VC 11/12 – T2
Image Formation

Mestrado em Ciência de Computadores

Mestrado Integrado em Engenharia de Redes e Sistemas Informáticos

Miguel Tavares Coimbra

U. PORTO FC

Outline

- ‘Computer Vision’?
- The Human Visual System
- Image Capturing Systems

Acknowledgements: Most of this course is based on the excellent courses offered by Prof. Shree Nayar at Columbia University, USA and by Prof. Srinivasa Narasimhan at CMU, USA. Please acknowledge the original source when reusing these slides for academic purposes.

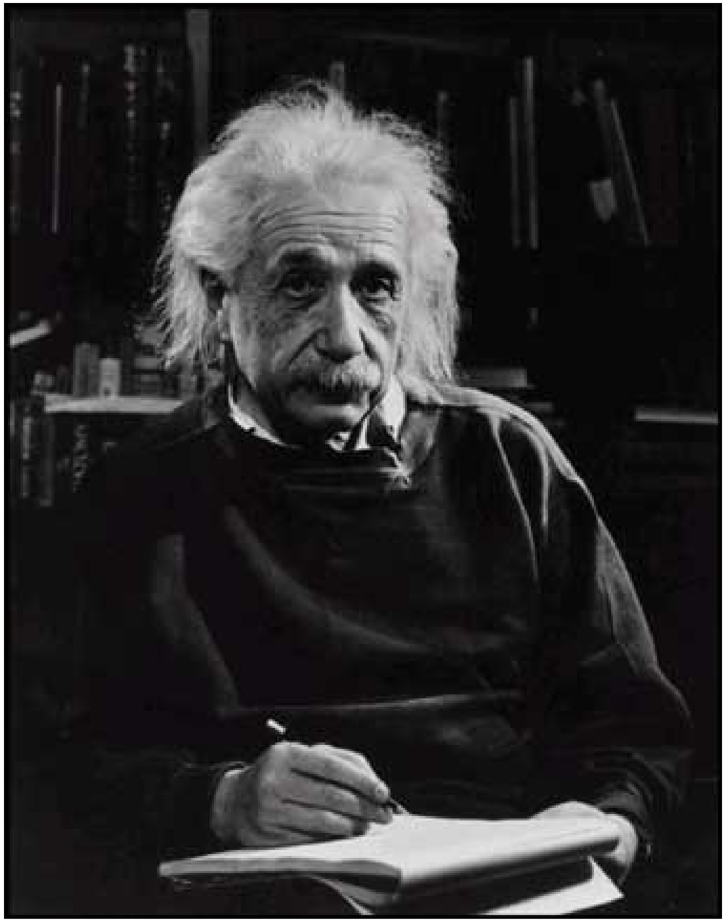
Topic: Computer Vision?

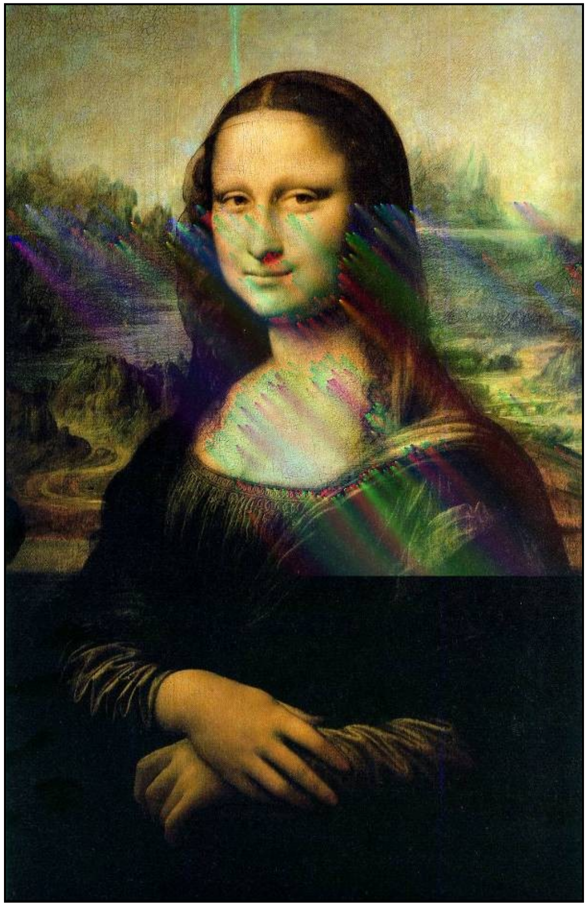
- ‘Computer Vision’?
- The Human Visual System
- Image Capturing Systems

A Picture is Worth 1000 Words

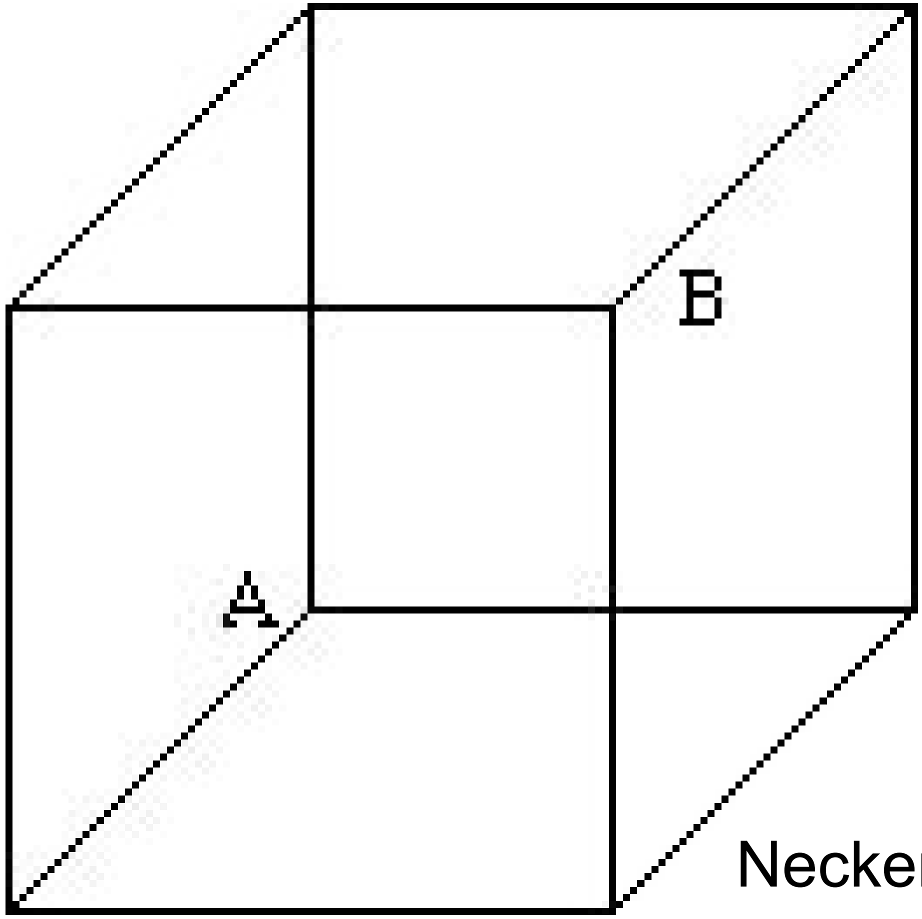
A Picture is Worth 100.000 Words

A Picture is Worth a Million Words





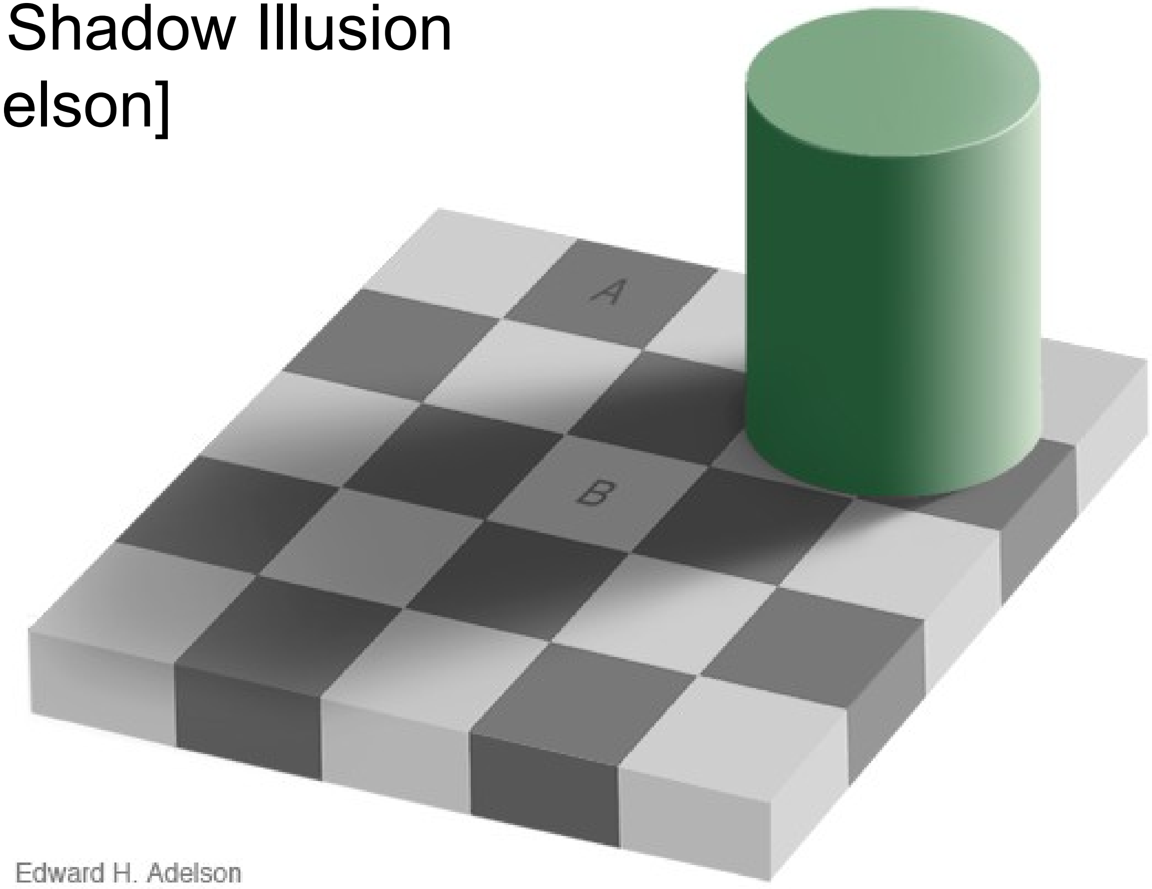
A Picture is Worth a ...?



Necker's Cube Reversal

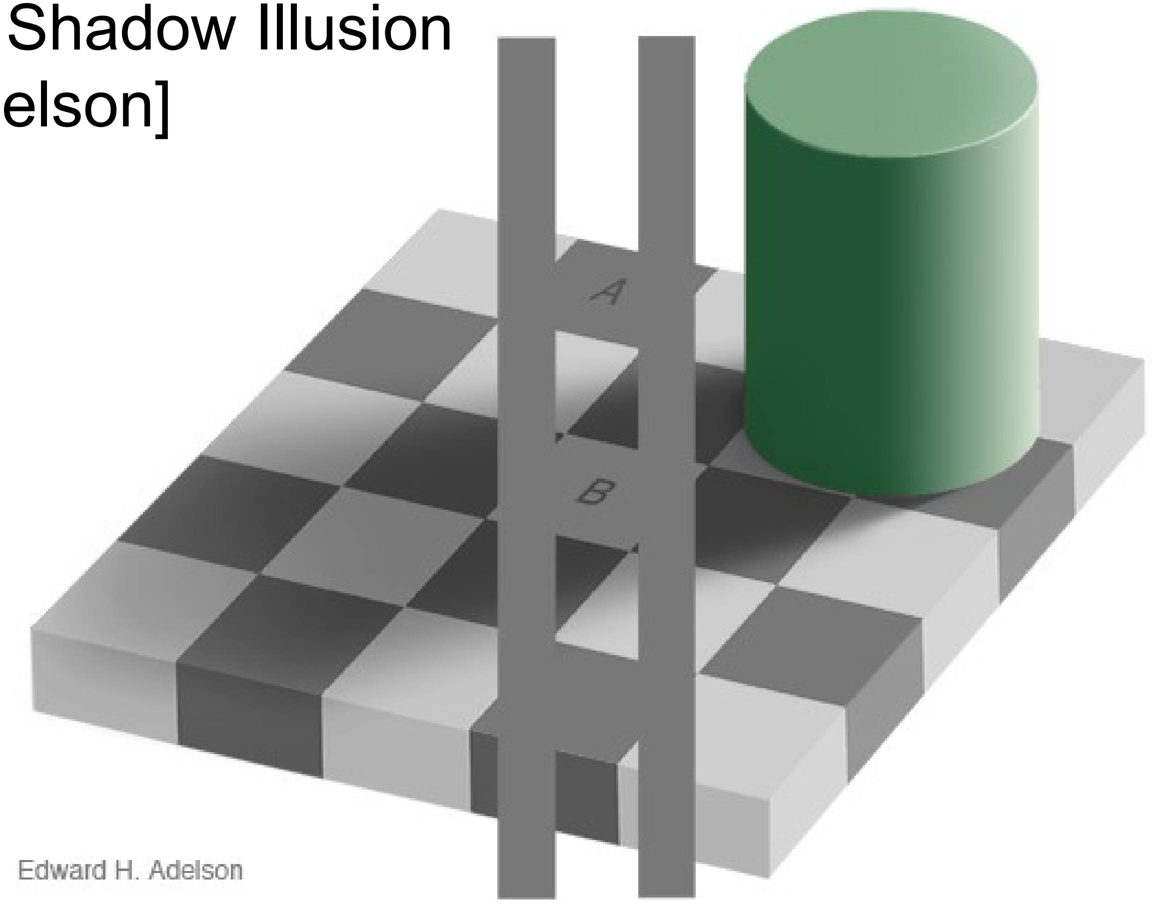
A Picture is Worth a ...?

Checker Shadow Illusion
[E. H. Adelson]



A Picture is Worth a ...?

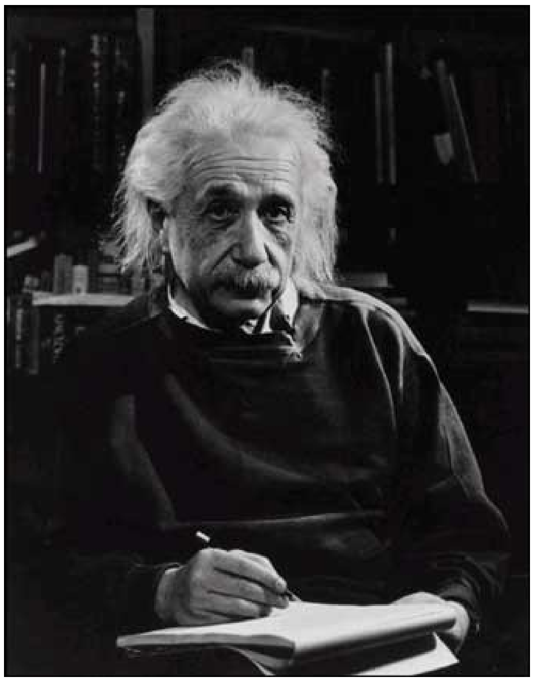
Checker Shadow Illusion
[E. H. Adelson]



Human Vision

- Can do amazing things like:
  - Recognize people and objects
  - Navigate through obstacles
  - Understand mood in the scene
  - Imagine stories
- But:
  - Suffers from Illusions
  - Ignores many details
  - Ambiguous description of the world
  - Doesn't care about accuracy of world

Digital Images



What we see

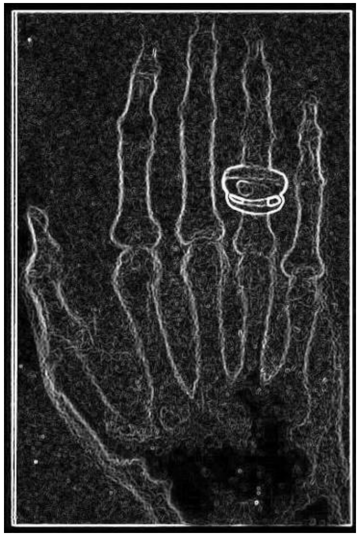
0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
6	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0

What a computer sees

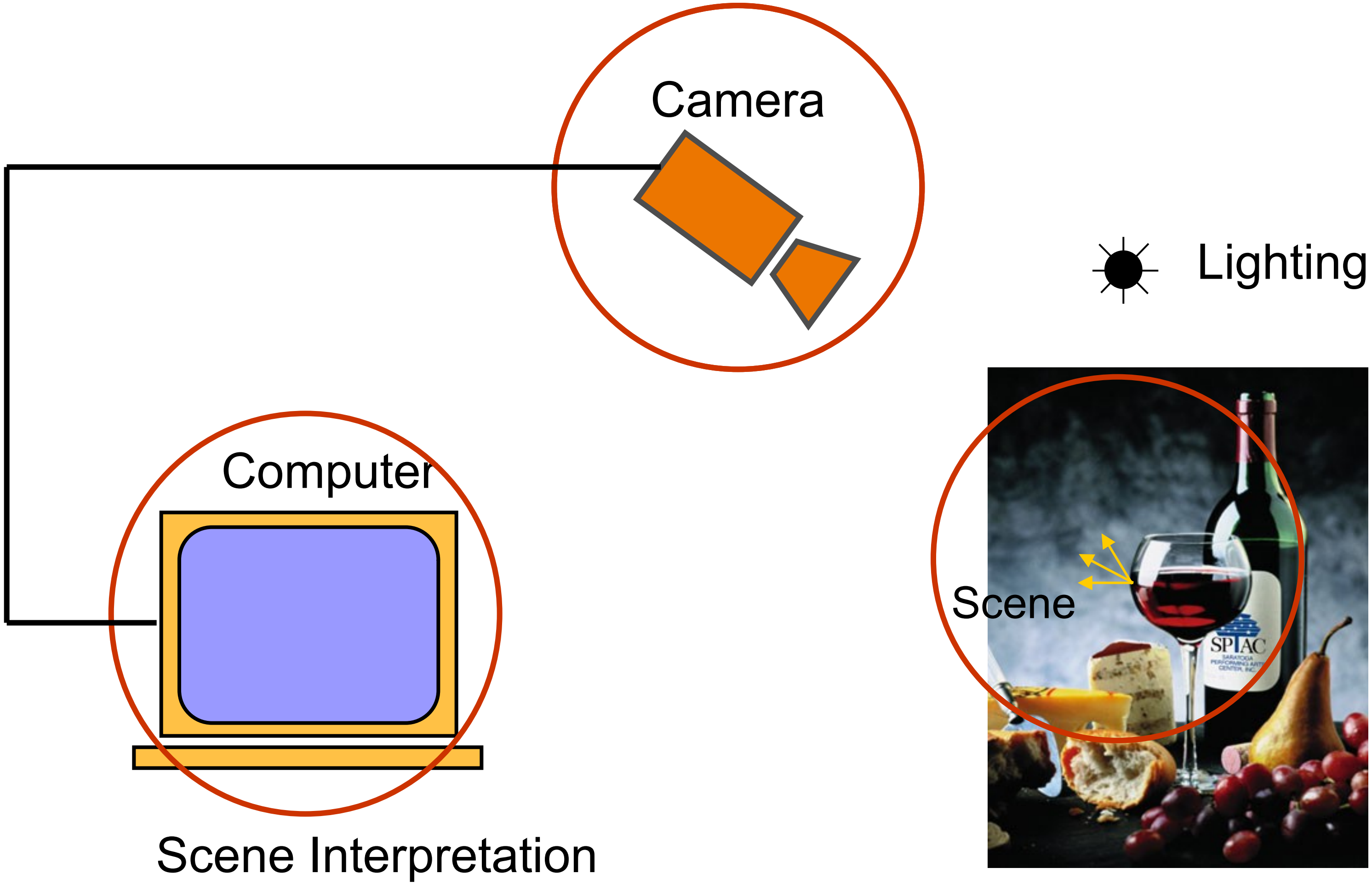
Computer Vision

"The goal of Computer Vision is to make useful decisions about real physical objects and scenes based on sensed images",

Shapiro and Stockman, "Computer Vision", 2001



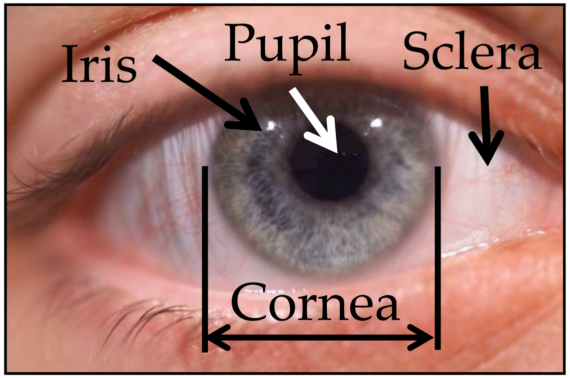
Components of a Computer Vision System

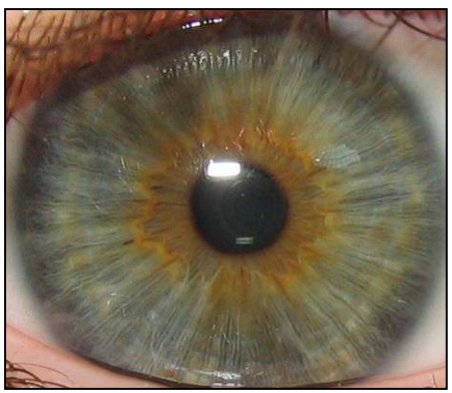


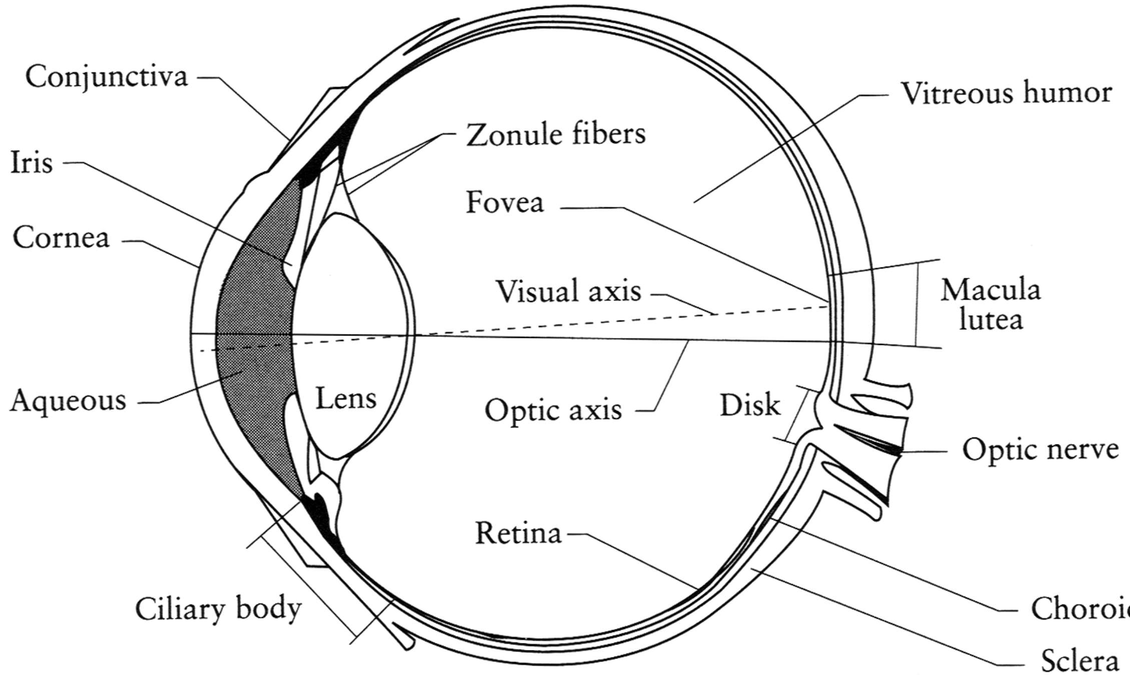
Topic: The Human Visual System

- ‘Computer Vision’?
- The Human Visual System
- Image Capturing Systems

Our Eyes



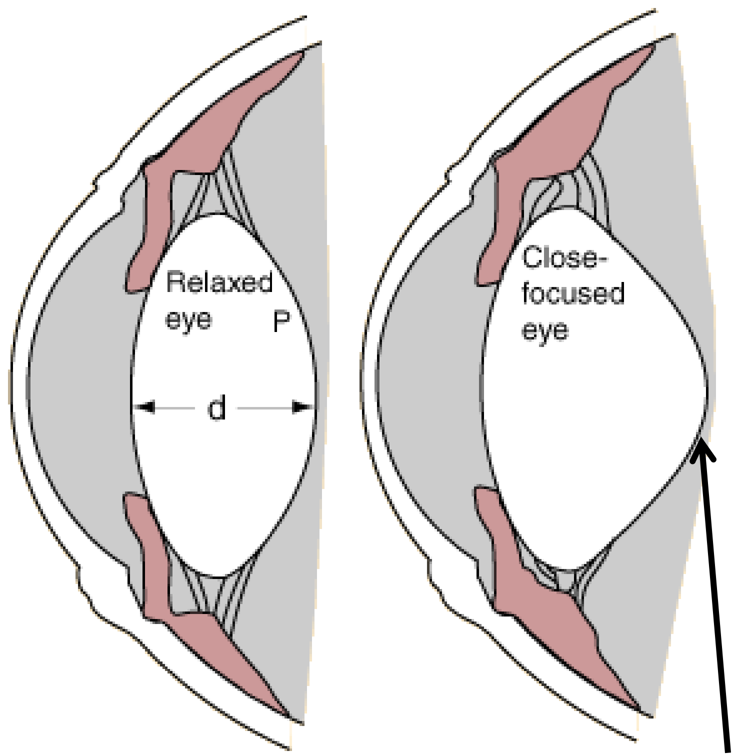




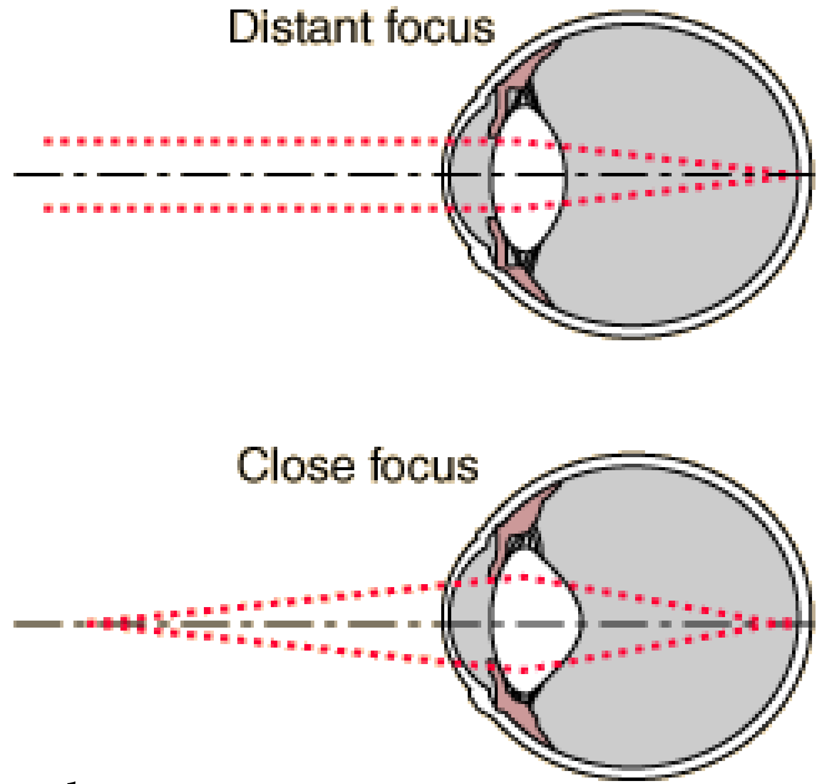
-Iris is the diaphragm that changes the aperture (pupil)

-Retina is the sensor where the fovea has the highest resolution

Focusing

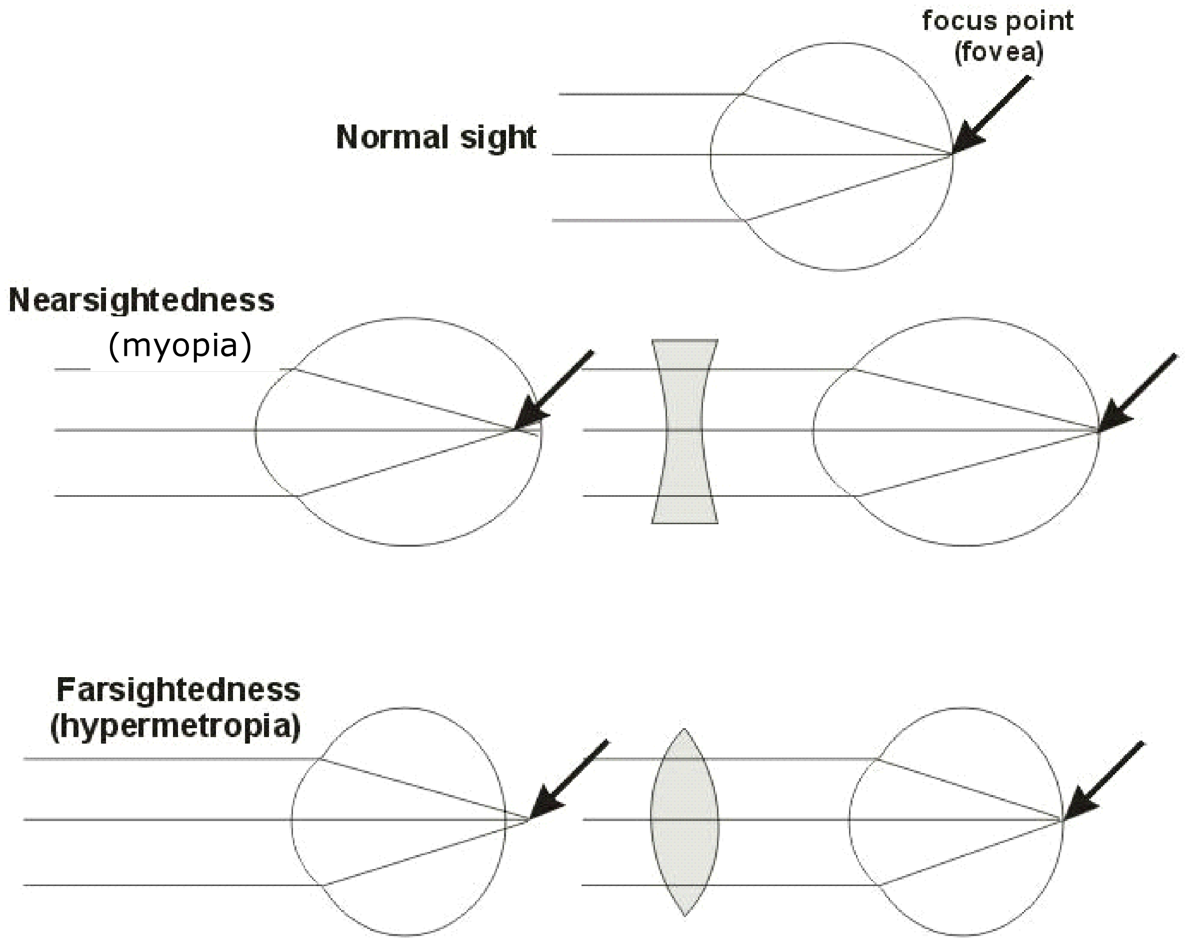


shorter focal length

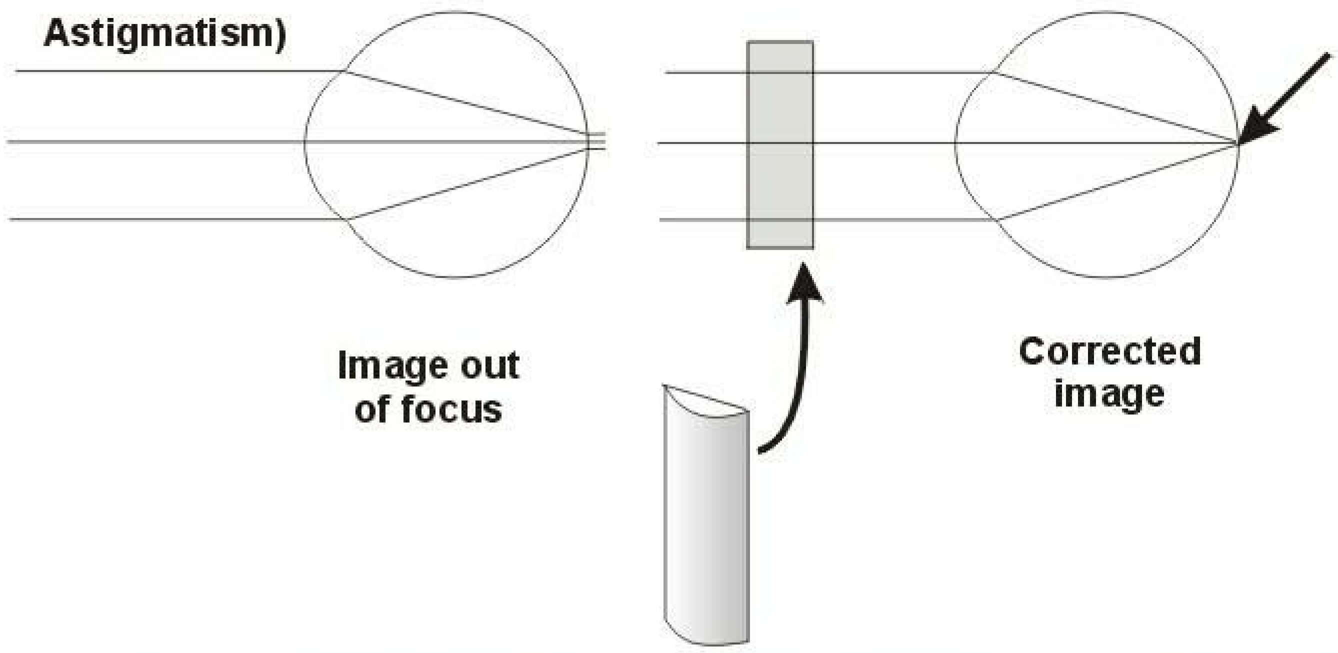


Changes the focal length of the lens

Myopia and Hyperopia



Astigmatism



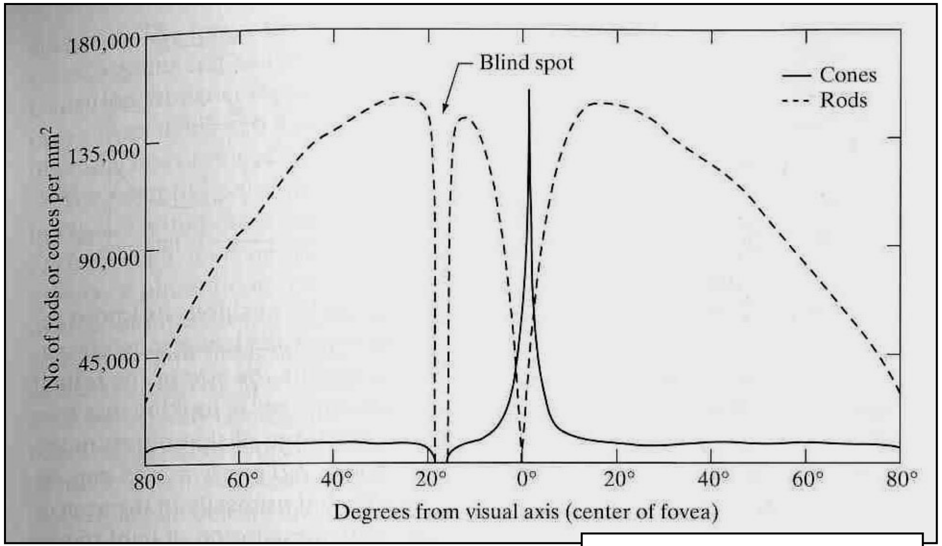
The cornea is distorted causing images to be un-focused on the retina.

Blind Spot in the Eye

Close your right eye and look directly at the “+”

Colour

- Our retina has:
  - **Cones** – Measure the frequency of light (colour)
    - 6 to 7 millions
    - High-definition
    - Need high luminosity
  - **Rods** – Measure the intensity of light (luminance)
    - 75 to 150 millions
    - Low-definition
    - Function with low luminosity



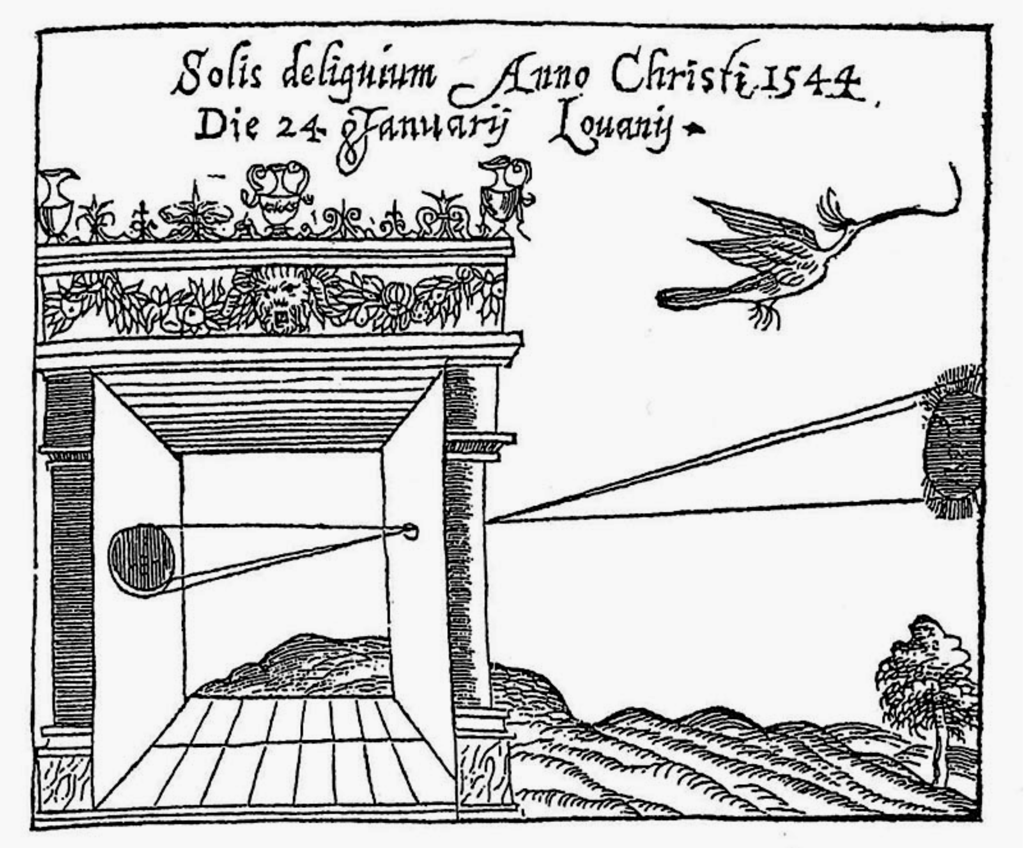
Gonzalez & Woods

We only see colour in the center of our retina!

Topic: Image Capturing Systems

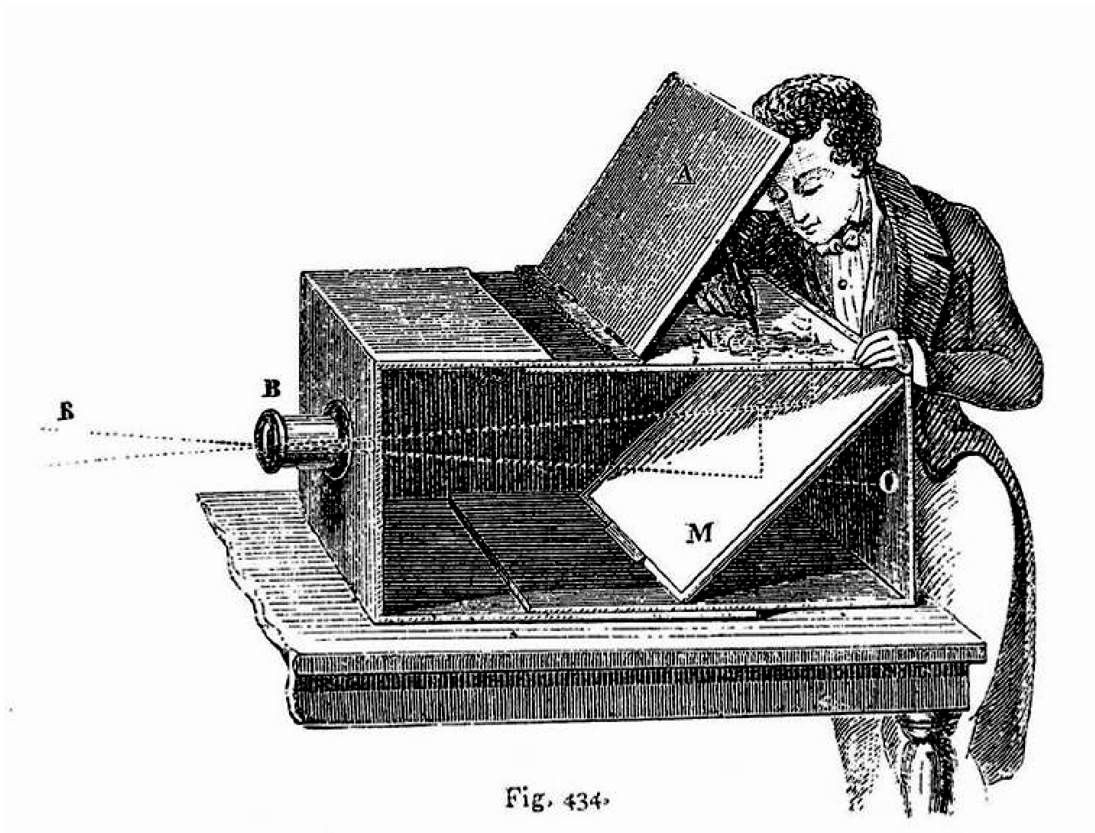
- ‘Computer Vision’?
- The Human Visual System
- Image Capturing Systems

A Brief History of Images



Camera Obscura, Gemma Frisius, 1544

A Brief History of Images

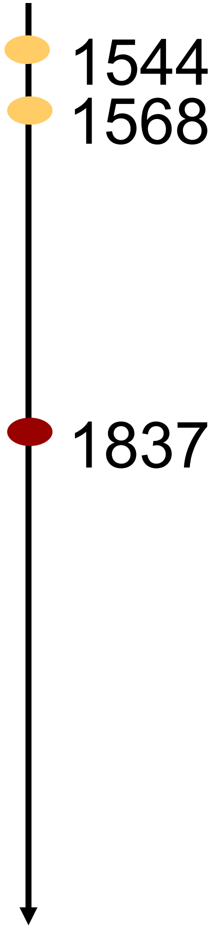


Lens Based Camera Obscura, 1568

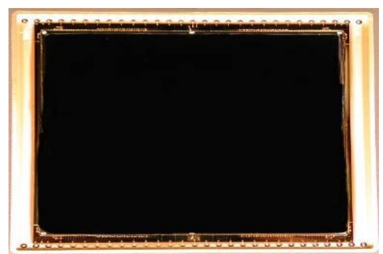
1544
1568

A Brief History of Images

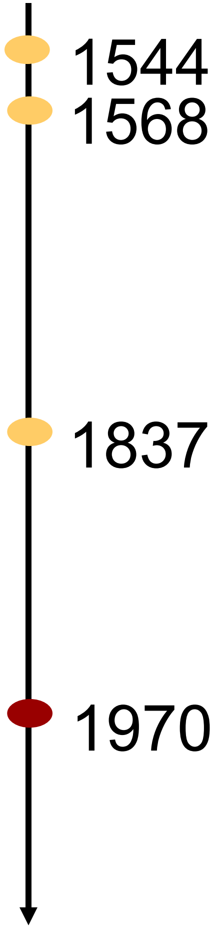
Still Life, Louis Jaques Mande Daguerre, 1837



A Brief History of Images



Silicon Image Detector, 1970

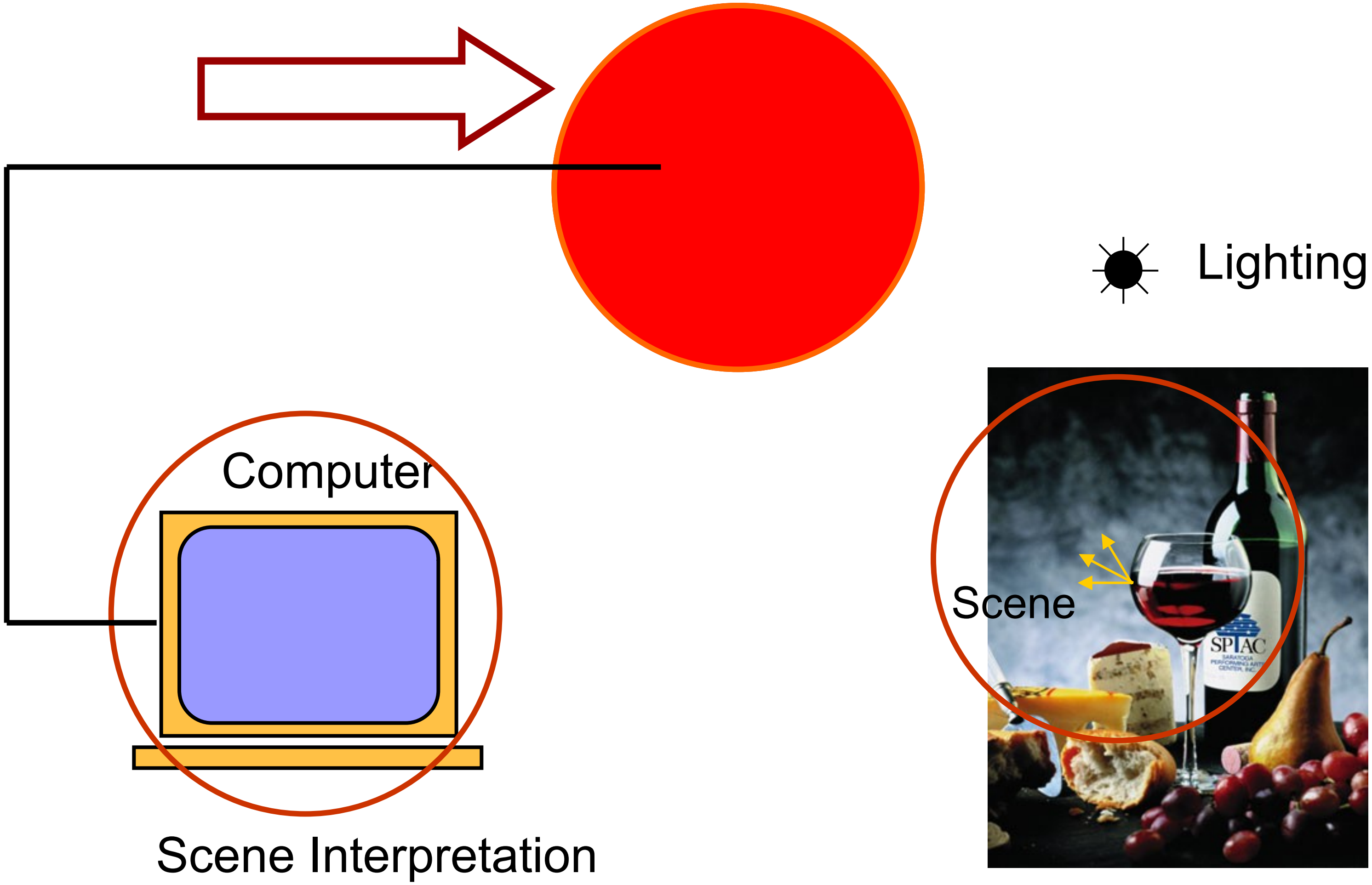


A Brief History of Images

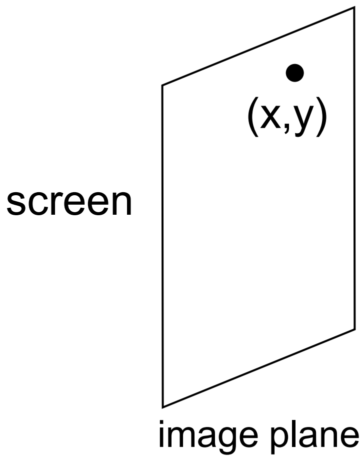
Digital Cameras

1544
1568
1837
1970
1995

Components of a Computer Vision System



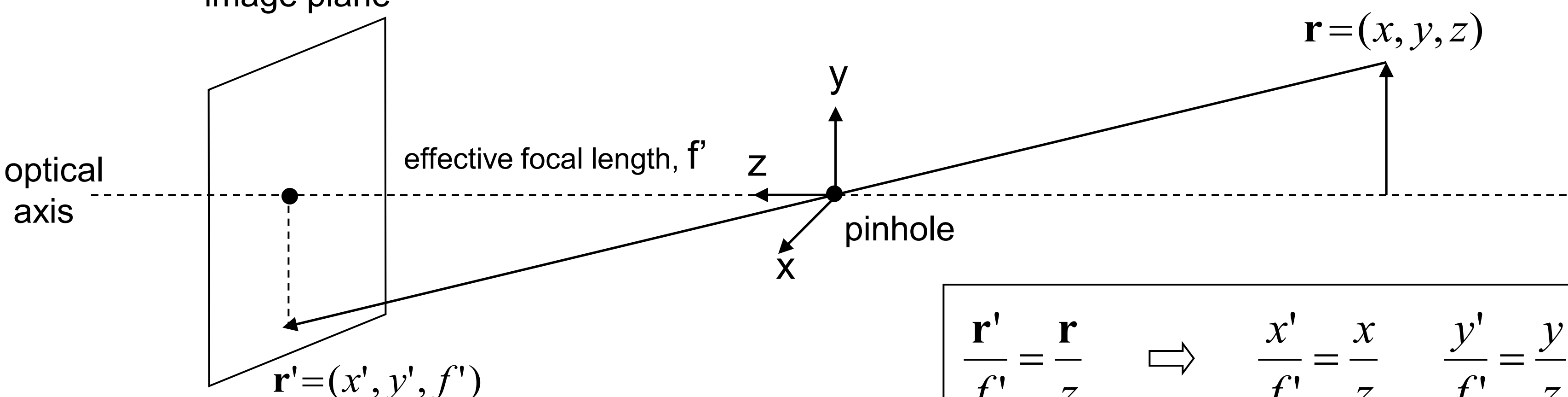
Pinhole and the Perspective Projection



scene

Is an image being formed on the screen?

YES!   But, not a “clear” one.

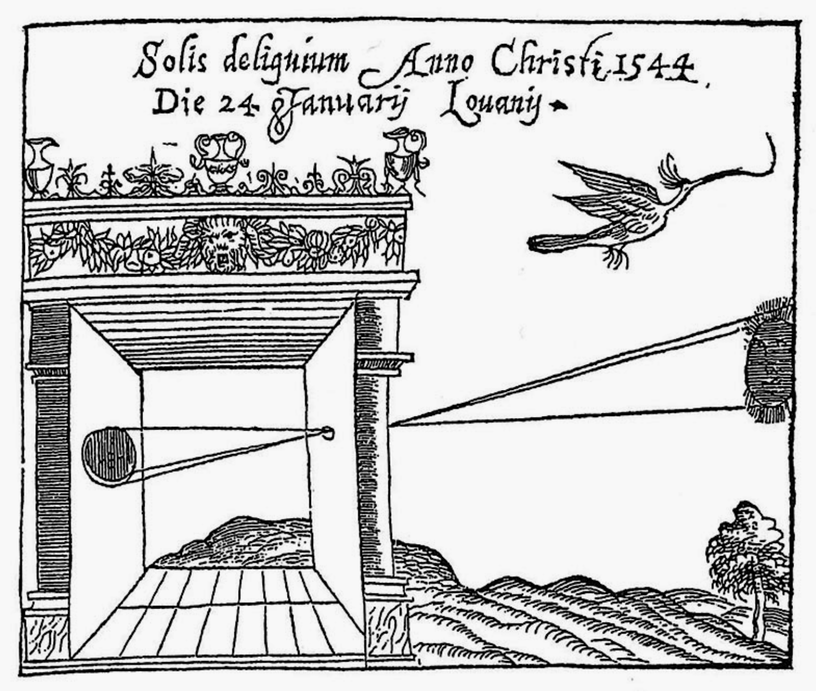


$$\frac{\mathbf{r}'}{f'}=\frac{\mathbf{r}}{z} \quad \Rightarrow \quad \frac{x'}{f'}=\frac{x}{z} \qquad \frac{y'}{f'}=\frac{y}{z}$$

Pinhole Camera

• Basically a pinhole camera is a box, with a tiny hole at one end and film or photographic paper at the other.

• Mathematically: out of all the light rays in the world, choose the set of light rays passing through a point and projecting onto a plane.



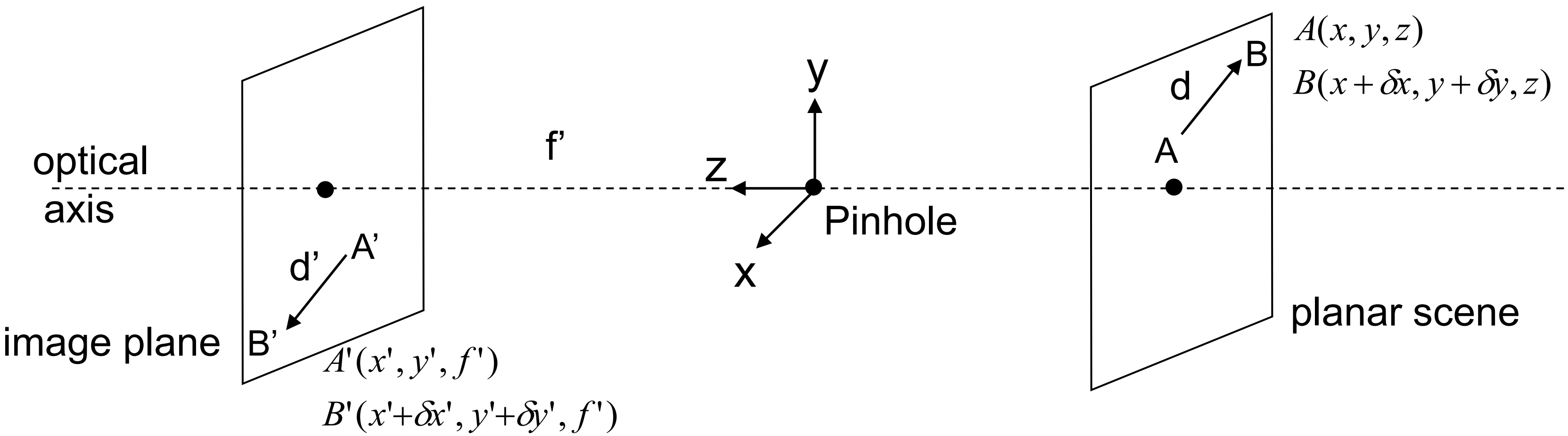
Pinhole Photography

©Charlotte Murray Untitled, 4" x 5" pinhole photograph, 1992

Image Size inversely proportional to Distance

Reading: http://www.pinholeresource.com/

Magnification



From perspective projection:

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

$$\frac{x'+\delta x'}{f'} = \frac{x+\delta x}{z} \qquad \frac{y'+\delta y'}{f'} = \frac{y+\delta y}{z}$$

Magnification:

$$m = \frac{d'}{d} = \frac{\sqrt{(\delta x')^2 + (\delta y')^2}}{\sqrt{(\delta x)^2 + (\delta y)^2}} = \frac{f'}{z}$$

$$\frac{Area_{image}}{Area_{scene}} = m^2$$

Problems with Pinholes

- Pinhole size (aperture) must be "very small" to obtain a clear image.
- However, as pinhole size is made smaller, less light is received by image plane.
- If pinhole is comparable to wavelength λ of incoming light, DIFFRACTION blurs the image!
- Sharpest image is obtained when:

pinhole diameter $d = 2\sqrt{f'\lambda}$

Example: If *f' = 50mm,*

= *600nm* (red),

d = 0.36mm

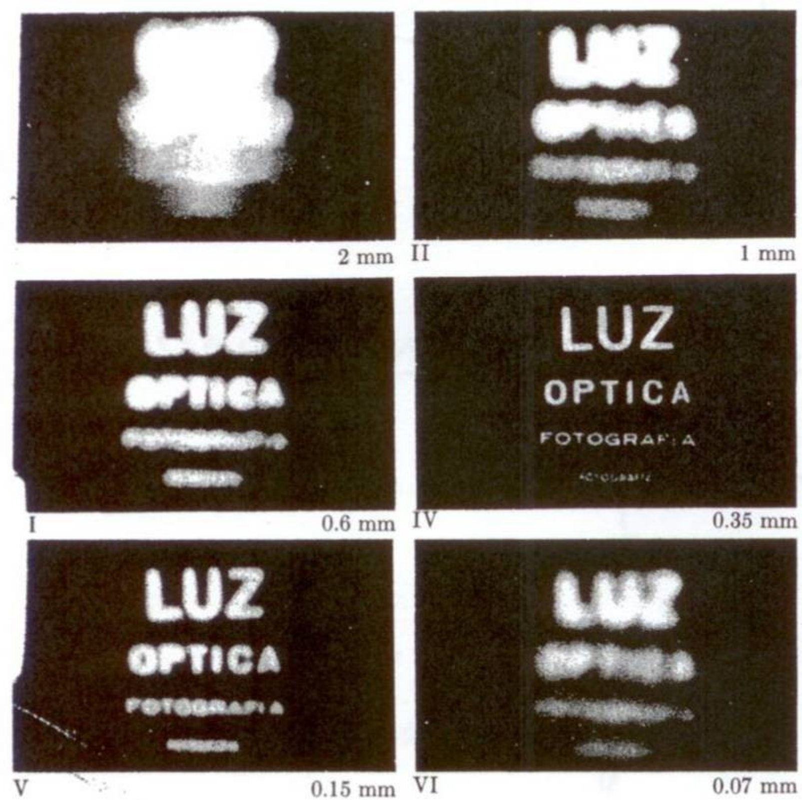
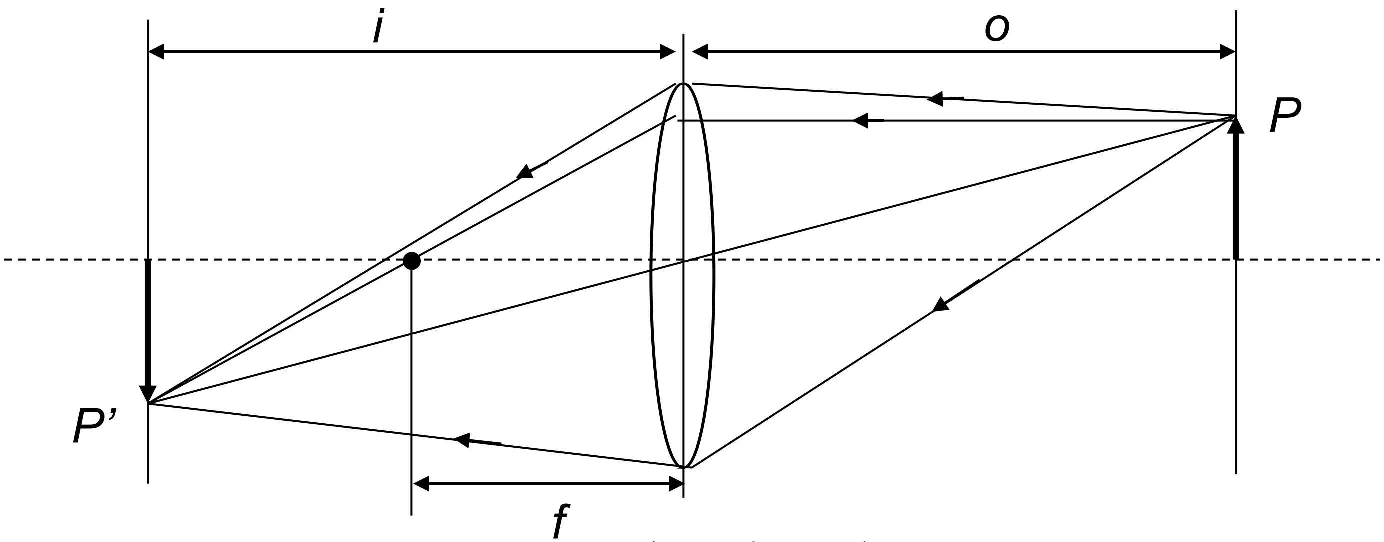


Fig. 5.96 The pinhole camera. Note the variation in image clarity as the hole diameter decreases. [Photos courtesy Dr. N. Joel, UNESCO.]

Image Formation using Lenses

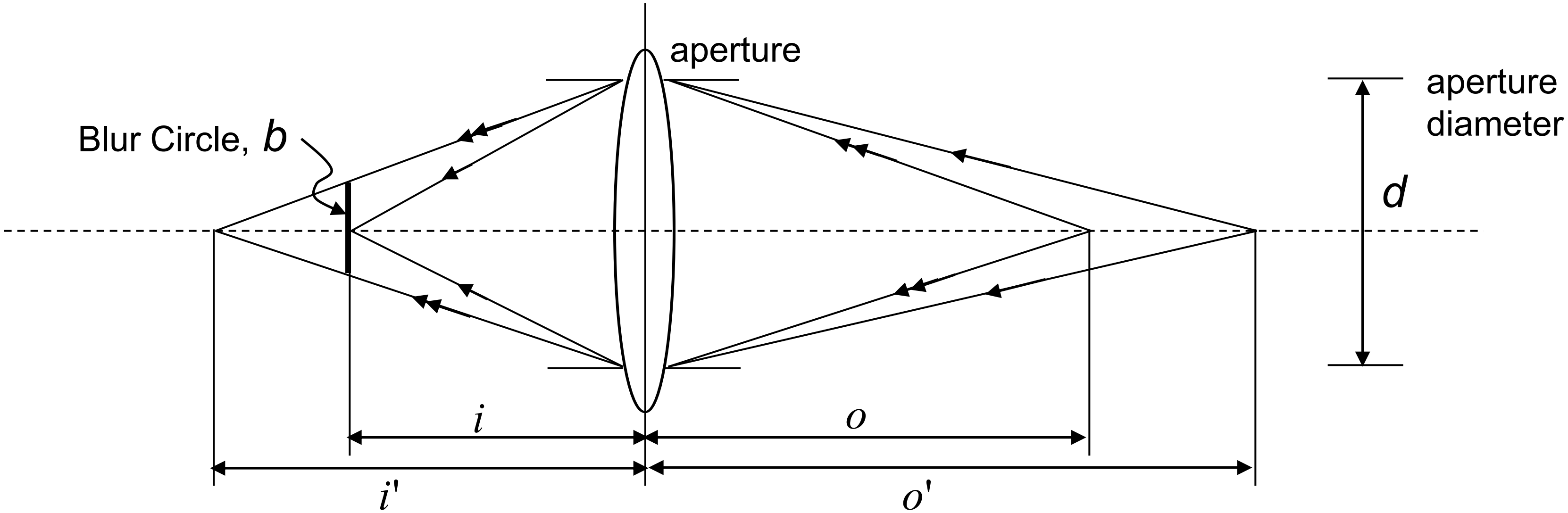
- Lenses are used to avoid problems with pinholes.
- Ideal Lens: Same projection as pinhole but gathers more light!



- Gaussian Thin Lens Formula: $\frac{1}{i} + \frac{1}{o} = \frac{1}{f}$

- *f* is the focal length of the lens – determines the lens's ability to refract light

Focus and Defocus



- Gaussian Law:

$$\frac{1}{i} + \frac{1}{o} = \frac{1}{f}$$

$$\frac{1}{i'} + \frac{1}{o'} = \frac{1}{f}$$

$$\Rightarrow \quad (i' - i) = \frac{f}{(o' - f)} \frac{f}{(o - f)} (o - o')$$

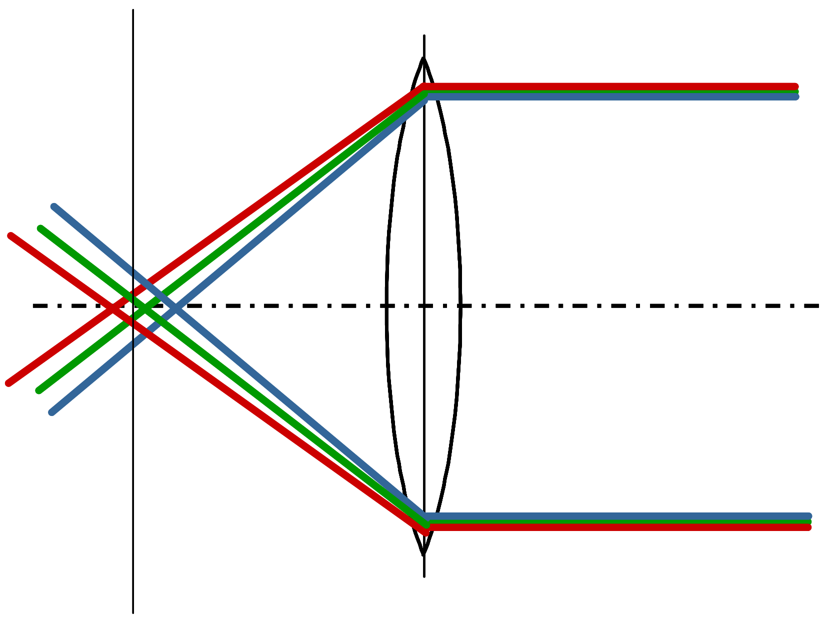
- In theory, only one scene plane is in focus.

Depth of Field

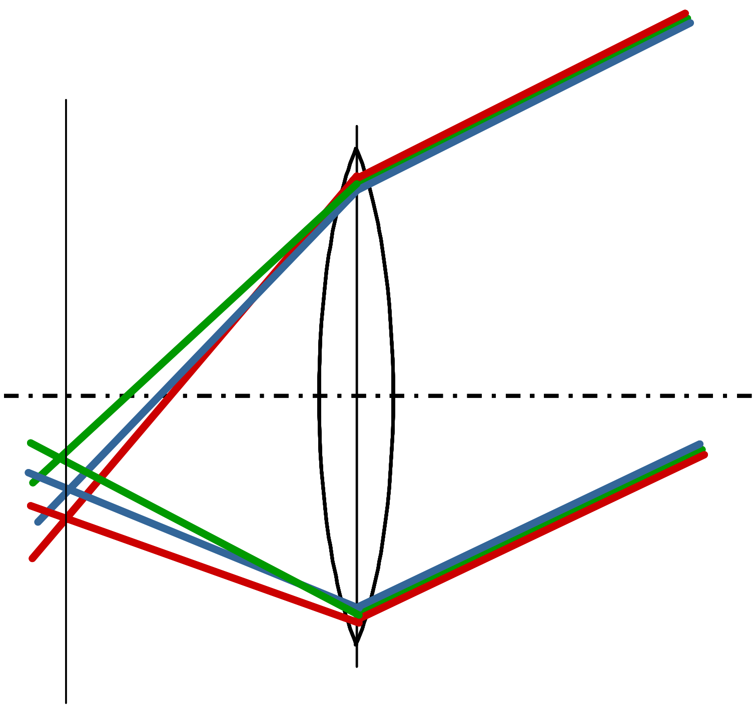
- Range of object distances over which image is <u>sufficiently well</u> focused.
- Range for which *blur circle* is less than the resolution of the sensor.

http://images.dpchallenge.com/images_portfolio/27920/print_preview/116336.jpg

Chromatic Aberration



longitudinal chromatic aberration (axial)



transverse chromatic aberration (lateral)

Image Sensors

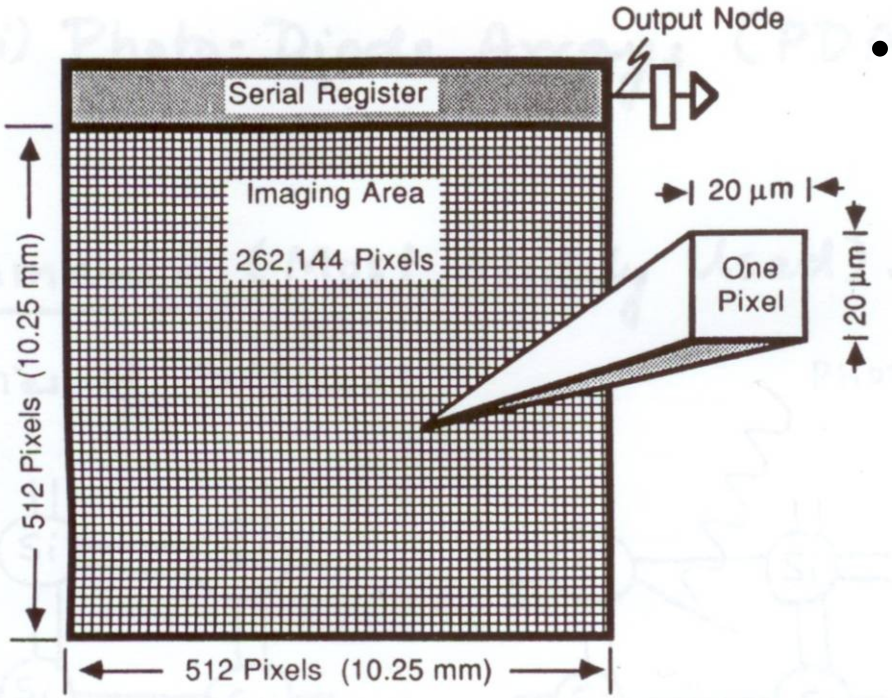
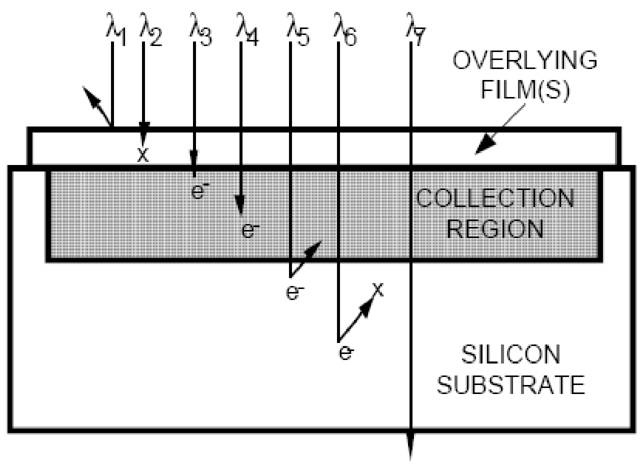


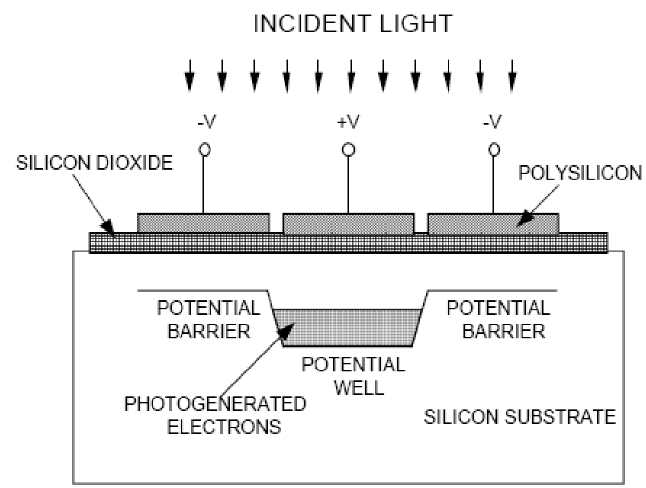
FIG. 4. Typical 512 × 512 CCD.

- Considerations
  - Speed
  - Resolution
  - Signal / Noise Ratio
  - Cost

Image Sensors

- Convert light into an electric charge



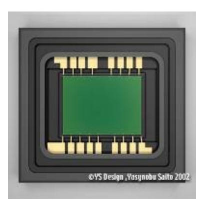


CCD (charge coupled device)

Higher dynamic range

High uniformity

Lower noise

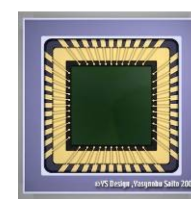


CMOS (complementary metal Oxide semiconductor)

Lower voltage

Higher speed

Lower system complexity

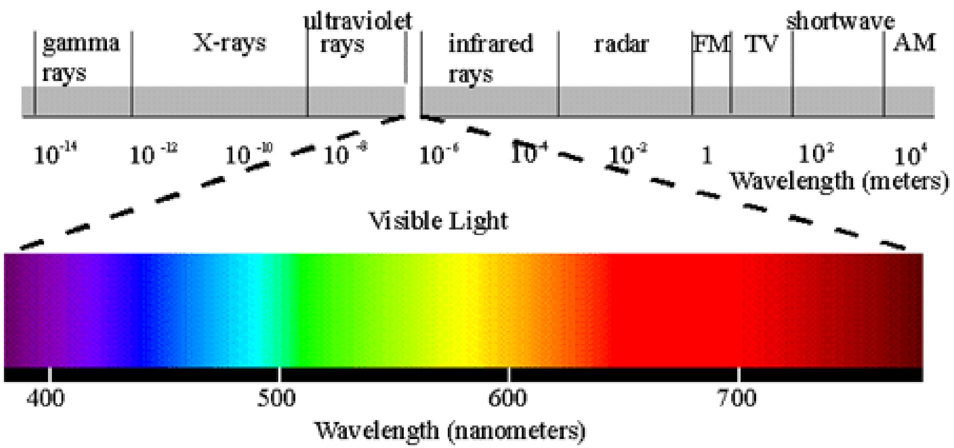


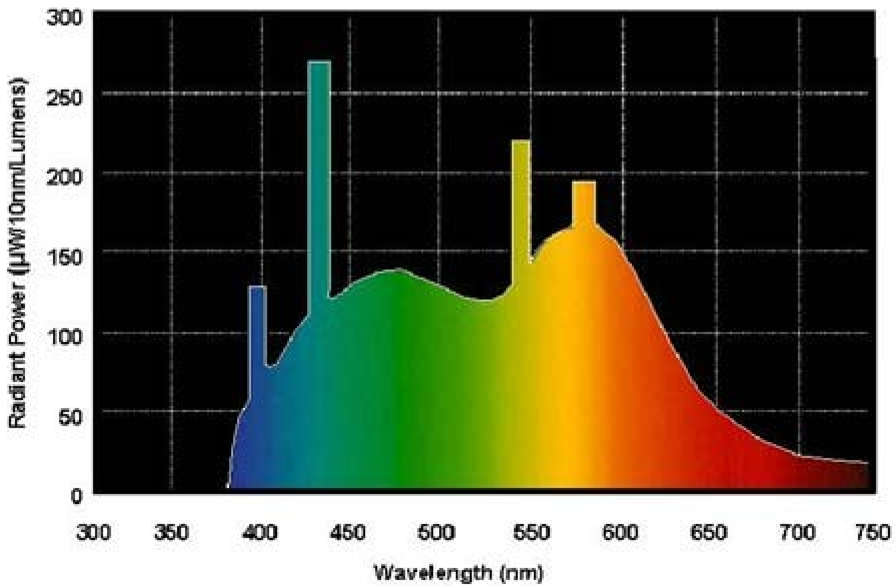
CCD Performance Characteristics

- Linearity Principle: Incoming photon flux vs. Output Signal
  - Sometimes cameras are made non-linear on purpose.
  - Calibration must be done (using reflectance charts)---covered later
- Dark Current Noise: Non-zero output signal when incoming light is zero
- Sensitivity: Minimum detectable signal produced by camera

Sensing Brightness

Incoming light has a **spectral distribution** $p(\lambda)$



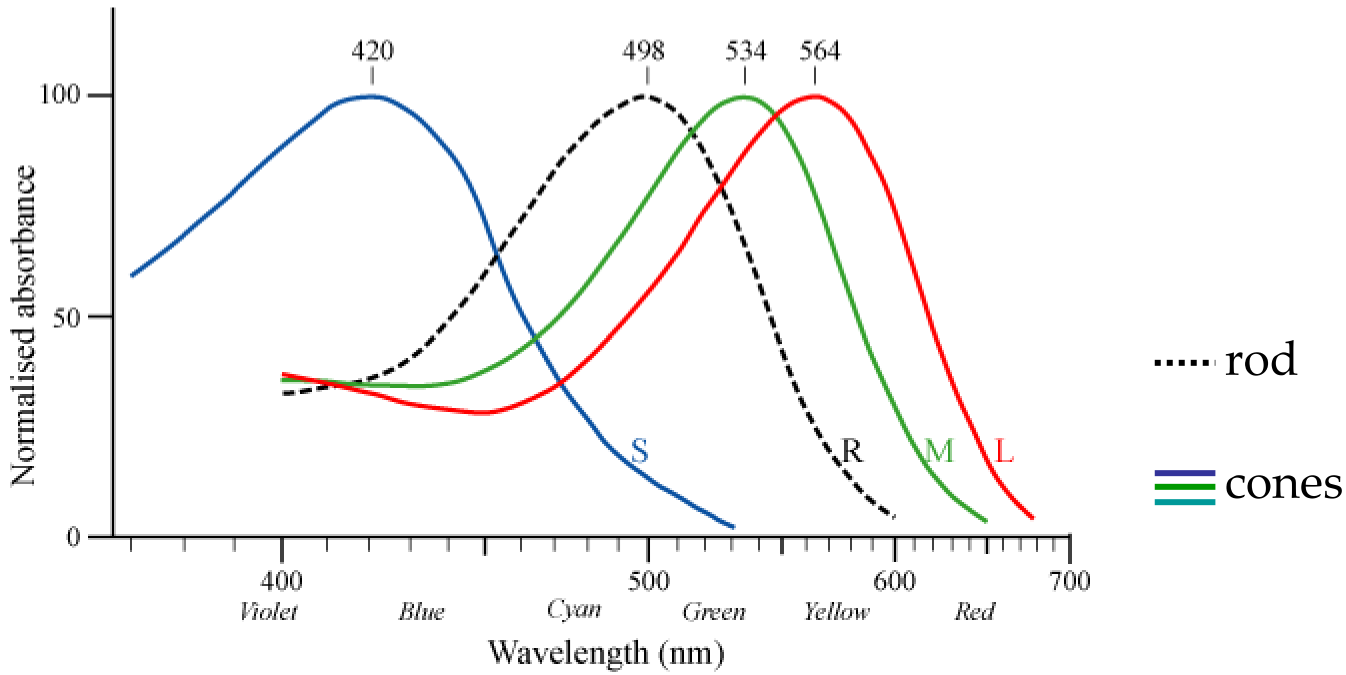


So the pixel intensity becomes

$$I = k \int_{-\infty}^{\infty} q(\lambda) p(\lambda) d\lambda$$

How do we sense colour?

- Do we have infinite number of filters?

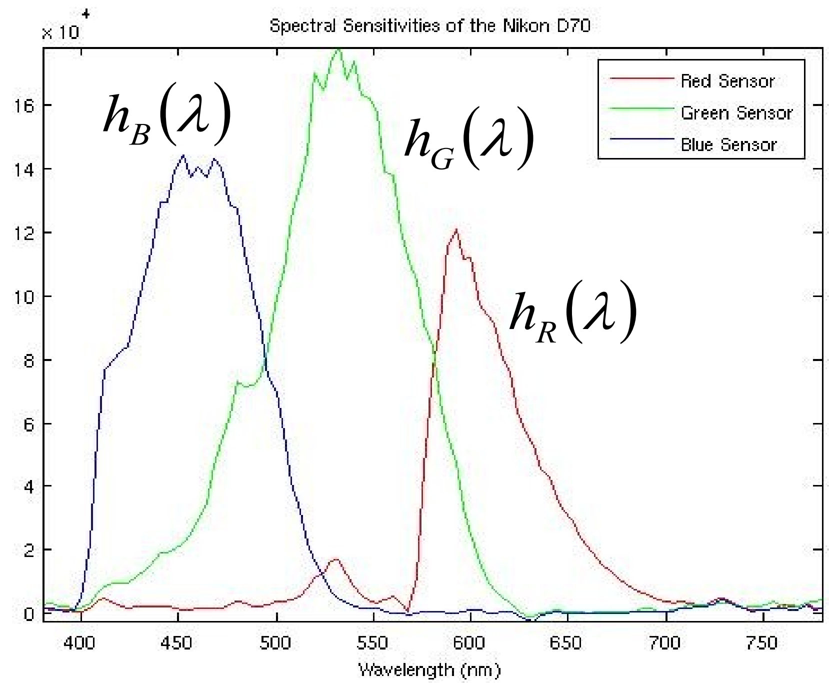


Three filters of different spectral responses

Sensing Colour

- Tristimulus (trichromatic) values (I_R, I_G, I_B)

Camera's spectral response functions: $h_R(\lambda), h_G(\lambda), h_B(\lambda)$

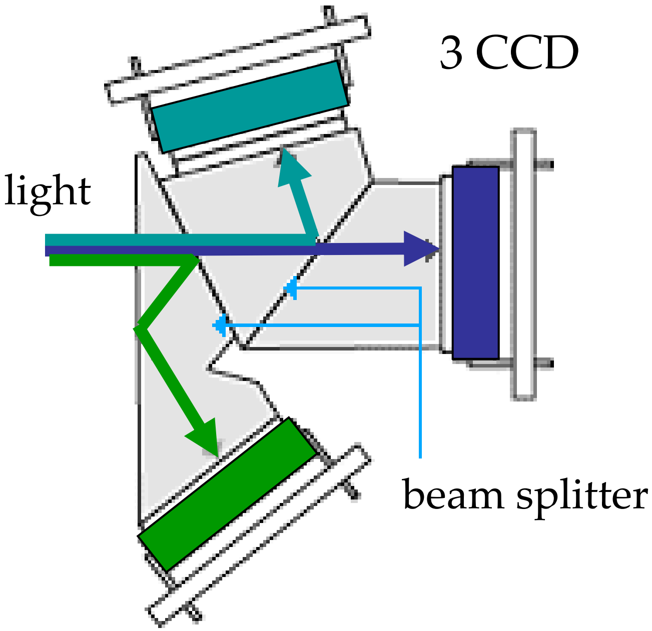


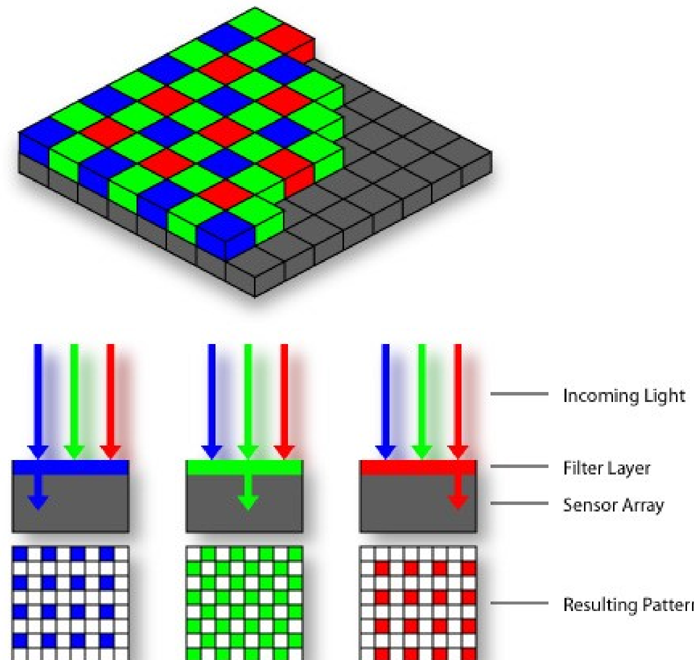
$$I_R = k \int_{-\infty}^{\infty} h_R(\lambda) p(\lambda) d\lambda$$

$$I_G = k \int_{-\infty}^{\infty} h_G(\lambda) p(\lambda) d\lambda$$

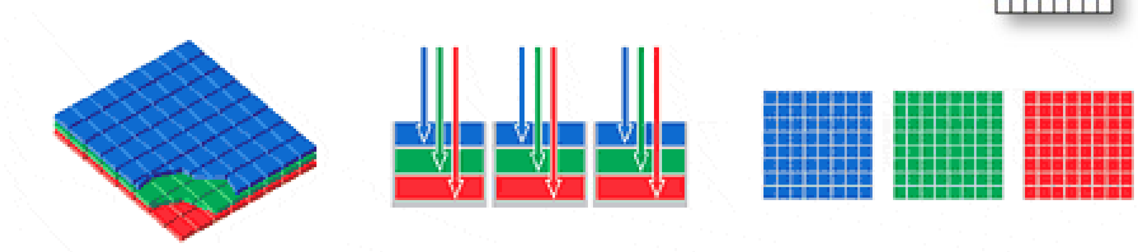
$$I_B = k \int_{-\infty}^{\infty} h_B(\lambda) p(\lambda) d\lambda$$

Sensing Colour





Bayer pattern



Foveon X3™

Resources

- J.C. Russ – Chapters 1 and 2
- L. Shapiro, and G. Stockman – Chapter 1
- "Color Vision: One of Nature's Wonders" in http://www.diycalculator.com/sp-cvision.shtml