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.

(1) 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.

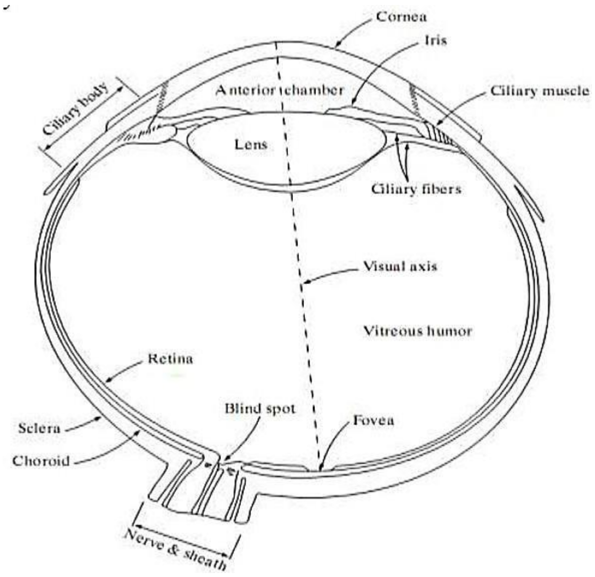


Fig 3.1 Simplified diagram of a cross section of the human eye

Source: Rafael C. Gonzalez, Richard E. Woods, Digital Image Processing‘, Pearson, Third Edition, 2010., Page-36

Image Formation in the Eye:

The principal difference between the lens of the eye and an ordinary optical lens is that the former is flexible. As illustrated in Fig. 1.4, the radius of curvature of the anterior surface of the lens is greater than the radius of its posterior surface. The shape of the lens is controlled by tension in the fibers of the ciliary body. To focus on distant objects, the controlling muscles cause the lens to be relatively flattened. Similarly, these muscles allow the lens to become thicker in order to focus on objects near the eye. The distance between the center of the lens and the retina (called the focal length) varies from approximately 17 mm to about 14 mm, as the refractive power of the lens increases from its minimum to its maximum.

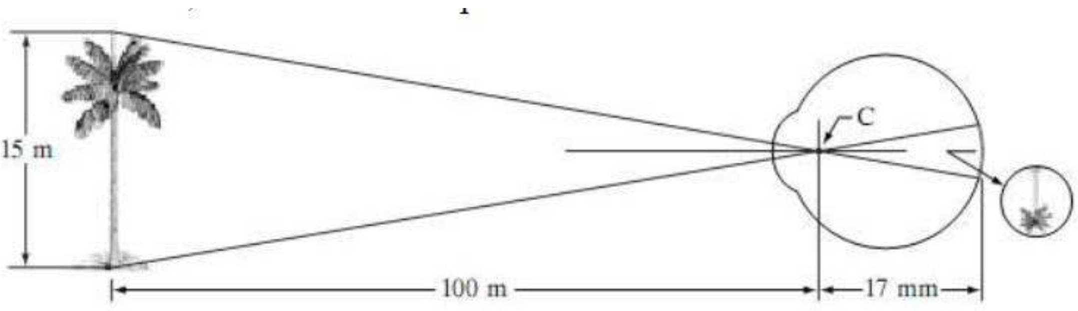


Fig : 3.2Graphical representation of the eye looking at a palm tree Point C is the optical center of the lens.

(Source: Rafael C. Gonzalez, Richard E. Woods, D igital Image Processing‘, Pearson, Third Edition, 2010., Page No. -38)

When the eye focuses on an object farther away than about 3 m, the lens exhibits its lowest refractive power. When the eye focuses on a nearby object, the lens is most strongly refractive. This information makes it easy to calculate the size of the retinal image of any object , 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

(2)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.

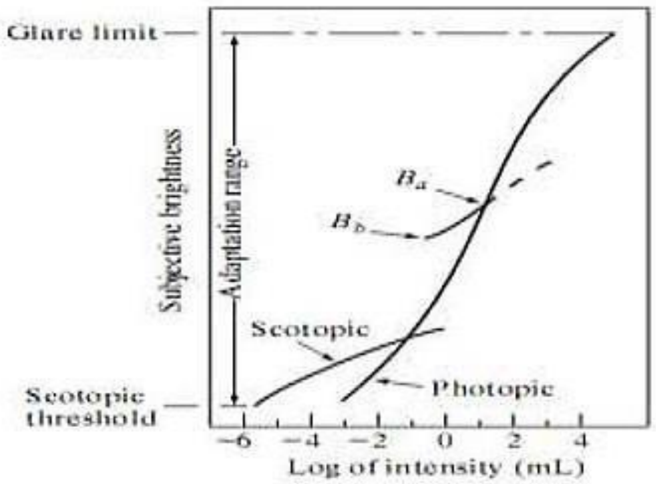


Fig 3.3: Range of subjective brightness sensations showing a particular adaptation level

(Source: Rafael C. Gonzalez, Richard E. Woods, D igital Image Processing‘, Pearson, Third Edition, 2010- Page no.-39)

presenting image processing 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 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. Figure 1.5, a plot of light intensity versus subjective brightness, illustrates this characteristic to the glare limit. The long solid curve represents the range of intensities to which the visual system can adapt. Inphotopic 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.

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

Image Sampling and Quantization:

To create a digital image, we need to convert the continuous sensed data into digital from. This involves two processes – sampling and quantization. An image may be continuous with respect to the x and y coordinates and also in amplitude. To convert it into digital form we have to sample the function in both coordinates and in amplitudes.

Digitalizing the coordinate values is called sampling. Digitalizing the amplitude values is called quantization. There is a continuous the image along the line segment AB. To simple this function, we take equally spaced samples along line AB. The location of each samples is given by a vertical tick back (mark) in the bottom part. The samples are shown as block squares superimposed on function the set of these discrete locations gives the sampled function.

In order to form a digital, the gray level values must also be converted (quantized) into discrete quantities. So we divide the gray level scale into eight discrete levels ranging from eight level values. The continuous gray levels are quantized simply by assigning one of the eight discrete gray levels to each sample. The assignment it made depending on the vertical proximity of a simple to a vertical tick mark.

Starting at the top of the image and covering out this procedure line by line produces a two dimensional digitalimage.

Digital Image definition:

A digital image f(m,n) described in a 2D discrete space is derived from an analog image f(x,y) in a 2D continuous space through a sampling process that is frequently referred to as digitization. The mathematics of that sampling process will be described in subsequent Chapters. For now we will look at some basic definitions associated with the digital image. The effect of digitization is shown in figure.

The 2D continuous image f(x,y) is divided into N rows and M columns. The intersection of a row and a column is termed a pixel. The value assigned to the integer coordinates (m,n) with m=0,1,2..N-1 and n=0,1,2…N-1 is f(m,n). In fact, in most cases, is actually a function of many variables including depth, color and time (t).

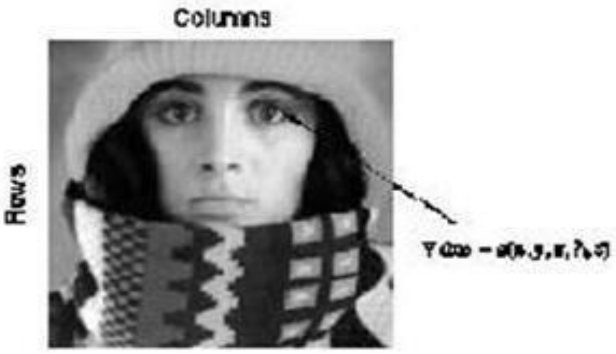


Fig2.1: Image in spatial Co- ordinates
Source: Anil K. Jain, Fundamentals of Digital Image Processing‘, Pearson, 2002.

There are three types of computerized processes in the processing of image

1) Low level process -these involve primitive operations such as image processing to reduce noise, contrast enhancement and image sharpening. These kind of processes are characterized by fact the both inputs and output are images.

2) Mid level image processing - it involves tasks like segmentation, description of those objects to reduce them to a form suitable for computer processing, and classification of individual objects. The inputs to the process are generally images but outputs are attributes extracted from images.

3) High level processing – It involves “making sense” of an ensemble of recognized objects, as in image analysis, and performing the cognitive functions normally associated with vision.

Representing Digital Images:

The result of sampling and quantization is matrix of real numbers. Assume that an image f(x,y) is sampled so that the resulting digital image has M rows and N Columns. The values of the coordinates (x,y) now become discrete quantities thus the value of the coordinates at orgin become 9X,y) =(o,o) The next

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

Coordinates value along the first signify the image along the first row. it does not mean that these are the actual values of physical coordinates when the image was sampled.

Thus the right side of the matrix represents a digital element, pixel or pel. The matrix can be represented in the following form as well. The sampling process may be viewed as partitioning the xy plane into a grid with the coordinates of the center of each grid being a pair of elements from the Cartesian products Z2 which is the set of all ordered pair of elements (Zi, Zj) with Zi and Zj being integers from Z. Hence f (x,y) is a digital image ifgray level (that is, a real number from the set of real number R) to each distinct pair of coordinates (x,y). This functional assignment is the quantization process. If the gray levels are also integers, Z replaces R, the and a digital image become a 2D function whose coordinates and she amplitude value are integers. Due to processing storage and hardware consideration, the number gray levels typically is an integer power of 2. L=2k

Then, the number, b, of bites required to store a digital image is B=M *N* k When M=N, the equation become b=N2*k. When an image can have 2k gray levels, it is referred to as “k- bit”. An image with 256 possible gray levels is called an “8- bit image” (256=28).

Spatial and Gray level resolution:

Spatial resolution is the smallest discernible details are an image. Suppose a chart can be constructed with vertical lines of width w with the space between the also having width W, so a line pair consists of one such line and its adjacent space thus. The width of the line pair is 2w and there is 1/2w line pair per unit distance resolution is simply the smallest number of discernible line pair unit distance.

Gray levels resolution refers to smallest discernible change in gray levels. Measuring discernible change in gray levels is a highly subjective process reducing the number of bits R while repairing the spatial resolution constant creates the problem of false contouring. It is caused by the use of an insufficient number of gray levels on the smooth areas of the digital image .

It is called so because the rides resemble top graphics contours in a map. It is generally quite visible in image displayed using 16 or less uniformly spaced gray levels

IMAGE SENSING AND ACQUISITION:

The types of images in which we are interested are generated by the combination of an “illumination” source and the reflection or absorption of energy from that source by the elements of the “scene” being imaged. We enclose *illumination* and *scene* in quotes to emphasize the fact that they are considerably more general than the familiar situation in which a visible light source illuminates a common everyday 3-D (three-dimensional) scene.

For example, the illumination may originate from a source of electromagnetic energy such as radar, infrared, or X-ray energy. But, as noted earlier, it could originate from less traditional sources, such as ultrasound or even a computer-generated illumination pattern. Similarly, the scene elements could be familiar objects, but they can just as easily be molecules, buried rockformations,orahumanbrain.Wecouldevenimageasource,suchasacquiringimagesof the sun. Depending on the nature of the source, illumination energy is reflected from, or transmitted through, objects. An example in the first category is light reflected from a planar surface. An example in the second category is when X-rays pass through a patient‟s body for the purpose of generating a diagnostic X-ray film. In some applications, the reflected or transmitted energy is focused onto a photo converter (e.g., a phosphor screen), which converts the energy into visible light. Electron microscopy and some applications of gamma imaging use this approach. The idea is simple: Incoming energy is transformed into a voltage by the combination of input electrical power and sensor material that is responsive to the particular type of energy being detected. The output voltage waveform is the response of the sensor(s), and a digital quantity is obtained from each sensor by digitizing its response. In this section, we look at the principal modalities for image sensing and generation.

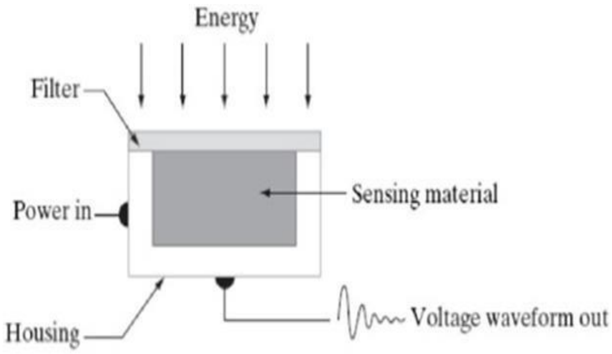


Fig2.2: Image Acquisition using a Single sensor

Source: Anil K. Jain, Fundamentals of Digital Image Processing‘, Pearson, 2002.

Image Acquisition using a Single sensor:

The components of a single sensor. Perhaps the most familiar sensor of this type is the photodiode, which is constructed of silicon materials and whose output voltagewaveformisproportionaltolight.Theuseofafilterinfrontofasensorimprovessele ctivity.example, a green (pass) filter in front of a light sensor favors light in the green band of the color spectrum. As a consequence, the sensor output will be stronger for green light than for other components in the visible spectrum.

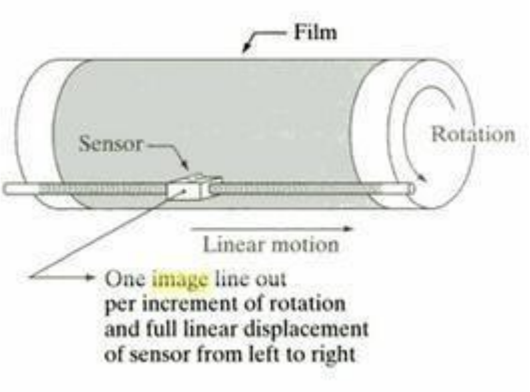


Fig2.3: Image Acquisition using a Single sensor

Source: Anil K. Jain, Fundamentals of Digital Image Processing‘, Pearson, 2002.

In order to generate a 2-D image using a single sensor, there has to be relative displacements in both the x- and y-directions between the sensor and the area to be imaged. Figure shows an arrangement used in high-precision scanning, where a film negative is mounted onto a drum whose mechanical rotation provides displacement in one dimension. The single sensor is mounted on a lead screw that provides motion in the perpendicular direction. Since mechanical motion can be controlled with high precision, this method is an inexpensive (but slow) way to obtain high-resolution images. Other similar mechanical arrangements use a flat bed, with the sensor moving in two linear directions. These types of mechanical digitizers sometimes are referred to as micro densitometers

Image Acquisition using a Sensor strips:

A geometry that is used much more frequently than single sensors consists of an in-line arrangement of sensors in the form of a sensor strip, shows. The strip provides imaging elements in one direction. Motion perpendicular to the strip provides imaging in the other direction. This is the type of arrangement used in most flat bed scanners. Sensing devices with 4000 or more in-line sensors are possible. In-line sensors are used routinely in airborne imaging applications, in which the imaging system is mounted on an aircraft that flies at a constant altitude and speed over the geographical area to be imaged. One dimensional imaging sensor strips that respond to various bands of the electromagnetic spectrum are mounted perpendicular to the direction of flight. The imaging strip gives one line of an image at a time, and the motion of the strip completes the other dimension of a two-dimensional image. Lenses or other focusing schemes are used to project area to be scanned onto the sensors. Sensor strips mounted in a ring configuration are used in medical and industrial imaging to obtain cross-sectional (“slice”) images of 3-Dobjects.

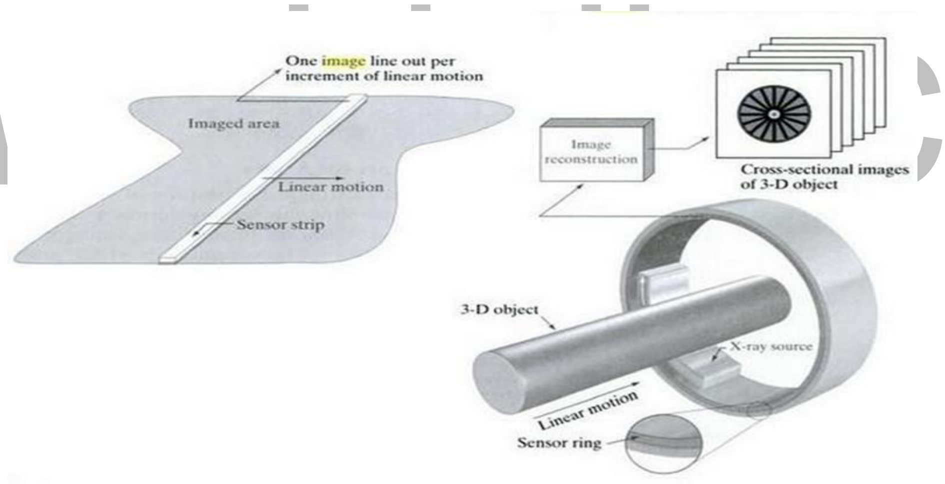


Fig 2.4 : Image Acquisition using linear strip and circular strips.

Source: Anil K. Jain, undamentals of Digital Image Processing‘, Pearson, 2002.

Image Acquisition using a Sensor Arrays:

The individual sensors arranged in the form of a 2-D array. Numerous electromagnetic and some ultrasonic sensing devices frequently are arranged in an array format. This is also the predominant arrangement found in digital cameras. A typical sensor for these cameras is a CCD array, which can be manufactured with a broad range of sensing properties and can be packaged in rugged arrays of elements or more. CCD sensors are used widely in digital cameras and other light sensing instruments. The response of each sensor is proportional to the integral of the light energy projected onto the surface of the sensor, a property that is used in astronomical and other applications requiring low noise images. Noise reduction is achieved by letting the sensor integrate the input light signal over minutes or even hours. The two dimensional, its key advantage is that a complete

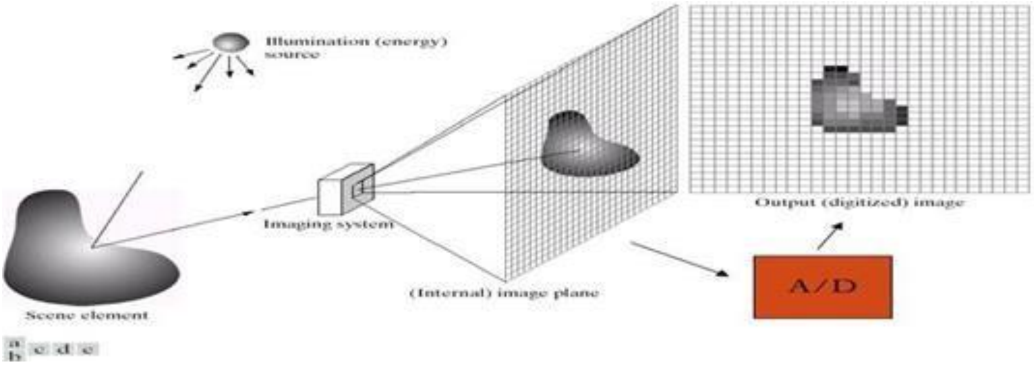


Fig: An example of the digital image acquisition process. (a) Energy ("illumination") source. (b) An element of a scene. (c) Imaging system. (d) Projection of the scene onto the image plane. (e) Digitized image.

Source 2.5: . Kenneth R. Castleman, igital Image Processing‘, Pearson, 2006.

Image can be obtained by focusing the energy pattern onto the surface of the array. Motion obviously is not necessary, as is the case with the sensor arrangements This figure shows the energy from an illumination source being reflected from a scene element, but, as mentioned at the beginning of this section, the energy also could be transmitted through the scene elements. The first function performed by the imaging system is to collect the incoming energy and focus it onto an image plane. If the illumination is light, the front end of the imaging system is a lens, which projects the viewed scene onto the lens focal plane. The sensor array, which is coincident with the focal plane, produces outputs proportional to the integral of the light received at each sensor. Digital and analog circuitry sweep these outputs and convert them to a video signal, which is then digitized by another section of the imaging system.

Relationship between Pixels:

We consider several important relationships between pixels in a digital image.

NEIGHBORS OF A PIXEL

A pixel *p* at coordinates (*x*,*y*) has four *horizontal* and *vertical* neighbors whose coordinates are givenby:

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

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

This set of pixels, called the 4-*neighbors* or *p*, is denoted by $N_4(p)$. Each pixel is one unit distance from (x,y) and some of the neighbors of p lie outside the digital image if (x,y) is on the border of the image. The four *diagonal* neighbors of *p* have coordinates and are denoted by $N_D(p)$.

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

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

These points, together with the 4-neighbors, are called the 8-neighbors of p, denoted by $N_8(p)$

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

As before, some of the points in $N_D(p)$ and $N_8(p)$ fall outside the image if (x,y) is on the border of the image.

ADJACENCY AND CONNECTIVITY

Let v be the set of gray –level values used to define adjacency, in a binary image, v={1}.

In a gray-scale image, the idea is the same, but *V* typically contains more elements, for example, $V = \{180, 181, 182, \ldots, 200\}$.

If the possible intensity values 0 – 255, *V* set can be any subset of these 256 values. 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$N_4(P)$
- 8- Adjacency – two pixel P and Q with value from V are 8 –adjacency if A is in the set$N_8(P)$
- M-adjacency –two pixel P and Q with value from V are m – adjacency if (i) Q is in $N_4(p)$ or (ii) Q is in $N_D(q)$ and the set $N_4(p) \cap N_4(q)$ has no pixel whose values are from V.
- Mixed adjacency is a modification of 8-adjacency. It is introduced to eliminate the ambiguities that often arise when 8-adjacency is used.
- Forexample:

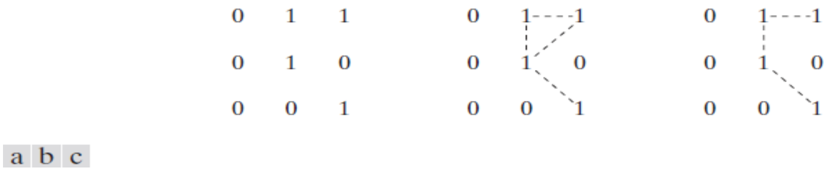


Fig:1.8(a) Arrangement of pixels; (b) pixels that are 8-adjacent (shown dashed) to the center pixel; (c) ***m***-adjacency.

Types of Adjacency:

- In this example, we can note that to connect between two pixels (finding a path between two pixels):

  -In 8-adjacency way, you can find multiple paths between twopixels

- While, in m-adjacency, you can find only one path between twopixels

- So, m-adjacency has eliminated the multiple path connection that has been generated by the8-adjacency.
- Two subsets $S_1$and $S_2$are adjacent, if some pixel in $S_1$is adjacent to some pixel in S_2.

A Digital Path:

- A digital path (or curve) from pixel *p* with coordinate (x,y) to pixel *q* with coordinate (s,t) is a sequence of distinct pixels with coordinates (x_0,y_0), (x_1,y_1), …, (x_n, y_n) where $(x_0,y_0) = (x,y)$ and $(x_n, y_n) = (s,t)$ and pixels (x_i, y_i) and (x_{i-1}, y_{i-1}) are adjacent for $1 \leq i \leq n$
- n is the length of the path
- If $(x_0,y_0) = (x_n, y_n)$, the path is closed.

  We can specify 4-, 8- or m-paths depending on the type of adjacency specified.

  - Return to the previous example:

Fig:1.8 (a) Arrangement of pixels; (b) pixels that are 8-adjacent(shown dashed) to the center pixel; (c) m-adjacency.

In figure (b) the paths between the top right and bottom right pixels are 8-paths. And the path between the same 2 pixels in figure (c) is m-path

Connectivity:

- Let *S* represent a subset of pixels in an image, two pixels*p* and *q* are said to be connected in *S* if there exists a path between them consisting entirely of pixels in *S*

- For any pixel *p* in *S*, the set of pixels that are connected to it in *S* is called a *connected component* of *S*. If it only has one connected component, then set *S* is called a *connected set*.

Region and Boundary:

- REGION: Let *R* be a subset of pixels in an image, we call *R* a region of the image if *R* is a connected set.
- BOUNDARY: The *boundary* (also called *border* or *contour*) of a 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 in the image. This extra definition is required because an image has no neighbors beyond its borders. Normally, when we refer to a region, we are referring to subset of an image, and any pixels in the boundary of the region that happen to coincide with the border of the image are included implicitly as part of the region boundary.

COLOR IMAGE FUNDAMENTALS

According to the theory of the human eye, all colors are seen as variable combinations of the three so-called primary colors red (*R*), green (*G*), and blue (*B*). The following specific wavelength values to the primary colors:

Blue(B) = 435.8 nm

Green (*G*)= 546.1 nm

Red (*R*) = 700.0 nm

The primary colors can be added to produce the secondary colors

Magenta (red + blue), **cyan** (green + blue), and **yellow** (red + green), see Figure

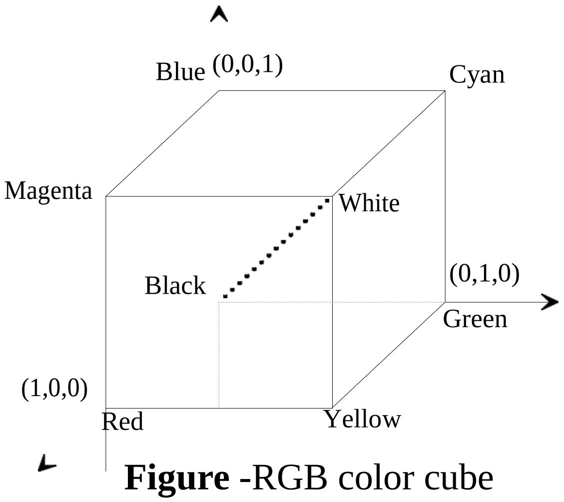


Figure -RGB color cube

(Source: Rafael C. Gonzalez, Richard E. Woods, Digital Image Processing‘, Pearson, Third Edition, 2010- Page no-402)

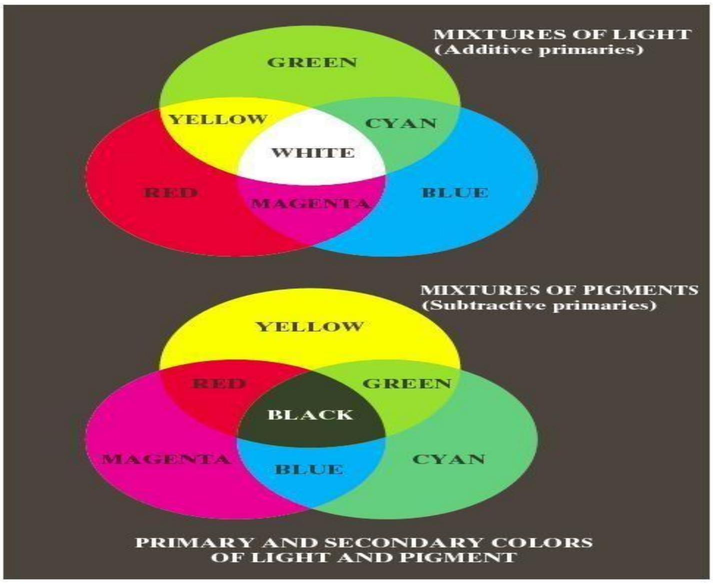


Fig: RGB image and the components of its corresponding HSI image: (b) hue, (c) saturation, and (d) intensity

(Source: Rafael C. Gonzalez, Richard E. Woods, Digital Image Processing‘, Pearson, Third Edition, 2010. – Page no. -403)

Mixing all three primary colors results in white. Color television reception is based on this three color system with the additive nature of light.

There are several useful color models: RGB, CMY, YUV, YIQ, and HSI.

1. RGB color model

The colors of the RGB model can be described as a triple (R, G, B), so that R, G, B .The RGB color space can be considered as a three-dimensional unit cube, in which each axis represents one of the primary colors, see Figure. Colors are points inside the cube defined by its coordinates. The primary colors thus are red=(1,0,0), green=(0,1,0), and blue=(0,0,1). The secondary colors of RGB are cyan=(0,1,1), magenta=(1,0,1) and yellow=(1,1,0).

The nature of the RGB color system is additive in the sense how adding colors makes the image brighter. **Black is at the origin, and white is at the corner** where R=G=B=1. The gray scale extends from black to white along the line joining these two points. Thus a shade of gray can be described by (x,x,x) starting from black=(0,0,0) to white=(1,1,1).

The colors are often normalized as given in equation .

This normalization guarantees that r+g+b=1.

ADDITIVE COLOUR MIXING

In order to generate suitable colour signals it is necessary to know definite ratios in which red, green and blue combine form new colours. Since R,G and B can be mixed to create any colour including white, these are called primary colours.

R + G + B = White colour
R – G – B = Black colour
The primary colours R, G and B combine to form new colours
R + G = Y
R + B = Magenta (purplish)
G + B = Cyan (greenish blue)

HSI COLOR MODEL

The HSI model consists of *hue* (*H*), *saturation* (*S*), and *intensity* (*I*).

Intensity corresponds to the luminance component (*Y*) of the YUV and YIQ models.

Hue is an attribute associated with the dominant wavelength in a mixture of light waves, i.e. the dominant color as perceived by an observer.

Saturation refers to relative purity of the amount of white light mixed with hue. The advantages of HSI are:

The intensity is separated from the color information (the same holds for the YUV and YIQ models though).

The hue and saturation components are intimately related to the way in which human beings perceive color.

Digital Image Fundamentals:

The field of digital image processing refers to processing digital images by means of digital computer. 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 digitalimage.

An image is a two-dimensional function that represents a measure of some characteristic such as brightness or color of a viewed scene. An image is a projection of a 3-D scene into a 2D projection plane.

An image may be defined as a two-dimensional function ***f(x,y),*** where x and y are spatial (plane) coordinates, and the amplitude of ***f***at any pair of coordinates **(x,y)** is called the intensity of the image at thatpoint.

Fig 1.1: Fundamentals of Digital Image Processing System

Source: Tutorials point

The term ***gray level*** is used often to refer to the intensity of monochrome images. Color images are formed by a combination of individual 2-D images.

For example: The RGB color system, a color image consists of three (red, green and blue) individual component images. For this reason many of the techniques developed for monochrome images can be extended to color images by processing the three component images individually.

An image may be continuous with respect to the x- and y- coordinates and also in amplitude. Converting such an image to digital form requires that the coordinates, as well as the amplitude, be digitized.

APPLICATIONS OF DIGITAL IMAGE PROCESSING

Since digital image processing has very wide applications and almost all of the technical fields are impacted by DIP, we will just discuss some of the major applications of DIP.

Digital image processing has a broad spectrum of applications, such as

- Remote sensing via satellites and otherspacecrafts
- Image transmission and storage for businessapplications
- Medicalprocessing,
- RADAR (Radio Detection andRanging)
- SONAR(Sound Navigation and Ranging)and
- Acoustic image processing (The study of underwater sound is known as underwater acousticsor hydroacoustics.)
- Robotics and automated inspection of industrial

parts. Images acquired by satellites are useful in trackingof

- Earthresources;
- Geographicalmapping;
- Prediction of agriculturalcrops,
- Urban growth and weathermonitoring
- Flood and fire control and many other

environmentalapplications. Space image applicationsinclude:

- Recognition and analysis of objects contained in images obtained from deep space-probemissions.
- Image transmission and storage applications occur in broadcasttelevision
- Teleconferencing
- Transmission of facsimile images(Printed documents and graphics) for office automation

Communication over computer networks

- Closed-circuit television based security monitoring systems and
- In military

Medicalapplications:

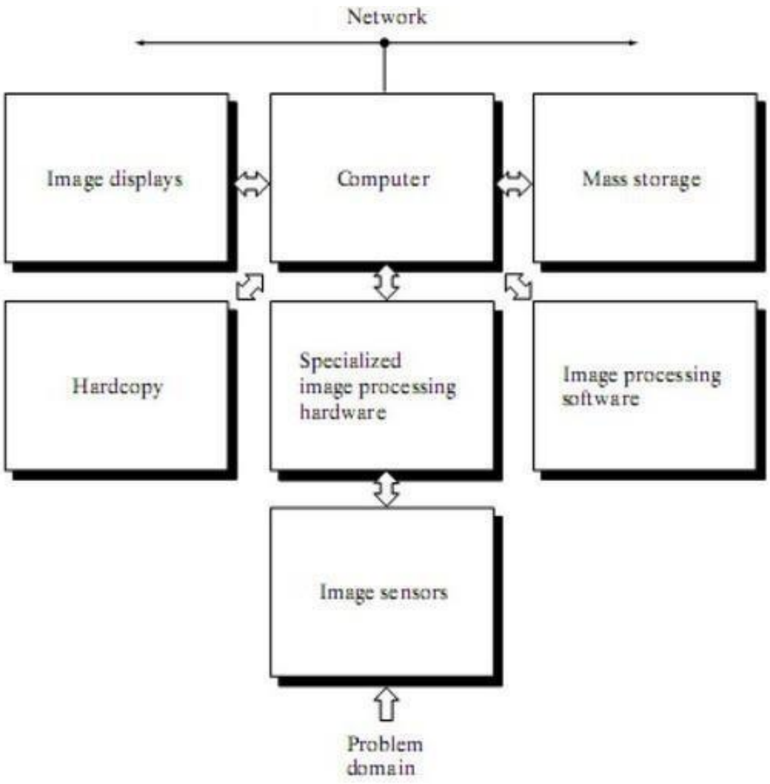
- Processing of chest X-rays
- Cineangiograms
- Projection images of transaxial tomography and
- Medical images that occur in radiology nuclear magnetic resonance(NMR)
- Ultrasonic scanning

IMAGE PROCESSING TOOLBOX (IPT) is a collection of functions that extend the capability of the MATLAB numeric computing environment. These functions, and the expressiveness of the MATLAB language, make many image-processing operations easy to write in a compact, clear manner, thus providing a ideal software prototyping environment for the solution of image processing problem.

COMPONENTS OF IMAGE PROCESSING SYSTEM:

Computer: It is a general purpose computer and can range from a PC to a supercomputer depending on the application. In dedicated applications, sometimes specially designed computer are used to achieve a required level of performance

Software: It consists of specialized modules that perform specific tasks a well designed package also includes capability for the user to write code, as a minimum,



utilizes the specialized module. More sophisticated software packages allow the integration of these modules.

Mass storage: This capability is a must in image processing applications. An image of size 1024 x1024 pixels, in which the intensity of each pixel is an 8- bit quantity requires one Megabytes of storage space if the image is not compressed .Image processing applications falls into three principal categories of storage

i) Short term storage for use during processing

ii) On line storage for relatively fast retrieval

iii) Archival storage such as magnetic tapes and disks

Image display: Image displays in use today are mainly color TV monitors. These monitors are driven by the outputs of image and graphics displays cards that are an integral part of computer system.

Hardcopy devices: The devices for recording image includes laser printers, film cameras, heat sensitive devices inkjet units and digital units such as optical and CD ROM disk. Films provide the highest possible resolution, but paper is the obvious medium of choice for written applications.

Networking: It 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 bandwidth.

FUNDAMENTAL STEPS IN DIGITAL IMAGE PROCESSING:

There are two categories of the steps involved in the image processing –

1. Methods whose outputs are input are images.
2. Methods whose outputs are attributes extracted from those images.

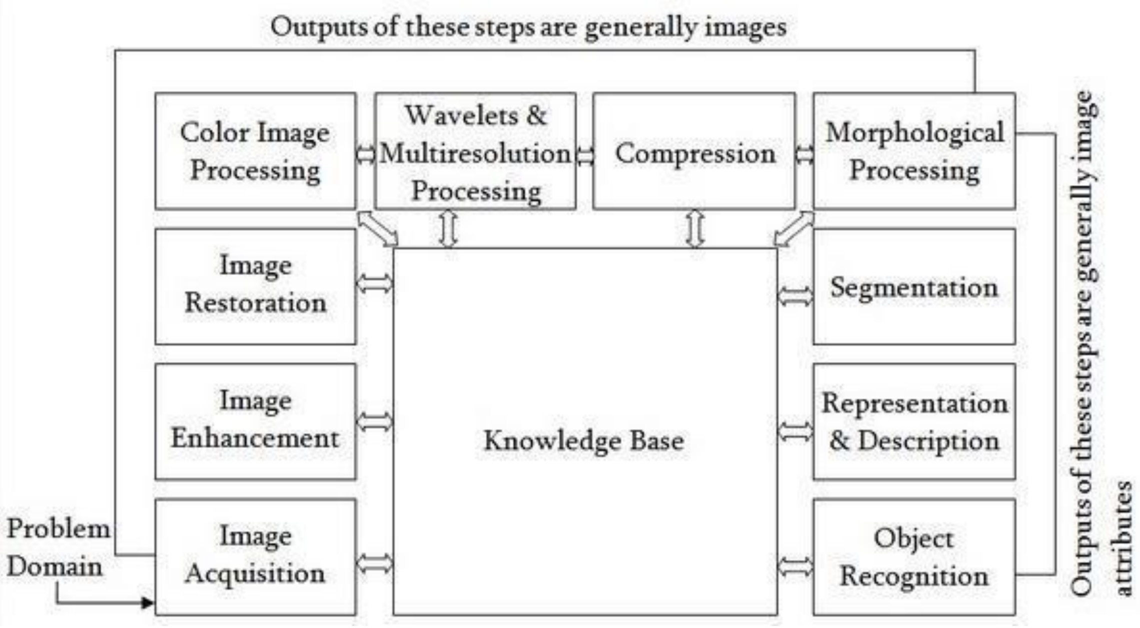


Fig 1.3 : Fundamental Steps in Digital Image Processing
Source: Rafael C. Gonzalez, Richard E. Woods, igital Image Processing‘, Pearson, Third Edition, 2010.

Image acquisition: It could be as simple as being given an image that is already in digital form. Generally the image acquisition stage involves processing such scaling.

Image Enhancement: It is among the simplest and most appealing areas of digital image processing. The idea behind this is to bring out details that are obscured or simply to highlight certain features of interest in image. Image enhancement is a very subjective area of image processing.

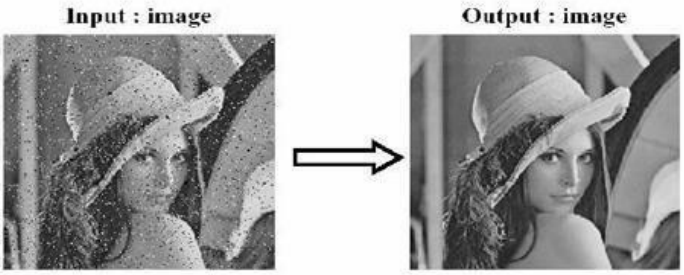


Figure: 1.4 Image Enhancement using Homomorphic Filtering
Source: Rafael C. Gonzalez, Richard E. Woods, igital Image Processing‘, Pearson, Third Edition, 2010.

Image Restoration: It deals with improving the appearance of an image. It is an objective approach, in the sense that restoration techniques tend to be based on mathematical or probabilistic models of image processing. Enhancement, on the other hand is based on human subjective preferences regarding what constitutes a “good” enhancement result.

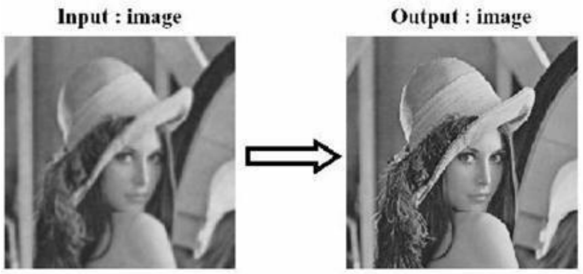


Figure: 1.5 Image Enhancement using Homomorphic Filtering
Source: Rafael C. Gonzalez, Richard E. Woods, igital Image Processing', Pearson, Third Edition, 2010

Color image processing: It is an area that is been gaining importance because of the use of digital images over the internet. Color image processing deals with basically color models and their implementation in image processing applications.

Wavelets and Multiresolution Processing: These are the foundation for representing image in various degrees of resolution.

Compression: It deals with techniques reducing the storage required to save an image, or the bandwidth required to transmit it over the network. It has to major approaches a) Lossless Compression b) Lossy Compression

Morphological processing: It deals with tools for extracting image components that are useful in the representation and description of shape and boundary of objects. It is majorly used in automated inspection applications.

Representation and Description: It always follows the output of segmentation step that is, raw pixel data, constituting either the boundary of an image or points in the region itself. In either case converting the data to a form suitable for computer processing is necessary.

Recognition: It is the process that assigns label to an object based on its descriptors. It is the last step of image processing which use artificial intelligence of software.

Knowledge base:

Knowledge about a problem domain is coded into an image processing system in the form of a knowledge base. This knowledge may be as simple as detailing regions of an image where the information of the interest in known to be located. Thus limiting search that has to be conducted in seeking the information. The knowledge base also can be quite complex such interrelated list of all major possible defects in a materials inspection problems or an image database containing high resolution satellite images of a region in connection with change detection application.

A Simple Image Model:

An image is denoted by a two dimensional function of the form f{x, y}. The value or amplitude of f at spatial coordinates {x,y} is a positive scalar quantity whose physical meaning is determined by the source of the image. When an image is generated by a physical process, its values are proportional to energy radiated by a physical source. As a consequence, f(x,y) must be nonzero and finite; that is o<f(x,y) <co The function f(x,y) may be characterized by two components- The amount of the source illumination incident on the scene being viewed.

(a) The amount of the source illumination reflected back by the objects in the scene These are called illumination and reflectance components and are denoted by i(x,y) an r (x,y) respectively.

The functions combine as a product to form f(x,y). We call the intensity of a monochrome image at any coordinates (x,y) the gray level (l) of the image at that point l= f (x, y.)

L min ≤ l ≤ Lmax Lmin is to be positive and

Lmax must be finite

Lmin=iminrmin

Lmax=imaxrmax

The interval [Lmin, Lmax] is called gray scale. Common practice is to shift this interval numerically to the interval [0, L-l] where l=0 is considered black and l= L-1 is considered white on the gray scale. All intermediate values are shades of gray of gray varying from black to white.

Discrete Cosine Transform (Dct) :

The discrete cosine transform (DCT) helps separate the image into parts (or spectral sub- bands) of differing importance (with respect to the image's visual quality). The DCT is similar to the discrete Fourier transform: it transforms a signal or image from the spatial domain to the frequency domain.

The general equation for a 1D (*N* data items) DCT is defined by the following equation:

$$F(u) = \left(\frac{2}{N}\right)^{\frac{1}{2}} \sum_{i=0}^{N-1} \Lambda(i).cos\left[\frac{\pi.u}{2.N}(2i+1)\right] f(i)$$

and the corresponding *inverse* 1D DCT transform is simple $F^{-1}(u)$, i.e.: where

$$\Lambda(i) = \begin{cases} \frac{1}{\sqrt{2}} & \text{for} \xi = 0 \\ 1 & \text{otherwise} \end{cases}$$

The general equation for a 2D (*N* by *M* image) DCT is defined by the following equation:

$$F(u,v) = \left(\frac{2}{N}\right)^{\frac{1}{2}} \left(\frac{2}{M}\right)^{\frac{1}{2}} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} \Lambda(i).\Lambda(j).cos\left[\frac{\pi.u}{2.N}(2i+1)\right] cos\left[\frac{\pi.v}{2.M}(2j+1)\right].f(i,j)$$

and the corresponding *inverse* 2D DCT transform is simple F^{-}

$$\Lambda(\xi) = \begin{cases} \frac{1}{\sqrt{2}} & \text{for} \xi = 0 \\ 1 & \text{otherwise} \end{cases}$$

$^{1}(u,v)$, i.e.: where

The basic operation of the DCT is as follows:

- The input image is N by M;
- f(i,j) is the intensity of the pixel in row i and column j;
- F(u,v) is the DCT coefficient in row k1 and column k2 of the DCT matrix.

- For most images, much of the signal energy lies at low frequencies; these appear in the upper left corner of the DCT.
- Compression is achieved since the lower right values represent higher frequencies, and are often small - small enough to be neglected with little visible distortion.
- The DCT input is an 8 by 8 array of integers. This array contains each pixel's gray scale level;
- 8 bit pixels have levels from 0 to 255.

Properties of the Cosine Transform

1. The cosine transform is real and orthogonal, that is, $C = C^*$ $C^{-1} = C^T$

2. The cosine transform is not the real part of the unitary DFT. This can be seen by inspection of C and the DFT matrix F. (Also see Problem) However, the cosine transform of a sequence is related to the DFT of its symmetric extension (see Problem)

3. The cosine transform is a fast transform. The cosine transform of a vector of N elements can be calculated in O (N $\log_2$N) operations via an N-point FFT [19]. To show this we define a new sequence u (n) by reordering the even and odd elements of u (n) as

4. The basis vectors of the cosine transform (that is, rows of C) are the eigenvectors of the symmetric tridiagonal matrix Q_c, defined as

$$Q_c = \begin{pmatrix} 1-\alpha & -\alpha & 0 \\ -\alpha & 1 & 1-\alpha \\ 0 & \alpha & 1-\alpha \end{pmatrix}$$

5. The N X N cosine transform is very close to the KL transform of a first-order stationary Markov sequence of length N whose covariance matrix is given by when the correlation parameter ρ is close to 1. The reason is that R^{-1} is a symmetric tridiagonal matrix, which for a scalar

$$\begin{pmatrix} 1-\rho\alpha & -\alpha & 0 \end{pmatrix}$$

$$\beta^2 R^{-1} = \begin{vmatrix} -\alpha & 1 & 1-\alpha \\ 0 & \alpha & 1-\rho\alpha \end{vmatrix}$$

This gives the approximation $\beta^2 R^{-1} \cong Q_C$ for $\rho \cong 1$

Therefore, if we calculate the N-point inverse FFT of the sequence (k) □ (k) exp (j k/2N), we can also obtain the inverse DCT in O (N log N) operations. Direct algorithms that do not require FFT as an intermediate step, so that complex arithmetic is avoided, are also possible [18]. The computational complexity of the direct as well as the FFT based methods is about the same.

6. The cosine transform has excellent energy compaction for highly correlated data..

The basis vectors of the cosine transform (that is, rows of C) are the eigenvectors of the sy Therefore, if we calculate the N-point inverse FFT of the sequence (k) (k) exp (j k/2N), we can also obtain the inverse DCT in O (N log N) operations. Direct algorithms that do not require FFT as an intermediate step, so that complex arithmetic is avoided, are also possible [18]. The computational complexity of the direct as well as the FFT based methods is about the same.

Hence the eigenvectors of R and the eigenvectors of Q_c, that is, the cosine transform, will be quite close. These aspects are considered in greater depth in Section on sinusoidal transforms.

This property of the cosine transform together with the fact that it is a fast transform has made it a useful substitute for the KL transform of highly correlated first order Markov sequences.

The two dimensional Discrete Fourier Transform:

Discrete Fourier transform (DFT) is the array given by

$$F(u,v) = \frac{1}{MN} \sum^{M-1} \sum^{N-1} f(x,y) e^{-j2\pi(ux/M+vy/N)}$$

$$x=0 \quad y=0$$

$$u = 0,\square, M-1,\ v = 0\square, N-1$$

A function represented in either of these forms and can be completely reconstructed via an inverse process with no loss of information.

$$\mathcal{F}\{f(x)\} = F(u) = \int_{-\infty}^{\infty} f(x)\exp(-j2\pi ux)dx \qquad j=\sqrt{-1}$$

1-D Fourier Transformation and its Inverse

If there is a single variable, continuous function f(x) , then Fourier transformation F (u) may be given as and the reverse process to recover f(x) from F(u) is

$$F(u) = R(u) + jI(u) \qquad F(u) = |F(U)|e^{j\phi(u)}$$

$$|F(u)| = [R^2(u)+I^2(u)]^{\frac{1}{2}} \quad \text{or} \quad \phi(u) = \tan^{-1}\left[\frac{I(u)}{R(u)}\right]$$

F(u) in polar coordinates

Fourier Transformation and its Inverse . The Fourier Transform of a two dimensional continuous function f(x,y) (an image) of size M * N is given by

$$\mathcal{F}\{f(x,y)\} = F(u,v) = \int\int_{-\infty} f(x,y)\exp[-j2\pi(ux+vy)]dxdy$$

Inverse Fourier transformation is given by equation

$$\mathcal{F}^{-1}\{F(u,v)\} = f(x,y) = \int\int_{-\infty}^{\infty} F(u,v)\exp[\,j2\pi(ux+vy)]dudv$$

Where (u,v) are frequency variables.
Preprocessing is done to shift the origin of F(u,v) to frequency coordinate (m/2,n/2)

which is the center of the M*N area occupied by the 2D-FT. It is known as frequency rectangle.

$$\mathcal{F}^{-1}\{F(u)\} = f(x) = \int_{-\infty}^{\infty} F(u)\exp[\,j2\pi ux\,]\,du$$

Fourier transformation of a discrete function of one variable f(x), x=0, 1, 2, m-1 is given by

$$F(u)=\frac{1}{N}\sum_{x=0}^{N-1} f(x)\exp\left[-j2\pi ux/N\right] \quad \text{for u=0,1,2,…..,N-1}$$

to obtain f(x) from F(u)

$$f(x)=\sum_{u=0}^{N-1} F(u)\exp\left[j2\pi ux/N\right] \qquad \text{for x=0,1,2,…….,N-1}$$

The above two equation (e) and (f) comprise of a discrete Fourier transformation pair. According to Euler's formula

. $e^{jx} = \cos x + j \sin x$

Substituting these value to equation (e)

F(u)=Σf(x)[cos 2πux/N+jSin 2πux/N] for u=0,1,2,……,N- 1

Now each of the m terms of F(u) is called a frequency component of transformation "The Fourier transformations eparates a function into various components, based on frequency components. These components are complex quantities.

$$F(u) = R(u) + jI(u) \qquad F(u) = |F(U)|e^{j\phi(u)}$$

or

$$|F(u)| = [R^2(u) + I^2(u)]^{1/2} \qquad \phi(u) = \tan^{-1}\left[\frac{I(u)}{R(u)}\right]$$

F(u) in polar coordinates

The Fourier Transform of a two dimensional continuous function f(x,y) (an image) of size M * N is given by

$$\mathcal{F}\{f(x,y)\} = F(u,v) = \int\!\!\int_{-\infty} f(x,y)\exp[-j2\pi(ux+vy)]dxdy$$

Inverse Fourier transformation is given by equation

$$\mathcal{F}^{-1}\{F(u,v)\} = f(x,y) = \int\!\!\int_{-\infty}^{\infty} F(u,v)\exp[j2\pi(ux+vy)]dudv$$

Where (u,v) are frequency variables.

Preprocessing is done to shift the origin of F(u,v) to frequency coordinate (m/2,n/2) which is the center of the M*N area occupied by the 2D-FT. It is known as frequency rectangle.