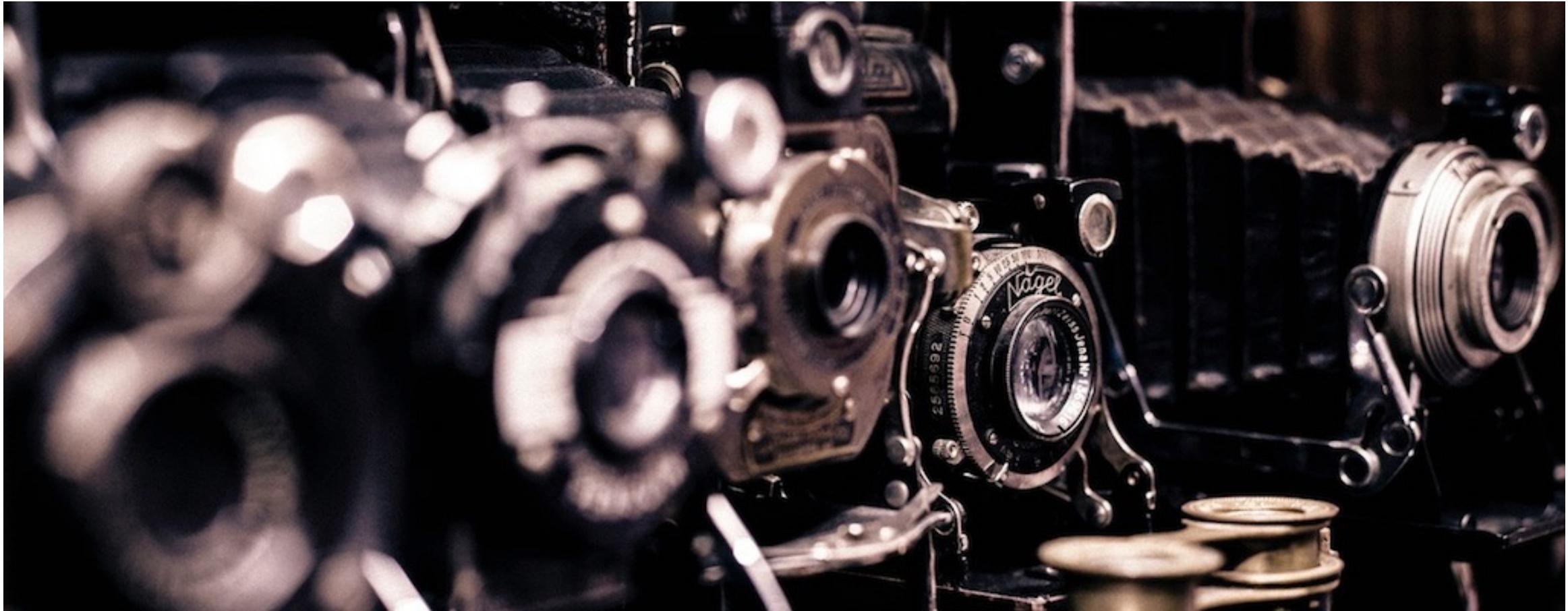


Geometric camera models



Overview of today's lecture

- Some motivational imaging experiments.
- Pinhole camera.
- Accidental pinholes.
- Camera matrix.

Slide credits

Most of these slides were adapted from:

- Kris Kitani (15-463, Fall 2016).

Some slides inspired from:

- Fredo Durand (MIT).

Some motivational imaging experiments

Let's say we have a sensor...



digital sensor
(CCD or CMOS)

... and an object we like to photograph

real-world
object



digital sensor
(CCD or CMOS)

What would an image taken like this look like?

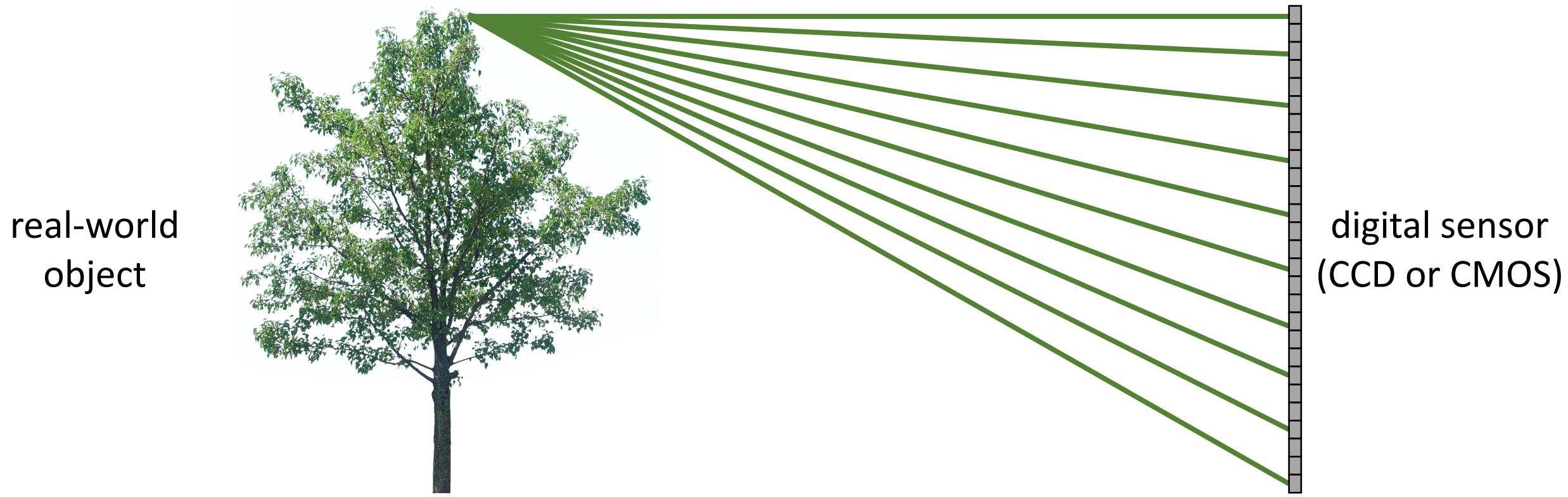
Bare-sensor imaging

real-world
object

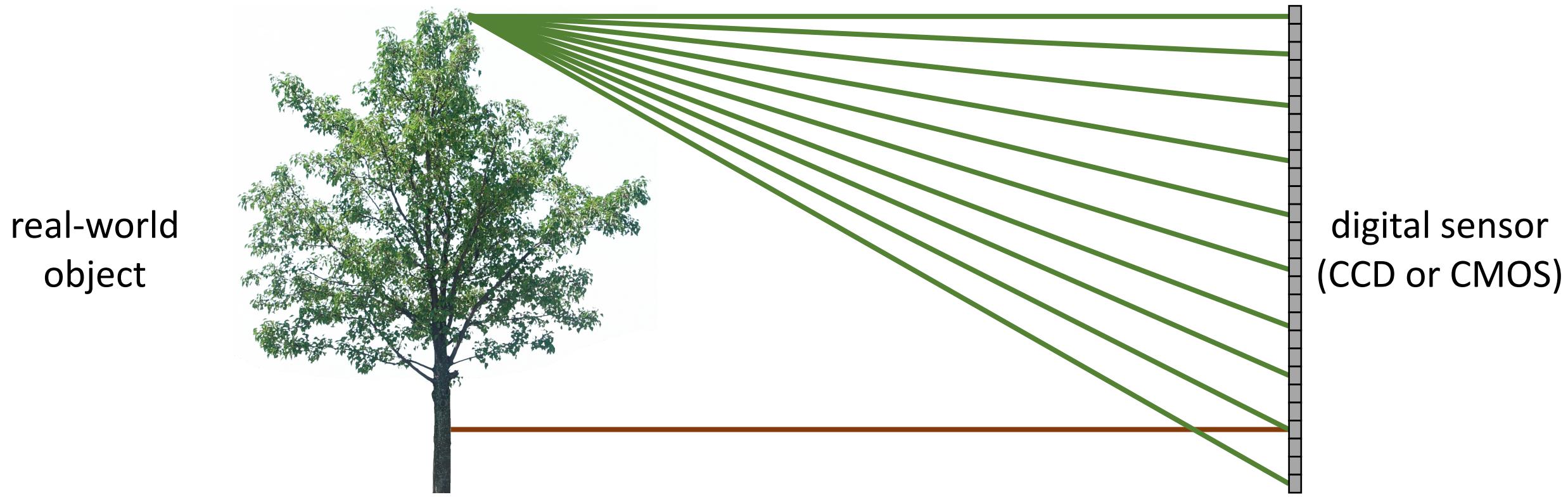


digital sensor
(CCD or CMOS)

Bare-sensor imaging

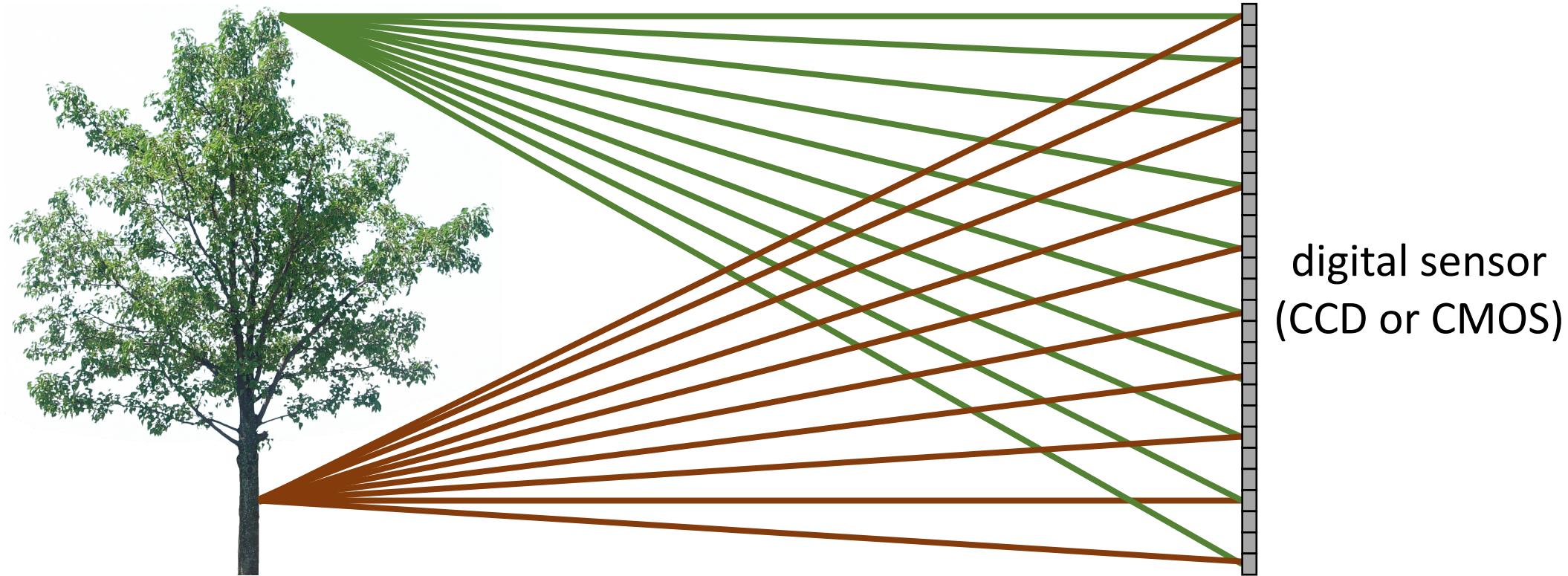


Bare-sensor imaging



Bare-sensor imaging

real-world
object



All scene points contribute to all sensor pixels

What does the
image on the
sensor look like?

Bare-sensor imaging



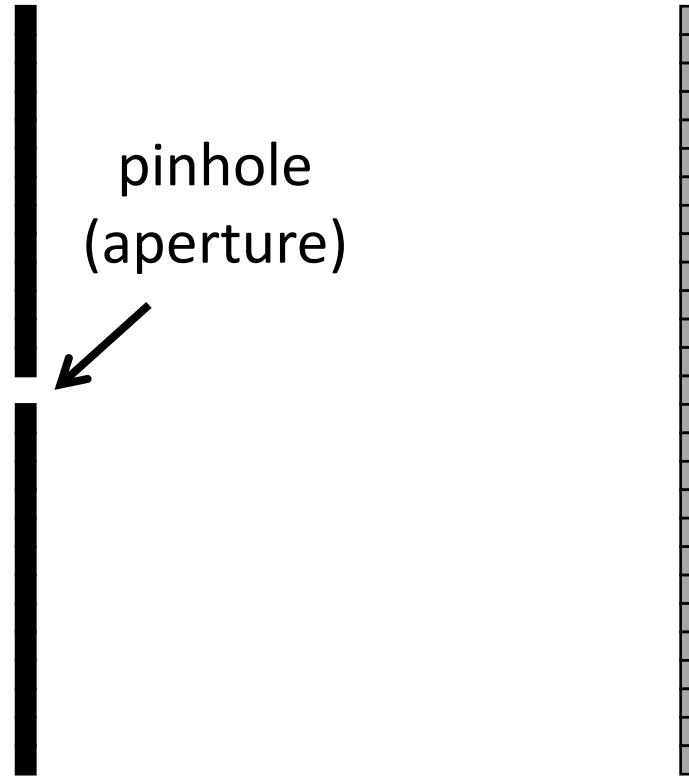
All scene points contribute to all sensor pixels

Let's add something to this scene

real-world
object

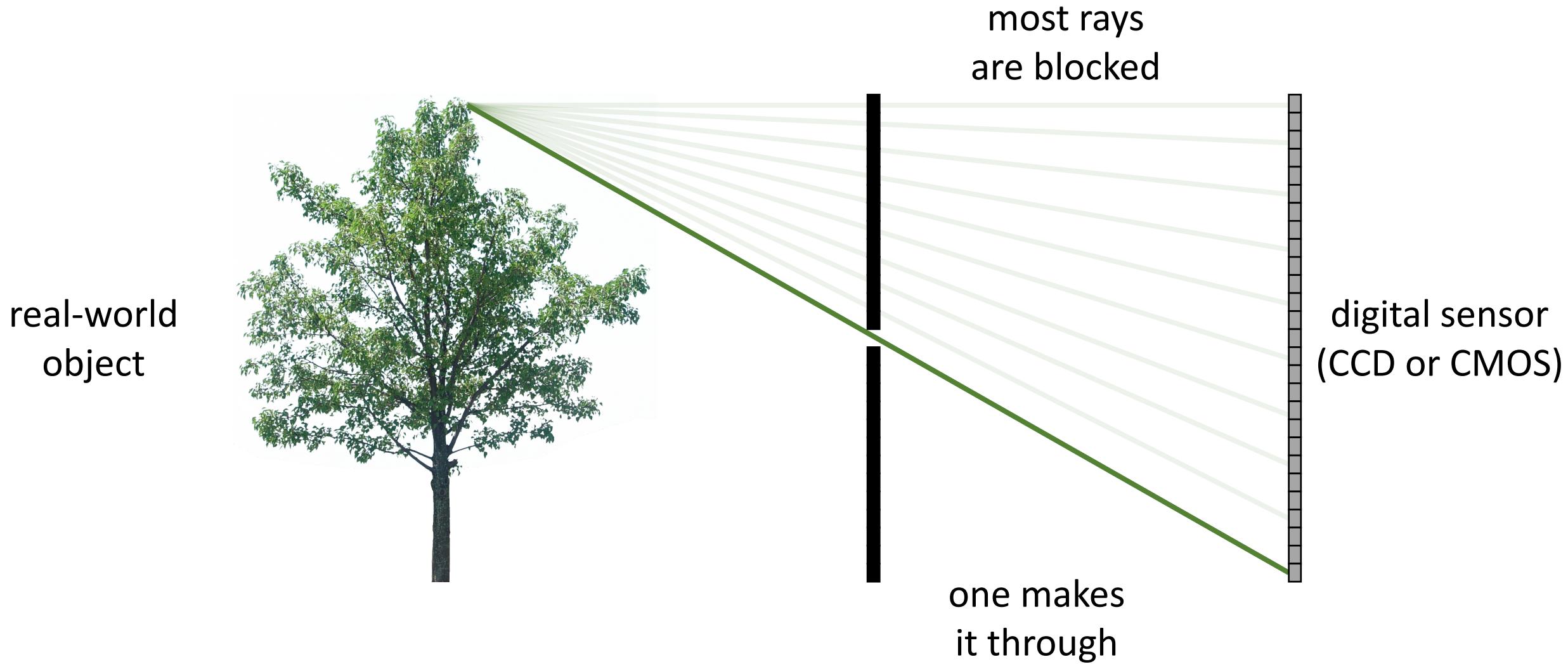


barrier (diaphragm)

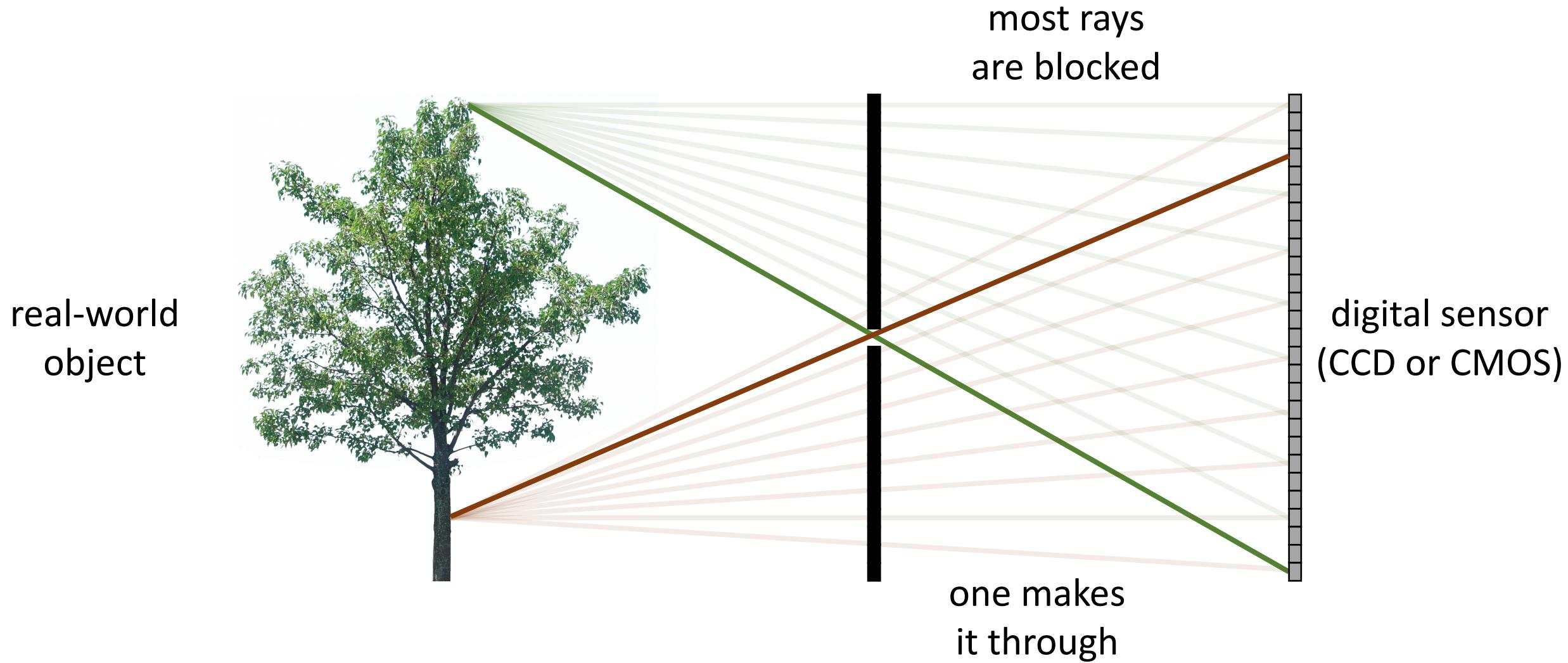


What would an image taken like this look like?

Pinhole imaging

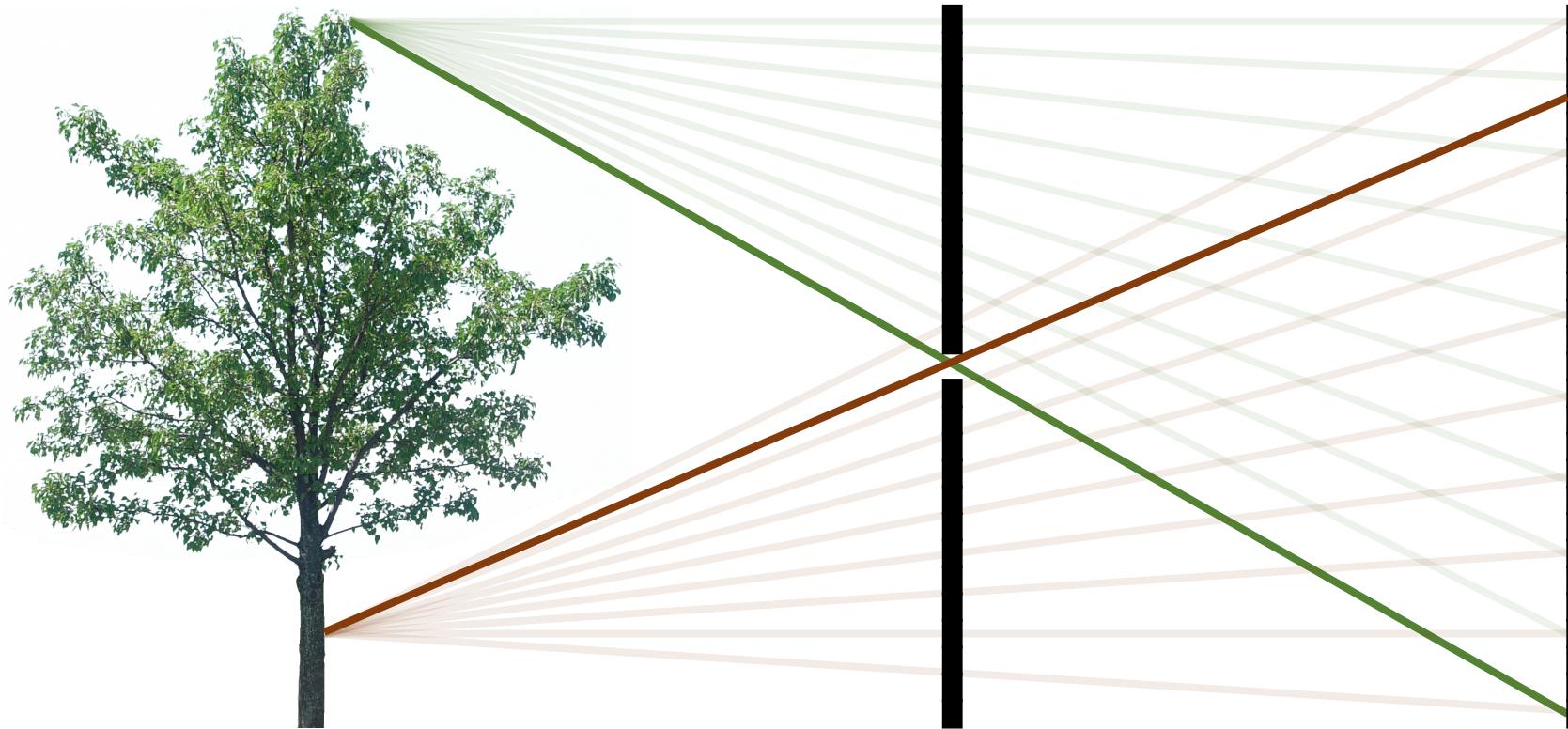


Pinhole imaging



Pinhole imaging

real-world
object

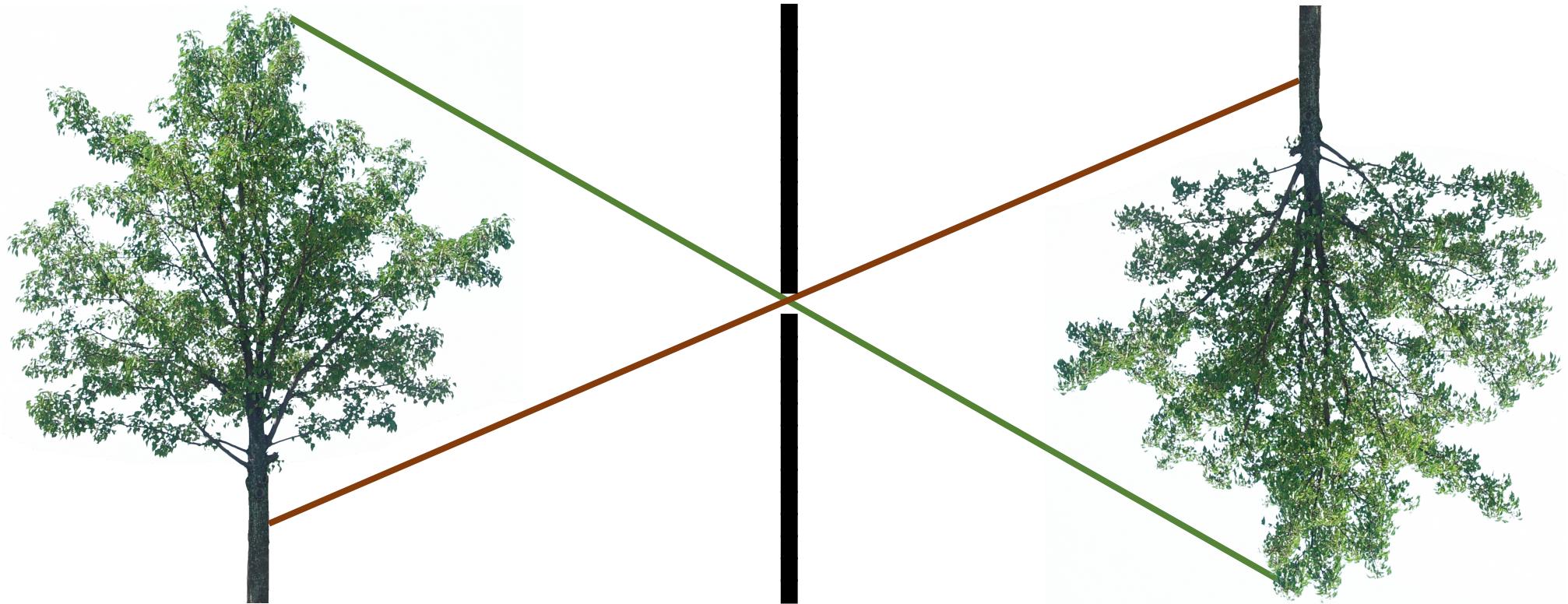


Each scene point contributes to only one sensor pixel

What does the
image on the
sensor look like?

Pinhole imaging

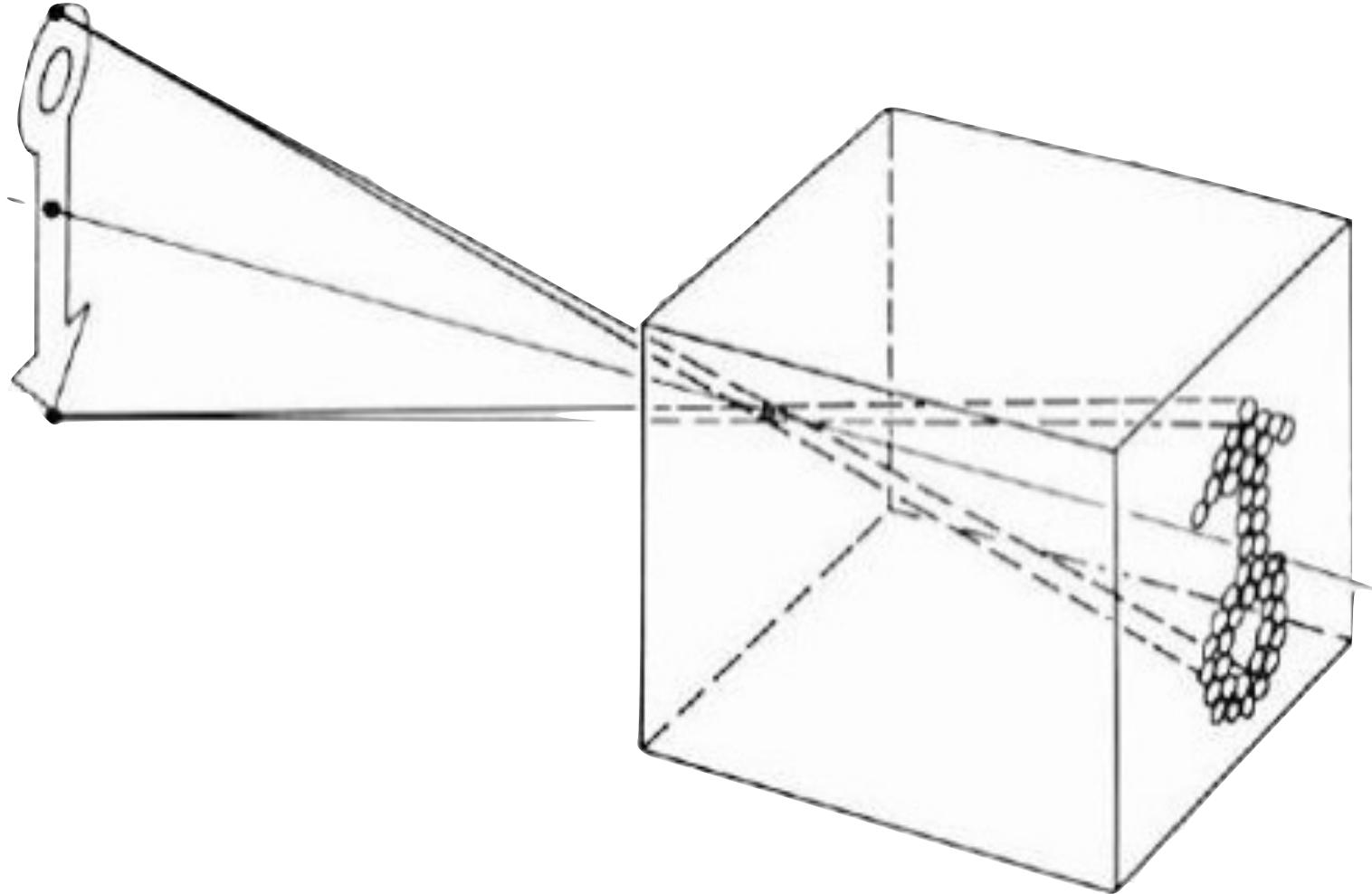
real-world
object



copy of real-world object
(inverted and scaled)

Pinhole camera

Pinhole camera a.k.a. camera obscura



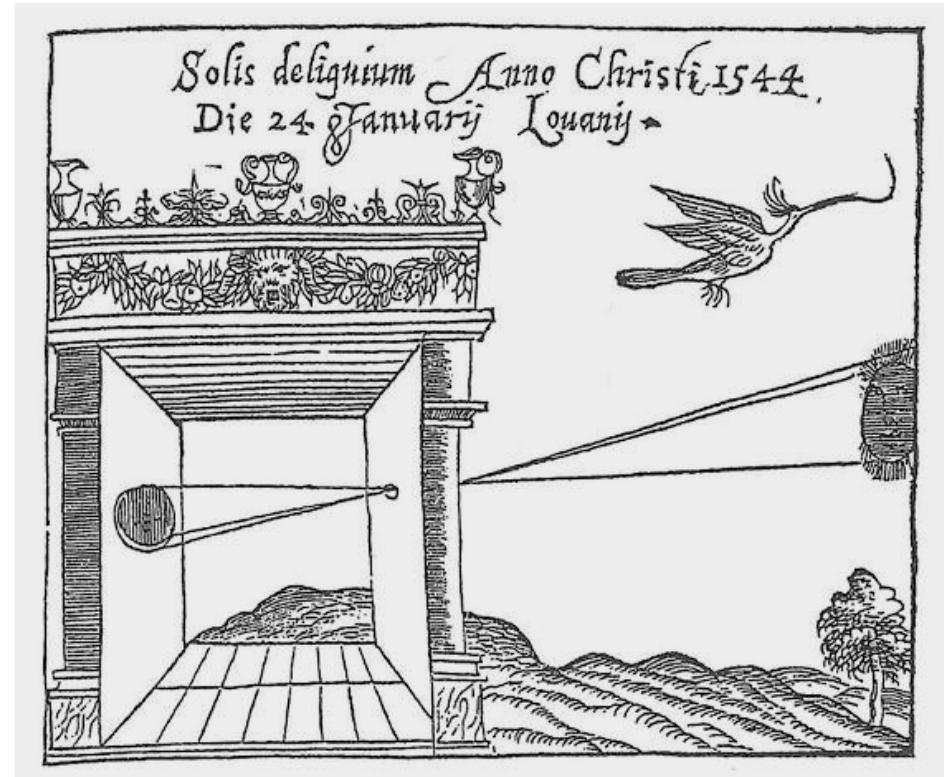
Pinhole camera a.k.a. camera obscura

First mention ...



Chinese philosopher Mozi
(470 to 390 BC)

First camera ...



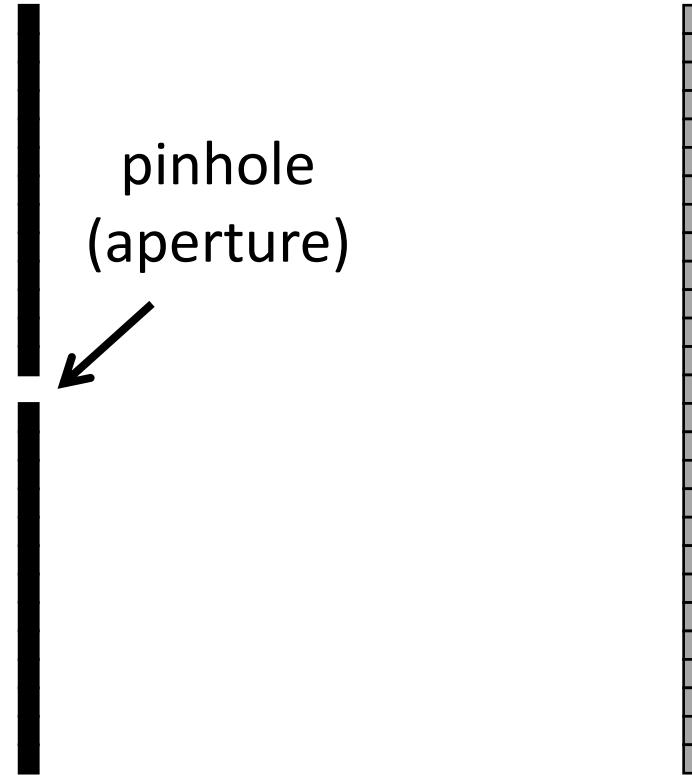
Greek philosopher Aristotle
(384 to 322 BC)

Pinhole camera terms

real-world
object



barrier (diaphragm)



digital sensor
(CCD or CMOS)

Pinhole camera terms

real-world
object



barrier (diaphragm)

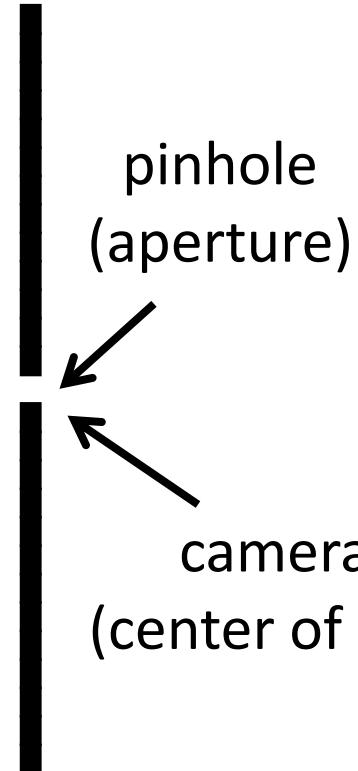
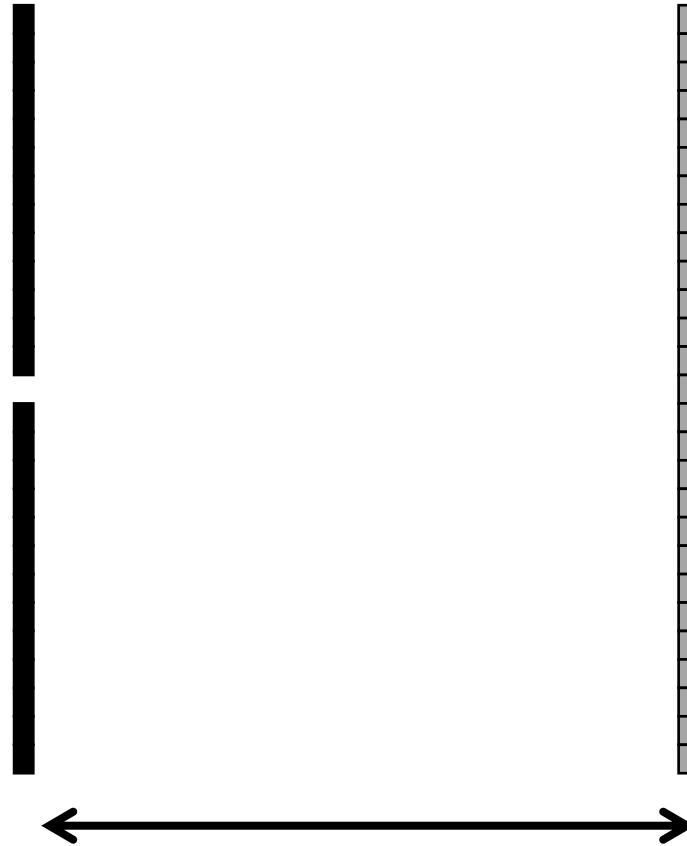


image plane

digital sensor
(CCD or CMOS)

Focal length

real-world
object

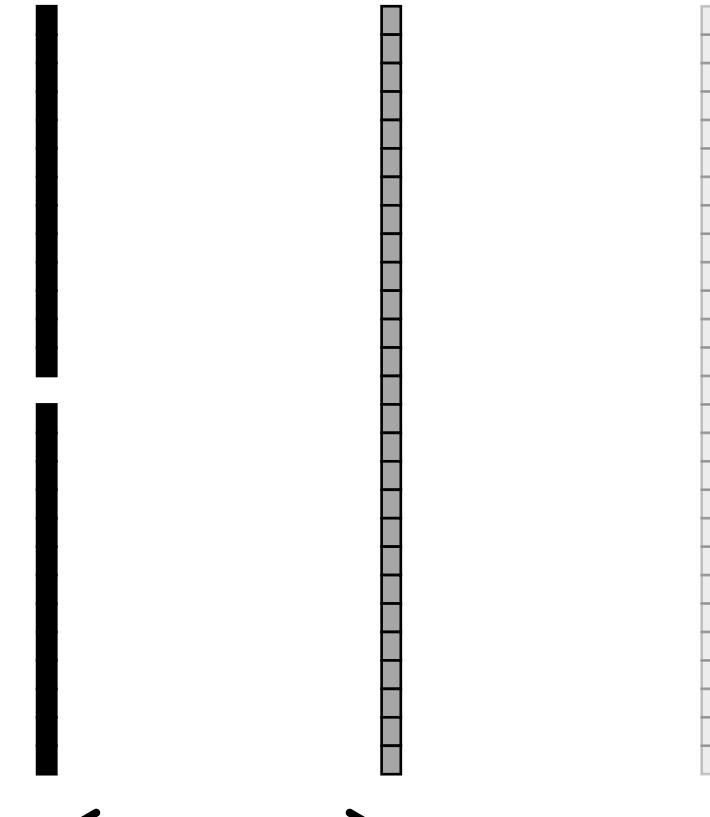


focal length f

Focal length

What happens as we change the focal length?

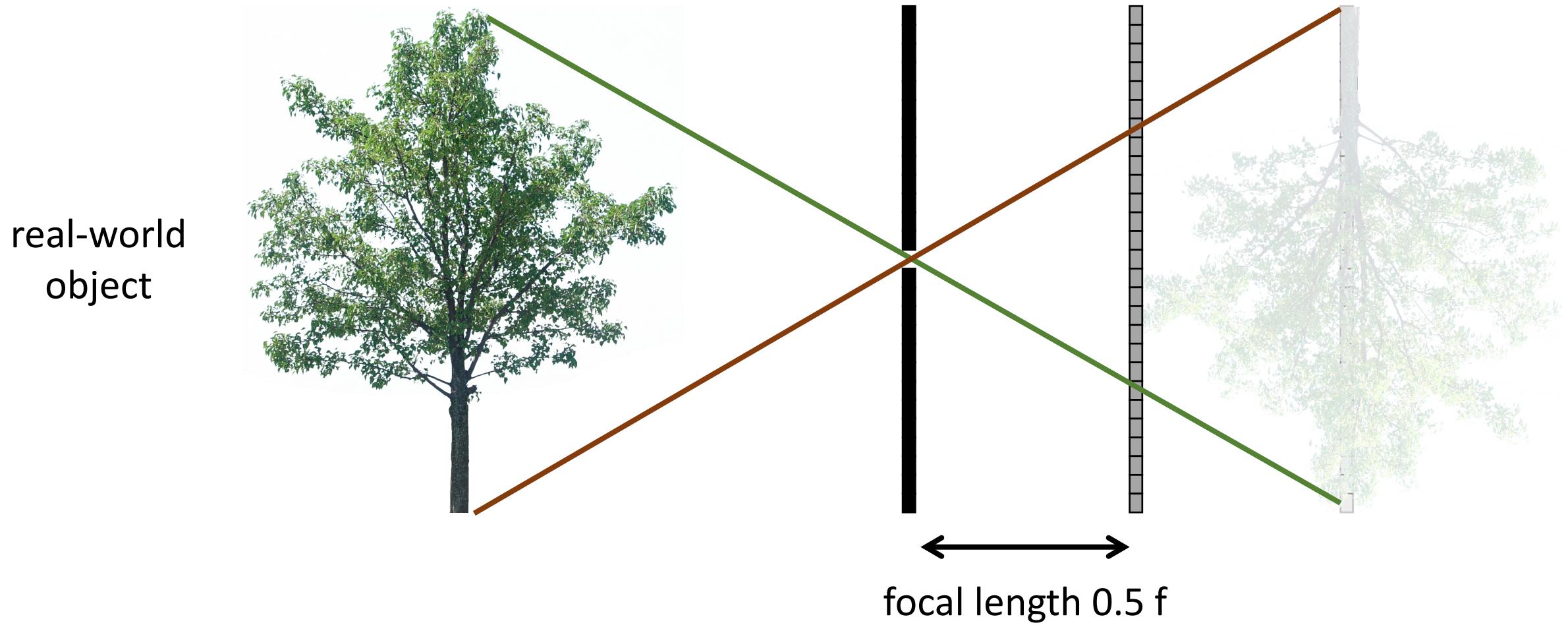
real-world
object



focal length $0.5 f$

Focal length

What happens as we change the focal length?

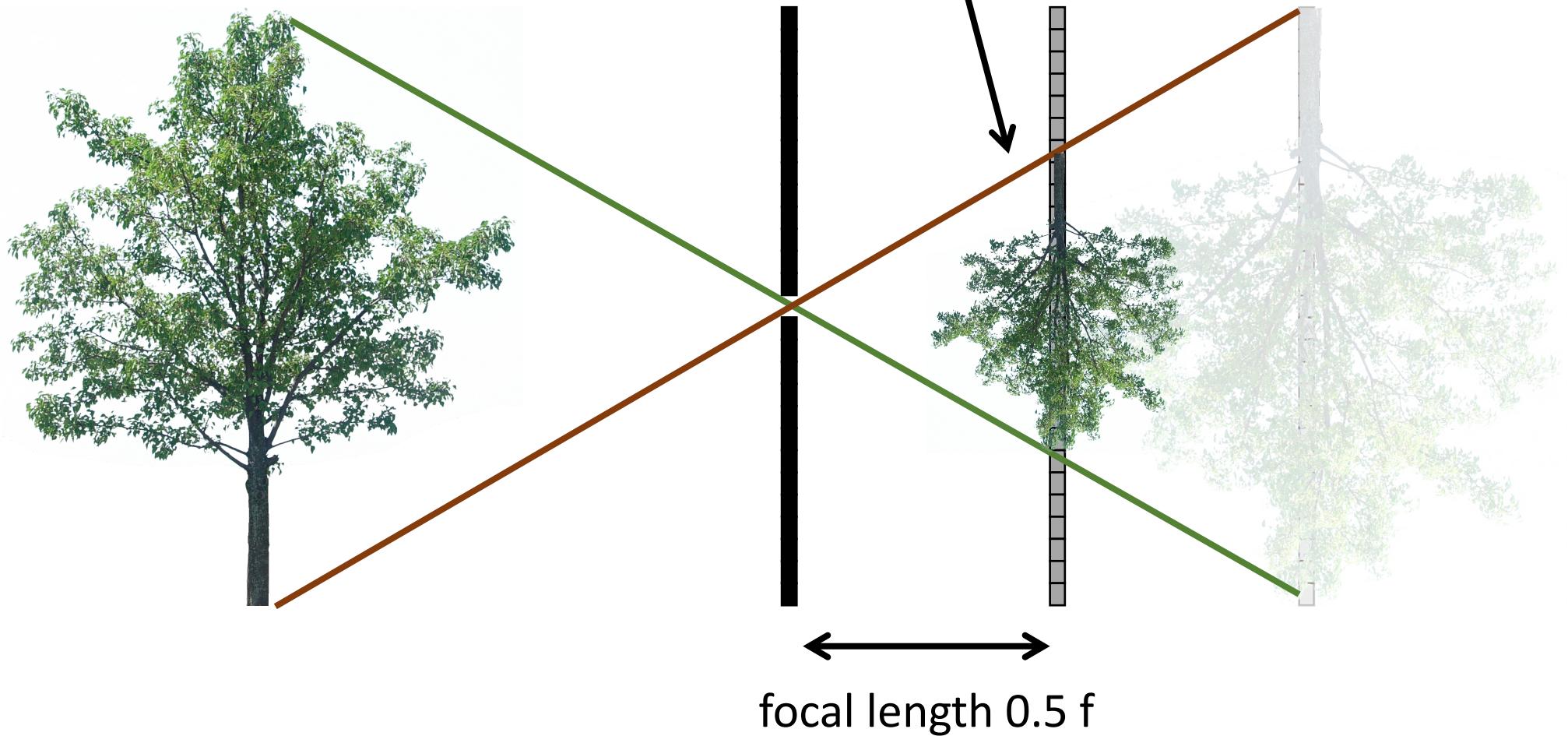


Focal length

What happens as we change the focal length?

object projection is half the size

real-world
object



Pinhole size

real-world
object



pinhole
diameter



Ideal pinhole has infinitesimally small size

- In practice that is impossible.

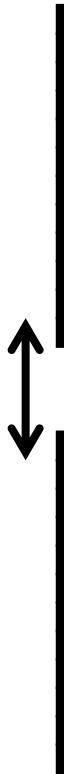
Pinhole size

What happens as we change the pinhole diameter?

real-world
object



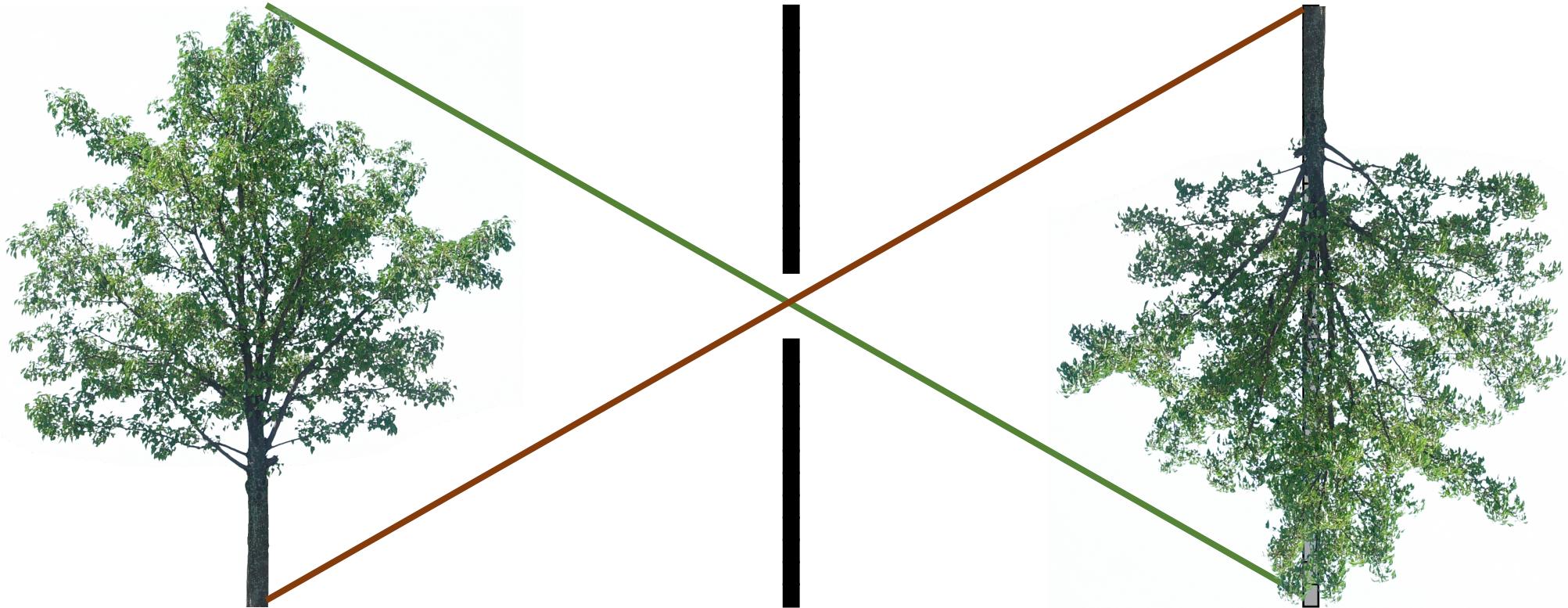
pinhole
diameter



Pinhole size

What happens as we change the pinhole diameter?

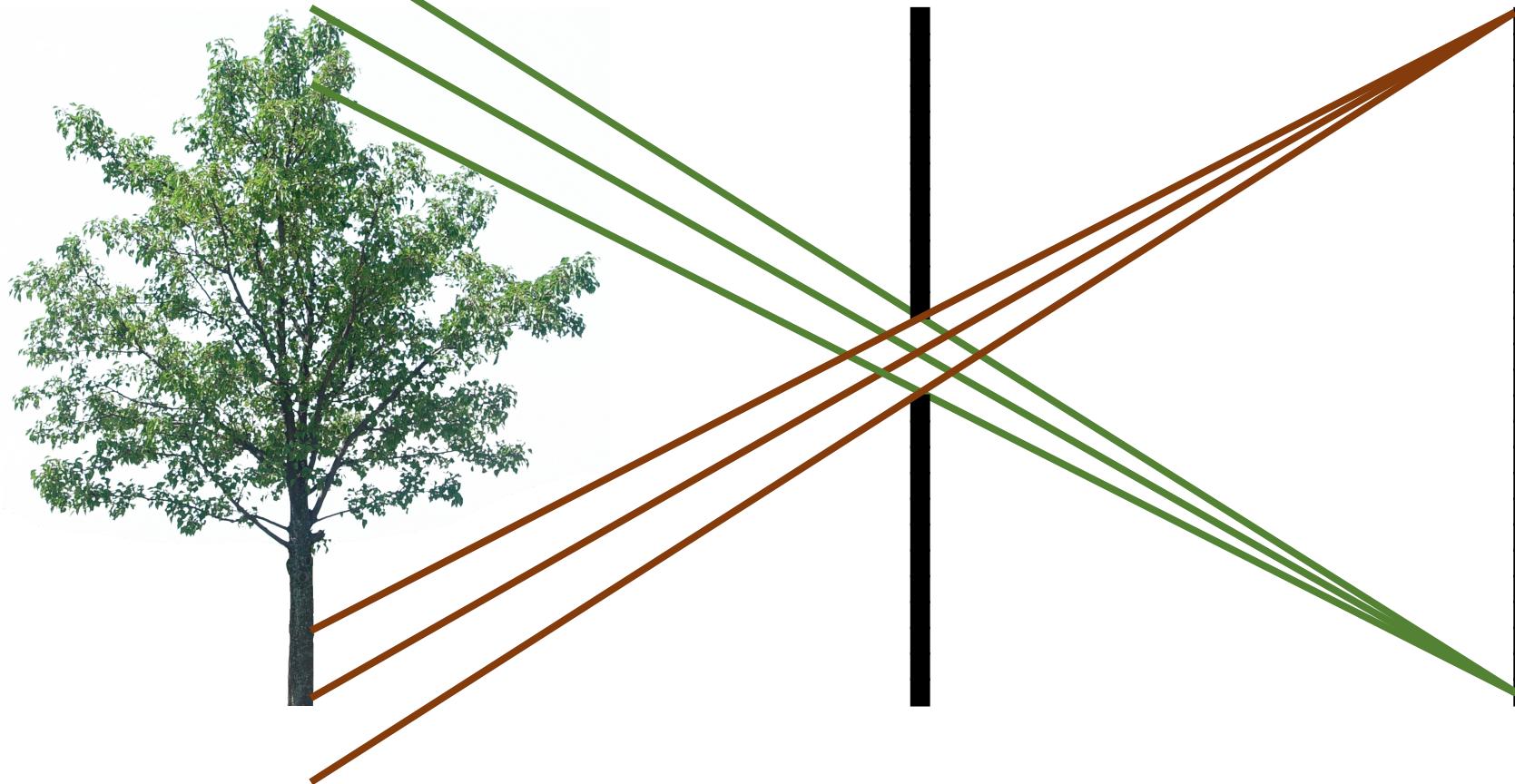
real-world
object



Pinhole size

What happens as we change the pinhole diameter?

real-world
object

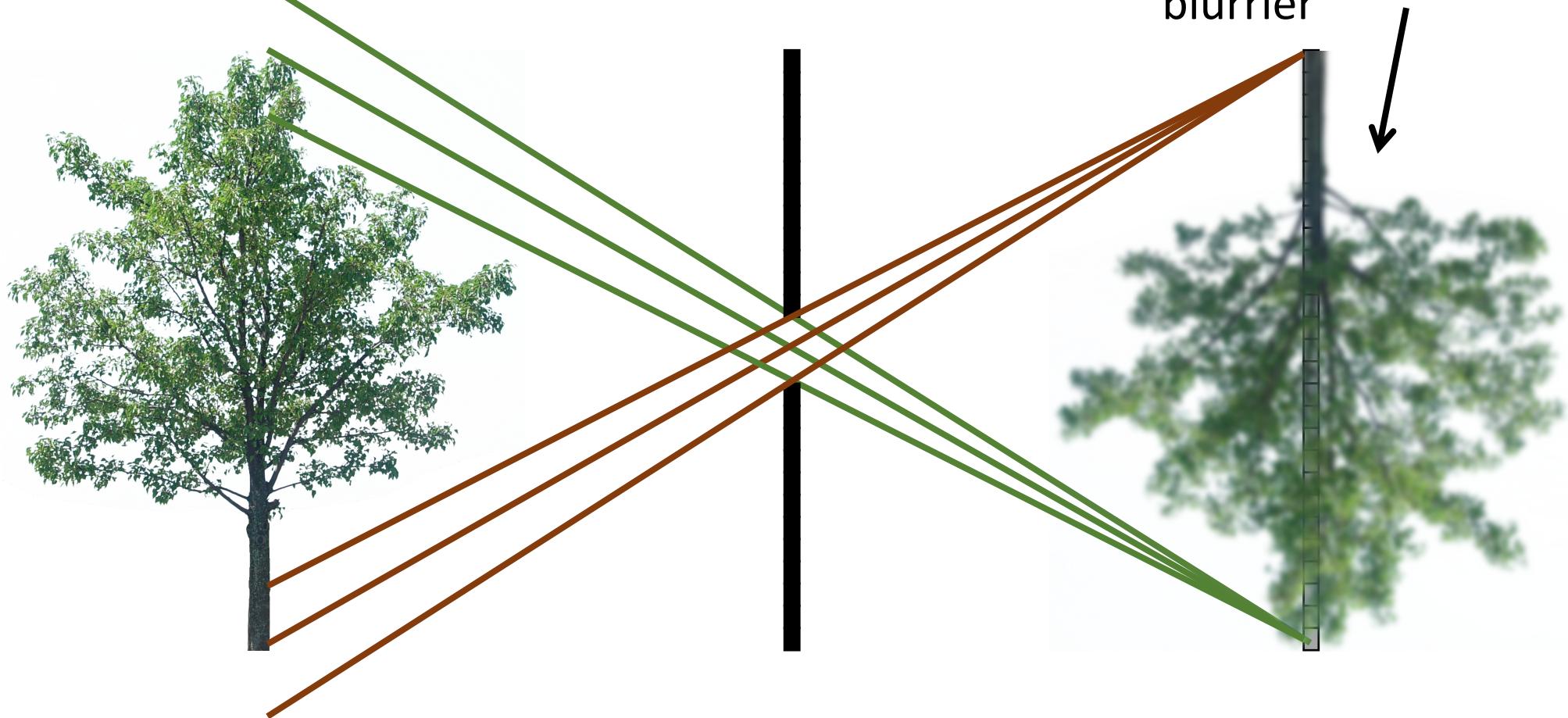


Pinhole size

What happens as we change the pinhole diameter?

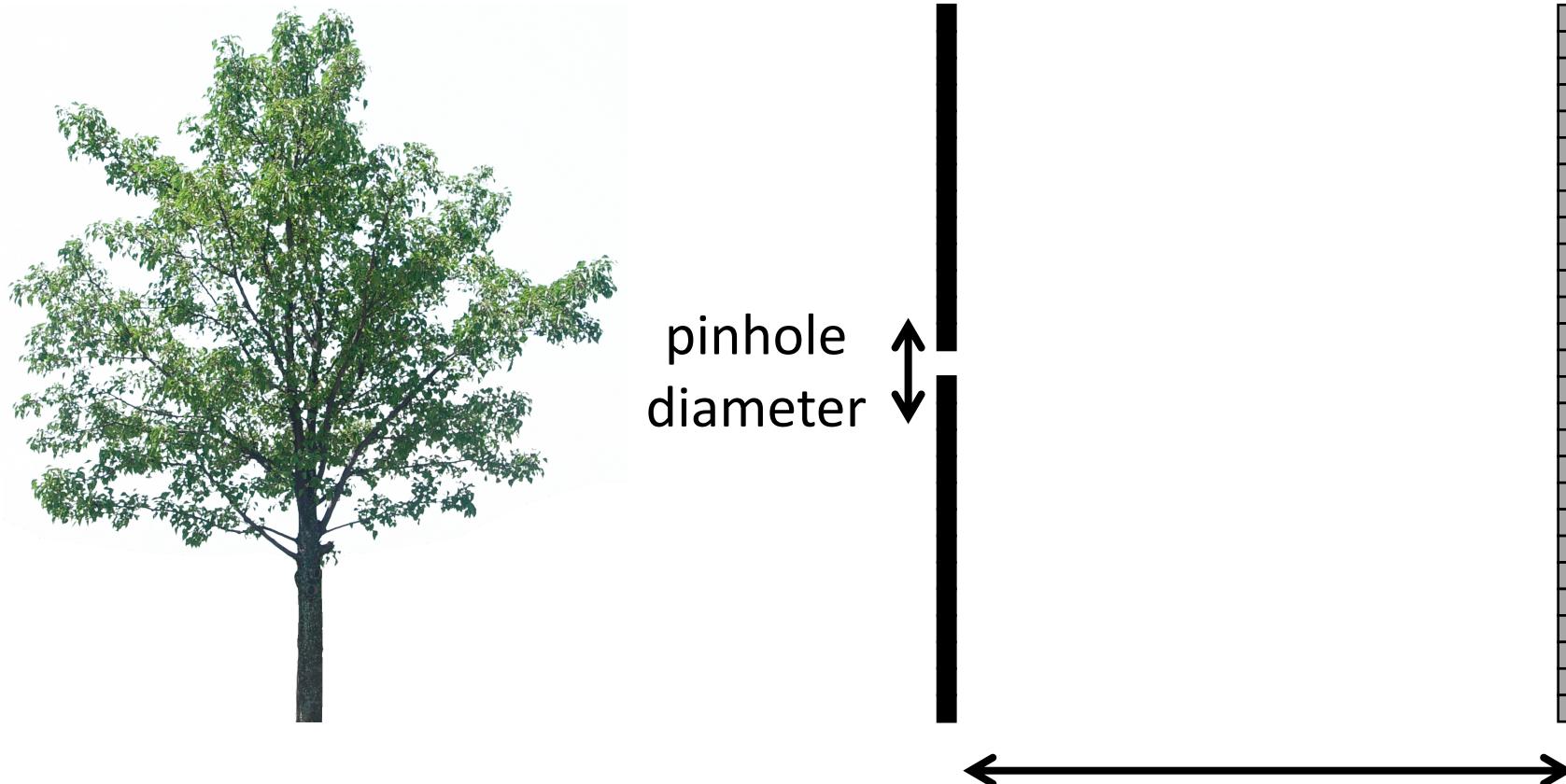
real-world
object

object projection becomes
blurrier



What about light efficiency?

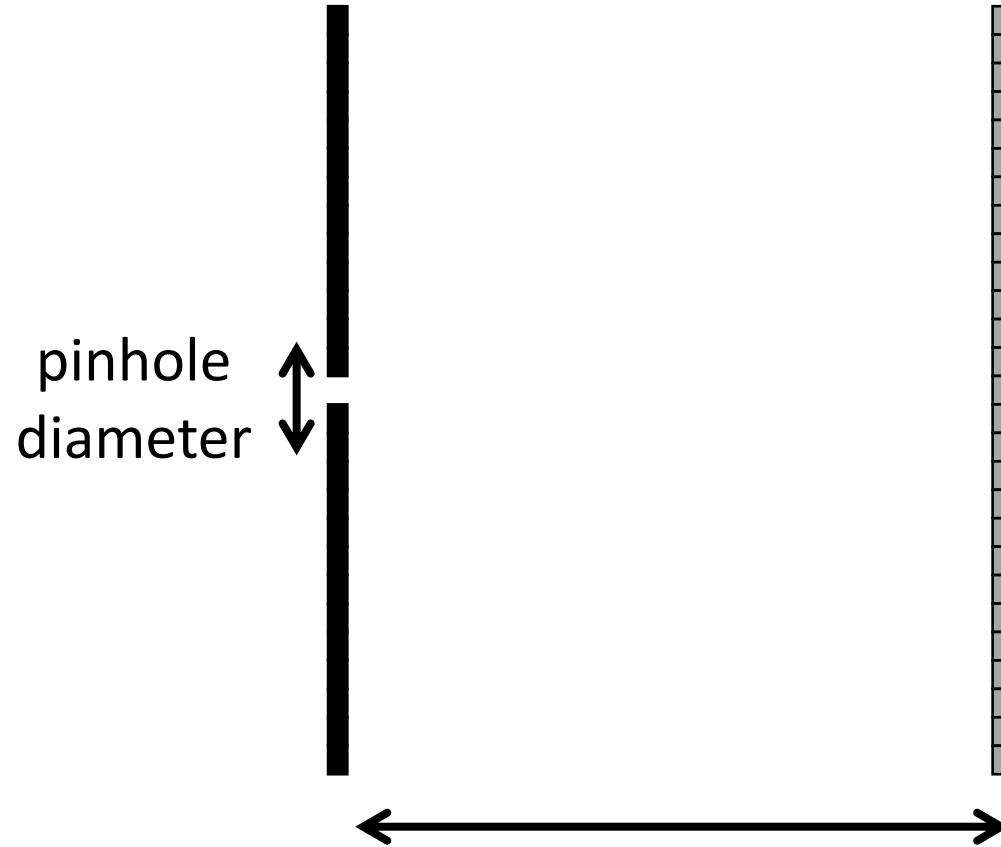
real-world
object



- What is the effect of doubling the pinhole diameter?
- What is the effect of doubling the focal length?

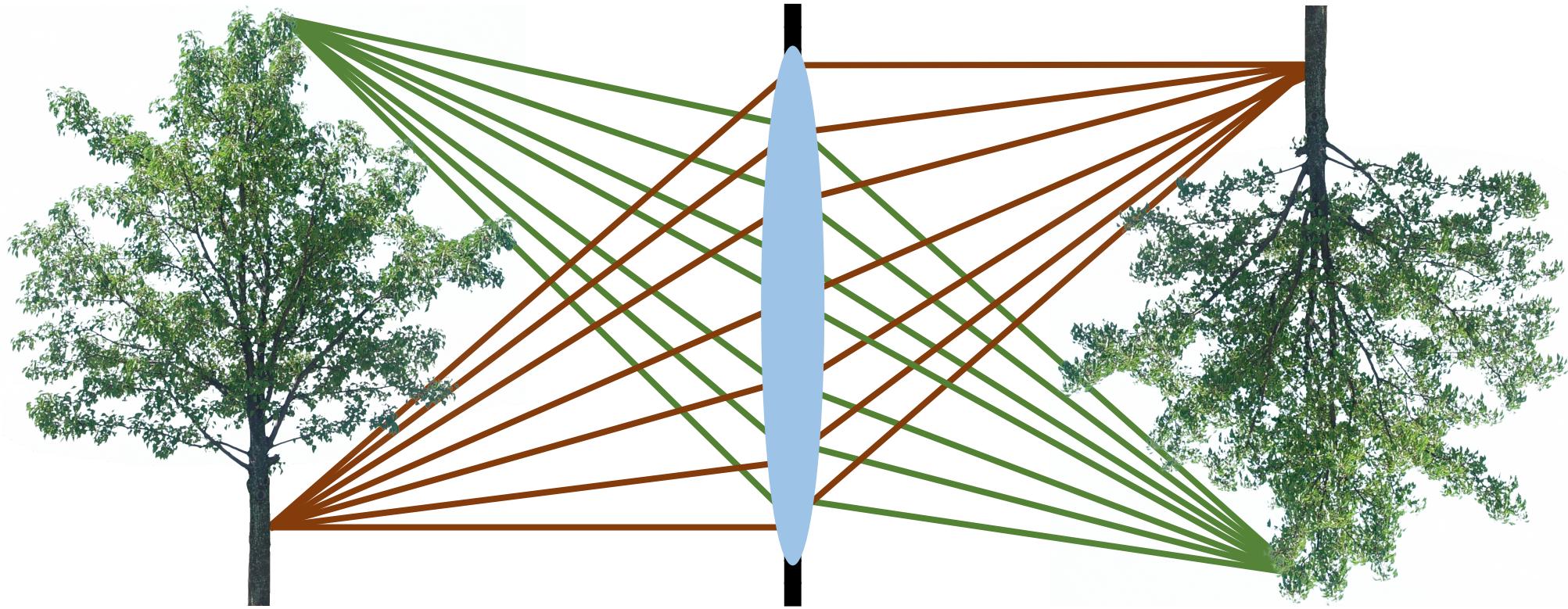
What about light efficiency?

real-world
object



- $2x$ pinhole diameter $\rightarrow 4x$ light
- $2x$ focal length $\rightarrow \frac{1}{4}x$ light

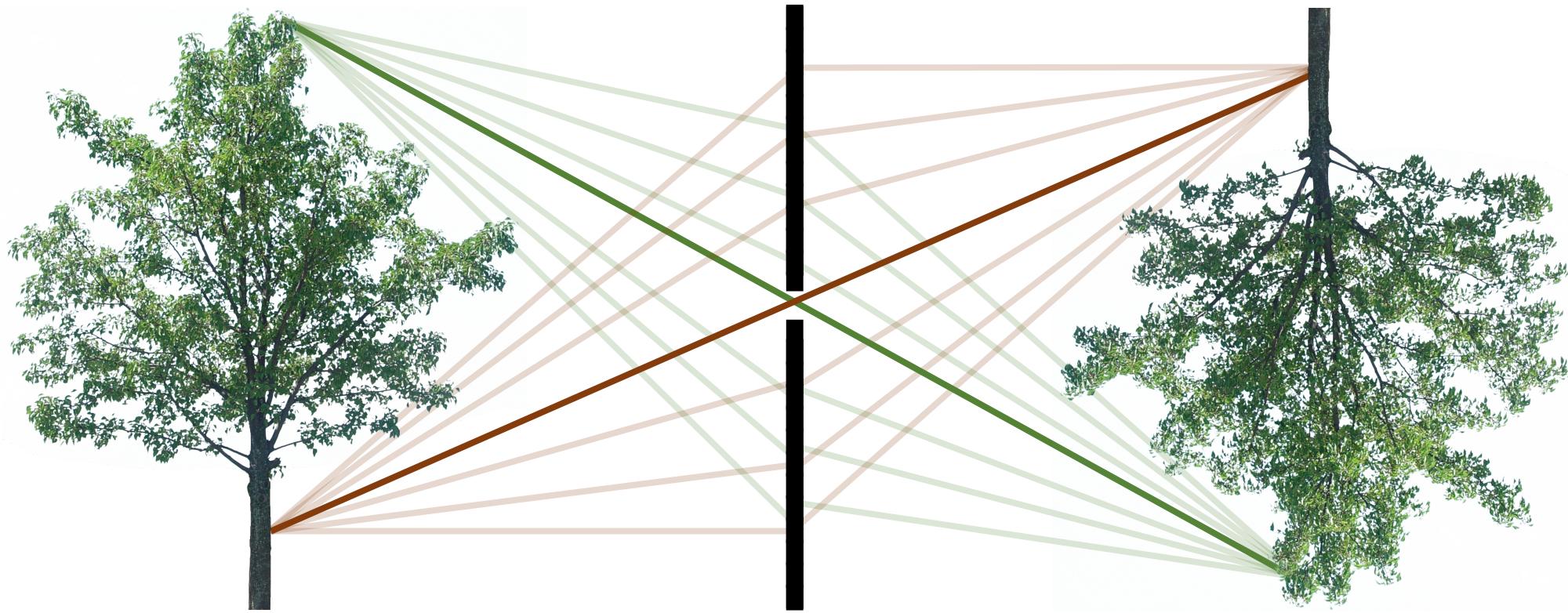
The lens camera



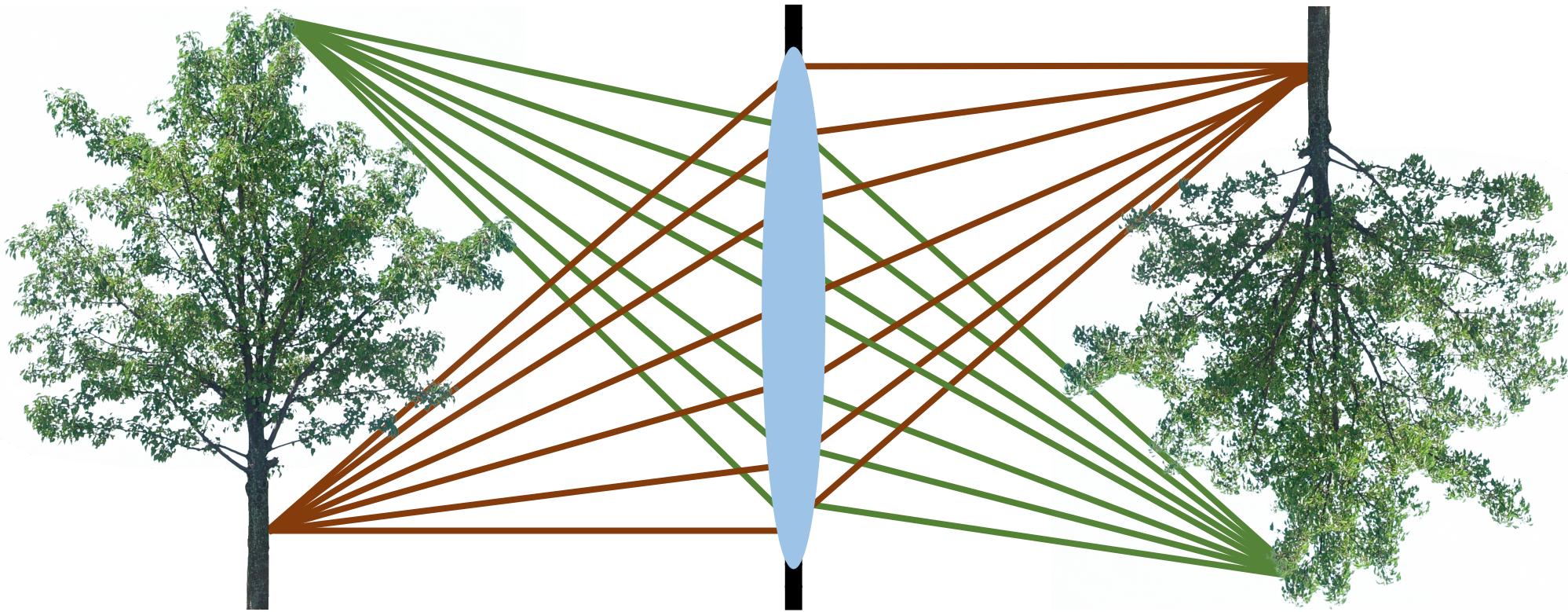
Lenses map “bundles” of rays from points on the scene to the sensor.

How does this mapping work exactly?

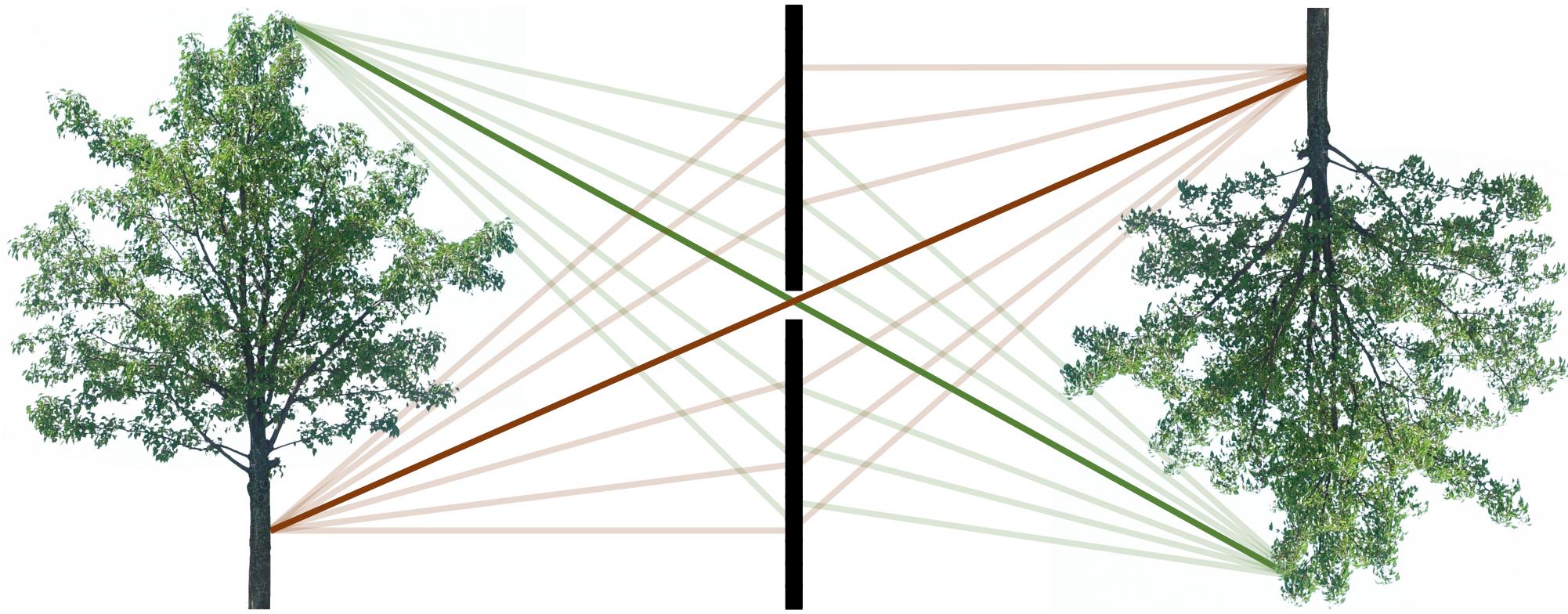
The pinhole camera



The lens camera

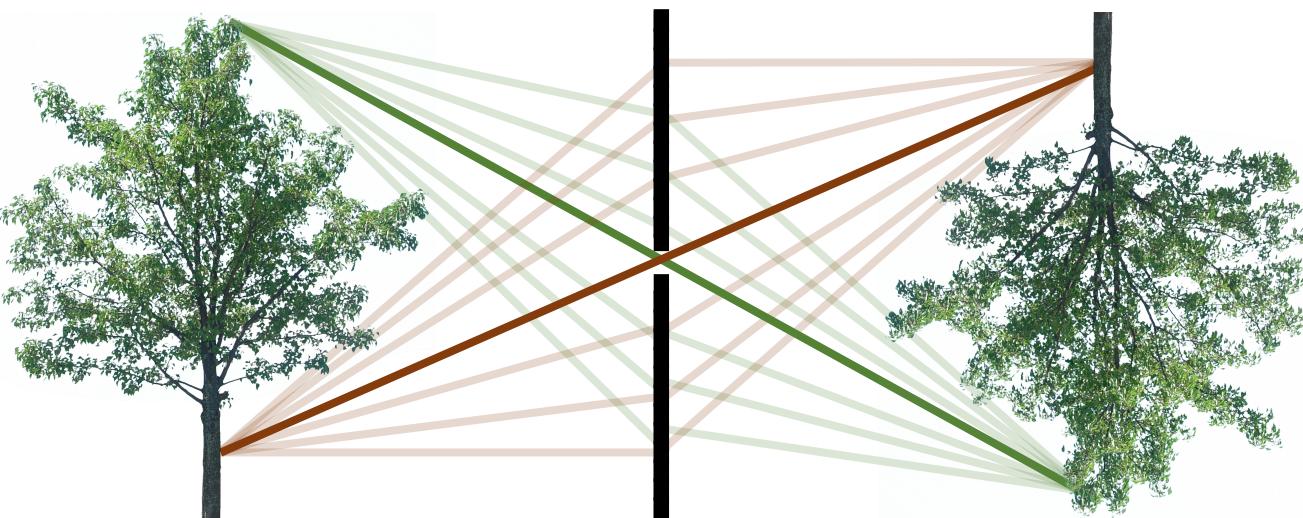
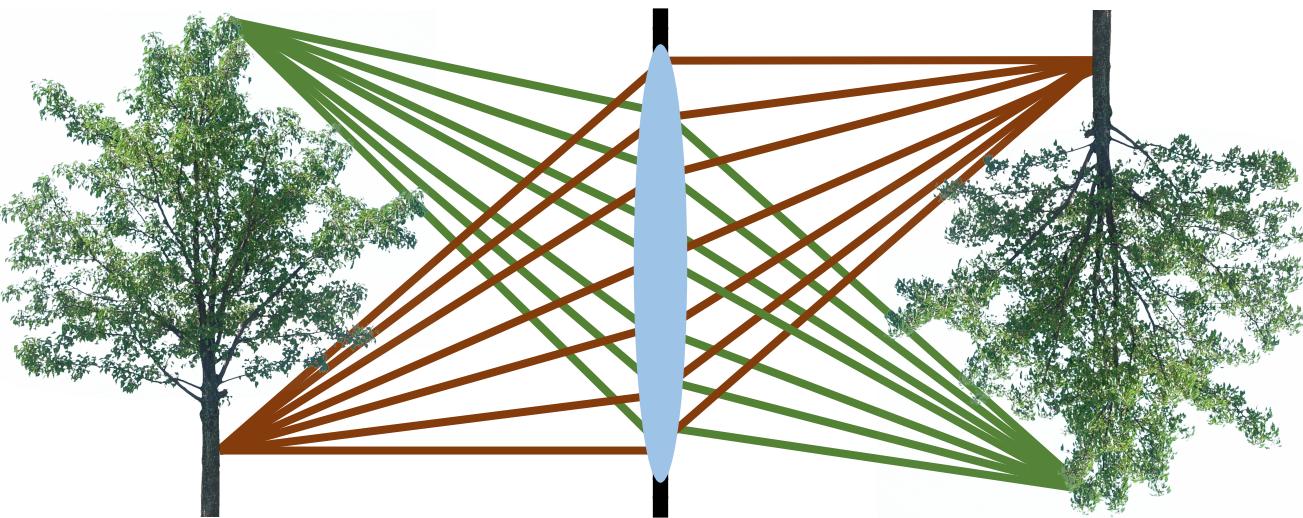


The pinhole camera



Central rays propagate in the same way for both models!

Describing both lens and pinhole cameras

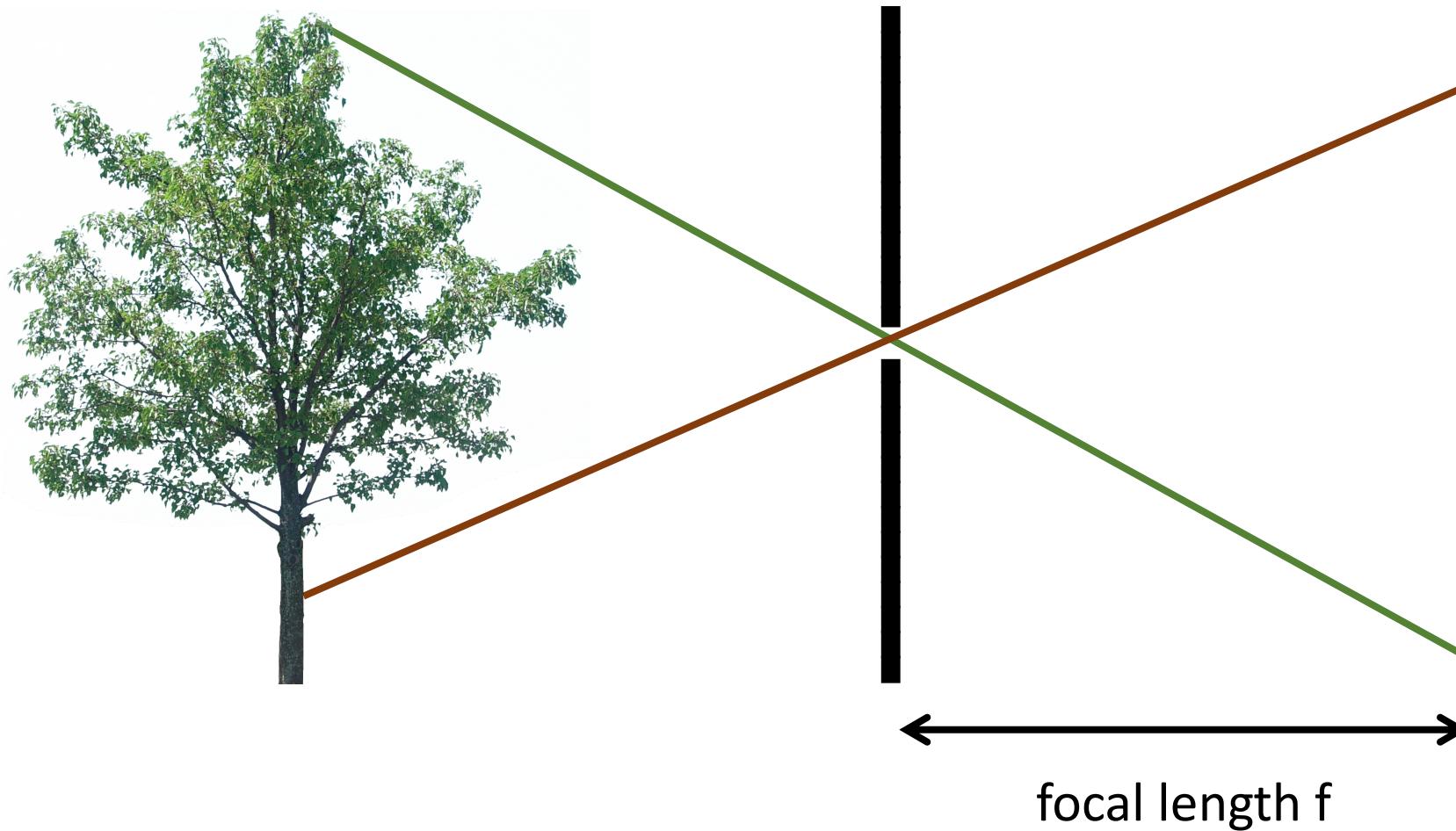


We can derive properties and descriptions that hold for both camera models if:

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

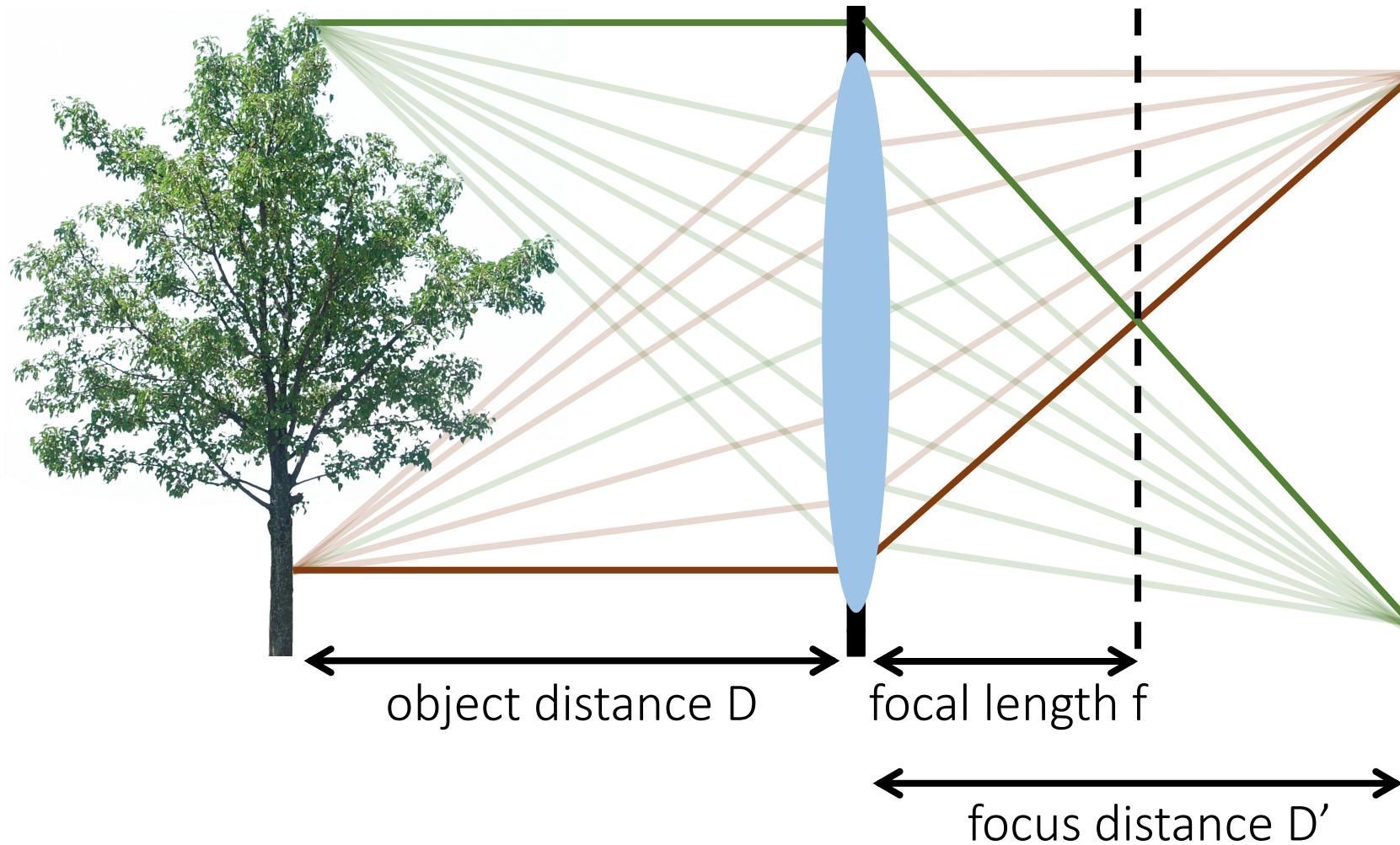
Important difference: focal length

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

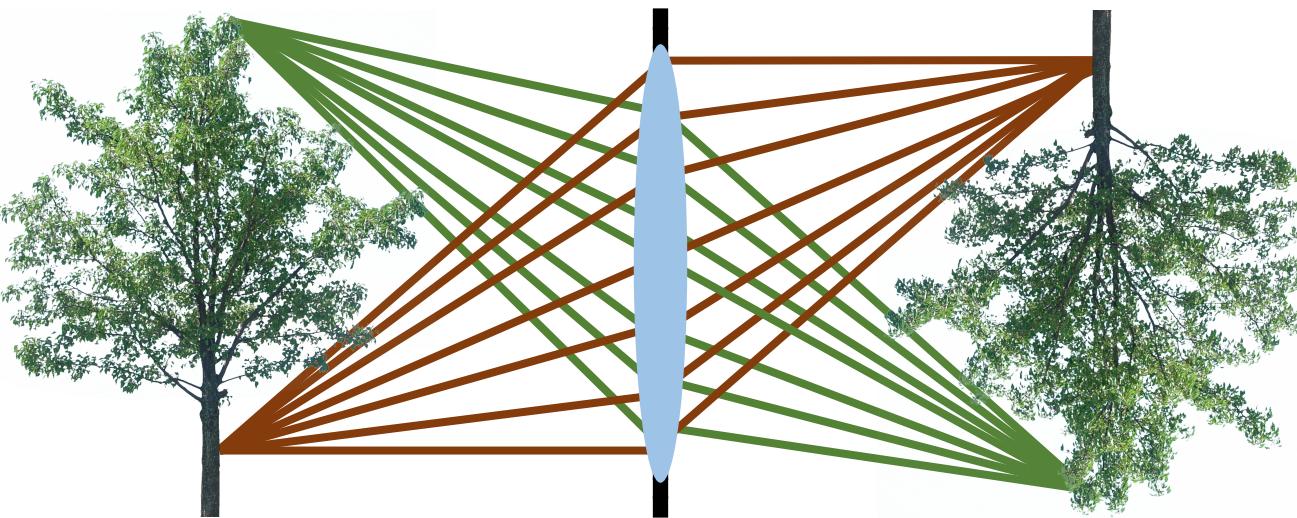


Important difference: focal length

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

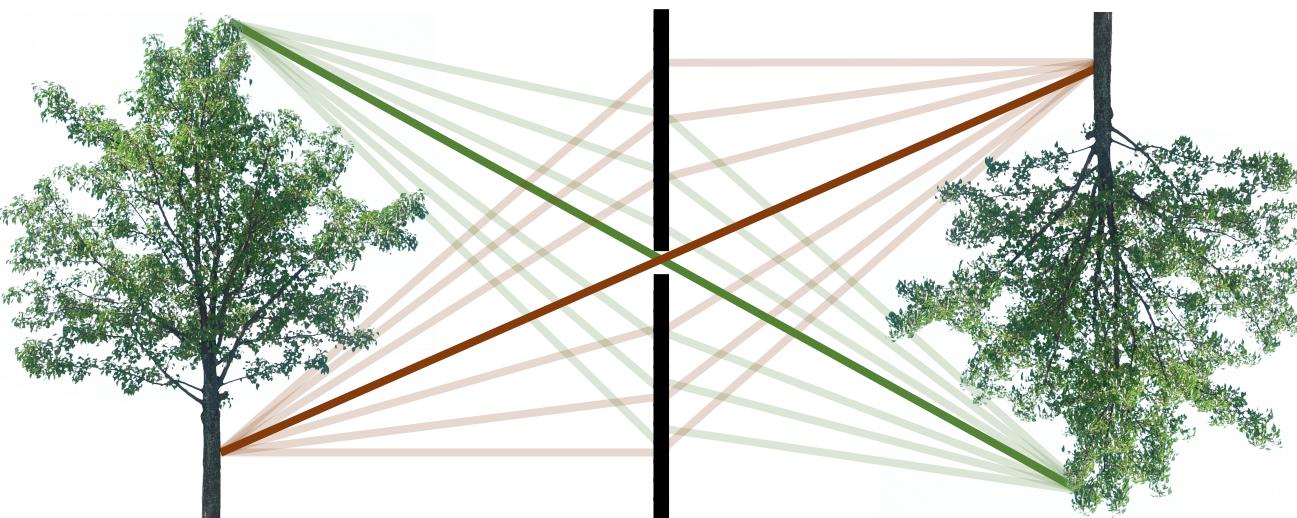


Describing both lens and pinhole cameras



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.



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.

Accidental pinholes





What does this image say about the world outside?



Accidental pinhole camera



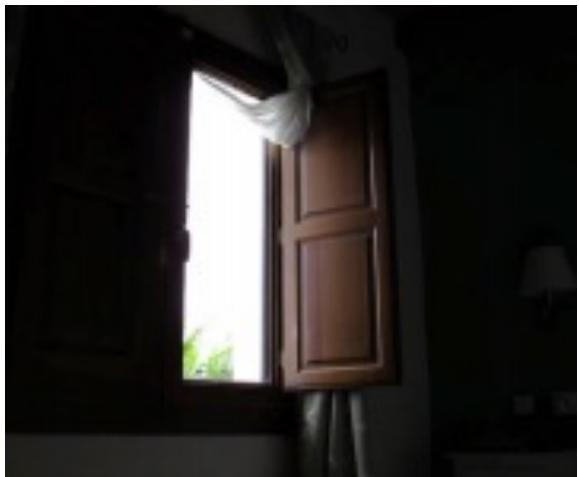
Antonio Torralba, William T. Freeman
Computer Science and Artificial Intelligence Laboratory (CSAIL)
MIT
torralba@mit.edu, billf@mit.edu

Accidental pinhole camera

projected pattern on the wall



upside down



window is an
aperture

window with smaller gap



view outside window



Pinhole cameras

What are we imaging here?



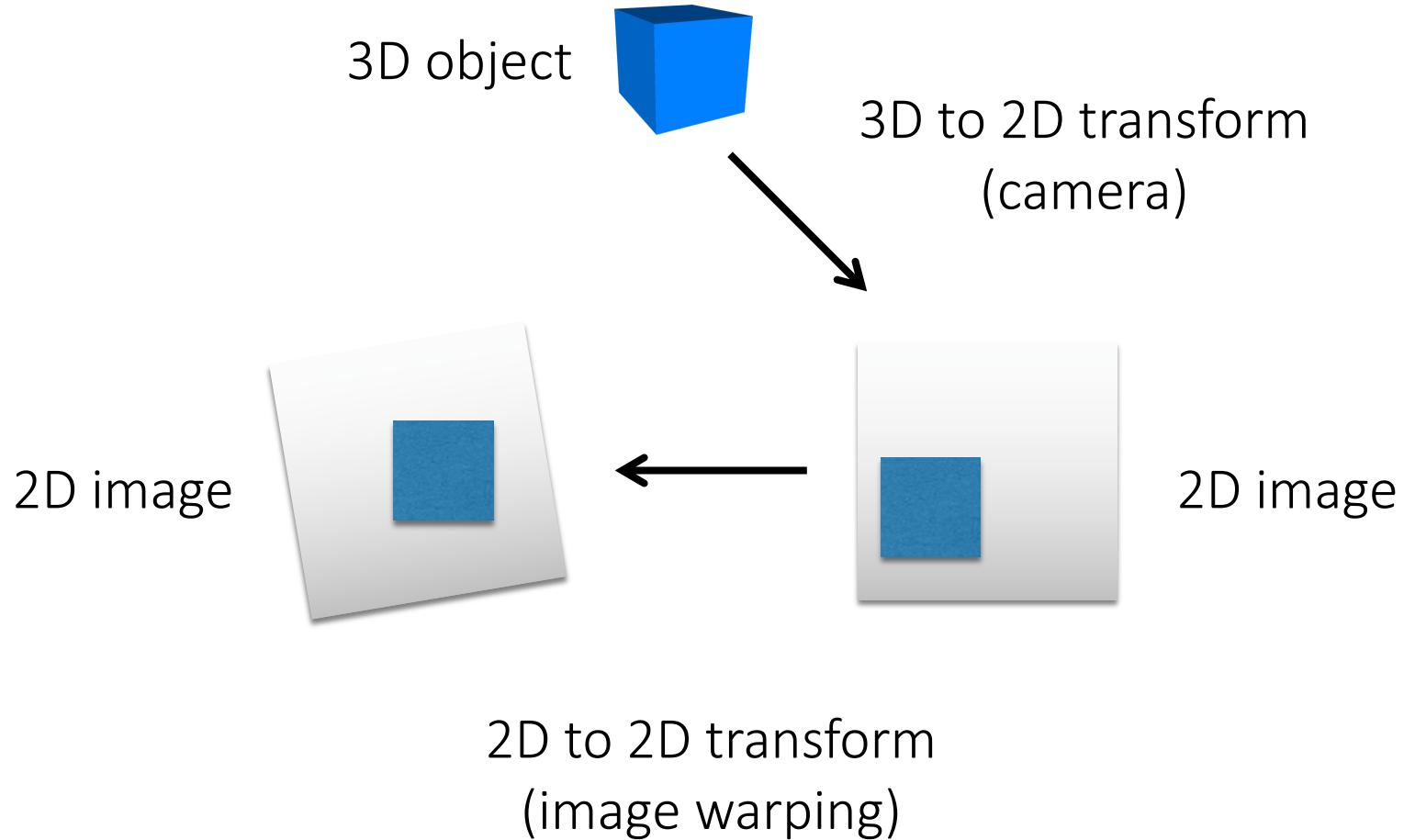
Camera matrix

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

A camera is a mapping from:

the 3D world

to:

a 2D image

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

homogeneous coordinates

2D image point

camera matrix

3D world point

The diagram illustrates the camera mapping equation $\mathbf{x} = \mathbf{P}\mathbf{X}$. It features two large, bold letters 'x' and 'X'. The letter 'x' is positioned above the equals sign, and the letter 'X' is positioned to its right. Above the letter 'x', the text "homogeneous coordinates" is written, with a black arrow pointing downwards towards the letter 'x'. Below the letter 'X', the text "3D world point" is written, with a black arrow pointing downwards towards the letter 'X'. To the left of the equals sign, the text "2D image point" is written, centered below the letter 'x'. To the right of the equals sign, the text "camera matrix" is written, centered between the letter 'x' and the letter 'X'.

What are the dimensions of each variable?

The camera as a coordinate transformation

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

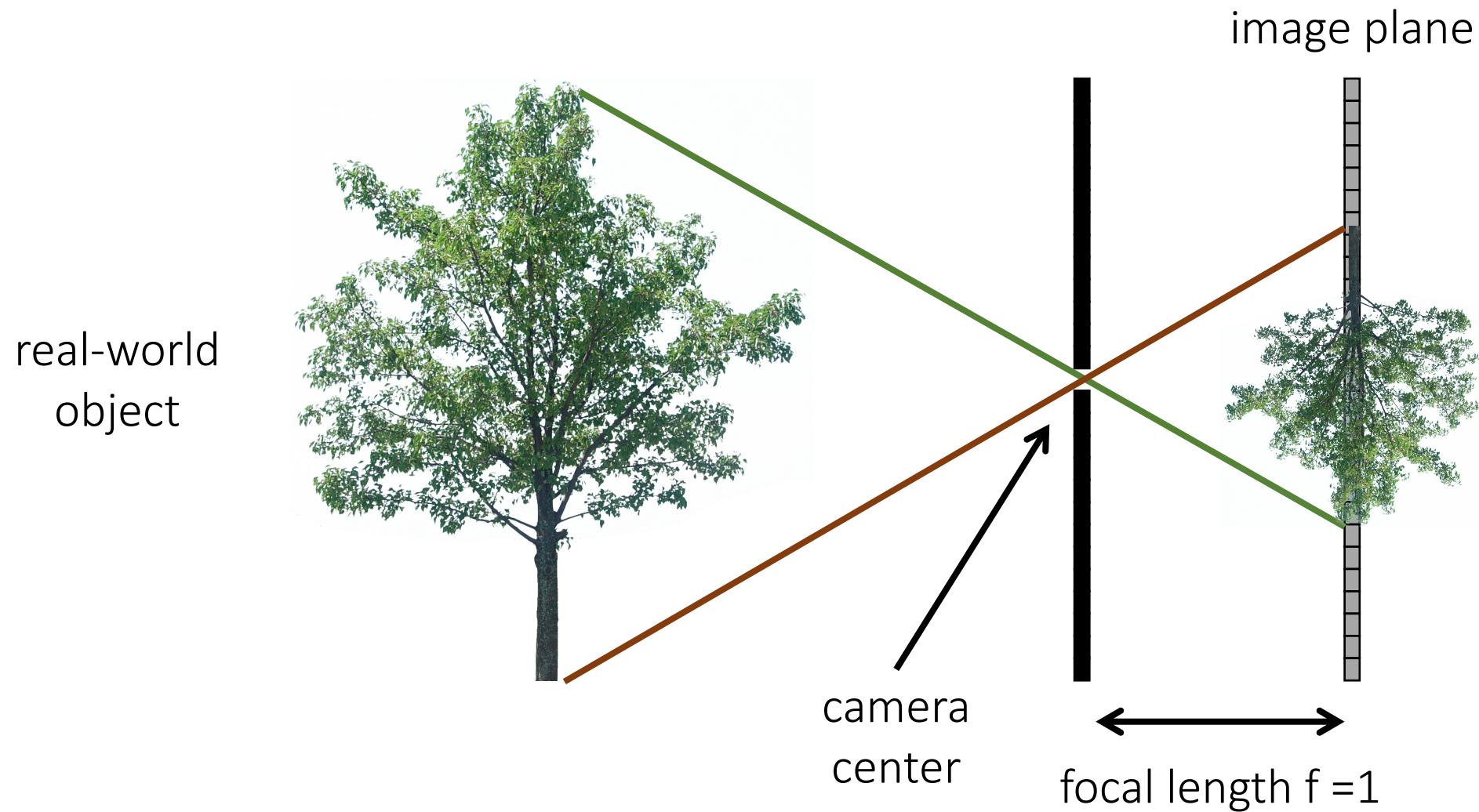
$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} p_1 & p_2 & p_3 & p_4 \\ p_5 & p_6 & p_7 & p_8 \\ p_9 & p_{10} & p_{11} & p_{12} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

homogeneous
image coordinates
 3×1

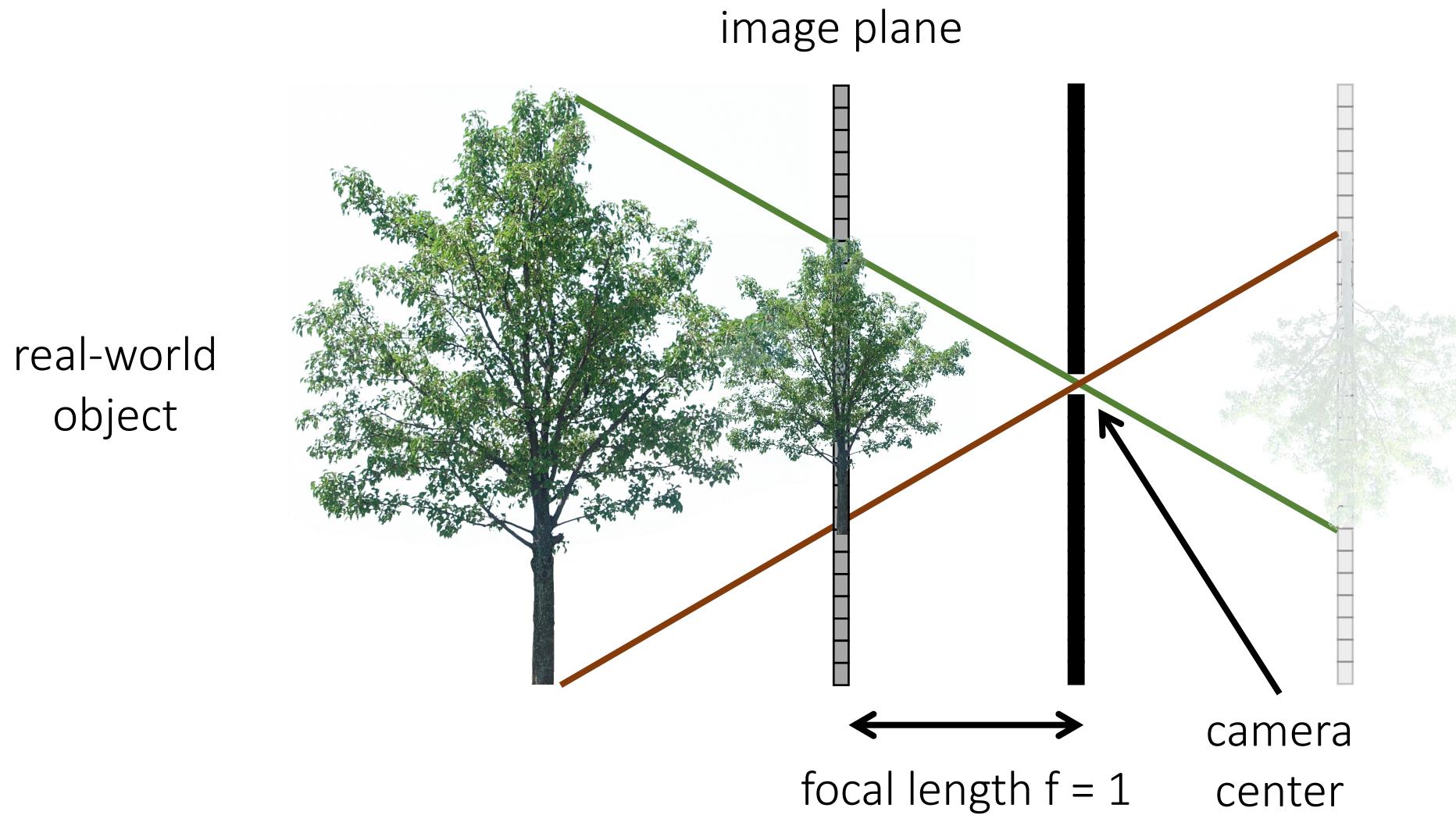
camera
matrix
 3×4

homogeneous
world coordinates
 4×1

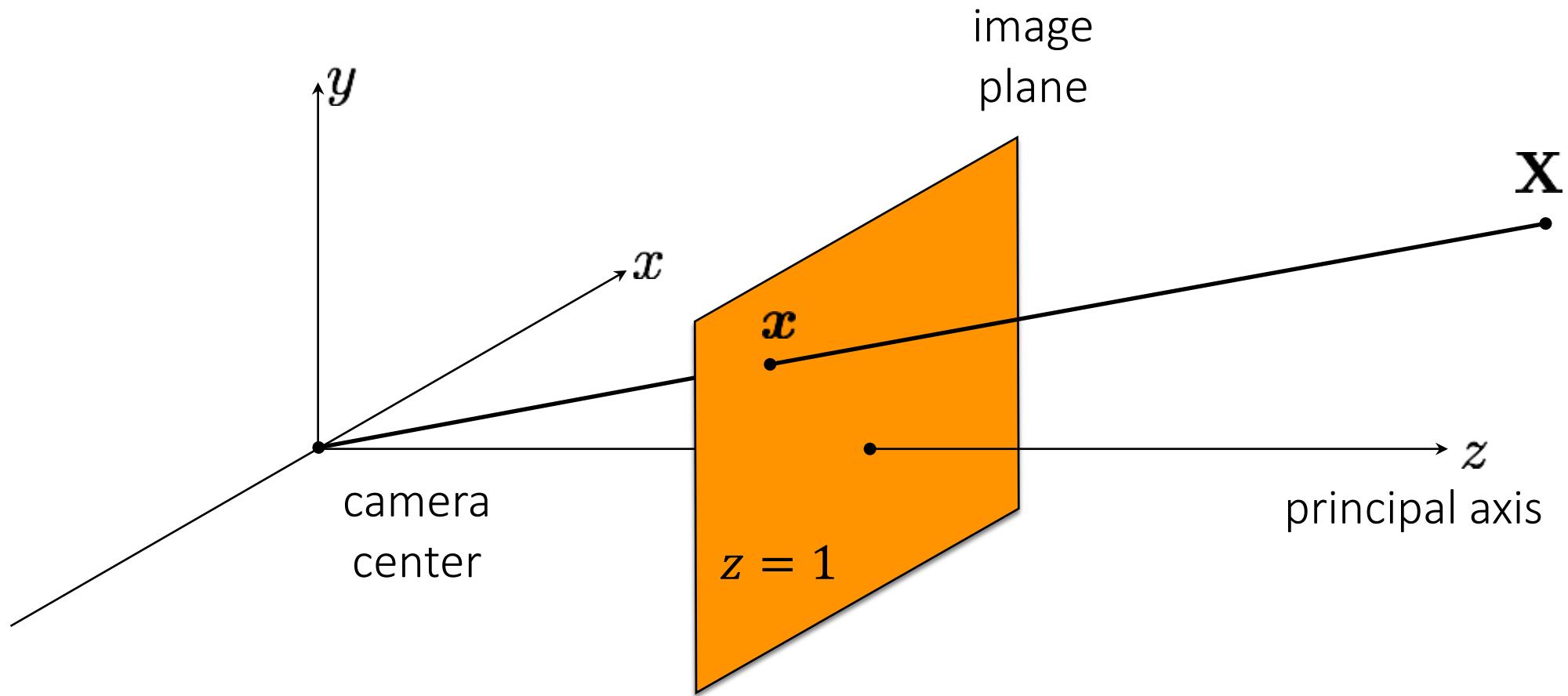
The pinhole camera



The (rearranged) pinhole camera

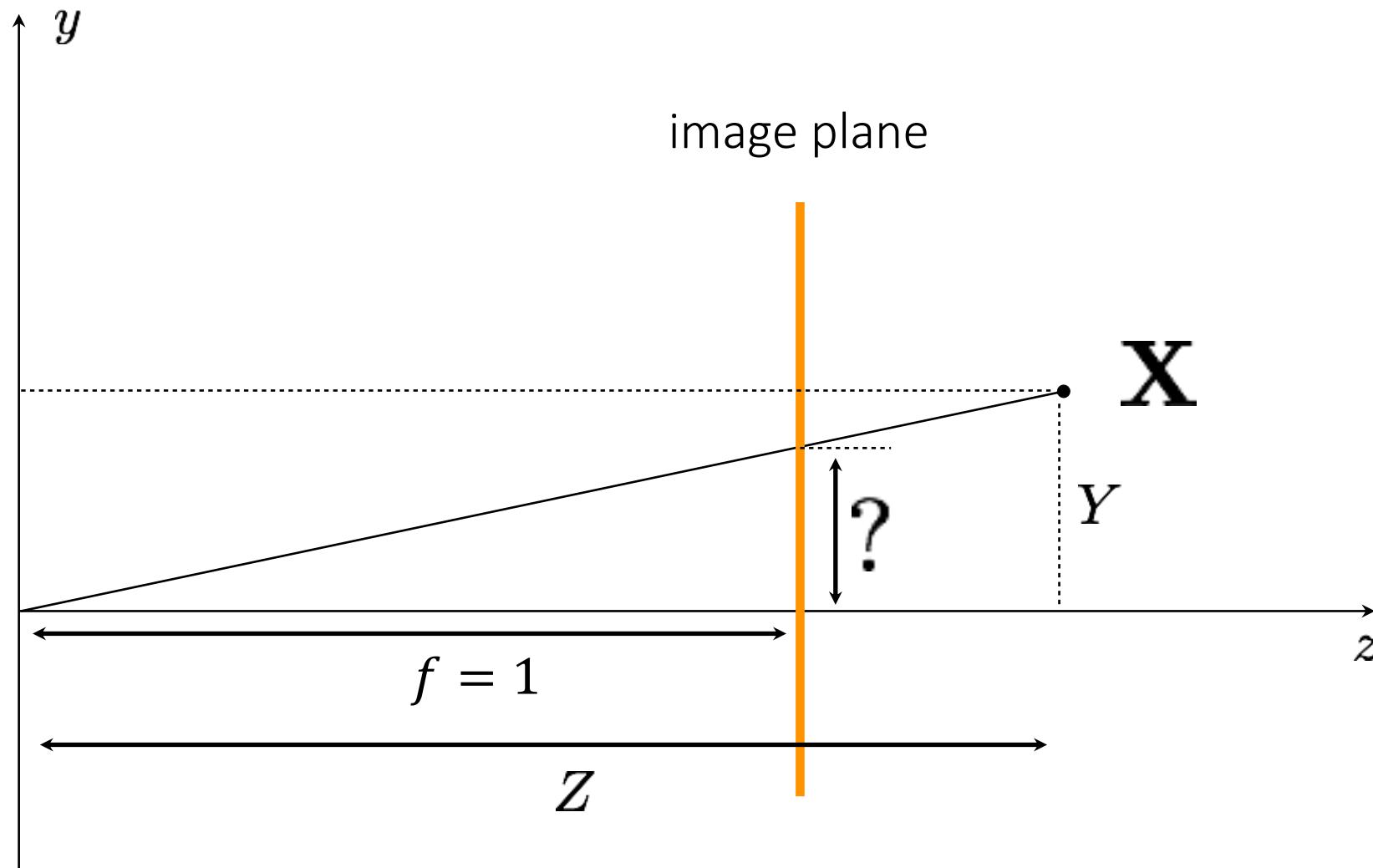


The (rearranged) pinhole camera



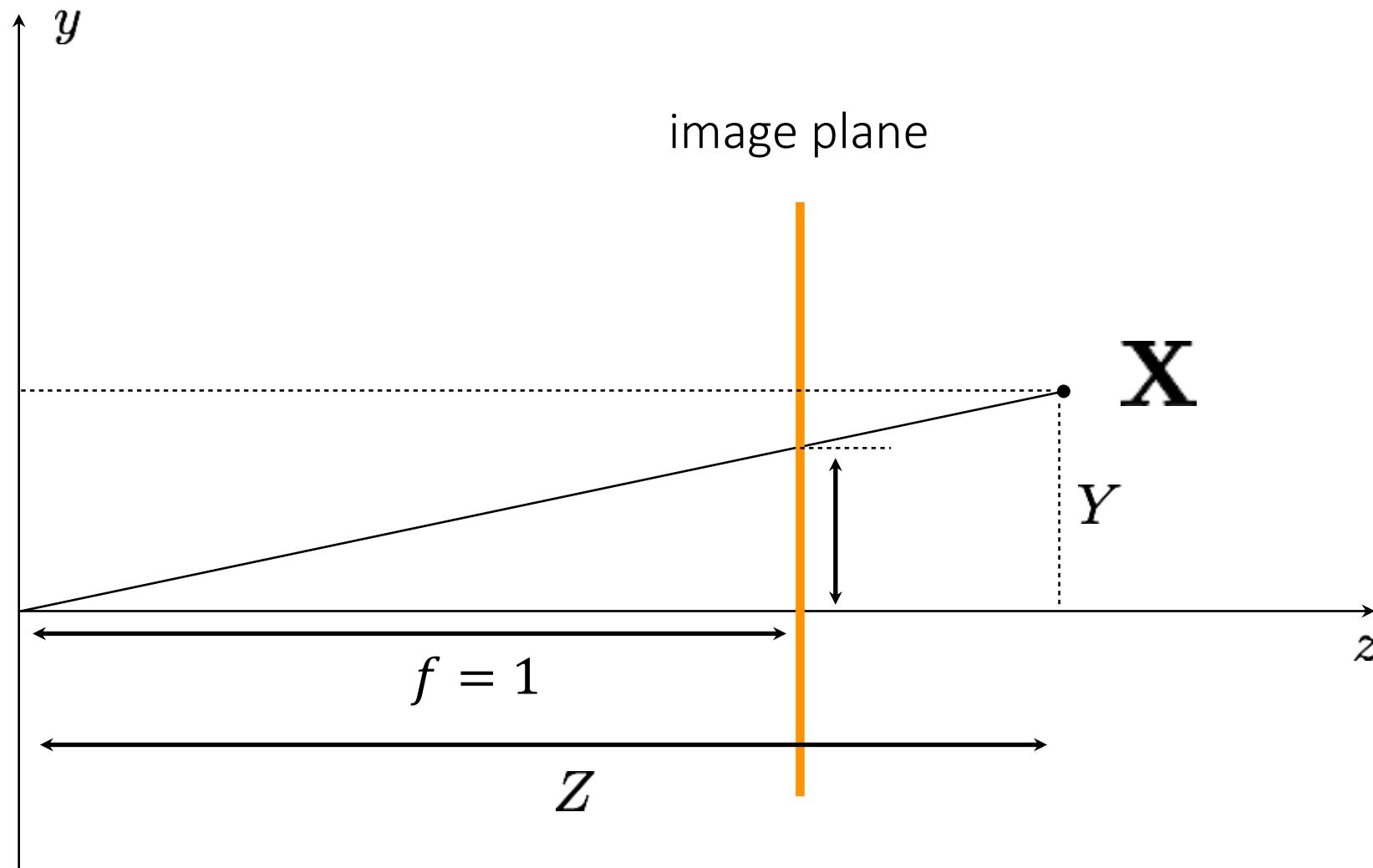
What is the equation for image coordinate x in terms of \mathbf{X} ?

The 2D view of the (rearranged) pinhole camera



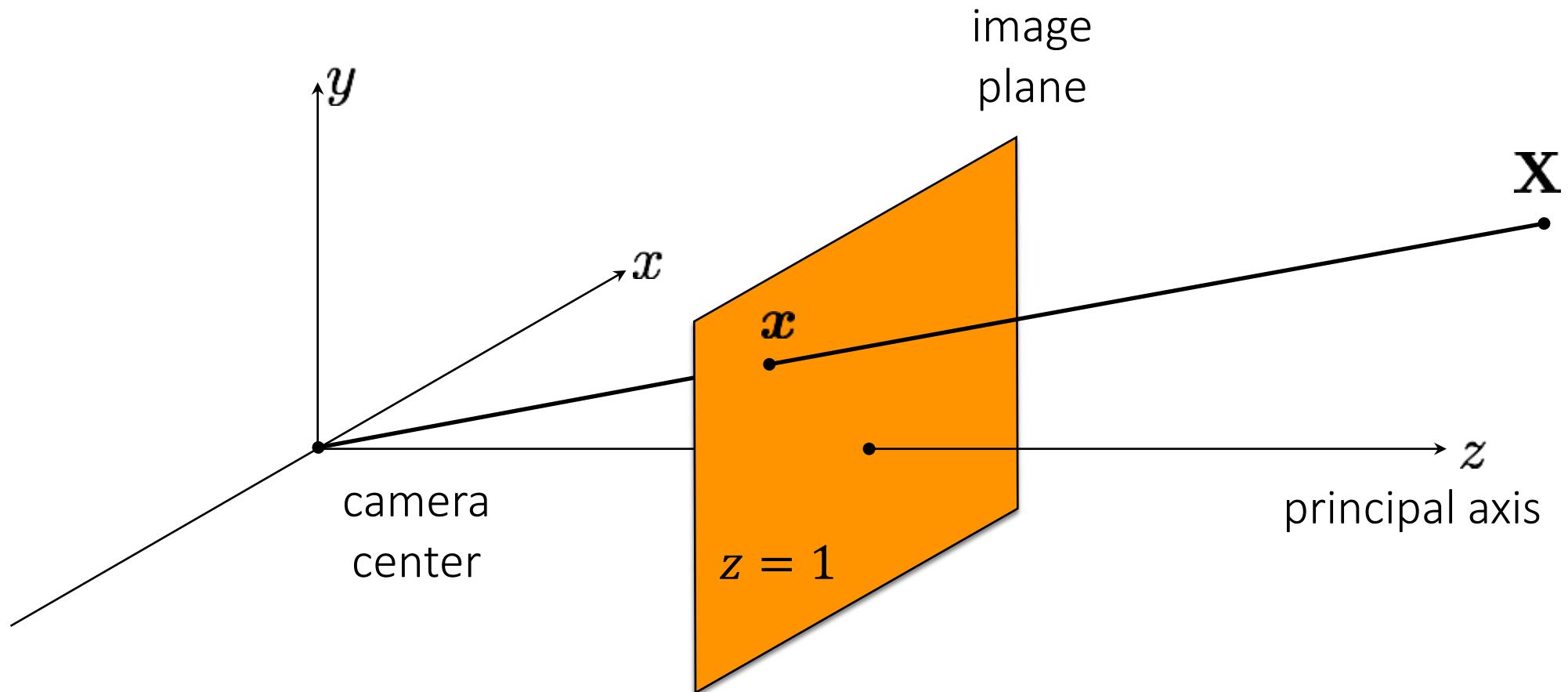
What is the equation for image coordinate x in terms of X ?

The 2D view of the (rearranged) pinhole camera



$$[X \quad Y \quad Z]^T \rightarrow [X/Z \quad Y/Z]$$

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

Relationship from similar triangles:

$$[X \quad Y \quad Z]^T \rightarrow [X/Z \quad Y/Z]$$

General camera model in *homogeneous coordinates*:

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

What does the pinhole camera projection look like?

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

The pinhole camera matrix

Relationship from similar triangles:

$$[X \quad Y \quad Z]^T \rightarrow [X/Z \quad Y/Z]$$

General camera model *in homogeneous coordinates*:

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

What does the pinhole camera 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

Relationship from similar triangles:

$$[X \quad Y \quad Z]^T \rightarrow [X/Z \quad Y/Z]$$

General camera model *in homogeneous coordinates*:

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

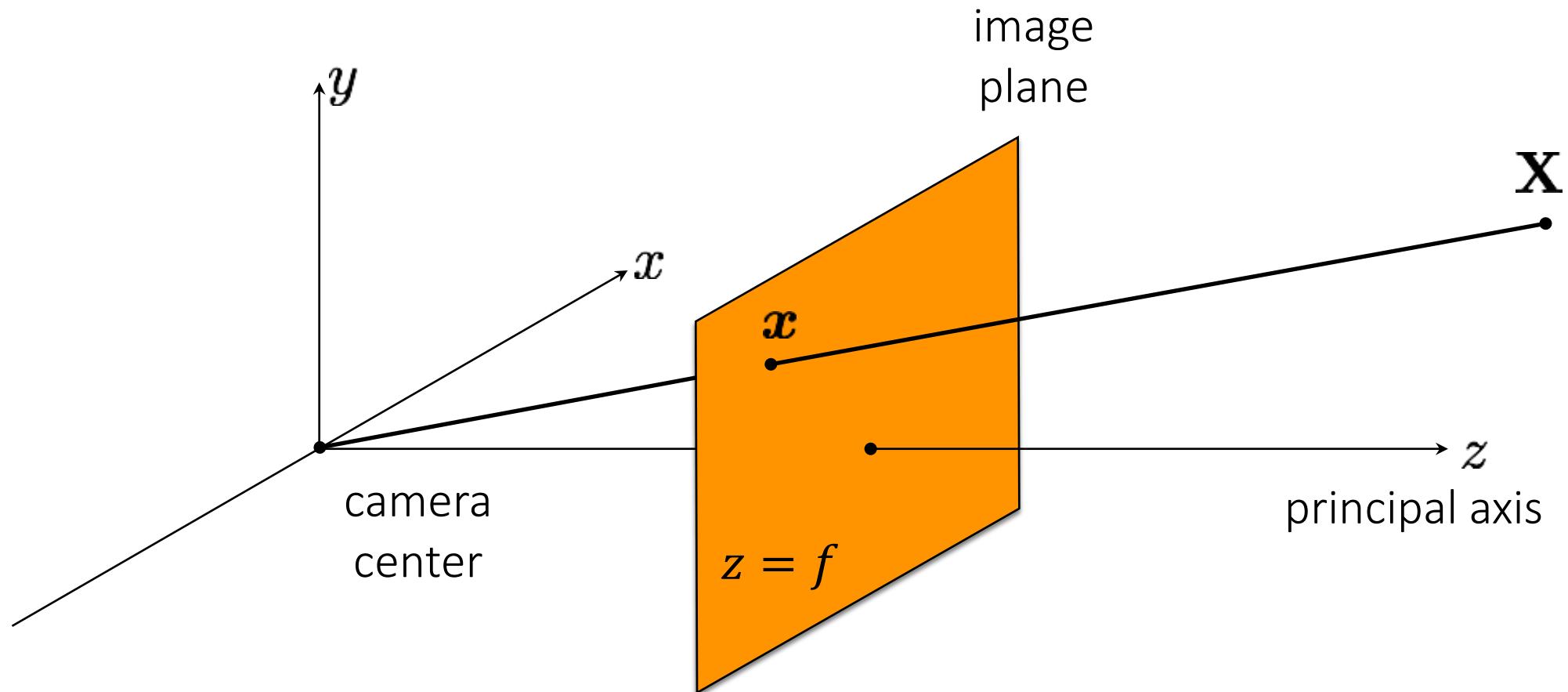
What does the pinhole camera projection look like?

The perspective
projection matrix

$$\mathbf{P} = \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array} \right] = [\mathbf{I} \quad | \quad \mathbf{0}]$$

alternative way to write
the same thing

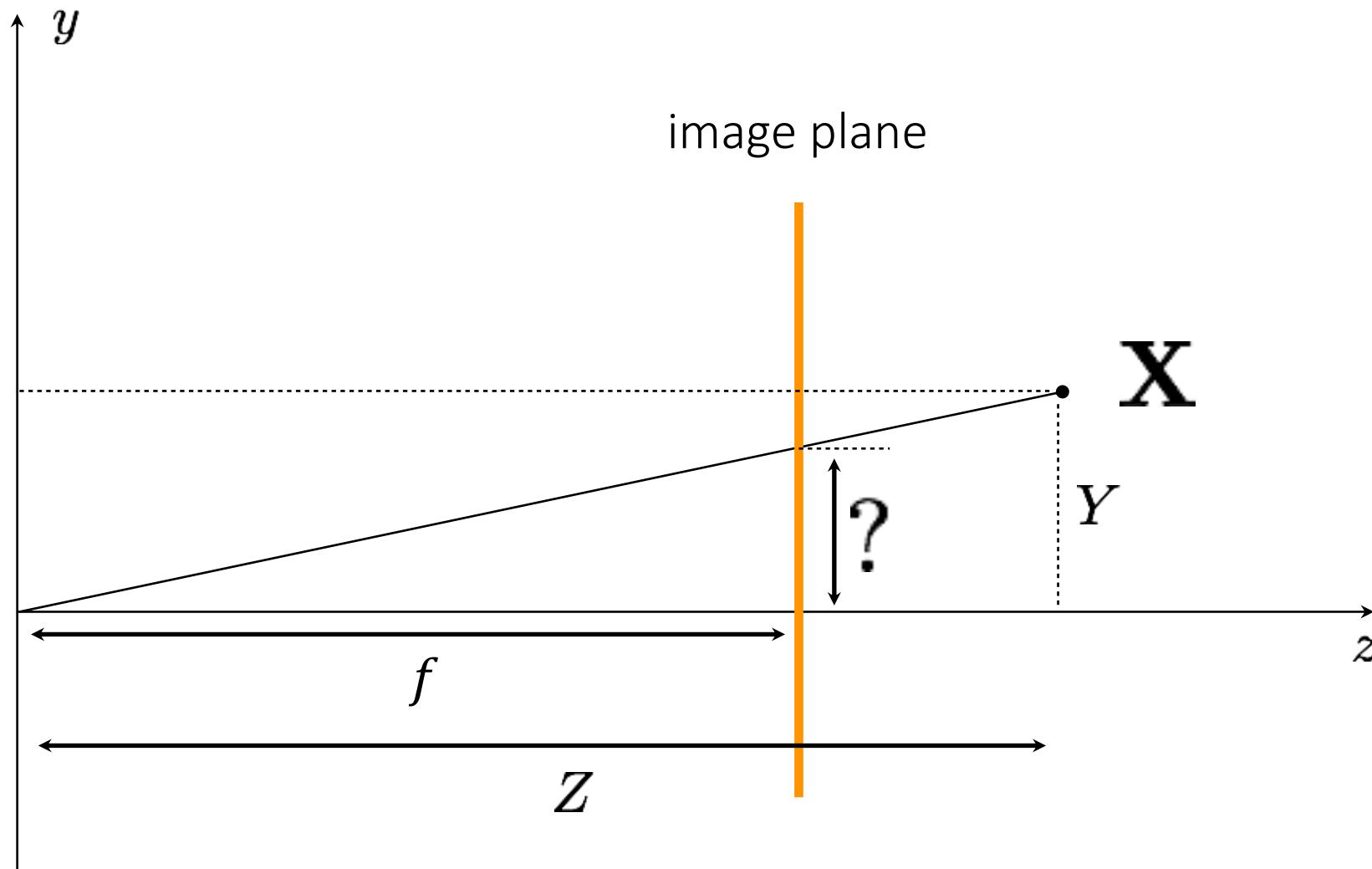
More general case: arbitrary focal length



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

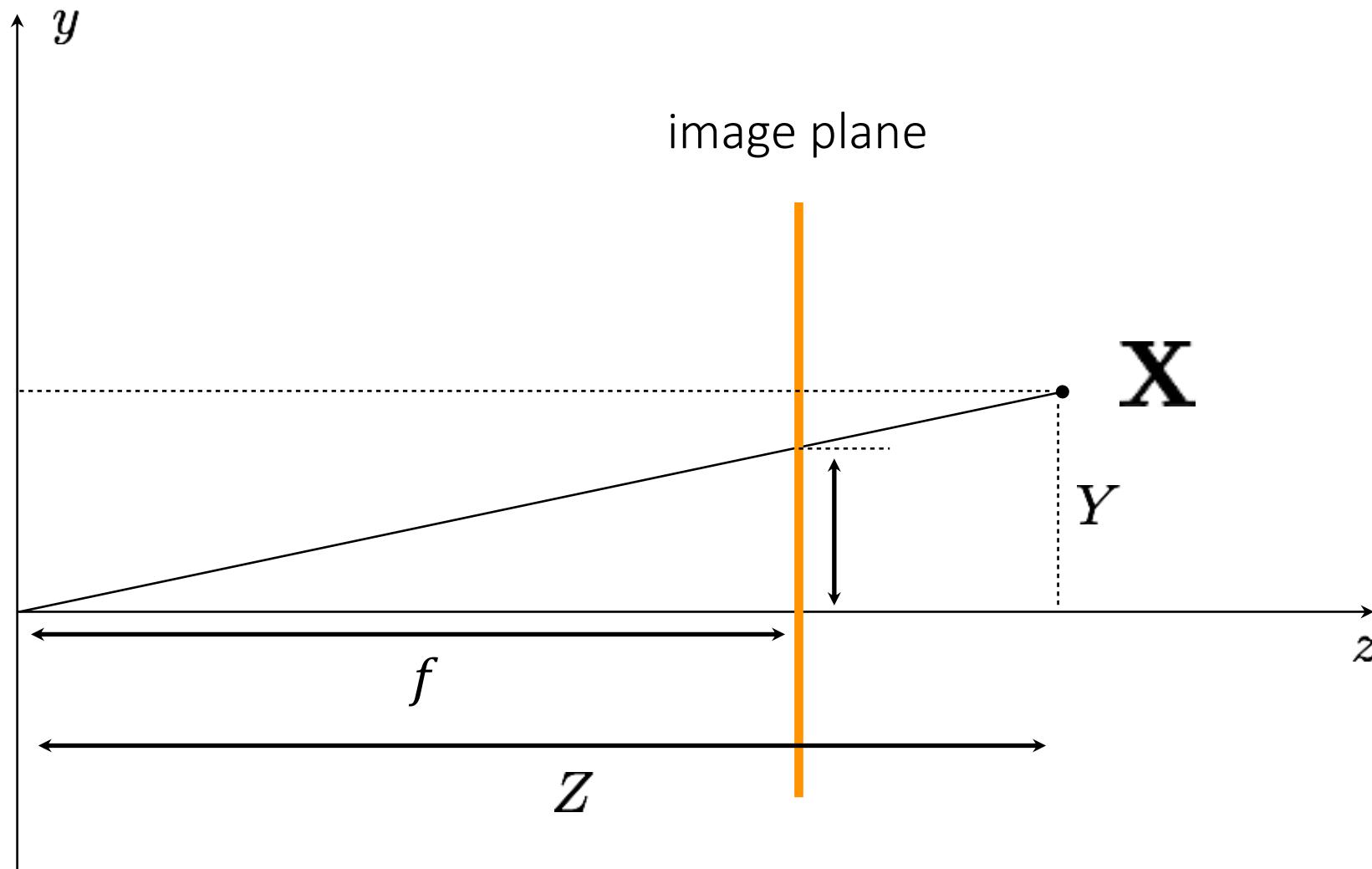
$$\mathbf{x} = \mathbf{P}\mathbf{X}$$

More general (2D) case: arbitrary focal length



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

More general (2D) case: arbitrary focal length



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

The pinhole camera matrix for arbitrary focal length

Relationship from similar triangles:

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

General camera model *in homogeneous coordinates*:

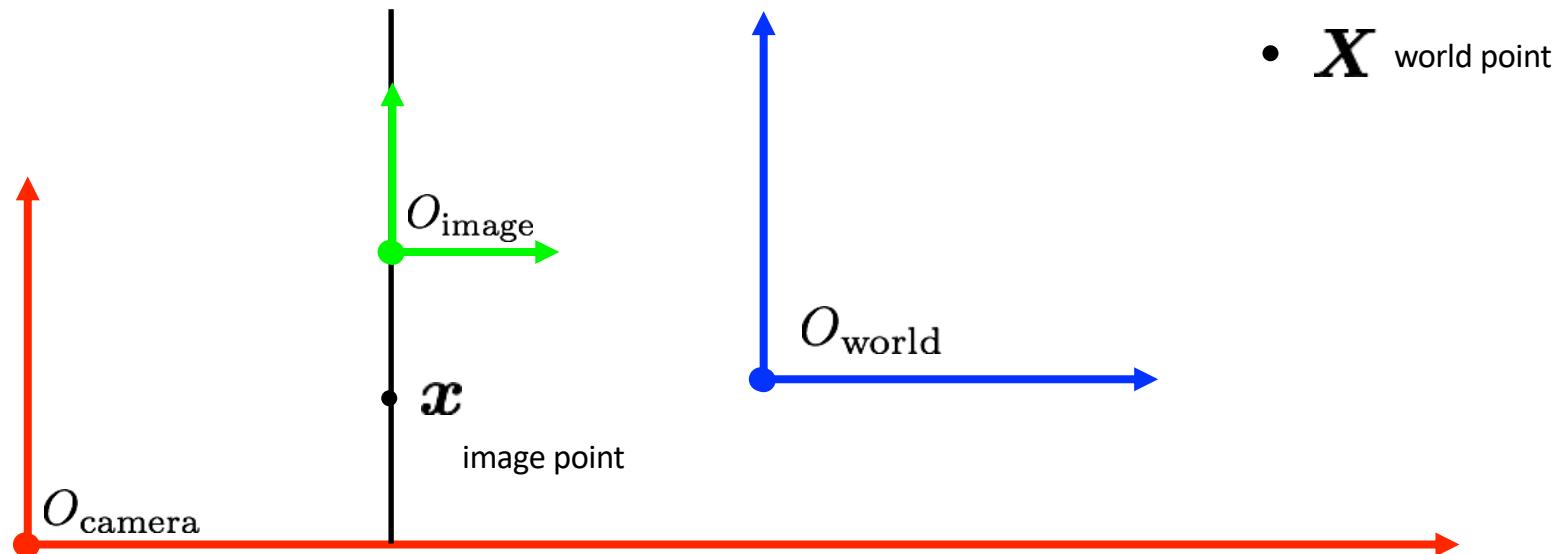
$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} p_1 & p_2 & p_3 & p_4 \\ p_5 & p_6 & p_7 & p_8 \\ p_9 & p_{10} & p_{11} & p_{12} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

What does the pinhole camera projection look like?

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

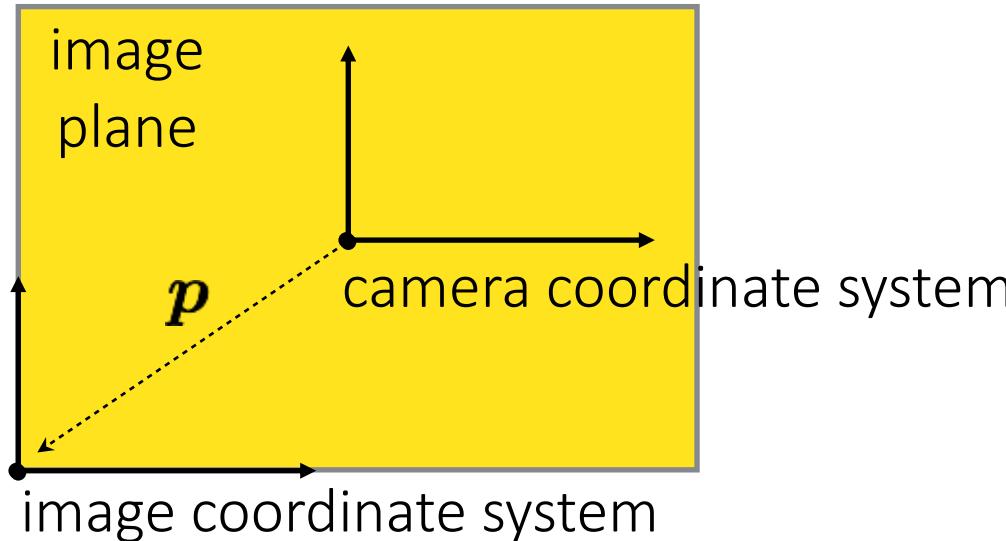
Generalizing the camera matrix

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



Generalizing the camera matrix

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

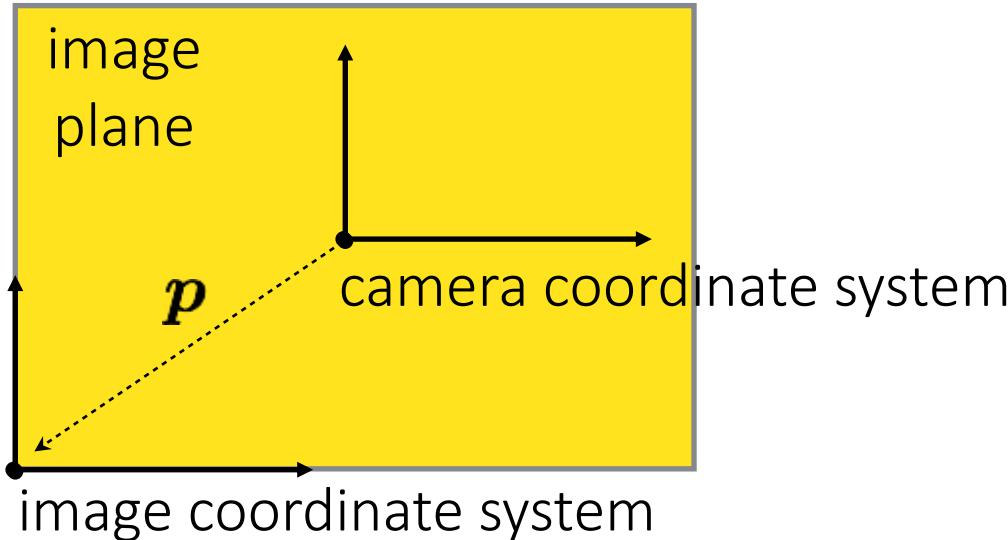


How does the camera matrix change?

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

Generalizing the camera matrix

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



How does the camera matrix change?

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

shift vector
transforming
camera origin to
image origin

Camera matrix decomposition

We can decompose the camera matrix like this:

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

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

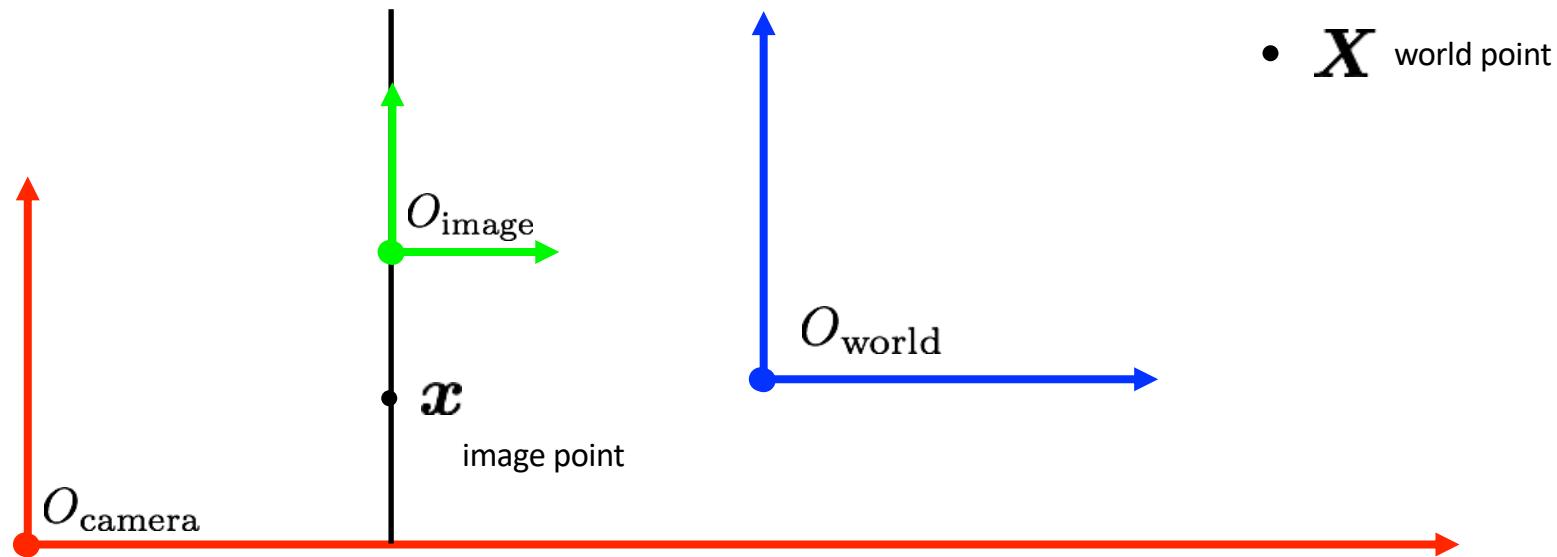
(homogeneous) perspective projection
from 3D to 2D, assuming image plane at
 $z = 1$ and shared camera/image origin

Also written as: $\mathbf{P} = \mathbf{K}[\mathbf{I}|0]$

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

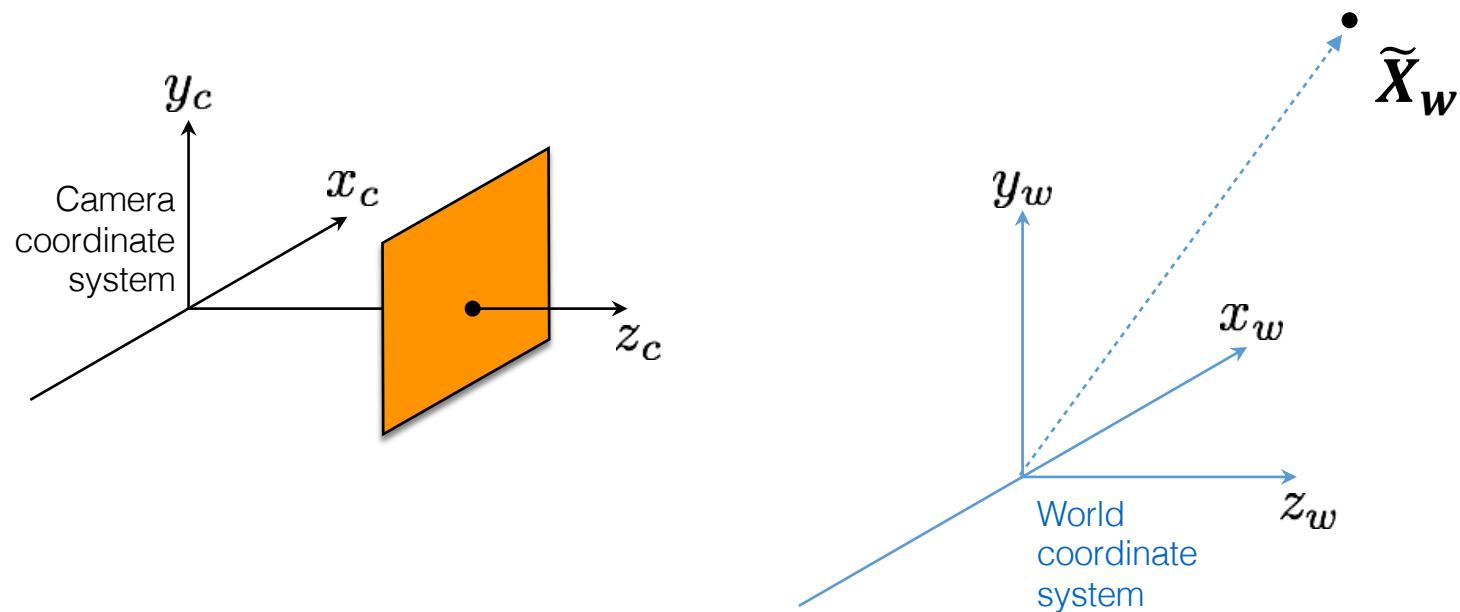
Generalizing the camera matrix

In general, there are *three*, generally different, coordinate systems.



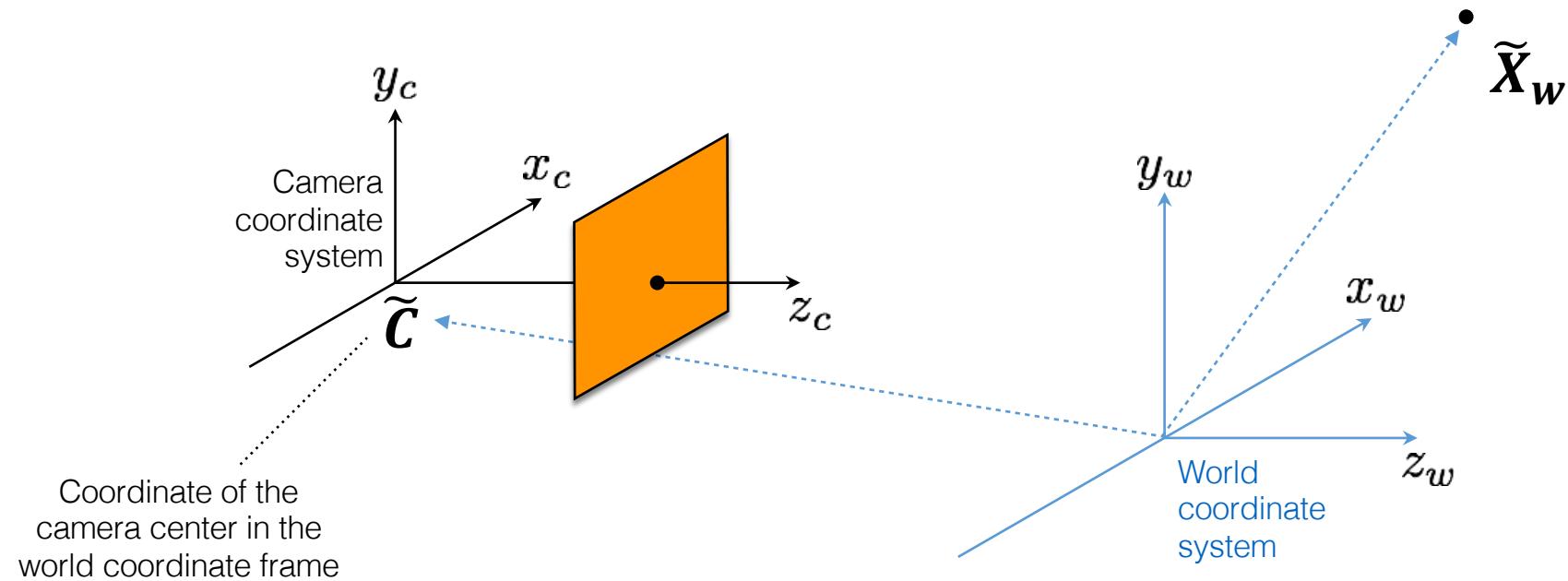
We need to know the transformations between them.

World-to-camera coordinate system transformation

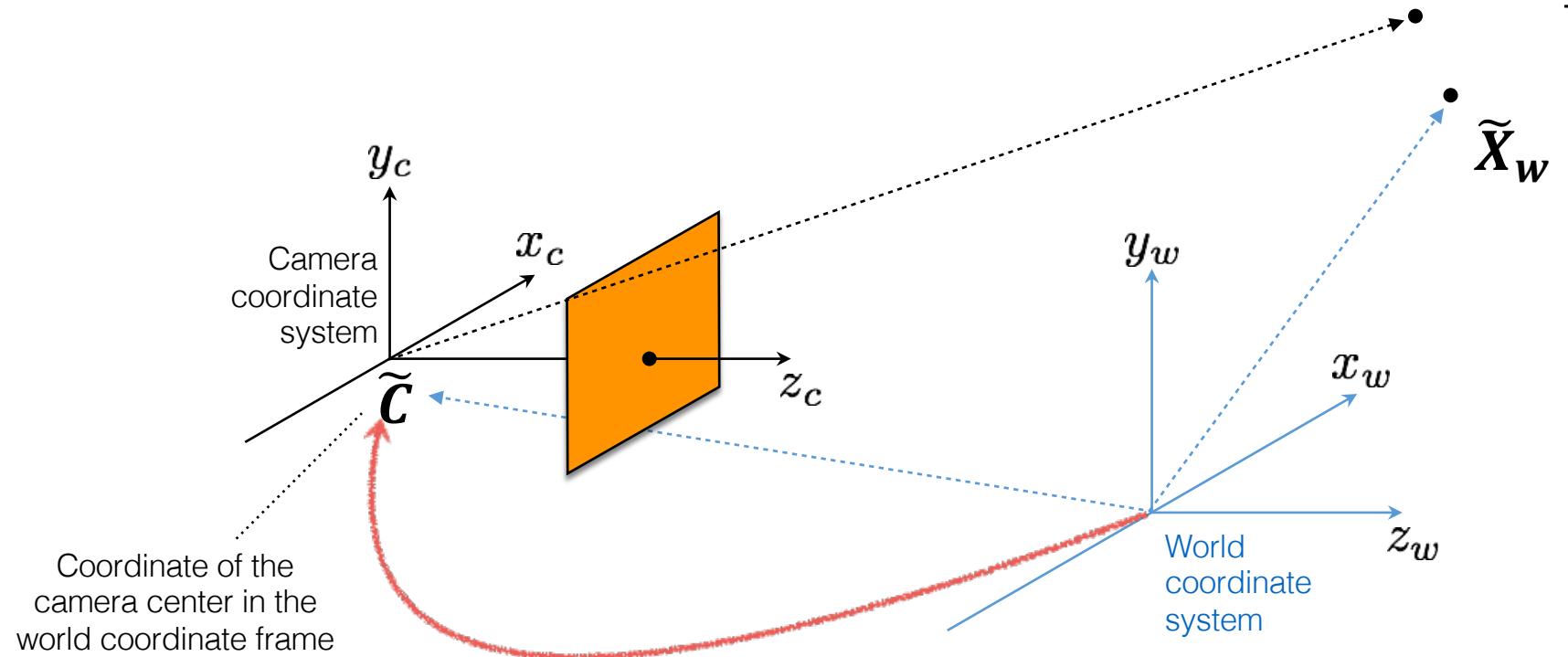


tilde means
heterogeneous
coordinates

World-to-camera coordinate system transformation



World-to-camera coordinate system transformation

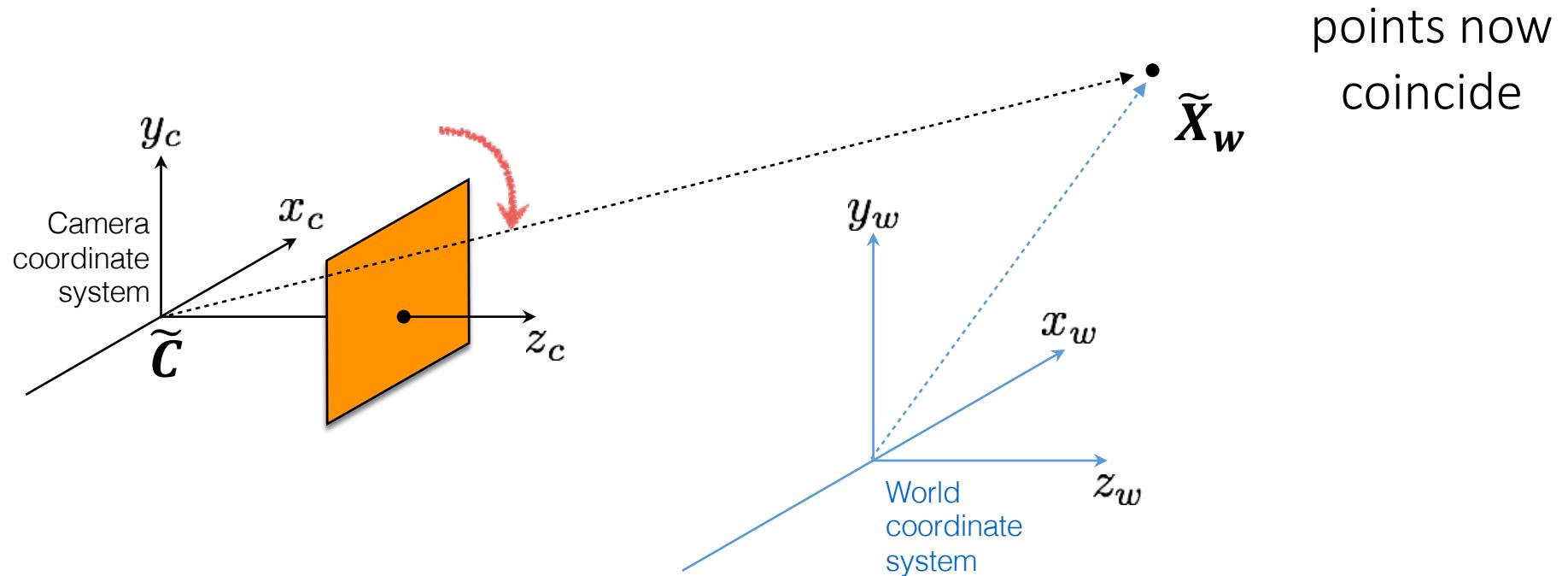


Why aren't
the points
aligned?

$$(\tilde{X}_w - \tilde{C})$$

translate

World-to-camera coordinate system transformation



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

rotate translate

Modeling the 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 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:

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R} & -\mathbf{R}\mathbf{C} \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad \text{or} \quad \mathbf{x}_c = \begin{bmatrix} \mathbf{R} & -\mathbf{R}\tilde{\mathbf{C}} \\ \mathbf{0} & 1 \end{bmatrix} \mathbf{x}_w$$

Incorporating the transform in the camera matrix

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

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

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:

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

intrinsic parameters (3×3):

correspond to camera
internals (image-to-image
transformation)

perspective projection (3×4):

maps 3D to 2D points
(camera-to-image
transformation)

extrinsic parameters (4×4):

correspond to camera
externals (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:

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

intrinsic parameters (3×3):
correspond to camera internals
(sensor not at $f = 1$ and origin shift)

extrinsic parameters (3×4):
correspond to camera externals
(world-to-image transformation)

General pinhole camera matrix

We can decompose the camera matrix like this:

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

(translate first then rotate)

Another way to write the mapping:

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

where $\mathbf{t} = -\mathbf{R}\mathbf{C}$

(rotate first then translate)

General pinhole camera matrix

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

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

intrinsic extrinsic
parameters parameters

$$\mathbf{R} = \left[\begin{array}{ccc} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \\ r_7 & r_8 & r_9 \end{array} \right] \quad \mathbf{t} = \left[\begin{array}{c} t_1 \\ t_2 \\ t_3 \end{array} \right]$$

3D rotation 3D translation

Recap

What is the size and meaning of each term in the camera matrix?

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

The diagram consists of a mathematical equation $\mathbf{P} = \mathbf{K}\mathbf{R}[\mathbf{I}] - \mathbf{C}$ centered on the page. Below the equation, there are four question marks: one under \mathbf{K} , one under \mathbf{R} , one under $[\mathbf{I}]$, and one under \mathbf{C} . Four black arrows originate from these question marks and point diagonally upwards towards their respective terms in the equation above.

Recap

What is the size and meaning of each term in the camera matrix?

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

3x3
intrinsics

Recap

What is the size and meaning of each term in the camera matrix?

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

3x3
intrinsics 3x3
3D rotation ? ?

Recap

What is the size and meaning of each term in the camera matrix?

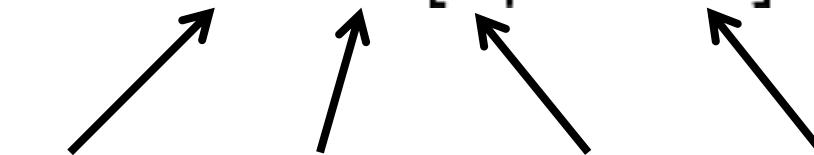
$$\mathbf{P} = \mathbf{K}\mathbf{R}[\mathbf{I}] - \mathbf{C}$$

The diagram illustrates the camera matrix equation $\mathbf{P} = \mathbf{K}\mathbf{R}[\mathbf{I}] - \mathbf{C}$. Four arrows point from the terms \mathbf{K} , \mathbf{R} , $[\mathbf{I}]$, and \mathbf{C} to their corresponding dimensions below the equation. The term \mathbf{K} is labeled "3x3 intrinsics". The term \mathbf{R} is labeled "3x3 3D rotation". The term $[\mathbf{I}]$ is labeled "3x3 identity". The term \mathbf{C} is labeled with a question mark "?".

Recap

What is the size and meaning of each term in the camera matrix?

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



3x3

3x3

3x3

3x1

intrinsics 3D rotation identity 3D translation

Quiz

The camera matrix relates what two quantities?

Quiz

The camera matrix relates what two quantities?

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

homogeneous 3D points to 2D image points

Quiz

The camera matrix relates what two quantities?

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

homogeneous 3D points to 2D image points

The camera matrix can be decomposed into?

Quiz

The camera matrix relates what two quantities?

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

homogeneous 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

More general camera matrices

The following is the standard camera matrix we saw.

$$\mathbf{P} = \begin{bmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix} \quad [\mathbf{R} \quad \vdots \quad -\mathbf{RC}]$$

More general camera matrices

CCD camera: pixels may not be square.

$$\mathbf{P} = \begin{bmatrix} \alpha_x & 0 & p_x \\ 0 & \alpha_y & p_y \\ 0 & 0 & 1 \end{bmatrix} \quad \left[\begin{array}{c|c} \mathbf{R} & -\mathbf{RC} \end{array} \right]$$

How many degrees of freedom?

More general camera matrices

CCD camera: pixels may not be square.

$$\mathbf{P} = \begin{bmatrix} \alpha_x & 0 & p_x \\ 0 & \alpha_y & p_y \\ 0 & 0 & 1 \end{bmatrix} \quad \left[\begin{array}{c|c} \mathbf{R} & -\mathbf{RC} \end{array} \right]$$

How many degrees of freedom?

10 DOF

More general camera matrices

Finite projective camera: sensor be skewed.

$$\mathbf{P} = \begin{bmatrix} \alpha_x & s & p_x \\ 0 & \alpha_y & p_y \\ 0 & 0 & 1 \end{bmatrix} \left[\begin{array}{c|c} \mathbf{R} & -\mathbf{RC} \end{array} \right]$$

How many degrees of freedom?

More general camera matrices

Finite projective camera: sensor be skewed.

$$\mathbf{P} = \begin{bmatrix} \alpha_x & s & p_x \\ 0 & \alpha_y & p_y \\ 0 & 0 & 1 \end{bmatrix} \left[\begin{array}{c|c} \mathbf{R} & -\mathbf{RC} \end{array} \right]$$

How many degrees of freedom?

11 DOF