Chapter 10

Color

As discussed elsewhere in this book, light has intensity and images have gray value; but light consists of a spectrum of wavelengths and images can include samples from multiple wavelengths called channels. Perceived color depends on three factors:

- Spectral reflectance of scene surfaces, which determines how surfaces reflect color
- Spectral content of ambient illumination, which is the color content of the light shining on surfaces
- Spectral response of the sensors in the imaging system

In this chapter, it is assumed that surfaces are opaque so that the position of a point (x, y, z) on a scene surface can be represented by the image plane coordinates (x', y') for the line of sight to the scene point. This viewer-centered coordinate system will be used consistently throughout this chapter, so the primes on image plane coordinates will be omitted.

10.1 Color Physics

As with shading, presented in Chapter 9, ambient illumination is reflected from surfaces onto the image plane; but with color the illumination has some distribution of wavelengths and the surface reflectance is different for each wavelength. For any wavelength λ, the scene radiance is

$$S(x, y, \lambda)E(\lambda), \tag{10.1}$$

where S is the spectral reflectance of the scene point and E is the spectral distribution of the irradiance on the scene point. In shading, it is assumed that image irradiance is equal to scene radiance since the effect of the imaging system on the image irradiance is a constant that is included in the albedo; but with color the imaging system will filter the scene radiance through one or more sensors, and each sensor has a different spectral response. The image irradiance for sensor k is

$$\rho_k(x, y) = \int_0^\infty R_k(\lambda)S(x, y, \lambda)E(\lambda)\, d\lambda. \tag{10.2}$$

It is not necessary for the wavelengths to be restricted to visible light. The sensors could respond to energy in the infrared, ultraviolet, or X-ray portions of the electromagnetic spectrum, for example. This is the reason why the radiometric quantities used in the equations presented in Chapter 9 on shading were based on energy rather than visible light.

The imaging system will create multiple images, called channels, one for each sensor in the imaging system. In color imagery, there are usually three sensors with spectral responses that cover the red, green, and blue regions of the visible spectrum (RGB space). The range of all possible colors that can be captured by an imaging system or reproduced on a display is in the positive octant of the RGB space. It is common to work with normalized RGB values, so the set of all possible colors is a unit cube with corners at $(0, 0, 0)$ and $(1, 1, 1)$ in the RGB space.

10.2 Color Terminology

The physics of how colored light is reflected from surfaces is based on spectral distributions, but a practical imaging system, including the human vision system, works with a small number of samples from the distribution of wavelengths. The infinite-dimensional vector space of spectral distributions is reduced to a finite-dimensional vector space of samples (Equation 10.2). For any finite set of spectral response functions $\{R_k(\lambda), k = 1, \ldots, p\}$, there is an infinite number of spectral distributions that are filtered to the same set of response values $\{\rho_k, k = 1, \ldots, p\}$. For example, the spectral distribution at wavelengths where $R_k(\lambda) = 0$ can be varied arbitrarily without affecting the value of the response. For color imagery, the perceptually meaningful differences in spectral distributions are captured by the quantities of hue, saturation, and brightness (luminance).

Hue is determined by the dominant wavelength in the spectral distribution of light wavelengths. The spectral hues are single wavelengths in the visible spectrum. For example, the primary colors (red, green, and blue) are located at specific positions in the visible spectrum. Nonspectral hues are mixtures of wavelengths; for example; purple is a mixture of red and blue.

Saturation is the magnitude of the hue relative to the other wavelengths:

$$S = \frac{s_1}{s_2}, \tag{10.3}$$

where s_1 is the amount of light at the dominant wavelength and s_2 is the amount of light at all wavelengths. For example, a deep red color has saturation close to 1; but as other wavelengths are added, the color approaches the distribution of white light, the proportion of red and hence the saturation is reduced, and the color is desaturated to a shade of pink.

The brightness is a measure of the overall amount of light that passes through all of the spectral response functions. You can think of the brightness as a scale factor that is applied to the entire spectral distribution. The hue is the location of the peak in the spectral distribution (or the location and relative magnitudes of two peaks in the case of nonspectral hues such as purple). The saturation is the height of the peak relative to the entire spectral distribution. The location and shape of the peak in the spectral distribution (hue and saturation) determine the characteristics of light that are normally thought of as color.

10.3 Color Perception

The CIE (Commission Internationale de l'Eclairage—the International Commission on Illumination) color system is based on three spectral curves for the CIE primaries. Colors are specified by the relative amounts of the CIE primaries X, Y, and Z that match a given color. The Y value is luminance, a measure of the amount of light at all wavelengths that corresponds to perceived brightness. The chromaticity values depend on the dominant wavelength and saturation, independent of the luminance:[1]

$$x = \frac{X}{X+Y+Z} \tag{10.4}$$
$$y = \frac{Y}{X+Y+Z} \tag{10.5}$$
$$z = \frac{Z}{X+Y+Z}. \tag{10.6}$$

Since $x + y + z = 1$, only two chromaticity values are needed. Colors are conveniently represented by the x and y chromaticities and luminance Y.

The x and y chromaticities represent the components of color independent of luminance. Two colors, such as dark green and light green, may appear different but may actually have the same relative distribution of wavelengths. If the spectral distribution is scaled by a constant, the color may appear lighter or darker, but the shape of the spectral distribution is not changed and the hue (dominant wavelength) and saturation (relative amount of the dominant wavelength) are not changed.

The perceptually significant chromaticities lie inside the arch-shaped region in Figure 10.1. White light is at the center of the chromaticity diagram. Draw a line from white at position W, through a particular color P, to a position H along the outer boundary of the chromaticity diagram. The hue is H, and the saturation S is the length of WP relative to WH. You can think of a color as an additive mixture of white light and pure spectral hue,

$$P = SH + (1 - S)W, \tag{10.7}$$

where the saturation S controls the relative proportions of white tint and hue.

Hue, saturation, and luminance are encoded in the RGB color values in a way that makes it hard to use hue and saturation in vision algorithms. For example, it may be easy to identify objects with different hues by setting thresholds on the range of hues (spectral wavelengths) that bracket the objects. But where are these thresholds in the RGB cube, what is the shape of

[1] Technically, the chromaticities are calculated from luminance and the other CIE values, but the chromaticities are normalized and used in a way that describes color independent of, or at least conceptually separate from, luminance [81].

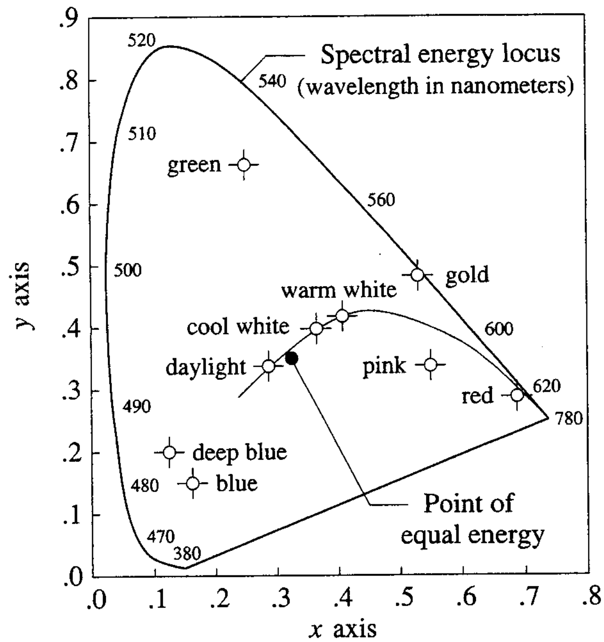

Figure 10.1: The diagram of CIE chromaticities. White light is at the center of the arch-shaped region. Fully saturated colors are along the outer edge of the diagram. The hue for a specific color is obtained by extending a line from white to the edge of the diagram passing through the color.

the surfaces that divide the color regions corresponding to different objects, and what is the formula for applying the thresholds to the RGB values in an image? These questions are hard to answer in the RGB color space but become simple when the RGB values are converted to hue, saturation, and luminance.

10.4 Color Processing

The HSI color model represents a color in terms of hue, saturation, and intensity. The intensity is the gray level of the pixels in a monochrome (black and white) image, such as the images used as examples for the machine vision algorithms presented in other chapters of this book.

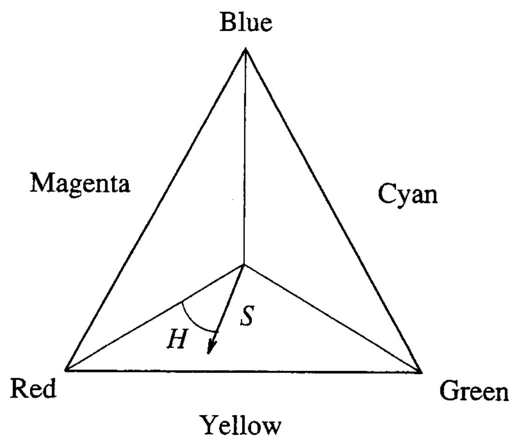

Figure 10.2: The HSI triangle represents the combinations of colors that can be represented by a linear combination of the primary colors at the corners of the triangle.

The HSI color triangle, shown in Figure 10.2, represents the combinations of hue and saturation that can be represented by combinations of the RGB primaries. The corners of the triangle correspond to the maximum amounts of each primary color (red, green, and blue) available from the imaging system. Achromatic (colorless) pixels are shades of gray, corresponding to equal amounts of the primary colors, and lie at the center of the HSI triangle.

The HSI solid adds the dimension of image intensity (Figure 10.3), with black at the bottom of the solid and white at the top. Shades of gray run along the axis of the solid. Each cross section of the solid is an HSI triangle with the proportions of the primary colors constrained to produce a particular intensity value. The HSI solid narrows to a point at the top and bottom because white and black can only be represented by unique combinations of the RGB primaries.

The RGB components of an image can be converted to the HSI color representation. Assume that the RGB components have been normalized to 1. This allows the derivation to be done in device-independent units. The intensity is the sum of the RGB values normalized to 1:

$$I = \frac{1}{3}(R + G + B). \tag{10.8}$$

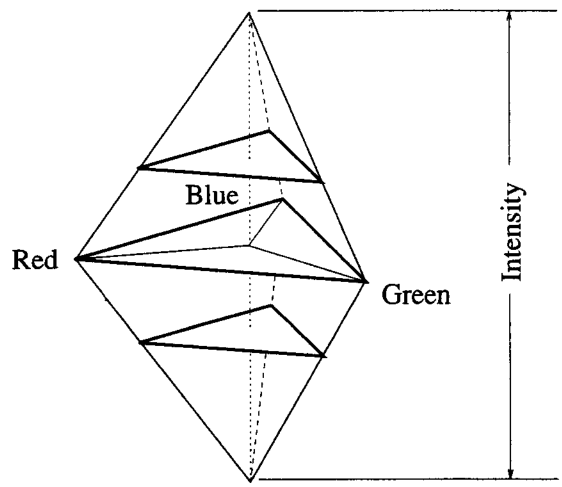

Figure 10.3: The HSI solid bounds the range of colors that can be represented by combinations of the primary colors.

The derivation of the formulas for hue and saturation begins by removing intensity from the RGB values:

$$r = \frac{R}{R+G+B} \tag{10.9}$$

$$g = \frac{G}{R+G+B} \tag{10.10}$$

$$b = \frac{B}{R+G+B}. \tag{10.11}$$

The locus of all possible values for r, g, and b is the triangle in the positive octant of rgb-space with corners $(1,0,0)$, $(0,1,0)$, and $(0,0,1)$. Let point P on the triangle denote the position in rgb-space of some color. Let $p = (r,g,b)$ be the vector to point P on the triangle, w be the vector to the point at the center of the triangle that represents white, and p_r be the vector to the corner of the triangle that corresponds to fully saturated red. The hue is the angle from vector $p_r - w$ to vector $p - w$, counterclockwise around the triangle as seen from the side of the triangle away from the origin. This is the right-hand rule with the thumb aligned along the normal to the triangle pointing away from the origin. The cosine of the hue is

$$\cos H = \frac{(p - w) \cdot (p_r - w)}{\|p - w\| \|p_r - w\|}. \tag{10.12}$$

Vector w to the center of the triangle is $(1/3, 1/3, 1/3)$. The magnitude of vector $p - w$ is

$$\|p - w\| = \sqrt{(r - 1/3)^2 + (g - 1/3)^2 + (b - 1/3)^2}. \tag{10.13}$$

Since $p_r = (1, 0, 0)$ and $w = (1/3, 1/3, 1/3)$, the magnitude of $p_r - w$ is

$$\|p_r - w\| = \sqrt{2/3}. \tag{10.14}$$

The dot product between $p - w$ and $p_r - w$ is

$$(p - w) \cdot (p_r - w) = \frac{2(r - 1/3) - (g - 1/3) - (b - 1/3)}{3}. \tag{10.15}$$

Divide this dot product by $\|p - w\|$ and $\|p_r - w\|$, substitute Equations 10.9 through 10.11, and simplify to get the formula for computing the hue from the R, G, and B values:

$$\cos H = \frac{2R - G - B}{2\sqrt{(R - G)^2 + (R - B)(G - B)}}. \tag{10.16}$$

In order to have the value for hue in the range from 0 to 360 degrees, it is necessary to subtract H from 360 when $B/I > G/I$. Even though this derivation began with normalized RGB values, the formula in Equation 10.16 will work with RGB values on any scale since the scale factors in the numerator and denominator will cancel.

The saturation is the distance on the triangle in the rgb-space from white relative to the distance from white to the fully saturated color with the same hue. Fully saturated colors are on the edges of the triangle. Let d_p be the distance from point W at the center of the triangle to point P for some color, and let d_q be the distance to point Q at the edge of the triangle along the line from W to Q passing through P. The saturation is the ratio of the distances, d_p/d_q. The formula for the saturation is

$$S = 1 - \frac{3}{R + G + B} \min(R, G, B). \tag{10.17}$$

The derivation is provided by Gonzalez and Woods [91].

Equations 10.16, 10.17, and 10.18 can be used to convert the RGB image from a color image acquisition system to the HSI representation for further processing. The hue is not defined when the saturation is zero, that is, for any colors along the axis of the HSI solid. Saturation is not defined when the intensity $I = 0$.

The transformation from RGB to HSI is used to convert color images into a form appropriate for machine vision. Gray-level algorithms can be performed on the I component of the HSI representation. Segmentation can be performed on the H component to discriminate between objects with different hues. However, hue is not reliable for discrimination when the saturation is low, so the segmentation algorithms must be modified to leave pixels with low saturation unassigned. Region growing algorithms can use thresholds on hue to form core regions, leaving unassigned any pixels that have low saturation or are outside the threshold boundaries. The algorithms for growing core regions by assigning neighboring pixels are unchanged.

More general algorithms can divide the HSI solid into regions using thresholds on hue, saturation, and intensity. These thresholds are easier to formulate and apply in the HSI representation than in the RGB representation provided by the imaging system.

10.5 Color Constancy

The color content of outdoor light varies considerably, yet people are able to correctly perceive the colors of objects in the scene independent, for the most part, from the color of the ambient illumination. This phenomenon is called *color constancy.*

Ambient light has the spectral distribution $E(\lambda)$, which describes the power at each wavelength. Assume that scene surfaces are opaque, so that scene coordinates can be specified using the coordinates of the corresponding point (x, y) on the image plane. The fraction of light at wavelength λ reflected from the surface point at location (x, y) is $S(x, y, \lambda)$. The light arriving at each location in the image is determined by the spectral distribution of the ambient light that falls on the scene surfaces and the fraction of light reflected at various wavelengths:

$$S(x, y, \lambda)E(\lambda). \tag{10.18}$$

Assume that there are p sensors at each image location (x, y) and each sensor has a different spectral response function. The spectral response of sensor k is $R_k(\lambda)$. Each sensor at location (x, y) samples a different distribution of light:

$$\rho_k(x, y) = \int_0^\infty R_k(\lambda) S(x, y, \lambda) E(\lambda)\, d\lambda. \tag{10.19}$$

The information about the color (spectral reflectance) of a point on the scene surface corresponding to location (x, y) in the image plane is encoded in the values $\rho_1(x, y), \rho_2(x, y), \ldots, \rho_p(x, y)$ obtained from the p sensors at that location. Color constancy is the problem of recovering the spectral reflectance $S(x, y, \lambda)$ of scene surfaces from the sensor responses $\{\rho_k(x, y), k = 1, \ldots, p\}$ independent of the spectral distribution of the ambient light $E(\lambda)$.

In principle, the color constancy problem can be formulated as an inverse problem in a finite-dimensional linear space and solved using matrix methods. Assume that the surface reflectance is a linear combination of basis functions,

$$S(x, y, \lambda) = \sum_{i=1}^{n} \sigma_i(x, y) S_i(\lambda). \tag{10.20}$$

The number n of basis functions is the number of degrees of freedom in the surface reflectance. Assume that the basis functions $S_i(\lambda)$ are known. Linear models with as few as three basis functions may be sufficient to account for typical surface reflectances.

Represent the ambient light by a linear model with m degrees of freedom:

$$E(\lambda) = \sum_{j=1}^{m} \epsilon_j E_j(\lambda). \tag{10.21}$$

Assume that the spectral distributions $E_j(\lambda)$ of the basis lights are known. Only three or four basis lights are needed to model natural daylight under a wide variety of weather conditions and times of day.

The color determination problem can be represented in matrix notation. Combine the m values of ϵ_j into a column vector ϵ, and combine the n values of σ_i into a column vector σ. Substitute the column vectors into Equation 10.19 to yield a matrix model for each pixel in the image:

$$\rho = \Lambda_\epsilon \sigma. \tag{10.22}$$

The lighting matrix Λ_ϵ is a p by n matrix, and the ij entry is

$$\int_0^\infty R_i(\lambda) S_j(\lambda) E(\lambda)\, d\lambda. \tag{10.23}$$

If the ambient illumination is known, then the lighting matrix Λ_ϵ is known. If the number of sensors equals the number of degrees of freedom in the surface reflectivity, $p = n$, then the lighting matrix can be inverted to obtain the coefficients of the surface spectral reflectivity σ at each point in the image which characterizes the color of the corresponding points on the scene surfaces.

If the ambient illumination is not known, then solving the problem will require more sensors than the number of degrees of freedom in surface reflectance. Since it was assumed that the ambient illumination is the same for all points in the scene, the information at multiple scene points can be used to estimate the ambient illumination. Suppose $p = n+1$. From s different spatial locations, $sp = s(n+1)$ different measurements can be obtained. There are sn unknowns for the surface reflectance and m unknowns for the ambient light spectrum. It is necessary to sample at $s > m$ locations to have more data than unknowns. This analysis suggests that the problem of color constancy can be solved without knowing the ambient illumination if there are $n+1$ sensors.

The matrix Λ_ϵ maps an n-dimensional surface space into an $(n+1)$-dimensional sensor space. For example, if $p = 3$ and $n = 2$, then the subspace is a plane in a three-dimensional space. This suggests the following two-step algorithm for determining surface reflectivity independent of scene illumination:

1. Determine the plane (subspace) spanning the data points in the sensor space to recover the ambient light vector ϵ.

2. Determine the lighting matrix Λ_ϵ from the ambient light vector ϵ and invert it to obtain the surface reflectivity vector σ.

10.6 Discussion

This chapter has barely scratched the surface of the topic of color vision. The intention has been to introduce the minimum number of formulas and concepts required to generalize the material on shading (Chapter 9) to cover colored light, reflectance, and imaging. The brief presentation on color terminology and the CIE color model provides an introduction to color theory sufficient to allow the reader to pursue the topic in other sources. The presentation of the HSI color model gives the reader some glimpse of the transformations that can be applied to multispectral imagery to allow the segmentation and edge detection algorithms covered elsewhere in this text to be used. The key concept in segmenting multispectral imagery is to find a transformation that reduces the dimensionality (number of channels) and makes it easy to use thresholds to determine core regions.

Image processing algorithms have been developed for correcting differences in the color capabilities of various displays and printers, and realistic graphics renderings require approximations to the spectral distributions for ambient illumination and spectral reflectivity. There are many image enhancement techniques that use color to highlight interesting features in images. Since this text is not concerned with the generation, reproduction, or enhancement of images, many interesting topics in color imagery have been omitted; the reader is encouraged to pursue the topic in the many excellent texts on image processing and computer graphics that are widely available.

Further Reading

There are many sources in computer graphics that provide very readable accounts of color theory [81]. Image processing texts provide extensive coverage of color representations, including the CMY color model used in color printing and the YIQ color model used in broadcast television. The HSI color model and the derivation of the formulas for converting RGB values to HSI values was adapted from Gonzalez and Woods [91]. The material on color constancy was adapted from the work of Maloney and Wandell (see Wandell [246]).

Exercises

10.1 What determines color at a point in an image? What is the role of the illumination in color images?

10.2 How is a color image represented in an image? Why are the particular frequencies of the electromagnetic spectrum selected for use in color vision? Name the basic colors used in defining characteristics at a point.

10.3 Define Hue, Saturation, and Brightness. Which of these is important in characterizing color at a point? Why?

10.4 How can you compute HSI characteristics from RGB characteristics? Why would you want to convert data in one representation to the other?

10.5 In machine vision, which color representation is more suitable? Justify your answer.

10.6 Design an edge detector that will detect prominent edges in a color image. Apply this to edge images. How will you combine the outputs of the different channels to provide you with edges in color images?

10.7 What are subtractive and additive models of color? Which one will be more suitable for displays, photos, and printers? Why?

10.8 What is the HSI solid? Where and how is it used in color processing?

10.9 Define and explain color constancy. Can machine vision systems display color constancey?

10.10 Why has color not been used much in machine vision? Do you think its application will increase? If so, what will be the leading application?