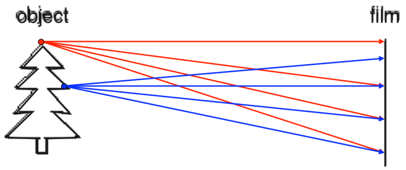
Geometric Transformations

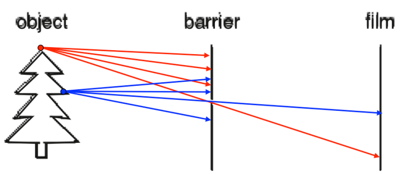
Let's design a camera



- Idea 1: put a piece of film in front of an object
- Do we get a reasonable image?

Slide by Steve Seitz

Pinhole camera



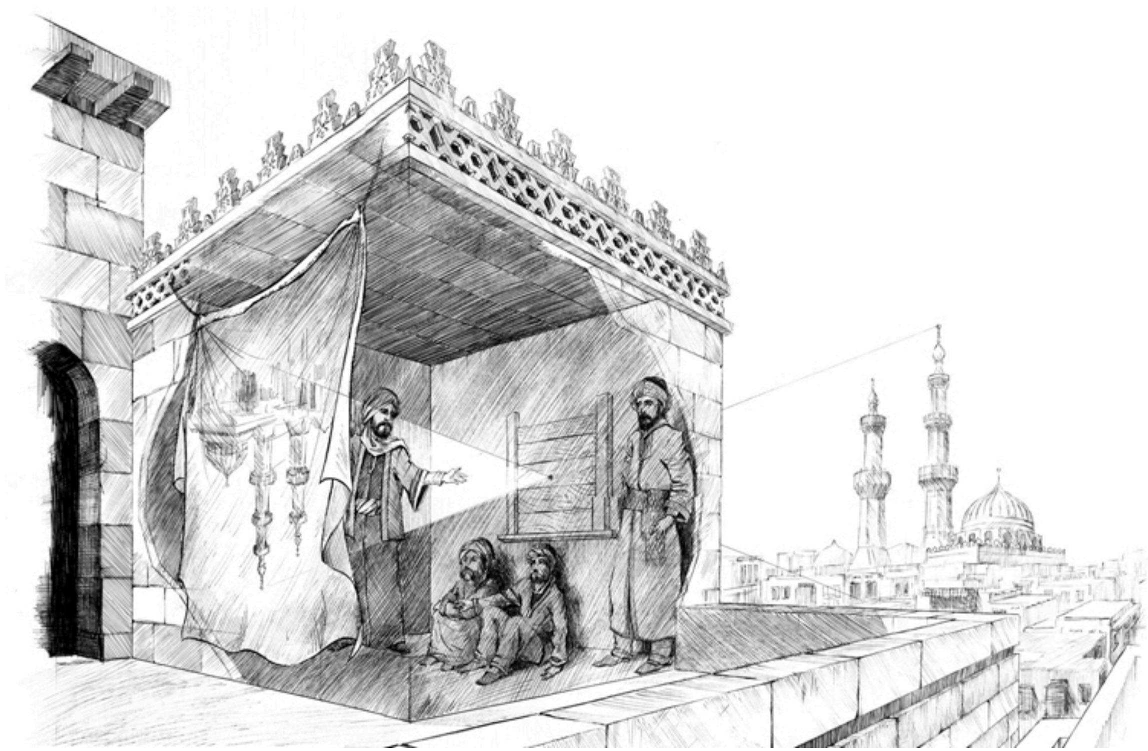
- Add a barrier to block off most of the rays
  - This reduces blurring
  - The opening is known as the **aperture**

Slide by Steve Seitz

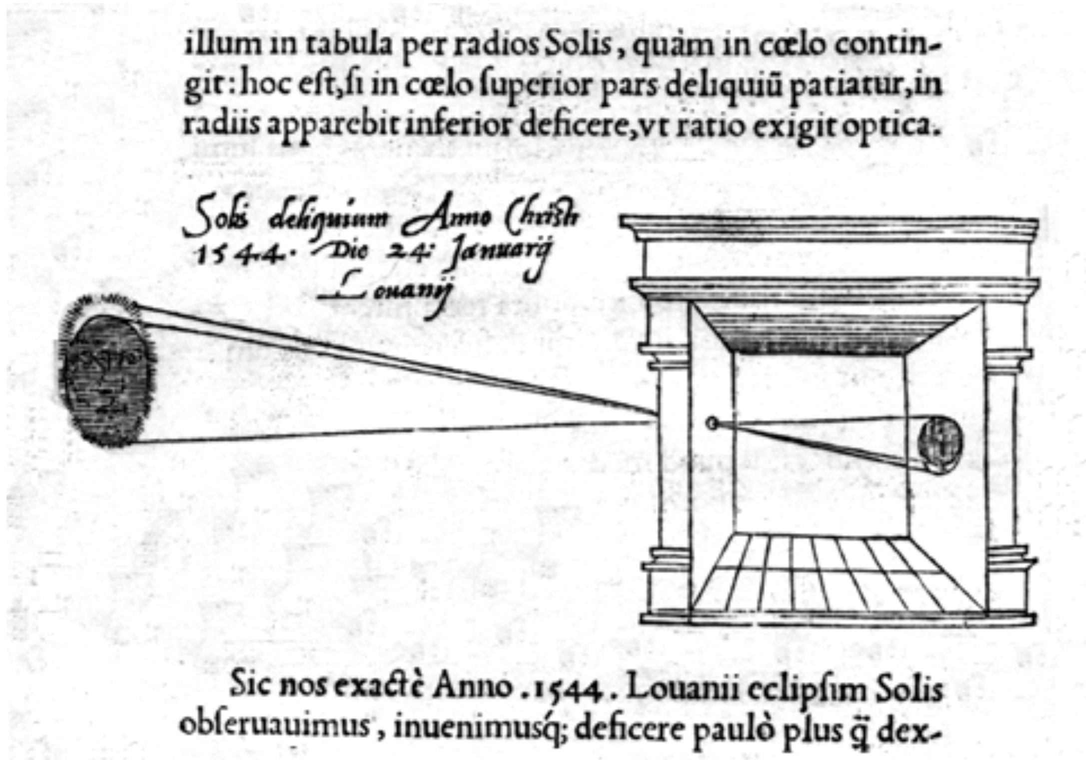
Camera Obscura

First Idea:Mo-Ti , China (470 BC to 390 BC)

First built: Ibn Al-Haytham or Alhazen, Iraq/Egypt (965 to 1039AD)



Camera Obscura



illum in tabula per radios Solis, quàm in cœlo contin-
git: hoc eſt, ſi in cœlo ſuperior pars deliquiũ patiatur, in
radiis apparebit inferior deficere, vt ratio exigit optica.

Solis deliquium Anno Christi
1544. Die 24. Januarij
Louanij

Sic nos exactè Anno .1544. Louanii eclipſim Solis
obſeruauimus, inuenimuſq; deficere paulò plus q̃ dex-

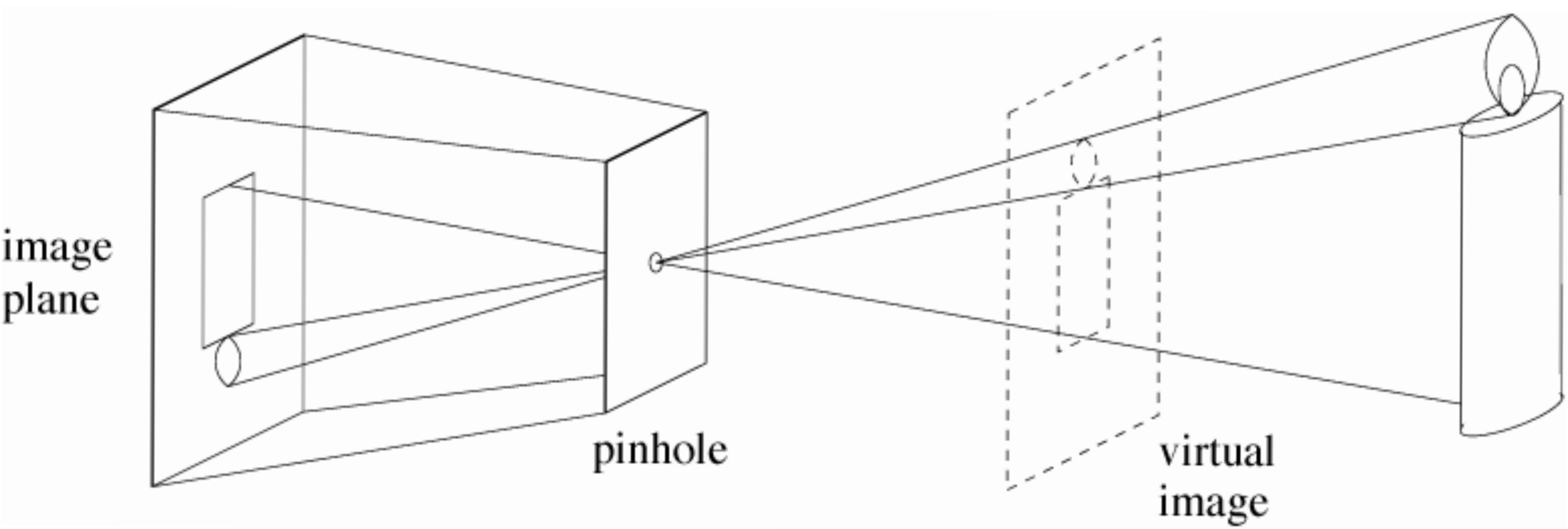
"When images of illuminated objects ... penetrate through a small hole into a very dark room ... you will see [on the opposite wall] these objects in their proper form and color, reduced in size ... in a reversed position, owing to the intersection of the rays".

Da Vinci

http://www.acmi.net.au/AIC/CAMERA_OBSCURA.html (Russell Naughton)

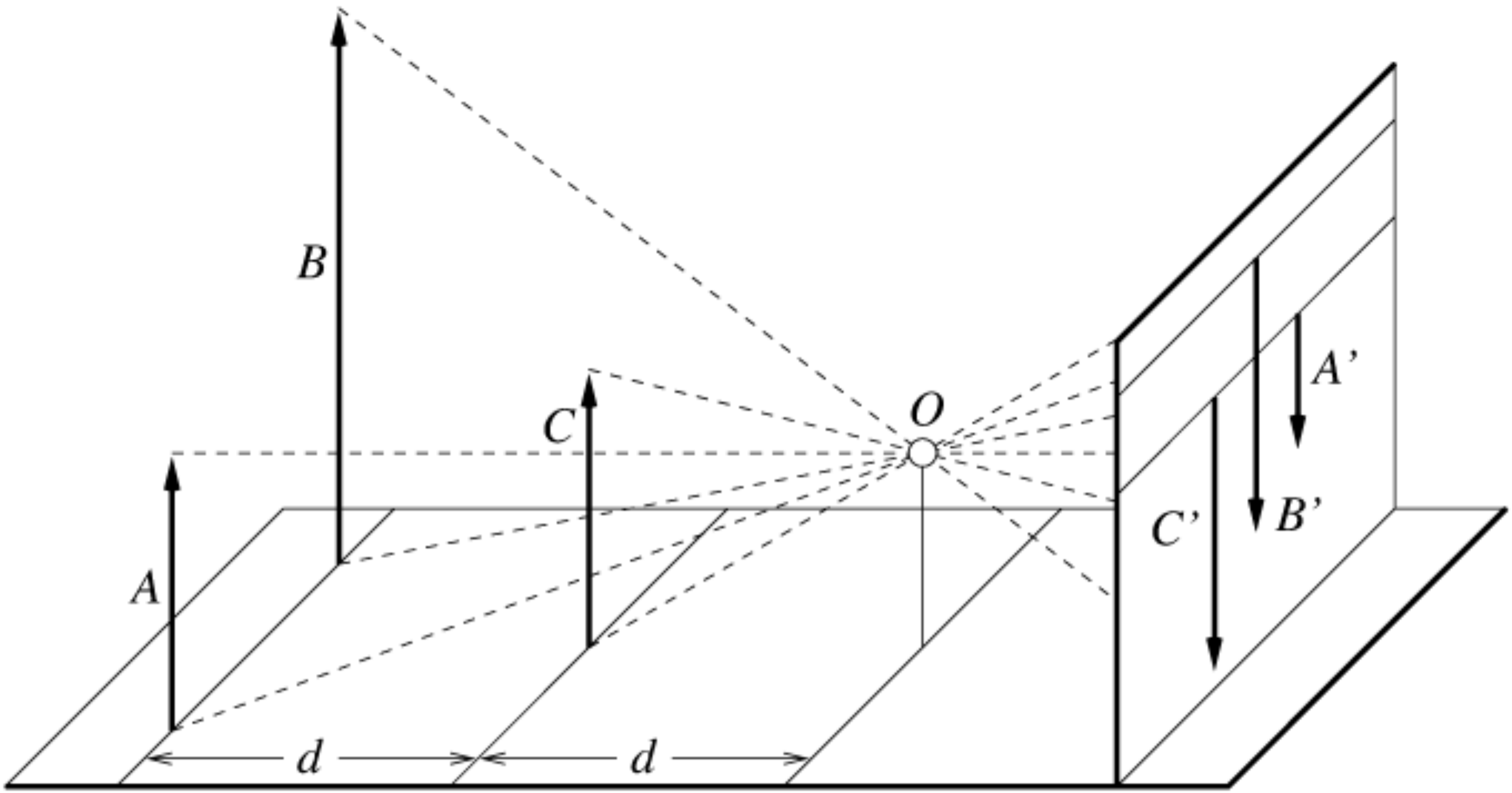
Pinhole cameras

- Abstract camera model - box with a small hole in it
- Pinhole cameras work in practice



(Forsyth & Ponce)

Distant objects are smaller

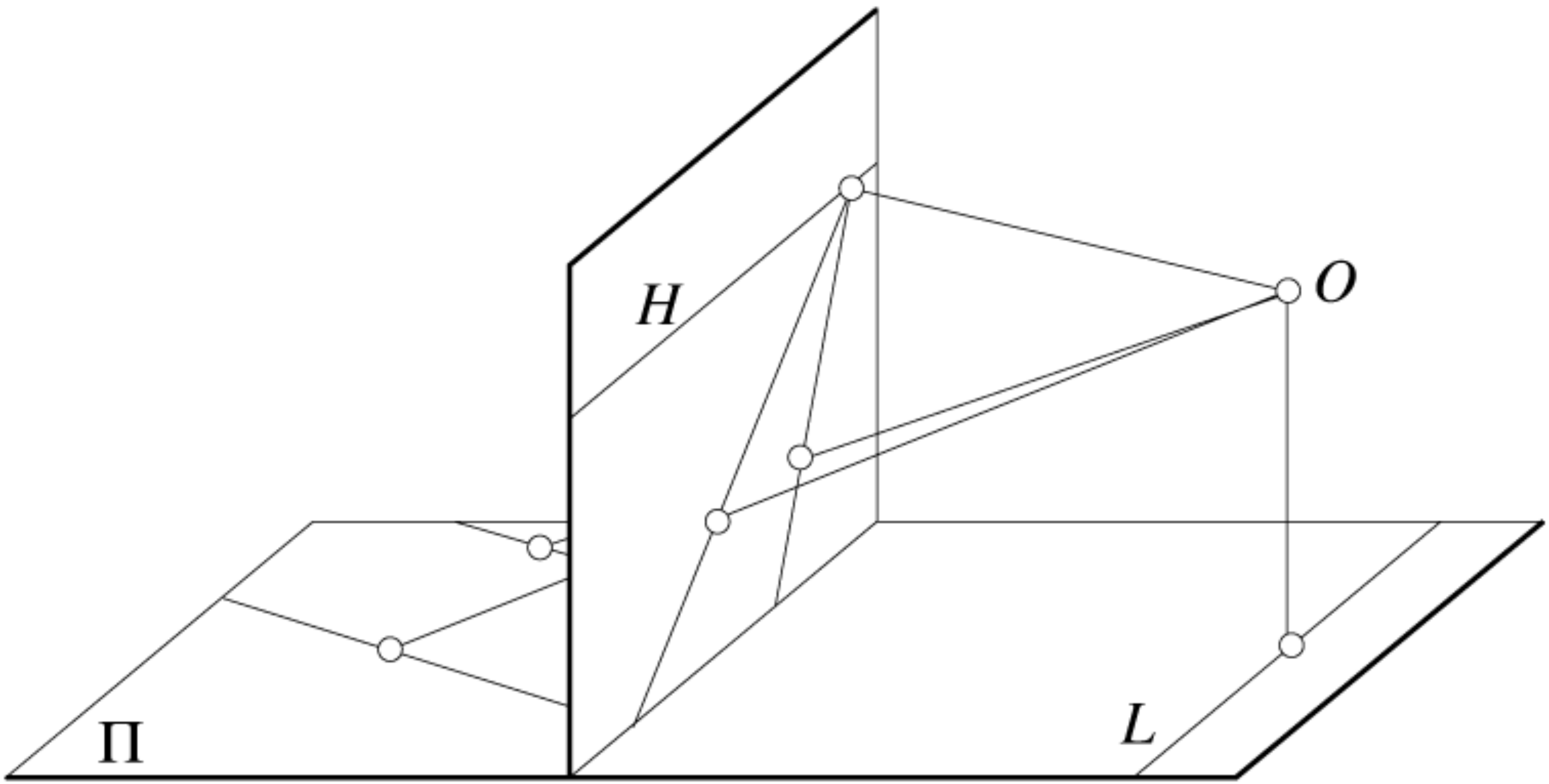


(Forsyth & Ponce)

Parallel lines meet

Parallel lines meet

Common to draw image plane *in front* of the focal point.
Moving the image plane merely scales the image.



(Forsyth & Ponce)

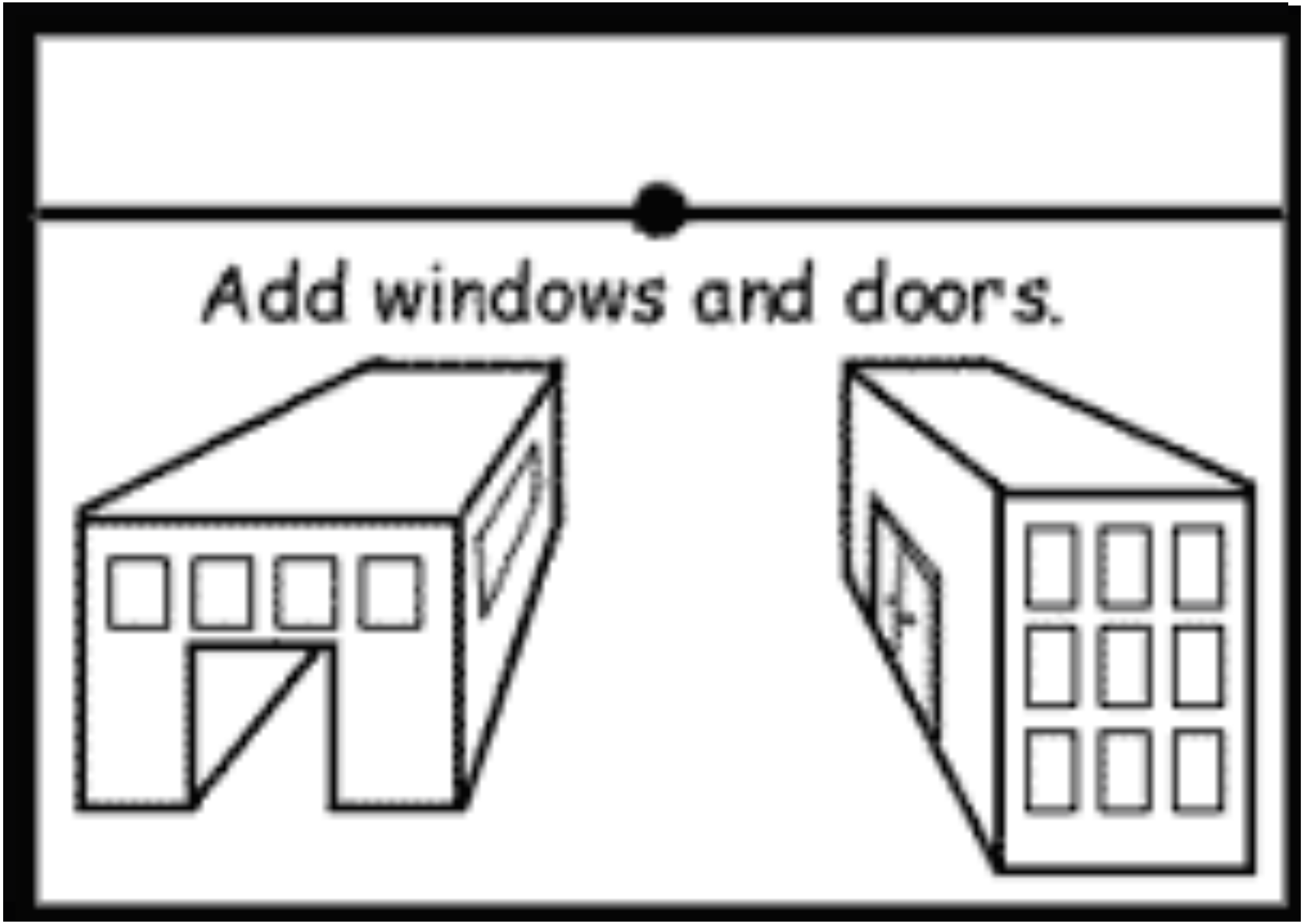
Vanishing points

- Each set of parallel lines meets at a different point
  - The *vanishing point* for this direction
- Sets of parallel lines on the same plane lead to *collinear* vanishing points.
  - The line is called the *horizon* for that plane

Properties of Projection

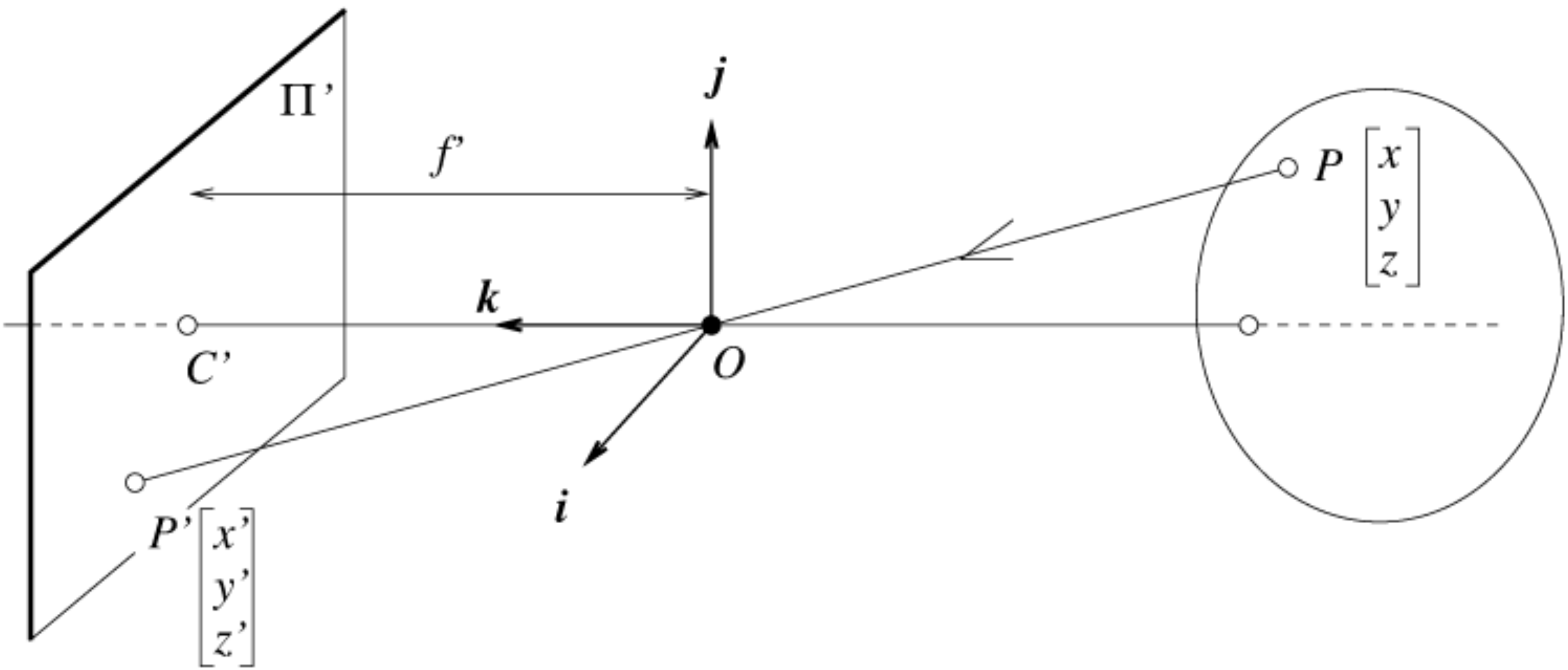
- Points project to points
- Lines project to lines
- Planes project to the whole image or a half image
- Angles are not preserved
- Degenerate cases
  - Line through focal point projects to a point.
  - Plane through focal point projects to line
  - Plane perpendicular to image plane projects to part of the image (with horizon).

Take out paper and pencil



http://www.sanford-artedventures.com/create/tech_1pt_perspective.html

The equation of projection



(Forsyth & Ponce)

The equation of projection

- Cartesian coordinates:
  - We have, by similar triangles, that

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

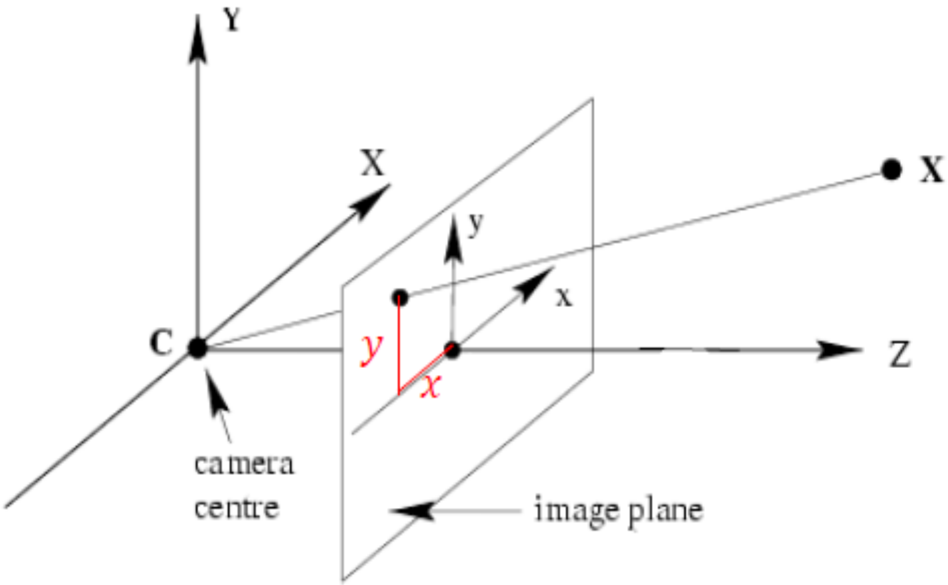
$$y' = f'\frac{y}{z}$$

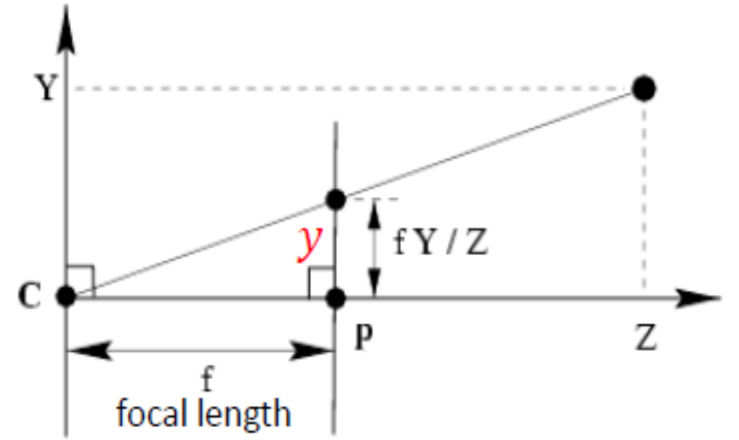
$$(x, y, z) \rightarrow (f'\frac{x}{z}, f'\frac{y}{z}, f')$$

  - Ignore the third coordinate, and get

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

Pinhole camera model – in maths

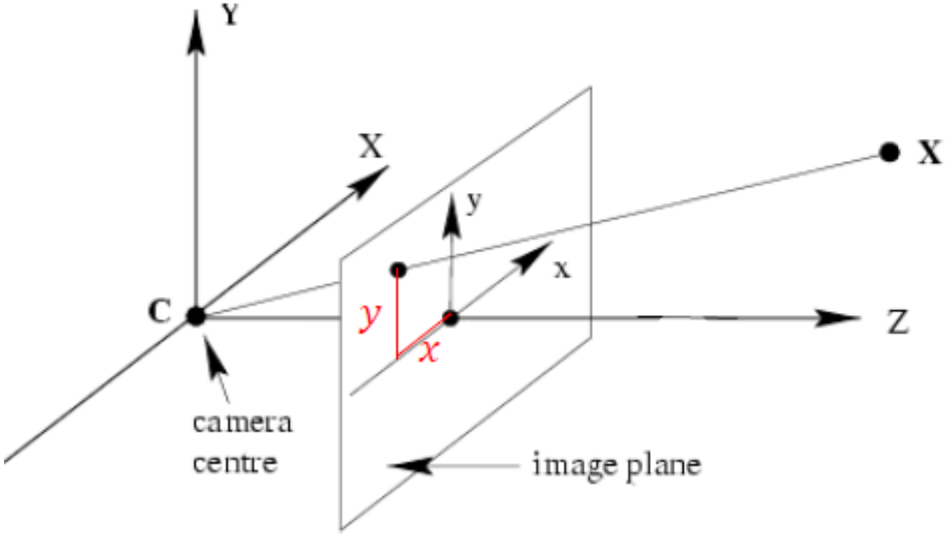


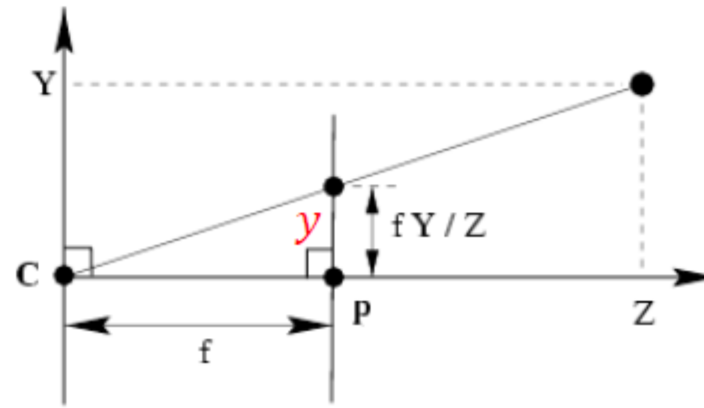


- Similar trinagles: $\frac{y}{f} = \frac{Y}{Z}$
- That gives: $y = f\frac{Y}{Z}$ and $x = f\frac{X}{Z}$

That gives: $\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$

Pinhole camera model – in maths





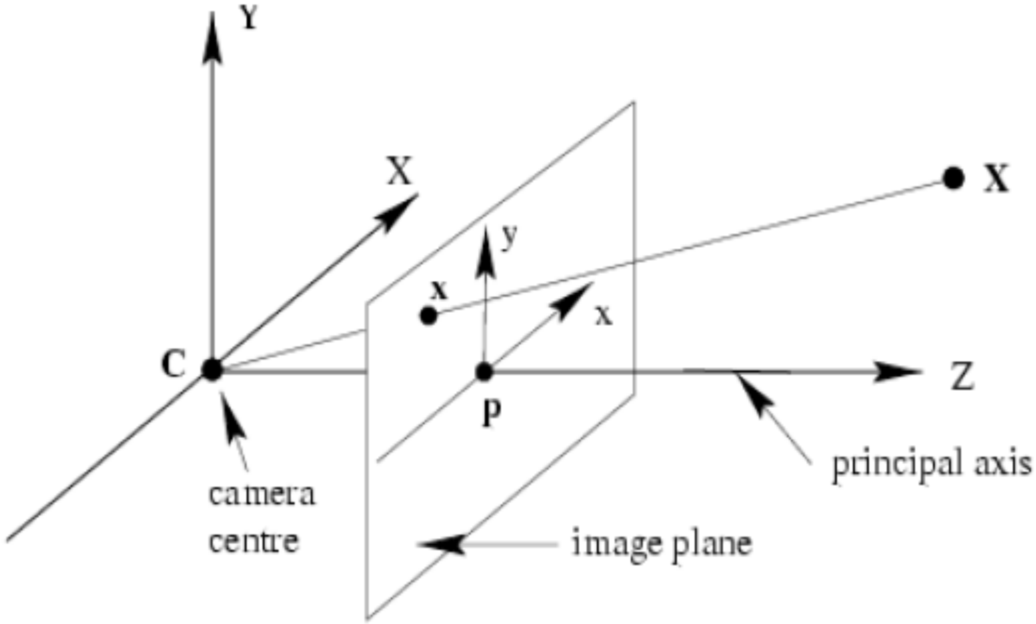
That gives: $$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\text{Calibration matrix } \boldsymbol{K}} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

In short $\boldsymbol{x} = \boldsymbol{K}\,\tilde{\mathbf{X}}$ (here $\tilde{\mathbf{X}}$ means inhomogeneous coordinates)

Intrinsic Camera Calibration means we know $\boldsymbol{K}$ (we do that later)

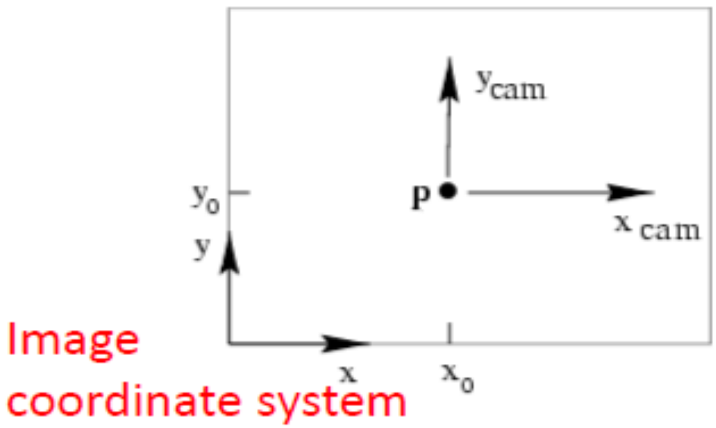
We can go from image points into the 3D world: $\tilde{\boldsymbol{X}} = \boldsymbol{K}^{-1}\,\boldsymbol{x}$

Pinhole camera - definitions



- Principal axis: line from the camera center perpendicular to the image plane
- Normalized (camera) coordinate system: camera center is at the origin and the principal axis is the z-axis
- Principal point (p): point where principal axis intersects the image plane (origin of normalized coordinate system)

Principal Point



Principal point (p_x, p_y)

- Camera coordinate system: origin is at the principal point
- Image coordinate system: origin is in the corner
  In practice: principal point in center of the image

Adding principal point into $\boldsymbol{K}$

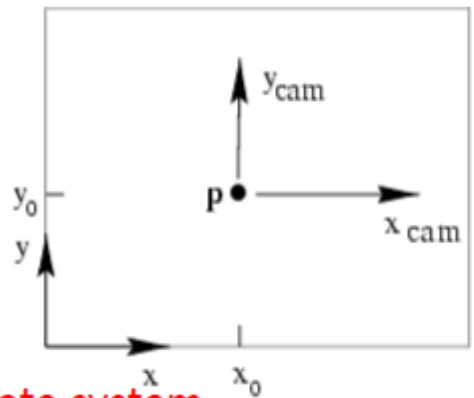


Image coordinate system

Principal point (p_x, p_y)

Projection with principal point : $y = f\frac{Y}{Z} + p_y = \frac{fY + Zp_y}{Z}$ and $x = f\frac{X}{Z} + p_x = \frac{fX + Zp_x}{Z}$

That gives: $\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$

Adding principal point into $\boldsymbol{K}$

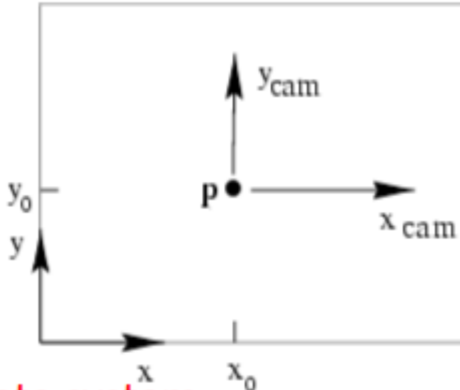


Image coordinate system

Principal point (p_x, p_y)

Projection with principal point : $y = f\frac{Y}{Z} + p_y = \frac{fY+Zp_y}{Z}$ and $x = f\frac{X}{Z} + p_x = \frac{fX+Zp_x}{Z}$

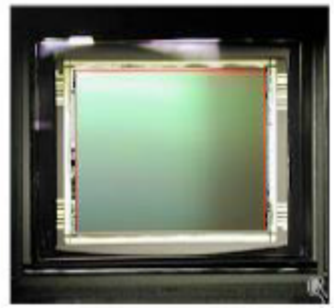
That gives: $\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$

Intrinsic matrix, K

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

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

Pixel Size and Shape



- m_x pixels per unit (m,mm,inch,...) in horizontal direction
- m_y pixels per unit (m,mm,inch,...) in vertical direction
- s' skew of a pixel
- *In practice (close to): m=1 s = 0*

That gives:
$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} m_x & s' & 0 \\ 0 & m_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

Simplified to:
$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} f & s & p_x \\ 0 & mf & p_y \\ 0 & 0 & 1 \end{bmatrix}}_{\text{Final calibration matrix } \boldsymbol{K}} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

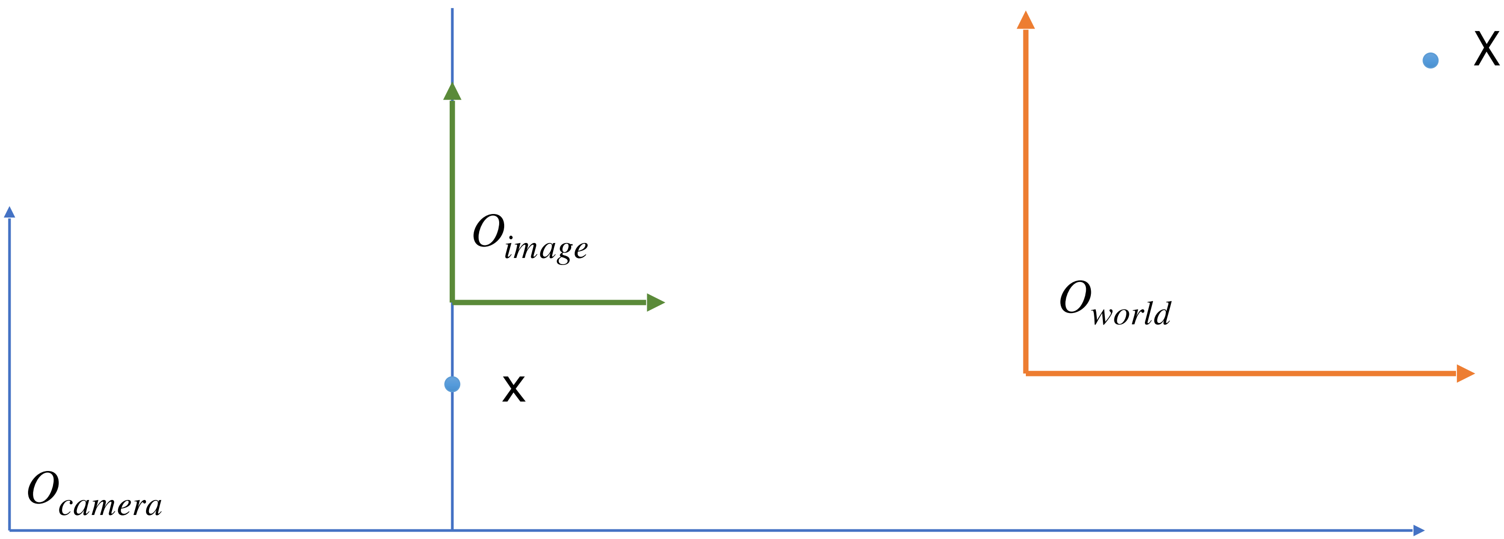
$\boxed{f \text{ now in units of pixels}}$

Camera intrinsic parameters - Summary

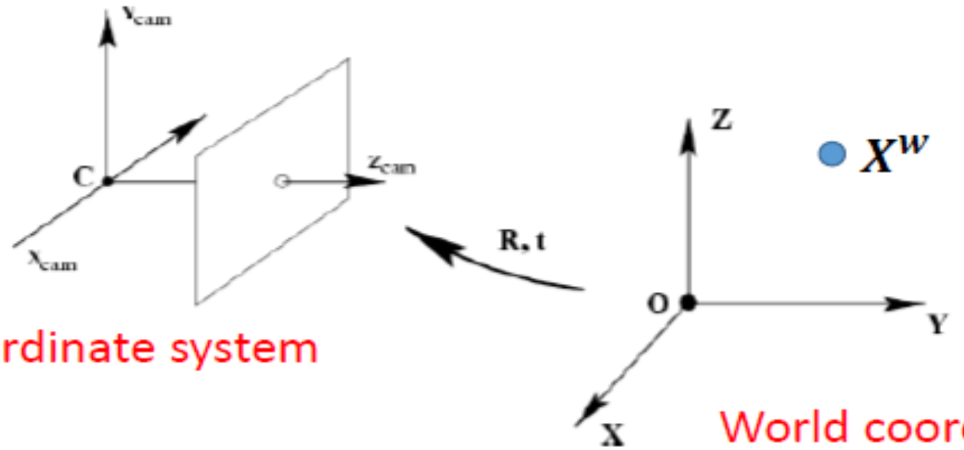
- <u>Intrinsic parameters</u>
  - Principal point coordinates (p_x, p_y)
  - Focal length f
  - Pixel magnification factors m
  - Skew (non-rectangular pixels) s

$$\boldsymbol{K} = \begin{bmatrix} f & s & p_x \\ 0 & mf & p_y \\ 0 & 0 & 1 \end{bmatrix}$$

Three different coordinate systems



Putting the camera into the world



camera coordinate system

World coordinate system

Given a 3D homogenous point X^w in world coordinate system

1) Translate from world to camera coordinate system:

$$\tilde{X}^{c'} = \tilde{X}^w - \tilde{C}$$

$$\tilde{X}^{c'} = \underbrace{(I_{3\times 3} \mid -\tilde{C})}_{3\times 4 \text{ matrix}} X^w$$ where $I_{3\times 3}$ is 3x3 identity matrix

2) Rotate world coordinate system into camera coordinate system

$$\tilde{X}^c = R\,(I_{3\times 3} \mid -\tilde{C})\, X^w$$

3) Apply camera matrix

$$x = K\,R\,(I_{3\times 3} \mid -\tilde{C})\, X$$

Camera extrinsic (or external) parameters

- Transform a point from the world coordinate to the camera's coordinate system
- Translation and rotation

$$X_c = R(X_w - C_w)$$

$$X_c = RX_w - RC_w$$

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \begin{bmatrix} R & -RC_w \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} R & -RC_w \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} R & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} I & -C_w \\ 0 & 1 \end{bmatrix}$$

Camera extrinsic (or external) parameters

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \begin{bmatrix} R & -RC_w \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} R & -RC_w \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} R & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} I & -C_w \\ 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} R & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & 0 \\ r_{21} & r_{22} & r_{23} & 0 \\ r_{31} & r_{32} & r_{33} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} I & -C_w \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & -c_x \\ 0 & 1 & 0 & -C_y \\ 0 & 0 & 1 & -C_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Camera extrinsic (or external) parameters

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & 0 \\ r_{21} & r_{22} & r_{23} & 0 \\ r_{31} & r_{32} & r_{33} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & -C_x \\ 0 & 1 & 0 & -C_y \\ 0 & 0 & 1 & -C_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = R(I_{3\times 3} \mid -C_w) \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = K \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} \qquad K = \begin{bmatrix} f & s & p_x \\ 0 & mf & p_y \\ 0 & 0 & 1 \end{bmatrix}$$

Summary

$$\begin{bmatrix} X_w \\ Y_w \\ z_w \end{bmatrix} \longrightarrow \begin{bmatrix} X_c \\ Y_c \\ z_c \end{bmatrix} \longrightarrow \begin{bmatrix} x \\ y \end{bmatrix}$$

world coordinates

camera coordinates

pixel coordinates

Camera matrix

- Camera matrix $\boldsymbol{P}$ is defined as:

$$\boldsymbol{x} = \underbrace{\boldsymbol{K}\,\boldsymbol{R}\,(\boldsymbol{I}_{3\times3} \mid -\tilde{\boldsymbol{C}})}_{\boldsymbol{P}\,(3\times4)\text{ camera matrix has 11 DoF}}\,\boldsymbol{X}$$

- In short we write: $\boldsymbol{x} = \boldsymbol{P}\,\boldsymbol{X}$

Camera matrix

- Camera matrix $\boldsymbol{P}$ is defined as:

$$\boldsymbol{x} = \underbrace{\boldsymbol{K}\,\boldsymbol{R}\,(\boldsymbol{I}_{3\times 3} \mid -\tilde{\boldsymbol{C}})}_{\boldsymbol{P}\ (3\times 4)\ \text{camera matrix has 11 DoF}}\,\boldsymbol{X}$$

- In short we write: $\boldsymbol{x} = \boldsymbol{P}\,\boldsymbol{X}$

Image of a Point

Homogeneous coordinates of a 3-D point p

$$\boldsymbol{X} = [X, Y, Z, W]^T \in \mathbb{R}^4, \quad (W = 1)$$

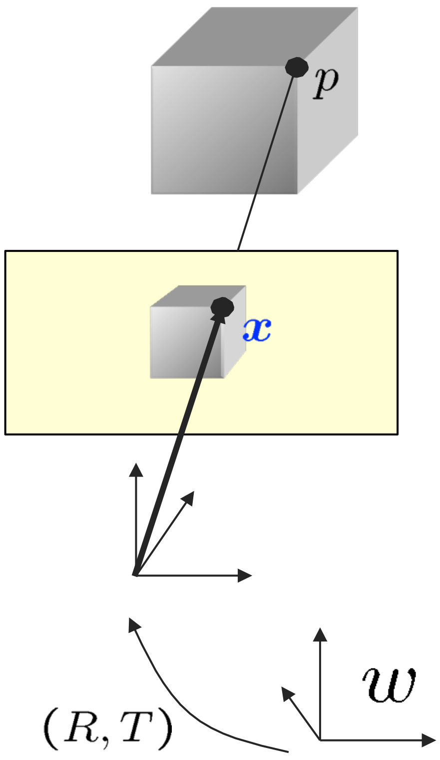
Homogeneous coordinates of its 2-D image

$$\boldsymbol{x} = [x, y, z]^T \in \mathbb{R}^3, \quad (z = 1)$$

Projection of a 3-D point to an image plane

$$\lambda x = PX$$

$$\lambda \in \mathbb{R}, \ \mathrm{P} = [R, T] \in \mathbb{R}^{3\times 4}$$



Camera parameters - Summary

- Camera matrix P has 11 DoF

$$\boldsymbol{x} = \boldsymbol{P}\,\boldsymbol{X}$$

$$\boldsymbol{x} = \boldsymbol{K}\,\boldsymbol{R}\,(\boldsymbol{I}_{3\times3}\,|\,-\tilde{\boldsymbol{C}})\,\boldsymbol{X}$$

$$\boldsymbol{K} = \begin{bmatrix} f & s & p_x \\ 0 & mf & p_y \\ 0 & 0 & 1 \end{bmatrix}$$

- <u>Intrinsic parameters</u>
  - Principal point coordinates (p_x, p_y)
  - Focal length f
  - Pixel magnification factors m
  - Skew (non-rectangular pixels) s

- <u>Extrinsic parameters</u>
  - Rotation $\boldsymbol{R}$ (3DoF) and translation $\tilde{\mathbf{C}}$ (3DoF) relative to world coordinate system