Ummie Habiba

Module-1

Digital Image Fundamentals

1.1 What is Digital Image Processing?

* An image may be defined as a two-dimensional functio $f(x,y)$, where x & y are spatial (plane) coordinates, and th amplitude of f at any pair of coordinates (x,y) is called the intensity or gray level of the image at that point.

* When x, y, & the amplitude values of f are all finite, discre quantities, then the image is called a digital image.

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

* A digital image is composed of a finite number of elemen each of which has a particular location & value. These elements are referred to as picture elements, image elements pels and pixels.

* Pixel is the term most widely used to denote the elements of a digital image.

* Vision is most advanced of our senses & images play th most important role in human perception. However, who c limited to visual band of EM spectrum, imaging machin cover almost the entire EM spectrum, ranging from gamm to radio waves. They can operate on images generate by sources that humans are not accustomed to associat with images. These include ultrasound, electron microscop & computer-generated images. Thus, digital image processing

Subscribe EC Academy on YouTube

* Image processing is a discipline in which both the input & o/p of a process are images.

* The area of image analysis lies in b/w image processing & computer vision.

* Image processing involves 3 types of computerized processes low-, mid- & high-level processes.

→ Low-level processes involve image processing to reduce noise, contrast enhancement & image sharpening. Both the i/p's & o/p's are images in low-level processing.

→ Mid-level processing involves tasks such as segmentation (partitioning an image into regions or objects), description of those objects to reduce them to a form suitable for computer processing & classification (recognition) of individual objects. In this, i/p's are images, but o/p's are attributes extracted from those images (e.g. edges, contours, & the identity of individual objects).

→ Higher-level processing involves "making sense" of an ensemble of recognized objects and performing the cognitive functions normally associated with vision.

* A simple example for digital image processing is that consider an area of automated analysis of text. The processes of acquiring an image of the area containing the text, preprocessing that image, extracting (segmenting) the individual characters, describing the characters in form suitable for computer processing & recognizing the individual characters. Making sense of content of page

Subscribe EC Academy on YouTube

may be viewed as being in the domain of image analysis & even computer vision depending on the level of complexity.

1.2 The Origins of Digital Image Processing:-

* The first application of digital images was in newspaper industry, when pictures were first sent by submarine cable b/w London & New York.

Introduction of Bartlane cable picture tx^n system in early 1920s reduced the time required to transport a picture o/c the Atlantic from more than a week to less than 3 hrs. Specialized printing equipment coded pictures for cable tx^n & then reconstructed at the receiving end.

The initial problems in improving visual quality of the early pictures were related to selection of printing procedure & distribution of intensity levels.

Later in 1921, the printing method based on photographic reproduction made from tapes perforated at the telegraph receiving terminal. The improved image was evident both in tonal quality & in resolution.

The early Bartlane systems were capable of coding images in 5 distinct levels of gray. This capability was increased to 15 levels in 1929.

Digital images require so much storage & computational power that progress in field of digital image processing has been dependent on development of digital computers & of supporting technologies that include data storage, display & transmission.

The invention of VLSI & ULSI in 1980s has added

Subscribe EC Academy on YouTube

advantages for ...

The invention of CAT (computerized Axial Tomography) also called computerized Tomography (CT) is most important event in application of image processing in medical diagnosis.

CAT is a process in which a ring of detectors encircles an object (or patient) & an X-ray source, concentric with detector ring, rotates about the object. Thus from 1960s until the present, the field of DIP has grown vigorously.

The continuing decline in the ratio of computer price to performance & the expansion of networking & comm'n BW via the WWW & the Internet have created lot of opportunities for continued growth of Digital image processing.

.3 Examples of Fields that use Digital Image Processing:-

Almost all areas are impacted by DIP, now-a-days. Here only few of these applications are explained.

One of the simplest ways to develop a basic understanding of the extent of image processing applications is to categorize images according to their source (e.g., visual, X-ray & so on).

The principal energy source for images in use today is the electromagnetic energy spectrum. Other sources of energy include acoustic, ultrasonic & electronic. Synthetic images, used for modeling & visualization are generated by computer.

Subscribe EC Academy on YouTube

EM waves can be conceptualized as propagating sinusoidal waves of varying wavelengths or they can be thought of as a stream of massless particles each traveling in a wavelike pattern & moving at the speed of light. Each massless particle contains a certain amount (or bundle) of energy. Each bundle of energy is called a photon.

If spectral bands are grouped according to energy per photon, we obtain the spectrum shown in the fig. ranging from gamma rays to radio waves.

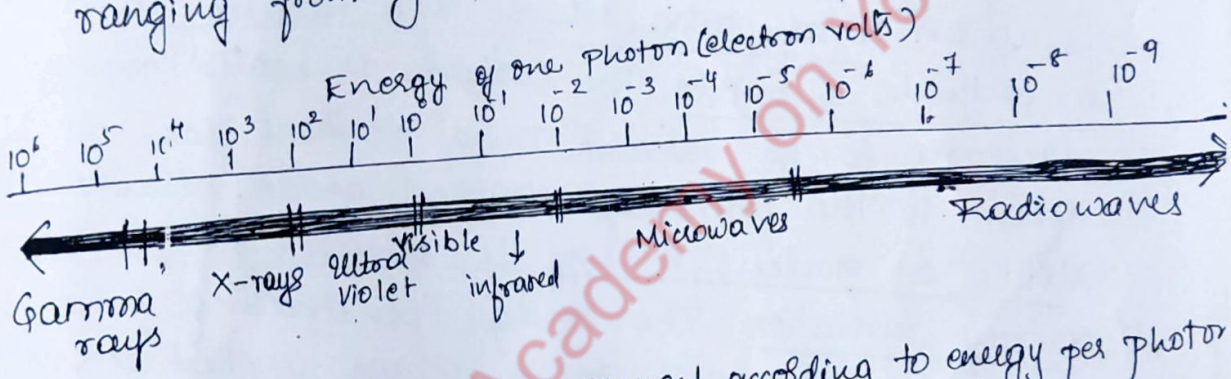


Fig:- The EM spectrum arranged according to energy per photon

1.3.1 Gamma-Ray Imaging:-

* Major uses of imaging based on gamma rays include nuclear medicine and astronomical observations.
* In nuclear medicine, the approach is to inject a patient with a radioactive isotope that emits gamma rays as it decays. Images are produced from the emissions collected by gamma ray detectors.
* Another major modality of nuclear imaging called positron emission tomography (PET). In this, the patient is given a radioactive isotope that emits positrons as it decays. When a positron meets an electron, both

Subscribe EC Academy on YouTube

are detected & a tomographic image is created using the basic principles of tomography.

3.2 X-ray Imaging:—

* X-rays are the oldest sources of EM radiation used for imaging. The best known use of X-rays is medical diagnostics & also used in industry & areas like astronomy.

* X-rays for medical & industrial imaging are generated using an X-ray tube, which is a vacuum tube with a cathode & anode. The cathode is heated causing free electrons to be released. These electrons flow at high speed to the positively charged anode. When electrons strike a nucleus, energy is released in the form of X-ray radiation. The energy (penetrating power) of X-rays is controlled by a voltage applied across the anode & the no. of X-rays is controlled by a current applied to the filament in the cathode.

* In digital radiography, digital images are obtained by one of two methods: (1) by digitizing X-ray films or (2) by having the X-rays that pass thr the patient fall directly onto devices (such as phosphor screen) that convert X-rays to light. The light signal intern is captured by a light-sensitive digitizing system.

* Angiography is another major application in an area called contrast-enhancement radiography. This procedure is used to obtain images (called angiograms) of blood vessels. A catheter (a small, flexible, hollow tube) is inserted, for ex, into an artery or vein in the groin.

Subscribe EC Academy on YouTube

The catheter is threaded into the blood vessels & guided to area to be studied. When the catheter reaches the site under investigation, an x-ray contrast medium is injected thr' the catheter. This enhances contrast of the blo vessels & enables the radiologist to see any irregularities or blockages.

* The best known of all uses of x-rays in medical imaging computerized axial tomography (CAT). CAT image is a "slice" taken perpendicularly through the patient. Numerous slice are generated as the patient is moved in a longitudinal direction. The ensemble of such images constitutes a 3-D rendition of inside of patient with longitudin resolution being proportional to the no. of slice imag taken..

1.3.3 Imaging in the Ultraviolet Band:-

* Applications of ultraviolet 'light' are varied. They inclu lithography, industrial inspection, microscopy, lasers, biological imaging & astronomical observations.

* UV light is used in fluorescence microscopy. When a photon of ultraviolet radiation collides with an electro in an atom of a fluorescent material, it elevates the electron to a higher energy level. Subsequently, the exci electron relaxes to a lower level & emits light in the form of a lower-energy photon in the visible (red) ligh region.

* The basic task of fluorescence microscope is to use a excitation light to irradiate a prepared specimen & t to separate the much weaker radiating fluorescen light from the brighter excitation light. Thus, only visible light reaches the eye or other detector

Subscribe ECAcademy on YouTube

The [illegible] fluorescing [illegible] against a dark background with sufficient contrast to permit detection. The darker the background of the nonfluorescing material, the more efficient the instrument.

1.3.4 Imaging in the visible & Infrared Bands:-

* The infrared band is often used in conjunction with visual imaging.
* One of most visual processing is remote sensing which usually includes several bands in the visual & infrared regions of the spectrum.
* Table below shows the thematic bands in NASA's LANDSAT satellite.

Band No.	Name	Wavelength (μm)	Chr's & Uses
1.	Visible blue	0.45-0.52	Max. water penetration
2.	Visible green	0.52-0.60	Good for measuring plant vigor
3.	Visible red	0.63-0.69	vegetation discrimination
4.	Near infrared	0.76-0.90	Biomass & shoreline mapping
5.	Middle infrared	1.55-1.75	Moisture content of soil & vegetation
6.	Thermal infrared	10.4-12.5	Soil moisture thermal mapping
7.	Middle infrared	2.08-2.35	Mineral mapping

Subscribe EC Academy on YouTube

1.3.5 Imaging in Microwave Band:-

* The dominant appln in imaging in microwave band is radar.
* The unique feature of imaging radar is its ability to collect data over virtually any region at any time, regardless of weather or ambient lighting conditions. Some radar waves can penetrate clouds & also ~~under~~ through vegetation, ice & extremely dry sand.
* Radar is the only way to explore inaccessible regions of Earth's surface. An imaging radar works like a flash camera in that it provides its own illumination to illuminate an area on the ground & take a snapshot image. Instead of camera lens, a radar uses an antenna & digital computer processing to record its images. In a radar image, one can see only the microwave energy that was reflected back toward the radar antenna.

1.3.6 Imaging in Radio Band:-

* The major appln's of imaging in radio band are in medicine & astronomy.
* In medicine, radio waves are used in magnetic resonance imaging (MRI). This technique places a patient in a powerful magnet & passes radio waves thro' his or her body in short pulses. Each pulse causes a responding pulse of radio waves to be emitted by patient's tissues. The location from which these signals originate & their strength are ... computer which produces a 2-D

Subscribe EC Academy on YouTube

picture of a section of interest. MRI can provide pictures in any plane.

1.3.7 Examples in which Other Imaging Modalities are used:-

* Other Imaging modalities are acoustic imaging, electron microscopy & synthetic (computer-generated) imaging.
* Imaging using "sound" finds appln in geological exploration, industry & medicine and also for marine acquisition.
* Geological appln's use sound in low end of sound spectrum (hundreds of Hz) while imaging in other areas use ultrasound (millions of Hz).
* Ultrasound imaging finds appln's in manufacturing medicine, obstetrics & in sex determination of a baby.
* A transmission electron microscope (TEM) works like a slide projector. A scanning electron microscope (SEM) works much like a TV camera.
* Images that are not obtained from physical objects are generated by computer. Fractals are striking examples of computer generated images. A fractal is an iterative reproduction of a basic pattern according to some mathematical rules.
* A more structured approach to image generation by computer lies in 3-D modelling. Images of this type can be used for medical training & for other appln's such as criminal forensics & special effects.

Subscribe EC Academy on YouTube

1.4 Fundamental Steps in Digital Image Processing:

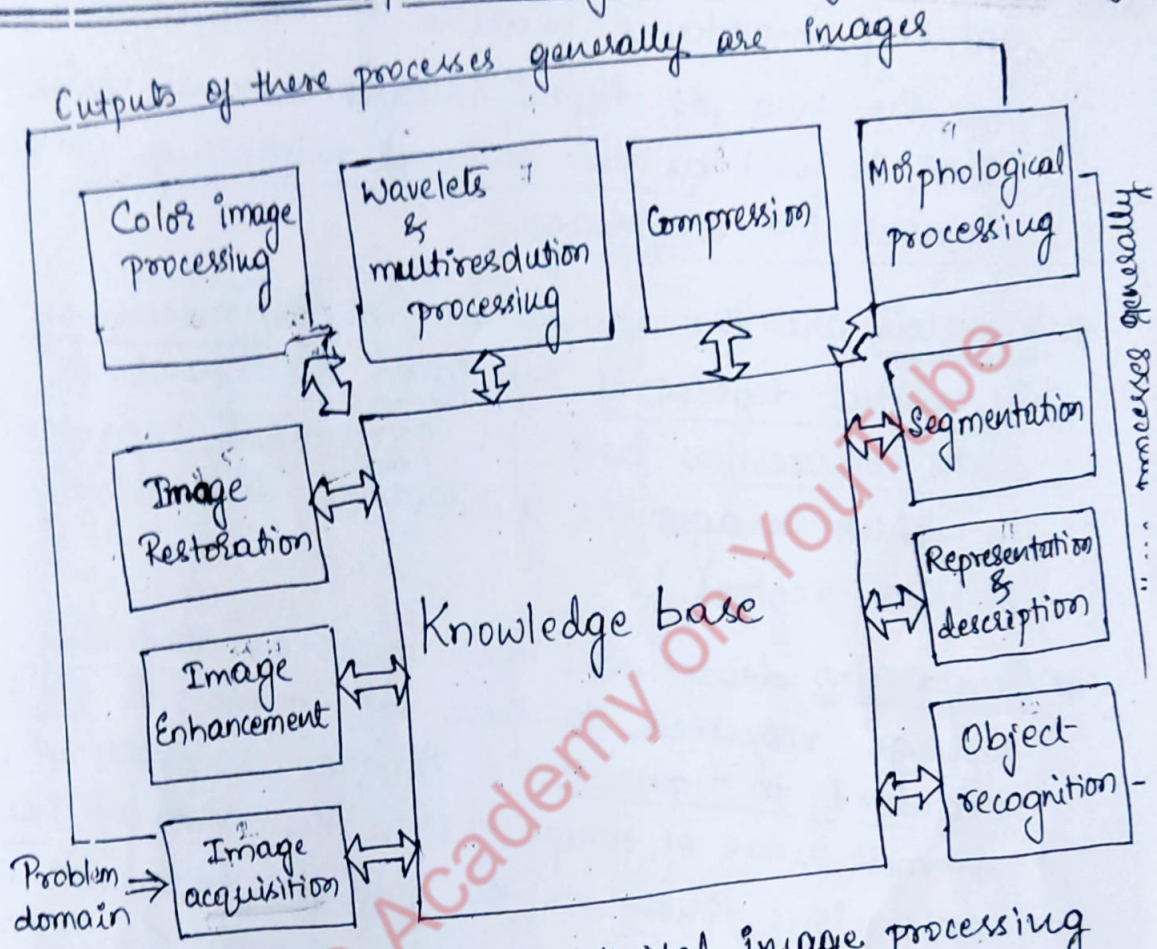


Fig: Fundamental Steps in digital image processing

* Image acquisition is the first process shown in fig. Acquisition could be as simple as being given an image that is already in digital form. This stage involves preprocessing, such as scaling.

* Image enhancement is the simplest process and is use to bringout detail that is obscured or to highlight certain features of interest in an image. Enhancement techniques are based on the Fourier Transforms.

* Image Restoration deals with improving the appearance of an image. Restoration techniques tend to be based on mathematical or probabilistic models of image degradation.

Subscribe EC Academy on YouTube

* Color image processing is an area that has been gaining importance because of significant increase in the use of digital images over the internet. Color is used as the basis for extracting features of interest in an image.

* Wavelets are the foundation for representing images in various degrees of resolution. Also used for image data compression and for pyramidal representation, in which images are subdivided successively into smaller regions.

* Compression deals with techniques for reducing the storage required to save an image, or the bandwidth required to transmit it. Image compression is familiar to most users of computers in the form of image file extensions, such as the jpg file extension used in the JPEG (Joint Photographic Experts Group) image compression method.

* Segmentation procedures partition an image into its constituent parts or objects. In general, autonomous segmentation is one of the most difficult tasks in digital image processing. A rugged segmentation brings the process a long way toward successful solution of imaging problems that require objects to be identified individually. Weak or erratic segmentation always guarantee eventual failure. In general, the more accurate the segmentation, the more likely recognition is to succeed.

* Representation & description always follows the o/p

Subscribe EC Academy on YouTube

of a segmentation stage, which usually is raw pixel data, constituting either the boundary of a region or all the points in the region itself.

In either case, converting the data to a form suitable for computer processing is necessary.

Boundary representation is appropriate when the focus is on external shape characteristics su as corners & inflections.

Regional representation is appropriate whe the focus is on internal properties such as textu or skeletal shape.

In some applications these representation complement each other.

Choosing a representation is only part of the solution for transforming raw data into form suitable for subsequent computer process

Description, also called feature selection, de with extracting attributes that result in some quantitative information of interest or basic differentiating one class of objects from anoth

* Recognition is the process that assigns a label t an object based on its descriptors.

* Knowledge about a problem domain is coded i an image processing system in the form of a knowledge database. This may be as simplex detailing regions of an image where the informat of interest is known to be located, or it can be complex as an interrelated list of all major

Subscribe EC Academy on YouTube

possible defects in a materials inspection problem or an image database containing high-resolution satellite images of a region in connection with change-detection applications.

1.5 Components of an Image Processing System:—

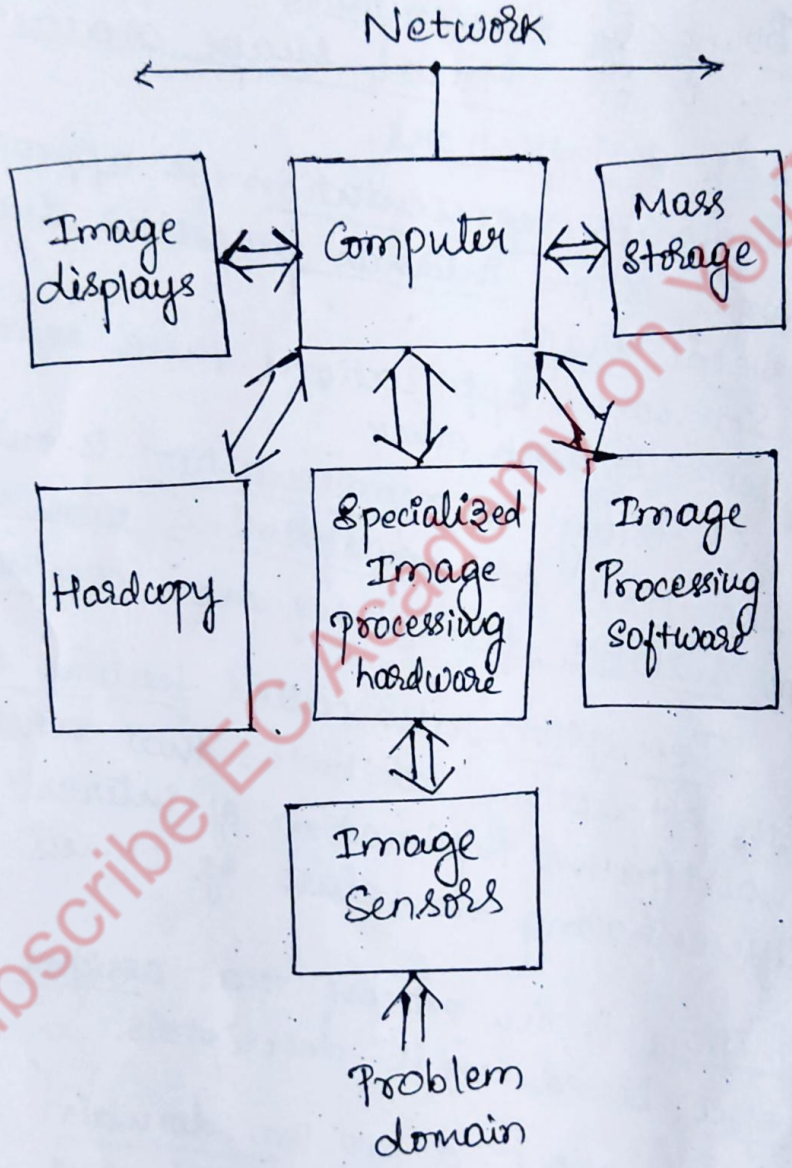


Fig:- Components of a general-purpose image processing system.

Subscribe EC Academy on YouTube

* With reference to sensing, two elements are require to acquire digital images. The first is a physical device that is sensitive to the energy radiated by the object we wish to image. The second called a digitizer, is a device for converting the o/p of the physical sensing device into digital form.

Ex:- In a digital video camera, the sensors produce an electrical output proportional to light intensit The digitizer converts these outputs to digital data

* Specialized image processing hardware consists of the digitizer plus hardware that performs arithmetic logic unit (ALU) in parallel on entire images. Ex:- ALU is used in averaging images as quickly as they are digitized, for the purpose of noise reduction. This type of hardware sometimes is called a front-end subsystem and its distinguishing characteristic is speed.

* The computer used in an image processing syst is a general-purpose computer and can range from a PC to a supercomputer.

* Software consists of specialized modules that perform specific tasks. A well-designed packag includes the capability for user to write code that as a minimum utilizes the specialized modu & more sophisticated software packages allow the integration of those modules.

* Mass storage capability is a must in DIP applica

Subscribe EC Academy on YouTube

An image of size 1024 × 1024 pixels, in which the intensity of each pixel is an 8-bit quantity, requires one megabyte of storage space if the image is not compressed. Thus when dealing with thousands or even millions, of images, providing adequate storage in an image processing system can be a challenge.

Digital storage for image processing applications falls into three principal categories:

(1) Short-term storage for use during processing
(2) On-line storage for relatively fast recall
(3) archival storage, characterized by infrequent access.

Storage is measured in bytes, Kbytes, Mbytes, Gbytes & Tbytes.

One method of providing short-term storage is computer memory. Another is frame buffers that store one or more images & can be accessed rapidly at video rates. Computer method allows instantaneous image zoom, as well as scroll (vertical) and pan (horizontal shifts). Frame buffers are used in specialized image processing hardware unit.

On-line storage takes the form of magnetic disks or optical-media storage.

Archival storage is characterized by massive storage requirements but infrequent need for access. Magnetic tapes & optical disks housed in "jukeboxes" are usual media for archival applications.

Subscribe EC Academy on YouTube

* Image displays in use today are mainly color TV monit
Monitors are driven by o/p's of image & graphics
display cards that are an integral part of compute
system. In some cases, it is necessary to have
stereo displays and these are implemented in
the form of headgear containing two small displa
embedded in goggles worn by the user.

* Hardcopy devices for recording images include
laser printers, film cameras, heat-sensitive device
inkjet units & digital units such as optical &
CD-ROM disks.

* Networking is a default function in any computer
system in use today. Because of large amount of da
inherent in image processing applications, the key
consideration in image tx^n is bandwidth. In
dedicated n/w's, this typically is not a problem
but communications with remote sites via the
Internet are not always as efficient. This situa
is improving quickly as a result of optic fiber
& other broadband technologies.

1.6 Elements of Visual Perception:—

Although the digital image processing field is
built on a foundation of mathematical & probabilisti
formulations, human intuition & analysis play a central
role in the choice of one technique versus another an
this choice is made based on subjective, visual
judgements. Hence, developing a basic understanding of
... is the first step in understand

Subscribe EC Academy on YouTube

the image processing. From a practical point of view, the comparison between how human and electronic imaging adapt to changes in illumination and resolution is important to study.

1.6.1 Structure of the Human Eye:-

9

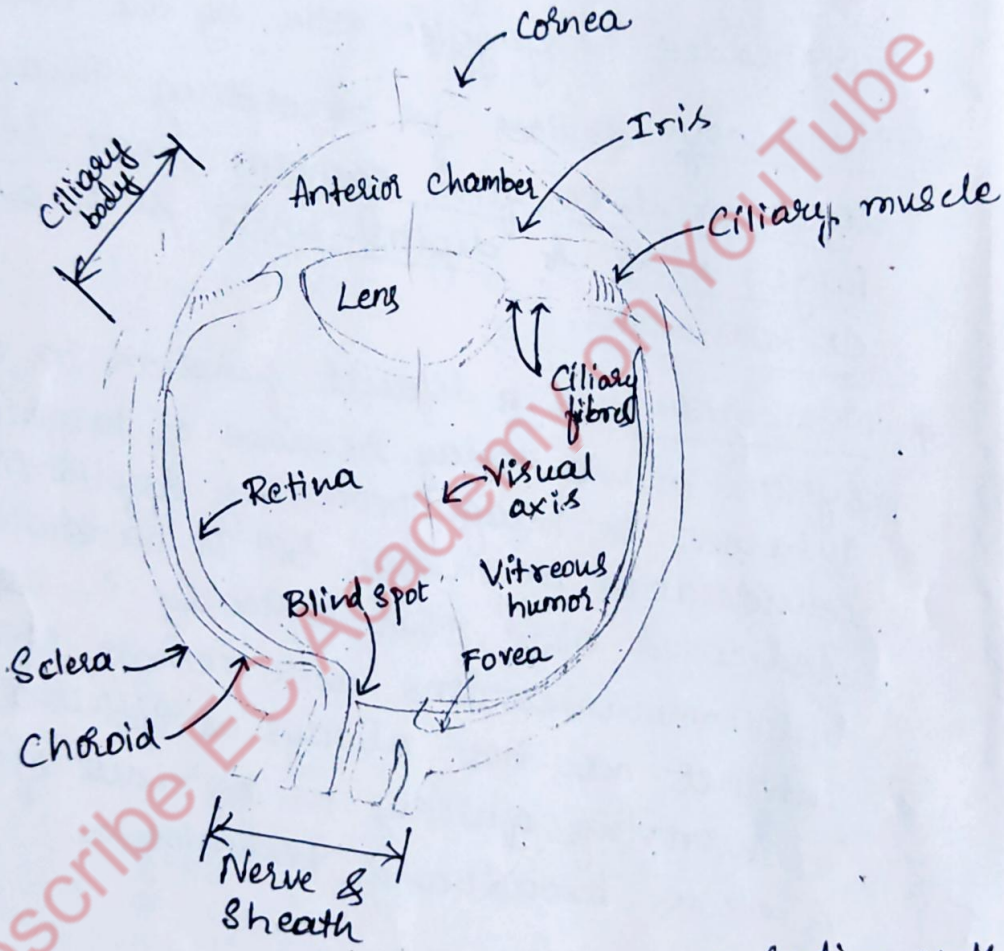


Fig: Simplified diagram of a cross section of the human eye

* The eye is nearly a sphere, with an average diameter of approximately 20mm.
* Three membranes enclose the eye: the 'cornea' & 'Sclera' outer cover; the choroid & the 'retina'.

Subscribe EC Academy on YouTube

* The cornea is a tough, transparent tissue that cover the anterior surface of the eye.
* Continuous with cornea, the Sclera is an opaque membrane that encloses the remainder of the optic globe.
* The choroid lies directly below the sclera. This memb contains a network of blood vessels that serve a the major source of nutrition to the eye. The Choroid coat is heavily pigmented and hence hely to reduce the amount of extraneous light enterin the eye & backscatter within the optical globe. At its anterior extreme, the choroid is divided into the ciliary body & the iris diaphragm. Th iris contracts or expands to control the amou of light that enters the eye. The central openi of the iris (the pupil) varies in diameter from 2 to 8mm. The front of the iris contains the visi pigment of the eye, whereas the back contain black pigment.
* The lens is madeup of concentric layers of fibr cells & is suspended by fibers that attach to the ciliary body. It contains 60 to 70% water, 6% & more protein.
* 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 focuss light from an object is imaged on the retina. Pattern vision is afforded by the distribution of discrete light receptors over the surface of retina.

Subscribe EC Academy on YouTube

There are two classes of receptors: cones & rods.
The cones in each eye number between 6 & 7 million. They are located primarily in the central portion of the retina, called the fovea, & are highly sensitive to color. 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.

Rods are 75 to 150 million in number are distributed over the retinal surface. Rods serve to give a general, overall picture of the field of view. They are not involved in color vision & are sensitive to low levels of illumination. For example, objects that appear brightly colored in day-light may appear as colorless when seen by moonlight because only the rods are stimulated. This phenomenon is known as scotopic or dim-light vision

* Fig below shows the density of rods & cones for a cross section of right eye passing thr' the region of emergence of the optic nerve from the eye.

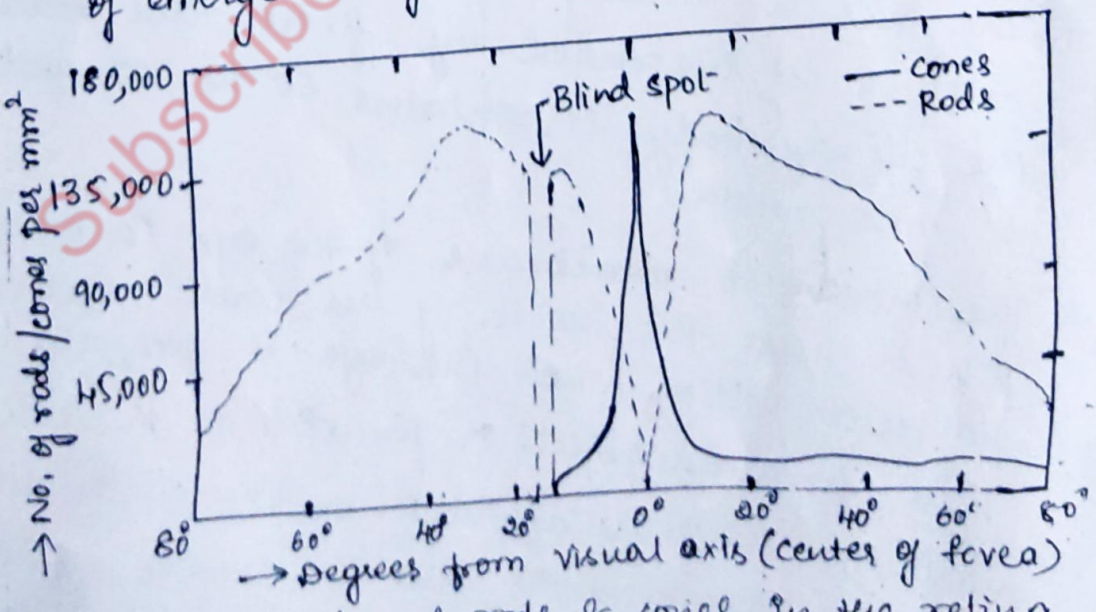


Fig: Distribution of rods & cones in the retina.

Subscribe EC Academy on YouTube

* The absence of receptors in this area results in so-called blind spot. Except this region, the distribution of receptors is radially symmetric about the fovea. Receptor density is measured degrees from the fovea. The cones are most dense in the center of retina. The rods increase in densi from center out to $\cong 20^\circ$ off axis & then decrea in density out to extreme periphery of the retina.

1.6.2 Image Formation in the Eye:-

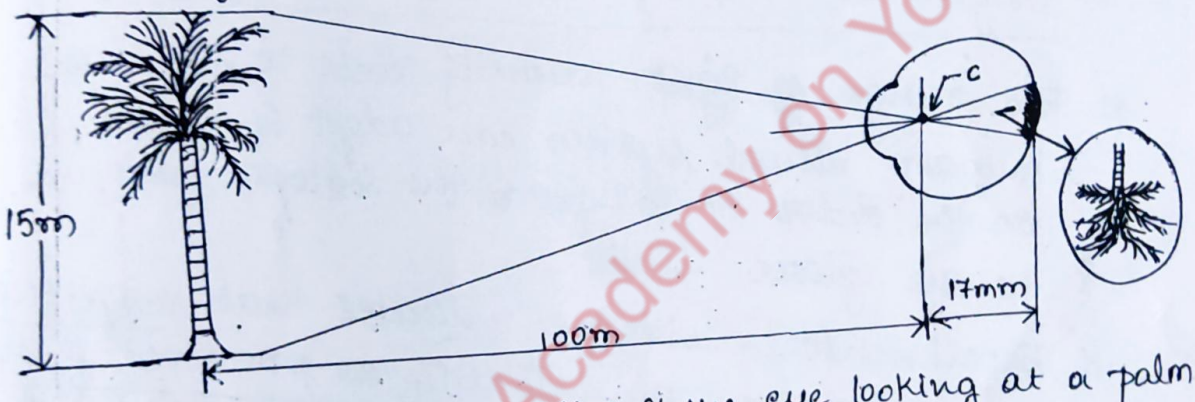


Fig: Graphical representation of the eye looking at a palm tree
Point C is the optical center of the lens.

The distance between the center of the lens & the retina (called the focal length) varies from approximate 17mm to about 14mm, as the refractive power of the le increases from its minimum to its maximum.

When the eye focusses on an object farther awa than about 3m, the lens exhibits its lowest refractiv power. When the eye focusses on a nearby object, the lens is most strongly refractive. This informati makes it easy to calculate the size of the retinal imag of any object.

In fig. the observer is looking at a tree 15m high at a distance of 100m. If h is the height in mm,

Subscribe EC Academy on YouTube

of that object in the retinal image, the geometry of fig. yields $\frac{15}{100} = \frac{h}{17} \Rightarrow h = 2.55mm$.

As indicated, 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.

1.6.3 Brightness Adaption & Discrimination:-

* The range of light intensity levels to which the human visual system can adapt is enormous - on the order of 10^{10} - from the scotopic threshold to the glare limit.
* Experimental evidence indicates that subjective brightness i.e., the intensity as perceived by the human visual system is a logarithmic function of the light intensity incident on the eye.

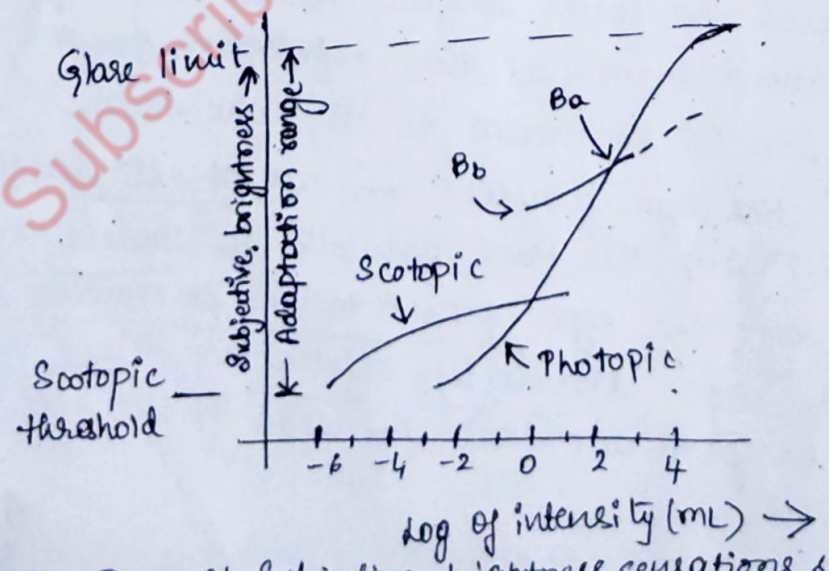


Fig:- Range of subjective brightness sensations showing a particular adaptation level.

Subscribe EC Academy on YouTube

* Fig. below shows a plot of light intensity versus subjective brightness. The long curve represents the range of intensities to which the visual system can adapt. In photopic vision alone, the range is about 10^6. 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 logscale) as the double branches of the adaptation curve in this range show.

* The visual system accomplishes large variation by changing in its overall sensitivity. This phenomenon is known as <u>brightness adaptation level</u>, which may correspond, for example to brightness B_a in fig.

The short intersecting curve represents the range of subjective brightness that the eye can perceive when adapted to this level. This range is rather restricted having a level B_b at & below which all stimuli are perceived as indistinguishable blacks. The upper (dashed) portion of the curve if extended too far, loses its meaning because much higher intensities would simply raise the adaptation level higher than B_a.

* A basic experiment is setup to <u>determine the capability of human visual system for brightness discrimination</u>. Consider a opaque glass illuminated from behind by a light source whose intensity I can be varied. To this field is added an increment of illumination ΔI in the form of a short-duration flash that appears as a circle in the center of the uniformly illuminated field.

Subscribe EC Academy on YouTube

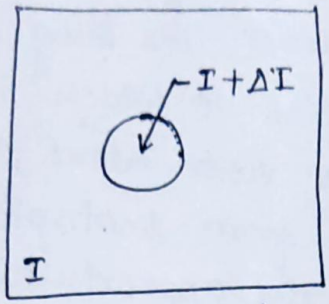


Fig:- Basic experimental setup used to characterize brightness discrimination.

* If ΔI is not bright enough, the subject says "no", indicating no perceivable change. As ΔI gets stronger, the subject may give a positive response of "yes", indicating a perceived change. Finally, when ΔI is strong enough, the subject will give a response of "yes" all the time.

* The quantity $\Delta I_c / I$ is called the Weber ratio. where ΔI_c is the increment of illumination discriminable 50% of the time with background illumination I.

A small value of $\Delta I_c / I$, means that a small percentage change in intensity is discriminable. This represents "good" brightness discrimination.

A large value of $\Delta I_c / I$ means that a large percentage change in intensity is required. This represents "poor" brightness discrimination.

* A plot of $\log \Delta I_c / I$, as a function of $\log I$ has the general shape as shown in fig.

Subscribe EC Academy on YouTube

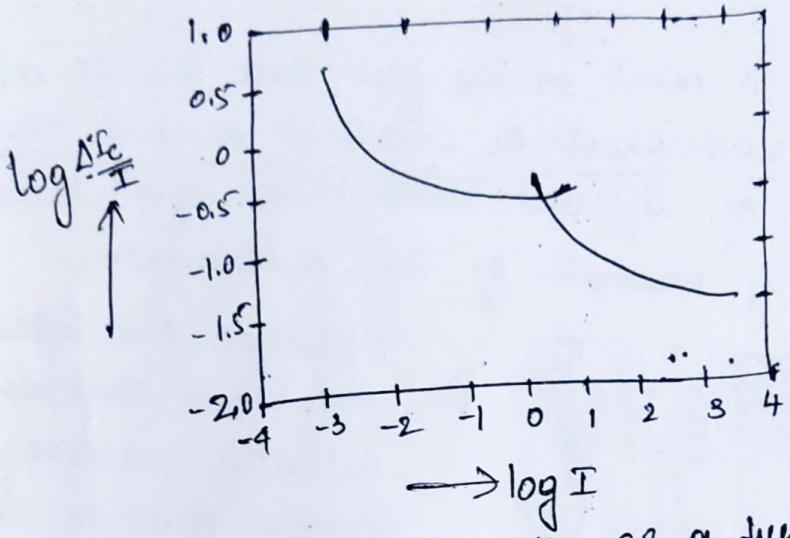


Fig:- Typical Weber ratio as a function of intensity

This curve shows that brightness discrimination is p at low levels of illumination & it improves significant as background illumination increases.

The two branches in the curve reflects + fact that at low levels of illumination vision is carried out by activity of rods, whereas at high level vision is the function of cones.

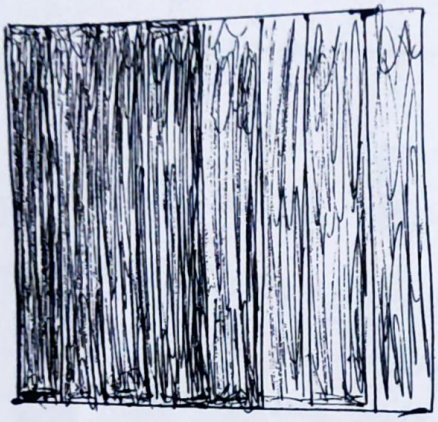
If the background illumination is held const & the intensity of other source, instead of flashing the observer can see a total of one to two dozen different intensity changes. This means, a person sees no. of different intensities at any one point in a monochrome image. Thus the eye is capable of detecting objectionable contouring effects in monochrome images whose overall intens is represented by fewer than approximately two dozen levels.

* Two phenomena clearly demonstrate that perceived brightness is not a simple function of intensit

Subscribe EC Academy on YouTube

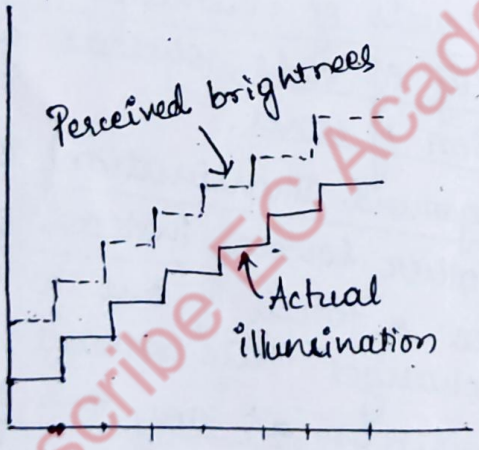
ii) Simultaneous contrast.

* The first is based on the fact that the visual system tends to undershoot or overshoot around the bounda of regions of different intensities. Fig(a) below show the striking example of this phenomenon.



fig(a)

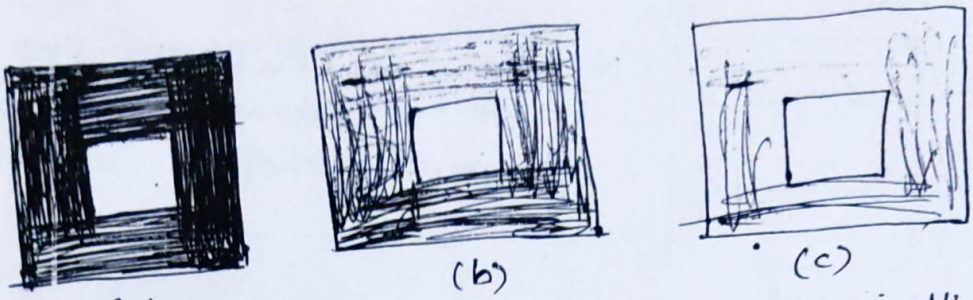
Although the intensity of stripes is constant, we actually perceive a bright pattern that is strongly scalloped, especially near boundaries. [fig(b)]. These seemingly scalloped ba are called Mach bands.



fig(b)

* The second phenomenon called simultaneous con is related to fact that a region's perceived bright does not depend simply on its intensity as fig(c) demonstrates. All the center squares have exactly same intensity. However they appear to the eye become darker as the background gets lighter.

Subscribe EC Academy on YouTube



(a) (b) (c)

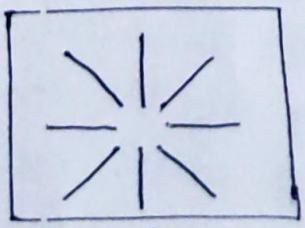
fig(1): Examples of simultaneous contrast. All the inn squares have the same intensity, but they appear progressively darker as the background becomes lighter.

* A more familier example is a piece of paper that seems white when lying on a desk, but can appea totally black when used to shield the eyes while looking directly at a bright sky.

* Other examples of human perception phenomena are optical illusions, in which the eye fills in nonexisti information or wrongly perceives geometrical propertie of objects. Some examples are shown in fig. below

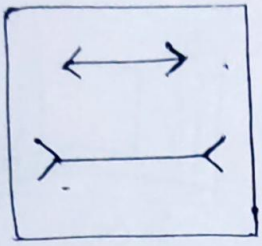
(a)

← The outcome of a square is see clearly in spite of the fact that no lines defining such a figure are part of the image.



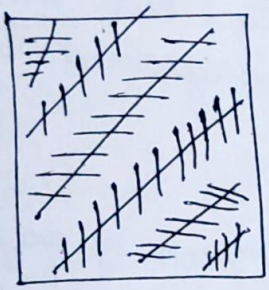
← The same effect with a circle. Ve that how just a few lines are sufficient to give the illusion o a complete circle.

Subscribe EC Academy on YouTube



(c)

← The two horizontal line segments are of same length, but one appears shorter than the other.



(d)

← All lines are oriented at 45° are equidistant and parallel. The crosshatching creates the illusion that those lines are far from being parallel.

Optical illusions are a characteristic of the human visual system that is not fully understood.

1.7 Image Sensing and Acquisition:-

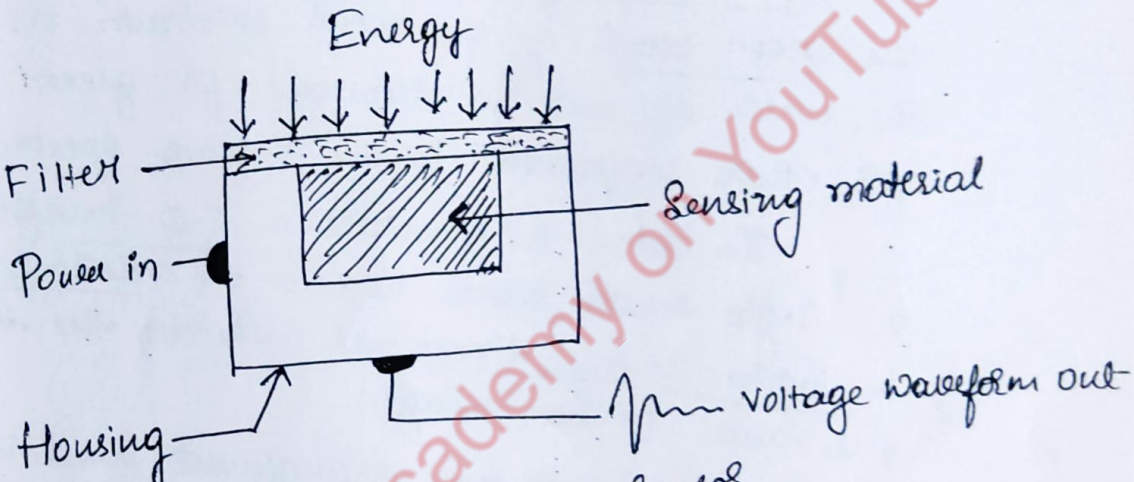
* The types of images in which we are interested are generated by the combination of an "illumination" source & the reflection or absorption of energy from that source by the elements of the "scene" being imaged.

* Fig. below shows the three principal sensor arrangements used to transform illumination energy into digital images.

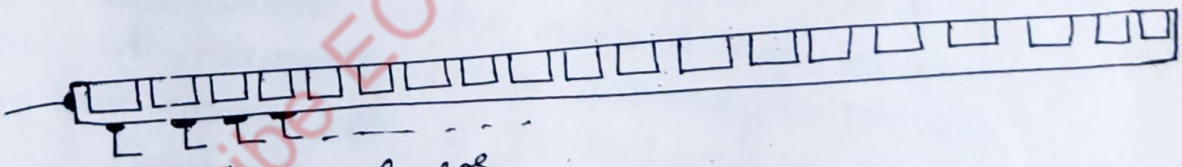
The idea is simple: Incoming energy is transformed into a voltage by the combination of input electrical power & sensor material that

Subscribe EC Academy on YouTube

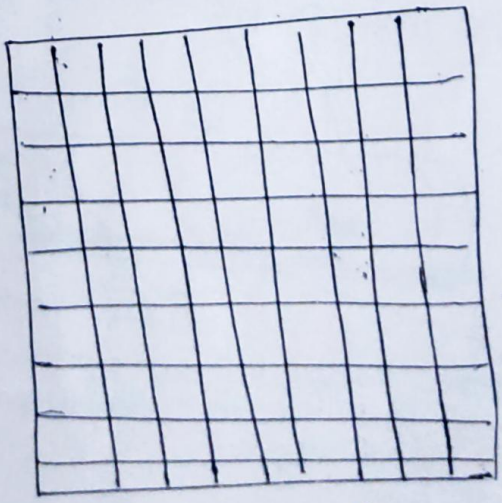
~~responsive~~ responsive to particular type of energy being detected. The o/p voltage waveform is the response of the sensor(s) & a digital quantity is obtained from each sensor by digitizing its response.



Fig(1)(a): Single Imaging sensor



fig(1) (b) Line Sensor



fig(1). (c) Array Sensor

Subscribe EC Academy on YouTube

3

1.7.1 Image Acquisition Using a Single Sensor :—

Fig (a) in the previous page shows the components of a single sensor. The most familiar sensor of this type is the photodiode, whose o/p voltage waveform is proportional to light. The use of a filter in front of a sensor improves selectivity. For 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 o/p will be stronger for green light than for other components in the visible spectrum.

In order to generate a 2-D image using a single sensor, there has to be relative displacements in both x- & y-directions between the sensor & the area to be imaged.

Fig. 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.

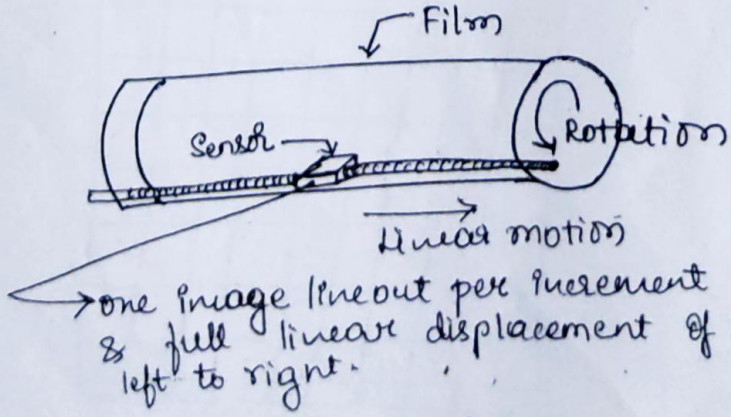


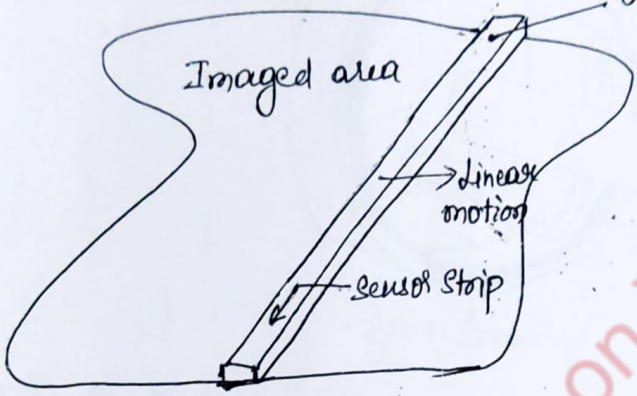
Fig (ii): Combining a single sensor with motion to generate a 2-D image.

Subscribe EC Academy on YouTube

1.7.2 Image Acquisition using Sensor Strips:-

8

* fig(i)(b) shows a sensor strip. The strip provides imagi elements in one direction. Motion perpendicular to t strip provides imaging in the other direction a shown in fig(iii)(a).



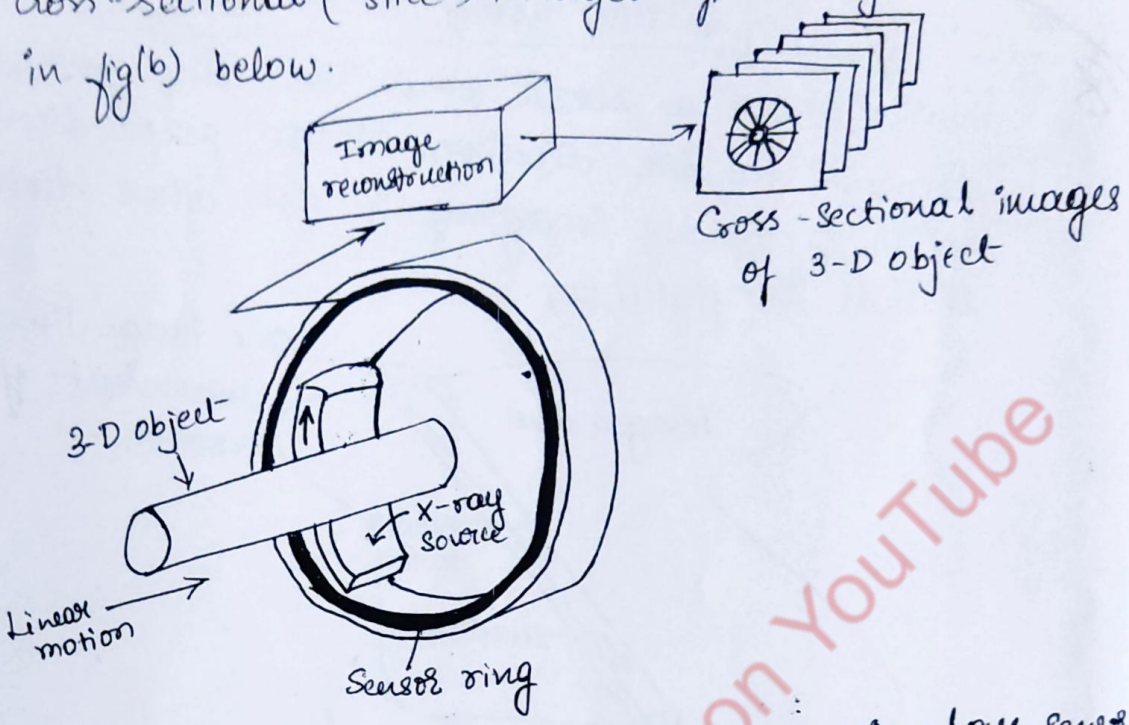
fig(iii)(a): Image acquisition using a linear sensor str

* Sensing devices with 4000 or more in-line sensor are possible. In-line sensors are used in airborne imaging applications, in which the imaging system i mounted on an aircraft that flies at a constant altitude & speed over geographical area to be imaged. One-dimensional imaging sensor strip that respond to various bands of electromagnetic strips that respond to various bands of EM spectrum are mounted ⊥r to the direction of flig The imaging strip gives one line of an image at a time, & the motion of the strip completes the other dimension of a two-dimensional imag lenses or other focussing schemes are used to project the area to be scanned onto the sensors.

* Sensor strips mounted in a ring configuration a used in medical & industrial imaging to obtain

Subscribe EC Academy on YouTube

cross-sectional ("slice") images of 3D objects, as shown in fig(b) below.



fig(iii)(b): Image acquisition using a circular sensor strip

A rotating X-ray source provides illumination, and the portion of the sensors opposite the source called the X-ray energy that pass through the object. This is the basis for medical & industrial CAT imaging.

The o/p of the sensors must be processed by reconstruction algorithms whose objective is to transform the sensed data into meaningful cross-sectional images.

A 3-D digital volume consisting of stacked images is generated as the object is moved in a direction perpendicular to the sensor ring.

Other modalities of imaging based on the CAT principle include magnetic resonance imaging (MRI) & Positron emission tomography (PET). The illumination sources, sensors & types of images are different but very similar to basic imaging approach shown in fig(b)

Subscribe EC Academy on YouTube

1.7.3 Image Acquisition Using Sensor Arrays :-

The principal manner in which array sensors are used is shown in fig.

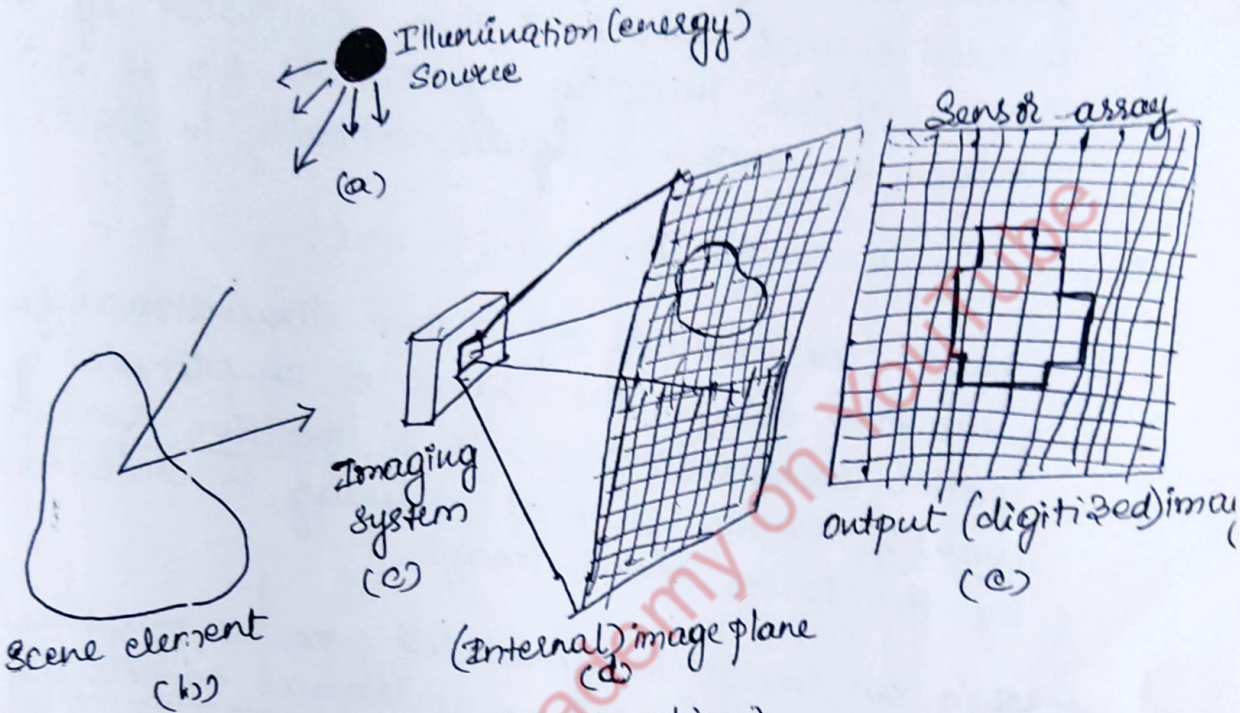


Fig (iv): (a) Energy source (illumination)
(b) An element of a scene
(c) Imaging system
(d) Projection of the scene onto the image plane
(e) Digitized image.

This figure shows the energy from an illumination sour being reflected from a scene element but energy could also be transmitted through scene elements.

The first function performed by the imaging system shown in fig(c) is to collect the incoming energy & focus it onto an image plane.

If the illumination is light, the front end the imaging system is a lens, which projects the viewed scene onto the lens focal plane, as fig(d) sh

Subscribe EC Academy on YouTube

The sensor array which is coincident with the focal plane, produces o/p's proportional to the integral of the light received at each sensor. Digital & analog circuitry sweep these o/p's & convert them to a video signal, which is then digitized by another section of the imaging system. The o/p is a digital image, as shown diagrammatically in fig(e).

1.7.4 A Simple Image Formation model :—

* Images are 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 from a physical process, its values are proportional to energy radiated by a physical source (e.g., EM waves). As a consequence, $f(x,y)$ must be nonzero and finite i.e., $0 < f(x,y) < \infty$ —— (1)

* The function $f(x,y)$ may be 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.
Appropriately, these are called the illumination & reflectance components and are denoted by $i(x,y)$ & $r(x,y)$ respectively.
The two functions combine as a product to the form $f(x,y)$: $f(x,y) = i(x,y)\, r(x,y)$ —— (2)

Subscribe EC Academy on YouTube

where $0 < i(x,y) < \infty$ — (3)

& $0 < r(x,y) < 1$ — (4)

* The values given in equations (3) & (4) are theoretical bounds. The following average numerical figures illustrate some typical ranges of $i(x,y)$ for visible light.

On a clear day, the sun may produce in excess 90,000 lm/m² of illumination on the surface of the earth. This figure decreases to less than 10,000 lm/m² on a cloudy day. On a clear evening, a full moon yields about 0.1 lm/m² of illumination. The typical illumination level in a commercial office is about 1000 lm/m².

Similarly, the following are some typical values of $r(x,y)$: 0.01 for black velvet, 0.65 for stainless steel, 0.80 for flat-white wall paint, 0.90 for silver-plated metal and 0.93 for snow.

* The intensity of a monochrome image at any coordinates (x_0, y_0) is called the gray level (l) of the image at that point

i.e., $l = f(x_0, y_0)$ — (5)

From equations (2) through (4), it is evident that 'l' lies in the range,

$$L_{min} \leq l \leq L_{max} \quad \text{— (6)}$$

* In theory, L_{min} is to be positive & L_{max} is to be finite.
In practice, $L_{min} = i_{min} r_{min}$ & $L_{max} = i_{max} r_{max}$.

Using the preceeding average office illumination & range of reflectance values as guidelines, we expect

Subscribe EC Academy on YouTube

$L_{min} \approx 10$ & $L_{max} \approx 1000$ to be typical limits for indoor values in the absence of additional illumin. The interval $[L_{min}, L_{max}]$ is called Gray scale. This interval is shifted numerically to the interval $[0, L-1]$, where $l=0$ is considered black & $l=L-1$ is considered white on the gray scale. All intermediate values are shades of gray varying from black to white.

1.8 Image Sampling and Quantization:-

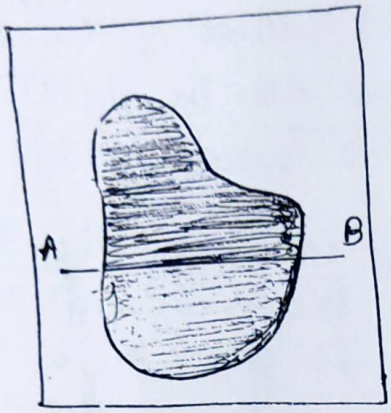
* Our objective is to generate digital images from sensed data.
* The output of most sensors is a continuous voltage waveform whose amplitude & spatial behavior are related to the physical phenomenon being sensed.
* To create a digital image, we need to convert the continuous sensed data into digital form. This involves two processes: Sampling & quantization.

1.8.1 Basic concepts in Sampling and Quantization:-

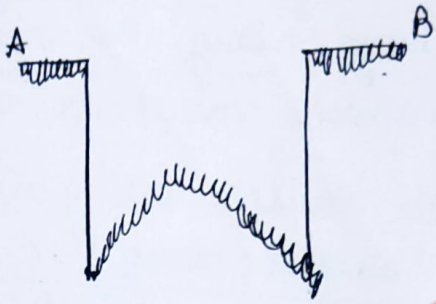
The basic idea behind sampling and quantization is illustrated in fig. below.

Fig(a) shows a continuous image $f(x,y)$ which we want to convert to digital form. An image may be continuous w.r.t. the x- & y- coordinates and also in amplitude.

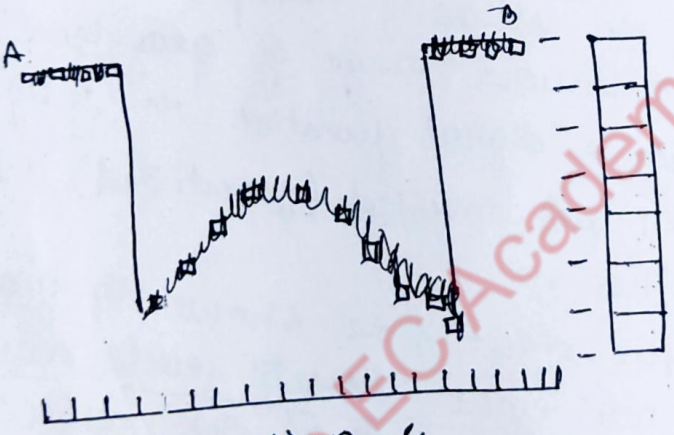
Subscribe EC Academy on YouTube



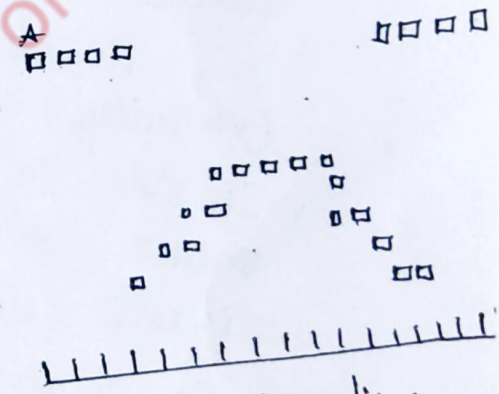
(a) Continuous image



(b) A scanline from A to B the continuous image, used to illustrate the conce of sampling & Quantizati



(c) Sampling & quantization



(d) Digital scan line

Fig: Generating a digital image.

To convert it to digital form, we have to sample the function in both coordinates & in ampli

[Digitizing the coordinate values is called 'Sampling'. Digitizing the amplitude values is called 'quantization'.]

Subscribe EC Academy on YouTube

* The one-dimensional function shown in fig(b) is a plot of amplitude (gray level) values of the continuous image along the line segment AB in fig(a). The random variations are due to image noise.

* To sample this function, we take equally spaced samples along line AB as shown in fig(c). The location of each sample is given by a vertical tick mark in the bottom part of the figure. The samples are shown as small white squares superimposed on the function. The set of these discrete locations gives the sampled function.

However, the values of samples still span (vertically) a continuous range of gray-level values. In order to form a digital function, the gray-level values also must be converted (quantized) into discrete quantities.

In Fig(c), the right side shows the gray-level scale divided into eight discrete levels, ranging from black to white. The vertical tick marks indicate the specific value assigned to each of the eight gray levels. The continuous gray levels are quantized simply by assigning one of the eight discrete gray levels to each sample. The assignment is made depending on the vertical proximity of a sample to a vertical tick mark. The digital samples resulting from both sampling & quantization are shown in fig(d).

Subscribe EC Academy on YouTube

Starting at the top of the image & carrying out this procedure line by line produces a two-dimensional digital image.

* Sampling the image in the above manner assumes that we have a continuous image in both coordinal directions as well as in amplitude.

In practice, 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, the o/p of the sensor is quantized in the manner discribed above.

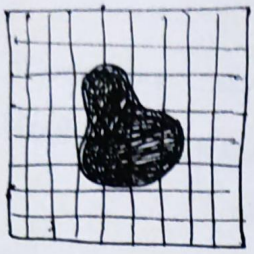
However, sampling is accomplished by selecting the no. of individual mechanical increment at which we activate the sensor to collect data. Mechanical motion can be made very exact so there is almost no limit as to how fine we can sample an image. However, practical limits are establis by imperfections in the optics used to focus on the sensor an illumination spot that is inconsistent with the fine resolution achievable with mechanical displacements.

When a sensing strip is used for image acquisition, the no. of sensors in the strip establish the sampling limitations in one image direction. Quantization of the sensor outputs completes the process of generating a digital image.

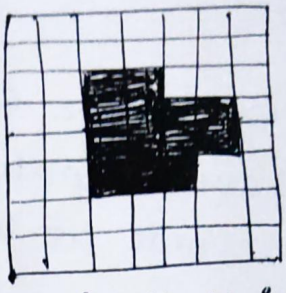
When a sensing array is used for image ..... there is no motion & the no. of

Subscribe EC Academy on YouTube

sensors in the array establishes the limits of sampling in both directions. Fig. below illustrates this concept.



(a) Continuous image projected on a sensor array



(b) Result of image sampling & Quantization.

Clearly, the quality of a digital image is determined to a large degree by the no. of samples and discrete gray levels used in sampling & quantization.

1.8.2 Representing Digital Images :-

* The result of sampling & quantization is a matrix of real numbers.
* Assume that an image $f(x, y)$ is sampled so that the resulting digital image has M rows & N columns. The values of coordinates (x, y) now become discrete quantities.
* The values of the coordinates at the origin are $(x, y) = (0, 0)$. The next coordinate values along the first row of image are represented as $(x, y) = (0, 1)$.
* Fig. below shows the coordinate convention to represent digital images.

Subscribe EC Academy on YouTube

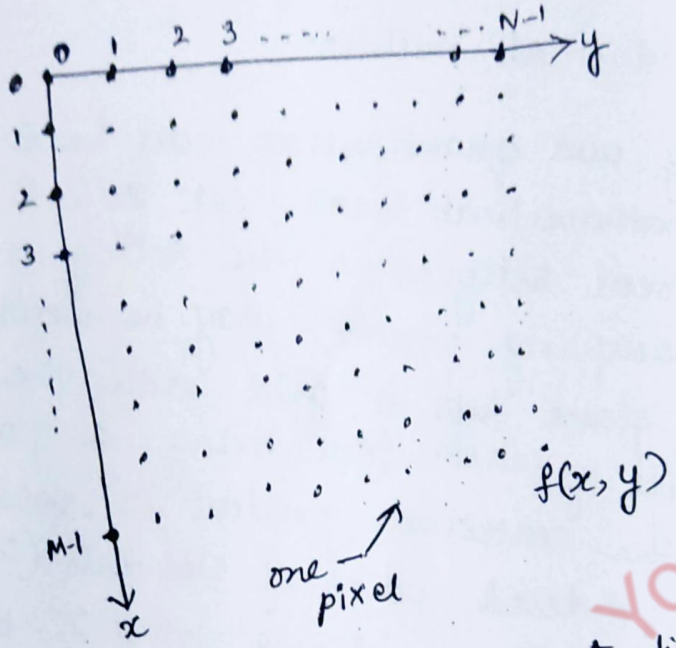


Fig: Coordinate convention to represent digital images.

* The notation used $f(x,y)$ allows us to write the complete $M \times N$ digital image in the following matrix form:

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

Each element of this matrix array is called an image element, picture element, pixel or pel. The right side of the eqn (1) represents a digital image.

* A traditional matrix notation to denote a digital image and its elements are given below:

$$A = \begin{bmatrix} a_{0,0} & a_{0,1} & \cdots & a_{0,N-1} \\ a_{1,0} & a_{1,1} & \cdots & a_{1,N-1} \\ \vdots & & & \vdots \\ a_{M-1,0} & a_{M-1,1} & \cdots & a_{M-1,N-1} \end{bmatrix} \quad \text{--- (2)}$$

Subscribe EC Academy on YouTube

Clearly, $a_{ij} = f(x=i, y=j) = f(i,j)$. so eqns (1) & (2) are identical matrices.

* Sampling and quantization can be expressed in more mathematical terms. Let Z & R denote the set of real integers & the set of real numbers. The sampling process may be viewed as partioning the xy plane into a grid, with the coordinates of the center of each grid being a pair of elements from the cartesian product Z^2, which is the set of all ordered pairs of elements (z_i, z_j) with z_i & z_j being integers from Z. Hence $f(x,y)$ is a digital image if (x,y) are integers from Z^2 & f is a function that assigns a gray-level value to each distinct pair of coordinates (x,y). The functional assignment is the quantization process. If the gray levels are integers, Z replaces R & a digital image then becomes a 2-D function whose coordinates & amplitude values are integers.

* The no. of gray levels typically is an integer power of 2: $L = 2^k$ — (3)

The discrete levels are equally spaced and they are integers in the interval $[0, L-1]$. The range of values spanned by the gray scale is called the dynamic range of an image.

An image with low dynamic range ...

Subscribe EC Academy on YouTube

to have a dull, washed out gray look. The image with high dynamic range spans a significant portion of the gray scale.

The no. of bits 'b' required to store a digitized image is,

$$b = M \times N \times K \quad \text{—— (4)}$$

where $M = N$, this equation becomes,

$$b = N^2 K \quad \text{—— (5)}$$

* If an image has 2^K gray levels, then that type of image is referred to as "K-bit image". For example an image with 256 possible gray-level values is called an 8-bit image.

8.3 Spatial and Gray-Level Resolution:—

(6)

* Sampling is the principal factor of determining the spatial resolution of an image.

Spatial resolution is the smallest detail in a image. Suppose a chart is constructed with vertical lines of width 'W'. A line pair consists of one such line & its adjacent space. Thus, the width of a line pair is 2W and there are 1/2W line pairs per unit distance. A widely used definition of resolution is simply the smallest number of discernible line pairs per unit distance. For example, 100 line pairs per millimeter.

* Gray-level resolution refers to the smallest discernible change in gray-level. Due to hardware consideratio

Subscribe EC Academy On YouTube

the no. of gray levels is usually an integer power of 2.

An L-level digital image of size MXN as having a spatial resolution of MXN pixels and gray-level resolution of L levels.

* Consider an image of size 1024X1024 pixels whose gray levels are represented by 8-bits. The subsampling was accomplished by deleting the appropriate no. of rows and columns from the original image.

1.8.4 Aliasing and Moire' Patterns:-

* Functions whose area under the curve is finite can be represented in terms of sines and cosines of various frequencies. The sine/cosine component with the highest frequency determines the highest "frequency content" of the function.

* Suppose this highest frequency is finite & the function is of unlimited duration (these functions are called band-limited functions). Then, the Shanon sampling theorem tells that, if the function is sampled at a rate equal to or greater than twice its highest frequency, it is possible to recover completely the original function from its samples.

* If the function is undersampled, then a phenomenon called aliasing corrupts the sampled image.

Subscribe EC Academy on YouTube

corruption is in the form of additional frequency components being introduced into the sampled function. These are called aliased frequencies.

Sampling rate in images is the number of samples taken (in both spatial directions) per unit distan

* The effect of aliased frequencies can be seen und the right conditions in the form of so called "Moire patterns".

* There is one special case of significant importan in which a function of infinite duration can be sampl over a finite interval without violating the sampling theorem. When a function is periodic, it may b sampled at a rate equal to or exceeding twi its highest frequency & it is possible to recon the functions from its samples provided that th sampling captures exactly an integer no. of peri of function. This special case allows us to illust vividly the Moire effect.

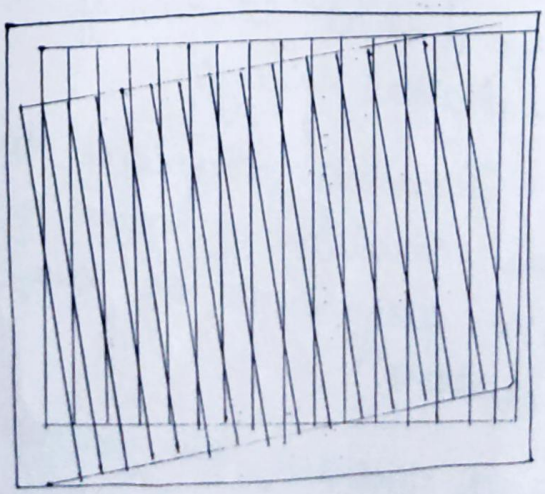


Fig: Illustration of the Moire pattern effect.

Subscribe EC Academy on YouTube

* Fig. shows two identical periodic patterns of equal spaced vertical bars, rotated in opposite directions then superimposed on each other by multiplying the two images.

* A Moire pattern, caused by a break-up of the periodicity is seen in fig as a 2-D sinusoidal waveform running in vertical direction.

A similar pattern can appear when images are digitized (eg., scanned) from a printed page which consists of periodic ink dots.

1.8.5 Zooming and Shrinking Digital Images :-

Zooming & shrinking are related to image sampling and quantization because Zooming may be viewed as oversampling, while shrinking may be viewed as undersampling.

The key difference b/w these two operations and sampling & quantizing an original continuous image is that Zooming & shrinking are applied to a digital image.

Zooming requires two steps:

i) The creation of new pixel locations &

ii) the assignment of gray levels to those new locations.

Let us consider an example. Suppose that we have an image of size 500X500 pixels and

SubscribeECAcademy on YouTube

we want to enlarge it 1.5 times to 750x750 pixels.

One of the easiest way to zoom an image is to lay an imaginary 750x750 grid over the Original image. Obviously, the spacing in the grid would be less than one pixel because we are fitting it over a smaller image. In order to perform gray-level assignment for any point in the overlay, we look for the closest pixel in the original image & assign its grey level to the new pixel in the grid.

When we are done with all points in the overlay grid, we simply expand it to the original specified size to obtain the zoomed image. This method of gray-level assignment is called 'nearest neighbor interpolation'.

'Pixel replication' is applicable when we want to increase the size of an image an integer number of times. For example, to double the size of an image, we can duplicate each column. This doubles the image size in horizontal direction. Then, we duplicate each row of the enlarged image to double the size in the vertical direction. The same procedure is used to enlarge the image by any integer number of times.

Duplication is just done the required number of times to achieve the desired size. The gray-level assignment of each pixel is

Subscribe EC Academy on YouTube

pre-determined by the fact that new locations are exact duplicates of old locations.

Although nearest neighbor interpolation is fast, it has the undesirable feature that it produces a checkerboard effect that is particularly objectionable at high factors of magnification.

A slightly more sophisticated way of accomplishing grey-level assignments is bilinear interpolation using the four nearest neighbors of a point.

Let (x', y') denote the coordinates of a point in the zoomed image, & let $v(x', y')$ denote the gray level assigned to it. For bilinear interpolation, the assigned gray level is given by, $\boxed{v(x', y') = ax' + by' + cx'y' + d}$

where the four coefficients are determined from four equations in four unknowns that can be written using the four nearest neighbors of point (x', y').

Image shrinking is done in a similar manner as described for zooming. The equivalent process of pixel replication is row-column deletion. For example, to shrink an image by one-half, we delete every other row & column.

We can use the zooming grid analogy to visualize the concept of shrinking by a noninteger factor, except that we now expand

Subscribe EC Academy on YouTube

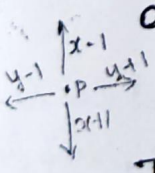
the grid to fit over the original image, do gray-level nearest neighbor or bilinear interpolation & then shrink the grid back to its original specified size. To reduce possible aliasing effects, it is a good idea to blur an image slightly before shrinking it.

It is possible to use more neighbors for interpolation. Using more neighbors implies fitting the points with a more complex surface, which generally gives smoother results. This is an important consideration in image generation for 3D graphics and in medical image processing where bilinear interpolation is the method of choice.

1.9 Some Basic Relationships between Pixels:-

1.9.1 Neighbors of a Pixel:

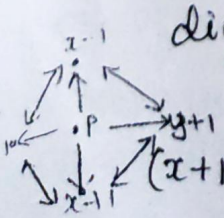
A pixel 'p' at coordinates (x, y) has four horizontal and vertical neighbors whose coordinates are given by



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

This set of pixels called the 4-neighbors of p, is denoted by $N_4(p)$.

Each pixel is a unit distance from (x, y) & some of the neighbors of p lie outside the digital image if (x, y) is on the border of the image.



The 4 diagonal neighbors of p have coordinates, $(x+1, y+1), (x+1, y-1), (x-1, y+1), (x-1, y-1)$ and are

Subscribe EC Academy on YouTube

denoted by $N_D(p)$.

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

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.

1.9.2 Adjacency, Connectivity, Regions and Boundaries:-

* Connectivity between pixels is a fundamental concept that simplifies the definition of numerous digital image concepts, such as regions and boundaries.

* If two pixels are connected, it must be determined if they are neighbors and if their gray levels satisfy a specified criterion of similarity. For example, in a binary image with values 0 & 1, two pixels may be 4-neighbors, but they are said to be connected only if they have the same value.

* Let V be the set of gray-level values used to define adjacency. In a binary image, $V = \{1\}$ if we are referring to adjacency of pixels with value 1.

In a grey-scale image, set 'V' contains more elements. For example, in the adjacency of pixels with a range of possible grey-level values 0 to 255, set V could be any subset of 256 values.

Subscribe FC Academy on YouTube

* We consider 3 types of adjacency:

(a) <u>4-adjacency</u>: Two pixels p & q with values from V are 4-adjacent if q is in the set $N_4(p)$.

(b) <u>8-adjacency</u>: Two pixels p & q with values from V are 8-adjacent if q is in the set $N_8(p)$.

(c) <u>m-adjacency (mixed adjacency)</u>: Two pixels p & q with values from V are m-adjacent if

(i) q is in $N_4(p)$ or

(ii) q is in $N_D(p)$ & the set $N_4(p) \cap N_4(q)$ has no pixels whose values are from V.

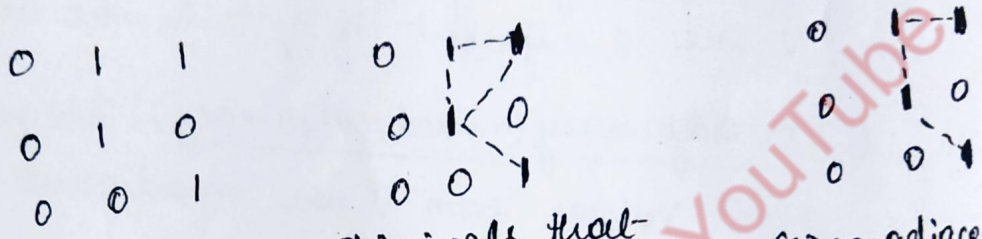
* Mixed adjacency is a modification of 8-adjacency. It is introduced to eliminate the ambiguities that arise when 8-adjacency is used. For example, consider the pixel arrangement shown in fig(a) for V = {1}. The three pixels at the top of fig(b) show multiple 8-adjacency, as indicated by the dashed lines. This ambiguity is removed by using m-adjacency as shown in fig(c). Two image subsets S_1 & S_2 are adjacent if some pixel in S_1 is adjacent to some pixel in S_2.

* A (digital) path (or curve) from pixel p with coordinates (x, y) to pixel q with coordinates (s, t) is a sequence of distinct pixels with coordinates

Subscribe EC Academy on YouTube

$$(x_0, y_0), (x_1, y_1), \ldots (x_n, y_n)$$

where $(x_0, y_0) = (x, y)$, $(x_n, y_n) = (s, t)$ & pixels (x_i, y_i) & (x_{i-1}, y_{i-1}) are adjacent for $1 \le i \le n$. In this case, n is the length of the path. If $(x_0, y_0) = (x_n, y_n)$, the path is a closed path.



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

1. Let 'S' represent a subset of pixels in an image. Two pixels p & 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 <u>connected component</u> of S. If it only has one connected component, then set S is called a <u>connected set</u>.

Let 'R' be a subset of pixels in an image. We call R a region of the image if R is a connected set. 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 (which we recall is a rectangular set of pixels), then its

Subscribe EC Academy on YouTube

boundary is defined as the set of pixels in the first & last rows and columns of the image. This extra definition is required because an image has no neighbors beyond its border.

Normally, when we refer to a region, we are referring to a subset of an image, & 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.

Edges are formed from pixels with derivative values that exceed a preset threshold. Edge is based on a measure of gray-level discontinuity at a point.

It is possible to link edge points into edge segments, & sometimes these segments are linked in such a way that correspond to boundaries, but this is not always the case.

The one exception in which edges & boundaries correspond is in binary images. Depending on the type of connectivity & edge operators used, the edge extracted from a binary region will be same as the region boundary. This is intuitive. Conceptually, it is helpful to think of edges as intensity discontinuities & boundaries as closed paths.

Subscribe EC Academy on YouTube

1.9.3 Distance Measures:—

* For pixels p, q & z with coordinates (x,y), (s,t) & (v,w) respectively, D is a distance function or metric if

(a) $D(p,q) \geq 0$ $[D(p,q) = 0$ iff $p = q]$

(b) $D(p,q) = D(q,p)$ &

(c) $D(p,z) \leq D(p,q) + D(q,z)$.

* The Euclidean distance between p & q is defined as,

$$D_e(p,q) = [(x-s)^2 + (y-t)^2]^{1/2} \quad \text{—— (1)}$$

* For this distance measure, the pixels having a distance less than or equal to some value r from (x,y) are the points contained in a disk of radius r centered at (x,y).

* The D_4 distance (also called city-block distance) (or) Manhattan distance between p & q is defined as

$$D_4(p,q) = |x-s| + |y-t| \quad \text{—— (2)}$$

The pixels having a D_4 distance from (x,y) less than or equal to some value r form a diamond centred at (x,y).

For example, the pixels with D_4 distance ≤ 2 from (x,y) [the center point] form the following contours of constant distance:

Subscribe EC Academy on YouTube

$$\begin{array}{ccccc} & & 2 & & \\ & 2 & 1 & 2 & \\ 2 & 1 & 0 & 1 & 2 \\ & 2 & 1 & 2 & \\ & & 2 & & \end{array}$$

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

* The D_8 distance (also called chessboard distance) between p & q is defined as

$$D_8(p, q) = \max(|x-s|, |y-t|) \quad \text{— (3)}$$

The pixels with D_8 distance from (x, y) less than or equal to some value r form a square centered at (x, y). For example, the pixels with D_8 distance ≤ 2 from (x, y) [center point] form the following contours of constant distance:

$$\begin{array}{ccccc} 2 & 2 & 2 & 2 & 2 \\ 2 & 1 & 1 & 1 & 2 \\ 2 & 1 & 0 & 1 & 2 \\ 2 & 1 & 1 & 1 & 2 \\ 2 & 2 & 2 & 2 & 2 \end{array}$$

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

Note that D_4 & D_8 distances between p & q are independent of any paths that might exist between the points because these distances involve only the coordinates of the points.

Subscribe EC Academy on YouTube

If we elect to consider m-adjacency, however the D_m distance between two points is defined as the shortest m-path between the points. In this case, the distance between two pixels will depend on the values of the pixels along the path, as well as the values of their neighbors.

For instance, consider the following arrangements of pixels & assume that P, P_2 & P_4 have value 1 & that P_1 & P_3 can have a value of 0 or 1:

1.

$$\begin{array}{ccc} & P_3 & P_4 \\ P_1 & P_2 & \\ P & & \end{array}$$

Suppose that we consider adjacency of pixels valued 1 (i.e., $V = \{1\}$). If P_1 & P_3 are 0, the length of the shortest m-path (the D_m distance) between P & P_4 is 2.

If P_1 is 1, then P_2 & P will no longer be m-adjacent & the length of the shortest m-path becomes 3 (the path goes thru the points $PP_1P_2P_4$). Similar comments apply if P_3 is 1 (& P_1 is 0); in this case, the length of shortest m-path also is 3. Finally, if both P_1 & P_3 are 1, the length of the shortest m-path between P & P_4 is 4. In this case, the path goes through the sequence of points $PP_1P_2P_3P_4$.

Subscribe EC Academy on YouTube

1.9.4 Image operations on a Pixel Basis:-

Several methods are defined to perform the operations between images such as dividing one image by another. W.K.T images were represented in the form of matrices.

When we refer to an operation like "dividing one image by another", we mean specifically that the division is carried out between corresponding pixels in the two images.

For example, if f and g are images, the first element of the image formed by "dividing" f by g is simply the first pixel in f divided by the first pixel in g; the assumption is that none of the pixels in g have value 0. Other arithmetic & logic operations are similarly defined between corresponding pixels in the images involved.

1.10 Linear and Nonlinear operations:-

Let 'H' be an operator whose i/p & o/p are images. 'H' is said to be a linear operator, if for any two images f & g and any two scalars a and b,

$$H(af+bg) = aH(f) + bH(g) \quad \text{— (1)}$$

i.e., the result of applying a linear operator to the sum of two images is identical to applying the operator to the images individually,

Subscribe EC Academy on YouTube

multiplying the results by the appropriate constants & then adding those results.

32 For ex:- an operator whose function is to compute the sum of K images is a linear operator. An operator that computes the absolute value of the difference of two images is not. An operator that fails the test of above equation is by definition nonlinear.

Linear operations are exceptionally important in image processing because they are based on a significant body of well-understood theoretical & practical results.

Although nonlinear operations sometimes offer better performance, they are not always predictable & for the most part are not well understood theoretically.

Problems:-

(1) Consider the image segment shown below. Compute the length of the shortest 4, 8 & m-paths between pixels P & Q where $V = \{1, 2\}$.

	4	2	3	2	(q)
	3	3	1	3	
	2	3	2	2	
(P)	2	1	2	3	

Subscribe EC academy on YouTube

Solr:-

$$\begin{array}{lcccc} & 4 & 2 & 3 & 2\ (q) \\ & 3 & 3 & 1 & 3 \\ & & & \uparrow & \\ & 2 & 3 & 2 & 2 \\ & & & \uparrow & \\ (p) & 2 & \rightarrow 1 \rightarrow & 2 & 3 \end{array}$$

∴ A shortest 4-path (D_4) does not exist because the pixel '1' in the second row doesnot have the neighbouring pixel {1, 2}.

$$\begin{array}{lcccc} & 4 & 2 & 3 & 2\ (q) \\ & 3 & 3 & 1 & 3 \\ & & & \uparrow & \\ & 2 & 3 & 2 & 2 \\ (p) & 2 & \rightarrow 1 & 2 & 3 \end{array}$$

∴ The shortest 8-path (D_8) is of length 4. i.e., $D_8 = 4$.

$$\begin{array}{lcccc} & 4 & 2 & 3 & 2\ (q) \\ & 3 & 3 & 1 & 3 \\ & & & \uparrow & \\ & 2 & 3 & 2 & 2 \\ & & & \uparrow & \\ (p) & 2 & \rightarrow 1 \rightarrow & 2 & 3 \end{array}$$

∴ The shortest m-path (D_m) is of length 5 i.e., $D_m = 5$

$N_4(p) \cap N_4(q)$

$\{2, 1\} \cap \{3, 3\}$

= empty

Subscribe EC Academy on YouTube

(2) Consider the image segment shown:

i) Let $V = \{0, 1\}$ & compute the lengths of shortest 4-, 8- & m-path b/w p & q. If a particular path does not exist b/w these two points explain why.

ii) Repeat for $V = \{1, 2\}$

	3	1	2	1 (q)
	2	2	0	2
	1	2	1	1
(p)	1	0	1	2

1.

Soln:- i)

	3	1	2	1 (q)
	2	2	0	2
	1	2	1	1
(p)	1	0	1	2

$\therefore D_4$ path does not exist.

	3	1	2	1 (q)
	2	2	0	2
	1	2	1	1
(p)	1	0	1	2

$\therefore D_8 = 4$

	3	1	2	1 (q)
	2	2	0	2
	1	2	1	1
(p)	1	0	1	2

$\therefore D_m = 5$

ii)

	3	1	2	1 (q)
	2	2	0	2
	1	2	1	1
(p)	1	0	1	2

$\therefore D_4 = 6$

	3	1	2	1
	2	2	0	2
	1	2	1	1
	1	0	1	2

$\therefore D_8 = 4$

	3	1	2	1
	2	2	0	2
	1	2	1	1
	1	0	1	2

$\therefore D_m = 6$

Subscribe EC Academy on YouTube

③ For $V = \{2, 3, 4\}$, compute the lengths of shortest 4, 8 & m-path b/w p & q in the following image.

$$\begin{array}{cccccc} & 3 & 4 & 1 & 2 & 0 \\ & 0 & 1 & 0 & 4 & 2 \ (q) \\ & 2 & 2 & 3 & 1 & 4 \\ (p) & 3 & 0 & 4 & 2 & 1 \\ & 1 & 2 & 0 & 3 & 4 \end{array}$$

Soln:-

$$\begin{array}{cccccc} & 3 & 4 & 1 & 2 & 0 \\ & 0 & 1 & 0 & 4 & 2 \ (q) \\ & 2 \rightarrow & 2 \rightarrow & 3 & 1 & 4 \\ & \uparrow & & & & \\ (p) & 3 & 0 & 4 & 2 & 1 \\ & 1 & 2 & 0 & 3 & 4 \end{array}$$

$\therefore$ Shortest 4-path doesnot exist b/w p & q because pixel '3' in the third row has no neighbours with values $\{2, 3, 4\}$.

$\therefore D_4 =$ does not exist

$$\begin{array}{cccccc} & 3 & 4 & 1 & 2 & 0 \\ & 0 & 1 & 0 & 4 \rightarrow & 2 \ (q) \\ & 2 & 2 \rightarrow & 3 \nearrow & 1 & 4 \\ (p) & 3 \nearrow & 0 & 4 & 2 & 1 \\ & 1 & 2 & 0 & 3 & 4 \end{array}$$

$\therefore$ Shortest 8-path is of length 4.

i.e., $D_8 = 4$

$$\begin{array}{cccccc} & 3 & 4 & 1 & 2 & 0 \\ & 0 & 1 & 0 & 4 \rightarrow & 2 \ (q) \\ & 2 \rightarrow & 2 \rightarrow & 3 \nearrow & 1 & 4 \\ & \uparrow & & & & \\ (p) & 3 & 0 & 4 & 2 & 1 \\ & 1 & 2 & 0 & 3 & 4 \end{array}$$

$\therefore$ Shortest m-path is of length 5.

i.e., $D_m = 5$

Subscribe EC Academy on YouTube

(4) Let the set of gray levels used to define connectivity be {94, 95, 96, 97} & compute the shortest D_4 & D_8 distances b/w pixels p & q for the image segment shown in fig. Indicate the shortest path by double lines.

(p)	96	97	94	97
	98	98	100	96
	99	97	98	95
(q)	97	96	97	96

Soln:-

1.

(p) 96 → 97 → 94 → 97
↓
98 98 100 96
↓
99 97 98 95
↓
(q) 97 ← 96 ← 97 ← 96

$\therefore$ $D_4 = 9$

(p) 96 → 97 → 94 97
↘
98 98 100 96
↓
99 97 98 95
↙
(q) 97 ← 96 ← 97 96

$\therefore$ $D_8 = 7$

(5) Let p & q are the two pixels at coordinates (100, 120) & (130, 160) respectively. Compute i) Euclidean distance ii) Chess board distance iii) Manhattan distance.

Soln:- Given:- $P(\overset{x}{100}, \overset{y}{120})$ & $q(\overset{s}{130}, \overset{t}{160})$

i) $D_e(p,q) = [(x-s)^2 + (y-t)^2]^{1/2}$

$= \sqrt{[(100-130)^2 + (120-160)^2]}$

$= \sqrt{(-30)^2 + (-40)^2} = \sqrt{900+1600} = \sqrt{2500}$

$D_e(p,q) = 50$

Subscribe EC Academy on YouTube

Energy of a photon, $E = h\nu$ where 'h' is planck's constant.

Sampling:- digitizing the 2-dimensional spatial coordinate values.

Quantization:- digitizing the amplitude values (brightness level)

Zooming:- Increasing the no. of pixels in an image so that the image appears larger.

Non-adjacency:-

* q & p are 4-adjacent (or)
* p & q are diagonally adjacent & do not have any common 4-adjacent neighbors.
* they cannot be both (i) & (ii).

Edge: a path of one or more pixels that separate two regions of significantly different gray levels.

	0	
0	X	0
	0	

$N_4(P)$

i.e set of horizontal & vertical neighbors

0		0
	X	
0		0

$N_D(P)$

set of 4 diagonal neighbors.

0	0	0
0	X	0
0	0	0

$N_8(P)$

union of 4-neighbors & diagonal neighbors.

Subscribe EC Academy on YouTube

Adjacency:- If the pixels are connected which belongs to a set 'V' then it is said to be 4-adjacent, 8-adjacent & m-adjacent (or) Two pixels that are neighbor & have the same grey-level are adjacent.

11

Path:- defined has sequence of distinct-pixels.

Image Acquisition:- Capturing of the image

Spatial resolution:- The smallest discernible changes in the gray-level of the image.

Neighbours:-

Connectivity of pixels:- Two pixels are said to be connected if there is a path b/w them that lies completely within subset 's' of an image.

Standard Formats of Storing compressed Image:-

1) PGM (Portable Gray Map)
2) PNG (Portable Network Graphics)
3) GIF (Graphic Interchange Format)
4) JPEG (Joint Photographic Experts group)
5) TIFF (Tagged Image File Format)
6) FITS (Flexible Image Transport System)

Radiance:- The total amount of energy that flow from the light source. (can be measured)

Luminance:- The amount of energy an observer perceives from a light source (can be measured)

Brightness:- a subjective descriptor of light perception

perceived quantity of light emitted (cannot be measured

Subscribe ECE Academy on YouTube

Module-2

2.1 Spatial Domain:-

* Image enhancement approaches fall into two broad categories: i) Spatial domain methods &
  ii) Frequency domain methods

* The term Spatial domain refers to the image plane itse and approaches in this category are based on direct manipulation of pixels in an image.

* Frequency domain processing techniques are based o modifying the Fourier transform of an image.

* The term Spatial domain refers to the aggregate of pixels composing an image. Spatial domain methods are procedures that operate directly on these pixel Spatial domain processes will be denoted by the expression,

$$g(x,y) = T[f(x,y)] \quad -(1)$$

where $f(x,y)$ is the i/p image
$g(x,y)$ is the processed image
T is an operator on f defined over som neighborhood of (x,y).

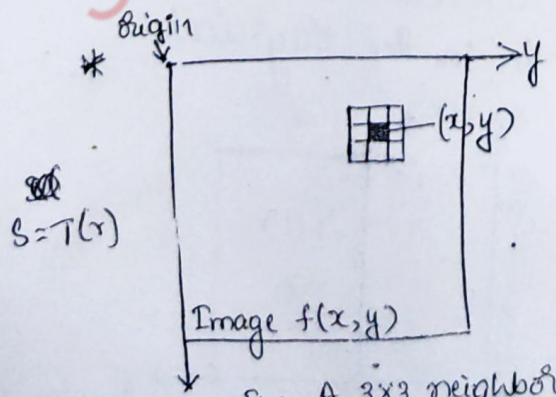


Fig:- A 3x3 neighborhood about a point (x,y) in an image.

* The principal approach in defin a neighborhood about a point (x, is to use a square or rectangul subimage area centered at (x as fig., shows.

The center of the subimage is moved from

Subscribe EC Academy on YouTube

pixel to pixel starting, say, at the top of left corner. The operator T is applied at each location (x,y) to yield the output, g, at that location. The process utilizes only the pixels in the area of the image spanned by the neighborhood.

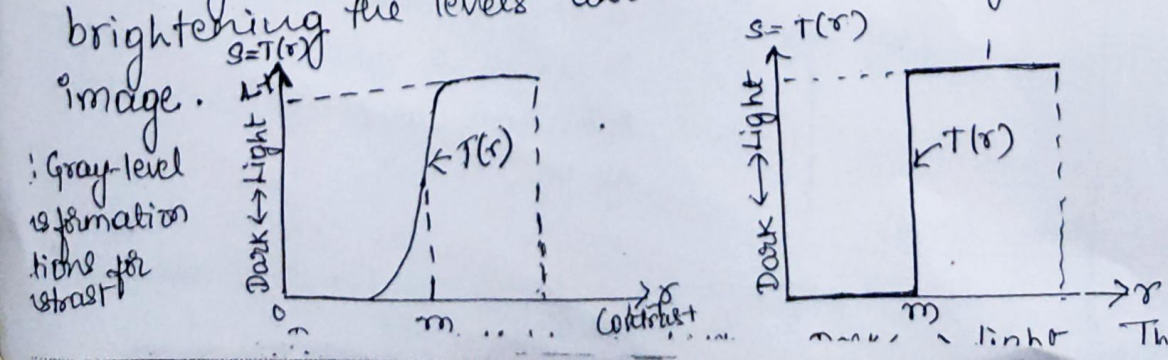
Although other neighborhood shapes, such as approximations to a circle sometimes are used, square & rectangular arrays are by far the most predominant because of their ease of implementation.

The simplest form of T is when the neighborhood is of size 1×1 (i.e., a single pixel). In this case, g depends only on the value of f at (x,y) & T becomes a gray-level (also called intensity or mapping) transformation function of the form,

$$s = T(r) \quad \text{—— (9)}$$

where r & s are variables denoting the gray-level of $f(x,y)$ & $g(x,y)$ at at any point (x,y).

For example, if T(r) has the form shown in fig(a), the effect of this transformation would be to produce an image of higher contrast than the original by darkening the levels below m & brightening the levels above m in the original image.



: Gray-level transformation functions for contrast

Subscribe EC Academy on YouTube

In this technique, known as contrast stretching, the values of r below m are compressed by the transformation function into a narrow range of s, toward black.

The opposite effect takes place for values of r above m. In the limiting case shown in fig(b), T(r) produces a two-level (binary) image. A mapping of this form is called a thresholding function.

Enhancement at any point in an image depends only on the gray level at that point, techniques in this category often are referred to as point processing.

* Larger neighborhoods allow considerably more flexibility. The general approach is to use a function of the values of f in a predefined neighborhood of (x,y) to determine the value of g at (x,y).

One of the principal approaches in this formulation is based on the use of so-called masks (also referred to as filters, kernels, templates or windows).

Consider a small mask (3x3), 2-D array in which the values of the mask coefficients determine the nature of the process, such as image sharpening. Enhancement techniques based on this type of approach often are referred to as mask processing or filtering.

2.2 Some Basic Gray Level Transformations:-

Gray-level transformation functions are the simplest technique among all image enhancement techniques.

Subscribe EC Academy on YouTube

The values of pixels, before and after processing, will be denoted by r & s, respectively. These values are related by an expression of the form $s = T(r)$, where T is a transformation that maps a pixel value r into a pixel value s.

Since we are dealing with digital quantities, values of the transformation function typically are stored in a one-dimensional array and the mappings from r to s are implemented via table lookups. For an 8-bit environment, a lookup table containing the value of T will have 256 entries.

The three basic types of functions used frequently for image enhancement are:

i) linear (negative & identity transformations)
ii) logarithmic (log & inverse-log transformations) &
iii) power-law (n^{th} power & n^{th} root transformations).

The identity function is the trivial case in which output intensities are identical to input intensities. It is included in the graph only for completeness.

i) <u>Image Negatives:</u>—

The negative of an image with gray levels in the range $[0, L-1]$ is obtained by using the negative transformation shown in fig, which is given by the expression, $\boxed{s = L - 1 - r}$ —— (1)

Subscribe EC Academy on YouTube

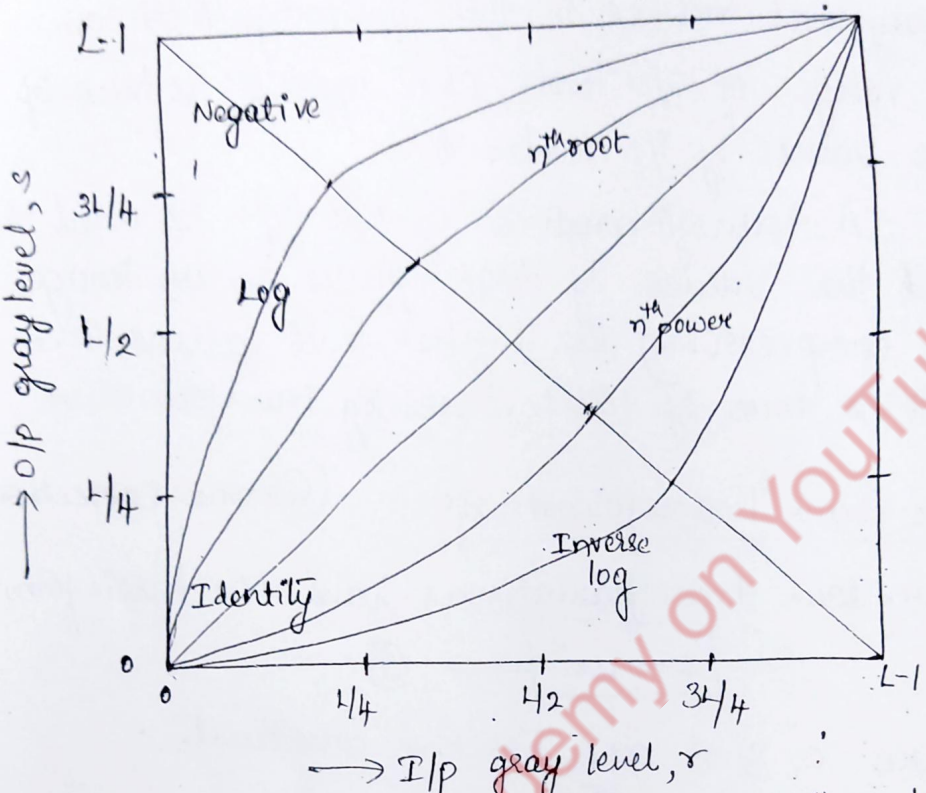


Fig:- Some basic gray-level transformation functions used for image enhancement.

Reversing the intensity levels of an image in this manner produces the equivalent of a photographic negative. This type of processing is particularly suitable for enhancing white or gray detail embedded in dark regions of an image, especially when the black areas are dominant in size.

[Refer to example fig 3.4 in text book pg. no: 79]

ii) Log Transformations :- The general form of the log transformation shown in the above fig. is

$$\boxed{s = c \log(1+r)} \quad \text{——} \; (2)$$

where c is a constant & it is assumed that $r \geq 0$.

Subscribe EC Academy on YouTube

The shape of the log curve in fig. shows that this transformation maps a narrow range of low gray-level values in the i/p image into a wider range of o/p levels. The opposite is true of higher values of i/p levels.

A transformation of this type is used to expand the values of dark pixels in an image while compressing the higher-level values. The opposite is true of the inverse log transformation.

3) Power-Law Transformations :— (Gamma Correction)

Power-law transformations have the basic form

$$\boxed{s = cr^{\gamma}} \quad \text{—— (3)}$$

where c & γ are positive constants.

sometimes eqn (3) is written as,

$s = c(r+\varepsilon)^{\gamma}$, to account for an offset i.e., a measurable o/p when i/p is zero.

However, offsets typically are an issue of display calibration & as a result they are normally ignored in eqn (3). Plots of s versus r for various values of γ are shown in fig.

As in the case of log transformation, power-law curves with fractional values of γ map a narrow range of dark i/p values into a wider range of o/p values, with the opposite being true for higher values of i/p levels.

Subscribe EC Academy on YouTube

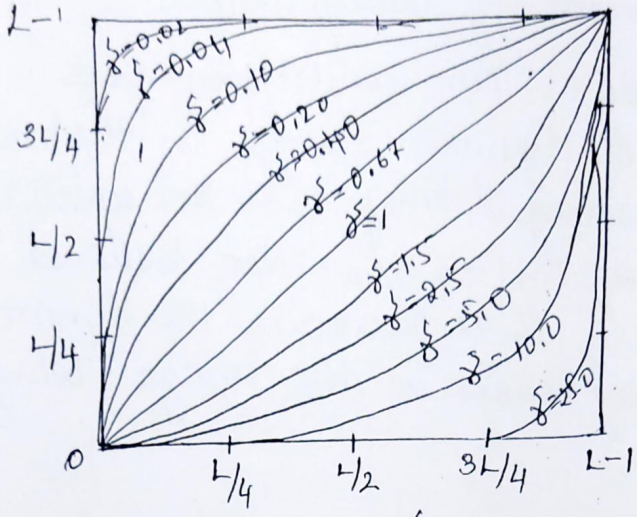


Fig:- Plots of eqn $s = cr^{\gamma}$ for various values of γ (c=1 in all cases)

The curves generated with values of $\gamma > 1$ have exactly the opposite effect as those generated with values of $\gamma < 1$. Finally, we note that eqn (3) reduces to the identity transformation when $c = \gamma = 1$.

A variety of devices used for image capture, printing & display respond according to a power law. By convention, the exponent in the power-law equn is referred to as gamma. The process used to correct this power-law response phenomena is called gamma correction.

iv) Piecewise-Linear Transformation Functions:—

Last three methods discussed uses piecewise linear functions. The principal advantage of piecewise linear functions over the types of functions is that it can be arbitrarily complex. The principal disadvantage of piecewise functions is that their specification requires considerably more user input.

Subscribe EC Academy on YouTube

Contrast stretching:-

* One of the simplest piecewise linear functions is a contrast-stretching transformation.

* Low-contrast images can result from poor illumination, lack of dynamic range in the image sensor or even wrong setting of a lens aperture during image acquisition. The idea behind contrast stretching is to increase the dynamic range of the gray levels in the image being processed.

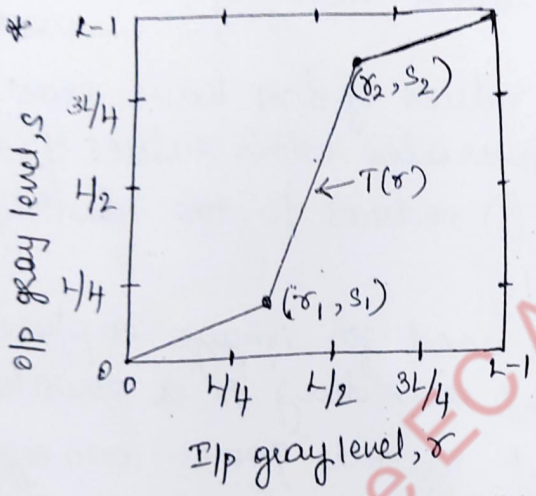


Fig:- form of transformation function.

Fig. shows a typical transformation used for contrast stretching. The locations of points (r_1, s_1) and (r_2, s_2) control the shape of the transformation function.

If $r_1 = s_1$ & $r_2 = s_2$, the transformation is a linear function that produces no changes in gray levels.

If $r_1 = r_2$, $s_1 = 0$ & $s_2 = L-1$, the transformation becomes a thresholding function that creates a binary image. Intermediate values of (r_1, s_1) & (r_2, s_2) produce various degrees of spread in the gray levels of the o/p image, thus affecting its contrast.

In general, $r_1 \le r_2$ & $s_1 \le s_2$ is assumed so that the function is single valued and monotonically increasing. This condition preserves the order of

Subscribe EC Academy on YouTube

gray levels, thus preventing the creation of intensity artifacts in the processed image.

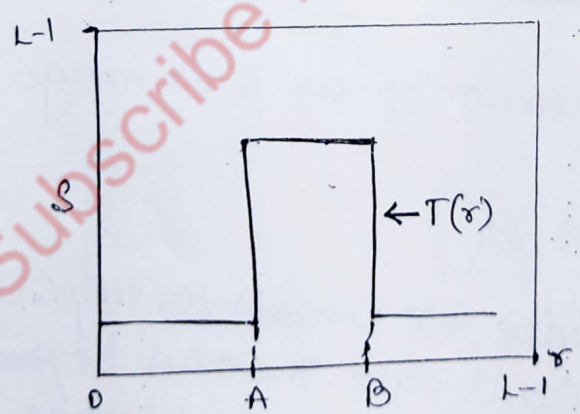
[Ex:- refer fig. 3.10 in text]

Gray-level Slicing:

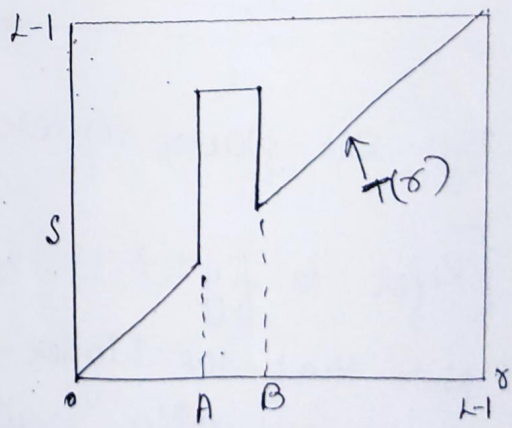
* Highlighting a specific range of graylevels in an im is often desired.

* Applications include enhancing features such as masses of water in satellite imagery & enhancin flaws in x-ray images.

* There are several ways of doing level slicing, bu most of them are variations of two basic theme
  -→ One approach is to display a high value f all graylevels in the range of interest and a low value for all other gray levels. This transfor -tion shown in fig (a), produces a binary image.



fig(a): This transformation highlights range [A,B] of gray levels & reduces all others to a constant level.



fig(b): This transformatio highlights range [A,B] preserves all other level

Subscribe ECAcademy on YouTube

Bit plane slicing:

* Instead of highlighting gray-level ranges, highlighting the contribution made to total image appearance by specific bits might be desired.

* Suppose that each pixel in an image is represented by 8-bits. Imagine that the image is composed of eight 1-bit planes, ranging from bit-plane 0 for the LSB & bit-plane 7 for the MSB. Fig. below illustrates these ideas.

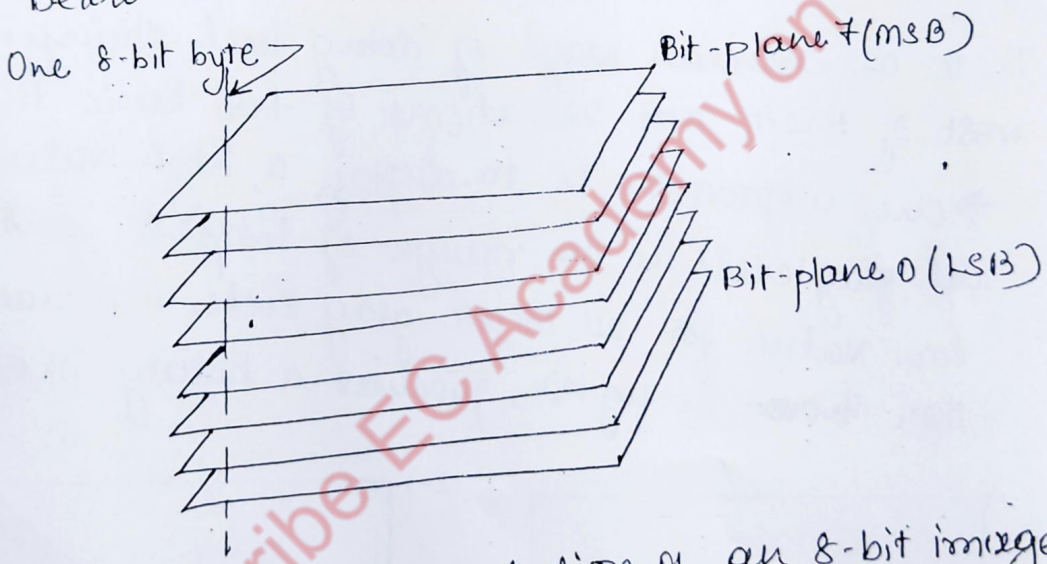


Fig:- Bit-plane representation of an 8-bit image.

[Refer to fig: 3.13 & 3.14]

Note that the higher-order bits contain the majority of the visually significant data. The other bit planes contribute to more subtle details in the image.

Separating a digital image into its bit planes is useful for analyzing the importance played by each bit of the image, which intern helps in determining

Subscribe EC Academy on YouTube

the adequacy of no. of bits used to quantize each pix
Also, this type of decomposition is useful for image compression.

In terms of bit-plane extraction for an 8-image, the image for bit-plane 7 can be obtained b processing the i/p image with a thresholding gray level transformation function that

(1) maps all levels in the image b/w 0 & 127 to one
& (2) maps all levels b/w 129 & 255 to another.

2.3 Histogram Processing:-

* The histogram of a digital image with grayle in the range $[0, L-1]$ is a discrete function $\boxed{h(r_k) = n_k}$, where r_k is the kth gray level & n_k is the no. of pixels in the image having graylevel r_k

* A histogram can be normalized by dividing ea of its values by the total no. of pixels in the image, denoted by n. Thus, a normalized histo is given by $\boxed{p(r_k) = n_k/n}$, for $k = 0, 1, \ldots, L-1$.

$p(r_k)$ gives an estimate of the probability of occurrence of gray level r_k.

The sum of all components of a normalized histogram is equal to 1.

* Histogram are the basis for numerous spatial domain processing techniques, such as image enhancement, image compression & segmentation Histograms are simple to calculate in Software

Subscribe EC Academy on YouTube

& also lend themselves to economic hardware implementations, thus making them a popular tool for real-time image processing.

These histogram plots are simple plots of $h(r_K) = n_K$, versus r_K or $p(r_K) = n_K/n$ versus r_K.

[Refer to fig 3.15, pg. no: 90 in text book]

i) Histogram Equalization:-

* Let variable 'r' represent the gray levels of the image to be enhanced. Consider the transformations of the form,

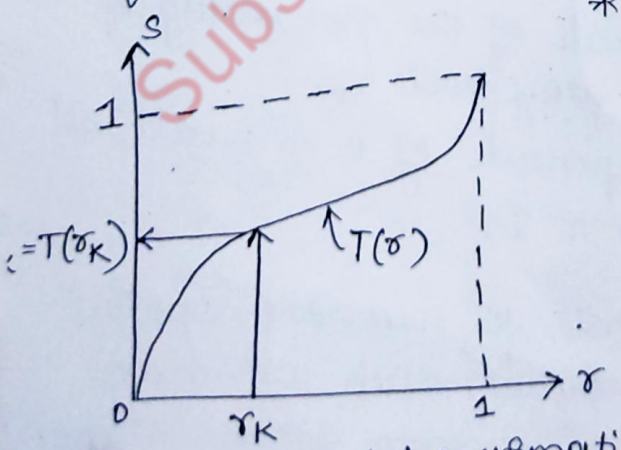
$$\boxed{S = T(r)}, \quad 0 \le r \le 1 \quad \text{--- (1)}$$

* Assume that the transformation function T(r) satisfies the following conditions:

(a) T(r) is single-valued and monotonically increasing in the interval $0 \le r \le 1$; and

(b) $0 \le T(r) \le 1$ for $0 \le r \le 1$.

* Fig. below gives an example of a transformation function that satisfies these two conditions:



* The inverse transformation from s back to r is denoted by

$$\boxed{r = T^{-1}(s)}, \quad 0 \le s \le 1$$

Fig:- A gray-level transformation function i.e both single valued & monotonically increasing.

Subscribe EC Academy on YouTube

* Note that, even if $T(r)$ satisfies conditions (a) & it is possible that the corresponding inverse $T^{-1}(s)$ may fail to be single valued.

* Let $P_r(r)$ and $P_s(s)$ denote the PDF's of random variables r & s, respectively.

If $P_r(r)$ & $T(r)$ are known and $T^{-1}(s)$ satisfies condition (a), then the PDF $P_s(s)$ of the transformed variable 's' can be obtained using a simple formula

$$P_s(s) = P_r(r)\left|\frac{dr}{ds}\right| \quad \text{--- (3)}$$

* A transformation function of particular importance in image processing has the form,

$$s = T(r) = \int_0^r P_r(w)\,dw \quad \text{--- (4)}$$

where 'w' is a dummy variable of integration.

The right side of eqn (4) is recognized as the cumulative distribution function (CDF) of random variable r.

$$\frac{ds}{dr} = \frac{dT(r)}{dr}$$

$$= \frac{d}{dr}\left[\int_0^r P_r(w)\,dw\right] \quad \text{--- (5)}$$

$$\frac{ds}{dr} = P_r(r).$$

Substituting this result for $\frac{dr}{ds}$ into eqn (3), & assuming all probability values are positive, yields

$$P_s(s) = P_r(r)\left|\frac{dr}{ds}\right| = P_r(r)\left|\frac{1}{P_r(r)}\right| \quad \text{--- (6)}$$

$$P_s(s) = 1 \quad , \quad 0 \le s \le 1$$

Subscribe EC Academy on YouTube

From eqn (4), $T(r)$ depends on $p_r(r)$ but by eqn (6), the resulting $p_s(s)$ always is uniform, independent of the form of $p_r(r)$.

The probability of occurrence of graylevel r_K in an image is approximated by,

$$P_r(r_K) = \frac{n_K}{n} \quad K = 0, 1, 2, \ldots, L-1 \quad \text{—— (7)}$$

where n is the total no. of pixels in the image
n_K is the no. of pixels that have graylevel r_K.

L is the total no. of possible graylevels in the image.

The discrete version of the transformation function given in eqn (4) is,

$$S_K = T(r_K) = \sum_{j=0}^{K} P_r(r_j) \quad \text{—— (8)}$$

$$= \sum_{j=0}^{K} \frac{n_j}{n} \quad K = 0, 1, 2, \ldots, L-1$$

Thus, a processed (o/p) image is obtained by mapping each pixel with level r_K in the i/p image into a corresponding pixel with level S_K in the o/p image via eqn (8). The transformation (mapping) given in eqn (8) is called histogram equalization or histogram linearization. The transformation in eqn (8), satisfies conditions (a) & (b).

The inverse transformation from s back to r is denoted by, $r_K = T^{-1}(S_K)$ $K = 0, 1, 2, \ldots, L-1$

—— (9)

Subscribe EC Academy on YouTube

Eqn (9) satisfies conditions (a) & (b) only if none of the levels, r_k, $k = 0, 1, 2, \ldots, L-1$, are missing from i/p image.

The inverse transformation is used in histo-matching scheme.

[Refer to fig. 3.17, page no. 95 in textbook]

ii) Histogram Matching (specification):—

Histogram equalization automatically determines a transformation function that seeks to produce an o/p image that has a uniform histogram.

When automatic enhancement is desired, this is a good approach because the results from the technique are predictable & the method is simple to implement.

But there are applications in which attempting to base enhancement on a uniform histogram is not the best approach. In particular it is useful sometimes to be able to specify the shape of the histogram that we wish the process image to have.

The method used to generate a process image that has a specified histogram is called histogram matching or histogram specification.

Development of the method:—

Let r & z be continuous gray levels & let $p_r(r)$ & $p_z(z)$ denote their corresponding PDFs

Subscribe EC Academy on YouTube

Here, r & z denote the gray levels of the i/p & o/p (processed) images, respectively.

Let 's' be a random variable with the property,

$$s = T(r) = \int_0^r p_r(w)\,dw \quad \text{—— (1)}$$

where 'w' is a dummy variable of integration.

We define a random variable (z) with the property,

$$G(z) = \int_0^z p_z(t)\,dt = s \quad \text{—— (2)}$$

where 't' is a dummy variable of integration.

From above eqns (1) & (2), it follows that $G(z) = T(r)$ and therefore, that z must satisfy the condition,

$$z = G^{-1}(s) = G^{-1}[T(r)] \quad \text{—— (3)}$$

The transformation $T(r)$ can be obtained from eqn (1) once $p_r(r)$ has been estimated from the i/p image. Lly, the transformation function $G(z)$ can be obtained using eqn (2) because $p_z(z)$ is given.

Assuming that G^{-1} exists and that it satisfying conditions from eqns (1) thr' (3) show that an image with a specified PDF can be obtained from an i/p image by using the following procedure:

(1) Obtain the transformation function $T(r)$ using eqn (1)

Subscribe EC Academy on YouTube

(2) Use eqn (2) to obtain the transformation funcn $G(z)$.

(3) Obtain the inverse transformation function G^{-1}.

(4) Obtain the o/p image by applying eqn (3) to all the pixels in the i/p image. The result this procedure will be an image whose gray levels z, have the specified PDF $p_z(z)$.

The discrete formulation of eqn (1) is give by,

$$S_K = T(r_K) = \sum_{j=0}^{k} p_r(r_j) \quad \text{—(}$$

$$= \sum_{j=0}^{k} \frac{n_j}{n} \quad K = 0, 1, 2, \ldots, L-1$$

where 'n' is the total no. of pixels in the image,
'n_j' is the no. of pixels with graylevel r_j
'L' is the no. of discrete graylevels.

III^{ly}, the discrete formulation of eqn (2) is obtained given histogram $p_z(z_i)$; $i = 0, 1, 2, \ldots, L-1$, & has the form,

$$v_K = G(z_K) = \sum_{i=0}^{k} p_z(z_i) = S_K \quad K = 0, 1, 2, \ldots, L-1$$

Finally, the discrete version of eqn (3) is given by

$$Z_K = G^{-1}[T(r_K)] \quad K = 0, 1, 2, \ldots, L-1 \quad \text{— (6)}$$

& from eqn (4),

$$Z_K = G^{-1}(S_K) \quad K = 0, 1, 2, \ldots, L-1 \quad \text{— (7)}$$

Eqns (6) & (7) are the foundation for implementing histogram processing in digital images.

Subscribe EC Academy on YouTube

Implementation :-

Assume the following:

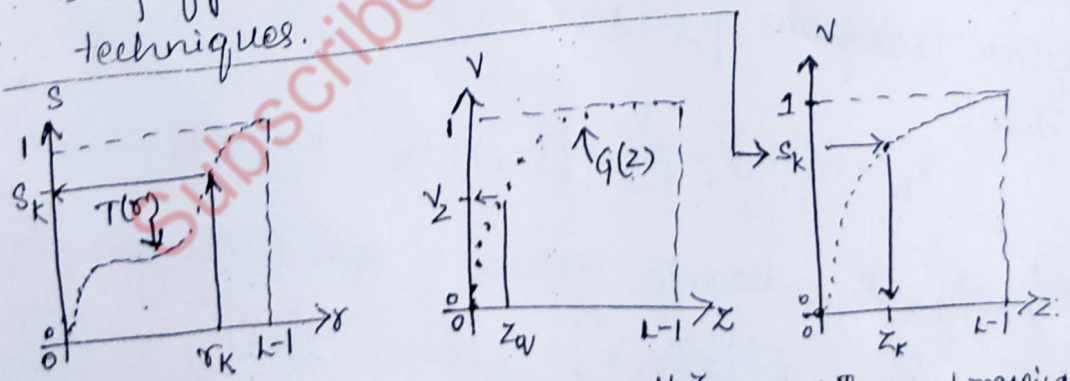
28

(1) Each set of gray levels $\{r_j\}$, $\{s_j\}$ & $\{z_j\}$, $j = 0, 1, \ldots, L-1$ is a one-dimensional array of dimension $L \times 1$.

(2) All mappings from r to s and from s to z are simple table lookups b/w a given pixel value & these arrays.

(3) Each of the elements of these arrays, for Ex, s_k contains two important pieces of information. The subscript k denotes the location of the element in the array & s denotes the value at that location

(4) We need to be concerned only with integer pixel values. For example, in case of an 8-bit image, $L = 256$ & the elements of each of the arrays are integers b/w 0 & 255. This implies that gray-level values lies in the interval $[0, L-1]$ instead of the normalized interval $[0, 1]$, which simplifies the development of histogram processing techniques.



(a) Graphical Interpretation of mapping from r_k to s_k via $T(r)$.

(b) Mapping of z_q to its corresponding value v_q via $G(z)$.

(c) Inverse mapping from s_k to its corresponding value of z_k.

Subscribe EC Academy on YouTube

The procedure for histogram matching may be summ as follows:

1) Obtain the histogram of the given image.

2) Use eqn, $S_K = T(r_K) = \sum_{j=0}^{K} P_r(r_j)$

$$= \sum_{j=0}^{K} \frac{n_j}{n} \quad K = 0, 1, 2, \ldots, L-1$$

to precompute a mapped level S_K for each level r_K.

3) Obtain the transformation function G from the giv $P_Z(Z)$ using eqn,

$$V_K = G(Z_K) = \sum_{i=0}^{K} P_Z(Z_i) = S_K \quad K = 0, 1, 2, \ldots, L-1$$

4) Precompute Z_K for each value of S_K using the iter scheme defined in connection with eqn,

$$(G(\hat{Z}) - S_K) \geq 0 \quad K = 0, 1, 2, \ldots, L-1$$

5) For each pixel in the original image, if the va of that pixel is r_K, map this value to its corresp level S_K; then map level S_K into the final level Z_K. Use the precomputed values from steps (2) & (4) these mappings.

<u>Local Enhancement:</u>

Subscribe EC Academy on YouTube

Local Enhancement

The histogram processing methods discussed in the previous two sections are *global*, in the sense that pixels are modified by a transformation function based on the gray-level content of an entire image. Although this global approach is suitable for overall enhancement, there are cases in which it is necessary to enhance details over small areas in an image. The number of pixels in these areas may have negligible influence on the computation of a global transformation whose shape does not necessarily guarantee the desired local enhancement. The solution is to devise transformation functions based on the gray-level distribution—or other properties—in the neighborhood of every pixel in the image. Although processing methods based on neighborhoods are the topic of Section 3.5, we discuss local histogram processing here for the sake of clarity and continuity. The reader will have no difficulty in following the discussion.

The histogram processing techniques previously described are easily adaptable to local enhancement. The procedure is to define a square or rectangular neighborhood and move the center of this area from pixel to pixel. At each location, the histogram of the points in the neighborhood is computed and either a histogram equalization or histogram specification transformation function is obtained. This function is finally used to map the gray level of the pixel centered in the neighborhood. The center of the neighborhood region is then moved to an adjacent pixel location and the procedure is repeated. Since only one new row or column of the neighborhood changes during a pixel-to-pixel translation of the region, updating the histogram obtained in the previous location with the new data introduced at each motion step is possible (Problem 3.11). This approach has obvious advantages over repeatedly computing the histogram over all pixels in the neighborhood region each time the region is moved one pixel location. Another approach used some times to reduce computation is to utilize nonoverlapping regions, but this method usually produces an undesirable checkerboard effect.

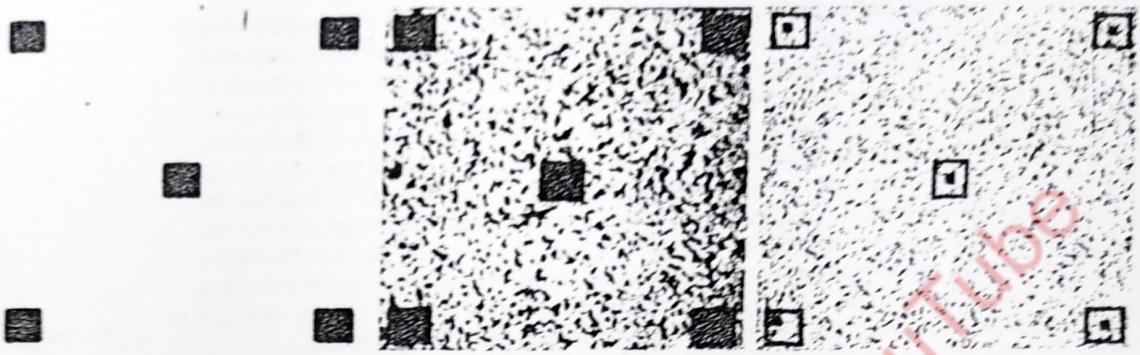
EXAMPLE 3.5: Enhancement using local histograms.

Figure 3.23(a) shows an image that has been slightly blurred to reduce its noise content (see Section 3.6.1 regarding blurring). Figure 3.23(b) shows the result of global histogram equalization. As is often the case when this technique is applied to smooth, noisy areas, Fig. 3.23(b) shows considerable enhancement of the noise, with a slight increase in contrast. Note that no new structural details were brought out by this method. However, local histogram equalization using a 7×7 neighborhood revealed the presence of small squares inside the larger dark squares. The small squares were too close in gray level to the larger ones, and their sizes were too small to influence global histogram equalization significantly. Note also the finer noise texture in Fig. 3.23(c), a result of local processing using relatively small neighborhoods.

Use of Histogram Statistics for Image Enhancement

Instead of using the image histogram directly for enhancement, we can use instead some statistical parameters obtainable directly from the histogram. Let r denote a discrete random variable representing discrete gray-levels in the range

Subscribe EC Academy on YouTube



a b c

FIGURE 3.23 (a) Original image. (b) Result of global histogram equalization. (c) Result of local histogram equalization using a 7 × 7 neighborhood about each pixel.

$[0, L-1]$, and let $p(r_i)$ denote the normalized histogram component corresponding to the ith value of r. As indicated previously in this section, we may view $p(r_i)$ as an estimate of the probability of occurrence of gray level r_i. The nth moment of r about its mean is defined as

$$\mu_n(r) = \sum_{i=0}^{L-1} (r_i - m)^n p(r_i) \tag{3.3-18}$$

where m is the mean value of r (its average gray level):

$$m = \sum_{i=0}^{L-1} r_i p(r_i). \tag{3.3-19}$$

It follows from Eqs. (3.3-18) and (3.3-19) that $\mu_0 = 1$ and $\mu_1 = 0$. The second moment is given by

$$\mu_2(r) = \sum_{i=0}^{L-1} (r_i - m)^2 p(r_i). \tag{3.3-20}$$

We recognize this expression as the variance of r, which is denoted conventionally by $\sigma^2(r)$. The standard deviation is defined simply as the square root of the variance. We will revisit moments in Chapter 11 in connection with image description. In terms of enhancement, however, we are interested primarily in the mean, which is a measure of average gray level in an image, and the variance (or standard deviation), which is a measure of average contrast.

We consider two uses of the mean and variance for enhancement purposes. The *global* mean and variance are measured over an entire image and are useful primarily for gross adjustments of overall intensity and contrast. A much more powerful use of these two measures is in local enhancement, where the *local* mean and variance are used as the basis for making changes that depend on image characteristics in a predefined region about each pixel in the image.

Subscribe EC Academy on YouTube

Let (x, y) be the coordinates of a pixel in an image, and let S_{xy} denote a neighborhood (subimage) of specified size, centered at (x, y). From Eq. (3.3-19) the mean value $m_{S_{xy}}$ of the pixels in S_{xy} can be computed using the expression

$$m_{S_{xy}} = \sum_{(s,t)\in S_{xy}} r_{s,t}\, p(r_{s,t}) \tag{3.3-21}$$

where $r_{s,t}$ is the gray level at coordinates (s, t) in the neighborhood, and $p(r_{s,t})$ is the neighborhood normalized histogram component corresponding to that value of gray level. Similarly, from Eq. (3.3-20), the gray-level variance of the pixels in region S_{xy} is given by

$$\sigma^2_{S_{xy}} = \sum_{(s,t)\in S_{xy}} [r_{s,t} - m_{S_{xy}}]^2 p(r_{s,t}). \tag{3.3-22}$$

The local mean is a measure of average gray level in neighborhood S_{xy}, and the variance (or standard deviation) is a measure of contrast in that neighborhood.

An important aspect of image processing using the local mean and variance is the flexibility they afford in developing simple, yet powerful enhancement techniques based on statistical measures that have a close, predictable correspondence with image appearance. We illustrate these characteristics by means of an example.

EXAMPLE 3.6: Enhancement based on local statistics.

Figure 3.24 shows an SEM (scanning electron microscope) image of a tungsten filament wrapped around a support. The filament in the center of the image and its support are quite clear and easy to study. There is another filament structure on the right side of the image, but it is much darker and its size and other features are not as easily discernable. Local enhancement by contrast manipulation is an ideal approach to try on problems such as this, where part of the image is acceptable, but other parts may contain hidden features of interest.

In this particular case, the problem is to enhance dark areas while leaving the light area as unchanged as possible since it does note require enhancement. We can use the concepts presented in this section to formulate an enhancement method that can tell the difference between dark and light and, at the same time, is capable of enhancing only the dark areas. A measure of whether an area is relatively light or dark at a point (x, y) is to compare the local average gray level $m_{S_{xy}}$ to the average image gray level, called the global mean and denoted M_G. This latter quantity is obtained by letting S encompass the entire image. Thus, we have the first element of our enhancement scheme: We will consider the pixel at a point (x, y) as a candidate for processing if $m_{S_{xy}} \leq k_0 M_G$, where k_0 is a positive constant with value less than 1.0. Since we are interested in enhancing areas that have low contrast, we also need a measure to determine whether the contrast of an area makes it a candidate for enhancement. Thus, we will consider the pixel at a point (x, y) as a candidate for enhancement if $\sigma_{S_{xy}} \leq k_2 D_G$, where D_G is the global standard deviation and k_2 is a positive constant. The value of this constant will be greater than 1.0 if we are interested in enhancing light areas and less than 1.0 for dark areas. Finally, we need to restrict

Subscribe ECAcademy on YouTube

the lowest values of contrast we are willing to accept, otherwise the procedure would attempt to enhance even constant areas, whose standard deviation is zero. Thus, we also set a lower limit on the local standard deviation by requiring that $k_1 D_G \leq \sigma_{S_{xy}}$, with $k_1 < k_2$. A pixel at (x, y) that meets all the conditions for local enhancement is processed simply by multiplying it by a specified constant, E, to increase (or decrease) the value of its gray level relative to the rest of the image. The values of pixels that do not meet the enhancement conditions are left unchanged.

A summary of the enhancement method is as follows. Let $f(x, y)$ represent the value of an image pixel at any image coordinates (x, y), and let $g(x, y)$ represent the corresponding enhanced pixel at those coordinates. Then

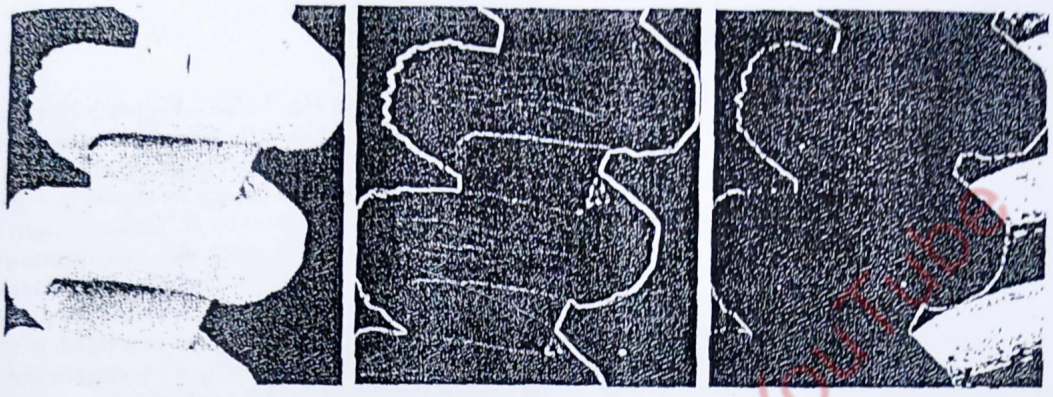
$$g(x, y) = \begin{cases} E \cdot f(x, y) & \text{if } m_{S_{xy}} \leq k_0 M_G \quad \text{AND} \quad k_1 D_G \leq \sigma_{S_{xy}} \leq k_2 D_G \\ f(x, y) & \text{otherwise} \end{cases}$$

where, as indicated previously, E, k_0, k_1, and k_2 are specified parameters; M_G is the global mean of the input image; and D_G is its global standard deviation.

Normally, making a successful selection of parameters requires a bit of experimentation to gain familiarity with a given image or class of images. In this case, the following values were selected: $E = 4.0$, $k_0 = 0.4$, $k_1 = 0.02$, and $k_2 = 0.4$. The relatively low value of 4.0 for E was chosen so that, when it was multiplied by the levels in the areas being enhanced (which are dark), the result would still tend toward the dark end of the scale, and thus preserve the general visual balance of the image. The value of k_0 was chosen as somewhat less than half the global mean since it is obvious by looking at the image that the areas that require enhancement definitely are dark enough to be below half the global mean. A similar analysis led to the choice of values for k_1 and k_2. Choosing these constants is not a difficult task in general, but their choice

FIGURE 3.24 SEM image of a tungsten filament and support, magnified approximately 130×. (Original image courtesy of Mr. Michael Shaffer, Department of Geological Sciences, University of Oregon, Eugene).

Subscribe EC Academy on YouTube



a b c

FIGURE 3.25 (a) Image formed from all local means obtained from Fig. 3.24 using Eq. (3.3-21). (b) Image formed from all local standard deviations obtained from Fig. 3.24 using Eq. (3.3-22). (c) Image formed from all multiplication constants used to produce the enhanced image shown in Fig. 3.26.

definitely must be guided by a logical analysis of the enhancement problem at hand. Finally, the choice of size for the local area should be as small as possible in order to preserve detail and keep the computational burden as low as possible. We chose a small (3×3) local region.

Figure 3.25(a) shows the values of $m_{S_{xy}}$ for all values of (x, y). Since the value of $m_{S_{xy}}$ for each (x, y) is the average of the neighboring pixels in a 3×3 area centered at (x, y), we expect the result to be similar to the original image, but

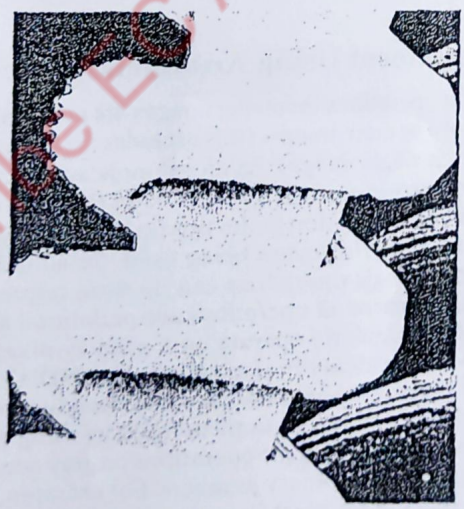


FIGURE 3.26 Enhanced SEM image. Compare with Fig. 3.24. Note in particular the enhanced area on the right side of the image.

Subscribe EC Academy on YouTube

slightly blurred. This indeed is the case in Fig. 3.25(a). Figure 3.25(b) shows in image formed using all the values of $\sigma_{s_{xy}}$. Similarly, we can construct an image out the values that multiply $f(x, y)$ at each coordinate pair (x, y) to form $g(x, y)$. Since the values are either 1 or E, the image is binary, as shown in Fig. 3.25(c). The dark areas correspond to 1 and the light areas to E. Thus, any light point in Fig. 3.25(c) signifies a coordinate pair (x, y) at which the enhancement procedure multiplied $f(x, y)$ by E to produce an enhanced pixel. The dark points represent coordinates at which the procedure did not to modify the pixel values.

The enhanced image obtained with the method just described is shown in Fig. 3.26. In comparing this image with the original in Fig. 3.24, we note the obvious detail that has been brought out on the right side of the enhanced image. It is worthwhile to point out that the unenhanced portions of the image (the light areas) were left intact for the most part. We do note the appearance of some small bright dots in the shadow areas where the coil meets the support stem, and around some of the borders between the filament and the background. These are undesirable artifacts created by the enhancement technique. In other words, the points appearing as light dots met the criteria for enhancement and their values were amplified by factor E. Introduction of artifacts is a definite drawback of a method such as the one just described because of the nonlinear way in which they process an image. The key point here, however, is that the image was enhanced in a most satisfactory way as far as bringing out the desired detail. ■

It is not difficult to imagine the numerous ways in which the example just given could be adapted or extended to other situations in which local enhancement is applicable.

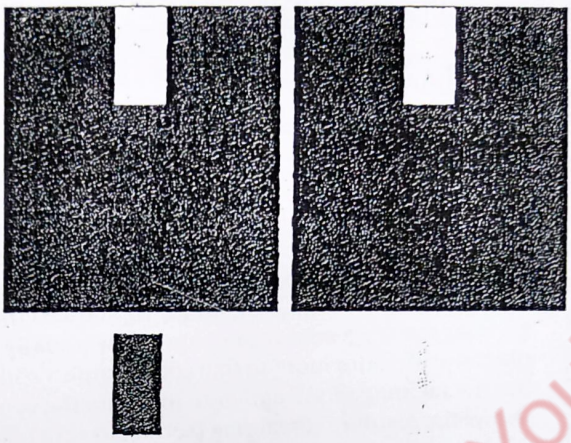
3.4 Enhancement Using Arithmetic/Logic Operations

Arithmetic/logic operations involving images are performed on a pixel-by-pixel basis between two or more images (this excludes the logic operation NOT, which is performed on a single image). As an example, subtraction of two images results in a new image whose pixel at coordinates (x, y) is the difference between the pixels in that same location in the two images being subtracted. Depending on the hardware and/or software being used, the actual mechanics of implementing arithmetic/logic operations can be done sequentially, one pixel at a time, or in parallel, where all operations are performed simultaneously.

Logic operations similarly operate on a pixel-by-pixel basis[1]. We need only be concerned with the ability to implement the AND, OR, and NOT logic operators because these three operators are *functionally complete*. In other words, any other logic operator can be implemented by using only these three basic functions. When dealing with logic operations on gray-scale images, pixel values are processed as strings of binary numbers. For example, performing the NOT operation on a black, 8-bit pixel (a string of eight 0's) produces a white pixel

[1] Recall that, for two binary variables a and b: aANDb yields 1 only when both a and b are 1; otherwise the result is 0. Similarly, aORb is 0 when both variables are 0; otherwise the result is 1. Finally, if a is 1, NOT (a) is 0, and vice versa.

Subscribe EC Academy on YouTube



a b c
d e f

FIGURE 3.27 (a) Original image. (b) AND image mask. (c) Result of the AND operation on images (a) and (b). (d) Original image. (e) OR image mask. (f) Result of operation OR on images (d) and (e).

(a string of eight 1's). Intermediate values are processed the same way, changing all 1's to 0's and vice versa. Thus, the NOT logic operator performs the same function as the negative transformation of Eq. (3.2-1). The AND and OR operations are used for *masking*; that is, for selecting subimages in an image, as illustrated in Fig. 3.27. In the AND and OR image masks, light represents a binary 1 and dark represents a binary 0. Masking sometimes is referred to as *region of interest* (ROI) processing. In terms of enhancement, masking is used primarily to isolate an area for processing. This is done to highlight that area and differentiate it from the rest of the image. Logic operations also are used frequently in conjunction with morphological operations, as discussed in Chapter 9.

Of the four arithmetic operations, subtraction and addition (in that order) are the most useful for image enhancement. We consider division of two images simply as multiplication of one image by the reciprocal of the other. Aside from the obvious operation of multiplying an image by a constant to increase its average gray level, image multiplication finds use in enhancement primarily as a masking operation that is more general than the logical masks discussed in the previous paragraph. In other words, multiplication of one image by another can be used to implement gray-level, rather than binary, masks. We give an example in Section 3.8 of how such a masking operation can be a useful tool. In the remainder of this section, we develop and illustrate methods based on subtraction and addition for image enhancement. Other uses of image multiplication are discussed in Chapter 5, in the context of image restoration.

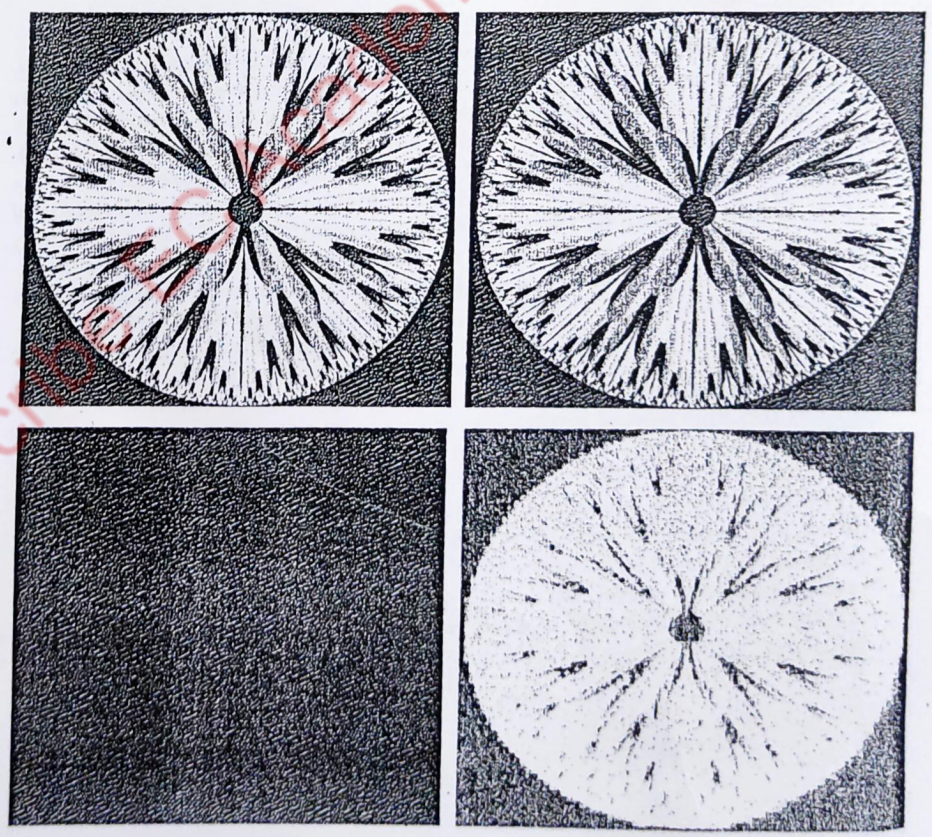
Subscribe EC Academy on YouTube

3.4.1 Image Subtraction

The difference between two images $f(x, y)$ and $h(x, y)$, expressed as

$$g(x, y) = f(x, y) - h(x, y), \tag{3.4-1}$$

is obtained by computing the difference between all pairs of corresponding pixels from f and h. The key usefulness of subtraction is the enhancement of *differences* between images. We illustrate this concept by returning briefly to the discussion in Section 3.2.4, where we showed that the higher-order bit planes of an image carry a significant amount of visually relevant detail, while the lower planes contribute more to fine (often imperceptible) detail. Figure 3.28(a) shows the fractal image used earlier to illustrate the concept of bit planes. Figure 3.28(b) shows the result of discarding (setting to zero) the four least significant bit planes of the original image. The images are nearly identical visually, with the exception of a very slight drop in overall contrast due to less variability of the gray-level values in the image of Fig. 3.28(b). The pixel-by-pixel difference between these two images is shown in Fig. 3.28(c). The differences in pixel values are so small that the difference image appears nearly black when displayed on an 8-bit



a b
c d

FIGURE 3.28 (a) Original fractal image. (b) Result of setting the four lower-order bit planes to zero. (c) Difference between (a) and (b). (d) Histogram-equalized difference image. (Original image courtesy of Ms. Melissa D. Binde, Swarthmore College, Swarthmore, PA).

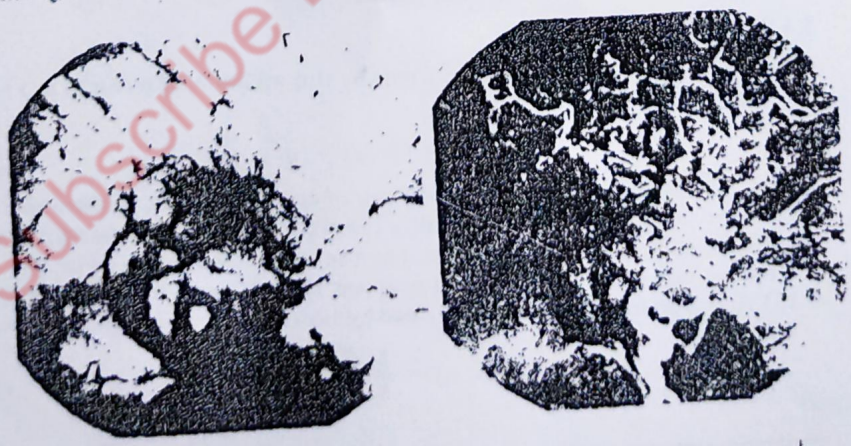
Subscribe to Codeacademy on YouTube

display. In order to bring out more detail, we can perform a contrast stretching transformation, such as those discussed in Sections 3.2 or 3.3. We chose histogram equalization, but an appropriate power-law transformation would have done the job also. The result is shown in Fig. 3.28(d). This is a very useful image for evaluating the effect of setting to zero the lower-order planes.

EXAMPLE 3.7: Use of image subtraction in mask mode radiography.

One of the most commercially successful and beneficial uses of image subtraction is in the area of medical imaging called *mask mode radiography*. In this case $h(x, y)$, the *mask*, is an X-ray image of a region of a patient's body captured by an intensified TV camera (instead of traditional X-ray film) located opposite an X-ray source. The procedure consists of injecting a contrast medium into the patient's bloodstream, taking a series of images of the same anatomical region as $h(x, y)$, and subtracting this mask from the series of incoming images after injection of the contrast medium. The net effect of subtracting the mask from each sample in the incoming stream of TV images is that the areas that are different between $f(x, y)$ and $h(x, y)$ appear in the output image as enhanced detail. Because images can be captured at TV rates, this procedure in essence gives a movie showing how the contrast medium propagates through the various arteries in the area being observed.

Figure 3.29(a) shows an X-ray image of the top of a patient's head prior to injection of an iodine medium into the bloodstream. The camera yielding this image was positioned above the patient's head, looking down. As a reference point, the bright spot in the lower one-third of the image is the core of the spinal column. Figure 3.29(b) shows the difference between the mask (Fig. 3.29a) and an image taken some time after the medium was introduced into the bloodstream. The bright arterial paths carrying the medium are unmistakably enhanced in Fig. 3.29(b). These arteries appear quite bright because they are not subtracted out (that is, they are not part of the mask image). The overall background is much darker than that in Fig. 3.29(a) because differences between areas of little change yield low values, which in turn appear as dark shades of gray in the difference image. Note, for instance, that the spinal cord, which is bright in Fig. 3.29(a), appears quite dark in Fig. 3.29(b) as a result of subtraction.



a b

FIGURE 3.29 Enhancement by image subtraction. (a) Mask image. (b) An image (taken after injection of a contrast medium into the bloodstream) with mask subtracted out.

Subscribe EC Academy on YouTube

A few comments on implementation are an order before we leave this section. In practice, most images are displayed using 8 bits (even 24-bit color images consists of three separate 8-bit channels). Thus, we expect image values not to be outside the range from 0 to 255. The values in a difference image can range from a minimum of −255 to a maximum of 255, so some sort of scaling is required to display the results. There are two principal ways to scale a difference image. One method is to add 255 to every pixel and then divide by 2. It is not guaranteed that the values will cover the entire 8-bit range from 0 to 255, but all pixel values definitely will be within this range. This method is fast and simple to implement, but it has the limitations that the full range of the display may not be utilized and, potentially more serious, the truncation inherent in the division by 2 will generally cause loss in accuracy.

If more accuracy and full coverage of the 8-bit range are desired, then we can resort to another approach. First, the value of the minimum difference is obtained and its negative added to all the pixels in the difference image (this will create a modified difference image whose minimum values is 0). Then, all the pixels in the image are scaled to the interval [0, 255] by multiplying each pixel by the quantity 255/Max, where Max is the maximum pixel value in the modified difference image. It is evident that this approach is considerably more complex and difficult to implement.

Before leaving this section we note also that change detection via image subtraction finds another major application in the area of segmentation, which is the topic of Chapter 10. Basically, segmentation techniques attempt to subdivide an image into regions based on a specified criterion. Image subtraction for segmentation is used when the criterion is "changes." For instance, in tracking (segmenting) moving vehicles in a sequence of images, subtraction is used to remove all stationary components in an image. What is left should be the moving elements in the image, plus noise.

3.4.2 Image Averaging

Consider a noisy image $g(x, y)$ formed by the addition of noise $\eta(x, y)$ to an original image $f(x, y)$; that is,

$$g(x, y) = f(x, y) + \eta(x, y) \tag{3.4-2}$$

where the assumption is that at every pair of coordinates (x, y) the noise is uncorrelated[1] and has zero average value. The objective of the following procedure is to reduce the noise content by adding a set of noisy images, $\{g_i(x, y)\}$.

If the noise satisfies the constraints just stated, it can be shown (Problem 3.15) that if an image $\bar{g}(x, y)$ is formed by averaging K different noisy images,

$$\bar{g}(x, y) = \frac{1}{K}\sum_{i=1}^{K} g_i(x, y) \tag{3.4-3}$$

[1] Recall that the variance of a random variable x with mean m is defined as $E[(x - m)^2]$, where $E\{\cdot\}$ is the expected value of the argument. The covariance of two random variables x_i and x_j is defined as $E[(x_i - m_i)(x_j - m_j)]$. If the variables are *uncorrelated*, their covariance is 0.

Subscribe EC Academy on YouTube

then it follows that

$$E\{\bar{g}(x, y)\} = f(x, y) \tag{3.4-4}$$

and

$$\sigma^2_{\bar{g}(x,y)} = \frac{1}{K}\sigma^2_{\eta(x,y)} \tag{3.4-5}$$

where $E\{\bar{g}(x, y)\}$ is the expected value of $\bar{g}$, and $\sigma^2_{\bar{g}(x,y)}$ and $\sigma^2_{\bar{\eta}(x,y)}$ are the variances of $\bar{g}$ and η, all at coordinates (x, y). The standard deviation at any point in the average image is

$$\sigma_{\bar{g}(x,y)} = \frac{1}{\sqrt{K}}\sigma_{\eta(x,y)}. \tag{3.4-6}$$

As K increases, Eqs. (3.4-5) and (3.4-6) indicate that the variability (noise) of the pixel values at each location (x, y) decreases. Because $E\{\bar{g}(x, y)\} = f(x, y)$, this means that $\bar{g}(x, y)$ approaches $f(x, y)$ as the number of noisy images used in the averaging process increases. In practice, the images $g_i(x, y)$ must be registered (aligned) in order to avoid the introduction of blurring and other artifacts in the output image.

EXAMPLE 3.8: Noise reduction by image averaging.

An important application of image averaging is in the field of astronomy, where imaging with very low light levels is routine, causing sensor noise frequently to render single images virtually useless for analysis. Figure 3.30(a) shows an image of a galaxy pair called NGC 3314, taken by NASA's Hubble Space Telescope with a wide field planetary camera. NGC 3314 lies about 140 million light-years from Earth, in the direction of the southern-hemisphere constellation Hydra. The bright stars forming a pinwheel shape near the center of the front galaxy have formed recently from interstellar gas and dust. Figure 3.30(b) shows the same image, but corrupted by uncorrelated Gaussian noise with zero mean and a standard deviation of 64 gray levels. This image is useless for all practical purposes. Figures 3.30(c) through (f) show the results of averaging 8, 16, 64, and 128 images, respectively. We see that the result obtained with $K = 128$ is reasonably close to the original in visual appearance.

We can get a better appreciation from Fig. 3.31 for how reduction in the visual appearance of noise takes place as a function of increasing K. This figure shows the difference images between the original [Fig. 3.30(a)] and each of the averaged images in Figs. 3.30(c) through (f). The histograms corresponding to the difference images are also shown in the figure. As usual, the vertical scale in the histograms represents number of pixels and is in the range $[0, 2.6 \times 10^4]$. The horizontal scale represents gray level and is in the range $[0, 255]$. Notice in the histograms that the mean and standard deviation of the difference images decrease as K increases. This is as expected because, according to Eqs. (3.4-3) and (3.4-4), the average image should approach the original as K increases. We can also see the effect of a decreasing mean in the difference images on the left column of Fig. 3.31, which become darker as the K increases.

Scanned by CamScanner

Module-3

Image Restoration

* Restoration, attempts to reconstruct or recover an image that has been degraded by using a priori knowledge of the degradation phenomenon.

* Thus restoration techniques are oriented toward modeling the degradation & applying the inverse process in order to recover the original image.

* Some restoration techniques are best formulated in the spatial domain, while others are better suited for the frequency domain. For example, spatial processing is applicable when the only degradation is additive noise. On the other hand, degradations such as image blur are difficult to approach in the spatial domain using small masks. Therefore in this case, frequency domain filters based on various criterion of optimal are used.

Note: Spatial domain refers to image plane itself & based on direct manipulation of pixels in an image.
Frequency domain techniques are based on modifying the Fourier Transform of an image

3.1 A Model of the Image Degradation / Restoration Process:-

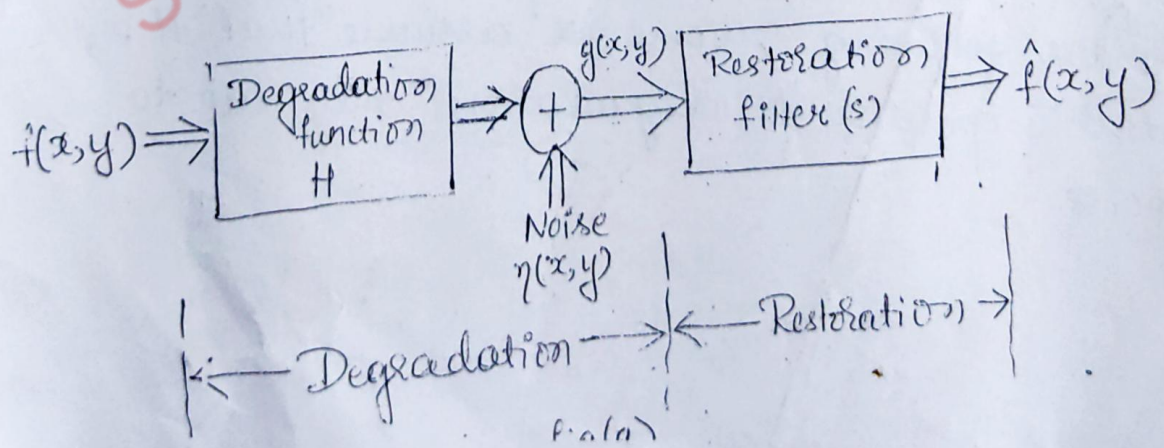


Fig.(a)

Subscribe EC Academy on YouTube

Fig. shows that the degradation function H together with an additive noise term $\eta(x,y)$, operates on an input image $f(x,y)$ to produce a degraded image $g(x,y)$.

If H is a linear, position invariant process, then the degraded image is given in the spatial domain by,

$$g(x,y) = h(x,y) * f(x,y) + \eta(x,y) \quad \text{--- (1)}$$

where $h(x,y) \rightarrow$ spatial representation of the degradation function

the symbol '*' indicates convolution.

The convolution in spatial domain is equal to multiplication in the frequency domain.

So eqn (1) in frequency domain can be represented as,

$$G(u,v) = H(u,v)F(u,v) + N(u,v) \quad \text{--- (2)}$$

where the terms in capital letters are the Fourier transforms of the corresponding terms in eqn (1).

* In the following sections, we assume that H is the identity operator & degradations only due to noise.

Subscribe EC Academy on YouTube

3.2 Noise Models :-

* The principal sources of noise in digital images arise during image acquisition (digitization) and/or transmission.

* The performance of the image sensors is affecte by a variety of factors such as environmental conditions during image acquisition & by the quality of the sensing elements themselves.

* For example, in acquiring images with a CCD camera, light levels & sensor temperature are major factors affecting the amount of noise in resulting image.

* Images are corrupted during transmission due to interference in the channel used for tx^n. For example, an image tx^d using a wireless n/k might be corrupted as a result of lightning or other atmospheric disturbance.

3.2a Spatial & Frequency Properties of Noise :-

Spatial properties refers to spatial characteris of noise.

Frequency properties refer to the frequency cont of noise in Fourier sense. For example, when the Fourier spectrum of noise is constant, the noise usu is called white noise.

Subscribe EC Academy on YouTube

3.2b Some important Noise Probability Density functions:-

The spatial noise descriptor is the statistical behavior of the gray-level values in the noise component of the model. These may be considered random variables, characterized by a probability density function (PDF).

The following are the most common PDFs found in image processing applications.

PDF → defines the probability distribution of noise

1) Gaussian Noise:-

Because of its mathematical tractability in both the spatial & frequency domains, Gaussian noise models (also called normal noise model) are frequently used in practice.

The PDF of a Gaussian random variable, z is given by,

$$P(z) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-(z-\mu)^2/2\sigma^2} \quad \text{——} \quad (3)$$

where z represents gray level

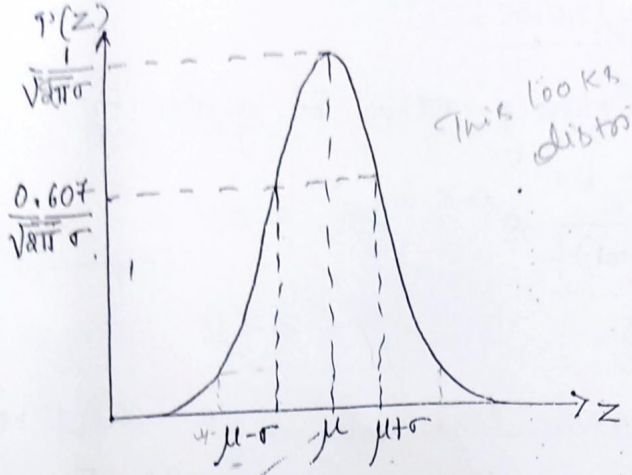
mu ← μ is the mean of average value of z

σ is the standard deviation.

sigma ← σ^2 is the variance of z

A plot of this function is shown in fig(b) below. When z is described by eqn (3), approximately 70% of its values will be in the range $[(\mu-\sigma), (\mu+\sigma)]$ & about 95% will be in range $[(\mu-2\sigma), (\mu+2\sigma)]$.

Subscribe EC Academy on YouTube



This looks as normal distribution.

→ Most common noise model.

→ Other models can also be approximated as Gaussian noise model

→ mathematically easy to calculate

fig(b): Gaussian Probability density function

2) Rayleigh noise:—

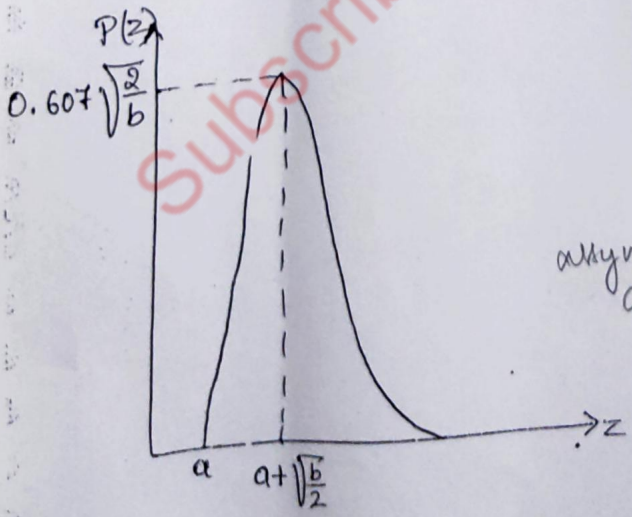
The PDF of Rayleigh noise is given by,

$$p(z) = \begin{cases} \frac{2}{b}(z-a)e^{-(z-a)^2/b} & \text{for } z \geq a \\ 0 & \text{for } z < a \end{cases} \quad \text{—— (4)}$$

The mean & variance of this density are given by

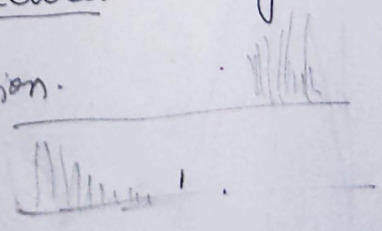
$$\mu = a + \sqrt{\pi b/4} \quad \text{—— (5)}$$

$$\& \quad \sigma^2 = \frac{b(4-\pi)}{4} \quad \text{—— (6)}$$



The Rayleigh density is quite useful for approximating skewed histograms.

asymmetric distribution.



fig(c): Rayleigh-density

Subscribe EC Academy on YouTube

3) Erlang (Gamma) noise :-

The PDF of Erlang noise is given by,

$$P(z) = \begin{cases} \dfrac{a^b z^{b-1}}{(b-1)!} e^{-az} & \text{for } z \geq 0 \\ 0 & \text{for } z < 0 \end{cases} \quad \text{——— (7)}$$

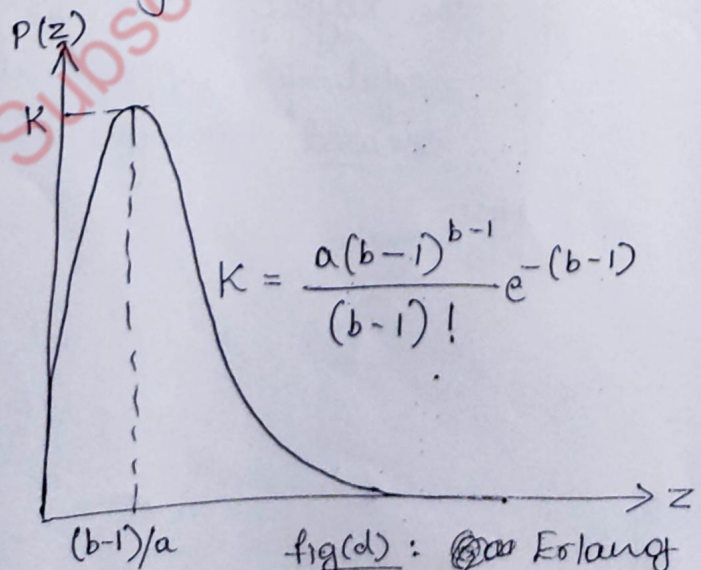
where the parameters are such that $a > 0$, b is a positive integer & "!" indicates factorial.

The mean & variance of this density are given by,

$$\mu = \frac{b}{a} \quad \text{— (8)}$$

$$\& \quad \sigma^2 = \frac{b}{a^2} \quad \text{— (9)}$$

eqn (7) is often referred to as the gamma density only when the denominator is the gamma function $\Gamma(b)$. When the denominator is as shown, the density is more appropriately called the Erlang density.



fig(d) : Erlang density (Gamma)

Subscribe EC Academy on YouTube

4) Exponential noise:-

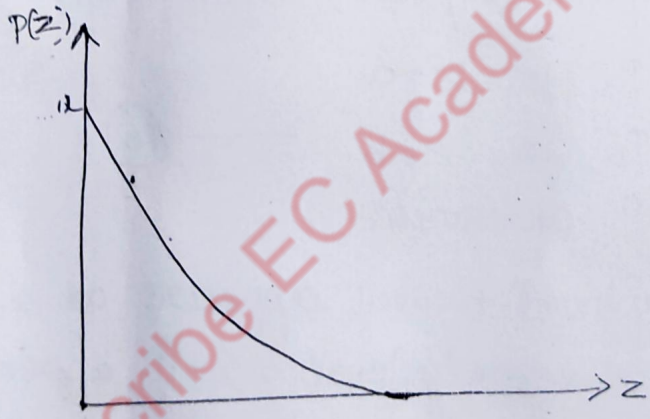
The PDF of exponential noise is given by,

$$p(z) = \begin{cases} ae^{-az} & \text{for } z \geq 0 \\ 0 & \text{for } z < 0 \end{cases} \quad \text{—— (10)}$$

Where $a > 0$. The mean & variance of this density function are $\mu = \frac{1}{a}$ —— (11)

$$\& \ \sigma^2 = \frac{1}{a^2} \quad \text{—— (12)}$$

This PDF is a special case of the Erlang PDF, with $b = 1$.



fig(e): Exponential density function

5) Uniform noise:-

The PDF of uniform noise is given by,

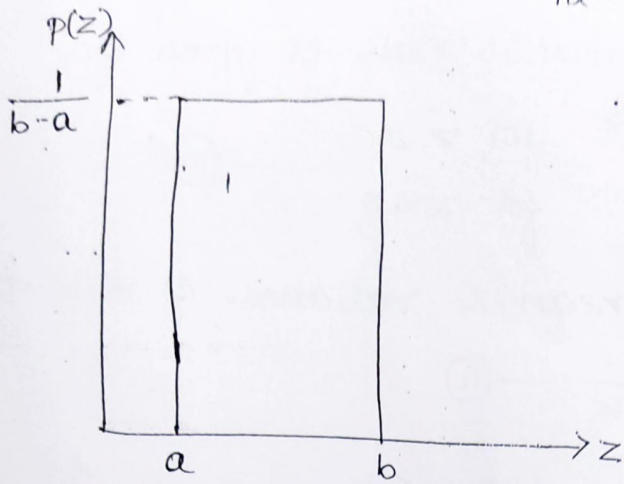
$$p(z) = \begin{cases} \frac{1}{b-a} & \text{if } a \leq z \leq b \\ 0 & \text{otherwise.} \end{cases} \quad \text{—— (13)}$$

The mean of this density function is given by,

$$\mu = \frac{a+b}{2} \quad \text{—— (14)}$$

Subscribe EC Academy on YouTube

& its variance by, $\sigma^2 = \frac{(b-a)^2}{12}$ ——— (15)



fig(f): Uniform density function.

6) Impulse (salt-and-pepper) noise :—

The PDF of impulse (bipolar) noise is given by,

$$p(z) = \begin{cases} P_a & \text{for } z = a \\ P_b & \text{for } z = b \\ 0 & \text{otherwise} \end{cases} \quad \text{——— (16)}$$

If $b > a$, gray-level b will appear as a light dot in the image. Conversely, level a will appear like a dark dot. If either P_a or P_b is zero, the impulse noise is called unipolar.

If neither probability is zero, & especially if they are approximately equal, impulse noise values will resemble salt-&-pepper granules randomly distributed over the image. For this reason, bipolar impulse noise also is called salt-and-pepper noise. Shot & Spike noise are the terms also used to refer to this type of noise.

Subscribe EC Academy on YouTube

Noise impulses can be − or +.

Since impulse corruption is large compared with the strength of image signal, impulse noise is digitized as extreme (pure black or white) values in an image. Thus, a & b are "saturated" values in the sense that they are equal to the minimum & maximum allowed values in the digitized imag As a result, negative impulses appear as black (pepper) points in an image & positive impulses appear white (salt) noise. For an 8-bit image, this means, a = 0 (black) & b = 255 (white).

fig (g): Impulse noise PDF.

—— * —————— * ——

Note :-

* Gaussian noise arises in an image due to factors such as electronic ckt noise & sensor noise due to poor illumination & /or high temperature.

* Rayleigh density is helpful in range imaging.
used in digital cameras

Subscribe EC Academy on YouTube

* The exponential & gamma densities find application in laser imaging.

* Impulse noise is found in situations where quick transients, such as faulty switching, take place during imaging.

* The uniform density is the least descriptive of practical situations. It is useful as the basis for numerous random number generators that are used in simulations.

3.2.c Periodic Noise:-

* Periodic noise in an image arises typically from electrical & electromechanical interference during image acquisition.

* This is the only type of spatially dependent noise.

* Periodic noise can be reduced significantly via frequency domain filtering.

Ex:- Refer to image fig. 5.5(a), which is severely corrupted by (spatial) sinusoidal noise of various frequencies. The Fourier transform of a pure sinusoid is a pair of conjugate impulses located at the conjugate frequencies of the sinewave. In fig(b), the impulses appearing in an approximate circle because the frequency values are so arranged

Subscribe EC Academy on YouTube

3.2d Estimation of Noise Parameters:-

* The parameters of periodic noise are estimated by inspection of the Fourier spectrum of the image, as it tends to produce frequency spikes that often can be detected even by visual analysis.

* Automated analysis is possible when some knowled is available about the general location of frequency components of the interference.

* If the imaging system is available, one simple way to study the chr's of system noise is to capture a set of images of "flat" environments. For example, in case of optical sensor, this is as simple as imaging a solid gray board that illuminated uniformly. The resulting image are good indicator of s/m noise

* The simplest way to use the data from image strips for calculating the mean & variance of the gray levels. Consider a strip (subimage) denoted by S. we can use sample approximations from basic statistics:

$$\mu = \sum_{z_i \in S} z_i \, p(z_i) \quad \text{—— (i)}$$

$$\& \quad \sigma^2 = \sum_{z_i \in S} (z_i - \mu)^2 \, p(z_i) \quad \text{—— (ii)}$$

where z_i's are gray-level values of the pixels in S & $p(z_i)$ are corresponding normalised histogram values. The shape of the histogram identifies th. closest PDF match. The heights of the peaks corresponding to black & white pixels are the estimates of P_a & P_b.

Note:- Refer to fig. 5.6 in textbook.

Subscribe EC Academy on YouTube

3.3 Restoration in the presence of Noise only - Spatial filtering:-

* When the degradation in an image is only due to noise, then eqn's (1) & (2) becomes

$$g(x,y) = f(x,y) + \eta(x,y) \quad \text{--- (1a)}$$

$$\& \; G(u,v) = F(u,v) + N(u,v) \quad \text{--- (2a)}$$

* In case of periodic noise, it is possible to estimate $N(u,v)$ from the spectrum $G(u,v)$. i.e., $N(u,v)$ can be subtracted from $G(u,v)$ to obtain an estimate of the original image.

* Spatial filtering is the method of choice in situations when only additive noise is present.

1) Mean Filters:-

i) Arithmetic mean filter:-

* This is the simplest of the mean filters.
* Let S_{xy} represent the set of coordinates in a rectangular subimage window of size $m \times n$, centered at point (x,y).
* The arithmetic mean filtering process computes the average value of the corrupted image $g(x,y)$ in the area defined by S_{xy}.

Subscribe EC Academy on YouTube

* The value of restored image $\hat{f}$ at any point (x, y) is simply the arithmetic mean computed using the pixel in the region defined by S_{xy}.

In other words,

Summation → addition.

$$\hat{f}(x,y) = \frac{1}{mn} \sum_{(s,t) \in S_{xy}} g(s,t) \quad \text{——— (i)}$$

* A mean filter simply smooths local variations in an image. Noise is reduced as a result of blurring.

ii) <u>Geometric mean filter</u>:—

An image restored using a geometric mean filter is given by the expression,

Pie → ~~[illegible]~~ Multiplication.

$$\hat{f}(x,y) = \left[\prod_{(s,t) \in S_{xy}} g(s,t) \right]^{\frac{1}{mn}} \quad \text{——— (ii)}$$

Here, each restored pixel is given by the product of the pixels in the subimage window, raised to the power $\frac{1}{mn}$.

A geometric mean filter achieves smoothing comparable to the arithmetic mean filter, but it tends to lose less image detail in the process.

iii) <u>Harmonic mean filter</u>:— This <u>operation is given</u> by,

$$\hat{f}(x,y) = \frac{mn}{\sum_{(s,t) \in S_{xy}} \frac{1}{g(s,t)}} \quad \text{——— (iii)}$$

Subscribe EC Academy on YouTube

The harmonic mean filter works well for salt noise but fails for pepper noise. It does well also with other types of noise like Gaussian noise.

iv) Contraharmonic mean filter:-

$$\hat{f}(x,y) = \frac{\sum_{(s,t)\in S_{xy}} g(s,t)^{Q+1}}{\sum_{(s,t)\in S_{xy}} g(s,t)^{Q}} \quad \text{——— (iv)}$$

where Q is the order of the filter.

This filter is well suited for reducing the effects of salt-and-pepper noise. For positive values of Q, it eliminates pepper noise. For negative values of Q, it eliminates salt noise. It cannot do both simultaneously.

The contraharmonic filter reduces to arithmetic mean filter if Q=0, & to the harmonic mean filter if Q=-1.
(Refer to fig. 5.8 & 5.9)

2) Order-Statistics Filters:-

These are spatial filters whose response is based on ordering (ranking) the pixels in the image area encompassed by the filter. The response of filter at any point is determined by the ranking result.

Subscribe EC Academy on YouTube

i) Median filter :—

Median filter replaces the value of a pixel by the median of the gray levels in the neighborhood of that pixel.

$$\hat{f}(x,y) = \underset{(s,t)\in S_{xy}}{\text{median}}\{g(s,t)\} \quad \text{——(i)}$$

Median filter provide excellent noise-reduction capabilities, with less blurring than linear smoothing filters of similar size.

Median filters are effective in the presence of both bipolar & unipolar impulse noise.

ii) Min and Max filters :—

The median represents the 50th percentile of a ranked set of no.'s

Using 100th percentile results in max filter given by,

$$\hat{f}(x,y) = \max_{(s,t)\in S_{xy}}\{g(s,t)\} \quad \text{——(ii)}$$

This filter is useful for finding the brightest points in an image. Pepper noise can be reduced by the filter since it has very low values.

The 0th percentile filter is the min filter :

$$\hat{f}(x,y) = \min_{(s,t)\in S_{xy}}\{g(s,t)\} \quad \text{——(iii)}$$

This filter is useful for finding darkest points in an image. Also it reduces salt noise.

Subscribe EC Academy on YouTube

(iii) Midpoint filter:-

This filter computes the midpoint b/w the maximum & minimum values in the area encompassed by the filter:

$$\hat{f}(x,y) = \frac{1}{2}\left[\max_{(s,t)\in S_{xy}}\{g(s,t)\} + \min_{(s,t)\in S_{xy}}\{g(s,t)\}\right] \quad \text{—— (iv)}$$

This filter combines order statistics & averaging. This filter works best for randomly distributed noise, like Gaussian or Uniform noise.

iv) Alpha-trimmed mean filter:-

Suppose that we delete the $\frac{d}{2}$ lowest & the $d/2$ highest grey-level values of $g(s,t)$ in the neighborhood S_{xy}.

Let $g_r(s,t)$ represent the remaining $mn-d$ pixels. A filter formed by averaging these remaining pixels is called an alpha-trimmed mean filter:

$$\hat{f}(x,y) = \frac{1}{mn-d}\sum_{(s,t)\in S_{xy}} g_r(s,t) \quad \text{—— (v)}$$

where the value of d can range from 0 to $mn-1$.

When $d=0$, alpha trimmed filter reduces to the arithmetic mean filter. If we choose $d=(mn-1)/2$, the filter becomes a median filter.

Subscribe EC Academy on YouTube

For other values of d, the alpha-trimmed filter is useful in situations involving multiple types of noise such as combination of salt-and-pepper and Gaussian noise.

3.3c Adaptive Filters:-

i) Adaptive, local noise reduction filter:-

* The simplest statistical measures of a random varial are its mean & variance. The mean gives a measure of average gray level in the region ove which mean is computed, & the variance give a measure of average contrast in that region.

* Our filter is to operate on a local region, S_{xy}. The response of the filter at any point (x,y) on which the region is centered is to be based on four quantities:

(a) $g(x,y)$, the value of noisy image at (x,y)

(b) σ_η^2, the variance of noise corrupting $f(x,y)$ to form $g(x,y)$.

(c) m_L, the local mean of the pixels in S_{xy}

(d) σ_L^2, the local variance of the pixels in S_{xy}.

* An adaptive expression for obtaining $\hat{f}(x,y)$ may be written as,

$$\hat{f}(x,y) = g(x,y) - \frac{\sigma_\eta^2}{\sigma_L^2}\left[g(x,y) - m_L\right]$$

[Refer to example fig-5.13]

Subscribe EC Academy on YouTube

ii) Adaptive median filter:—

* The median filter performs well, as long as the spatial density of impulse noise is not large i.e., P_a & P_b, less than 0.2.

* The adaptive median filtering can handle impulse noise with probabilities even larger than these.

* An additional benefit of adaptive median filter is that it seeks to preserve detail while smoothing nonimpulse noise.

* The o/p of the filter is a single value used to replace the value of the pixel at (x, y), the particular point on which the window S_{xy} is centered at a given time.

* Let Z_{min} = minimum graylevel value in S_{xy}
  Z_{max} = maximum graylevel value in S_{xy}
  Z_{med} = median of gray levels in S_{xy}
  Z_{xy} = gray level at coordinates (x, y)
  S_{max} = maximum allowed size of S_{xy}.

* The adaptive median filtering algorithm works in two levels, denoted level A & level B, as follows:

Level A: $A1 = Z_{med} - Z_{min}$
$A2 = Z_{med} - Z_{max}$
If $A1 > 0$ AND $A2 < 0$, Go to level B
Else increase the window size

Subscribe EC Academy on YouTube

If window size $\leq S_{max}$ repeat level A
Else output Z_{xy}.

Level B: $B1 = Z_{xy} - Z_{min}$

$B2 = Z_{xy} - Z_{max}$

If $B1 > 0$ AND $B2 < 0$, output Z_{xy}
Else output Z_{med}.

This algorithm is used for 3 main purposes:

1) to remove salt-and-pepper (impulse) noise,
2) to provide smoothing
3) to reduce distortion, such as excessive thinning or thickening of object boundaries.

3.4 Periodic Noise Reduction by Frequency Domain Filtering:

is the noise added to the image due to electrical/electro mechanical interference during image acquisition.

– involves removing particular range freq from an image.

The Band reject, Band pass & notch filters are use as tools for periodic noise reduction or removal.

i) Band reject Filters:—

* This removes or attenuate a band of frequencies about the origin of the Fourier Transform.
* An ideal band reject filter is given by

$$\underbrace{H(u,v)}_{\text{o/p of filter}} = \begin{cases} 1 & \text{if } \underbrace{D(u,v)}_{\text{freq in original image}} < D_0 - \frac{W}{2} \\ 0 & \text{if } D_0 - \frac{W}{2} \leq D(u,v) \leq D_0 + \frac{W}{2} \\ 1 & \text{if } D(u,v) > D_0 + \frac{W}{2} \end{cases} \quad \text{—— (i)}$$

Subscribe ECAcademy on YouTube

where $D(u,v)$ is the distance from the origin of the centered frequency rectangle.

W is the width of the band

D_0 is its radial center.

* Similarly, a Butterworth bandreject filter of order n is given by

$$H(u,v) = \frac{1}{1 + \left[\frac{D(u,v)W}{D^2(u,v) - D_0^2}\right]^{2n}} \quad \text{--- (ii)}$$

& a Gaussian bandreject filter is given by,

$$H(u,v) = 1 - e^{-\frac{1}{2}\left[\frac{D^2(u,v) - D_0^2}{D(u,v)W}\right]^2} \quad \text{--- (iii)}$$

ii) <u>Bandpass Filters:-</u>

* A bandpass filter performs the opposite operation of a bandreject filter.

* So, the transfer function $H_{bp}(u,v)$ of a bandpass filter is obtained from a corresponding bandreject filter with transfer function $H_{br}(u,v)$ by using the eqn,

$$H_{bp}(u,v) = 1 - H_{br}(u,v) \quad \text{--- (1)}$$

Subscribe FC Academy on YouTube

iii) Notch Filters :—

* A notch filters (or passes) frequencies in predef neighborhoods about a center frequency.

* The transfer function of an ideal notch reje filter of radius D_0 with centres at (u_0, v_0) & by symmetry at $(-u_0, -v_0)$ is

$$H(u,v) = \begin{cases} 0 & \text{if } D_1(u,v) \le D_0 \text{ or } D_2(u,v) \le D_0 \\ 1 & \text{otherwise} \end{cases} \quad \text{— (1)}$$

where $D_1(u,v) = \left[(u - M/2 - u_0)^2 + (v - N/2 - v_0)^2\right]^{1/2}$ — (2)

& $D_2(u,v) = \left[(u - M/2 + u_0)^2 + (v - N/2 + v_0)^2\right]^{1/2}$ — (3)

* For Butterworth notch reject filter of order n. is given by,

$$H(u,v) = \frac{1}{1 + \left[\frac{D_0^2}{D_1(u,v) D_2(u,v)}\right]^n} \quad \text{— (4)}$$

* A Gaussian notch reject filter has the form,

$$H(u,v) = 1 - e^{-\frac{1}{2}\left[\frac{D_1(u,v) D_2(u,v)}{D_0^2}\right]} \quad \text{— (5)}$$

* Note that these 3 filters become high pass filters if $u_0 = v_0 = 0$.

Subscribe EC Academy on YouTube

* The notch filters that pass the frequencies are given by,

$$H_{np}(u,v) = 1 - H_{nr}(u,v) \quad \text{--- (6)}$$

iv) Optimum Notch Filtering:-

* Images derived from electro-optical scanners are sometimes corrupted by coupling & amplification of low-level signals in the scanner electronic circuitry. The resulting images tend to contain pronounced, 2-D periodic structures superimposed on the scene data with patterns more complex than earlier.

* When several interference components are present, the methods in the preceding sections are not always acceptable because they may remove too much image information in the filtering process.

* The first step is to extract the principal frequency components of interference pattern. The FT of interference noise pattern is given by the expression,

$$N(u,v) = H(u,v)\,G(u,v) \quad \text{--- (1)}$$

* After a particular filter has been selected, the corresponding pattern in spatial domain is obtained from, $n(x,y) = \mathcal{F}^{-1}\{H(u,v)\,G(u,v)\}$ --- (2)

Subscribe EC Academy on YouTube

$$\hat{f}(x,y) = g(x,y) - w(x,y)\,\eta(x,y) \quad \text{--- (3)}$$

where $w(x,y)$ is a weighting or modulation funct

* Consider a neighborhood of size $(2a+1)$ by $(2b+1)$ about a point (x,y).

$$\sigma^2(x,y) = \frac{1}{(2a+1)(2b+1)} \sum_{s=-a}^{a} \sum_{t=-b}^{b} \left[\hat{f}(x+s, y+t) - \bar{\hat{f}}(x,y)\right]^2$$

where $\bar{\hat{f}}$ is the avg. value of $\hat{f}$ in neighbood,

$$\bar{\hat{f}}(x,y) = \frac{1}{(2a+1)(2b+1)} \sum_{s=-a}^{a} \sum_{t=-b}^{b} \hat{f}(x+s, y+t) \quad \text{--- (5)}$$

Substituting (3) in (4),

$$\sigma^2(x,y) = \frac{1}{(2a+1)(2b+1)} \sum_{s=-a}^{a} \sum_{t=-b}^{b} \Big\{ \big[g(x+s, y+t) - w(x+s, y+t)\,\eta(x+s, y+t) \big] - \big[\bar{g}(x,y) - \overline{w(x,y)\,\eta(x,y)} \big] \Big\}^2 \quad \text{--- (6)}$$

Assuming $w(x,y)$ remains constant over the neighborhood gives the approximation,

$$w(x+s, y+t) = w(x,y) \quad \text{--- (7)}$$

for $-a \le s \le a$ & $-b \le t \le b$. This assumption also results in

$$\overline{w(x,y)\,\eta(x,y)} = w(x,y)\,\bar{\eta}(x,y) \quad \text{--- (8)}$$

Subscribe EC Academy on YouTube

(6) $\Rightarrow$ $$\sigma^2(x,y) = \frac{1}{(2a+1)(2b+1)} \sum_{s=-a}^{b} \sum_{t=-b}^{b} \Big\{ \big[g(x+s, y+t) - w(x,y)\,\eta(x+s, y+t) \big] - \big[\bar{g}(x,y) - w(x,y)\,\bar{\eta}(x,y) \big] \Big\}^2 \quad \text{--- (9)}$$

To minimize $\sigma^2(x,y)$,

$$\frac{\partial \sigma^2(x,y)}{\partial w(x,y)} = 0 \quad \text{--- (10)}$$

The result is,

$$w(x,y) = \frac{\overline{g(x,y)\,\eta(x,y)} - \bar{g}(x,y)\,\bar{\eta}(x,y)}{\bar{\eta}_2(x,y) - \bar{\eta}^{2}(x,y)} \quad \text{--- (11)}$$

To obtain the restored image $\hat{f}(x,y)$, we compute $w(x,y)$ from (11) & then make use of (3). As $w(x,y)$ is assumed to be constant in a neighborhood, comp this function for every value of x & y in the image unnecessary. Instead, $w(x,y)$ is computed for one point in each nonoverlapping neighborhood (preferab the center point) & then used to process all the ima points contained in that neighborhood.

Subscribe EC Academy on YouTube

3.5 Linear, Position-Invariant Degradations:-

* The input-output relationship before the restoration stage is expressed as,

$$g(x,y) = H[f(x,y)] + \eta(x,y) \quad \text{--- (1)}$$

Let assume that $\eta(x,y)=0$, so that

$$g(x,y) = H[f(x,y)].$$

* H is <u>linear</u> if,

$$H[a f_1(x,y) + b f_2(x,y)] = aH[f_1(x,y)] + bH[f_2(x,y)] \quad \text{--- (2)}$$

where a & b are scalars and $f_1(x,y)$ & $f_2(x,y)$ are any two i/p images.

* If $a=b=1$, eqn (2) becomes

$$H[f_1(x,y) + f_2(x,y)] = H[f_1(x,y)] + H[f_2(x,y)] \quad \text{--- (3)}$$

which is called property of <u>additivity</u>, which say that, if H is a linear operator, the response to a sum of two inputs is equal to the sum of the two responses.

* with $f_2(x,y)=0$, eqn (2) becomes,

$$H[a f_1(x,y)] = a H[f_1(x,y)] \quad \text{--- (4)}$$

which is called the property of <u>homogeneity</u>. It say that the response to a constant multiple of any input is equal to the response to that input multiplied by the same constant.

* Thus a linear operator possesses both the properties of additivity & property of <u>homogeneity</u>.

Subscribe EC Academy on YouTube

i) Estimation by Image Observation:-

* Suppose we are given a degraded image without any knowledge about the degradation function H. One way to estimate this function is to gather information from the image itself.

* For example, if the image is blurred, we can look at a small section of the image containing simple structures like part of an object & the background.

In order to reduce the effect of noise in our observation, we look for areas of strong signal content.

Using sample gray levels of the object & background, we can construct an unblurred image of the same size & characteristics as the observed subimage.

Let the observed subimage be denoted by $g(x,y)$ & let the constructed subimage be denoted by $\hat{f}_s(x,y)$. Then, assuming the effect of noise is negligible because of our choice of a strong-signal area, it follows that

$$H_s(u,v) = \frac{G_s(u,v)}{\hat{F}_s(u,v)} \quad \text{——— (1)}$$

From this function, we can deduce the complete function $H(u,v)$ by making use of fact that we are assuming position invariance.

Subscribe EC Academy on YouTube

ii) Estimation by Experimentation :—

* If equipment similar to equipment used to acquire the degraded image is available, then it is possible, to obtain an accurate estimate of the degradation.

* Image similar to degraded image can be acquired with various system settings until they are degrad as closely as possible to the image we wish to restore.

* Then the idea is to obtain the impulse response of the degradation by imaging an impulse (small dot of light) using the same system settings. As noted, a linear, space-invariant system is decribed completely by its impulse response.

* An impulse is simulated by a bright dot of light, as bright as possible to reduce the effect of noise. Then, recalling that the Fourier transform of an impulse is a constant, it follows from eqn (17) that,

$$H(u,v) = \frac{G(u,v)}{A} \quad \text{—— (2)}$$

where $G(u,v)$ is the Fourier Transform of the observed image &

A is ~~the~~ constant describing the strength of the impulse.

Subscribe EC Academy on YouTube

iii) Estimation by Modeling:-

* Degradation modeling has been used for many years because of insight it affords into the image restoration problem.

* In some cases, the model can even take into account environmental conditions that cause degradations.

* For example, a degradation model proposed by Hufnagel & Stanley is based on physical characteristics of atmospheric turbulence. This model has a familiar form:

$$\boxed{H(u,v) = e^{-K(u^2+v^2)^{5/6}}} \quad \text{——— (1)}$$

where K is a constant that depends on the nature of the turbulence.

The exception of the exponent 5/6 power in eqn, the above eqn has the same form as the Gaussian lowpass filter. In fact, Gaussian LPF is used sometimes to model mild, uniform blurring.

* Another major approach in modeling is to derive a mathematical model starting from basic principles.

* Consider an image $f(x,y)$ undergoes planar motion & that $x_0(t)$ and $y_0(t)$ are the time varying components of motion in x- & y- directions

Subscribe EC Academy on YouTube

respectively.

The total exposure at any point of the recordi medium is obtained by integrating the instantane exposure over the time interval during which the imaging system shutter is open.

If T is the duration of exposure, it follows that

$$g(x,y) = \int_0^T f\left[x - x_0(t),\ y - y_0(t)\right] dt \quad \text{—} (2)$$

where $g(x,y)$ is blurred image.

From (1), the FT of eqn (2) is,

$$G(u,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)\, e^{-j2\pi(ux+vy)}\, dx\, dy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left[\int_0^T f\left[x - x_0(t),\ y - y_0(t)\right] dt\right] e^{-j2\pi} \quad \text{—} (3)$$

Reversing the order of integration allows eqn (3) to be expressed in the form,

$$G(u,v) = \int_0^T \left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \underbrace{f\left[x - x_0(t),\ y - y_0(t)\right]} e^{-j2\pi(ux+vy)}\, dx\, dy\right] d \quad \text{—} (4)$$

$$G(u,v) = \int_0^T F(u,v)\, e^{-j2\pi\left[u x_0(t) + v y_0(t)\right]} dt$$

$$= F(u,v) \int_0^T e^{-j2\pi\left[u x_0(t) + v y_0(t)\right]} dt \quad \text{—} (5)$$

Thus $F(u,v)$ is independent of t.

By defining, $H(u,v) = \int_0^T e^{-j2\pi\left[u x_0(t) + v y_0(t)\right]} dt$

Subscribe EC Academy on YouTube

Eqn (5) may be expressed as,

$$G(u,v) = H(u,v)\,F(u,v) \quad \text{—— (7)}$$

Suppose that image in acquisition undergoes uniform linear motion in x-direction only, at rate given by $x_0(t) = at/T$. When $t = T$, the image has been displaced by a total distance a. With $y_0(t) = 0$, eqn (6) yields.

$$H(u,v) = \int_0^T e^{-j2\pi u x_0(t)}\,dt$$

$$= \int_0^T e^{-j2\pi u\, at/T}\,dt \quad \text{—— (8)}$$

$$= \frac{T}{\pi u a}\sin(\pi u a)\,e^{-j\pi u a}$$

H vanishes at values of u given by $u = n/a$ where n is an integer. If we allow y to vary by $y_0 = bt/T$, then degradation function becomes

$$H(u,v) = \frac{T}{\pi(ua+vb)}\sin[\pi(ua+vb)]\,e^{-j\pi(ua+vb)} \quad \text{—— (9)}$$

Subscribe EC Academy on YouTube

3.7 Inverse Filtering:-

* The simplest approach to restoration is direct inverse filtering, where we compute an estimate $\hat{F}(u,v)$ of the transform of the original image simply by dividing the transform of the degraded image $G(u,v)$ by the degradation function:

$$\hat{F}(u,v) = \frac{G(u,v)}{H(u,v)} \quad \text{--- (1)}$$

WKT, $G(u,v) = H(u,v)F(u,v) + N(u,v)$ --- (2)

Substituting for $G(u,v)$ in eqn (1) from (2),

$$\hat{F}(u,v) = \frac{H(u,v)F(u,v) + N(u,v)}{H(u,v)}$$

$$\hat{F}(u,v) = F(u,v) + \frac{N(u,v)}{H(u,v)} \quad \text{--- (3)}$$

3.8 Minimum Mean Square Error (Weiner) Filtering:-

* This method is founded on considering images & noise as random processes and the objective is to find an estimate $\hat{f}$ of the uncorrupted image f such that the mean square error b/w them is minimized. This error measure is given by,

$$e^2 = E\{(f-\hat{f})^2\} \quad \text{--- (1)}$$

where $E\{\cdot\}$ is the expected value of the argument.

Subscribe EC Academy on YouTube

* The minimum of the error function in eqn ① is given in the frequency domain by the expression,

$$\hat{F}(u,v) = \left[\frac{H^*(u,v)\, S_f(u,v)}{S_f(u,v)\,|H(u,v)|^2 + S_\eta(u,v)}\right] G(u,v)$$

$$(u,v) = \frac{G(u,v) - N(u,v)}{H(u,v)}$$

$$= \left[\frac{H^*(u,v)}{|H(u,v)|^2 + S_\eta(u,v)/S_f(u,v)}\right] G(u,v) \quad \text{—— ②}$$

$$(u,v) - \hat{F}(u,v) = \frac{H(u,v)}{N(u,v)}$$

$$= \left[\frac{1}{H(u,v)} \cdot \frac{|H(u,v)|^2}{|H(u,v)|^2 + S_\eta(u,v)/S_f(u,v)}\right] G(u,v)$$

where the fact that the product of a complex quantity with its conjugate is equal to the magnitude of the complex quantity squared. This result is known as the wiener filter, named after N. wiener.

The filter which consists of the terms inside the brackets, also is commonly referred to as the minimum mean square error filter or the least squares error filter.

$H(u,v)$ = degradation function

$H^*(u,v)$ = complex conjugate of $H(u,v)$

$|H(u,v)|^2 = H^*(u,v) \cdot H(u,v)$

$S_\eta(u,v) = |N(u,v)|^2$ = power spectrum of noise

$S_f(u,v) = |F(u,v)|^2$ = power spectrum of the undegraded image.

Subscribe EC Academy on YouTube

* An approach used frequently when the spectrum $|N(u,v)|^2$ is not known or cannot be estimated is to approximate eqn ② by the expression,

$$\hat{F}(u,v) = \left[\frac{1}{H(u,v)} \frac{|H(u,v)|^2}{|H(u,v)|^2 + K}\right] G(u,v) \quad \text{--- ③}$$

where K is a specified constant.

3.9 Constrained Least Squares Filtering:-

The disadvantage of the wiener filter is that the power spectra of the undegraded image $S_f(u,v)$ & $S_\eta(u,v)$ noise must be known.

We can express the equation,

$g(x,y) = H[f(x,y)] + \eta(x,y)$ in vector matrix form

as, $g = Hf + \eta$ --- ①

To find the minimum of a criterion function, c, defined as,

$$C = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \left[\nabla^2 f(x,y)\right]^2 \quad \text{--- ②}$$

subject to constraint

$$\| g - H\hat{f} \|^2 = \|\eta\|^2 \quad \text{--- ③}$$

where $\|w\|^2 \triangleq w^T w$ is the Euclidean vector norm

$\hat{f}$ is the estimate of the ungeaded image

∇^2 is the Laplacian operator

The frequency domain solution to this optimization problem is given by,

Subscribe EC Academy on YouTube

$$\hat{F}(u,v) = \left[\frac{H^*(u,v)}{|H(u,v)|^2 + \gamma |P(u,v)|^2}\right] G(u,v) \quad \text{--- (4)}$$

23

γ is a parameter must be adjusted so that the constraint in (3) is satisfied & $P(u,v)$ is the Fourier transform of the function,

$$p(x,y) = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix} \quad \text{--- (5)}$$

A procedure for computing γ by iteration is as follows: Define a "residual" vector r as

$$r = g - H\hat{f} \quad \text{--- (6)}$$

Since $\hat{F}(u,v)$ in (4) is a function of γ, then r also is a function of this parameter. It can be shown that

$$\phi(\gamma) = r^T r$$
$$= \|r\|^2 \quad \text{--- (7)}$$

is a monotonically increasing function of γ. To adjust γ, $\|r\|^2 = \|\eta\|^2 \pm a$ --- (8)

where a is an accuracy factor. In view of eqn (6) if $\|r\|^2 = \|\eta\|^2$, the constraint in eqn (3) will be strictly satisfied.

Since $\phi(\gamma)$ is monotonic, finding the desired value of γ is not difficult. One approach is to

1) Specify an initial value of γ.

Subscribe EC Academy on YouTube

2) Compute $\|r\|$

3) Stop if eqn (8) is satisfied; otherwise return to step 2 after increasing γ if $\|r\|^2 < \|\eta\|^2 - \alpha$ or decreasing γ if $\|r\|^2 > \|\eta\|^2 + \alpha$. Use the new value of γ in eqn (4) to recompute the optimum estimate $\hat{F}(u,v)$.

Other procedures, such as a Newton-Raphson algorithm, can be used to improve the speed of convergence.

In order to use this algorithm, we need the quantities $\|r\|^2$ & $\|\eta\|^2$. To compute $\|r\|^2$, we note from eqn (6) that,

$$R(u,v) = G(u,v) - H(u,v)\hat{F}(u,v) \quad \text{--- (9)}$$

from which we obtain $r(x,y)$ by computing the inverse transform of $R(u,v)$. Then

$$\|r\|^2 = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} r^2(x,y) \quad \text{--- (10)}$$

To compute $\|\eta\|^2$, first consider the variance of the noise over the entire image, which can be estimated by sample-average method:

$$\sigma_\eta^2 = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [\eta(x,y) - m_\eta]^2 \quad \text{--- (11)}$$

where $$m_\eta = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \eta(x,y) \quad \text{--- (12)}$$

is the sample mean.

Subscribe EC Academy on YouTube

With reference to the form of eqn (9), the double summation in eqn (11) is equal to $\|\eta\|^2$. This gives us the expression,

$$\|\eta\|^2 = MN\left[\sigma_\eta^2 + m_\eta^2\right] \quad \text{——— (13)}$$

Thus, we can implement an optimum restoration algorithm by having knowledge of only the mean & variance of the noise.

Subscribe EC Academy on YouTube

Module-4

4.1 COLOR Image Processing

* The use of color in image processing is motivated by two principal factors:
  i) Color is a powerful descriptor that often simplifies object identification & extraction from a scene.
  ii) Humans can discern thousands of color shades & intensities, compared to about only two dozen shad of gray. This factor is particularly important in manual image analysis.

* Color image processing is divided into 2 major areas: full-color and pseudo-color processing.

  In Full-color processing, the images are acquired with a full-color sensor such as a color TV camera or color scanner.

  In pseudocolor processing, color is assigned to a particular monochrome intensity or range of intensities. Until recently, most digital color image processing was done at the pseudocolor level.

  Full-color image processing techniques are now used in a broad range of applications including publishing, visualization & the internet.

Subscribe EC Academy on YouTube

4.1.1 Color Fundamentals:—

* The process followed by the human brain in perceiving & interpreting color is a physiopsychological phenomenon, which is not yet fully understood.

* Sir Issac Newton discovered that when a beam of sunlight passes through a glass prism, the emerging beam of light is not white but consists of a continuous spectrum of colors ranging from violet at one end to red at the other.

Fig. below shows the color spectrum may be divided into 6 broad regions: violet, blue, green, yellow, orange & red.

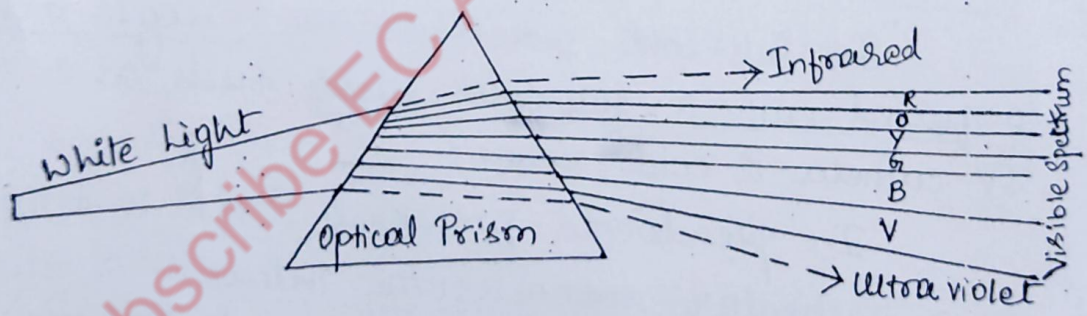


Fig: Color Spectrum seen by passing white light through a prism.

When viewed in full color (EM spectrum), no color in the spectrum ends abruptly, but rather each color blends smoothly into the next.

Subscribe EC Academy on YouTube

* Basically, the colors that humans and some other animals perceive in an object are determined by the nature of the light reflected from the object.

Visible light is composed of a relatively narrow band of frequencies in the EM spectrum. A body that reflects light that is balanced in all visible wavelengths appears white to the observer.

However, a body that favors reflectance in a limited range of the visible spectrum exhibits some shades of color. For example, green objects reflect light with wavelengths primarily in the 500 to 570 nm range while absorbing most of the energy at other wavelengths.

* Achromatic light is what viewers see on a black & white television set. If the light is achromatic (void of color), its only attribute is its intensity, or amount. Gray level refers to a scalar measure of intensity that ranges from black to grays & finally to white.

* Chromatic light spans the EM spectrum from approximately 400 to 700 nm. Three basic quantities are used to describe the quality of a chromatic light source: radiance, luminance & brightness.

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

Subscribe EC Academy on YouTube

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

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

* Cones are the sensors in the human eye responsible for color vision. Experimental evidence has established that the 6 to 7 million cones in the human eye can be divided into three principal sensing categories namely Red, Green & Blue.

Approximately 65% of all cones are sensitive to red light, 33% are sensitive to green light & only 2% are sensitive to blue.

* Fig. below shows average experimental curves detailing the absorption of light by the R, G & B cones in the eye.

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

Subscribe EC Academy on YouTube

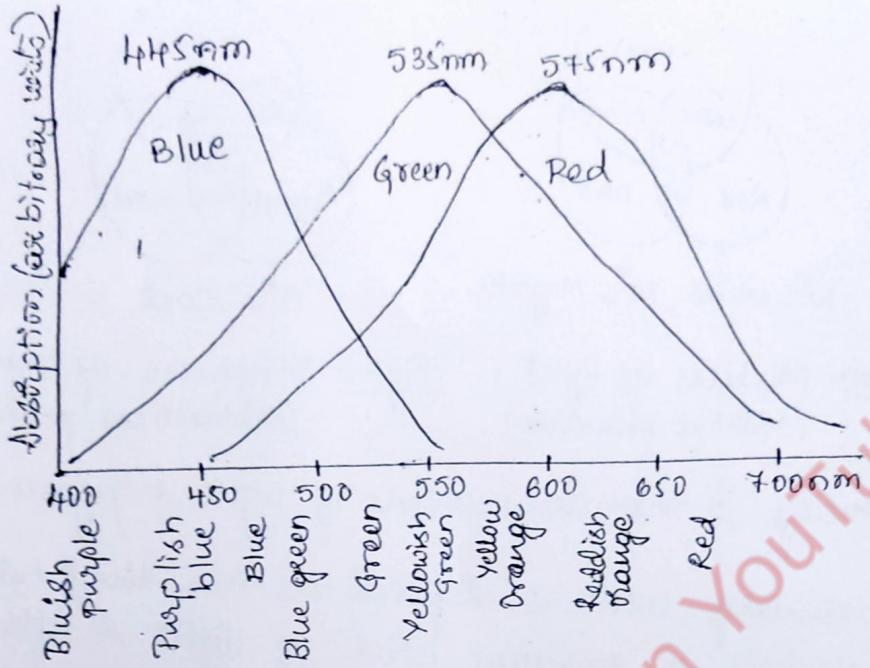


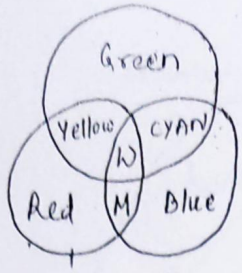
Fig:- Absorption of light by red, green & blue cones in the human eye as a function of wavelength.

* For the propose of standardization, the CIE (Commission Internationale de l'Eclairage - the International Commission on Illumination) designated in 1931, the following specific wavelength values to the three primary colors: blue = 435.8 nm, green = 546.1 nm & red = 700 nm. This std. was set before the detailed experimental curves shown in the above fig. Thus, the CIE std's correspond only approximately with experimental data.

* These three std. primaries, when mixed in various intensity proportions, can produce all visible colors.

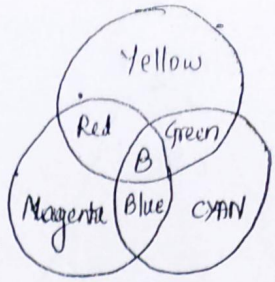
* The primary colors can be added to produce the secondary colors of light - magenta (red + blue) cyan (green + blue) & yellow (red + green). Mixing the 3 primaries or a secondary with its opposite primary color, in the right intensities produces white light. This result is shown in fig. below.

Subscribe EC Academy on YouTube



W^* = White, M^* = Magenta

Fig(a): Mixtures of light (Additive primaries)



B^* = Black

Fig(b): Mixtures of pigments (subtractive primaries)

Fig: Primary & Secondary colors of light & pigments.

* A primary color is defined as one that subtracts or absorbs a primary color of light & reflects or transmits the other two. Therefore, the primary colors of pigments are magenta, cyan & yellow & the secondary colors are R, G & B. A proper combination of 3 pigment primaries or a secondary with its opposite primary, produces black.

* Color television reception is an example of the additive nature of light colors.

The interior of many color TV tubes is composed of a large array of traingular dot patterns of electron-sensitive phosphor. When excited, each dot in a triad is capable of producing light in one of the primary colors.

The effect, viewed on TV receiver is that the 3 primary colors from each phosphor triad are "added" together & received by the color-sensitive cones in the eye as a full-color image. Thirty successive image changes per second in all three colors complete the illusion of a continuous image display on the screen.

Subscribe EC Academy on YouTube

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

Brightness is the chromatic notion of intensity.

Hue is the dominant wavelength in a mixture of light waves. Hue represents dominant color as perceived by an observer. Thus, when we call an object red, orange or yellow, we are specifying its hue.

Saturation refers to relative purity or the amount of white light mixed with a hue. The pure spectrum colors are fully saturated. Colors such as pink (Red + white) & lavender (violet + white) are less saturated, with the degree of saturation being inversely proportional to the amount of white light added.

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

* The amounts of R, G & B needed to form any particular color are called tristimulus values and are denoted by X, Y & Z respectively. A color is then specified by its trichromatic coefficients defined as,

$$x = \frac{X}{X+Y+Z} \quad \text{--- (1)}$$

$$y = \frac{Y}{X+Y+Z} \quad \text{--- (2)}, \quad z = \frac{Z}{X+Y+Z} \quad \text{--- (3)}$$

It is noted from these equations that

$$x + y + z = 1 \quad \text{--- (4)}$$

Subscribe EO Academy on YouTube

* An approach for specifying colors is to use the CIE chromaticity diagram shown in the fig, which shows color composition as a function of x (red) & y (green).

For any value of x & y, the corresponding value of z (blue) is obtained from eqn (4) by noting that $z = 1 - (x+y)$. The point marked green in fig, for example, has approximately 62% green & 25% red content. From eqn (4), the composition of blue is approximately 13%.

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

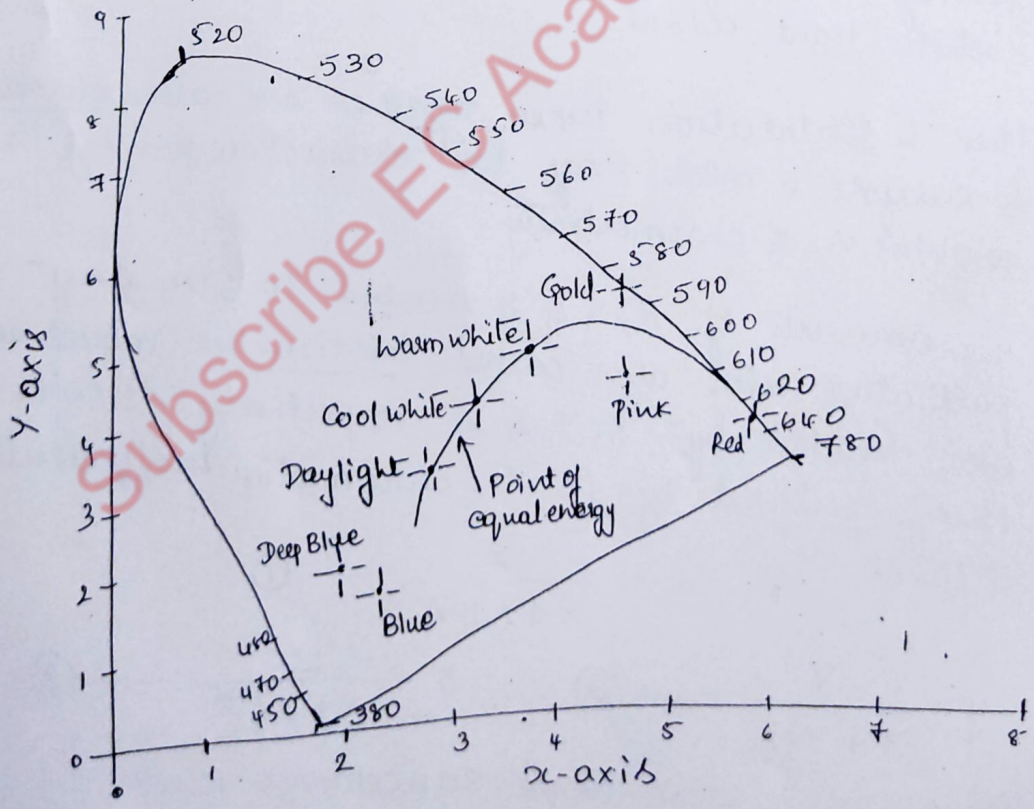


Fig:- C.I.E Chromaticity Diagram.

Subscribe EC Academy on YouTube

* Spectrum shows the pure colors. Any point not actually on the boundary but within the diagram represents some mixture of spectrum colors. 25

The point of equal energy shown in fig corresponds to equal fractions of the three primary colors; it represents the CIE std for white light.

Any point located on the boundary of the chromaticity chart is fully saturated. As a point leaves the boundary & approaches the point of equal energy, more white light is added to the color & it becomes less saturated. Saturation at the point of equal energy is zero.

* The chromaticity diagram is useful for color mixing because a straight-line segment joining any two points in the diagram defines all the different color variations that can be obtained by combining these two colors additively.

Ex:- A St. line drawn from Red to green point is shown in fig. If there is more red light than green light, the exact point representating the new color will be on the line segment, but it will be closer to red point than to the green point. If a line drawn from point of equal energy to any point on the boundary of the chart, define all the shades of that particular spectrum color.

* To determine the range of colors that can be obtained from any three given colors in chromaticity diagram

Subscribe EC Academy on YouTube

we simply draw connecting lines to each of the three color points. The result is a traingle, & any color inside the triangle can be produced by various combinations of three initial colors. A Δ^{le} with vertices at any three fixed colors cannot enclose the entire color region in fig.

* The Δ^{le} shown in fig below shows a typical range of colors (called the color gamut) produced by RGB monitors.

The irregular region inside the Δ^{le} is representative of color gamut of today's high-quality color printing devices. This irregularity is due to + & − of color mixing.

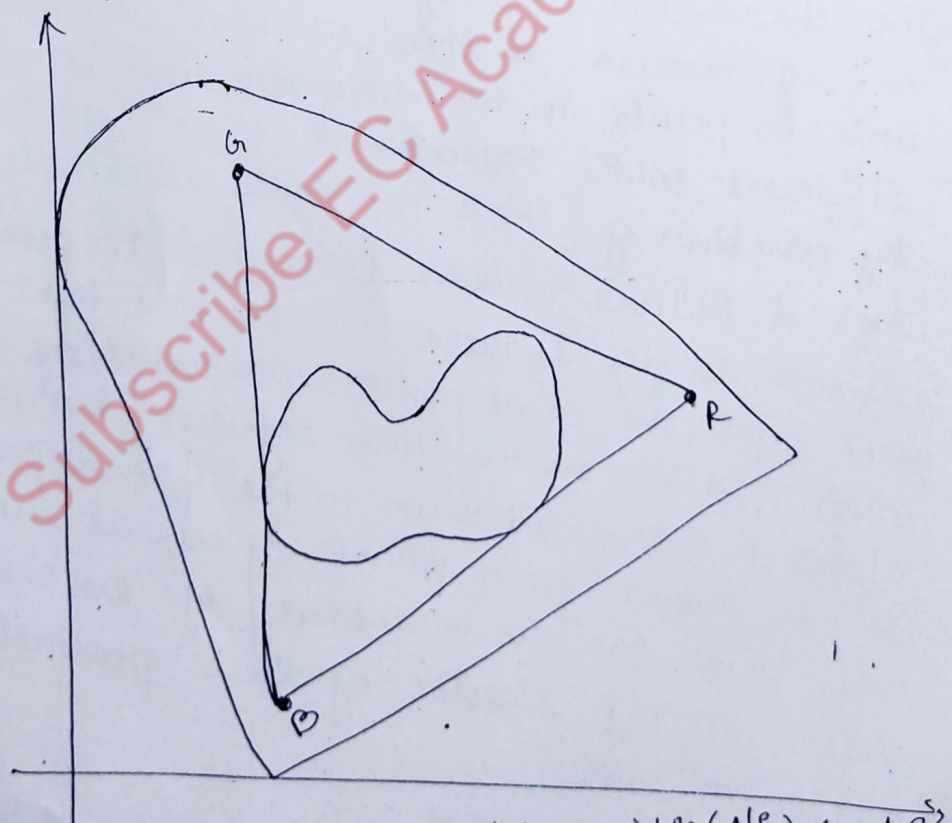


Fig:- Typical color gamut of color monitors (Δ^{le}) & color printing devices (Irregular region).

Subscribe EC Academy on YouTube

4.1.2 Color Models :—

* also called as color space or color system.
* The purpose of a color model is to facilitate the specification of colors in some standard.
* A color model is a specification of a coordinal system & a subspace within that system where each color is represented by a single point.
* Most color models in use today are oriented either toward hardware (such as for color monito & printers) or toward applications where color manipulation is a goal (such as in the creatio of color graphics for animation).
* In terms of digital image processing, the hardware-oriented models most commonly used in practice are
  - → the RGB model for color monitors & color video cameras.
  - → the CMY (cyan, magenta, yellow) & CMYK (cyan, Magenta, Yellow, black) models for color printing &
  - → the HSI (hue, saturation, intensity) model which corresponds closely with the way humans describe & interpret color.

cyan → sky-an
magenta → mah-jen-tah

* The HSI model also has the advantage that it decouples the color & gray-scale information in an image, making it suitable for many of the gray-scale techniques.

Subscribe EC Academy on YouTube

i) The RGB Color Model:-

* In the RGB model, each color appears in its primary spectral components of red, green & blue.

* This model is based on Cartesian coordinate system.

* The color subspace of interest is the cube shown in fig. below.

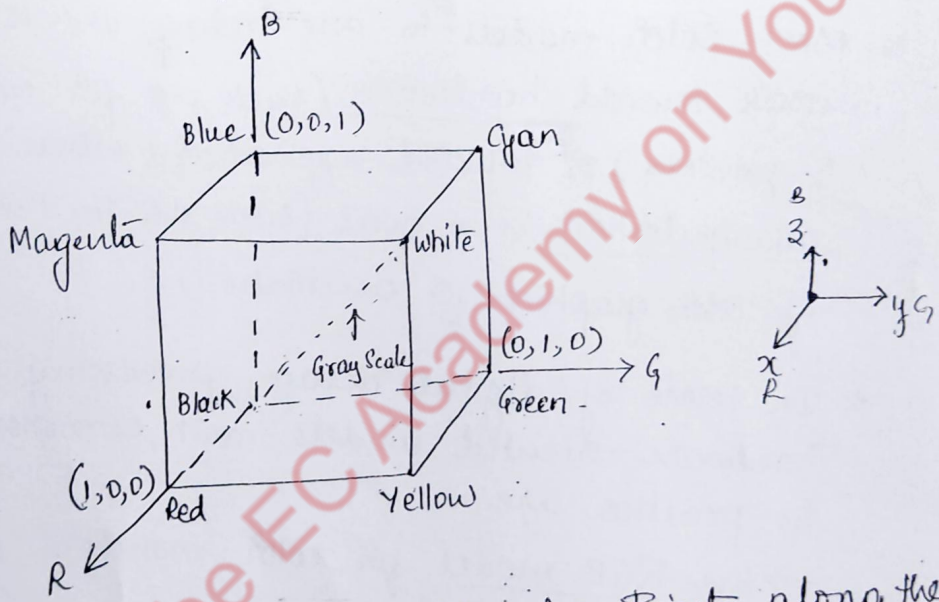


Fig: Schematic of RGB color cube. Points along the main diagonal have gray values, from black at the origin to white at point (1,1,1).

* In the above cube, RGB values are at three corners; cyan, magenta & yellow are at three corners; black is at the origin; & white is at the corner farthest from the origin. The gray scale (points of equal RGB values) extends from black to white along the line joining these two points.

Subscribe EC Academy on YouTube

* The different colors in this model are points on or inside the cube, and are defined by vectors extending from the origin.

* For convenience, the assumption is that all col values have been normalized so that the cut shown in fig is the unit cube. That is, all values of R, G & B are assumed to be in the rang [0, 1].

* Image represented in RGB color model consist of three component images, one for each primary color. When fed into an RGB monitor, these three images combine on the phosphor screen to produce a composite color image. The no. of bits used to represent each pixel in RGB spa is called the pixel depth.

Ex:- Consider an RGB image in which R, G & images is an 8-bit image. Under these conditio each RGB color pixel [i.e., a triplet of values (R, G, is said to have a depth of 24-bits (3 image planes times the no. of bits per plane). The term full-color image is used often to denote a 24-bit RGB color image. The total no. of color in a 24-bit RGB image is $(2^8)^3 = 16,777,216$.

* A convenient way to view these colors is to generate color planes (faces or cross sections of the cube). This is accomplished by fixing one of th three colors & allowing the other two to vary. [illegible] a cross-sectional plane thr' the

Subscribe EC Academy on YouTube

center of the cube & parallel to the GB-plane is the plane (127, G, B) for G, B = 0, 1, 2, ..., 255.

* Fig. below shows that an image of the cross-sectional plane is viewed simply by feeding the three individual component images into a color monitor.

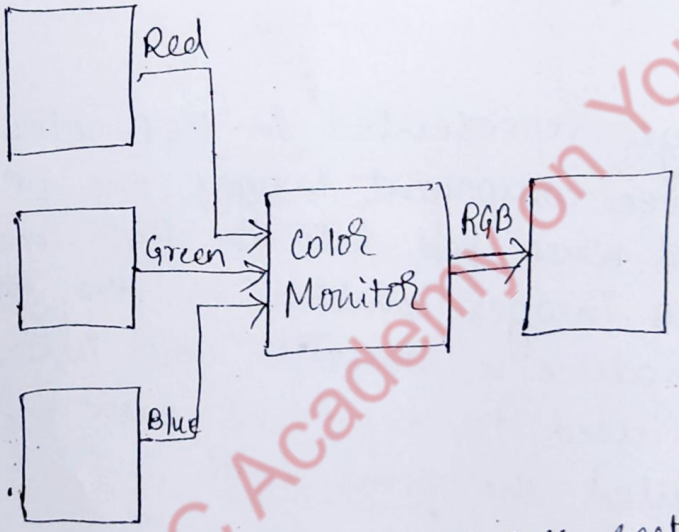


Fig:- Generating RGB image of cross-sectional color plane (127, G, B).

* While high-end display cards & monitors provide a reasonable rendition of the colors in a 24-bit RGB image, many systems in use today are limited to 256 colors.

* Given the variety of systems in current use, it is of considerable interest to have a subset of colors that are likely to be reproduced faithfully, reasonably independent of viewer hardware capabilities. This subset of colors

Subscribe EC Academy on YouTube

is called the set of safe RGB colors or the set of all-systems-safe colors. In Internet applications, th are called safe web colors or safe browser colors.

* On the assumption that 256 colors is the minimu number of colors that can be reproduced faithfu by any system in which a desired result is like to be displayed, it is useful to have an accept Std. notation to refer to these colors.

* Forty of these 256 (2^8) colors are known to be processed differently by various operating syste leaving only 216 colors that are common to mo systems. These 216 colors have become the de facto standard for safe colors, especially in Internet applications. They are used whenever it is desired that the colors viewed by most people appear the same.

* Each of 216 safe colors is formed from three RGB values as before, but each value can only be 0, 51, 102, 153, 204 or 255. Thus, RGB triplets of these values give us $(6)^3 = 216$ possible values.

* It is customary to express these values in the hexagonal no. system, as shown in table below:

No. System	Color Equivalents					
Hex	00	33	66	99	CC	FF
Decimal	0	51	102	153	204	255

Subscribe EC Academy on YouTube

W.K.T the hex numbers 0, 1, 2, ... 9, A, B, C, D, E, F correspond to decimal numbers 0, 1, 2, ..., 9, 10, 11, ..., 15. Also that $(0)_{16} = (0000)_2$ & $(F)_{16} = (1111)_2$. Thus for example, $(FF)_{16} = (255)_{10} = (11111111)_2$.

* Since it takes three numbers to form an RGB color, each safe color is formed from three of the two digit hex numbers shown in table. For example, the purest red is FF 0000. The values 000000 & FFFFFF represent black & white.

The brightest red in decimal notation has Red = 255 (FF) & G = B = 0.

* The 216 safe colors, organized in descending RGB values are shown in fig. 6.10(a) (Reft textbook page 293) and fig 6.10 (b) shows the hexcodes for all the possible gray colors in a 256-color RGB system. The grays from the safe color group, $(KKKKKK)_{16}$ for K = 0, 3, 6, 9, C, F are shown underlined in fig 6.10 (b).

Each plane in fig. 6.10(a) has a total of 36 colors, so the entire surface of the safe-color cube is covered by 216 different colors.

Subscribe EC Academy on YouTube

4.1.3 The CMY & CMYK Color Models:-

* The CMY i.e., CYAN, Magenta & Yellow are the secondary colors of light (or) the primary colors of pigments.

* When a surface coated with CYAN pigment is illuminated with white light, no red light is reflected from the surface. i.e., cyan subtracts red light from reflected white light, which itself is composed of equal amounts of red green & blue light.

$$\text{CYAN} = (\text{green} + \text{blue})$$

* Most devices that deposit colored pigments on paper, such as color printers & copiers, require CMY data i/p or perform an RGB to CMY conversion internally. This conversion is performed using the simple operation

$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad \text{—— (5)}$$

where the assumption is that all color values have been normalized to the range [0, 1]. Eqn (1) demonstrates that light reflected from a surface coated with pure cyan doesnot contain red i.e., $C = 1 - R$ in the eqn.

* Similarly, pure Magenta doesnot reflect green

Subscribe EC Academy on YouTube

& pure yellow doesnot reflect blue.

Magenta (R & Blue)

Yellow (red + green)

31

* Eqn (1) reveals that RGB values can be obtained easily from a set of CMY values by subtracting the individual CMY values from 1.

* In image processing, this color model is used in connection with generating hardcopy output.

* Equal amounts of the pigment primaries, cyan, magenta & yellow should produce black. In practice, combining these colors for printing produces a muddy-looking black. So, in order to produce true black, a fourth color, black, is added, giving rise to the CMYK color model. Thus, when publishers talk about "four-color printing" they are referring to the three colors of the CMY color model plus black.

4.1.4 The HSI Color Model:-

* When humans view a color object, we describe it by its hue, saturation & brightness.

Hue is a color attribute that describes a pure color (pure yellow, orange or red).

Saturation gives a measure of the degree to which a pure color is diluted by

Subscribe EC Academy on YouTube

white light.

Brightness is a subjective descriptor that is practically impossible to measure. It embodies the achromatic notion of intensity & is one of the 35 key factors in describing color sensation.

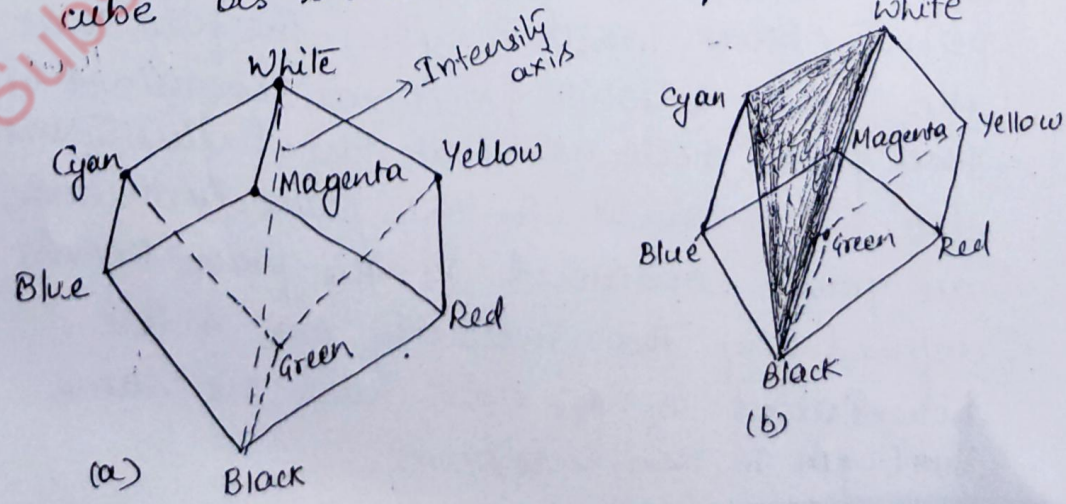
* The HSI color model, decouples the intensity compone from the color-carrying information (hue & saturati in a color image.

HSI model is an ideal tool for developing image processing algorithms based on color descripti that are natural & intuitive to humans.

* RGB is ideal for image color generation but its use for color description is much more limited

* An RGB color image can be viewed as three monochrome intensity images (representing red, green & blue) & so that we should be able to extract intensity from an RGB image.

* This becomes clear if we consider a color cube as shown in the fig(a) below.



Interrelationship between the RGB & HSI color models.

Subscribe EC Academy on YouTube

The intensity (gray scale) is along the line joining the black (0,0,0) vertex with the white vertex (1,1,1) directly and is vertical.

Thus, if we wanted to determine the intensity component of any color point in fig, we would simply pass a plane $\perp^r$ to the intensity axis & containing the color point.

The intersection of the plane with the intensity axis would give us a point with intensity value in the range [0,1]. The saturation (purity) of a color increases as a function of distance from the intensity axis. In fact, the saturation of points on the intensity axis is zero, as evidenced by the fact that all points along this axis are gray.

* In order to see how hue can be determined also from a given RGB point, consider fig(b) which shows a plane defined by three points (black, white & cyan). The fact that the black & white points are contained in the plane tells us that the intensity axis also is contained in the plane. Furthermore, all points contained in the plane segment defined by the intensity axis & the boundaries of the cube have the same hue (cyan in this case).

Subscribe EC Academy on YouTube

* We would arrive at the same conclusion that all colors generated by three colors lie in the triangle defined by those colors. If two of those points are black & white & the third is a color point, all points on the triangle would have the same hue because the black & white components cannot change the hue. By rotating the shaded plane about the vertical intensity axis, we would obtain different hues.

From these concepts we arrive at the conclusion that the hue, saturation & intensity values required to form the HSI space can be obtained from the RGB color cube. i.e., we can convert any RGB point to a corresponding point is the HSI color model by working out the geometric formulas describing the reasoning.

* The HSI color space model as shown in fig (b) is that the HSI space is represented by a vertical intensity axis & the locus of color points that lie on planes $\perp^r$ to this axis. As the planes move up & down the intensity axis, the boundaries defined by the intersection of each plane with the faces of the cube have either a triangular or hexagonal shape. This can be visualized much more readily by looking at the cube down its gray-scale axis, as shown in fig 2(a).

Subscribe EC Academy on YouTube

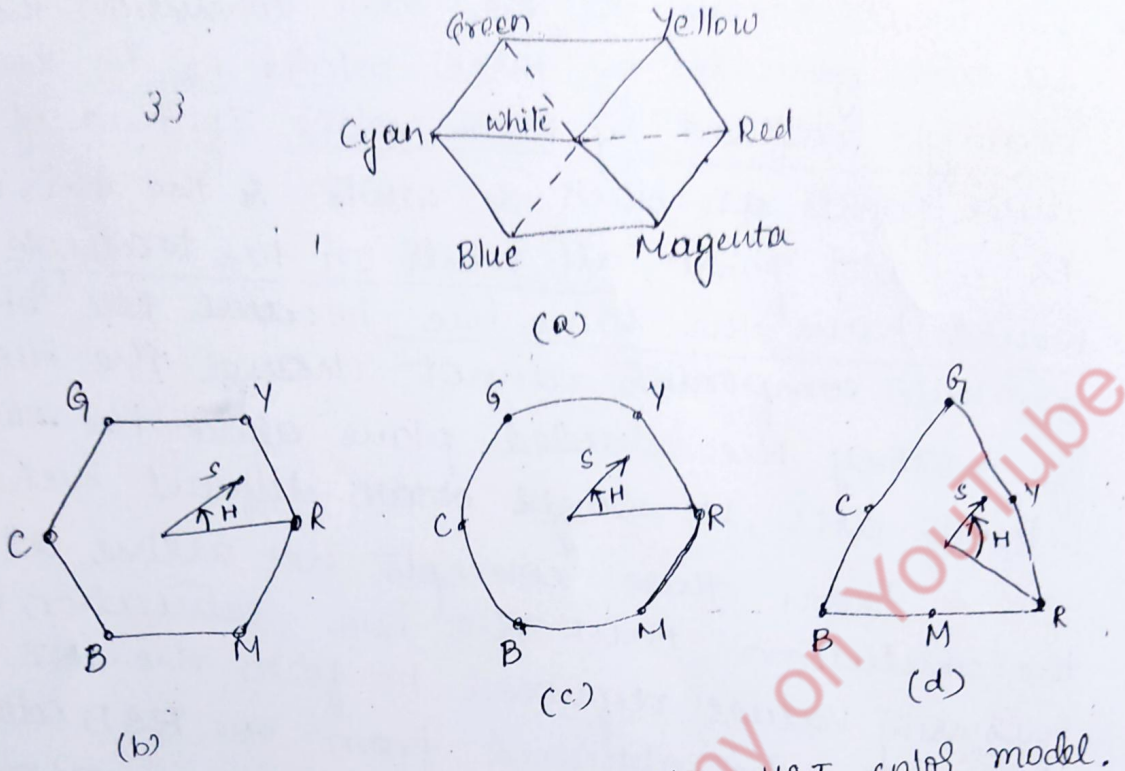


Fig2: Hue & saturation in the HSI color model. The dot is an arbitrary color point. The angle from the red axis gives the hue, & the length of the vector is the saturation. The intensity of all colors in any of these planes is given by the position of the plane on the vertical intensity axis.

In this plane, the primary colors are separated by 120°. The secondary colors are 60° from the primaries, which means that the angle b/w secondaries also is 120°.

Fig(b) shows the same hexagonal shape & an arbitrary color point (shown as a dot). The hue of the point is determined by an angle from some reference point. Usually

Subscribe EC Academy on YouTube

an angle of 0° from the red axis designates 0 hue, & the hue increases counterclockwise fro there. The saturation (distance from the vertical axis) is the length of the vector from the origin to the point. Note that the origin is defined by the intersection of the color plane with the vertical intensity axis.

The important components of HSI color spa are vertical intensity axis, the length of the vecto to a color point & the angle this vector mar with the red axis. Therefore it is not unusual to see the HSI planes defined in terms of hexagon, a traingle or even a circle as show in fig (c) & (d). The shape chosen realy doesnot matter, since any one of these shapes can be warped into one of the other two by a geometric transformatio

Fig. shows the HSI model based on color traingles & also on circles.

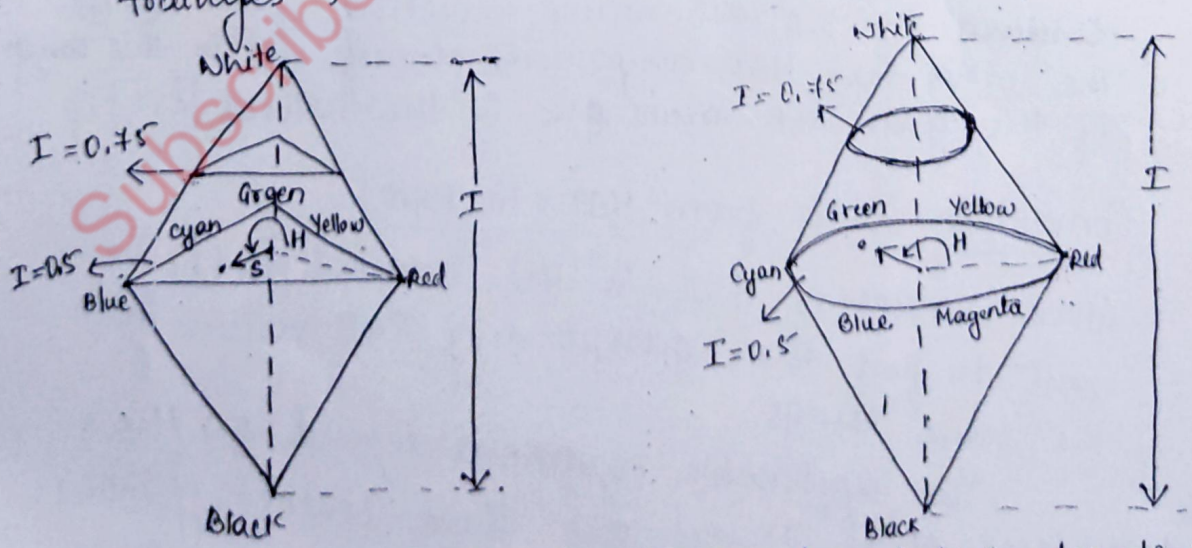


Fig: The HSI color model based on (a) triangular & (b) circular color pla The Δ^{ls} & O^{les} are $\perp^r$ to vertical intensity axis.

Subscribe EC Academy on YouTube

Converting colors from RGB to HSI :-

Given an image in RGB color format, the H component of each RGB pixel is obtained using the eqn,

$$H = \begin{cases} \theta & \text{if } B \le G \\ 360^\circ - \theta & \text{if } B > G \end{cases} \quad \text{——— (6)}$$

with

$$\theta = \cos^{-1}\left\{ \frac{\frac{1}{2}\left[(R-G)+(R-B)\right]}{\left[(R-G)^2+(R-B)(G-B)\right]^{1/2}} \right\}$$

The saturation component is given by,

$$S = 1 - \frac{3}{(R+G+B)}\left[\min(R,G,B)\right] \quad \text{——— (7)}$$

Finally, the intensity component is given by,

$$I = \frac{1}{3}(R+G+B) \quad \text{——— (8)}$$

It is assumed that the RGB values have been normalized to the range [0,1] & that angle θ is measured w.r.to red axis of the HSI space.

Hue can be normalized to the range [0,1] by dividing by 360° all values resulting from eqn (6). The other two HSI components already are in this range if the given RGB values are in the interval [0,1].

Converting colors from HSI to RGB :-

Given values of HSI in the interval [0,1], we want to find the corresponding RGB values in the same range.

The applicable equations depend on the values of H. There are three sectors of interest corresponding to 120° intervals in the separation of

Subscribe EC Academy on YouTube

primaries. We begin by multiplying H by 360°, which returns the hue to its original range of [0°, 360°]

RG sector ($0° \leq H < 120°$):— when H is in this sector, RGB components are given by equations,

$$B = I(1-S) \quad \text{——— (9)}$$

$$R = I\left[1 + \frac{S\cos H}{\cos(60° - H)}\right] \quad \text{——— (10)}$$

$$\& \quad G = 3I - (R+B) \quad \text{——— (11)}$$

GB sector ($120° \leq H < 240°$):— If the given value of H is in this sector, we first subtract 120° from

$$H = H - 120° \quad \text{——— (12)}$$

Then the RGB components are

$$R = I(1-S) \quad \text{——— (13)}$$

$$G = I\left[1 + \frac{S\cos H}{\cos(60° - H)}\right] \quad \text{——— (14)}$$

$$\& \quad B = 3I - (R+G) \quad \text{——— (15)}$$

BR sector ($240° \leq H \leq 360°$):— If H is in this range, we subtract 240° from it:

$$H = H - 240° \quad \text{——— (16)}$$

The RGB components are: $G = I(1-S)$ ——— (17)

$$B = I\left[1 + \frac{S\cos H}{\cos(60° - H)}\right] \quad \text{——— (18)}$$

$$\& \quad R = 3I - (G+B) \quad \text{——— (19)}$$

Subscribe EC Academy on YouTube

Manipulating HSI components images:-

* consider an image composed of primary & secondary RGB colors. The H, S & I components of this image can be obtained by using the eqns (6) thr' eqn (8). The Hue, saturation & intensity images were generated.

* To generate hue of the image, the gray-level values correspond to angles. For example, red corresponds to 0°. Thus the red region is mapped to a black region in hue image.

* The gray levels are scaled to [0, 255] for display which corresponds to saturation.

* The graylevels are average in intensities images.

[Refer to fig. 6.16 [page 302] in textbook.]

* To convert the HSI image back to RGB image, equations (9) thr' eqn's (19) are used.

[Refer to fig 6.17 [page 303] in textbook]

Subscribe EC Academy on YouTube

4.1.5 Pseudocolor Image Processing:-

* Pseudocolor (also called false color) image processi consists of assigning colors to gray values based on a specified criterion.

* The term pseudo or false color is used to differentiate the process of assigning colors to monochrome images from the processes associated with true color images.

* The principal use of pseudocolor is for human visualization & interpretation of gray-scale event in an image or sequence of images. i.e., huma can discern thousands of color shades & intensiti compared to only two dozen or so shades of gray.

4.1.5.1 Intensity Slicing:-

* The technique of intensity (also called density) slicing & color coding is one of the simplest examples of pseudocolor image processing.

* If an image is interpreted as a 3-D function (intensity v/s spatial coordinates), the method can be viewed as one of the placing planes parallel to the coordinate plane of the image; each plane then "slices" the function in the area of intersection.

Subscribe EC Academy on YouTube

Fig. below shows an example of using a plane at $f(x,y) = l_i$ to slice the image function into two levels.

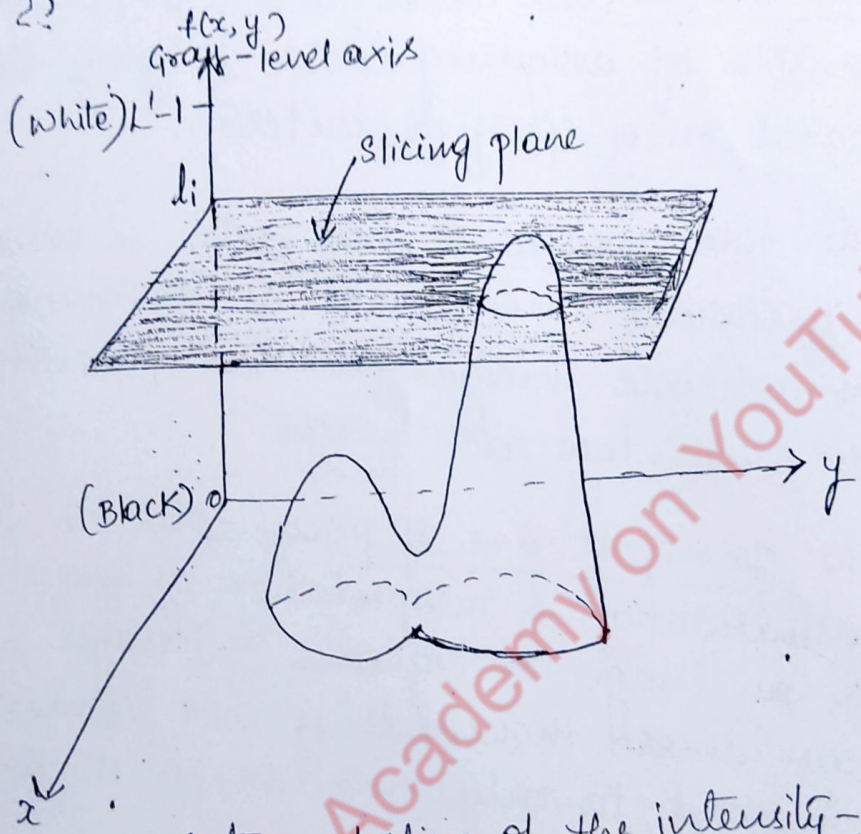


Fig (1). Geometric interpretation of the intensity-slicing technique.

If a different color is assigned to each side of the plane shown in fig, any pixel whose gray level is above the plane will be coded with one color & any pixel below the plane will be coded with the other. Levels that lie on the plane itself may be arbitrarily assigned one of the two colors. The result is a two-color image whose relative appearance can be controlled by moving the slicing plane

Subscribe EC Academy on YouTube

up & down the gray-level axis.

* In general, the technique may be summarized as follows: Let $[0, L-1]$ represents the grayscale let level l_0 represent black $[f(x,y)=0]$ & the level l_{L-1} represent white $[f(x,y)=L-1]$.

Suppose that 'P' planes $\perp^r$ to the intensity axis are defined at levels $l_1, l_2, \ldots, l_P$. Then, assuming that $0 < P < L-1$, the 'P' planes partition the gray scale into $P+1$ intervals, $V_1, V_2, \ldots, V_{P+1}$. Gray-level to color assignments are made according to the relation,

$$f(x,y) = c_K \quad \text{if } f(x,y) \in V_K \quad \text{—— (1)}$$

where C_K is the color associated with the Kth intensity interval V_K defined by the partitioning planes at $l = K-1$ & $l = K$.

The idea of planes is useful primarily for a geometric interpretation of the intensity-slicing technique.

Fig(ii) shows an alternative representation that defines the same mapping as in fig(i).

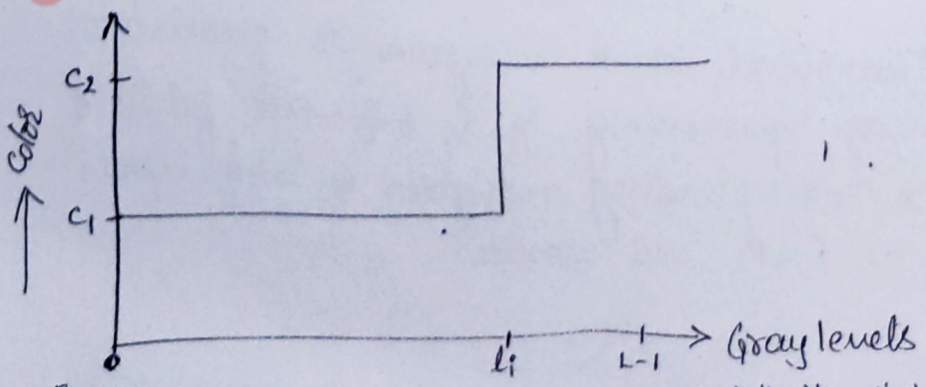


Fig (ii): An alternative representation of intensity-slicing technique

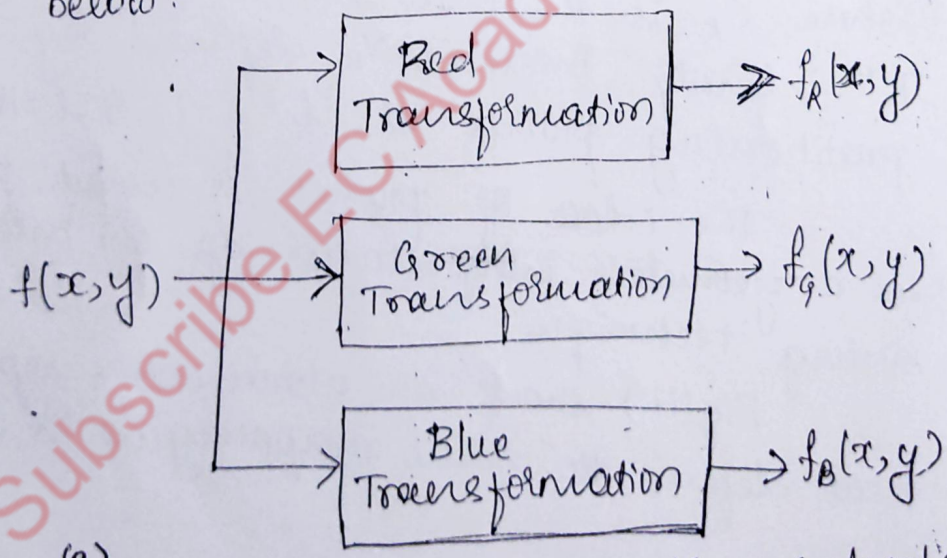
Subscribe EC Academy on YouTube

According to mapping function shown in fig(ii) any i/p graylevel is assigned one of two colors, depending on whether it is above or below the value of l_i. When more levels are used, the mapping function takes on a staircase form.

[Refer to examples given in textbook with images. page 305, 306, 307]

4.1.5.2 Gray level to Color Transformations:-

* The method widely used for Gray level to color transformations are shown in fig. below.



(a)

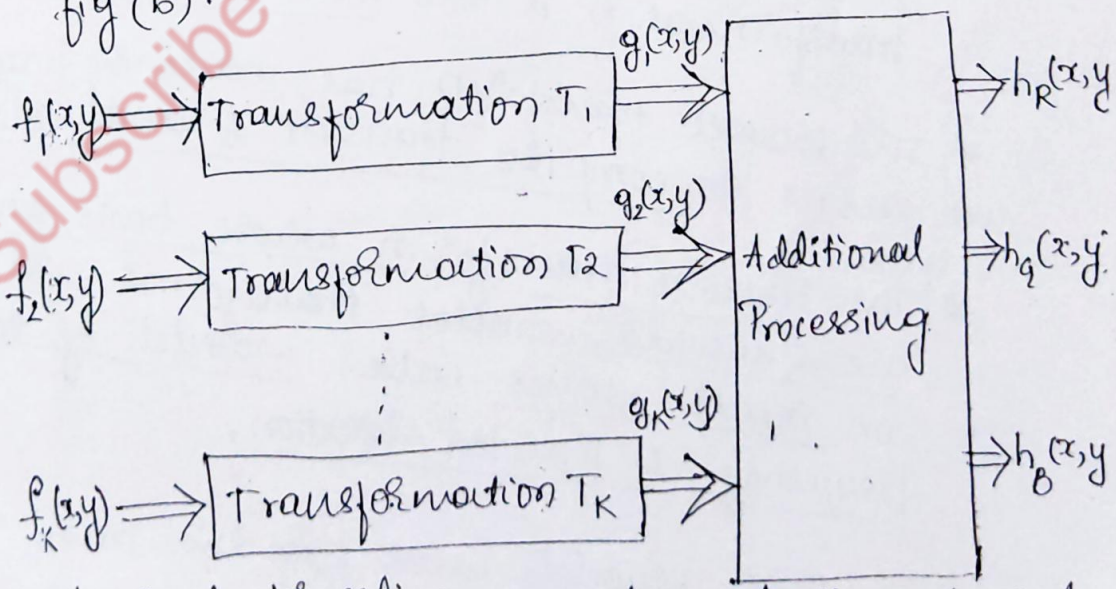
Fig: Functional block diagram for pseudocolor image processing. f_R, f_G & f_B are fed into the corresponding red, green & blue inputs of an RGB color monitor.

Subscribe EC Academy on YouTube

* Basically, the idea underlying this approach is to perform three independent transformations on the gray-level of any input pixel. The 3 results are then fed separately into the red, green & blue channels of a color television monitor. This method produces a composite image whose color content is modulated by the nature of the transformation functions. Note that these are transformations on the gray-level values of an image & are not the functions of position.

[Refer to fig^s. 6.24 & 6.25 in textbook]

* Fig(a) is the approach based on a single monochrome image. But often, it is of interest to combine several monochrome images into a single color composite, as shown in fig (b).



Fig(b): A pseudocolor coding approach used when several monochrome images are available.

Subscribe EC Academy on YouTube

* A frequent use of this approach is in multispectral image processing, where different sensors produce individual monochrome images, each in a different spectral band. The types of additional processes shown in fig (b) can be techniques such as color balancing, combining images & selecting the three images for display based on knowledge about response characteristics of the sensors used to generate the images.

4.2 WAVELETS:

* Although the Fourier transform has been the mainstay of transform-based image processing since 1950s, a more recent transformation is the wavelet transform.

* The wavelet transform has made the images even easier to compress, transmit & analyze.

* Unlike The Fourier transform whose basic functions are sinusoids, wavelet transforms are based on small waves called wavelets of varying frequency & limited duration.

* In 1987, wavelets were first shown to be the foundation of a powerful new approach to

Subscribe EC Academy on YouTube

signal processing & analysis called multiresolution theory.

* Multiresolution theory incorporates & unifies techniques from a variety of disciplines, including subband coding from signal processing, quadrature mirror filtering from digital speech recognition & pyramidal image processing.

* As its name implies, multiresolution theory is concerned with the representation & analysis of signals (or images) at more than one resolution.

4.2.1 Background:-

* When we look at images, generally we see connected regions of similar texture & grey level that combine to form objects. If the objects are small in size or low in contrast, we normally examine them at high resolutions. If both small & large objects - or low & high contrast objects - are present simultaneously, it can be advantageous to study them at several resolutions. This is the fundamental motivation for multiresolution processing.

4.2.1a Image Pyramids:-

* A powerful but conceptually simple structure for representing images at more than one

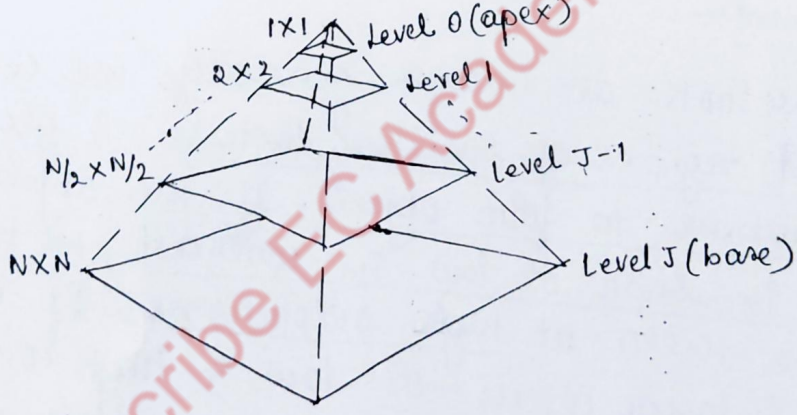
Subscribe FC Academy on YouTube

resolution is the image pyramid.

* An image pyramid is a collection of decreasing resolution images arranged in the shape of a pyramid.

* It is originally devised for machine vision & image compression applications.

* Fig(a) below shows a pyramidal image structure where the base of the pyramid contains a high-resolution representation of the image being processed, & the apex contains a low-resolution approximation. So as we move up both size & resolution decrease.



Fig(a): A pyramidal image structure

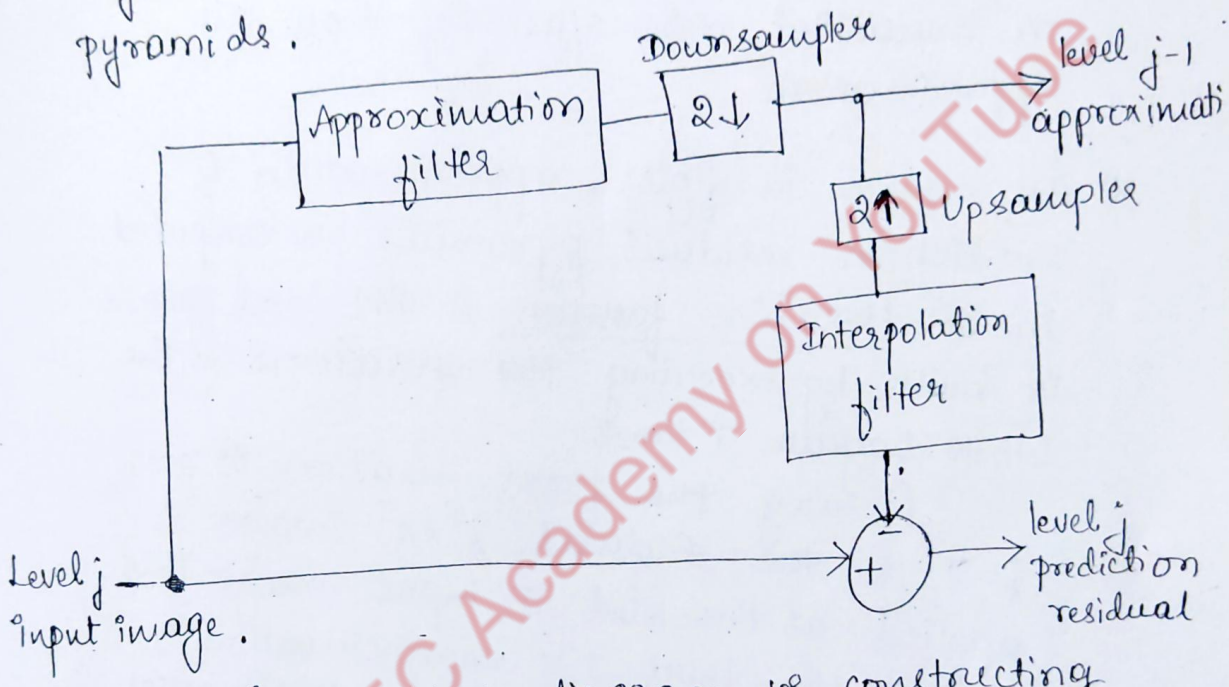
* The base level J is size $2^J \times 2^J$ or NXN, where $J = \log_2 N$, intermediate level j is size $2^j \times 2^j$, where $0 \le j \le J$. Fully populated pyramids are composed of J+1 resolution levels from $2^J \times 2^J$ to $2^0 \times 2^0$, but most pyramids are truncated to P+1 levels, where $j = J-P, \ldots, J-2, J-1, J$ & $1 \le P \le J$.

Subscribe EC Academy on YouTube

* The total number of elements in a (P+1) level pyramid for P>0 is

$$N^2\left(1+\frac{1}{(4)^1}+\frac{1}{(4)^2}+\cdots+\frac{1}{(4)^P}\right) \leq \frac{4}{3}N^2$$

* Fig(b) shows a simple system for constructing image pyramids.



Fig(b): System block diagram for constructing image pyramids.

* The level j-1 approximation output is used to create approximation pyramids which contain one or more approximations of the original image.

* The level j prediction residual output of fig(b) is used to build prediction residual pyramids. These pyramids contain a low-resolution approximation of the original image at level J-P & information for the construction of P higher-resolution approximations at the other levels.

Subscribe EC Academy on YouTube

* The information at level j is the difference between the level j approximation of the corresponding approximation pyramid & an estimate of that approximation based on the level $j-1$ prediction residual. This difference can be coded & therefore stored & transmitted more efficiently than the approximation.

* As shown in fig(b), approximation & prediction residual pyramids are computed in an iterative fashion. A P+1 level pyramid is built by executing the operations in the block diagram P times.

During the first iteration or pass, $j = J$ & the original $2^J \times 2^J$ image is applied as the level J input image. This produces the level J−1 approximation & level J prediction residual results. For passes $j = J-1, J-2, \ldots, J-P+1$, the previous iteration's level $j-1$ approximation output is used as the input. Each pass is composed of three sequential steps:

1) Compute a reduced-resolution approximation of the i/p image. This is done by filtering input & downsampling the filtered result by a factor of 2. A variety of filtering operations can be used, including neighbourhood averaging, which produces

Subscribe EC Academy on YouTube

a Gaussian pyramid, or no filtering, which results in a subsampling pyramid.

2) Upsample the output of the previous step - again by a factor of 2 - & filter the result. This creates a prediction image with the same resolution as the input. By interpolating intensities b/w the pixels of the step 1 o/p, the interpolation filter determines how accurately the prediction approximates the input to step 1. If the interpolation filter is omitted, the prediction is an unsampled version of step 1 o/p & the blocking effects of pixel replication may become visible.

3) Compute the difference between the prediction of step 2 & the input to step 1. This difference labeled the level j prediction residual, can be later used to reconstruct progressively the original image. In the absence of quantization error, a prediction residual pyramid can be used to generate the corresponding approximation pyramid, including the original image, without error.

Executing this procedure 'P' times produces two intimately related P+1 level approximation & prediction residual pyramids. The level j-1 approximation outputs are used to populate

Subscribe EC Academy on YouTube

the approximate pyramid; the level of prediction residual outputs are placed in the prediction residual pyramid. If a prediction residual pyramid is not needed, steps 2 & 3 together with the upsampler, interpolation filter, & summer of fig(b), can be omitted.

[Refer to fig. 7.3 in textbook]

4.2.1b Subband coding:-

* Another important imaging technique which ties to multiresolution analysis is subbandcoding.

* In subband coding, an image is decomposed into a set of band-limited components, called subbands, which can be reassembled to reconstruct the original image without error.

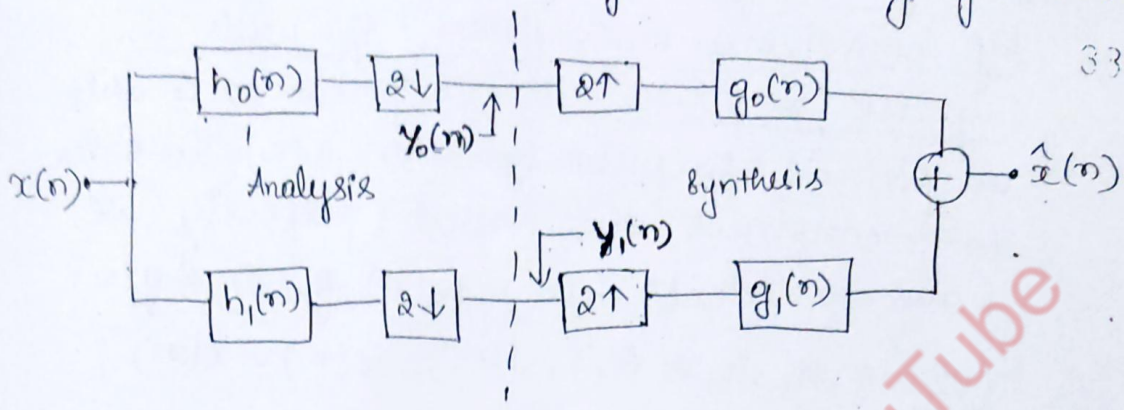
* This technique is used for speech & image compression, where each subband is generated by bandpass filtering the i/p.

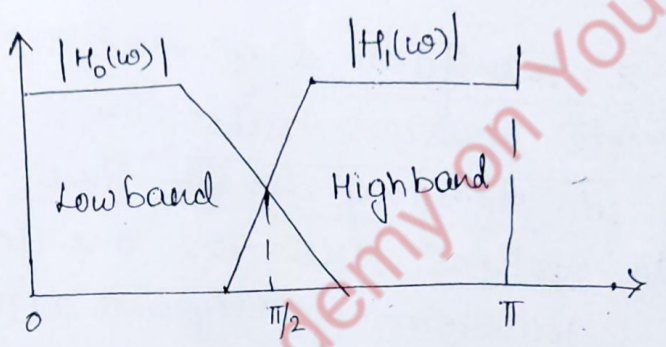
* Since the BW of resulting subbands is smaller than that of original image, the subbands can be downsampled without loss of information.

* Reconstruction of original image is accomplished by upsampling, filtering & summing the individual subbands.

Subscribe EC Academy on YouTube

* Fig (c) below shows the principal components of a two-band subband coding & decoding system.





Fig(d): its spectrum splitting properties.

* The i/p to system is a one-dimensional, band-limited discrete-time signal $x(n)$ for $n = 0, 1, 2, \ldots$; The o/p sequence $\hat{x}(n)$ is formed through the decomposition of $x(n)$ into $y_0(n)$ & $y_1(n)$ via analysis filters $h_0(n)$ & $h_1(n)$, & subsequent recombination via synthesis filters $g_0(n)$ & $g_1(n)$.

* Filters $h_0(n)$ & $h_1(n)$ are half-band digital filters whose idealized transfer chr's H_0 & H_1 are shown in fig(d). Filter H_0 is a low-pass filter whose o/p is an approximation of $x(n)$; filter H_1 is a high pass filter whose o/p is the high-frequency

Subscribe EC Academy on YouTube

or detail part of $x(n)$.

* All filtering is performed in the time domain by convolving each filter's i/p with its impulse response - its response to a unit amplitude impulse function $\delta(n)$. In order to reconstruct the input perfectly, we have to select $h_0(n)$, $h_1(n)$, $g_0(n)$ & $g_1(n)$ (or H_0, H_1, G_0 & G_1), i.e., $\hat{x}(n) = x(n)$.

* The Z-transform, a generalization of the discrete Fourier transform, is the ideal tool for studying discrete-time, sampled-data systems like the one in fig(c). The Z-transform of sequence $x(n)$ for $n = 0, 1, 2, \ldots$ is

$$X(z) = \sum_{-\infty}^{\infty} x(n) z^{-n} \quad \text{— (1)}$$

where z is a complex variable.

* Downsampling by a factor of 2 in the time domain corresponds to the simple Z-domain operation,

$$x_{down}(n) = x(2n) \Longleftrightarrow X_{down}(z) = \frac{1}{2}\left[X(z^{1/2}) + X(-z^{1/2})\right] \quad \text{— (2)}$$

* Upsampling again by a factor of 2, is defined by the transform pair

$$x^{up}(n) = \begin{cases} x(n/2) & n = 0, 2, 4, \ldots \\ 0 & \text{otherwise} \end{cases} \Longleftrightarrow X^{up}(z) = X(z^2) \quad \text{— (3)}$$

Subscribe EC Academy on YouTube

* If sequence $x(n)$ is downsampled & subsequently upsampled to yield $\hat{x}(n)$, eqn's (2) & (3) combine to yield

$$\hat{X}(z) = \frac{1}{2}\left[X(z) + X(-z)\right] \quad \text{—(4)}$$

where $\hat{x}(n) = z^{-1}\left[\hat{X}(z)\right]$ is the resulting downsampled-upsampled sequence. The $X(-z)$ term in this equation is the z-transform of an aliased or modulated version of sequence $x(n)$. Its inverse z-transform is

$$z^{-1}\left[X(-z)\right] = (-1)^n x(n) \quad \text{—(5)}$$

* In accordance with eq (4), we can express the system's o/p as,

$$\hat{X}(z) = \frac{1}{2} G_0(z)\left[H_0(z)X(z) + H_0(-z)X(-z)\right] + \frac{1}{2} G_1(z)\left[H_1(z)X(z) + H_1(-z)X(-z)\right] \quad \text{—(6)}$$

where, for example, the output of filter $h_0(n)$ in fig(c) is defined by the transform pair,

$$h_0(n) * x(n) = \sum_K h_0(n-K)\,x(K) \Leftrightarrow H_0(z)X(z) \quad \text{—(7)}$$

* As with Fourier transforms, convolution in the time domain (or spatial) domain is equivalent to multiplication in the z-domain. Rearranging terms in eqn (6), we then get,

$$\hat{X}(z) = \frac{1}{2}\left[H_0(z)G_0(z) + H_1(z)G_1(z)\right]X(z) + \frac{1}{2}\left[H_0(-z)G_0(z) + H_1(-z)G_1(z)\right]X(-z) \quad \text{—(8)}$$

Subscribe EC Academy on YouTube

* For error-free reconstruction of the input, $\hat{x}(n) = x(n)$ & $\hat{X}(z) = X(z)$. Thus, we impose the following conditions.

$$H_0(-z)G_0(z) + H_1(-z)G_1(z) = 0 \quad \text{--- (9)}$$

$$H_0(z)G_0(z) + H_1(z)G_1(z) = 2 \quad \text{--- (10)}$$

* Eqn (9) eliminates aliasing by forcing the second term of eqn (8) to zero; Eqn (10) eliminates amplitude distortion by reducing the first term to $X(z)$. Both can be incorporated into the single matrix expression,

$$\begin{bmatrix} G_0(z) & G_1(z) \end{bmatrix} H_m(z) = \begin{bmatrix} 2 & 0 \end{bmatrix} \quad \text{--- (11)}$$

where analysis modulation matrix $H_m(z)$ is

$$H_m(z) = \begin{bmatrix} H_0(z) & H_0(-z) \\ H_1(z) & H_1(-z) \end{bmatrix} \quad \text{--- (12)}$$

Assuming $H_m(z)$ is nonsingular i.e., it has a common left & right inverse), we can transpose eqn (11) & left multiply by inverse $(H_m^T(z))^{-1}$ to get,

$$\begin{bmatrix} G_0(z) \\ G_1(z) \end{bmatrix} = \frac{2}{\det(H_m(z))} \begin{bmatrix} H_1(-z) \\ -H_0(-z) \end{bmatrix} \quad \text{--- (13)}$$

where $\det(H_m(z))$ denotes the determinate of $H_m(z)$.

In the above eqn (13), $G_1(z)$ is a function of $H_0(-z)$, while $G_0(z)$ is a

Subscribe EC Academy on YouTube

function of $H_1(-z)$. The analysis & synthesis filters are cross-modulated.

For FIR filters, the determinate of this modulation matrix is a pure delay i.e.,
$\det(H_m(z)) = \alpha z^{-(2K+1)}$. Thus the exact form of the cross-modulation is a function of α. The term $z^{-(2K+1)}$ changes for only overall delay of the filter. Ignoring the delay, letting $\alpha = 2$ & taking the inverse Z-transform of eqn (13), we get-

$$\left.\begin{aligned} g_0(n) &= (-1)^n h_1(n) \\ g_1(n) &= (-1)^{n+1} h_0(n) \end{aligned}\right\} \quad \text{--- (14)}$$

If $\alpha = -2$, the resulting expressions are sign reversed:

$$\left.\begin{aligned} g_0(n) &= (-1)^{n+1} h_1(n) \\ g_1(n) &= (-1)^n h_0(n) \end{aligned}\right\} \quad \text{--- (15)}$$

Thus, FIR synthesis filters are cross-modulated copies of the analysis filters with one & only one being sign reversed.

Eqns (9) thru' (13) can also be used to demonstrate the biorthogonality of the analysis & synthesis filters. To do this, let $P(z)$ be the product of the lowpass analysis & synthesis filter transfer functions. Substituting for G_0 from eqn (13), we get-

$$P(z) = G_0(z) H_0(z) = \frac{2}{\det(H_m(z))} \cdot H_0(z) H_1(-z) \quad \text{--- (16)}$$

Since $\det(H_m(z)) = -\det(H_m(-z))$, product $G_1(z) H_1(z)$ can be similarly defined as,

Subscribe EC Academy on YouTube

$$G_1(z)H_1(z) = \frac{-2}{\det(H_m(z))} H_0(-z)H_1(z) = P(-z) \quad \text{--- (17)}$$

Thus, $G_1(z)H_1(z) = P(-z) = G_0(-z)H_0(-z)$ & eqn (10) becomes $G_0(z)H_0(z) + G_0(-z)H_0(-z) = 2$ --- (18)

Taking inverse Z-transform,

$$\sum_K g_0(K)\, h_0(n-K) + (-1)^n \sum_K g_0(K)\, h_0(n-K) = 2\delta(n)$$

where $\delta(n) = 1$ if $n=0$ & 0 otherwise. Since the odd indexed terms cancel, additional simplification yields,

$$\sum_K g_0(K)\, h_0(2n-K) = \langle g_0(K), h_0(2n-K)\rangle = \delta(n) \quad \text{--- (19)}$$

By starting from eqns (9) & (10) and expressing G_0 & H_0 as a function of G_1 & H_1, we can also show that,

$$\langle g_1(K), h_1(2n-K)\rangle = \delta(n) \quad \text{--- (20)}$$

$$\langle g_0(K), h_1(2n-K)\rangle = 0$$

$$\& \ \langle g_1(K), h_0(2n-K)\rangle = 0$$

which in combination with eqn (19), establishes the general expression,

$$\langle h_i(2n-K), g_j(K)\rangle = \delta(i-j)\,\delta(n), \quad i,j = \{0,1\} \quad \text{--- (21)}$$

Filter banks satisfying this condition are called <u>biorthogonal</u>.

Subscribe EC Academy on YouTube

Table below gives three general solutions to eqns (9) & (10). While each satisfies the biorthogonality requirement of eqn (21), they are generated in different ways & define unique classes of perfect reconstruction filters.

Filter	QMF	CQF	Orthogonal
$H_0(z)$	$H_0^2(z) - H_0^2(-z) = 2$	$H_0(z)H_0(z^{-1}) + H_0(-z)H_0(-z^{-1}) = 2$	$G_0(z^{-1})$
$H_1(z)$	$H_0(-z)$	$z^{-1}H_0(-z^{-1})$	$G_1(z^{-1})$
$G_0(z)$	$H_0(z)$	$H_0(z^{-1})$	$G_0(z)G_0(z^{-1}) + G_0(-z)G_0(-z^{-1}) = 2$
$G_1(z)$	$-H_0(-z)$	$zH_0(-z)$	$-z^{-2k+1}G_0(-z^{-1})$

Table: Perfect reconstruction filter families.

Table shows the classic results from the filter bank literature – namely, quadrature mirror filters (QMFs) & Conjugate quadrature filters (CQFs). The filters in column three, which are later used in the development of the fast wavelet transform are called orthonormal. They move a step beyond biorthogonality & require

$$\langle g_i(n), g_j(n+2m)\rangle = \delta(i-j)\,\delta(m), \quad i,j = \{0,1\} \qquad (22)$$

which defines orthonormality for perfect reconstruction filter banks.

Taking the inverse z-transform of the appropriate entries from column 3 of Table, we get that

Subscribe EC Academy on YouTube

$$g_1(n) = (-1)^n g_0(2K - 1 - n)$$
$$h_i(n) = g_i(2K - 1 - n), \; i = \{0, 1\} \quad \text{—— (23)}$$

where h_0, h_1, g_0 & g_1 are the impulse responses of the defined orthonormal filters.

The one-dimensional filters in table can be used as two-dimensional separable filters for the processing of images.

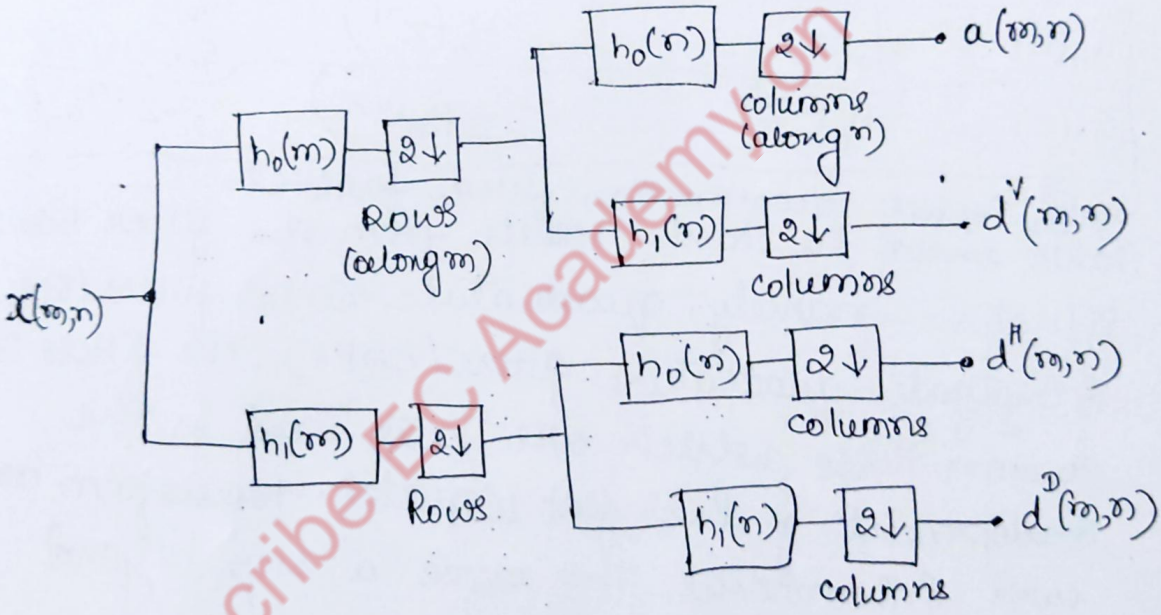


Fig: A two-dimensional four-band filter bank for subband image coding.

As can be seen in fig., separable filters are first applied in one dimension (eg: horizontally) & then in the other (eg., vertically).

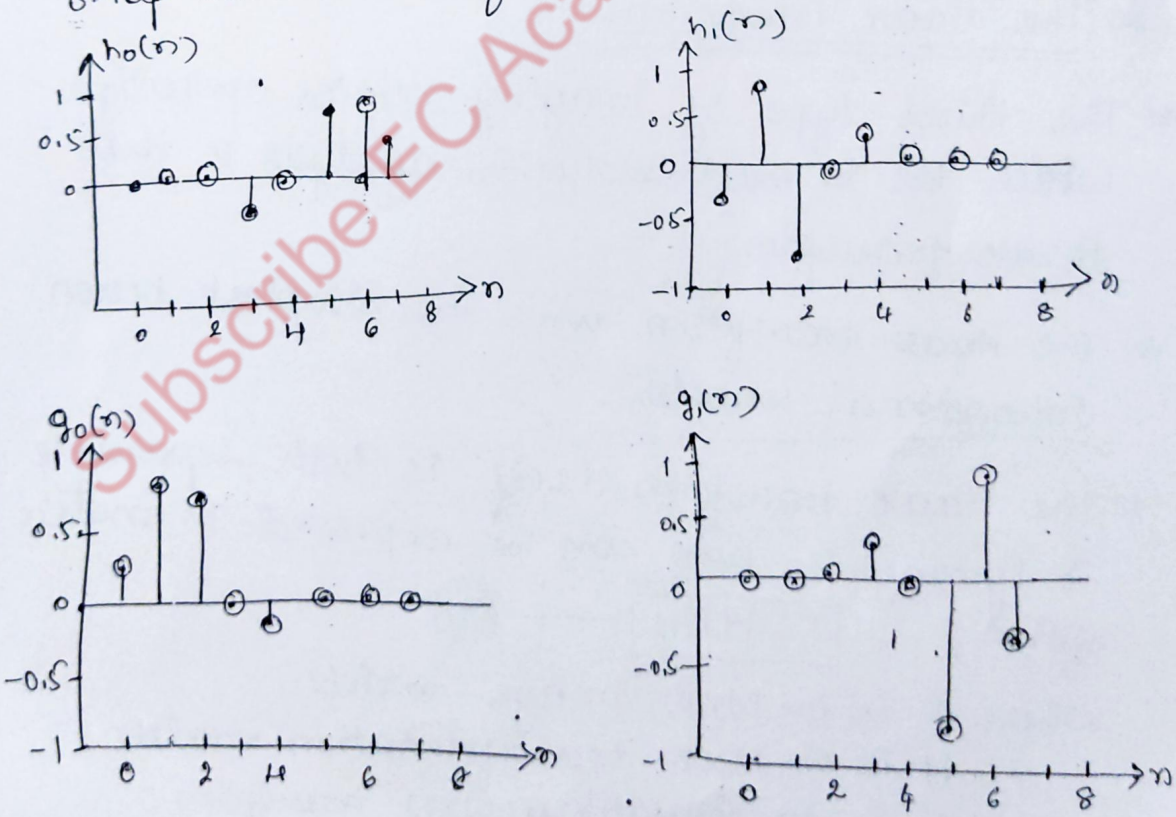
Downsampling is performed in two stages - once before the second filtering operation to reduce the overall no of computations. The resulting filtered

Subscribe EC Academy on YouTube

o/p's denoted as $a(m,n)$, $d^V(m,n)$, $d^H(m,n)$ & $d^D(m,n)$ are called the approximation, vertical detail, horizontal detail & diagonal detail subbands of the image, respectively. One or more of these subbands can be split into four smaller subbands, which can be split again & so on.

Note:- h_0, h_1, g_0 & g_1 are the impulse responses of the defined orthonormal filters. Examples include the Smith & Barnwell filter, Daubechies filters, & the Vaidyanathan & Hoang filter.

* Fig. below shows the impulse responses of four 8-tap orthonormal filters.



Subscribe EC Academy on YouTube

The coefficients of lowpass filter $h_0(n)$ for $0 \le n \le 7$ are -0.01059740, 0.03288301, 0.03084138, -0.18703481, -0.02798376, 0.63088076, 0.71484657 & 0.23037781 (Daubechies filter).

The coefficients of the remaining orthonormal filters can be computed using eqn (23).

The graphs shows numerically that the filters are both biorthogonal (satisfy eqn (21)) & orthogonal (satisfy eqn 22). They also fulfill eqns (9) & (10) and support error-free reconstruction of the decomposed input.

[Refer to fig:7.7 of page no. 361]

4.2. The Haar Transform:-

* The third type of imaging-related operation which ties to multiresolution analysis is the Haar transform.

* The Haar transform uses the simplest known orthonormal wavelets.

* The Haar transform itself is both separable & symmetric and can be expressed in matrix form,

$$\boxed{T = HFH} \quad \text{—— (24)}$$

where F is an NXN image matrix
H is an NXN transformation matrix
T is the resulting NXN transform

Subscribe EC Academy on YouTube

* For Haar transform, transformation matrix H contains the Haar basis functions, $h_K(z)$. They are defined over the continuous, closed interval $z \in [0,1]$ for $K = 0, 1, 2, \ldots, N-1$ where $\boxed{N = 2^n}$. To generate H, we define the integer K such that $K = 2^p + q - 1$, where $0 \le p \le n-1$, $q = 0$ or 1 for $p = 0$ & $1 \le q \le 2^p$ for $p \ne 0$. Then the Haar basis functions are

$$h_0(z) = h_{00}(z) = \frac{1}{\sqrt{N}}, \quad z \in [0,1] \quad \text{— (25)}$$

&

$$h_K(z) = h_{pq}(z) = \frac{1}{\sqrt{N}} \begin{cases} 2^{p/2} & (q-1)/2^p \le z < (q-0.5)/2^p \\ -2^{p/2} & (q-0.5)/2^p \le z < q/2^p \\ 0 & \text{otherwise}, z \in [0,1] \end{cases} \quad \text{— (26)}$$

* The ith row of an $N \times N$ Haar transformation contains the elements of $h_i(z)$ for $z = 0/N, 1/N, 2/N, \ldots, (N-1)/N$. If $N = 4$, for example, K, q & p assume the values

K	p	q
0	0	0
1	0	1
2	1	1
3	1	2

& the 4×4 transformation matrix H_4 is

$$H_4 = \frac{1}{\sqrt{4}} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ \sqrt{2} & -\sqrt{2} & 0 & 0 \\ 0 & 0 & \sqrt{2} & -\sqrt{2} \end{bmatrix} \quad \text{— (27)}$$

Subscribe EC Academy on YouTube

In a similar manner, the 2x2 transformation matrix, H_2 is

$$H_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad \text{—— (28)}$$

Its basis functions define the only two-tap FIR filter bank that will satisfy the QMF prototype filter specification in row 1 & column 1 of Table. The coefficients of corresponding QMF analysis filters $h_0(n)$ & $h_1(n)$ are the elements of the first & second rows of matrix H_2, respectively.

[Refer to fig: 7.8 in textbook page. no: 363]

4.2.2 Multiresolution Expansions:—

* Thus the three well-known imaging techniques namely Image Pyramids, sub-band coding and the Haar transformation played an important role in the development of a unique mathematical theory called multiresolution analysis (MRA).

* In MRA, a scaling function is used to create a series of approximations of a function or image, each differing by a factor of 2 from its nearest neighboring approximations. Additional functions, called wavelets are then used to encode the difference in information b/w

Subscribe EC Academy on YouTube

adjacent approximations.

① Find 2×2 Haar matrix H_2.

Soln:- Since 2×2 transformation matrix,

here $\boxed{N=2} \Rightarrow N=2^n$

$2=2^n \Rightarrow \boxed{n=1}$

$$H_2 = \begin{array}{c} \\ K=0 \\ K=1 \end{array} \begin{array}{c} \begin{array}{cc} z=\frac{0}{2} & z=\frac{1}{2} \end{array} \\ \begin{bmatrix} h_0\left(\frac{0}{2}\right) & h_0\left(\frac{1}{2}\right) \\ h_1\left(\frac{0}{2}\right) & h_1\left(\frac{1}{2}\right) \end{bmatrix} \end{array}$$

Since $z = \frac{0}{N}, \frac{1}{N}, \frac{2}{N}, \ldots, \frac{N-1}{N}$

& $K = 0, 1, 2, \ldots, N-1$

$K = 2^p + q - 1$

K	p	q
0	0	0
1	0	1
2	1	1
3	1	2

$$H_2 = \begin{array}{c} \\ K=0 \\ K=1 \end{array} \begin{array}{c} \begin{array}{cc} z=\frac{0}{2} & z=\frac{1}{2} \end{array} \\ \begin{bmatrix} h_0\left(\frac{0}{2}\right) & h_0\left(\frac{1}{2}\right) \\ h_1\left(\frac{0}{2}\right) & h_1\left(\frac{1}{2}\right) \end{bmatrix} \end{array}$$

$h_0\left(\frac{0}{2}\right) = \frac{1}{\sqrt{N}} = \frac{1}{\sqrt{2}}$, $h_0\left(\frac{1}{2}\right) = \frac{1}{\sqrt{2}}$

$\frac{q-1}{2^p} \le z < \frac{q-\frac{1}{2}}{2^p} \Rightarrow \frac{1-1}{2^0} \le z < \frac{1-\frac{1}{2}}{2^0} \Rightarrow \frac{0}{1} \le z \le \frac{1/2}{1} \Rightarrow 0 \le z \le \frac{1}{2}$

$\frac{q-\frac{1}{2}}{2^p} \le z < \frac{q}{2^p} \Rightarrow \frac{1-\frac{1}{2}}{2^0} \le z \le \frac{1}{2^0} \Rightarrow \frac{\frac{1}{2}}{1} \le z \le \frac{1}{1} \Rightarrow \frac{1}{2} \le z \le 1$

Subscribe EC Academy on YouTube

$$\therefore h_2 = \frac{1}{\sqrt{N}} \begin{cases} 2^{p/2} & 0 \le z < 1/2 \\ -2^{p/2} & \frac{1}{2} \le z < 1 \\ 0 & \end{cases}$$

$$h_1\left(\frac{0}{2}\right) = h_1(0) = \frac{2^{p/2}}{\sqrt{N}} = \frac{2^{0/2}}{\sqrt{2}} = \frac{2^0}{\sqrt{2}} = \underline{\underline{\frac{1}{\sqrt{2}}}}$$

$$h_1\left(\frac{1}{2}\right) = h_1\left(\frac{1}{2}\right) = \frac{-2^{p/2}}{\sqrt{N}} = \frac{-2^{0/2}}{\sqrt{2}} = \frac{-2^0}{\sqrt{2}} = \underline{\underline{\frac{-1}{\sqrt{2}}}}$$

$$\therefore H_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} = \underline{\underline{\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}}}$$

(2) Find 4X4 Haar matrix H_4.

<u>Soln:-</u> $N=4$ & $n=2$

$$H_4 = \begin{array}{c} \\ K=0 \\ K=1 \\ K=2 \\ K=3 \end{array} \begin{array}{c} \begin{array}{cccc} z=0/4 & z=\frac{1}{4} & z=\frac{2}{4} & z=\frac{3}{4} \end{array} \\ \begin{bmatrix} h_0(0/4) & h_0(1/4) & h_0(2/4) & h_0(3/4) \\ h_1\left(\frac{0}{4}\right) & h_1\left(\frac{1}{4}\right) & h_1\left(\frac{2}{4}\right) & h_1\left(\frac{3}{4}\right) \\ h_2\left(\frac{0}{4}\right) & h_2\left(\frac{1}{4}\right) & h_2\left(\frac{2}{4}\right) & h_2\left(\frac{3}{4}\right) \\ h_3\left(\frac{0}{4}\right) & h_3\left(\frac{1}{4}\right) & h_3\left(\frac{2}{4}\right) & h_3\left(\frac{3}{4}\right) \end{bmatrix} \end{array}$$

a) For $K=0$, $h_0\left(\frac{0}{4}\right) = h_0\left(\frac{1}{4}\right) = h_0\left(\frac{2}{4}\right) = h_0\left(\frac{3}{4}\right) = \frac{1}{\sqrt{N}} = \underline{\underline{\frac{1}{\sqrt{4}}}}$

b) For $K=1$, <u>interval 1</u>: $\frac{q-1}{2^p} \le z < \frac{q-\frac{1}{2}}{2^p} \Rightarrow \frac{1-1}{2^0} \le z < \frac{1-\frac{1}{2}}{2^0}$

$p=0$, $q=1$

$$\Rightarrow \frac{0}{1} \le z \le \frac{1/2}{1}$$

$$\Rightarrow \underline{\underline{0 \le z \le \frac{1}{2}}}$$

Subscribe EC Academy on YouTube

interval 2: $\frac{q-\frac{1}{2}}{2^p} \le z < \frac{q}{2^p} \Rightarrow \frac{1-\frac{1}{2}}{2^0} \le z \le \frac{1}{2^0}$

$$\Rightarrow \frac{\frac{1}{2}}{1} \le z \le \frac{1}{1}$$

$$\Rightarrow \frac{1}{2} \le z \le 1$$

$$h_1\left(\frac{0}{4}\right) = h_1(0) = \frac{2^{p/2}}{\sqrt{4}} = \frac{2^{0/2}}{\sqrt{4}} = \frac{2^0}{\sqrt{4}} = \frac{1}{\sqrt{4}}$$

$$h_1\left(\frac{1}{4}\right) = \frac{2^{p/2}}{\sqrt{N}} = \frac{2^{0/2}}{\sqrt{4}} = \frac{1}{\sqrt{4}}$$

$$h_1\left(\frac{2}{4}\right) = h_1\left(\frac{1}{2}\right) = \frac{-2^{p/2}}{\sqrt{N}} = \frac{-2^{0/2}}{\sqrt{4}} = \frac{-1}{\sqrt{4}}$$

$$h_1\left(\frac{3}{4}\right) = \frac{-2^{p/2}}{\sqrt{N}} = \frac{-2^{0/2}}{\sqrt{4}} = \frac{-1}{\sqrt{4}}$$

$\Rightarrow$ For $k=2$, $p=1$, $q=1$

interval 1: $\frac{q-1}{2^p} \le z < \frac{q-\frac{1}{2}}{2^q} \Rightarrow \frac{1-1}{2^1} \le z < \frac{1-\frac{1}{2}}{2^1}$

$$\Rightarrow 0 \le z < \frac{1}{4}$$

interval 2: $\frac{q-\frac{1}{2}}{2^p} \le z < \frac{q}{2^p} \Rightarrow \frac{1-\frac{1}{2}}{2^1} \le z < \frac{1}{2^1}$

$$\Rightarrow \frac{1}{4} \le z < \frac{1}{2}$$

$$h_2\left(\frac{0}{4}\right) = h_2(0) = \frac{2^{p/2}}{\sqrt{N}} = \frac{2^{1/2}}{\sqrt{4}} = \frac{\sqrt{2}}{\sqrt{4}}$$

$$h_2\left(\frac{1}{4}\right) = \frac{-2^{p/2}}{\sqrt{N}} = \frac{-2^{1/2}}{\sqrt{4}} = \frac{-\sqrt{2}}{\sqrt{4}}$$

$h_2\left(\frac{2}{4}\right) = h_3\left(\frac{1}{2}\right) = 0$

since $z = \frac{1}{2}$ is neither in interval 1 nor in interval 2.

olog
mpor
tion
& th

-or
ing

is
y re
al
mpl

-D inf
we w

rite

or em

Subscribe EC Academy on YouTube

$h_2\left(\frac{3}{4}\right) = h_2\left(\frac{3}{4}\right) = \underline{\underline{0}}$ as $z = \frac{3}{4}$ is neither in interval 1 nor in interval 2.

d) for $K = 3$, $\left.\begin{matrix} P = 1 \\ q = 2 \end{matrix}\right\}$

h interval 1: $\frac{q - \frac{1}{2}}{2^P} \le z \le \frac{q}{2^q} \Rightarrow \frac{2 - \frac{1}{2}}{2^1} \le z < \frac{2 - \frac{1}{2}}{2^1}$

$$\Rightarrow \underline{\underline{\frac{1}{2} \le z < \frac{3}{4}}}$$

h

$$\frac{q - \frac{1}{2}}{2^P} \le z < \frac{q}{2^P} \Rightarrow \frac{2 - \frac{1}{2}}{2^1} \le z < \frac{2}{2^1}$$

$$\Rightarrow \underline{\underline{\frac{3}{4} \le z < 1}}$$

2) f $h_3\left(\frac{0}{4}\right) = \underline{\underline{0}}$ $h_3\left(\frac{1}{4}\right) = \underline{\underline{0}}$ $h_3\left(\frac{2}{4}\right) = h_3\left(\frac{1}{2}\right) = \frac{2^{P/2}}{\sqrt{N}}$

o/n

$$= \frac{2^{1/2}}{\sqrt{4}} = \underline{\underline{\frac{\sqrt{2}}{\sqrt{4}}}}$$

$$h_3\left(\frac{3}{4}\right) = \frac{-2^{P/2}}{\sqrt{4}} = \frac{-2^{1/2}}{\sqrt{4}} = \underline{\underline{\frac{-\sqrt{2}}{\sqrt{4}}}}$$

$$H_4 = \frac{1}{\sqrt{4}} \underline{\underline{\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ \sqrt{2} & -\sqrt{2} & 0 & 0 \\ 0 & 0 & \sqrt{2} & \sqrt{2} \end{bmatrix}}}$$

g) f

) for

Subscribe EC Academy on YouTube

4.3 Morphological Image Processing:—

4.3.1 Preliminaries:—

* In Image processing, mathematical morphology is used as a tool for extracting image component that are useful in representation & description of region shape such as boundaries, skeletons & the convex hull.

* Morphological techniques are also used for pre- or post processing such as morphological filtering, thinning & pruning.

* The language of mathematical morphology is set theory. Sets in mathematical morphology represent objects in an image. For example, the set of all black pixels in a binary image is a complete morphological description of the image.

a) Some Basic Concepts from Set Theory:—

Let 'A' be a set in Z^2 where 'Z' is a 2-D integer space. If $a=(a_1, a_2)$ is an element of A, then we write

$$a \in A \quad \text{——} \quad (1)$$

Similarly, if a is not an element of 'A; we write

$$a \notin A \quad \text{——} \quad (2)$$

The set with no elements is called null or empty set & is denoted by the symbol $\emptyset$.

Subscribe EC Academy on YouTube

A set is specified by the contents of two braces: { }. The elements of the sets are the coordinates of pixels representing objects or other features of interest in an image.

If every element of a set A is also an element of another set B, then A is said to be a subset of B, denoted as,

$$A \subseteq B \quad \text{——— (3)}$$

The union of two sets A & B is denoted by

$$C = A \cup B \quad \text{——— (4)}$$

The intersection of two sets A & B is denoted by,

$$D = A \cap B \quad \text{——— (5)}$$

Two sets A & B are said to be disjoint or mutually exclusive if they have no common elements.

In this case, $A \cap B = \emptyset$ ——— (6)

The complement of a set A is the set of elements not contained in A:

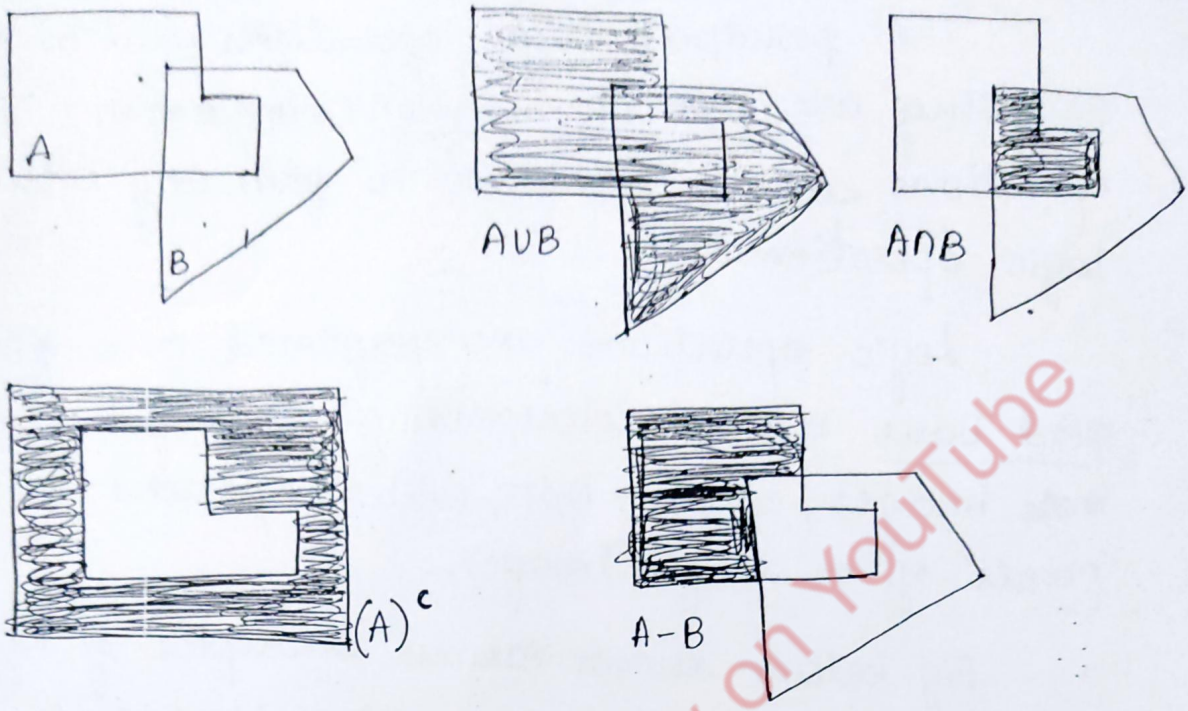
$$A^c = \{ w \mid w \notin A \} \quad \text{——— (7)}$$

The difference of two sets A & B, denoted A − B, is defined as,

$$A - B = \{ w \mid w \in A, w \notin B \} = A \cap B^c \quad \text{——— (8)}$$

Fig. below illustrates the examples for above set theory. The result of set operation indicated in each figure is shown in gray.

Subscribe EC Academy on YouTube



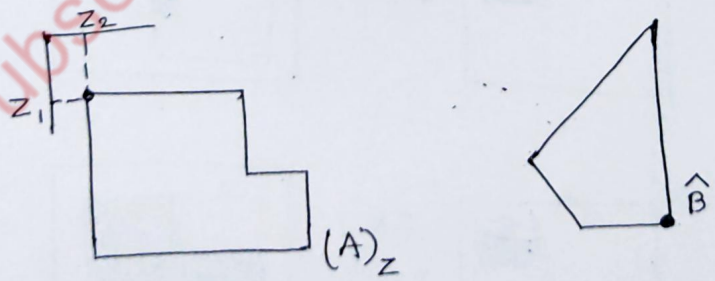
The <u>reflection</u> of set B, denoted $\hat{B}$ is defined as,

$$\hat{B} = \{ w \mid w = -b, \text{ for } b \in B \} \quad \text{—— (9)}$$

The <u>translation</u> of set A by point $z = (z_1, z_2)$, denote $(A)_z$ is defined as,

$$(A)_z = \{ c \mid c = a + z, \text{ for } a \in A \} \quad \text{—— (10)}$$

<u>Ex:-</u>



b) <u>Logic operations involving Binary Images:-</u>

* Logic operations provide a powerful complement to implementation of image processing algorithms

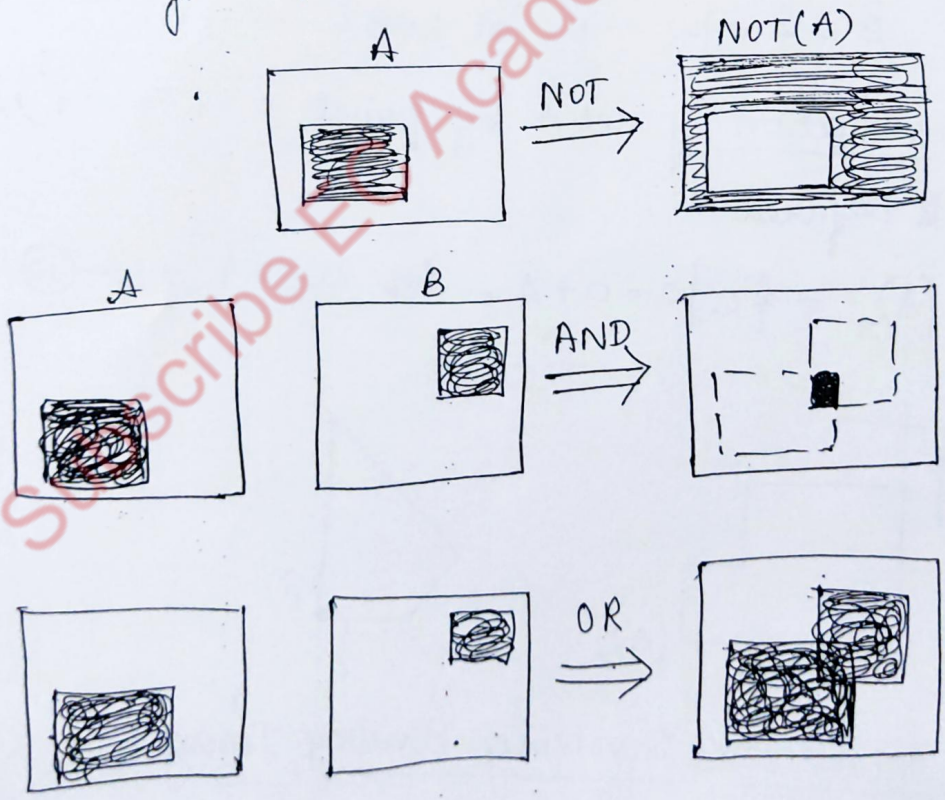
Subscribe EC Academy on YouTube

based on morphology.

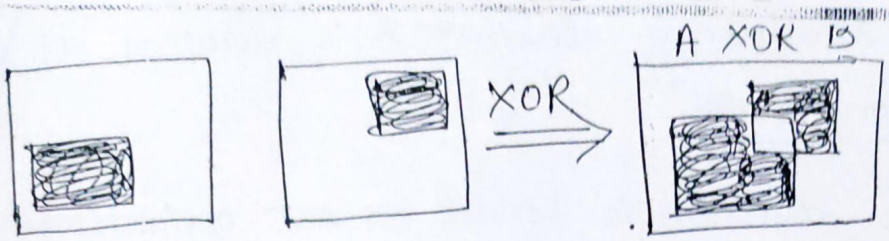
The principal logic operations used in image processing are AND, OR and NOT (COMPLEMENT). These operations can be combined to form any other logic operation.

Logic operations are performed on a pixel by pixel basis between corresponding pixels of two or more images, except NOT, which operates on the pixels of a single image).

Fig. below shows various examples of logic operations involving images, where black indicates a binary 1 & white indicates a 0.



Subscribe EC Academy on YouTube



4.3.2 Dilation & Erosion:-

4.3.2a Dilation:- Expands the boundary of A (I-Image).

* With A & B as sets in Z^2, the dilation of A by B, deno A $\oplus$ B, is defined as

$$A \oplus B = \{ z \mid (\hat{B})_z \cap A \neq \phi \} \quad \text{—— (1)}$$

this equation is based on obtaining the reflection of B about its origin & shifting this reflection by z. The dilation of A by B then is the set of all displacements z, such that $\hat{B}$ & A overlap by atleast one element.

* Based on above interpretation, eqn (1) may be rewri as,

$$A \oplus B = \{ z \mid [(\hat{B})_z \cap A] \subseteq A \} \quad \text{—— (2)}$$

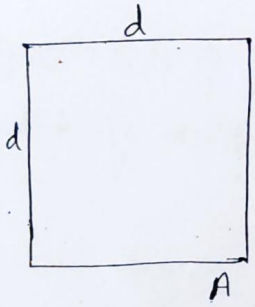
Set B is commonly referred to as the <u>structuring element</u> in dilation, as well as in other morphological operations.

* The preceding definition has a distinct advantage over other formulations in that it is more intuit

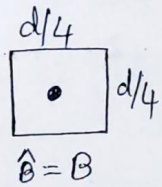
Subscribe EC Academy on YouTube

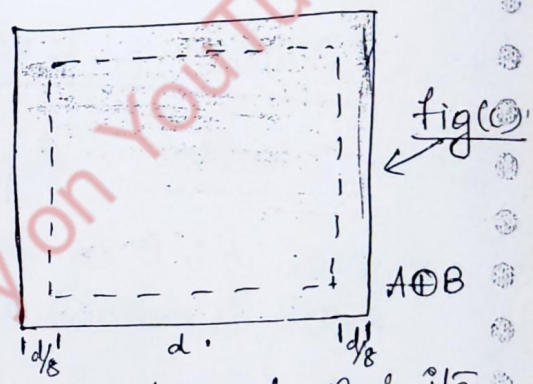
when the structuring element B is viewed as a convolution mask.

* Although dilation is based on set operations, convolution is based on arithmetic operations, the basic process of "flipping" B about its origin & then successively displacing it so that it slides over set A.



fig(a): Set A

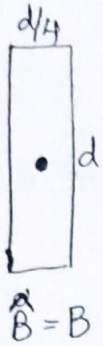




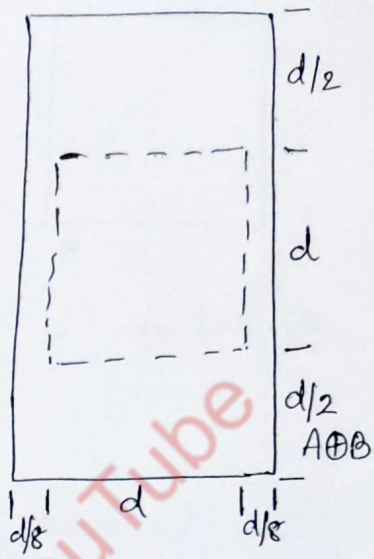
fig(b): Structuring element B & its reflection. The dark dot denotes the origin of the element. In this case the structuring element & its reflection are equal because B is symmetric w.r.to its origin.

In fig(c), ~~the dashed~~ shows the original set for reference, & the solid line shows the limit beyond which any further displacements of the origin of $\hat{B}$ by z would cause the intersection of $\hat{B}$ and A to be empty. Therefore, all points inside this boundary constitute the dilation of A by B.

Subscribe EC Academy on YouTube



fig(d): shows a structuring element designed to achieve more dilation vertically than horizontally.



fig(e): shows the dilation of A achieved with this element.

* One of the simplest applications of dilation is for bridging gaps. Consider an image with broken characters. Let the maximum length of the breaks is known to be two pixels. A simple structuring element can be used for repairing the gaps. Thus the gaps have been bridged.

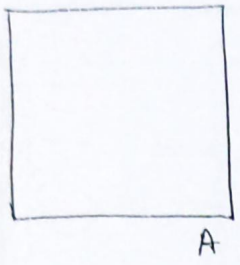
4.3.2b Erosion:— Contracts the boundary of A. (I)

For sets A & B in Z^2 the erosion of A by B, denoted $A \ominus B$ is defined as,

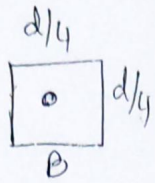
$$A \ominus B = \{ z \mid (B)_z \subseteq A \} \quad \text{——— (3)}$$

The above equation indicates that the erosion of A by B is the set of all points z such that B, translat by z, is contained in A.

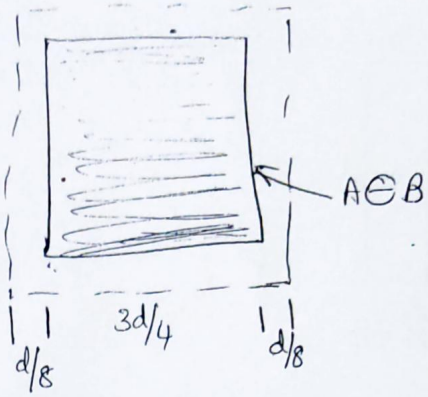
Subscribe EC Academy on YouTube



(a) Set A.

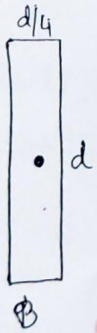


(b) Square Structuring element

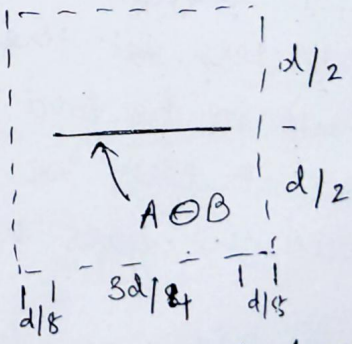


(c) Erosion of A by B, shown shaded

The boundary of the shaded region shows the limit beyond which further displacement of the origin of B would cause this set to cease being completely contained in A. Thus, the locus of points within this boundary i.e., the shaded region constitutes the erosion of A by B.



(d) Elongated Structuring element



(e) Erosion of A by this element.

Note that the original set was eroded down to line.

* Dilation & erosion are duals of each other w.r.to set complementation & reflection,

i.e., $(A \ominus B)^c = A^c \oplus \hat{B}$ ——— (4)

Subscribe EC Academy on YouTube

Proof:- From definition of erosion, we have

$$(A \ominus B)^c = \{z \mid (B)_z \subseteq A\}^c$$

If set $(B)_z$ is contained in set A, then $(B)_z \cap A^c = \phi$ in which case the preceding equation becomes,

$$(A \ominus B)^c = \{z \mid (B)_z \cap A^c = \phi\}^c$$

But the complement of the set of z's that satisfy $(B)_z \cap A^c = \phi$ is the set of z's such that $(B)_z \cap A^c \neq \phi$.

Thus $$(A \ominus B)^c = \{z \mid (B)_z \cap A^c \neq \phi\}$$
$$= A^c \oplus \hat{B}.$$

* One of the simplest use of erosion is for eliminatin irrelevant detail (in terms of size) from a binary image

[Refer to fig 9.7 page no. 527]

4.3.3 Opening & Closing:-

* Dilation expands an image & erosion shrinks it
* Opening generally smoothes the contour of an object, breaks narrow isthmuses & eliminates thin protrusions.
* Closing also tends to smooth sections of contour but, as opposed to opening, it generally fuses narrow breaks & long thin gulfs, eliminates small holes & fills gaps in the contour.

Subscribe EC Academy on YouTube

* The opening of set A by structuring element B, denoted $A \circ B$, is defined as

$$A \circ B = (A \ominus B) \oplus B \quad \text{—— (1)}$$

Thus the opening A by B is the erosion of A by B followed by a dilation of the result by B.

* Similarly, the closing of set A by structuring element B, denoted $A \cdot B$ is defined as,

$$A \cdot B = (A \oplus B) \ominus B \quad \text{—— (2)}$$

Thus closing of A by B is simply the dilation of A by B, followed by the erosion of the result by B.

[Refer to fig. 9.8 & 9.9 & 9.11 pg. no: 528 & 529]

9.3.4 The Hit-or-Miss Transformation:-

* It is a basic tool for shape detection.

$$A \circledast B = (A \ominus X) \cap [A^c \ominus (W - X)] \quad \text{—— (1)}$$

$$A \circledast B = (A \ominus B_1) \cap (A^c \ominus B_2) \quad \text{—— (2)}$$

$$A \circledast B = (A \ominus B_1) - (A \oplus \hat{B}_2) \quad \text{—— (3)}$$

[Refer to fig 9.12 page no. 533] as example.

The set of points (coordinates) at which, simultaneously, B_1 found a match ("hit") in A & B_2 found a match in A^c.

Subscribe EC Academy on YouTube

4.3.5 Some Basic Morphological Algorithms:-

1) **Boundary Extraction**:- Set of points on the boundary of set A. (I).

$$\beta_1(A) = A - (A \ominus B)$$

[Refer fig. 9.13 & 9.14]

2) **Region Filling**:- Fills a region in A, given a point p the region (II)

$$X_K = (X_{K-1} \oplus B) \cap A^c \quad K = 1, 2, 3, \ldots$$

[Refer to ex 9.15 & 9.16]

3) **Extraction of connected components**:- Finds a connec component Y in A, given a point in Y.

$$X_K = (X_{K-1} \oplus B) \cap A \quad K = 1, 2, 3, \ldots$$

Refer to fig 9.17 & 9.18

4) **Convex Hull**:-

A set A is said to be convex if the straig line segment joining any two points in A lies entire within A. The convex hull H of an arbitrary set S is the smallest convex set containing S. The set difference H-S is called the convex deficiency of S

Let B^i, $i = 1, 2, 3, 4$ represent the 4 structurin elements. The procedure is

$$X_K^i = (X_{K-1} \otimes B^i) \cup A \quad i = 1, 2, 3, 4 \ \& \ K = 1, 2, 3, \ldots \quad \text{—(4)}$$

with $X_0^i = A$.

Subscribe EC Academy on YouTube

Now let $D^i = X^i_{conv}$, where "conv" indicates convergence i.e., $X^i_K = X^i_{K-1}$.

Then the convex hull of A is

$$C(A) = \bigcup_{i=1}^{4} D^i \quad \text{——— (5)}$$

5) Thinning:- Thins set A. The eqn's (6a) & (6b) give the basic definition of thinning. Eqns (7) & (8) denote thinning by a sequence of structuring elements. This method is normally used in practice.

$$A \otimes B = A - (A \circledast B) \quad \text{——— (6a)}$$

$$= A \cap (A \circledast B)^c \quad \text{——— (6b)}$$

$$\{B\} = \{B^1, B^2, B^3, \ldots, B^n\} \quad \text{——— (7)}$$

$$A \otimes \{B\} = ((\cdots((A \otimes B^1) \otimes B^2) \cdots) \otimes B^n). \quad \text{——— (8)}$$

6) Thickening:- Thickens set A.

$$A \odot B = A \cup (A \circledast B) \quad \text{——— (9)}$$

$$A \odot \{B\} = ((\cdots((A \odot B^1) \odot B^2) \cdots) \odot B^n) \quad \text{——— (10)}$$

7) Skeletons:- Finds the Skeleton $S(A)$ of set A.

$$S(A) = \bigcup_{K=0}^{K} S_K(A) \quad \text{——— (11)}$$

With subsets ← $S_K(A) = (A \ominus KB) - (A \ominus KB) \circ B$ ——— (12)

$$(A \ominus KB) = (\cdots(A \ominus B) \ominus B) \ominus \cdots) \ominus B \quad \text{——— (13)}$$

$$K = \max\{K \mid (A \ominus KB) \neq \phi\} \quad \text{——— (14)}$$

$(\ominus KB)$ denotes K^{th} iteration of successive erosion of A by B

Subscribe EC Academy on YouTube

can be reconstructed using subsets $S_K(A)$,

$$A = \bigcup_{K=0}^{K} (S_K(A) \oplus KB)$$

'K' is the value of iterative step

8) Pruning :-

Pruning methods are essential complement to thinning & skeletonizing algorithms because these procedures tend to leave parasitic components that need to be "cleaned up" by post processing

Ex:- A common approach in the automated recognition of hand-printed characters is to analyze the shape of the skeleton of each character

$$X_1 = A \otimes \{B\}$$

$$X_2 = \bigcup_{K=1}^{8} (X_1 \circledast B^K)$$

$$X_3 = (X_2 \oplus H) \cap A$$

$$X_4 = X_1 \cup X_3$$

Subscribe EC Academy on YouTube

4

$$\begin{bmatrix} 4 & 7 \\ 2 & 6 \end{bmatrix}^{-1} = \frac{1}{|A|} \operatorname{adj}[A]$$

$$= \frac{1}{4\times6 - 7\times2} \begin{bmatrix} 6 & -7 \\ -2 & 4 \end{bmatrix}$$

$$= \frac{1}{24-14} \begin{bmatrix} 6 & -7 \\ -2 & 4 \end{bmatrix}$$

$$= \frac{1}{10} \begin{bmatrix} 6 & -7 \\ -2 & 4 \end{bmatrix}$$

$$= \begin{bmatrix} 0.6 & -0.7 \\ -0.2 & 0.4 \end{bmatrix}$$

Soln:-

(a) $T = H_2 F H_2$

$$= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \times \begin{bmatrix} 3 & -1 \\ 6 & 2 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & \end{bmatrix}$$

$$= \frac{1}{2} \begin{bmatrix} 3+6 & -1+2 \\ 3-6 & -1-2 \end{bmatrix} \begin{bmatrix} 1 & \\ & \end{bmatrix}$$

$$= \frac{1}{2} \begin{bmatrix} 9 & 1 \\ -3 & -3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$= \frac{1}{2} \begin{bmatrix} 9+1 & 9-1 \\ -3-3 & -3+3 \end{bmatrix}$$

$$= \frac{1}{2} \begin{bmatrix} 10 & 8 \\ -6 & 0 \end{bmatrix}$$

$$\boxed{T = \begin{bmatrix} 5 & 4 \\ -3 & 0 \end{bmatrix}}$$

) (a) Compute the Haar transform of the 2x2 image

$$F = \begin{bmatrix} 3 & -1 \\ 6 & 2 \end{bmatrix}$$

(b) The inverse Haar transform is $F = H^{-1} T H^{-1}$, where T is the Haar transform & H^{-1} denotes the matrix inverse of Haar Transformation matrix H. Find H_2^{-1} for H_2 & use it to compute the inverse Haar transform of result in (a).

Subscribe EC Academy on YouTube

(b) $F = H^{-1} T H^{-1}$

$= H_2^{-1} T H_2^{-1}$

$= \frac{-1}{\sqrt{2}} \begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 5 & 4 \\ -3 & 0 \end{bmatrix} \times \frac{-1}{\sqrt{2}} \begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix}$

$= \frac{1}{2} \begin{bmatrix} -5+3 & -4+0 \\ -5-3 & -4+0 \end{bmatrix} \begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix}$

$= \frac{1}{2} \begin{bmatrix} -2 & -4 \\ -8 & -4 \end{bmatrix} \begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix}$

$= \frac{1}{2} \begin{bmatrix} 2+4 & 2-4 \\ 8+4 & 8-4 \end{bmatrix}$

$= \frac{1}{2} \begin{bmatrix} 6 & -2 \\ 12 & 4 \end{bmatrix}$

$F = \begin{bmatrix} 3 & -1 \\ 6 & 2 \end{bmatrix}$

Thus from inverse Haar Transform, the 2x2 image F in (a) is retrieved back.

$H_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$

$H_2 = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$

$H_2^{-1} = \frac{1}{|H_2|} \cdot \text{adj}[H_2]$

$= \frac{1}{\left[-\frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}}\right]} \begin{bmatrix} -1/\sqrt{2} & -1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$

$= \frac{1}{\frac{-2}{\sqrt{2}}} \begin{bmatrix} -1/\sqrt{2} & -1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$

$= \frac{-1}{\sqrt{2}} \begin{bmatrix} -1/\sqrt{2} & -1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$

$\cancel{= \frac{1}{2} \begin{bmatrix} -\sqrt{2} & -1 \\ -\sqrt{2} & 1 \end{bmatrix}}$

$H_2^{-1} = \frac{-1}{\sqrt{2}} \begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix}$

Subscribe EC Academy on YouTube

Properties of Haar matrix :-

1) H is real

$$H_4 = \frac{1}{\sqrt{4}}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ \sqrt{2} & \sqrt{2} & 0 & 0 \\ 0 & 0 & \sqrt{2} & \sqrt{2} \end{bmatrix} \qquad H_4^* = \frac{1}{\sqrt{4}}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ \sqrt{2} & \sqrt{2} & 0 & 0 \\ 0 & 0 & \sqrt{2} & \sqrt{2} \end{bmatrix}$$

$$= H$$

$\therefore$ As $H = H^* \Rightarrow H$ is real.

2) H is not symmetrical

$$H^T = \frac{1}{\sqrt{4}}\begin{bmatrix} 1 & 1 & \sqrt{2} & 0 \\ 1 & 1 & -\sqrt{2} & 0 \\ 1 & -1 & 0 & \sqrt{2} \\ 1 & -1 & 0 & -\sqrt{2} \end{bmatrix} \neq H$$

As $H \neq H^T \Rightarrow H$ is not symmetrical

3) H is unitary

$$H^T = \frac{1}{\sqrt{4}}\begin{bmatrix} 1 & 1 & \sqrt{2} & 0 \\ 1 & 1 & -\sqrt{2} & 0 \\ 1 & -1 & 0 & \sqrt{2} \\ 1 & -1 & 0 & -\sqrt{2} \end{bmatrix} = H$$

Subscribe EC Academy on YouTube

$$HH^{*T} = \frac{1}{\sqrt{4}} \times \frac{1}{\sqrt{4}} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ \sqrt{2} & -\sqrt{2} & 0 & 0 \\ 0 & 0 & \sqrt{2} & -\sqrt{2} \end{bmatrix} \begin{bmatrix} 1 & 1 & \sqrt{2} & 0 \\ 1 & 1 & -\sqrt{2} & 0 \\ 1 & -1 & 0 & \sqrt{2} \\ 1 & -1 & 0 & -\sqrt{2} \end{bmatrix}$$

$$= \frac{1}{4} \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = I$$

As $HH^{*T} = I \Rightarrow H$ is unitary.

Properties:

1. H is real, not symmetrical & Unitary.
2. The basis vectors of Haar transform is sequency ord
3. It has poor energy compaction for images.
4. It is used in edge detection applications.
5. Haar Transform is a very fast algorithm. It can b implemented in O(N) operations for N dimensiona vector.

SubscribeEC Academy on YouTube