University of Baghdad
College of Education for Pure Science Ibn Al-Haitham
Department of Physics

Advance Digital Image Processing Lectures for (Ph.D.)

Second Semester (2023/2022)

Prof. Dr. Taghreed Abdulhameed Naji

Lecture 1

(Digital Image Fundamentals)

Outlines:

1. Structure of the Human Eye
2. Image Formation in the Eye
3. Brightness Adaption and Discrimination
4. Digital Image Processing Systems
5. Light and the electromagnetic spectrum
6. Visible light imaging
7. Imaging outside the Visible Range of the EM Spectrum
8. Image Sensing and Acquisition
9. Image Sampling and Quantization
10. Representing Digital Images
11. Spatial and Intensity Resolution
12. **Image Quantization***
13. Image Interpolation
14. Some Basic Relationships Between Pixels
15. Image Representation
16. Digital Image File Formats

References:

1. Chapter 2, Rafael C. Gonzalez and Richard E. Woods, 2018, "Digital Image Processing", 4^{th} Ed, Pearson Education Limited, New York.
2. Chapter 3, Scott E Umbaugh, 2018, "Digital Image Processing and Analysis Applications with MATLAB and CVIPtools", 3th Ed, Taylor & Francis.

* Chapter 3, only "Image Quantization" Scott E Umbaugh, 2018, "Digital Image Processing and Analysis", 3th Ed, Taylor & Francis.

1. Elements of visual perception

Although the field of digital image processing is built on a foundation of mathematics, human intuition and analysis often play a role in the choice of one technique versus another, and this choice often is made based on subjective, visual judgments. Thus, developing an understanding of the basic characteristics of human visual perception as the first step in this course of study is appropriate, by the elementary mechanics of how images are formed and perceived by humans.

1.1 The Human Visual System

The Human Visual System (HVS) has two primary components:

- Eye.
- Brian.

1.2 Structure of the Human Eye

The eye is nearly a sphere with an average of approximately 20 mm in diameter. The eye is enclosed with three membranes:

a) The cornea and sclera

b) The choroid: It has two parts:

(1) Iris Diaphragms

(2) Ciliary body.

(c) Retina:. The two major classes of receptors are:

(1) cones: they are in the number about 6 to 7 million. These are located in the central portion of the retina called the fovea. These are highly sensitive to color. Human can resolve fine details with these cones because each one is connected to its own nerve end. Cone vision is called photopic or bright light vision.

(2) Rods: they are very much in number from 75 to 150 million and are distributed over the entire retinal surface. The large area of distribution and the fact that several roads are connected to a single nerve give a general overall picture of the field of view. They are not involved in the color vision and are sensitive to low level of illumination. Rod vision is called is scotopic or dim light vision. The absent of reciprocators is called blind spot. Figure 1, illustrates a simplified cross section of the human eye.

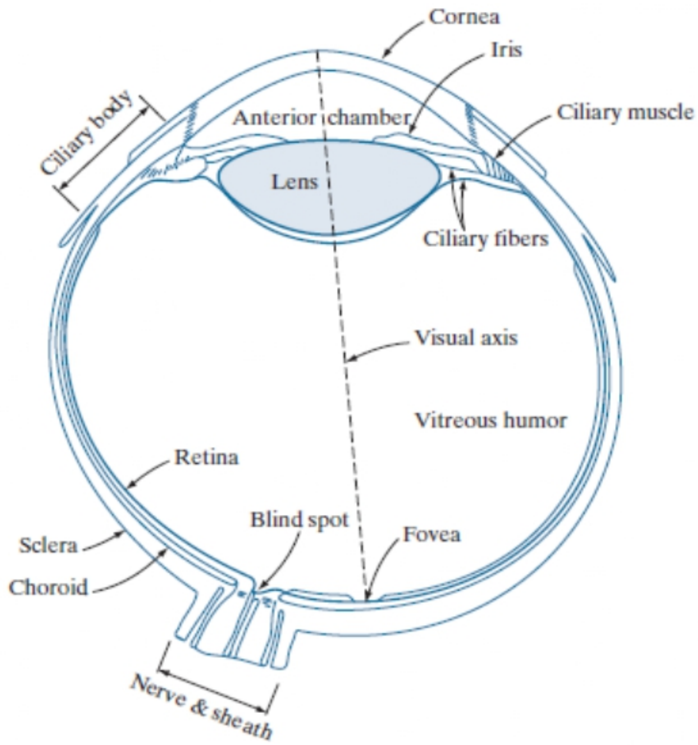


Figure 1. Simplified diagram of a cross-section of the human eye.

1.3 Image Formation in the Eye

The major in an ordinary photographic camera, the lens has a fixed focal length. Focusing at various distances is achieved by varying the distance between the lens and the imaging plane, where the film (or imaging chip in the case of a digital camera) is located. In the human eye, the converse is true; the distance between the center of the lens and the imaging sensor (the retina) 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 by flattening or thickening the lens for distant or near objects, respectively. The distance between the center of the lens and the retina along the visual axis is approximately 17 mm. The range of focal lengths is approximately 14 mm to 17 mm, the latter taking place when the eye is relaxed and focused at distances greater than about 3 m. The geometry in figure 2, illustrates how to obtain the dimensions of an image formed on the retina. For example, suppose that a person is looking at a tree 15 m high at a distance of 100 m. Letting h denote the height of that object in the retinal image, the geometry of figure 3 yields (15/100) = (h/17) or h = 2.5 mm. The retinal image is focused primarily on the region of the fovea. Perception then takes place by the relative excitation of light receptors, which transform radiant energy into electrical impulses that ultimately are decoded by the brain.

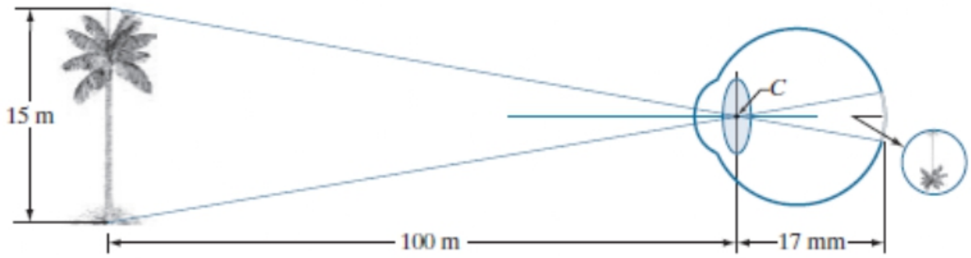


Figure 2. Graphical representation of the eye looking at a palm tree. Point C is the focal center of the lens.

1.4 Brightness Adaption and Discrimination

Digital image is displayed as a discrete set of intensities. The range of light intensity levels to which the human visual system can adopt is enormous (on the order of 10^6) from scotopic threshold to the glare limit. Experimental evidences indicate that subjective brightness is a logarithmic function of the light intensity incident on the eye. Figure 3 illustrates a plot of light intensity versus subjective brightness.

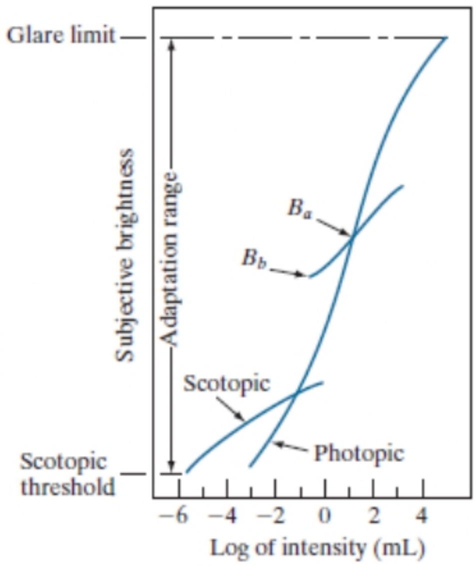


Figure 3. Range of subjective brightness sensations showing a particular adaptation level, B_a

Subjective brightness: means intensity as perceived by the human visual system.

Brightness adaptation: means the human visual system can operate only from scotopic to glare limit. It can't operate over the range simultaneously. It accomplishes this large variation by changes in its overall intensity.

Two phenomena demonstrate that perceived brightness is not a simple function of intensity. The first is based on the fact that the visual system tends to undershoot or overshoot around the boundary of regions of different intensities. Figure 4, shows a striking example of this phenomenon. Although the intensity of the stripes is constant, we actually perceive a brightness pattern that is strongly scalloped near the boundaries. These perceived scalloped bands are called Mach bands after Ernst Mach, who first described the phenomenon in 1865.

Mach bands: are one of the many visual phenomena which occur when two grey images are placed adjacent to each other differing only in illumination.

(هي إحدى الظواهر المرئية العديدة التي تحدث عند وضع صورتين رماديتين متجاورتين مع بعضهما البعض يختلفان فقط في الاضاءة)

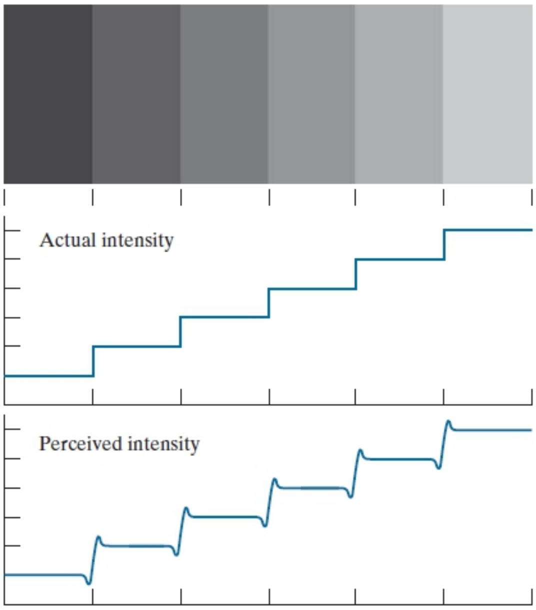


Figure 4. Illustration of **the Mach band effect**. Perceived intensity is not a simple function of actual intensity.

The second phenomenon, called simultaneous contrast, is that a region's perceived brightness does not depend only on its intensity, as figure 5 demonstrates. All the center squares have exactly the same intensity, but each appears to the eye to become darker as the background gets lighter. A more familiar example is a piece of paper that looks white when lying on a desk, but can appear totally black when used to shield the eyes while looking directly at a bright sky.

Simultaneous Contrast:

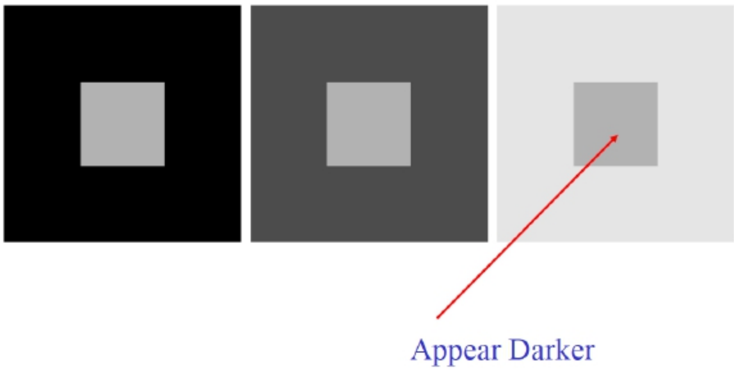


Figure 5. Examples of simultaneous contrast. All the inner squares have the same intensity, but they appear progressively darker as the background becomes lighter.

2. Digital Image Processing Systems

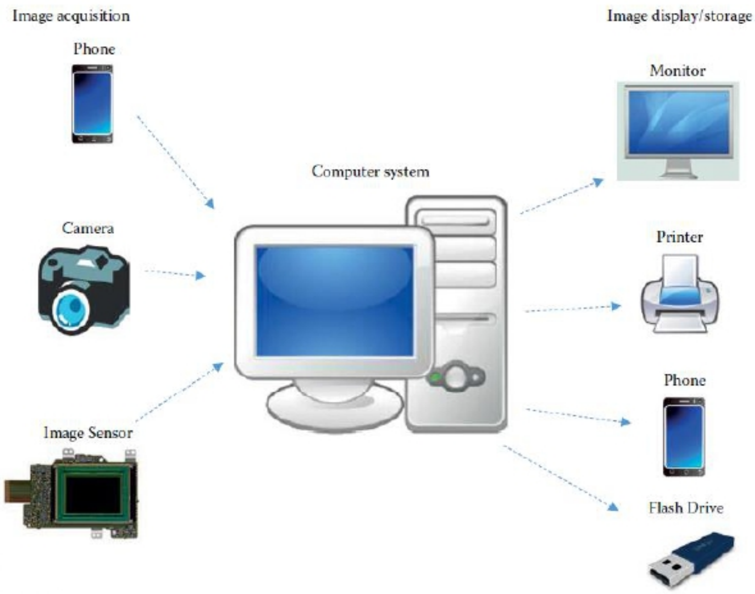


Figure 6. Digital Image Processing System Hardware.

The process of transforming a standard analog video signal into a digital image is called digitization. This transformation is necessary because the standard video signal is in analog (continuous) form, and the computer requires a digitized or sampled version of that continuous signal.

In figure 7 (a & b), we see the typical image on a display device and the electrical signal that corresponds to one line of video information (one row of image data). The horizontal sync pulse between each line of information is the control signal for the hardware to designate and tell the display hardware to end of a line, and start a new line. After one frame has been displayed, a longer synchronization pulse, called the vertical synch pulse, tells the display hardware to start a new field or frame.

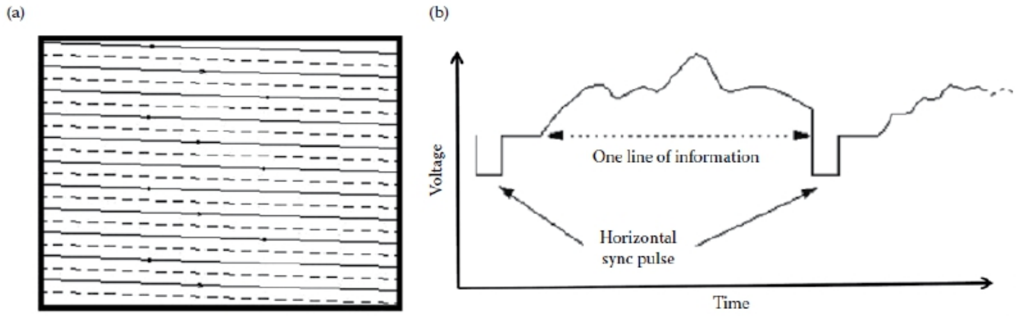


Figure 7. The Video Signal. (a) One frame, and (b) the analog video signal.

The analog video signal is converted to a digital image by sampling the continuous signal at a fixed rate. In figure 8, we see one line of a video signal being sampled (digitized) by converting the value of the voltage at each instant is converted into a number that is stored, corresponding to the brightness of the image at that point. The image brightness at a point depends on both the properties of the object and the lighting conditions in the scene. The computer can store it and process it as a digital image.

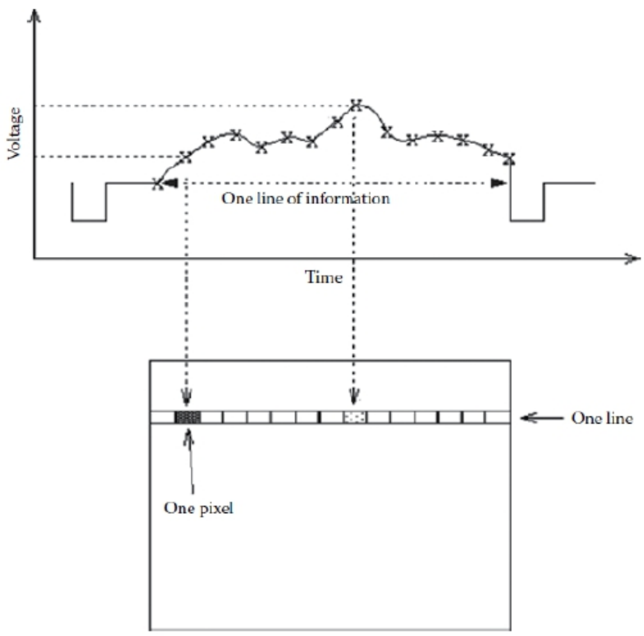


Figure 8. Digitizing (sampling) an analog video signal. The sampled voltage corresponds to the brightness of the image at that point.

The SDTV display device uses a 4:3 aspect ratio, and 16:9 aspect ratio. The aspect ratio is the width-to-height ratio of the display device, as shown in figure 10.

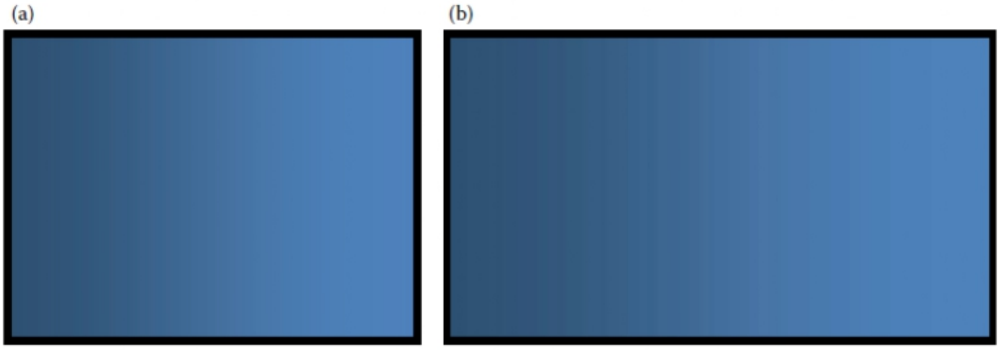


Figure 9. Aspect Ratio for (a) 4:3 standard-definition television, and (b) 16:9 high-definition television.

Digital image processing involves taking the digital image, I(r,c) (the brightness of the image at the point), and performing various imaging operations, often by applying image processing software, to modify the image data. The different levels and various types of processing can be illustrated by the hierarchical image pyramid, as shown in figure 10. In this figure, the image operations are on the left and the corresponding image representation is on the right. As we traverse this pyramid from the bottom up we get increasingly higher levels of information representation. At the very lowest level, we deal with the very large number of individual pixels, where we may perform some low-level preprocessing. The next level up is the neighborhood, which typically consists of a single pixel and the surrounding pixels, and we may to perform some preprocessing operations at this level. As we ascend the pyramid, we get higher and higher level representations of the image, and consequently, a reduction in the amount of data. At the highest level are the objects as the human visual system sees them.

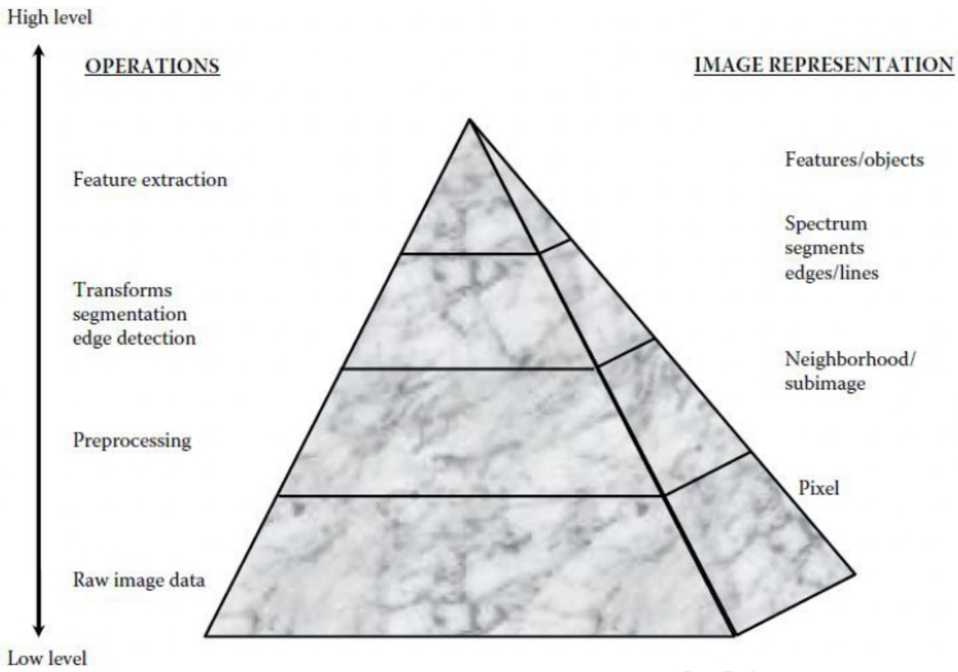


Figure 10. The pyramid of image processing operations representation

3. Light and the electromagnetic spectrum

In 1666, Sir Isaac Newton discovered that 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 figure 11, shows the range of colors we perceive in visible light is a small portion of the electromagnetic spectrum. On one end of the spectrum are radio waves with wavelengths billions of times longer than those of visible light. On the other end of the spectrum are gamma rays with wavelengths millions of times smaller than those of visible light.

The electromagnetic spectrum can be expressed in terms of wavelength, frequency, or energy. Wavelength (λ) and frequency (υ) are related by the expression:

$$\lambda = \frac{c}{\upsilon} \qquad (1)$$

Where c is the speed of light (2.998 * 10^8 m/s). Figure 11 shows a schematic representation of one wavelength. The energy of the various components of the electromagnetic spectrum is given by the expression:

$$E = h\nu \qquad (2)$$

Where h is Planck's constant. The units of wavelength are meters, with the term's microns (denoted mm and equal to 10^{-6} m) and nanometers (denoted nm and equal to 10^{-9} m) being used just as frequently. Frequency is measured in

Hertz (Hz), with one Hz being equal to one cycle of a sinusoidal wave per second. A commonly used unit of energy is the electron-volt. Electromagnetic waves can be visualized as propagating sinusoidal waves with wavelength (λ), or they can be thought of as a stream of massless particles each traveling in a wavelike pattern and moving at the speed of light. Each massless particle contains a certain amount (or bundle) of energy, called a photon.

The colors perceived in an object are determined by the nature of the light reflected by the object. A body that reflects light relatively balanced in all visible wavelengths appears white to the observer. The object 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.

Light that is void of color is called **monochromatic** (or **achromatic**) light. The only attribute of monochromatic light is its intensity. The intensity of monochromatic light is perceived to vary from black to gray and finally to white. The term **gray level** is used commonly to denote **monochromatic intensity**. The range of values of monochromatic light from black to white is called the **gray scale**, and monochromatic images are frequently referred to as gray scale images.

Chromatic (color) light spans the electromagnetic energy spectrum from approximately 0.43 to 0.79 mm. Three other quantities are used to describe a chromatic light source: radiance, luminance, and brightness. Radiance is the total amount of energy that flows from the light source, and it is usually measured in watts (W). Luminance, measured in lumens (lm), gives a measure of the amount of energy an observer perceives from a light source. The brightness is a subjective descriptor of light perception that is practically impossible to measure. It is the achromatic notion of intensity.

If a sensor can be developed that is capable of detecting energy radiated in a band of the electromagnetic spectrum, we can image events of interest in that band. The wavelength of an electromagnetic wave required to an object must be of the same size as, or smaller than, the object. For example, a water molecule has a diameter on the order of 10^{-10} m. Thus, to study these molecules, we would need a source capable of emitting energy in the far (high energy) ultraviolet band or soft (low-energy) X-ray bands. Although imaging is based predominantly on energy from electromagnetic wave radiation, this is not the only method for generating images.

Sensors may also respond to acoustical energy, as in ultrasound images. The images are created by radar (radio detection and ranging), sound energy, or lasers. Other sources of digital images are electron beams for electron microscopy and software for generating synthetic images used in graphics and visualization.

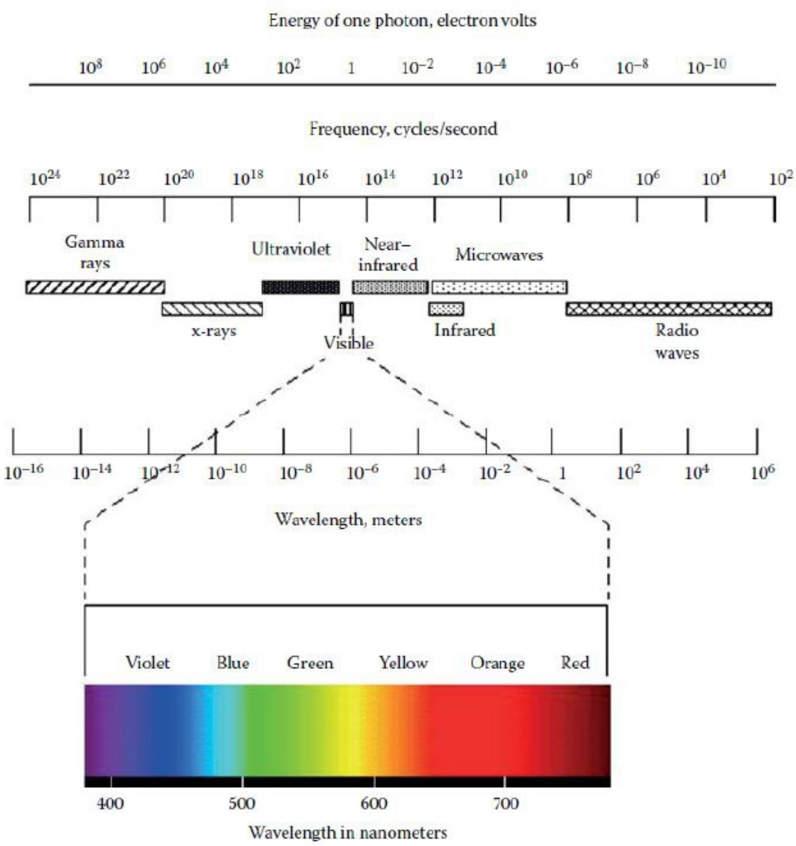


Figure 11. The electromagnetic (EM) spectrum. Higher frequencies have higher energy and are more dangerous. The visible spectrum is only a small part of the entire EM spectrum.

The sources for images:

1. Electromagnetic (EM) energy spectrum
2. Acoustic
3. Ultrasonic
4. Electronic
5. Synthetic images produced by computer

4. Visible Light Imaging

The basic model for visible light imaging is shown in figure 12. Here the light source emits light that is reflected from the object, and focused by the lens onto the image sensor. The sensor responds to the light energy by converting it into electrical energy which is then measured. This measurement is proportional to the incident energy and determines the brightness of the image at that point.

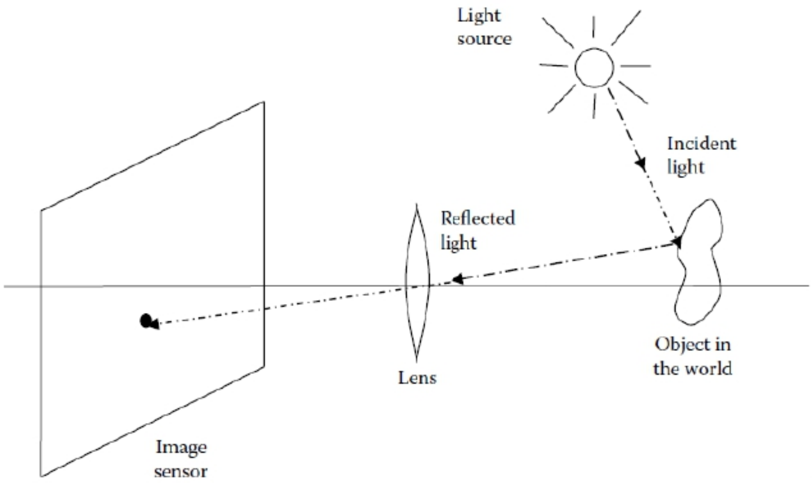


Figure 12 Model for visible light imaging.

In imaging two terms are necessary to define brightness; irradiance, and radiance. Figure 13 illustrates the difference between these two terms. Irradiance is the amount of light falling on a surface, such as an image sensor, it is measured in power per unit area. While radiance is the amount of light reflected or emitted from a surface into a solid unit angle. The units used for these two measures are different:

$$\text{Irradiance} = \frac{Power}{Area}$$

$$Radiance = \frac{Power}{(Area)(Soild\ Angle)}$$

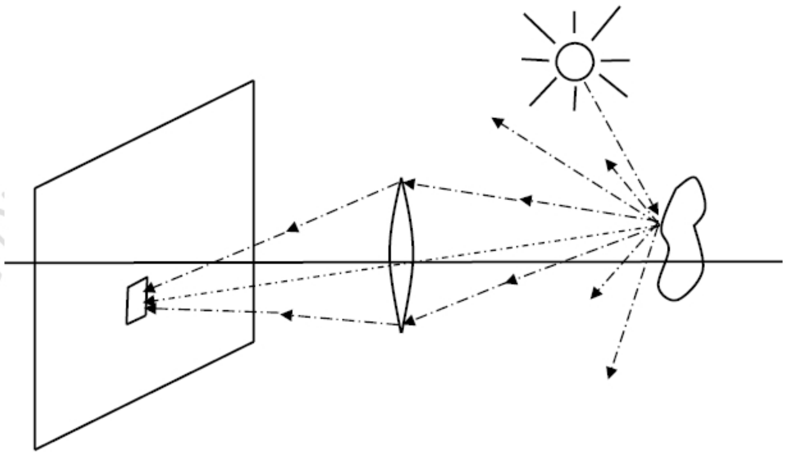


Figure 13. Irradiance and radiance.

A lens is necessary to focus light in a real imaging system. In figure 14, shows the relationship between points in the world and points in the image. The relationship between the distance of the object in the world (a) and the image plane (b) is defined by the lens equation:

$$\frac{1}{a}+\frac{1}{b}=\frac{1}{f}$$

Where f is the focal length of the lens. In this figure, three rays of light are shown. The one through the center of the lens goes straight through to the image plane, and, if the system is in focus, the other rays will meet at that point. If the object is moved closer to the lens, the single point will become a blur circle; the diameter of the circle is given by the blur equation:

$$c=\frac{d}{b'}|b'-b|$$

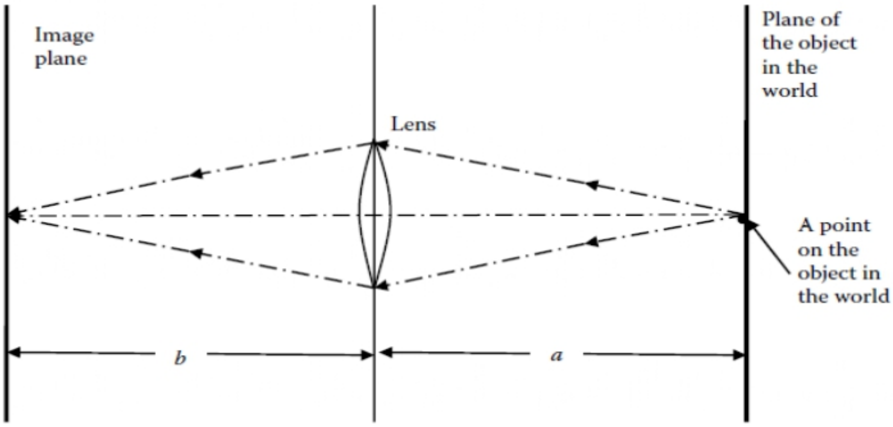


Figure 14. Relationship between points in the world and points in the image

Application of the lens equation shows the object is actually focused behind the image plane. The blur equation defines the amount of blur.

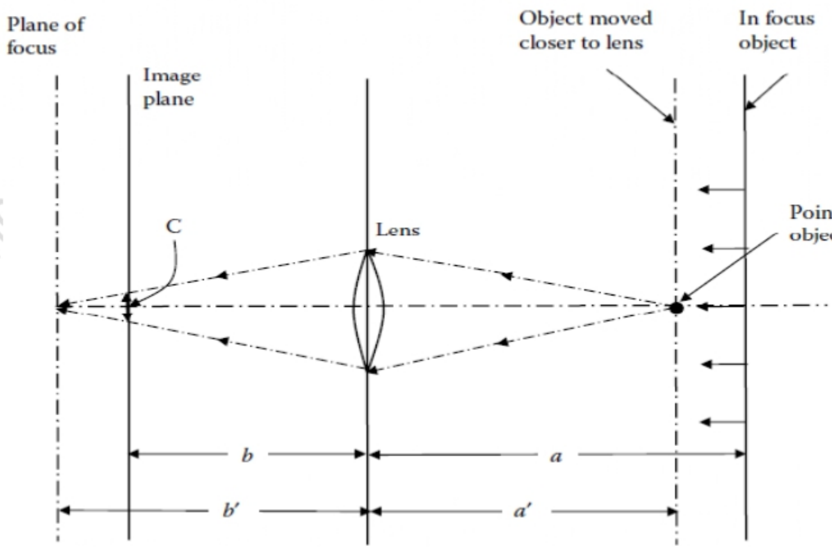


Figure 15. The blur circle from a poorly focused lens.

A real object typically does not appear in a single plane, so some blurring will occur. The conditions that will allow an object to be focused sufficiently well, will be determined by the spatial resolution of the imaging device. If the blur circles are equal to or smaller than the device resolution, the object will be focused sufficiently well. In addition to change the effective diameter Deffective, of the lens. The f-number (or f-stop) is defined as the ratio of the focal length to the lens diameter.

$$f\text{-}number = \frac{f}{D_{effective}}$$

Another important parameter of an imaging device is the field of view (FOV). The FOV is the amount of the scene that the imaging device actually (sees), that is, the angle of the cone of directions from which the device will create the image. Figure 16 shows that the FOV can be defined as:

$$FOV = 2\varphi,$$
$$\text{where } \varphi = \tan^{-1}\left(\frac{d/2}{f}\right)$$

With d is the diagonal size of the image sensor and f is the focal length of the lens. The FOV for an imaging system depends on both focal length of the lens, f, and the size of image sensor, d. For a fixed size image sensor, a wider FOV requires a lens with a shorter focal length.

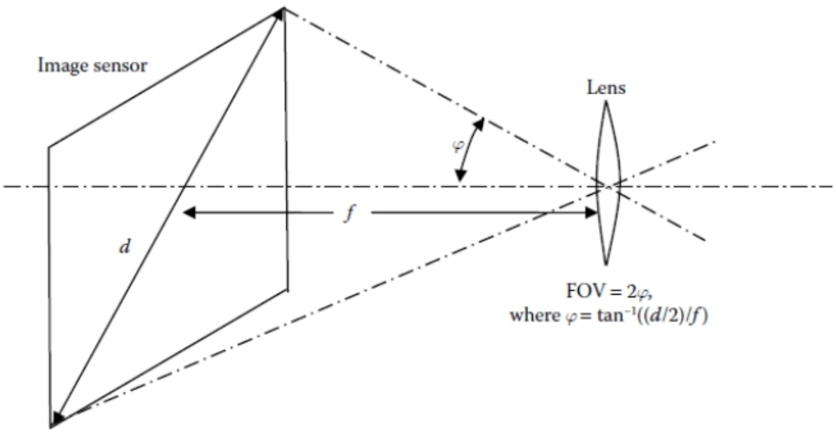


Figure 16. Field of View (FOV)

Real lenses do not typically consist of a single lens, but are multiple lenses aligned together, due to the fact that a single lens will have various types of distortions, called aberrations. The effect of these aberrations can be mitigated by aligning multiple lenses of varying types and sizes to create a compound lens. One of the negative effects of a compound lens is the vignetting effect. This effect, shown in figure 17, causes the amount of energy that actually makes it through the lens to the image plane to decrease as we move farther away from the center of the image. This effect can be avoided by using only the center portion of the lens. A compound lens causes less light on the edges of the image to get through to the image sensor. This has the effect of decreasing brightness as we move away from the center of the image.

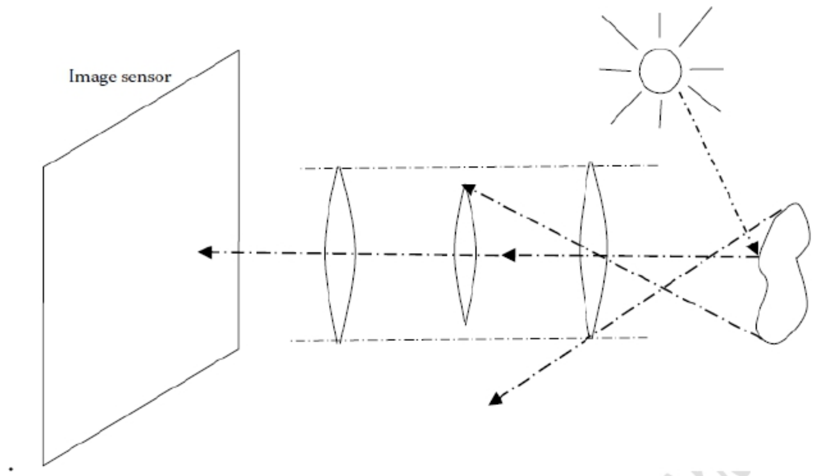


Figure 17. The vignetting effect

5. Imaging outside the Visible Range of the EM Spectrum

Imaging with gamma rays is performed by measuring the rays as they are emitted from the object. In nuclear medicine using positron emission tomography, a patient is injected with a radioactive isotope and as it decays, gamma rays are detected and measured. X-rays are used in medical diagnostics by using film that responds to x-ray energy. The x-rays are passed through the patient and recorded on the film. X-rays are also used in computerized tomography (CT) where a ring of detectors encircles the patient and is rotated various angles to obtain two-dimensional "slices" which can be assembled into a three-dimensional image. Fluorescence microscopy works by using dyes that emit visible light when ultraviolet (UV) light is beamed upon it. Examples of x-ray and UV images are shown in figure 18.

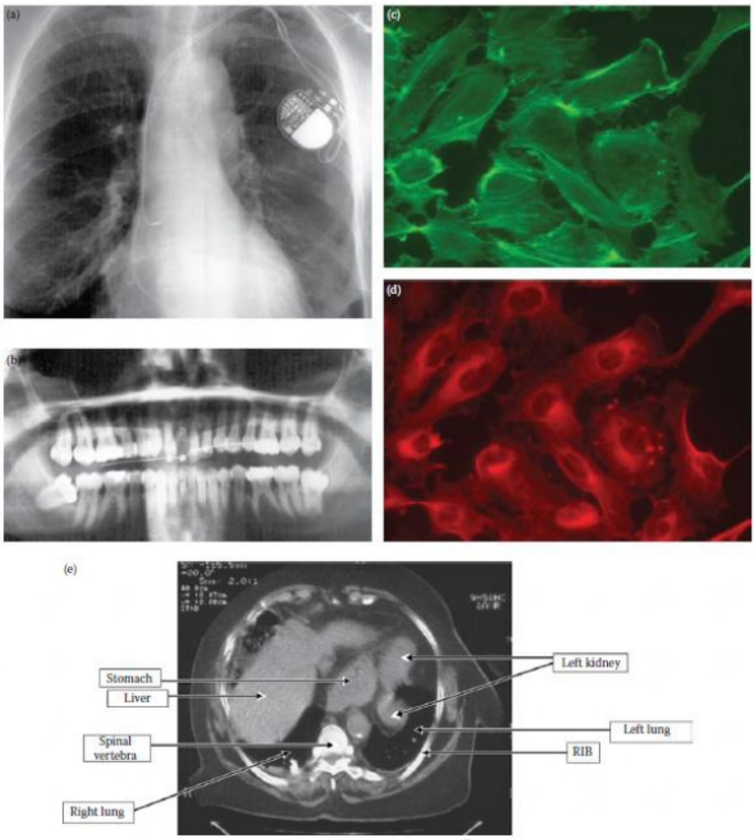


Figure 18. X-ray and UV Images. (a) X-ray of a chest, (b) dental x-ray, (c and d) fluorescence microscopy images of cells, generated by emitting visible light when illuminated by ultraviolet light, (e) one "slice" of a computerized tomography (CT) image of a patient's abdomen,

UV imaging is used in industrial applications, microscopy, and astronomy. The imaging systems use short UV, < 300 nm wavelengths.

Infrared (IR) images are used in satellite imaging (remote sensing), as features of interest, for example, moisture content and mineral mapping, are found in the IR spectral bands (Figure 19). Infrared images can be divided into four primary spectral ranges—near IR, 780 nm – 1.3 μm, middle-wave IR, 3–5 μm, long-wave IR, 7–14 μm, and very long-wave IR ≥ 30 μm.

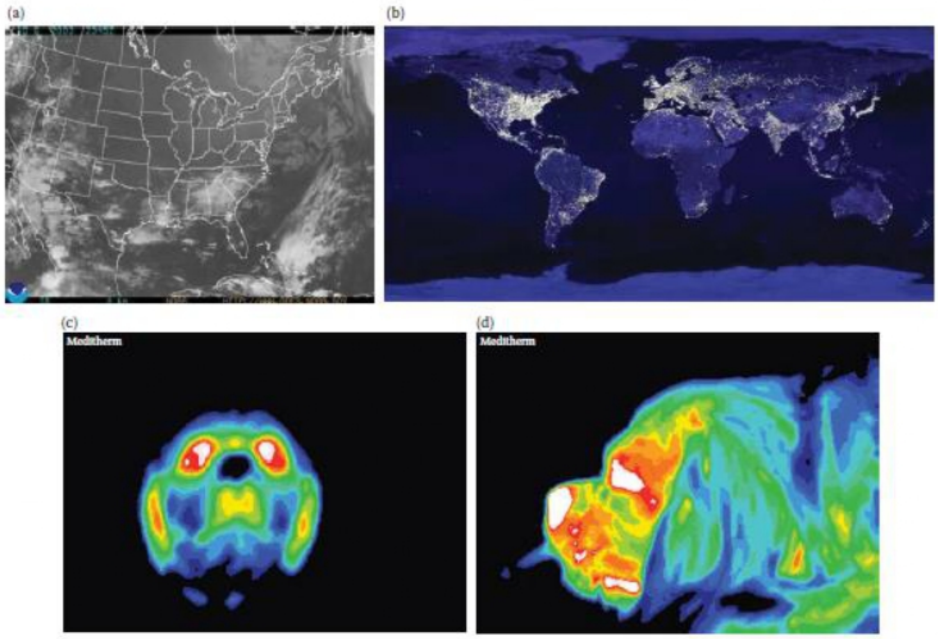


Figure 19. Infrared Images. (a) Infrared satellite image showing water vapor, (b) Infrared satellite imagery in the near infrared band, (c and d) Thermographic images being used to determine their efficacy in diagnosing brain diseases in canines.

Multispectral images, which include IR bands, are used in weather analysis (Figure 20 a). Microwave images are used most often in radar applications, where the primary requirement is the capability to acquire information even through clouds or other obstacles, regardless of lighting conditions. In the radio band of the EM spectrum, applications are primarily in astronomy and medicine.

In medicine, magnetic resonance imaging (MRI) works by sending radio waves through a patient in short pulses in the presence of a powerful magnetic field. The body responds to these pulses by emitting radio waves,which are measured to create an image of any part of the patient's body (Figure 202 b). MRI systems have excellent contrast resolution, multiple images are taken at different angles and assembled to create a 3D image. They are much better at showing subtle differences among the soft tissues and organs of the body that are not easily viewed on conventional x-ray or CT films.

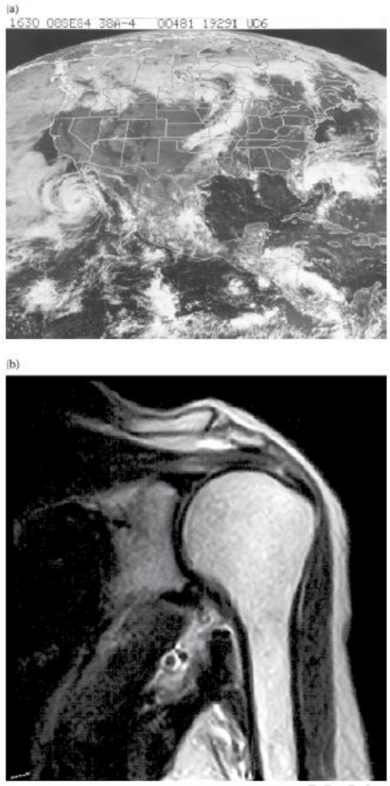


Figure 20. Multispectral and Radio Wave Images. (a) Multispectral Geostationary Operational Environmental Satellite (GOES) image of North America, showing a large tropical storm, (b) magnetic resonance image (MRI) of a patient's shoulder

5.1 Acoustic Imaging

Acoustic imaging operates by sending out pulses of sonic energy (sound) at various frequencies and then measuring the reflected waves. The time it takes for the reflected signal to appear contains distance information, and the amount of energy reflected contains information about the object's density and material. Then the measured information is used to create a two- or three-dimensional image. Acoustic imaging is used in biological systems, for example, bats use it to see, and in man-made systems, such as the sonar used in submarines and medicine. One common use in medicine is to follow the development of the unborn baby inside the womb, the health (and gender) of the baby can be determined (see figure 21.). Because ultrasonic imaging allows us to see inside opaque objects, it is also commonly used in manufacturing to detect defects in materials.

Figure 21. (a) Standard ultrasound image of a baby showing the head, arm, and body, (b) standard ultrasound image showing face and eyes, (c) newer 3D ultrasound image showing baby face and arm, (d) 3D ultrasound of baby yawning

5.2 Electron Imaging

Electron microscopes are used in applications that require extremely high magnification. Standard light microscopes can magnify up to 2000 times, but electron microscopes can magnify up to 10 million times. These microscopes function by producing a focused beam of electrons, which is used to image a specimen similar to the way a light beam is used in a standard microscope.

5.3 Laser Imaging

Lasers are used to create range images, which contain information about the distance of a point in the world (dimension parameters) to the image sensor. (يستعمل لمقدرات المدى)

5.4 Computer-Generated Images

Computers can be used to create images for a myriad of applications, include the engineering, medicine, education, movies, art, and games, and many other applications. Computer graphics images are shown in figure 22 a and b.

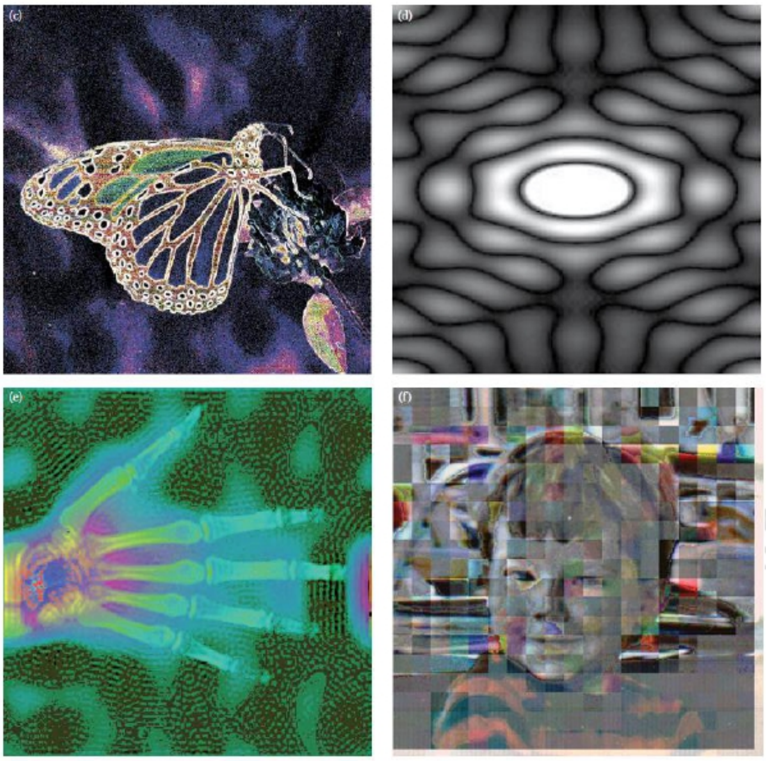


Figure 22. Computer Generated Images. (c) an image of a butterfly processed by edge detection, (d) Fourier transform spectrum image of an ellipse, (e) X-ray image of a hand processed by frequency domain pseudo-color, (f) error image from an image compressed with block truncation coding

6. Image sensing and acquisition

Most of the images in which we are interested are generated by the combination of an illumination source and the reflection of energy from the target (object) by the elements of the scene being imaged.
The illumination may originate from a source of electromagnetic energy, such as a radar, infrared, or X-ray system. But it could originate from other sources, such as ultrasound or even a computer-generated illumination pattern. The reflected energy is focused on a photo converter that converts the energy into visible light. Figure 23 shows the three principal sensor arrangements used to transform incident energy into digital images. The Incoming energy is transformed into a voltage by a combination of the input electrical power and sensor material that is responsive to the type of energy being detected. The output voltage waveform is the response of the sensor, and a digital quantity is obtained by digitizing that response.

The types of components in an image processing system:

- The first is a physical device (sensor) that is sensitive to the energy radiated by the object we wish to convert to an image.
- The second is a digitizer that is used for converting the output of a physical sensing device into digital form.

- الأول هو جهاز مادي (مستشعر)حساس للطاقة التي يشعها الكائن الذي نرغب في تحويله إلى صورة
- الثاني هو جهاز التحويل الرقمي الذي يستخدم لتحويل خرج جهاز الاستشعار المادي إلى شكل رقمي.

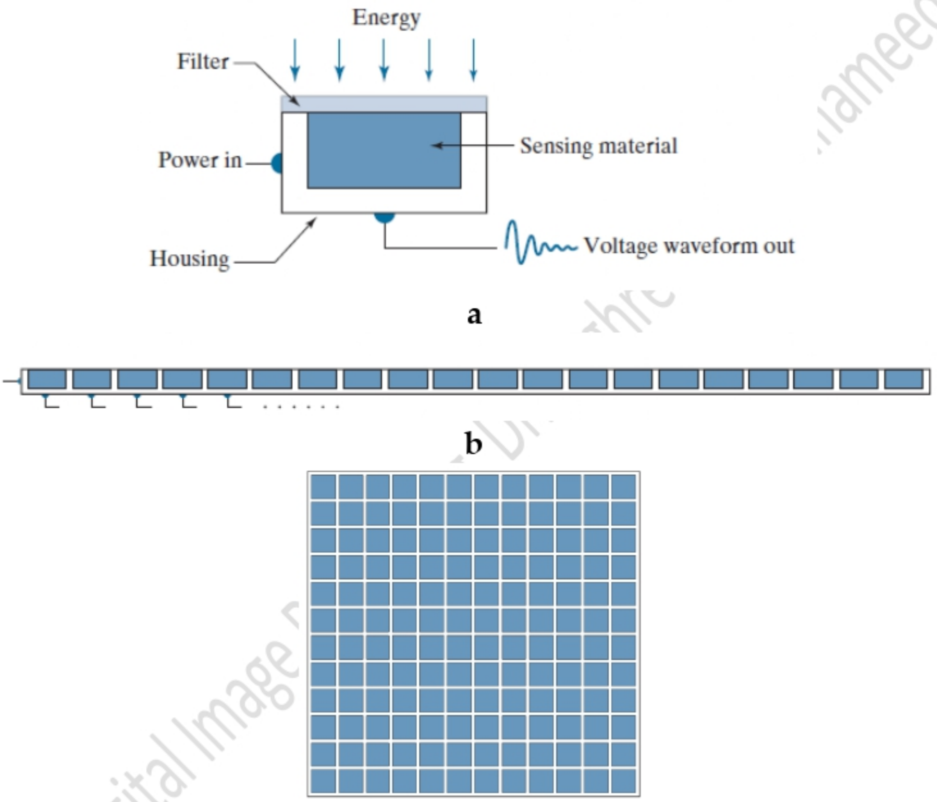


Figure 23. (a) Single sensing element, (b) line sensor, and (c) array sensor

The principal modalities for image sensing and generation are the following:

6.1 Image acquisition using a single sensing element

Figure 23 (a) shows the components of a single sensing element. A familiar sensor of this type is the photodiode, which is constructed of silicon materials and whose output is a voltage proportional to light intensity. Using a filter in front of a sensor improves its selectivity. In order to generate a 2-D image using a single sensing element, figure 24 shows an arrangement used in scanning. **This method is an inexpensive way to obtain high-resolution images because**

mechanical motion can be controlled with high precision. The main disadvantages of this method are that it is slow and not readily portable.

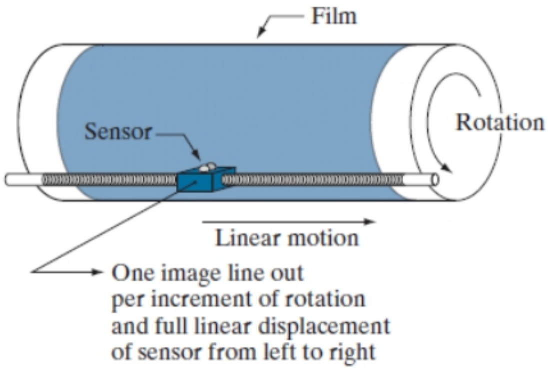


Figure 24. Combining a single sensing element with the mechanical motion to generate a 2-D image.

6.2 Image acquisition using sensor strips

This type used more frequently than single sensors is an in-line sensor strip, as in figure 23 (b). The strip provides imaging elements in one direction.. **With a line scanner, speed and resolution can be increased, while cost is minimized.**

Sensor strips are used in medical and industrial imaging to obtain cross-sectional (slice) images of 3-D objects, as imaging by; tomography (CAT), magnetic resonance, (MRI), and positron emission tomography (PET).

6.3 Image acquisition using sensor arrays

Figure 23 (c) shows individual sensing elements arranged in the form of a 2-D array. Electromagnetic and ultrasonic sensing devices frequently are arranged in this manner. The array sensor is the primary type used in digital cameras, and the typical sensing element is a charge-coupled device (CCD). The sensing substance will output electrons to form electrical energy.

The key advantage of this arrangement is that a complete image can be obtained by focusing the energy pattern onto the surface of the array. Motion is not necessary, as is the case with the other sensor arrangements. Figure 14 shows the principal manner in which array sensors are used. This figure shows the energy from an illumination source is reflected from a scene. The first function performed by the imaging system in figure 25 (c) is to collect the incoming energy and focus it onto an image plane, as in figure 25 (d). Digital and analog circuitry sweep these outputs and convert them to an analog signal, which is then digitized by another section of the imaging system. The output is a digital image, as shown diagrammatically in figure 25 (e).

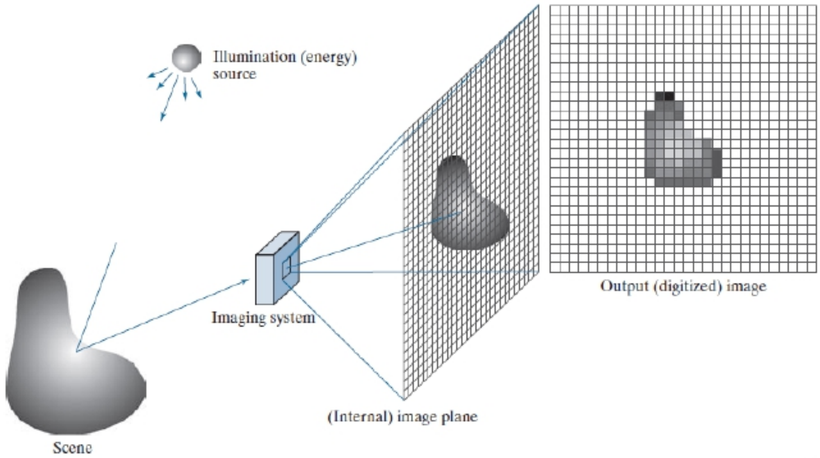


Figure 25. An example of digital image acquisition, (a) illumination (energy) source, (b) a scene, (c) imaging system, (d) projection of the scene onto the image plane, and (e) digitized image

7. A simple image formation model

As a consequence, f (x, y) must be nonnegative and finite; that is:

$$0 \leq f(x, y) < \infty \quad (1)$$

Function f (x, y) is characterized by two components:
(1) The amount of source illumination incident on the scene being viewed.
(2) The amount of illumination reflected by the objects in the scene.

These are called the illumination and reflectance components, and are denoted by i(x, y) and r(x, y), respectively. The two functions combine as a product to form f (x, y):

$$f(x, y) = i(x, y)r(x, y) \quad (2)$$

Where:

$$0 \leq i(x, y) < \infty, \text{ and } \ 0 \leq r(x, y) \leq 1 \quad (3)$$

The reflectance is bounded by 0 (total absorption) and 1 (total reflectance). The nature of $i(x, y)$ is determined by the illumination source, and $r(x, y)$ is determined by the characteristics of the imaged objects.

The intensity (gray level) of a monochrome image at any coordinates (x, y) be denoted by:

$$l = f(x, y) \quad (4)$$

From eqs. (2) through (3) it is evident that l lies in the range;

$$L_{min} \leq l < L_{max} \quad (5)$$

In theory, the requirement on L_{min} is that it be nonnegative, and on L_{max} that it be finite. **In practice**, $L_{min} = i_{\min} r_{\min} =$ and $L_{max} = i_{\max} r_{\max}$.

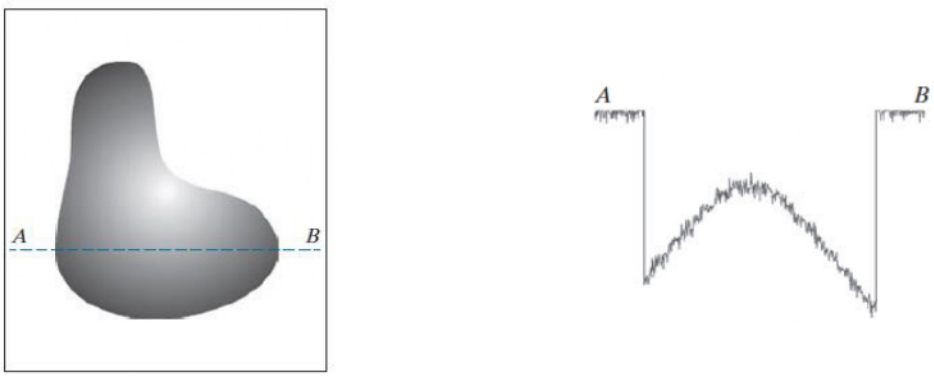
The units of these quantities are lum/m^2. The interval [$L_{\min}$, $L_{\max}$] is called the intensity (or gray) scale. This interval numerically to the interval [0, 1], or [0, C], where l = 0 is considered black and l = 1 (or C) is considered white on the scale. All intermediate values are shades of gray varying from black to white.

8. Image sampling and quantization

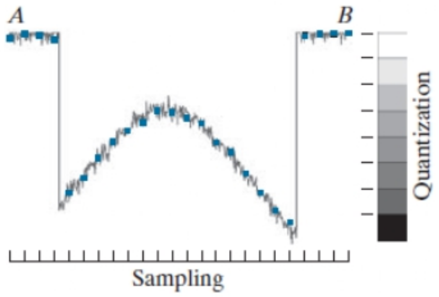
The output of most sensors is a continuous voltage waveform whose amplitude and spatial behavior are related to the physical phenomenon being sensed. To create a digital image, we need to convert the continuous sensed data into a digital format. This requires two processes: sampling and quantization.

8.1 Basic concepts in sampling and quantization

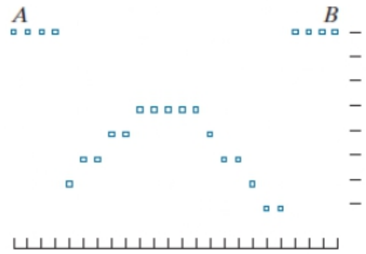
Figure 26 (a) shows a continuous image f with respect to the x- and y-coordinates, and also in amplitude that we want to convert to digital form. To digitize it, we have to sample the function in both coordinates and also in amplitude. Digitizing the coordinate values is called sampling. Digitizing the amplitude values is called quantization.



a b



c



d

Figure 26. (a) Continuous image, (b) a scan line showing intensity variations along line AB in the continuous image, (c) sampling and quantization, and (d) digital scan line (The black border in (a) is included for clarity, it is not part of the image).

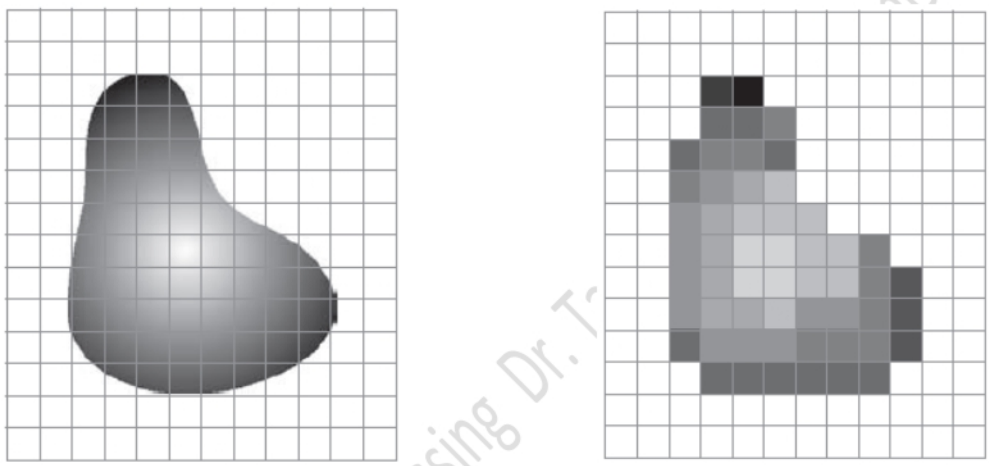
The one-dimensional function in figure 26 (b) is a plot of amplitude (intensity level) values of the continuous image along the line segment AB in figure 26 (a). The random variations are due to image noise. To sample this function, we take equally spaced samples along line AB, as shown in figure 26 (c). The samples are shown as small dark squares superimposed on the function, and their (discrete) spatial locations are indicated by corresponding tick marks at the bottom of the figure. The set of dark squares constitutes the sampled function. 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 vertical gray bar in figure 26(c) depicts 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, depending on the vertical proximity of a sample to a vertical tick mark. The digital samples resulting from both sampling and quantization are shown as white squares in figure 26 (d). Starting at the top of the continuous image and carrying out this procedure downward, line by line produces a two-dimensional digital image.

The accuracy achieved in quantization is highly dependent on the noise content of the sampled signal. The method of sampling is determined by the sensor arrangement used to generate the image.

When an image is generated by a single sensing element combined with mechanical motion. However, spatial sampling is accomplished by selecting the number of individual mechanical increments at which the sensor collects data. Mechanical motion can be very exact so, but it is almost no limit on how fine we can sample an image using this approach. In practice, limits on sampling accuracy are determined by other factors, such as the quality of the optical components used in the system.

When a sensing strip is used for image acquisition, the number of sensors in the strip establishes the samples in the resulting image in one direction, and mechanical motion establishes the number of samples in the other. Quantization of the sensor outputs completes the process of generating a digital image.

When a sensing array is used for image acquisition, no motion is required. The number of sensors in the array establishes the limits of sampling in both directions. Quantization of the sensor outputs is as explained previously. Figure 27 illustrates this concept. Figure 27 (a & b) shows a continuous image projected onto the plane of a 2-D sensor and the image after sampling and quantization. **The quality of a digital image is determined to a large degree by the number of samples and discrete intensity levels used in sampling and quantization.**



a b

Figure 27. (a) Continuous image projected onto a sensor array, and (b) Result of image sampling and quantization.

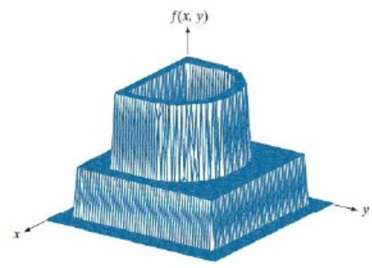
Digitizing the coordinate values is called sampling.
Digitizing the amplitude values is called quantization.

9. Representing digital images

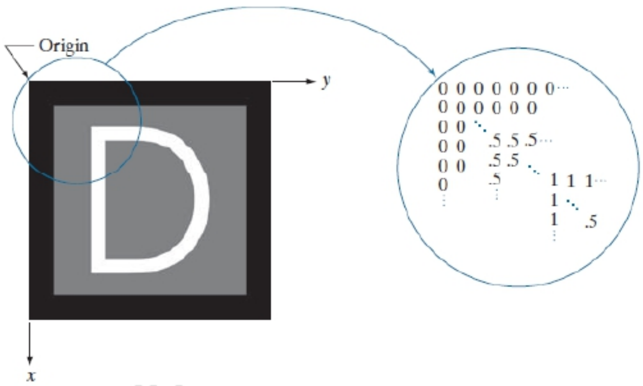
The continuous image sampled into a digital image f (x, y), containing M rows and N columns, where (x, y) are discrete coordinates, with integer values, (x = 0, 1, 2,…,M − 1 and y = 0, 1, 2,…, N − 1). The value of the digital image at the origin is f (0,0), and its value at the next coordinates along the first row is f (0,1). Here, the notation (0, 1) is used to denote the second sample along the first row. The value of a digital image at any coordinates (x, y) is denoted f (x, y), where x and y are integers. The real plane spanned by the coordinates of an image is called the **spatial domain**, with x and y being referred to as spatial variables or spatial coordinates. The equation form can be represented of an (M × N) numerical array as:

$$f(x,y)=\begin{bmatrix} f(0,0) & f(0,1) & \cdots & f(0,N-1) \\ f(1,0) & f(1,1) & \cdots & f(1,N-1) \\ \vdots & \vdots & & \vdots \\ f(M-1,0) & f(M-1,1) & \cdots & f(M-1,N-1) \end{bmatrix} \tag{6}$$

The right side of this equation is a digital image represented as an array of real numbers. Each element of this array is called an image element, picture element, pixel, or pel.



a



b c

Figure 28. (a) Image plotted as a surface, (b) image displayed as a visual intensity array, and (c) image shown as a 2-D numerical array (The numbers 0, .5, and 1 represent black, gray, and white, respectively.)

Image digitization requires that the values of M, N, and the number L of discrete intensity levels have to be positive integers. The digital storage leads to the number of intensity levels L, being an integer power of two, that is:

$$L = 2^k \tag{7}$$

Where k is an integer. The discrete levels are equally spaced and they are integers in the range [0, L − 1]. **The range of values spanned by the grayscale is referred to as the dynamic range**. The **dynamic range** of an imaging system is defined to be the ratio of the maximum measurable intensity to the minimum detectable intensity level in the system. The upper limit is determined by saturation and the lower limit by noise (see figure 29). The dynamic range establishes the lowest and highest intensity levels that a system can represent and, consequently, that an image can have. Closely associated with this concept is **image contrast**, which we define as the difference in intensity between the highest and lowest intensity levels in an image. When an appreciable number of pixels in an image has a high dynamic range, we can expect the image to have high contrast. Conversely, an image with a low dynamic range typically has a dull, washed-out gray look. The number, b, of bits required to store a digital image is:

$$b = M \times N \times k \tag{8}$$

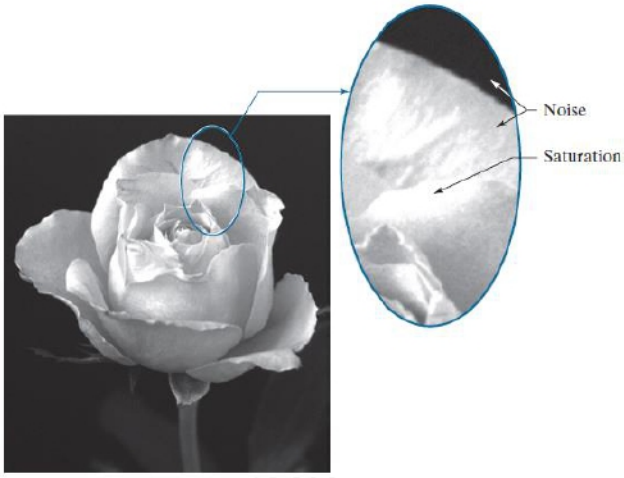


Figure 29. An image exhibiting saturation and noise.

What mean the Gray Level?

Gray level refers to a scalar measure of intensity that ranges from black to grays and finally to white.

10. Spatial and intensity resolution

The **resolution** is of two types: spatial and intensity (level Gray level).

1. **Spatial resolution**

It is the smallest discernible detail in an image. It is determined by the sampling process. The spatial resolution of a digital image reflects the amount of details that one can see in the image (i.e. the ratio of pixel size to the size of the image display). If an image is spatially sampled at (M×N) pixels, then the larger (M×N) the finer the observed details.

Measuring spatial resolution:

Since the spatial resolution of the images refers to the clarity, so in the various devices there are several measures that we can rely on in measuring the spatial resolution, which are:

- Lines per unit distance or per inch: used in laser printers.
- Dots (pixels) per unit distance or dots (pixels) per inch (dpi): used in screens (monitors) and mobile devices.

A widely used definition of image resolution is the largest number of discernible line pairs per unit distance (e.g.,100 line pairs per mm).
Dots per unit distance is a measure of image resolution used in the printing and publishing industry. this measure usually is expressed as dots per inch (dpi). To give you an idea of quality, newspapers are printed with a resolution of 75 dpi, magazines at 133 dpi, glossy brochures at 175 dpi, and the book page at which you are presently looking was printed at 2400 dpi.
The measures of spatial resolution must be stated with respect to spatial units. Image size does not gave the complete truth. For example, to say that an image has a resolution of (1024 × 1024) pixels is not a meaningful statement without stating the spatial dimensions encompassed by the image.
Image size is helpful only in making comparisons between imaging capabilities. For instance, a digital camera with a 20-megapixel CCD imaging chip can be expected to have a higher capability to resolve detail than an 8-megapixel camera, assuming that both cameras are equipped with comparable lenses and the comparison images are taken at the same distance.

When the spatial resolution of the digital image is reduced, the degradation is visible in most features of the image. It is possible to improve the result by the choice of interpolation methods used.

<u>Intensity resolution (gray level)</u>

- It refers to the smallest discernible change in gray level. It is determined by the quantization process. The number of gray levels is usually an integer power of 2. The most common number is 8 bits, however, 16 bits is used in some applications where enhancement of specific gray-level ranges is necessary.
- Represents the number of bits needed to represent a single pixel of an image Bit Per Pixel
- Assuming the number of bits K then the number of possible levels of gray in the image is 2^k, whereas the range of gray levels is $[0\text{-}(2^k\text{-}1)]$.
- Assuming that the image is represented by 8 bits, it means that the range of gray levels is [0-255].
- If a few bits are used to represent the pixels, we get a false contour problem, which means poor image colors (poor colors).
- The discernible changes in intensity are influenced also by noise and saturation values, and by the capabilities of human perception to analyze and interpret details in the context of an entire scene.

False contouring effect: is caused using an insufficient number of intensity levels in smooth areas of a digital image. It appears in the image as false edges, or lines as a result of the gray level quantization method.

These <u>two parameters</u> (**Spatial and intensity resolution**) interact in **determining perceived image quality**.

11. Image Quantization

Image quantization is the process of reducing the image data by removing some of the detailed information by mapping a group of data points to a single point. This can be done by:

1. Gray Level reduction (reduce pixel values themselves I(r, c).
2. Spatial reduction (reduce the spatial coordinate (r, c).

(التكميم يعني حذف معلومة لان بدون تكميم تصبح اشارة مستمرة)

• Gray Level Quantization (Reduction)

The simplest method of gray level reduction is **Thresholding**. We select a threshold gray level and set everything above that value equal to "1" (255 for 8-bit data), and everything below the threshold equal to "0". This effectively turns a gray_level image into a binary (two-level) image and is often used as a preprocessing step in the extraction of object features, such as shape, area, or perimeter.

A more method (هناك طريقة أخرى لتقليل مستوى اللون الرمادي) of gray level reduction is the process of taking the data and reducing the number of bits per pixel. This can be done very efficiently by masking the lower bits via an AND operation. Within this method, the numbers of bits that are masked determine the number of gray levels available.

Example:

We want to reduce 8_bit information containing 256 possible gray_level values down to 32 possible values. This can be done by ANDing each 8-bit value with the bit string **11111000**.

This is equivalent to dividing by eight (2^3), corresponding to the lower three bits that we are masking and then shifting the result left three times. [Gray _level in the image 0-7 are mapped to 0, gray_level in the range 8-15 are mapped to 8, and so on].We can see that by masking the lower three bits we reduce 256 gray levels to 32 gray levels:

$$256 \div 8 = 32$$

The general case requires us to mask k bits, where 2^k is divided into the original gray-level range to get the quantized range desired. Using this method, we can reduce the number of gray levels to any power of 2: 2,4,6,8, 16, 32, 64, or 128.

- Image quantization by masking to 128 gray levels, this can be done by ANDing each 8-bit value with bit string 11111110 (2^1).
- Image quantization by masking to 64 gray_level. This can be done by ANDing each 8-bit value with bit string 11111100 (2^2).

The AND based method maps the quantized gray-level values to the low end of each range; alternately, if we want to map the quantized gray-level values to the high end of each range, we use an OR operation. The number of "1" bits in the OR mask determine how many quantized gray levels are available.

As the number of gray levels decreases, we can see an increase in a phenomenon called contouring. Contouring appears in the image as false edges, or lines as a result of the gray _level quantization method.

Figure 30. **False Contouring**. (a) Original 8-bit image, 256 gray levels, (b) quantized to 7 bits, 128 gray levels, (c) quantized to 6 bits, 64 gray levels, (d) quantized to 5 bits, 32 gray levels, (e) quantized to 4 bits, 16 gray levels, (f) quantized to 3 bits, 8 gray levels, (g) quantized to 2 bits, 4 gray levels, and (h) quantized to 1 bit, 2 gray levels.

This **false contouring effect** can be visually improved upon by using an **IGS** (**improved gray-scale**) **quantization** method. In this method (IGS) the improvement will be by adding a small random number to each pixel before quantization, which results in a more visually pleasing appearance.

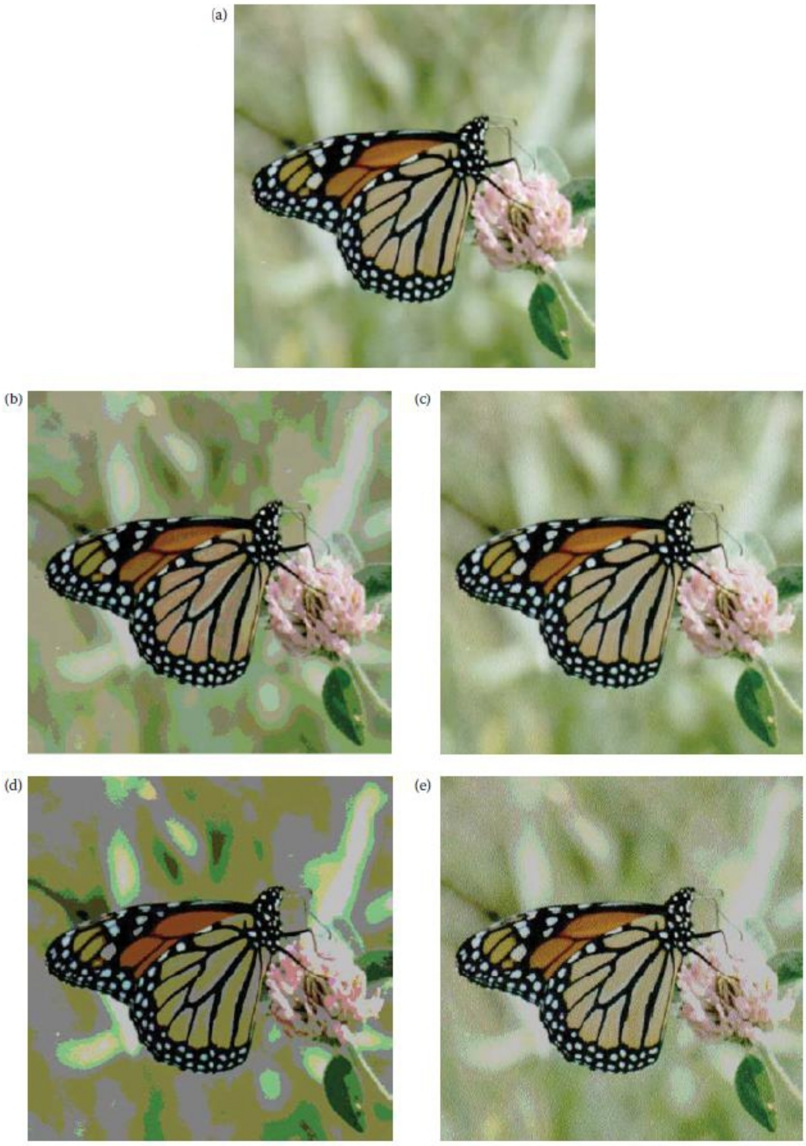


Figure 31. **IGS Quantization.** (a) Original image, (b) uniform quantization to 8 gray levels per color band, (c) IGS quantization to 8 gray levels per band, (d) uniform quantization to 4 gray levels per color band, and (e) IGS quantization to 4 gray levels per band

Improved Gray Scale (IGS) quantization method: is used to improve the false contouring effect can be visually upon by adding a small random number to each pixel before quantization, which results in a more visually pleasing appearance.

- ## Spatial Quantization

Quantization of the spatial coordinates, **spatial quantization**, results in **reducing the actual size of the image**. This is accomplished by taking groups of pixels that are spatially adjacent and mapping them to one pixel. This can be done in one of three ways:

(1) Averaging,
(2) Median,
(3) Decimation.

For the first method, averaging, the mean value of all the pixels in each group is found by summing the values and dividing by the number of pixels in the group. With the second method, median, the pixel values are sorted from lowest to highest and then the middle value is selected. The third approach, decimation, also known as (subsampling), entails simply eliminating some of the data. For example, to reduce the image by a factor of 2, every other row and column is removed.

The averaging method result causes blurs the image. The median and decimation methods produce some visible artifacts. To improve the image quality when applying the decimation technique, we use averaging spatial filter, this type of filtering is called (anti-aliasing) filtering on the original image before using decimation method.

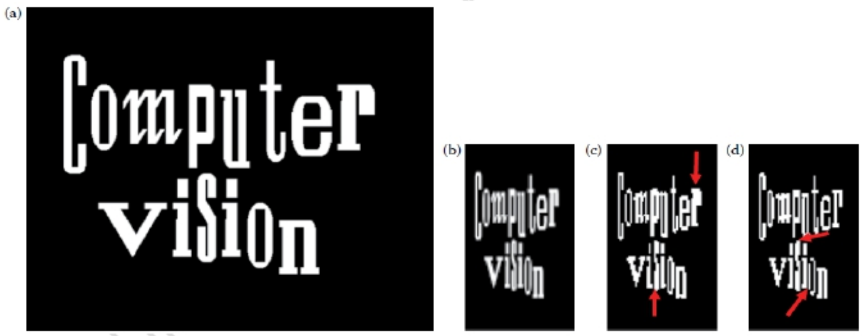


Figure 32. Spatial Reduction. (a) Original 512×512 image, (b) spatial reduction to 64×128 via averaging, (c) spatial reduction to 64×128 via median method, note the space in the “s” and “r” is filled in, and (d) spatial reduction to 64×128 via decimation method, note the “o” is split and “s” is filled in.

Figure 33. Decimation and Anti-aliasing Filter. (a) Original 512×512 image, (b) result of spatial reduction to 128×128 via decimation, and (c) result of spatial reduction to 128×128 via decimation, but the image was first preprocessed by a 4×4 averaging filter, an anti-aliasing filter. Note that the “s” is still clear and the “o” is not so jagged.

12. Image interpolation

Interpolation is used in tasks such as zooming, shrinking, rotating, and geometrically correcting digital images. Our principal objective is to introduce interpolation and apply it to image resizing (shrinking and zooming), which are basically image resampling methods.
Interpolation is the process of using known data to estimate values at unknown locations.

Interpolation methods:

1. Nearest neighbor interpolation
2. Bilinear interpolation
3. Bicubic interpolation

Zooming may be said oversampling and shirking may be called as under sampling these techniques are applied to a digital image. These are two steps of zooming;
i) Creation of new pixel locations
ii) Assignment of gray level to those new locations.

In order to perform intensity assignment for any point in the overly, we look for the closet pixel in the original image and assign its intensity to the new pixel in the grid. This method is named as nearest neighbor interpolation. It assigns to each new location the intensity of its nearest neighbor in the original image

Pixel replication: is a special case of nearest neighbor interpolation, it is applicable if we want to increase the size of an image an integer number of times.

For increasing the size of an image as double, we can duplicate each column. This doubles the size of the image horizontal direction. To increase the assignment of each vertical direction we can duplicate each row. The gray level assignment of each pixel is determined by the fact that new locations are exact duplicates of old locations.

Drawbacks

Although this approach is simple and fast, it has the tendency to produce undesirable features (artifacts) that it produces a severe distortion of straight edges or the so-called (checkboard) that is not desirable.
A more suitable approach is bilinear interpolation, in which we use the four nearest neighbors to estimate the intensity at a given location. Bilinear interpolation gives much better results than nearest neighbor interpolation, with

a modest increase in computational burden. The next level of complexity is a bicubic interpolation, which involves the sixteen nearest neighbors of a point. Bicubic interpolation does a better job of preserving fine detail than its bilinear counterpart. It is the standard used in commercial image editing applications, such as Adobe Photoshop.

Shrinking is done in a similar manner. The equivalent process of pixel replication is row/column deletion. Shrinking leads to the problem of aliasing.

13. Some basic relationships between pixels

Content:

1. Digital path
2. Regions
3. Boundaries.
4. Distance Measures

Several important relationships between pixels in a digital image are considered.

1. Neighbors of a Pixel

A pixel p at coordinates (x, y) has the following neighbors:

1. N_4 (p): 4-neighbors of p.

- Any pixel p (x, y) has two vertical and two horizontal neighbors, given by:

  (x+1,y), (x-1, y), (x, y+1), (x, y-1)
- This set of pixels are called the 4-neighbors of P, and is denoted by N_4(P)
- Each of them is at a unit distance from P.

2. N_D(p)

- This set of pixels, called 4-neighbors and denoted by N_D (p).
- N_D(p): four diagonal neighbors of p have coordinates:

  (x+1,y+1), (x+1,y-1), (x-1,y+1), (x-1,y-1)
- Each of them are at Euclidean distance from P.

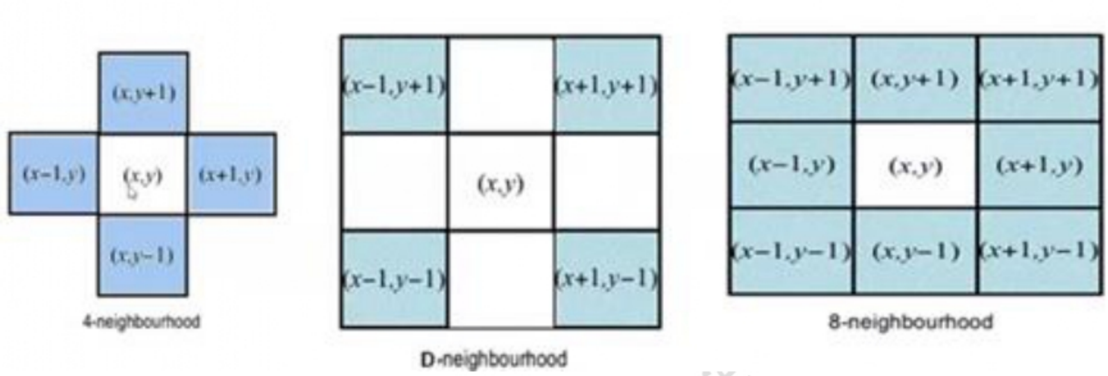
3. N_8 (p): 8-neighbors of p.

- N_4 (P) and N_D (p) together are called 8-neighbors of p, denoted by **N_8**(p).
- $N_8 = N_4 \cup N_D$

Note: Some of the points in the N4 , ND and N8 may fall outside image when P lies on the border of image..

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

N_8 (p)



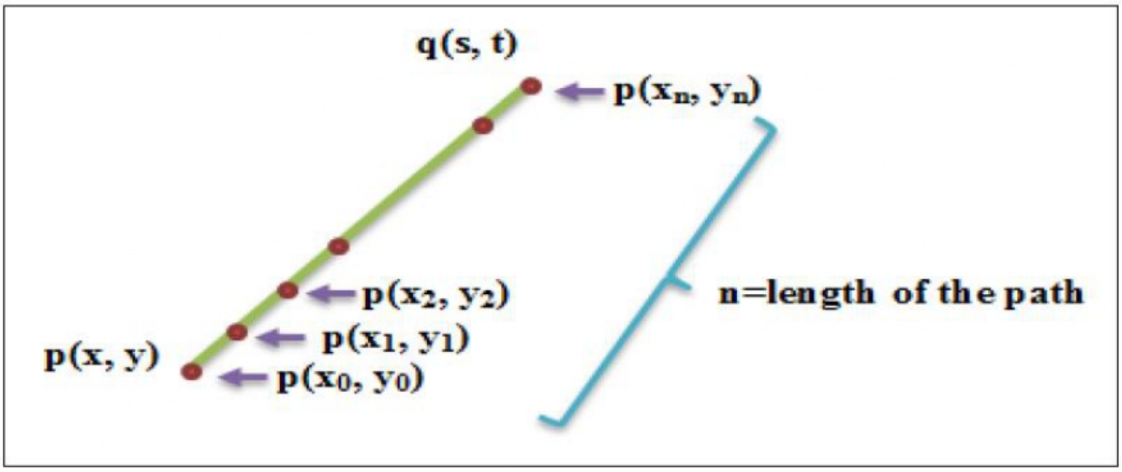
Digital path

Path

A (digital) path (or curve) from pixel p with coordinates (x_0, y_0) to pixel q with coordinates (x_n , y_n) is a sequence of distinct pixels with coordinates (x_0, y_0), (x_1, y_1), …, (x_n ,y_n)

Here n is the length of the path.

We can define 4-, 8-, and m-paths based on the type of adjacency used.



Example : consider the image segment shown in figure (a). Compute length of shortest -4, shortest -8, shortest –m paths between pixels p & q where v={1,2}

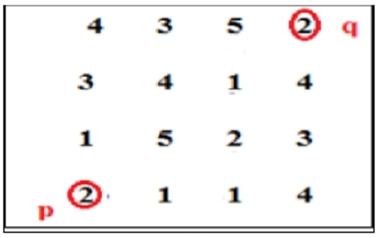
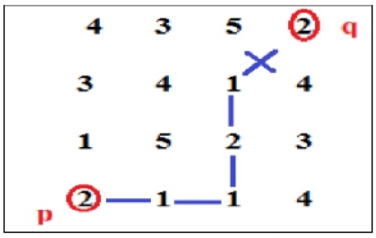
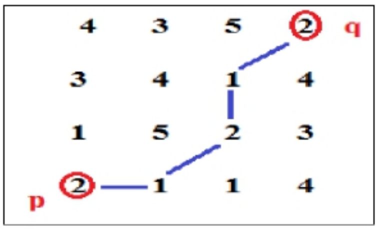


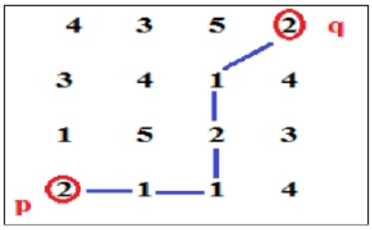
figure (a)



Shortest-4 Path does not exist



shortest-8 path = 4



shortest–m Path = 5

Example : consider the image segment shown in figure (a). Compute length of shortest -4, shortest -8, shortest –m paths between pixels p & q where v={2, 3}

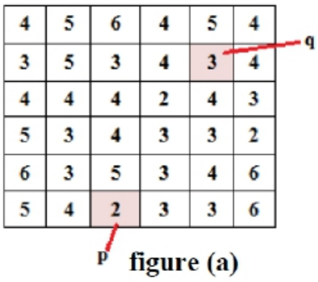
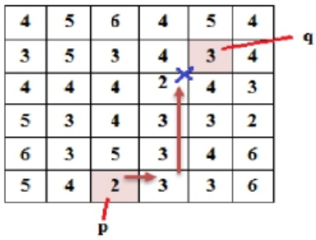
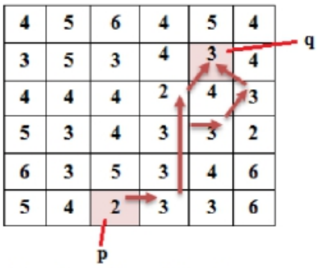


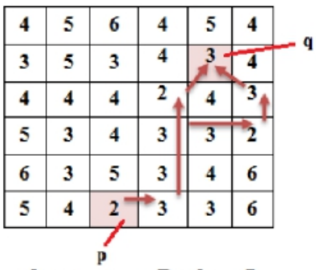
figure (a)



Shortest-4 Path does not exist



shortest-8 path = 5
= 6



shortest–m Path = 5
=7

Region

Connectivity

Connectivity: **two pixels are said to be connected** if their exists a path between them.

Let 'S' represent subset of pixels in an image.

Two pixels **p & q** are said to be connected in 'S' if their exists a path between them consisting entirely تماما of pixels in 'S'.

Connected component and region

Let S represent a **subset of pixels in an image** • For every pixel **p** in **S**, the set of pixels in S that are connected to p is called a connected component of S.

- If S has only one connected component, then S is called Connected Set.
- We call R a region of the image if R is a connected set.
- Two regions, Ri and Rj are said to be adjacent if their union forms a connected set.
- Regions that are not to be adjacent are said to be disjoint.

Let *R* be a subset of pixels in an image, we call *R* a region of the image if *R* is a connected set.

000000
010010
011010
010110
000000

Regions and boundaries

a) Boundary (or border)

The **boundary** of the region **R** is the set of pixels in the region that have **one** or **more neighbors** that are **not** in R.

If R happens to be an entire image, then its boundary is defined as the set of pixels in the first and last rows and columns of the image.

b) Foreground and background

An image contains K disjoint regions, R_k , k = 1, 2, …, K.

Let R_u denote the union of all the K regions, and let $(R_u)^c$ denote its complement.

All the points in Ru is called foreground;

All the points in $(R_u)^c$ is called background.

Distance measures

Assuming there are three image points (p, q, and z), with coordinates (x, y), (s, t), and (v, w). A distance measure is normally conducted for evaluating how close these three pixels are and how they are related. D is a distance function or metric if:

a) *D (p, q) ≥ 0 (D (p, q) = 0 iff p = q)),*

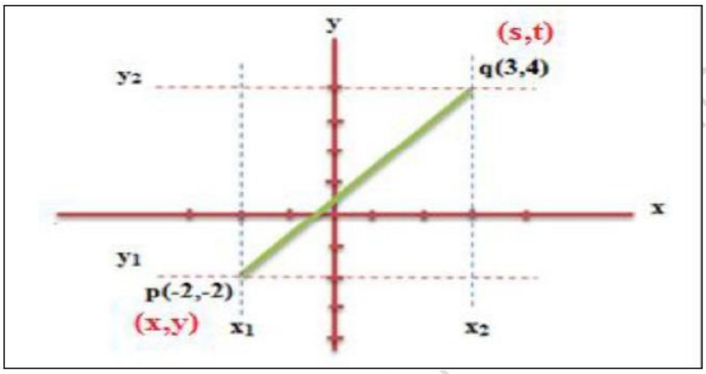
b) *D (p, q) = D (q, p), and*

c) *D (p, z) ≤ D (p, q) + D (q, z).*

A number of distance measurements have been commonly used for this purpose:

- **Euclidean distance** between two 2-D points p(x , y)and q(s , t) is defined as:

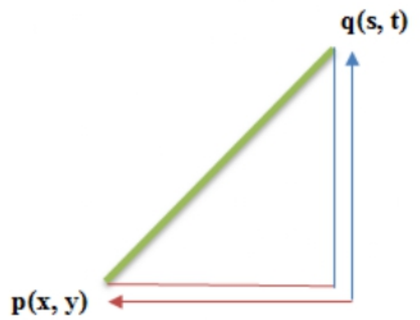
$$D(p,q) = \sqrt{(x-s)^2 + (y-t)^2} \text{.......(1)}$$



Where: $x=x_1$, $y=y_1$ $s=x_2$, $t=y_2$

- **City-block distance** between two 2-D points p(x, y)and q(s, t)can be calculated as follows:

$$D(p,q) = |x-s| + |y-t| \text{.....................(2)}$$



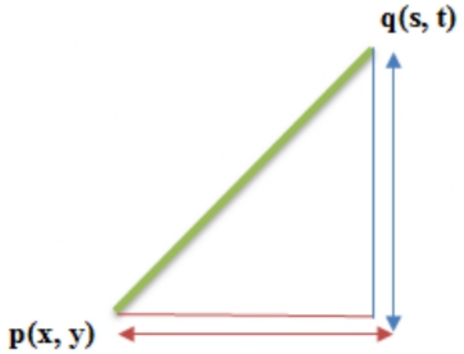
Example: The pixels with D_4 distance ≤**2** from (x, y) (the center point) from the following contours of constant distance:

2	1	2
1	1	1
2	1	2

The pixels with D_4 =1 are the 4-neighbors of (x, y).

- **<u>Chessboard distance:</u>** The D_8 distance (called the chessboard distance) between p and q is defined as:

$$D(p,q) = max(|x-s|,|y-t|)\ldots\ldots\ldots\ldots\ldots(3)$$



In this case, the pixels having a D_8 distance from (x, y) less than or equal to some value r from a square centered at (x, y).

<u>Example</u>: The pixels with D8 distance ≤ from (x, y) (the center point) from the following contours of constant distance:

2	2	2	2	2
2	1	1	1	2
2	1	1	1	2
2	1	1	1	2
2	2	2	2	2

The pixels with D_8 =1 are the 8-nieghbors of (x, y).

14. Image Representation

The digital image, I(r,c), is represented as a two-dimensional array of data, where each pixel value corresponds to the brightness of the image at the point (r,c). A two-dimensional array like our image model, I(r,c), is referred to as a matrix, and one row or column is called a vector. This image model is for monochrome ("one color"), or grayscale, image data, but there are other types of image data. Typically, these are multiband images, such as color and multispectral, and they can be modeled by a different I(r,c) function corresponding to each separate band of brightness information.

The image types are: (1) binary, (2) grayscale, (3) color, and (4) multispectral.

14.1 Binary Images

Binary images are the simplest type of images and can take on two values, typically black and white, or "0" and "1." A binary image is referred to as a 1 bit per pixel image, because it takes only 1 binary digit to represent each pixel. Binary images are often created from grayscale images via a threshold operation, where every pixel above the threshold value is turned white ("1"), and those below it are turned black ("0"). Although in this process much information is lost, the resulting image file is much smaller making it easier to store and transmit.

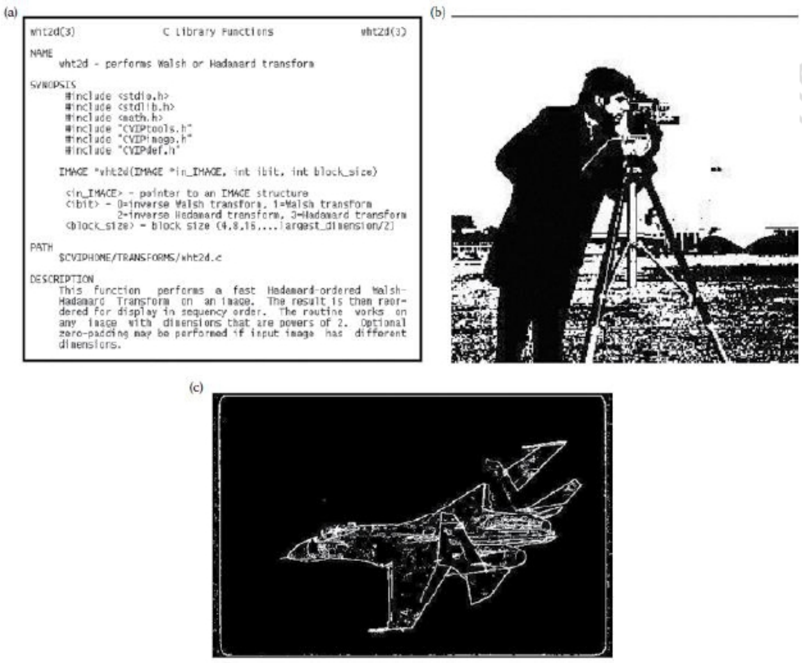


Figure 34. Binary Images. (a) Binary text, (b) image created by a threshold operation, (c) edge detection and threshold operation.

14.2 Grayscale Images

Grayscale images are referred to as monochrome ("one color") images. They contain brightness information only, no color information. The number of bits used for each pixel determines the number of different brightness levels available. The typical image contains 8 bits per pixel data, which allows for 256 different brightness (gray) levels, values of 0–255. Additionally, the 8-bit representation is typical due to the fact that the byte, which corresponds to 8-bits of data, is the standard small unit in the world of digital computers. Figure 35 shows typical monochrome, grayscale, or gray-level, images. In applications requiring higher brightness resolution, such as medical imaging or astronomy, 12 or 16 bits per pixel representations are used. These extra brightness levels only become useful when the image is zoomed, that is, a small section of the image is made much larger.

The small details may be discernable that would be missing without this additional brightness resolution. to be useful, this also requires a higher level of spatial resolution, which means more samples for the same area and more pixels.

The light energy is typically divided into different bands, where each band refers to a specific subsection of the visible image spectrum.

Figure 35. Monochrome, Grayscale or Gray-Level Images

14.3 Color Images

Color images can be modeled as three-band monochrome image data, where each band of data corresponds to a different color. The actual information stored in the digital image data is the brightness information in each spectral band. Typical color images are represented as RGB images. Using the 8-bit monochrome standard as a model, the corresponding color image has 24 bits per pixel 8 bits for each of the three color bands, RGB. Figure 36 shows the three individual color bands of a typical RGB color image. Figure 36 b shows the image bands combined to create the color image, and Figure 36 c illustrates that, in addition to referring to a row or column of an image as a vector, a single pixel's RGB values are referred to as a color pixel vector (R,G,B). For many applications, RGB color information is transformed into a mathematical space that decouples the brightness information from the color information, referred as a color model, a color transform, or mapping into another color space. Where the image information consists of a one-dimensional brightness, or luminance space, and a two-dimensional color space. The two-dimensional color space does not contain any brightness information, but typically contains information regarding the relative amounts of the different colors.

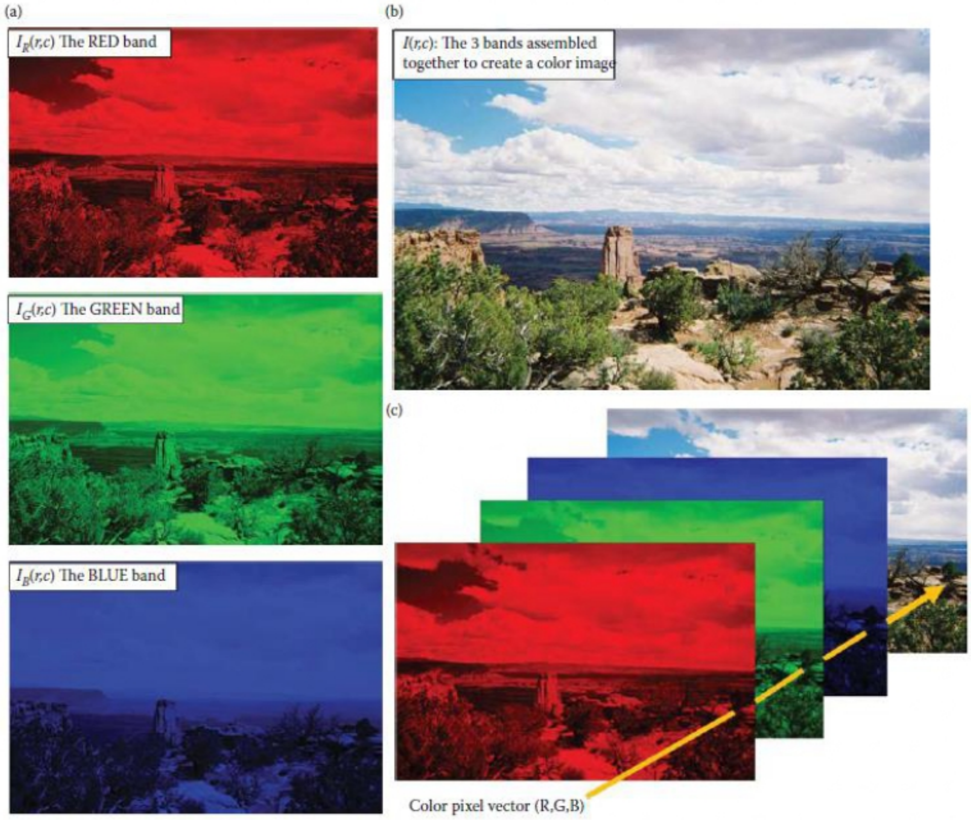


Figure 36. Color image representation.

The Hue/Saturation/Lightness (HSL) color transform allows for a description of colors (see figure 37). The lightness, also referred to as intensity or value, is the brightness of the color, and the hue is a color; for example, green, blue, or orange. The saturation is a measure of the amount of white in the color; for example, pink is red with more white, so it is less saturated than pure red. Most people can relate to this method of describing color, for example, a deep, bright orange would have a large intensity (bright), a hue of orange, and a high value of saturation (deep). We can picture this color in our minds, but if the color is defined in terms of its RGB components, R = 245, G = 110, and B = 20, most people have no idea how this color appears. As the HSL color space was developed based on heuristics relating to human perception, various algorithmics are available to transform RGB pixel values into the HSL color space. Most of these are algorithmic in nature and are geometric approximations for mapping the RGB color cube into some HSL color space (see figure 38). Equations for mapping RGB to HSL are given below. These equations assume that the RGB values are normalized to lie between 0 and 1. The normalization is often done by dividing the RGB values by their sum, but other normalization methods are possible; for example, dividing by the maximum of (R,G,B). The max and min values in the equations below are respectively the largest and smallest of the RGB normalized values.

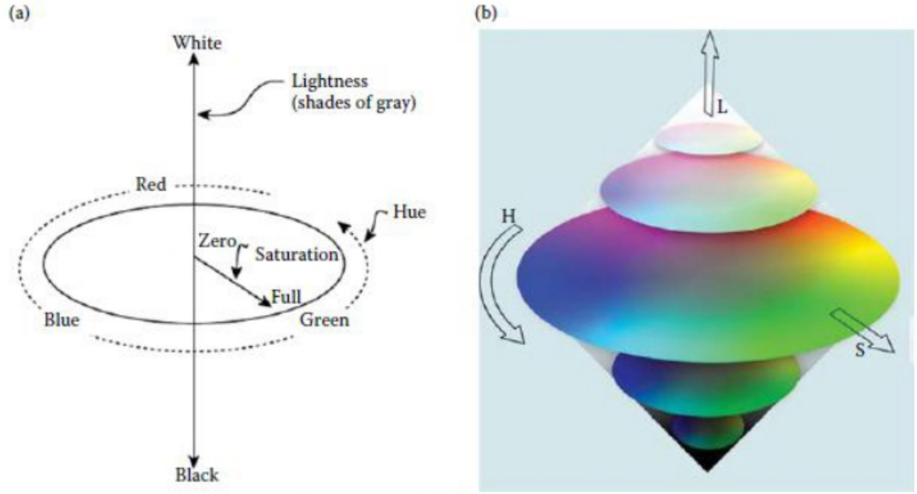


Figure 37. Hue/Saturation/Lightness (HSL) Color Space. (a) Schematic representation of the HSL color space, (b) color representation of the HSL color space.

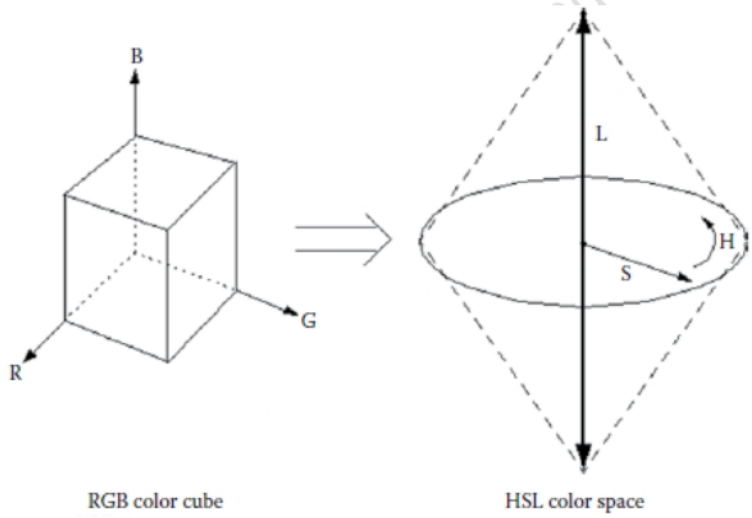


Figure 38. Red/Green/Blue to Hue/Saturation/Lightness mapping.

$$Hue = \begin{cases} 0 & \text{if } max = min \\ 60^\circ \times \dfrac{g-b}{max-min} + 360^\circ & \text{if } max = r \\ 60^\circ \times \dfrac{b-r}{max-min} + 120^\circ & \text{if } max = g \\ 60^\circ \times \dfrac{r-g}{max-min} + 240^\circ & \text{if } max = b \end{cases}$$

$$Lightness = L = \frac{1}{2}(max + min)$$

$$Saturation = \left\{ \begin{array}{ll} 0 & \text{if } max = min \\ \dfrac{max-min}{max+min} = \dfrac{max-min}{2L} & \text{if } L \leq 1/2 \\ \dfrac{max-min}{2-(max+min)} = \dfrac{max-min}{2-2L} & \text{if } L > 1/2 \end{array} \right\}$$

If the maximum and minimum RGB values are equal, then the concepts of hue and saturation are not meaningful because the color is gray, it is essentially a monochrome pixel, so hue and saturation are set equal to 0. In other cases, the value of hue is usually modulo 360°, to lie between 0° and 360°.

14.4 Multispectral Images

Multispectral images typically contain information outside the normal human perceptual range. They may include IR, UV, x-ray, microwave, other bands in the EM spectrum, laser, or acoustic signals. These are not images in the usual sense, as the information represented is not directly visible by the human visual system. However, the information is often represented in visual form by mapping the different spectral bands to RGB components. Sources for these types of images include: satellite systems, sonar systems, various types of airborne radar, IR imaging systems, and medical diagnostic imaging systems. The number of bands into which the data is divided is strictly a function of the sensitivity of the imaging sensors used to capture the images. In the visible spectrum can be divided into three bands because this mimics the human visual system. The older satellites currently in orbit collect image information in seven to nine spectral bands; typically, three are in the visible spectrum and one or more in the IR region, and some have sensors that operate in the microwave range. The newest satellites have sensors that collect image information in 30 or more bands.

15. Digital Image File Formats

Many different types of image file formats are according to applications with varying requirements. The standard file formats are widely available. The bitmap (BMP) format is commonly used in Microsoft Windows based machines. Most imaging and graphics programs in this environment support the BMP format. Another commonly used format is JPEG (Joint Photographic Experts Group). This file format is capable of high degrees of image compression, so is typically used on the internet to reduce bandwidth requirements, meaning you do not need to wait forever for images to appear.

In computer graphics, types of image data are divided into two primary categories: bitmap and vector. Bitmap images, or raster images, can be represented by our image model, I(r,c), where the pixel data and the corresponding brightness values are stored in a specified file format. Vector images refer to methods of representing lines, curves, and shapes by storing only the key points. These key points are sufficient to define the shapes, and the process of turning these into an image is called **rendering**. These types of images contain both **header information** and the pixel data itself. The image file header is a set of parameters normally found at the start of the file and must contain information regarding: (1) the number of rows, height, (2) the number of columns, width, (3) the number of color or spectral bands, (4) the number of bits per pixel, and (5) the file type. Additionally, with some of the more complex file formats, the header may contain information about the type of compression used and any other necessary parameters to create the image, I(r,c).