Fundamentals of Multimedia

3rd Edition + supplementary material

Chapter 4 :

Color in Image and Video

- This chapter explores:
  - several issues in the use of color, since color is vitally important in multimedia programs

- in this chapter we shall consider the following topics:
  - Color Science
  - Color Models in Images
  - Color Models in Video.

Color in Image and Video

- 4.1 Color Science
- 4.2 Color Models in Images
- 4.3 Color Models in Video

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 (ياقوت) 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.

4.1 Color Science

- **Spectrophotometer:** device used to measure visible light, by reflecting light from a diffraction grating (prism منشور) (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.

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

4.1 Color Science

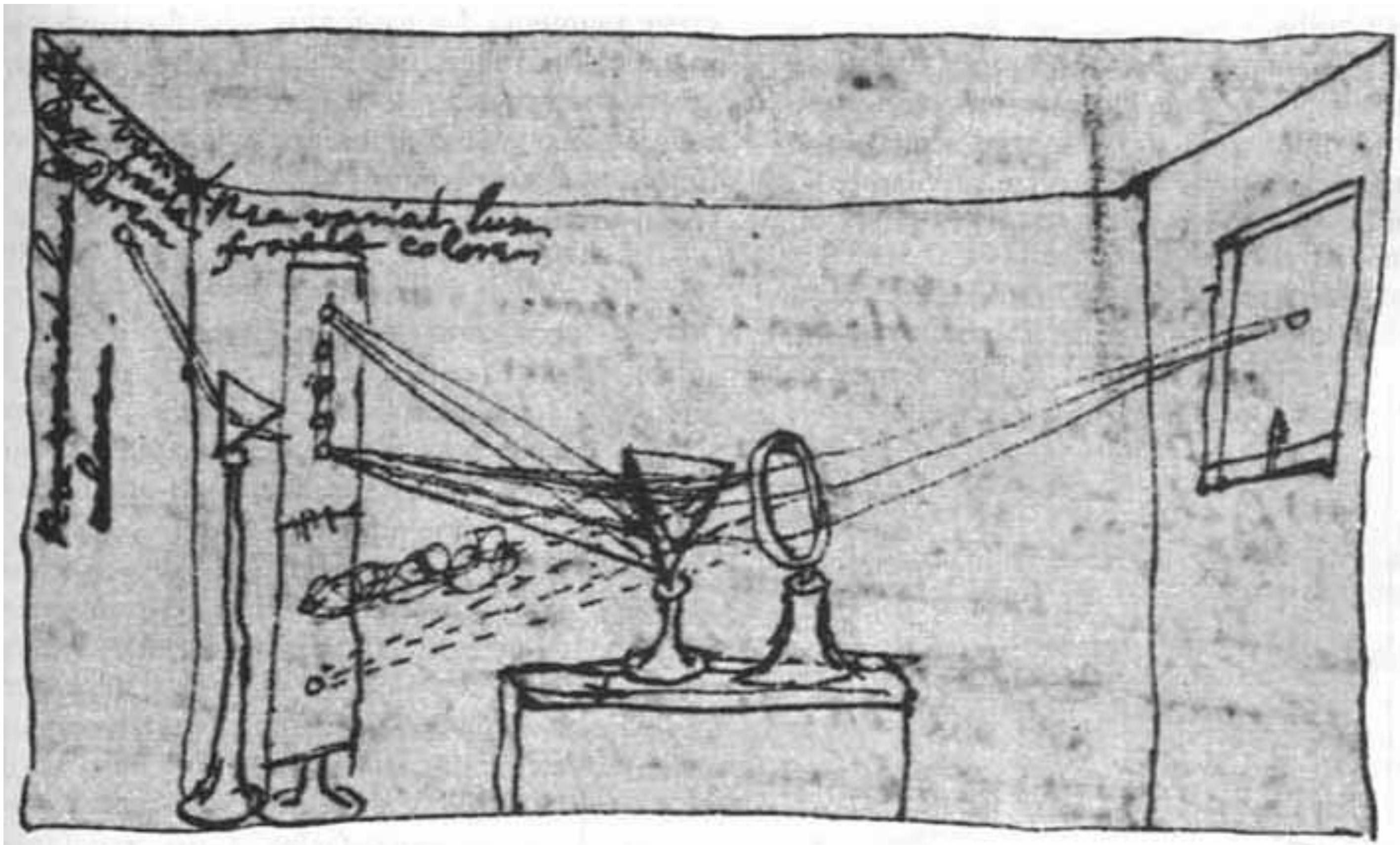
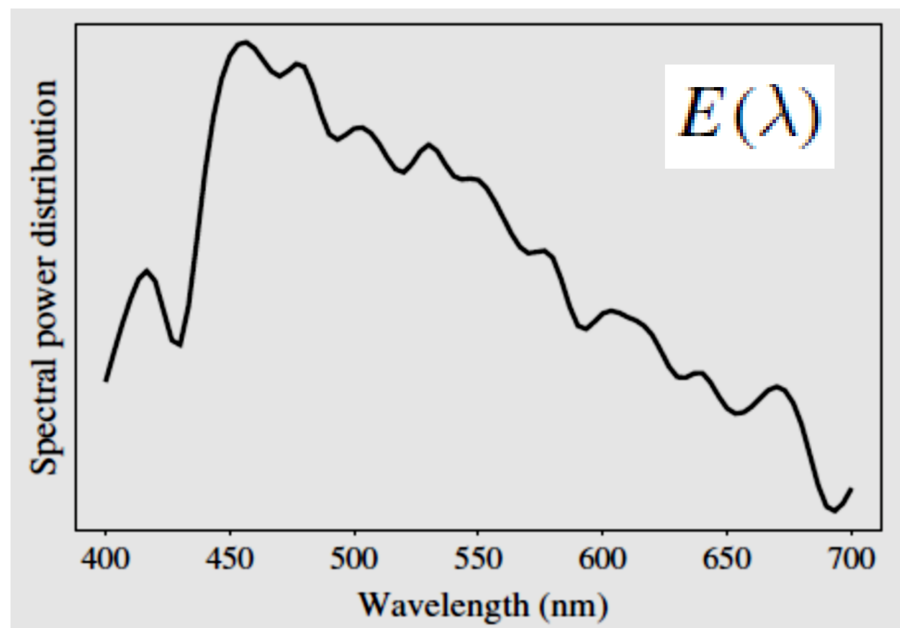


Fig. 4.1: Sir Isaac Newton's experiments.

4.1 Color Science

- Fig. 4.2 (See Book) 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(* λ *)*.

Note: In practice, measurements are used that effectively sum up voltage in a small wavelength range, say 5 or 10 nm, so such plots usually consist of segments joining function values every 10 nm. That means, also, that such profiles are actually stored as vectors;



4.1 Color Science

- **Human Vision**
- *The eye works like a camera, with the lens focusing an image* onto the retina شبكية (upside-down and left-right reversed).

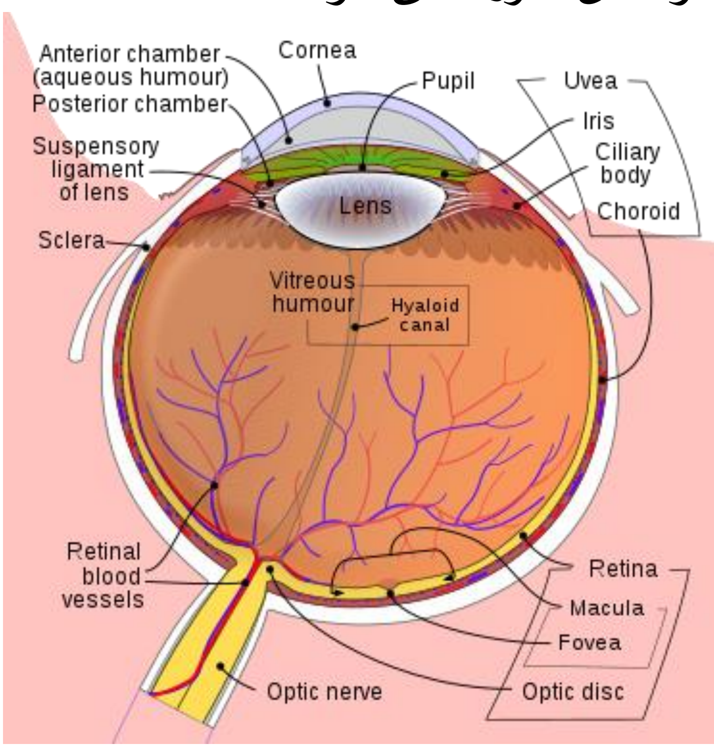
- *The retina consists of an array of rods and three kinds of* cones.
- *The rods come into play when light levels are low and produce* a image in shades of gray ("all cats are gray at night!").

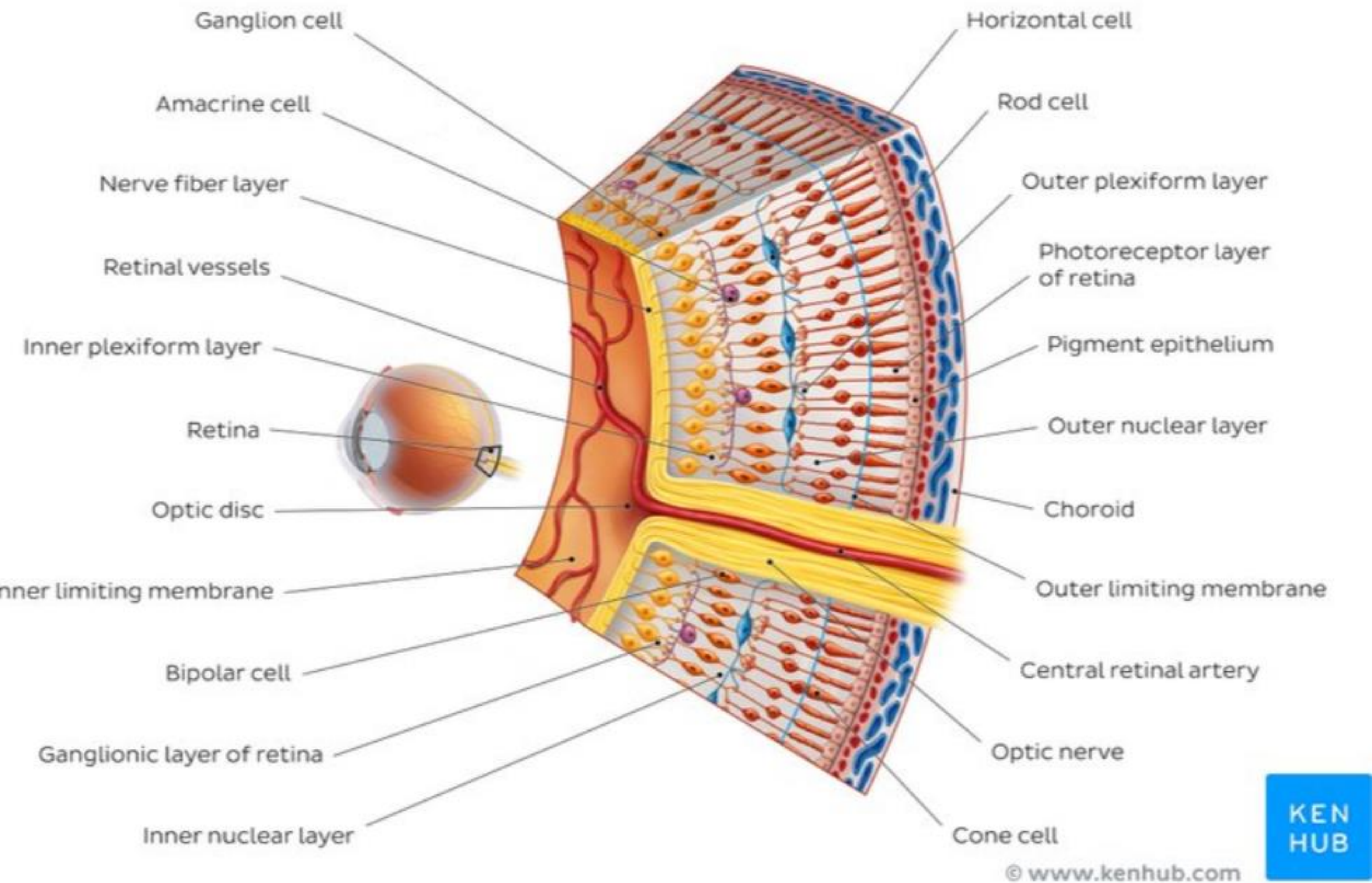
- *For higher light levels, the cones each produce a signal. Because of* their differing pigments, the three kinds of cones are most sensitive to red *(R), green (G), and blue (B) light.*

- *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.

نگاهی به آناتومی چشم انسان

- در کالبدشناسی چشم انسان، **گودهٔ مرکزی** (انگلیسی: Fovea centralis) به فرورفتگی ریزی در مرکز لکهٔ زرد گفته می‌شود که حاوی یاخته‌های مخروطی Cone است که به علت افزایش طولشان به یاخته‌های استوانه‌ای Rod شبیه شده‌اند.[۱]
- در مرکز شبکیه، ناحیه کوچکی به نام لکه زرد وجود دارد که در آن فقط نورون‌های مخروطی یافت می‌شود. در نواحی دورتر از لکه زرد به‌تدریج از تعداد نورون‌های مخروطی کاسته شده و بر تعداد نورون‌های استوانه‌ای افزوده می‌شود.




Ganglion cell
Amacrine cell
Nerve fiber layer
Retinal vessels
Inner plexiform layer
Retina
Optic disc
nner limiting membrane
Bipolar cell
Ganglionic layer of retina
Inner nuclear layer
Horizontal cell
Rod cell
Outer plexiform layer
Photoreceptor layer of retina
Pigment epithelium
Outer nuclear layer
Choroid
Outer limiting membrane
Central retinal artery
Optic nerve
Cone cell
KEN HUB
© www.kenhub.com

4.1 Color Science

- **Spectral Sensitivity of the Eye** حساسيت طيفي چشم
- *The sensitivity of our receptors is a function of wave*-length (Fig. 4.3).
- *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.
- 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 (λ) and is formed as the sum of the response* curves for Red, Green, and Blue.

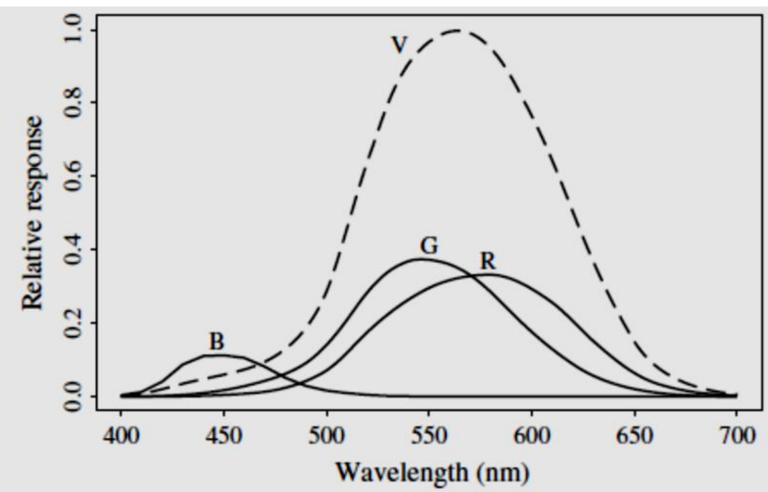


Fig. 4.3 R,G, and B cones, and luminous-efficiency curve V(λ)

4.1 Color Science

- **Spectral طيفي Sensitivity of the Eye**
- The eye is most sensitive to light in the middle of the visible Spectrum طيف.
- The eye has about 6 million cones, but the proportions of *R, G, and B cones* are different.
- They likely are present in the ratios 40:20:1
- So the achromatic channel produced by the cones is thus something like $2R + G + B/20$.
- The spectral sensitivity function:

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

4.1 Color Science

Image Formation

- In most situations, we actually image light that is reflected from a surface.
- Surfaces reflect different amounts of light at different wavelengths, and dark surfaces reflect less energy than light surfaces.
- then the reflected light filtered by the eye's cone

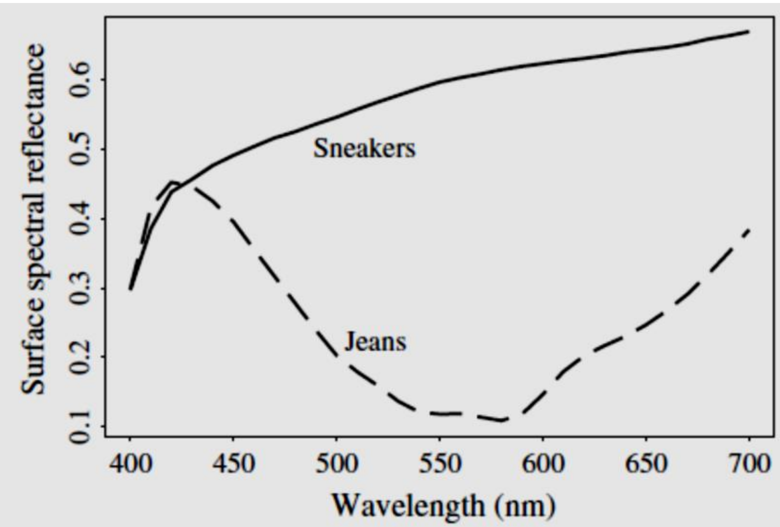


Fig. 4.4 Surface spectral reflectance functions $S(\lambda)$ for objects

4.1 Color Science

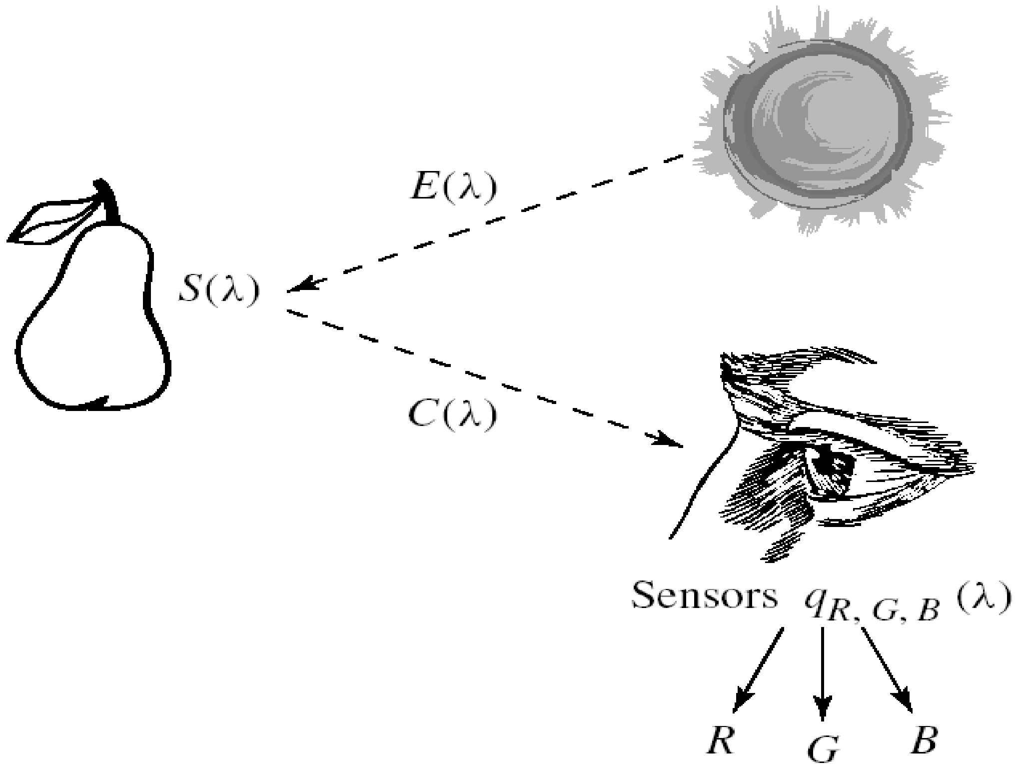


Fig. 4.5: Image formation model.

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

4.1 Color Science

4.1.5.Camera Systems

- Camera systems are made in a similar fashion; a good 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.*

4.1.6. Gamma Correction

- In TV systems, The light emitted is actually roughly proportional to the voltage raised to a power; this power is called "gamma," with symbol γ. The value of gamma is around 2.2.
- We can precorrect for this situation by actually applying the inverse transformation before generating TV voltage signals. It is customary to append a prime to signals that are "gamma-corrected" by raising to the power (1/γ) before transmission.

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

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

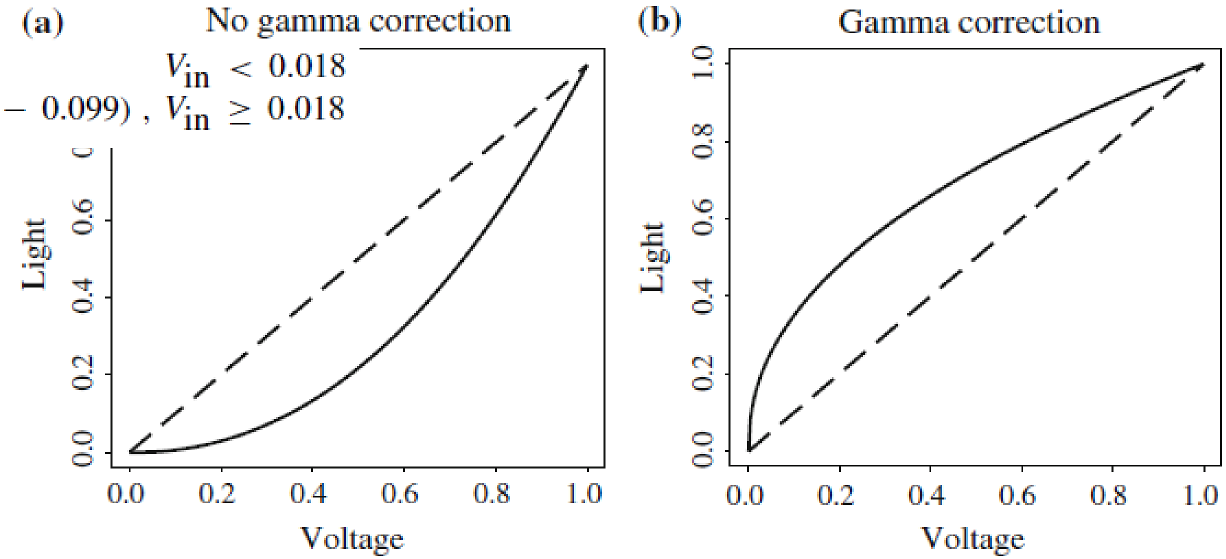


Fig. 4.6 **a** Effect of putative standard CRT (mimiced by an actual modern display) on light emitted from screen (voltage is normalized to range 0..1). **b** Gamma correction of signal

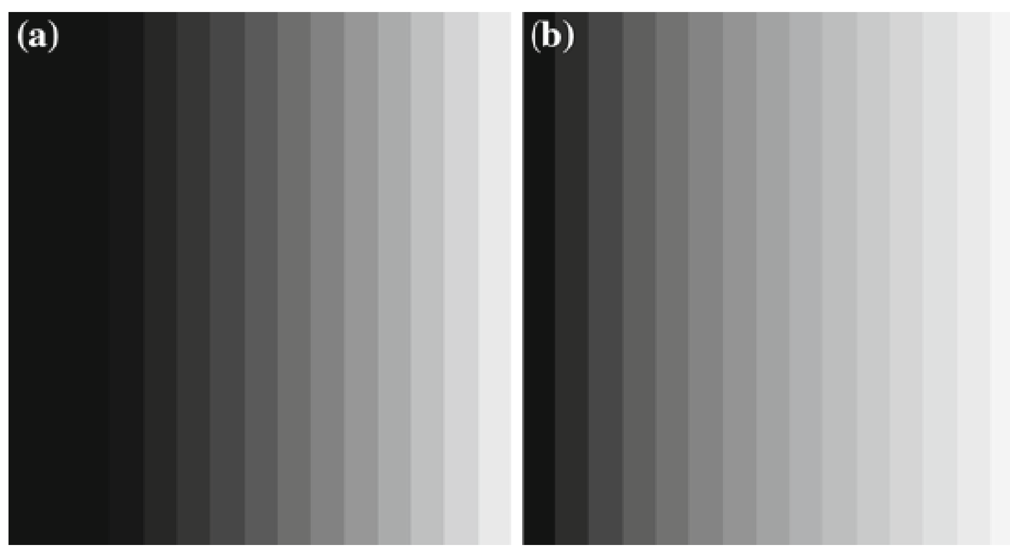


Fig. 4.7 **a** Display of ramp from 0 to 255, with no gamma correction. **b** Image with gamma correction applied

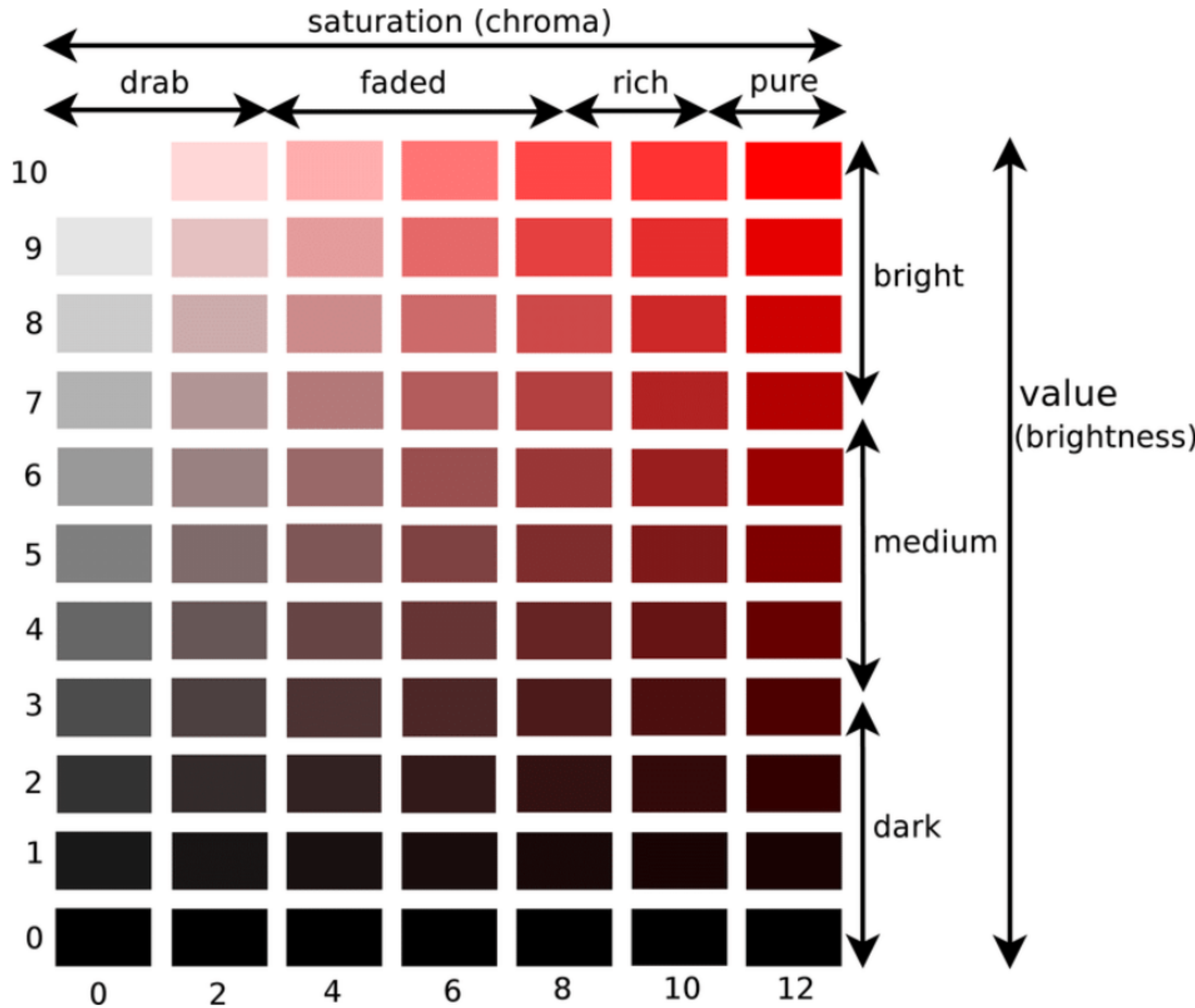
- بخش های 4.1.7 الی 4.1.15 جزو درس محسوب نمی شود.

4.1.16.Munsell color system

The Munsell color system specifies colors based on hue, chroma (saturation) and brightness (value).

The image shows the combinations of brightness and chroma for the color Red.

Munsell system divides the hue space into 10 up to 40 different colors.



Category	Red	Red Yellow	Yellow	Yellow Green	Green	Green Blue	Blue	Blue Purple	Purple	Purple Red

4.2 Color Models in Images

- RGB is the most used color model to store images.
- We expect to be able to use 8 bits per color channel for color that is accurate enough.
- 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 since gamma correction results in many fewer available integer levels.
- RGB is also called **device-dependent** since the sensitivity of each camera filter could result in considerably false color reproduction.
- Luminance-based color systems are less device dependent as being closer to human-perception.

4.2.2 Multi-sensor Cameras

- Using cameras with more than three sensors, i.e., more than three color filters so that images are closer to full spectrum.
- Using a rotating filter, which places a different color filter in the light path over a quick series of shots.
- To extend the camera's sensitivity into the infrared.

4.2.3 Camera-Dependent/independent Color

- CIE color coordinates that are meant to be *camera independent*, i.e., human-perception oriented.
- The proposed set of axes:
  - $L*$, human visual system percepts of *Lightness*
  - $a*$ (redness-greenness) or hue $h*$, meaning a magnitude-independent notion of color;
  - $b*$ (yellowness-blueness) or chroma $c*$, meaning the purity (vividness) of a color.

HSV Color model

- H stands for hue;
- S stands for "saturation" of a color, defined by chroma divided by its
- luminance—the more desaturated the color is the closer it is to gray;
- V stands for "value," meaning a correlate of brightness as perceived by humans.

$$M = \max\{R, G, B\}$$

$$m = \min\{R, G, B\}$$

$$V = M$$

$$S = \begin{cases} 0 & \text{if } \mathrm{V} = 0 \\ \left(V - m\right)/V & \text{if } \mathrm{V} > 0 \end{cases}$$

$$H = \begin{cases} 0 & \text{if } S = 0 \\ 60(G - B)/(M - m) & \text{if } (M = R \text{ and } G \geq B) \\ 60(G - B)/(M - m) + 360 & \text{if } (M = R \text{ and } G < B) \\ 60(B - R)/(M - m) + 120 & \text{if } M = G \\ 60(R - G)/(M - m) + 240 & \text{if } M = B \end{cases}$$

4.2.4 Subtractive Color: CMY Color Model

- RGB is in fact an additive color system: Color=R+G+B
- Subtractive color system: primaries that amount to −red (cyan or C), −green (magenta or M), and −blue (yellow or Y) (absence of these colors)
- Suitable for printers since printed on a white background.

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

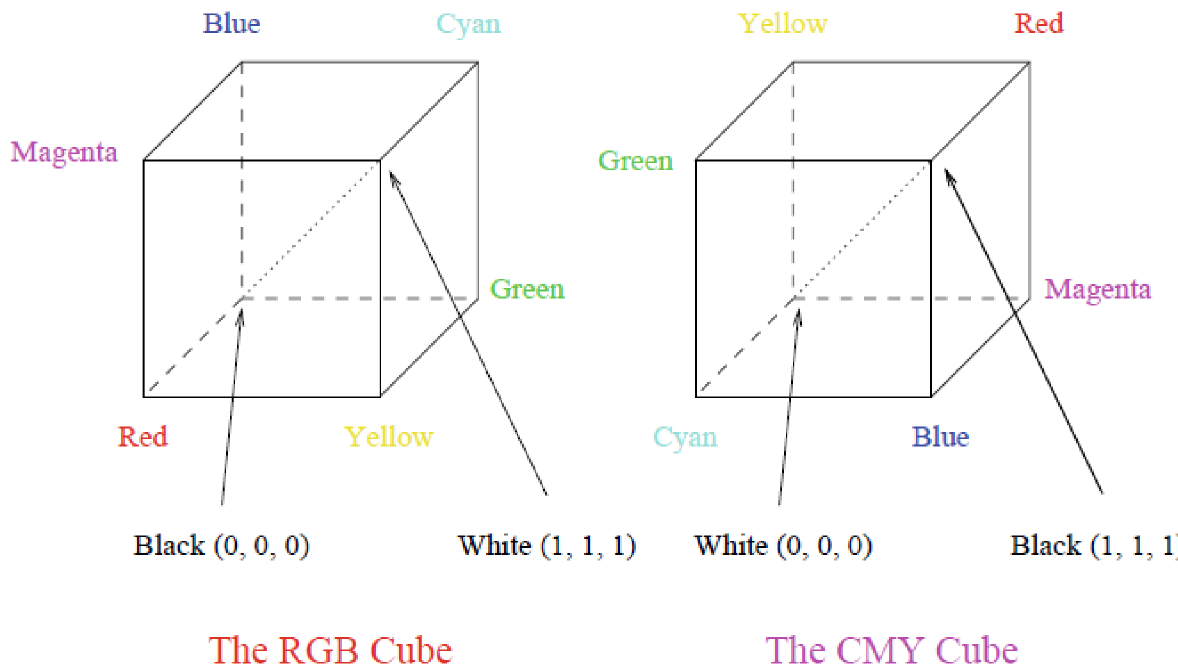


Fig. 4.15 RGB and CMY color cubes

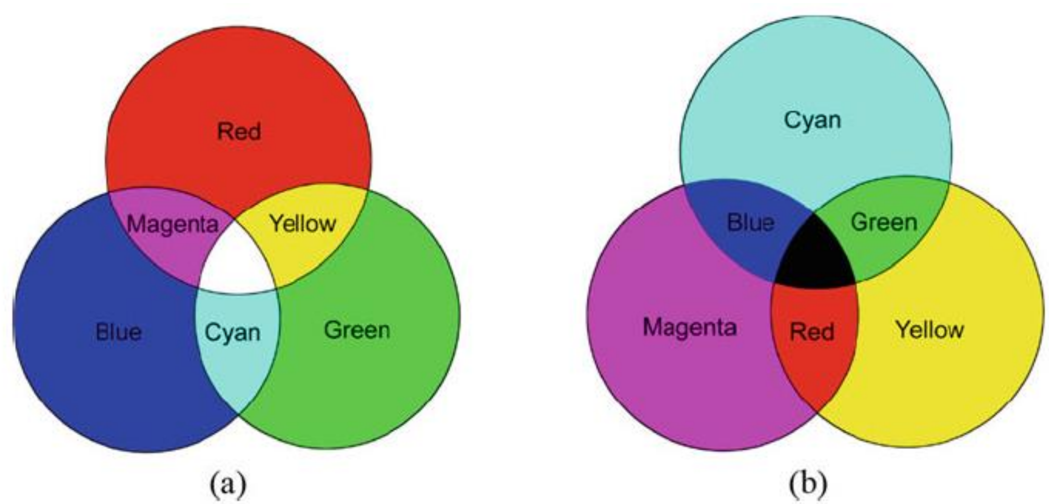


Fig. 4.16 Additive and subtractive color. **a** RGB is used to specify additive color. **b** CMY is used to specify subtractive color

4.2.6 Undercolor Removal: CMYK System

- C, M, and Y are supposed to mix to black.
- However, more often they mix to a muddy brown if their **proportion** does not precisely hold.
- A simple approach to producing sharper printer colors is to calculate that part of the three-color mix that would be black: K=min(C,M,Y)

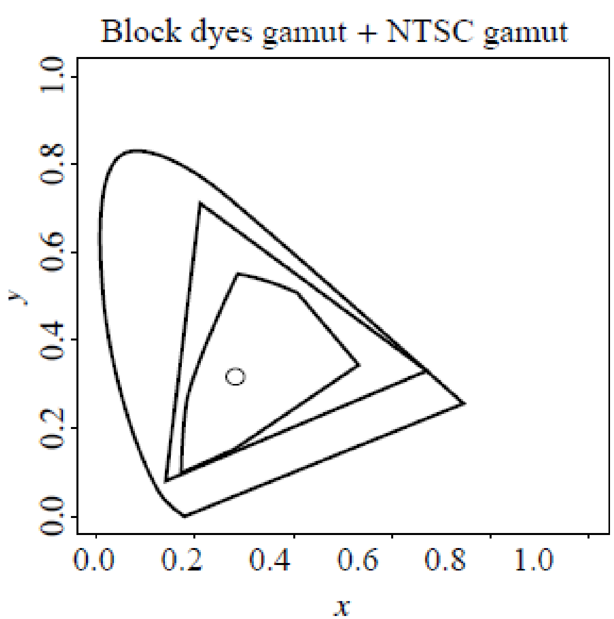
$$K \equiv \min\{C, M, Y\}$$

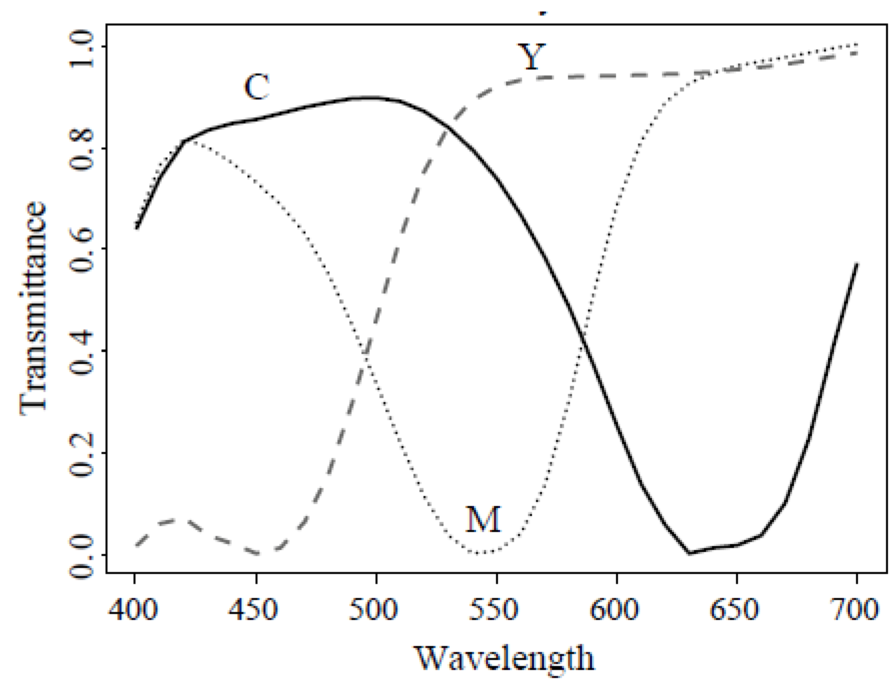
$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} \Rightarrow \begin{bmatrix} C - K \\ M - K \\ Y - K \end{bmatrix}$$

4.2.8 Color production in multi-ink printers

- A cyan printing ink does not completely blocks red light **in practice.** So is for magenta and yellow.
- This leads to "crosstalk" between the color channels and errors in producing colors and limiting achievable colors in printing.

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





4.3 Color Models in Video

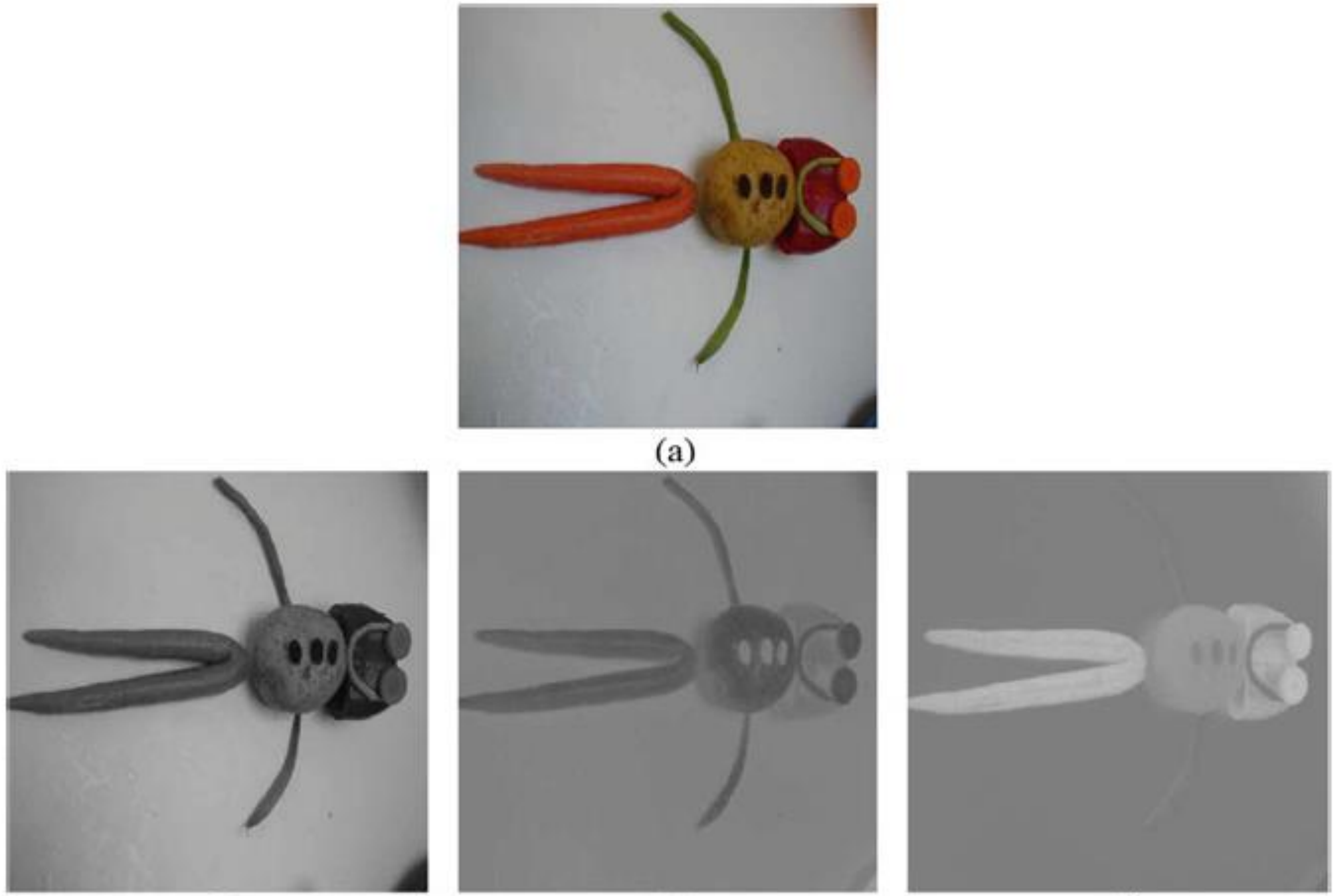
- Typically, some version of the luminance is combined with color information in a single signal.
- NTSC coding: uses YIQ
- PAL or SECAM codings: uses a matrix transform called YUV
- The luma Y ` is the CIE luminance (**brightness**) value Y that is gamma-corrected.
- ***chrominance*** refers to the difference between a color and a reference white at the same luminance. It can be represented by the color *differences U*, *V*:

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

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

Then the chrominance signal is composed of U and V as the composite signal

$$
C = U \cdot \cos(\omega t) + V \cdot \sin(\omega t) \tag{4.35}
$$



(a)

(b) (c) (d)

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

4.3.3 YIQ Color Model

- It is thought that, U and V are do not capture the most-to-least hierarchy of human vision sensitivity.
- Although U and V nicely define the color differences, they do not best correspond to actual human perceptual color sensitivities.
- In NTSC, I and Q are used instead. is roughly the orange–blue direction and Q roughly corresponds to the purple–green direction.

$$\begin{aligned} I &= 0.492111(R' - Y')\cos 33^\circ - 0.877283(B' - Y')\sin 33^\circ \\ Q &= 0.492111(R' - Y')\sin 33^\circ + 0.877283(B' - Y')\cos 33^\circ \end{aligned} \tag{4.36}$$

$$\begin{bmatrix} Y' \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.595879 & -0.274133 & -0.321746 \\ 0.211205 & -0.523083 & 0.311878 \end{bmatrix} \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix}$$

- Question: why are these I and Q components in gray scale instead of being colorful?

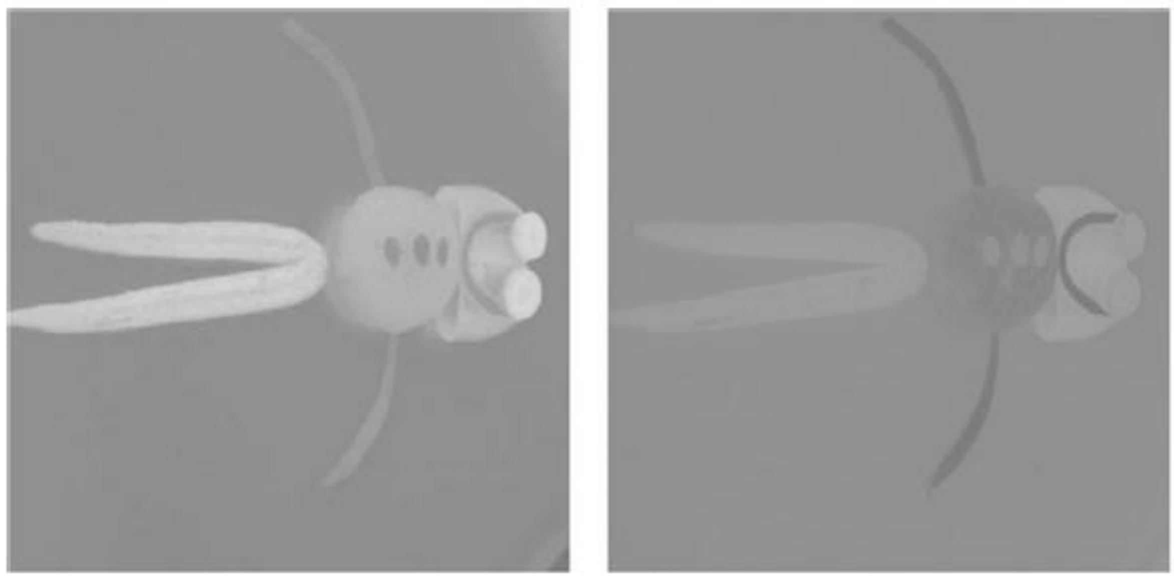

(a) (b)

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

4.3.4 YCbCr Color Model

- The international standard for component (three-signal, studio quality) *digital* video is officially "Recommendation ITU-R BT.601-4" (known as Rec. 601).
- This standard uses another color space, *YCbCr*, often simply written YCbCr.
- The YCbCr transform is closely related to the YUV transform.

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

$$
\begin{bmatrix} Y' \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.168736 & -0.331264 & 0.5 \\ 0.5 & -0.418688 & -0.081312 \end{bmatrix} \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} + \begin{bmatrix} 0 \\ 0.5 \\ 0.5 \end{bmatrix} \tag{4.40}
$$

End of Chapter 4