B.TECH.

Seventh Semester Examination, 2008-09

DIGITAL IMAGE PROCESSING

Time : 3 Hours *Total Marks : 100*

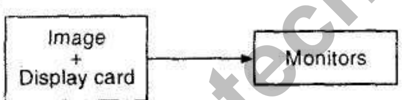
Note : (1) Attempt all questions
*(2) All questions carry **equal** marks*

Q.1. Attempt any four parts of the following:

Q.1 (*a*). Differentiate between image and scene. How is an image displayed on a display unit? Explain.

Ans. An ***image*** may be defined as a two-dimensional function, $f(x, y)$, where x and y are spatial plane coordinates and the amplitude of f at any pair of coordinates (x, y) is called the intensity of gray level of the ***scene*** at that point. So, image is a finite element of an infinite scene.

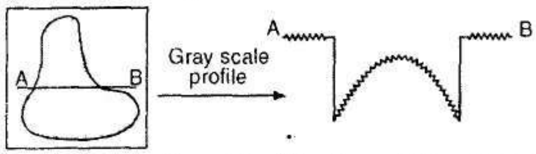
How Image is displayed : Monitors (or display unit) as driven by the outputs of image and graphics display cards that are an intergral part of the computer system.



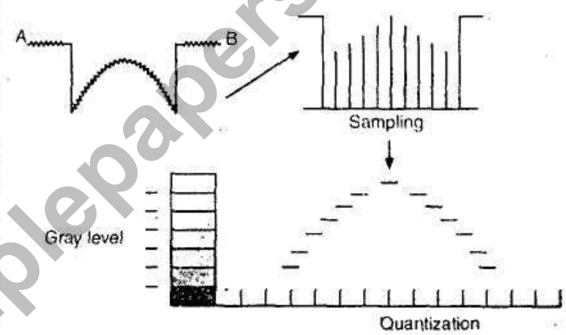
In few cases, there are requirements for image display applications that cannot be displayed by display card.

Q.1 (*b*). Explain the sampling and quantisation of images in details with the help of suitable example.

Ans. Digitizing the co-ordinate values is called sampling. Digitizing the amplitude values is called quantizatim. As shown below an image (a continuous image) is scanned along a scan line from A to B in the continious image.



A scan line from A to B in the continous image, show variation of gray values also called as gray scale profile.



Q.1 (*c*). Discuss the Histogram specification.

Ans. Histogram specification : This is method used to generate a processed image that has a specified histogram. This process is also called as histogram matching. As against of histogram equalization technique, histogram specification to not use uniform histogram for image enhancement, instead it specify the shape of the histogram that we wish the processed image to hard.

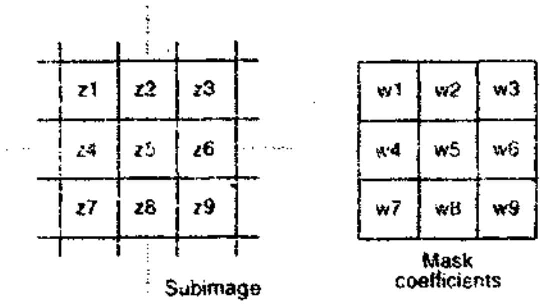
Histogram specification is a trial-and error processor. That is one can use guidelines learned from the problem at hand. So in general, Here are no rules for specifying histograms, and must resort to analysis on a case-by-case basic for any given enhancement task.

Q.1 (*d*). Discuss the image enhancement in spatial domain and also discuss the spatial domain methods.

Ans. 1. Mask -processing function : (Filter) As shown below there is a pixel $f(x, y)$, $z5$ and their neighborhood pixel, a 3 × 3 matrix (w1, w2 ... w9) also called as MASK, KERNEL. TEMPLATE or WINDOW. The term spatial domain refers to aggregate of pixel composing an image in spatial domain, image processing operate directly over these pixel as function given below :

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

Where $f(x, y)$ is the input image, $g(x, y)$ is the processed image, and T is an operation on f, defined over some neighborhood of (x, y) using mask coefficients.



Neighborhood of (x, y) are 3 × 3 matrix pixel around it, *i.e.*, neighborhood of pixel (x, y) is given by a MASK and resultant pixel value is given be SUM of product of $z5 = z = w_1 z_1 + w_2 z_2 + ... + w9\ z9$.

2. Gray-level Transformation function: Neighborhood of (x, y) is NONE, *i.e.*, no neighbour pixel.

(We call this gray-level Transformation function or intensity mapping function).

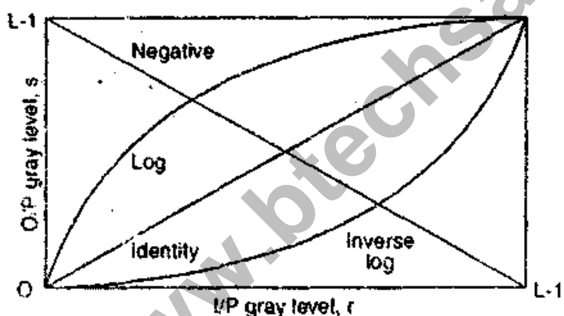
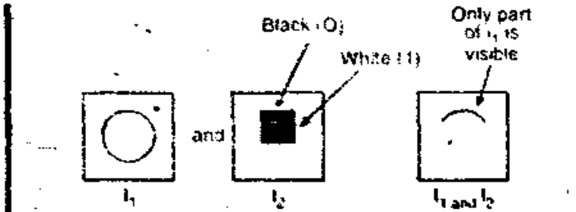


Fig. I/P gray level, r

3. Use of Arithmetic and logical operation between images : In addition T can operate on a set of input image, such as performing the pixel-by-pixel sum of images for noise reduction ($I = I1 + I2 + I3 + ... In$).

Arithmetic operation: Gray level of two images are added or multiplied with constant factor.

Logical Operation : Two images are logical operated. (I_1 AND I_2) implies only those region in I_2 will be marked and will be in output.



Q.1 (*e*) Explain the contrast stretching with the help of example.

Ans. Contrast stretching comes under piece-wise linear transformation function. Low-contrast image can result from poor illumination which in-turn cause lackness of dynamic range "All value of grey value are used in image". Thus contrast stretching is used to improve dynamic range. Here either "Lower range is suppressed while higher range is elevated" OR "Lower range is elevated while higher range is suppressed". Figure is discribed below.

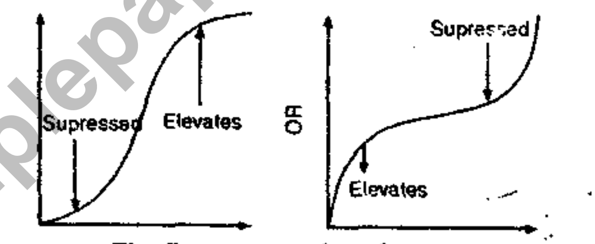


Fig. Contrast stretching function.

Fig. (*a*) will be used for dark shade image while figure (*b*) function will be used for right shade image in gray scale.

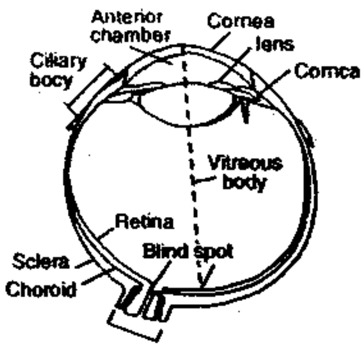
Q.1 (*f*) Explain visual perception and also discuss the visual system and elements of visual perception.

Ans. The elements that govern visual perception are:

1. Structure of eyes
2. Image fomation
3. Brightness adaptation and Discrimination

Basically all comes under physiophy scological behaviour of human being or living animal.

1. Structure of eyes: There are two classes of receptors : Cones and Rods in our eyes, area of distribution and the fact that several rods are connected to a single nerve end reduce the amount of detail discernible of these receptors. Rods severe to give a general, overall picture of the field of view. They are not involved in color vision and are sensitive to low levels of illumination. For example, objects that appear brightly colored in daylight, when seen by moonlight appear as colorless forms because only the rods are simulated. This phenomenon is known as scotopic or dim-light vision.



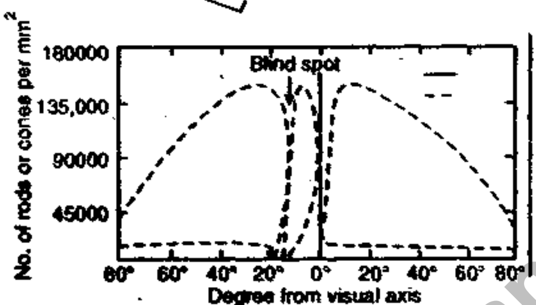
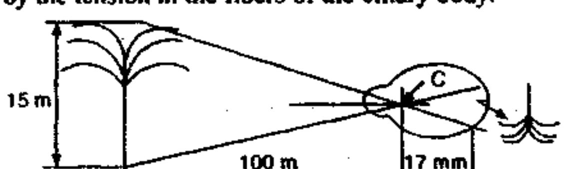


Fig. Degree from visual axis

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

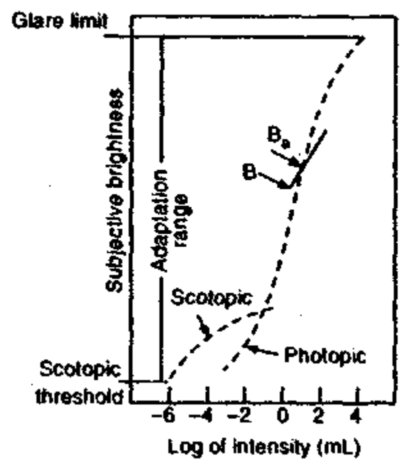
2. Image formation: The principal difference between the lens of the eye and an ordinary optical lens is that the former is flexible. As illustrated in figure the radius of the curvature of the anterior surface of the lens is greater than the radius of its posterior surface. The shape of the lens is controlled by the tension in the fibers of the ciliary body.



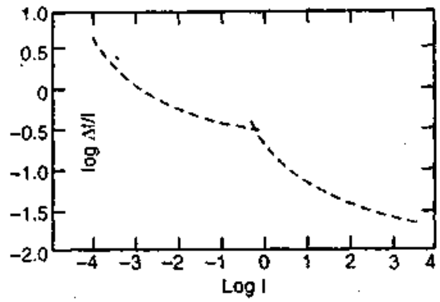
- The focus on distant objects the controlling muscles cause the lens to be relatively flattened. Similarly, these muscles allow the lens to become thicker in order to focus on objects near the eye.
- The distance between the focal center of the lens and the retina varies from approximately 17 mm to about 14 mm, as the refractive power of the lens increases from its minimum to its maximum.
- When the eye focuses on an object farther away than about 3 m, the lens exhibits its lowest refractive power, and when the eye focuses on a nearby object the lens is most strongly refractive.

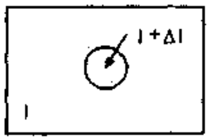
3. Brightness Adaptation and Discrimination: Figure below showing two graph and on experiment that explains brightness adaptation and discrimination of human visual system.

As in figure, because digital images are displayed as a discrete set of brightness points, the eye's ability to discriminate between different brightness 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 pow (10) from the scotopic (darkness) threshold to the glare (too bright) limit. Considerable experimental evidence indicates that subjective brightness (brightness as perceived by the human visual system) is a logarithmic function of the light intensity incident on the eye. ΔI intensity change that eye can perceive is high in photopic



region while it is less in scotopic region, that's why log $\Delta I/Ic$ "weberr ratio" reduces when eye adjust





from scotopic region to photopic region as shown in fig.

Q.2. Attempt any four parts of the following:

(*a*) Describe the filtering process with its advantages and disadvantages.

Ans. Filters are mask in frequency domain denoted by $H(u, v)$. In spatial domain filters are NXN convolution matrix. Correspondence between filter in frequency domain and spatial domain is given as:

$$g(x, y) = f(x, y) \otimes h(x, y)$$
$$DFT \downarrow\uparrow_{IDFT} \quad \downarrow \uparrow \qquad \downarrow \uparrow$$
$$G(u, v) = F(u,v) \times H(u, v)$$

Filters in frequency domain are discrete fourier transform of convolution mask in spatial domain.

$$H(u, v) \rightleftharpoons h(x, y)$$

So, by multiplication of filter function $H(u, v)$ with $F(u, v)$ we can achieve filtering in frequency domain.

For example: Lowpass filtering, it is known that edges and other sharp transitions (such as noise) in gray level of an image is due to the contribution of high-frequency component is spatial domain or high-frequency content in fourier domain. Hence by removing a specified range of high-frequency content in fourier domain with use of low pass filter as :

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

Advantage of Filtering Process (Frequency domain):

- Image enhancement is possible, like removal of periodic noise.
- Help in analysis and understanding of image.

Disadvantage of Filtering Process (Frequency domain):

- Overhead of calculation of Fourier Transform and Inverse Fourier Transform.
- In practical application spatial domain filter are better than frequency domain filter, as filters are implemented directly using firmware/hardware based on spatial domain.

Q.2 (*b*) What is the purpose of using short-time Fourier transform and differentiate it with the fractional fourier transforms.

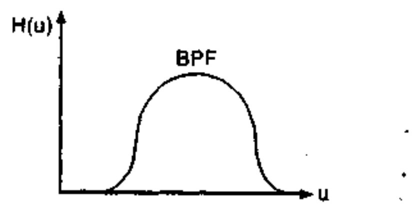
Ans. The short-time Fourier transform (STFT), or alternatively short-term Fourier tranform is a Fourier-related transform used to determine the sinusoidal frequency and phase content of the local sections of a signal as it changes over time.

STFT core frequently used to analyze music. The height of each bar repersents the amplitude of frequencies within that band. The depth dimension represents time, where each new bar was a separate distinct transform. Audio engineers use to find frequencies which may be more or less resonant in the space where the signal was recorded. This information can be used for equalization or turning other audio effects.

In the area of harmonic analysis, the fractional Fourier transform (FRFT) is a linear transformation generalizing the Fourier Transform. It can be thought of a the Fourier transform to the n-th power where n need not be an integer thus, it can transform a function to an intermediate domain between time and frequency. Its application range from filter disign and signal analyser and pattern recognition.

Q.2 (c) Discuss the operation of Bandpass filter.

Ans. A bandpass filter alternates very low and very high frequencies, but retains a middle range band of frequencies.



Bandpass filter can be used to enhance edges (suppressing low frequency) while reducing noise at the same time. There are applications in which it is of interest to process specific bands of frequencies or small regions of frequency rectangle. Filters in the first category are called band reject or bandpass filters, respectively. Filters in second category are called notch filters.

A bandpass filter is oppared from bandreject filter as

$$\boxed{H_{BP}(u, v) = 1 - H_{BR}(u, v)}$$

Q.2 (*d*) Discuss the point spread fuunction (PSF) of blurring.

Ans. The point-spread function (PSF) model of blurring : Most blurring processes can be approximated by convolution integrals, also known as Fredholm integral equations of the first kind. The blurring is characterized by a Point-Spread Function (PSF) or impulse response. The PSF is the output of the imaging system for an input point source. All the blurring processes considered in this are linear and have a spatially invariant PSF.

For discrete image processing, the convolution integral is replaced by a sum. The blurry image $x(n,m)$ is obtained from the original image $s(n, m)$ by this convolution:

$$x(n, m) = \sum_{a=-\infty}^{+\infty} \sum_{b=-\infty}^{+\infty} s(n+a, m+b)h(-a, -b)$$

The function $h(n, m)$ is the discrete point spread function for the imaging system. Also of interest is the Discrete Fourier Transform (DFT) representation of the point-spread function, given by

$$H(u, v) = \sum_{n=0}^{N-1} \sum_{b=0}^{M-1} h(n, m) e^{-2pj\left(\frac{un}{N} + \frac{vm}{M}\right)}$$

for $\quad u = \lfloor -N/2 \rfloor + 1, \ldots, \lfloor N/2 \rfloor$

and $\quad v = \lfloor -M/2 \rfloor + 1, \ldots, \lfloor M/2 \rfloor$

$H(u, v)$ gives a set of coefficients for plane waves of various frequencies and orientations. These plane waves, called spatial frequency components, reconstruct the PSF exactly when multiplied by the coefficients $H(u, v)$ and summed. The function $H(u, v)$ is referred to as the *transfer function*, or *system frequency response*. By examining $|H(u, v)|$, one can quickly determine which spatial frequency components are passed or attenuated by the imaging system.

As an example, consider this 3 × 3 mask which can be used to model small amounts of blurring :

$$\frac{1}{15}\begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 1 \end{bmatrix} \qquad \ldots(3)$$

The DFT of this mask is:

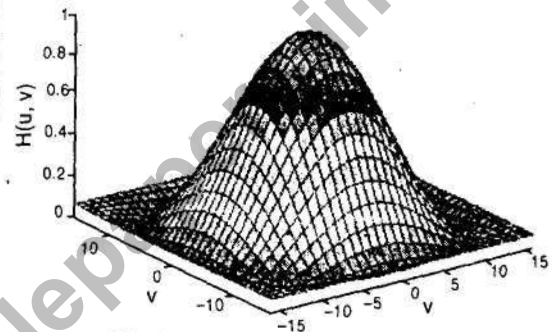


***Fig.** $|H(u, v)|$ for a 3 × 3 blurring mask, with N=M=33*

$$H(u, v) = \frac{1}{15}\left[3 + 4\cos\left(\frac{2\pi u}{N}\right) + 4\cos\left(\frac{2\pi v}{M}\right) + 4\cos\left(\frac{2\pi u}{N}\right)\cos\left(\frac{2\pi v}{M}\right)\right]$$

Figure shows a plot of $|H(u, v)|$. Near $(u, v) \approx (0, 0)$, the transfer function has $|H(u, v)| \approx 1$.

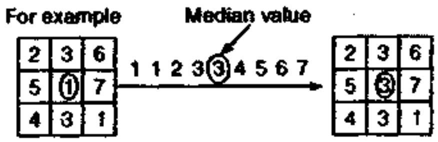
This indicates that low-frequency components are passed. Near the perimeter of the plot, $|H(u, v)| \approx 0$, meaning that high frequency components are blocked.

Q.2 (*e*) Explain the purpose of different order statistics filters.

Ans. Order-statisfies filters are non linear spatial filters whose response is passed on ordering (ranking) the pixels contained in the image area encompassed by filter, and then replacing the value of the center pixel with the value determined by the ranking result.

The best-known example in this category is median filter, which as its name implies, replace the value of a pixel by the median of gray levels in the neighborhood of that pixel (the original value of the pixel is included in the computation of the median).

For exmaple



Median filters are particularly effective in the pressence of impulse noise, also called salt-and-pepper noise because of its appearance as white and black dots superimposed on an image.

Other filter in this clour are Max., Min. filter.

Q.2 (*f*) Show that Fourier transform and its inverse are linear process.

Ans. Central to the study of linear systems and the Fourier transform is the concept of an impulse and its shifting property. A unit impulse of a continuous variable t located at $t = 0$, denoted $\delta(t)$, is *defined* as

$$\delta(t) = \begin{cases} \infty \text{ if } t = 0 \\ 0 \text{ if } t \neq 0 \end{cases} \qquad ...(1)$$

and is constrained also to satisfy the identity

$$\int_{-\infty}^{\infty} \delta(t)\, dt = 1 \qquad ...(2)$$

Physically, if we interpret t as time, an impulse may be viewed as a spike of infinity amplitude and zero duration, having unit area. An impulse has the so called *sifting property* with respect to integration,

$$\int_{-\infty}^{\infty} f(t)\,\delta(t)\, dt = f(0) \qquad ...(3)$$

provided that $f(t)$ is continuous at $t = 0$, a condition typically satisfied in practice. Sifting simply yields the *value* of the function $f(t)$ at the *location* of the impulse (*i.e.*, the origin, $t = 0$, in the previous equation). A more general statement of the sifting property involves an impulse located at an arbitrary point t_0, denoted $\delta(t - t_0)$. In this case, the sifting property becomes

$$\int_{-\infty}^{\infty} f(t)\,\delta(t - t_0)dt = f(t_0) \qquad ...(4)$$

which yields the value of the function at the impulse location, t_0. For instance, if $f(t) = \cos(t)$, using the impulse $\delta(t - \pi)$ in equation yields the result $f(\pi) = \cos(\pi) = -1$. The power of the sifting concept will becomes quite evident shortly.

Let x represent a *discrete* variable. The **unit discrete impulse**, $\delta(x)$, serves the same purposes in the context of discrete systems as the impulse $\delta(t)$does when working with continuous variables. It is defined as

$$\delta(x) = \begin{cases} 1 & x = 0 \\ 0 & x \neq 0 \end{cases} \qquad ...(5)$$

Clearly, this definition also satisfies the discrete equivalent of equation (2)

$$\sum_{x=-\infty}^{\infty} \delta(x) = 1 \qquad ...(6)$$

The sifting property for discrete variables has the form

$$\sum_{x=-\infty}^{\infty} f(x)\,\delta(x) = f(0) \qquad ...(7)$$

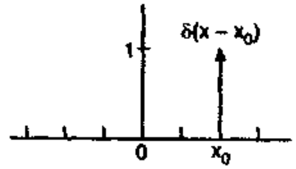
or, more generally using a discrete impulse located at $x = x_0$,

$$\sum_{x=-\infty}^{\infty} f(x)\,\delta(x - x_0) = f(x_0) \qquad ...(8)$$

As before, we see that the sifting property simply yields the value of the function at the location of the impulse. Figure shows the unit discrete impulse diagrammatically. Unlike its continuous counterpart, the discrete impulse is an ordinary function.

Of particular interest later in this section is an impulse train, $s_{\Delta T}(t)$, defined as the sum of infinitely many *periodic* impulses ΔT units apart:

$$s_{\Delta T}(t) = \sum_{x=-\infty}^{\infty} \delta(t - n\Delta T) \qquad ...(9)$$



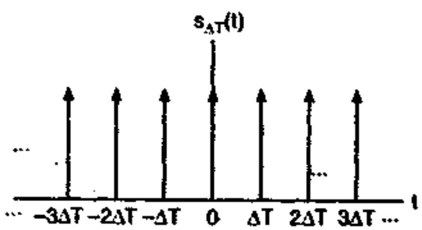


Figure shows an impulse train. The impulses can be continuous or discrete.

Q.3. Attempt any two parts of the following:

(a) Explain the rescaling and resampling. Differentiate between thinning and thickening.

Ans. Resampling: For any 2-D interpolation in image processing an image resizing is required (Zooming and shrinxing). Zooming may be viewed as over-sampling, while shrinking may be viewed as underr-sampling.

For instance, to double the size of an image, we duplicate each column. This doubles the image size in horizontal direction. Then, we duplicate each row of the enlarged image to double the size in the vertical direction.

Rescaling : It is a geometric transformations that modify the spatial relctorshps etween pixels in an image. These transformation is called "rubber-sheet" transformation because they may be viewed as analogous to "printing" an image on a sheet of rubber and stretching the sheet according to a predefined set of rules.

Q.3 (b) Explain pseudocolor image processing with examples.

Ans.

Thinning	Thickening
The thinning of a set A by structuring element B, denoted $A \otimes B$, can be defined in terms of hit-or-miss transform $A \otimes B = A - (A \otimes B)$ $= A \cap (A \otimes B)$ Let $\{B\} = \{B', B^2, ...B^h\}$ where B_1 is rotated version of B^{i-1}. Using above sets, thining can be defined by a sequence of structuring element as	Thickening is the morphological dual of thinning and is defined by the expression $A \odot B = A \bigcup (A \otimes B)$ where B is a structuring element suitable for thickening. Thickening can be defined as a sequential operation
$A \otimes \{B\} = ((... ((A \otimes B') \otimes B^2) ...) \otimes B^h)$	$A \odot \{B\} = ((...((A \odot B^2)...) \odot B^n)$

Q.3 (c) Describe the convex hull method. Explain the segmentation of image processing.

Ans. A set A is said to be convex if the straight line segment joining any two points in A lies entirely within A. The convex hull H-S is called convex deficincey of S.

Simple algorithm for obtaining the convex hell $c(A)$ of a set A.

Let B^i, $i = 1, 2, 3, 5$ respect the four structuring elements in figure below.

B^1 B^2 B^3 B^4

The procedure consists of implementing the equation

$$X_K^i = (X_{K-1} \otimes B_i) \quad ...(1)$$
$$i = 1, 2, 3, 4$$
$$K = 1, 2, 3 ...$$

with $X_0^i = A$, when the procedure converger (*i.e.*, when $X^i_K = X^i_{K-1}$), then convex hull of A is

$$C(A) = \bigcup_{i=1}^{4} D^i$$

and where, $D_i = X^i_K$

Image segmentation : Segmentation is the first step towards productging the descriptions on an input image. Here input and output are 'still' images, but output is an abstract representation of the input. Segmentation technique basically divides the spatial domain, on which the image is defined, into 'meaningful' parts or regions. This meaningful region may be complete object or may be part of it. The segmentation algorithm makes systematic use of physcally measured images feature, such as historgram and probability distributions function function to extract region's within images.

The performance of segmentation algorithm is measured based on the 'meaning' associated with the extracted regions.

For example : pdf generated from a meaningful region within a image will always remain same. So by finding pdf within a region we can find out meaningful region.

Property satisfied by regions: Let the spatial domain on which the image is defined as D (rectangular array of size M × N)

$D = \{(x, y)\}, x = 0, 1, 2, ... M-1,$
$y = 0, 1, 2, ... N-1\}$

Say, there are n regions ($R_1, R_2, R_3 ....., R_n$), then property are

1. $\bigcup_{i=1}^{n} R_i = D$ ⇒ Every pixel of the image domain is mapped to one region or other.
2. $R_i \bigcap R_j = \phi$ ⇒ It ensures that a pixel is mapped to only one region
3. Prop (R_i) True ⇒ It indicates that regions are defined based on same property.
4. $R_i \bigcup R_j$ = False, if R_i and R_j are adjacent ⇒ Maximality of each region is assured.

Theresholding selection methods : There all three types of thresholding methods LOCAL, DYNAMIC and GLOBAL. The choice of threshold gray level (t) depends on two aspects : coordinate of pixel (x, y) and value of pixel $P(x, y)$.

Thus, $t = t(x, y, P(x, y))$.

Local threshold : If 't' depends on the pixel value, $P(x, y)$.

Dynamic threshold : If 't' depends both on pixel position (x, y) as well as on the pixel value $P(x, y)$.

Global threshold: (Also called as postion-independent threshold). If 't' is independent of pixel position.

Q. 4. Attempt any two parts of the following:

(a) Discuss the geometric transformation of images using spline interpolation.

Ans. Geometric Transformation of Images using Spline-Interpolation : To perform geometric transforms on discrete images such as a rotation or zooming we need to first fit the discrete data to a continuous function. This can be done using splines. B-splines can be used to interpolate and form the continuous image from the discrete samples.

Here we illustrate the use of splines in two-dimensions for transforming discrete images. When we apply a transformation (such as rotation or zooming) it becomes necessary to know the image intensity at a location in between the sample points. This is an interpolation problem and splines come in handy.

Image Transformation : Let $[x, y] = G([u, v])$ be the transformation (rotation, scaling) of the domain, where $[u, v]$ describe the transformed plane and $[x, y]$ the original image plane. Then our problem is to find the samples of the transformed image $g([m, n])$ from the samples of original image $f([k, l])$. Here is how we go about it:

- Calculate $c([k, l])$ from the pixel values $f([k, l])$
- For each $[m, n]$ on transformed image find corresponding source location $[x, y]$
- Compute $f([x, y])$ using the spline model, now $g([m, n]) = f([x, y[)$

Example 1: In figure 1 we look at the rotation of an image by certain angle and compare the fits using $\phi^n([m, n])$ $n = 0$ (nearest neighbour), $n = 1$ (bilinear), $n = 3$ (bicubic).

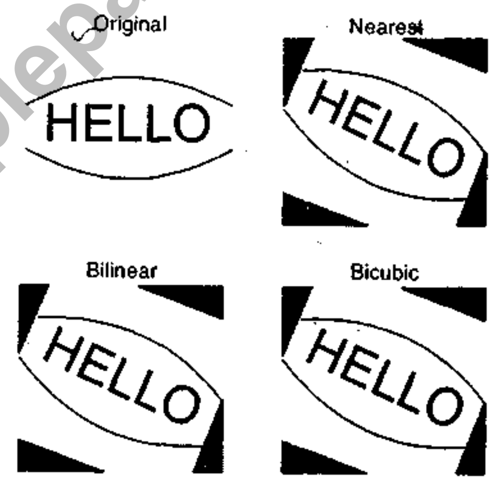
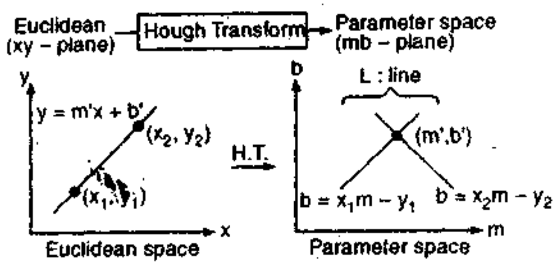


Figure 1 : Illustration of how the jagged edges appear in nearest neighbour interpolation and they progressively smooth out in higher order spline-interpolation. Bicubic spline interpolation is highly popular as cubic appear smooth to the human eye.

Q. 4 (b) Explain the Hough transforms. Differentiate between region merging and region splitting.

Ans. Human transform: This method of edge linking determines whether or not a set of edge pixels lie on a line. This method then links the edge pixel by producing line or curve.



In the mb-plane, the line is represented by point (m', b'). It can be seen that infinite number of lines pass through (x_1, y_1) in the xy-plane which correspond to line $b = x_i m - y_i$ in mb – plane. Similar expression is true for (x_2, y_2). Thus for any point on $y = m' x + b'$, there is line in mb-plane and the point when these all line intersect in mb-plane is denoted by (m', b'). Therefore, the problem of detecting line passing through a given set of points reduces to finding the point of intersection, of lines on parameter space, *i.e.*, lines in mb'-plane is point Euclidean space.

Region based Thresholding: This approach is sequential or iterative in nature. There are two sub-category : Region growing and region merging approaches. First one starts from single pixel called seed pixel and end up with region around seed pixel and second approach merger adjacent homogeneour region till a single homogenous region.

Second approch of (Region based Thresholding) is called ***Region splitting and Merging***, here following procedure is followed:

1. Split into four disjoint guadrants any region R_i for which $Q(R_i)$ =FALSE.
2. When no further splitting is possible, merge any adjacent regions R_i and R_K for which $Q(R_i \cup R_K)$ = True.
3. Stop when no further merging is possible.

Q.4 (*c*) Discuss stereo imaging. Explain the various thresolding algorithms.

Ans. Computer vision is an imitation of human visual systemand steregineging is process/principle on which computer vision is based upon.

Two objective of stereo imaging are:

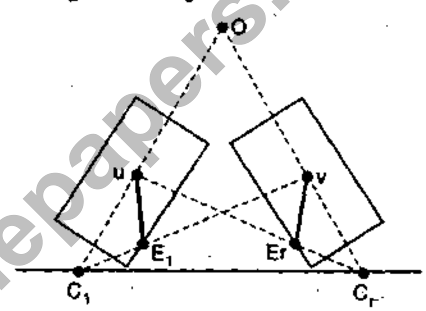
(*i*) Corresponding between pair of stero image

(*ii*) Recovery of depth information.

First objective is a problem of recovery of depth information. This can be done when first objective is achieved *i.e.*, "correspondence is established". With help of simple geometrcal formula depth is recovered by knowing physical position of eyes (visual system).

The process of establishment for correspondence in stereo imaging: Establishment of the correspondence is described as a problem of "finding conjugate points". First, stereo image pair, left camera (eye) image and right camera (eye) image are captured by bincoular imaging system.

In second step, the same entity in bouthversion of image is detected *i.e.*, finding same landmarks. Let the entity identified is object '*O*' in real world and '*u*' is the point in left image *L* and '*v*' is the point in right image *R* for object '*O*'.



As shown above, let C_l and C_r are two optical centers of the left and right images *L* and *R*, respectively. The projection of images point *u* at the right image plore *R* is called as epipole E_r. Similarly E_l for v. The line joining E_r and v is the epipolar line of the image point *u* in *L* image. The corresponding point of v in the right image plane *R* will always lie on epipolar line of v. These search space reduces from 2-*D* space to 1-*D* epipolar lines.

Hence, epipolar line can be used for tracing out the correspondence.

Various thresholding algorithms are multi-level thresholding, local thresholding and region based thresholding.

Multi-level Thresholding : It is extension to any thresholding methods where image contains more then 2-regions *i.e.*, $(n > 2)$. Thus gray level histogram of the image is expected to contain *n* number of peaks (nodes) each of which corresponds to distinct type of region. Thus then are $(n - 1)$ valleys in the histogram and it is required to select gray levels corresponding to bottom of these valleys as threshold for image segmentation.

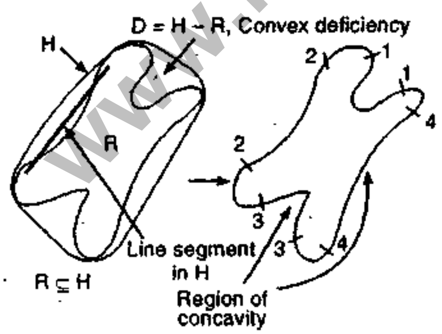
Local Threshlding : In case the gray level ground overlap each other significantly. That is object in image have a good local contrast but global contrast is bad, hence object is not clearly visualized. This object can be segmented out by dividing the whole image into smaller ones so that the variation in illumination over each of these sub-images is negligible. Finally, each sub-image is segmented independently, and the segmented sub-images are put together in appropriate order to get segmented vision of original image.

Region based Thresholding: This approach is sequential or iterative in natural. There are two sub-category: region growing and region merging approaches. First one starts from single pixel called seed pixel and end up with a region around seed pixel and second approach merges adjacent homogeneous region till a single homogeneous region.

Q.5. Attempt any four parts of the following :

(*a*) Describe recursive boundary splitting.

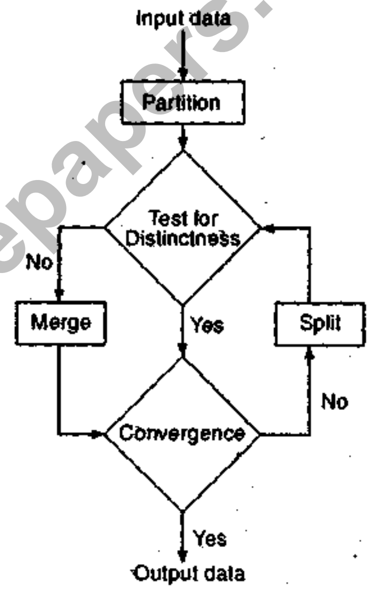
Ans. A region *R* is convex if and only if any two points from *R*, the whole line segment defined by its end-points is inside the region *R*. The convex hull of a region is the smallest convex region *H* which satisfies the condition *R* is a subset of *H*. The convex hull has some special properties in digital data which do not exist in the continuous case. For instance, concave parts can appear and disappear in digital data due to rotation, and therefore, the convex hull is not rotation invariant in digital space. The convex hull can be used to describe region shape properties and can be used to build a tree structure of region concavity.



Q.5 (*b*) What is clustering?

Ans. A general clustering algorithm is based on a *split* and *merge* technique, as shown in figure below. Using a *similarity measure* (e.g., the dot product of two vectors, the weighted Euclidean distance, etc.), the input vectors can be partitioned into subset, each of which should be sufficiently distinct. Subsets which do not meet this criterion are merged. This procedure is repeated on all of the subsets until no further spliting of subsets occurs or until some *stopping criteria* is met. The basic of groups (or classes) so that a predefined set of criteria are satisfied. Criteria usually include inter-class and intra-class distance, density points within class, etc.

General clustering algorithm:



Q.5 (*c*). Discuss patterns and pattern clauses.

Ans. Suppose that we wanted to distinguish football players and jockeys by their physical characteristics. Height would be one characteristic, and by itself would probably be a good discrimination. What about football and basketball players? Height alone might not be enough. Let's add weight into the decision. If we plot a sampling of jockeys, football players and basketball players along axes of height and weight, the jockeys would be clustered down in the low-height section, and the basketball players might be clustered in the taller-than-football-player height but slightly-less-than-football-player weight section.

Clearly, there will be some overlap. Many football players could be playing basketball, and many do indeed play both.

Even if we can't tell perfectly, we can make reasonable guesses based on these statistical properties. This is statistical pattern recognition.

The quantities that measure in order to make our decision are called features. The items that we are trying to classify the patterns are called classes.

The vector space spanned by the features is called a features space.

Q.5 (*d*). Explain the tree grammars and tree automata.

Ans. An object may be represented by chain-code of its boundary, or by its skeleton approximated with line and arc segments. One of the most prevalent methodology used for recognizing objects based on such features is syntactic to change. The synthatic approach to pattern recognition problems provides a powerful method for describing a large set of complex pattern recognition problems provides a powerful method for describing a large set of complex patterns by using small sets simple pattern primitives and of grammatical rules that are recursive in nature. The technique has been adopted from formal language theory, and is sometimes called linguistic approach. Basic concept of syntactic pattern recognition is the specification of structural features, called primitives and a set of rules, called production rules. These rules govern interconnection between the said primitives. Thus they form a grammar.

For example grammar for string recognitions or classification.

A grammar may be defined as a 4-tuple:

$$G = \{T, V, s, P\}$$

Where

- *t* — The finite set of terminal symbols or alphabets or constants (i.e., structural primitives) that represents the final substitution phase of the production of sentences (i.e., patterns or objects).
- v — The finite set of non-terminal symbols or variables that represent the intermediate constructions; these variables are distinct from *T*.
- s — A special distinguished variable in V that is the sentence (i.e., pattern or object) to be constructed. So it is the starting symbol.
- *P* — The finite set of production rules defining how symbols can be combined to form the sentences (*i.e.*, patterns or objects).

A language $L(G)$ is the set of all sentences that can be produced by the grammar G. In the present context, language corresponds to a class. Finally, the parsing of a sequence of symbols enables us to determine whether the sequence is a sentence in the language. Thus in syntactic pattern recognition gives a set of grammars each of which correspnds to unique class and a sequence of primitives we parse the sequence using each grammar to determine whether the sequence is a valid pattern of the corresponding class.

Production rules are the most critical parts of the grammar or the system. String grammars are characterized primarily by the form of their productions. By putting certain restrictions on the form of the rules we can define different classes of grammars. Of particular interest in syntactic pattern recognition are the *regular grammars*, whose production rules have the following characteristics. The left-hand side of each rule contains only a single non-terminal symbol, and the right-hand side must contain either a terminal symbol by itself or a terminal symbol followed by a non-terminal symbol. However, two other grammars may be of our interest-context-sensitive and *context-free*. In case of context-sensitive grammar, length of left-hand side of rule is less than or equal to that of the right-hand side, and left-hand side may have more than one non-terminal symbol. On the other hand, in context-free grammar left-hand side of all the rules have only one non-terminal symbol. However, in both the cases, right-hand side of rule may have any sequence of symbols. Thus in the case of regular grammars productions are always of the form $A \rightarrow aB$ and $A \rightarrow a$ with $A, B \in V$, and $a \in T$, and in context-free grammar, production are of the form $A \rightarrow a$ with $A \in V$, and $a \in (T \cup V)^* - \phi$; that is a can be any string composed of terminals and non-terminals, except the empty string.

Q.5 (*e*). Discuss the training algorithms.

Ans. A tranining algorithm that maximizes the margin between the training patterns and the decision boundary is presented. The technique is applicable to a wide variety of classification functions, including perceptrons, polynomials, and Radial basis functions. The effective number of parameters is adjusted automatically to match the complexity of the problem. The solution is expressed as a linear combination of supporting patterns. These are the subset of training patterns that are closest to the decision boundary. Bounds on the generalization performance based on the leave-one-out method and the VC-dimension are given. Experimental results on optical character recognition problems demonstrate the good generalization obtained when compared with other learning algorithms.

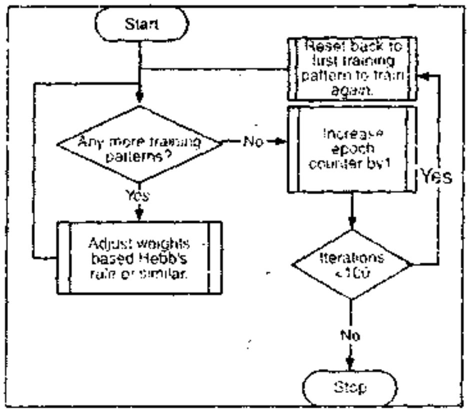


Fig. Unsupervised Training

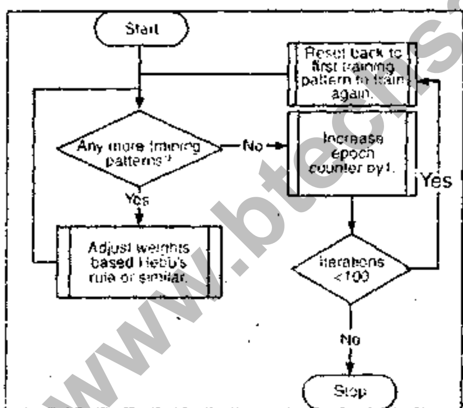


Fig. Supervised training

Q.5 (f). Explain the flow-diagram of image analysis and understanding methods.

Ans. Objective of description process is to identify meaningful segments out of a scene.

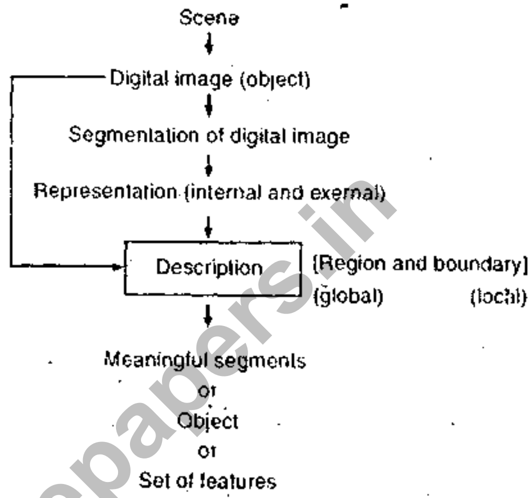


Fig.

After the choice of [internal, external or both] forms of representation. Next step is to convert the representation into description.

Representation $\xrightarrow{\text{Converted to}}$ Description

Self features or description is a pattern space with reduce dimensionality such that it is only necessary for recognition of the object.

OR

Feature is a translation of an object in an image into its abstract notion using attributes or dimension.

The objective of above conversion is to choose descriptors [attributes or dimension] that captures essential [minimum or reduce] differences between objects, or classes of objects while maintaining as much independence as possible to change in factors such as location, size and orientation.

— — —