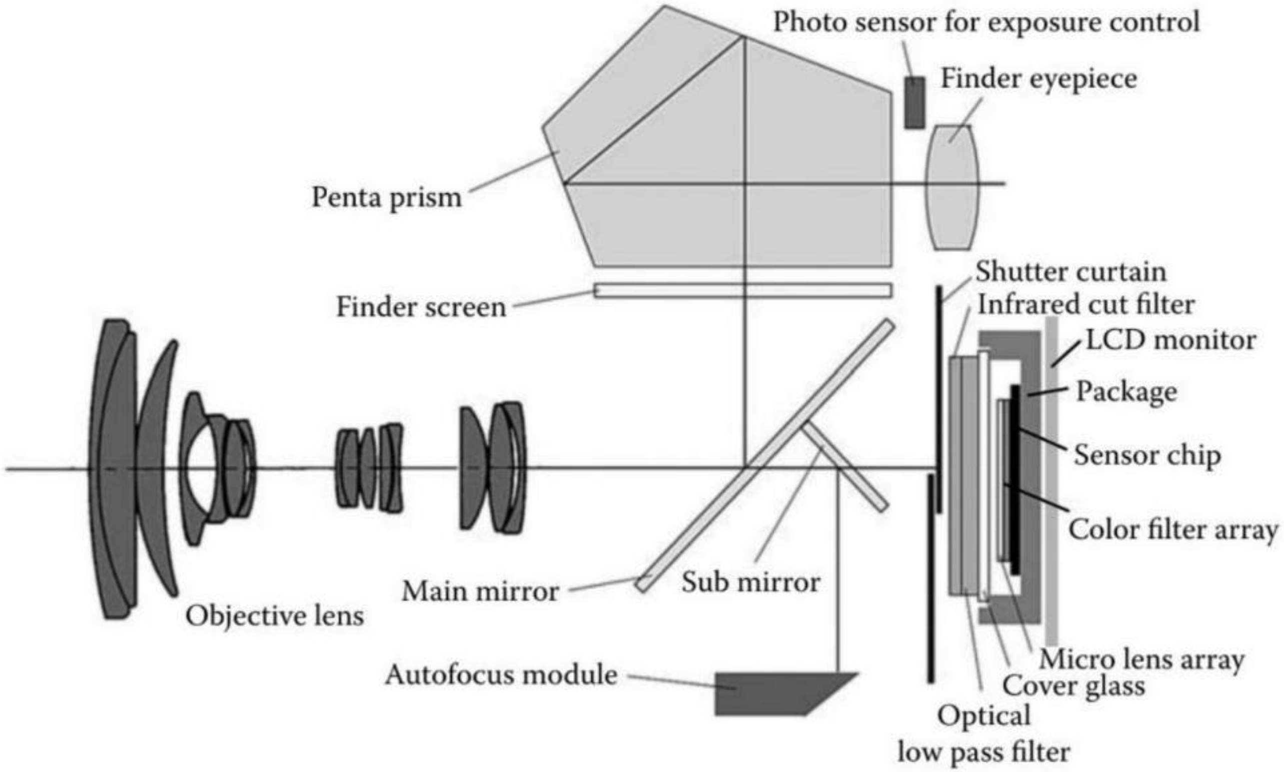


Digital Image Fundamental

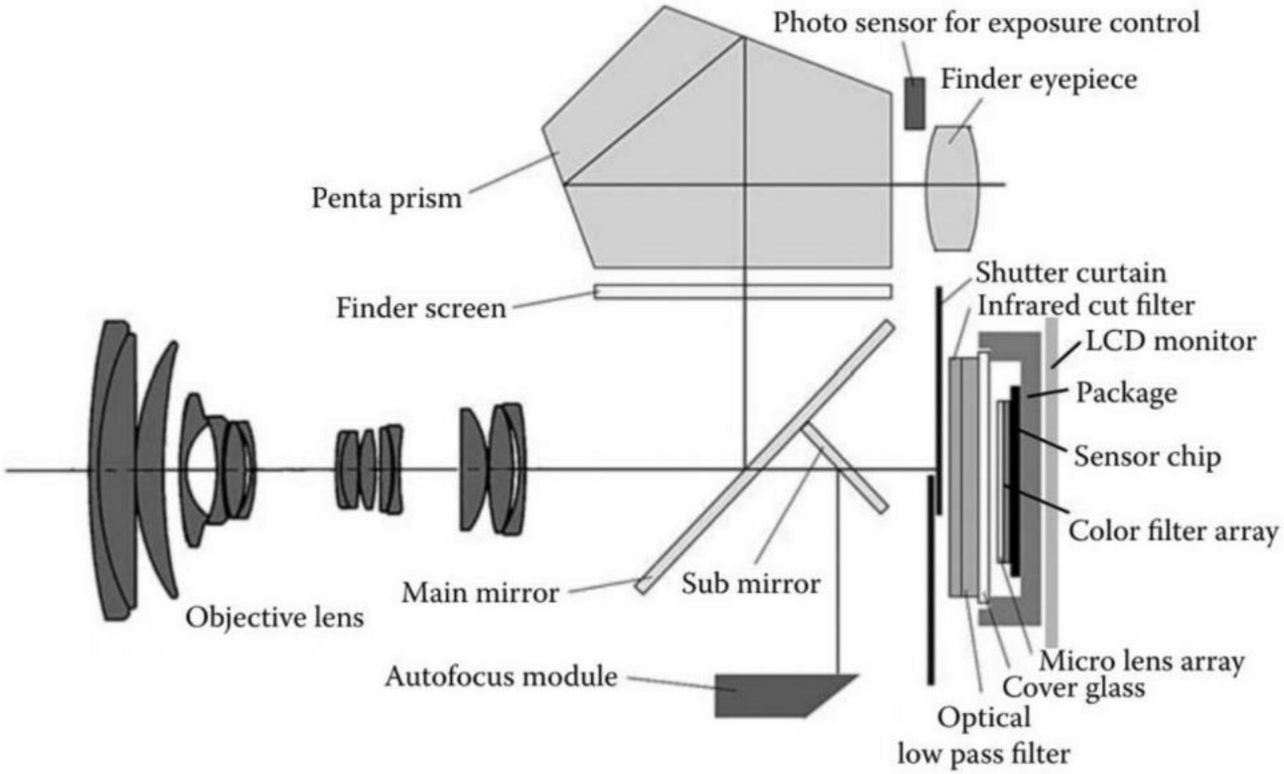
Fall 2024

Yi-Ting Chen

slido

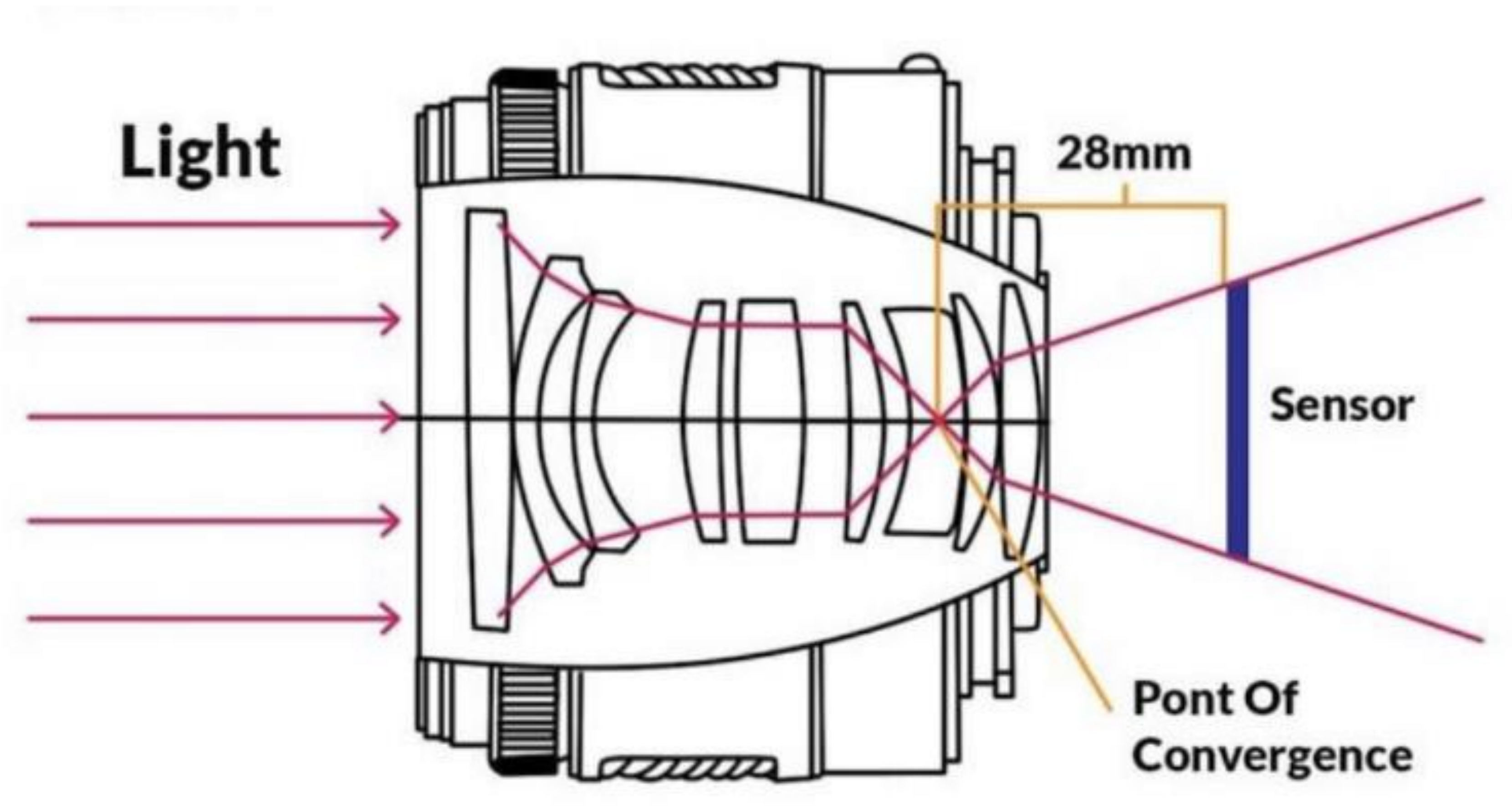
Source: Toyoda, “Digital still cameras at a glance”

Inside a digital single lens reflex (DSLR)



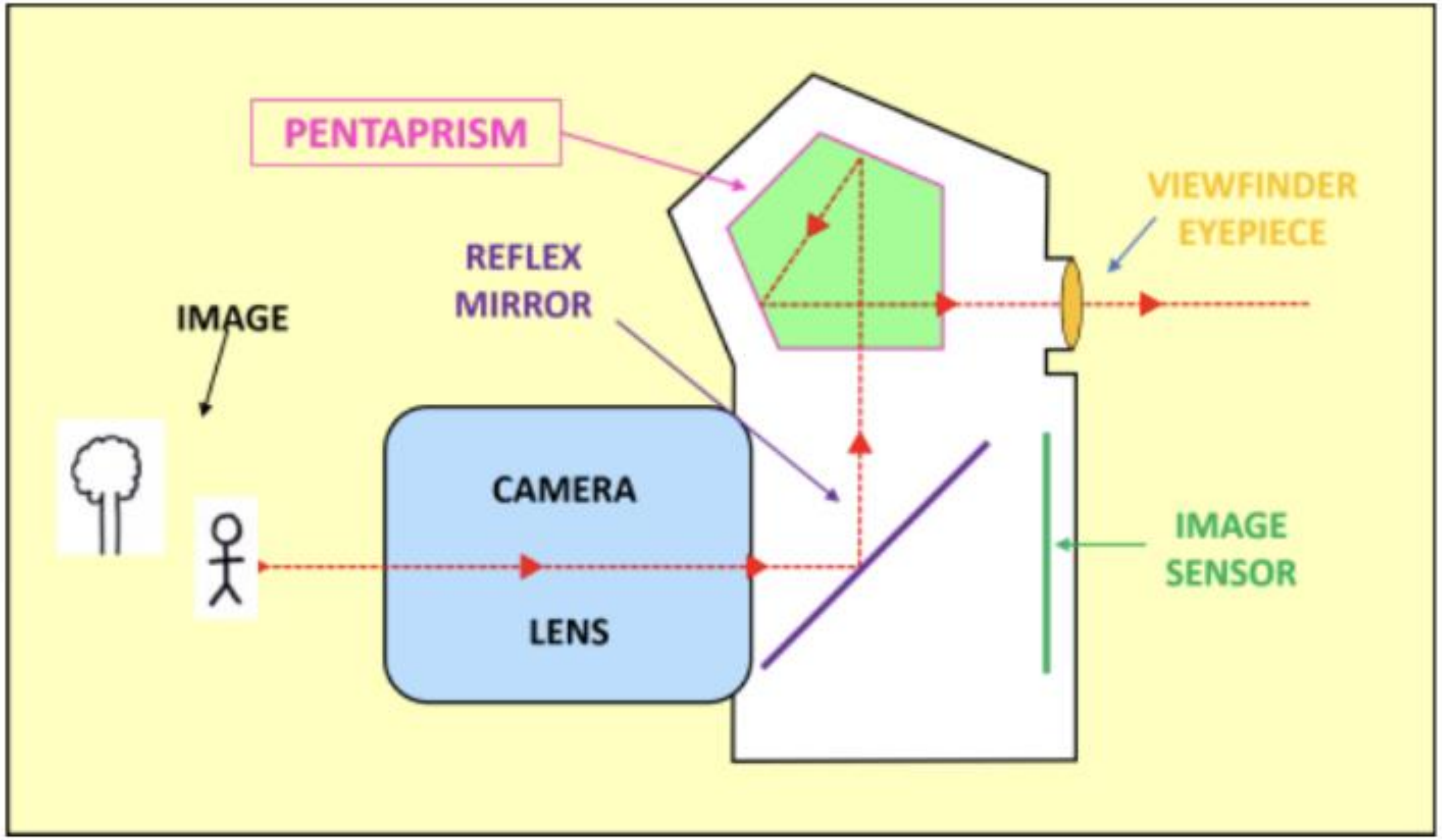
Source: Toyoda, “Digital still cameras at a glance”

Lens



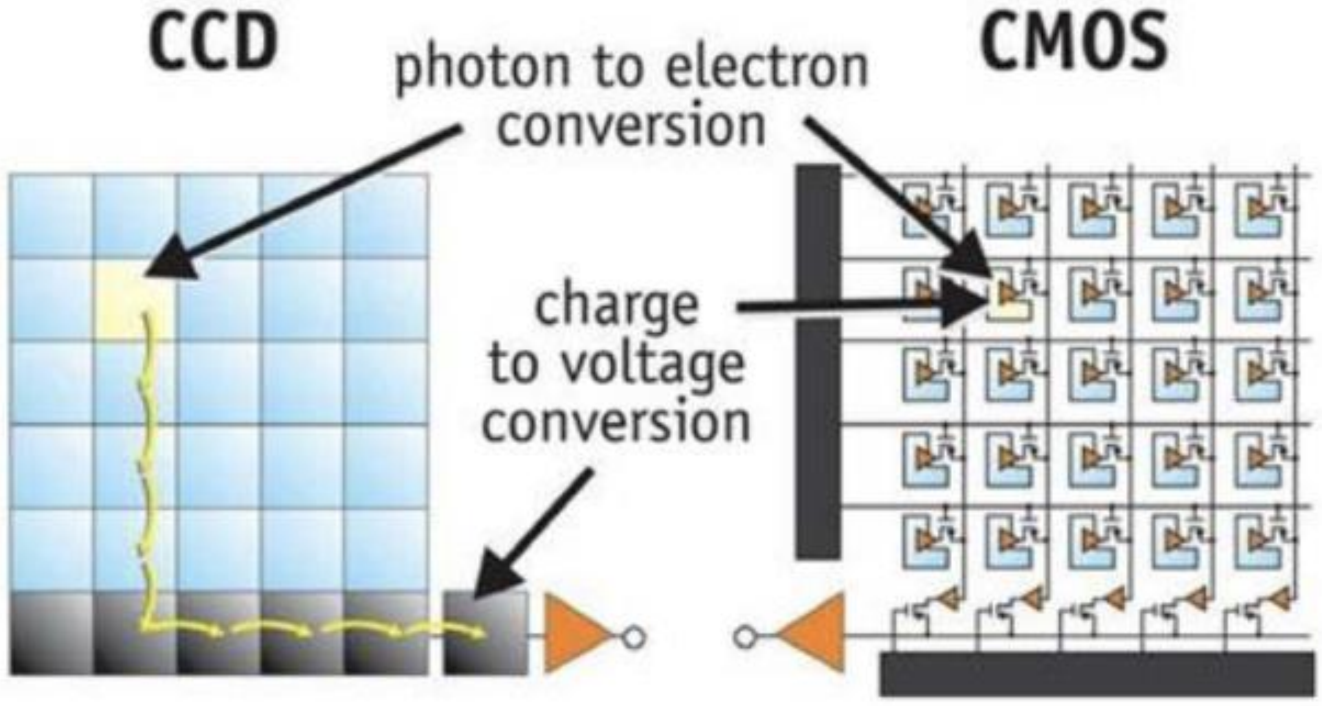
https://thesmartphonephotographer.com/smartphone-camera-focal-length/

The optical path from the aperture to the viewer



https://www.houseofmaths.co.uk/2017/10/parallax-pentaprisms-and-photography/

Image Sensors



https://meroli.web.cern.ch/lecture_cmos_vs_ccd_pixel_sensor.html

CCD

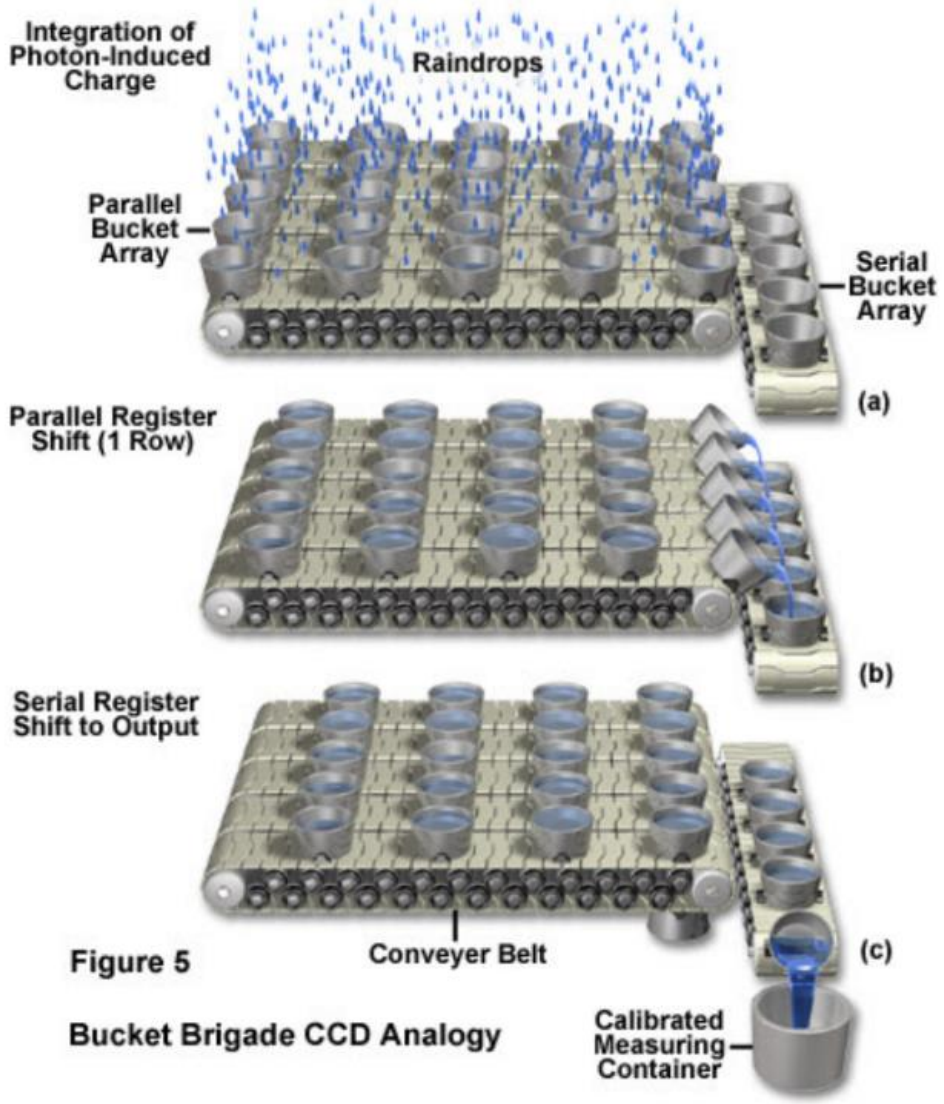
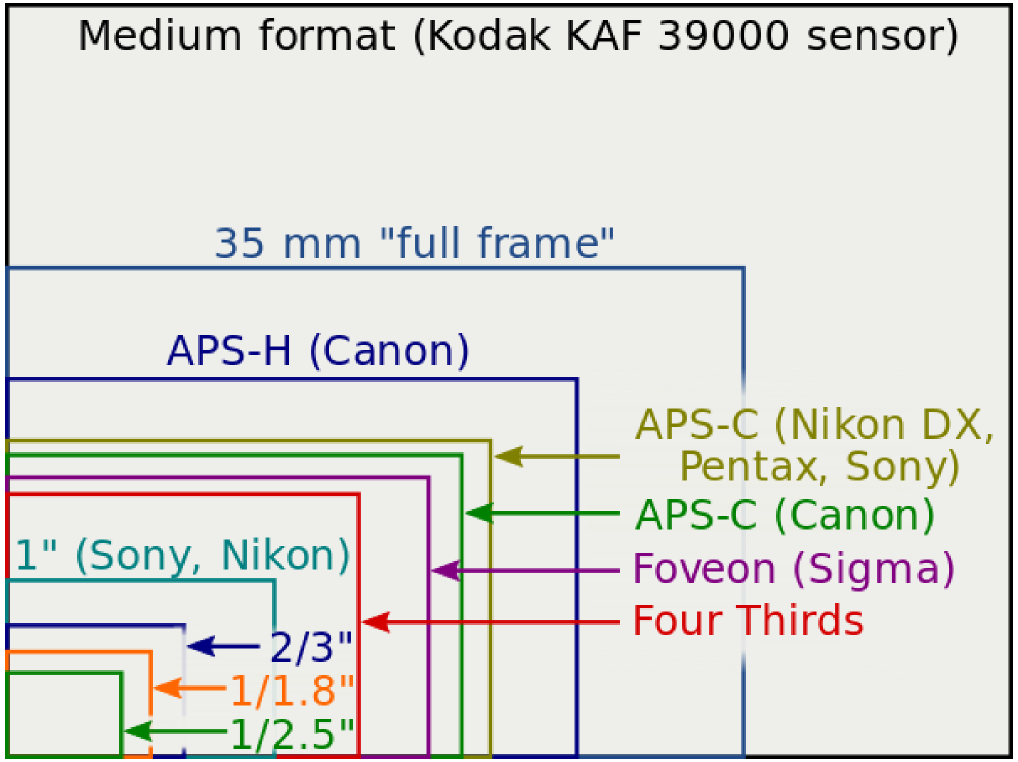

Integration of Photon-Induced Charge
Raindrops
Parallel Bucket Array
Serial Bucket Array
(a)
Parallel Register Shift (1 Row)
(b)
Serial Register Shift to Output
(c)
Figure 5
Conveyer Belt
Bucket Brigade CCD Analogy
Calibrated Measuring Container

Image Sensors



film camera

https://www.letsgojp.com/archives/527854

https://en.wikipedia.org/wiki/Image_sensor_format

Develop your film in a darkroom

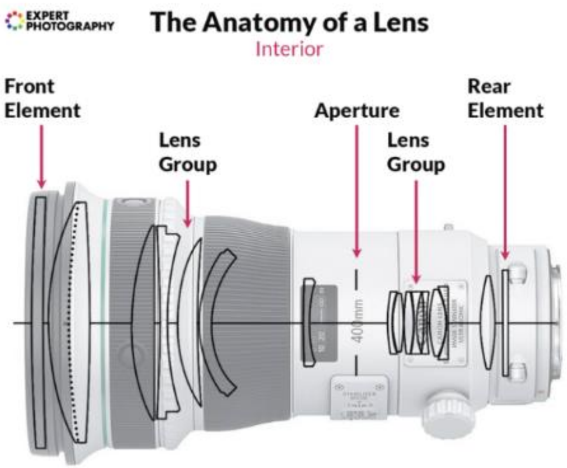
https://www.shutterstock.com/search/darkroom-process

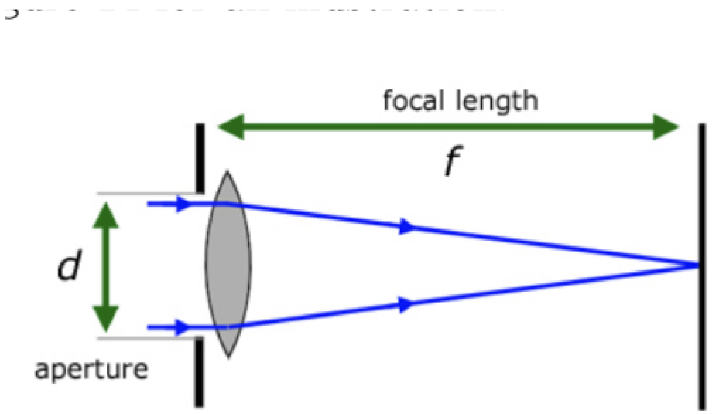
How to use a camera to produce a high-quality image?

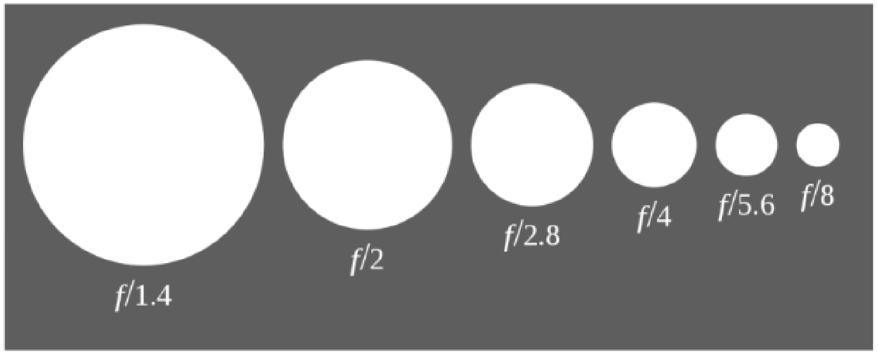
- F-number
- Shutter speed
- Sensor gain and ISO

F-number: How much light you can collect

- Aperture: controls the amount of light that can reach the image sensor
- F-number is the ratio of the focal length to the diameter of the aperture



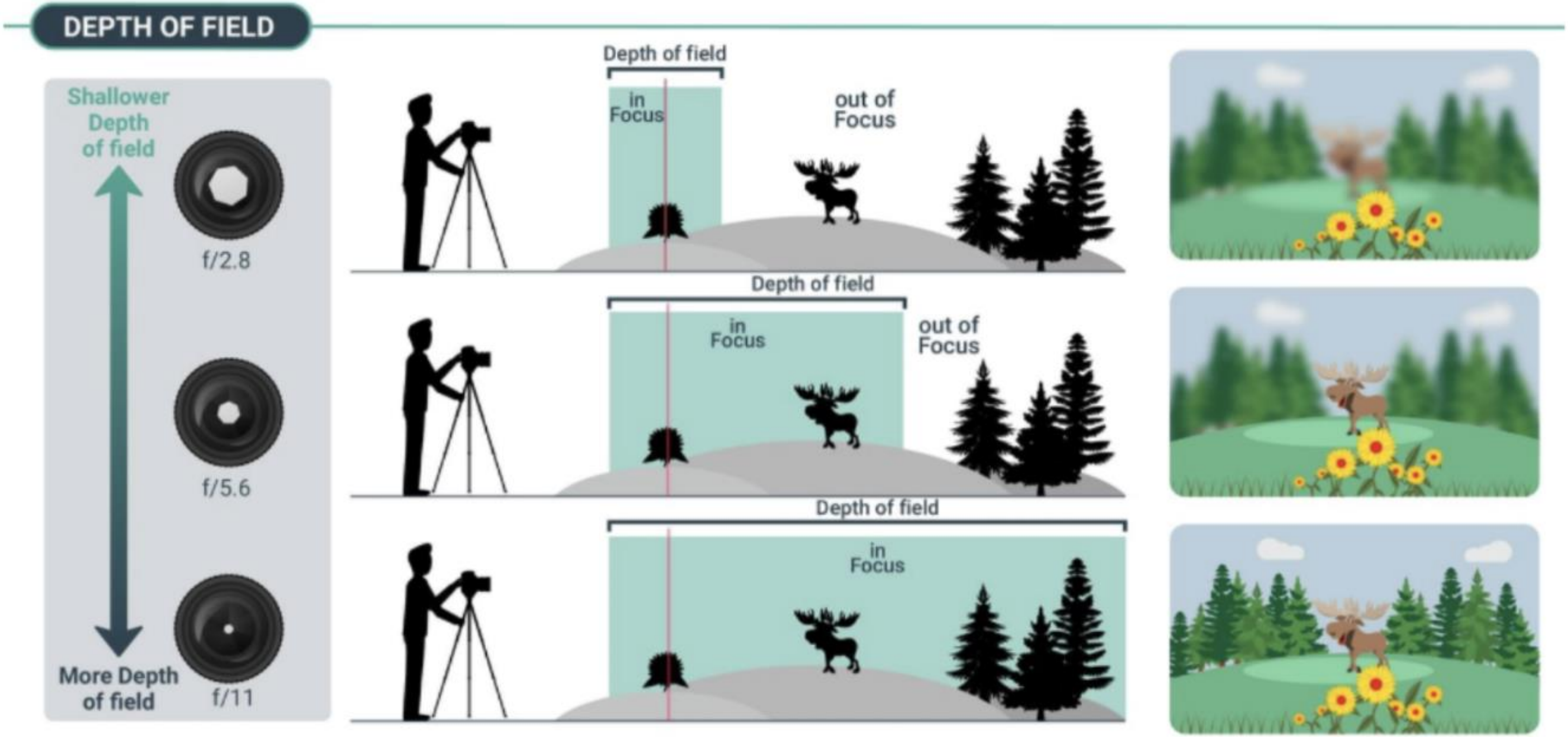




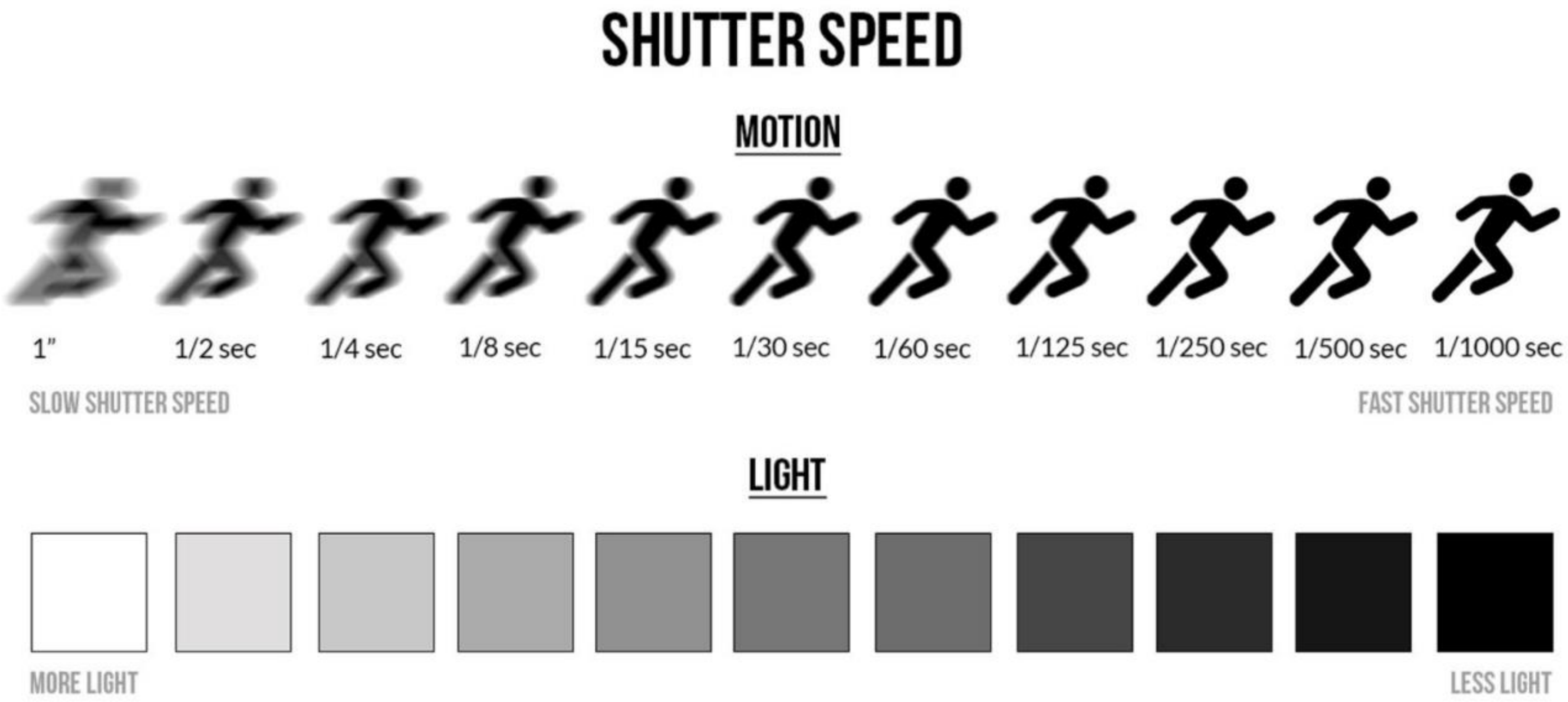
(a) changing the f-number

(b) changing the focal length

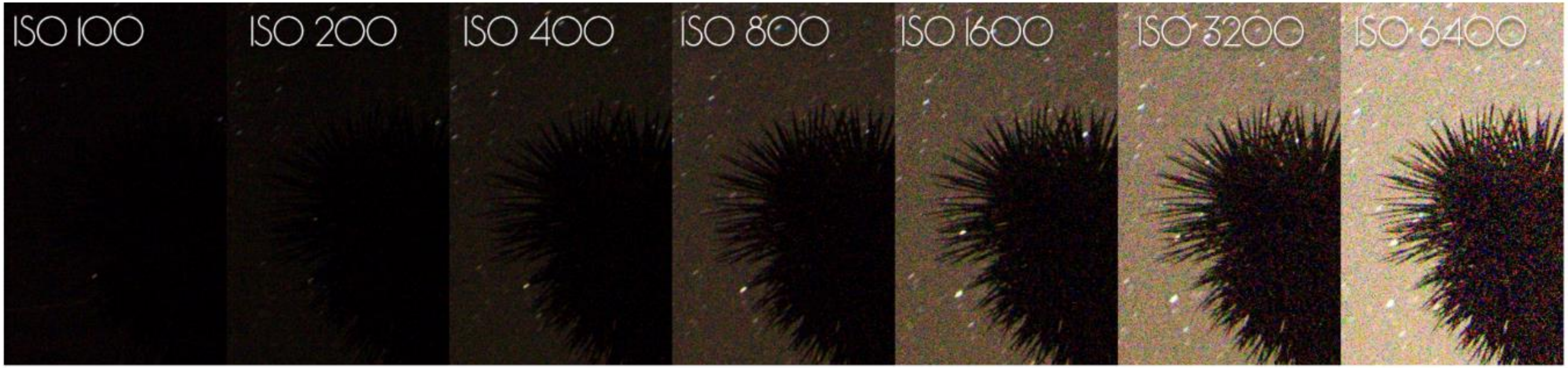
https://www.theverge.com/2012/1/2/2663464/camera-buyers-guide



https://capturetheatlas.com/depth-of-field-photography/



https://karltayloreducation.com/class/shutter-speeds-2/



https://www.lonelyspeck.com/


100mm
f/8.0
1/200
ISO 1600

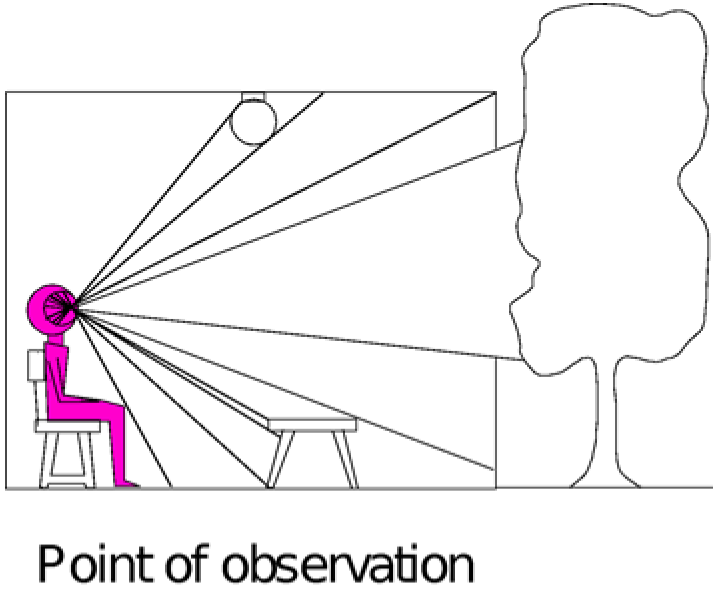

100 mm
f/2.8
1/320
ISO 400

https://unsplash.com/photos/butterfly-perched-on-flower-at-daytime-BtoUjLUtPnQ

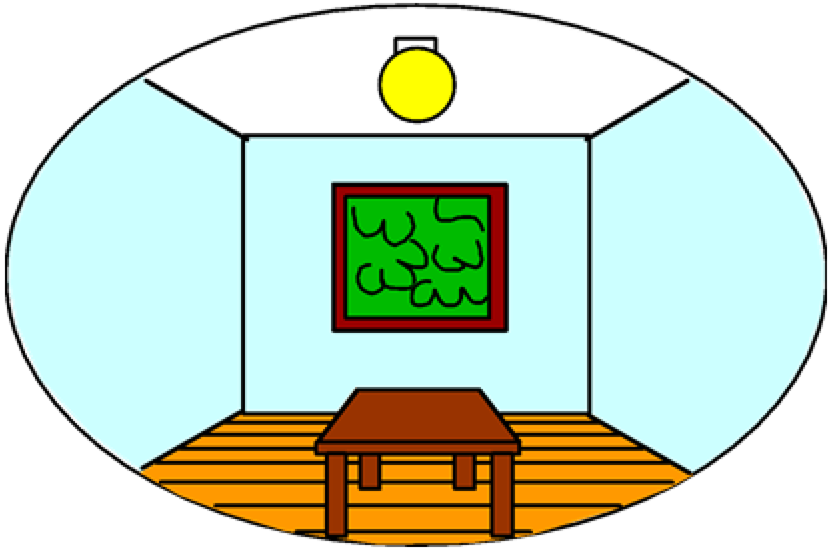
Any questions?

Image Formation

- Projection of 3D scene onto 2D plane
- Understand the **geometric** and **photometric** relation between the scene and its image



3D scene



2D image

What is an image?

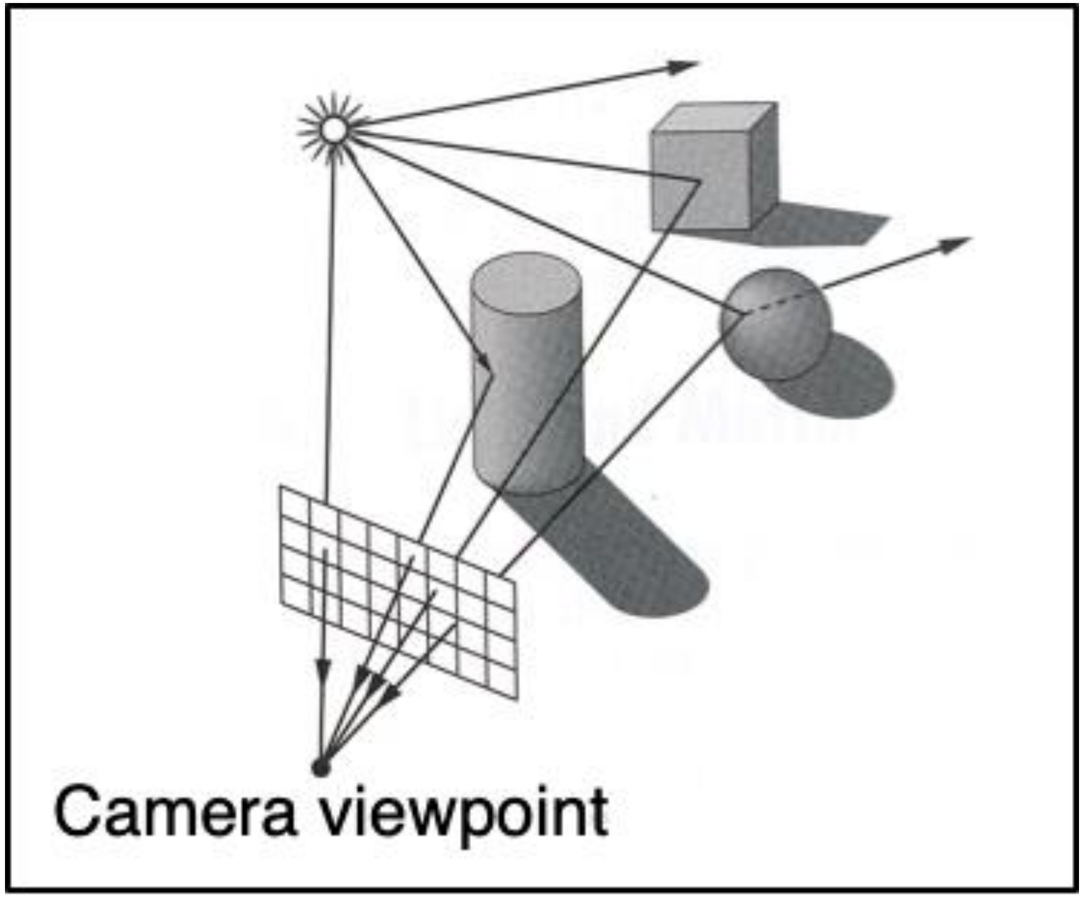
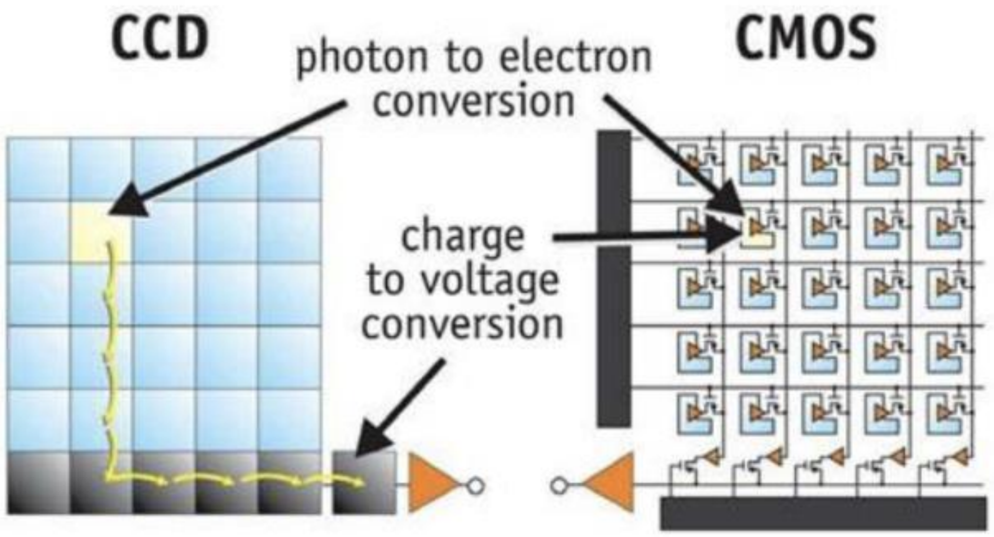
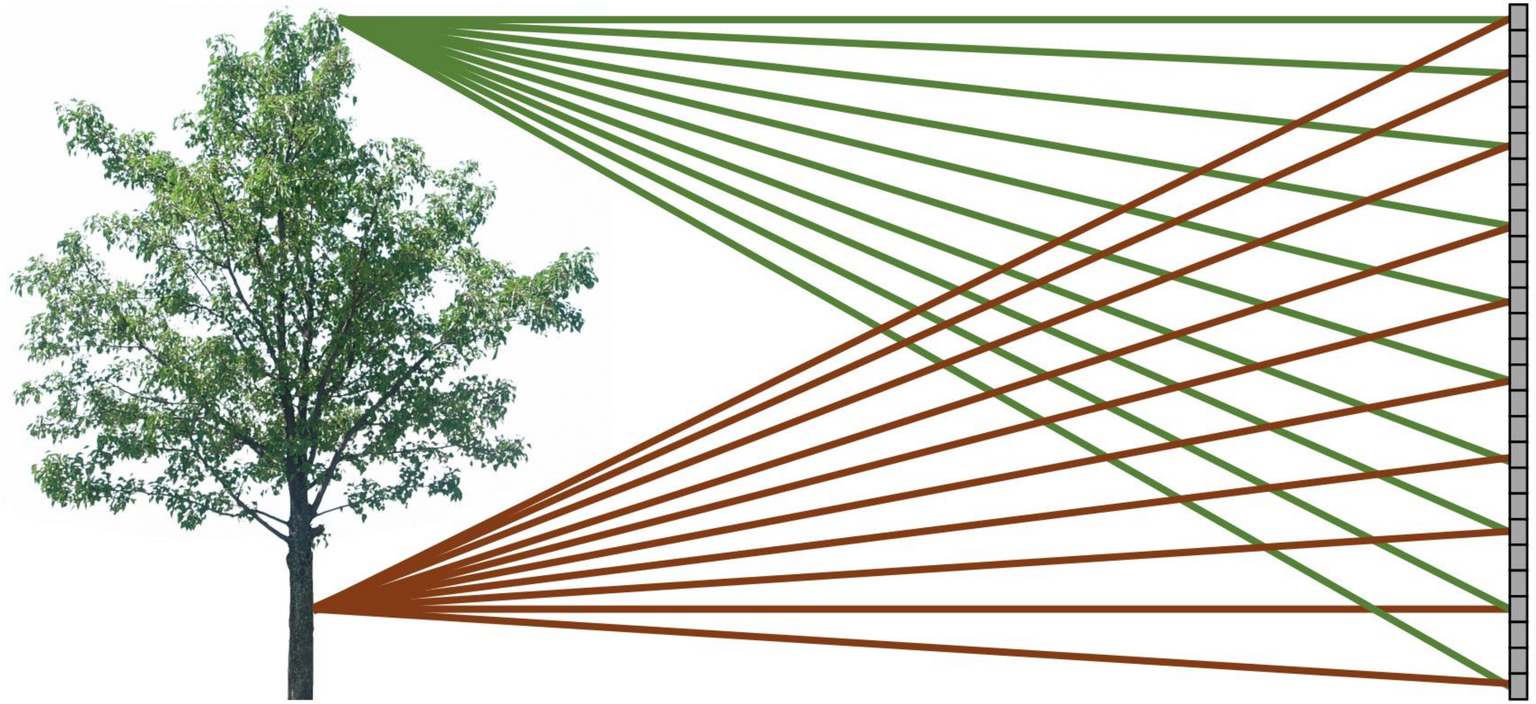


Image is a collection of pixels that are samples of radiance arriving at a camera viewpoint

Imaging

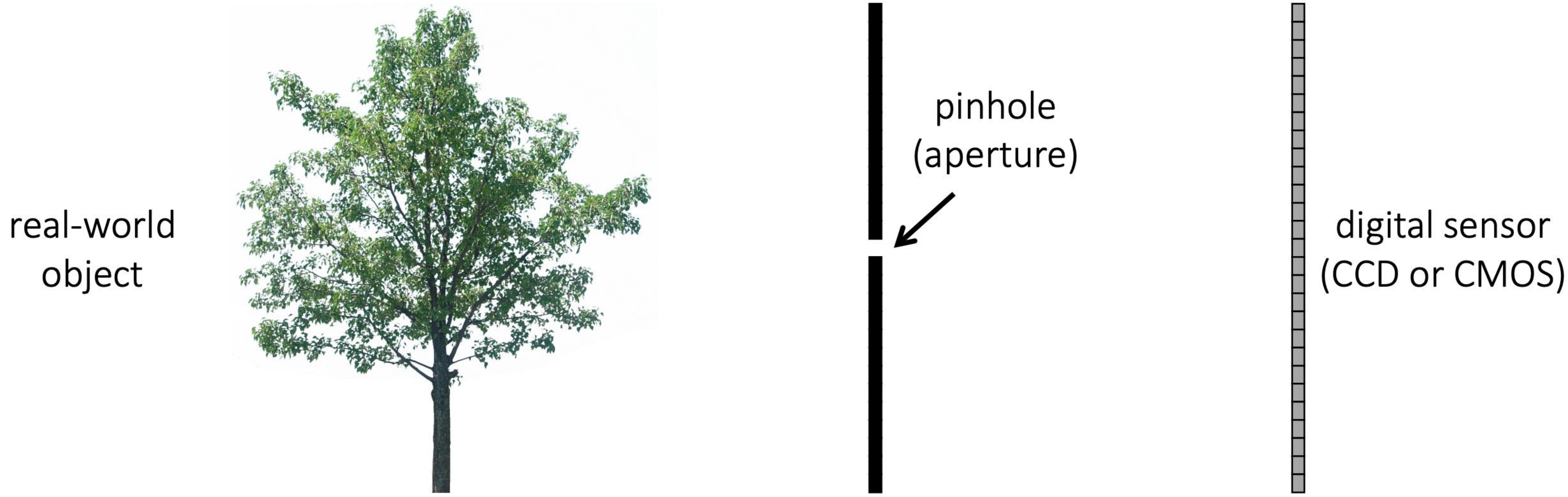

CCD
photon to electron conversion
CMOS
charge to voltage conversion


real-world
object
digital sensor
(CCD or CMOS)

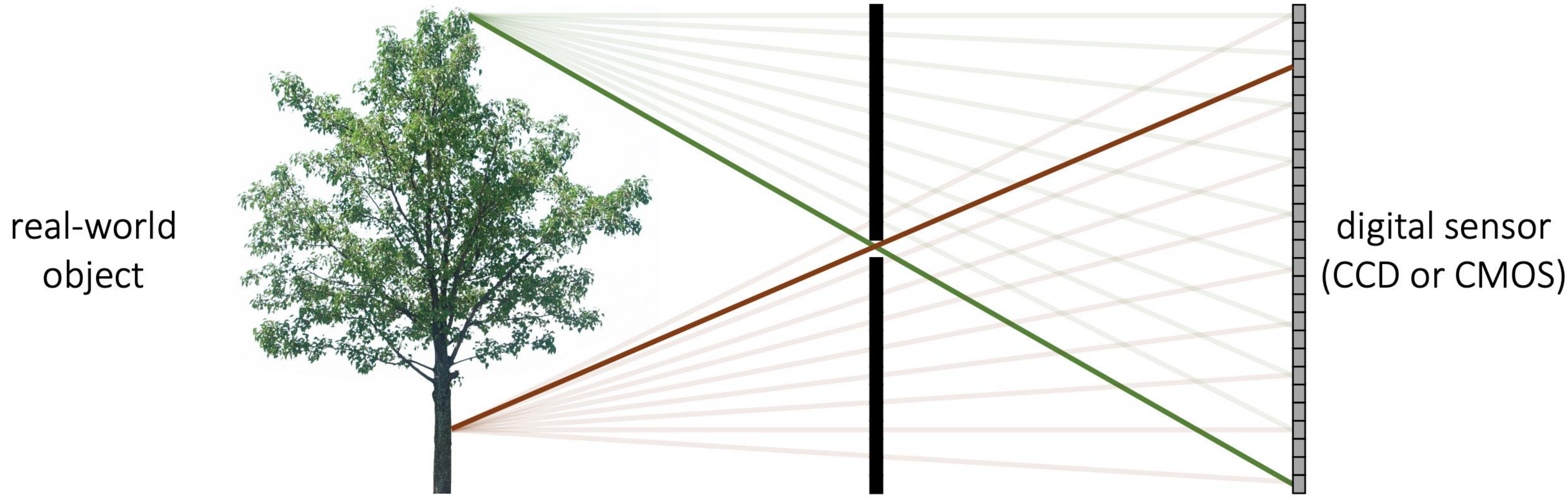


All scene points contribute to all sensor pixels

All scene points contribute to all sensor pixels



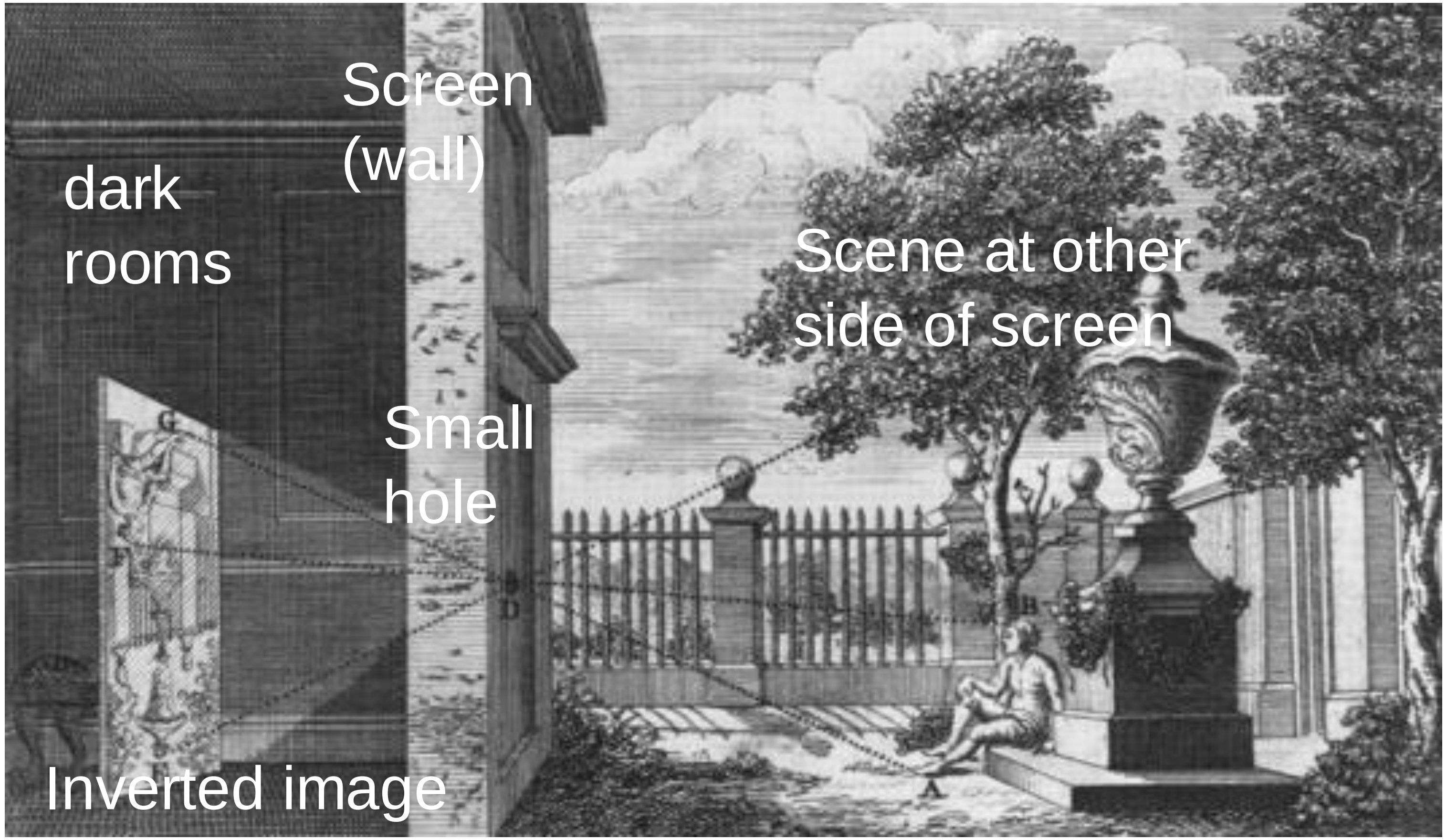
What would an image taken like this look like?



Each scene point contributes to only one sensor pixel


real-world
object
copy of real-world object
(inverted and scaled)

Pinhole Image



- Natural optical phenomenon
- Known back to 500 BC discovered by pioneers in China and Greece
- Used for art creation (16 century)
  - Leonardo da Vinci: first record of camera obscura

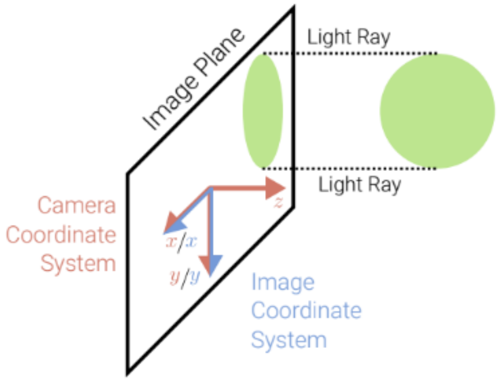
https://robertacucchiaro.wordpress.com/2012/10/24/abelardo-morells-camera-obscura-series/

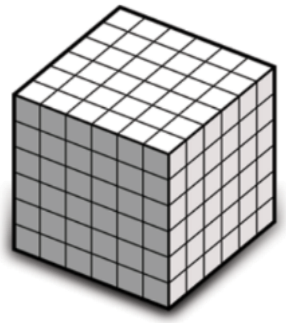
Any questions?

Geometric Image Formation

Projection Models

Orthographic Projection





Opto Engineering Telecentric Lens

Canon 800mm Telephoto Lens

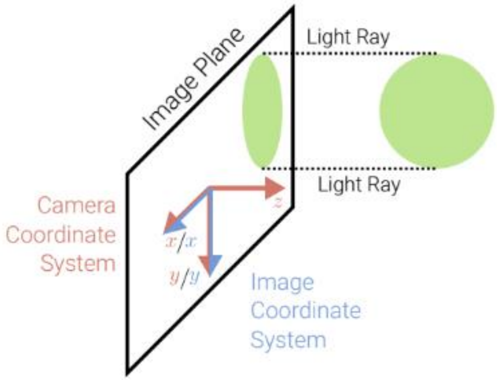
Pros:
- Accurage representation of dimensions
- Multiple views for clarity
- Useful in CAD and 3D modeling

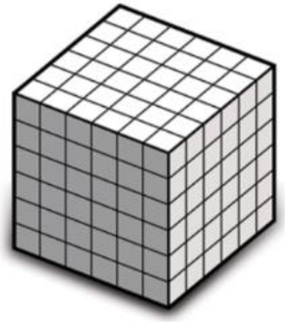
Cons:
- Lack of depth perception
- Require multiple views
- Ambiguity in Shape and Structure

Geometric Image Formation

Projection Models

Orthographic Projection

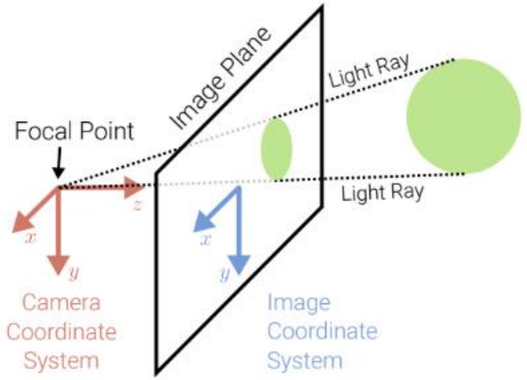


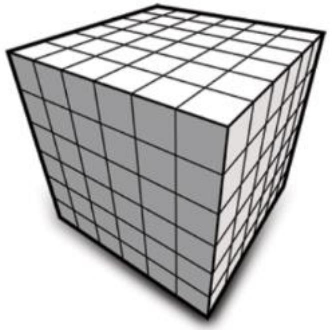


Opto Engineering Telecentric Lens

Canon 800mm Telephoto Lens

Perspective Projection

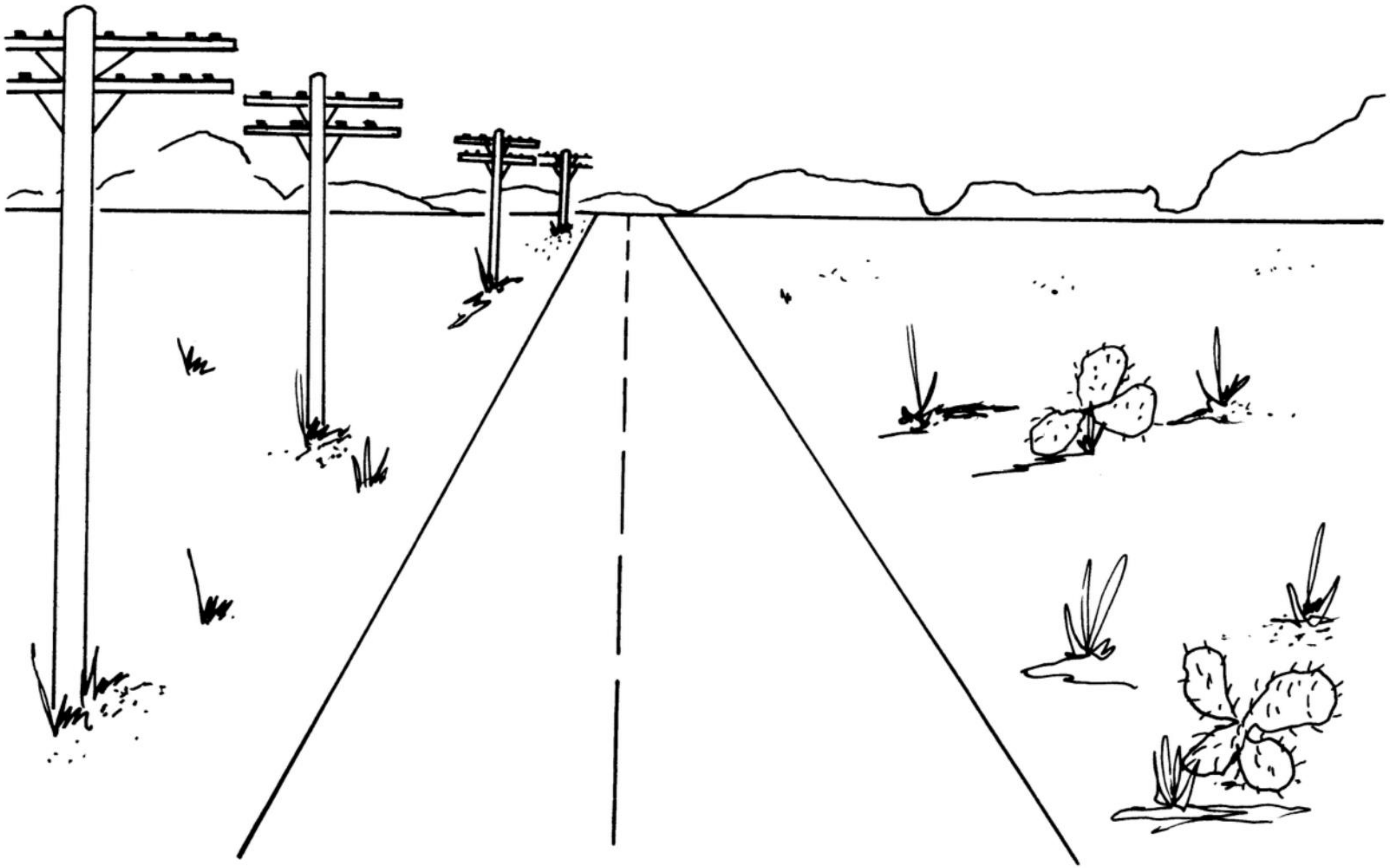


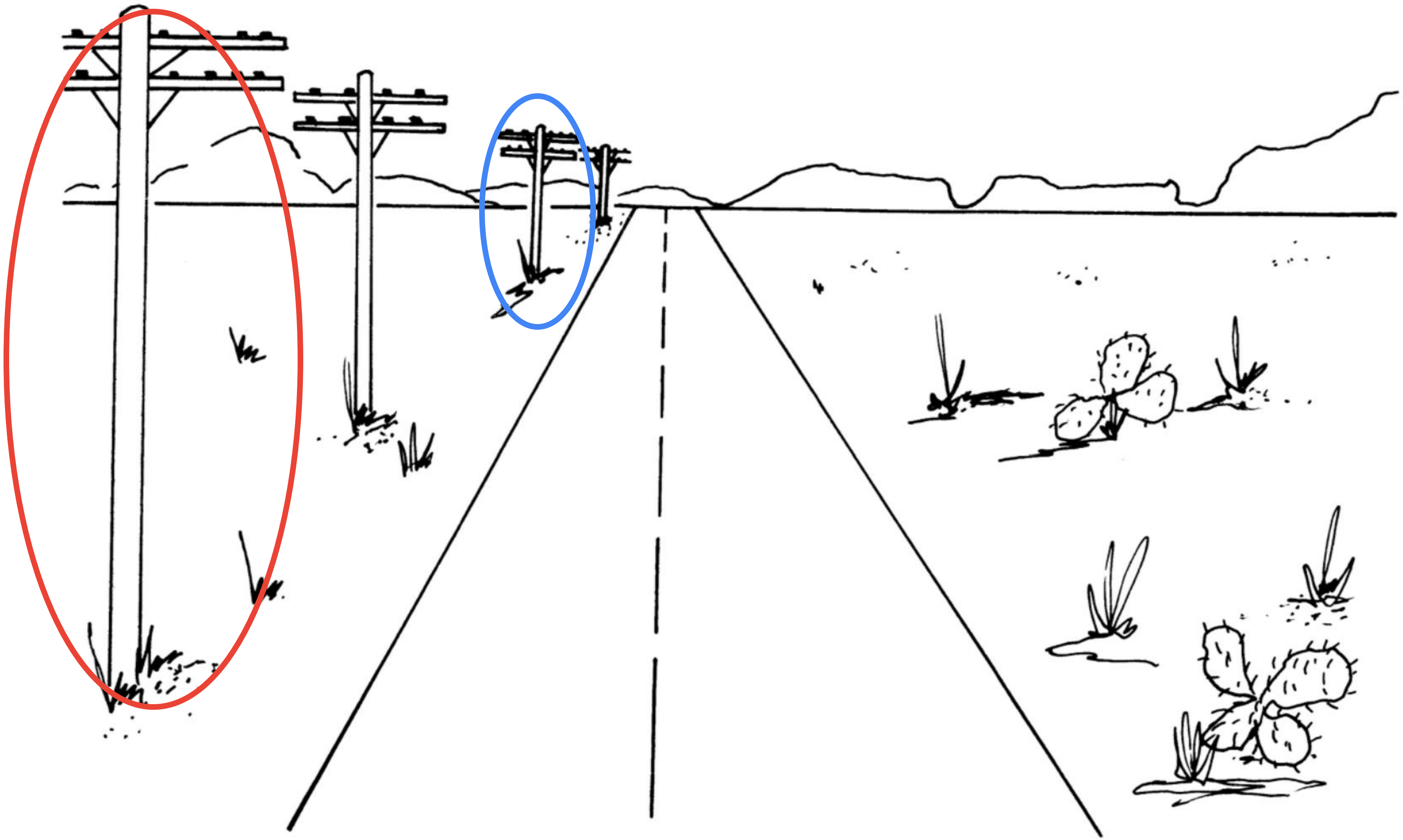


Nikon AF-S Nikkor 50mm

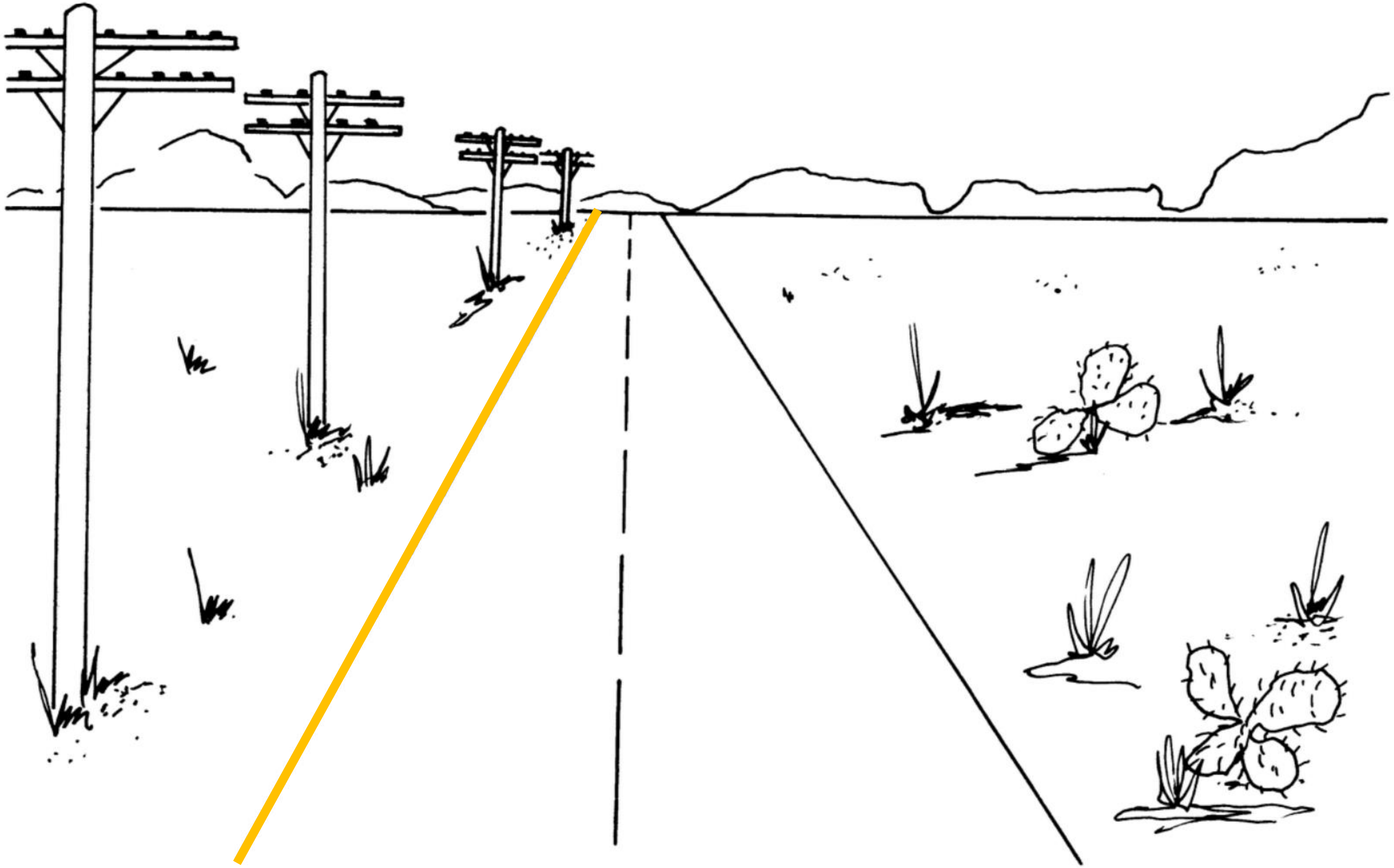
Sony DSC-RX100 V

Samsung Galaxy S20

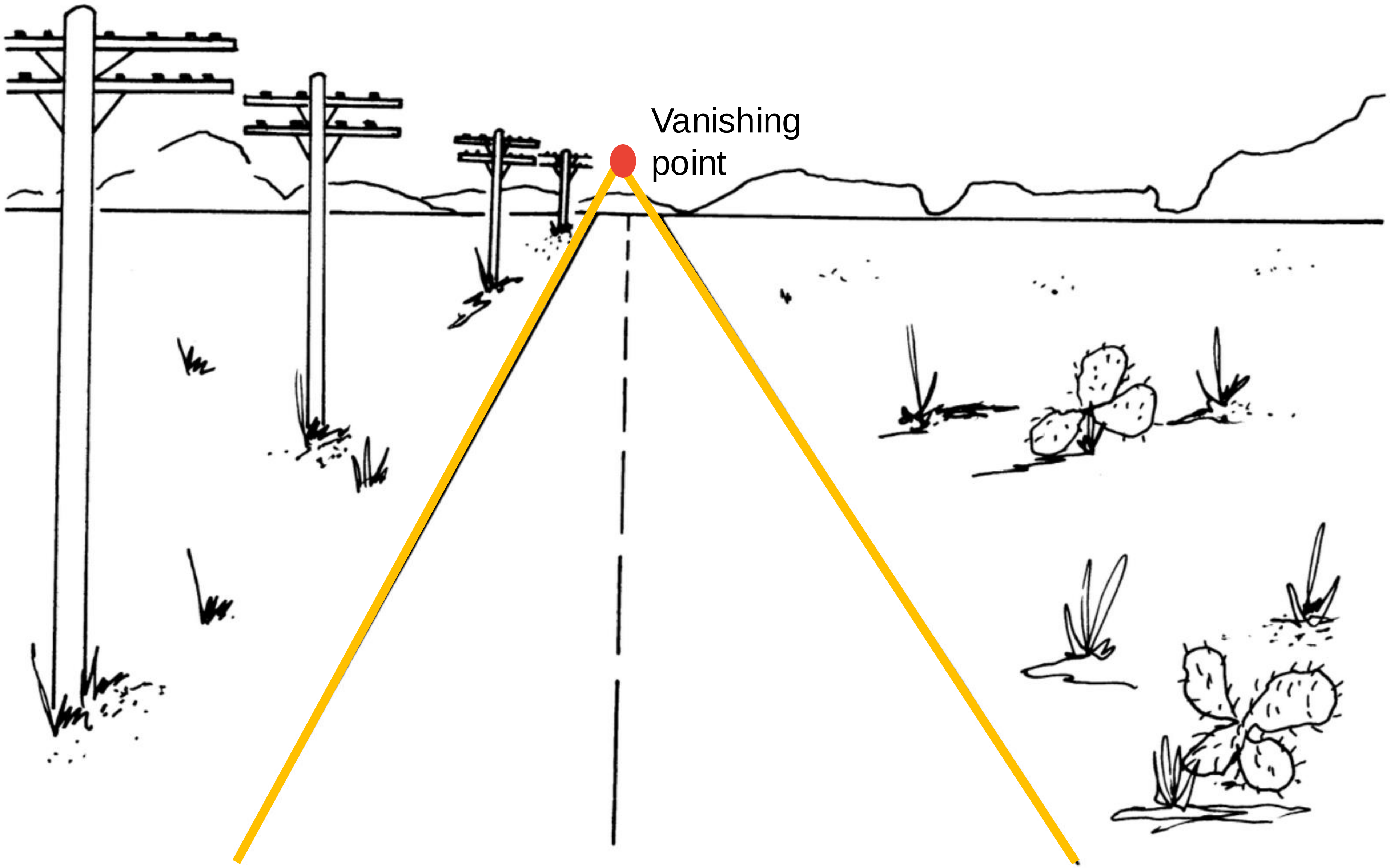




Size of object ~ 1/distance



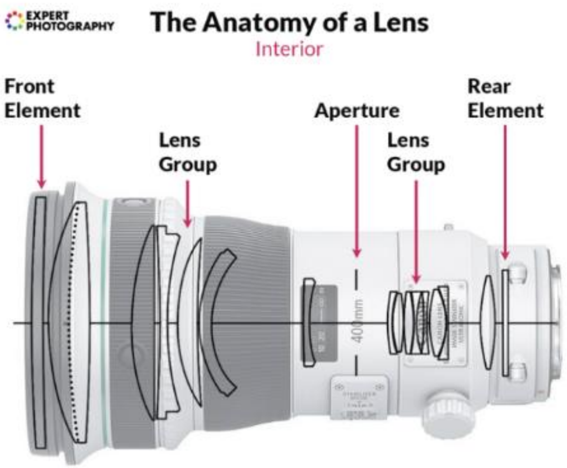
A 3D straight line remains a straight in 2D

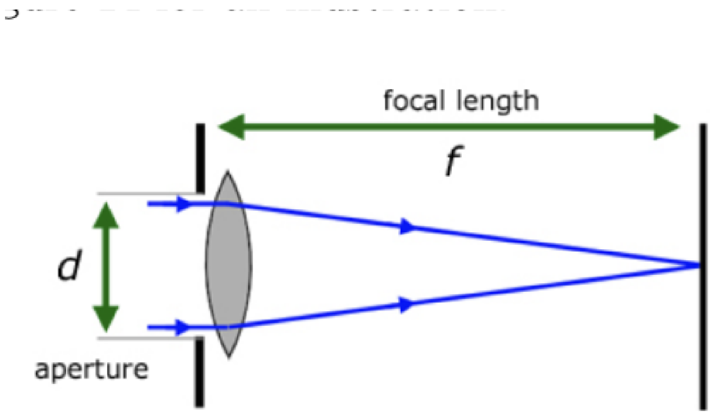


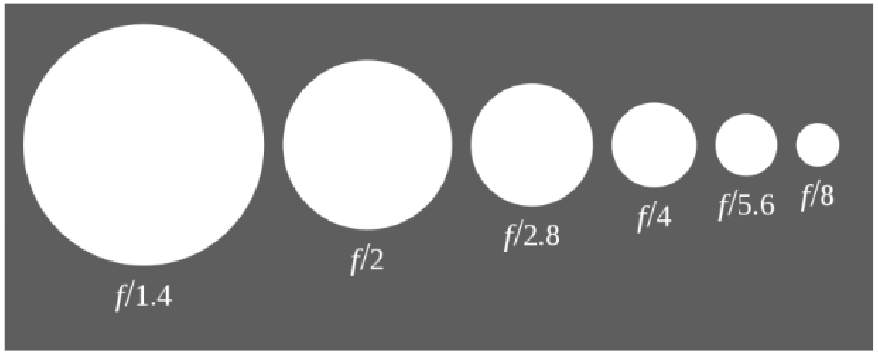
Parallel lines converge at vanishing points

F-number: How much light you can collect

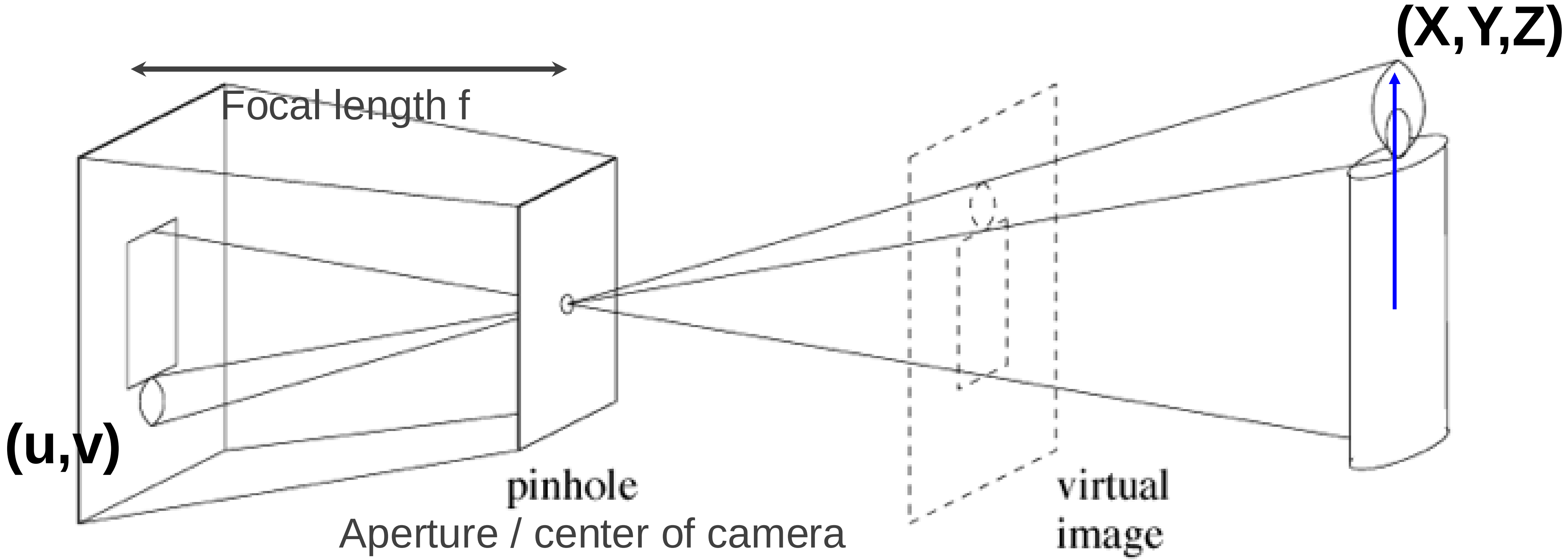
- Aperture: controls the amount of light that can reach the image sensor
- F-number is the ratio of the focal length to the diameter of the aperture





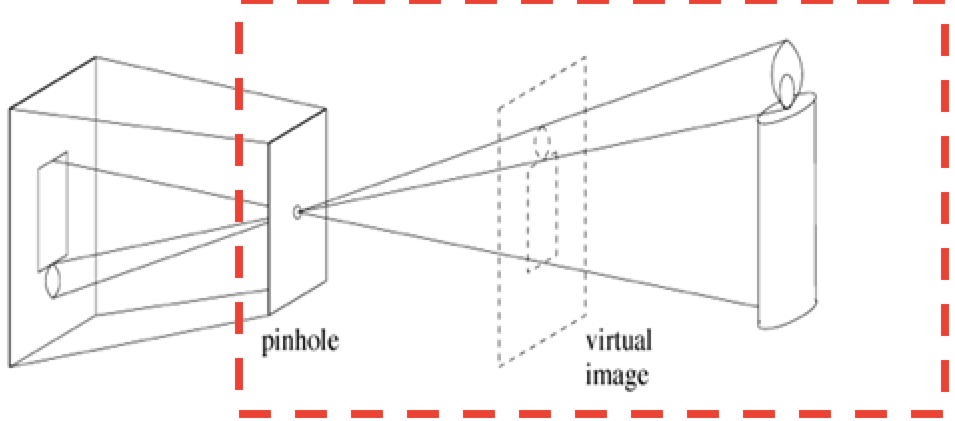


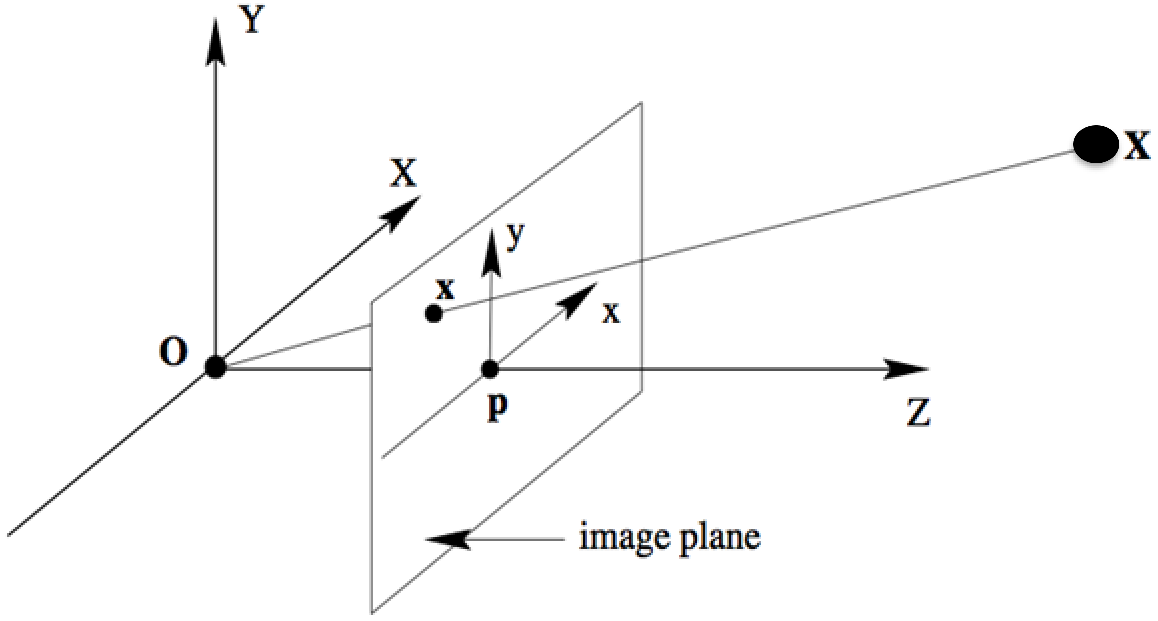
Pinhole Camera Model



How to compute a 2D pixel location (u,v) image from a 3D location (X,Y,Z)?

Perspective Projection



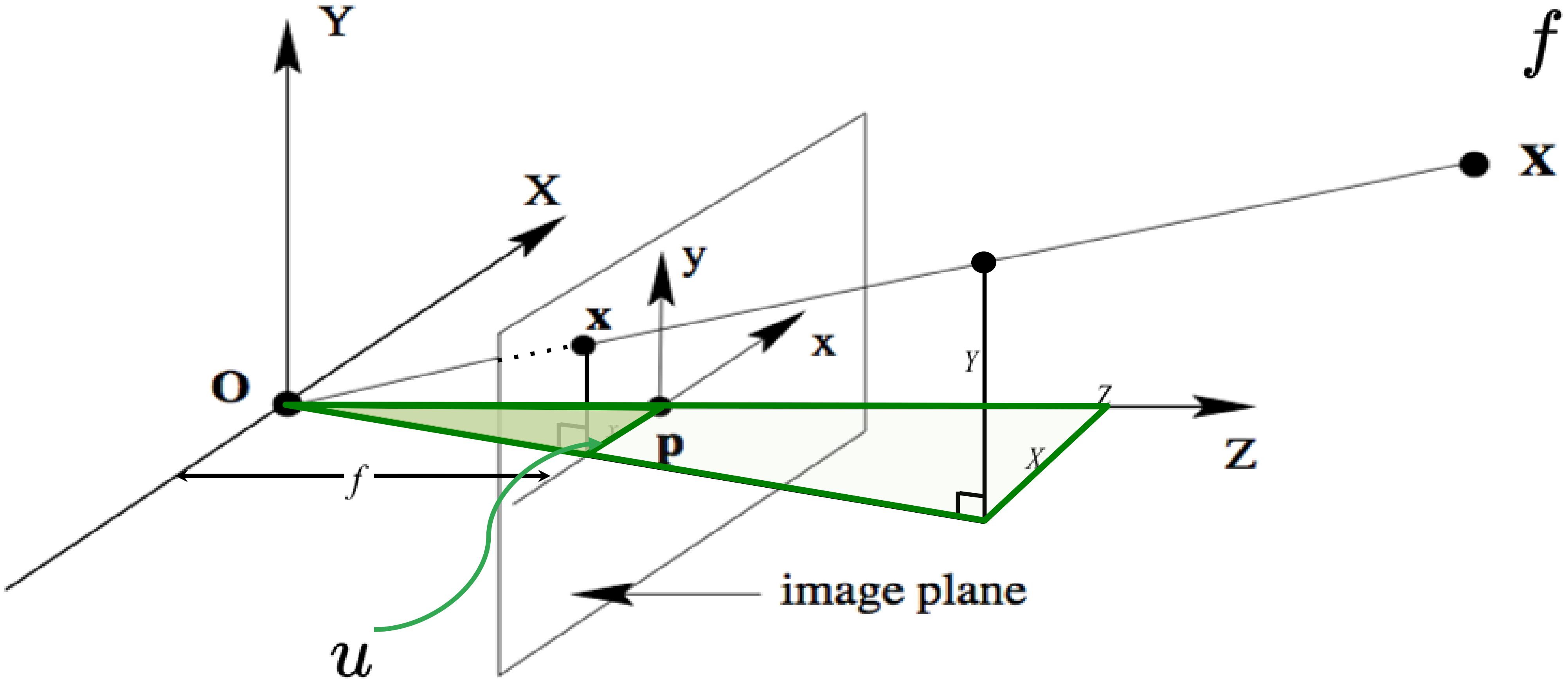


$$x = \begin{bmatrix} u \\ v \end{bmatrix}$$

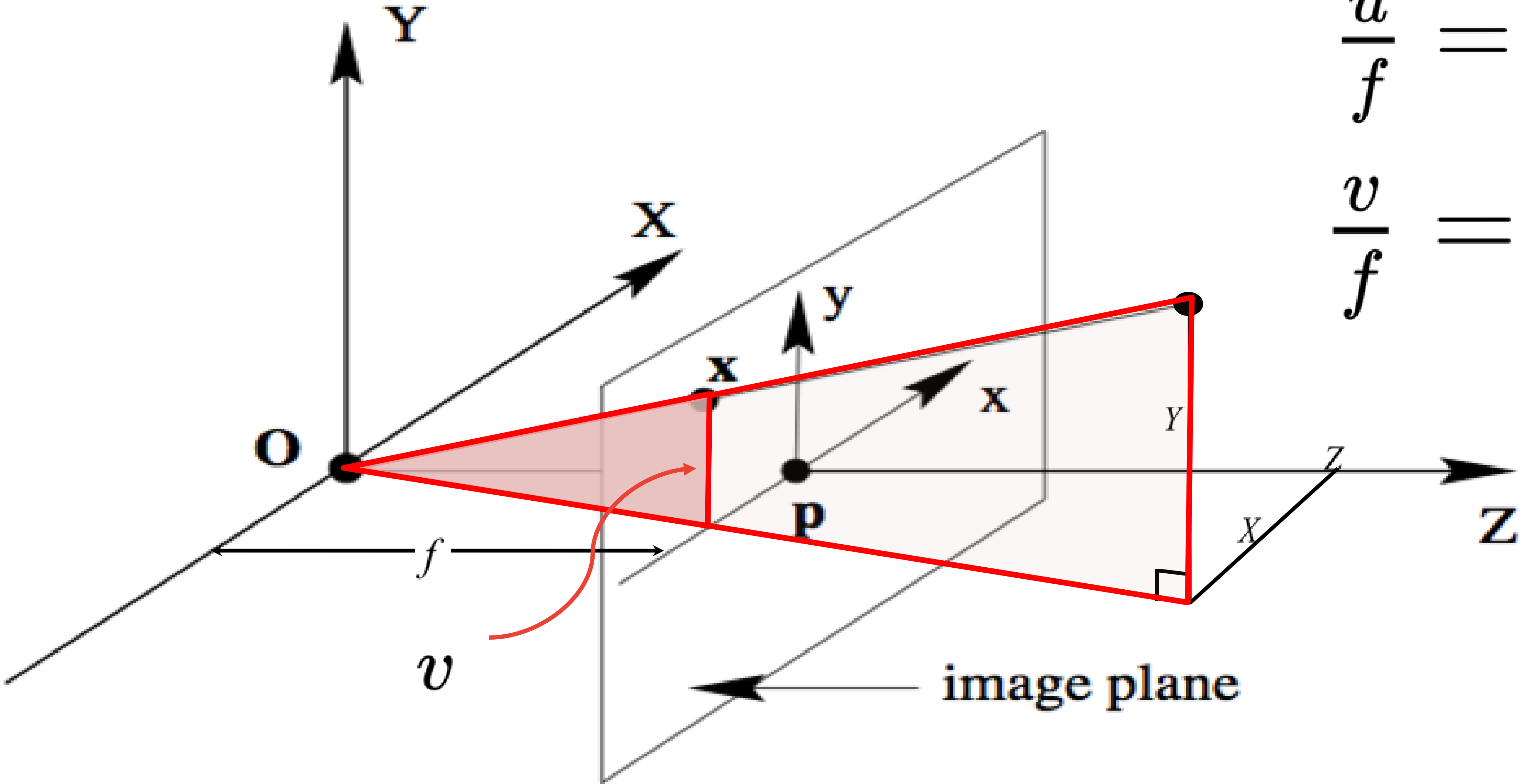
$$\mathbf{X} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

Perspective Projection

$$\frac{u}{f} = \frac{X}{Z}$$



Perspective Projection



$$\frac{u}{f} = \frac{X}{Z}$$

$$\frac{v}{f} = \frac{Y}{Z}$$

Pinhole Camera Model

$$(x, y) \Rightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \qquad \begin{bmatrix} x \\ y \\ w \end{bmatrix} \Rightarrow (x/w, y/w)$$

$$\frac{v}{f} = \frac{Y}{Z} \qquad \frac{u}{f} = \frac{X}{Z}$$

$$u = f\frac{X}{Z} \qquad v = f\frac{Y}{Z}$$

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} f\frac{X}{Z} \\ f\frac{Y}{Z} \end{bmatrix} \qquad \begin{bmatrix} f\frac{X}{Z} \\ f\frac{Y}{Z} \\ 1 \end{bmatrix} \qquad \begin{bmatrix} fX \\ fY \\ Z \end{bmatrix}$$

Non-linear!

Can we do something so that we can have the same transformation but without the division?

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

Homogeneous coordinate!

Homogeneous Coordinates

2D homogeneous coordinates

$$(x, y) \Rightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

3D homogeneous coordinates

$$(x, y, z) \Rightarrow \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

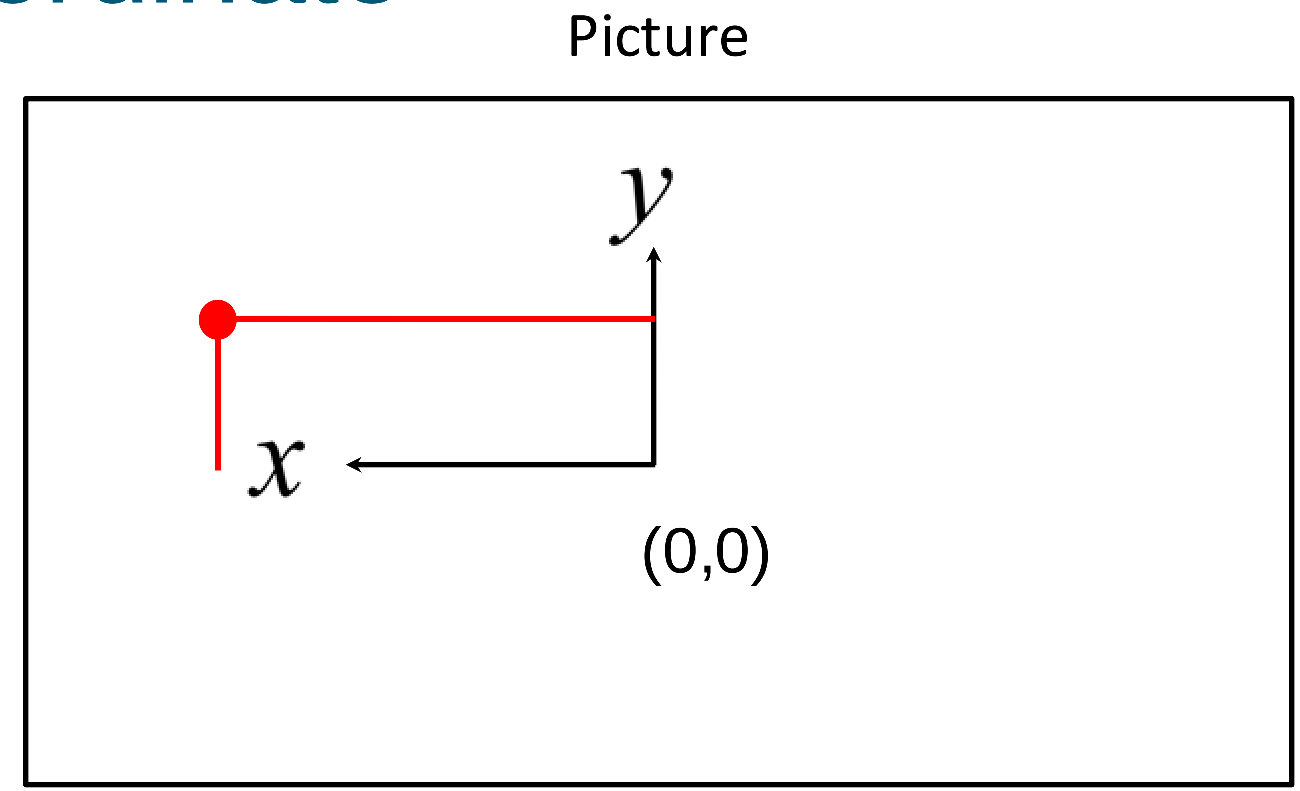
Converting *from* homogeneous coordinates

$$\begin{bmatrix} x \\ y \\ w \end{bmatrix} \Rightarrow (x/w, y/w)$$

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} \Rightarrow (x/w, y/w, z/w)$$

Image Coordinate

u = 300
v = 100



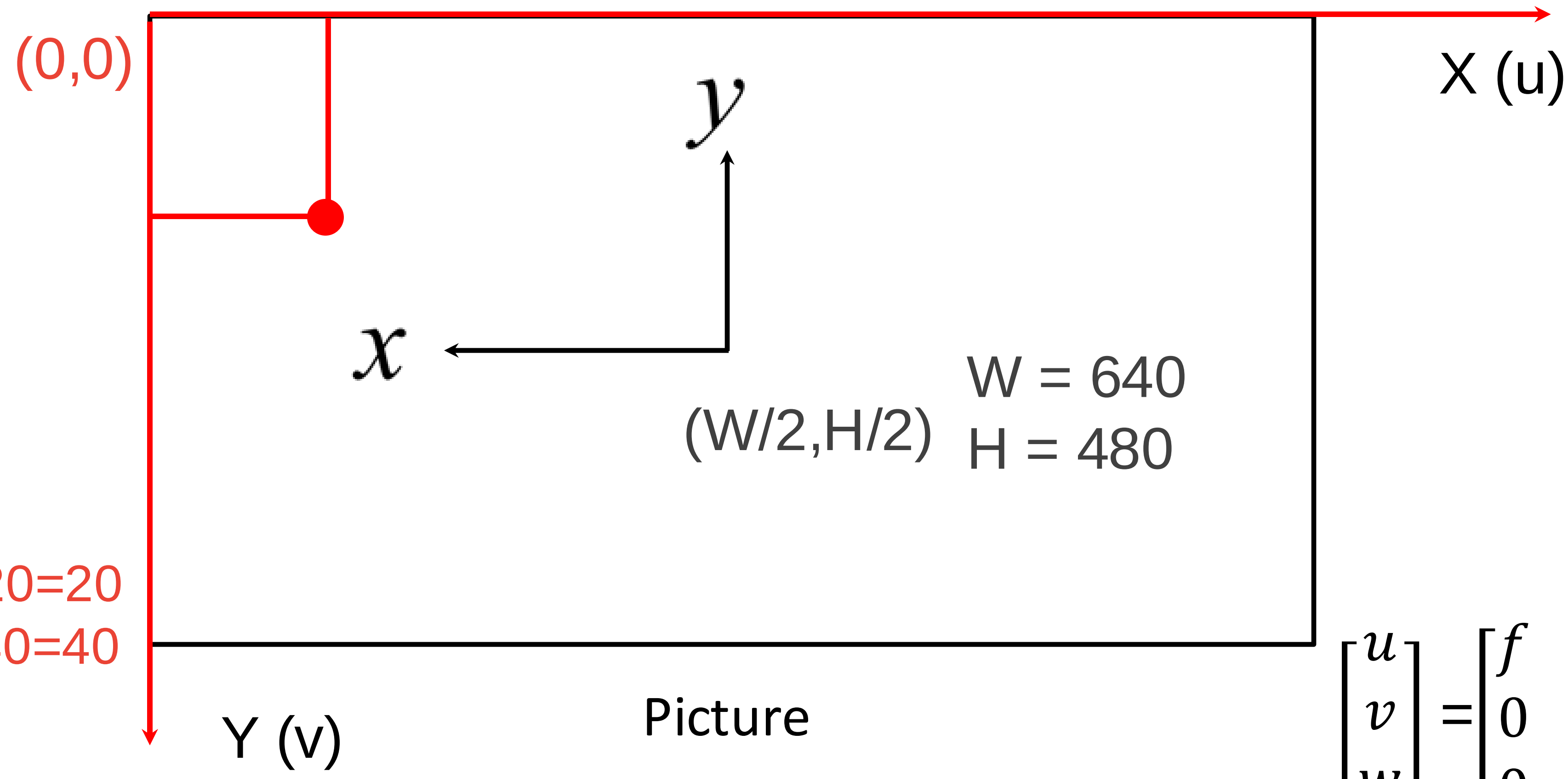
So far, we use 2D coordinate conventions that are consistent with the 3D Camera coordinate. If your application uses a different 2D coordinate, you'll need to further transform the (u,v)

Image Coordinate (Example 1)

Cols indexed by u
Rows indexed by v

u = 300
v = 100

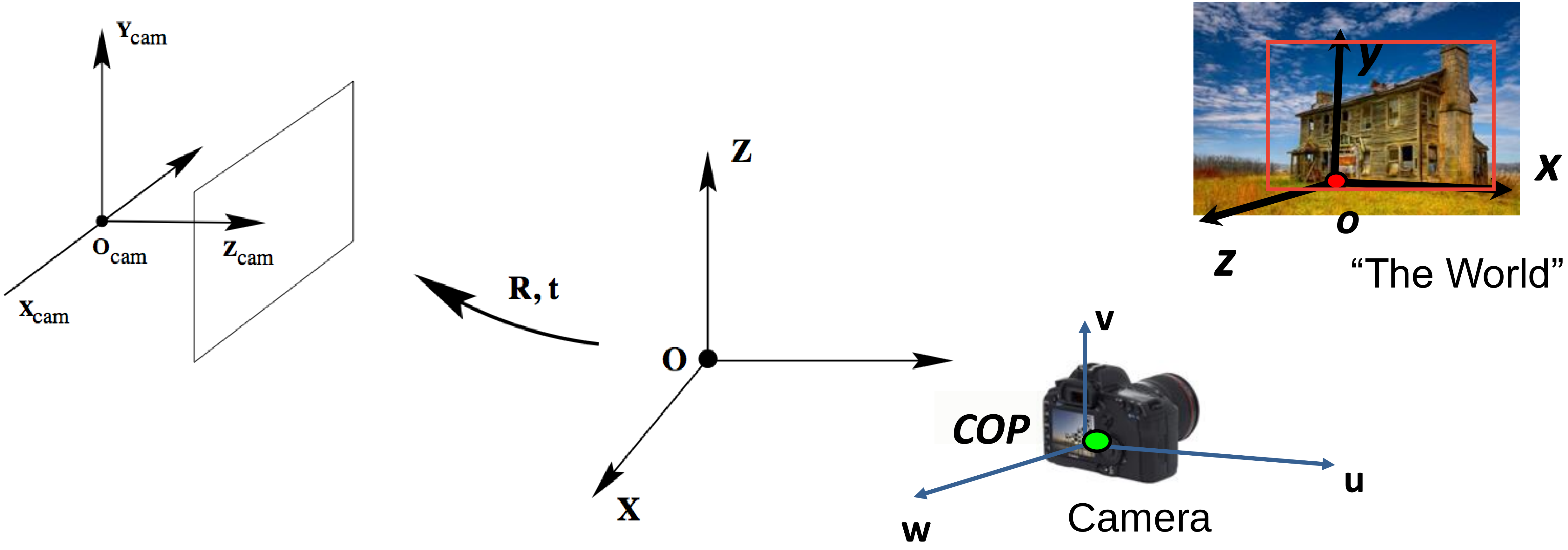
u’ = -1*300+320=20
v’ = -1*100+240=40



$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} f & 0 & u_0 \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

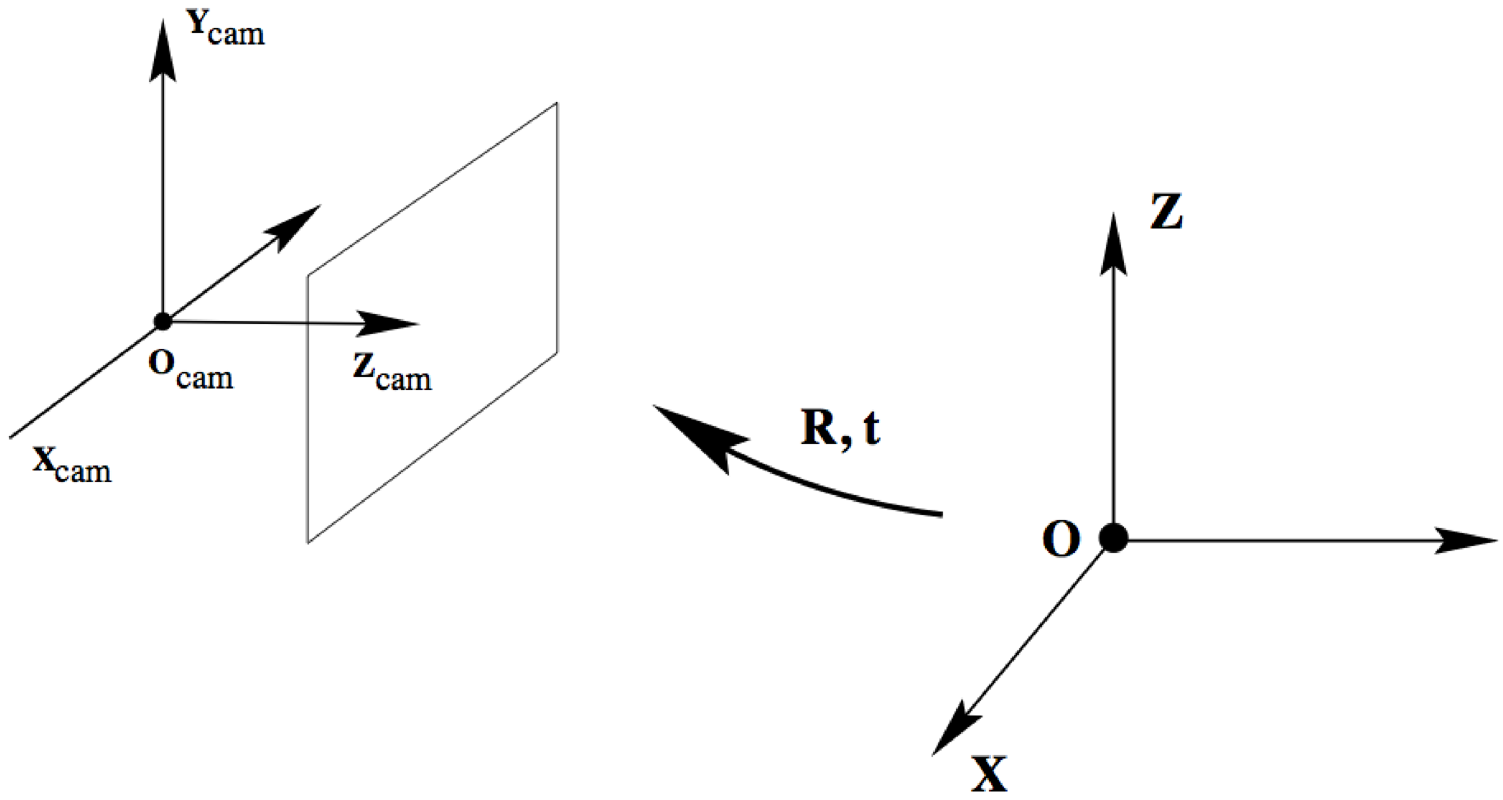
Image Coordinate Convention used in Computer Vision

World Coordinate to Camera Coordinate



To apply the pinhole model, objects in the scene must be expressed in camera coordinates

World Coordinate to Camera Coordinate



How to representation this transformation?

$^{?}T_{?}$ $\quad$ $^{Camera}T_{world}$

$$\begin{pmatrix} X_{cam} \\ Y_{cam} \\ Z_{cam} \\ 1 \end{pmatrix} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Change coordinate frames from world frame to camera frame

Camera Projection Matrix

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \underbrace{\underbrace{\mathbf{K}}_{\text{intrinsics}} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{\text{projection}} \underbrace{\underbrace{\left[\begin{array}{c|c} I & \boldsymbol{t} \\ \hline \mathbf{0} & 1 \end{array}\right]}_{\text{translation}} \times \underbrace{\left[\begin{array}{c|c} R & \mathbf{0} \\ \hline \mathbf{0} & 1 \end{array}\right]}_{\text{rotation}}}_{\text{Camera Extrinsic Matrix [R|t]}} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

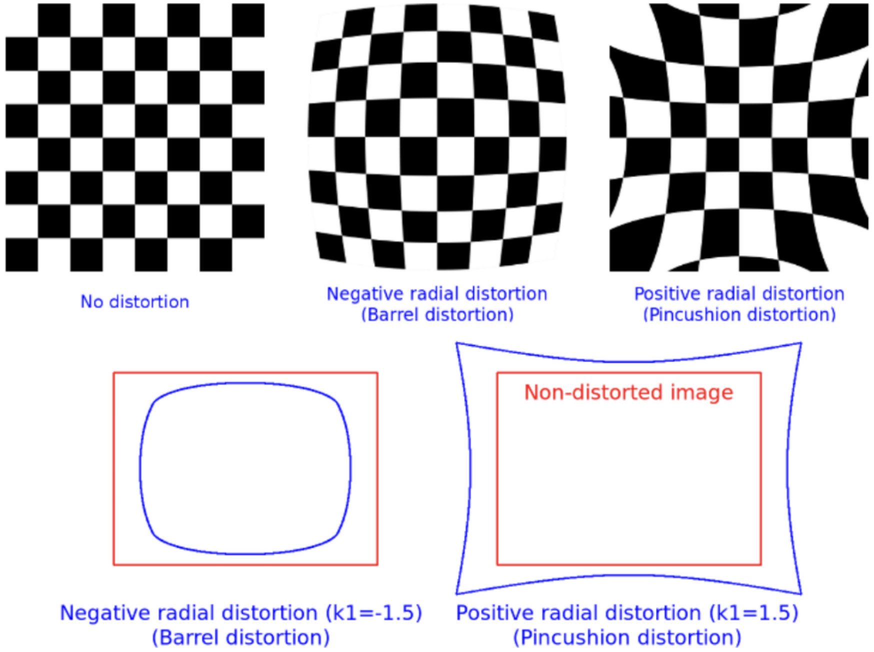
$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} f & 0 & u_0 \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

Camera Projection Matrix **P**

$$\mathbf{x} = \mathbf{P}\mathbf{X}$$

Lens Distortion

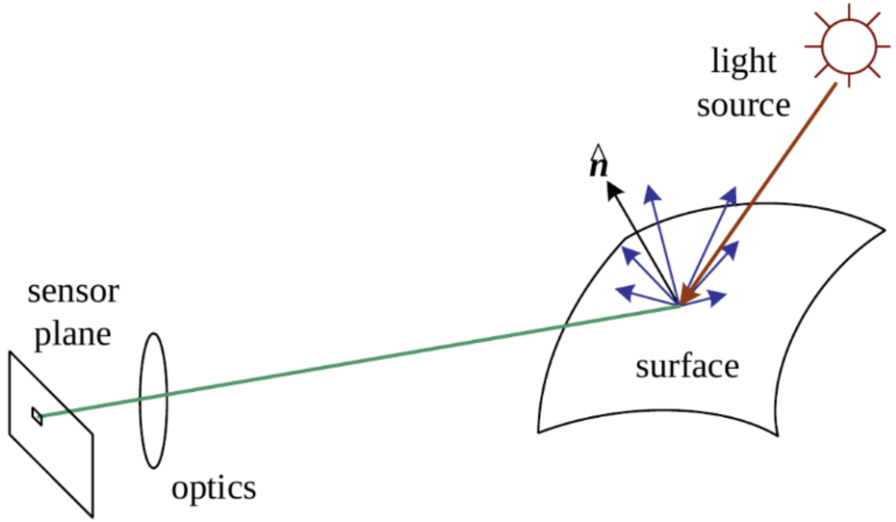
- The assumption of linear projection (straight lines remain straight) is violated in practice due to the properties of the camera lens which introduces distortions.



Any questions?

Photometric Image Formation

- Concern about how an image is formed in terms of **pixel intensities** and **colors**
- Light is emitted by one or more light sources and reflected or refracted (once or multiple times) at surfaces of objects (or media) in the scene

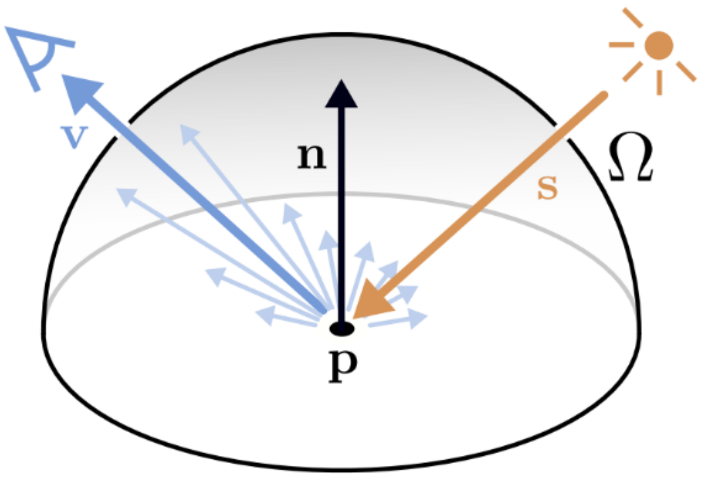


Rendering Equation

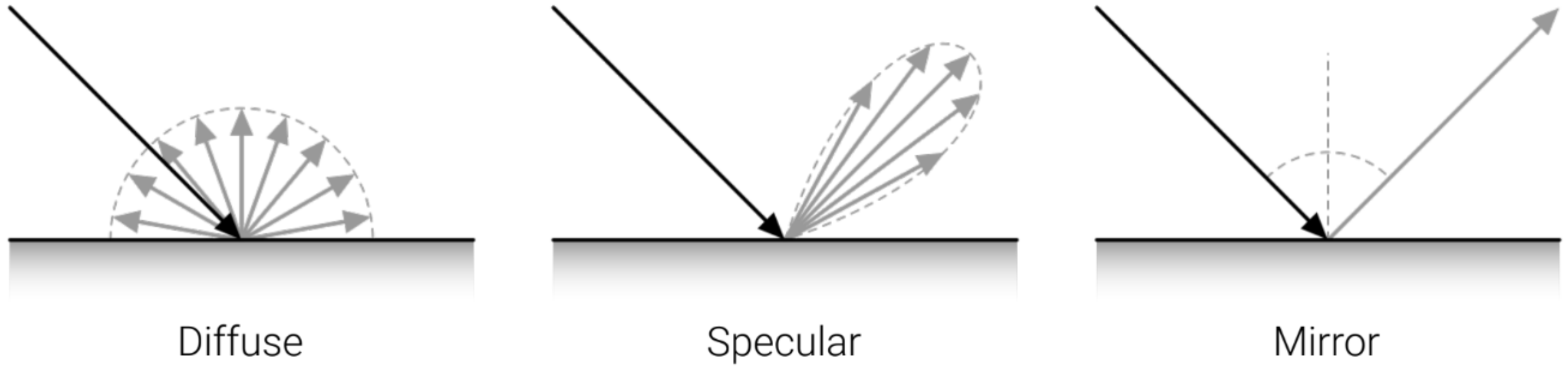
Let $\mathbf{p} \in \mathbb{R}^3$ denote a 3D surface point, $\mathbf{v} \in \mathbb{R}^3$ the viewing direction and $\mathbf{s} \in \mathbb{R}^3$ the incoming light direction. The **rendering equation** describes how much of the light L_{in} with wavelength λ arriving at $\mathbf{p}$ is reflected into the viewing direction $\mathbf{v}$:

$$L_{\text{out}}(\mathbf{p}, \mathbf{v}, \lambda) = L_{\text{emit}}(\mathbf{p}, \mathbf{v}, \lambda) + \int_{\Omega} \text{BRDF}(\mathbf{p}, \mathbf{s}, \mathbf{v}, \lambda) \cdot L_{\text{in}}(\mathbf{p}, \mathbf{s}, \lambda) \cdot (-\mathbf{n}^\top \mathbf{s})\, d\mathbf{s}$$

- Ω is the unit hemisphere at normal $\mathbf{n}$
- The bidirectional reflectance distribution function $\text{BRDF}(\mathbf{p}, \mathbf{s}, \mathbf{v}, \lambda)$ defines how light is reflected at an opaque surface.
- $L_{\text{emit}} > 0$ only for light emitting surfaces

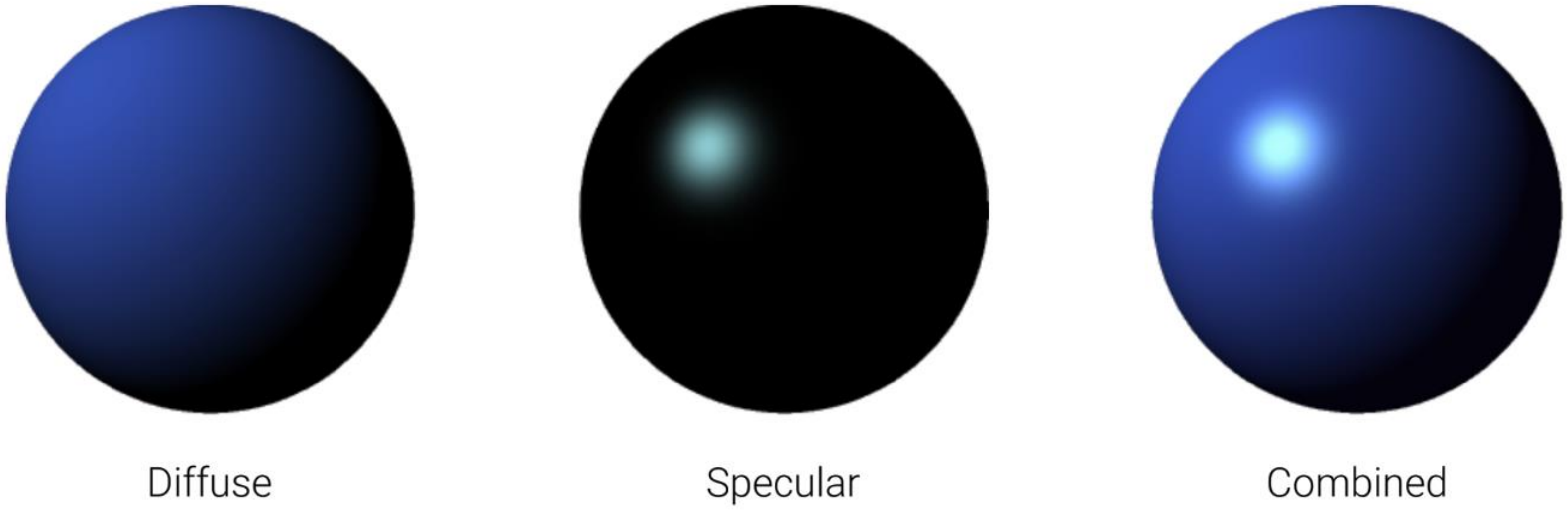


Diffuse and Specular Reflection



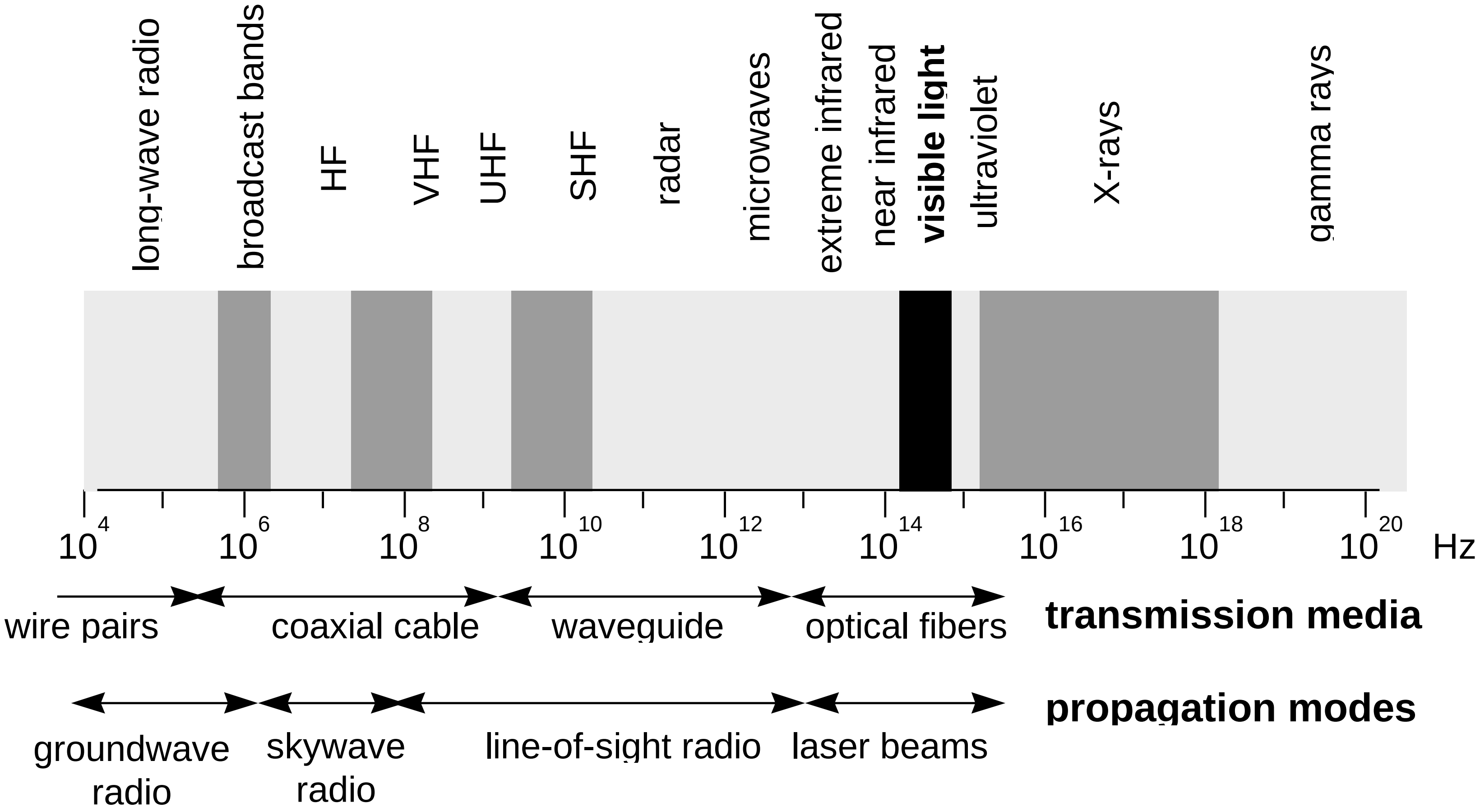
- Typical BRDFs have a **diffuse** and a **specular** component
- The diffuse (=constant) component scatters light uniformly in all directions
- This leads to shading, i.e., smooth variation of intensity wrt. surface normal
- The specular component depends strongly on the outgoing light direction

Diffuse and Specular Reflection

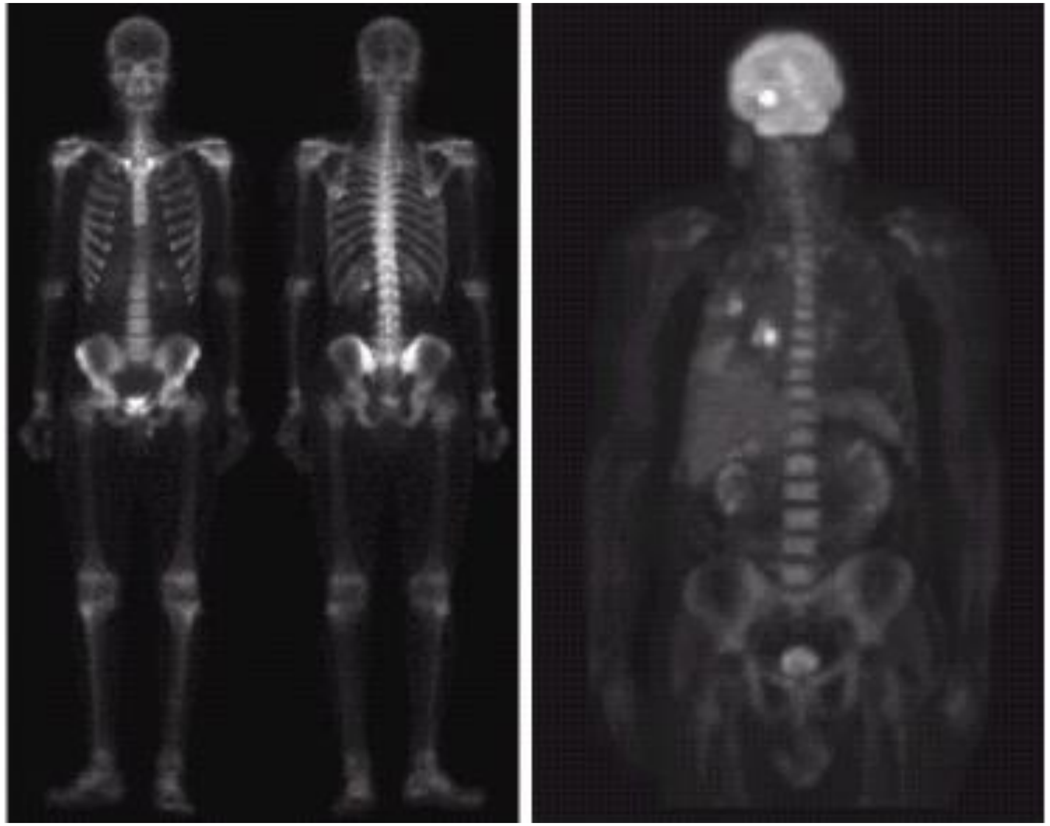


- Typical BRDFs have a **diffuse** and a **specular** component
- The diffuse (=constant) component scatters light uniformly in all directions
- This leads to shading, i.e., smooth variation of intensity wrt. surface normal
- The specular component depends strongly on the outgoing light direction

Electromagnetic Spectrum



X-rays



Gamma-Ray Imaging

X-ray Imaging

Visible Light

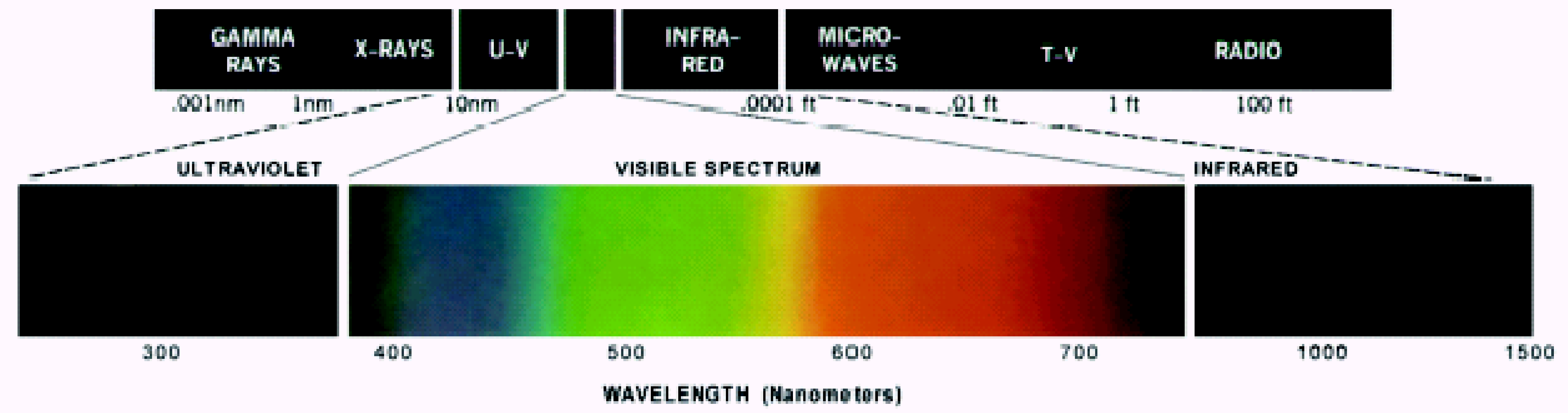


FIGURE 6.2 Wavelengths comprising the visible range of the electromagnetic spectrum. (Courtesy of the General Electric Co., Lamp Business Division.)

What is color?

Color is a sensation experienced by a human being in response to a visual stimulus.

What is color?

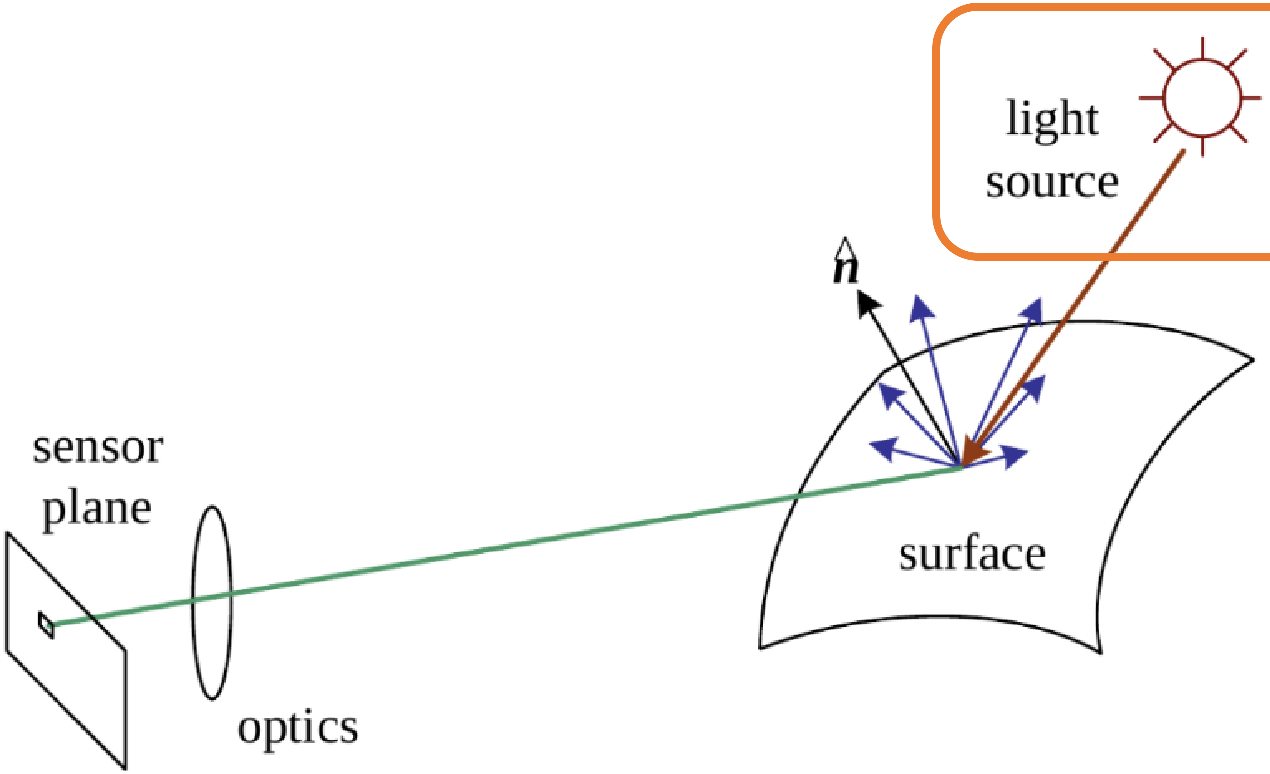
Color is a sensation experienced by a human being in response to a visual stimulus.

Sensation

- To understand ”image,” we must understand how humans perceive visual stimulus
  - Physiology: Measure chemical, electrical, or magnetic changes
    - Examples: neural probes, functional magnetic resonance
  - Psychophysics: Measure cognitive response
    - “I can see A”(detection)
    - “A matches B”
    - “A looks brighter than B”(discrimination)

Visual Stimulus

Model of light source:
We will talk about it next week



Any questions?

Image Sensing Pipeline

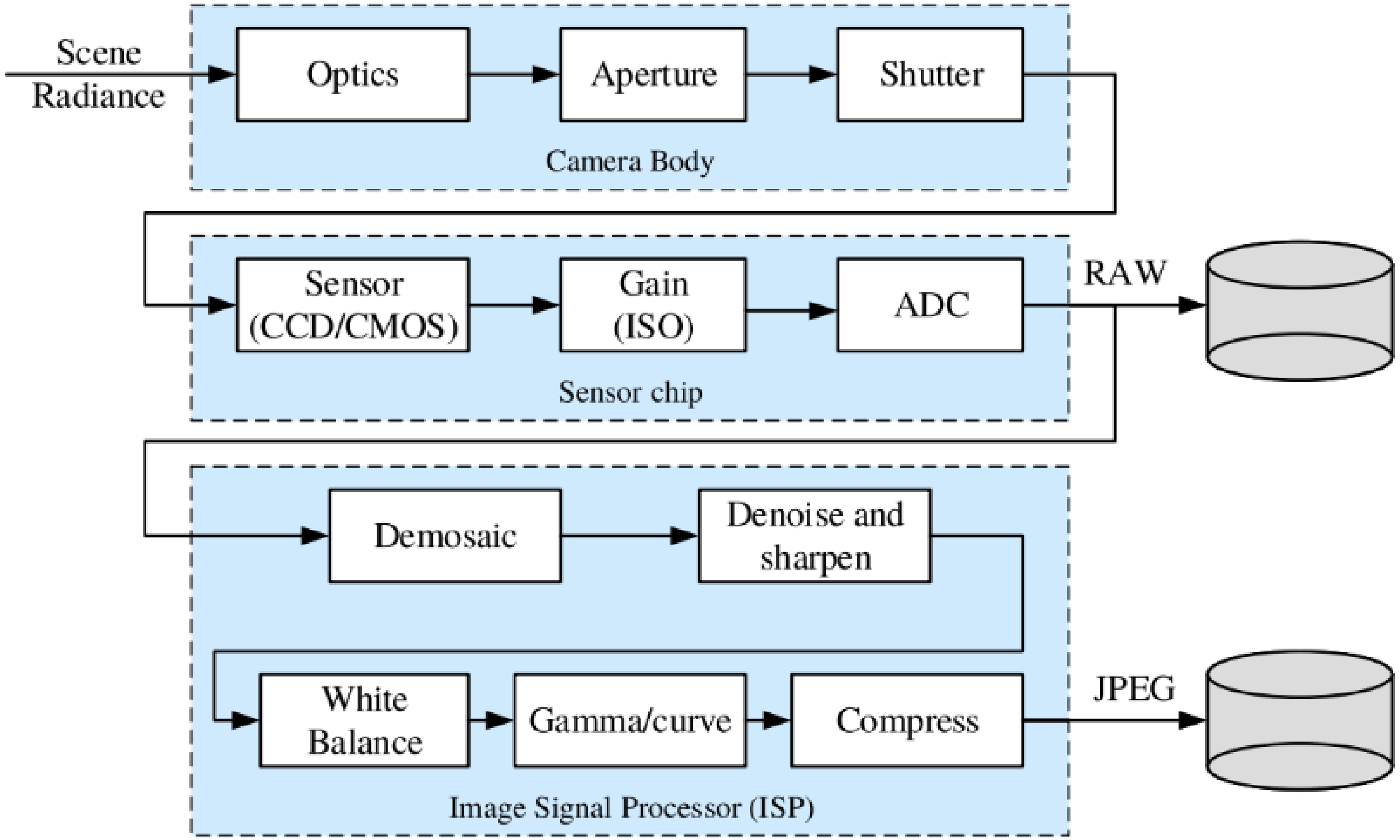
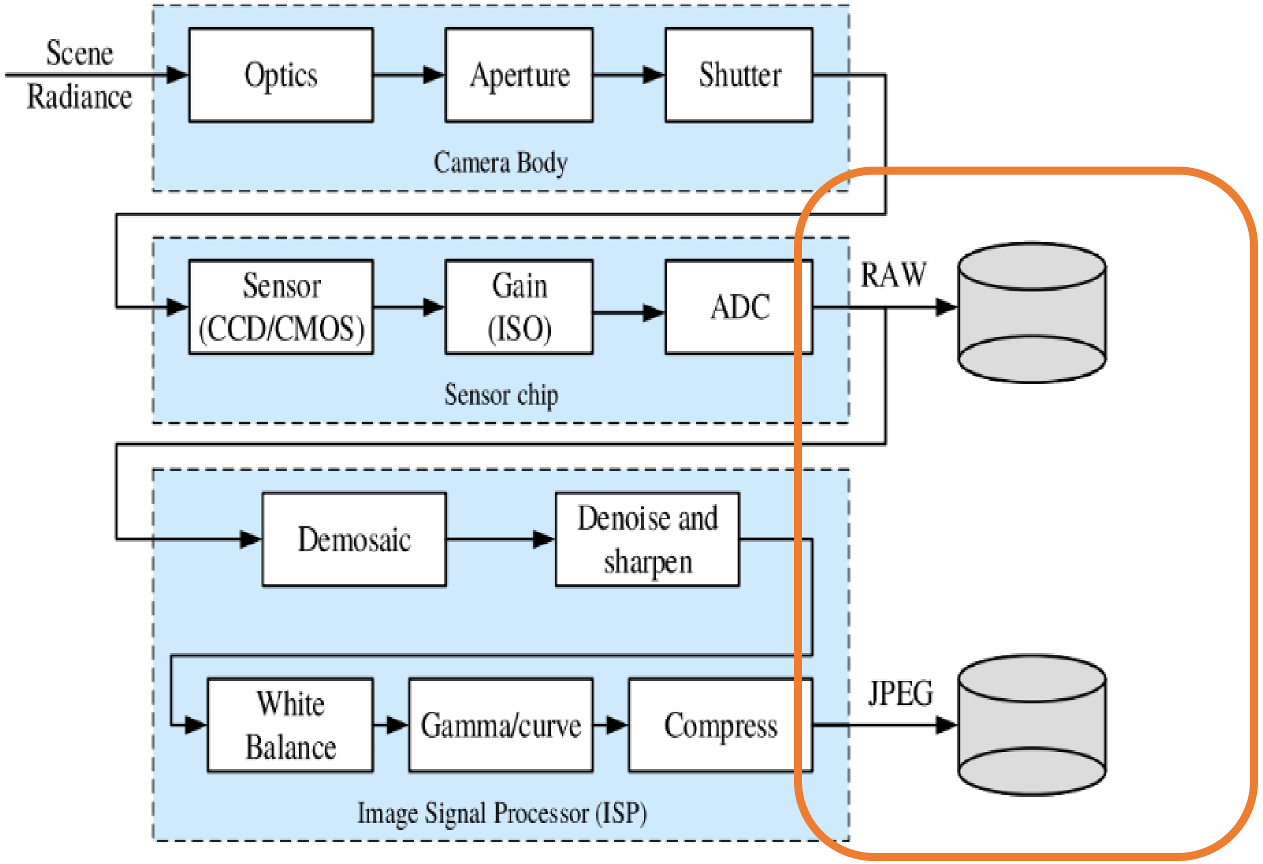


Image File Format

Name	Type	Usage
Tagged image file format	*.TIF	DOS, UNIX, and Macintosh images
Graphical interchange format	*.GIF	CompuServe graphics format
Bit-mapped format	*.BMP	Microsoft Windows format
PNG	*.png	Portable Network Graphics
JPEG/JFIF	*.jpg	Compressed images
MPEG	*.mpg	Compressed motion images

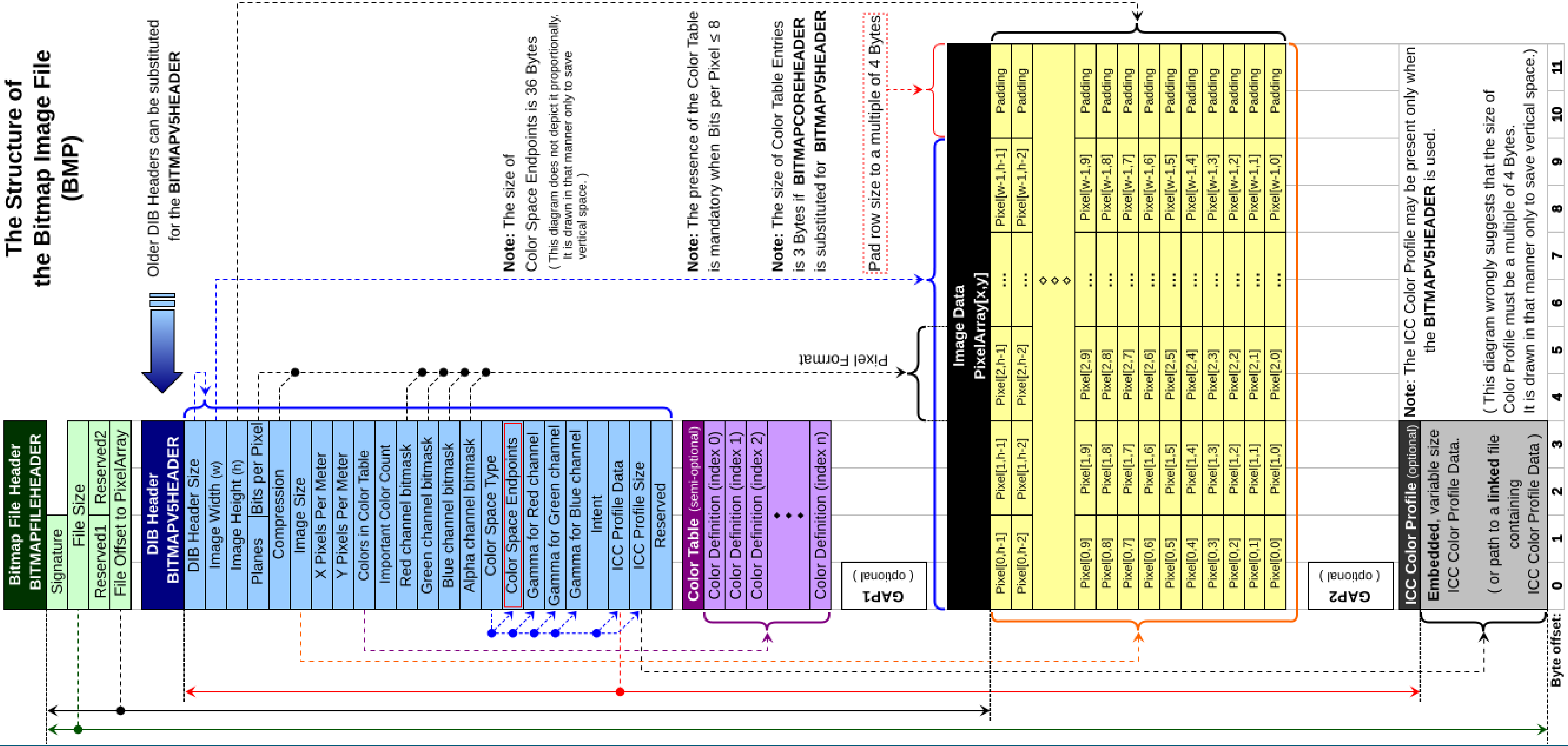


Example:
The BMP format is an uncompressed raster file designed to display high-quality images on Windows and store printable photos.

- **BMP File Header** Stores general information about the BMP file.
- **Bitmap Information** (DIB header) Stores detailed information about the bitmap image.
- **Color Palette** Stores the definition of the colors being used for indexed color bitmaps.
- **Bitmap Data** Stores the actual image, pixel by pixel

https://en.wikipedia.org/wiki/Image_file_formats

The BMP Format



https://en.wikipedia.org/wiki/BMP_file_format

Image Sensing Pipeline

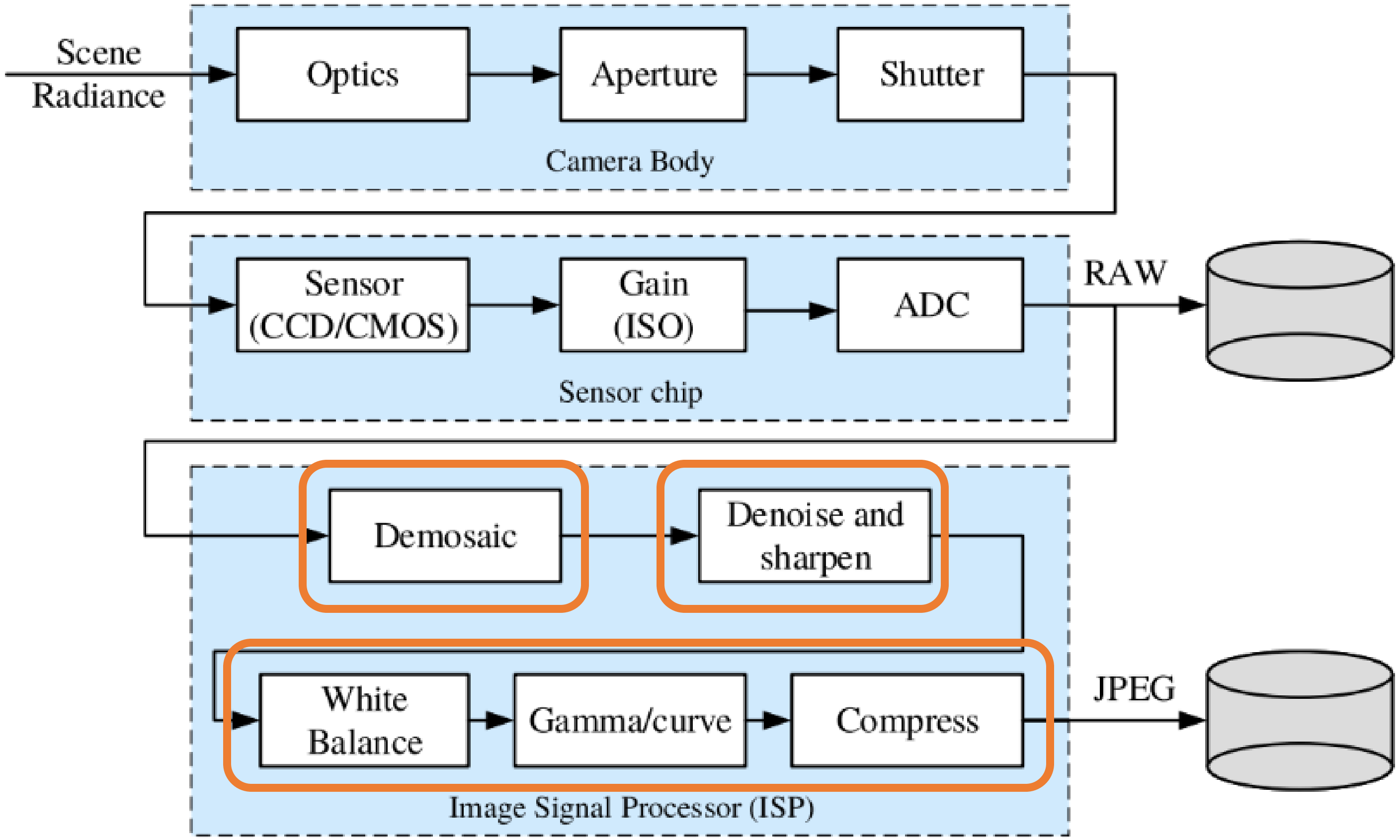


Image Sensing Pipeline + Image Display

Image Capturing Device

Interconnection Space or Processing

Image Display Device

Cameras

Scanners

Smart phones

Tablets

Personal Computer

Cloud

Projectors

Printers

Quick Summary

- Inside a camera
- Image formation
  - Geometric
  - Photometric
- What is color?
  - Will continue to discuss this topic next week