Module 1

What Is Digital Image Processing?

An image may be defined as a two-dimensional function, , where x and y are spatial (plane) coordinates, and the amplitude of f at any pair of coordinates (x, y) is called the intensity or gray level of the image at that point. When x, y, and the intensity values of f are all finite, discrete quantities, we call the image a digital image.

A digital image is composed of a finite number of elements, each of which has a particular location and value. These elements are called picture elements, image elements, pels, and pixels. Pixel is the term used most widely to denote the elements of a digital image.

Fundamental Steps in Digital Image Processing

The diagram does not imply that every process is applied to an image. Rather, the intention is to convey an idea of all the methodologies that can be applied to images for different purposes and possibly with different objectives.

Image acquisition is the first process gave some hints regarding the origin of digital image acquisition could be as simple as being given an image that is already in digital form. Generally, the image acquisition stage involves preprocessing, such as scaling.

Image enhancement is the process of manipulating an image so that the result is more suitable than the original for a specific application. Thus, for example, a method that is quite useful for enhancing X-ray images may not be the best approach for enhancing satellite images taken in the infrared band of the electromagnetic spectrum. There is no general "theory" of image enhancement. When an image is processed for visual interpretation, the viewer is the ultimate judge of how well a particular method works

Image restoration is an area that also deals with improving the appearance of an image. However, unlike enhancement, which is subjective, image restoration is objective, in the sense that restoration techniques tend to be based on mathematical or probabilistic models of image degradation. Enhancement, on the other hand, is based on human subjective preferences regarding what constitutes a "good" enhancement result.

Color image processing is an area that has been gaining in importance because of the significant increase in the use of digital images over the Internet.

Wavelets are the foundation for representing images in various degrees of resolution.

Compression, as the name implies, deals with techniques for reducing the storage required to save an image, or the bandwidth required to transmit it. Compression, as the name implies, deals with techniques for reducing the storage required to save an image, or the bandwidth required to transmit it. Although storage technology has improved significantly over the past decade, the same cannot be said for transmission capacity.

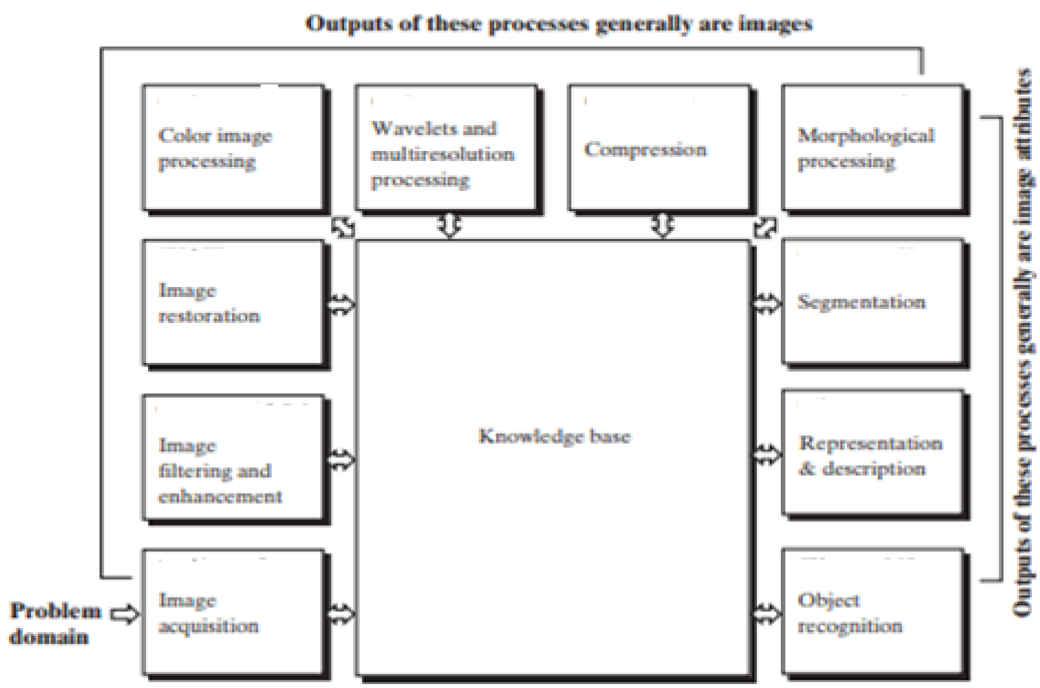
Morphological processing deals with tools for extracting image components that are useful in the representation and description of shape.

Segmentation procedures partition an image into its constituent parts or objects. In general, autonomous segmentation is one of the most difficult tasks in digital image processing. A rugged segmentation procedure brings the process a long way toward successful solution of imaging problems that require objects to be identified individually. On the other hand, weak or erratic segmentation algorithms almost always guarantee eventual failure.

Representation and description almost always follow the output of a segmentation stage, which usually is raw pixel data, constituting either the boundary of a region (i.e., the set of pixels separating one image region from another) or all the points in the region itself.

The first decision that must be made is whether the data should be represented as a boundary or as a complete region**. Boundary representation** is appropriate when the focus is on external shape characteristics, such as corners and inflections. **Regional representation** is appropriate when the focus is on internal properties, such as texture or skeletal shape. In some applications, these representations complement each other.

Recognition is the process that assigns a label (e.g., "vehicle") to an object based on its descriptors

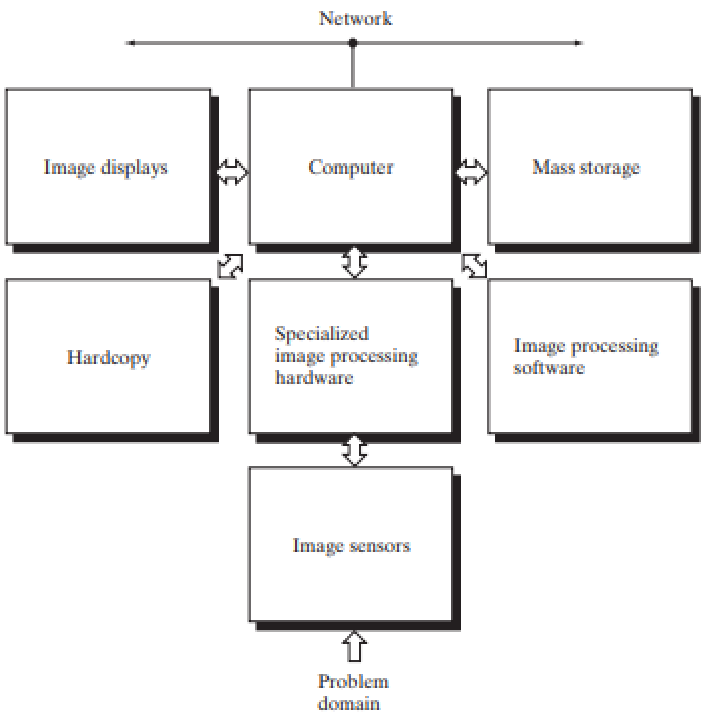


Fundamental steps in digital image processing

Components of an Image Processing System

Although large-scale image processing systems still are being sold for massive imaging applications, such as processing of satellite images, the trend continues toward miniaturizing and blending of general-purpose small computers with specialized image processing hardware. Figure 1.24 shows the basic components comprising a typical general-purpose system used for digital image processing.

The function of each component is discussed in the following



Components of a general-purpose image processing system.

Image sensing refers to sensing, two elements are required to acquire digital images. The first is a physical device that is sensitive to the energy radiated by the object we wish to image. The second, called a digitizer, is a device for converting the output of the physical sensing device into digitalform.

Specialized image processing hardware usually consists of the digitizer just mentioned, plus hardware that performs other primitive operations, such as an arithmetic logic unit (ALU) that performs arithmetic and logical operations in parallel on entire images. One example of how an ALU is used is in averaging images as quickly as they are digitized, for the purpose of noise reduction. This type of hardware sometimes is called a front-end subsystem, and its most distinguishing characteristic is speed.

The **computer** in an image processing system is a general-purpose computer and can range from a PC to a supercomputer. In dedicated applications, sometimes custom computers are used to achieve a required level of performance, but our interest here is on general-purpose image processing systems. In these systems, almost any well-equipped PC-type machine is suitable for off-line image processing tasks.

Software for image processing consists of specialized modules that perform specific tasks. A well-designed package also includes the capability for the user to write code that, as a minimum, utilizes the specialized modules. More sophisticated software packages allow the integration of those modules and general-purpose software commands from at least one computer language.

Mass storage capability is a must in image processing applications. An image of size pixels, in which the intensity of each pixel is an 8-bit quantity, requires one megabyte of storage space if the image is not compressed. When dealing with thousands, or even millions, of images, providing adequate storage in an image processing system can be a challenge. Digital storage for image processing applications falls into three principal categories:

(1) Short-term storage for use during processing,→ One method of providing short-term storage is computer memory. Another is by specialized boards, called frame buffers, that store one or more images and can be accessed rapidly, usually at video rates (e.g., at 30 complete images per second). The latter method allows virtually instantaneous image zoom, as well as scroll (vertical shifts) and pan (horizontal shifts). Frame buffers usually are housed in the specialized image processing hardware unit.

(2) On-line storage for relatively fast recall, generally takes the form of magnetic disks or optical-media storage. The key factor characterizing on-line storage is frequent access to the stored data.

(3) Archival storage, characterized by massive storage requirements but infrequent need for access. Magnetic tapes and optical disks housed in "jukeboxes" are the usual media for archival applications.

Storage is measured in bytes (eight bits), Kbytes (one thousand bytes), Mbytes (one million bytes), Gbytes (meaning giga, or one billion, bytes), and Tbytes (meaning tera, or one trillion, bytes).

Image displays in use today are mainly color (preferably flat screen) TV monitors. Monitors are driven by the outputs of image and graphics display cards that are an integral part of the computer system. In some cases, it is necessary to have stereo displays, and these are implemented in the form of headgear containing two small displays embedded in goggles worn by the user.

Hardcopy devices for recording images include laser printers, film cameras, heat-sensitive devices, inkjet units, and digital units, such as optical and CDROM disks. Film provides the highest possible resolution, but paper is the obvious medium of choice for written material.

Networking is almost a default function in any computer system in use today. Because of the large amount of data inherent in image processing applications, the key consideration in image transmission is bandwidth. In dedicated networks, this typically is not a problem, but communications with remote sites via the Internet are not always as efficient.

Elements of Visual Perception:

Although the digital image processing field is built on a foundation of mathematical and probabilistic formulations, human intuition and analysis play a central role in the choice of one technique versus another, and this choice often is made based on subjective, visual judgments

Structure of the Human Eye:

Figure 1.3 shows a simplified horizontal cross section of the human eye. The eye is nearly a sphere, with an average diameter of approximately 20 mm.

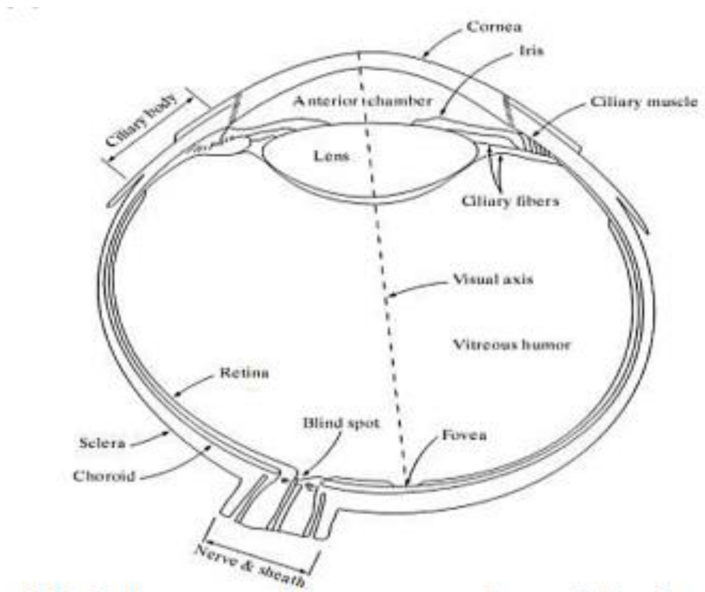
Three membranes enclose the eye: the cornea and sclera outer cover; the choroid; and the retina. The cornea is a tough, transparent tissue that covers the anterior surface of the eye. Continuous with the cornea, the sclera is an opaque membrane that encloses the remainder of the optic globe. The choroid lies directly below the sclera. This membrane contains a network of blood vessels that serve as the major source of nutrition to the eye. Even superficial injury to the choroid, often not deemed serious, can lead to severe eye damage as a result of inflammation that restricts blood flow. The choroid coat is heavily pigmented and hence helps to reduce the amount of extraneous light entering the eye and the backscatter within the optical globe.

At its anterior extreme, the choroid is divided into the ciliary body and the iris diaphragm. The latter contracts or expands to control the amount of light that enters the eye. The central opening of the iris (the pupil) varies in diameter from approximately 2 to 8 mm. The front of the iris contains the visible pigment of the eye, whereas the back contains a black pigment.The lens is made up of concentric layers of fibrous cells and is suspended by fibers that attach to the ciliary body. It contains 60 to 70%water, about 6%fat, and more protein than any other tissue in the eye.

The innermost membrane of the eye is the retina, which lines the inside of the wall's entire posterior portion. When the eye is properly focused, light from an object outside the eye is imaged on the retina. Pattern vision is afforded by the distribution of discrete light receptors over the surface of the retina. There are two classes of receptors: cones and rods.

The cones in each eye number between 6 and 7 million. They are located primarily in the central portion of the retina, called the fovea, and are highly sensitive to color. Humans can resolve fine details with these cones largely because each one is connected to its own nerve end. Muscles controlling the eye rotate the eyeball until the image of an object of interest falls on the fovea. Cone vision is called photopic or bright-light vision.

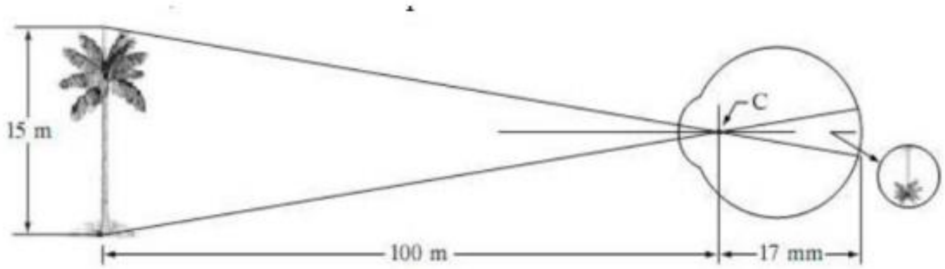
The number of rods is much larger: Some 75 to 150 million are distributed over the retinal surface. The larger area of distribution and the fact that several rods are connected to a single nerve end reduce the amount of detail discernible by these receptors. Rods serve to give a general, overall picture of the field of view. They are not involved in color vision and are sensitive to low levels of illumination. For example, objects that appear brightly colored in daylight when seen by moonlight appear as colorless forms because only the rods are stimulated. This phenomenon is known as scotopic or dim-light vision



Simplified diagram of a cross section of the human eye.

Image Formation in the Eye:

In an ordinary camera the lens has a fixed focal length, and focusing at various distances is achieved by varying the distance between the lens and the imaging plane, where the film is located. In human eye the distance between the lens and the imaging region is fixed, and the focal length needed to achieve proper focus is obtained by varying the shape of the lens. The fibers in the ciliary body accomplish this, flattering or thickening the lens for distant or near objects respectively. The distance between the lens and retina along the visual axis is approximately 17mm. the range of focal lengths is approximately 14mm to 17mm, later when eye is relaxed and focused at distances greater than about 3m.



Graphical representation of the eye looking at a palm tree Point C is the optical center of the lens.

For example, the observer is looking at a tree 15 m high at a distance of 100 m. If h is the height in mm of that object in the retinal image, the geometry of Fig.1.4 yields 15/100=h/17 or h=2.55mm. The retinal image is reflected primarily in the area of the fovea. Perception then takes place by the relative excitation of light receptors, which transform radiant energy into electrical impulses that are ultimately decoded by the brain.

Brightness Adaptation and Discrimination:

Because digital images are displayed as a discrete set of intensities, the eye's ability to discriminate between different intensity levels is an important consideration in presenting imageprocessing results. The range of light intensity levels to which the human visual system can adapt is enormous—on the order of 1010—from the scotopic threshold to the glare limit.

Experimental evidence indicates that subjective brightness (intensity as perceived by the human visual system) is a logarithmic function of the light intensity incident on the eye. The below Figure, a plot of light intensity versus subjective brightness, illustrates this characteristic. The long solid curve represents the range of intensities to which the visual systems can adapt.In photopic vision alone, the range is about 106. The transition from scotopic to photopic vision is gradual over the approximate range from 0.001 to 0.1 millilambert (–3 to –1 mL in the log scale), as the double branches of the adaptation curve in this range show.

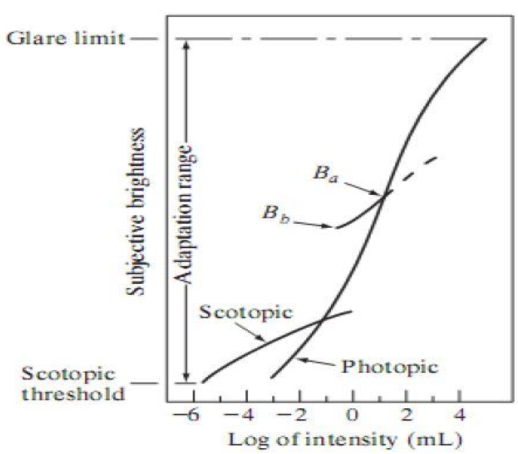


Figure: Range of subjective brightness sensations showing a particular adaption level

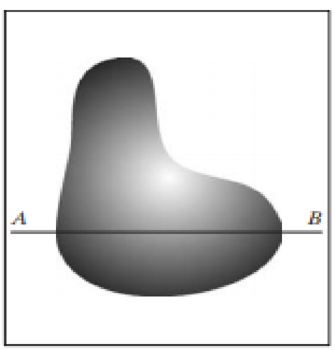
The essential point in interpreting the impressive dynamic range depicted is that the visual system cannot operate over such a range simultaneously. Rather, it accomplishes this large variation by changes in its overall sensitivity, a phenomenon known as brightness adaptation. The total range of distinct intensity levels it can discriminate simultaneously is rather small when compared with the total adaptation range. For any given set of conditions, the current sensitivity level of the visual system is called the brightness adaptation level, which may correspond, for example, to brightness B_a in Fig. The short intersecting curve represents the range of subjective brightness that the eye can perceive when adapted to this level. This range is rather restricted, having a level B_b at and below which all stimuli are perceived as indistinguishable blacks. The upper (dashed) portion of the curve is not actually restricted but, if extended too far, loses its meaning because much higher intensities would simply raise the adaptation level higher than B_a.

Image Sampling and Quantization

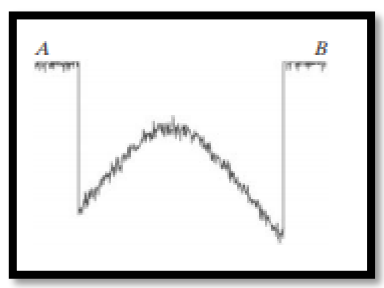
To create a digital image, we need to convert the continuous sensed data into digital form. This involves two processes: sampling and quantization.

Basic Concepts in Sampling and Quantization

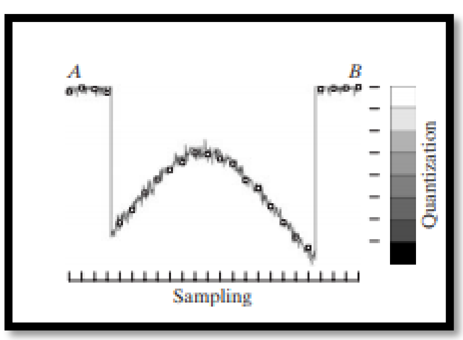
The basic idea behind sampling and quantization is illustrated in Fig. 2.16. Figure 2.16(a) shows a continuous image f that we want to convert to digital form. An image may be continuous with respect to the x- and y-coordinates, and also in amplitude. To convert it to digital form, we have to sample the function in both coordinates and in amplitude. Digitizing the coordinate values is called sampling. Digitizing the amplitude values is called quantization.



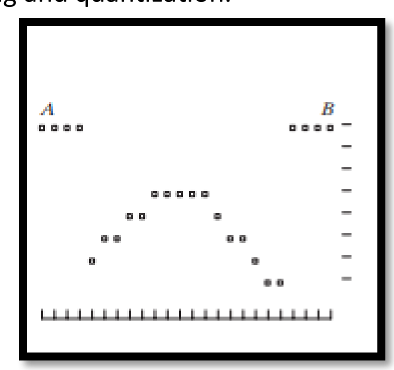
(a) Continuous image.



(b) A scan line from A to B in the continuous image, used to Illustrate the concepts of sampling and quantization.



(c) Sampling and quantization.



(d) Digital scan line.

The one-dimensional function in Fig. (b) is a plot of amplitude (intensity level) values of the continuous image along the line segment AB in Fig. (a). The random variations are due to image noise. To sample this function, we take equally spaced samples along line AB, as shown in Fig.

(c). The spatial location of each sample is indicated by a vertical tick mark in the bottom part of the figure.

However, the values of the samples still span (vertically) a continuous range of intensity values. In order to form a digital function, the intensity values also must be converted (quantized) into discrete quantities. The right side of Fig. (c) shows the intensity scale divided into eight discrete intervals, ranging from black to white. The vertical tick marks indicate the specific value assigned to each of the eight intensity intervals. The continuous intensity levels are quantized by assigning one of the eight values to each sample. The digital samples resulting from both sampling and quantization are shown in Fig. (d). Starting at the top of the image and carrying out this procedure line by line produces a two-dimensional digital image.

Color fundamentals

Color of an object is determined by the nature of the light reflected from it. When a beam of sunlight passes through a glass prism, the emerging beam of light is not white but consists instead of a continuous spectrum of colors ranging from violet at one end to red at the other. As Fig. shows, the color spectrum may be divided into six broad regions: violet, blue, green, yellow, orange, and red. When viewed in full color no color in the spectrum ends abruptly, but rather each color blends smoothly into the next.

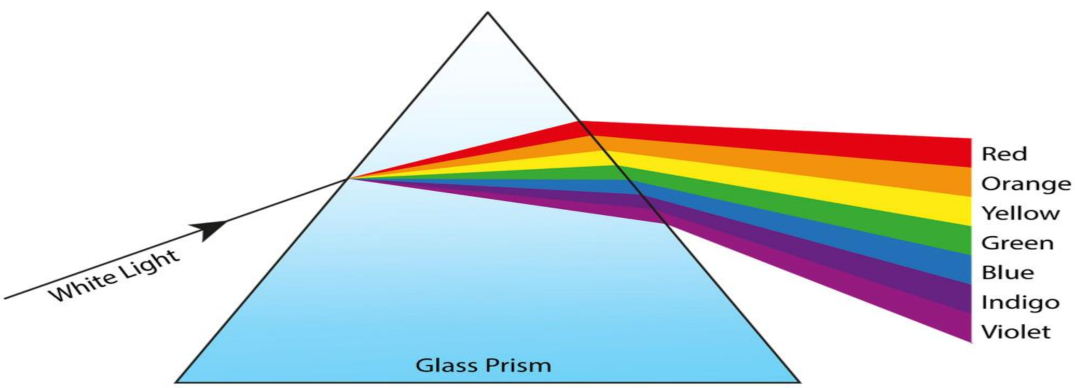


Fig: Color spectrum seen by passing white light through a prism

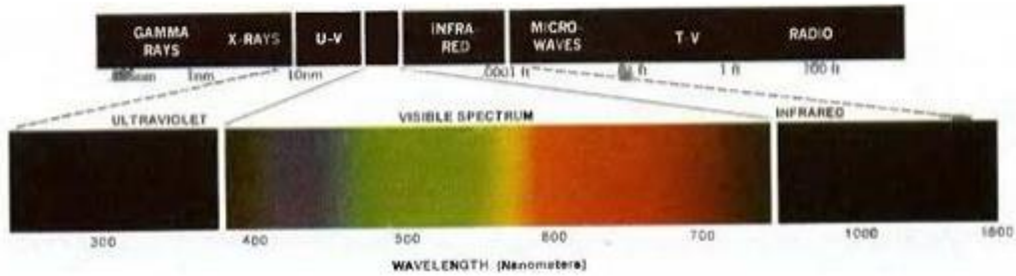


Fig: Wavelength comprising the visible range of electromagnetic spectrum

As illustrated in Figure, visible light is composed of a relatively narrow band of frequencies in the electromagnetic spectrum. A body that reflects light that is balanced in all visible wavelengths appears white to the observer. However, a body that favors reflectance in a limited range of the visible spectrum exhibits some shades of color. For example, green objects reflect light with wavelengths primarily in the 500 to 570 nm range while absorbing most of the energy at other wavelengths. Characterization of light is central to the science of color. If the light is achromatic (void of color), its only attribute is its intensity, or amount. Achromatic light is what viewers see on a black and white television set. Three basic quantities are used to describe the quality of a chromatic light source: radiance, luminance, and brightness.

Radiance: Radiance is the total amount of energy that flows from the light source, and it is usually measured in watts (W).

Luminance: Luminance, measured in lumens (lm), gives a measure of the amount of energy an observer perceives from a light source

Brightness: Brightness is a subjective descriptor that is practically impossible to measure. It embodies the achromatic notion of intensity and is one of the key factors in describing color sensation.

Cones are the sensors in the eye responsible for color vision. Detailed experimental evidence has 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. Approximately65% 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 shows average experimental curves detailing the absorption of light by the red, green, and blue cones in the eye. Due to these absorption characteristics of the human eye, colors arc seen as variable combinations of the so- called primary colors red (R), green (G), and blue (B).The primary colors can be added to produce the secondary colors of light --magenta (red plus blue), cyan (green plus blue), and yellow (red plus green). Mixing the three primaries or a secondary with its opposite primary color, in the right intensities produces white light.

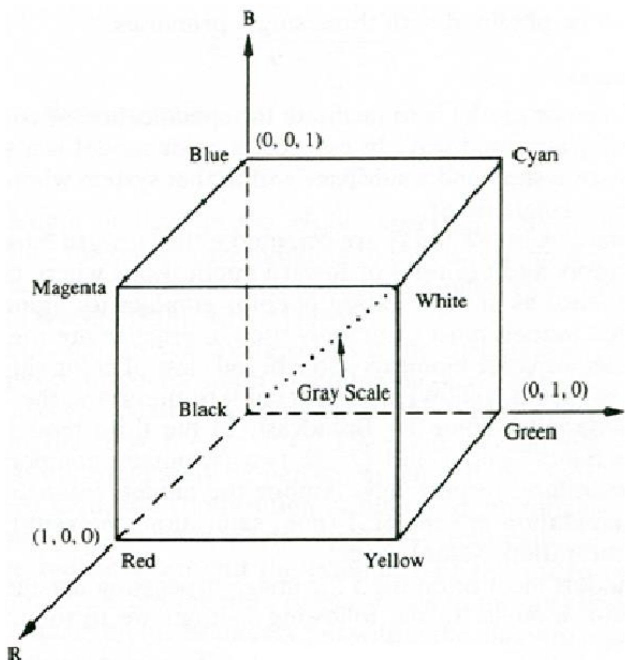
The characteristics generally used to distinguish one color from another are brightness, hue, and saturation. Brightness embodies the chromatic notion of intensity. Hue is an attribute associated with the dominant wavelength in a mixture of light waves. Hue represents dominant color as perceived by an observer. Saturation refers to the relative purity or the amount of white light mixed with a hue. The pure spectrum colors are fully saturated. Colors such as pink(red and white) and lavender (violet and 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.

Color Models

The purpose of a color model (also called color space or color system) is to facilitate the specification of colors in some standard, generally accepted way. In essence, a color model is a specification of a coordinate system and a subspace within that system where each color is represented by a single point.

The RGB Color Model:

In the RGB model, each color appears in its primary spectral components of red, green, and blue. This model is based on a Cartesian coordinate system. The color subspace of interest is the cube shown in Fig, in which RGB values are at three corners; cyan, magenta, and yellow are at three other corners; black is at the origin; and white is at the corner farthest from the origin. In this model, the gray scale (points of equal RGB values) extends from black to white along the line joining these two points. The different colors in this model arc points on or inside the cube, and are defined by vectors extending from the origin. For convenience, the assumption is that all color values have been normalized so that the cube shown in Fig. is the unit cube.That is, all values of R, G. and B are assumed to be in the range [0, 1].



Consider an RGB image in which each of the red, green, and blue images is an 8-bit image. Under these conditions each RGB color pixel [that is, a triplet of values (R, G, B)] is said to have a depth of 24 bits C image planes times the number of bits per plane). 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 (28) 3 = 16,777,216.

RGB is ideal for image color generation (as in image capture by a color camera or image display in a monitor screen), but its use for color description is much more limited

The CMY and CMYK color model.

Cyan, magenta, and yellow are the secondary colors of light or, alternatively, the primary colors of pigments. For example, when a surface coated with cyan pigment is illuminated with white light, no red light is reflected from the surface. That is, cyan subtracts red light from reflected white light, which itself is composed of equal amounts of red, green, and blue light. Most devices that deposit colored pigments on paper, such as color printers and copiers, require CMY data input or perform an RGB to CMY conversion internally. This conversion is performed using the simple operation (1) where, again, the assumption is that all color values have been normalized to the range [0, 1]. Equation (1) demonstrates that light reflected from a surface coated with pure cyan does not contain red (that is, C = 1 — R in the equation).

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

Similarly, pure magenta does not reflect green, and pure yellow does not reflect blue. Equation (1) also reveals that RGB values can be obtained easily from a set of CMY values by subtracting the individual CMY values from 1. As indicated earlier, in image processing this color model is used in connection with generating hardcopy output, so the inverse operation from CMY to RGB generally is of little practical interest.

Equal amounts of the pigment primaries, cyan, magenta, and yellow should produce black. In practice, combining these colors for printing produces a muddy-looking black.

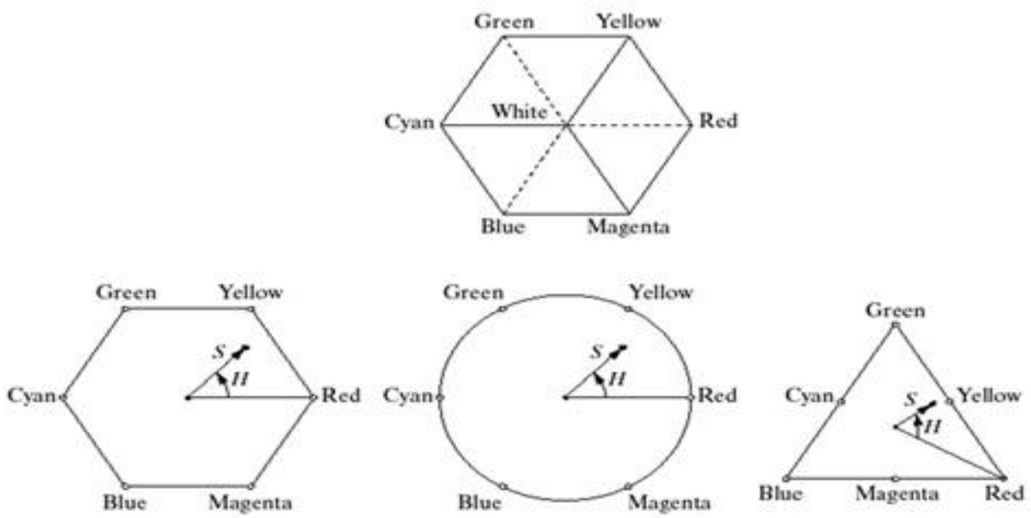
HSI color Model.

When humans view a color object, we describe it by its hue, saturation, and brightness. Hue is a color attribute that describes a pure color (pure yellow, orange, or red), whereas saturation gives a measure of the degree to which a pure color is diluted by white light. Brightness is a subjective descriptor that is practically impossible to measure. It embodies the achromatic notion of intensity and is one of the key factors in describing color sensation. Intensity (gray level) is a most useful descriptor of monochromatic images.

This quantity definitely is measurable and easily interpretable. The HSI (hue, saturation, intensity) color model, decouples the intensity component from the color-carrying information (hue and saturation) in a color image. As a result, the HSI model is an ideal tool for developing image processing algorithms based on color descriptions that are natural and intuitive to humans. In Fig 1.15 the primary colors are separated by 120°. The secondary colors are 60° from the primaries, which means that the angle between secondaries is also 120°. Figure 5.4(b) shows the same hexagonal shape and an arbitrary color point (shown as a dot).

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.

Note that the origin is defined by the intersection of the color plane with the vertical intensity axis. The important components of the HSI color space are the vertical intensity axis, the length of the /vector to a color point, and the angle this vector makes with the red axis.



a
b c d

FIGURE 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.

Some basic relationship between pixels

Neighbor of a pixel:

A pixel p at coordinate (x , y) has four horizontal and vertical neighbor whose coordinate can be given by

(x+1 , y) (x-1 , y) (x , y + 1) (x ,y-1)

This set of pixel called the 4-neighbours of p is denoted by N_4(p) ,Each pixel is a unit distance from (x,y) and some of the neighbors of P lie outside the digital image of (x,y) is on the border if the image . The four diagonal neighbor of P have coordinated

(x+1, y+1),(x+1,y+1),(x-1,y+1),(x-1,y-1)

And are deported by nd (p) .these points, together with the 4-neighbours are called 8 – neighbors of P denoted by N_s (p).

Adjacency:

Let v be the set of intensity

values used to define adjacency, in a binary image, v= {1} if we are reference to adjacency of pixel with value. Three types of adjacency 4- Adjacency – two pixel P and Q with value from V are 4 –adjacency if A is in the set n4(P) 8- Adjacency – two pixel P and Q with value from V are 8 –adjacency if A is in the set n8(P) M-adjacency –two pixel P and Q with value from V are m – adjacency if (i) Q is in n4 (p) or (ii)Q is in nd (q) and the set N4(p) U N4(q) has no pixel whose values are from V

Distance measures:

For pixel p,q and z with coordinate (x.y) ,(s,t) and (v,w) respectively D is a distance function or metric

if D [p . q] ≥ O {D[p . q] = O iff p=q}

D [p . q] = D [p . q]

and D [p . q] ≥ O {D[p . q]+D(q , z)

The Education Distance between p and is defined as

D_e (p,q) = I $_y$ – t I

The D4 Education Distance between p and is defined as De (p,q) = Iy – t I