Johns Hopkins Engineering

Computer Vision

Image Formation

JOHNS HOPKINS
WHITING SCHOOL
of ENGINEERING

Announcements

- Lecture 1 slides and video uploaded
- Linear algebra review slides posted
- Quiz on Tuesday, in class, 5 questions based on lectures from this week

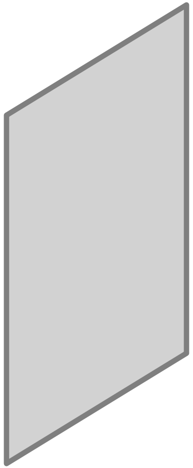
Today’s Topics

- Complex terrain navigation research
- Image formation and optics
- Leave time for anyone who wants to discuss projects

Image Formation and Optics

- **Image:** Projection of 3D scene onto 2D plane
- We need to understand **Geometric** and **Radiometric** relations between the scene and its image
- This minilecture will cover the basic properties of pinhole projection

Image Formation



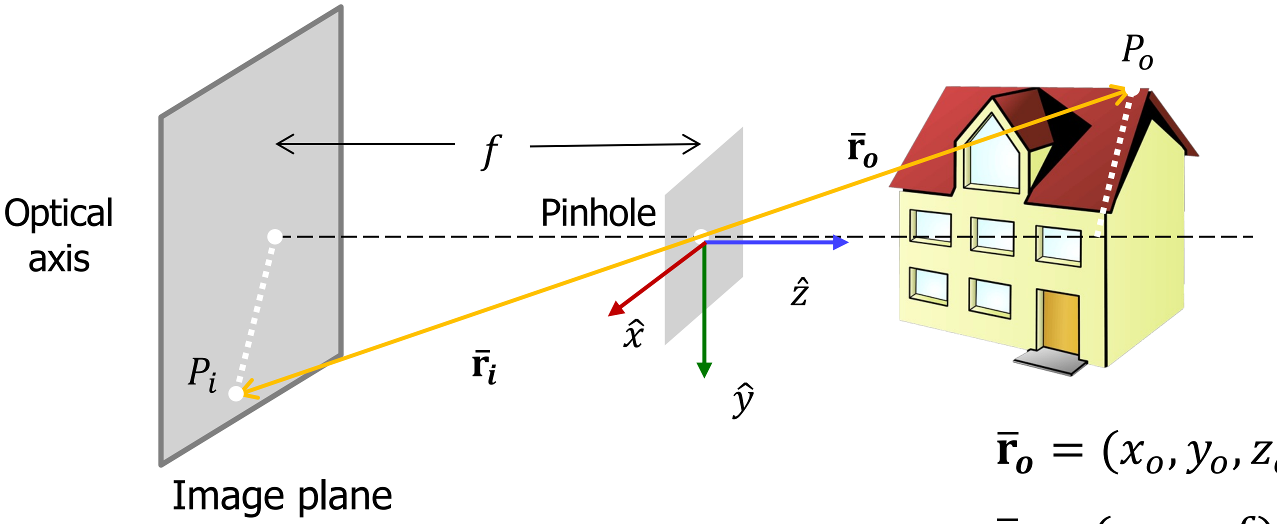
Screen

Scene

Is an image being formed on the screen?

Yes! But not a “clear” one.

Perspective Imaging with Pinhole



$$\bar{\mathbf{r}}_o = (x_o, y_o, z_o)$$

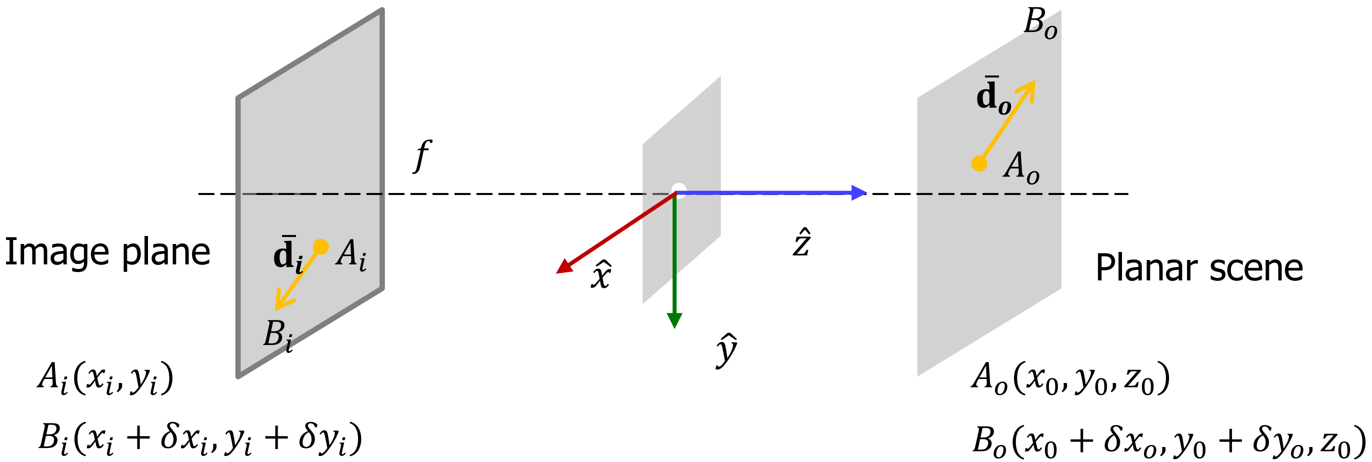
$$\bar{\mathbf{r}}_i = (x_i, y_i, f)$$

Using similar triangles:

$$\boxed{\frac{\bar{\mathbf{r}}_i}{f} = \frac{\bar{\mathbf{r}}_o}{z_o}} \rightarrow \boxed{x_i = f\frac{x_o}{z_o}, y_i = f\frac{y_o}{z_o}}$$

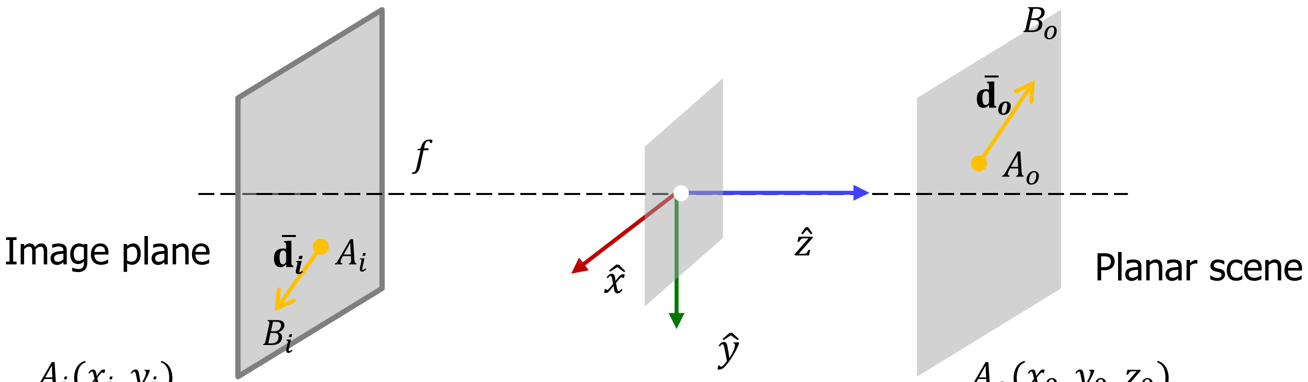
f: Effective Focal Length

Image Magnification



Magnification: $$|m| = \frac{\|\bar{\mathbf{d}}_i\|}{\|\bar{\mathbf{d}}_o\|} = \sqrt{{\delta x_i}^2 + {\delta y_i}^2} \Big/ \sqrt{{\delta x_o}^2 + {\delta y_o}^2}$$

Image Magnification



$A_i(x_i, y_i)$

$B_i(x_i + \delta x_i, y_i + \delta y_i)$

$A_o(x_0, y_0, z_0)$

$B_o(x_0 + \delta x_o, y_0 + \delta y_o, z_0)$

From Perspective Projection:

$$\frac{x_i}{f} = \frac{x_o}{z_o} \quad \text{and} \quad \frac{y_i}{f} = \frac{y_o}{z_o} \quad \text{------------------------} \quad \text{(A)}$$

$$\frac{x_i + \delta x_i}{f} = \frac{x_o + \delta x_o}{z_o} \quad \text{and} \quad \frac{y_i + \delta y_i}{f} = \frac{y_o + \delta y_o}{z_o} \quad \text{--------} \quad \text{(B)}$$

Image Magnification

$$\delta x_i = f\frac{\delta x_o}{z_o} \quad \text{and} \quad \delta y_i = f\frac{\delta y_o}{z_o}$$

Magnification:

$$|m| = \frac{\|\bar{\mathbf{d}}_i\|}{\|\bar{\mathbf{d}}_o\|} = \sqrt{{\delta x_i}^2 + {\delta y_i}^2} \Big/ \sqrt{{\delta x_o}^2 + {\delta y_o}^2} = \left|\frac{f}{z_o}\right|$$

$$\boxed{m = \frac{f}{z_o}}$$

Image Magnification

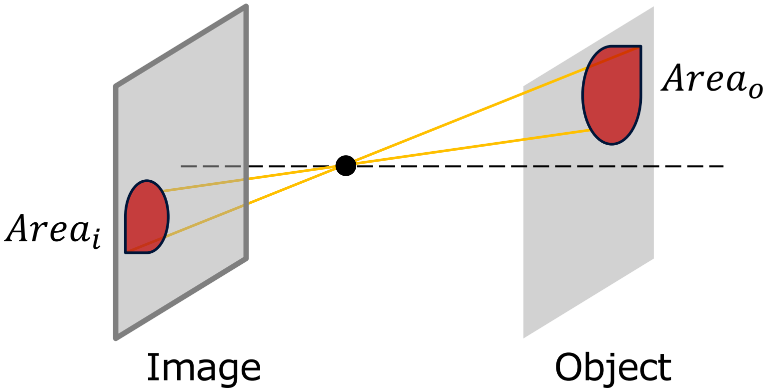
$$m = \frac{f}{z_o}$$

Image size **inversely proportional** to depth

Image Magnification

- Notes:
  - We assumed planar surface
  - m can be assumed to be constant if the range of scene depth Δz is much smaller than the average scene depth $\tilde{z}$

$$\frac{Area_i}{Area_o} = m^2$$



Vanishing Point

What if we do have a lot of changes in the depth of the scene?

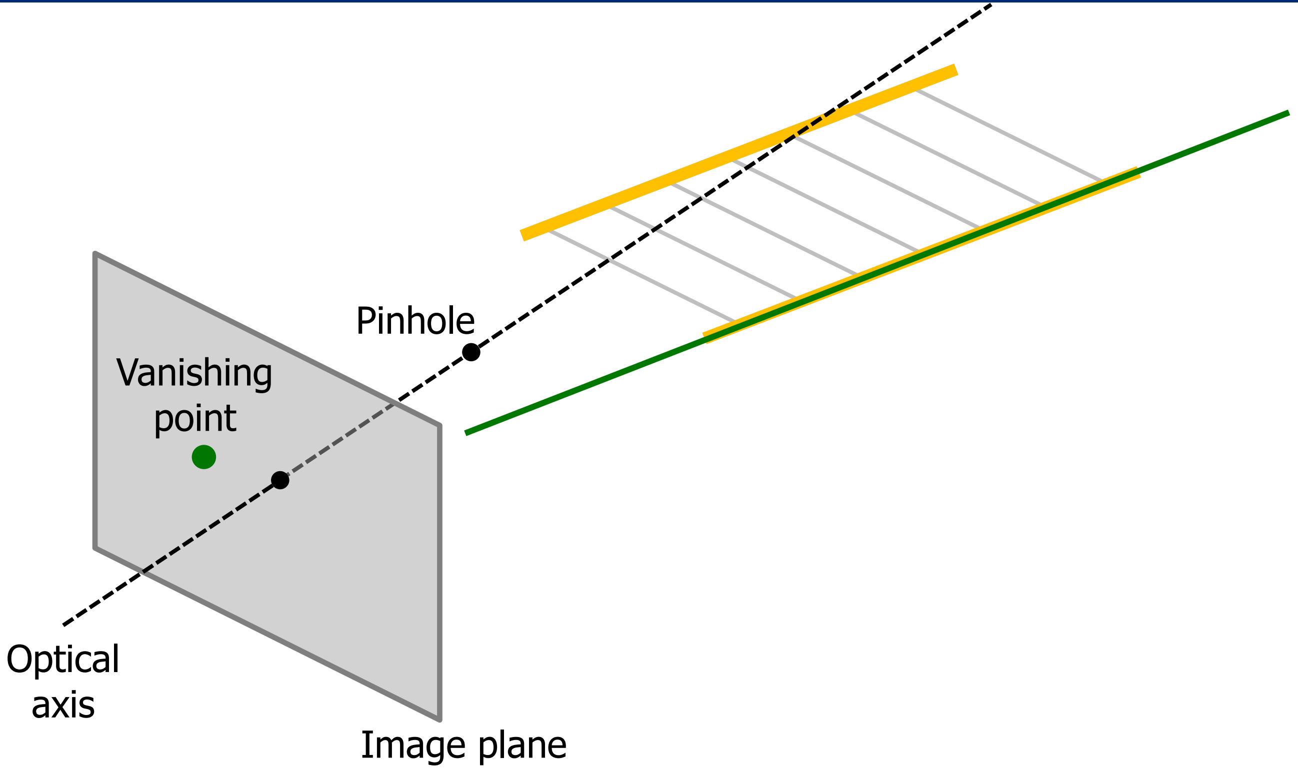
Vanishing Point

Parallel straight lines converge at a single image point

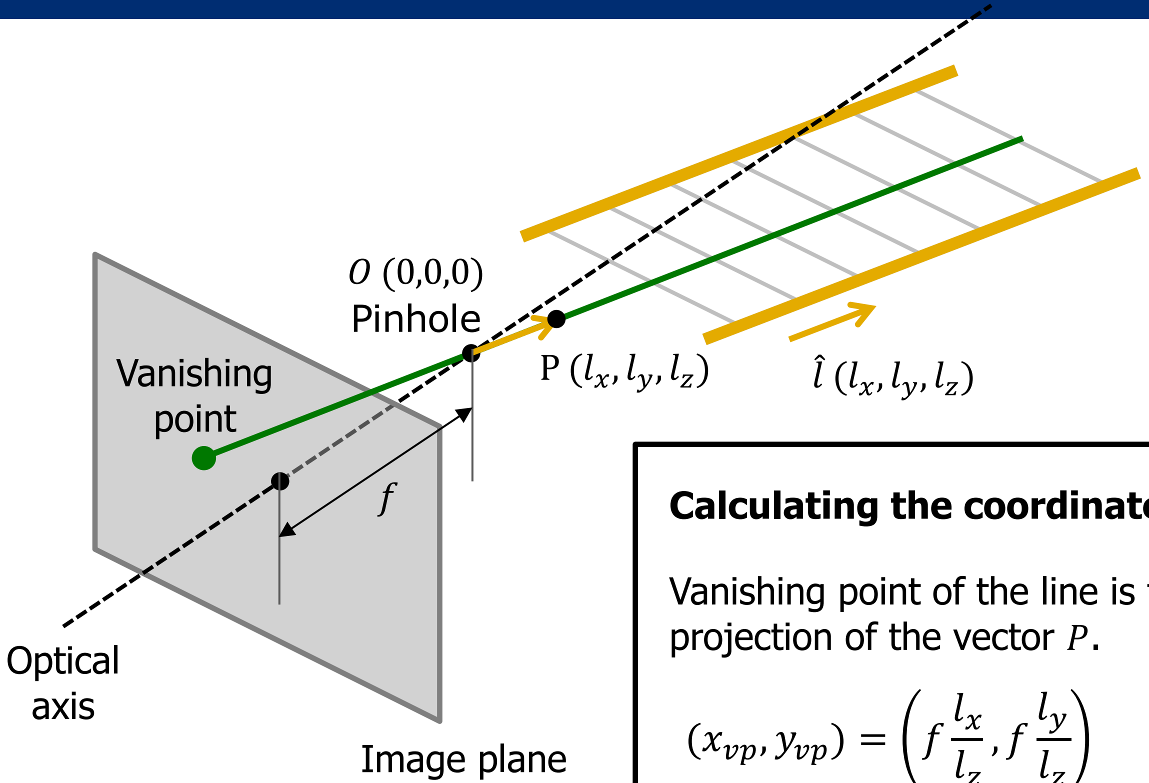
Vanishing Point

Location of Vanishing Point depends on the **orientation** of parallel straight lines.

Finding the Vanishing Point



Finding Vanishing Point



Calculating the coordinates:

Vanishing point of the line is the projection of the vector P.

$$(x_{vp}, y_{vp}) = \left(f\frac{l_x}{l_z}, f\frac{l_y}{l_z}\right)$$

Summary

- **Image:** Projection of 3D scene onto 2D plane
- We need to understand **Geometric** and **Radiometric** relations between the scene and its image
- Essential concepts in this lecture:
  - Pinhole projection and how to express it
  - Properties of projection: magnification and vanishing point

Johns Hopkins Engineering

Computer Vision

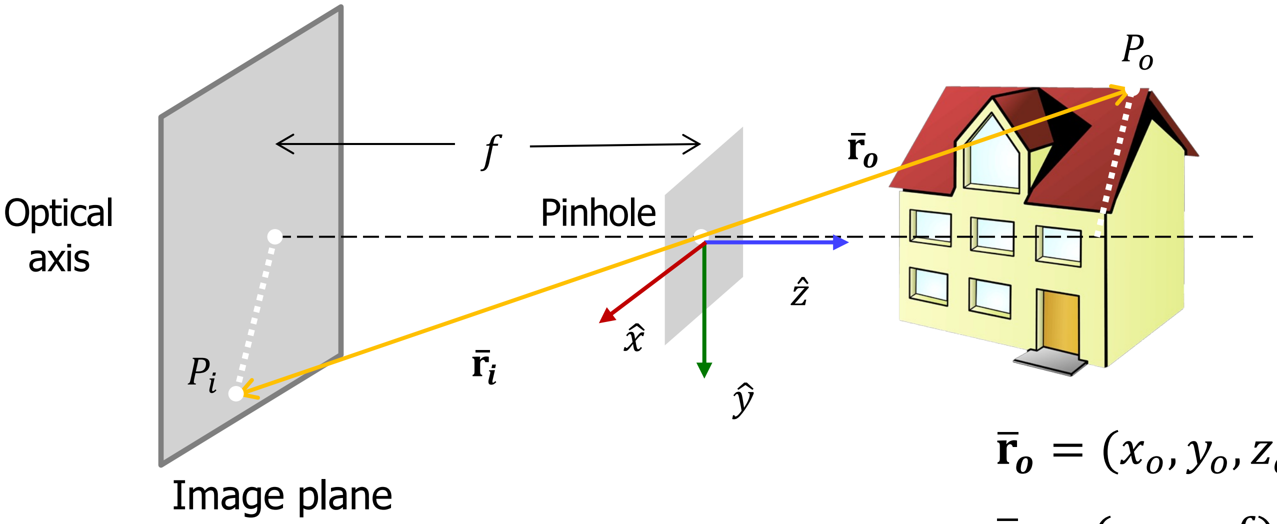
Image Formation

JOHNS HOPKINS
WHITING SCHOOL
of ENGINEERING

Image Formation and Optics

- **Image:** Projection of 3D scene onto 2D plane
- We need to understand **Geometric** and **Radiometric** relations between the scene and its image
- This minilecture will cover the basic concepts of lenses and image formation

Perspective Imaging with Pinhole



$$\bar{\mathbf{r}}_o = (x_o, y_o, z_o)$$

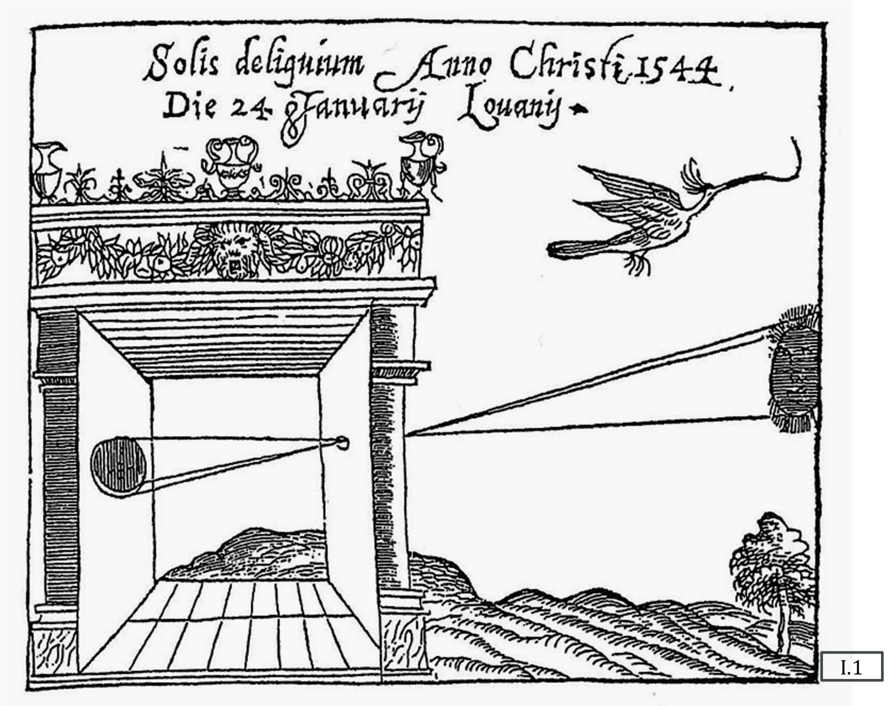
$$\bar{\mathbf{r}}_i = (x_i, y_i, f)$$

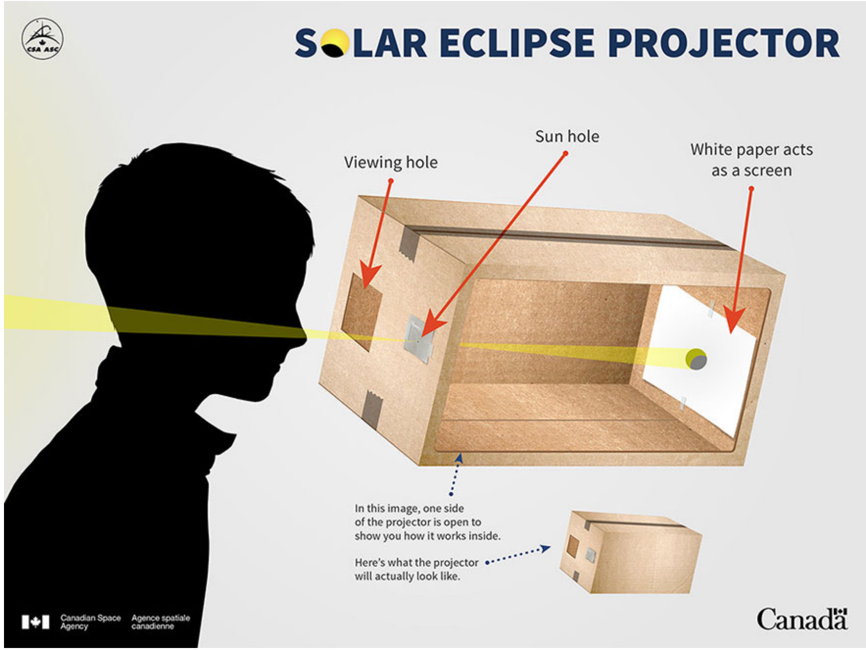
Using similar triangles:

$$\boxed{\frac{\bar{\mathbf{r}}_i}{f} = \frac{\bar{\mathbf{r}}_o}{z_o}} \rightarrow \boxed{x_i = f\frac{x_o}{z_o},\quad y_i = f\frac{y_o}{z_o}}$$

f: Effective Focal Length

This is an Old Idea: Camera Obscura





"Dark Chamber"

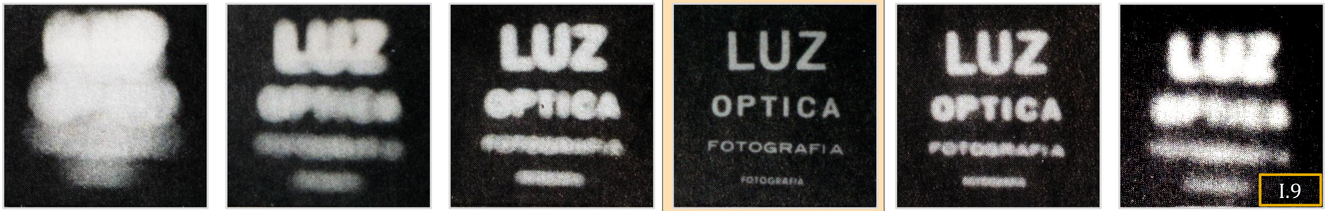
Frisius 1545

Pinhole Eye of *Nautilus pompilius*

I.2

I.3

What is the Ideal Pinhole Size?

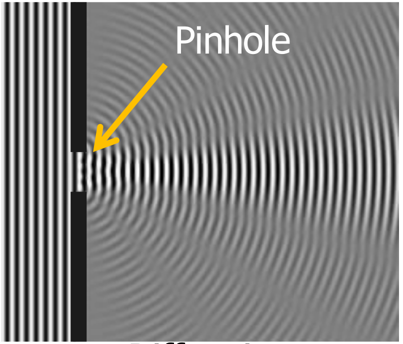


2 mm | 1 mm | 0.6 mm | 0.35 mm | 0.15 mm | 0.07 mm

The pinhole must be tiny,
but if it's too tiny it will cause diffraction.

Ideal pinhole diameter: $\boxed{d \approx 2\sqrt{f\lambda}}$

f: effective focal length
λ: wavelength



Diffraction

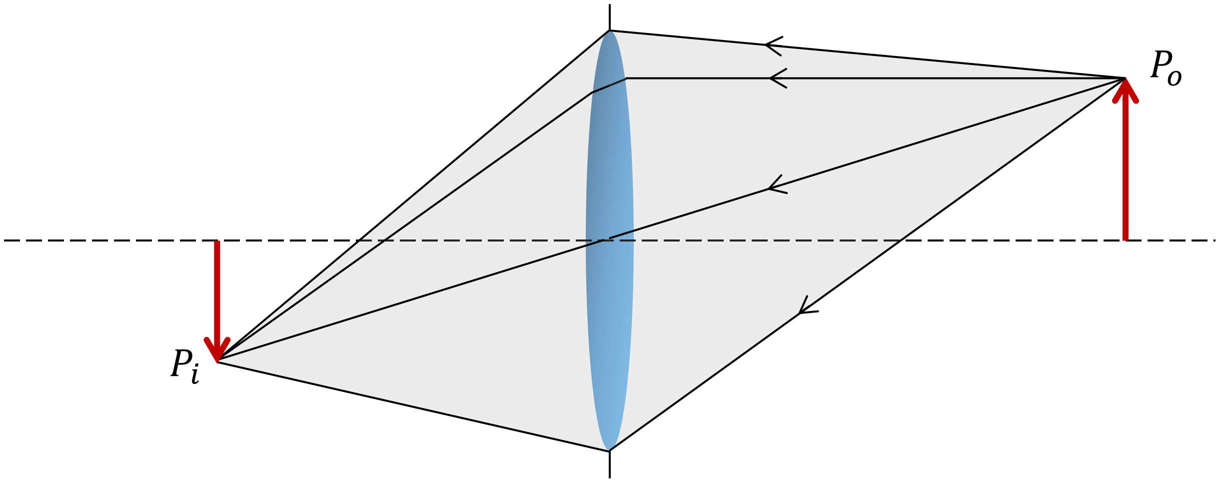
What about Exposure Time?

Pinholes pass less light and hence require **long exposures** to capture bright images.

$f = 73\ mm,\ d = 0.2\ mm,$
Exposure, $\boldsymbol{T = 12\ s}$

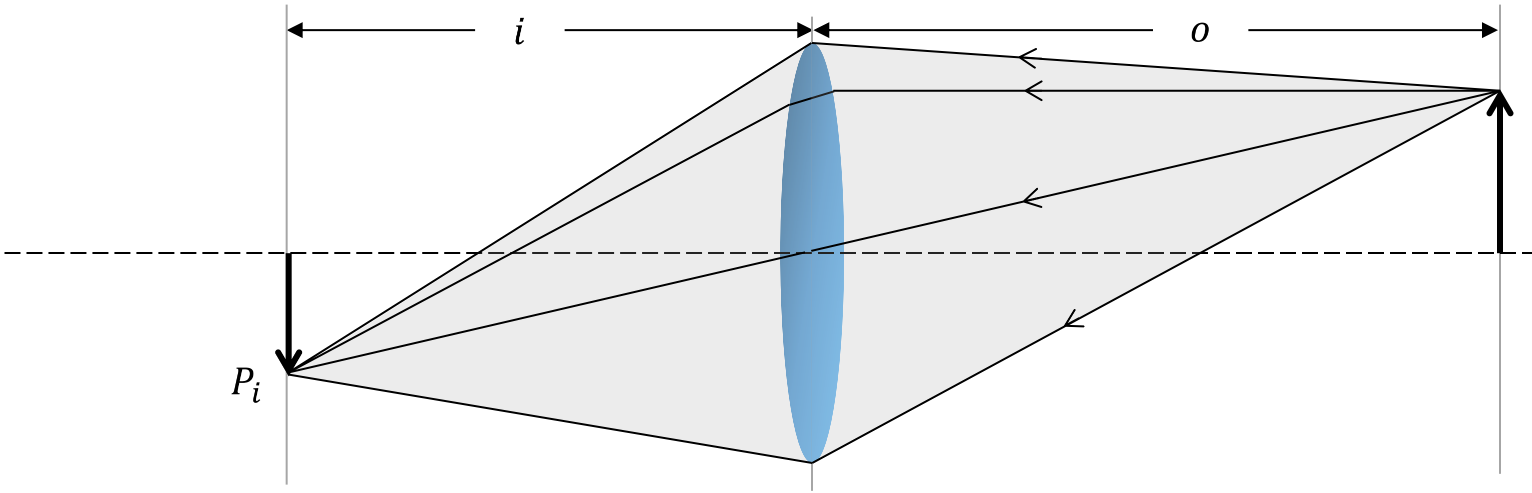
Lenses

- Same projection as pinhole but gather more light!



Focal length (f) determines the lens' bending power

Gaussian Lens (Thin Lens) Law



f: focal length

i: image distance

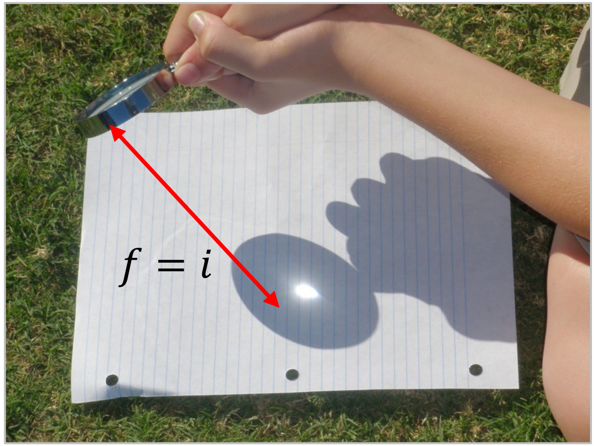
o: object distance

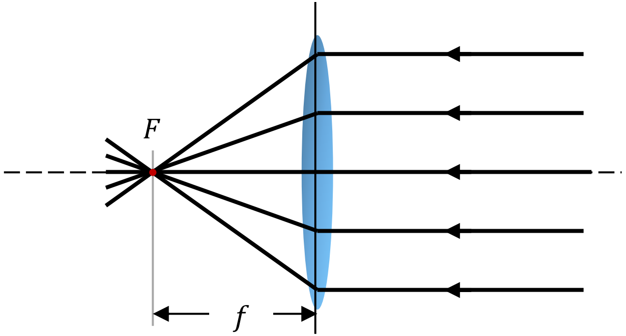
$$\boxed{\frac{1}{i} + \frac{1}{o} = \frac{1}{f}}$$

Example: If $f = 50mm$ & $o = 300mm$, then image distance $i = 60mm$

How to Find the Focal Length

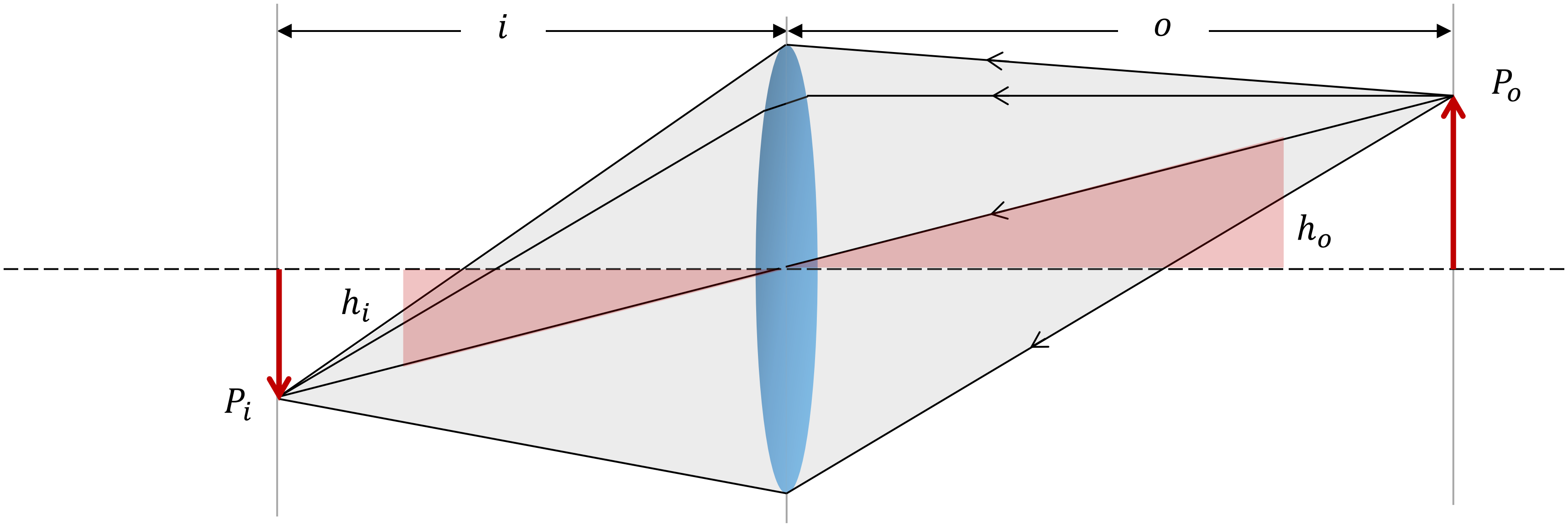
$$\frac{1}{i} + \frac{1}{o} = \frac{1}{f} \quad \Rightarrow \quad \text{If } o = \infty\text{, then } f = i$$





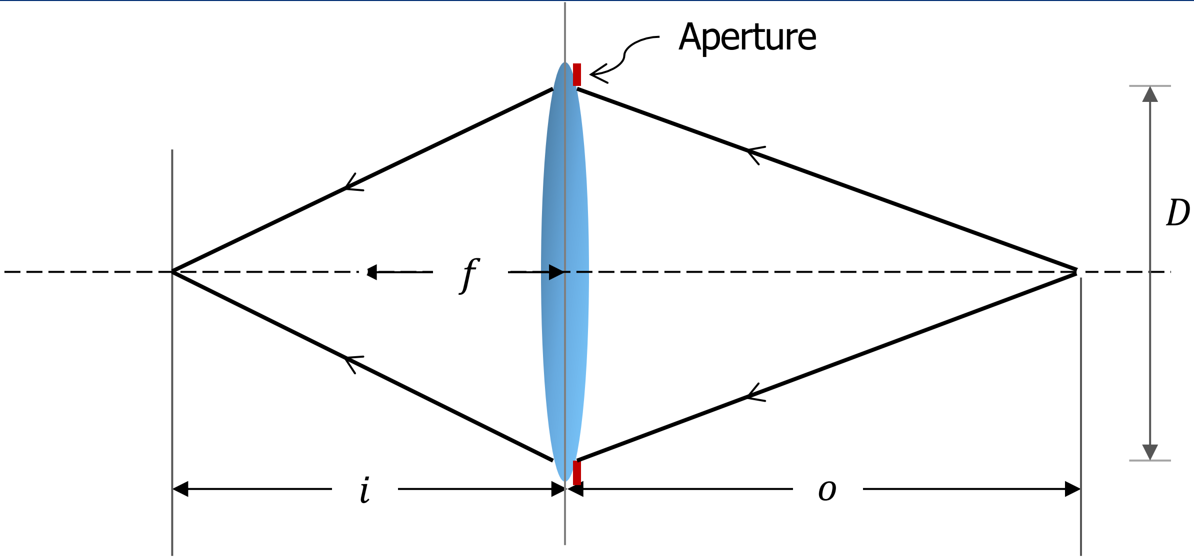
Focal length: Distance at which incoming rays that are parallel to the optical axis converge.

Image Magnification

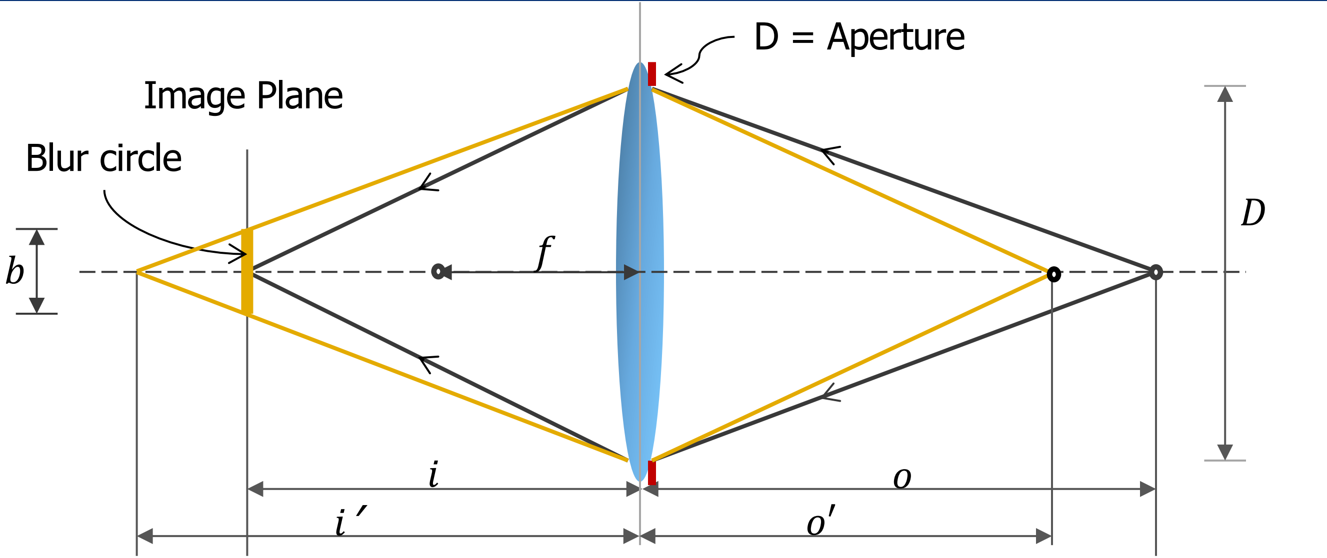


Magnification: $$\boxed{m = \frac{h_i}{h_o} = \frac{i}{o}}$$

Blur Circle (Defocus)



Blur Circle (Defocus)



From similar triangles:

$$\frac{b}{D} = \frac{|i' - i|}{i'}$$

Blur circle diameter:

$$\boxed{b = \frac{D}{i'}|i' - i|} \quad \boxed{b \propto D}$$

Focusing

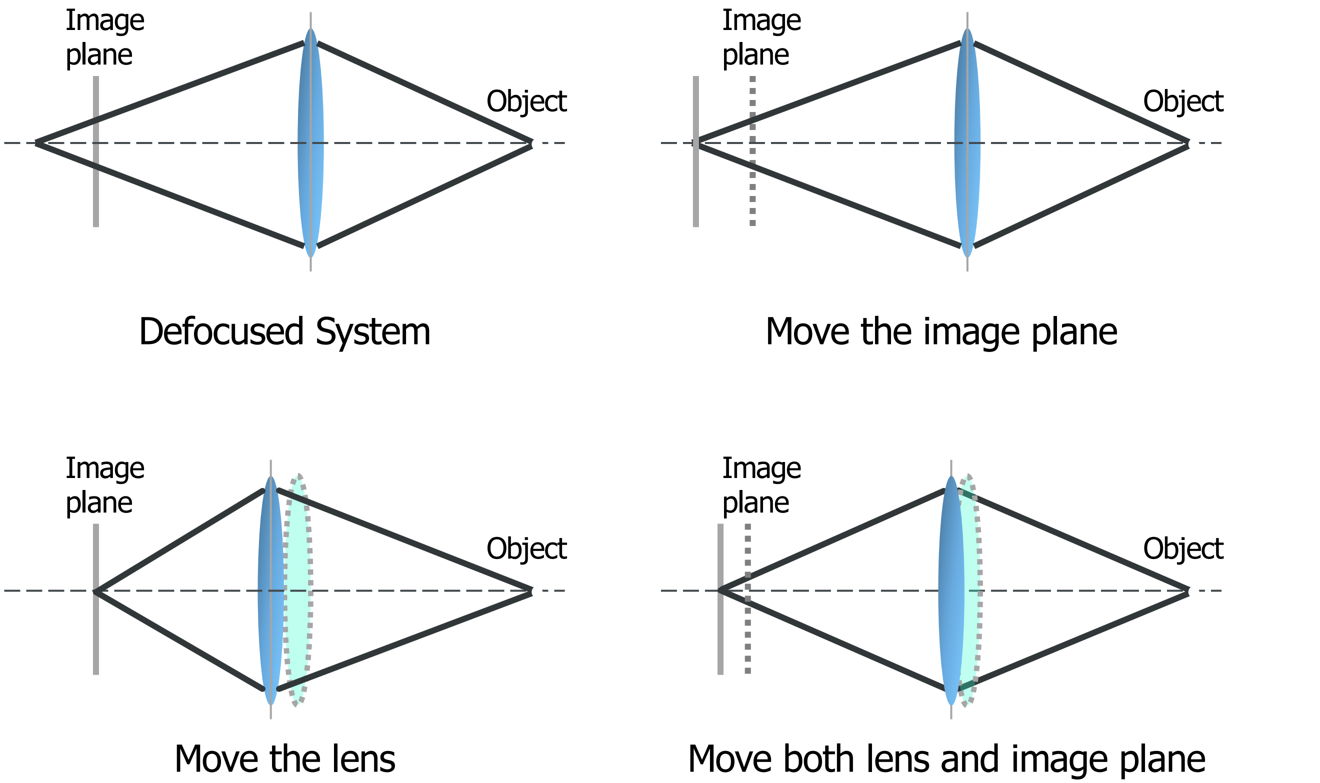
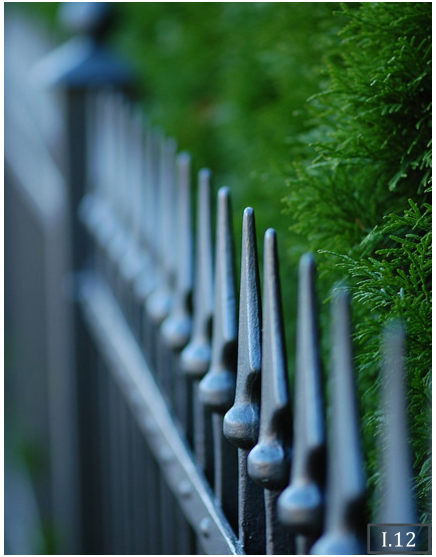


Image
plane
Object
Defocused System
Image
plane
Object
Move the image plane
Image
plane
Object
Move the lens
Image
plane
Object
Move both lens and image plane

Depth of Field (DoF)

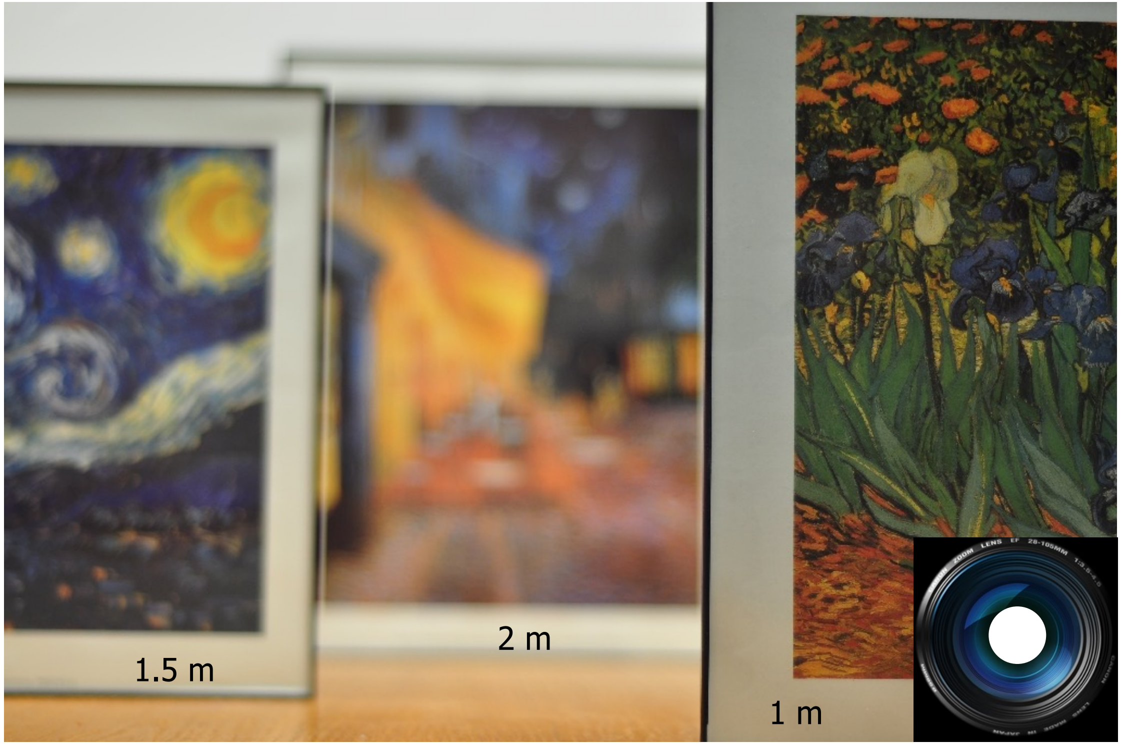
I.11



I.12

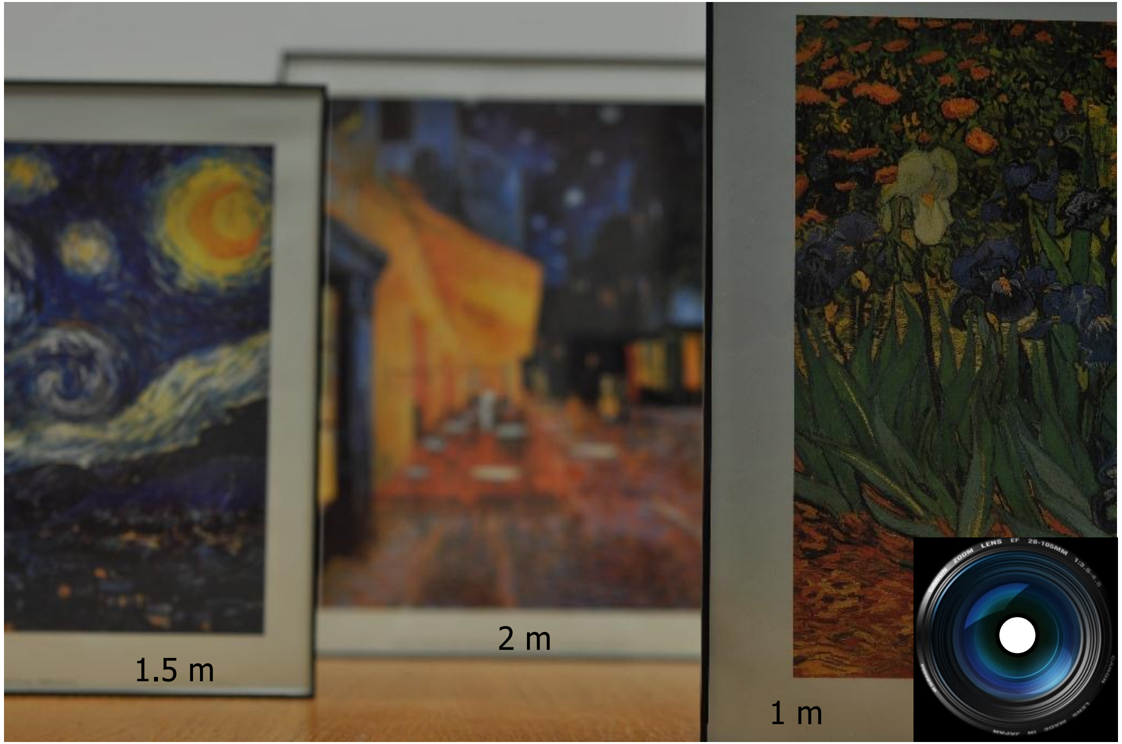
Range of object distances $(o - o')$ over which the image is "sufficiently well" focused. i.e., Range $(o - o')$ for which blur circle b is less than pixel size.

Aperture Size: DOF vs. Brightness



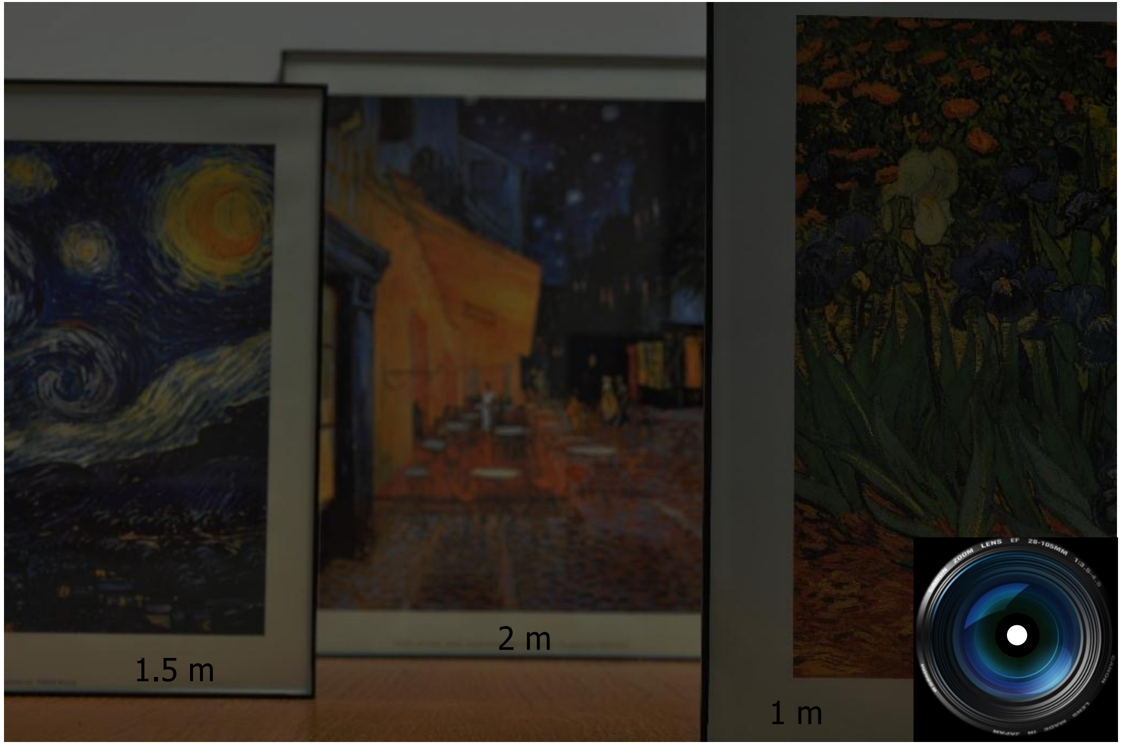
Focal Length 50 mm, Focus = 1 m, Aperture D = **25 mm**

Aperture Size: DOF vs. Brightness



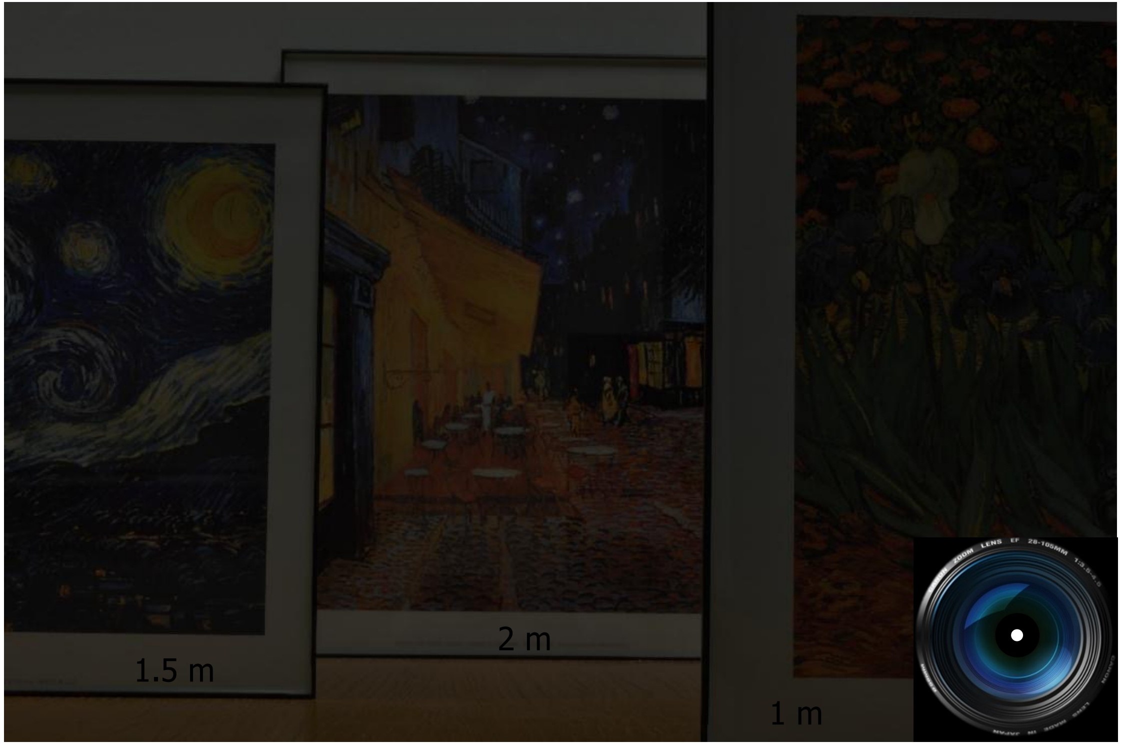
Focal Length 50 mm, Focus = 1 m, Aperture D = **12.5 mm**

Aperture Size: DOF vs. Brightness



Focal Length 50 mm, Focus = 1 m, Aperture D = **6.25 mm**

Aperture Size: DOF vs. Brightness



Focal Length 50 mm, Focus = 1 m, Aperture D = **3.125 mm**

Aperture Size: DOF vs. Brightness

- **Large** Aperture
  - **Bright** Image or **Small** Exposure Time
  - **Shallow** Depth of Field
- **Small** Aperture
  - **Dark** Image or **Long** Exposure Time
  - **Large** Depth of Field

f-number

DOF is a function of both focus **distance** and aperture **diameter** d

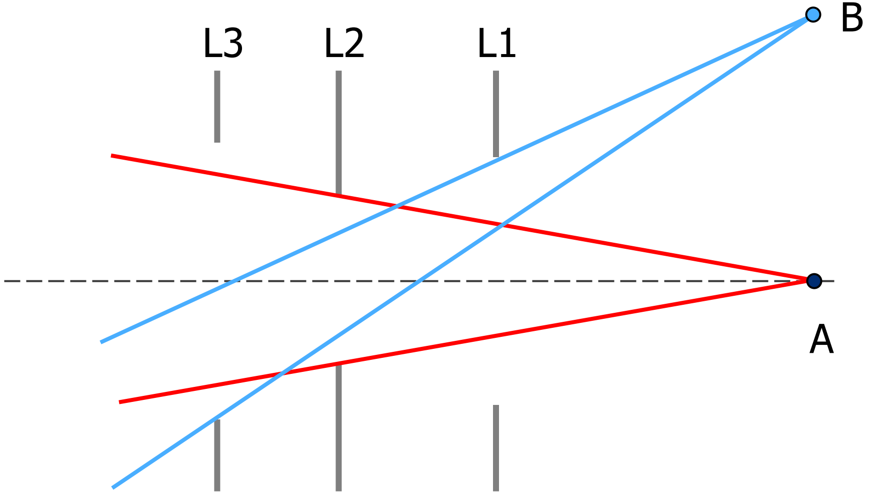
f-number is a common number associated with cameras, which is usually denoted *f/#*

$f/\# = N = f\,/\,d$

f: focal length
d: aperture diameter

This is usually written as f/#, where # is the aperture diameter in millimeters (e.g., *f*/1.4, *f*/2, *f*/2.8, …, *f*/22). We interpret these numbers by noticing that dividing the focal length by the f-number gives the aperture diameter *d*.

Vignetting



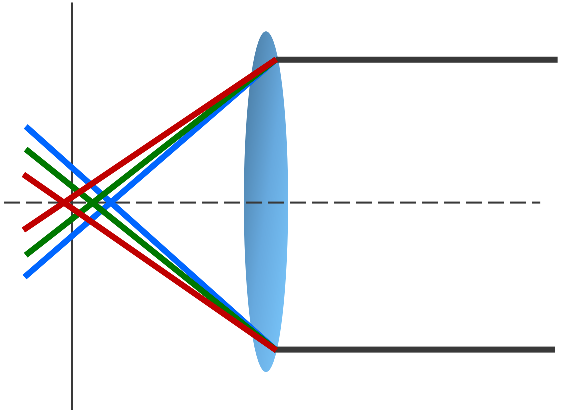
More light passes through L3 from point A than point B. Results in a smooth fall-off of brightness from A to B.

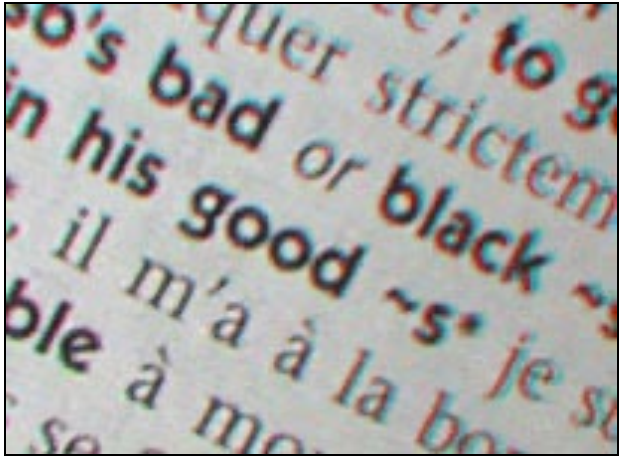
Vignetting

Brightness fall-off (Vignetting)
in image of a White Wall

Brightness fall-off (Vignetting)
in image of a Natural Scene

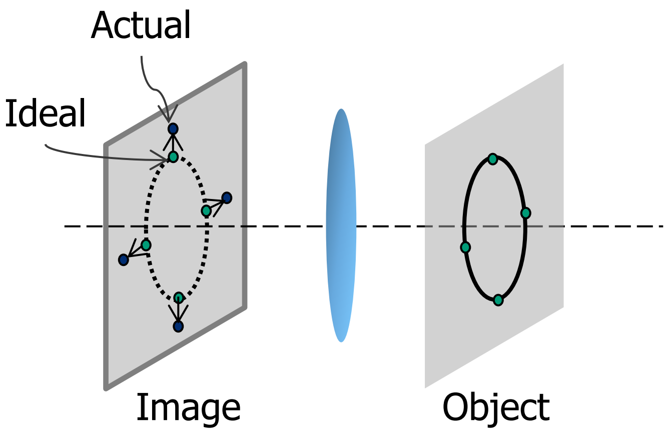
Chromatic Aberration



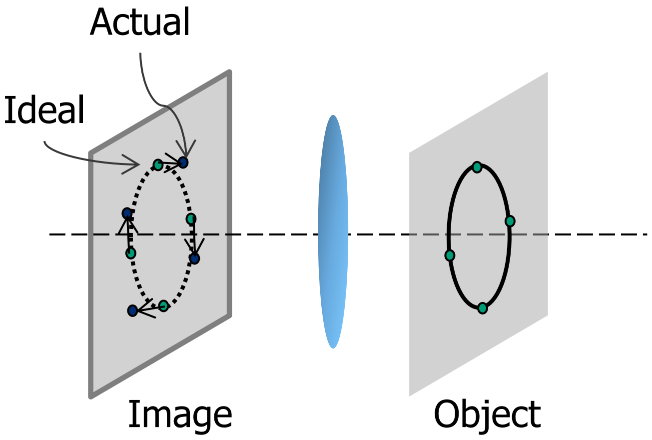


Refractive index (and hence focal length) of lens is different for different wavelengths.

Geometric Distortion



Radial distortion



Tangential distortion

Due to lens imperfections

When Geometric Distortion is Useful

Fisheye Lens

Radial Distortion in Fisheye Lens

Summary

- **Image**: Projection of 3D scene onto 2D plane
- We need to understand **Geometric** and **Radiometric** relations between the scene and its image
- Essential concepts in this lecture:
  - The reason for lenses
  - Basic lens properties: focal length, blur
  - How lenses and exposure interact
  - How lenses diverge from pinhole model

Johns Hopkins Engineering

Computer Vision

Image Formation

JOHNS HOPKINS
WHITING SCHOOL
of ENGINEERING

The Digital Camera

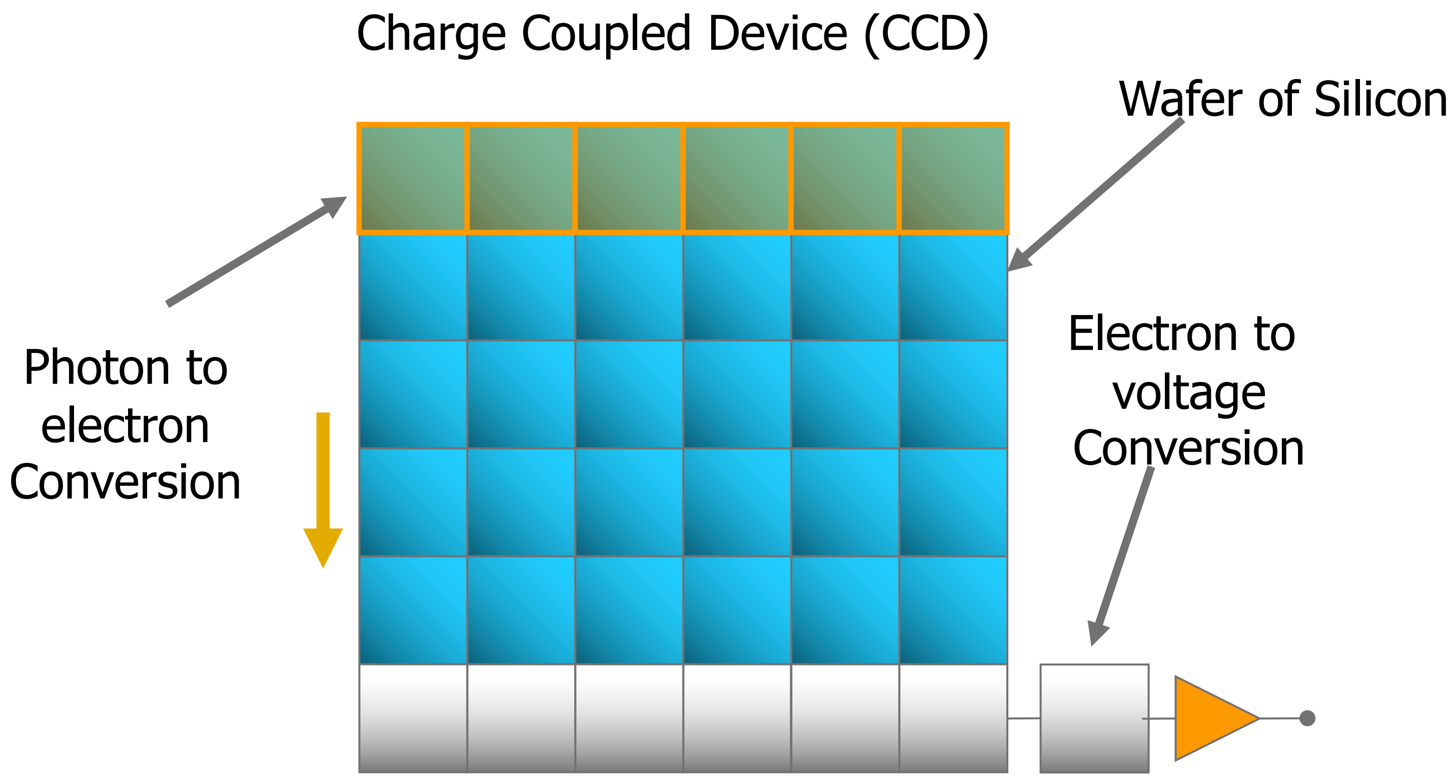
- **Image:** Projection of 3D scene onto 2D plane
- We need to understand **Geometric** and **Radiometric** relations between the scene and its image.
- This minilecture will cover the basic concepts of imaging using modern digital cameras

The Digital Camera

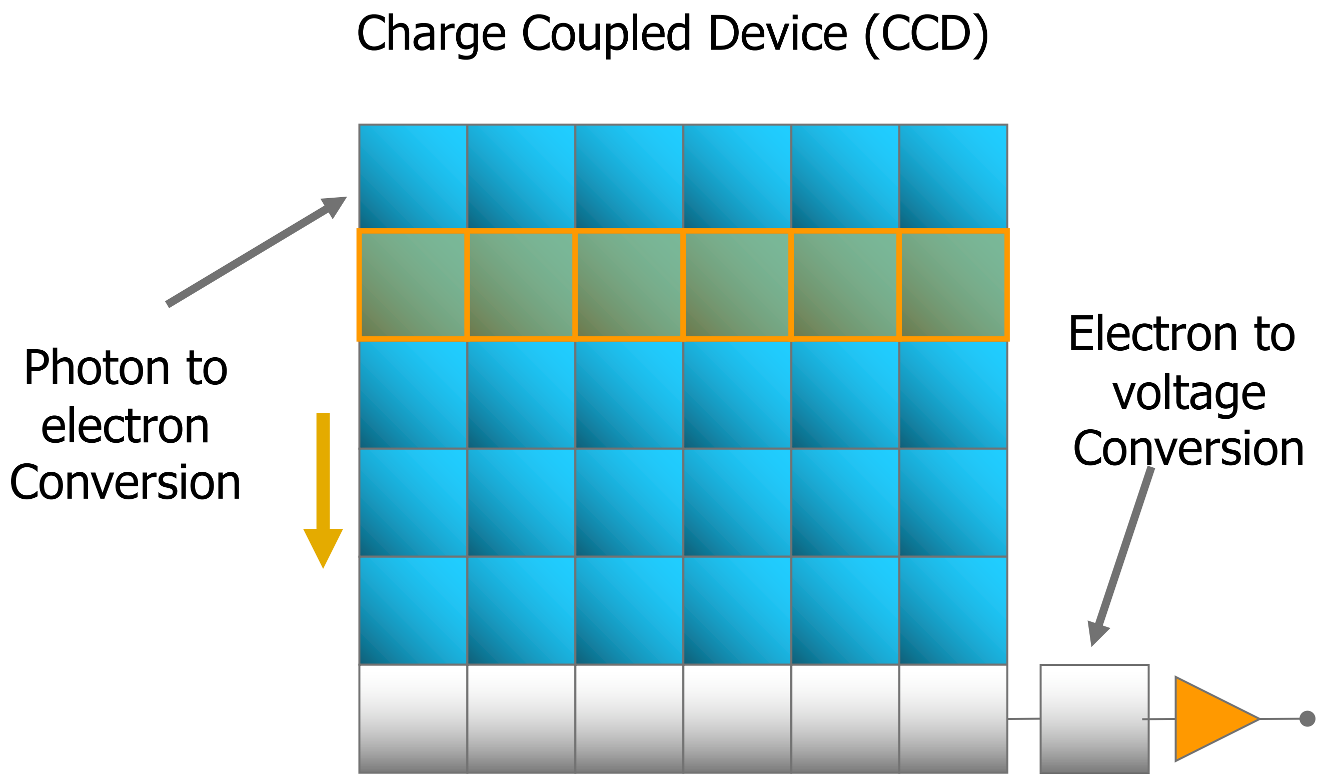
Two main kinds of sensors:

- **Charge-Coupled Device (CCD)**: photons are accumulated in each active well during exposure and all charges are transferred from well-to-well until they are deposited at the amplifiers, and this is then passed to the Analog-to-Digital Converter (ADC).
- **Complimentary Metal Oxide on Silicon (CMOS)**:

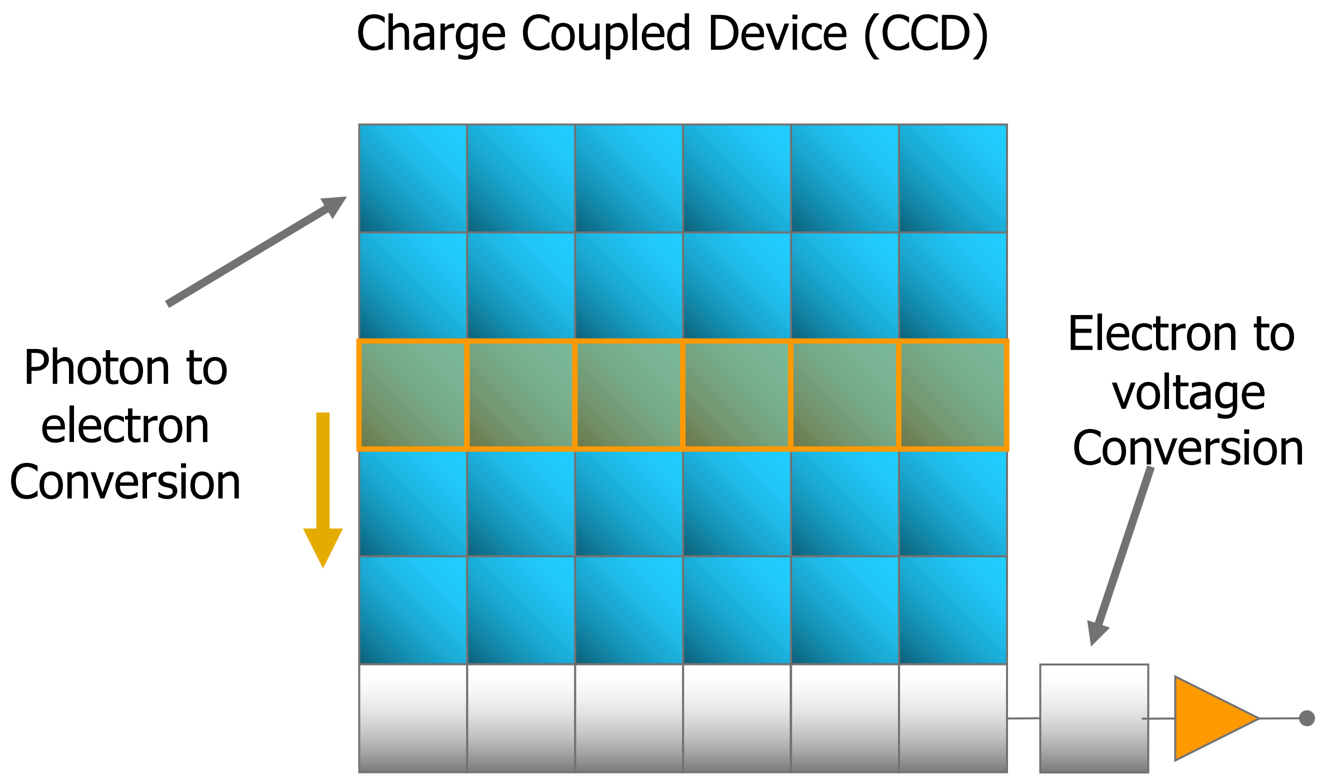
Popular Types of Image Detectors



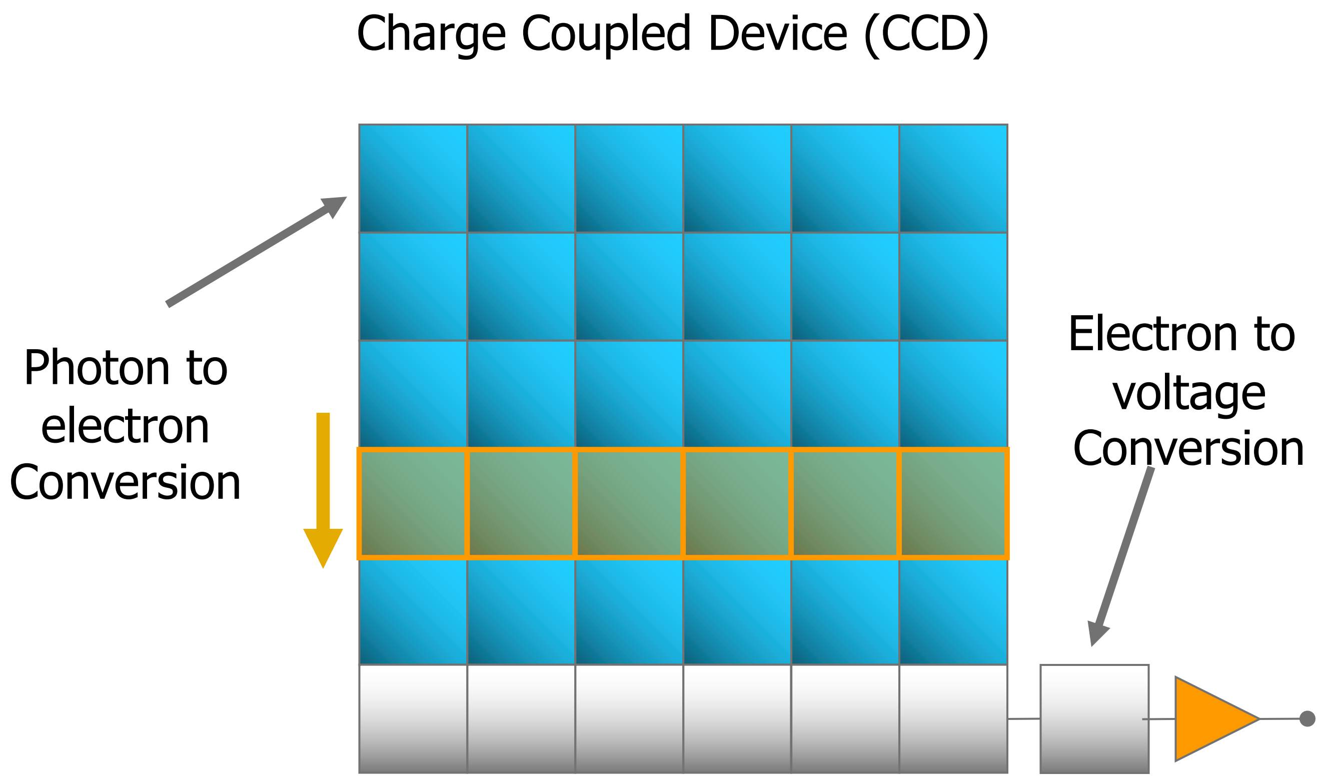
Popular Types of Image Detectors



Popular Types of Image Detectors

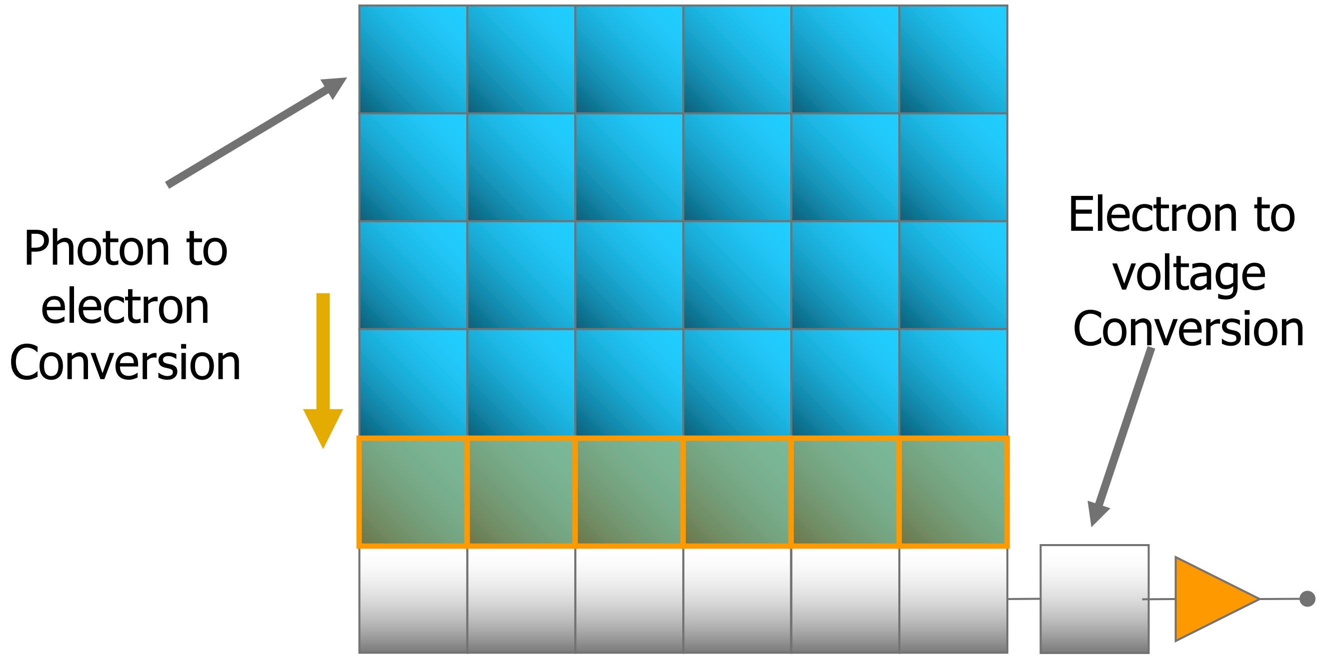


Popular Types of Image Detectors



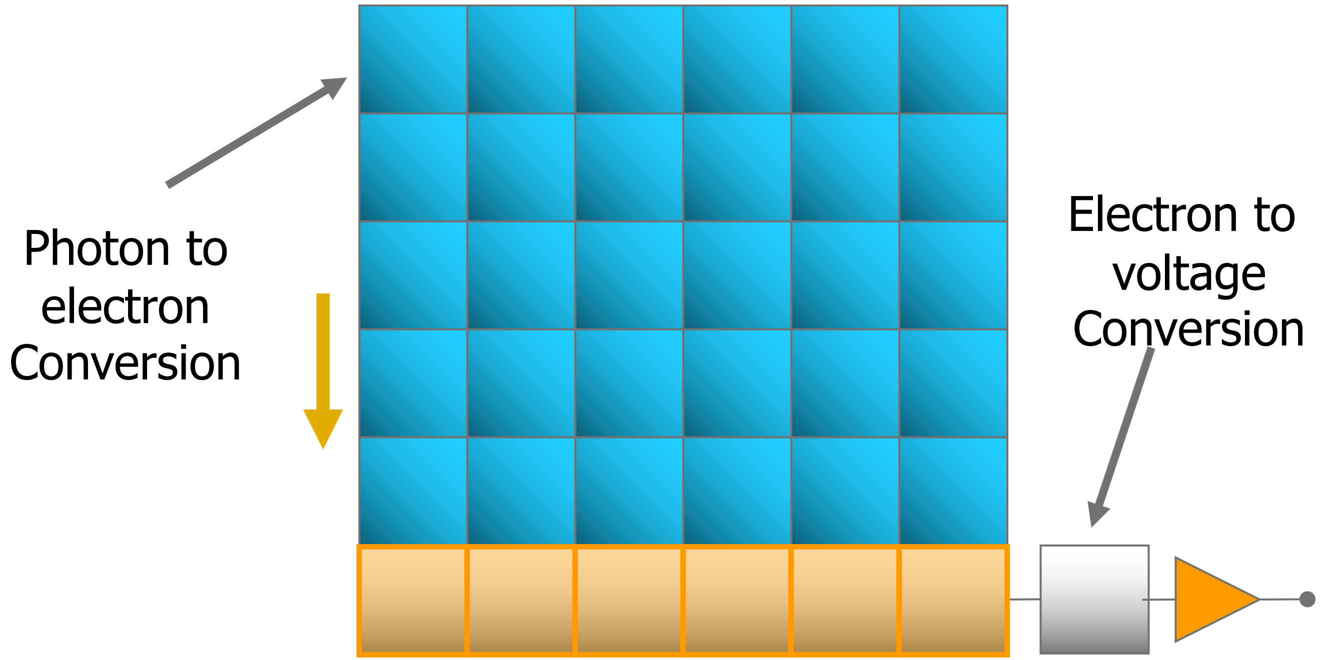
Popular Types of Image Detectors

Charge Coupled Device (CCD)

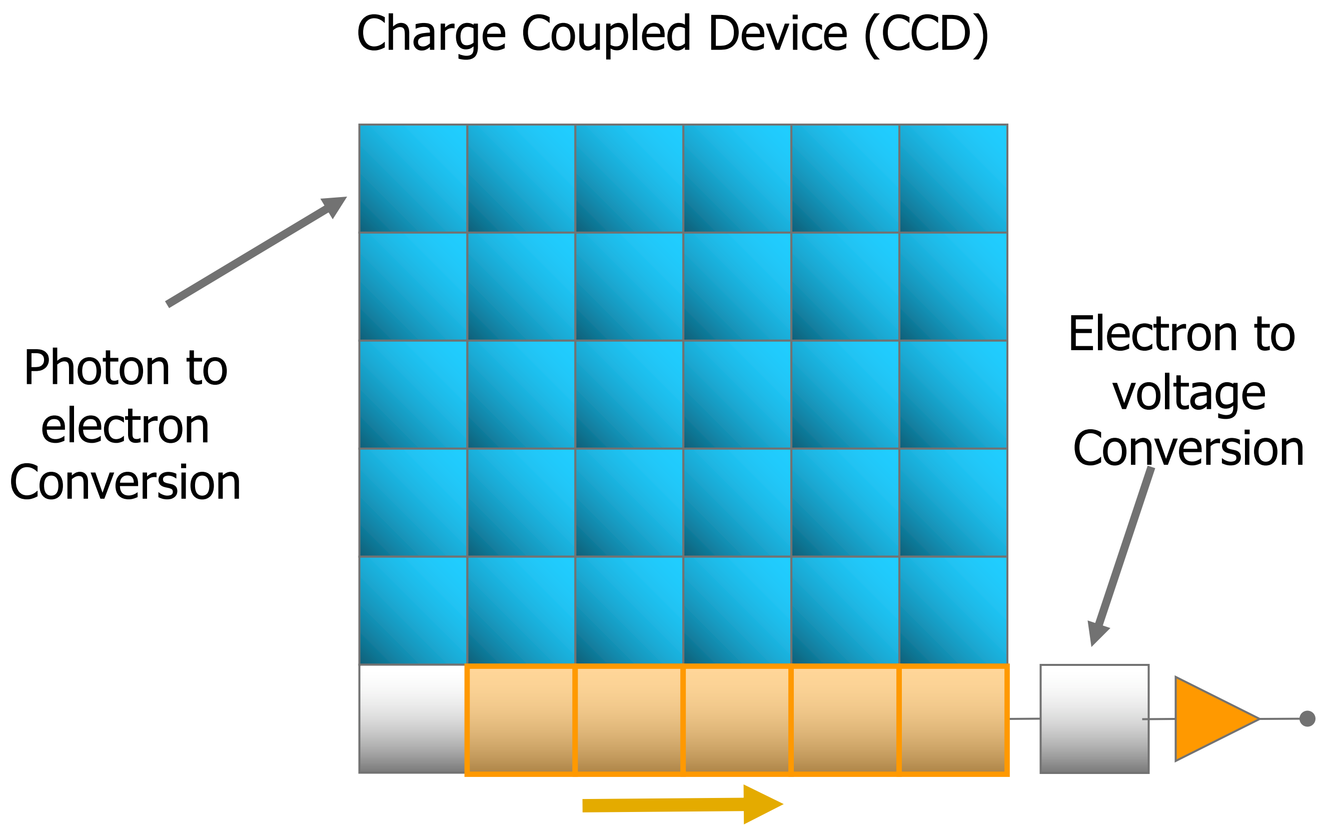


Popular Types of Image Detectors

Charge Coupled Device (CCD)

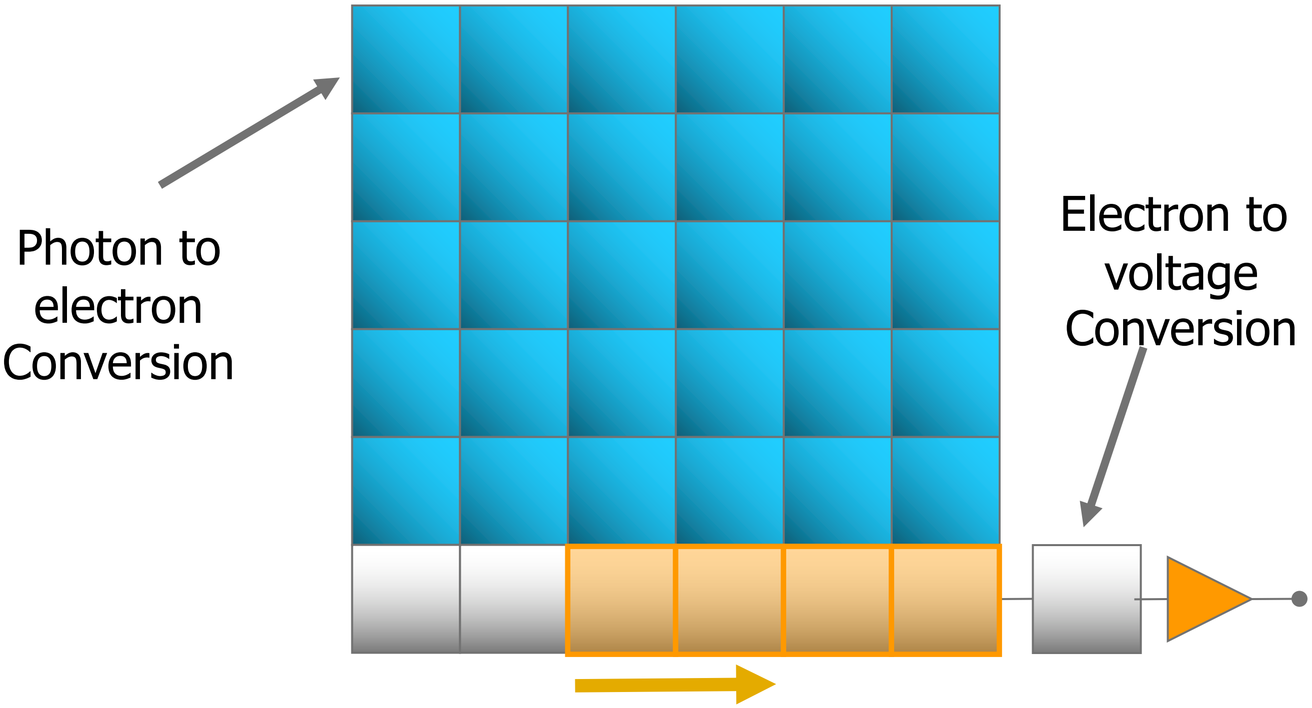


Popular Types of Image Detectors

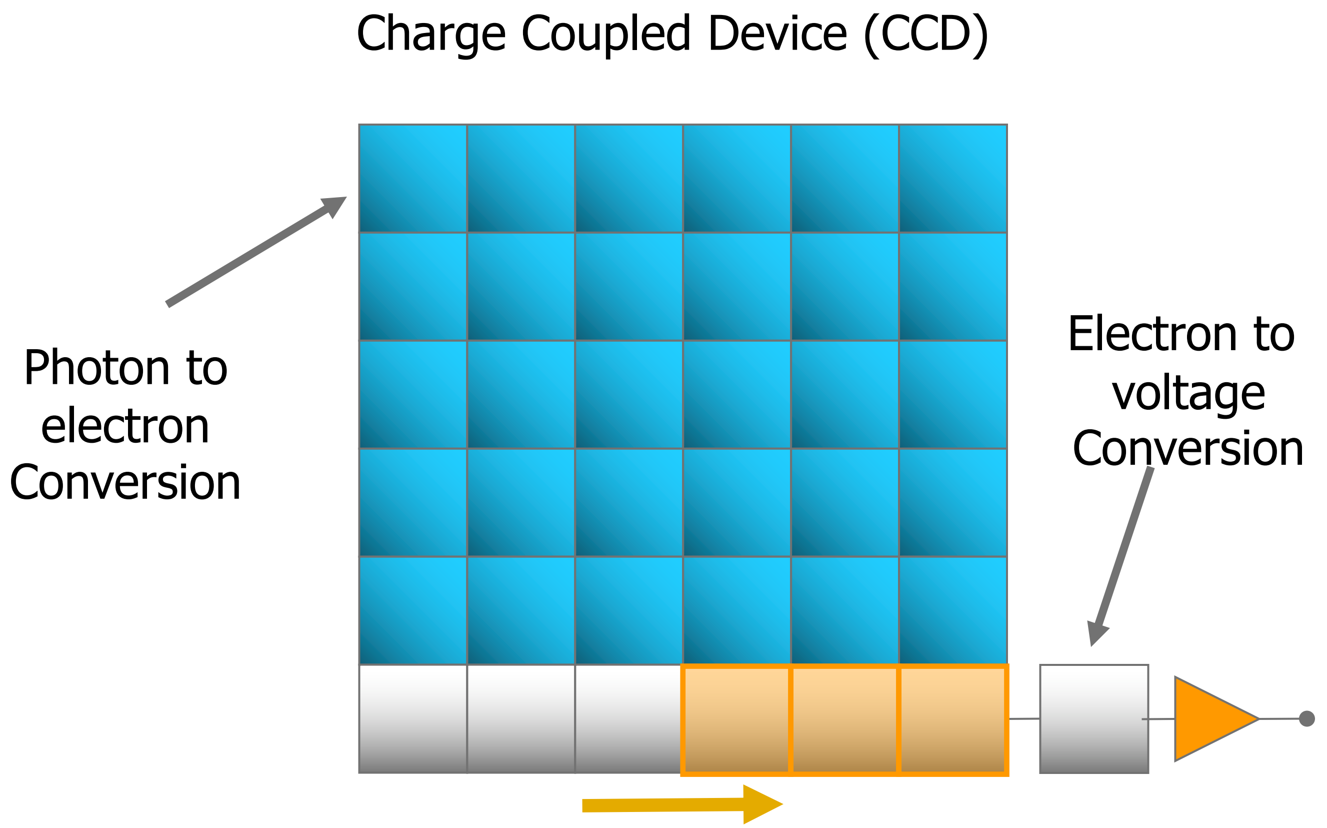


Popular Types of Image Detectors

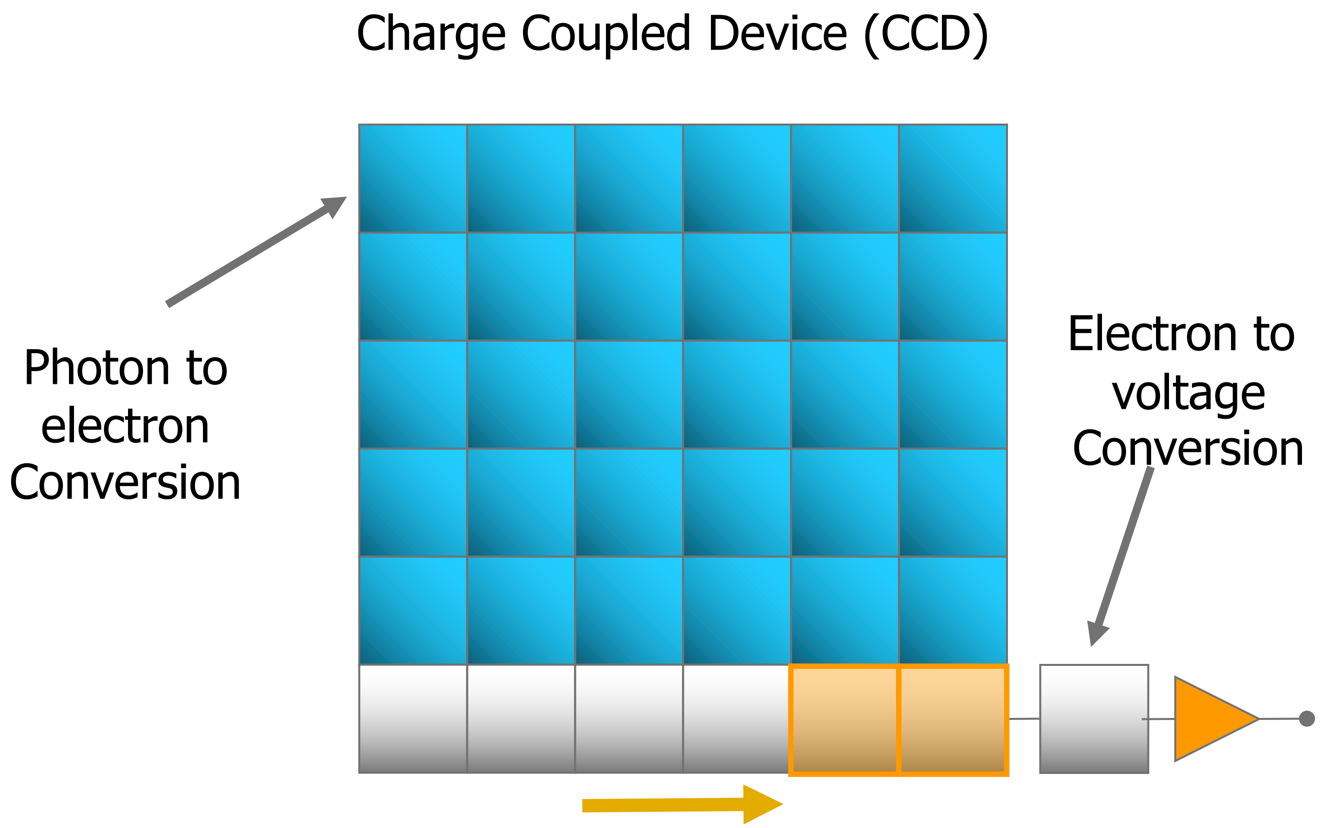
Charge Coupled Device (CCD)



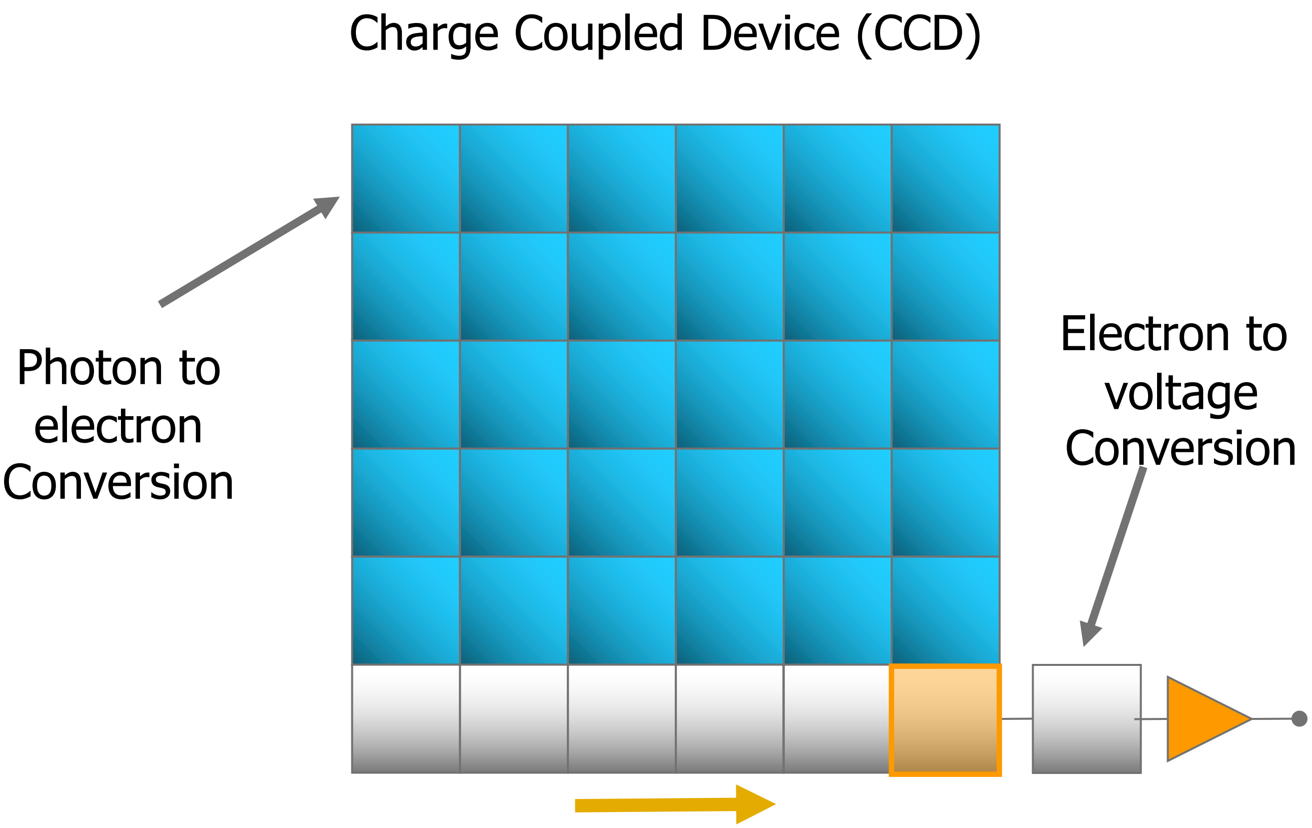
Popular Types of Image Detectors



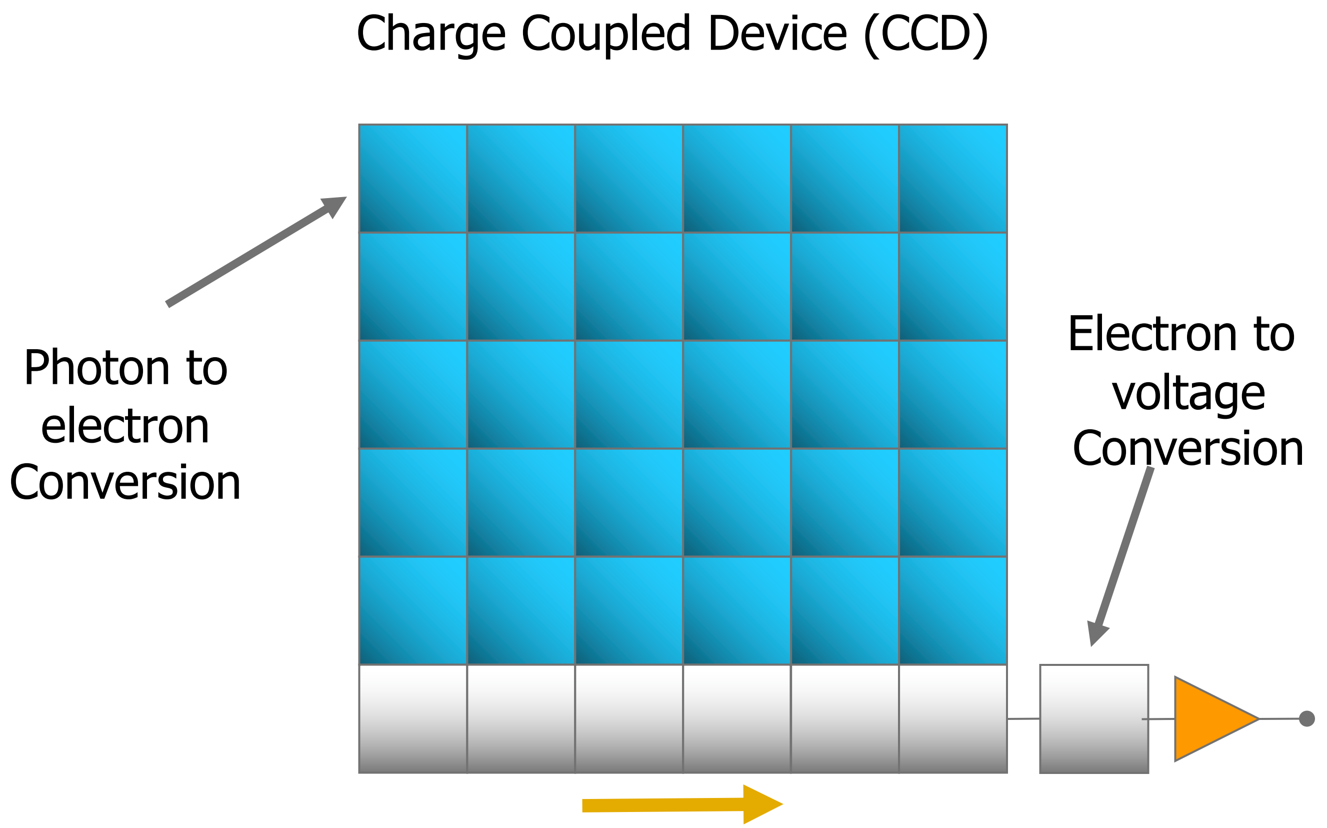
Popular Types of Image Detectors



Popular Types of Image Detectors



Popular Types of Image Detectors



The Digital Camera

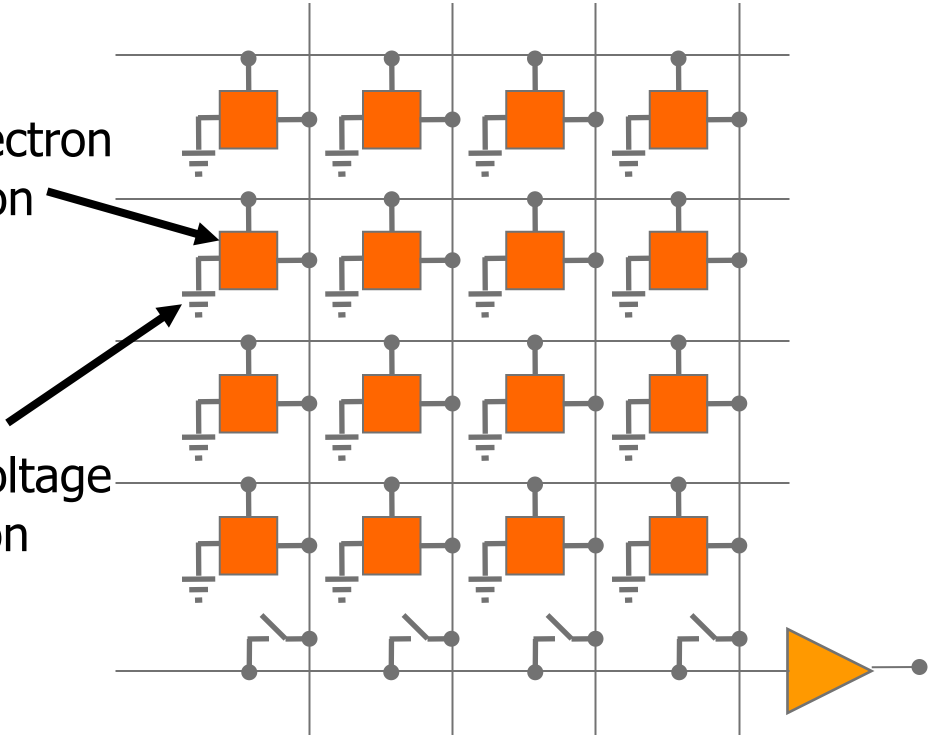
Two main kinds of sensors:

- **Charge-Coupled Device (CCD)**: photons are accumulated in each active well during exposure and all charges are transferred from well-to-well until they are deposited at the amplifiers, and this is then passed to the Analog-to-Digital Converter (ADC).
- **Complimentary Metal Oxide on Silicon (CMOS)**: photons hit the sensor directly and affect the conductivity of the photodetector, which can be selectively gated to control exposure duration and locally amplified before being read out.

Popular Types of Image Detectors

Complimentary Metal-Oxide Semiconductor (CMOS)

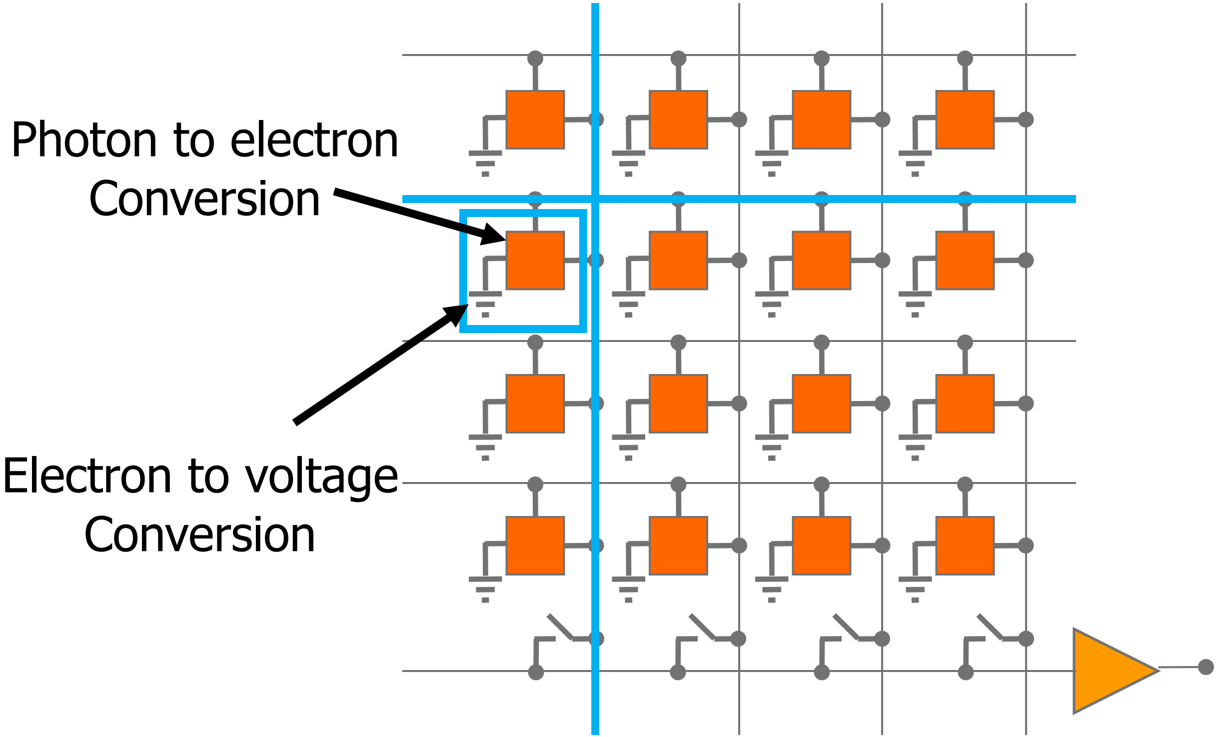
Photon to electron Conversion



Electron to voltage Conversion

Popular Types of Image Detectors

Complimentary Metal-Oxide Semiconductor (CMOS)

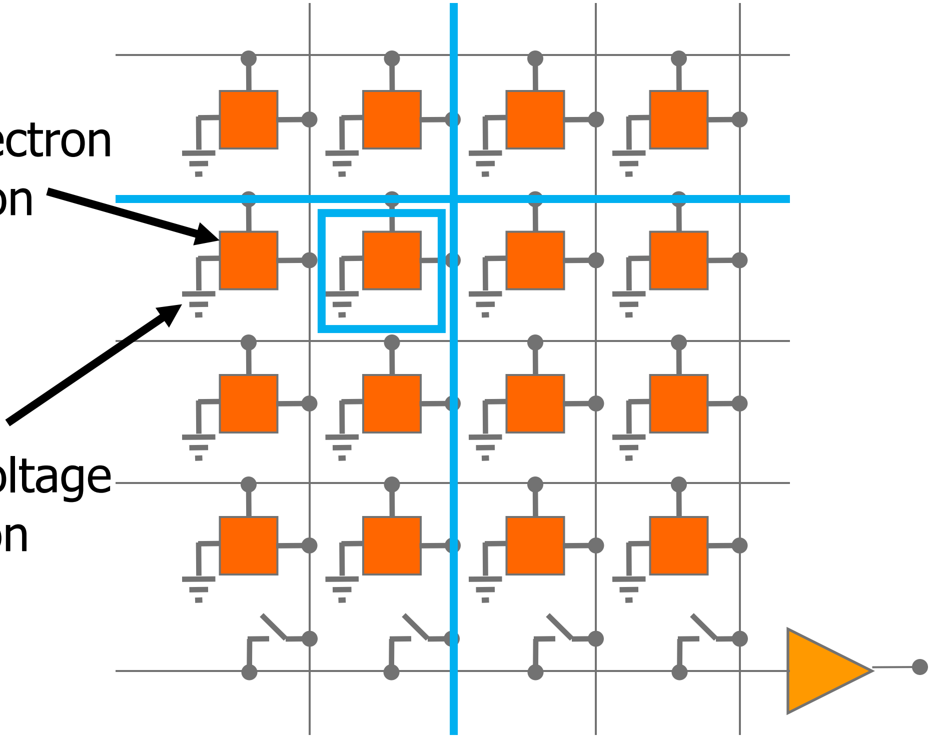


Popular Types of Image Detectors

Complimentary Metal-Oxide Semiconductor (CMOS)

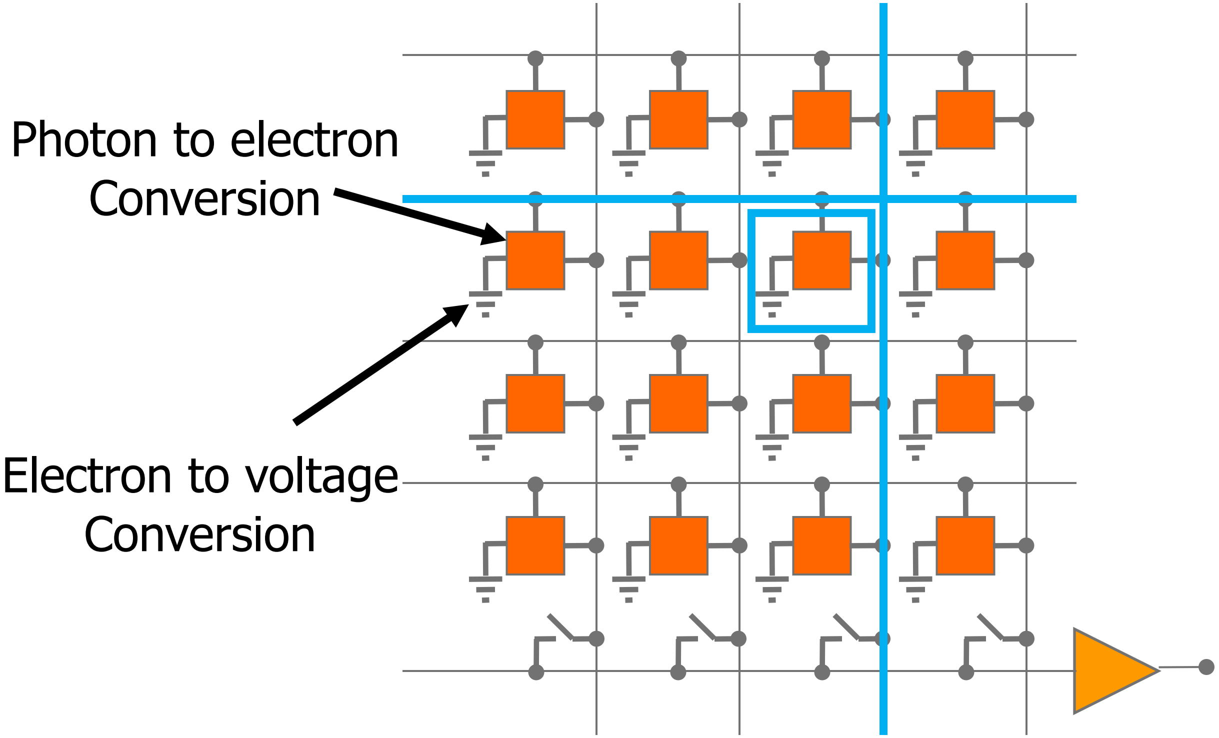
Photon to electron Conversion

Electron to voltage Conversion



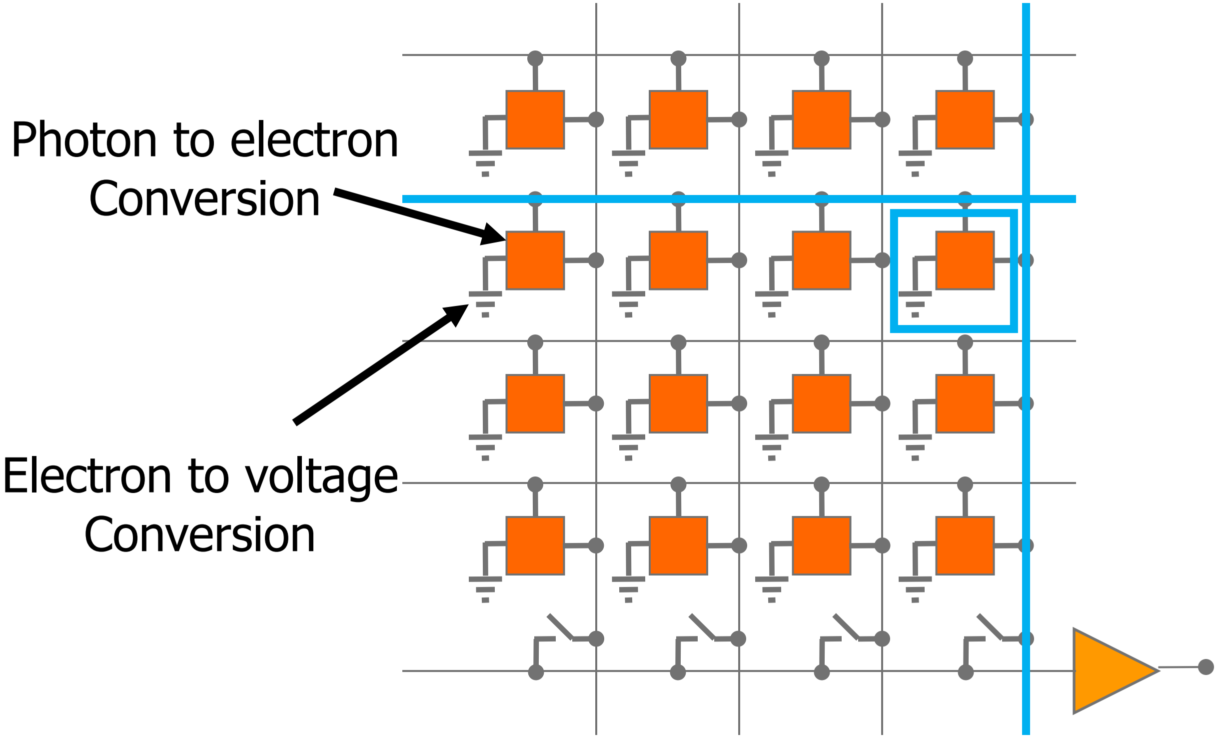
Popular Types of Image Detectors

Complimentary Metal-Oxide Semiconductor (CMOS)



Popular Types of Image Detectors

Complimentary Metal-Oxide Semiconductor (CMOS)



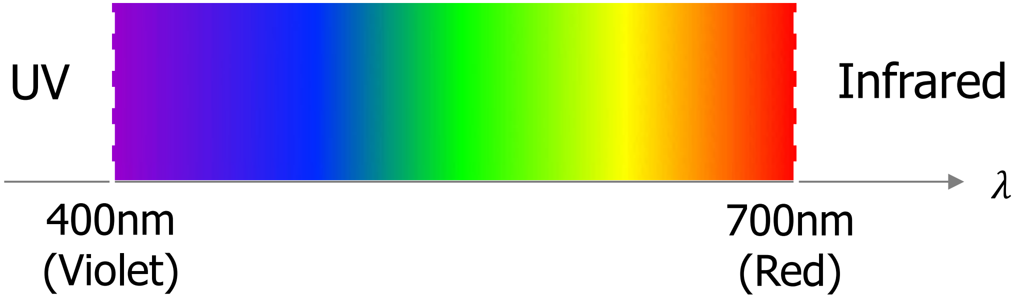
Comparison: CCD vs. CMOS

	CCD	CMOS
Signal Output	Separate circuit to convert photons to voltage	Convert photons to voltage within each pixel
Noise	Low	Moderate
Dynamic Range	High	Moderate
Uniformity	High	Low to Moderate
Windowing	Limited	Extensive
Power Consumption	Moderate	Low

What is "Color"?

Human response to different wavelengths

Visible light:



Do We recover Spectral Distribution $p(\lambda)$?

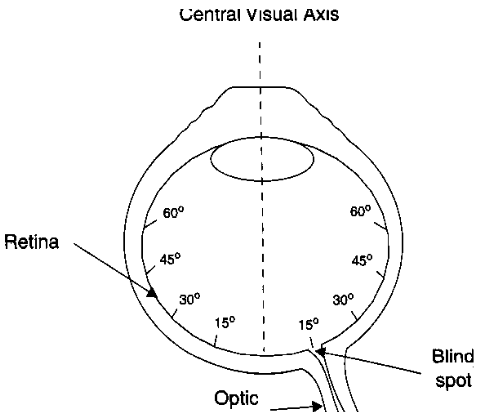
Sensors in the human eye: Rods & Cones
Neurochemical Sensors (3 types)

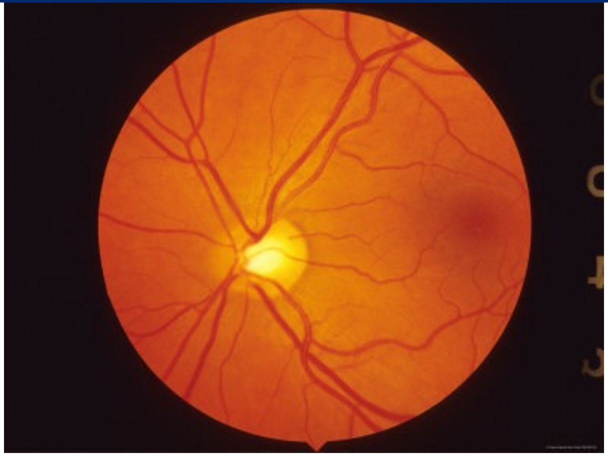
Color Measurement

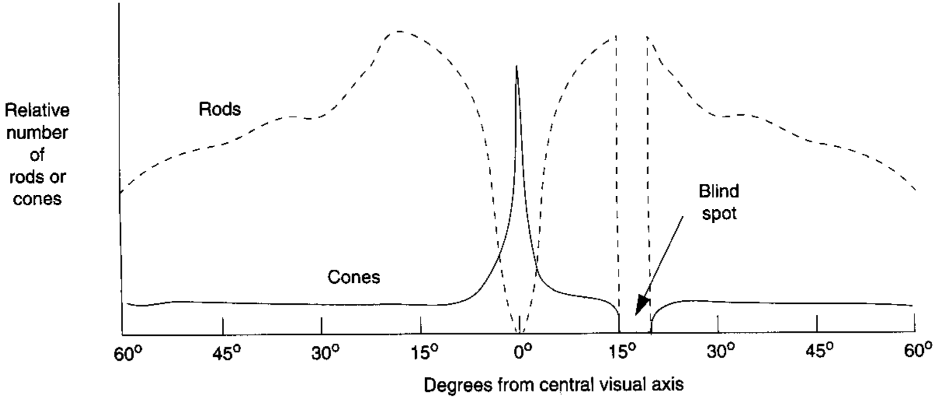
- Let λ denote wavelength
- Let $E(\lambda)$ denote the spectral power at a given wavelength
- Let $\rho_k(\lambda)$ denote the responsiveness of a sensor k to a given wavelength of light
- Then we can compute the "response" r_k of k as

$$r_k = \int \rho(\lambda) E(\lambda) d\lambda$$

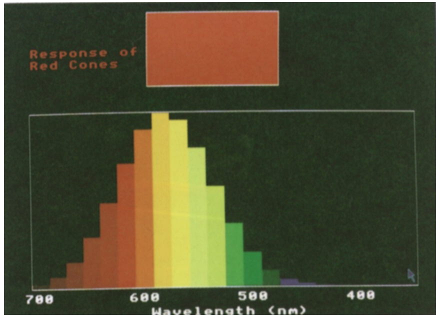
Example: The Human Eye



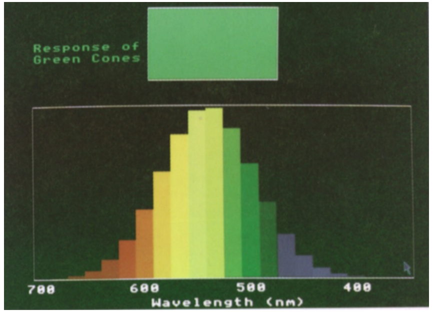




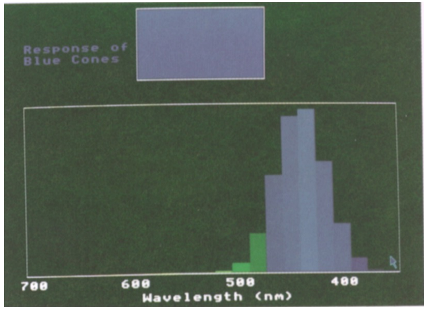
Example: The Human Eye



"Red" cone

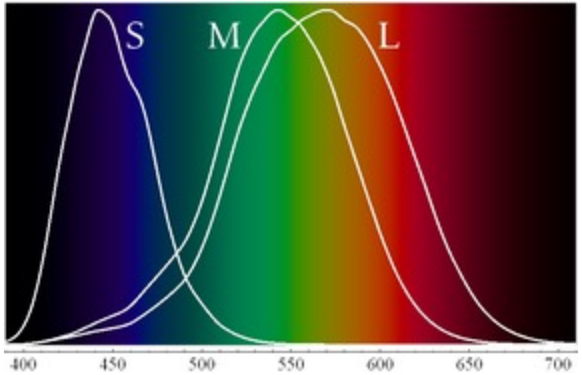


"Green" cone



"Blue" cone

Principle of univariance: cones give the same amount of response to different wavelengths - a single cone cannot distinguish color. Output of cone is obtained by summing probability of absorption over wavelengths



The Mixing of Colors

Young's Experiment on Color Mixture

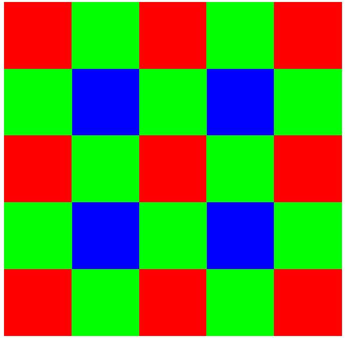
Human Sensation of nearly all colors can be produced using 3 wavelengths!

$$(\lambda_r \quad \lambda_g \quad \lambda_b) = (\ 650\ 530\ 410\ )nm$$

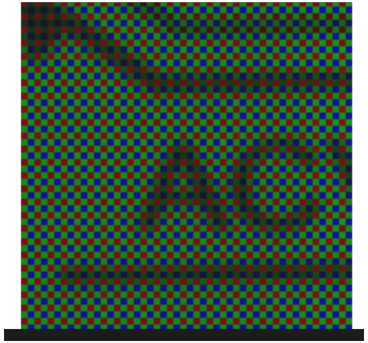
Hence, cameras and displays often use 3 filters:

(red, green, blue)

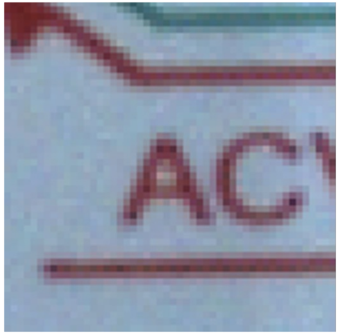
Sensing Color Using Color Mosaic



Bayer Pattern
(Color Filter Mosaic)



Raw Image



Interpolated Image

Color Filled in by Interpolation (**Demosaicing**)

High Dynamic Range: Multiple Exposures

Assume Camera Response $f(\cdot)$ is Linear

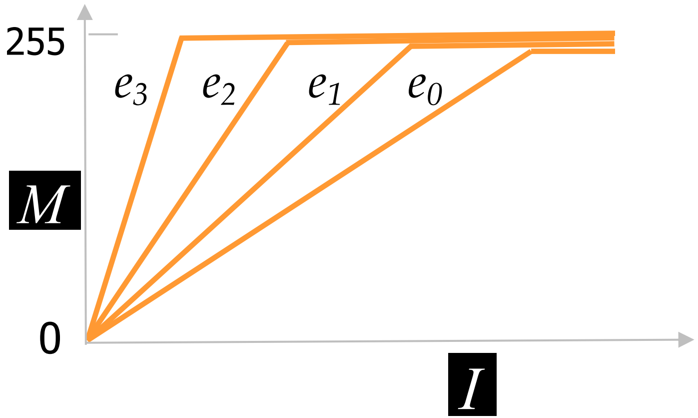
with e_0 e_1 e_2 e_3

$$M_0 = \min(e_0 \cdot \mathrm{I}, 255)$$

$$M_1 = \min(e_1 \cdot \mathrm{I}, 255)$$

$$M_2 = \min(e_2 \cdot \mathrm{I}, 255)$$

$$M_3 = \min(e_3 \cdot \mathrm{I}, 255)$$

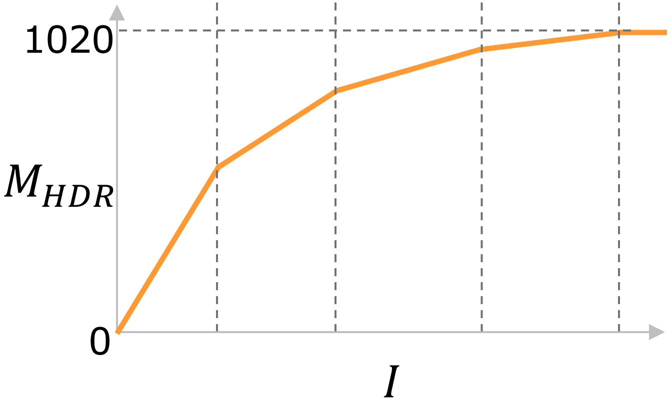


Nayar 2000

High Dynamic Range: Multiple Exposures

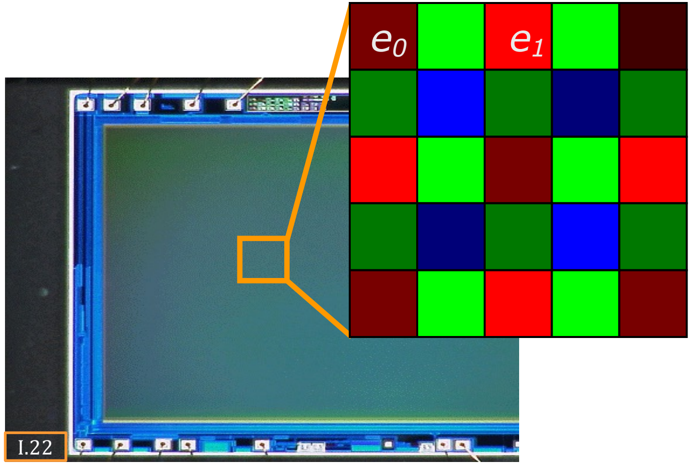
Aggregate Image: $M_{HDR} = M_0 + M_1 + M_2 + M_3$

Camera Response $f(\cdot)$ for Aggregate Image:

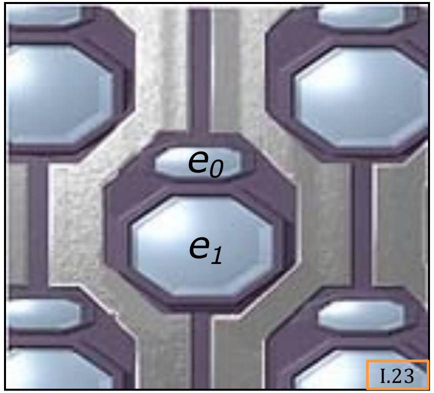


Nayar 2000

High Dynamic Range: Single Shot



Assorted Pixels:
Spatially Varying Color & Exposure



SuperCCD SR, FujiFilm:
Pixels with Subpixels

Summary

- **Image:** Projection of 3D scene onto 2D plane
- We need to understand **Geometric** and **Radiometric** relations between the scene and its image
- Essential concepts in this lecture:
  - The two different types of cameras
  - How color is represented
  - How color images are formed

JOHNS HOPKINS

WHITING SCHOOL
of ENGINEERING