B.TECH.

SEVENTH SEMESTER EXAMINATION 2009-10

DIGITAL IMAGE PROCESSING

Time : 3 Hours | *Total Marks : 100*

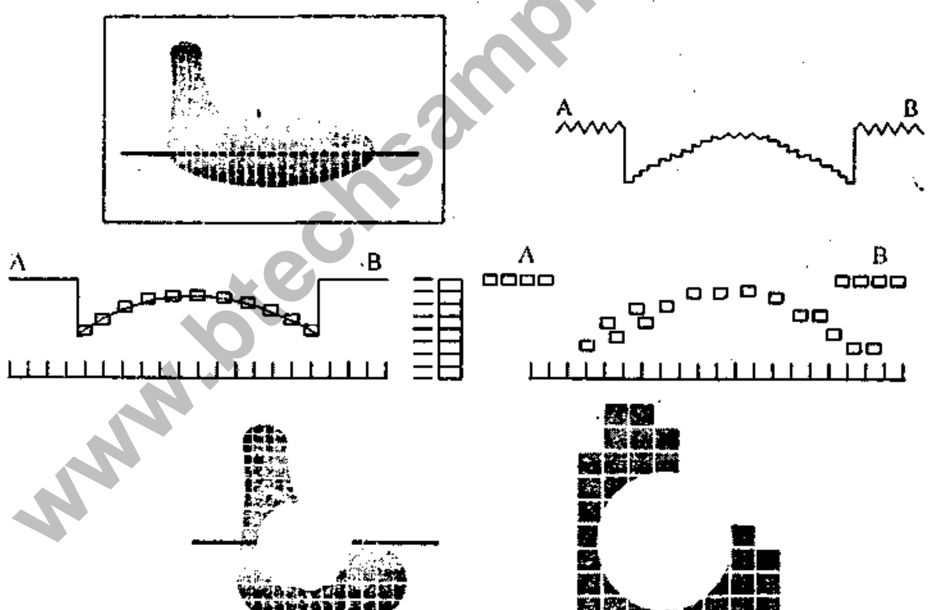
Note: (i) Attempt all questions.

(ii) Each question carries equal marks.

Q.1. Attempt any four parts of the following: **(5×4=20)**

Q.1. (a) Explain the sampling and quantisation of images with the help of suitable diagram.

Ans. Digitizig the co-ordinate values is called sampling. Digitizing the amplitude values is called quantization. As shown below an image (a Continuous image) is scanned along a scan line from A to B in the continuous image. Sampling and quantization against grey scale strip and finally digitally scaned line is shown below:



Digitizing the amplitude values

(b) Discuss the Visual system and elements of visual perception.

Ans. The three elements that govern visual perception are:

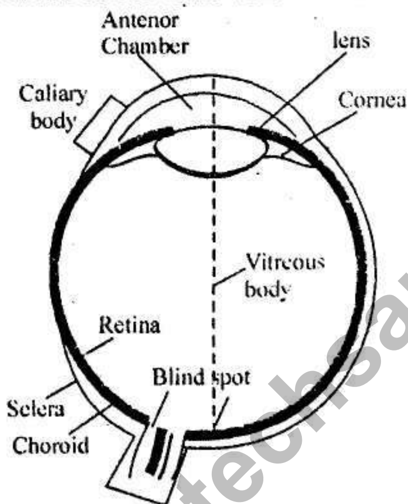
1. Structure of eyes
2. Image formation
3. Brightness Adaptation and Discrimination

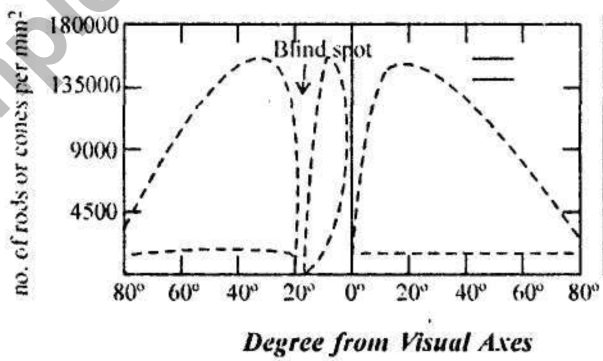
Basically all comes under physiophycological behaviour of human being or living animal.

1. Structure of eyes: There are two classes of receptors: Cones and Rods in our eyes. The rods is much larger : some 75 to 150 million are distributed over the retinal surface. The large 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 sever 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.

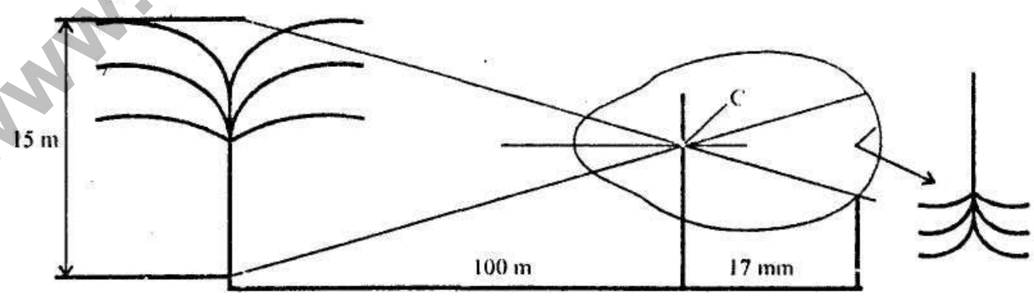
For example, objects that appear brightly colored in daylight, when seen by moonlight appear as colorless forms because only the rods are simulated. This phenomenon is known as scotopic or dim-light vision.

The Cones in each eye number 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 principle 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.

(c) Discuss the Histogram specification.

Ans. Histogram equalization automatically determines a transformation functin that seeks to produce an output image that has a uniform histogram. When automatic enhancement is desired, this is a good approach because the results from this techique are predictable and the method is simple to implement. We show in this section that there are applications in which attempting to base enhancement on a uniform histogram is not the best approach. In particular, it is useful sometimes to be able to specify the shape of the histogram that we wish the processed image to have. The method used to generate a processed image that has a specified histogram is called histogram matching or histogram specification. Let us return for a moment to continuous gray level r and z (considered continuous random variables), and let $pr(r)$ and $pz(z)$ denote their corresponding continuous probability density functions. In this notation, r and z denote the gray levels of the input and output (processed) images, respectively. We can estimate $pr(r)$ from the given input image, while $pz(z)$ is the specified probability density function that we wish the outpt image to have.

Let s be a random variable with the property

$$s = T(r) = \int_0^r p_r(w)dw \quad \text{...(i)}$$

where w is a dummy variable of integration. We recognize this expression as the continuous version of histogram equalization. Suppose next that we define a random variable z with the property

$$G(z) = \int_0^z p_z(t)dt = s \quad \text{...(ii)}$$

where t is a dummy variable of integration. It then follows from these two equations that $G(z) = T(r)$ and, therefore, that z must satisfy the condition

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

The transformation $T(r)$ can be obtained from eq. (*i*) once $pr(r)$ has been estimated from the input image. Similarly, the transformation function $G(z)$ can be obtained using eq. (*ii*) because $pz(z)$ is given.

Assuming that $G{-1}$ exists and that it satisfies conditions (*a*) and (*b*). Eq. (*i*) and (*iii*) show that an image with a specified probability density function can be obtained from an input image by using the following procedure: (1) Obtain the transformation function $T(r)$ using eq. (*i*). (2) Use eq. (*ii*) to obtain the transformatioon function $G(z)$. (3) Obtain the inverse transformation function $G{-1}$. (4) Obtain the output image by applying eq. (*ii*) to all the pixels in the input image. The result of this procedure will be an image whose gray levels, z have the specified probability density function $pz(z)$.

Although the procedure just described is straightforward in principle, it is seldom possible in practice to obtain analytical expression for $T(r)$ and for $G{-1}$. Fortunately, this problem is simplified considerably in the case of discrete values. The price we pay is the same as in histogram equalization, where only an approximation to the desired histogram is achievable. In spite of this, however, some very useful results can be obtained even with crude approximations. Which we repeat here for convenience:

$$s_k = T(r_k) = \sum_{j=0}^{k} p_r(r_j) = \sum_{j=0}^{k} \frac{n_j}{n} \qquad k = 0, 1, 2, ..., L-1 \quad (iv)$$

where n is the total number of pixels in the image, nj is the number of pixels with gray level rj, and L is the number of discrete gray levels. Similarly, the discrete formulation of eq. (*ii*) is obtained from the given histogram pz Azi B, $i = 0, 1, 2, p, L - 1$, and has the form

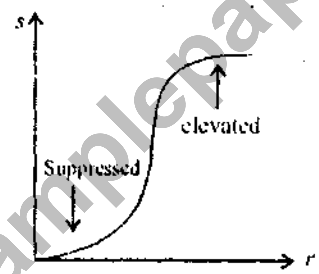
$$v_k = G(z_k) = \sum_{j=0}^{k} p_z(z_i) = s_k \qquad k = 0, 1, 2, \ldots, L-1 \qquad \ldots(v)$$

or from eq. (*iii*) $$z_k = G^{-1}(s_k) \qquad k = 0, 1, 2, \ldots, L-1 \qquad \ldots(vi)$$

Eqs. (*iii*) through (*vi*) are the foundation for implementing histogram matching for digital images. Eq. (*iii*) is a mapping from the levels in the original image into corresponding levels sk based on the histogram of the original image, which we compute from thepixels in the image. Eq. (*iv*) computers a transformation function G from the given histogram $pz(z)$. Finally, eq. (*v*) or its equivalent, eq. (*vi*), gives us (an approximatio of) the desired levels of the image with the histogram. The first two equations can be implemented easily because all the quantities are known. Implementation of eq. (*vi*) is straightforward, but requires additional explanation.

(d) Explain the contrast stretching with the help of example.

Ans. Contrast stretching: Lower range is suppressed while higher range is elevated.



(e) Explain the digital processing of camera images.

Ans.

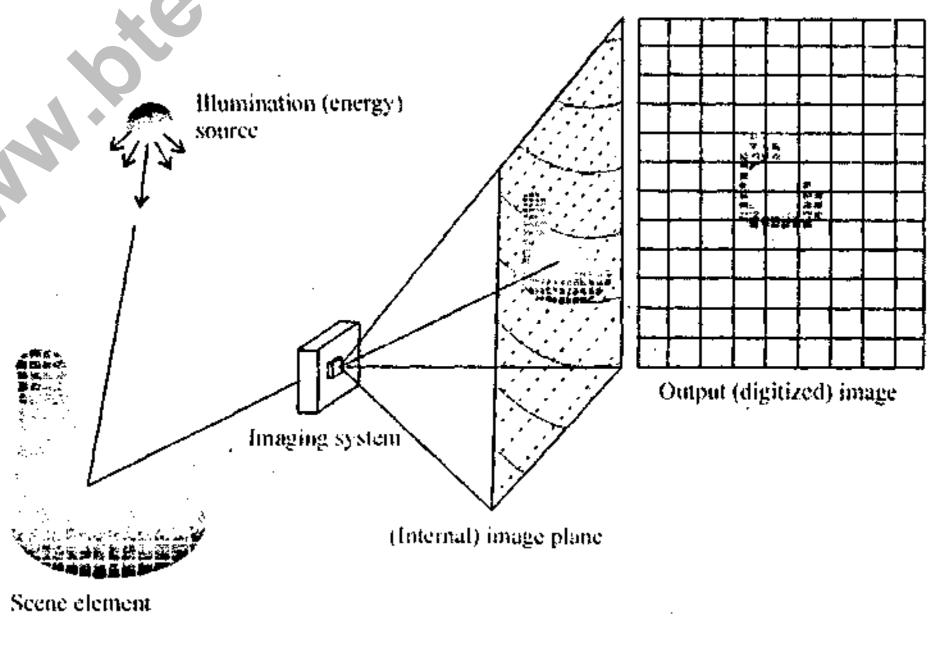


Image by two-dimensional function of the form $f(x, y)$. The value or amplitude of f at spatial coordinates (x, y) is a positive scalar quantity whose physical meaning is determined by the source of the image. Most of the images in which we are interested in this book are monochromatic images, whose values are said to span the gray scale. When an image is generated from a physical process, its values are proportional to energy radiated by a physical source (e.g. electromagnetic waves). As a consequence, $f(x, y)$ must be nonzero and finite; that is

$$0 < f(x, y) < q \qquad \text{...}(i)$$

The function $f(x, y)$ may be characterized by two components: (1) the amount of soruce illumination incident on the scene being viewed, and (2) the amount of illumination reflected by the objects in the scene. Approximately, these are called the illumination and reflectance components and are denoted by $i(x,y)$ and $r(x, y)$, respectively. The two functions combine as a product to form $f(x, y)$:

$$f(x, y) = i(x, y)\, r(x, y) \qquad \text{...}(ii)$$

where $0 < i(x, y) < \infty$ ...(*iii*)

and $0 < r(x, y) < 1$

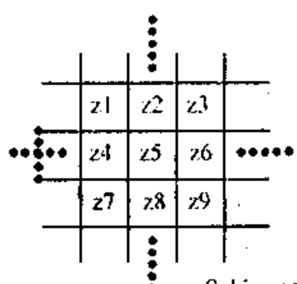
Equation (*iii*) indicates that reflectance is boiunded by 0 (total absorption) and 1 (total reflectance). The nature of $i(x, y)$ is determined by the illumination source, and $r(x, y)$ is determined by the characteristics of the imaged objects. It is noted that these expressions also are applicable to images formed via transmission of the illumination through a medium, such as a chest X-ray. In this case, we would deal with a transmissivity instead of a reflectivity function, but the limits would be the same as in eq. (*iii*), and the image function formed would be modeled as the product in eq. (*i*).

(f) Explain spatial domain methods.

Ans. 1. Mask-Processing function: (Filter) as shown below there is a pixel $f(x, y)$, $z5$ and their neighbourhood 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 gieven below:

$$g(x, y) \qquad T[f(x, y)]$$

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.



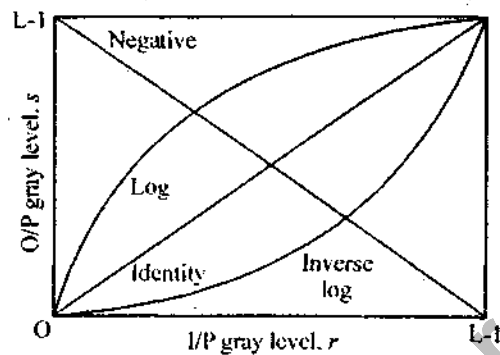
Subimage

w1	w2	w3
w4	w5	w6
w7	w8	w9

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 by SUM of product of $z5 = z = w1z1 + w2z2 + \ldots + w9z9$.

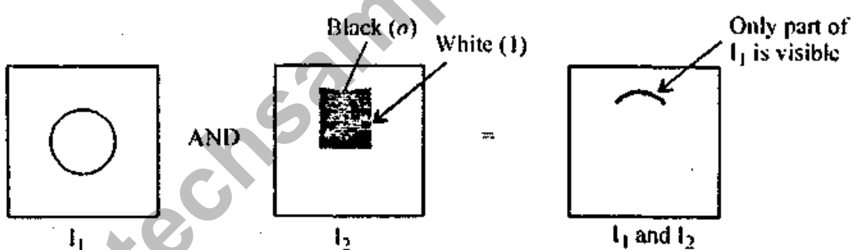
2. Gray-level Transformation function: Neighborhood of (x, y) is NONE i.e. no neighbour pixel. (We call this gray-level Transformation function or intensity mapping function).



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 ($I = I_1 + I_2 + I_3 + \dots I_n$).

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

Logical operation: Two images are logically operated, (I_1 and I_2) implies only those region in I_2 will be marked and will be in output.



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

(a) Explain the concept of filtering and its advantage.

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

$$\begin{array}{cccc} g(x, y) & f(x, y) & \otimes & h(x, y) \\ \text{DFT} \downarrow\uparrow \text{IDFT} & \text{DFT} \downarrow\uparrow \text{IDFT} & & \text{DFT} \downarrow\uparrow \text{IDFT} \\ G(u, v) & F(u, v) & \bullet & H(u, v) \end{array}$$

The filters in frequency domain are Discrete Fourier transform of convolution mask in spatial domain.

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

For example: Lowpass filtering, it is know that edges and other sharp transitions (such as noise) in gray level of an image is due to the contribution of the high-frequency component is spatial domain or high-frequency content in fourier domain.

Hence by removing or attenuating a specified range of high-frequency content in Fourier domain with use of lowpass filter as:

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

(Low pass filter)

results into smoothing or blurring of image in spatial domain when $G(u, v)$ is inverse transformed into $g(x, y)$

$$G(u, v) \xrightarrow{IDFT} g(x, y) \qquad \text{(Output image without high frequency component)}$$

Advantages of Filtering:

1. Noise reduction (arithmetic and mean filters for Gaussian noise).
2. Smoothing of false contours.
3. Reduction of irrelevant detail in an image.
4. Reduction of noise with less blurring (Median filter for salt and pepper noise).
5. Used for finding brightest points in an image (Max filter).
6. Used for finding darkest points in an image (Min filter).
7. Enhancing small detaisl preserving the background tonality.

(b) Discuss low-pass filter and high-pass filter in brief.

Ans. Image filtering is a kind of image enhancement techniques that involves a rate of change in gray-level values within an image. Image filtering cna be perfonned either in spatial domain or frequency domain.

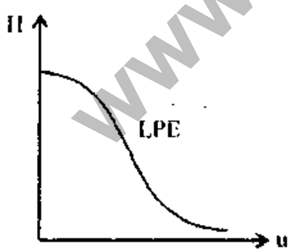
Convolution process is requied in performing image filtering in the spatial domain while multiplying Fourier transforms performs image filtering in the frequency domain.

In other words, image filtering is process to emphasis features of interest in an image. These features (required) are based on the rate of change in gray-level values which are categorized as:

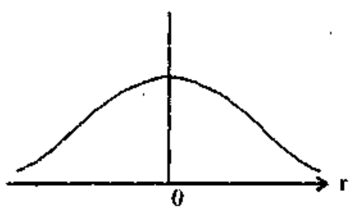
1. Low frequency component or feature "Texture".
2. High frequency component or feature "Edges".

Here frequency is regarded as how rapidly the gray-level value is changing is space (2D). As state above, LOW frequency components are responsible for the slowly varying characteristics of an image such as over all contrast and average intensity also called "Texture". While the HIGH frequency component characterize edges and other sharp detaisl in an image.

Low Pass Filter: A low-pass filter attenuates high frequency and retains low frequency unchanged. The result in the spatial domain is equivalent to that of a smoothing filter; as blocked high frequency correspond to sharp intensity changes, i.e. fine scale details and noise.



Frequency Domain

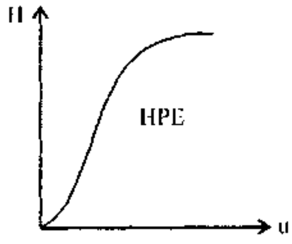


Spatial Domain

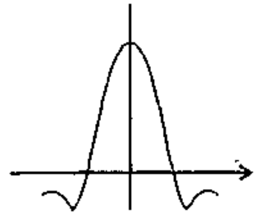
−1	−1	−1
−1	−1	−1
−1	−1	−1

(mask)

High-pass Filter: A high-pass filter, results into edge enhancement or edge detection in the spatial domain, as edges contain high frequency. Areas of constant gray level consist of mainly low frequency and then for they are supressed.



Frequency Domain



Spatial Domain

−1	−1	−1
−1	−1	−1
−1	−1	−1

(mask)

Mask

(c) Explain the Minimum Mean-square Error Restoratin.

Ans. The wiener filter exploits the statistical properties of the image and can be used to resotre image in the presence of blur as well as noise. This method is founded on considering images and noise as random variables and the objective is to find an estimate f_1 of the uncorrupted image f such that the mean square error between them is minimized. This error measure is given by:

$$e^2 = E\{(f - f_1)^2\}$$

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

The minimum of the error function is given in the frequency domain by the expression:

$$\begin{aligned} F(u, v) &= [H^*(u, v)S1(u, v)/S1(u, v)|H(u, v)|^2 + S2(u, v)\ G(u, v)] \\ &= [H^*(u, v)/|H(u, v)|^2 + S2(u, v)/S1(u, v)\ G(u, v)] \\ &= [H(u, v)|^2 + H(u, v)\ (|H(u, v)|^2 + S2(u, v)/S1(u, v))\ G(u, v)] \end{aligned}$$

Where we have used the fact that the product of a complex quantity with its conjugate is equal to the magnitude of the complex quantity squared.

This result is known as the Wiener filter.

The filter which consists of the terms inside the bracket is commonly referred to the minimum mean square error filter or the least square error filter.

The terms used in the eq. are as follows:

$H(u, v)$ = degradation function

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

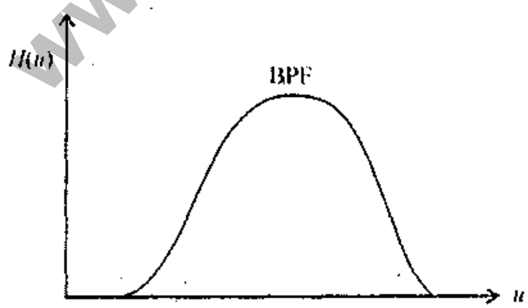
$|H(u, v)|^2 = H^*(u, v)H(u, v)$

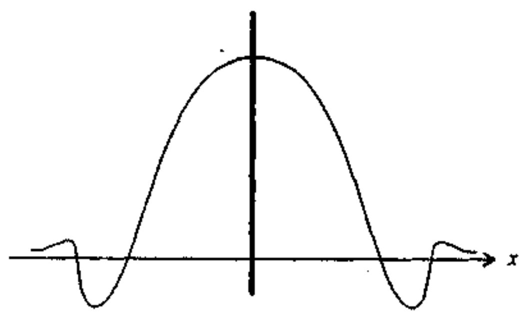
$S2(u, v) = |N(u, v)|^2$ – power spectrum of the noise

$S1(u, v) = |F(u, v)|^2$ = power spectrum of the undegraded image.

(d) Explain the operation of Band-pass filters.

Ans. Band-pass filter: A band-pass filter attenuates very low and very high frequencies, but retains a middle rane band of frequencies. Band-pass filter can be used to enhance edges (suppressing low frequency) while reducing noise at the same time.





–1	–2	–1
–2	–4	–2
–1	–2	–1

(mask)

(e) Explain the basis of filtering in Frequency Domain and Frequency Domain method.

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

$$\begin{array}{ccccc} g(x, y) & & f(x, y) & \otimes & h(x, y) \\ \text{DFT} \downarrow\uparrow \text{IDFT} & & \text{DFT} \downarrow\uparrow \text{IDFT} & & \text{DFT} \downarrow\uparrow \text{IDFT} \\ G(u, v) & & F(u, v) & \bullet & H(u, v) \end{array}$$

The filters in frequency domain are Discrete Fourier transform of convolution mask in spatial domain.

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

For example: Lowpass filtering, it is know that edges and other sharp transitions (such as noise) in gray level of an image is due to the contribution of the high-frequency component is spatial domain or high-frequency content in fourier domain.

Hence by removing or attenuating a specified range of high-frequency content in Fourier domain with use of lowpass filter as:

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

(Low pass filter)

results into smoothing or blurring of image in spatial domain when $G(u, v)$ is inverse transformed into $g(x, y)$

$$G(u, v) \xrightarrow{IDFT} g(x, y) \quad \text{(Output image without high frequency component)}$$

Q.2.(*f*) Discuss the common source of blurring and noise.

Ans. the principle sources of noise in digital images arise during image acquisition (digitization) and/or transmission. The performance of imaging sensors is affected b a variety of factors, such as environmental conditions during image acquisition, and by the quality of the sensing elements themselves.

For instance, in acquiring images with a CCD camera, light levels and sensor temperature are major factors affecting the amount of noise in the resulting image. Images are corrupted during transmission principally due to interference in the channel used for transmission.

For example, an image transmitted using a wireless network might be corrupter as a result of lightning. The important parameters that define the spatial characteristic of noise is correlated with the image. Frequency properties refer to the frequency content of noise in the Fourier sense (i.e., as opposed to the electromagnetic spectrum).

For example, when the Fourier spectrum of noise is constant, the noise is usually called white noise. This terminology is a carry over from the physical properties of white light, which contains nearly all frequencies in the visible specturm in equal proportions. Noise is independent of spatial coordinates, and that is uncorrelated with respect to the image itself (that is, there is no correlation betwen pixel values and the values of noise components).

Noise in a image are of two types:

Correlated noise: These noise are periodic in nature, chances of existence of these types of noises are rare and can be easily filtered out to restore.

Examples are:

- Due to electrical interference.
- Due to source/Sensor interferenee.
- Halftone distortion/Moire pattern.

Uncorrelated Noise: These noise are Random in nature, chances of existence of these types of noises are common and they can't be filtered out to restore compltely.

Examples are:

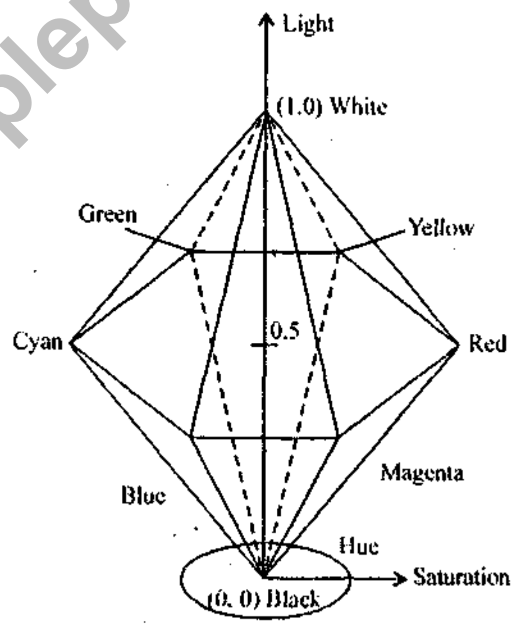
- Quantum noise in CCD arrays.
- Neuronal noise in a *retina*.
- Quantization noise in digital image.

Q.3. (*a*) Explain the HSI (Hue Saturation Intensity) colour model. Discuss Image Smoothing too.

Ans. HSI Model: User (human) oriented color model is based on some of the color sensing properties of human visual system. HSI color model decouples intensity (I) information from color. Hue (H) and Saturation (S) components are related to way in which human (user) perceive color.

HUE	Pure Color (Yellow, Orange ... etc.)
SATURATION	Degree to which ure color is diluted by white light
INTENSITY	Brightness

Hence, Hue represents dominant color as "perceived" by an observer, that dominate wavelength. Saturation refers to the relative purity or amount of white light mixed with a hue (pure color). The pure color's (e.g. yellow) are fully saturated. Pink, for example is less saturated as Red is mixed with some white light.



HSI Color Model

Image Smothing: The filters are used for blurring (removal of small details from an image prior to object extraction, bridging of small gaps in lines or curves) and noise reduction.

(a) Mean Filters: The response is simply the average of the pixels contained in the neighborhood of the filter mask.

(*i*) Arithmetic Mean Filter: The response is simply the arithmetic mean computed using the pixels in the region defined by S_{xy} where S_{xy} represents the set of coordinate in a rectangular sub image window of size *mXn*, centred at point (x, y).

Mathematical formula:

$$g(x, y) = 1/mn \sum_{(s,t) \in S_{xy}} g(s, t)$$

Where $g(x, y)$ is image.

Mask

1/9 X

1	2	1
2	4	2
1	2	1.

***(ii)* Geometric Mean Filter:** The response time is given by the product of the pixels in the subimage raised to the power 1/mn.

Mathematical formula:

$$g(x, y) = \left[\prod_{(s,t) \in sxy} g(s, t)\right]^{1/mn}$$

***(iii)* Weighted average filter:** In this filter, pixels are multiplied by different coefficients weighing the center point the highest and then reducing the value of the coefficients as a funtion of increasing distance form the origin.

Mathematical formula:

$$g(x, y) = \sum_{s=-a}^{a} \sum_{t=-b}^{b} w(s, t)\, f(xts, ytt) / \sum_{s=-a}^{a} \sum_{t=-b}^{b} w(s, t)$$

Mask

1/16 X

1	2	1
2	4	2
1	2	1

Applications of smoothing linear filters:

- Noise reduction (arithmetic and mean filters for Gaussian noise).
- Smoothing of false contours.
- Reduction of irrelevant detail in an image.

Disadvantage of smoothing linear filters:

- Blurring of edges
- Blurring of small image details (less in geometric mean than arithmetic mean).

(b) Ordered Statistics Filters: The response is based on ordering (ranking) the pixels contained in the image area encompassed by the filter and then replacing the value of the center pixel with the value determined by the ranking result.

(i) Median Filter: It replaces the value of a pixel by the median of the gray levels in the neighborhood of the pixel. The function is to force points with distinct gray levels to be more like their neighbors.

Mathematical Formula: The median ξ, of a set of values is such that half the values in the set are less than or equal to ξ, and half are greater than or equal to ξ. In a 3*3 neighborhood the median is the 5th largest value in a 5*5 neighborhood the 15th largest value

$$R = \text{median } \{zk/k - 1, 2, \ldots, 9\}$$

***(ii)* Min Filter:** The 0^{th} percentile filter used for finding the darkest points in an image. It replaces the value of a pixel by the minimum value of the gray levels in the neighborhood of that pixal.

Mathematical Formula:

$$R = \min \{zk/k = 1, 2, \ldots, 9\}$$

(iii) Max filter: The 100th percentile filter used for finding the brightest points in an image. It replaces the value of a pixel by the maximumvalue of the grey levels in the neighborhood of the pixel.

Mathematical Formula:

$$R = \max\{zk/k = 1, 2, ..., 9\}$$

Applications of smoothing non-linear filter:

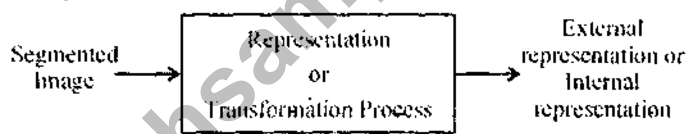
- Reduction of noise with less blurring (Median filter for salt and pepper noise).
- Used for finding brightest ponts in an image (Max filter).
- Used for finding darkest points in an image (Min filter).

Disadvantages of smoothing non-linear filter:

- Repeated passes of median filter tend to blur the image.
- Max filter also removes some dark pixels from the borders of the dark objects.
- Min filter also removes some light (white points) points around the border of light objects.

Q.3.(b) Explain the feature extraction and differentiate it with the segmentation of the image processing.

Ans. Feature Extraction: The feature extraction aspect of image analysis to identify inherent characteristics or features of objects fourn within an image. These characteristics are used to describe the object. The role of representation or transforming an image into another form is to make a form [representation, transform] that is suitable for further computer processing. A raw ditital image cannot be processed directly pixel by pixel as algorithm for such processing will be highly time and space consuming. So standard practice is to use schemes that compact's digital data (digital image) into another representation that are considerably more useful in computation of descriptor for an image.



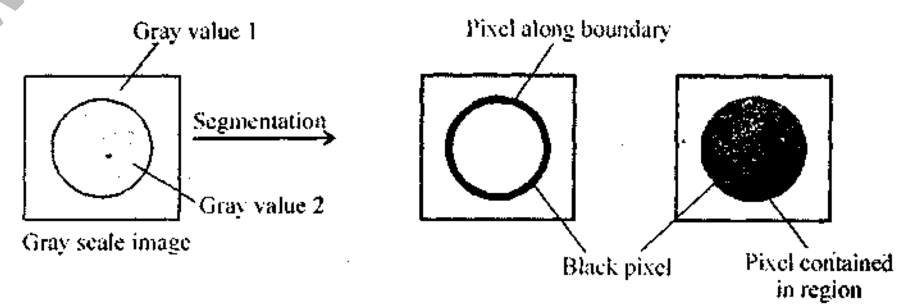
- Pixel along Boundary (External Data).
- Pixel contained in Region (Internal Data).

Representation a region involves two choice:

1. Represent region in terms of its external characteristics [its boundary].
2. Represent region in terms of internal characteristics [pixel comprising the region].

The choice of representation scheme depends on the task, which data is useful [weather boundary pixel or region pixel] for computer processing.

For example:



In above figure, pixel along boundary [external feature] represent external characteristics and boundary can be described by feature such as its [length, orientation, number of concavities etc.] i.e., when representation has to focus on shape. Pixel contained in region [internal feature] is selected when the primary focus is on regional properties such as [color, texture etc.]

Image Segmentation: Segmentaiton is the first step towards producing 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 physically measured image feature, such as histogram and probability distributions 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.

Thresholding selection methods: There all three types of thresholding methods LOCA, 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, $t = \text{I}(x, y; P(x, y))$

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

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

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

We know that the process of image segmentation divides an image into separate regions. Once an image has been segmented, it is often desirable to represent an object using other than the coordinates of the pixel defining the objects. Since edges play an important role in the recognitiion of objects, contour description methods have been developed that completely describes an object.

Q.3.(c) Explain the Morphological image processing in detail.

Ans. Morphological image processing is a type of processing in which the spatial form or structure of objects within an image are modified. Dilation, erosion and skeletonization are three fundamental morphological operations. Wtih dilation, an object grows uniformly in spatial extent, whereas with erosion and object shrinks uniformly. Skeletonization results in a stick figure representation of an object.

A dilated by B, denoted by $A \oplus B$ is defined as:

$A \oplus B = [p \subset D]\ p = a + b$ for some $a \subset A$, $b \in B$

Where B is called the structuring element, Dilatin can also be defined as Union of translate i.e.

or

eroded by B, denoted by $A \ominus B$, is defined as:

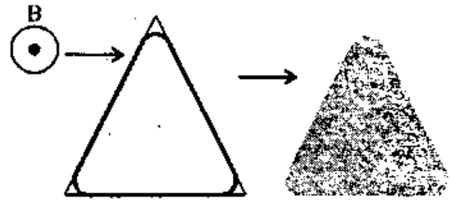
$$A \ominus B = [p - b \in A \text{ and for every } b \in B]$$

Opening: A opened by B, denoted by A o B and is defined as:

$$AoB = (A \ominus B) \oplus B$$

That is, erosion followed by dilation with same structuring element 'B' to set A.

A Trnaslate of B in A.



(a) (b) (c)

Structuring element B "rolling" along inner boundary of A. Heavy line in (b) is the outer boundary of the opening. Shaded portion (c) is complete opening.

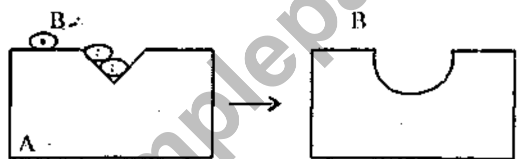
From above example, it is clear that opening suppresses sharp capes and also removes objects smaller than the structuring elements and cuts the narrow isthmuses.

Closing: A closed by B, denoted by $A \bullet B$, is defined as

$$A \bullet B = (A \oplus B) \ominus B$$

That is, dilation followed by erosion with same structuring element 'B' to set A

Example:



(a) (b)

Hit-and-Miss Transform: Suppose there are two structuring elements B and C such that $B \cap C = \varphi$, then hit-and-miss transorm of set A by B and C, denoted $A \otimes (B, C)$, is defined as

$$A \oplus B = (A \ominus B) \cup (A \ominus C)$$

Here, $[B \cap C = \phi]$ implies that B & C set has no element in common or B is forground and C is background. When set A is transformed using (B, C) through "hit-and-miss" we are looking for (B, C) in set A. So, after this transform output will be those points or area where (B, C) exists in set A.

Extensive Operation are those in which output is superset to input.

$$A \bullet B \supseteq A$$

This implies that, closing B extensive oepration and dilation too.

Similarly, anti-extensive operation are those in which output is subset of input

$$A \circ B \subseteq A$$

This implies opening as anti-extensive operation and erosion too.

Q.4.(a) *Explain Geometric transformation of images using Spline Transformation.*

Ans. 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 fro the discrete samples.

When we apply a transformation (such as rotation or zooming) it becomes necessary to know the image intensity at a locatioin in between the sample points. This is an interpolation problem and splines come in handy.

Spline Tensor Product Basis: We extend the ID B-spline basis to 2D using tensor products $\varphi n([x, y])$ defined as

$$\varphi^n(x - k, y - l) = \beta^n(x - k)\, \beta^n(y - l)$$

Using these bases we can interpolate a discrete image $f([k, l])$ and come up with a continuous function $f([x, y])$ over R2.

$$f([x, y]) = \sum_k \left(\sum_l (c([k, l])\beta^n(x - k)\, \beta^n(y - l)) \right)$$

The coefficients $c([k, l])$ are obtained by fitting the discrete image to the 2D spline basis.

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

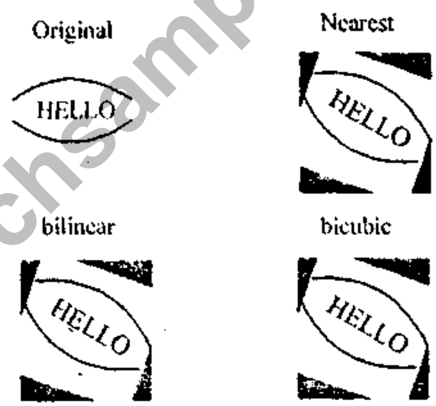


Fig. *Illustration of how the jagged edges appear in nearest neighbour interpolation and they progressively smooth out in higher order spline-interpolation. Bicubic spline interpolation is highly popular as cubic splines appear smooth to the human eye.*

Q.4.(b) Discuss various thresholding algorithms.

Ans. 1. Bi-level Thresholding: Bi-level thresholding is employed on images which have bimodel histograms. In bi-level threholding, the object and background form two different groups with distinct gray levels. In these cases, the shapes of the histograms are bimodal with peaks corresponding to the object and background regions and a valley in between. The valley point is usually chosen as the threshold. In bimodel thresholding, all gray values greater than threshold T are assigned the object label and all other gray values are assigned are background label, thus, separating the object pixels from the background pixels. Thresholding thus is a transformation of an input image A into a segmented output image B as follows:

(a) $b_{ij} = 1$ for $a_{ij} \geq T$.

(b) $b_{,,} = 0$ for at, < T, where T is the threshold

A simple iterative algorithm for theshold selection in a birndal image is presented below:

Step 1. Choose an initial threshold $T \leftarrow T_0$.

Step 2. Partition the image using *T* in two regions background and force-ground (object).

Step 3. Compute mean gravy values μ_1 and μ_2 of background and object regions respectively.

Step 4. Compute the new threshold T ←

$$\frac{\mu_1 + \mu_2}{2};$$

Step 5. Repeat Steps 2 to 4 until there is no change of *T*.

Otsu's Global Thresholding Algorithm: This type of thresholding is global thresholding. It stores the intensities of the pixels in an array. The threshld is calculated by using total mean and variance. Based on this threshold value each pixel is set to either 0 or 1, i.e., background or foreground. Thus, here the change of image takes place only once.

The following formulas are used to calculate the total mean and variance.

The pixels are divided into 2 classes,

C1 with gray levels [1, .., *t*] and

C2 with gray leels [t + 1, ..., L].

The probability distribution for the two classes is

$$C_i : p_1/w_1(t), \ldots p_1/w_1(t)$$

and $$C_i : p_{01}/w_2(t), \ldots p_i/w_i(t)$$

$$\text{Where}w_i(t) = \sum_{i-1}^{t} p_i \text{ and } w_2(t) = \sum_{i-t}^{t} p_i$$

Also, the means for the two classes are:

$$\mu_1 = \sum tp / w_i(t)$$

and $$\mu_2 = \sum tp_2 / w_i(t)$$

Using Discriminant Analysis, Otsu defined the beween-class variance of the thresholded image as

$$\sigma_g^2 = w_2(\mu_i - \mu_t)^2 + w_2(\mu_2 - \mu_t)^2$$

For bi-level thresholding, Otsu verified that the optimal threshold t* is chosen so that the between-class variance B is maximized; that is

$$t^* = \underset{i=1}{Arg \max} \left\{\sigma_g^2(t)\right\}$$

Region Merging and Splitting: A segmentation algorithm can produce too many small regions because of fragmentatioin of a single large region in the scene. In such a situation, the smaller regions need to be merged based on similarity and compactness of the smaller regions. A simple region merging algorithm is presented below.

Step 1. Segment the image into R_1, R_2, ..., R, using a set of thresholds.

Step 2. Create a region adjacency graph (RAG) from the segmented description of the image.

Step 3. For every R_i, i = 1, 2, ..., m, identify all R_j, j # i from the RAG such that R_i is adjacent to R_i.

Step 4. Compute an approporiate similarity measure S, between R, and R,, for all i and j.

Step 5. If S,, > T, then merge R, and R,.

Step 6. Repeat steps 3 to 5 until there is no region to be merged according to the similarity criteria.

Q.4(c) Write short notes on following:

(i) Edge Relaxation

(ii) Border Training

Ans. (i) Edge Relaxation: A process similar to thresholding with hysteresis is edge relaxation. The idea of edge relaxation is not simply and pixels if they are next to other edge pixels but to consider the context as well.

Let's consider this question of whether or not a pixel between two sets of edge pixels is itself an edge pixel. One way of determining this is to look at the magnitude of the intervening pixel: if it is relatively high, but less than the threshold used to determine that its neighbors are edge pixels, it's probably an edge. Of course, we can also check the similarity of the gradient magnitude and gradient orientation, just like we did with edge linking. We can also use this not just to interpolate between edge pixels, but to extrapolate from them as well. Suppose that we have two adjacent edge pixels followed my a slightly sub-threshold one (with similar gradient magniude and orientation).

Again, it's likely that this is really an edge pixel. We can add these possible pixels to the set of edge pixels and repeat the process. Supposing that these are now really edge pixels, there may be othe near-misses that we might want to allow as edge pixels. In a sense, we are successively relaxing the criteria used to determine edge pixels, taking into account not just the properties of the pixel in question but of its neighbors as well. This process is called edge relaxation.

In general, the term relaxation applies to any technique such as this that iteratively reevaluates pixel classification.

(ii) Border Tracing: Tracing the border of a surface patch deals with arranging the border locations in such an order that a closed line which doesnot intersect is obtained, when the ordered border ponts are connected in the x_1, y_1, z_1 coordinate system.

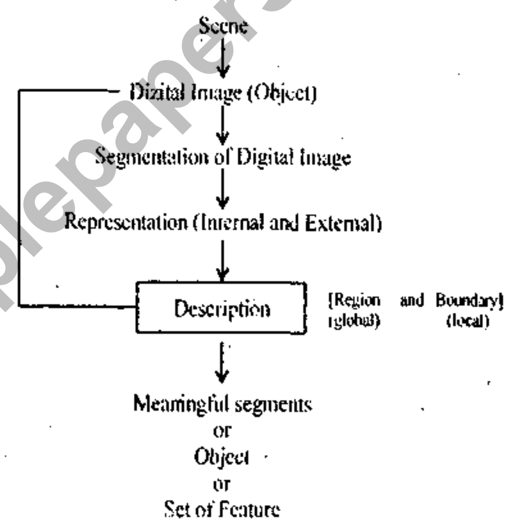
The case of a merged planar patch is considered in. If the neighboring patches are also planar, the border of the patch is a poly-line consisting of vertices and non-occuluded or occluded edges. The corresponding vertices and edges between the original patches that have been merged are first determined by matching vertex angles, edge lengths, and edge labels. The border of the merged patch is then traced out by followng both borders around. "When neither edge is occluded then both edges are followed, when one edge is occluded the other is followed, and when both edges are occluded the outermost is followed."

The border locations are first determined independently on each map. Then the locations given by different maps are compared and the locations are removed which are actually inside the patch when all the views are considered. In case the same piece of the border has been measured from several viewpoints, only single border locations are stored for the tracing. The locations left are then traced on multiple domains by switching between the domains whenever the border locations end on a domain and continue on another one.

After this, we have improved the definition of a border location so that there has to be at least one truly inside location also in the neighborhood of the border location. This makes the border smoother so that the tracing does not get stuck into thin ledges having the width of a single location. It is also checked tha the new piece of line does ot intersect the line traced so far. We may also take a couple of steps backwards during the tracing and select another location than previously.

Q.5. (a) Explain flow diagram of image analysis and understanding methods.

Ans.



A raw digital image cannot be processed directly pixel by pixel as algorithm for such processing will be highly time and space consuming. So, standard practice is to use schemes that compact's digital data (digital image) into another representation that are considerably more useful in computation of descriptor for an image.

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

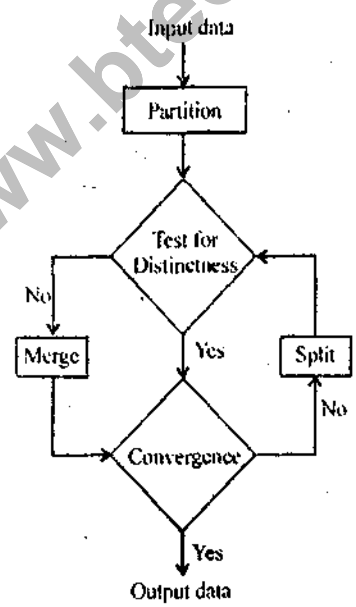
Representation $\xrightarrow{\text{Converted to}}$ Description

In the next step, objects are assigned to the classes based on the invariant features associated with them, so object of the same class must process similar feature and those which belong to different class posses different feature. Therefore, the set of features that distinguishes objects of different classes and is comon to objects of the same classes is the key for classificatioin and recognition.

Q.5.(b) What's clustering? Discuss types of 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 merge. This procedure is repeated on all of the subsets until no further splitting of subsets occurs or until some stopping criteria is met. The basic objective of clustering technique is to divide the data points of the feature space into a number of groups (or classes) so that a predefined set of criteria are satisfied. Criteria usually include interclass and intra-class distance, density points within class, etc.

General clustering algorithm:



Types of Clustering: Data clustering algorithms can be hierarchical. Hierarchical lgorithms find successive clusters using previously established clusters. These algorithms can be either agglomerative ("bottom-up") or divisive ("top-down").

Two-way clustering, co-clustering or biclustering are clustering methods where not only the objects are clustered but also the features of the objects, i.e., if the data is represented in a data matrix, the rows and columns are clustered simultaneously.

Density-based clustering algorithms are devised to discover arbitrary-shaped clusters. In this approach, a cluster is regarded as a region in which the density of data objects eceeds a threshold. DBSCAN and OPTICS are two typical algorithms of this kind.

Q.5.(c) Explain feature extraction and feature detectors.

Ans. Feature Extraction: Feature extraction involves simplifying the amount of resources required to describe a large set of data accurately. When performing analysis of complex data one of the major problems stems from the number of vaiables involved. Analysis with a large number of variables generally requires a large amount of memory and computation power or a classification algorithm which overfits the training sample and generalizes poorly to new samples. Feature extractioin is a general term for methods of constructing combinations of the variables to get around these problems while still describing the data with sufficient accuracy.

Best results are achieved when an expert constructs a set of application-dependent features. Nevertheless, if no such expert knowledge is available general dimensionality reduction techniques may help.

Feature Detectors: Feature detectors are used to extract various features out of an image. Various features extracted are:

1. Edges
2. Corner/Interest Points
3. Blobs/Regions of interest
4. Ridges

Feature detector	Edge	Corner	Blob
Canny	X		
Sobel	X		
Harris & Stephens/Plessey	X		X
SUSAN	X		X
Shi & Tomasi	X		
Level curve curvature	X		
FAST	X		
Laplacian of Gaussian	X		X
Difference of Gaussians	X		X
Determinant of Hessian	X		X
MSER	X		
PCBR			X
Grey-level blobs			X

Q.5.(d) Discuss geometric invariants.

Ans. The features that are invariant to geometric structure of that feature representation, (i.e., Rotation, translation etc.) are called geometric invariants.

For example, to make chain code geometrically invariant, we may get the first difference of the cod and then may choose the one that gives the smallest number on rotating the sequence.

Example: If 4-directional chain code is given by 10103322 then difference is 3133030.

Circular chain code: 33133030

Rotation of circular chain code: 03033133

Normalized chain codes are exact only if the boundaries are invariant to rotation and scale change.

Q.5.(e) Discuss brief object recognition.

Ans. World object are considered to be made up of patterns. A pattern is essentially an arrangement. It is characterized by the order of the elements of which it is made. A pattern is an arrangement of descriptors. The name feature is used often to denote a description. A pattern class is a family of patterns that share some common properties. Recognition or pattern recognition is typical characteristic of human being. The pattern means something that is set up as an ideal to be imitiated. For example, Shape 'A' is shown to us and we try to imitate that. When second time similar shape 'A' is shown the human being identifies it as 'A' again.

This identification step is called recognition and the shape we imitate is termed as patterns. Thus, the pattern recognition means identification of the ideal object, Similarly, pattern recognition by machine involves techniques for assigning patterns to their respective classes automatically and with or little human intervention as possible. That is, objects are assigned to the classes based on the invariant features associated with them, so object of the same class must possess similar feature and those which belong to different class posses different feature. Therefore, the set of features that distinguishes objects of different classes and is common to objects of the same classes is the key for classification and recognition.

Q.5.(f) Discuss multi-level feature processing.

Ans. In order to enable multilevel image representation, it has been proposed that a dot area, which has a fixed and is composed fo plurality of dots, be defined as a unit area for the tone production process, and the ont representation be made by changing the number of recording dots and the number of non-recording dots contained in the unit area. Examples of this type of image representatiion are a tone production method employing a density pattern, and a dither method.

However, the conventional multilevel image representation has a disadvantage in that as the number of tone levels increases, the resolution is deteriorated. An increase in the number of tone levels is suitable for images of photographs, but, on the other hand, it is unsuitable for character images and line images when it is desired to reproduce images with a high resolution. Therefore, it is preferable to represent character and line images by using a bi-level signal.

Accordingly, it is a general object of the present invention to provide an improved image processing method in which the above described problems are eliminated.

Another, more specific object of multi-level feature processing is to provide a digital image processing method capable of producing a high-quality halftone image having a high gradation and high resolution in whichone dot is represented by a plurality of tone levels.