Working with Light

Working with Light for Computer Graphics

Physics/Optics:

- Light is emitted from a light source
  - e.g. the sun, a light bulb, computer monitor, cell phone, etc.
- That emitted light impacts various objects, where it may be reflected or absorbed
  - This reflection/absorption modifies the light
    - e.g. creating color, brightness, dullness/shininess, highlights, etc.
- In addition, light may pass (transmit) through a material and (in doing so) be bent, scattered, etc.
  - e.g. prism, stained glass windows, water, etc.

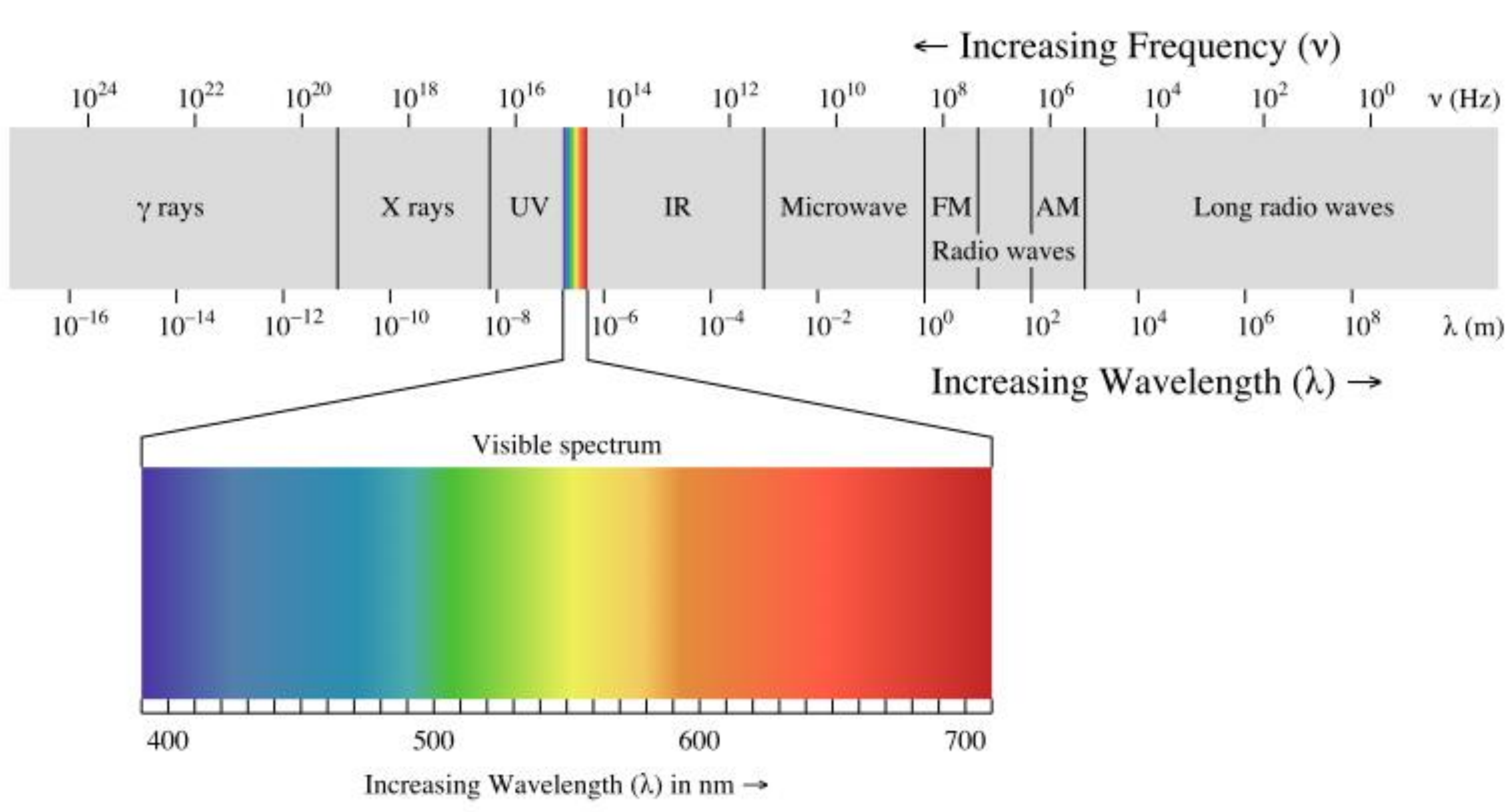
Human Perception:

- Eventually, some light enters your eyes creating a signal
- Your brain creates an image based on the signals it gets from your eyes

Software/Hardware:

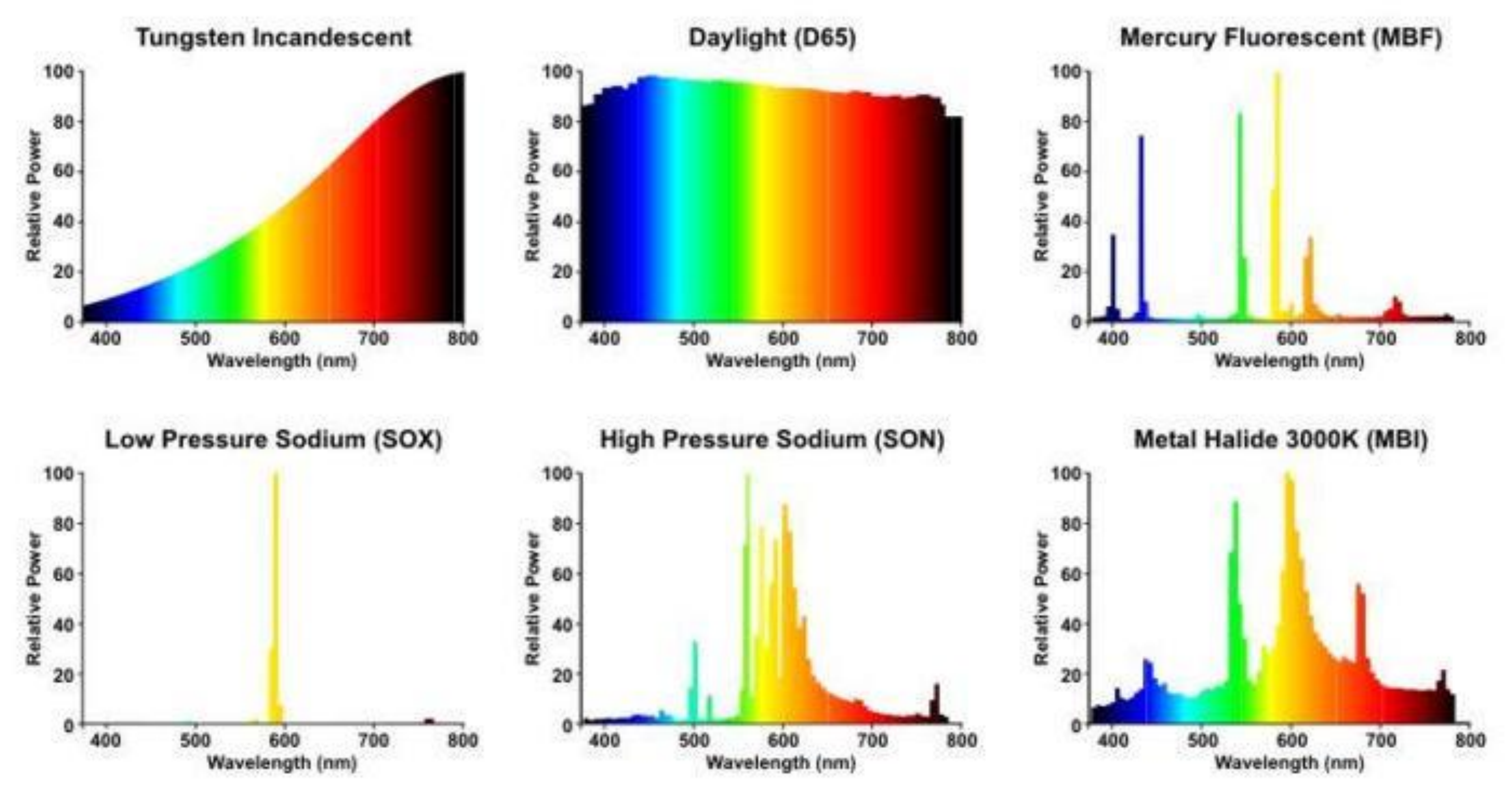
- Understanding the physics of light (i.e. optics) is important, in order to be able to work with it
- Understanding human perception allows for MANY optimizations/simplifications in both software/hardware
- The images we create ARE NOT intended to duplicate reality, only to fool humans into believing such

Electromagnetic Spectrum

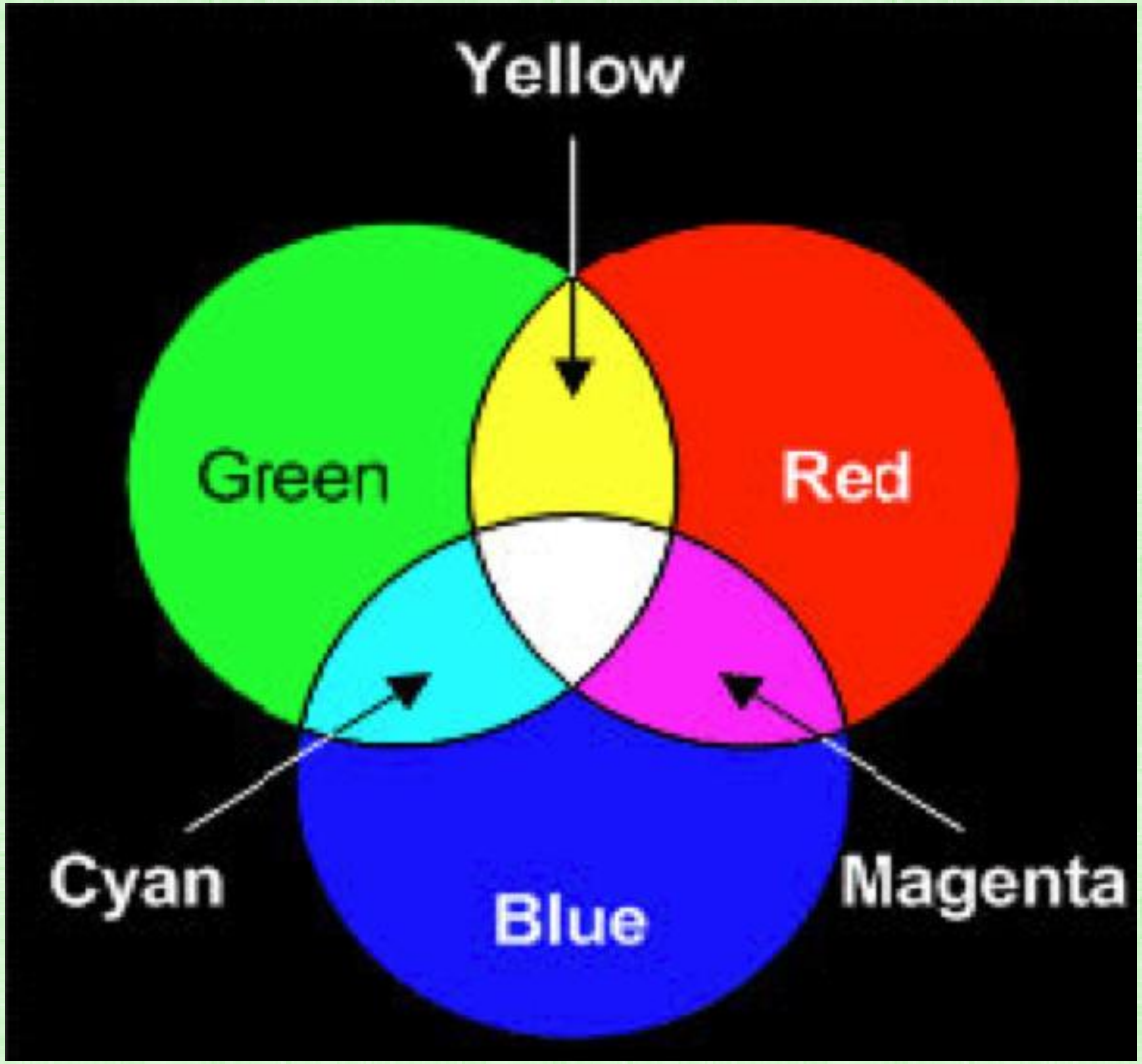


The human eye can only see wavelengths between about 400 nm to 700 nm, so we focus on those

Relative Power Distribution of Lights



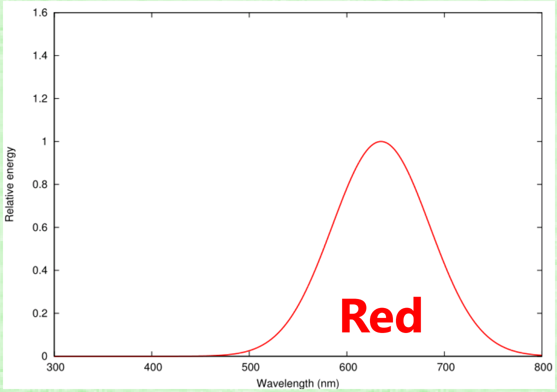
Adding Light Energy (human perception)



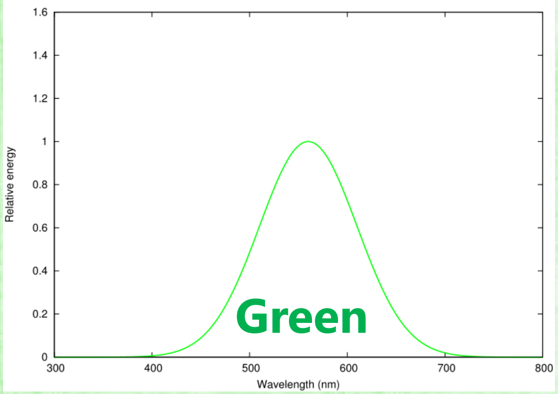
The human eye perceives combinations of light energy

Adding Light Energy (physics)

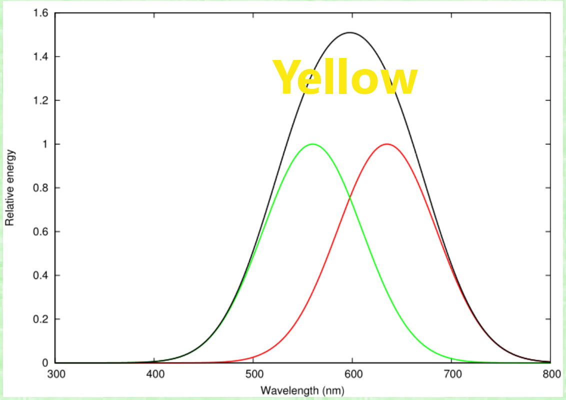
Energy is additive (per wavelength): $\boldsymbol{E(\lambda) = E_1(\lambda) + E_2(\lambda)}$

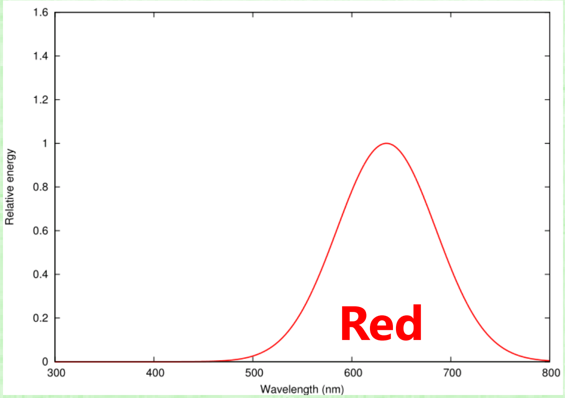


+

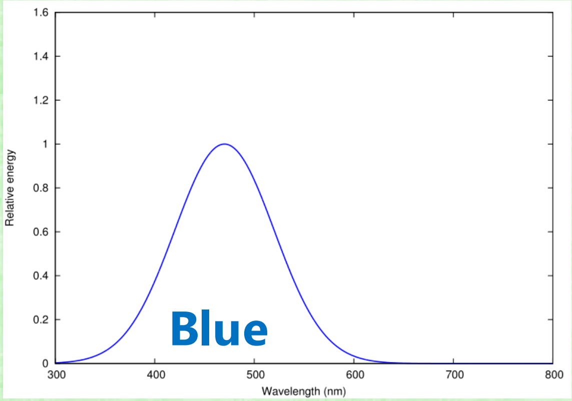


=

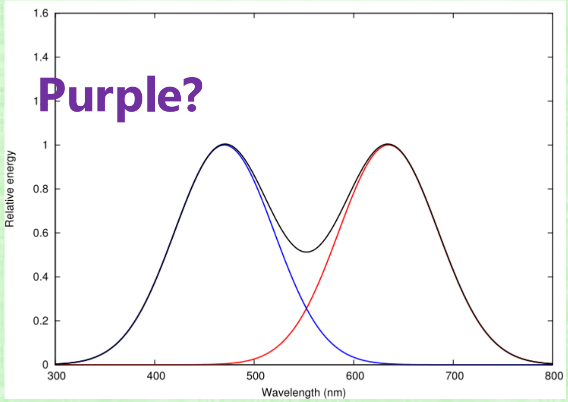




+

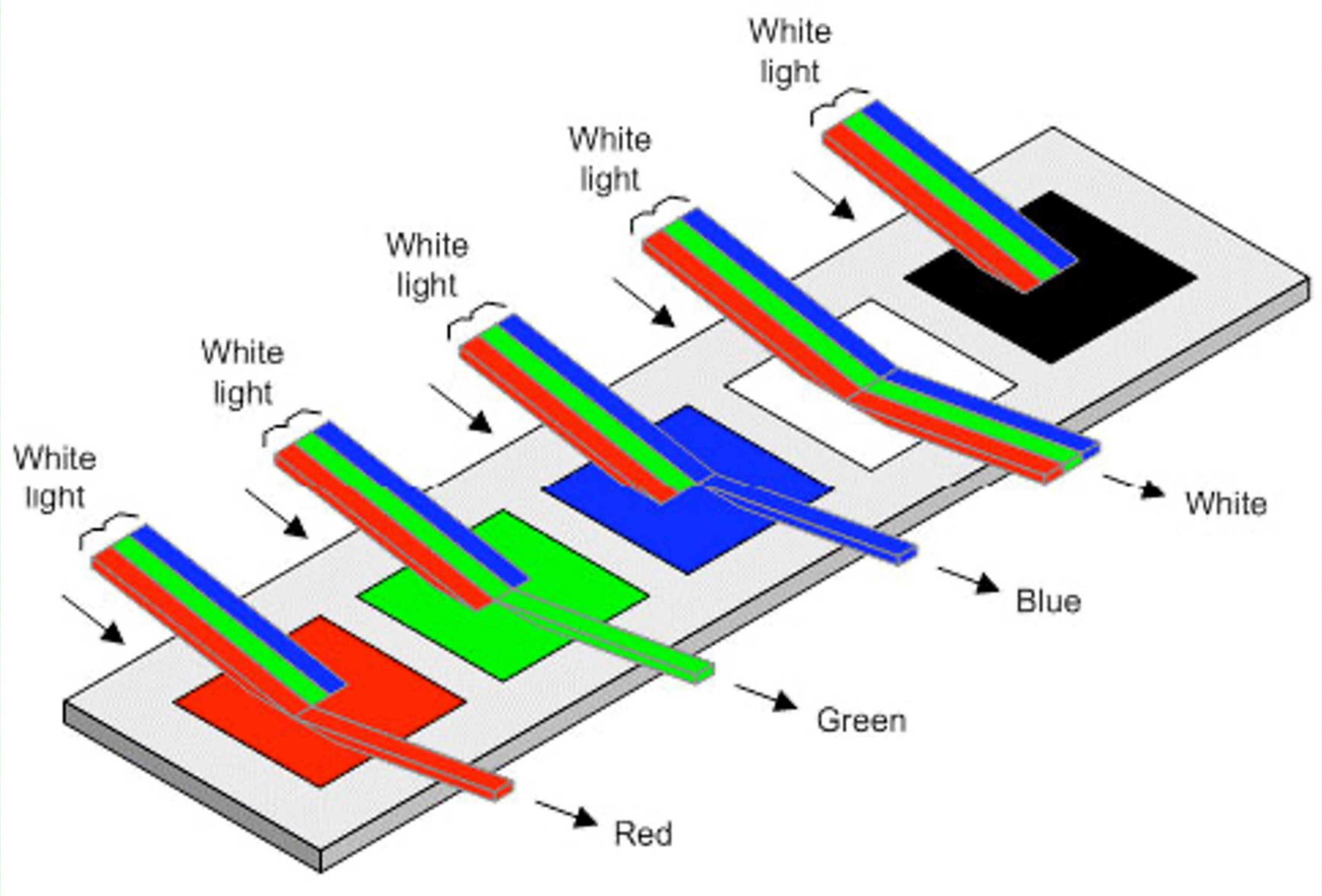


=



human perception & physics seem to disagree here

Absorbing & Reflecting Light Energy (human perception)



Shining white light on different colored surfaces

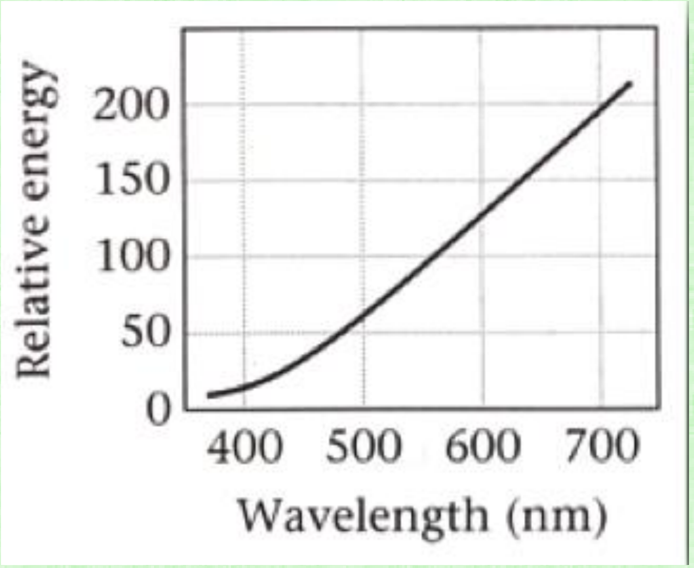
Absorbing & Reflecting Light Energy (physics)

Light energy is either reflected or absorbed (per wavelength): $\mathbf{r(\lambda) + a(\lambda) = 1}$

$$\mathbf{0 \le r(\lambda), a(\lambda) \le 1}$$

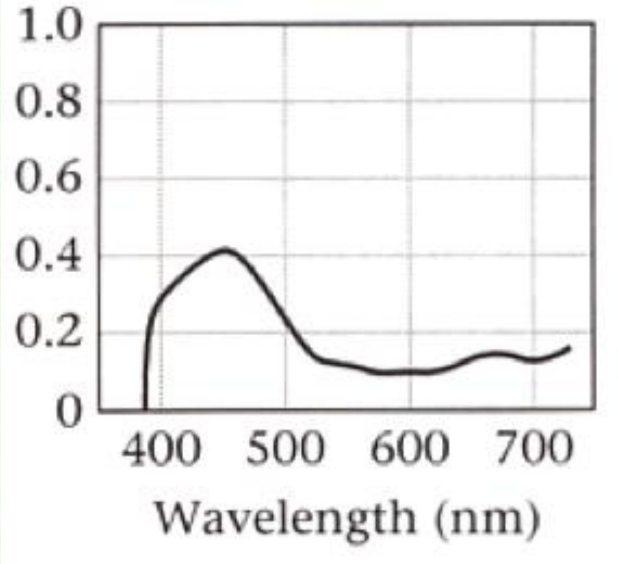
Reflected/Absorbed light energy (per wavelength): $\boldsymbol{Reflected(\lambda) = E(\lambda)r(\lambda)}$

Reflected/Absorbed light energy (per wavelength): $\boldsymbol{Absorbed(\lambda) = E(\lambda)a(\lambda)}$



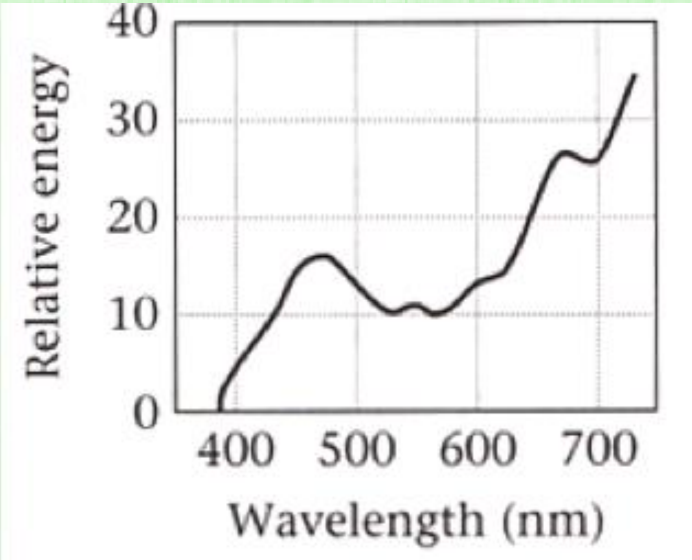
Incoming Energy

×



Surface Reflectance

=



Reflected Energy

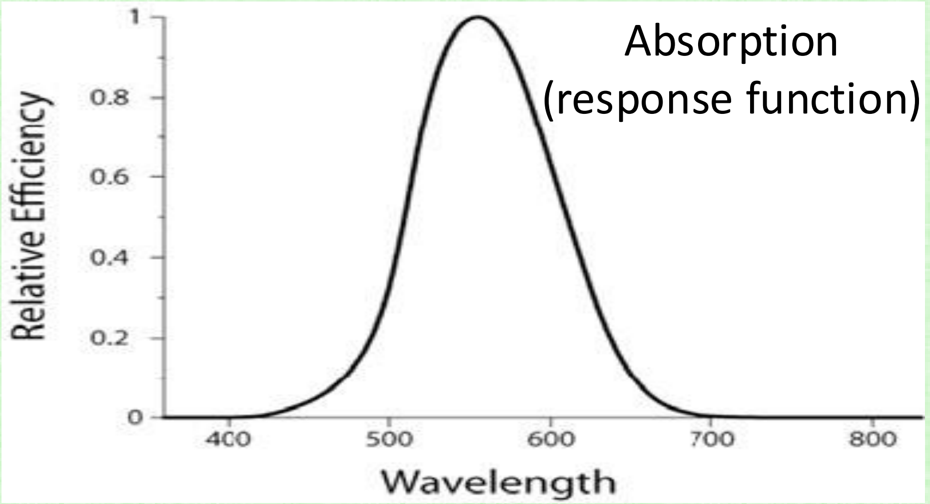
Sensor Absorption

Sensors absorb light (per unit time) and create a signal (per unit time)

- In order to be small (whether biological or mechanical), they end up being highly specialized
- This specialization leads to an entire sensor creating only one signal (per unit time)
- All the various incoming wavelengths of light are converted into just one signal (per unit time)
- Absorbed light loses its wavelength information!

Signal power (energy per second):

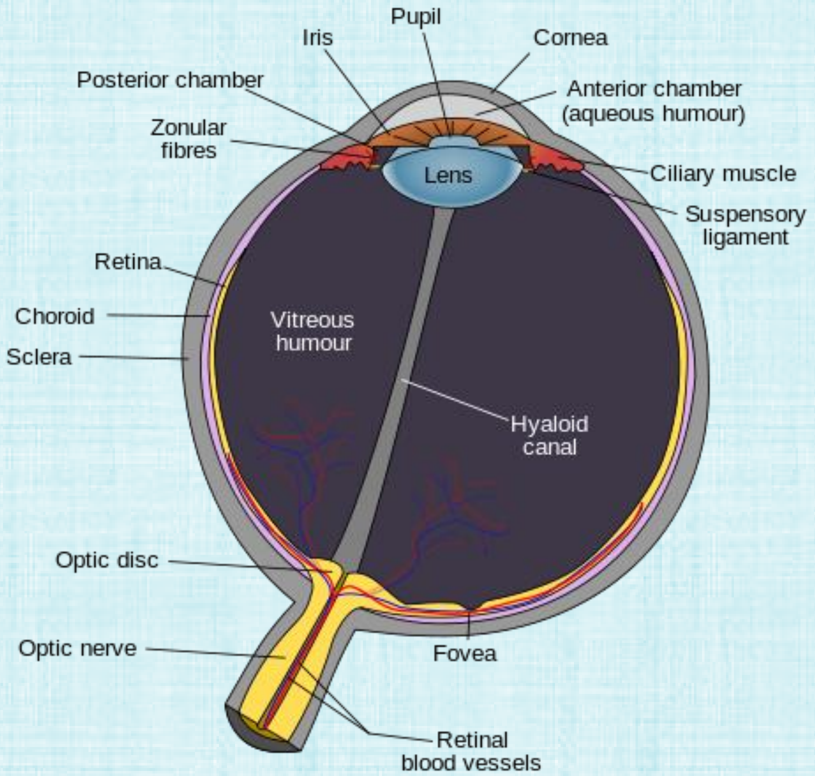
$$P = \int E(\lambda)a(\lambda)d\lambda$$

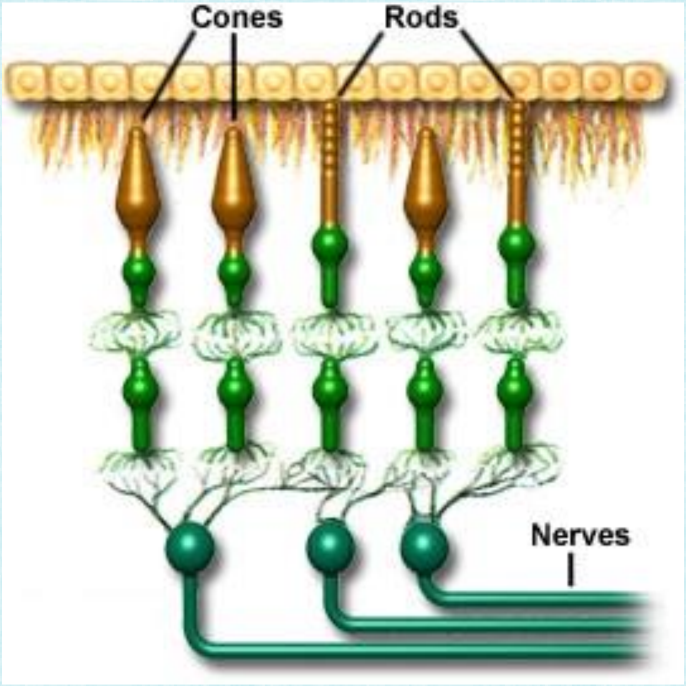


Not all wavelengths contribute equally to the final signal

Sensors in the Human Eye

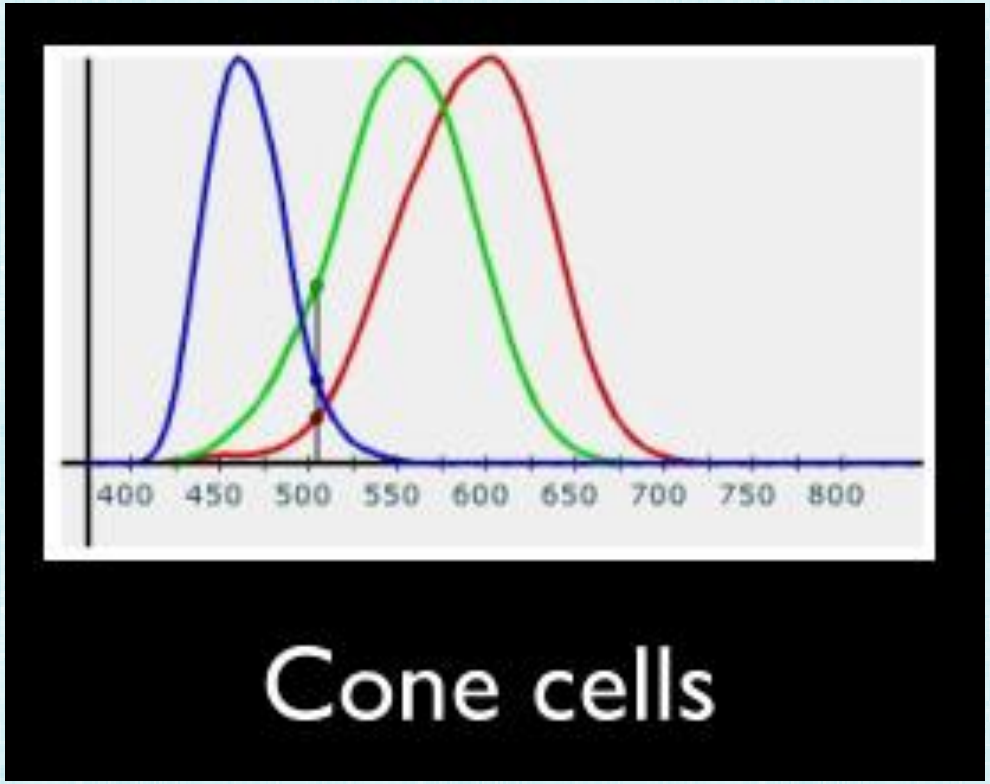
- The eye has 4 different kinds of sensors: 3 types of cones, 1 rod
- Proteins in the cone/rod cells absorb photons changing the cell membrane potential
- At <u>night</u>, cones are under-saturated (no/low/noisy signal), and rods produce most of the understandable signal
- During the <u>day</u>, the rod signals are over-saturated (maxed out), and we see primarily with cones



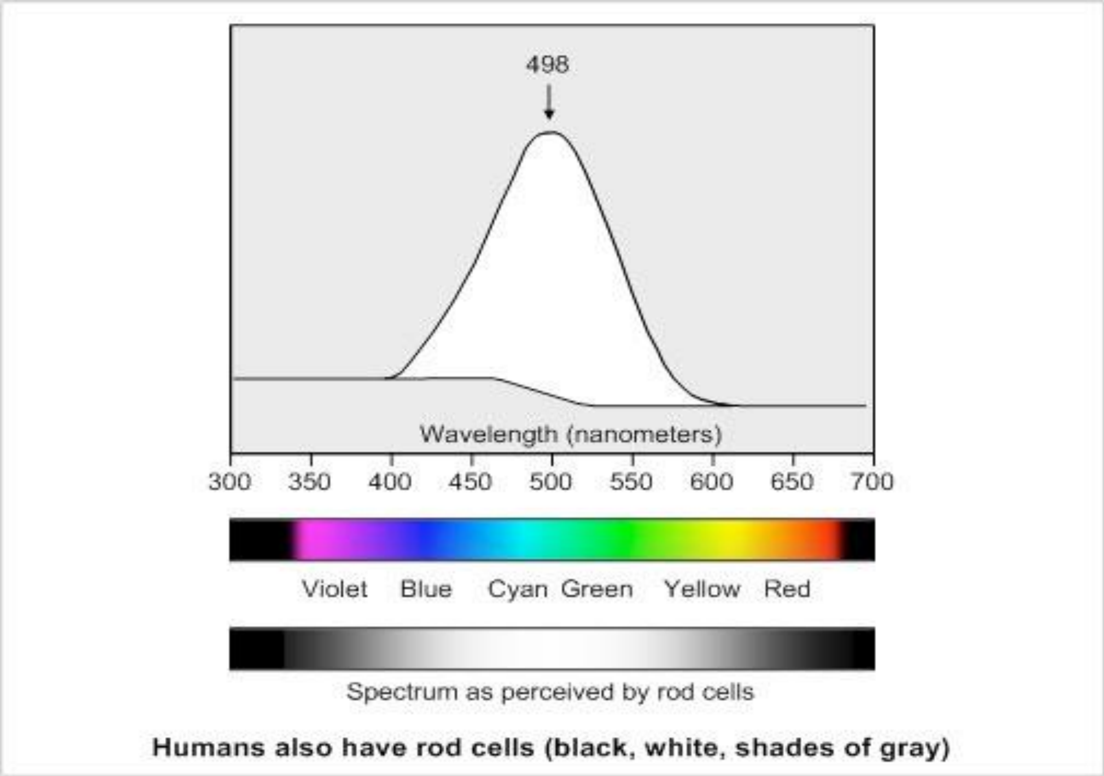


Response Functions for Cone/Rod Sensors

- The cone response functions vary based on the type of cone (red/green/blue)
- The rod sensor is interpreted as a gray-scale intensity (from black to white)



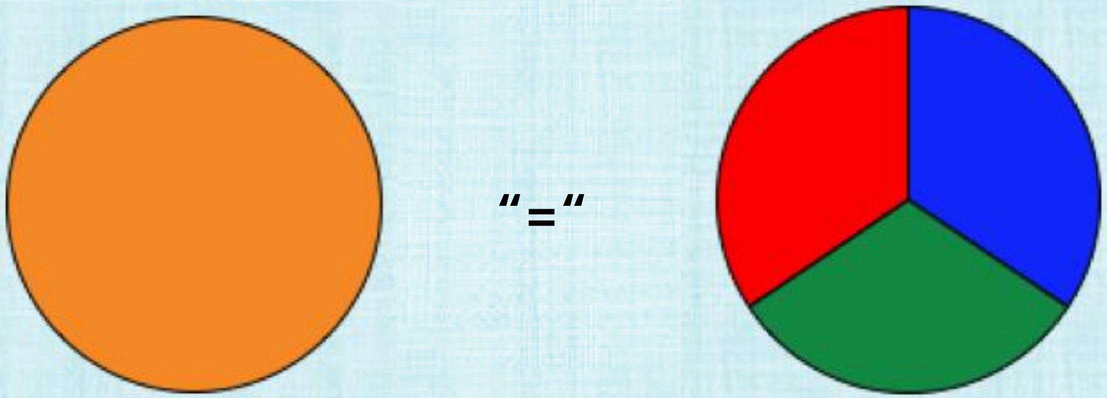
Note the similarity in red/green
(regarding red/green colorblindness)



At night, the single signal from rods is interpreted as a shade of gray

Trichromatic Theory

- Given any human perceived “color” or any single wavelength of visible light:
  - Can adjust the brightness of 3 single wavelength lasers (e.g. R = 700 nm, G = 546 nm, B = 435 nm) to fool a human observer into “mistakenly” thinking that the laser combination exactly matches that “color”
  - This is doable because each of the 3 cones can only send one signal (i.e., a 3-dimensional basis)

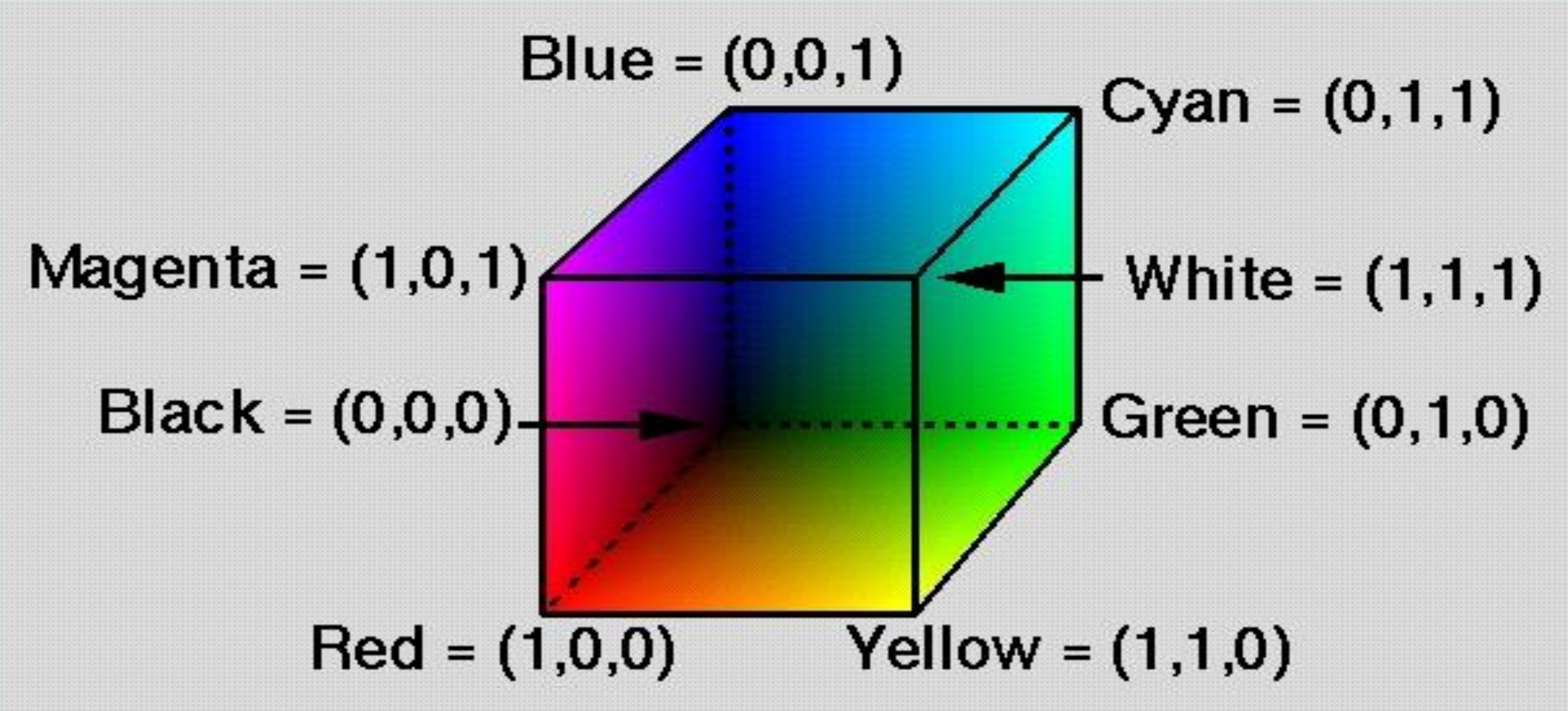


C **“=“** **R “+” G “+” B**

- Thus, only 3 signals are required for images, cameras, printers, displays, etc.
- Image formats store values in 3 channels: R, G, B

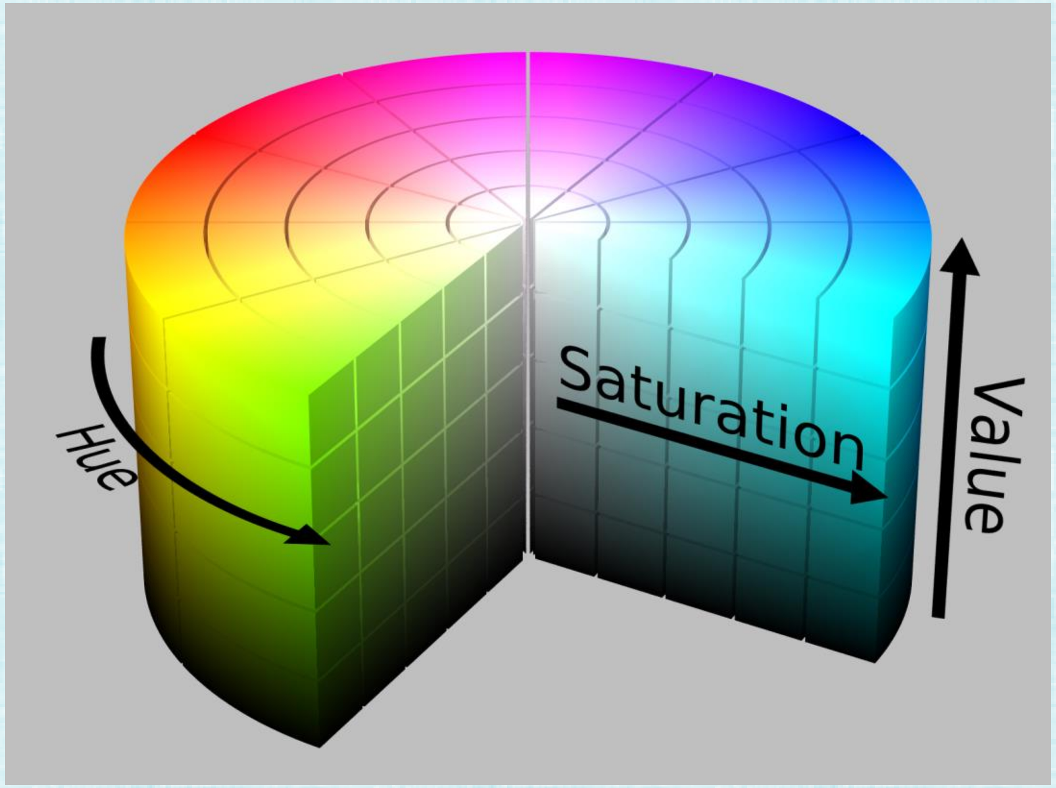
3D Color Space

- Map each primary color (Red, Green, Blue) to the unit distance along the x, y, z axes
- Black at (0,0,0), White at (1,1,1)
- The resulting RGB Color Cube represents all possible colors



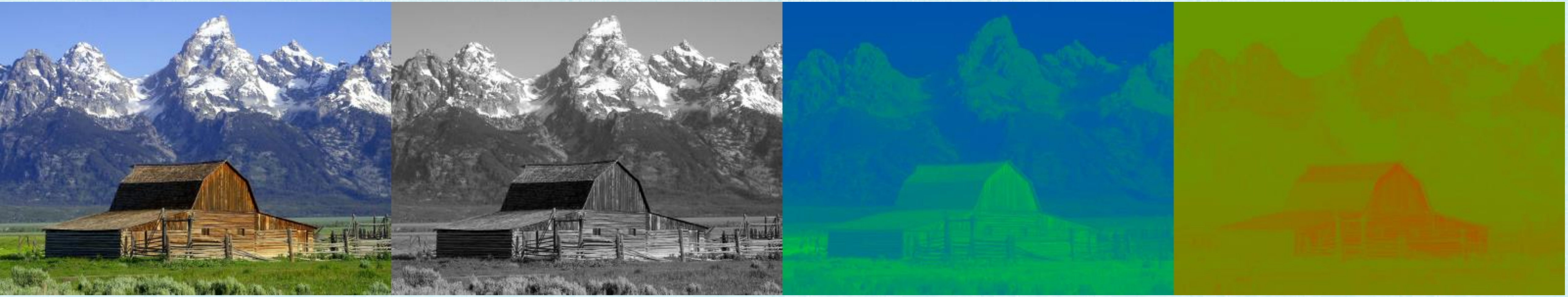
Cylindrical HSV Color Space

- A better 3D color space for <u>user interfaces</u> is based on Hue, Saturation, and Value (HSV)
- <u>Hue:</u> rainbow of colors (“wavelength”)
- <u>Saturation:</u> intensity for a particular color (“amount” of the color blended in with white)
- <u>Value:</u> lightness or darkness of a particular color



Luminance and Chrominance (YUV)

- Another 3D color space:
  - 1 luminance (Y) channel, 2 chrominance (UV) channels
- Black and White televisions used Y only, which perceptually holds the most spatial details
- Can compress more aggressively in U & V than in Y



Original Y U V

Interchangeability of 3D Color Spaces

- Can typically (as long as the mapping is linear) map back and forth between 3D color spaces via matrix multiplication, using an appropriate matrix and its inverse
- For example:

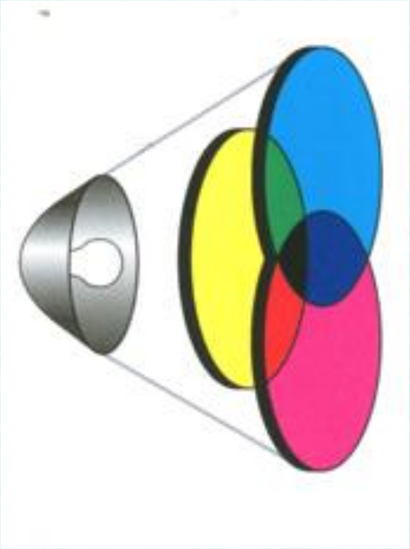
$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} .299 & .587 & .114 \\ -.14713 & -.28886 & .436 \\ .615 & -.51499 & -.10001 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

- Aside: note how important the Green channel is for the details in Y, as well as how unimportant the Blue channel is for those detail

Additive vs. Subtractive Color Spaces

Additive:
- Superimposed colored lights (e.g. smartphone display)
- Add spectra (wavelength by wavelength)
- R + G + B = white

Subtractive:
- Sequence of color filters (e.g. ink pigments or paint)
- Multiply by all absorption coefficients (wavelength by wavelength)
- R + G + B = black



Printers (CMYK)

- Printers use subtractive color spaces
- Ink partially or entirely masks/filters/absorbs colors on a white background, reducing the light that would otherwise be reflected

- Cyan, Magenta, Yellow (CMY) are the three primary colors of the subtractive color model

- Equal mixtures of C, M, Y would ideally produce all shades of gray
  - However, in practice, ink mixtures do not give perfect grays
  - In addition, it's difficult to get perfect alignment of the 3 inks
- Thus, most fine details are printed with the Key color (= K = black)
  - This also reduces ink bleeding, reduces the time for ink to dry, and saves money on colored ink

Limited Spatial Resolution

- Sensors require non-zero surface area, and each sends only one signal (per unit time)
- This limits the number of signals per square inch (based on how closely sensors are packed together)

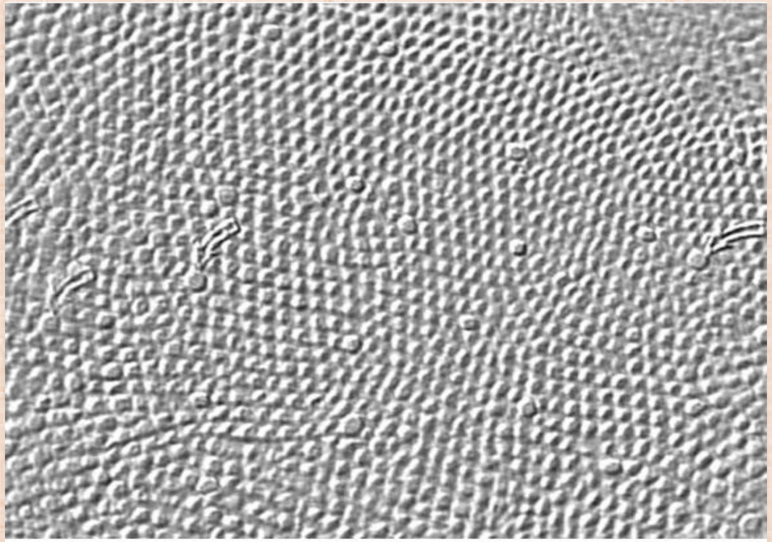
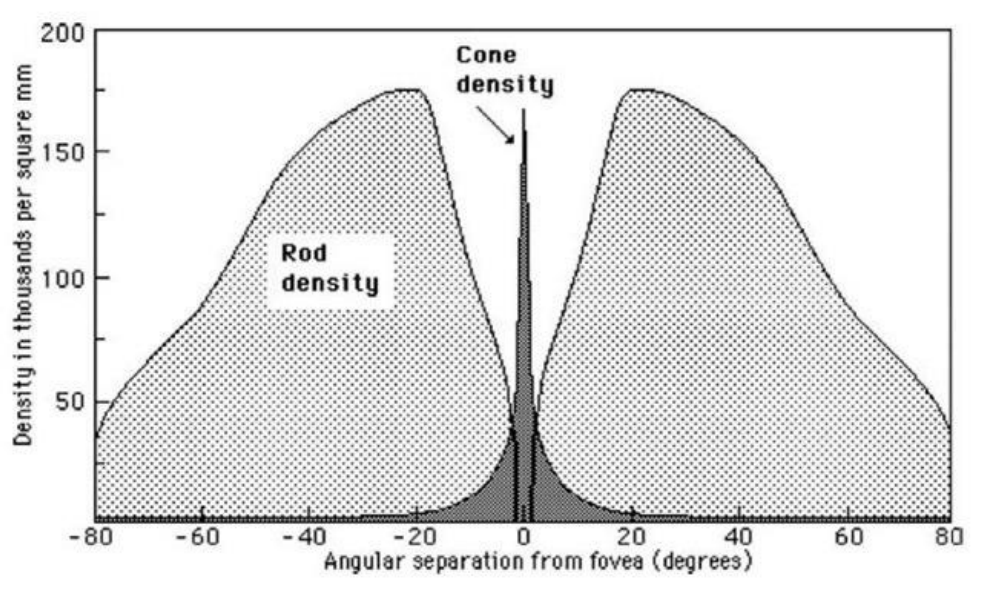


Fig. 13. Tangential section through the human fovea. Larger cones (arrows) are blue cones.



- Cones are densely packed near the center of the retina (the fovea), giving maximum detail for whatever the eye is looking directly at
- Rods have almost zero density at the fovea, which is why astronomers look out of the “side” of their eye

Distance Matters

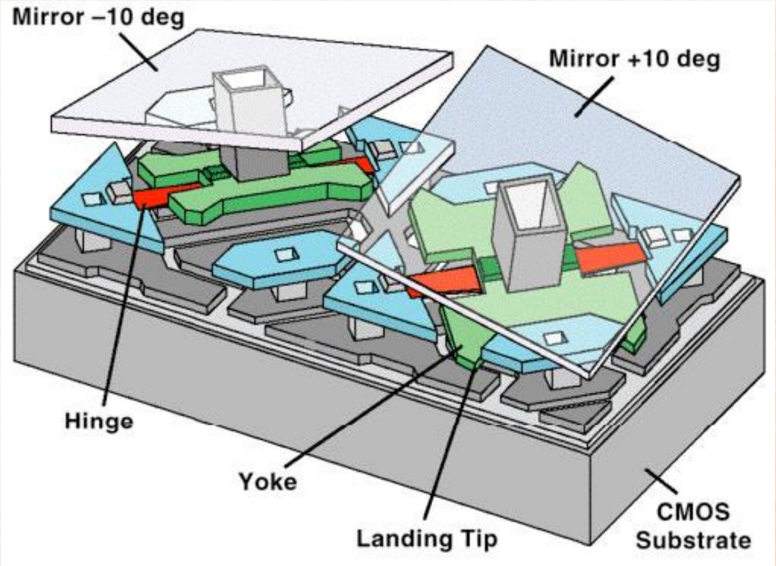
- Closer/farther away objects project onto larger/smaller areas on the cones, meaning that more/less cones receive light signals from the object
- Thus, closer objects can be seen in higher spatial detail than farther away objects

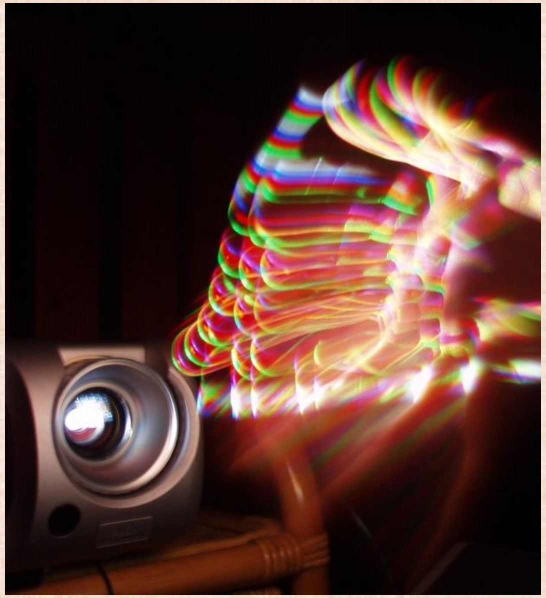
Resolution: 2048x1080
Size: 13.7m diagonal
4.29 dots per inch (dpi)

- A far away cinema screen requires lower resolution (4.29 dots per inch, or dpi) than a closely viewed cell phone (300-500 pixels per inch, or ppi)
- The number of cones per (image) feature is comparable between cinema screens and cell phones, given the differing distance of the observer

Projectors

- Making very large displays for far away viewers is difficult; so, projectors are important

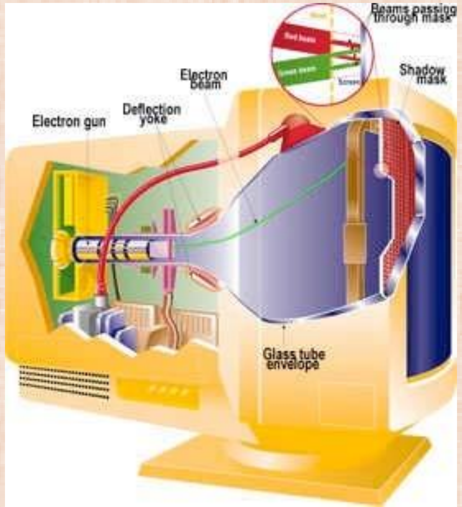




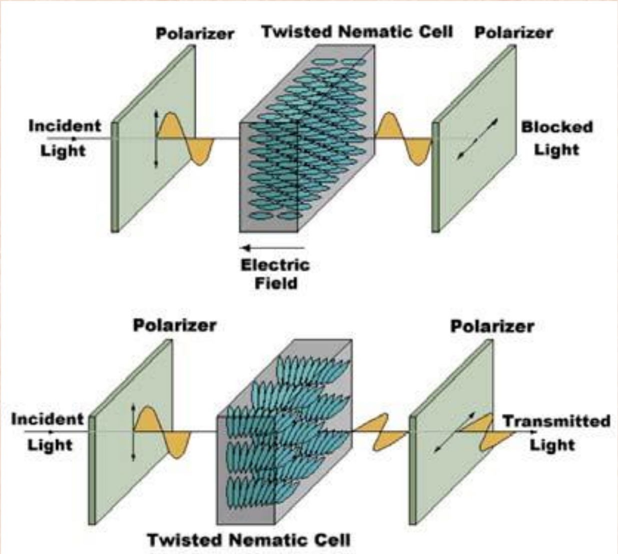
- A Digital Micro-Mirror Device (DMD) is the core component in Digital Light Processing (DLP) projectors
- Each mirror corresponds to one pixel, and has two states; it can either reflect the light into or out of the "pupil" of the projector
- Rapidly toggling a mirror between these two states produces brighter/dimmer light, controlled by the ratio of on-time to off-time

Display Technology

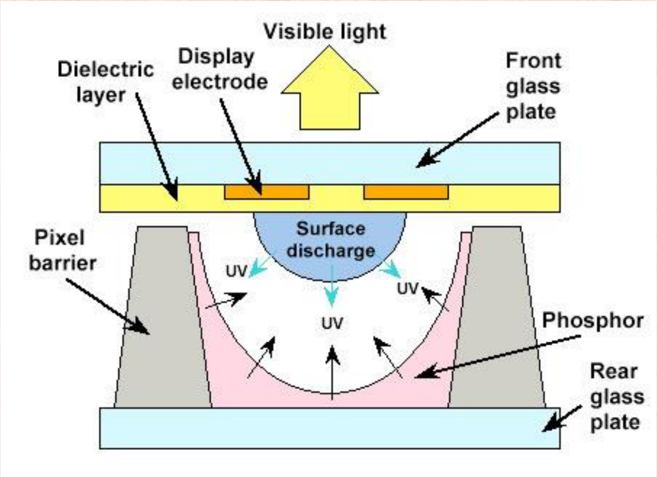
- Significant efforts have been spent on improving display (spatial) resolution, for close-up viewers (who can see more details)



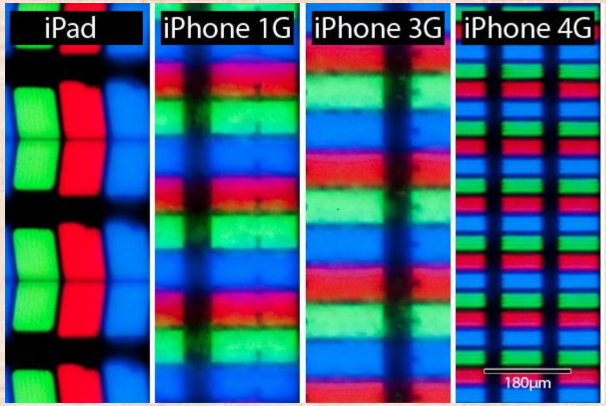
cathode ray tube (CRT)



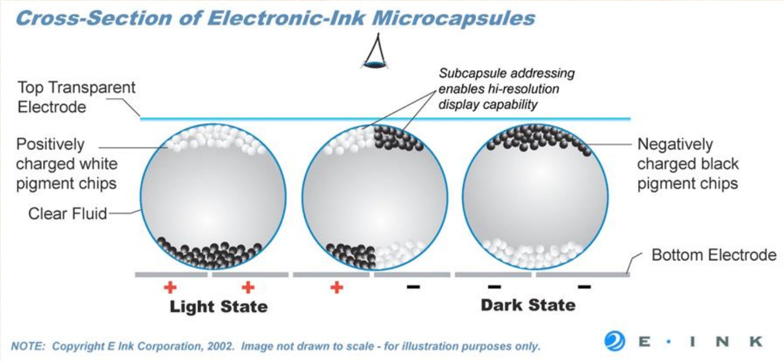
liquid crystal display (LCD)



plasma



iPhone/iPad LCD

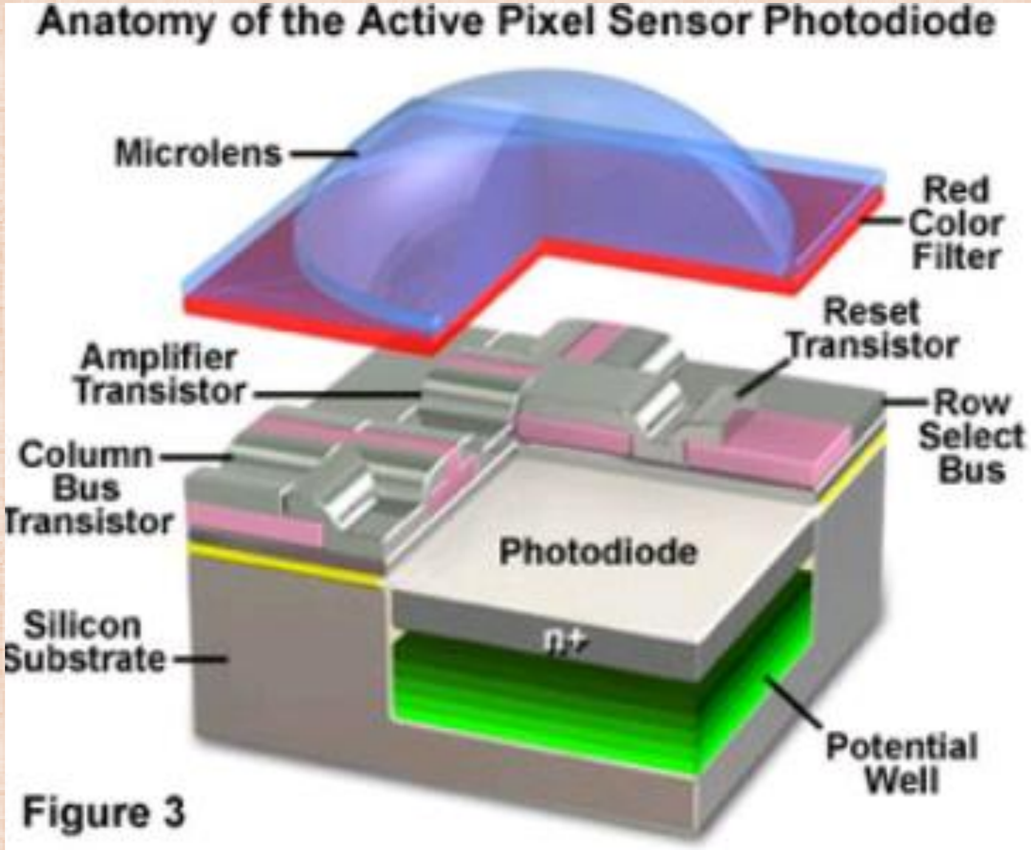


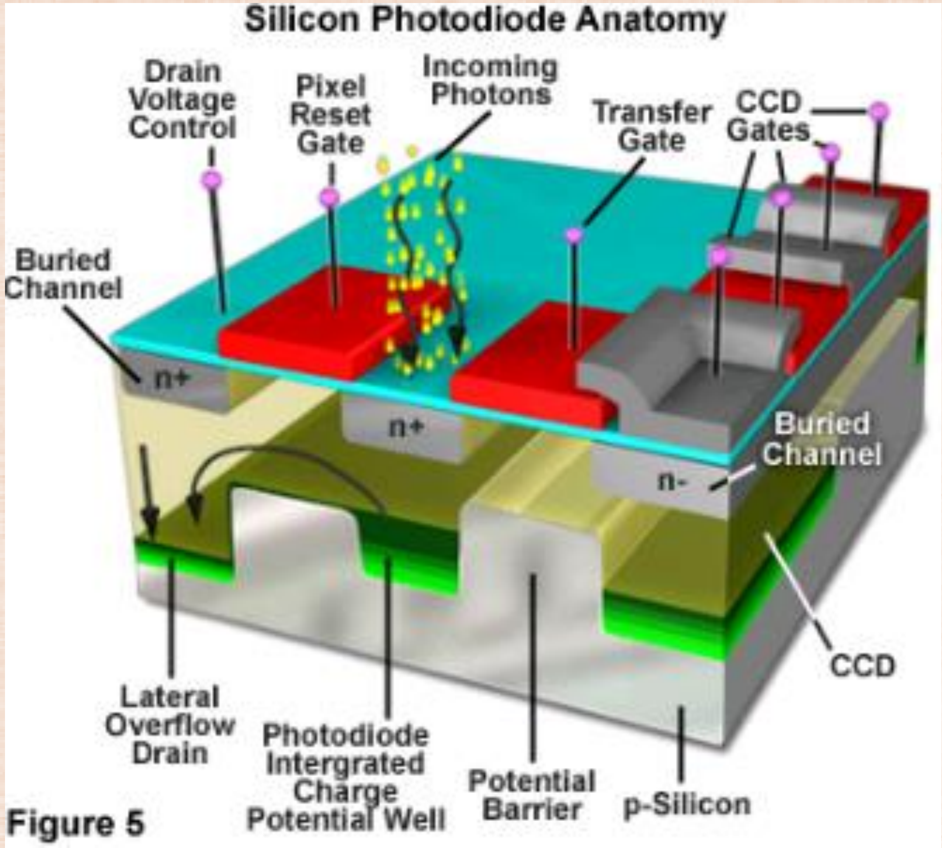
Electronic ink, ebook readers

Display Technology

Camera Resolution

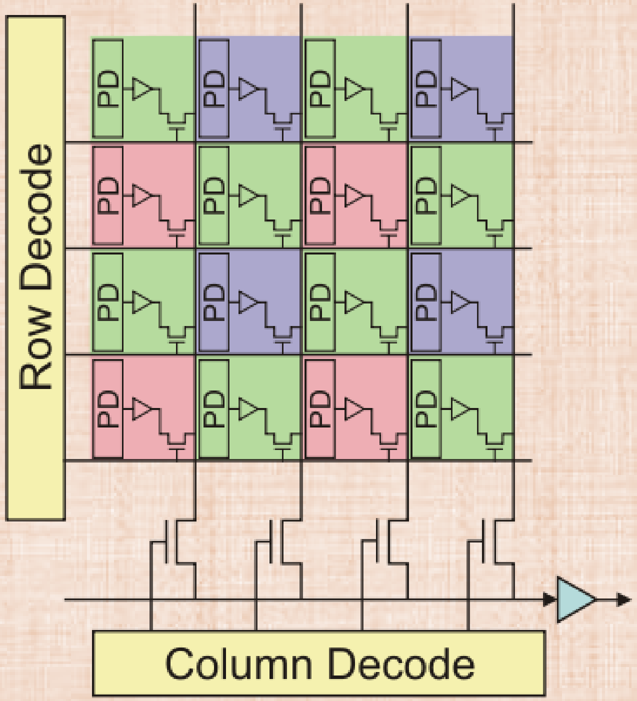
- Camera sensors use the photelectric effect to generate an electron when hit by a photon (with some efficiency/probability)
- They are quite complex, and (like cones) take up physical space
- This limits the resolution they (like cones) can capture



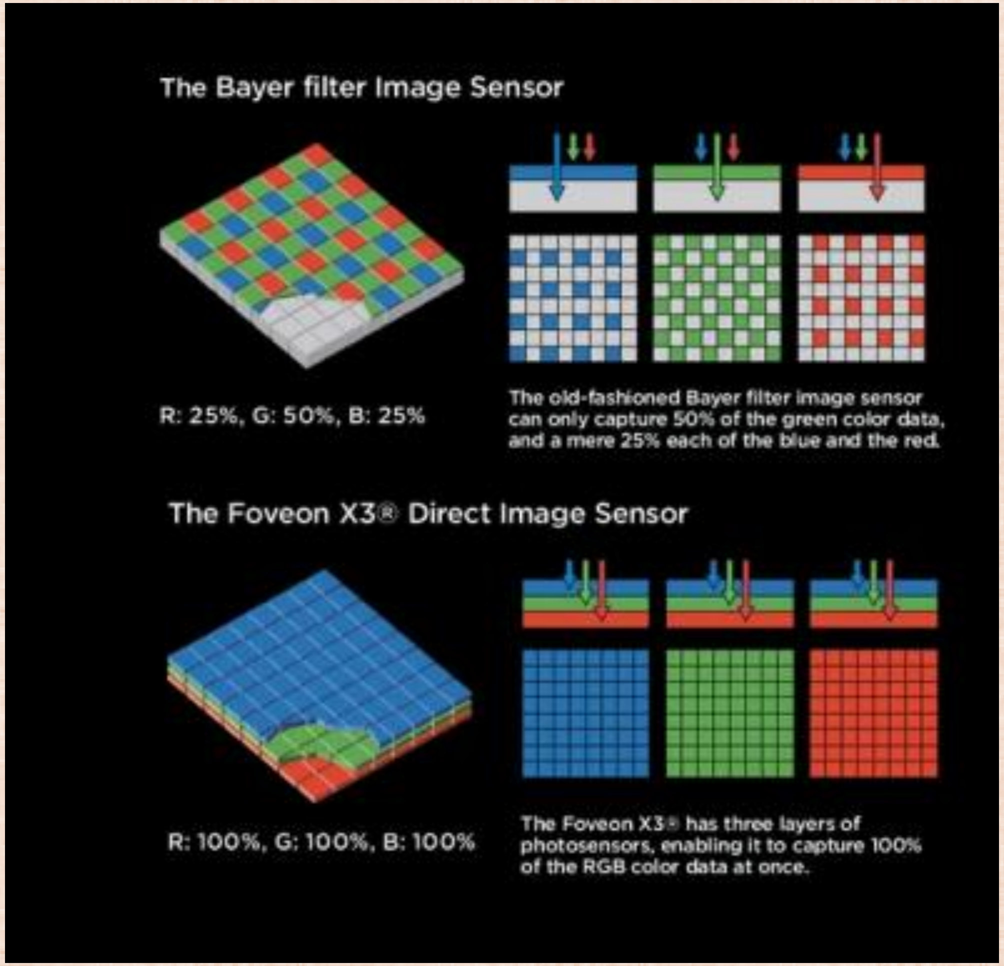


Camera Resolution

- Each camera sensor records incoming light energy per second (power)
- Each captures only one signal (per unit time) for its entire 2D spatial area
- Color filters limit incident light to a particular color (the same sensor is used for every color)



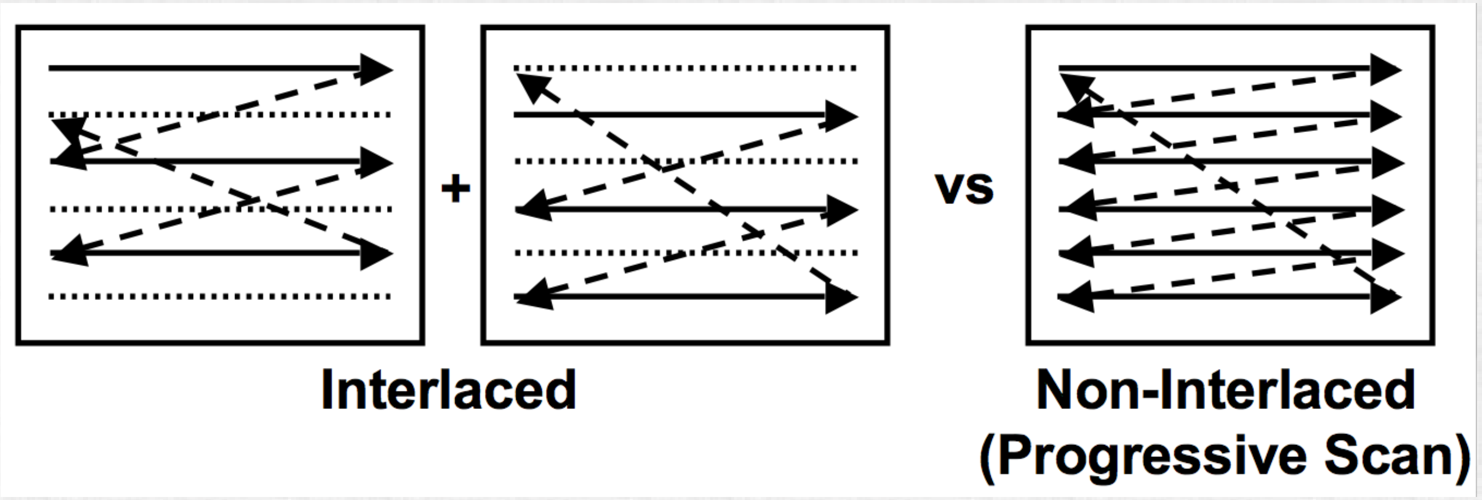
Note the doubled number of Green sensors, due to that color's importance in capturing spatial details



Aside: Temporal Resolution

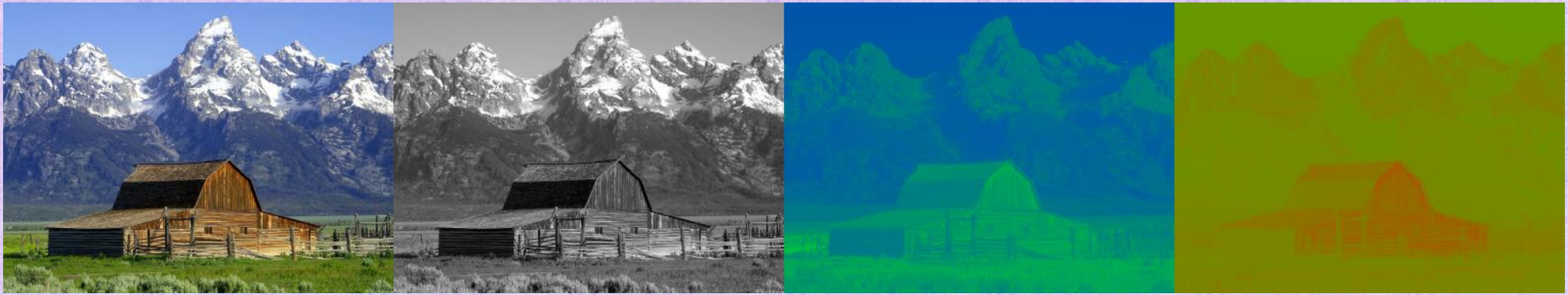
- For moving images (animations), at least 16 Hz is required for humans to *not* interpret them as a series of still images
  - Movies are recorded at 24 images per second
  - TV broadcasts have 30 images per second

- Even though motion seems continuous at 24-30 images per second, the brightness may still seem to flicker
  - Thus, movies are refreshed at 48 or 72 Hz (each image is projected 2 or 3 times)
  - Computer monitors refresh at 60-80 Hz (or more) independent of what is being displayed
  - TVs (used to) use interlacing to approximate 60 Hz, showing half of each frame at a time
- Flicker fusion threshold: frequency at which an intermittent light stimulus appears to be steady to the observer



Brightness (Luminance)

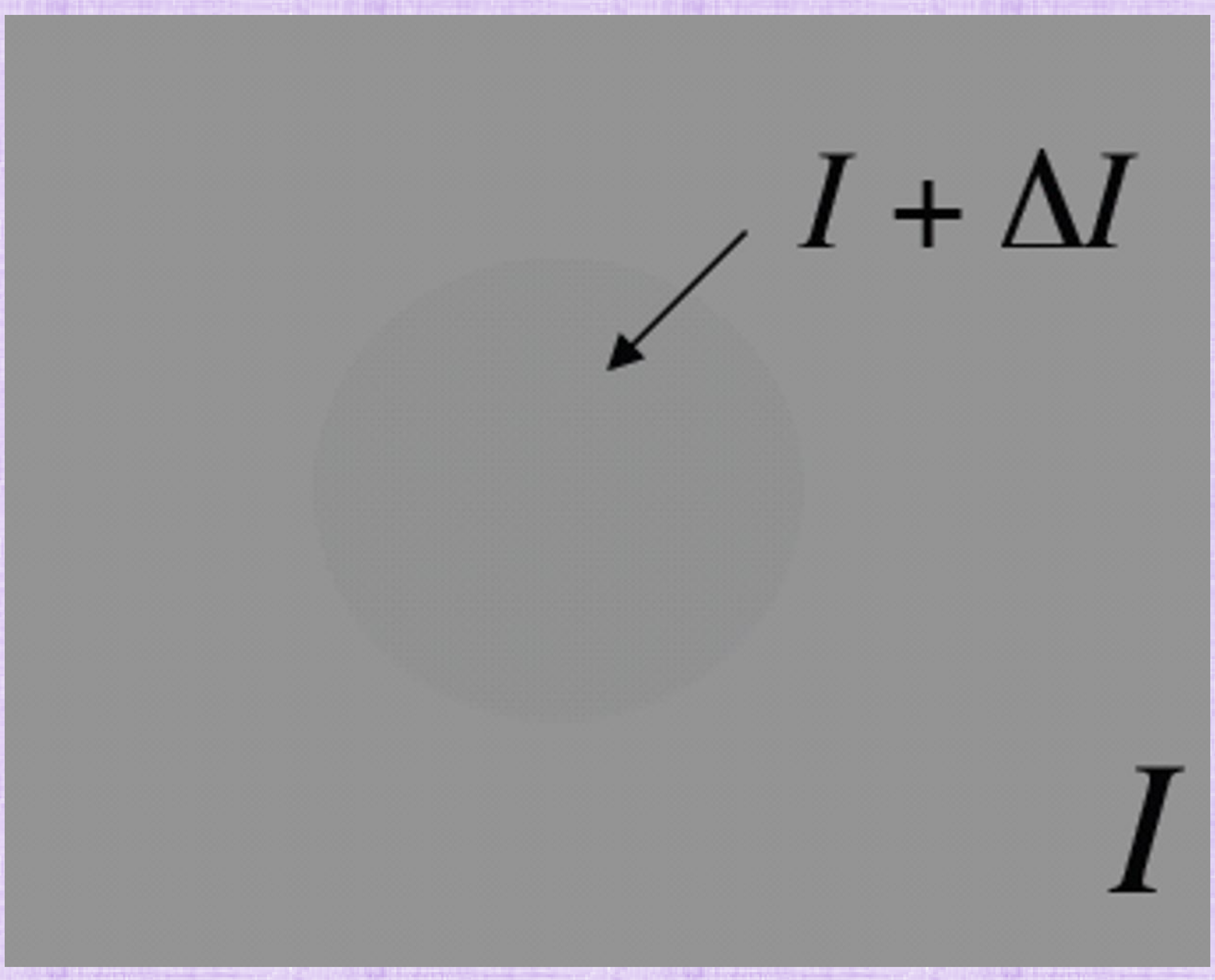
- The human eye is more sensitive to spatial variations in brightness (gray-scale) than to spatial variations in color
- The 3 images on the right add together to give the image on the left
- Notice which image on the right has the most spatial details



original = gray-scale + color 1 + color 2

Brightness Discrimination Experiment

- Changing the brightness (intensity) by 1 to 2% makes it just noticeable to most observers



Discretizing Brightness

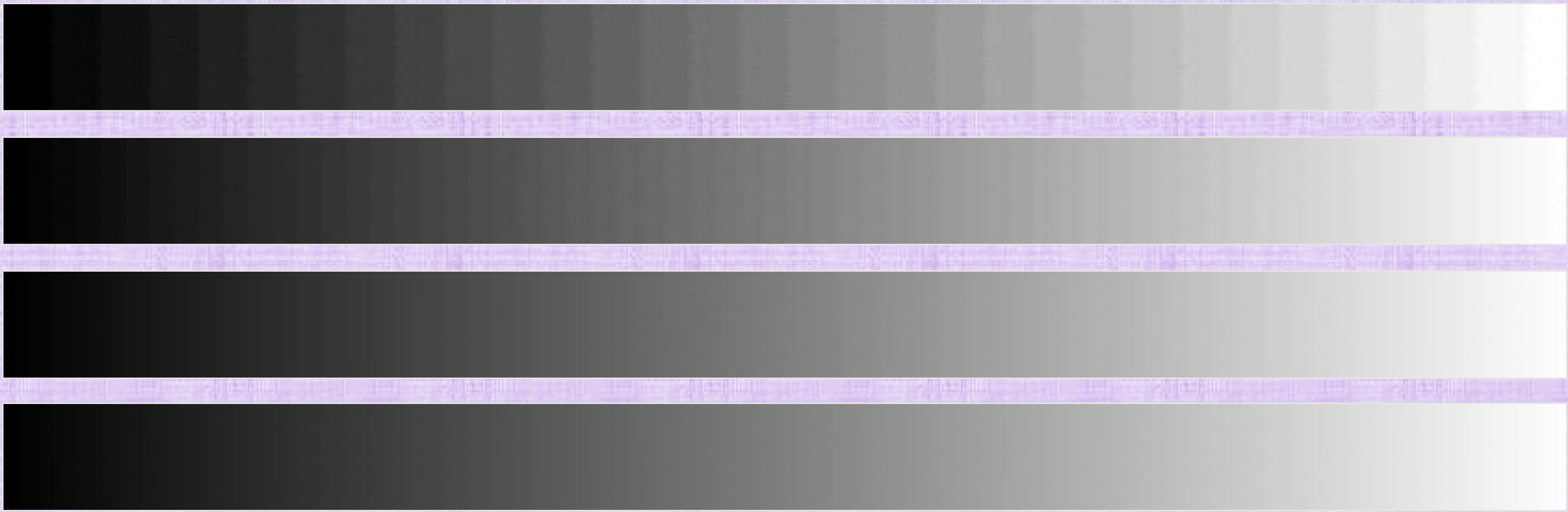
- Since our eyes can perceive small changes in brightness, we need many levels for brightness
  - Otherwise, changing brightness by the smallest amount looks discontinuous
- We typically use 256 levels for brightness
  - Each color channel (R, G, B) stores an integer ranging from 0 to 255
- High Dynamic Range (HDR) image formats use an even larger range than 0-255

32 levels

64 levels

128 levels

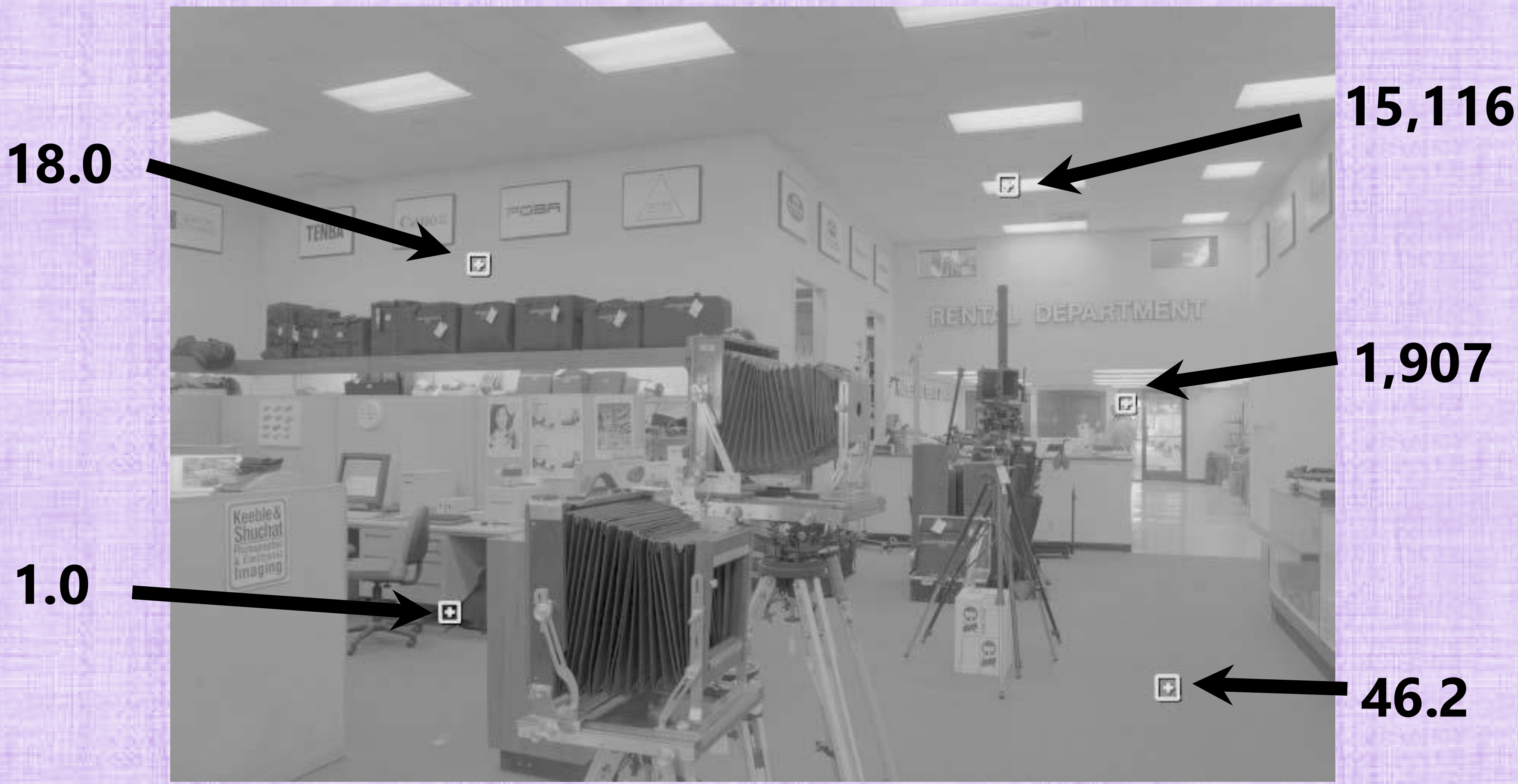
256 levels



Dynamic Range

- World:
  - Possible: 100,000,000,000:1 (from the sun to pitch black)
  - Typical real-world scenes: 100,000:1
- Human Eye:
  - Static: 100:1
  - Dynamic: 1,000,000:1 (as the eye moves, it adaptively adjusts exposure by changing the pupil size)
- Media:
  - Newsprint: 10:1, Glossy print: 60:1
  - Samsung F2370H LCD monitor: static 3,000:1, dynamic 150,000:1
    - Static contrast ratio: luminance ratio between the brightest white and darkest black in a single image
    - Dynamic contrast ratio: luminance ratio between brightest white possible (on any image) and the darkest black possible (on any image) on the same device
  - A monitor's contrast ratio is measured in a dark room
  - The effective contrast ratio typically drops from 3,000:1 to less than 200:1 in office lighting conditions

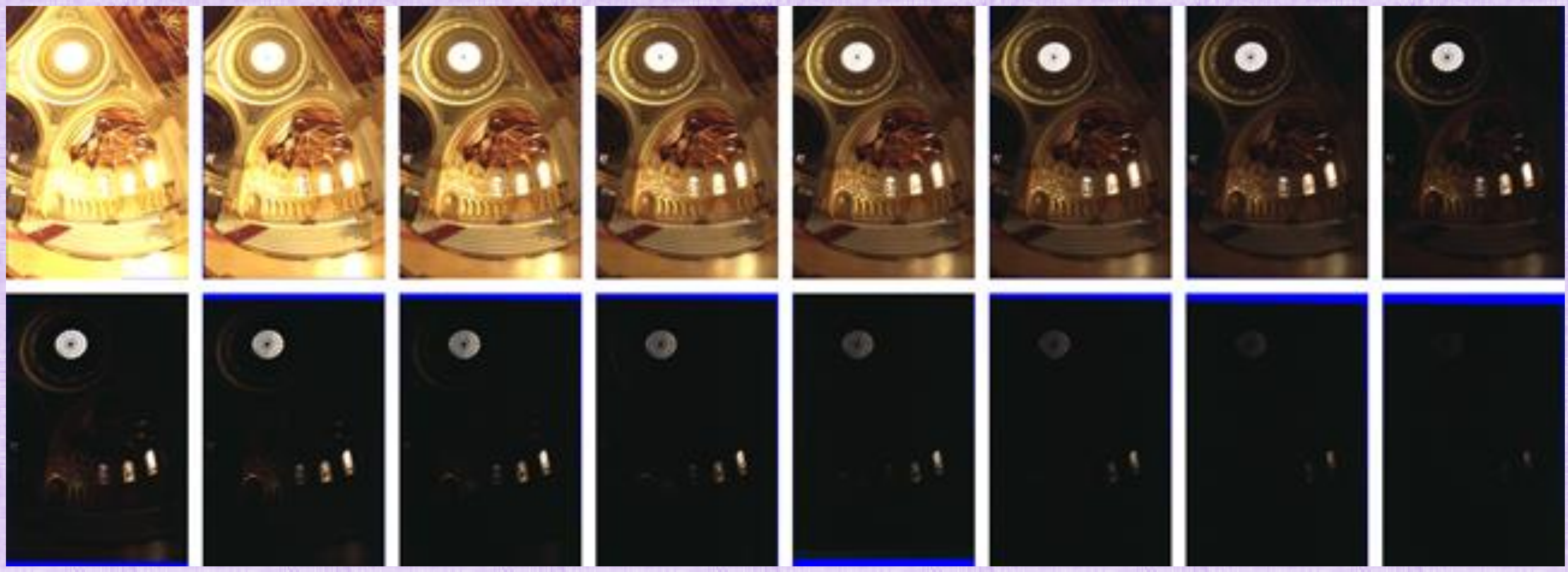
Dynamic Range (an example)



The labels refer to relative brightness

Dynamic Range (an example)

- 16 photographs of the Stanford Church
  - half as much light hits the film in each successive picture
- No image captures all the details in both the darkest and the brightest regions
  - some pixels oversaturate and others get no signal at all



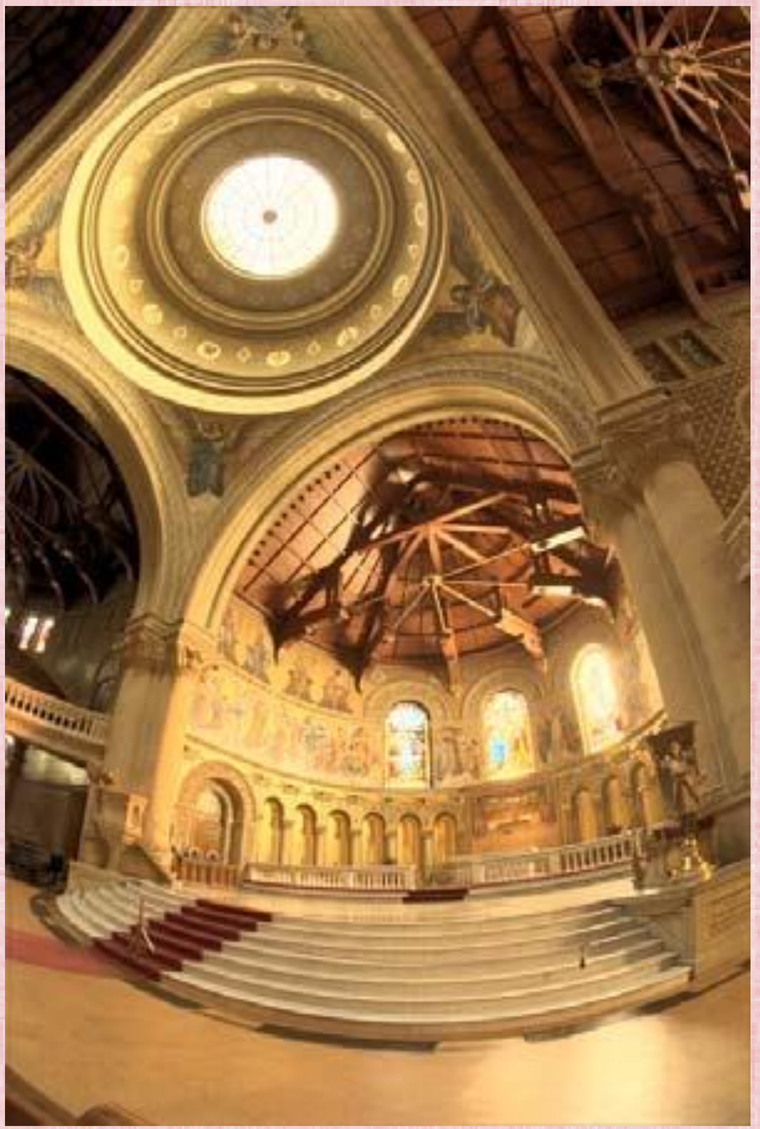
From Debevec and Malik, [High Dynamic Range Photographs]

Tone Mapping

- Compositing all the information from all the images gives a result with a High Dynamic Range (i.e., 0-X with X >> 255)
- That range is too large for the standard image format (since X > 255)
- Solution #1: Linearly rescale/compress the values so that X=255
  - Small intensity differences are quantized (a range of values map to the same integer), and relative differences (and thus details) are lost
- Solution #2: Use a logarithmic map to rescale/compress
  - Information is still quantized, but in a way that exploits human "perceptual space"
- Solution #3: Other approaches…
  - E.g., Local operators: map each pixel value based on surrounding pixel values (human vision is sensitive to local contrast)

Linear vs. Logarithmic Compression

Linear



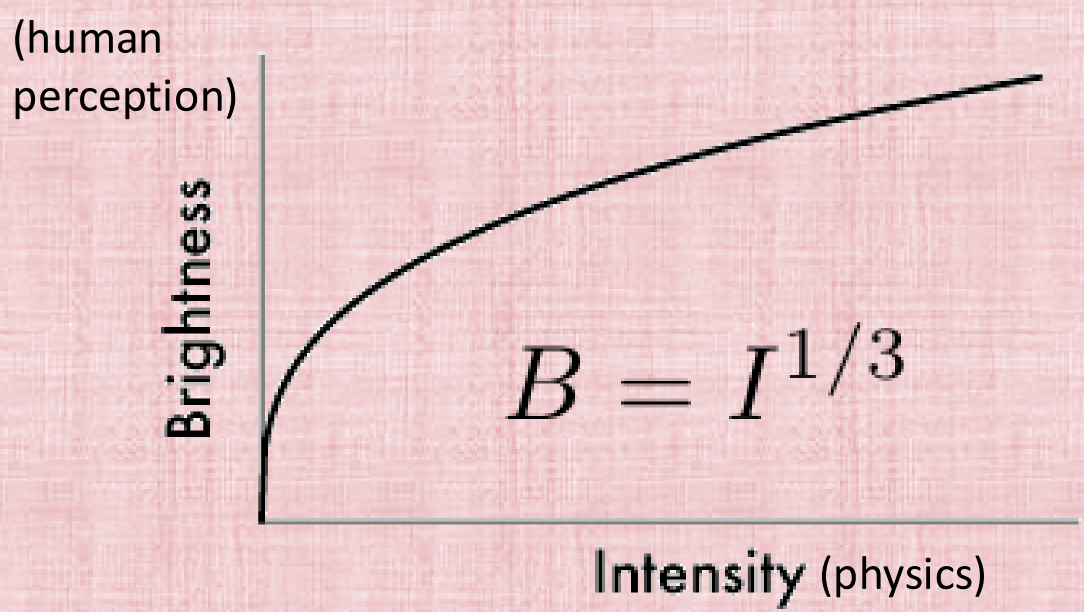
Logarithmic

Human Perception of Intensities

- Humans perceive brightness differences better at lower (as opposed to higher) light intensities

$$S = I^p$$

Sense	Exponent
Brightness	0.33
Loudness	0.60
Length	1.00
Heaviness	1.45



- Logarithmic compression uses more resolution for the more-important lower intensities in the image (and less resolution for the less-important higher intensities)
- This gives less quantization in the lower intensities than in the higher intensities, and is thus <u>more optimal for human consumption</u>

Gamma Encoding and Correction

- More bits are allocated to the lower intensity (darker) regions of the image (where the human eye is more sensitive) than to the higher intensity (brighter) regions (where the human eye is less sensitive)
- Gamma correction reverses the process, converting gamma encoded images back to their original brightness/luminance