



Part II: Models

Nicola Conci
nicola.conci@unitn.it

Why do we need models?

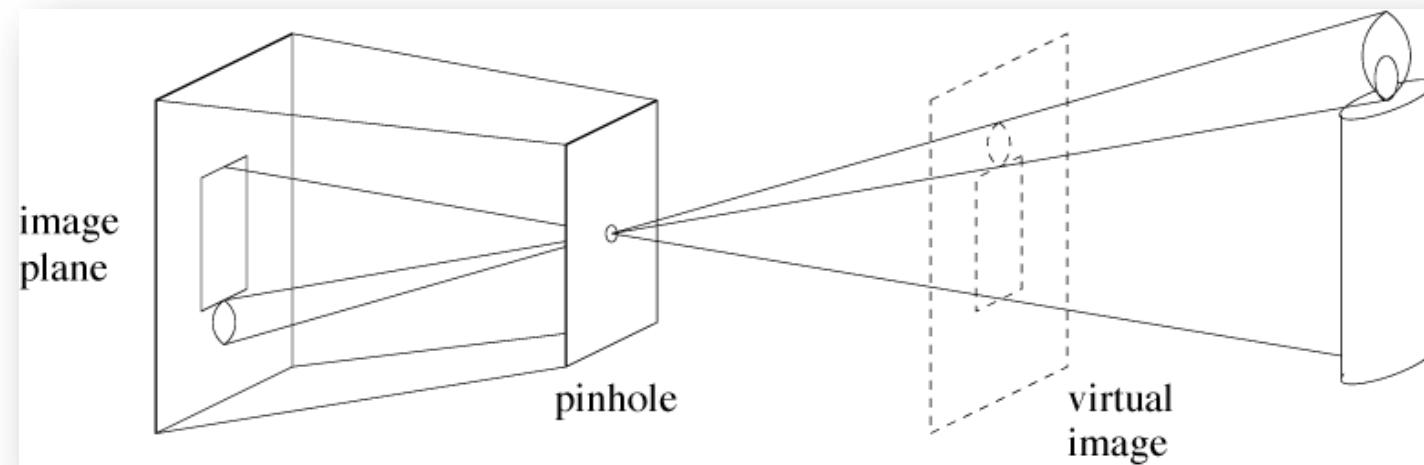


- They represent a good approximation of the real world
- They allow describing events through a parametric representation
- Parameters can be extracted and used for processing

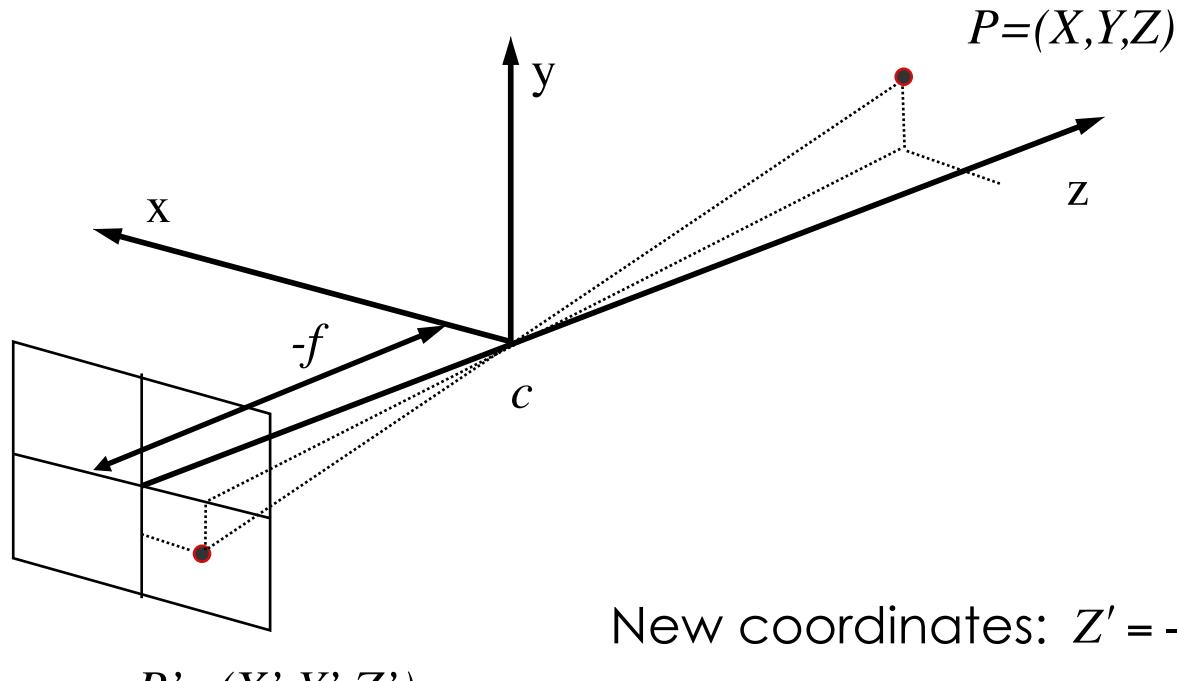
The pinhole camera model



- It's an abstract model
- It's the most common model



The pinhole camera model



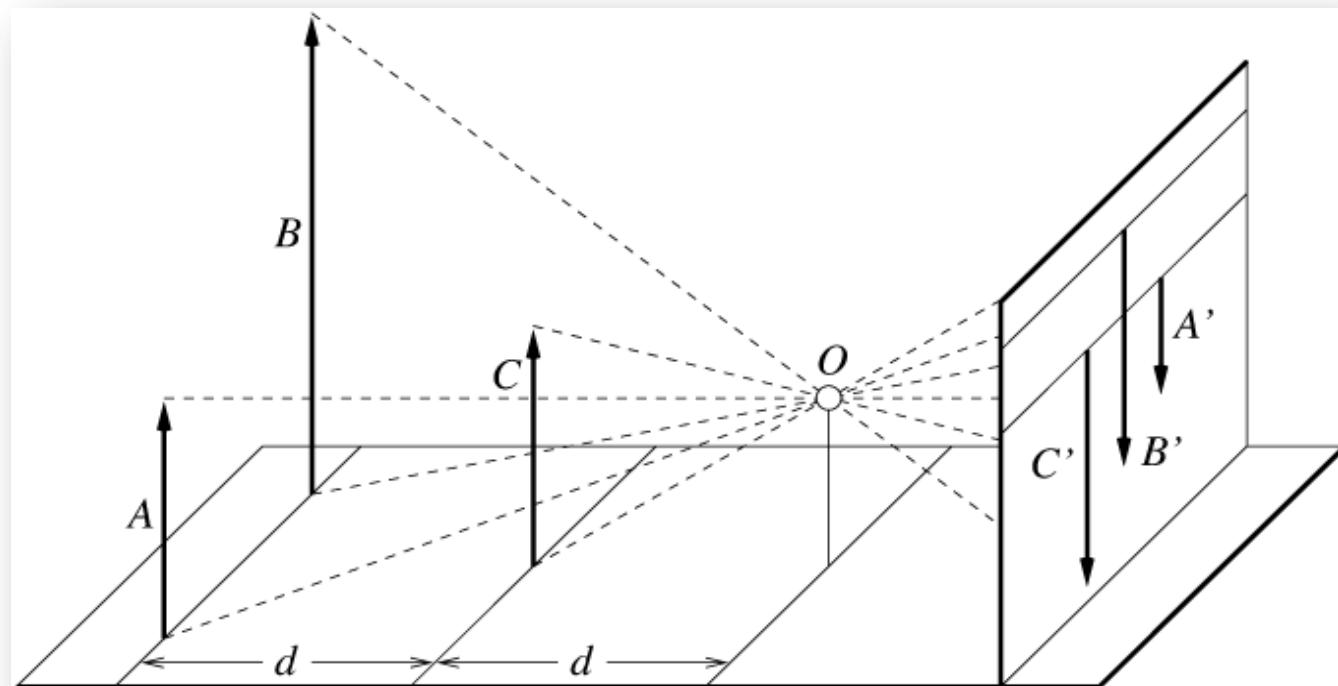
New coordinates: $Z' = -f$, $X' = -f \frac{X}{Z}$, $Y' = -f \frac{Y}{Z}$
 $x = -X'$, $y = -Y'$

From the camera: $(X, Y, Z) \rightarrow (x, y, f) = (f \frac{X}{Z}, f \frac{Y}{Z}, f)$

Features of the pinhole model



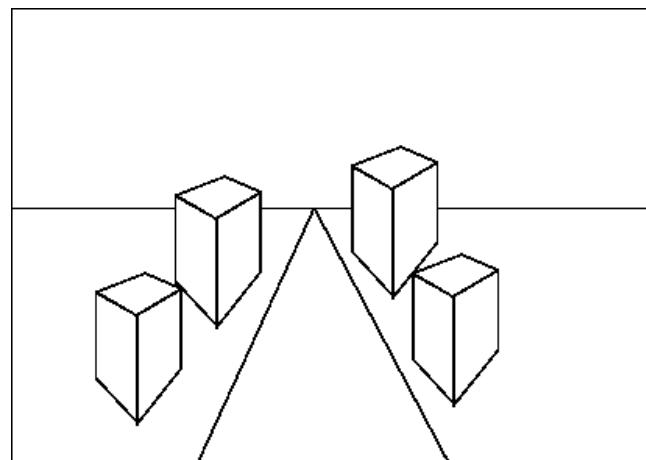
- If the object is far, it appears smaller



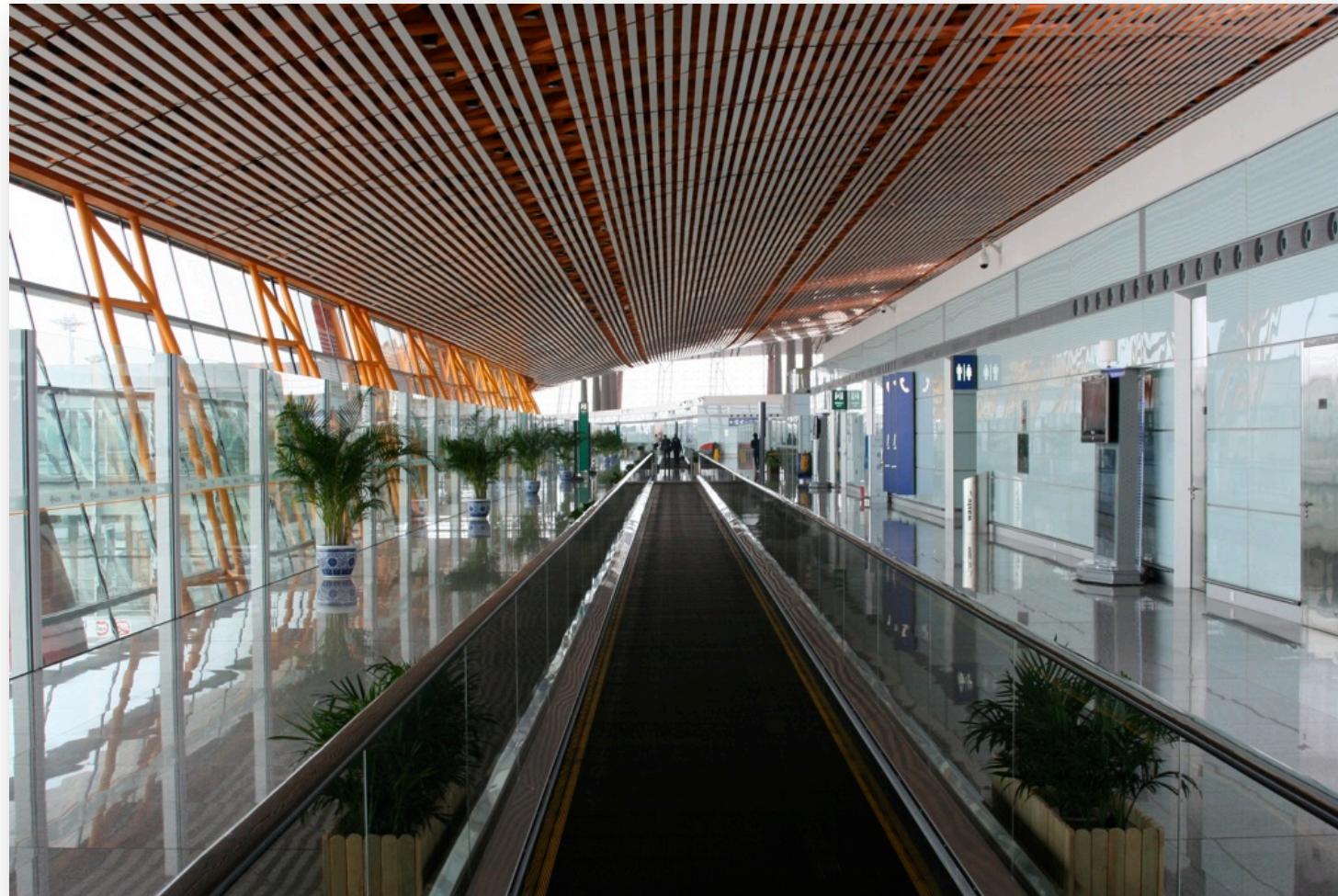
Features of the pinhole model



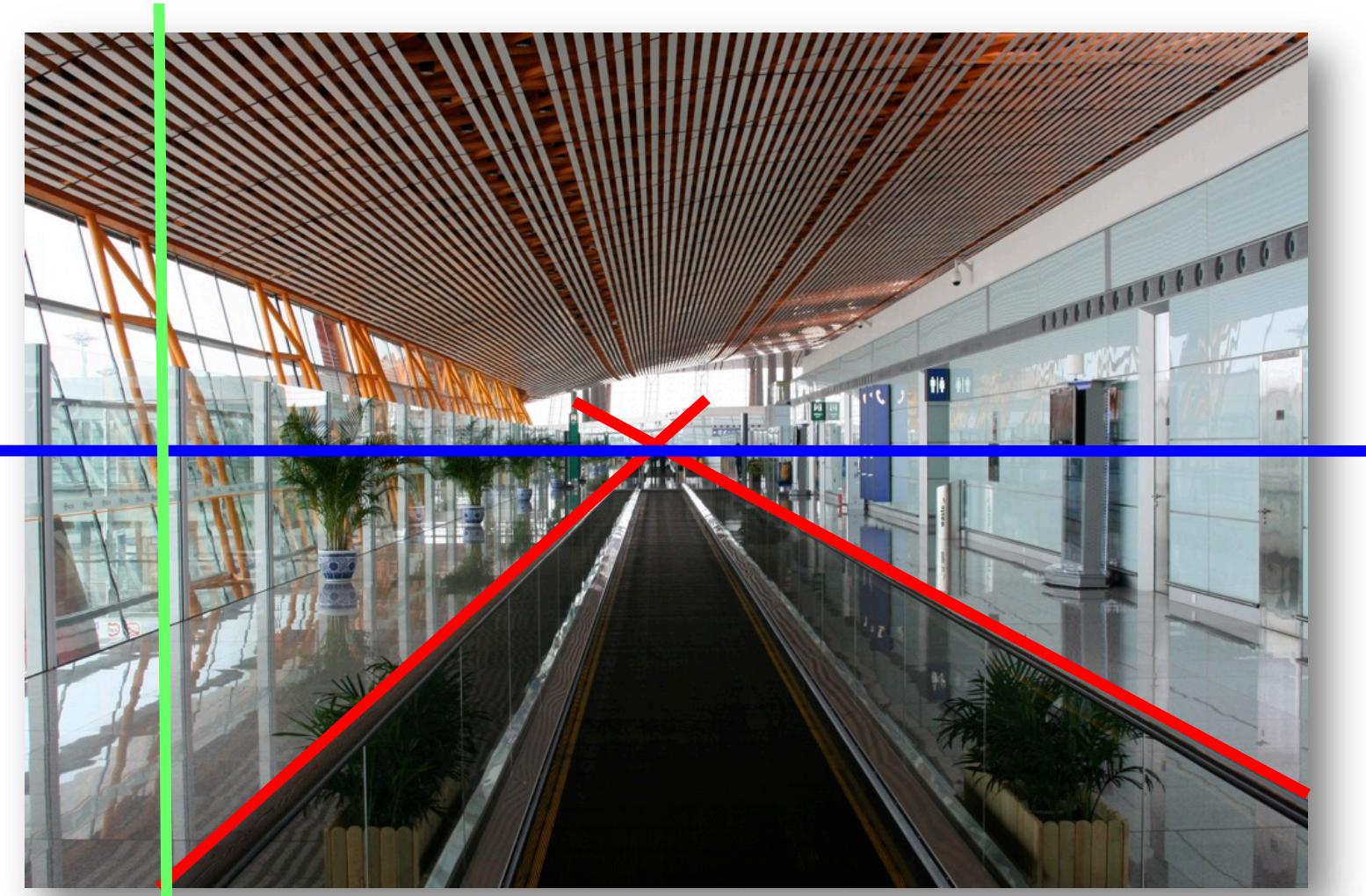
- Parallel lines converge to a single point
- Parallel lines on the same plane lead to collinear vanishing points.
- The line is called the *horizon* for that plane
- Vertical lines are perpendicular to the horizon



Vanishing Points and Horizon



Vanishing Points and Horizon



Projections



- Capturing devices record 2D projections of 3D (+ time) scenes:

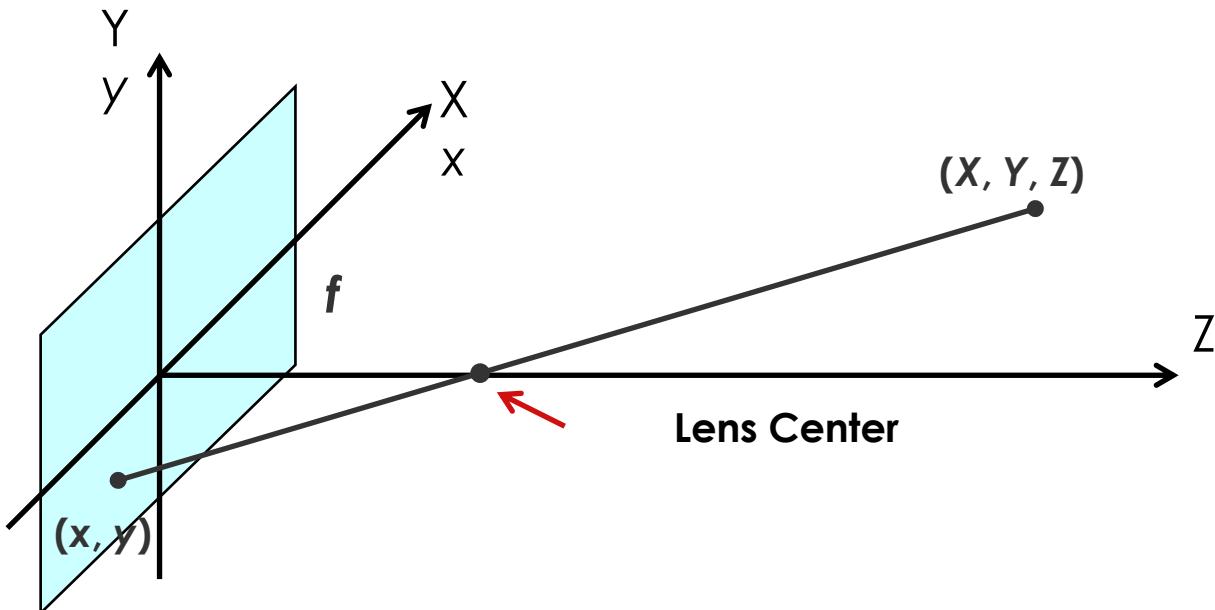
$$f : \Re^4 \rightarrow \Re^3$$
$$(X, Y, Z, t) \rightarrow (x, y, t)$$

- (X, Y, Z) , (x, y) , and t , are continuous variables.
- Two types of projection:
 - Perspective (central)
 - Orthographic (parallel)

Perspective projection



- Using the pinhole model, all rays pass through the center of projection, which corresponds to the lens.
- In the picture, the center is between the object and the image plane
- The image plane corresponds to the (X, Y) plane of the 3-D space.



Perspective projection



- From the pairs of triangles:

$$(x,0,0), (0,0,f), (0,0,0); \text{ and } (X,0,Z), (0,0,f), (0,0,Z)$$

$$(y,0,0), (0,0,f), (0,0,0); \text{ and } (Y,0,Z), (0,0,f), (0,0,Z)$$

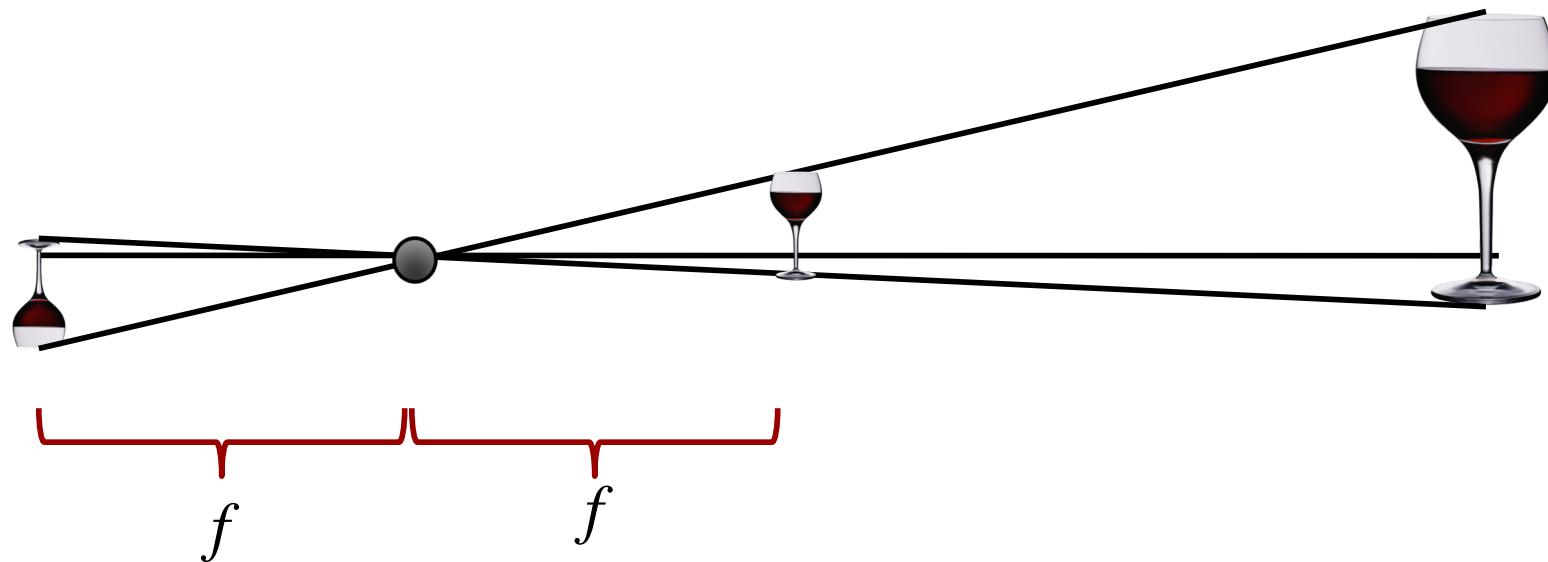
- We obtain the following formulation:

$$\begin{aligned}\frac{x}{f} &= \frac{X}{Z-f} \quad \text{and} \quad \frac{y}{f} = \frac{Y}{Z-f} \Rightarrow \\ \Rightarrow x &= \frac{fX}{Z-f} \quad \text{and} \quad y = \frac{fY}{Z-f}\end{aligned}$$

Perspective projection and the pinhole model

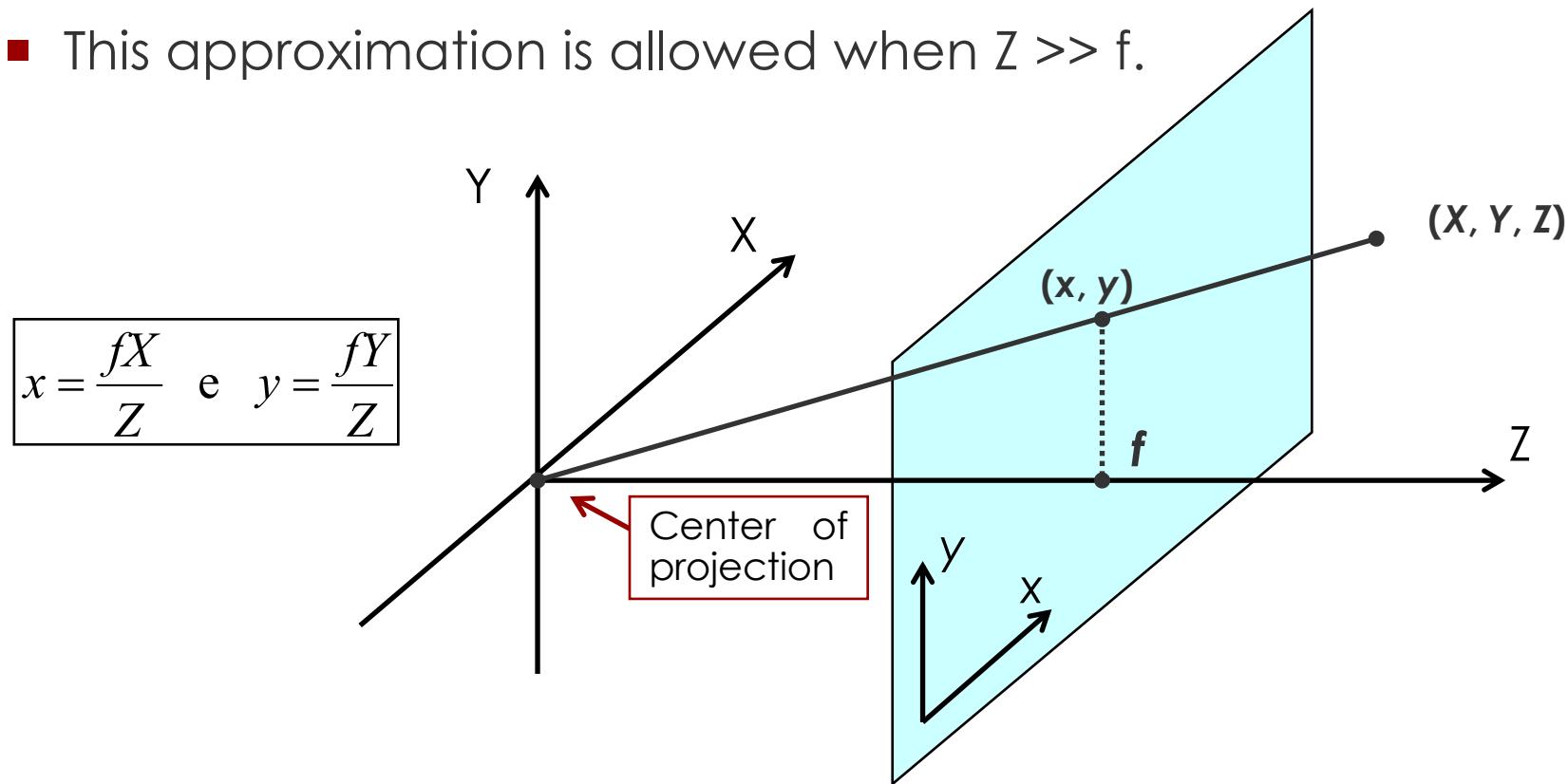


- For simplicity we usually consider the image plane on the same side of the “real world”, to avoid the picture flip



Simplified perspective projection

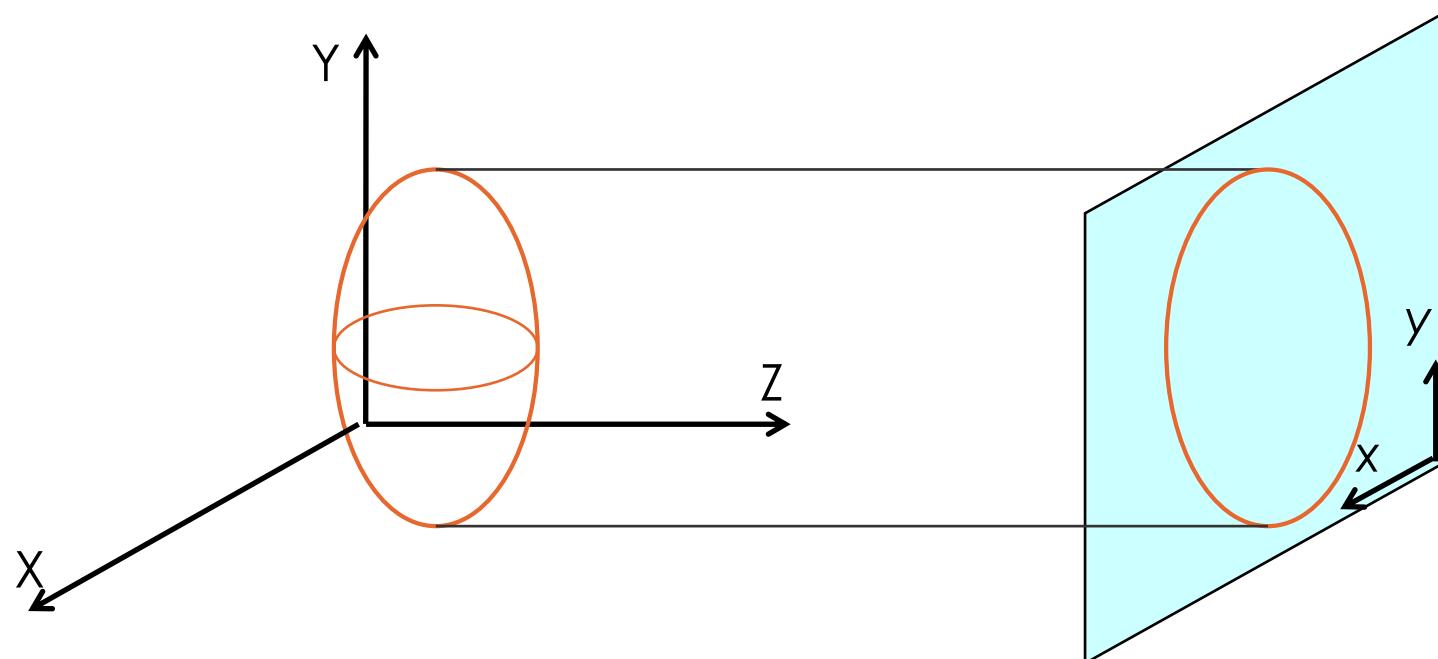
- The center of projection corresponds to the origin of the 3D space.
- The plane (X,Y) is parallel to (x,y)
- This approximation is allowed when $Z \gg f$.



Orthographic projection



- It is assumed that all rays originated from the 3D object, and from the scene in general, are parallel among each other.
- In the drawing, the image plane is parallel to (X, Y)



Orthographic projection



- Assuming that the image plane is parallel to (X,Y), the orthographic projection can be simply described in Cartesian coordinates:

$$x = X \quad \text{and} \quad y = Y$$

- Or in form of a matrix:

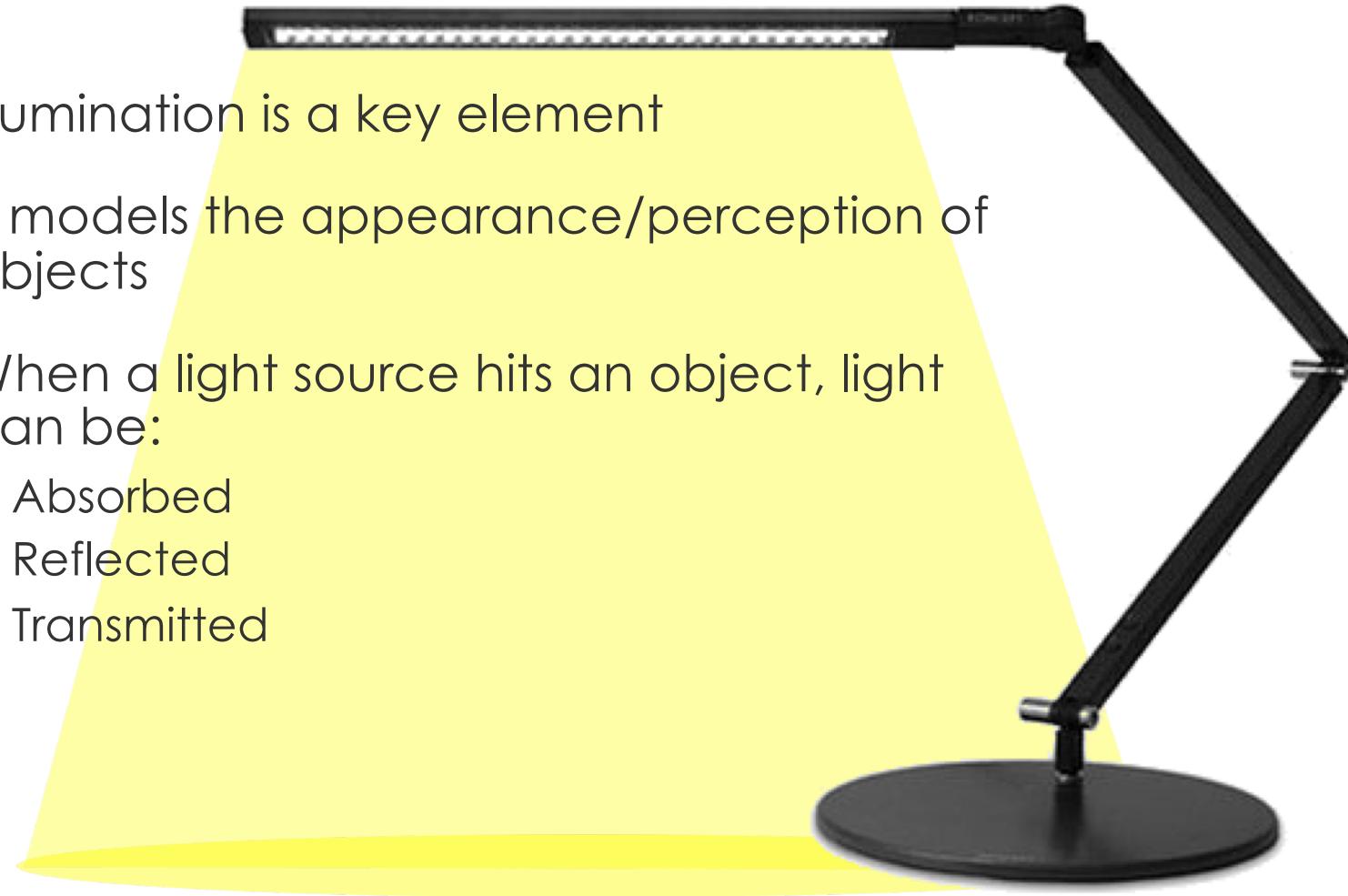
$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

- Notes:
 - The **distance** of the object from the camera **does not affect** the intensity of the image projected onto the 2D plane.
 - It is a **good approximation** when the distance of the object is much bigger than the **depth** of the object itself

Illumination models



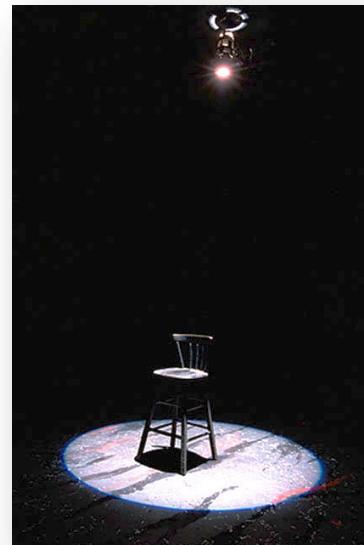
- Illumination is a key element
- It models the appearance/perception of objects
- When a light source hits an object, light can be:
 - Absorbed
 - Reflected
 - Transmitted



Illumination models



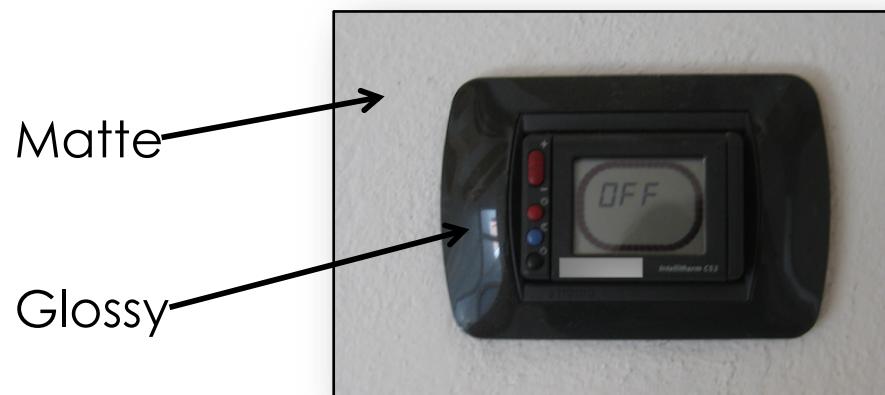
- It's complex to model
 - We perceive objects because they reflect light in specific wavelengths
 - Reflection can be:
 - Specular (more energy is concentrated in the light source direction)
 - Diffuse (constant in all directions, and the position of the observer is irrelevant)



Illumination



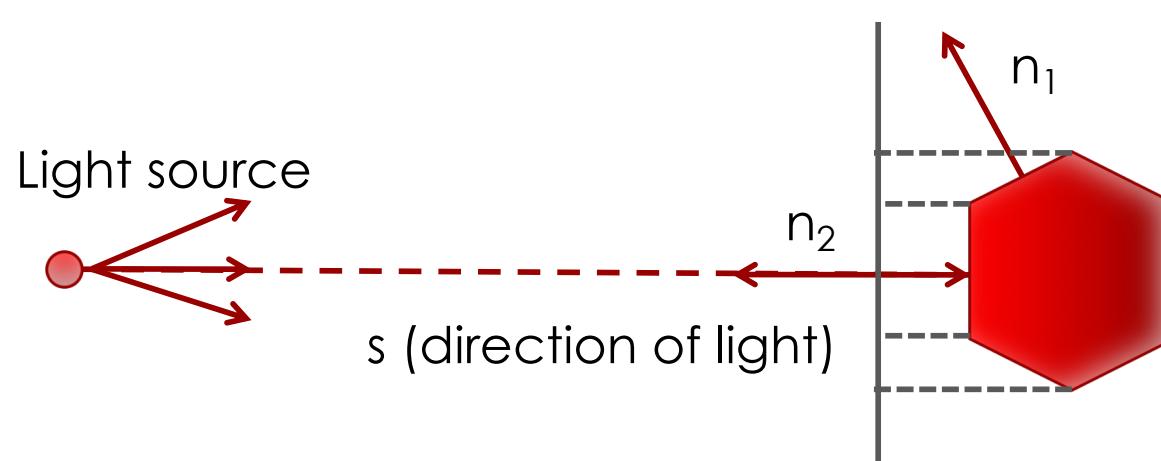
- Surfaces vary in *specularity*
- Some of them (matte) reflect light uniformly in all directions
- Glossy objects reflect light in specific directions
- It also depends on the distance and the inclination of the light source



Illumination from one light source



- Problem: determine how the surface is irradiated by the light source
- Assumption: light is far, we can assume all rays can be represented by a single unit vector s (orthographic projection)
- For each surface element (btw dashed lines) the light is irradiated considering the cosine of the angle between the surface normal and the light direction
- $i = n \cdot s$



Lambertian surface



- Model for diffuse reflection
- The specular component is neglected
- The luminance of the surface is the same regardless of the viewing angle
- Possible when the surface is rough enough w.r.t. the light wavelength
- → each surface element reflects light evenly in all directions

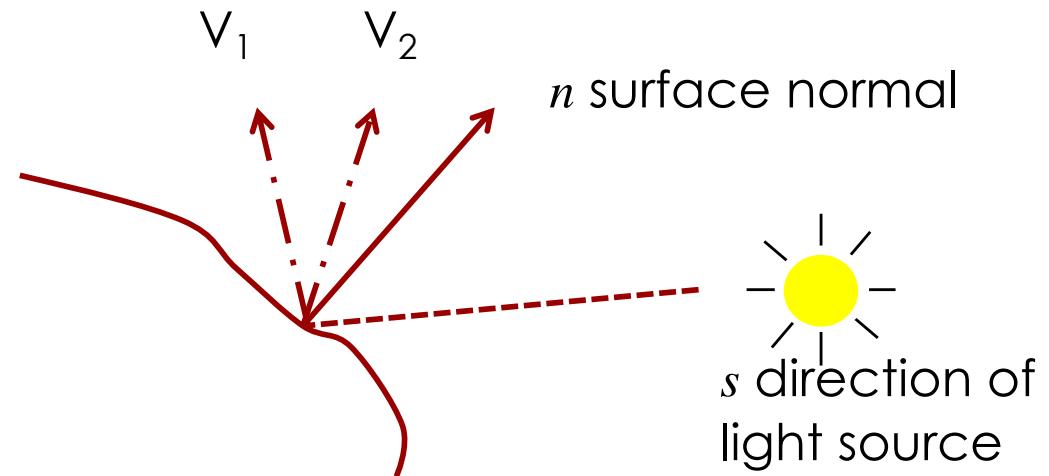


Reflectance model



$$I = \rho n \cdot s = |n| |s| \cos \alpha$$

- n, s unit vectors
- ρ is the albedo
 - Ratio of the reflected illumination to the total illumination
 - In general intrinsic property of the surface
 - Not true in all cases, since some surfaces may reflect light differently depending on the view angle
 - An element is not visible if $n \cdot v < 0$, with v the angle of the viewer
- The pixel in image $I(r,c)$ depends on:
 - The light source direction s
 - The normal of the element direction n



Remarks: Cameras and Lenses

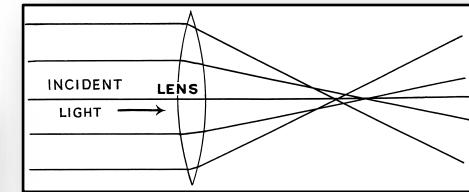
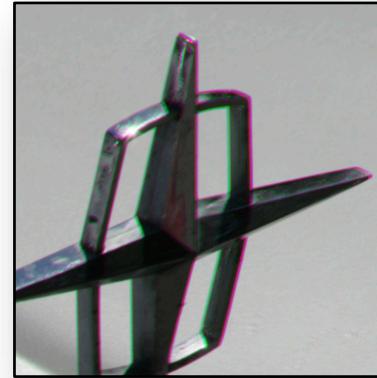


- Cameras are equipped with lenses (no pinhole in real life!)
- This means that the object is **on focus** if the distance from the center of the camera and the image plane obeys to the *thin lens* equation
- If not, we have aberrations
- We typically assume the object is on focus, but let's see what this means first!

Typical issues with lenses



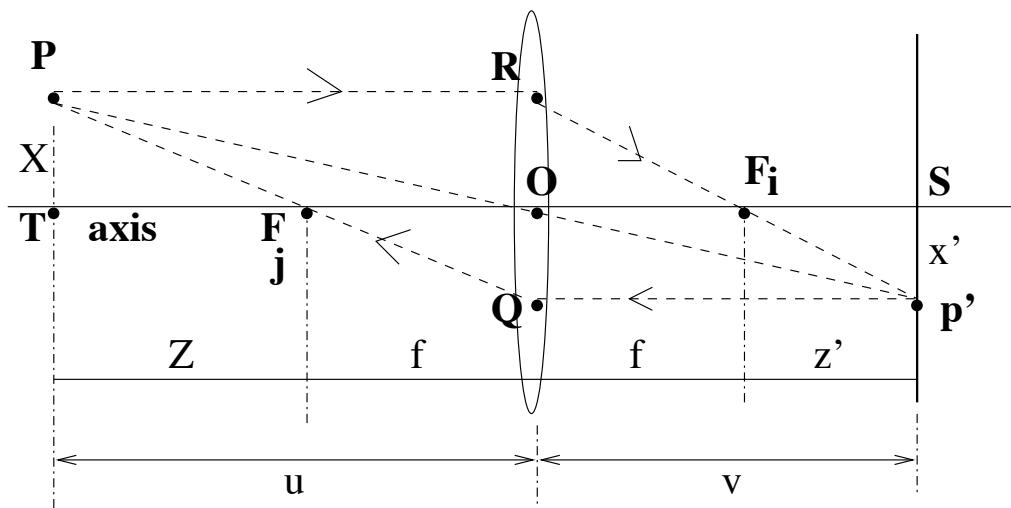
- Spherical aberration (causes blurring)
- Chromatic aberrations
- Vignetting (two lenses, dirt, cheap cameras)
- Barrel distortion (short focal length – wide angle lenses)



Using lenses



- Pinhole is an abstract model
- Thin lens



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

$$\frac{X}{Z+f} = \frac{x'}{z'+f}$$

Substituting X: $f^2 = Zz'$

where: $Z = u - f$, $z' = v - f$

leads to: $uv = f(u+v)$

dividing by (uvf): $\frac{1}{f} = \frac{1}{u} + \frac{1}{v}$

Notice: $u \rightarrow \infty \Rightarrow \frac{1}{f} = \frac{1}{v}$

Drawing taken from Shapiro/Stockman

CV and MM 2019-2020

FOCUS

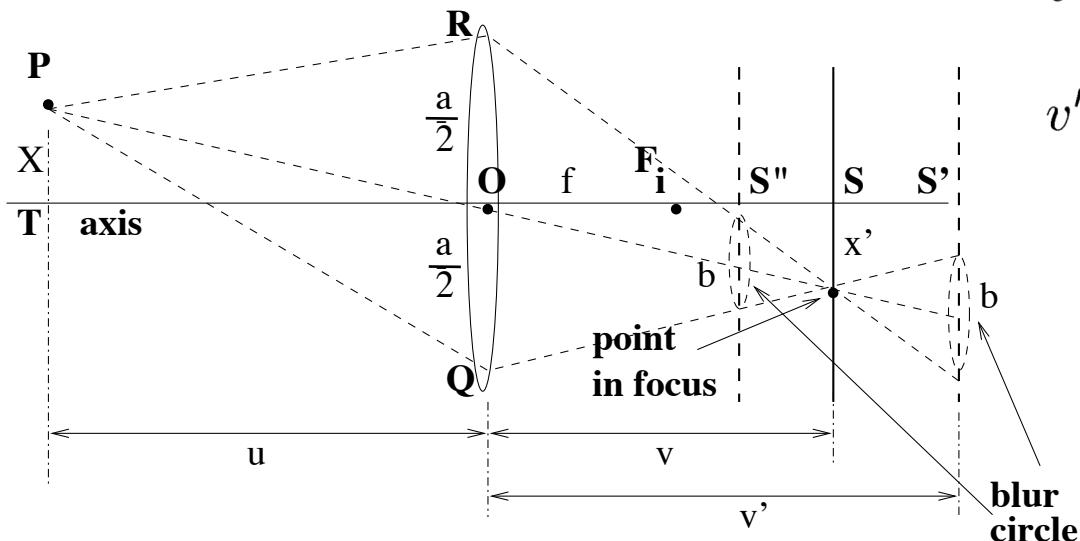


- If we move the image plane, point p' is out of focus $\rightarrow v$ changes $\rightarrow v'$
 - If P is moved $\rightarrow u$ changes $\rightarrow u'$
-
- Result: the image is blurred on the image plane.
 - Instead of a point I see a circle!

Focus



- Considering that the blur can be acceptable if the circle is within b , what is the range of u ?



$$v' = \frac{a+b}{a}v : \text{in case } v' > v$$

$$v' = \frac{a-b}{a}v : \text{in case } v' < v$$

$$u_n = \frac{u(a+b)}{bu} - \frac{a}{a+bu}$$

$$u_r = \frac{u(a-b)}{a - \frac{bu}{f}}$$



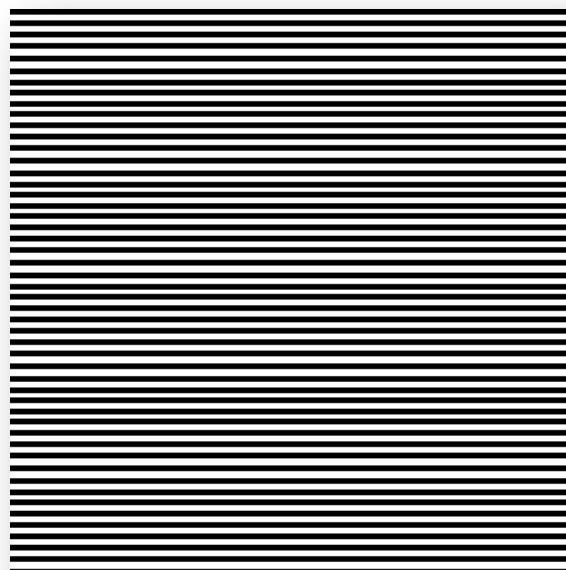
Notice that:

- In general $u > f \Rightarrow u_n < u$
- If f becomes smaller, u_n is closer to the camera
- $u_r > u$
- If f becomes smaller, u_r is farther from the camera
- If $u \rightarrow \infty$ rays are parallel and converge to the camera center
- Depth of field: difference between the far and near planes limiting b

Resolution and blur



- A camera with a NxM CCD can detect $N/2$ horizontal lines (one pixel left between two lines)
- If blur is larger than 1 pixel, the image will be grey
- Depending on the problem you have to solve, the appropriate lens must be chosen



Resolving power



- $R_p = 1/(2\Delta)$ [lines/mm]
- Δ = pixel spacing (inches or mm)
- Example:
 - $CCD_{size} = 10\text{mm square, } 500 \times 500\text{px}$
 - $R_p = 1/[2 \times (2 \times 10^{-2})] = 25$ [lines/mm]
- Example:
 - Spacing of cones in the fovea = $\Delta = 10^{-4}$ inches
 - $R_p = 5 \times 10^3$
 - $f = 20\text{mm}=0.8\text{in.}$ (diameter of the eye)
 - Subtended angle $\theta = 2.5 \times 10^{-4}$ rad = 1 min of arc = $1/60$ deg = 0.016 deg
 - Human eye can see a pencil stroke 0.5mm wide at 2m distance.

Autofocus



- What is it?
 - Capability of focusing a specific portion of the image
- How?
 - Active
 - Passive
 - Combination of the two

Active autofocus



- Mainly used for point-and-shoot cameras
- Polaroid in 1986: SONAR
- Today: Infrared, up to 6m
- Example:
 - an infrared signal is sent and the time between sending and receiving is computed
- Issues
 - Obstacles
 - Glossy and bright surfaces

Passive autofocus



- More expensive SLR (single-lens reflex) cameras
- Distance computed using image analysis
 - Take a strip of pixels and analyze the distribution
 - If values are too similar object is out of focus
 - If contrast is high, object is on focus
 - Problems with flat surfaces
 - Good cameras compute the metric on the vertical and horizontal axes

What camera do I have?



- Is there an infrared emitter?
- Otherwise:
 - Go outdoors and point at an area of the sky with no clouds (or in general a flat surface).
 - Press the shutter button halfway down. If you get a "focus okay" indication, it's an active autofocus system.
 - If you get a "focus not okay" indication, it's a passive autofocus system. The CCD cannot find any contrast in a blue sky, so it gives up.

The whole system

