Image Formation

CSE P576

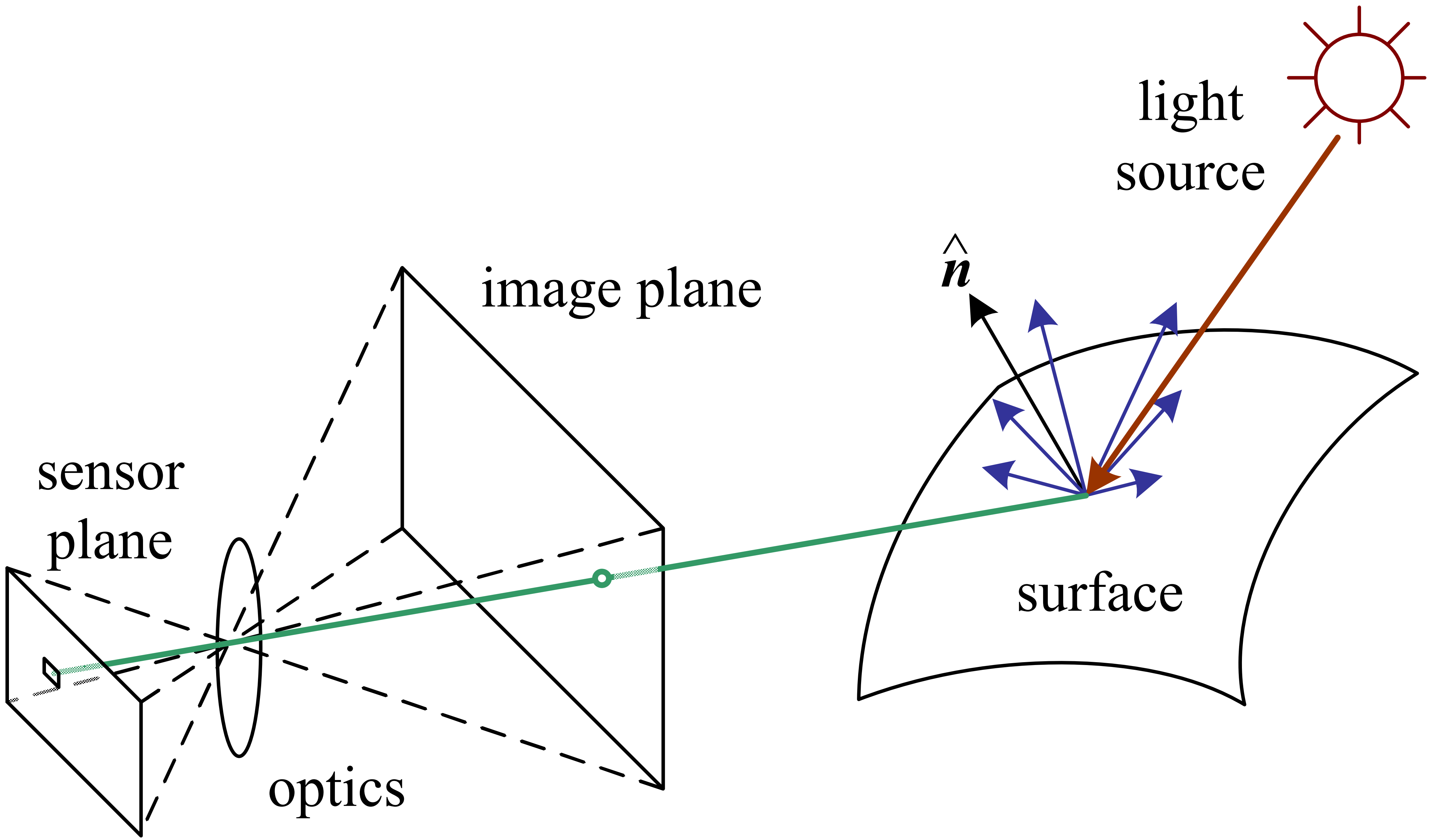
Dr. Matthew Brown

Image Formation

- Light, Optics, Sensing
- The Digital Camera

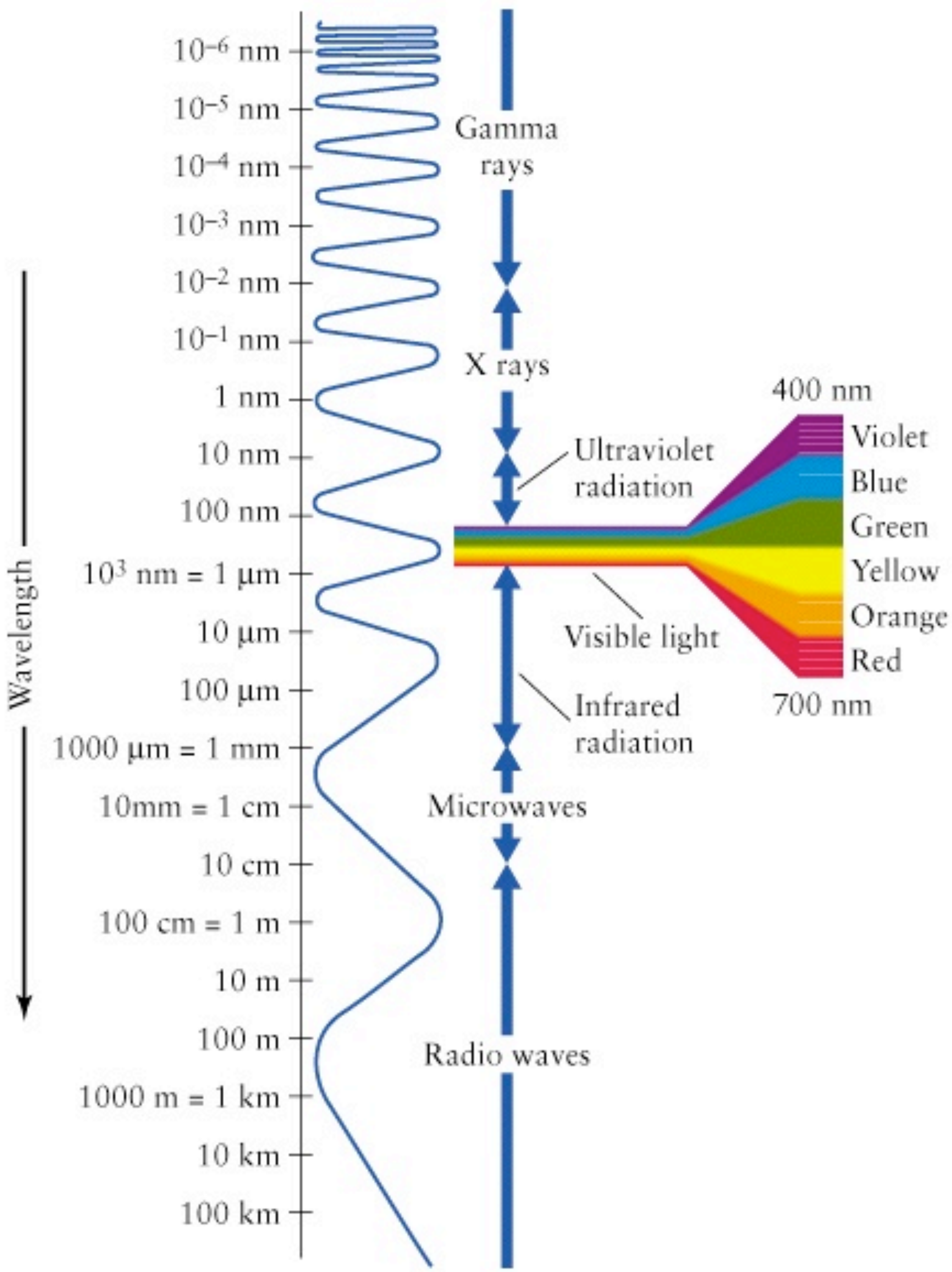
[Szeliski Chapter 2]

Image Formation



[Szeliski]

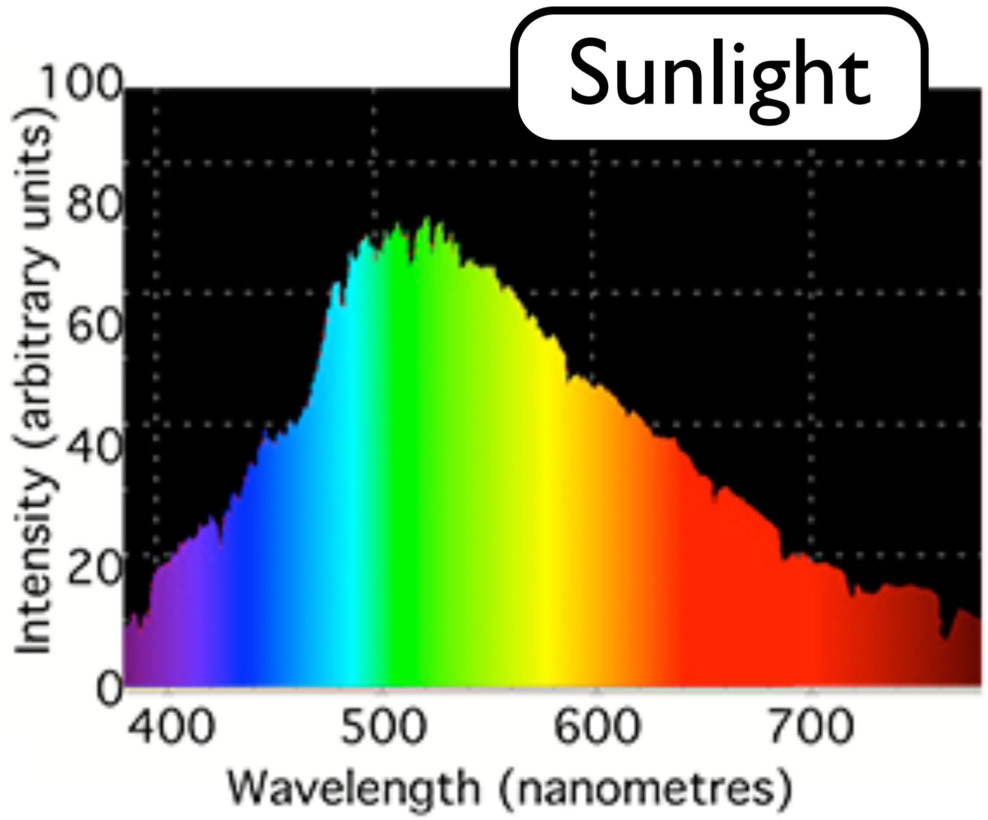
Light and Colour

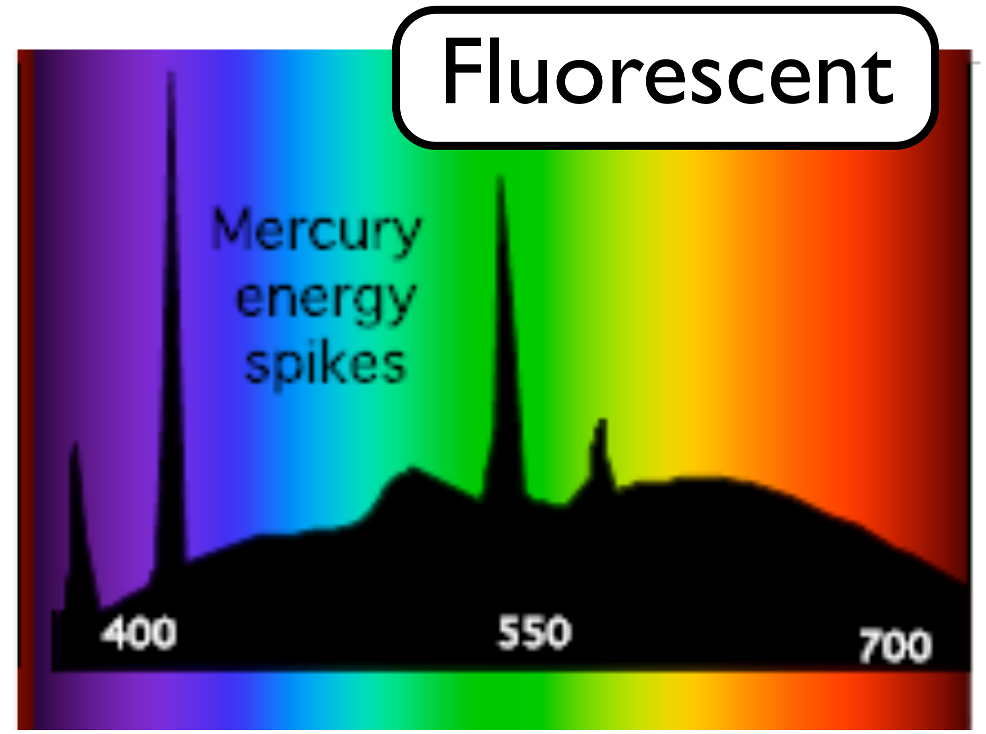


- Light is electromagnetic radiation in the 400-700nm band
- This is the peak in the spectrum of sunlight passing through the atmosphere
- Newton's Prism experiment showed that white light is composed of all frequencies
- Black is the absence of light!

Spectral Power Distribution

- The spectral distribution of energy in a light ray determines it's colour

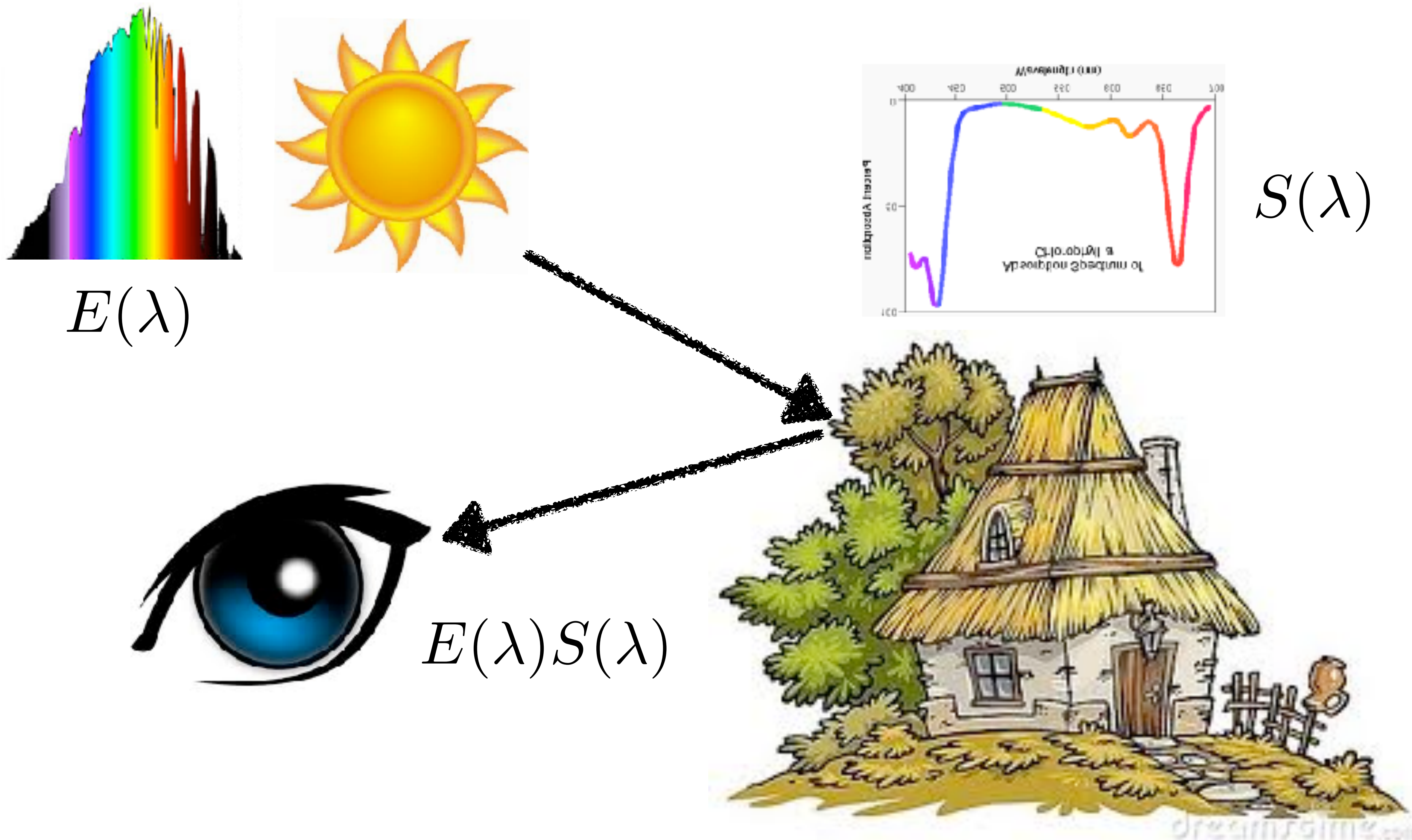




- Surfaces reflect light energy according to a spectral distribution as well
- The combination of incident spectra and reflectance spectra determine the light colour

[scratchapixel.com]

Spectral Reflectance Example



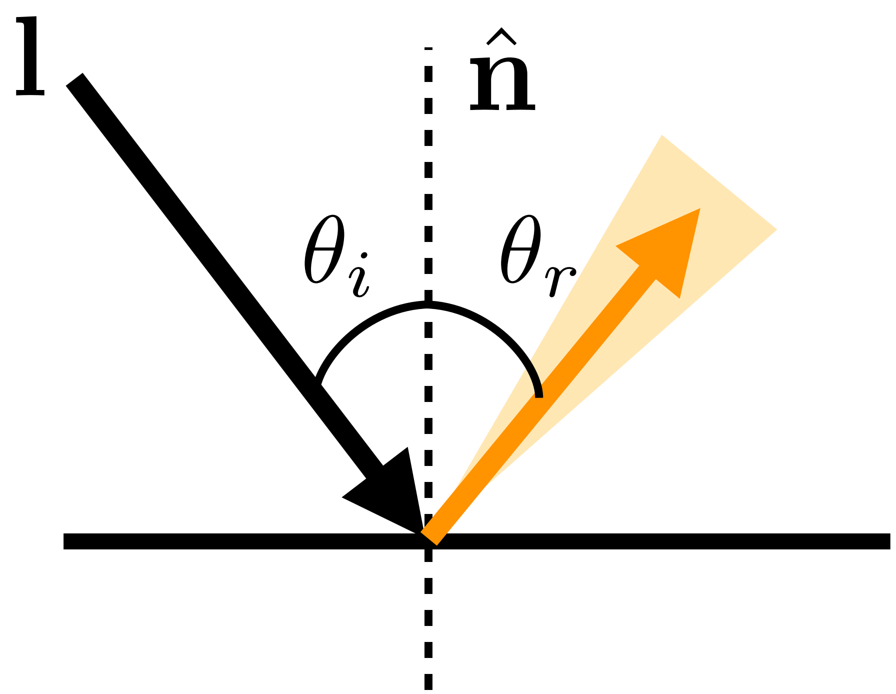
Surface Reflectance

- Reflected intensity also depends on geometry: surface orientation, viewer position, shadows, etc.

It also depends on surface properties, e.g., diffuse or specular

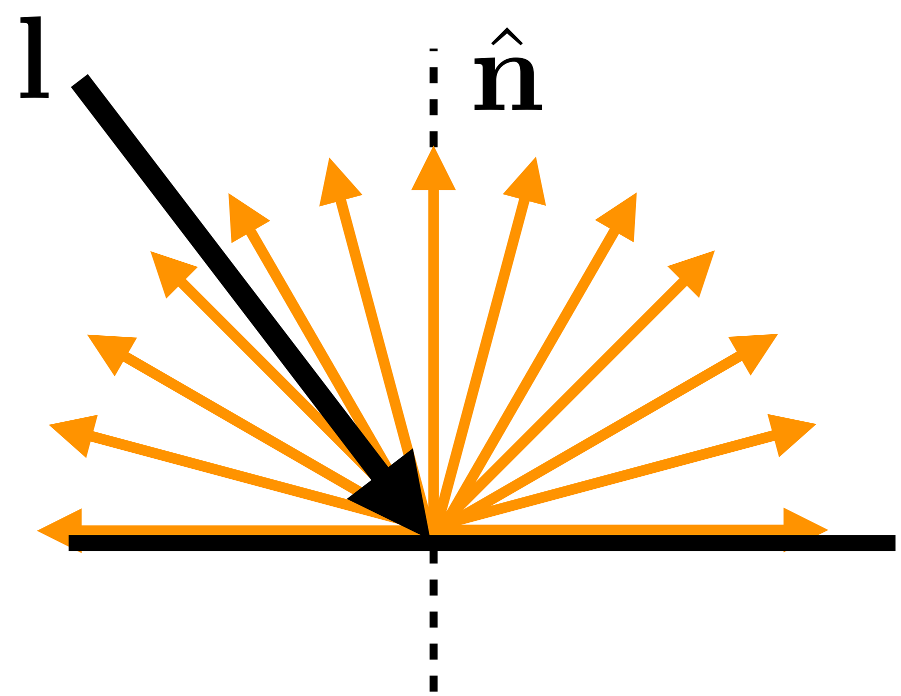
Diffuse and Specular

- A pure **mirror** reflects light along a line symmetrical about the surface normal
- A pure **diffuse** surface scatters light equally in all directions



Pure Mirror Reflection

$$\theta_i = \theta_r$$

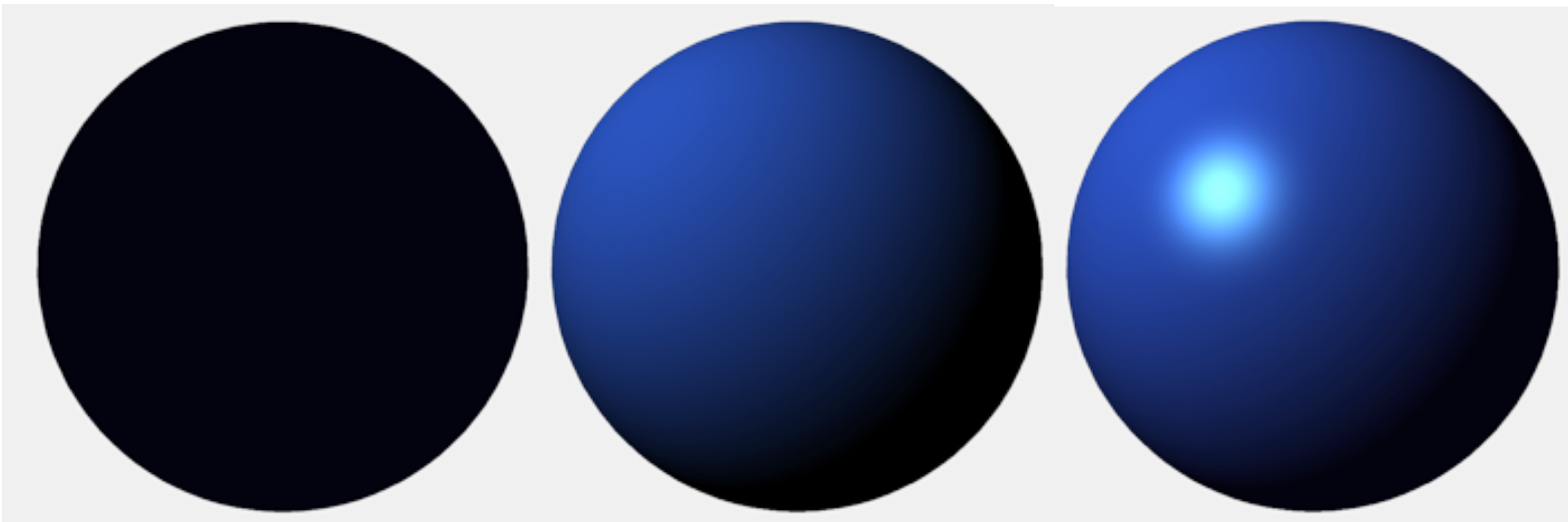


Lambertian Reflection (Diffuse)

Specular surfaces directly reflect over a small angle

Diffuse and Specular

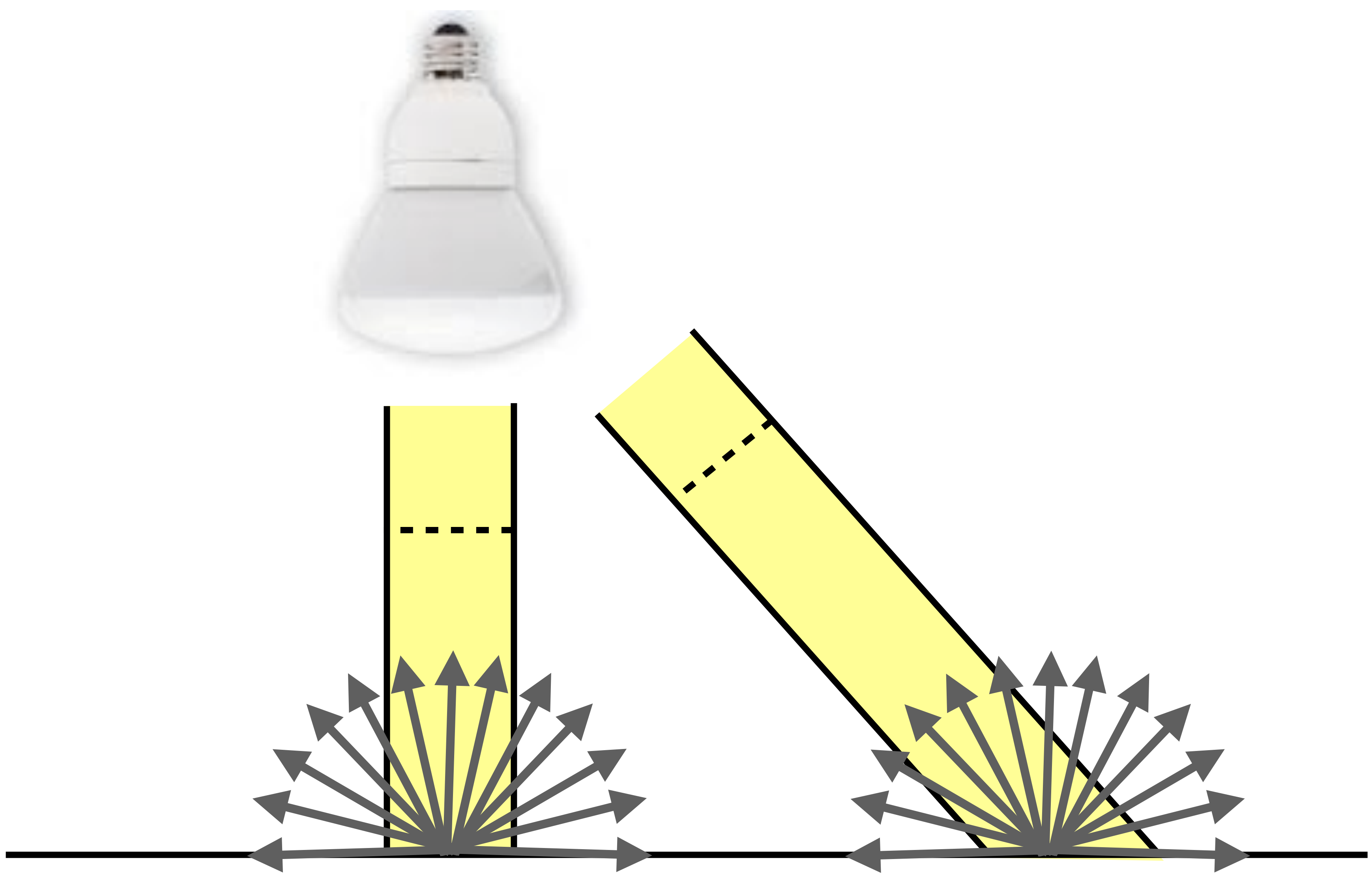
- A sphere lit with ambient, +diffuse, +specular reflectance



Ambient +Diffuse +Specular

Diffuse Reflection

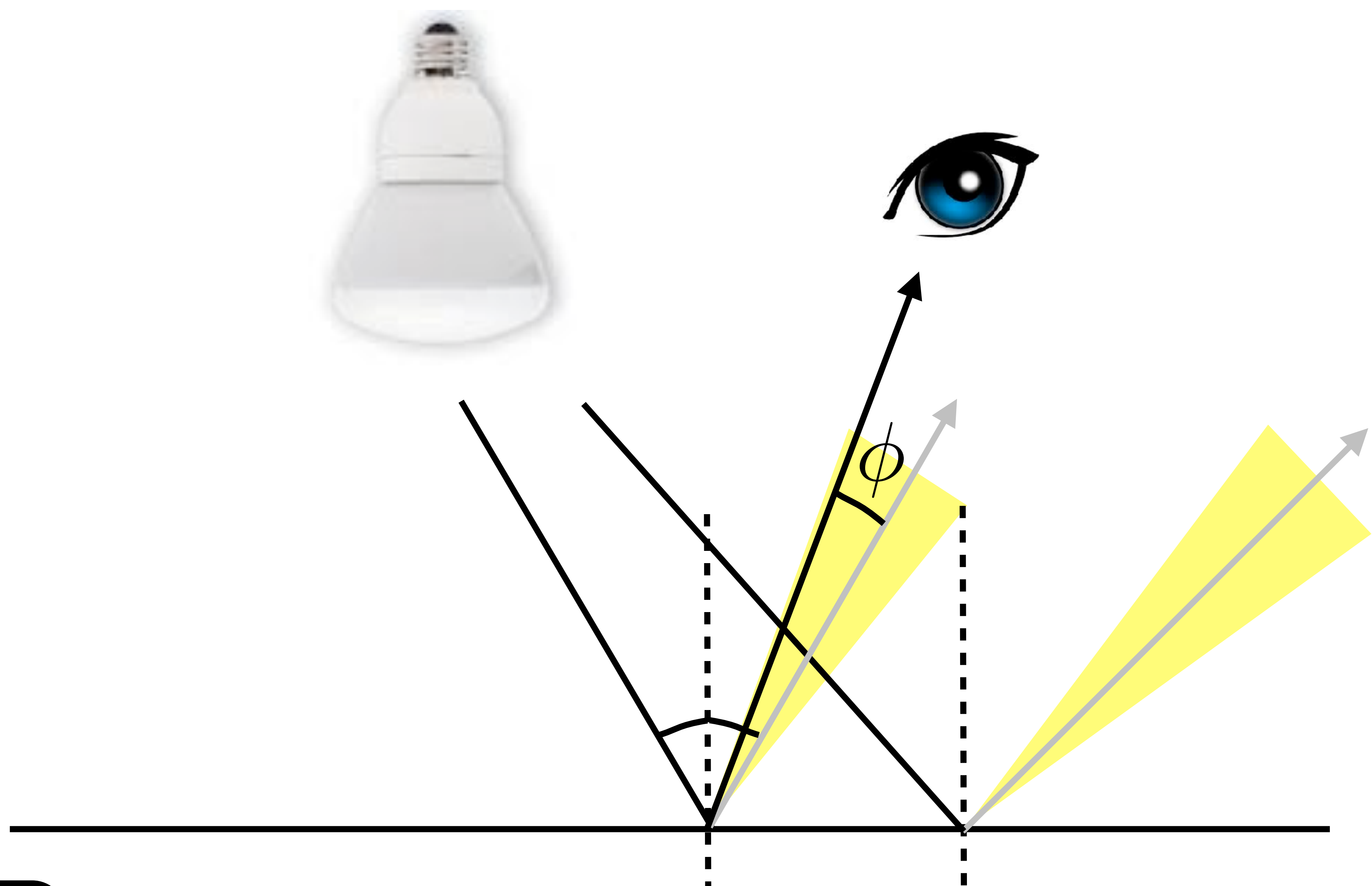
- Light is reflected equally in all directions (Lambertian surface)
- But the amount of light reaching unit surface area depends on the angle between the light and the surface...



1.1

Specular Reflection

- Light reflected strongly around the mirror reflection direction
- Intensity depends on viewer position

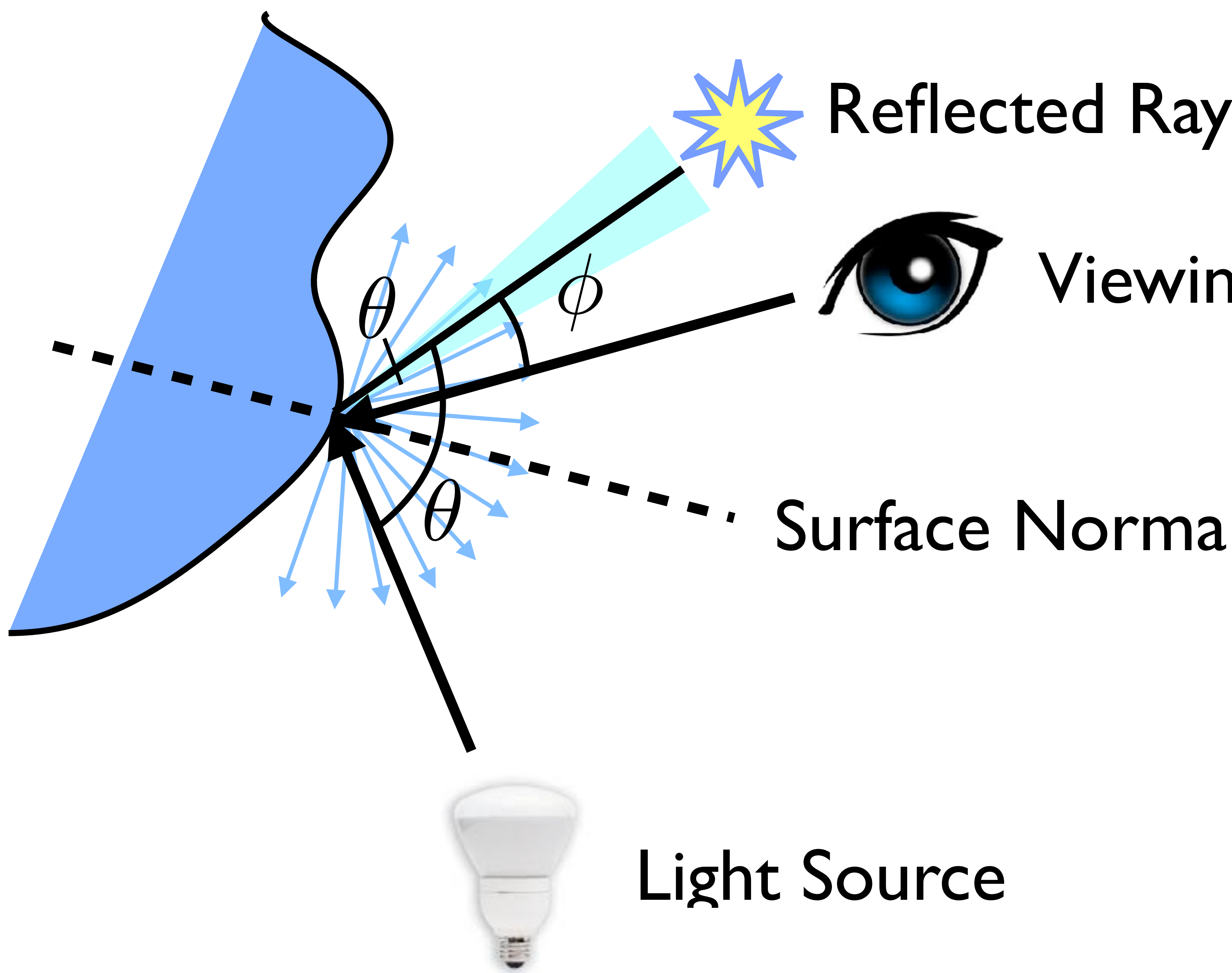


1.2

Phong Illumination Model

- Includes ambient, diffuse and specular reflection

$$I = k_a i_a + k_d i_d \cos\theta + k_s i_s \cos^{\alpha}\phi$$



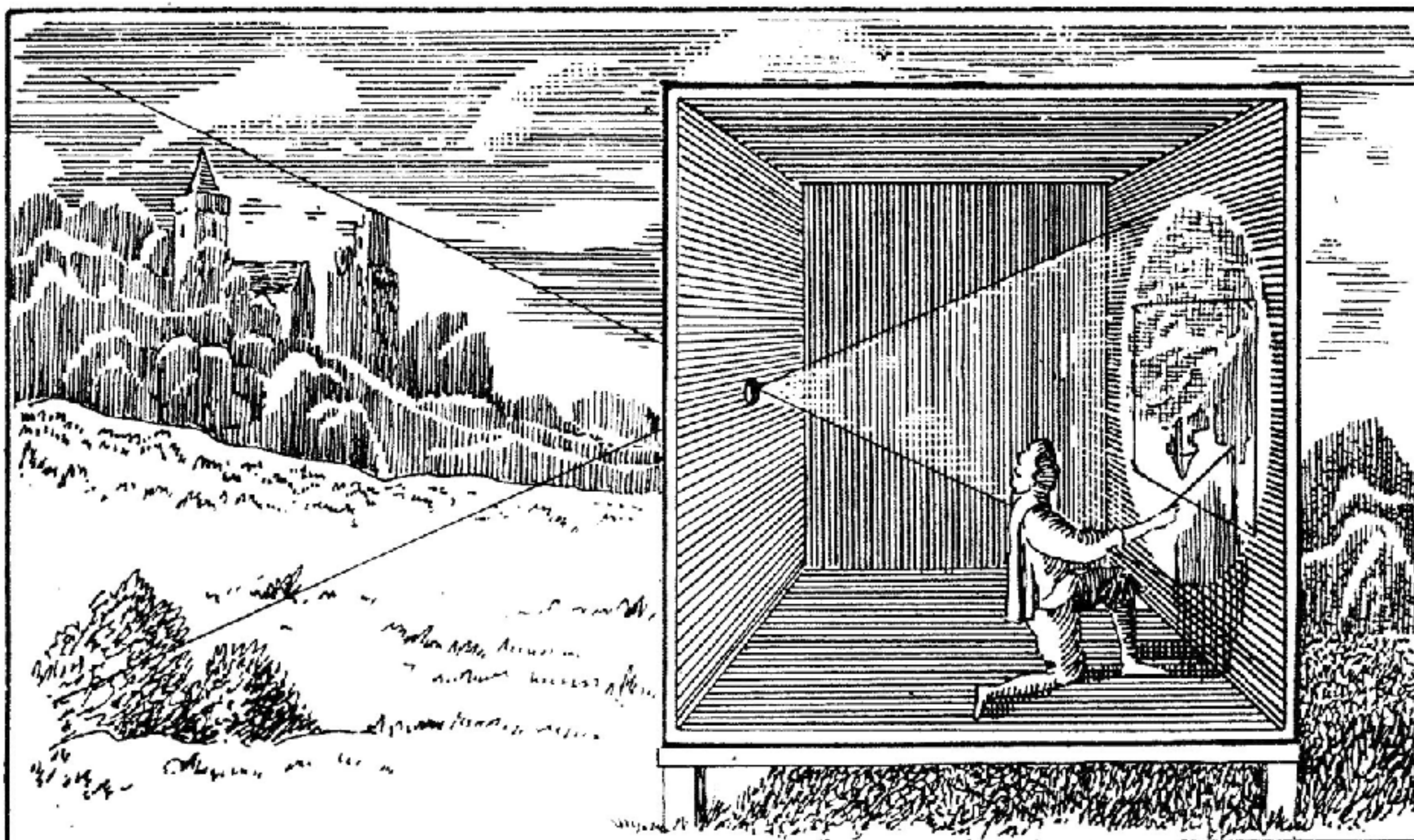
Reflectance in Vision

Reflectance in Vision

- More complex models than Phong are possible with reflected intensity at a given ray an arbitrary function of the surface geometry and lights, see Szeliski 2.2.2 (BRDFs)
- For Computer Vision, understanding reflection can help us to infer shape, e.g., **shape from shading** and **photometric stereo,** we will revisit this later in the course

[Hertzmann Seitz 2003]

Optics



Camera Obscura = “dark room”

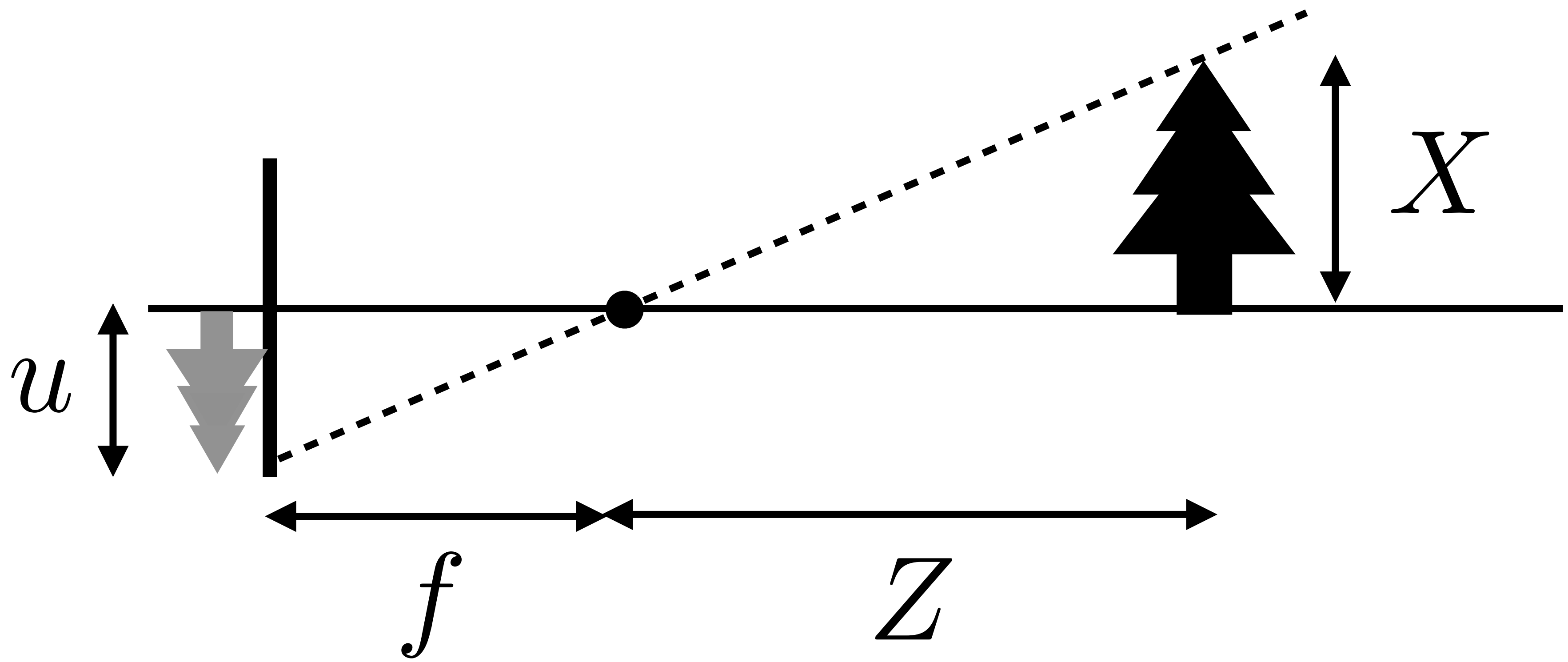
Clifton Observatory

[P. Chapman]

A working camera obscura open to the public

Pinhole Camera

- All rays pass through a single point (the pinhole)

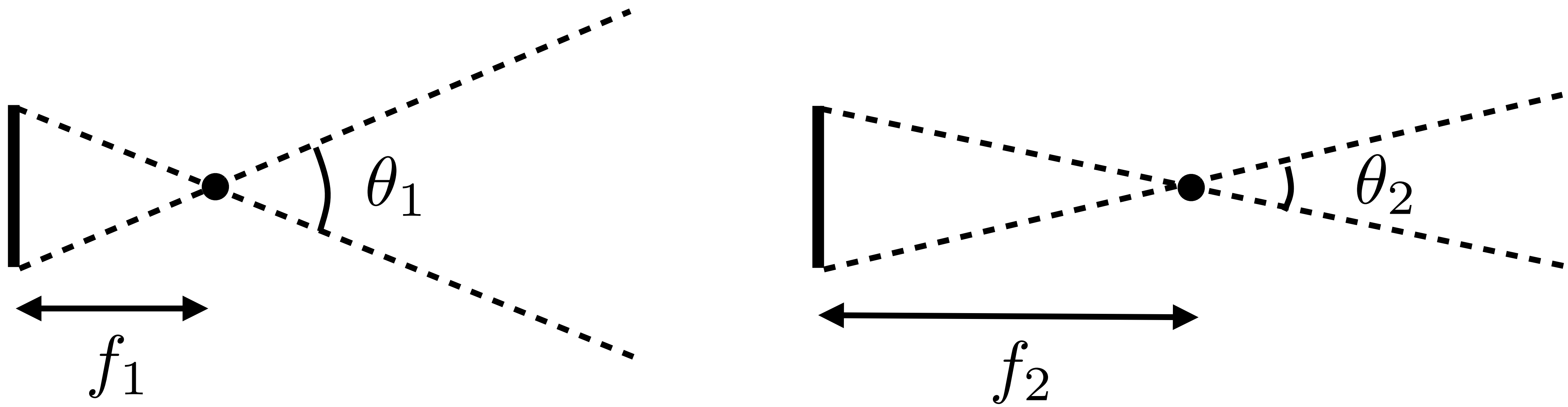


- Similar triangles

 1.3

Focal Length

- For a fixed sensor size, focal length determines the field of view (fov)



Q: What is the field of view of a full frame (35mm) camera with 50mm lens? 100mm lens?

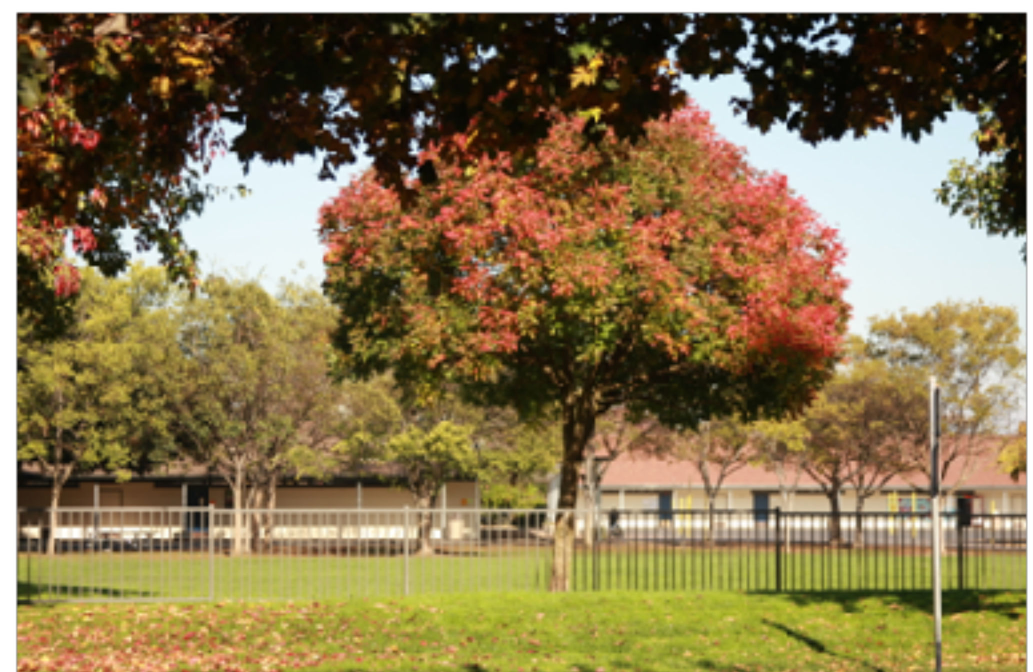
1.4

Focal Length

28 mm

35 mm

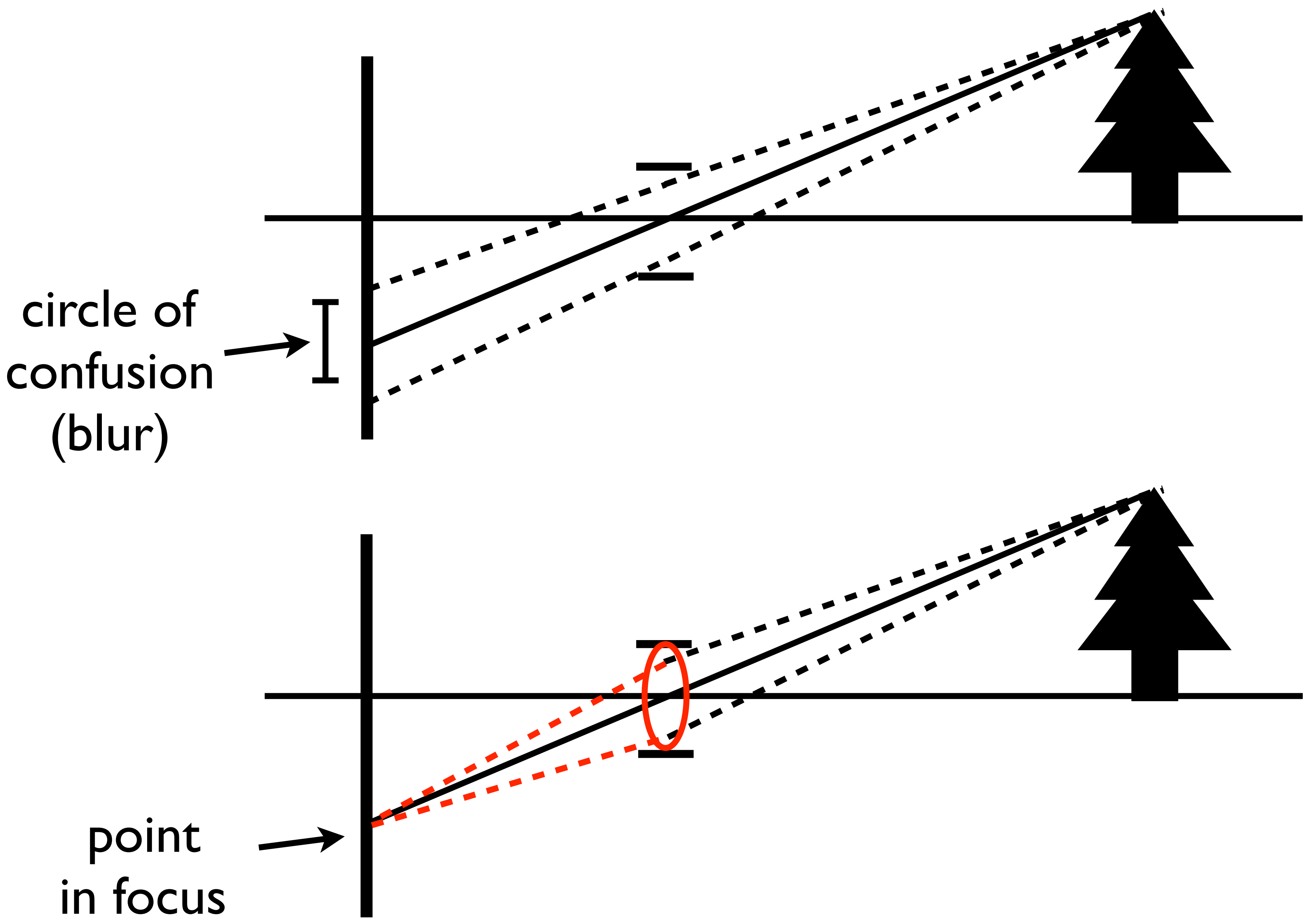
50 mm



70 mm

Finite Aperture

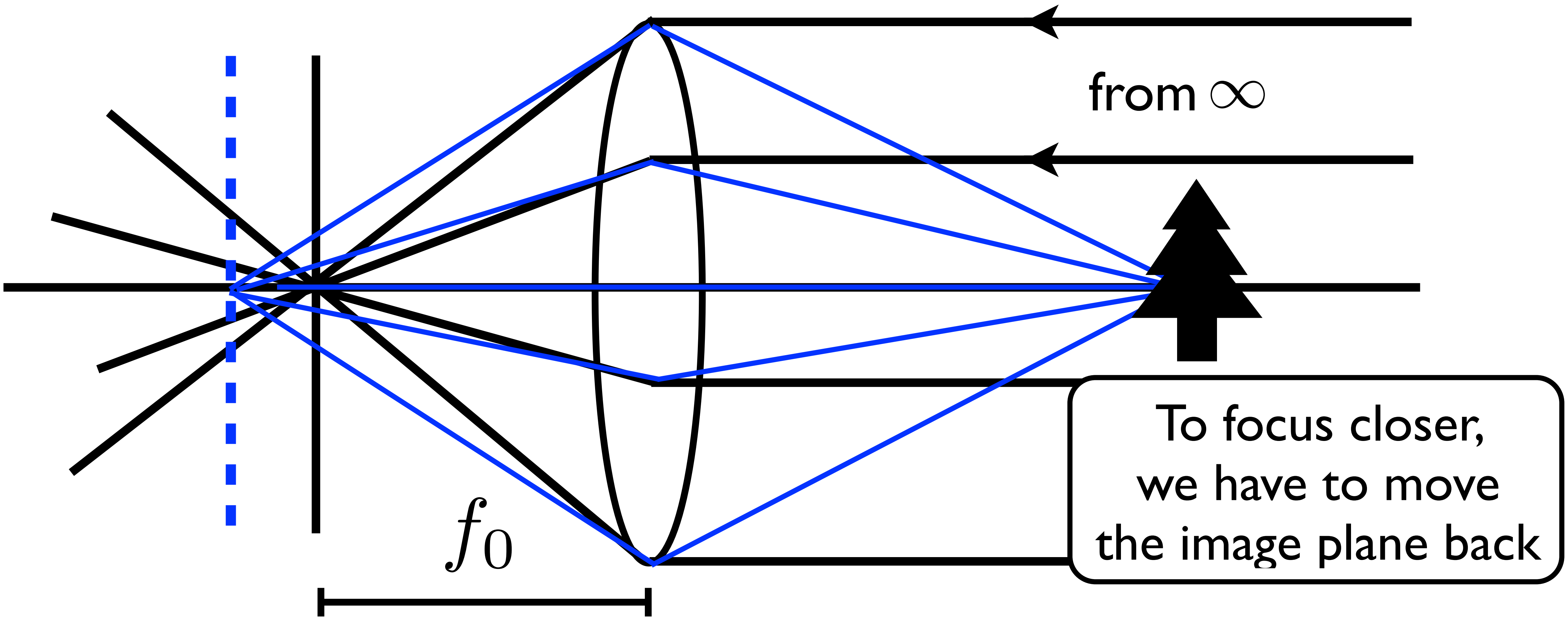
- A real camera must have a finite aperture to get enough light, but this causes **blur** in the image



Solution: use a **lens** to focus light onto the image plane

Lens Basics

- A lens focuses rays from infinity at the focal length of the lens
- Points passing through the centre of the lens are not bent

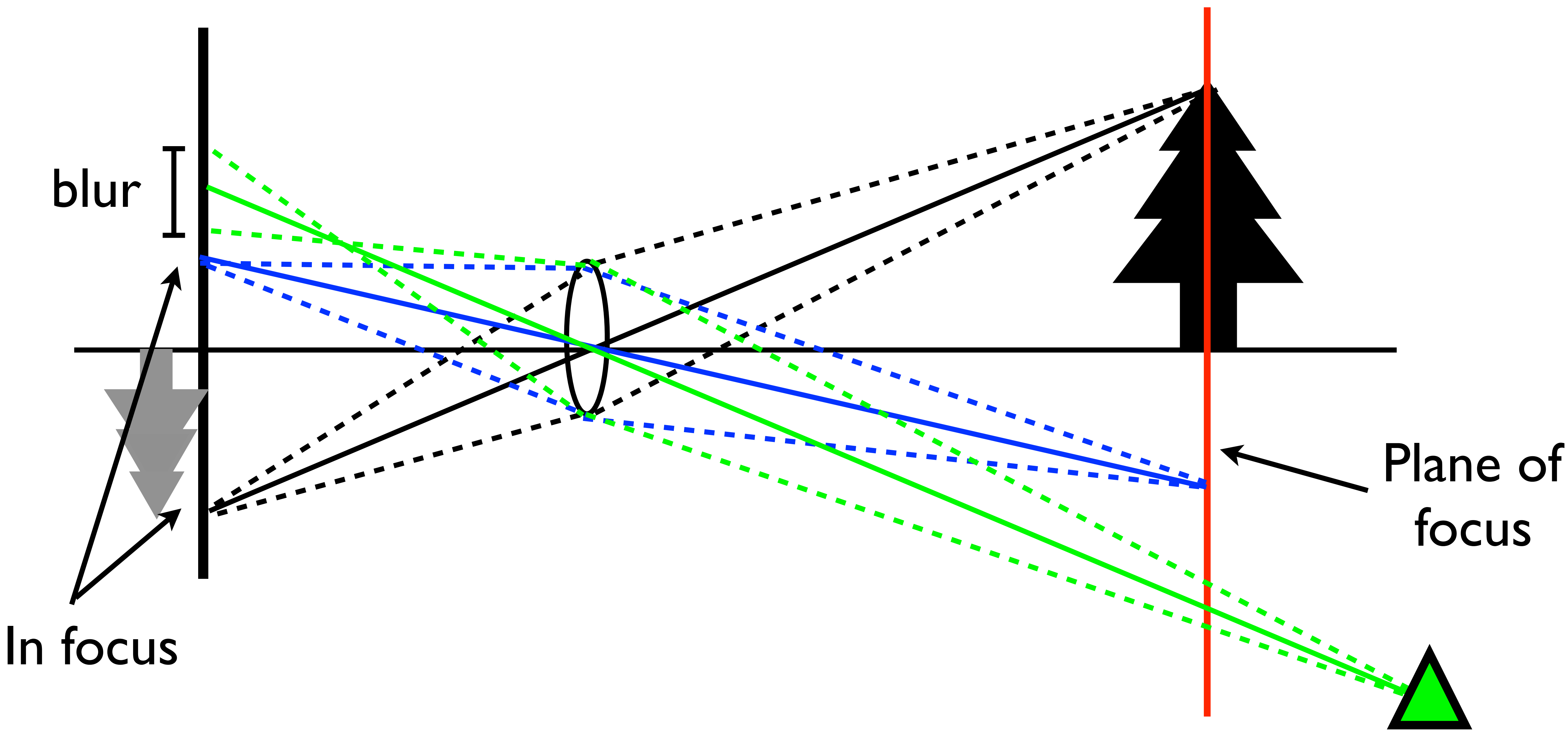


- We can use these 2 properties to find the lens equation

1.5

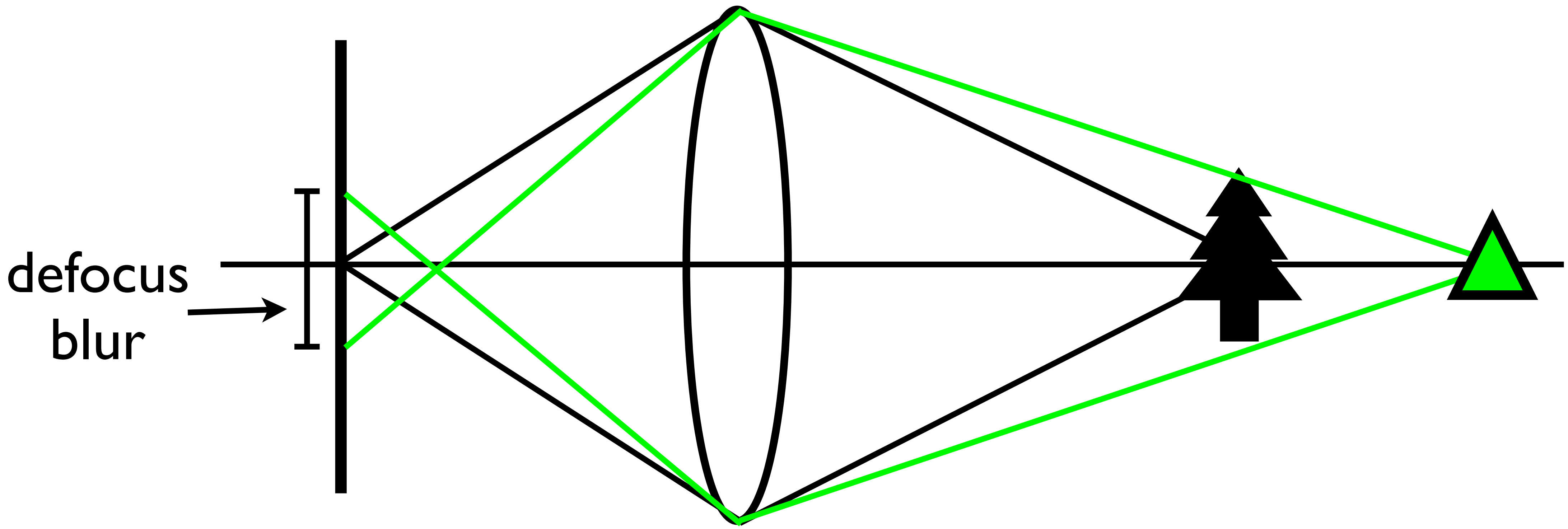
Lens Basics

- Note that lenses focus all rays from a **plane** in the world

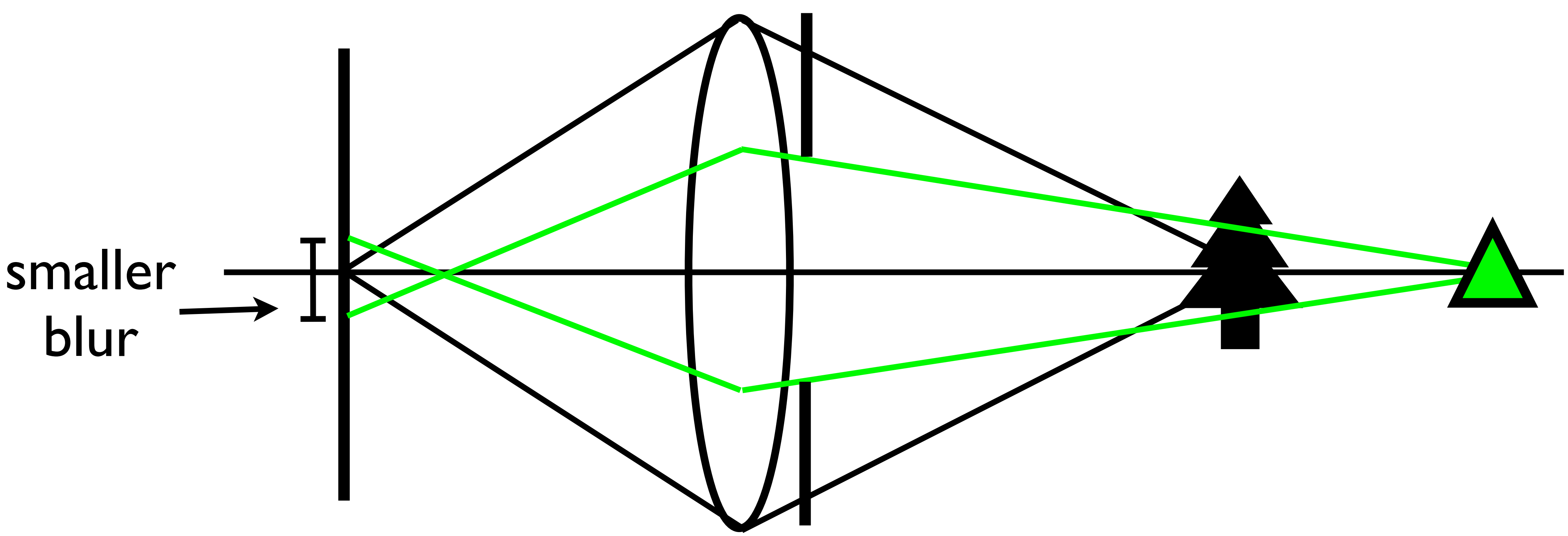


- Objects off the plane are blurred depending on distance

Effect of Aperture



Smaller aperture ⇒ smaller blur, larger **depth of field**

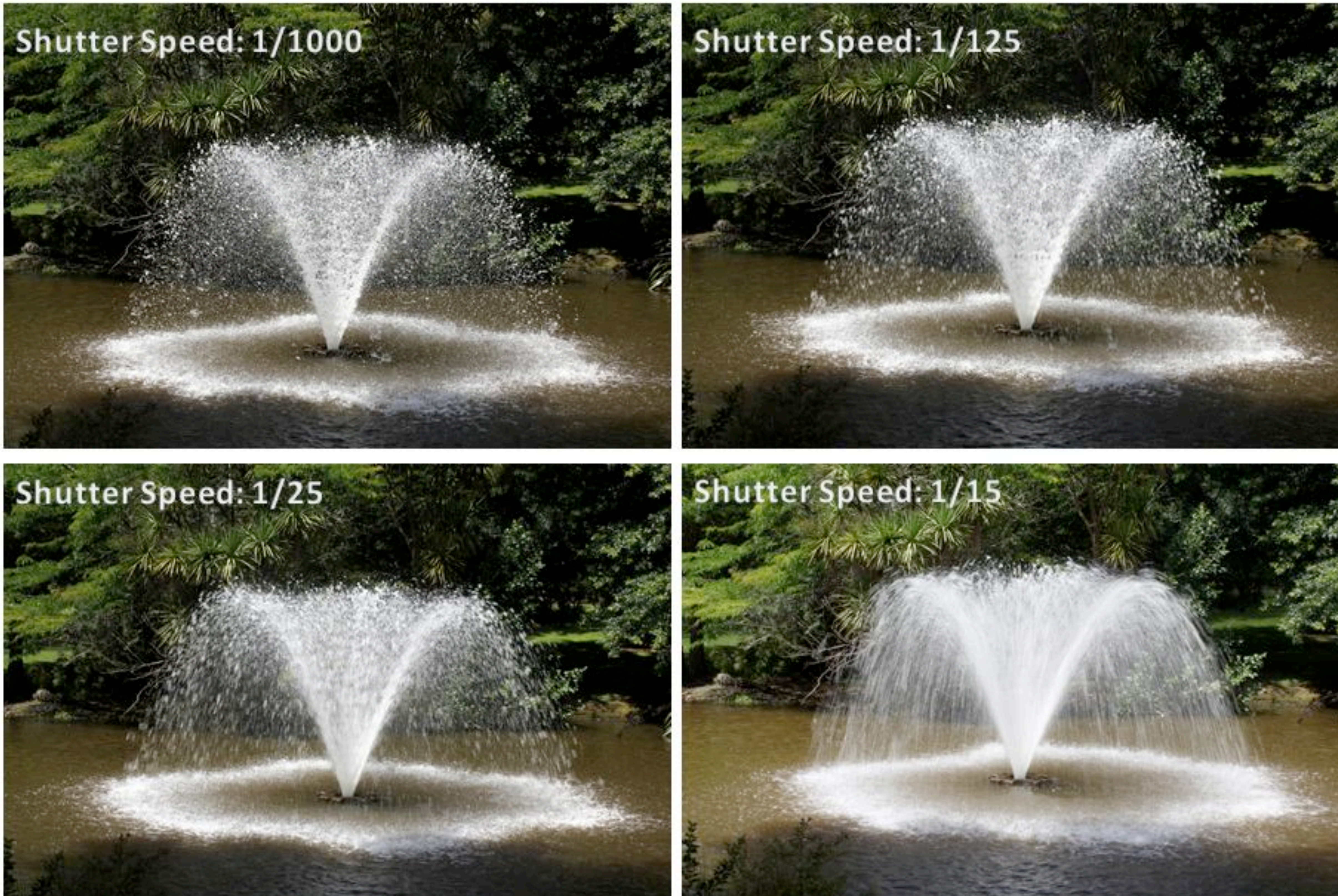


Depth of Field

- Photographers use large apertures to give small depth of field

Aperture size = f/N, ⇒ large N = small aperture

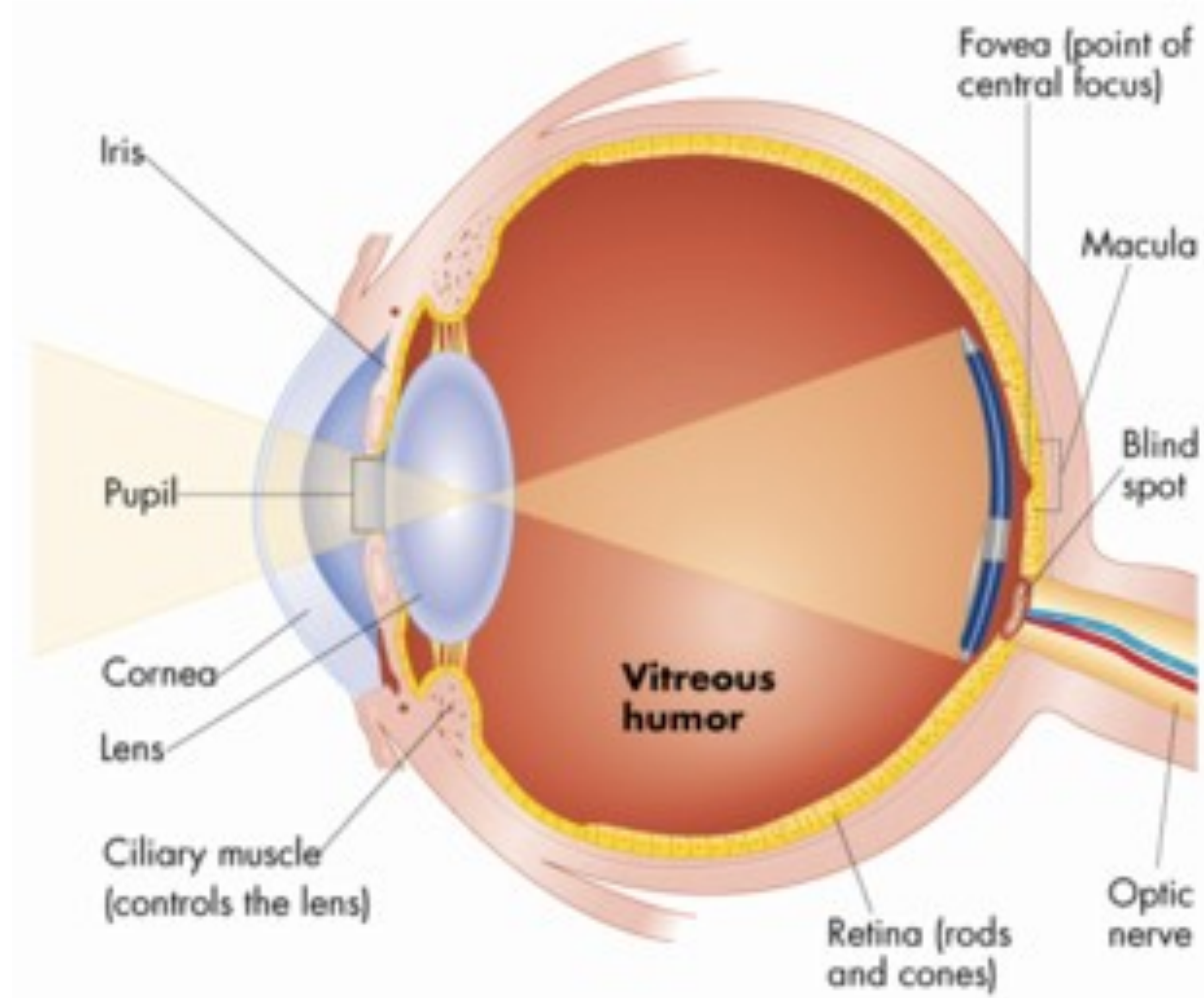
Shutter Speed



Real Lenses

- Multiple stages of positive and negative elements with differing refractive indices
- Deal with issues such as chromatic aberration (different colours bent by different amounts), vignetting (light fall off at image edge) and sharp imaging across the zoom range

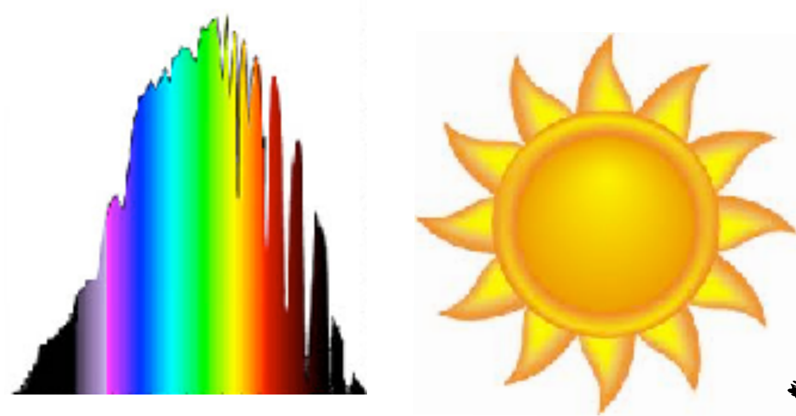
Sensors



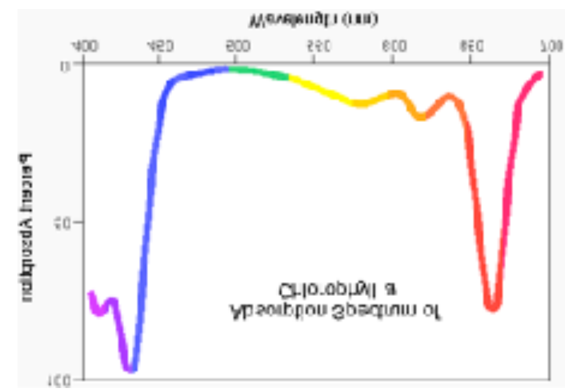
CMOS (or CCD)

Retina

Colour Perception



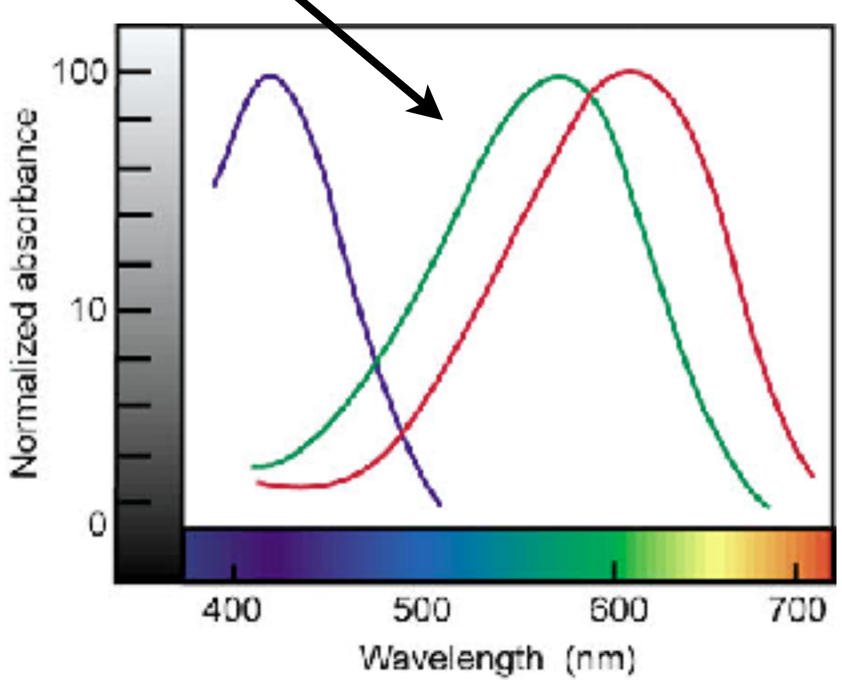
$E(\lambda)$



$S(\lambda)$

Cone responses

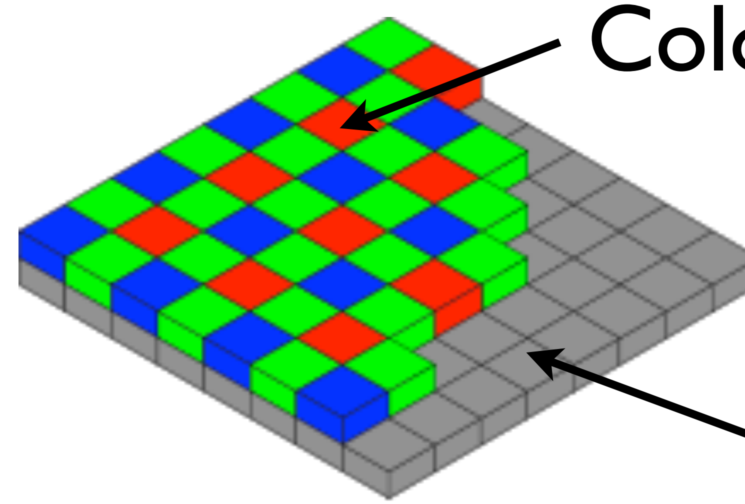
$R_{red}(\lambda)$



Cone excitation (multiply and add):

$$\rho_{red} = \int R_{red}(\lambda)E(\lambda)S(\lambda)\ d\lambda$$

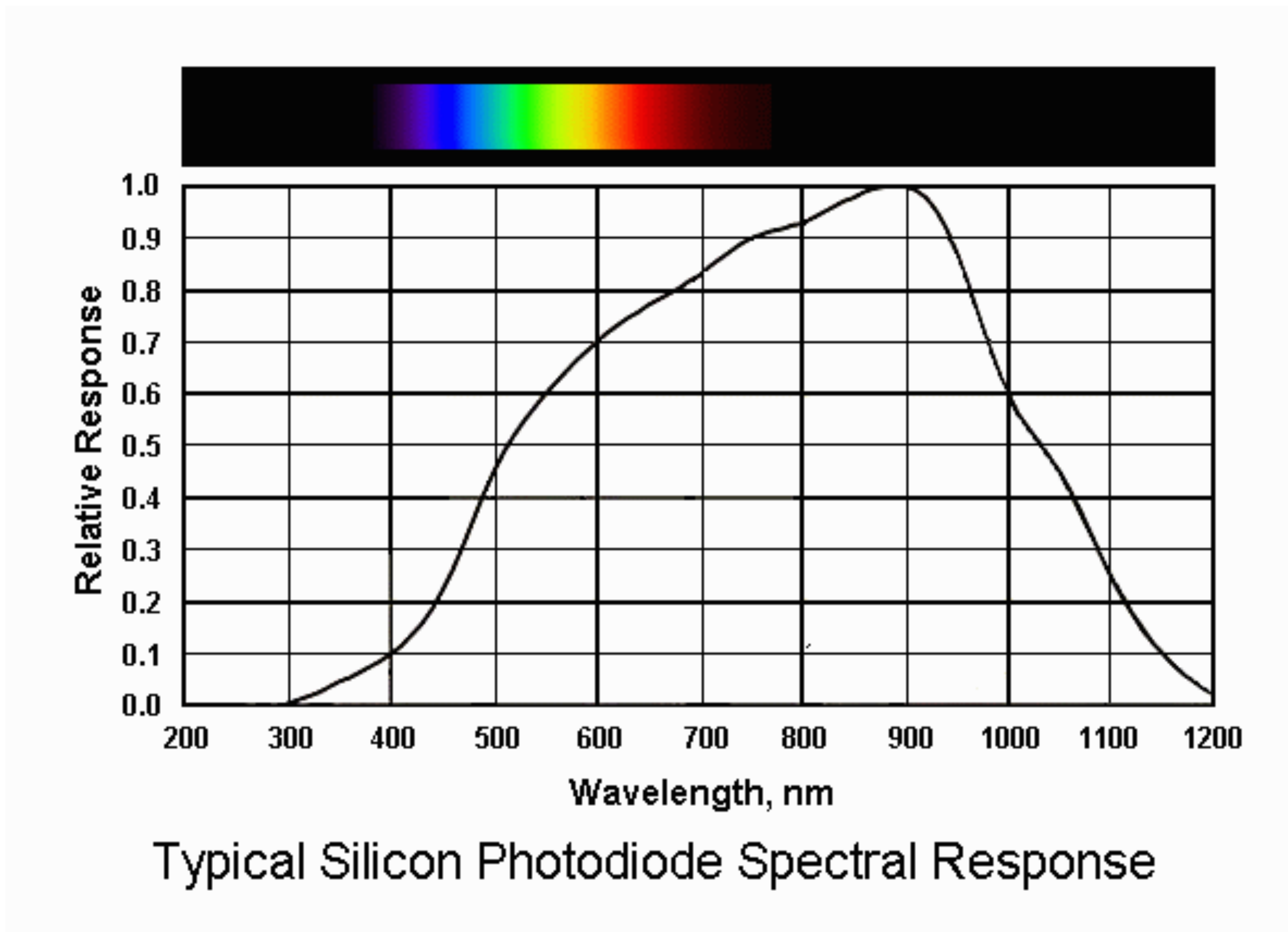
Digital Sensor



- Analogue image is sampled by a CCD or CMOS sensor
- RGB pixels arranged in a "Bayer" pattern
- Demosaicing gives an RGB value per pixel

Silicon Sensitivity

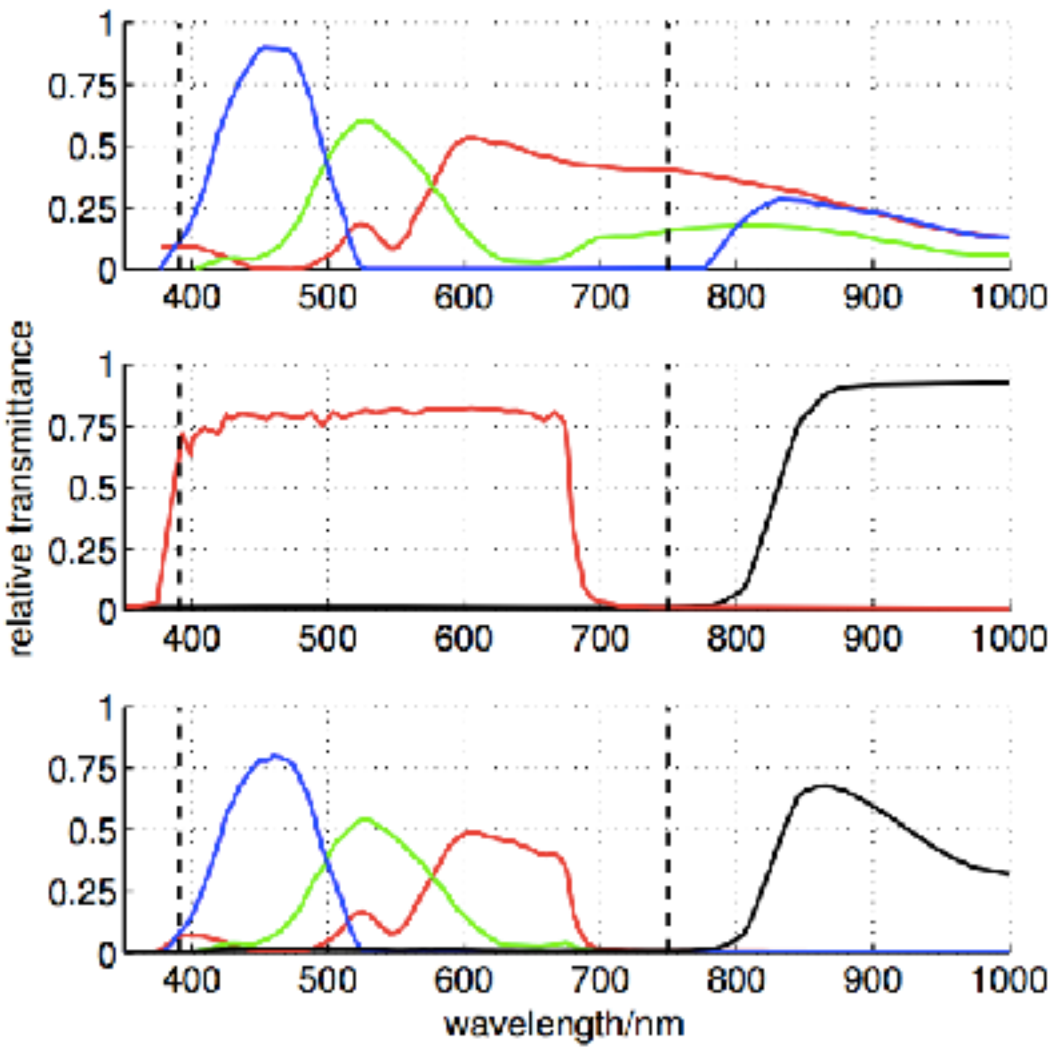
- Silicon is sensitive beyond the visible spectrum → cameras have and extra filter called a "hot mirror" to block NIR light



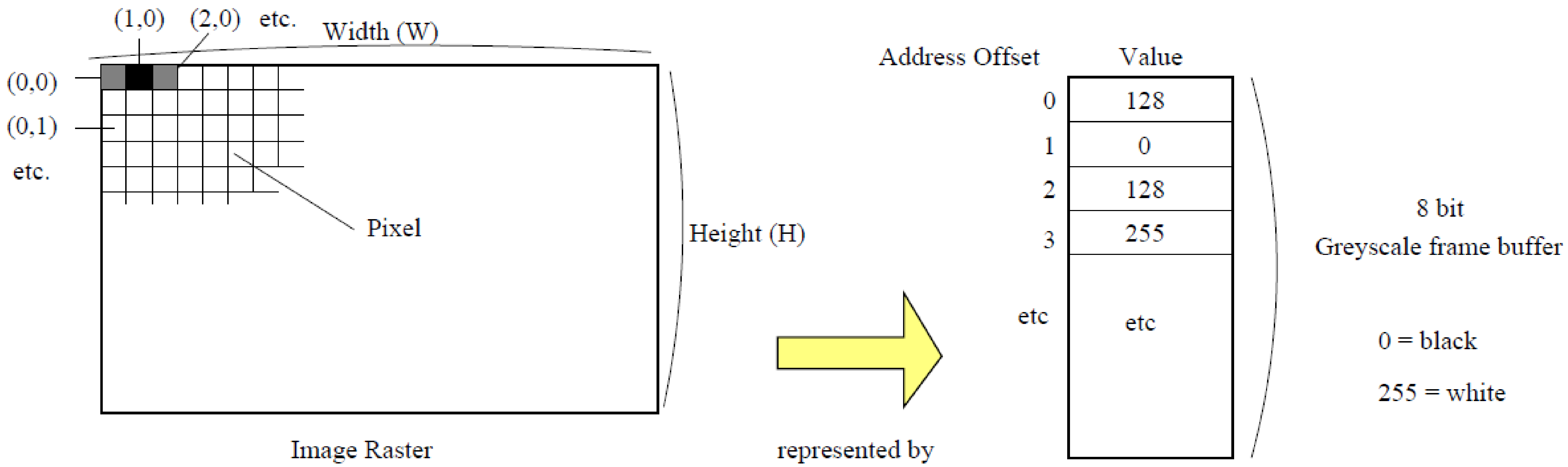
[arduino.cc]

Near Infrared Photography

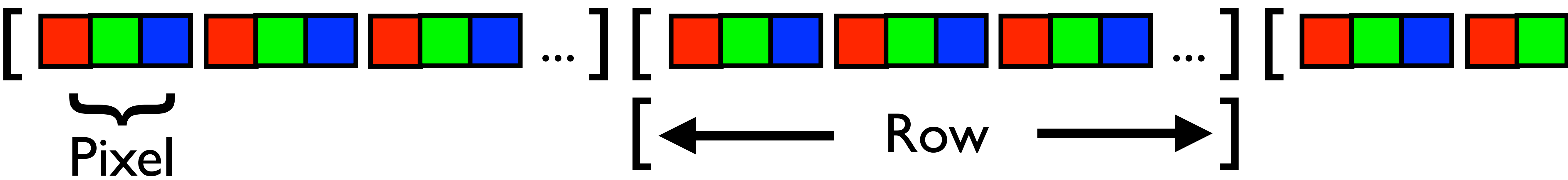
- Removing the blocking filter allows us to take NIR photos



The Digital Image



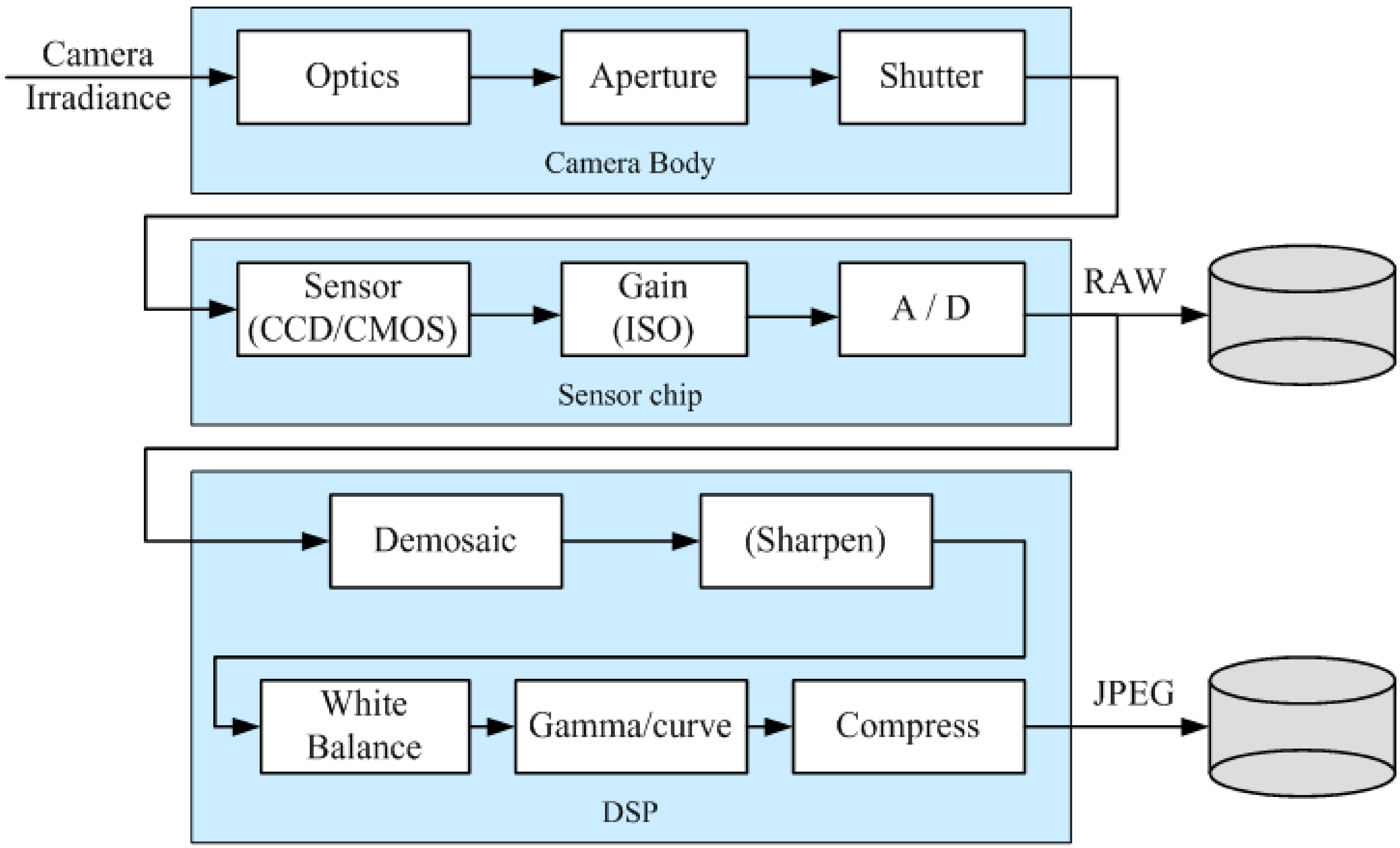
- e.g., arranged in memory with RGB pixels stored in rows:



- Many other possibilities, e.g., BGR, RGBA pixels, row/column major ordering, and rows or columns aligned to power of 2 boundaries

Digital Camera Processing

- Main stages in a digital camera



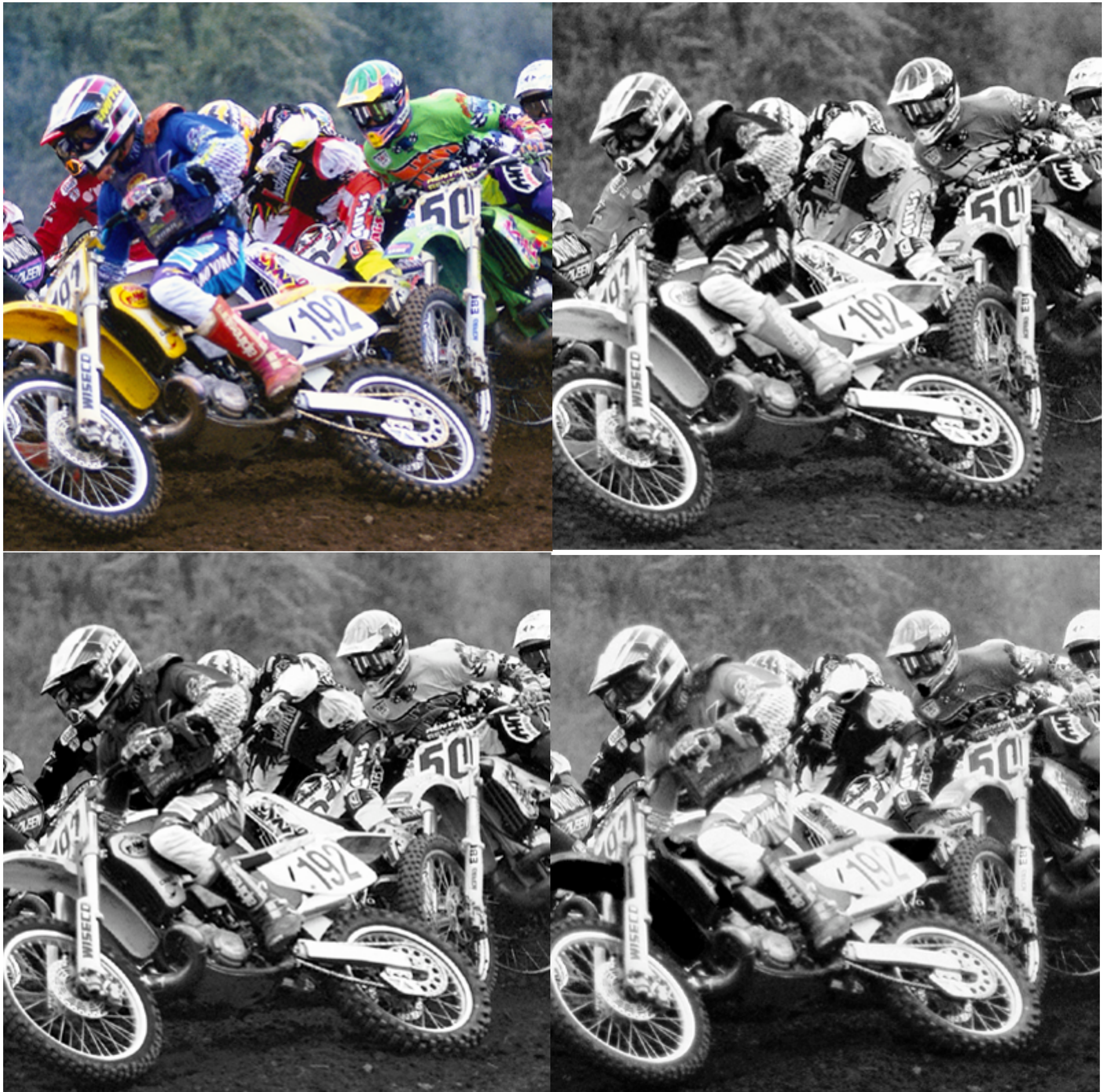
[Szeliski 2.3]

White Balance

- Humans are good at adapting to global illumination conditions: you would still describe a white object as white whether under blue sky or candle light.
- However, when the picture is viewed later, the viewer is no longer correcting for the environment and the illuminant colour typically appears too strong.
- **White balancing** is the process of correcting for the illuminant

1.6

- A simple white balance algorithm is to assume the scene is grey on average "greyworld", state of the art methods use learning, e.g., Barron ICCV 2015



Red

Green

Blue

Y

Cr

Cb

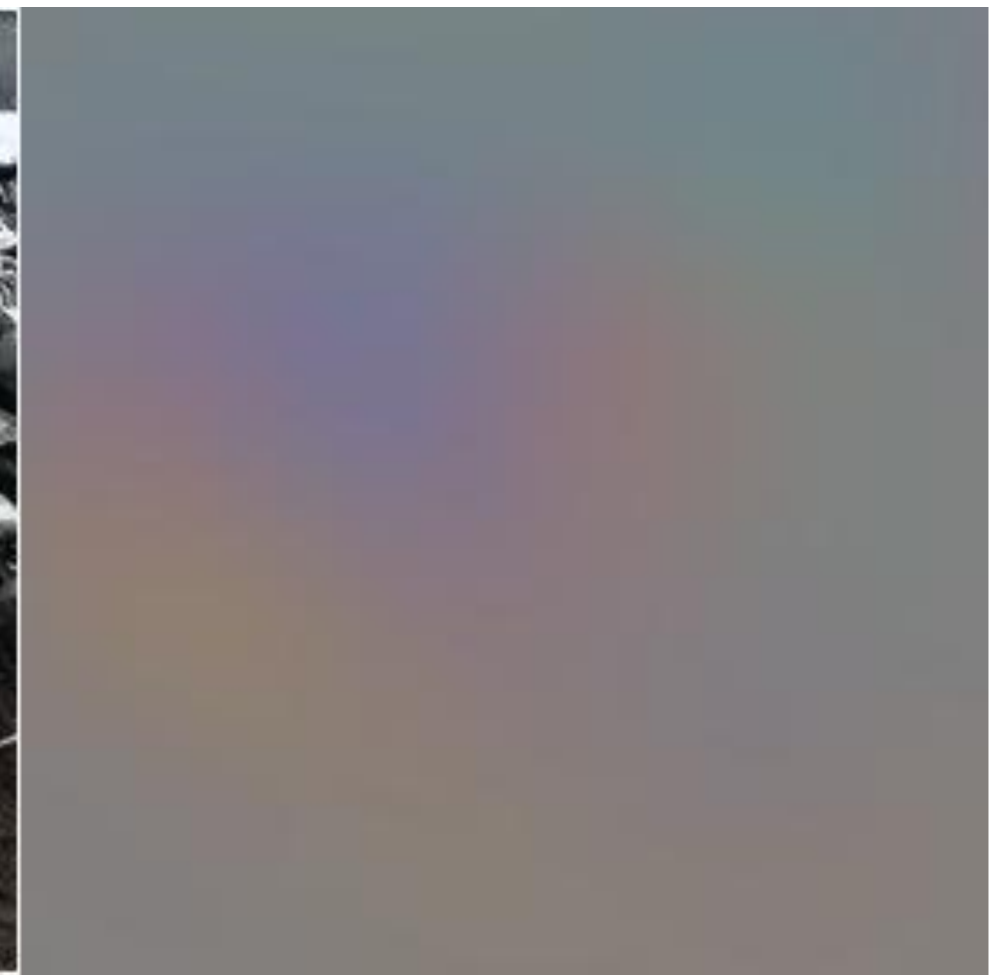
sigma = 1.0

sigma = 2.0

sigma = 4.0

sigma = 8.0

sigma = 16.0



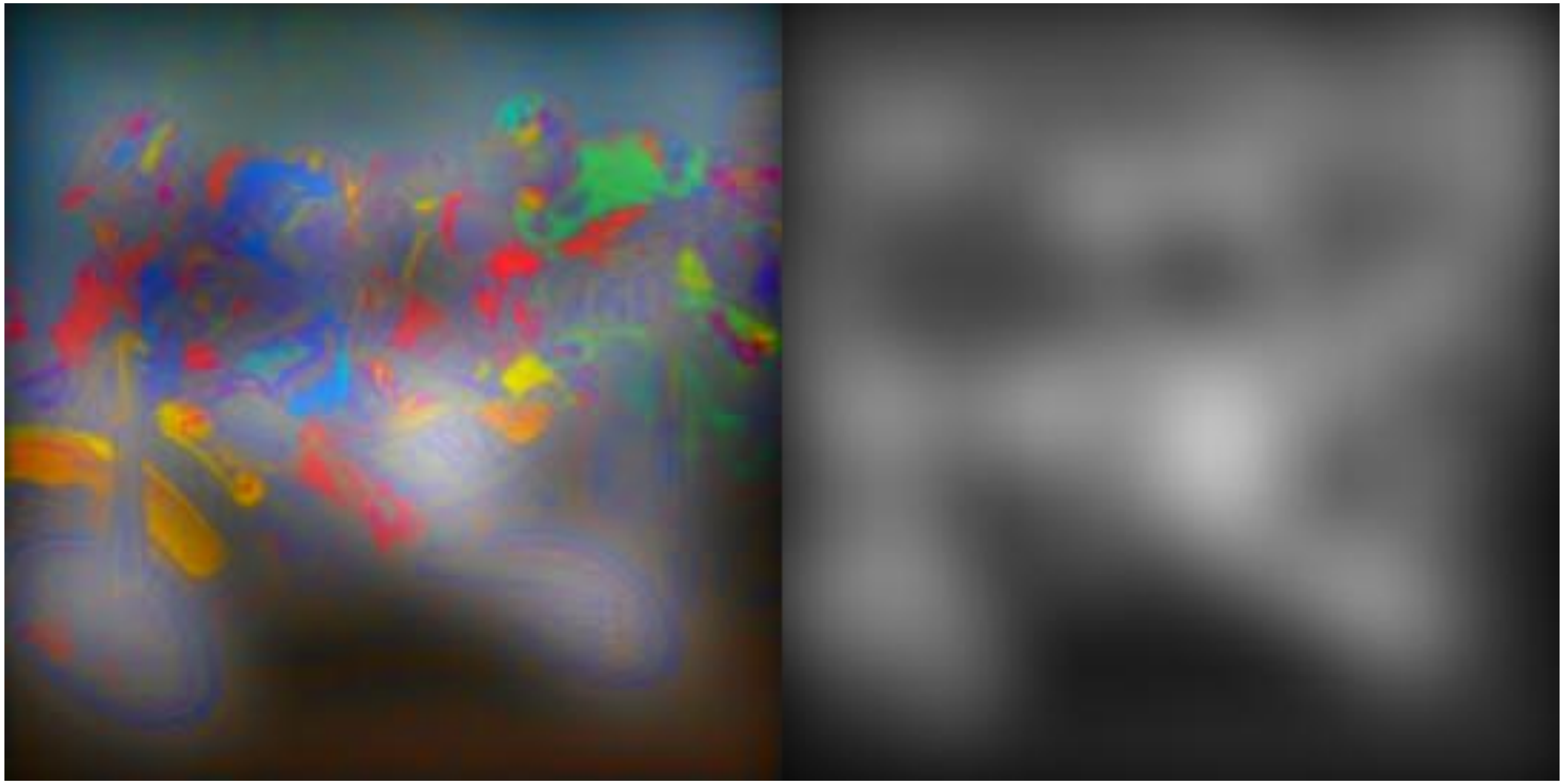
sigma = 32.0

sigma = 1.0

sigma = 2.0

sigma = 4.0

sigma = 8.0



sigma = 16.0

Compressibility of Chrominance

Original

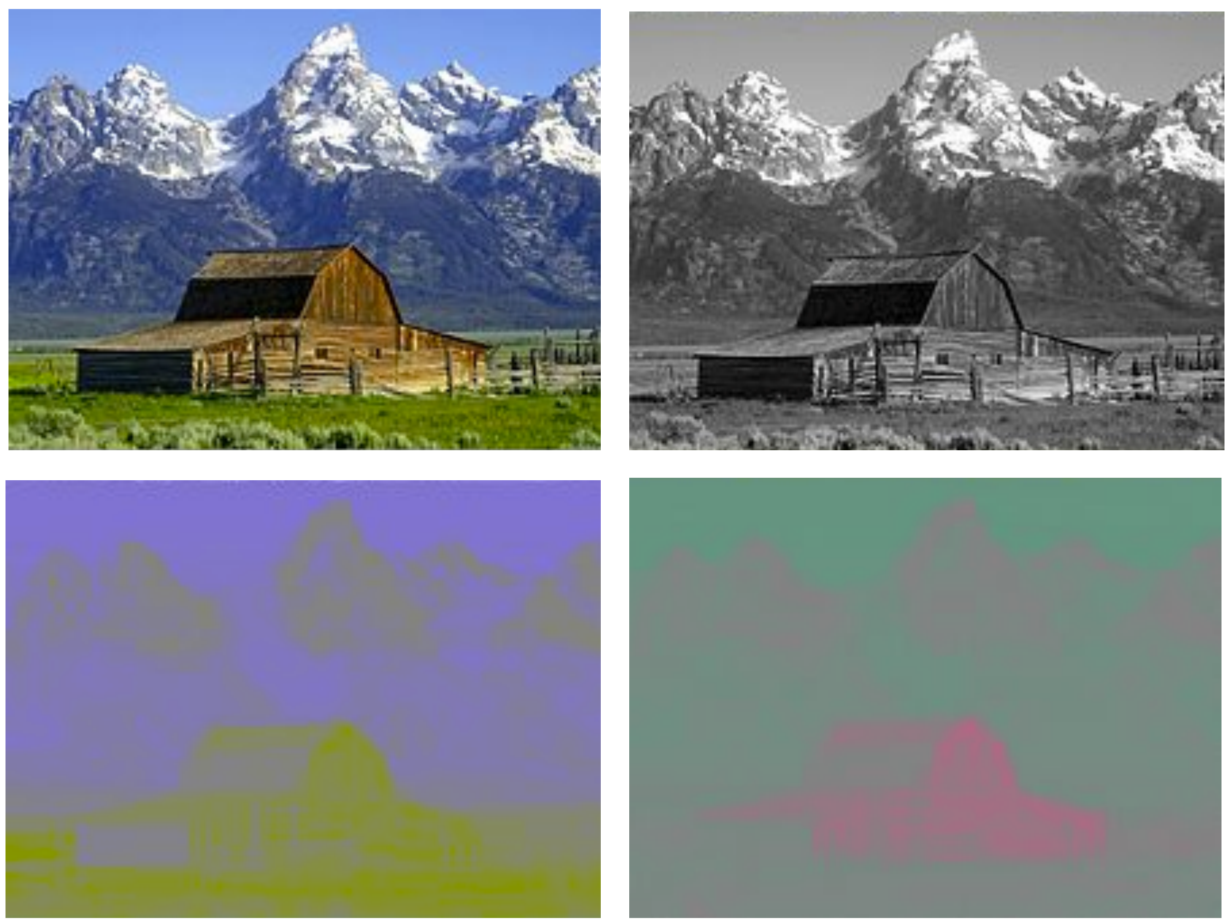
Chrominance
64 x subsampled
(scaled 8 x in
each direction)

Luminance
64 x subsampled

→use YCrCb and fewer bits for Cr and Cb

YCbCr

- Separates luminance (Y) from chrominance (Cb, Cr)
- Chrominance can be compressed more (e.g. 1/2 size in JPG)



$$Y' = 16 + 65.5R' + 128.6G' + 25.0B'$$
$$Cb = 128 - 37.8R' - 74.2G' + 112B'$$
$$Cr = 128 + 112.0R' - 93.8G' - 18.2B'$$

- Linear transform of RGB
- Primes = gamma correction

YCbCr is used for image and video coding. Human vision uses a similar transform (opponent colours) and we have more rods than cones

Gamma Correction

- Equal steps in luminance ≠ equal in perceived brightness

linear luminance (raw)	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
equal brightness steps	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0

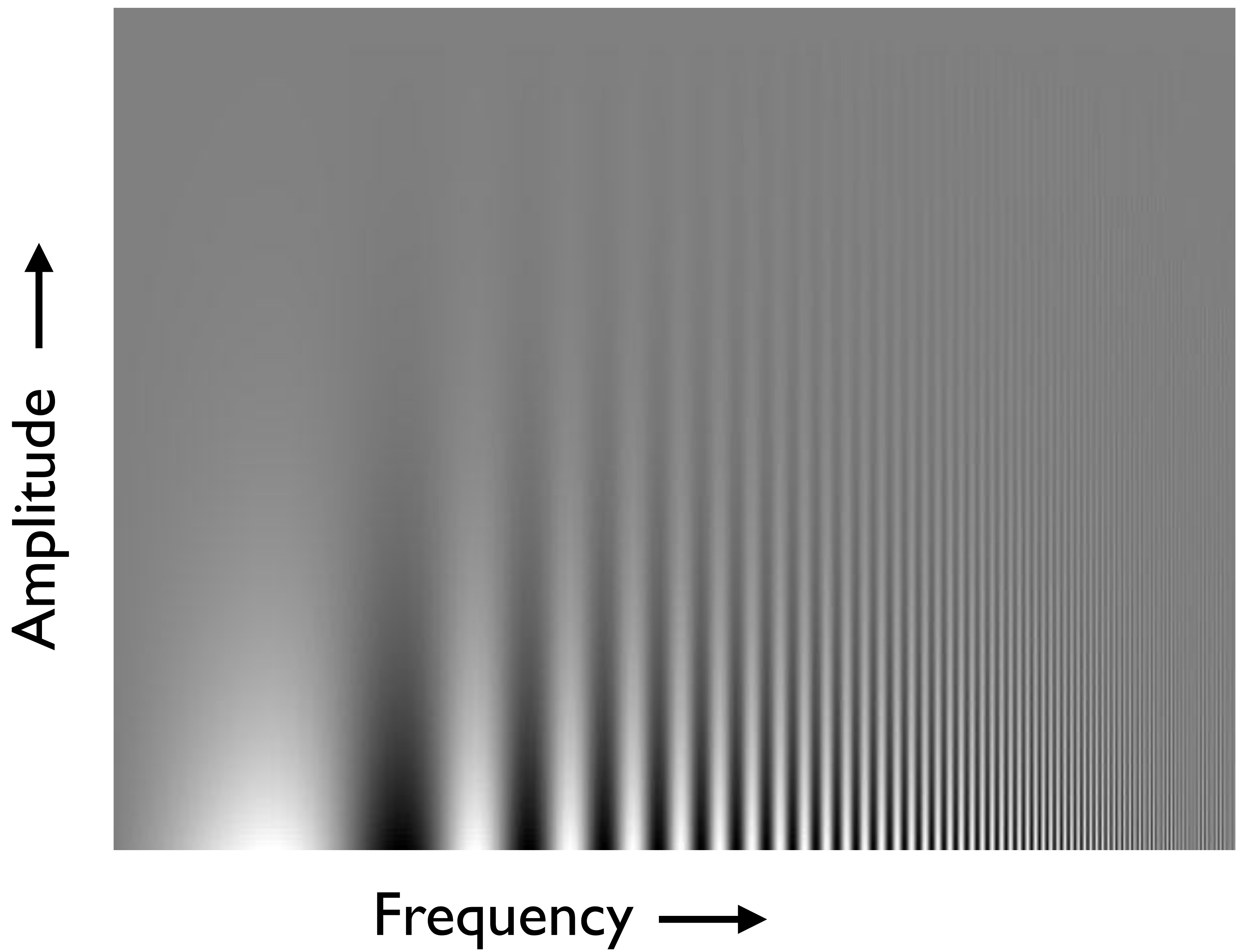
- Equal steps in perceived brightness are achieved by increasingly large steps in luminance (sensor counts)
- Human brightness perception follows a power law:

$$L = V^{\gamma}$$

- Using raw sensor counts wastes bits as we can't differentiate the large values → use **gamma corrected encoding** that allocates more bits to smaller values

Contrast Sensitivity

- Human visual system is most sensitive to mid-frequencies

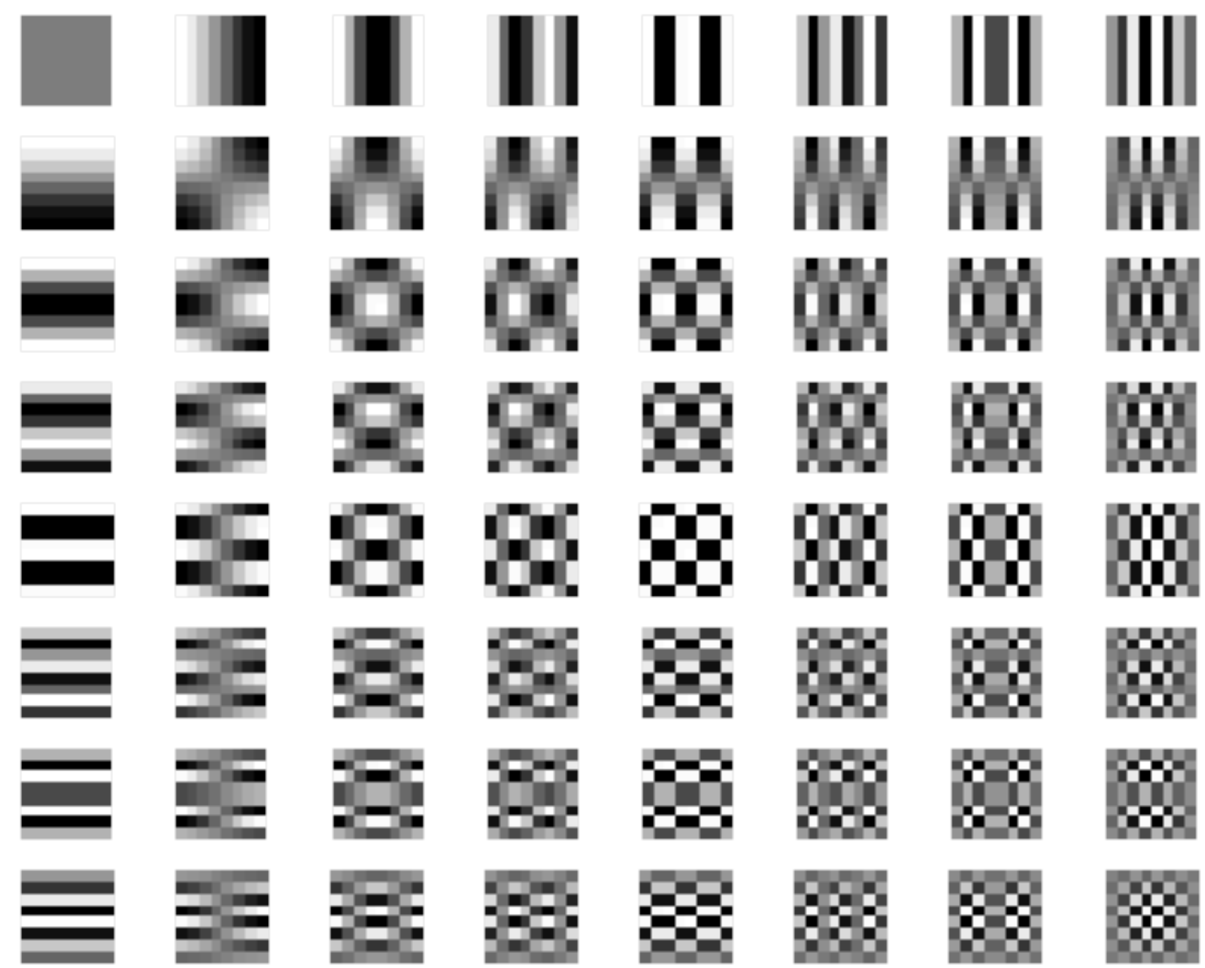


Discrete Cosine Transform

- Basis functions

$$X(m,n) = \alpha_m \alpha_n \sum_{k=0}^{K-1} \sum_{l=0}^{L-1} x(k,l) \cos\left[\frac{(2k+1)m\pi}{2K}\right] \cos\left[\frac{(2l+1)n\pi}{2L}\right]$$

- Energy is concentrated in the low frequency components
- Efficient algorithm to compute (similar to FFT)



8x8 basis functions

Coefficient Quantisation

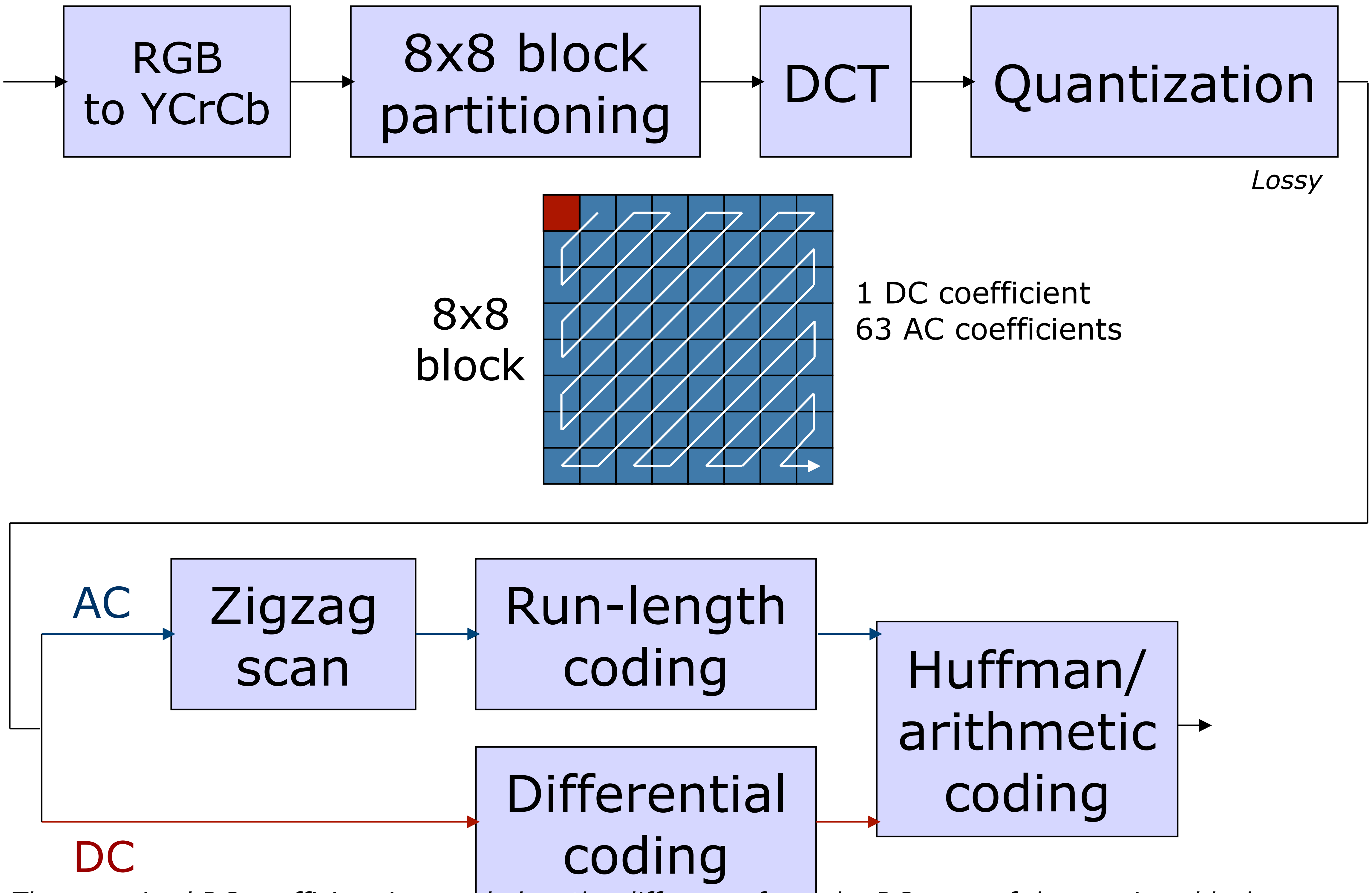
$$F^Q[u,v] = Round\left(\frac{F[u,v]}{Q[u,v]}\right)$$

- DCT coefficients F(u, v) are quantised according to a quantisation table
- High frequencies are less important (high factor)
- Quantisation table entries determine the "lossiness" of the compression

$Q[u,v]$

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

JPEG Compression



The quantized DC coefficient is encoded as the difference from the DC term of the previous block to account for the strong correlation between adjacent CD coefficient

Next Lecture

- Filtering and Pyramids
- Features and Matching