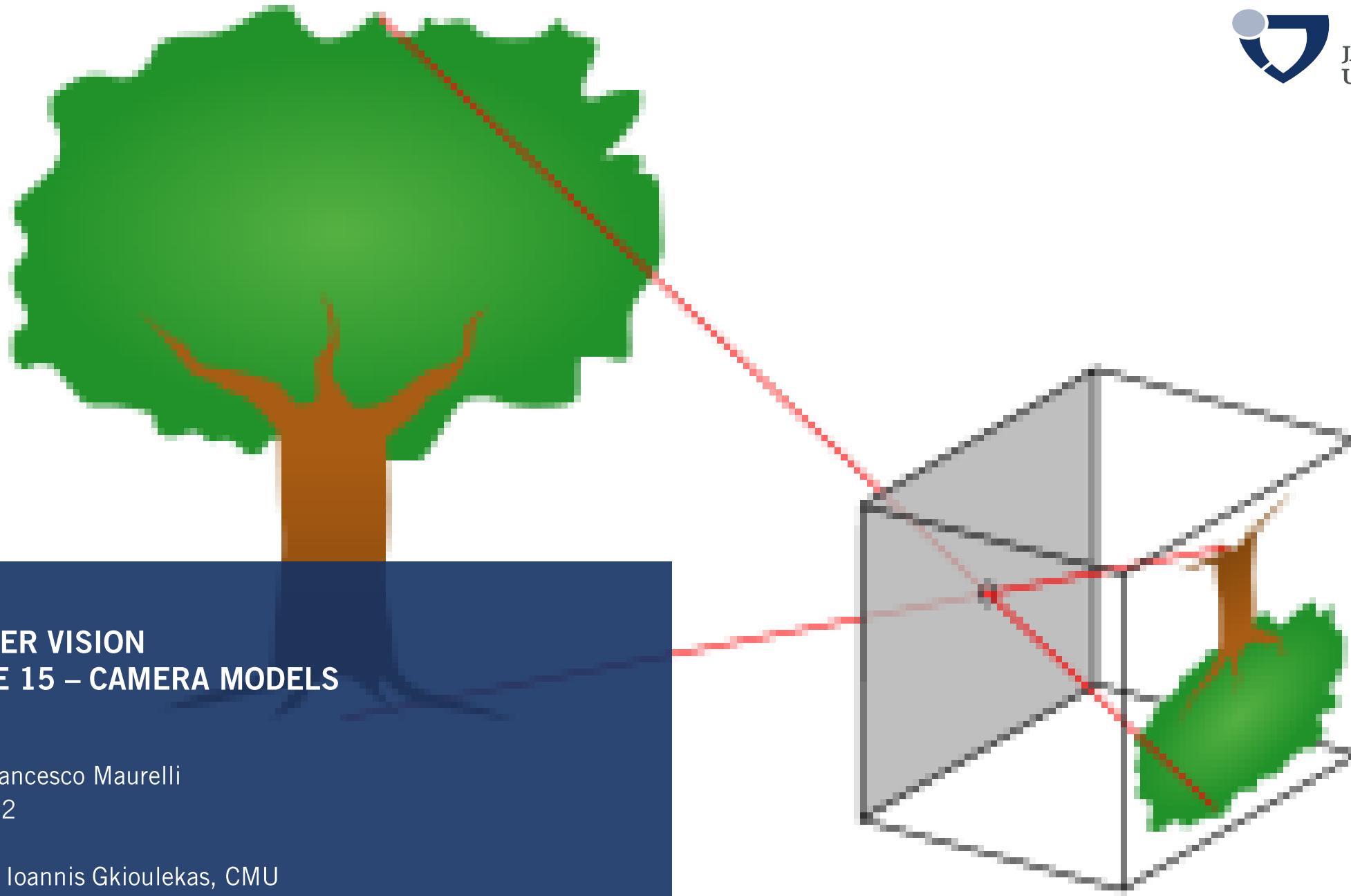


COMPUTER VISION LECTURE 15 – CAMERA MODELS

Prof. Dr. Francesco Maurelli
2019-10-22

Courtesy of Ioannis Gkioulekas, CMU
Fei-Fei Li, Stanford

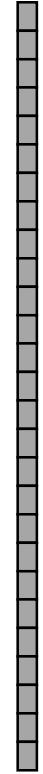


Overview of today's lecture

- Pinhole camera.
- Accidental pinholes.
- Camera matrix.
- Perspective.
- Other camera models.
- Pose estimation.

... 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

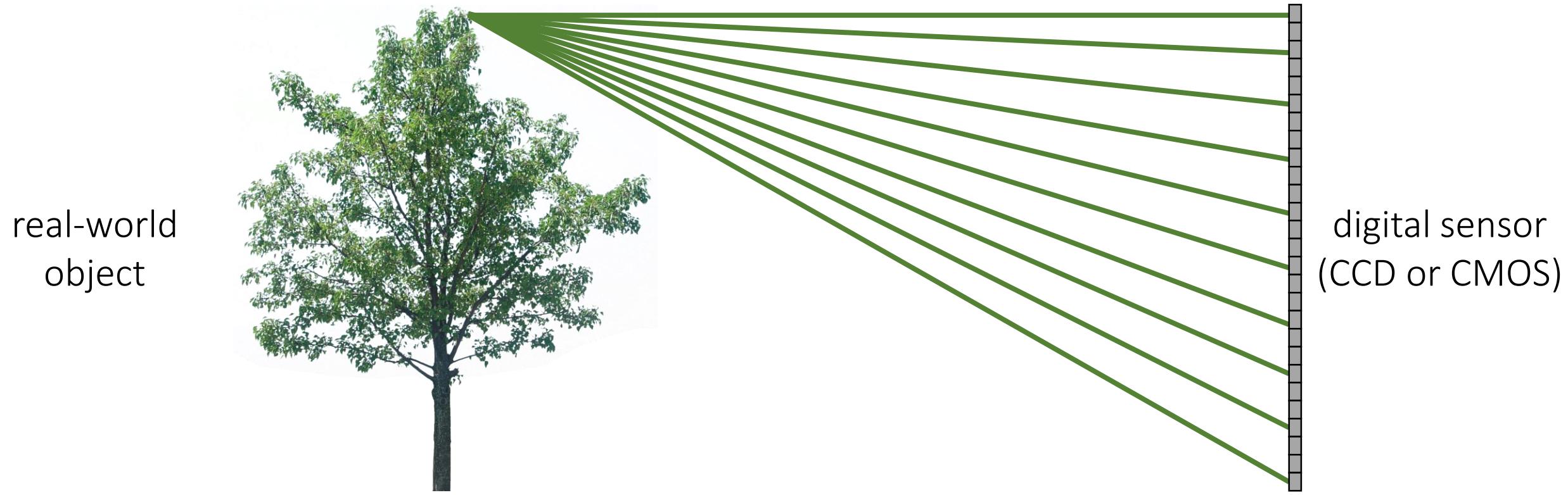
real-world
object



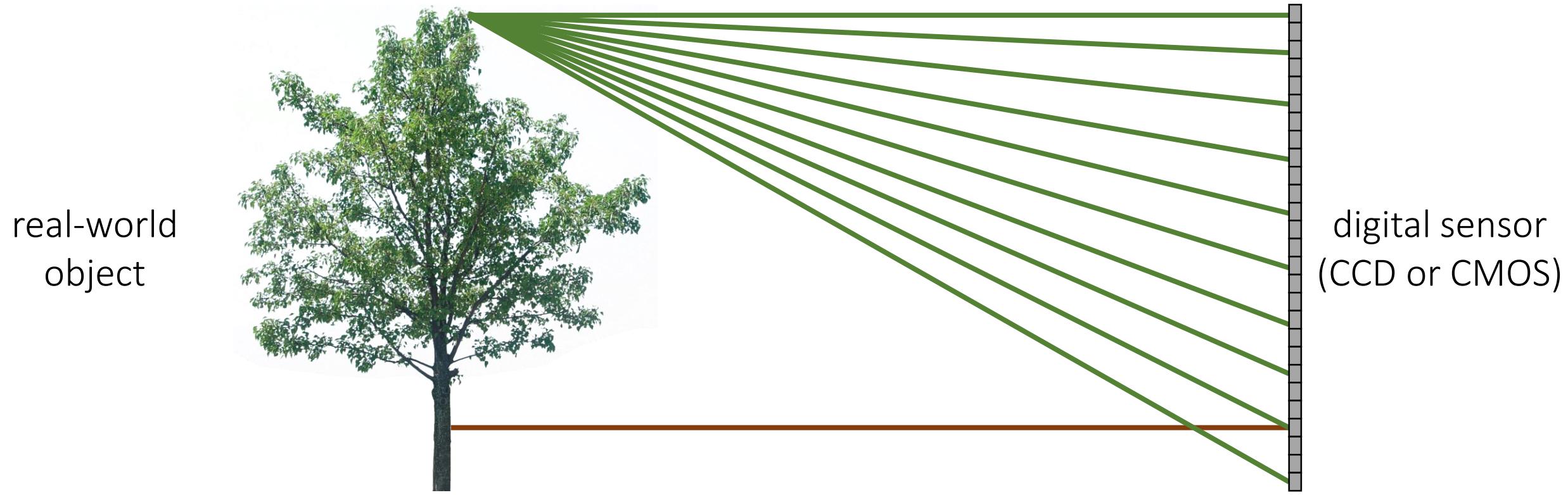
digital sensor
(CCD or CMOS)



Bare-sensor imaging

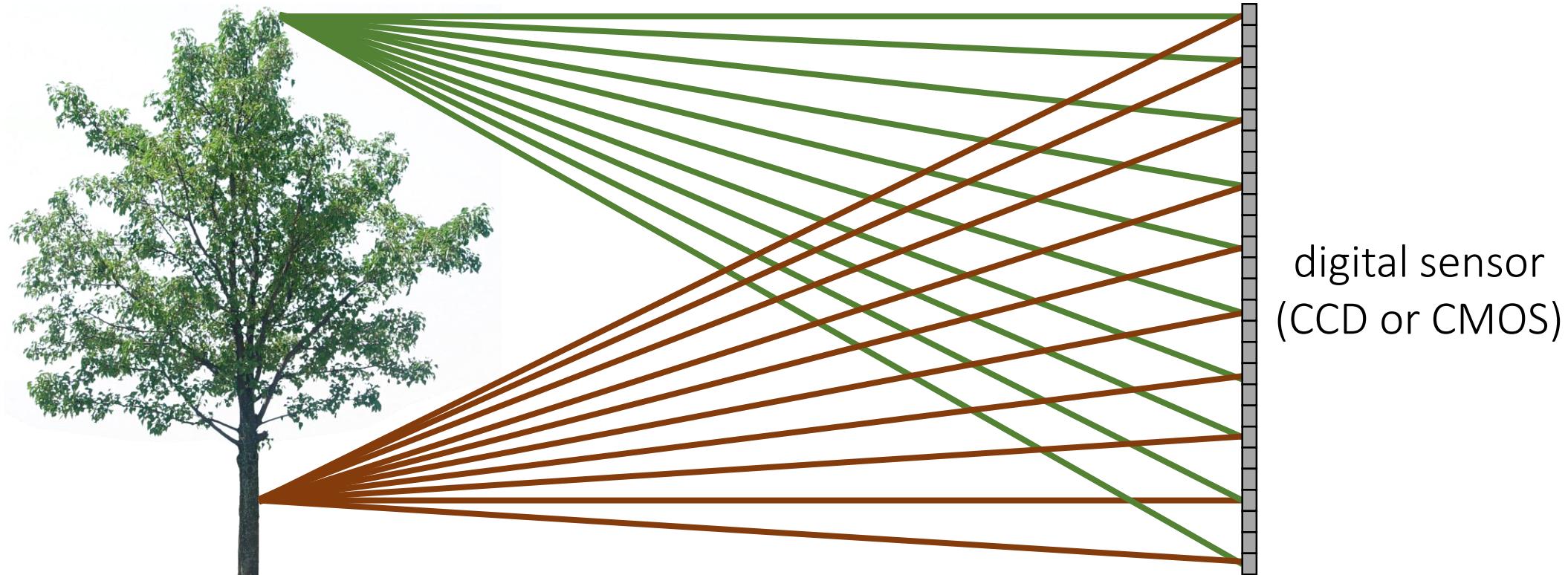


Bare-sensor imaging



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



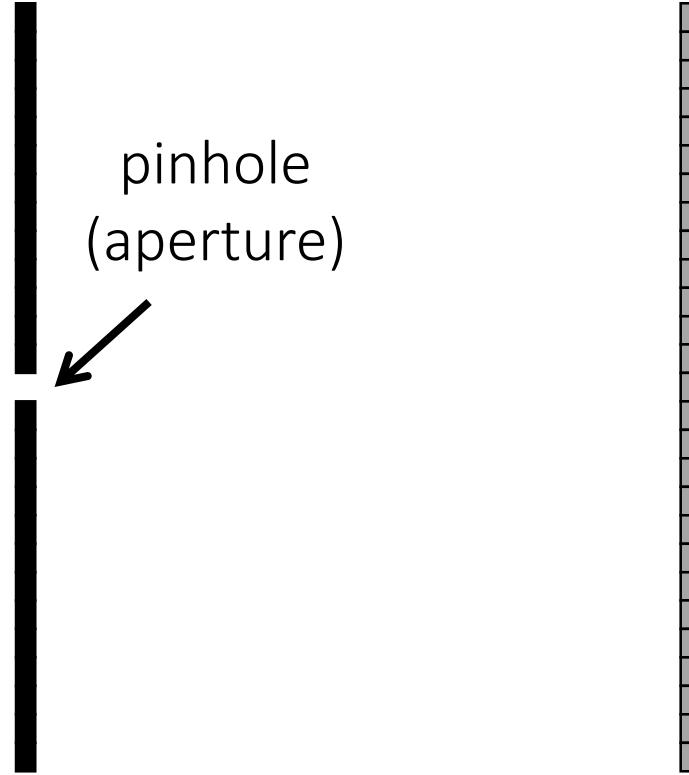
All scene points contribute to all sensor pixels

Let's add something to this scene

real-world
object

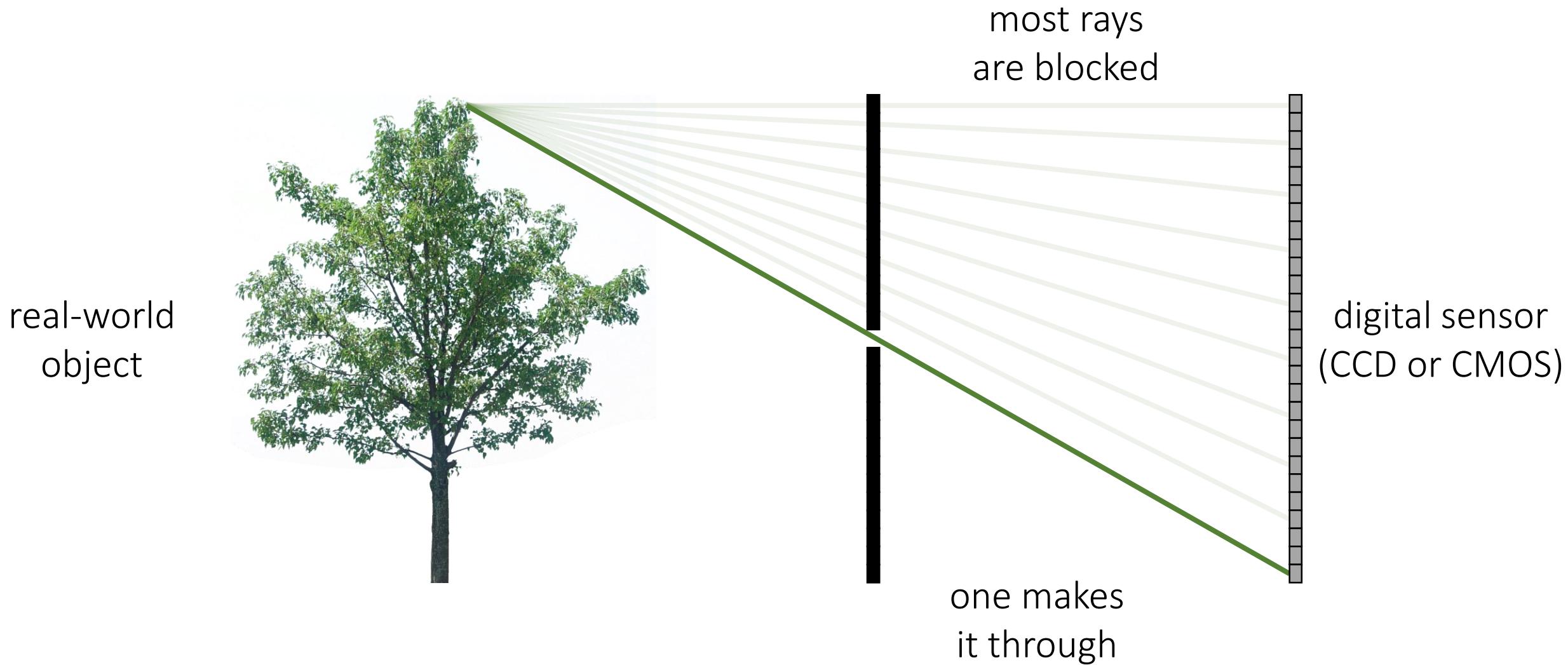


barrier (diaphragm)

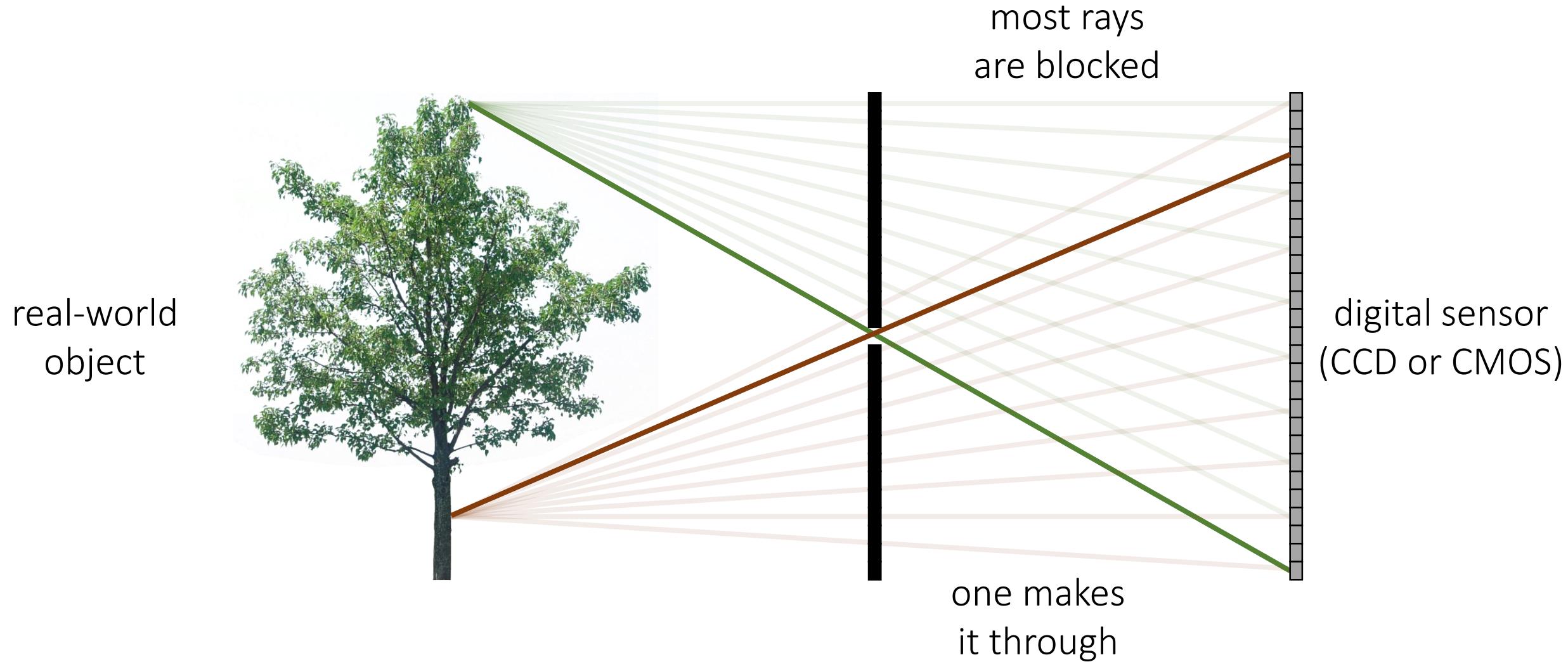


What would an image taken like this look like?

Pinhole imaging

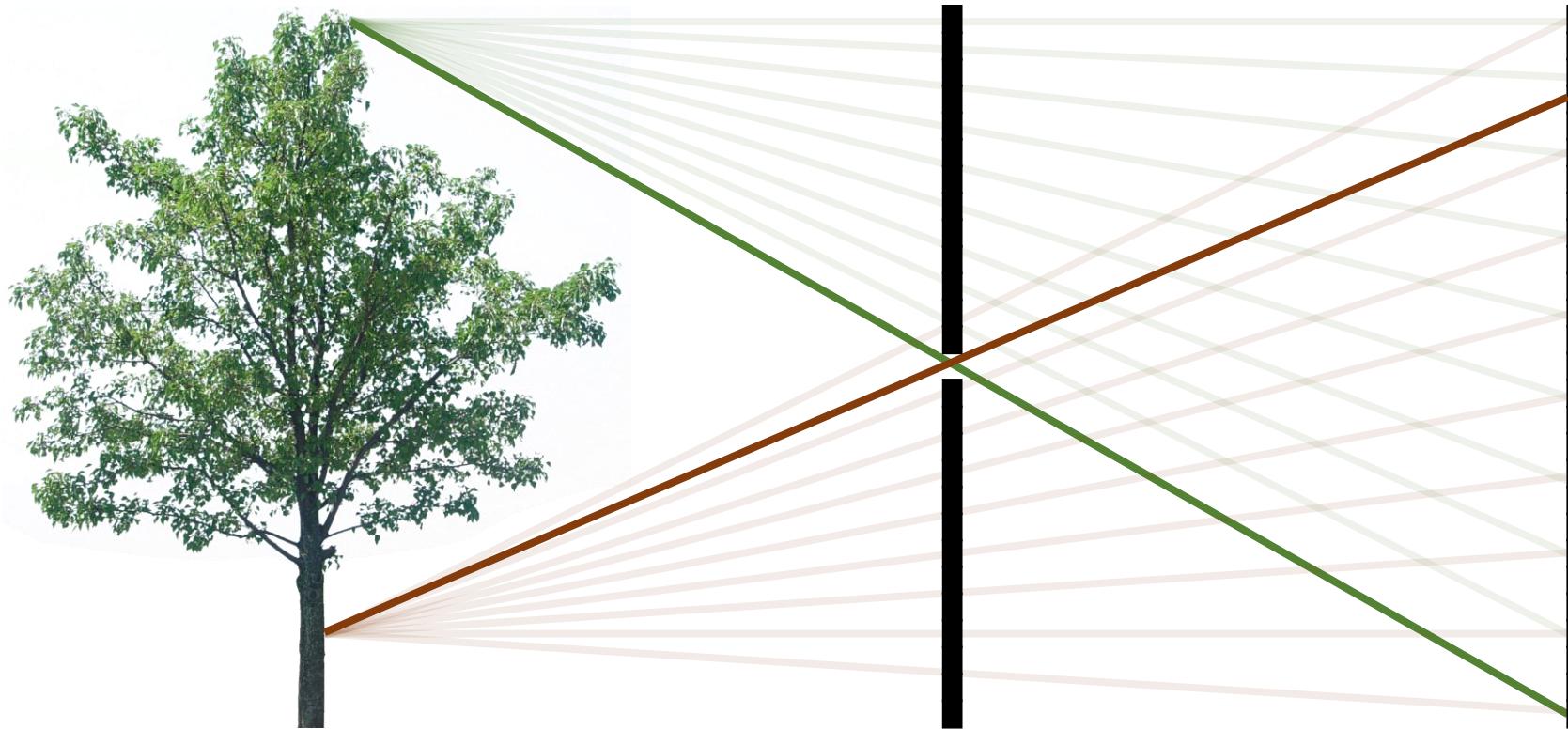


Pinhole imaging



Pinhole imaging

real-world
object



Each scene point contributes to only one sensor pixel

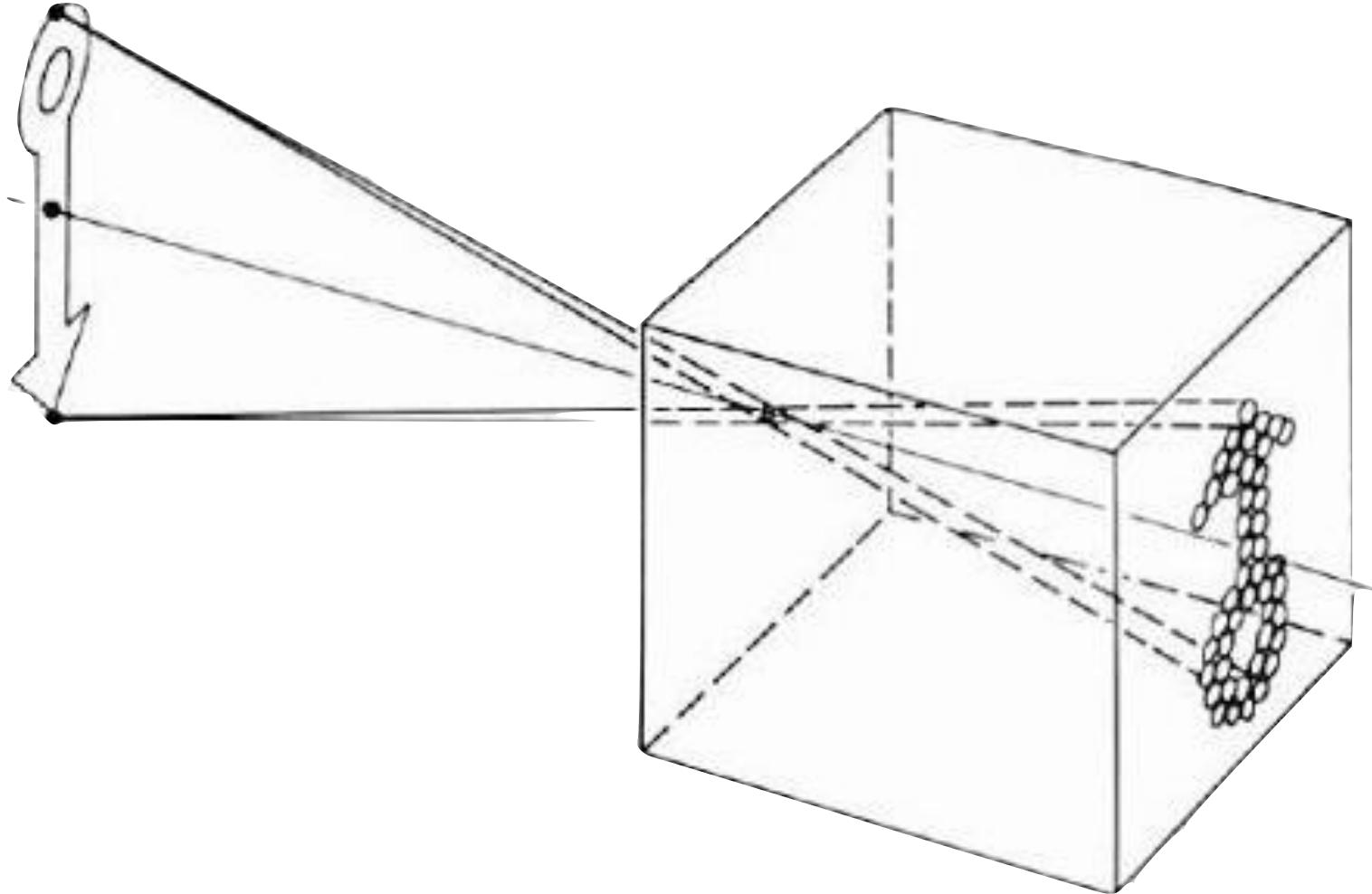
What does the
image on the
sensor look like?

Pinhole imaging



Pinhole camera

Pinhole camera a.k.a. camera obscura



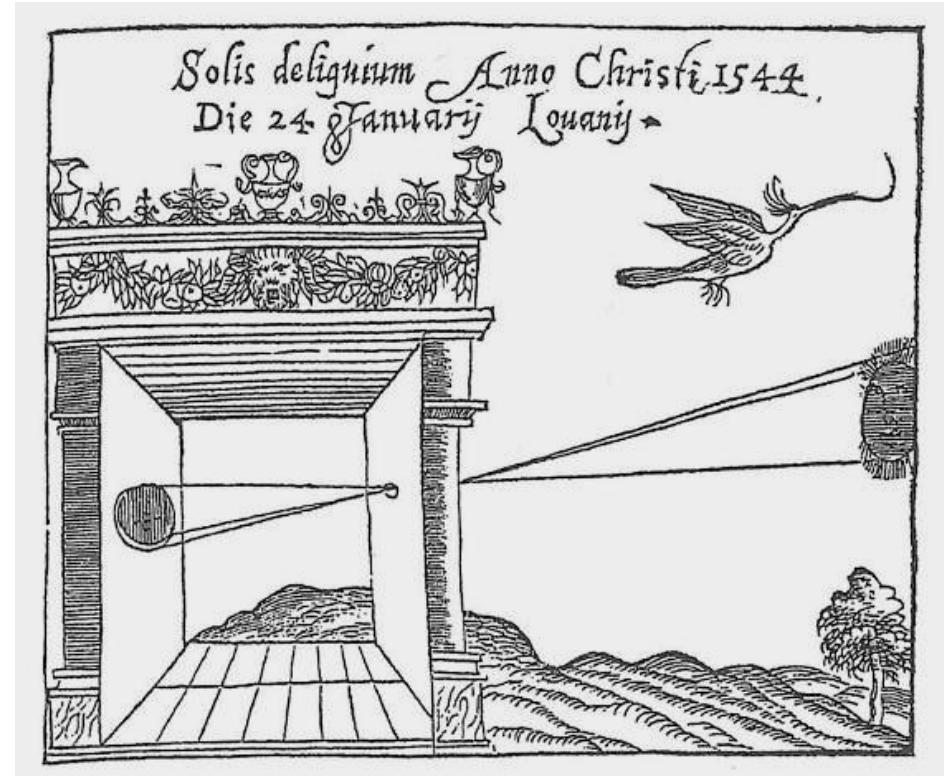
Pinhole camera a.k.a. camera obscura

First mention ...



Chinese philosopher Mozi
(470 to 390 BC)

First camera ...



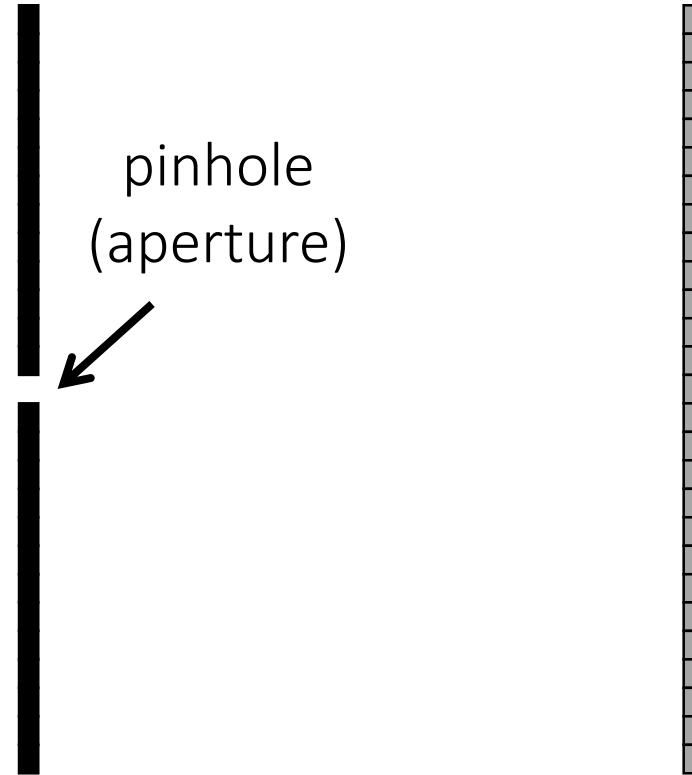
Greek philosopher Aristotle
(384 to 322 BC)

Pinhole camera terms

real-world
object



barrier (diaphragm)



digital sensor
(CCD or CMOS)

Pinhole camera terms

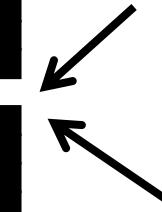
real-world
object



barrier (diaphragm)



pinhole
(aperture)



camera center
(center of projection)

image plane

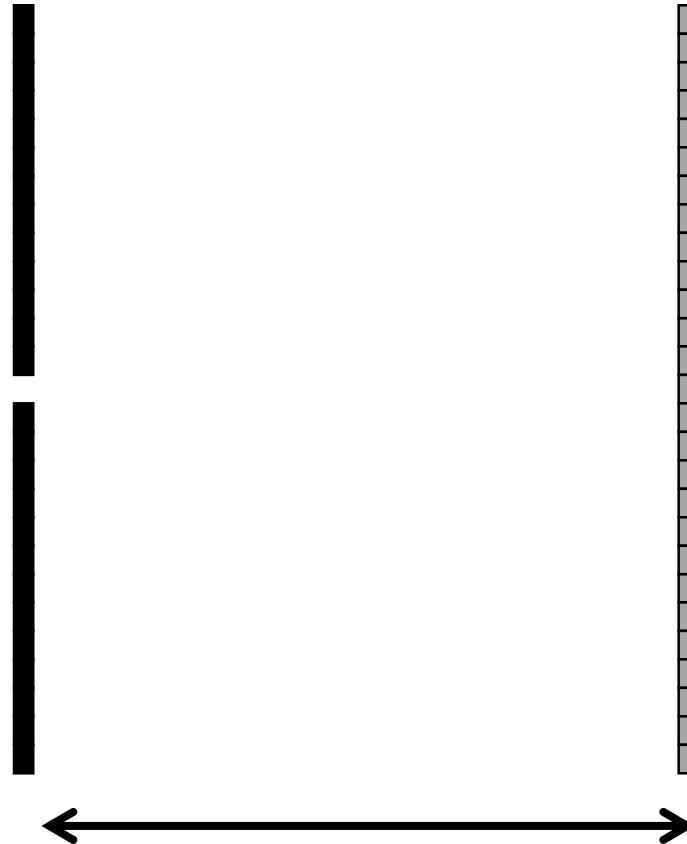


digital sensor
(CCD or CMOS)



Focal length

real-world
object

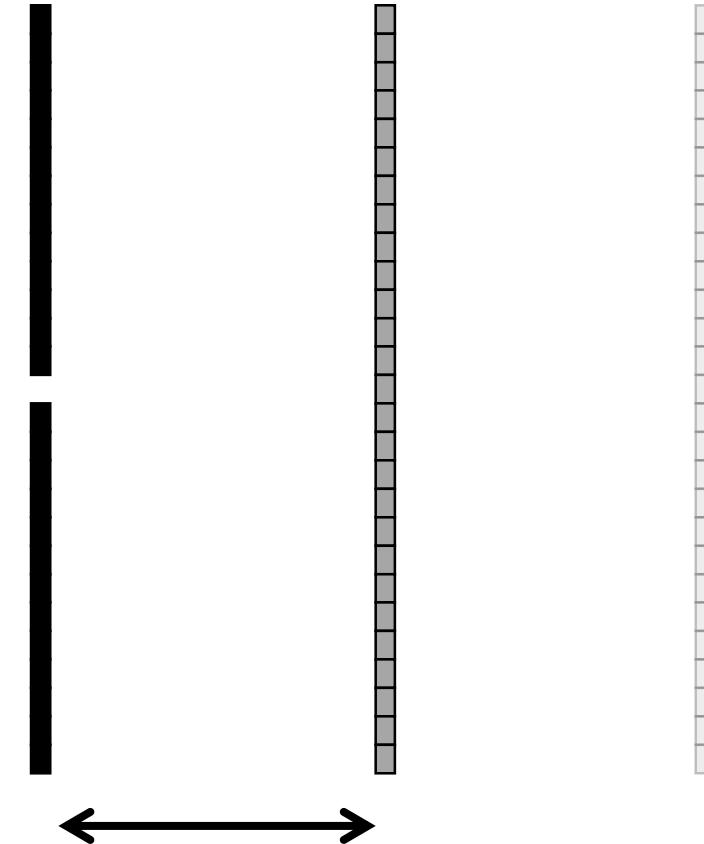


focal length f

Focal length

What happens as we change the focal length?

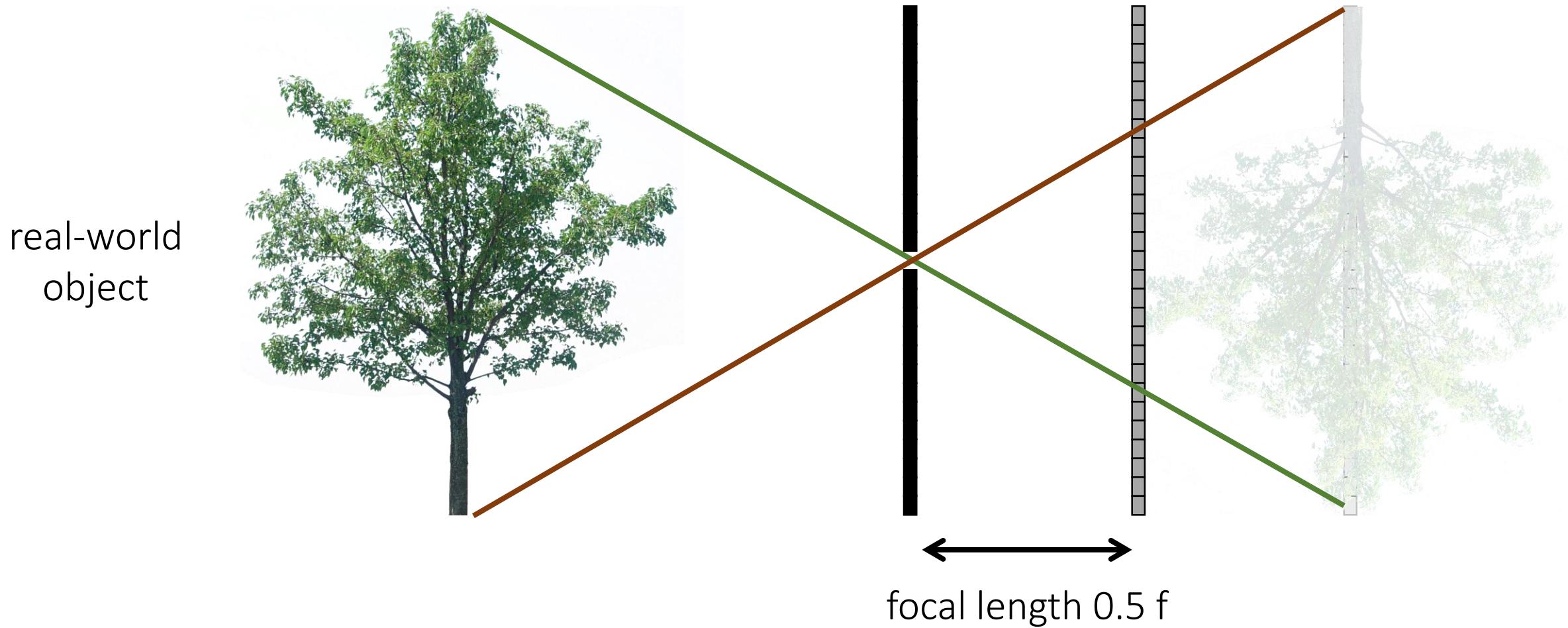
real-world
object



focal length $0.5 f$

Focal length

What happens as we change the focal length?

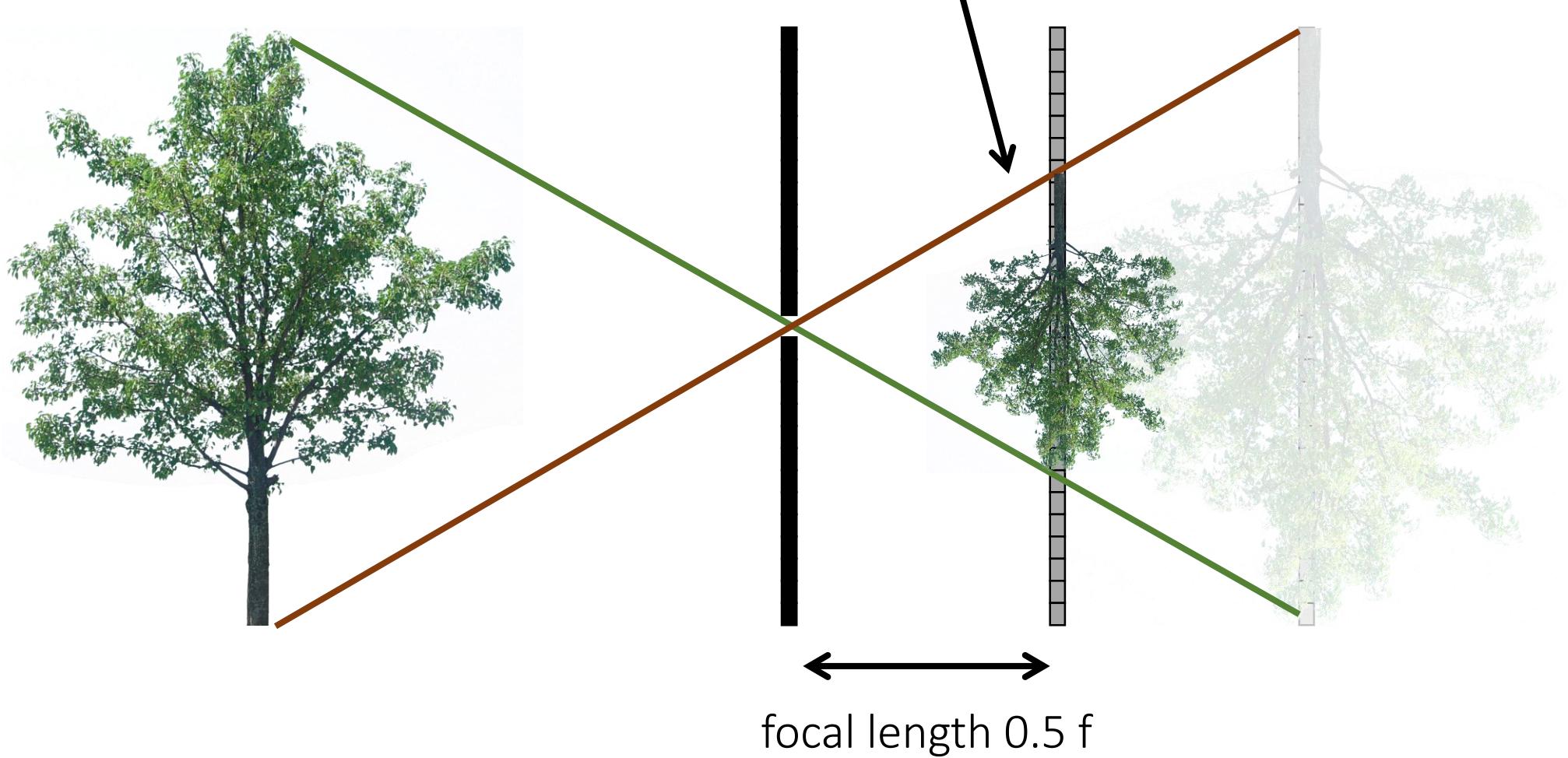


Focal length

What happens as we change the focal length?

real-world
object

object projection is half the size



Pinhole size

real-world
object



pinhole
diameter



Ideal pinhole has infinitesimally small size
In practice that is impossible

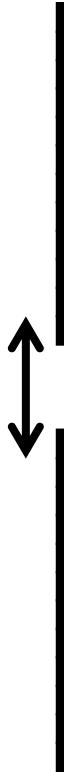
Pinhole size

What happens as we change the pinhole diameter?

real-world
object



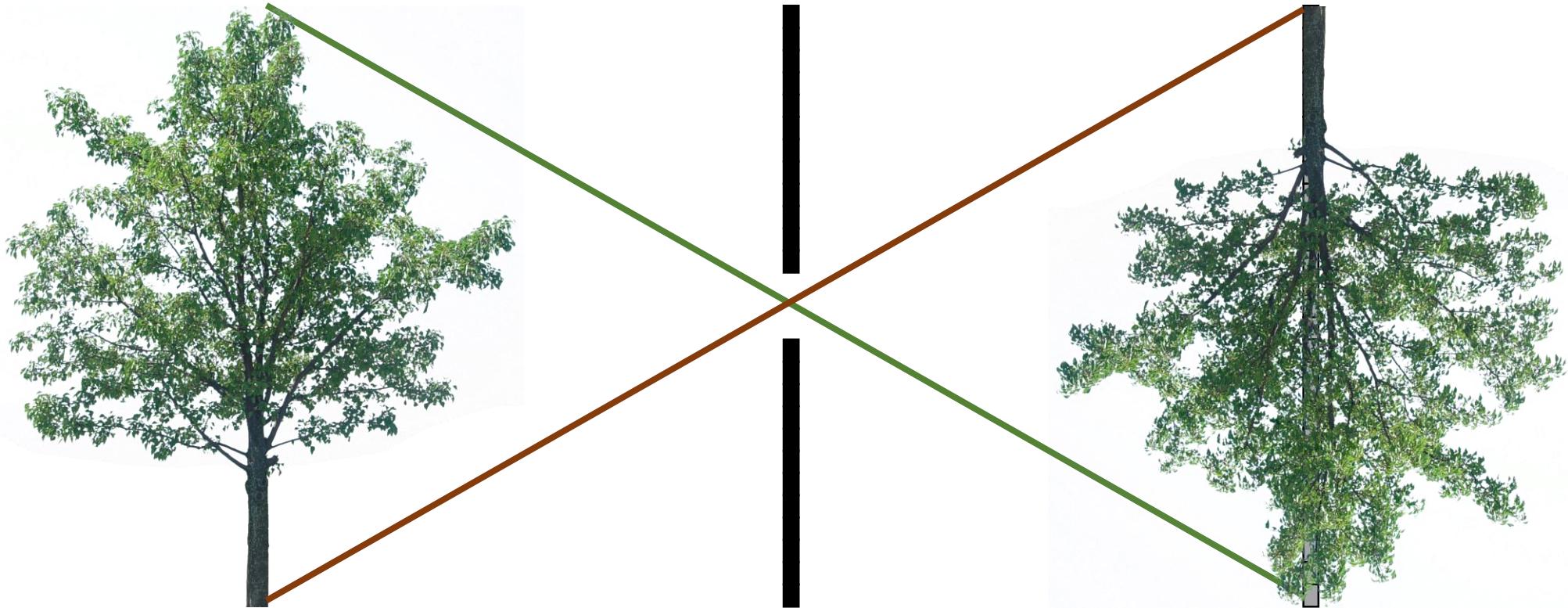
pinhole
diameter



Pinhole size

What happens as we change the pinhole diameter?

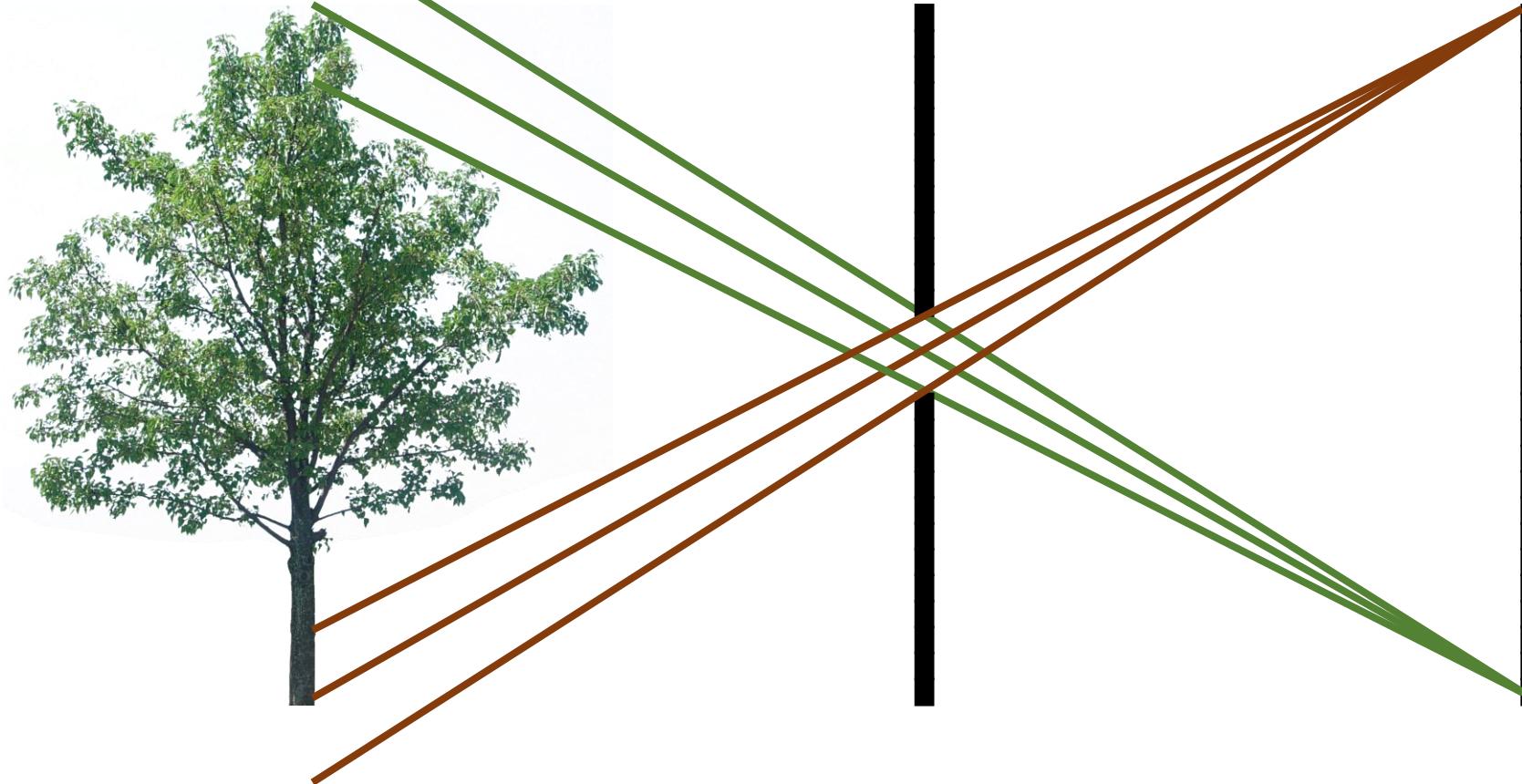
real-world
object



Pinhole size

What happens as we change the pinhole diameter?

real-world
object

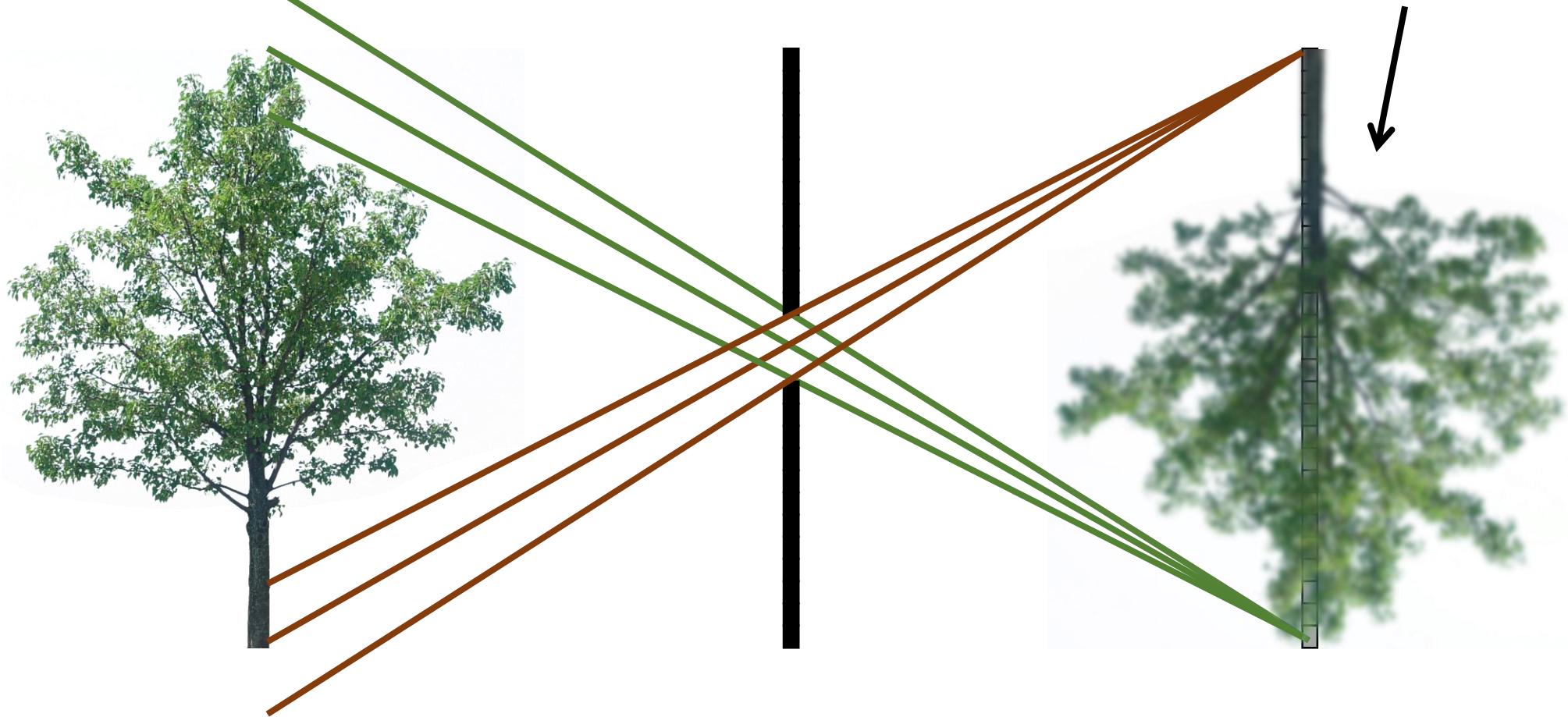


Pinhole size

What happens as we change the pinhole diameter?

object projection becomes blurrier

real-world
object

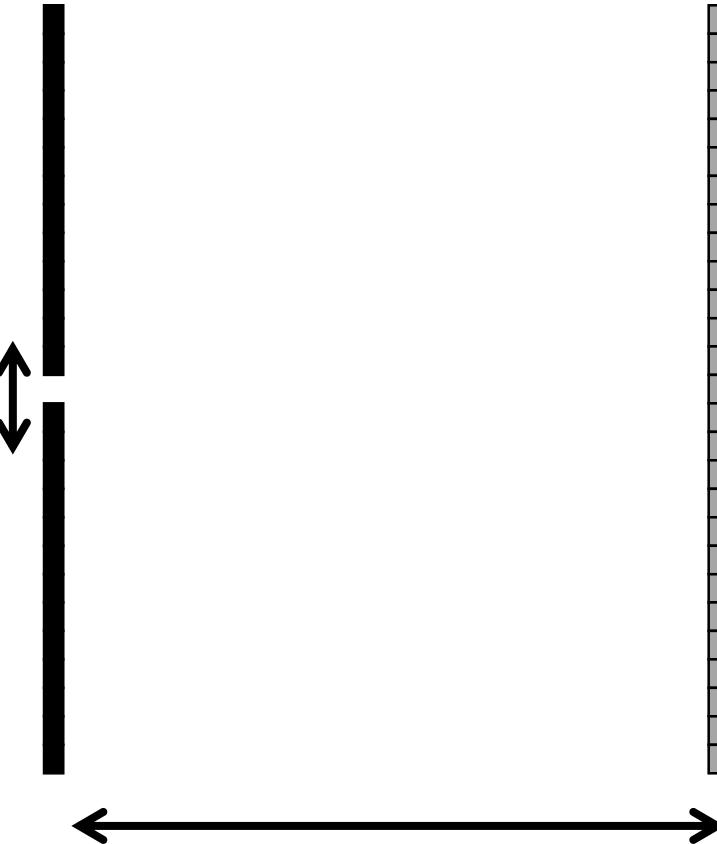


What about light efficiency?

real-world
object



pinhole
diameter



focal length f

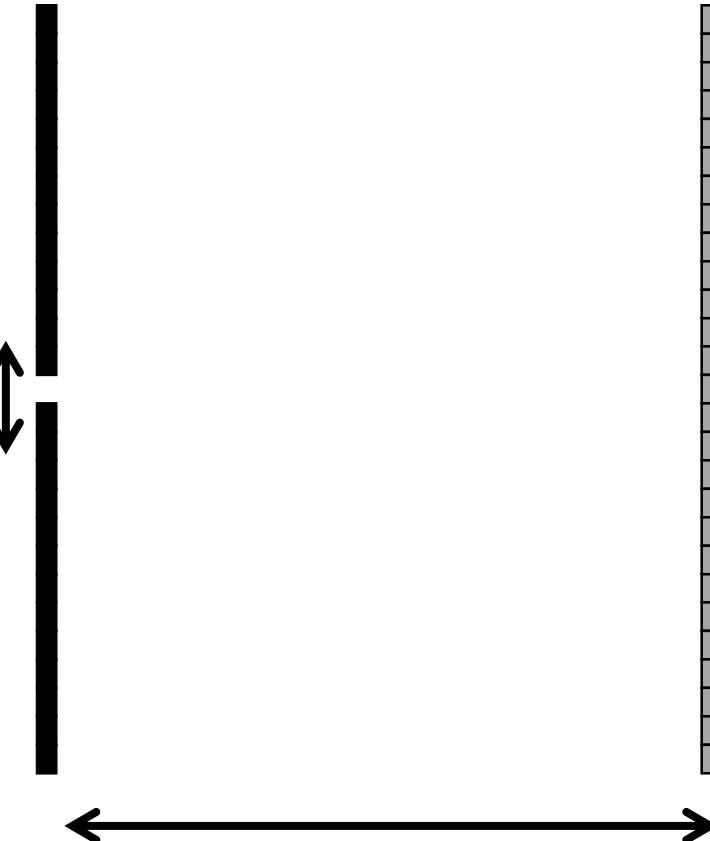
What is the effect of doubling the pinhole diameter?

What about light efficiency?

real-world
object



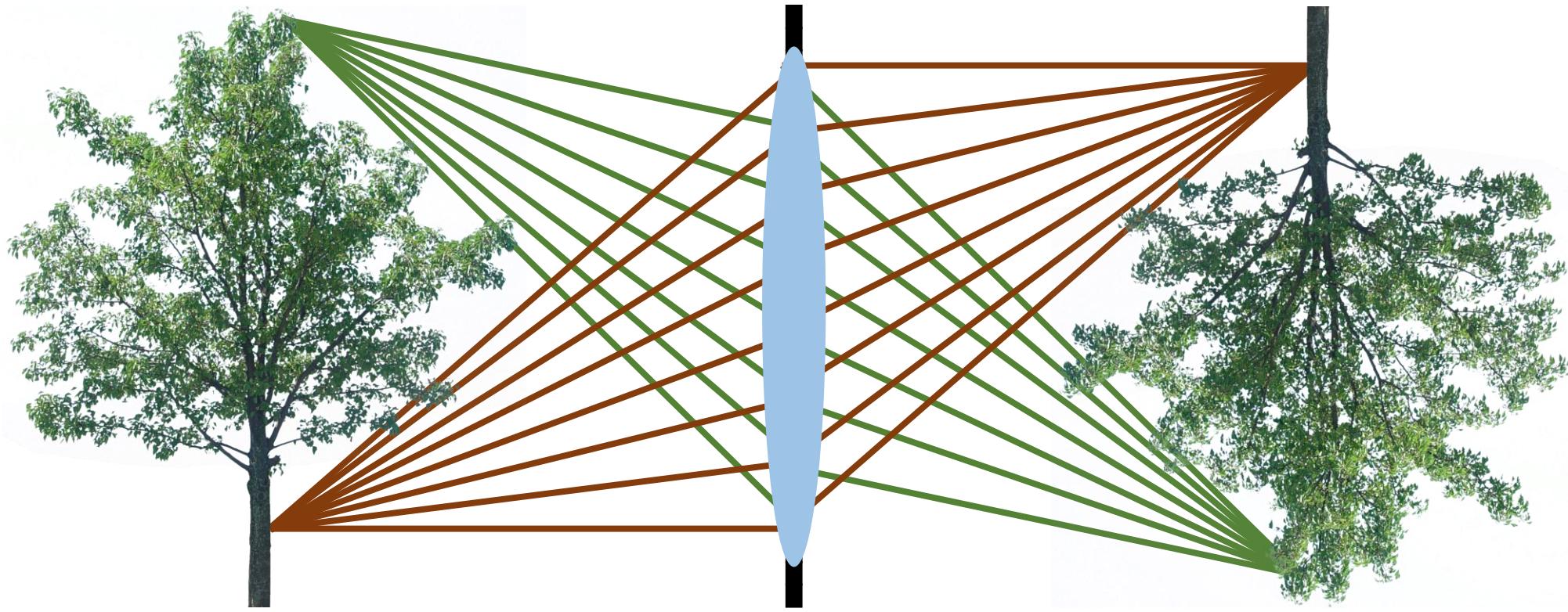
pinhole
diameter



focal length f

$2 \times$ pinhole diameter $\rightarrow 4 \times$ light

In practice



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

How does this mapping work exactly?

Accidental pinholes

What does this image say about the world outside?



Accidental pinhole camera



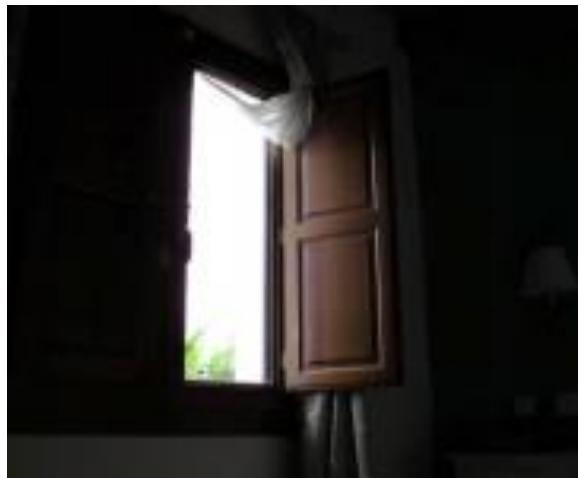
Antonio Torralba, William T. Freeman
Computer Science and Artificial Intelligence Laboratory (CSAIL)
MIT
torralba@mit.edu, billf@mit.edu

Accidental pinhole camera

projected pattern on the wall



upside down



window is an
aperture

window with smaller gap



view outside window



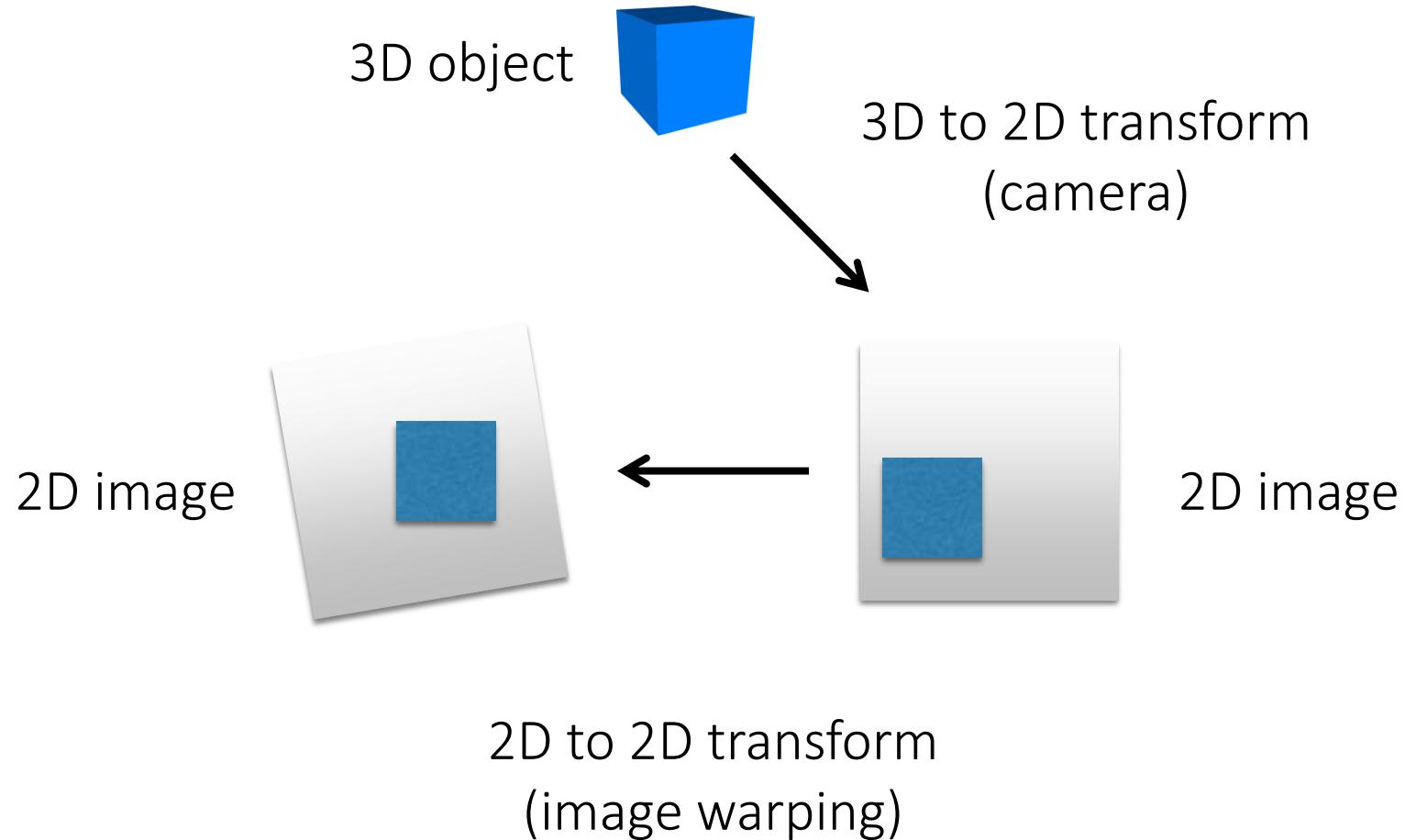
Camera matrix

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

homogeneous coordinates

$$\mathcal{C} = \mathbf{P} \mathbf{X}$$

2D image
point

camera
matrix 3D world
 point

What are the dimensions of each variable?

The camera as a coordinate transformation

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

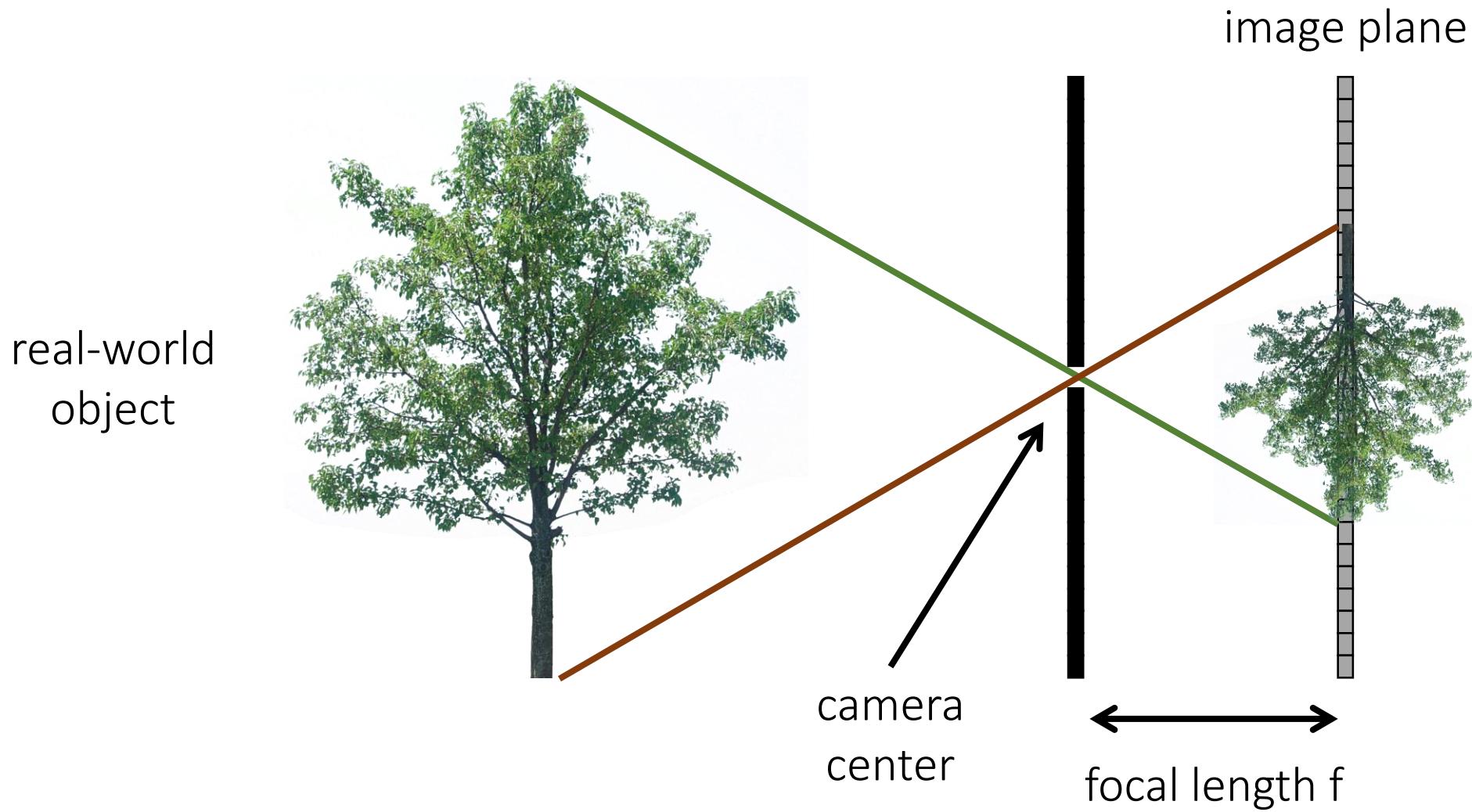
$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} p_1 & p_2 & p_3 & p_4 \\ p_5 & p_6 & p_7 & p_8 \\ p_9 & p_{10} & p_{11} & p_{12} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

homogeneous
image coordinates
 3×1

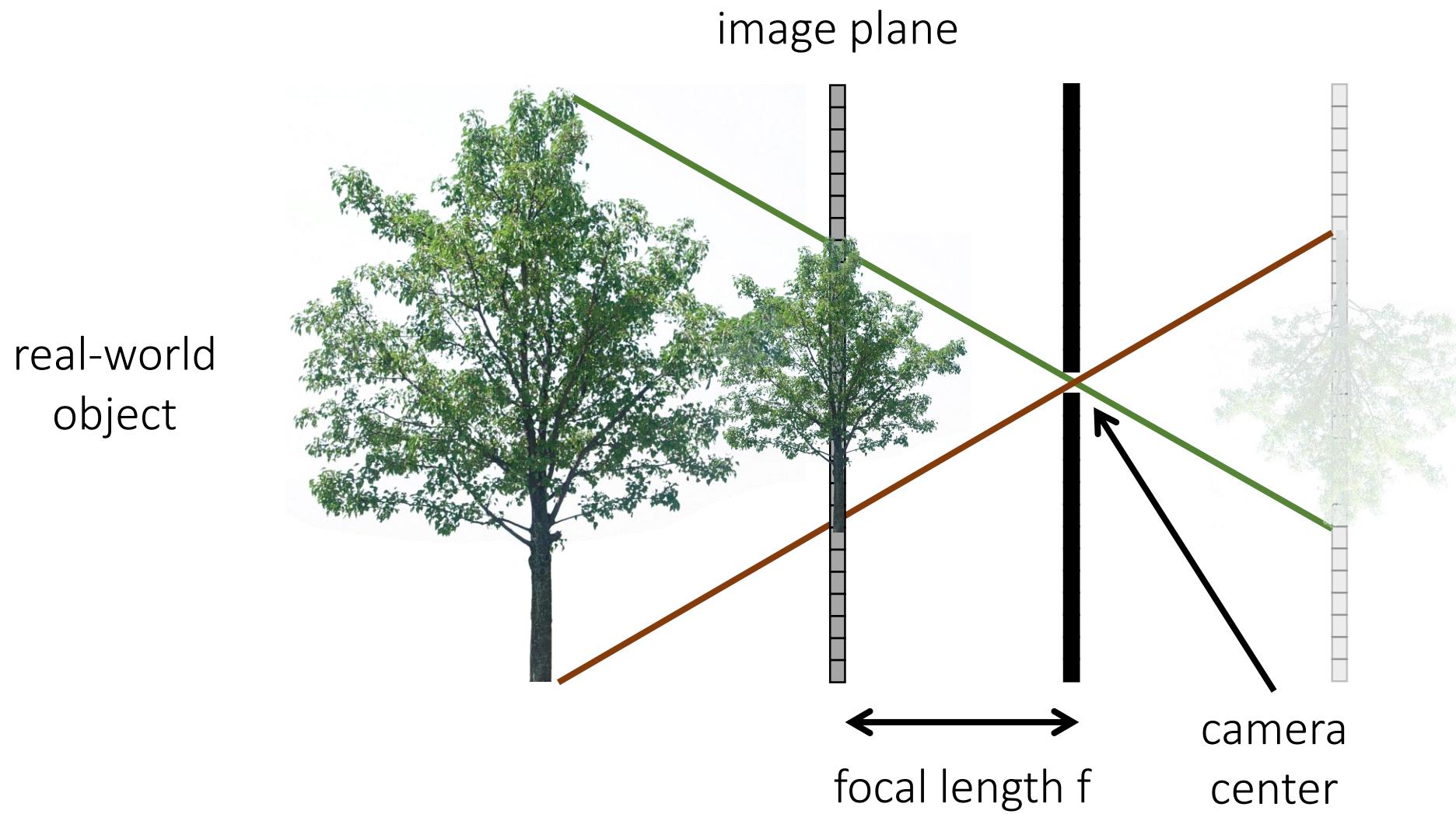
camera
matrix
 3×4

homogeneous
world coordinates
 4×1

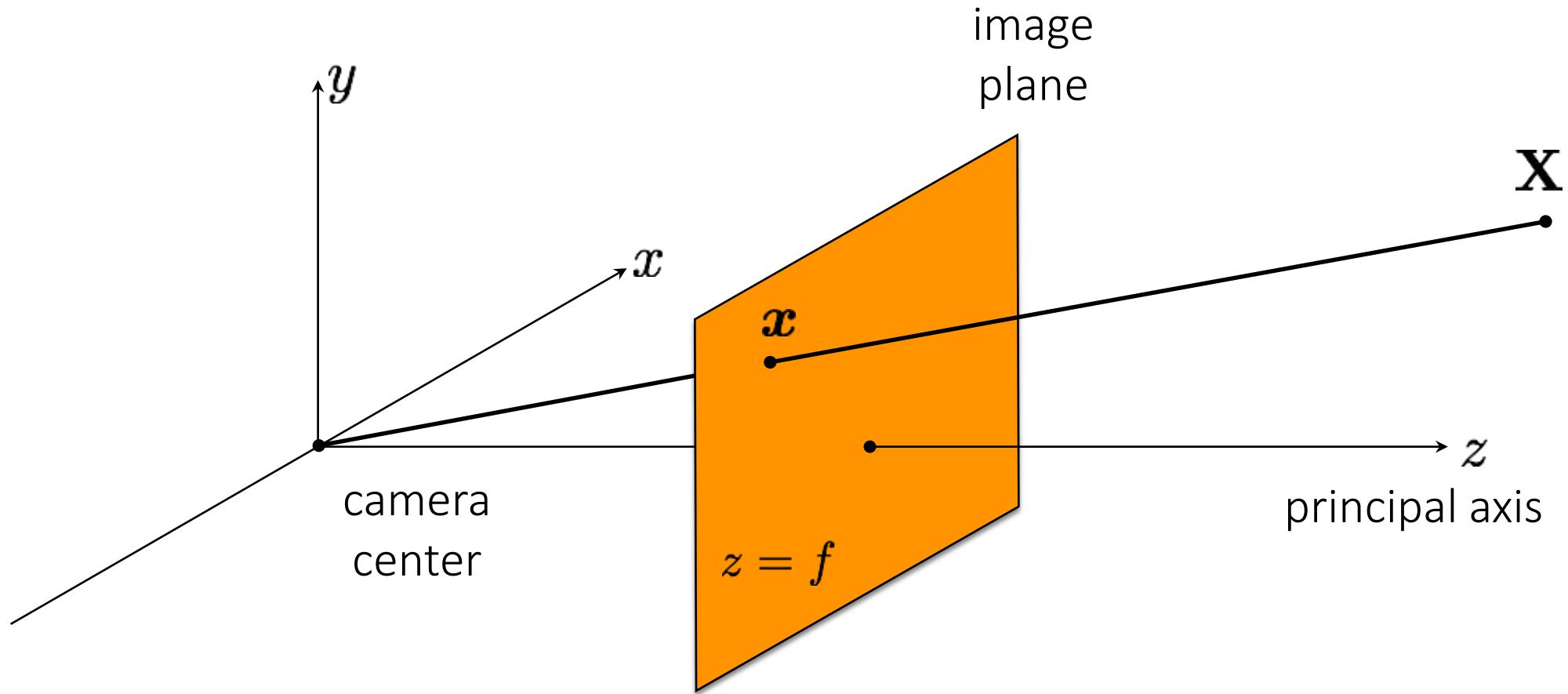
The pinhole camera



The (rearranged) pinhole camera

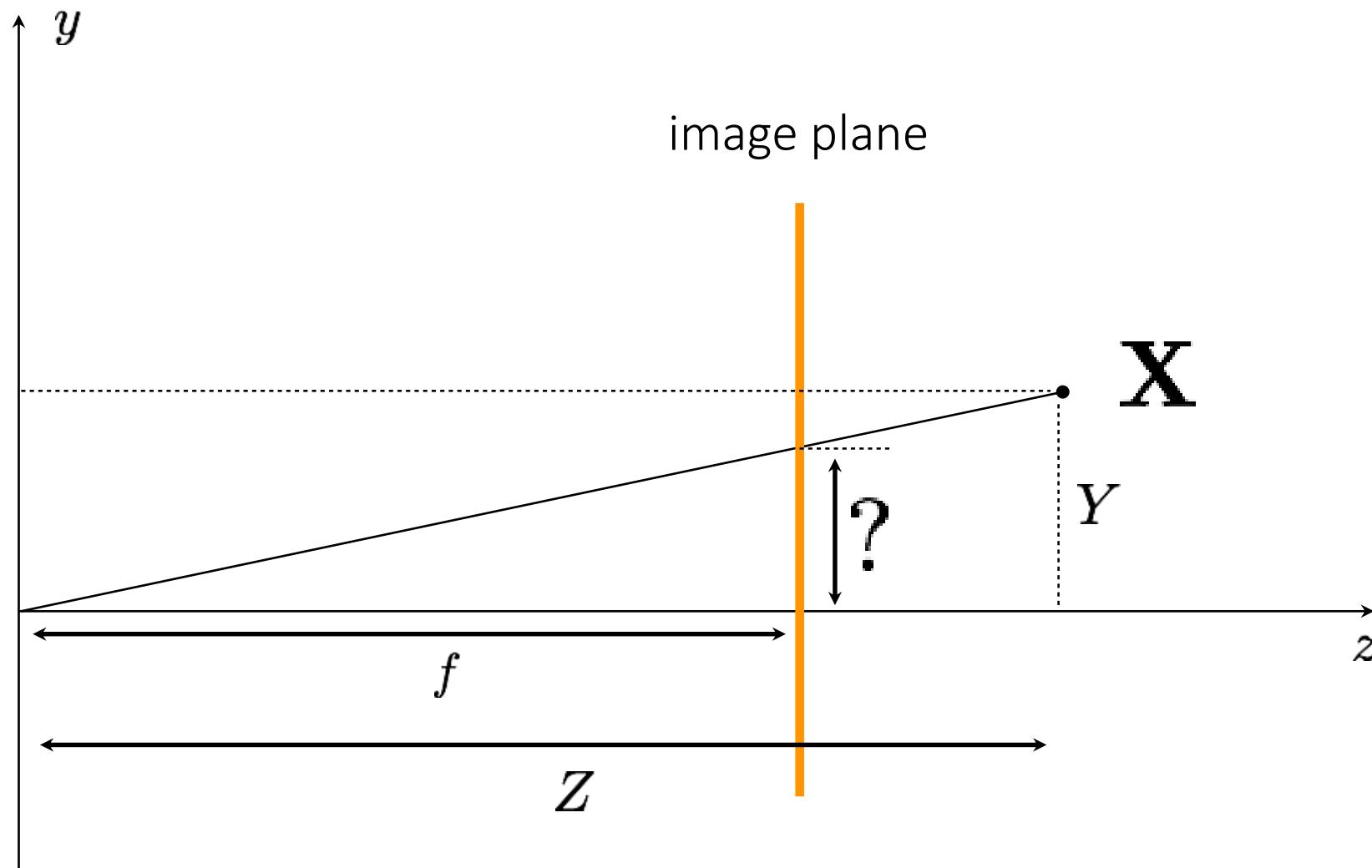


The (rearranged) pinhole camera



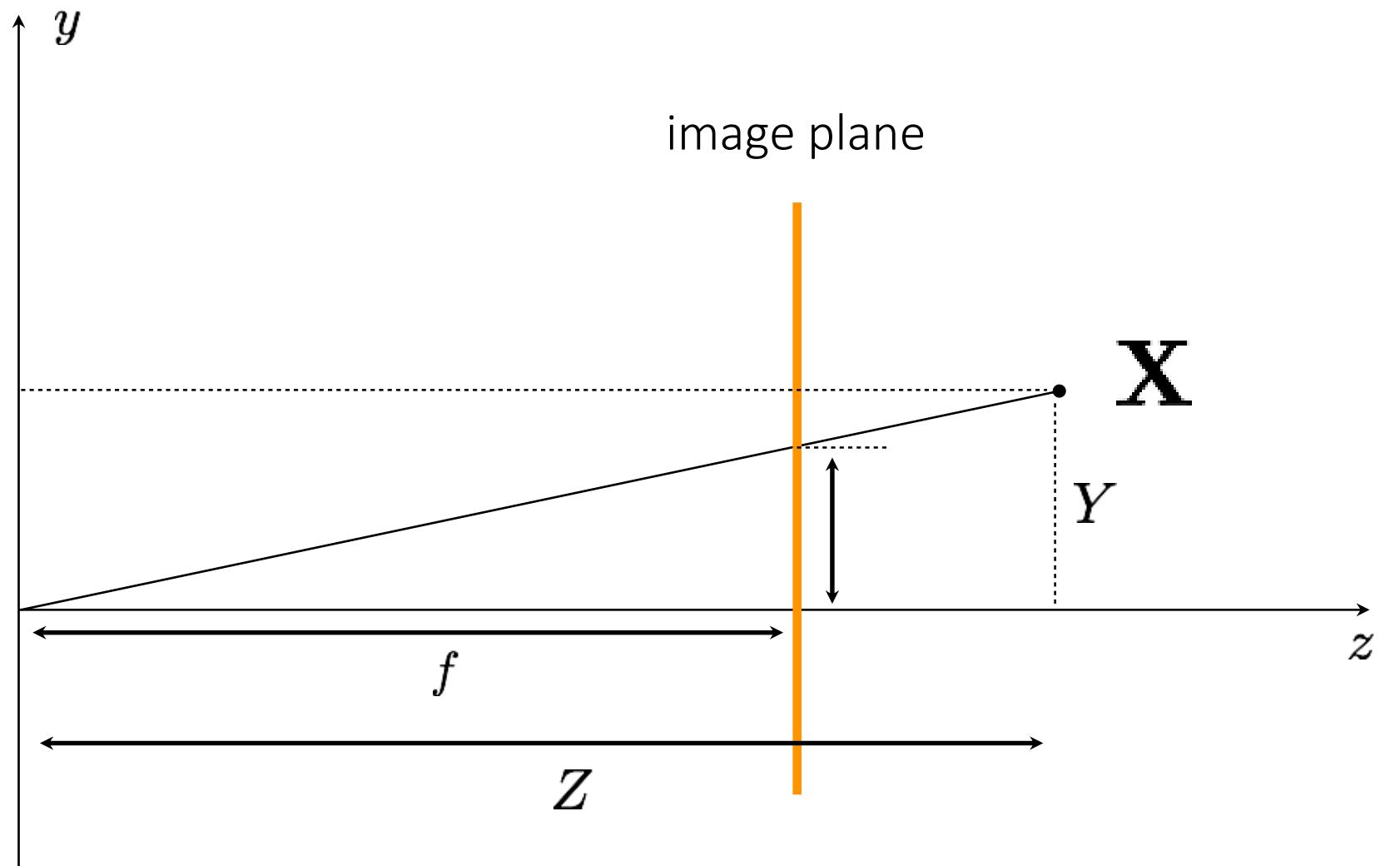
What is the equation for image coordinate \mathbf{x} in terms of \mathbf{X} ?

The 2D view of the (rearranged) pinhole camera



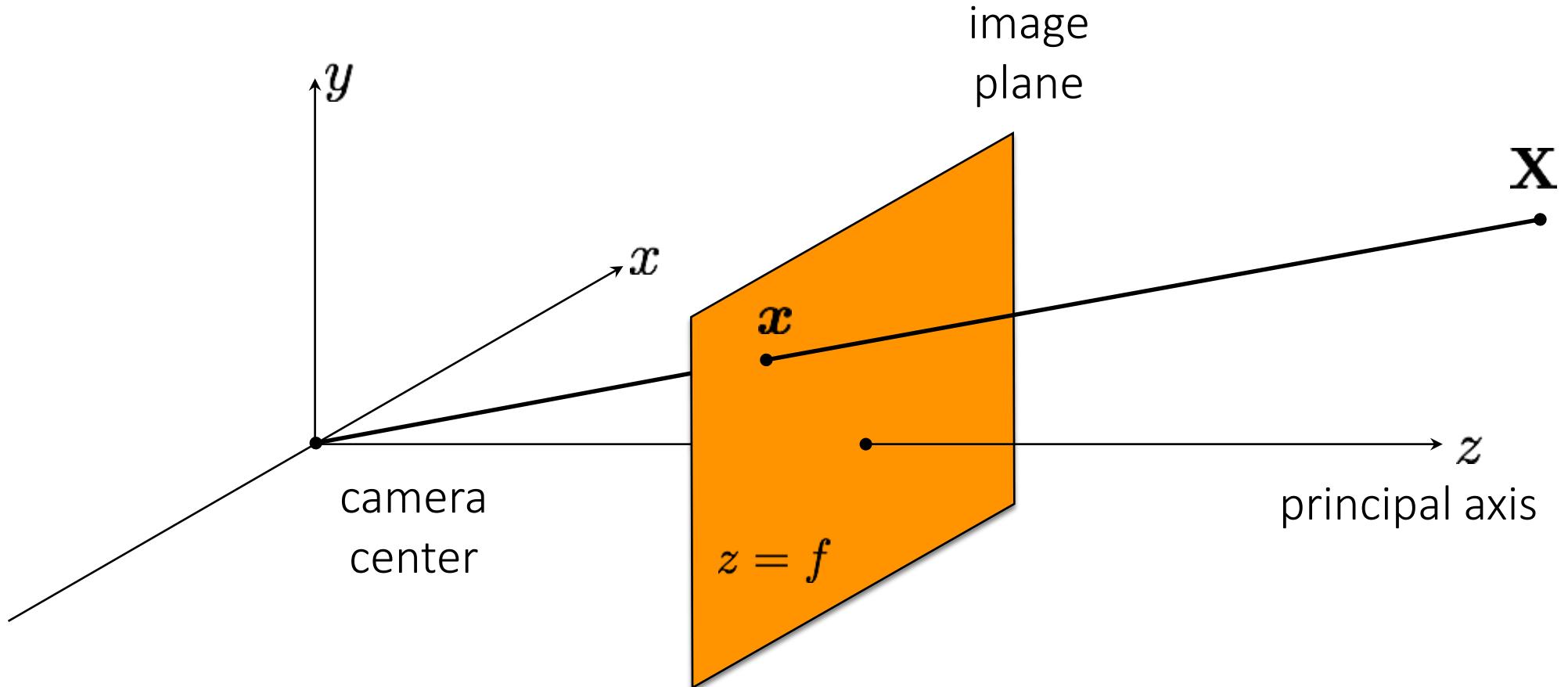
What is the equation for image coordinate x in terms of X ?

The 2D view of the (rearranged) pinhole camera



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

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:

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

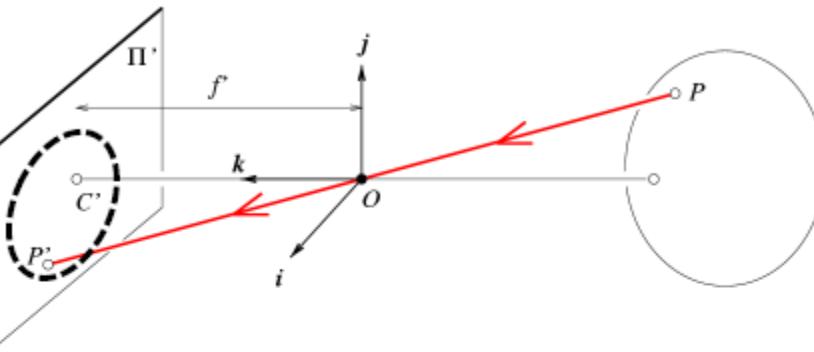
General camera model:

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

What does the pinhole camera projection look like?

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

Relating a real-world point to a point on the camera



In homogeneous coordinates:

$$P' = \begin{bmatrix} f & x \\ f & y \\ z \end{bmatrix} = \underbrace{\begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_M \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

ideal world

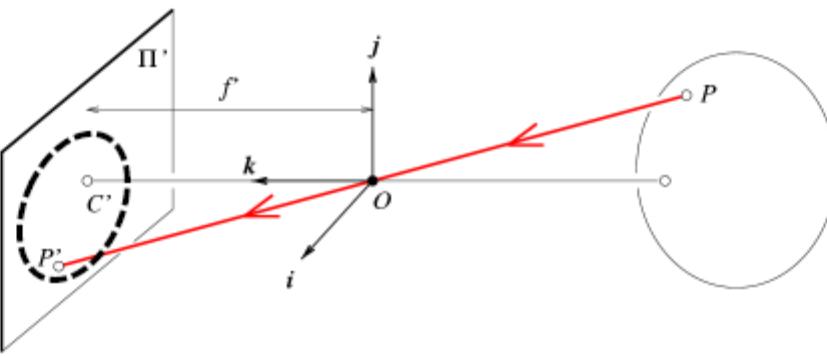
Intrinsic Assumptions

- Unit aspect ratio
- Optical center at (0,0)
- No skew

Extrinsic Assumptions

- No rotation
- Camera at (0,0,0)

Relating a real-world point to a point on the camera



In homogeneous coordinates:

$$P' = \begin{bmatrix} f x \\ f y \\ z \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = K[I \ 0]P$$

K

Intrinsic Assumptions

- Unit aspect ratio
- Optical center at (0,0)
- No skew

Extrinsic Assumptions

- No rotation
- Camera at (0,0,0)

Remove assumption: known optical center

Intrinsic Assumptions

- Optical center at $(0,0)$
- **Optical center at (u_0, v_0)**
- Square pixels
- No skew

Extrinsic Assumptions

- No rotation
- Camera at $(0,0,0)$

$$P' = K \begin{bmatrix} I & 0 \end{bmatrix} P \xrightarrow{\text{w}} w \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & u_0 \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

Remove assumption: square pixels

Intrinsic Assumptions

- Optical center at (u_0, v_0)
- ~~Square pixels~~
- **Rectangular pixels**
- No skew

Extrinsic Assumptions

- No rotation
- Camera at $(0,0,0)$

$$P' = K \begin{bmatrix} I & 0 \end{bmatrix} P \rightarrow w \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha & 0 & u_0 & 0 \\ 0 & \beta & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

Remove assumption: non-skewed pixels

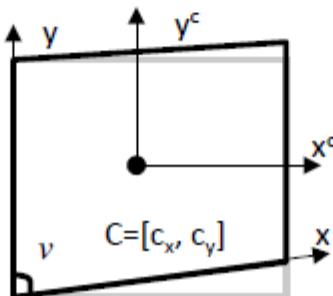
Intrinsic Assumptions

- Optical center at (u_0, v_0)
- Rectangular pixels
- ~~No skew~~
- Small skew

Extrinsic Assumptions

- No rotation
- Camera at $(0,0,0)$

$$P' = K \begin{bmatrix} I & 0 \end{bmatrix} P \rightarrow w \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha & s & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$



Remove assumption: non-skewed pixels

Intrinsic Assumptions

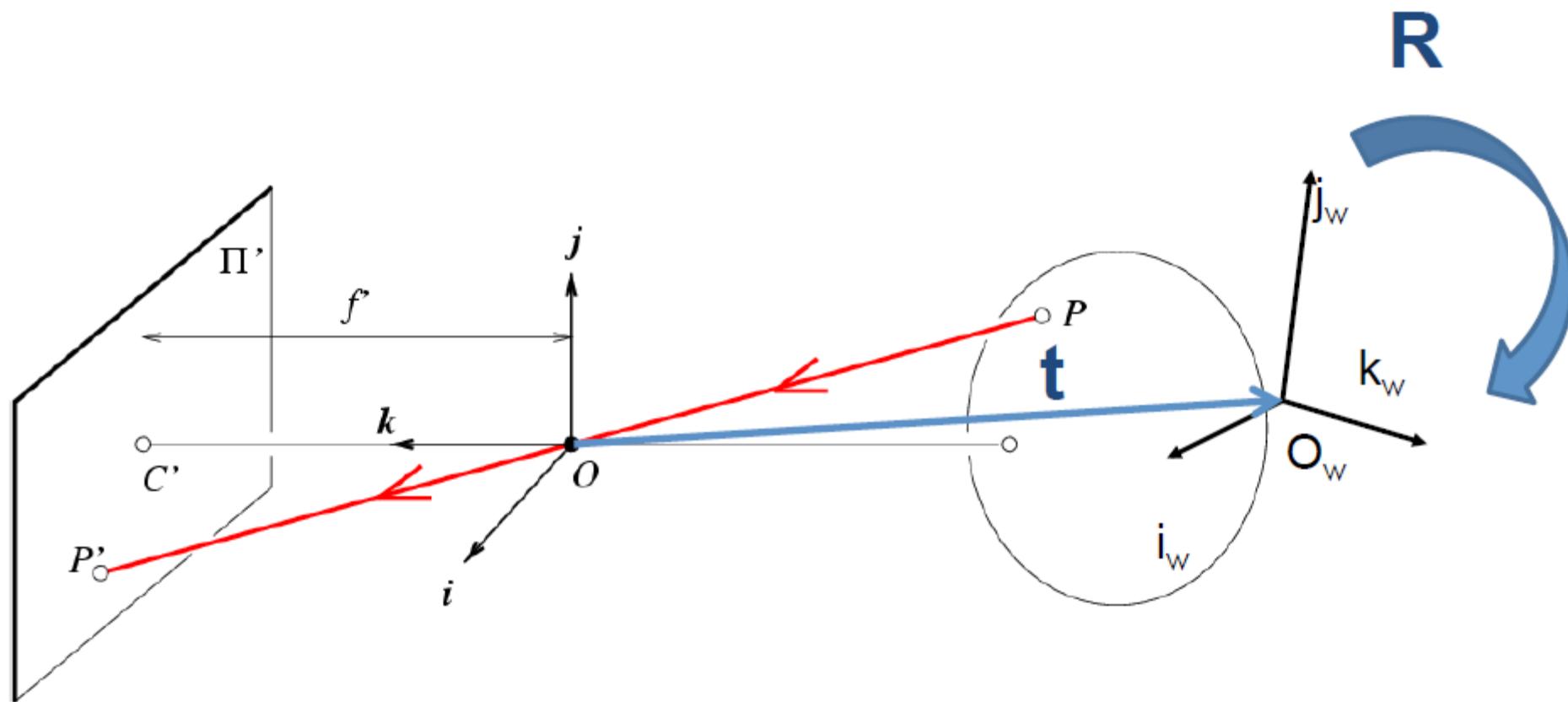
- Optical center at (u_0, v_0)
- Rectangular pixels
- Small skew

Extrinsic Assumptions

- No rotation
- Camera at $(0,0,0)$

$$P' = K \begin{bmatrix} I & 0 \end{bmatrix} P \xrightarrow{\text{Intrinsic parameters}} w \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha & s & u_0 & 0 \\ 0 & \beta & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

Real world camera:
Translate + Rotate



Remove assumption: allow translation

Intrinsic Assumptions

- Optical center at (u_0, v_0)
- Rectangular pixels
- Small skew

Extrinsic Assumptions

- No rotation
- Camera at $(0,0,0)$ $\rightarrow (t_x, t_y, t_z)$

$$P' = K \begin{bmatrix} I & \bar{t} \end{bmatrix} P \quad \xrightarrow{\text{blue arrow}} \quad w \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha & 0 & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

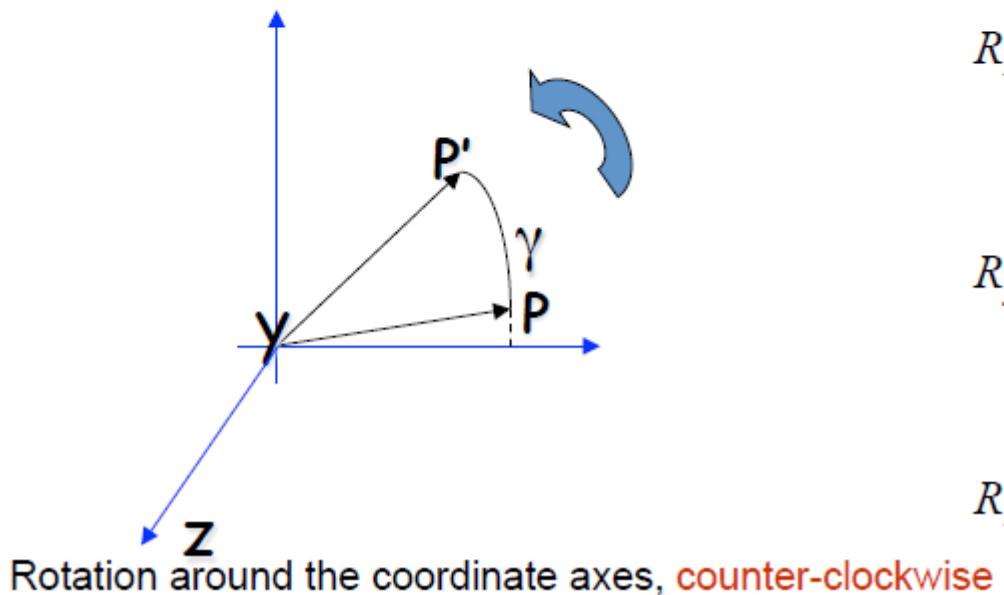
Remove assumption: allow rotation

Intrinsic Assumptions

- Optical center at (u_0, v_0)
- Rectangular pixels
- Small skew

Extrinsic Assumptions

- ~~No~~ rotation
- Camera at (t_x, t_y, t_z)



Rotation around the coordinate axes, **counter-clockwise**

$$R_x(\alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{bmatrix}$$

$$R_y(\beta) = \begin{bmatrix} \cos \beta & 0 & \sin \beta \\ 0 & 1 & 0 \\ -\sin \beta & 0 & \cos \beta \end{bmatrix}$$

$$R_z(\gamma) = \begin{bmatrix} \cos \gamma & -\sin \gamma & 0 \\ \sin \gamma & \cos \gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

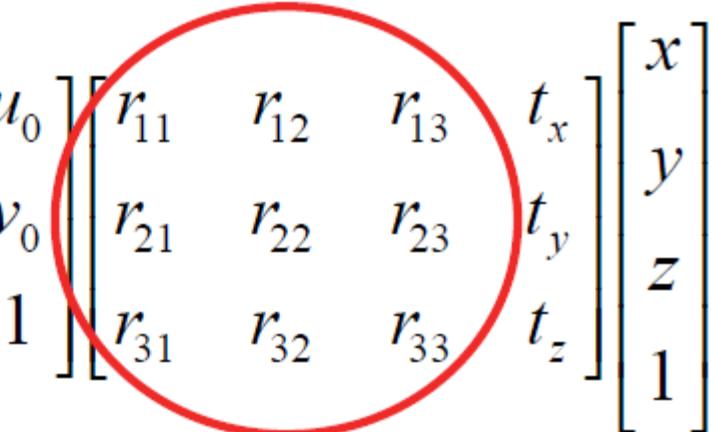
Remove assumption: allow rotation

Intrinsic Assumptions

- Optical center at (u_0, v_0)
- Rectangular pixels
- Small skew

Extrinsic Assumptions

- ~~No~~ rotation
- Camera at (t_x, t_y, t_z)

$$P' = K \begin{bmatrix} R & \bar{t} \end{bmatrix} P \xrightarrow{\text{blue arrow}} w \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha & s & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} t_x \\ t_y \\ t_z \\ 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$


A generic projection matrix

Intrinsic Assumptions

- Optical center at (u_0, v_0)
- Rectangular pixels
- Small skew

Extrinsic Assumptions

- Allow rotation
- Camera at (t_x, t_y, t_z)

$$P' = K \begin{bmatrix} R & \bar{t} \end{bmatrix} P \xrightarrow{\text{w}} w \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha & s & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

A generic projection matrix

Intrinsic Assumptions

- Optical center at (u_0, v_0)
- Rectangular pixels
- Small skew

Extrinsic Assumptions

- Allow rotation
- Camera at (t_x, t_y, t_z)

$$P' = K \begin{bmatrix} R & \bar{t} \end{bmatrix} P \xrightarrow{\text{blue arrow}} w \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha & s & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

Degrees of freedom??

A generic projection matrix

Intrinsic Assumptions

- Optical center at (u_0, v_0)
- Rectangular pixels
- Small skew

Extrinsic Assumptions

- Allow rotation
- Camera at (t_x, t_y, t_z)

$$P' = K \begin{bmatrix} R & \bar{t} \end{bmatrix} P \xrightarrow{\text{5}} w \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha & s & u_0 \\ 0 & p & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

Degrees of freedom??

Quiz

The camera matrix relates what two quantities?

Quiz

The camera matrix relates what two quantities?

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

Quiz

The camera matrix relates what two quantities?

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

3D points to 2D image points

Quiz

The camera matrix relates what two quantities?

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

3D points to 2D image points

The camera matrix can be decomposed into?

Quiz

The camera matrix relates what two quantities?

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

3D points to 2D image points

The camera matrix can be decomposed into?

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

Quiz

The camera matrix relates what two quantities?

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

3D points to 2D image points

The camera matrix can be decomposed into?

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

intrinsic and extrinsic parameters

Perspective

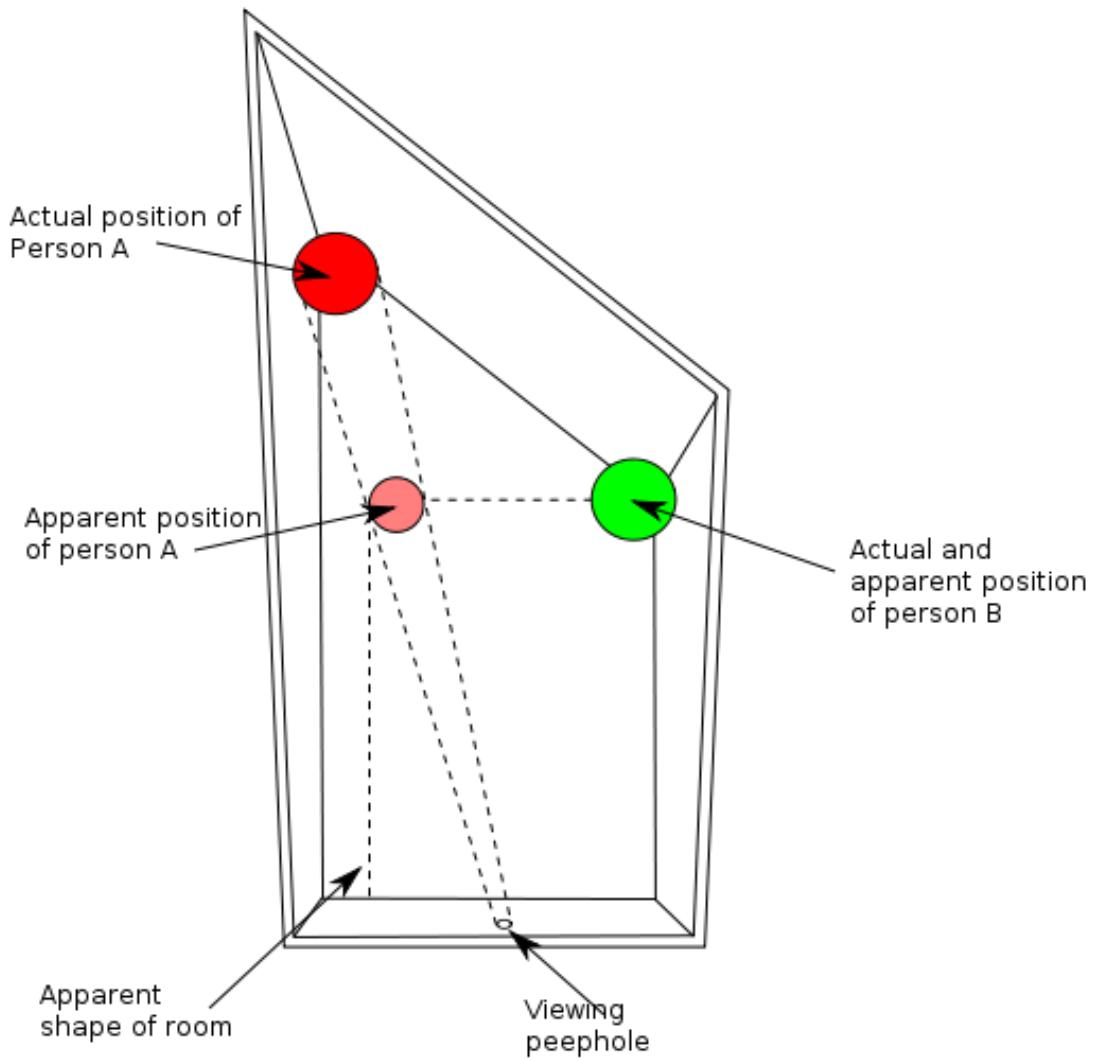
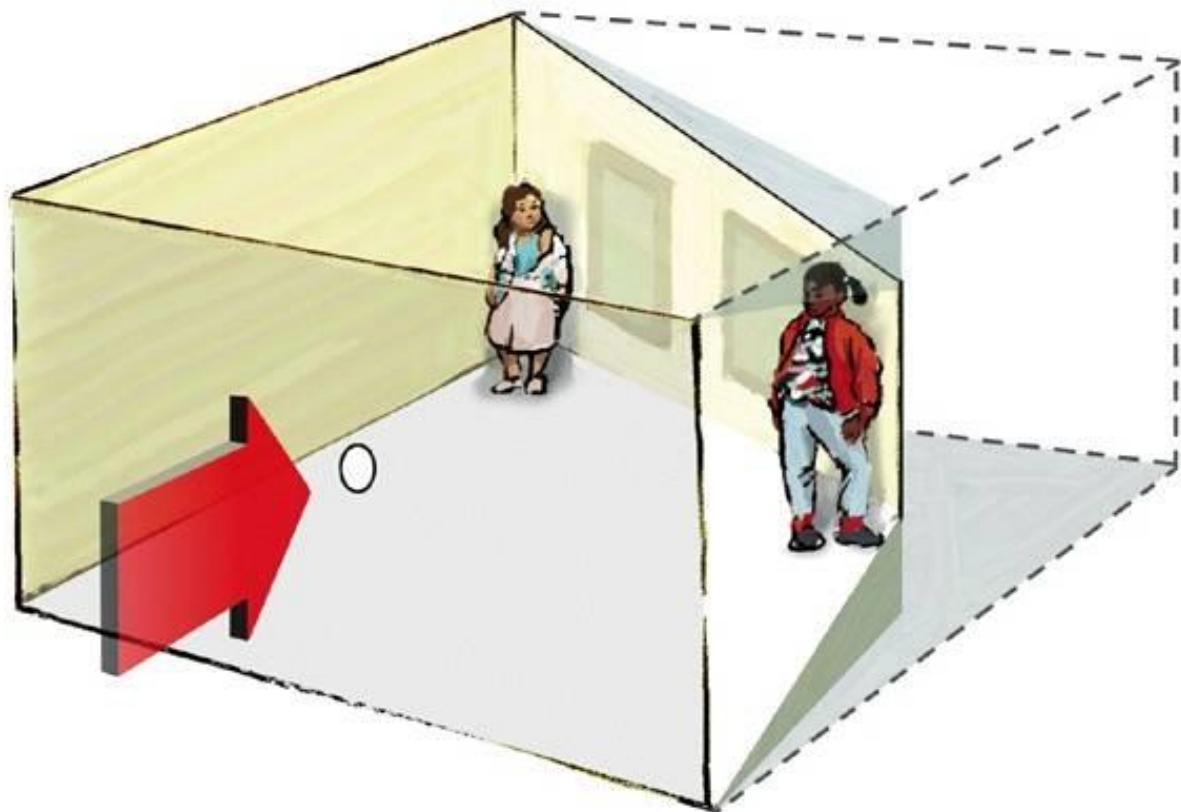
Forced perspective



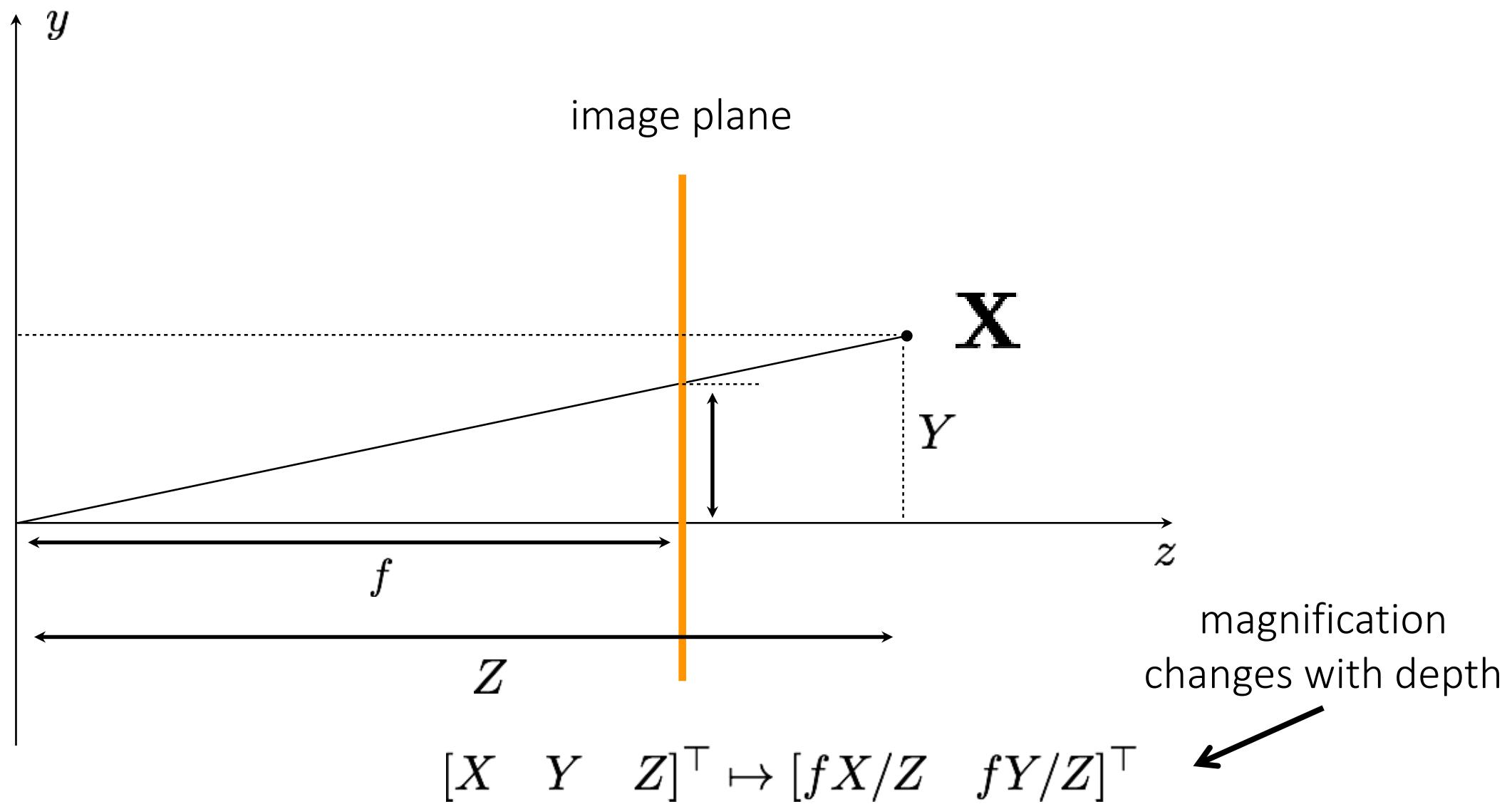
The Ames room illusion



The Ames room illusion

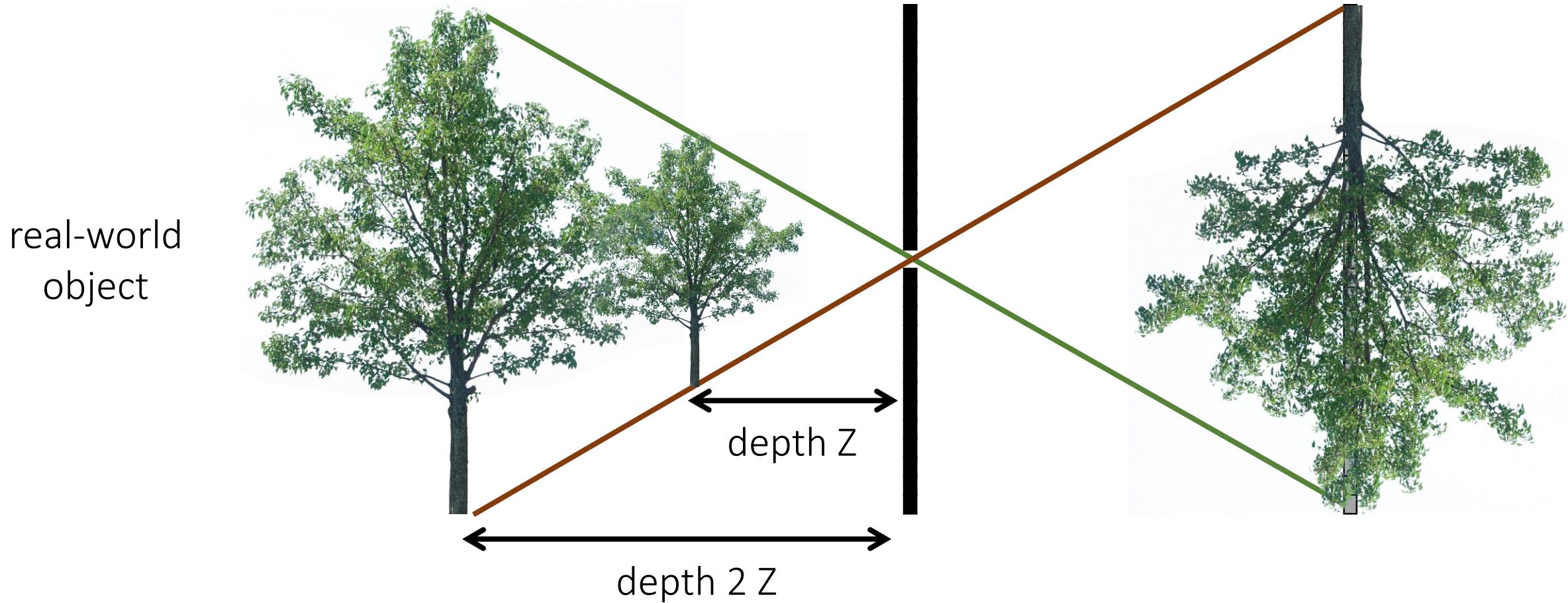


The 2D view of the (rearranged) pinhole camera

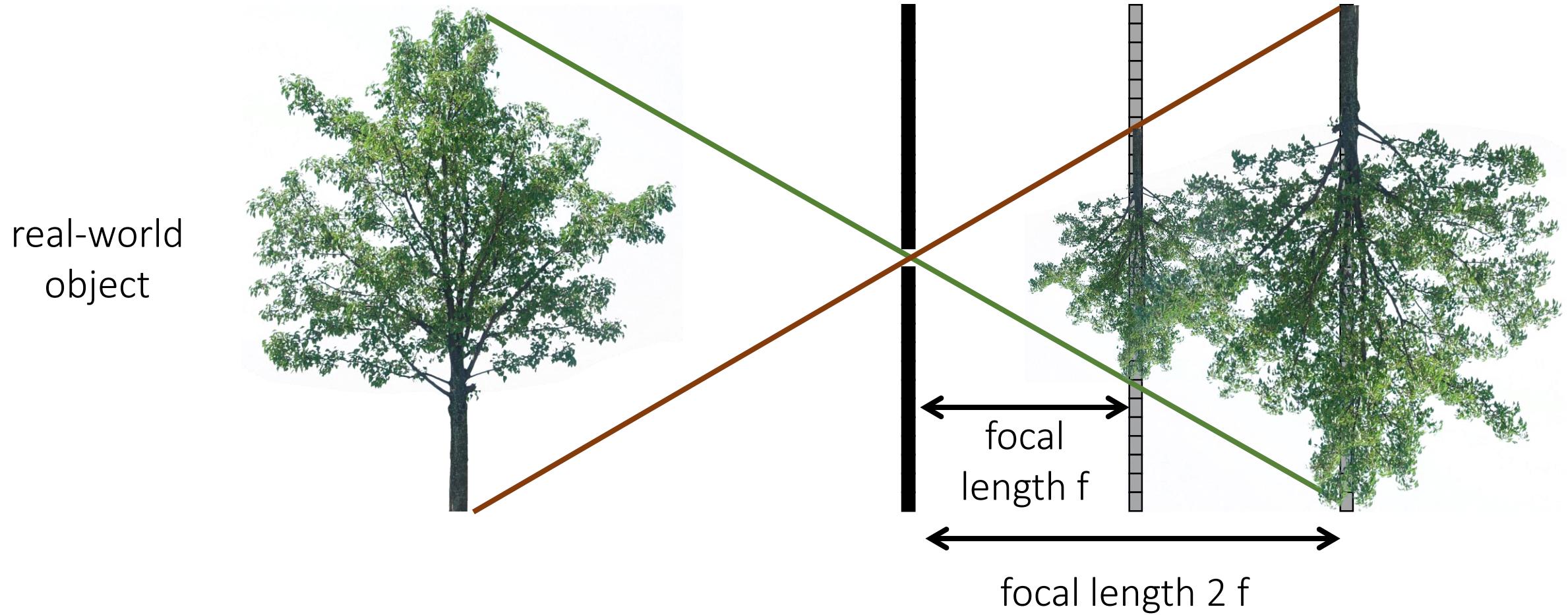


Magnification depends on depth

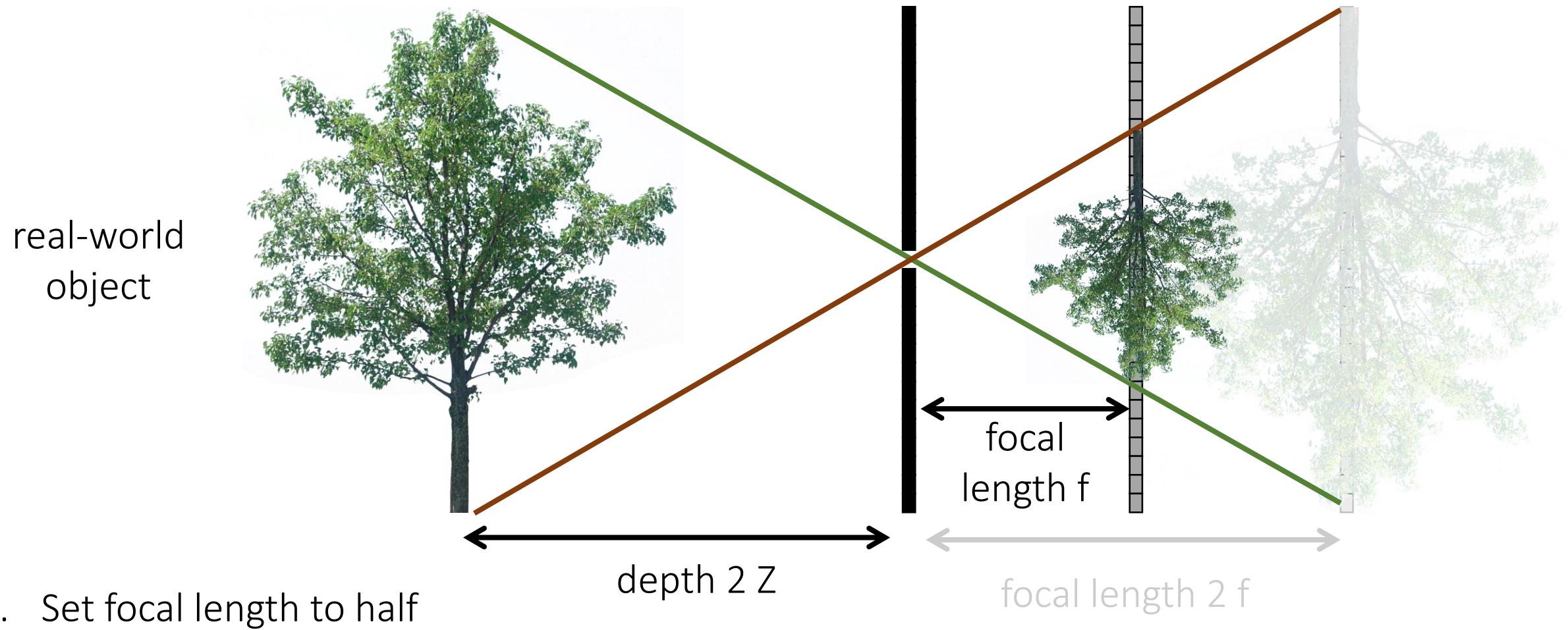
What happens as we change the focal length?



Magnification depends on focal length

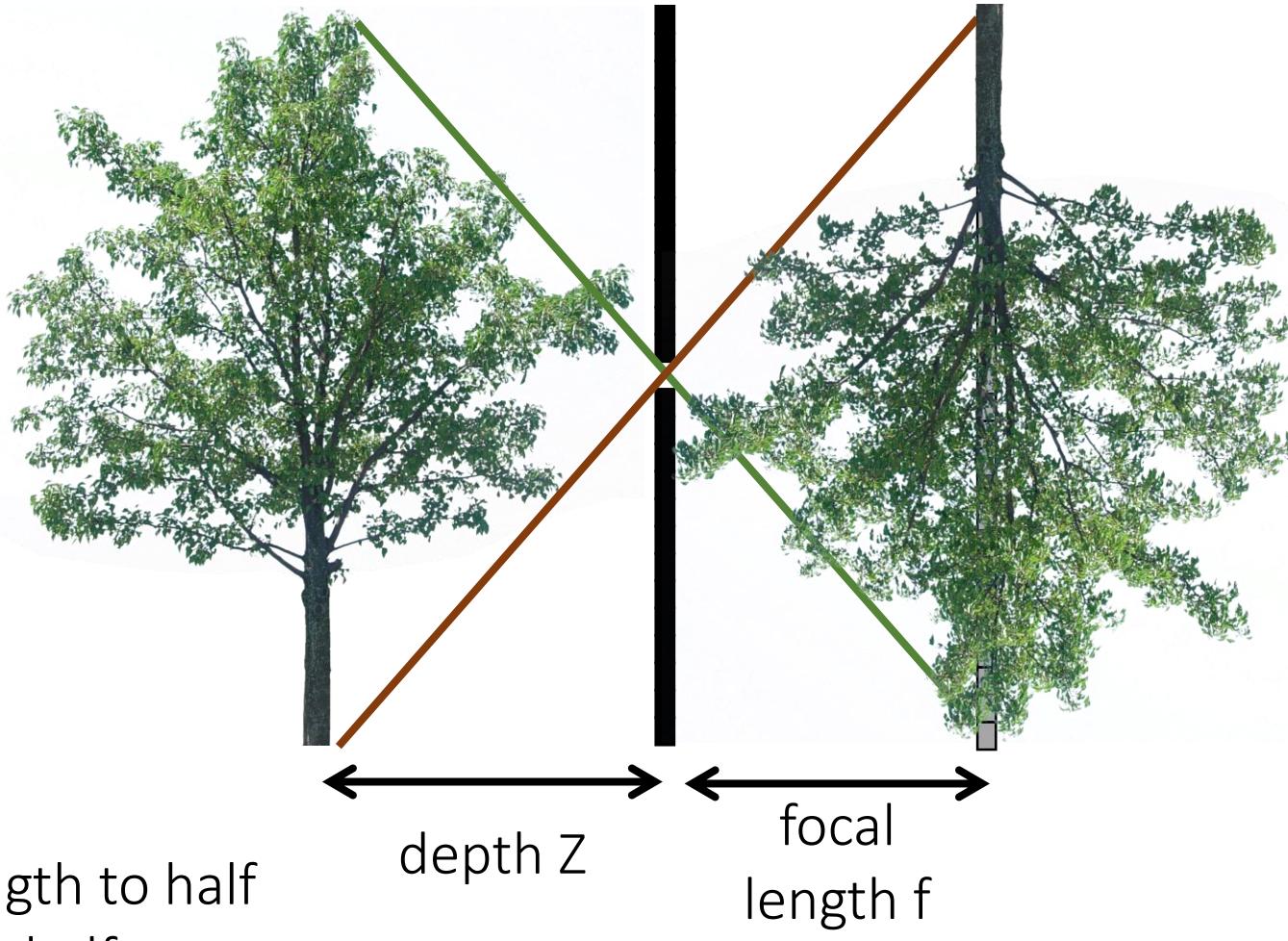


What if...



What if...

real-world
object



Is this the same image as
the one I had at focal
length $2f$ and distance $2Z$?

1. Set focal length to half
2. Set depth to half

Perspective distortion



long focal length

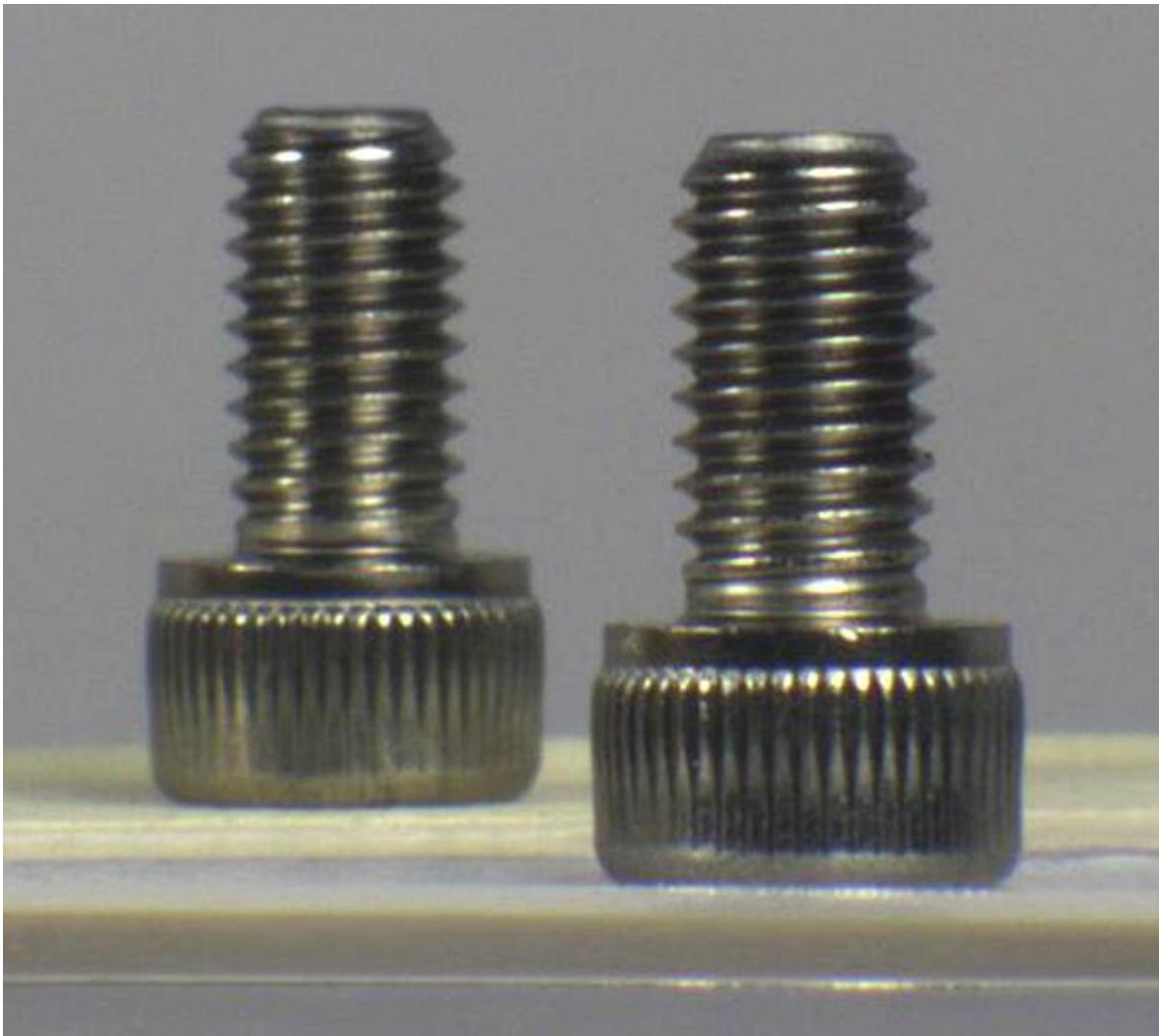


mid focal length

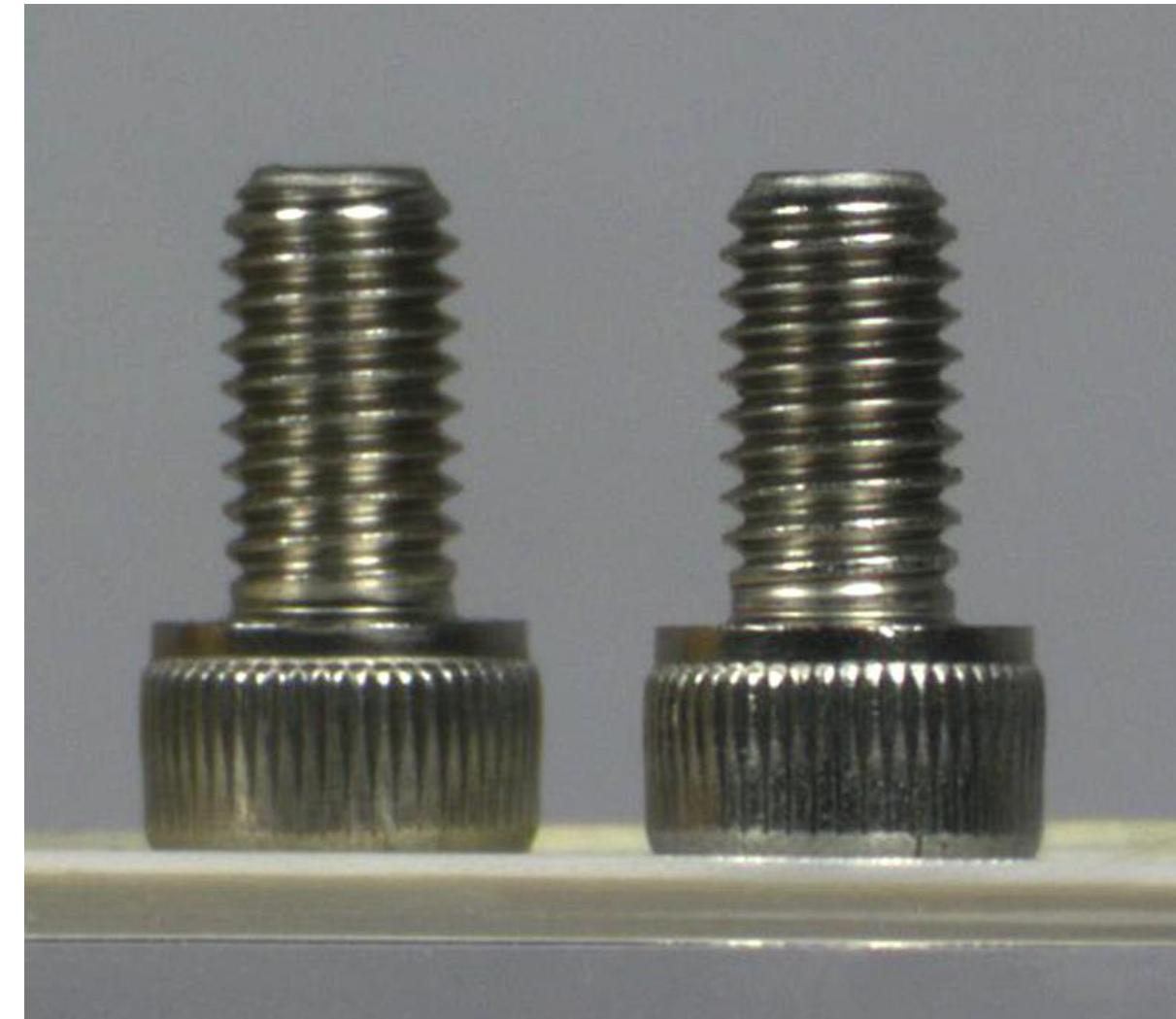


short focal length

Different cameras

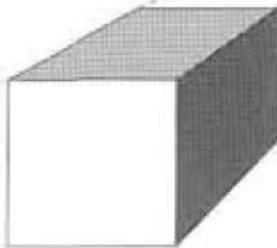


Perspective camera

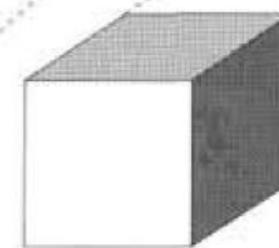


Orthographic camera

Camera is
close to object
small focal length



perspective



weak perspective

Camera is
far from object
large focal length

increasing focal length →

increasing distance from camera →



When can you assume a weak perspective camera model?



When can you assume a weak perspective camera model?

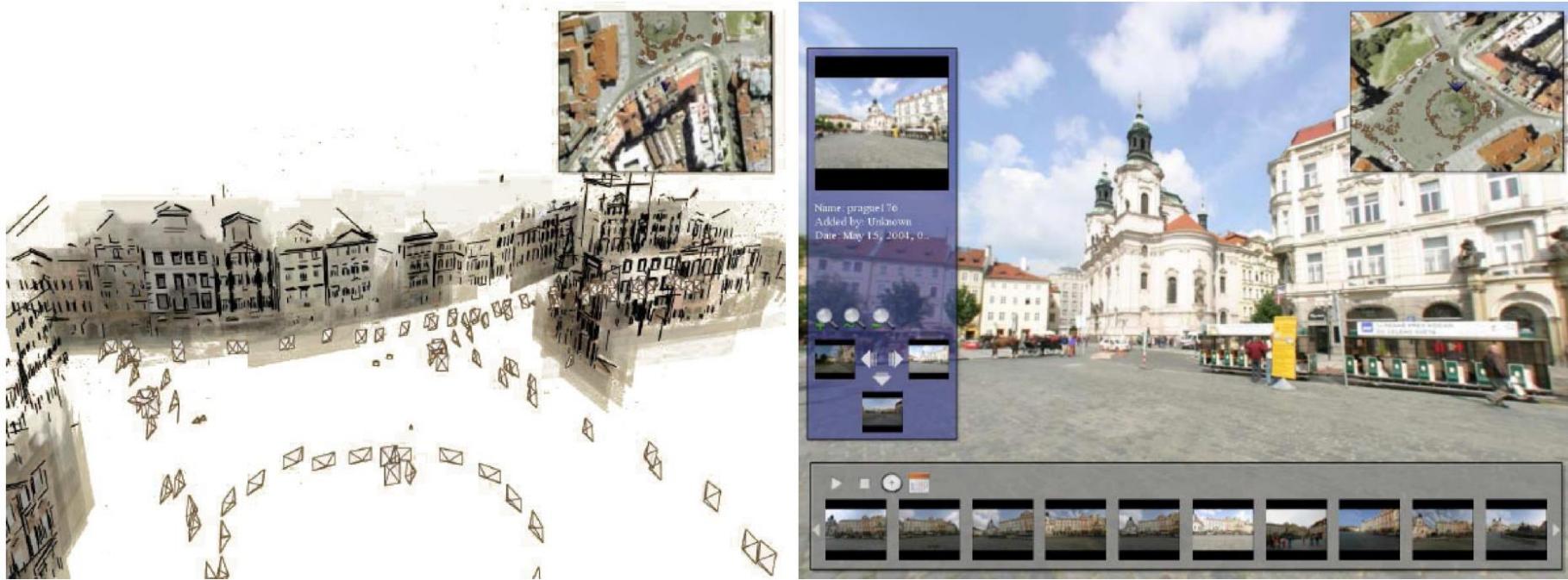


all the mountains are roughly 'far away'

Pose estimation

	Structure (scene geometry)	Motion (camera geometry)	Measurements
Pose Estimation	known	estimate	3D to 2D correspondences
Triangulation	estimate	known	2D to 2D correspondences
Reconstruction	estimate	estimate	2D to 2D correspondences

Pose Estimation



Given a single image,
estimate the exact position of the photographer

3D Pose Estimation

(Resectioning, Geometric Calibration, Perspective n-Point)

Given a set of matched points

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

point in 3D
space point in the
image

and camera model

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

projection
model

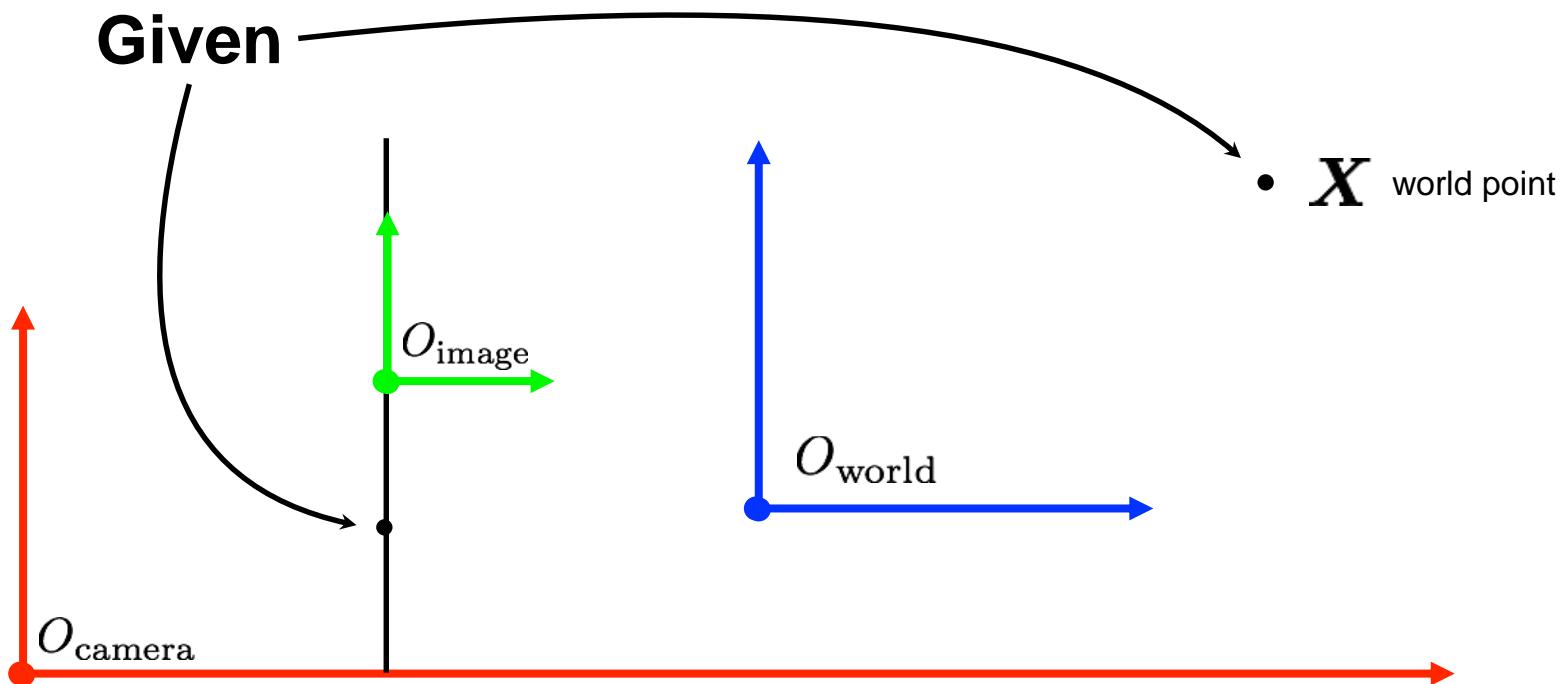
parameters

Camera
matrix

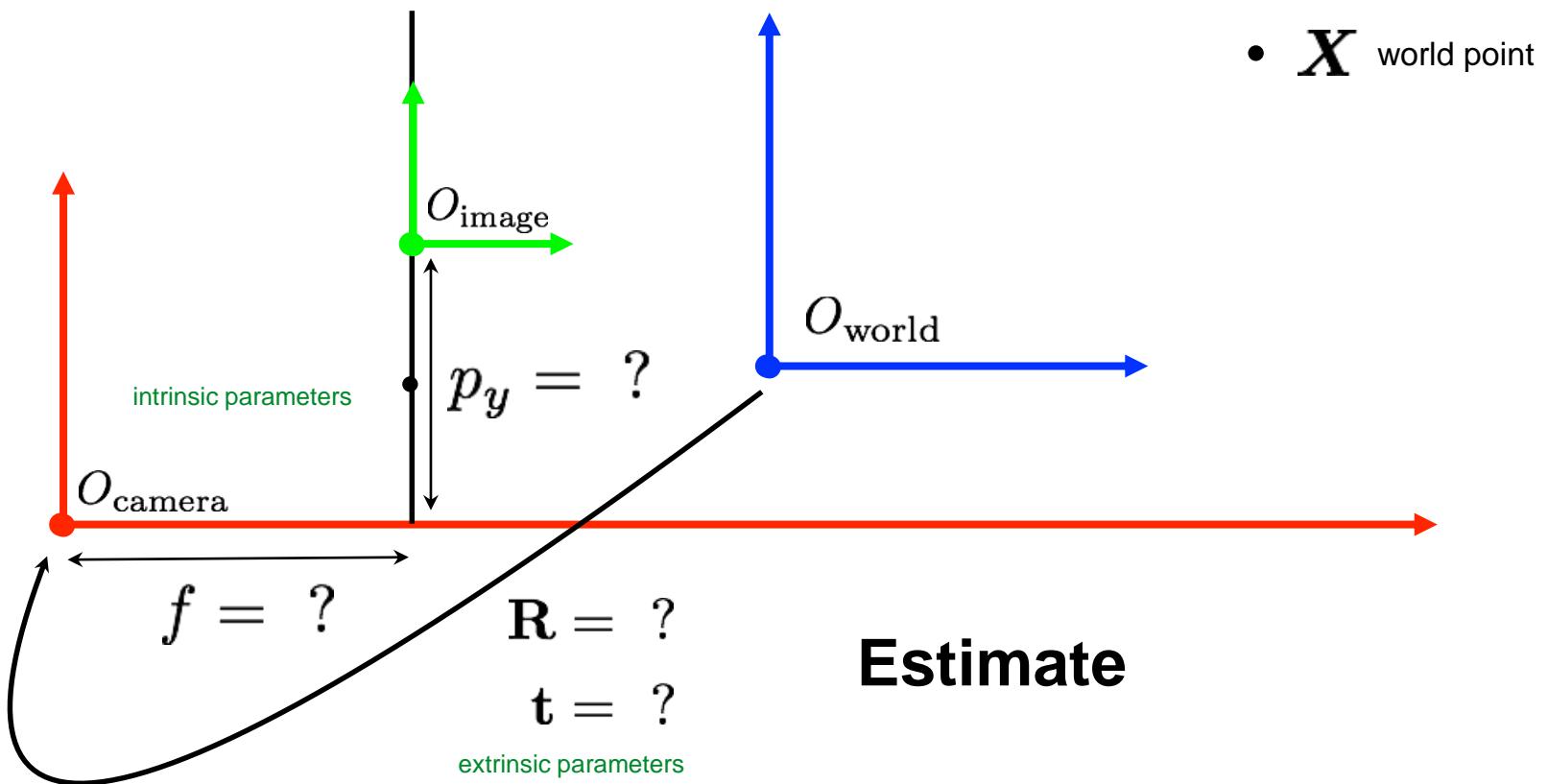
Find the (pose) estimate of

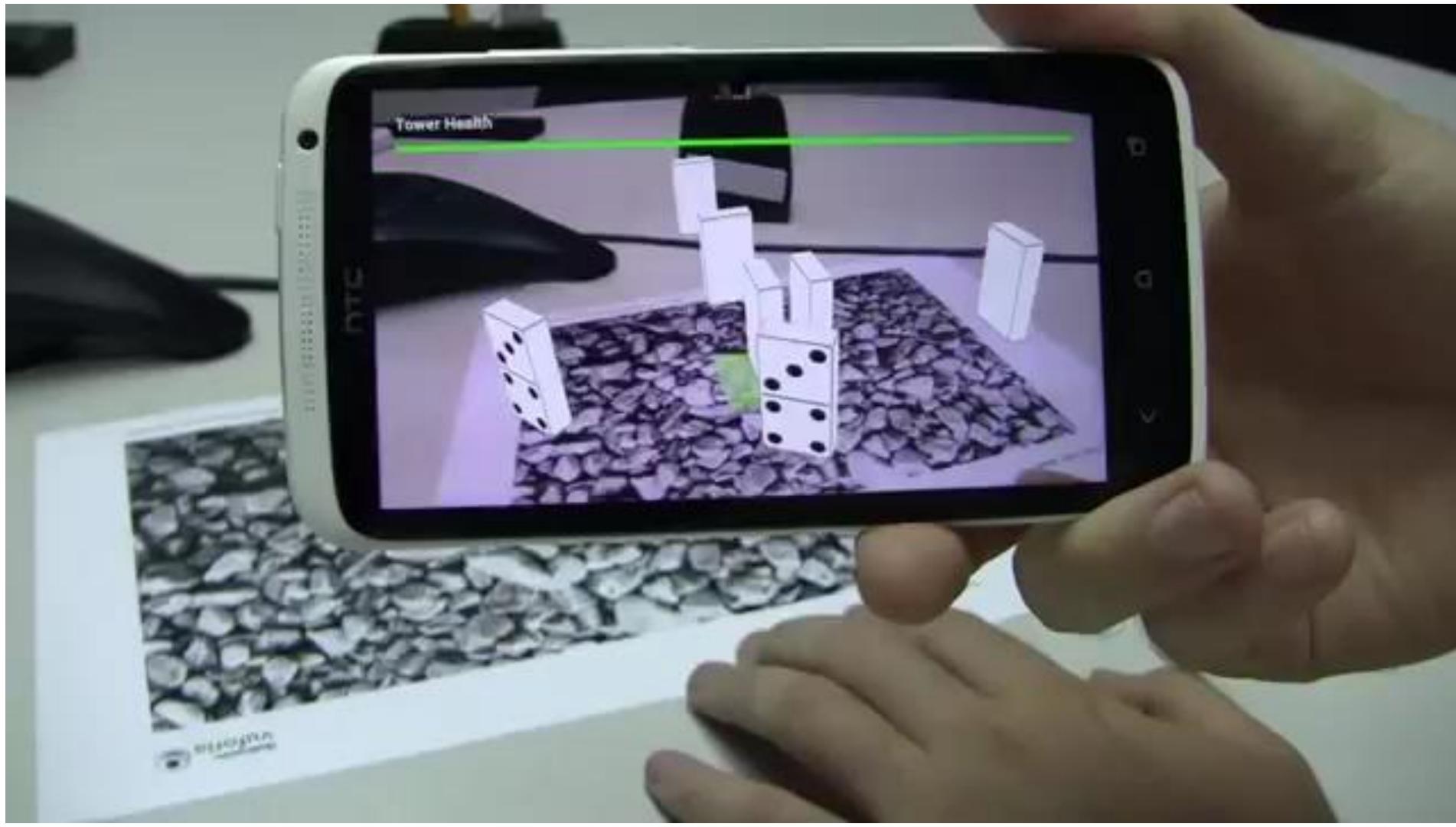
P

What is Pose Estimation?

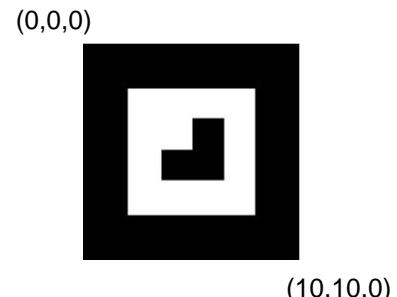


What is Pose Estimation?

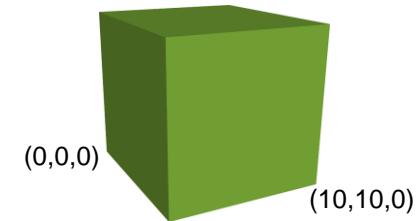




3D locations of planar marker features are known in advance



3D content prepared in advance



Simple AR program

1. Compute point correspondences (2D and AR tag)
2. Estimate the pose of the camera \mathbf{P}
3. Project 3D content to image plane using \mathbf{P}





PEPSI MAX
PRESENTS

References

Basic reading:

- Szeliski textbook, Section 2.1.5, 6.2.

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.
- Torralba and Freeman, “Accidental Pinhole and Pinspeck Cameras,” CVPR 2012.
the eponymous paper discussed in the slides.