

3D Vision & Stereo I

coordinate systems, calibration, stereo geometry, triangulation



CSC420

David Lindell

University of Toronto

cs.toronto.edu/~lindell/teaching/420

Slide credit: Kris Kitani, Srinivasa Narasimhan, Fredo Durand, Yannis Gkioulekas (CMU 16-385)

Announcements

- HW 2 due tomorrow!
- HW 3 is out (due Nov 13)
- next week: virtual lecture will be posted on Quercus.
No in-person lecture!
- midterm on Oct 24!

Midterm Information (Oct 24)

Updated 27 seconds ago by David Lindell

Hi all,

The midterm for the course will be held on October 24th in lieu of the tutorial hour.

- Section LEC0101: SF 3202 1-2pm
- Section LEC0201: SF 3202 3-4pm

Please arrive promptly at the hour so that we can distribute the exam booklets and begin and end on time.

The exam will be a combination of multiple choice and long answer. You will write directly on the exam booklet, so bring a pencil and eraser. The exam is closed book and no notes are allowed.

Also, you must abide by the honor code — do not discuss the content of the midterm exam with any other students until after the grades have been released (obviously we have students taking the exam at different times, so please abide by this out of fairness to everyone).

Please comment on this post with any questions.

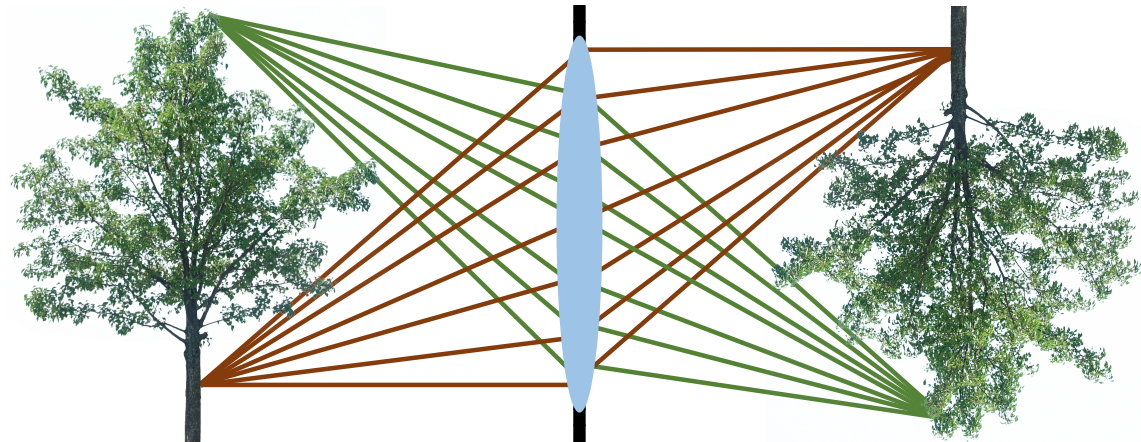
David

Course Overview

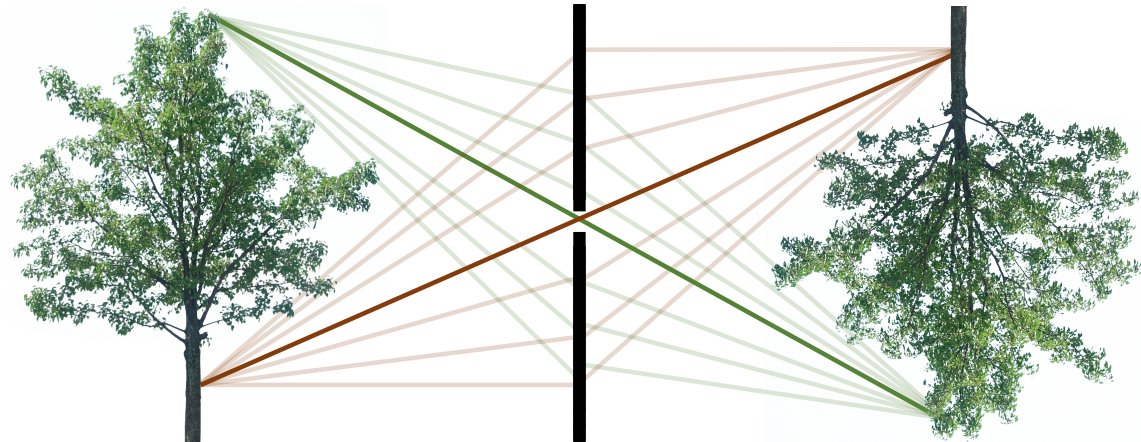


Pinhole and lens cameras

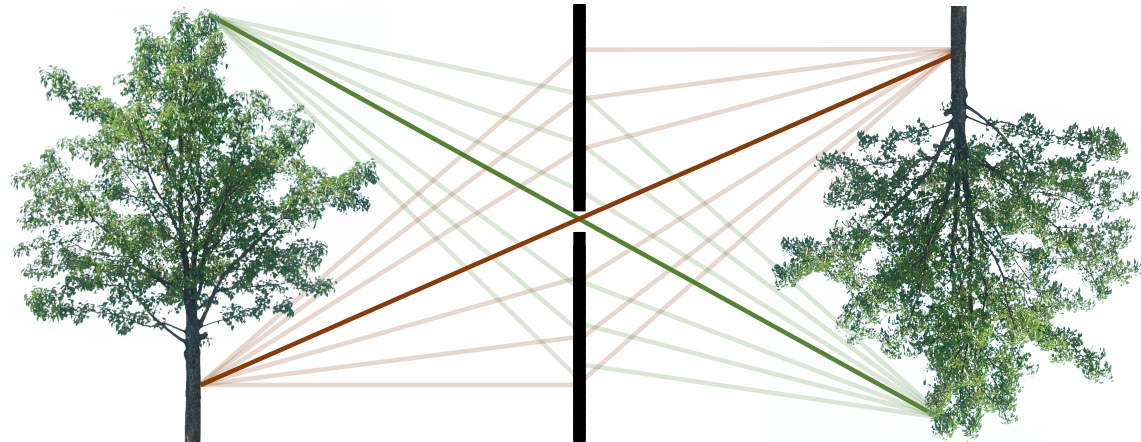
The lens camera



The pinhole camera

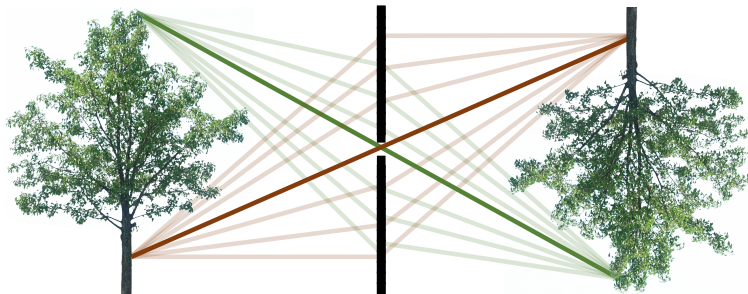
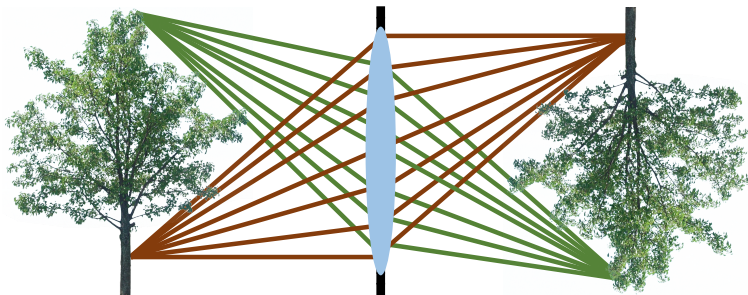


The pinhole camera



Central rays propagate in the same way for both models!

Describing both lens and pinhole cameras

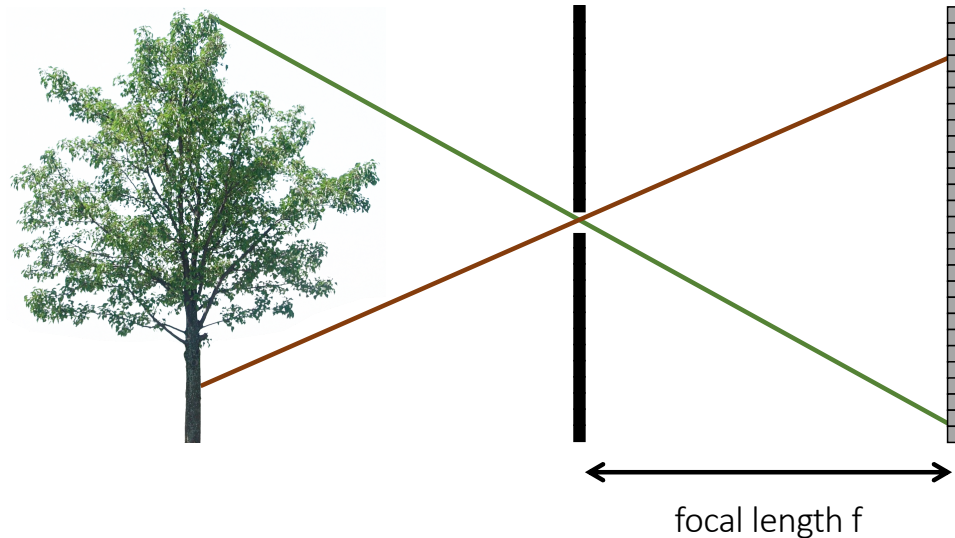


We can derive properties and descriptions that hold for both camera models if:

- We use only central rays.
- We assume the lens camera is in focus.

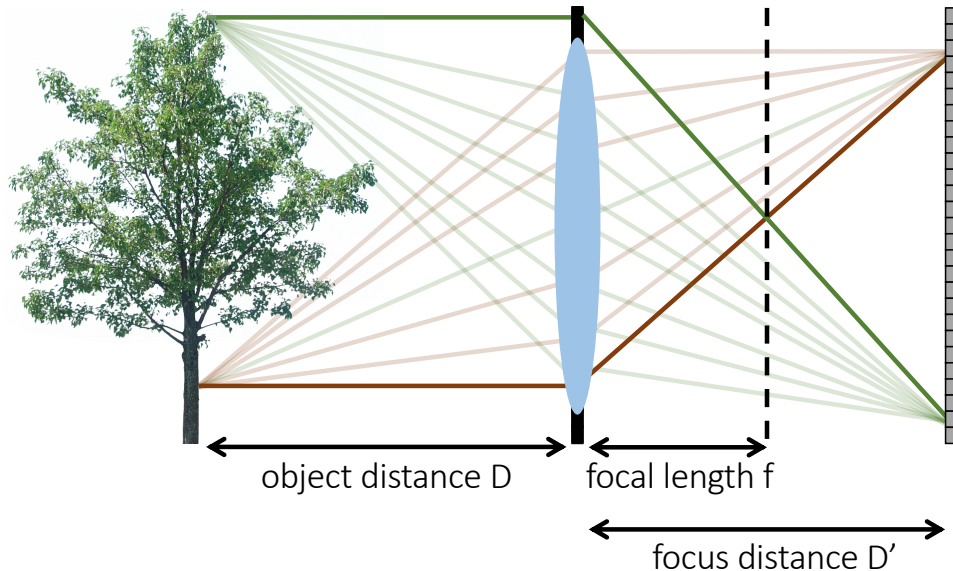
Important difference: focal length

In a pinhole camera, focal length is distance between aperture and sensor

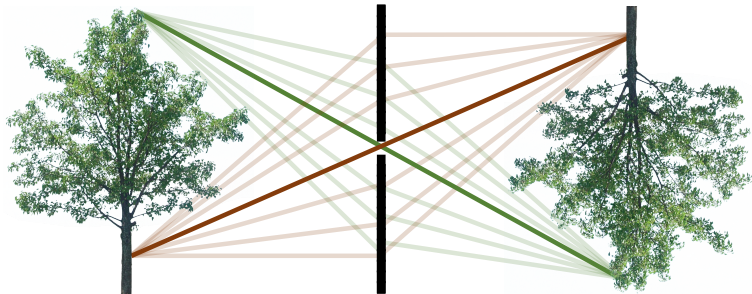
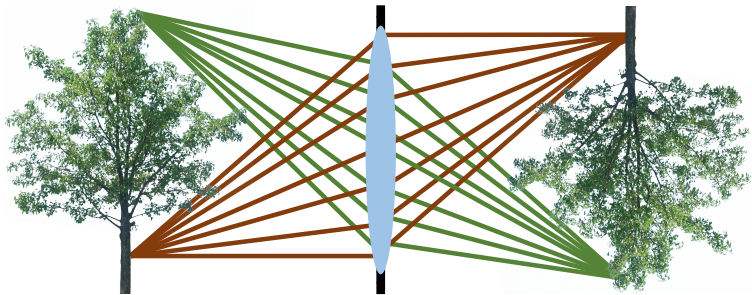


Important difference: focal length

In a lens camera, focal length is distance where parallel rays intersect



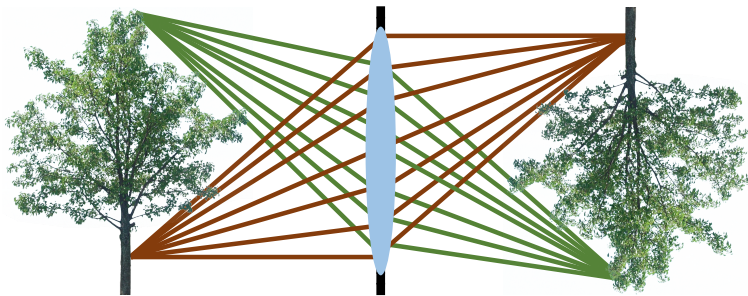
Describing both lens and pinhole cameras



We can derive properties and descriptions that hold for both camera models if:

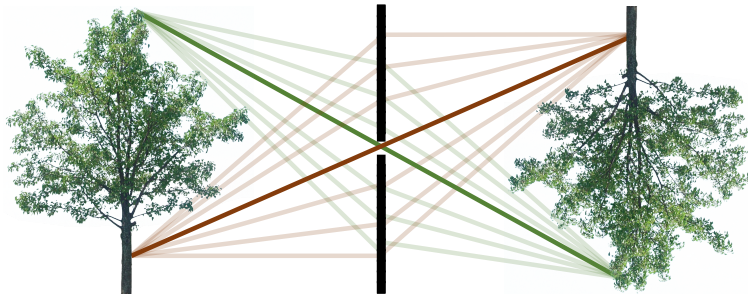
- We use only central rays.
- We assume the lens camera is in focus.
- We assume that the focus distance of the lens camera is equal to the focal length of the pinhole camera.

Describing both lens and pinhole cameras



We can derive properties and descriptions that hold for both camera models if:

- We use only central rays.
- We assume the lens camera is in focus.
- We assume that the focus distance of the lens camera is equal to the focal length of the pinhole camera.



Remember: *focal length* f refers to different things for lens and pinhole cameras.

- In this lecture, we use it to refer to the aperture-sensor distance, as in the pinhole camera case.

Camera matrix

The camera as a coordinate transformation

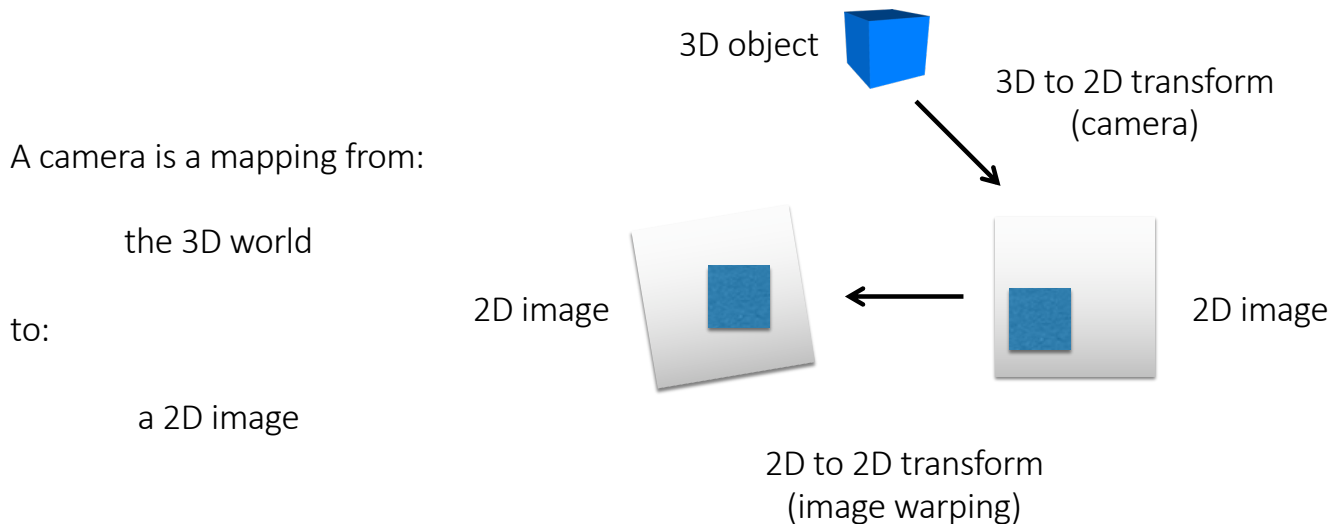
A camera is a mapping from:

the 3D world

to:

a 2D image

The camera as a coordinate transformation



The camera as a coordinate transformation

A camera is a mapping from:

the 3D world

to:

a 2D image

homogeneous coordinates

The diagram shows the equation $x = PX$ in a large, bold, italicized font. Above the equation, the text "homogeneous coordinates" has two arrows pointing down to the x and X terms. Below the equation, the terms are labeled: "2D image point" under x , "camera matrix" under P , and "3D world point" under X .

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

2D image point camera matrix 3D world point

What are the dimensions of each variable?

Reminder: 2D homogeneous coordinates

heterogeneous
coordinates

homogeneous
coordinates

$$\begin{bmatrix} x \\ y \end{bmatrix} \Rightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

← add a 1 here

- Represent 2D point with a 3D vector

Reminder: 2D homogeneous coordinates

heterogeneous
coordinates

homogeneous
coordinates

$$\begin{bmatrix} x \\ y \end{bmatrix} \Rightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \stackrel{\text{def}}{=} \begin{bmatrix} ax \\ ay \\ a \end{bmatrix}$$

- Represent 2D point with a 3D vector
- 3D vectors are only defined up to scale

Reminder: 2D homogeneous coordinates

Conversion:

- heterogeneous $\rightarrow$ homogeneous

$$\begin{bmatrix} x \\ y \end{bmatrix} \Rightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

- homogeneous $\rightarrow$ heterogeneous

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} \Rightarrow \begin{bmatrix} x/z \\ y/z \end{bmatrix}$$

Scale invariance:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = a \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Reminder: 2D homogeneous coordinates

Conversion:

- heterogeneous $\rightarrow$ homogeneous

$$\begin{bmatrix} x \\ y \end{bmatrix} \Rightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

- homogeneous $\rightarrow$ heterogeneous

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} \Rightarrow \begin{bmatrix} x/z \\ y/z \end{bmatrix}$$

Scale invariance:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = a \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Special points:

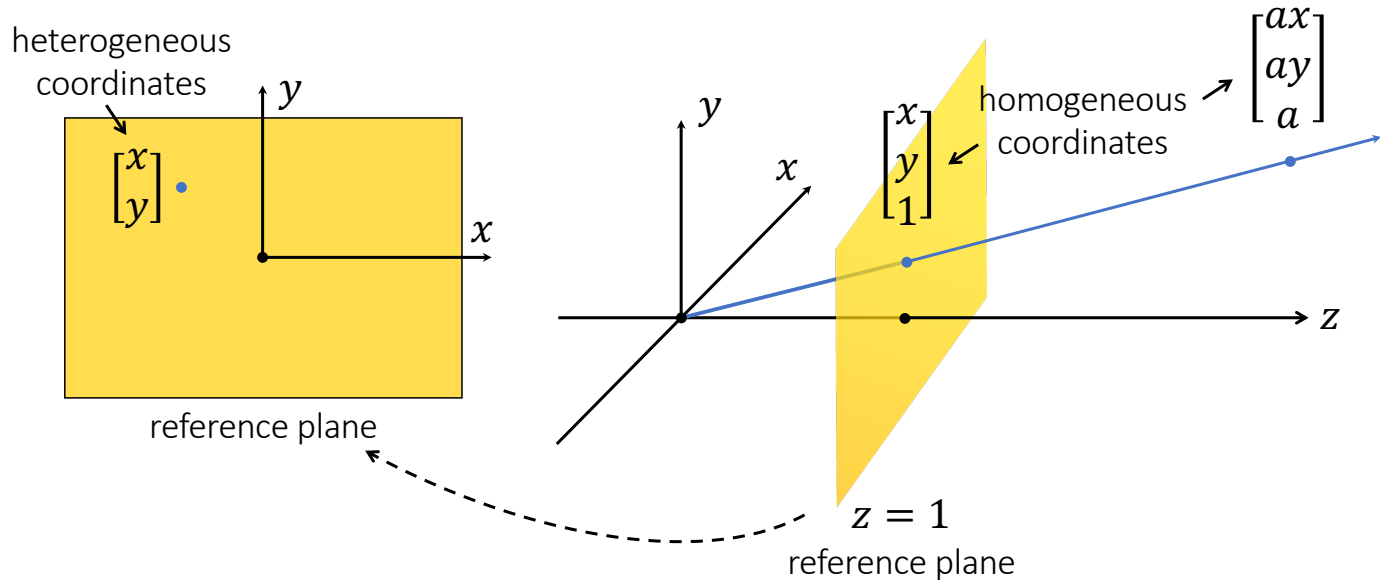
- point at infinity

$$\begin{bmatrix} x \\ y \\ 0 \end{bmatrix}$$

- undefined

$$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Reminder: 2D projective geometry



Through the scale invariance property, homogeneous coordinates map all points on a line passing through the origin to the point where this line intersects the reference plane.

Reminder: 3D homogeneous coordinates

heterogeneous
coordinates

homogeneous
coordinates

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \Rightarrow \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \stackrel{\text{def}}{=} \begin{bmatrix} aX \\ aY \\ aZ \\ a \end{bmatrix}$$

- Represent 3D point with a 4D vector
- 4D vectors are only defined up to scale

Reminder: notation

heterogeneous coordinates

homogeneous coordinates

2D
coordinates

2D vector $\tilde{\mathbf{x}} = \begin{bmatrix} x \\ y \end{bmatrix}$

3D vector $\mathbf{x} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$

3D
coordinates

3D vector $\tilde{\mathbf{X}} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$

4D vector $\mathbf{X} = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$

The camera as a coordinate transformation

A camera is a mapping from:

the 3D world

to:

a 2D image

homogeneous coordinates

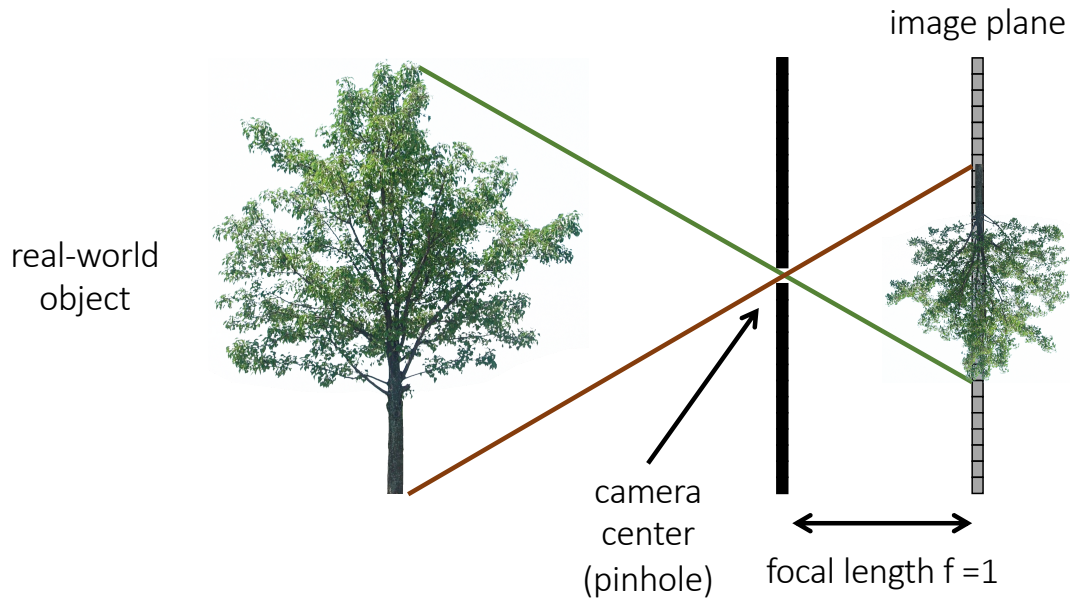
The diagram shows the equation $x = PX$ in a large, bold, italicized font. Above the equation, the text "homogeneous coordinates" is centered. Two arrows point downwards from this text: one points to the variable x and the other points to the variable X . Below the equation, the terms are labeled: "2D image point" is centered under x ; "camera matrix" is centered under P ; and "3D world point" is centered under X .

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

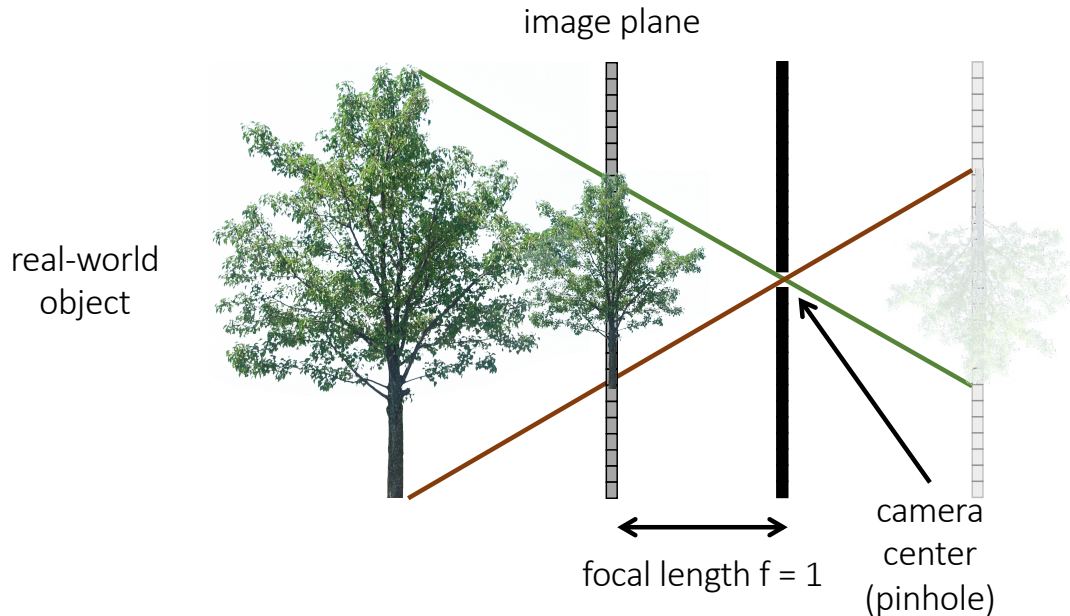
2D image point camera matrix 3D world point

What does this transformation look like?

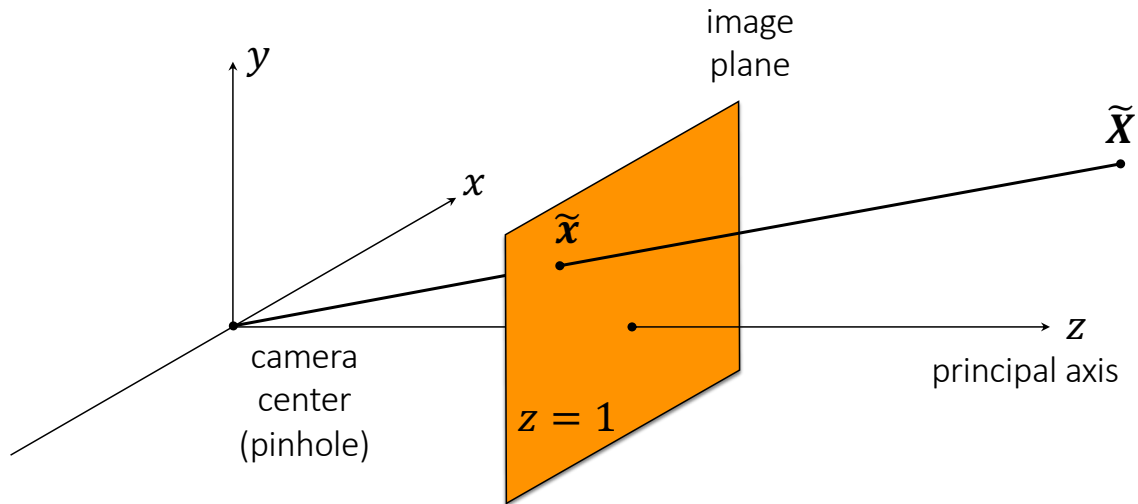
The pinhole camera



The (rearranged) pinhole camera

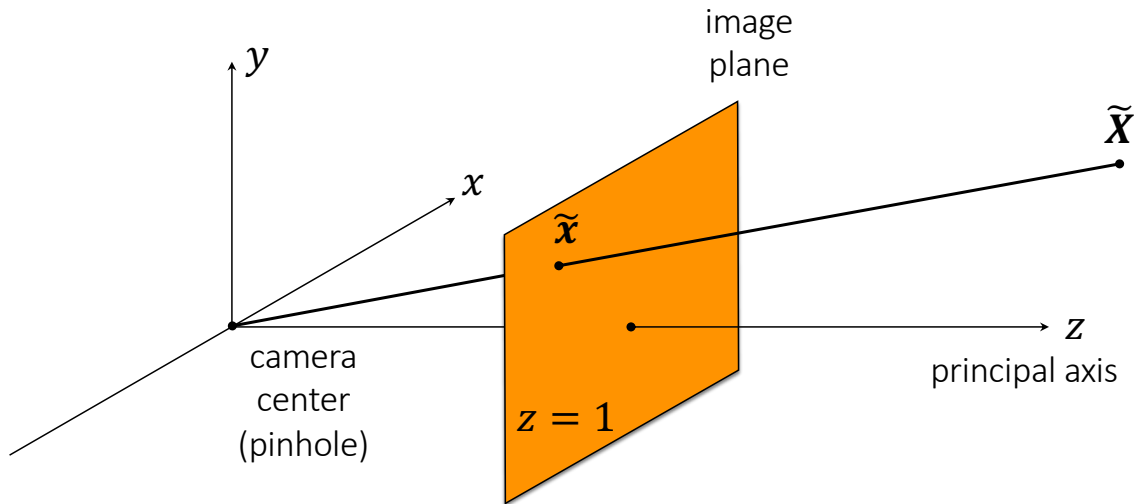


The (rearranged) pinhole camera



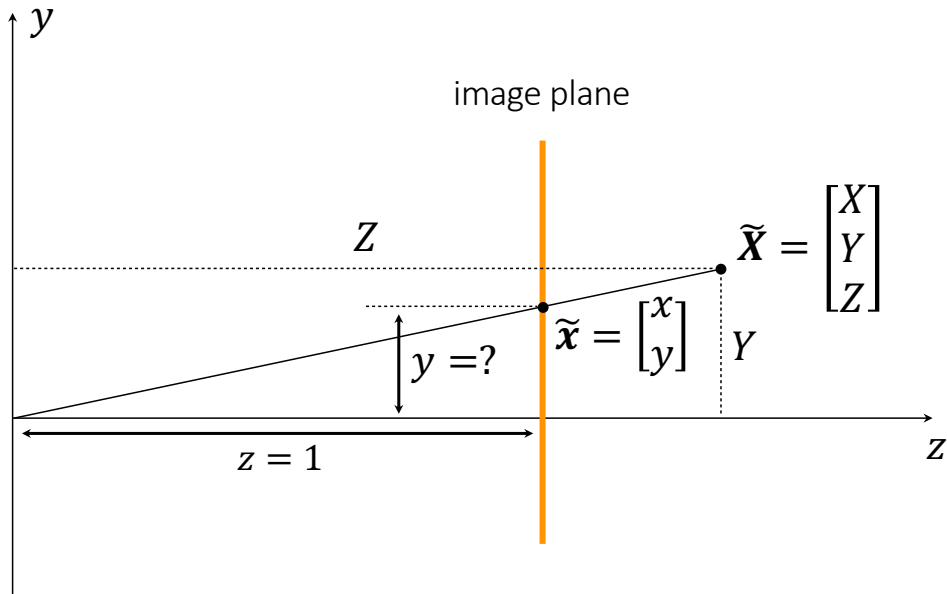
Where did we see a similar picture?

The (rearranged) pinhole camera



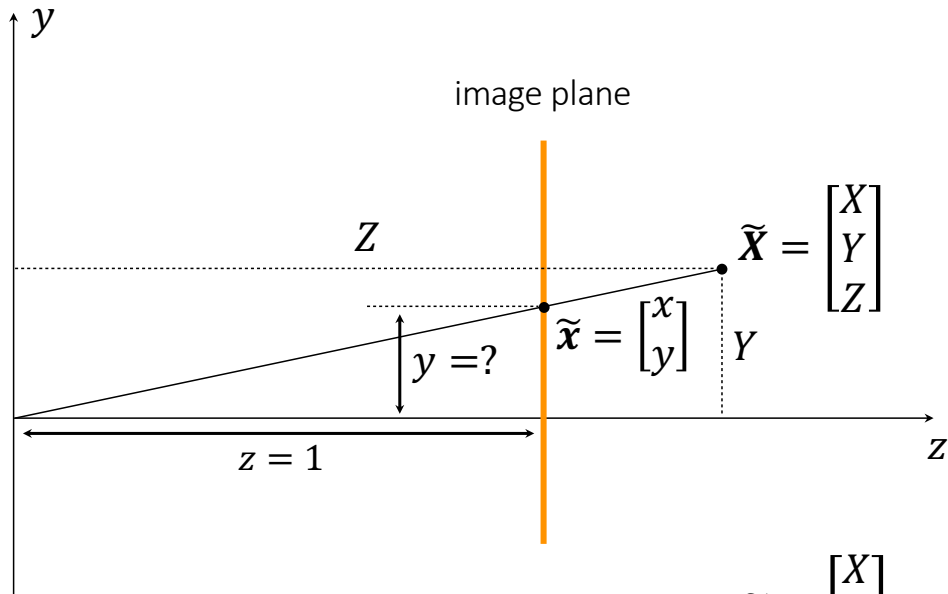
What is the equation for image coordinate $\tilde{\mathbf{x}}$ in terms of $\tilde{\mathbf{X}}$?

The 2D view of the (rearranged) pinhole camera



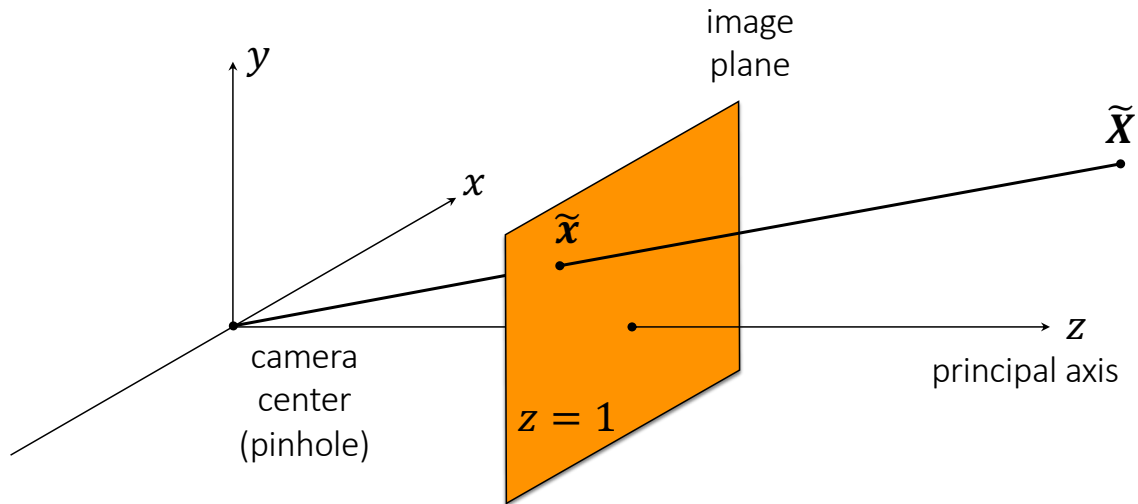
What is the equation for image coordinate $\tilde{\mathbf{x}}$ in terms of $\tilde{\mathbf{X}}$?

The 2D view of the (rearranged) pinhole camera



Using similar triangles: $\tilde{\mathbf{X}} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \rightarrow \tilde{\mathbf{x}} = \begin{bmatrix} X/Z \\ Y/Z \end{bmatrix}$

The (rearranged) pinhole camera



What is the camera matrix $\mathbf{P}$ for a pinhole camera?

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

The pinhole camera matrix

Camera projection relationship expressed:

- in *heterogeneous coordinates*
- in *homogeneous coordinates*

$$\tilde{\mathbf{X}} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \rightarrow \tilde{\mathbf{x}} = \begin{bmatrix} X/Z \\ Y/Z \end{bmatrix}$$

$$\mathbf{X} = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \rightarrow \mathbf{x} = ?$$

The pinhole camera matrix

Camera projection relationship expressed:

- in *heterogeneous coordinates*
- in *homogeneous coordinates*

$$\tilde{\mathbf{X}} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \rightarrow \tilde{\mathbf{x}} = \begin{bmatrix} X/Z \\ Y/Z \end{bmatrix}$$

$$\mathbf{X} = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \rightarrow \mathbf{x} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

General camera model in *homogeneous coordinates*:

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

What does the pinhole camera projection look like?

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

The pinhole camera matrix

Camera projection relationship expressed:

- in *heterogeneous coordinates*
- in *homogeneous coordinates*

$$\tilde{\mathbf{X}} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \rightarrow \tilde{\mathbf{x}} = \begin{bmatrix} X/Z \\ Y/Z \end{bmatrix}$$

$$\mathbf{X} = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \rightarrow \mathbf{x} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

General camera model in *homogeneous coordinates*:

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

What does the pinhole camera projection look like?

The perspective
projection matrix

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

The pinhole camera matrix

Camera projection relationship expressed:

- in *heterogeneous coordinates*
- in *homogeneous coordinates*

$$\tilde{\mathbf{X}} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \rightarrow \tilde{\mathbf{x}} = \begin{bmatrix} X/Z \\ Y/Z \end{bmatrix}$$

$$\mathbf{X} = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \rightarrow \mathbf{x} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

General camera model in *homogeneous coordinates*:

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

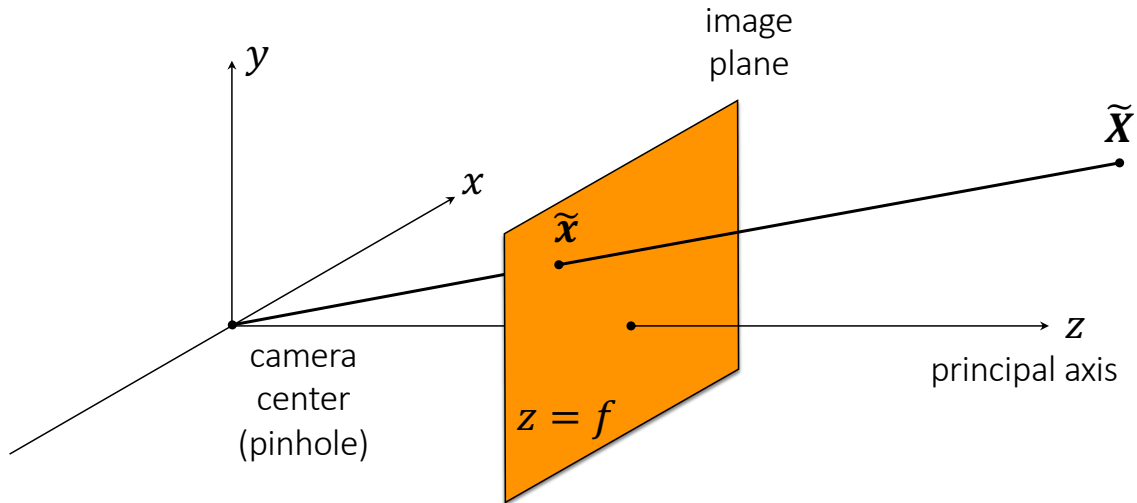
What does the pinhole camera projection look like?

The *perspective projection matrix*

$$\mathbf{P} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = [\mathbf{I} \mid \mathbf{0}]$$

alternative way to write the same thing

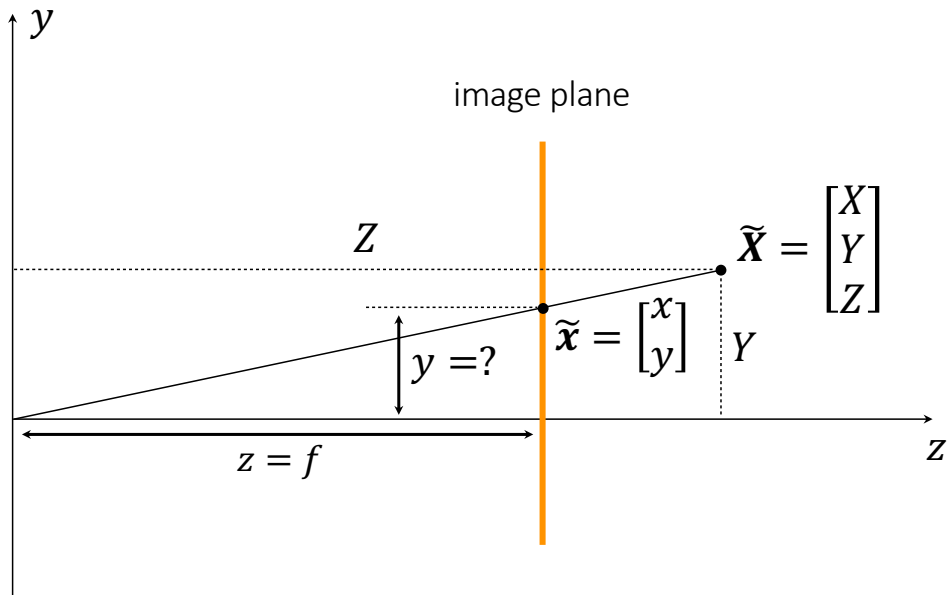
More general case: arbitrary focal length



What is the camera matrix $\mathbf{P}$ for a pinhole camera?

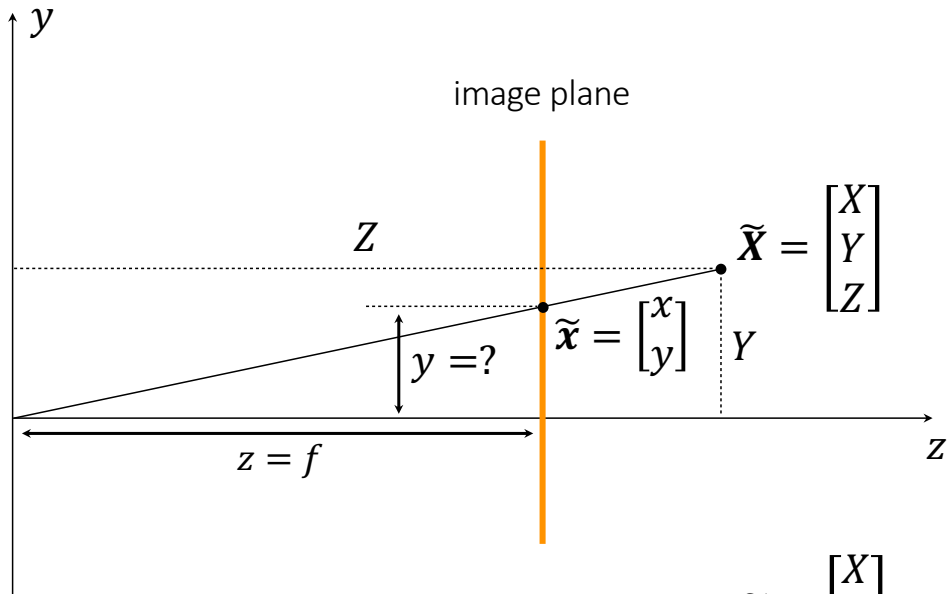
$$\mathbf{x} = \mathbf{P}\mathbf{X}$$

More general (2D) case: arbitrary focal length



What is the equation for image coordinate $\tilde{\mathbf{x}}$ in terms of $\tilde{\mathbf{X}}$?

More general (2D) case: arbitrary focal length



Using similar triangles: $\tilde{\mathbf{X}} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \rightarrow \tilde{\mathbf{x}} = \begin{bmatrix} fX/Z \\ fY/Z \end{bmatrix}$

The pinhole camera matrix for arbitrary focal length

Camera projection relationship expressed:

- in *heterogeneous coordinates*
- in *homogeneous coordinates*

$$\tilde{\mathbf{X}} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \rightarrow \tilde{\mathbf{x}} = \begin{bmatrix} fX/Z \\ fY/Z \end{bmatrix}$$

$$\mathbf{X} = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \rightarrow \mathbf{x} = \begin{bmatrix} fX/Z \\ fY/Z \end{bmatrix}$$

General camera model in *homogeneous coordinates*:

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

What does the pinhole camera projection look like?

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

The pinhole camera matrix for arbitrary focal length

Camera projection relationship expressed:

- in *heterogeneous coordinates*
- in *homogeneous coordinates*

$$\tilde{\mathbf{X}} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \rightarrow \tilde{\mathbf{x}} = \begin{bmatrix} fX/Z \\ fY/Z \end{bmatrix}$$

$$\mathbf{X} = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \rightarrow \mathbf{x} = \begin{bmatrix} fX/Z \\ fY/Z \\ Z \end{bmatrix}$$

General camera model in *homogeneous coordinates*:

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

What does the pinhole camera projection look like?

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

The pinhole camera matrix for arbitrary focal length

Camera projection relationship expressed:

- in *heterogeneous coordinates*
- in *homogeneous coordinates*

$$\tilde{\mathbf{X}} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \rightarrow \tilde{\mathbf{x}} = \begin{bmatrix} fX/Z \\ fY/Z \end{bmatrix}$$

$$\mathbf{X} = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \rightarrow \mathbf{x} = \begin{bmatrix} fX/Z \\ fY/Z \\ Z \end{bmatrix}$$

General camera model in *homogeneous coordinates*:

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

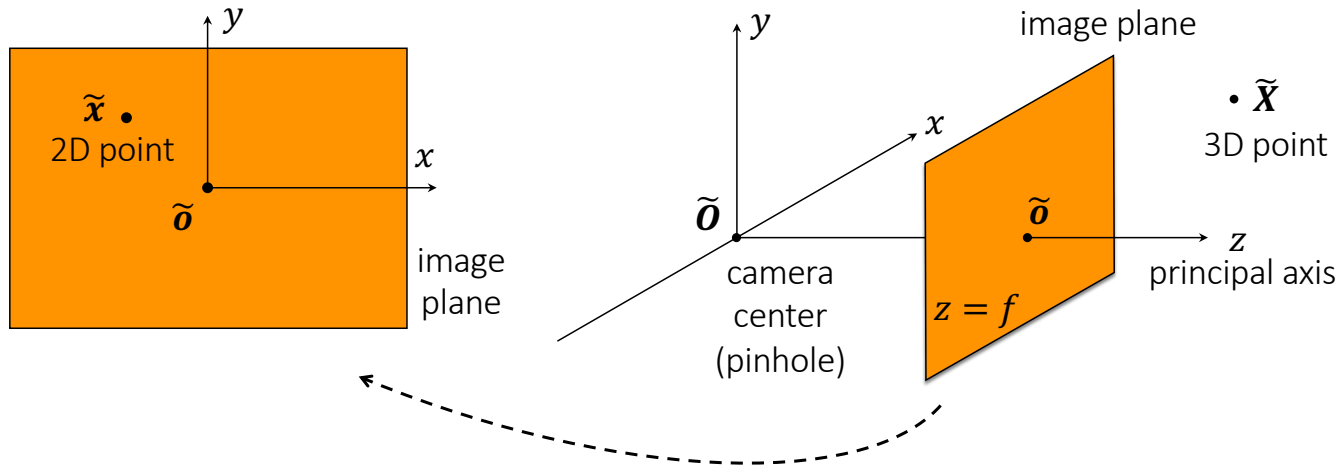
What does the pinhole camera projection look like?

Equivalently we
can write:

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

combination of perspective
projection and a 2D scaling
transformation

Generalizations: coordinate systems

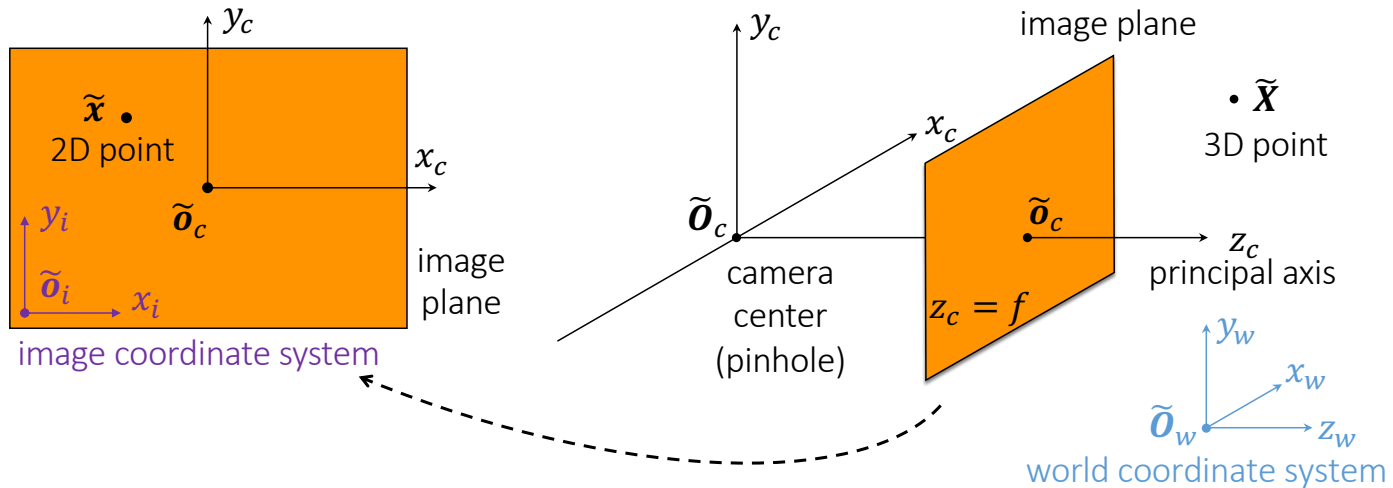


2D camera coordinate system

3D camera coordinate system

- A camera introduces two related coordinate systems, in 3D (world), and in 2D (image plane).

Generalizations: coordinate systems



2D camera coordinate system

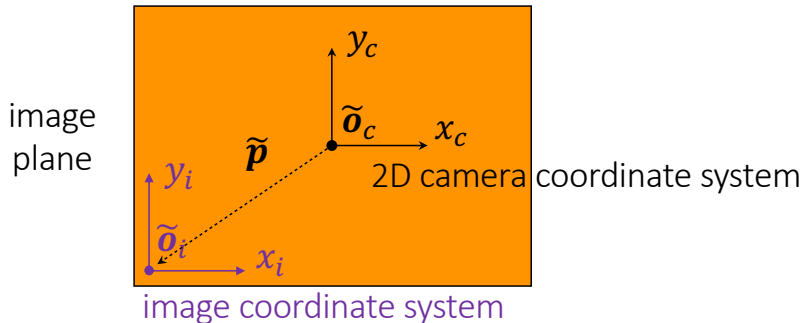
3D camera coordinate system

- A camera introduces two related coordinate systems, in 3D (world), and in 2D (image plane).
- These coordinate systems may be different from the coordinate systems of our application.

Generalization: image coordinate system

In particular, the camera origin and image origin may be different.

- Can you think of a case when this happens?



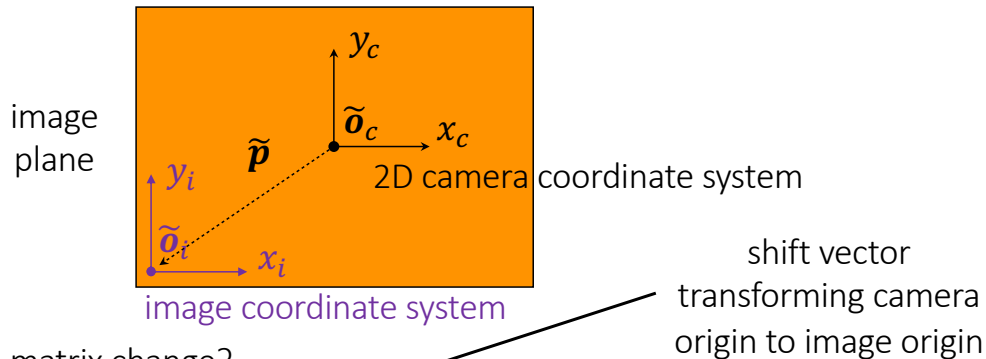
How does the camera matrix change?

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

Generalization: image coordinate system

In particular, the camera origin and image origin may be different.

- Can you think of a case when this happens?



How does the camera matrix change?

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

Camera matrix decomposition

We can decompose the camera matrix like this:

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

What does each part of the matrix represent?

Camera matrix decomposition

We can decompose the camera matrix like this:

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



(homogeneous) transformation
from 2D to 2D, accounting for non-
unit focal length and origin shift

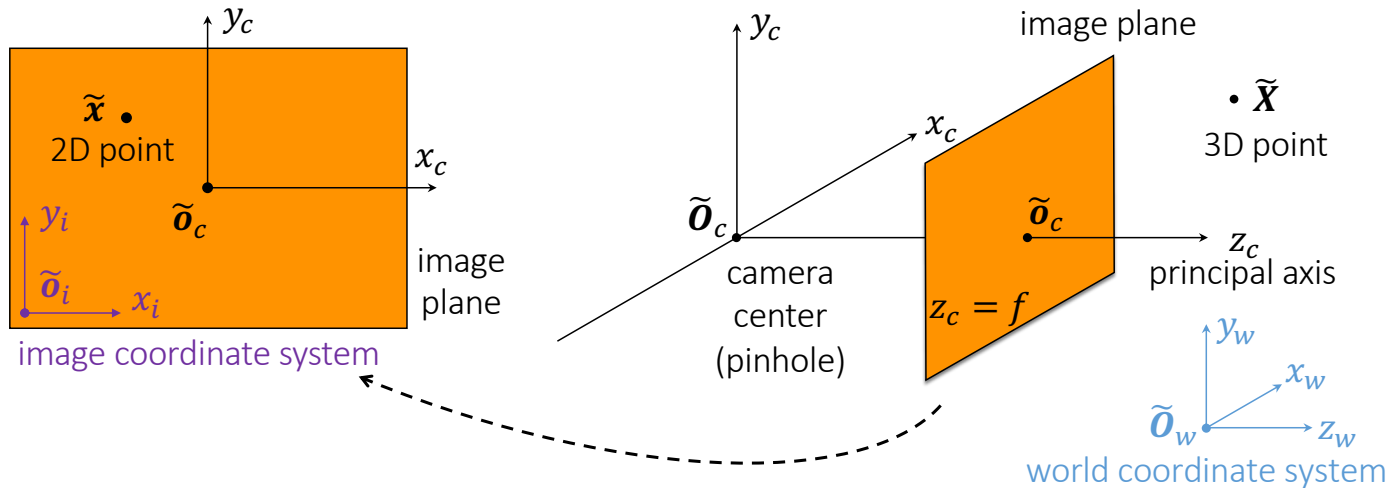


(homogeneous) perspective projection
from 3D to 2D, assuming image plane at
 $z = 1$ and shared camera/image origin

Also written as:

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

Generalizations: coordinate systems

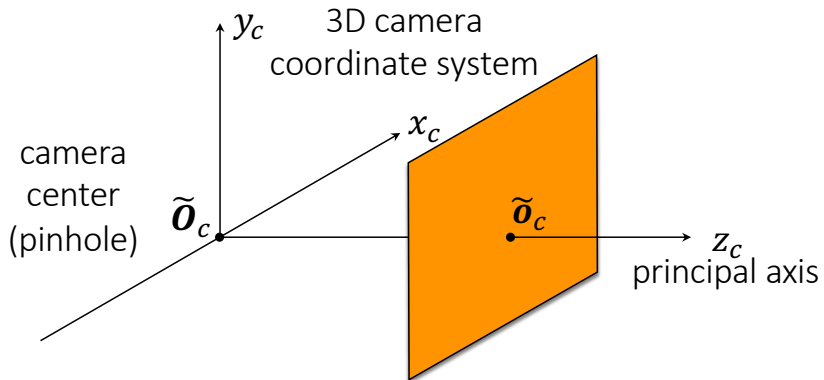


2D camera coordinate system

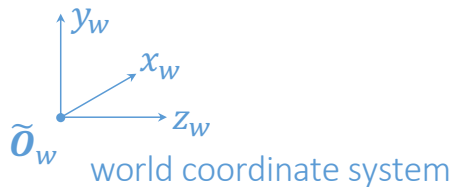
3D camera coordinate system

- A camera introduces two related coordinate systems, in 3D (world), and in 2D (image plane).
- These coordinate systems may be different from the coordinate systems of our application.

World-to-camera coordinate system transformation



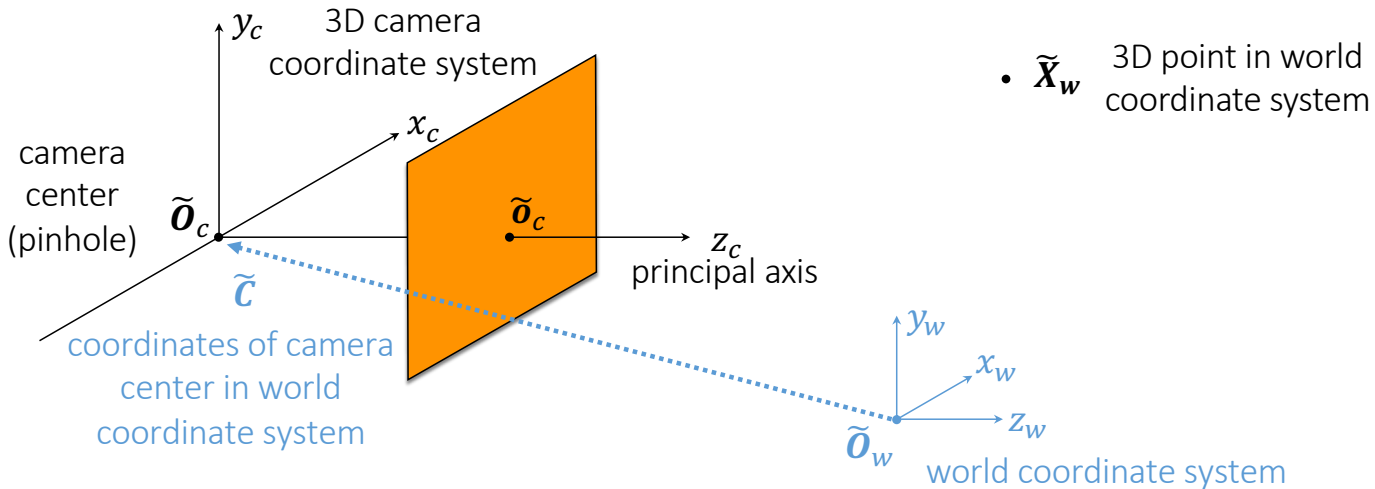
- $\tilde{\mathbf{X}}_w$ 3D point in world coordinate system



How do we express $\tilde{\mathbf{X}}$ in the 3D camera coordinate system?

$$\tilde{\mathbf{X}}_w$$

World-to-camera coordinate system transformation

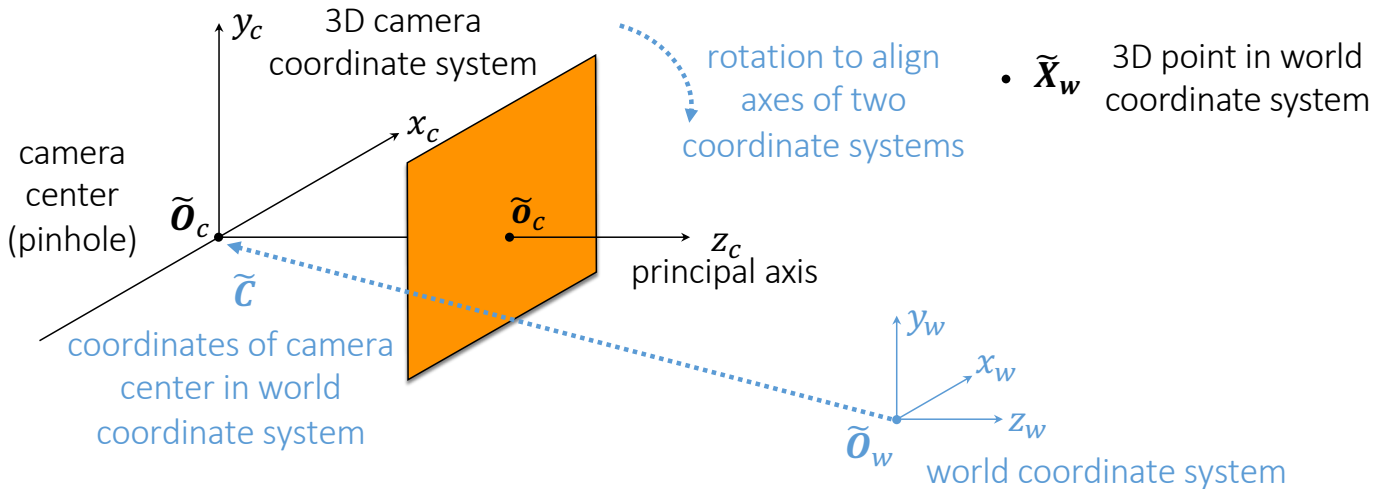


How do we express $\tilde{\mathbf{X}}$ in the 3D camera coordinate system?

$$\tilde{\mathbf{X}}_w - \tilde{\mathbf{c}}$$

translate

World-to-camera coordinate system transformation



How do we express $\tilde{X}$ in the 3D camera coordinate system?

$$R \cdot (\tilde{X}_w - \tilde{c})$$

rotate translate

Modeling the 3D coordinate system transformation

In heterogeneous coordinates, we have:

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

How do we write this transformation in homogeneous coordinates?

Modeling the 3D coordinate system transformation

In heterogeneous coordinates, we have:

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

In homogeneous coordinates, we have:

$$\mathbf{X}_c = \begin{bmatrix} \mathbf{R} & -\mathbf{R}\tilde{\mathbf{C}} \\ \mathbf{0} & 1 \end{bmatrix} \mathbf{X}_w$$

Incorporating the transform in the camera matrix

The previous camera matrix is for homogeneous 3D coordinates in camera coordinate system:

$$\mathbf{x} = \mathbf{P}\mathbf{X}_c = \begin{bmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix} [\mathbf{I} \mid \mathbf{0}]\mathbf{X}_c$$

We also just derived:

$$\mathbf{X}_c = \begin{bmatrix} \mathbf{R} & -\mathbf{R}\tilde{\mathbf{C}} \\ \mathbf{0} & 1 \end{bmatrix} \mathbf{X}_w$$

Putting it all together

We can write everything into a single projection:

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

The camera matrix now looks like:

$$\mathbf{P} = \begin{bmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix} [\mathbf{I} \mid \mathbf{0}] \begin{bmatrix} \mathbf{R} & -\mathbf{R}\tilde{\mathbf{C}} \\ \mathbf{0} & 1 \end{bmatrix}$$



intrinsic parameters (3 x 3):

correspond to camera

internals (2D image-to-image
transformation)

Putting it all together

We can write everything into a single projection:

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

The camera matrix now looks like:

$$\mathbf{P} = \begin{bmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix} [\mathbf{I} \mid \mathbf{0}] \begin{bmatrix} \mathbf{R} & -\mathbf{R}\tilde{\mathbf{C}} \\ \mathbf{0} & 1 \end{bmatrix}$$



intrinsic parameters (3 x 3):
correspond to camera
internals (2D image-to-image
transformation)



perspective projection (3 x 4):
maps 3D to 2D points
(camera-to-image
transformation)

Putting it all together

We can write everything into a single projection:

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

The camera matrix now looks like:

$$\mathbf{P} = \begin{bmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix} [\mathbf{I} \mid \mathbf{0}] \begin{bmatrix} \mathbf{R} & -\mathbf{R}\tilde{\mathbf{C}} \\ \mathbf{0} & 1 \end{bmatrix}$$

intrinsic parameters (3 x 3):
correspond to camera
internals (2D image-to-image
transformation)

perspective projection (3 x 4):
maps 3D to 2D points
(camera-to-image
transformation)

extrinsic parameters (4 x 4):
correspond to camera
externals (3D world-to-camera
transformation)

Putting it all together

We can write everything into a single projection:

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

The camera matrix now looks like:

$$\mathbf{P} = \begin{bmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix} [\mathbf{R} \mid -\mathbf{R}\tilde{\mathbf{C}}]$$

It is common to combine the perspective projection and extrinsics in one matrix.



intrinsic parameters (3 x 3):
correspond to camera
internals (2D image-to-image
transformation)



extrinsic parameters (3 x 4):
correspond to camera externals (3D
world-to-camera transformation)
and perspective projection

The pinhole camera matrix

More compactly, we can write the pinhole camera matrix as:

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

where

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

2D Euclidean transform

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

3D rotation

$$\mathbf{t} = -\mathbf{R}\tilde{\mathbf{C}} = \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix}$$

3D translation

intrinsic parameters

extrinsic parameters

More general pinhole camera matrices

The following is the standard pinhole camera matrix we saw.

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

How many degrees of freedom does this matrix have?

More general pinhole camera matrices

The following is the standard pinhole camera matrix we saw.

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

How many degrees of freedom does this matrix have?

- 9 degrees of freedom (3 for intrinsics, 3 for rotation, 3 for translation).

More general pinhole camera matrices

The following is the standard pinhole camera matrix we saw.

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

How many degrees of freedom does this matrix have?

- 9 degrees of freedom (3 for intrinsics, 3 for rotation, 3 for translation).

We can get more general pinhole cameras with more degrees of freedom by generalizing the intrinsics matrix, while leaving everything else the same..

More general pinhole camera matrices

CCD camera: pixels may not be square.

$$\mathbf{P} = \begin{bmatrix} a_x & 0 & p_x \\ 0 & a_y & p_y \\ 0 & 0 & 1 \end{bmatrix} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array} \right] \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix}$$

How many degrees of freedom does this matrix have?

More general pinhole camera matrices

CCD camera: pixels may not be square.

$$\mathbf{P} = \begin{bmatrix} a_x & 0 & p_x \\ 0 & a_y & p_y \\ 0 & 0 & 1 \end{bmatrix} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array} \right] \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix}$$

How many degrees of freedom does this matrix have?

- 10 degrees of freedom (4 for intrinsics, 3 for rotation, 3 for translation).

More general pinhole camera matrices

Finite projective camera: sensor may be skewed.

$$\mathbf{P} = \begin{bmatrix} a_x & s & p_x \\ 0 & a_y & p_y \\ 0 & 0 & 1 \end{bmatrix} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array} \right] \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix}$$

How many degrees of freedom does this matrix have?

More general pinhole camera matrices

Finite projective camera: sensor may be skewed.

$$\mathbf{P} = \begin{bmatrix} a_x & s & p_x \\ 0 & a_y & p_y \\ 0 & 0 & 1 \end{bmatrix} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array} \right] \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix}$$

How many degrees of freedom does this matrix have?

- 11 degrees of freedom (5 for intrinsics, 3 for rotation, 3 for translation).

Can we get a *perspective projection* camera with more degrees of freedom?

More general pinhole camera matrices

Finite projective camera: sensor may be skewed.

The finite projective camera is the most general camera implementing perspective projection.

$$\mathbf{P} = \begin{bmatrix} a_x & s & p_x \\ 0 & a_y & p_y \\ 0 & 0 & 1 \end{bmatrix} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array} \right] \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix}$$

How many degrees of freedom does this matrix have?

- 11 degrees of freedom (5 for intrinsics, 3 for rotation, 3 for translation).

Can we get a *perspective projection* camera with more degrees of freedom?

- No, as the entire camera matrix $\mathbf{P}$ has 12 elements (3x4) and is defined up to scale.

More general pinhole camera matrices

Finite projective camera: sensor may be skewed.

The finite projective camera is the most general camera implementing perspective projection.

$$\mathbf{P} = \begin{bmatrix} a_x & s & p_x \\ 0 & a_y & p_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & | & 0 \\ 0 & 1 & 0 & | & 0 \\ 0 & 0 & 1 & | & 0 \end{bmatrix} \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix}$$

How many degrees of freedom does this matrix have?

- 11 degrees of freedom (5 for intrinsics, 3 for rotation, 3 for translation).

Can we get a *perspective projection* camera with more degrees of freedom?

- No, as the entire camera matrix $\mathbf{P}$ has 12 elements (3x4) and is defined up to scale.

Perspective distortion

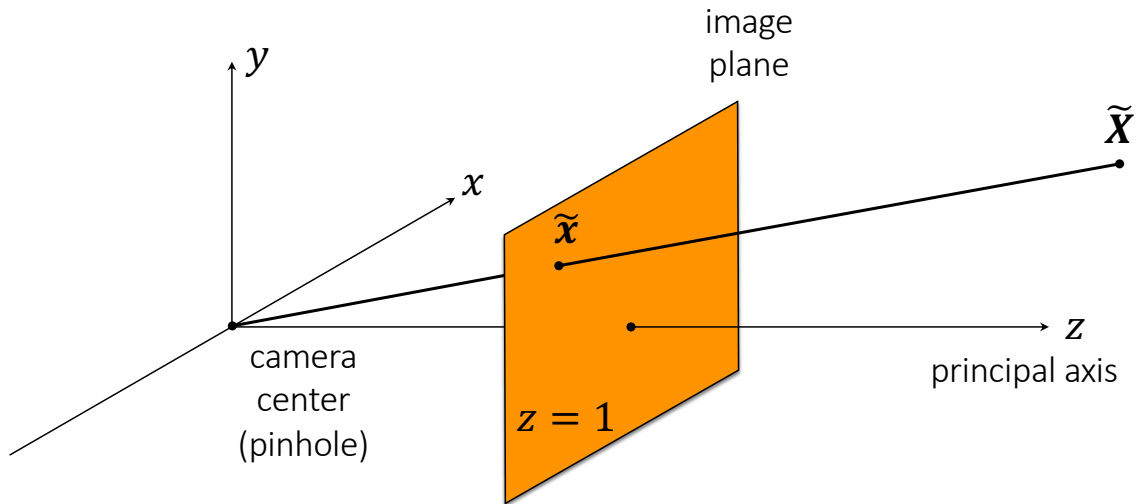
Finite projective camera

Let's ignore intrinsics and extrinsics for now.

$$\mathbf{P} = \begin{bmatrix} a_x & s & p_x \\ 0 & a_y & p_y \\ 0 & 0 & 1 \end{bmatrix} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array} \right] \begin{bmatrix} \mathbf{R} \\ \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{t} \\ 1 \end{bmatrix}$$

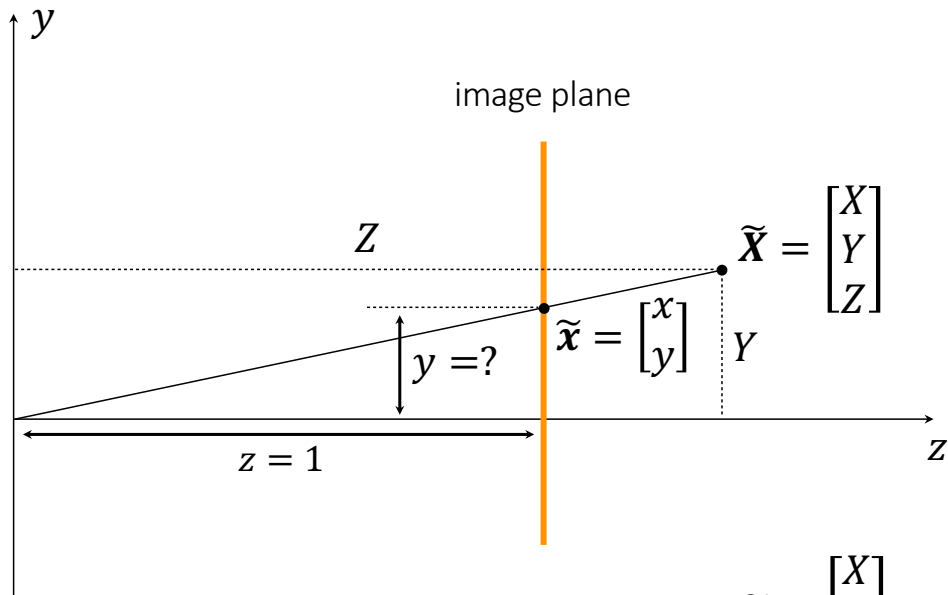
What is the effect of the perspective projection matrix?

The (rearranged) pinhole camera



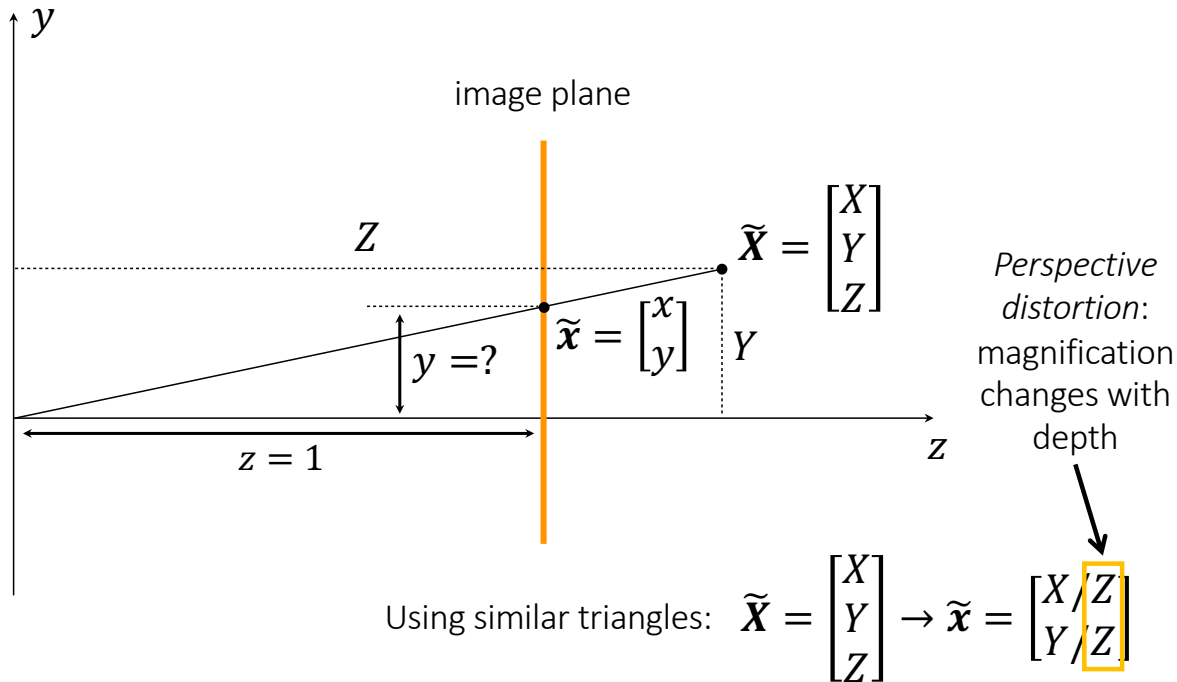
What is the equation for image coordinate $\tilde{\mathbf{x}}$ in terms of $\tilde{\mathbf{X}}$?

The 2D view of the (rearranged) pinhole camera



Using similar triangles: $\tilde{\mathbf{X}} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \rightarrow \tilde{\mathbf{x}} = \begin{bmatrix} X/Z \\ Y/Z \end{bmatrix}$

The 2D view of the (rearranged) pinhole camera



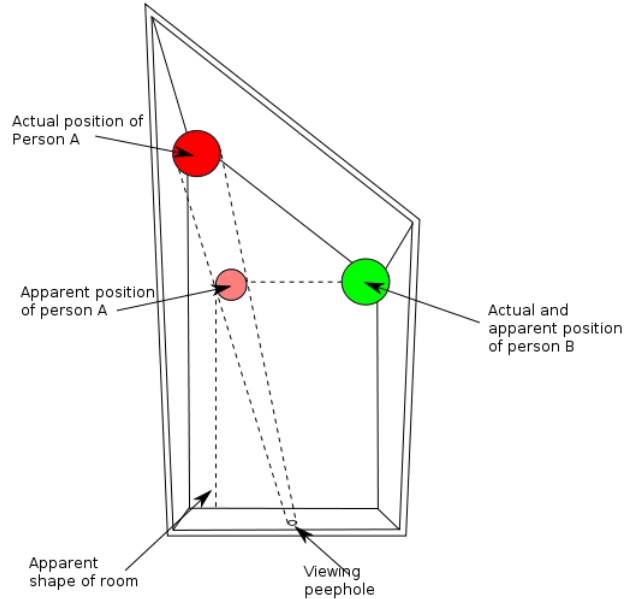
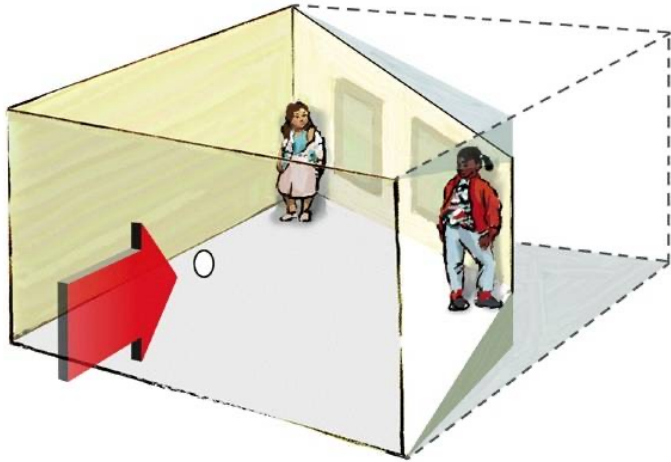
Forced perspective



The Ames room illusion



The Ames room illusion



The arrow illusion

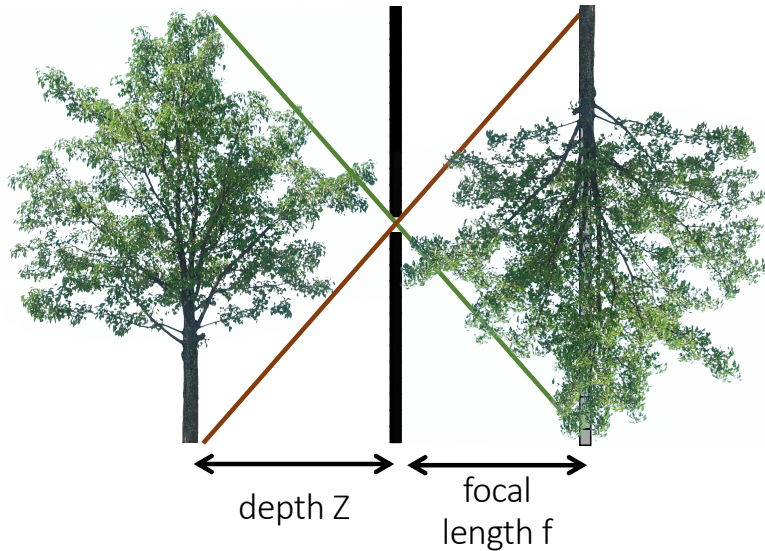


Is there a camera without perspective distortion?

Other camera models

What if...

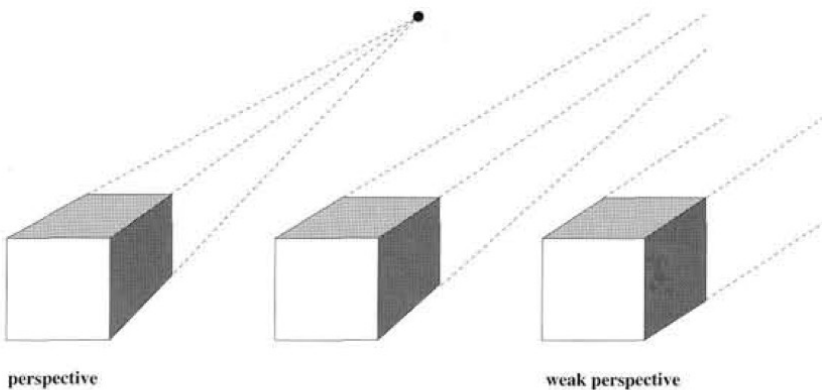
real-world
object



... we continue increasing Z
and f while maintaining
same magnification?

$$f \rightarrow \infty \text{ and } \frac{f}{Z} = \text{constant}$$

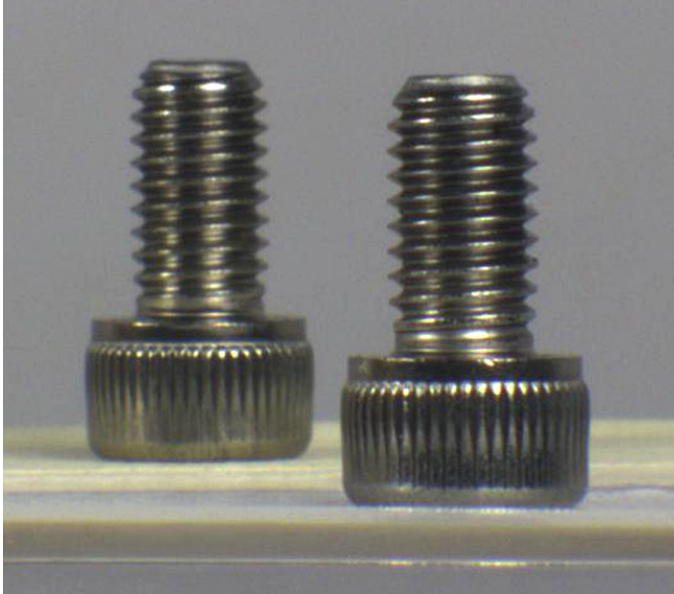
Perspective camera:
camera is *close* to
object and has
small focal length



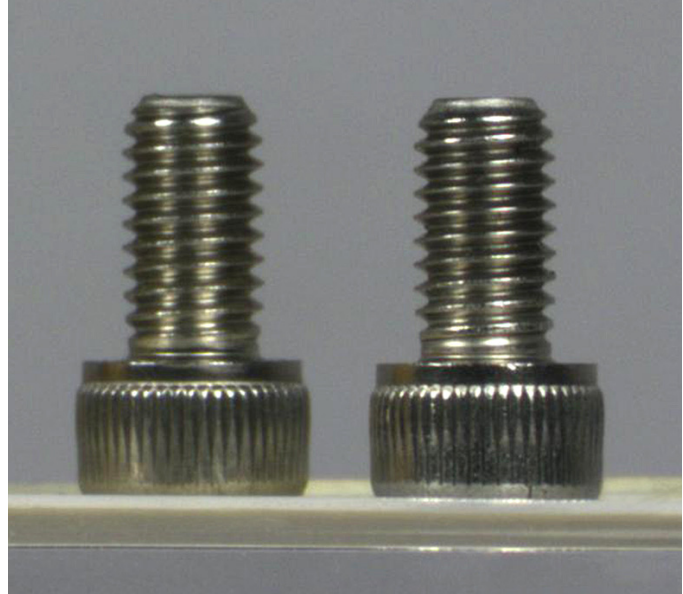
Weak-perspective
camera: camera is *far*
from object and has
large focal length



Different cameras

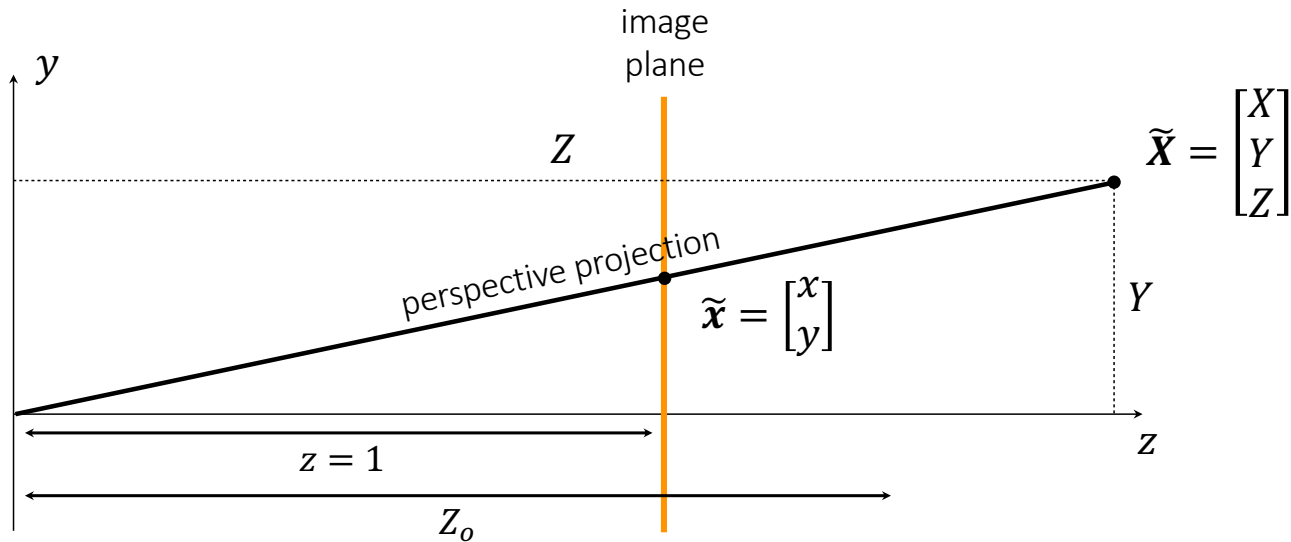


perspective camera



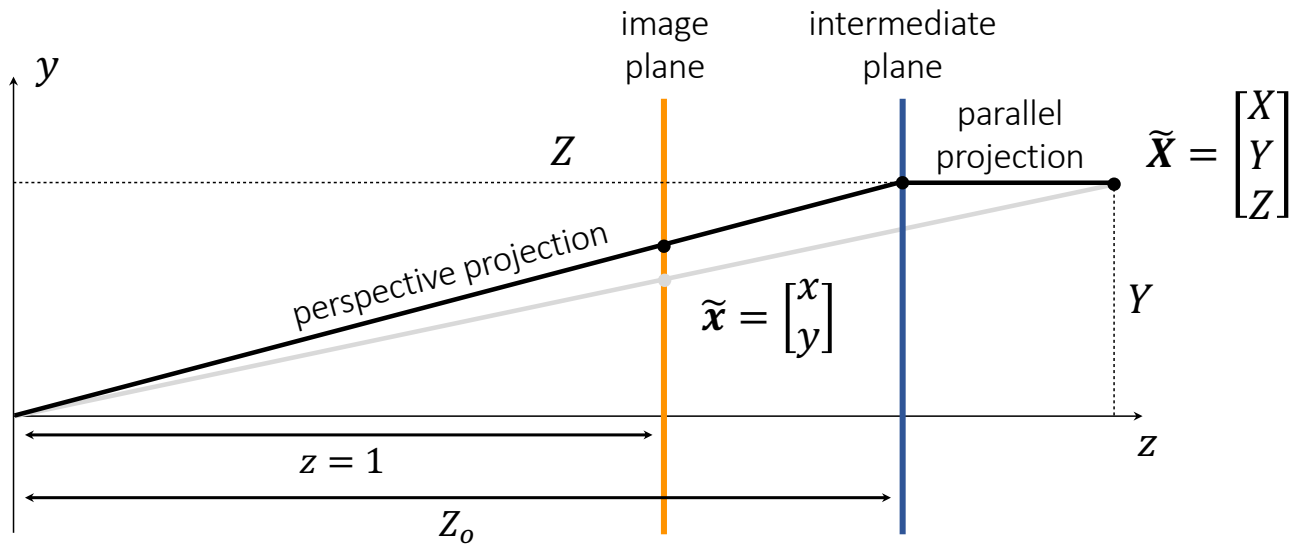
weak perspective camera

Perspective versus weak-perspective camera



perspective projection $\tilde{\mathbf{X}} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \rightarrow \tilde{\mathbf{x}} = \begin{bmatrix} X/Z \\ Y/Z \end{bmatrix}$

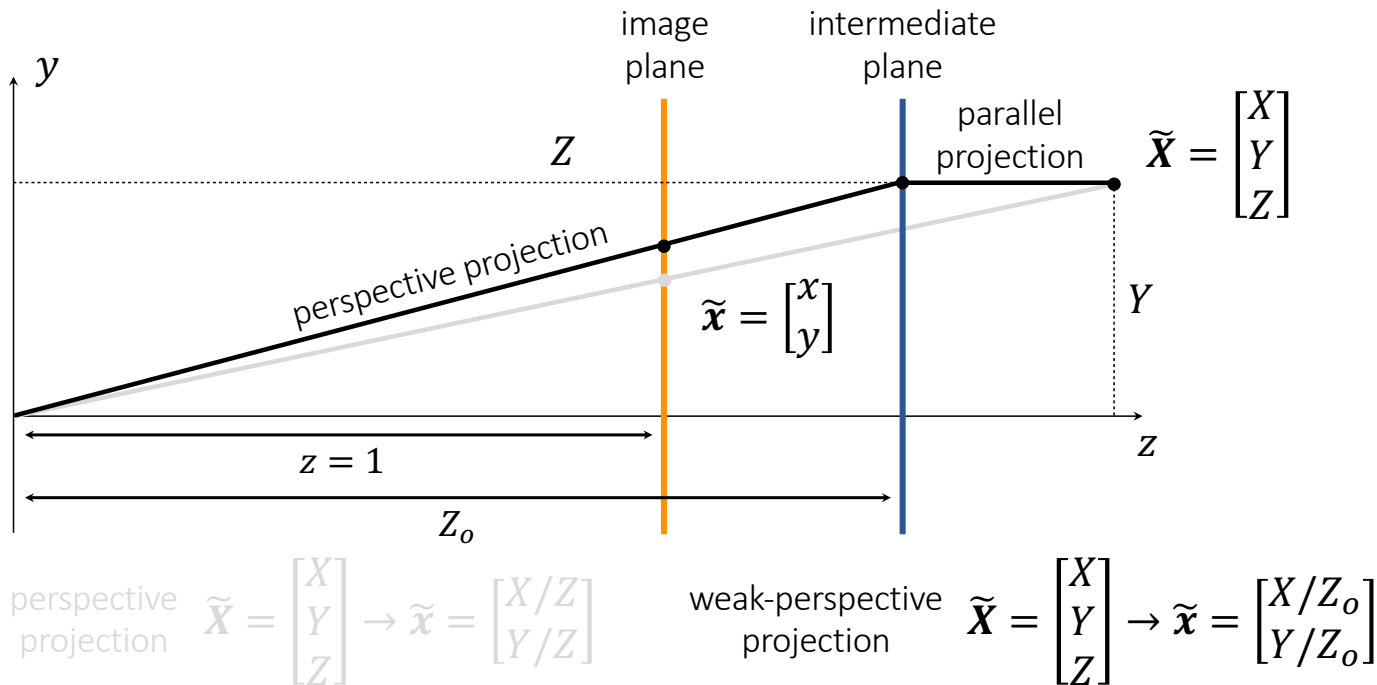
Perspective versus weak-perspective camera



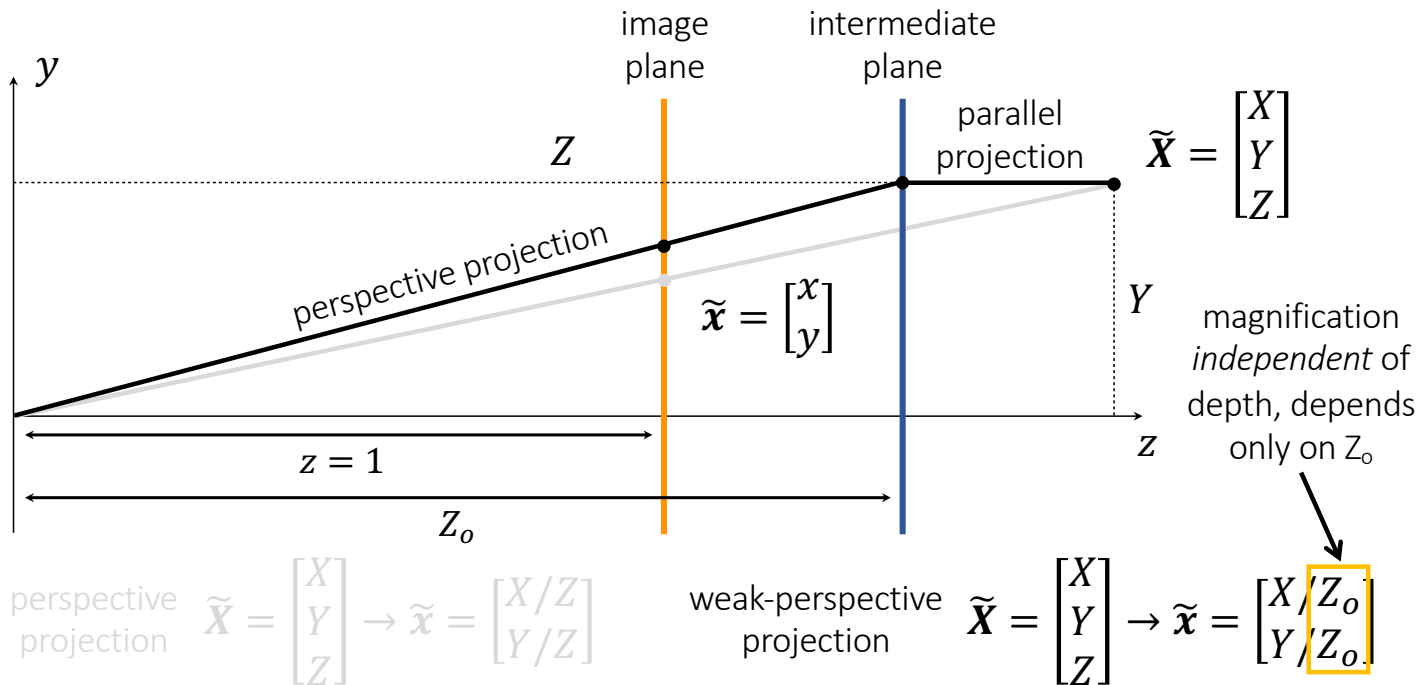
perspective projection $\tilde{\mathbf{X}} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \rightarrow \tilde{\mathbf{x}} = \begin{bmatrix} X/Z \\ Y/Z \end{bmatrix}$

weak-perspective projection $\tilde{\mathbf{X}} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \rightarrow \tilde{\mathbf{x}} = ?$

Perspective versus weak-perspective camera



Perspective versus weak-perspective camera



Comparing camera projection matrices

Let's ignore intrinsics and extrinsics for now.

- The *perspective projection matrix* can be written as:

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

- What would the *weak-perspective projection matrix* look like?

Comparing camera projection matrices

Let's ignore intrinsics and extrinsics for now.

- The *perspective projection matrix* can be written as:

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

- The *weak-perspective projection matrix* can be written as:

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

Comparing camera matrices

Let's now incorporate intrinsics and extrinsics.

- The *finite projective camera matrix* can be written as:

$$\mathbf{P} = \mathbf{K} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix}$$

- What would the matrix of the so-called *affine camera* look like?

Comparing camera matrices

Let's now incorporate intrinsics and extrinsics.

- The *finite projective camera matrix* can be written as:

$$\mathbf{P} = \mathbf{K} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix}$$

Change only the projection matrix, and use the exact same matrices for intrinsics and extrinsics.

- The *affine camera matrix* can be written as:

$$\mathbf{P} = \mathbf{K} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & Z_o \end{bmatrix} \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix}$$

Special case: orthographic projection

Let's now incorporate intrinsics and extrinsics.

- The *finite projective camera matrix* can be written as:

$$\mathbf{P} = \mathbf{K} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix}$$

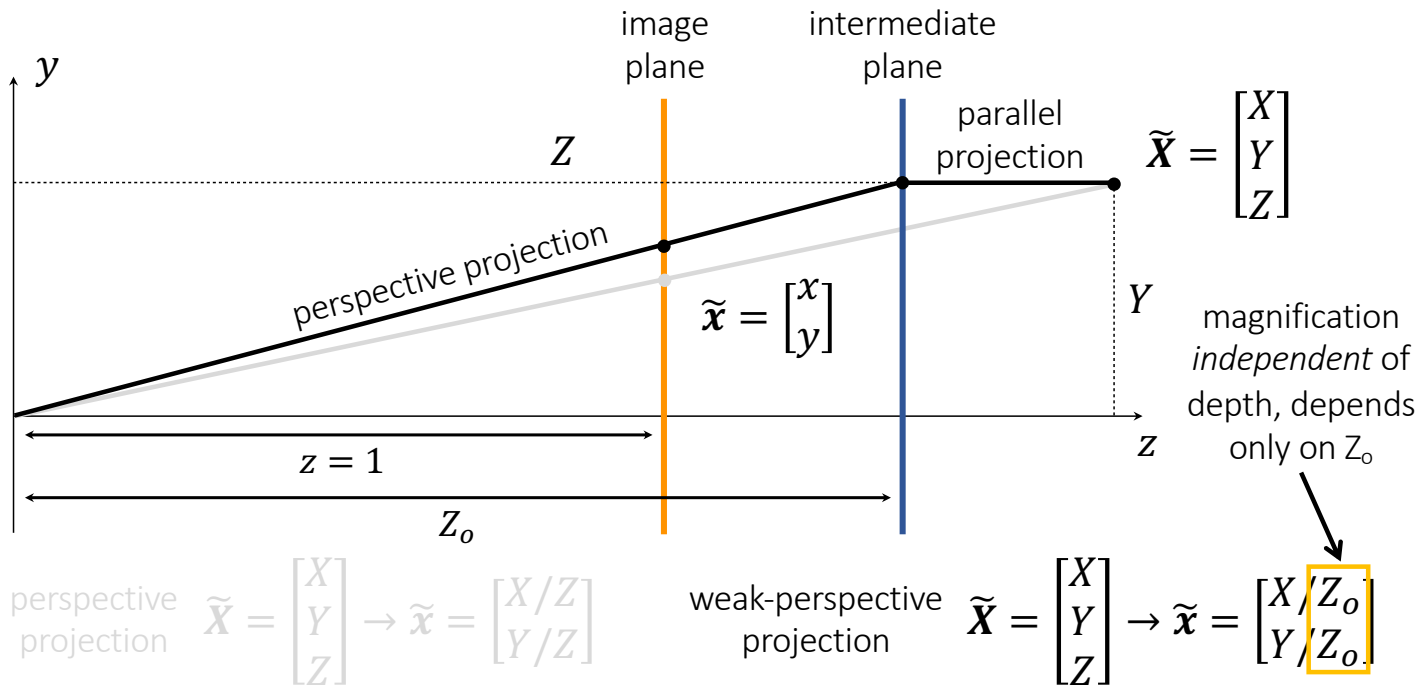
Change only the projection matrix, and use the exact same matrices for intrinsics and extrinsics.

- The *affine camera matrix* can be written as:

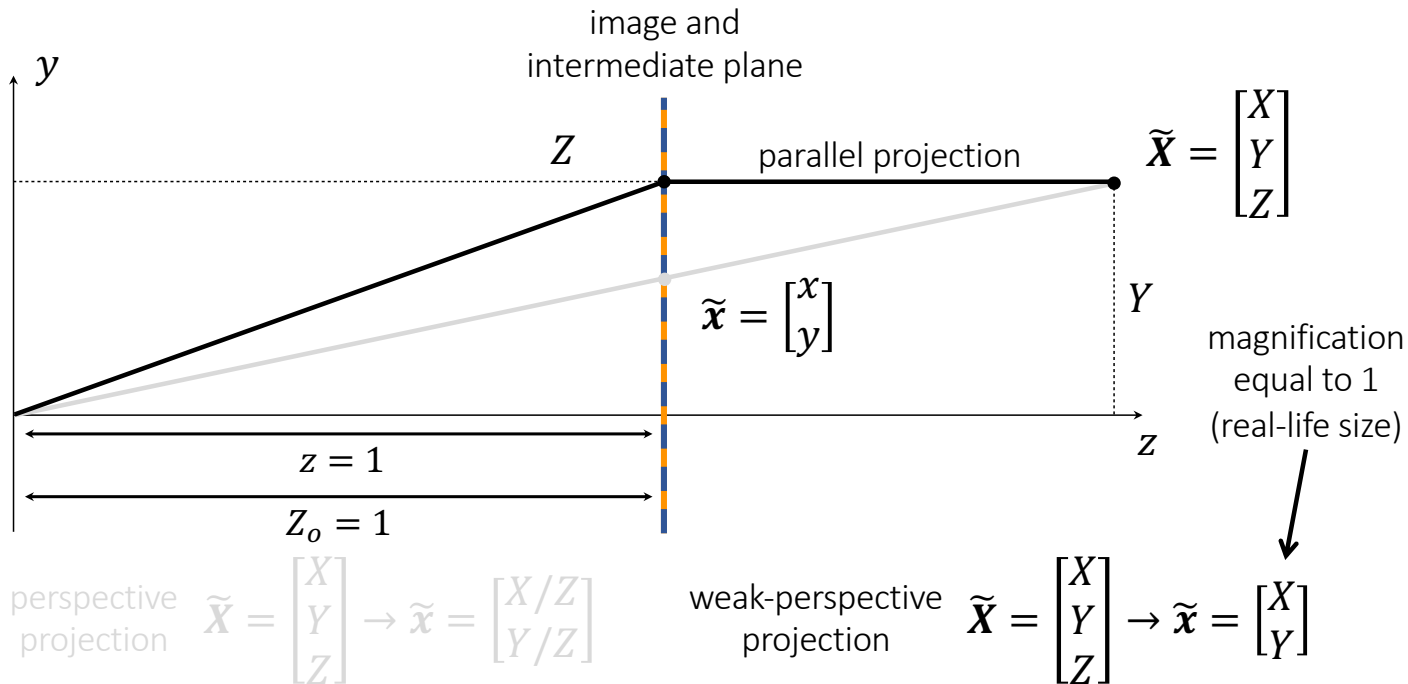
What's the effect of setting $Z_o = 1$?

$$\mathbf{P} = \mathbf{K} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix}$$

Perspective versus weak-perspective camera



Perspective versus orthographic camera



When can we assume a weak-perspective camera?

When can we assume a weak-perspective camera?

1. When the scene (or parts of it) is very far away.

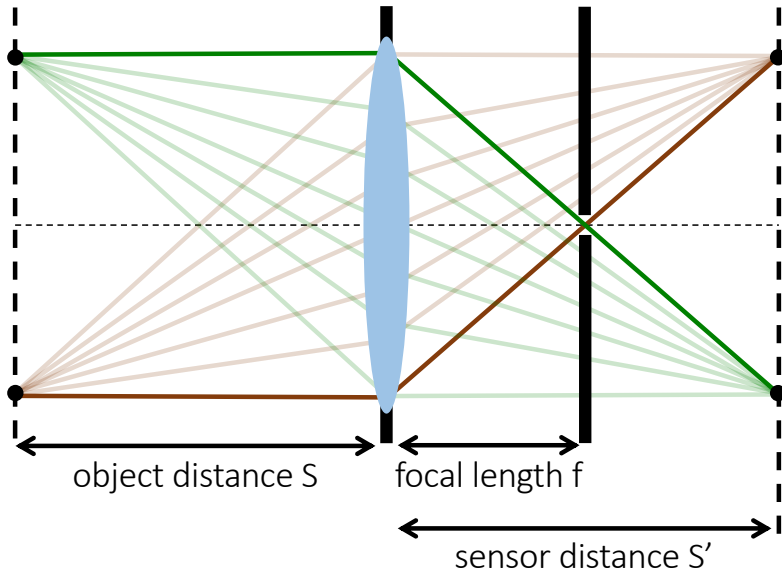


Weak-perspective projection applies to the mountains.

When can we assume a weak-perspective camera?

2. When we use a telecentric lens.

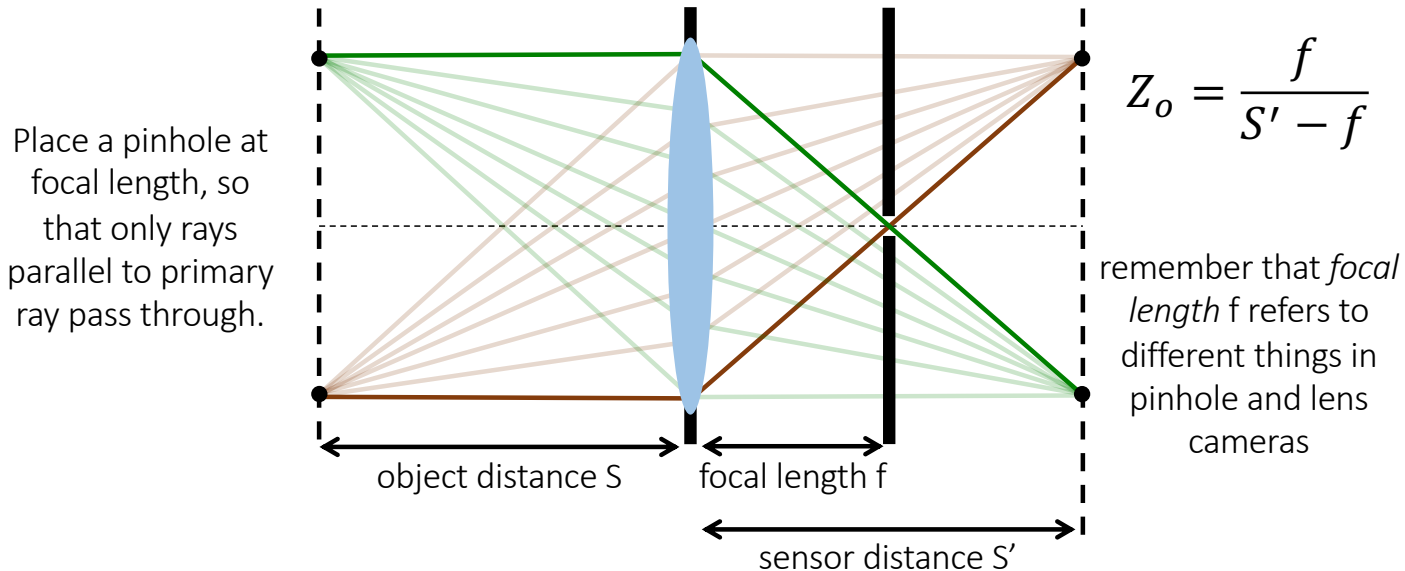
Place a pinhole at focal length, so that only rays parallel to primary ray pass through.



What does Z_o equal in this case?

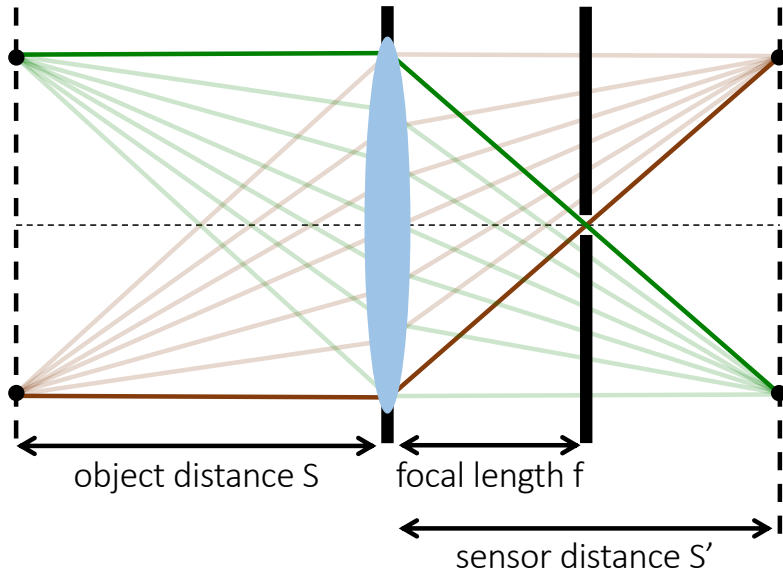
When can we assume a weak-perspective camera?

2. When we use a telecentric lens.



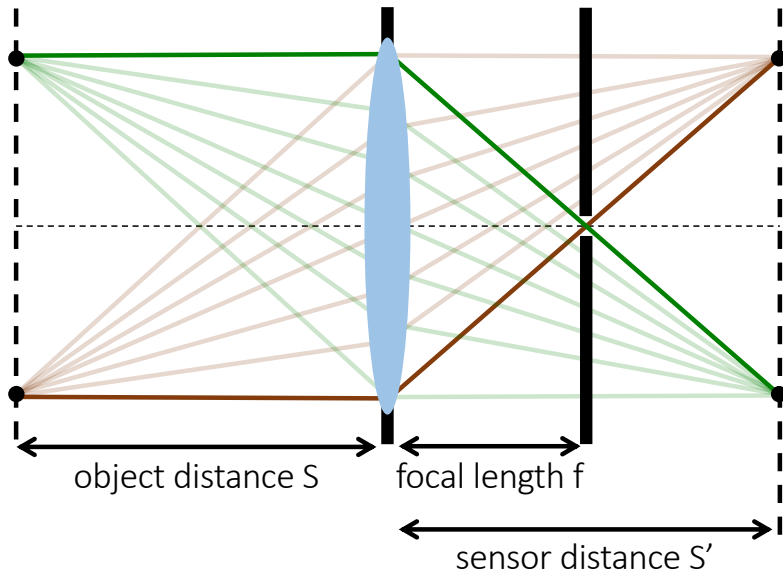
Orthographic projection using a telecentric lens

How do we make the telecentric lens act as an orthographic camera?



Orthographic projection using a telecentric lens

How do we make the telecentric lens act as an orthographic camera?

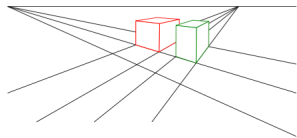


We set the sensor distance as:

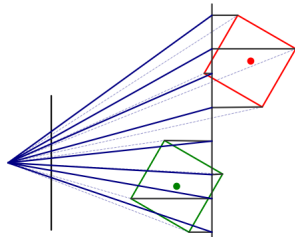
$$S' = 2f$$

in order to achieve unit magnification.

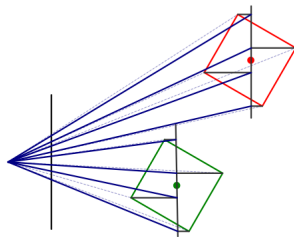
Many other types of cameras



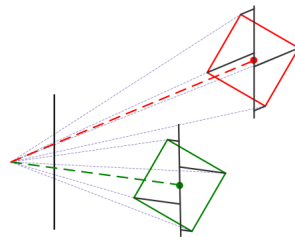
(a) 3D view



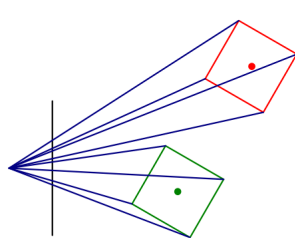
(b) orthography



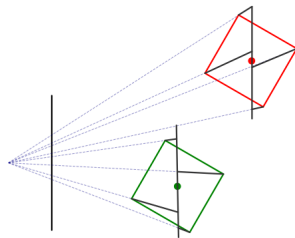
(c) scaled orthography



(d) para-perspective



(e) perspective



(f) object-centered

Geometric camera calibration

Geometric camera calibration

Given a set of matched points

$$\{\mathbf{X}_i, \mathbf{x}_i\}$$

point in 3D
space

point in the
image

and camera model

$$\mathbf{x} = \mathbf{f}(\mathbf{X}; \mathbf{p}) = \mathbf{P}\mathbf{X}$$

projection
model

parameters

Camera
matrix

Find the (pose) estimate of

P

We'll use a **perspective** camera
model for pose estimation

Mapping between 3D point and image points

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

What are the unknowns?

Mapping between 3D point and image points

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

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \text{---} & \mathbf{p}_1^\top & \text{---} \\ \text{---} & \mathbf{p}_2^\top & \text{---} \\ \text{---} & \mathbf{p}_3^\top & \text{---} \end{bmatrix} \begin{bmatrix} | \\ \mathbf{X} \\ | \end{bmatrix}$$

Heterogeneous coordinates

$$x' = \frac{\mathbf{p}_1^\top \mathbf{X}}{\mathbf{p}_3^\top \mathbf{X}} \quad y' = \frac{\mathbf{p}_2^\top \mathbf{X}}{\mathbf{p}_3^\top \mathbf{X}}$$

(non-linear relation between coordinates)

How can we make these relations linear?

How can we make these relations linear?

$$x' = \frac{\mathbf{p}_1^\top \mathbf{X}}{\mathbf{p}_3^\top \mathbf{X}} \quad y' = \frac{\mathbf{p}_2^\top \mathbf{X}}{\mathbf{p}_3^\top \mathbf{X}}$$

Make them linear with algebraic manipulation...

$$\mathbf{p}_2^\top \mathbf{X} - \mathbf{p}_3^\top \mathbf{X} y' = 0$$

$$\mathbf{p}_1^\top \mathbf{X} - \mathbf{p}_3^\top \mathbf{X} x' = 0$$

Now we can setup a system of linear equations
with multiple point correspondences

$$\mathbf{p}_2^\top \mathbf{X} - \mathbf{p}_3^\top \mathbf{X} y' = 0$$

$$\mathbf{p}_1^\top \mathbf{X} - \mathbf{p}_3^\top \mathbf{X} x' = 0$$

How do we proceed?

$$\mathbf{p}_2^\top \mathbf{X} - \mathbf{p}_3^\top \mathbf{X} y' = 0$$

$$\mathbf{p}_1^\top \mathbf{X} - \mathbf{p}_3^\top \mathbf{X} x' = 0$$

In matrix form ...

$$\begin{bmatrix} \mathbf{X}^\top & \mathbf{0} & -x' \mathbf{X}^\top \\ \mathbf{0} & \mathbf{X}^\top & -y' \mathbf{X}^\top \end{bmatrix} \begin{bmatrix} \mathbf{p}_1 \\ \mathbf{p}_2 \\ \mathbf{p}_3 \end{bmatrix} = \mathbf{0}$$

How do we proceed?

$$\mathbf{p}_2^\top \mathbf{X} - \mathbf{p}_3^\top \mathbf{X} y' = 0$$

$$\mathbf{p}_1^\top \mathbf{X} - \mathbf{p}_3^\top \mathbf{X} x' = 0$$

In matrix form ...

$$\begin{bmatrix} \mathbf{X}^\top & \mathbf{0} & -x' \mathbf{X}^\top \\ \mathbf{0} & \mathbf{X}^\top & -y' \mathbf{X}^\top \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} = \mathbf{0}$$

For N points ...

$$\begin{bmatrix} \mathbf{X}_1^\top & \mathbf{0} & -x' \mathbf{X}_1^\top \\ \mathbf{0} & \mathbf{X}_1^\top & -y' \mathbf{X}_1^\top \\ \vdots & \vdots & \vdots \\ \mathbf{X}_N^\top & \mathbf{0} & -x' \mathbf{X}_N^\top \\ \mathbf{0} & \mathbf{X}_N^\top & -y' \mathbf{X}_N^\top \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} = \mathbf{0}$$

*How do we solve
this system?*

Solve for camera matrix by

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{A}\mathbf{x}\|^2 \text{ subject to } \|\mathbf{x}\|^2 = 1$$

$$\mathbf{A} = \begin{bmatrix} \mathbf{X}_1^\top & \mathbf{0} & -x' \mathbf{X}_1^\top \\ \mathbf{0} & \mathbf{X}_1^\top & -y' \mathbf{X}_1^\top \\ \vdots & \vdots & \vdots \\ \mathbf{X}_N^\top & \mathbf{0} & -x' \mathbf{X}_N^\top \\ \mathbf{0} & \mathbf{X}_N^\top & -y' \mathbf{X}_N^\top \end{bmatrix} \quad \mathbf{x} = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix}$$

SVD!

Solve for camera matrix by

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{A}\mathbf{x}\|^2 \text{ subject to } \|\mathbf{x}\|^2 = 1$$

$$\mathbf{A} = \begin{bmatrix} \mathbf{X}_1^\top & \mathbf{0} & -x' \mathbf{X}_1^\top \\ \mathbf{0} & \mathbf{X}_1^\top & -y' \mathbf{X}_1^\top \\ \vdots & \vdots & \vdots \\ \mathbf{X}_N^\top & \mathbf{0} & -x' \mathbf{X}_N^\top \\ \mathbf{0} & \mathbf{X}_N^\top & -y' \mathbf{X}_N^\top \end{bmatrix} \quad \mathbf{x} = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix}$$

Solution $\mathbf{x}$ is the column of $\mathbf{V}$
corresponding to smallest singular
value of

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\top$$

Solve for camera matrix by

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{A}\mathbf{x}\|^2 \text{ subject to } \|\mathbf{x}\|^2 = 1$$

$$\mathbf{A} = \begin{bmatrix} \mathbf{X}_1^\top & \mathbf{0} & -x' \mathbf{X}_1^\top \\ \mathbf{0} & \mathbf{X}_1^\top & -y' \mathbf{X}_1^\top \\ \vdots & \vdots & \vdots \\ \mathbf{X}_N^\top & \mathbf{0} & -x' \mathbf{X}_N^\top \\ \mathbf{0} & \mathbf{X}_N^\top & -y' \mathbf{X}_N^\top \end{bmatrix} \quad \mathbf{x} = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix}$$

Equivalently, solution $\mathbf{x}$ is the Eigenvector corresponding to smallest Eigenvalue of

$$\mathbf{A}^\top \mathbf{A}$$

Now we have:

$$\mathbf{P} = \begin{bmatrix} p_1 & p_2 & p_3 & p_4 \\ p_5 & p_6 & p_7 & p_8 \\ p_9 & p_{10} & p_{11} & p_{12} \end{bmatrix}$$

Are we done?

Almost there ...

$$\mathbf{P} = \begin{bmatrix} p_1 & p_2 & p_3 & p_4 \\ p_5 & p_6 & p_7 & p_8 \\ p_9 & p_{10} & p_{11} & p_{12} \end{bmatrix}$$

How do you get the intrinsic and extrinsic parameters from the projection matrix?

Decomposition of the Camera Matrix

$$\mathbf{P} = \left[\begin{array}{ccc|c} p_1 & p_2 & p_3 & p_4 \\ p_5 & p_6 & p_7 & p_8 \\ p_9 & p_{10} & p_{11} & p_{12} \end{array} \right]$$

Decomposition of the Camera Matrix

$$\mathbf{P} = \left[\begin{array}{ccc|c} p_1 & p_2 & p_3 & p_4 \\ p_5 & p_6 & p_7 & p_8 \\ p_9 & p_{10} & p_{11} & p_{12} \end{array} \right]$$

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

Decomposition of the Camera Matrix

$$\mathbf{P} = \left[\begin{array}{ccc|c} p_1 & p_2 & p_3 & p_4 \\ p_5 & p_6 & p_7 & p_8 \\ p_9 & p_{10} & p_{11} & p_{12} \end{array} \right]$$

$$\begin{aligned} \mathbf{P} &= \mathbf{K}[\mathbf{R}|\mathbf{t}] \\ &= \mathbf{K}[\mathbf{R}|-\mathbf{R}\mathbf{c}] \\ &= [\mathbf{M}|-\mathbf{M}\mathbf{c}] \end{aligned}$$

Decomposition of the Camera Matrix

$$\mathbf{P} = \left[\begin{array}{ccc|c} p_1 & p_2 & p_3 & p_4 \\ p_5 & p_6 & p_7 & p_8 \\ p_9 & p_{10} & p_{11} & p_{12} \end{array} \right]$$

$$\begin{aligned}\mathbf{P} &= \mathbf{K}[\mathbf{R}|\mathbf{t}] \\ &= \mathbf{K}[\mathbf{R} | -\mathbf{R}\mathbf{c}] \\ &= [\mathbf{M} | -\mathbf{M}\mathbf{c}]\end{aligned}$$

Find the camera center $\mathbf{C}$

What is the projection of the camera center?

Find intrinsic $\mathbf{K}$ and rotation $\mathbf{R}$

Decomposition of the Camera Matrix

$$\mathbf{P} = \left[\begin{array}{ccc|c} p_1 & p_2 & p_3 & p_4 \\ p_5 & p_6 & p_7 & p_8 \\ p_9 & p_{10} & p_{11} & p_{12} \end{array} \right]$$

$$\begin{aligned}\mathbf{P} &= \mathbf{K}[\mathbf{R}|\mathbf{t}] \\ &= \mathbf{K}[\mathbf{R} | -\mathbf{R}\mathbf{c}] \\ &= [\mathbf{M} | -\mathbf{M}\mathbf{c}]\end{aligned}$$

Find the camera center $\mathbf{c}$

$$\mathbf{P}\mathbf{c} = \mathbf{0}$$

How do we compute the camera center from this?

Find intrinsic $\mathbf{K}$ and rotation $\mathbf{R}$

Decomposition of the Camera Matrix

$$\mathbf{P} = \left[\begin{array}{ccc|c} p_1 & p_2 & p_3 & p_4 \\ p_5 & p_6 & p_7 & p_8 \\ p_9 & p_{10} & p_{11} & p_{12} \end{array} \right]$$

$$\begin{aligned} \mathbf{P} &= \mathbf{K}[\mathbf{R}|\mathbf{t}] \\ &= \mathbf{K}[\mathbf{R} | -\mathbf{R}\mathbf{c}] \\ &= [\mathbf{M} | -\mathbf{M}\mathbf{c}] \end{aligned}$$

Find the camera center $\mathbf{c}$

$$\mathbf{P}\mathbf{c} = \mathbf{0}$$

SVD of $\mathbf{P}$!

*$\mathbf{c}$ is the Eigenvector corresponding to
smallest Eigenvalue*

Find intrinsic $\mathbf{K}$ and rotation $\mathbf{R}$

Decomposition of the Camera Matrix

$$\mathbf{P} = \left[\begin{array}{ccc|c} p_1 & p_2 & p_3 & p_4 \\ p_5 & p_6 & p_7 & p_8 \\ p_9 & p_{10} & p_{11} & p_{12} \end{array} \right]$$

$$\begin{aligned} \mathbf{P} &= \mathbf{K}[\mathbf{R}|\mathbf{t}] \\ &= \mathbf{K}[\mathbf{R} | -\mathbf{R}\mathbf{c}] \\ &= [\mathbf{M} | -\mathbf{M}\mathbf{c}] \end{aligned}$$

Find the camera center $\mathbf{c}$

$$\mathbf{P}\mathbf{c} = \mathbf{0}$$

SVD of $\mathbf{P}$!

$\mathbf{c}$ is the Eigenvector corresponding to
smallest Eigenvalue

Find intrinsic $\mathbf{K}$ and rotation $\mathbf{R}$

$$\mathbf{M} = \mathbf{K}\mathbf{R}$$

*Any useful properties of $\mathbf{K}$
and $\mathbf{R}$ we can use?*

Decomposition of the Camera Matrix

$$\mathbf{P} = \left[\begin{array}{ccc|c} p_1 & p_2 & p_3 & p_4 \\ p_5 & p_6 & p_7 & p_8 \\ p_9 & p_{10} & p_{11} & p_{12} \end{array} \right]$$

$$\begin{aligned} \mathbf{P} &= \mathbf{K}[\mathbf{R}|\mathbf{t}] \\ &= \mathbf{K}[\mathbf{R} | -\mathbf{R}\mathbf{c}] \\ &= [\mathbf{M} | -\mathbf{M}\mathbf{c}] \end{aligned}$$

Find the camera center $\mathbf{c}$

$$\mathbf{P}\mathbf{c} = \mathbf{0}$$

SVD of $\mathbf{P}$!

$\mathbf{c}$ is the Eigenvector corresponding to
smallest Eigenvalue

Find intrinsic $\mathbf{K}$ and rotation $\mathbf{R}$

$$\mathbf{M} = \mathbf{K}\mathbf{R}$$

right upper
triangle orthogonal

*How do we find $\mathbf{K}$
and $\mathbf{R}$?*

Decomposition of the Camera Matrix

$$\mathbf{P} = \left[\begin{array}{ccc|c} p_1 & p_2 & p_3 & p_4 \\ p_5 & p_6 & p_7 & p_8 \\ p_9 & p_{10} & p_{11} & p_{12} \end{array} \right]$$

$$\begin{aligned} \mathbf{P} &= \mathbf{K}[\mathbf{R}|\mathbf{t}] \\ &= \mathbf{K}[\mathbf{R} | -\mathbf{R}\mathbf{c}] \\ &= [\mathbf{M} | -\mathbf{M}\mathbf{c}] \end{aligned}$$

Find the camera center $\mathbf{c}$

$$\mathbf{P}\mathbf{c} = \mathbf{0}$$

SVD of $\mathbf{P}$!

*$\mathbf{c}$ is the Eigenvector corresponding to
smallest Eigenvalue*

Find intrinsic $\mathbf{K}$ and rotation $\mathbf{R}$

$$\mathbf{M} = \mathbf{K}\mathbf{R}$$

QR decomposition

Geometric camera calibration

Given a set of matched points

$$\{\mathbf{X}_i, \mathbf{x}_i\}$$

point in 3D
space

point in the
image

*Where do we get these
matched points from?*

and camera model

$$\mathbf{x} = \mathbf{f}(\mathbf{X}; \mathbf{p}) = \mathbf{P}\mathbf{X}$$

projection
model

parameters

Camera
matrix

Find the (pose) estimate of

P

We'll use a **perspective** camera
model for pose estimation

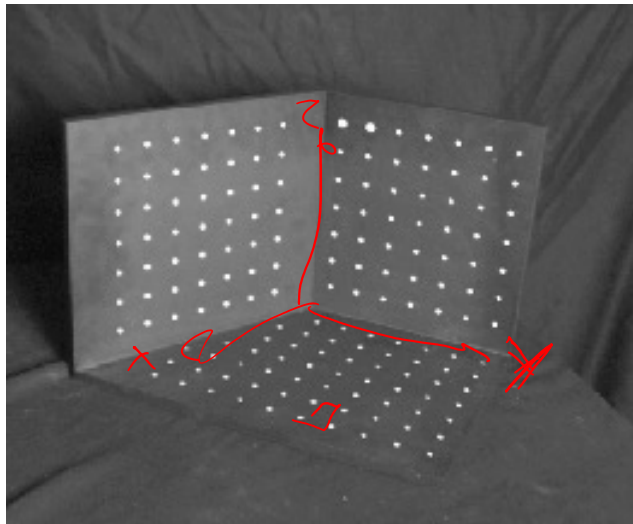
Calibration using a reference object

Place a known object in the scene:

- identify correspondences between image and scene
- compute mapping from scene to image

Issues:

- must know geometry very accurately
- must know 3D->2D correspondence



Geometric camera calibration

Advantages:

- Very simple to formulate.
- Analytical solution.

Disadvantages:

- Doesn't model radial distortion.
- Hard to impose constraints (e.g., known f).
- Doesn't minimize the correct error function.

Geometric camera calibration

Advantages:

- Very simple to formulate.
- Analytical solution.

Disadvantages:

- Doesn't model radial distortion.
- Hard to impose constraints (e.g., known f).
- Doesn't minimize the correct error function. why?

Geometric camera calibration

Advantages:

- Very simple to formulate.
- Analytical solution.

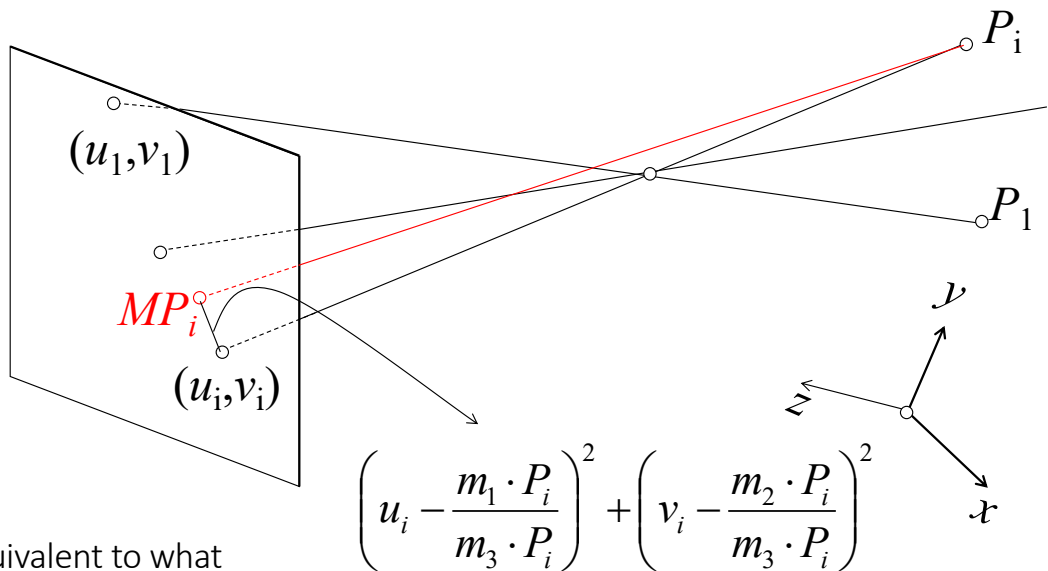
Disadvantages:

- Doesn't model radial distortion.
- Hard to impose constraints (e.g., known f).
- Doesn't minimize the correct error function.

For these reasons, *nonlinear methods* are preferred

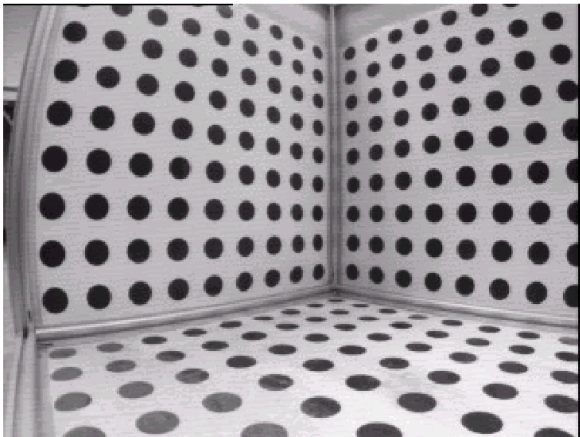
- Define error function E between projected 3D points and image positions
 - E is nonlinear function of intrinsics, extrinsics, radial distortion
- Minimize E using nonlinear optimization techniques

Minimizing reprojection error

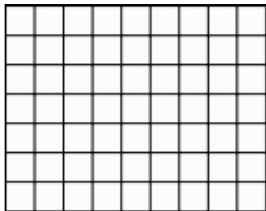


Is this equivalent to what we were doing previously?

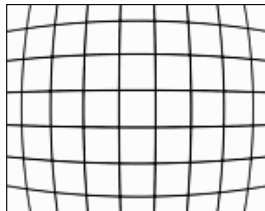
Radial distortion



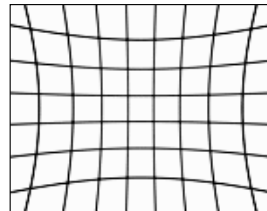
What causes this distortion?



no distortion

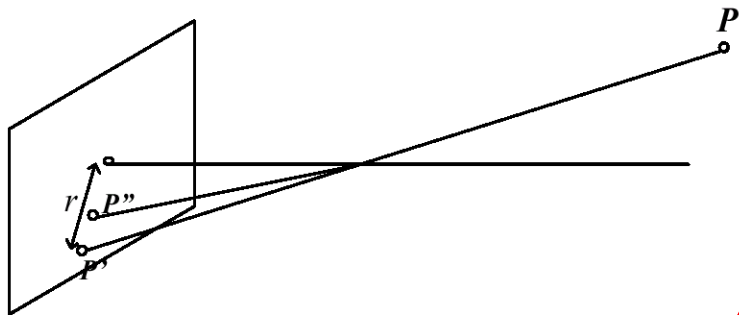


barrel distortion



pincushion distortion

Radial distortion model



Ideal:

$$x' = f \frac{x}{z}$$

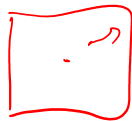
$$y' = f \frac{y}{z}$$

Distorted:

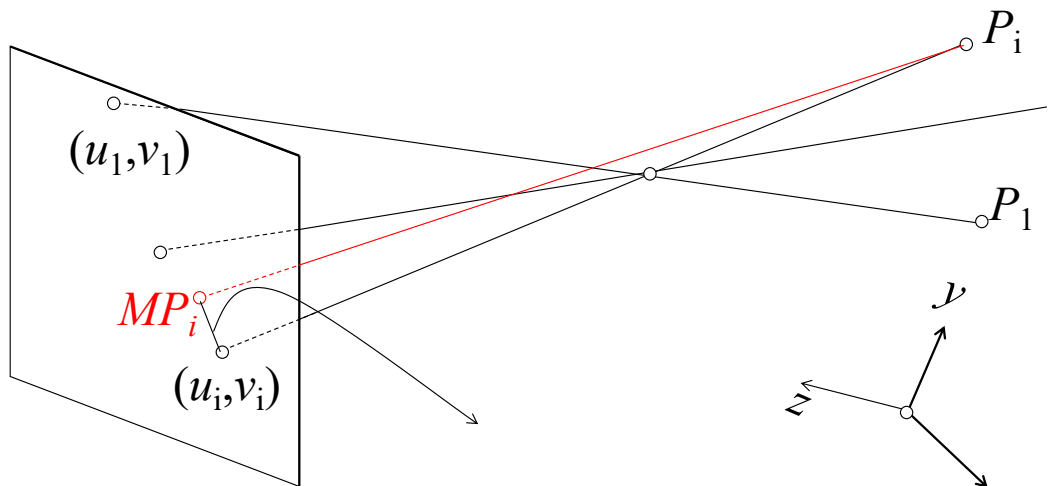
$$x'' = \frac{1}{\lambda} x'$$

$$y'' = \frac{1}{\lambda} y'$$

$$\lambda = 1 + k_1 \overbrace{r^2}^{\sqrt{x'^2 + y'^2}} + k_2 r^4 + \dots$$



Minimizing reprojection error with radial distortion



Add distortions to
reprojection error:

$$\left(u_i - \frac{1}{\lambda} \frac{m_1 \cdot P_i}{m_3 \cdot P_i}\right)^2 + \left(v_i - \frac{1}{\lambda} \frac{m_2 \cdot P_i}{m_3 \cdot P_i}\right)^2$$

Correcting radial distortion

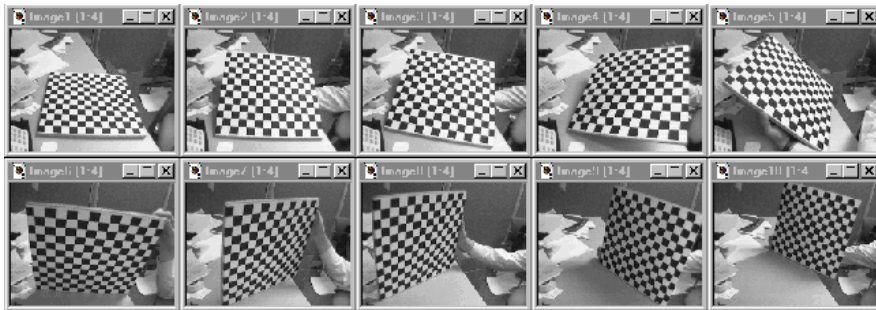


before



after

Alternative: Multi-plane calibration



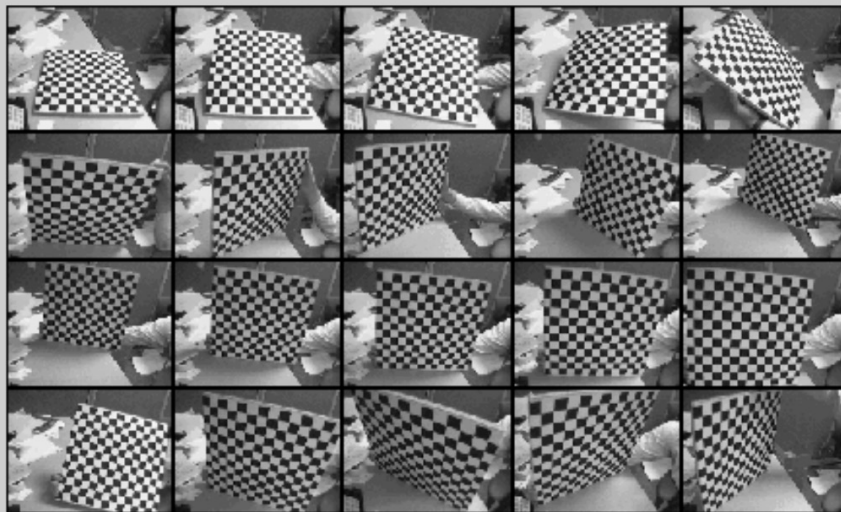
Advantages:

- Only requires a plane
- Don't have to know positions/orientations
- Great code available online!
 - Matlab version: http://www.vision.caltech.edu/bouguetj/calib_doc/index.html
 - Also available on OpenCV.

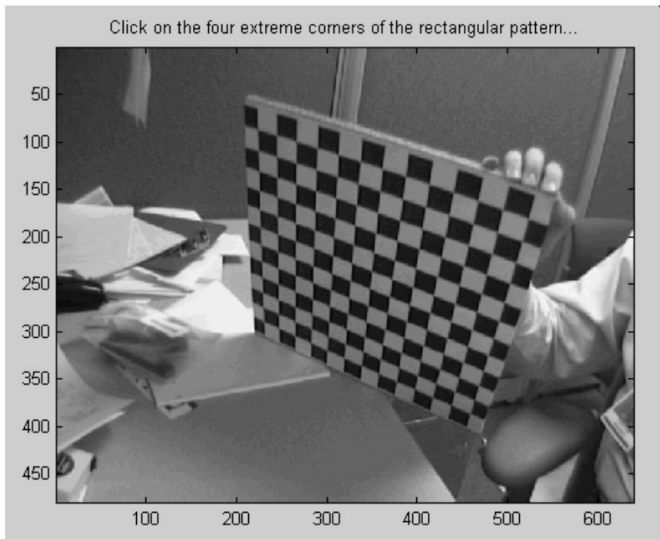
Disadvantage: Need to solve non-linear optimization problem.

Step-by-step demonstration

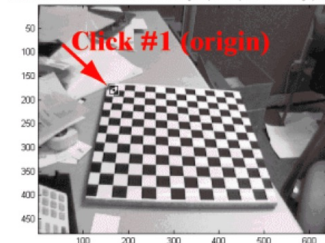
Calibration images



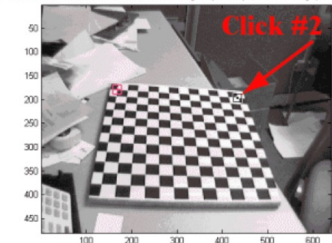
Step-by-step demonstration



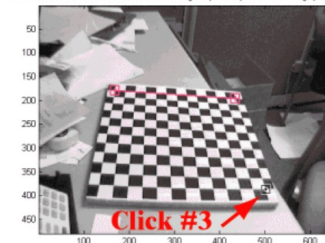
Click on the four extreme corners of the rectangular pattern (first corner = origin) ... Image 1



Click on the four extreme corners of the rectangular pattern (first corner = origin) ... Image 1



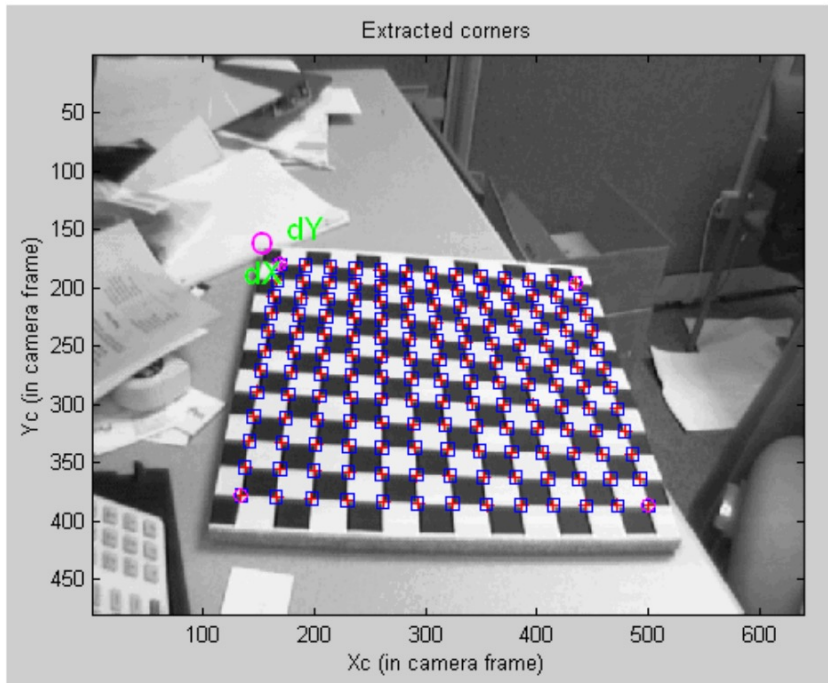
Click on the four extreme corners of the rectangular pattern (first corner = origin) ... Image 1



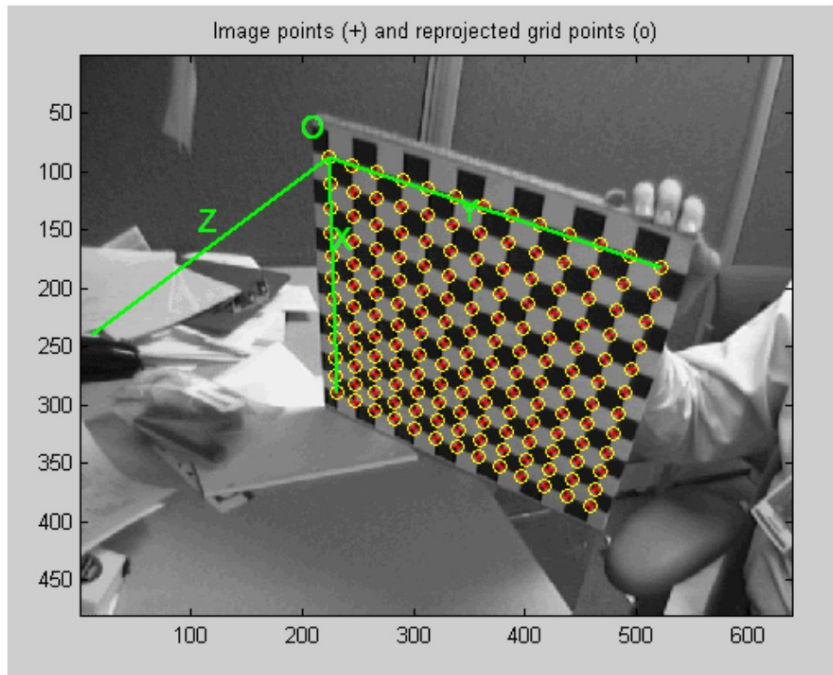
Click on the four extreme corners of the rectangular pattern (first corner = origin) ... Image 1



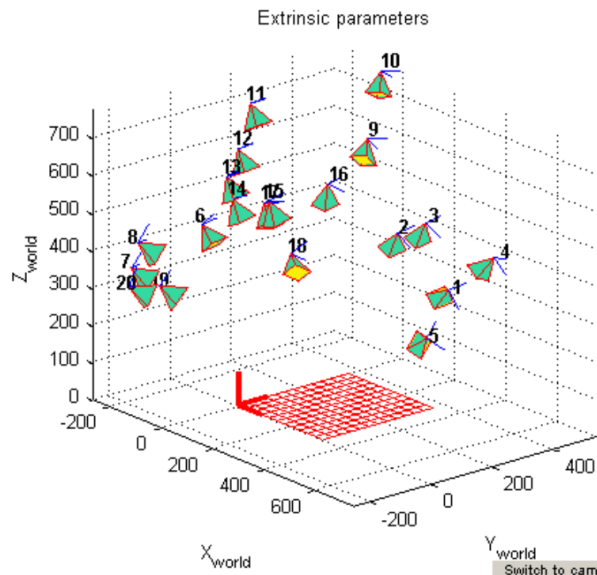
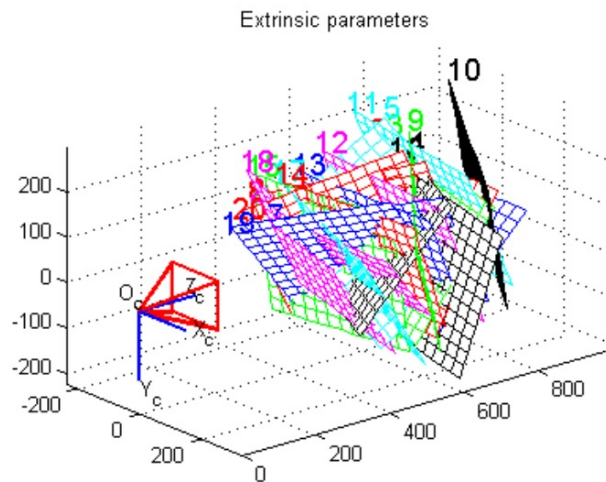
Step-by-step demonstration



Step-by-step demonstration



Step-by-step demonstration



Switch to camera-centered view

What does it mean to “calibrate a camera”?

What does it mean to “calibrate a camera”?

Many different ways to calibrate a camera:

- Radiometric calibration. ← lecture 5-ish
- Color calibration. ← lecture 7-ish
- Geometric calibration. ← lecture 19 (this lecture)
- Noise calibration. ← lecture 6-ish
- Lens (or aberration) calibration. ← lecture 12-ish, (maybe) later lecture

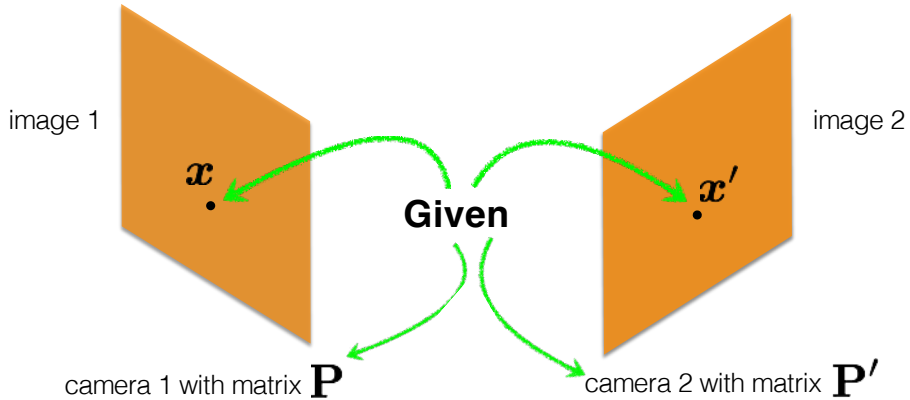
Two-view geometry

Overview

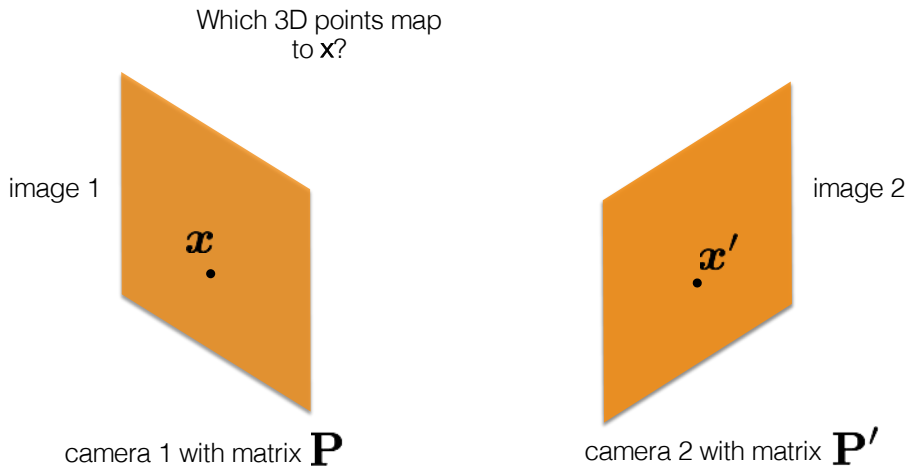
- Triangulation.
- Epipolar geometry.
- Next time:
 - Essential matrix.
 - Fundamental matrix.
 - 8-point algorithm.

Triangulation

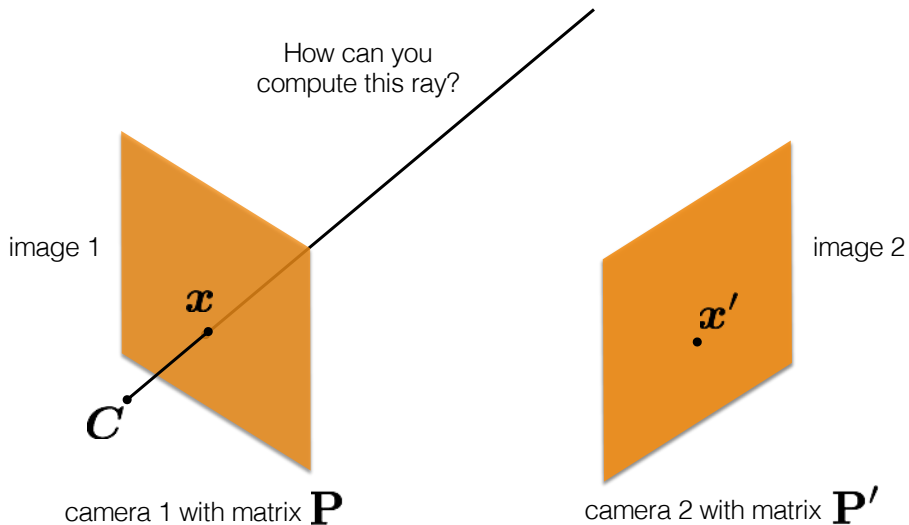
Triangulation



Triangulation



Triangulation



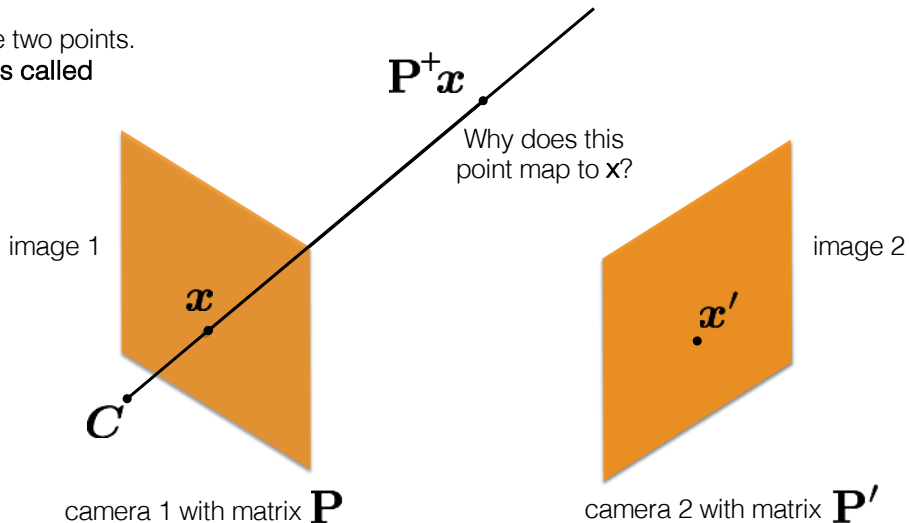
Triangulation

Create two points on the ray:

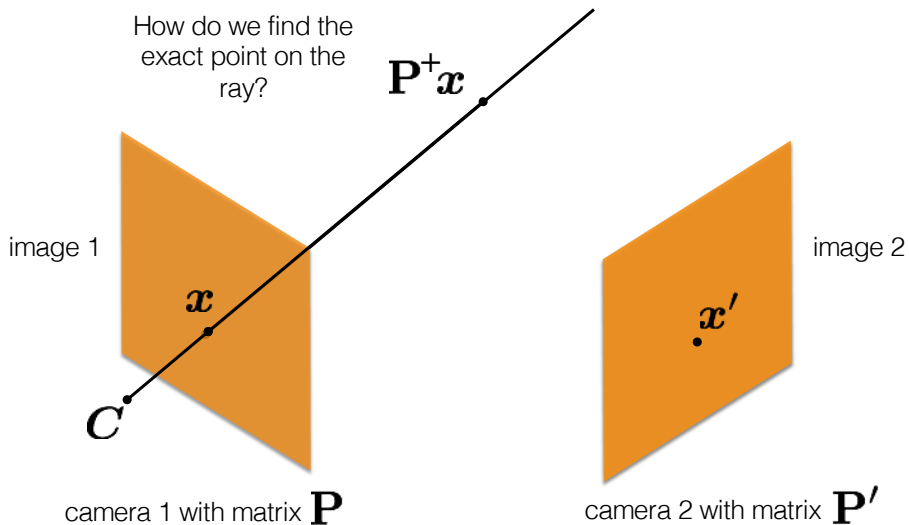
- 1) find the camera center; and
- 2) apply the pseudo-inverse of $\mathbf{P}$ on $\mathbf{x}$.

Then connect the two points.

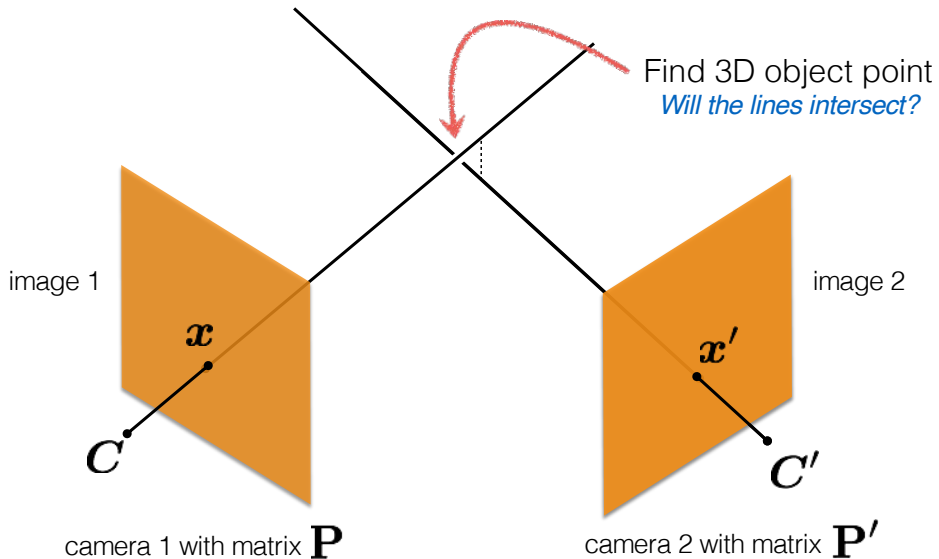
**This procedure is called
backprojection**



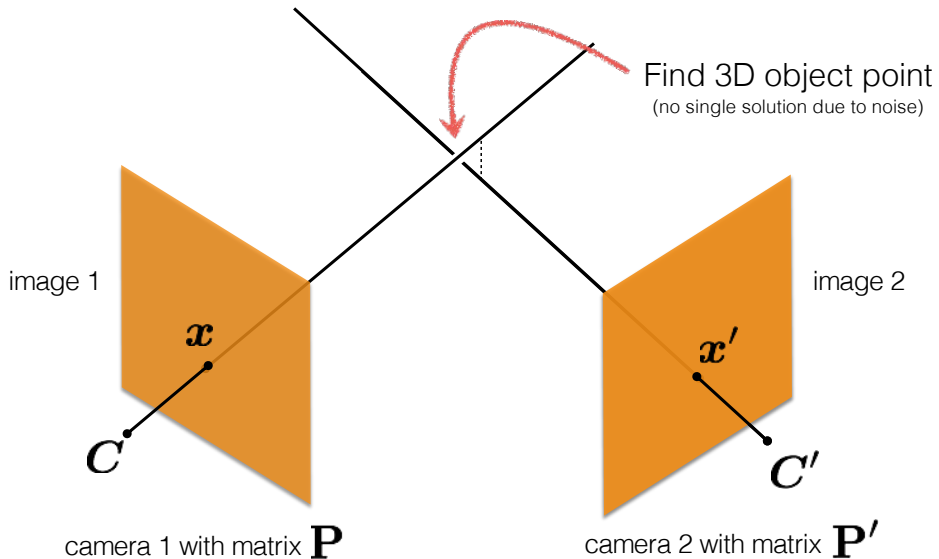
Triangulation



Triangulation



Triangulation



Triangulation

Given a set of (noisy) matched points

$$\{\mathbf{x}_i, \mathbf{x}'_i\}$$

and camera matrices

$$\mathbf{P}, \mathbf{P}'$$

Estimate the 3D point

$$\mathbf{X}$$

$$\underset{\text{known}}{\mathbf{x}} = \underset{\text{known}}{\mathbf{P}} \mathbf{X}$$

*Can we compute $\mathbf{X}$ from a single
correspondence $\mathbf{x}$?*

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

(homogeneous
coordinate)

This is a similarity relation because it involves homogeneous coordinates

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

(homogeneous
coordinate)

Same ray direction but differs by a scale factor

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

How do we solve for unknowns in a similarity relation?

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

(homogeneous
coordinate)

Also, this is a similarity relation because it involves homogeneous coordinates

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

(inhomogeneous
coordinate)

Same ray direction but differs by a scale factor

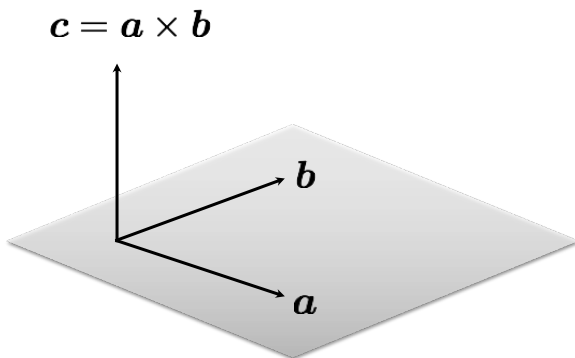
$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \alpha \begin{bmatrix} p_1 & p_2 & p_3 & p_4 \\ p_5 & p_6 & p_7 & p_8 \\ p_9 & p_{10} & p_{11} & p_{12} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

How do we solve for unknowns in a similarity relation?

Linear algebra reminder: cross product

Vector (cross) product

takes two vectors and returns a vector perpendicular to both



$$\mathbf{c} \cdot \mathbf{a} = 0$$

$$\mathbf{c} \cdot \mathbf{b} = 0$$

$$\mathbf{a} \times \mathbf{b} = \begin{bmatrix} a_2b_3 - a_3b_2 \\ a_3b_1 - a_1b_3 \\ a_1b_2 - a_2b_1 \end{bmatrix}$$

cross product of two vectors in
the same direction is zero
vector

$$\mathbf{a} \times \mathbf{a} = \mathbf{0}$$

remember this!!!

Linear algebra reminder: cross product

Cross product

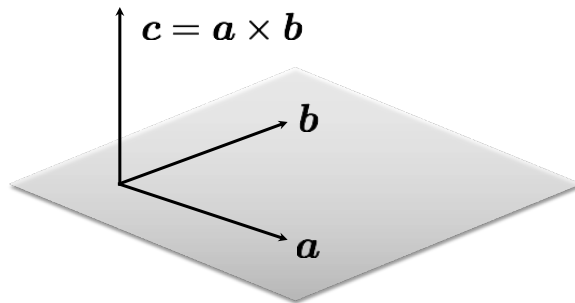
$$\mathbf{a} \times \mathbf{b} = \begin{bmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{bmatrix}$$

Can also be written as a matrix multiplication

$$\mathbf{a} \times \mathbf{b} = [\mathbf{a}]_{\times} \mathbf{b} = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

Skew symmetric

Compare with: dot product



$$c \cdot a = 0$$

$$c \cdot b = 0$$

dot product of two orthogonal vectors is (scalar) zero

Back to triangulation

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

Same direction but differs by a scale factor

How can we rewrite this using vector products?

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

Same direction but differs by a scale factor

$$\mathbf{x} \times \mathbf{P} \mathbf{X} = \mathbf{0}$$

Cross product of two vectors of same direction is zero
(this equality removes the scale factor)

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

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \alpha \begin{bmatrix} \text{---} & \mathbf{p}_1^\top & \text{---} \\ \text{---} & \mathbf{p}_2^\top & \text{---} \\ \text{---} & \mathbf{p}_3^\top & \text{---} \end{bmatrix} \begin{bmatrix} | \\ \mathbf{X} \\ | \end{bmatrix}$$

Do the same after first
expanding out the
camera matrix and points

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \alpha \begin{bmatrix} \mathbf{p}_1^\top \mathbf{X} \\ \mathbf{p}_2^\top \mathbf{X} \\ \mathbf{p}_3^\top \mathbf{X} \end{bmatrix}$$

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \times \begin{bmatrix} \mathbf{p}_1^\top \mathbf{X} \\ \mathbf{p}_2^\top \mathbf{X} \\ \mathbf{p}_3^\top \mathbf{X} \end{bmatrix} = \begin{bmatrix} y\mathbf{p}_3^\top \mathbf{X} - \mathbf{p}_2^\top \mathbf{X} \\ \mathbf{p}_1^\top \mathbf{X} - x\mathbf{p}_3^\top \mathbf{X} \\ x\mathbf{p}_2^\top \mathbf{X} - y\mathbf{p}_1^\top \mathbf{X} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Using the fact that the cross product should be zero

$$\mathbf{x} \times \mathbf{P}\mathbf{X} = \mathbf{0}$$

$$\begin{bmatrix} yp_3^\top \mathbf{X} - p_2^\top \mathbf{X} \\ p_1^\top \mathbf{X} - xp_3^\top \mathbf{X} \\ xp_2^\top \mathbf{X} - yp_1^\top \mathbf{X} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Third line is a linear combination of the first and second lines.
(x times the first line plus y times the second line)

One 2D to 3D point correspondence give you  equations

Using the fact that the cross product should be zero

$$\mathbf{x} \times \mathbf{P}\mathbf{X} = \mathbf{0}$$

$$\begin{bmatrix} yp_3^\top \mathbf{X} - p_2^\top \mathbf{X} \\ p_1^\top \mathbf{X} - xp_3^\top \mathbf{X} \\ xp_2^\top \mathbf{X} - yp_1^\top \mathbf{X} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Third line is a linear combination of the first and second lines.
(x times the first line plus y times the second line)

One 2D to 3D point correspondence give you 2 equations

$$\begin{bmatrix} yp_3^\top \mathbf{X} - p_2^\top \mathbf{X} \\ p_1^\top \mathbf{X} - xp_3^\top \mathbf{X} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Remove third row, and
rearrange as system on
unknowns

$$\begin{bmatrix} yp_3^\top - p_2^\top \\ p_1^\top - xp_3^\top \end{bmatrix} \mathbf{X} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\mathbf{A}_i \mathbf{X} = \mathbf{0}$$

Now we can make a system of linear equations
(two lines for each 2D point correspondence)

Concatenate the 2D points from both images

Two rows from camera
one

Two rows from camera
two

$$\begin{bmatrix} y\mathbf{p}_3^\top - \mathbf{p}_2^\top \\ \mathbf{p}_1^\top - x\mathbf{p}_3^\top \\ y'\mathbf{p}_3'^\top - \mathbf{p}_2'^\top \\ \mathbf{p}_1'^\top - x'\mathbf{p}_3'^\top \end{bmatrix} \mathbf{X} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

sanity check! dimensions?

$$\mathbf{A}\mathbf{X} = \mathbf{0}$$

How do we solve homogeneous linear system?

Concatenate the 2D points from both images

$$\begin{bmatrix} y\mathbf{p}_3^\top - \mathbf{p}_2^\top \\ \mathbf{p}_1^\top - x\mathbf{p}_3^\top \\ y'\mathbf{p}_3'^\top - \mathbf{p}_2'^\top \\ \mathbf{p}_1'^\top - x'\mathbf{p}_3'^\top \end{bmatrix} \mathbf{X} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

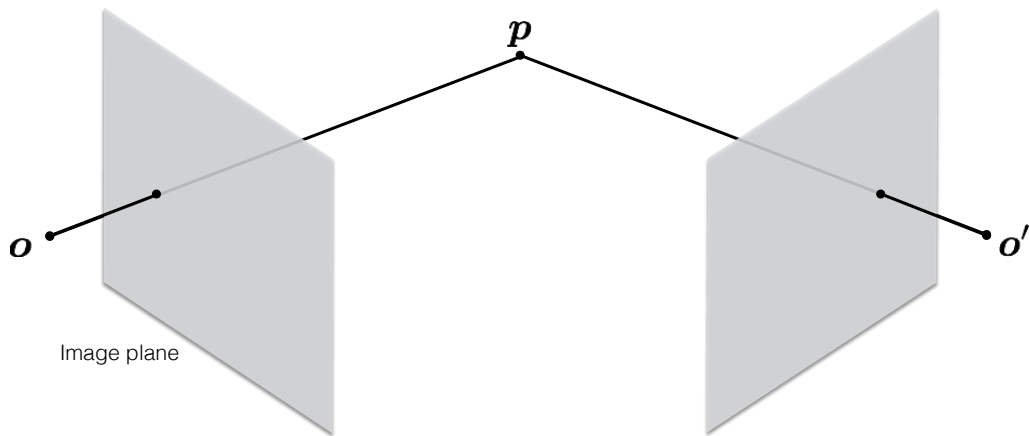
$$\mathbf{A}\mathbf{X} = \mathbf{0}$$

How do we solve homogeneous linear system?

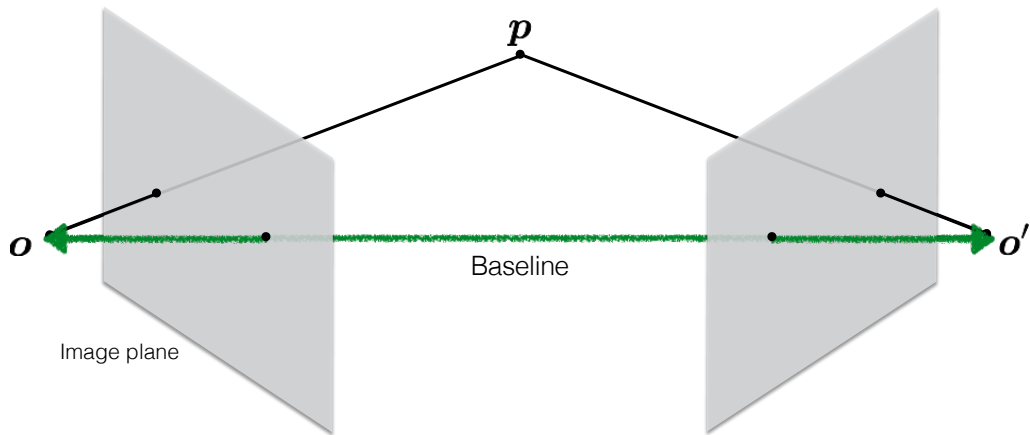
S V D !

Epipolar geometry

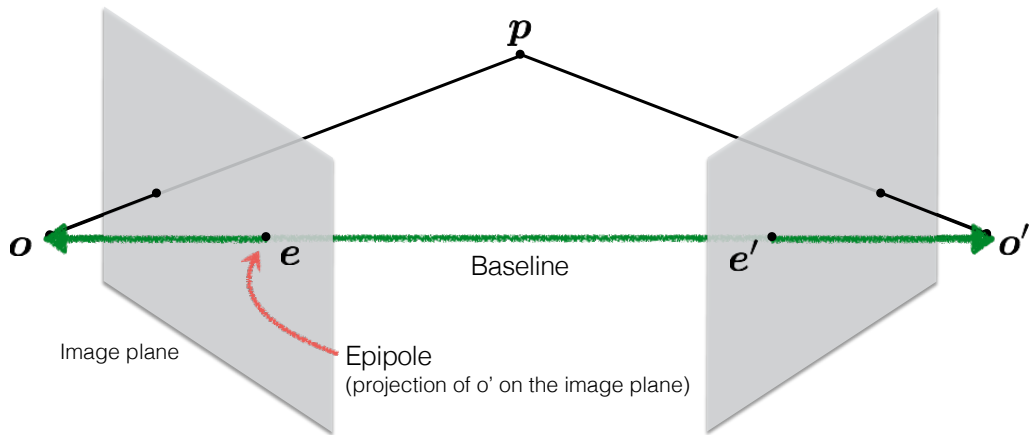
Epipolar geometry



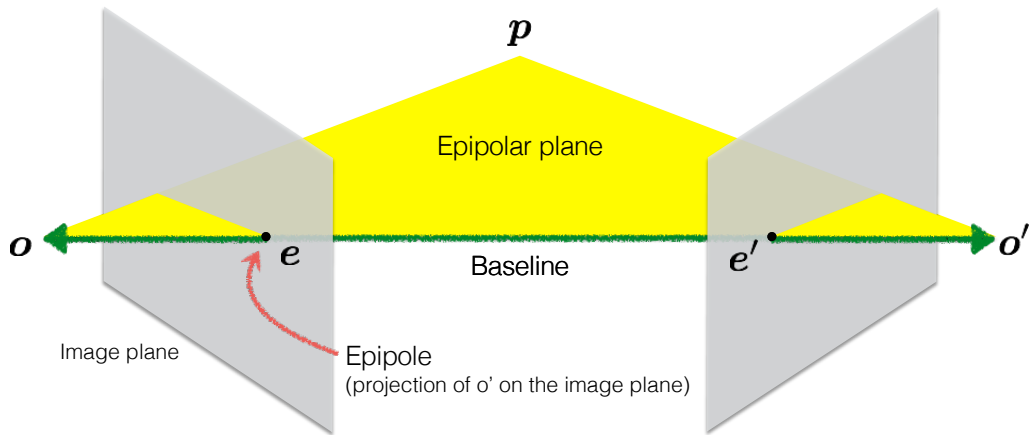
Epipolar geometry



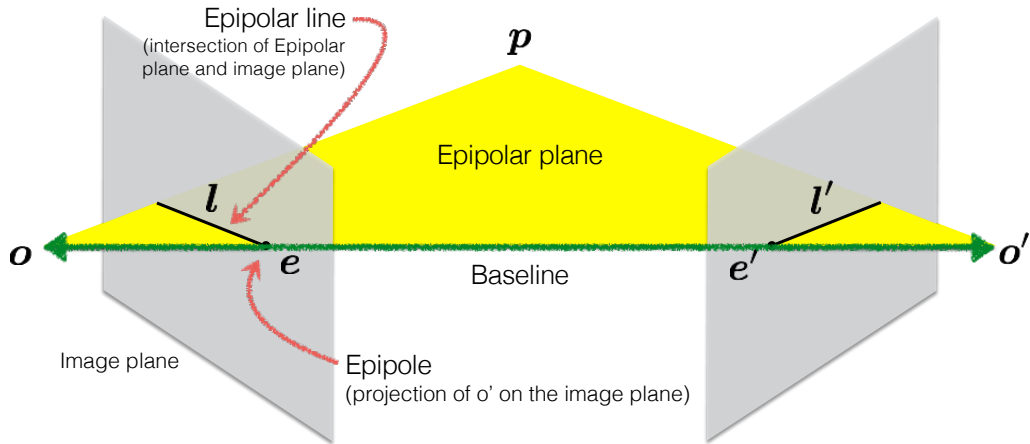
Epipolar geometry



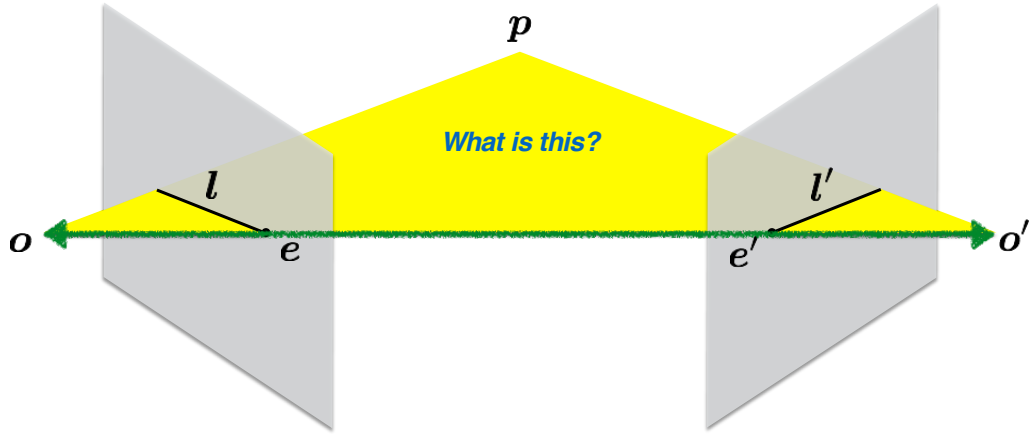
Epipolar geometry



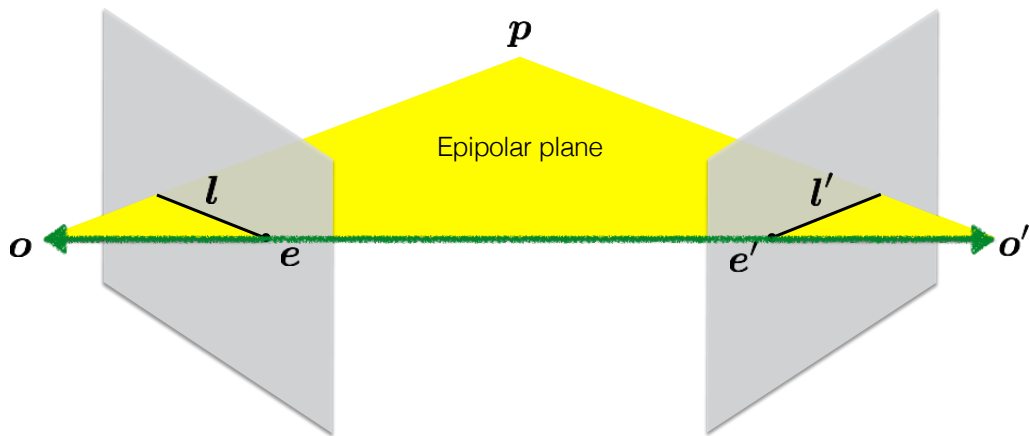
Epipolar geometry



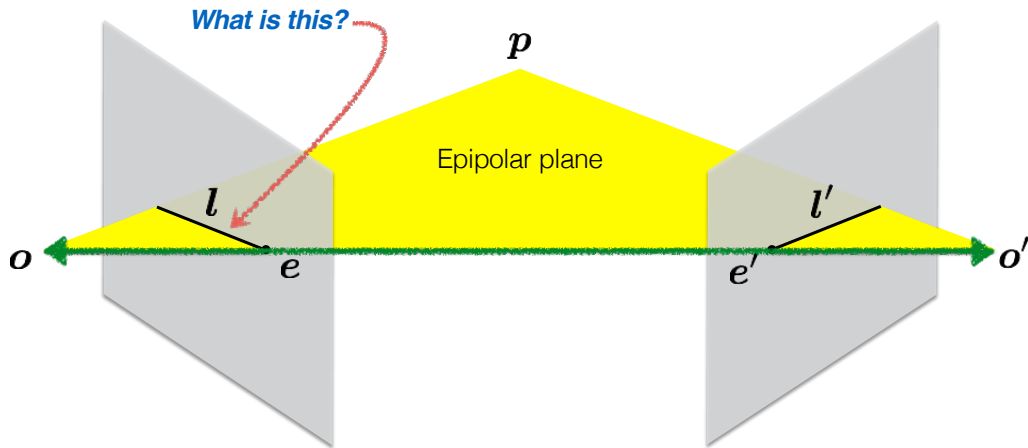
Quiz



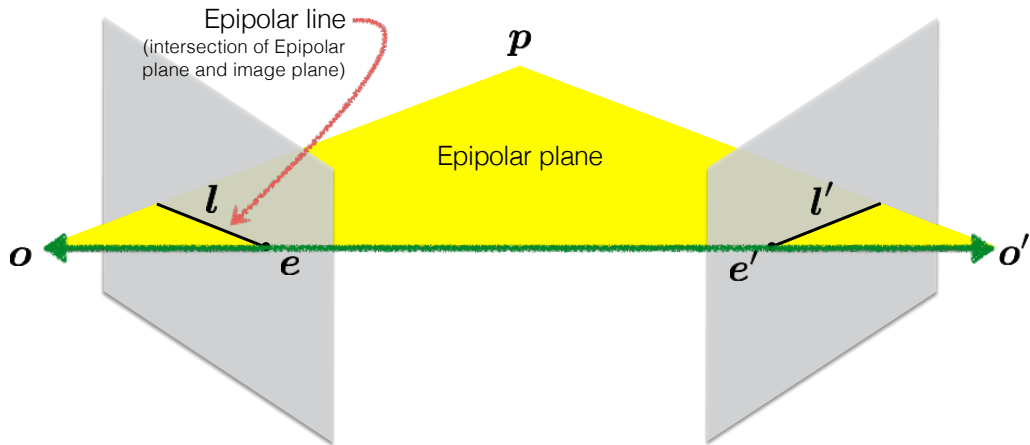
Quiz



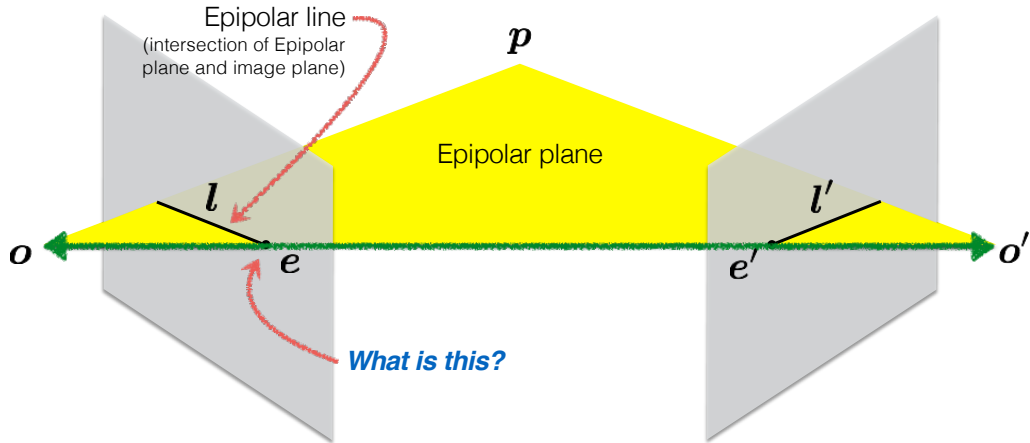
Quiz



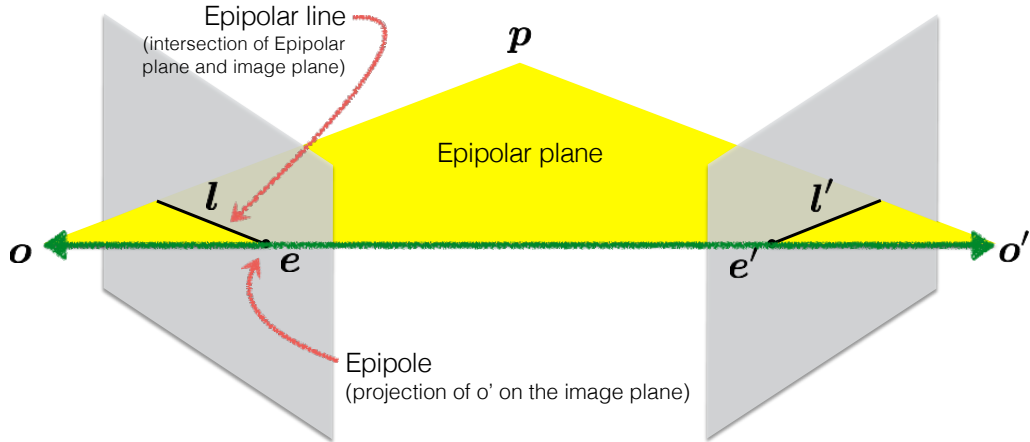
Quiz



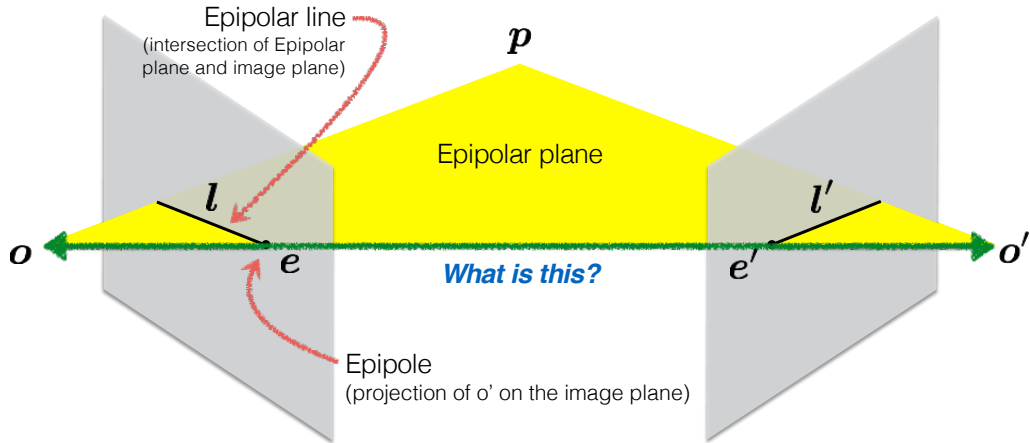
Quiz



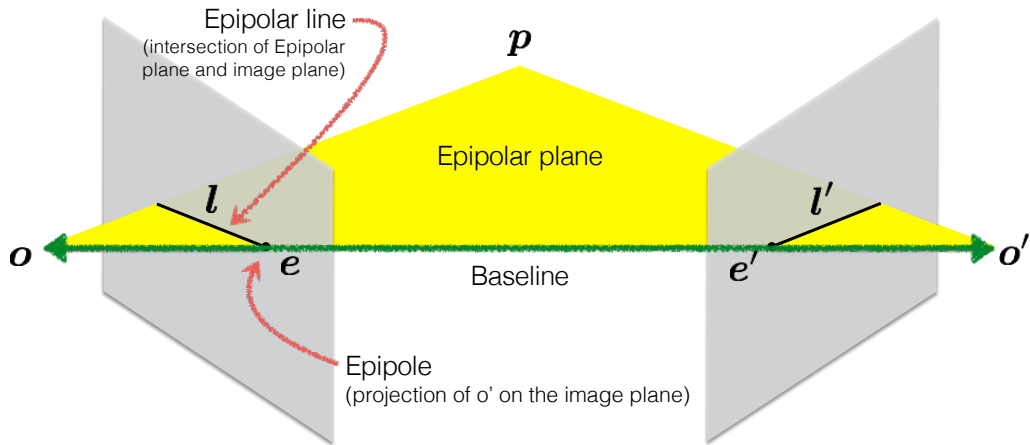
Quiz



Quiz

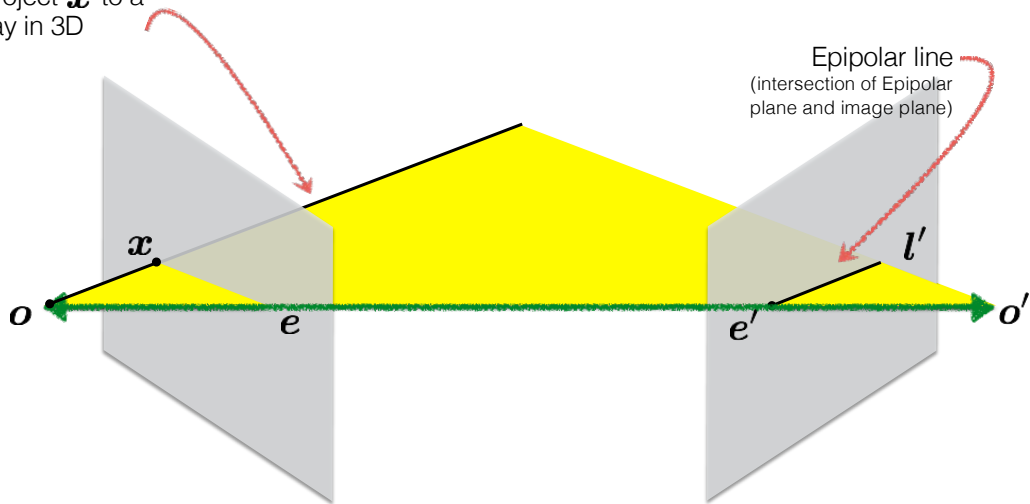


Quiz



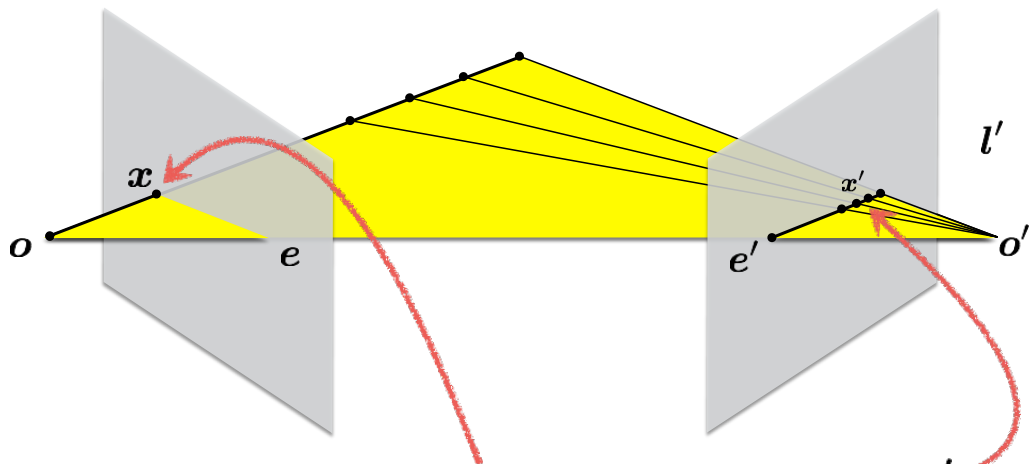
Epipolar constraint

Backproject $\mathbf{x}$ to a
ray in 3D

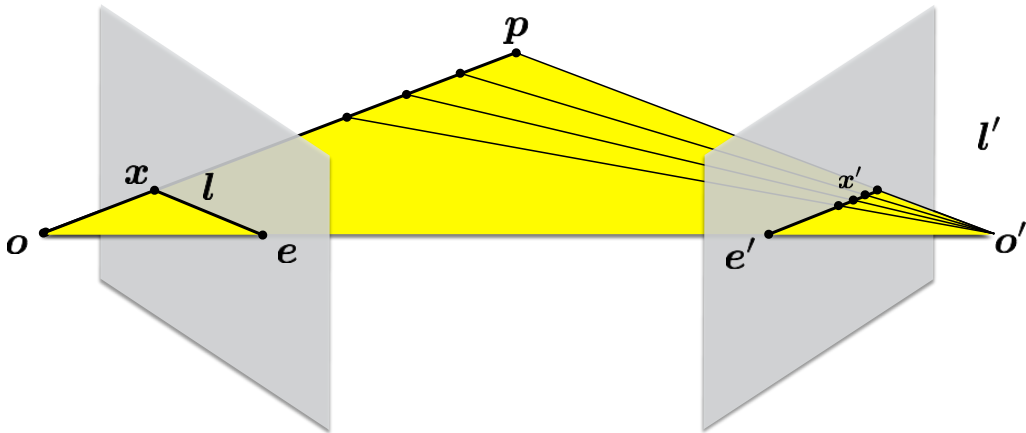


Another way to construct the epipolar plane, this time given $\mathbf{x}$

Epipolar constraint



Potential matches for $\mathbf{x}$ lie on the epipolar line l'



The point $\mathbf{x}$ (left image) maps to a _____ in the right image

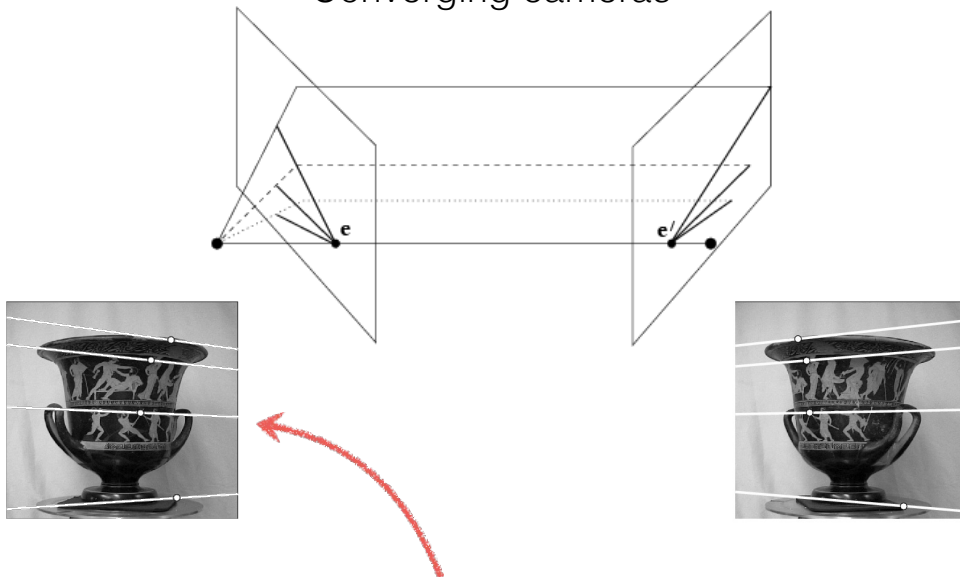
The baseline connects the _____ and _____

An epipolar line (left image) maps to a _____ in the right image

An epipole $\mathbf{e}$ is a projection of the _____ on the image plane

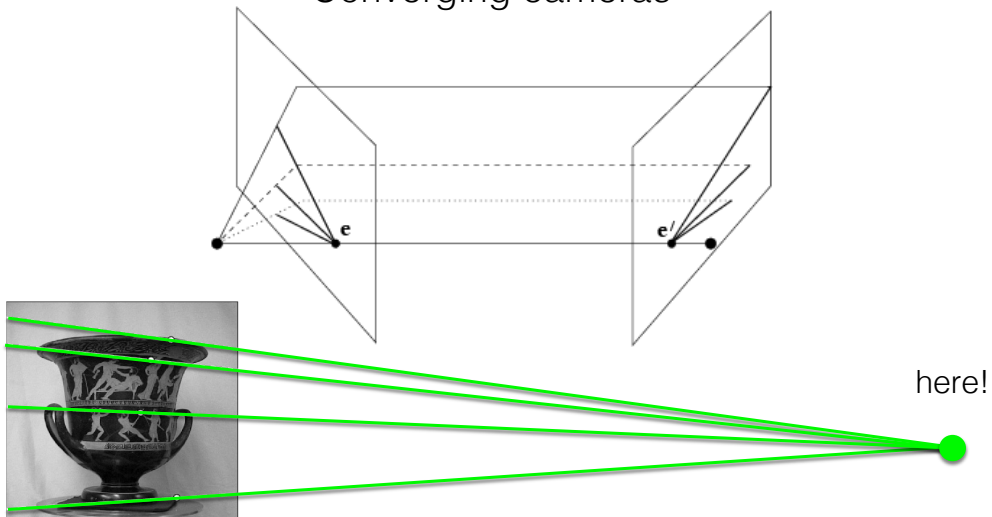
All epipolar lines in an image intersect at the _____

Converging cameras



Where is the epipole in this image?

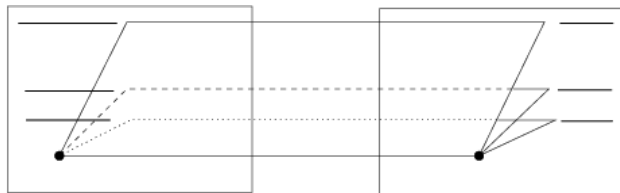
Converging cameras



Where is the epipole in this image?

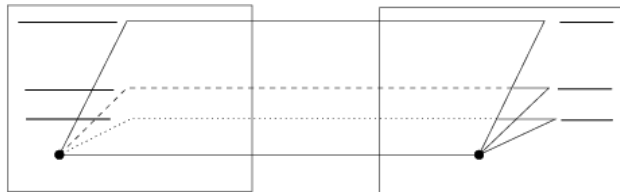
It's not always in the image

Parallel cameras



Where is the epipole?

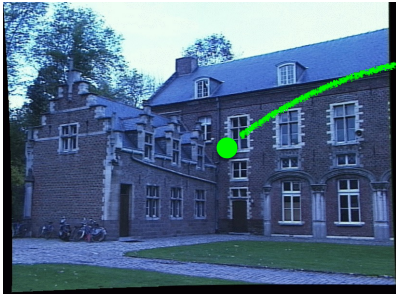
Parallel cameras



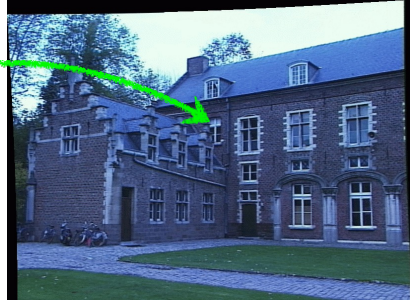
epipole at infinity

The epipolar constraint is an important concept for stereo vision

Task: Match point in left image to point in right image



Left image

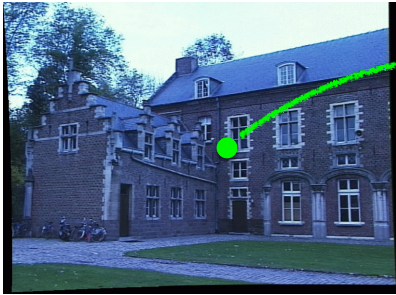


Right image

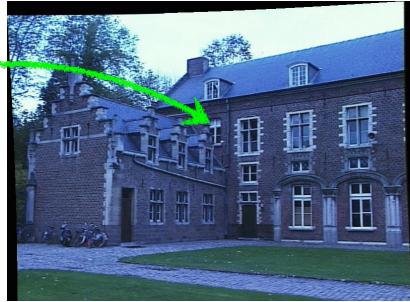
How would you do it?

The epipolar constraint is an important concept for stereo vision

Task: Match point in left image to point in right image



Left image



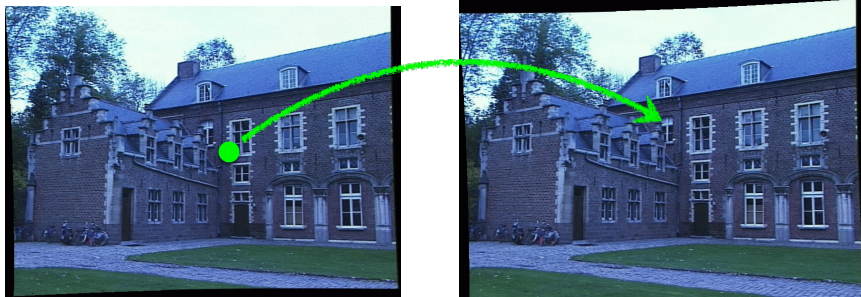
Right image

Want to avoid search over entire image

Epipolar constraint reduces search to a single line

The epipolar constraint is an important concept for stereo vision

Task: Match point in left image to point in right image



Left image

Right image

Want to avoid search over entire image

Epipolar constraint reduces search to a single line

How do you compute the epipolar line?

Next Time: Essential & Fundamental Matrices,
8-point algorithm, active stereo



References

Basic reading:

- Szeliski textbook, Section 8.1 (not 8.1.1-8.1.3), Chapter 11, Section 12.2.
- Hartley and Zisserman, Section 11.12.

References

Basic reading:

- Szeliski textbook, Section 2.1.5, 6.2.
- Bouguet, “Camera calibration toolbox for Matlab,” available at http://www.vision.caltech.edu/bouguetj/calib_doc/

The main resource for camera calibration in Matlab, where the screenshots in this lecture were taken from. It also has a detailed of the camera calibration algorithm and an extensive reference section.

Additional reading:

- Hartley and Zisserman, “Multiple View Geometry in Computer Vision,” Cambridge University Press 2004.
Chapter 6 of this book has a very thorough treatment of camera models.
- Gortler, “Foundations of 3D Computer Graphics,” MIT Press 2012.
Chapter 10 of this book has a nice discussion of pinhole cameras from a graphics point of view.
- Zhang, “A flexible new technique for camera calibration,” PAMI 2000.
The paper that introduced camera calibration from multiple views of a planar target.