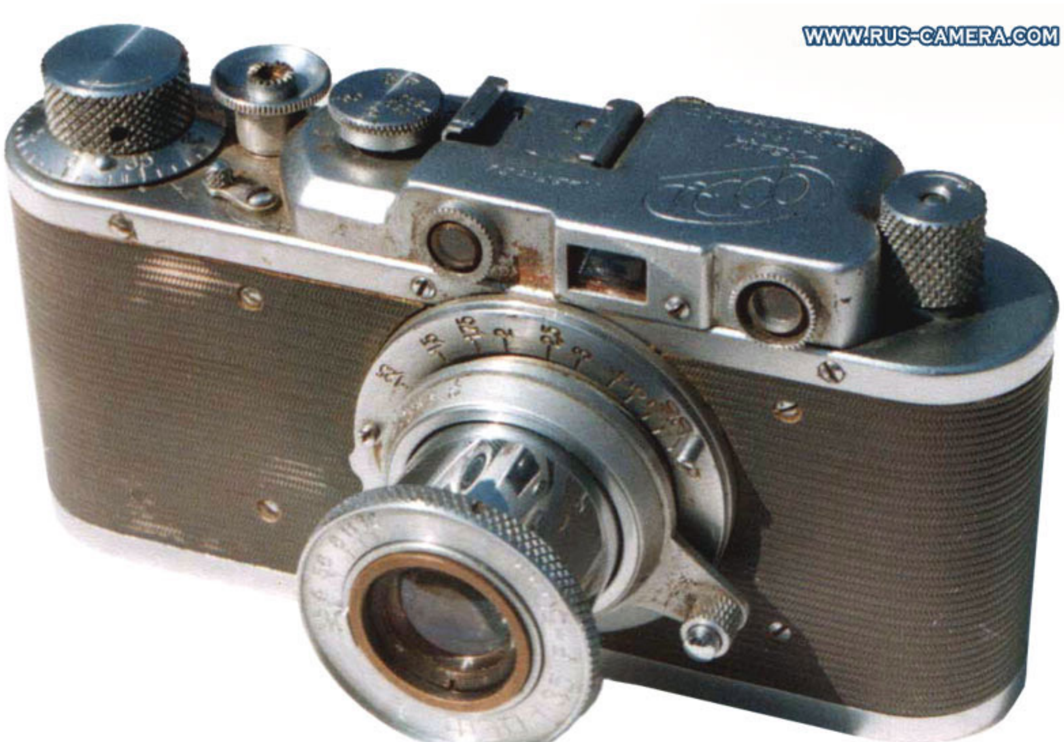
Cameras

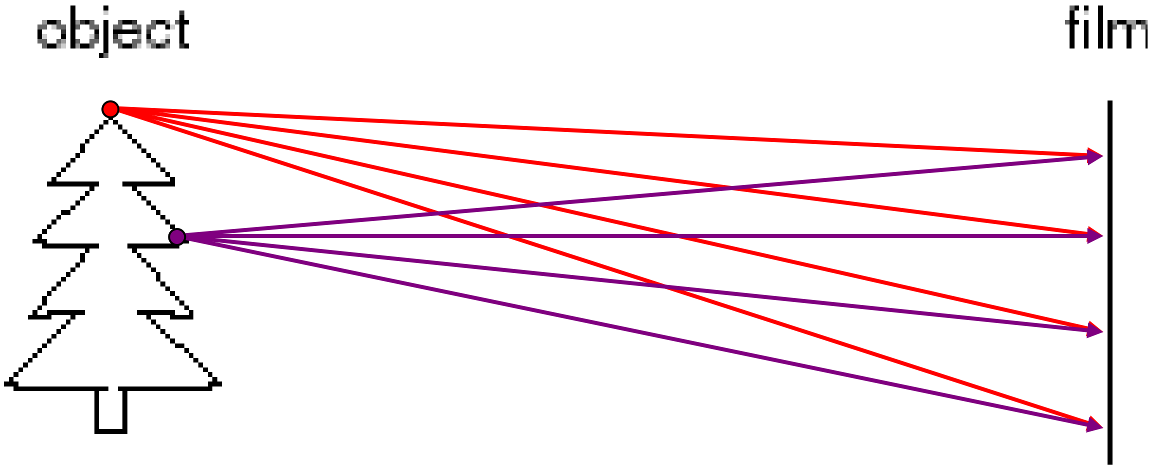


Source: S. Lazebnik

Reading

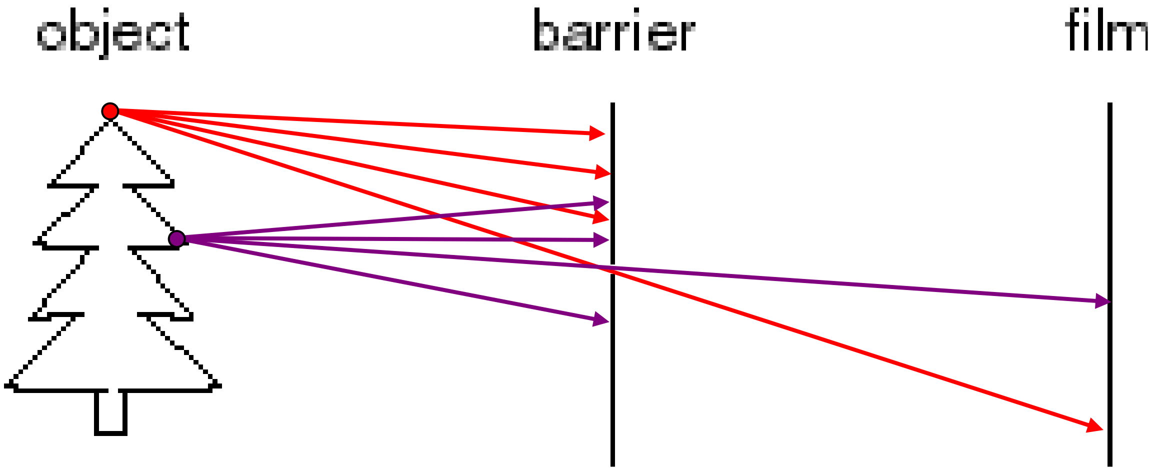
- Szeliski 2.1.3-2.1.6

Image formation



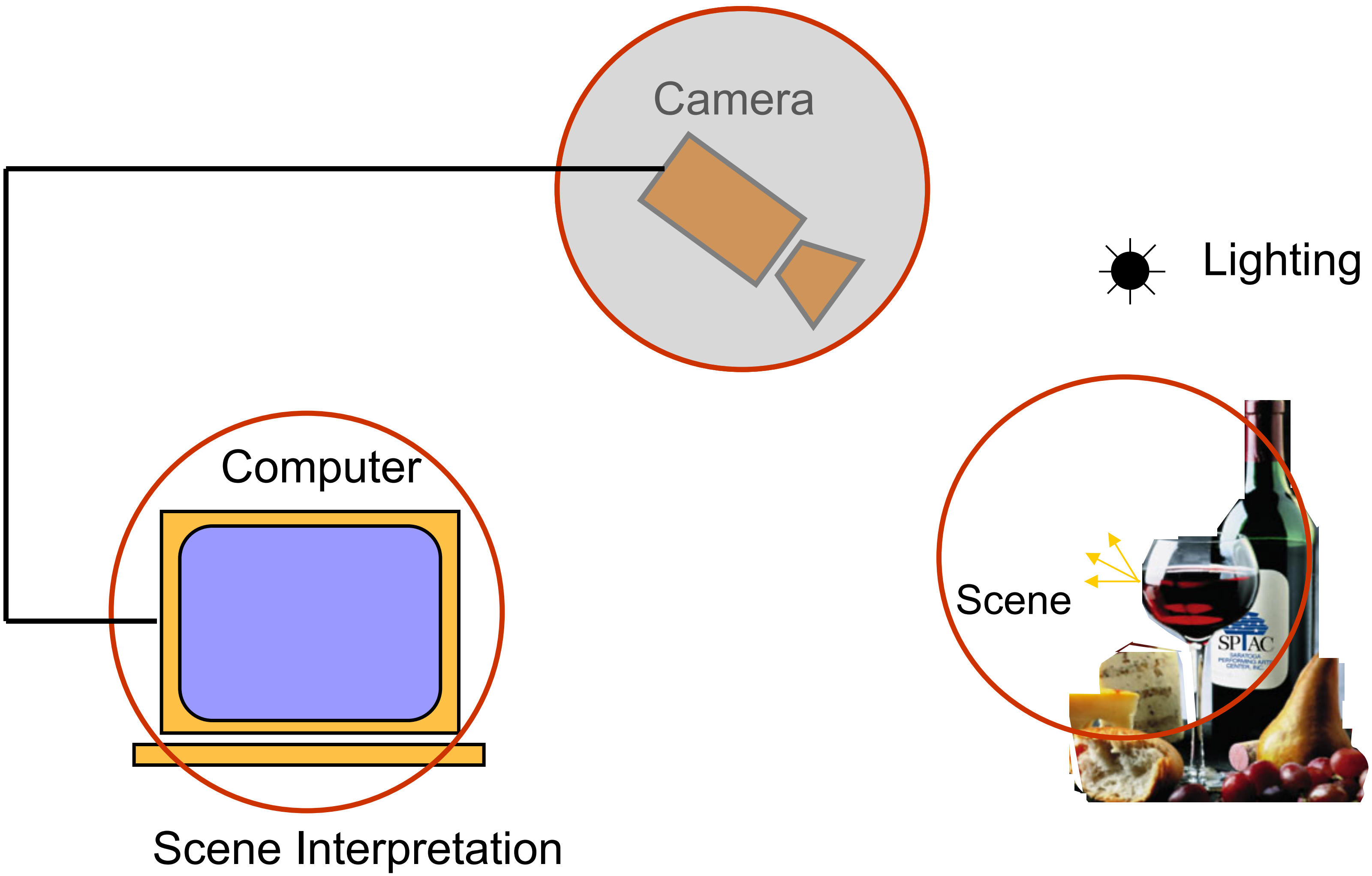
- Let’s design a camera
  - Idea 1: put a piece of film in front of an object
  - Do we get a reasonable image?

Pinhole camera

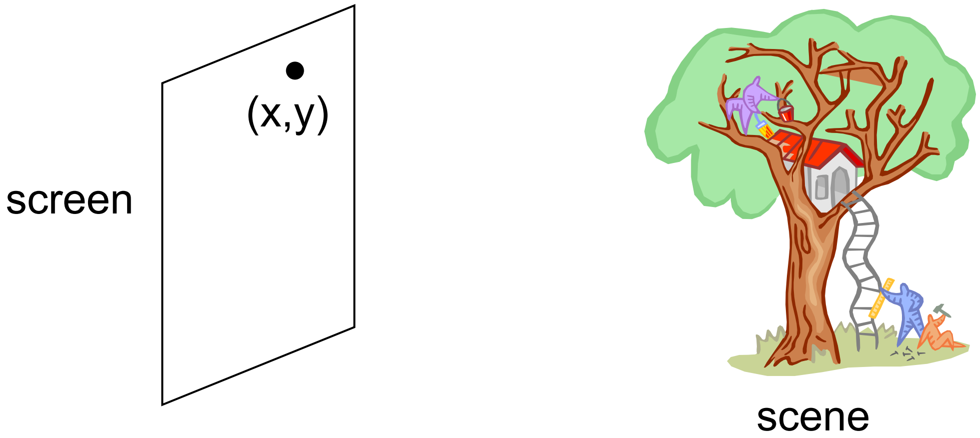


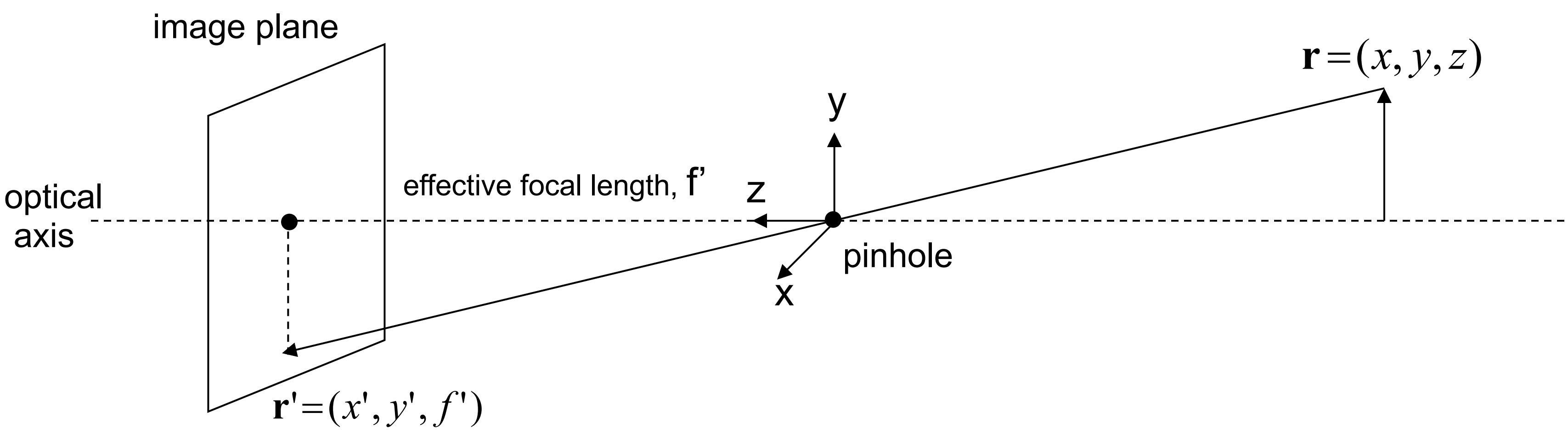
- Add a barrier to block off most of the rays
  - This reduces blurring
  - The opening known as the **aperture**
  - How does this transform the image?
  - The image is upside-down

Components of a Computer Vision System



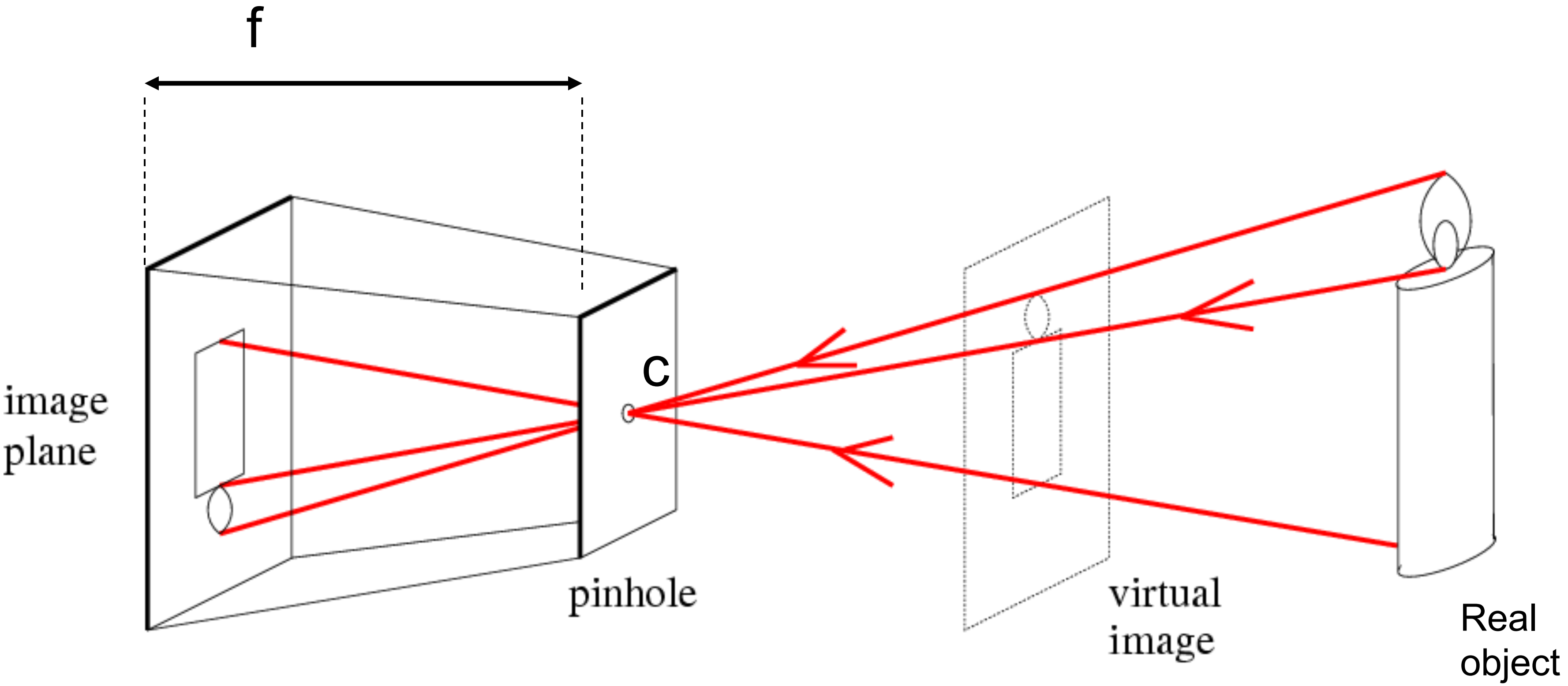
Pinhole and the Perspective Projection





$$\frac{\mathbf{r}'}{f'} = \frac{\mathbf{r}}{z} \quad \Rightarrow \quad \frac{x'}{f'} = \frac{x}{z} \qquad \frac{y'}{f'} = \frac{y}{z}$$

Pinhole camera model

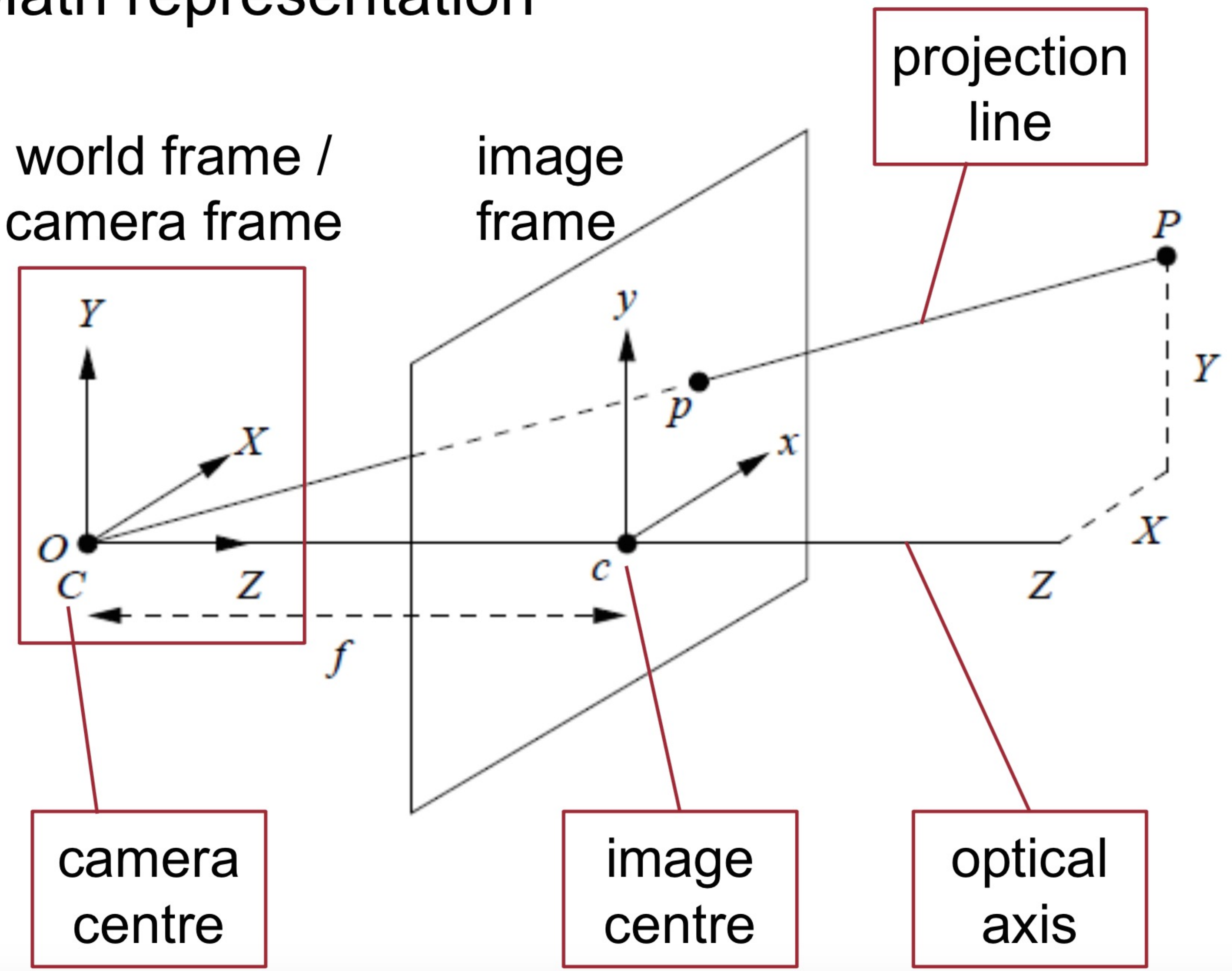


f = Focal length

c = Optical center of the camera

Figure from Forsyth

- Math representation



Pinhole camera model

- 3D point $P = (x, y, z)$ projects to 2D image point $P' = (x', y')$

- By symmetric

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

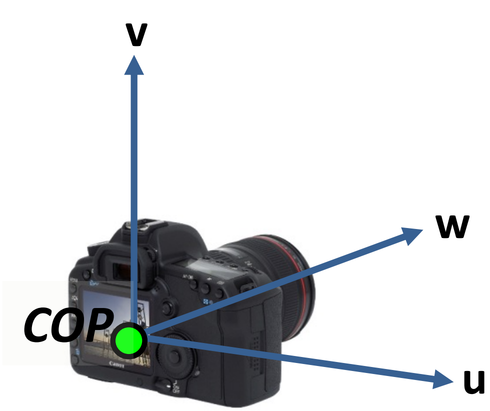
  i.e.,

$$P' = (x', y') = (f\frac{x}{z}, f\frac{y}{z})$$

- Simplest form of perspective projection

Camera parameters

- How can we model the geometry of a camera?



Camera

"The World"

Two important coordinate systems:

1. *World* coordinate system
2. *Camera* coordinate system

Camera parameters

- To project a point (x,y,z) in *world* coordinates into a camera
- First transform (x,y,z) into *camera* coordinates
- Need to know
  - Camera position (in world coordinates)
  - Camera orientation (in world coordinates)
- The formation of image frame
  - Need to know camera *intrinsics*

Intrinsic Parameters

- In the image frame, denote location of $c\ (principle\ point)$ image plane as c_x and c_y
- Image principle point:

Intersection between the camera optical axis and image plane

- Then

$$P' = (x', y') = (f\frac{x}{z} + c_x, f\frac{y}{z} + c_y)$$

Intrinsic Parameters

- Points in digital image are expressed as in pixels
- Points in image plane are represented in physical measurement (e.g., centimeter)
- The mapping between digital image and image plan can be something like $\frac{pixels}{cm}$
- We can use two parameters, *k* and *l*, to describe the mapping. If $k = l$, then the camera has "square pixels".
- The equation now becomes:

$$P' = (x', y') = (fk\frac{x}{z} + c_x, fl\frac{y}{z} + c_y)$$

$$= (\alpha\frac{x}{z} + c_x, \beta\frac{y}{z} + c_y)$$

Modeling projection

Homogeneous coordinates to the rescue!

$$(x, y) \Rightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

homogeneous image coordinates

$$(x, y, z) \Rightarrow \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

homogeneous scene coordinates

Converting *from* homogeneous coordinates

$$\begin{bmatrix} x \\ y \\ w \end{bmatrix} \Rightarrow (x/w, y/w)$$

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} \Rightarrow (x/w, y/w, z/w)$$

Intrinsic Parameters

$$\boxed{P' = (x', y') = (\alpha \frac{x}{z} + c_x, \beta \frac{y}{z} + c_y)}$$

In matrix form:

$$P' = \begin{bmatrix} \alpha & 0 & c_x & 0 \\ 0 & \beta & c_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = MP$$

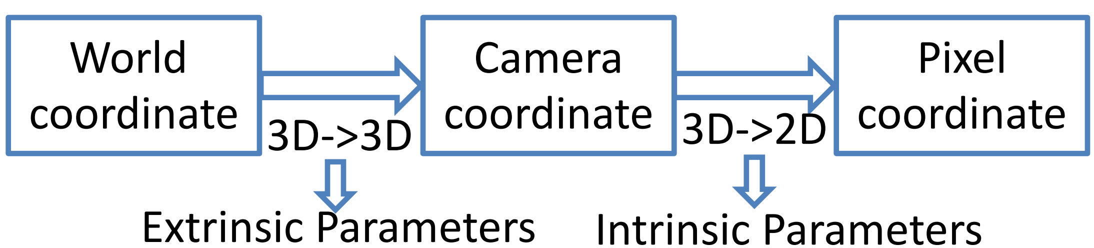
$$P' = MP = \begin{bmatrix} \alpha & 0 & c_x \\ 0 & \beta & c_y \\ 0 & 0 & 1 \end{bmatrix} [I \quad 0]P = K[I \quad 0]P$$

K: Camera matrix (or calibration matrix)

Extrinsic Parameters

- What if the information about the 3D world is available in a different coordinate system?
- We need to relate the points from world reference system to the camera reference system
- Given a point in world reference system P_w, the camera coordinate is computed as

$$P = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} P_w$$



Projection Matrix

- Combining intrinsic and extrinsic parameters, we have

$$P' = \underbrace{K}_{\text{intrinsic parameters}}\overbrace{[R \quad T]}^{\text{extrinsic parameters}}P_w = MP_w$$

- *K* changes as the type of camera changes
- Extrinsic parameters are independent of camera

Where does all this lead?

- Given an arbitrary camera, we may not have access to intrinsic parameters
- The problem of estimating intrinsic and extrinsic parameters is known as camera calibration

A pinhole camera has focal length 5mm. Each pixel is 0.02mm×0.02mm and the image principle point is at pixel (500,500). Pixel coordinate start at (0,0) in the upper-left corner of the image.

(a) Show the 3×3 camera matrix for this camera.
(b) Assume the world coordinate system is aligned with camera coordinate system (i.e., their origins are the same and their axes are aligned), and the origins are at the camera's pinhole, show the 3×4 projection matrix.
(c) What is the projection of a 3D scene point (100, 150, 800) into image coordinates?