

Computer Vision

Course 7

The Hit-or-Miss Transformation (nimerește sau ratează)

The morphological hit-or-miss transformation is a basic tool for shape detection. Consider the set A from Figure 9.12 consisting of three shapes (subsets) denoted C , D , and E . The objective is to locate one of the shapes, say, D .

Let the origin of each shape be located at its center of gravity. Let D be enclosed by a small window, W . The *local background* of D with respect to W is defined as the set difference $(W-D)$ (Figure 9.12(b)). Figure 9.12(c) shows the complement of A . Fig. 9.12(d)

shows the erosion of A by D . Figure 9.12(e) shows the erosion of the complement of A by the local background set ($W-D$).

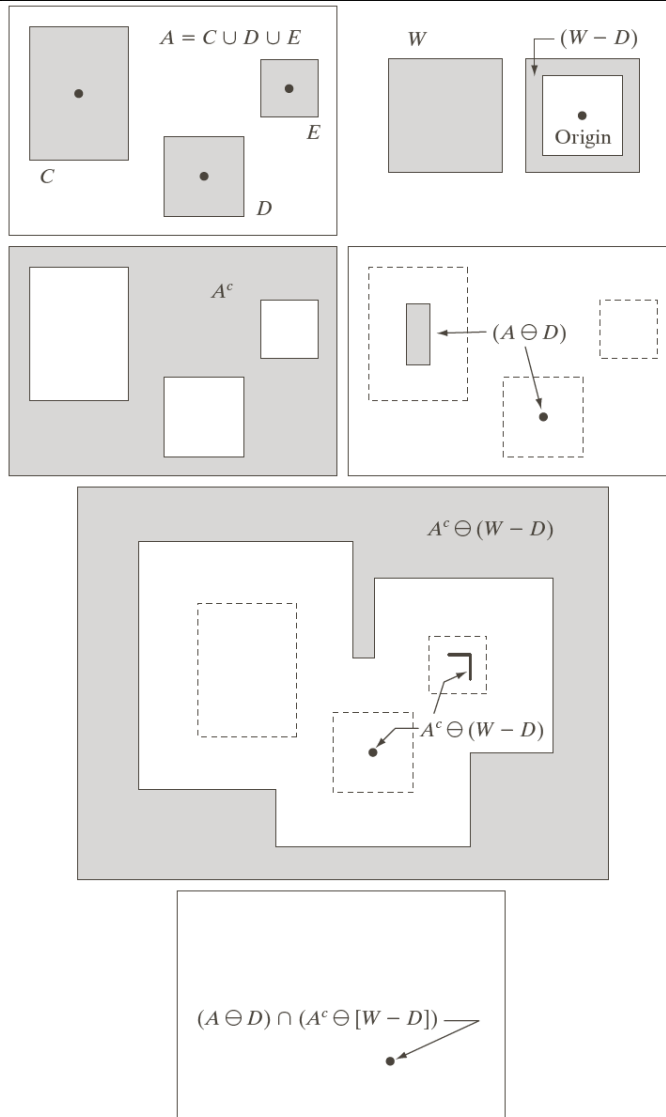
From Figures 9.12(d) and (e) we can see that the set of location for which D exactly fits inside A is the intersection of the erosion of A by D and the erosion of A^c by ($W-D$) as shown in Figure 9.12(f).

If B denotes the set composed of D and its background, the match (or the set of matches) of B in A , denoted $A \Upsilon B$ is:

$$A \Upsilon B = (A \odot D) \cap [A^c \odot (W - D)]$$

Computer Vision

Course 7



a b
c d
e
f

FIGURE 9.12
 (a) Set A . (b) A window, W , and the local background of D with respect to W , $(W - D)$. (c) Complement of A . (d) Erosion of A by D . (e) Erosion of A^c by $(W - D)$. (f) Intersection of (d) and (e), showing the location of the origin of D , as desired. The dots indicate the origins of C , D , and E .

We can generalize the notation by letting $\mathbf{B} = (\mathbf{B}_1, \mathbf{B}_2)$, where \mathbf{B}_1 is the set formed from elements of \mathbf{B} associated with an object and \mathbf{B}_2 is the set of elements of \mathbf{B} associated with the corresponding background ($\mathbf{B}_1=D, \mathbf{B}_2=W-D$) in the preceding example).

$$\mathbf{A} \Upsilon \mathbf{B} = (\mathbf{A} \odot \mathbf{B}_1) \cap [\mathbf{A}^c \odot \mathbf{B}_2]$$

The set $\mathbf{A} \Upsilon \mathbf{B}$ contains all the (origin) points at which, simultaneously, \mathbf{B}_1 found a match (“hit”) in \mathbf{A} and \mathbf{B}_2 found a match in \mathbf{A}^c . Taking into account the definition and properties of erosion we can rewrite the above relation as:

$$A \Upsilon B = (A \odot B_1) - (A \oplus \bar{B}_2)$$

The above three equations for $A \Upsilon B$ are referred as the *morphological hit-or-miss transform*.

Some Basic Morphological Algorithms

When dealing with binary images, one of the principal applications of morphology is in extracting image components that are useful in the representation and the description of shape. We consider morphological algorithms for extracting boundaries, connected components, the convex hull, and the skeleton of a region.

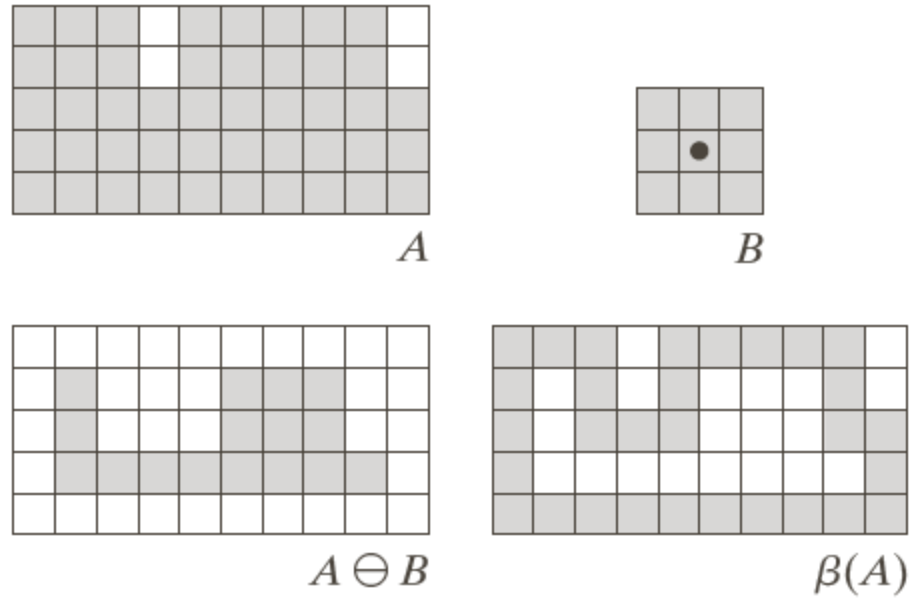
The images are shown graphically with 1s shaded and 0s in white.

Boundary Extraction

The boundary of a set A , denoted $\beta(A)$, can be obtained by first eroding A by B and then performing the set difference between A and its erosion.

$$\beta(A) = A - (A \odot B)$$

where B is a suitable structuring element.



a	b
c	d

FIGURE 9.13 (a) Set A . (b) Structuring element B . (c) A eroded by B . (d) Boundary, given by the set difference between A and its erosion.



a b

FIGURE 9.14

(a) A simple binary image, with 1s represented in white. (b) Result of using Eq. (9.5-1) with the structuring element in Fig. 9.13(b).

Filling Holes

A *hole* may be defined as a background region surrounded by a connected border of foreground pixels. We present an algorithm based on set dilation, complementation, and intersection for filling holes in an image.

Let A denote a set whose elements are δ -connected boundaries, each boundary enclosing a background region (i.e. a hole). Given a point in each hole, the objective is to fill all the holes with 1 s.

We form an array, X_0 , of 0 s (the same size as the array containing A), except at the location in X_0 corresponding to the

given point in each hole, which is set to I . The following procedure fills all the holes with I s:

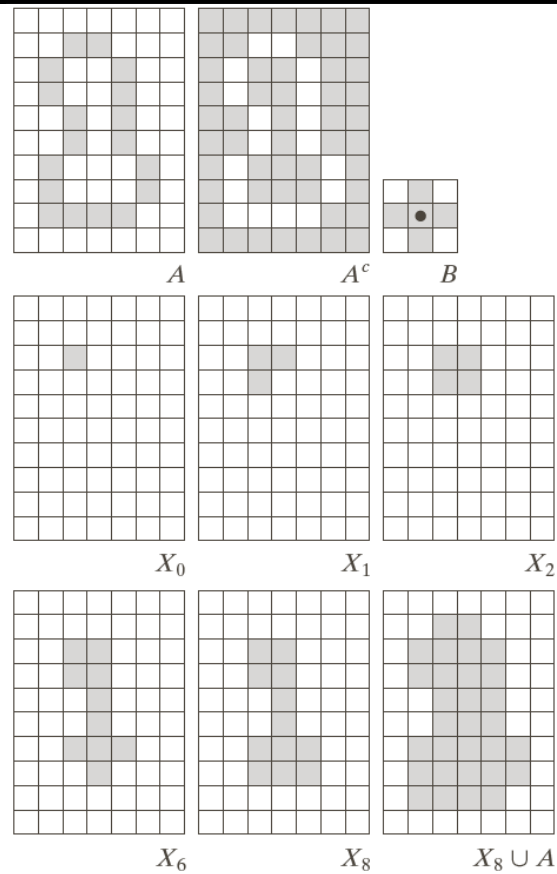
$$X_k = (X_{k-1} \oplus B) \cap A^c \quad , \quad k = 1, 2, 3, \dots$$

where B is the symmetric structuring element in Figure 9.15(c).

The algorithm terminates at iteration step k if $X_k = X_{k-1}$. The set X_k then contains all the filled holes. The set union of X_k and A contains all the filled holes and their boundaries.

Computer Vision

Course 7



a	b	c
d	e	f
g	h	i

FIGURE 9.15 Hole filling. (a) Set A (shown shaded). (b) Complement of A . (c) Structuring element B . (d) Initial point inside the boundary. (e)–(h) Various steps of Eq. (9.5-2). (i) Final result [union of (a) and (h)].

Extraction of Connected Components

Extraction of connected components from binary images is important in many automated image analysis applications.

Let A be a set containing one or more connected components. Form an array X_0 (of the same size as the array containing A) whose elements are 0 s (background values), except at each location known to correspond to a point in each connected component in A , which we set to 1 (foreground value). The objective is to start with X_0 and find all the connected components. The procedure that accomplishes this task is the following:

$$X_k = (X_{k-1} \oplus B) \cap A, \quad k = 1, 2, 3, \dots$$

where B is a suitable structuring element. The procedure terminates when $X_k = X_{k-1}$, with X_k containing all connected components of the input image.

Computer Vision

Course 7

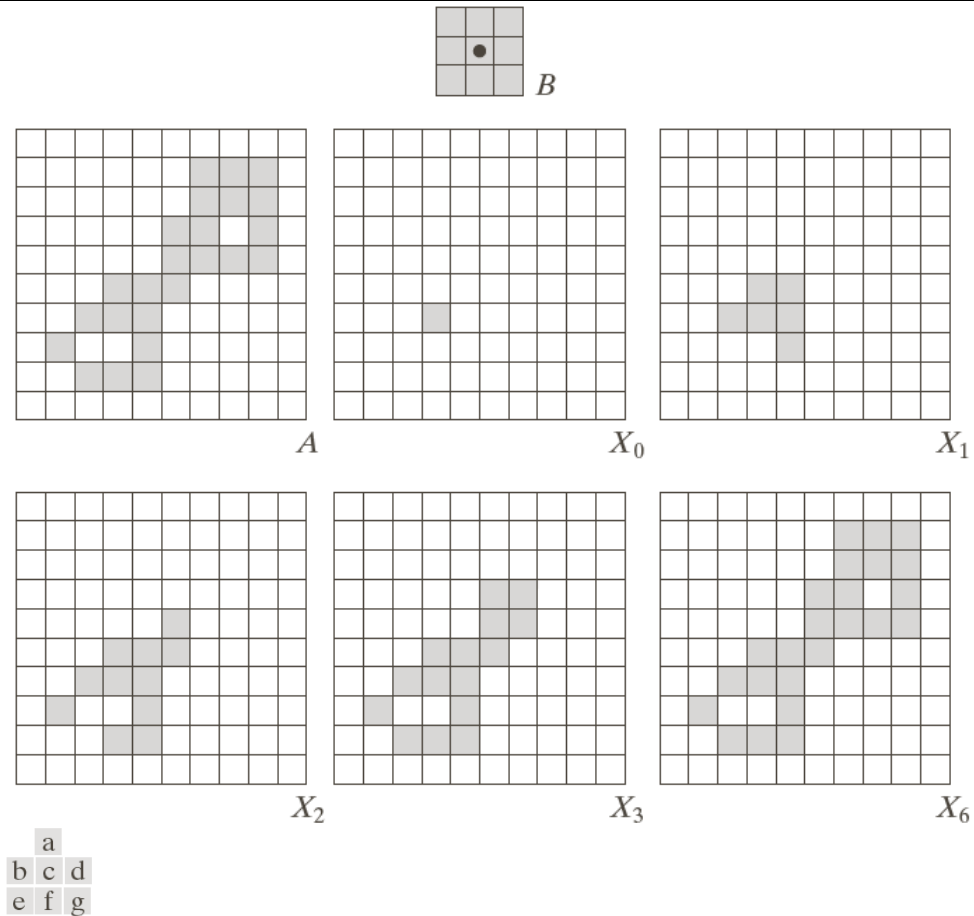
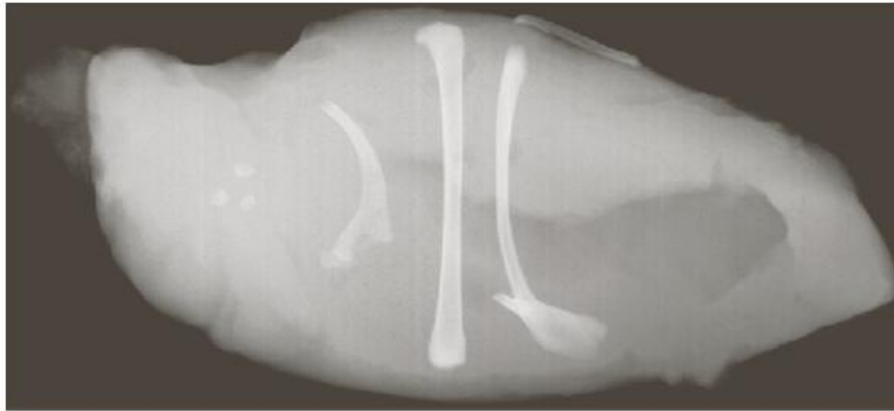


FIGURE 9.17 Extracting connected components. (a) Structuring element. (b) Array containing a set with one connected component. (c) Initial array containing a 1 in the region of the connected component. (d)–(g) Various steps in the iteration of Eq. (9.5-3).

Course 7



a
b
c d

FIGURE 9.18

(a) X-ray image of chicken file with bone fragments.

(b) Thresholded image. (c) Image eroded with a 5×5 structuring element of 1s.

(d) Number of pixels in the connected components of (c).

(Image courtesy of NTB

Elektronische Geraete GmbH, Diepholz, Germany,

www.ntbxray.com.)

Connected component	No. of pixels in connected comp
01	11
02	9
03	9
04	39
05	133
06	1
07	1
08	743
09	7
10	11
11	11
12	9
13	9
14	674
15	85

Figure 9.18(a) shows an *X*-ray image of a chicken breast that contains bone fragment. It is of considerable interest to be able to detect such objects in processed food before packing and/or shipping. In this case, the density of the bones is such that their normal intensity values are different from the background. This makes extraction of the bones from the background a simple matter by using a single threshold. The result is the binary image in Figure 9.18(b). We can erode the thresholded image so that only objects of „significant” size remain. In this example, we define as significant any object that remains after erosion with a **5×5**

structuring element of *1s*. The result of erosion is shown in Figure 9.18(c). The next step is to analyse the objects that remain. We identify these objects by extracting the connected components in the image. There are a total of *15* connected components, with four of them being of dominant size. This is enough to determine that significant undesirable objects are contained in the original image.

Color Image Processing Color Image Processing

Color Image Processing

Color is very important characteristic of an image that in most cases simplifies object identification and extraction from a scene. Human eye can discern thousands of color shades and intensities and only two dozen shades of gray.

Color image processing is divided in 2 major areas: *full-color* (images acquired with a full-color sensor) and *pseudo-color* (gray images for which color is assigned) processing.

<https://youtube.com/shorts/pS5Cvpz7-i4?si=rbWW1pT0Hcsejhmo>

The colors that humans can perceive in an object are determined by the nature of the light reflected from the object. Visible light is composed of a relatively

Computer Vision

Course 7

narrow band of frequencies in the electromagnetic spectrum (390nm to 750nm). A body that reflects light that is balanced in all visible wavelengths appears white to the observer. A body that favors reflectance in a limited range of the visible spectrum exhibits some shades of color.

For example, blue objects reflect light with wavelengths from 450 to 475 nm, while absorbing most of the energy of other wavelengths.

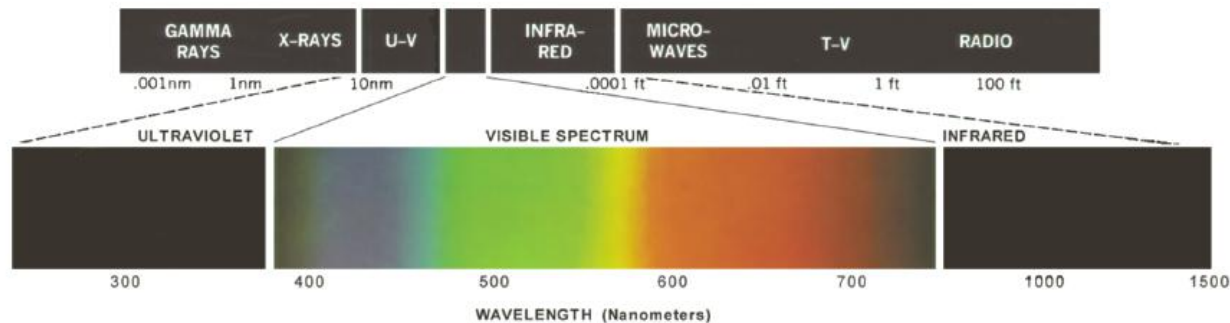


FIGURE 6.2 Wavelengths comprising the visible range of the electromagnetic spectrum. (Courtesy of the General Electric Co., Lamp Business Division.)

How to characterize light? If the light is achromatic (void of color) its only attribute is its *intensity* (or amount) – determined by levels of gray (black-grays-white).

Chromatic light spans the electromagnetic spectrum from approximately 400 to 720 nm. Three basic quantities are used to describe the quality of a chromatic light source: radiance, luminance, and brightness.

- *Radiance* is the total amount of energy that flows from the light source (usually measured in watts).
- *Luminance* (measured in lumens – lm) gives a measure of the amount of energy an observer perceives from a light source. For example, the light emitted from a source operating in the infrared region of the spectrum could

have significant energy (radiance), but an observer would hardly perceive it (the luminance is almost zero).

- *Brightness* is a subjective descriptor that cannot be measured, it embodies the achromatic notion of intensity and is a factor describing color sensation.

Cones are the sensors in the eye responsible for color vision. It has been established that the 6 to 7 million cones in the human eye can be divided into three principal sensing categories, corresponding roughly to red, green, and blue. Approximately 65% of all cones are sensitive to red light, 33% are sensitive to green light, and only about 2% are sensitive to blue (but the blue cones are the most sensitive).

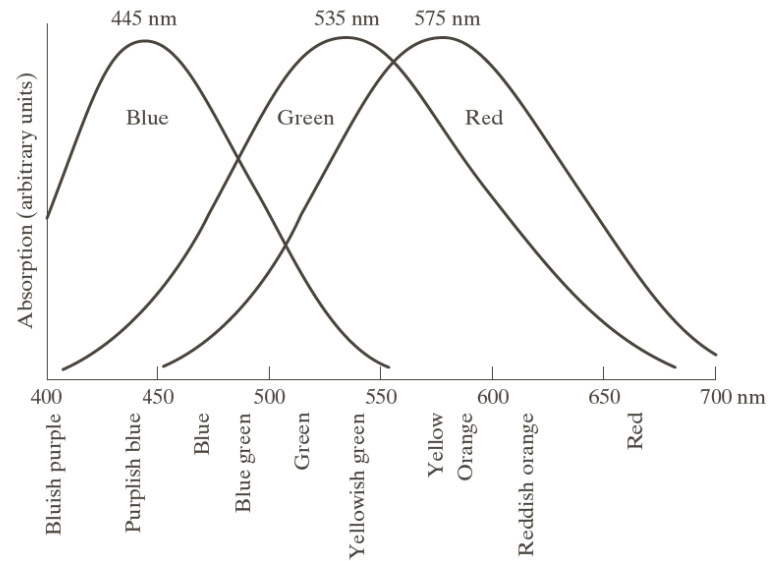


FIGURE 6.3
Absorption of light by the red, green, and blue cones in the human eye as a function of wavelength.

Due to these absorption characteristics of the human eye, colors are seen as variable combinations of the so-called *primary colors*: red (R), green (G), and blue (B).

Computer Vision

Course 7

For the purpose of standardization, the CIE (Commission Internationale de l'Eclairage) designated in 1931 the following specific wavelength values to the three primary colors: blue= 435.8 nm, green = 546.1 nm, and red=700 nm. The CIE standards correspond only approximately with experimental data.

These three standard primary colors, when mixed in various intensity proportions, can produce all visible colors.

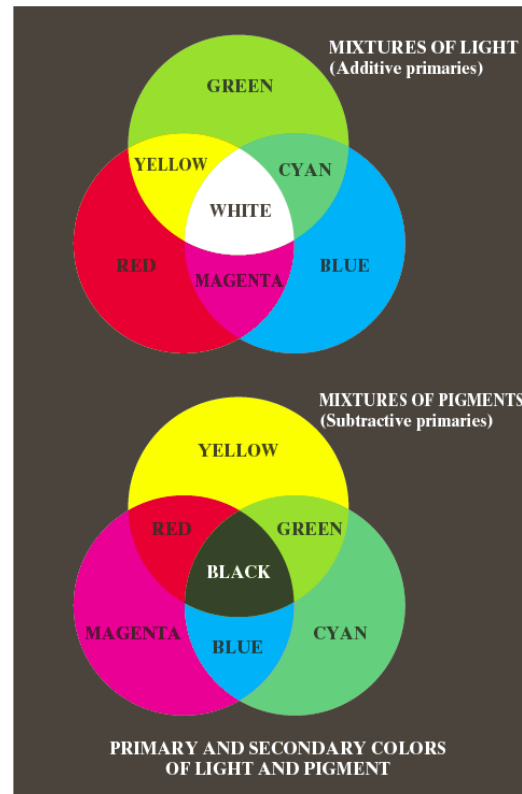
The primary colors can be added to produce the *secondary* colors of light – magenta (red+blue), cyan (green+blue), and yellow (red+green). Mixing the three primaries, or a secondary with its opposite primary color in the right intensities produces white light.

We must differentiate between the primary colors of light and the primary colors of pigments. A primary color for pigments is one that subtracts or absorbs a

Computer Vision

Course 7

primary color of light and reflects or transmits the other two. Therefore, the primary colors of pigments are magenta, cyan, and yellow, and the secondary colors are red, green, and blue.



a
b

FIGURE 6.4
Primary and secondary colors of light and pigments. (Courtesy of the General Electric Co., Lamp Business Division.)

The characteristics usually used to distinguish one color from another are *brightness*, *hue*, and *saturation*.

Brightness embodies the achromatic notion of intensity. *Hue* is an attribute associated with the dominant wavelength in a mixture of light waves. Hue represents dominant color as perceived by an observer (when we call an object to be red, orange or yellow we refer to its hue). *Saturation* refers to the relative purity or the amount of white light mixed with a hue. The pure spectrum colors are fully saturated. Color such as pink (red+white) and lavender (violet+white) are less saturated, with the degree of saturation being inversely proportional to the amount of white light added.

Hue and saturation taken together are called *chromaticity*, and therefore a color may be characterized by its brightness and chromaticity.

Computer Vision

Course 7

The amounts of red, green, and blue needed to form any particular color are called the *tristimulus* values and are denoted X , Y and Z , respectively. A color is specified by its *trichromatic coefficients*, defined as:

$$x = \frac{X}{X + Y + Z}$$

$$y = \frac{Y}{X + Y + Z}$$

$$z = \frac{Z}{X + Y + Z}$$

$$x + y + z = 1$$

For any wavelength of light in the visible spectrum, the tristimulus values needed to produce the color corresponding to that wavelength can be obtained from the existing curves or tables.

Another approach for specifying colors is to use the CIE *chromaticity diagram*, which shows color composition as a function of x (red) and y (green); z (blue) is obtained from relation $z = 1-x-y$.

Computer Vision

Course 7

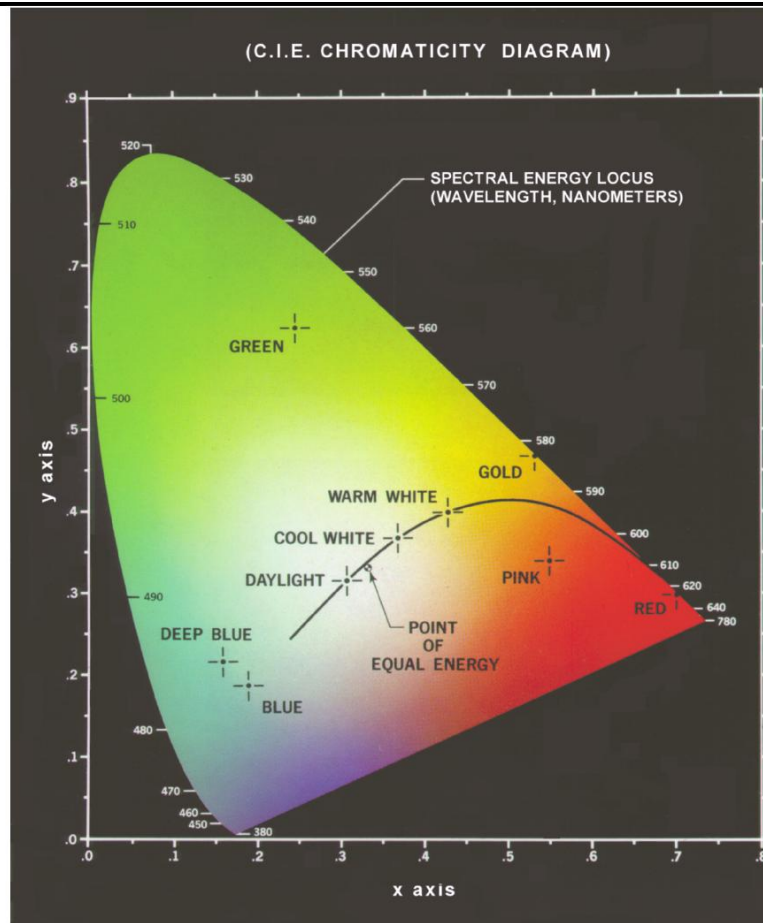


FIGURE 6.5
Chromaticity diagram.
(Courtesy of the
General Electric
Co., Lamp
Business
Division.)

Computer Vision

Course 7

The positions of the various spectrum colors (from violet at 380 nm to red at 780 nm) are indicated around the boundary of the tongue-shaped chromaticity diagram.

The chromaticity diagram is useful for color mixing because a straight-line segment joining any two points in the diagram defines all the different color variation that can be obtained by combining these two colors. This procedure can be extended to three colors: to triangle determined by the three color-points on the diagram embodies all the possible colors that can be obtained by mixing the three colors.

Computer Vision

Course 7

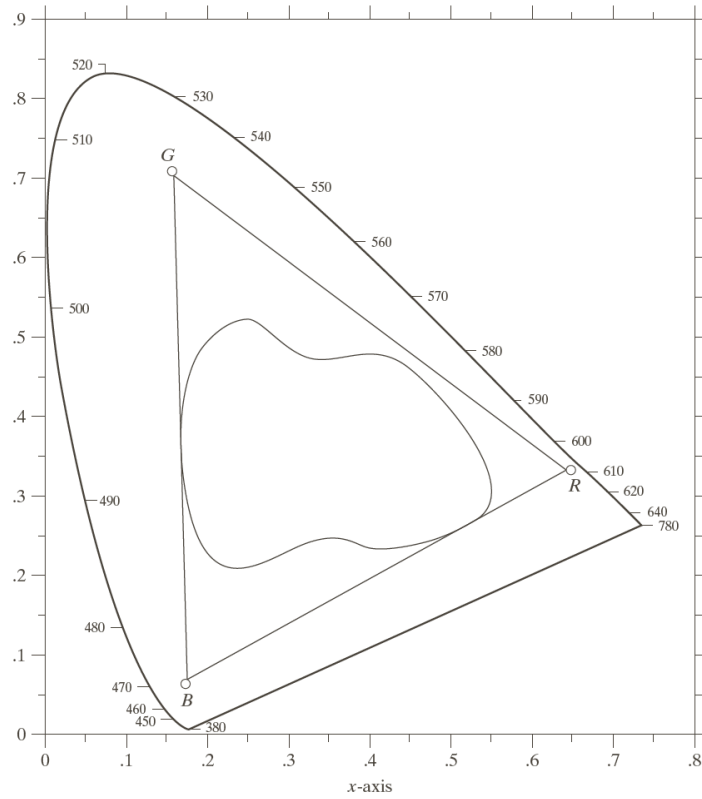


FIGURE 6.6
Typical color gamut of color monitors (triangle) and color printing devices (irregular region).

Color Models

A *color model* (*color space* or *color system*) is a specification of a coordinate system and a subspace within that system where each color is represented by a single point.

<http://www.colorcube.com/articles/models/model.htm>

Most color models in use today are oriented either toward hardware (color monitors or printers) or toward applications where color manipulation is a goal.

The most commonly used hardware-oriented model is **RGB** (red-green-blue) – for color monitors, color video cameras.

The **CMY** (cyan-magenta-yellow) and **CMYK** (cyan-magenta-yellow-black) models are in use for color printing.

The **HSI** (hue-saturation-intensity) model correspond with the way humans describe and interpret colors. The HSI model has the advantage that it decouples the color and gray-scale information in an image, making it suitable for using the gray-scale image processing techniques.

The RGB Color Model

In the RGB model, each color appears decomposed in its primary color components: red, green, blue. This model is based on a Cartesian coordinate system. The color subspace of interest is the unit cube (Figure 6.7), in which the primary and the secondary colors are at the corners; black is at the origin, and white is at the corner farthest from the origin.

Computer Vision

Course 7

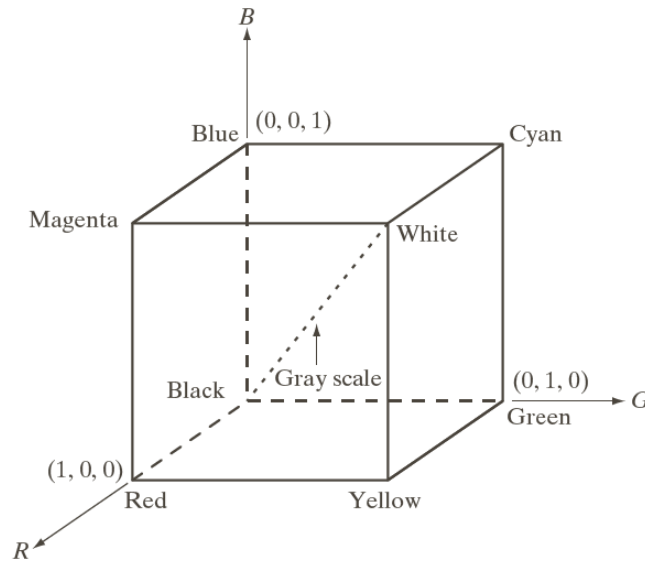


FIGURE 6.7
Schematic of the RGB color cube. Points along the main diagonal have gray values, from black at the origin to white at point $(1, 1, 1)$.

The gray scale (point of equal RGB values) extends from black to white along the line joining these two points. The different colors in this model are points on or inside the cube, and are defined by vectors extending from the origin.

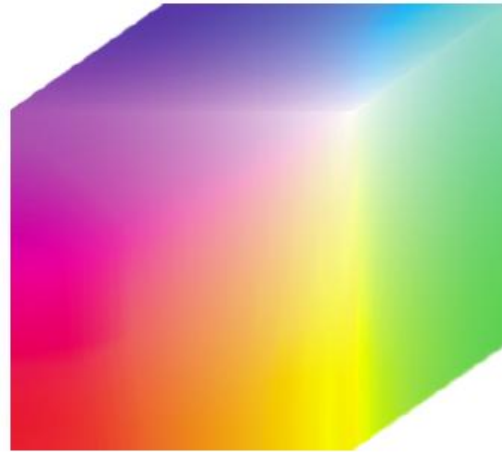


FIGURE 6.8 RGB
24-bit color cube.

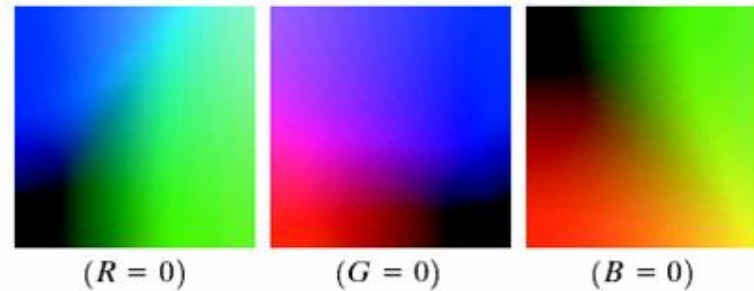
Images represented in the RGB color model consist of three component images, one for each primary color. The number of bits used to represent each pixel in RGB space is called the *pixel depth*. Consider an RGB image in which each of the red, green, and blue images are an 8-bit image. In this case, each RGB color pixel has a depth of 24 bits. The term *full-color* image is used often to denote a 24-bit RGB color image. The total number of colors in a 24-bit RGB image is

Computer Vision

Course 7

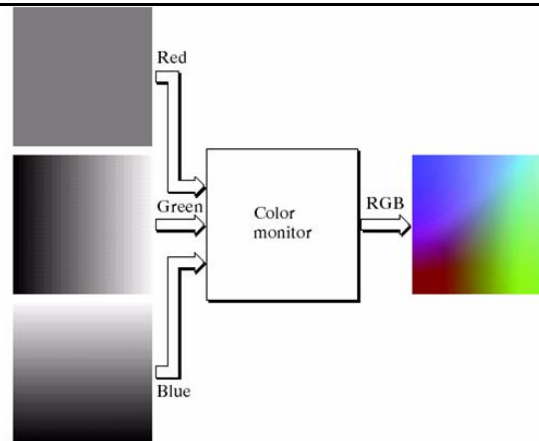
$$(2^8)^3 = 16.777.216$$

A convenient way to view these colors is to generate color planes (faces or cross sections of the cube).



A color image can be acquired by using three filters, sensitive to red, green, and blue.

FIGURE 6.9
Generating
the RGB image of
the cross-sectional
color plane
(127, G , B).



Because of the variety of systems in use, it is of considerable interest to have a subset of colors that are likely to be reproduced faithfully, reasonably independently of viewer hardware capabilities. This subset of colors is called the set of *safe RGB colors*, or the set of *all-systems-safe colors*. In Internet applications, they are called *safe Web colors* or *safe browser colors*.

We assume that 256 colors is the minimum number of colors that can be reproduced faithfully by any system. Forty of these 256 colors are known to be

Computer Vision

Course 7

processed differently by various operating system. We have 216 colors that are common to most systems, and are the safe colors, especially in Internet applications. Each of the 216 safe colors has a RGB representation with:

$$R, G, B \in \{0, 51, 102, 153, 204, 255\}$$

We have $(6)^3=216$ possible color values. It is customary to express these values in the hexagonal number system.

Number System	Color Equivalents					
Hex	00	33	66	99	CC	FF
Decimal	0	51	102	153	204	255

TABLE 6.1

Valid values of each RGB component in a safe color.

Each safe color is formed from three of the two digit hex numbers from the above table. For example purest red is FF0000. The values 000000 and FFFFFFFF represent black and white respectively.

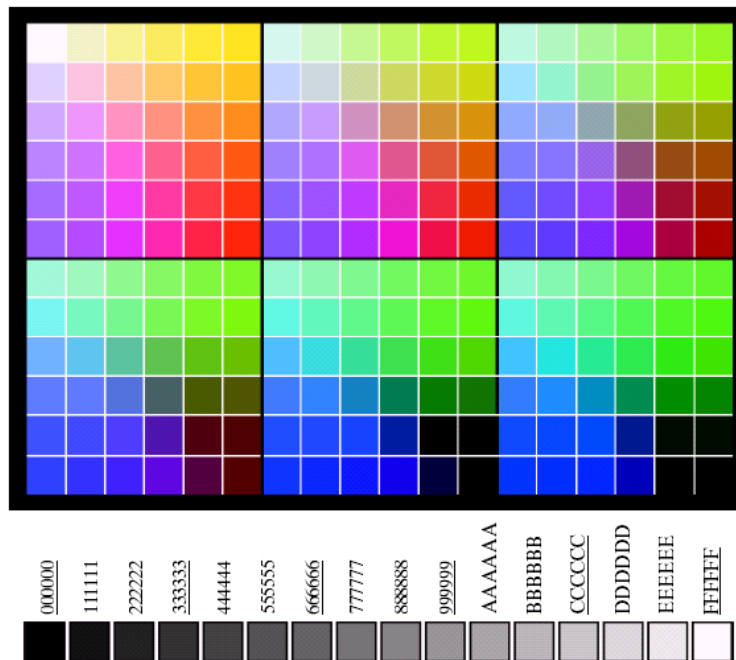
Computer Vision

Course 7

Figure 6.10(a) shows the 216 safe colors, organized in descending RGB values.

Figure 6.10(b) shows the hex codes for all the possible gray colors in the 216 safe color system.

Figure 6.11 shows the RGB safe-color cube.



a
b

FIGURE 6.10

(a) The 216 safe RGB colors.

(b) All the grays in the 256-color RGB system

(grays that are part of the safe color group are shown

underlined).

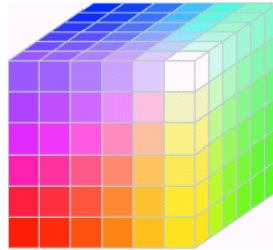


FIGURE 6.11 The RGB safe-color cube.

<http://www.techbomb.com/websafe/>

The CMY and CMYK Color Models

Cyan, magenta, and yellow are the secondary colors of light but the primary color of pigments. For example, when a surface coated with yellow pigment is illuminated with white light, no blue light is reflected from the surface. Yellow subtracts blue light from reflected white light (which is composed of equal amounts of red, green, and blue light).

Most devices that deposit color pigments on paper, such as color printers and copiers, require CMY data input and perform RGB to CMY conversion. Assuming that the color values were normalized to range $[0,1]$, this conversion is:

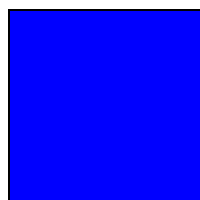
$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

From this equation we can easily deduce, that pure cyan does not reflect red, pure magenta does not reflect green, and pure yellow does not reflect blue.

Equal amount of pigments primary, cyan, magenta, and yellow should produce black. In practice, combining these colors for printing produces a muddy-looking black. In order to produce true black (which is the predominant color in printing), a fourth color, *black*, is added, giving rise to the **CMYK** color model.

Computer Vision

Course 7



The HSI Color Model

The RGB, CMY, and other similar color models are not well suited for describing colors in terms that are practical for human interpretation.

We (humans) describe a color by its hue, saturation and brightness. Hue is a color attribute that describes a pure color, saturation gives a measure of the degree to which a pure color is diluted by white light and brightness is a subjective descriptor that embodies the achromatic notion of intensity.

The HSI (hue, saturation, intensity) color model, decouples the intensity component from the color information (hue and saturation) in a color image.

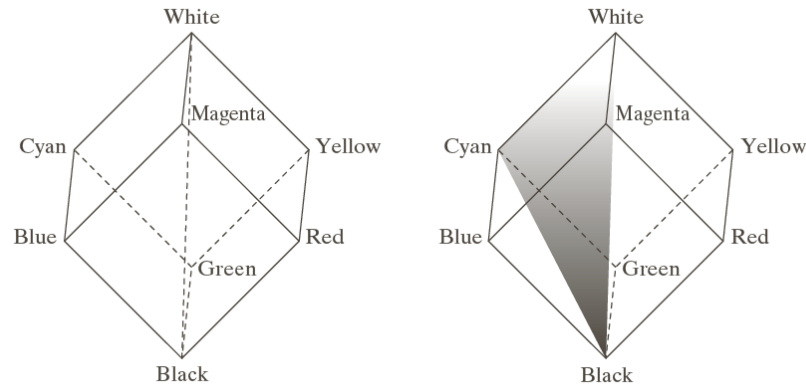
What is the link between the RGB color model and HSI color model? Consider again the RGB unit cube. The intensity axis is the line joining the black and the white vertices. Consider a color point in the RGB cube. Let P be a plane

Computer Vision

Course 7

perpendicular to the intensity axis and containing the color point. The intersection of this plane with the intensity axis gives us the intensity of the color point. The saturation (purity) of the considered color point increases as a function of distance from the intensity axis (the saturation of the point on the intensity axis is zero).

In order to determine how hue can be linked to a given RGB point, consider a plane defined by black, white and cyan. The intensity axis is also included in this plane. The intersection of this plane with the RGB-cube is a triangle. All point contained in this triangle would have the same hue (i.e. cyan).



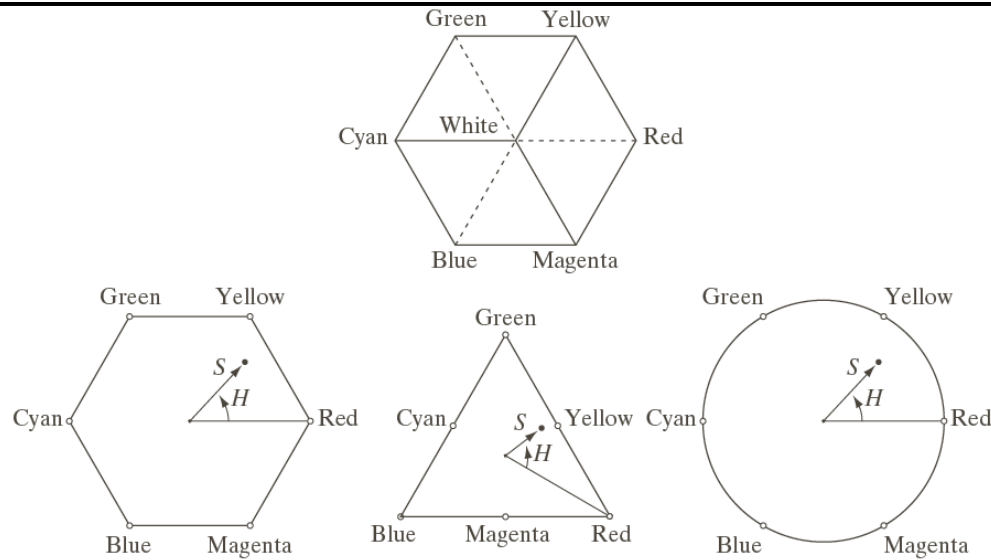
a b

FIGURE 6.12
Conceptual relationships between the RGB and HSI color models.

The HSI space is represented by a vertical intensity axis and the locus of color points that lie on planes perpendicular to this axis. As the planes move up and down the intensity axis, the boundary defined by the intersection of this plane with the faces of the cube have either triangular or hexagonal shape.

Computer Vision

Course 7



a
b c d

FIGURE 6.13 Hue and saturation in the HSI color model. The dot is an arbitrary color point. The angle from the red axis gives the hue, and the length of the vector is the saturation. The intensity of all colors in any of these planes is given by the position of the plane on the vertical intensity axis.

In the plane shown in Figure 6.13(a) primary colors are separated by 120° . The secondary colors are 60° from the primaries. The hue of the point is determined

by an angle from some reference point. Usually (but not always) an angle of 0° from the red axis designates 0 hue, and the hue increases counterclockwise from there. The saturation (distance from the vertical axis) is the length of the vector from the origin to the point. The origin is defined by the intersection of the color plane with the vertical intensity axis.

Converting colors from RGB to HSI

$$H = \begin{cases} \theta & \text{if } B \leq G \\ 360 - \theta & \text{if } B > G \end{cases}$$
$$\theta = \arccos \left\{ \frac{[(R - G) + (R - B)]}{2\sqrt{[(R - G)^2 + (R - B)(G - B)]}} \right\}$$

$$S = 1 - \frac{3}{R + G + B} \min\{R, G, B\}$$

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

It is assumed that the RGB values have been normalized to the range $[0,1]$ and that angle θ is measured with respect to the red axis of the HSI space in Figure 6.13. Hue can be normalized to the range $[0,1]$ by dividing it to 360° . The other two HSI components are in this range if the RGB values are in the interval $[0,1]$.

$R=100, G=150, B=200 \rightarrow H=210^\circ, S=1/3, I=150/255=0.588$



Converting colors from HSI to RGB

Given values of HSI we now want to find the corresponding RGB values in the same range.

RG sector ($0^\circ \leq H < 120^\circ$)

$$B = I(1 - S)$$

$$R = I \left[1 + \frac{S \cos H}{\cos(60^\circ - H)} \right]$$

$$G = 3I - (R + B)$$

GB sector ($120^\circ \leq H < 240^\circ$)

$$R = I(1 - S)$$

$$H = H - 120^\circ, G = I \left[1 + \frac{S \cos H}{\cos(60^\circ - H)} \right]$$

$$B = 3I - (R + G)$$

BR sector ($120^\circ \leq H < 240^\circ$)

$$G = I(1 - S)$$

$$H = H - 240^\circ, B = I \left[1 + \frac{S \cos H}{\cos(60^\circ - H)} \right]$$

$$R = 3I - (G + B)$$

Lab color space

The effectiveness of such transformations is judged ultimately in print. The transformations are developed and evaluated on monitors. It is necessary to have a high degree of consistency between the monitors and the output devices. This is best accomplished with a *device-independent color model* that relates the color gamut of the monitors and output devices, as well as any other devices being used, to one another. The model of choice for many *color management systems* (CMS) is the CIE L*a*b* model, also called CIELAB. The L*a*b* color components are given by the following equations:

Computer Vision

Course 7

$$L^* = 116 \cdot h\left(\frac{Y}{Y_w}\right) - 16$$

$$a^* = 500 \cdot \left[h\left(\frac{X}{X_w}\right) - h\left(\frac{Y}{Y_w}\right) \right]$$

$$b^* = 200 \cdot \left[h\left(\frac{Y}{Y_w}\right) - h\left(\frac{Z}{Z_w}\right) \right]$$

$$h(q) = \begin{cases} \sqrt[3]{q} & q > 0.008856 \\ 7.787q + \frac{16}{116} & q \leq 0.008856 \end{cases}$$

X_W, Y_W, Z_W are reference tristimulus values – typically the white of a perfectly reflecting diffuser under CIE standard D65 illumination ($x = 0.3127$, $y = 0.33290$, $z = 1 - x - y$).

The $L^*a^*b^*$ color space is *colorimetric* (i.e. colors perceived as matching are encoded identically), *perceptually uniform* (i.e. color differences among various hues are perceived uniformly), and *device independent*. Like the HSI system, the $L^*a^*b^*$ system is an excellent decoupler of intensity (represented by lightness L^*) and color (represented by a^* for red minus green and b^* for green minus blue), making it useful in both image manipulation (tone and contrast editing) and image compression applications.

Other color spaces

YIQ – for the NTSC (National Television System Committee) television system in US

Y – luminance

I (in-phase), Q (quadrature) – chrominance

YUV – for the PAL (Phase Alternation Line) and SECAM (Séquentiel Couleur à Mémoire) television system in Europe

(I, Q) – obtained by rotating (U, V)

YC_bC_r – digital video transmission

More about color spaces in: Andreas Koschan, Mongi Abidi, *Digital Color Image Processing*, Wiley, 2008

Color Difference

RGB, CMY, Lab – Euclidean distance

HSI - $F_1=(H_1, S_1, I_1)$, $F_2=(H_2, S_2, I_2)$

$$d_{\text{HSI}}(F_1, F_2) = \sqrt{(\Delta I)^2 + (\Delta C)^2}, \quad \Delta I = |I_1 - I_2|$$

$$\Delta C = \sqrt{S_1^2 + S_2^2 - 2S_1S_2 \cos \theta}$$

$$\theta = \begin{cases} |H_1 - H_2| & \text{if } |H_1 - H_2| \leq \pi \\ 2\pi - |H_1 - H_2| & \text{if } |H_1 - H_2| > \pi \end{cases}$$