TU Delft 3Dgeoinfo

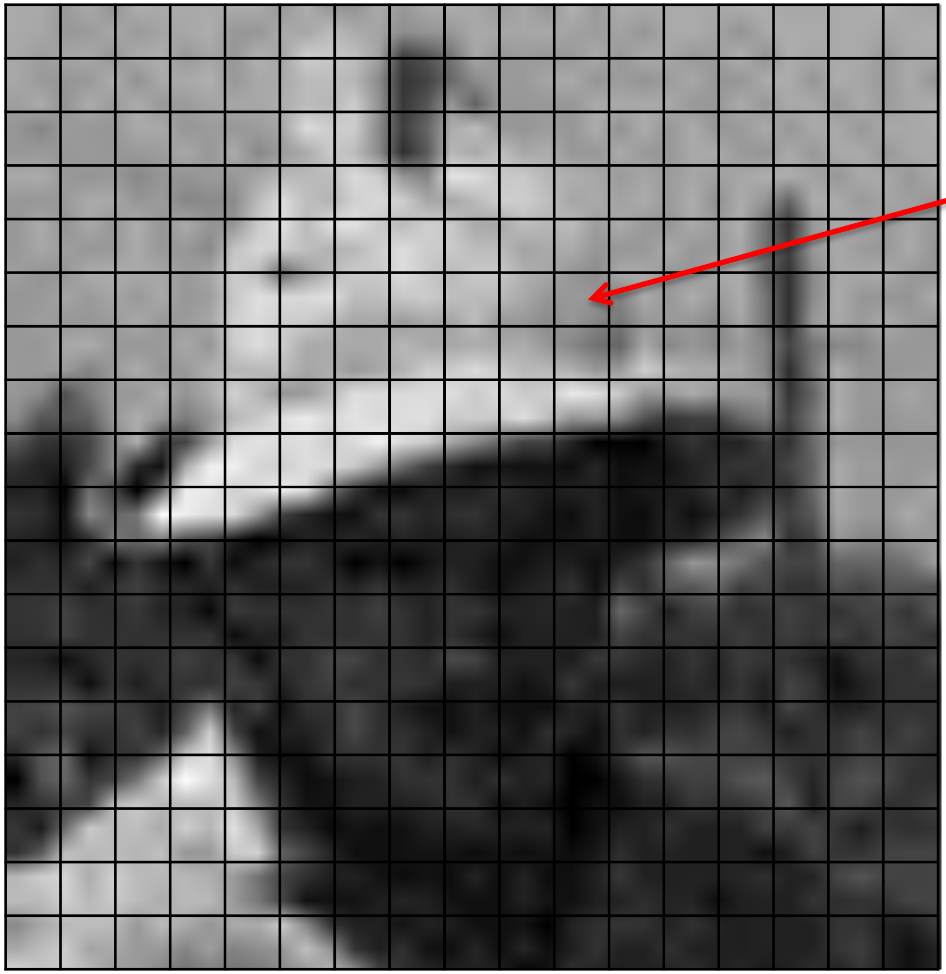
GEO1016
Photogrammetry and 3D Computer Vision

Lecture 2
Camera Models

Liangliang Nan

What is an image?

What is an image?

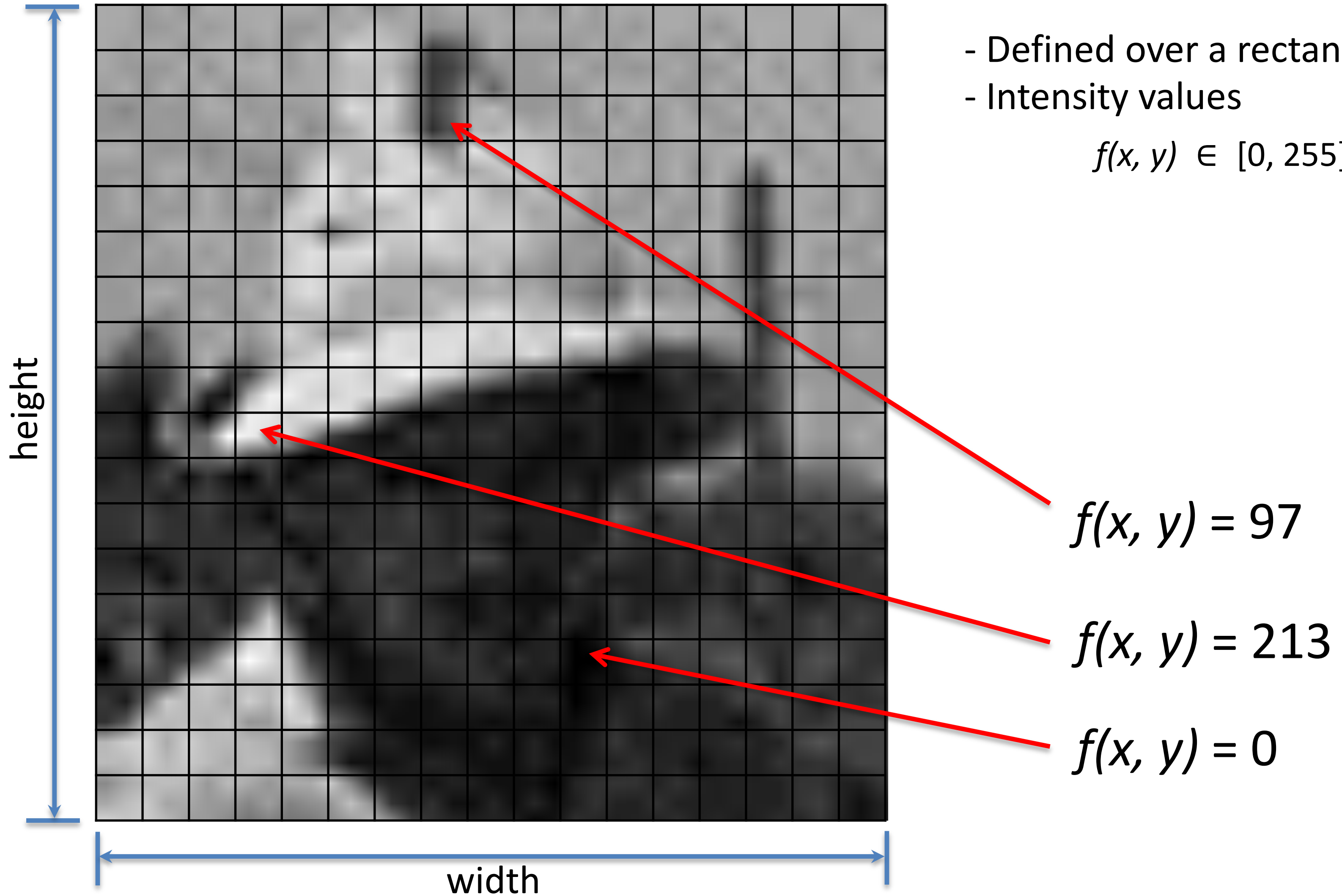


$$P = f(x, y)$$

$$f : R^2 \Rightarrow R$$

Pixel

What is an image?

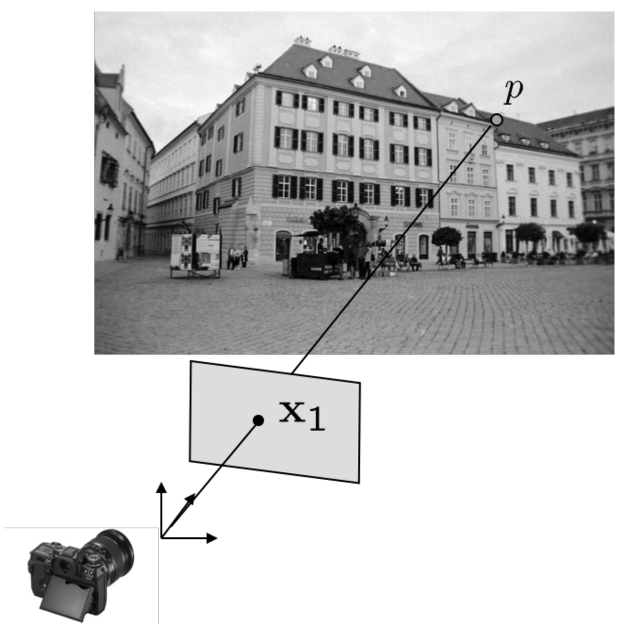


- Defined over a rectangle
- Intensity values

$f(x, y) \in [0, 255]$

What is an image?

- Projection of the scene on the image plane
- Digital (discrete) image
  - A matrix of integer values

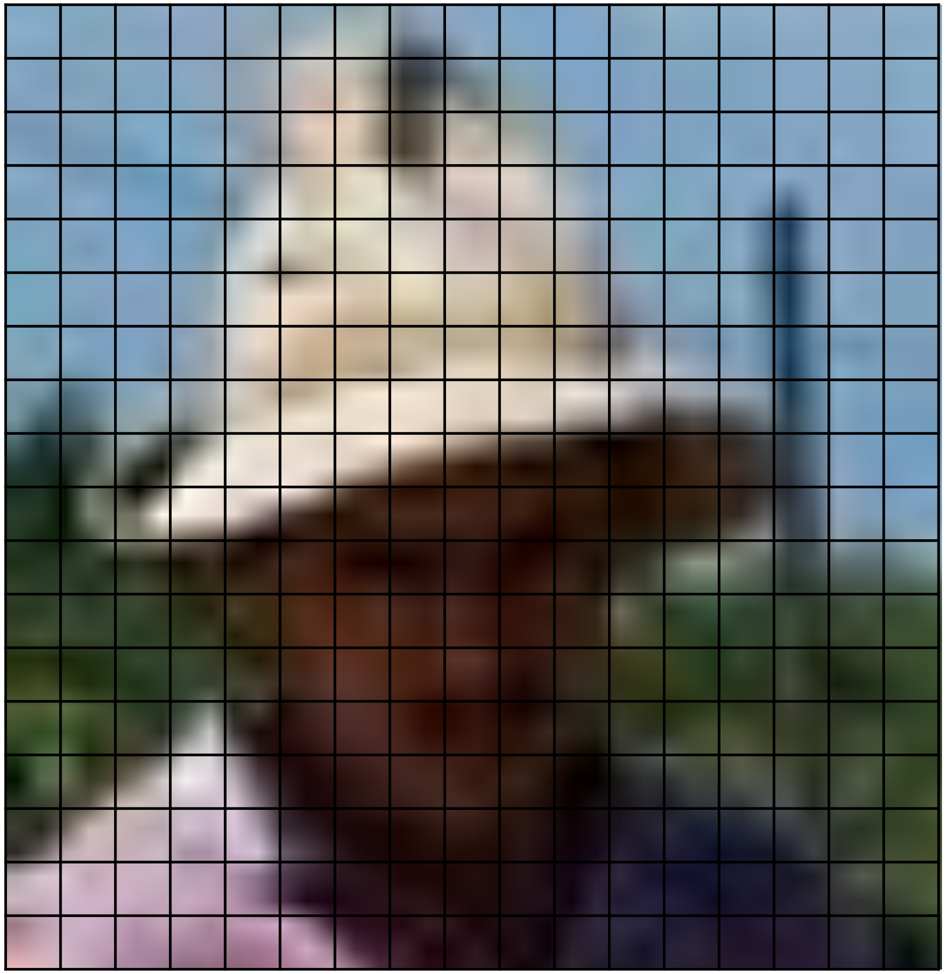


$j \rightarrow$

$i \downarrow$

62	79	23	119	120	105	4	0
10	10	9	62	12	78	34	0
10	58	197	46	46	0	0	48
176	135	5	188	191	68	0	49
2	1	1	29	26	37	0	77
0	89	144	147	187	102	62	208
255	252	0	166	123	62	0	31
166	63	127	17	1	0	99	30

What is an image?



A color image: R, G, B channels

$$f(x,y)=\begin{bmatrix} r(x,y) \\ g(x,y) \\ b(x,y) \end{bmatrix}$$

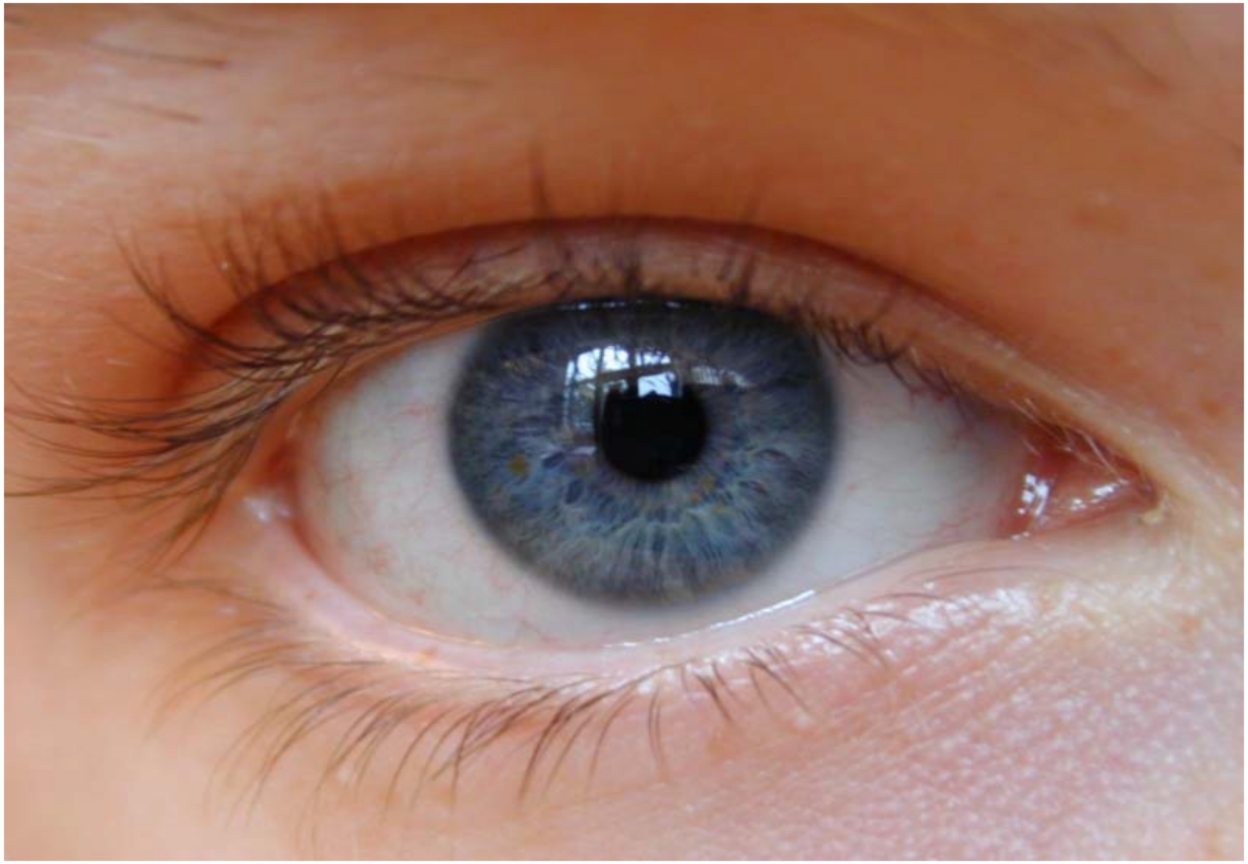
"vector-valued" function

Today's Agenda

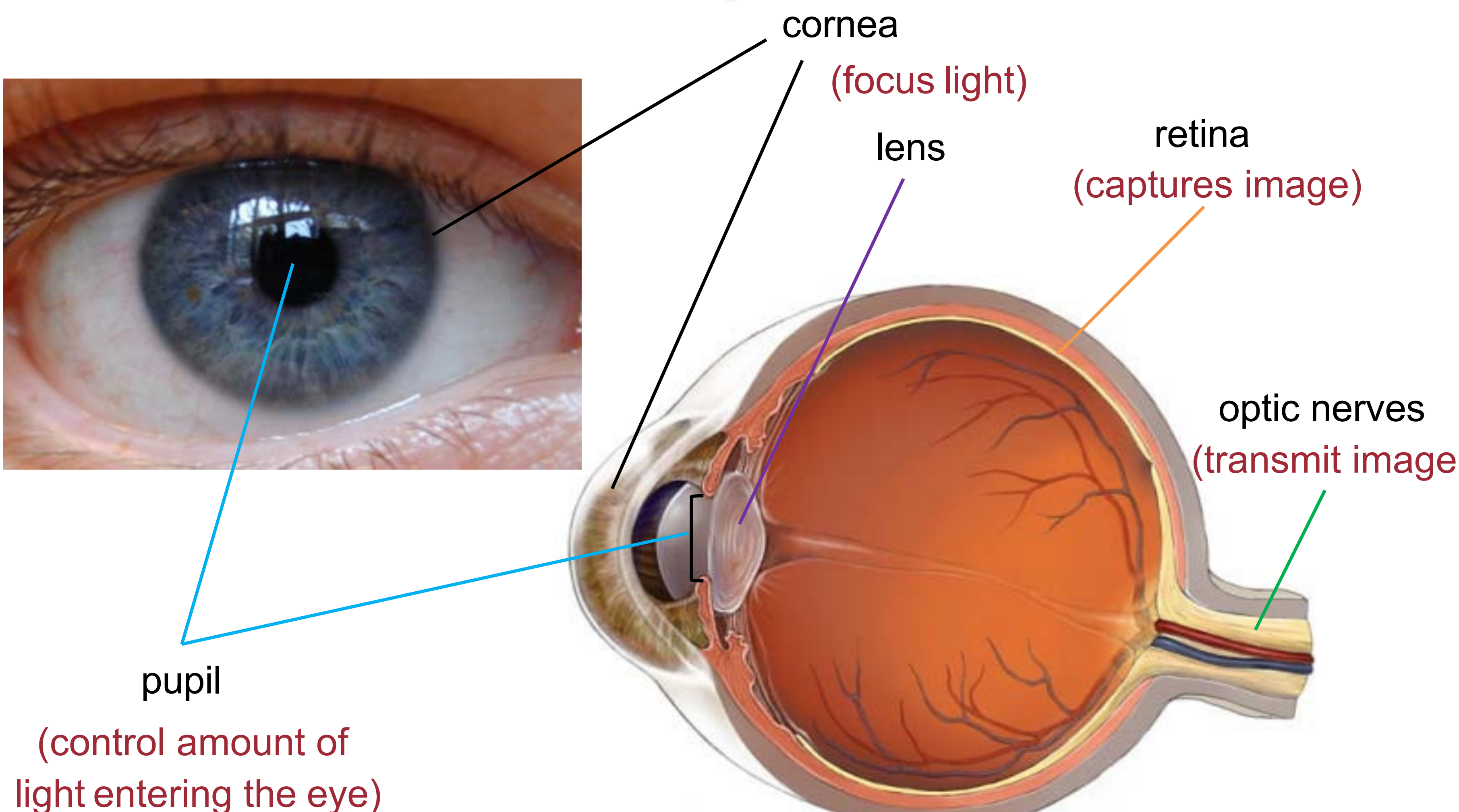
- Images
- Camera Models

Through our eyes…

- We see the world

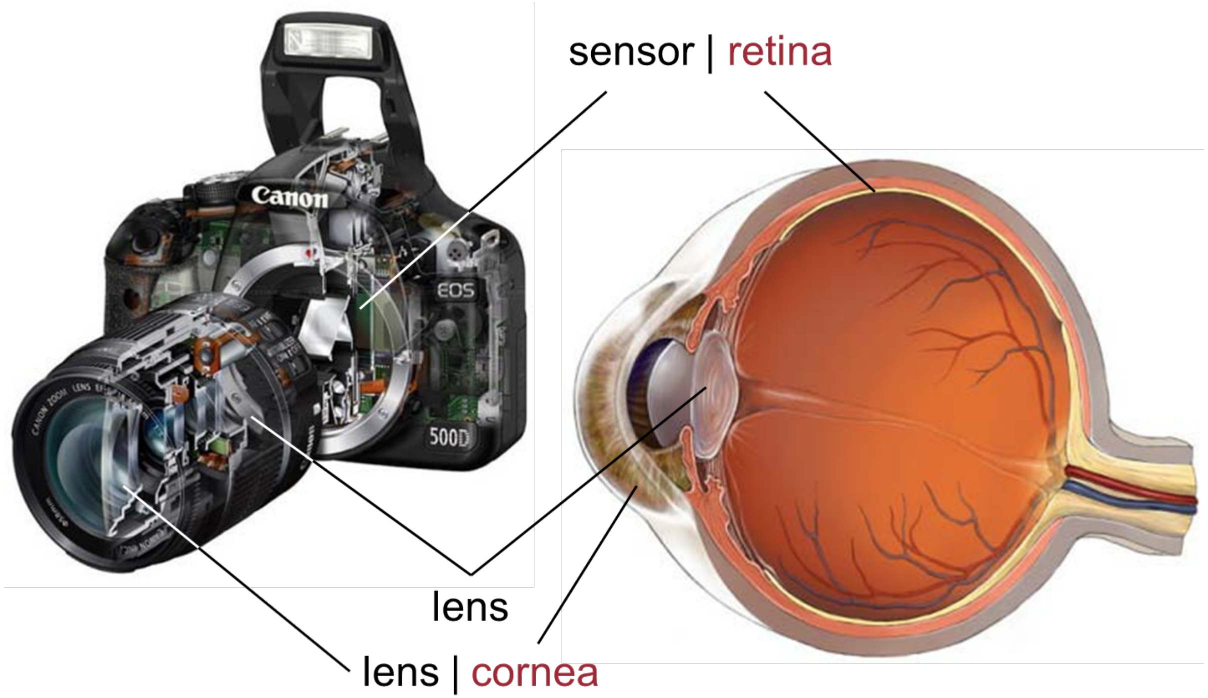


Through our eyes…



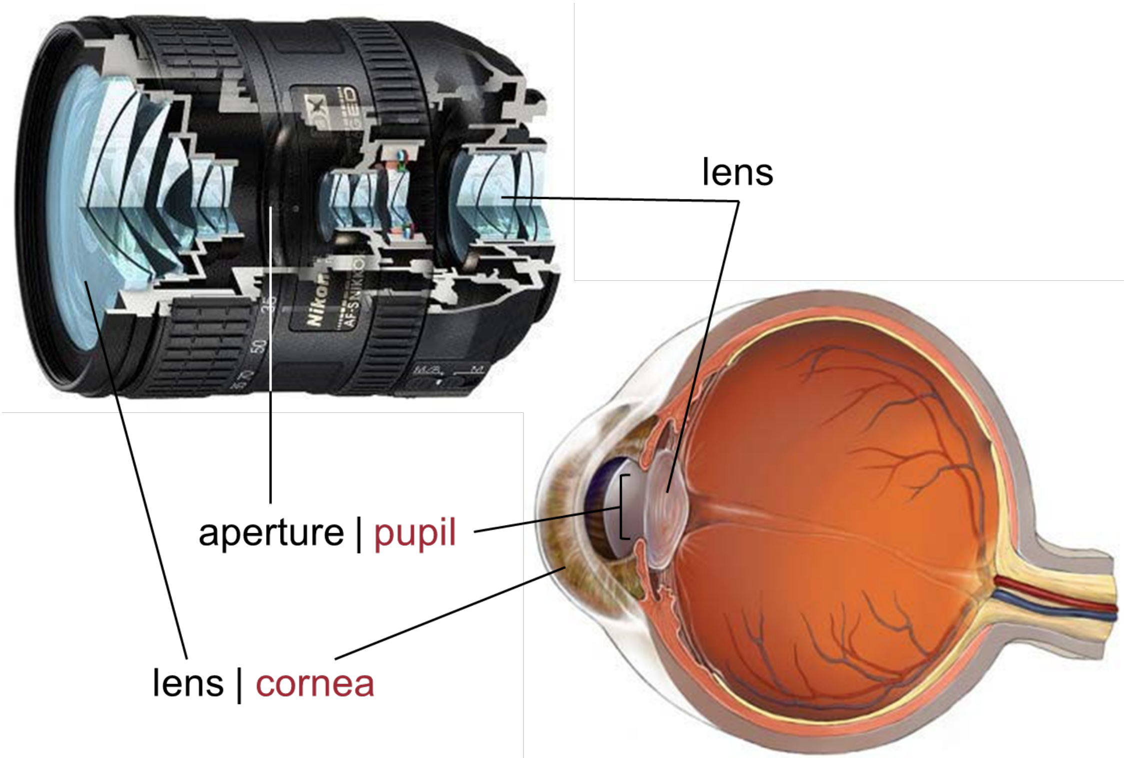
Through their eyes…

- Camera is structurally the same the eye
  - Lens does similarly as our lens and cornea
  - Sensor receives the light signals to form images



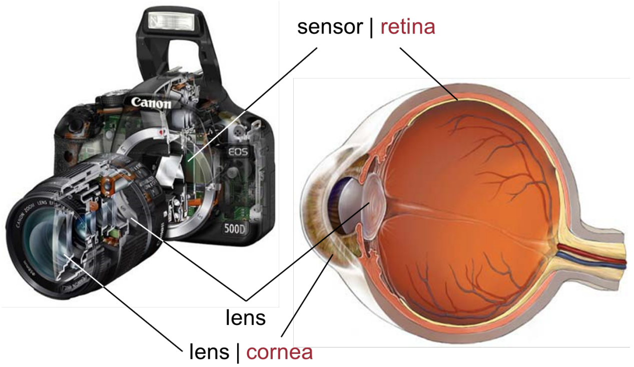
Through their eyes…

- Camera is structurally the same the eye
  - Aperture controls the amount of light



Camera vs. eye

- Similarities
  - – Image focusing
  - – Light adjustment
- Differences (to name a few)
  - – Lens focus
    - Camera: lens moves closer/further from the film
    - Eye: lens changes shape to focus
  - – Sensitivity to light
    - Camera: A film is uniformly sensitive to light
    - Eye: retina is not; has greater sensitivity in dark



Imaging…

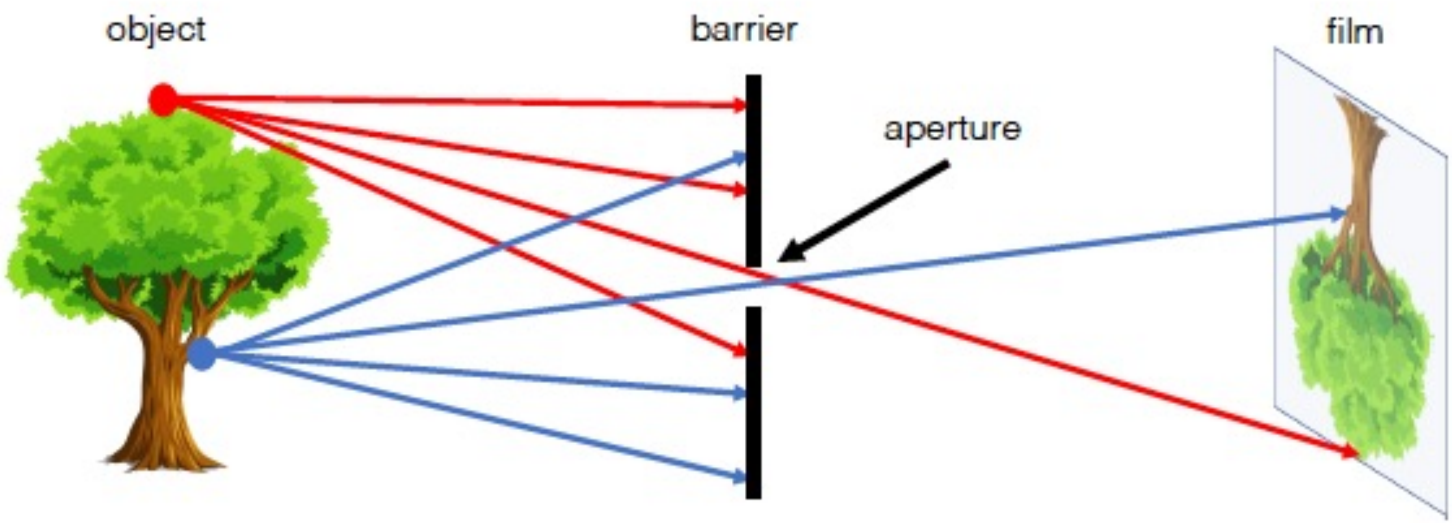
- Images are 2D projections of real-world scenes
- Images capture two kinds of information:
  - Geometric: points, lines, curves, etc.
  - Photometric: intensity, color.
- Complex 3D-2D relationships
  - Camera models approximate relationships.

Camera models

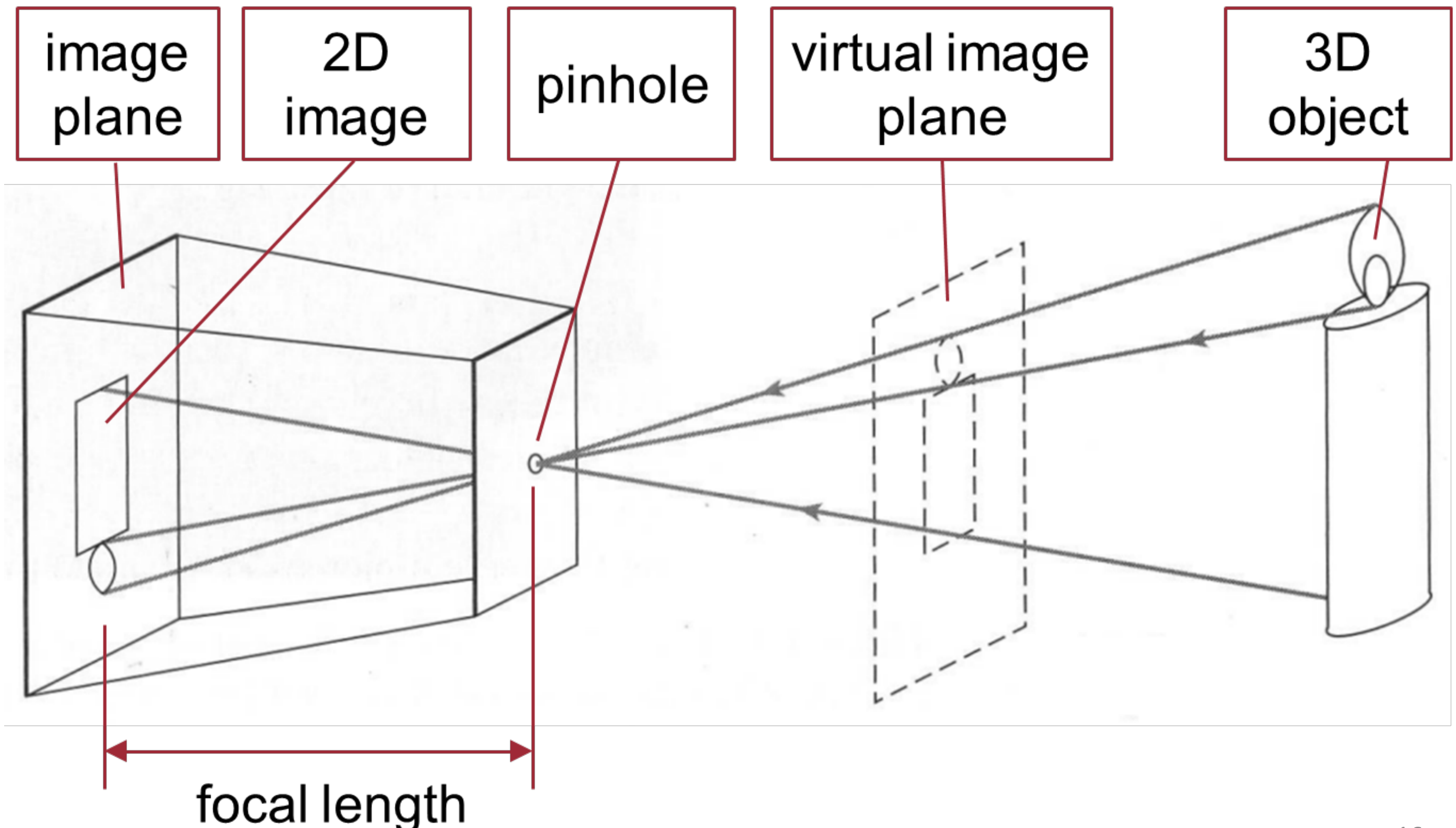
- Pinhole camera model
- Perspective projection model
  - Most commonly used model

Pinhole camera model

- One-to-one mapping

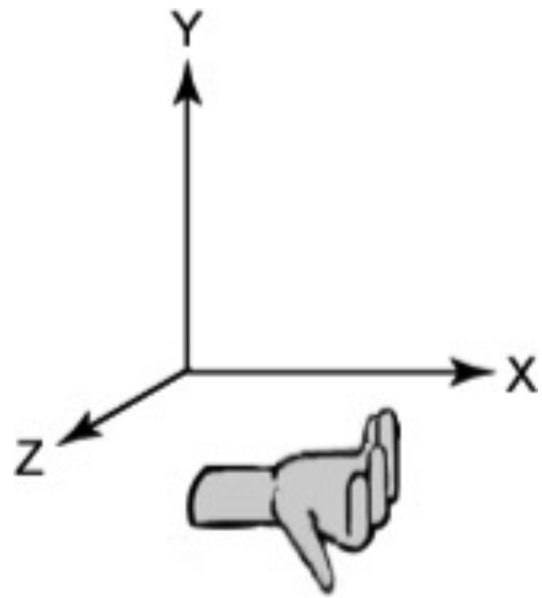


Pinhole camera model



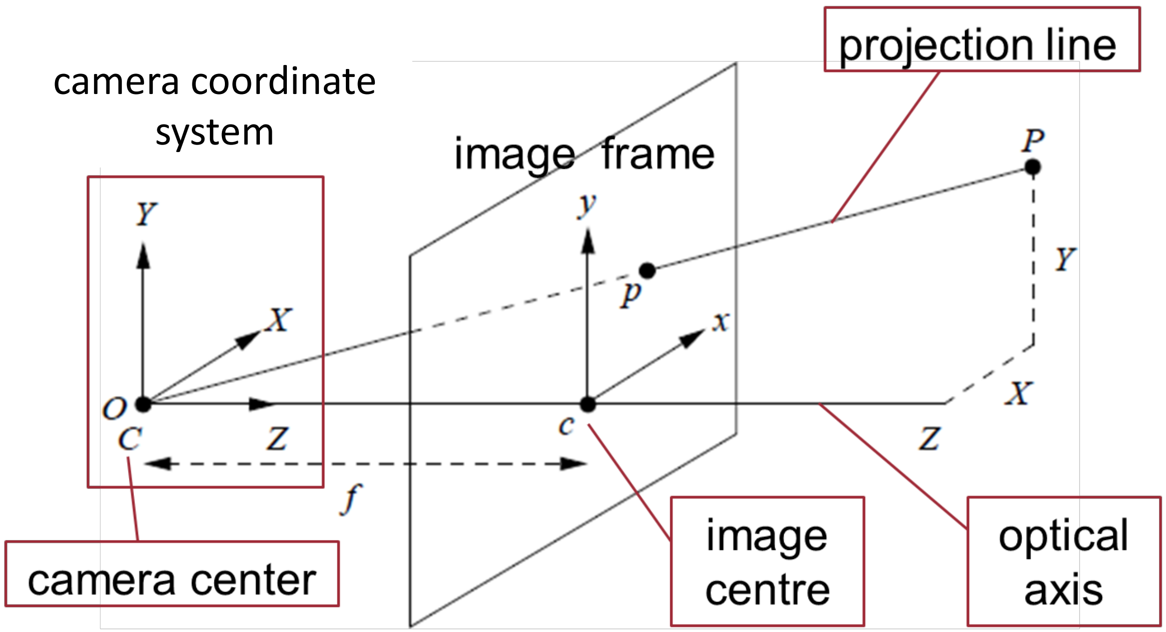
Pinhole camera model

- Right-handed coordinate system



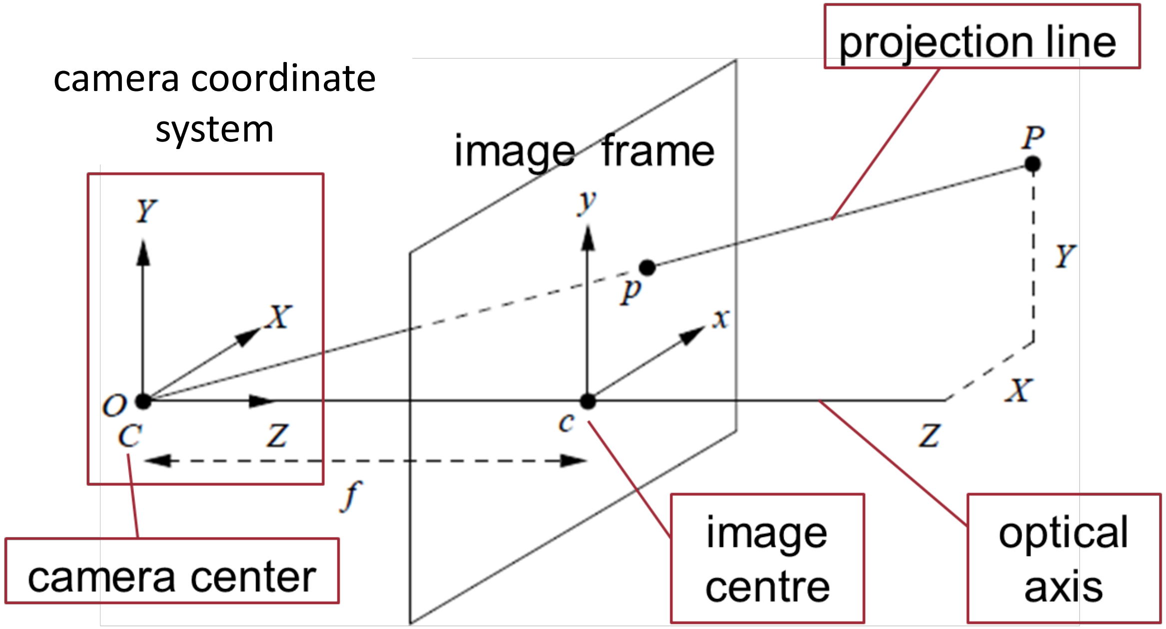
Pinhole camera model

- 3D point $\mathbf{P} = (X, Y, Z)^\mathrm{T}$ projected to 2D image point $\mathbf{p} = (x, y)^\mathrm{T}$.



Pinhole camera model

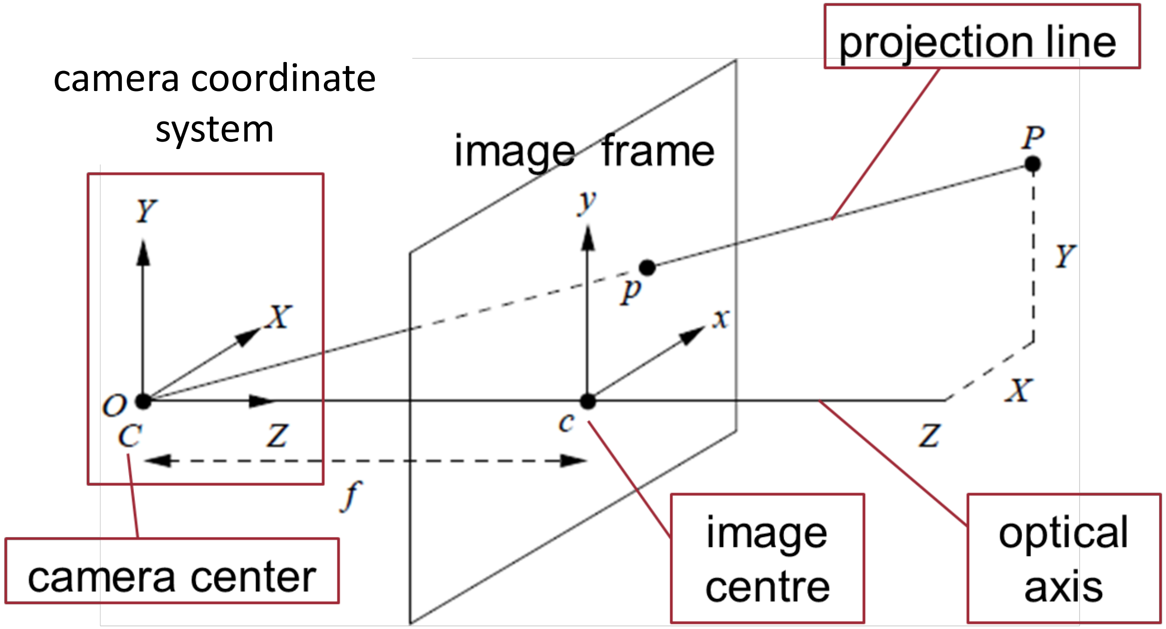
- 3D point $\mathbf{P} = (X, Y, Z)^T$ projected to 2D image point $\mathbf{p} = (x, y)^T$.



$x = ?$ $y = ?$

Pinhole camera model

- 3D point $\mathbf{P} = (X, Y, Z)^{\mathsf{T}}$ projected to 2D image point $\mathbf{p} = (x, y)^{\mathsf{T}}$.



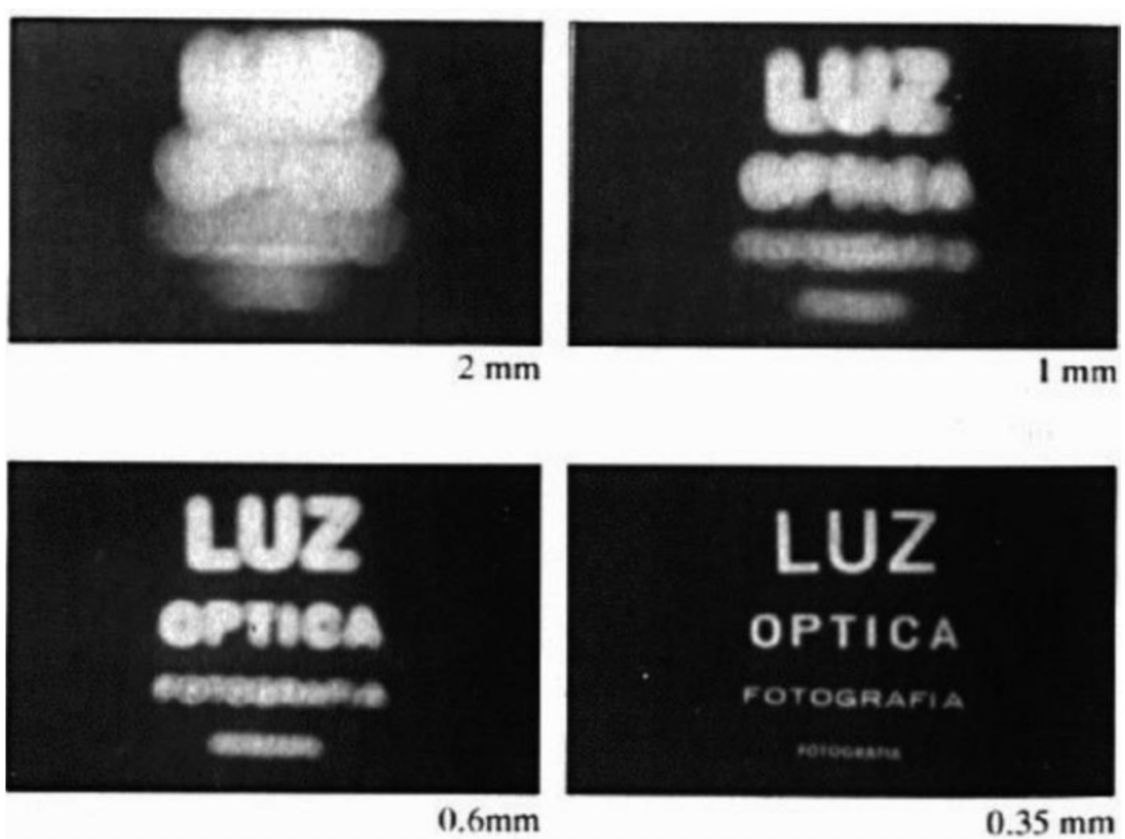
$$\frac{X}{Z} = \frac{x}{f}, \quad \frac{Y}{Z} = \frac{y}{f}$$

$$x = f\frac{X}{Z}, \quad y = f\frac{Y}{Z}$$

Simplest form of perspective projection

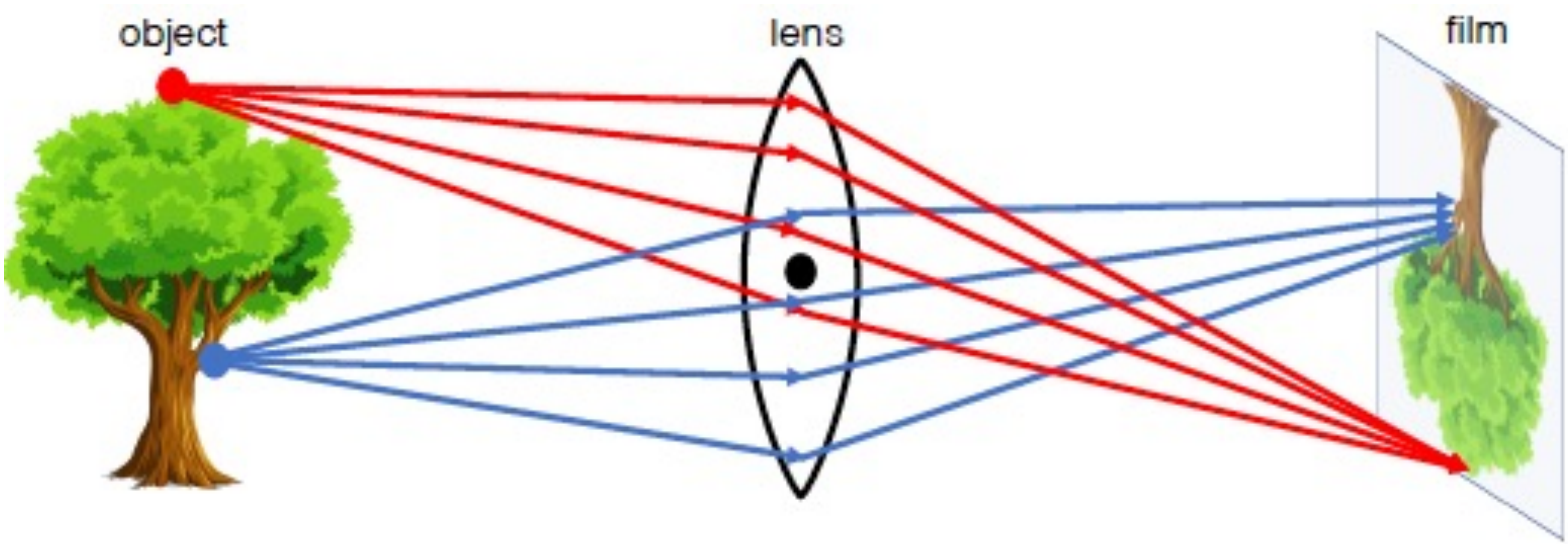
Pinhole camera model

- Assumption: aperture is a single point.



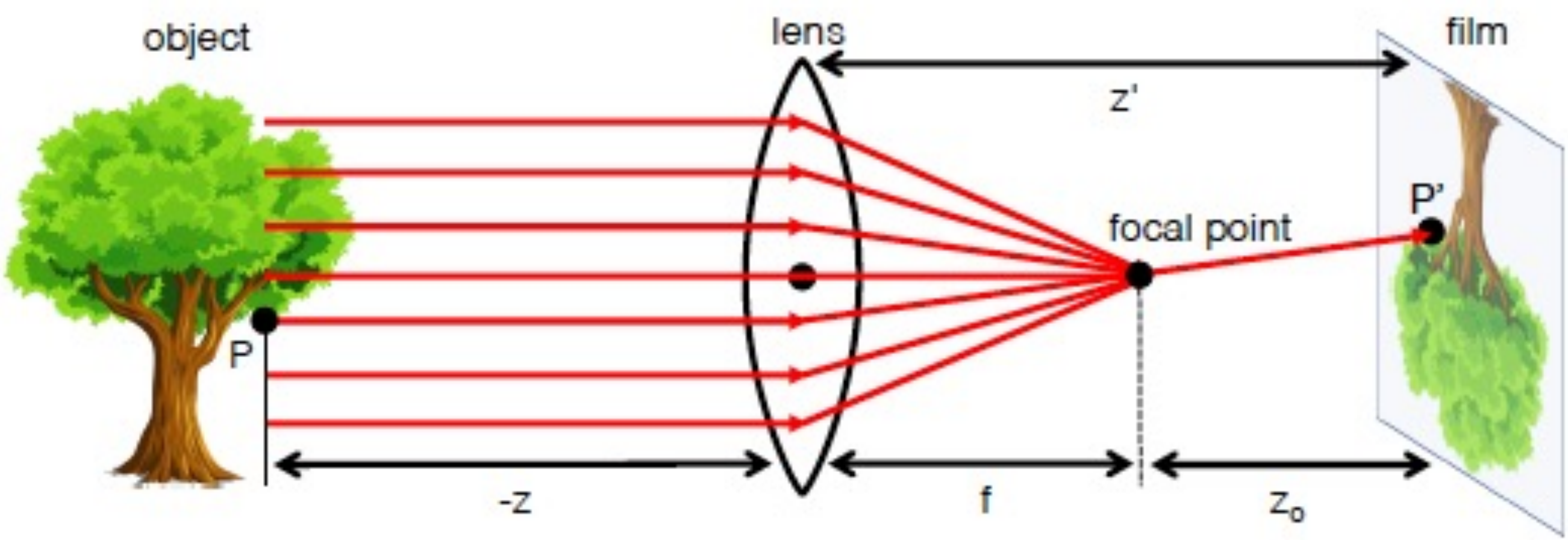
Perspective projection model

- Sharpness vs. brightness?
- Lens
  - Refract light and converge to a singe point



Perspective projection model

- Sharpness vs. brightness?
- Lens
  - Focus parallel light rays to the focal point



Perspective projection model

- Pinhole camera model

$$x = f\frac{X}{Z}, \quad y = f\frac{Y}{Z}$$

- Camera sensor's pixels not exactly square

$$x = kf\frac{X}{Z}, \quad y = lf\frac{Y}{Z}$$

  - x, y : coordinates (pixels)
  - k, l : scale parameters (pixels/mm)
  - f : focal length (m or mm)

Perspective projection model

$$x = kf\frac{X}{Z}, \quad y = lf\frac{Y}{Z}$$

- x, y : coordinates
- k, l : scale parameters
- f : focal length

- We can rewrite (in pixels)

$$f_x = kf \qquad f_y = lf \quad \Rightarrow \quad x = f_x\frac{X}{Z}, \quad y = f_y\frac{Y}{Z}$$

Perspective projection model

$$x = kf\frac{X}{Z}, \quad y = lf\frac{Y}{Z}$$

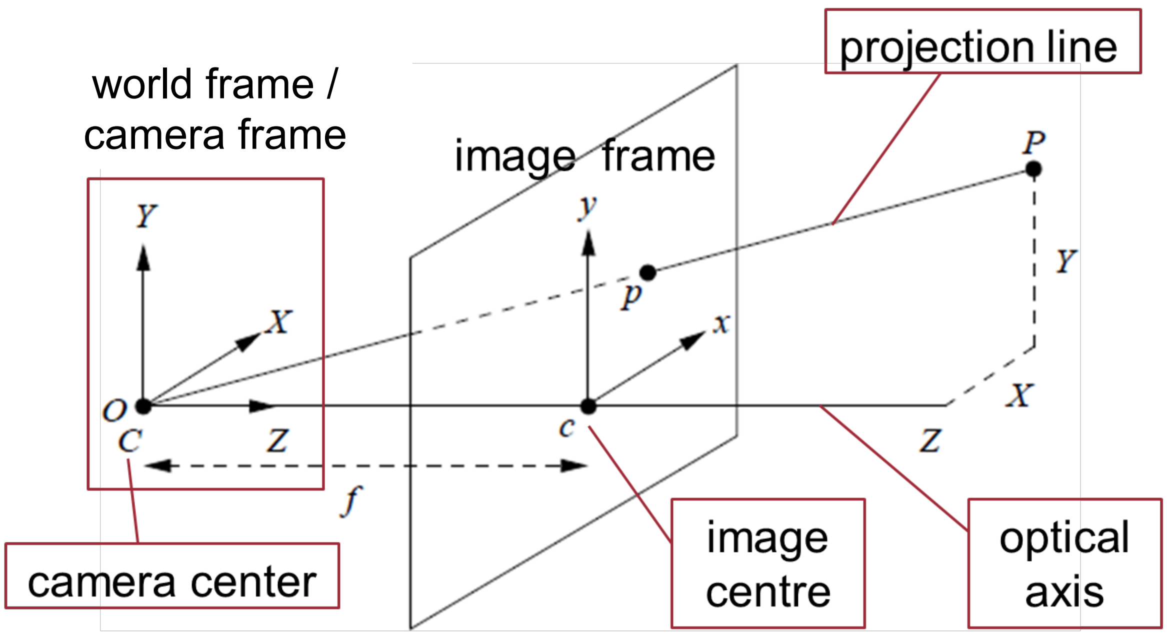
- x, y : coordinates
- k, l : scale parameters
- f : focal length

- We can rewrite (in pixels)

$$f_x = kf \qquad f_y = lf \quad \Rightarrow \quad x = f_x\frac{X}{Z}, \quad y = f_y\frac{Y}{Z}$$

- Image center (or principal point) c may not be at origin
  - Pinhole: the origin of the image coordinate system is located at the image center

Perspective projection model



We must consider: an unknown translation between the origin of the digital image coordinate system and the origin of the image plane

Perspective projection model

$$x = kf\frac{X}{Z}, \quad y = lf\frac{Y}{Z}$$

- x, y : coordinates
- k, l : scale parameters
- f : focal length

- We can rewrite (in pixels)

$$f_x = kf \qquad f_y = lf \qquad \Rightarrow \qquad x = f_x\frac{X}{Z}, \quad y = f_y\frac{Y}{Z}$$

- Image center (or principal point) c may not be at origin
  – Denote location of c in image plane as c_x, c_y.

Perspective projection model

$$x = kf\frac{X}{Z}, \quad y = lf\frac{Y}{Z}$$

- x, y : coordinates
- k, l : scale parameters
- f : focal length

- We can rewrite (in pixels)

$$f_x = kf \qquad f_y = lf \quad \Rightarrow \quad x = f_x\frac{X}{Z}, \quad y = f_y\frac{Y}{Z}$$

- Image center (or principal point) c may not be at origin
  - Denote location of c in image plane as c_x, c_y.

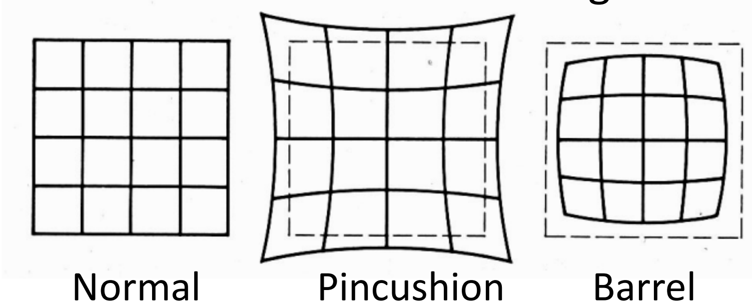
$$x = f_x\frac{X}{Z} + c_x, \quad y = f_y\frac{Y}{Z} + c_y$$

Perspective projection model

$$x = f_x \frac{X}{Z} + c_x, \quad y = f_y \frac{Y}{Z} + c_y$$

- Image frame may not be exactly rectangular

Common aberration: referred to as radial distortion
Reason: different portions of the lens have differing focal lengths.



Perspective projection model

$$x = f_x \frac{X}{Z} + c_x, \quad y = f_y \frac{Y}{Z} + c_y$$

- Image frame may not be exactly rectangular
  - Let θ denote skew angle between x- and y-axis

$$x = f_x \frac{X}{Z} - f_x \cot\theta \frac{Y}{Z} + c_x, \quad y = \frac{f_y}{\sin\theta} \frac{Y}{Z} + c_y$$

What if $\theta = \pi/2$?

Intrinsic Parameters

$$x = f_x \frac{X}{Z} - f_x \cot\theta \frac{Y}{Z} + c_x, \quad y = \frac{f_y}{\sin\theta} \frac{Y}{Z} + c_y$$

- Combine all the parameters

$$\tilde{\mathrm{x}} = \frac{1}{Z}\mathbf{K}\mathbf{X}, \quad \mathbf{K} =$$

Intrinsic Parameters

$$x = f_x \frac{X}{Z} - f_x \cot\theta \frac{Y}{Z} + c_x, \quad y = \frac{f_y}{\sin\theta}\frac{Y}{Z} + c_y$$

- Combine all the parameters

$$\tilde{\mathbf{x}} = \frac{1}{Z}\mathbf{K}\mathbf{X}, \quad \mathbf{K} = \begin{bmatrix} f_x & -f_x \cot\theta & c_x \\ 0 & \dfrac{f_y}{\sin\theta} & c_y \\ 0 & 0 & 1 \end{bmatrix}$$

Intrinsic parameter matrix

$$\tilde{\mathbf{x}} = [x, y, 1]^\top$$

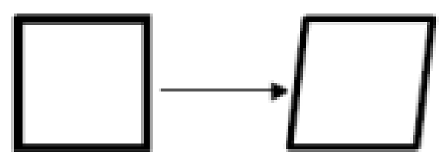
homogeneous coordinates

Intrinsic Parameters

- For simplicity, people use a simpler form of $\mathbf{K}$

$$\mathbf{K} = \begin{bmatrix} f_x & -f_x \cot\theta & c_x \\ 0 & \dfrac{f_y}{\sin\theta} & c_y \\ 0 & 0 & 1 \end{bmatrix} \Rightarrow \mathbf{K} = \begin{bmatrix} f_x & s & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}$$

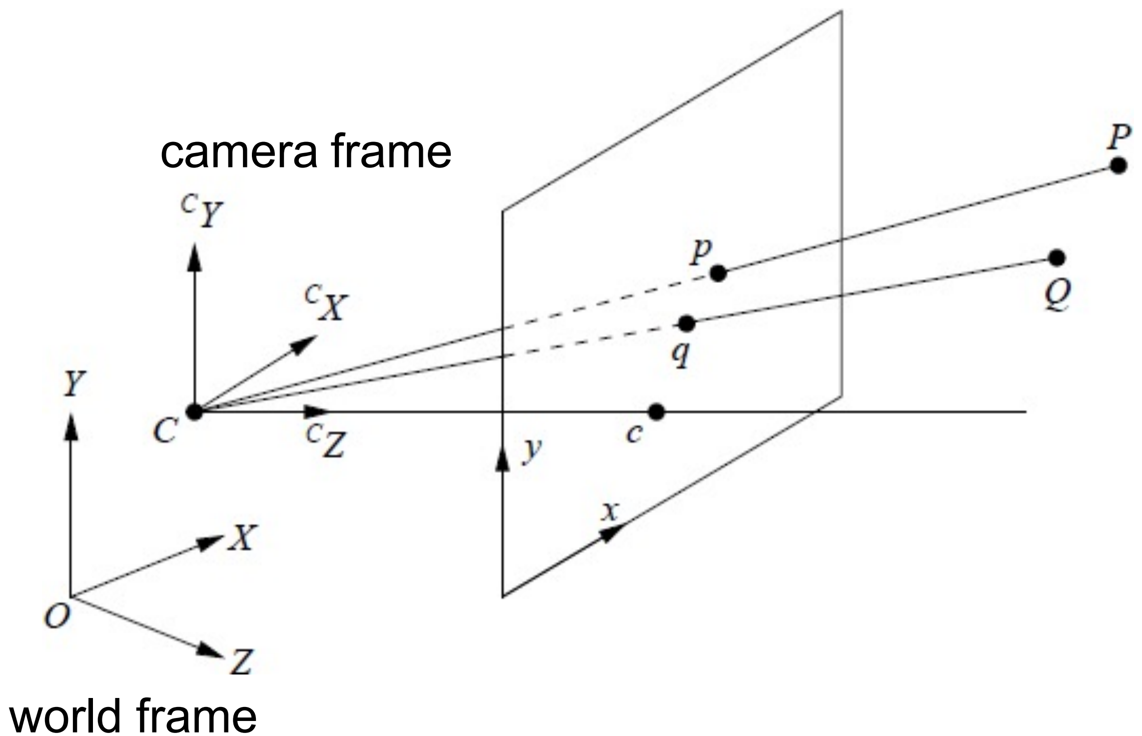
s: skew parameter



- Internal characteristics
  - – focal length, skew, distortion, and image center.

Extrinsic Parameters

- Camera frame is not aligned with world frame
- Camera can move and rotate



Extrinsic Parameters

- Camera frame is not aligned with world frame
- Camera can move and rotate
- Rigid transformation between them

Camera frame

$$\boxed{{}^{C}\mathbf{X}} = \boxed{{}^{C}_{W}\mathbf{R}}\boxed{{}^{W}\mathbf{X}} + \boxed{{}^{C}_{W}\mathbf{T}}$$

World frame

1 3 2 4

1. Coordinates of 3D scene point in camera frame.
2. Coordinates of 3D scene point in world frame.
3. Rotation matrix of world frame in camera frame.
4. Position of world frame's origin in camera frame.

From 3D points to pixels

- Combine intrinsic and extrinsic parameters

$$\boxed{\tilde{\mathbf{x}} = \frac{1}{Z}\mathbf{K}\mathbf{X}} \qquad {}^{C}\mathbf{X} = {}^{C}_{W}\mathbf{R}\,{}^{W}\mathbf{X} + {}^{C}_{W}\mathbf{T} \qquad \rho = \frac{1}{Z}$$

$$\rho\,\tilde{\mathbf{x}} = \mathbf{K}\,{}^{C}\mathbf{X} = \mathbf{K}\left({}^{C}_{W}\mathbf{R}\,{}^{W}\mathbf{X} + {}^{C}_{W}\mathbf{T}\right)$$

- Use a simpler notation

$$\tilde{\mathbf{x}} = \mathbf{K}(\mathbf{R}\mathbf{X} + \mathbf{T}) = \mathrm{MX}$$

Note: in many text books, the 1/Z (rho) is absorbed in K and M.

M: projection matrix

Summary Camera Models

- Simplest camera model: pinhole model.
- Most commonly used model: perspective model.
- Intrinsic parameters:
  - Focal length, principal point (image center), skew factor
- Extrinsic parameters:
  - Camera rotation and translation.

Further reading :

R. Szeliski. *Computer Vision: Algorithms and Applications*. Springer, 2010.

- Camera models: Section 2.1.5
- Lens distortion: Section 2.1.6

Next Lecture

- **Camera Calibration**