Chapter 2

Projective Geometry and Camera Models

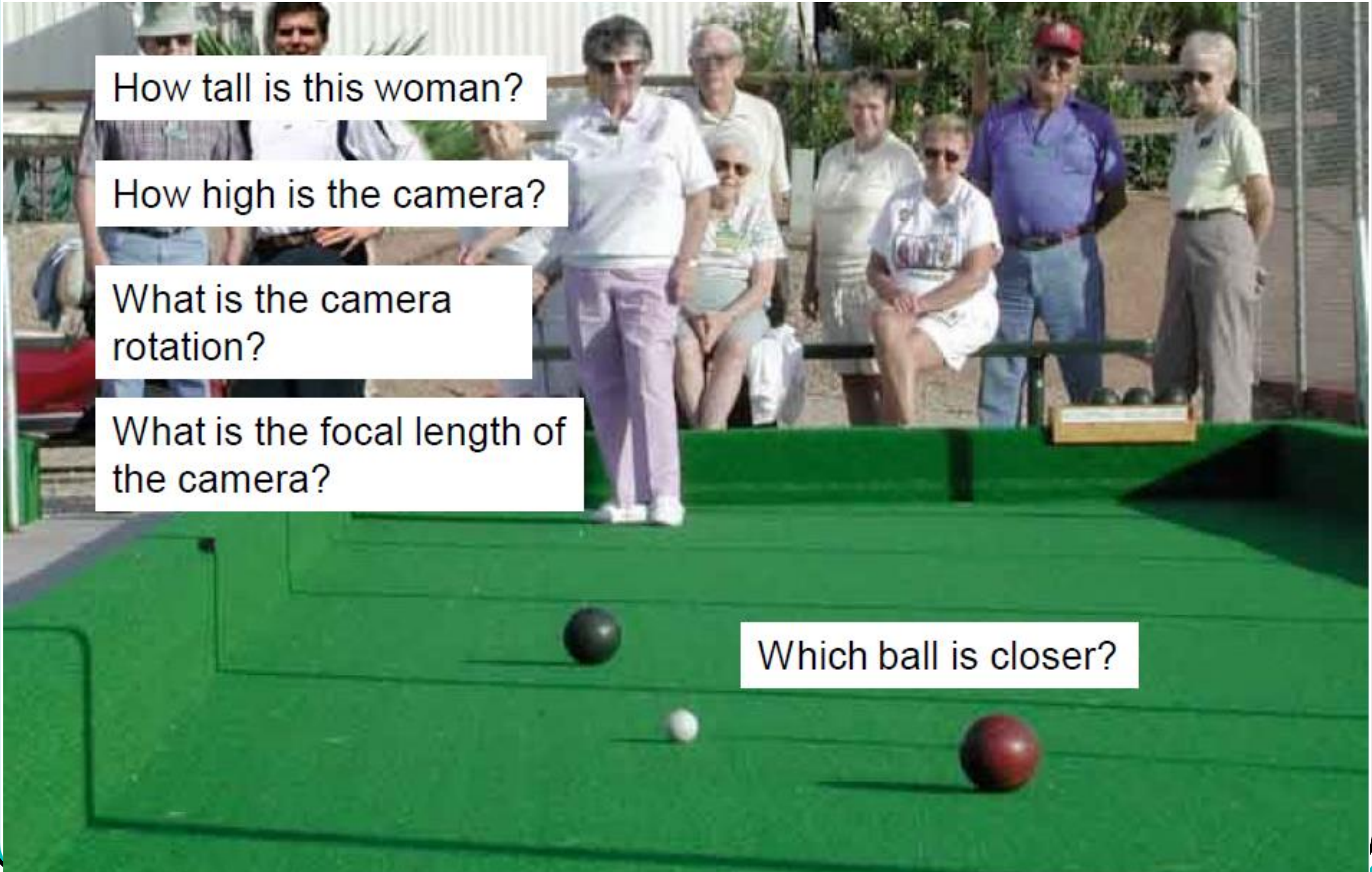
James Hays, Brown University

Contents

Mapping between image and world coordinates

- Projective geometry
  - Vanishing points and lines
- Pinhole camera model
- Cameras & lenses
- Projection matrix

Camera and World Geometry

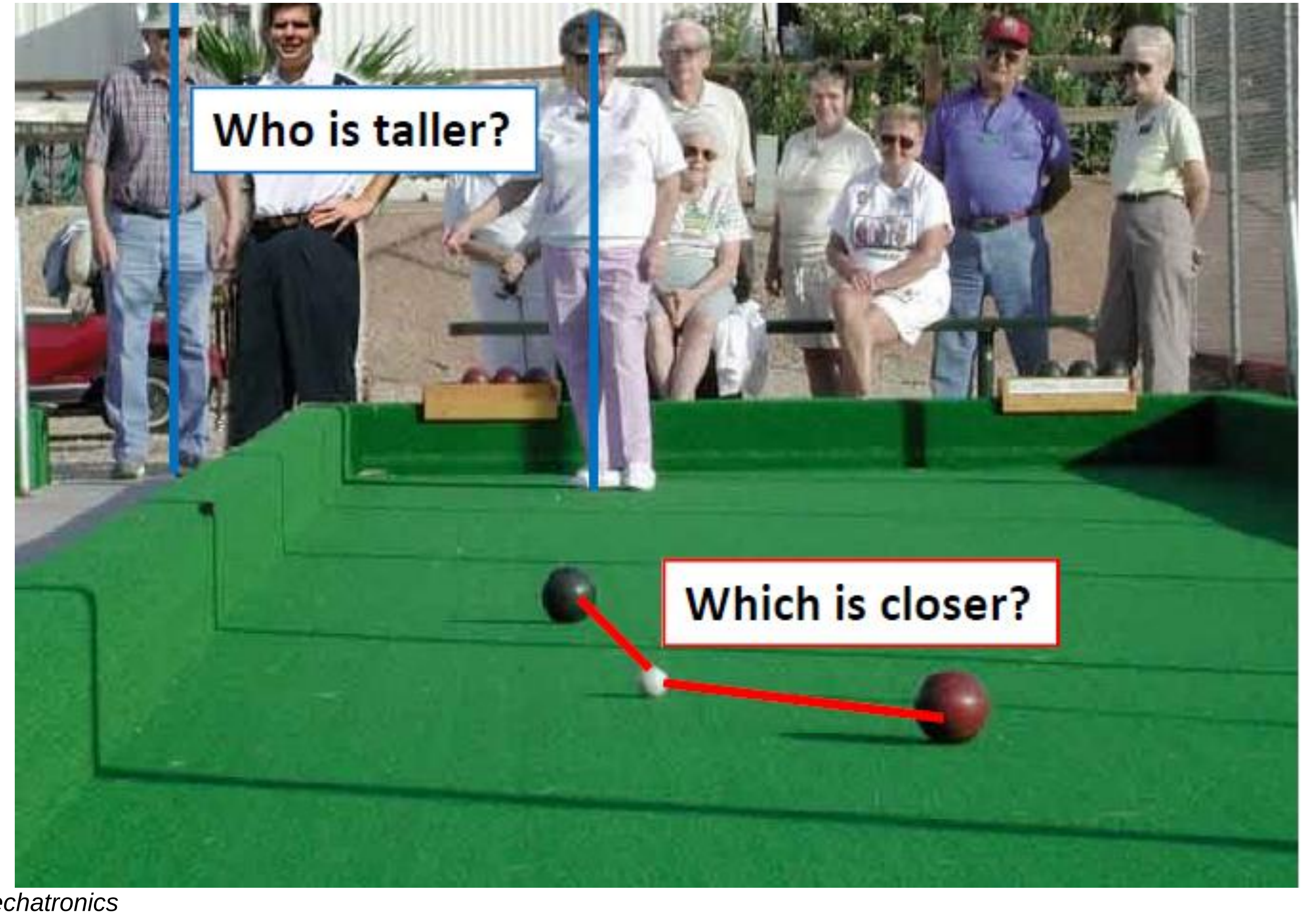


Projection can be tricky…

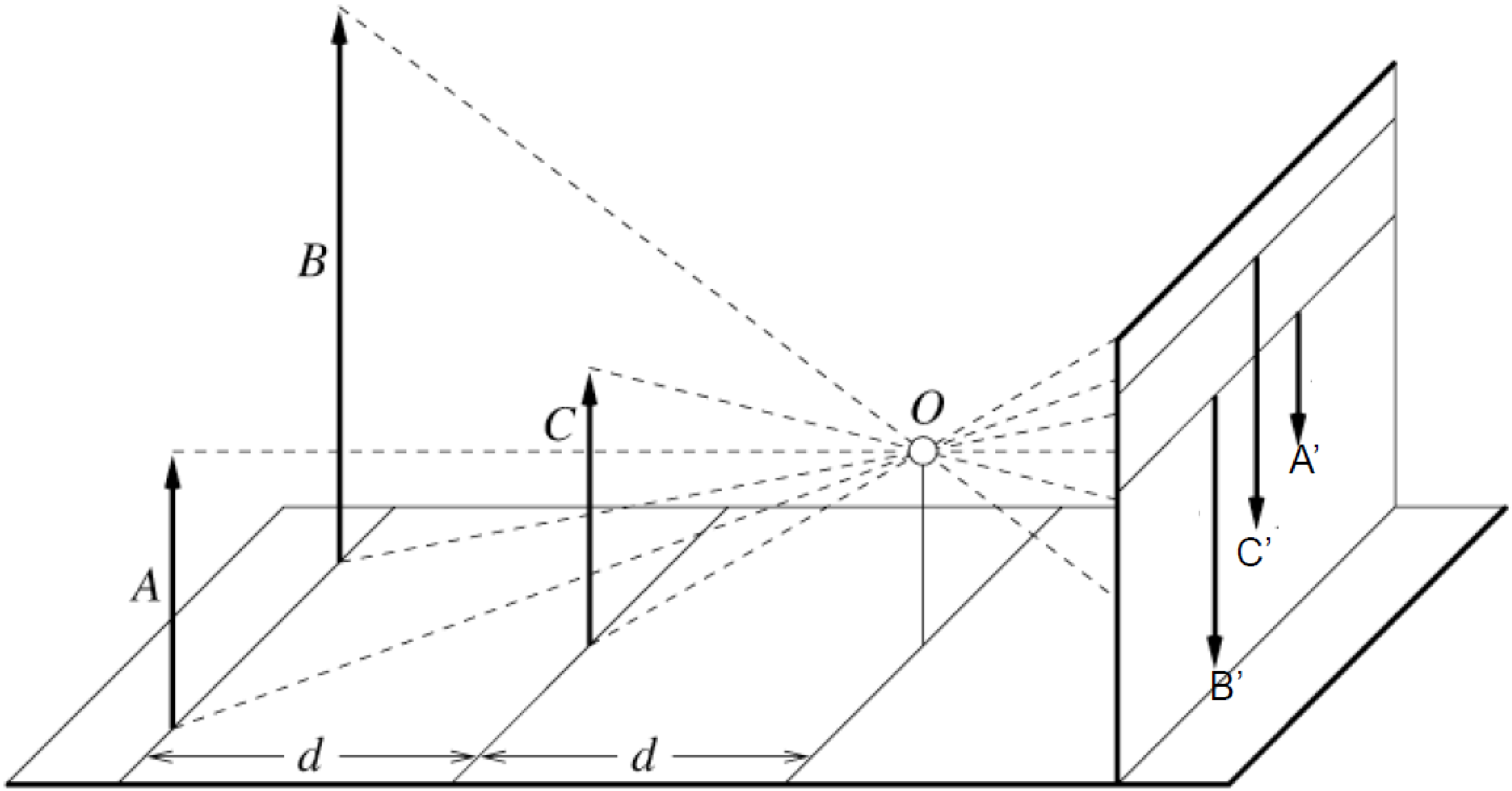
Projection can be tricky…

Projective Geometry

- What is lost?
  - Length

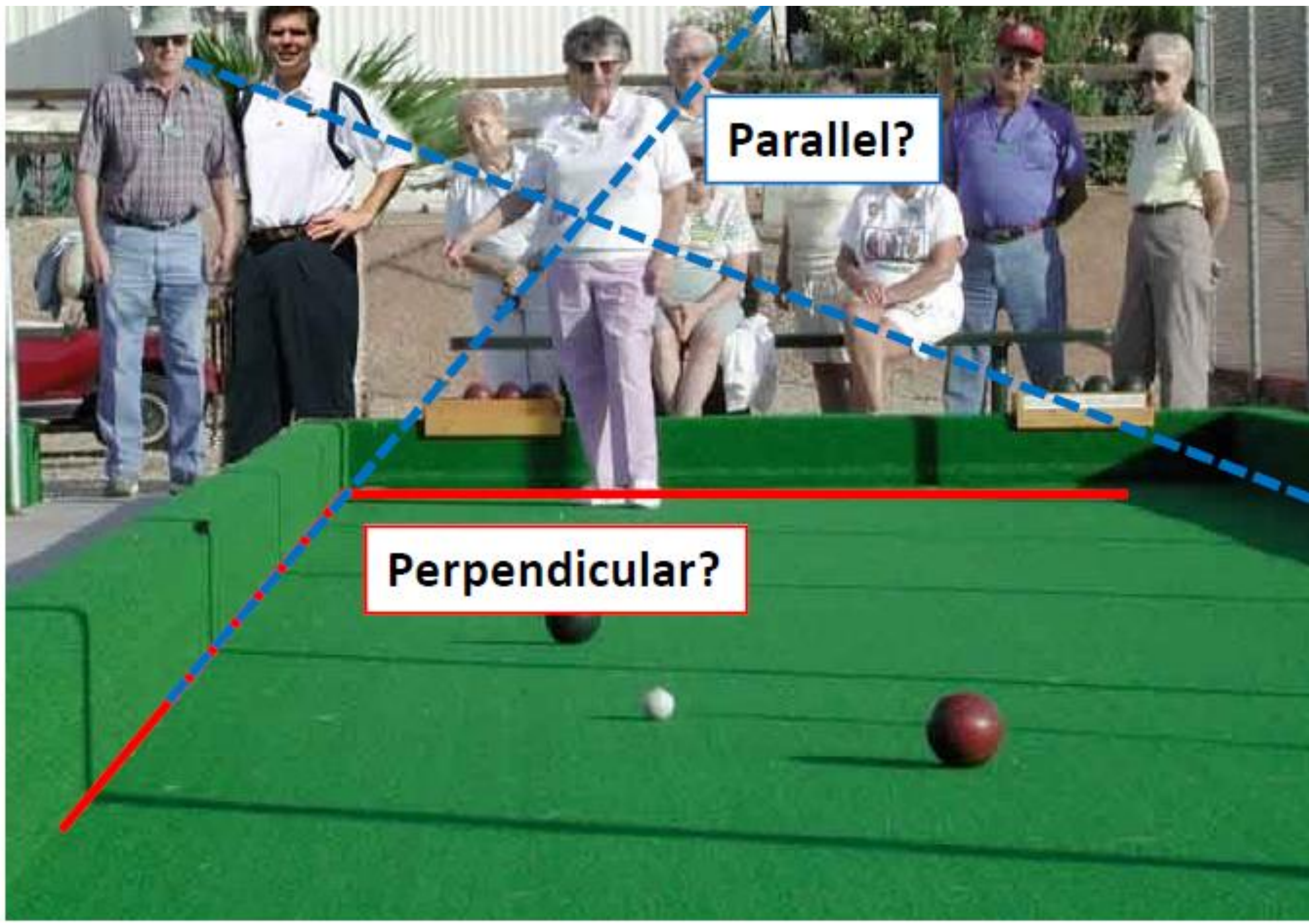


Length is not preserved



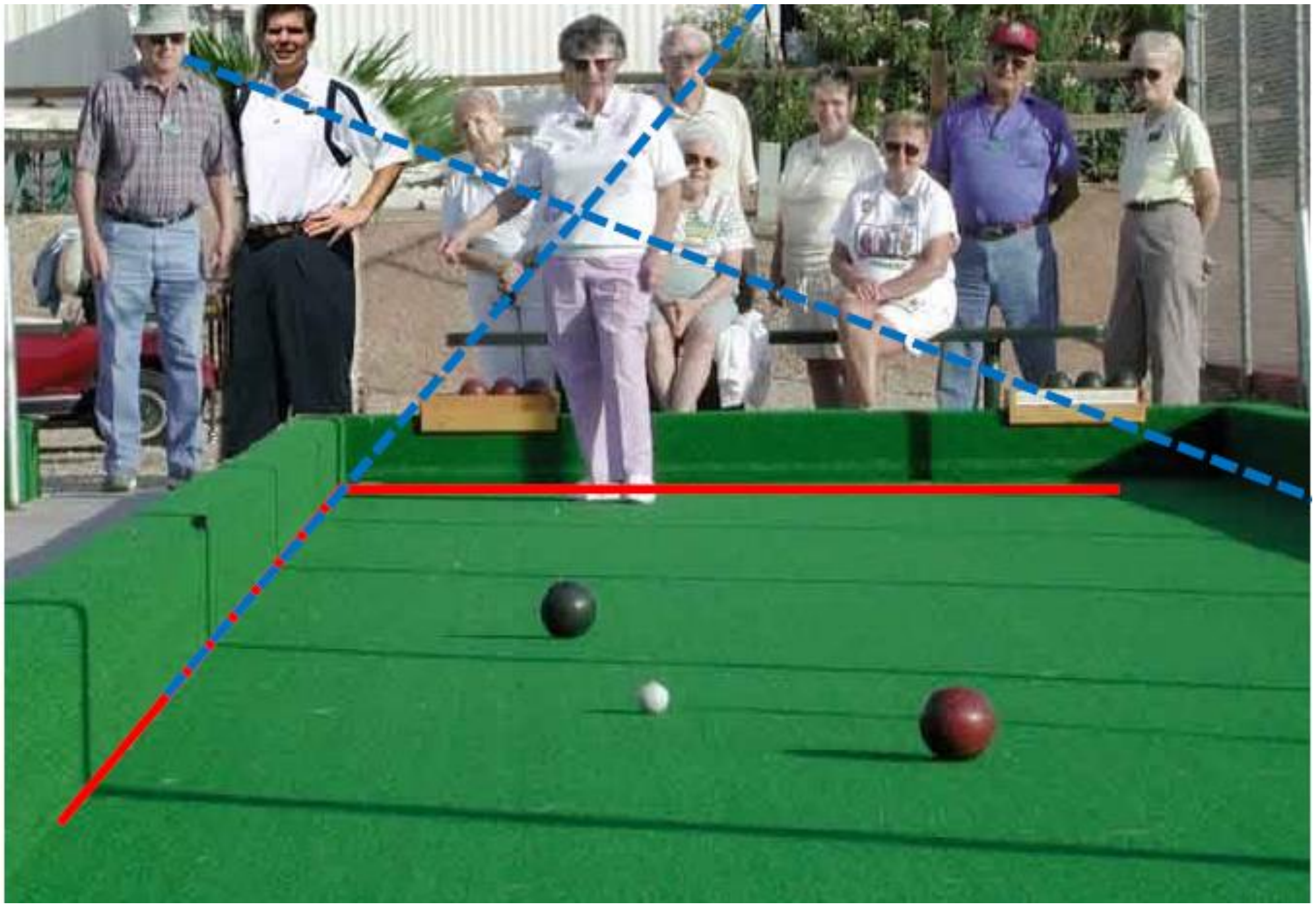
Projective Geometry

- What is lost?
  - Length
  - Angles



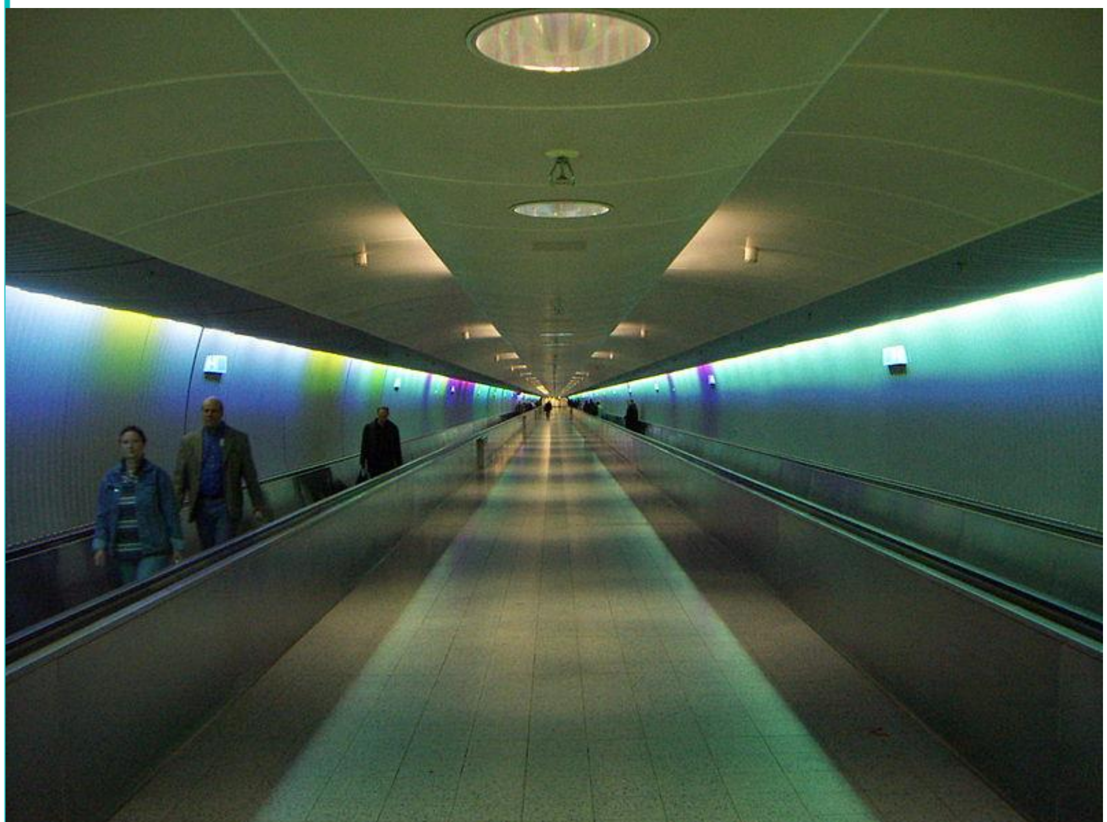
Projective Geometry

- What is preserved?
  - Straight lines are still straight

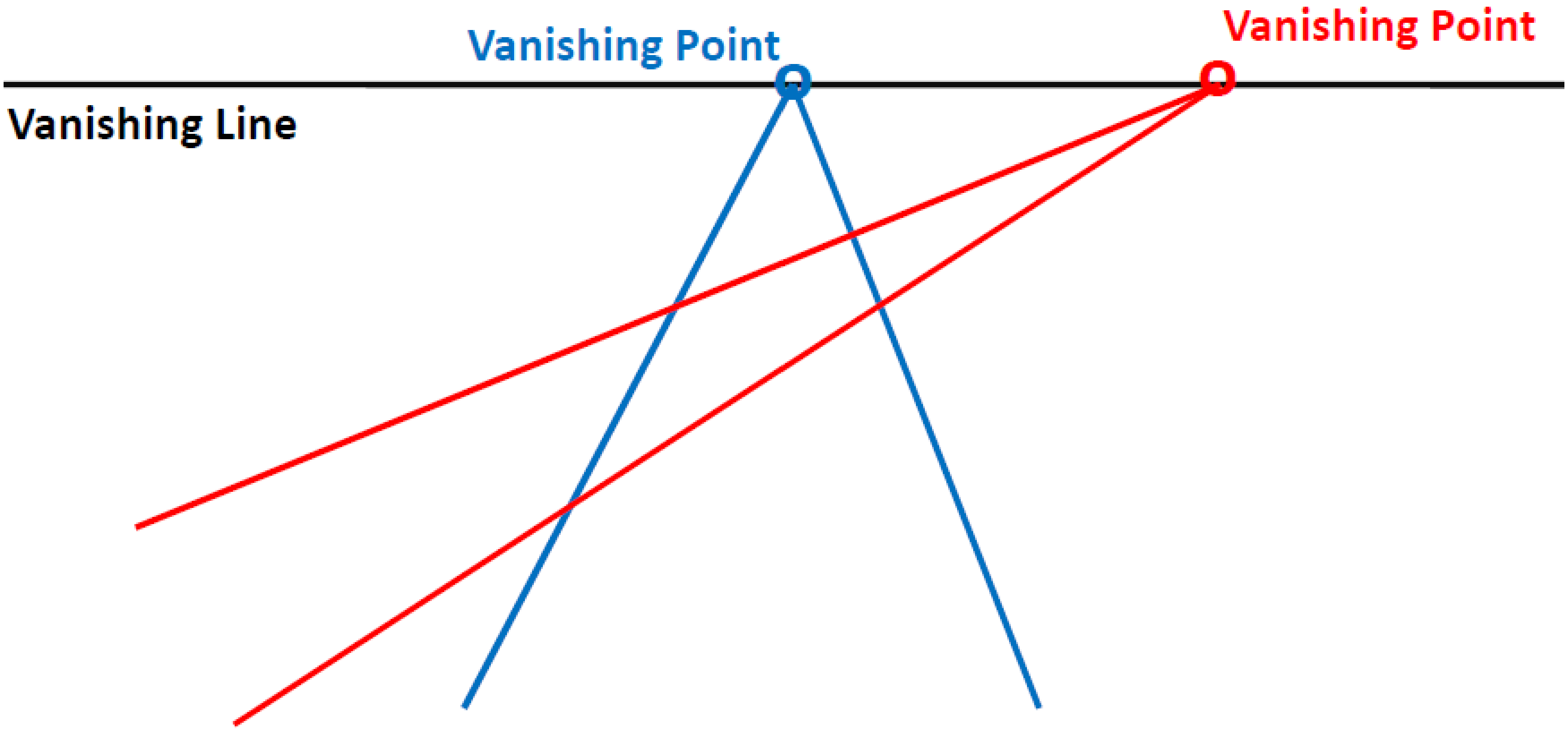


Vanishing points and lines

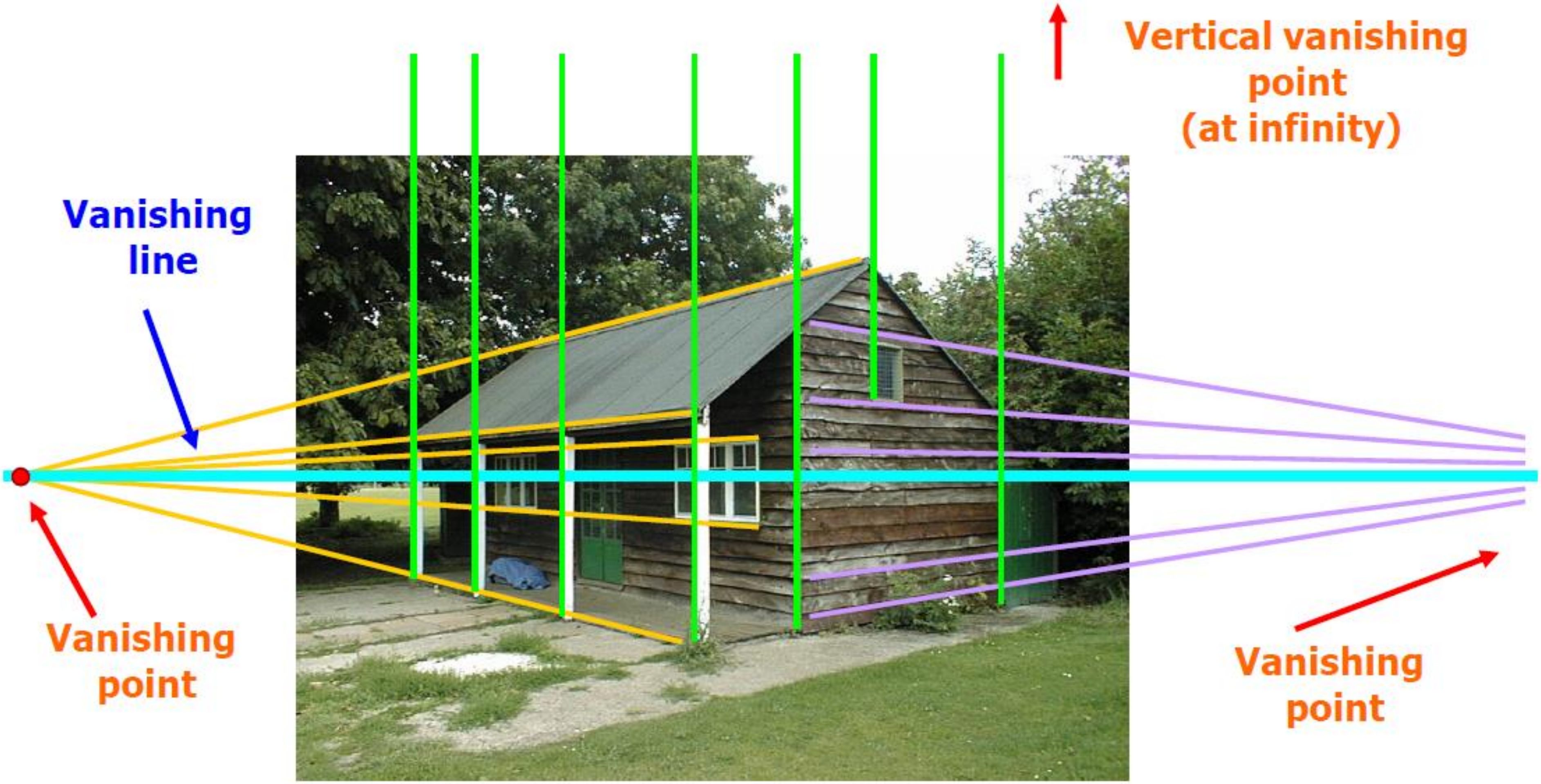
- Parallel lines in the world intersect in the image at a "vanishing point".



Vanishing points and lines

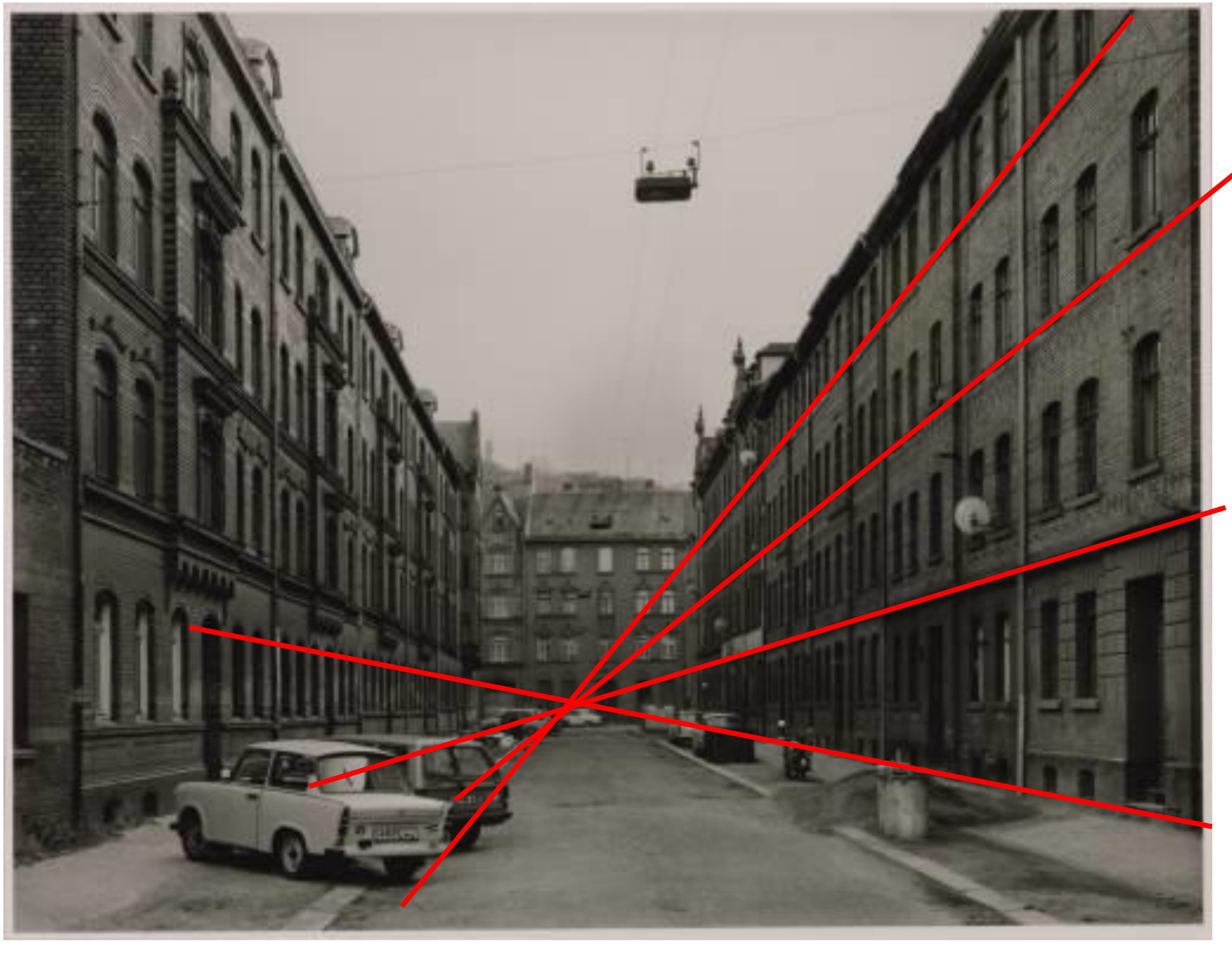


Vanishing points and lines

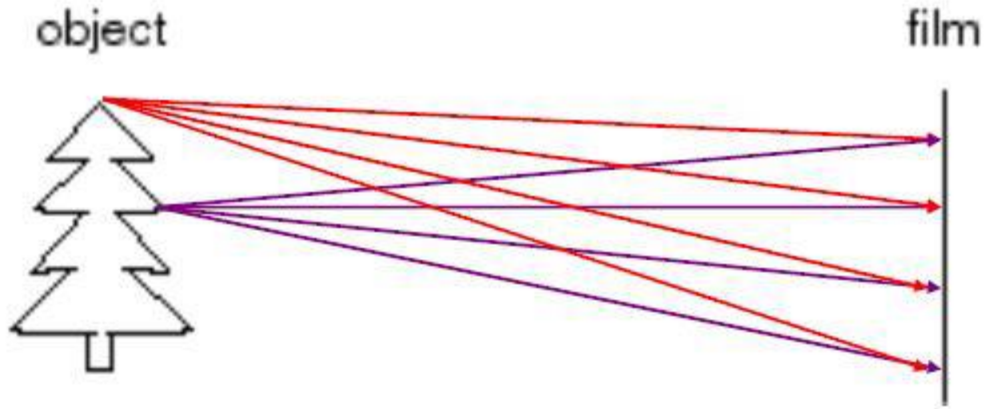


Vanishing points and lines

Note on estimating vanishing points

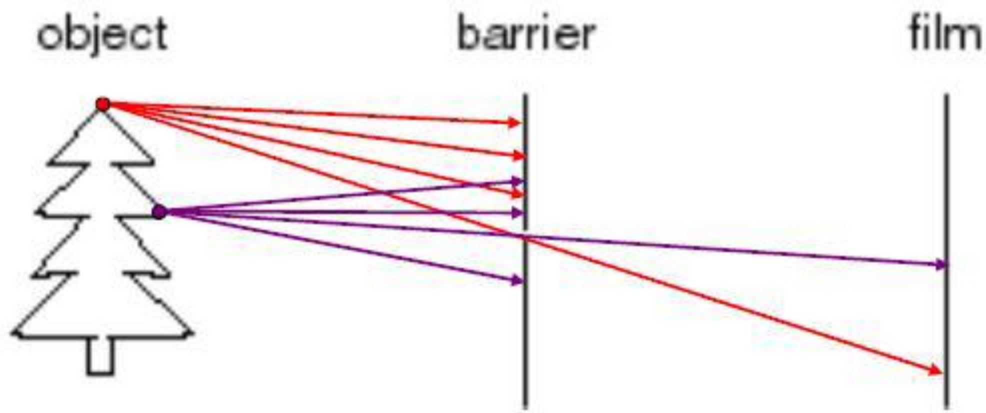


How do we see the world?



- 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**

Camera obscura: the pre-camera

- Known during classical period in China and Greece (e.g. Mo-Ti, China, 470BC to 390BC)

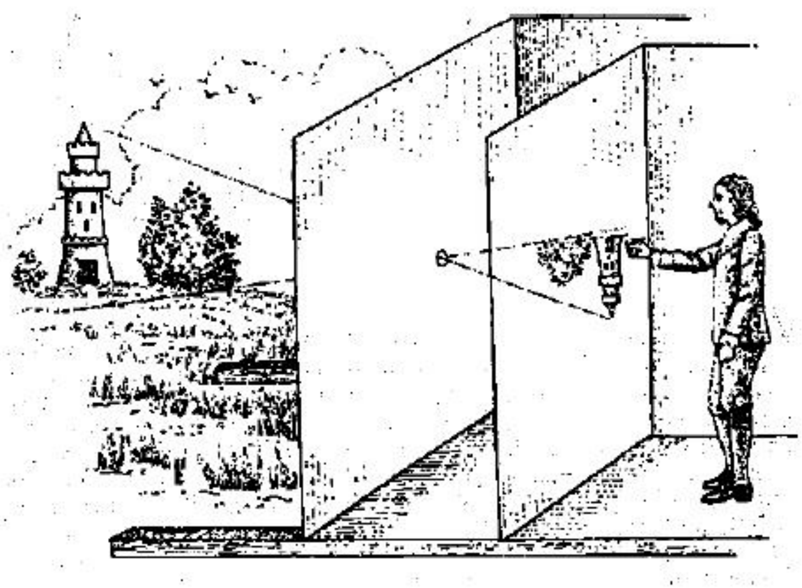
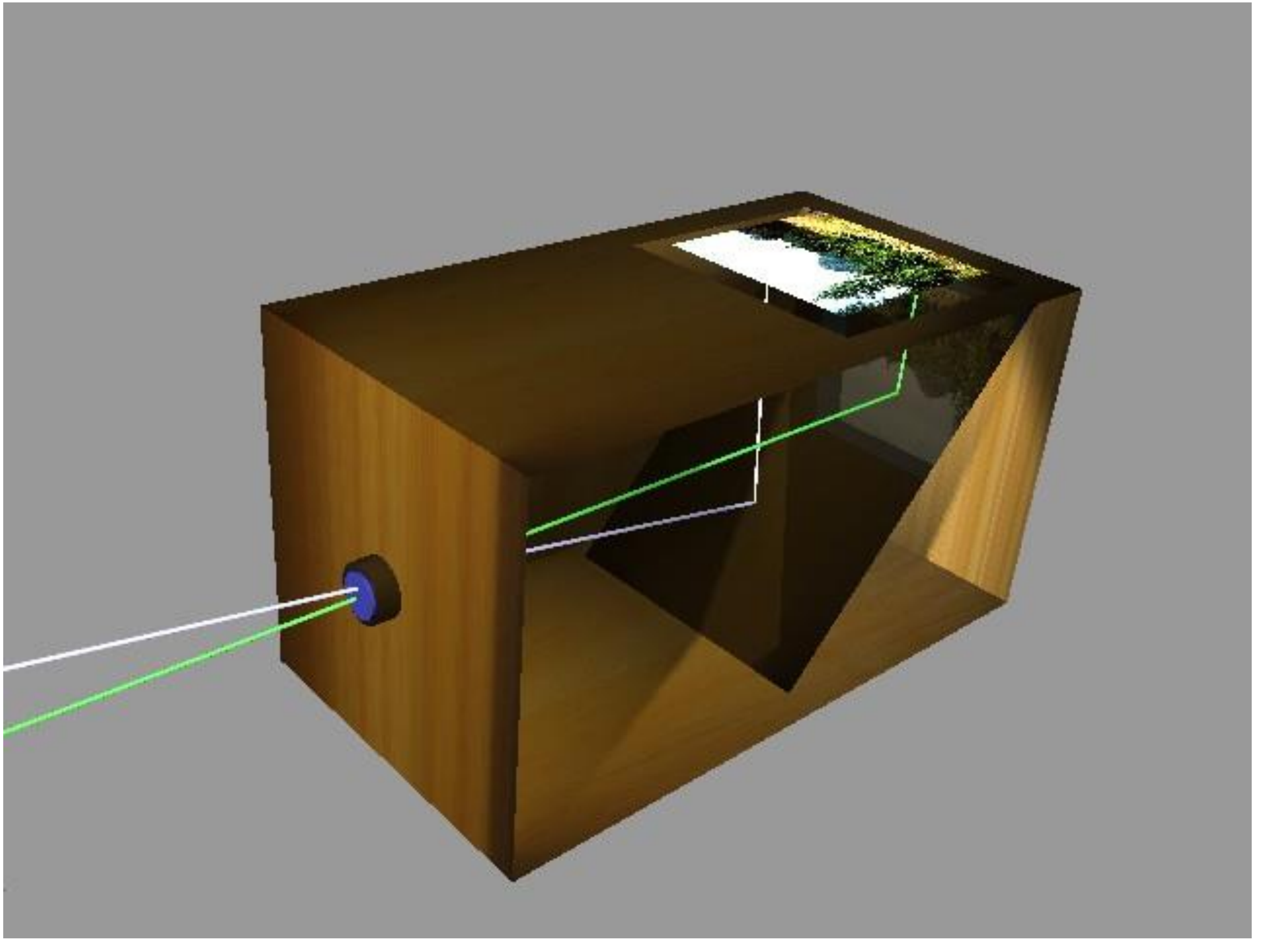
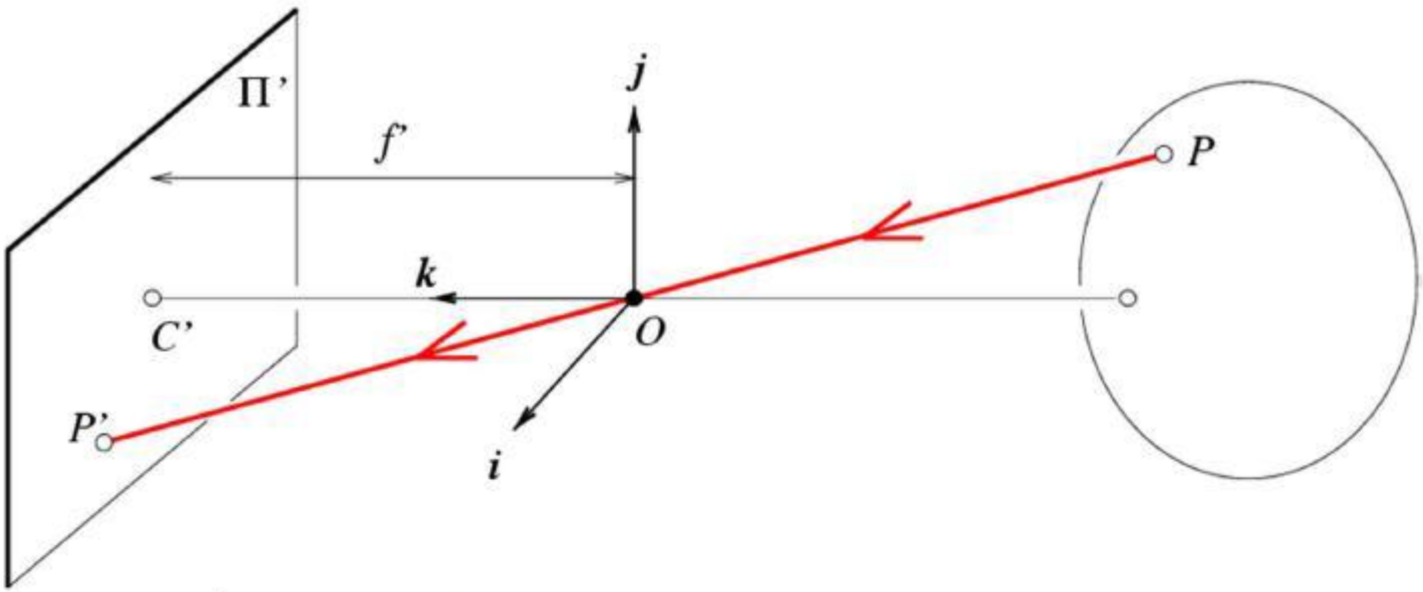


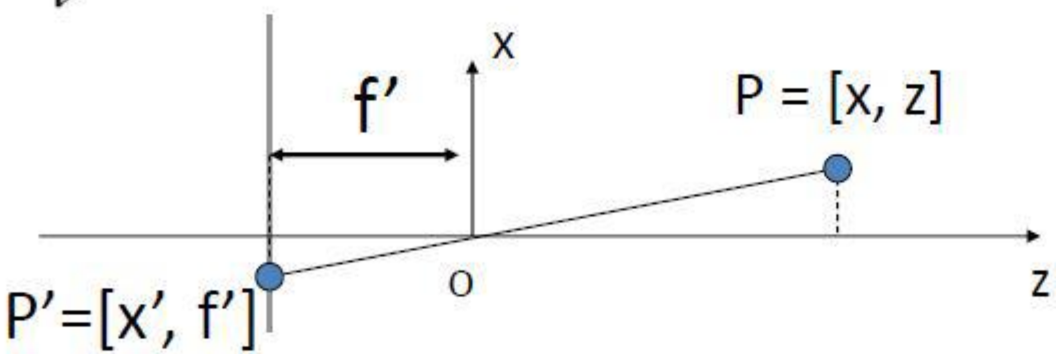
Illustration of Camera Obscura

Freestanding camera obscura at UNC Chapel Hill



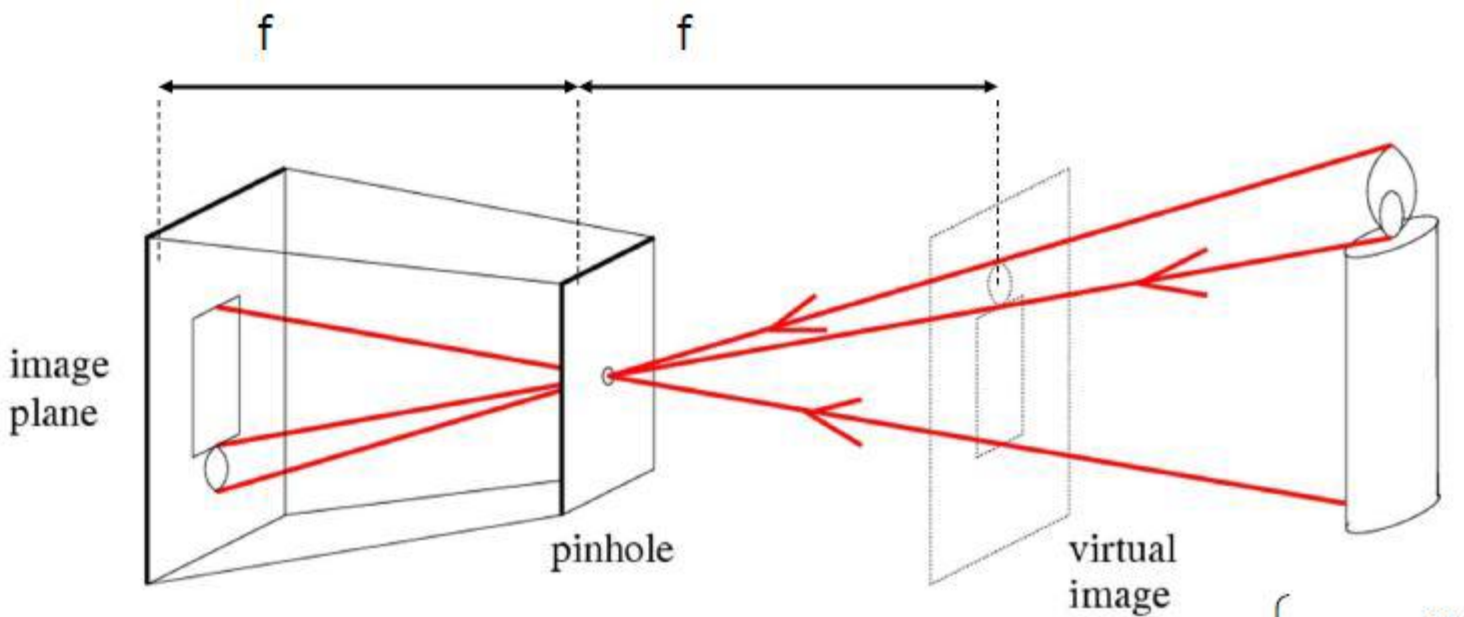
Pinhole camera





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

Pinhole camera

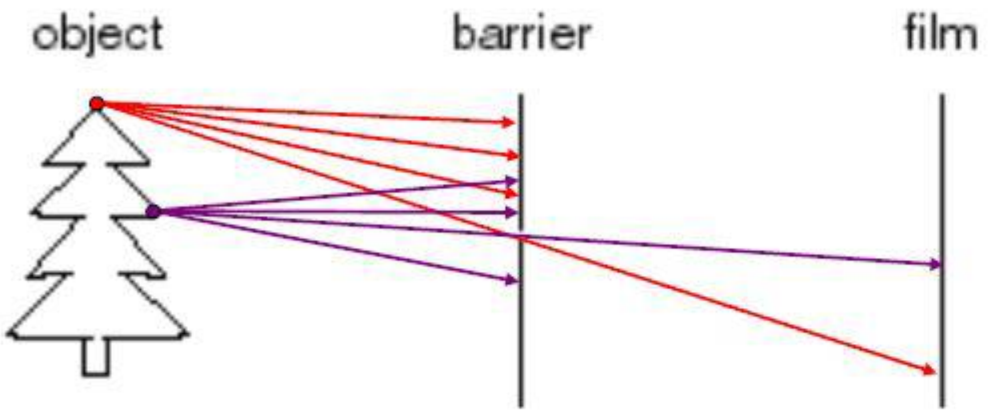


$$\begin{cases} x' = f\,\dfrac{x}{z} \\ y' = f\,\dfrac{y}{z} \end{cases}$$

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

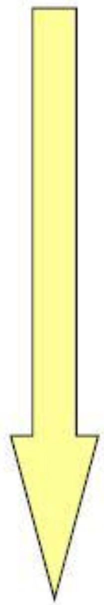
Pinhole camera

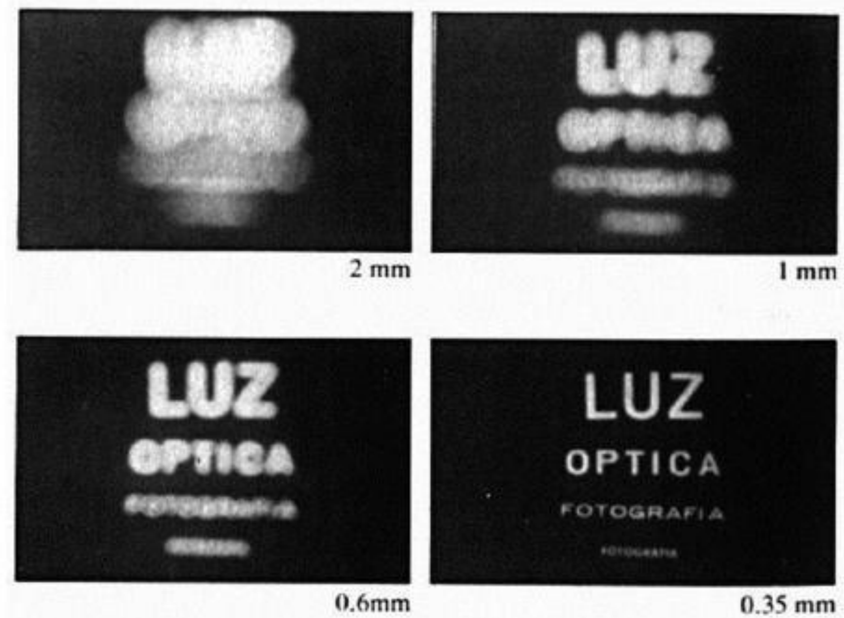
Is the size of the aperture important?



Cameras & Lenses

Shrinking aperture size





- Rays are mixed up

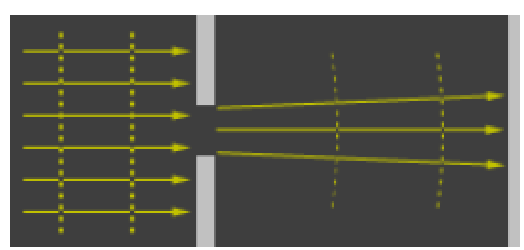
-Why the aperture cannot be too small?

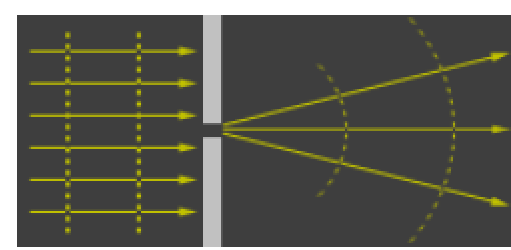
-Less light passes through

-Diffraction effect

Adding lenses!

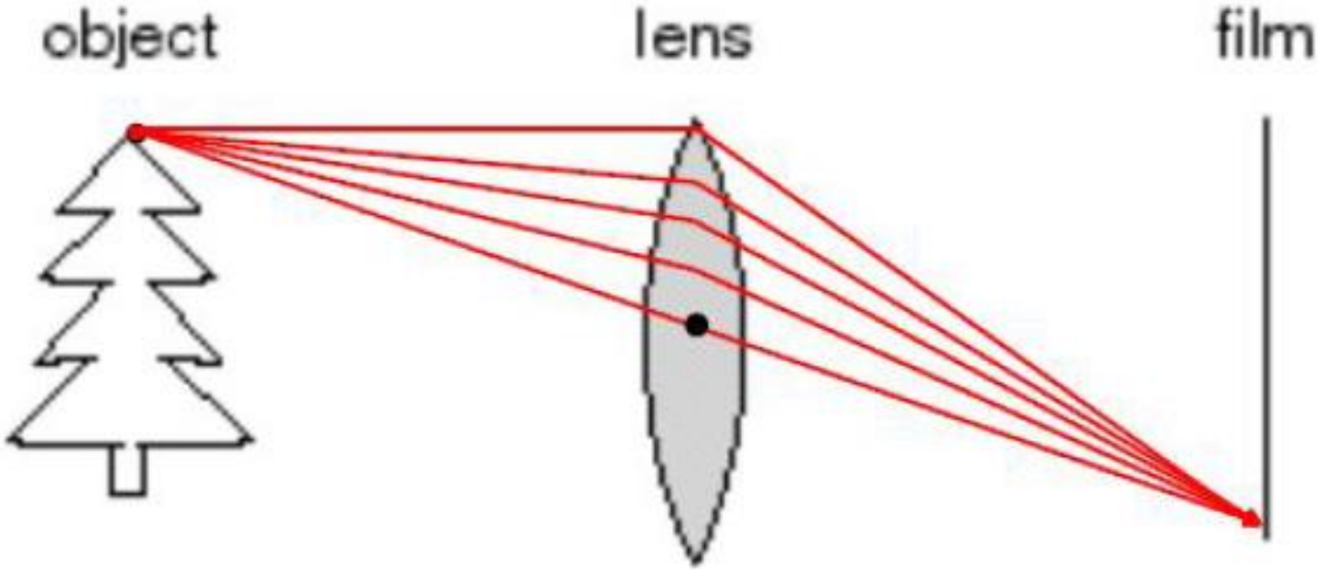
Parallel light rays which pass through a small aperture begin to diverge and interfere with one another. This becomes more significant as the size of the aperture decreases relative to the wavelength of light passing through, but occurs to some extent for any size of aperture or concentrated light source.





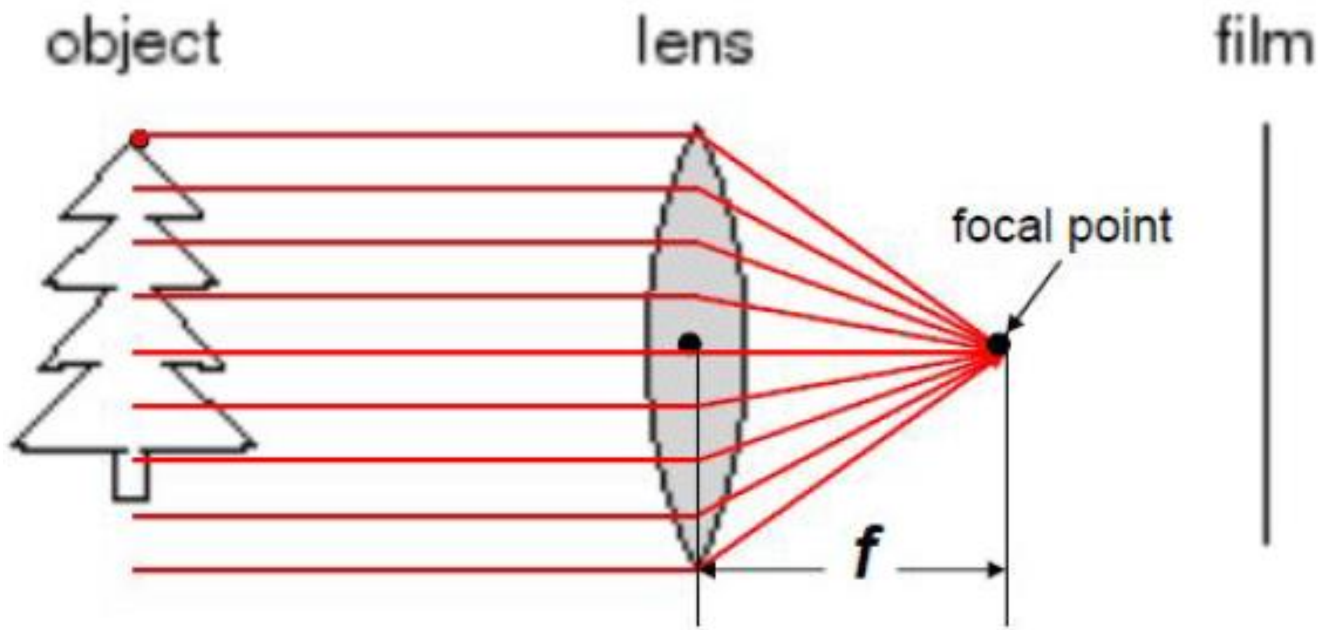
Cameras and Lenses

- **A lens focuses light onto the film**



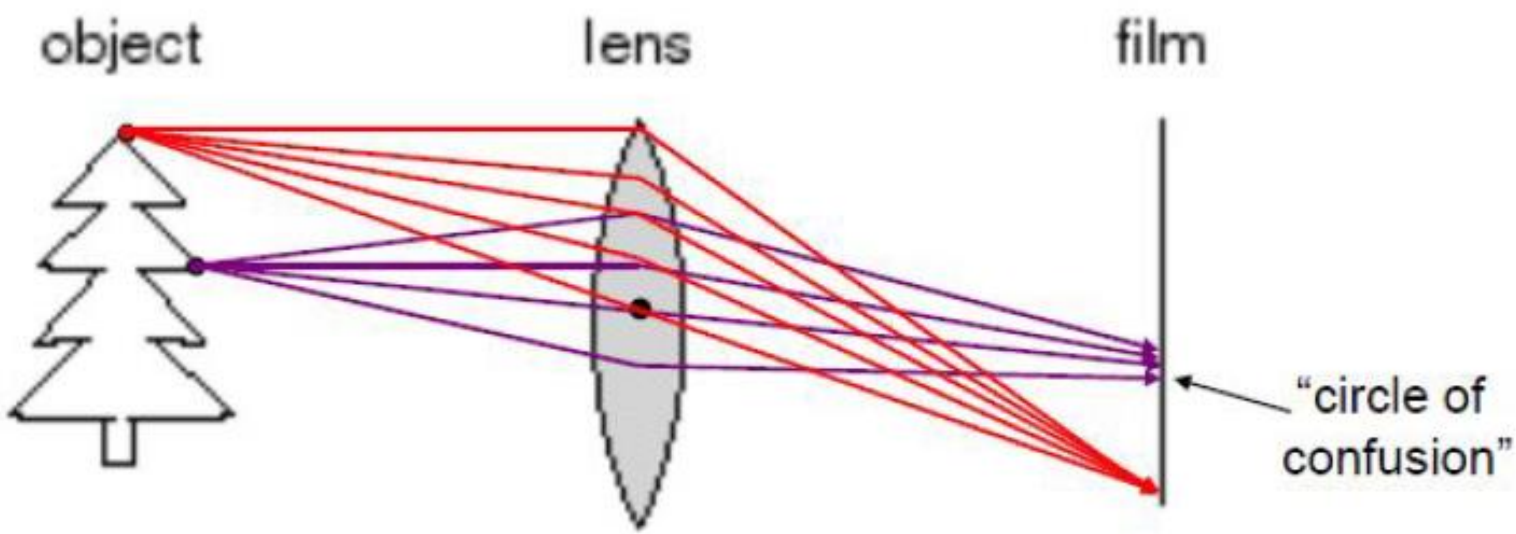
Cameras and Lenses

- **A lens focuses light onto the film**
  - **Rays passing through the center are not deviated.**
  - **All parallel rays converge to one point on a plane located at the *focal length f*.**



Cameras and Lenses

- A lens focuses light onto the film
  - There is a specific distance at which objects are “in focus” [other points project to a “circle of confusion” in the image].



In optics the refractive index or index of refraction of a substance or medium is a measure of the speed of light in that medium

***n* = speed of light in a vacuum / speed of light in medium**

http://en.wikipedia.org/wiki/Refractive_index#Typical_values

Cameras and Lenses

- **Laws of geometric optics:**
  - **Light travels in straight lines in homogeneous medium.**
  - **Reflection upon a surface: incoming ray, surface normal, and reflection are co-planar.**
  - **Refraction: when a ray passes from one medium to another.**

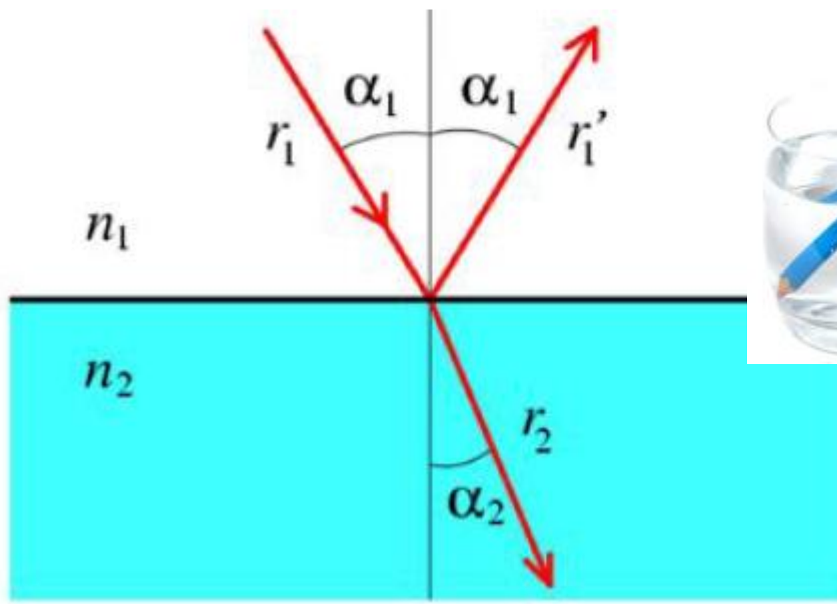
Snell's law

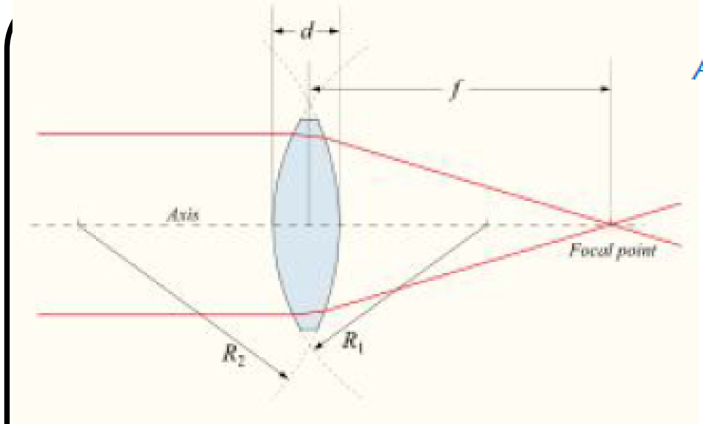
$$n_1 \sin \alpha_1 = n_2 \sin \alpha_2$$

α_1 = incident angle

α_2 = refraction angle

n_i = index of refraction





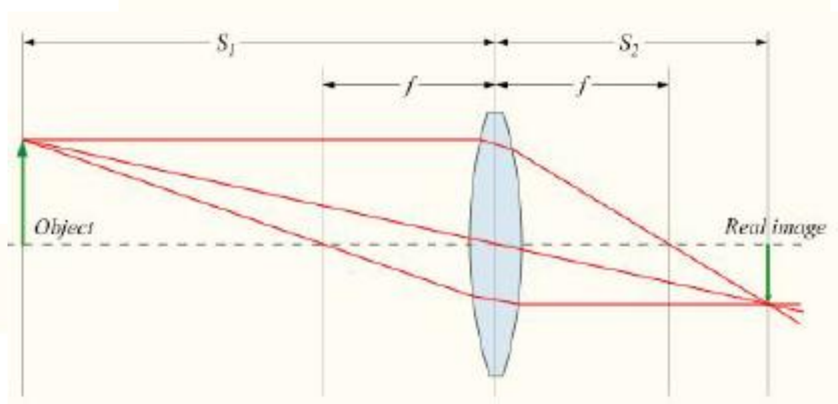
Positive (converging) lens

A lens can be considered a **thin lens** if d << f.

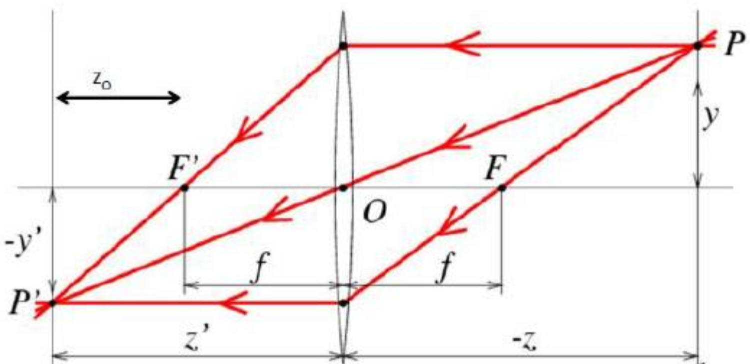
Thin lens equation

If d is small compared to R_1 and R_2, then the thin lens approximation can be made. For a lens in air, f is then given by

$$\frac{1}{f} \approx (n-1)\left[\frac{1}{R_1} - \frac{1}{R_2}\right]$$



Thin Lenses



$$\boxed{\begin{aligned} z' &= f + z_o \\ f &= \frac{R}{2(n-1)} \end{aligned}}$$

Snell's law:

$n_1 \sin \alpha_1 = n_2 \sin \alpha_2$ ⇨

Small angles:

$n_1 \alpha_1 \approx n_2 \alpha_2$

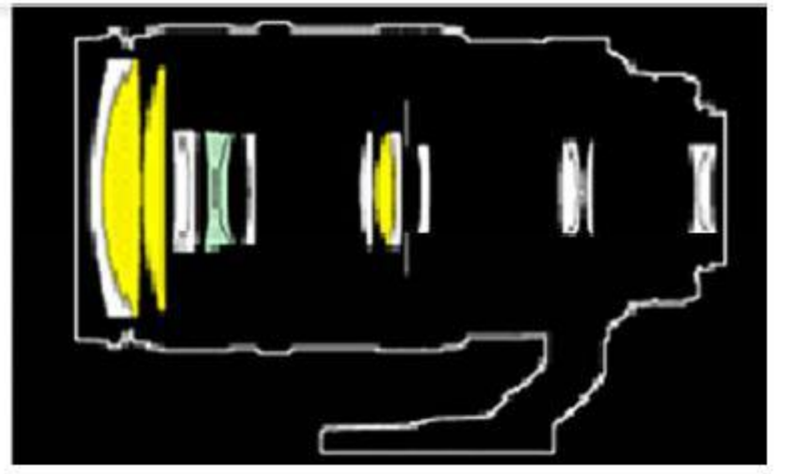
$n_1 = n$ (lens)

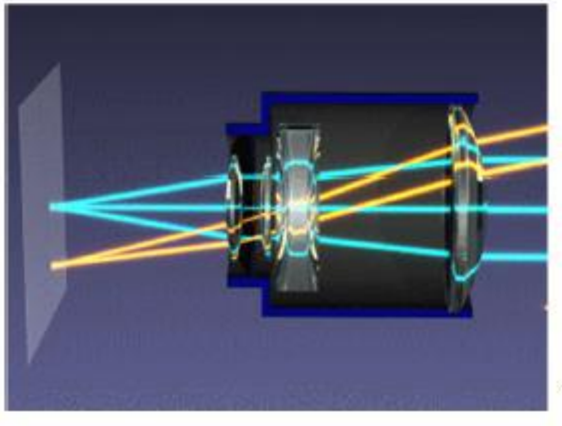
$n_1 = 1$ (air)

⇨

$$\begin{cases} x' = z'\dfrac{x}{z} \\ y' = z'\dfrac{y}{z} \end{cases}$$

Cameras & Lenses





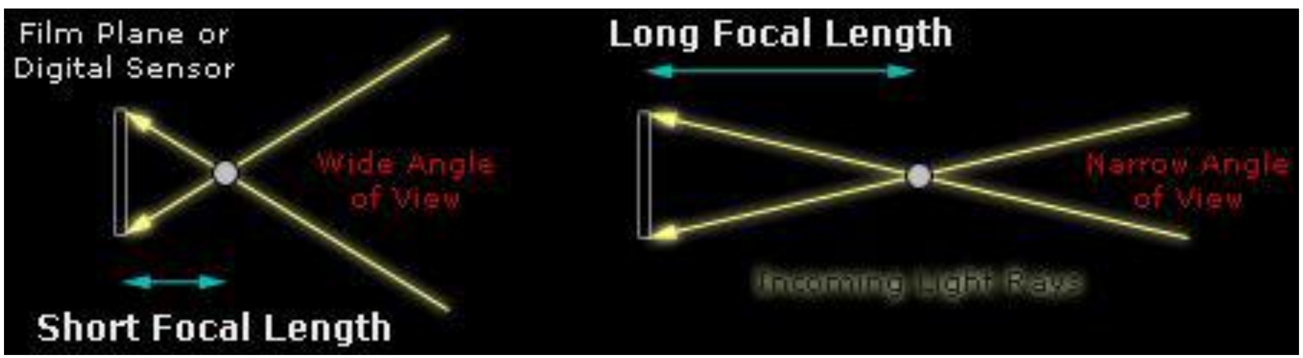
Source wikipedia

28 mm lens

50 mm lens

70 mm lens

210 mm lens

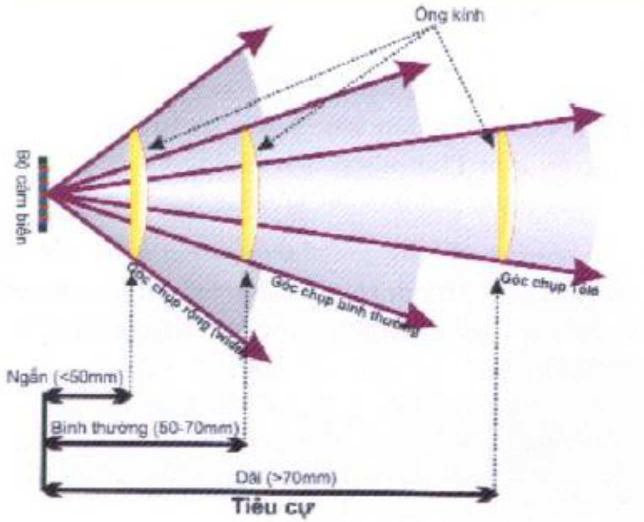


Lens is described by its focal length, which is the distance in millimeters (mm) between the lens and the image it forms on the sensor or film http://photography-cameras.org/camera-lenses/focal-length-of-camera-lens

Tiêu cự của máy ảnh là thông số cho biết góc nhìn của máy ảnh, nghĩa là khoảng phạm vi mà máy ảnh có thể "thâu tóm" được

http://www.chuphinh.vn/may-chup-hinh-ky-thuat-so/cac-yeu-to-ky-thuat/75-tieu-cu.html

Hình 13: Minh họa phạm vi thu hình của các góc chụp

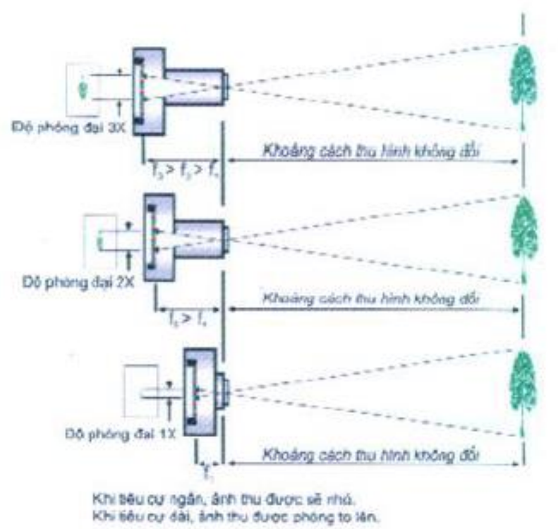


Hình 14: Máy ảnh Sony DSC-S750

f = 5.8 – 17.4 mm Với thông số này cho biết máy ảnh có khả năng thay đổi tiêu cự từ 35 mm (góc chụp rộng – wide) đến 105 mm (góc chụp – tele).

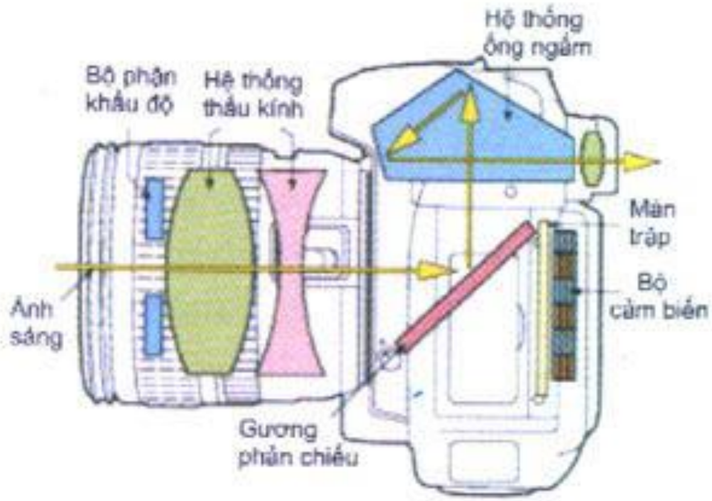
Optical 3X: Lấy 105/35 (hay lấy 17.4/5.8) ta được kết quả là 3 (thông số zoom)

Hình 16: Ảnh hưởng của tiêu cự trên độ lớn của ảnh chụp

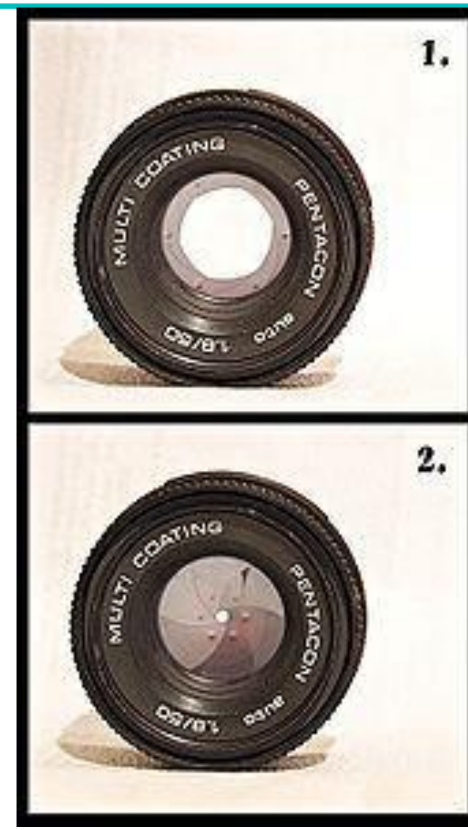


odel

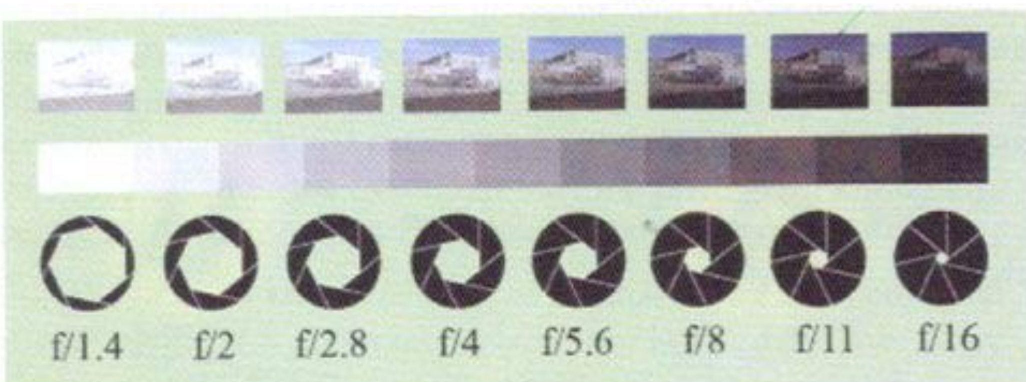
Bộ phận khẩu độ điều tiết lượng ánh sáng đi qua hệ thống thấu kính.



Hình 17: Khẩu độ (Aperture)



Large (top) and small (bottom) apertures

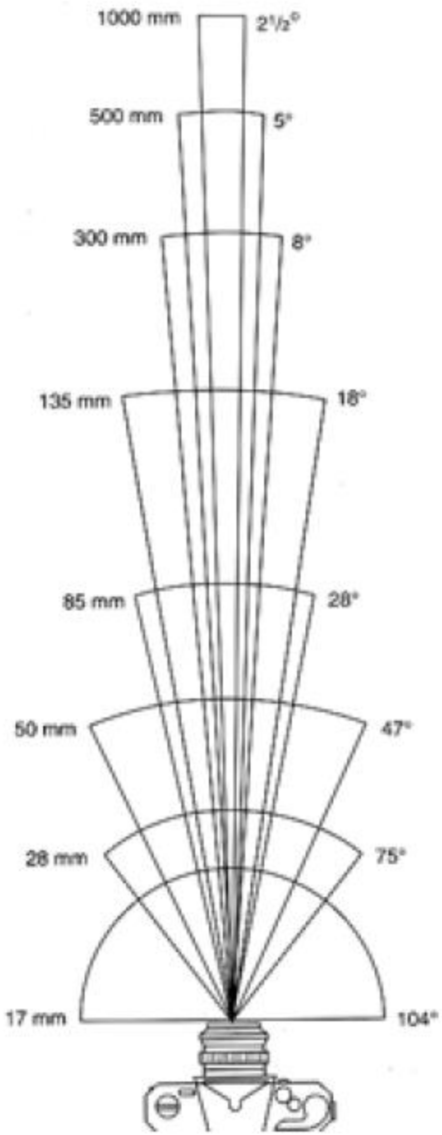


Hình 18: Khẩu độ với cường độ ánh sáng tương ứng

Khẩu độ không hoạt động độc lập, mà nó phải có kết hợp chặc chẽ với độ dài tiêu cự để có được hiệu ứng ánh sáng tốt nhất.

Máy ảnh số đã có chế độ đo sáng tự động, các chế độ chụp cài sẵn. Những chương trình này đã được lập trình sẵn và dựa vào hệ thống đo sáng, máy sẽ lựa chọn khẩu độ phù hợp với khung cảnh tại thời điểm chụp.

Field of View (Zoom, focal length)

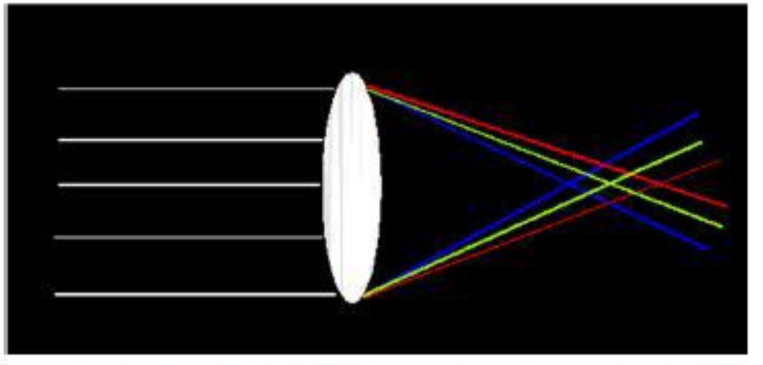


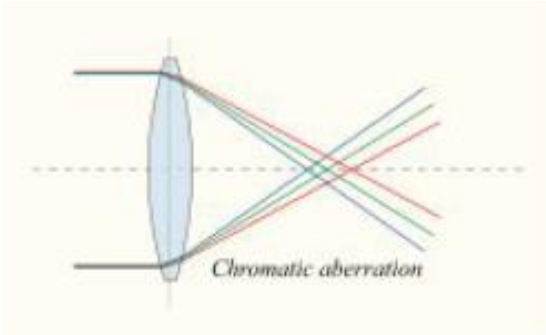
From London and Upton

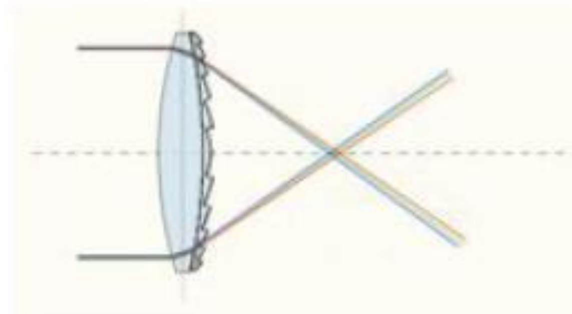
Issues with lenses: Chromatic Aberration

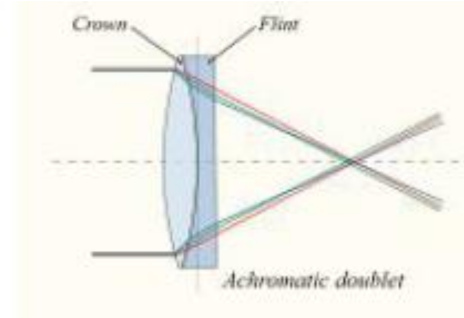
- Lens has different refractive indices for different wavelengths: causes color fringing

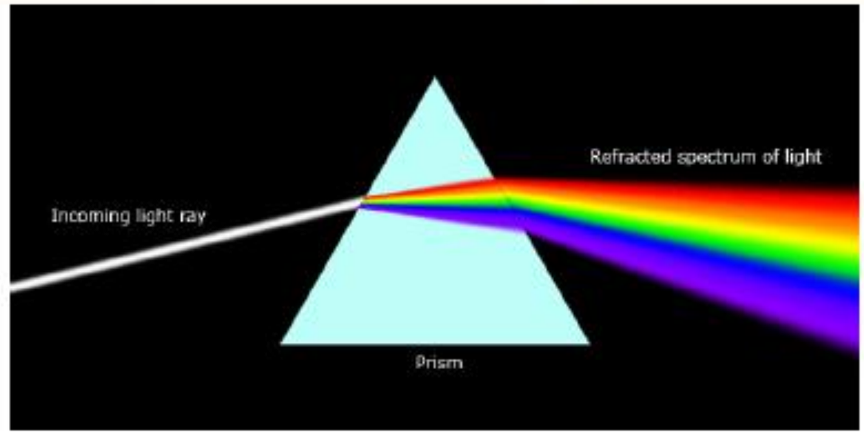
$$f = \frac{R}{2(n-1)}$$

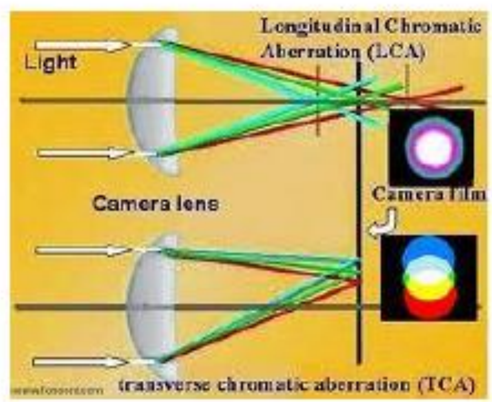



Chromatic aberration




Crown
Flint
Achromatic doublet

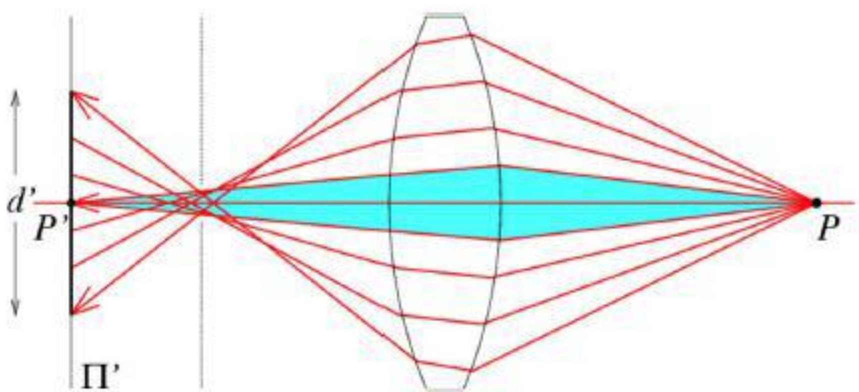

Refracted spectrum of light
Incoming light ray
Prism

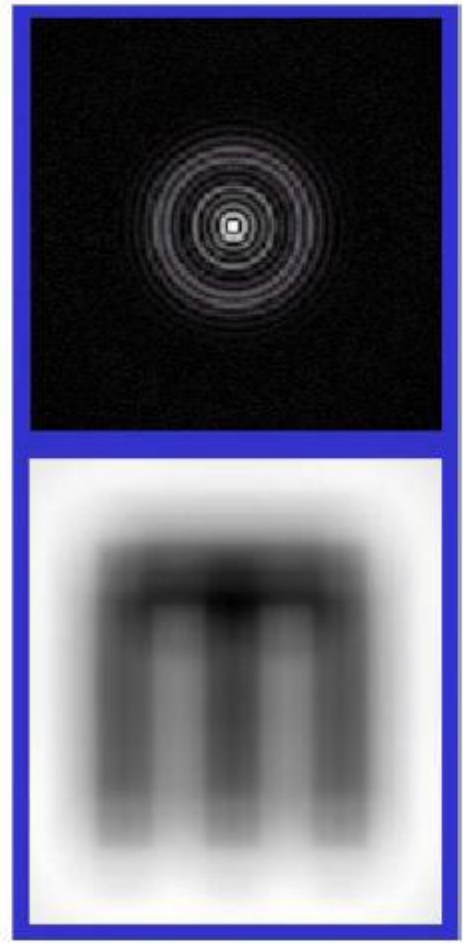

Longitudinal Chromatic
Aberration (LCA)
Light
Camera lens
Camera film
transverse chromatic aberration (TCA)

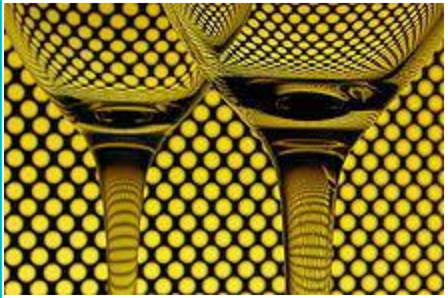

ANIMATING
ANIMATING

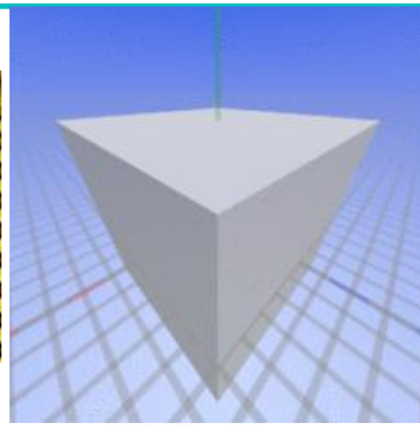
Issues with lenses: Chromatic Aberration

- Rays farther from the optical axis focus closer





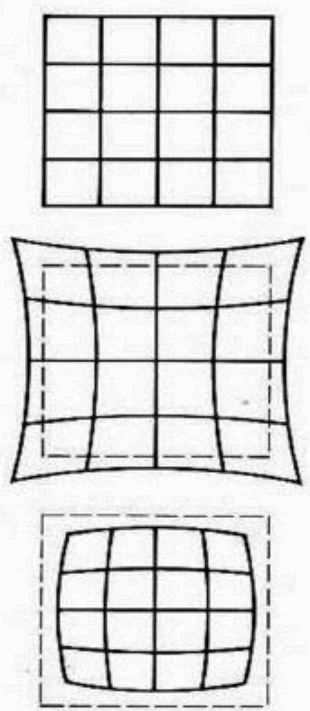




This simulation shows how adjusting the angle of view of a camera, while varying the camera distance, keeping the object in frame, results in vastly differing images. At narrow angles, large distances, light rays are nearly parallel, resulting in a "flattened" image. At wide angles, short distances, the object appears distorted.

Issues with lenses: Chromatic Aberration

– Deviations are most noticeable for rays that pass through the edge of the lens



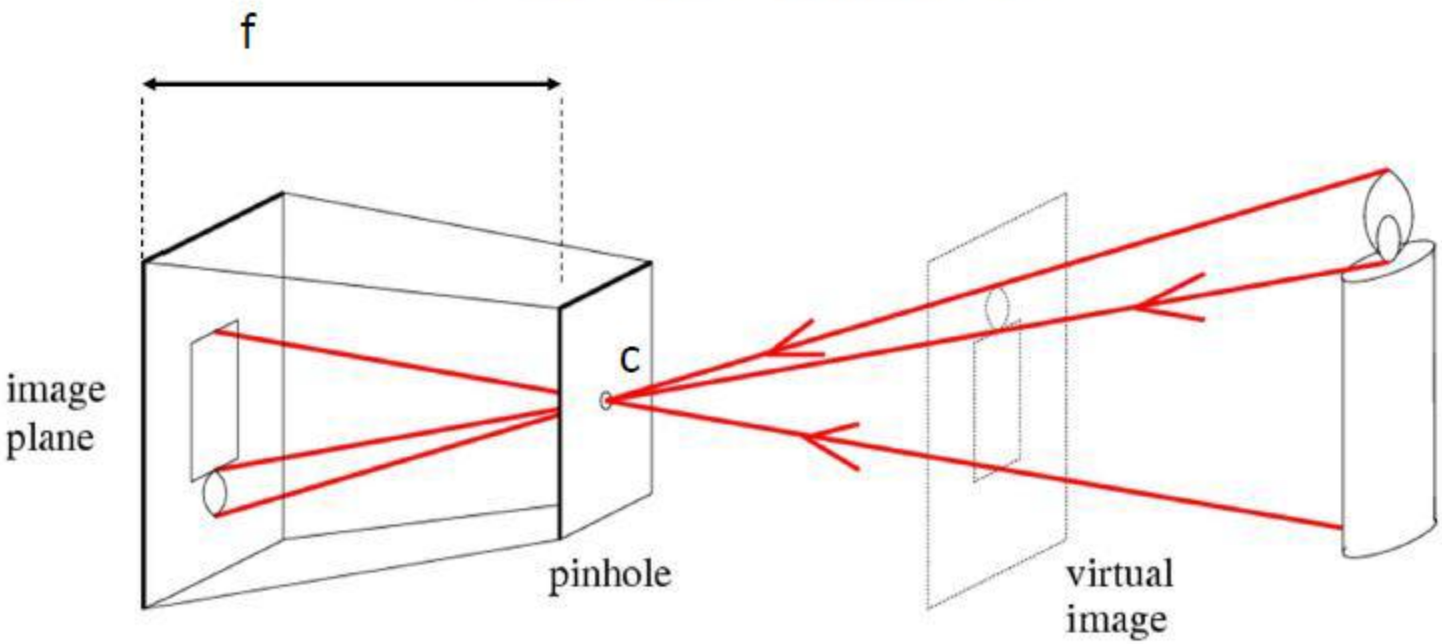
No distortion

Pin cushion

Barrel (fisheye lens)

Image magnification decreases with distance from the optical axis

Pinhole camera



f = focal length
c = center of the camera

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

$$\Re^3 \xrightarrow{E} \Re^2$$

Pinhole camera

Is this a linear transformation?

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

No — division by z is nonlinear!

How to make it linear?

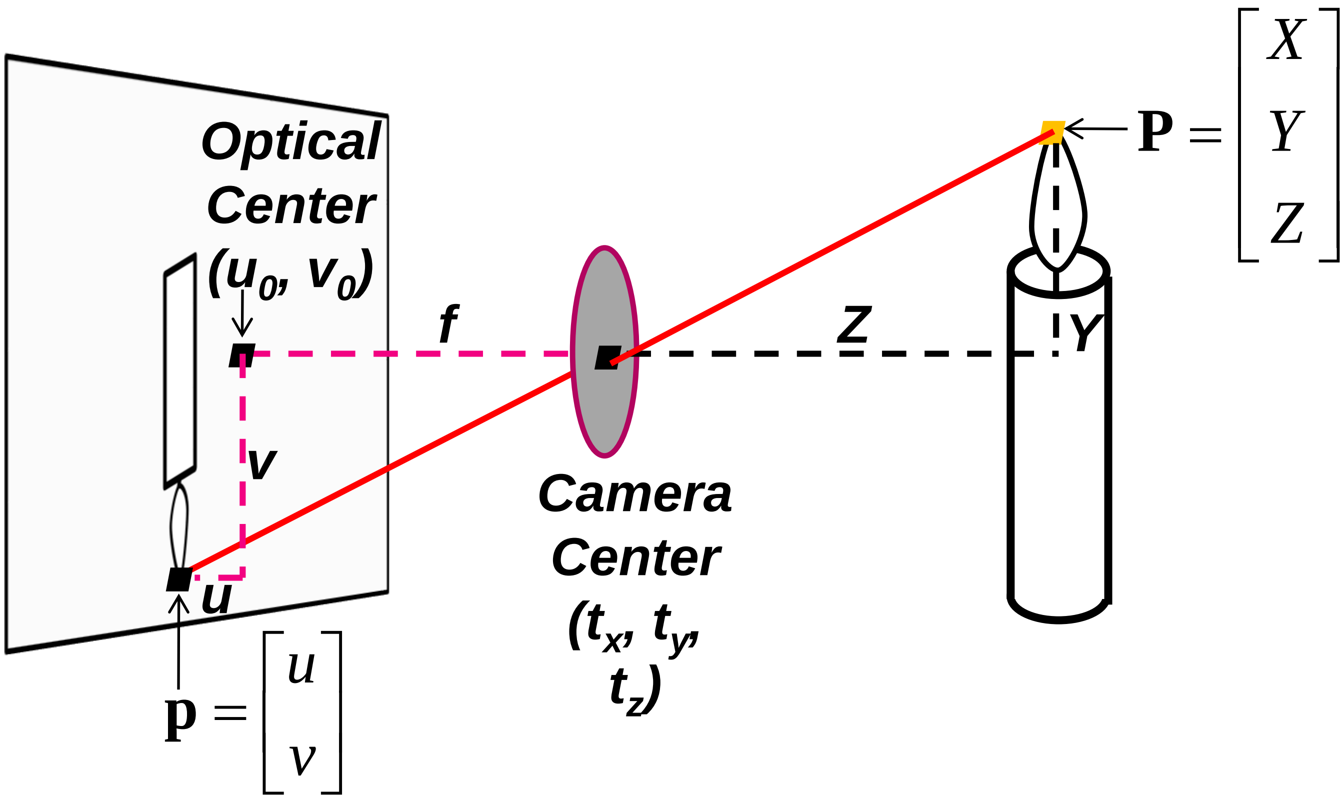
Mathematically, for a linear system, F, defined by F(x) = y, where x is some sort of stimulus (input) and y is some sort of response (output), the superposition (i.e., sum) of stimuli yields a superposition of the respective responses:

$$F(x_1 + x_2 + \ldots) = F(x_1) + F(x_2) + \ldots$$

In the field of electrical engineering, where the x and y signals are allowed to be complex-valued (as is common in signal processing), a linear system must satisfy the superposition property, which requires the system to be additive and homogeneous

$$F(x_1 + x_2) = F(x_1) + F(x_2) \qquad F(ax) = aF(x)$$

Projection
World coordinates ➡ Image coordinates



Homogeneous coordinates

Conversion

Converting to homogeneous coordinates

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

homogeneous image coordinates

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

homogeneous scene coordinates

Converting from homogeneous coordinates

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

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

Homogeneous coordinates

Perspective Projection Transformation:

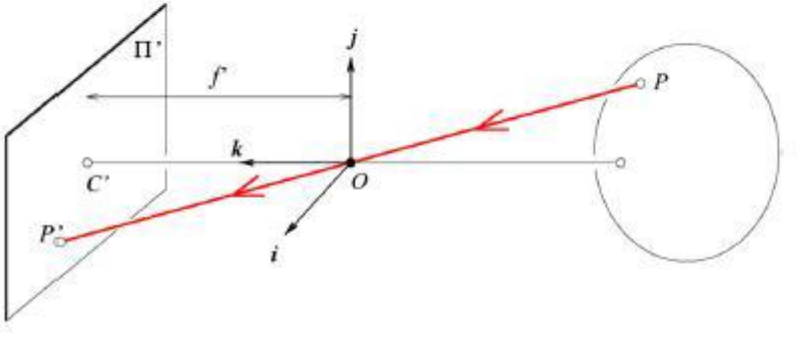
$$P' = \begin{bmatrix} f\,x \\ f\,y \\ z \end{bmatrix} = \underbrace{\begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{\text{M}} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

"Projection matrix"

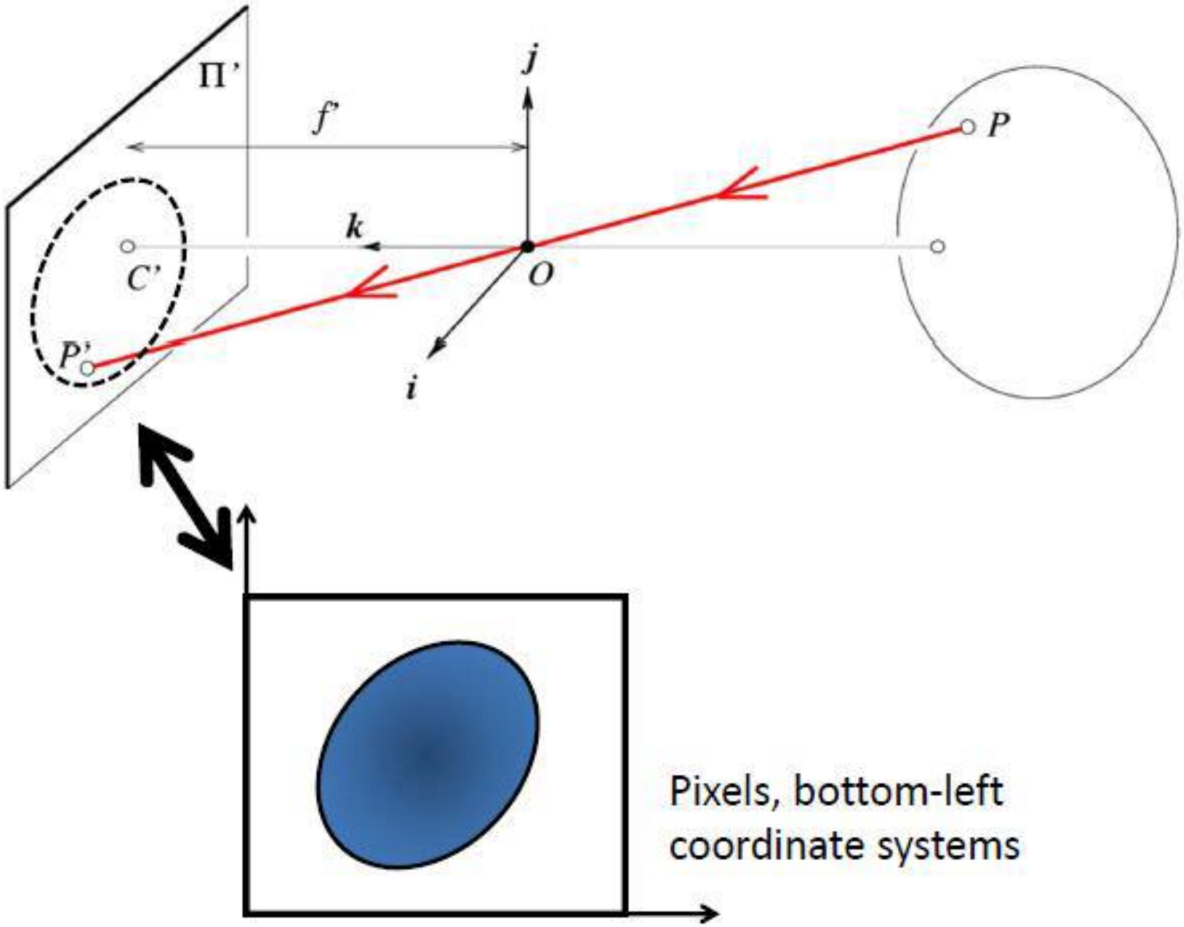
$$P' = M\,P$$

$$\Re^4 \xrightarrow{\text{H}} \Re^3$$

$$P_i' = \begin{bmatrix} f\dfrac{x}{z} \\ f\dfrac{y}{z} \end{bmatrix}$$

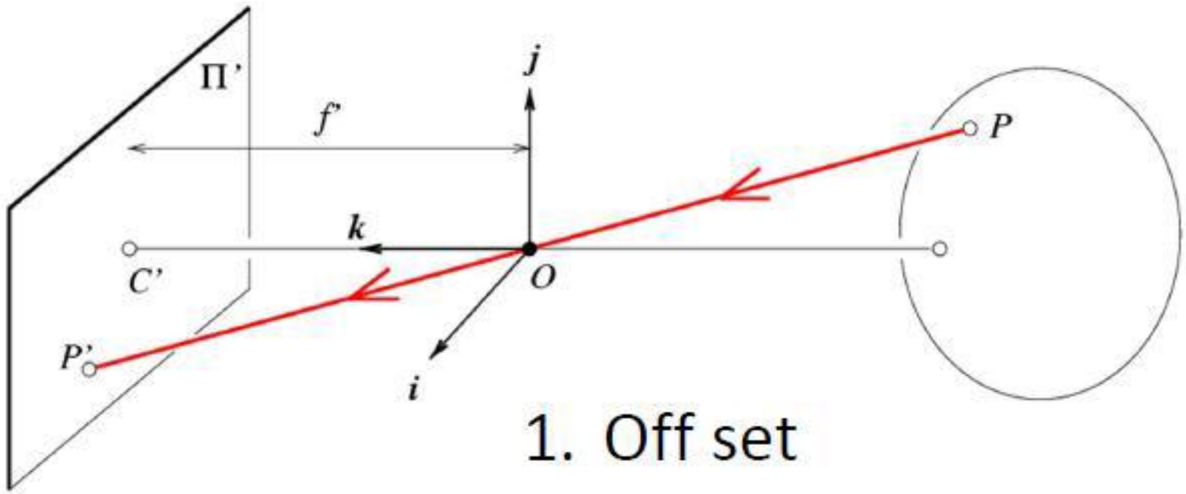


From retina plane to images

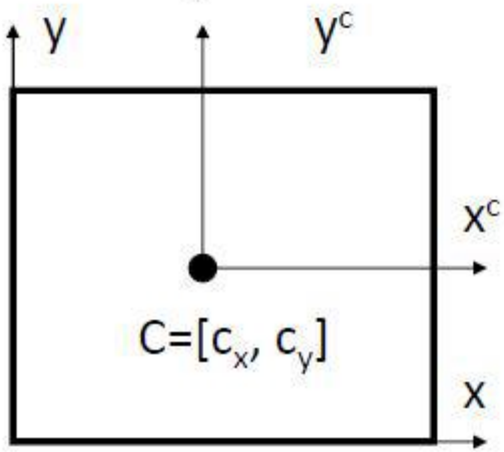


Pixels, bottom-left coordinate systems

From retina plane to images

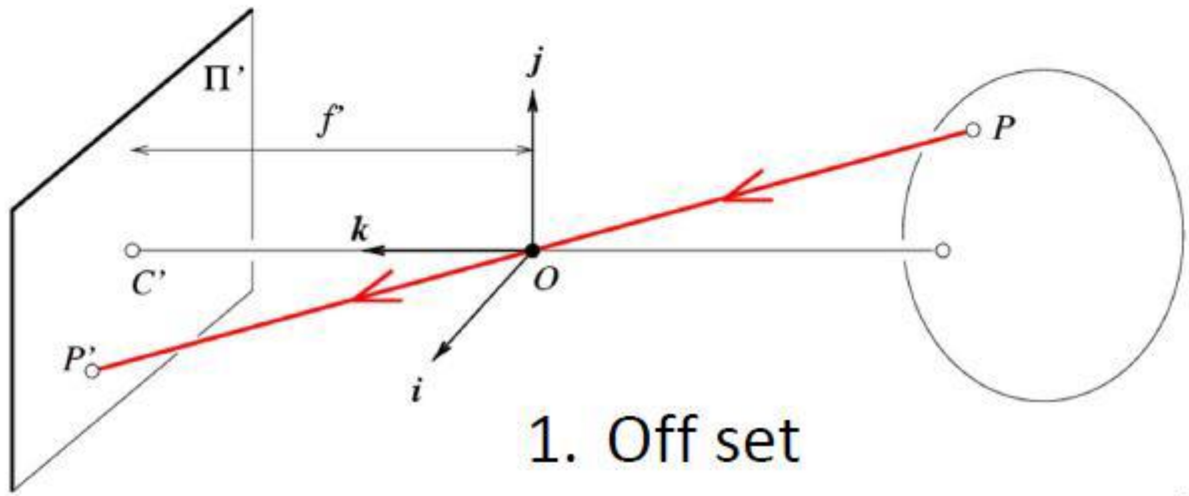


1. Off set

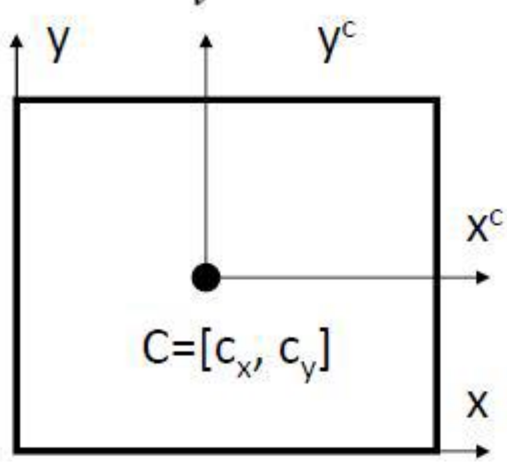


$$(x, y, z) \rightarrow (f\frac{x}{z} + c_x, f\frac{y}{z} + c_y)$$

From retina plane to images



1. Off set
2. From metric to pixels



$$(x, y, z) \rightarrow (\underbrace{f\;k}_{\alpha}\frac{x}{z} + c_x, \underbrace{f\;l}_{\beta}\frac{y}{z} + c_y)$$

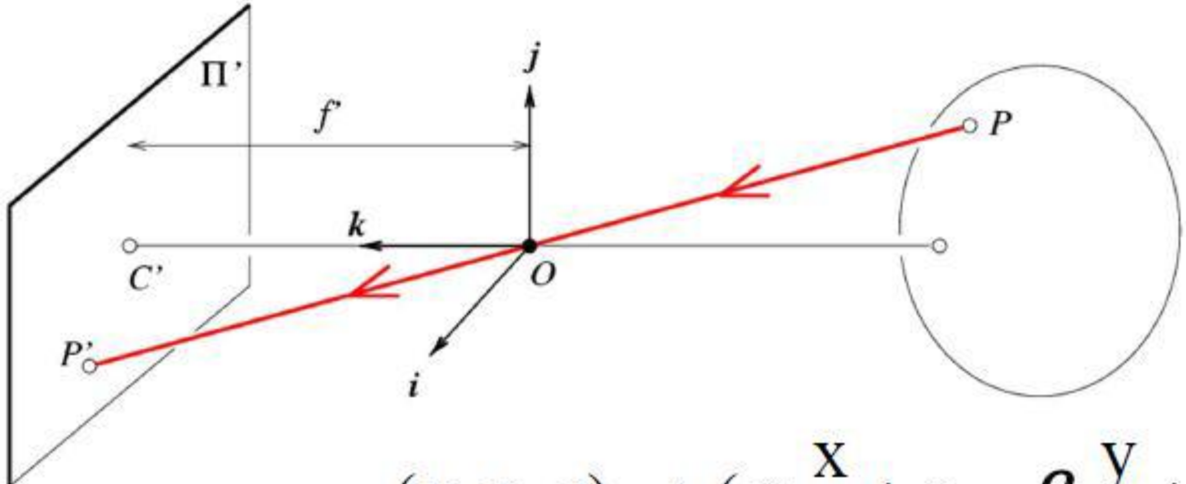
Units: k,l : pixel/m

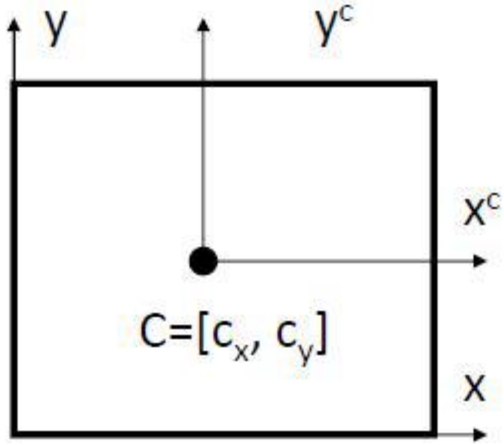
f : m

Non-square pixels

α, β : pixel

From retina plane to images

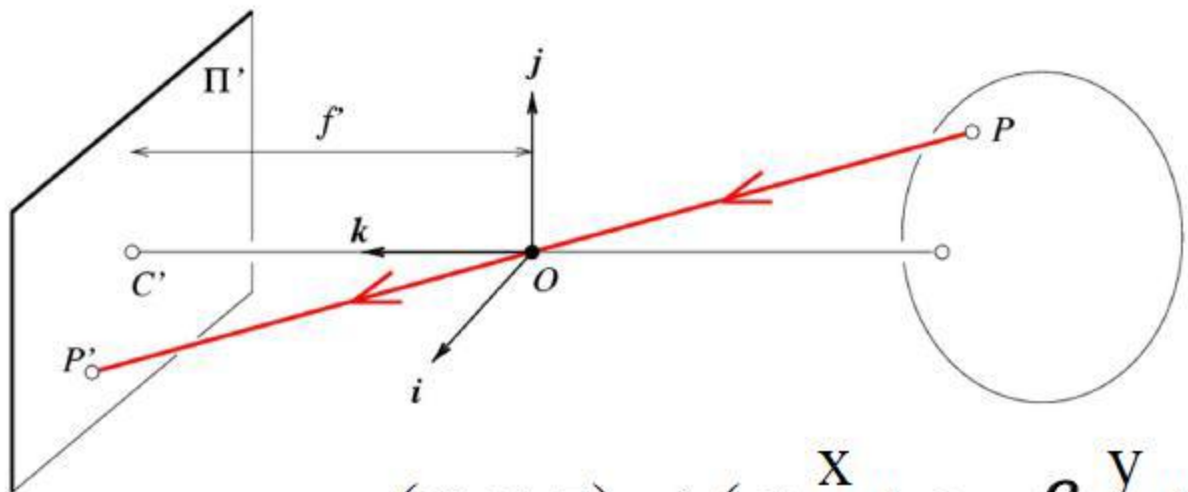


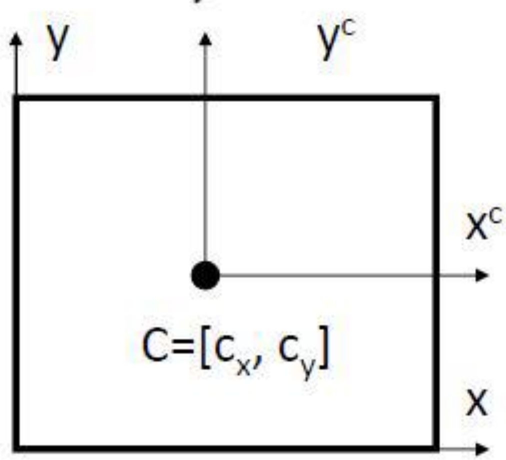


$$(\mathrm{x}, \mathrm{y}, \mathrm{z}) \rightarrow \left(\boldsymbol{\alpha} \frac{\mathrm{x}}{\mathrm{z}} + \mathrm{c}_{\mathrm{x}}, \boldsymbol{\beta} \frac{\mathrm{y}}{\mathrm{z}} + \mathrm{c}_{\mathrm{y}}\right)$$

Matrix form?

Camera matrix

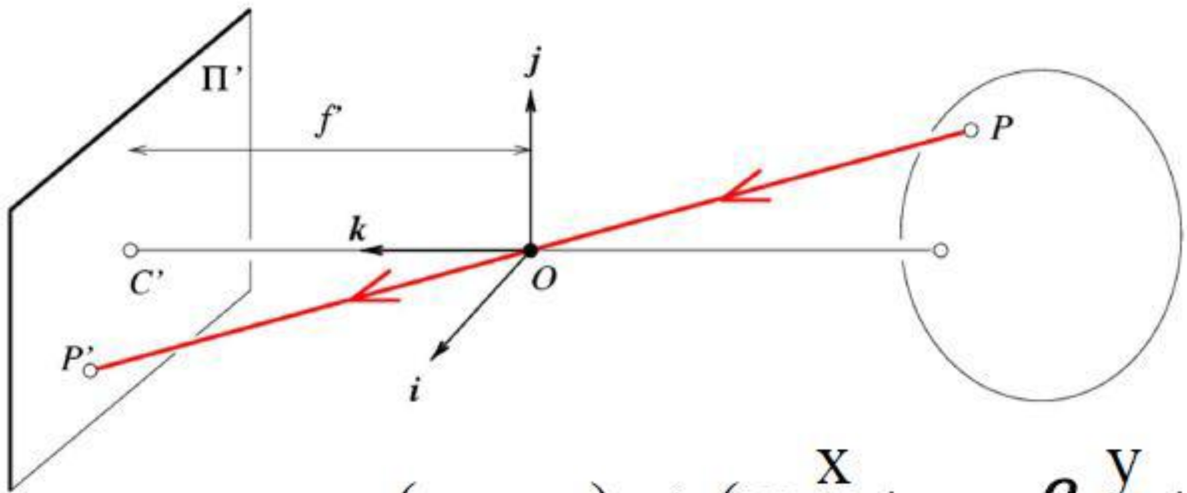


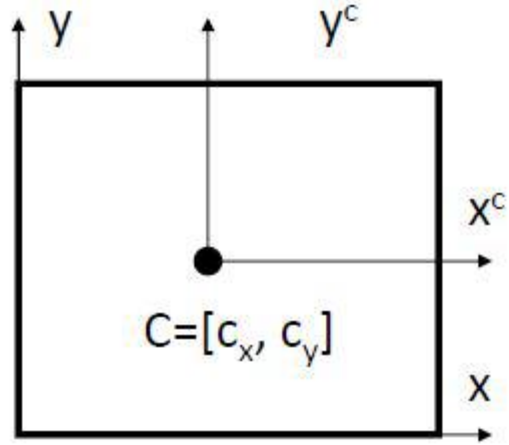


$$(\mathrm{x},\mathrm{y},\mathrm{z}) \rightarrow \left(\boldsymbol{\alpha}\frac{\mathrm{x}}{\mathrm{z}}+\mathrm{c}_{\mathrm{x}},\ \boldsymbol{\beta}\frac{\mathrm{y}}{\mathrm{z}}+\mathrm{c}_{\mathrm{y}}\right)$$

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

Camera matrix

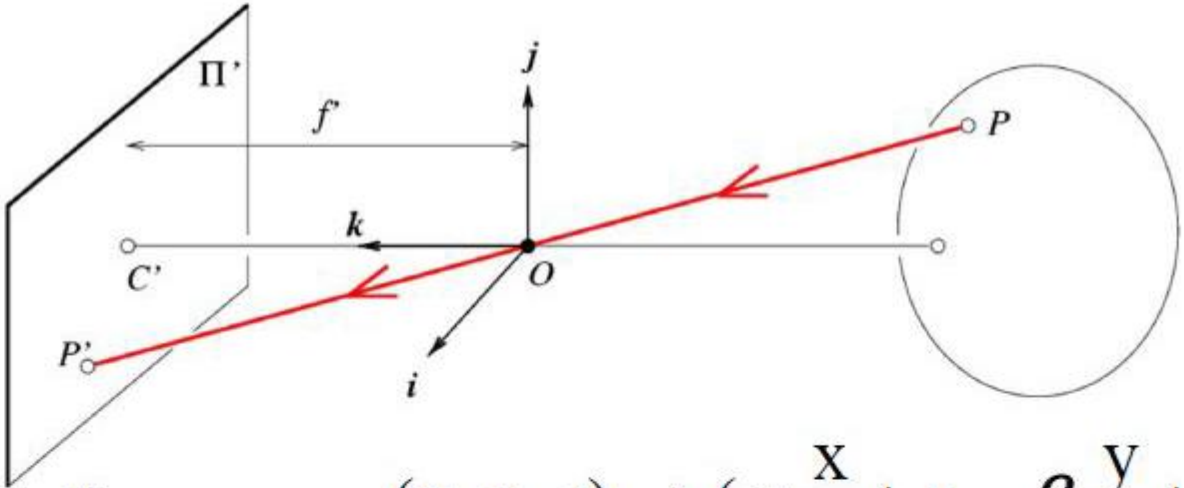


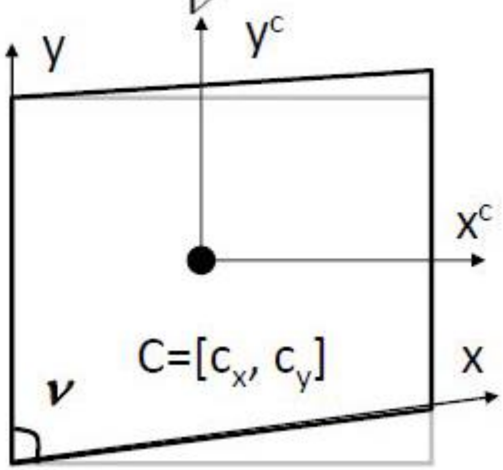


$$(x, y, z) \rightarrow (\boldsymbol{\alpha}\frac{x}{z} + c_x, \boldsymbol{\beta}\frac{y}{z} + c_y)$$

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

Camera matrix





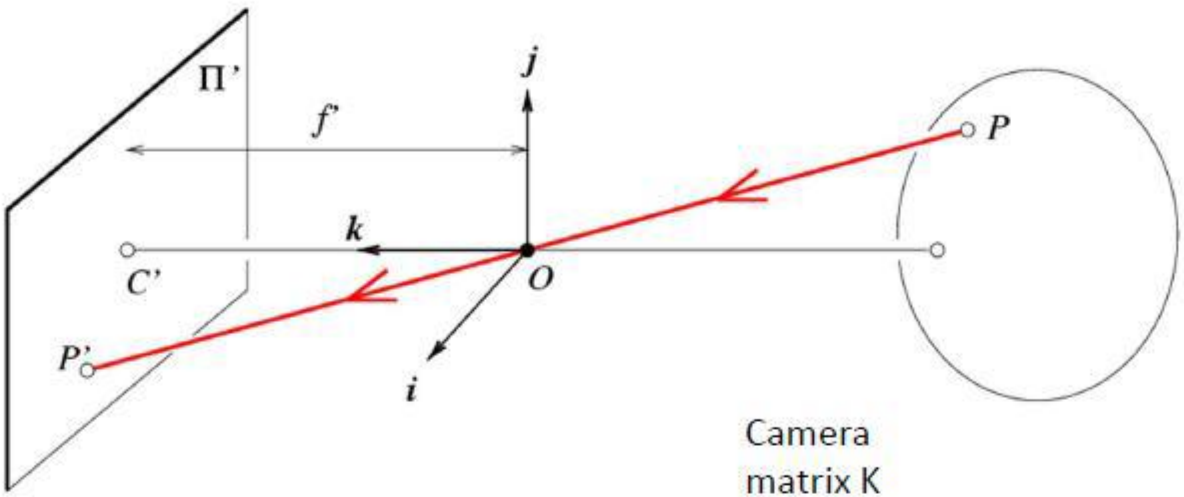
$$(x, y, z) \rightarrow (\boldsymbol{\alpha}\frac{x}{z} + c_x, \boldsymbol{\beta}\frac{y}{z} + c_y)$$

Skew parameter

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

Finally, the camera coordinate system may be skewed due to manufacturing error, so that angle θ between two image axes is not equal to 90º.

Camera matrix



$$P' = M\,P$$

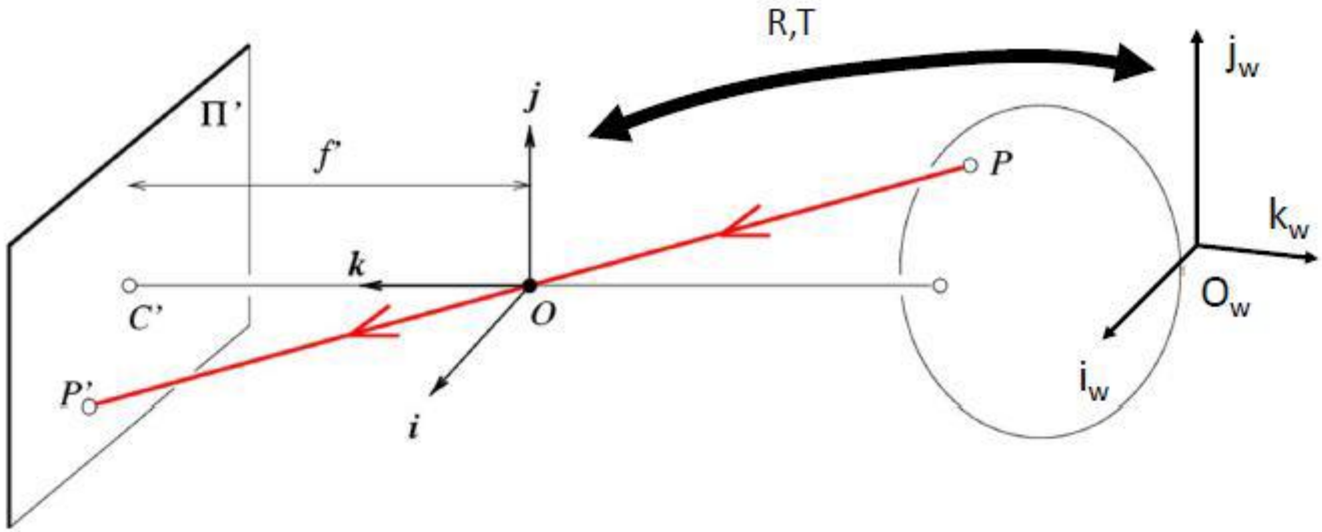
$$= K\begin{bmatrix} I & 0 \end{bmatrix} P$$

Camera matrix K

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

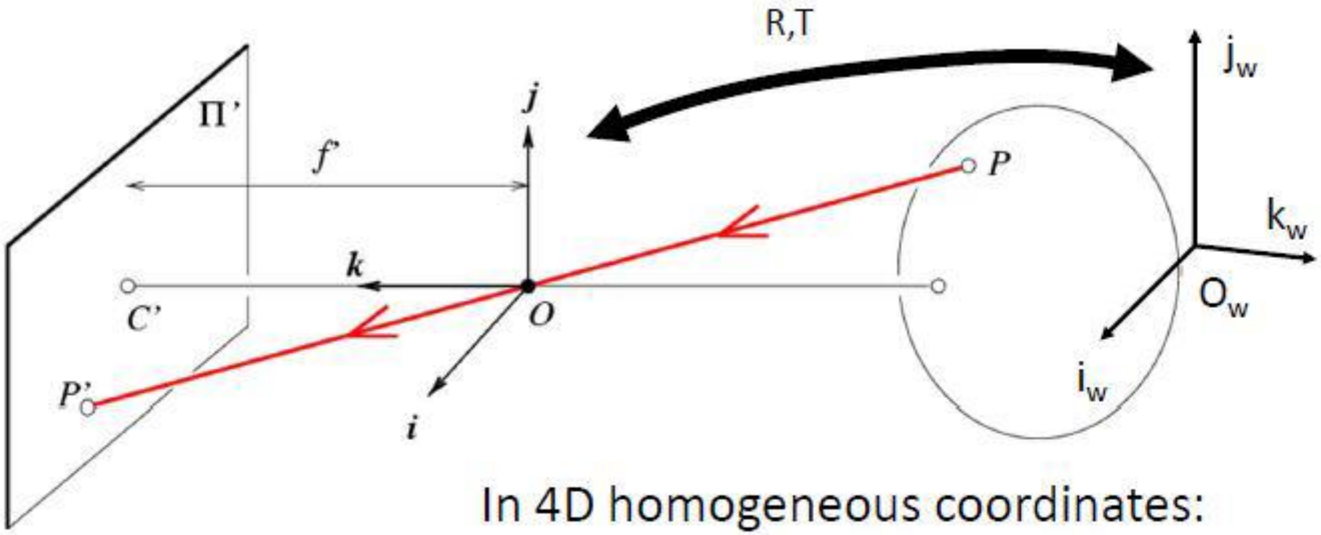
K has 5 degrees of freedom!

Camera & world reference system



- The mapping is defined within the camera reference system
- What if an object is represented in the world reference system?

Camera & world reference system



In 4D homogeneous coordinates:

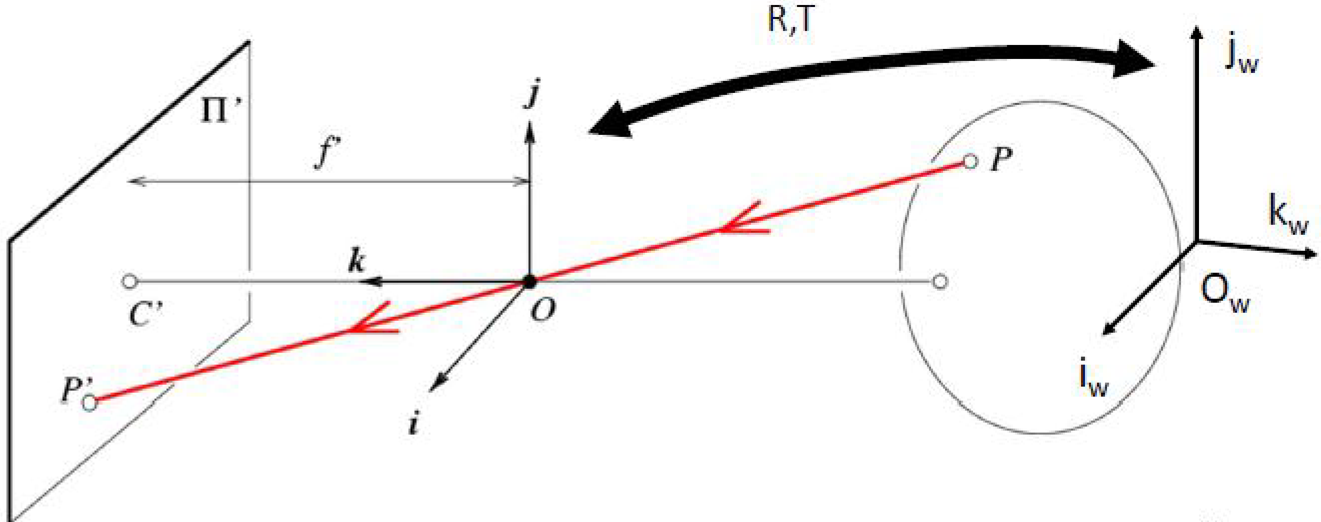
$$P = \begin{bmatrix} R & T \end{bmatrix} P_w$$

$$P' = M\,P_w = \boxed{K}\boxed{\begin{bmatrix} R & T \end{bmatrix}} P_w$$

Internal parameters

External parameters

Projective cameras



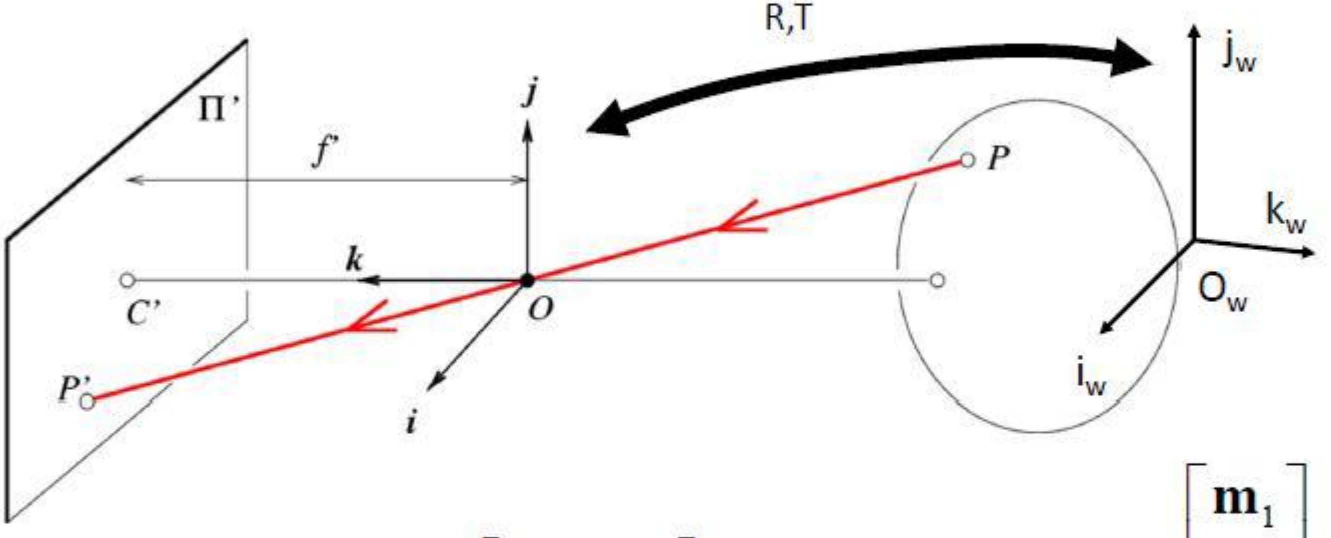
$$P'_{3\times1} = M\,P_w = K_{3\times3}\begin{bmatrix} R & T \end{bmatrix}_{3\times4} P_{w\,4\times1}$$

$$K = \begin{bmatrix} \alpha & s & c_x \\ 0 & \beta & c_y \\ 0 & 0 & 1 \end{bmatrix}$$

How many degrees of freedom?

5 + 3 + 3 =11!

Projective cameras



$$P'_{3\times1} = M\ P_w = K_{3\times3}\begin{bmatrix} R & T \end{bmatrix}_{3\times4}\ P_{w4\times1} \qquad M = \begin{bmatrix} \mathbf{m}_1 \\ \mathbf{m}_2 \\ \mathbf{m}_3 \end{bmatrix}$$

$$(x, y, z)_w \rightarrow \left(\frac{\mathbf{m}_1 P_w}{\mathbf{m}_3 P_w}, \frac{\mathbf{m}_2 P_w}{\mathbf{m}_3 P_w}\right)$$

M is defined up to scale!
Multiplying M by a scalar
won't change the image

Theorem (Faugeras, 1993)

$$M = K\begin{bmatrix} R & T \end{bmatrix} = \begin{bmatrix} KR & KT \end{bmatrix} = \begin{bmatrix} A & b \end{bmatrix}$$

$$A = \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \mathbf{a}_3 \end{bmatrix}$$

$$K = \begin{bmatrix} \alpha & s & c_x \\ 0 & \beta & c_y \\ 0 & 0 & 1 \end{bmatrix}$$

$$\boldsymbol{\alpha} = \mathrm{f\ k};$$
$$\boldsymbol{\beta} = \mathrm{f\ l}$$

Let $\mathcal{M} = (\mathcal{A} \quad \boldsymbol{b})$ be a 3×4 matrix and let $\boldsymbol{a}_i^T$ $(i = 1, 2, 3)$ denote the rows of the matrix $\mathcal{A}$ formed by the three leftmost columns of $\mathcal{M}$.

- A necessary and sufficient condition for $\mathcal{M}$ to be a perspective projection matrix is that $\mathrm{Det}(\mathcal{A}) \neq 0$.
- A necessary and sufficient condition for $\mathcal{M}$ to be a zero-skew perspective projection matrix is that $\mathrm{Det}(\mathcal{A}) \neq 0$ and
$$(\boldsymbol{a}_1 \times \boldsymbol{a}_3) \cdot (\boldsymbol{a}_2 \times \boldsymbol{a}_3) = 0.$$
- A necessary and sufficient condition for $\mathcal{M}$ to be a perspective projection matrix with zero skew and unit aspect-ratio is that $\mathrm{Det}(\mathcal{A}) \neq 0$ and
$$\begin{cases} (\boldsymbol{a}_1 \times \boldsymbol{a}_3) \cdot (\boldsymbol{a}_2 \times \boldsymbol{a}_3) = 0, \\ (\boldsymbol{a}_1 \times \boldsymbol{a}_3) \cdot (\boldsymbol{a}_1 \times \boldsymbol{a}_3) = (\boldsymbol{a}_2 \times \boldsymbol{a}_3) \cdot (\boldsymbol{a}_2 \times \boldsymbol{a}_3). \end{cases}$$