

IT6005 - Digital Image Processing

UNIT - I

Digital Image Fundamentals.

Introduction - Origin

Steps in Digital Image processing

Components

Elements of Visual perception

Image Sensing and Acquisition

Image Sampling and Quantization

Relationships between pixels

color models.

1. Introduction:-

Interest in digital image processing methods stems from two principal application areas.

1. Improvement of pictorial information for human perception / Interpretation
2. Processing of image data for storage, transmission and representation for autonomous machine perception.

1. What is Digital Image Processing?

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

When x, y and the intensity values of 'f' are all finite, discrete quantities, we call the image is a digital image

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

⇒ Digital image is composed of a finite number of elements, each of which has a particular location and value.

⇒ These elements are called "picture elements" / "image elements" / "pels" and "pixels".

⇒ "pixel" is the term used most widely to denote the elements of digital image.

Vision is the most advanced of our senses. Humans, who are limited to the visual band of the electromagnetic (EM) spectrum, imaging machines cover almost the entire (EM) spectrum, ranging from gamma to radio waves.

These include ultrasound, electron microscopy, and computer-generated images. Thus, digital image processing encompasses a wide and varied field of application.

Three types of Computerized processes

1. Low-level process
2. Mid-level process
3. High-level process.

1. Low-level process:-

* primitive operations:

Such as, image preprocessing to reduce noise, contrast enhancement and image sharpening.

* characterized by,

Both inputs and outputs are images.

2. Mid-level process:-

* primitive operations:-

on images involves task such as Segmentation (partitioning an image into regions/objects), description of those objects to reduce them to form suitable for computer processing, and classification (recognition) of individual objects.

* characterized by:

Its inputs generally are images,

its outputs are attributes extracted from these images.

3. High-level process:-

* primitive operations:-

"making sense" of an ensemble of recognized objects as an image analysis and at far end of the process, performing the

(2)
Cognitive functions normally associated with vision.

* Characterized by,

its inputs generally an attributes of an image
its outputs making sense / understanding

② Origin of Digital Image processing:-

- * one of the first applications of digital images was in the newspaper industry, when pictures were first sent by "Submarine Cable" between London and New York.
- * Eastline Cable picture transmission system in 1920's reduced the time required to transport a picture across the Atlantic from more than week to < than three hours.
- * Specialized printing equipment coded pictures for cable transmission and then reconstructed them at the receiving end.
- * The printing method above said was abandoned at the end of 1921.
- * A new technique based on photographic reproduction made from tapes perforated at the telegraph receiving terminal. In received images an improvement of tonal quality and in resolution.
- * The early Eastline System were introduced to capable of coding images in five distinct levels of gray.
- * This capability was increased to 15 levels in 1929.
- * During this period, introduction of a system for developing a film plate via light beams that were modulated by the coded picture tape improved the reproduction process considerably.
- * The first computer powerful enough to carry out meaningful image processing tasks appeared in the early 1960's.

Examples of Fields that use Digital Image Processing

The areas of application of digital image processing are so varied that some form of organization is desirable in attempting to capture the breadth of this field.

Categorize the images according to their source (x-ray, gamma). The principal energy source for images in use today is the electromagnetic energy spectrum. Other important sources of energy include acoustic, ultrasonic and electronic. Synthetic images, used for modeling and visualization, are generated by computers.

Images based on radiation from EM spectrum are the most familiar, especially images in the x-ray and visual bands of the spectrum.

1. Gamma-Ray Imaging:
 - imaging based on gamma rays
 - nuclear medicine & astronomical observations, etc.,
2. X-Ray Imaging:
 - imaging based on x-rays.
 - Medical diagnostics, Industry, astronomy, etc.,
3. Imaging in the Ultraviolet Band.
 - imaging based on ultraviolet light source
 - Lithography, Industrial inspection, microscopy, lasers, biological imaging astronomical observations, etc.,
4. Imaging in Visible and Infrared Bands.
 - imaging based on visible light sources and infrared light sources.
 - pharmaceuticals & microinspection to materials characterization, Remote sensing, etc.,

(3)

- Infrared imaging:- Nighttime Lights of the world data set., automated visual inspection of manufactured goods.

5. Imaging in the Microwave Band:-

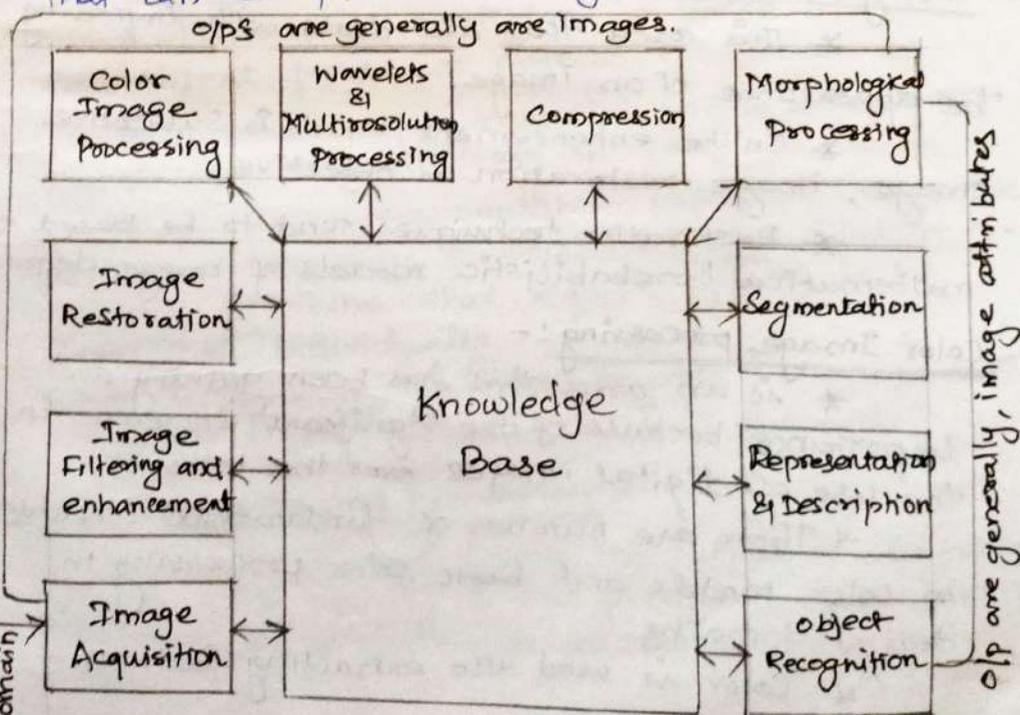
- imaging based on microwave band spectrum.
- Radar imaging, (Explore inaccessible regions of the Earth's surface)

6. Imaging in the Radio Band:-

- imaging based on Radio band spectrum.
- medicine and astronomy (MRI)

③ Steps in Digital Image processing

Computerized process methods whose ip and op are images, and methods whose ip may be images but whose op are attributes extracted from those images. This organization is summarized in below diagram. This diagram does not imply that every process is applied to an image, rather, the intention is to convey an idea of all the methodologies that can be applied to images for different purposes.



1. Image Acquisition:-

- * Origin of digital images
- * acquisition could be as simple as being given an image that is already in digital form. This stage involves preprocessing, such as scaling.

2. Image Enhancement:-

- * Is the process of manipulating an image, so, that the result is more suitable than the original
- * The Enhancement techniques are problem oriented
- * Eg: quite suitable for enhancing X-ray images, may not be the best approach for enhancing satellite images taken in the infrared band of EM spectrum.
- * There is no general theory of image enhancement
- * The image is processed for visual interpretation, the viewer is the ultimate judge of how well a particular method works.
- * In the field of image processing generally find enhancement applications visually appealing, interesting and relatively simple to understand.

3. Image Restoration:-

- * This area that also deals with improving the appearance of an image.
- * Unlike enhancement, which is subjective image, image restoration is objective.
- * Restoration techniques tend to be based on mathematical/probabilistic models of image degradation

4. Color Image processing:-

- * is an area that has been gaining an importance because of the significant increase in the use of digital images over the internet.
- * There are number of fundamental concepts in color models and basic color processing in digital domains.
- * Color is used also extracting features of interest in an image.
- * Some of the color models - B&W, Gray Scale, RGB, CMY, CMYA, etc.,

5. Wavelets: -

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

* Mostly used for image data compression and for pyramidal representation, in which images are subdivided successively into smaller regions.

6. Compression: -

* Deals with techniques for reducing the storage required to save an image, the bandwidth required to transmit it.

* Image compression is familiar to most users of computers in the form of image file extensions, such as jpg file extension used in the JPEG (Joint Photographic Experts Group) Image compression Standard.

7. Morphological processing: -

* Deals with tools for extracting image components that are useful in the representation and description of shape.

* A transition from processes that output images to processes that output image attributes.

8. Segmentation: -

* Procedures partition an image into its constituent parts or objects.

* Autonomous Segmentation is one of the most significant, difficult task in DIP.

* A rugged Segmentation procedure brings the process a long way toward successful solution of imaging problems that require objects to be identified individually.

* In general, the most accurate the Segmentation, the most likely recognition is to succeed.

9. Representation and Description: -

* almost always follow the output of a Segmentation stage, which usually is raw pixel data, constituting either the boundary of region or all points in the region itself.

* Converting the data to a form suitable for computer processing.

* The decision must be made is whether the data should be represented as a boundary/complete region

* Boundary Representation - focus is on external shape characteristics such as corners and inflections.

* Regional Representation - focus is on internal properties, such as texture/skeletal shape.

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

Description:-

* Also, called feature selection, deals with extracting attributes that result in some quantitative information of interest/basic for differentiating one class of objects from another.

10. Recognition:-

* is the process that assigns a label (eg. Vehicle) to an object based on its descriptors.

* Coverage of DIP with the development of methods for recognition of individual objects.

11. Knowledge Base:-

* Knowledge about a problem domain is coded into an image processing system in the form of a knowledge data base.

* This knowledge may be simple as detailing regions of an image where the information of interest is known to be located.

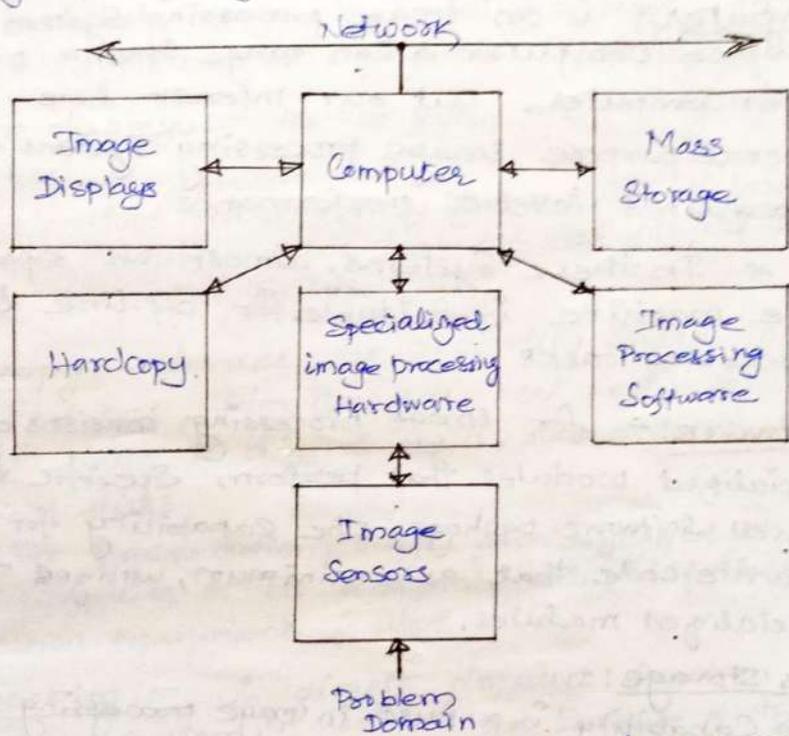
* The knowledge base, also can be quite complex, such as an interrelated list of all major possible defects in a materials inspection problem (or) An image database containing high resolution satellite images of a region in connection with change-detection applications.

* Knowledge Base, as also controls the interaction between modules

④ Components of an Image Processing System.

Large-scale Image Processing Systems still are being sold for massive imaging applications, such as processing of satellite images, the trend continues toward minimizing and blending of general-purpose small computers with specialized image processing hardware.

The below figure shows the basic components comprising a typical general-purpose system used for digital image processing.



1. Sensing:- two elements are required to acquire digital images. The (1) physical device that is sensitive to the energy radiated by the object we wish to image.
- (2) Digitizer, is a device for converting the output of the physical sensing device into general digital form.

* In a digital video camera, the sensors produce an electrical output proportional to light intensity

* Digitizer converts these outputs to digital data.

2. Specialized Image processing hardware: - usually consist of the digitizer + hardware that perform other primitive operations such as, ALU ALU that perform Arithmetic & Logical operations in parallel on entire images.

* Averaging Images as quickly as they are digitized, for noise reduction. This is sometimes called front-end subsystem.

* This unit performs functions that require fast data throughputs. (digitizing & averaging 30 fr/sec)

3. Computer: is an image processing system in general - purpose computer & can range from a PC to a super computer. But our interest here is on general-purpose image processing systems to achieve a required level of performance.

* In these systems, almost well equipped PC type machine is suitable for off-line image processing tasks.

4. Software: - for image processing consists of specialized modules that perform specific tasks. A well software package, the capability for the user to write code that, as a minimum, utilizes the specialized modules.

5. Mass Storage: -

↳ Capability is a must in image processing application

ex: An image size (1024 x 1024) pixels } intensity of each } quantity }
pixel is 8-bit } require to }
1MB of }
Storage space.

- if the image is not compressed →

- if even millions of images are to be stored in DIP Sp

* Digital image processing systems - Digital Storage falls in to three principal categories.

1. Short term storage for using while processing.
2. On-line storage for relatively fast recall
3. Archival storage - characterized by infrequent access.

* Storage measured in Bytes (8 bits)

KBytes (1000 bytes)

MBytes (1 million bytes)

GBytes (01 billion bytes)

TBytes (1 Trillion bytes)

→ 1 Short term storage - Computer memory, frame buffers. - Store one or more images can accessed.

2. Online storage - takes the form of magnetic disks or optical-media storage. - frequent access to the stored data.

3. Archival storage - by massive storage - infrequent need for access. - Magnetic tapes / optical disks.

6. Image displays: - In use today are mainly color TV monitors

* Monitors are driven by the outputs of image & graphics display cards that are an integral part of the computer system.

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

* For presentation, images are displayed on film transparencies or in a digital medium if image projection equipment is used.

8. Networking: - is almost default function in any computer system in use today. Because of the large amount of data inherent in image processing applications, the key consideration in image txn. is bandwidth.

* In dedicated networks typically is not a problem.

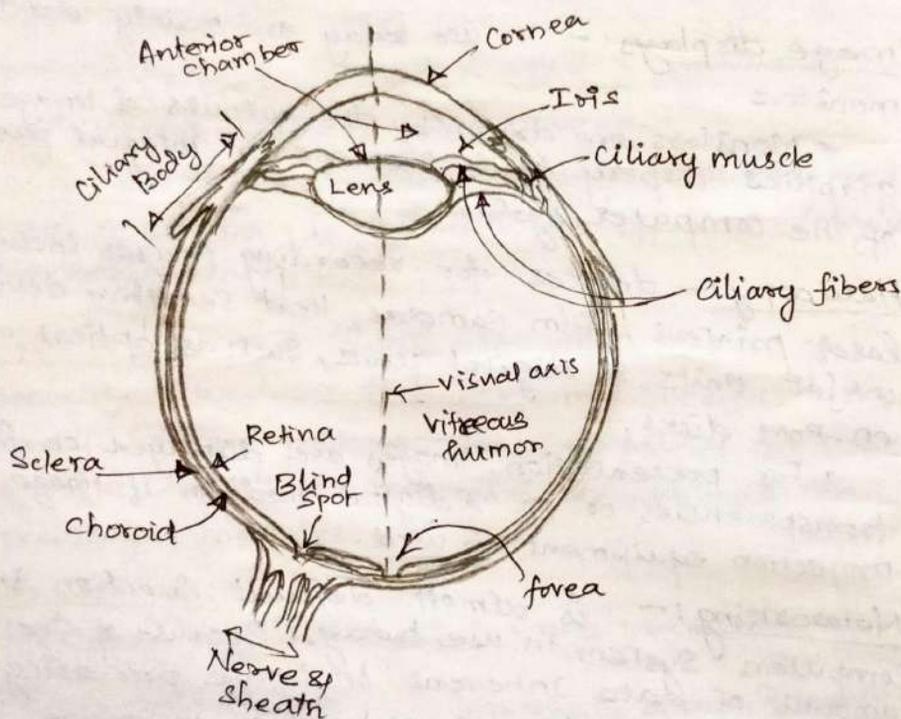
* But in communications with remote sites via the Internet are not always as efficient.

⑤. Elements of Visual Perception:-

Although the field of digital image processing is built on a foundation of mathematical and probabilistic formulations.

Human intuitions and analysis play a central role in the choice of one technique versus another and based on subjective, visual judgements. Hence understanding of human visual perception as a first step. In particular, our interest is in the mechanics and parameters related to how images are formed and perceived by humans.

Structure of the Human Eye:-



- * Above diagram cross section of the Human eye.
- * Eye is nearly sphere - avg. diameter 20 mm
- * Three membrane encloses the eye.
 1. Cornea & Sclera - outer cover
 2. choroid
 3. Retina

Cornea: * is a tough, transparent tissue, that covers the anterior surface of the eye.

Sclera * is an opaque membrane that encloses the remainder of the optic globe.

Choroid * lies directly below the sclera. This membrane contains a network of blood vessels that serve as the major source of nutrition to the eye.
* choroid coat is heavily pigmented and hence helps to reduce the amount of extraneous light entering the eye & the back scatter within the optic globe.

* In its anterior extreme, the choroid is divided into ciliary body and the Iris.

* It contracts or expands to control the amount of light that enters the eye.

* The front of the iris contains the visible pigments of the eye, whereas the back contains a black pigment.

* Lens - made up of concentric layers of fibrous cells and is suspended by fibers that attach to the ciliary body.

- It contains - 60 to 70% water,

6% fat

more protein than any other tissue in the eye.

- The lens absorbs approximately 8% of the visible light spectrum with relatively higher absorption at shorter wavelengths.

Retina:-

The innermost membrane of eye, which lines the inside of the wall's entire posterior portion, when the eye is properly focused, light from an object outside the eye is imaged on the retina.

* Distribution of discrete light receptors over the surface of the retina.

* There two classes of receptors:

1. Cones 2. Rods.

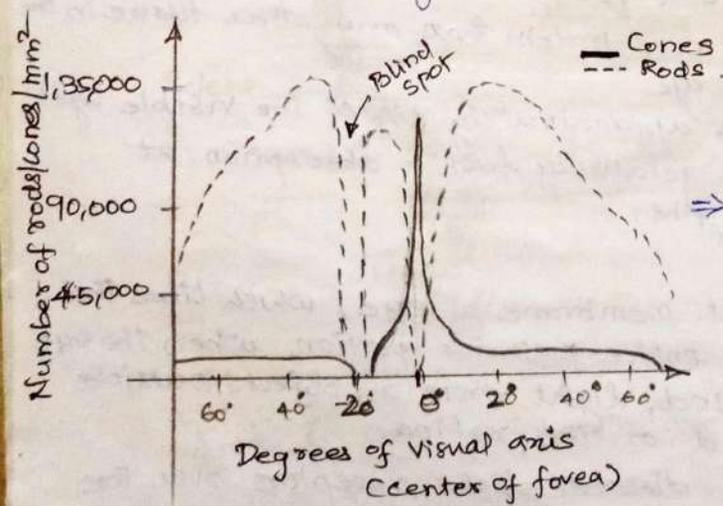
* Cones - 6-7 million

- located primarily in the central portion of the retina, called the 'fovea'.

- highly sensitive to color.
- Human can resolve fine details of with these cones largely, because each one is connected to its own nerve end.
- muscles controlling the eye rotate the eyeball until the image of an object of interest falls on the fovea.
- cone vision is called "photopic" (or) bright-light vision.

* Rods:

- 75 to 150 million are distributed over the retinal surface.
- larger area of distribution and several rods are connected to single nerve end reduce the amount of detail discernible by these receptors.
- Rods - give overall picture of the field of view
- It is not involved in color vision and
- Sensitive to low levels of illumination.
- moon light vision - colorless forms, because only the rods are stimulated.
- This phenomenon is known as "Scotopic" / dim-light vision.



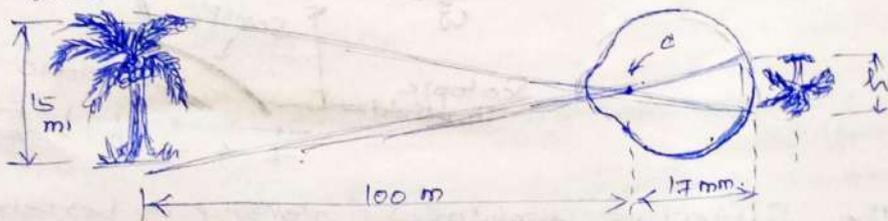
⇒ Shows the density of rods & cones for a cross section of the right eye.

- * The absence of receptors in this area results in so-called "blind spot"
- * Receptors density is measured in degrees from the fovea
- * cones - most dense in the center of the retina

- * rods increase in density from the center out to approximately 20° off axis.
- * The fovea itself is a circular indentation in the retina of about 1.5 mm in diameter.
- * The density of cones in that area of the retina is approximately 150,000 elements/mm².
- * Based on these approximations, the number of cones in the region of highest acuity in the eye is about 337,000 elements.
- * The ability of humans to integrate intelligence and experience with vision makes these types of number comparisons somewhat superficial.

Image Formation in the Eye:-

- * ordinary photographic camera, the lens has a fixed focal length, and focusing at various distances is achieved by varying the distance between the lens and the imaging plane, where the film is located.



The fibers in ciliary body accomplish this, flattening (or) thickening the lens for distant / near objects respectively.

- * The distance between the center of the lens and the retina along the visual axis is approximately 17 mm.

* The range of focal lengths is approximately 14 mm - 17 mm.

* The above figure, to obtain the dimensions of an image formed on the retina, suppose that a person is looking at a tree 15 m height at a distance of 100 m.

- * let h - height of object in the retina image.

geometry yields $\frac{15}{100} = \frac{h}{17} \Rightarrow h = 2.5 \text{ mm}$.

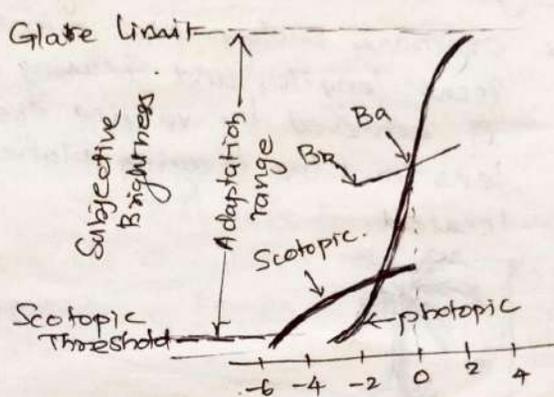
- * Retinal image is focused primarily on the region of the fovea.

* Perception then takes place by the relative excitation of light receptors, which transform radiant energy into electrical impulses that ultimately are decoded by the brain.

Brightness Adaptation and Discrimination

Digital images are displayed as a discrete set of intensities, the eye's ability to discriminate b/w different intensity levels is an important consideration in presenting image processing results.

The range of light intensity levels to which the human visual system can adapt is enormous - on the order of 10^6 - from the scotopic threshold to glare limit.



- * The Subjective Brightness (intensity as perceived by the human visual system). is a logarithmic function of the light intensity incident on the eye.
- * The long solid curve represents the range of intensities to which the visual system can adapt.
- * The visual system cannot operate over the such a range simultaneously. Rather, it accomplishes this large variations by changing its overall sensitivity a phenomenon known as brightness adaptation.
- * For any given set of conditions, the current sensitivity level of the visual system is called the brightness adaptation level.
- * The ability of the eye to discriminate between changes in light intensity at any specific adaptation level is also considerable interest.

* A classic experiment used to determine the capability of the human visual system for brightness discrimination consists of having a subject look at a flat, uniformly illuminated area large enough to occupy the entire field of view.

* opaque glass, that is illuminated from behind by a light source whose intensity 'I' can be varied.

* to this field is added an increment of illumination, ΔI ,

* If ΔI is not bright enough, the subject says "no", indicating no perceivable change.

* If ΔI is stronger, the subject may give a +ve response of "yes" indicating a perceived change.

* The quantity $\Delta I_c / I$, where I_c - incremental of illumination
 I - background illumination.

is called "Weber ratio"

* $\Delta I_c / I = \text{Small}$; means small % change in intensity.
discriminable.

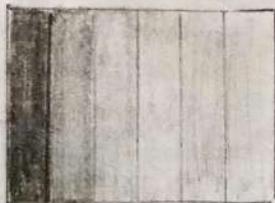
This represents "good" brightness discrimination.

$\Delta I_c / I = \text{large}$; large % change in intensity is required.

This represents "poor" brightness discrimination.

\Rightarrow Two phenomena clearly demonstrate that perceived brightness is not a simple function of intensity.

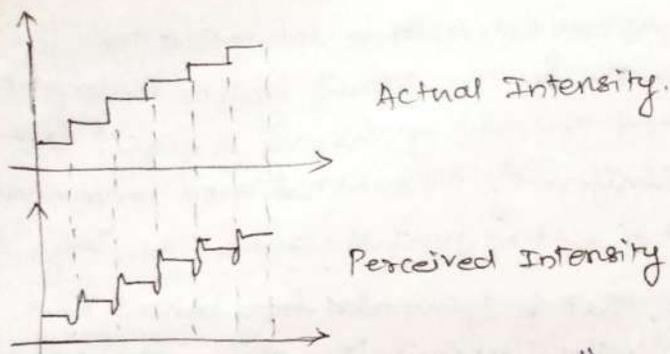
(1). Based on the fact that the visual system tends to undershoot (or) overshoot around the boundary of regions of different intensity.



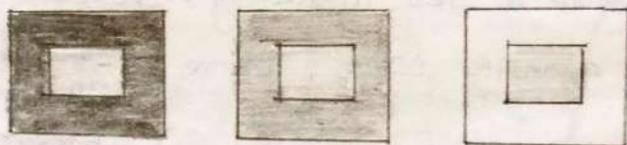
* Although the intensity of the stripes is constant.

* We perceive a brightness pattern that is strongly S-shaped near the boundaries.

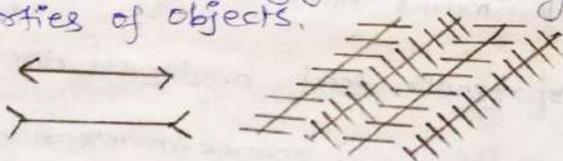
* These S-shaped bands are called Mach Bands.



- (2) Second phenomenon called "Simultaneous Contrast" is related to the fact that regions perceived brightness does not depend simply on its intensity.
- * Below figure, all the Center Squares have exactly the same intensity.
 - * However they appear to the eye to become darker as the background gets lighter.



Other examples of human perception phenomena are optical illusions, in which the eye fills in non-existing information (os) wrongly perceives geometrical properties of objects.



⑥ Image Sensing and Acquisition :-

Most of The images are generated by the combination of an "illumination" source and the reflection (or absorption) of energy from that source by the elements of the "Scene" being imaged.

- * The illumination may originate from a source of EM energy such as Radar, infrared, X-ray system
- * Depending on the nature of source, illumination energy is reflected from, or transmitted through, objects.
- * Ex. 1. light is reflected from a planar surface.
- Ex. 2. when X-rays pass through a patient's body for the purpose of generating diagnostic X-ray film.

other applications like, reflected (or) transmitted energy is focussed on to a photo converter.

⇒ Three principal sensor arrangement used to transform illumination energy into digital images.

- * Idea is, incoming energy is transformed into a voltage by the combination of i/p electrical power and sensor material that is responsive to the particular type of energy being detected.
- * The output voltage waveform is the response of the sensor(s) and a digital quantity is obtained from each sensor by digitizing its responses.

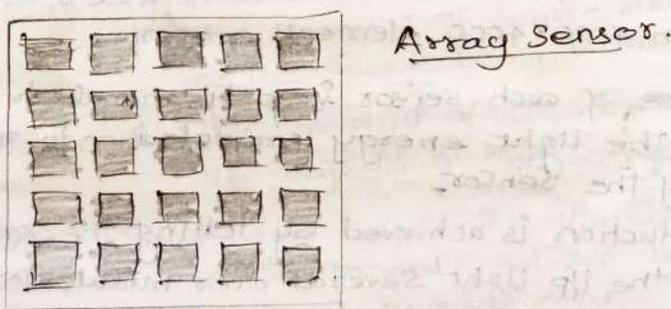
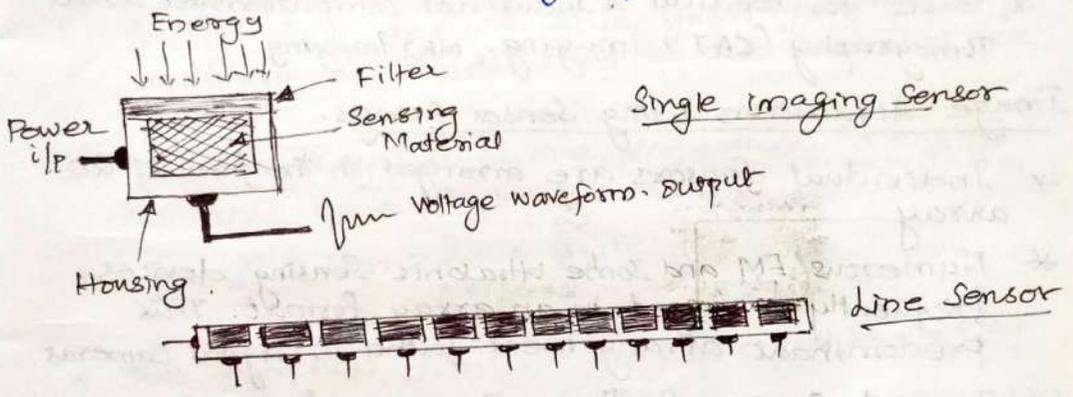


Image Acquisition Sensing using Single Sensor.

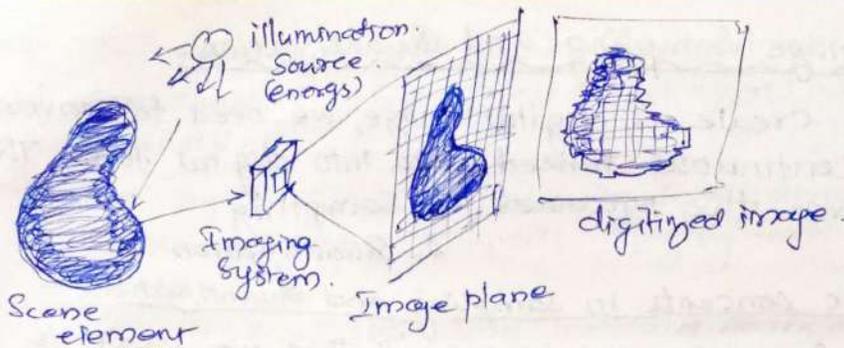
- * familiar sensor of this type is the photodiode.
- * Constructed of silicon materials & whose o/p voltage waveform is proportional to light.
- * Filter - in front of the sensor improves Selectivity.
- * In order to generate a 2D image using a single sensor, relative displacement in both x & y axis between the sensor and the area to be imaged.
- * Another example, moving mirrors are used to control the outgoing beam in scanning pattern and to direct the reflected laser signal onto the sensors.

Image Acquisition using Sensor Strips:-

- * The Strip provides imaging elements in one direction, motion perpendicular to the Strip provides imaging in the other direction. This type of arrangement used in most flat bed Scanners.
- * Airborne imaging applications.
- * Sensor Strips mounted in a ring configuration are used in medical and industrial imaging to obtain cross-sectional (slice) images of 3D objects.
- * Basis for medical & Industrial computerized axial Tomography (CAT) imaging, MRI imaging.

Image Acquisition using Sensor Arrays.

- * Individual Sensors are arranged in the form of 2D array
- * Numerous EM and some ultrasonic Sensing devices frequently arranged in an array format. This predominant arrangement found in Digital Cameras.
- * Typical Sensor for these Cameras is a CCD array. Packaged in 4000x4000 elements are more.
- * The response of each sensor is proportional to the integral of the light energy projected onto the Surface of the Sensor.
- * Noise reduction is achieved by letting the sensor integrate the light several over minutes (or) even hours.
- * The principal manner in which the array sensors the energy from an illumination source being reflected from a scene element.
- * The first f_1 performed by imaging system - collect the incoming energy and focus it onto an image plane.
- * The second f_2 . The sensor array which is coincident with 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 an analog signal, which is then digitized by another section of the imaging system.



A simple Image formation model:-

- * Images by 2D functions of the form $f(x, y)$.
- * The value/amplitude of 'f' at spatial coordinates (x, y) is a positive scalar quantity, whose physical meaning is determined by the source of the image.
- * The intensity values are proportional to energy radiated by a physical source.
- * $f(x, y)$ is non zero & finite, i.e.,

$$0 < f(x, y) < \infty$$
- * The function $f(x, y)$ is 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.
- * The above are called the 'illumination' and 'reflectance' components and are denoted by $i(x, y)$ & $r(x, y)$ respectively.
- * The two functions combine as a product to form $f(x, y)$.

$$f(x, y) = i(x, y) r(x, y)$$

where, $\begin{cases} 0 < i(x, y) < \infty \text{ and} \\ 0 < r(x, y) < 1 \end{cases}$
- The $r(x, y)$, that reflections is bounded by 0 (total absorption) and 1 (total reflectance)
- * $i(x, y)$ is determined by the illumination source, $r(x, y)$ is determined by the characteristics of the imaged objects.

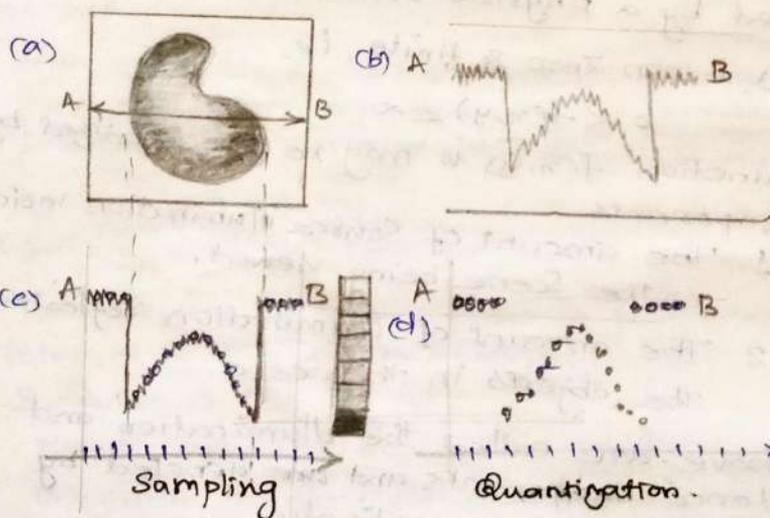
⑦ Image Sampling and Quantization.

To create a digital image, we need to convert the continuous sensed data into digital form. This involves two processes: 1. Sampling
2. Quantization.

Basic concepts in Sampling and Quantization: -

A continuous image 'f' that we want to convert to digital form. An image may be continuous with respect to the x and y coordinates and also in amplitude.

- * To convert it to digital form, we have to sample the function in both coordinates and in amplitude.
- * Digitizing the coordinates values is called Sampling.
- * Digitizing the amplitude values is called quantization.



- * The one dimensional function is a plot of amplitude (intensity level) values of the continuous image along the line segment AB in above figure.
- * To sample this function we take equally spaced samples along the line AB.
- * The spatial location of each sample is indicated vertical line in the bottom of the figure.
- * Samples are shown small white circles superimposed on the function.
- * The set of these discrete location gives the sampled function.

* In order to form a digital f_d , the intensity values are must be converted (quantized) into discrete quantities.

* The fig(c) shows the intensity scale divided into eight discrete intervals, ranging from black to white.

* The vertical tick line indicate the specific value assigned to each of the eight intensity intervals.

* The continuous intensity levels are quantized by assigning one of the eight values to each sample.

⇒ The method of sampling is determined by the sensor arrangement used to generate the image.

⇒ when an image is generated by single sensing element combined with mechanical motion.

* output of the sensor is quantized,

* Spatial sampling is accomplished by selecting the number of individual mechanical increments at which we activate the sensor to collect data.

⇒ when a sensing strip is used for imaging, the number of sensors in the strip establishes the sampling function limitations in one image direction. Mechanical motion in other direction can be controlled more accurately.

* The sampling limits established by the number of sensors in the other.

* Quantization of the sensor outputs completes the process of generating a digital image.

⇒ clearly, the quality of a digital image is determined to a large degree by the number of samples and discrete intensity levels used in sampling and quantization.

Representing Digital Images:-



⇒ let $f(s, t)$ represent a continuous image f_c of two continuous variables s & t

⇒ Now converting $f(s, t)$ into digital image by sampling and quantization.

* Now Suppose we sample continuous image into 2D array $f(x, y)$, containing M rows and N columns. where, (x, y) discrete coordinates.

* i.e, $x = 0, 1, 2, \dots, M-1$
 $y = 0, 1, 2, \dots, N-1$

Thus at value of digital image at the origin is, $f(0, 0)$, and next coordinates $f(0, 1)$

where $f(0, 1)$ - Second sample along the first row.
- It does not mean the physical coordinates.

* In general the value of the image at any coordinates (x, y) is denoted $f(x, y)$

* The section of the real plane spanned by the coordinates of an image is called the Spatial domain, where x & y being referred as Spatial variables (or) Spatial coordinates.

* Image displays allow us to view results at a glance.

* Numerical arrays are used for processing and algorithm development.

* In equation form, the representation of an $M \times N$ numerical array as,

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

- The Right side is a matrix of real numbers

- Each element of this matrix is called an image element, picture element, pixel (or) pel.

- in terms of traditional matrix notation to denote DIP

$$A = \begin{bmatrix} a_{0,0} & a_{0,1} & \dots & a_{0,N-1} \\ \vdots & \vdots & \dots & \vdots \\ a_{M-1,0} & a_{M-1,1} & \dots & a_{M-1,N-1} \end{bmatrix}$$

* This digitization process requires that decisions be made regarding the values for M, N and for the number L of discrete intensity levels.

- * Due to storage and quantizing hardware considerations, the number of intensity levels typically is an integer power of 2.

$$L = 2^k$$

- * The number, 'b' of bits required to store a digitized image is,

$$b = M \times N \times k \quad ; \quad \text{if } M = N \\ \text{then } b = N^2 k$$

Image Interpolation:-

Interpolation is a basic tool used extensively in tasks such as Zooming, Shrinking, rotating and geometric corrections.

- * Interpolation apply it to image resizing, which are basically image resampling methods.

- * Fundamentally, interpolation is the process of using known data to estimate values at unknown locations

let

$$500 \times 500 \text{ pixels} \rightarrow 1.5 \text{ times} \rightarrow 750 \times 750 \text{ pixels}$$

- * The above process method is called "nearest neighbor interpolation"

i.e, it assigns to each new location the intensity of its nearest neighbor in the original image.

- * A more suitable approach is "Bilinear interpolation"

In which we use the four nearest neighbors to estimate the intensity at a given location.

let (x, y) denote the coordinates of the location to which we want to assign an intensity value, and let $v(x, y)$ denote the intensity value.

for bilinear interpolation, the assigned value is obtained using the equation,

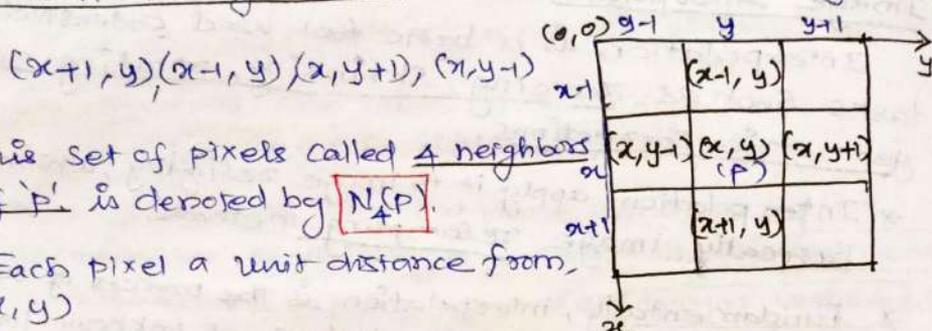
$$v(x, y) = ax + by + cxy + d$$

⑧ Relationships between pixels:-

- * An image is denoted by $f(x, y)$
- * Referring particular pixel, we use lowercase letters, such as 'p' and 'q'

Neighbors of a pixel

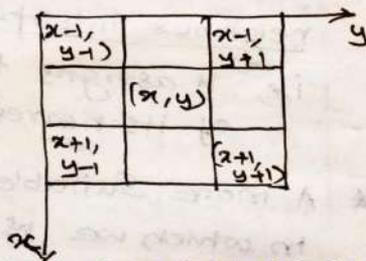
- * A pixel p at coordinates (x, y) has four horizontal and vertical neighbors whose coordinates are given by,



- This set of pixels called 4 neighbors of 'p' is denoted by $N_4(p)$.
- Each pixel a unit distance from, (x, y)
- Some of the neighbors location of p lie outside the digital image, if (x, y) is on the border of the image.
- * Four diagonal neighbors of 'p' have coordinates

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

are denoted by $N_D(p)$



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

Adjacency:- Two pixels are connected, if they are neighbors and their gray levels satisfy some specified criterion of similarity

- * let V be the set of intensity values used to define adjacency.
- * Binary image, $V = \{1\}$ if we referring to adjacency of pixels with value 1.
- * Gray scale image:

V - typically contains more elements $\{0, \dots, 255\}$

In the adjacency of pixels with range of possible intensity values 0-255, set 'V' could be any subset of 256 elements/values.

Consider Three types of Adjacency.

1. 4-adjacency: - two pixels 'p' & 'q' with values from V are 4-adjacent if 'q' is in the set $N_4(p)$
 $\Rightarrow q \in N_4(p)$

2. 8-Adjacency: Two pixels 'p' and 'q' with values from V are 8-adjacent if 'q' is in the Set of $N_8(p)$ i.e., $q \in N_8(p)$

3. m-adjacency (Mixed adjacency). Two pixels p and q with values from V are

m-adjacent if $\rightarrow q \in N_4(p)$

- (i) $q \in N_4(p) \leftarrow$ (i) 'q' is in $N_4(p)$, (or)
 (ii) $q \in N_D(p)$ and the Set $N_4(p) \cap N_4(q)$ has no pixels whose values are from V .
 (iii) $q \in N_D(p) \& N_4(p) \cap N_4(q) = \emptyset$ (empty)

(Mixed Adjacency - is a modification of 8-adjacency

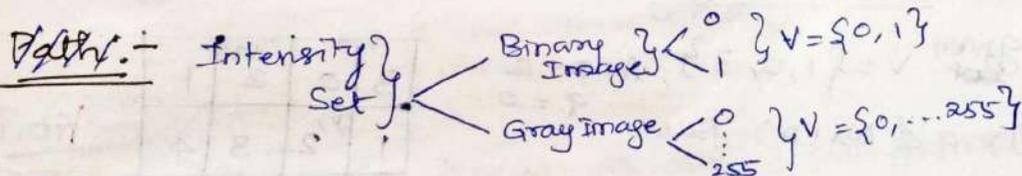
"introduced to eliminate the ambiguities that often arise when 8-adjacency is used.

Ex. Pixel arrangement as shown below. for

$$V = \{1\}$$

0	1	1
0	1	0
0	0	0

— m-adjacent
 — 8-adjacent



let x

Example:- 4 adjacency:

given $V = \{1, 2, 3\}$ pixel arrangement is

1	4	5
1 ^(p)	2	3
1	3	2 ^(q)

\Rightarrow let $p=2$
 $q=2$

$$\Rightarrow N_4(p) = \begin{array}{|c|c|c|} \hline 1 & 4 & 5 \\ \hline 1 & 2 & 3 \\ \hline & 3 & \\ \hline \end{array}$$

$$\therefore q \notin N_4(p)$$

~~but~~ let $p=2$ (center pixel.)
 $q=3$

Then, $q=3$ is available in $N_4(p)$.

\therefore p & q are 4 adjacency pixels.
 i.e., 2 & 3.

Ex. 8 adjacency

Given value of set $V = \{1, 2, 3\}$

1	2	3
1	4	5
3	2	1

\Rightarrow let $p=2$
 $q=5$

Condition 1

$q \in N_8(p)$

$5 \in N_8(p)$ Yes it is present

Condition 2

$q \in V$ (given set)?

$5 \notin \{1, 2, 3\}$ No

\therefore P & q are not 8-adjacency

\Rightarrow let $p=2$; $q=3$

Condition

$q \in N_8(p)$

$q \in V$ (intensity value set)

Disadvantage of 8-adjacency is, it will create multiple paths \rightarrow So Introduced m-adjacency.

Ex. m-adjacency :-

\Rightarrow Given ~~set~~ $V = \{1, 2, 3\}$, let $p=2$
 $q=0$

(i) $q \notin N_4(p)$ false.

(ii) $q \in N_D(p)$ yes.

but $q \notin V$ $\{0 \notin \{1, 2, 3\}\}$

\therefore it is also false.

$q \notin N_D(p)$

\therefore pixel P & q, are not m-adjacency.

3	0	2 ^(q)	1
1	2 ^(p)	3	4
4	1	0 ^(q)	3
1	1	3	2

\Rightarrow let $p=2$
 $q=2$

Condition

(i) is it $q \in N_4(p)$, no its not belongs to $N_4(p)$
 $\neq 2 \notin N_4(2)$

So it is false.

Condition

(ii) $q \in N_D(p) \Rightarrow$ yes it is belongs to $N_D(p)$
and $q \in V$ yes its belongs.

and also we want to check the condition (in addition to)

Intersection.

$$\Rightarrow N_4(P) \cap N_4(Q) = \phi$$

$N_4(P)$ in the diagram is, bounded. \odot

Then $N_4(Q)$ in the diagram double bounded.

3	\odot	2 ^Q	1
1	2 ^P	\odot	4
4	1	0	3
1	1	3	2

$$\therefore N_4(P) \cap N_4(Q) = \{0, 3\} \notin V \Rightarrow \notin V$$

In intensity value set $V = \{1, 2, 3\}$, Common pixel between $N_4(P)$ & $N_4(Q)$ is present i.e., '3'.

$$\therefore N_4(P) \cap N_4(Q) \neq \phi$$

$$\Rightarrow q = 1$$

$$p = 2$$

Condition (i) $q \in N_4(P)$
yes it is belongs.

Condition (ii) $q \in N_4(P)$.

No

and. $N_4(P) \cap N_4(Q) = \phi$. Yes there is no common pixels between $N_4(P)$ & $N_4(Q)$.

\therefore pixels P & Q are m-adjacency.

3	\odot	2	1
\odot	2 ^P	\odot	4
4	1	0	3
1	1	3	2

Path:-

\Rightarrow A path is also known as digital path/curve.

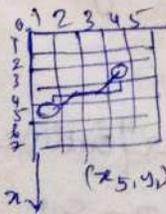
\Rightarrow A path from pixel 'p', with coordinate (x, y) to pixel 'q' with coordinates (s, t) is defined as the sequence of different pixels with coordinates

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

$$\text{where, } (x_0, y_0) = (x, y)$$

$$(x_n, y_n) = (s, t) \text{ and.}$$

(x_i, y_i) & (x_{i-1}, y_{i-1}) are adjacent for $1 \leq i \leq n$



Path length:-

path length is the no. of pixels present in a path. It is given by the value of 'n'

Closed path:-

In a path, if $(x_0, y_0) = (x_n, y_n)$ i.e., the first and last pixel are the same.

Types:-

According to adjacency present, paths can be classified as,
i) 4-path, (ii) 8-path (iii) m-path.

Connectivity:-

Connectivity between pixels is a fundamental concept of DIP.

Condition:-

Two pixels are said to be connected if

- (i) They are neighbors (and)
- (ii) Their gray levels satisfy the specified similarity criterion. (If their gray levels are equal).

Definition:-

Let S' represent subset of pixels in an image.

Two pixels p & q are said to be connected in S' , if there exists a path b in S' which consists of all the pixels in S' .

Connected Component:-

For any pixel p in S' , the set of pixels that are connected to p is called a "Connected Component of S' ".

Connected Set
If the pixel p has only one connected component, then set S' is called a "Connected Set."

Region:-

Let R subset of pixels in an image.

If R is a connected set, it is called a region of the image.

Boundary:-

* Boundary is also known as border / contour / outline

* The boundary of a region R is defined as the set of pixels in the region, that has one (or) more neighbors which are not in the same region R .

If R is an entire image, its boundary is defined as the set of pixels in the first & last rows & columns of the image.

Distance Measures:-

Various distance measures are used to determine the distance b/w different pixels

Conditions:-

Three pixels - P, Q, Z .

$$P(x, y), Q(s, t), Z(v, w)$$

for these 3 pixels D - is a Distance fn / metric.

if (a) $D(P, Q) \geq 0$; $D(P, Q) = 0$
if $P = Q$

(b) $D(P, Q) = D(Q, P)$

(c) $D(P, Z) \leq D(P, Q) + D(Q, Z)$

Important Distance Measures:-

- (i) ~~Euclidean~~ Euclidean distance $D_e(P, Q)$
- (ii) City-Block (or) D_A Distance
- (iii) chess board (or) D_B Distance
- (iv) D_m distance.

1. Euclidean distance:- $D_e(P, Q)$

The euclidean distance between two pixels P, Q is defined as,

$$D_e(P, Q) = \sqrt{(x-s)^2 + (y-t)^2}$$

The two pixels locations are
 $P(x, y)$
 $Q(s, t)$

The pixels having euclidean distance $\leq r$ from (x, y) form a disk shape with radius r centered at (x, y) .

2. City-Block / D_A Distance:-

$$D_A(P, Q) = |x-s| + |y-t|$$

* $D_A(P, Q) \leq r$ (Some value).

* ~~It~~ form a diamond shape centered at (x, y)

3. Chessboard / D8 distance! - b/n p & q pixels.

$$D_8(p, q) = \max(|x-s|, |y-t|)$$

- * The pixels with D_8 distance $\leq r$ (Some value)
- * If form a square centered (x, y)

```

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
    
```

4. Dm distance! -

D_m distance ^{between two points} is defined as the "Shortest m-path" between the [^]points. It considers m-adjacency.

Here, distance b/n two pixels depends on,

- * The value of the pixels along the path and
- * The values of their neighbors.

Ex. Consider the pixel arrangement given below,

0	P_3	P_4
P_1	P_2	0
P	0	0

assume the pixel values are

$$P = P_2 = P_4 = 1 \quad \& \quad P_1 = P_3 = 0 \text{ (or) } 1.$$

If adjacency of pixel values is 1!

i.e., $V = \{1\}$ is considered.

Case (1) : if $P_1 = P_3 = 0$.

$\Rightarrow D_m$, distance b/n P & P_4 is

$$\Rightarrow \boxed{2}$$

```

0 0 1
0 1 0
1 0 0
    
```

Case (2) : - if $P_1 = 1$; $P_3 = 0$.

$\Rightarrow P_2$ & P are not m-adjacent,

$\therefore D_m$ distance b/n P & P_4 is,

$$\Rightarrow \boxed{3}$$

```

0 0 1
1 1 0
1 0 0
    
```

Case (3) - if $P_1 = 0$; $P_3 = 1$,

```

0 1 1
0 1 0
1 0 0
    
```

P_2 & P_4 are not m-adjacent,

$\therefore D_m$ distance b/n P & P_4 is,

$$\boxed{3}$$

Case (4)

$$P_1 = P_3 = 1,$$

$$\begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

→ Dm distance bln.

$$P \& P_4 = 4.$$

9. Color Models:

Fundamentals of color Image Processing,

Two broad categories.

(i) Pseudo color processing.

- color is assigned to particular monochromatic intensity
- outdated method.

(ii) Full color processing.

- using full color sensors,
eg. color TV camera, color scanner, etc.
- widely used nowadays.

Color Spectrum divided into six broad regions.

1. violet, blue, green, yellow, orange, and red.

Characteristics of Light:

describing by

(1) Achromatic / Monochromatic light

- light seen on B&W TV
- characterized by one attributes called its Intensity (or amount of gray level).

(2) Chromatic light

quality is described by the three basic quantities.

* Radiance - total amount of energy that flows from light source.
(W)

* Luminance - It is measure of the amt. of energy perceived from a light source by an observer.
(Lumence)

* Brightness - It is one of the key factors describing Color Sensation, which is achromatic notion of intensity.
↓
It is subjective descriptor

(3) Colors

The characteristics that are used to differentiate one color from other.

* Brightness: It gives chromatic notion of Intensity.

* Hue: It represents the dominant color perceived by an observer.

(Eg: object is in red color, it specifies the Hue)

* Saturation: It refers to the relative purity / the amount of white light mixed with color.

The degree of Saturation - \propto amount of white light added.

⇒ eg. pink = (red + white)

lavender = (violet + white) are less saturated

⇒ Hue & Saturation taken together are called = Chromaticity.

∴ Color characterized by brightness & Chromaticity

Primary & Secondary Colors:-

Primary $\left\{ \begin{array}{l} \rightarrow R \\ \rightarrow G \\ \rightarrow B \end{array} \right.$ Can produce when added with different proportions to produce secondary color.

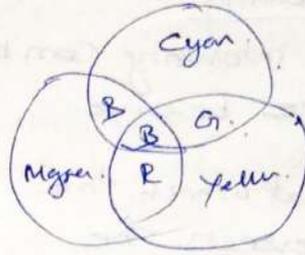
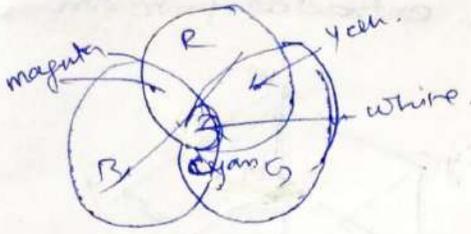
Magenta (R+B)

Cyan (G+B)

Yellow (R+G)

Mixture of Light

Primary color combination. to produce Secondary colors.



Color Models

No. of color models are use today,

(i) Hardware-oriented

- RGB
- CMY (Cyan, Magenta, Yellow)
- CMYK (Cyan, Magenta, yellow, Black).
for color printing.
- HSI (Hue, Saturation, Intensity) model.

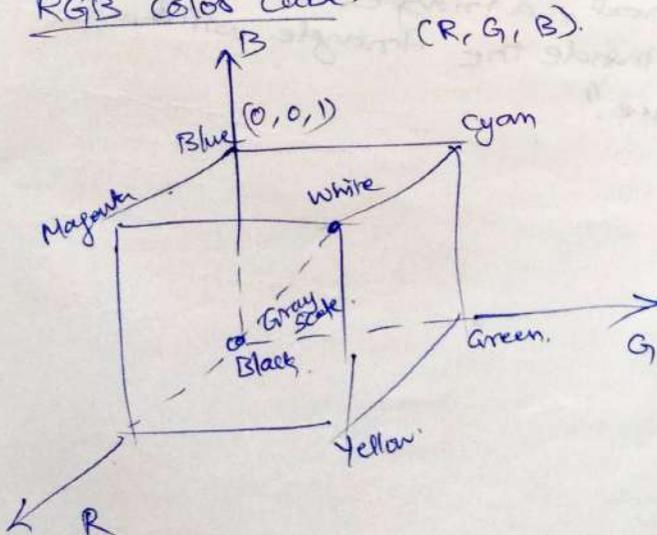
(ii) Application oriented. Color models.

RGB Color model.

all colors appear in their primary Spectral Components. of R, G, B

Images - its color represented by Three Components RGB.

RGB Color Cube.



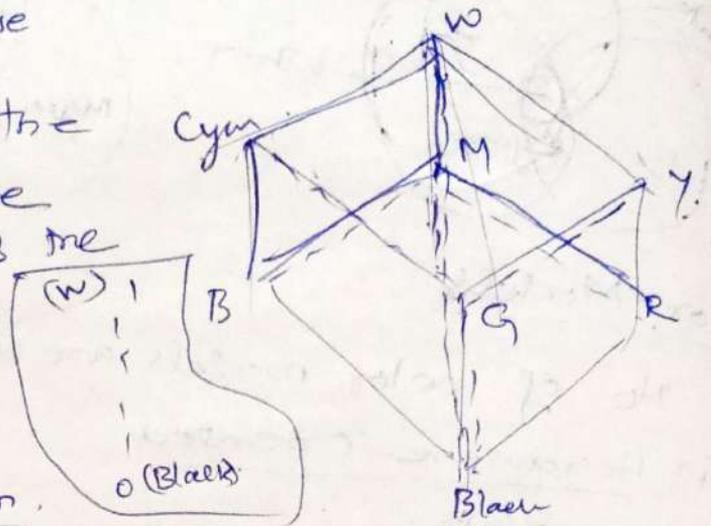
HSI

(1) To find intensity

The intensity can be extracted from an RGB image because

The point at which the plane intersects the intensity axis gives the intensity value -

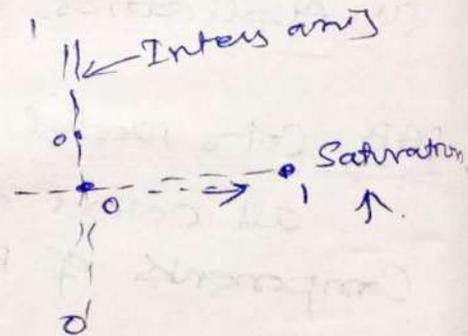
(0 - 1).



(2) To find Saturation

* all points on the intensity axis are gray which means that the Saturation, i.e. purity of points on the axis is zero.

* when the distance of color from the intensity axis increases, the Saturation of that color also increases



(3) To find Hue

" If three points namely, Black white & any color are joined, a triangle is formed. all the points inside the triangle will have the same hue."

UNIT-II
IMAGE ENHANCEMENT

Spatial Domain: -

Grey level Transformations - Histogram Processing - Basics of Spatial Filtering - Smoothing and Sharpening Spatial Filtering.

Frequency Domain: -

Introduction to Fouries transform - Smoothing and Sharpening frequency Domain filters - Ideal, Butterworth and Gaussian Filters.

1. Spatial Domain: -

Gray level transformations.

Introduction: -

Image enhancement is the process of highlighting certain features of interest.

- * Image enhancement - is a process on image so that the result is more suitable than the original image
- * details is in dark to brighten.
- * Enhancement is "subjective process" - means that human perception decides the best method from the obtained results.

Examples of Image enhancement.

- * Contrast Enhancement
- * Edge Enhancement
- * Noise Filtering.
- * Sharpening.
- * Magnifying.

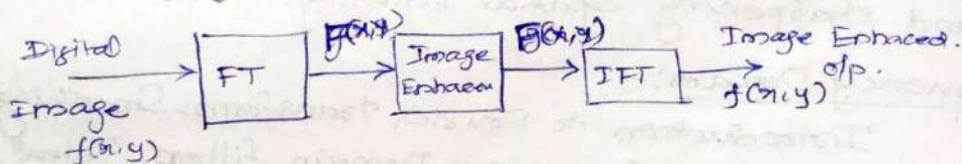
Methods.

Image Enhancement approaches,

- (1) Spatial Domain methods: → Spatial domain means Image plane itself) → This method, direct manipulation of pixels in an image.

2. Frequency Domain Methods

Converting an image from Spatial Domain to frequency Domain by Fourier Transforms and these methodologies/techniques are based on modifying the Fourier Transform of an image.



Spatial Domain Methods - Fundamentals:

The pixels which compose an image are referred as the spatial domain. The technique that directly operate on these pixels are known as the spatial domain methods.

The process is denoted as,

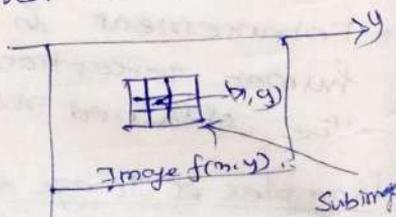
$$g(x,y) = T[f(x,y)]$$

↑
↑ i/p Image
 Processed Image

T - operator of f

Defining a Neighborhood

To define a neighborhood about a point/pixel (x,y) a square/rectangular subimage area which is centered at (x,y) is used



Procedure:-

- (1) Center of the subimage is moved pixel to pixel, starting at the top left corner of the image.
- (2) at each location the operator is applied to produce the output at that location (x,y) .

Gray level Transformation function, $T(r)$

$$s = T(r)$$

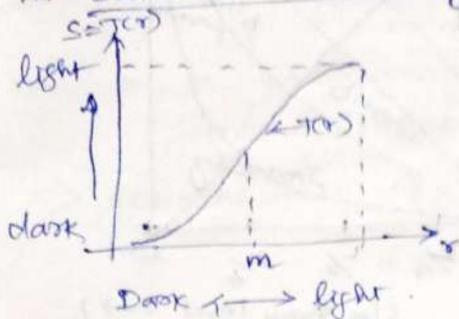
r - Gray level of $f(x,y)$ at (x,y)

s - Gray level of $g(x,y)$ at (x,y)

Based on the shape of $T(r)$ there are two categories of techniques which are used for Contrast Enhancement

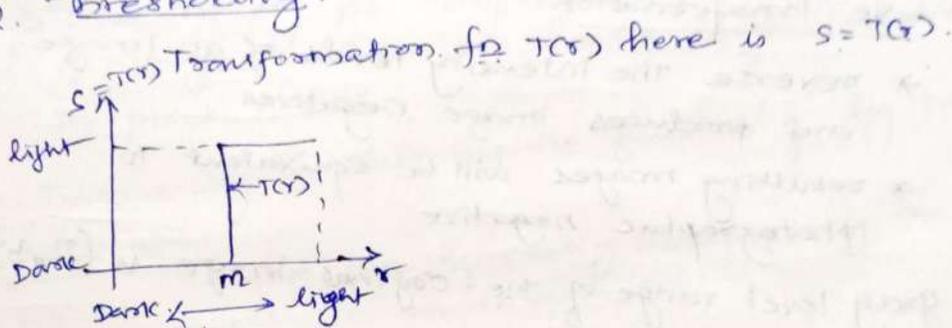
1. Contrast stretching.
2. Thresholding.

1. Contrast stretching.



- * Thus levels below 'm' are darkened
- * levels above 'm' are brightened.
- * results in higher contrast image than the original image.

2. Thresholding:-



3. Mask processing / filtering:-

A mask is a two dimensional array / subimage. A 3x3 mask also referred as a filters / kernels / templates (or) window.

- * The values of the mask coefficients are selected according to the process done.
- * The center of the mask coefficients moved from pixel to pixel.

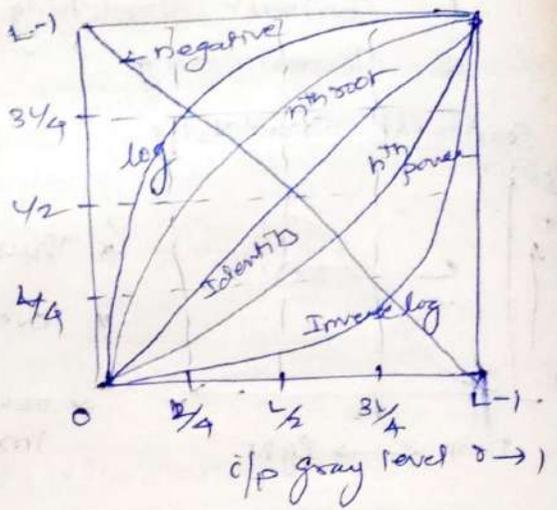
Gray level transformations.

$$s = T(r)$$

Types:

- (i) Linear transformation
 - (a) -ve "
 - (b) Identity transformation
- (ii) Logarithmic transformation.
 - (a) log transformation
 - (b) anti log / inverse log transformation
- (iii) Power-law
 - (a) nth power
 - (b) nth root
- (iv) piecewise - Linear transformation function.

Basic
Gray
level
transform
func
↑
s
o/p
gray
level



2.3.1. -ve Transformation.

- * reverse the intensity levels of an image and produces image negative.
- * resulting images will be equivalent to photographic negative.

⇒ gray level range of the original image is $(0, L-1)$

$$s = L - 1 - r$$

Eq: $[0 - 8]$ $L=9$

i/p
Image:

0	1	3
5	2	4
6	8	7

$$s = L - 1 - r$$

$$= 9 - 1 - 0$$

$$s = 8$$

$$r = 1$$

$$s = 7$$

$$r = 2$$

$$s = 6$$

$$r = 3$$

$$s = 5$$

o/p image

8	7	5
3	6	4
2	0	1

2. Log Transformation

* a narrow range of a gray-level values in the low ip Image into a wider range of o/p levels and.

* a wider range of high gray-level values into a narrow range of o/p level.

⇒ general form of this transformation is given by.

$$s = c \log(1+r) \quad r \geq 0$$

$c = \text{constant}$

⇒ Thus the log fn compresses the dynamic range of images which are having large variations in pixel values.

Inverse log fn. Just opposite function of log transform.

Application used for spreading/compressing the gray levels in an image.

(3) power law Transformations:-

The basic form is given by,

$$s = c r^{\gamma}$$

also denoted by

$$s = c (r + \epsilon)^{\gamma}$$

c & $\gamma \rightarrow$ +ve constants.

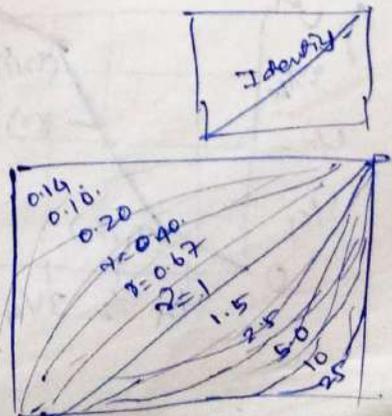
$\epsilon \rightarrow$ Represents offset

when

$c = \gamma = 1 \Rightarrow \therefore s = r$ Identity transformation.

Power law curves

By varying the value of γ and setting $c=1$



- * Curves with fractional values of $\gamma = 0.67, 0.40, 0.20$. map a narrow range of dark i/p values into a wider range of o/p values.
- * Also they map a wide range of higher values into a narrow range of o/p values.
- * $\gamma > 1$ curve opposite effect of $\gamma < 1$

Applic

Gamma Correction.
 General purpose contrast manipulation.
 Enhancing images with washed out appearance.

Gamma Correction.

The ~~ex~~ Exponent γ is known as gamma.

Piece-wise linear transform γ

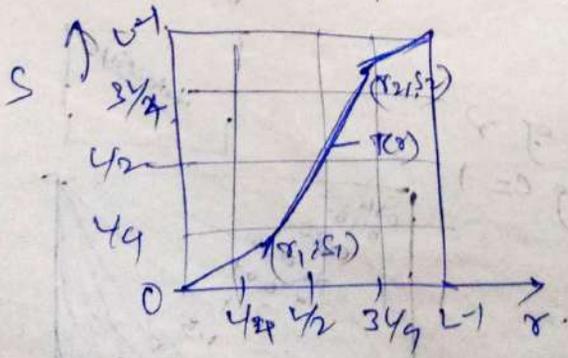
opposite to above methods.
 completely different approach.

Three types

- (1) Contrast stretching.
- (2) Gray level slicing.
- (3) Bit plane slicing.

Contrast stretching.

Converting low-contrast to high contrast.
 due to poor illumination.



location of the pixel (r_1, s_1) (r_2, s_2) decides transformation function.

if $r_1 = s_1$ & $r_2 = s_2$.

The transformation is a linear f_m .
it will not change.

if $r_1 = r_2$, $s_1 = 0$ & $s_2 = L-1$.

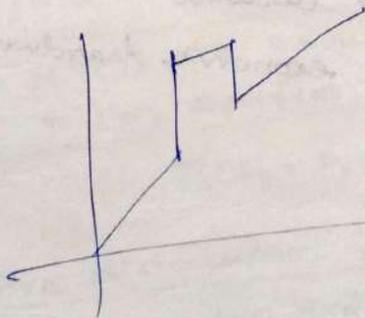
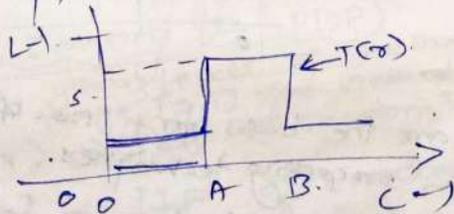
Transformation becomes thresholding f_m . It produces a binary (00) two level image.

if intermediate values of (r_1, s_1) & (r_2, s_2) , change the contrast of the image by producing different levels of gray level spread in the original image.

\therefore Transformation f_m $r_1 \leq r_2$ & $s_1 \leq s_2$ is preferred

Gray level slicing

Two basic approaches to perform gray-level slicing



Histogram Processing.
 Histogram of an image is a plot/graph drawn between gray level values (0-255) in x-axis & no. of pixels having the corresponding gray levels in y-axis.
 $h(r_k) = n_k$
 $r_k \Rightarrow k^{\text{th}}$ gray level.
 $n_k \Rightarrow$ Number of pixels in the image having gray level r_k .

Normalized Histogram.
 is obtained by dividing each value of histogram by the total no. of pixels in the image, (0-L-1)

$$P(r_k) = \frac{n_k}{n}; k=0, 1, \dots, L-1$$

$n =$ total No. of pixels in the image. $(L-1 \times L-1)$

$P(r_k)$ - gives the probability of occurrence of gray level r_k .

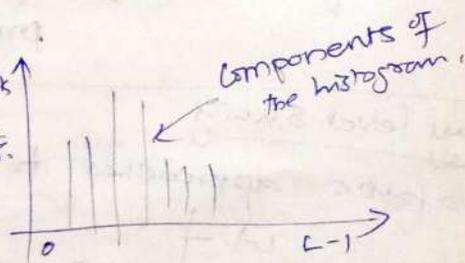
also sum of all components of a normalized histogram is equal to 1.

Ex. Histogram is

in which.

$$f(r_k) = n_k$$

$$P(r_k) = \frac{n_k}{n}$$



Adv

- (1) Histograms are the basis for a no. of spatial domain processing techniques.
- (2) They are simple to calculate in software.
- (3) Histograms provide economic hardware implementation.

Appli

- Image Enhance
- Compression
- Segmentation.

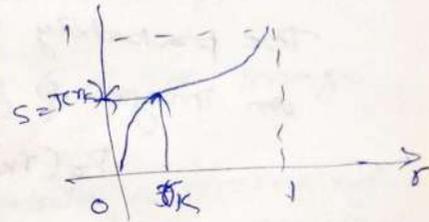
Histogram equalization / Linearization.

" is a process of automatically determining a transformation function which produces an output image with a uniform histogram. This transformation f_{tr} is different for continuous and discrete.

Histogram eq'n. for Continuous f_{tr}

Let r represent the gray levels of the image to be enhanced. r is normalized to the interval $[0, 1]$ where $r=0$ represent black and $r=1$ represent white.

* Transformation function.
 $s = T(r) \quad 0 \leq r \leq 1$



* Inverse Transform.

$$r = T^{-1}(s) \quad 0 \leq s \leq 1$$

* Inverse transformation from s to r is denoted by

* Probability Density f_{tr} (PDF)
Since the gray levels of the image which are in the range $[0, 1]$ are viewed as random variables.

Let $P_r(r) \Rightarrow$ PDF of random variable r .

$P_s(s) \rightarrow$ PDF of random variable s .

By elementary probability theory.

If $P_r(r)$ & $T(r)$ are known and $T^{-1}(s)$ satisfies condition

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

* Cumulative distribution function (CDF)
The CDF of random variable r is given by,

$$s = T(r) = \int_0^r P_r(w) dw$$

$w \rightarrow$ dummy variable of integration.

To find $P_s(s)$,

The above equation, derive w.r.t. r .

$$\frac{ds}{dr} = \frac{dT(r)}{dr} = \frac{d}{dr} \left[\int_0^r P_r(w) \cdot dw \right] = P_r(r)$$

Leibniz's Rule:- Derivative of a definite integral with respect to its upper limit = Integral evaluated at that limit

$$\frac{ds}{dx} = P_0(x) \cdot \text{Integral evaluated at that limit}$$

$$P_0(s) = P_0(r) \left| \frac{dr}{ds} \right| =$$

$$= P_0(r) \left[\frac{1}{P_0(r)} \right]$$

$$P_0(s) = 1 \quad ; \quad 0 \leq r \leq 1$$

Histogram Equalization for Discrete values

Let, The pixel values are in the interval $[0, L-1]$
The probability of occurrence of gray level r_k in an image is given by.

$$P_0(r_k) = \frac{n_k}{n} \quad ; \quad k = 0, 1, 2, \dots, L-1$$

where, n = total no. of pixels in the image.

n_k = Number of pixels having gray level r_k .

Transformation function:-

The discrete form of the transformation fn

$$S_k = T(r_k) = \sum_{j=0}^k P_r(r_j)$$

$$= \sum_{j=0}^k \frac{n_j}{n} \quad ; \quad \text{where } k = 0, 1, \dots, L-1$$

Thus mapping each pixel with level r_k in the i/p image into corresponding pixel with level S_k in the o/p image using above eqn.

The plot of $P_0(r_k)$ versus r_k is known as histogram and transformation (or) mapping given above eqn. is called histogram equalization / linearization.

Adv

- (i) It is very simple to implement
- (ii) The ~~if~~ information, it needs can be obtained directly from the given images and no additional parameter specification are required.
- (iii) The results from this technique are predictable.
- (iv) It is fully automatic.

Histogram Matching (or) Histogram Specification.

- * Histogram equalization is not suitable for some applications.
- * Such cases, histogram shape of the o/p image may be specified.
- * The method used to generate a processed image that has a specified histogram is called histogram matching/specification.

Implementation.

Let $r \rightarrow$ Gray levels of the i/p image.

$z \rightarrow$ Gray levels of the o/p image.

$P_r(r) \Rightarrow$ Continuous PDF of r .

$P_z(z) =$ " " " z .

To obtain image with specified PDF.

* $P_z(z)$ is the specified PDF that can be obtained by

1) Obtain transformation $T(r)$ using the equation,

$$S = T(r) = \int_0^r P_r(w) dw.$$

2) Obtain the transformation $f_z G(z)$ using the equation,

$$G(z) = \int_0^z P_z(t) dt = S.$$

3) Get the inverse transformation $f_r^{-1} G^{-1}$.

1) Apply the following eqs to all the pixels in the c/p image to get the o/p image.

$$z = G^{-1}(s)$$

$$= G^{-1}[T(r)]$$

Procedure for Histogram Matching: -

The overall procedure of histogram matching is follow

- 1) Obtain the histogram of the given image.
- 2) pre compute a mapped level "s_k" for each level "r_k" using the equation below.

$$s_k = T(r_k) = \sum_{j=0}^k P_r(r_j)$$

$$= \sum_{j=0}^k \frac{n_j}{n} ; k = 0, 1, 2, \dots, L-1$$

- 3) Get the transformation f_z = G_z from the given p_z(z) using the equation

$$v_k = G_z(z_k) = \sum_{j=0}^k p_z(z_j) = s_k$$

- 4) pre compute z_k for each value of s_k using the iterative scheme defined by,

$$[G(z) - s_k] \geq 0 \quad ; \quad k = 0, 1, 2, \dots, L-1$$

where

z = Smallest integer in the interval (0 - L-1)

- 5) This step perform two mapping using the pre computed values from (2) & (4).

Local Histogram Processing.

Global approaches.

⇒ In global histogram processing methods, the transformation function which modifies the pixel is based on the gray-level content of an entire image.

* These methods are suitable only for overall enhancement.

* Ex. (i) Histogram Equalization.

(ii) Histogram Matching.

Procedure: local histograms processing has 5 steps.

(i) ~~Square~~ Square/rectangular neighborhood is defined.

(ii) The center of this neighborhood area is moved from pixel to pixel.

(iii) at each location, the histogram of the points in the neighborhood is computed and either histogram equalization/specification transformation function is obtained.

(iv) Now the gray-level of the neighborhood center pixel is mapped by the obtained transformation function.

(v) The center of the neighborhood region is moved to an adjacent pixel location and the steps (i) to (iv) are repeated.

Histogram Statistics for Image Enhancement.

Some statistical parameters obtained directly from the histograms can be used, instead of Image Enhancement.

(i) Moment

(ii) Mean.

(iii) Variance.

(iv) Standard Deviation.

(1) Moment.

If 'r' is a random variable representing discrete gray-level in the range $[0 \rightarrow L-1]$,

The n^{th} moment of 'r' about its mean is,

$$\mu_n(r) = \sum_{i=0}^{L-1} (r_i - m)^n P(r_i) \quad \text{--- (1)}$$

where $P(r_i)$ = probability of occurrence of gray level r_i .

m - avg. gray level of r (mean of r)

$$m = \sum_{i=0}^{L-1} r_i P(r_i) \quad \text{--- (2)}$$

from eq'n (1) & (2).

0th moment of r ; $\mu_0 = 1$

1st " " r ; $\mu_1 = 0$

2nd " " r ; μ_2 is known as Variance of

'r' is given by

$$\sigma^2(r) = \mu_2(r) = \sum_{i=0}^{L-1} (r_i - m)^2 P(r_i)$$

(2) Mean

where S_{xy} = Subimage of Specified size centered at (x, y)

$m_{S_{xy}}$

$$m_{S_{xy}} = \sum_{(s,t) \in S_{xy}} r_{s,t} P(r_{s,t})$$

$r_{s,t}$ \Rightarrow gray level at coordinate (s, t) in the neighborhood.

$P(r_{s,t})$ \Rightarrow Normalized neighborhood

(3) Variance:

$$\sigma_{S_{xy}}^2 = \sum_{(s,t) \in S_{xy}} [r_{s,t} - m_{S_{xy}}]^2 P(r_{s,t})$$

(4) Standard Deviation.

$$\sqrt{\sigma_{S_{xy}}^2}$$

Procedure to perform Histogram Equalization.

1. Find the running sum of the histogram values.
2. Normalize the values from Step(1), by dividing by the total no. of pixels.
3. Multiply the values from Step(2) by the maximum gray-level value and round.
4. Map the gray level values to the results from Step(3) using a one to one correspondence.

Ex: 1 Perform Histogram Equalization of an image.

$$\begin{bmatrix} 4 & 4 & 4 & 4 & 4 \\ 3 & 4 & 5 & 4 & 3 \\ 3 & 5 & 5 & 5 & 3 \\ 3 & 4 & 5 & 4 & 3 \\ 4 & 4 & 4 & 4 & 4 \end{bmatrix}$$

Solution.

from the image pixel. max. Value is 5, so we need min of 3 bits to represent the number.

∴ There are eight possible gray levels from 0 to 7.

$$\Rightarrow 000, 001, 010, 011, \dots, 111$$

$$[0, 1, 2, \dots, 7]$$

Q. Histogram of the I/p image is given below.

Gray level	0	1	2	3	4	5	6	7
No. of pixels	0	0	0	6	14	5	0	0

1. Compute the running sum known as cumulative frequency distribution.

Gray level.	0	1	2	3	4	5	6	7
No. of pixel.	0	0	0	6	14	5	0	0
Cumulative sum.	0	0	0	6	20	25	25	25

2. Divide the running sum. obtained in step(1). by total no. of pixels. In this case, total no. of pixel $5 \times 5 = 25$

Running sum / 25	0	1	2	3	4	5	6	7
	0/25	0/25	0/25	6/25	20/25	25/25	25/25	25/25

3. Multiply the result obtained in Step (2) by the maximum-gray-level value.

\therefore Maximum-gray level in this case = 7

\Rightarrow Gray level \Rightarrow 0 1 2 3 4 5 6 7

Multiply the above result by maximum gray level \Rightarrow $\frac{0}{25} \times 7$ $\frac{0}{25} \times 7$ $\frac{0}{25} \times 7$ $\frac{6}{25} \times 7$ $\frac{20}{25} \times 7$ $\frac{25}{25} \times 7$ $\frac{25}{25} \times 7$ $\frac{25}{25} \times 7$

④ The result is rounded to closest integer to get the following table.

gray level \Rightarrow 0 1 2 3 4 5 6 7

rounded to closest integer \Rightarrow 0 0 0 2 6 7 7 7

⑤ Mapping of gray level by a one to one correspondence.

Original gray level \Rightarrow 0 1 2 3 4 5 6 7

Histogram equalized values \Rightarrow 0 0 0 2 6 7 7 7

i/p image \Rightarrow Histogram equalized images

original \Rightarrow $\begin{bmatrix} 4 & 4 & 4 & 4 & 4 \\ 3 & 4 & 5 & 4 & 3 \\ 3 & 5 & 5 & 5 & 3 \\ 3 & 4 & 5 & 4 & 3 \\ 4 & 4 & 4 & 4 & 4 \end{bmatrix} \Rightarrow \begin{bmatrix} 6 & 6 & 6 & 6 & 6 \\ 2 & 6 & 7 & 6 & 2 \\ 2 & 7 & 7 & 7 & 2 \\ 2 & 6 & 7 & 6 & 2 \\ 6 & 6 & 6 & 6 & 6 \end{bmatrix}$

2. Describe histogram equalization for the following image, Segment size of 5x5? write the inference on image. Segment before and after equalization.

20 20 20 18 16
 15 15 16 18 15
 15 15 19 15 17
 16 17 19 18 16
 20 18 17 20 15

⇒ 1 Original Histogram.

Gray Value	0	1	2	3	4	5	6	7
No. of Pixels	7	4	3	4	2	5	0	0

2 Cumulative Histogram

Gray Value	0	1	2	3	4	5	6	7
Histogram Counter	7	7+4	$\frac{11+3}{14}$	18	20	25	25	25
	7	7+4	11+3	14+4	18+2	20+5	25+0	25+0

3 Gray value transformation (T)

Gray Value	0	1	2	3	4	5	6	7
Transformed	2	3	4	5	6	7	7	7

from ②
 H.C. Cont. ① $\frac{7}{25} \times 7 = 1.96 \approx 2$
 ② $\Rightarrow \frac{14}{25} \times 7 = 3.92 \approx 4$
 ① $\frac{11}{25} \times 7 = 3.08 \approx 3$

4 Histogram Size.

Gray Val	0	1	2	3	4	5	6	7
	$\frac{7}{25} \times 31$	$\frac{11}{25} \times 31$	$\frac{14}{25} \times 31$	$\frac{18}{25} \times 31$	$\frac{20}{25} \times 31$	$\frac{25}{25} \times 31$	31	31
	9	14	17	22	25	31	31	31

from ② given size 5x5 image
 w.c.T $L = 2^k$, Here $k=5$ length is 0 to L
 $L = 2^5 = 32$; $0 \rightarrow 31$

5

	Equalized Histogram.							
Gray Value	0	1	2	3	4	5	6	7
No. of pixels	0	0	7	4	3	4	2	5

Compare
Answer 1

$$\begin{vmatrix} 20 & 20 & 20 & 18 & 16 \\ 15 & 15 & 16 & 18 & 15 \\ 15 & 15 & 19 & 15 & 17 \end{vmatrix} = \begin{vmatrix} 4 & 4 & 4 & 4 & 0 \\ 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 3 & 0 & 7 \\ 0 & 7 & 3 & 4 & 0 \\ 4 & 4 & 7 & 4 & 0 \end{vmatrix}$$

Basics of Spatial Filtering :-

~~Spatial averaging~~

Spatial Convolution :-

used to perform the operation of zooming.

Median filter

Compute the median value of the marked pixels

$$\begin{vmatrix} 1 & 5 & 7 \\ 2 & 4 & 6 \\ 3 & 2 & 1 \end{vmatrix} \Rightarrow$$

Step 1. The pixel arrange in ascending.
1 1 2 2 3 3 4 5 6 7

2. median value of the ordered pixel is computed.
i.e, 3

it is replaced by center of the original

$$\begin{vmatrix} 1 & 5 & 7 \\ 2 & 4 & 6 \\ 3 & 2 & 1 \end{vmatrix} \Rightarrow \begin{vmatrix} 1 & 5 & 7 \\ 2 & 3 & 6 \\ 3 & 2 & 1 \end{vmatrix}$$

Q3) Compute the median value of the marked pixel.

$$\begin{vmatrix} 18 & 22 & 33 & 25 & 32 & 24 \\ 34 & 28 & 24 & 32 & 26 & 23 \\ 22 & 19 & 32 & 31 & 28 & 26 \end{vmatrix}$$

$$\Rightarrow \begin{vmatrix} 18 & 22 & 33 & 25 & 32 & 24 \\ 34 & 24 & 31 & 31 & 26 & 23 \\ 22 & 19 & 32 & 31 & 28 & 26 \end{vmatrix}$$

$$f - \bar{f} = fs$$

$$\bar{f} = fs - f$$

$$\bar{f} = f - fs$$

High boost filtering

$$I_{\text{high boost}} = A I_{\text{original}} + I_{\text{high pass}}$$

$$= (A W_{\text{all pass}} + W_{\text{high pass}}) * I_{\text{original}}$$

$$= W_{\text{high boost}} * I_{\text{original}}$$

where $W_{\text{high boost}} = A W_{\text{all pass}} + W_{\text{high pass}}$ is the high boost convolution kernel, and 'A' is constant.

$$W_{\text{high boost}} = A W_{\text{all pass}} + W_{\text{high pass}}$$

$$= A \begin{vmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{vmatrix} + \begin{vmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{vmatrix} = \begin{vmatrix} 0 & -1 & 0 \\ -1 & A+4 & -1 \\ 0 & -1 & 0 \end{vmatrix}$$

$$= A \begin{vmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{vmatrix} + \begin{vmatrix} \phi & -1 & \phi \\ -1 & 8 & -1 \\ \phi & -1 & \phi \end{vmatrix} = \begin{vmatrix} \phi & -1 & \phi \\ -1 & A+8 & -1 \\ \phi & -1 & \phi \end{vmatrix}$$

Unsharp Masking

$$f_s(m, y) = f(m, y) - \bar{f}(m, y) \rightarrow$$

\downarrow original image \downarrow blurred image

$$f_{\text{hp}}(m, y) = A f(m, y) - \bar{f}(m, y)$$

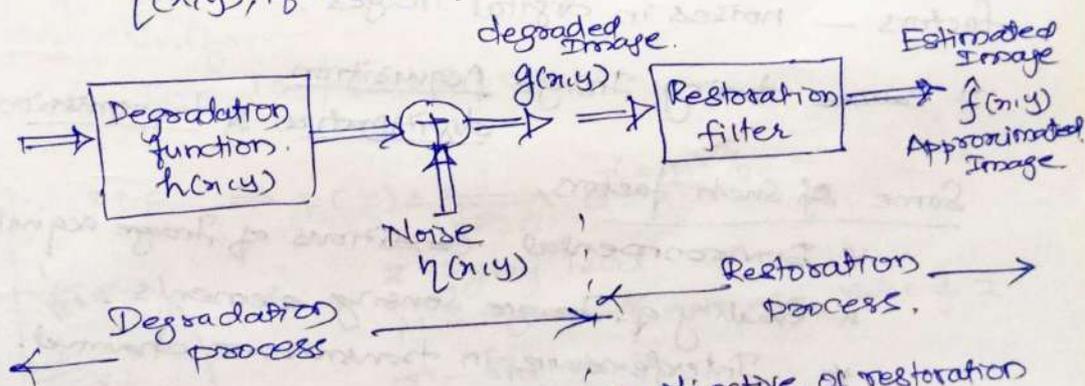
IMAGE RESTORATION AND SEGMENTATION

Image Restoration / Degradation Models.

- Noise Models - Mean filters - Order statistics - Adaptive filters - Band Reject filters - Band pass filters - Notch filters - Optimum notch filtering - Inverse filtering - Wiener filtering. Segmentation: Detection of Discontinuities - Edge Linking and Boundary detection - Region Based Segmentation - Morphological processing Erosion & dilation.

Degradation Models:-

- * Degraded Image $g(x,y)$
- * Some knowledge about the degradation $H(u,v)$.
- * Some knowledge about the additive noise $\eta(x,y)$, if needed.



- * It produces the degraded form of (Additive noise) of the original Image. $\Rightarrow g(x,y)$.
- * Objective of restoration process is to obtain Estimate/Approximation $\hat{f}(x,y)$ of original Image.
- * The approximation should be as close as possible to the original input image. $\Rightarrow \hat{f}(x,y)$.

Spatial Domain Representation: -

Let $H \Rightarrow$ linear, position invariant process,
The degraded image represented in the spatial domain is,

$$g(x, y) = h(x, y) * f(x, y) + \eta(x, y).$$

$h(x, y) \Rightarrow$ Spatial representation of degradation function.

$*$ \Rightarrow represents convolution process

Frequency Domain Representation: -

$$G(u, v) = H(u, v) \cdot F(u, v) + N(u, v).$$

Fourier transform of $h(x, y)$, $f(x, y)$, $g(x, y)$
& $\eta(x, y)$.

Noise Models.

factors - noises in digital images.

It causes during Image Acquisition,
digitization & transmission.

Some of such factors,

- * Environmental conditions of Image acquisition
- * Quality of Image Sensing elements.
- * Interference in transmission channel.

Freq. Properties of Noise.

freq. content of noise defines freq. properties of noise. to find these properties.

$$N \cdot \underline{F}^T \rightarrow N$$

White Noise.

If all freqs. of $f(x, y)$ are in equal proportions, its Fourier spectrum is said to be constant

Thus freq. spectrum of noise is constant is called white noise.

~~thus.~~

Spatial properties of noise.

assumed that no correlation b/w pixel values and the noise component values in an image.

This implies that the noise is uncorrelated with an image.

∴ noise is independent of spatial coordinates

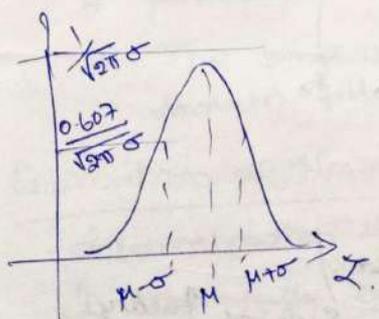
Noise PDF

The Statistical behaviour of the gray level values in the noise component is considered as random variable.

No Imp.

1. Gaussian/Normal Noise
2. Rayleigh Noise.
3. Erlang/Gamma.
4. Exponential
5. Uniform
6. Impulse / salt & pepper.

1. PDF : $\Rightarrow P(z) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(z-\mu)^2}{2\sigma^2}}$



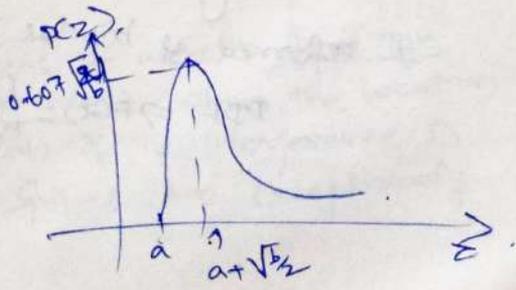
$z \rightarrow$ gray level.
 μ - mean of average value of z .
 σ - std. devia.
 $\sigma^2 \rightarrow$ Variance.

2. $P(z) = \begin{cases} \frac{2}{b}(z-a)e^{-\frac{(z-a)^2}{b}} & \text{for } z \geq a \\ 0 & z < a \end{cases}$

mean

$$\mu = a + \sqrt{\frac{\pi b}{4}}$$

$$\sigma^2 = \frac{b(4-\pi)}{4}$$



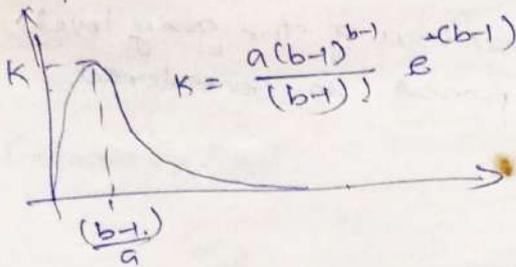
(3) PDF

$$f(z) = \begin{cases} \frac{a^b z^{b-1}}{(b-1)!} e^{-az} & z \geq 0 \\ 0 & z < 0 \end{cases}$$

$a, b \Rightarrow +ve \text{ integer } a > 0$

!.

$$M = b/a \quad ; \quad \sigma^2 = \frac{b}{a^2}$$

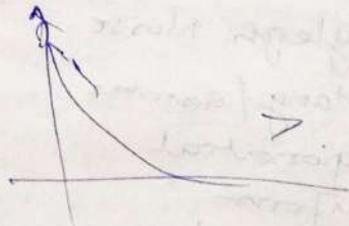


3.

$$f(z) = a e^{-az}$$

$$M = 1/a$$

$$\sigma^2 = 1/a^2$$

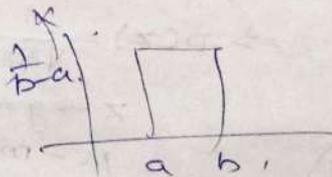


4. Uniform noise

$$f(z) = \frac{1}{b-a}$$

$$M = \frac{a+b}{2}$$

$$\sigma^2 = \frac{(b-a)^2}{12}$$



5. Impulse / Salt & pepper Noise :-

* Impulse (∞) short noise (spike noise.)

* occurred in quick transitions such as faulty switching. takes place during imaging.

* referred as bipolar impulse noise

$$\text{PDF} \Rightarrow f(x) = \begin{cases} P_a & \text{for } x = a \\ P_b & \text{for } x = b \\ 0 & \text{otherwise.} \end{cases}$$

based on PDF, there are three cases

Case (i) $b > a$

gray level $b \rightarrow$ appears like light dot

gray level $a \rightarrow$ appears like a dark dot in the image

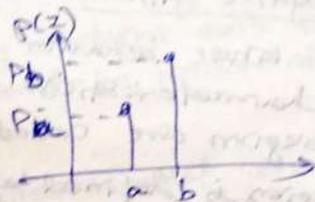
Case (ii) $P_a(\infty) P_b = 0$

now impulse noise is called unipolar noise

Case (iii) $P_a(\infty) P_b \neq 0$ & $P_a = P_b$

The impulse noise values in an image look like salt & pepper granules spread over the image. Therefore it is called Salt & Pepper noise.

PDF of Impulse



Periodic Noise:-

- * It is generated during image acquisition
- * Caused by electrical / electromechanical interference
- * periodic noise are spatially dependent
- * This kind of noise can be reduced in frequency domain filtering.

Estimation of Noise parameters

determination of various noise parameters

- (1) Carefully examining the Fourier Spectrum, the parameters of periodic noise can be detected by visual analysis.
- (2) Automated analysis.
Automated analysis is used to find the periodicity of noise components, when the location of the frequency components of the interference is known or when the noise spikes are exceptionally pronounced.

- (3). Parameters of PDF can also be estimated from the sensor specification. But only a partial estimation is possible.
- (4) If the imaging system is available, the system noise characteristics can be known by capturing a set of images with flat (or) plain environment. The system noise can be studied with the use of these images.
- (5) If the images captured by the sensor is the only available thing, ~~the~~

3.6. Adaptive filters.

- * The filters whose behavior changes based on statistical characteristics of the image inside the filter region are called "adaptive filters".
- * The filter region is defined by a window S_{xy} of size $m \times n$.
- * But complexity of adaptive filter is higher than other filter.
- * Two ~~is~~ important adaptive filters are.
 - (i) Adaptive, local noise reduction filter.
 - (ii) Adaptive median filter.

1. Adaptive local noise reduction filter.

Simplest statistical measure of random variables are

- (i) mean \rightarrow measure of avg. gray level.
- (ii) variance \rightarrow measure of avg. contrast.

The above two used by the adaptive filters. Since it closely related to the appearance of the image.

Filter Response:

The filter operates on a local region S_{xy} .

Its response at any center point (x, y) is based on four quantities given by,

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

(2) σ_n^2 - The variance of noise corrupting $f(x,y)$.
+ to ~~forming~~ form $g(x,y)$.

(3) m_L - The local mean of the pixels in S_{xy} .

(4) σ_L^2 - The local variance of one pixels in S_{xy} .

Requirements.

filter's behavior is required,

(i) if $\sigma_n^2 = 0$ - The filter should return simply the value of $g(x,y)$. This is called the zero-noise case. In which $g(x,y) = f(x,y)$.

(ii) if σ_L^2 is high relative to σ_n^2 - The filter should return a value close to $g(x,y)$. A high local variance σ_L^2 is usually related with edges and therefore, it should be preserved.

(iii) if $\sigma_n^2 = \sigma_L^2$; The filter should return the arithmetic mean value of the pixels in S_{xy} . This condition occurs when the local area has the same properties as the overall image. Therefore, local noise is to be reduced simply by ~~simple~~ averaging.

Restored Image:

The adaptive local noise filter is,

$$\hat{f}(x,y) = g(x,y) \cdot \frac{\sigma_n^2}{\sigma_L^2} [g(x,y) - m_L]; \quad \sigma_n^2 \leq \sigma_L^2$$

parameters computed from the pixels in S_{xy} at each center location (x,y) .

when $\sigma_n^2 > \sigma_L^2 \rightarrow$ the ~~ratio~~ of this two eqn. in eqn (B) is set to '1' to avoid negative gray level.

2. Adaptive Mean Filter:

* adaptive mean filter also works in a rectangular window area S_{xy} .

* This is called adaptive because of its increase the size of S_{xy} during filter operation. depending on certain specified conditions.

Algorithm

I_{min} \rightarrow min. gray level value in S_{xy} .

I_{max} \rightarrow Max. gray level value in S_{xy} .

I_{med} - Median " " "

I_{xy} \rightarrow Gray level at coordinate (xy)

S_{man} \rightarrow Man. allowed size of S_{xy} .

* Three main purpose of this algorithm.

(i) to remove salt & pepper.

(ii) to provide smoothing of noise other than impulse.

(iii) to reduce distortion, such as excessive thinning or thickening of objective boundaries.

It works in two levels.

Level A, Level B.

Level A

aim. is to determine whether the median filter o/p (I_{med}) is an impulse (black/white) or not.

Algor

Level A: $A_1 = I_{med} - I_{min}$.

$A_2 = I_{med} - I_{max}$.

if $A_1 > 0$ & $A_2 < 0$.

Go to level B.

Else

Increase the window size.

If window size $\leq S_{man}$,

repeat level A.

else

o/p I_{xy} .

i). If the condition

$I_{min} < I_{med} < I_{max}$.

- then I_{med} can't be impulse,

- In this case, it goes to level B & test to see

Level B

Algorithm

Level B: $B_1 = I_{xy} - I_{min}$.

$$B_2 = I_{xy} - I_{max}$$

if $B_1 > 0$ & $B_2 < 0$,

o/p $\rightarrow I_{xy}$.

Else
o/p I_{med} .

Here, (1) If condition $B_1 > 0$ & $B_2 < 0$ is True.

I_{xy} is can't be an impulse. In this case, the algorithm o/p is the unchanged pixel value I_{xy} . This reduces distortion.

Periodic Noise Reduction.

Multiresolution Processing:-

In the Image,

if both small & large objects,
ie, low & high contrast objects,
are present simultaneously. It is easy to
analyse & process the images at ~~any~~
many resolutions.

It is basis for multiresolution processing

⇒ Thus Multi Resolution theory deals with the
process of the representation and analysis
of images / signals at more than one
~~resol~~ resolution.

⇒ Because, some features that are not detected
at one resolution, may be detected at another
resolution.

⇒ Multiresolution Theory utilizes techniques from
different areas such as,

- (1) Subband Coding from signal processing
- (2) Quadrature mirror filtering from
digital speech processing
- (3) Pyramidal Image processing.

1. Subband Coding:-

There are three very important image-related
operations that are used in multiresolution
processing,

(1) Image Pyramids.

- * There are simple structure to represent
images at more than one resolution.
- * An Image Pyramid,
is a collection of decreasing
resolution images that are arranged
in the shape of a pyramid.

(2). The Haar Transform.

The basis function of the transform are the oldest and simplest known orthonormal wavelets.

(3) Subband Coding:-

- * An Image is decomposed into a set of band limited components called subbands.
- * The decomposition is in such a way that the subbands can be reassembled to reconstruct the original image without error.
- * The decomposition and reconstruction process are performed by digital filters.

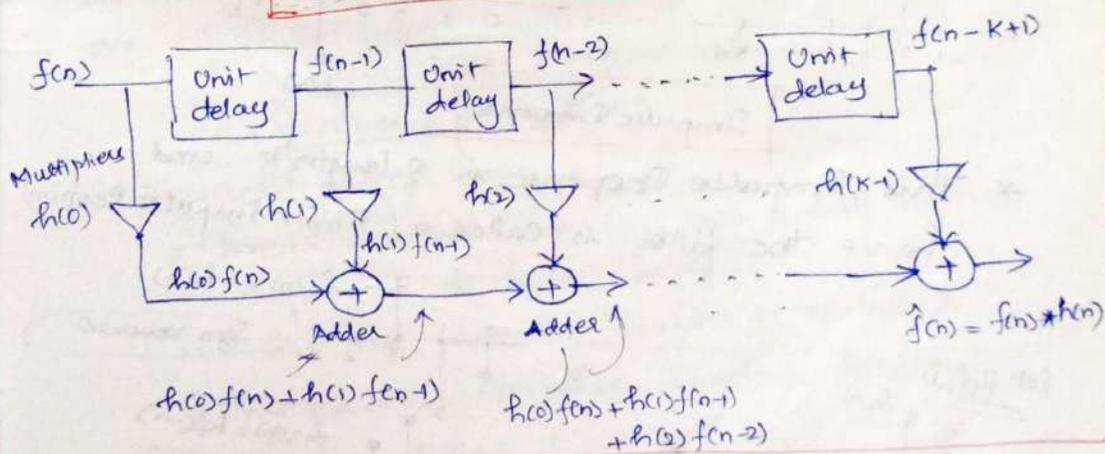
Digital filters

These are basic units of subband coding systems.

⇒ A simple digital filter is shown below.

⇒ It is constructed using three basic components,

- (1) Unit delays.
- (2) Multipliers
- (3) Adders.



* $(k-1)$ unit delays are connected / right shifted versions of the i/p sequence $f(n) = f(n-0)$.

* The delayed sequences are.

$$\Rightarrow f(n-0), f(n-1), f(n-2) \dots f(n-k+1)$$

* These delayed Sequences are multiplied by the constants $h(0), h(1), \dots, h(k-1)$

* Results are added to produce the o/p Sequence

$$f(n) = \sum_{k=-\infty}^{\infty} h(k) \cdot f(n-k) = f(n) * h(n) \quad (\text{Convolution})$$

\Rightarrow k -multiplication constants = Called "filter coefficients".

\Rightarrow It is said to be order of k .

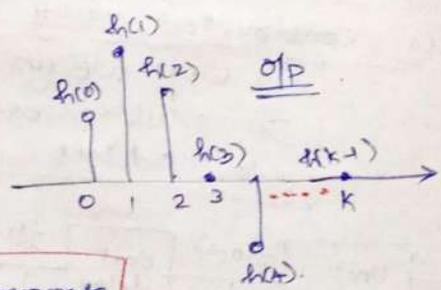
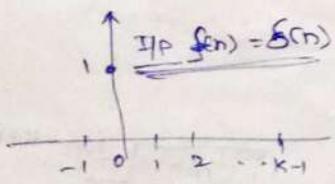
Finite Impulse Response (FIR) filter.

* If of the filter is a unit discrete impulse 'Secs.' Then,

* It's response becomes,

$$f(n) = \sum_{k=-\infty}^{\infty} h(k) \cdot \delta(n-k) = h(n).$$

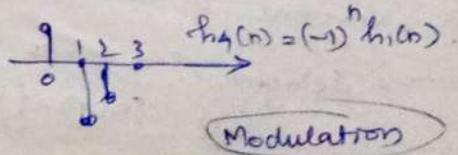
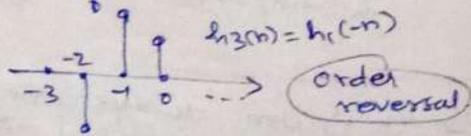
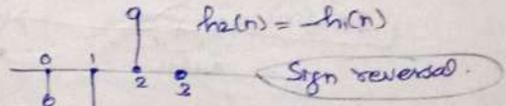
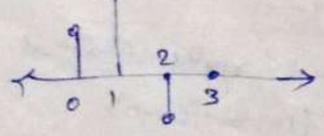
for example



Impulse Response

* This Impulse Response is of length 'k' and hence the filter is called a (Finite Impulse Response)

let ifp is:



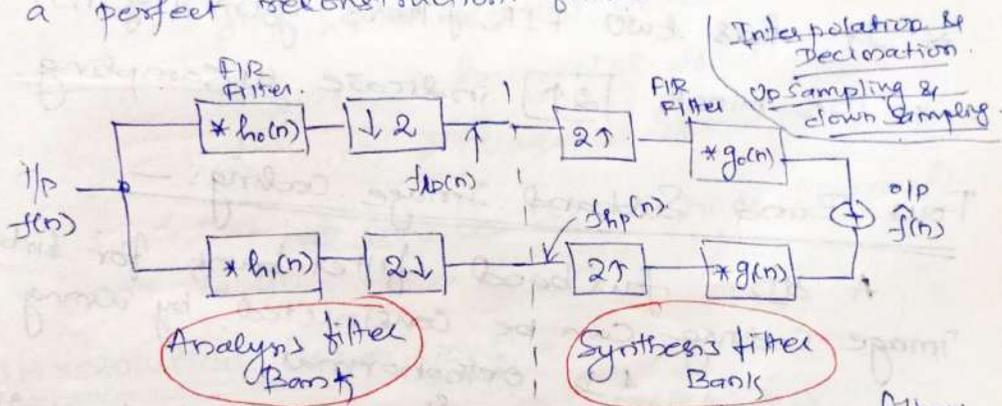
Two Band Subband Image Coding:-

* It is constructed using two filter banks.
 Filter bank \rightarrow Collection of two (or) more filters.

* Perfect Reconstruction.

The goal of the Subband Coding is to select the filters $h_0(n)$ & $h_1(n)$, $f_0(n)$ & $f_1(n)$ in such a way that, the i/p & o/p of the Subband coding and decoding system are identical.
 i.e., $\hat{f}(n) = f(n)$.

If it is done, the resulting system said to be a "perfect reconstruction filter".



In two filter banks, each containing two FIR filters are,

- (1) Analysis filter bank
- (2) Synthesis filter bank.

(1.) Analysis filter bank.

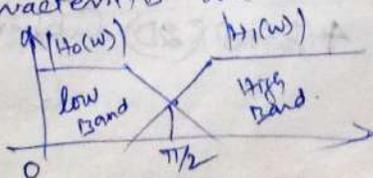
* break i/p sequence $f(n)$ into 2 half length sequences $f_0(n)$ & $f_1(n)$.

* these two are the subbands of the i/p.

* This process is called "Spectrum Splitting"

* two FIR filters $h_0(n)$ & $h_1(n)$.

* $h_0(n)$ & $h_1(n) \rightarrow$ half band filters with transfer characteristics are shown.



* $h_0(n) = \text{LPF}$,

its of Subband $f_p(n)$ is called the approximation of $f(n)$.

* $h_1(n) = \text{HPF}$

$f_1(n)$ - high freq. (os)
detail part of $f(n)$

1/11/17
Vengayur,
Sathish
Naveen,
Shravan
Vijayalakshmi
Parthodi
P. Shalini

(b) Synthesis filter bank:-

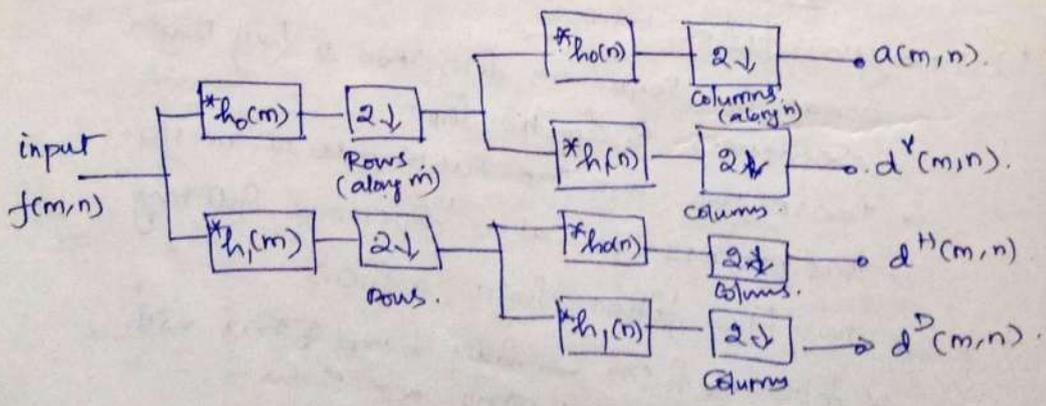
* This filter bank is used to combine the subbands $f_p(n)$ & $f_q(n)$ to produce the output $\hat{f}(n)$.

* It has two FIR filters $g_0(n)$ & $g_1(n)$.

* The blocks $\boxed{2\uparrow}$ indicate upsampling.

Four Band Subband Image Coding:-

A 2D four band filter bank for subband image coding can be constructed by using 1D orthogonal & biorthogonal filters as 2D separable filters for image processing.



A 4 Band (2D) Subband Coding System.

Here,

- (1) The separable filters are first applied in 1D and then in the other dimension.
Such as, vertically first, horizontally etc.,
- (2) down Sampling $\lfloor \frac{\cdot}{2} \rfloor$ performed 2 times, This helps to reduce the no. of computations.
- (3) 4 subbands obtained. as a result of filtering the ip image.

They are

- $a(m,n) \rightarrow$ approximation
- $d^v(m,n) \rightarrow$ Vertical detail
- $d^h(m,n) \rightarrow$ Horizontal detail
- $d^d(m,n) \rightarrow$ diagonal detail.

(4) Further that can be split into 4 subbands and so on.

Multi resolution Expansions! -

~~is~~ Multi Resolution Analysis (MRA) is a

- mathematical framework in which,
- * a scaling function is used to create a series of approximation of a function or image.
- * Each of these approximation, differs in resolution. by a factor of 2 from its nearest neighboring approximation.
- * Then this difference in information b/w adjacent approximation is encoded by using additional functions known as Wavelets.

3 Important fn., play role in MRA.

- (1) Series expansions (2) Scaling functions
- (3) Wavelet functions.

1. Series Expansion

⇒ Let a signal (or) function = $f(x)$

⇒ It is expressed as a linear combination of expansion fr as,

$$f(x) = \sum_k \alpha_k \psi_k(x);$$

Expansion set $\{\psi_k(x)\}$.

where

k - Integer index
finite/infinite

α_k - real valued
Expansion coeff

$\psi_k(x)$ - real valued
Expansion fr
(or)
Basis function.

* closed Span of the function.

$$V = \text{Span} \left\{ \psi_k(x) \right\}$$

* Computation of α_k Coefficients

There is a set of dual functions denoted as $\{\psi_k(x)\}$ for any space V corresponding to expansion set $\{\psi_k(x)\}$ and. This can be used to compute the α_k coefficients for any $f(x) \in V$ of dual.

⇒ taking integration of the product of $\psi_k(x)$ & $f(x)$.
The coefficients as computed as,

$$\alpha_k = \left\langle \overline{\psi_k(x)}, f(x) \right\rangle \\ = \int \overline{\psi_k(x)}^* f(x) dx.$$

These have three possible forms.

Case 1

i.e., Orthogonal basis for V i.e.,

$$\left\langle \psi_j(x), \psi_k(x) \right\rangle = \delta_{jk} = \begin{cases} 0, & j \neq k \\ 1, & j = k \end{cases}$$

∴ basis = dual.

$$\psi_k(x) = \overline{\psi_k(x)}$$

Now α_k computed as

$$\alpha_k = \left\langle \psi_k(x), f(x) \right\rangle$$

$\overline{\psi_k(x)}$ - dual
 $\psi_k^*(x)$ - conjugate
 $f(x)$ - function.
 $\psi_k(x)$ - basis.

Case 2

If the expansion f_n form an orthogonal basis for V
 then, $\langle \psi_j(x) \cdot \psi_k(x) \rangle = 0, j \neq k$.

α_k computed as,

$$\alpha_k = \int \psi_k^*(x) \cdot f(x) dx$$

Case 3

\Rightarrow If the expansion set is spanning set,
 such expansion f_n & their duals are said to be
 overcomplete (or) redundant

\Rightarrow These expansion functions form a frame in which

$$\|f(x)\|^2 \leq \sum_k |\langle \psi_k(x) | f(x) \rangle|^2 \leq B \|f(x)\|^2$$

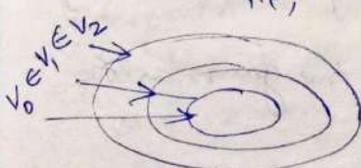
2. Scaling f_n

Let the function $\phi(x) \in V_0$, such that the set of
 f_n consisting of $\phi(x)$ & its integer translates
 $\{\phi(x-k) | k \in \mathbb{Z}\}$ form a basis for the space V_0 which
 is termed scaling f_n (or) father f_n .

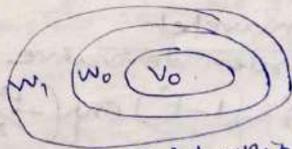
* The subspaces V_j are nested which implies
 $V_j \subseteq V_{j+1}$. It's possible to decompose V_{j+1} in terms
 of V_j & W_j ,

\Rightarrow Here W_j is the orthogonal complement of V_j
 in V_{j+1} .

i.e., $V_j \oplus W_j = V_{j+1}$ $\oplus \Rightarrow$ direct sum



Nested subspaces spanned by the scaling f_n .



Relationship b/n spaces spanned by the scaling & wavelet f_n .

from above fig (2) it is obvious that W_j is the difference
 b/n V_{j+1} & V_j .

also, $V_0 \oplus W_0 = V_1$

$V_0 \oplus W_0 \oplus W_1 = V_2$

also $W_j \perp V_j$, it is noted that \oplus direct sum of W_j is equal
 to L^2 which is,

$$\bigcup_{j=-\infty}^{\infty} V_j = \bigoplus_{j=-\infty}^{\infty} W_j = L^2$$

If spaces V_0 & W_0 are subspaces of V_1 , then the scaling fn. $\phi(x)$ & wavelet fn. $\psi(x)$ can be expressed as in terms of basis fn. of V_1 as,

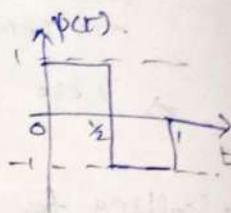
$$\left. \begin{aligned} \phi(x) &= \sum_k h_k \phi(2x-k) \\ \psi(x) &= \sum_k g_k \phi(2x-k) \end{aligned} \right\} \text{These are } 2 \text{ scale \& relations}$$

Ex. of wavelets

1. Haar wavelet

It is bipolar fn.

$$\psi(t) = \begin{cases} 1 & \text{when } 0 < t < \frac{1}{2} \\ -1 & \text{when } \frac{1}{2} < t < 1 \\ 0 & \text{otherwise} \end{cases}$$



2. ~~Haar~~ Daubechies wavelet

are a family of orthonormal, compactly supported scaling & wavelet fn. ~~that~~

3. Morlet wavelet

obtained by multiplying the Fourier basis with a Gaussian window.

$$\text{i.e., } \psi(t) = \exp(j\omega_0 t) \cdot \exp(-t^2/2)$$

Spectrum of Morlet wavelet,

$$\Phi(\omega) = \sqrt{\pi/2} \left[\exp\left(-\frac{(\omega-\omega_0)^2}{2}\right) + \exp\left(-\frac{(\omega+\omega_0)^2}{2}\right) \right]$$

4. Mexican-hat wavelet

It is 2nd order derivative of the gaussian fn.

$$\psi(t) = (1-t^2) \exp(-t^2/2)$$

5. Shannon wavelet

$$\psi(t) = \frac{\sin(2\pi t) - \sin(\pi t)}{\pi t}$$

Wavelet-

- * a wave is an oscillating fn. of time/space that is periodic. ~~the wave~~
- * The wave is an infinite length continuous fn. in time/space.
- * In contrast, wavelets, are localized waves.
- * A wavelet is a waveform of an effectively limited duration that has an average value of zero.

A fn. $\phi(x)$ can be called a wavelet if it poses the following properties.

- (1) The fn. integrate to zero, or equivalently, its FT denoted as $\phi(\omega)$ is zero at the origin.

$$\int_{-\infty}^{\infty} \phi(x) dx = 0 \quad \text{--- (1)}$$

This implies $\phi(\omega)|_{\omega=0} = 0$ in the freq. domain

- (2) It is square integrable | equivalently has finite energy, $\int_{-\infty}^{\infty} |\phi(x)|^2 dx < \infty$. --- (2)

- (3) The FT $\phi(\omega)$ must satisfy the admissibility condition given by,

$$C_{\phi} = \int_{-\infty}^{\infty} \frac{|\phi(\omega)|^2}{|\omega|} d\omega < \infty \quad \text{--- (2)}$$

Interpretation of above eqn.

\Rightarrow a family of wavelets can be generated by dilating & translating the mother wavelet $\phi(x)$, which is given by

$$\psi_{(a,b)}(x) = \frac{1}{\sqrt{a}} \phi\left(\frac{x-b}{a}\right)$$

Here, a - Scale parameter
 b - Shift parameter.

Wavelet Transform.

1D CWT

$$W_f(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) \cdot \psi\left(\frac{t-b}{a}\right) dt$$

2D CWT

$$\frac{1}{\sqrt{a}} \iint f(x,y) \psi\left(\frac{x-m}{a}, \frac{y-n}{a}\right) dx dy$$

$m, n \rightarrow$ shifting parameter
 $a \rightarrow$ scaling parameter

Image Compression

⇒ Digital images are compressed. of an enormous amount of data.

Reduction of size of the image data for both storing and transmission of digital are becoming increasingly important as find more applications.

⇒ Image compressor is a mapping from a higher dimensional space to a lower dimensional space.

⇒ The basic goal of image compression is to represent an image with minimum number of bits of an acceptable image quality.

Need for Image Compression.

Storage & too.

for ex

to store an image of size 1600×1200 . colour image then the space required to store the image is,

$$1200 \times 1600 \times 8 \times 3 = 46,080,000 \text{ bits}$$

$$= 5,760,000 \text{ bytes}$$

$$= 5.76 \text{ Mbytes.}$$

let max. space available is

floppy disk is 1.44 Mb.

if we have 4 floppy $1.44 \text{ Mb} \times 4 = 5.76 \text{ Mbytes.}$

⇒ i.e, min. of 4 floppy disk is required for storage of RGB image 1600×1200 .

⇒ The amount of data transmitted through the internet doubles every year, and a large portion of that data comprises of images.

Redundancy in Images

- * Images are highly coherent -
that means redundant information.
- * Compression is achieved through,
 1. redundancy
 2. irrelevancy reduction.
- * Redundancy means duplication,
&
irrelevancy means that part of the image information that will not be noticed by the human visual system.

Classification of Redundancy in images:-

classified into

1. ~~Statistical~~

1. Statistical redundancy

2. psy^{ch} chorisual redundancy

↓
Interspatial redundancy

↓
Coding redundancy

↓
Spatial redundancy

↓
Temporal redundancy

I Statistical redundancy

(1) Interspatial: due to correlation b/w neighboring pixels in an image.

(Spatial, Geometric, Interframe redundancy)
ie, means neighboring pixel are not statistically independent.

(2) Coding: associated with the representation of information.

= the information is represented in the form of codes.

eg, Huffman code, Arithmetic code

Data Redundancy:

Some representations may have data which ~~do~~ do not convey any useful information, therefore

These representations are said to contain data redundancy.

Spatial redundancy

* represents the Statistical Correlation b/w neighboring pixels in an image.
⇒ Spatial redundancy implies that there is a relationship b/w neighboring pixels in an image.
i.e., It is not necessary to represent each pixel in an image independently.

Temporal redundancy

⇒ is a Statistical correlation b/w. pixels from successive frame. In a video sequence.

⇒ The temporal redundancy is called Intraframe redundancy.

⇒ Motion compensated predictive coding is employed to reduce temporal redundancy.

Psycho visual Redundancy

is associated with the characteristics of human visual system. (HVS)

In HVS, visual information is not perceived equally.

Some information more important than other information.

If less data is used to represent less important visual information, perception will not be affected.

This implies that visual information is psychovisually redundant.

Relative Data redundancy

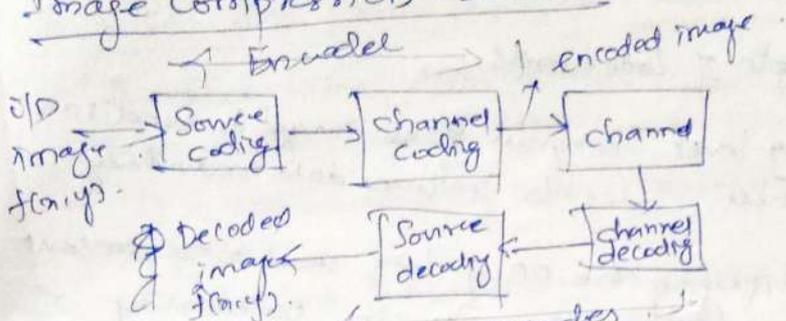
Let, two data sets that represent the same information,

- n_1 & n_2 are the no. of information-carrying units (bits) in the data set

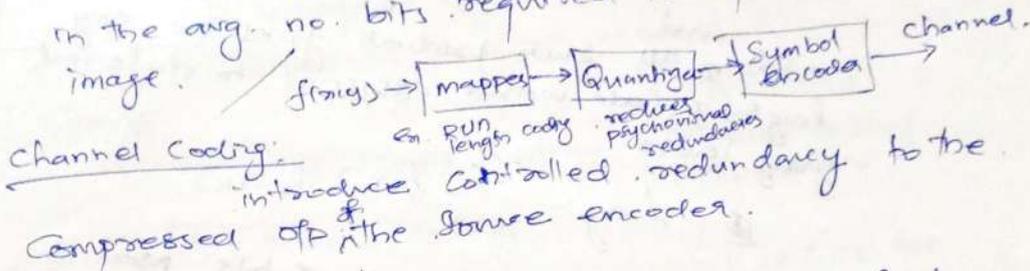
Now, Relative D. Redun (R_D) of the 'I' data set is given by

$$R_D = 1 - \frac{1}{CR} \quad \left| \begin{array}{l} CR - \text{Compression ratio} \\ CR = \frac{n_1}{n_2} \end{array} \right.$$

Image Compression Scheme



Source coding reduces the entropy, i.e., means decreasing in the avg. no. bits required to represent the image.



Channel coding: Compressed output of the source encoder. introduce controlled redundancy to the purpose is to protect the comm system against noise & other Txn. errors. in the channel.

Classification of Image Compression Scheme

lossless compression scheme

- * Preferred in the case of medical image compression.
- * The reconstructed image exactly resembles the original image without any loss of information.
- * But compression ratio is usually less.

lossy compression scheme

- * preferred as in the case of multimedia applications.
- * The reconstructed image expense the quality of the original image. i.e., loss the information.
- * High compression ratio. Can be obtained

* if $f(x,y) = f(x,y)$ it is called error free / information preserving / lossless compression.

- Based on n_1 & n_2 values, can be 3 different cases.
- (1) $n_1 = n_2 \Rightarrow CR = 1$ & $RD = 0 \Rightarrow$ First Set No Redundant data.
 - (2) $n_1 > n_2 \Rightarrow CR = \infty$ & $RD = 1 \Rightarrow$ First Set contains highly redundant data.
 - (3) $n_1 < n_2 \Rightarrow CR = 0$ & $RD = \infty \Rightarrow$ I data Set has more data than the II set. This is undesirable.

Entropy Redundancy

Avg length of code words.

→ Gray level histogram of an image is used to construct codes to reduce data redundancy.

⇒ Multiplying the no. of bits used to represent each gray level with the probability of occurrence of that gray level

and adding all such products gives the avg length of code words assigned to different gray levels.

$$\text{Avg Length} = \sum_{k=0}^{L-1} l(r_k) \cdot P_r(r_k)$$

$l(r_k)$ → No. of bits used to represent each gray level

$P_r(r_k)$ → Probability of occurrence of gray level r_k

* This is the avg. no. of bits needed to represent each pixel

* Probability of occurrence of gray level ' r_k '

$$P_r(r_k) = \frac{n_k}{n} ; k = 0, 1, \dots, L-1$$

L → No. of gray levels

n_k → No. of times the k th gray level appears in the image

n → total no. of pixels

⇒ Total no. of bits

required to code $M \times N$ image is,

$$L_{\text{total}} = M \cdot N \cdot \text{Avg Length}$$

Lossless (or) Error Free Compression.

- * Is acceptable data reduction method since there is no loss of data.
- * It is applicable to both Binary & Gray Scale images.
- * All lossless compression techniques are formed by two independent components.
 - (1) Interpixel redundancies.
 - (2) Coding redundancies.

Methods

- (1) Variable-length coding. — Reduce coding redundancies.
- (2) Bit plane coding.
- (3) LZW coding.
- (4) Lossless predictive coding. } reduce the interpixel redundancies.

1. Variable length coding:-
assigne. Shortest possible code words, to the most probable gray levels, and vice versa.

Three types of coding schemes are available to implement variable-length coding namely,

- (1) Huffman code
- (2) Near optimal variable length codes.
- (3) Arithmetic coding.

1. Huffman coding:-

- * It provides smallest no. of codes, symbols per source symbol.
- * It gives ~~the~~ removing coding redundancy.

Coding procedure:-

1. Consider six source symbols, with probabilities given below are to be coded.
 $\{a_1, a_2, a_3, a_4, a_5, a_6\} = \{0.1, 0.4, 0.06, 0.1, 0.04, 0.3\}$.

Step 1

The probability of symbols.

To order the probabilities of the symbol (given) in the descending order.

Now, the source symbol is,

$$\{a_2, a_6, a_1, a_4, a_3, a_5\}$$
$$= \{0.4, 0.3, 0.1, 0.1, 0.06, 0.04\}$$

Step 2

A source reduction is created by adding the lowest probabilities, i.e., bottom two probabilities, into a single symbol.

This symbol is known as a compound symbol. Then the two probabilities are replaced by the compound symbol & its probability in the next source reduction.

⇒ Step 1 & Step 2 are repeated until a source with only two symbols is obtained. i.e.,

Original Source		Source reduction			
Symbol	Probability	1	2	3	4
a_2	0.4	0.4	0.4	0.4	→ 0.6
a_6	0.3	0.3	0.3	0.3	
a_1	0.1	0.1	→ 0.2	→ 0.3	} 0.4
a_4	0.1	0.1			
a_3	0.06	→ 0.1	} 0.1	} 0.3	} 0.4
a_5	0.04				

Huffman source reduction.

Step 3

Each reduced source is coded. It starts from the smallest source obtained in the last step and goes back to the original source.

The minimal length binary codes used are 0 & 1.

Ex 10.2

Original Source		Code Assignment Steps - (Source reduction)				Word Length
Symbol	Probability	Code word	1	2	3	
a2	0.4	1	0.4 (1)	0.4 (1)	0.4 (1)	1
a6	0.3	00	0.3 (00)	0.3 (00)	0.3 (00)	2
a1	0.1	011	0.1 (011)	0.2 (01)	0.3 (01)	3
a4	0.1	0100	0.1 (0100)	0.1 (01)		4
a3	0.06	01010	0.1 (01)	0.1 (01)		5
a5	0.04	01011	0.1 (01)	0.1 (01)		5

1. Average Code length :-

It is defined as,

$$L_{avg} = \sum_{k=0}^{L-1} l(r_k) \cdot Pr(r_k)$$

~~L~~ $L = 6$

$l(r_k)$ = No. of bits used to represent each gray level.

$Pr(r_k)$ = Probability of occurrence of gray level r_k .

$$\therefore L_{avg} = (0.4)(1) + 0.3(2) + 0.1(3) + 0.1(4) + 0.06(5) + 0.04(5)$$

$$L_{avg} = 2.2 \text{ bits/symbol}$$

2. Entropy

It is defined as,

$$H(x) = - \sum_{k=0}^{L-1} Pr(r_k) \cdot \log_{10} [Pr(r_k)] \quad / \quad H(x) = - \sum_{k=0}^{N-1} P_k \log_2 (P_k)$$

$$H(x) = 0.4 \cdot \log_{10}(0.4) + 0.3 \log_{10}(0.3) + 0.1 \log_{10}(0.1) + 0.1 \log_{10}(0.1) + 0.06 \log_{10}(0.06) + 0.04 \log_{10}(0.04)$$

$$= \frac{\text{result}}{\log 2} = \frac{\text{result}}{0.3010} = \frac{0.643}{\log 2} = 2.13$$

In case computing by \log_{10} .

If computing by \log_2

means no need to divide the result by $\log 2$

3. Determination of efficiency: (η)

$$\eta = \frac{H(Z)}{L_{avg}} \times 100 = \frac{2.13}{2.2} \times 100 =$$

4. Redundancy:-

$$R = 1 - \eta = 1 - 0.97 = 0.03$$

Ex 2 a source emits letters from an alphabet and obtains the Huffman code for "COMMITTEE"

\Rightarrow Probability of Symbols = $\frac{\text{total no. of symbols}}{\text{total no.}}$

Symbol	Probability
C	$\frac{1}{9}$
O	$\frac{1}{9}$
M	$\frac{2}{9}$
I	$\frac{1}{9}$
T	$\frac{2}{9}$
E	$\frac{2}{9}$

\Rightarrow descending order

M	2
T	2
E	2
C	1
O	1
I	1

$$L_{avg} = 2.553$$

$$H(Z) = 2.5034 \text{ bits/symbol}$$

$$\eta = 97.97\%$$

Drawbacks

(1) It has the limitation that only one symbol can be coded at a time.

\therefore This coding is difficult. no. of symbols is to be coded.

(2) for 'k' source symbols, it requires source reduction and k-2 codewords. \therefore The computational

Ex. 3 For the image shown below compute the compression ratio that can be achieved using Huffman code.

$$\begin{bmatrix} 3 & 3 & 3 & 2 \\ 2 & 3 & 3 & 3 \\ 3 & 2 & 2 & 2 \\ 2 & 1 & 1 & 0 \end{bmatrix}$$

$$P(0) = \frac{1}{16} = 0.0625$$

$$P(1) = \frac{2}{16} = 0.125$$

$$P(2) = \frac{6}{16} = 0.375$$

$$P(3) = \frac{7}{16} = 0.4375$$

Sym	Prob	Code
3	0.4375	1
2	0.375	00
1	0.125	010
0	0.0625	011

$$L_{Avg} = 1.75 \text{ bits} = 1.75 ; n_1 = \frac{2}{1.75}$$

$$H(2) = 1.67$$

$$\eta = 95.4\% \quad R = 0.046$$

$$CR = \frac{n_1}{n_2} = \frac{2}{1.75} = 1.14$$

0	-00
1	-01
2	-10
3	-11

Non-Binary Huffman Code:

- * As a non-binary Huffman code, consider the ternary Huffman code.
- * The No. of nodes to be combined initially is given by the formula.

$$J = 2 + (N-2) \cdot \text{mod}(D-1)$$

where,

J = No. of symbols to be combined initially.

N = No. of probabilities.

D = No. of counts.

Ex. from the ternary Huffman code for the word.

"COMMITTEE"

⇒ The No. of nodes to be combined initially is given by,

$$J = 2 + (N-2) \text{ Mod}(D-1)$$

where N = 6.

D = 3

i.e., Code word composed of (0, 1, 2)

C	-0	Total 6 Symbols
O	-1	
M	-2	
I	-3	
T	-4	
E	-5	

$$\therefore J = 2 + (6-2) \text{ Mod}(3-1)$$

$$= 2 + 4 \times \text{Mod} 2$$

$$= 2 + 0 = 2$$

	Probability		code					Word length			
M	2/9	00	3/9	1	6/9	0		2			
T	2/9	01	2/9	} 00	3/9	1		2			
E	2/9	02	2/9						} 01		
C	1/9	10	} 2/9	} 02			2				
O	1/9	11									2
I	1/9	12									

Arithmetic entropy

$$H(S) = - \sum_{k=0}^{N-1} P_k \log_2 P_k$$

$$= 2.5034 \text{ bits/symbol}$$

$$L_{avg} = \sum_{k=0}^{N-1} P_k l_k$$

$$= 1.9958 \text{ bits/symbol}$$

$$\eta = \frac{H(S)}{L_{avg}} = 1.2578$$

$$\approx 125.18\%$$

Arithmetic Coding:-

Ex A coin

In this, the interval from $[0-1]$ is divided according to the probability of the occurrences of the intensities.
+ It is not generating individual codes for each character but perform arithmetic operation on a block of data based on probabilities.

Ex

A source emits a four symbols $\{A, B, C, D\}$ with the probabilities $\{0.4, 0.2, 0.1, 0.3\}$ respectively, Construct arithmetic coding to encode & decode the word 'DAD'

Sol

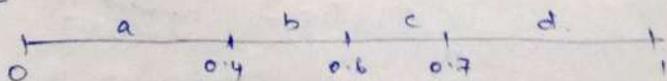
The symbol probabilities & their corresponding sub-ranges are tabulated,

Symbol	A	B	C	D
Prob.	0.4	0.2	0.1	0.3
Sub range	0.0-0.4	0.4-0.6	0.6-0.7	0.7-1.0

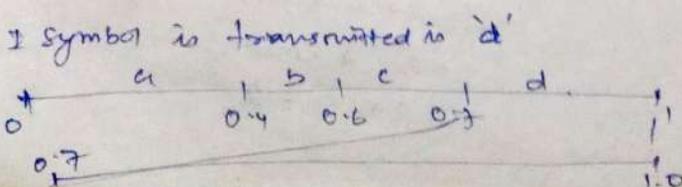
Encoder

of our objective is encode word 'DAD'

Step 1



Step 2



The formula that is used to compute the low & high range for each Symbol within the interval is:

$$\text{low} = \text{low} + \text{range} \times (\text{low-range}) \quad \text{Symbols range}$$

$$\text{high} = \text{low} + \text{range} \times (\text{high-range})$$

To find the interval

for a to compute the interval for the Symbol a in the interval 0.7 to 1.0

$$\Rightarrow \text{range} = 1.0 - 0.7 = \underline{0.3}$$

$$\Rightarrow \text{low} = \text{low} + \text{range} (\text{low-range})$$

$$= 0.7 + (0.3) (0.0)$$

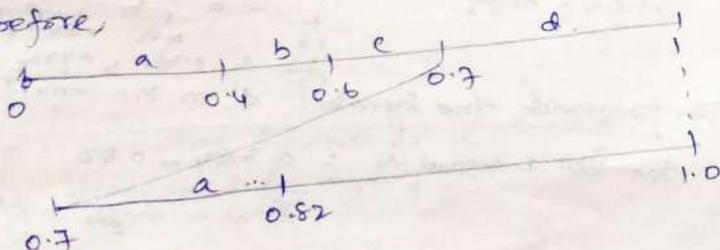
$$\therefore \text{low} = \underline{0.7}$$

$$\Rightarrow \text{high} = \text{low} + \text{range} (\text{high-range})$$

$$= 0.7 + 0.3 (0.4)$$

$$= \underline{0.82}$$

Therefore,



To find the interval for b

$$\text{range} = 0.3$$

$$\text{low} = \text{low} + \text{range} (\text{low-range})$$

$$= 0.7 + 0.3 (0.4) = \underline{0.82}$$

$$\text{high} = \text{low} + \text{range} (\text{high-range})$$

$$= 0.7 + 0.3 (0.6) = \underline{0.88}$$

To find the interval for c

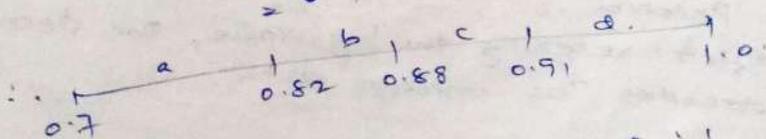
$$\text{range} = 0.3$$

$$\text{low} = \text{low} + \text{range} (\text{low-range})$$

$$= 0.7 + 0.3 (0.6) = \underline{0.88}$$

$$\text{high} = \text{low} + \text{range} (\text{high-range})$$

$$= 0.7 + 0.3 (0.7) = \underline{0.91}$$



Step The Symbol to be transmitted is 'a'
Hence the interval is 0.7 to 0.82

to find the interval for a

here, range = $0.82 - 0.7 = \underline{0.12}$

low = 0.7

\therefore range = 0.12

low = low + range (low-range) = $0.7 + 0.12(0.0)$

high = low + range (high-range) = $0.7 + 0.12(0.4)$
 $= 0.748$

for b

low = $0.7 + 0.12 \times (0.4) = 0.748$

high = $0.748 + 0.12 \times (0.6) = 0.772$

for c

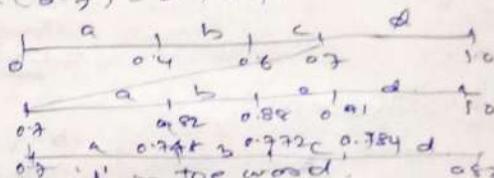
low = $0.7 + 0.12 \times (0.6) = 0.772$

high = $0.772 + 0.12 \times (0.7) = 0.784$

for d

low = 0.784

high = 0.82



Step 4

To transmit the symbol 'd' in the word.

the 'd' interval is = $0.784 - 0.82$
 $\underline{\underline{\text{low} - \text{high}}}$

Step 5

to compute the tag value.

The tag is the avg of the lower & higher intervals.

i.e., low = 0.784
 high = 0.82

\therefore tag = $\frac{0.784 + 0.82}{2} = \underline{\underline{0.802}}$

Decoding Procedure

- * The tag value & symbol probabilities will be sent to the Receiver.
- * After receiving the tag value, the decoder decodes the encoded data.

Symbol	a	b	c	d
Prob.	0.4	0.2	0.1	0.3
sub-range	0.0-0.4	0.4-0.6	0.6-0.7	0.7-1.0

→ The tag received is 0.802

→ The initial interval is assumed $5/10 - 1$

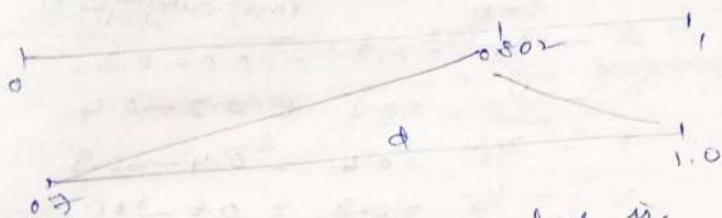


the tag value 0.802 is lies b/n 0.7 - 1.0.

⇒ hence the corresponding decoder symbol is d.

* The decoded symbol in step 1 is d &

hence the new interval is fixed is 0.7 - 1.0.

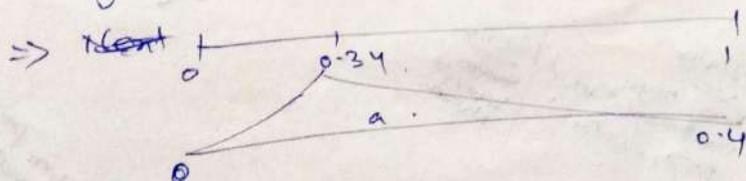


⇒ The new tag value is computed is,

$$\text{tag}^* = \frac{\text{tag} - \text{low}}{\text{range}}$$

$$= \frac{0.802 - 0.7}{0.3} = 0.34$$

⇒ the decoded symbol is 'd' because the new tag lies in the range of 0.0 - 0.4.



⇒ The new tag is,

$$\text{tag}^* = \frac{\text{tag} - \text{low}}{\text{range}} = \frac{0.34 - 0}{0.4} = 0.85$$

⇒ This tag 0.85 lies b/n 0.7 to 1.0.

∴ hence the decoded symbol is 'd'.

by following steps.

That the decoded word is dad.

Ex. 2

Generate the tag using arithmetic ~~coding~~ ^{Coding} procedure to transmit the word INDIA

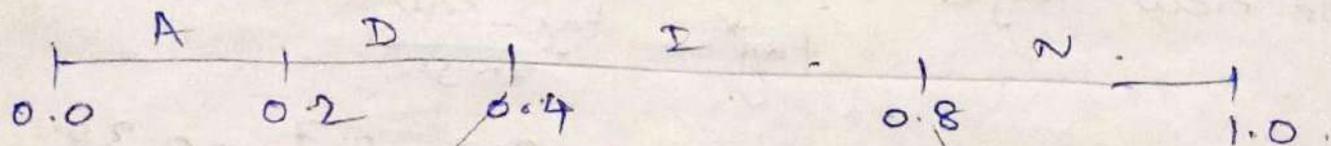
(1) let lower interval is = 0.

higher is = 1.

(2) Probability

We have the letters of I, N, D, A.
write in alphabetical order.

	Prob.	range/cumulative density to
A	$\frac{1}{5} = 0.2$	$= 0.0 \rightarrow 0.2$
D	$\frac{1}{5} = 0.2$	$= 0.2 \rightarrow 0.4$
I	$\frac{2}{5} = 0.4$	$= 0.4 \rightarrow 0.8$
N	$\frac{1}{5} = 0.2$	$= 0.8 \rightarrow 1.0$



(3) First letter to transmit in INDIA is 'I'.

MPEG Encoder / DPCM / DCT Coder.

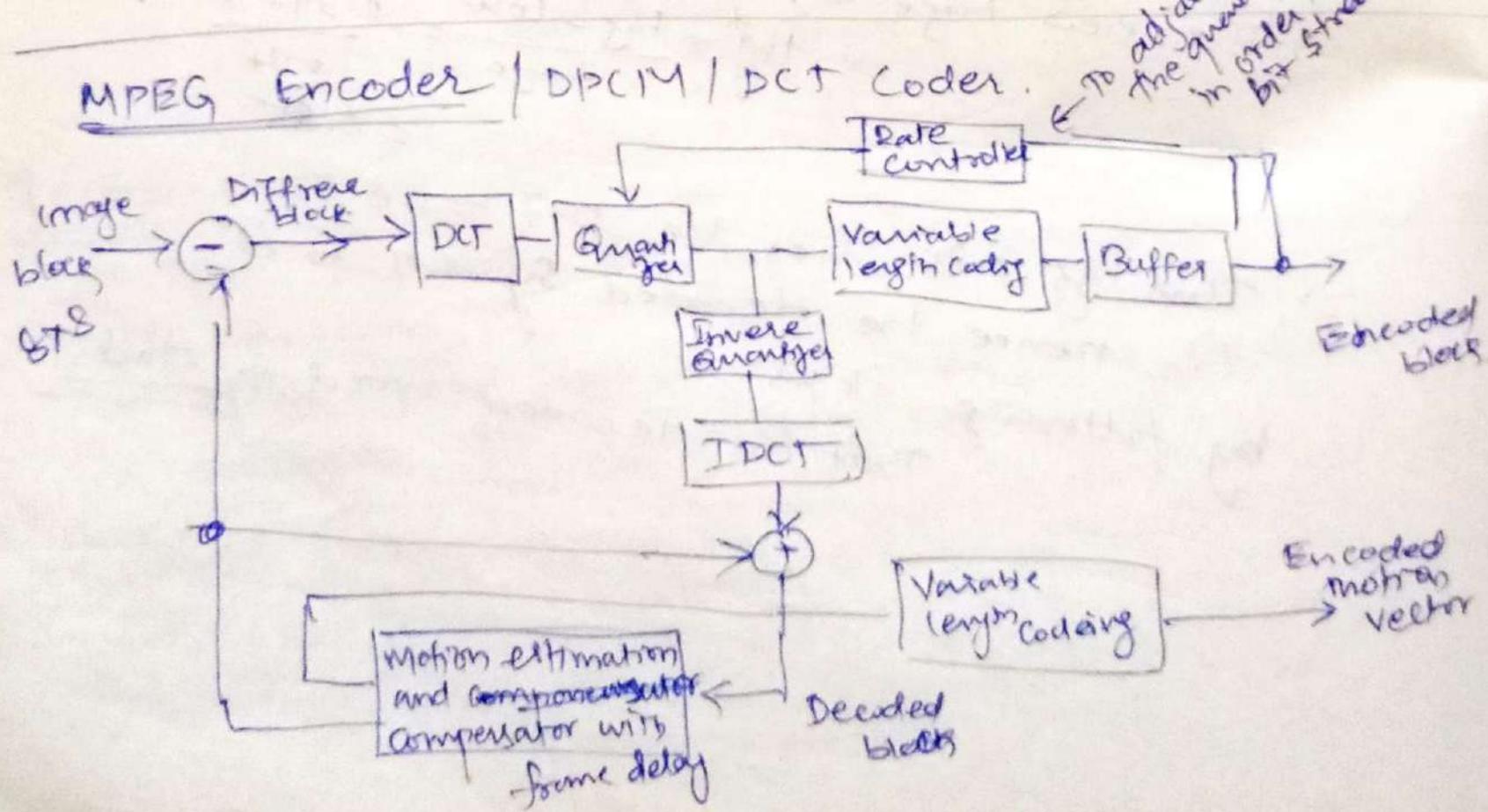


IMAGE REPRESENTATION AND RECOGNITION.

Boundary Representation - chain code - polygonal approximation, signature, boundary segments -
Boundary description - shape number - Fourier Descriptor, moments - Regional Descriptors - topological feature, Texture - patterns & pattern classes - Recognition based on matching.

Image Representation: -

- * Representation is the process of characterizing the quality of quantity represented by each pixel.
 - * This is done after segmentation of an image.
 - * The output of image segmentation is a group of pixels.
- ∴ To change this "raw pixel" data into a form suitable for further computer processing, this representation is used.

Two basic types of representation are possible,

1. Boundary (or) External Representation:

- * It represents a region in terms of its external characteristics i.e., its boundary.
- * It is selected when importance is given to shape characteristics.

Ex Corners, Inflections

2. Regional (or) Internal Representation:

- * It represents a region in terms of its internal characteristics.
- * It is chosen when the main focus is on regional properties.

Ex Texture.

⊗ Some cases both types of representation may be used.

Descriptors :- Descriptions.

* Not only representation makes the data useful to a computer.

* That is description of the region, based on the chosen representation, is also needed to complete the process.

Obj:

description process is to "capture" the needed difference b/w. objects / classes of objects while maintaining as much independence as possible to changes in factors such as ~~area~~
location,
size,
orientation.

features

Some features used to describe the boundary of a region are,

1. length
2. Orientation of the straight line joining the extreme points.
3. No. of concavities i.e., curves in the boundary etc.

feature selected for description should be insensitive to the following variations.
Size, Translation & Rotation.

Types

(1) Boundary descriptors.

(2) Regional descriptors

(3) Relational descriptors. (used for both boundaries & Regions)

Boundary Representation :-

It represents a region in terms of its external characteristics. Therefore it is also known as external representation.

1. Chain Codes
2. Polygonal approximations
3. Boundary Segments
4. Signatures
5. Skeletons

1. Chain Codes :-

Chain codes represents a boundary by using a connected sequence of straight line segments with specified length & direction.

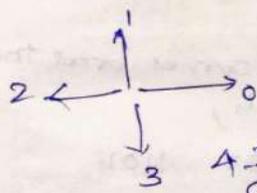
It is represented based on two types of connectivity of the segments, known as,

- (a) 4-Connectivity
- (b) 8-Connectivity.

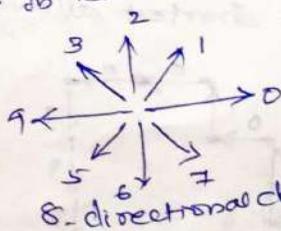
Direction Numbers

The direction of each segments in the chain code is given a number.

The numbering scheme is shown below in fig.



4-Directional chain code



8-directional chain code

Procedure

- * A grid format with equal spacing in X & Y direction is used
- * This grid, a chain code can be created by following a boundary in clockwise direction.
- * Then each segment connecting a pair of pixels is assigned a direction.

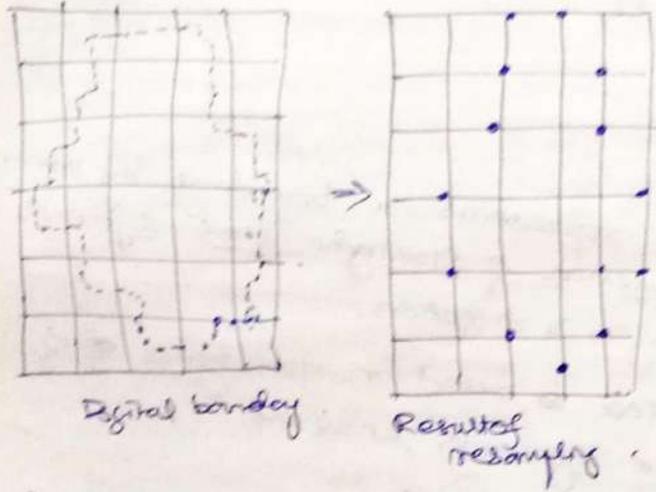
drawbacks

- (1) The resulting chain of codes is very long.
- (2) Noise / Imperfect segmentation may produce small disturbances in the boundary.

Resampling

above ~~drawbacks~~ drawbacks overcome by resampling.

1. The boundary is resampled by selecting a larger grid spacing as in fig 1.



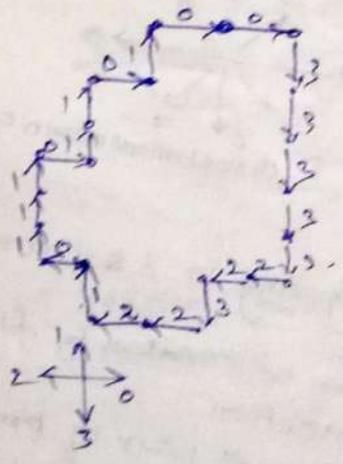
(2) Then the boundary is traversed, i.e. tracked and the same time, boundary points are assigned to each node of the large grid.

Chain codes generation.

To represent the resampled boundary, two types of codes can be generated.

- (1) 4-directional chain code (2) 8-directional

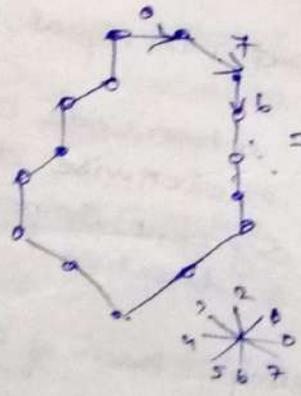
The shortest allowable 4 paths / 4-code is shown.



Starting point is at top, left corner and the chain code is,

003332...1101

||| shortest allowable 8-paths / 8-code



Chain code.

=> 0766...1212

Normalization

Chain codes should be normalized w.r. to its starting point, since it depends on the starting point.

Magnitude Normalization.

Consider the chain code as a sequence of direction ~~no.~~ numbers. & redefine the starting point.

The results sequence of ~~no.~~ numbers forms an integer of min. magnitude.

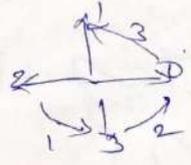
(iii) Rotation Normalization

The first difference of the chain code is used to normalized rotation.

I difference = No. of direction changes to anticlockwise direction, which separate two adjacent elements of the code.

Ex. Consider the chain code: 03321

Chain code : 0 3 3 2 1
 diff : 3 3 0 3 3



I diff bin 1 & 0 \Rightarrow 3 bin 0 & 3 = 3
 bin 3 & 3 = 0 bin 3 & 2 = 3
 bin 2 & 1 = 3

19/9/17

Polygonal Approximations.

polygon is used to approximate a digital banday.

obj:

To capture the 'essence' of boundary shape with fewer possible no. of polygonal segments.

Perfect (This approximation)

when each pair of adjacent points in the boundary is assigned to a segment in the polygon.

No. of segments in the polygon } = No. of points in the boundary.

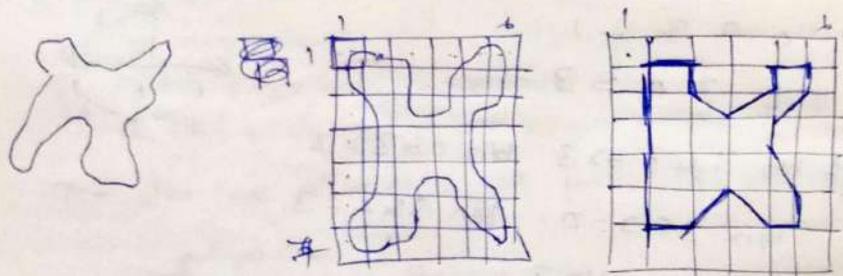
Three Important types of polygonal approximation

1. Minimum perimeter polygons
2. Merging technique
3. Splitting "

1. Min. Perimeter polygons (Perimeter = ^{outer} boundary)

Steps

1. Enclose the boundary using Sampling grid having a set of concatenated cells. Cells shows - inside & outside ^{walls} of the ~~bar~~ boundary.
2. Assume the boundary as a rubber band, which placed in bin the two walls.
3. Imagine that the rubber band is shrunk so that it touches the inside the walls.
⇒ The shape taken after shrinking gives the Min Perimeter polygon.



also called rubber band approximation

error $\Rightarrow \sqrt{2} \cdot d$; where d - distance b/w the lines in the sampling grid.

2. Merging Technique

* based on average error

Procedure

- (1) Set a threshold for least square error LSE
- (2) Merge some points along the boundary. This is done until the LSE line fit of the points merged so far..
- (3) ~~When~~ when it exceeds, store the parameters of the line & set the error.
- (4) Repeat step (1) & (3) by merging new points until the error again exceeds the threshold.

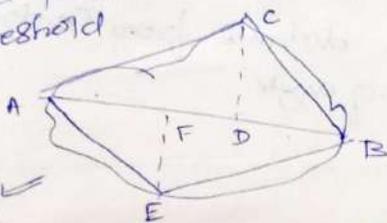
3. Splitting Techniques

Procedure

1. Subdivide the segment successively into two parts, until a specified criterion is satisfied.
2. The resulting vertices are joined in the form of the polygon.

Specific criteria

The max. perpendicular distance from a boundary segment to the line joining its two end points should not exceed a preset threshold.



AB - two segments by line AB.

EF, CD - \perp^r distance from the line about its farthest points.

obtained vertices are A, B, C, D.

Boundary Segments

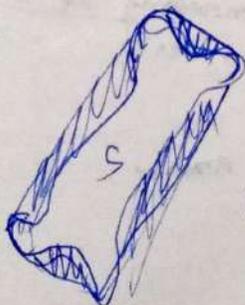
- * In which boundary is decomposed into segments.
- * It is used when the boundary has one or more concavities i.e., curves that carry shape information.

Adv.

- (i) reduce the complexity of the boundary &
- (ii) simplify the description process.

Convex. Hull: H.

a Set 'A' is said to be convex.



Proof

- (1) The convex deficiency of the set, i.e., object 'S' is defined, which is the shaded region.
- (2) The contour, i.e., outline of S is followed and points at which there is a transition into or out of the convex deficiency are marked.
- (3) These points are the partitioning points that give the segmented boundary.

Signature

* 1. a signature is a 1-D fn. representation of a boundary

* It is used to reduce the original 2D boundary to a 1D fn. so, that the description becomes easier.

* Basic idea

- (1) remove the dependency on size.
- (2) Preserve the shape.

Distance Vs Angle Signature:-

Many way to generate the signature.

One way is to plot the distance from Centroid to the boundary as a fn. of angle.

Boundary Descriptors -

Simple Descriptors -

(A) Length

* It gives a no. of pixels along a boundary.

$$\Rightarrow \text{length} = (\text{No. of vertical \& horizontal components}) \times \sqrt{2} \left[\text{No. of diagonal components} \right]$$

(2) Diameter

* It gives the diameter of the boundary is expressed as,

$$\text{Dia}(B) = \max_{i,j} [D(P_i, P_j)]$$

where $D \rightarrow$ distance measure
 $P_i, P_j \rightarrow$ points on boundary

(3) Major axis

as the line segment connecting the two extreme points of its diameter.

(A) Minor Axis

as a line \perp to Major axis.