Camera Models

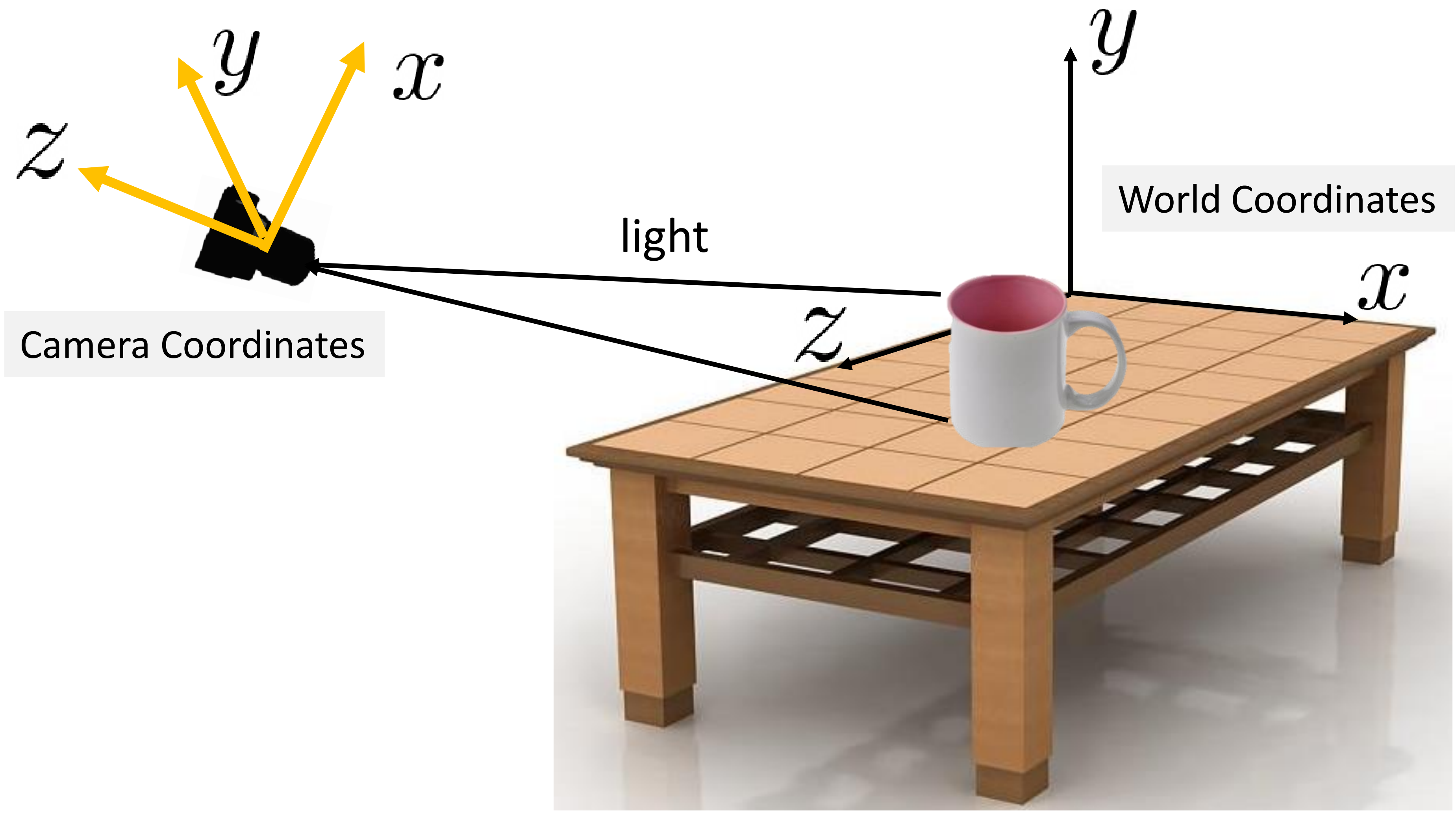
CS 6384 Computer Vision

Professor Yu Xiang

The University of Texas at Dallas

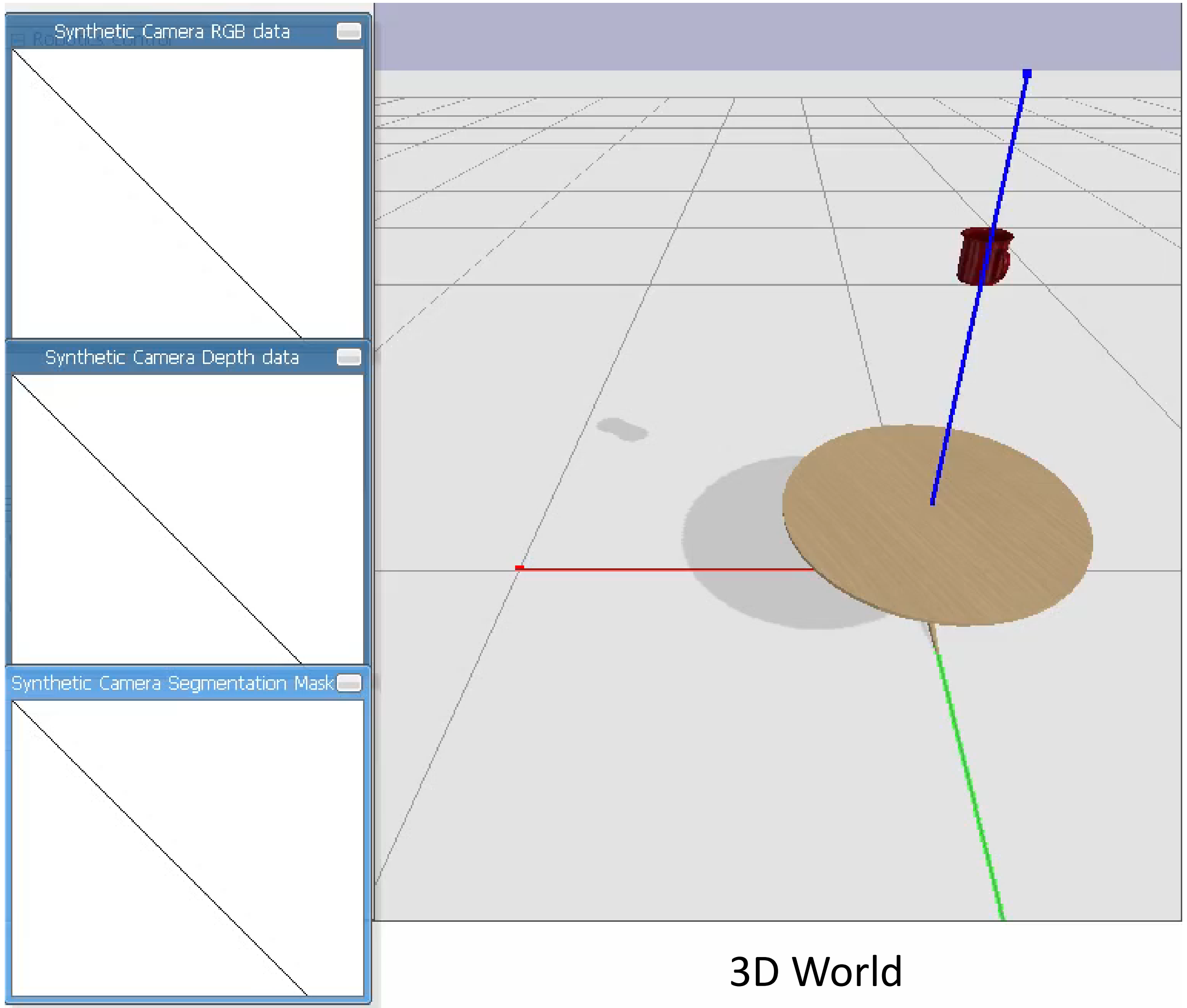
Some slides of this lecture are courtesy Silvio Savarese

A Camera in the 3D World



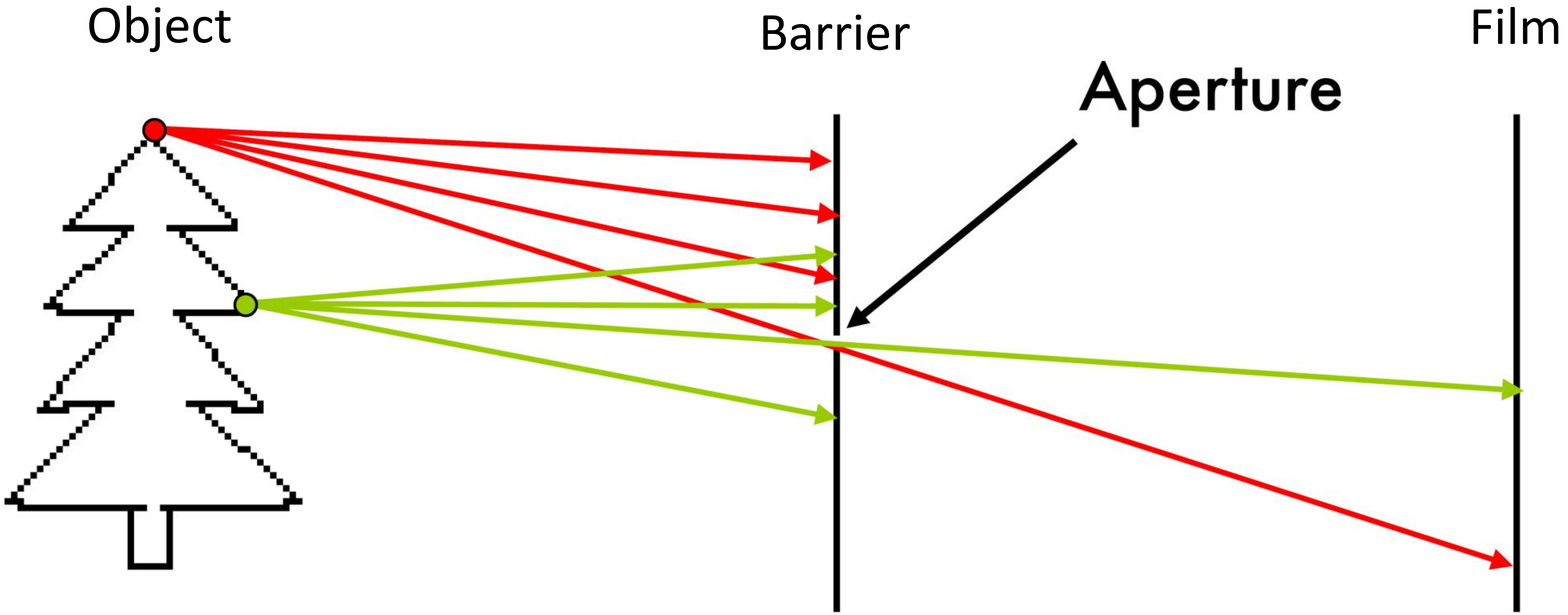
PyBullet with a Camera

Camera View

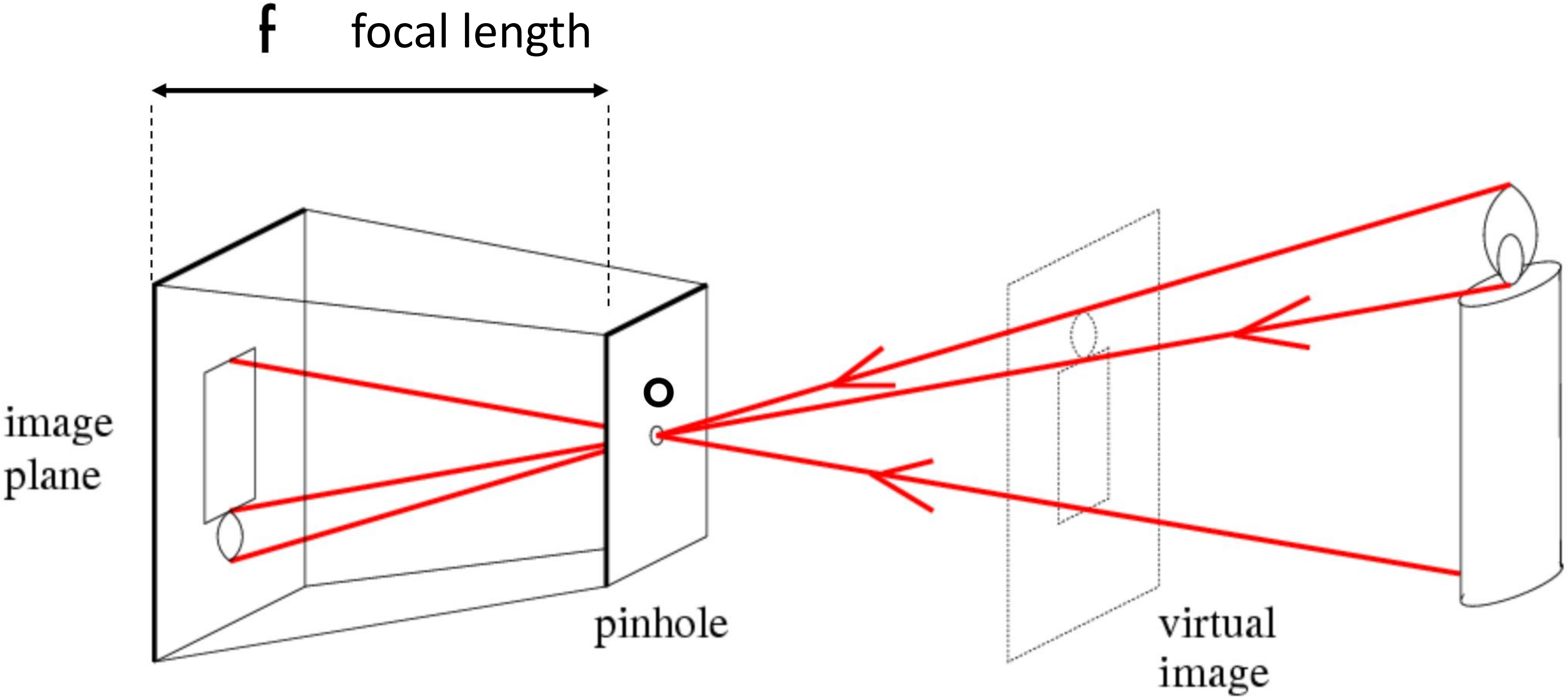


3D World

Pinhole Camera



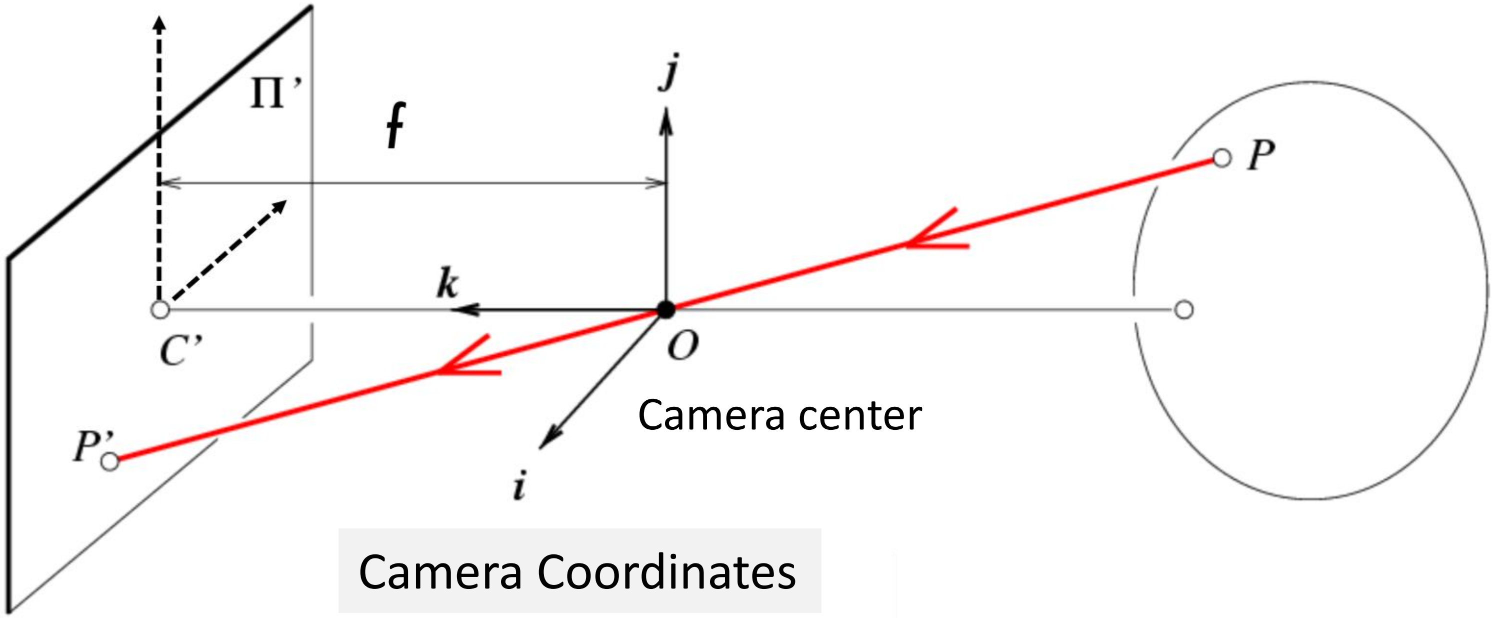
Pinhole Camera

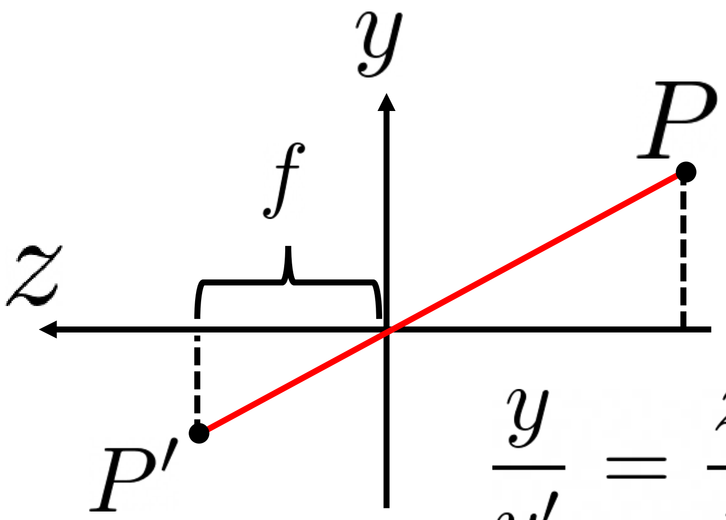


Rotate the image plane by 180°

Cannot be implemented in practice
Useful for theoretic analysis

Central Projection in Camera Coordinates





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

Camera coordinates $P = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \rightarrow P' = \begin{bmatrix} x' \\ y' \end{bmatrix}$

$$z' = f$$

$$\begin{cases} x' = f\,\dfrac{x}{z} \\ y' = f\,\dfrac{y}{z} \end{cases}$$

Central Projection with Homogeneous Coordinates

$$P = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \rightarrow P' = \begin{bmatrix} x' \\ y' \end{bmatrix}$$

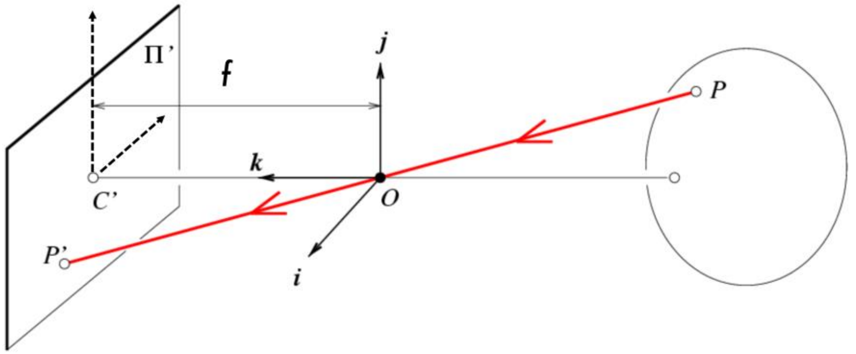
$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} \longrightarrow \begin{bmatrix} f\frac{x}{z} \\ f\frac{y}{z} \end{bmatrix}$$

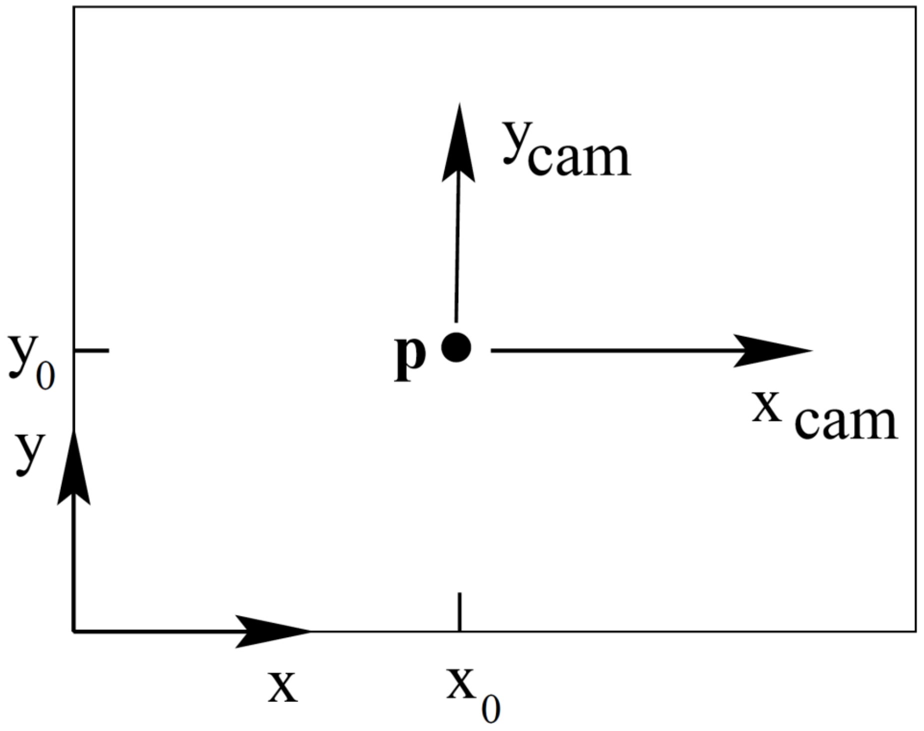
Central projection

$$\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \longrightarrow \begin{bmatrix} fx \\ fy \\ z \end{bmatrix} = \begin{bmatrix} f & & & 0 \\ & f & & 0 \\ & & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

3x4 matrix

Principal Point Offset





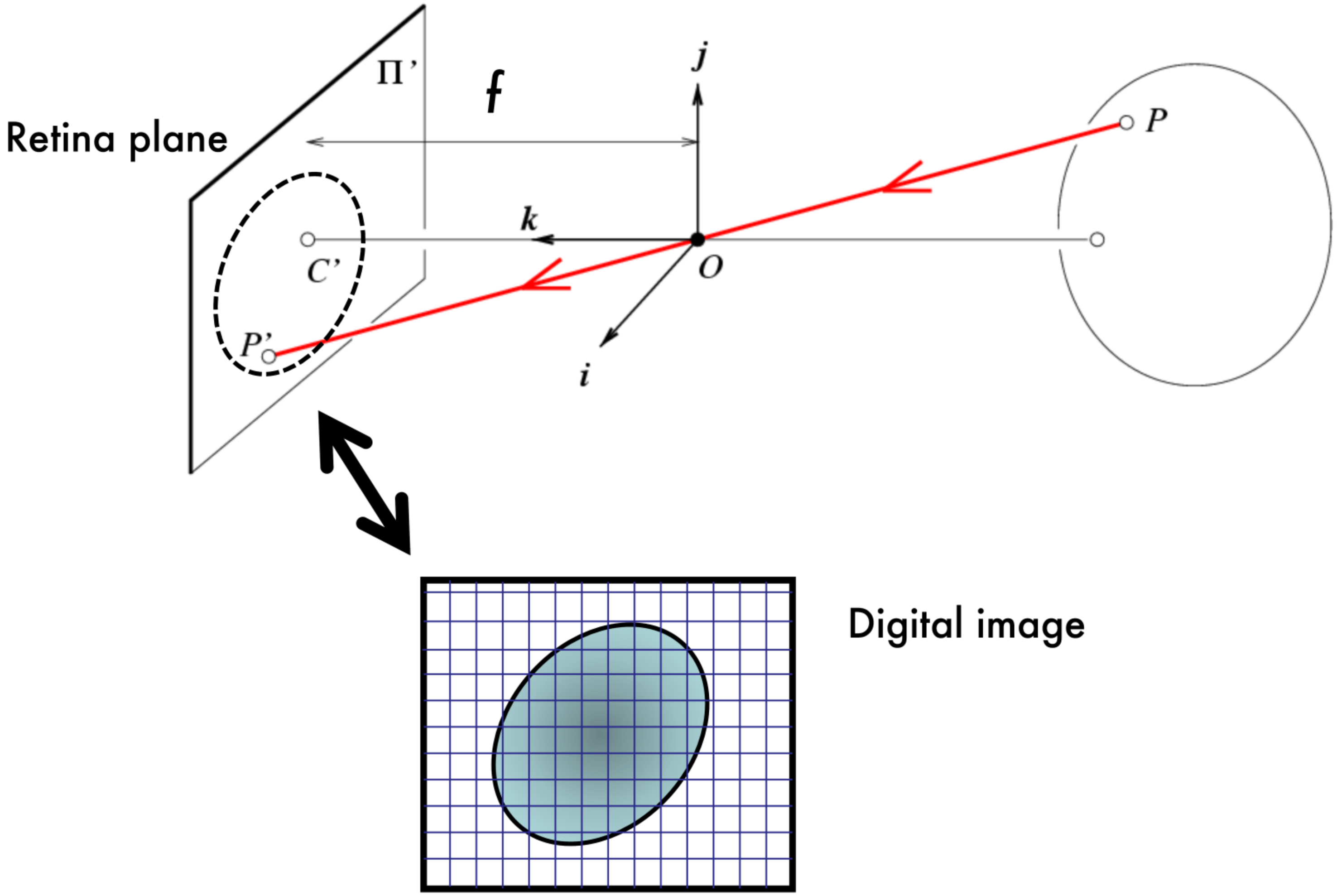
Principle point: projection of the camera center

Principal point $\mathbf{p} = (p_x, p_y)$

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} \longrightarrow \begin{bmatrix} f\frac{x}{z} + p_x \\ f\frac{y}{z} + p_y \end{bmatrix}$$

$$\begin{bmatrix} f & & p_x & 0 \\ & f & p_y & 0 \\ & & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

From Metric to Pixels



Pixels, bottom-left coordinate systems

From Metric to Pixels

- Metric space, i.e., meters

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

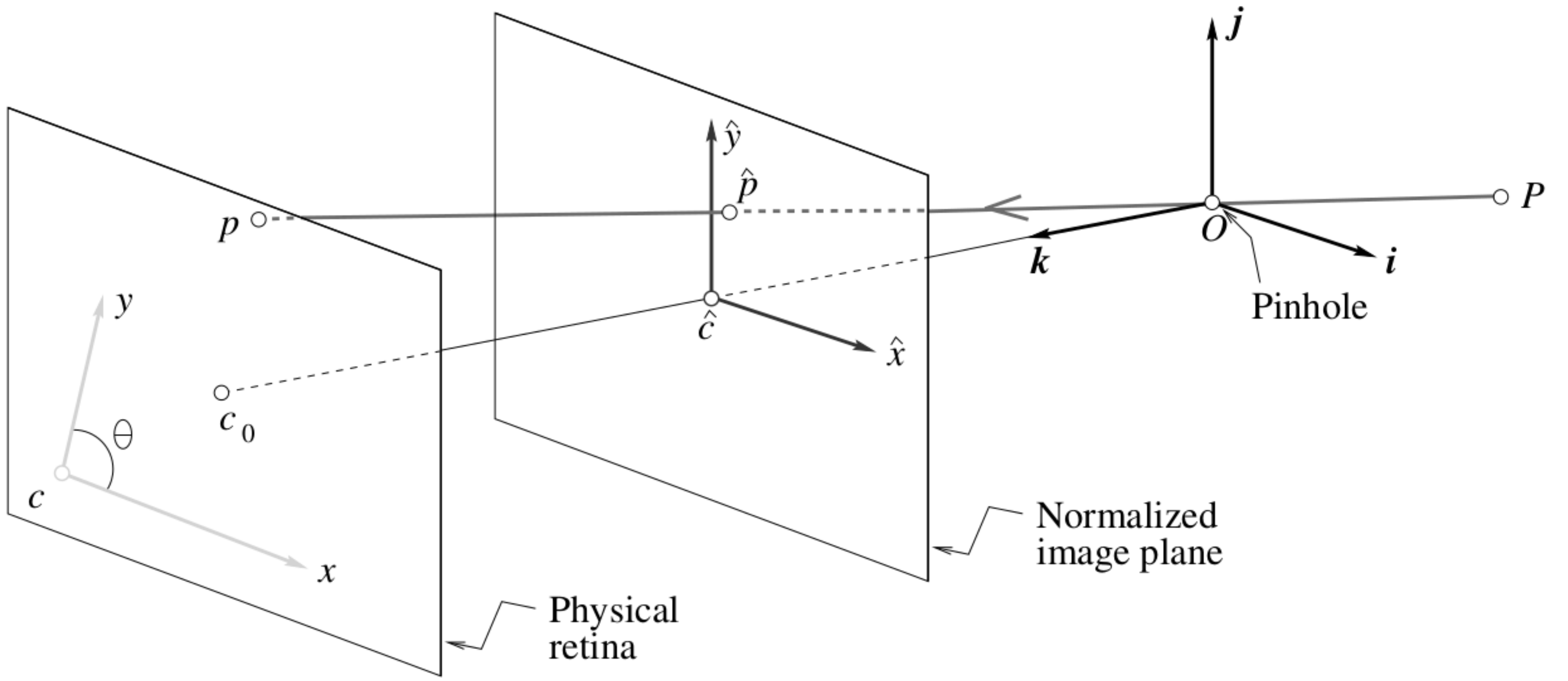
- Pixel space

$$\begin{bmatrix} \alpha_x & & x_0 & 0 \\ & \alpha_y & y_0 & 0 \\ & & 1 & 0 \end{bmatrix}$$

$$\alpha_x = f m_x$$
$$\alpha_y = f m_y$$
$$x_0 = p_x m_x$$
$$y_0 = p_y m_y$$

m_x, m_y Number of pixel per unit distance

Axis Skew



The skew parameter will be zero for most normal cameras.

$$\begin{bmatrix} \alpha_x & & x_0 & 0 \\ & \alpha_y & y_0 & 0 \\ & & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \longrightarrow \begin{bmatrix} \alpha_x \frac{x}{z} + x_0 \\ \alpha_y \frac{y}{z} + y_0 \end{bmatrix}$$

$$\begin{bmatrix} \alpha_x & -\alpha_x \cot(\theta) & x_0 & 0 \\ & \frac{\alpha_y}{\sin(\theta)} & y_0 & 0 \\ & & 1 & 0 \end{bmatrix}$$

https://blog.immenselyhappy.com/post/camera-axis-skew/

Camera Intrinsics

$$\begin{bmatrix} \alpha_x & -\alpha_x \cot(\theta) & x_0 & 0 \\ & \frac{\alpha_y}{\sin(\theta)} & y_0 & 0 \\ & & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

Camera intrinsics

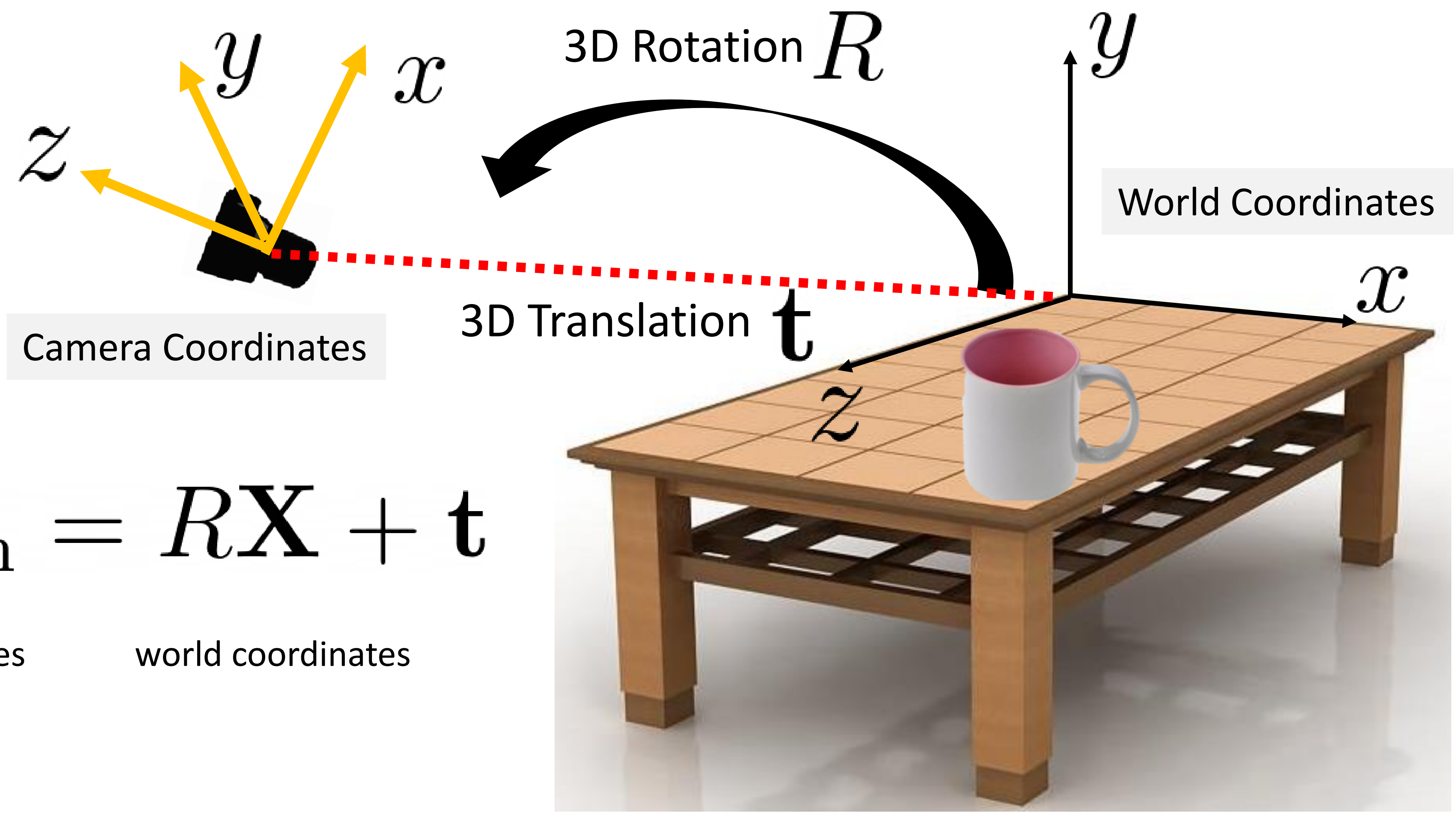
$$K = \begin{bmatrix} \alpha_x & s & x_0 \\ & \alpha_y & y_0 \\ & & 1 \end{bmatrix}$$

$$\mathbf{x} = K[I|\mathbf{0}]\mathbf{X}_{\text{cam}}$$

3x1 3x3 3x4 4x1

Homogeneous coordinates

Camera Extrinsics: Camera Rotation and Translation



$$\mathbf{X}_{\text{cam}} = R\mathbf{X} + \mathbf{t}$$

camera coordinates world coordinates

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

- Homogeneous coordinates

$$\mathbf{x} = K[I|\mathbf{0}]\mathbf{X}_{\text{cam}}$$
$$= K[R|\mathbf{t}]\mathbf{X}$$

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

3x1 3x3 3x4 4x1

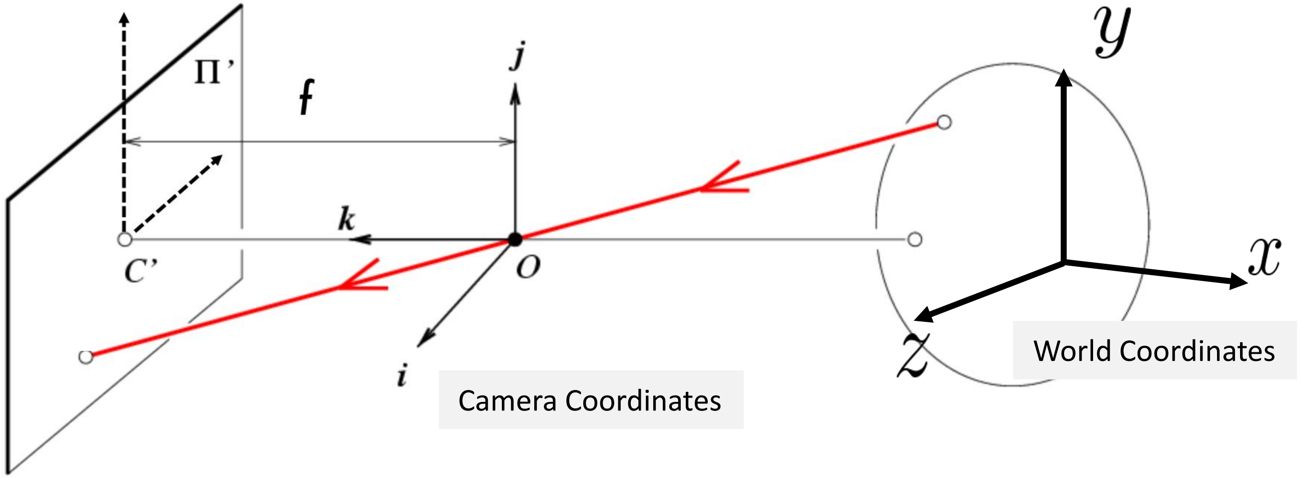
Image coordinates

Camera intrinsics

Camera extrinsics: rotation and translation

World coordinates

Back-projection to a Ray in the World Coordinate



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

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

- The camera center O is on the ray
- $P^+\mathbf{x}$ is on the ray

$$P^+ = P^T(PP^T)^{-1}$$

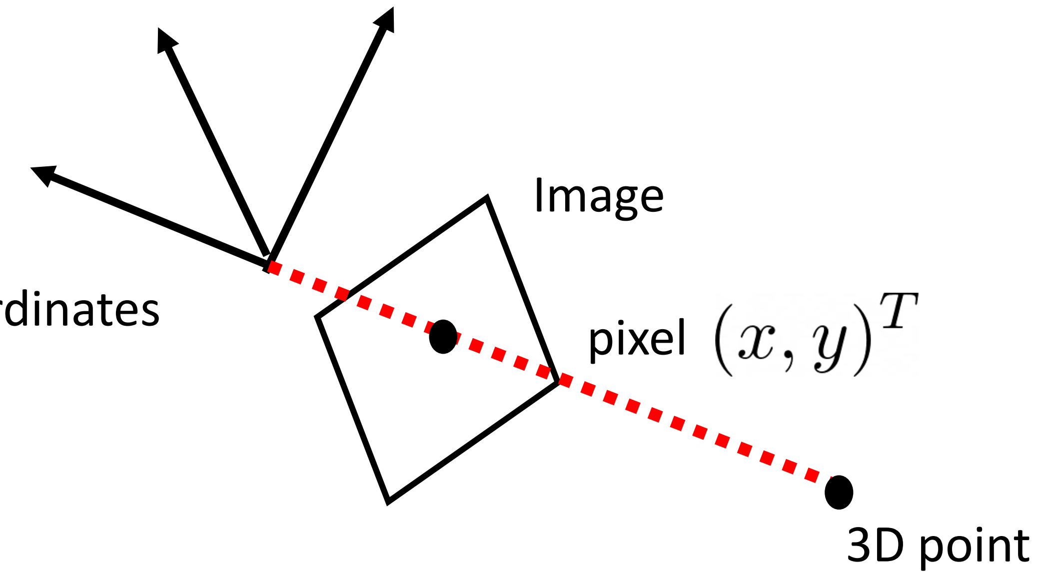
Pseudo-inverse

The ray can be written as

$$\mathbf{X}(\lambda) = (1-\lambda)P^+\mathbf{x} + \lambda O$$

- A pixel on the image backprojects to a ray in 3D

Back-projection to a 3D Point in Camera Coordinates



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

$$\mathbf{x} = K[I|\mathbf{0}]\mathbf{X}_{\text{cam}}$$

$$\begin{bmatrix} f_x & 0 & p_x \\ 0 & f_y & p_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

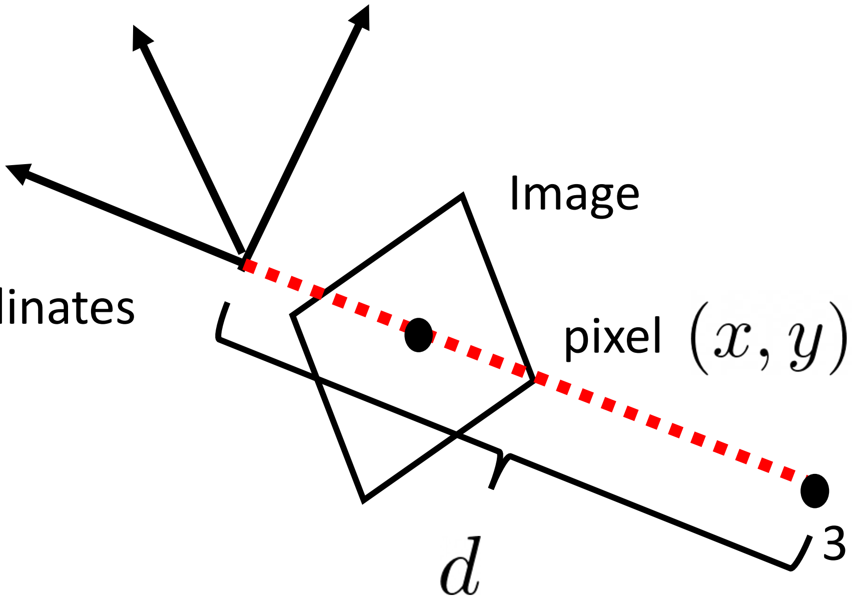
$$x = f\frac{X}{Z} + p_x \qquad \frac{X}{Z} = \frac{x - p_x}{f}$$

$$y = f\frac{Y}{Z} + p_y \qquad \frac{Y}{Z} = \frac{y - p_y}{f}$$

3D Point $\begin{bmatrix} \frac{x-p_x}{f} Z \\ \frac{y-p_y}{f} Z \\ Z \end{bmatrix}$

We need to know the depth of the pixel

Back-projection to a 3D Point in Camera Coordinates



3D camera coordinates $\begin{bmatrix} d\frac{x-p_x}{f_x} \\ d\frac{y-p_y}{f_y} \\ d \end{bmatrix}$

Equivalently

$$\mathbf{x} = K[I|\mathbf{0}]\mathbf{X}_{\text{cam}}$$

$$K^{-1}\mathbf{x}$$

3D point with depth d : $dK^{-1}\mathbf{x}$

$$K = \begin{bmatrix} f_x & 0 & p_x \\ 0 & f_y & p_y \\ 0 & 0 & 1 \end{bmatrix}$$

3D Point $\begin{bmatrix} \frac{x-p_x}{f}Z \\ \frac{y-p_y}{f}Z \\ Z \end{bmatrix}$

The Pinhole Camera Model

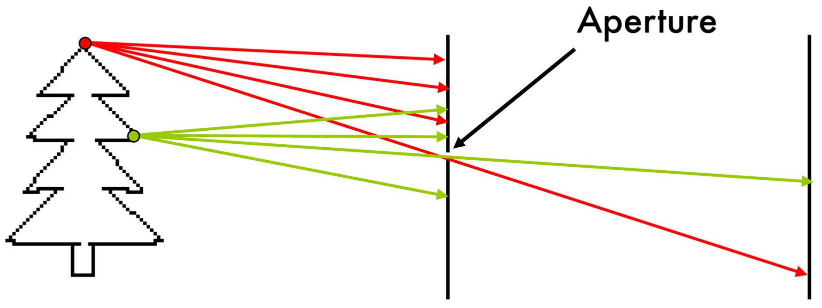
- Camera projection matrix: intrinsics and extrinsics

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

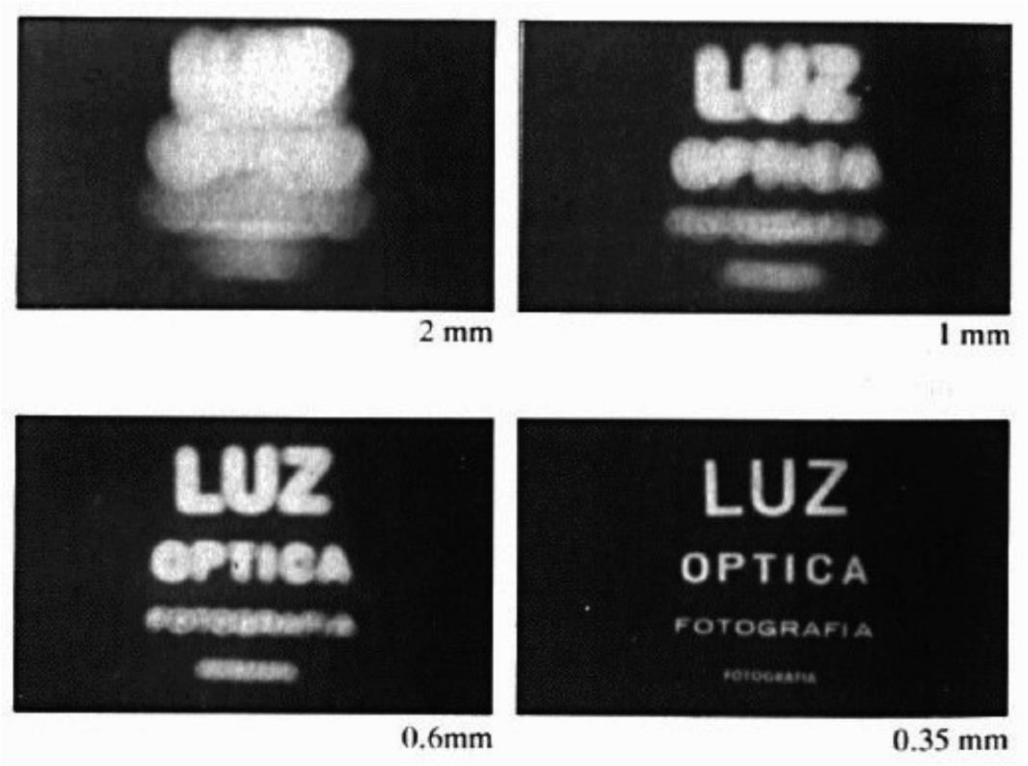
3x3 — Camera intrinsics

3x4 — Camera extrinsics: rotation and translation

Aperture Size of Pinhole Camera



Shrinking aperture size



What happen if the aperture is too small?

- Less light passes through
- Adding lenses

Lenses

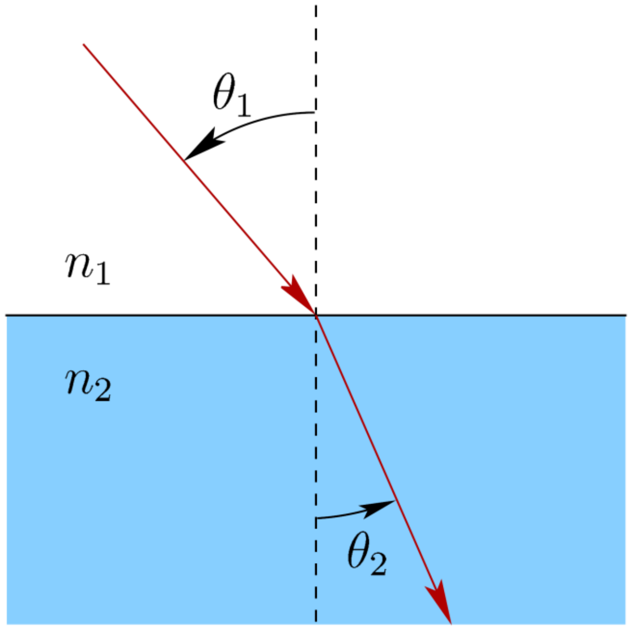
(a) (b)

Figure 4.8: (a) The earliest known artificially constructed lens, which was made between 750 and 710 BC in ancient Assyrian Nimrud. It is not known whether this artifact was purely ornamental or used to produce focused images. Picture from the British Museum. (b) A painting by Conrad con Soest from 1403, which shows the use of reading glasses for an elderly male.

Snell's Law

- How much rays of light bend when entering and exiting a transparent material
- Refractive index of a material $n = \frac{c}{s}$ ← c: Speed of light in a vacuum; s: Speed of light in the medium
  - Air n = 1.000293, water n = 1.33
  - Crown glass n = 1.523
- Snell's Law

$$n_1 < n_2$$

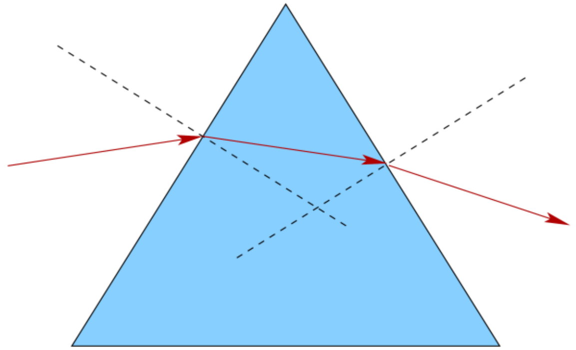


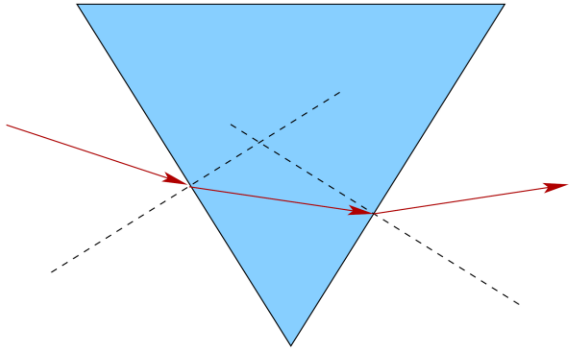
$$n_1 \sin \theta_1 = n_2 \sin \theta_2$$

$$\theta_2 = \sin^{-1}\left(\frac{n_1 \sin \theta_1}{n_2}\right)$$

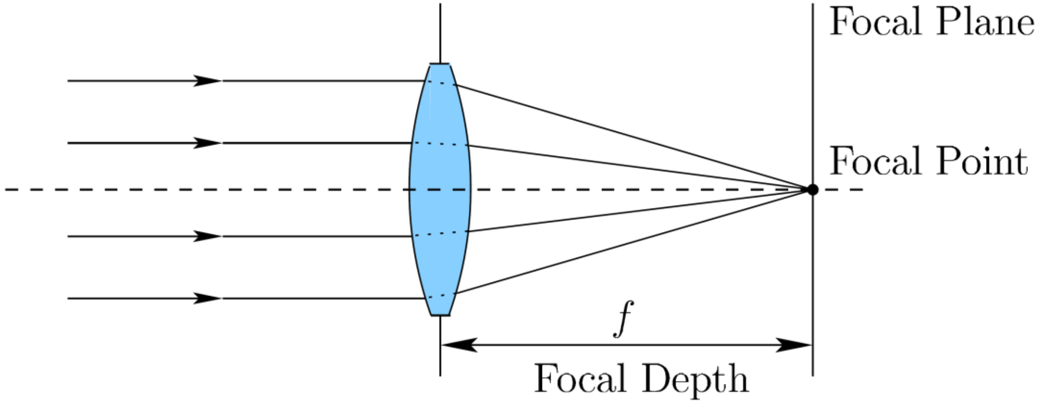
Convex Lenses

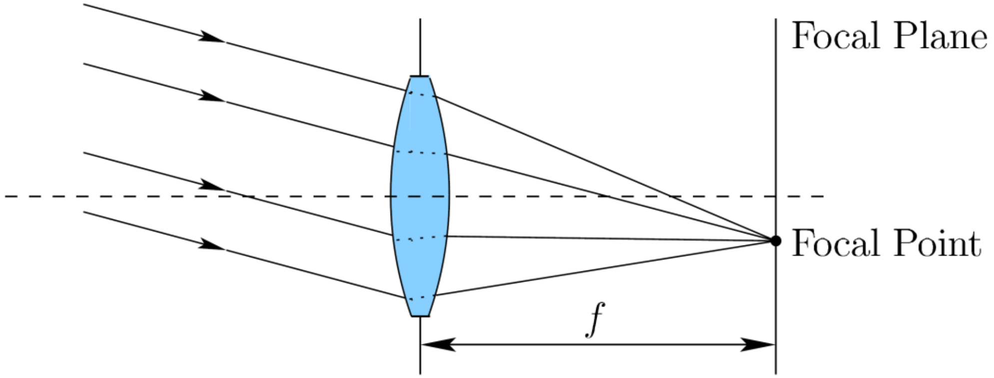
- Prisms





- A simple convex lens

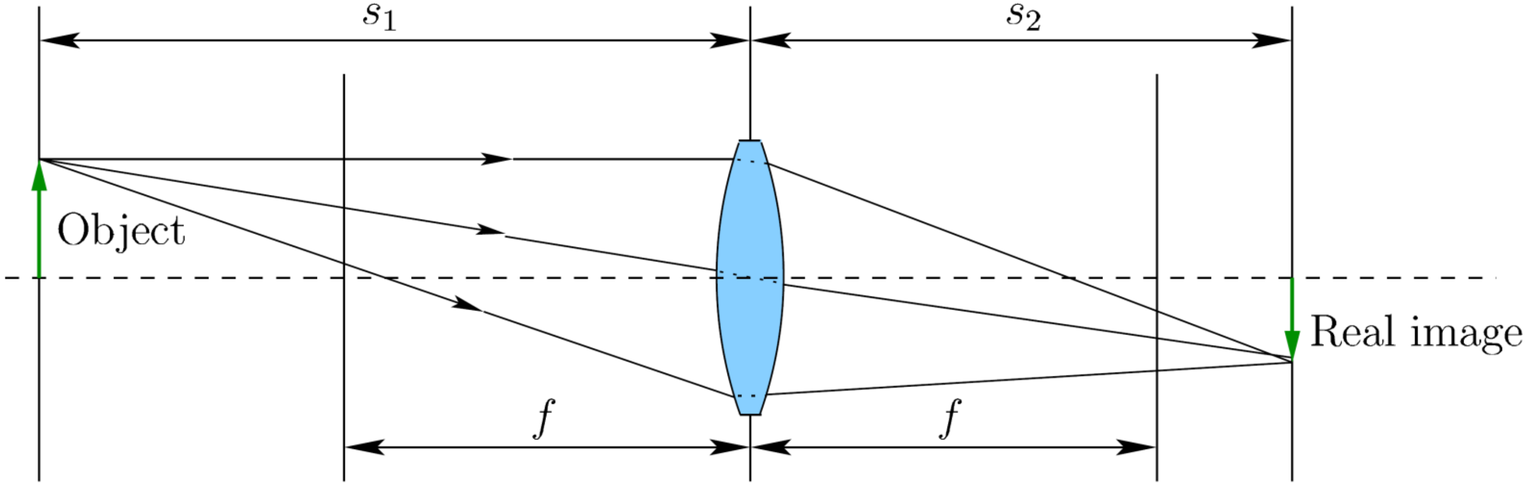




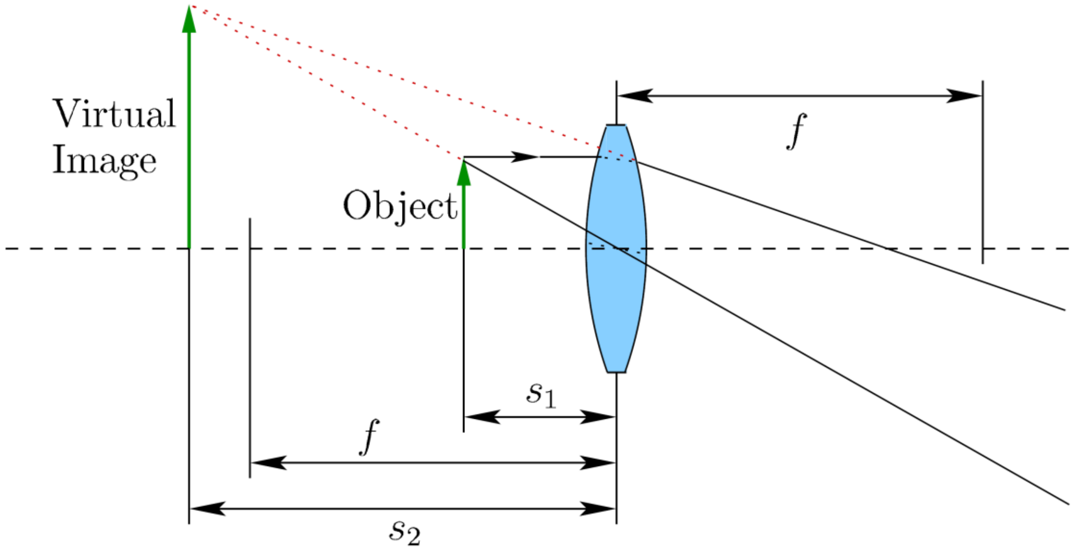
Convex Lenses

- Objects in distance
  - Cameras

$$\frac{1}{s_1} + \frac{1}{s_2} = \frac{1}{f}$$

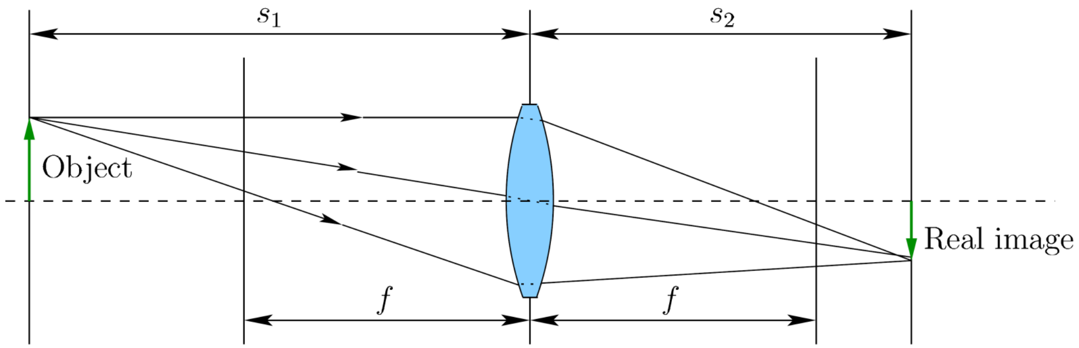


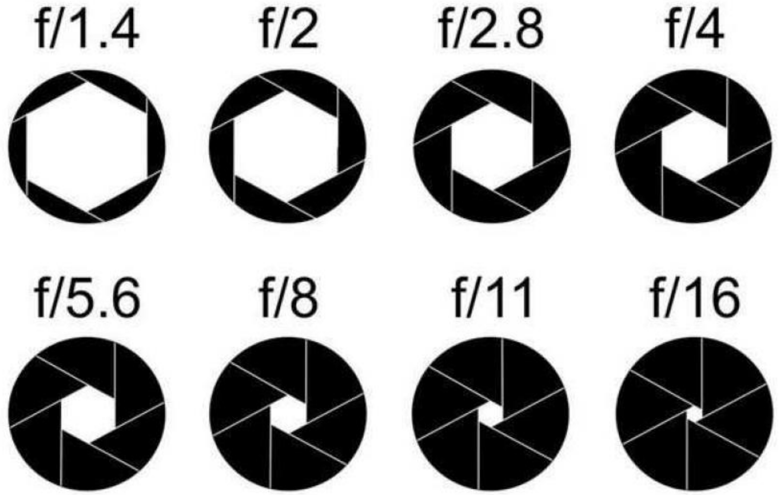
- Objects very close
  - Magnification
  - VR headsets



Controllable Aperture

- In the pinhole case, all depths are "in focus", but there may not enough lights
- When using a convex lens, it focuses objects at a single depth





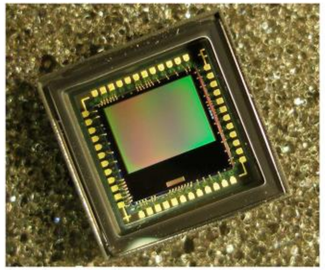
Behind the lens

Lens -> pinhole

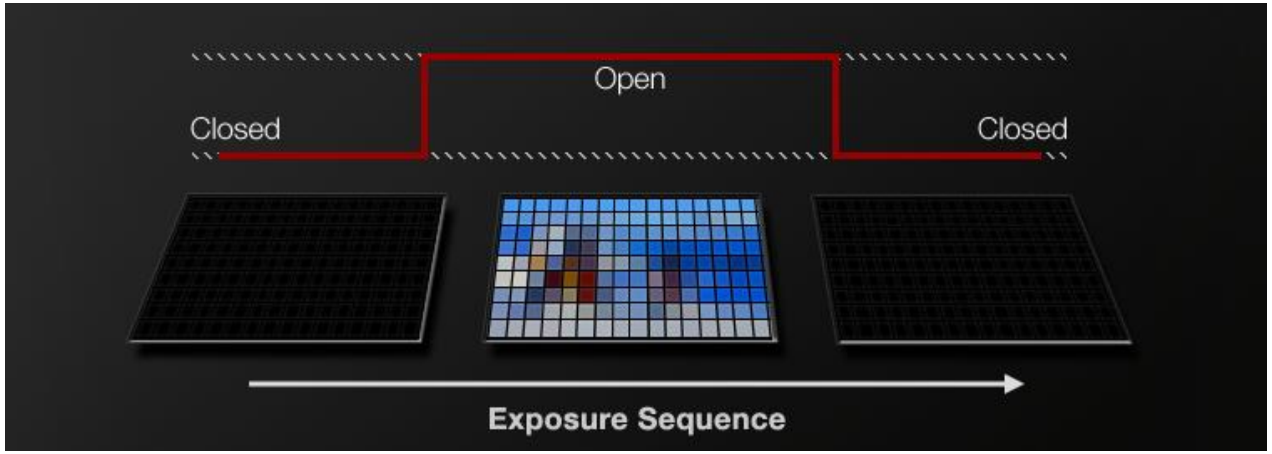
Figure 4.34: A spectrum of aperture settings, which control the amount of light that enters the lens. The values shown are called the *focal ratio* or *f-stop*.

Shutters

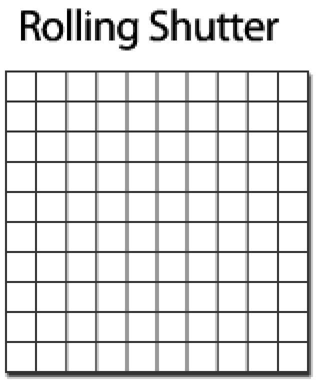
- Collecting photons for each pixel



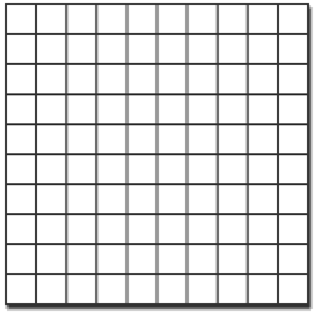
CMOS sensors

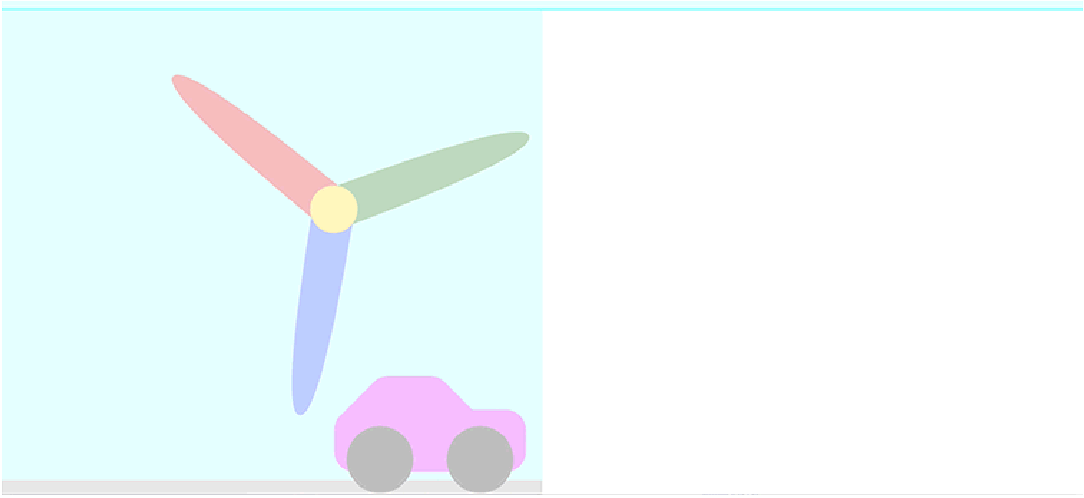


- Rolling shutter vs. Global shutter



Global shutter





Rolling shutter

Chromatic Aberration

- The speed of light through a medium depends on the wavelength
  - Solution: combining convex and concave lenses of different materials

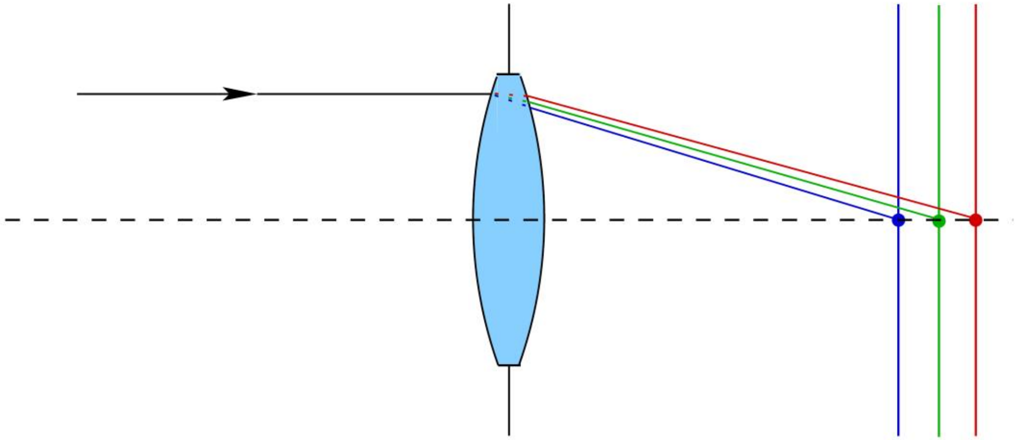
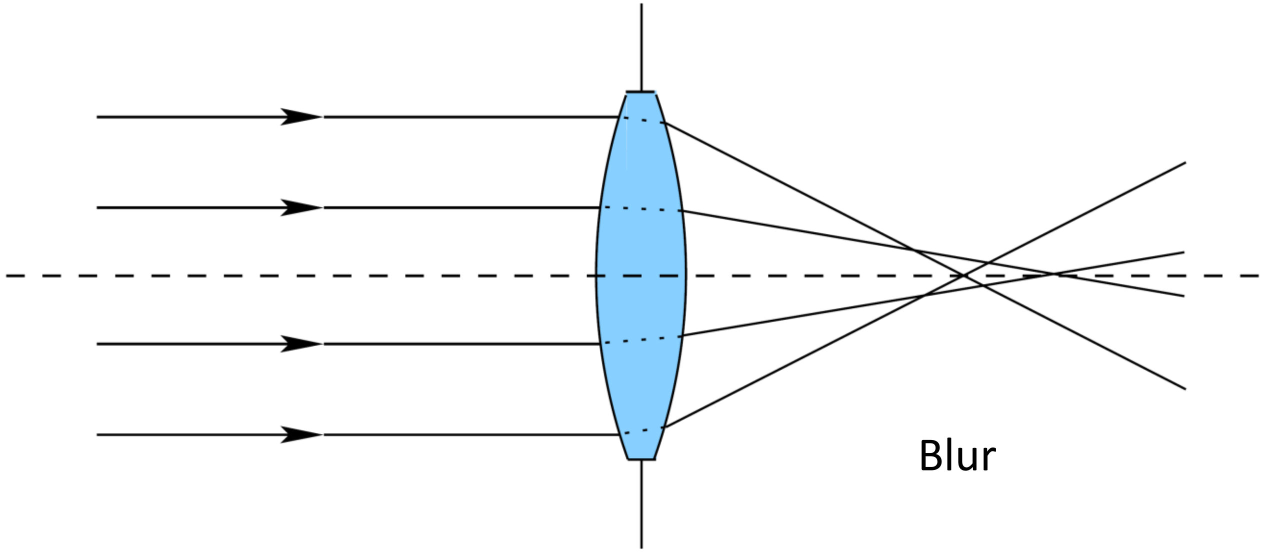


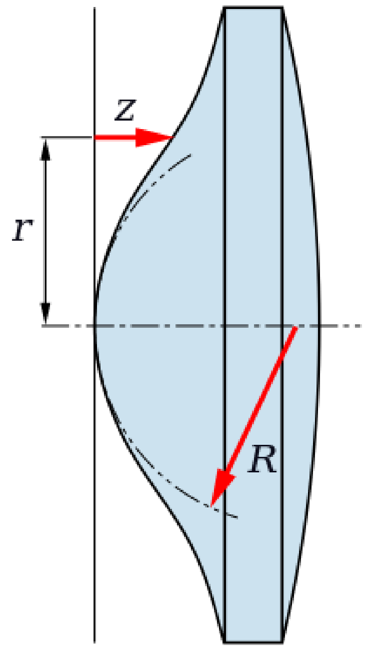
Figure 4.17: Chromatic aberration is caused by longer wavelengths travelin
quickly through the lens. The unfortunate result is a different focal plane f
wavelength or color.

Figure 4.18: The upper image is properly focused whereas the lower image suffers from chromatic aberration. (Figure by Stan Zurek, license CC-BY-SA-2.5.)

Spherical Aberration

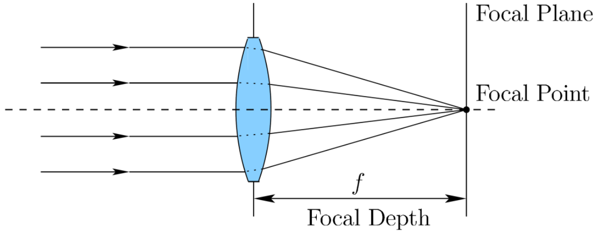
- Rays further away from the lens center being refracted more than rays near the center



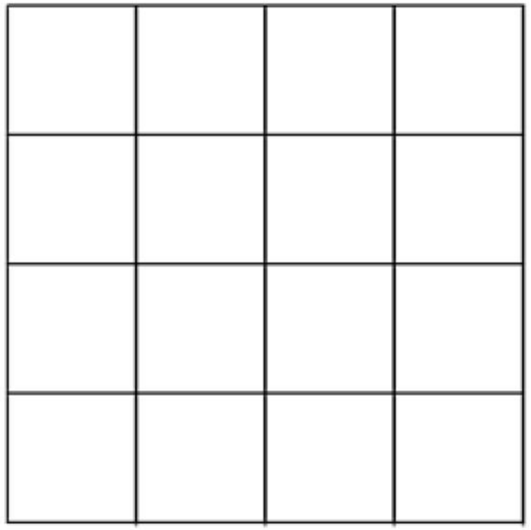


Aspheric lens

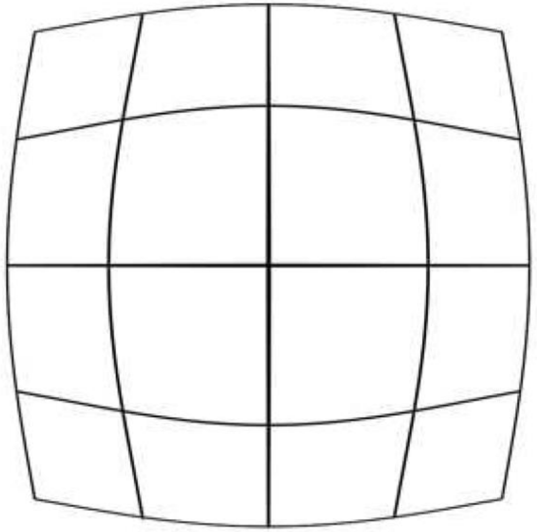
Optical Distortion



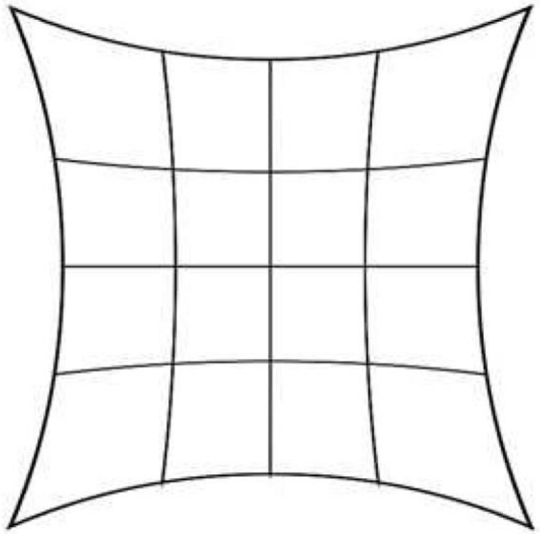
- The variation of refractive index towards the outer extremities of a rotational symmetric lens can cause magnification changes in the image space, depending on the distance from the principal axis.



No distortion



Barrel distortion (wide-angle lenses)



pincushion distortion (telephoto-lenses)

Angular Field of View (AFOV)

$$\mathrm{AFOV} = 2 \times \tan^{-1}\left(\frac{H}{2f}\right)$$

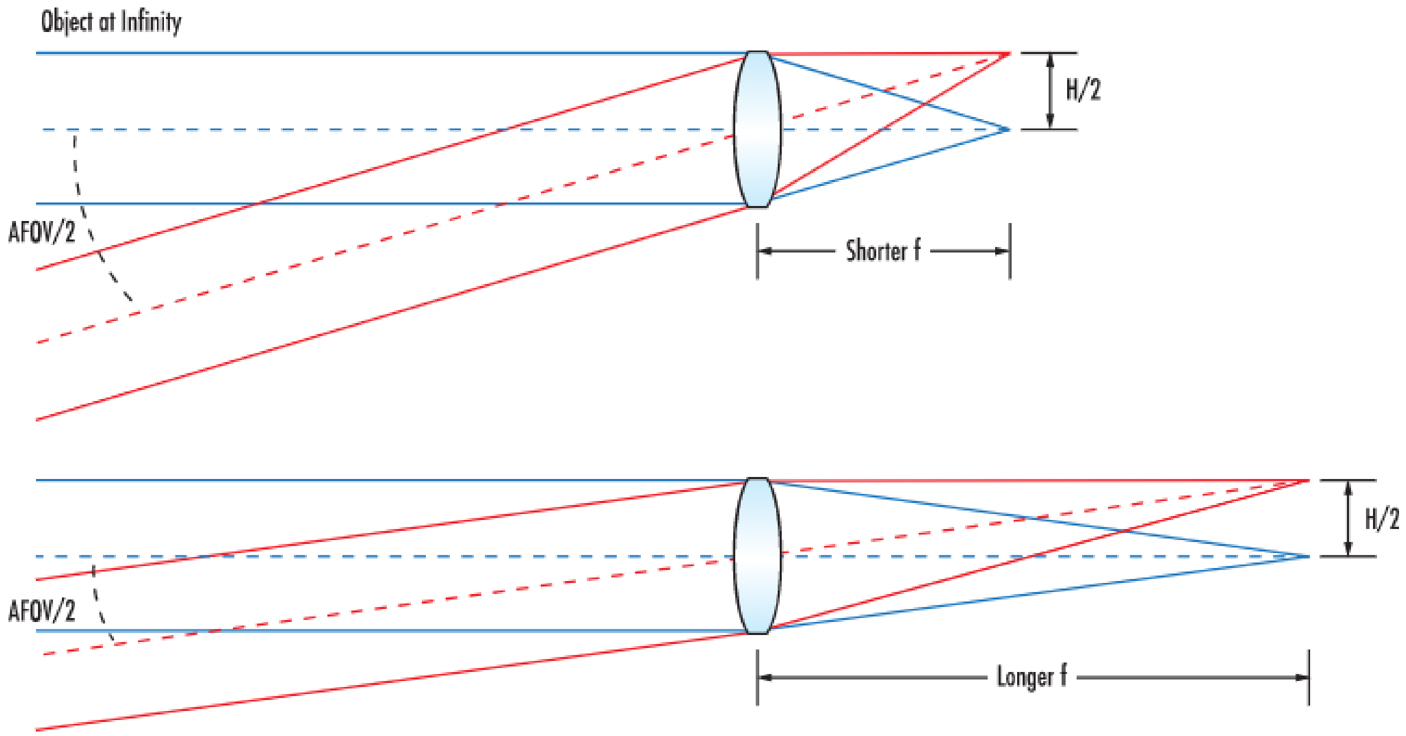


Figure 1: *For a given sensor size, H, shorter focal lengths produce wider AFOV's.*

Barrel distortion (wide-angle lenses)

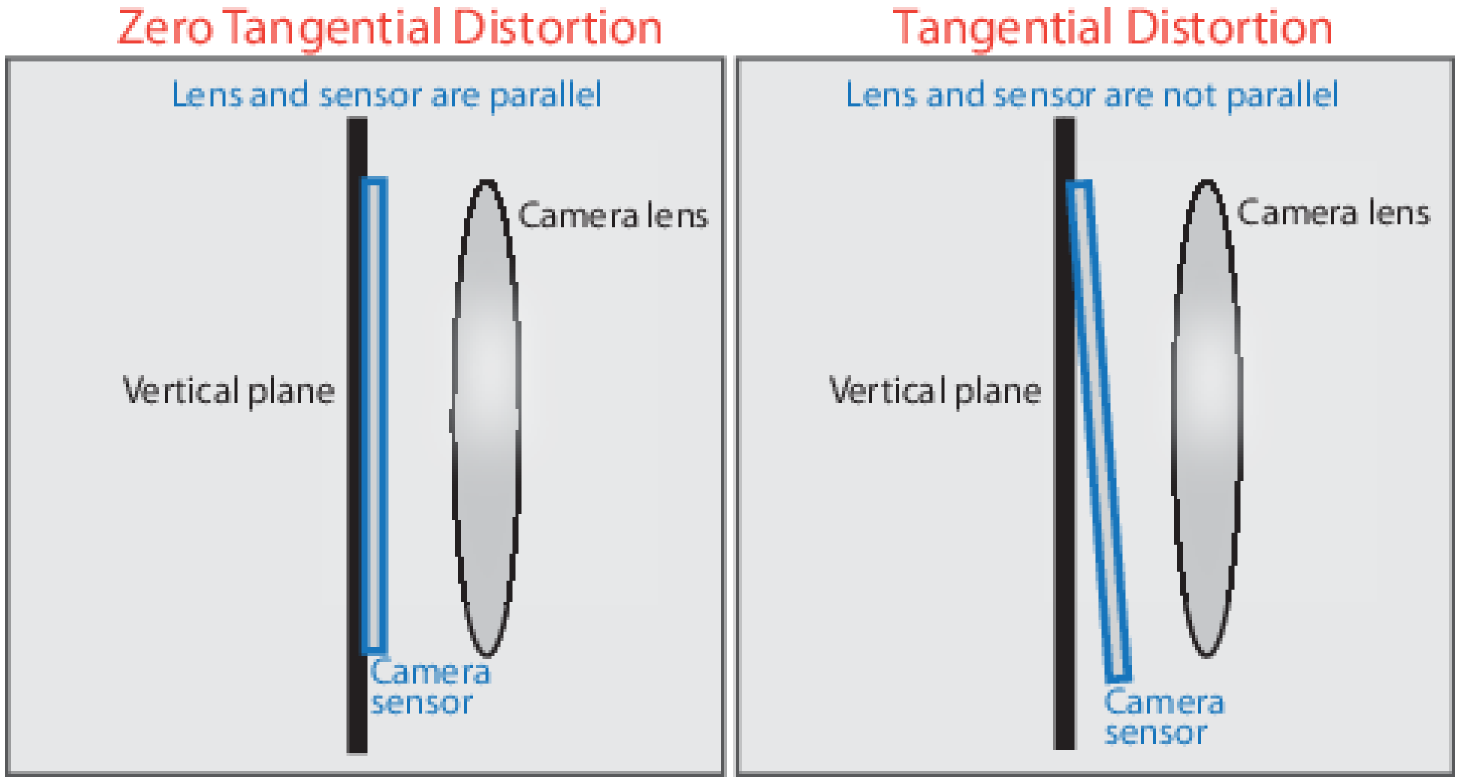
Pincushion distortion (telephoto-lenses)

Barrel Distortion of Fisheye Cameras

Figure 4.21: An image with barrel distortion, taken by a fish-eyed lens. (Image by Wikipedia user Ilveon.)

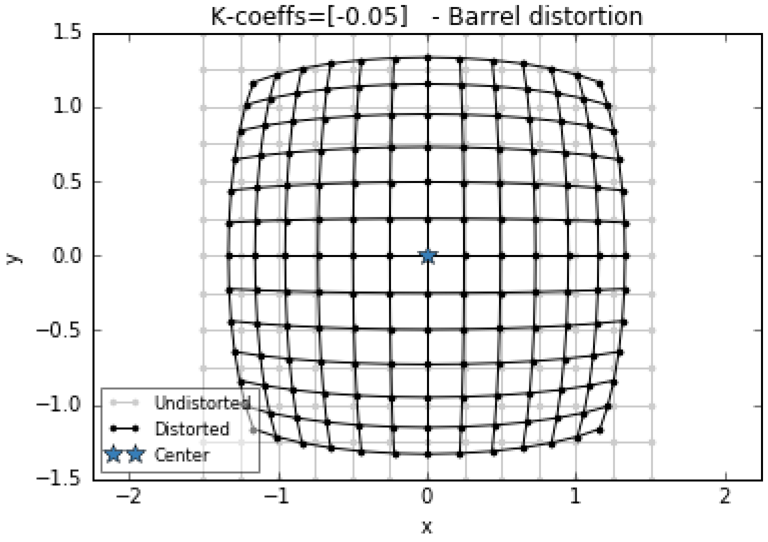
Tangential Distortion

- Camera sensor mis-alignment during the manufacturing process



Distortion Correction

- The Brown-Conrady distortion model [Wikipedia]



$$x_u = x_d + (x_d - x_c)(K_1 r^2 + K_2 r^4 + \cdots) + (P_1(r^2 + 2(x_d - x_c)^2) + 2P_2(x_d - x_c)(y_d - y_c))(1 + P_3 r^2 + P_4 r^4 \cdots)$$

$$y_u = y_d + (y_d - y_c)(K_1 r^2 + K_2 r^4 + \cdots) + (2P_1(x_d - x_c)(y_d - y_c) + P_2(r^2 + 2(y_d - y_c)^2))(1 + P_3 r^2 + P_4 r^4 \cdots),$$

where

- $(x_d,\ y_d)$ is the distorted image point as projected on image plane using specified lens;
- $(x_u,\ y_u)$ is the undistorted image point as projected by an ideal pinhole camera;
- $(x_c,\ y_c)$ is the distortion center;
- K_n is the n^{th} radial distortion coefficient;
- P_n is the n^{th} tangential distortion coefficient; and
- $r = \sqrt{(x_d - x_c)^2 + (y_d - y_c)^2}$, the Euclidean distance between the distorted image point and the distortion center.[3]

Use calibration tools to estimate these coefficients

Summary: Camera Models

- Camera projection matrix: intrinsics and extrinsics

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

- Lens distortion
  - Radial distortion coefficients $K_1, K_2, K_3, \ldots$
  - Tangential distortion coefficients $P_1, P_2, P_3, \ldots$

Further Reading

- Section 2.1, Computer Vision, Richard Szeliski
- Multiview Geometry in Computer Vision, Richard Hartley and Andrew Zisserman, Chapter 6, Camera Models
- Stanford CS231A: Computer Vision, From 3D Reconstruction to Recognition, Lecture 2 https://web.stanford.edu/class/cs231a/syllabus.html
- Image formation by lenses https://courses.lumenlearning.com/physics/chapter/25-6-image-formation-by-lenses/
- Distortion (Wikipedia) https://en.wikipedia.org/wiki/Distortion_(optics)