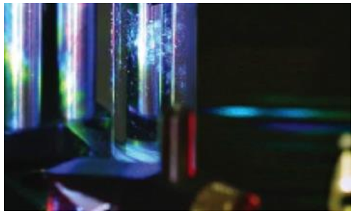
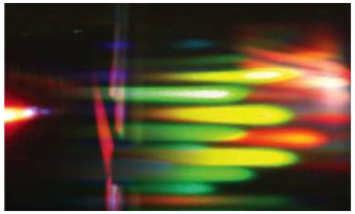
Imperial College London

COMP70058 Computer Vision
Lecture 2 – Image Formation

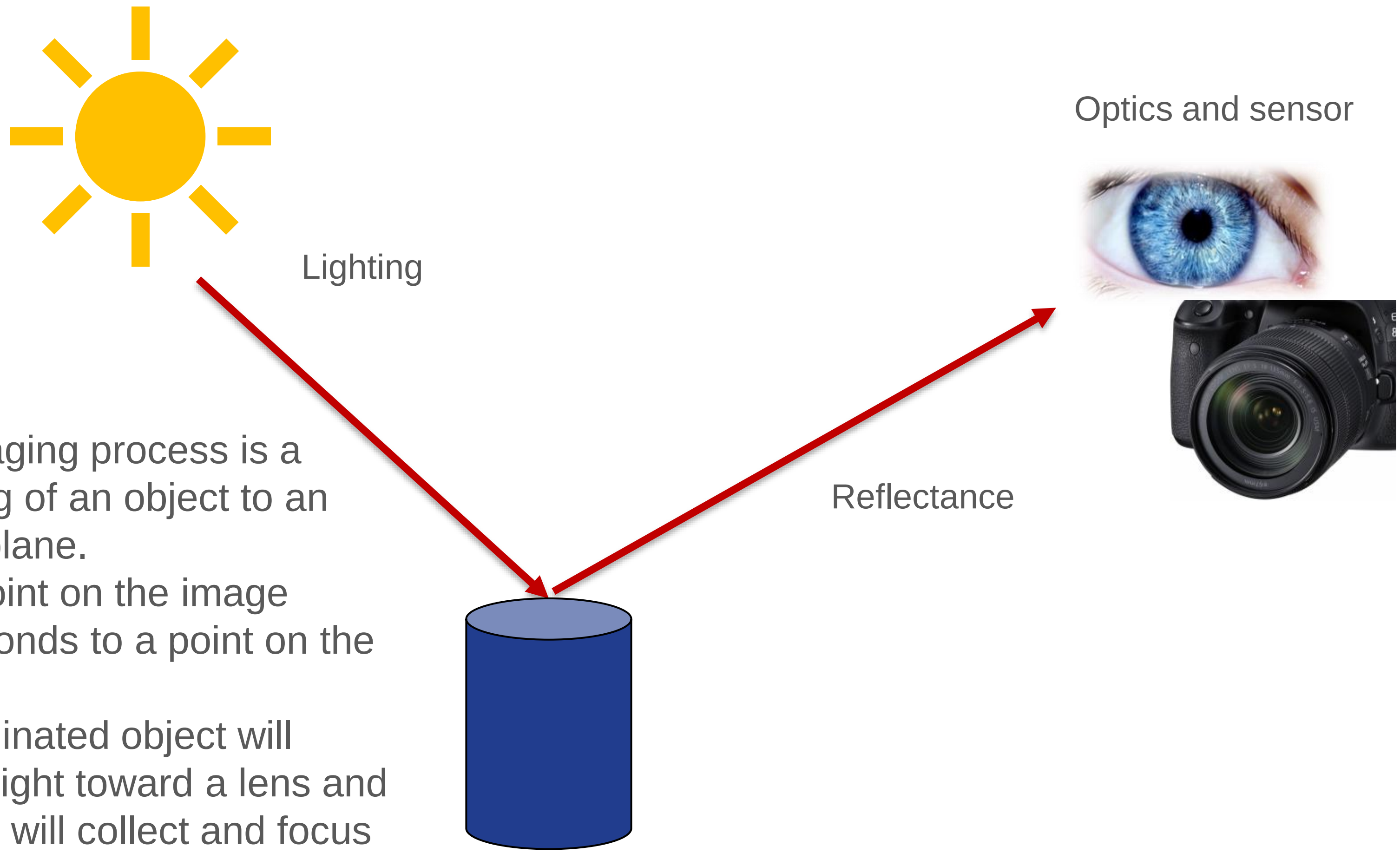

Stamatia (Matina) Giannarou, PhD
The Hamlyn Centre for Robotic Surgery

stamatia.giannarou@imperial.ac.uk

The Hamlyn Centre
for Robotic Surgery

Contents

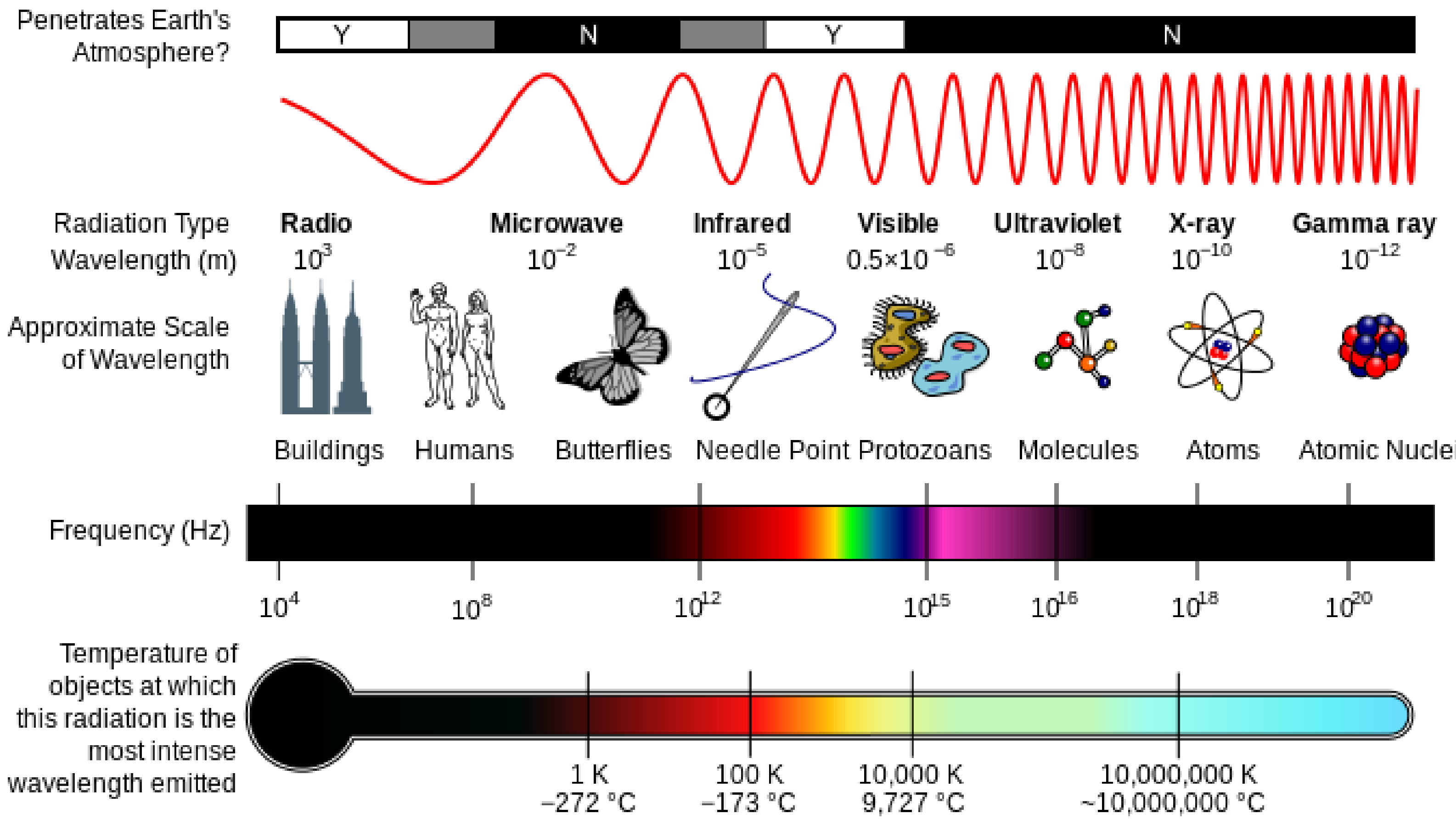
Photometric image formation



- The imaging process is a mapping of an object to an image plane.
- Each point on the image corresponds to a point on the object.
- An illuminated object will scatter light toward a lens and the lens will collect and focus the light to create the image.
- Images are made up of discrete intensity values.

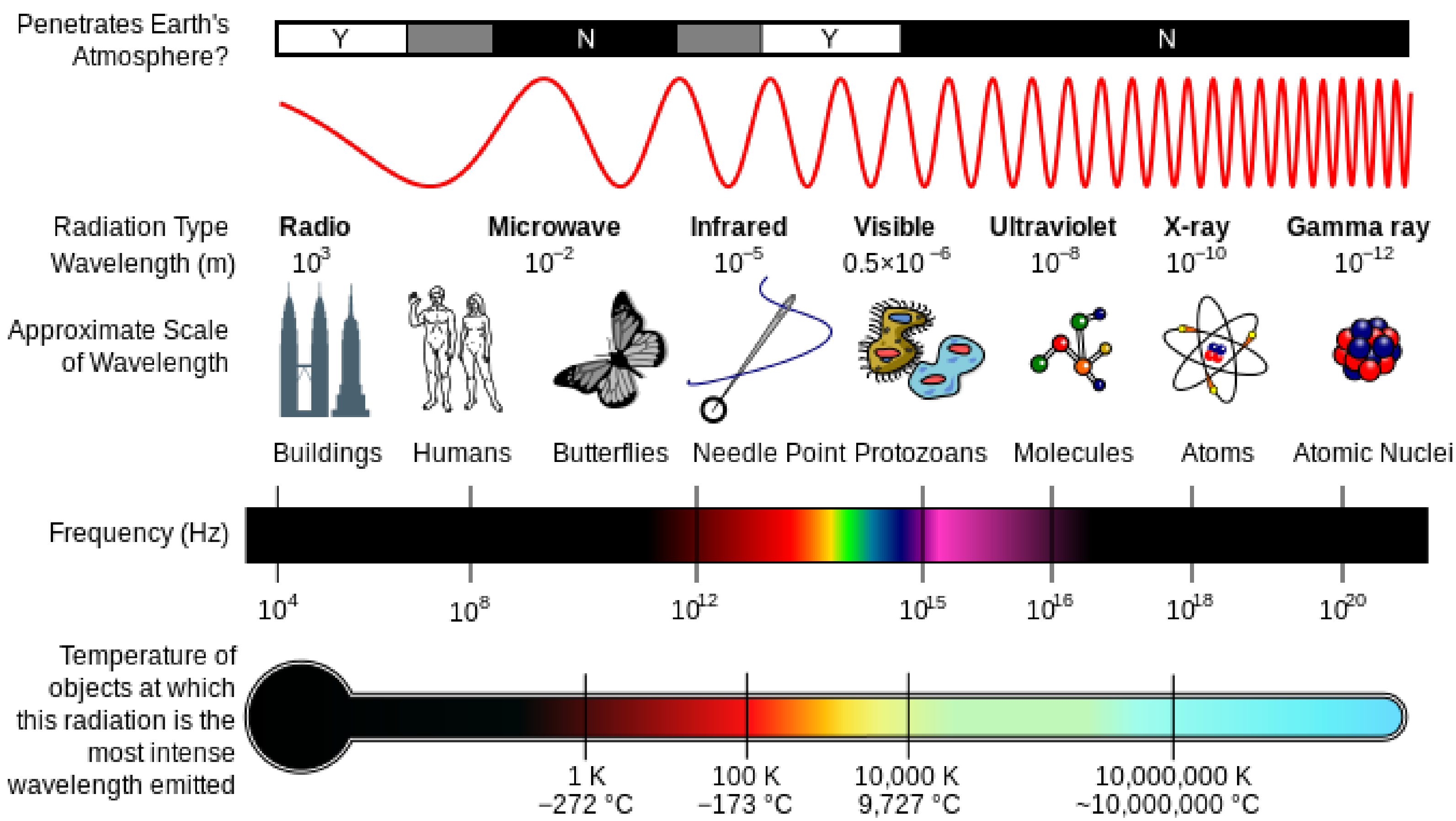
Photometric Image Formation - Lighting

- Electromagnetic (EM) radiation refers to the waves of the electromagnetic field, propagating through space, carrying electromagnetic radiant energy.



Photometric Image Formation - Lighting

Light is electromagnetic radiation within a certain portion of the electromagnetic spectrum.

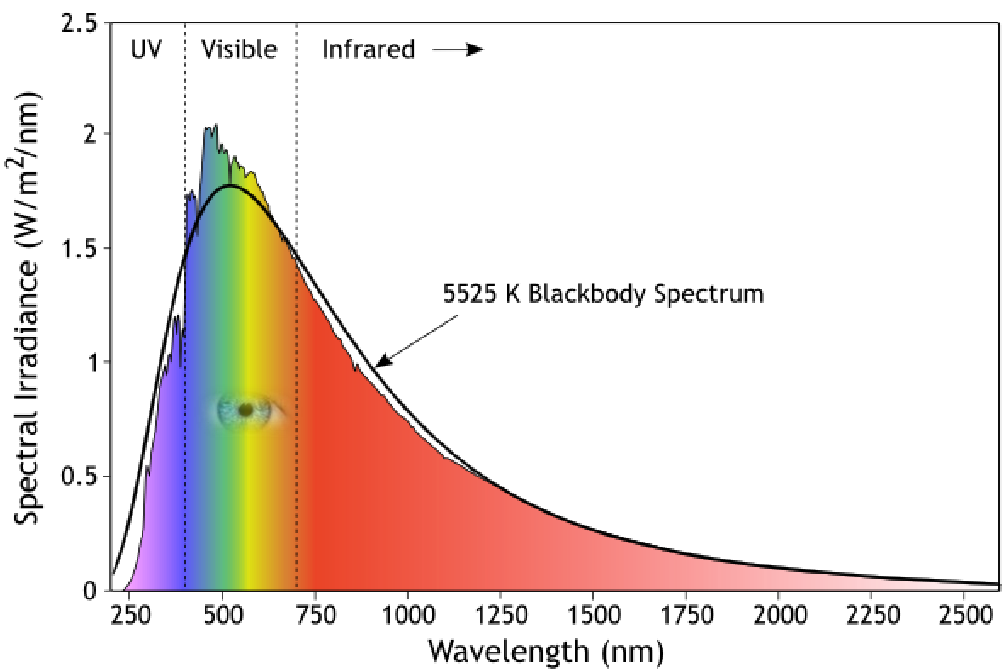


Lighting

- Point light source
  - Originates from a single location in space
  - A small light bulb or the sun remotely
  - Location, intensity and spectrum

- Area light source
  - More complicated
  - Ceiling lights in the classroom

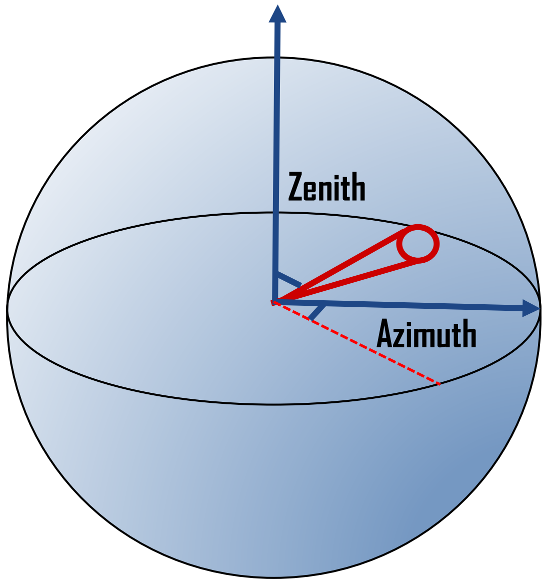
Spectrum of solar radiation



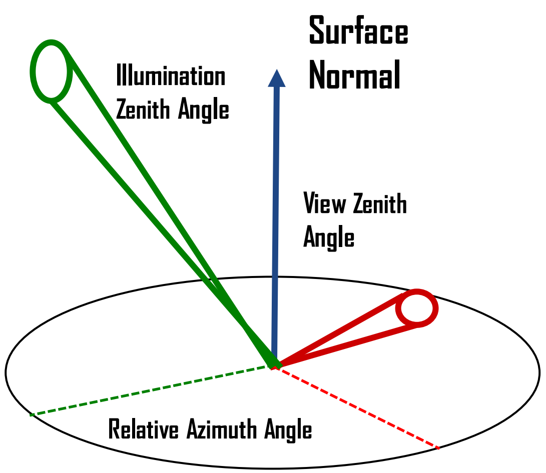
Source: Samuel J. Fogarty

Reflectance

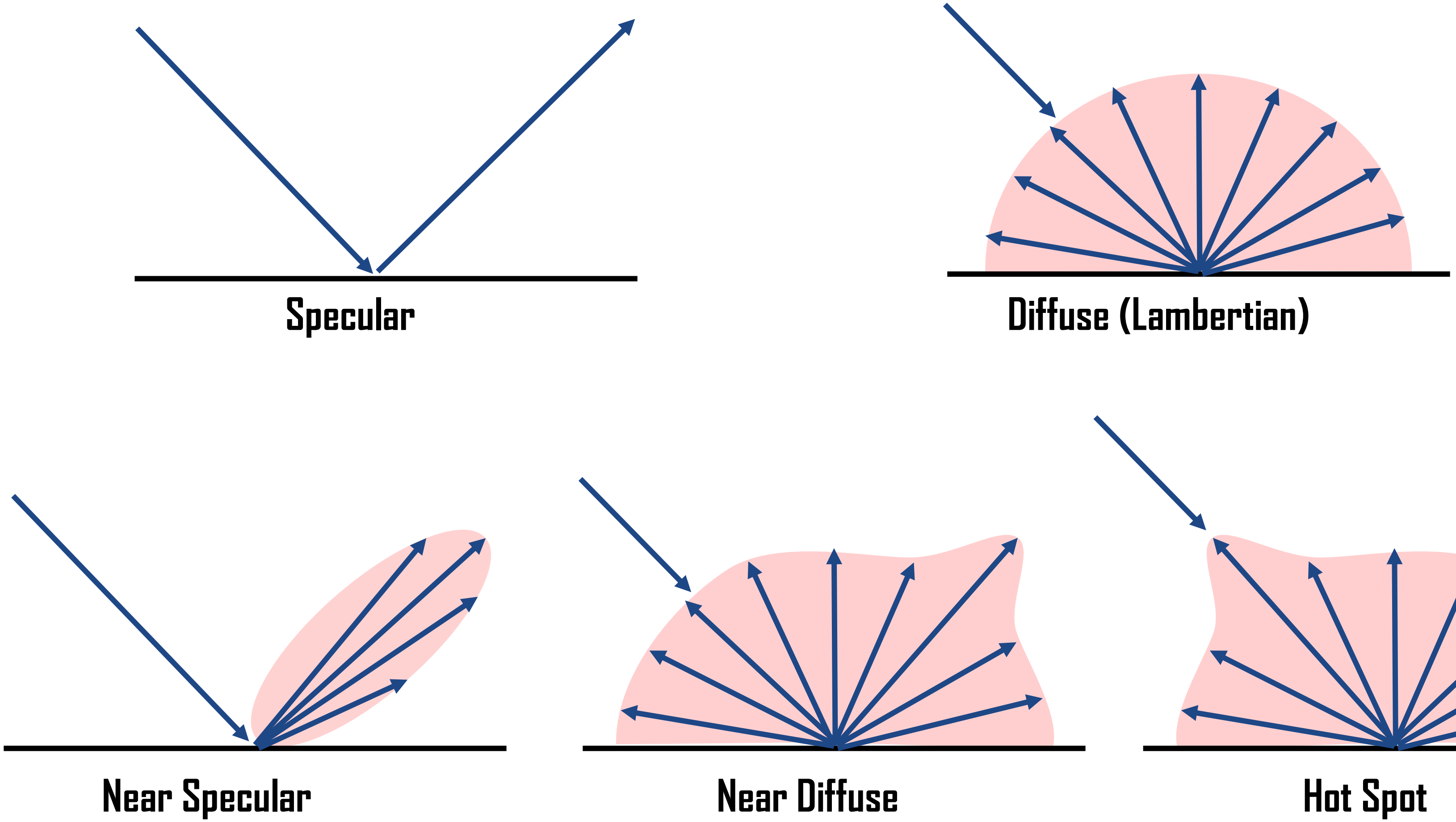
- Bidirectional reflectance distribution function (BRDF)
- A four dimensional function that describes how much of each wavelength of light arriving at an incident direction is emitted in a reflected direction
- Each direction is defined with respect to the surface normal and parameterised by azimuth and zenith angles



$$\text{BRDF: } f_r(\underbrace{\theta_i, \varphi_i}_{\text{incident direction}}, \underbrace{\theta_r, \varphi_r}_{\text{reflected direction}}, \underbrace{\lambda}_{\text{wavelength}}) = \frac{dL_r}{dE_i} \quad \begin{matrix}\text{output power} \\ \text{input power}\end{matrix}$$



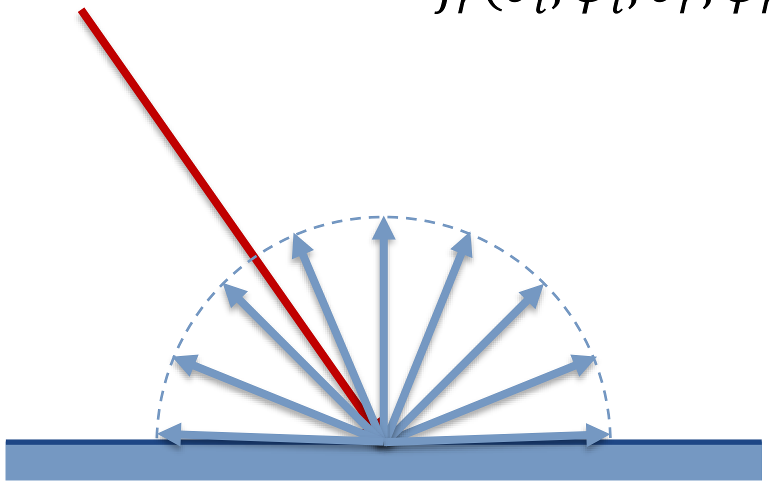
Different reflectance models

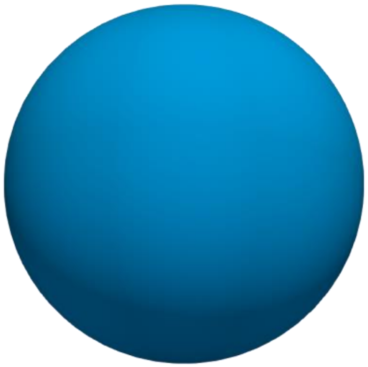


Diffuse reflection model

- Light is scattered uniformly in all directions.
- It is the phenomenon we most normally associate with shading.
- The amount of light depends on the angle between the incident light direction and the surface normal.
- The BRDF is constant:

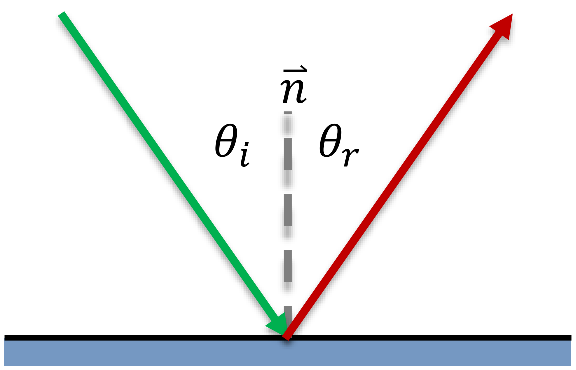
$$f_r(\theta_i, \varphi_i, \theta_r, \varphi_r, \lambda) = f_r(\lambda)$$

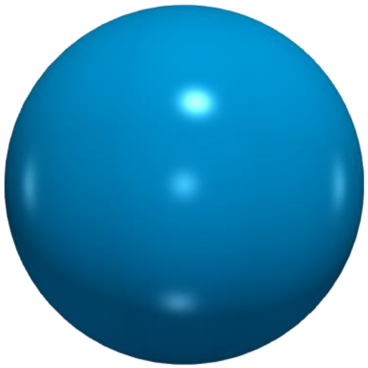




Specular reflection model

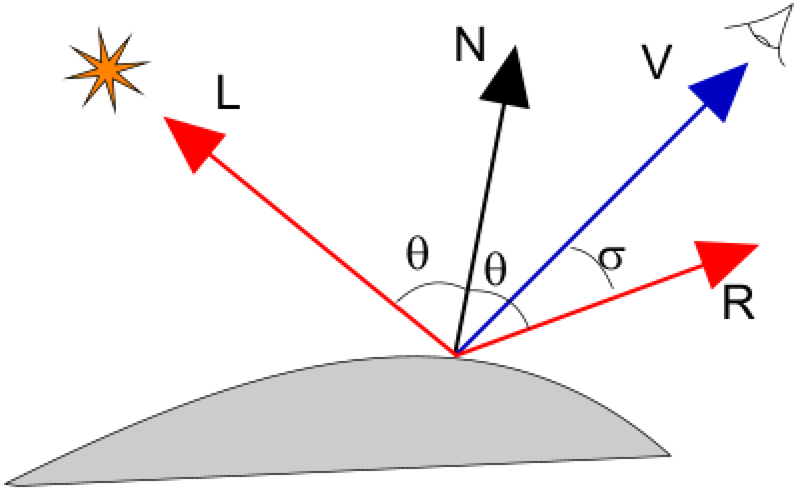
- Reflection in a mirror-like fashion.
- The reflection strongly depends on the direction of the incident and out-going light.
- The reflection and incident directions are symmetric with respect to the surface normal $\vec{n}$: $\theta_r = \theta_i$

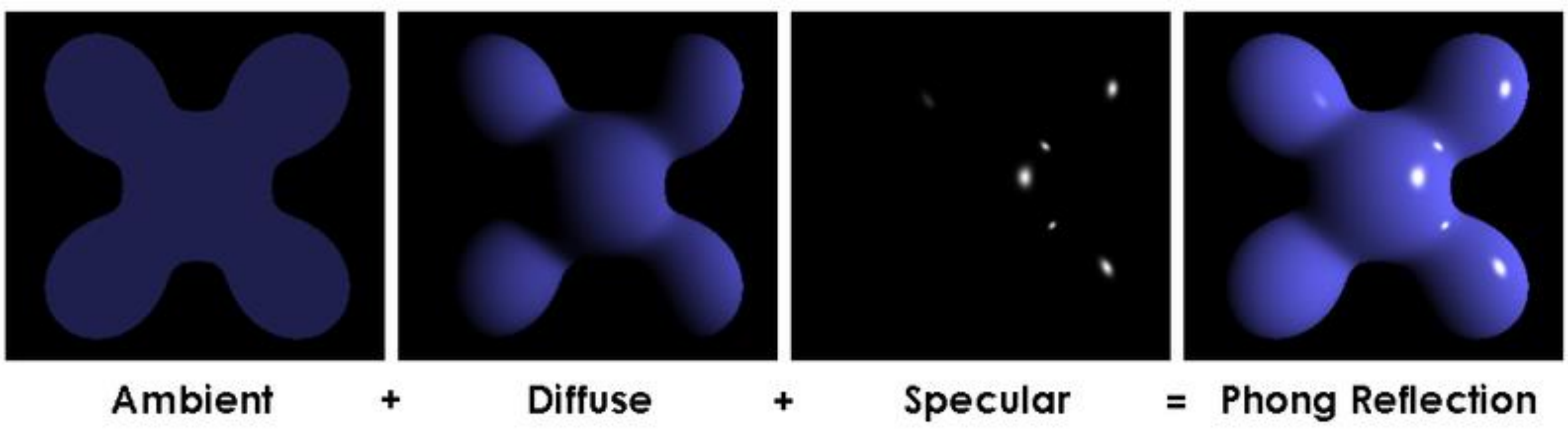




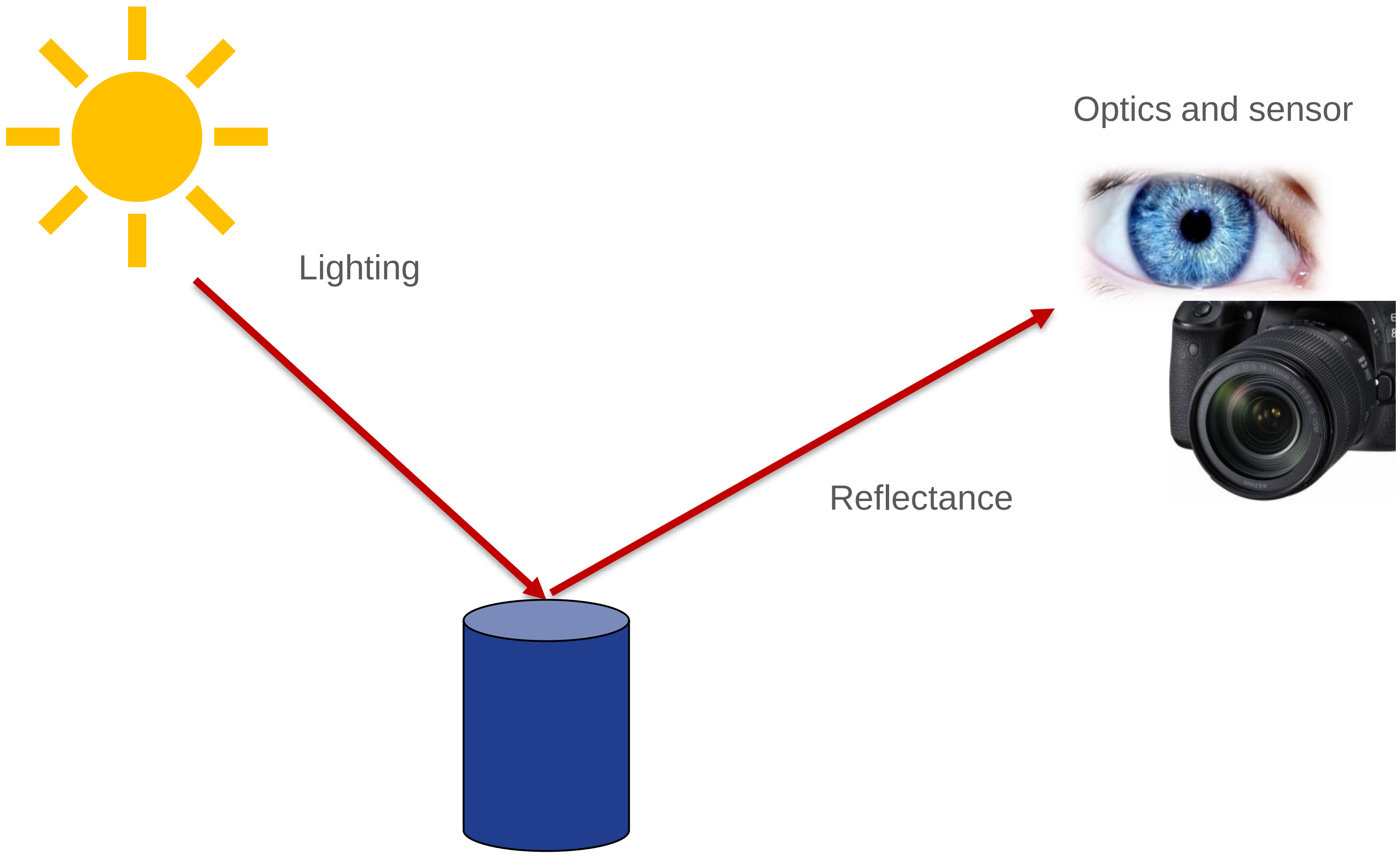
Phong shading

- For most of the cases, what we see is a combination of diffuse reflection, specular reflection and ambient illumination.
- The ambient illumination accounts for general illumination which may be complicated to model, such as
  - inter-reflection between walls in a room
  - distant sources, e.g. blue sky, sunny outdoor





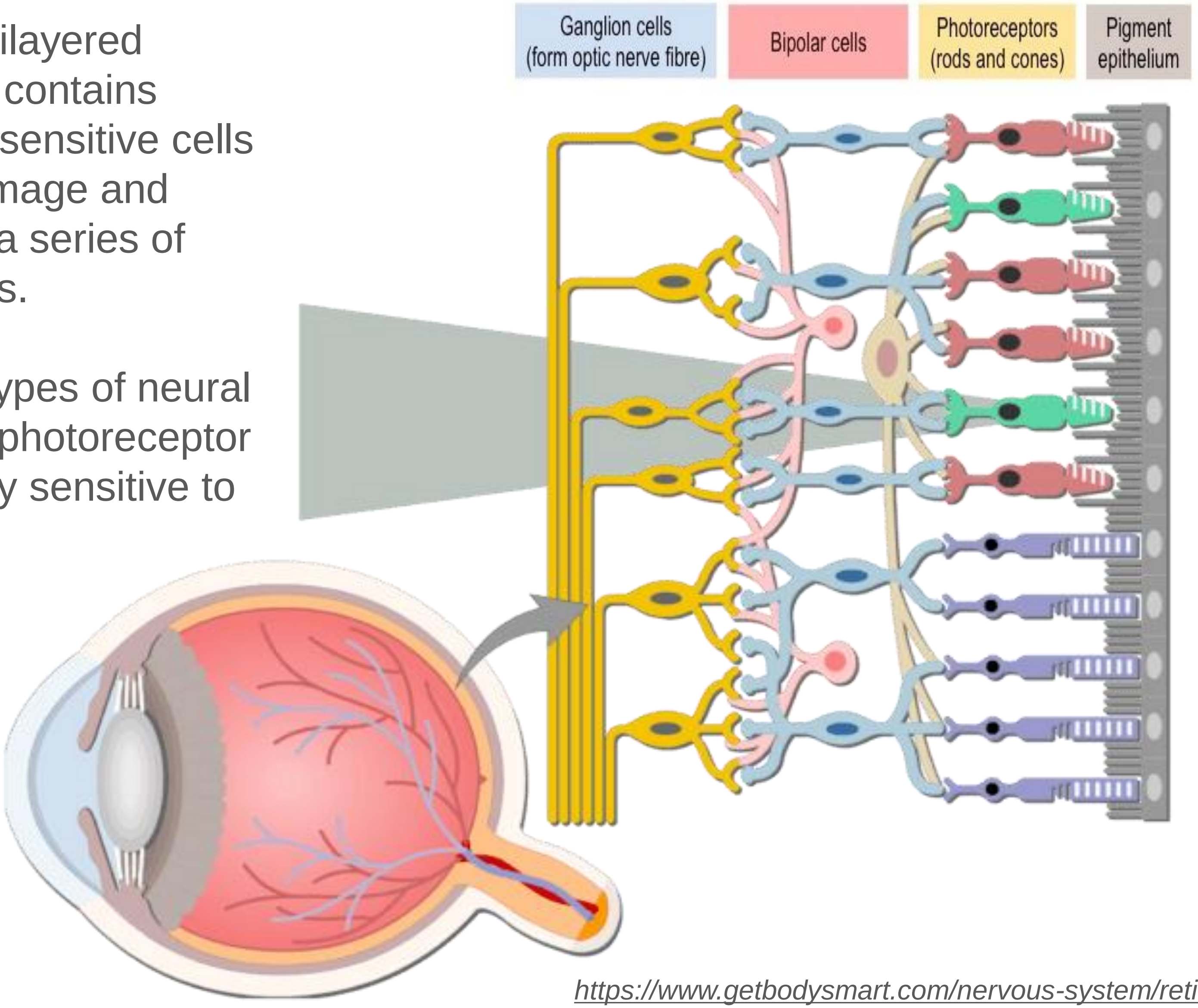
Photometric image formation



Human vision

- Retina is a multilayered membrane that contains millions of light-sensitive cells that detect an image and translate it into a series of electrical signals.

- There are two types of neural cells which are photoreceptor cells, i.e. directly sensitive to light.



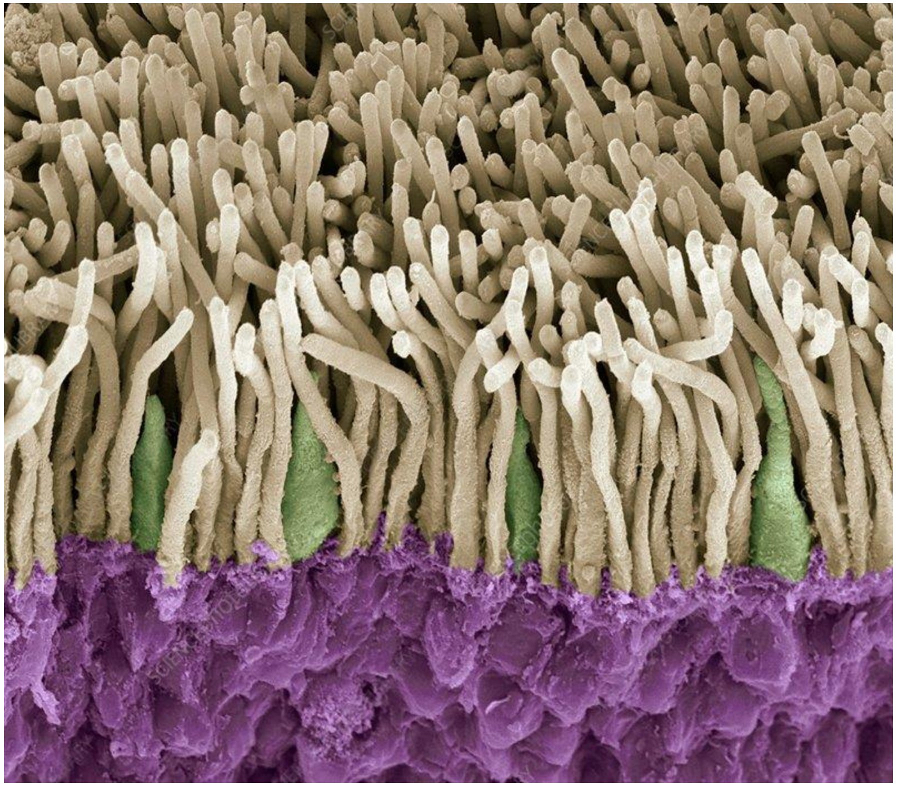
https://www.getbodysmart.com/nervous-system/retina

Human vision

- Retina is a multilayered membrane that contains millions of light-sensitive cells that detect an image and translate it into a series of electrical signals.

- Rod cells
  - Provide black-and-white vision
  - Detect motion
  - Function in dim light
  - They are the primary source of visual information at night

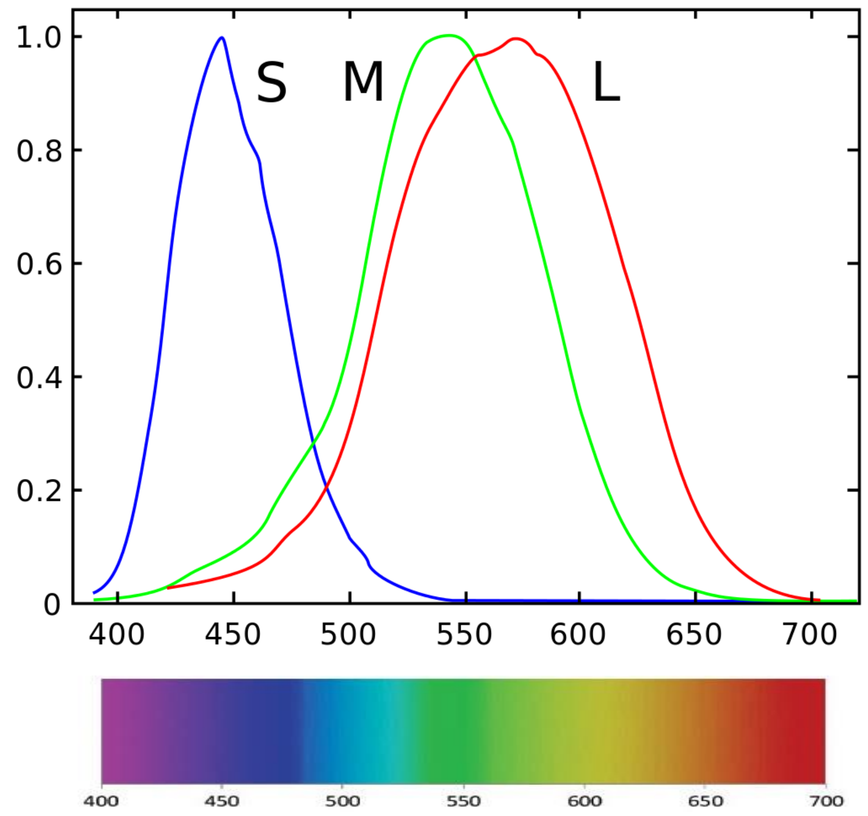
- Cone cells
  - Responsible for colour vision
  - Function in bright light
  - Three types of cones sensitive to three different spectra, resulting in trichromatic colour vision.



https://www.sciencephoto.com/media/186019/view/retina-rods-and-cones-sem

Human vision – Cone cells

- The cones are labeled according to the ordering of the wavelengths of the peaks of their spectral sensitivities: short (S), medium (M), and long (L) cone types.
- A range of wavelengths of light stimulates each of these receptor types to varying degrees.
- The brain combines the information from each type of receptor to give rise to different perceptions of different wavelengths of light.
- The RGB colour model therefore, is a convenient means for representing colour, but is not directly based on the types of cones in the human eye.



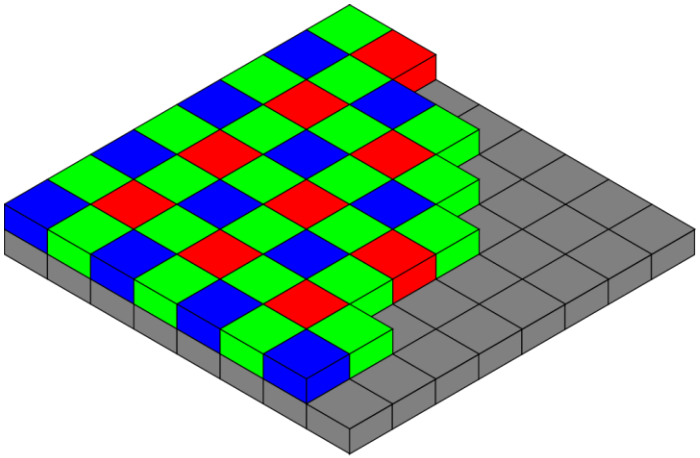
Camera sensors

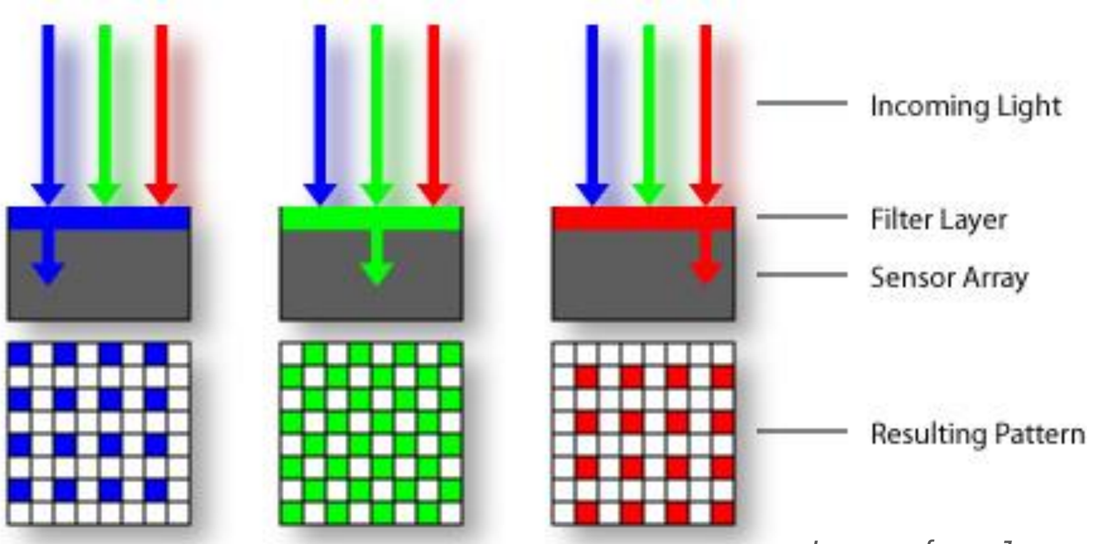
- Two common types of sensors
  - CCD (charged-coupled device)
    - Used in professional, medical, and scientific applications where high-quality image data is required.
  - CMOS (complementary metal-oxide semiconductor)
    - Used by most smart phone cameras.

- When light from a scene falls on a camera sensor, it is collected by a matrix of small wells called pixels.
- The pixels may be photodiodes or photocapacitors, which generate a charge proportional to the amount of incident light, spatially restricting and storing it.
- The information from these photosites is collected, organized, and transferred to a monitor to be displayed.
- The sensor resolution is defined as the number of pixels per millimeter.
- The size of the pixels affects the resolution, contrast and signal-to-noise ratio (SNR).

Camera sensors

- Most cameras use a colour filter array (CFA), where alternating sensors are covered by different coloured filters.
- The most commonly used pattern is the Bayer pattern which places green filters over half of the sensors and red and blue filters over the remaining ones.
- 50% green, 25% red, 25% blue.

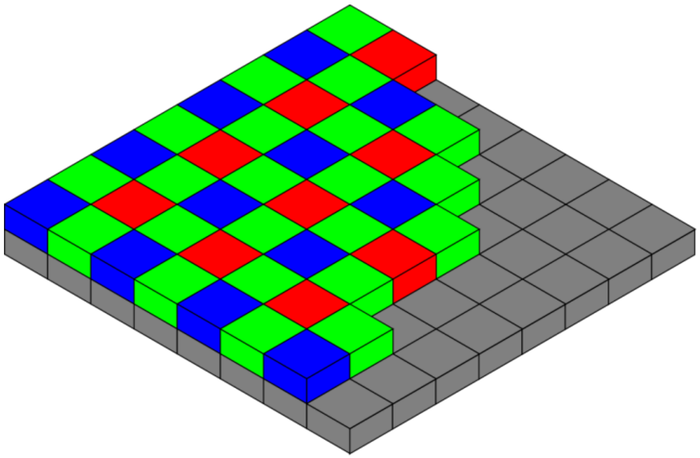


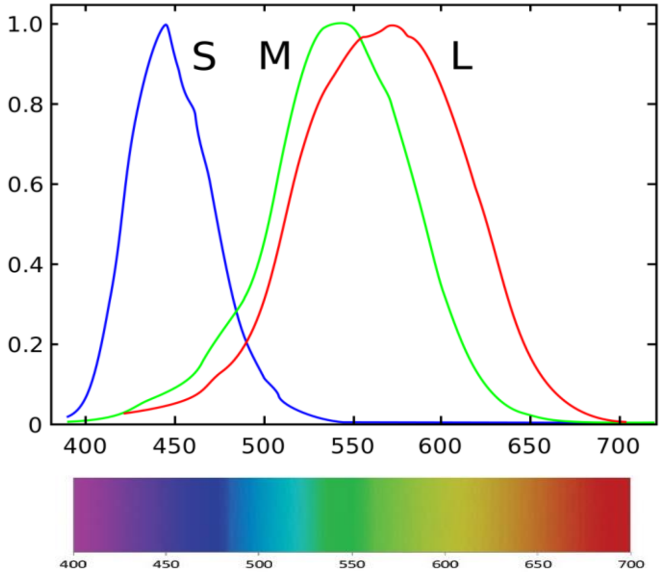


Images from James Tompkin's slides

Camera sensors

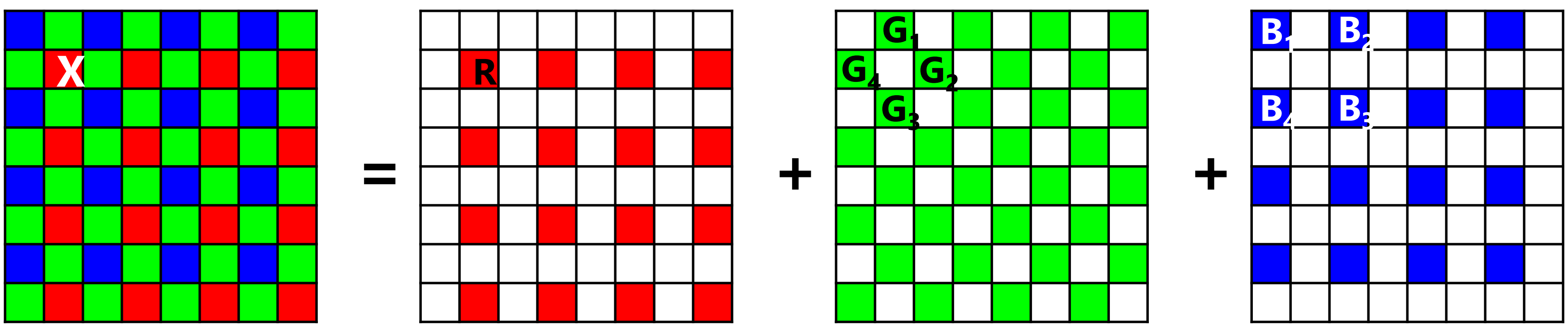
- Most cameras use a colour filter array (CFA), where alternating sensors are covered by different coloured filters.
- The most commonly used pattern is the Bayer pattern which places green filters over half of the sensors and red and blue filters over the remaining ones.
- 50% green, 25% red, 25% blue.
- The reason that there are twice as many green filters as red and blue is because the luminance signal is mostly determined by green values and the visual system is much more sensitive to high frequency detail in luminance than in chrominance.





Bayer filter

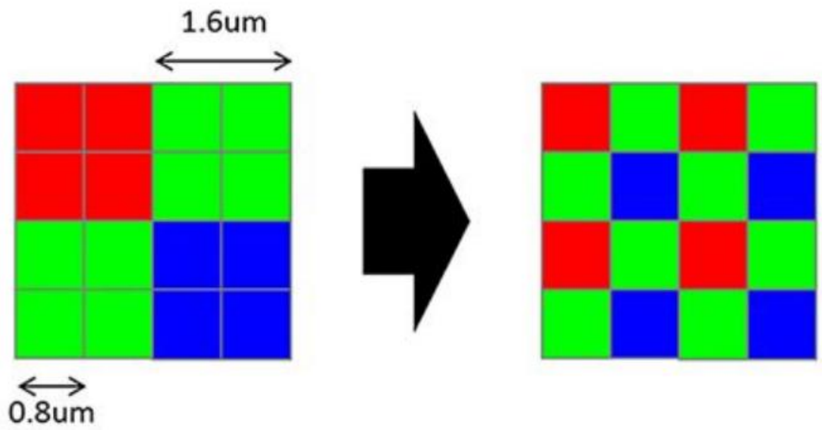
- Only one colour is available at each pixel.
- The other two colours can be interpolated from neighbouring pixels.
- By interpolation, at each pixel, we can get (R,G,B) values.
- The process of interpolating the missing colour values so that we have valid RGB values for all the pixels is known as *demosaicing*.

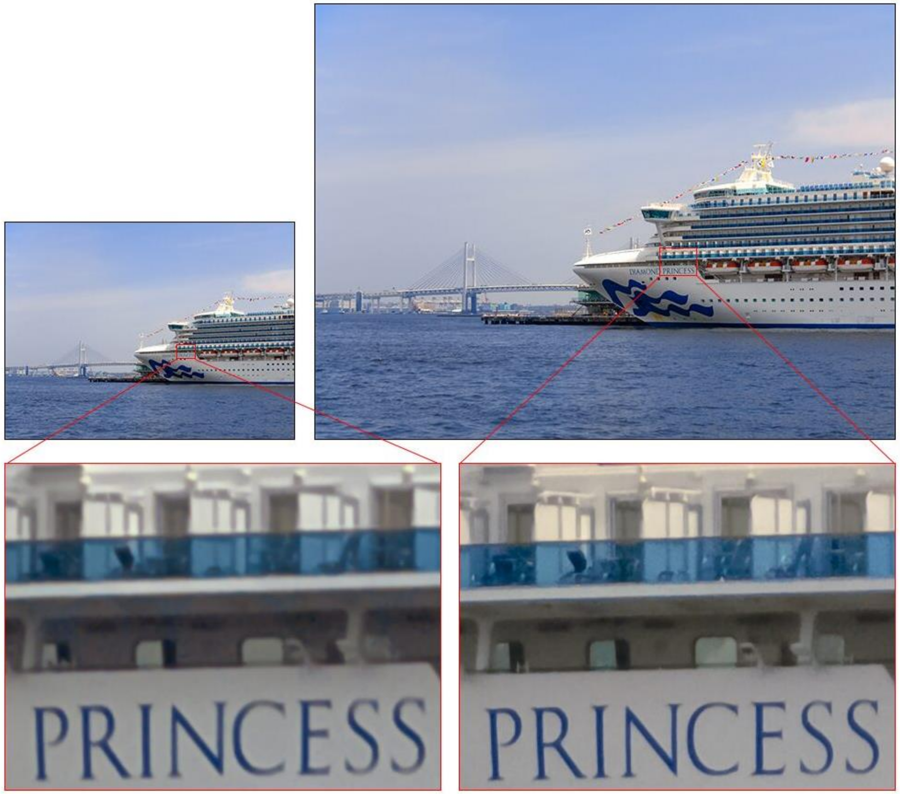


T. Sakamoto et al. Software pixel interpolation for digital still cameras suitable for a 32-bit MCU. IEEE TCE, 1998.

Camera sensors

- Sony IMX586 CMOS image sensor for smartphones (as Jul 2018)
- 8,000 (H) x 6,000 (V) pixels, i.e. 48MP
- Each pixel size: 0.8μm x 0.8μm
- Quad Bayer colour filter array, where the adjacent 2x2 pixels come in the same colour
- The new sensor delivers both high sensitivity and high resolution.



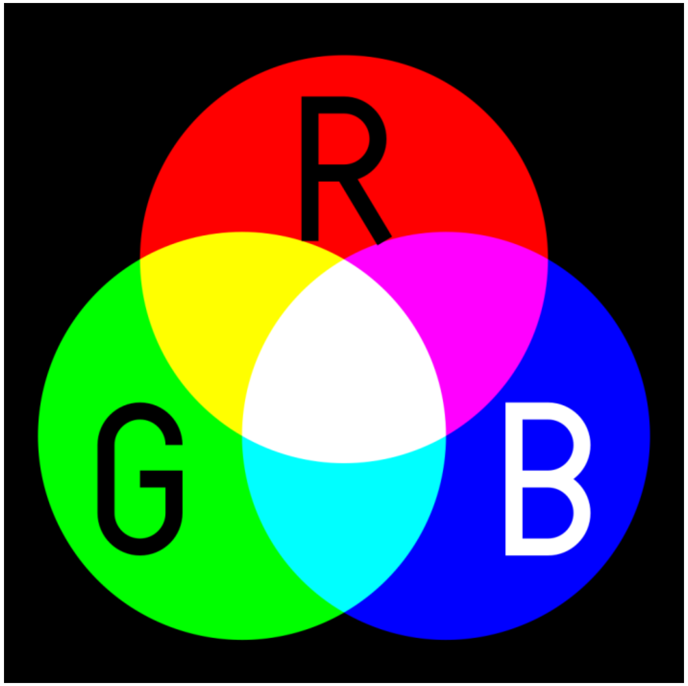


Conventional image (12 effective megapixels)

IMX586 image (48 effective megapixels*2)

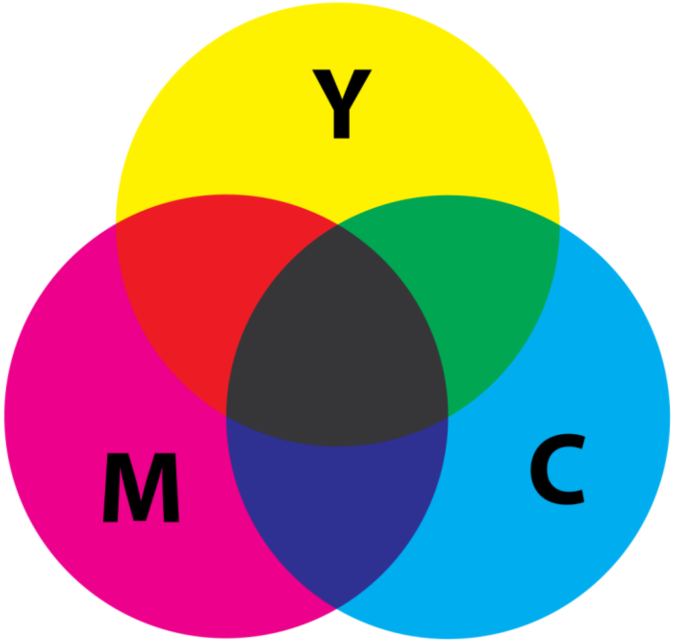
Source: https://www.sony.net/SonyInfo/News/Press/201807/18-060E/index.html

Colour models



Additive colour model

- Start from black (e.g. computer monitor)
- Add a spectrum of light in mixing



Subtractive colour model

- Start from white (e.g. white paper)
- Subtract a spectrum of light
- The subtractive colours are called subtractive because pigments in the paint absorb certain wavelengths in the colour spectrum.

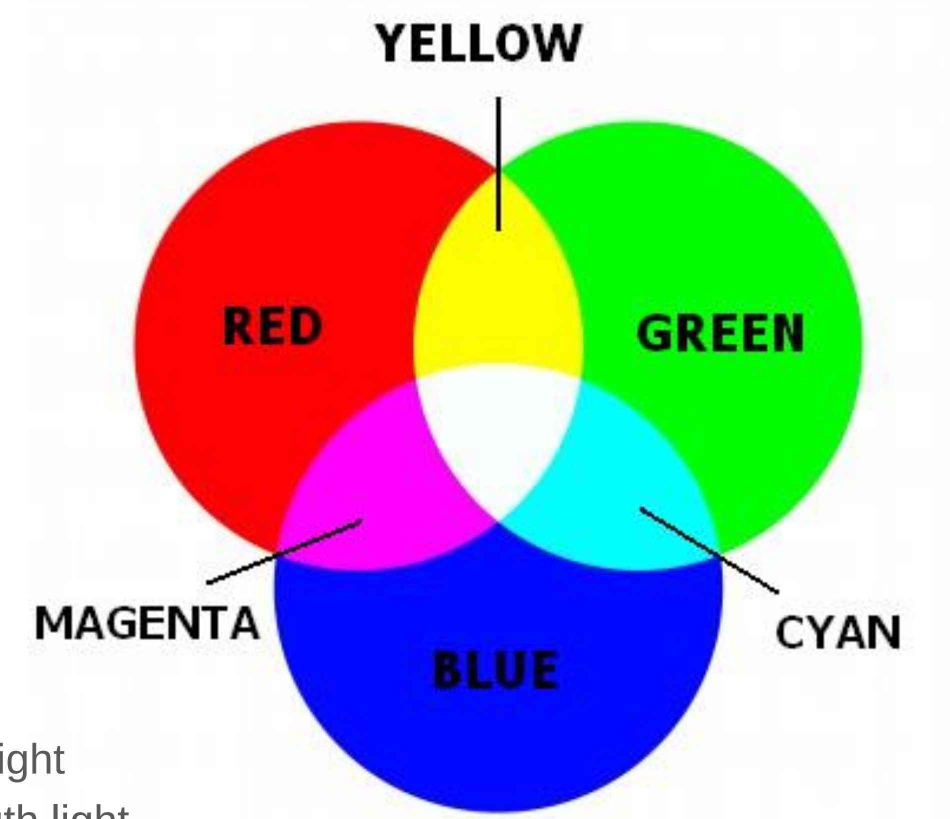
Colour spaces

- Primary colours
  - Most visible colours can be represented by mixing three colours (R, G, B)

- The existence of three primaries is a result of the tri-stimulus nature of the human visual system

- Our eyes have 3 types of cone cells
  - **S** cell responding to short wavelength light
  - **M** cell responding to medium wavelength light
  - **L** cell responding to long wavelength light

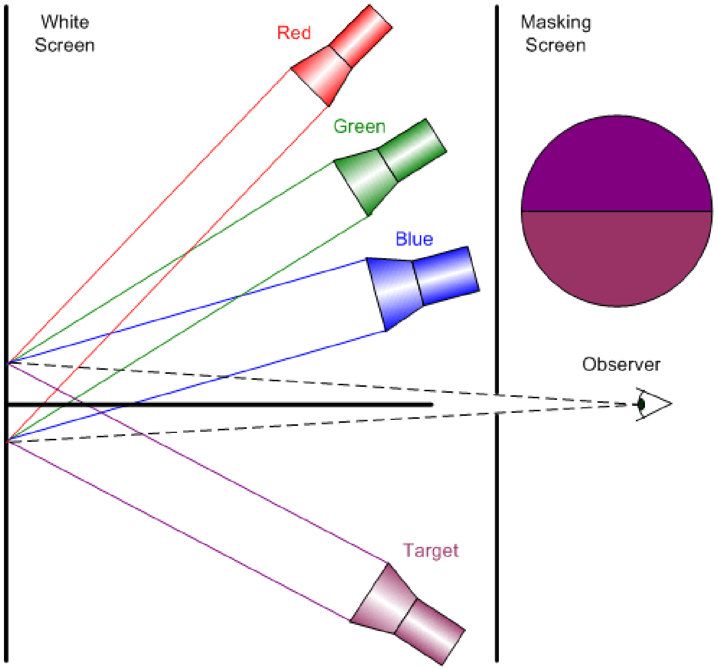
- The camera sensors consist of R, G, B sensors



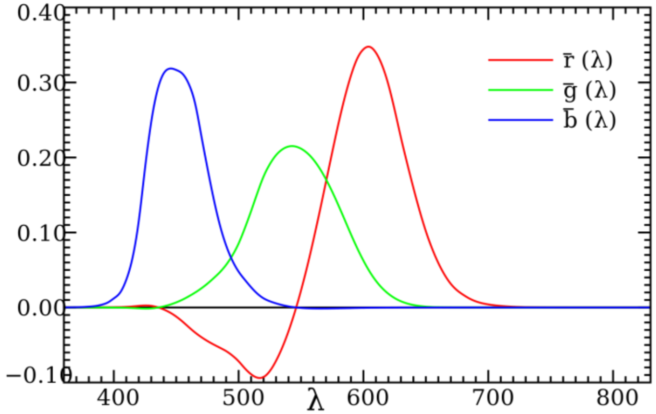
CIE RGB colour space

- CIE (Commission internationale de l'éclairage), or International Commission on Illumination standardized the RGB representation by performing colour matching experiments using the primary colours of red (700.0nm wavelength), green (546.1nm), and blue (435.8nm).

- Colour matching experiment
  - The observer needs to adjust the values of the three primary lights (red: 700nm; green: 546.1 nm; blue: 435.8 nm) until they produce a colour indistinguishable from the target light.
  - For certain pure spectra, a negative amount of red light has to be added.
  - This means that it is impossible to display certain perceptual colors using CIE RGB emissive displays



Any color = r*R + g*G + b*B



https://en.wikipedia.org/wiki/CIE_1931_color_space

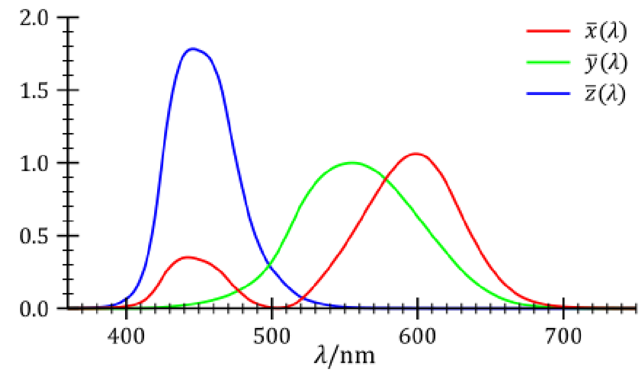
CIE 1931 XYZ colour space

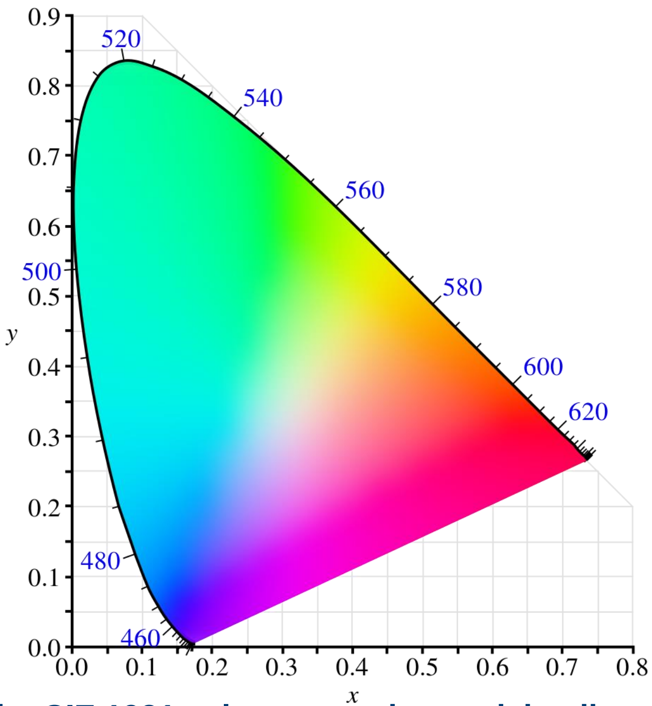
- To avoid these negative RGB values, and to have one component that describes the perceived brightness, "imaginary" primary colours (X,Y,Z) and corresponding colour-matching functions were formulated.
- The Y parameter is a measure of the luminance (perceived relative brightness) of a colour.
- The x and y parameters are the chromaticity coordinates being functions of all three tristimulus values (XYZ).

$$x = \frac{X}{X+Y+Z}$$

$$y = \frac{Y}{X+Y+Z}$$

$$z = \frac{Z}{X+Y+Z} = 1 - x - y$$



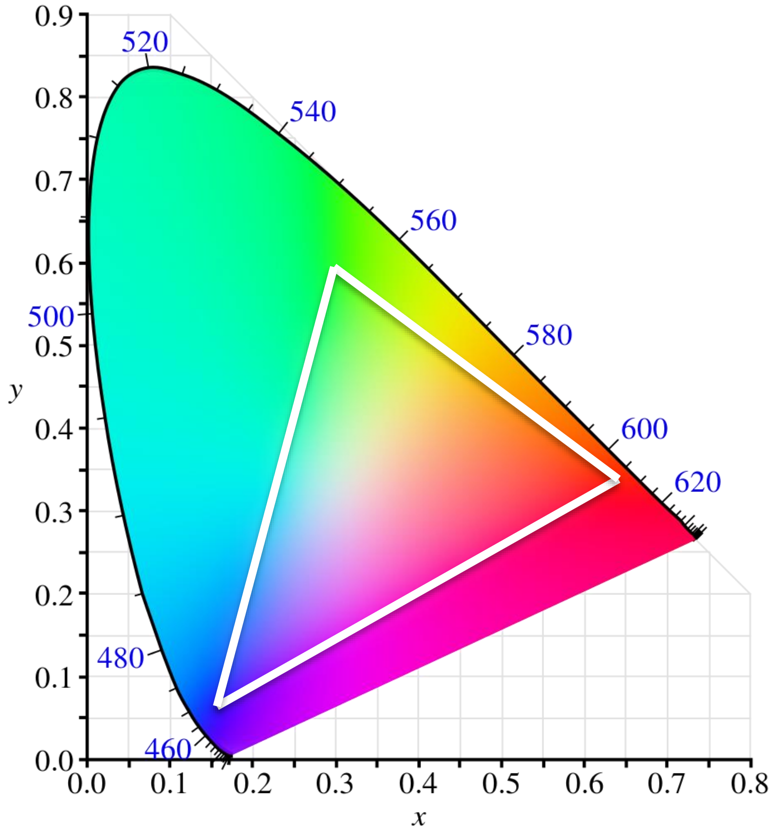


The CIE 1931 color space chromaticity diagram.

sRGB colour space

- sRGB: Standard RGB space
- Created by HP and Microsoft in 1996 for use on monitors, printers and internet
- It is only a subset (gamut) of all the colours that we can see.
- The projectors, laptops that you use may not produce all the colours.

sRGB definition	x	y
Red	0.6400	0.3300
Green	0.3000	0.6000
Blue	0.1500	0.0600



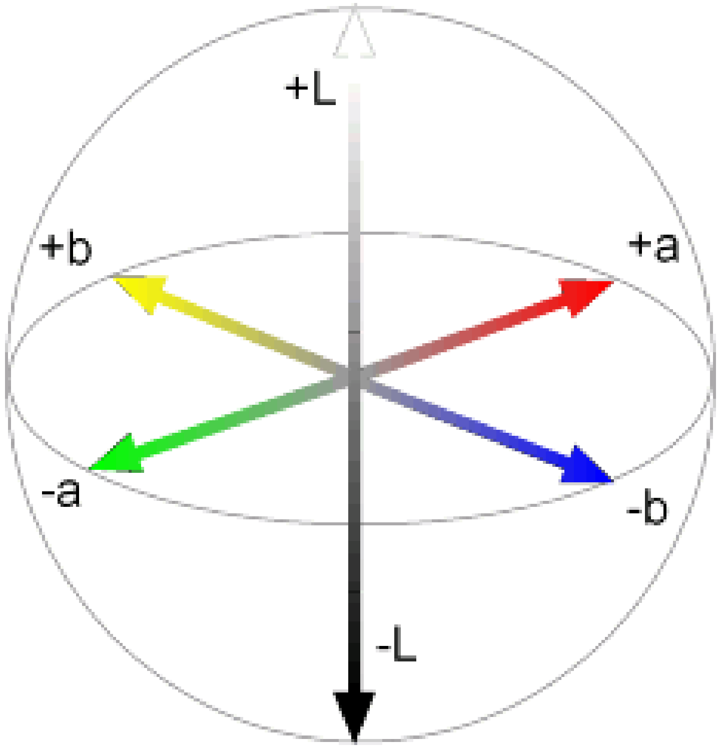
sRGB colour triangle

L*a*b* color spaces

- The response of the human visual system is roughly logarithmic and we can perceive relative luminance differences of about 1%.

- While the XYZ colour space has the ability to separate luminance from chrominance, it does not actually predict how well humans perceive differences in colour or luminance.

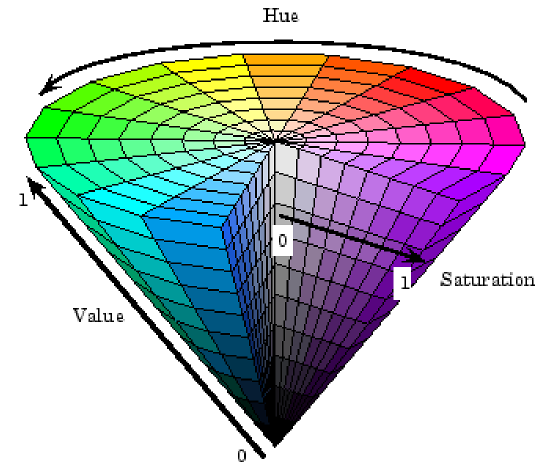
- The CIE L*a*b* colour space is a non-linear re-mapping of the XYZ space where differences in luminance or chrominance are more perceptually uniform.



- The L*a*b* colour space expresses colour as three values:
  - L* for the lightness from black (0) to white (100)
  - a* from green (−) to red (+)
  - b* from blue (−) to yellow (+)

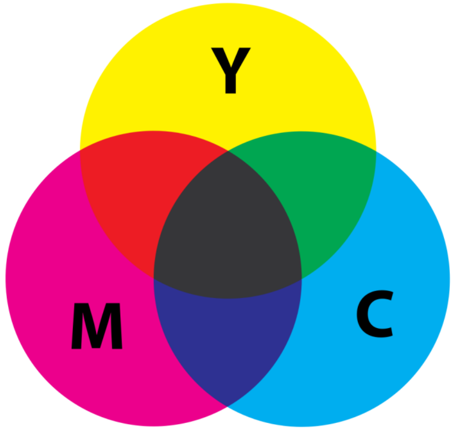
Other colour spaces

- HSV which is a projection of the RGB colour cube onto a non-linear chroma angle, a radial saturation percentage, and a luminance-inspired value.
  - Hue: values in [0, 1], colours
  - Saturation: values in [0, 1], unsaturated to fully saturated (no white component)
  - Value: values in [0, 1], brightness



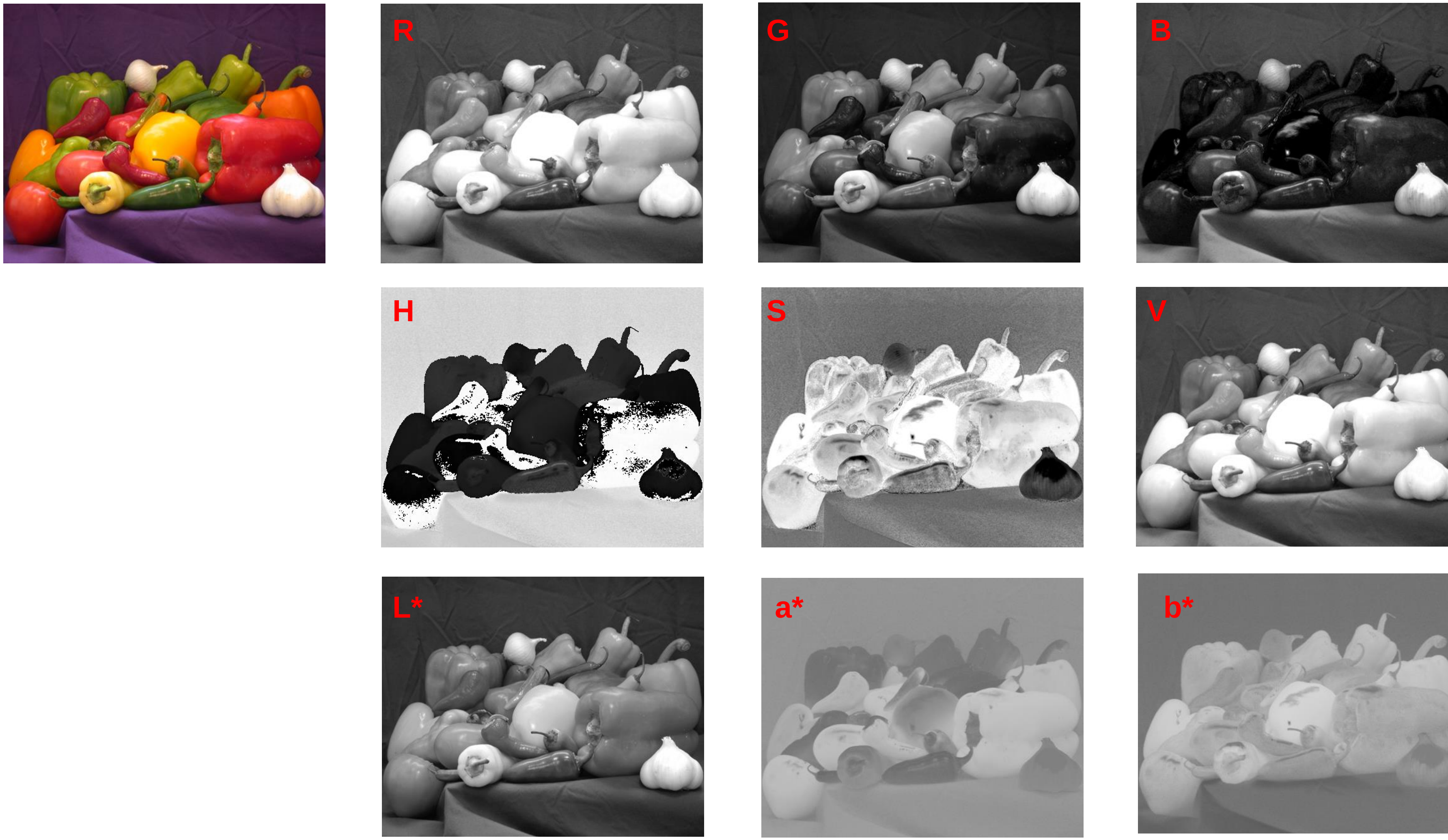
HSV colour space

- CMYK, used for printing
  - Cyan
  - Magenta
  - Yellow
  - Key (black)



CMY primary colours

Colour spaces

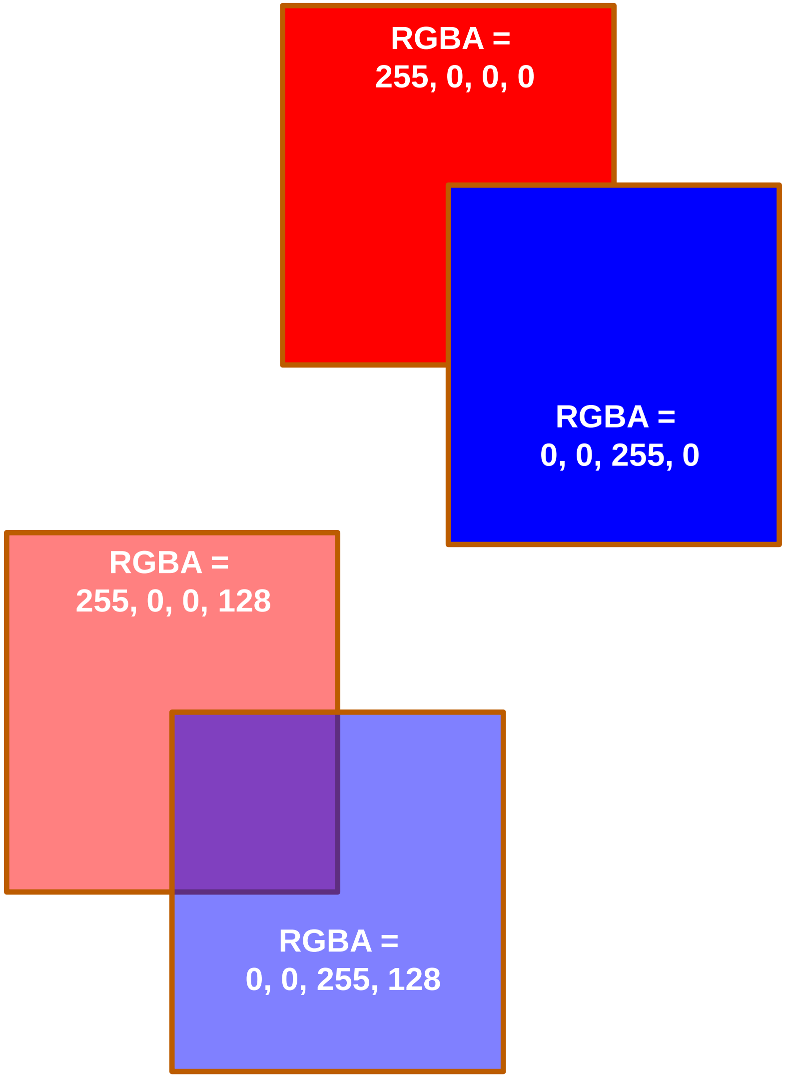


Colour representation in computer

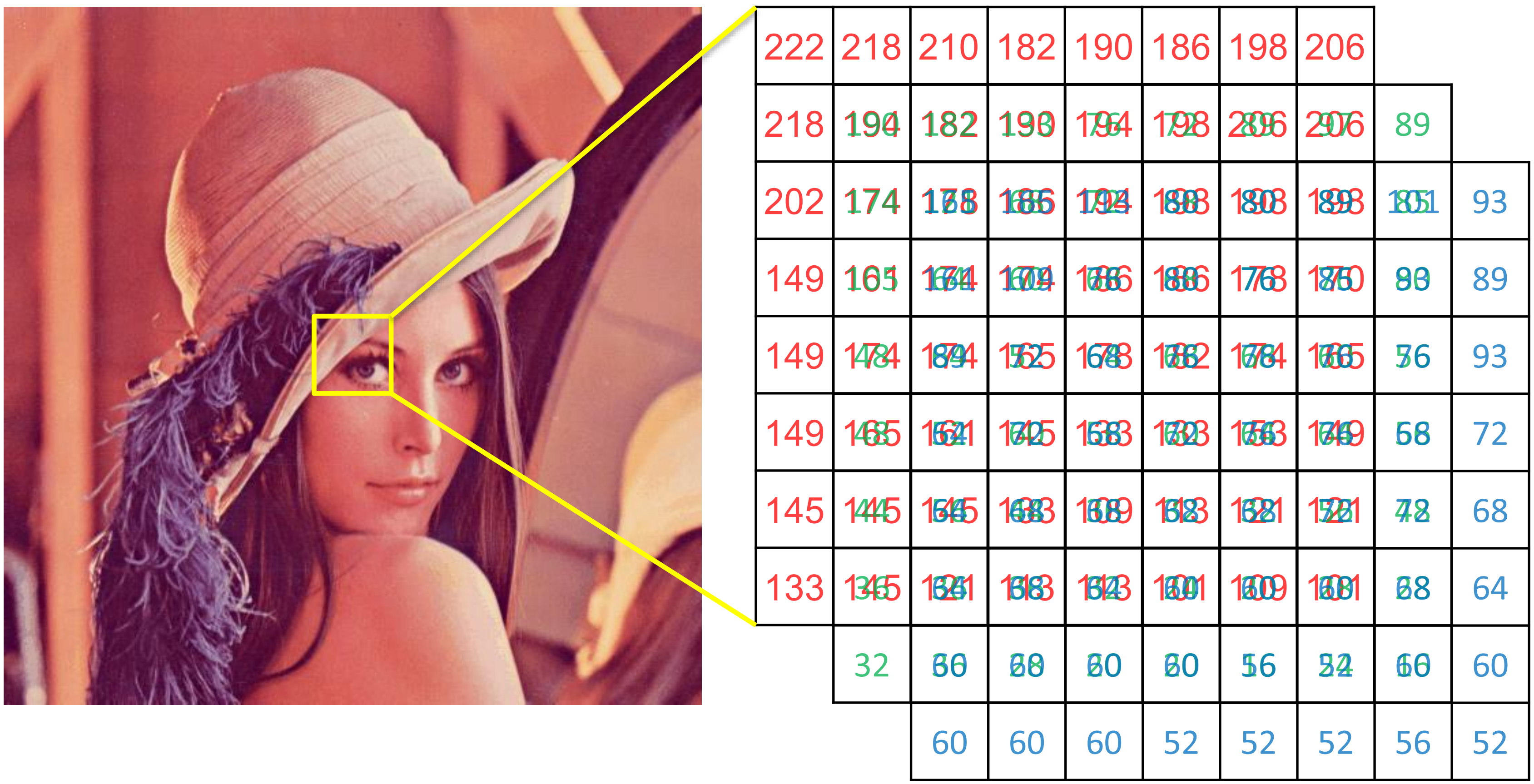
- RGB or RGBA
  - Red
  - Green
  - Blue
  - Alpha channel: transparency

- Grey
  - Greyscale intensity (R = G = B)

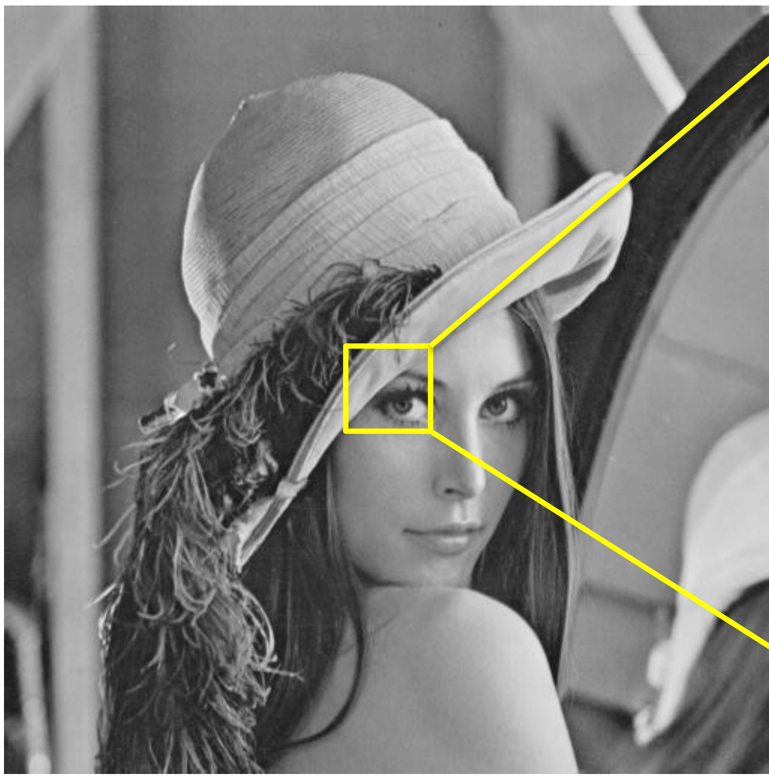
- Representation
  - 8 bits, i.e. 0 to 255 (most common)
  - 16 bits (camera raw files)



Colour representation in computer



Grayscale representation

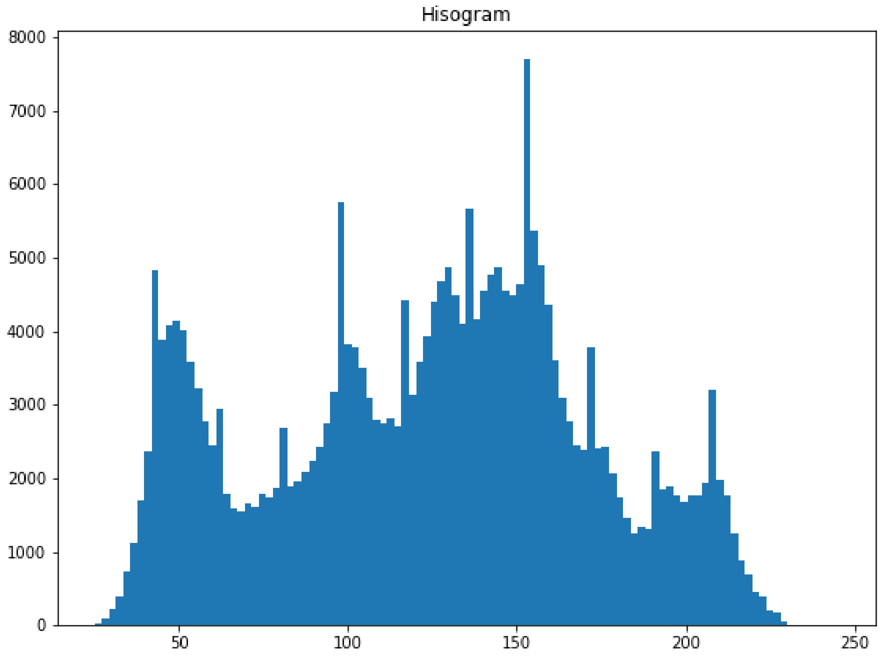


199	192	158	111	110	123	130	130
189	149	108	111	113	120	126	125
130	100	98	108	113	113	114	120
85	100	96	104	108	107	101	94
85	95	98	96	100	103	100	96
79	94	87	77	69	70	87	84
77	80	72	71	60	52	59	64
68	67	63	58	53	51	54	52

Histogram

- Histogram divides the intensity range into a number of bins and shows number of pixels or frequencies in each bin.

- It is an intuitive way to show the intensity distribution.



Conclusions

- Photometric Image Formation
- Light
- Reflectance
- Optics and sensors
- Colour spaces
- Image representation