MAIA - MedicAl Imaging and Applications

Advanced Image Analysis

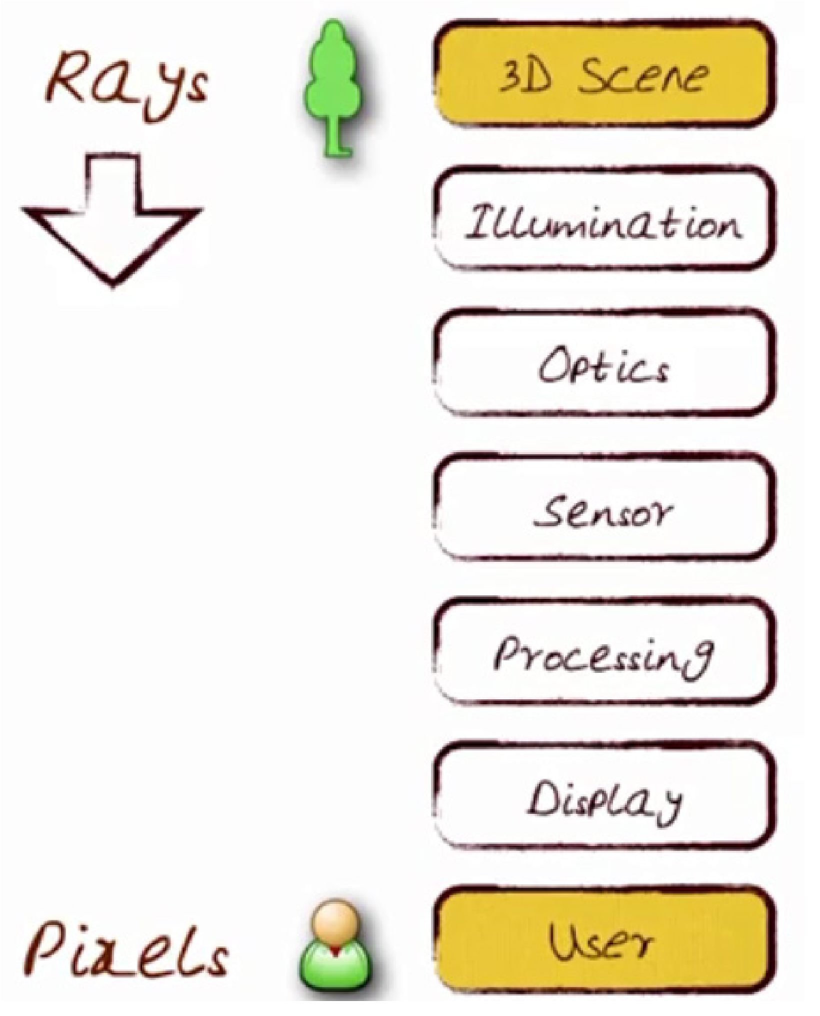
Lectures on Advanced Color Image Processing

B4 - Digital Image Processing: The Camera

March/2018

Adrian Galdran - Post-Doctoral Researcher
Biomedical Imaging Lab – INESC TEC Porto (Portugal)
http://bioimglab.inesctec.pt/

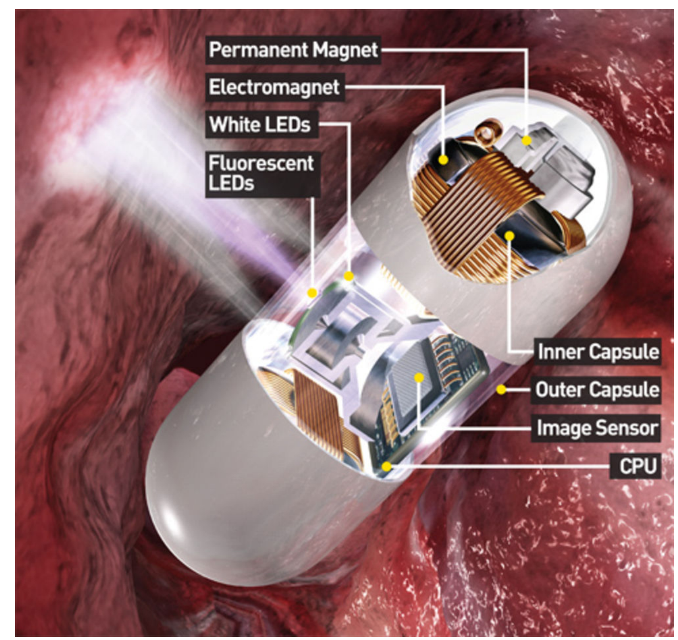
How do we go from rays to pixels?



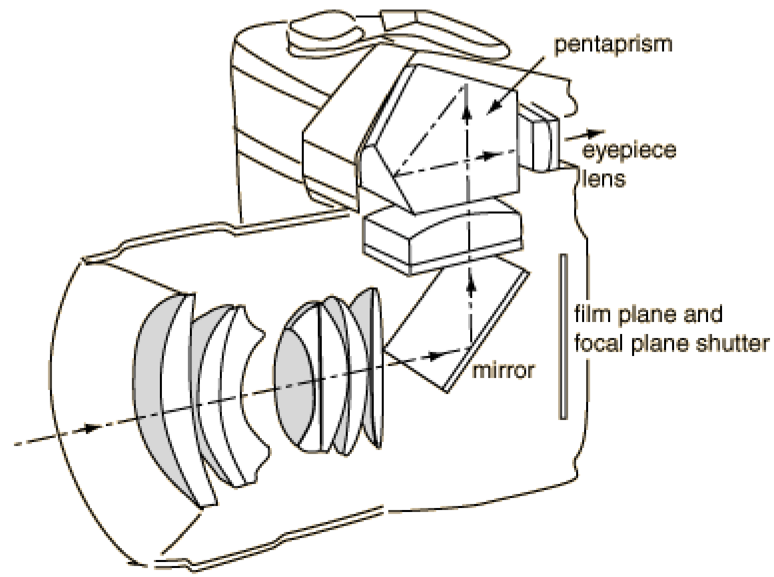
- **B1 - Human Vision System: The Retina and the Brain**
  1. **Basics of Human Perception.**
  2. **Perception and Color. Measuring Color.**
  3. **Color Spaces. Perceptually Uniform Color Spaces.**
  4. **Application: Color Image Segmentation. Superpixels.**
  5. **Introduction to the Retinex model for Human Perception.**
  6. **Introduction to Computational Image Processing.**
  7. **Application: A first example - Image Composition.**

- **B2 - Digital Image Processing: The Camera**
  1. **The Image Acquisition Pipeline.**
  2. **Cameras: Lenses, Focus, Exposure, White Balance, Sensors, Output.**
  3. Unconventional Image Acquisition: Panorama.
  4. Application: Panorama Creation.

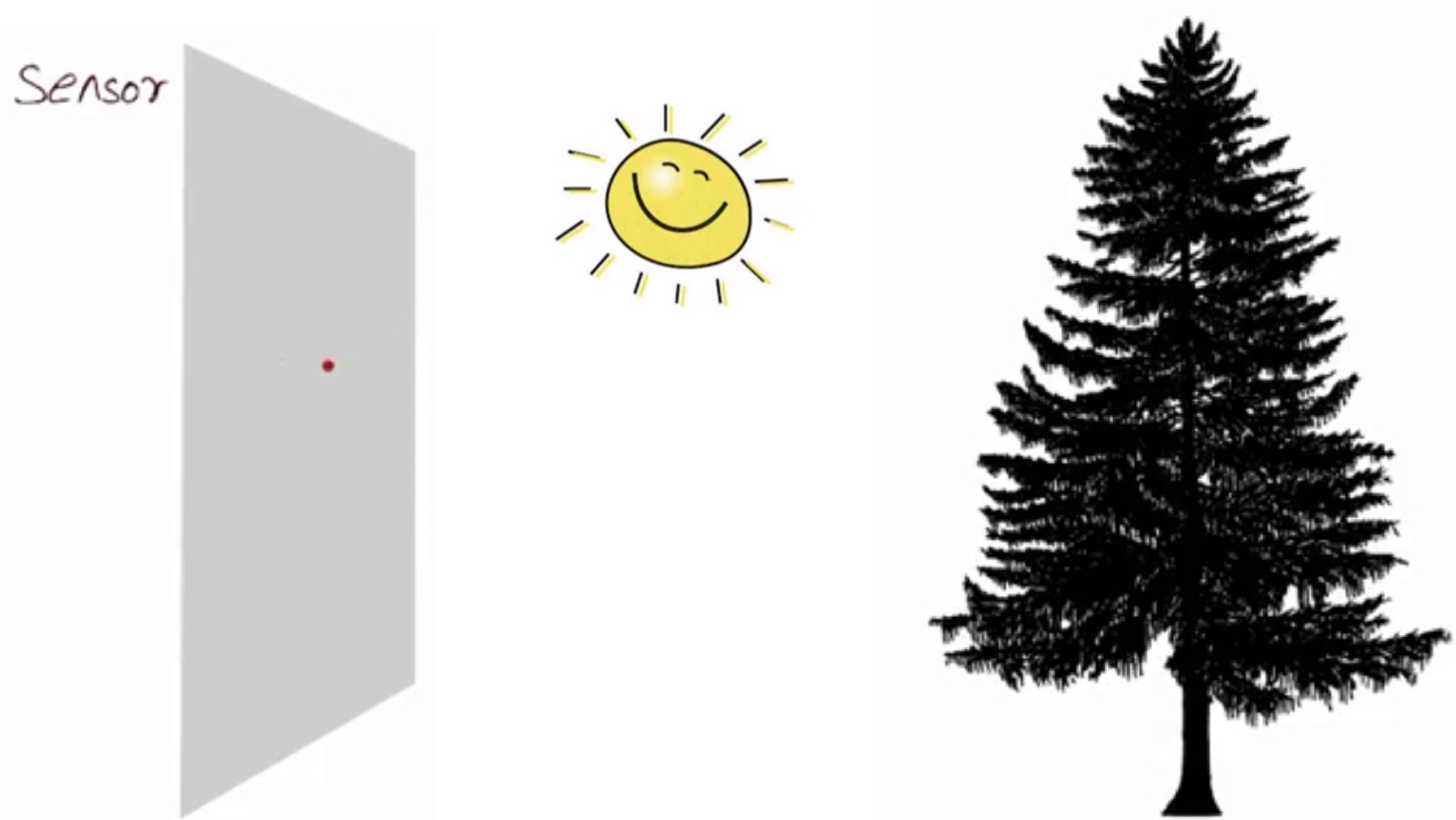
Cameras and their evolution



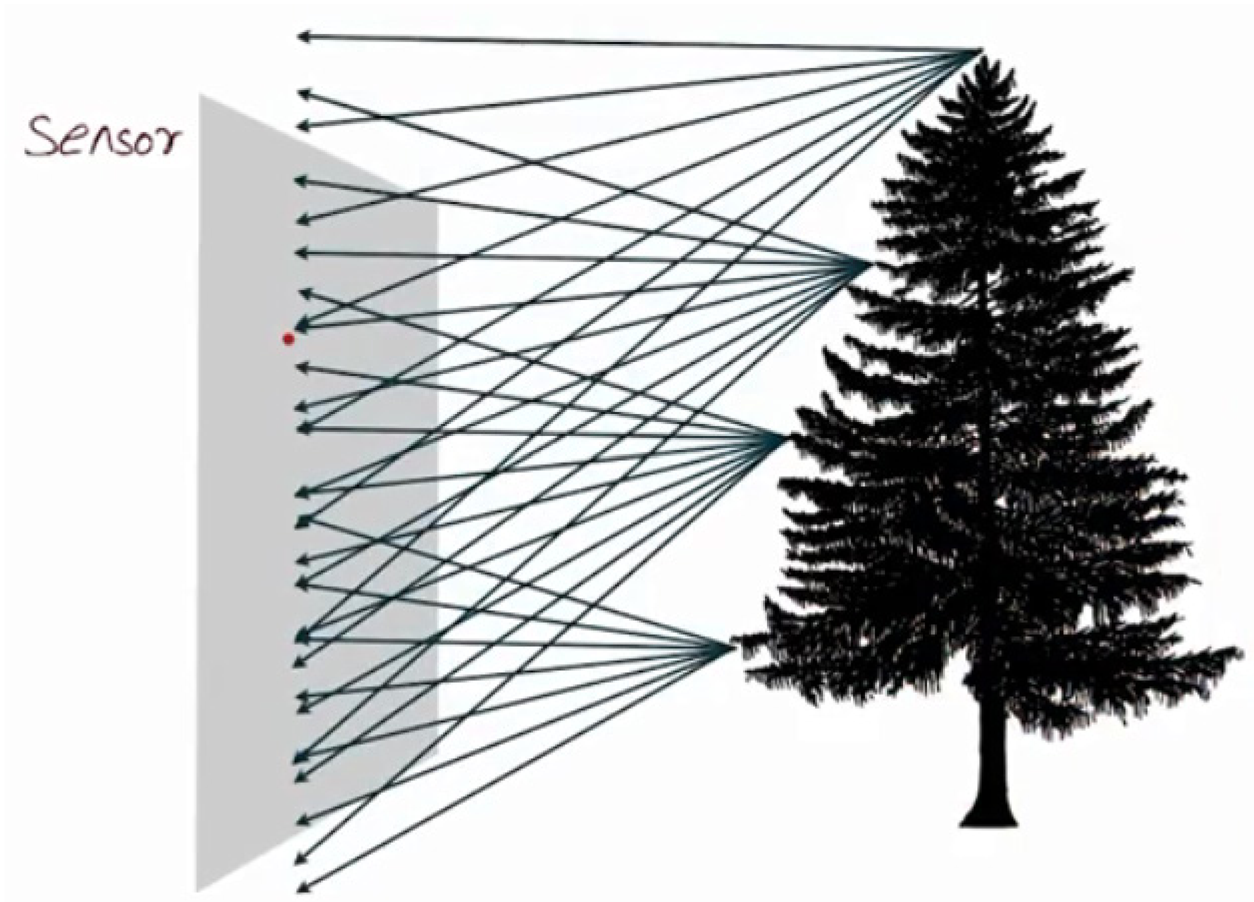
Single Lens Reflex Cameras



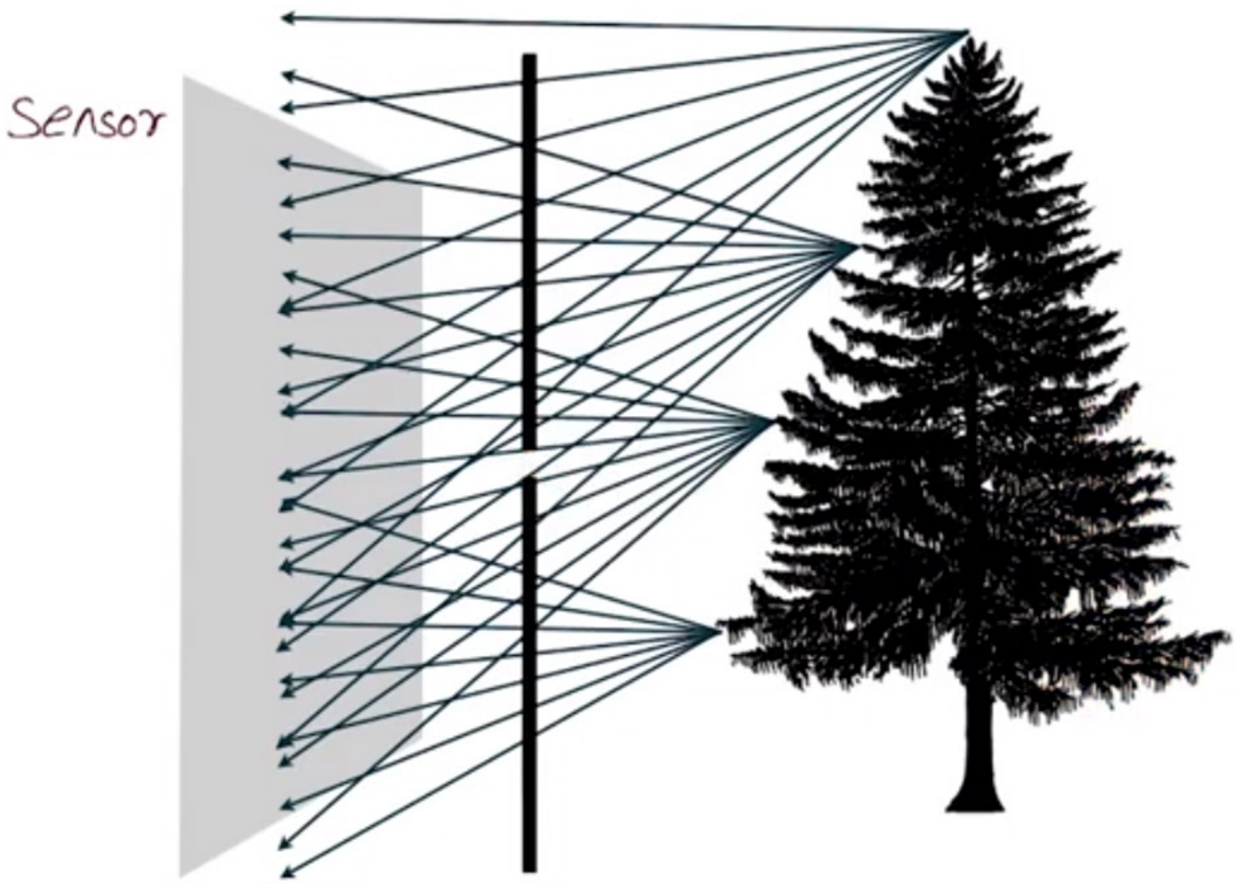
No Optics Model - Pinhole Camera



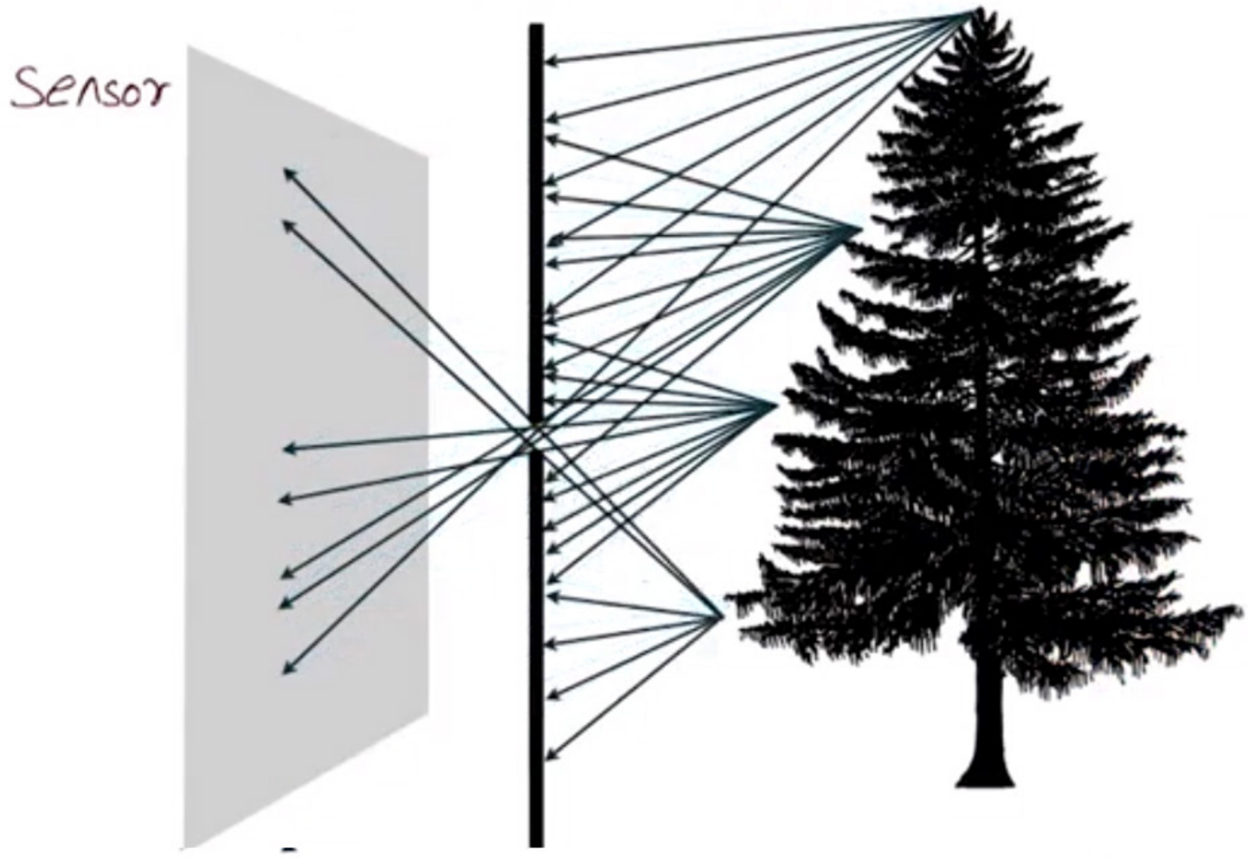
No Optics Model



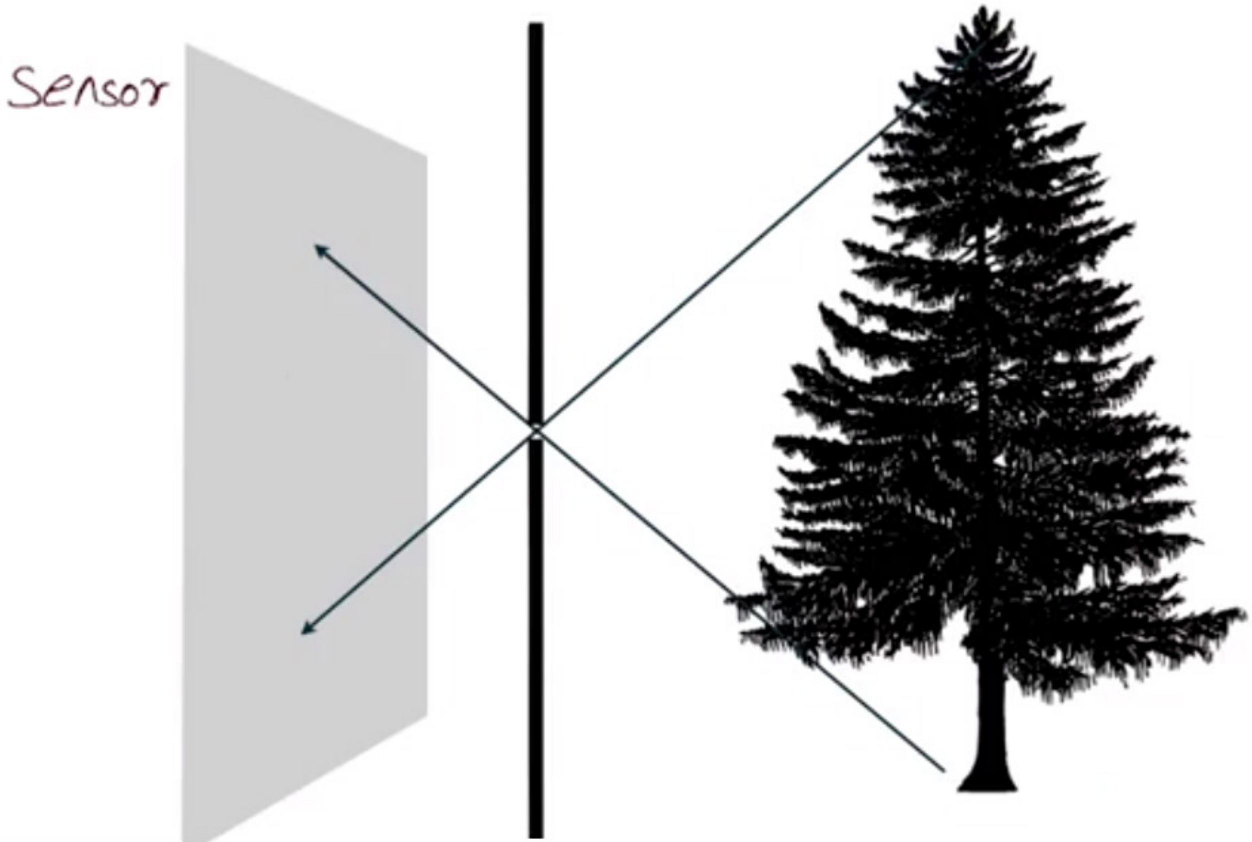
No Optics Model - Pinhole Camera



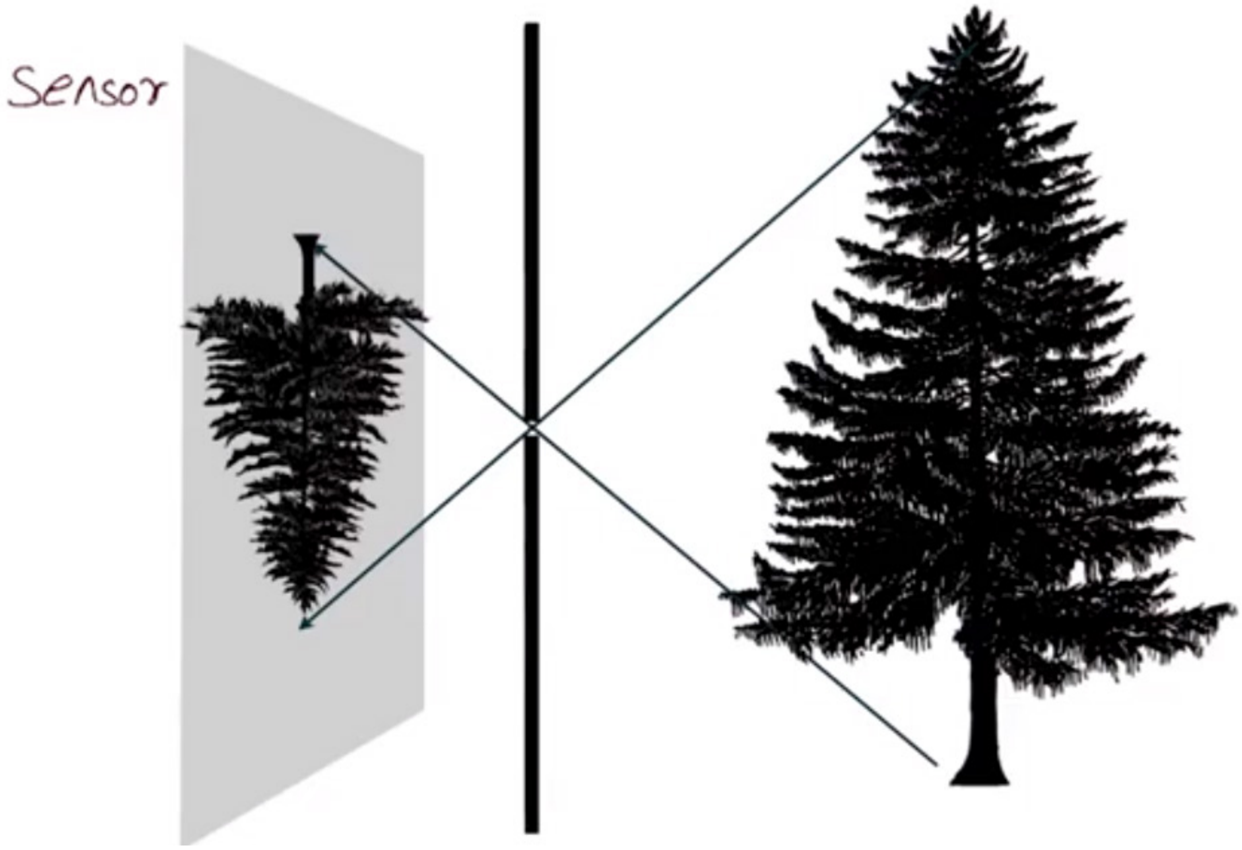
No Optics Model - Pinhole Camera



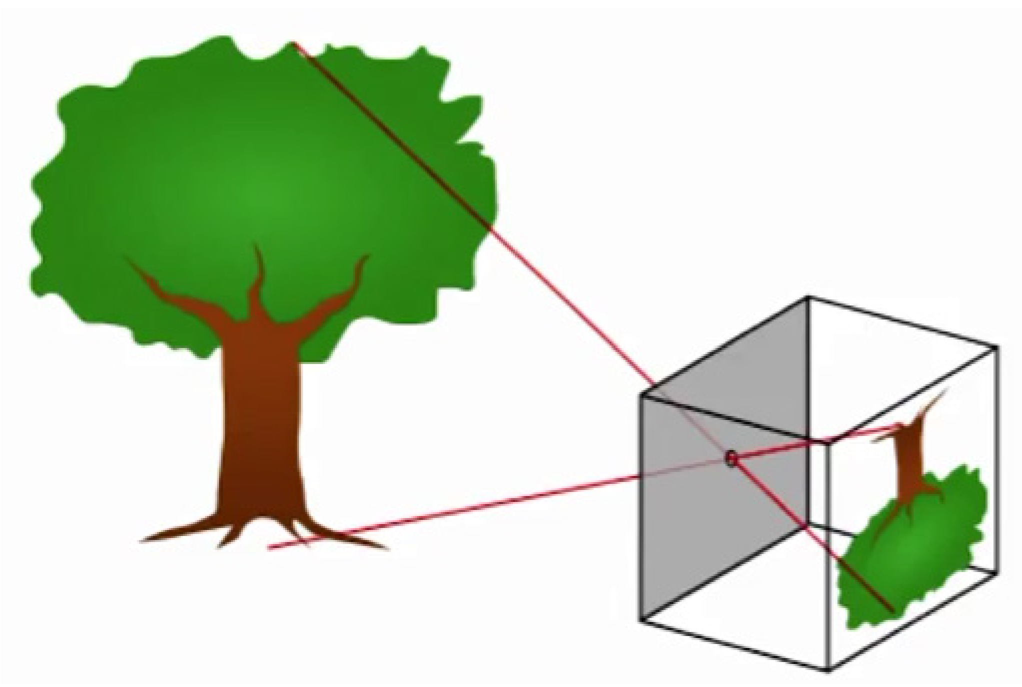
No Optics Model - Pinhole Camera

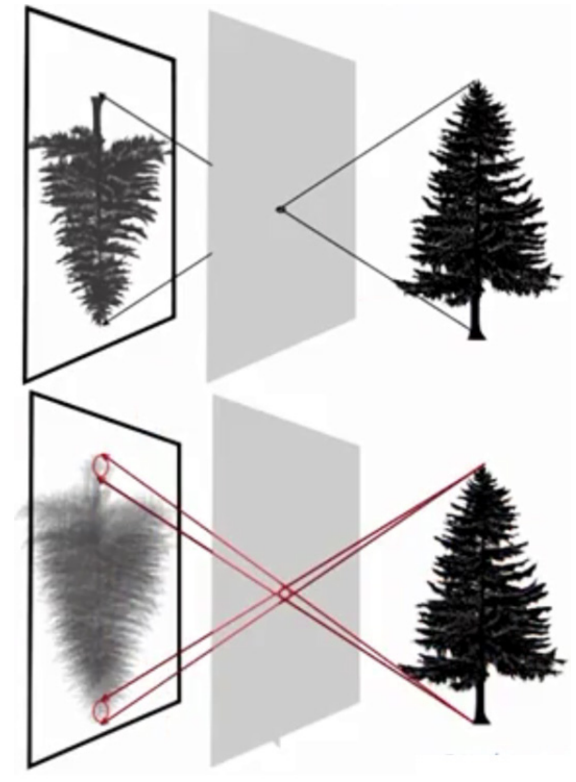


No Optics Model - Pinhole Camera

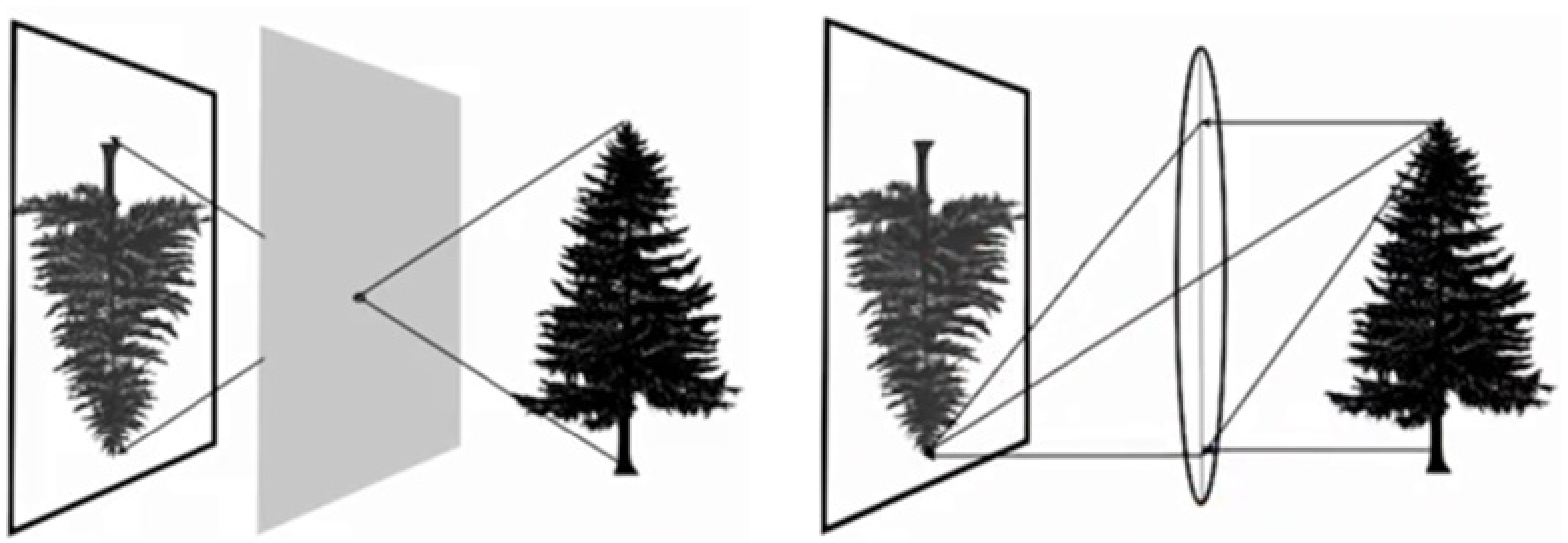


No Optics Model - Pinhole Camera



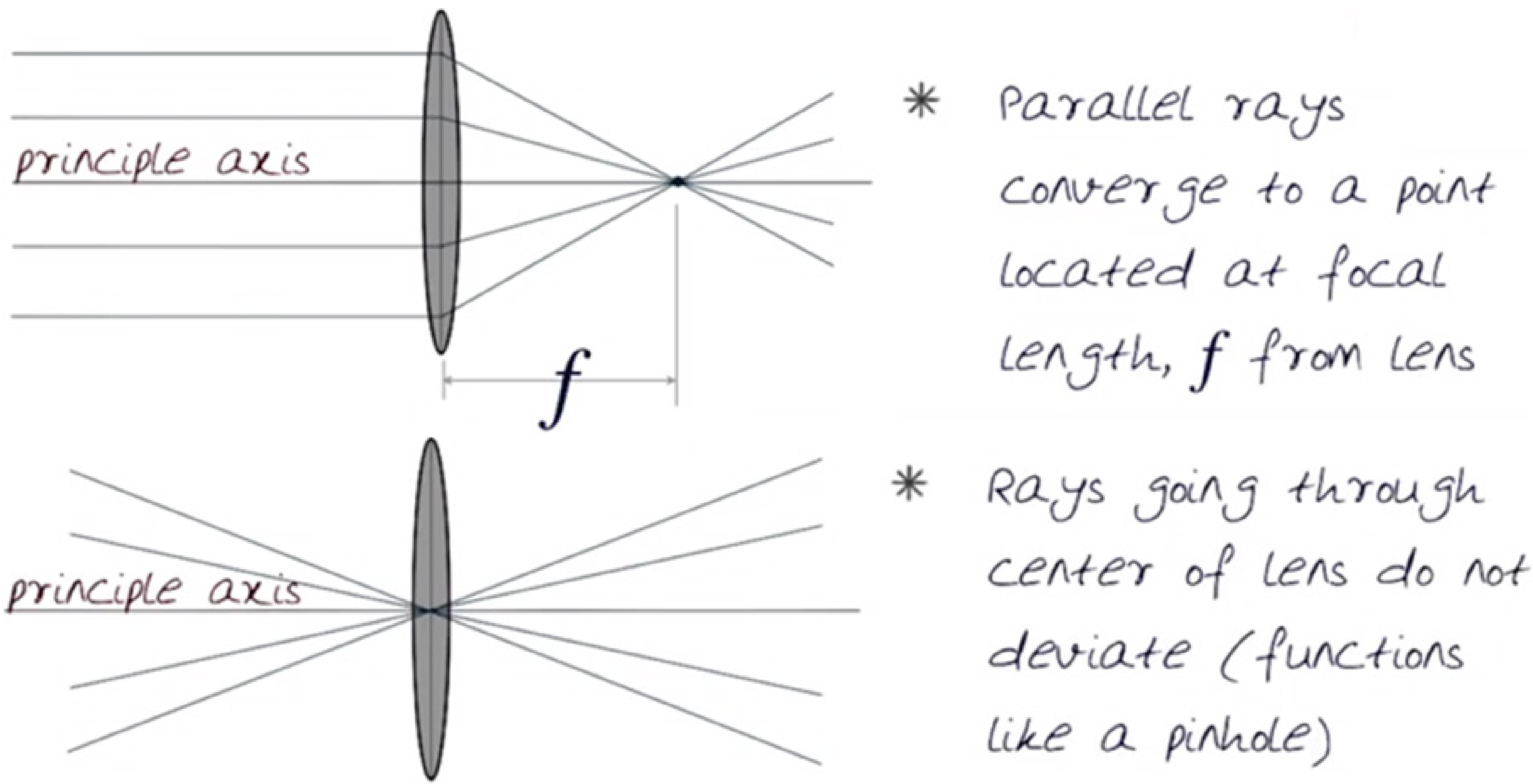


Optics - From Pinhole to Lenses

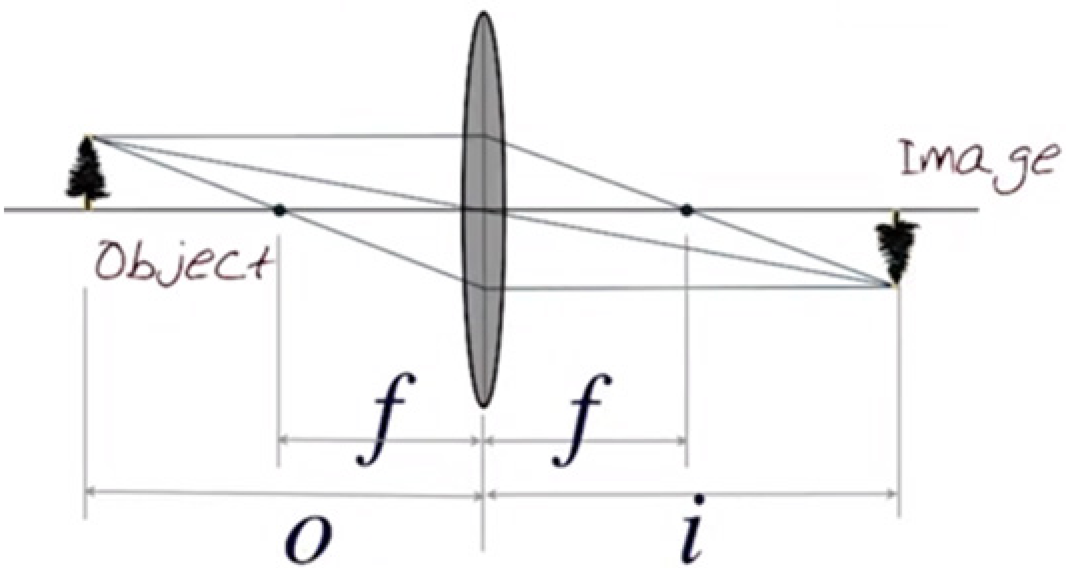


Optics - From Pinhole to Lenses

Lens: specially shaped pieces of glass or other material with refraction index substantially higher than that of air.



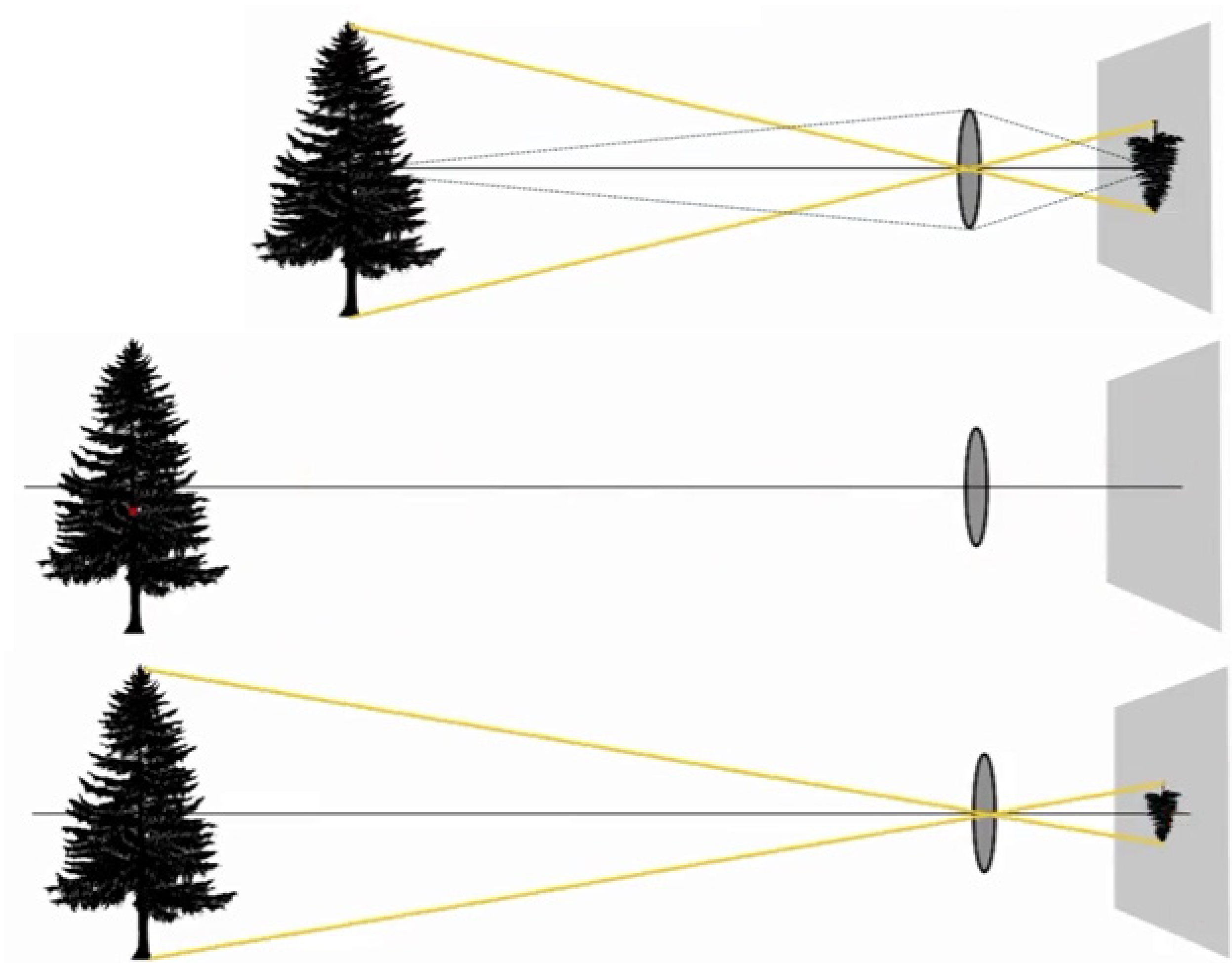
The Lens Equation



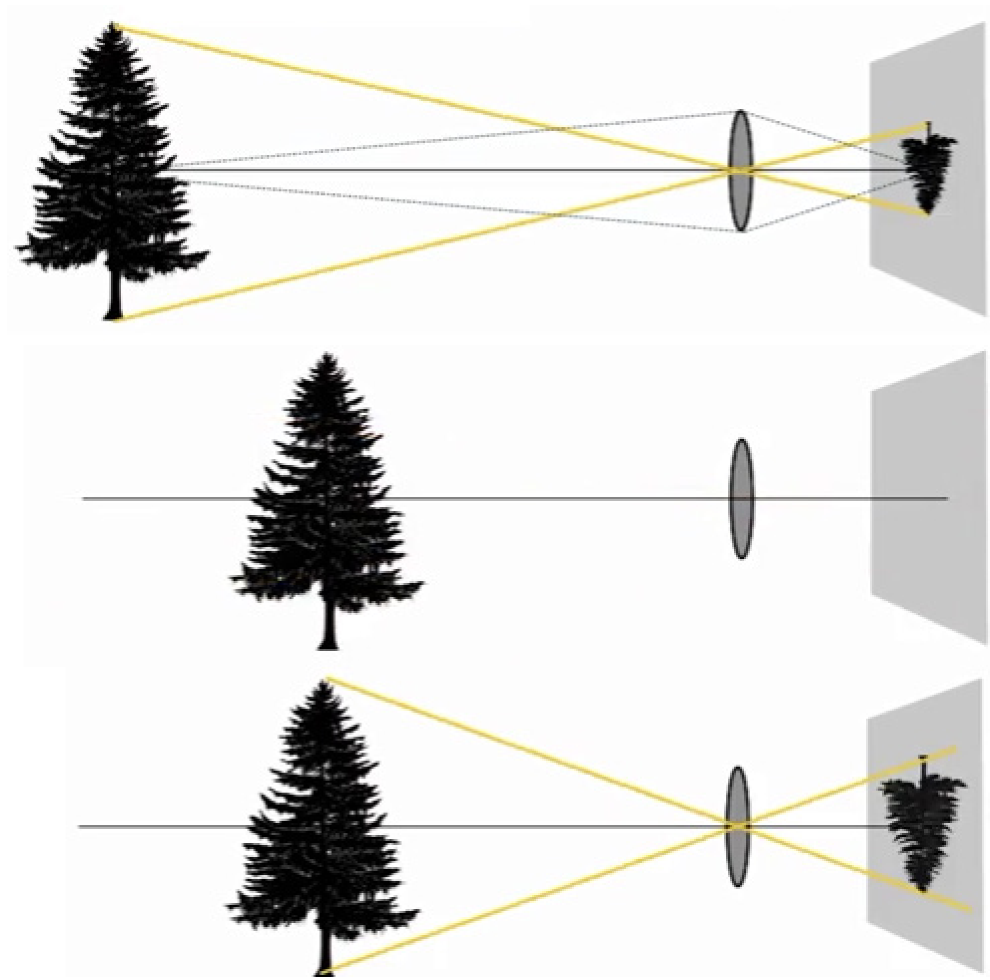
* Rays from points on a plane parallel to the lens, focus on a plane parallel to the lens on the other side (and upside down)

* Lens Equation $$\frac{1}{o} + \frac{1}{i} = \frac{1}{f}$$

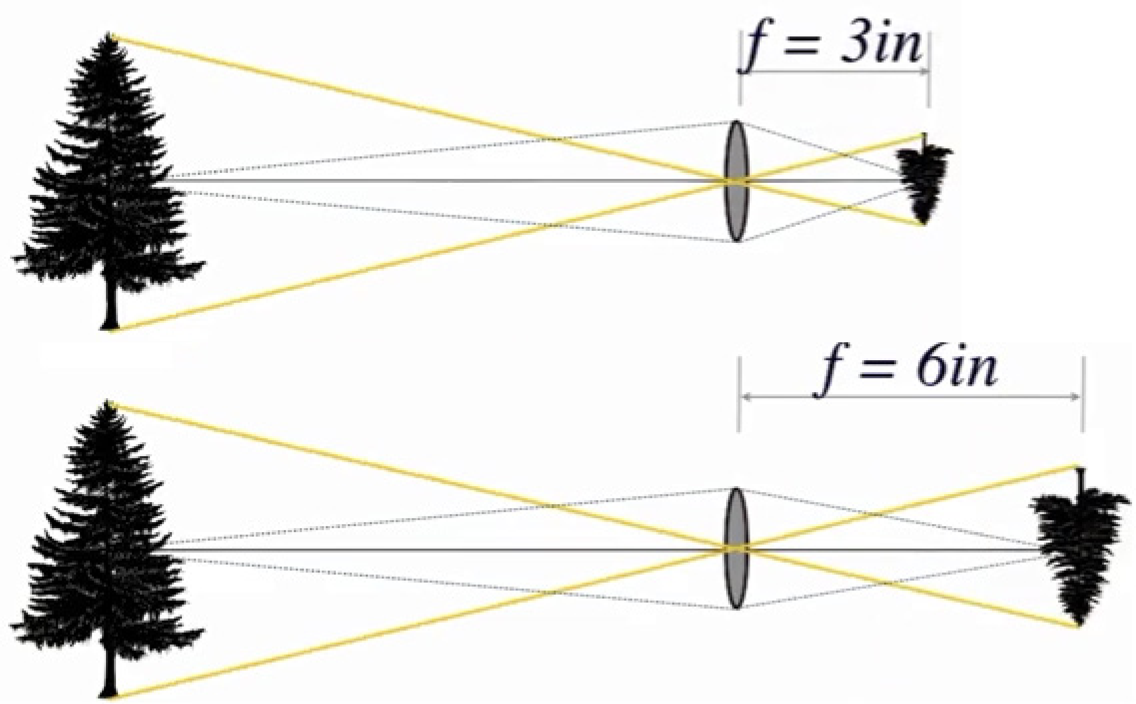
Playing with distances



Playing with distances



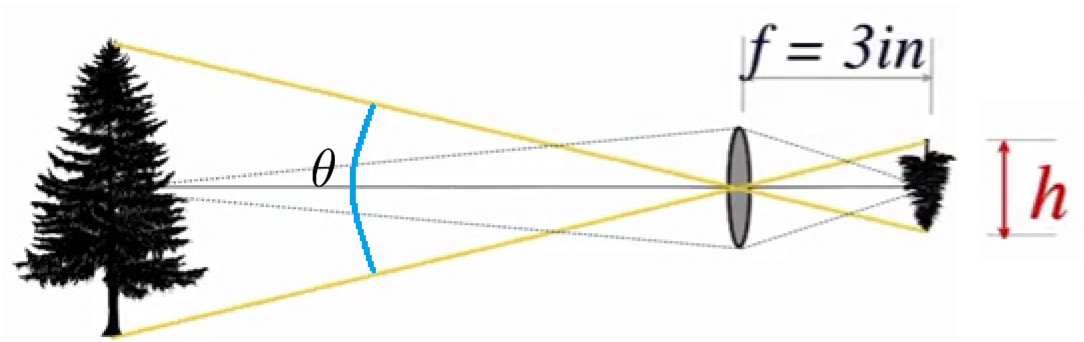
Optical Zoom

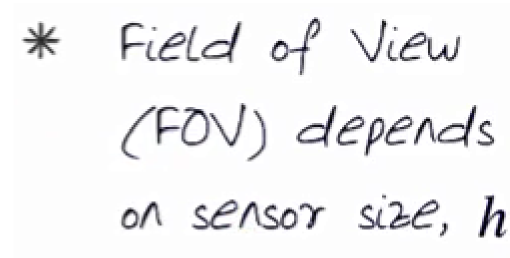


* longer or shorter f

* focal lengths are specific to lenses

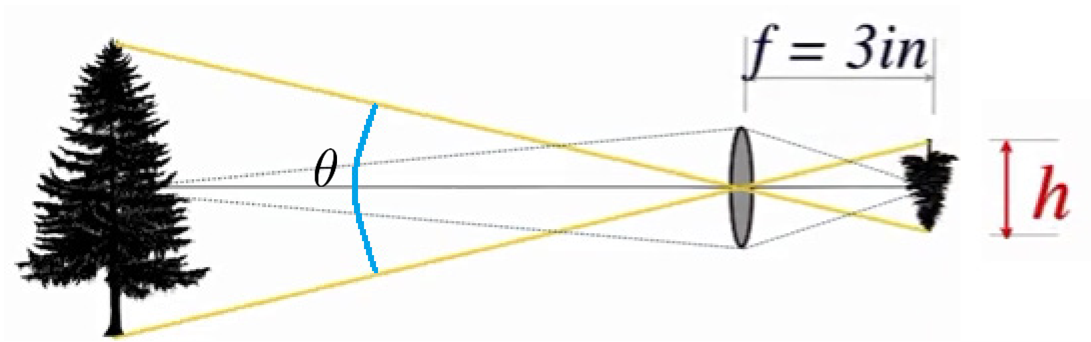
Field of View





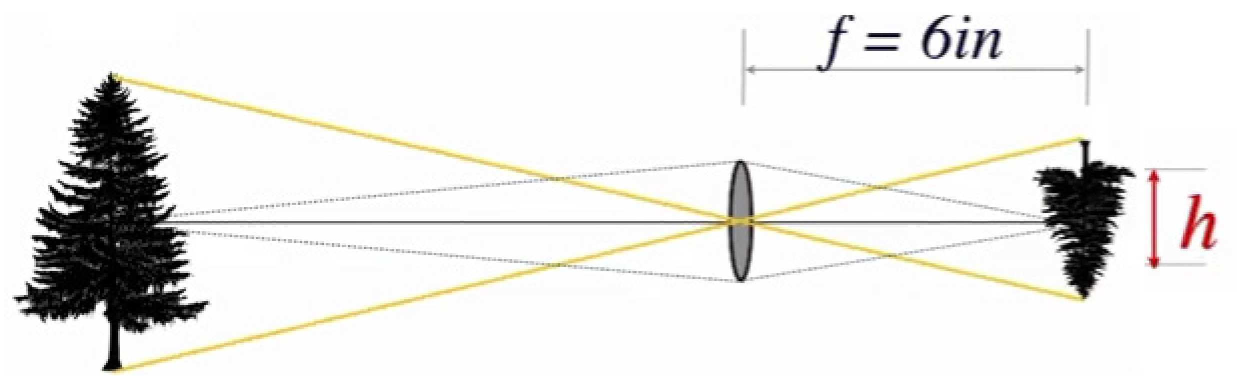
QUIZ!
How can you compute the field of view as a function of sensor size and focal length?
Link: https://goo.gl/forms/RPNXfjZjve4Qd1Gn2

Field of View



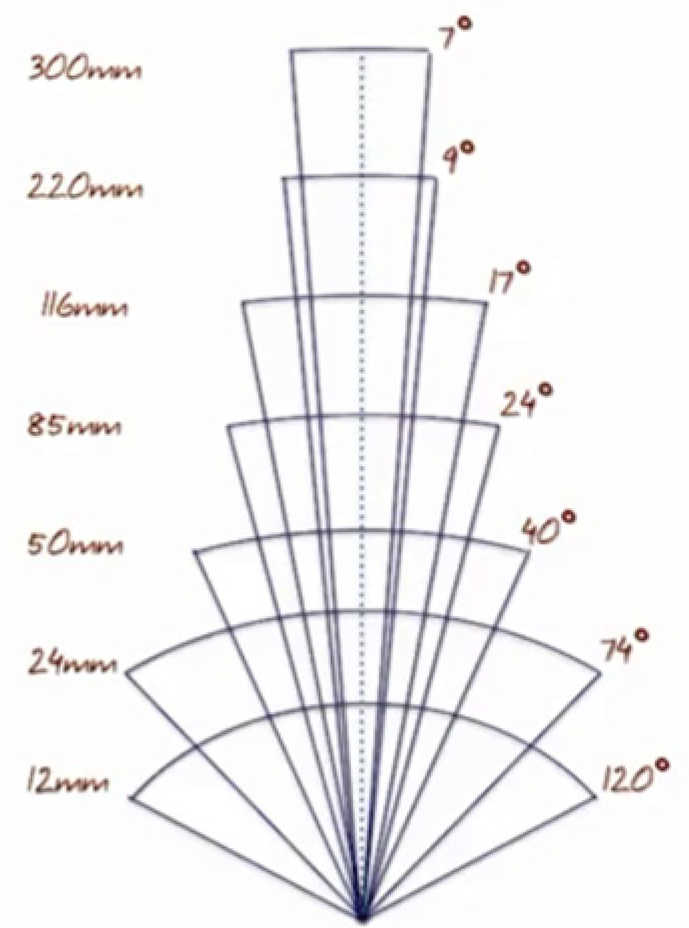
* Field of View (FOV) depends on sensor size, h

QUIZ!
How can you compute the field of view as a function of sensor size and focal length?

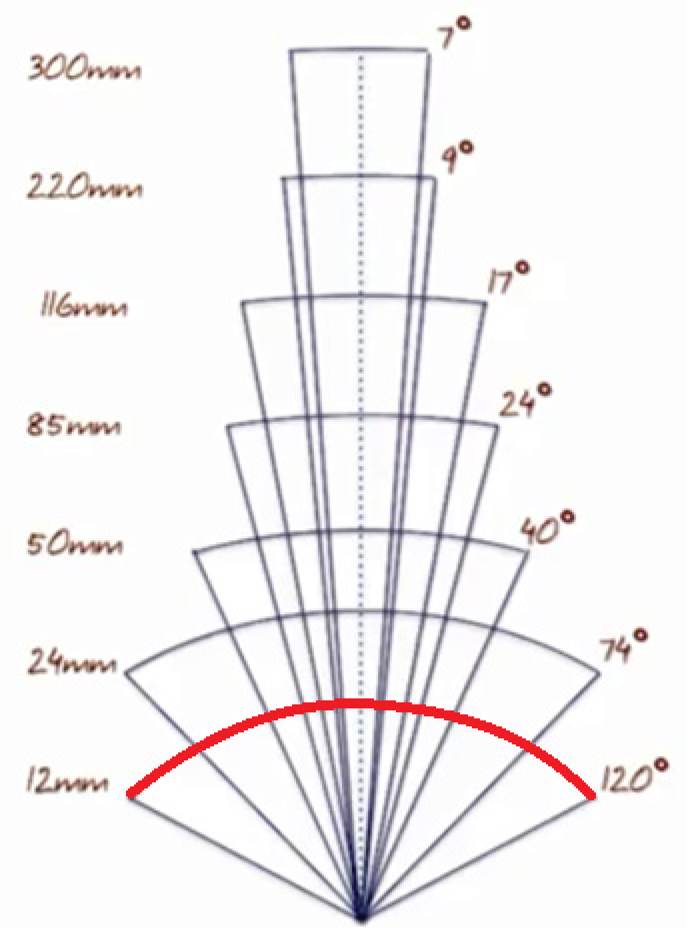


* Sensor size small, FOV is small too

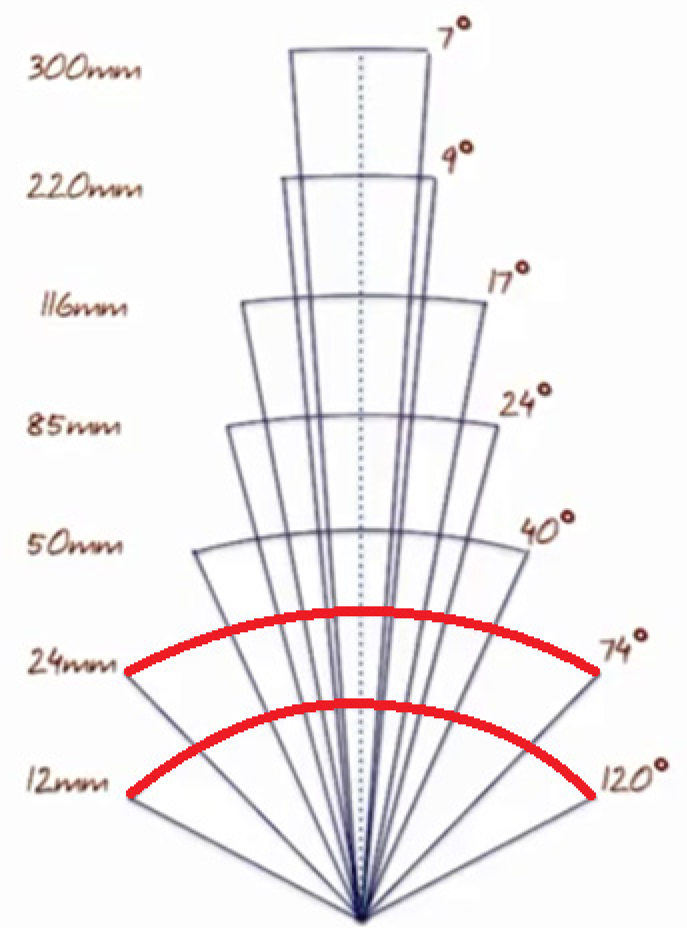
Optical Zoom 2



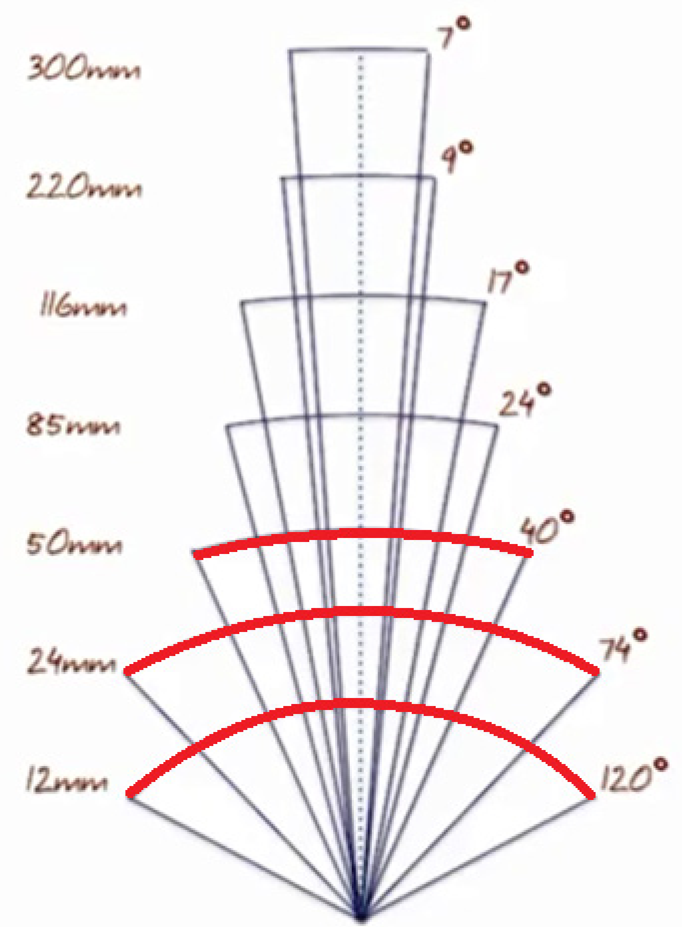
Optical Zoom 2



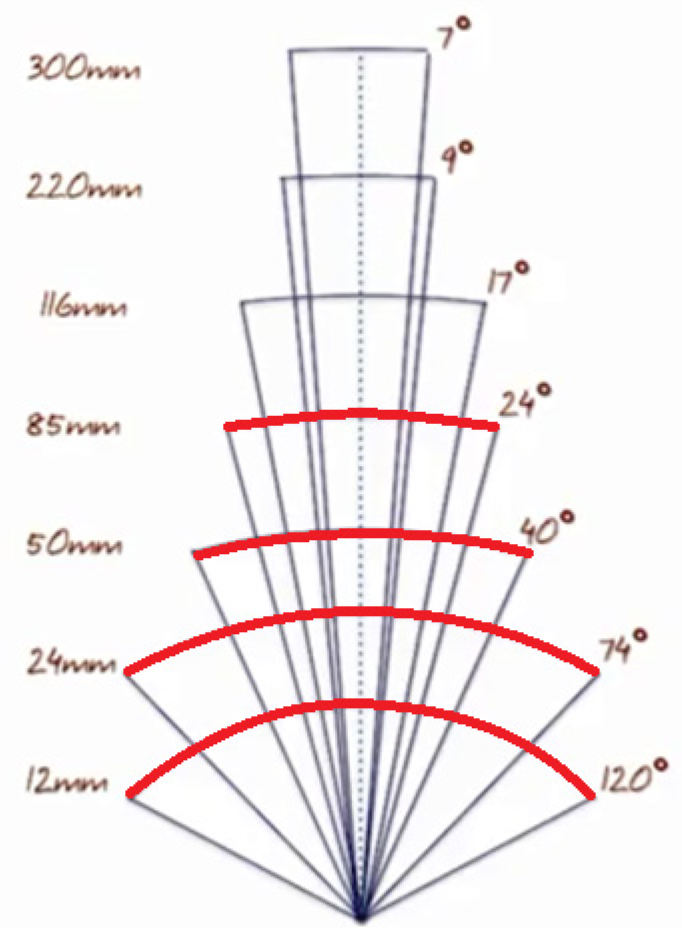
Optical Zoom 2



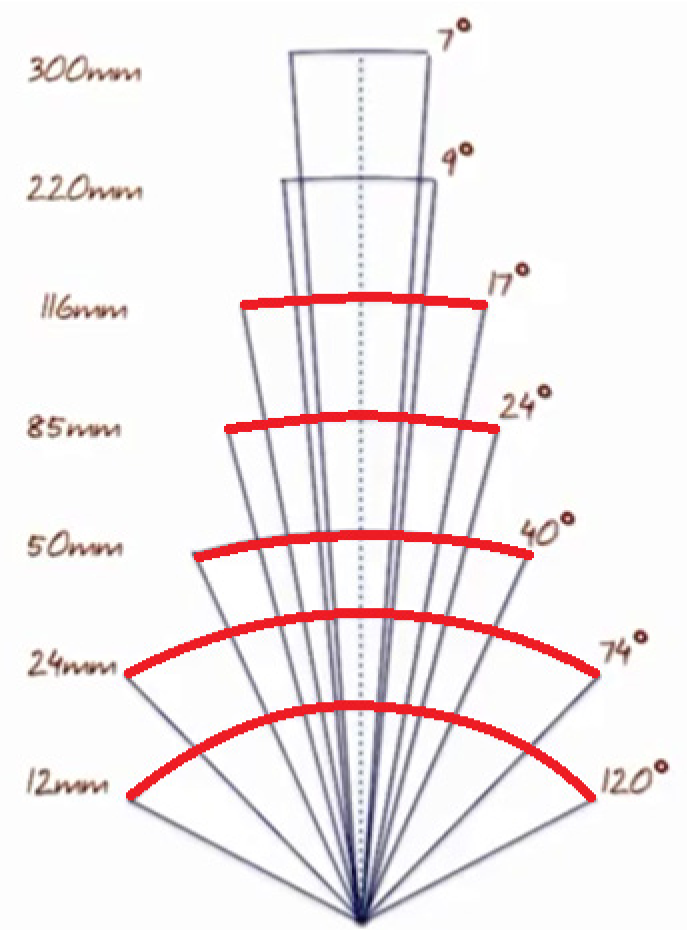
Optical Zoom 2



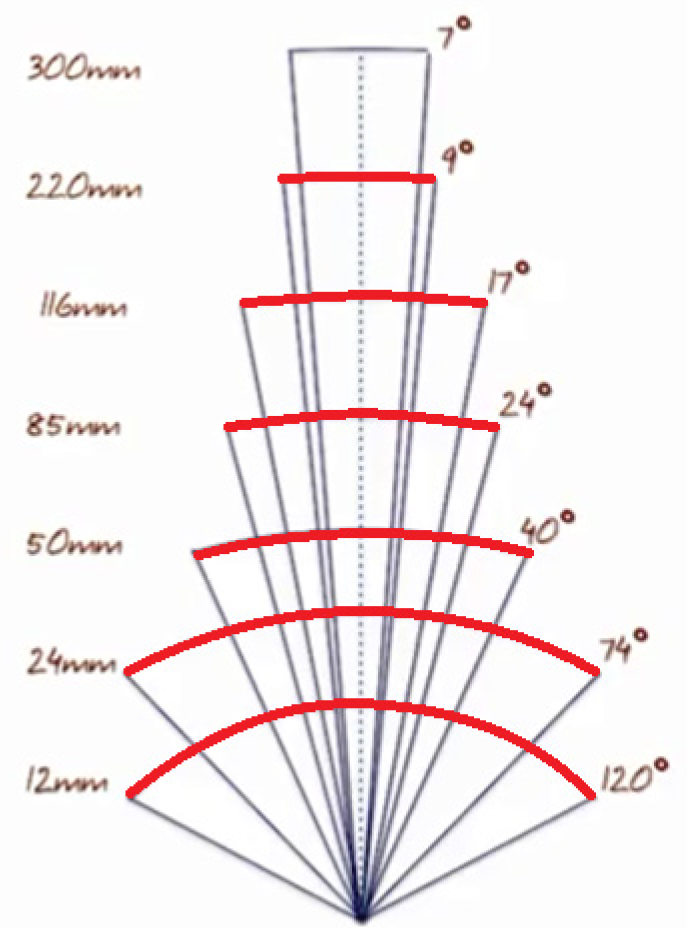
Optical Zoom 2



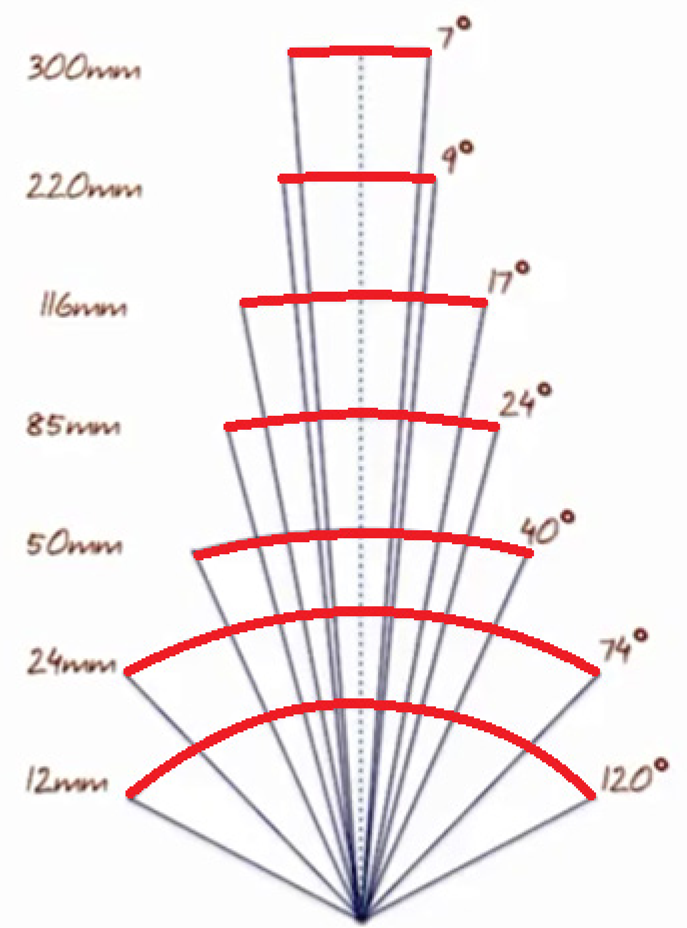
Optical Zoom 2



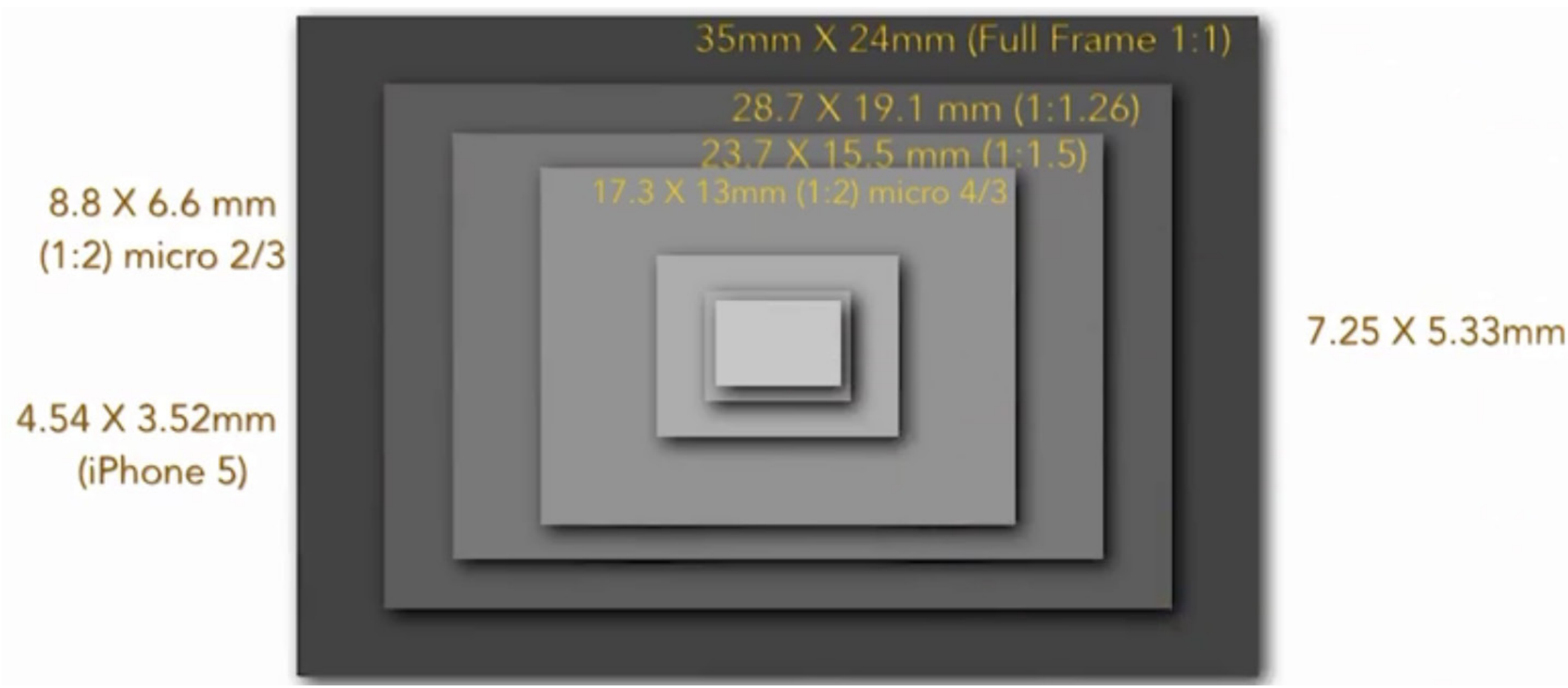
Optical Zoom 2



Optical Zoom 2



Common Sensor Sizes



Focal Length and Perspective

f = 18mm, 35mm sensor

Distance to 1st Subject = 0.5m

Distance to 2nd Subject = 2.0m

f = 180mm, 35mm sensor

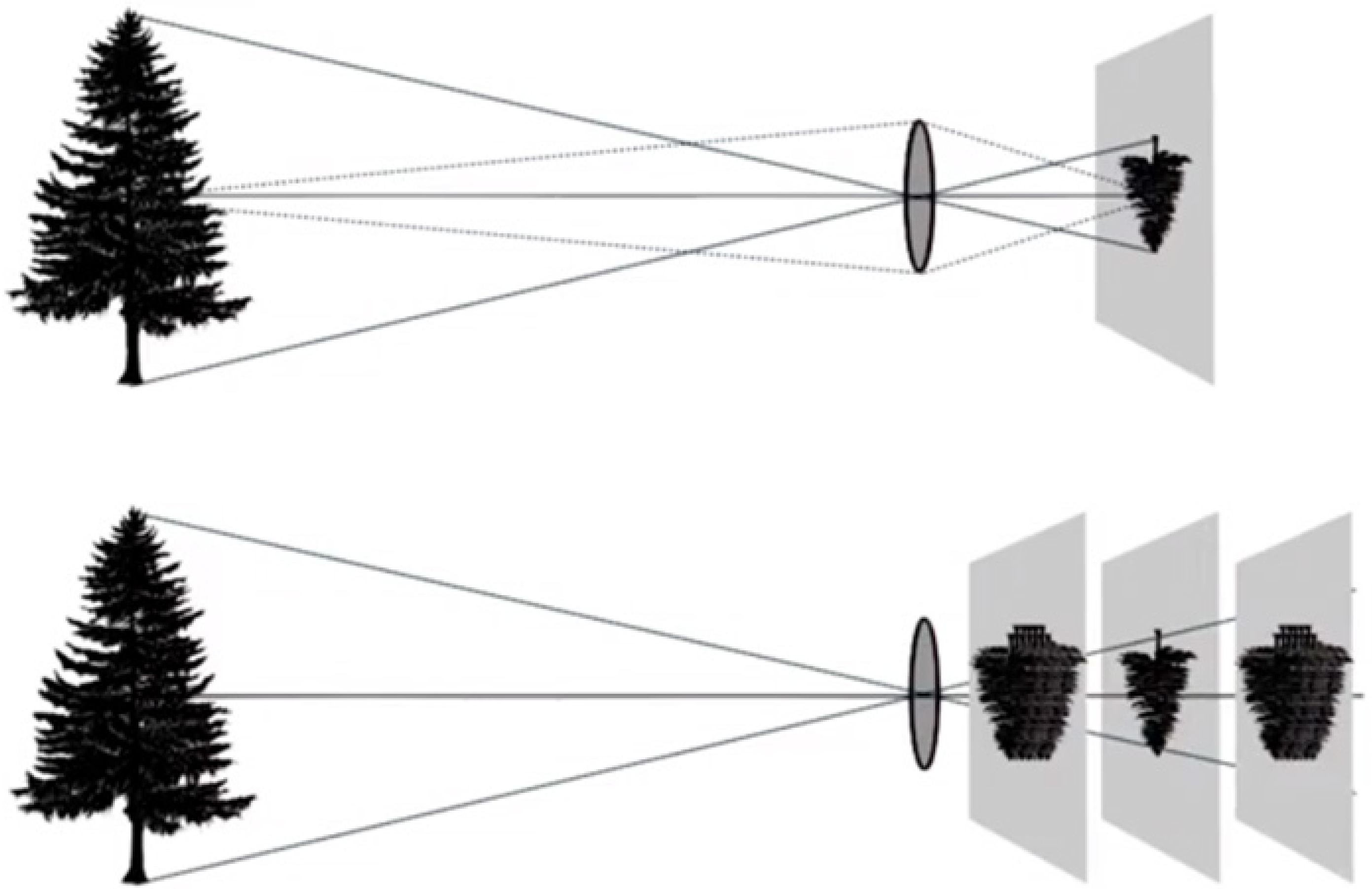
Distance to 1st Subject = 3.0m

Distance to 2nd Subject = 4.5m

* Changing focal length allows us to move back, and still capture the scene
* Changing viewpoint causes perspective changes
* See the "vertigo" effect from Hitchcock movies

https://www.wired.com/2012/12/how-to-make-a-hobbit-with-forced-perspective/

How are images focused?



Exposure

Exposure = Irradiance X Time

$$H = E \times T$$

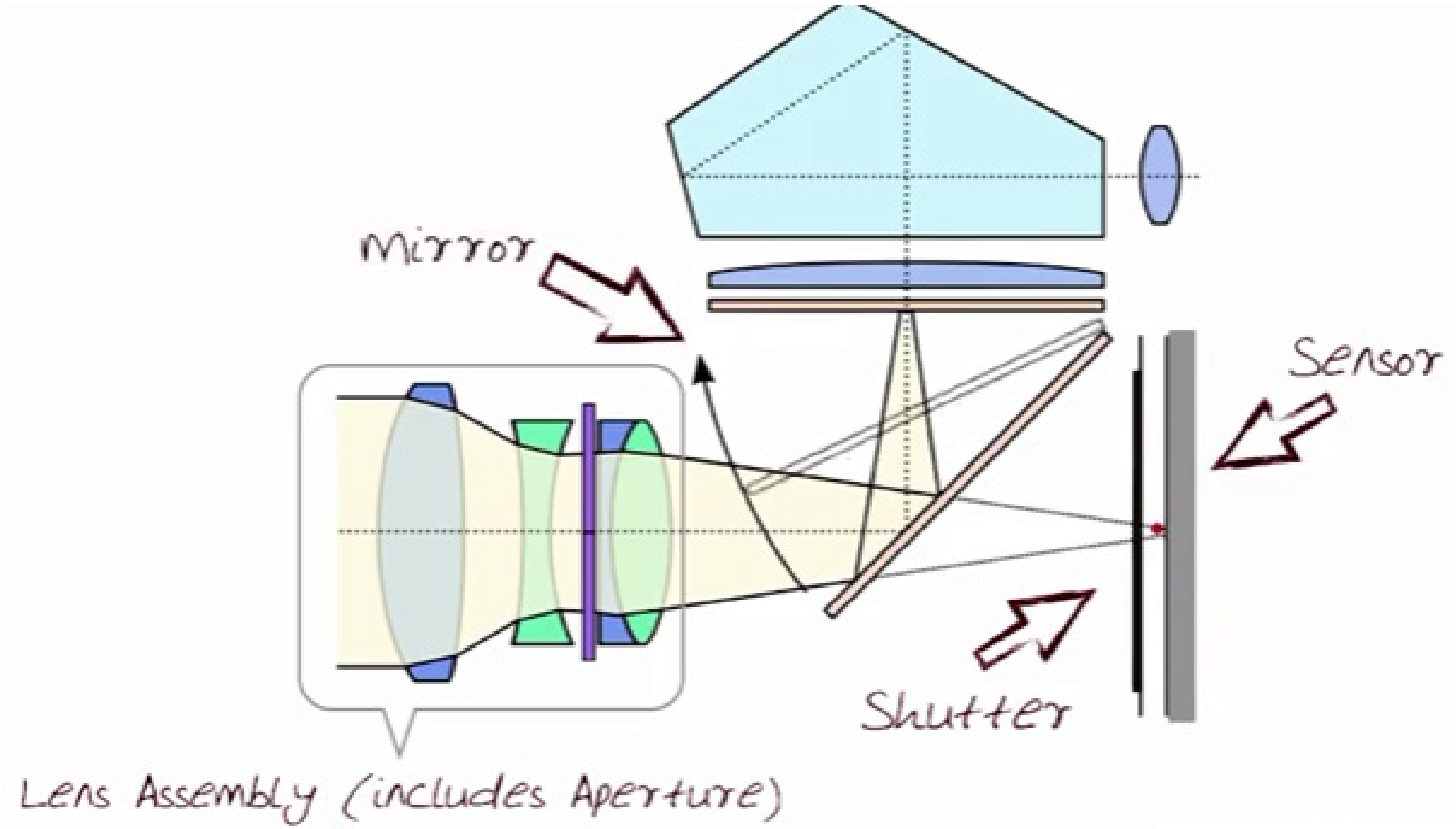
Irradiance (E)

* Amount of light falling on a unit area of sensor per second
* Controlled by lens **aperture**

Exposure Time (T)

* How long the **shutter** is kept open

Exposure



https://www.youtube.com/watch?v=ptfSW4eW25g

Shutter Speed

* Amount of time the sensor is exposed to light
* Usually denoted in fractions of a second (1/2000, 1/1000, ..., 1/250, ..., 1/60, ..., 1/15, ..., 15, 30, Bulb)

Shutter Speed

Shutter Speed

Shutter Speed

Shutter Speed

Shutter Speed

Shutter Speed

https://www.youtube.com/watch?v=yr3ngmRuGUc

Aperture

Irradiance on Sensor → The amount of light captured is proportional to the Area of the Aperture (opening):

$$Area = \pi \left(\frac{f}{2N}\right)^2$$

* **f** is the the focal length. What is the diameter of the Aperture?
* Aperture Number **N** usually written as **f/N** [F-number]
* Aperture Number gives irradiance irrespective of the lens in use

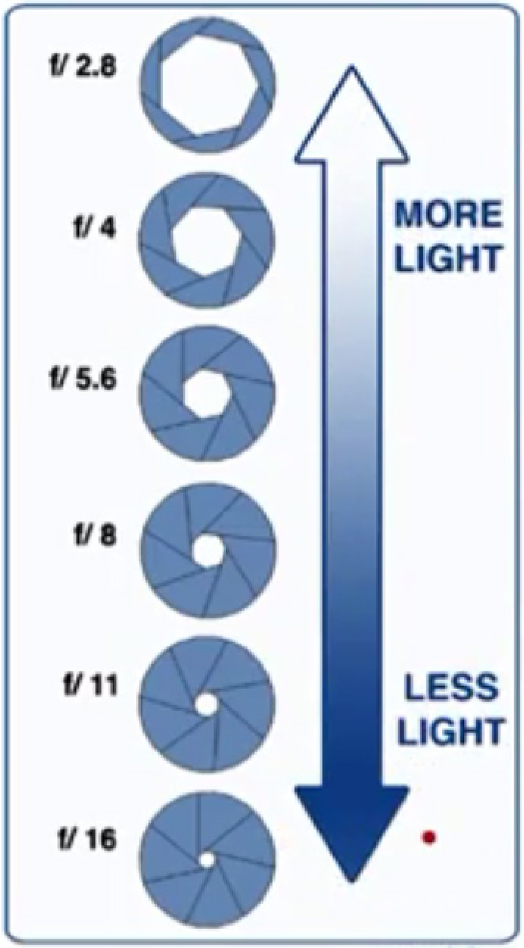
Aperture

Aperture

$$Area = \pi \left(\frac{f}{2N} \right)^2$$

* Doubling N reduces A by 2X, and therefore reduces light by 4X
  * from f/2.8 to f/5.6 cuts light by 4X
  * to cut light by 2X, increase N by $\sqrt{2}$

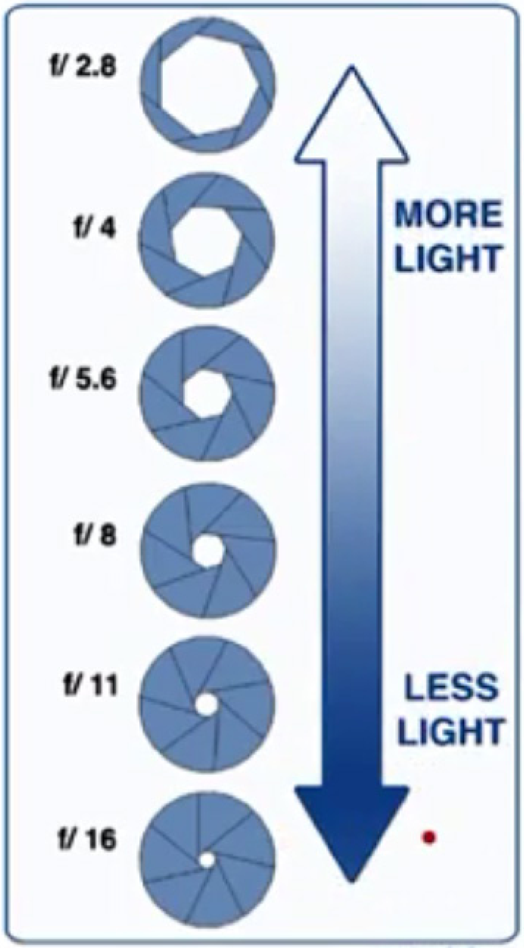
Aperture




MULTI COATING PENTACON
1.8/50 auto

Aperture


MULTI COATING PENTACON
1.8/50 auto

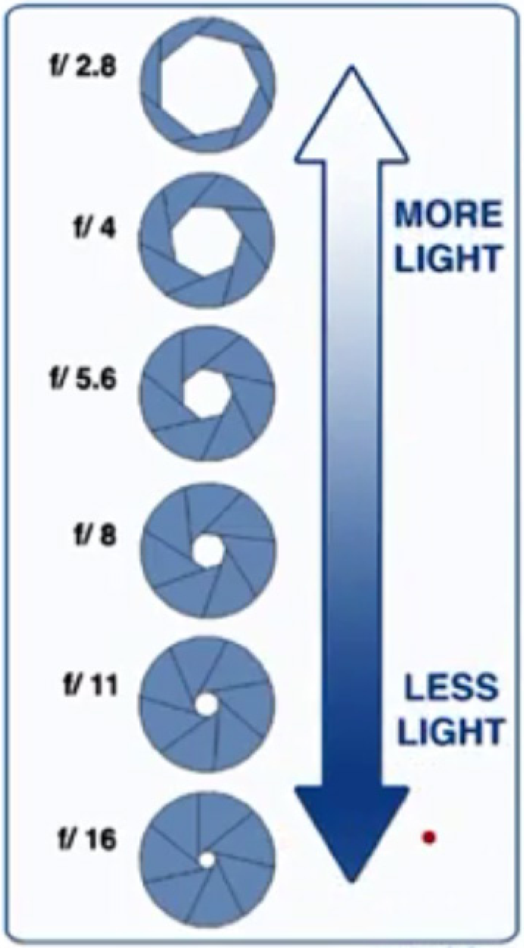

f/ 2.8
f/ 4
f/ 5.6
f/ 8
f/ 11
f/ 16
MORE LIGHT
LESS LIGHT


f/4.0
f/3.5


MULTI COATING
PENTACON
1.8/50 auto

Aperture


MULTI COATING
PENTACON
1.8/50 auto

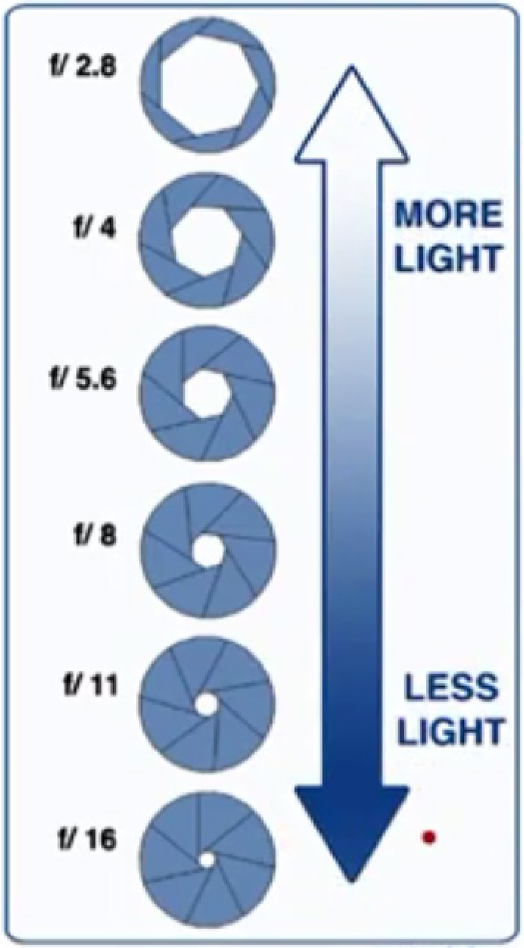

f/ 2.8
f/ 4
f/ 5.6
f/ 8
f/ 11
f/ 16
MORE
LIGHT
LESS
LIGHT


f/5.6
f/4.0
f/3.5


MULTI COATING
PENTACON
auto 1.8/50

Aperture


MULTI COATING
PENTACON
auto 1.8/50

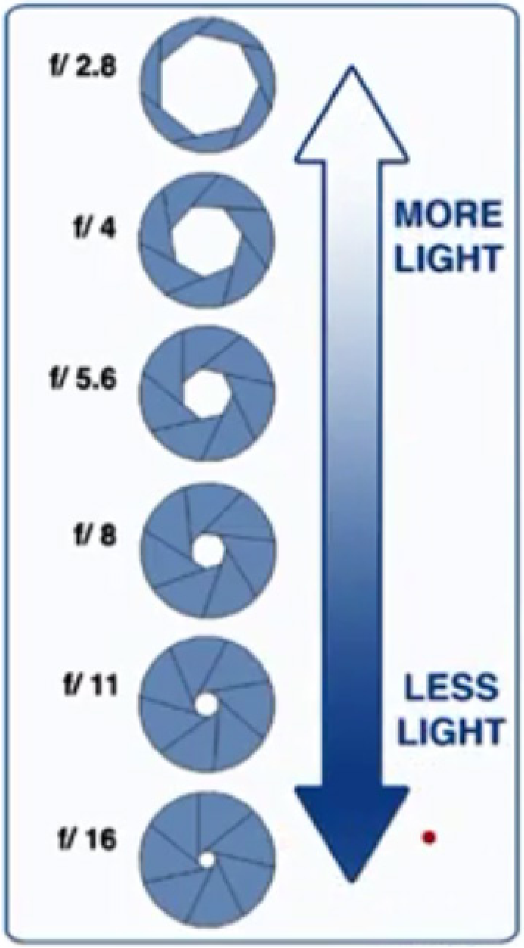

f/ 2.8
f/ 4
f/ 5.6
f/ 8
f/ 11
f/ 16
MORE LIGHT
LESS LIGHT


f/8
f/5.6
f/4.0
f/3.5


MULTI COATING
PENTACON
1.8/50 auto

Aperture


MULTI COATING
PENTACON
1.8/50 auto

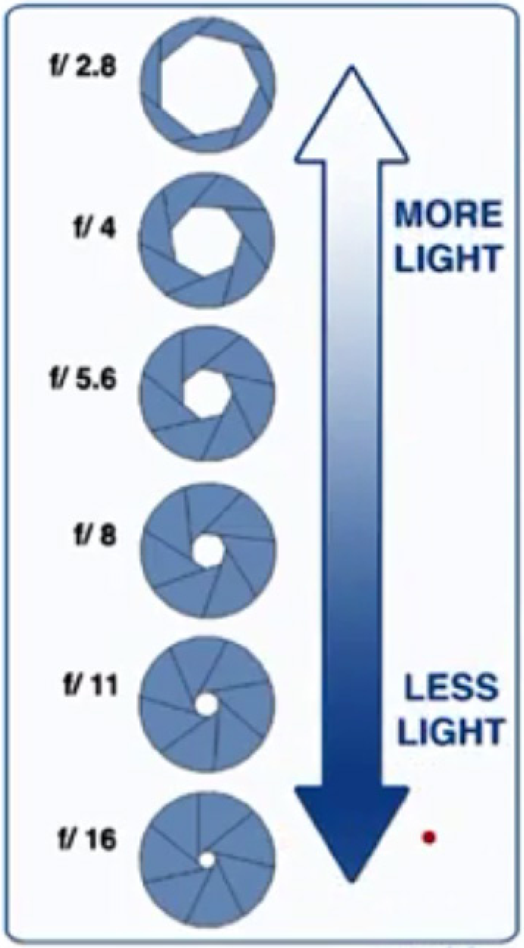

f/ 2.8
f/ 4
f/ 5.6
f/ 8
f/ 11
f/ 16
MORE LIGHT
LESS LIGHT


f/11
f/8
f/5.6
f/4.0
f/3.5


MULTI COATING
PENTACON
1.8/50
auto

Aperture


MULTI COATING
PENTACON
1.8/50
auto

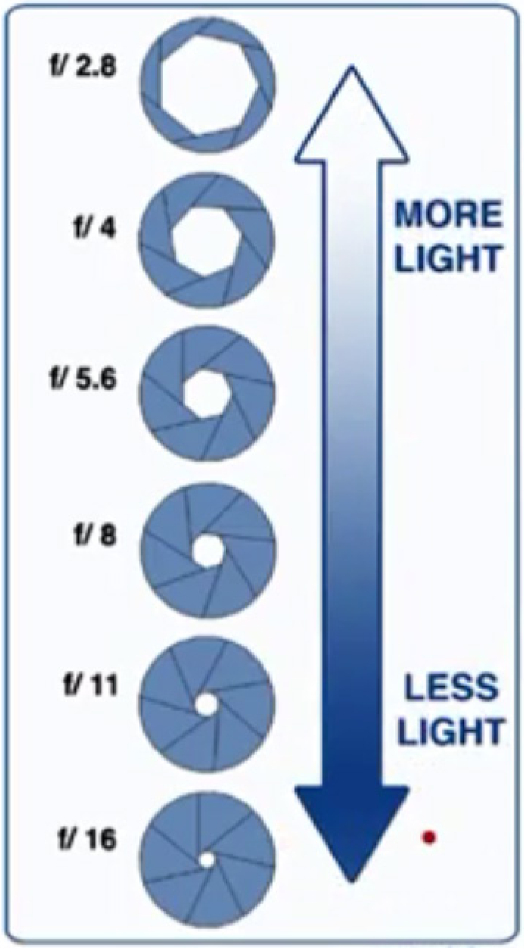

f/ 2.8
f/ 4
f/ 5.6
f/ 8
f/ 11
f/ 16
MORE LIGHT
LESS LIGHT


f/14
f/11
f/8
f/5.6
f/4.0
f/3.5


MULTI COATING
PENTACON
auto 1.8/50

Aperture

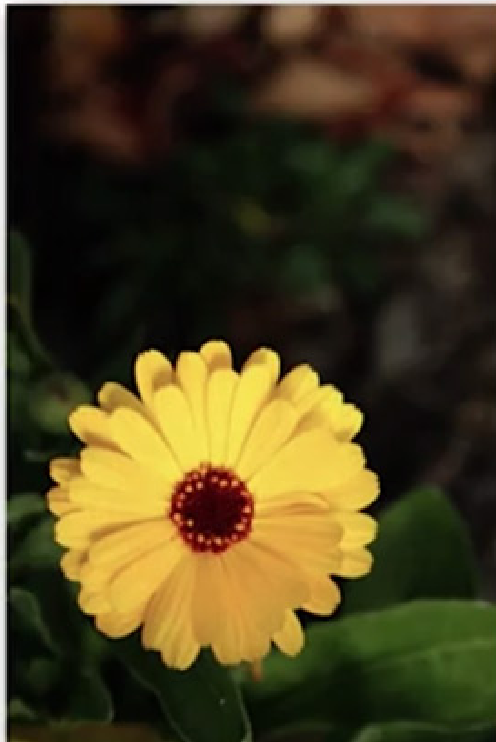

MULTI COATING
PENTACON
auto 1.8/50


f/ 2.8
f/ 4
f/ 5.6
f/ 8
f/ 11
f/ 16
MORE LIGHT
LESS LIGHT


f/18
f/14
f/11
f/8
f/5.6
f/4.0
f/3.5

Sensor Gain - ISO

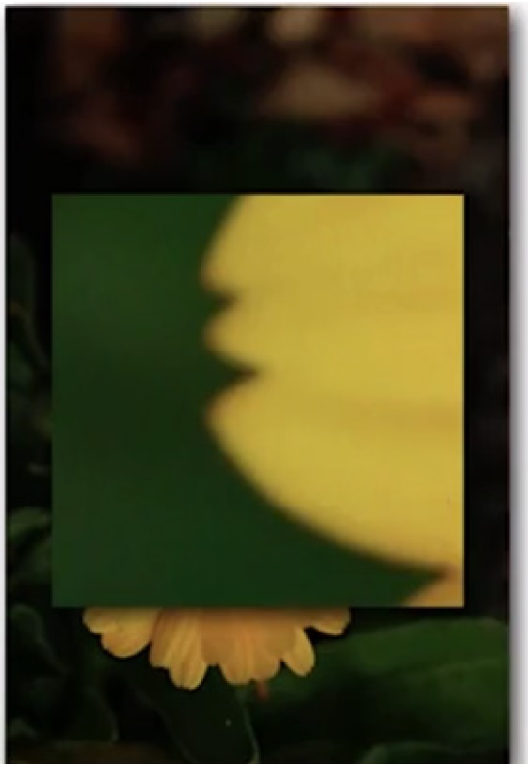
ISO 100



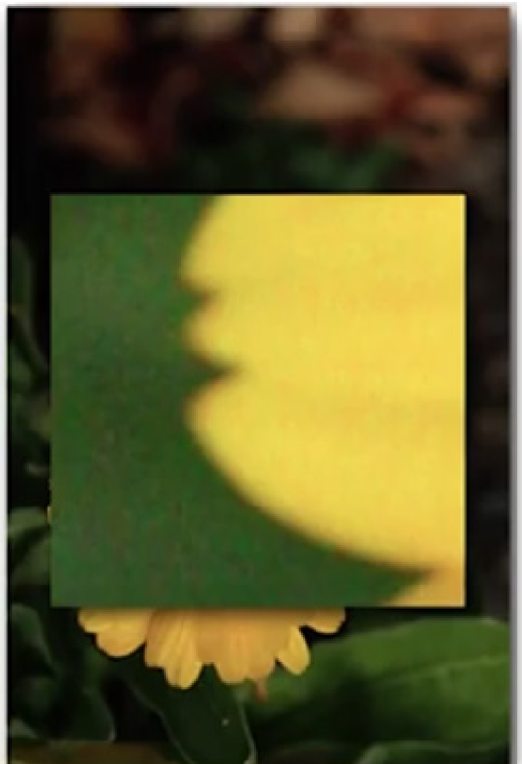
ISO 1600

- Third Variable in getting the right Exposure
- Film: Sensitivity vs. Grain (of film)
- Digital: Sensitivity vs. Noise (of sensor)
- Linear: 200 ISO needs half the light of 100 ISO

Sensor Gain - ISO



ISO 100



ISO 1600

* Third Variable in getting the right Exposure
* Film: Sensitivity vs. Grain (of film)
* Digital: Sensitivity vs. Noise (of sensor)
* Linear: 200 ISO needs half the light of 100 ISO

White Balance

Human Color Constancy:

QUIZ!
How would you define color constancy?

Reproducing this behavior in our cameras is a problem called
Machine Color Constancy

But wait!

QUIZ!
If humans have color constancy, why to try to reproduce it in digital devices?

Assumptions:

- All objects in the scene are **flat**.
- All objects are considered **Lambertian** (no specularities, diffuse reflection)
- Scene uniformly illuminated with a **single global illuminant**.

White Balance

$$R = \int_{380}^{740} r(\lambda) \times I(\lambda) \times S(\lambda)\ d\lambda$$

$$G = \int_{380}^{740} g(\lambda) \times I(\lambda) \times S(\lambda)\ d\lambda$$

$$B = \int_{380}^{740} b(\lambda) \times I(\lambda) \times S(\lambda)\ d\lambda$$

$$\longrightarrow$$

$$R = I(\lambda_R)S(\lambda_R)$$
$$G = I(\lambda_G)S(\lambda_G)$$
$$B = I(\lambda_B)S(\lambda_B)$$

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} I(\lambda_R) & 0 & 0 \\ 0 & I(\lambda_G) & 0 \\ 0 & 0 & I(\lambda_B) \end{bmatrix} \begin{bmatrix} S(\lambda_R) \\ S(\lambda_G) \\ S(\lambda_B) \end{bmatrix}$$

$$S(\lambda_R) = \frac{R}{I(\lambda_R)},\ S(\lambda_G) = \frac{G}{I(\lambda_G)},\ S(\lambda_B) = \frac{B}{I(\lambda_B)}$$

White Balance

$$\begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} = \begin{bmatrix} \frac{1}{I(\lambda_R)} & 0 & 0 \\ 0 & \frac{1}{I(\lambda_G)} & 0 \\ 0 & 0 & \frac{1}{I(\lambda_B)} \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

1. **Estimate illuminant components**
2. **Invert for white-balanced triplet**

Automatic White Balance

Gray World

Colors in the scene are sufficiently varied, and the average of reflectances is gray. In other words, reflectances are uniformly distributed over $[0, 1]$. For each waveband, **impose average 0.5**.

White Patch

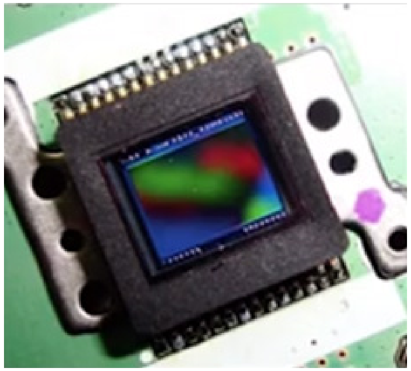
The brightest object in a scene becomes the reference white, it is perceived as white. An estimate of illuminant color is obtained by finding the brightest pixel in the scene. **Impose white presence**.

Sensors

Photographic Film

External Flash Disk

* Film and Digital Cameras are the same
  * There have been significant improvements in actuators, and lenses
* Difference is how light is trapped and preserved
* Chemical process for Film, and Electronic for Digital capture the moment in Time and Space

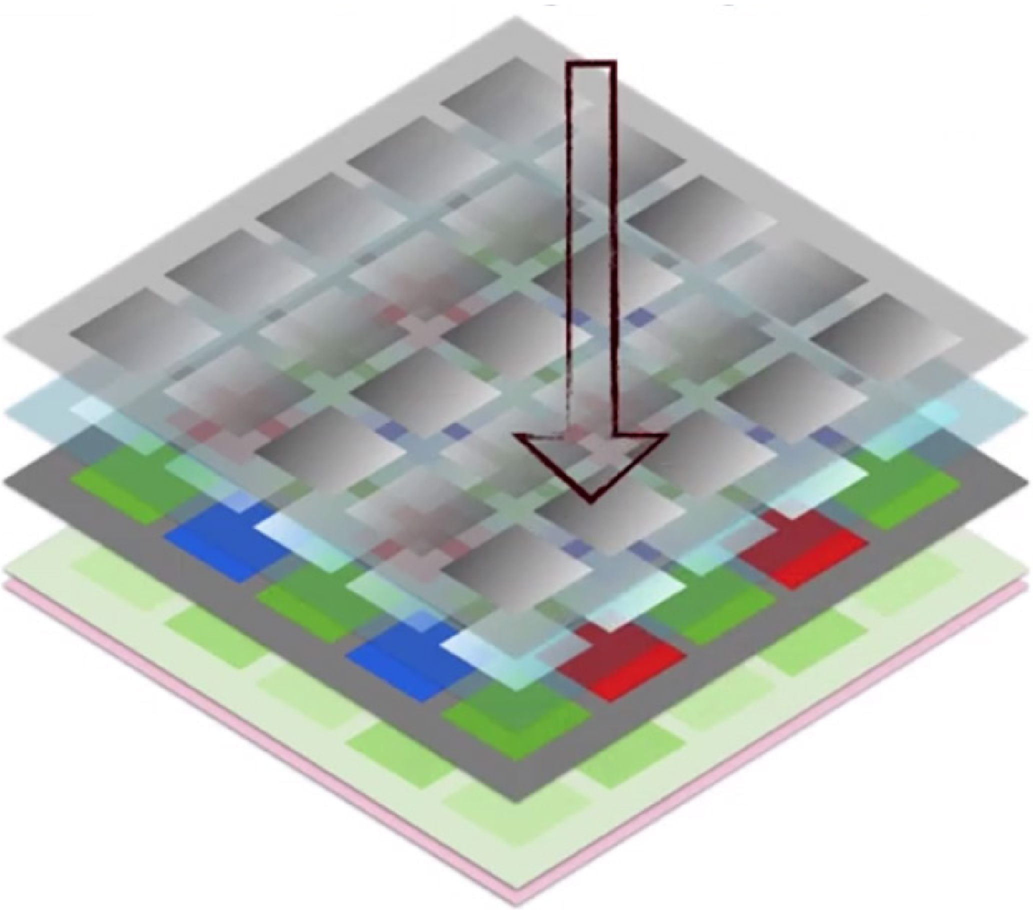


CCD

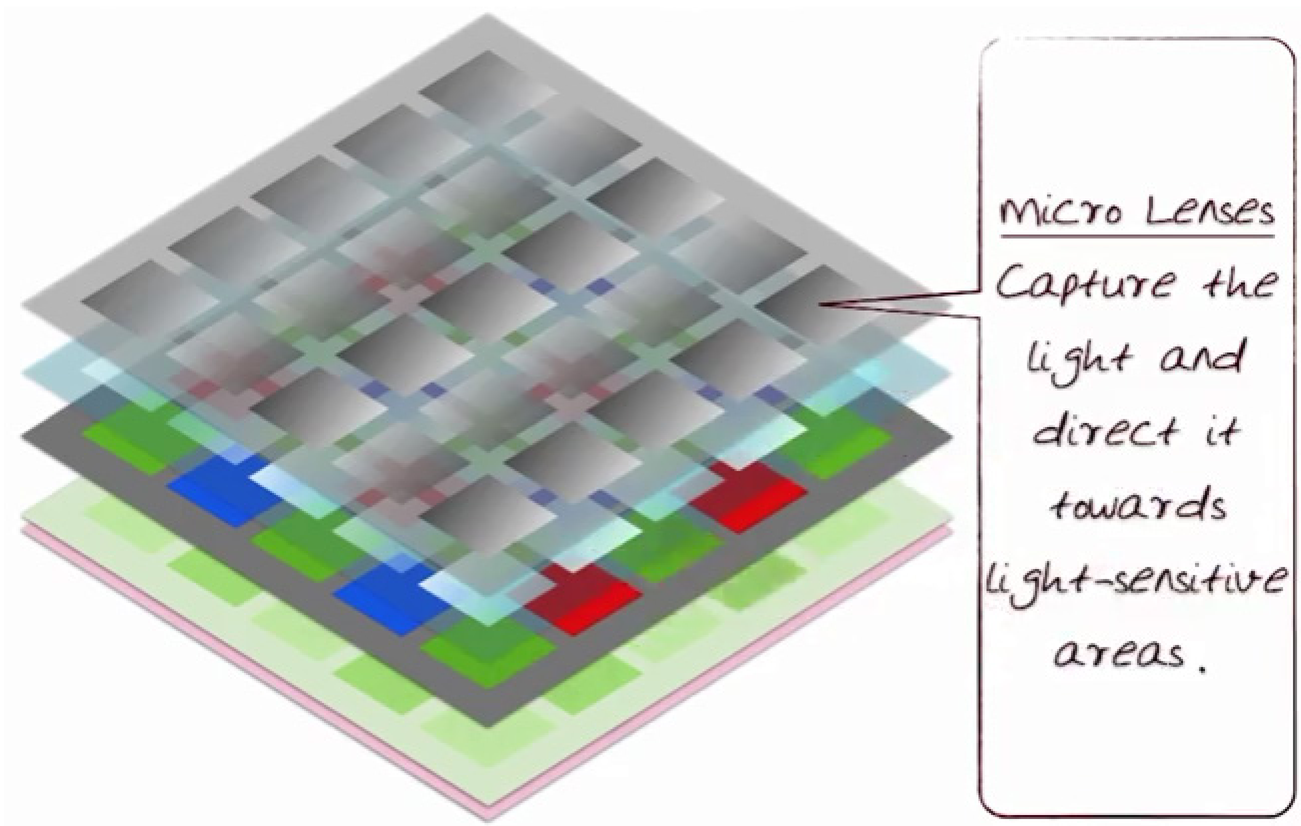
Digital Sensors - CCD

* CCD: Charge-Coupled Device, a device for converting electrical charge, into a digital value
* Pixels are represented by capacitors, which convert and store incoming photons as electron charges
* Willard Boyle and George E. Smith, 1969 (Won a Noble Prize in Physics in 2009).

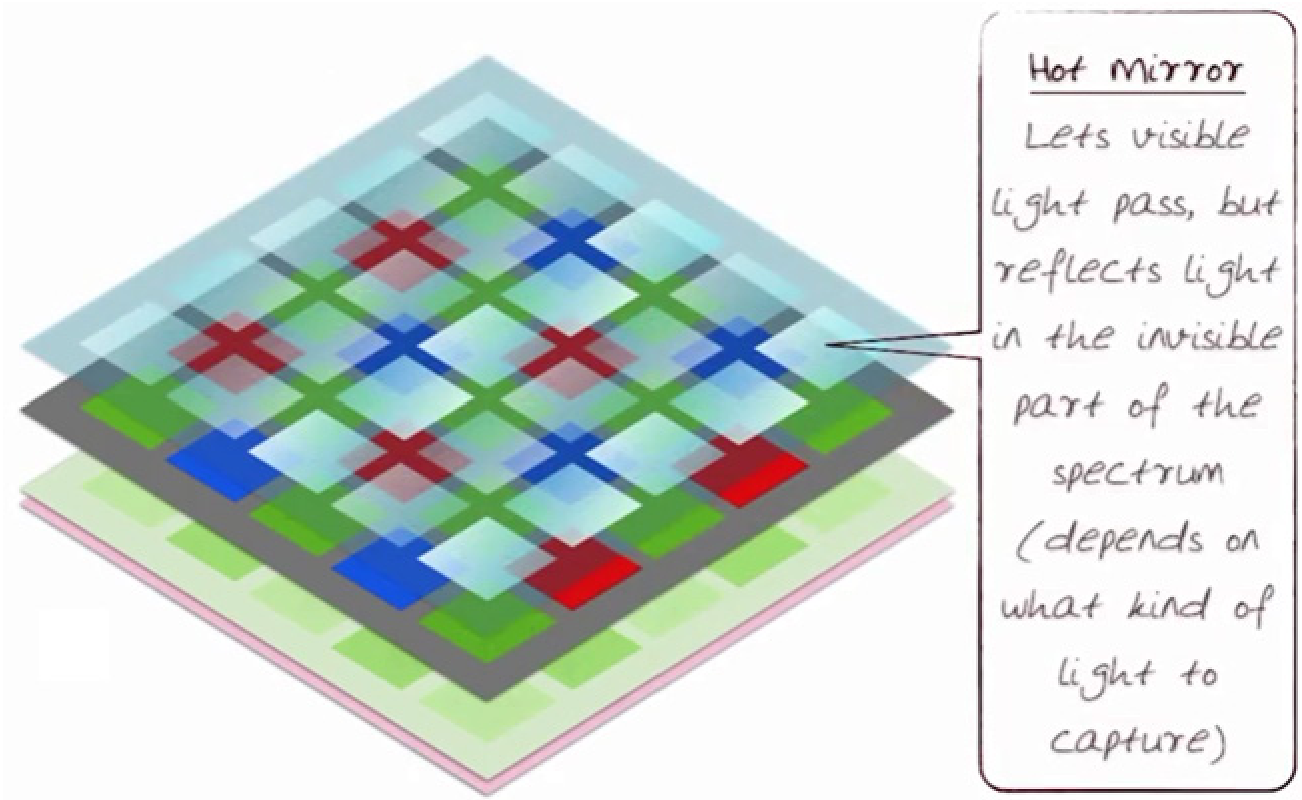
Digital Sensors - CCD



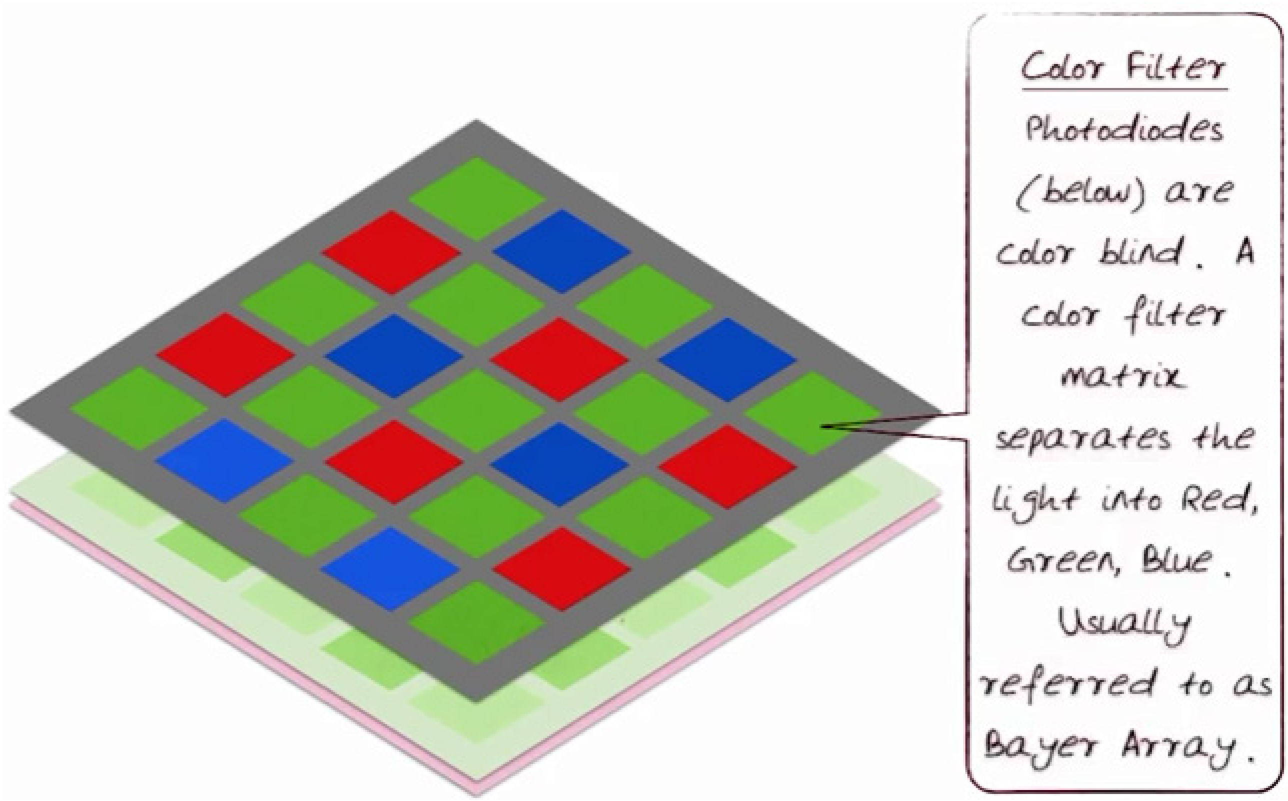
Digital Sensors - CCD



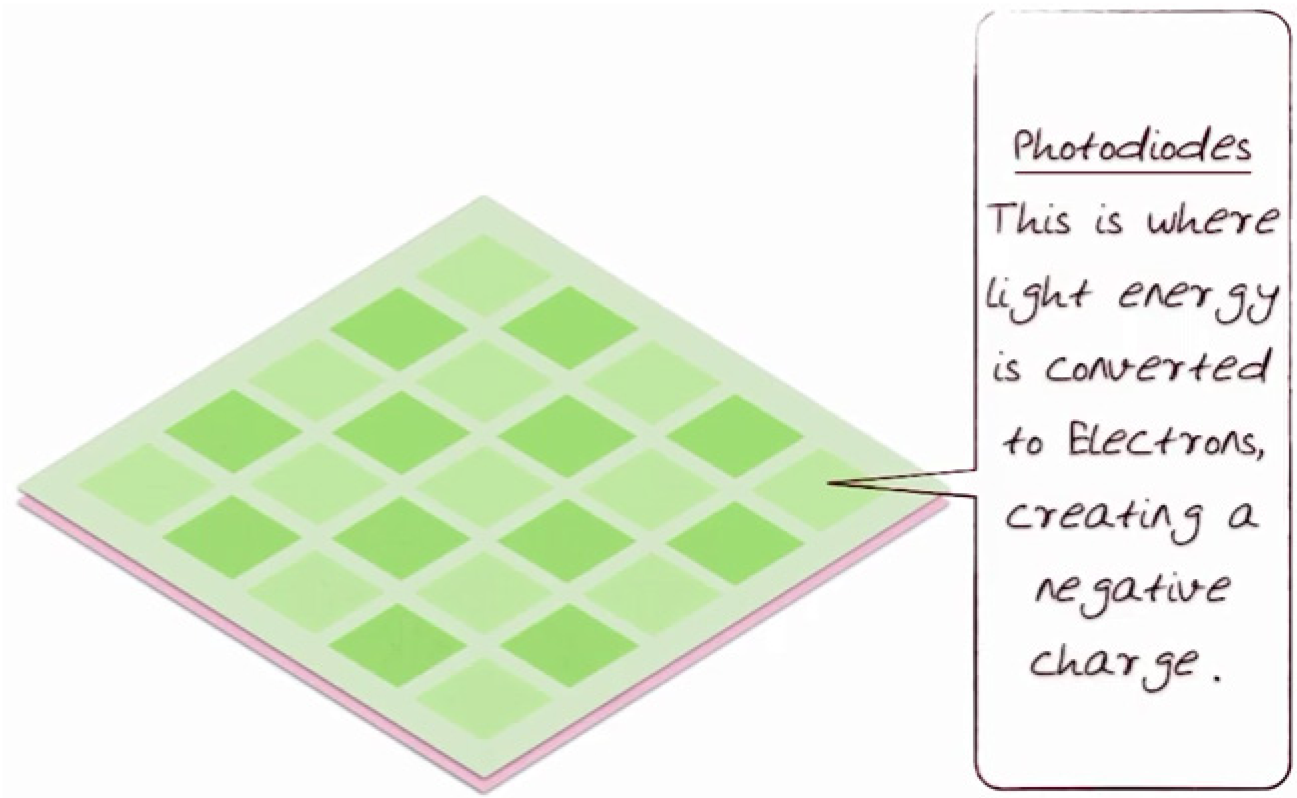
Digital Sensors - CCD



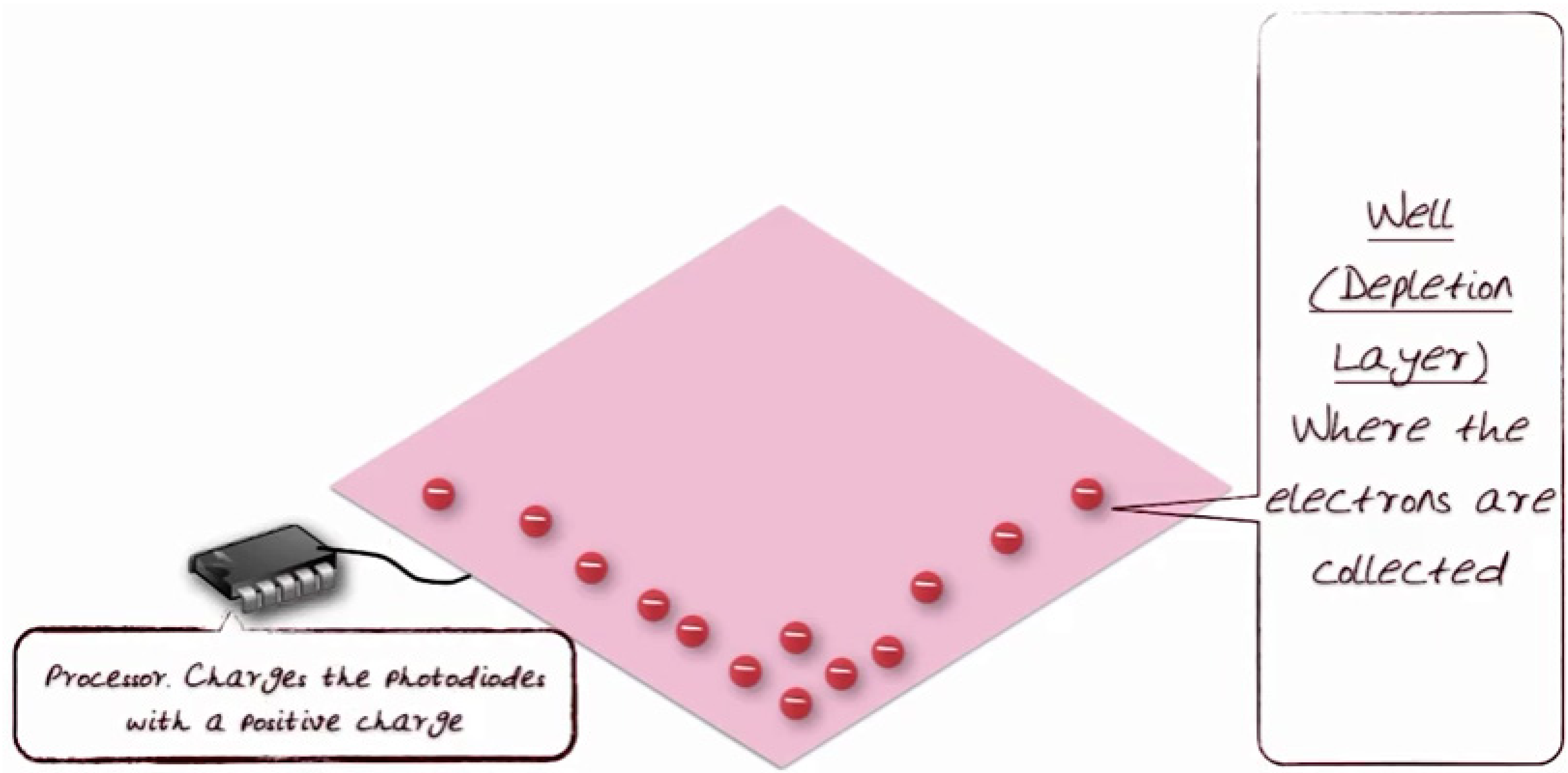
Digital Sensors - CCD



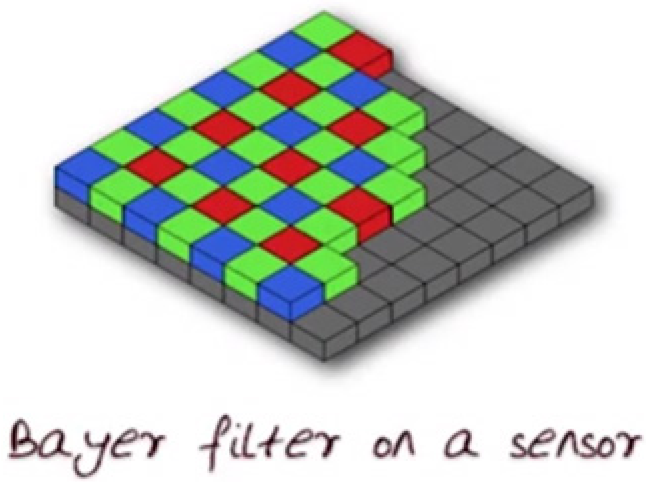
Digital Sensors - CCD

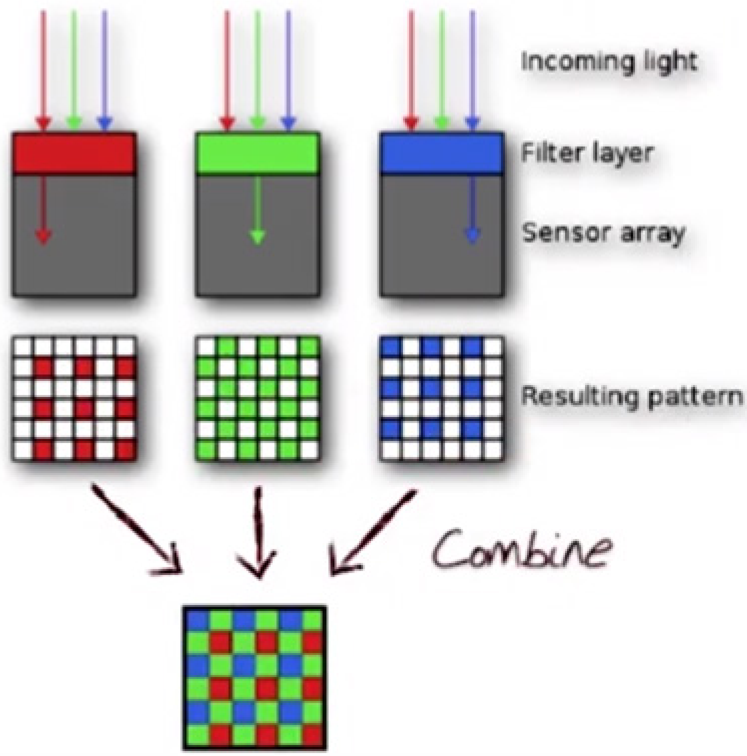


Digital Sensors - CCD

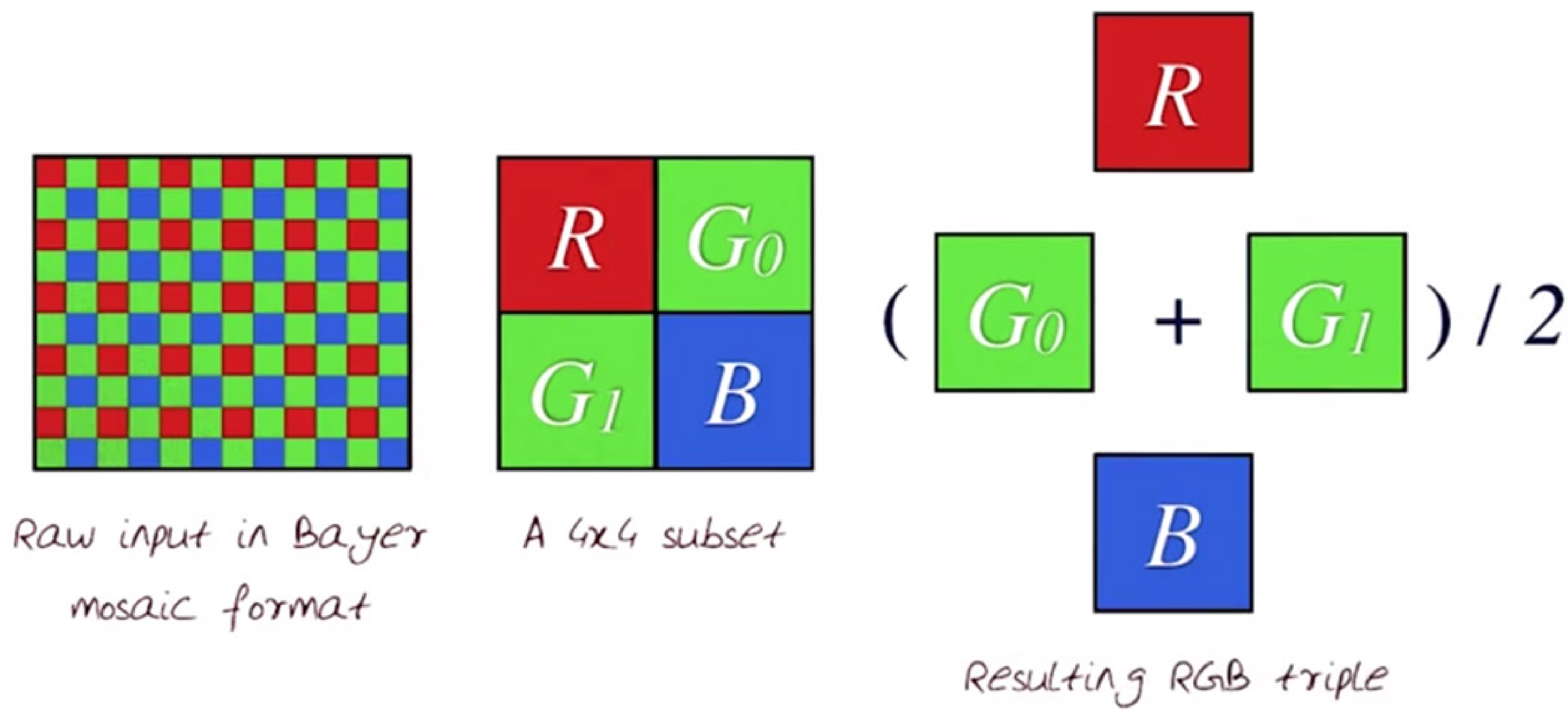


Bayer Filtering

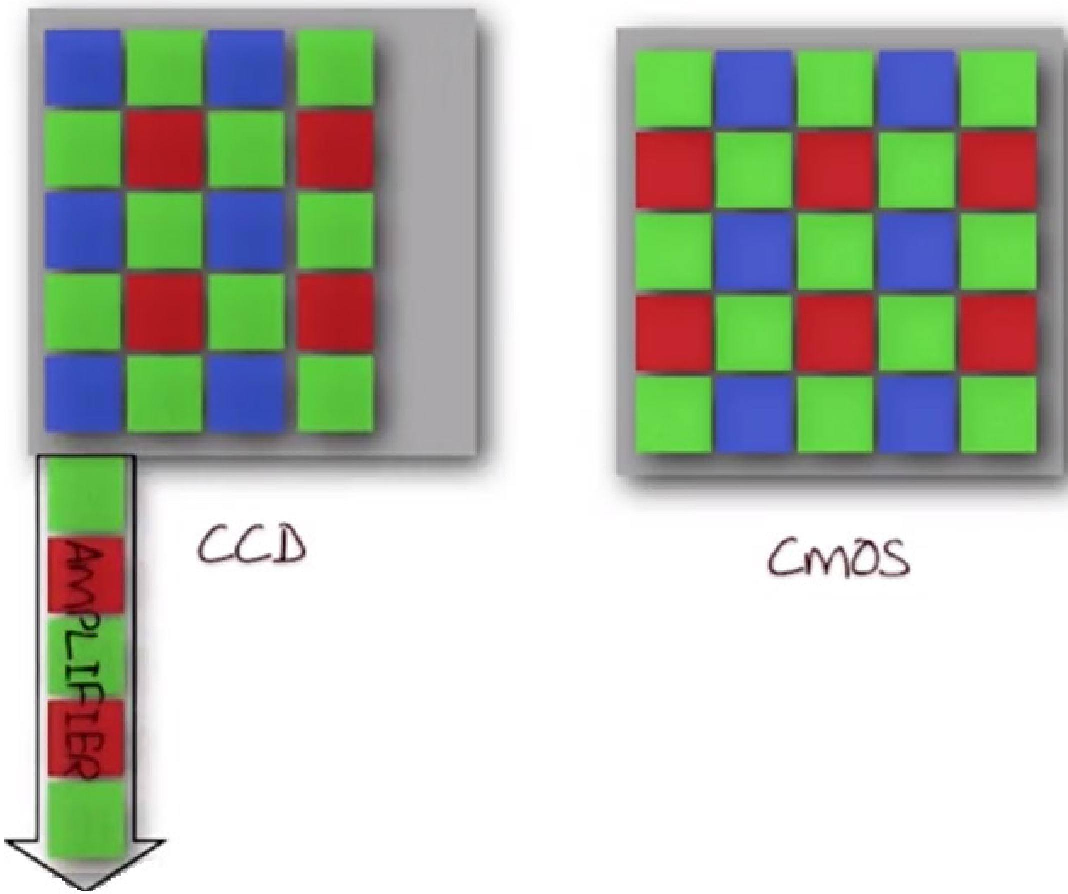




Bayer to RGB: Demosaicking

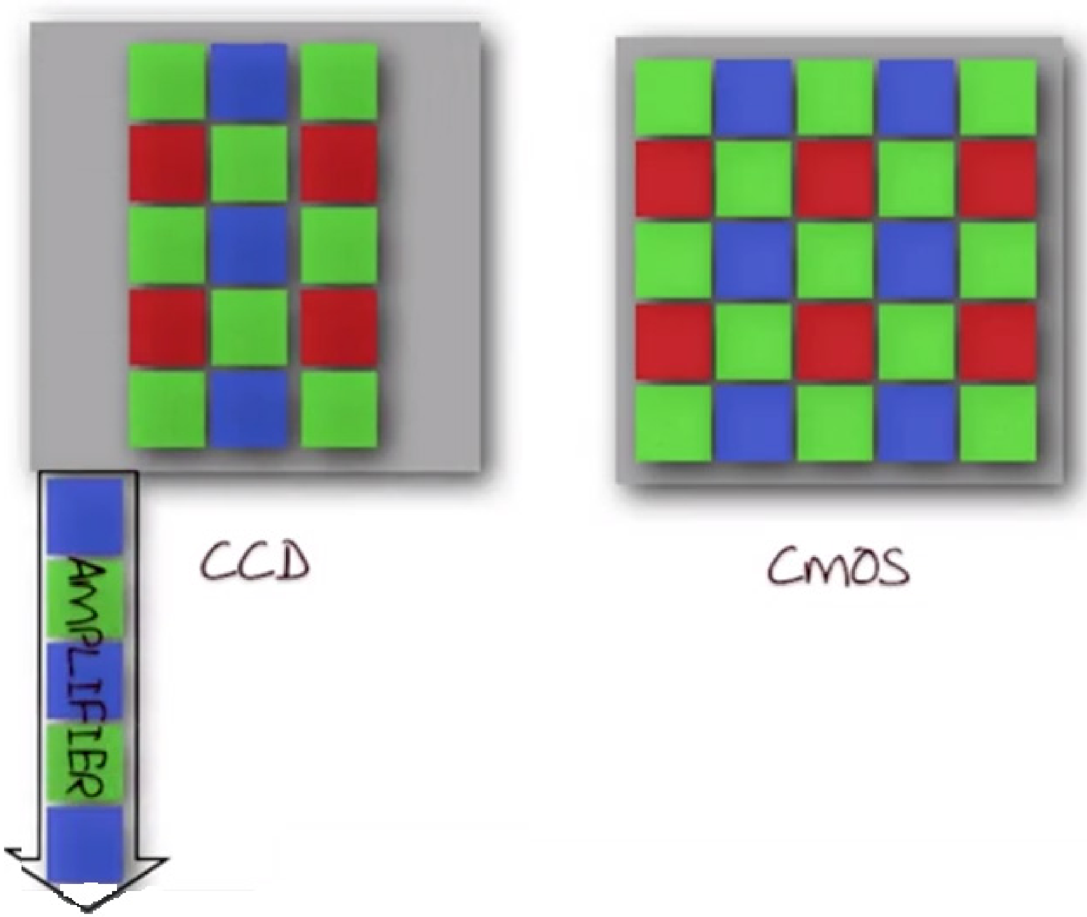


Digital Sensors - CMOS



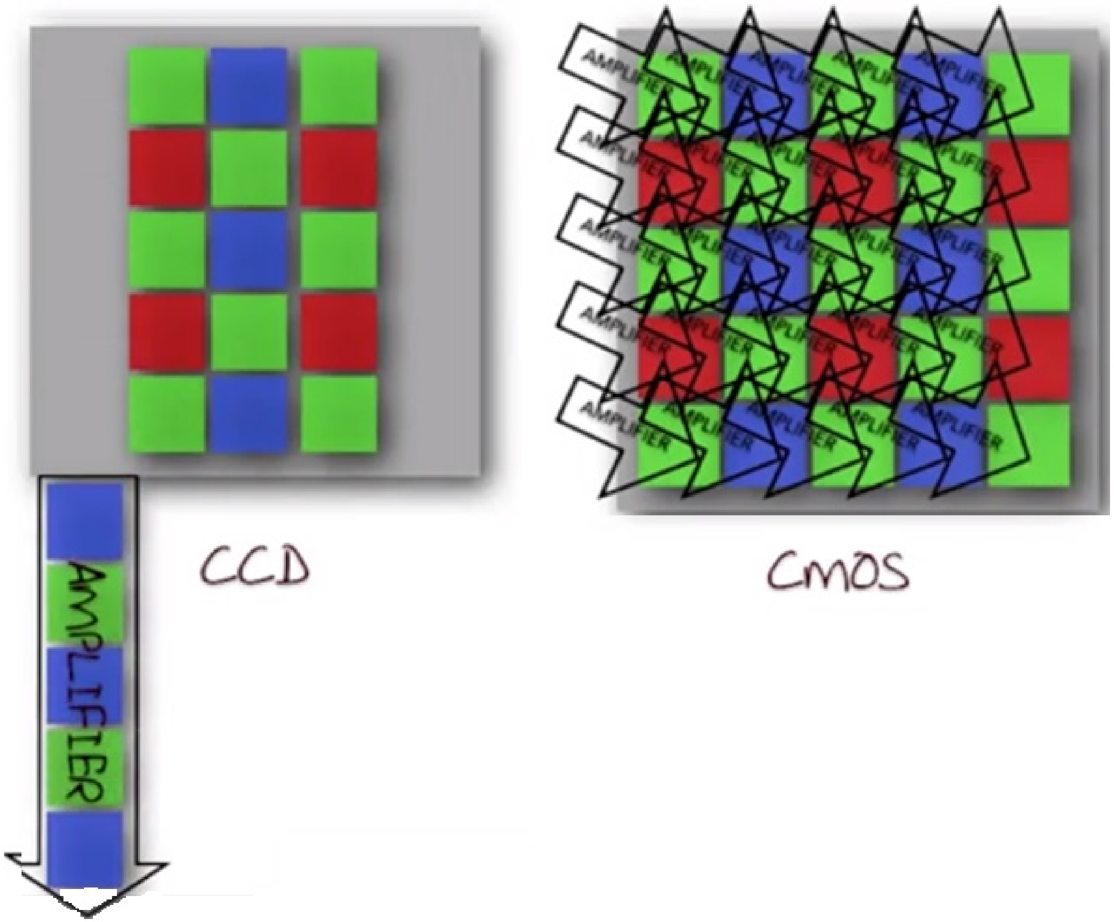
- CMOS: Complementary Metal Oxide Semiconductor
- Photosites in CCD are passive and do no "work"

Digital Sensors - CMOS



- CMOS: Complementary Metal Oxide Semiconductor
- Photosites in CCD are passive and do no "work"

Digital Sensors - CMOS



- CMOS: Complementary Metal Oxide Semiconductor
- Photosites in CCD are passive and do no "work"

Recording in RAW

* Contains minimally processed data from the sensor
* Image encoded in a device-dependent colorspace
* Captures radiometric characteristics of the scene
* Viewable image from the camera's sensor data
* Like a photographic negative
    * has a wider dynamic range or color; preserves most of the information of the captured image

Color Transformation

For the camera to accurately capture colors matching our perception, the color triplets obtained by the camera sensor should correspond to the cone responses of the human visual system.

However, for a series of practical reasons, emulating cone responses is not practical for image capture.

Still, our goal is that the stimulus the scene would have produced in the human visual system are estimated as accurately as possible.

$$\boxed{\textbf{RGB}} \longrightarrow \boxed{\textbf{XYZ}}$$

QUIZ!
Do you remember anything about this transformation?

Color Transformation

$$\boxed{\textbf{RGB}} \longrightarrow \boxed{\textbf{XYZ}}$$

Transform the (R,G,B) values of the sensor into (X,Y,Z) tristimulus values.

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \underbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}}_{\textbf{Colorimetric matrix}} \begin{bmatrix} R \\ G \\ B \end{bmatrix} = A \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

This process would imply computing a colorimetric matrix for every single triplet (R,G,B), really unpractical.

Color Transformation

Cameras carry a unique (optimal) colorimetric matrix, computed as follows:

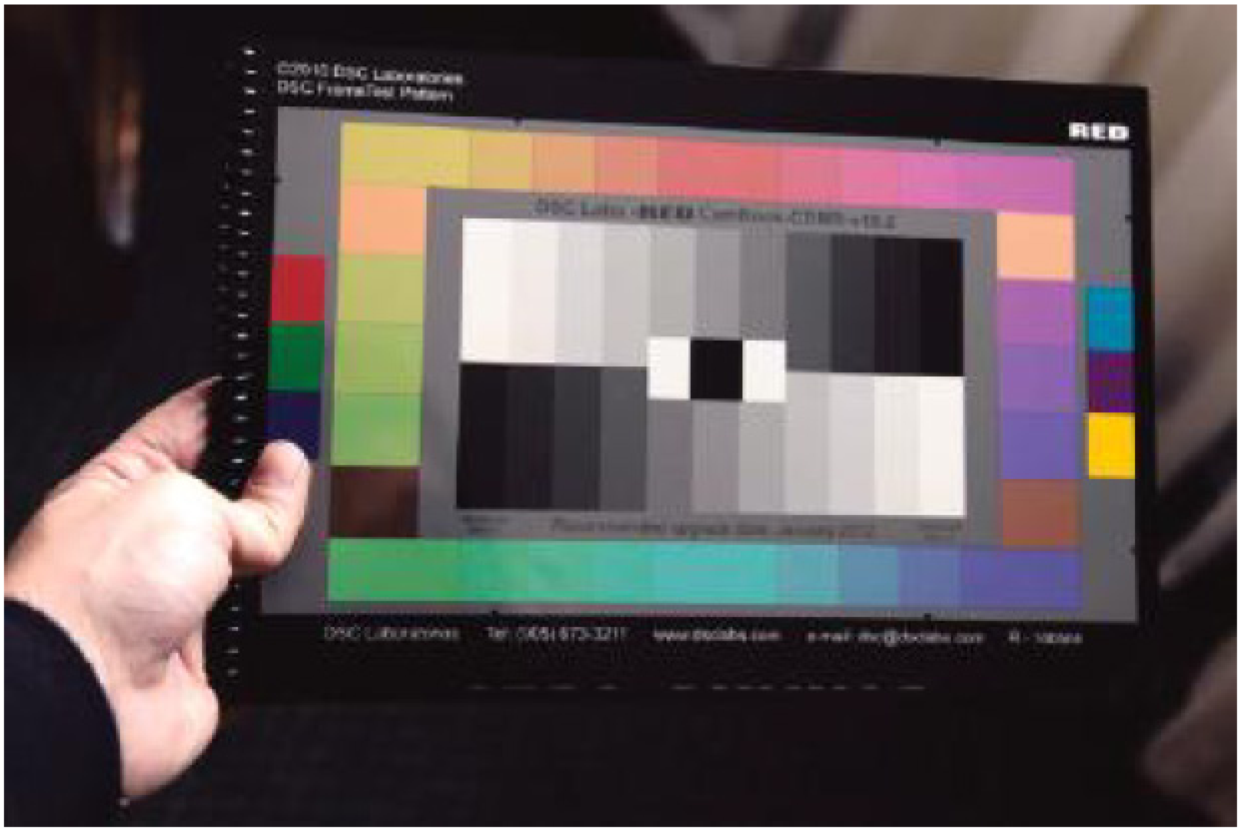
- Build a set of n test patches of representative or important colors.
- Under controlled conditions, with a known illuminant, measure the tristimulus values of the patches with a tristimulus colorimeter obtaining (Xi, Yi, Zi), for $1 < i < n$.
- Under the same conditions, use the camera to measure the (R, G, B) values of the patches, obtaining (Ri, Gi, Bi), for $1 < i < n$.
- Compute an optimal matrix A that minimizes the color difference between the target stimulus and corresponding RGB2XYZ transformed points.

Ideally a different matrix should be used for each different scene illuminant.

The above process finds the best colorimetric matrix for the calibration patch set under a given illuminant, and many cameras (most consumer models) use only this one matrix.

Some cameras come with several pre-set matrices computed under different illuminations. e.g., fluorescent light, daylight, halogen, etc.

Color Transformation



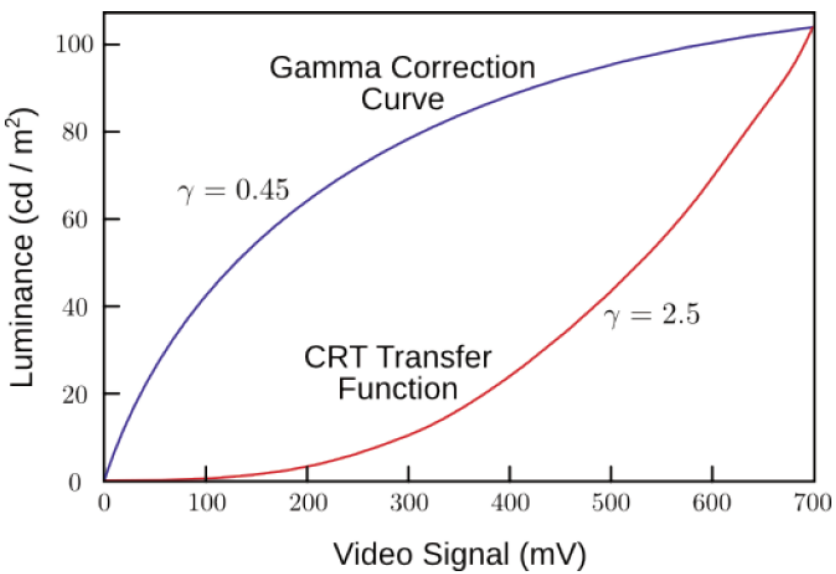
Color Transformation

A last step that is given before projection into desired display is **Gamma Correction**.

At the onset of broadcast television it was observed that CRTs produced luminance as a non-linear, power function of the device's voltage input.

$$\boxed{L = V^{\gamma},\ \gamma \equiv 2.5}$$

For correct luminance reproduction on a TV set, the camera's luminance signal needed to be non-linearly scaled with a power function with the inverse exponent of the CRT's power function.



But why is gamma correction still used today, when CRTs have become obsolete?

Humans have a perceptual response to luminance that is also non-linear: perceived lightness is roughly the 0,42 power of luminance. Differences in the dark parts of an image are more noticeable than differences of the same amount on bright parts of the same image. We want more bits in dark areas!

Outline of these lectures - Week 1

- **B1 - Human Vision System: The Retina and the Brain**
  1. **Basics of Human Perception.**
  2. **Perception and Color. Measuring Color.**
  3. **Color Spaces. Perceptually Uniform Color Spaces.**
  4. **Application: Color Image Segmentation. Superpixels.**
  5. **Introduction to the Retinex model for Human Perception.**
  6. **Introduction to Computational Image Processing.**
  7. **Application: A first example - Image Composition.**

- **B2 - Digital Image Processing: The Camera**
  1. **The Image Acquisition Pipeline.**
  2. **Cameras: Lenses, Exposure, Focus, White Balance, Sensors, Output.**
  3. **Unconventional Image Acquisition: Panorama.**
  4. Application: Panorama Creation.