Chapter 4
Color in Image and Video

4.1 Color Science

4.2 Color Models in Images

4.3 Color Models in Video

4.4 Further Exploration

4.1 Color Science

Light and Spectra

- Light is an electromagnetic wave. Its color is characterized by the wavelength content of the light.

  (a) Laser light consists of a single wavelength: e.g., a ruby laser produces a bright, scarlet-red beam.

  (b) Most light sources produce contributions over many wavelengths.

  (c) However, humans cannot detect all light, just contributions that fall in the "visible wavelengths".

  (d) Short wavelengths produce a blue sensation, long wavelengths produce a red one.

- **Spectrophotometer**: device used to measure visible light, by reflecting light from a diffraction grating (a ruled surface) that spreads out the different wavelengths.

- Figure 4.1 shows the phenomenon that white light contains all the colors of a rainbow.

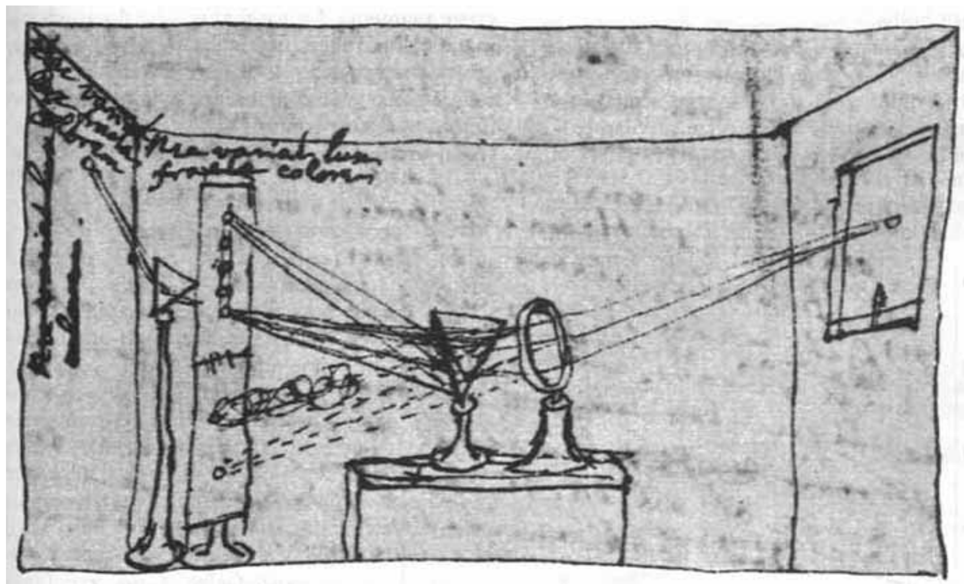

Fig. 4.1: Sir Isaac Newton's experiments.

- Visible light is an electromagnetic wave in the range 400 nm to 700 nm (where nm stands for nanometer, 10^{-9} meters).

- Fig. 4.2 shows the relative power in each wavelength interval for typical outdoor light on a sunny day. This type of curve is called a Spectral Power Distribution (**SPD**) or a **spectrum**.

- The symbol for wavelength is λ. This curve is called $E(\lambda)$.

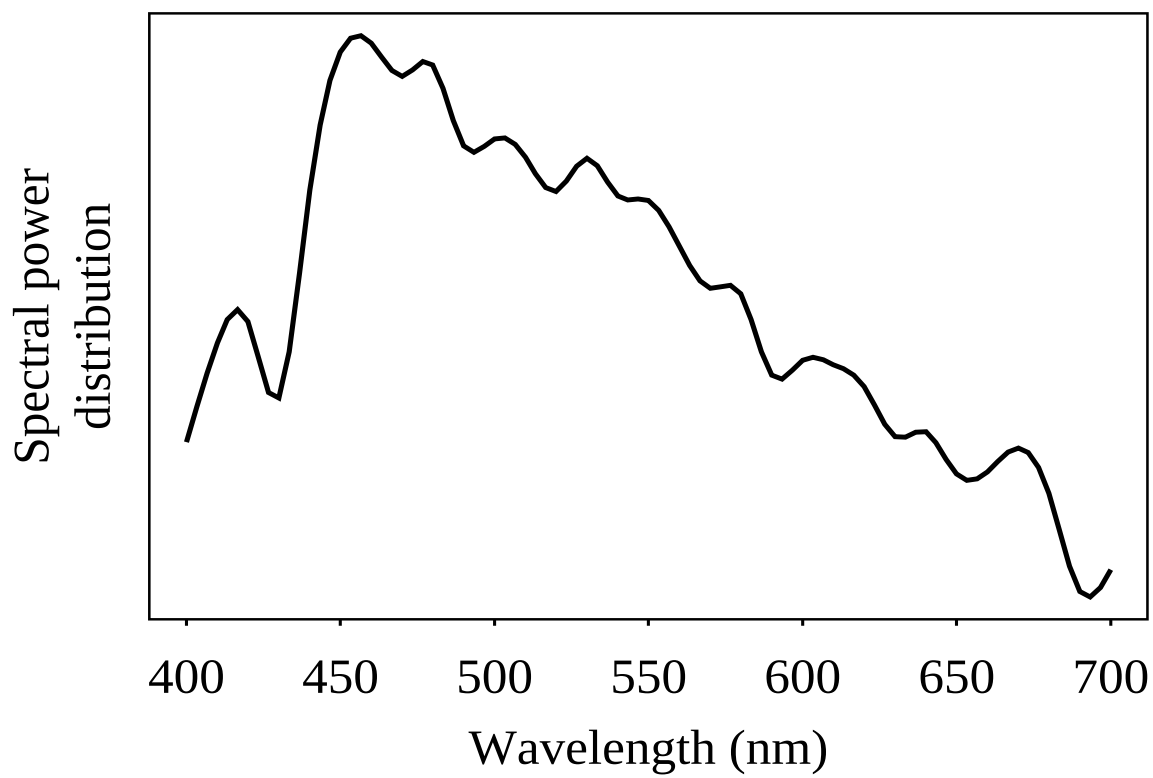

Fig. 4.2: Spectral power distribution of daylight.

Human Vision

- The eye works like a camera, with the lens focusing an image onto the retina (upside-down and left-right reversed).
- The retina consists of an array of *rods* and three kinds of cones.
- The rods come into play when light levels are low and produce a image in shades of gray ("all cats are gray at night!").
- For higher light levels, the cones each produce a signal. Because of their differing pigments, the three kinds of cones are most sensitive to red (R), green (G), and blue (B) light.
- It seems likely that the brain makes use of *differences* R-G, G-B, and B-R, as well as combining all of R, G, and B into a high-light-level achromatic channel.

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 (Fig. 4.3 below).
- The Blue receptor sensitivity is not shown to scale because it is much smaller than the curves for Red or Green — Blue is a late addition, in evolution.
  - Statistically, Blue is the favorite color of humans, regardless of nationality — perhaps for this reason: Blue is a latecomer and thus is a bit surprising!
- Fig. 4.3 shows the overall sensitivity as a dashed line — this important curve is called the luminous-efficiency function.
  - It is usually denoted $V(\lambda)$ and is formed as the sum of the response curves for Red, Green, and Blue.

- The rod sensitivity curve looks like the luminous-efficiency function $V(\lambda)$ but is shifted to the red end of the spectrum.

- The achromatic channel produced by the cones is approximately proportional to $2R + G + B/20$.

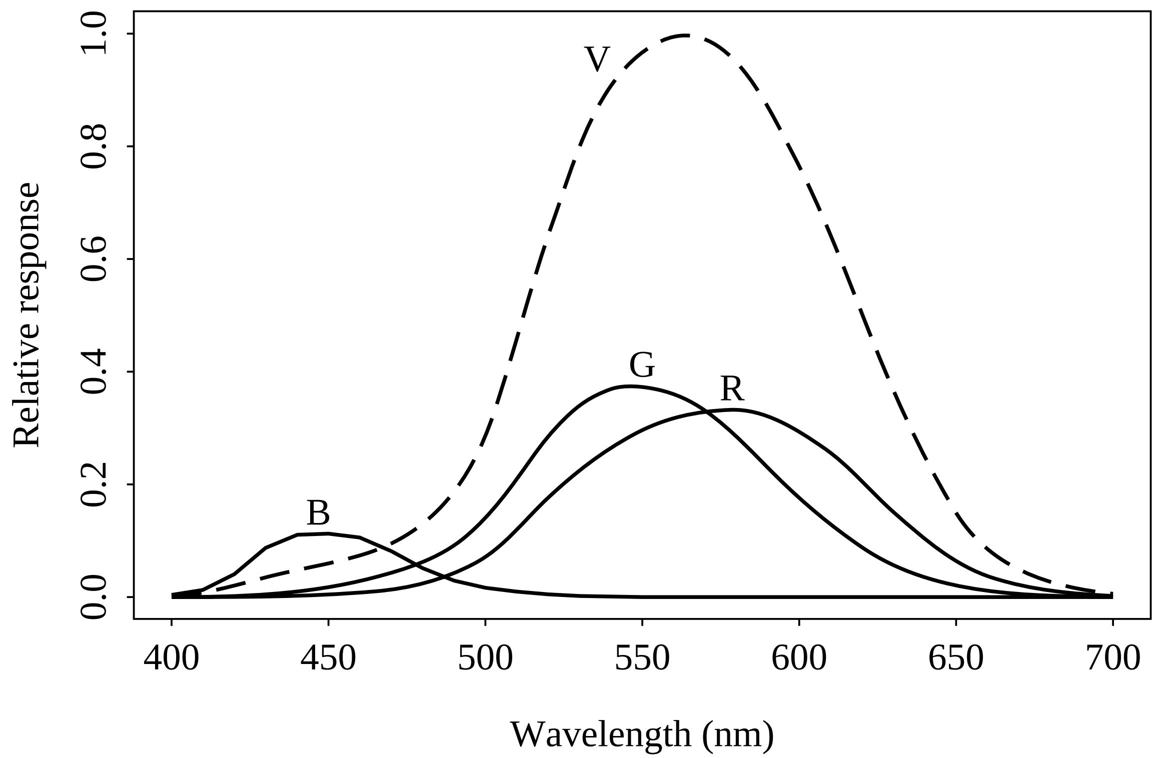

Fig. 4.3: R,G, and B cones, and Luminous Efficiency curve V(λ).

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

$$\boldsymbol{q}(\lambda) = (q_R(\lambda), q_G(\lambda), q_B(\lambda))^T \tag{4.1}$$

- 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. 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.

- We can succinctly write down this idea in the form of an integral:

$$
\begin{aligned}
R &= \int E(\lambda)\, q_R(\lambda)\, d\lambda \\
G &= \int E(\lambda)\, q_G(\lambda)\, d\lambda \\
B &= \int E(\lambda)\, q_B(\lambda)\, d\lambda
\end{aligned}
\tag{4.2}
$$

Image Formation

- Surfaces reflect different amounts of light at different wavelengths, and dark surfaces reflect less energy than light surfaces.

- Fig. 4.4 shows the surface spectral reflectance from (1) orange sneakers and (2) faded bluejeans. The reflectance function is denoted $S(\lambda)$.

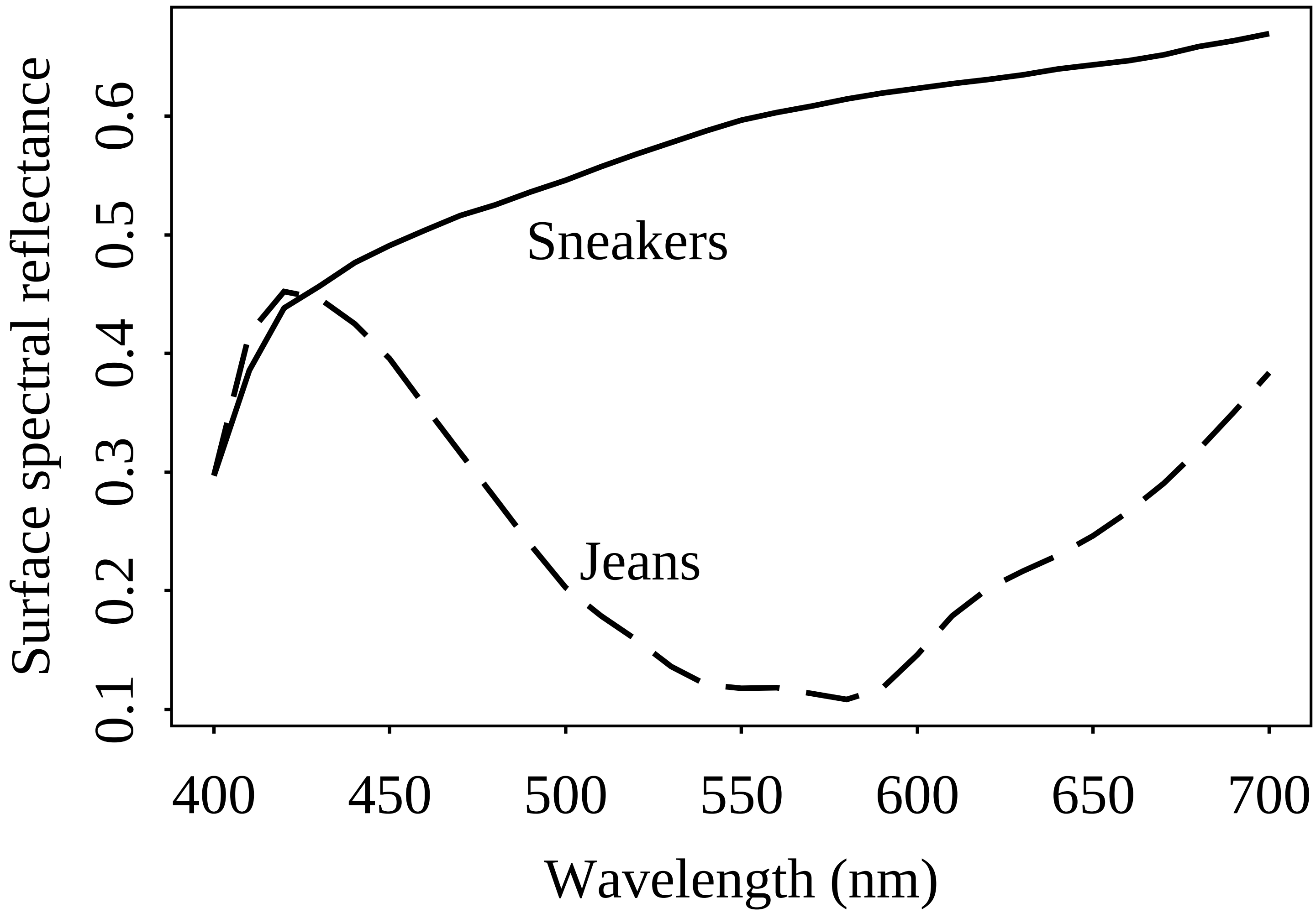

Fig. 4.4: Surface spectral reflectance functions $S(\lambda)$ for objects.

- Image formation is thus:
  - Light from the illuminant with SPD $E(\lambda)$ impinges on a surface, with surface spectral reflectance function $S(\lambda)$, is reflected, and then is filtered by the eye's cone functions $\boldsymbol{q}\,(\lambda)$.
  - Reflection is shown in Fig. 4.5 below.
  - The function $C(\lambda)$ is called the *color signal* and consists of the product of $E(\lambda)$, the illuminant, times $S(\lambda)$, the reflectance:

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

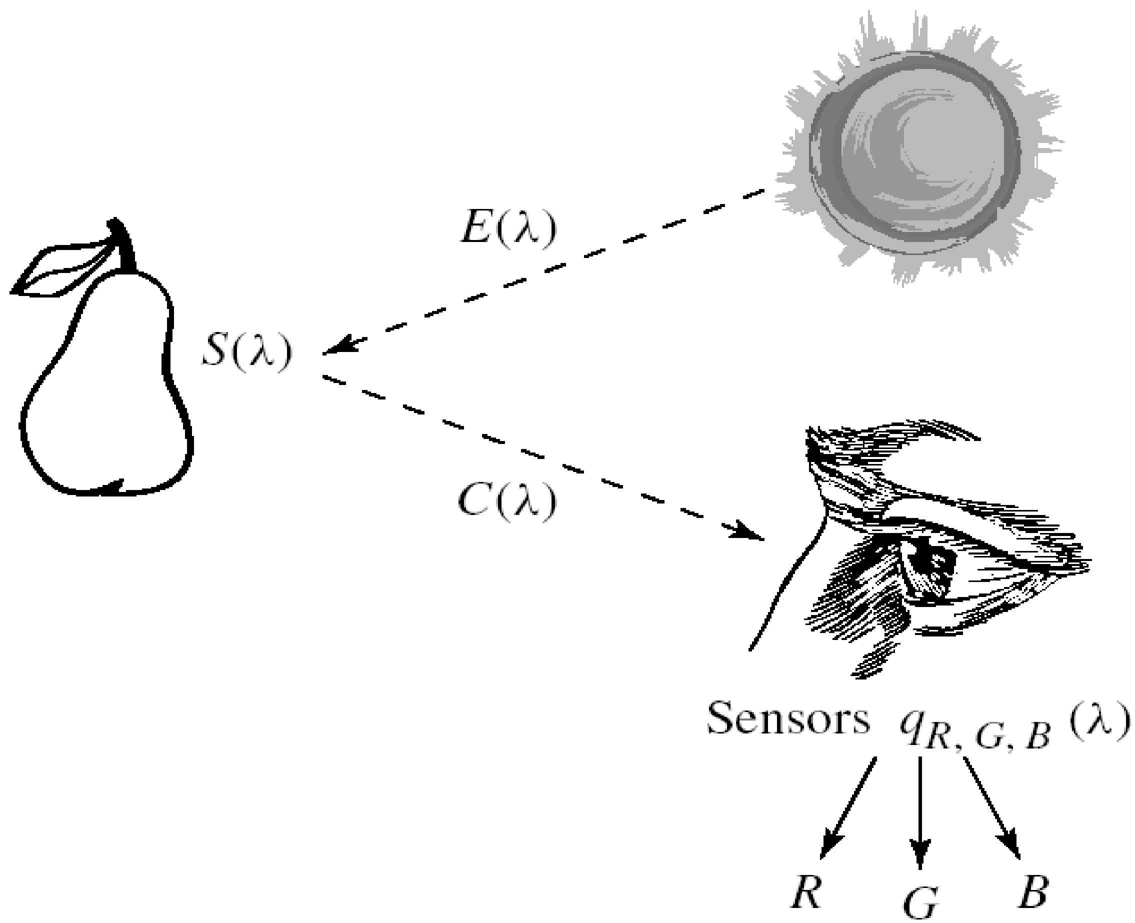

Fig. 4.5: Image formation model.

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

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

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

$$B = \int E(\lambda)\, S(\lambda)\, q_B(\lambda)\, d\lambda \qquad (4.3)$$

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.

  (a) 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.

  (b) It is customary to append a prime to signals that are **gamma-corrected** by raising to the power $(1/\gamma)$ before transmission. Thus we arrive at **linear signals**:

$$R \rightarrow R' = R^{1/\gamma} \Rightarrow (R')^\gamma \rightarrow R \tag{4.4}$$

- Fig. 4.6(a) shows light output with no gamma-correction applied. We see that darker values are displayed too dark. This is also shown in Fig. 4.7(a), which displays a linear ramp from left to right.

- Fig. 4.6(b) shows the effect of pre-correcting signals by applying the power law $R^{1/\gamma}$; it is customary to normalize voltage to the range [0,1].

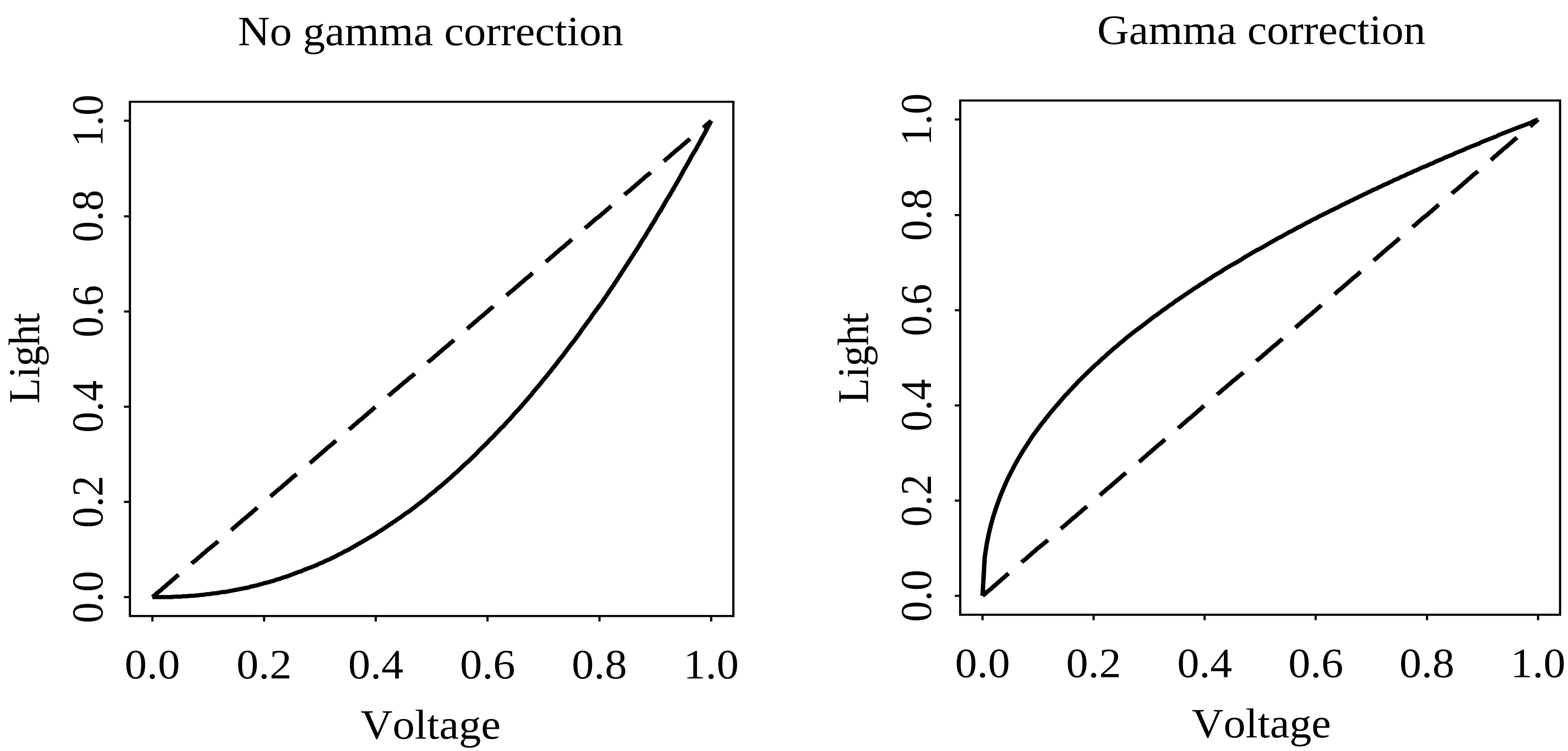

Fig. 4.6: (a): Effect of CRT on light emitted from screen (voltage is normalized to range 0..1). (b): Gamma correction of signal.

- The combined effect is shown in Fig. 4.7(b). Here, a ramp is shown in 16 steps from gray-level 0 to gray-level 255.

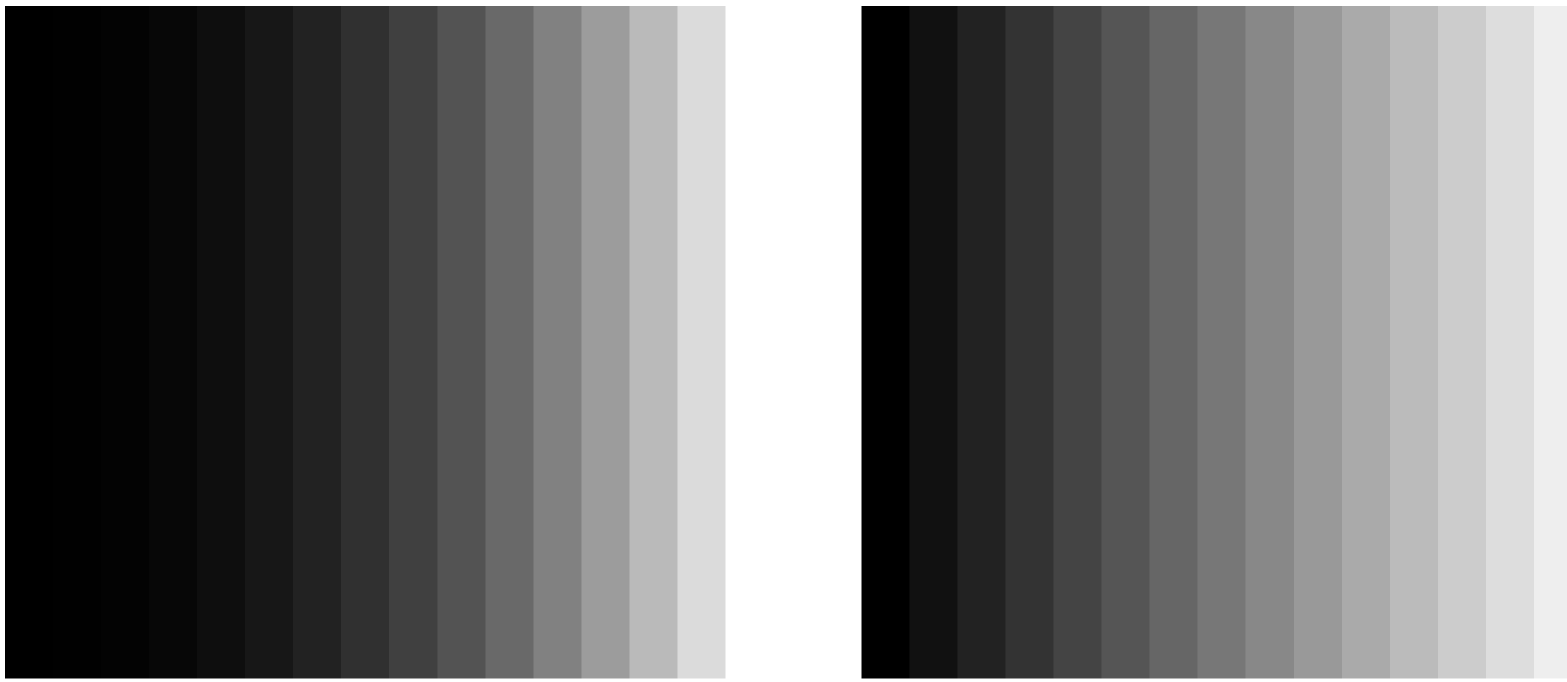

Fig. 4.7: (a): Display of ramp from 0 to 255, with no gamma correction. (b): Image with gamma correction applied

- 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:

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

Color-Matching Functions

- Even without knowing the eye-sensitivity curves of Fig.4.3, a technique evolved in psychology for matching a combination of basic R, G, and B lights to a given shade.
- The particular set of three basic lights used in an experiment are called the set of **color primaries**.
- To match a given color, a subject is asked to separately adjust the brightness of the three primaries using a set of controls until the resulting spot of light most closely matches the desired color.
- The basic situation is shown in Fig.4.8. A device for carrying out such an experiment is called a **colorimeter**.

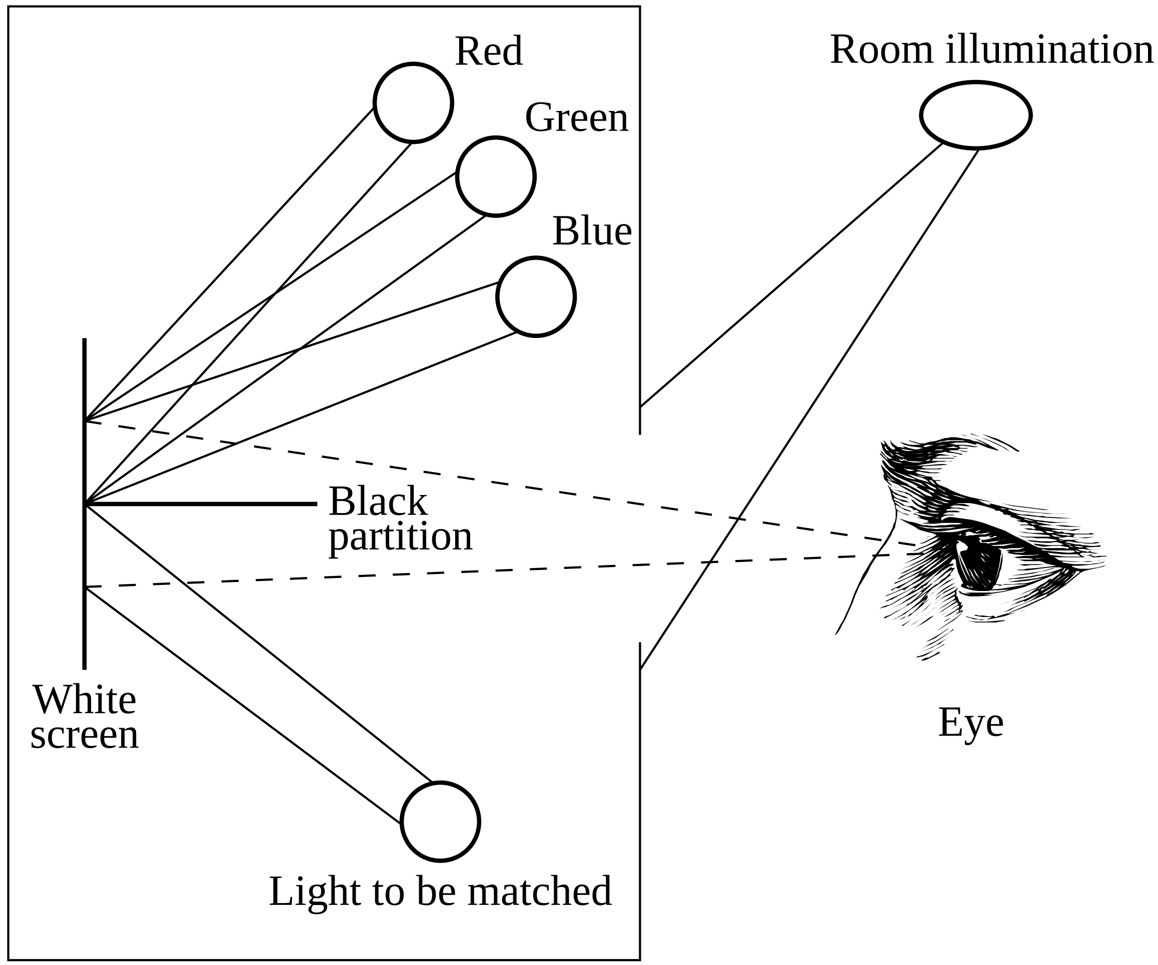

Fig. 4.8: Colorimeter experiment.

- The amounts of R, G, and B the subject selects to match each single-wavelength light forms the *color-matching curves*. These are denoted $\bar{r}(\lambda)$, $\bar{g}(\lambda)$, $\bar{b}(\lambda)$ and are shown in Fig. 4.9.

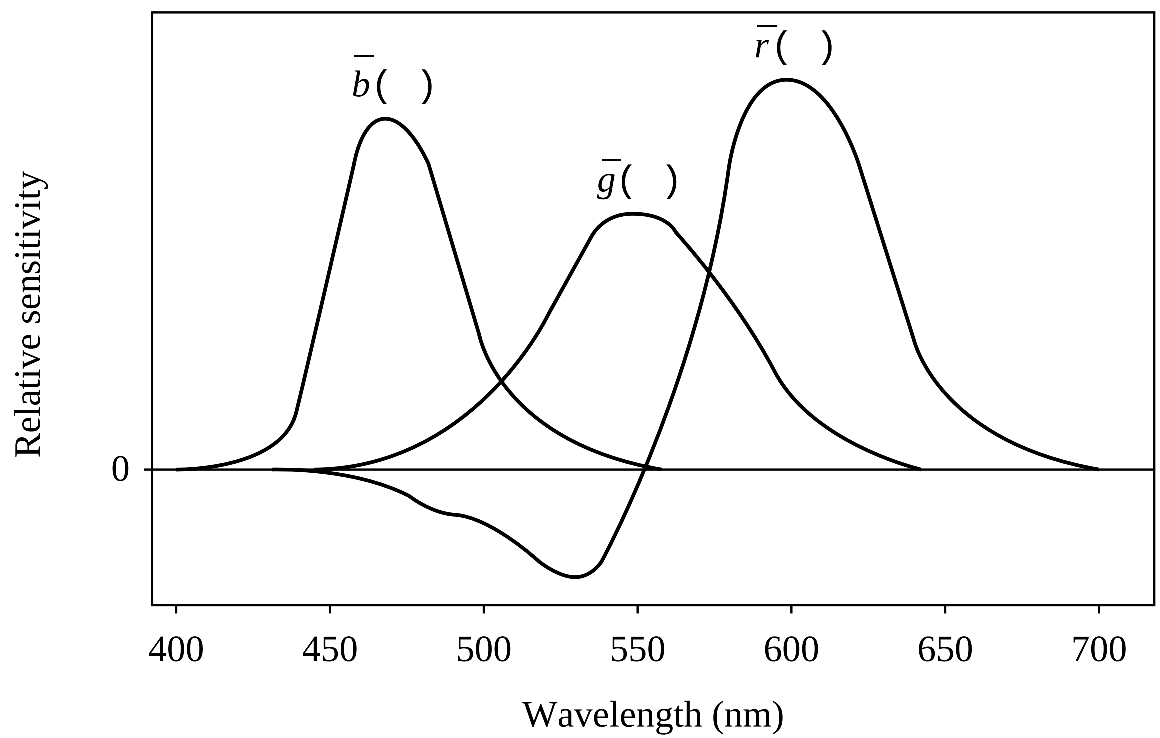

Fig. 4.9: CIE RGB color-matching functions $\bar{r}(\lambda), \bar{g}(\lambda), \bar{b}(\lambda)$.

CIE Chromaticity Diagram

- Since the $\bar{r}(\lambda)$ color-matching curve has a negative lobe, a set of fictitious primaries were devised that lead to color-matching functions with only positives values.

  (a) The resulting curves are shown in Fig. 4.10; these are usually referred to as *the* color-matching functions.

  (b) They are a 3×3 matrix away from $\bar{r}, \bar{g}, \bar{b}$ curves, and are denoted $\bar{x}(\lambda), \bar{y}(\lambda), \bar{z}(\lambda)$.

  (c) The matrix is chosen such that the middle standard color-matching function $\bar{y}(\lambda)$ exactly equals the luminous-efficiency curve $V(\lambda)$ shown in Fig. 4.3.

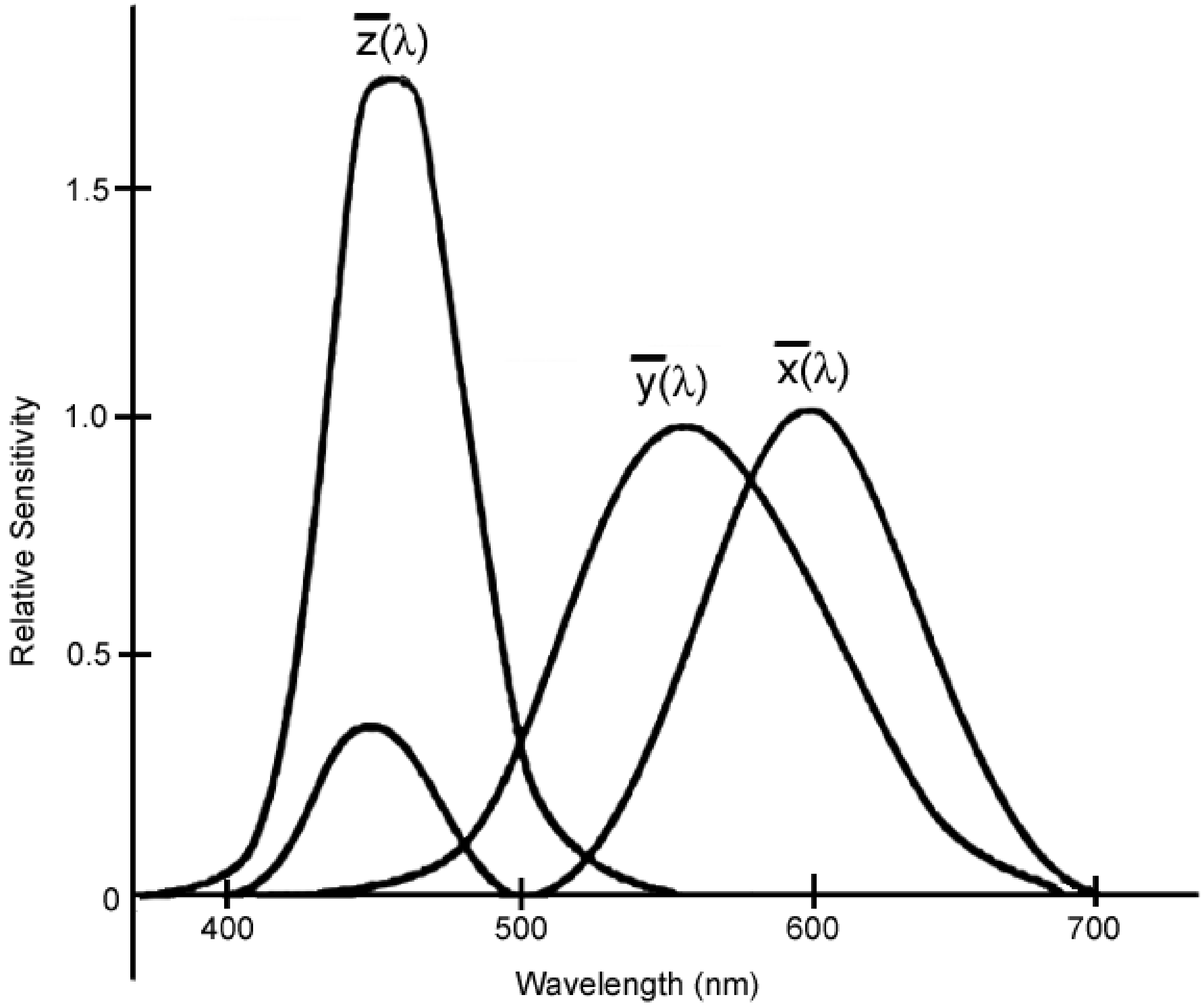

Fig. 4.10: CIE standard XYZ color-matching functions $\bar{x}(\lambda), \bar{y}(\lambda), \bar{z}(\lambda)$.

- For a general SPD $E(\lambda)$, the essential "colorimetric" information required to characterize a color is the set of *tristimulus values* X, Y, Z defined in analogy to (Eq. 4.2) as (Y == **luminance**):

$$
\begin{aligned}
X &= \int E(\lambda)\, \bar{x}(\lambda)\, d\lambda \\
Y &= \int E(\lambda)\, \bar{y}(\lambda)\, d\lambda \\
Z &= \int E(\lambda)\, \bar{z}(\lambda)\, d\lambda
\end{aligned}
\tag{4.6}
$$

- 3D data is difficult to visualize, so the CIE devised a 2D diagram based on the values of (X, Y, Z) triples implied by the curves in Fig. 4.10.

- We go to 2D by factoring out the magnitude of vectors (X, Y, Z); we could divide by $\sqrt{X^2 + Y^2 + Z^2}$, but instead we divide by the sum $X + Y + Z$ to make the **chromaticity**:

$$\begin{aligned} x &= X/(X+Y+Z) \\ y &= Y/(X+Y+Z) \\ z &= Z/(X+Y+Z) \end{aligned} \tag{4.7}$$

- This effectively means that one value out of the set (x, y, z) is redundant since we have

$$x + y + z = \frac{X+Y+Z}{X+Y+Z} \equiv 1 \tag{4.8}$$

so that

$$z = 1 - x - y \tag{4.9}$$

- Effectively, we are projecting each tristimulus vector (X, Y, Z) onto the plane connecting points $(1, 0, 0)$, $(0, 1, 0)$, and $(0, 0, 1)$.
- Fig. 4.11 shows the locus of points for monochromatic light

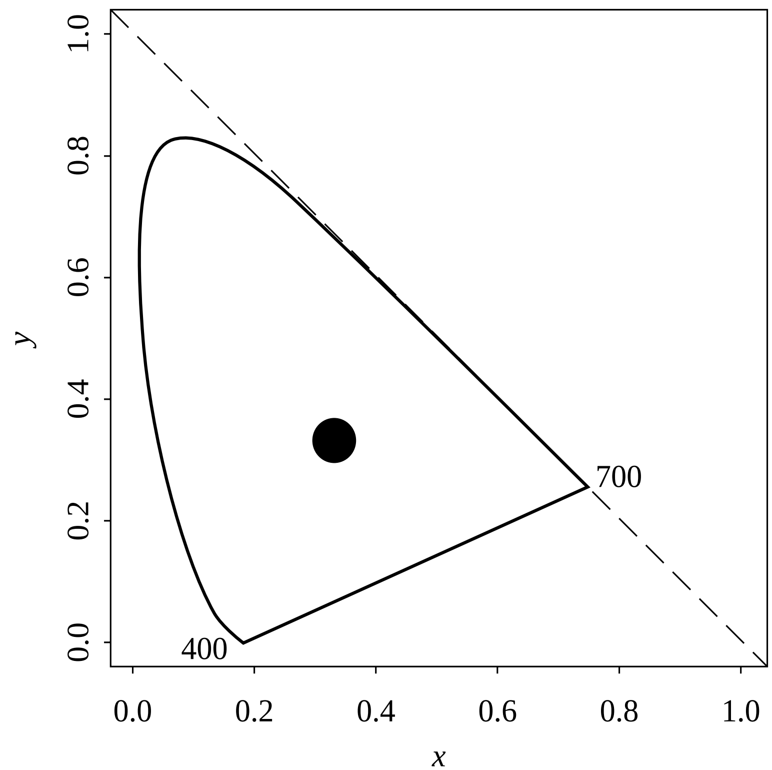

Fig. 4.11: CIE chromaticity diagram.

(a) The color matching curves each add up to the same value — the area under each curve is the same for each of $\bar{x}(\lambda), \bar{y}(\lambda), \bar{z}(\lambda)$.

(b) For an $E(\lambda) = 1$ for all λ, — an "equi-energy white light" — chromaticity values are $(1/3, 1/3)$. Fig. 4.11 displays a typical actual white point in the middle of the diagram.

(c) Since $x, y \leq 1$ and $x + y \leq 1$, all possible chromaticity values lie below the dashed diagonal line in Fig. 4.11.

- The CIE defines several "white" spectra: illuminant A, illuminant C, and standard daylights D65 and D100. (Fig. 4.12)

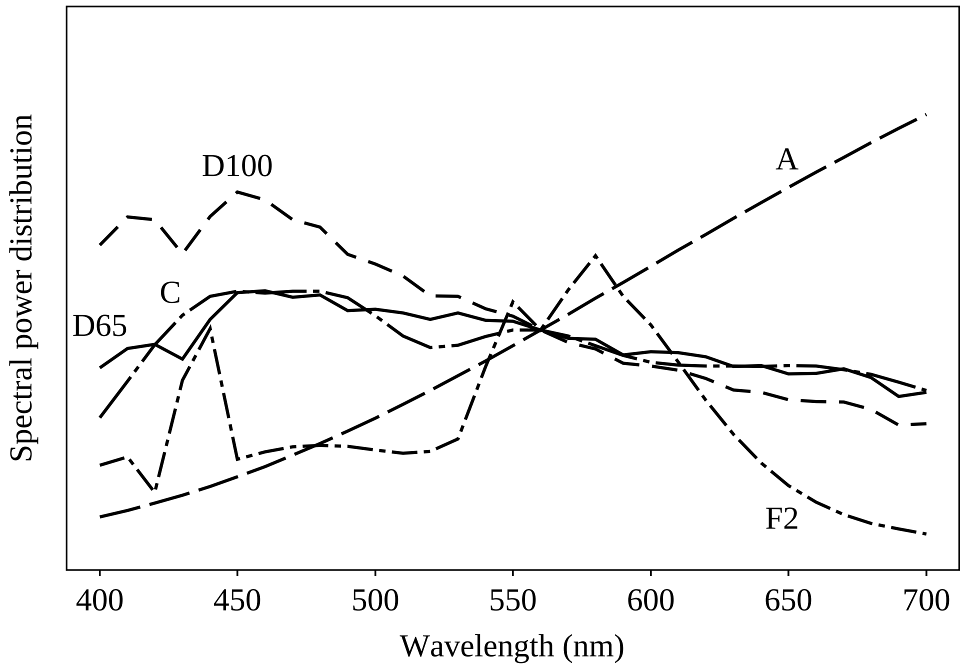

Fig. 4.12: Standard illuminant SPDs.

- Chromaticities on the spectrum locus (the "horseshoe" in Fig. 4.11) represent "pure" colors. These are the most "saturated". Colors close to the white point are more unsaturated.

- The chromaticity diagram: for a mixture of two lights, the resulting chromaticity lies on the straight line joining the chromaticities of the two lights.

- The "dominant wavelength" is the position on the spectrum locus intersected by a line joining the white point to the given color, and extended through it.

Color Monitor Specifications

- Color monitors are specified in part by the white point chromaticity that is desired if the RGB electron guns are all activated at their highest value (1.0, if we normalize to [0,1]).

- We want the monitor to display a specified white when when $R'=G'=B'=1$.

- There are several monitor specifications in current use (Table 4.1).

Table 4.1: Chromaticities and White Points of Monitor Specifications

System	Red		Green		Blue		White Point	
	x_r	y_r	x_g	y_g	x_b	y_b	x_W	y_W
NTSC	0.67	0.33	0.21	0.71	0.14	0.08	0.3101	0.3162
SMPTE	0.630	0.340	0.310	0.595	0.155	0.070	0.3127	0.3291
EBU	0.64	0.33	0.29	0.60	0.15	0.06	0.3127	0.3291

Out-of-Gamut Colors

- For any (x, y) pair we wish to find that RGB triple giving the specified (x, y, z): We form the z values for the phosphors, via $z = 1 - x - y$ and solve for RGB from the phosphor chromaticities.

- We combine nonzero values of R, G, and B via

$$\begin{bmatrix} x_r & x_g & x_b \\ y_r & y_g & y_b \\ z_r & z_g & z_b \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \tag{4.10}$$

- If (x, y) [color without magnitude] is *specified*, instead of derived as above, we have to invert the matrix of phosphor (x, y, z) values to obtain RGB.
- What do we do if any of the RGB numbers is *negative*? — that color, visible to humans, is out-of-gamut for our display.

1. One method: simply use the closest in-gamut color available, as in Fig. 4.13.

2. Another approach: select the closest complementary color.

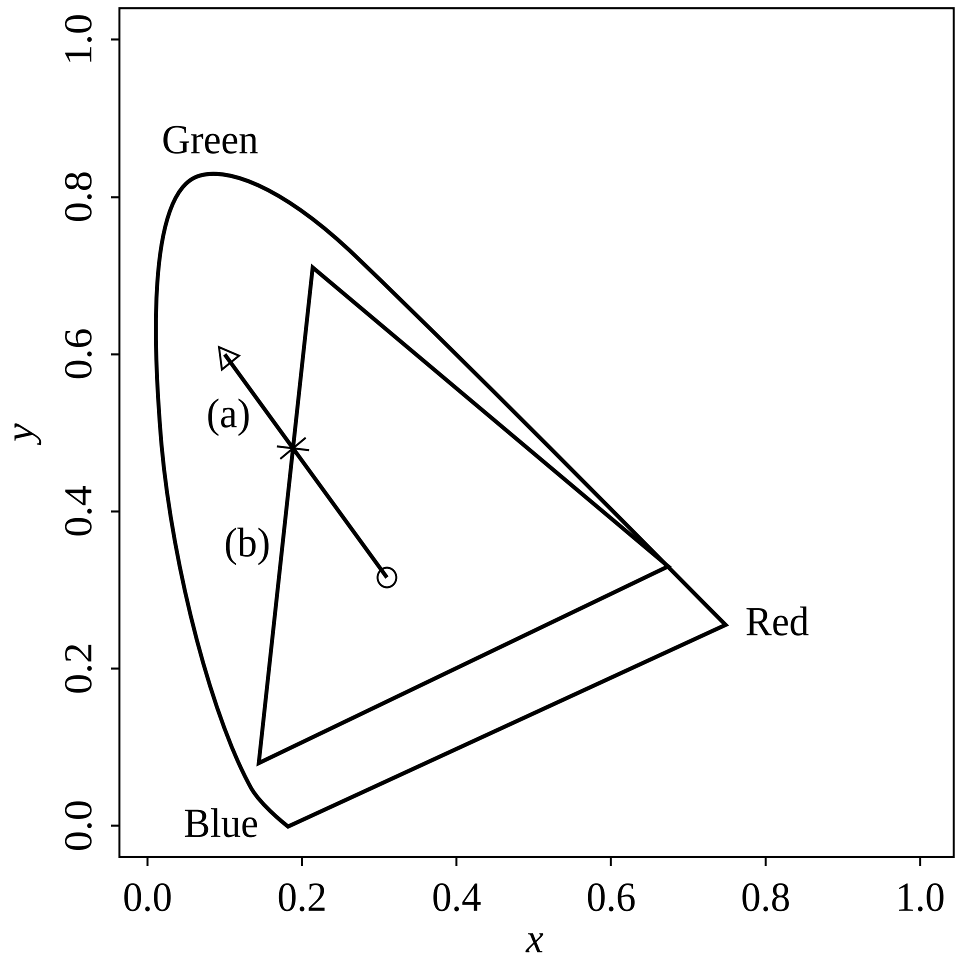

Fig. 4.13: Approximating an out-of-gamut color by an in-gamut one. The out-of-gamut color shown by a triangle is approximated by the intersection of (a) the line from that color to the white point with (b) the boundary of the device color gamut.

- **Grassman's Law**: (Additive) color matching is linear. This means that if we match $color1$ with a linear combinations of lights and match $color2$ with another set of weights, the combined color $color1 + color2$ is matched by the sum of the two sets of weights.

- Additive color results from self-luminous sources, such as lights projected on a white screen, or the phosphors glowing on the monitor glass. (Subtractive color applies for printers, and is very different).

- Fig. 4.13 above shows the triangular gamut for the NTSC system, drawn on the CIE diagram — a monitor can display only the colors inside a triangular gamut.

White Point Correction

- **Problems**:

  (a) One deficiency in what we have done so far is that we need to be able to map tristimulus values XYZ to device RGBs including magnitude, and not just deal with chromaticity xyz.

  (b) Table 4.1 would produce incorrect values:

  - E.g., consider the SMPTE specifications. Setting $R = G = B = 1$ results in a value of X that equals the sum of the x values, or $0.630 + 0.310 + 0.155$, which is 1.095. Similarly the Y and Z values come out to 1.005 and 0.9. Now, dividing by $(X + Y + Z)$ this results in a chromaticity of $(0.365, 0.335)$, rather than the desired values of $(0.3127, 0.3291)$.

- To correct both problems, first take the white point magnitude of Y as unity:

$$Y\,(\text{white point}) \;=\; 1 \tag{4.11}$$

- Now we need to find a set of three correction factors such that if the gains of the three electron guns are multiplied by these values we get exactly the white point XYZ value at $R = G = B = 1$.

- Suppose the matrix of phosphor chromaticities $x_r, x_g, ...$ etc. in Eq. (4.10) is called $\boldsymbol{M}$. We can express the correction as a diagonal matrix $\boldsymbol{D} = diag(d_1, d_2, d_3)$ such that

$$XYZ_{\text{white}} \equiv \boldsymbol{M} \ \boldsymbol{D} \ (1,1,1)^T \tag{4.12}$$

- For the SMPTE specification, we have $(x, y, z) = (0.3127, 0.3291, 0.3582)$ or, dividing by the middle value — $XYZ_{\text{white}} = (0.95045,\ 1\ ,\ 1.08892)$. We note that multiplying $\boldsymbol{D}$ by $(1,1,1)^T$ just gives $(d_1, d_2, d_3)^T$ so we end up with an equation *specifying* $(d_1, d_2, d_3)^T$:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_{\text{white}} = \begin{bmatrix} 0.630 & 0.310 & 0.155 \\ 0.340 & 0.595 & 0.070 \\ 0.03 & 0.095 & 0.775 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix} \tag{4.13}$$

- Inverting, with the new values XYZ_{white} specified as above, we arrive at

$$(d_1, d_2, d_3) = (0.6247, 1.1783, 1.2364) \tag{4.14}$$

These are large correction factors.

XYZ to RGB Transform

- Now the 3×3 transform matrix from XYZ to RGB is taken to be

$$\boldsymbol{T} = \boldsymbol{M}\,\boldsymbol{D} \tag{4.15}$$

even for points other than the white point:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \boldsymbol{T} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{4.16}$$

- For the SMPTE specification, we arrive at:

$$\boldsymbol{T} = \begin{bmatrix} 0.3935 & 0.3653 & 0.1916 \\ 0.2124 & 0.7011 & 0.0866 \\ 0.0187 & 0.1119 & 0.9582 \end{bmatrix} \tag{4.17}$$

- Written out, this reads:

$$\begin{aligned} X &= 0.3935 \cdot R + 0.3653 \cdot G + 0.1916 \cdot B \\ Y &= 0.2124 \cdot R + 0.7011 \cdot G + 0.0866 \cdot B \\ Z &= 0.0187 \cdot R + 0.1119 \cdot G + 0.9582 \cdot B \end{aligned} \tag{4.18}$$

Transform with Gamma Correction

- Instead of linear R, G, B we usually have nonlinear, gamma-corrected R', G', B' (produced by a camcorder or digital camera).

- To transform XYZ to RGB, calculate the linear RGB required, by inverting Eq. (4.16) above; then make nonlinear signals via gamma correction.

- Nevertheless this is not often done as stated. Instead, the equation for the Y value is used as is, but applied to nonlinear signals.
  (a) The only concession to accuracy is to give the new name Y' to this new Y value created from R', G', B'.
  (b) The significance of Y' is that it codes a descriptor of brightness for the pixel in question.

- Following the procedure outlined above, but with the values in Table 4.1 for NTSC, we arrive at the following transform:

$$\begin{aligned} X &= 0.607 \cdot R + 0.174 \cdot G + 0.200 \cdot B \\ Y &= 0.299 \cdot R + 0.587 \cdot G + 0.114 \cdot B \\ Z &= 0.000 \cdot R + 0.066 \cdot G + 1.116 \cdot B \end{aligned} \tag{4.19}$$

- Thus, coding for nonlinear signals begins with encoding the nonlinear-signal correlate of luminance:

$$Y' = 0.299 \cdot R' + 0.587 \cdot G' + 0.114 \cdot B' \tag{4.20}$$

L*a*b* (CIELAB) Color Model

- **Weber's Law**: Equally-perceived differences are proportional to magnitude. The more there is of a quantity, the more change there must be to perceive a difference.

- A rule of thumb for this phenomenon states that equally-perceived changes must be relative — changes are about equally perceived if the ratio of the change is the same, whether for dark or bright lights, etc.

- Mathematically, with intensity I, change is equally perceived so long as the change $\frac{\Delta I}{I}$ is a constant. If it's quiet, we can hear a small change in sound. If there is a lot of noise, to experience the same difference the change has to be of the same proportion.

- For human vision, the CIE arrived at a different version of this kind of rule — **CIELAB** space. What is being quantified in this space is **differences** perceived in color and brightness.

- Fig. 4.14 shows a cutaway into a 3D solid of the coordinate space associated with this color difference metric.

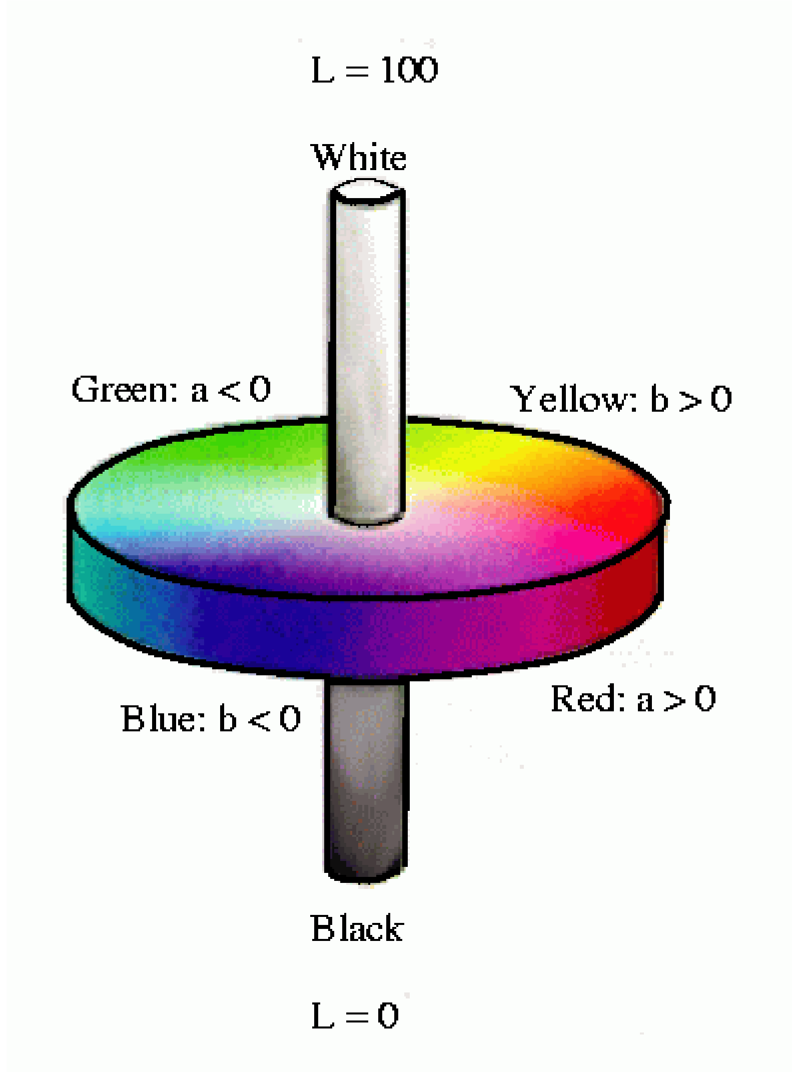

Fig. 4.14: CIELAB model.

- CIELAB:

$$\Delta E = \sqrt{(L*)^2 + (a*)^2 + (b*)^2} \tag{4.21}$$

$$L^* = 116 \left(\frac{Y}{Y_n}\right)^{(1/3)} - 16$$

$$a^* = 500 \left[\left(\frac{X}{X_n}\right)^{(1/3)} - \left(\frac{Y}{Y_n}\right)^{(1/3)}\right]$$

$$b^* = 200 \left[\left(\frac{Y}{Y_n}\right)^{(1/3)} - \left(\frac{Z}{Z_n}\right)^{(1/3)}\right] \tag{4.22}$$

with X_n, Y_n, Z_n the XYZ values of the white point. Auxiliary definitions are:

$$\text{chroma} = c^* = \sqrt{(a^*)^2 + (b^*)^2}\,, \text{ hue angle} = h^* = \arctan \frac{b^*}{a^*}$$

More Color Coordinate Schemes

- Beware: gamma correction or not is usually ignored.
- Schemes include:

  a) CMY — Cyan (C), Magenta (M) and Yellow (Y) color model;

  b) HSL — Hue, Saturation and Lightness;

  c) HSV — Hue, Saturation and Value;

  d) HSI — Hue, Saturation and Intensity;

  e) HCI — C=Chroma;

  f) HVC — V=Value;

  g) HSD — D=Darkness.

4.2 Color Models in Images

- Colors models and spaces used for stored, displayed, and printed images.

- **RGB Color Model for CRT Displays**

  1. We expect to be able to use 8 bits per color channel for color that is accurate enough.
  2. However, in fact we have to use about 12 bits per channel to avoid an aliasing effect in dark image areas — contour bands that result from gamma correction.
  3. For images produced from computer graphics, we store integers proportional to intensity in the frame buffer. So should have a gamma correction LUT between the frame buffer and the CRT.
  4. If gamma correction is applied to floats before quantizing to integers, before storage in the frame buffer, then in fact we can use only 8 bits per channel and still avoid contouring artifacts.

Subtractive Color: CMY Color Model

- So far, we have effectively been dealing only with **additive color**. Namely, when two light beams impinge on a target, their colors add; when two phosphors on a CRT screen are turned on, their colors add.

- But for ink deposited on paper, the opposite situation holds: yellow ink *subtracts* blue from white illumination, but reflects red and green; it appears yellow.

1. Instead of red, green, and blue primaries, we need primaries that amount to -red, -green, and -blue. I.e., we need to *subtract* R, or G, or B.

2. These subtractive color primaries are Cyan (C), Magenta (M) and Yellow (Y) inks.

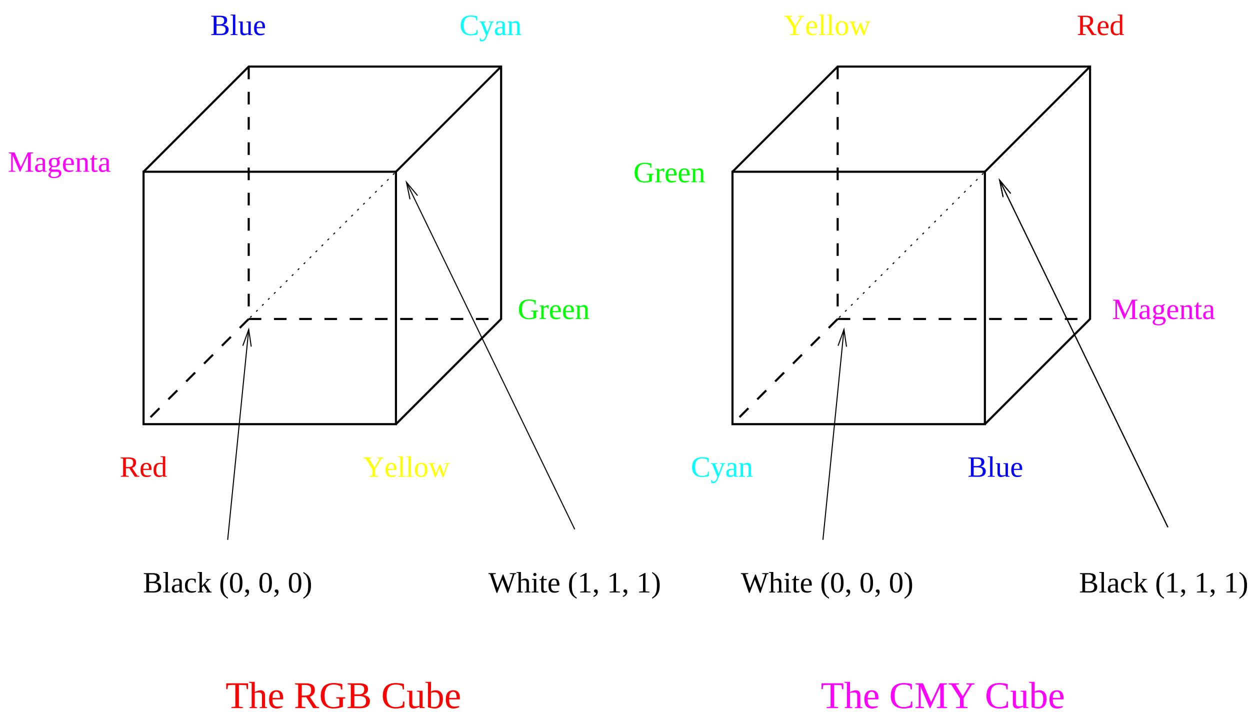

Fig. 4.15: RGB and CMY color cubes.

Transformation from RGB to CMY

- Simplest model we can invent to specify what ink density to lay down on paper, to make a certain desired RGB color:

$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{4.24}$$

  Then the inverse transform is:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} C \\ M \\ Y \end{bmatrix} \tag{4.25}$$

Undercolor Removal: CMYK System

- **Undercolor removal**:. Sharper and cheaper printer colors: calculate that part of the CMY mix that would be black, remove it from the color proportions, and add it back as real black.

- The new specification of inks is thus:

$$K \equiv min\{C, M, Y\} \tag{4.26}$$

$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} \Rightarrow \begin{bmatrix} C-K \\ M-K \\ Y-K \end{bmatrix}$$

- Fig. 4.16: color combinations that result from combining primary colors available in the two situations, additive color and subtractive color.

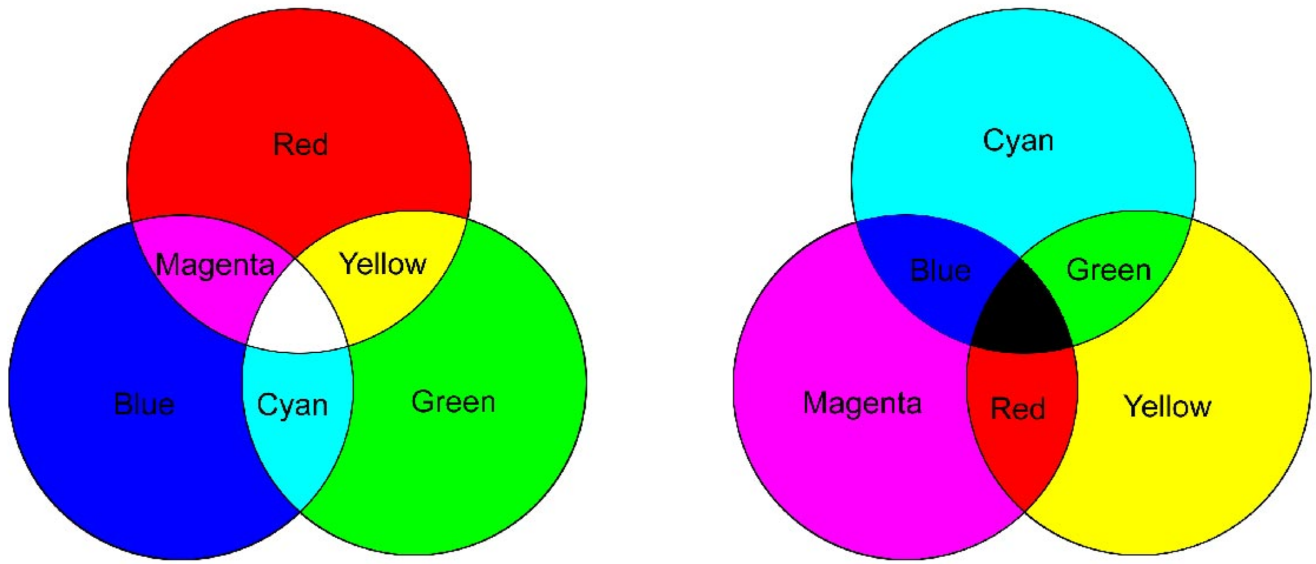

Fig. 4.16: Additive and subtractive color. (a): RGB is used to specify additive color. (b): CMY is used to specify subtractive color

Printer Gamuts

- Actual transmission curves overlap for the C, M, Y inks. This leads to "crosstalk" between the color channels and difficulties in predicting colors achievable in printing.

- Fig. 4.17(a) shows typical transmission curves for real "block dyes", and Fig.4.17(b) shows the resulting color gamut for a color printer.

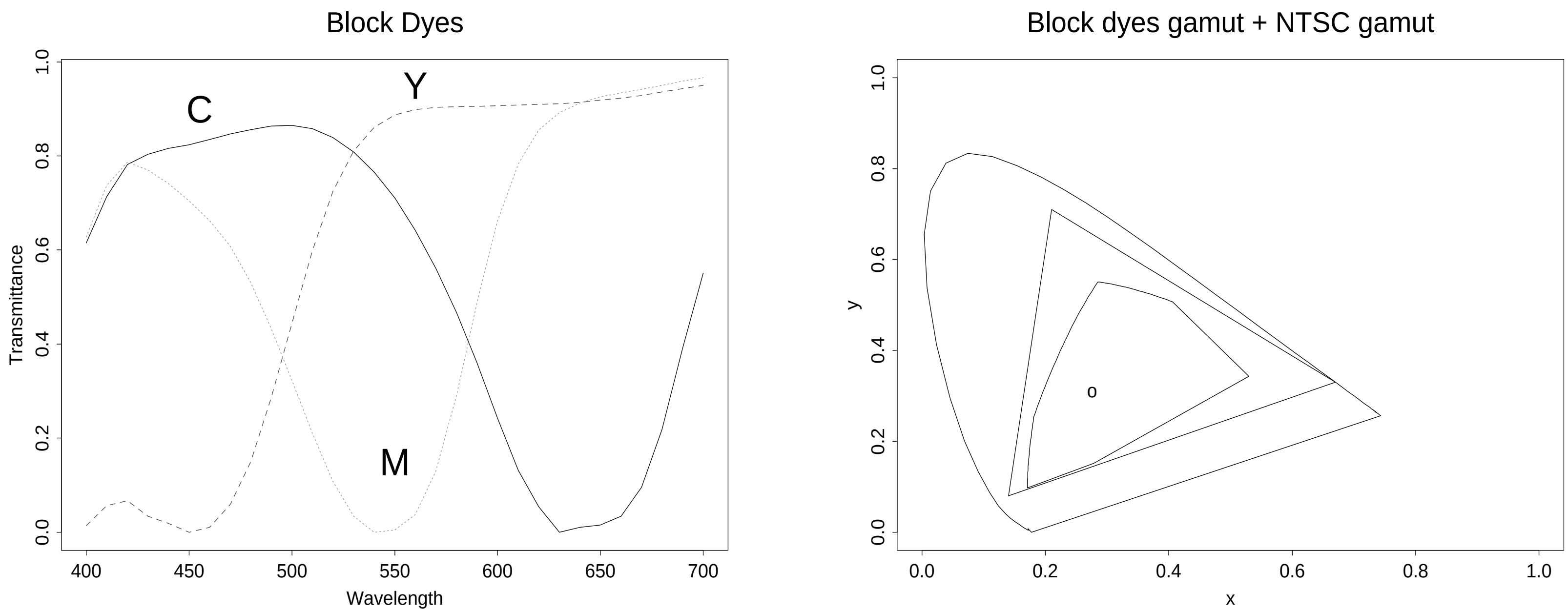

Fig. 4.17: (a): Transmission curves for block dyes. (b): Spectrum locus, triangular NTSC gamut, and 6-vertex printer gamut

4.3 Color Models in Video

- **Video Color Transforms**

  (a) Largely derive from older analog methods of coding color for TV. Luminance is separated from color information.

  (b) For example, a matrix transform method similar to Eq. (4.9) called YIQ is used to transmit TV signals in North America and Japan.

  (c) This coding also makes its way into VHS video tape coding in these countries since video tape technologies also use YIQ.

  (d) In Europe, video tape uses the PAL or SECAM codings, which are based on TV that uses a matrix transform called YUV.

  (e) Finally, digital video mostly uses a matrix transform called YCbCr that is closely related to YUV

YUV Color Model

(a) YUV codes a luminance signal (for gamma-corrected signals) equal to Y' in Eq. (4.20). the "luma".

(b) **Chrominance** refers to the difference between a color and a reference white at the same luminance. $\longrightarrow$ use color differences U, V:

$$U = B' - Y', \qquad V = R' - Y' \tag{4.27}$$

From Eq. (4.20),

$$\begin{bmatrix} Y' \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.144 \\ -0.299 & -0.587 & 0.886 \\ 0.701 & -0.587 & -0.114 \end{bmatrix} \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} \tag{4.28}$$

(c) For gray, $R' = G' = B'$, the luminance Y' equals to that gray, since $0.299 + 0.587 + 0.114 = 1.0$. And for a gray ("black and white") image, the chrominance (U, V) is zero.

(d) In the actual implementation U and V are rescaled to have a more convenient maximum and minimum.

(e) For dealing with composite video, it turns out to be convenient to contain U, V within the range $-1/3$ to $+4/3$. So U and V are rescaled:

$$\begin{aligned} U &= 0.492111\,(B' - Y') \\ V &= 0.877283\,(R' - Y') \end{aligned} \tag{4.29}$$

The chrominance signal = the composite signal C:

$$C = U\cdot cos(\omega t) + V\cdot sin(\omega t) \tag{4.30}$$

(f) Zero is not the minimum value for U, V. U is approximately from blue ($U > 0$) to yellow ($U < 0$) in the RGB cube; V is approximately from red ($V > 0$) to cyan ($V < 0$).

(g) Fig. 4.18 shows the decomposition of a color image into its Y', U, V components. Since both U and V go negative, in fact the images displayed are shifted and rescaled.

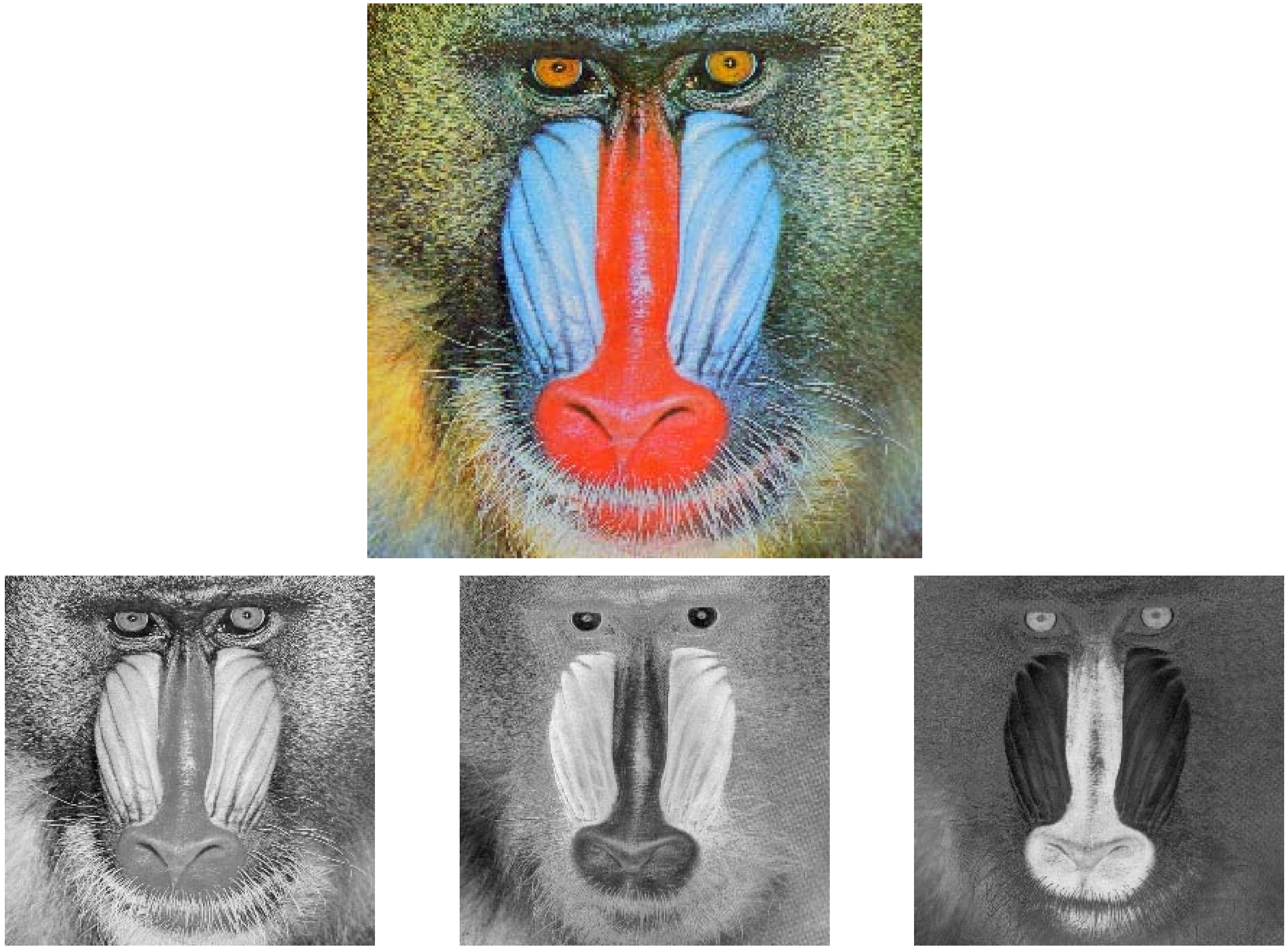

Fig. 4.18: $Y'UV$ decomposition of color image. Top image (a) is original color image; (b) is Y'; (c,d) are (U, V)

YIQ Color Model

- YIQ is used in NTSC color TV broadcasting. Again, gray pixels generate zero (I, Q) chrominance signal.

  (a) I and Q are a rotated version of U and V.

  (b) Y' in YIQ is the same as in YUV; U and V are rotated by 33°:

$$\begin{aligned} I &= 0.492111(R' - Y')\cos 33^\circ - 0.877283(B' - Y')\sin 33^\circ \\ Q &= 0.492111(R' - Y')\sin 33^\circ + 0.877283(B' - Y')\cos 33^\circ \end{aligned} \tag{4.31}$$

  (c) This leads to the following matrix transform:

$$\begin{bmatrix} Y' \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.144 \\ 0.595879 & -0.274133 & -0.321746 \\ 0.211205 & -0.523083 & 0.311878 \end{bmatrix} = \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} \tag{4.32}$$

  (d) Fig. 4.19 shows the decomposition of the same color image as above, into YIQ components.

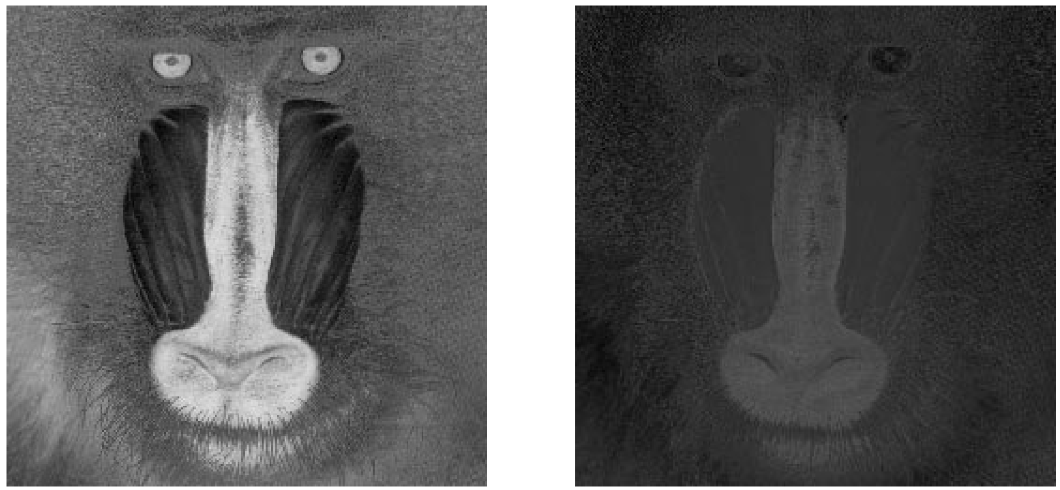

Fig.4.19: I and Q components of color image.

YCbCr Color Model

- The Rec. 601 standard for digital video uses another color space, YC_bC_r, often simply written YCbCr — closely related to the YUV transform.

  (a) YUV is changed by scaling such that C_b is U, but with a coefficient of 0.5 multiplying B'. In some software systems, C_b and C_r are also shifted such that values are between 0 and 1.

  (b) This makes the equations as follows:

$$\begin{aligned} C_b &= ((B' - Y')/1.772) + 0.5 \\ C_r &= ((R' - Y')/1.402) + 0.5 \end{aligned} \tag{4.33}$$

  (c) Written out:

$$\begin{bmatrix} Y' \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.144 \\ -0.168736 & -0.331264 & 0.5 \\ 0.5 & -0.418688 & -0.081312 \end{bmatrix} \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} + \begin{bmatrix} 0 \\ 0.5 \\ 0.5 \end{bmatrix} \tag{4.34}$$

(d) In practice, however, Recommendation 601 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', C_b, C_r in [0..255] via the transform:

$$\begin{bmatrix} Y' \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 65.481 & 128.553 & 24.966 \\ -37.797 & -74.203 & 112 \\ 112 & -93.786 & -18.214 \end{bmatrix} \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} + \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} \tag{4.35}$$

(f) The YCbCr transform is used in JPEG image compression and MPEG video compression.

4.4 Further Exploration

⟶ Link to Further Exploration for Chapter 4.

- Links in the Chapter 4 section of the "Further Exploration" directory on the text website include:
  - More details on gamma correction for publication on the WWW.
  - The full specification of the new sRGB standard color space for WWW applications.
  - A link to an excellent review of color transforms.
  - A Matlab script to exercise (and expand upon) the color transform functions that are part of the Image Toolbox in Matlab: the standard 'Lena' image is transformed to YIQ and to YCbCr.
  - The new MPEG standard, MPEG-7, will be discussed in Chapter 12. MPEG-7 includes six color spaces.