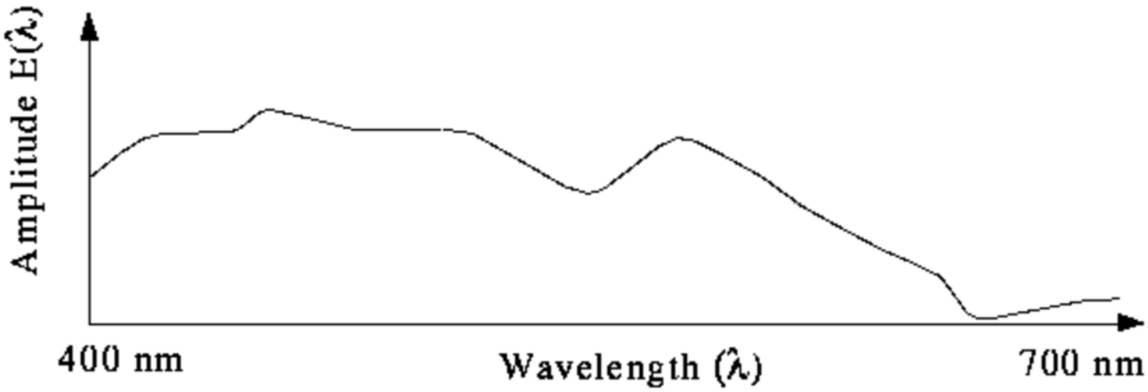
4. Color in Image and Video

4.1. Basics of Color

Light and Spectra

- Visible light is an electromagnetic wave in the 400 nm - 700 nm range.

  Most light we see is not one wavelength, it's a combination of many wavelengths.



- The profile above is called a *spectral power distribution* or *spectrum*.

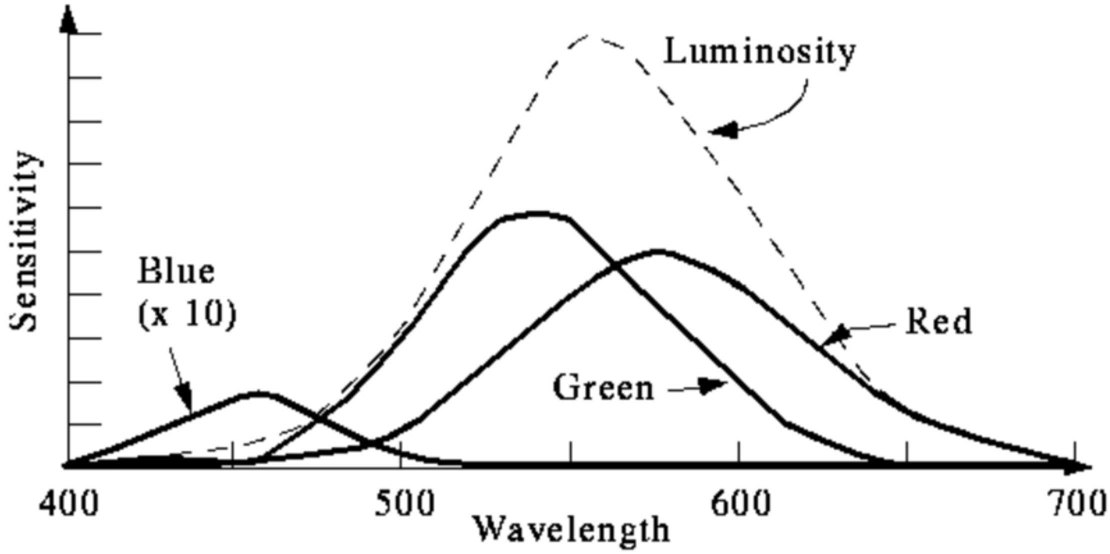
The Human Retina

- The eye is basically just a camera

  Each neuron is either a *rod* or a *cone*. Rods are not sensitive to color.

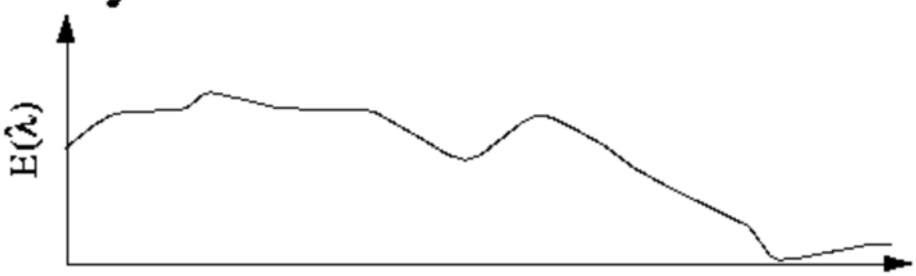
Cones and Perception

- Cones come in 3 types: red, green and blue. Each responds differently to various frequencies of light. The following figure shows the spectral sensitivity functions of the cones and the luminous-efficiency function of the human eye.



- The color signal to the brain comes from the response of the 3 cones to the spectra being observed. That is, the signal consists of 3 numbers:

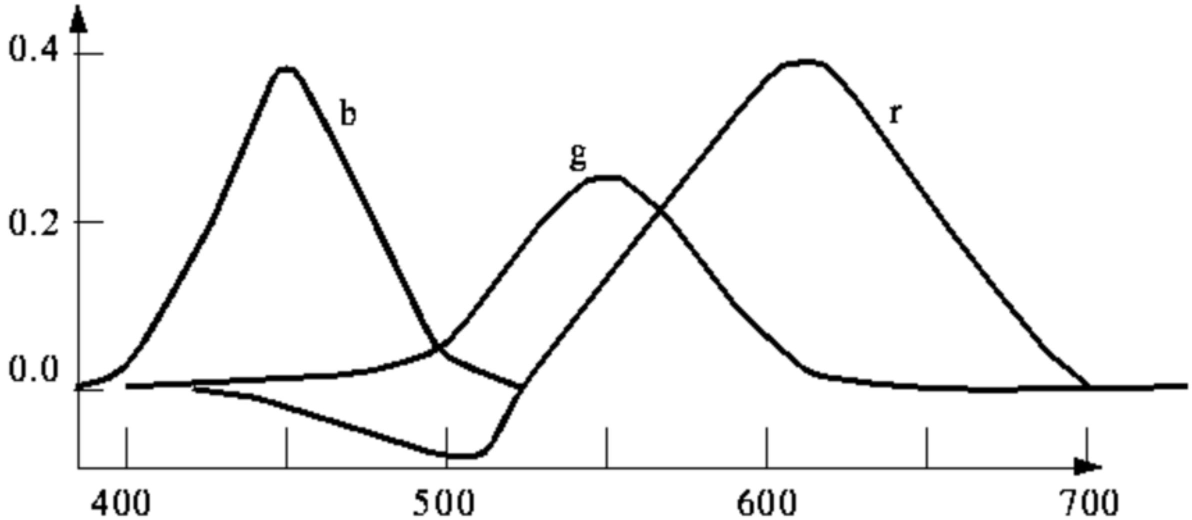
$$R = \int E(\lambda)\, S_R(\lambda)\, d\lambda$$
$$G = \int E(\lambda)\, S_G(\lambda)\, d\lambda$$
$$B = \int E(\lambda)\, S_B(\lambda)\, d\lambda$$



where *E* is the light (spectral power distribution) and *S* are the spectral sensitivity functions.

- A color can be specified as the sum of three colors. So colors form a 3 dimensional vector space.

- The following figure shows the amounts of three primaries needed to match all the wavelengths of the visible spectrum.

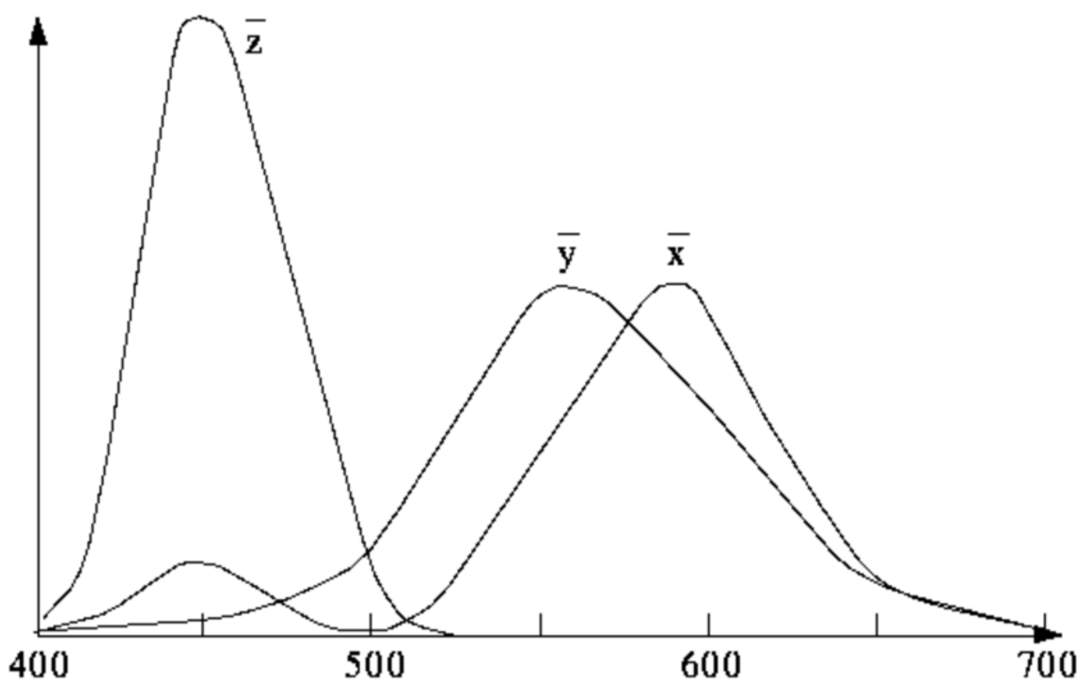


- The negative value indicates that some colors cannot be exactly produced by adding up the primaries.

CIE Chromaticity Diagram

- Q: Does a set of primaries exist that span the space with only positive coefficients?
- A: Yes, but no pure colors.

  In 1931, the CIE (Commission Internationale de L'Eclairage, or International Commission on Illumination) defined three standard primaries **(X, Y, Z)**. The **Y** primary was intentionally chosen to be identical to the luminous-efficiency function of human eyes.



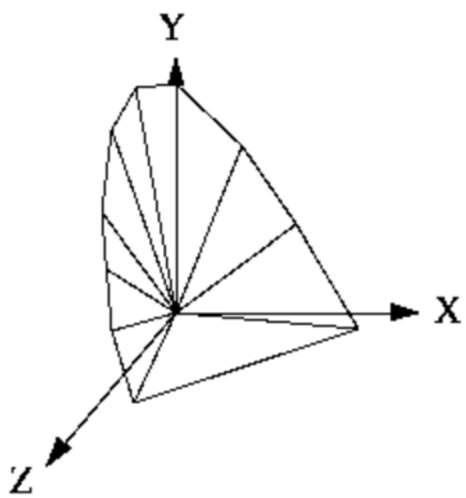
- The above figure shows the amounts of X, Y, Z needed to exactly reproduce any visible color.

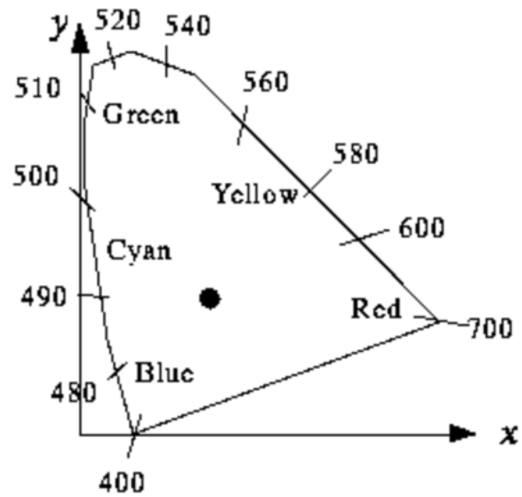
$$X = \int E(\lambda)\, \bar{x}(\lambda)\, d\lambda$$

$$Y = \int E(\lambda)\, \bar{y}(\lambda)\, d\lambda$$

$$Z = \int E(\lambda)\, \bar{z}(\lambda)\, d\lambda$$

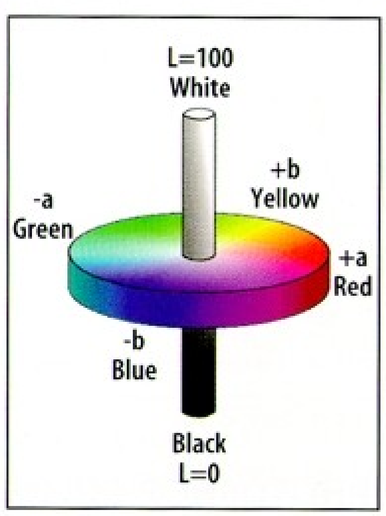
- All visible colors are in a "horseshoe" shaped cone in the X-Y-Z space. Consider the plane $X+Y+Z=1$ and project it onto the X-Y plane, we get the *CIE chromaticity diagram* as below.



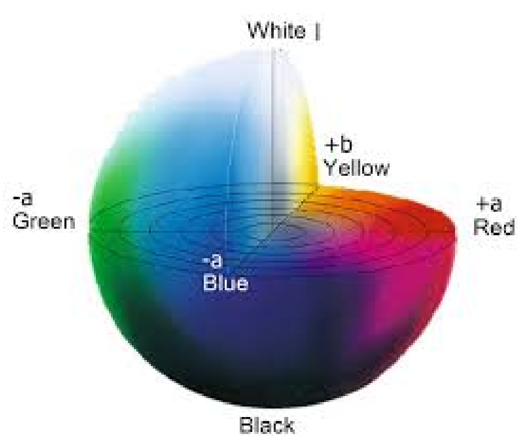


- The edges represent the "pure" colors (sine waves at the appropriate frequency)
- White (a blackbody radiating at 6447 kelvin) is at the "dot"
- When added, any two colors (points on the CIE diagram) produce a point on the line between them.
- Q: how can we find a color's complement on the CIE diagram?

L*a*b (Lab) Color Model



Lab model



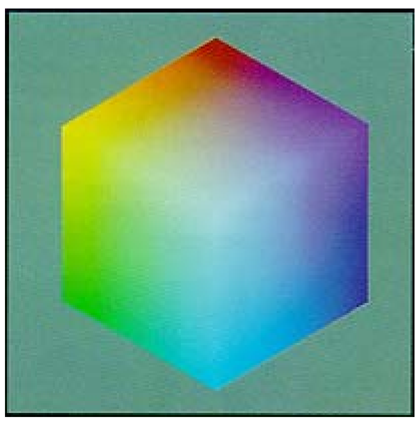
- A refined CIE model, named CIE L*a*b in 1976
- Luminance: L
  Chrominance: a -- ranges from green to red, b -- ranges from blue to yellow
- Used by *Photoshop*

4.2. Color Models in Images

- A color image is a 2-D array of (R,G,B) integer triplets. These triplets encode how much the corresponding phosphor should be excited in devices such as a monitor.

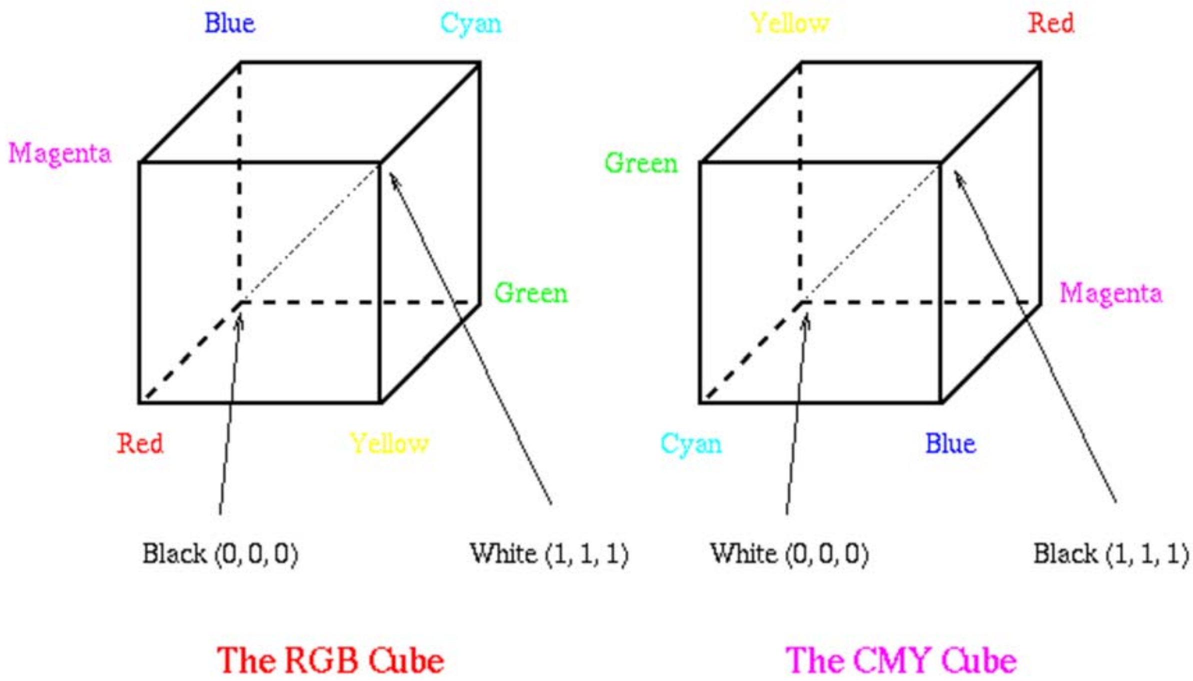
RGB Color Model for CRT Displays

- CRT displays have three phosphors (RGB) which produce a combination of wavelengths when excited with electrons.



CMY Color Model

- Cyan, Magenta, and Yellow (CMY) are complementary colors of RGB. They can be used as *Subtractive Primaries*.
- CMY model is mostly used in printing devices where the color pigments on the paper absorb certain colors (e.g., no red light reflected from cyan ink).



The RGB and CMY Cubes

Conversion between RGB and CMY:

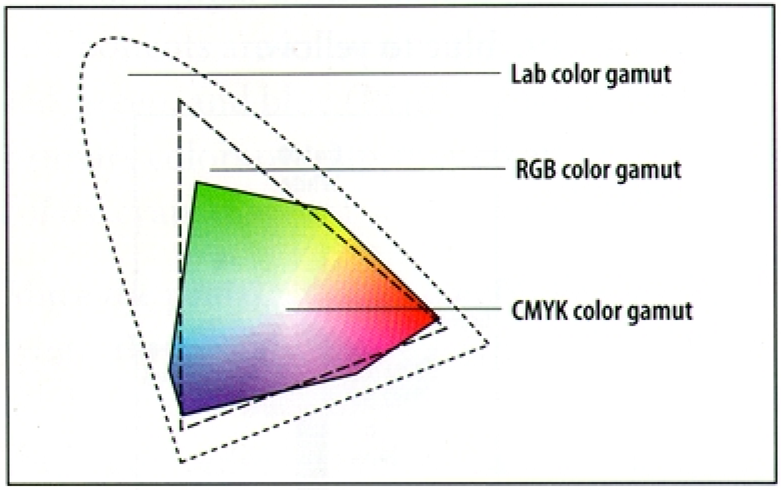
-- e.g., convert **White** from (1, 1, 1) in RGB to (0, 0, 0) in CMY.

$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} C \\ M \\ Y \end{bmatrix}$$

- Sometimes, an alternative CMYK model (K stands for *Black*) is used in color printing (e.g., to produce darker black than simply mixing CMY).
    - K := min (C, M, Y), C := C - K, M := M - K, Y := Y - K.

Comparison of Three Color Gamuts



- The *gamut* of colors is all colors that can be reproduced using the three primaries
- The Lab gamut covers all colors in visible spectrum
- The RGB gamut is smaller, hence certain visible colors (e.g. pure yellow, pure cyan) cannot be seen on monitors
- The CMYK gamut is the smallest (but not a straight subset of the RGB gamut)

4.3. Color Models in Video

- YIQ and YUV are the two commonly used color models in video

YUV Color Model

- Initially, for PAL analog video, it is now also used in CCIR 601 standard for digital video
- Y (luminance) is the CIE Y primary.

  $Y = 0.299R + 0.587G + 0.114B$

- *Chrominance* is defined as the difference between a color and a reference white at the same luminance. It can be represented by U and V -- the *color differences*.

  $U = B - Y$
  $V = R - Y$

  - If b/w image, then U = V = 0. --> No chrominance!
  - ** In actual PAL implementation:

    $U = 0.492 (B - Y)$
    $V = 0.877 (R - Y)$

- Sample YUV Decomposition:

Y **U** **V**

- Eye is most sensitive to Y. In PAL, 5.5 MHz is allocated to Y, 1.8 MHz each to U and V.

YCbCr Color Model

- The YCbCr model is closely related to the YUV, it is a scaled and shifted YUV.

  $Cb = (B - Y) / 1.772 + 0.5$
  $Cr = (R - Y) / 1.402 + 0.5$

- The chrominance values in YCbCr are always in the range of 0 to 1.

- YCbCr is used in JPEG and MPEG.

YIQ Color Model

- YIQ is used in NTSC color TV broadcasting, it is downward compatible with B/W TV where only Y is used.

- Although U and V nicely define the color differences, they do not align with the desired human perceptual color sensitivities. In NTSC, I and Q are used instead.

  I is the orange-blue axis, Q is the purple-green axis.
  I and Q axes are scaled and rotated R - Y and B - Y (by 33 degrees clockwise).

  $I = 0.877(R - Y) \cos 33 - 0.492(B - Y) \sin 33$

$Q = 0.877(R - Y) \sin 33 + 0.492(B - Y) \cos 33$

Namely,

$I = 0.736(R - Y) - 0.268(B - Y) = 0.596R - 0.275G - 0.321B$
$Q = 0.478(R - Y) + 0.413(B - Y) = 0.212R - 0.523G + 0.311B$

- The YIQ transform:

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.275 & -0.321 \\ 0.212 & -0.523 & 0.311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

- Eye is most sensitive to Y, next to I, next to Q.
  In NTSC broadcast TV, 4.2 MHz is allocated to Y, 1.5 MHz to I and 0.55 MHz to Q. For VCR, Y is cut down to 3.2 MHz and I to 0.63 MHz.

Summary

- Color images are encoded as triplets of values.
- RGB is an additive color model that is used for light-emitting devices, e.g., CRT displays
  CMY is a subtractive model that is used often for printers
- Two common color models in imaging are RGB and CMY, two common color models in video are YUV and YIQ.
- YUV uses properties of the human eye to prioritize information. Y is the black and white (luminance) image, U and V are the color difference (chrominance) images. YIQ uses similar idea.
- Besides the hardware-oriented color models (i.e., RGB, CMY, YUV, YIQ), HSB (Hue, Saturation, and Brightness) and HLS (Hue, Lightness, and Saturation) are also commonly used.

2D rotation : $\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$

The relationship between (U V) and (I Q):

$$V' = 0.877\,(R - Y)$$

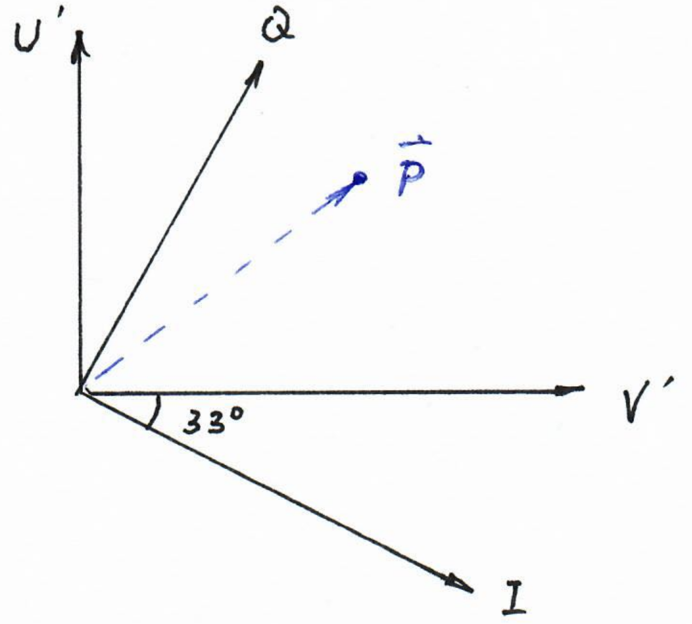
$$U' = 0.492\,(B - Y)$$

$$\begin{pmatrix} I \\ Q \end{pmatrix} = \begin{pmatrix} \cos 33^\circ & -\sin 33^\circ \\ \sin 33^\circ & \cos 33^\circ \end{pmatrix} \begin{pmatrix} V' \\ U' \end{pmatrix}$$

$\Downarrow$

$$I = V' \cos 33^\circ - U' \sin 33^\circ$$

$$Q = V' \sin 33^\circ + U' \cos 33^\circ$$



$$Y \approx 0.3R + 0.6G + 0.1B$$

$$U = B - Y$$
$$\approx -0.3R - 0.6G + 0.9B$$

$$V = R - Y$$
$$\approx 0.7R - 0.6G - 0.1B$$

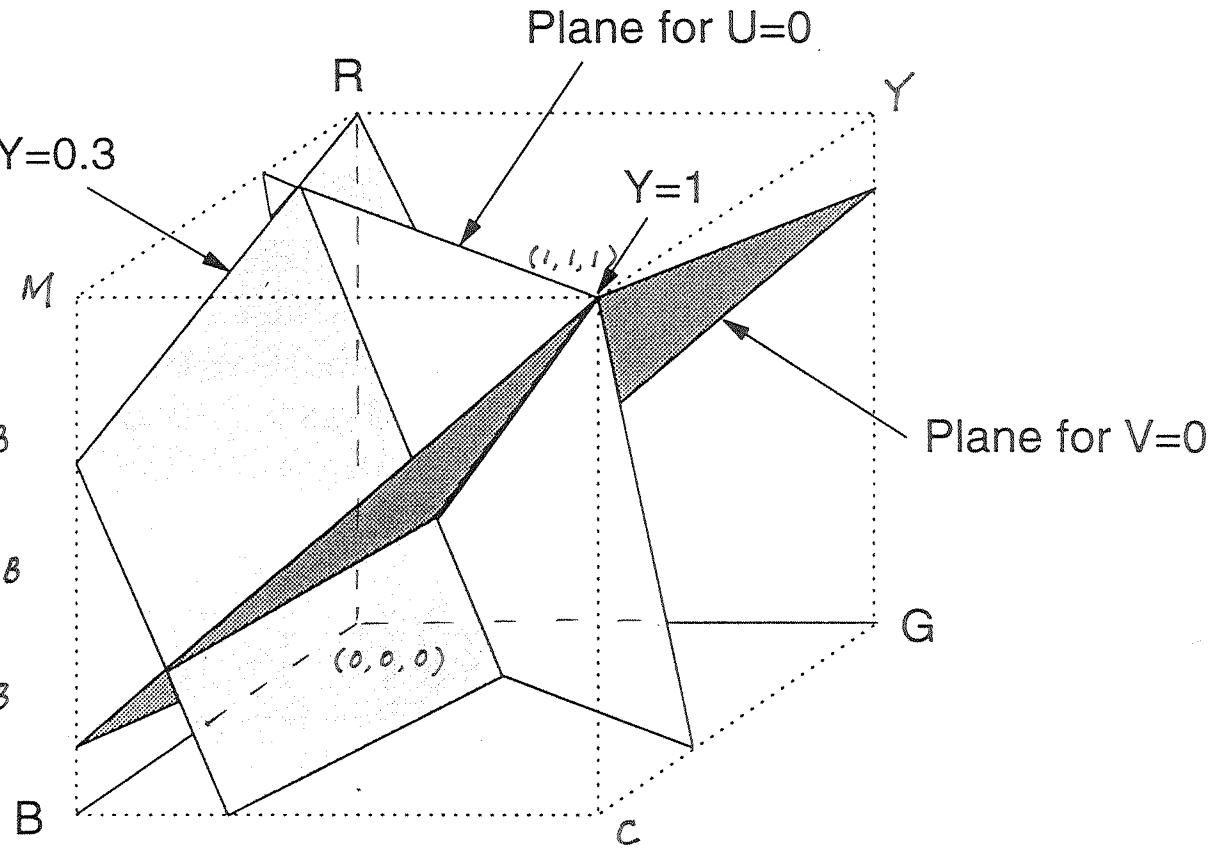


Figure 2-5. Relationship between the RGB and YUV coordinate systems

RGB-YUV Conversion

The basic definition of the *color differences* U, V is as below:

$$\begin{aligned} U &= B - Y \\ V &= R - Y \end{aligned} \tag{1}$$

Therefore,

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.299 & -0.587 & 0.886 \\ 0.701 & -0.587 & -0.114 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{2}$$

One goes backwards, from (Y, U, V) to (R, G, B), by inverting the matrix in Eq. (2).

It real implementations, however, the following is used:

$$\begin{aligned} U &\approx 0.492(B - Y) \\ V &\approx 0.877(R - Y) \end{aligned} \tag{3}$$

As a result, the following matrices should be used in converting RGB to YUV, and in converting YUV back to RGB, e.g., in the JPEG codec steps.

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.14713 & -0.28886 & 0.436 \\ 0.615 & -0.51499 & -0.10001 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{4}$$

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1.13983 \\ 1 & -0.39465 & -0.58060 \\ 1 & 2.03211 & 0 \end{bmatrix} \begin{bmatrix} Y \\ U \\ V \end{bmatrix} \tag{5}$$