Image Formation

Dr. Gerhard Roth

COMP 4102A

Winter 2015

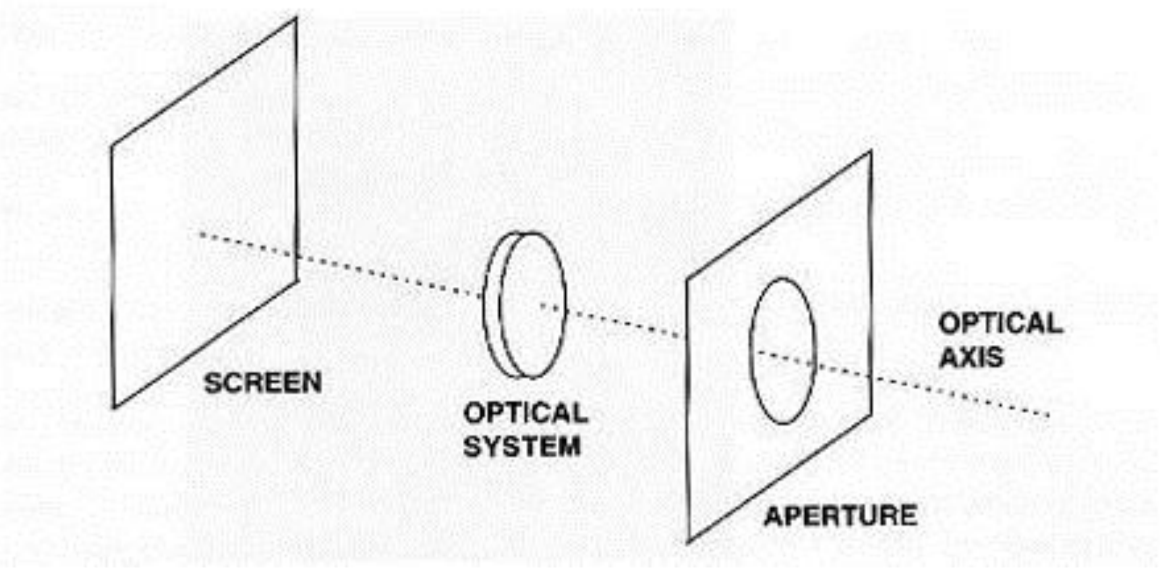
Version 3

Image Formation

- Two type of images
  - Intensity– image encodes light intensities (passive sensor)
  - Range (depth)– image encodes shape and distance
    - Created from processing passive images or by an active sensor
- Intensity image is a function of three things
- Optical parameters of the lens
  - Lens type, focal length, field of view, angular apertures
- Photogrammetric (Radiometric) parameters
  - Type, direction and intensity of the illumination
  - Reflectance properties of the viewed surface
  - Characteristics of the image sensor
- Geometric parameters
  - Type of projection, position and orientation of camera

Elements of a real imaging device

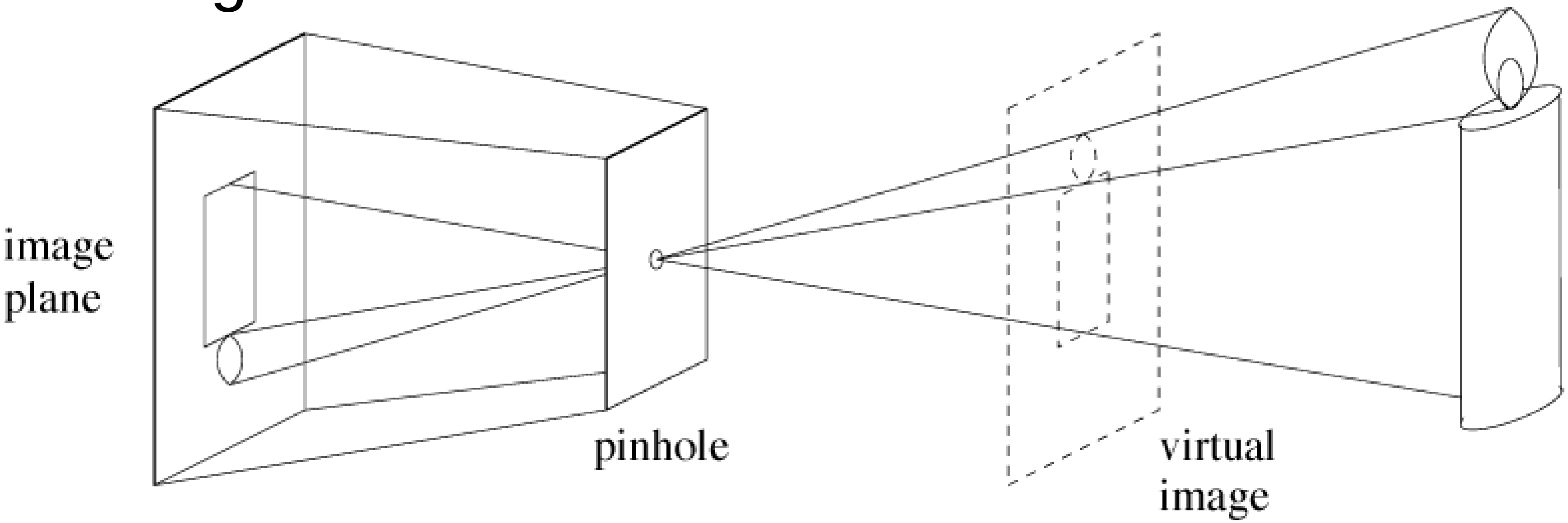
- Light rays coming from outside world and falling on the photoreceptors in the retina.
- Aperture lets in light and size can vary



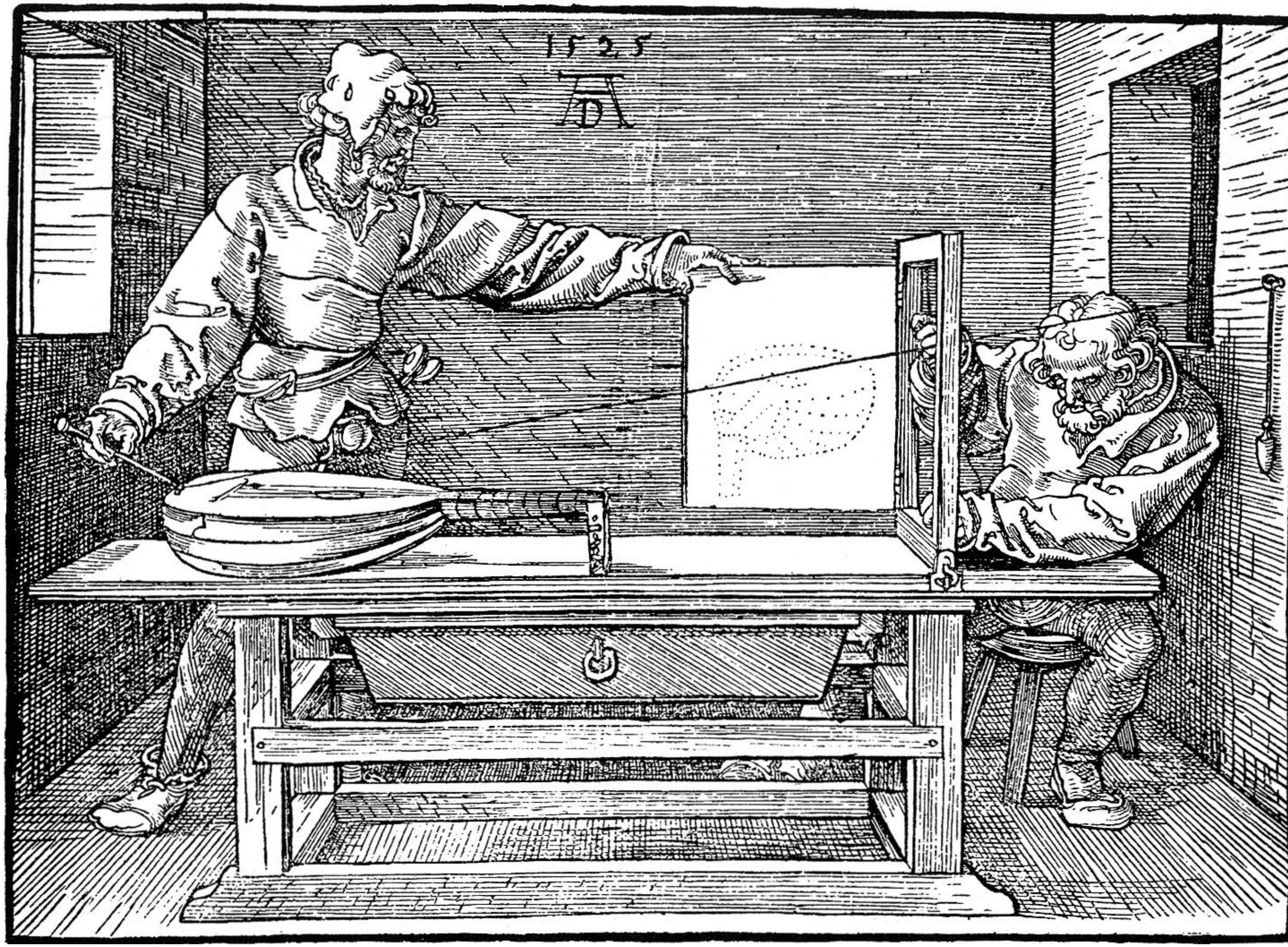
Screen represents any sensor that can capture light such as a photographic plate or a film negative or an electronic sensor, a 2d array of pixels.
Aperture usually opened for a small amount of time.

Pinhole camera

- Pinhole is the aperture in this case
- Changing pinhole size changes amount of light that is let in

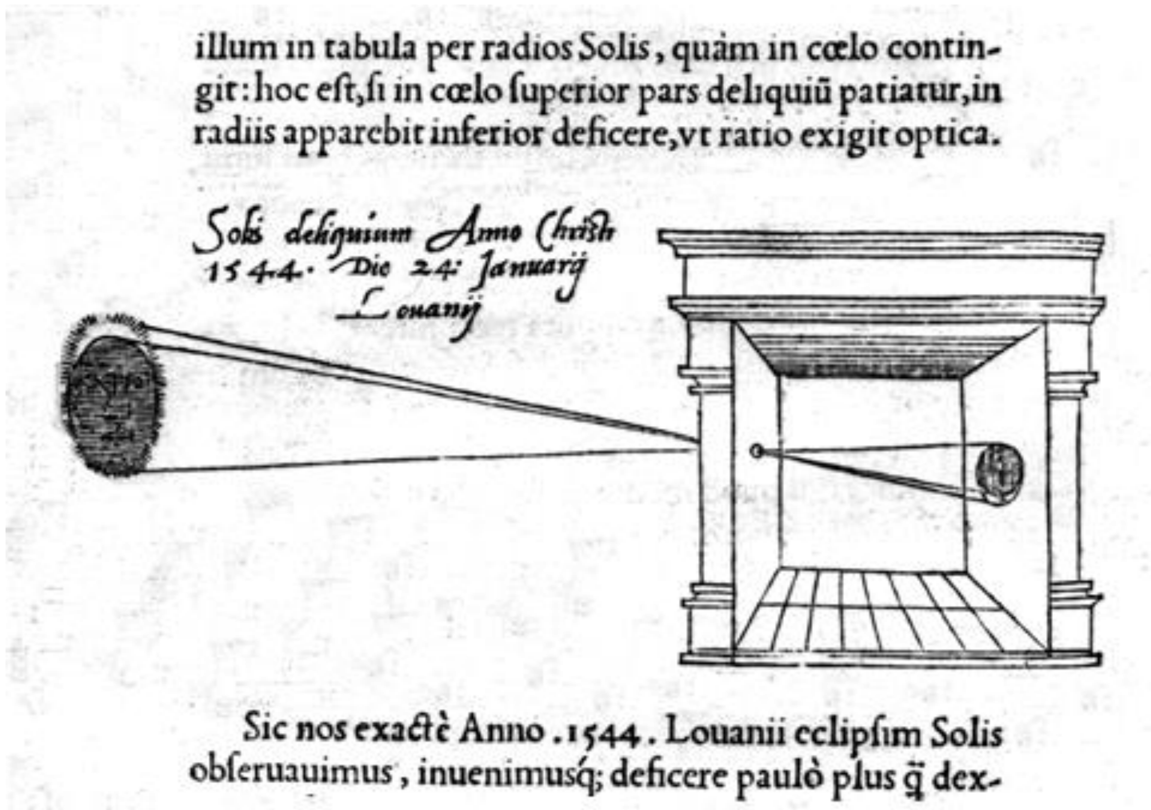


Perspective Projection



Draughtsman Drawing a Lute, Albrecht Dürer, 1525

Camera Obscura



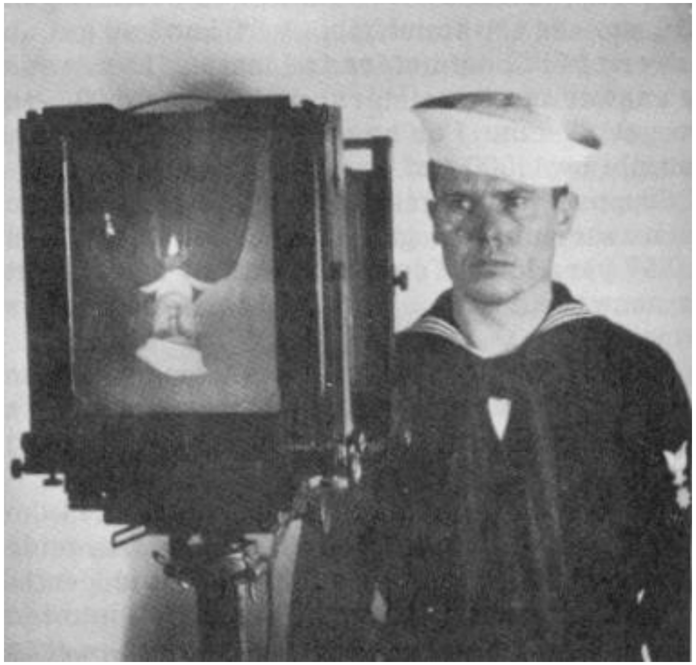
Camera Obscura, Reinerus Gemma Frisius, 1544

Camera Obscura: Latin ‘dark chamber’

Camera Obscura

Contemporary artist Madison Cawein rented studio space in an old factory building where many of the windows were boarded up or painted over. A random small hole in one of those windows turned one room into a camera obscura.

Photographic Camera

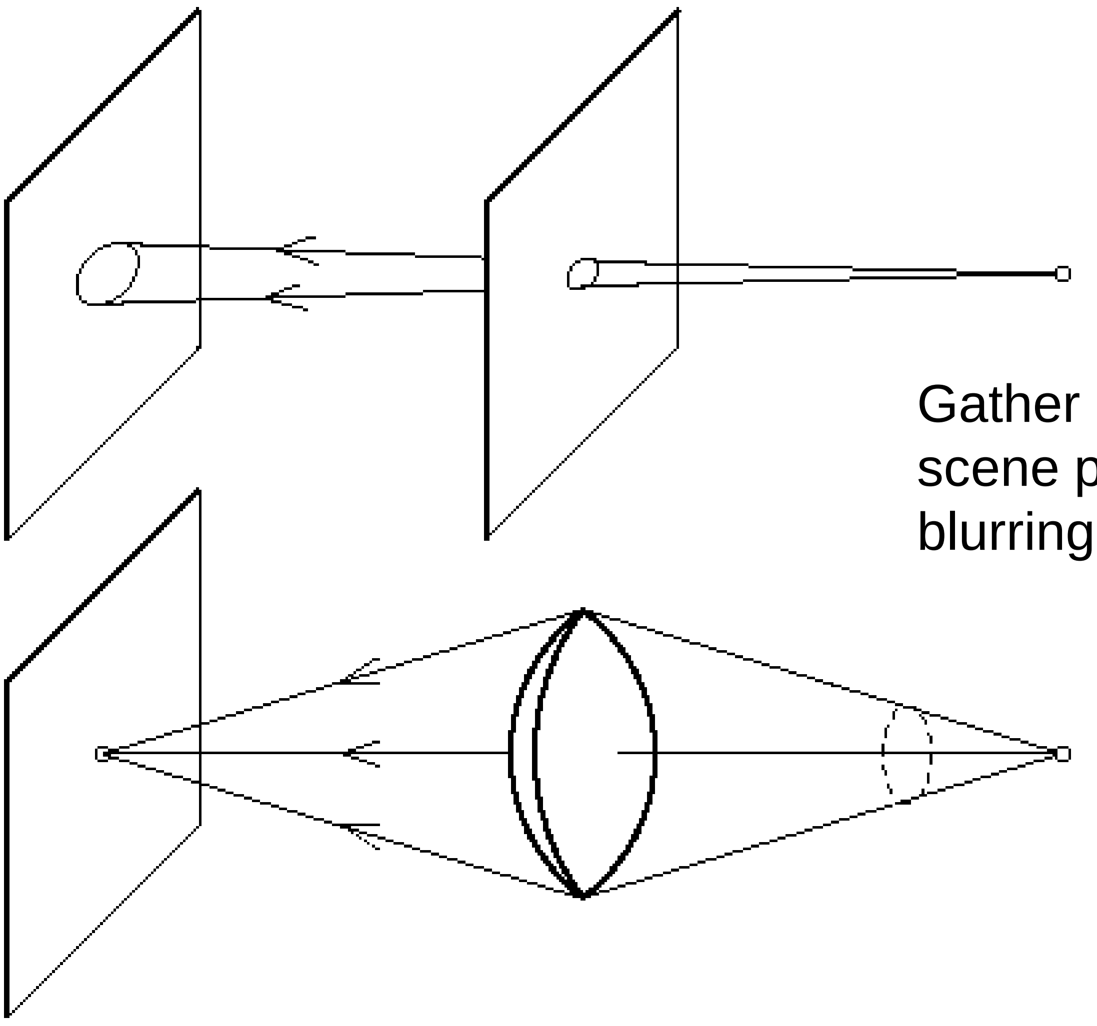


Photographic camera: Joseph Nicéphore Niepce, 1816

First Photograph

First photograph on record, *la table servie*, obtained by Niepce in 1822.

Why Lenses?



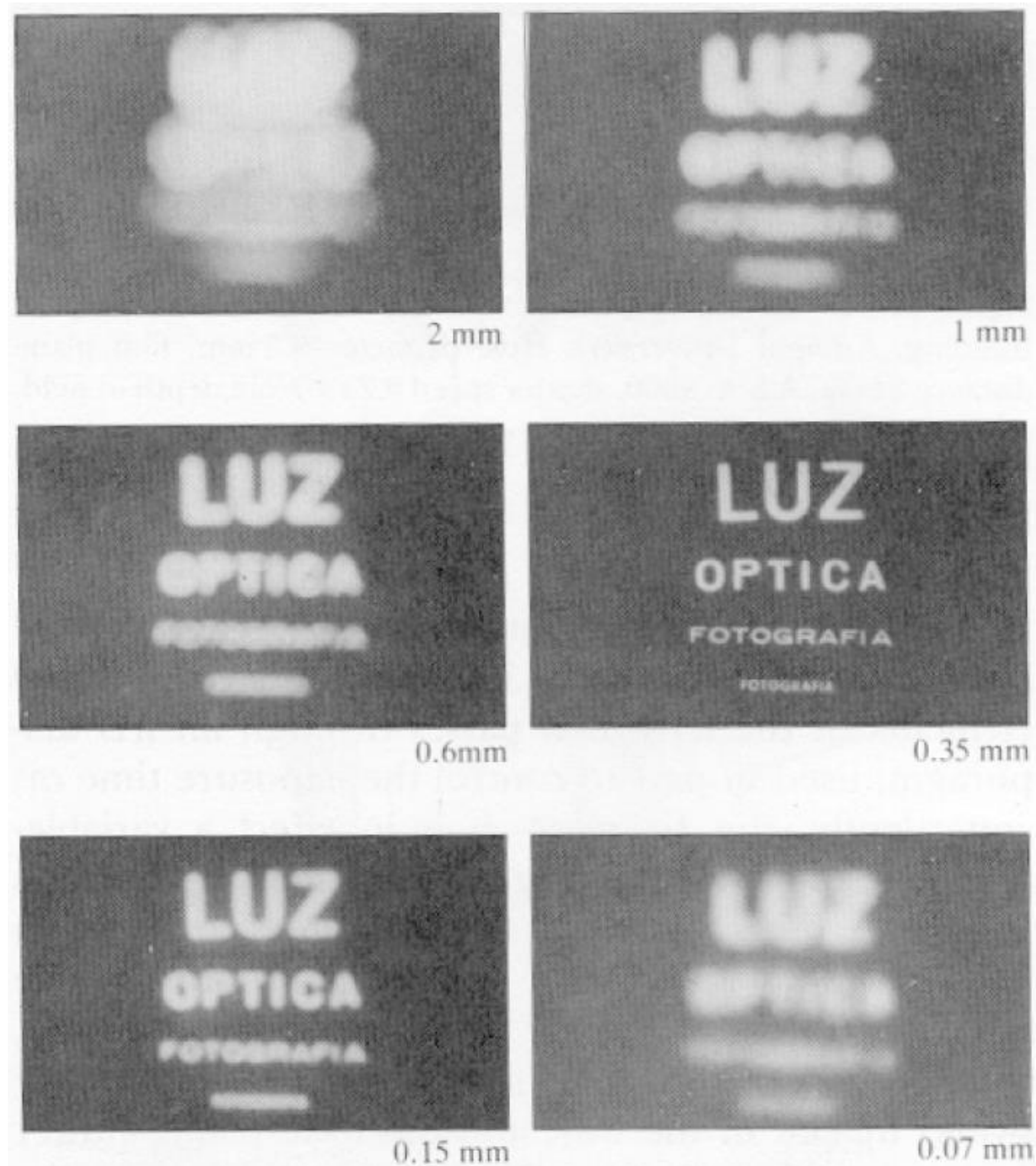
Gather more light from each scene point and also reduce blurring

Why Lenses?

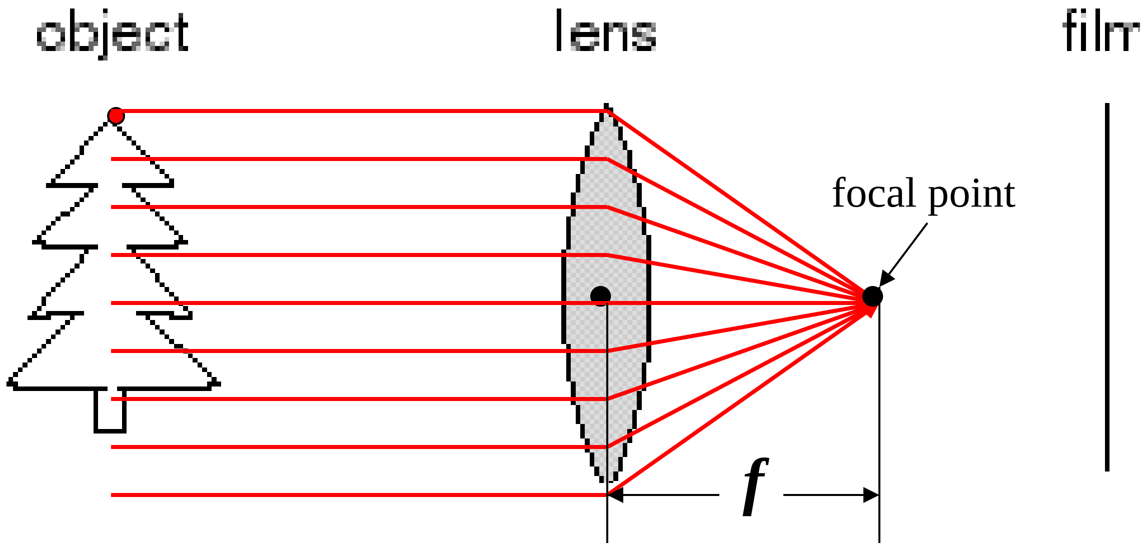
•Pinhole too big - many directions are averaged, blurring the image

•Pinhole too small - diffraction effects blur the image

•Generally, pinhole cameras are *dark*, because a very small set of rays from a particular point hits the screen.



Adding a lens

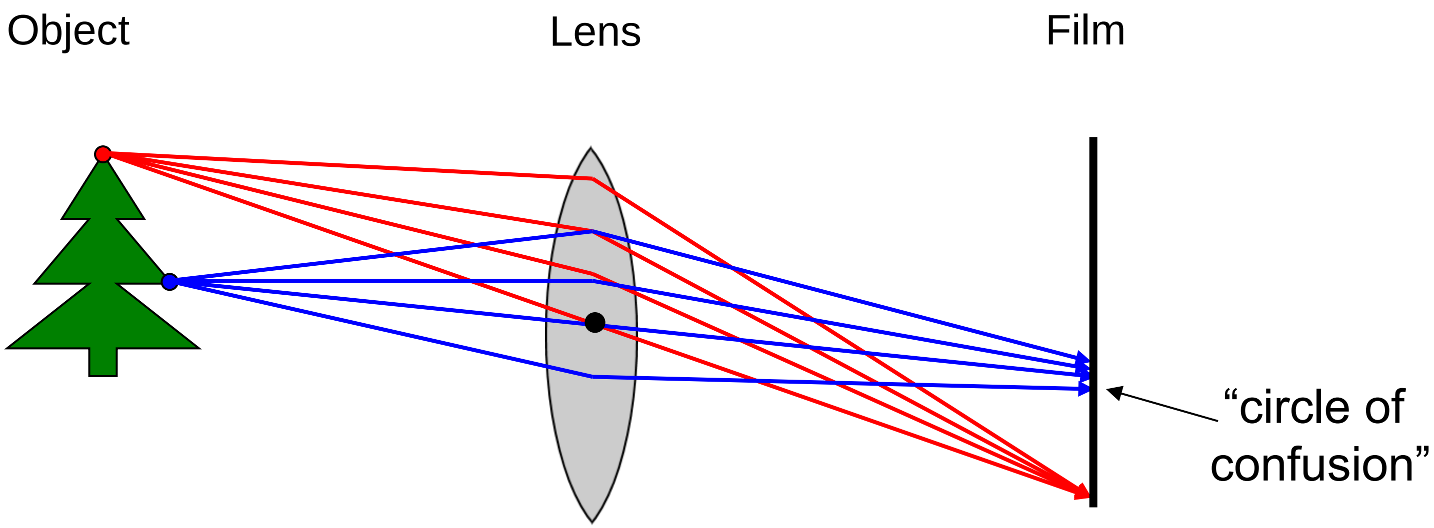


A lens focuses light onto the film

- Thin lens model:
  - Rays passing through the center are not deviated (pinhole projection model still holds)
  - All parallel rays converge to one point on a plane located at the *focal length f*

Slide by Steve Seitz

Adding a lens



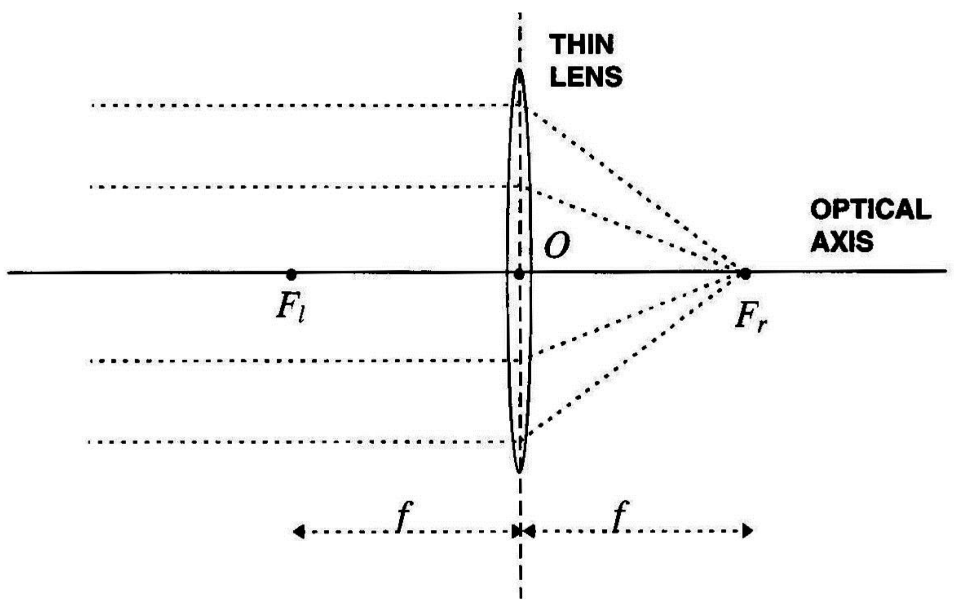
A lens focuses light onto the film

- There is a specific distance at which objects are “in focus”
  - other points project to a “circle of confusion” in the image
- Changing the shape of the lens changes this distance

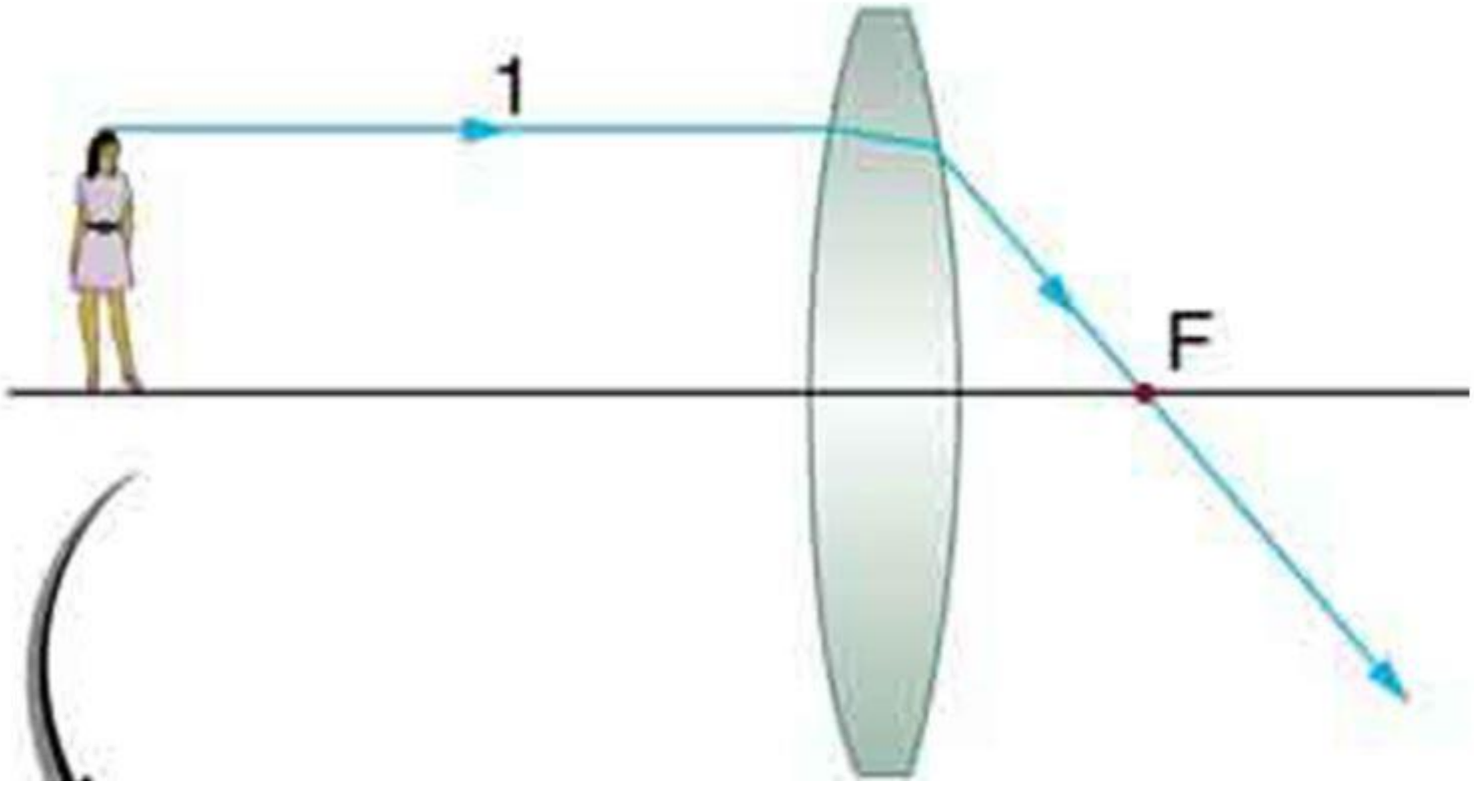
Camera with Lens - Thin Lens Model

- Lens thickness small compared to focal length
- Basic properties

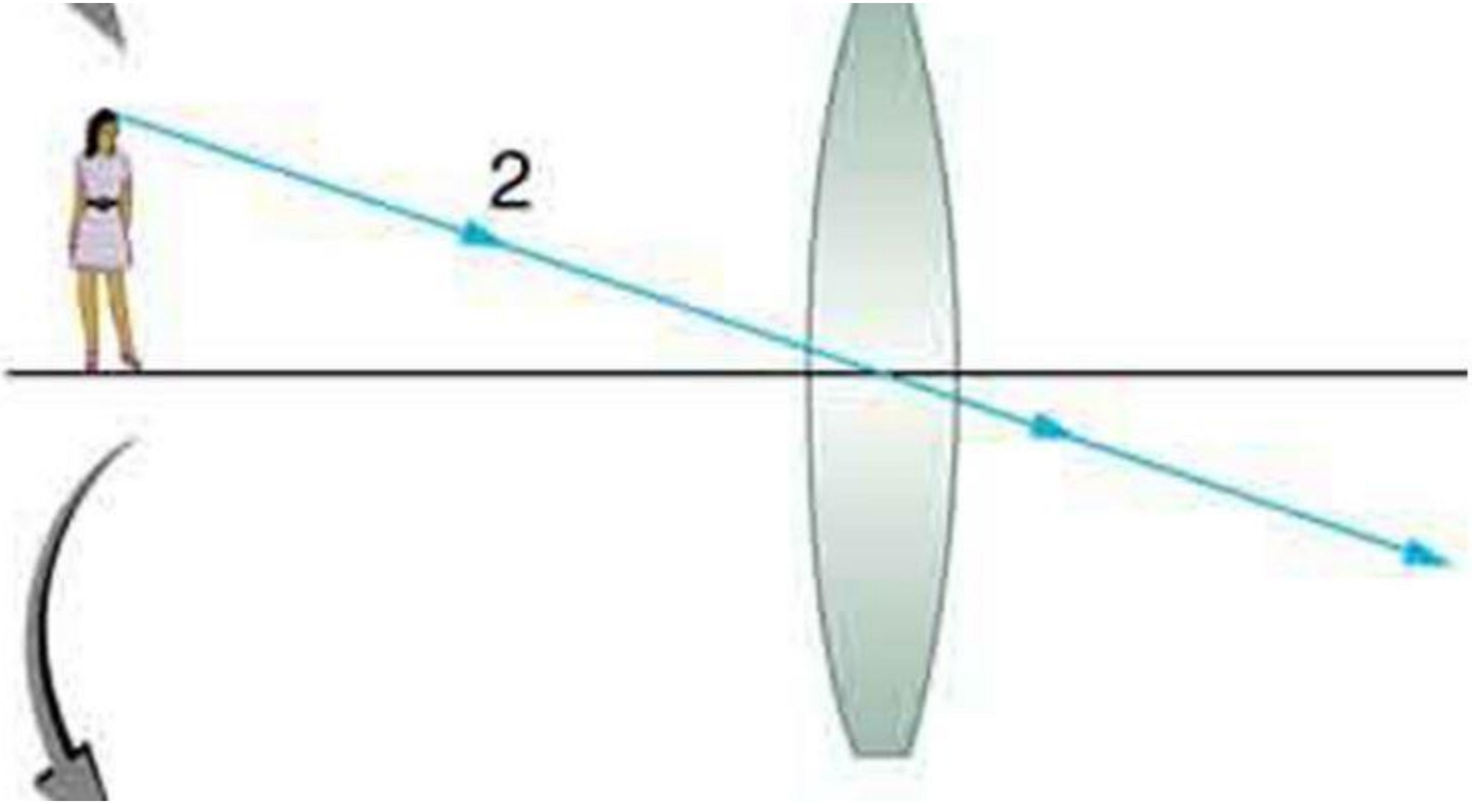
1. Any ray entering the lens parallel to the axis on one side goes through the focal point on the other side.
2. Any ray entering the lens from the focal point on one side emerges parallel to the axis on the other side.



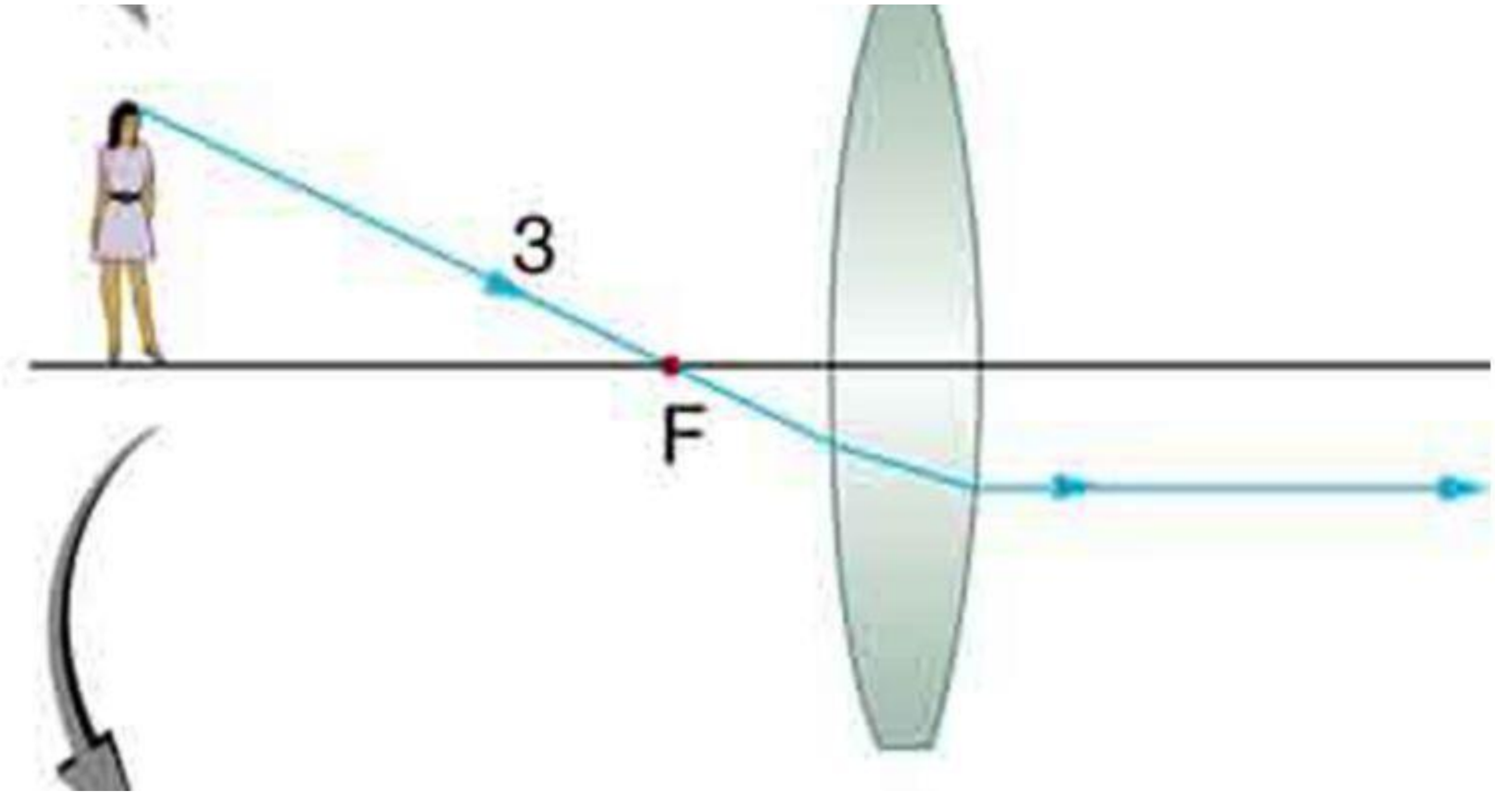
Thin Lens



Thin Lens



Thin Lens



Fundamental Equation of Thin Lenses

Object Lens

Focal Point

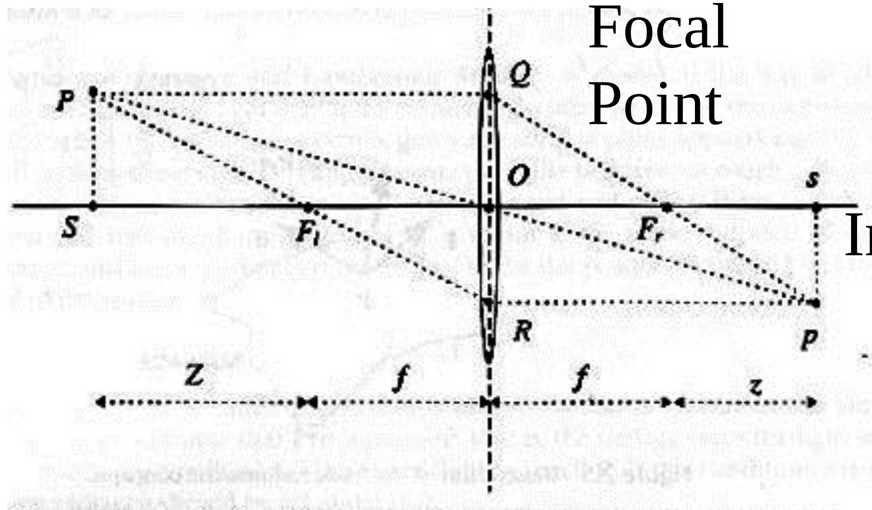


Image sensor

$$\hat{Z} = Z + f \qquad \hat{z} = z + f$$

Effective Focal length

$$\frac{1}{\hat{Z}} + \frac{1}{\hat{z}} = \frac{1}{f}$$

Any point satisfying this equation is in focus

Proof uses similar triangles: PSFl~ORFl and QOFr~spFr and fact that |PS| = |QO| and |sp| = |OR|

Effect of changing value of Z on z

- Look at equations

$$\frac{1}{\widehat{Z}} + \frac{1}{\widehat{z}} = \frac{1}{f} \qquad \widehat{Z} = Z + f \qquad \widehat{z} = z + f$$

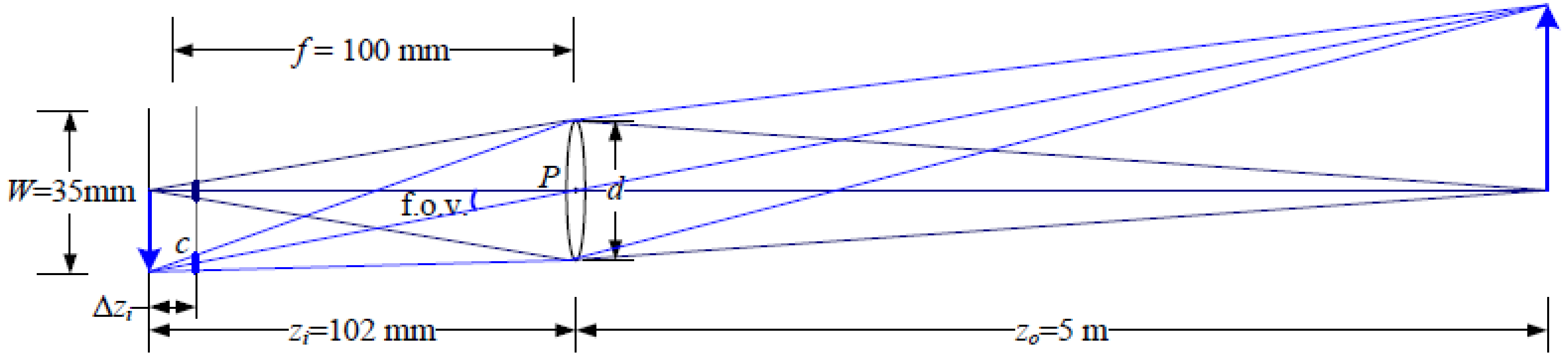
- As you move farther from lens Z increases
- This affects value of z, which is where this point is in focus on the other side of lens
- If Z goes to infinity then z goes to zero
- If Z goes to zero then z goes to infinity

Thin lens applet:

http://www.phys.hawaii.edu/~teb/java/ntnujava/Lens/lens_e.html

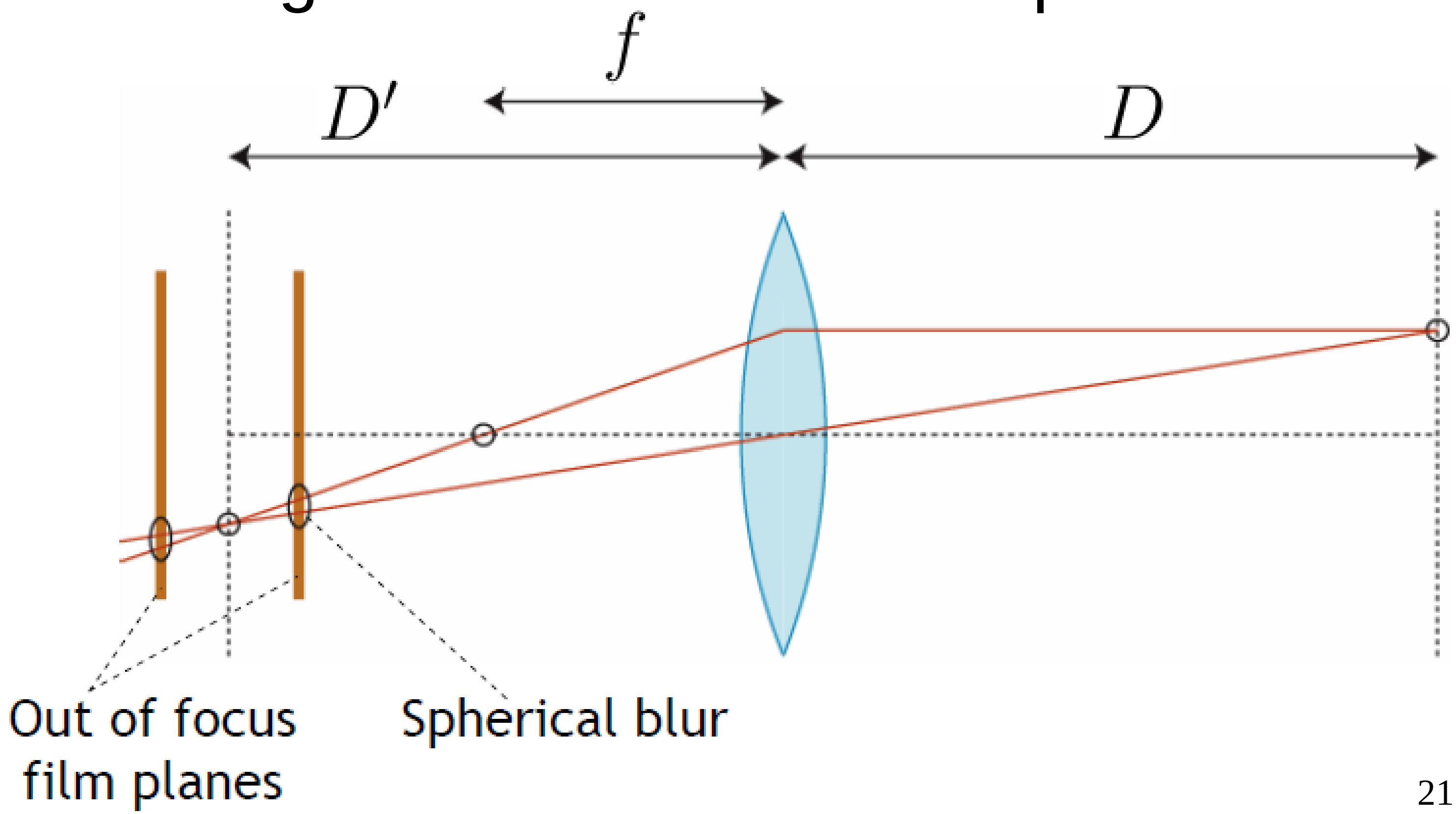
Thin Lenses

- As the point goes to infinity the focal point approaches f, the value for a pin hole camera
- For a lens we can adjust focus ring to move the lens and aperture ring to change aperture
- Both of these adjustments affect what is called the depth of field (explained by model)

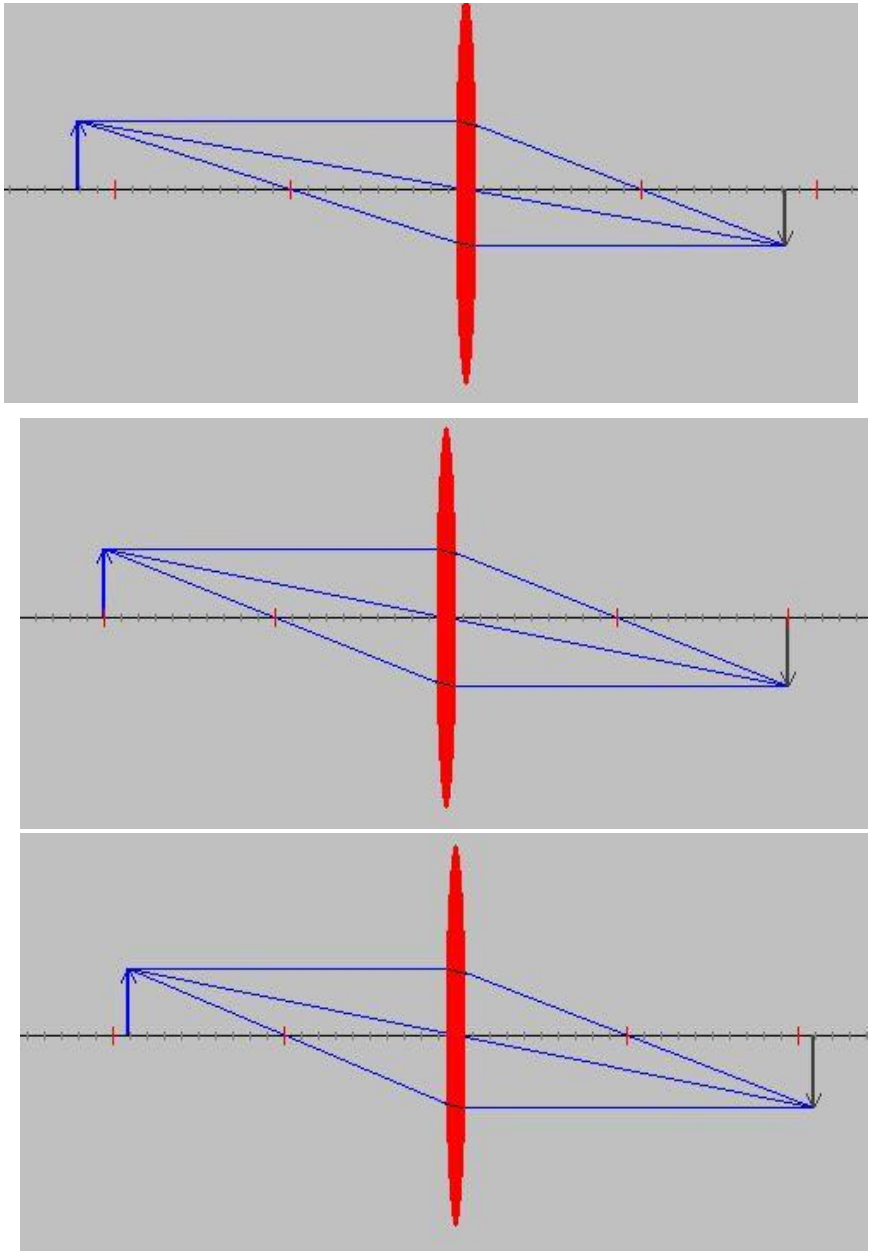


Depth of field

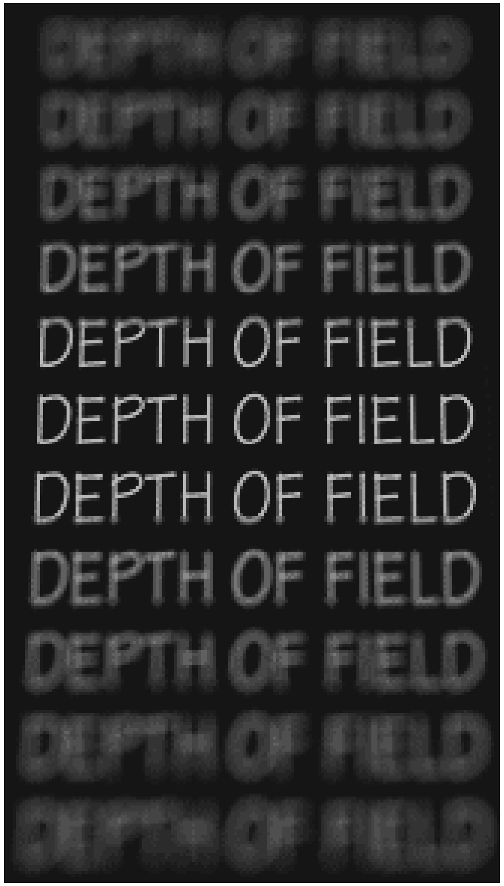
- Point is in focus over a given distance Z
  - This range of Z is called depth of field
  - Depth of field changes with the lens focal length f
- In focus region has less than one pixel of blur



Depth of field



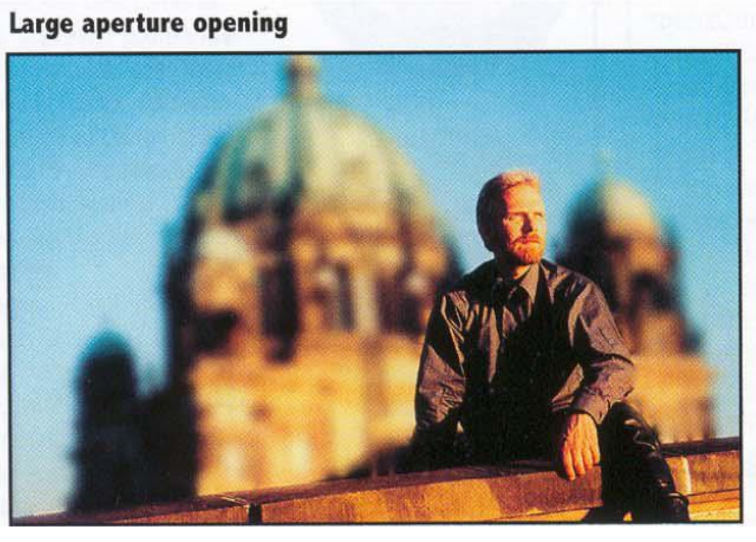
Depth of field

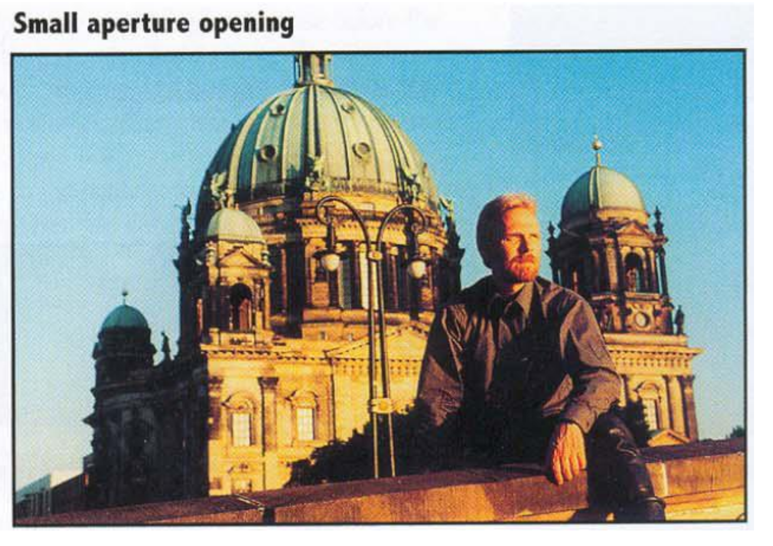


- Pin hole camera has infinite depth of field
- Thin lens implies there is a finite depth of field
- Can change depth of field by changing lens or aperture

Aperture size – also affects dof

- Change aperture size =changes depth of field
- Blurriness of out of focus objects depends on the aperture size
- Larger aperture means smaller depth of field but it also lets in more light





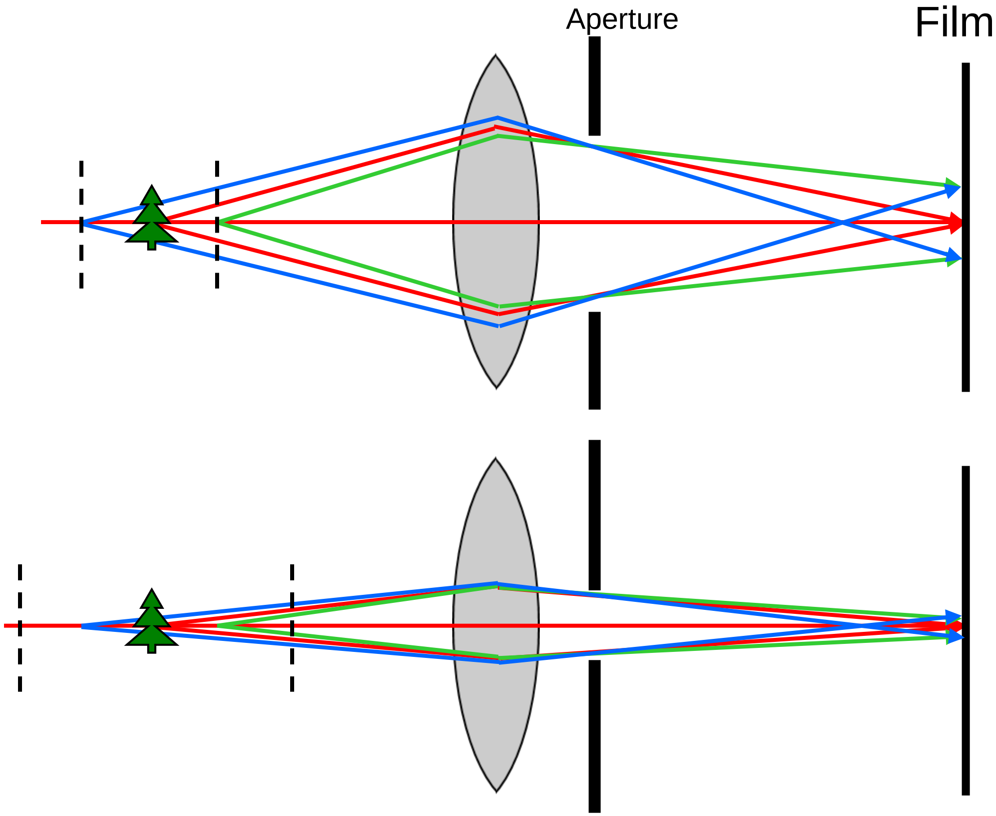
Varying the aperture

Large apeture = small DOF

Small apeture = large DOF

Nice Depth of Field effect

Depth of field



f / 5.6

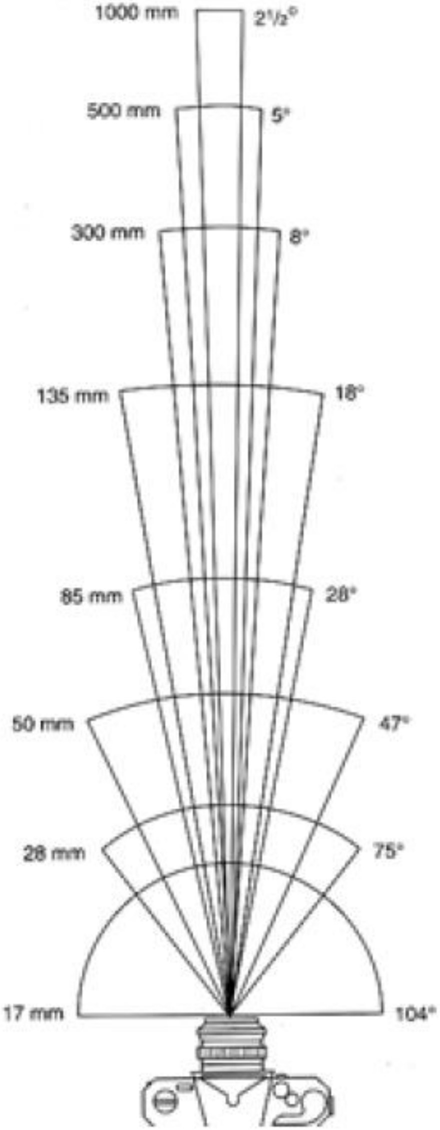
f / 32

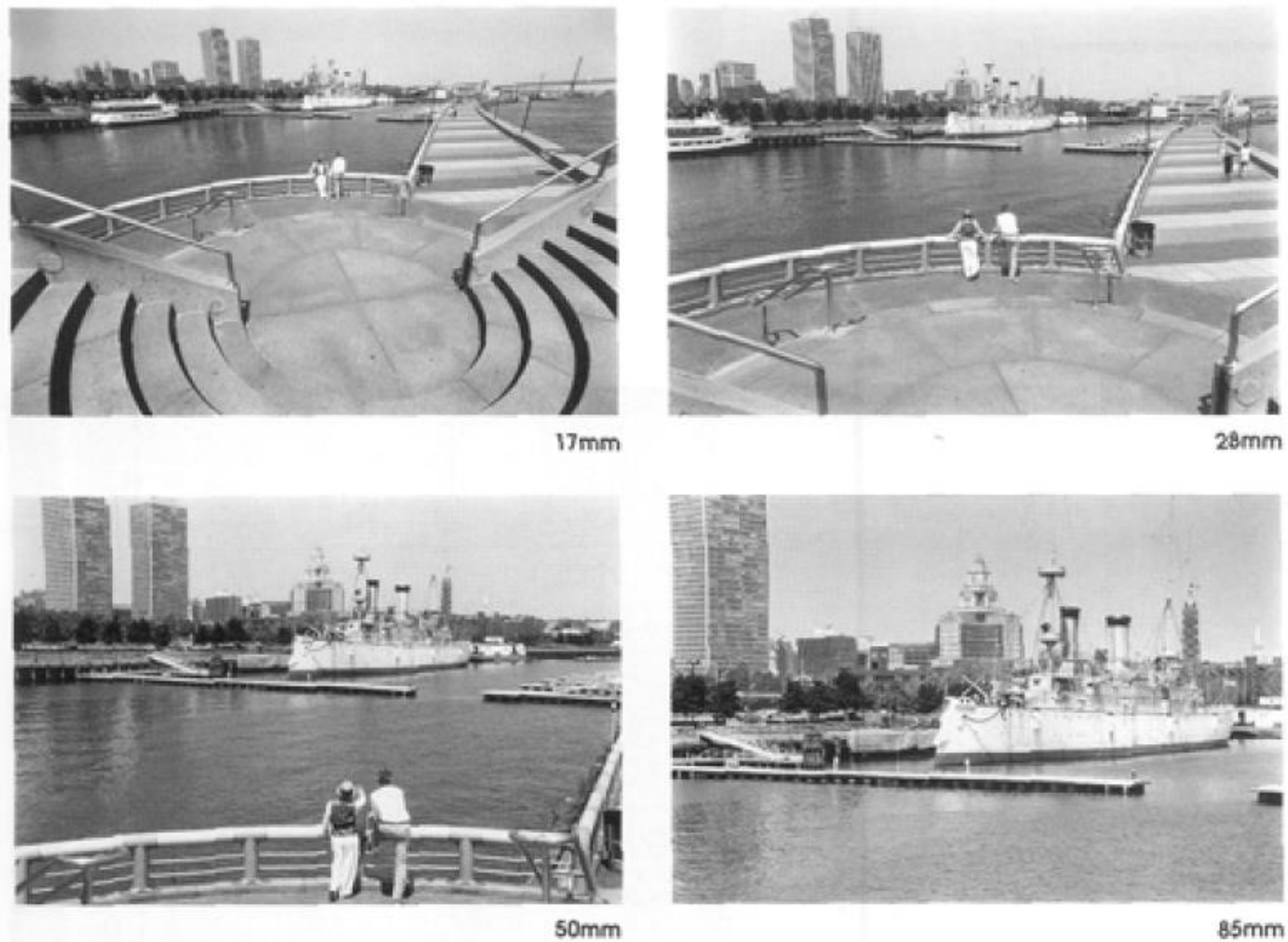
Changing the aperture size affects depth of field

- A smaller aperture increases the range in which the object is approximately in focus

Flower images from Wikipedia http://en.wikipedia.org/wiki/Depth_of_field

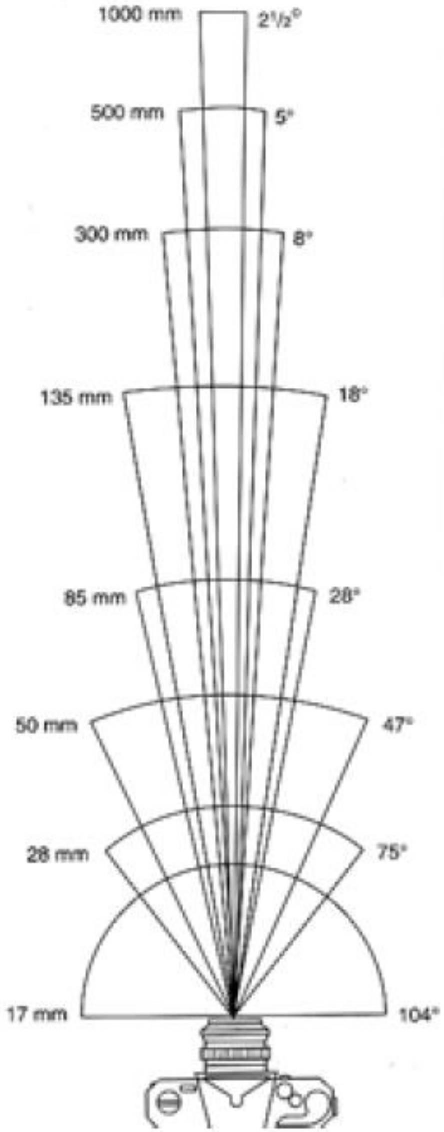
Field of View (Zoom)

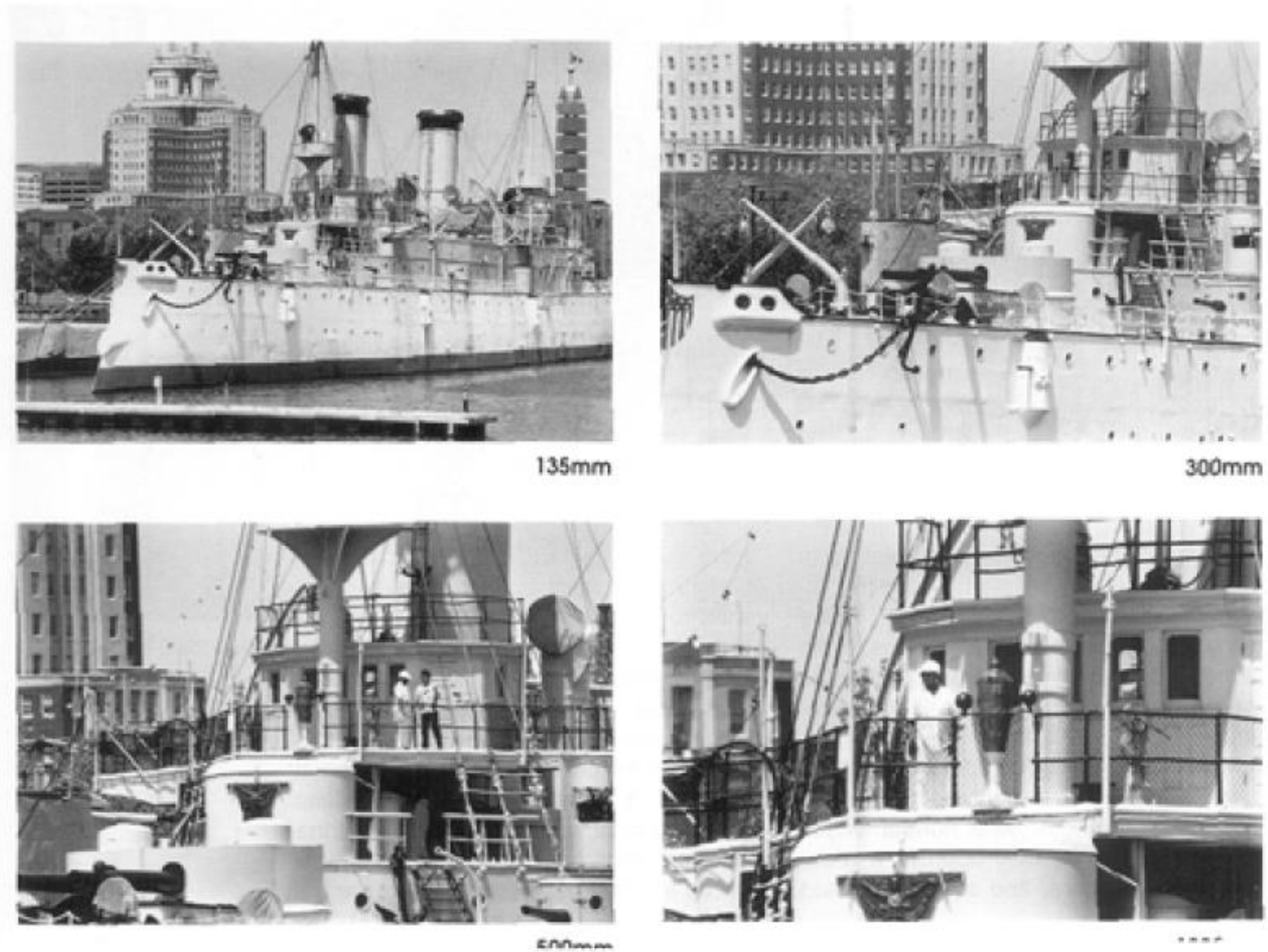




From London and Upton

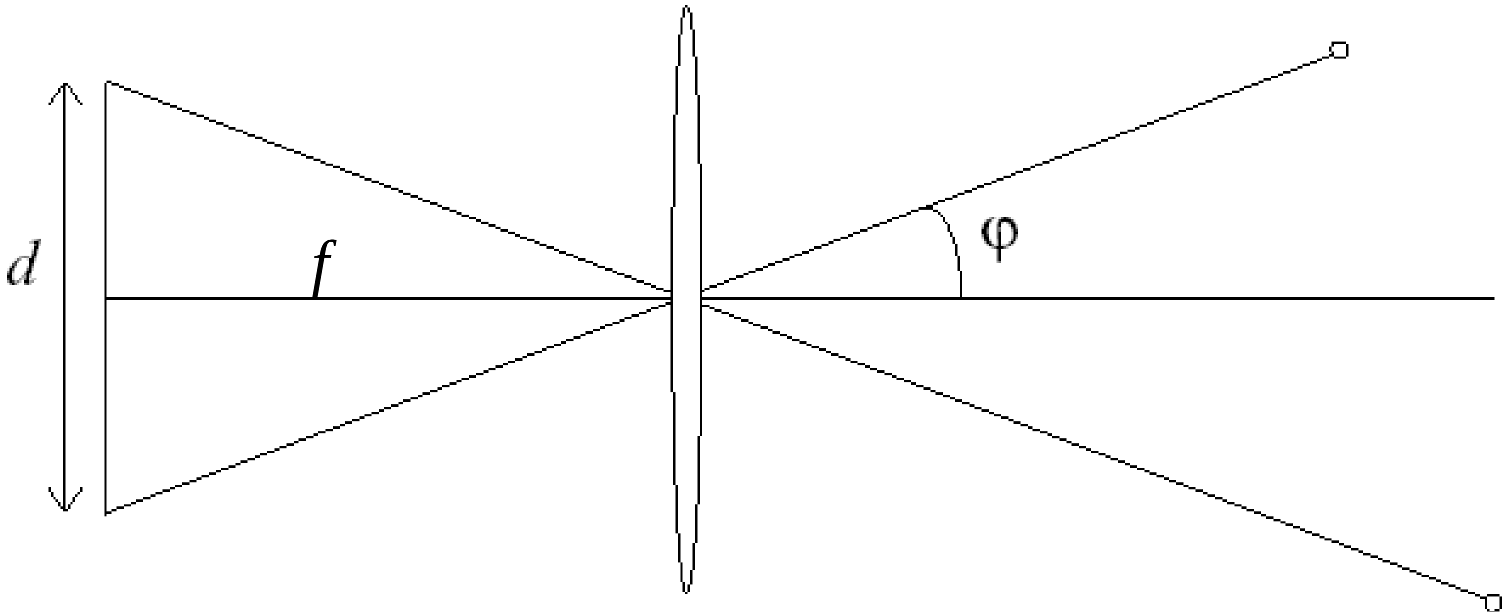
Field of View (Zoom)





From London and Upton

FOV depends of Focal Length



Size of field of view governed by size of the camera retina:

$$\varphi = \tan^{-1}\left(\frac{d}{2f}\right)$$

Smaller FOV = larger Focal Length

Field of View / Focal Length

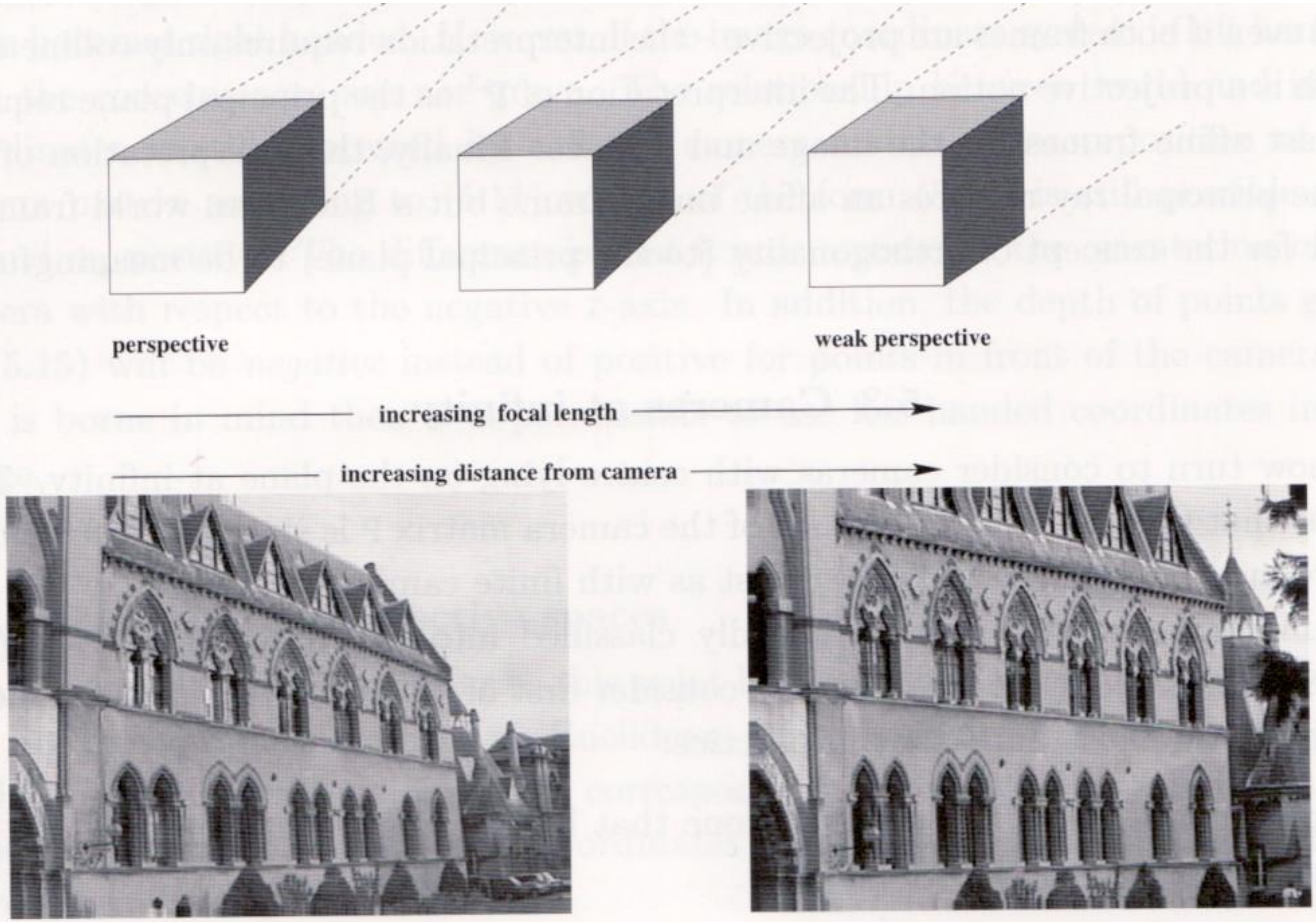
Large FOV, small f
Camera close to car

Small FOV, large f
Camera far from the car

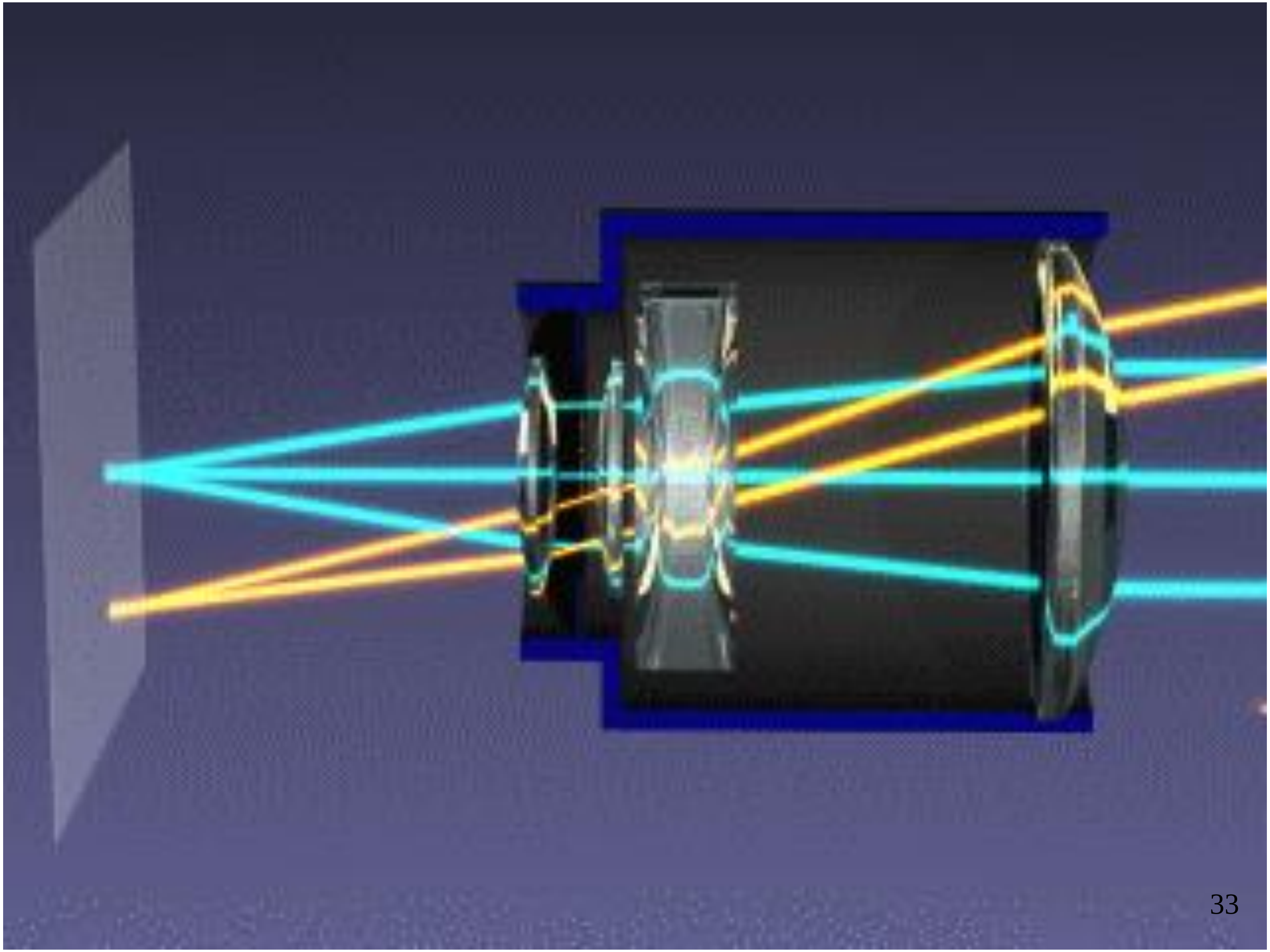
Small field of view has wide angle,
but more perspective distortion

Effect of change in focal length

Small f is wide angle, large f is telescopic



Zoom Lens



Camera parameters

Focus – Shifts the depth that is in focus. Controlled by focus ring. This is a ring on lens elements which moves the lens body.

Focal length – Adjusts the zoom, i.e., wide angle or telephoto lens. Internally a mechanical assembly of lens elements. A fixed focal length lens only has one lens element.

Aperture – Adjusts the depth of field and the amount of light let into the sensor. Controlled by changing the f-stop.

Exposure time – How long an image is exposed. The longer an image is exposed the more light, but could result in motion blur.

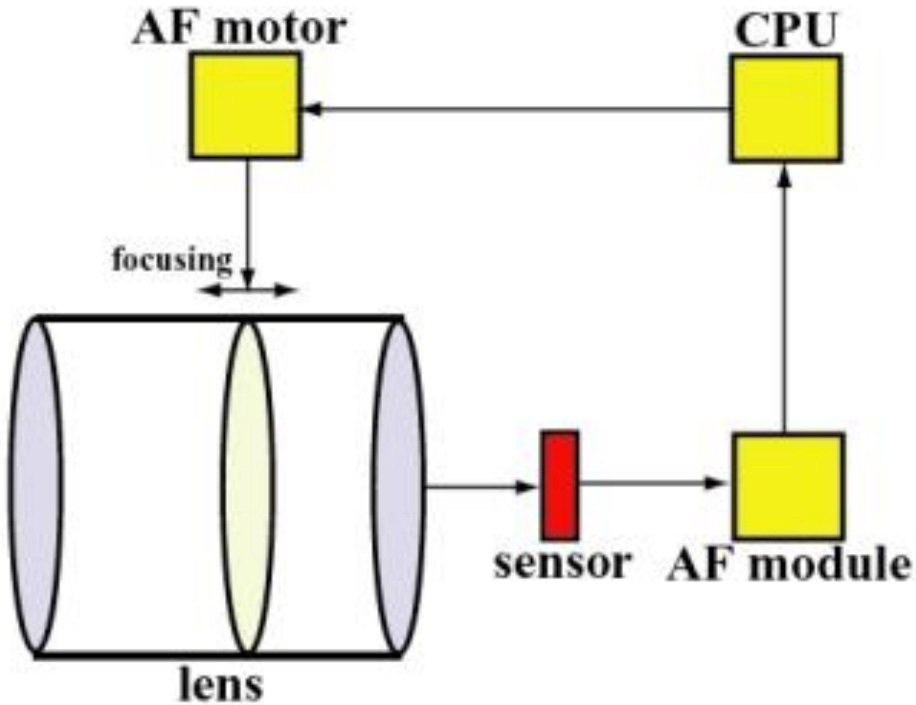
ISO – Adjusts the sensitivity of the “film”. Basically a gain function for digital cameras. Increasing ISO also increases noise.

Autofocus

- Uses sensor, control system and motor to focus on a selected point or area
  - Can get sharp images over large depth variation
- Intelligently adjust camera lens to maintain focus on an object (another definition)
- Two approaches, passive and active
- Active
  - Triangulation using an active sensor such as laser, ultrasound, or infrared light
- Passive
  - Phase detection (similar to stereo) to find depth
  - Contrast detection uses blur or lack of it to find depth

Passive Autofocus – Basic technology

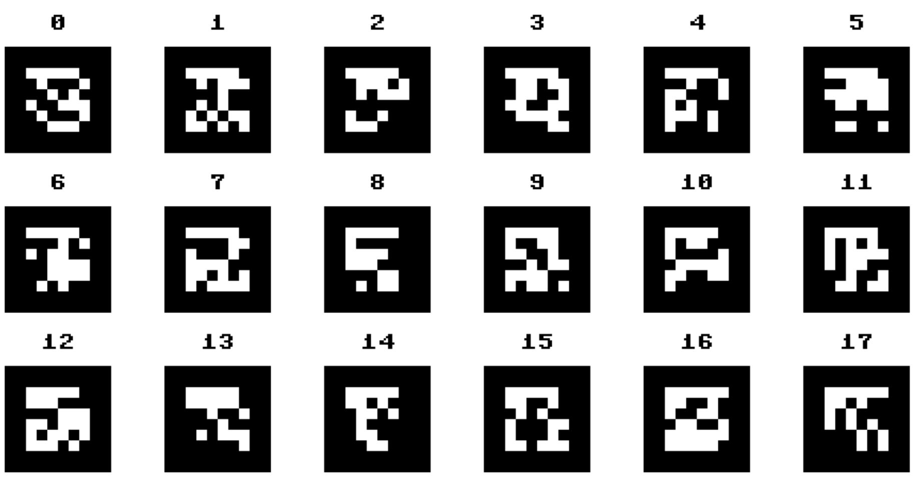
- Camera lens projects image onto sensor
- AF module gives portion of image to CPU in order to process the contrast information
- CPU controls focus motor to move lens



Autofocus

- On all high end cameras, and now on many low end cameras (webcams) and phones
- In Android can have fixed focus, autofocus (it does focus once), or continuous autofocus
- Most sophisticated image processing applications require an in-focus image
- Requires autofocus
  - QRTag and OCR (including my chess application)
- Not require autofocus
  - ARTag, a tag system with much less information
  - Such applications work on wider variety of devices

QRTag versus ARTag

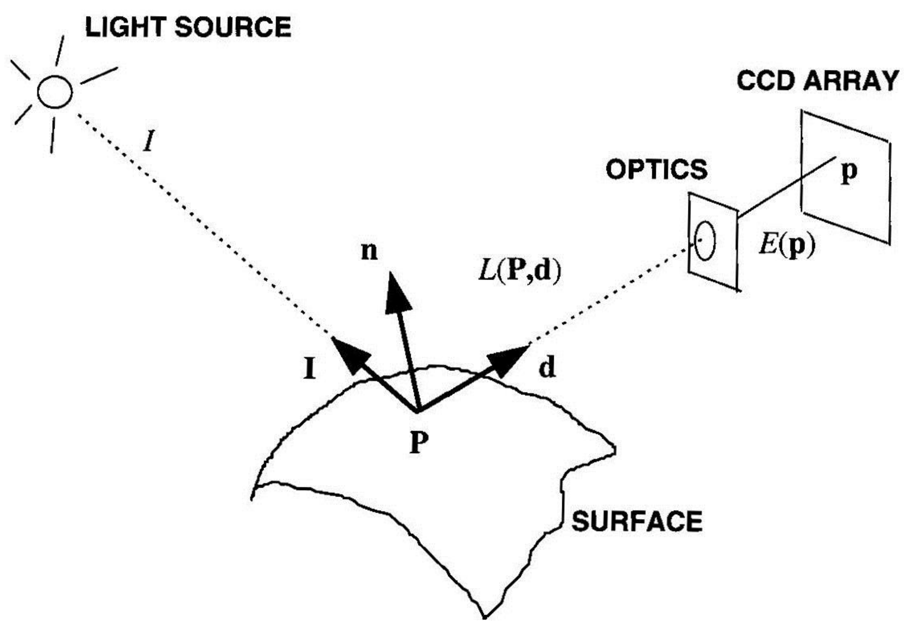


- QRTag
  - Lot of info, small regions
  - Can encode entire URL
- ARTag
  - Less info, large regions
  - Only 10 bits of encoding

Basic radiometry

Image Irradiance: the power of light, per unit area and at each point p of the image plane.

Scene (surface) Radiance: the power of the light, per unit area, ideally emitted by each point p of a surface in 3-D space in a given direction.

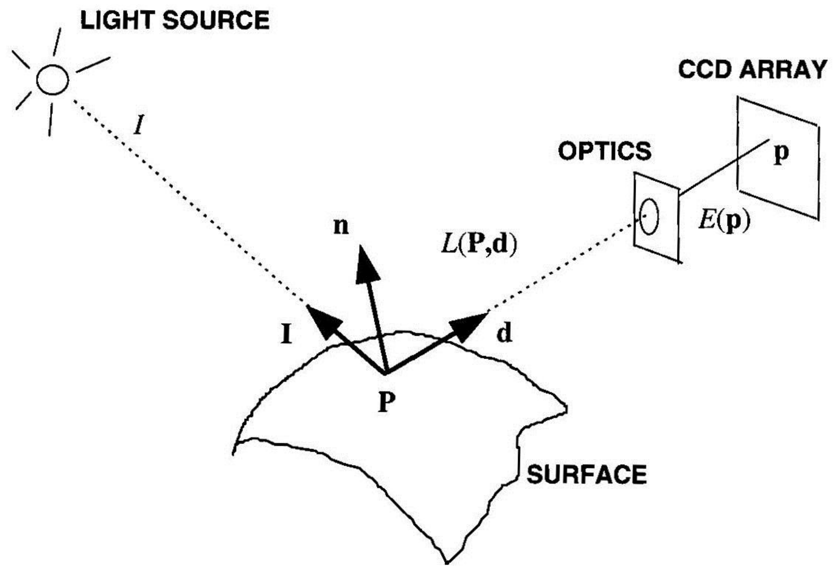


Surface Reflectance Model

- A model of the way in which the surface reflects incident light is called a surface reflectance model
- There are a number of different types of surface reflectance models
  - Fix the lighting, and the object and then move the camera while looking a single surface point
- The changes in appearance of that surface point defines the specularity
  - Plain sheet of paper is non-specular (no change)
  - Desktop is semi-specular (some change)
  - Mirror is very specular (a great deal of change)

Surface Reflectance for Lambertian

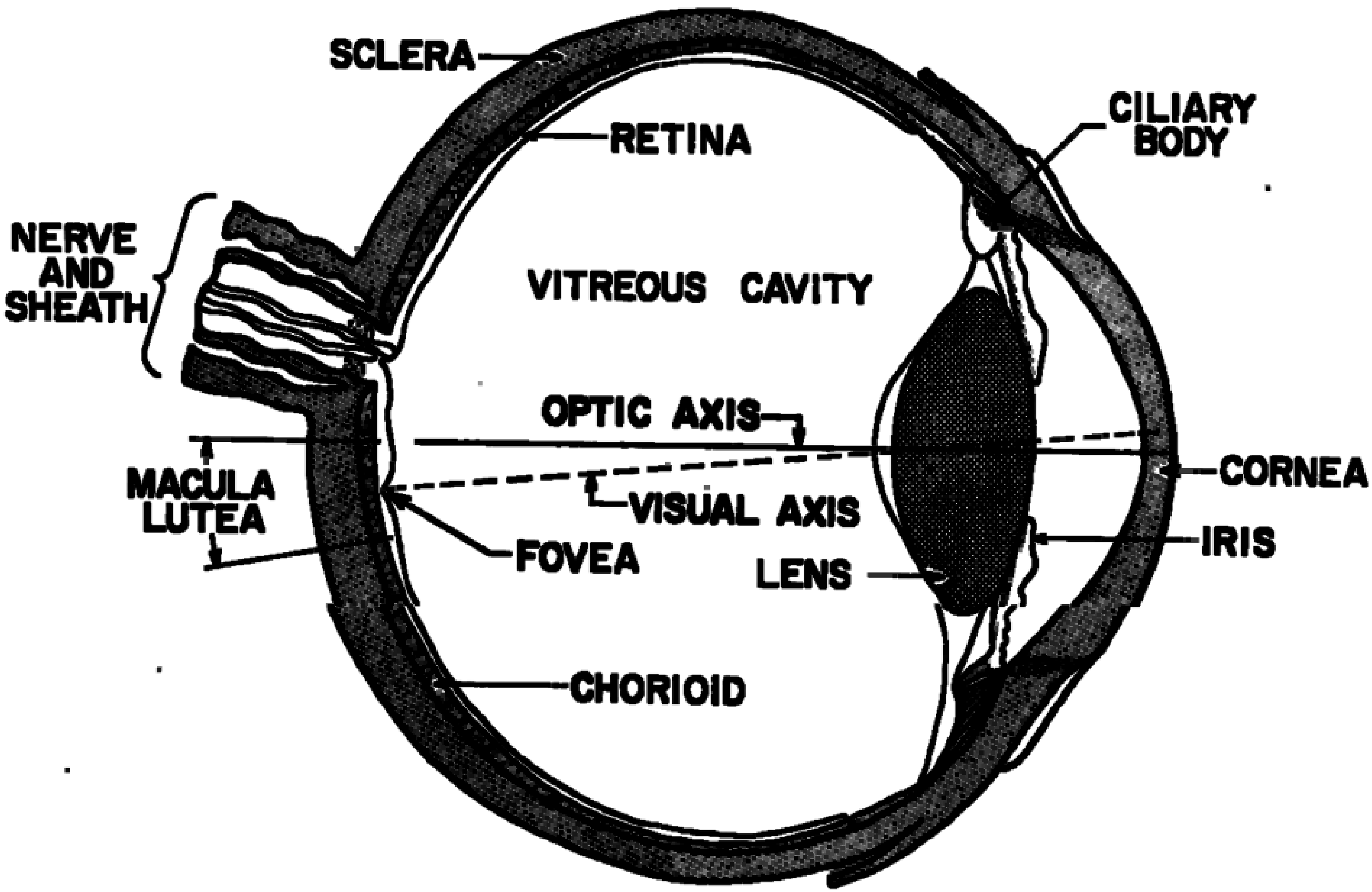
$$\mathbf{L} = \rho \mathbf{I}^T \mathbf{n}$$



ρ is called surface albedo and it depends on the surface material
And L is scene irradiance (no d vector term)

Lambertian model: each surface point appears equally bright from all viewing directions (no term with d). Non specular surface. **Specular model**: this is not true, looks brighter from some viewing directions (mirrors are very specular). These models are much more complex than the lambertain model (more parameters)

Human Eye


SCLERA
RETINA
CILIARY BODY
NERVE AND SHEATH
VITREOUS CAVITY
OPTIC AXIS
MACULA LUTEA
CORNEA
VISUAL AXIS
FOVEA
IRIS
LENS
CHORIOID

CCD (Charge-Coupled Device) Cameras

Small solid state cells convert light energy into electrical charge (sensing elements always rectangles and are usually square)

The image plane acts as a digital memory that can be read row by row by a computer

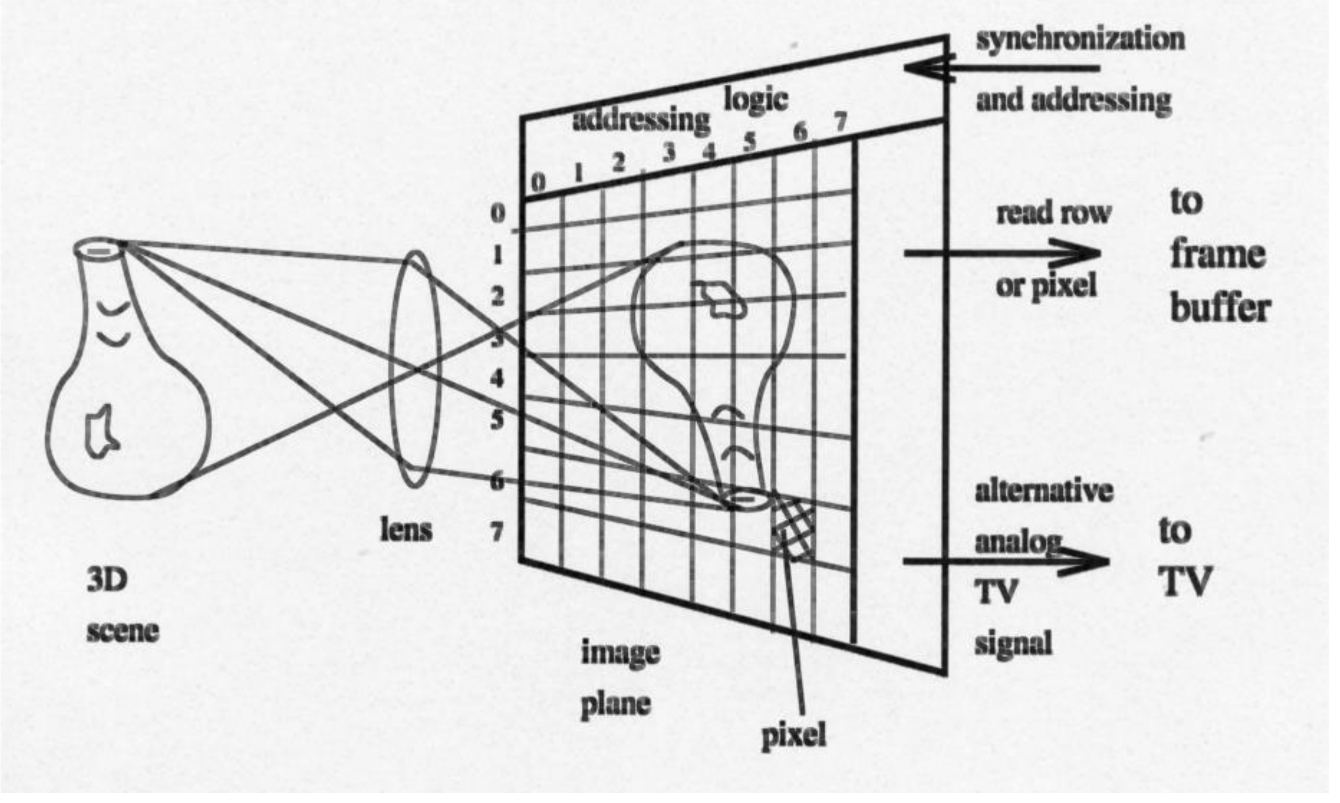
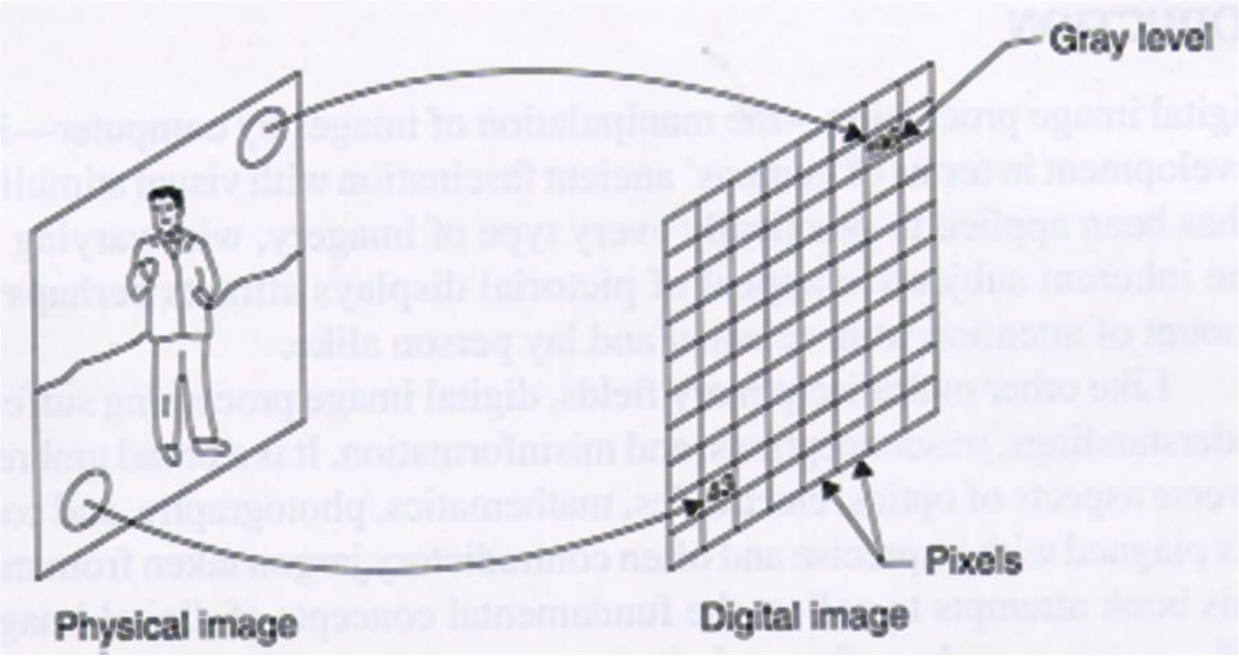


Image Digitization

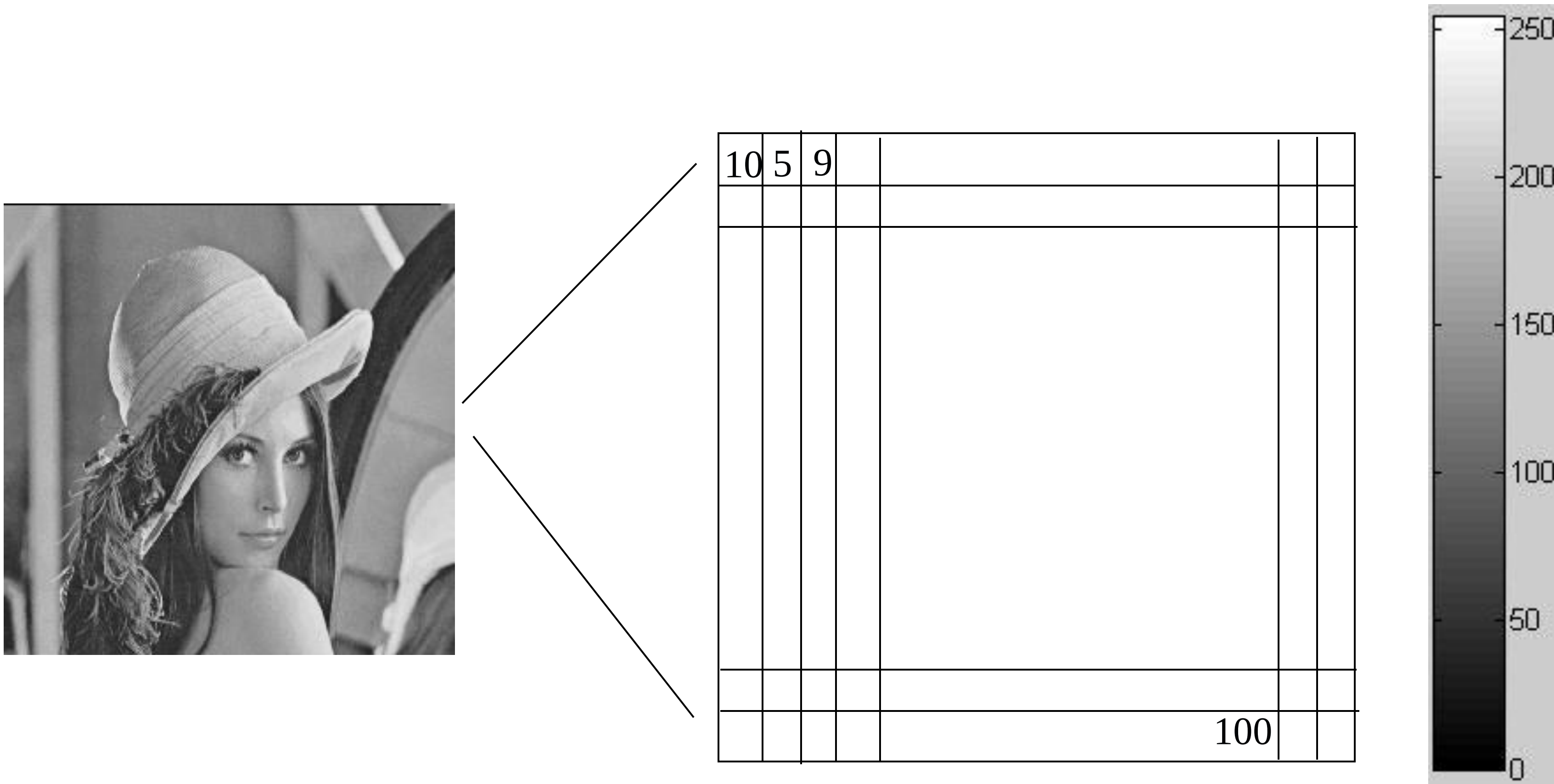


Sampling – measuring the value of an image at a finite number of points.

Quantization – representing the measured value at the sampled point, by an integer.

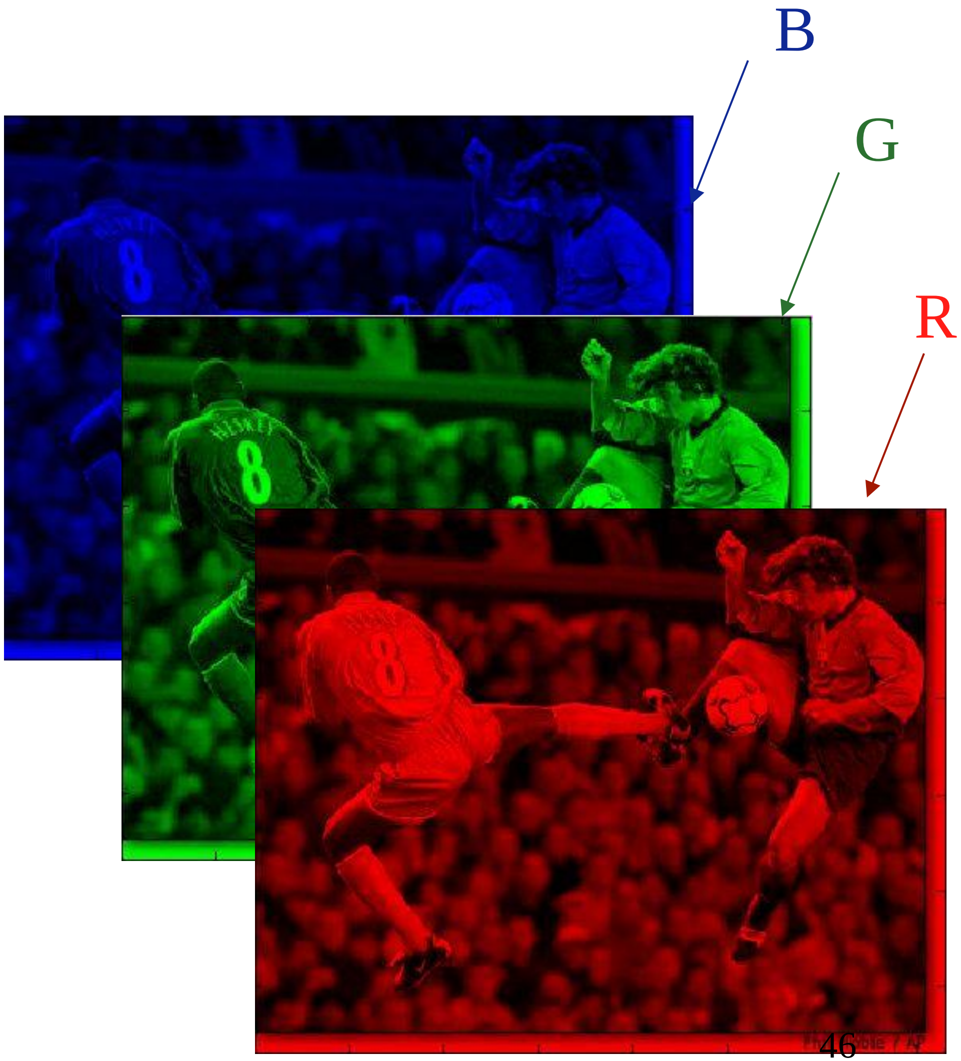
Pixel – picture element, usually in the range [0,255]

Grayscale Image



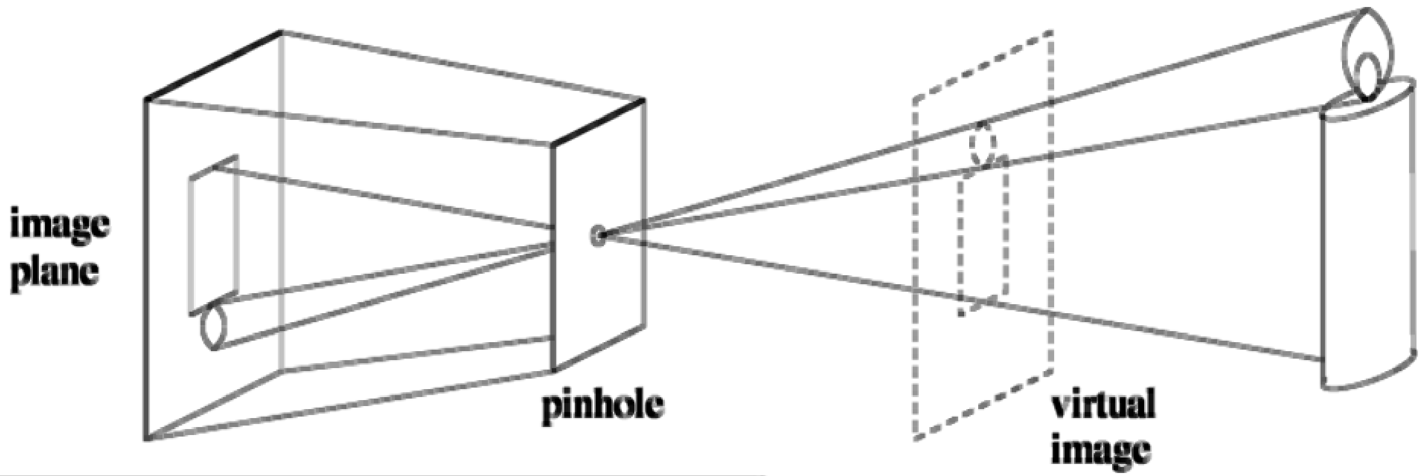
A digital image is represented by an integer array E of m-by-n.
E(i,j), a pixel, is an integer in the range [0, 255].

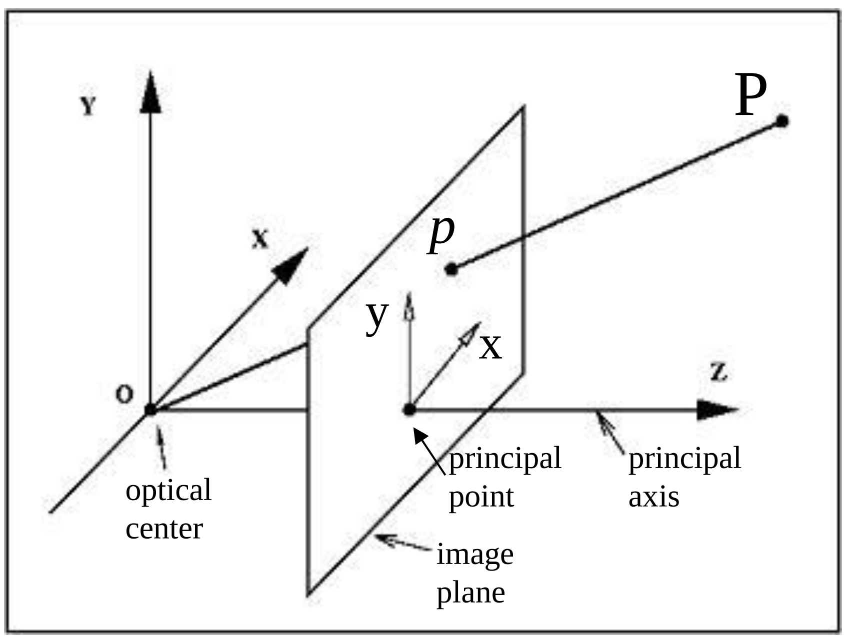
Color Image



Geometric Model of Camera

Perspective projection

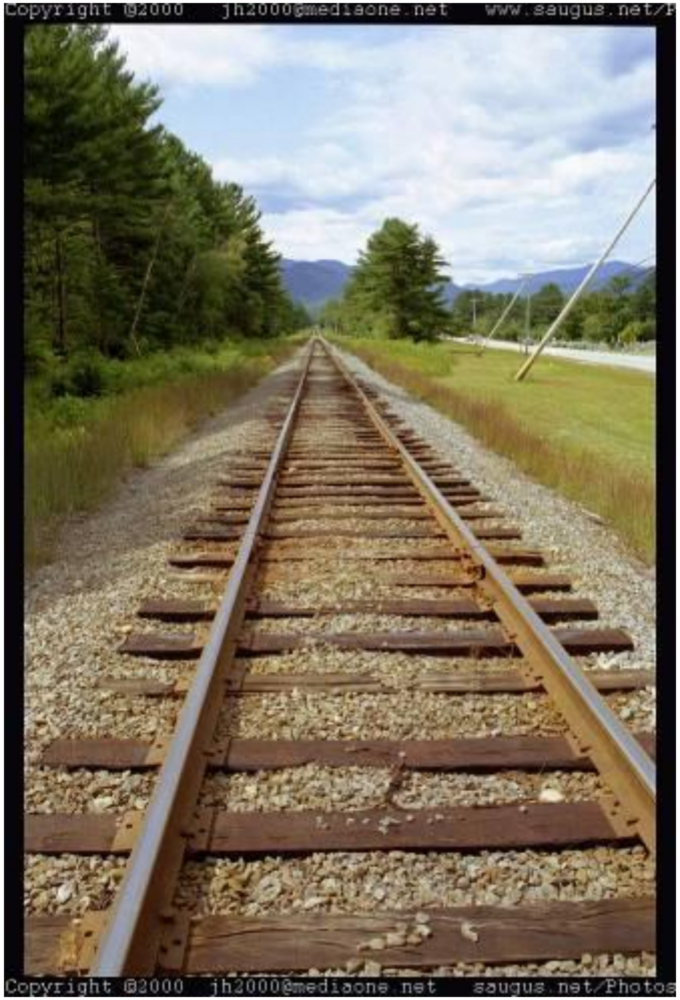




$$P(X,Y,Z) \longrightarrow p(x,y)$$

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

Funny things happen…



Parallel lines aren't…

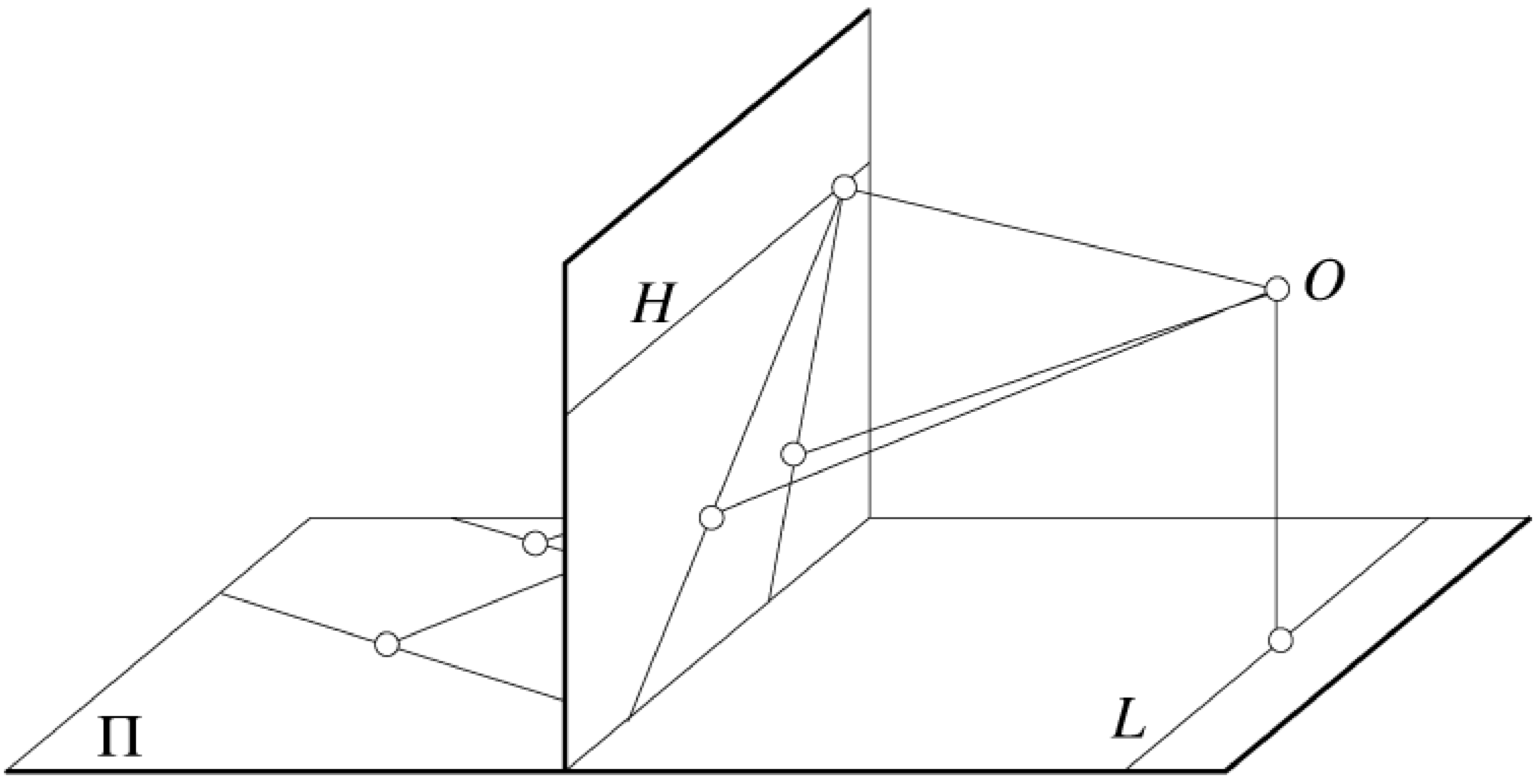


Figure by David Forsyth

Lengths can't be trusted...

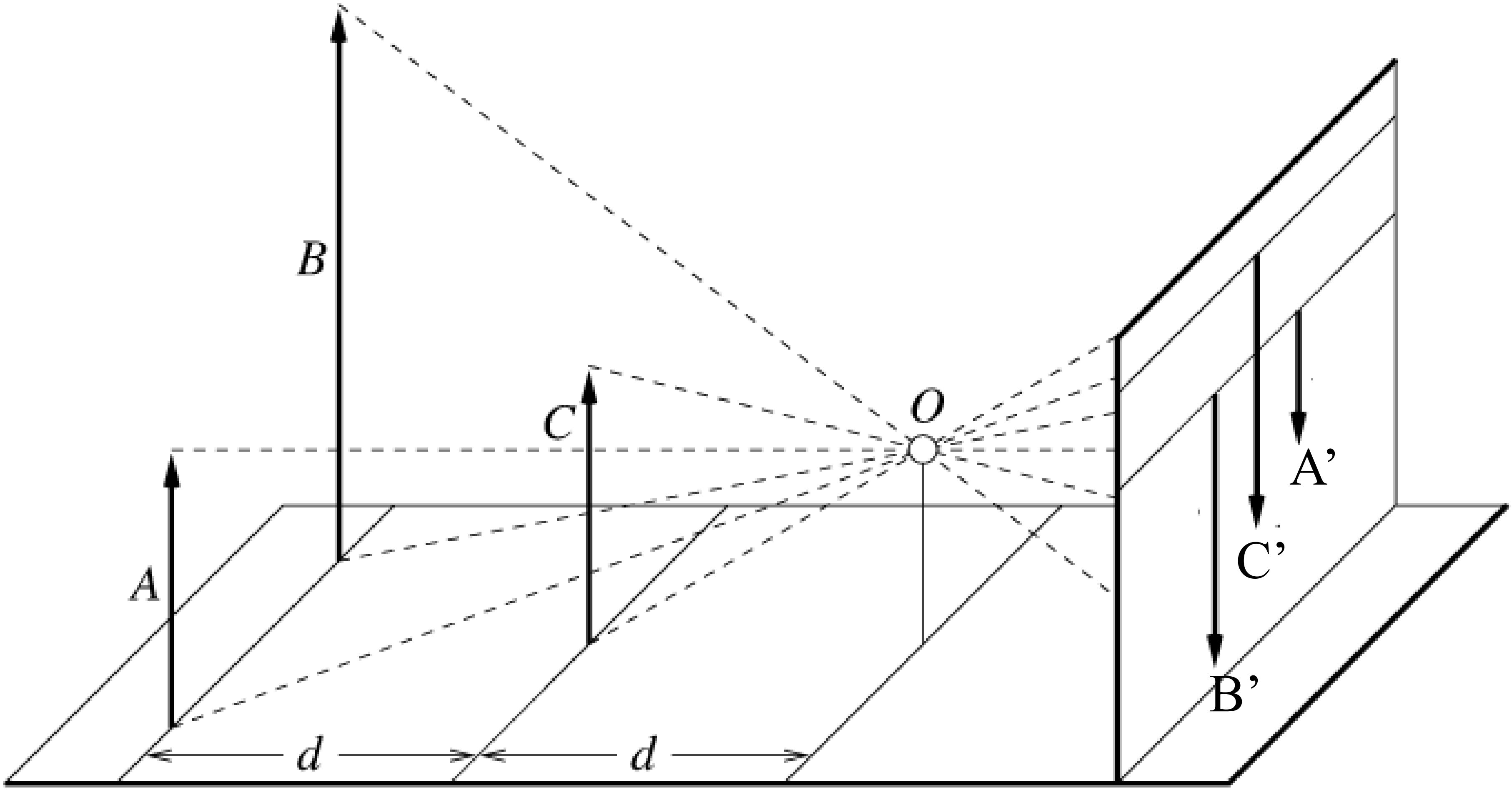


Figure by David Forsyth