2. For each pair of horizontally neighboring pixels (x, y), $(x, y+1)$ generate a pair of independent random numbers r, ϕ in the range $[0, 1]$.
3. Calculate

$$z_1 = \sigma \cos(2\pi\phi)\sqrt{-2 \ln r}\,, \qquad z_2 = \sigma \sin(2\pi\phi)\sqrt{-2 \ln r}\,. \tag{2.11}$$

(This is the Box-Muller transform which assumes that z_1, z_2 are independently normally distributed with zero mean and variance σ^2.)

4. Set $f'(x,y) = g(x,y) + z_1$ and $f'(x,y+1) = g(x,y+1) + z_2$, where g is the input image.
5. Set

$$f(x,y) = \begin{cases} 0 & \text{if } f'(x,y) < 0\,, \\ G-1 & \text{if } f'(x,y) > G-1\,, \\ f'(x,y) & \text{otherwise,} \end{cases} \tag{2.12}$$

$$f(x,y+1) = \begin{cases} 0 & \text{if } f'(x,y+1) < 0\,, \\ G-1 & \text{if } f'(x,y+1) > G-1\,, \\ f'(x,y+1) & \text{otherwise.} \end{cases} \tag{2.13}$$

6. Go to 3 until all pixels have been scanned.

The truncation performed by equations (2.12) and (2.13) will attenuate the Gaussian nature of the noise; this will become more marked for values of σ that are high relative to G. Other algorithms for noise generation may be found in [Pitas, 1993].

Equation (2.10) leads to a definition of **signal-to-noise ratio** (SNR); computing the total square value of the noise contribution

$$E = \sum_{(x,y)} \nu^2(x,y)$$

we compare this with the total square value of the observed signal

$$F = \sum_{(x,y)} f^2(x,y)\,.$$

The **signal-to-noise ratio** is then

$$\text{SNR} = \frac{F}{E} \tag{2.14}$$

(strictly, we are comparing the mean observation with the mean error—the computation is obviously the same). SNR represents a measure of image quality, with high values being 'good'.

Signal-to-noise is often expressed in the logarithmic scale, in decibels

The noise magnitude depends in many cases on the signal magnitude itself

$$f = g\,\nu\,. \tag{2.16}$$

This model describes **multiplicative noise**. An example of multiplicative noise is television raster degradation, which depends on TV lines; in the area of a line this noise is maximal, and between two lines it is minimal. Another example of multiplicative noise is the degradation of film material caused by the finite size of silver grains used in photosensitive emulsion.

Quantization noise occurs when insufficient quantization levels are used, for example, 50 levels for a monochromatic image. In this case false contours appear. Quantization noise can be eliminated simply, see Section 2.2.2.

Impulse noise means that an image is corrupted with individual noisy pixels whose brightness differs significantly from that of the neighborhood. The term **salt-and-pepper noise** is used to describe saturated impulsive noise—an image corrupted with white and/or black pixels is an example. Salt-and-pepper noise can corrupt binary images.

The problem of suppressing noise in images is addressed in Chapter 5. If nothing is known a priori about noise properties, local pre-processing methods are appropriate (Section 5.3). If the noise parameters are known in advance, image restoration techniques can be used (Section 5.4).

2.4 Color images

Human color perception adds a subjective layer on top of underlying objective physical properties—the wavelength of electromagnetic radiation. Consequently, color may be considered a psychophysical phenomenon.

Color has long been used in painting, photography and films to display the surrounding world to humans in a similar way to that in which it is perceived in reality. There is considerable literature on the variants in the naming of colors across languages, which is a very subtle affair [Kay, 2005]. The human visual system is not very precise in perceiving color in absolute terms; if we wish to express our notion of color precisely we would describe it relative to some widely used color which is used as a standard: recall, e.g., the red of a British public telephone box. There are whole industries which present images to humans— the press, films, displays, and hence a desire for color constancy. In computer vision, we have the advantage of using a camera as a measuring device, which yields measurements in absolute quantities.

Newton reported in the 17th century that white light from the sun is a spectral mixture, and used the optical prism to perform decomposition. This was a radical idea to propose at time; over 100 years later influential scientists and philosophers such as Goethe refused to believe it.

2.4.1 Physics of color

The electromagnetic spectrum is illustrated in Figure 2.23.

Only a narrow section of the electromagnetic spectrum is visible to a human, with wavelength λ from approximately 380 nm to 740 nm. Visible colors with the wavelengths shown in Figure 2.24 are called **spectral colors** and are those which humans see when

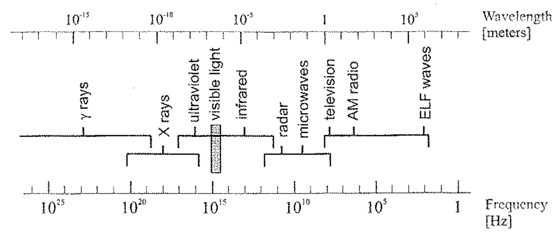


Figure 2.23: Division of the whole electromagnetic spectrum (ELF means Extremely Low Frequencies).

on the sky. Colors can be represented as combinations of the **primary colors**, e.g., red, green, and blue, which for the purposes of standardization have been defined as 700 nm, 546.1 nm, and 435.8 nm, respectively [Pratt, 1978], although this standardization does not imply that all colors can be synthesized as combinations of these three.

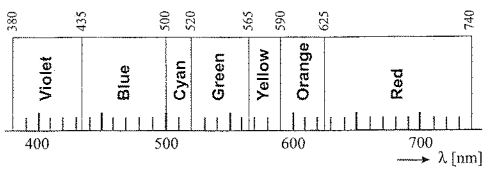


Figure 2.24: Wavelength λ of the spectrum visible to humans.

The intensity of irradiation for different wavelengths λ usually changes. This variation is expressed by a **power spectrum** (called also power spectrum distribution) $S(\lambda)$.

Why do we see the world in color? There are two predominant physical mechanisms describing what happens when a surface is irradiated. First, the **surface reflection** rebounds incoming energy in a similar way to a mirror. The spectrum of the reflected light remains the same as that of the illuminant and it is independent of the surface—recall that shiny metals 'do not have a color'. Second, the energy diffuses into the material and reflects randomly from the internal pigment in the matter. This mechanism is called **body reflection** and is predominant in dielectrics as plastic or paints. Figure 2.25 illustrates both surface reflection (mirroring along surface normal **n**) and body reflection. Colors are caused by the properties of pigment particles which absorb certain wavelengths from the incoming illuminant wavelength spectrum.

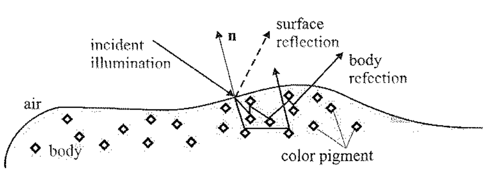


Figure 2.25: Observed color of objects is caused by certain wavelength absorptions by pigment

Most sensors used for color capture, e.g., in cameras, do not have direct access to color; the exception is a **spectrophotometer** which in principle resembles Newton's prism. Incoming irradiation is decomposed into spectral colors and intensity along the spectrum with changing wavelength λ is measured in a narrow wavelength band, for instance, by a mechanically moved point sensor. Actual spectrophotometers use diffraction gratings instead of a glass prism.

Sometimes, intensities measured in several narrow bands of wavelengths are collected in a vector describing each pixel. Each spectral band is digitized independently and is represented by an individual digital image function as if it were a monochromatic image. In this way, **multispectral images** are created. Multispectral images are commonly used in remote sensing from satellites, airborne sensors and in industry. Wavelength usually span from ultraviolet through the visible section to infrared. For instance, the LANDSAT 4 satellite transmits digitized images in five spectral bands from near-ultraviolet to infrared. Seven or a dozen wavelength bands are common.

2.4.2 Color perceived by humans

Evolution has developed a mechanism of indirect color sensing in humans and some animals. Three types of sensors receptive to the wavelength of incoming irradiation have been established in humans, thus the term **trichromacy**. Color sensitive receptors on the human retina are the **cones**. The other light sensitive receptors on the retina are the **rods** which are dedicated to sensing monochromatically in low ambient light conditions. Cones are categorized into three types based on the sensed wavelength range: S (short) with maximum sensitivity at $\approx$ 430 nm, M (medium) at $\approx$ 560 nm, and L (long) at $\approx$ 610 nm. Cones S, M, L are occasionally called cones B, G and R, respectively, but that is slightly misleading. We do not see red solely because an L cone is activated. Light with equally distributed wavelength spectrum looks white to a human, and an unbalanced spectrum appears as some shade of color.

The reaction of a photoreceptor or output from a sensor in a camera can be modeled mathematically. Let i be the specific type of sensor, $i = 1, 2, 3$, (the retinal cone type S, M, L in the human case). Let $R_i(\lambda)$ be the spectral sensitivity of the sensor, $I(\lambda)$ be the spectral density of the illumination, and $S(\lambda)$ describe how the surface patch reflects each wavelength of the illuminating light. The spectral response q_i of the i-th sensor, can be modeled by integration over a certain range of wavelengths

$$q_i = \int_{\lambda_1}^{\lambda_2} I(\lambda)\, R_i(\lambda)\, S(\lambda)\, \mathrm{d}\lambda\,. \qquad (2.17)$$

Consider the cone types S, M, L. How does the vector (q_S, q_M, q_L) represent the color or the surface patch? It does not according to equation (2.17) since the output from the photosensors depends on the three factors $I(\lambda)$, $S(\lambda)$ and $R(\lambda)$. Only the factor $S(\lambda)$ is related to the surface patch. Only in the ideal case, when the illumination is perfectly white, i.e., $I(\lambda) = 1$, can we consider (q_S, q_M, q_L) as an estimate of the color of the surface.

Figure 2.26 illustrates qualitatively the relative sensitivities of S, M, L cones. Measurements were taken with the white light source at the cornea so that absorption of wavelength in cornea, lens and inner pigments of the eye is taken into account [Wandell

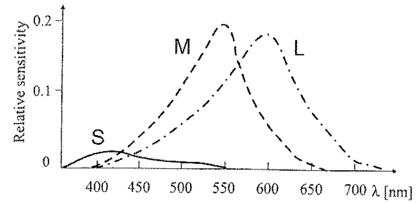


Figure 2.26: Relative sensitivity of S, M, L cones of the human eye to wavelength.

A phenomenon called **color metamer** is relevant. A metamer, in general, means two things that are physically different but perceived as the same. Red and green adding to produce yellow is a color metamer, because yellow could have also been produced by a spectral color. The human visual system is fooled into perceiving that red and green is the same as yellow.

Consider a color matching experiment in which someone is shown a pattern consisting of two adjacent color patches. The first patch displays a test light—a spectral color of certain wavelength. The second patch is created as an additive combination of three selected primary lights, e.g., colors red=645.2 nm, green=525.3 nm and blue=444.4 nm. The observer is asked to control the red, green and blue intensities until both patches look identical. This color matching experiment is possible because of the color metamer. The result of measurements (redrawn from [Wandell, 1995]) is in Figure 2.27. Negative lobes can be seen on the curves for red and green in this figure. This would seem to be impossible. For wavelengths exhibiting negative values the three additive lights do not perceptually match the spectral color because it is darker. If the perceptual match has to be obtained then the observer has to add the intensity to the patch corresponding to the spectral color. This increase of this intensity is depicted as a decrease in the color matching function. Hence the negative values.

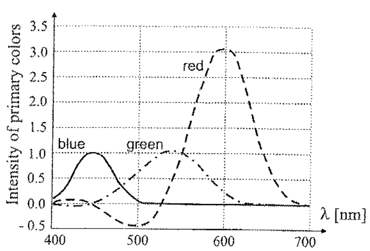


Figure 2.27: Color matching functions obtained in the color matching experiment. Intensities of the selected primary colors which perceptually match spectral color of given wavelength λ.

Human vision is prone to various illusions. Perceived color is influenced, besides the spectrum of the illuminant, by the colors and scene interpretation surrounding the observed color. In addition, eye adaptation to changing light conditions is not very fast and perception is influenced by adaptation. Nevertheless, we assume for simplicity that the spectrum of light coming to a point on the retina fully determines the color.

Since color can be defined by almost any set of primaries, the world community agreed on primaries and color matching functions which are widely used. The **color model** was introduced as a mathematical abstraction allowing us to express colors as tuples of numbers, typically as three or four values of color components. Being motivated by

Illumination, still acting in Lausanne, Switzerland) issued a technical standard called **XYZ color space**.

The standard is given by the three imaginary lights X=700.0 nm, Y=546.1 nm, Z=435.8 nm and by the color matching functions $X(\lambda)$, $Y(\lambda)$ and $Z(\lambda)$ corresponding to the perceptual ability of an average human viewing a screen through an aperture providing a 2° field of view. The standard is artificial because there is no set of physically realizable primary lights that would yield the color matching functions in the color matching experiment. Nevertheless, if we wanted to characterize the imaginary lights then, very roughly speaking, $X \approx$ red, $Y \approx$ green and $Z \approx$ blue. The CIE standard is an example of an absolute standard, i.e., defining unambiguous representation of color which does not depend on other external factors. There are more recent and more precise absolute standards: CIELAB 1976 (ISO 13665) and HunterLab (http://www.hunterlab.com). Later, we will also discuss relative color standards such as RGB color space. There are several RGB color spaces used—two computer devices may display the same RGB image differently.

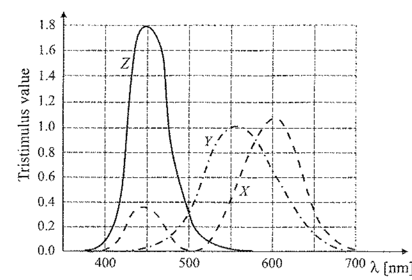
The XYZ color standard fulfills three requirements:

- Unlike the color matching experiment yielding negative lobes of color matching functions, the color matching functions of XYZ color space are required to be non-negative;
- The value of $Y(\lambda)$ should coincide with the brightness (luminance);
- Normalization is performed to assure that the power corresponding to the three color matching functions is equal (i.e., the area under all three curves is equal).

The resulting color matching functions are shown in Figure 2.28. The actual color is a mixture (more precisely a convex combination) of

$$c_X\, X + c_Y\, Y + c_Z\, Z\,, \tag{2.18}$$

where $0 \le c_X,\, c_Y,\, c_Z \le 1$ are weights (intensities) in the mixture. The subspace of colors perceivable by humans is called the color gamut and is demonstrated in Figure 2.29.



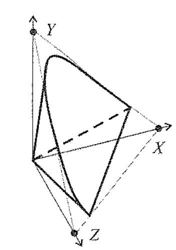


Figure 2.28: Color matching functions for the CIE standard from 1931. $X(\lambda)$, $Y(\lambda)$, $Z(\lambda)$ are color

Figure 2.29: Color gamut - a subspace of the X, Y, Z color space showing all

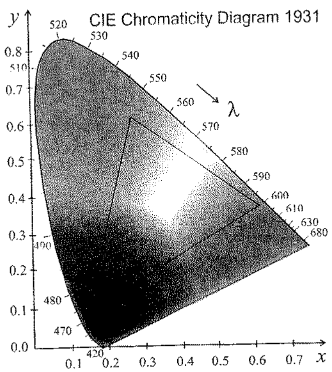


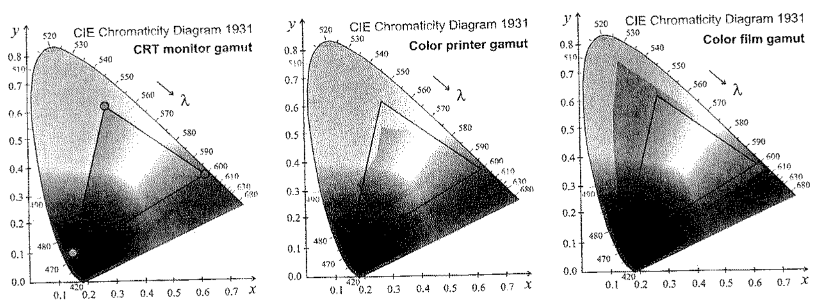
Figure 2.30: CIE chromaticity diagram is a projection of XYZ color space into a plane. The triangle depicts a subset of colors spanned by red, green, and blue. These are TV colors, i.e., all possible color which can be seen on a CRT display. *A color version of this figure may be seen in the color inset—Plate 1.*

3D figures are difficult to handle in publications, and so a planar view of a 3D color space is used. The projection plane is given by the plane passing through extremal points on all three axes, i.e., points X, Y, Z. The new 2D coordinates x, y are obtained as

$$x = \frac{X}{X+Y+Z}\,, \qquad y = \frac{Y}{X+Y+Z}\,, \qquad z = 1 - x - y\,.$$

The result of this plane projection is the CIE chromaticity diagram, see Figure 2.30. The horseshoe like subspace contains colors which people are able to see. All monochromatic spectra visible to humans map into the curved part of the horseshoe—their wavelengths are shown in Figure 2.30.

Display and printing devices use three selected real primary colors (as opposed to three syntactic primary colors of XYZ color space). All possible mixtures of these primary colors fail to cover the whole interior of the horseshoe in CIE chromaticity diagram. This situation is demonstrated qualitatively for three particular devices in Figure 2.31.



(a) CRT monitor (b) printer (c) film

... can be displayed using three typical display devices. *A color version*

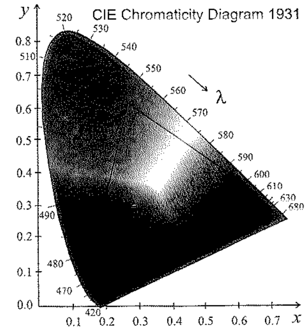
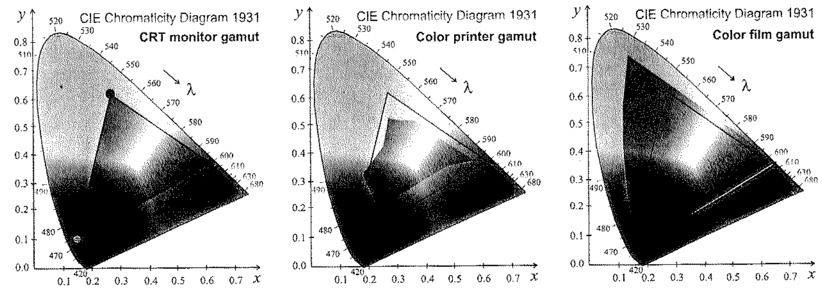


Plate 1: *Page 36, Figure 2.30.*



(a) CRT monitor (b) printer (c) film

Plate 2: *Page 36, Figure 2.31.*

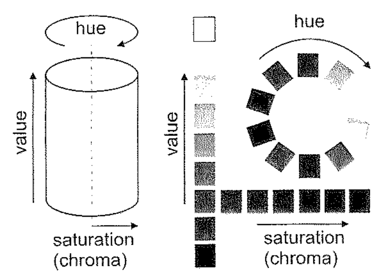


Plate 3: *Page 38, Figure 2.33.*

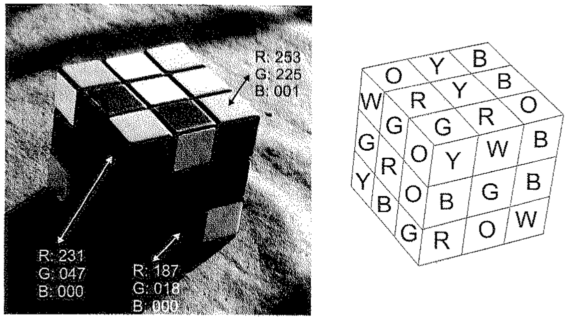
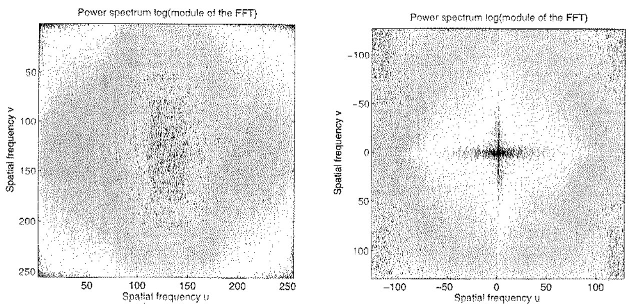
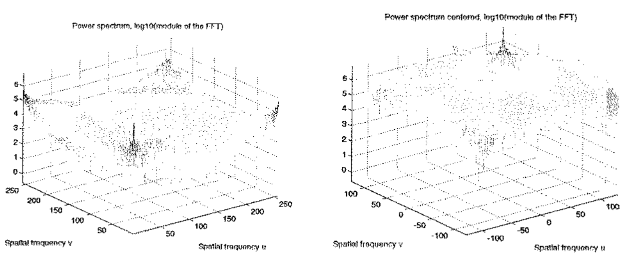


Plate 4: *Page 40, Figure 2.34.*



(a) (b)

Plate 5: *Page 61, Figure 3.7.*



(a) (b)

Plate 6: *Page 62, Figure 3.9.*

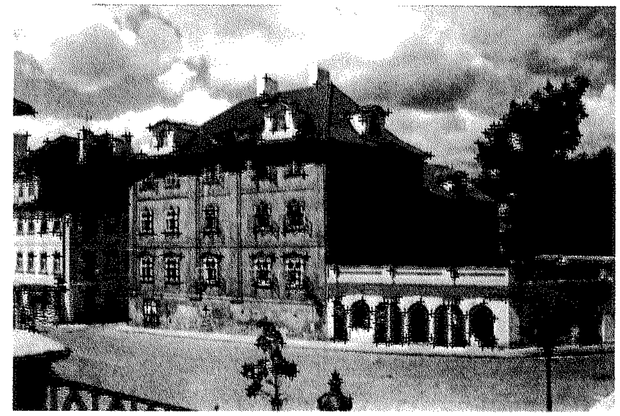


Plate 7: *Page 160, Figure 5.37.*

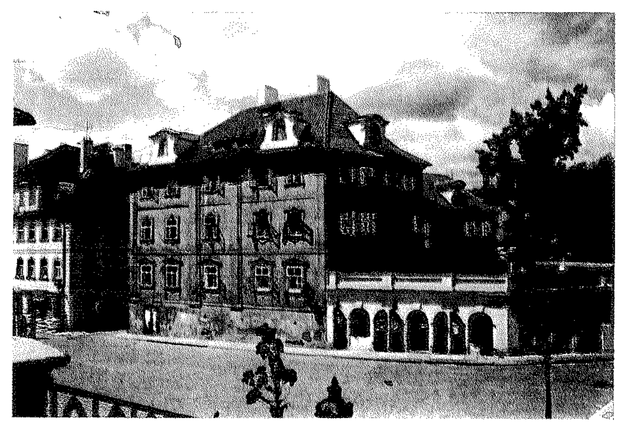


Plate 8: *Page 161, Figure 5.38.*

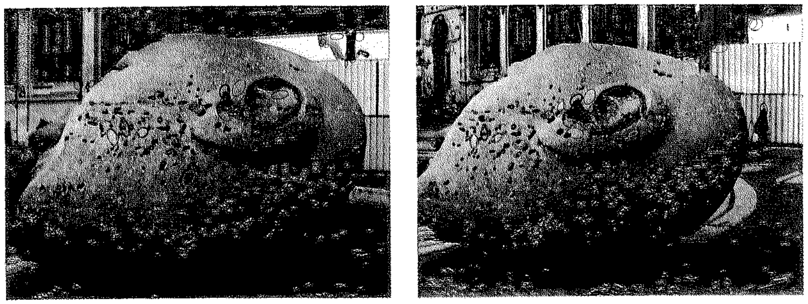


Plate 16: *Page 647, Figure 12.30.*

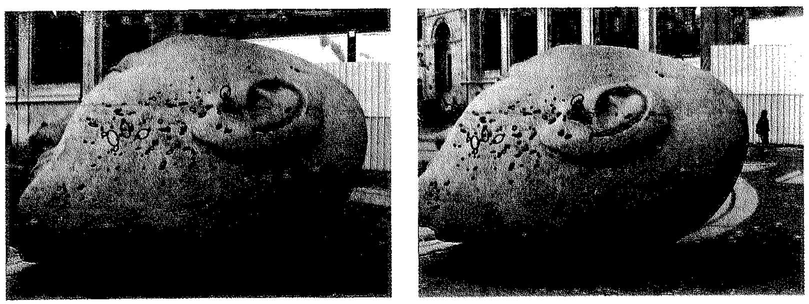


Plate 17: *Page 647, Figure 12.31.*

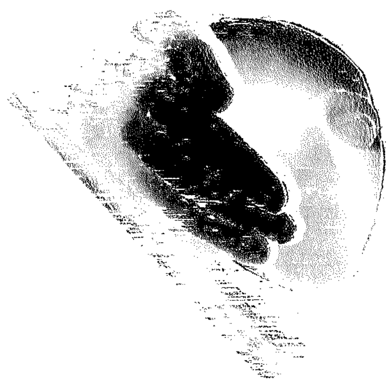


Plate 18: *Page 649, Figure 12.34.*

2.4.3 Color spaces

Several different primary colors and corresponding color spaces are used in practice, and these spaces can be transformed into each other. If the absolute color space is used then the transformation is the one-to-one mapping and does not lose information (except for rounding errors). Because color spaces have their own gamuts, information is lost if the transformed value appears out of the gamut. See [Burger and Burge, 2006] for a full explanation and for algorithms; here, we list several frequently used color spaces.

The **RGB** color space has its origin in color television where Cathode Ray Tubes (CRT) were used. RGB color space is an example of a relative color standard (as opposed to the absolute one, e.g., CIE 1931). The primary colors (R–red, G–green and B–blue) mimicked phosphor in CRT luminophore. The RGB model uses additive color mixing to inform what kind of light needs to be emitted to produce a given color. The value of a particular color is expressed as a vector of three elements—intensities of three primary colors, recall equation (2.18). A transformation to a different color space is expressed by a transformation by a 3×3 matrix. Assume that values for each primary are quantized to $m = 2^n$ values; let the highest intensity value be $k = m - 1$; then $(0, 0, 0)$ is black, (k, k, k) is (television) white, $(k, 0, 0)$ is 'pure' red, and so on. The value $k = 255 = 2^8 - 1$ is common, i.e., 8 bits per color channel. There are $256^3 = 2^{24} = 16,777,216$ possible colors in such a discretized space.

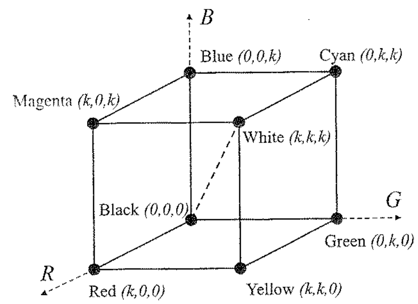


Figure 2.32: RGB color space with primary colors red, green, blue and secondary colors yellow, cyan, magenta. Gray-scale images with all intensities lie along the dashed line connecting black and white colors in RGB color space.

The RGB model may be thought of as a 3D co-ordinatization of color space (see Figure 2.32); note the secondary colors which are combinations of two pure primaries. There are specific instances of the RGB color model as sRGB, Adobe RGB and Adobe Wide Gamut RGB. They differ slightly in transformation matrices and the gamut. One of transformations between RGB and XYZ color spaces is

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 3.24 & -1.54 & -0.50 \\ -0.98 & 1.88 & 0.04 \\ 0.06 & -0.20 & 1.06 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} ,$$

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.41 & 0.36 & 0.18 \\ 0.21 & 0.72 & 0.07 \\ 0.02 & 0.12 & 0.95 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} . \quad (2.19)$$

The US and Japanese color television formerly used **YIQ** color space. The Y additive color mixing. This system stores a luminance value with two chrominance values, corresponding approximately to the amounts of blue and red in the color. This color space corresponds closely to the YUV color model in the PAL television norm (Australia, Europe, except France, which uses SECAM). YIQ color space is rotated 33° with respect to the YUV color space. The YIQ color model is useful since the Y component provides all that is necessary for a monochrome display; further, it exploits advantageous properties of the human visual system, in particular our sensitivity to **luminance**, the perceived energy of a light source.

The **CMY**—for Cyan, Magenta, Yellow—color model uses subtractive color mixing which is used in printing processes. It describes what kind of inks need to be applied so the light reflected from the white substrate (paper, painter's canvas) and passing through the inks produces a given color. CMYK stores ink values for black in addition. Black color can be generated from C, M, Y components but as it is abundant in printed documents, it is of advantage to have a special black ink. Many CMYK colors spaces are used for different sets of inks, substrates, and press characteristics (which change the color transfer function for each ink and thus change the appearance).

HSV – Hue, Saturation, and Value (also known as HSB, hue, saturation, brightness) is often used by painters because it is closer to their thinking and technique. Artists commonly use three to four dozen colors (characterized by the hue; technically, the dominant wavelength). If another color is to be obtained then it is mixed from the given ones, for example, 'purple' or 'orange'. The painter also wants colors of different saturation, e.g., to change 'fire brigade red' to pink. She will mix the 'fire brigade red' with white (and/or black) to obtain the desired lower saturation. The HSV color model is illustrated in Figure 2.33.

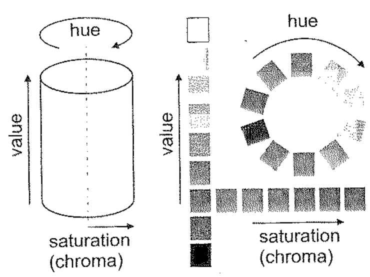


Figure 2.33: HSV color model illustrated as a cylinder and unfolded cylinder. *A color version of this figure may be seen in the color inset—Plate 3.*

HSV decouples intensity information from color, while hue and saturation correspond to human perception, thus making this representation very useful for developing image processing algorithms. This will become clearer as we proceed to describe image enhancement algorithms (for example, equalization Algorithm 5.1), which if applied to each component of an RGB model would corrupt the human sense of color, but which would work more or less as expected if applied to the intensity component of HSV (leaving the color information unaffected). HSL (hue, saturation, lightness/luminance), also known as HLS or HSI (hue, saturation, intensity) is similar to HSV. 'Lightness' replaces 'brightness'.

Models	Color spaces	Applications
Colorimetric	XYZ	Colorimetric calculations
Device oriented, nonuniform spaces	RGB, UIQ	Storage, processing, coding, color TV
Device oriented, Uniform spaces	LAB, LUV	Color difference, analysis
User oriented	HSL, HSI	Color perception, computer graphics

2.4.4 Palette images

Palette images (called also **indexed images**) provide a simple way to reduce the amount of data needed to represent an image. The pixel values constitute a link to a **lookup table** (also called a color table, color map, index register, **palette**). The lookup table contains as many entries as the range of possible values in the pixel, which is typically 8 bits ≡ 256 values. Each entry of the table maps the pixel value to the color, so there are three values, one for each of three color components. In the typical case of the RGB color model, values for red, green and blue are provided. It is easy to see that this approach would reduce data consumption to one-third if each of the RGB channels had been using 8 bits (plus size of the look up table). Many widely used image formats for raster images such as TIFF, PNG and GIF can store palette images.

If the number of colors in the input image is less than or equal to the number of entries in the lookup table then all colors can be selected and no loss of information occurs. Such images may be cartoon movies, or program outputs. In the more common case, the number of colors in the image exceeds the number of entries in the lookup table and a subset of colors has to be chosen, and a loss of information occurs.

This color selection may be done many ways. The simplest is to quantize color space regularly into cubes of the same size. In the 8 bit example, there would be $8 \times 8 \times 8 = 256$ such cubes. If there is, e.g., a green frog in green grass in the picture then there will not be enough shades of green available in the lookup table to display the image well. In such a case, it is better to check which colors appear in the image by creating histograms for all three color components and quantize them to provide more shades for colors which occur in the image frequently. If an image is converted to a palette representation then the nearest color (in some metric sense) in the lookup table is used to represent the color. This is an instance of **vector quantization** (see Section 14.4) which is widely used in analyzing large multi-dimensional datasets. It is also possible to view the occupation by the pixels of RGB space as a **cluster analysis** problem (see Section 9.2.5), susceptible to algorithms such as k-means (Algorithm 9.4).

The term **pseudocolor** is usually used when an original image is gray-level and is displayed in color; this is often done to exploit the color discriminatory power of human vision. The same palette machinery as described above is used for this purpose; a palette is loaded into the lookup table which visualizes the particular gray-scale image the best. It could either enhance local changes, or might provide various views of the image. Which palette to choose depends on the semantics of the image and cannot be derived from image statistics alone. This selection is an interactive process.

2.4.5 Color constancy

Consider the situation in which the same surface is seen under different illumination, e.g., for a Rubik's cube in Figure 2.34. The same surface colors are shown fully illuminated and in shadow. The human vision system is able to abstract to a certain degree from the illumination changes and perceive several instances of a particular color as the same. This phenomenon is called color constancy. Of course, it would be desirable to equip artificial perception systems based on photosensors with this ability too, but this is very challenging.

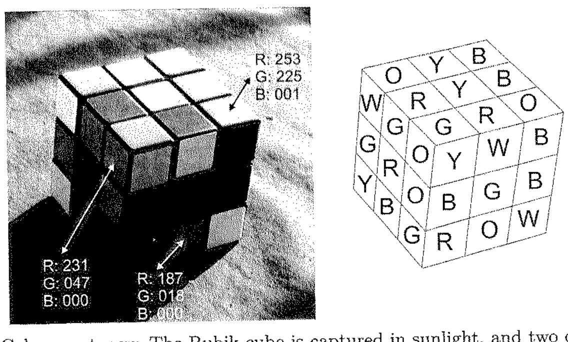


Figure 2.34: Color constancy: The Rubik cube is captured in sunlight, and two of three visible sides of the cube are in shadow. The white balance was set in the shadow area. There are six colors on the cube: R-red, G-green, B-blue, O-orange, W-white, and Y-yellow. The assignment of the six available colors to 3 × 9 visible color patches is shown on the right. Notice how different the same color patch can be: see RGB values for the three instances of orange. *A color version of this figure may be seen in the color inset—Plate 4.*

Recall equation (2.17) which models the spectral response q_i of the i-th sensor by integration over a range of wavelengths as a multiplication of three factors: spectral sensitivity $R_i(\lambda)$ of the sensor $i = 1, 2, 3$, spectral density of the illumination $I(\lambda)$, and surface reflectance $S(\lambda)$. A color vision system has to calculate the vector q_i for each pixel as if $I(\lambda) = 1$. Unfortunately, the spectrum of the illuminant $I(\lambda)$ is usually unknown.

Assume for a while the ideal case in which the spectrum $I(\lambda)$ of the illuminant is known. Color constancy could be obtained by dividing the output of each sensor with its sensitivity to the illumination. Let q_i' be the spectral response after compensation for the illuminant (called von Kries coefficients), $q_i' = \rho_i \, q_i$, where

$$\rho_i = 1 \Big/ \int_{\lambda_1}^{\lambda_2} I(\lambda)\, R_i(\lambda)\, \mathrm{d}\lambda\,. \tag{2.20}$$

Partial color constancy can be obtained by multiplying color responses of the three photosensors with von Kries coefficients ρ_i.

In practice, there are several obstacles that make this procedure intractable. First, the

q_i of the i-th sensor. Clearly the color constancy problem is ill-posed and cannot be solved without making additional assumptions about the scene.

Several such assumptions have been suggested in the literature. It can be assumed that the average color of the image is gray. In such a case, it is possible to scale the sensitivity of each sensor type until the assumption becomes true. This will result in an insensitivity to the color of the illumination. This type of color compensation is often used in automatic white balancing in video cameras.

Another common assumption is that the brightest point in the image has the color of the illumination. This is true when the scene contains specular reflections which have the property that the illuminant is reflected without being transformed by the surface patch.

The problem of color constancy is further complicated by the perceptual abilities of the human visual system. Humans have quite poor quantitative color memory, and also perform color adaptation. The same color is sensed differently in different local contexts.

2.5 Cameras: an overview

2.5.1 Photosensitive sensors

Photosensitive sensors most commonly found in cameras can be divided into two groups:

Sensors based on photo-emission principles explore the photoelectric effect. An external photon carried in incoming radiation brings enough energy to provoke the emission of a free electron. This phenomenon is exhibited most strongly in metals. In image analysis related applications, it has been used in photomultipliers and vacuum tube TV cameras.

Sensors based on photovoltaic principles became widely used with the development of semiconductors. The energy of a photon causes an electron to leave its valence band and changes to a conduction band. The quantity of incoming photons affects macroscopic conductivity. The excited electron is a source of electric voltage which manifests as electric current; the current is directly proportional to the amount of incoming energy (photons). This phenomenon is explored in several technological elements as a photodiode, an avalanche photodiode (an amplifier of light which has similar behavior from the user's point of view as the photomultiplier; it also amplifies noise and is used, e.g., in night vision cameras), a photoresistor, and Schottky photodiode.

There are two types of semiconductor photoresistive sensors used widely in cameras: CCDs (charge-coupled devices) and CMOS (complementary metal oxide semiconductor). Both technologies were developed in laboratories in the 1960s and 1970s. CCDs became technologically mature in the 1970s and became the most widely used photosensors in cameras. CMOS technology started being technologically mastered from about the 1990s. At the time of writing (2006) neither of these technologies is categorically superior to the other. The outlook for both technologies is good.

In a CCD sensor, every pixel's charge is transferred through just one output node to be converted to voltage, buffered, and sent off-chip as an analog signal. All of the pixel area can be devoted to light capture. In a CMOS sensor, each pixel has its own