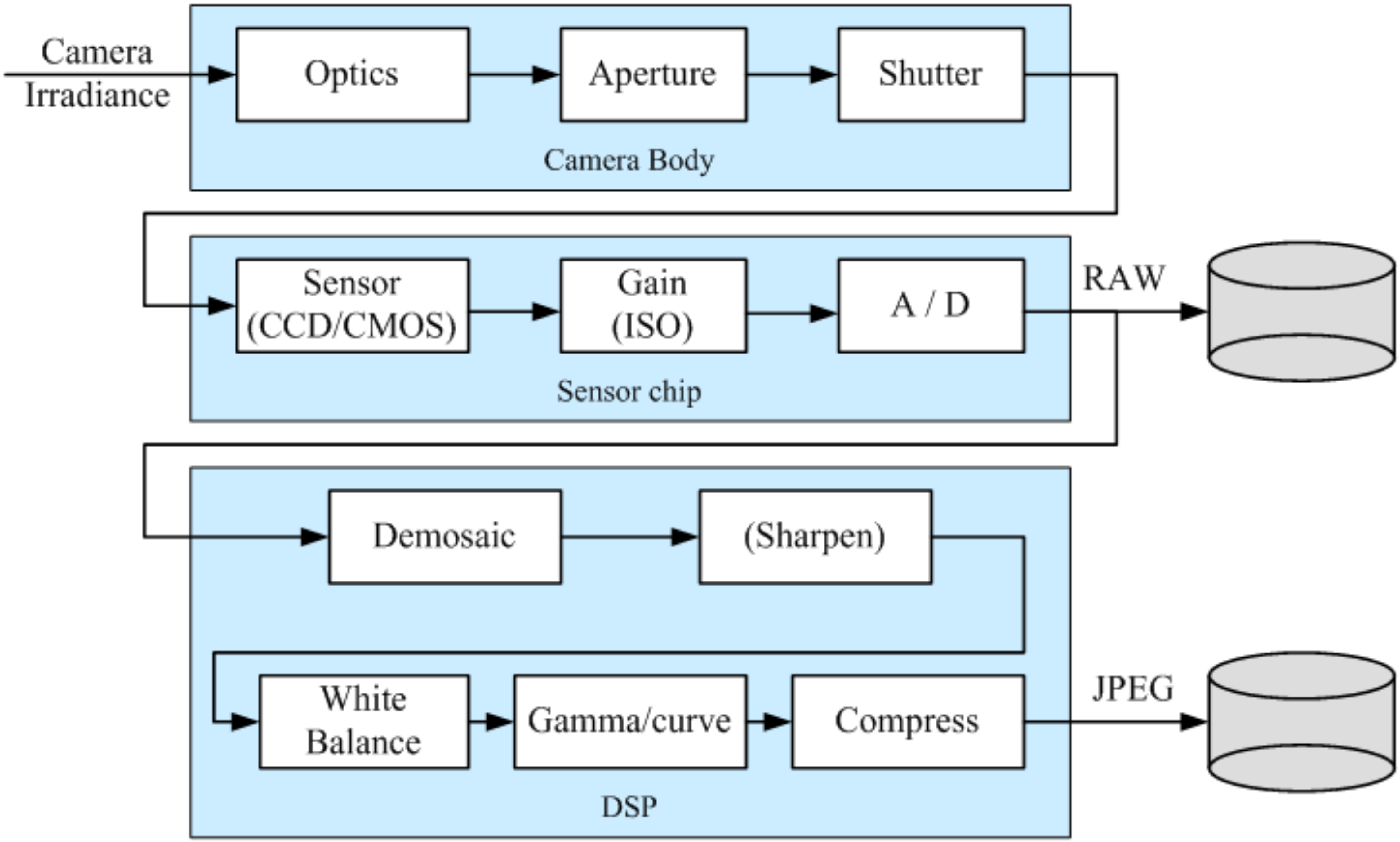
2.3 The Digital Camera

Image Sensing Pipeline (Simplified)



Outline

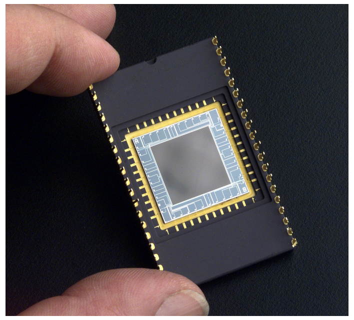
- The Sensor
- Sampling & Aliasing
- Colour Coding

Outline

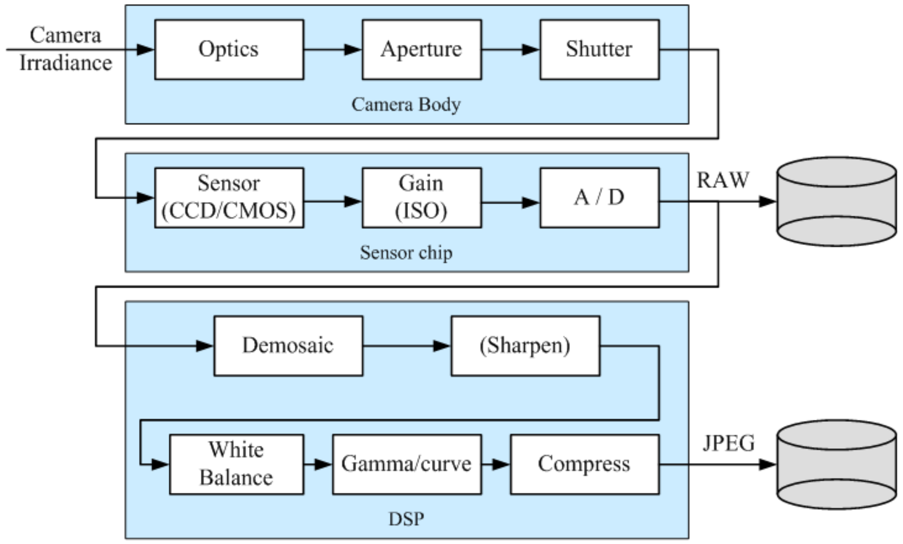
- **The Sensor**
- Sampling & Aliasing
- Colour Coding

Sensor

- CCD (Charge-Coupled Device)
  - Photons accumulated in each active well.
  - Then charge transferred from well to well ("bucket brigade") until deposited at sense amplifiers

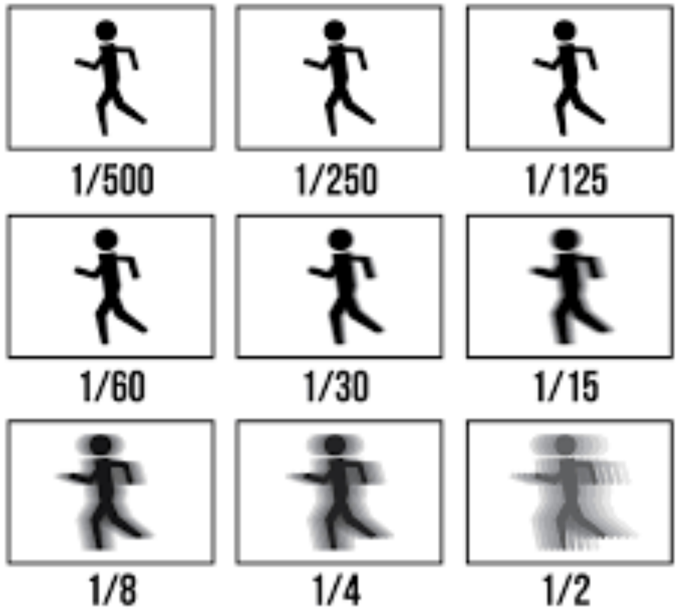


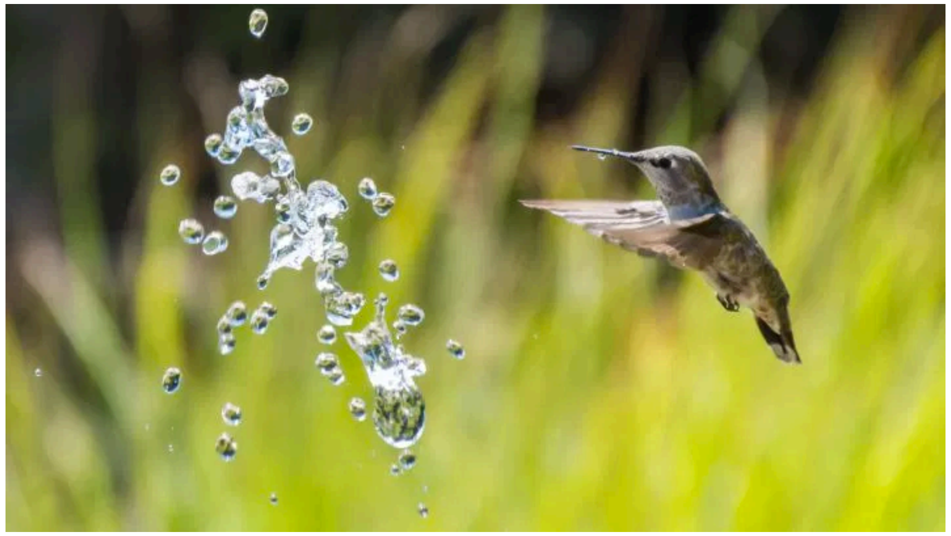
- CMOS (Complementary Metal Oxide Semiconductors)
  - photons directly affect conductivity of a photodetector
  - Each photodetector can be selectively gated and amplified
  - Read out using multiplexing scheme
- Most digital cameras now use CMOS.

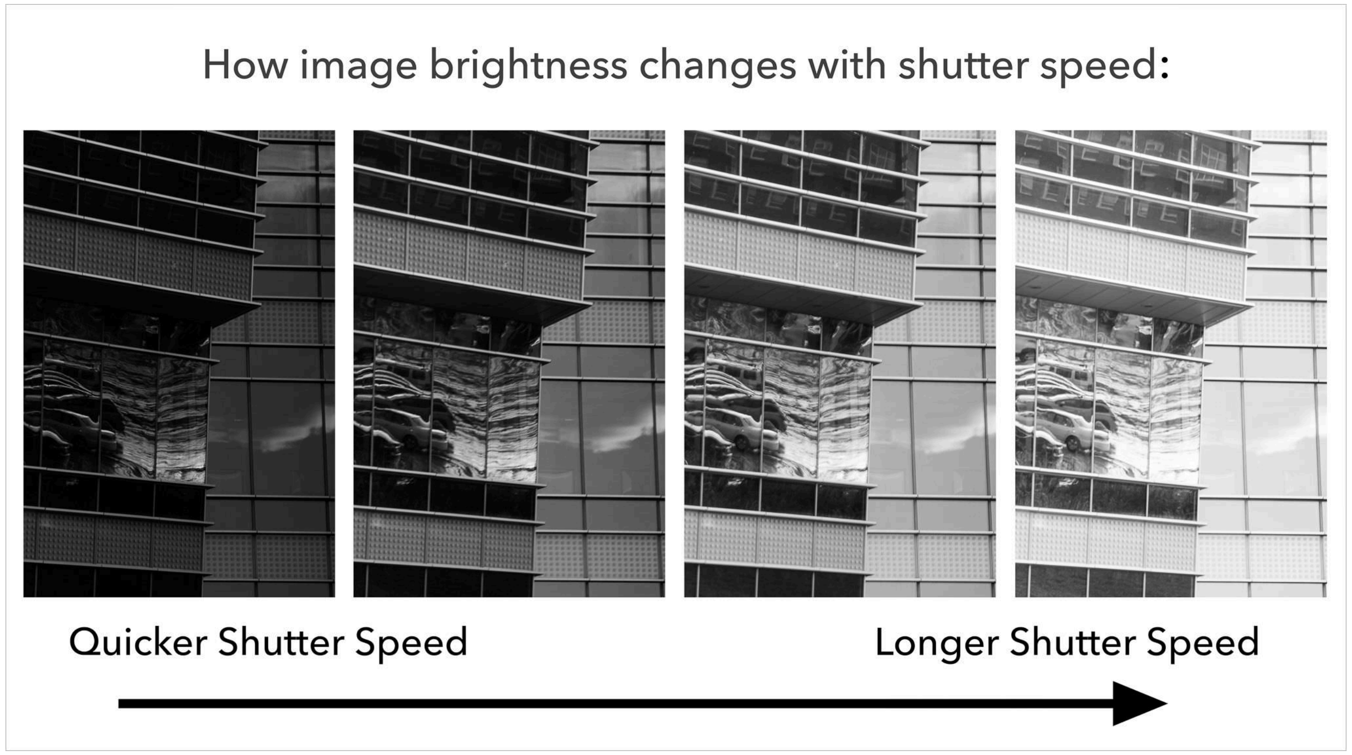


Shutter Speed

- Measured in fractions of a second (e.g., 1/125, 1/60, 1/30,…)
- Controls the amount of light integrated by the sensor
- Faster shutter speeds prevent 'camera shake' and reduce motion blur but will be noisier unless scene is well illuminated.
- Need to use a tripod for slower shutter speeds!





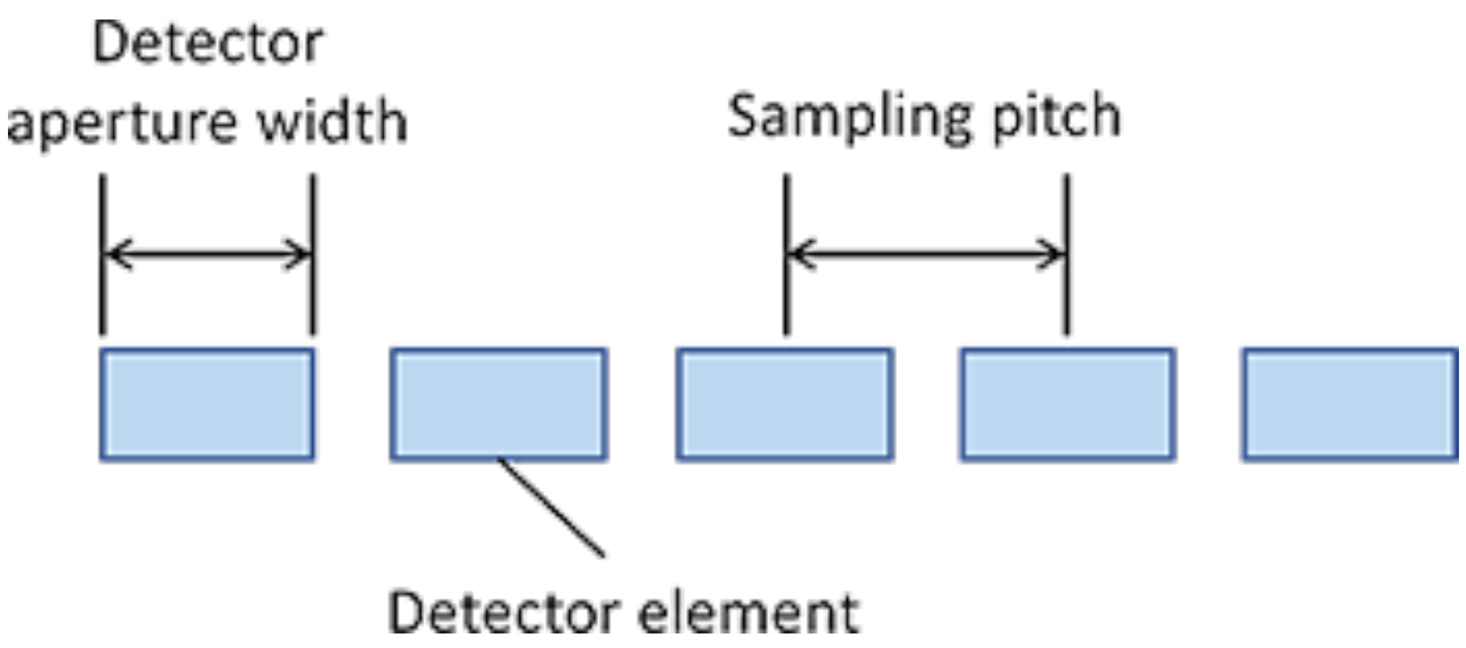


Sampling Pitch & Fill Factor

- Sampling pitch is the physical spacing between adjacent sensor cells.

- For a fixed chip size, smaller pitch means higher resolution (good!) but less light per pixel (bad!)

$$\text{Fill factor} \simeq \frac{\text{Detector aperture width}}{\text{Sampling pitch}}$$



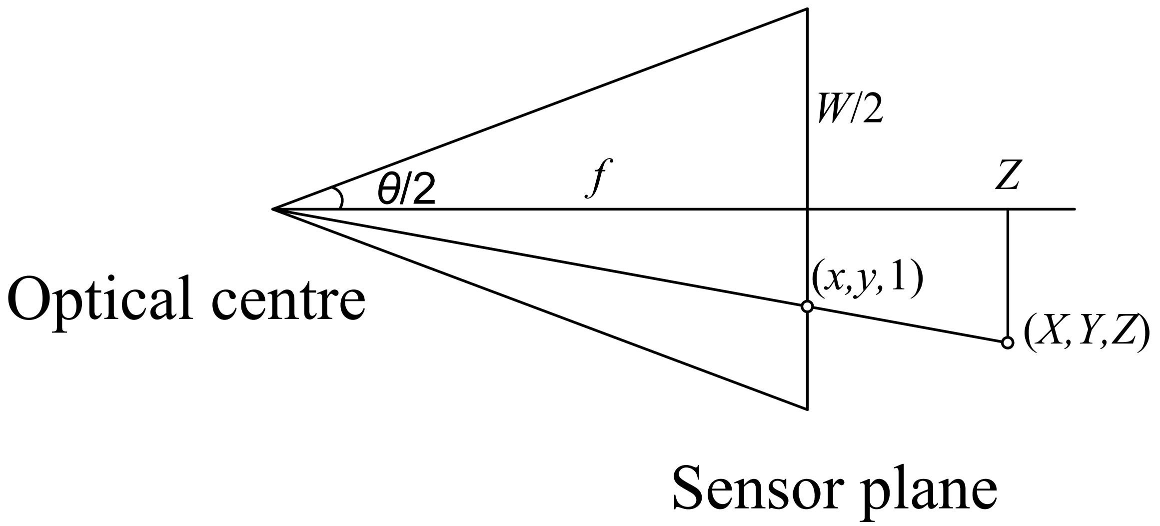
Chip Size

- Chip widths can vary from around ~1/4" to ~1.1"
- Generally this is less than the 35mm width of a standard film frame.
- Our understanding of focal lengths (e.g., a standard 50mm lens) is based on using 35mm film
- To adapt this to a digital camera we must scale by the ratio of the sensor widths.

Focal Lengths

- Focal length can be measured either in pixels or in mm.

$$\tan\frac{\theta}{2} = \frac{W}{2f} \quad \text{or} \quad f = \frac{W}{2}\left[\tan\frac{\theta}{2}\right]^{-1}$$

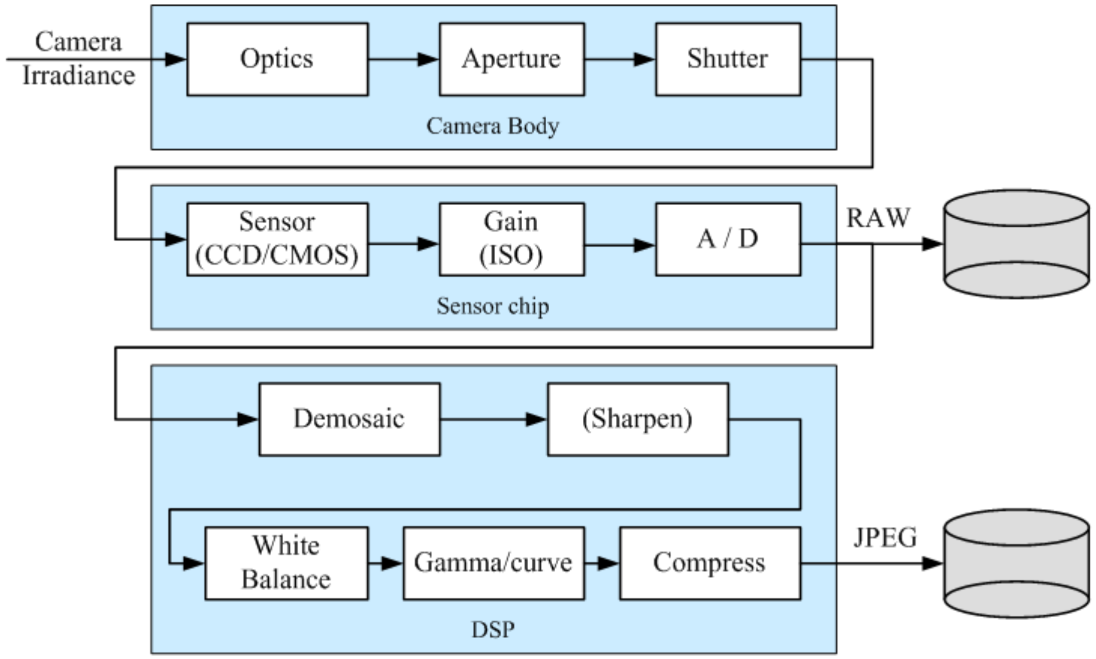


- Example: What focal length would give me the equivalent of a 50mm lens for the FLIR BlackFly S BFS-PGE-122S6C-C?

  - Sensor width: 1.1” = 27.94mm

Analog Gain

- May be controlled through automatic gain control logic
- Can also be adjusted through ISO setting
- Higher gain allows faster shutters speeds (less motion blur) and/or smaller apertures (greater depth of field).
- But at the expense of higher sensor noise!



Sensor Noise

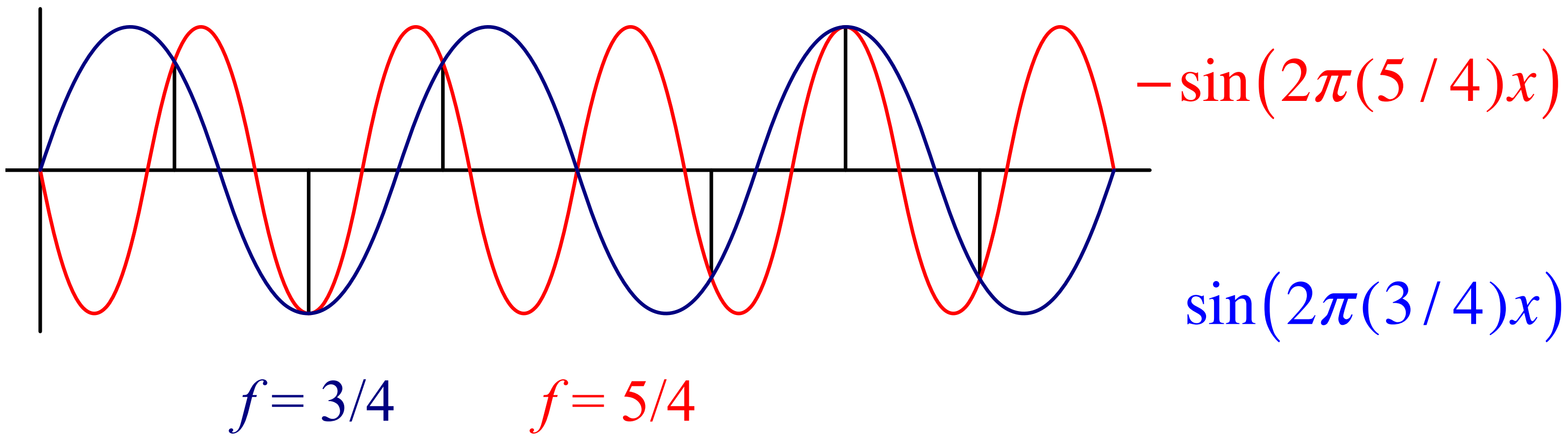
- May include
  - fixed pattern noise
  - dark current noise
  - shot noise
  - amplifier noise
  - quantization noise
- Increases with sensor gain
- Can be estimated (Assignment 1) by
  - Measuring variability when irradiance is constant
  - Differencing two images taken in rapid succession

Outline

- ❖ The Sensor
- ❖ **Sampling & Aliasing**
- ❖ Colour Coding

Sampling & Aliasing

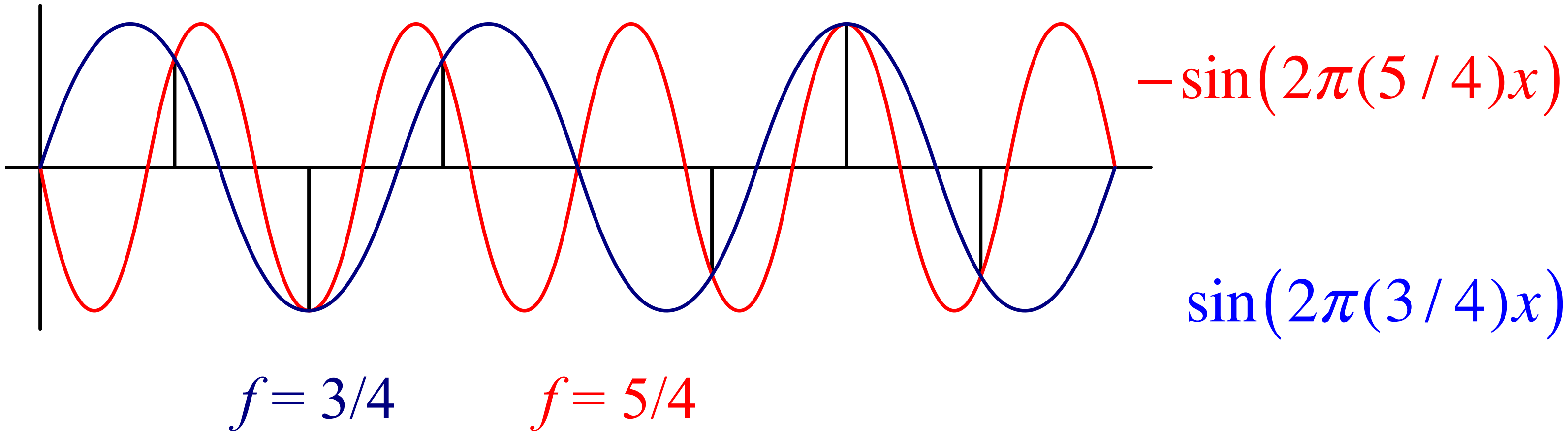
- The optical signal is continuous, containing arbitrarily high spatial frequencies.
- The sensor is spatially sampling this signal at discrete locations determined by the sampling pitch.
- If the image is not low-pass filtered, *aliasing* will result: high frequency content will be inextricably mixed with low frequency content in the digital image.
- Example: sampling rate $f_s = 2$



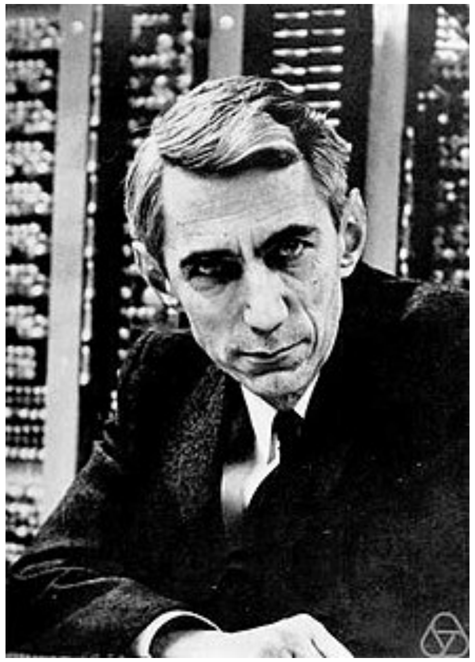
Nyquist Limit

- Shannon's sampling theorem: sampling rate must be at least twice the maximum frequency in the signal.

$$f_s \geq 2 f_{max}$$



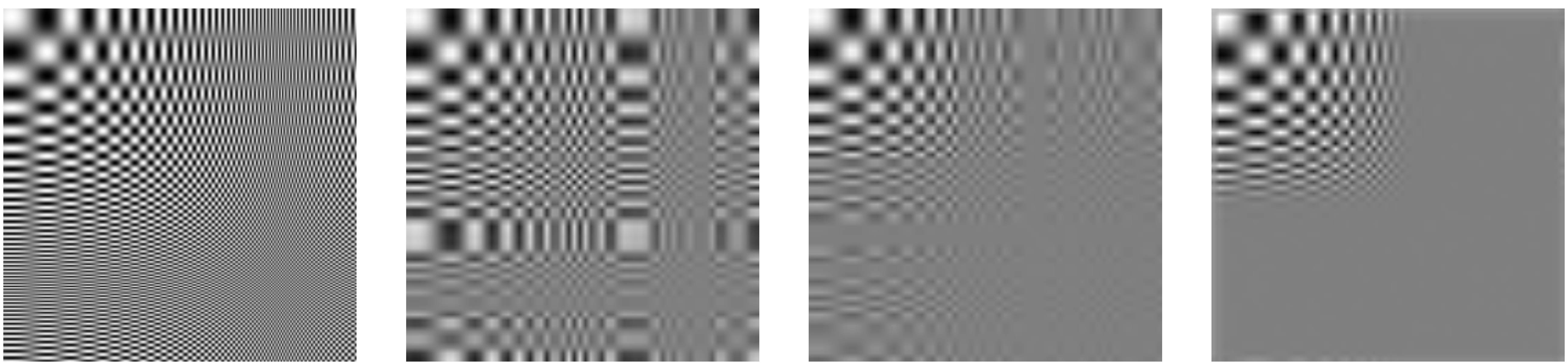
Harry Nyquist (1889 - 1976)



Claude Shannon (1916 - 2001)

Effect of the Fill Factor

- Each pixel is actually the result of integrating light over a small square, the size of which is determined by the sampling pitch and fill factor.
- This serves to attenuate high frequencies.
- However, the Fourier transform of this ‘boxcar’ filter falls only as 1/f, and thus high frequencies, while attenuated, are still present and cause aliasing.



Original Image | Boxcar with 25% fill factor | Boxcar with 100% fill factor | High-quality lowpass filter

Subsampled by a factor of 4

Point Spread Function (PSF)

- The pre-filtering of the optical signal is determined by:
  - The optical system (diffraction, focal blur)
  - The integration area (sampling pitch and fill factor)
  - Integrated optical anti-aliasing filters
- If together these filters adequately attenuate frequencies above the Nyquist limit, visible aliasing will be minimal.

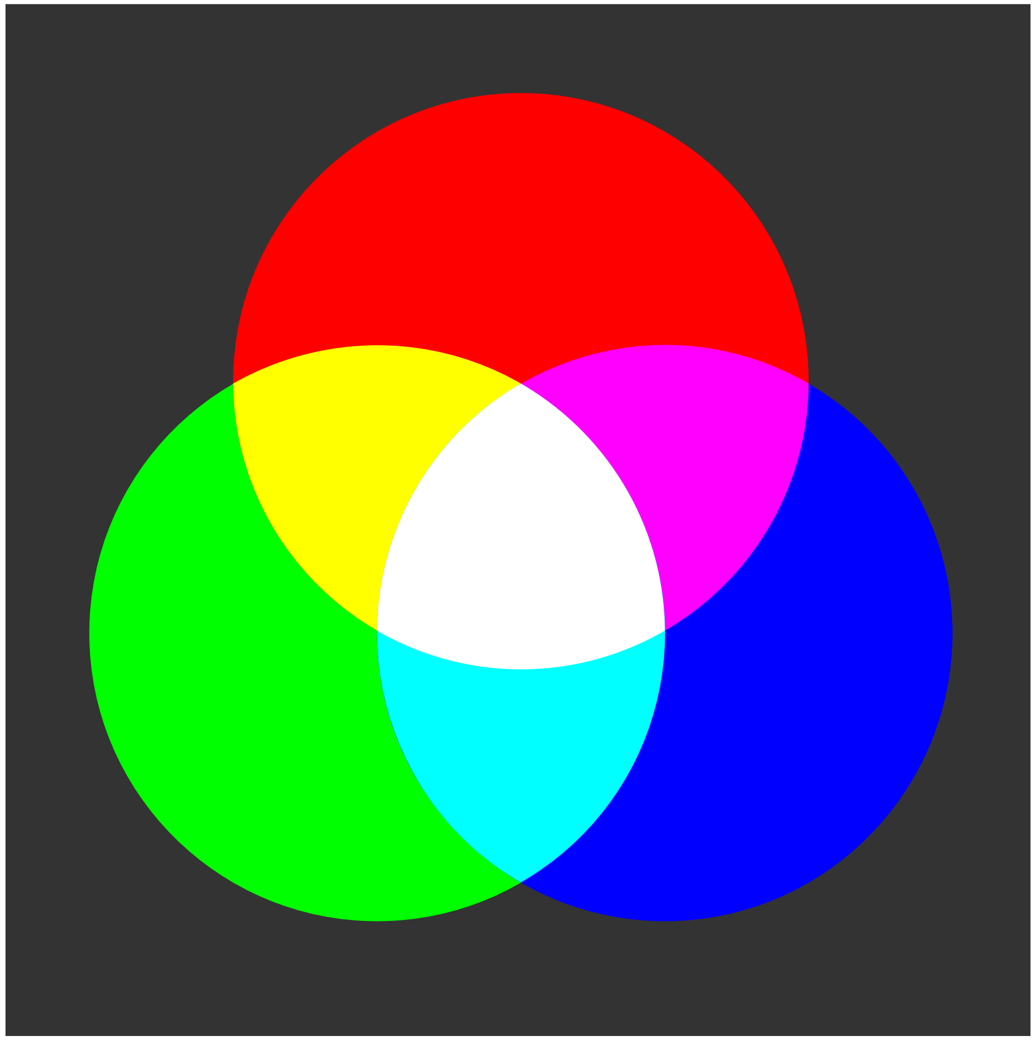
End of Lecture
Sept 19, 2018

Outline

- The Sensor
- Sampling & Aliasing
- **Colour Coding**

Colour Sampling

- Natural scenes reflect light rays over a wide continuum of wavelengths.
- Yet most colour cameras have only 3 discrete types of sensor elements tuned to 3 different colours (wavelengths): red, green and blue.
- Similarly, most colour displays have 3 distinct types of light-emitting elements, also emitting at red, green and blue wavelengths.
- Why is this? Why should this be sufficient?

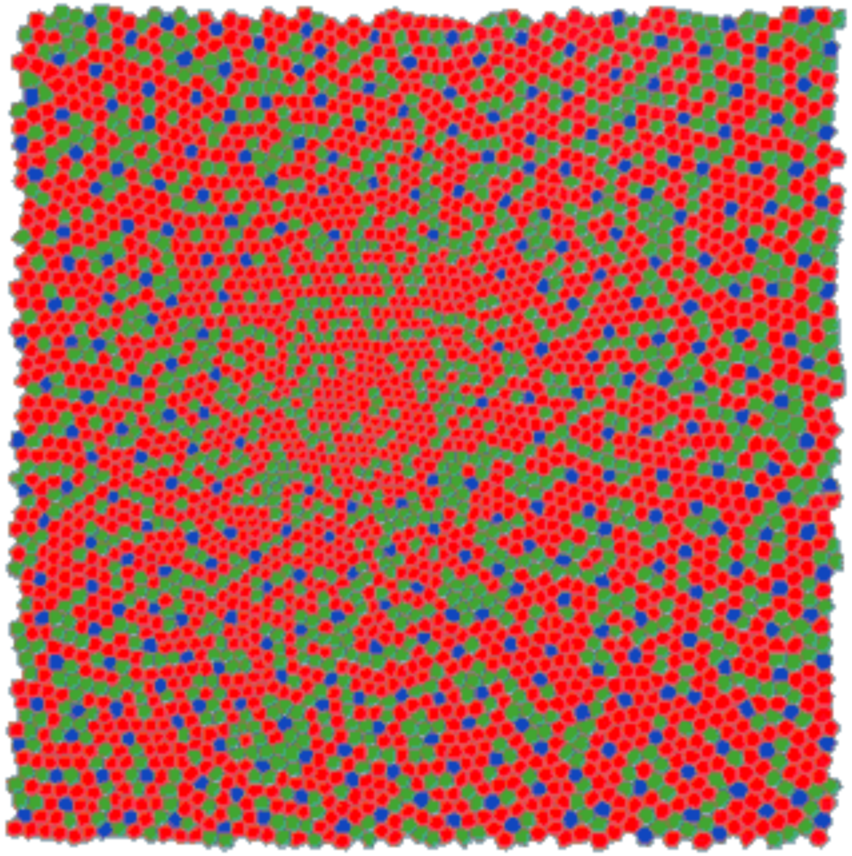


Trichromacy

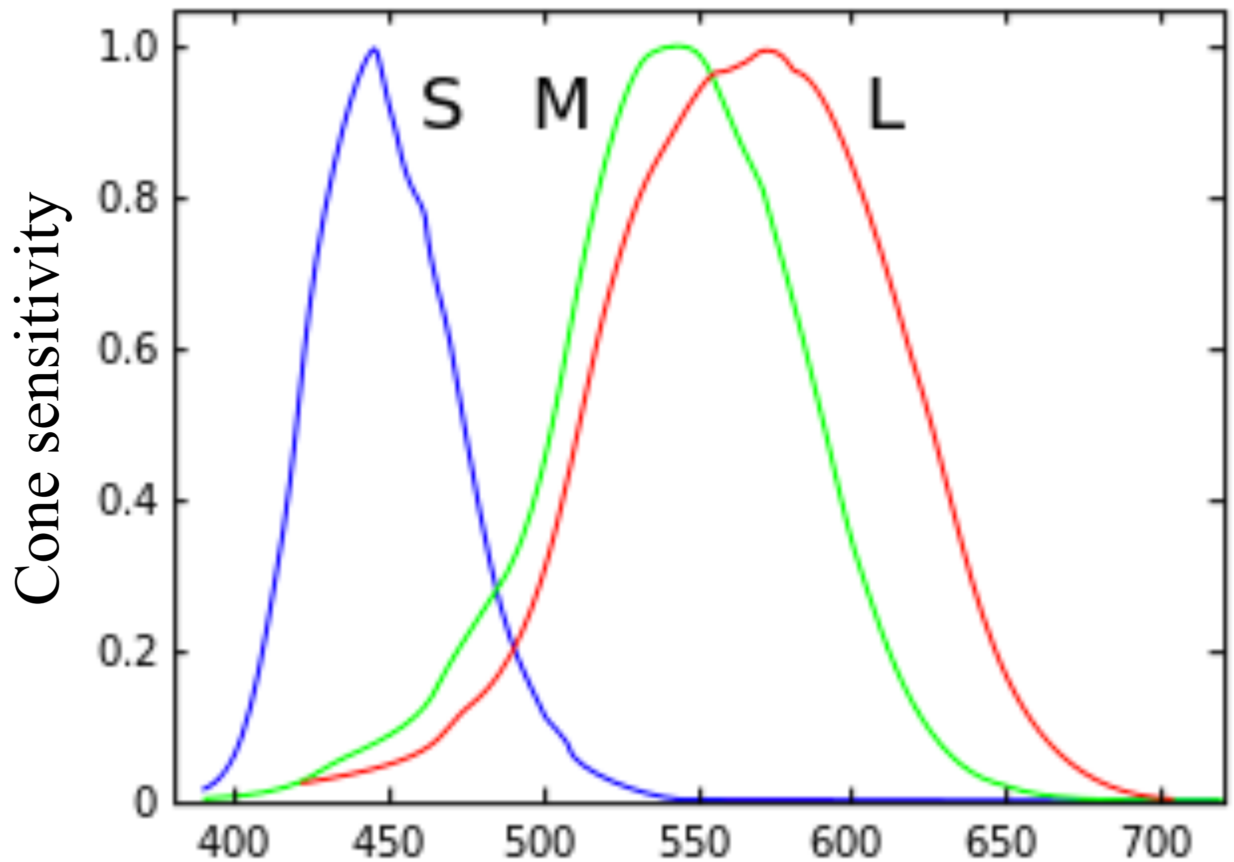
- The human retina has (at most) 3 distinct photoreceptive cone types, each tuned to a specific band of wavelengths.

- This means that human colour vision is 3-dimensional

- Any 3 colour vectors that span this 3D space are sufficient to generate the entire space of colours that we experience.



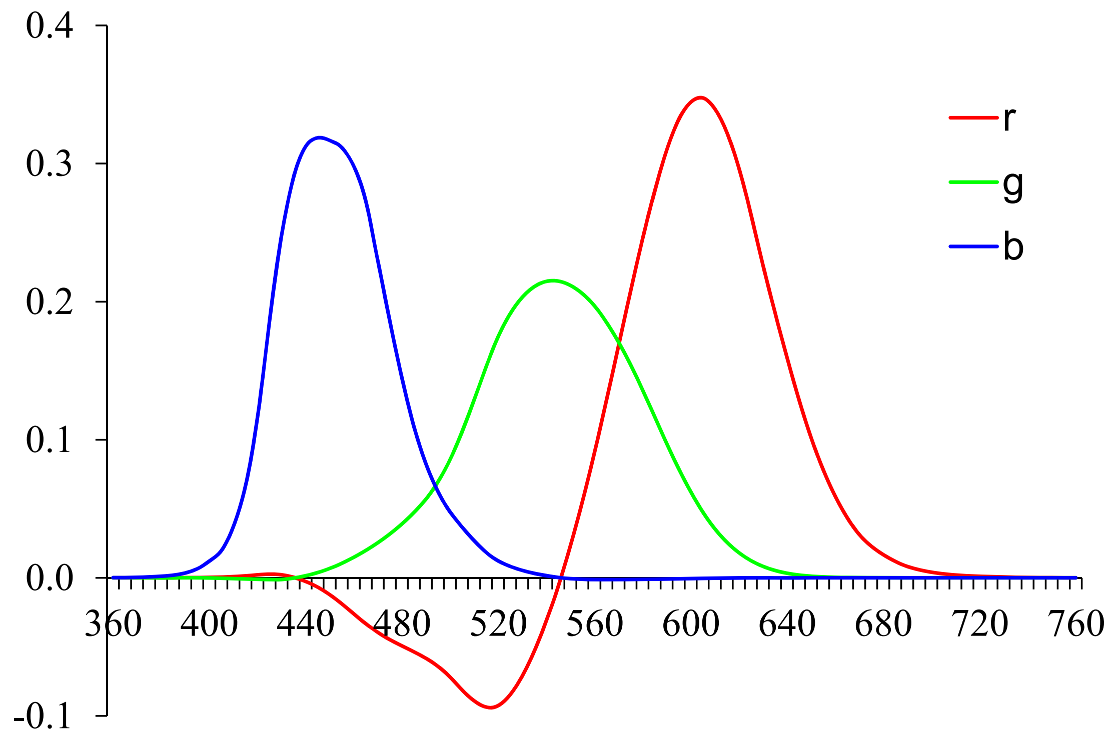
Human cone mosaic



Wavelength (nm)

CIE RGB Representation

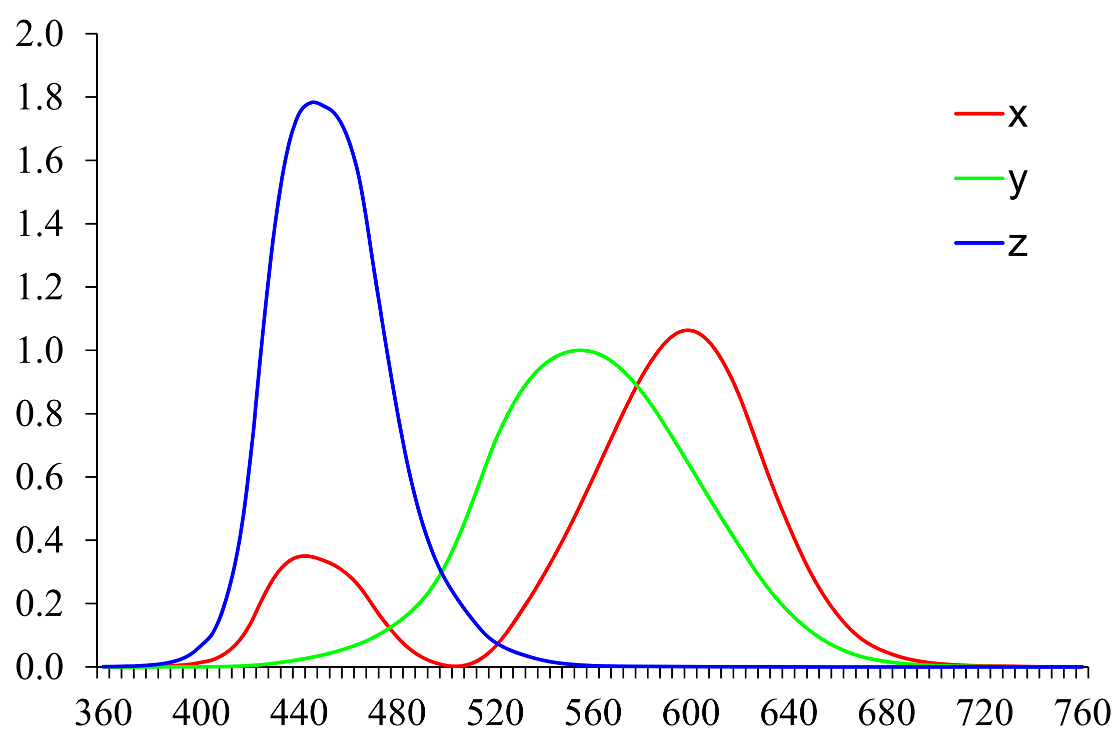
- Colour representation standard formed in 1931
- Based on human behavioural colour matching to monochromatic test colours.
- Subjects adjusted the relative amplitudes of 3 monochromatic primaries:
  - Red (700.0nm)
  - Green (546.1nm)
  - Blue (435.8nm)
- Note that reproducing pure spectra in the blue-green range requires a negative amount of red light!



CIE XYZ Representation

- To avoid this negative light problem, the CIE created a new *XYZ* standard based on a linear transformation of the RGB standard.
- In the XYZ representation, the Y channel corresponds to (achromatic) luminance.
- Note that, unlike the CIE RGB space, the XYZ dimensions are 'imaginary' primary colours having no physical reality.

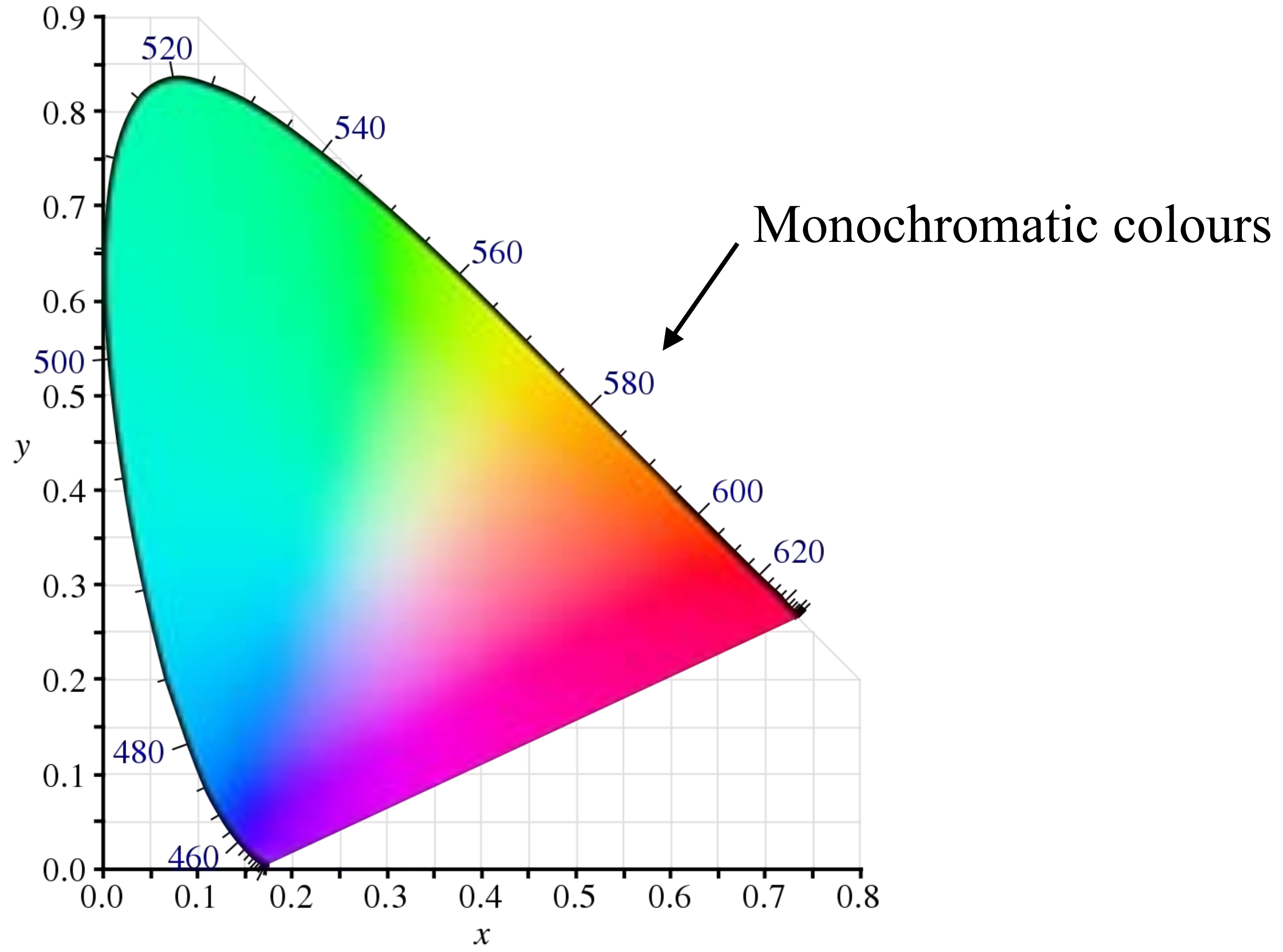
$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \frac{1}{0.17697} \begin{bmatrix} 0.49 & 0.31 & 0.20 \\ 0.17697 & 0.81240 & 0.01063 \\ 0.00 & 0.01 & 0.99 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$



In MATLAB:
- **rgb2xyz**(rgb)
- **xyz2rgb**(xyz)

Chromaticity Coordinates

$$x = \frac{X}{X+Y+Z}, \quad y = \frac{Y}{X+Y+Z}, \quad z = \frac{Z}{X+Y+Z}$$



L*a*b* Space

❖ Human luminance/colour sensitivity is roughly logarithmic

◉ We can perceive relative differences of about 1%.

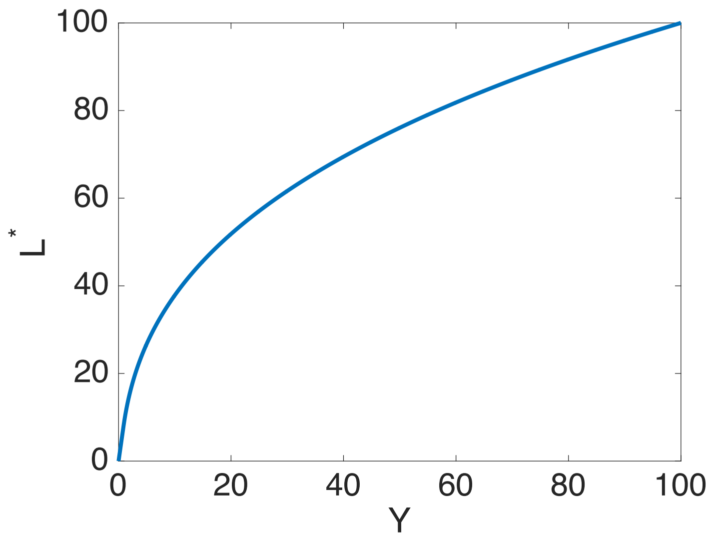
❖ Since XYZ space is linear with the amplitude of the stimulus, it does not predict human perception of colour and luminance differences.

❖ L*a*b space is a nonlinear remapping of XYZ space that renders differences in luminance and chrominance more perceptually uniform.

The L* component of *lightness* is defined as

NB: Error in textbook

$$L^* = 116 f\left(\frac{Y}{Y_n}\right) - 16$$



where Y_n is the luminance value for nominal white (Fairchild 2005) and

$$f(t) = \begin{cases} t^{1/3} & t > \delta^3 \\ t/(3\delta^2) + 2\delta/3 & \text{else,} \end{cases}$$

where $Y_n = 100,\ \delta = 6/29.$

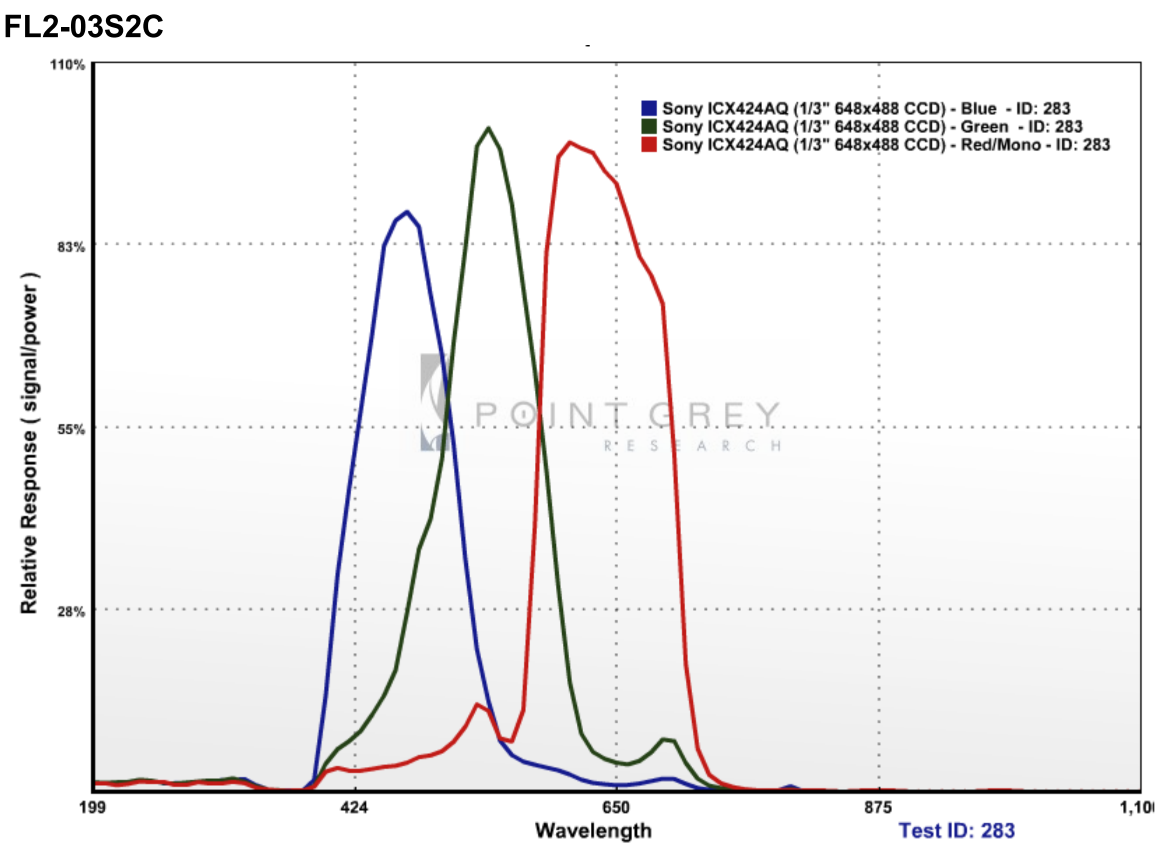
In MATLAB:

- **rgb2lab**(rgb)
- **lab2rgb**(lab)
- **xyz2lab**(xyz)
- **lab2xyz**(lab)

Colour Cameras

- Spectral sensitivities vary from camera to camera.
- It's the job of the camera firmware to convert these proprietary sensor responses to standard colour values.
- For some professional and scientific cameras, the manufacturer provides the spectral responses.

Spectral response curves for Point Grey Flea2

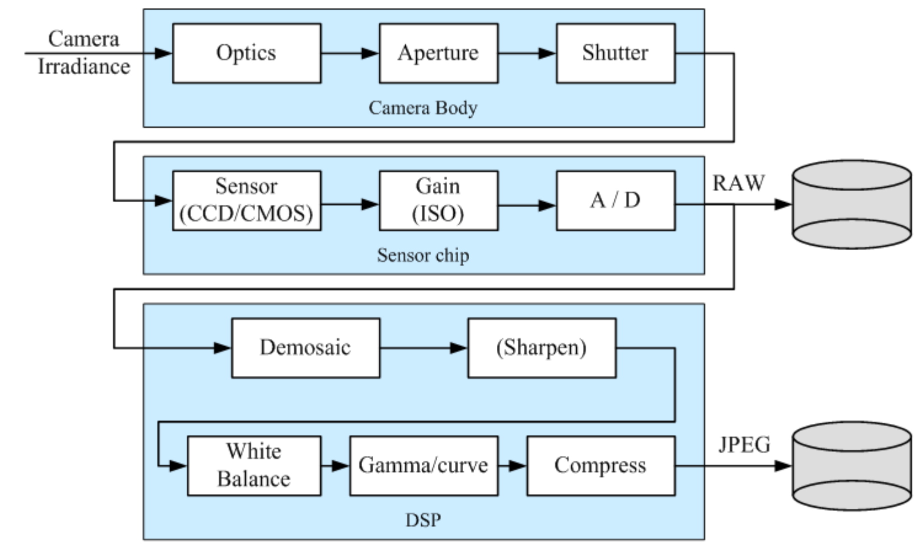


Colour Filter Arrays

- Colour camera sensors consist of a mosaic of sensing elements covered by different coloured filters.

- The most common design is the Bayer pattern, consisting of
  - 25% red
  - 50% green
  - 25% blue

G	R	G	R
B	G	B	G
G	R	G	R
B	G	B	G

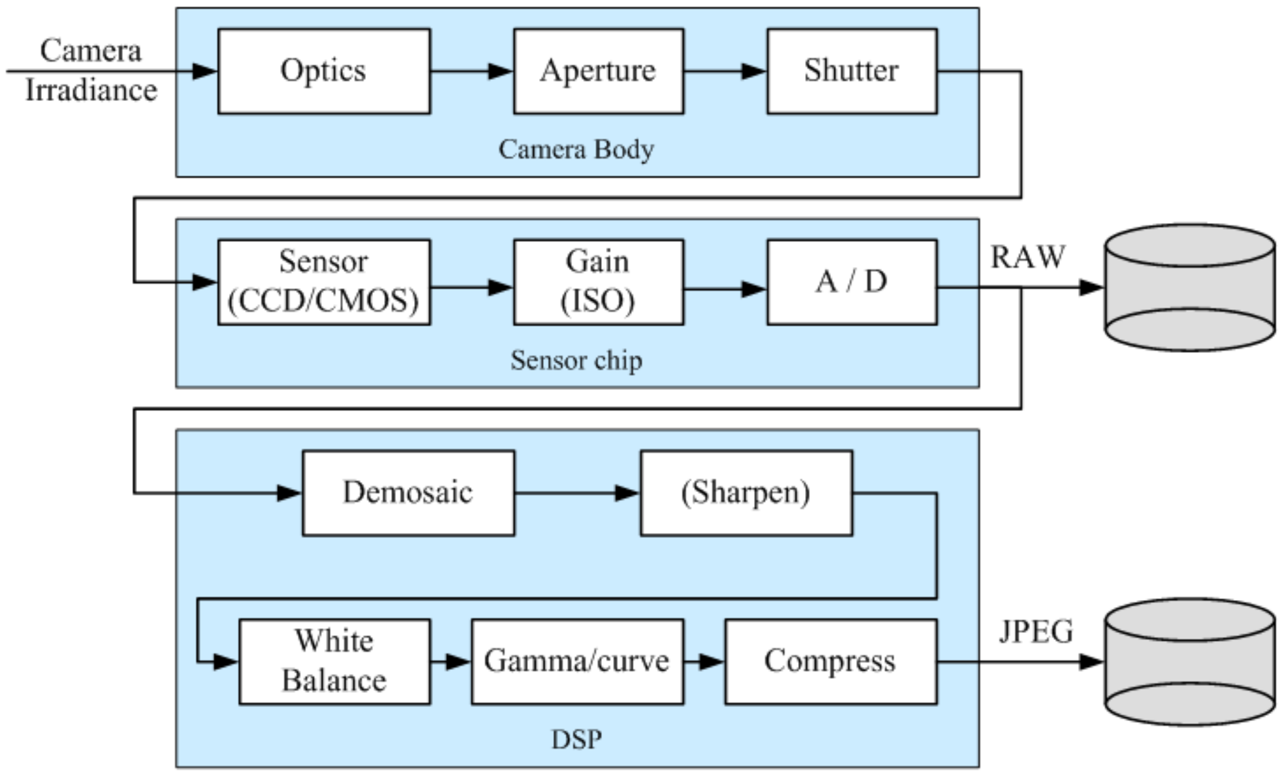


- The greater density of green elements reflects the fact that
  - perceived luminance depends primarily on the green channel
  - visual acuity is far greater for luminance than colour

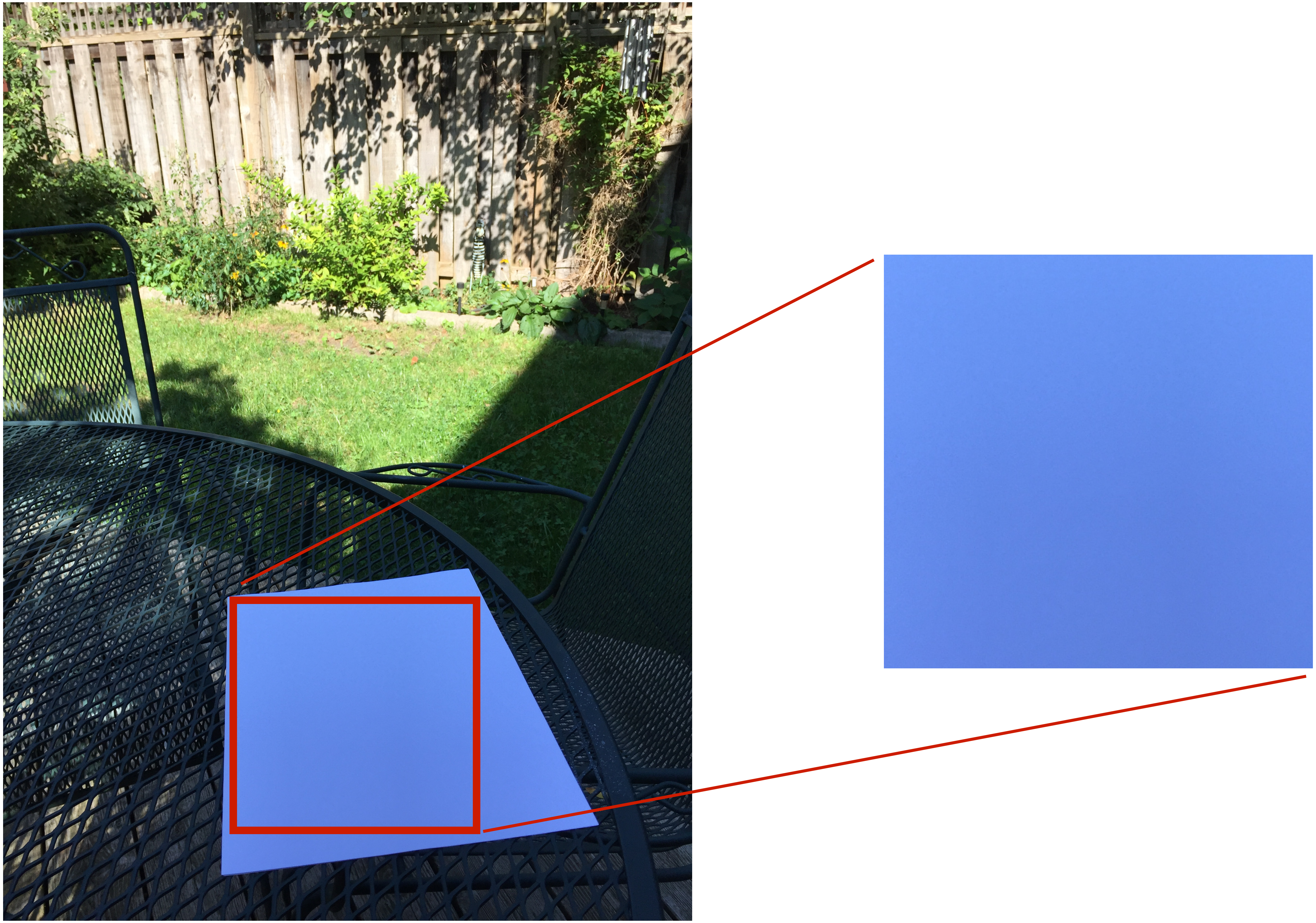
- Interpolation of missing colour values at each pixel known as *demosaicing*.

White Balance

- The colour of the irradiance received from a surface depends upon both the colour of the surface material and the colour of the illuminant.
- Standard colour systems assume a specific illuminant (e.g., daylight)
- If the illuminant deviates from this standard, the resulting photo (out of context) may look oddly coloured.
- White balance is an attempt to reduce this effect by moving the white point of the image closer to pure white (equal RGB values).
- Can achieve this by scaling the R, G and B values by different amounts (Assignment 1).



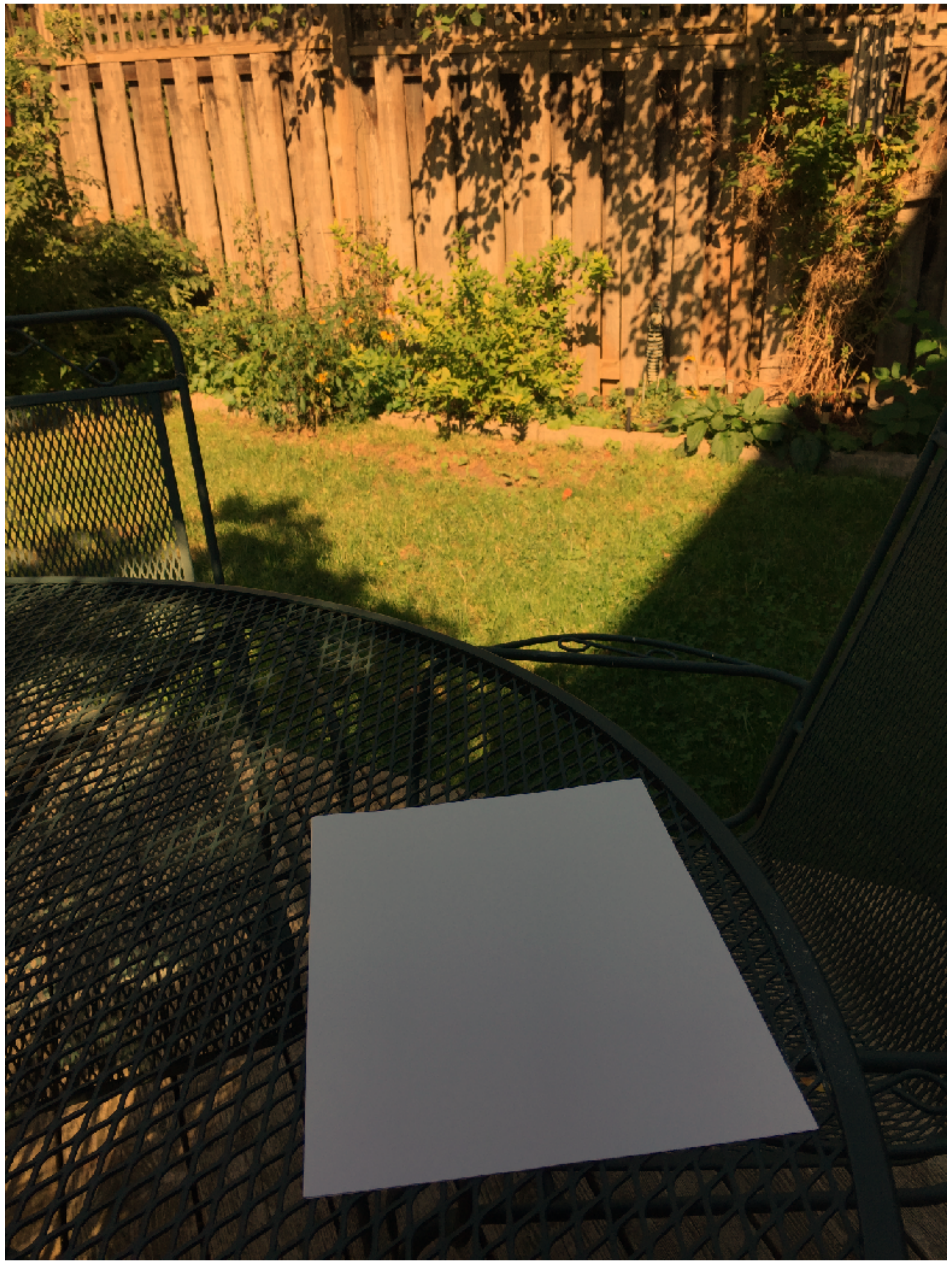
White Balance Example



White Balance Results

Original

White Balanced



Gamma

- Cameras typically compress the intensity (luminance) of pixel values through an inverse 'gamma function':

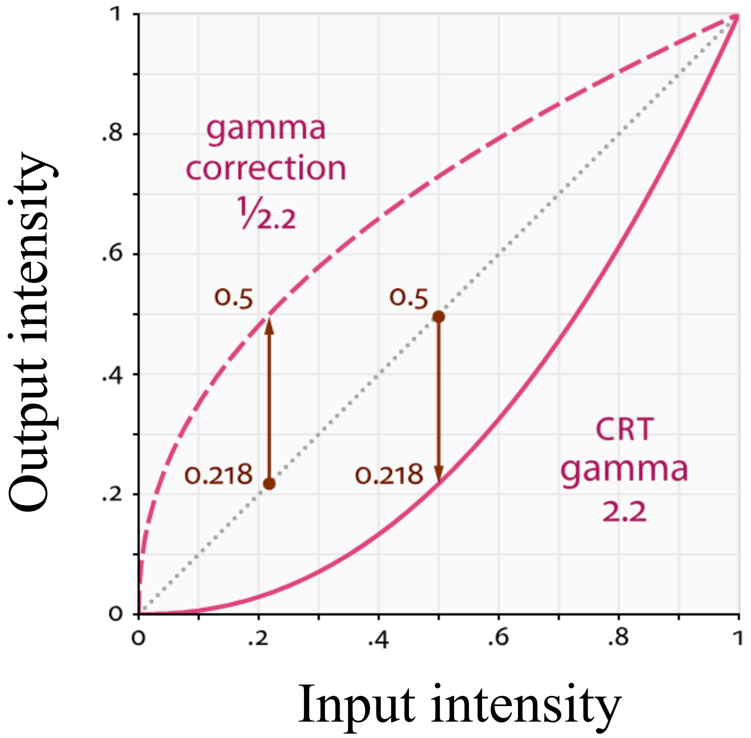
$$Y' = Y^{\frac{1}{\gamma}}$$

where $\gamma \simeq 2.2$.

- This roughly cancels the gamma function applied to RGB values by display systems prior to rendering:

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

- However the nonlinear relationship between encoded RGB values and physical intensities complicates physics-based computer vision algorithms, which often assume access to linear luminance values.



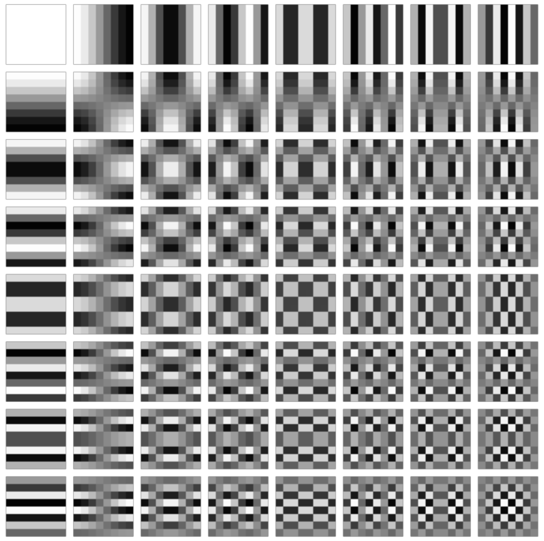
Compression

❖ All compression algorithms start by separating luma and chroma channels so that luma can be encoded with higher fidelity.

❖ Block transform stage then breaks image into disjoint blocks (e.g., 8 x 8 pixels) and codes each using a discrete cosine transform (DCT), which approximates an efficient coding (principal components) strategy.

❖ Resulting DCT coefficients then coded using a variation of Huffman coding.

❖ Video coding uses predictive (difference) encoding between frames, compensating for estimated motion in the image.



DCT Basis Functions

Outline

- The Sensor
- Sampling & Aliasing
- Colour Coding