GEO1016
Photogrammetry and 3D Computer Vision

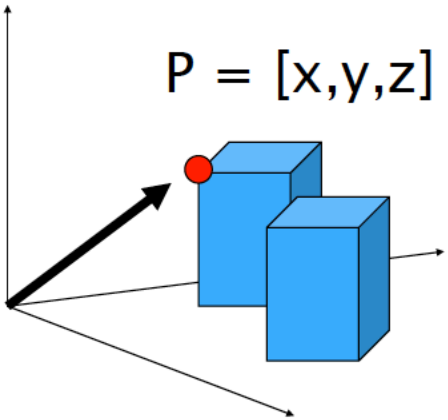
Lecture 2
Camera Models

Liangliang Nan

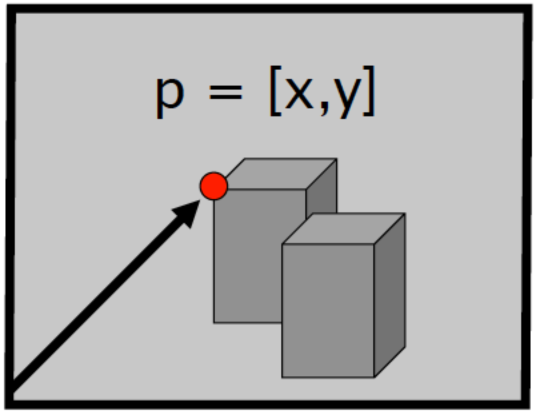
Today's Agenda

- Review of basic linear algebra
- Images
- Camera Models

Vectors (i.e., 2D and 3D vectors)



3D world

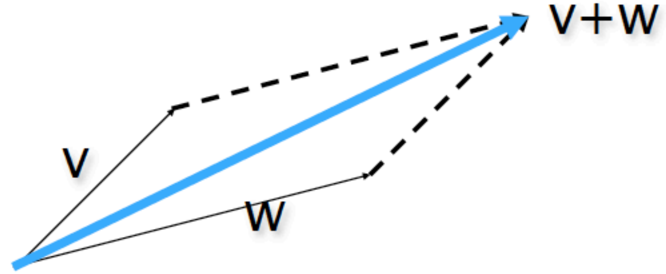


Image

Vector arithmetic

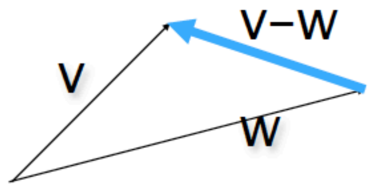
- Addition

$$\mathbf{v} + \mathbf{w} = (x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2)$$



- Subtraction

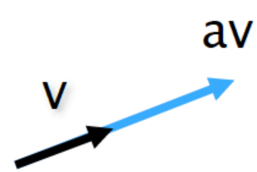
$$\mathbf{v} - \mathbf{w} = (x_1, x_2) - (y_1, y_2) = (x_1 - y_1, x_2 - y_2)$$



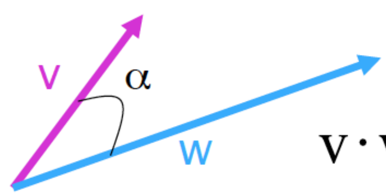
Vector arithmetic

- Scalar Product

$$a\mathbf{v} = a(x_1, x_2) = (ax_1, ax_2)$$



- Dot (inner) product



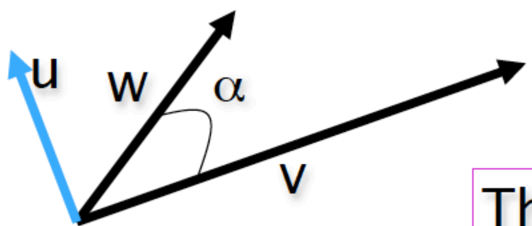
$$\mathrm{v} \cdot \mathrm{w} = (\mathrm{x}_1, \mathrm{x}_2) \cdot (\mathrm{y}_1, \mathrm{y}_2) = \mathrm{x}_1\mathrm{y}_1 + \mathrm{x}_2\mathrm{y}_2$$

The inner product is a SCALAR!

$$\mathrm{v} \cdot \mathrm{w} = (\mathrm{x}_1, \mathrm{x}_2) \cdot (\mathrm{y}_1, \mathrm{y}_2) = \| \mathrm{v} \| \cdot \| \mathrm{w} \| \cos\alpha$$

Vector arithmetic

- Vector (cross) Product

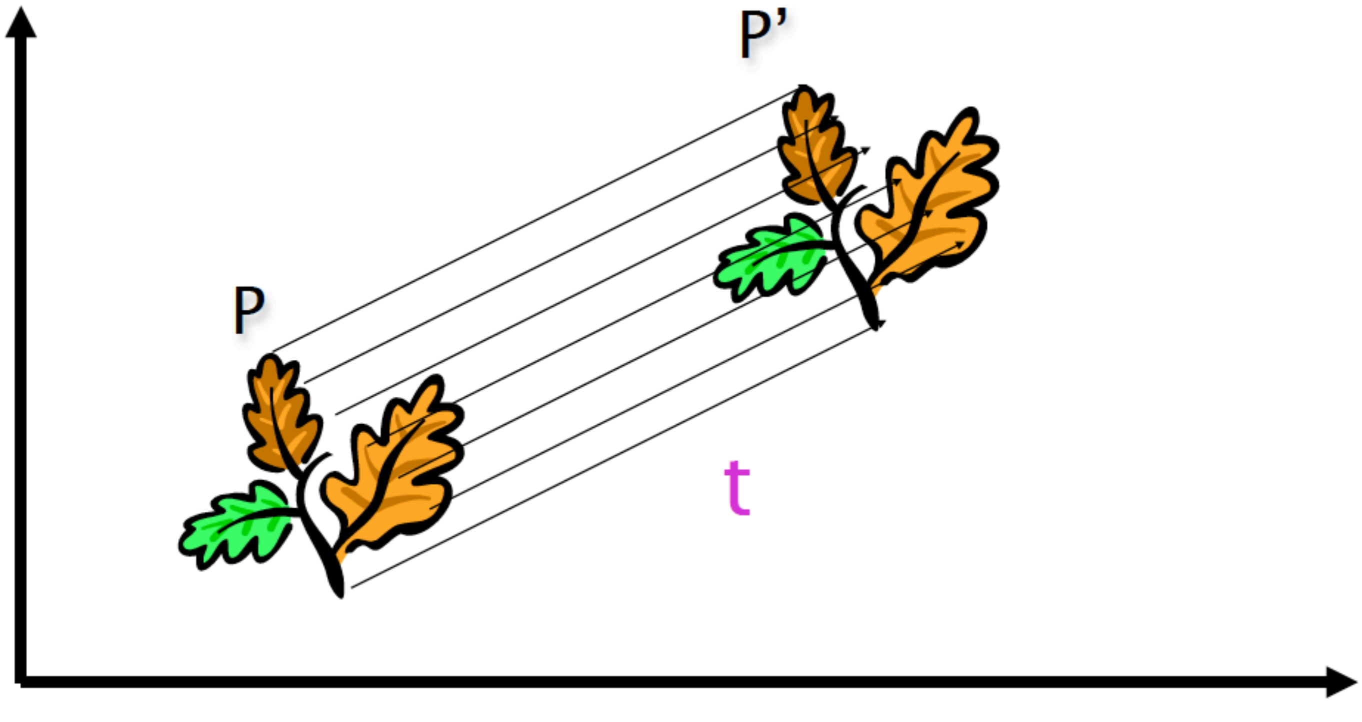


$$u = v \times w$$

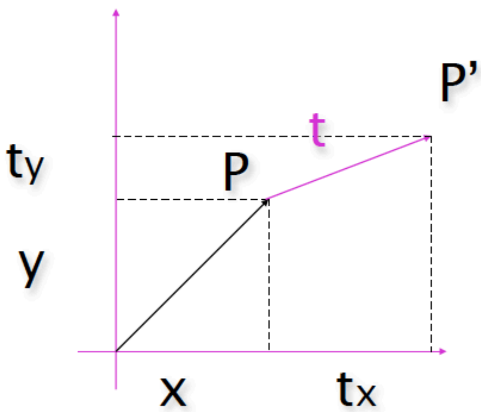
The cross product is a VECTOR!

$$\text{Magnitude:} ||u|| = ||v \times w|| = ||v|| ||w|| \sin \alpha$$

Translation



Translation



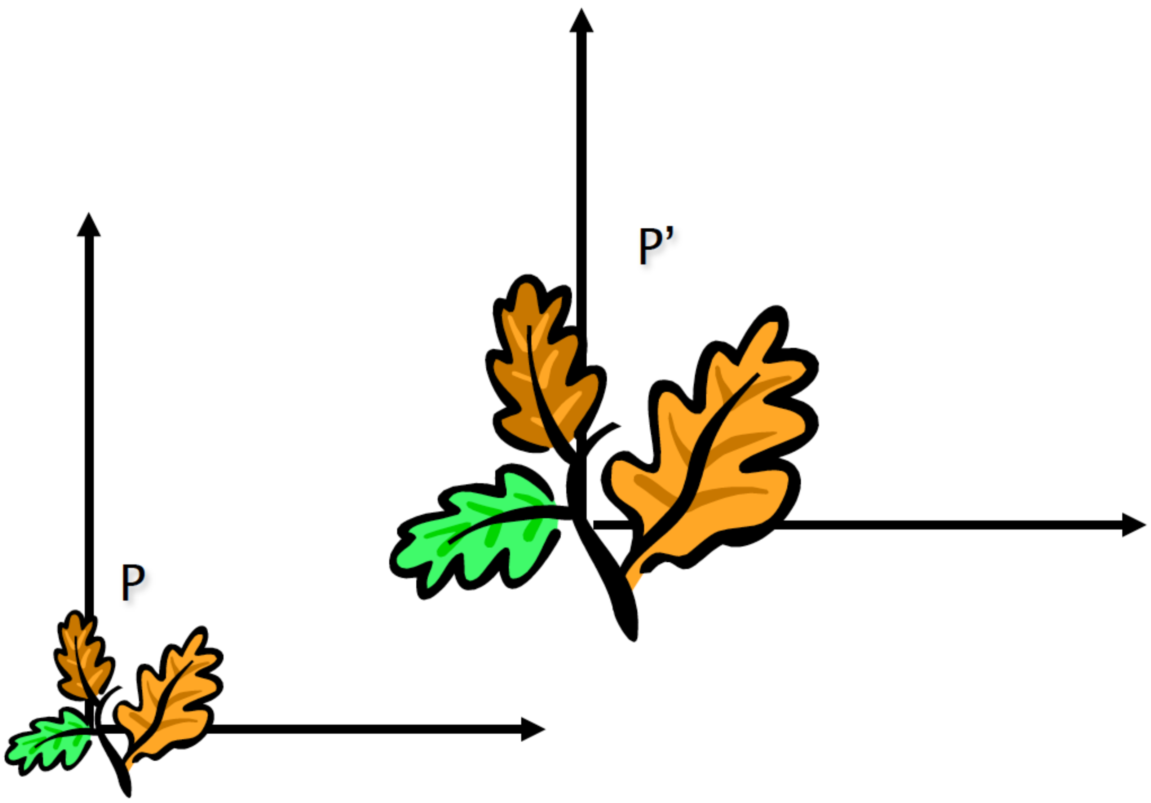
$$\mathbf{P} = (x, y) \rightarrow (x, y, 1)$$

$$\mathbf{t} = (t_x, t_y) \rightarrow (t_x, t_y, 1)$$

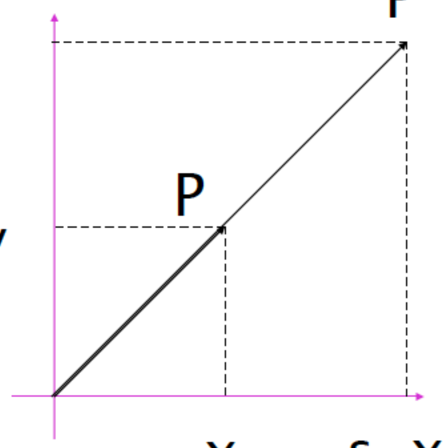
$$\mathbf{P'} \rightarrow \begin{bmatrix} x + t_x \\ y + t_y \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{I} & \mathbf{t} \\ 0 & 1 \end{bmatrix} \cdot \mathbf{P} = \mathbf{T} \cdot \mathbf{P}$$

Scaling



Scaling Equation



$$\mathbf{P} = (x, y) \rightarrow \mathbf{P}' = (s_x x, s_y y)$$

$$\mathbf{P} = (x, y) \rightarrow (x, y, 1)$$

$$\mathbf{P}' = (s_x x, s_y y) \rightarrow (s_x x, s_y y, 1)$$

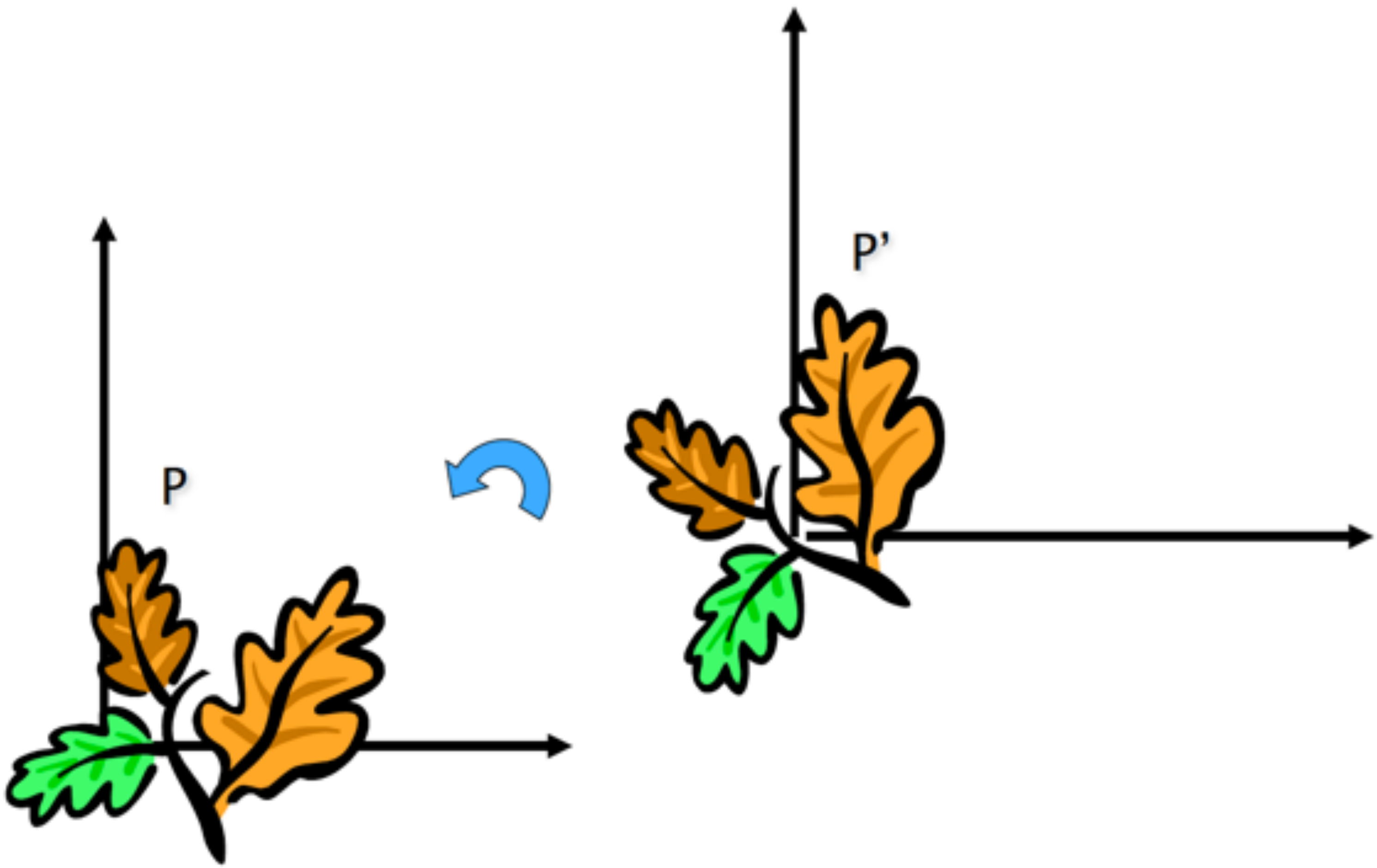
$$\mathbf{P}' \rightarrow \begin{bmatrix} s_x x \\ s_y y \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\mathbf{S}} \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{S}' & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{bmatrix} \cdot \mathbf{P} = \mathbf{S} \cdot \mathbf{P}$$

Scaling & Translation

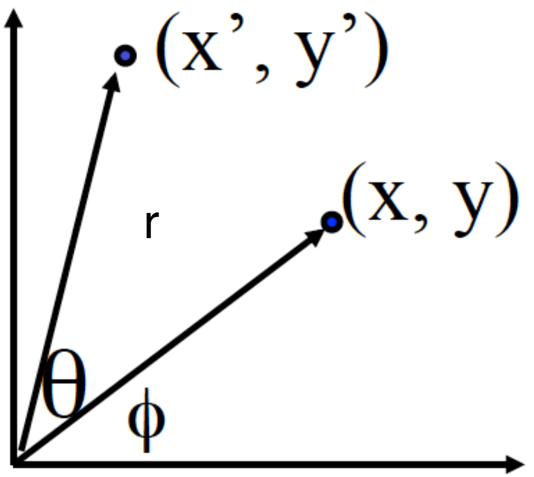
$$\mathbf{P''} = \mathbf{T} \cdot \mathbf{S} \cdot \mathbf{P} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} =$$

$$= \underbrace{\begin{bmatrix} s_x & 0 & t_x \\ 0 & s_y & t_y \\ 0 & 0 & 1 \end{bmatrix}}_{A} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Rotation



Rotation Equations



$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \underbrace{\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}}_{\mathbf{R}} \begin{bmatrix} x \\ y \end{bmatrix}$$

What is the inverse transformation

- Rotation by –θ

R has many interesting properties:

$$\mathbf{R}^{-1} = \mathbf{R}^{T} \qquad \mathbf{R} \cdot \mathbf{R}^{\mathrm{T}} = \mathbf{R}^{\mathrm{T}} \cdot \mathbf{R} = \mathbf{I} \qquad \det(\mathbf{R}) = 1$$

Translation + Rotation + Scaling

$$\mathbf{P}' = \mathbf{T} \cdot \mathrm{R} \cdot \mathbf{S} \cdot \mathbf{P} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} =$$

$$= \begin{bmatrix} \cos\theta & -\sin\theta & t_x \\ \sin\theta & \cos\theta & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} =$$

$$= \begin{bmatrix} R' & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} S & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} R'S & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

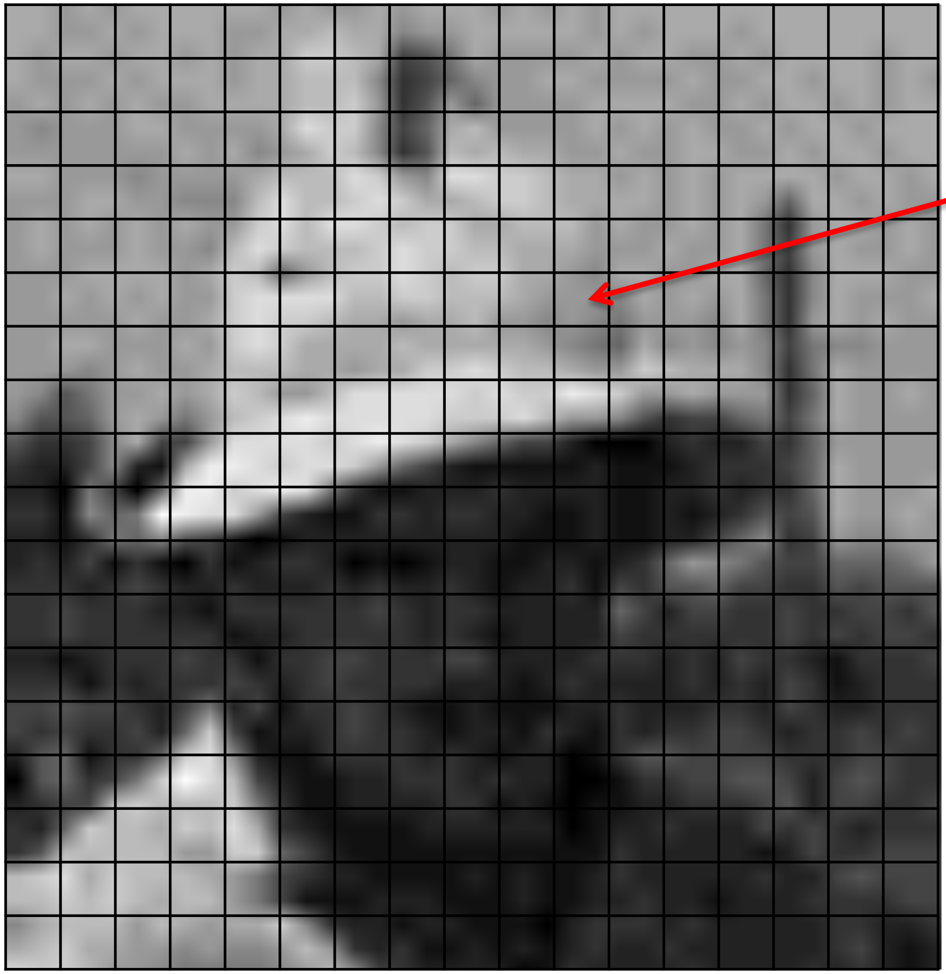
If $s_x = s_y$, this is a similarity transformation

Today’s Agenda

- Review of basic linear algebra
- Images
- Camera Models

What is an image?

What is an image?

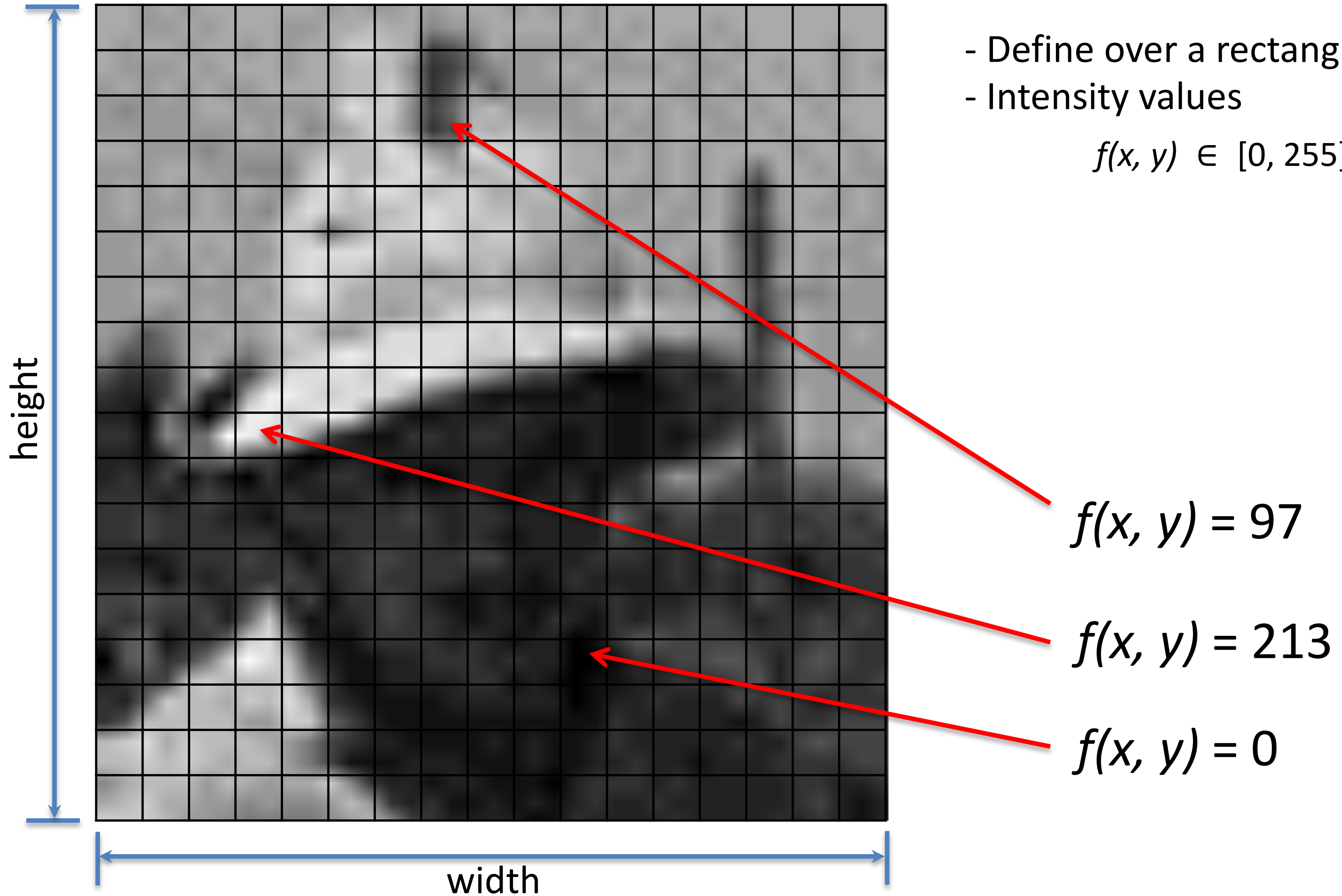


$$P = f(x, y)$$

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

Pixel

What is an image?

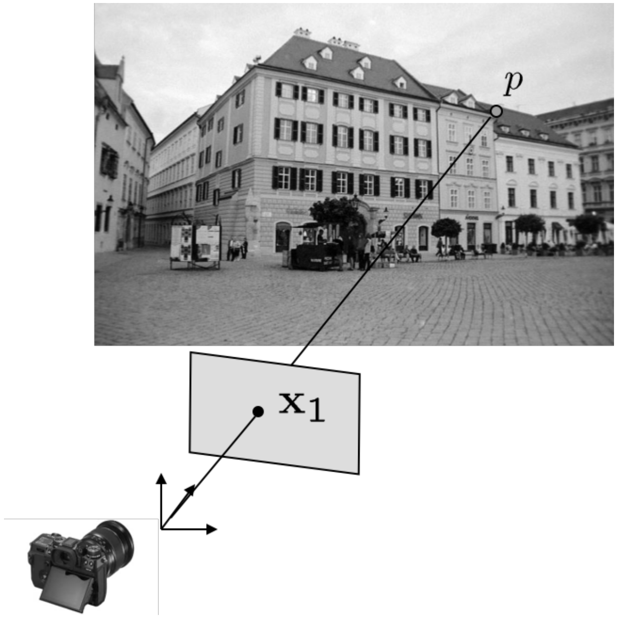


- Define over a rectangle
- Intensity values

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

What is an image?

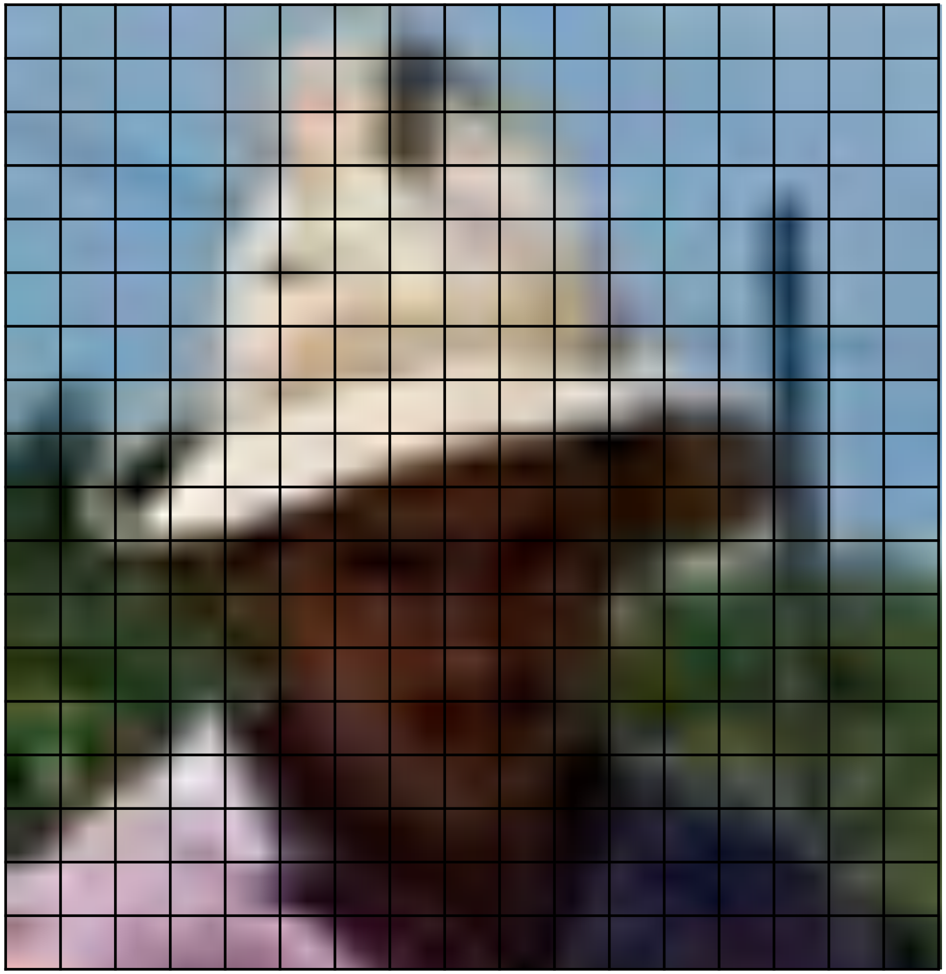
- Projection of the scene on the image plane
- Digital (discrete) image
  - A matrix of integer values



$j \rightarrow$, $i \downarrow$

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

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

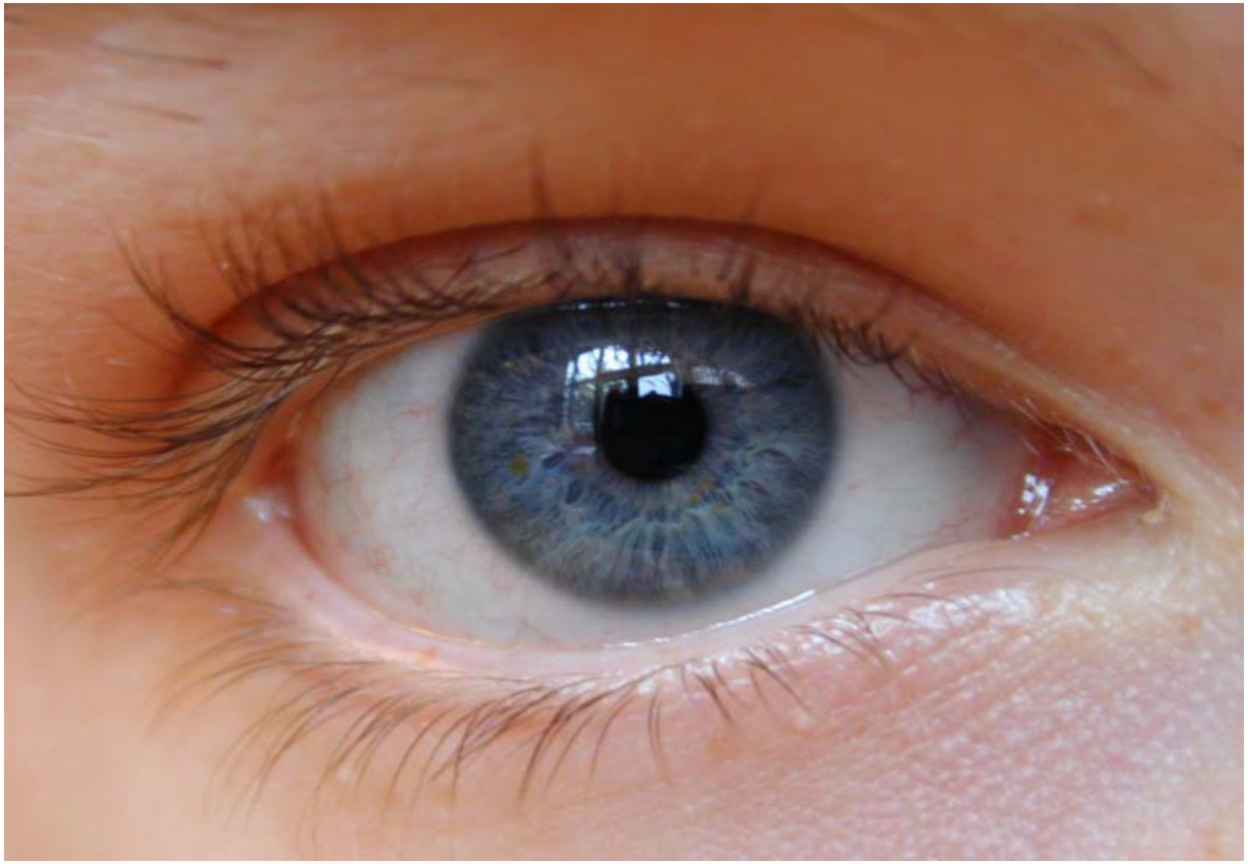
“vector-valued” function

Today’s Agenda

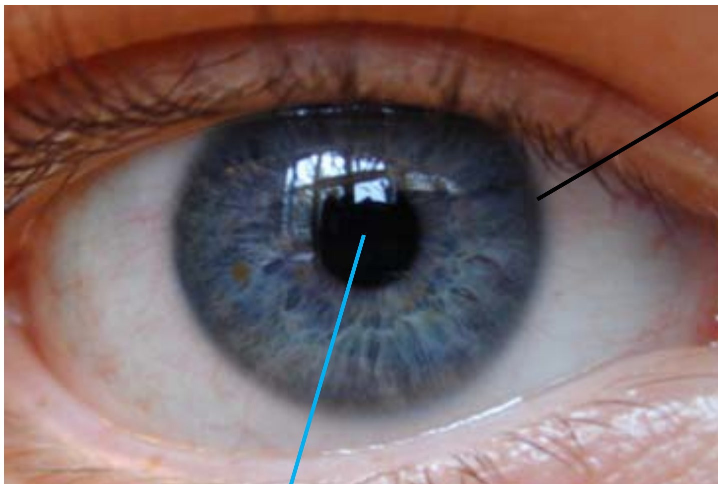
- Review of basic linear algebra
- Images
- Camera Models

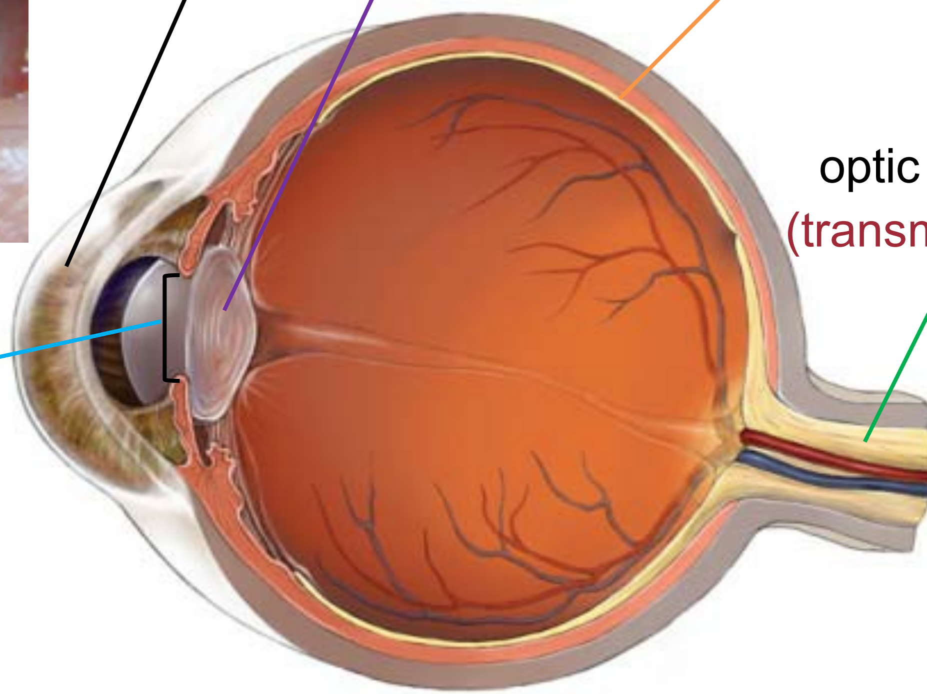
Through our eyes…

- We see the world



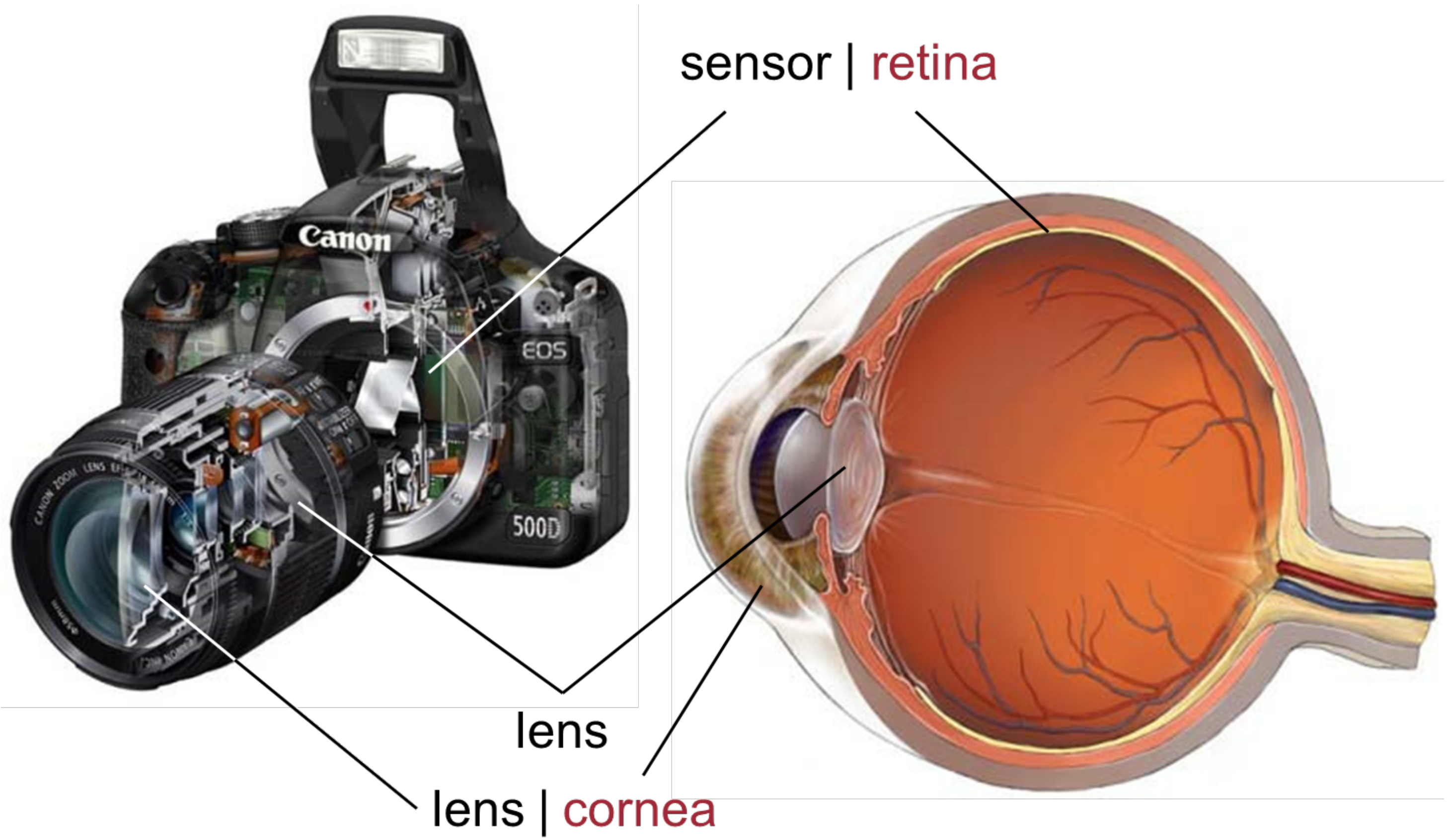
Through our eyes…





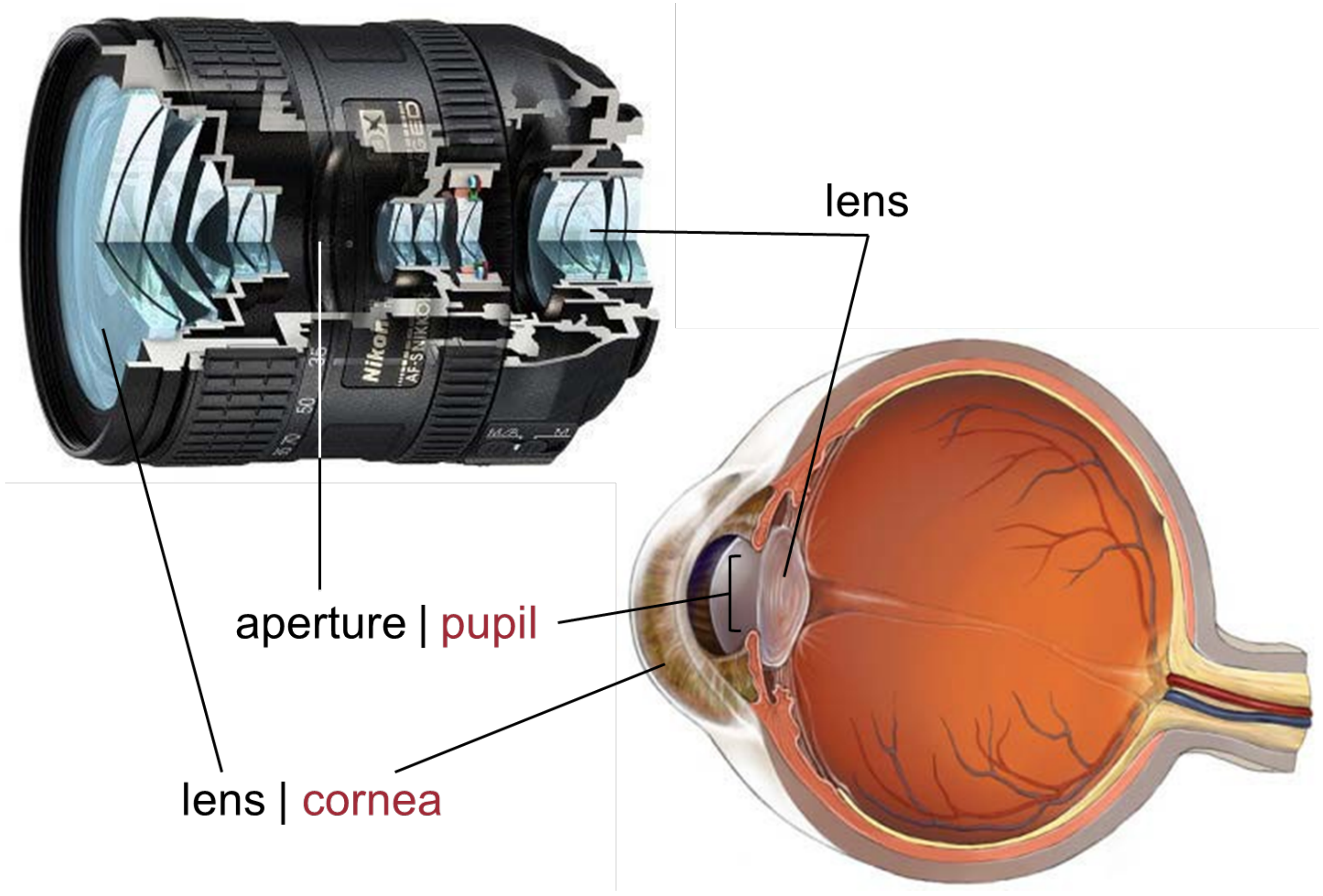
Through their eyes…

- Camera is structurally the same the eye.



Through their eyes…

- Camera is structurally the same the eye.



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

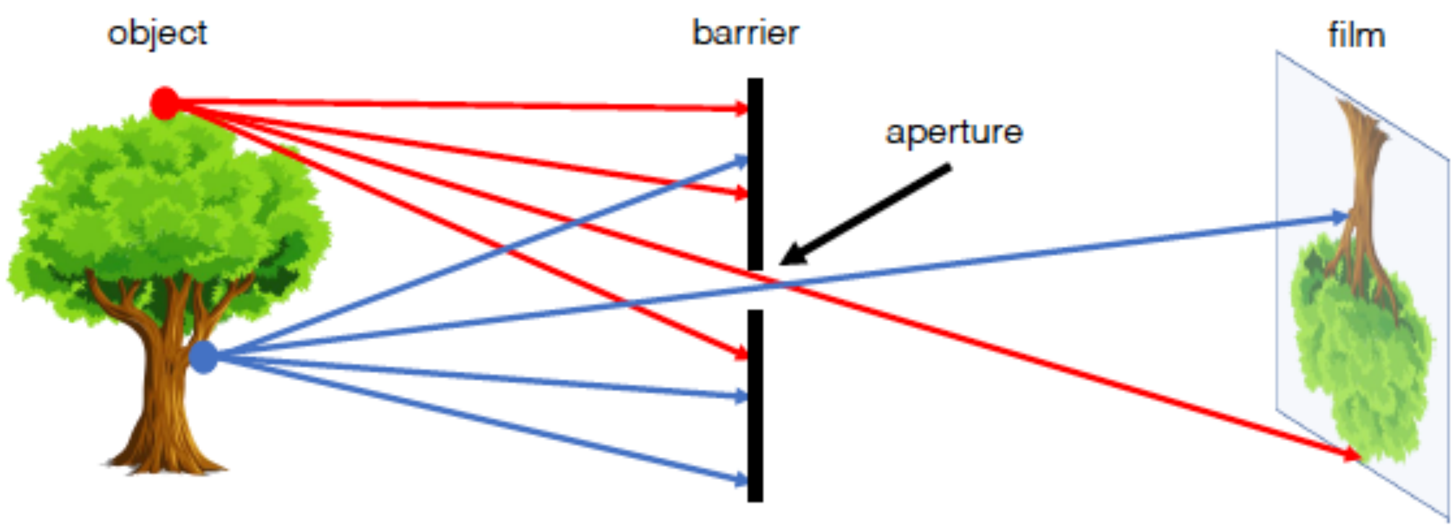
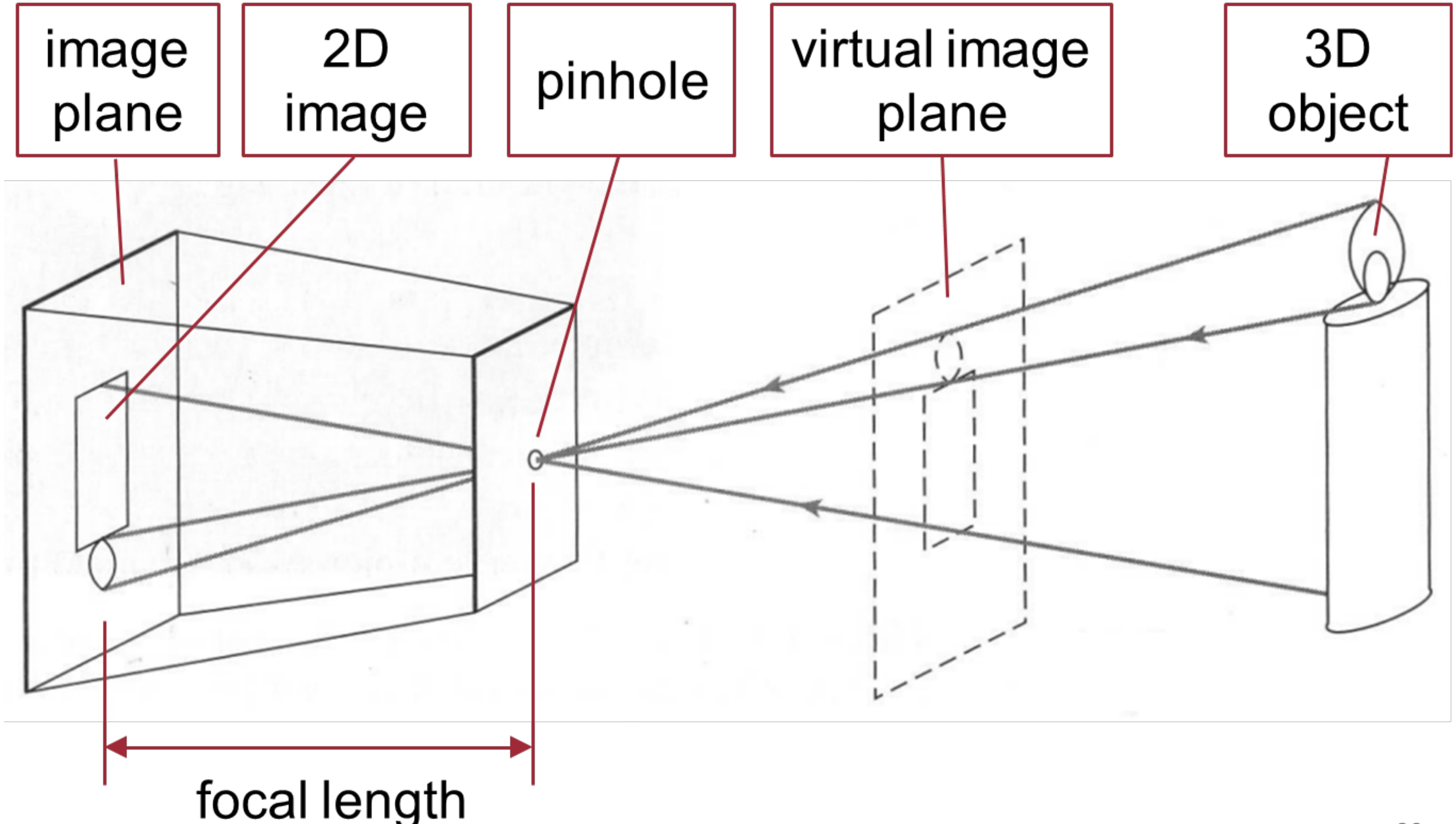


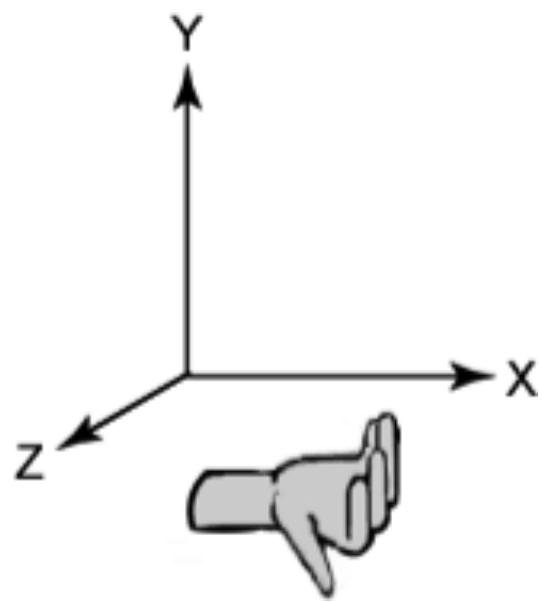
Figure 1: A simple working camera model: the pinhole camera model.

Pinhole camera model



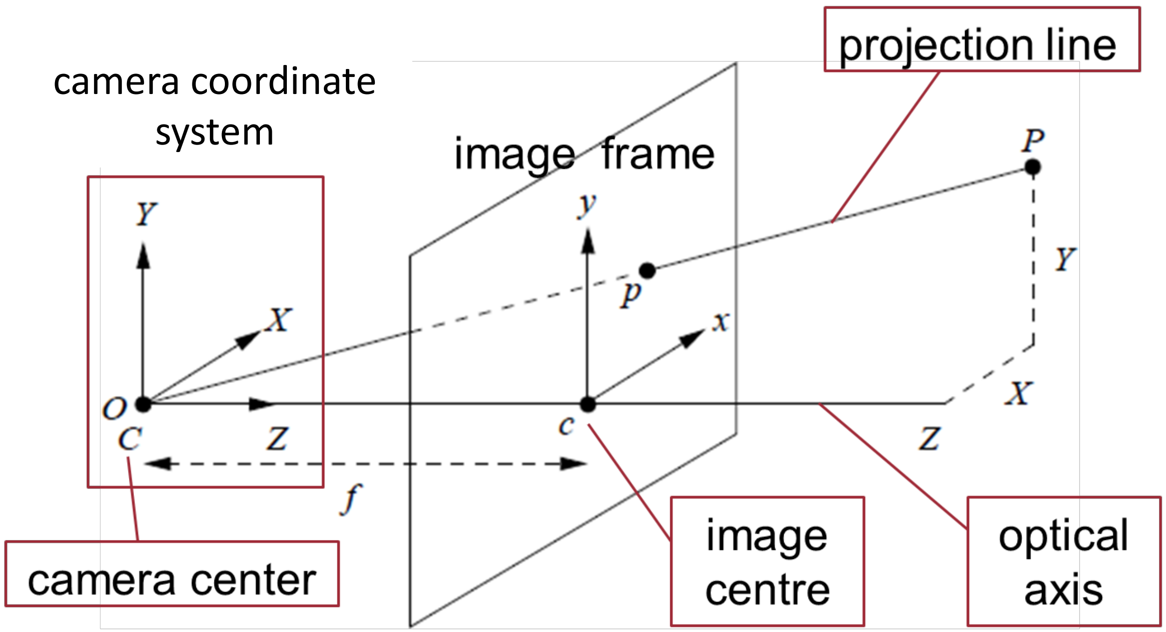
Pinhole camera model

- By default, use right-handed coordinate system



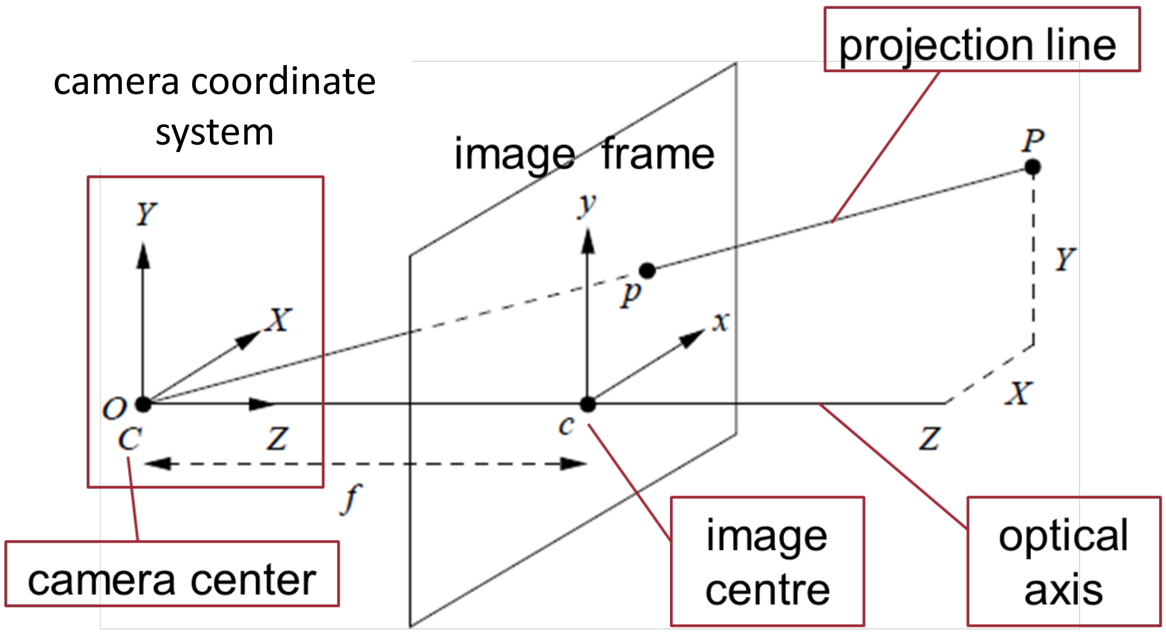
Pinhole camera model

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



Pinhole camera model

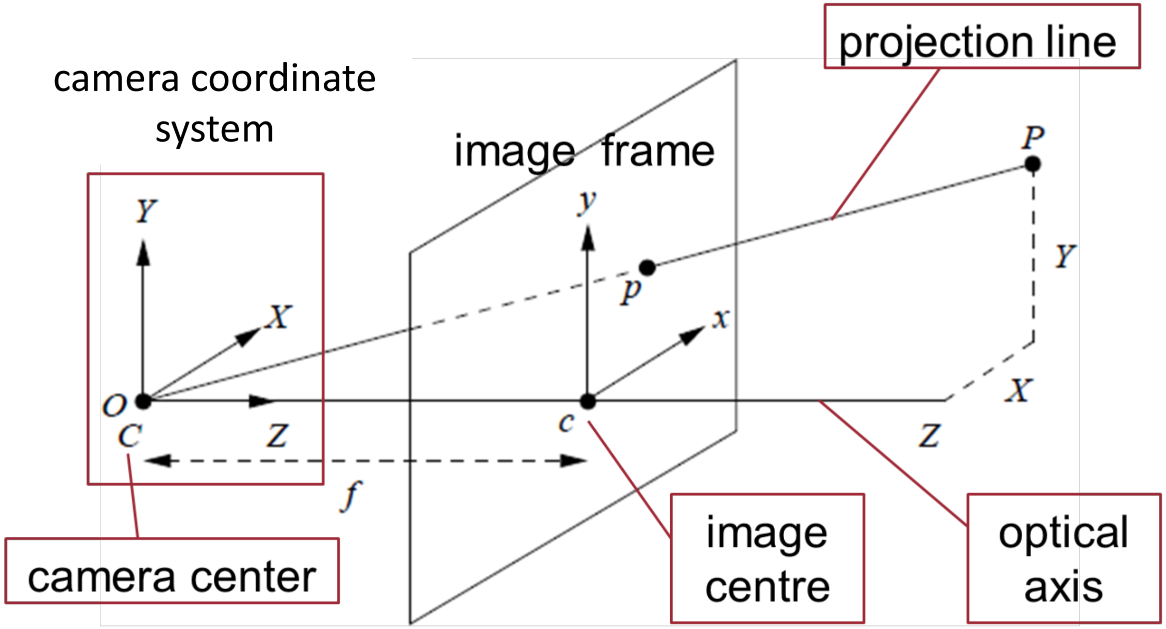
- 3D point $\mathbf{P} = (X, Y, Z)^{\mathrm{T}}$ projected to 2D image point $\mathbf{p} = (x, y)^{\mathrm{T}}$.



$x = ?$ $y = ?$

Pinhole camera model

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



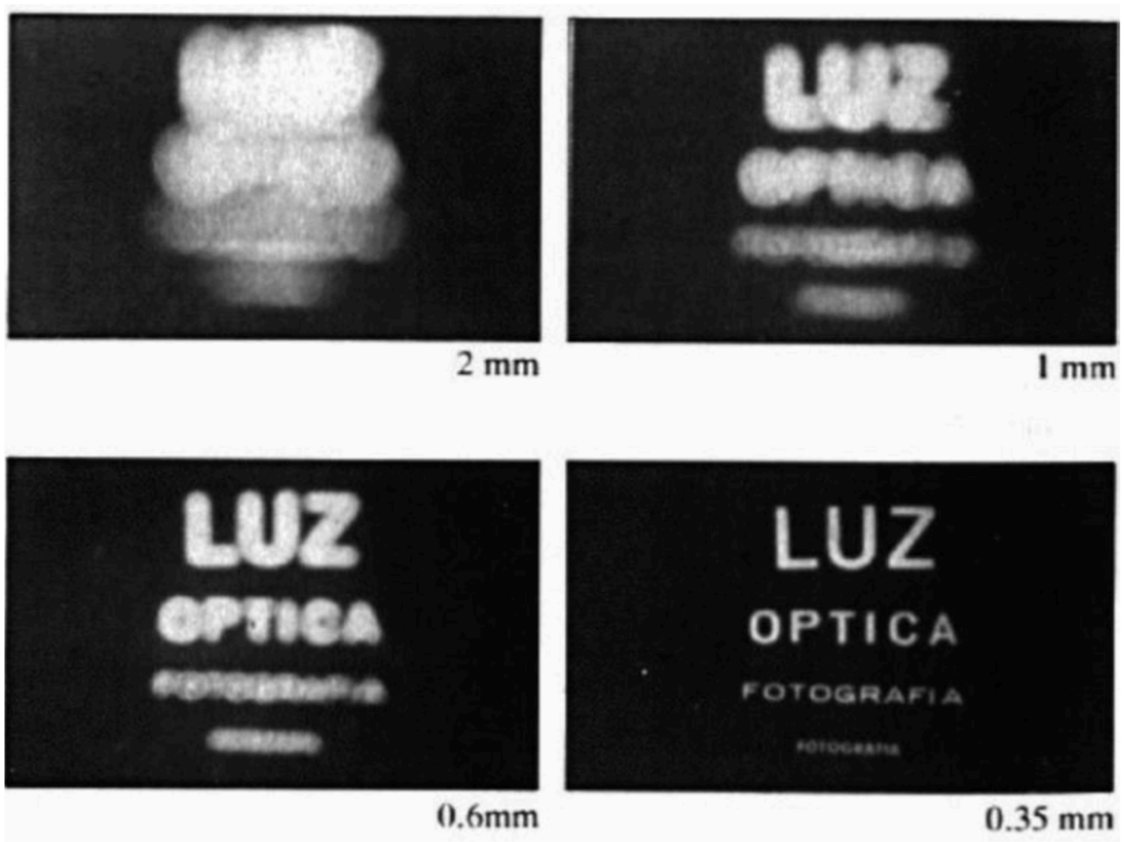
$$\frac{X}{Z} = \frac{x}{f}, \quad \frac{Y}{Z} = \frac{y}{f}$$

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

Simplest form of perspective projection

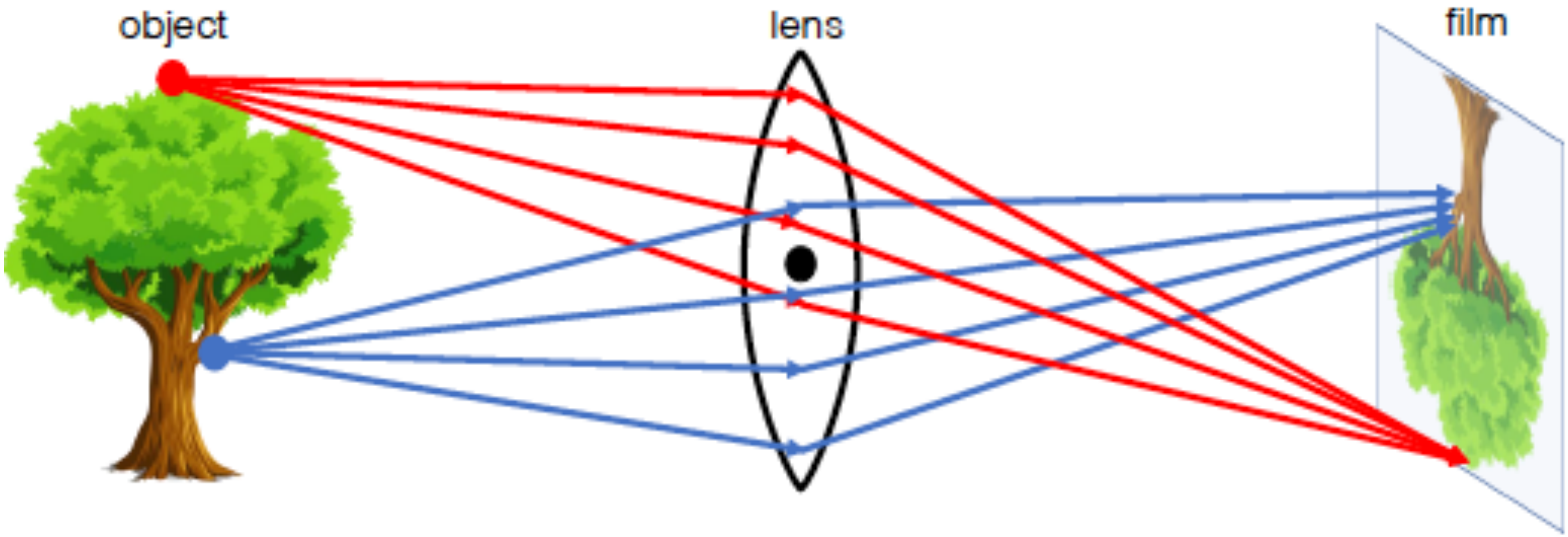
Pinhole camera model

- Assumption: aperture is a single point.



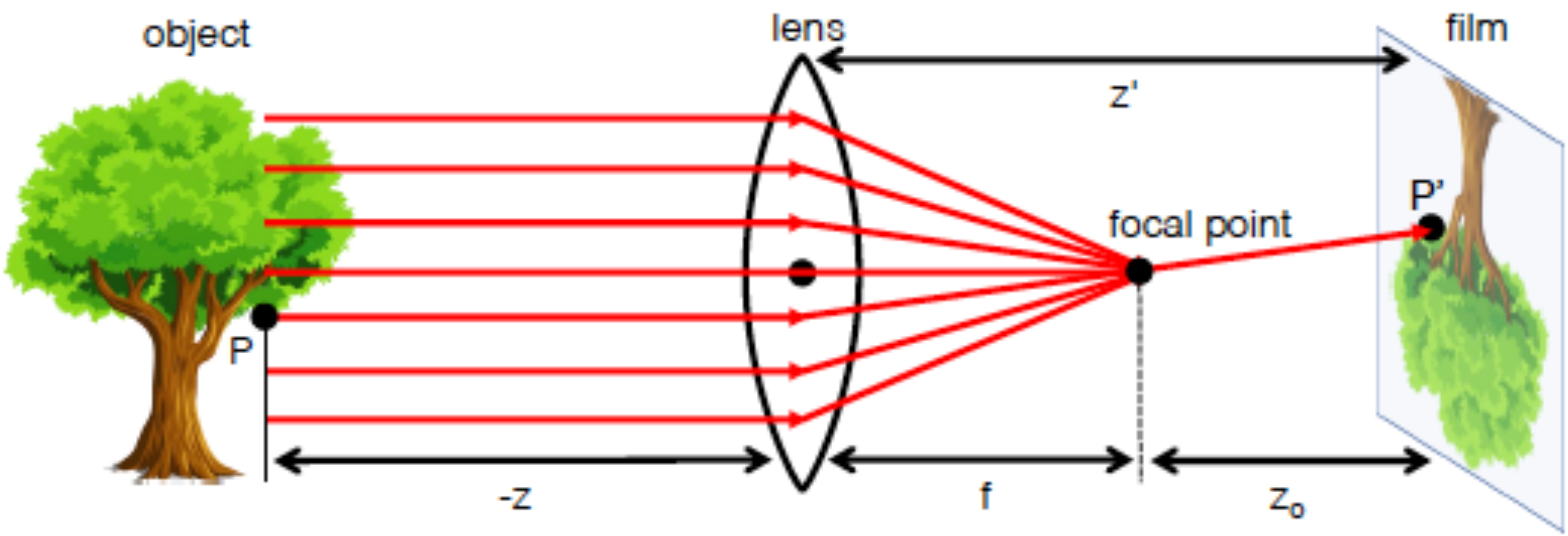
Perspective projection model

- Sharpness vs. brightness?
- Lens



Perspective projection model

- Sharpness vs. brightness?
- Lens



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/m)
  - 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

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

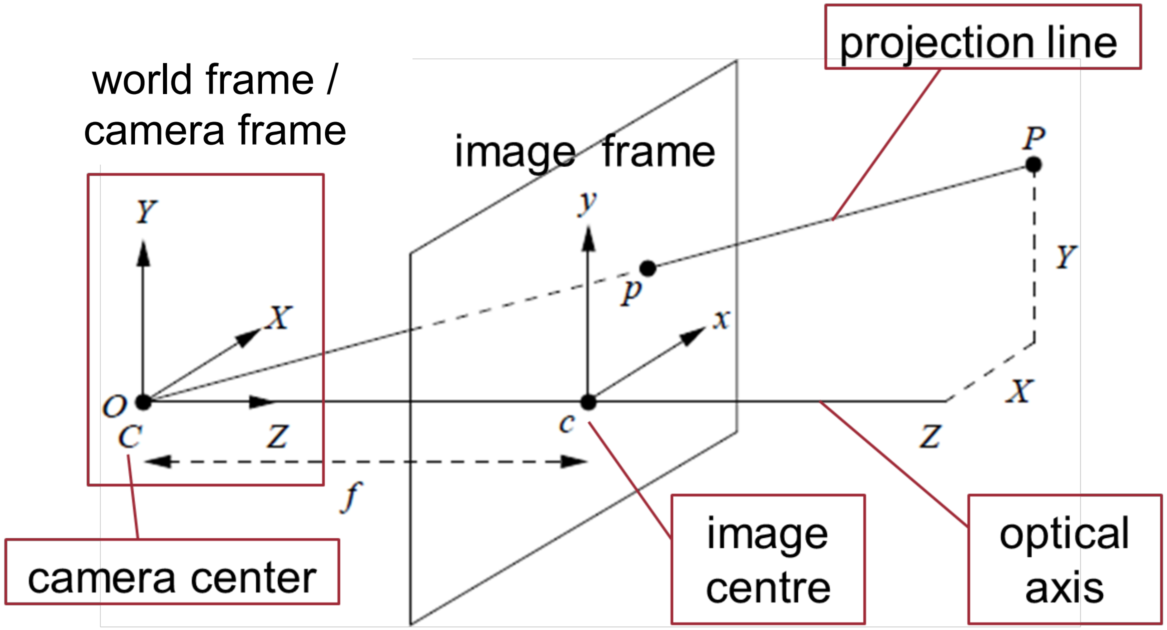
- x, y : coordinates
- k, l : scale parameters
- f : focal length

- 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

Perspective projection model



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.

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)

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

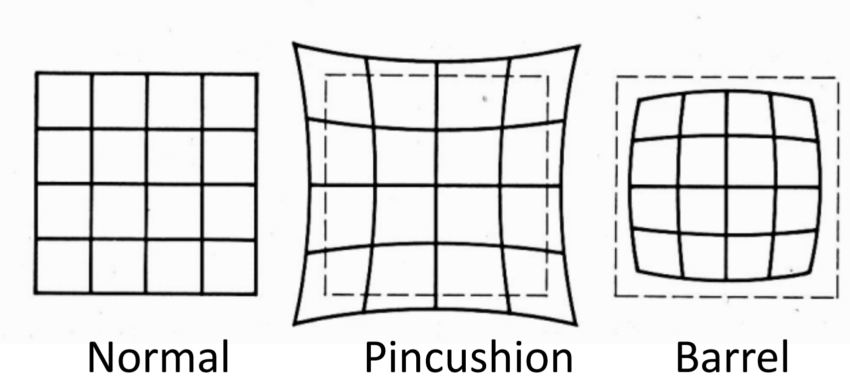
- Image center or principal point c may not be at origin
  - Denote location of c in image plane as c_x, c_y.

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

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



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

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

Intrinsic parameter matrix

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

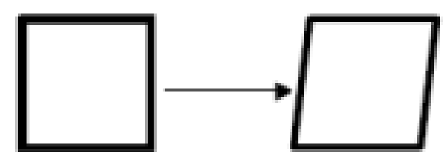
homogeneous coordinates

Intrinsic Parameters

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

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

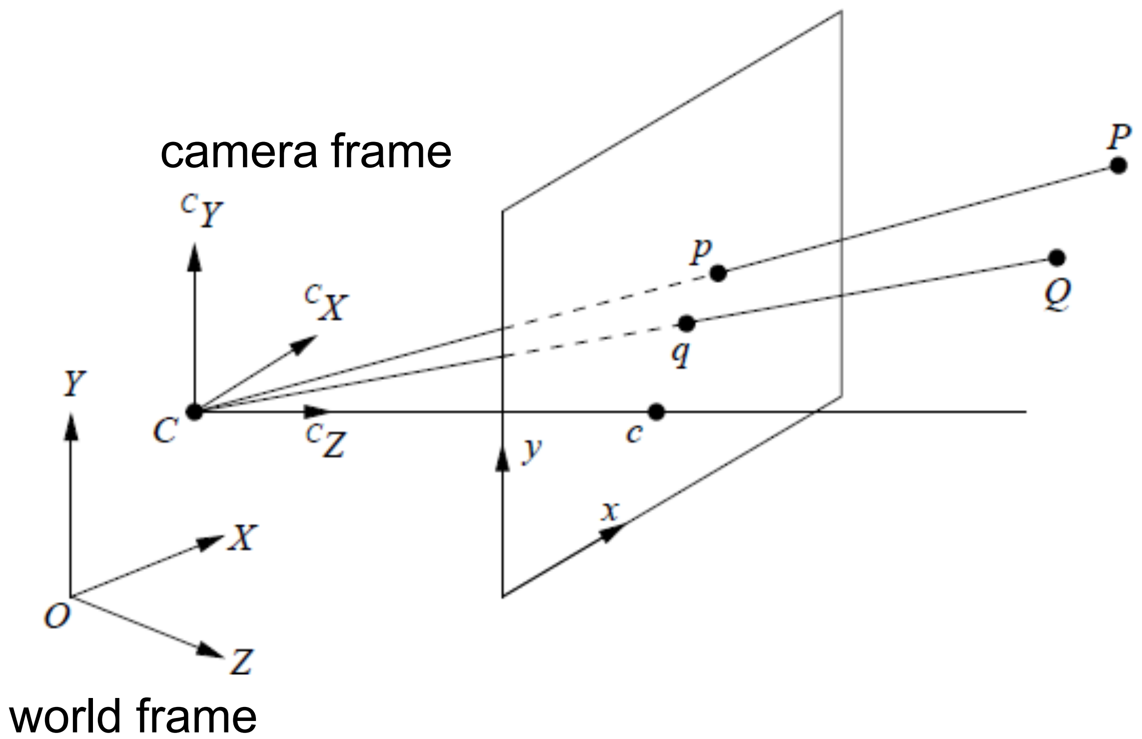
s: skew parameter



- 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

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

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

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

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

- Use a simpler notation

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

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**