Multimedia Systems

Media Representation: Image

Dr. Mojtaba Aajami

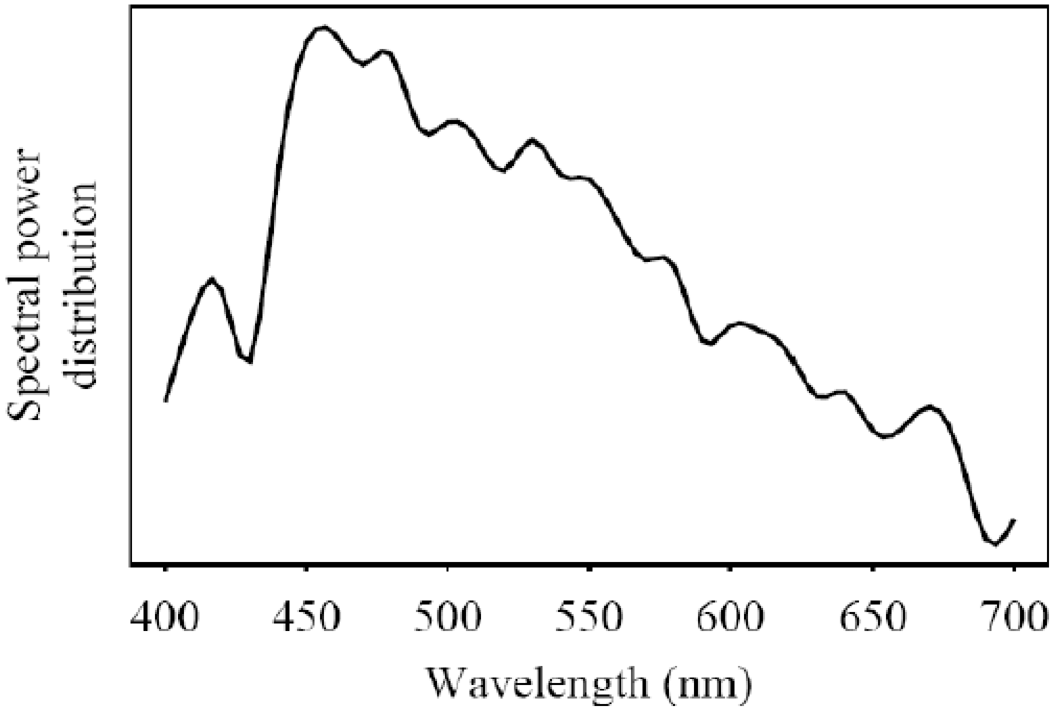
What is image/picture?

- A 2D signal
  - X*Y pixels after sampling
  - Captured by CCD/CMOS
- Each pixel has a color
- Color itself is a multi-dimensional vector, unless for black/white (0/1) or grayscale (scalar)

Color Science

- Light and Spectra
  - Light is an electromagnetic wave. Its color is characterized by the wavelength content of the light.
  - Laser light consists of a single wavelength
  - Short wavelengths produce a blue sensation, long wavelengths produce a red one.
  - Most light sources produce contributions over many wavelengths
  - However, humans cannot detect all light, just contributions that fall in the “visible wavelengths“
  - Electromagnetic waves in the range 400 nm to 700 nm are visible light.

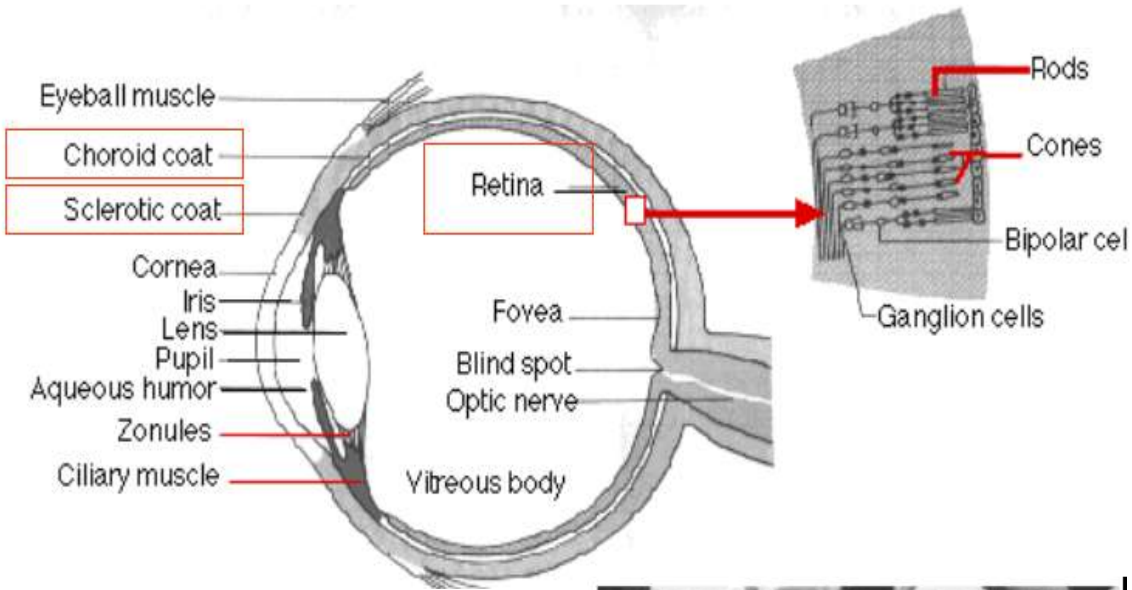
Color Science

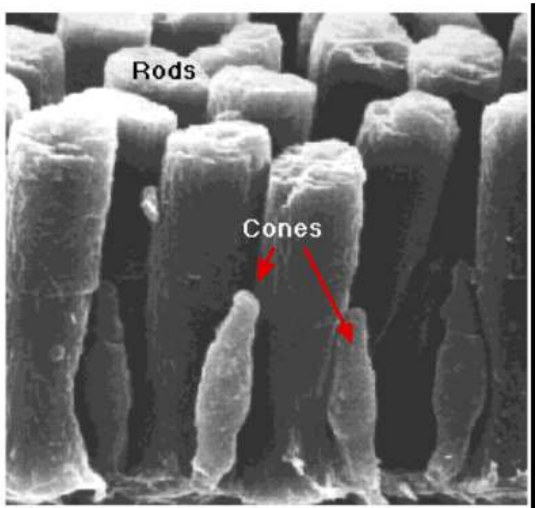


Spectral power distribution of daylight

Human Visual System

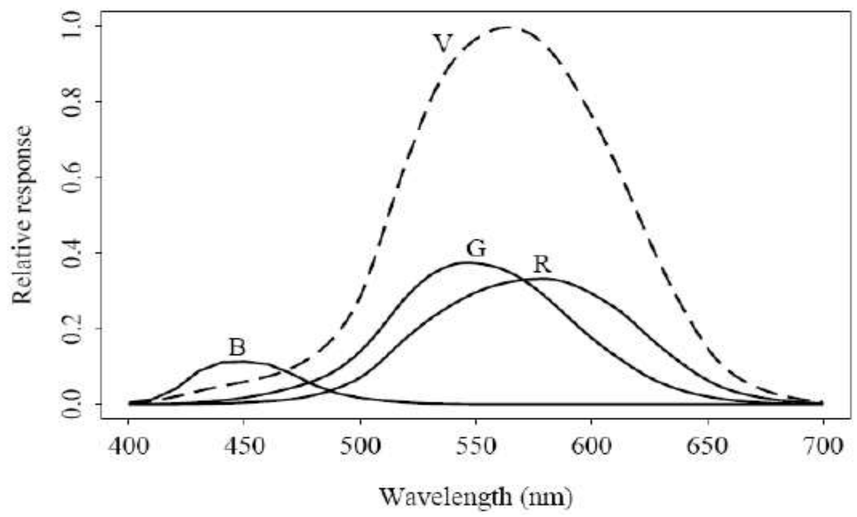
- Sclerotic coat
  - Includes cornea
- Choroid coat
  - Includes iris, pupil
- Retina: contains light receptors
  - Rods
  - Cones that absorb red light (long-wavelength)
  - Cones that absorb green light
  - Cones that absorb blue light (short wl)





Spectral Sensitivity of the Eye

- The eye is most sensitive to light in the middle of the visible spectrum.
- The sensitivity of our receptors is also a function of wavelength.
- The rod sensitivity curve looks like the luminous-efficiency function V(λ) but is shifted to the red end of the spectrum.



Spectral Sensitivity of the Eye

- These spectral sensitivity functions are usually denoted by letters other than "R,G,B"; here let's use a vector function q(λ), with components

$$q(\lambda) = (q_R(\lambda), q_G(\lambda), q_B(\lambda))^T$$

- The response in each color channel in the eye is proportional to the number of neurons firing.
- A laser light at wavelength λ would result in a certain number of neurons firing.

Spectral Sensitivity of the Eye

- An SPD is a combination of single frequency lights (like "lasers"), so we add up the cone responses for all wavelengths, weighted by the eye's relative response at that wavelength.

$$R = \int E(\lambda)\, q_R(\lambda)\, d\lambda$$

$$G = \int E(\lambda)\, q_G(\lambda)\, d\lambda$$

$$B = \int E(\lambda)\, q_B(\lambda)\, d\lambda$$

Image Formation

- Surfaces reflect different amounts of light at different wavelengths, and dark surfaces reflect less energy than light surfaces.
- Fig shows the surface spectral reflectance from orange sneakers and faded blue jeans. The reflectance function is denoted $S(\lambda)$.

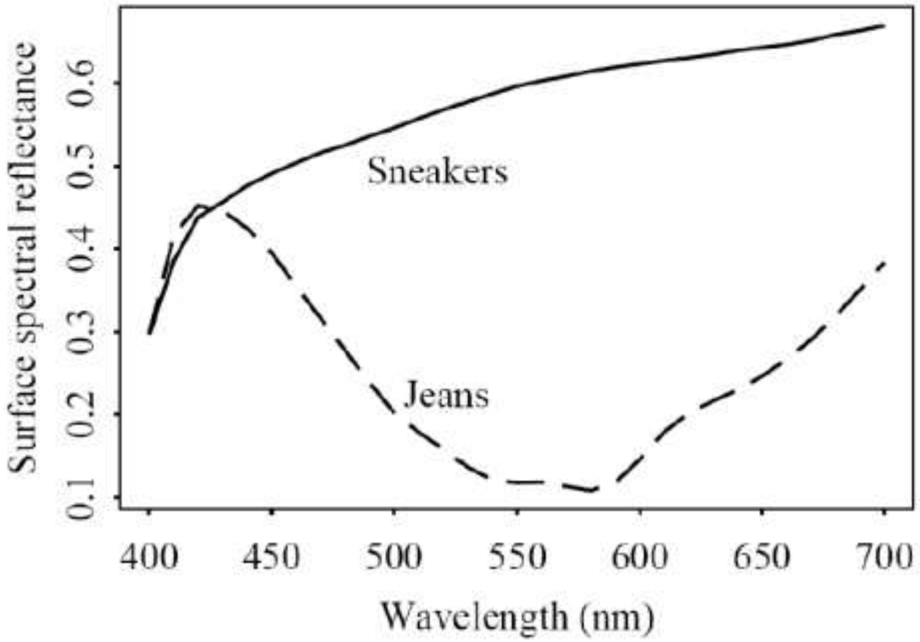


Image Formation Model

- Light from the illuminant with SPD E(λ) impinges on a surface, with surface spectral reflectance function S(λ), is reflected, and then is filtered by the eye's cone functions q(λ).
- The function C(λ) is called the color signal and consists of the product of E(λ), the illuminant, times S(λ), the reflectance:

$$C(\lambda) = E(\lambda)\, S(\lambda).$$

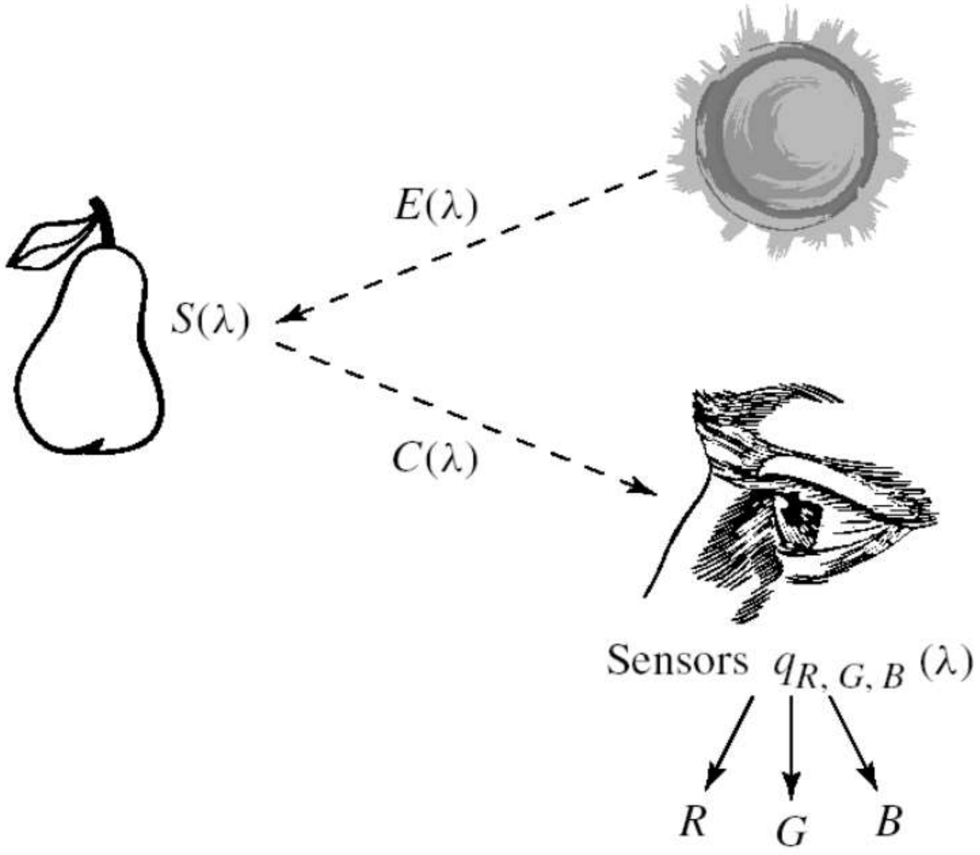


Image Formation Model

- The equations that take into account the image formation model are:

$$R = \int E(\lambda)\, S(\lambda)\, q_R(\lambda)\, d\lambda$$

$$G = \int E(\lambda)\, S(\lambda)\, q_G(\lambda)\, d\lambda$$

$$B = \int E(\lambda)\, S(\lambda)\, q_B(\lambda)\, d\lambda$$

Camera Systems

- Camera systems are made in a similar fashion; a studio quality camera has three signals produced at each pixel location (corresponding to a retinal position).
- Analog signals are converted to digital, truncated to integers, and stored. If the precision used is 8-bit, then the maximum value for any of R,G,B is 255, and the minimum is 0.
- However, the light entering the eye of the computer user is that which is emitted by the screen—the screen is essentially a self-luminous source. Therefore we need to know the light $E(\lambda)$ entering the eye.

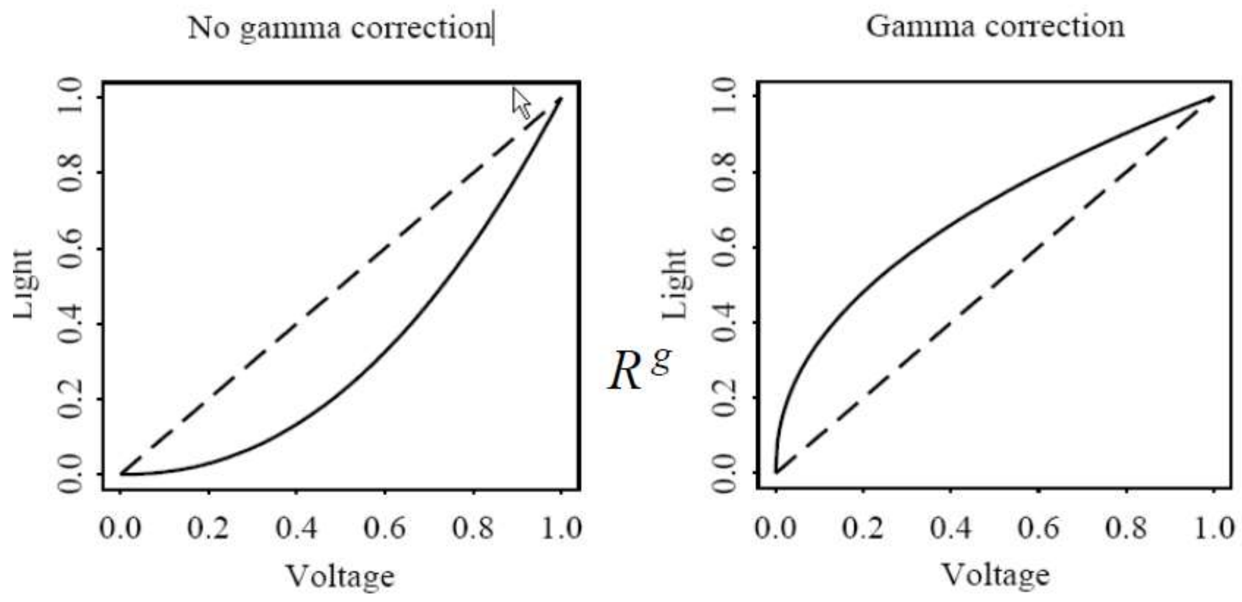
Gamma Correction

- The light emitted is in fact roughly proportional to the voltage *raised to a power*; this power is called **gamma**, with symbol γ.
- Thus, if the file value in the red channel is *R*, the screen emits light proportional to R^{γ}, with SPD equal to that of the red phosphor paint on the screen that is the target of the red channel electron gun. The value of gamma is around 2.2.
- It is customary to append a prime to signals that are **gamma corrected** by raising to the power (1/γ) before transmission. Thus we arrive at **linear signals**:

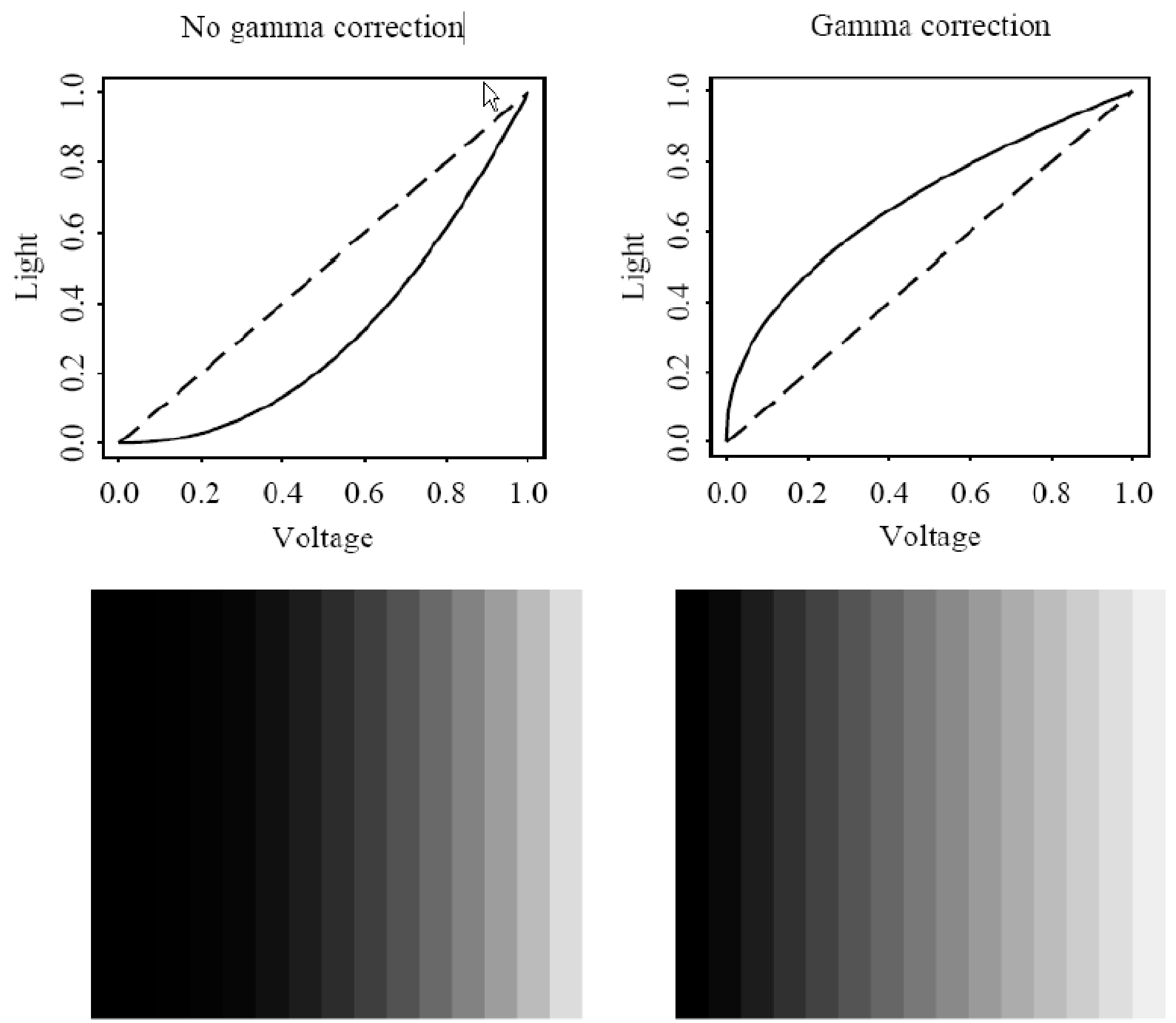
$$R \rightarrow R' = R^{1/\gamma} \Rightarrow (R')^{\gamma} \rightarrow R$$

Gamma Correction

- Left: light output from CRT with no gamma-correction applied. -- Darker values are displayed too dark.
- Right: pre-correcting signals by applying the power law



Gamma Correction



Gamma Correction

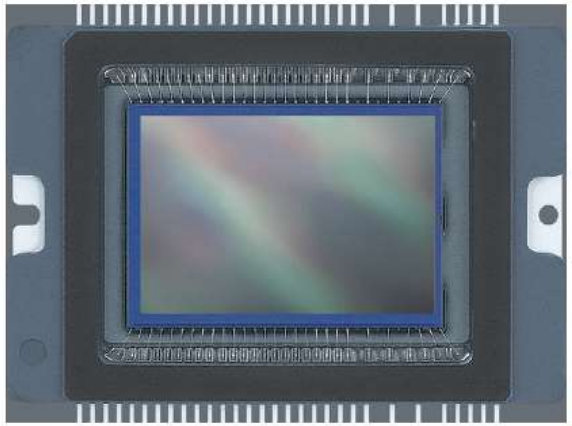
- A more careful definition of gamma recognizes that a simple power law would result in an infinite derivative at zero voltage — makes constructing a circuit to accomplish gamma correction difficult to devise in analog.

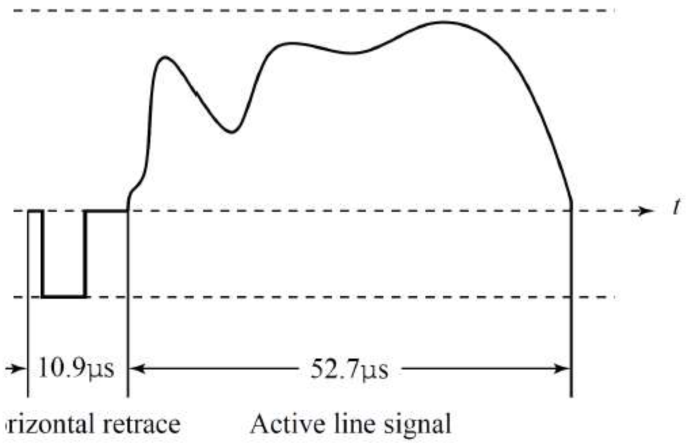
In practice a more general transform, such as $R \rightarrow R' = a \times R^{1/\gamma} + b$ is used, along with special care at the origin:

$$V_{out} = \begin{cases} 4.5 \times V_{in}, & V_{in} < 0.018 \\ 1.099 \times (V_{in}^{0.45} - 0.099), & V_{in} \geq 0.018 \end{cases}$$

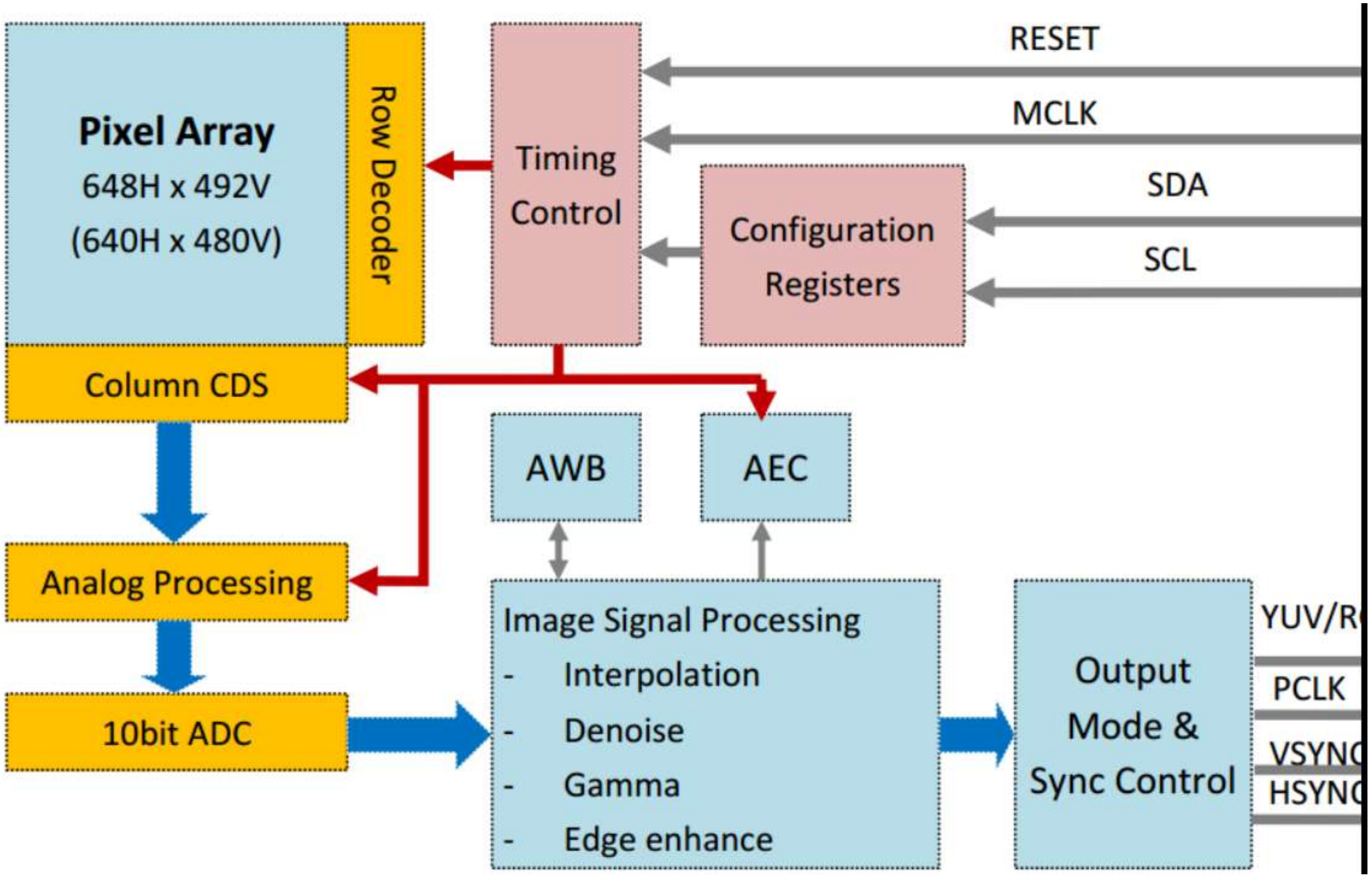
Correction--Camera

- CCD/CMOS are just sensor type.
- Usually 2d matrix array
- Scan through each point to produce electronical signal
- Digital camera has Analog to Digital conversion



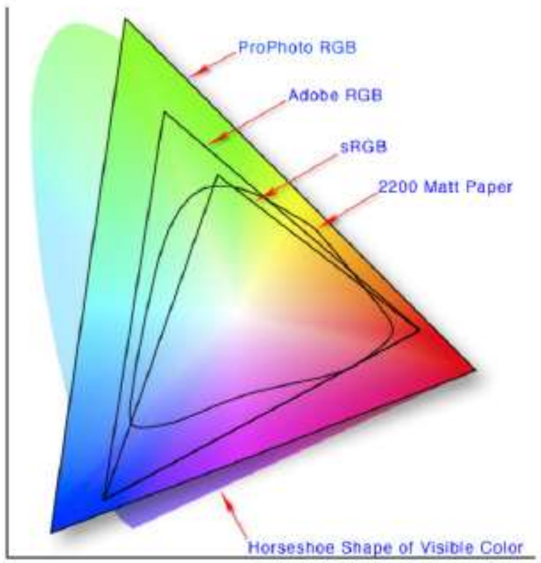


Correction--Camera



Color Space

- All color can be created by mixing basic components.
- Different ways of choosing basic color components.
- RGB Color Space
  - R, G, B components.
  - Usually 8 bits per components: [0, …, 255].
  - Widely used: BMP, TIFF, PPM …
- RGBA Color Space
  - RGB with Alpha (for transparency)
  - Used by PNG format

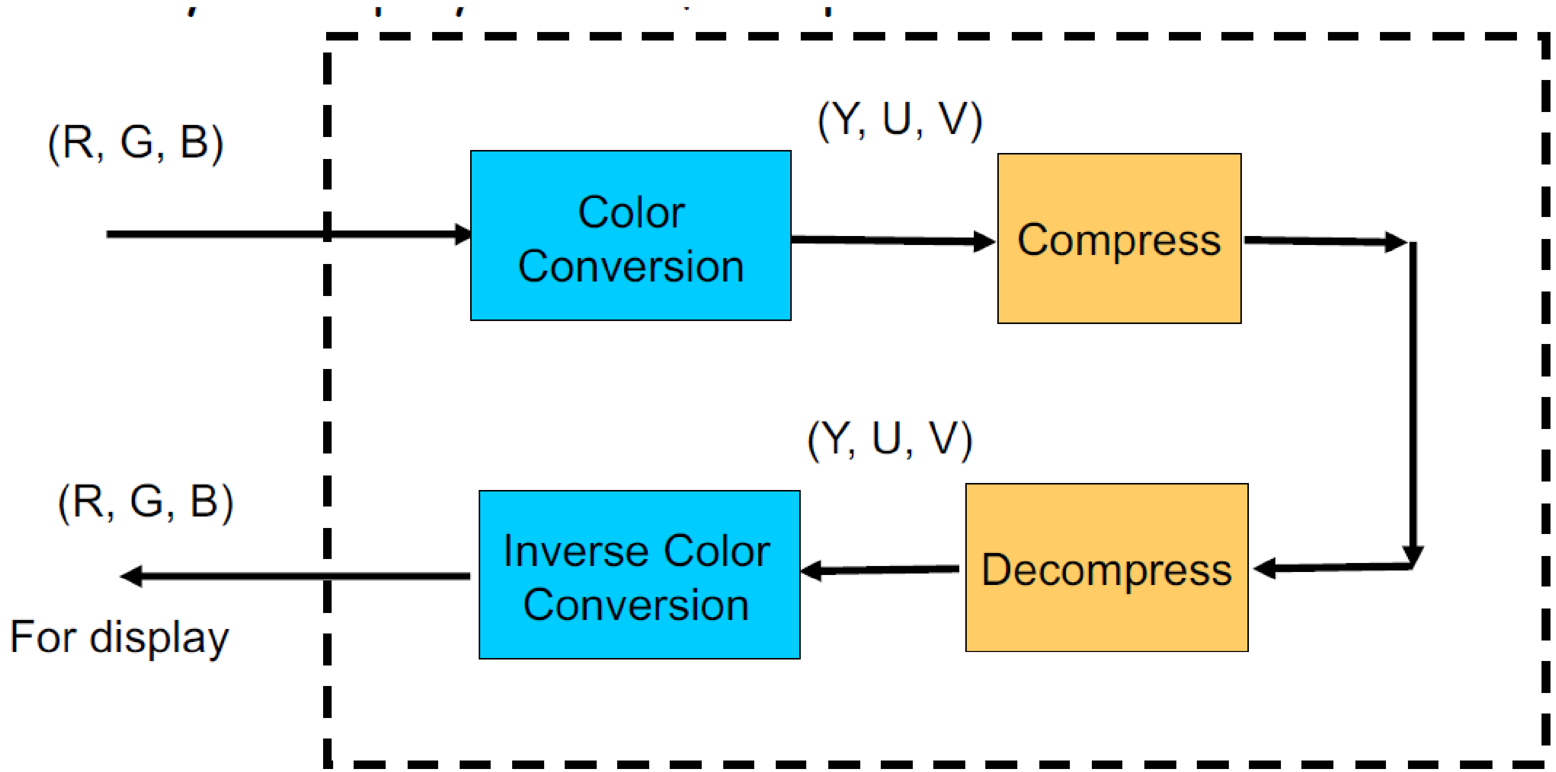


RGB Color Space

- RGB components of an image are strongly correlated.

Color Space: RGB to YUV

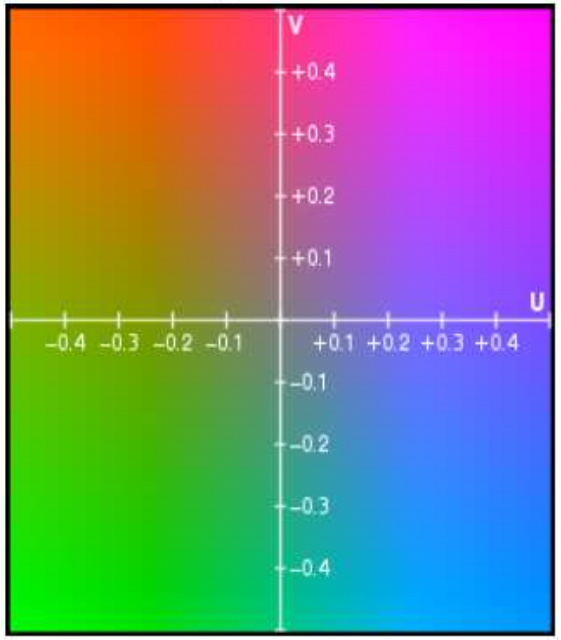
- Converting to other spaces
  - Why ? Display device, compression.



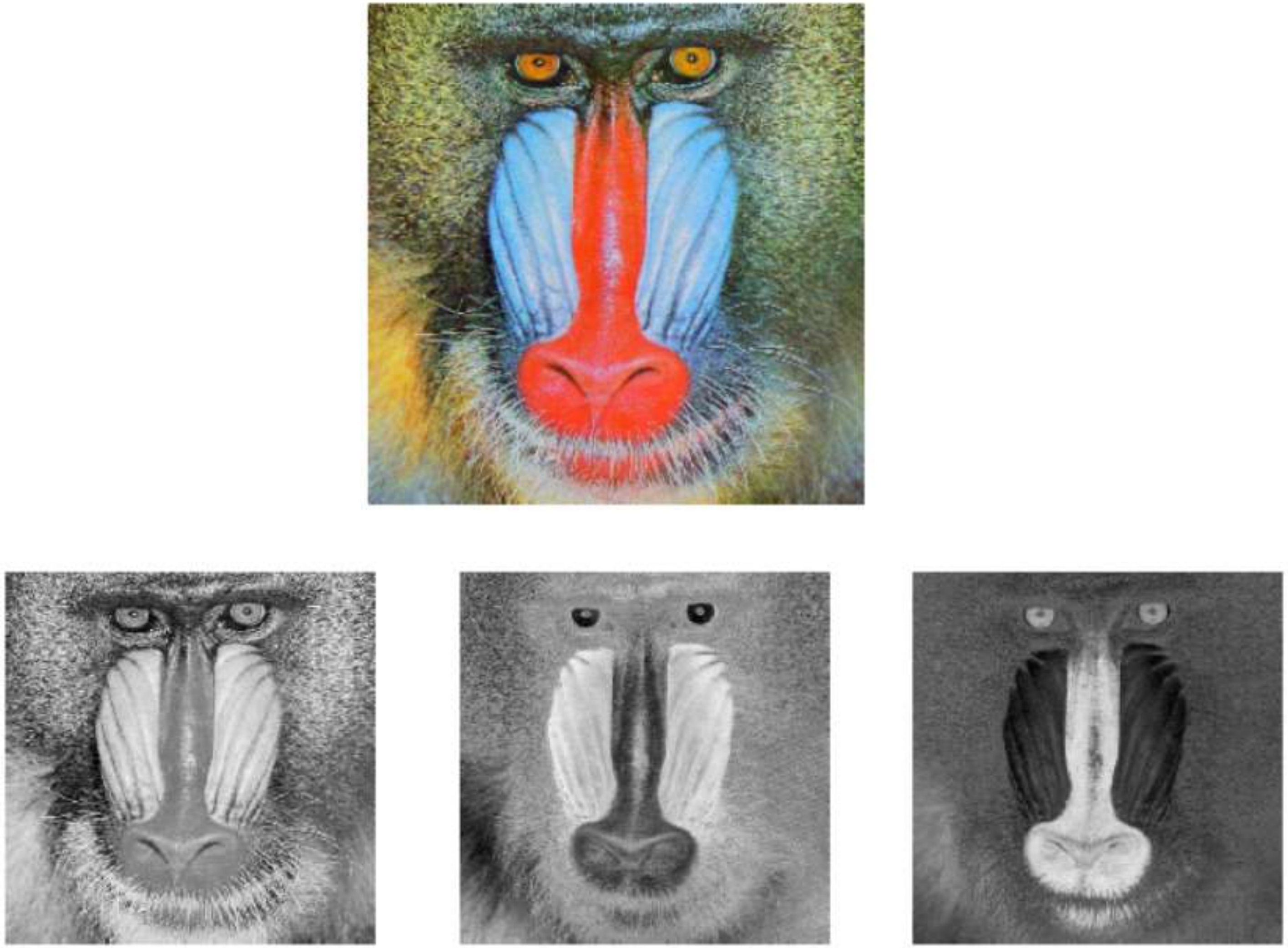
Other Color Spaces

- YUV color space (used by PAL TV system)
  - Y: Luminance component (brightness)
  - U, V: Chrominance components
    - the difference between a color and a reference
- YCrCb Space (used in image/video coding)
  - Derived from YUV
  - U,V shifted by 0.5
  - Components approximately uncorrelated

U-V color plane when Y=0.5



YUV Decomposition



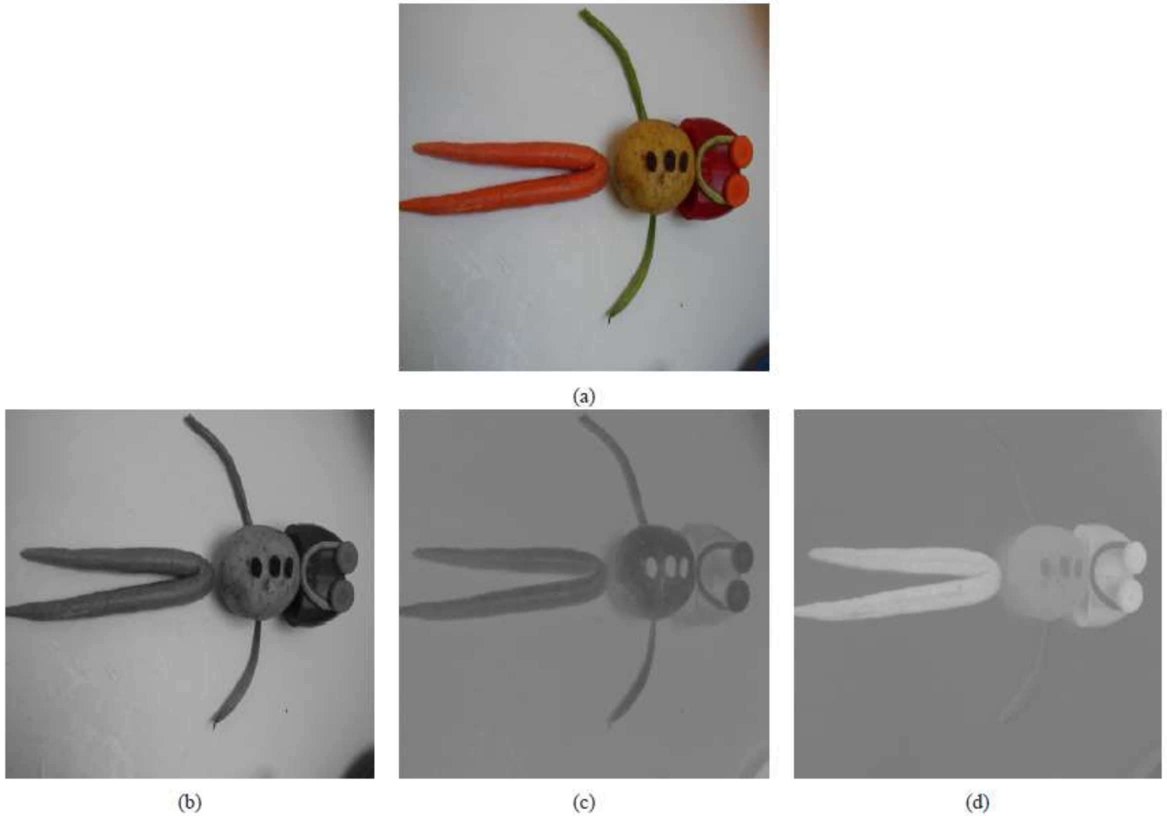
Color Space: YUV

- YUV codes a luminance signal Y(brightness)
- **Chrominance:**
  - The difference between a color and a reference white at the same luminance i.e. U, V

$$U = B' - Y', \qquad V = R' - Y'$$

$$\begin{bmatrix} Y' \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.299 & -0.587 & 0.886 \\ 0.701 & -0.587 & -0.114 \end{bmatrix} \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix}$$

Color Space: YUV



Y 'UV decomposition of color image. Top image (a) is original color image; (b) is Y '; (c,d) are (U, V)

YCbCr Color Model

- YUV is changed by scaling such that Cb is U, but with a coefficient of 0.5 multiplying B′.
- This makes the equations as follows:

$$C_b = ((B' - Y')/1.772) + 0.5$$
$$C_r = ((R' - Y')/1.402) + 0.5$$

Written out:

$$\begin{bmatrix} Y' \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.168736 & -0.331264 & 0.5 \\ 0.5 & -0.418688 & -0.081312 \end{bmatrix} \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} + \begin{bmatrix} 0 \\ 0.5 \\ 0.5 \end{bmatrix}$$

YCbCr Color Model

- In practice, the standard specifies 8-bit coding, with a maximum Y' value of only 219, and a minimum of +16. Cb and Cr have a range of ±112 and offset of +128. If R', G', B' are floats in [0.. + 1], then we obtain Y', Cb, Cr in [0..255] via the transform:

$$\begin{bmatrix} Y' \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 65.481 & 128.553 & 24.966 \\ -37.797 & -74.203 & 112 \\ 112 & -93.786 & -18.214 \end{bmatrix} \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} + \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix}$$

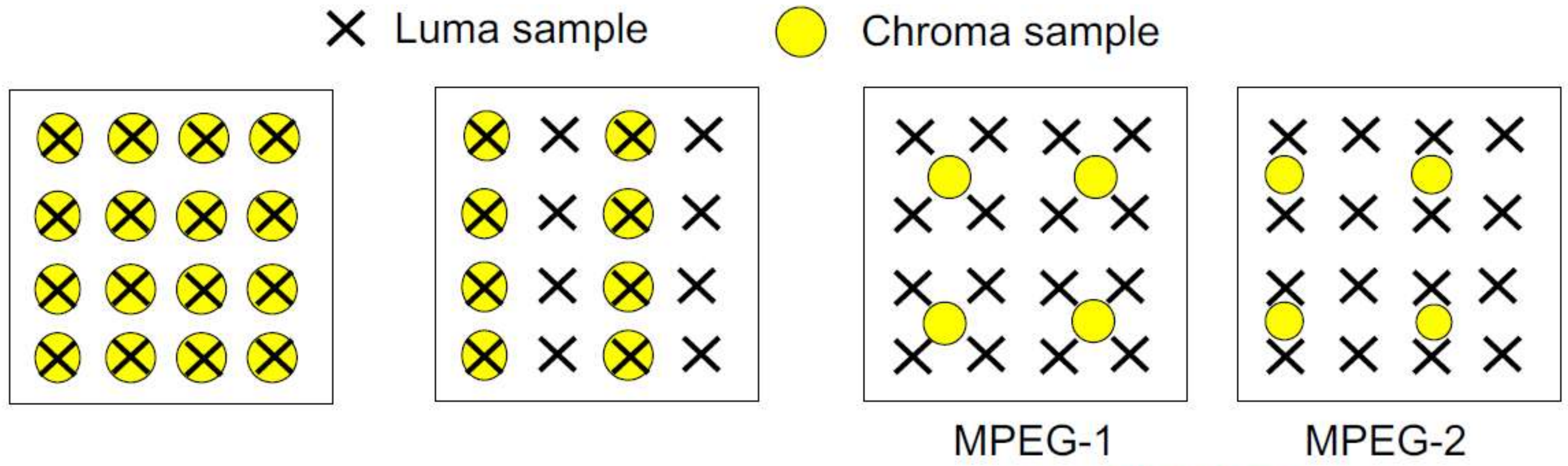
- The YCbCr transform is used in JPEG image compression and MPEG video compression.

RGB vs YCbCr

- Most information is in Y channel (brightness)
  - Cb and Cr are small → easier for compression
- Human eyes are not sensitive to color error
  - Don’t need high resolution for color component

Color Space: Down-sampling

- Down-sampling color components to improve compression



YUV 4:4:4
No downsampling
Of Chroma

YUV 4:2:2
- 2:1 horizontal downsampling of chroma components
- 2 chroma samples for every 4 luma samples

YUV 4:2:0
- 2:1 horizontal downsampling of chroma components
- 1 chroma sample for every 4 luma samples
- Widely used

Raw YUV Data File Format

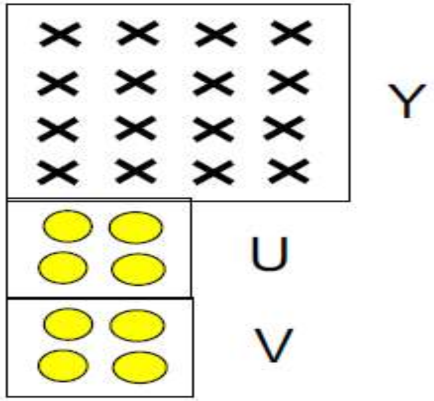
In YUV 4:2:0, number of U and V samples are 1/4 of the Y samples

YUV samples are stored separately:

Image: YYYY.....Y UU...U VV...V

(row by row in each channel)

Video: YUV of frame 1, YUV of frame 2, ...



CIF (Common Intermediate format):

- 352 x 288 pixels for Y, 176 x 144 pixels for U, V

QCIF (Quarter CIF): 176 x 144 pixels for Y, 88 x 72 pixels for U, V

CIF, and QCIF formats are widely used for video conference

Y: 176 x 144

U: 88 x 72

V: 88 x 72