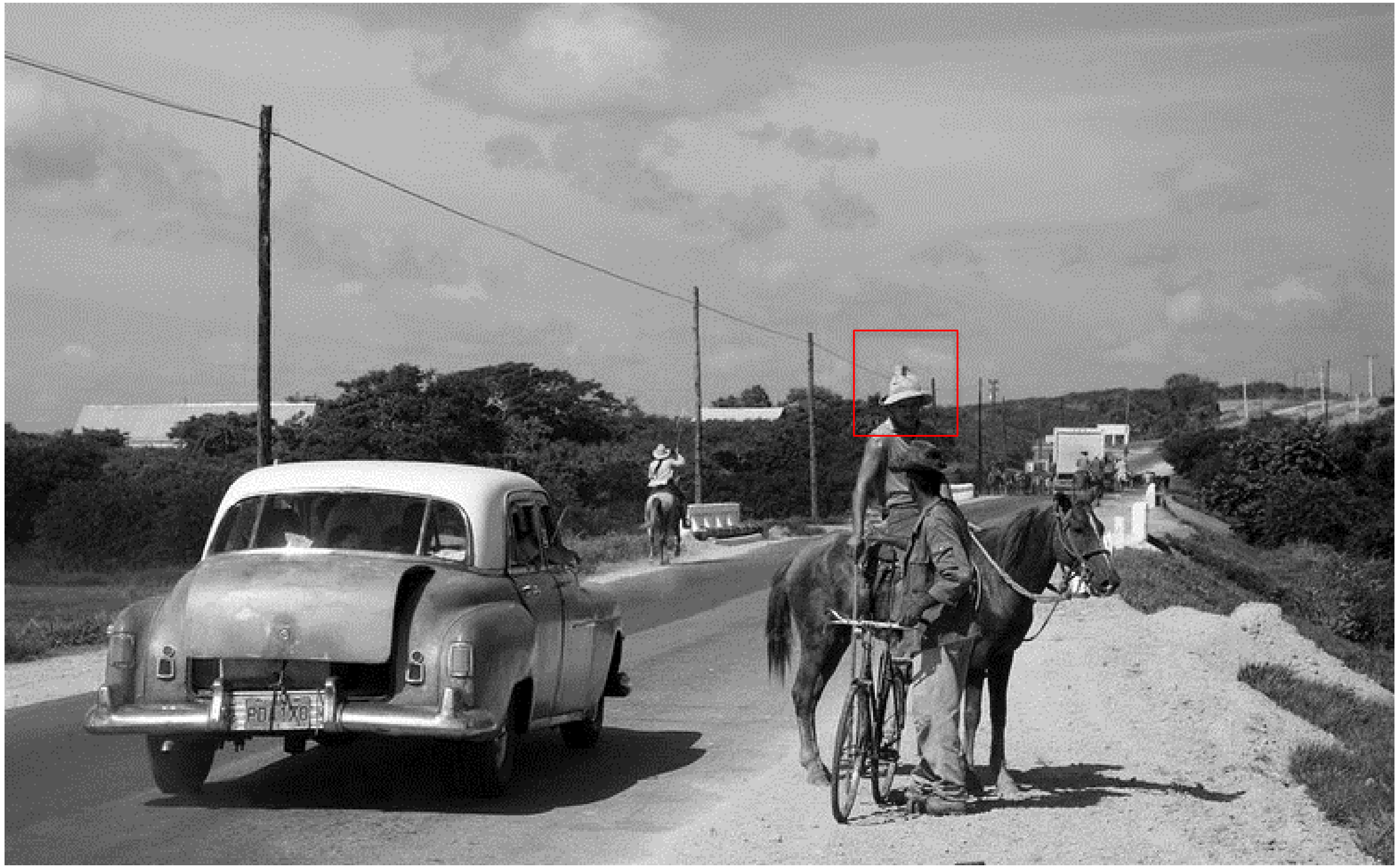
Lecture

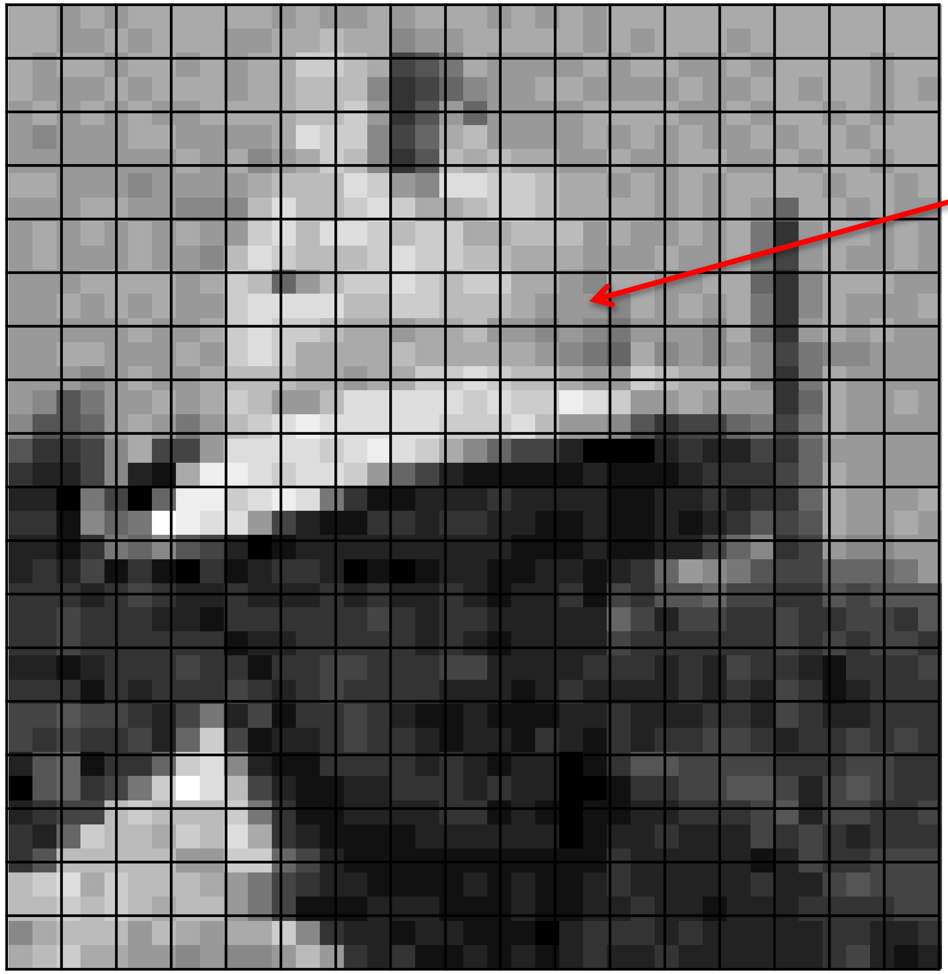
Camera Models

Liangliang Nan

What is an image?



What is an image?

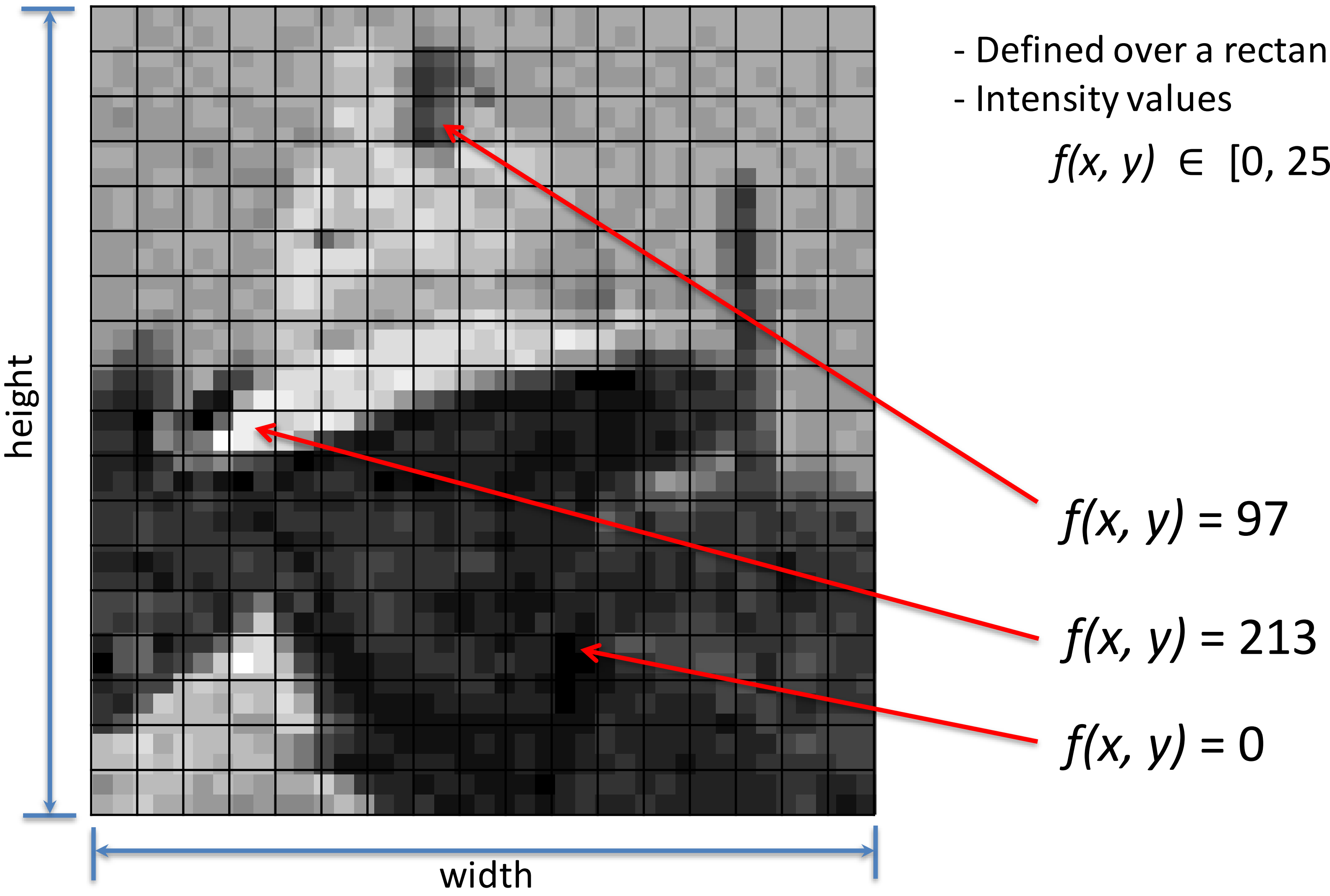


$$P = f(x,y)$$

$$f : R^2 \rightarrow R$$

Pixel

What is an image?

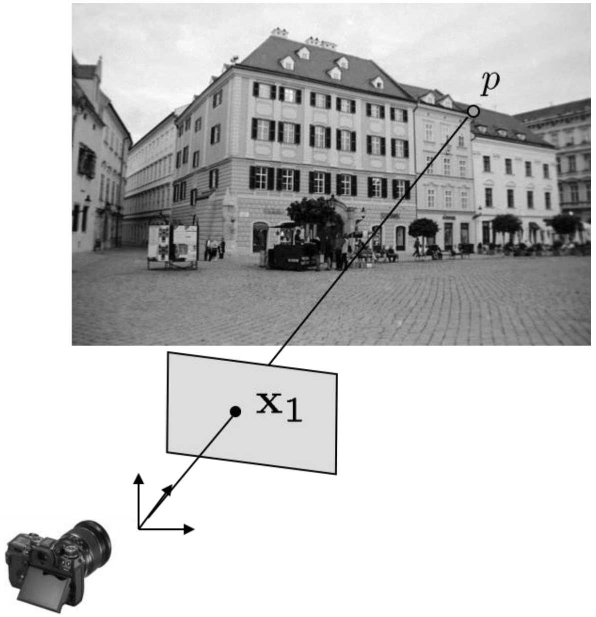


- Defined over a rectangle
- Intensity values

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

What is an image?

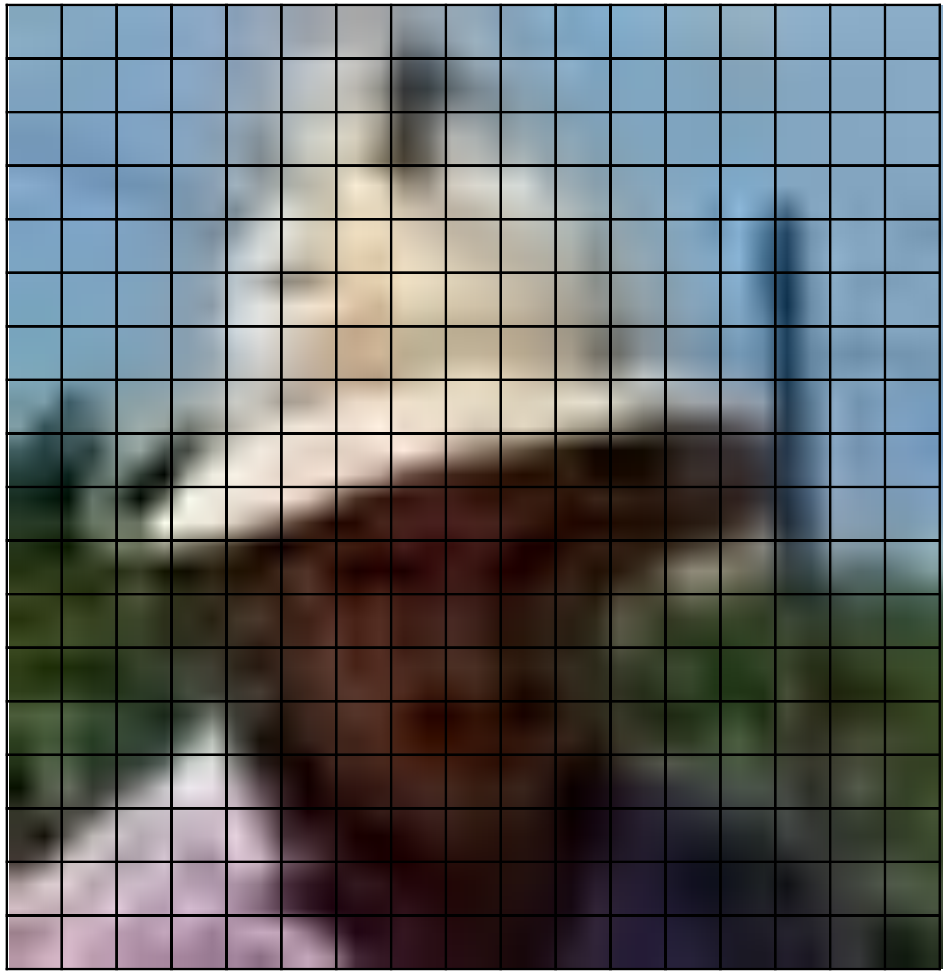
- Projection of the scene on the image plane
- Digital (discrete) image
  - A matrix of integer values



j →, i ↓

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

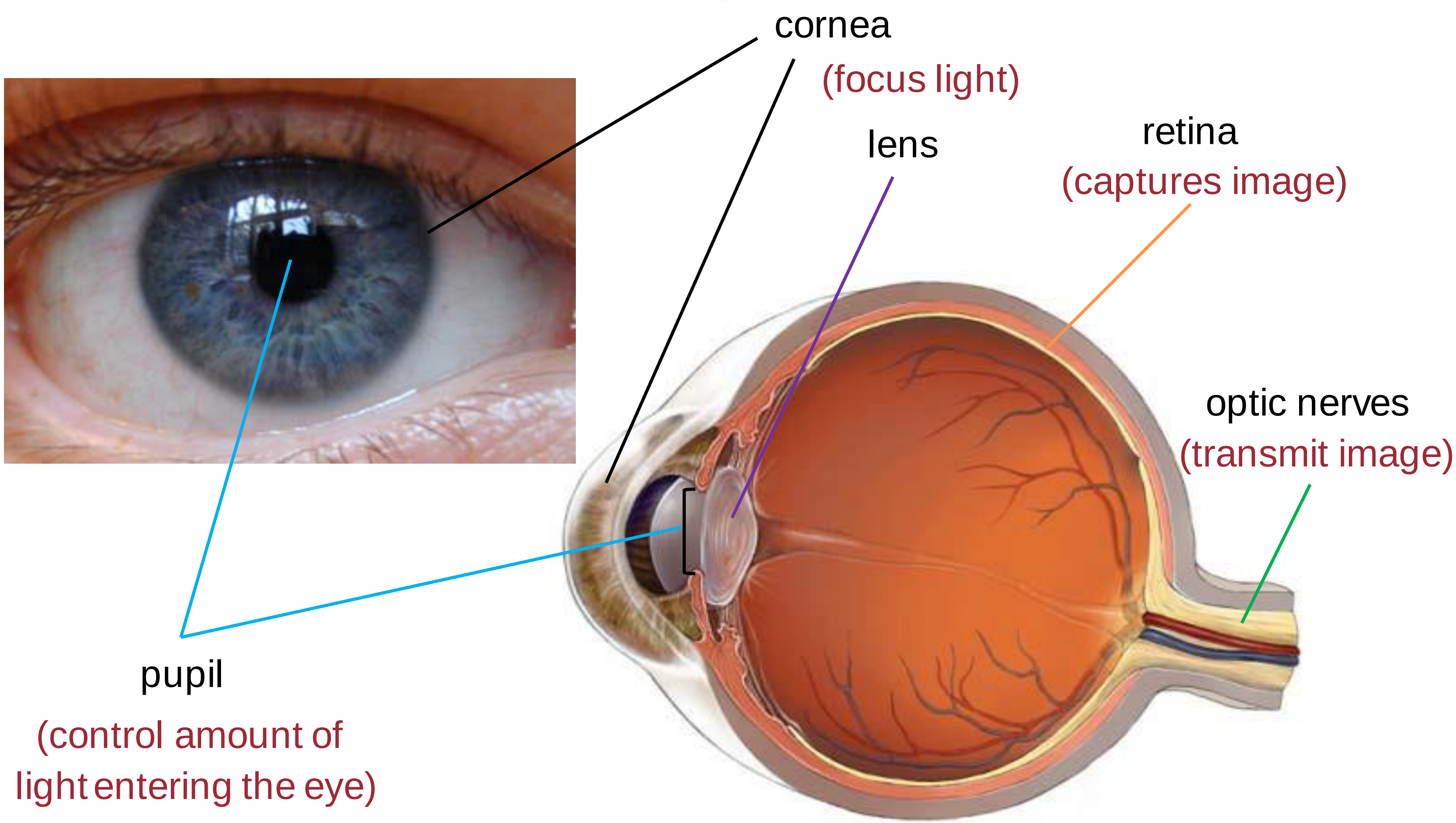
$$f(x,y) = \begin{bmatrix} r(x,y) \\ g(x,y) \\ b(x,y) \end{bmatrix}$$

“vector-valued” function

Today's agenda

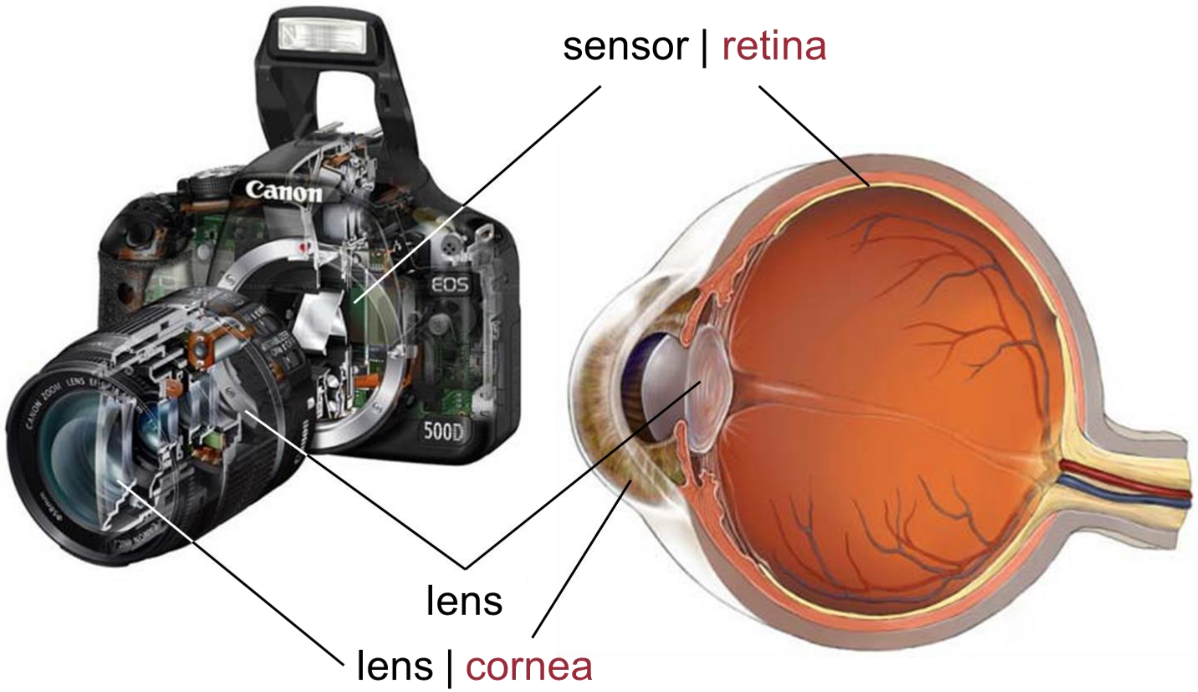
- Images
- Camera Models

We see the world through our eyes…



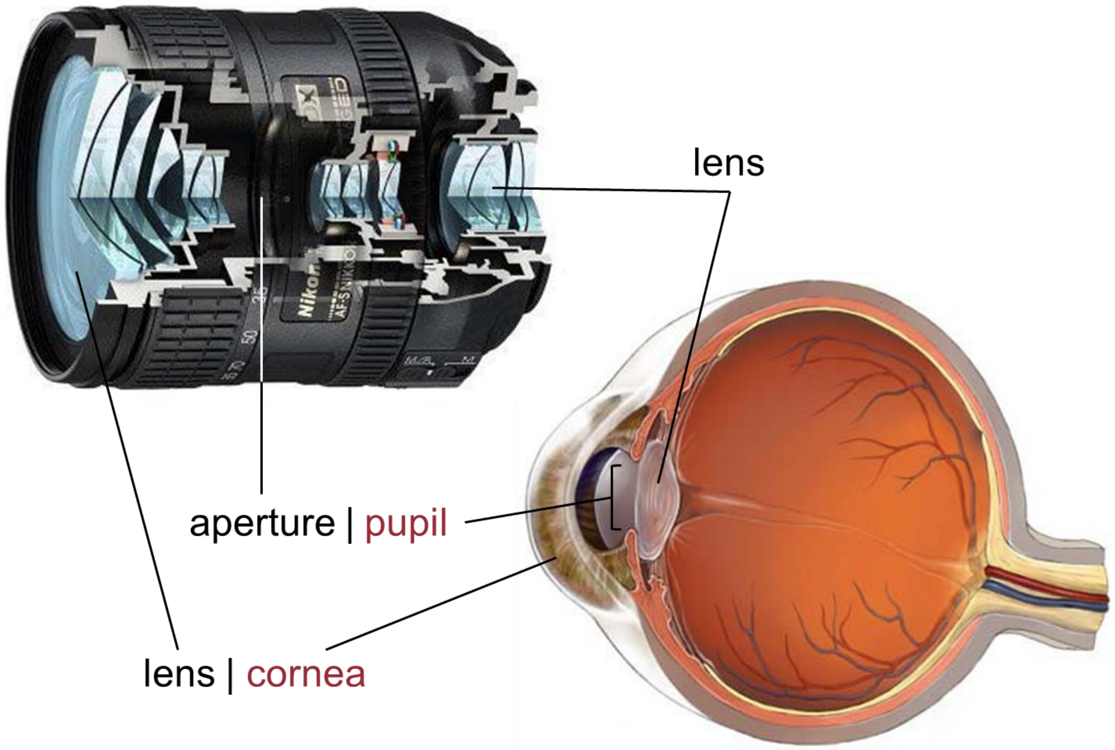
Machines see the world through cameras…

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



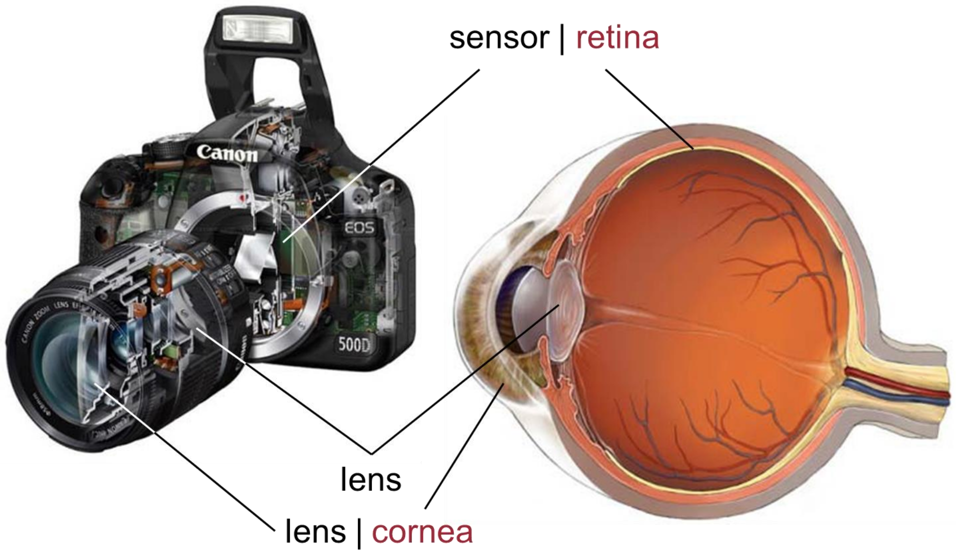
Machines see the world through cameras…

- Camera is structurally the same as our eyes
  - Lens does similarly to our lens and cornea
  - Sensor receives the light signals to form images
  - Aperture controls the amount of light



Camera vs. eye

- Similarities
  - Image focusing (cornea & lens)
  - Light adjustment (pupil & aperture)
- Differences (just 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 designed to be uniformly sensitive to light
    - Eye: retina is not; has greater sensitivity in dark



Imaging …

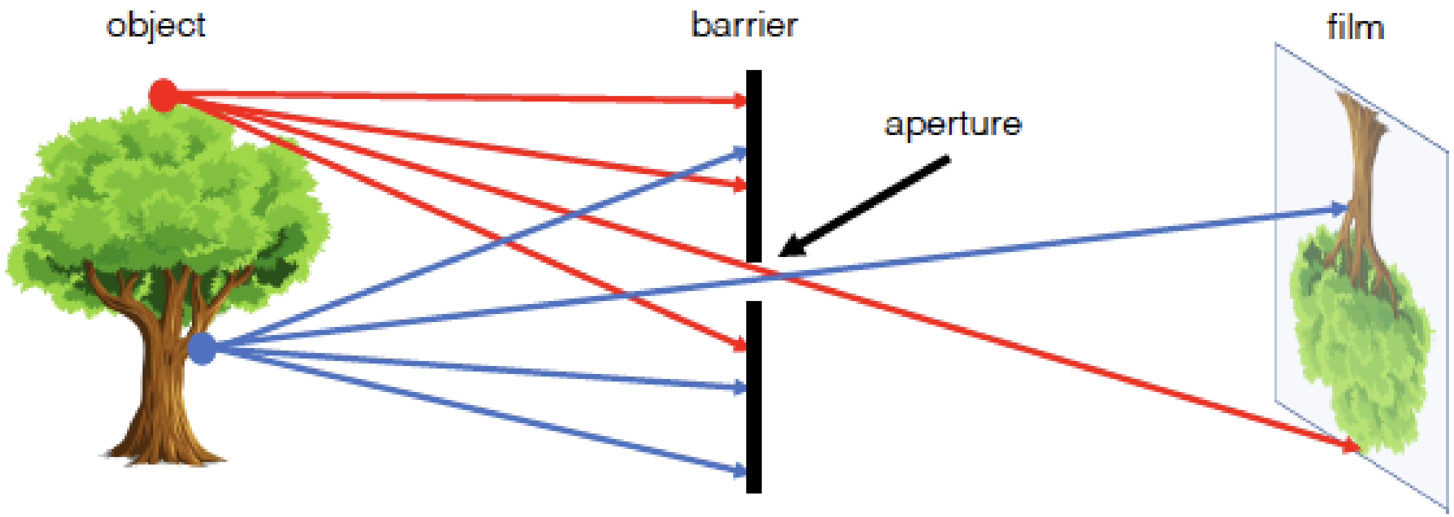
- Images are 2D projections of real-world scenes
- Images capture two kinds of information:
  - **Geometric**: points, lines, curves, etc.
  - **Photometric**: intensity, color
- Complex 3D-2D relationships
  - Camera models approximate relationships

Camera models

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

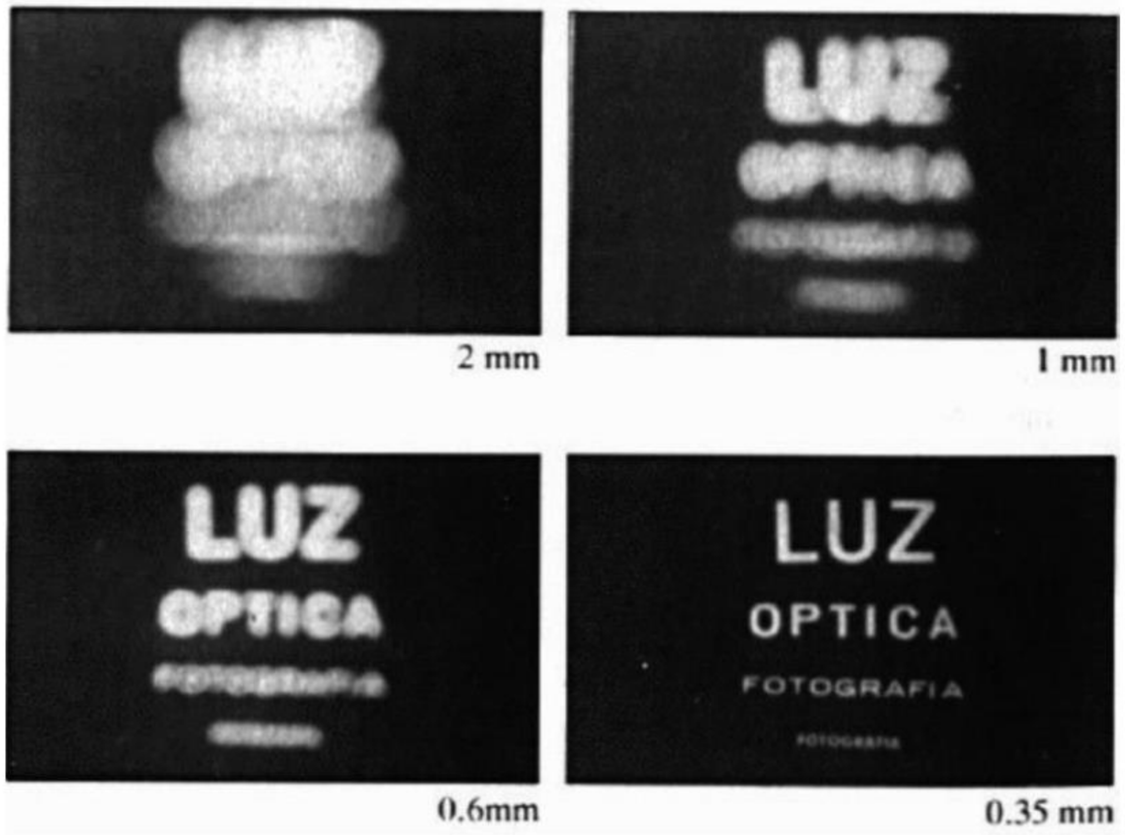
Pinhole camera model

- One-to-one mapping

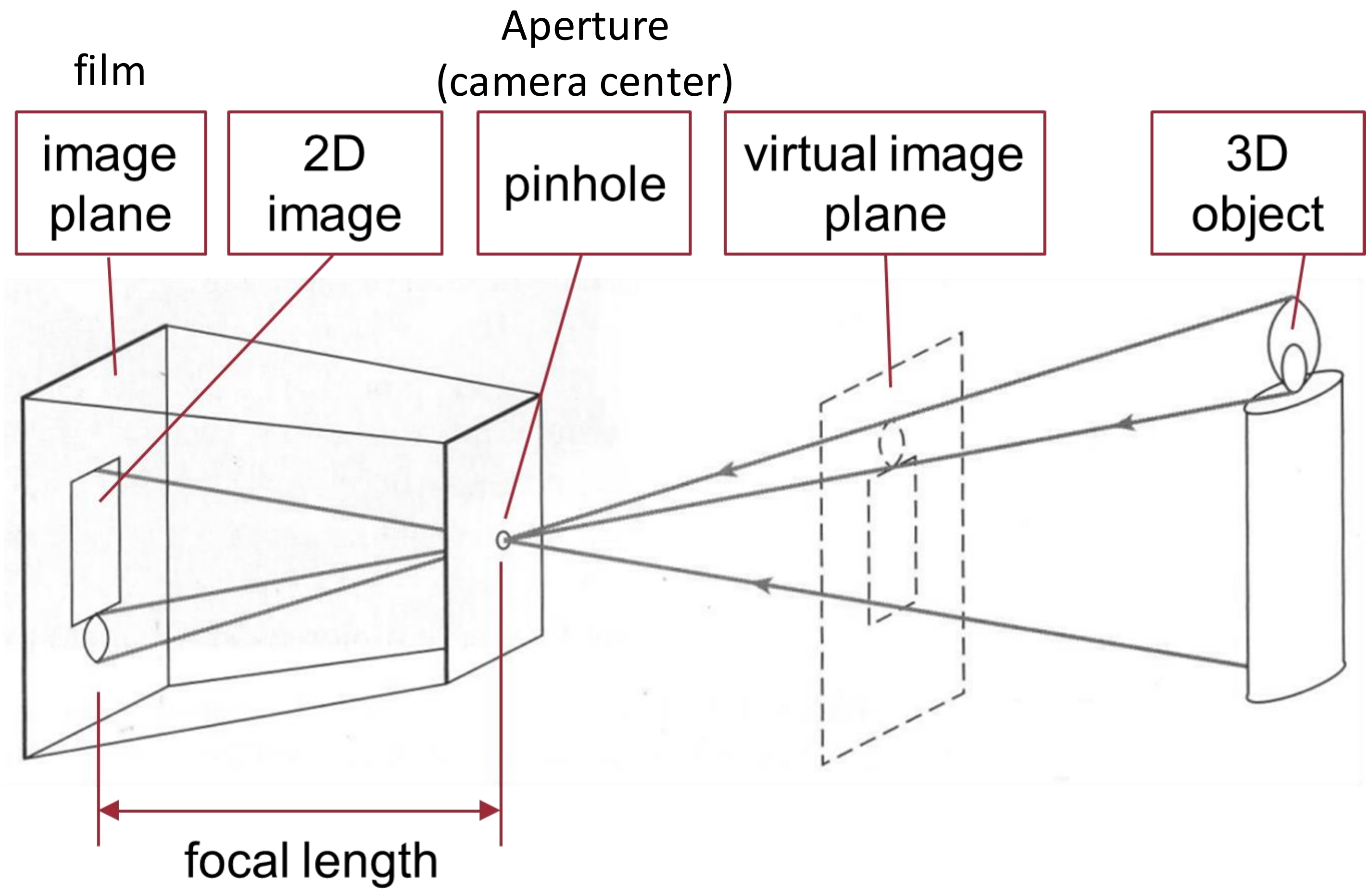


Pinhole camera model

- Assumption: aperture is a single point.



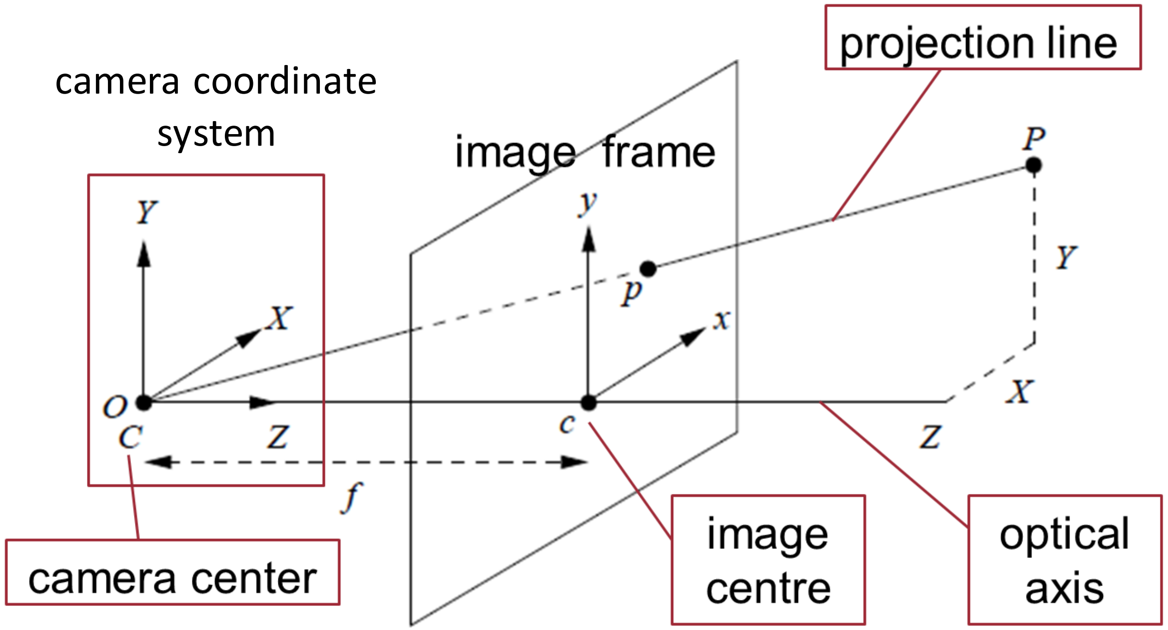
Pinhole camera model

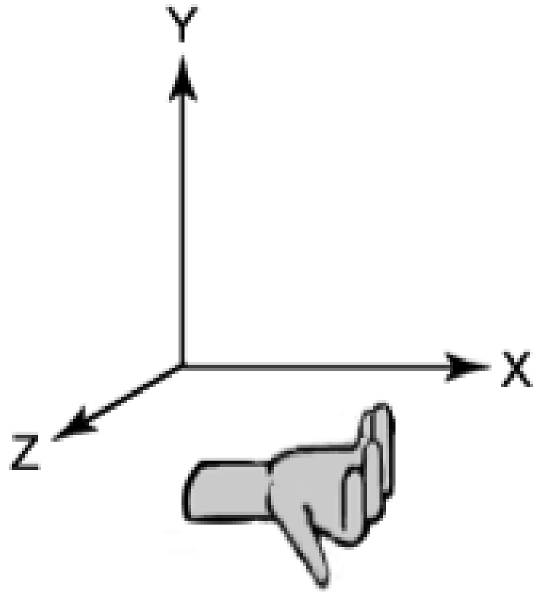


What is the transformation between images on the two image planes?

Pinhole camera model

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

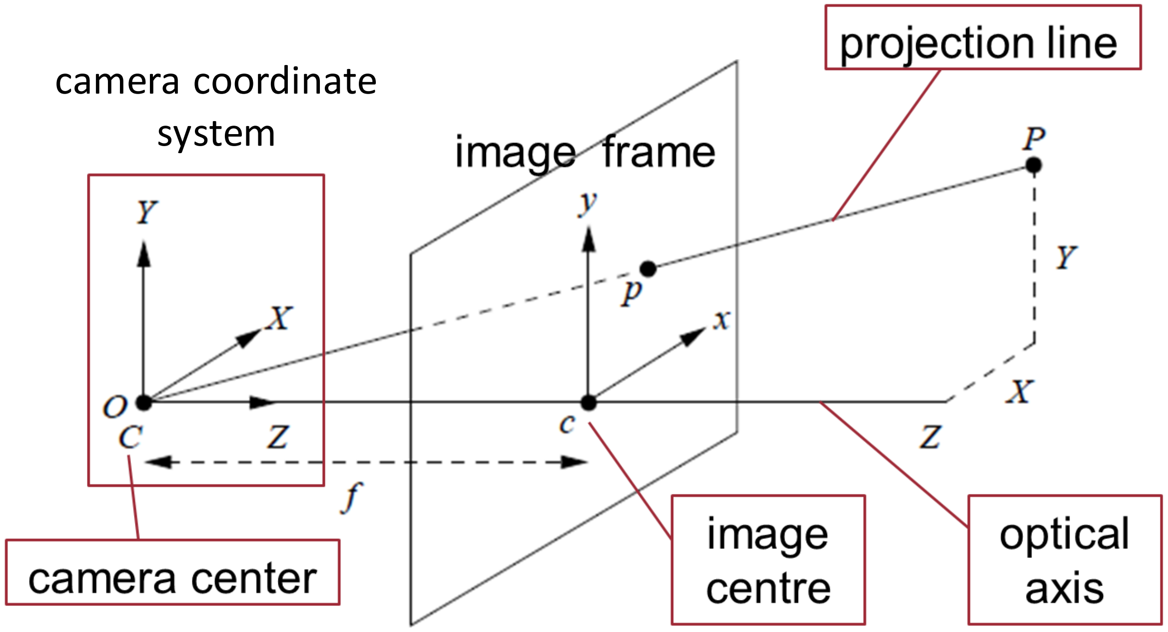




Right-handed coordinate system

Pinhole camera model

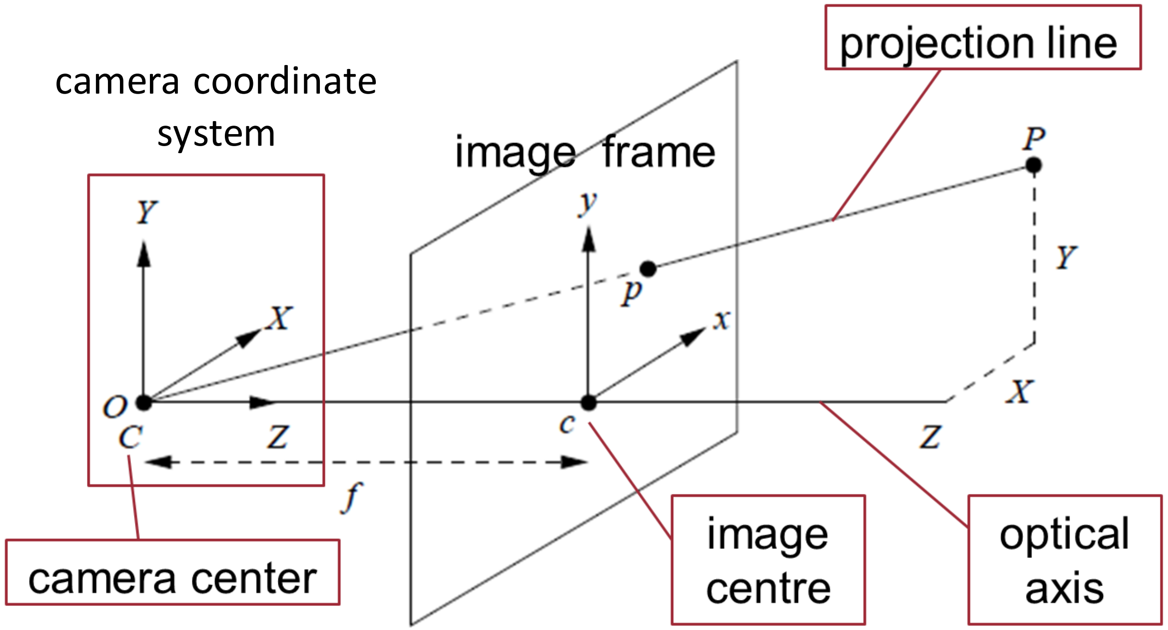
- 3D point $\mathbf{P} = (X, Y, Z)^T$ projected to 2D image $\mathbf{p} = (x, y)^T$



$x =$  $y =$ 

Pinhole camera model

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



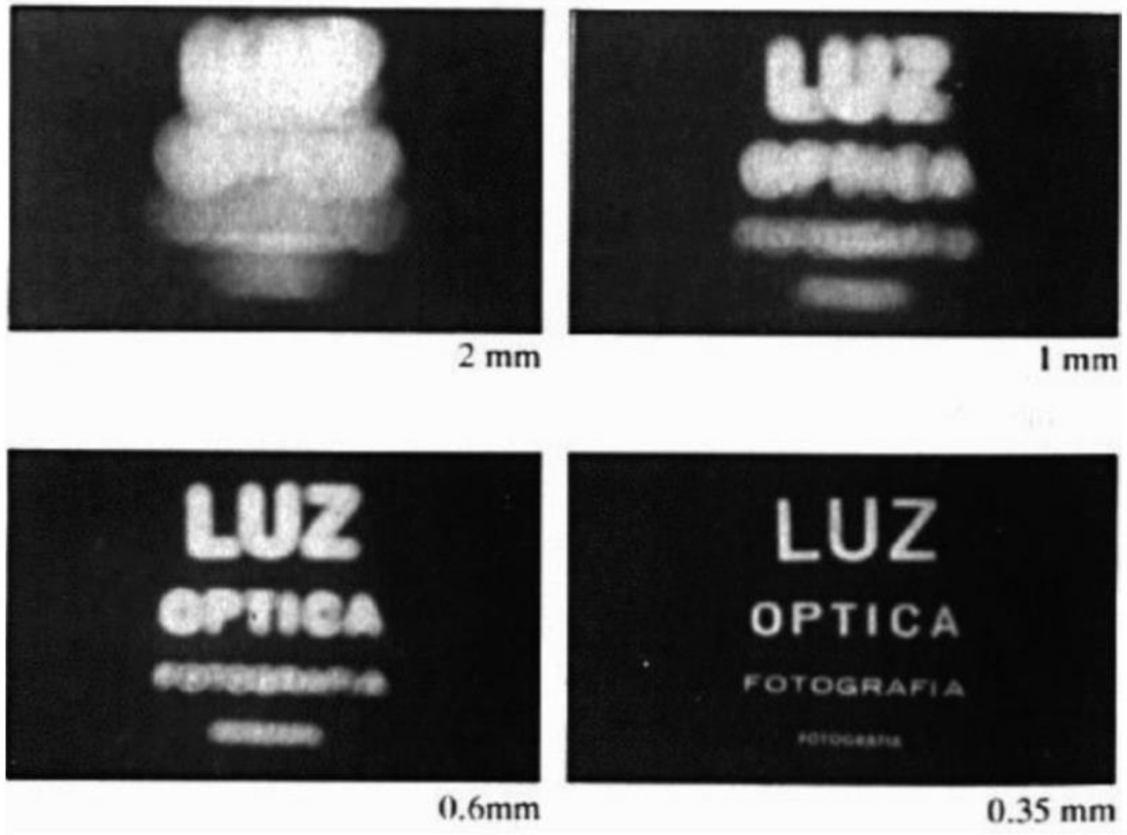
$$\frac{X}{Z} = \frac{x}{f}, \qquad \frac{Y}{Z} = \frac{y}{f}$$

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

Simplest form of perspective projection

Pinhole camera model

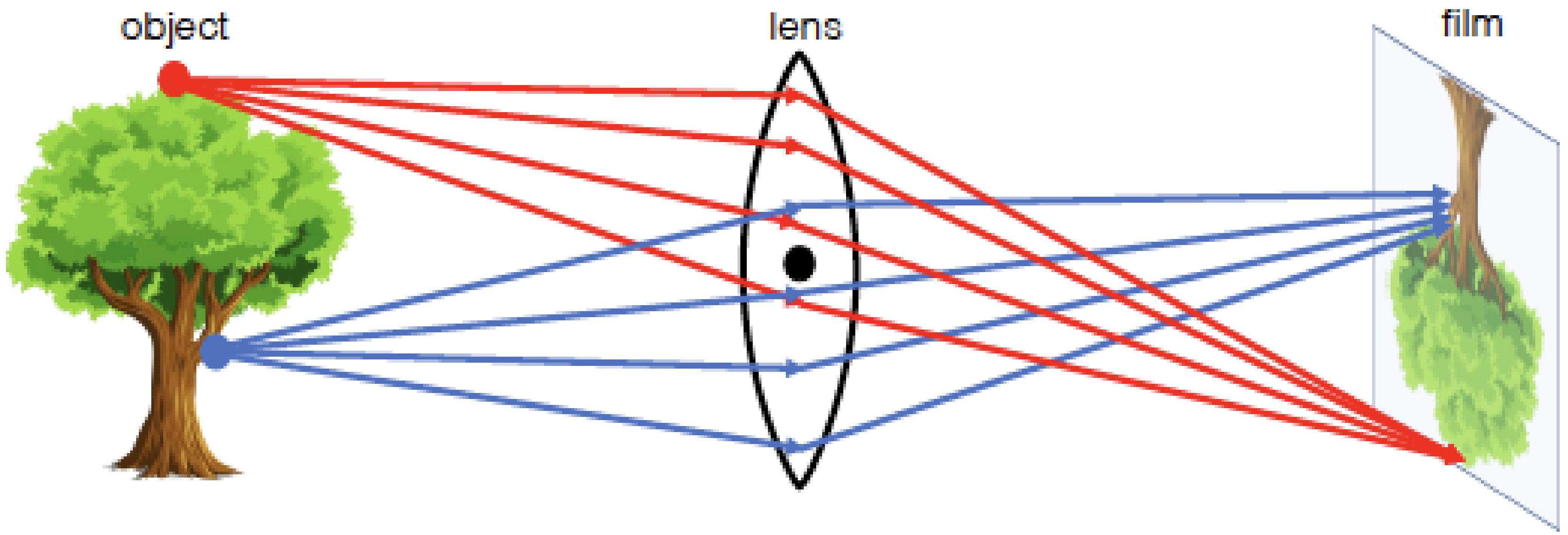
- Compromise between sharpness and brightness



Can we develop cameras that take images both sharp and bright?

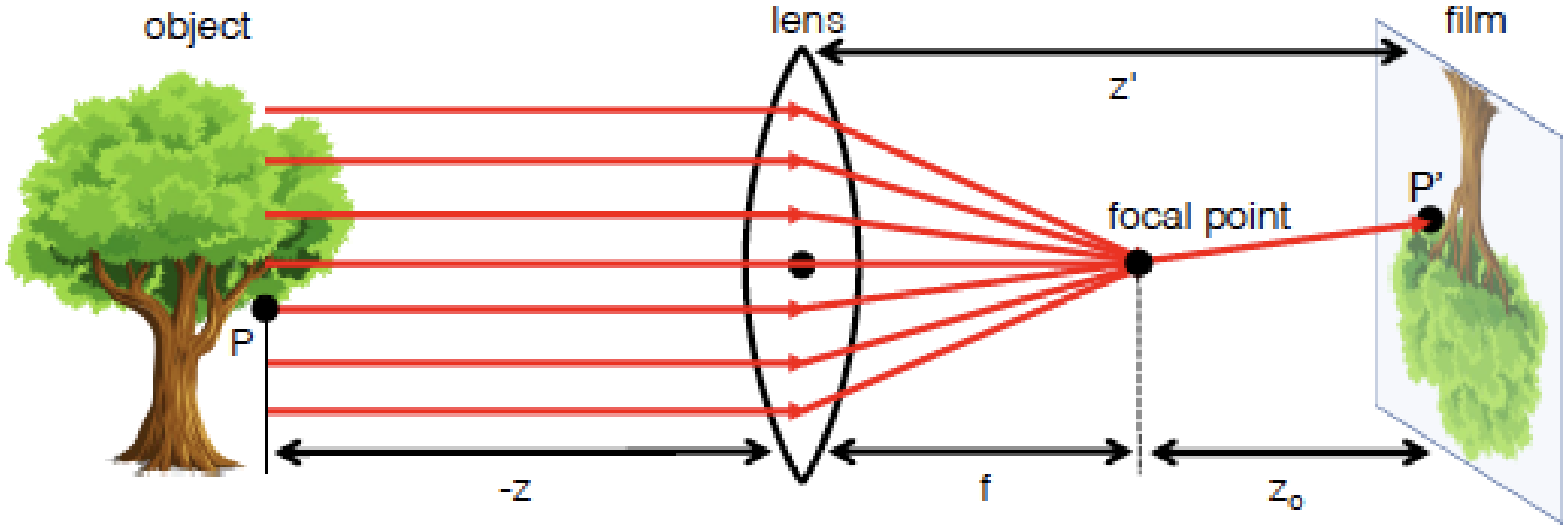
Perspective projection model

- Balancing sharpness and brightness using lenses
  - Replace the pinhole with a lens
  - The lens refracts light
  - The lens converges all rays of light of same points to a single point



Perspective projection model

- Lens
  - Converges all rays of light of same points to a single point
  - Focuses parallel light rays to the focal point



Perspective projection model

- Pinhole camera model

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

Image plane coordinates in physical measurements (*mm*)

- Change of unit: physical measurements -> pixels

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

What if $k = l$

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

Perspective projection model

- Pinhole camera model

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

Image plane coordinates in physical measurements (mm)

- Change of unit: physical measurements -> pixels

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

Denote $\alpha = kf, \beta = lf$

$$x = \alpha\frac{X}{Z}, \qquad y = \beta\frac{Y}{Z}$$

Perspective projection model

- Change of coordinate system
  - Image plane coordinates have origin at image center
  - Digital image coordinates have origin at top-left corner

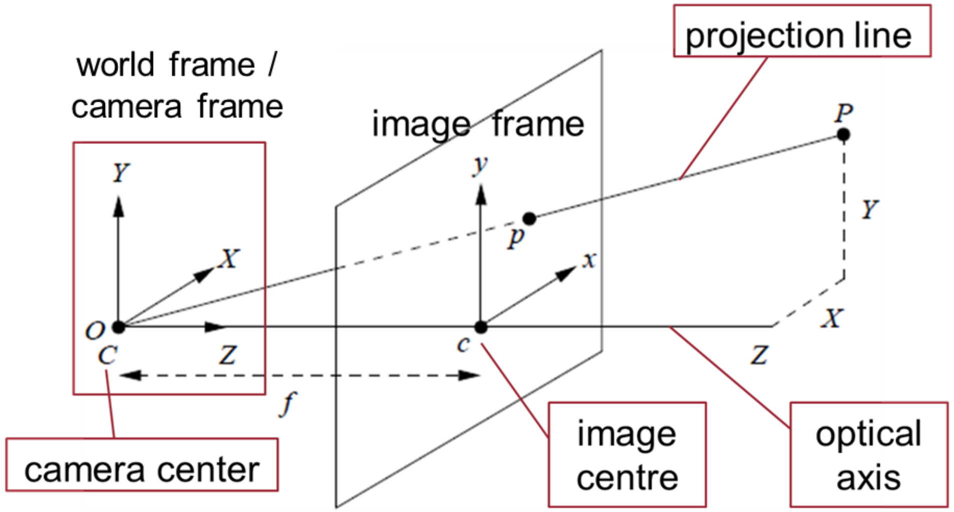


Image plane coordinates

Digital image coordinates

Perspective projection model

- Change of coordinate system
  - Image plane coordinates have origin at image center
  - Digital image coordinates have origin at top-left corner

$$x = \alpha \frac{X}{Z}, \qquad y = \beta \frac{Y}{Z}$$

Image center (principal point): (c_x, c_y)

Perspective projection model

- Change of coordinate system
  - Image plane coordinates have origin at image center
  - Digital image coordinates have origin at top-left corner

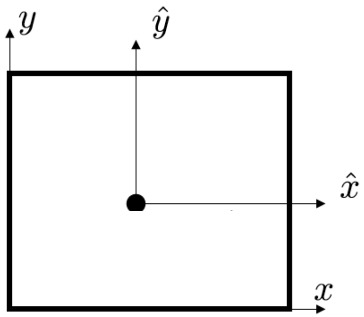
$$x = \alpha \frac{X}{Z}, \qquad y = \beta \frac{Y}{Z}$$

↓ Image center (principal point): (c_x, c_y)

$$x = \alpha \frac{X}{Z} + c_x, \qquad y = \beta \frac{Y}{Z} + c_y$$

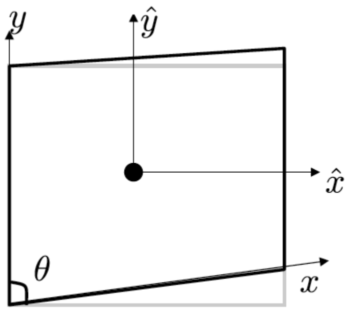
Perspective projection model

- Image frame may not be exactly rectangular
  - Due to sensor manufacturing errors



$$x = \alpha \frac{X}{Z} + c_x, \qquad y = \beta \frac{Y}{Z} + c_y$$

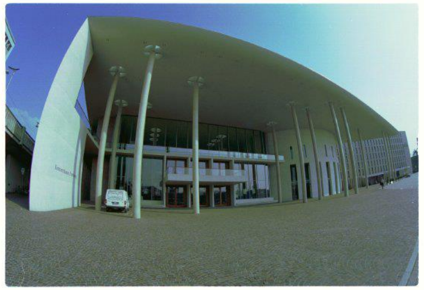
θ: skew angle between x- and y-axis

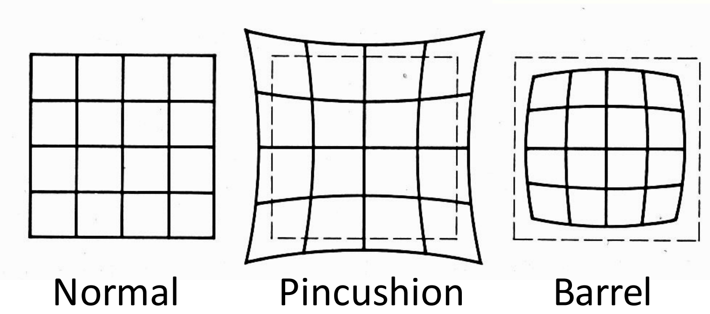


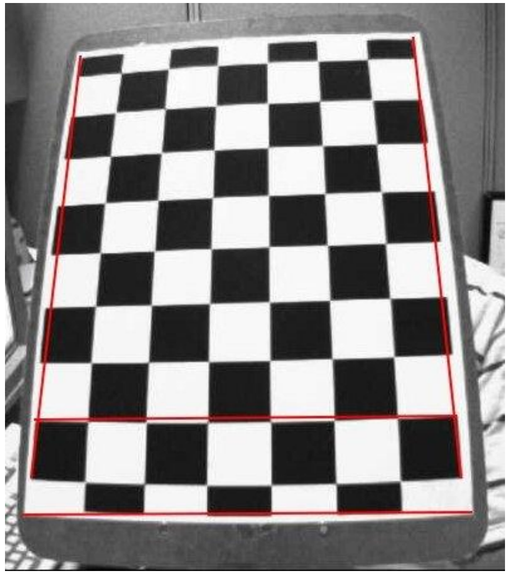
$$x = \alpha \frac{X}{Z} - \alpha \cot\theta \frac{Y}{Z} + c_x, \qquad y = \frac{\beta}{\sin\theta} \frac{Y}{Z} + c_y$$

Perspective projection model

- Other types of distortions
  - Common aberration: radial distortion
    - Different portions of the lens have differing focal lengths
    - Straight lines appear curved
    - Errors by radial distortion << scanning resolution of image





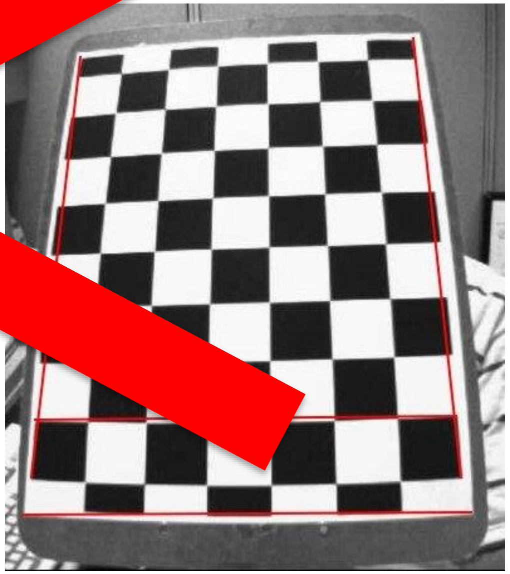


Perspective projection model

- Other types of distortions
  - Common aberration: radial distortion
    - Different portions of the lens have differing focal lengths
    - Straight lines appear curved
    - Errors by radial distortion << scanning resolution of image

$$x_{\text{corrected}} = x(1 + k_1 r^2 + k_2 r^4 + k_3 r^6)$$

$$y_{\text{corrected}} = y(1 + k_1 r^2 + k_2 r^4 + k_3 r^6)$$



Intrinsic parameters

$$x = \alpha \frac{X}{Z} - \alpha \cot\theta \frac{Y}{Z} + c_x, \qquad y = \frac{\beta}{\sin\theta} \frac{Y}{Z} + c_y$$

- Rewrite in matrix-vector product form

$$\mathbf{P} = [X, Y, Z]^{\mathrm{T}}, \quad \mathbf{p} = [x, y, 1]^{\mathrm{T}}$$

(homogeneous coordinates)

$$\mathbf{p} = K\mathbf{P}, \qquad K =$$

Intrinsic parameters

$$x = \alpha \frac{X}{Z} - \alpha \cot\theta \frac{Y}{Z} + c_x, \qquad y = \frac{\beta}{\sin\theta} \frac{Y}{Z} + c_y$$

- Rewrite in matrix-vector product form

$$\mathbf{P} = [X, Y, Z]^{\mathrm{T}}, \quad \mathbf{p} = [x, y, 1]^{\mathrm{T}}$$

(homogeneous coordinates)

$$\mathbf{p} = K\mathbf{P}, \qquad K = \begin{bmatrix} \alpha & -\alpha \cot\theta & c_x \\ 0 & \dfrac{\beta}{\sin\theta} & c_y \\ 0 & 0 & 1 \end{bmatrix}$$

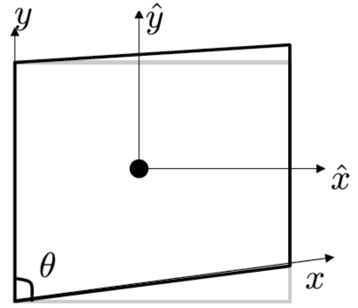
Intrinsic parameter matrix

Intrinsic parameters

- For simplicity, people use a simpler form of K

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

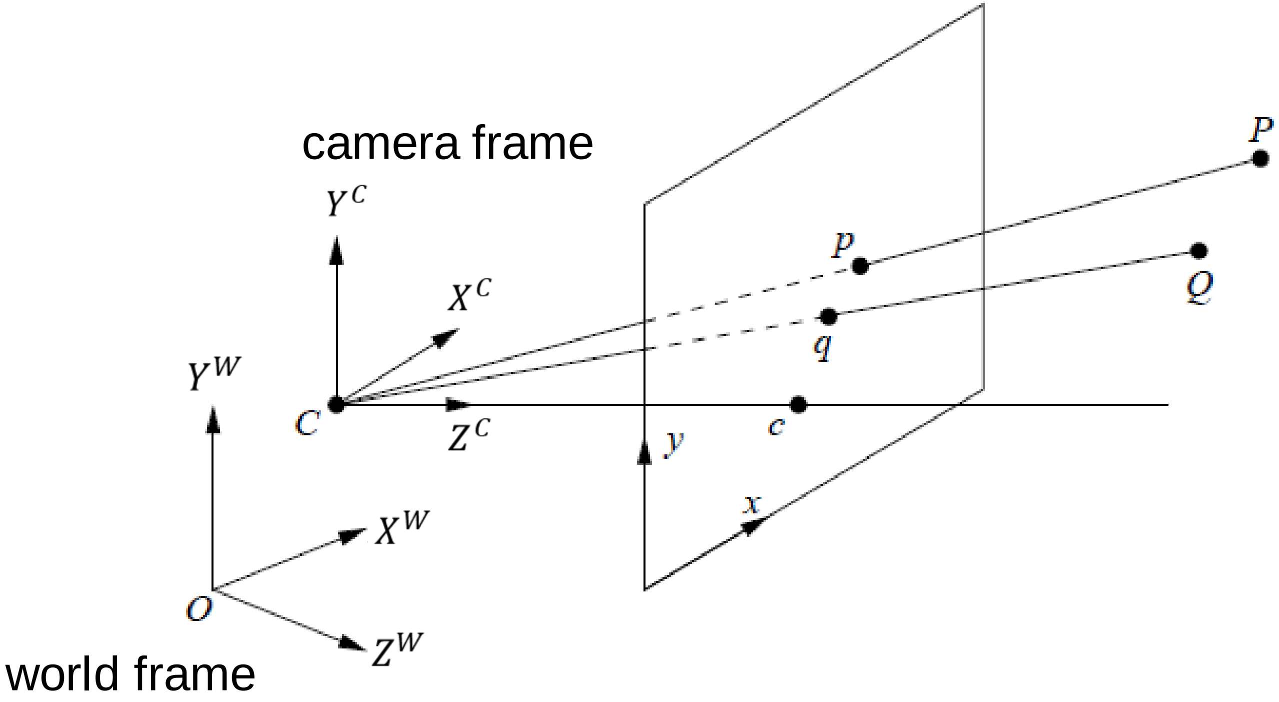
s: skew parameter



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

Extrinsic parameters

- Camera motion
  - World frame may not align with the camera frame
  - Camera can move and rotate



Extrinsic parameters

- Camera motion
  - World frame may not align with the camera frame
  - Camera can move and rotate

$$\mathbf{P}^{C} = R_{W}^{C} \mathbf{P}^{W} + \mathbf{t}_{W}^{C}$$

1 (for $\mathbf{P}^{C}$), 3 (for R_{W}^{C}), 2 (for $\mathbf{P}^{W}$), 4 (for $\mathbf{t}_{W}^{C}$)

Camera frame (superscript C of $\mathbf{t}_{W}^{C}$)

World frame (subscript W of $\mathbf{t}_{W}^{C}$)

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.

Complete camera model

- Combine intrinsic and extrinsic parameters

$$\mathbf{p} = K\mathbf{P}, \qquad \mathbf{P}^C = R_W^C \cdot \mathbf{P}^W + \mathbf{t}_W^C$$

$$\Downarrow$$

$$\mathbf{p} = K(R_W^C \cdot \mathbf{P}^W + \mathbf{t}_W^C)$$

- Use a simpler notation

$$\begin{aligned} \mathbf{p} &= K(R\mathbf{P} + \mathbf{t}) \\ &= K[R\ \mathbf{t}]\mathbf{P} \\ &= M\mathbf{P} \end{aligned}$$

$M = K[R\ \mathbf{t}]$: 3 x 4 projection matrix

Summary

- 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