

# Johns Hopkins Engineering

## Computer Vision

Image Formation

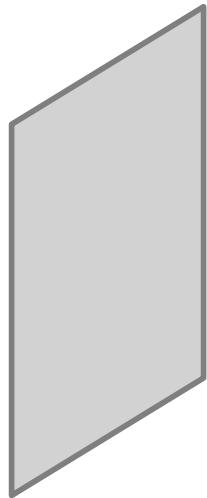


JOHNS HOPKINS  
WHITING SCHOOL  
*of* ENGINEERING

# 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 properties of pinhole projection

# Image Formation



Screen

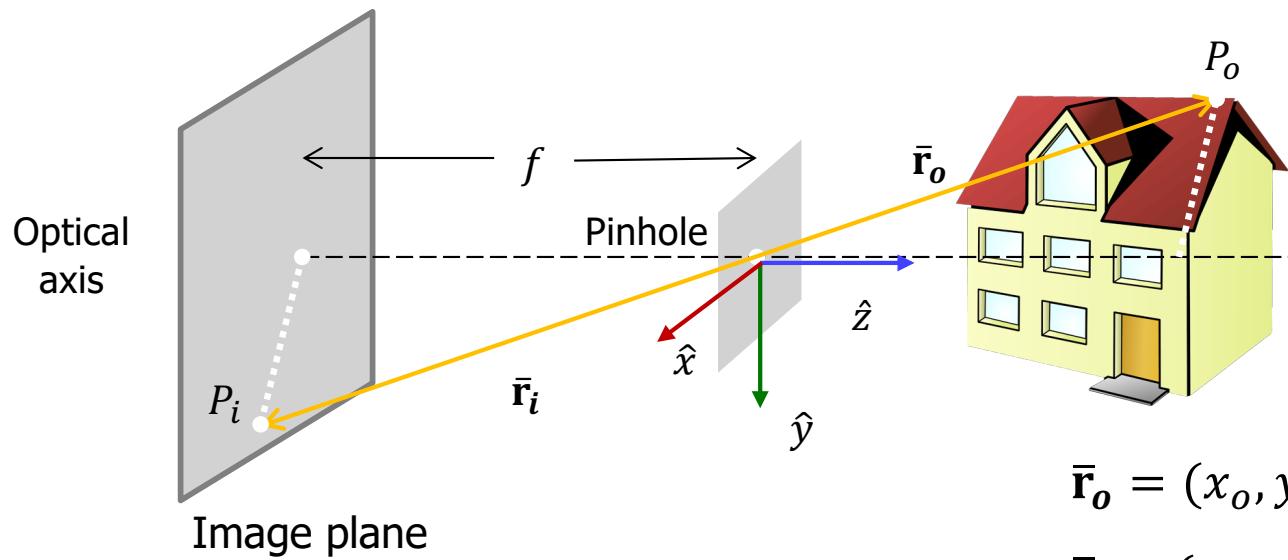


Scene

Is an image being formed on the screen?

**Yes! But not a “clear” one.**

# Perspective Imaging with Pinhole



$$\begin{aligned}\bar{r}_o &= (x_o, y_o, z_o) \\ \bar{r}_i &= (x_i, y_i, f)\end{aligned}$$

Using similar triangles:

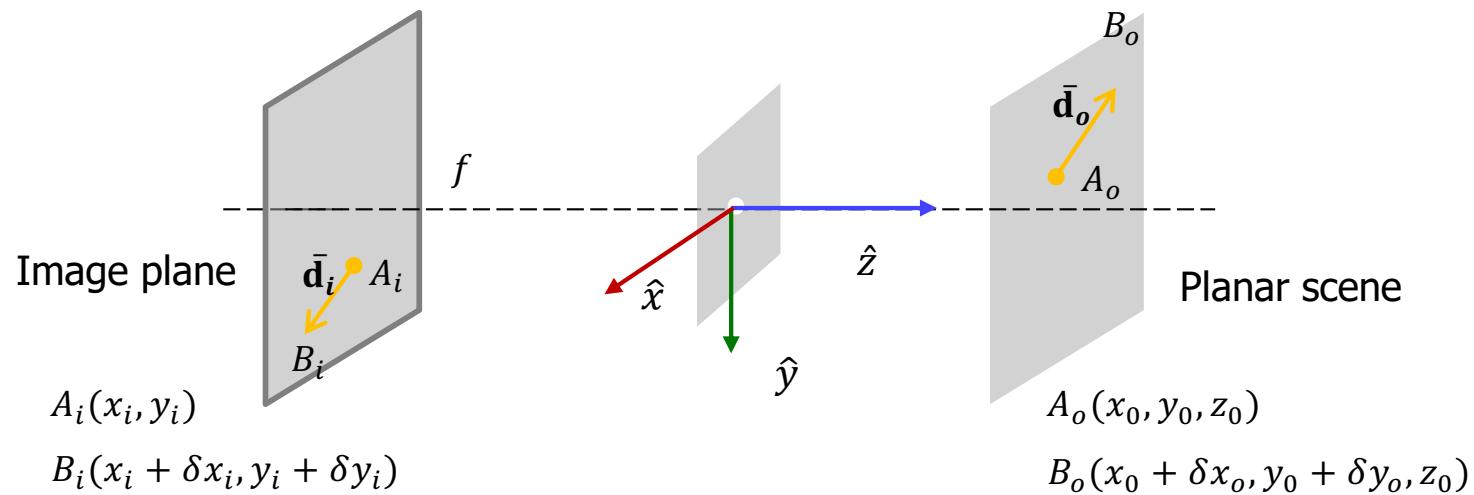
$$\frac{\bar{r}_i}{f} = \frac{\bar{r}_o}{z_o}$$

→

$$x_i = f \frac{x_o}{z_o}, \quad y_i = f \frac{y_o}{z_o}$$

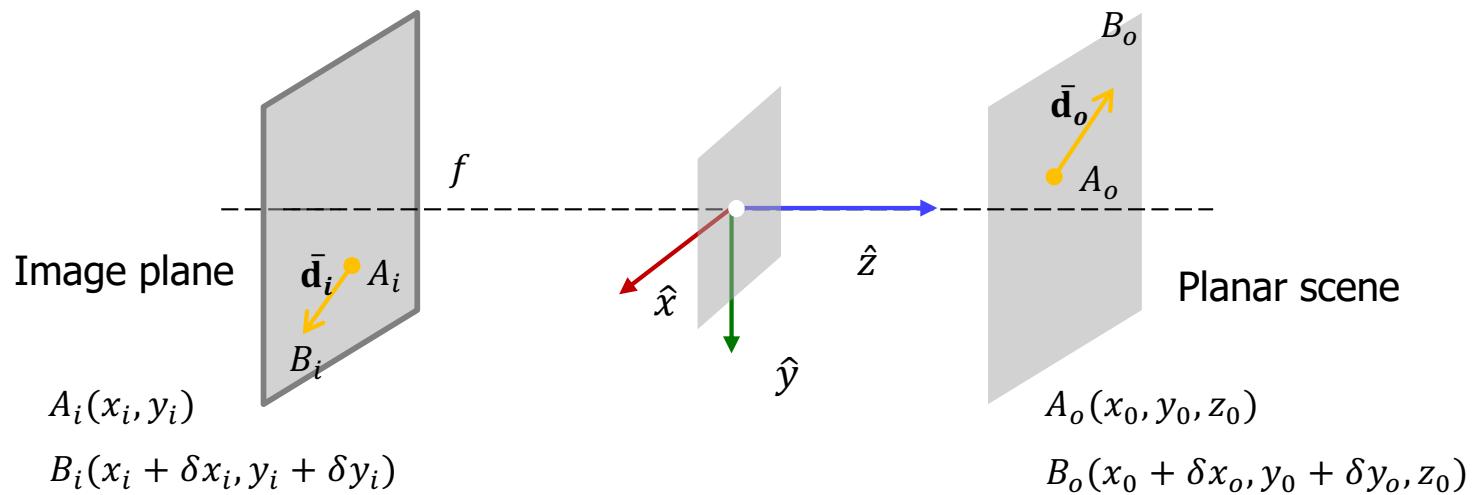
$f$ : Effective Focal Length

# Image Magnification



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

# Image Magnification



From Perspective Projection:

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

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

# Image Magnification

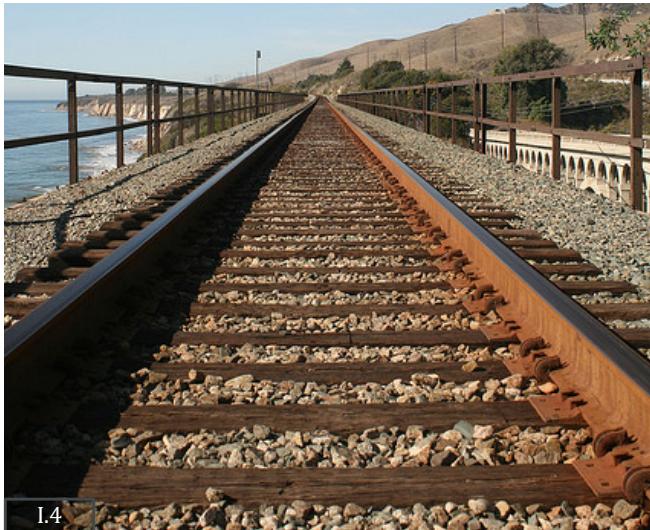
$$\delta x_i = f \frac{\delta x_o}{z_o} \quad \text{and} \quad \delta y_i = f \frac{\delta y_o}{z_o}$$

Magnification:

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

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

# Image Magnification



I.4



I.5

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

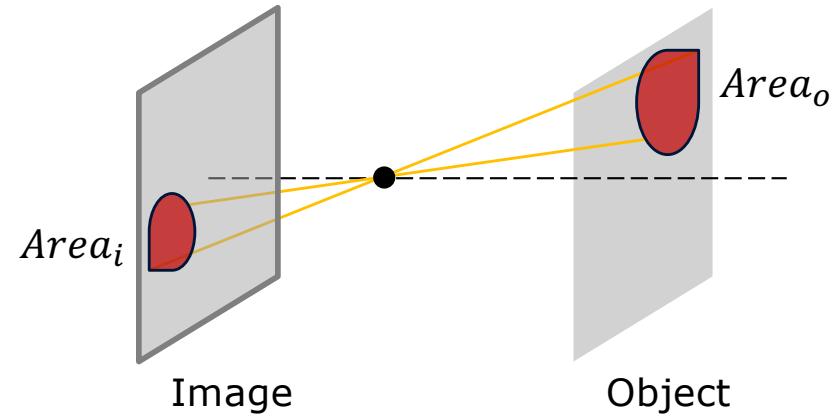
Image size **inversely proportional** to depth

# Image Magnification

## ■ Notes:

- $m$  can be assumed to be **constant** if the range of scene depth  $\Delta z$  is much smaller than the average scene depth  $\tilde{z}$

$$\frac{\text{Area}_i}{\text{Area}_o} = m^2$$



# Vanishing Point



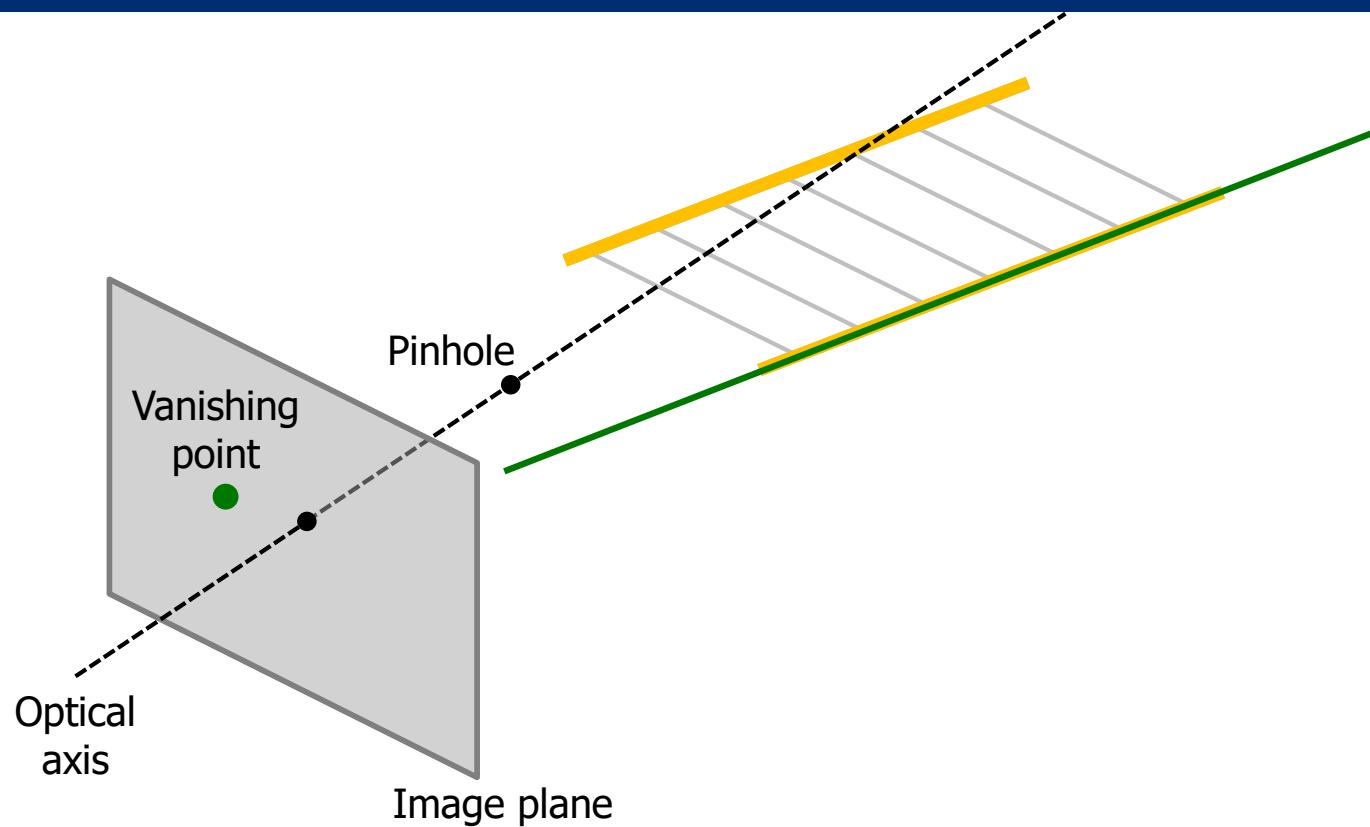
Parallel straight lines converge at a single image point

# Vanishing Point

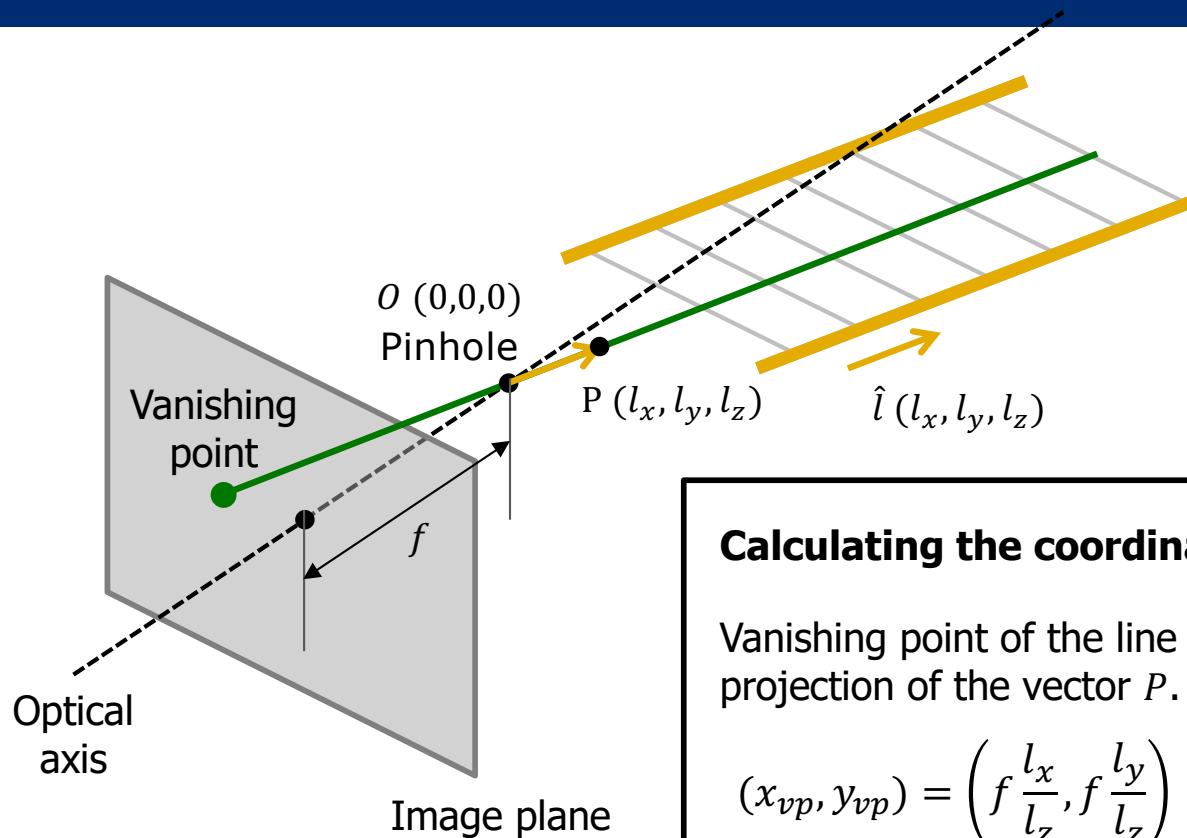


Location of Vanishing Point depends on  
the **orientation** of parallel straight lines.

# Finding the Vanishing Point



# Finding Vanishing Point



## Calculating the coordinates:

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

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

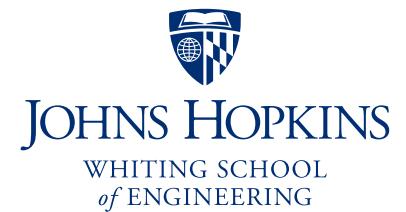
# Summary

- **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

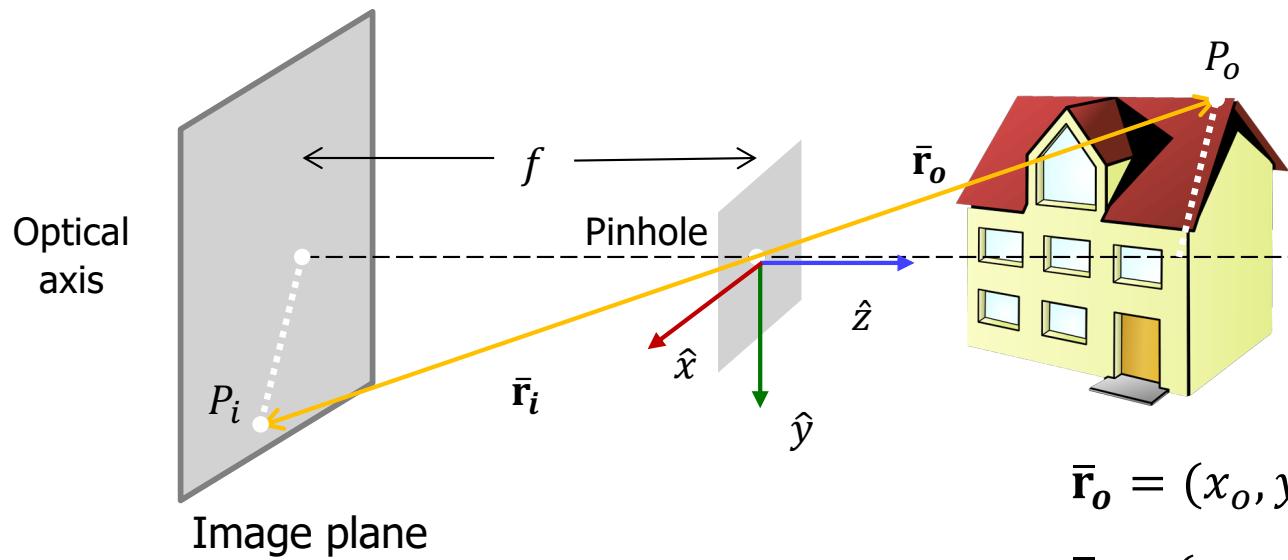
Image Formation



# 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



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

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

Using similar triangles:

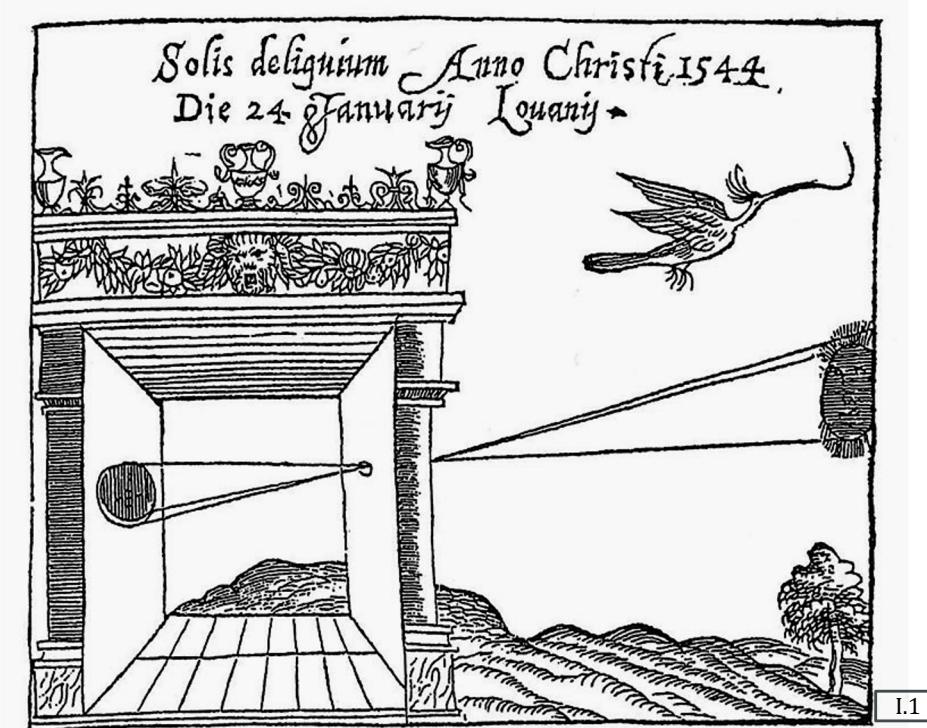
$$\frac{\bar{r}_i}{f} = \frac{\bar{r}_o}{z_o}$$

→

$$x_i = f \frac{x_o}{z_o}, \quad y_i = f \frac{y_o}{z_o}$$

$f$ : Effective Focal Length

# This is an Old Idea: Camera Obscura



“Dark Chamber”

# Pinhole Eye of *Nautilus pompilius*

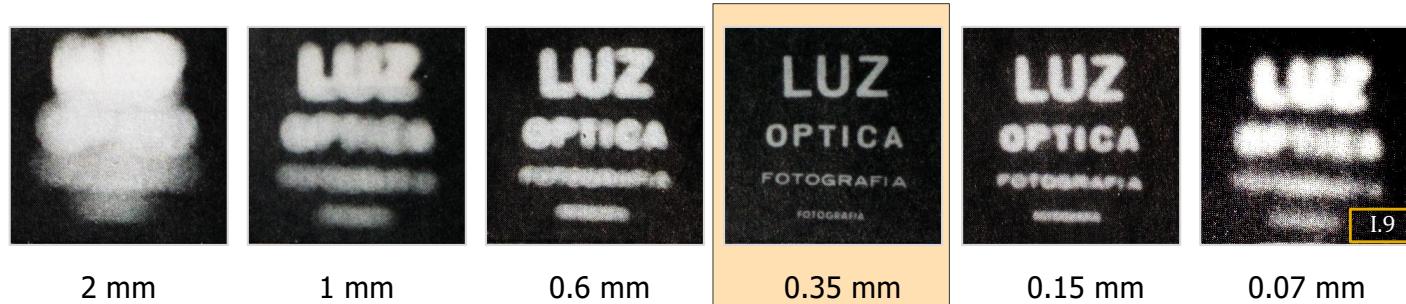


I.2



I.3

# What is the Ideal Pinhole Size?

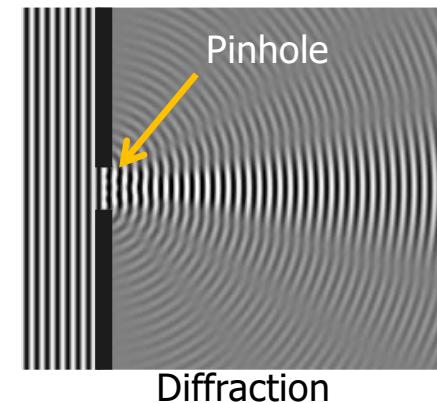


The pinhole must be tiny,  
but if it's too tiny it will cause diffraction.

Ideal pinhole diameter:

$$d \approx 2\sqrt{f\lambda}$$

$f$ : effective focal length  
 $\lambda$ : wavelength



# What about Exposure Time?

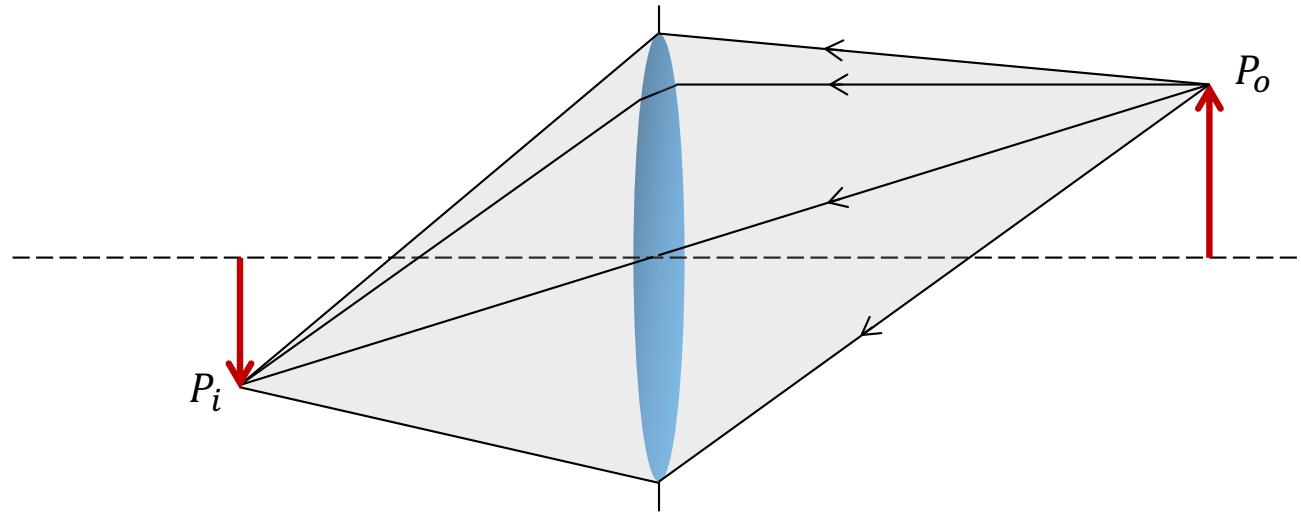
Pinholes pass less light and hence require **long exposures** to capture bright images.

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



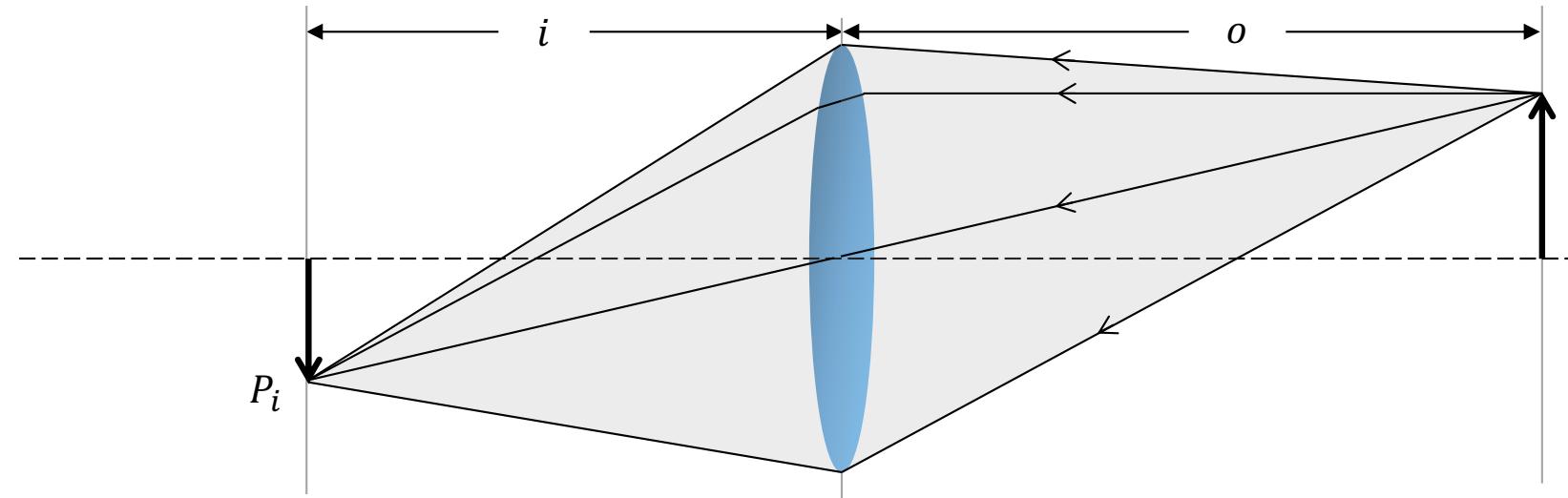
# Lenses

- 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

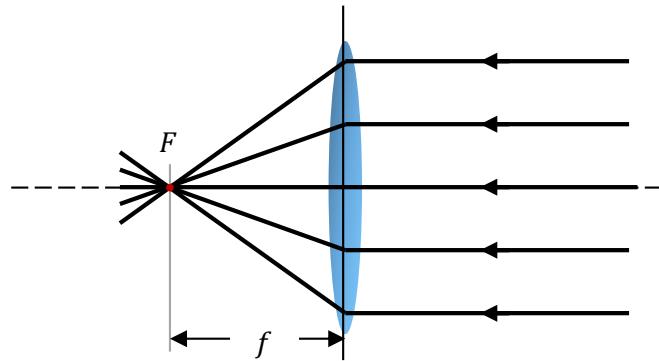
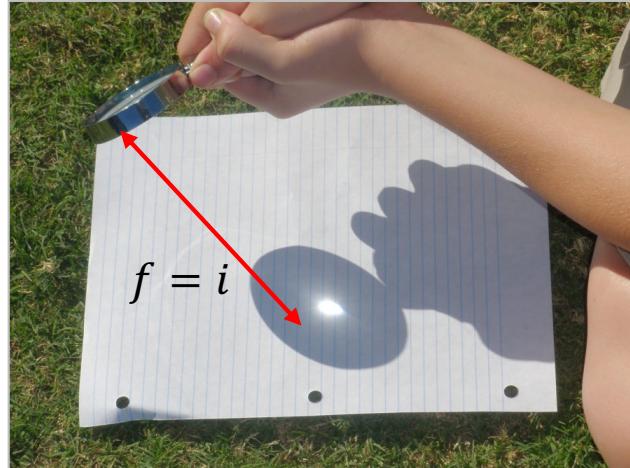
$o$ : object distance

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

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

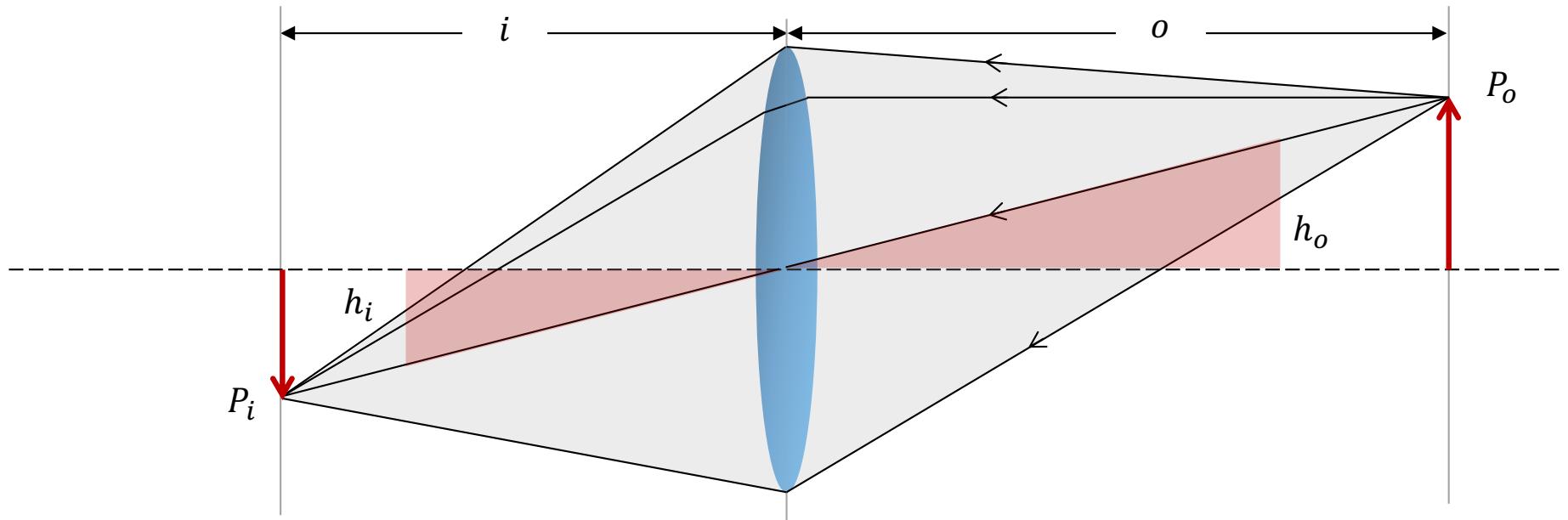
# How to Find the Focal Length

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



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

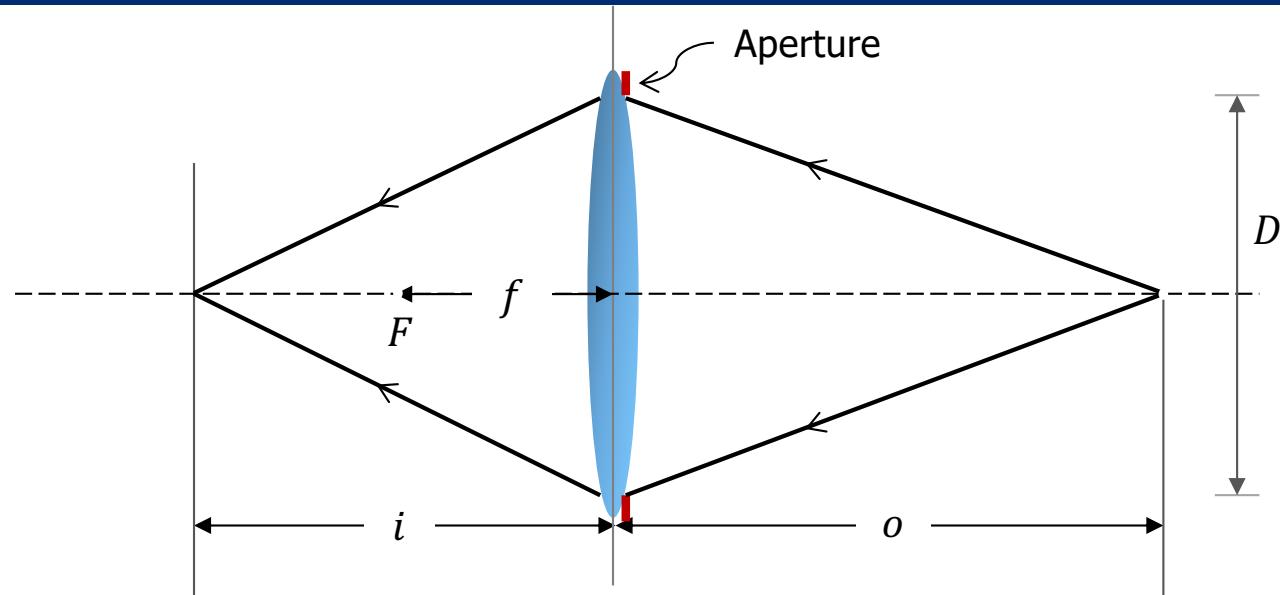
# Image Magnification



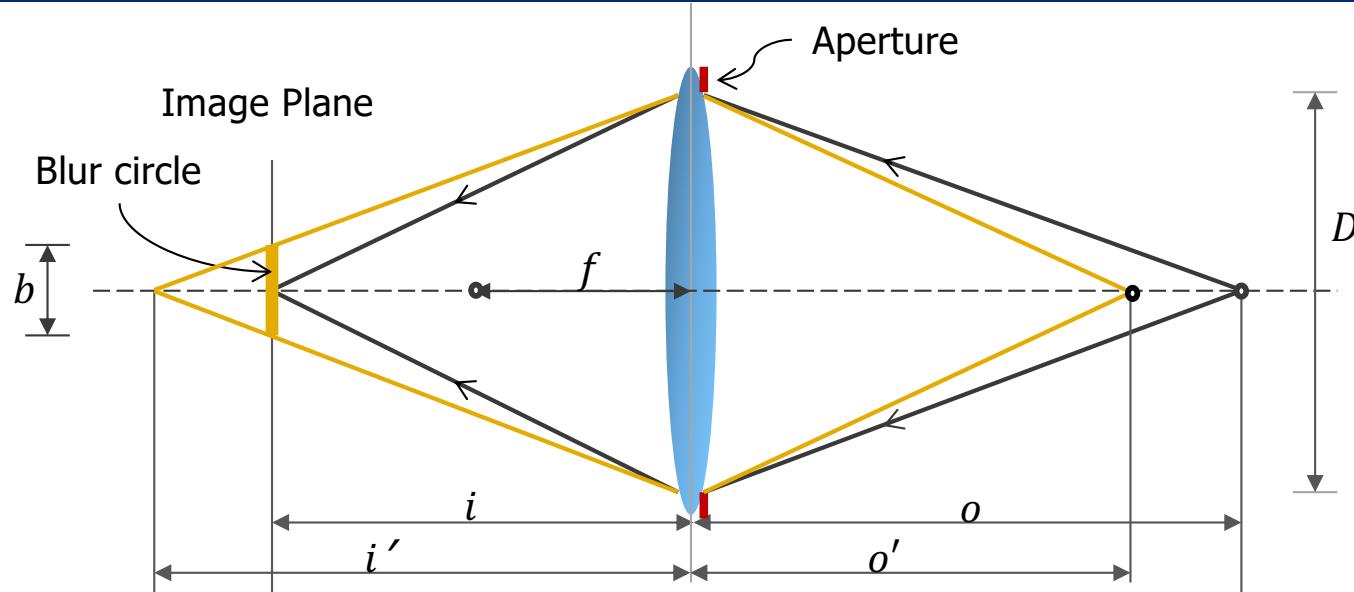
Magnification:

$$m = \frac{h_i}{h_o} = \frac{i}{o}$$

# Blur Circle (Defocus)



# Blur Circle (Defocus)



From similar triangles:

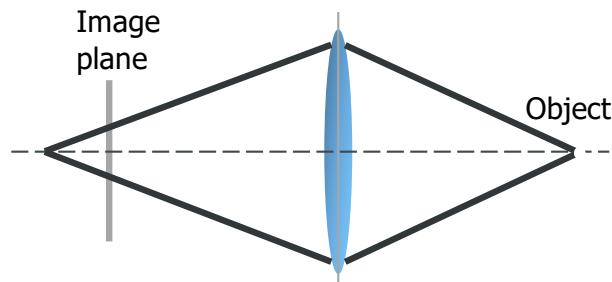
$$\frac{b}{D} = \frac{|i' - i|}{i'}$$

Blur circle diameter:

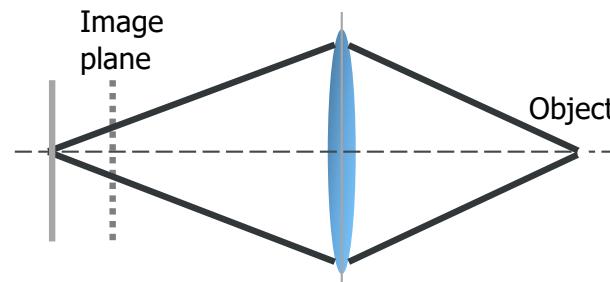
$$b = \frac{D}{i'} |i' - i|$$

$$b \propto D$$

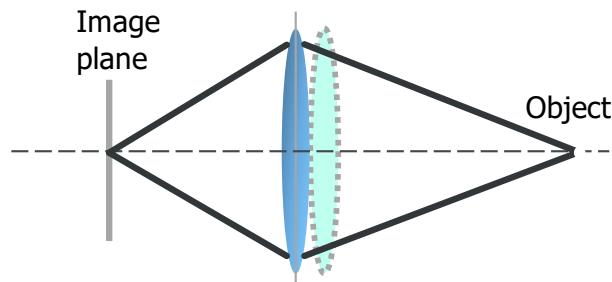
# Focusing



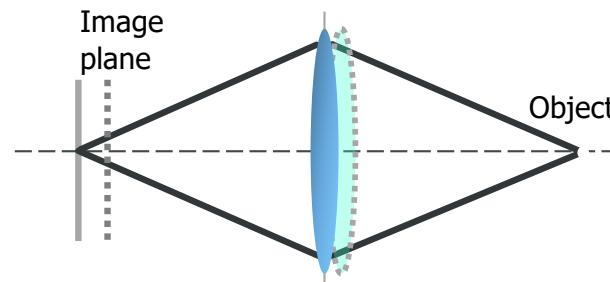
Defocused System



Move the image plane



Move the lens



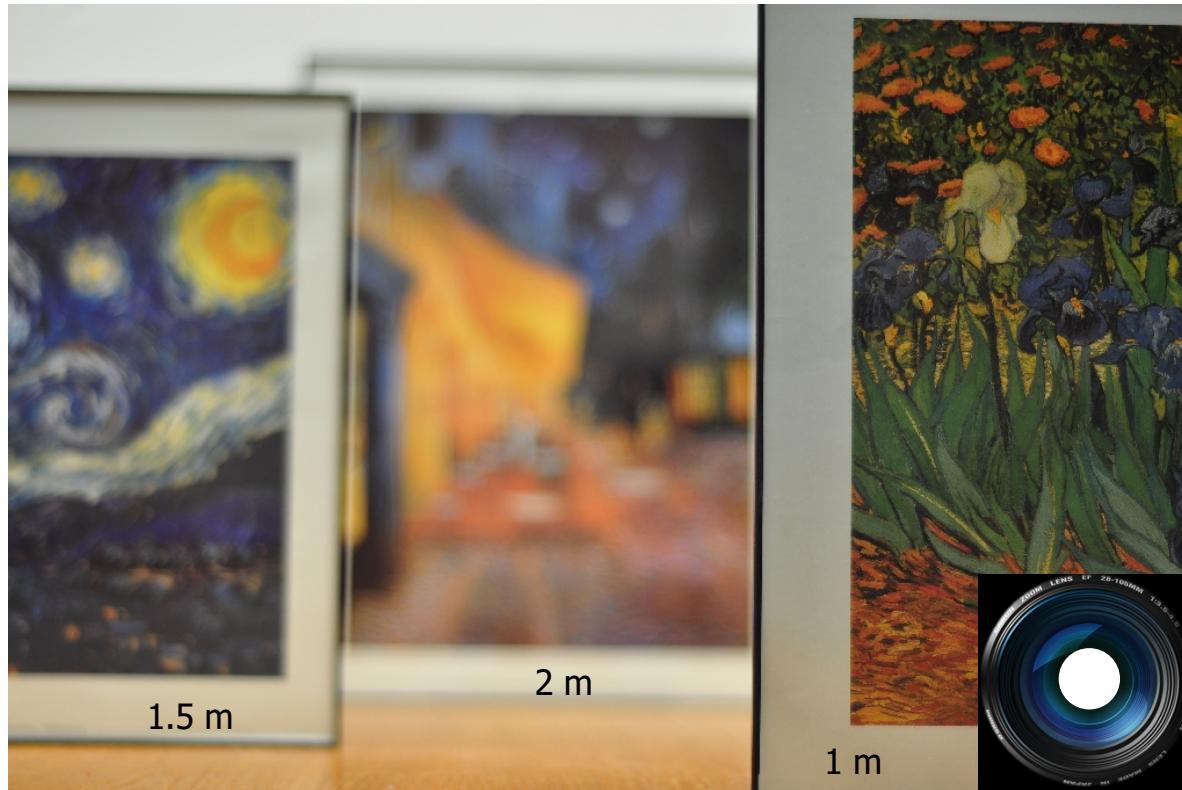
Move both lens and image plane

# Depth of Field (DoF)



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

# Aperture Size: DOF vs. Brightness



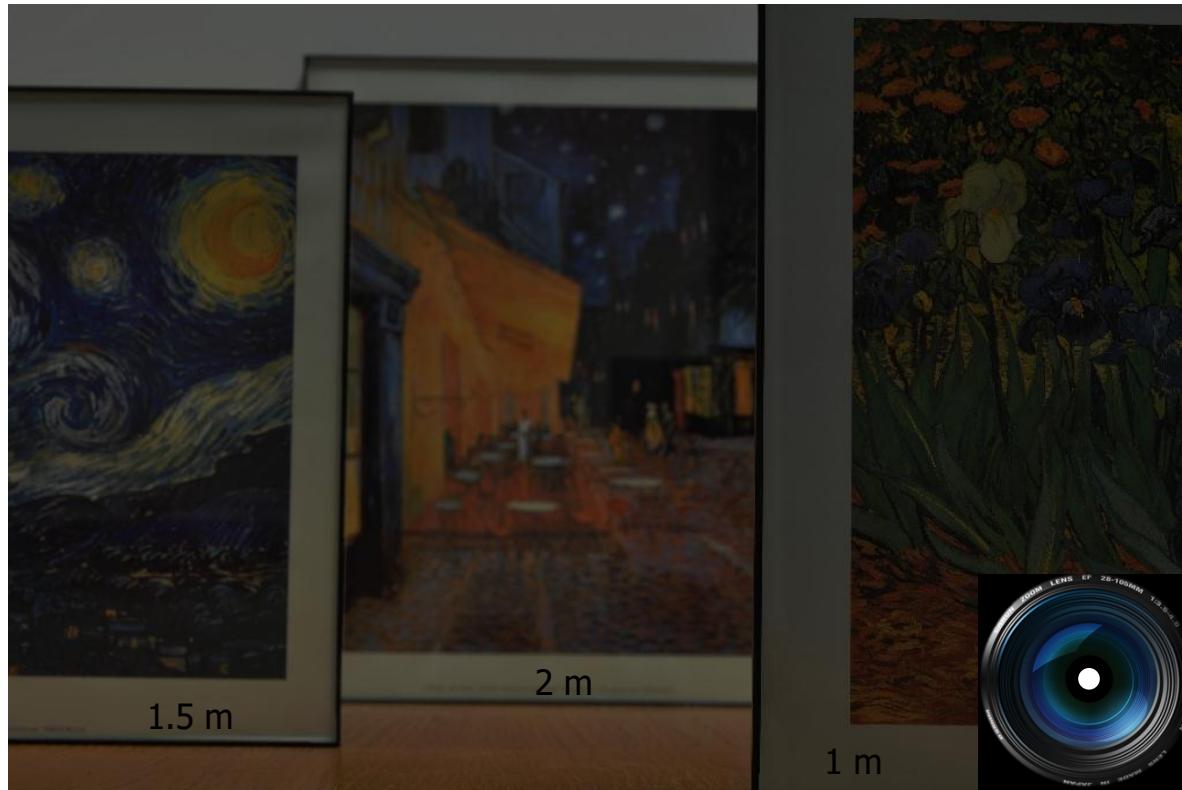
Focal Length 50 mm, Focus = 1 m, Aperture D = **25 mm**

# Aperture Size: DOF vs. Brightness



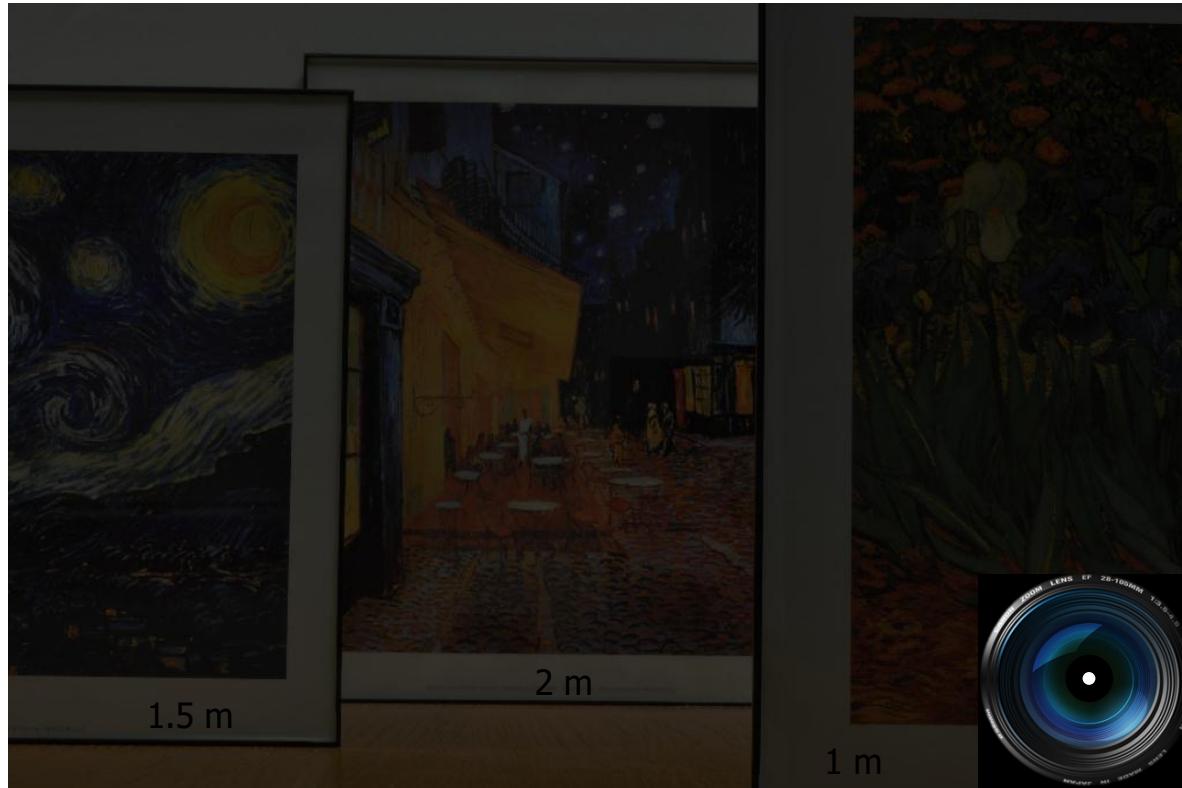
Focal Length 50 mm, Focus = 1 m, Aperture D = **12.5 mm**

# 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  $f/ \#$

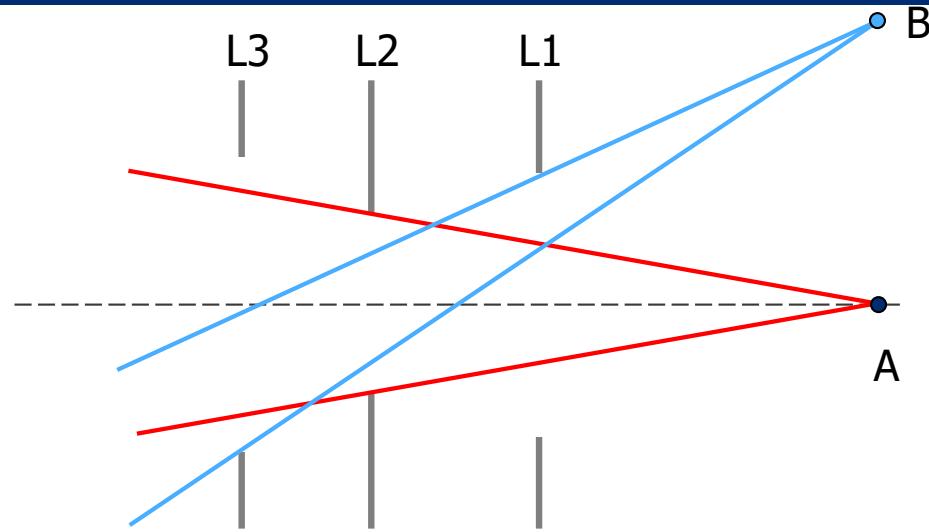
$$f/\# = N = f / d$$

$f$ : 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



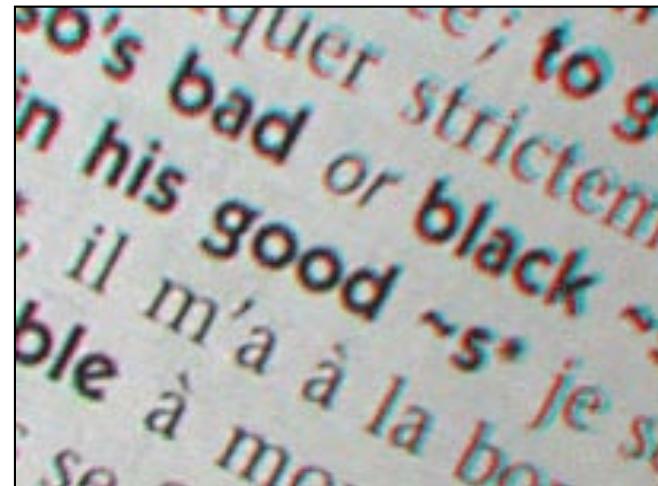
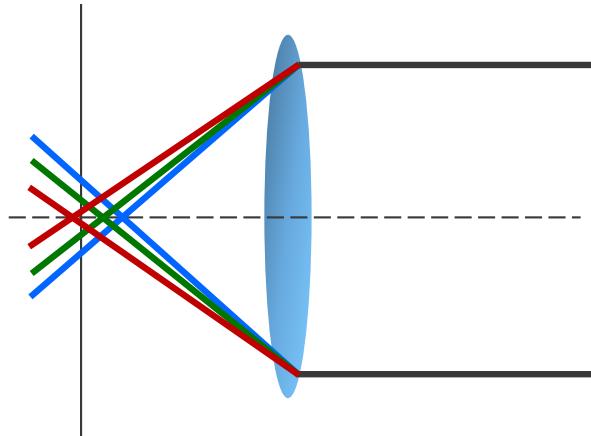
Brightness fall-off (Vignetting)  
in image of a White Wall



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

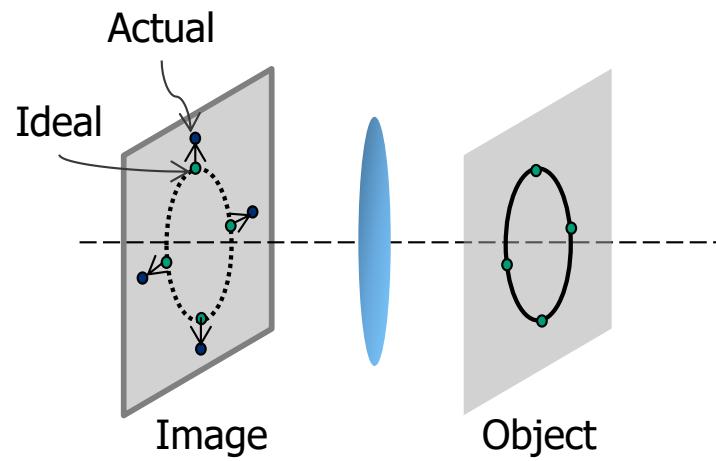
I.15

# Chromatic Aberration

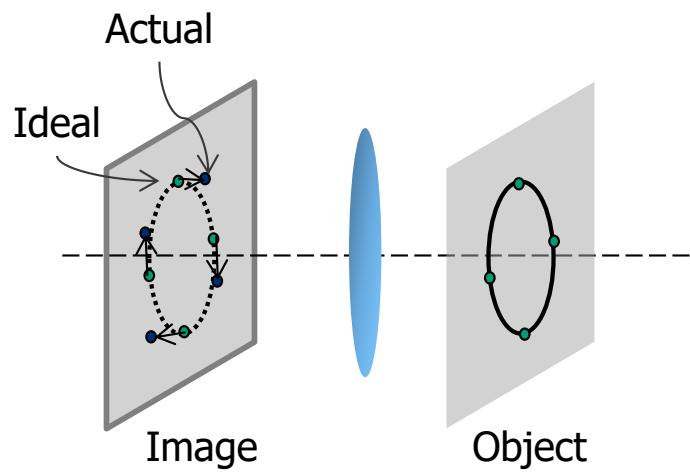


Refractive index (and hence focal length) of lens is different for different wavelengths.

# Geometric Distortion



**Radial** distortion



**Tangential** distortion

Due to lens imperfections

# When Geometric Distortion is Useful



Fisheye Lens



Radial Distortion in Fisheye Lens

# Geometric Distortion Correction



Radial (Barrel) distortion



Undistorted image

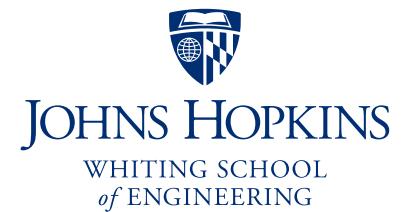
# Summary

- **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



# 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

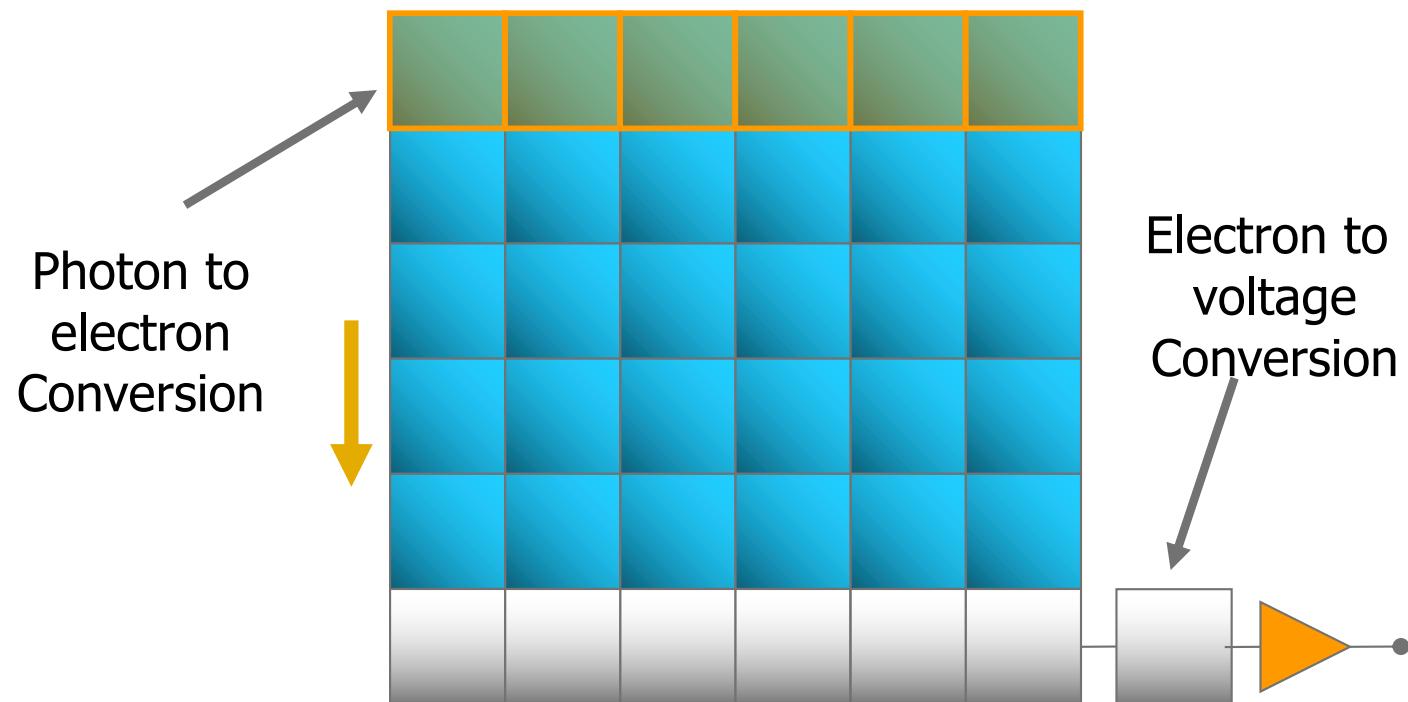
# The Digital Camera

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):**

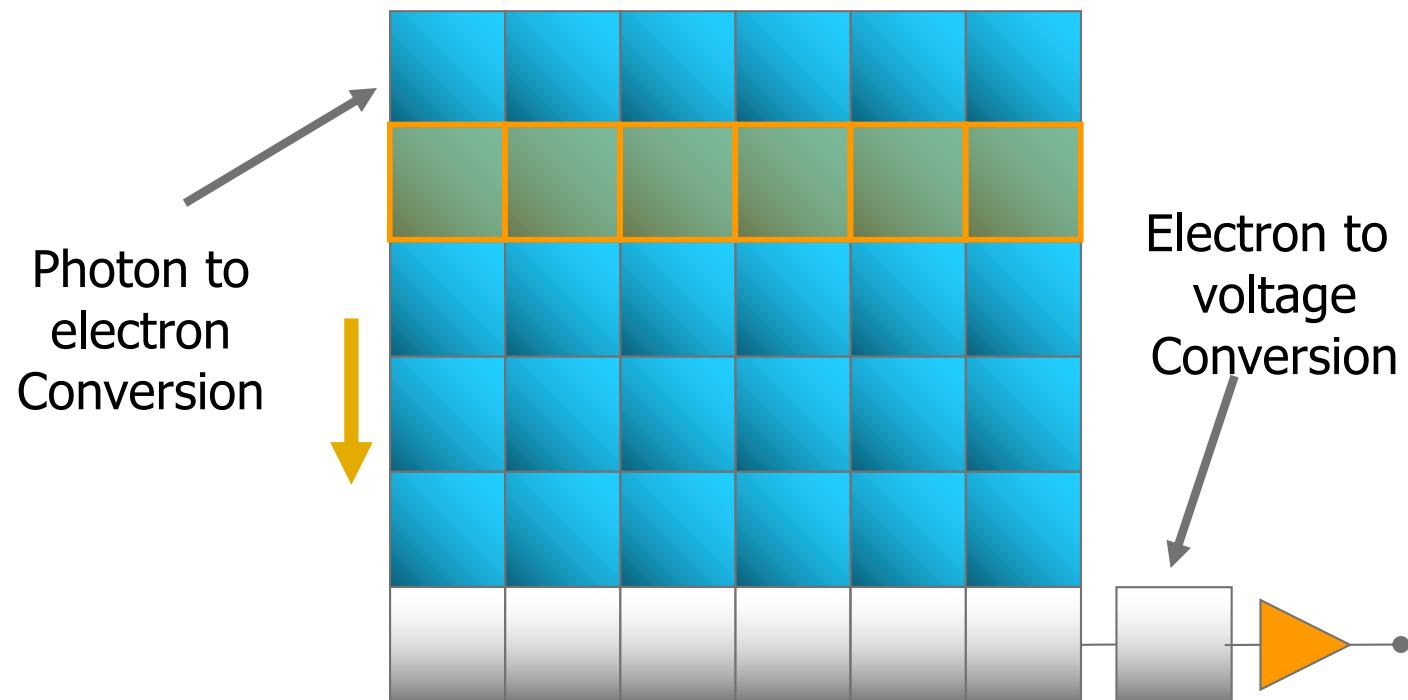
# Popular Types of Image Detectors

Charge Coupled Device (CCD)



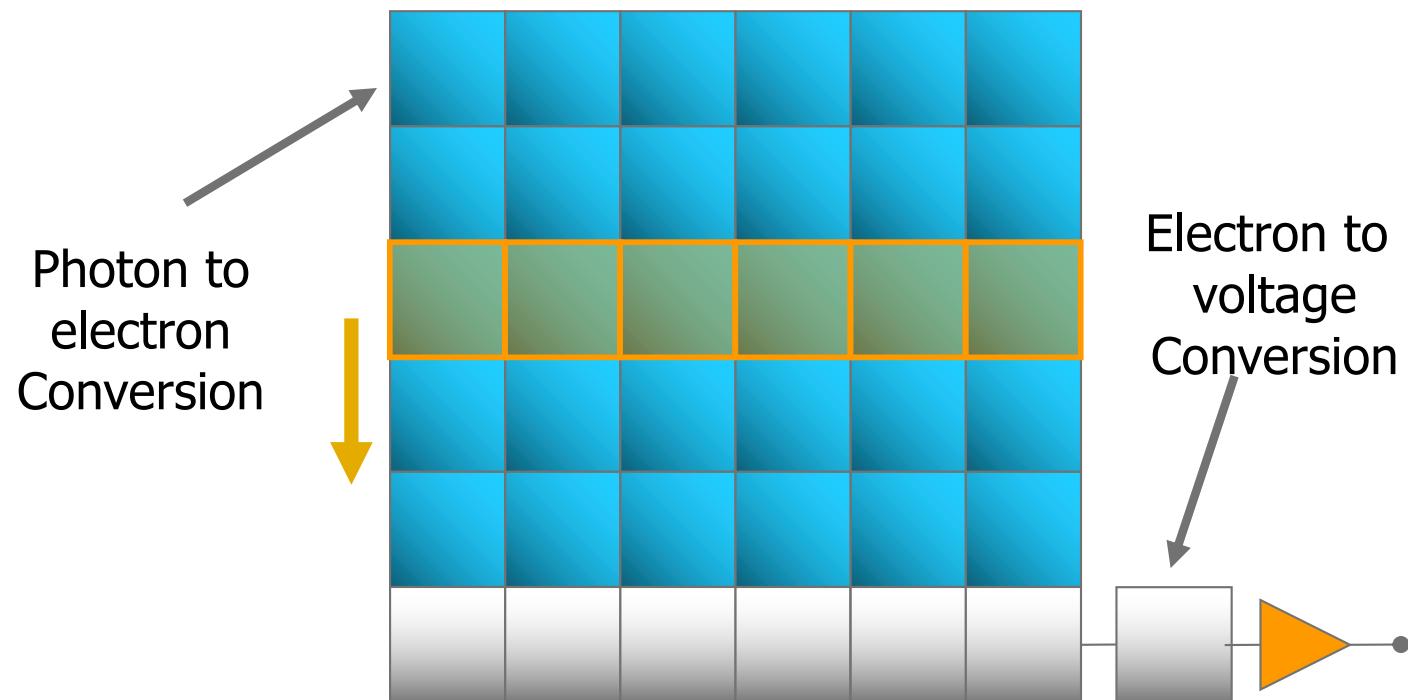
# Popular Types of Image Detectors

Charge Coupled Device (CCD)



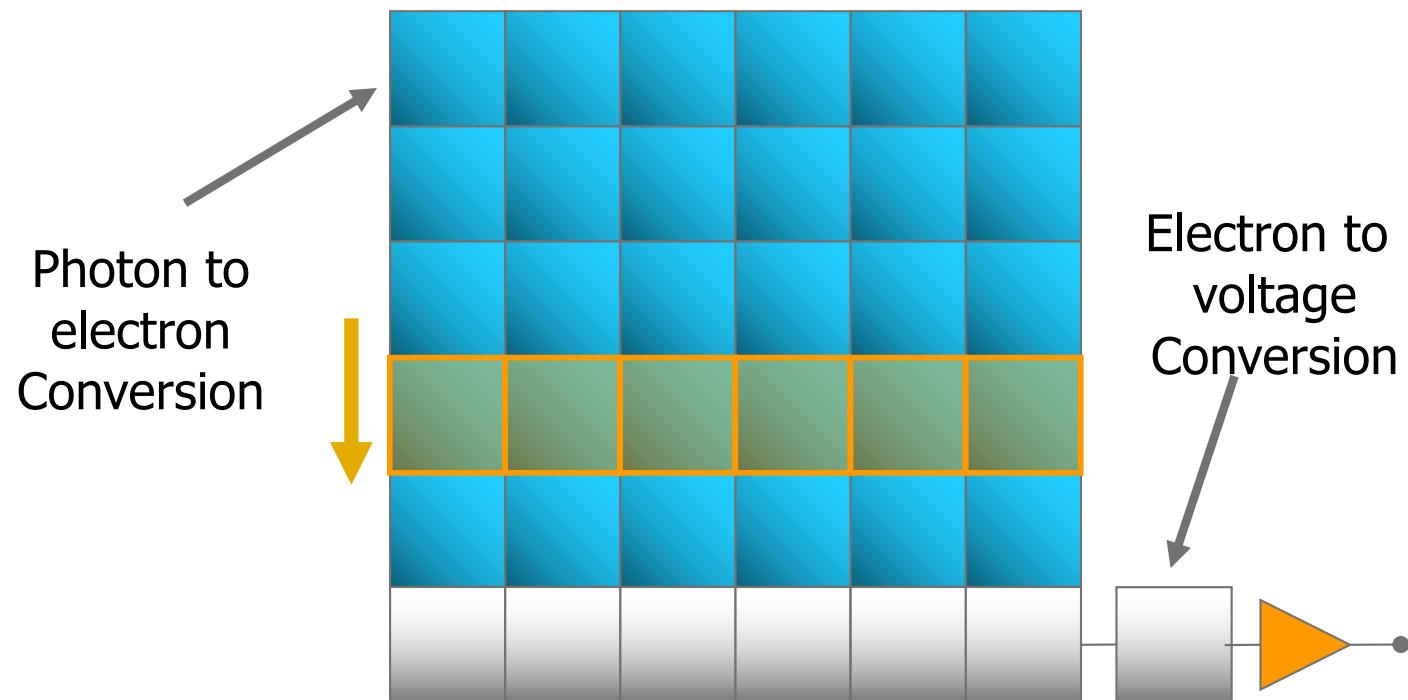
# Popular Types of Image Detectors

Charge Coupled Device (CCD)



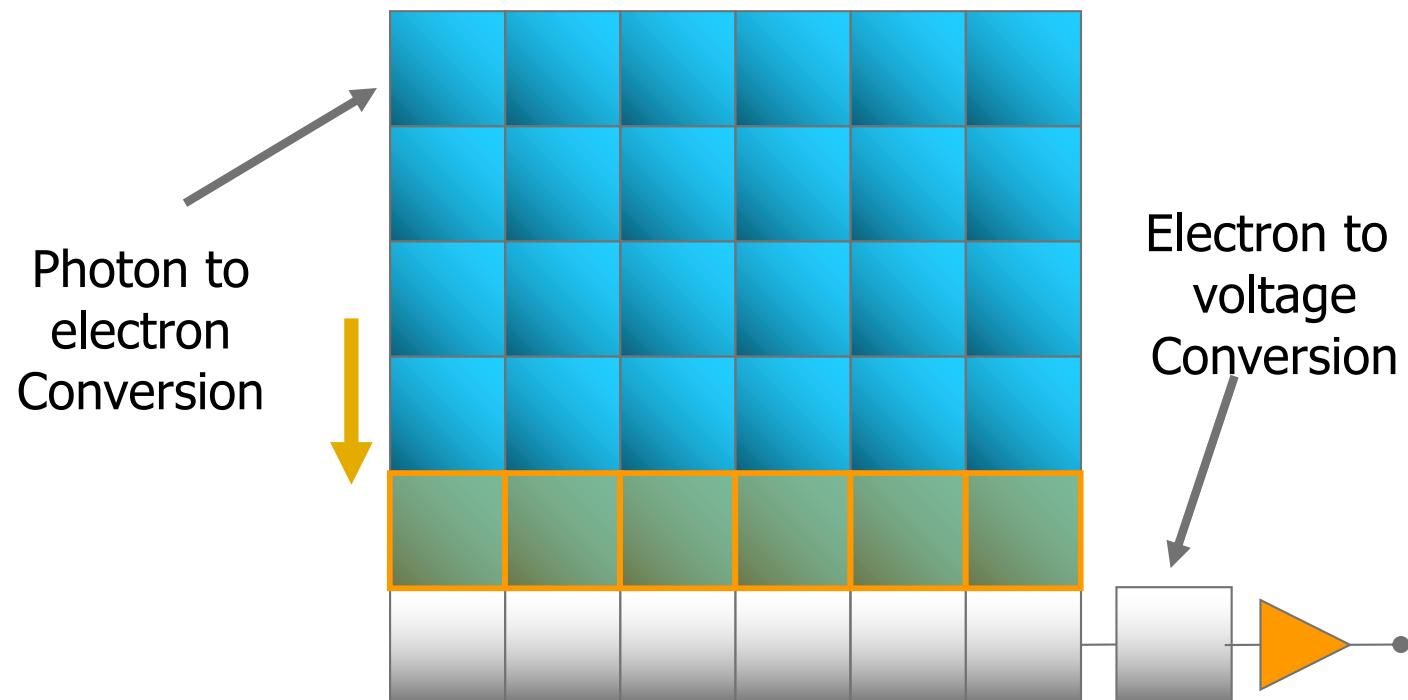
# Popular Types of Image Detectors

Charge Coupled Device (CCD)



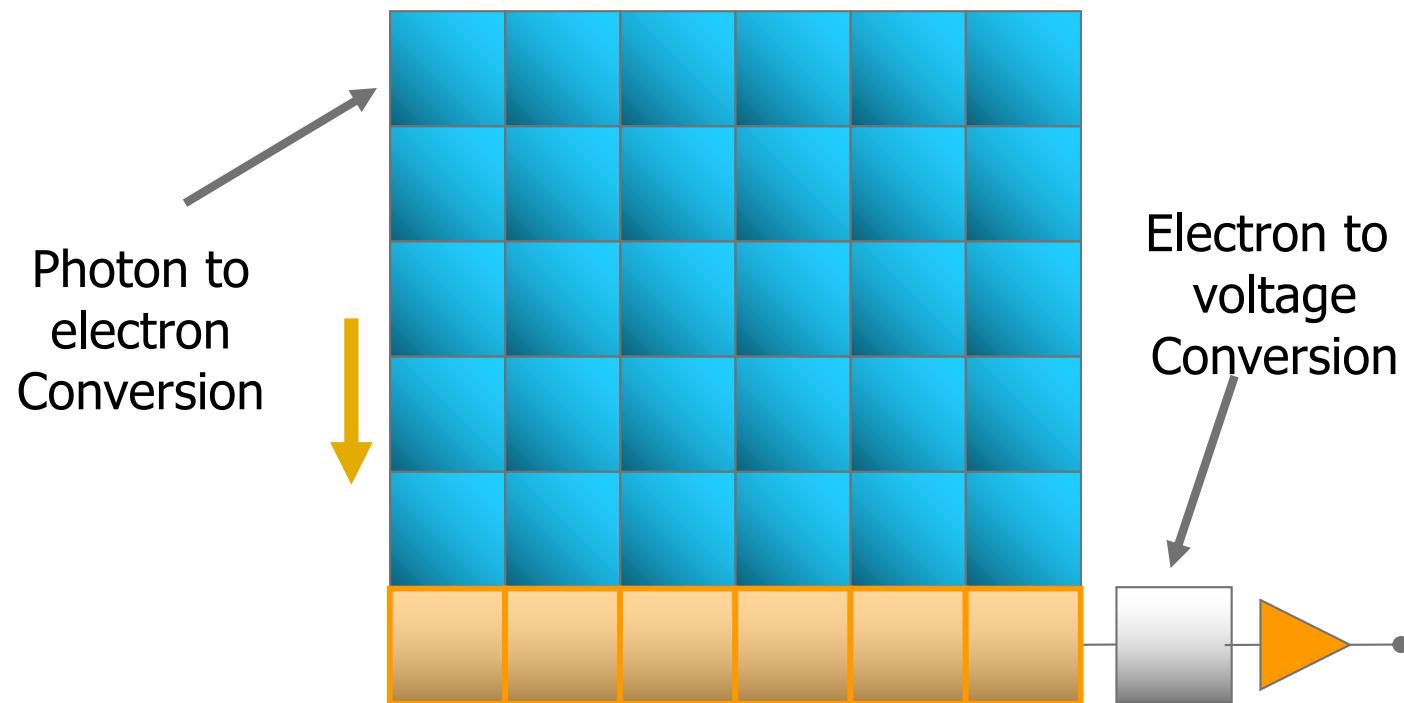
# Popular Types of Image Detectors

Charge Coupled Device (CCD)



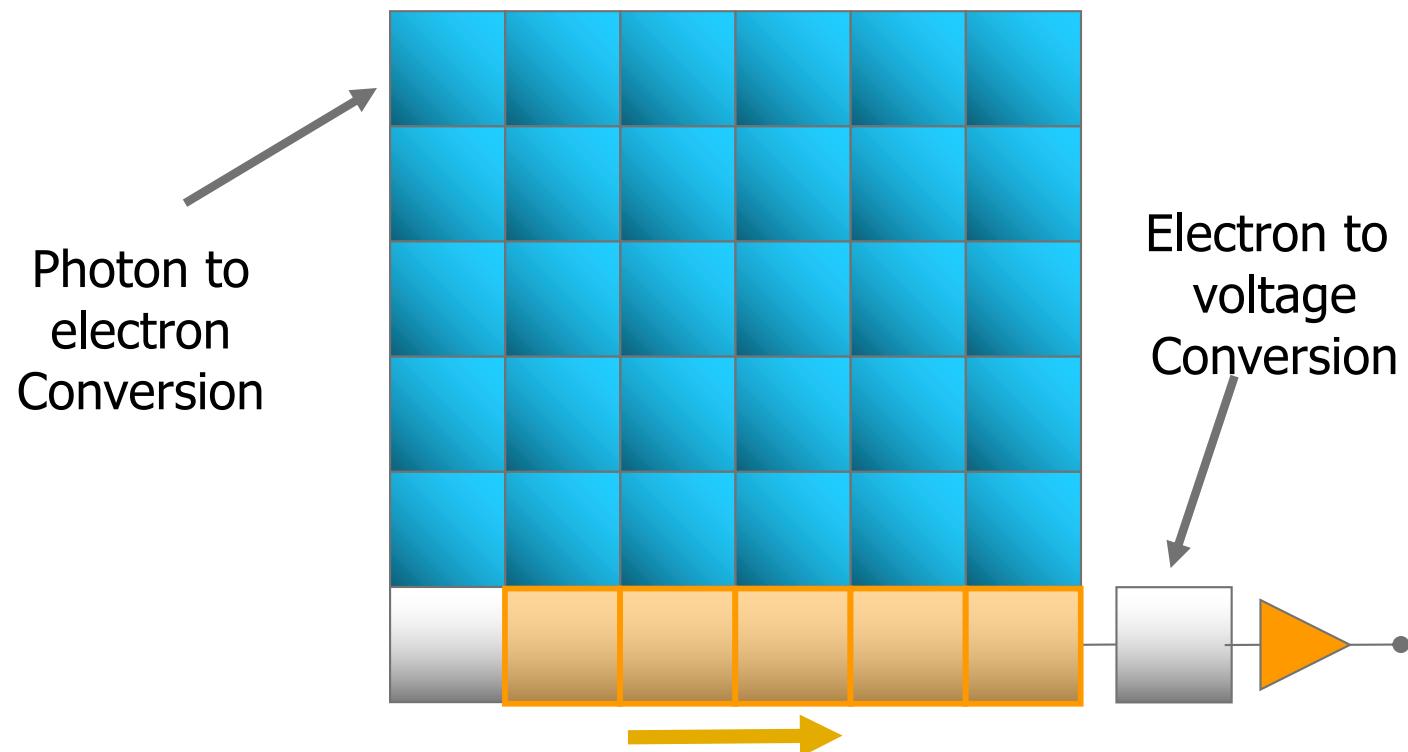
# Popular Types of Image Detectors

Charge Coupled Device (CCD)



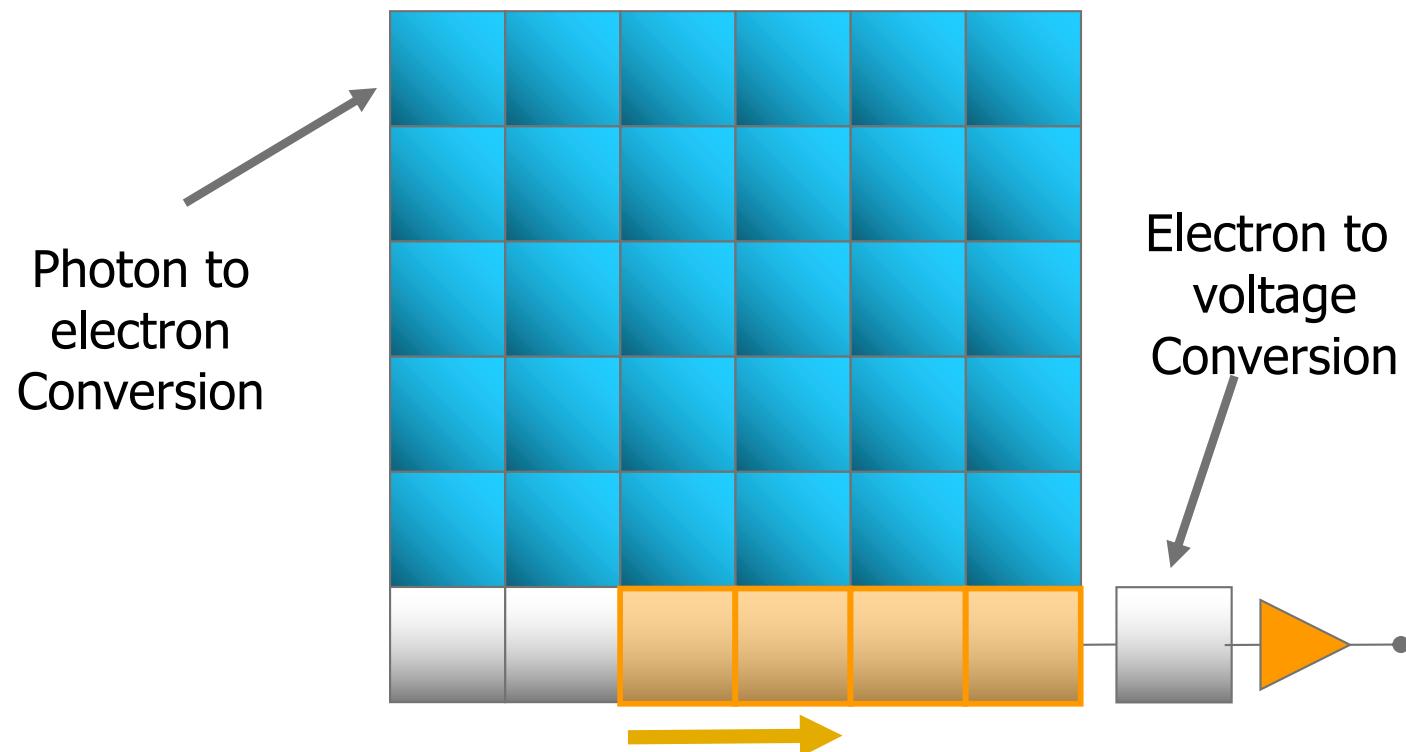
# Popular Types of Image Detectors

Charge Coupled Device (CCD)



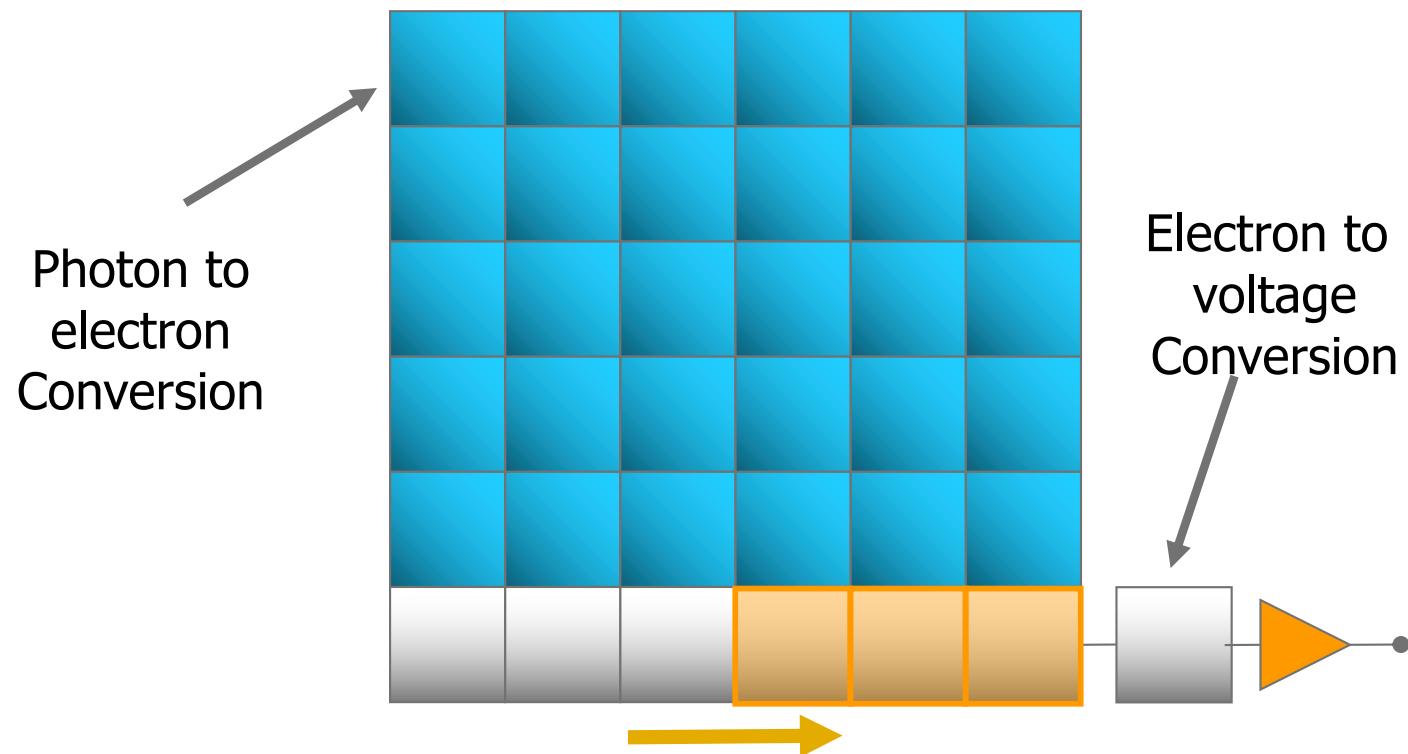
# Popular Types of Image Detectors

Charge Coupled Device (CCD)



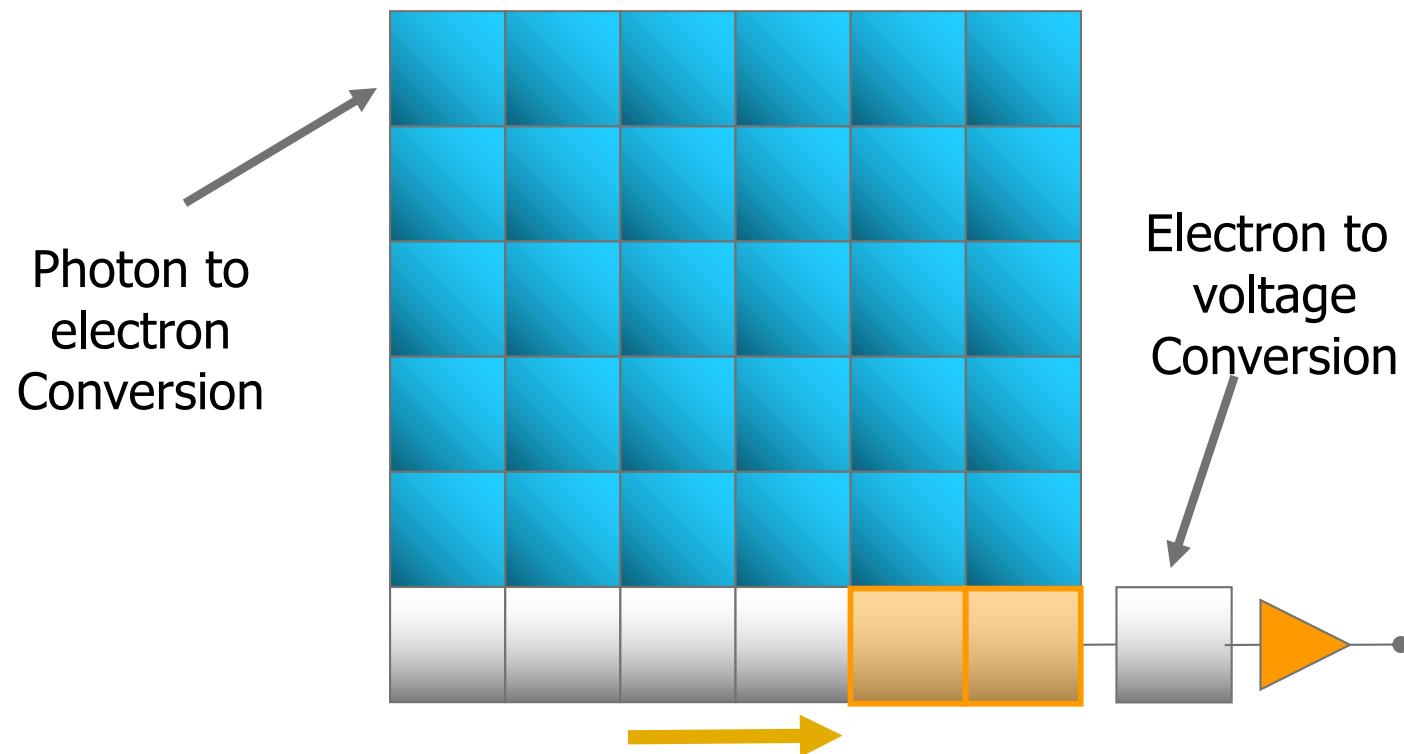
# Popular Types of Image Detectors

Charge Coupled Device (CCD)



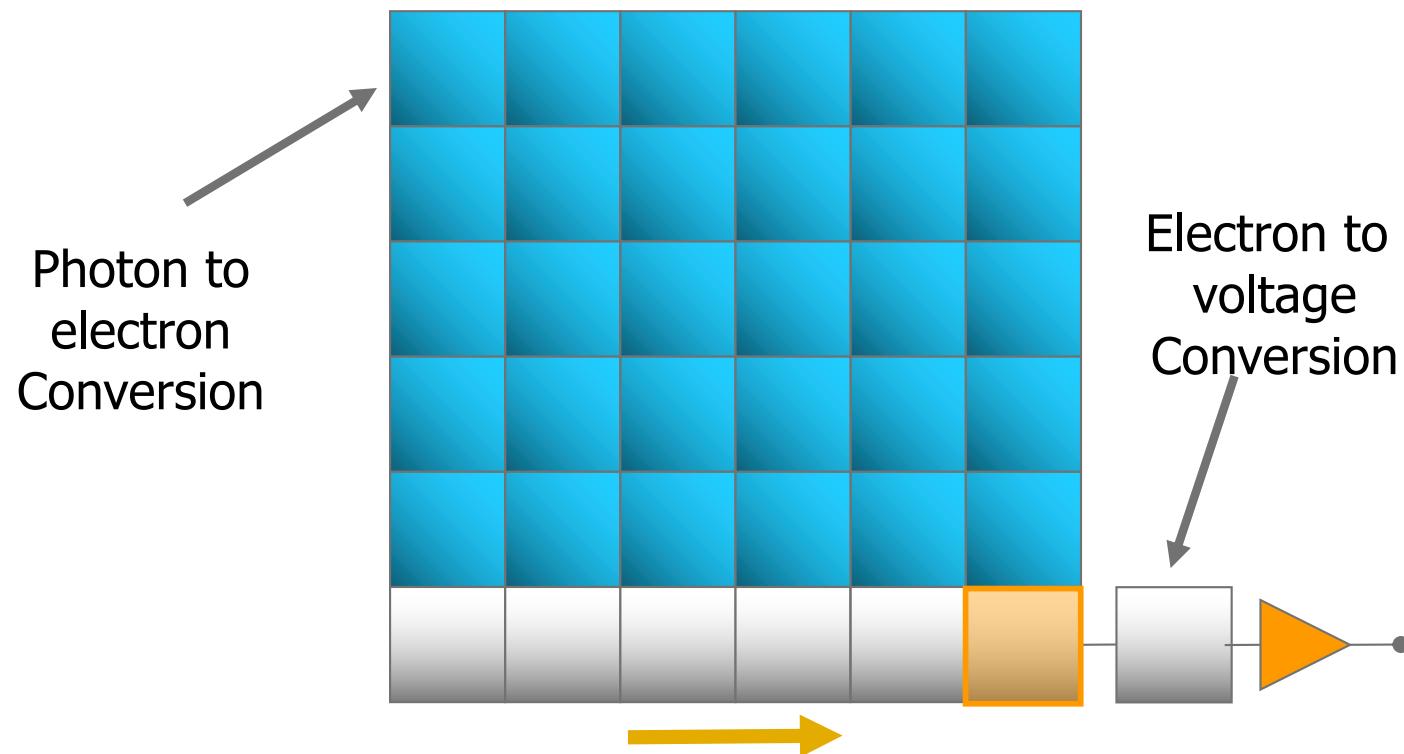
# Popular Types of Image Detectors

Charge Coupled Device (CCD)



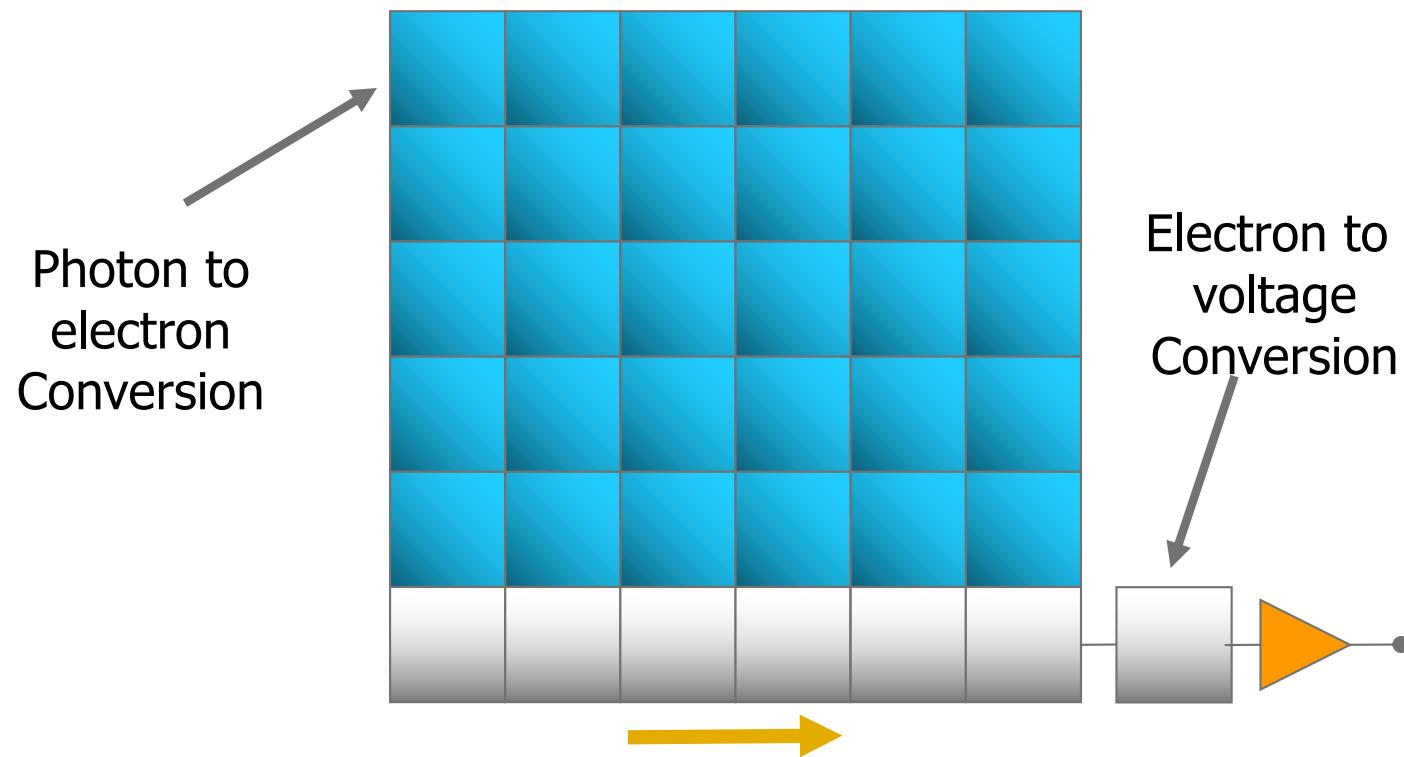
# Popular Types of Image Detectors

Charge Coupled Device (CCD)



# Popular Types of Image Detectors

Charge Coupled Device (CCD)



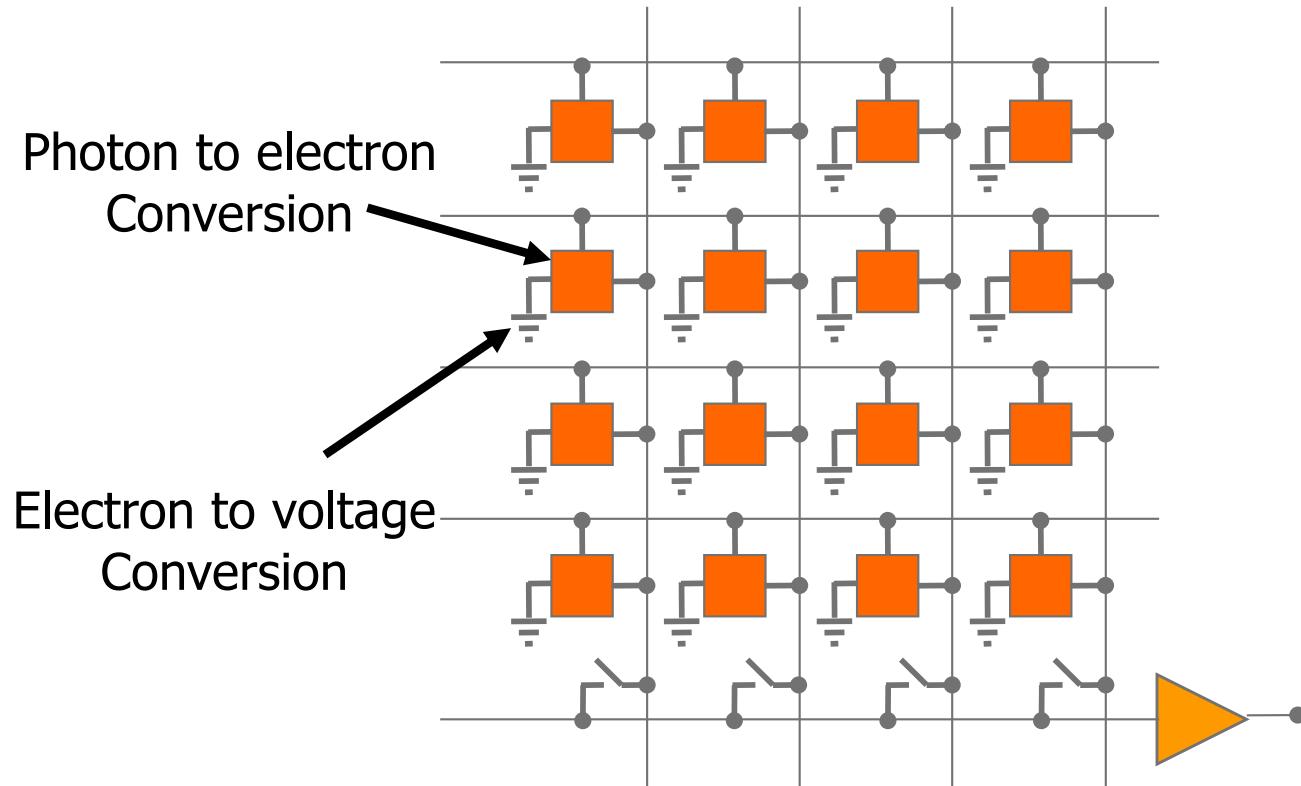
# The Digital Camera

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.

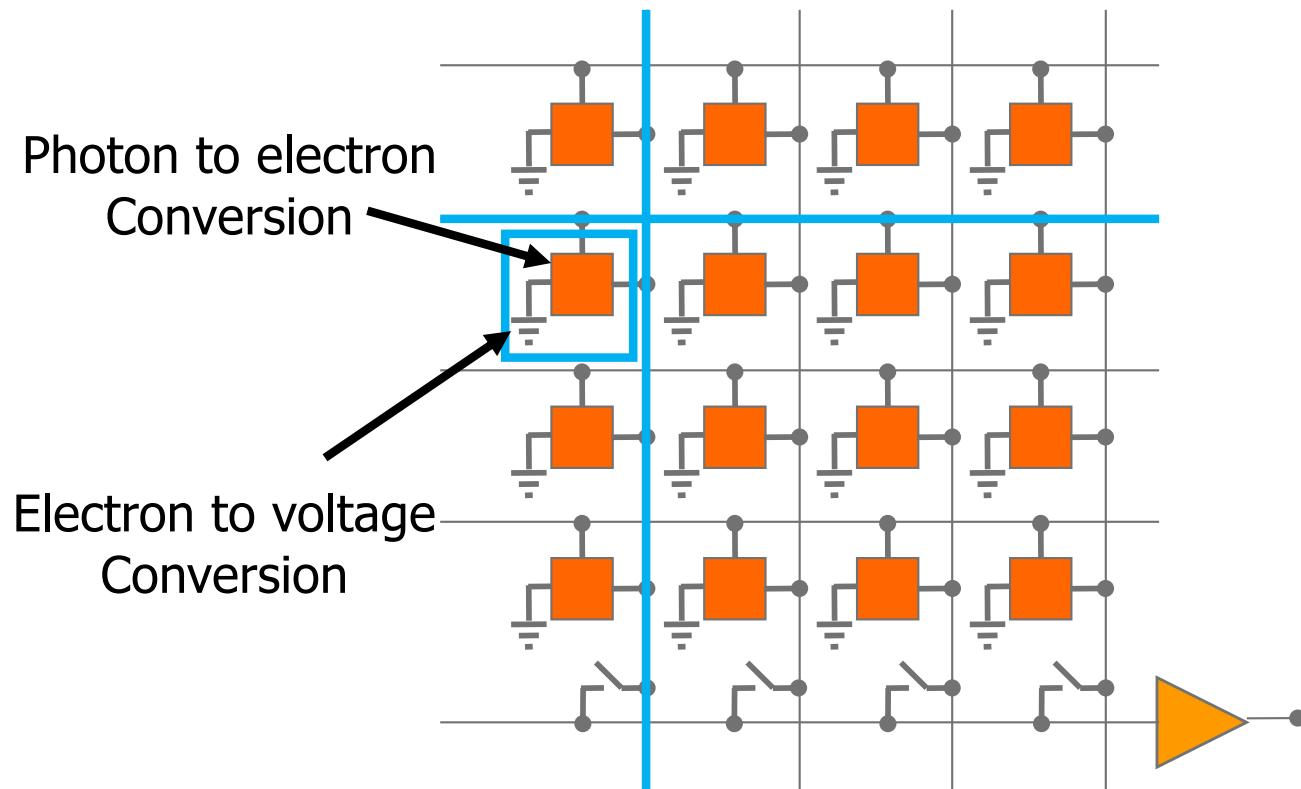
# Popular Types of Image Detectors

Complimentary Metal-Oxide Semiconductor (CMOS)



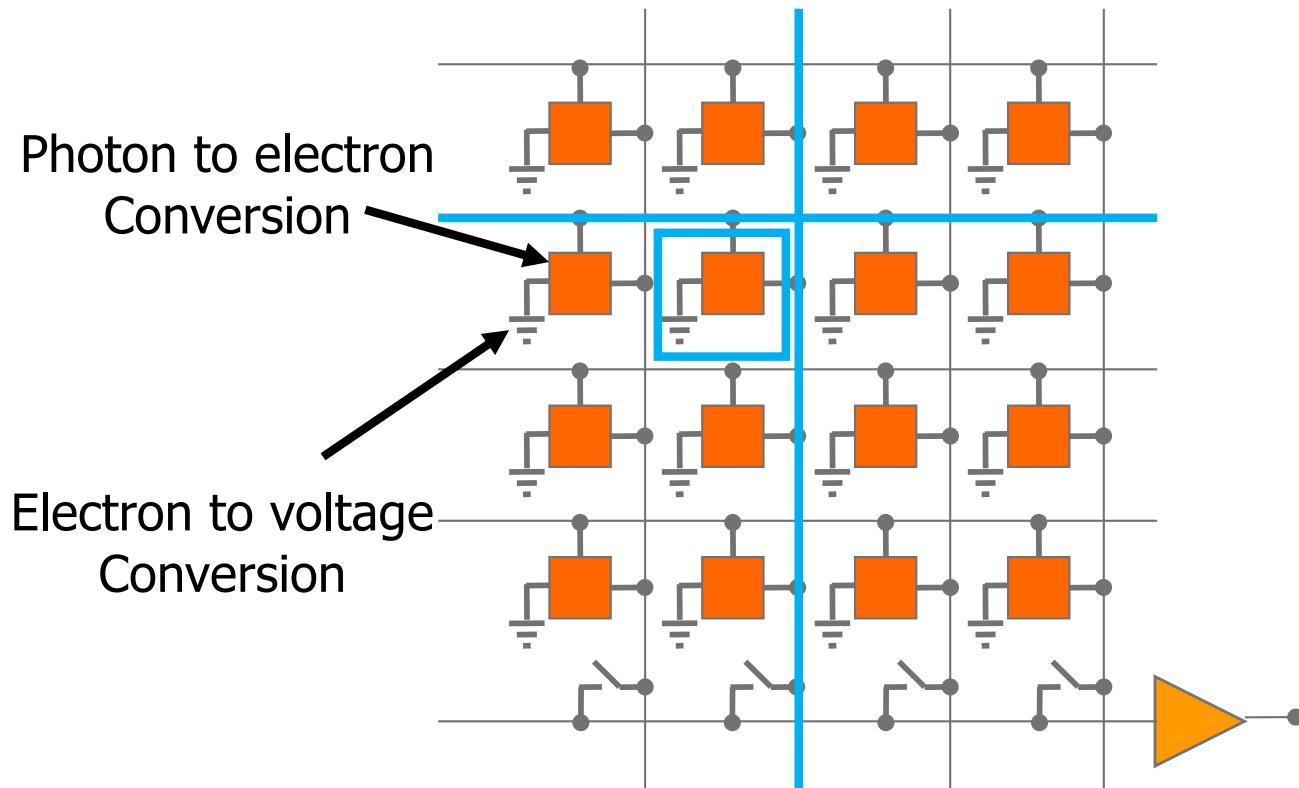
# Popular Types of Image Detectors

Complimentary Metal-Oxide Semiconductor (CMOS)



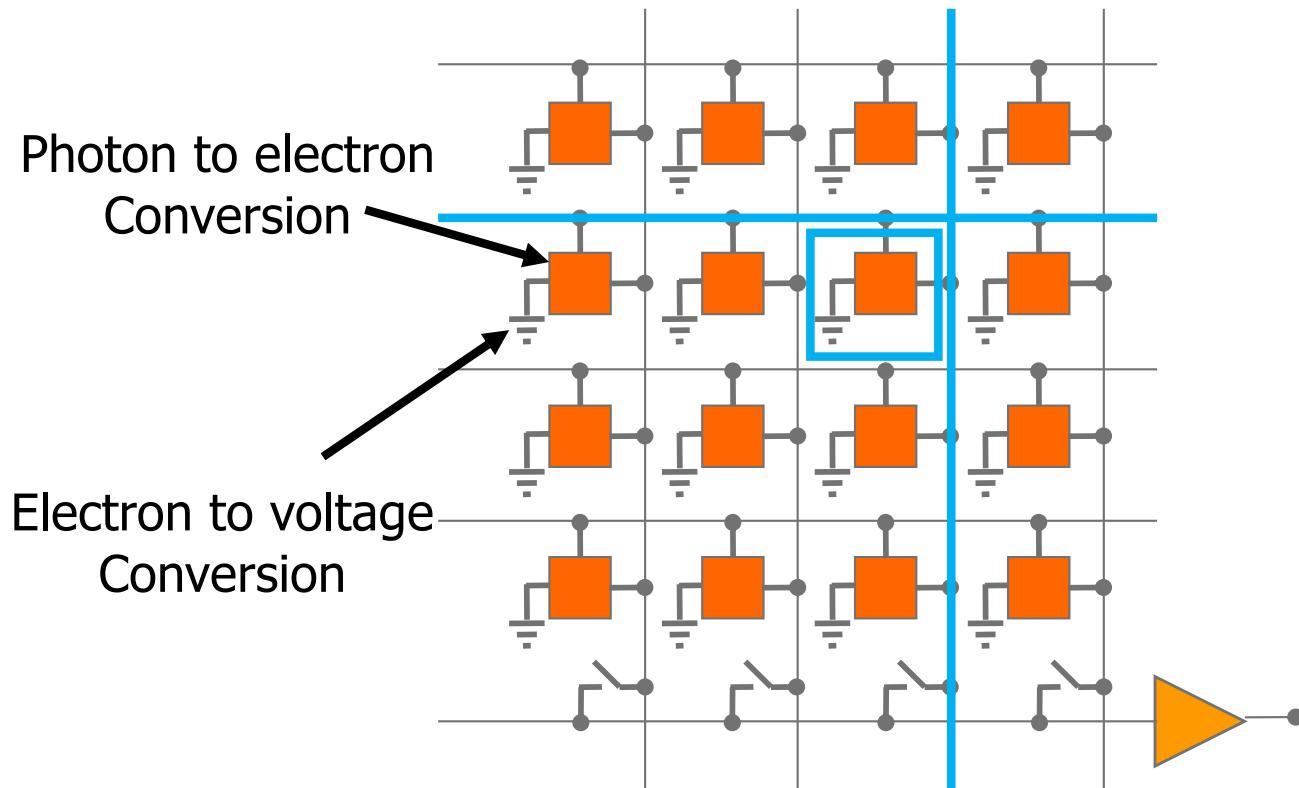
# Popular Types of Image Detectors

Complimentary Metal-Oxide Semiconductor (CMOS)



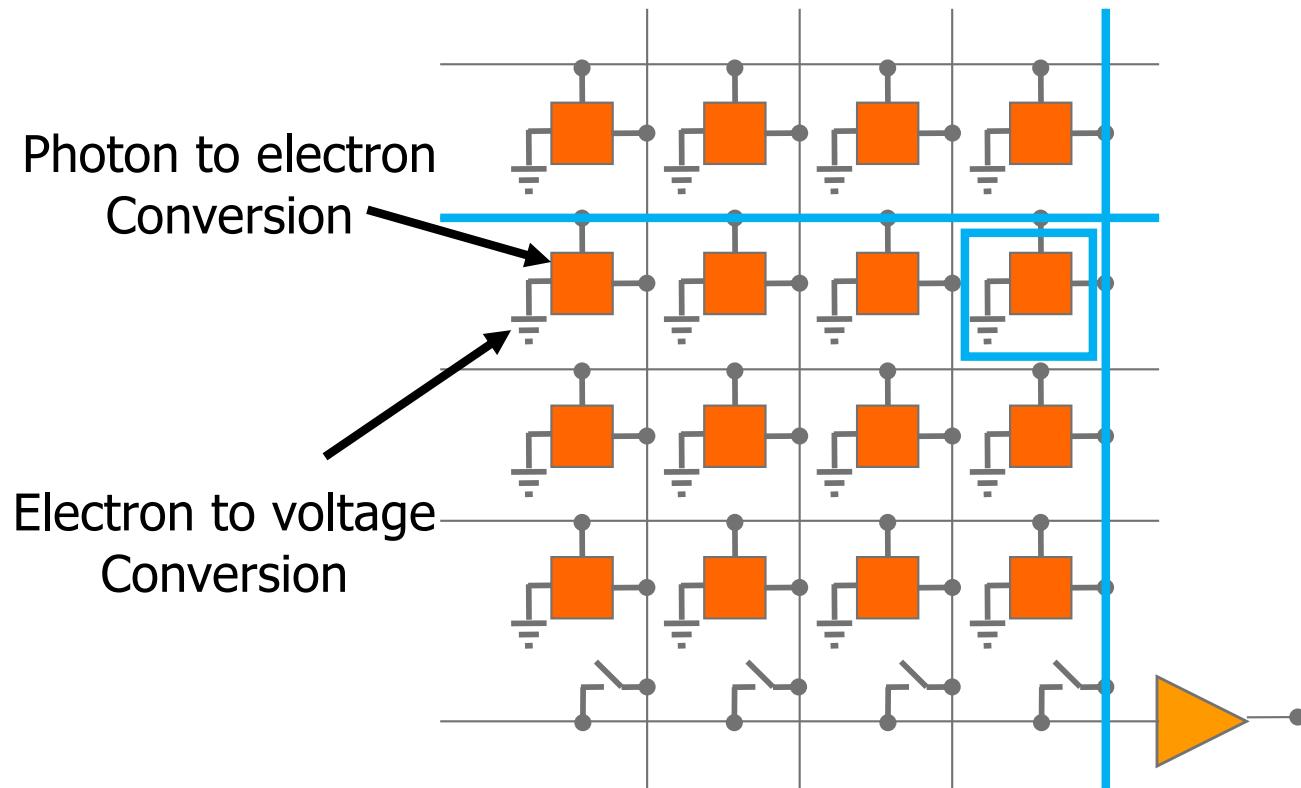
# Popular Types of Image Detectors

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

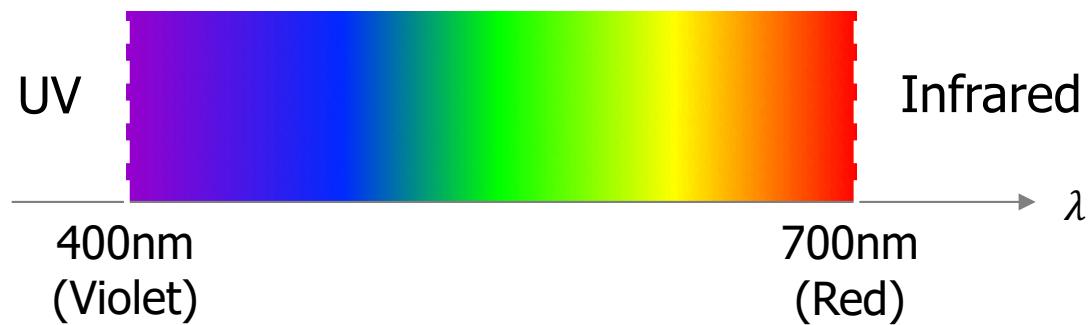
# What's Going on Here?



# What is “Color”?

Human response to different wavelengths

Visible light:



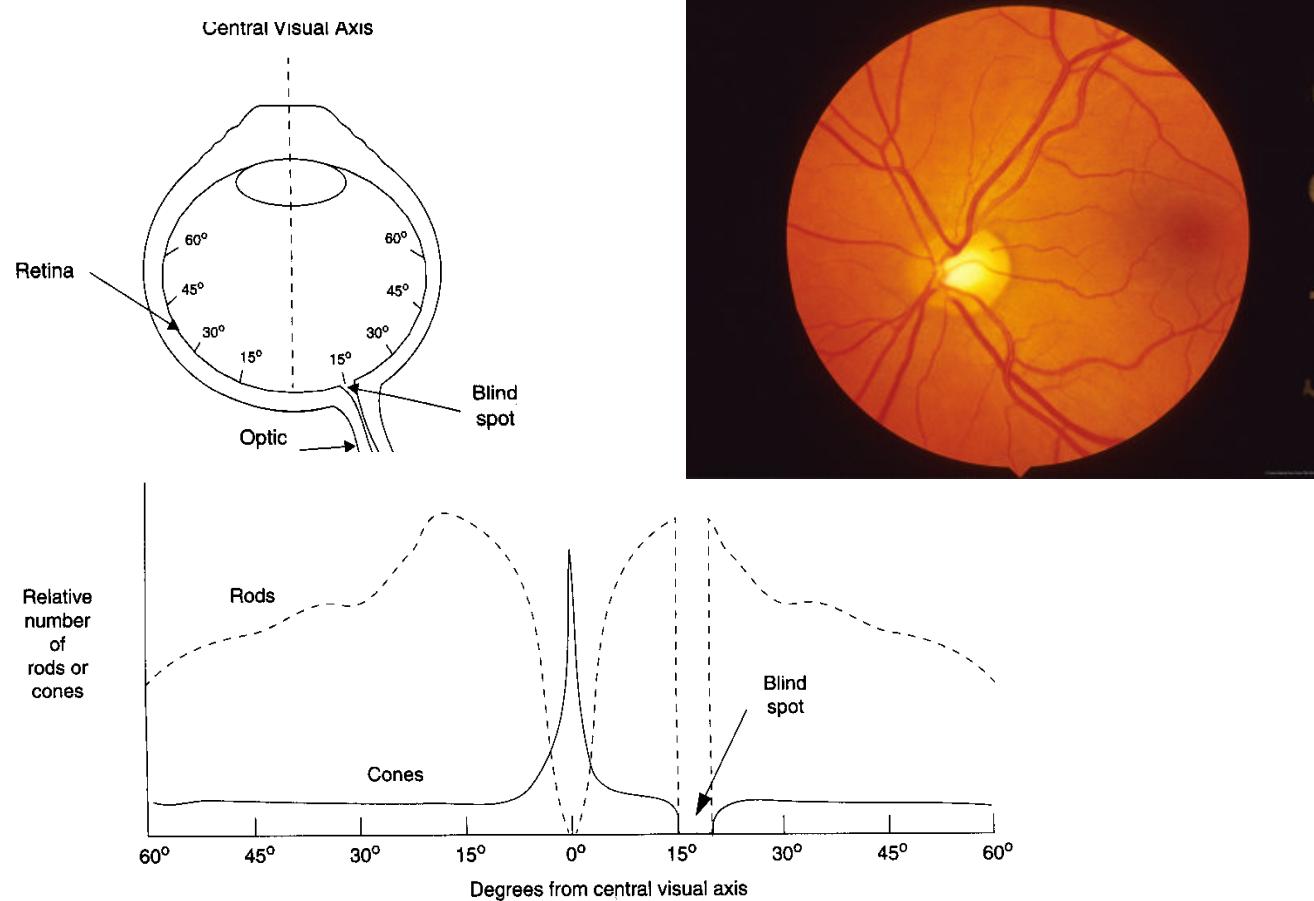
How do we recover Spectral Distribution  $E(\lambda)$ ?

# Color Measurement

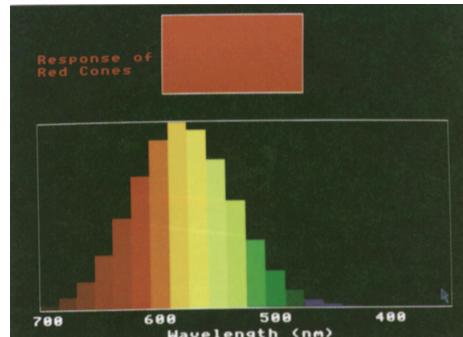
- Let  $\lambda$  denote wavelength
- Let  $E(\lambda)$  denote the spectral power at a given wavelength
- Let  $\rho_k(\lambda)$  denote the responsiveness of a sensor  $k$  to a given wavelength of light
- Then we can compute the “response”  $r_k$  of  $k$  as

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

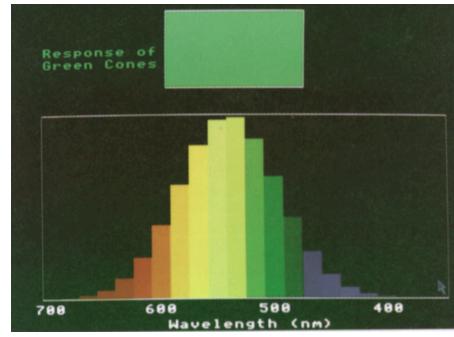
# Example: The Human Eye



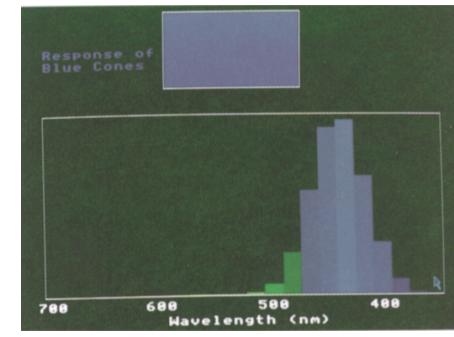
# Example: The Human Eye



"Red" cone

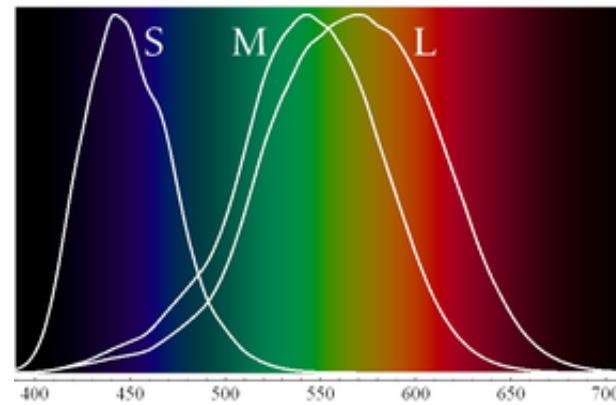


"Green" cone

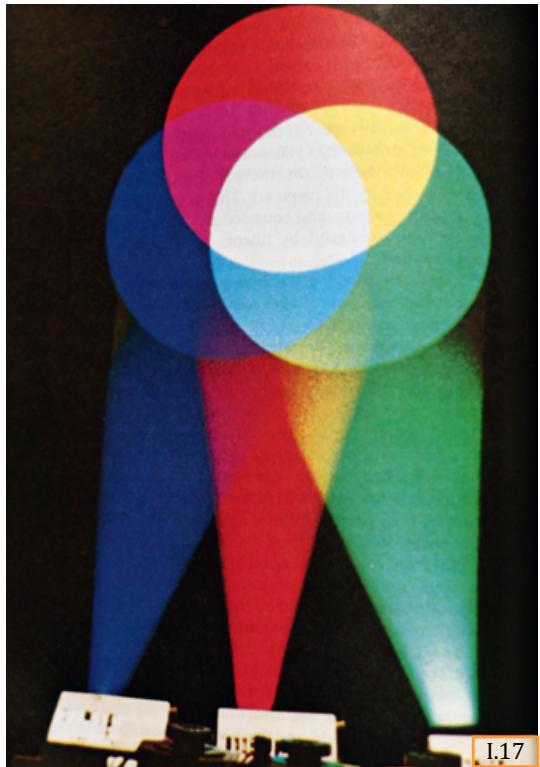


"Blue" cone

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

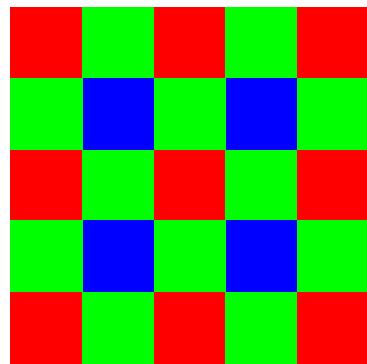
Human Sensation of nearly all colors can be produced using 3 wavelengths!

$$(\lambda_r \ \lambda_g \ \lambda_b) = (650 \ 530 \ 410) \text{ nm}$$

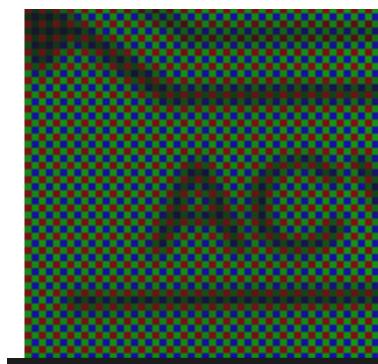
Hence, cameras and displays often use 3 filters:

(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

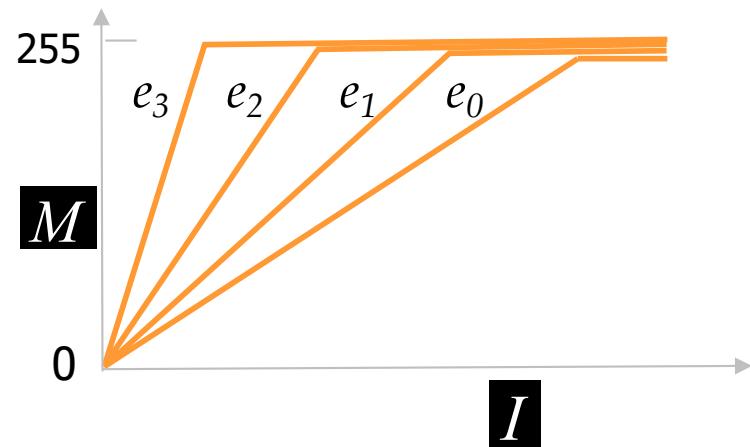


$$M_0 = \min(e_0 \cdot I, 255)$$

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

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

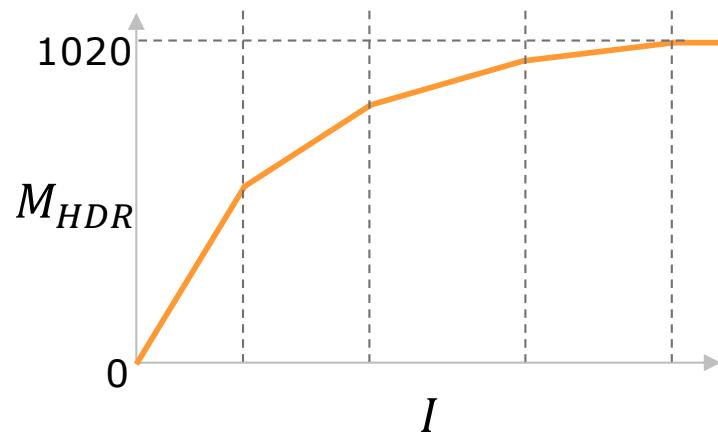
$$M_3 = \min(e_3 \cdot I, 255)$$



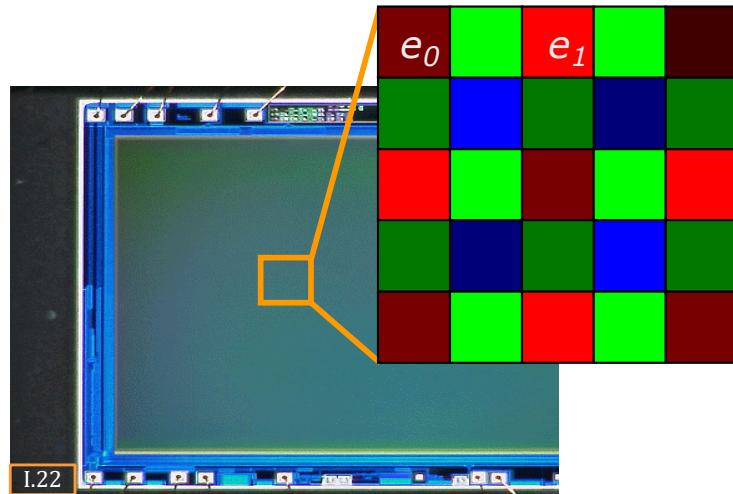
# High Dynamic Range: Multiple Exposures

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

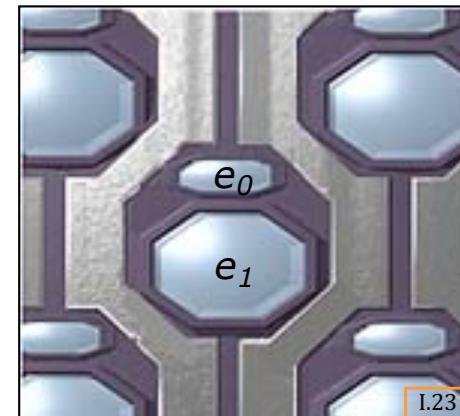
Camera Response  $f(\cdot)$  for Aggregate Image:



# High Dynamic Range: Single Shot



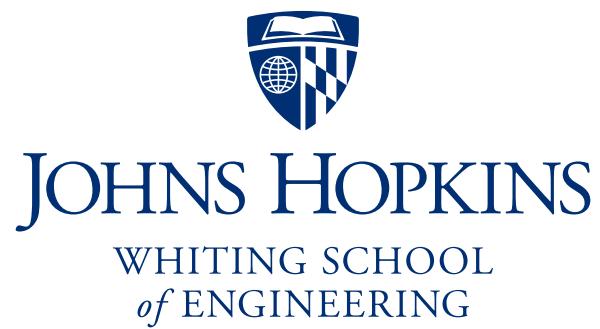
Assorted Pixels:  
Spatially Varying Color & Exposure



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



© The Johns Hopkins University 2020, All Rights Reserved.