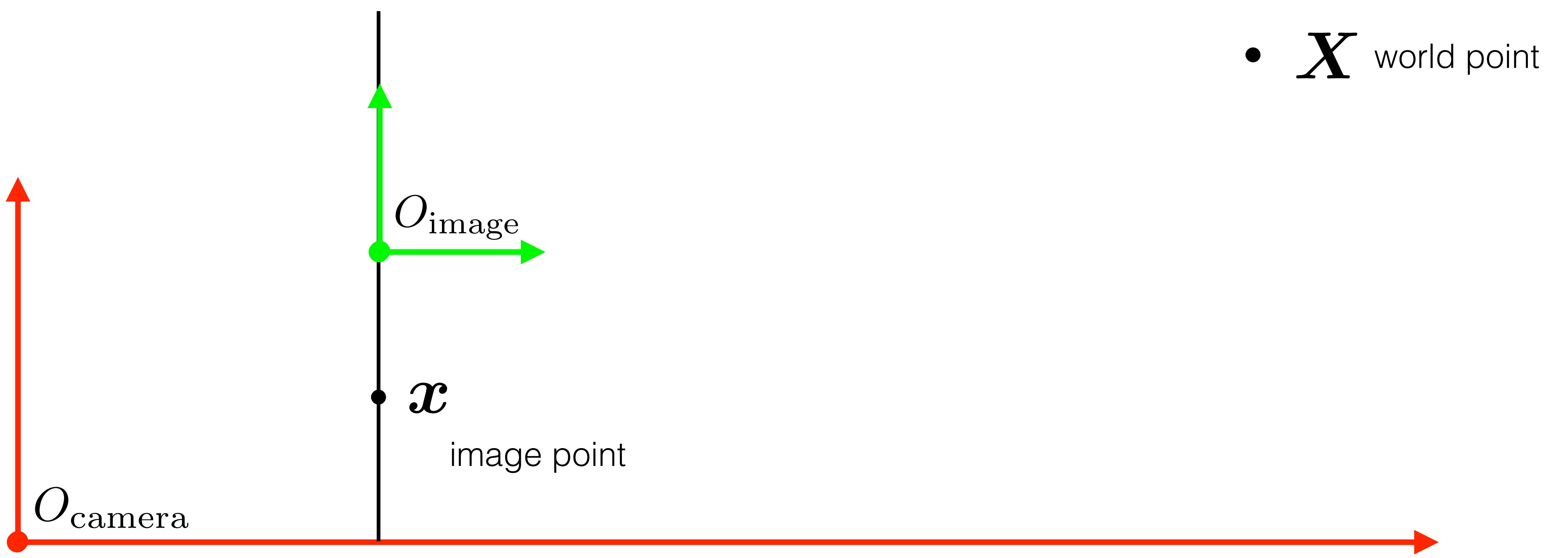




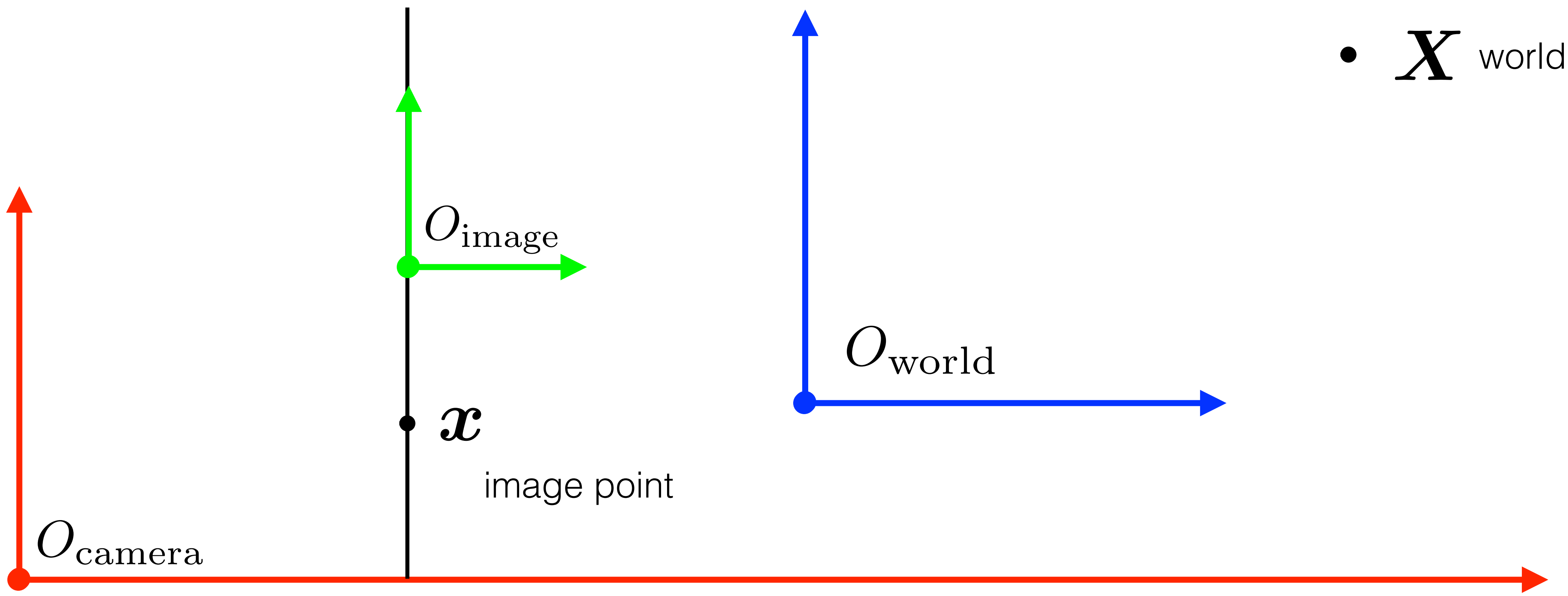

$$\mathbf{P} = \begin{bmatrix} f & 0 & p_x & 0 \\ 0 & f & p_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Accounts for different origins

In general, the camera and image sensor have **different** coordinate systems



In general, there are **three different** coordinate systems…




so you need the know the transformations between them

# Can be decomposed into two matrices

$$\mathbf{P} = \begin{bmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix} \left[\begin{array}{ccc:c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right]$$

(3 x 3) (3 x 4)

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

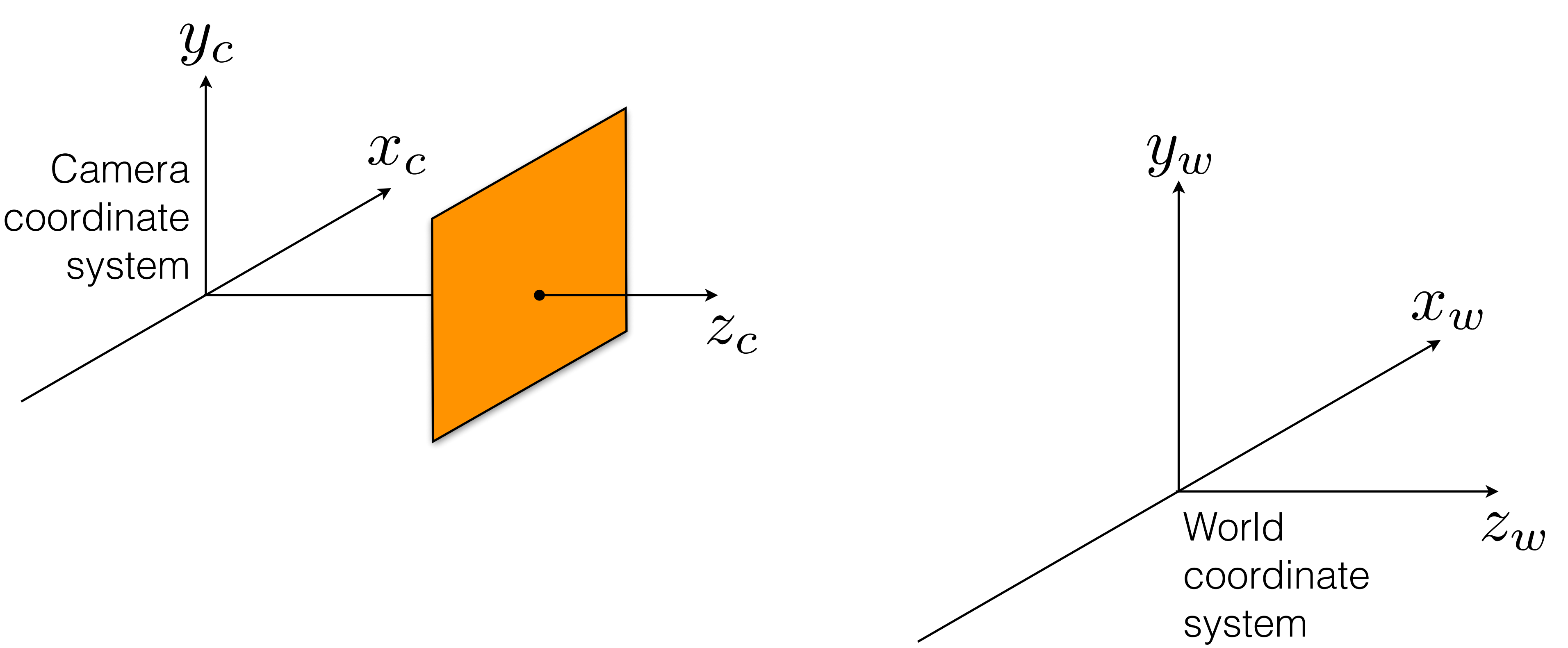
$$\mathbf{K} = \begin{bmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix}$$

calibration matrix

Assumes that the **camera** and **world** share the same coordinate system

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

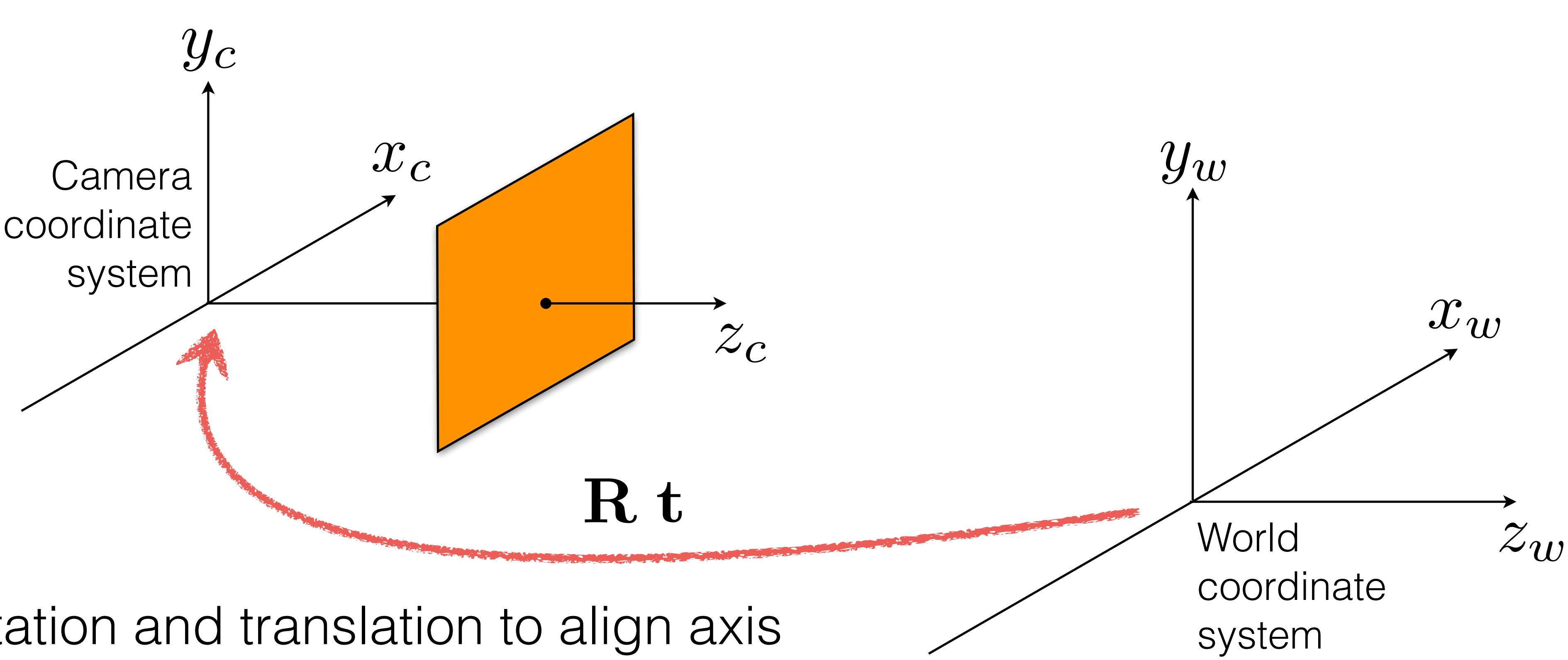
*What if they are different?*
*How do we align them?*



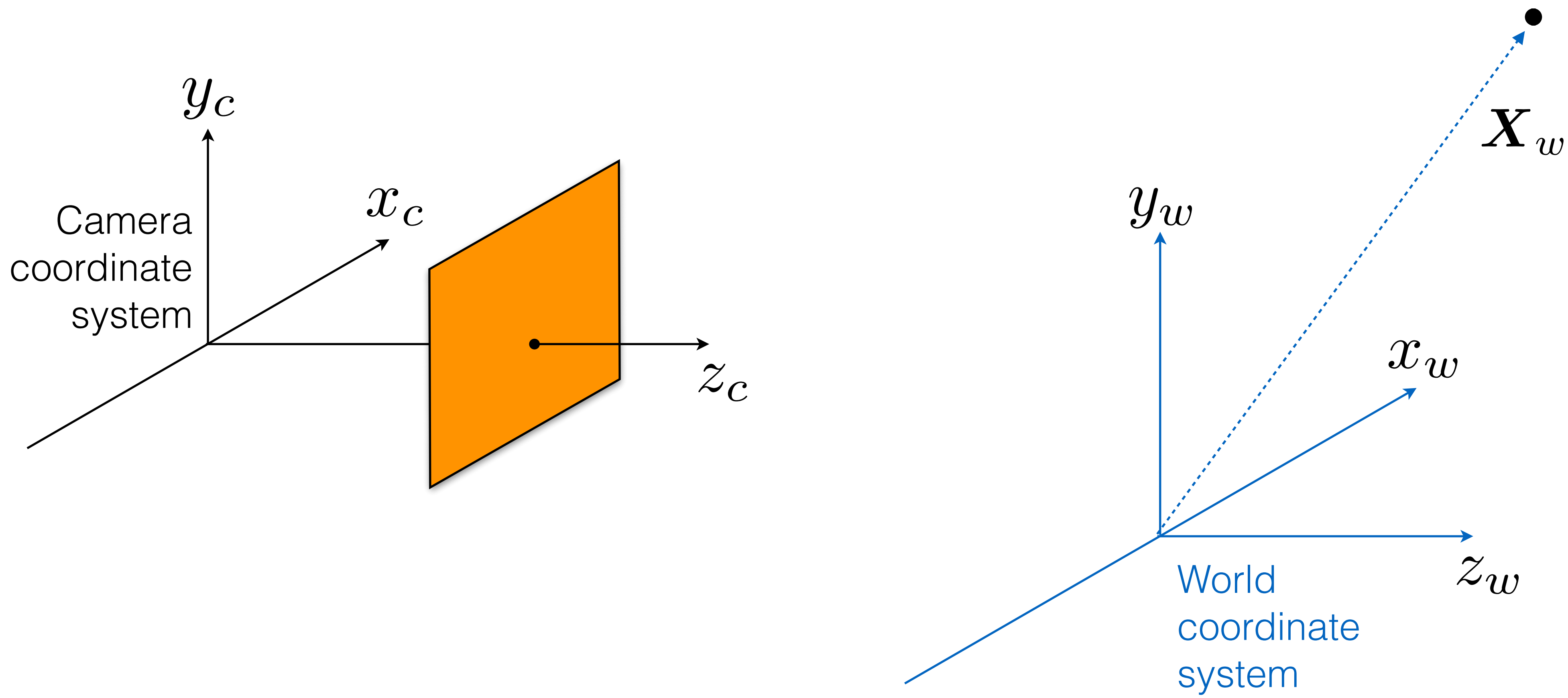
Assumes that the camera and world share the same coordinate system

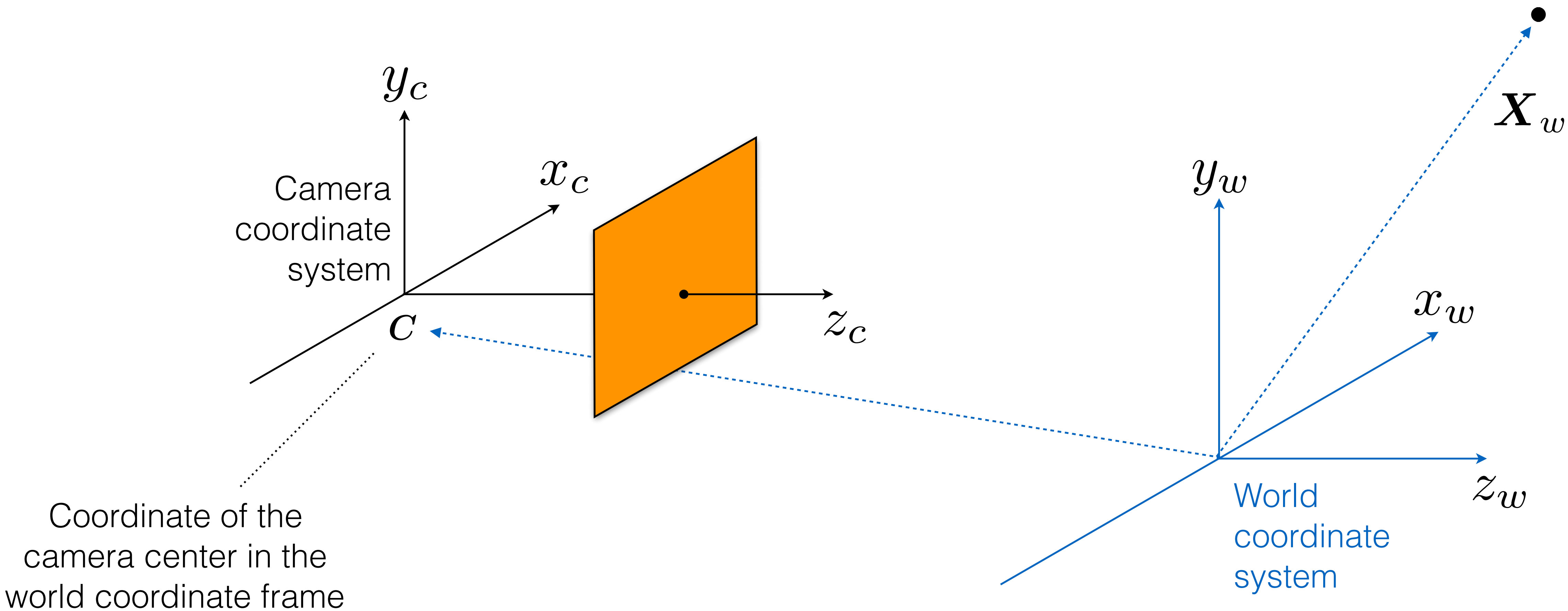
$$\mathbf{P} = \begin{bmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

*What if they are different?*
*How do we align them?*

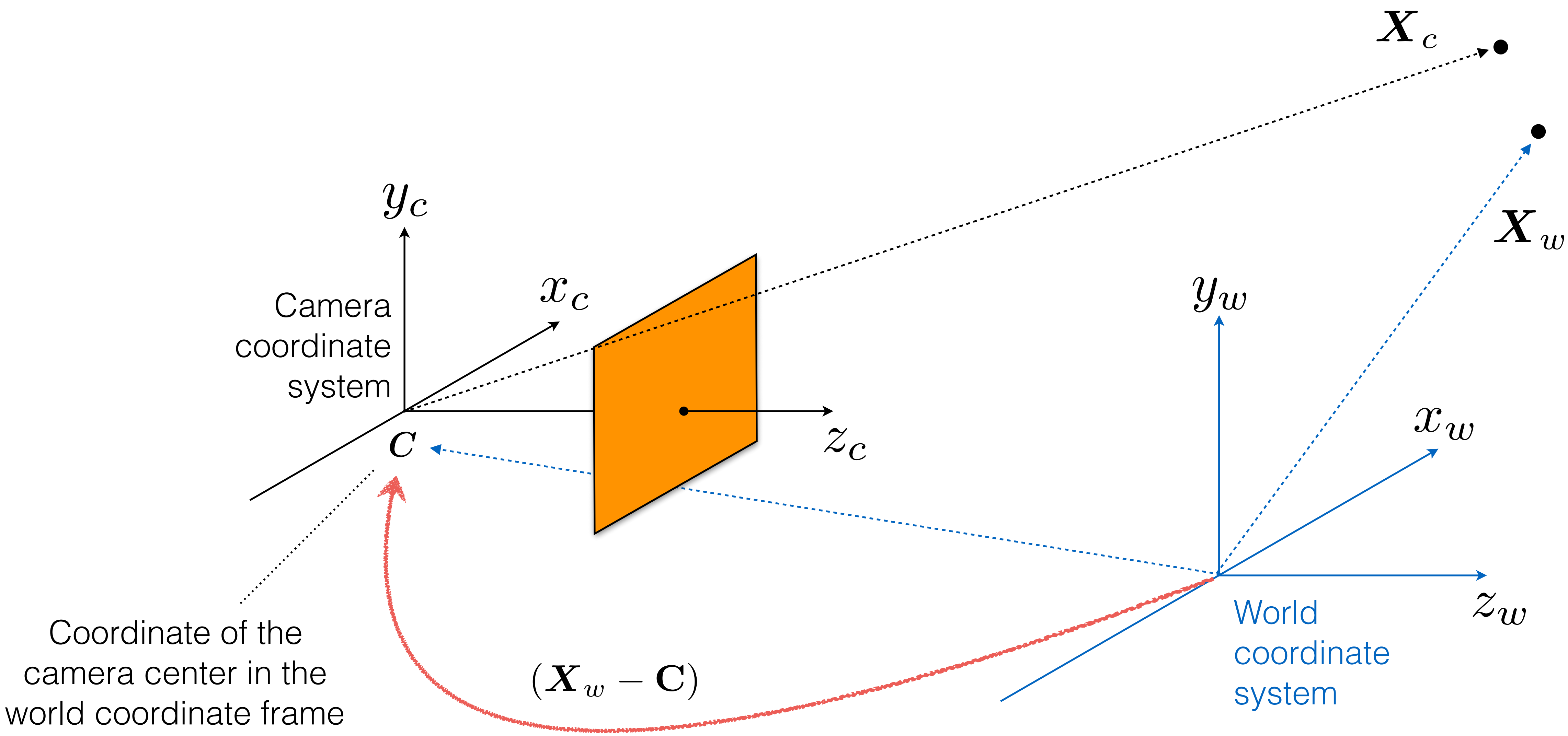



3R rotation and translation to align axis


$y_c$
$x_c$
Camera
coordinate
system
$z_c$
$\boldsymbol{X}_w$
$y_w$
$x_w$
$z_w$
World
coordinate
system

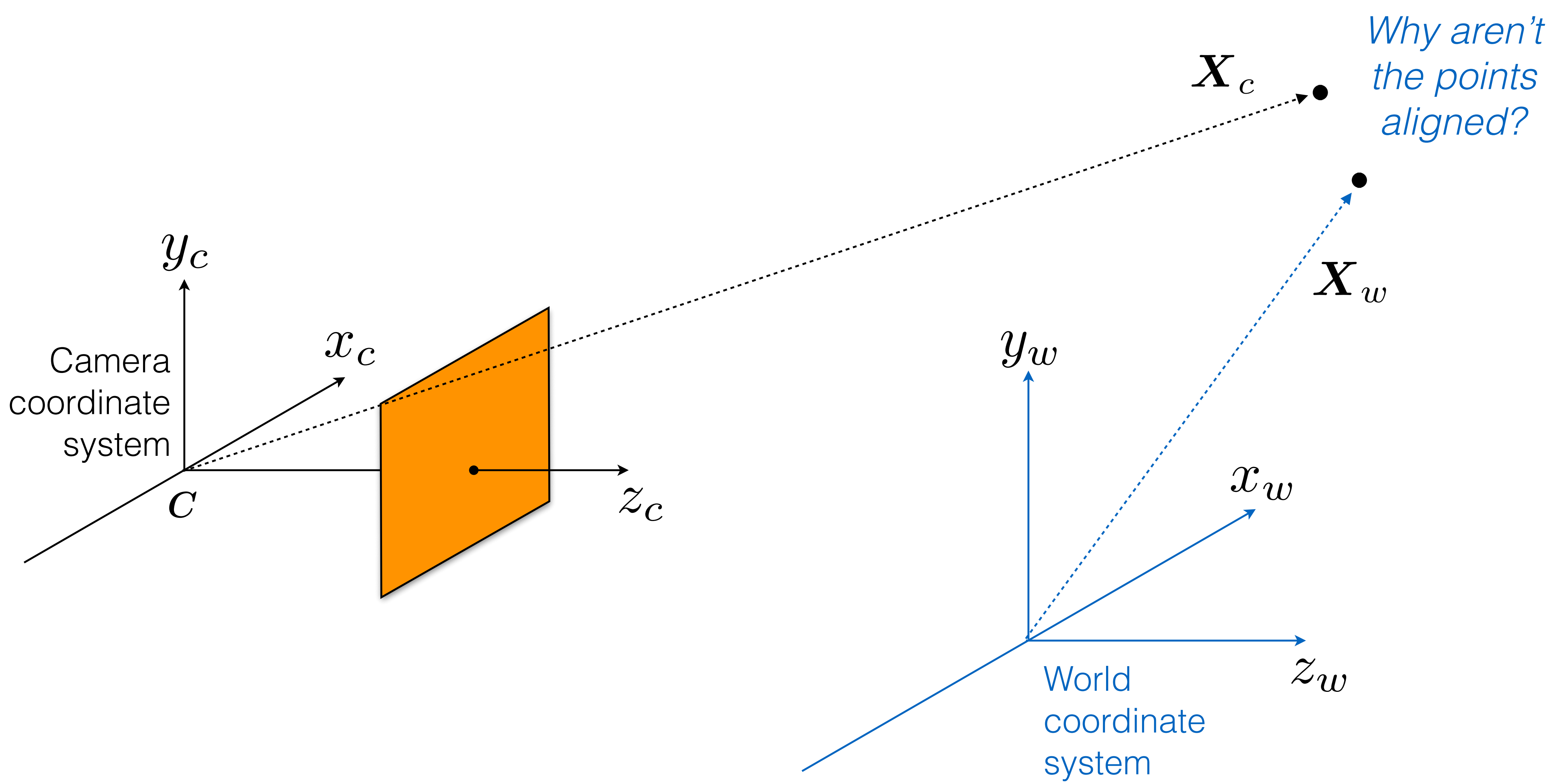


$y_c$
$x_c$
Camera coordinate system
$z_c$
$\boldsymbol{C}$
Coordinate of the camera center in the world coordinate frame
$\boldsymbol{X}_w$
$y_w$
$x_w$
$z_w$
World coordinate system




$$(\boldsymbol{X}_w - \mathbf{C})$$

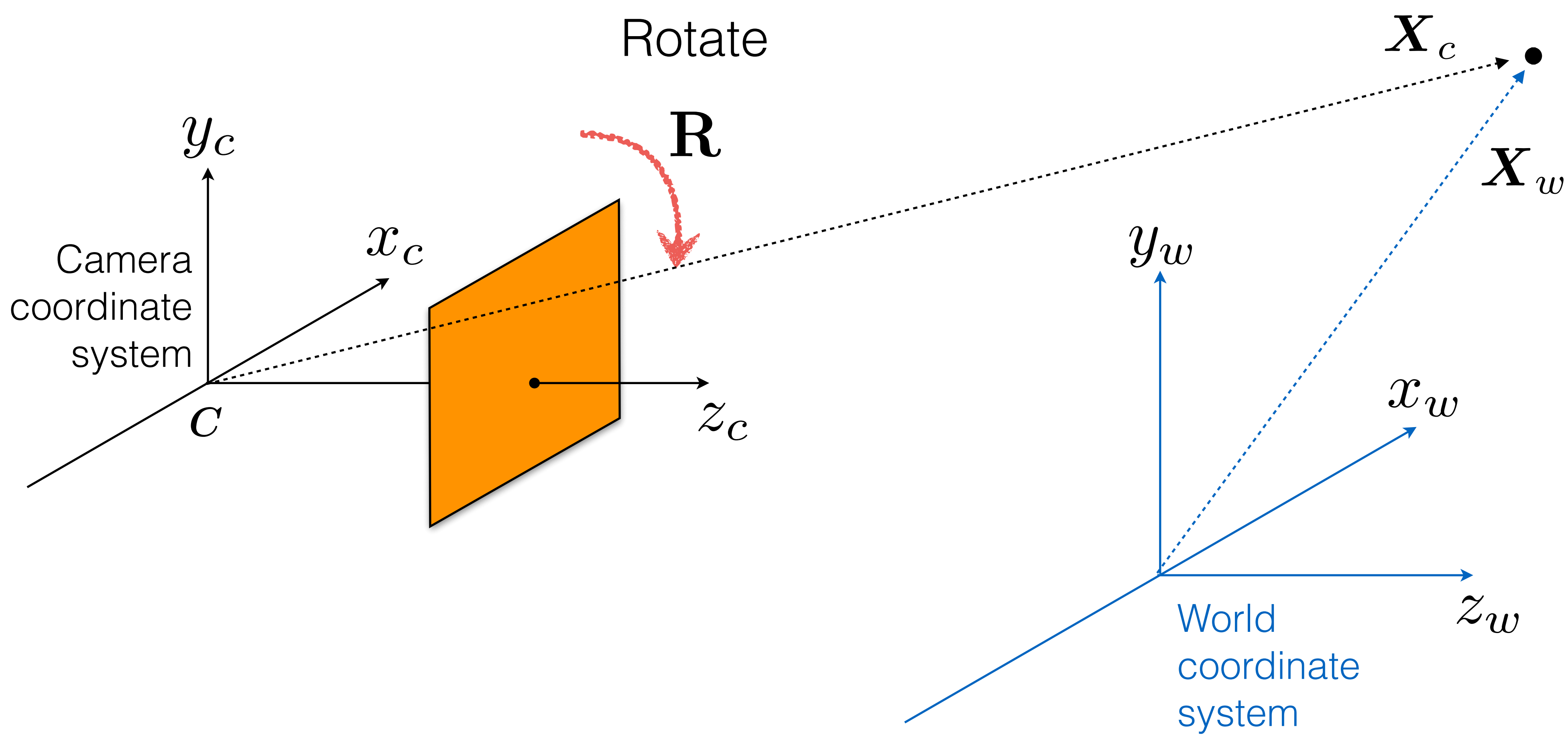
Translate




$$(X_w - \mathbf{C})$$

Translate

# *What happens to points after alignment?*




$$\mathbf{R}(\boldsymbol{X}_w - \mathbf{C})$$

Rotate Translate

In inhomogeneous coordinates:

$$\boldsymbol{X}_c = \mathbf{R}(\boldsymbol{X}_w - \mathbf{C})$$

Optionally in homogeneous coordinates:

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \underset{(4 \times 4)}{\begin{bmatrix} \mathbf{R} & -\mathbf{RC} \\ \mathbf{0} & 1 \end{bmatrix}} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

General mapping of a pinhole camera

$$\mathbf{P} = \mathbf{KR}[\mathbf{I} | -\mathbf{C}]$$

## Quiz

What is the meaning of each matrix of the camera matrix decomposition?

$$\mathbf{P} = \mathbf{K}\mathbf{R}[\mathbf{I}| - \mathbf{C}]$$

3x3
intrinsics

# Quiz

What is the meaning of each matrix of the camera matrix decomposition?

$$\mathbf{P} = \mathbf{K}\mathbf{R}[\mathbf{I} | -\mathbf{C}]$$

3x3
intrinsics

3x3
3D rotation

# Quiz

What is the meaning of each matrix of the camera matrix decomposition?

$$\mathbf{P} = \mathbf{K}\mathbf{R}[\mathbf{I} | -\mathbf{C}]$$

3x3
intrinsics

3x3
3D rotation

3x3
identity

# Quiz

What is the meaning of each matrix of the camera matrix decomposition?

$$\mathbf{P} = \mathbf{K}\mathbf{R}[\mathbf{I}| - \mathbf{C}]$$

| 3x3 intrinsics | 3x3 3D rotation | 3x3 identity | 3x1 3D translation |
| --- | --- | --- | --- |

General mapping of a pinhole camera

$$\mathbf{P} = \mathbf{KR}[\mathbf{I}| - \mathbf{C}]$$

(translate first then rotate)

Another way to write the mapping

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

where

$$\mathbf{t} = -\mathbf{RC}$$

(rotate first then translate)

# Quiz

The camera matrix relates what two quantities?

# Quiz

The camera matrix relates what two quantities?

$$\boldsymbol{x} = \mathbf{P}\mathbf{X}$$

# Quiz

The camera matrix relates what two quantities?

$$\boldsymbol{x} = \mathbf{P}\mathbf{X}$$

3D points to 2D image points

## Quiz

The camera matrix relates what two quantities?

$$\boldsymbol{x} = \mathbf{P}\mathbf{X}$$

3D points to 2D image points

The camera matrix can be decomposed into?

# Quiz

The camera matrix relates what two quantities?

$$\boldsymbol{x} = \mathbf{P}\mathbf{X}$$

3D points to 2D image points

The camera matrix can be decomposed into?

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

# Quiz

The camera matrix relates what two quantities?

$$\boldsymbol{x} = \mathbf{P}\mathbf{X}$$

3D points to 2D image points

The camera matrix can be decomposed into?

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

intrinsic and extrinsic parameters

# Generalized pinhole camera model

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

$$\mathbf{P} = \begin{bmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix} \left[\begin{array}{ccc:c} r_1 & r_2 & r_3 & t_1 \\ r_4 & r_5 & r_6 & t_2 \\ r_7 & r_8 & r_9 & t_3 \end{array}\right]$$

intrinsic parameters    extrinsic parameters

$$\mathbf{R} = \begin{bmatrix} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \\ r_7 & r_8 & r_9 \end{bmatrix} \qquad \mathbf{t} = \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix}$$

3D rotation    3D translation

*Why do we need **P**?*

to properly relate **world points** to **image points**
(by taking into account different coordinate systems)