IMAGE PROCESSING

NOTE:

MAKAUT course structure and syllabus of 6^{th} semester has been changed from 2021. **IMAGE PROCESSING** has been introduced as a new subject in present curriculum. The syllabus of this subject is almost same as **Digital Image Processing [EC 802B]**. Taking special care of this matter we are providing the relevant MAKAUT university solutions of **Digital Image Processing [EC 802B]**, so that students can get an idea about university questions patterns.

INTRODUCTION

Multiple Choice Type Questions

1. An image is a 2D array of [WBUT 2009, 2011, 2017]
   a) digital data
   b) electrical signals
   c) photographic objects
   d) light signals

Answer: (a)

2. A line sensor is used to [WBUT 2009]
   a) capture a scene
   b) capture a 3D image
   c) scan a 2D image
   d) none of these are true

Answer: (d)

3. What device is used to form an image on the film of a camera? [WBUT 2009]
   a) A p-n-p transistor
   b) A converging lens
   c) An Op-Amp
   d) A plane mirror

Answer: (a)

4. If an input image is $f(x, y)$ and a transform T is operated to get an processed image $g(x, y)$, we can write [WBUT 2009]
   a) $f(x, y) = T[g(x, y)]$
   b) $f(x, y) = T / g(x, y)$
   c) $g(x, y) = T[f(x, y)]$
   d) none of these are true

Answer: (c)

5. Euclidian distance of two points (x, y) and (s, t) of a two-dimensional space is [WBUT 2010]
   a) $\left[(x-s)^2 + (y-t)^2\right]^{\frac{1}{2}}$
   b) $|x-s| + |y-t|$
   c) Max$(|x-s|, |y-t|)$
   d) none of these

Answer: (b)

6. Coloured Model Names [WBUT 2011, 2014]
   a) The RGB colour model
   b) The CMY & CMYK colour models
   c) The HIS colour model
   d) (a) & (b) only
   e) (a), (c) & (d)

Answer: (e)

7. Digital image processing uses [WBUT 2011]
   a) Fuzzy set theory
   b) DFT
   c) DCT
   d) (b) & (c)
   e) (a), (b) & (c)

Answer: (d)

8. Find the odd one out w.r.t. DIP: [WBUT 2011]
   a) Arithmetic operation
   b) Softwares
   c) Vector & matrix operations
   d) Image transforms

Answer: (c)

9. Intensity range of 8-bit pixel image is [WBUT 2012]
   a) 0 to 7
   b) 0 to 15
   c) 0 to 31
   d) 0 to 255

Answer: (d)

10. A digital image is composed of a finite number of elements, each of which has a particular location and value. These elements are called [WBUT 2012]
   a) dot
   b) pixel
   c) point
   d) none of these

Answer: (b)

11. The total amount of energy that flows from the light source and it is usually measured in watts (W) is called [WBUT 2012, 2018]
   a) Radiance
   b) Luminance
   c) Reflectance
   d) None of these

Answer: (b)

12. Sampling of an image required for [WBUT 2012, 2019]
   a) quantization
   b) sharpening
   c) smoothing
   d) digitization

Answer: (d)

13. The common major of transmission of digital data is [WBUT 2013]
   a) bit rate
   b) baud rate
   c) frame per second
   d) none of these

Answer: (a)

14. HDTV stands for [WBUT 2013]
   a) High Definition Television
   b) High Level Digital Television
   c) (a) & (b) both
   d) None of these

Answer: (a)

15. Digital Image Processing deals with [WBUT 2013]
   a) analog signal
   b) digital signal
   c) discrete signal
   d) (b) & (c) both

Answer: (a)

16. Colour image processing is gaining importance because [WBUT 2014]
   a) It's more pleasant to watch
   b) It's cost effective
   c) It's easy to capture and represent
   d) Use of digital image over the internet has increased significantly

Answer: (d)

17. Which of the following is not an image file format? [WBUT 2015, 2017]
a) TIFF b) BMP c) GIF d) none of these
Answer: (d)

18. The smallest units of a digital image is represented by [WBUT 2015]
a) A one dimensional matrix b) Logic 0 or 1
c) Dot d) Pixel
Answer: (d)

19. Which of the following statement is true? [WBUT 2015]
a) The resolution of CMOS image sensor is better than CCD image
b) The resolution of CCD image sensor is better than CMOS image
c) CCD image sensor is cost effective than CMOS image
d) none of these
Answer: (b)

20. An example of volume image is [WBUT 2015]
a) A one dimensional image is b) A two dimensional image
c) A three dimensional image d) all of these
Answer: (c)

21. The spectrum of the visible light is [WBUT 2016]
a) 10-350 nm b) 380-760 nm
c) 760 nm and above d) none of these
Answer: (b)

22. The image acquisition process is used to [WBUT 2016]
a) store image b) transform image
c) display image d) none of these
Answer: (d)

23. Colour image is represented by [WBUT 2016]
a) 2-bit b) 8-bit c) 24-bit d) 64-bit
Answer: (c)

24. The photosensitive detector of the human eye is [WBUT 2016]
a) eye lens b) iris c) retina d) cornea
Answer: (c)

25. Through decimated by 2 operations the sampling rate is [WBUT 2016]
a) increased b) decreased c) none of these d) same
Answer: (b)

26. What is the range of subjective brightness for human? [WBUT 2017]
a) scotopic threshold to glare limit b) photopic threshold to glare limit
c) scotopic threshold to infinity d) photopic threshold to infinity
Answer: (b)

27. The image function $f(x,y)$ is characterized by two components: [WBUT 2017]
$f(x,y)=i(x,y).r(x,y)$ where
a) $0<i(x,y)<1$ & $0<r(x,y)<\infty$
b) $0<i(x,y)<1$ & $0<r(x,y)<1$
c) $0<i(x,y)<\infty$ & $0<r(x,y)<\infty$
d) $0<i(x,y)<\infty$ & $0<r(x,y)<1$

Answer: (d)

28. Each element of image matrix is called [WBUT 2019]
a) dots
b) coordinates
c) pixel
d) value

Answer: (c)

29. Intensity levels in 8-bit image are [WBUT 2019]
a) 128
b) 255
c) 256
d) 512

Answer: (c)

30. What would the time be at 3000K baud, a representative medium speed of a phone DSL (Digital Subscriber Line) connection? [WBUT 2019]
a) 3 sec
b) 3 mins
c) 3.5 sec
d) 3.5 mins

Answer: (c)

31. The process of extracting information from the image is called as [WBUT 2019]
a) image enhancement
b) image restoration
c) image analysis
d) image compression

Answer: (c)

32. CAT in imaging stands for [WBUT 2019]
a) Computer Aided Telegraphy
b) Computer Aided Topography
c) Computerized Axial Telegraphy
d) Computerized Axial Tomography

Answer: (d)

33. A pixel p at coordinate (x,y) has four horizontal and vertical neighbors whose coordinates are given by [WBUT 2019]
a) $(x-1,\ y-1),\ (x+1,\ y+1),\ (x+1,\ y),\ (x,\ y+1)$
b) $(x+1,\ y),\ (x,\ y+1),\ (x-1,\ y),\ (x,\ y-1)$
c) $(x+1,\ y+1),\ (x+1,\ y),\ (x,\ y+1),\ (x,\ y)$
d) $(x,\ y-1),\ (x-1,\ y)(x-1,\ y),\ (x-1,\ y-1)$

Answer: (b)

Short Answer Type Questions

1. Develop a procedure for computing the median of an $n\times n$ neighbourhood. Propose a technique for updating the median as the centre of neighbourhood if moved from pixel to pixel. [WBUT 2010]

Answer:
1st Part:
Numerically sort the n^2 values.
The median is $\xi = \left[\left(n^2+1\right)/2\right]^{th}$ largest value.

2nd Part:
Once the values have been sorted one time, we simply delete the values in the trailing edge of the neighbourhood and insert the values in the leading edge in the appropriate location in the sorted array.

2. What is pixel? Explain 4-neighbour and 8-neighbour of a pixel. Explain m-adjacency. **[WBUT 2010]**

Answer:
1st Part:
Picture element in short is called pixel. It is the smallest square element of an image that can be turned on or off in a monitor. Resolution of a monitor depends on the no. of pixels i.e., VGA monitor display 640×480 pixels / inch.

2nd Part:
A pixel p at coordinates (x, y) has four horizontal and vertical neighbours whose coordinate are given by $(x+1, y)$, $(x-1, y)$, $(x, y+1)$, $(x, y-1)$. This set of pixel called the 4-neighbours of p, is denoted by $N_4(p)$. Each pixel is at a distance (unit) from (x, y) and some of the neighbours of p lie outside the digital image if (x, y) is on the border of the image. The four diagonals neighbours of p have coordinates $(x+1, y+1)$, $(x+1, y-1)$, $(x-1, y+1)$, $(x-1, y-1)$ and are denoted by $N_D(p)$. These points together with the 4 neighbours are called 8-neighboure of p and are denoted by $N_8(p)$.

3rd Part:
Two pixels p and q with values from V are m-adjacent if q is in $N_4(p)$ or q is in $N_D(p)$ and the set $N_4(p) \cap N_4(q)$ has no pixels whose values are from V.

3. Explain image sensing and acquisition (using single sensor, sensor strip and sensor arrays). **[WBUT 2010]**

Answer:
Image sensing:
An image sensor is a device that converts an optical image to an electric signal. It is mostly used in digital cameras and other imaging devices.

Image acquisition is the creation of digital images, such as physical scene or of the interior structure of an object. The term is often assumed to imply or include the processing, compression, storage, printing and display of such images.

Image acquisition using a single sensor:
The most common sensor of this type is photodiode, which is constructed of silicon materials and whose output voltage waveform is proportional to light. In order to generate a 2D image using a single sensor, there have to be relative displacements in both the X- and Y-directions between the sensor and the area to be imaged. The single sensor is mounted on a lead screw that provides motion in the perpendicular direction.

Image acquisition using a sensor strips:
A geometry that is used much more frequently than single sensors consists of an in line arrangement of sensors in the form of a sensor strip. The strip provides imaging elements in one direction. Motion perpendicular to the strip provides imaging in the other direction. This type of arrangements used in **flat bed scanners**. Sensor strips mounted in a ring configuration are used in medical and industrial imaging to obtain cross-sectional (slice) images of 3D objects.

Image acquisition using sensor arrays:
Individual sensors can be arranged in the form of a 2d array. Numerous electromagnetic and some ultrasonic sensing devices are arranged frequently in an array format. This is also the predominant arrangement found in digital cameras. A typical sensor for these cameras is a CCD array, which can be manufactured with a broad range of sensing properties. CCD sensors are widely used in digital cameras and other light sensing instruments.

4. Explain Bilinear interpolation method. **[WBUT 2011, 2014]**

OR,

Explain bilinear interpolation method used for image zooming. **[WBUT 2015]**

Answer:
In Bi-linear interpolation method, we use the four nearest neighbours to estimate the intensity at a given location. Let (x, y) denotes the coordinates of the location to which we want to assign an intensity value and let $v(x, y)$ denote that intensity value. For bi-linear interpolation, the assigned value is obtained using the equation $v(x, y) = ax + by + cxy + d$ where the four co-efficients are determined from the four equations in four unknowns that can be written using the four nearest neighbours of a point. The intensity value assigned to point (x, y) is obtained using the equation

$$v(x, y) = \sum_{i=0}^{3} \cdot \sum_{j=0}^{3} a_{ij} x^i y^j$$

where the sixteen co-efficients are determined from the sixteen equations in sixteen unknowns that can be written using the sixteen nearest neighbours of point (x, y) which is **bi-cubic interpolation.** If the limits of both summation are 0 to 1, then bi-cubic transforms to bi-linear.

5. What is the resolution of an image? Compute the size of a 640 × 480 image at 240 pixels per inch. **[WBUT 2012]**

Answer:

1st Part:

The maximum number of points that can be displayed without overlap on CRT is its resolution.

2nd Part:

$$\text{Size of Image} = \frac{640}{240} \text{ by } \frac{480}{240} = \frac{8}{3} \text{ by 2 inches.}$$

6. What do you mean by Digitization? Explain its two important steps. **[WBUT 2013]**

Answer:

1st Part:

The method of converting an image which is continuous in space as well as in its value, into a discrete numerical form is called digitization.

2nd Part:

The two important steps of digitization are: sampling and quantization.

Taking measurements at regular space intervals, called sampling.

Mapping the measured intensity or value to one of finite number of discrete levels, called quantization.

7. Write down the key stages in Digital Image Processing & explain them. **[WBUT 2013, 2017]**

OR,

With neat sketch explain briefly about the components of image processing system. **[WBUT 2016]**

Answer:

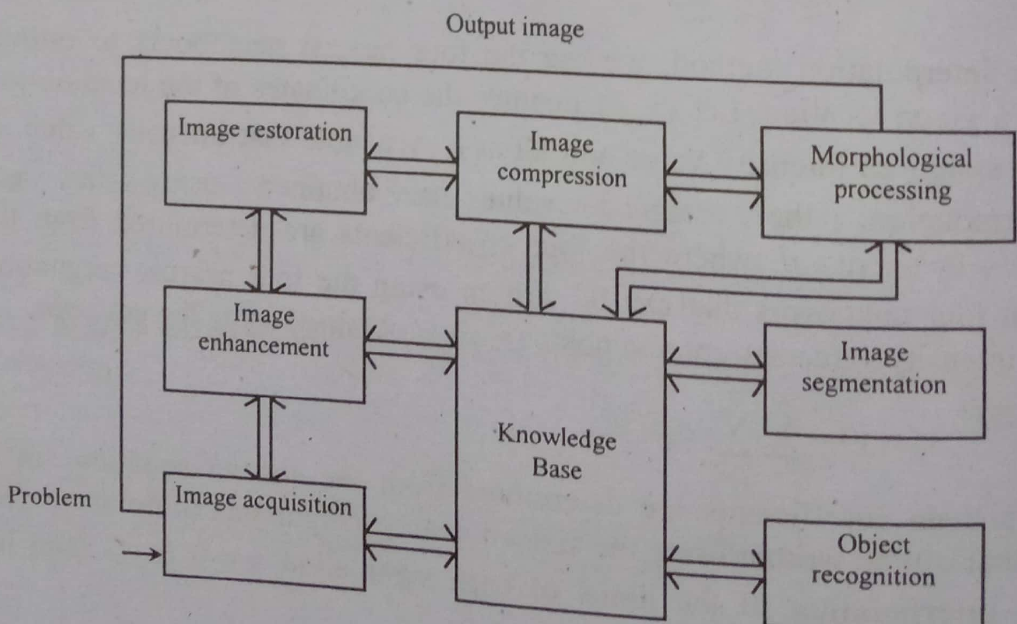
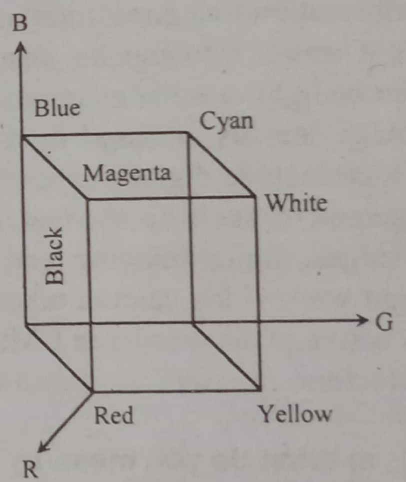


Image acquisition is the first step of digital image. This steps involves preprocessing such as scaling. Here sampling and quantization require to convert the continuous sensed data into a digital form.
Image enhancement technique convert the image to a form better suited for analysis by a human or machine. Common enhancement techniques are point operation and spatial operation, which includes contrast stretching, filtering, noise clipping etc.
Image restoration improves the quality of image acquired by optical means.
Image segmentation defined by a set of regions that are connected and non overlapping, so that each pixel in a segment in the image acquires a unique region label that indicates the region it belongs to.
Image compression helps to reduce the redundancies in raw image data in order to reduce the storage and communication bandwidth.

8. What is 8 bit colour image? For what purpose could it be used? Explain.
[WBUT 2013, 2016, 2018]

Answer:
In the RGB model, each color appears in its primary spectral components of red, green and blue. This model is based on Cartesian coordinate system. The different colors in this model are points on or inside the cube and are defined by vectors extending from the origin. For convenience, the assumption is that all color values have been normalized so that the cube is the unit cube i.e., all values of R, G and B are assumed to be in the range [0, 1]. Images represented in the RGB color model consist of three component images, one for each primary color. When fed into an RGB monitor, these three images combine on the screen to produce a composite color image. The no. of bits used to represent each pixel in RGB space is called pixel depth. Consider an RGB image in which each of the red, green and blue images is an 8-bit image. Under these conditions each RGB color pixel is said to have a depth of 24 bits (3 image planes times the number of bits per plane). The term full color image is used often to denote a 24 bit RGB color image. The total number of colors in a 24-bot RGB image is



$$\left(2^8\right)^3 = 16,777,216$$

Many monitor display use color maps with 8-bit index numbers, meaning that they can only display 256 different colors at any one time. Thus it is often wasteful to store more than 256 different colors in an image anyway, since it will not be possible to display them all on screen.
Because of this many image formats use 8-bit color maps to restrict the maximum number of different colors to 256. Using this method, it is only necessary to store 8-bit index into the color map for each pixel, rather than the full 24-bit color value.

9. What is Weber Ratio? Show the variation of Weber ratio for humans. **[WBUT 2014, 2017]**

Answer:
The ratio of the increment of illumination to background of illumination is called as weber ratio i.e. $\frac{\Delta_i}{i}$. If the ratio $\frac{\Delta_i}{i}$ is small, then small percentage of change in intensity is needed, i.e. good brightness adaptation. If the ratio $\frac{\Delta_i}{i}$ is large, then large percentage of change in intensity is needed i.e. poor brightness adaptation.

10. a) Differentiate between image and scene.
b) What are the image sensors and how are they used? **[WBUT 2016]**

Answer:
a) The Process of receiving and analyzing visual information by the human species is referred to as scene. Similarly the process of receiving and analyzing visual information or scene by digital computer is called image processing i.e. When scene is fed to the computer it is called image, because the computer stores and processes numerical image of a scene.

b) An image sensor or imaging sensor is a sensor that detects and conveys the information that constitutes an image. It does so by converting the variable attenuation of light waves into signals, small burst of current that convey the information. The waves can be light or other electromagnetic radiation.
Image sensors are used in electronic imaging devices of both analog and digital types, which include digital cameras, camera modules, medical imaging equipment, night vision equipment such as thermal imaging devices, radar, sonar and others. As technology changes, digital imaging tends to replace analog imaging. Early analog sensor for visible light were video camera tubes. Currently, used types are CCD (Charge coupled devices) or active pixel sensors in CMOS or NMOS technologies. Digital sensor include flat panel detectors.

11. a) What do you mean by aliasing in the context of image sampling? Explain.
b) What do you mean by the term 'image file format'? Mention some of the frequently used image file format. **[WBUT 2016]**

Answer:
a) Aliasing is a process in which high frequency components of a continuous function "masquerade" as lower frequencies in the sampled function. This is consistent with the common use of the term 'alias', which means "a false identity". Unfortunately, except for some special cases, aliasing is always present in sampled signals because, even if the original sampled function is band limited, infinite frequency components are introduced the moments limit the duration of the function. We conclude that aliasing is an inevitable fact of working with sampled records of finite length. In practice, the effects of aliasing can be reduced by smoothing the input function to attenuate its higher frequencies. This

process, called anti-aliasing, has to be done before the function is sampled because aliasing is a sampling issue that cannot be undone after the fact using computational techniques.

b) Images may be stored in a variety of file formats. Each file format is characteristics by a specific compression type and color depth. The choice of file formats would depend on the final image quality requires (e.g. some file formats support only 256 colors) and the import capabilities of the authoring system. The most popular file formats are: BMP (Bitmap), JPEG (Joint photographic expert group), TIFF (Tagged Image file format), GIF (Graphics Interchange format) etc. BMP supports RGB, Grayscale and Bitmap color modes. JPEG is commonly used to display photographs and other continuous tone images in HTML documents over the WWW and other on line services. TIFF is used to exchange files between applications and computer platforms. It is a flexible bitmap image format supported by virtually all paint, image editing and page layout applications.

12. a) Discuss one technique of image acquisition. **[WBUT 2017]**
b) What is meant by classification of image?

Answer:

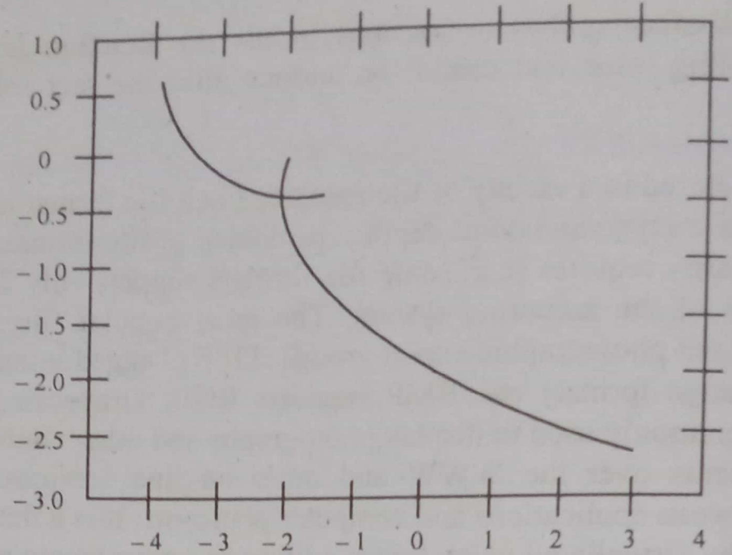
a) ***Refer to Question No. 3(2nd Part) of Short Answer Type Questions.***

b) A broad group of digital image processing techniques is directed towards image classification, the automated grouping of all or selected land cover features into summary categories. Classification approaches can be implemented to classify the total scene content into a limited number of major classes. Classification approaches can also be implemented to distinguish one or more specific classes of terrain (such as water bodies, paved surfaces, irrigated agriculture, forest cutting or other type of disturbances) within the landscape. The result of such classification can be used to spatially direct the efforts of subsequent digital operation or detailed visual interpretation.

13. Brightness discrimination is poor at low levels of illumination. Explain. **[WBUT 2018]**

Answer:

Brightness discrimination is poor (the weber ratio is large) at low levels of illumination and it improves significantly (the weber ratio decreases) as background illumination increases. At low levels of illumination vision is carried out by rods, whereas at high levels (showing better discrimination) vision is the function of the cones. This is shown below.



14. Define digital image and explain image and explain image pixel. What are the storage requirements for (500×500) & (1024×1024) binary images? [WBUT 2019]

Answer:

1st Part:

An image is said to be a digital image if it is a computer readable format so that it can be stored in a computer for further processing.

2nd & 3rd Part:

An image $f(x, y)$ is said to be digital image, if its spatial coordinate (x, y) and the amplitude values are all finite and discrete quantities.

A digital image is comprised of a finite number of elements called pixel or image pixel. Each image pixel has a specific location and value. Digital images require so much storage and computational power that progress in the field of digital image processing has been dependent on the development of digital computer and of supporting technologies that include data storage, display and transmission.

4th Part:

Storage requirements for $500 \times 500 = 500 \times 500 \times 2 = 500000$

and storage requirements for $1024 \times 1024 = 1024 \times 1024 \times 2 = 1024 \times 2048 = 2097152$

Long Answer Type Questions

1. Write short notes on the following:
a) Colour image processing **[WBUT 2011]**
b) Brightness adaptation **[WBUT 2014, 2017]**
c) Weber ratio **[WBUT 2018]**

Answer:

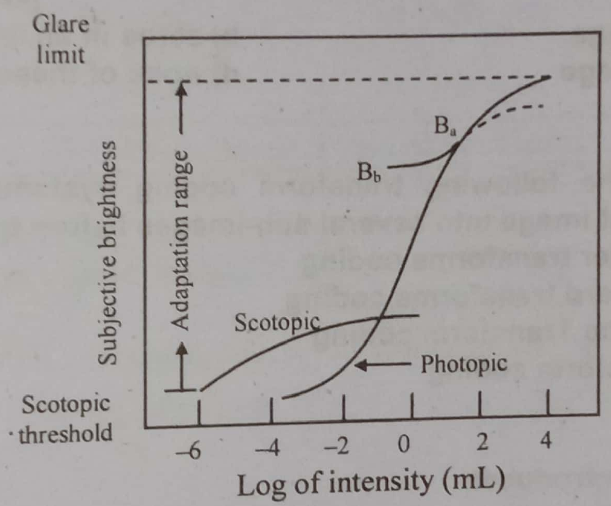
a) Colour image processing:

The use of colour in image processing is motivated by two principal factors. First, colour is a powerful descriptor that often simplifies object identification and extraction from a scene. Second, humans can discern thousands of colour shades and intensities, compared to about only two dozen shades of gray. This second factor is particularly important in manual image analysis.

Colour image processing is divided into two major areas: full colour and pseudo colour processing. In the first category, the images in question typically are acquired with a full-colour sensor, such as a colour TV camera. In the second category, the problem is one of assigning a colour to a particular monochrome intensity. Until relatively recently, most digital colour image processing was done at the pseudo colour level.

b) Brightness adaptation:

The range of light intensity levels to which the human visual system can adopt is enormous, on the order of 10^{10} from the scotopic threshold (minimum low light condition) to the glare limit (maximum light level condition). There is considerable experimental evidence indicating that subjective brightness, brightness as perceived by the human visual system, is a logarithmic function of the light intensity incident on the eye. This characteristics is illustrated below:



It is a plot of light intensity Vs. subjective brightness. The log solid curve represents the range of intensities to which the visual system can adapt. In photopic vision alone, the range is about 10^6. The transition from scotopic to photopic vision is gradual over the approximate range from 0.001 to 0.1 millilambert (–3 to –1 mL in the long scale), as illustrated by the double branches of the adaptation curve in this range. The visual system does not operate over such a wide dynamic range simultaneously. Rather it accomplishes this large variation by changes in its overall sensitivity, a phenomenon known as **Brightness adaptation.** The total range of intensity levels it can discriminate simultaneously is rather small compared with the total adaptation range.

c) Weber ratio: ***Refer to Question No. 9 of Short Answer Type Questions.***

DIGITAL IMAGE FORMATION

Multiple Choice Type Questions

1. If a function $f(x, y)$ is real and we have $F(u, v) = 2DFFT[f(x, y)]$. [WBUT 2009, 2011]
   a) $F(u, v)$ contains only real parts
   b) $F(u, v)$ contains only imaginary parts
   c) $F(u, v)$ contains both real and imaginary parts
   d) none of these are true

Answer: (c)

2. The classical Hough transform is concerned with the identification of [WBUT 2009, 2016, 2018]
   a) lines in an image
   b) zeros in an image
   c) poles in an image
   d) none of these are true

Answer: (a)

3. Which one of the following transform coding systems (usually) does not decompose the input image into several sub-images before transform? [WBUT 2010]
   a) Discrete Fourier transforms coding
   b) Walse-Hadamard transforms coding
   c) Discrete Cosine Transform coding
   d) Wavelet Transform coding

Answer: (d)

4. Faulty switching introduces [WBUT 2010]
   a) Gaussian noise
   b) Rayleigh noise
   c) Gamma noise
   d) Impulse noise

Answer: (d)

5. Poor illumination introduces [WBUT 2010]
   a) Gaussian noise
   b) Rayleigh noise
   c) Exponential noise
   d) Impulse noise

Answer: (a)

6. In the intensity distribution scale the background will of course be [WBUT 2014]
   a) Lower intensity value
   b) Higher intensity Value
   c) Medium intensity value
   d) None of these

Answer: (a)

7. The computation of Walsh coefficient involves [WBUT 2015, 2017]
   a) only subtraction
   b) only addition
   c) addition and subtraction
   d) none of these

Answer: (c)

8. The transform which is widely used in JPEG compression scheme is [WBUT 2015, 2017]
a) FFT
b) IDFT
c) Hadamard Transform
d) Discrete Cosine Transform

Answer: (d)

9. Which of the following uses 2 × 2 mask for edge detection? [WBUT 2017]
a) Sobel
b) Roberts
c) Prewitts
d) Kirsh

Answer: (b)

10. Which of the following transforms is used for line detection in image processing? [WBUT 2017]
a) hadamard
b) haugh
c) haar
d) slant

Answer: (b)

11. Which of the following image processing transforms does not satisfy the separability property? [WBUT 2017]
a) Walsh
b) Fourier
c) DCT
d) Hotelling

Answer: (d)

Short Answer Type Questions

1. a) What is image Sampling? [WBUT 2009, 2013]
b) Define saturation in digital image. [WBUT 2009]

Answer:

a) Digitization of spatial coordinates (x, y) is called image sampling. It is suitable for computer processing. An image function $f(x, y)$ must be digitized both spatially and in magnitude.

b) The saturation is determined by the excitation purity and depends on the amount of white light mixed with the hue. A pure hue is fully saturated, i.e., no white light mixed in. Hue and saturation together determine the chromaticity for a given colour. Hue and saturation tool controls colour. Changing the hue alters the balance of the colour. Changing the saturation alters the strength of the colour.

2. Write down the various 2D transforms. [WBUT 2009]

Answer:

The various 2D transform are:
Translation, scaling, rotation and shearing, reflection etc.
Translation: A translation moves an object to a different position on the screen.
We can translate a point by adding translation coordinate (t_x, t_y) to the original coordinate (x, y) to get the new coordinate (X', Y'). i.e., $X' = x + t_x$; $Y' = y + t_y$.

Scaling: Scaling change the size of the object. Let us assume that the original coordinates are (x, y) and the scaling factors are S_x and S_y in x and y direction. the new coordinate will be $X' = x \cdot S_x$ and $Y' = y \cdot S_y$

i.e., $$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \begin{bmatrix} S_x & 0 \\ 0 & S_y \end{bmatrix}$$

$$P' = P \cdot S$$

where S is the scaling matrix i.e., $\begin{bmatrix} S_x & 0 \\ 0 & S_y \end{bmatrix}$ and P' is the new coordinate (x', y') matrix.

Rotation: In rotation, we rotate the object at particular angle θ from its origin. The new coordinate position will be $\begin{bmatrix} X' & Y' \end{bmatrix} = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$

i.e., $P' = P \cdot R$

where R is the rotation matrix

i.e., $$R = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

P' is the new coordinate and P is the old coordinate of the point.

3. What are the basic steps involved in image geometrical transformation? Develop the homogeneous from of this transformation? **[WBUT 2014, 2018]**

Answer:

1st Part:

A geometric transformation consists of two basic operations

i) a spatial transformation of coordinates

ii) intensity interpolation that assigns intensity values to the spatially transformed pixels.

The transformation of coordinates may be expressed as

$$(x, y) = T\{(v, w)\}$$

where (v, w) are pixel coordinates in the original image and (x, y) are the corresponding pixel coordinates in the transformed image.

For example, the transformation

$$(x, y) = T\{(v, w)\} = \left(v/2, w/2\right)$$

shrinks the original image to half its size in both spatial directions.

2nd Part:

One of the most commonly used spatial coordinate transformation is the affine transformation which has the general form

$$[x \ y \ 1] = [v \ w \ 1]\, T = [v \ w \ 1] \begin{bmatrix} t_{11} & t_{12} & 0 \\ t_{21} & t_{22} & 0 \\ t_{31} & t_{32} & 1 \end{bmatrix}$$

This is the homogenous form of transformation. This transformation can scale, rotate, translate or sheer a set of coordinate points depending on the value chosen for the element of matrix T.

4. State the applications of image transforms. What is energy compaction of image Transform? **[WBUT 2014]**

Answer:

1st Part:

Image transforms usually refers to a class of unitary matrices used for representing images and are applicable to filtering, data compression, feature extraction and other analysis. Two dimensional transforms are applied to image enhancement, restoration, encoding and description.

2nd Part:

Many common unitary transforms tend to pack a large fraction of signal energy into a few transform co-efficient which is generally termed as energy compaction.

5. If an image is rotated by an angle $\frac{\pi}{4}$, will there be any change in the histogram of that image? Justify your answer. **[WBUT 2015, 2017]**

Answer:

Since rotation is a transformation, so transformation makes change in the histogram of an image. Thus a processed image after transformation is obtained by mapping each pixel in the input image with intensity say r_k into a corresponding pixel with level S_k in the output image i.e., $S_k = T(r_k) = (L-1)\sum_{j=0}^{k} P_r(r_j)$

$P_r(r_j)$ is the probability density function and intensity levels in an image may be viewed as random variables in the interval $(0,\ L-1)$.

6. Explain CMY and CMYK colour model. **[WBUT 2019]**

Answer:

CMY stands for Cyan, magenta and yellow are the secondary colours of light or alternatively the primary colours of pigments. For example, when a surface coated with cyan pigment is illuminated with white light, no red light is reflected from the surface i.e. cyan subtracts red light from reflected white light.

In order to produce true black (which is the predominant colour in printing) a fourth colour black is added, giving rise to the CMYK colour model. So, CMY plus black i.e. CMYK.

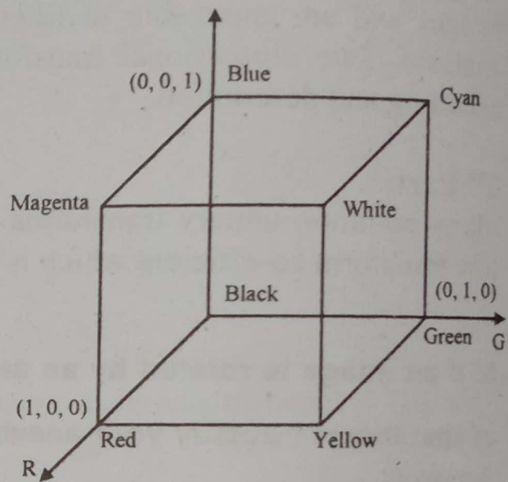
7. How is image represented in digital formats? **[WBUT 2019]**

Answer:
The method of converting an image which is continuous in space as well as its value into a discrete numerical form is called digitization. The steps for digital formal from analog formats are sampling and quantization.

Long Answer Type Questions

1. Briefly describe any three colour models. Write the conversion rules for converting RGB colour model to HSI colour model and vice-versa. How can a colour image be converted to gray scale image? **[WBUT 2009, 2010]**

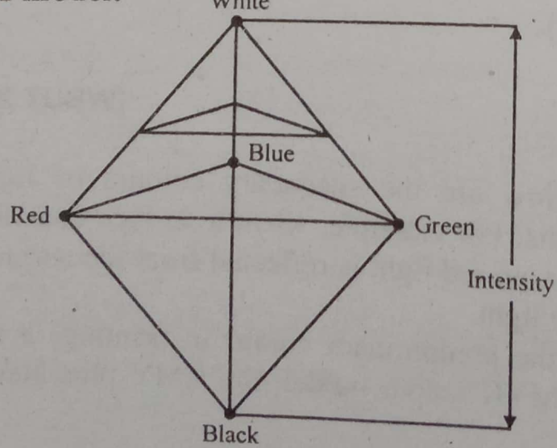
Answer:
1st Part:
RGB model:
In RGB model, an image consists of three independent image planes, one in each of the primary colours: red, green and blue, as shown below:

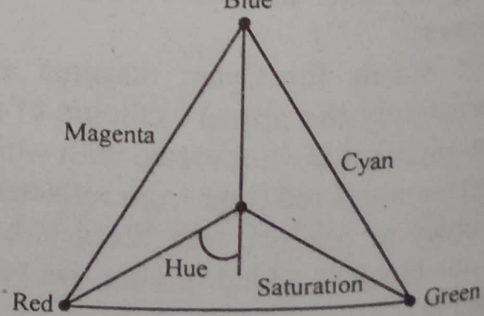


This is an additive model i.e., the colours present n the light add to form new colours and is appropriate for the mixing of coloured light. For example, red and green produce yellow, blue and green produces cyan and red and blue produces magenta and red, green and blue produces white. The RGB model is used for colour monitors and most video cameras.

CMY model:
The CMY (Cyan-Magenta-Yellow) is a subtractive model appropriate to absorption of colours, for example due to pigments in paints. The CMY model is used by printing devices and filters.

HSI model:





HSI (Hue-Saturation-Intensity) model is shown above. Hue is measured from red and saturation is given by distance from the axis. Colours on the surface of the solid are fully saturated i.e., pure colours and the grey scale spectrum is on the axis of the solid.

2nd Part:

RGB to HSI:

Suppose R, G and B are the red, green and blue values of a colour. The HSI intensity is given by the equation

$$I = (R + G + B)/3$$

Now, let m be the minimum value among R, G and B. The HSI saturation value of a colour is given by the equation

$$S = 1 - m/I \quad \text{if } I > 0 \text{, or}$$
$$S = 0 \quad \text{if } I = 0$$

To convert a colour's overall hue, H, to an angle measure,

$$H = \cos^{-1}\left[R - \frac{1}{2}G - \frac{1}{2}B\right] / \sqrt{R^2 + G^2 + B^2 - RG - RB - GB} \qquad \text{if } G \geq B$$

$$H = 360 - \cos^{-1}\left[R - \frac{1}{2}G - \frac{1}{2}B\right] / \sqrt{R^2 + G^2 + B^2 - RG - RB - GB} \quad \text{if } B > G$$

where the inverse cosine output is in degrees.

HSI to RGB:

If $H = 0$, then $R = I + 2IS$; $G = I - IS$; $B = I - IS$

If $0 < H < 120$, then

$$R = I + IS * \cos(H) / \cos(60 - H)$$
$$G = I + IS * [1 - \cos(H) / \cos(60 - H)]$$
$$B = I - IS$$

If $H = 120$, then $R = I - IS$; $G = I + 2IS$ & $B = I - IS$

If $120 < H < 240$, then

$$R = I - IS;\ G = I + IS * COS(H - 120) / COS(180 - H)$$
$$B = I + IS * [1 - \cos(H - 120) / \cos(180 - H)]$$

If $H = 240$, then $R = I - IS$, $G = I - IS$, $B = I + 2IS$ and if $240 < H < 360$, then

$$R = I + IS * [1 - \cos(H - 240) / \cos(300 - H)]$$
$$G = I - IS;\ B = I + IS * \cos(H - 240) / \cos(300 - H)$$

3rd Part:

There are three algorithms for converting colour image to gray scale image. These are **lightness, average** and **luminosity**.

The **lightness** method averages the most prominent and least prominent colours: $(\max(R,G,B)+\min(R,G,B))/R$.

It tends to reduce contrast

The **average** method simply averages the values: $(R+G+B)/3$.

The **luminosity** method is a more sophisticated version of the average method. It also averages the values, but it forms a weighted average to account for human perception. The formula for luminosity is $0.21R+0.72G+0.07B$. This method works best than other.

MATHEMATICAL PRELIMINARIES

Multiple Choice Type Questions

1. In a function $f(x, y)$ is finite in the space domain, the Fourier transform of $f(x, y)$ will be [WBUT 2009, 2011, 2018]
   a) finite
   b) infinite
   c) undefined
   d) zero

Answer: (a)

2. A wavelet transform is a special case of [WBUT 2009, 2011, 2016, 2018]
   a) Laplace transform
   b) Z-transform
   c) Fourier transform
   d) discrete cosine transform

Answer: (c)

3. Discrete cosine transform is a [WBUT 2013]
   a) Real Transform
   b) Imaginary Transform
   c) Both (a) and (b)
   d) None of these

Answer: (a)

4. If the DFT of $\{f(m,n) \to F(k,l)\}$ then for the sequence $f(m-m_0,n)$ the DFT becomes [WBUT 2016]
   a) $e^{(j\pi m_0 k)/N}.F(k,l)$
   b) $e^{-(j\pi m_0 k)/N}.F(k,l)$
   c) $e^{(j2\pi m_0 k)/N}.F(k,l)$
   d) $e^{-(j2\pi m_0 k)/N}.F(k,l)$

Answer: (c)

5. The kernel for 2D-DFT with $N = 4$ is calculated by the command [WBUT 2016]
   a) $dft2(4)$
   b) $fft2(4)$
   c) $dftmtx(4)$
   d) none of these

Answer: (b)

6. DCT is widely used for [WBUT 2016]
   a) image degradation
   b) image compression
   c) image restoration
   d) none of these

Answer: (b)

Short Answer Type Questions

1. Compare one dimension and two dimension DFT. [WBUT 2009]

Answer:
DFT (Discrete Fourier transform) is the sampled Fourier transform and therefore contain all frequencies forming an image. The number of frequencies corresponds to the number of pixels in the spatial domain image.

In 1D DFT, x goes up in step of 1 and that there are N samples, at values of x from 0 to $N-1$, i.e.,

$$F(u)=\frac{1}{N}\sum_{x=0}^{N-1} f(x)e^{-2\pi ixu/N}$$

In 2D DFT, for a square image of size $N\times M$ in x and y we have

$$F(u, v)=\frac{1}{NM}\sum_{x=0}^{N-1}\sum_{y=0}^{M-0} f(x, y)e^{-2\pi i}\left(\frac{xu}{N}+\frac{yv}{M}\right)$$

2D DFT are symmetric, periodic extensions separability, sampled Fourier transform in nature.

2D DFT can be implemented as two consecutive 1D DFT first in x direction, then in y direction or vice-versa

1D DFT are periodic in nature

2D DFT can be decomposed using 1D DFT as primitive.

In 1D DFT, the DFT and unitary DFT matrices are symmetric. The DFT or unitary DFT of a real sequence is conjugate symmetric about $\frac{N}{2}$.

2D DFT has a conjugate symmetry and also has a periodic extension.

2. Prove that imaginary part of a Fourier transform of an even function is zero.
[WBUT 2010]

Answer:

Let Fourier transform of $f(x)$ is $F(s)$, then

$$F(s)=\int_{-\infty}^{\infty} f(x)\cdot e^{isx}dx=\int_{-\infty}^{\infty} f(x)(\cos sx+i\sin sx)dx$$

$$=\int_{-\infty}^{\infty} f(x)\cdot\cos sx\cdot dx+i\int f(x)\sin sx\cdot dx$$

$\therefore$ The imaginary part of $F(s)$ is

$$=\int_{-\infty}^{\infty} f(x)\sin sx\cdot dx=\int_{0}^{\infty} f(x)\sin sx\cdot dx+\int_{-\infty}^{0} f(x)\cdot\sin sx\cdot dx$$

$$=\int_{0}^{\infty} f(x)\cdot\sin sx\cdot dx-\int_{\infty}^{0} f(-y)\sin(-sy)\cdot dy$$

$$\left[\because \text{ Let } y=-x \Rightarrow dy=-dx;\ \sin sx=\sin(-sy)=-\sin sy\right]$$

$$=\int_{0}^{\infty} f(x)\sin sx\cdot dx+\int_{\infty}^{0} f(y)\sin sy\,dy$$

$$\left[\because f(x) \text{ is even function } f(-x)=f(x)\right]$$

$$= \int_0^\infty f(x)\sin sx \cdot dx - \int_\infty^0 f(y)\sin sy\,dy$$

$$\left[\because \int_b^a f(x)\cdot dx = -\int_a^b f(x)\cdot dx\right]$$

$$= \int_0^\infty f(x)\sin sx \cdot dx - \int_0^\infty f(x)\sin sx \cdot dx = 0.$$

3. Show that the Fourier transform of the auto-correlation function of $f(x)$ is the power spectrum $|\tau(u)|^2$. **[WBUT 2010]**

Answer:

The Fourier transform of auto-correlation function $f(x)$ is

$$F[f(x)] = \int_{-\infty}^{\infty} f(x)\cdot e^{-j\omega x}\cdot dx$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x(t)\cdot x(t-x)e^{-j\omega x}\cdot dx$$

$$= \int_{-\infty}^{\infty} x(t)\cdot e^{-j\omega t}dt \int_{-\infty}^{\infty} x(t-x)\cdot e^{j\omega(t-x)}dx$$

$$= u\int_{-\infty}^{\infty} x(t-x)e^{j\omega(t-x)}\cdot dx$$

[$\because$ u is the energy density function]

Now, $t-x=n$ in the second integral, we have

$$F[f(x)] = u\int_{-\infty}^{\infty} x(n)e^{j\omega n}dx = u[-u] = |u|^2.$$

4. Discuss briefly about the usefulness of discrete cosine transform. **[WBUT 2011, 2015]**

Answer:

Discrete Cosine Transform (DCT) are used to convert data into the summation of a series of cosine waves oscillating at different frequencies. They are widely used in image and audio compression. They are very similar to Fourier transform, but DCT involves the use of just cosine functions and real co-efficients, whereas Fourier transformations make use of both sine and cosine and require the use of complex numbers. DCT are simpler to calculate. Both Fourier and DCT convert data from a spatial domain into a frequency domain and their inverse function convert things back.

5. What is Hough transform and where is it used? [WBUT 2011, 2014, 2015]

OR,

Discuss the Hough transform method for edge linking. [WBUT 2012, 2017]

OR

Explain Hough transformation and describe its application in image processing. [WBUT 2013]

Answer:

The Hough transform is a technique which can be used to isolate feature of a particular shape within an image. Because it requires that the desired features be specified in some parametric form. The classical Hough transform is most commonly used for the detection of regular curves such as lines, circles, ellipse etc. A generalized Hough transform can be employed in applications where a simple analytic descriptions of a features is not possible.

The Hough transform is particularly useful for computing global description of a features (where the number of solution classes need not be known as priori), given local measurements. The Hough transform can be used to identify the parameter of curve which best fits a set of given edge points. It is used to detect curves in pictures.

6. Define 4-adjacency, 8-adjancency and m-adjacency. Consider the two-image Subset S1 and S2 shown below:

		S1				S2			
0	0	0	0	0	0	0	1	1	0
1	0	0	1	0	0	1	0	0	1
1	0	0	1	0	1	1	0	0	0
0	0	1	1	1	0	0	0	0	0
0	0	1	1	1	0	0	1	1	1

For $v=\{1\}$, determine whether S1 and S2 are

(i) 4-connected
(ii) 8-connected
(iii) M-connected. [WBUT 2012]

Answer:

1st Part:

4-adjacency: Two pixels p and q with values from V are 4-adjacent of q is in the set $N_4(p)$.

8-adjancy: Two pixels p and q with values from V are 8-adjacent if q is in the set $N_8(p)$.

***m*-adjacency:** Two pixels p and q with values from V are m-adjacent of q is in $N_4(p)$ or q is in $N_D(p)$ and the set $N_4(p)\cap N_4(q)$ has no pixels whose values are from V.

2nd Part:

Let p and q be as shown below,

	S_1				S_2				
0	0	0	0	0	0	0	1	1	0
1	0	0	1	0	0	1	0	0	1
1	0	0	1	0	1^q	1	0	0	0
0	0	1	1	1_p	0	0	0	0	0
0	0	1	1	1	0	0	1	1	1

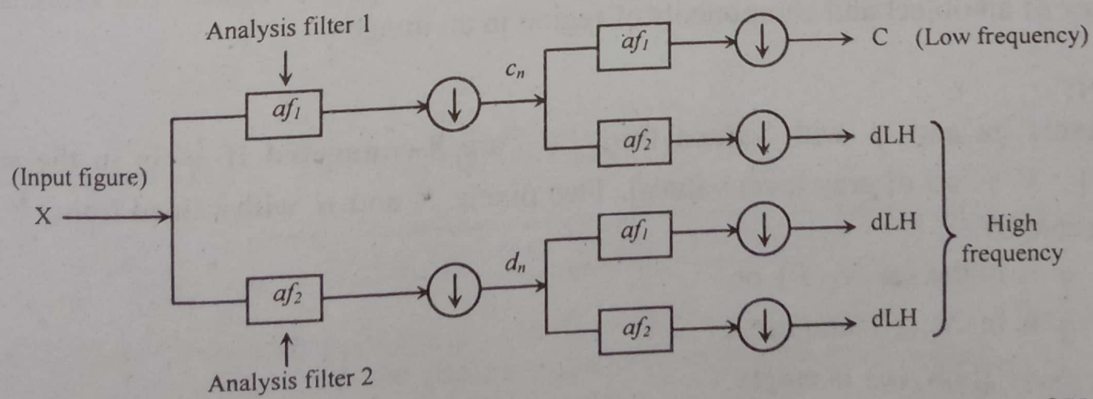
Fig: 1

Then, (i) S_1 and S_2 are not 4-connected because q is not in the set $N_4(p)$; (ii) S_1 and S_2 are 8-connected because q is in the set $N_8(p)$; and (iii) S_1 and S_2 are m-connected because (a) q is in $N_4(p)$ and (b) the set $N_4(p) \cap N_4(q)$ is empty.

7. Draw the schematic diagram of 2-D DWT synthesis filter bank structure for Haar Wavelet Transform and explain the components. **[WBUT 2013]**

Answer:

To use the wavelet transform for image processing we must implement a 2D vision of the analysis and synthesis filter banks. In the 2D case, the 1D analysis filter bank is final applied to the columns of the image and then applied to the news. If the image has N_1 rows and N_2 columns, then after applying the 1D analysis filter bank to each column, we have two sub-band images, each having $N_1/2$ rows and $N_2/2$ columns; after applying the 1D analysis filter bank to each row of both of the two sub-band images, we have four sub-band images each having $N_1/2$ rows and $N_2/2$ columns. This is illustrated in the diagram below. The 2D synthesis filter bank combines the four sub-band images to obtain the original image of size N_1 by N_2.



We denote X as input signal. The signal C represents the low frequency part of X, while the signal d represent the high frequency. The analysis filter bank finet filters X using a low pass and high pass filter. We denote the lowpass filter by af_1 (analysis filter 1) and the highpass filter by af_2 (analysis filter 2)

8. What is m-connectivity among pixels? Give an example. **[WBUT 2015, 2018]**

Answer:

Let v be a set of gray level values used to define connectivity, then two pixels p and q with values from v are m-connected if i) q is in $N_4(p)$ i.e., 4-connectivity 2) q is in $N_D(p)$ (i.e., 8-connectivity) and the set $N_4(p) \cap N_4(q)$ is empty.

Example:

$$\begin{array}{ccc} 0 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{array} \qquad \begin{array}{ccc} 0 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{array}$$

m - connectivity　　　　4 - connectivity

&

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

8 - connectivity

9. Define connectivity. What is the difference between 8-connectivity and m-connectivity? What are the different methods for calculating the distance of pixel? Explain with relevant example. **[WBUT 2019]**

Answer:

1st Part:

Connectivity between pixel is a fundamental concept that simplifies the definition of digital image concepts such as regions and boundaries. To establish of two pixels are connected it must be determined if they are neighbours and if their gray level satisfy the specified criterion of similarity connectivity between pixels issued for establishing boundary of an object and components of region in an image.

2nd Part:

Two pixels p and q with valued from 'V' are **8-connected** if q is in the set of $N_8(P)$ [$\because V$ is set of gray level values]. Two pixels P and q with valued from 'V' are 8-connected of

i) q is in the set $N_4(P)$ or

ii) q is in $N_D(P)$ and the set $N_4(P) \cap N_4(q)$ is empty.

m - connectivity is introduced to eliminate the multi-path connection that often arise when 8-connected is used.

3rd Part:
The different methods for calculating the distance of pixel are:
i) Euclidean distance
ii) D_4 distance (city block distance)
iii) D_8 distance (chessboard distance)

4th Part:
The Euclidean distance between P and q is defined as

$$D_4(P, q) = |x - S| + |Y - t|$$

where (x, y) and (S, t) are the coordinate of P and q.
D_4 distance is

$$D_4(P, q) = |x - S| + |Y - t|$$

and $D_8(P, q) = \max\{|x - s|, |y - t|\}$ city block distance.

Example:

		2		
	2	1	2	
2	1	0	1	2
	2	1	2	
		2		

The pixels with D_4 distance ≤ 2 from (x, y) i.e. centre point form the above contours of constant distance. The pixel with $D_4 = 1$ are the 4-neighbors of (x, y)

Example: Chessboard distance

2	2	2	2	2
2	1	1	1	2
2	1	0	1	2
2	1	1	1	2
2	2	2	2	2

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

Long Answer Type Questions

1. a) Discuss the global processing via the Hough transform. [WBUT 2011]
b) Explain the role of Discrete cosine transform in image processing. [WBUT 2011, 2014]

Answer:

a) A straight line at a distance s and orientation θ of the Fig. (a) can be represented as
$s = x\cos\theta + y\sin\theta$ .... (1)

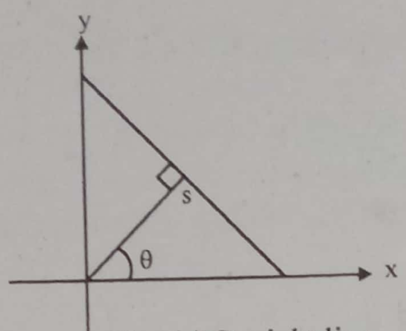


Fig: (a) Straight line

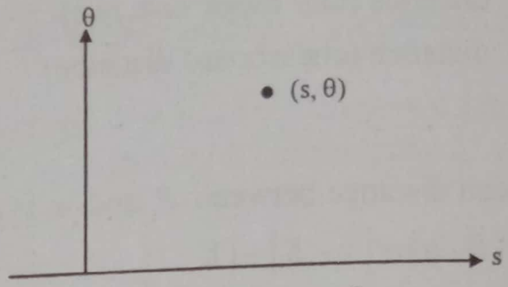


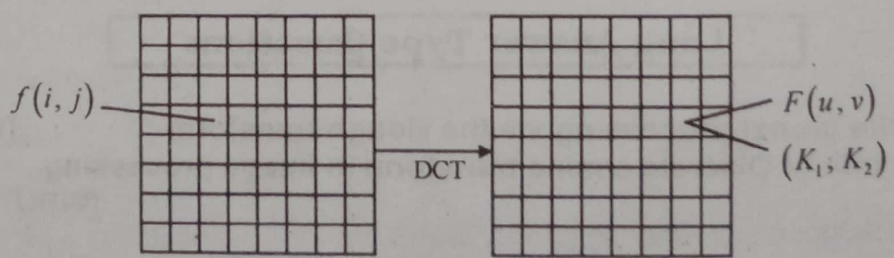
Fig: (b) Hough transform

The Hough transform of this line is just a point in the $(s,\ \theta)$ plane i.e., all the points on this line map into a single point (Fig. b). This fact can be used to detect straight lines in a given set of boundary points. Suppose we are given boundary points $(x_i,\ y_i)$ for $i = 1,\ 2,\ \ldots\ldots,\ N$. For some chosen quantized values of parameters s and θ, map each $(x_i,\ y_i)$ into the $(s,\ \theta)$ space and count $C(s,\ \theta)$, the number of edge points that map into the location $(s,\ \theta)$ i.e., set $C(s_k,\ \theta_\ell) = C(s_k,\ \theta_\ell) + 1$, if $x_i\cos\theta + y_i\sin\theta = s_k$ for $\theta = \theta_\ell$. Then the local maxima of $C(s,\ \theta)$ give the different straight line segments through the edge points. This two-dimensional search can be reduced to a one-dimensional search if the gradient θ_i at each edge location are also known. Differentiating both sides of Eqn. (1) w.r.t. x, we get

$$\frac{dy}{dx} = -\cot\theta = \tan\left(\frac{\pi}{2} + \theta\right)$$

Hence $C(s,\ \theta)$ need to be evaluated only for $\theta = \frac{\pi}{2} - \theta$. The Hough transform can be generalized to detect curves also other than straight lines.

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



The general equation for a one-dimensional DCT is defined by the equation

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

And the corresponding inverse one-dimensional DCT is simply $F'(u)$ i.e.,

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

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

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

and the corresponding inverse 2D DCT transform is simple $F^{-1}(u,v)$ i.e.,

$$\text{where}\quad A(\xi)=\begin{cases}\frac{1}{2} & \text{for}\quad \xi=0\\ 1 & \text{otherwise}\end{cases}$$

The basic operation of DCT is as follows:

- The input image is N by M.
- $f(i,j)$ is the intensity of the pixel in row i and column j.
- $F(u,v)$ is the DCT co-efficient in row k_1 and column k_2 of the DCT matrix.
- For most images, much of the signal energy lies at low frequency, these appear in the upper left corner of the DCT.
- Compression is achieved since the lower right values represent higher frequency and are often small.
- The DCT input is an 8×8 array of integer and the array contains each pixel's gray scale level and 8 bit pixels have levels from 0 to 255.

2. What is unitary transform? Define and compute the equations for unitary transform, Fourier transform, inverse Fourier transform for both 1-D and 2-D image. [WBUT 2014, 2018, 2019]

Answer:

A unitary transform is a specific type of linear transformation in which basic linear operation is exactly invertible and operator kernel satisfy certain orthogonality conditions. The forward unitary transform of the $N_1\times N_2$ image array $F(n_1,n_2)$ results in a $N_1\times N_2$ transformed image array as defined by

$$f(m_1,m_2)=\sum_{n_1=1}^{N_1}\sum_{n_2=1}^{N_2}F(n_1,n_2)A(n_1,n_2;m_1,m_2)$$

where $A(n_1,n_2;m_1,m_2)$ represents the forward transform kernel.

The one dimensional discrete fourier transform of a sequence $\{u(n), n = 0,.......N-1\}$ is defined as $V(K) = \sum_{n=0}^{N-1} u(n) W_N^{Kn}, K = 0,1,....N-1$ where $W_N \triangleq \exp\left\{\frac{-j2\pi}{N}\right\}$

The inverse transform is given by $u(n) = \frac{1}{N}\sum_{K_0}^{N-1} V(K) W_N^{-Kn}, n = 0,1,....N-1$

The two dimensional DFT of an $N \times N$ image $\{u(m,n)\}$ is a separable transform defined as

$$V(K,\ell) = \sum_{m=0}^{N-1}\sum_{n=0}^{N-1} u(m,n) W_N^{Km} W_N^{\ell n} \qquad 0 \le K, \ell \le N-1$$

and the inverse transform is

$$u(m,n) = \frac{1}{N^2}\sum_{K=0}^{N-1}\sum_{\ell=0}^{N-1} V(K,\ell) W_N^{-Km} W_N^{-\ell n} \qquad 0 \le m, n \le N-1$$

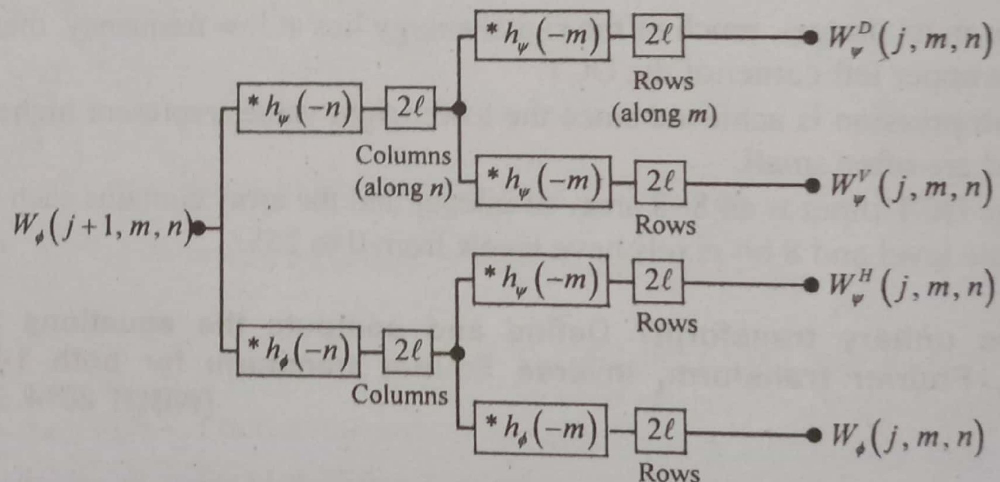
and the inverse unitary transform in two dimensional is

$$F(j,k) = \sum_{u=0}^{N-1}\sum_{v=0}^{N-1} f(u,v) B(j,k;u,v)$$

3. Draw the schematic of 2D DWT synthesis filter bank for Haar Wavelet transform and explain the components. **[WBUT 2019]**

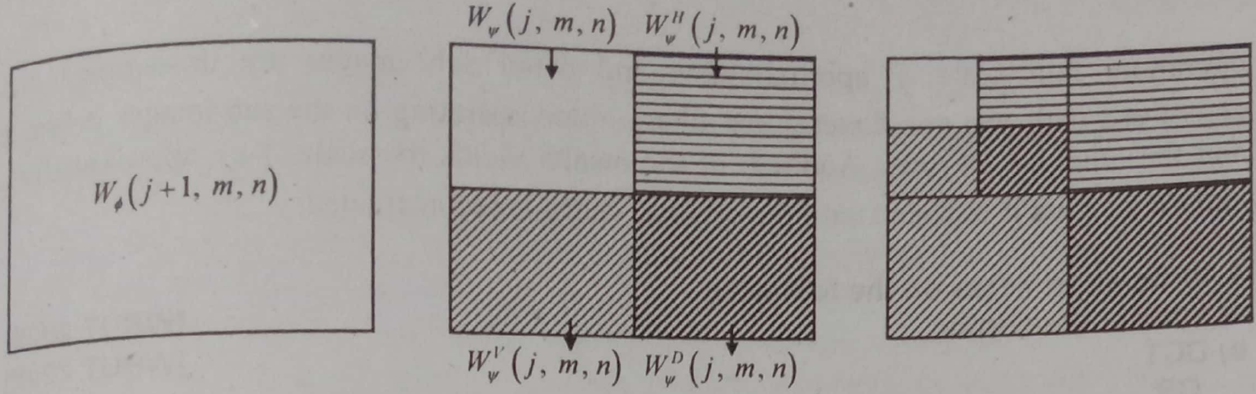
Answer:

Like the 1-D discrete wave let transform (DWT), the 2-D DWT can be implemented using digital filters and down samplers.

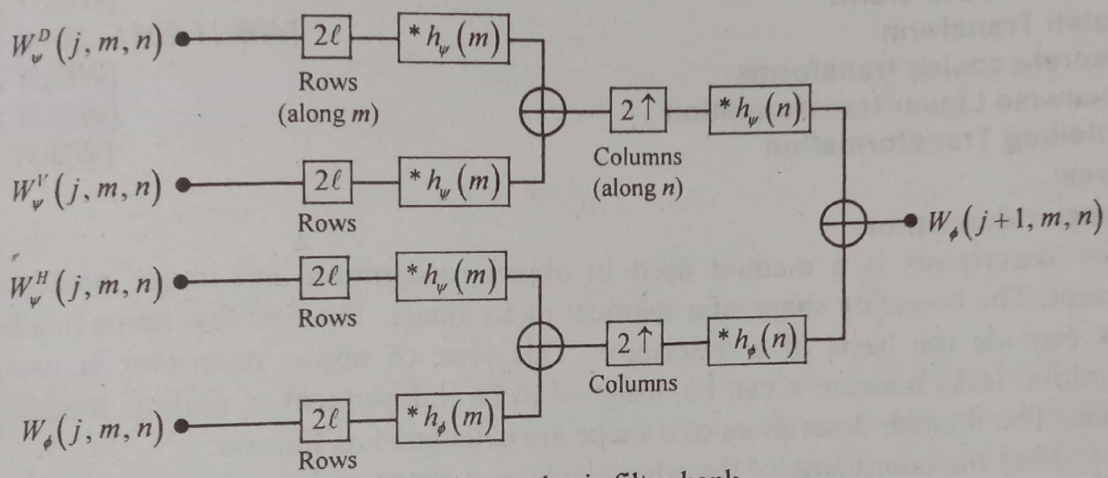


2D wavelet transform
(a) The analysis filter bank

With separable two dimensional scaling and wavelets functions, we simply take the 1-D FWT of the rows of $f(x, y)$, followed by the 1-D FWT of the resulting columns. Fig. shows the process in block diagram form



(b) The resulting decomposition



(c) The synthesis filter bank

Note that like its one dimensional counterpart, the 2-D FWT 'filters' the scale; i+1 approximation, co-efficient to construct the scale j approximation and detail co-efficient. In the two-dimensional case, however, we get three sets of detail co-efficient-The horizontal, vertical and diagonal details. The single-scale filter bank of Fig.(a) can be operated to produce a P scale transform in which scale J is equal to $J-1, J-2, \cdots\cdots, J-P$. As in the one-dimensional case, image $f(x, y)$ is used as the $W_\Psi(J, m, n)$ input. Convolving its rows with $h_\Psi(-n)$ and $h_\phi(-n)$ and down sampling its columns, we get two sub images whose horizontal resolution are reduced by a factor of 2. The high pass or detect component characterizes the image's high frequency information with vertical orientations; the low pass, approximation component contains its low frequency, vertical information. Both sub images are then filtered column wise and down sampled to yield four quarter-wise output sub images W_ϕ, W^H_ψ, W^V_ψ and W^D_ψ. These sub-images which are shown in the middle of (b), are the inner products of $f(x, y)$ and the two-dimensional scaling and wavelet functions followed by down sampling by two in each dimension. Two interactions of the filtering process produces the two-scale decompositions at the far right of Fig. (b).

Fig. (c) shows the synthesis filter bank that reverses the process just explained. As would be expected, the reconstruction algorithm is similar to the one-dimensional case. At each operation, four scale j approximation and detail sub images are un-sampled and convolved with two one dimensional filters – one operating on the sub images columns and the other on its rows. Addition of the results yields the scale $j+1$ approximation, and the process is repeated until the original image is reconstructed.

4. Write short notes on the following:

a) Fourier descriptor	**[WBUT 2009]**
b) DCT	**[WBUT 2009]**
OR,	
DCT in Image Compression	**[WBUT 2019]**
c) Hadamard transform	**[WBUT 2011]**
d) Walsh Transform	**[WBUT 2011, 2014, 2016]**
e) Discrete cosine transforms	**[WBUT 2012]**
f) Piecewise Linear transformation	**[WBUT 2015]**
g) Hotelling Transformation	**[WBUT 2016]**

Answer:

a) Fourier descriptor:

Fourier descriptors is a method used in object recognition and image processing to represent. The boundary shape of a segment in an image. The first few terms in a Fourier series provide the basis of a descriptor. This type of object descriptor is useful for recognition tasks because it can be designed to be independent of scaling, translation or rotation. The founder descriptors of a shape are calculated as follows:

1) Find the coordinate of the edge pixels of a shape and put them in a list in order, going clockwise around the shape.
2) Define a complex-valued vector using the coordinates obtained. Fourier descriptors inherit several properties from the fourier transform like translation invariance, scaling, rotation etc.

b) DCT:

Refer to Question No. 1(b) of Long Answer Type Questions.

c) Hadamard transform:

The Hadamard transform H_m is a $2^m \times 2^m$ matrix, the Hadamard matrix (scaled by a normalization factor), that transforms 2^m real numbers x_n into 2^m real numbers x_k. The Hadamard transform can be defined in two ways: recursively or by using the binary representation of the indices n and k. The lowest order of Hadamard matrix is 2 and the corresponding matrix is $H_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$.

If H_M be a Hadamard matrix of order M, then $H_{2M} = \begin{bmatrix} H_M & H_M \\ H_M & -H_M \end{bmatrix}$ is also Hadamard matrix of order $2M$.

Unlike the other transforms the elements of the basis vectors of the Hadamard transform take only the binary values $+-1$ and are therefore well suite for digital signal processing. The Hadamard transform H is real, symmetric and orthogonal i.e.,

$$H = H^{-1} = H^{T} = H^{*}$$

The Hadamard transform is a fast transform and has good to very good energy compaction for highly correlated images.
Hadamard transform is faster than sinusoidal transforms, since no multiplication is required; useful for digital hardware implementation of image processing algorithm.
Application areas are in image data compression, filtering and design of codes.

d) Walsh Transform:

When $N = 2^n$, the discrete Walsh transform of a function $f(x)$, denoted by $W(u)$ is obtained by substituting the Kernel

$$g(x,y) = \frac{1}{N}\prod_{i=0}^{n-1}(-1)bi(x)b(u)_{n-1-i}$$

$$W(u) = \frac{1}{N}\sum_{n=0}^{N-1} f(x)\prod_{i=0}^{n-1}(-1)bi(x)b(u)_{n-1-i}$$

Where $b_K(Z)$ is the K_{th} bit in the binary representation of Z. The array formed by the Walsh transformation Kernel is a symmetic matrix, whose rows and columns are orthogonal. Thus the inverse Kernel is identical to the forward Kernel except for a constant multiplicative factor of $1/N$

$$h(x,u) = \prod_{i=0}^{n-1}(-1)bi(x)b(n)_{n-1-i}$$

Wash transform kernel for N = 8

u \ x	0	1	2	3	4	5	6	7
0	+	+	+	+	+	+	+	+
1	+	+	+	+	–	–	–	–
2	+	+	–	–	+	+	–	–
3	+	+	–	–	–	–	+	+
4	+	–	+	–	+	–	+	–
5	+	–	+	–	–	+	–	+
6	+	–	–	+	+	–	–	+
7	+	–	–	+	–	+	+	–

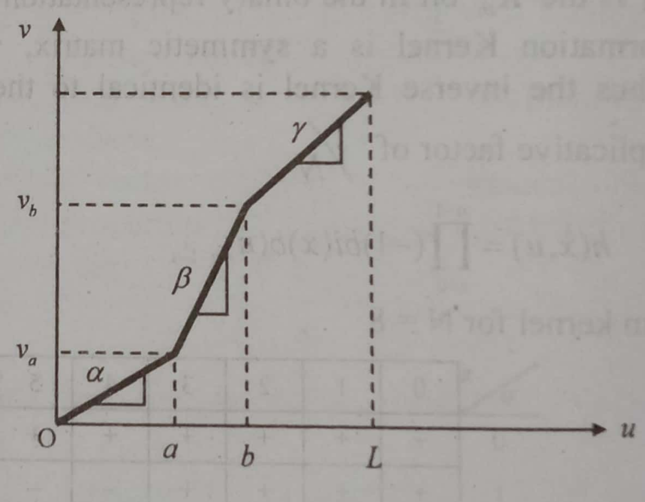
Thus the inverse Walsh transform is given by $f(x) = \sum_{u=0}^{N-1} W(u)\prod_{i=0}^{n-1}(-1)bi(x)b(u)_{n-1-i}$

Walsh transform consists of a series expansion of basic functions, whose values are +1 or −1.

e) Discrete cosine transforms:
Discrete Cosine Transform (DCT) is preferred for compression because it concentrates the largest percentage of the signal energy is a small percentage of the co-efficient, especially in the case of signals having high spatial correlation. DCT is a real transform in contrast to DFT (Discrete Fourier Transform) that is complex transform in general and thus, only real calculations are required for DCT computation. DCT is used in compression standards such as JPEG due to its very strong compression properties. There are chips that implement DCT in hardware.

f) Piecewise Linear transformation:
The principal advantage of piecewise linear transformation is that these function can be arbitrarily complex. But their specification requires considerably more user input.
Contrast stretching is the simplest piecewise linear transformation we may have various low contrast images and that might result due to various reasons such as lack of illumination, problem in imaging sensor or wrong setting of lens aperture during image acquisition. The idea behind contrast stretching is to increase the dynamic range of gray levels in the image being processed. Figure shows a typical contrast stretching transformation which can be expressed as

$$v = \begin{cases} \alpha u, & 0 \le u < a \\ \beta(u-a) + v_a, & a \le u < b \\ \gamma(u-b) + v_b, & b \le u < L \end{cases}$$



Contrast stretching transformation

The slope of the transformation is chosen greater than unity in the region of stretch.
For dark region stretch $\alpha > 1,\ a \simeq L/3;$

mid region stretch, $\beta > 1,\ b \simeq \frac{2}{3}L;$

bright region stretch, $\gamma > 1.$

The parameter a and b can be obtained by examining the histogram of the image.

g) Hotelling Transformation:

The Hotelling transform is a conventional image processing transformation. The equation is as follows. Suppose the points P_i are the feature points that we found, then covariance matrix is calculated by the mathematical relations. In this part, the eigen values and vectors are obtained. As the eigen values and vectors found, the geometric attack can be detected by some mathematical relations.

$$P_i = [x_i y_i]^T$$

$$m_o = \frac{1}{|0|}\sum_{i=1}^{|0|} P_i$$

$$C_o = \frac{1}{|0|}\sum_{i=1}^{|0|} P_i P_i^T - m_o m_o^T$$

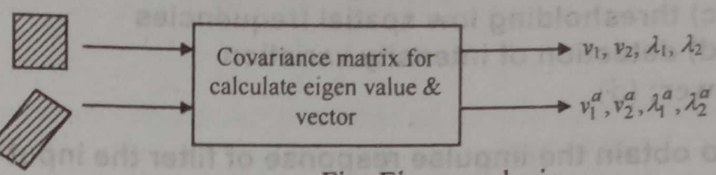


Fig: Eigen analysis

$$C_o e_j = \lambda_j e_j, j = 1,2$$

C_o is the covariance matrix calculated from the mean vector and summation from every feature points. After getting the covariance vectors, the eigen analysis can be used to find out the eigen vectors and values. Thus, eigen vectors and values represent the characteristic of one image. The above fig. shows a concept of eigen analysis. The eigen sets of original image and attacked ones can be used to determine what kind of geometric attack is.

IMAGE ENHANCEMENT

Multiple Choice Type Questions

1. Edge detection of an image broadly means [WBUT 2009, 2016, 2018]
   a) low spatial frequency enhancement
   b) high spatial frequency enhancement
   c) thresholding low spatial frequencies
   d) detection of intensity variation

Answer: (d)

2. To obtain the impulse response of filter the input impulse image should be like [WBUT 2010]
   a) a total white image of size $M \times N$
   b) a total black image of size $M \times N$
   c) a white dot in a centre of black image of $M \times N$
   d) a black dot in a centre of white image of $M \times N$

Answer: (c)

3. If the image is degraded by motion blur and added noise then .... Gives the best result [WBUT 2010]
   a) median filter
   b) inverse filter
   c) Wiener filter
   d) constraint least square filter

Answer: (c)

4. Diagonal edge can be detected by using which of the following masks? [WBUT 2010]

a)

0	1	0
1	–4	1
0	1	0

b)

1	1	1
1	1	1
1	1	1

c)

0	1	2
–1	0	1
–2	–1	0

d)

–1	–1	–1
–1	8	–1
–1	–1	–1

Answer: (c)

5. Which of the following grey level transformations produces image negative? **[WBUT 2010]**

a) $S = C\log(1+r)$
b) $S = L-1-r$
c) $S = Cr^k$
d) $S_k = \sum_{j=0}^{k} \frac{n_j}{n},\ k = 0, 1, 2, 3, \ldots\ldots(L-1)$

Answer: (b)

6. Wiener Filter is used for **[WBUT 2012, 2013, 2018]**

a) restoration
b) smoothening
c) sharpening
d) none of these

Answer: (d)

7. The effect, caused by the use of an insufficient number of gray levels in smooth areas of a digital image is called **[WBUT 2012, 2018]**

a) false counting
b) gray levels slicing
c) bit plane
d) thinning

Answer: (a)

8. Consider an image of size M × N with 64 gray levels. The total number of bits required to store this digitized image is **[WBUT 2012]**

a) M × N × 64
b) M × N × 63
c) M × N × 6
d) M × N × 8

Answer: (d)

9. Periodic noise happened due to **[WBUT 2012]**

a) infinite frequency
b) electrical & electromechanical interface during acquisition
c) ringing effect
d) none of these

Answer: (b)

10. Averaging filter is used for **[WBUT 2013, 2018]**

a) sharpening
b) contrast
c) brightness
d) smoothing

Answer: (d)

11. Which of the following is improved by histogram technique? **[WBUT 2013]**

a) Contrast
b) Sharpness
c) Brightness
d) Both (a) & (b)

Answer: (a)

12. Image restoration is a / an **[WBUT 2013]**

a) subjective process
b) objective process
c) (a) & (b) both
d) None of these

Answer: (d)

13. Measuring an intensity value of a fixed pixel including the effect of the neighborhood is called [WBUT 2014]
a) Averaging
b) Spatial filtering
c) (a) & (b) both
d) None of these

Answer: (c)

14. Usually frequency domain operations are [WBUT 2015]
a) Global operation
b) Mask operation
c) Point operation
d) none of the above

Answer: (b)

15. If the size of the mask for averaging is increased, the image will be [WBUT 2016]
a) noise free
b) blurred
c) degraded
d) none of these

Answer: (b)

16. Salt and pepper noise can be removed by [WBUT 2016]
a) weighted average filter
b) Gaussian filter
c) Median filter
d) high boost filter

Answer: (c)

17. The histogram equalization process [WBUT 2016]
a) blurs the image
b) fades the image
c) improves the brightness of the image
d) none of these

Answer: (c)

18. A spatial averaging filter in which all the coefficients are equal is called [WBUT 2018]
a) weighted averaging filter
b) median filter
c) box filter
d) none of these

Answer: (c)

19. What is the basis of numerous spatial domain processing? [WBUT 2019]
a) Transformation
b) Scaling
c) Histogram
d) None of these

Answer: (c)

20. The objective of the sharpening filter is [WBUT 2019]
a) highlight the intensity transition
b) highlight the low transition
c) highlight the bright transition
d) highlight the colour transition

Answer: (a)

21. Which is the image processing technique used to improve the quality of image for human viewing? [WBUT 2019]
a) Compression
b) Enhancement
c) Restoration
d) Analysis

Answer: (b)

22. **Histogram is the technique processed in** **[WBUT 2019]**
   a) **intensity domain**
   b) **frequency domain**
   c) **spatial domain**
   d) **undefined domain**

Answer: (c)

Short Answer Type Questions

1. What are called median filters? **[WBUT 2009]**

Answer:

Median filter is the best order statistics filter; it replaces the value of a pixel by the median of gray levels in the neighbourhood of the pixel.

$$f(x,\ y) = \underset{(S,\ t)\in S_{XY}}{\text{median}}\{g(S,\ t)\}$$

The original of the pixel is included in the computation of the median of the filter are quite possible because for certain types of random noise, provide excellent noise reduction capabilities with considerably less blurring then smoothing filters of similar size. These are effective for bi-polar and unipolar impulse noise.

2. Distinguish between image enhancement and image restoration. **[WBUT 2009, 2012, 2013, 2015, 2017, 2019]**

What is the equation for getting a negative image? **[WBUT 2012]**

Answer:

1st Part:

Enhancement technique is based primarily on the pleasing aspects it might present to the viewer. For example, contrast stretching. Whereas removal of image blur by applying a deblurrings function is considered as a restoration technique.

2nd Part:

The equation for getting a negative image is $S = L - 1 - r$.

3. What are image negatives? **[WBUT 2009, 2019]**

Answer:

The negative of an image with gray levels in the range $[0,\ L-1]$ is obtained by using the negative transformation, which is given by the expression $S = L - 1 - r$, where S is the output pixel and r is the input pixel.

4. Discuss the limiting effect of repeatedly applying a 3×3 spatial filter to a digital image. Ignore the border effects. **[WBUT 2010]**

Answer:

One of the easiest way to look at repeated applications of a spatial filter is to use superposition. Let $f(x,\ y)$ and $h(x,\ y)$ denote the image and the filter function respectively. Assuming square images of size $N \times N$ for convenience, we can express $f(x,\ y)$ as the sum of at most N^2 images, each of which has only one non-zero pixel (initially, we assume that N can be infinite). Then, the process of running $h(x, y)$ over $f(x, y)$ can be expressed as the following convolution:

$$h(x,y)*f(x,y)=h(x,y)*[f_1(x,y)+f_2(x,y)+........+f_N(x,y)]$$

Suppose the illustrative purposes that $f_i(x, y)$ has value 1 at its center, while the other pixels are valued 0 (fig. (a)). If $h(x, y)$ is a 3×3 mask of 1/9's (fig. (b)), then convolving $h(x, y)$ with $f_i(x, y)$ will produce an image, with a 3×3 array of 1/9's at its center and 0's elsewhere (fig. (c)). If $h(x, y)$ is now applied to this image, the resulting image will be as shown in fig. (d). Note that the sum of the non-zero pixels in both fig. (c) and fig. (d) is the same and equal to the value of the original pixel. Thus, it is intuitively evident that successive application of $h(x, y)$ will "diffuse" the non-zero value of $f_i(x, y)$. Since the sum remains constant, the value of the non-zero elements will become smaller and smaller, as the number of applications of the filter increases. The overall result is given by adding all the convolved $f_k(x, y)$, $k=1,2,......,N^2$.

It is noted that every iteration of blurring further diffuses the value outwardly from the starting point. In the limit, the values would get infinitely small, but because the average value remains constant, this would require an image of infinite spatial proportions.

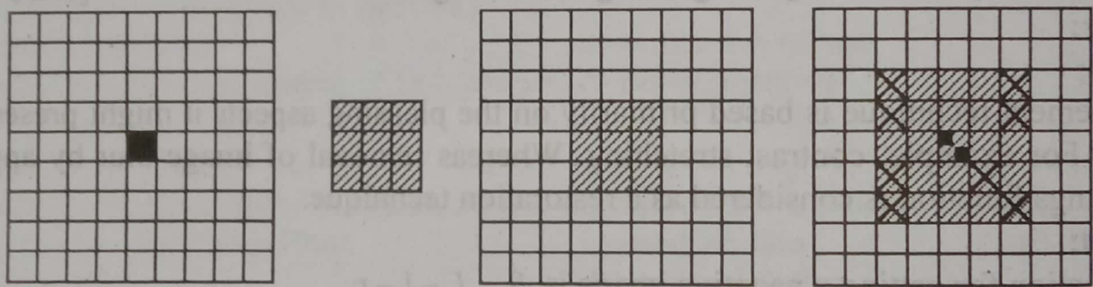


Fig: (a) Fig: (c) Fig: (d)

■ = 1, □ = 0, ▨ = 1/9, ◪ = 9/81 etc.

5. Suppose a digital image is subjected to histogram equalization. Show that a second pass of histogram will produce exactly the same result as the first pass?

[WBUT 2010]

Answer:

Let n be the total number of pixels and let n_{r_j} be the number of pixels in the input image with intensity value r_j. Then, the histogram equalization transformation is

$$s_k=T(r_k)=\sum_{j=0}^{k}\frac{n_{r_j}}{n}=\frac{1}{n}\sum_{j=0}^{k}n_{r_j}$$

Since every pixel (and no others) with value r_k is mapped to value s_k, it follows that $n_{s_k}=n_{r_k}$. A second pass of histogram equalization would produce values V_k according to the transformation

$$V_k = T(s_k) = \sum_{j=0}^{k} \frac{n_{s_j}}{n}, \text{ but } n_{s_j} = n_{r_j}$$

Then $$V_k = T(s_k) = \sum_{j=0}^{k} \frac{n_{r_j}}{n} = s_k$$

which shows that a second pass of histogram equalization would yield the same result as the first pass.

6. Image blurring caused by long term exposure to atmospheric turbulence can be modelled by the transfer function $H(u, v) = \frac{\exp\left(-\left(u^2+v^2\right)\right)}{2\sigma^2}$. Assume negligible noise.

What is the equation of Weiner filter to be used to reconstruct an image blurred by this type of degradation? **[WBUT 2010]**

Answer:

The explanation will be clearer if we start with one variable. We want to show that, if

$$H(u) = e^{-u^2} / 2\#^2$$

Then $$h(t) = 3^{-1}\left[H(u)\right] = \int_{-\infty}^{+\infty} e^{-u^2} / 2\#^2 e^{j2\pi ut} du = \sqrt{2\pi}\,\#^{-2\pi^2 \#^2 t^2}$$

We can express the integral in the preceding equation as

$$h(t) = \int_{-\infty}^{+\infty} e^{-\frac{1}{2\#^2}\left[u^2 - j4\pi\#^2 ut\right] du}$$

Making use of the identity

$$e^{\frac{-(2\pi)^2 \#^2 t^2}{2}} \cdot e^{\frac{(2\pi)^2 \#^2 t^2}{2}} = 1$$

In the preceding integral yields

$$h(t) = e^{\frac{-(2\pi)^2 \#^2 t^2}{2}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2\#^2}\left[u^2 - j4\pi\#^2 ut - (2\pi)^2 \#^2 t^2\right] du} = e^{\frac{-(2\pi)^2 \#^2 t^2}{2}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2\#^2}\left[u - j2\pi\#^2 t\right]^2 du}$$

Now, we make the change of variables

$$r = u - j2\pi\#^2 t$$

Then $dr = du$ and the preceding integral becomes,

$$h(t) = e^{\frac{-(2\pi)^2 \#^2 t^2}{2}} \int_{-\infty}^{+\infty} e^{-\frac{r^2}{2\#^2} dr}$$

Finally, we multiply and divide the right side of this equation by $\sqrt{2\pi}\,\#$ and obtain

$$h(t) = \sqrt{2\pi}\,\# e^{\frac{-(2\pi)^2 \#^2 t^2}{2}} \left[\frac{1}{\sqrt{2\pi}\,\#} \int_{-\infty}^{+\infty} e^{\frac{-r^2}{2\#^2} dr}\right]$$

The expression inside the bracket is recognized as the Gaussian probability density function whose value from $-\infty$ to $+\infty$ is 1.

Therefore,

$$h(t)=\sqrt{2\pi}\,\#e^{-(2\pi)^2\#^2 t^2}$$

with the preceding result as background we are now ready to show that

$$h(u,v)=\Im^{-1}\left[A\cdot e^{-(u^2+v^2)/2\#^2}\right]=A2\pi\#^2\,e^{-2\pi^2\#^2(t^2+z^2)}$$

Hence, the result.

7. Explain Unsharp masking and High boost filtering. Write the expression of Laplacian operator for an image of two variables. [WBUT 2011, 2014, 2015]

Answer:

1st part:

A process that has been used for many years by the printing and publishing industry to sharpen images consist of subtracting an un-sharp (smoothed) version of an image from the original image. This process, called un-sharp masking.

A slight generalization of un-sharp masking is called high boost filtering. The equation of un-sharp masking is $f_S(x,y)=f(x,y)-\tilde{f}(x,y)$ and the equation of high boost filtering is $f_{ub}(x,y)=Af(x,y)-\tilde{f}(x,y)$. The steps involved in unsharp masking and high boost filtering is,

a) Blur the original image
b) Subtract the blurred image from the original to get a image called as mask image.
c) Add the mask to the original

2nd part:

The widely used second derivative operator is the Laplacian edge detector.

Given an image matrix, the Laplacian of the image function is the second order partial derivatives along x and y, direction i.e. $\nabla^2 g=\dfrac{\partial^2 g}{\partial x^2}+\dfrac{\partial^2 g}{\partial y^2}$

In the digital approximation of the second order, partial derivative in x-direction is

$$\frac{\partial^2 g}{\partial x^2}=\frac{\partial\{g(r,c)-g(r,c-1)\}}{\partial x}=\frac{\partial g(r,c)}{\partial x}-\frac{\partial g(r,c-1)}{\partial x}$$
$$=g(r,c)-g(r,c-1)-g(r,c-1)+g(r,c-2)$$
$$=g(r,c)-2g(r,c-1)+g(r,c-2)$$

Similarly, the second order derivative along y-direction is given by

$$\frac{\partial^2 g}{\partial y^2}=g(r,c)-2g(r-1,c)+g(r-2,c)$$

Both the approximating equation are centered around the point $(r-1,c-1)$. Conveniently, this pixel could be replaced by (r,c). The corresponding laplacian mark for convolution are given by

0	1	0
1	−4	1
0	1	0

8. Explain the operation of a median filter. **[WBUT 2011, 2014, 2017]**

Answer:

The median filter as the name implies replaces the value of a pixel by the median of the intensity values in the neighbourhood of that pixel. Median filters are quite popular because for certain type of random noise, they provide excellent noise reduction capabilities with considerably less blurring than linear smoothing filters of similar size.

In order to perform median filtering at a point in an image, we first sort the values of the pixel in the neighbourhood, determine their median, and assign that value to the corresponding pixel in the filtered image. For example, in a 3×3 neighbourhood the median is the 5th largest value; when several values in the neighbourhood are same, all equal values are grouped. Thus the principal function of median filters is to force points with distinct intensity levels to be more like their neighbours. In fact, isolated clusters of pixel that are light or dark with respect to their neighbours and whose area is less than $m^2/2$ (one half the filter area) are eliminated by $m \times m$ median filter. Eliminated means forced to the median intensity of the neighbours. Larger clusters are effected considerably less.

9. What is the value of the marked pixel after a 5×5 median filtering? **[WBUT 2011]**

$$\begin{array}{ccccc} 2 & 1 & 3 & 4 & 5 \\ 1 & 1 & 0 & 2 & 3 \\ 2 & 0 & \underline{0} & 1 & 2 \\ 5 & 1 & 2 & 3 & 1 \\ 3 & 3 & 1 & 2 & 0 \end{array}$$

Answer:

In order to perform a median filter, the pixels have to be sorted either in ascending or descending order.

Step 1: Arrange the pixel in ascending order

0000 1111111 222222 3333 44 55

Step 2: To determine the median value

$$\cancel{0}\cancel{0}\cancel{0}\cancel{0}\ \cancel{1}\cancel{1}\cancel{1}\cancel{1}\cancel{1}\cancel{1}\cancel{1}\ \cancel{2}\,2\,\cancel{2}\cancel{2}\cancel{2}\cancel{2}\ \cancel{3}\cancel{3}\cancel{3}\cancel{3}\ \cancel{4}\cancel{4}\ \cancel{5}\cancel{5}$$

The median value is found to be 2. Hence the marked pixel $\underline{0}$ will be replaced by 2.

10. Briefly describe the smoothing linear spatial filtering. What is bit plane slicing method? **[WBUT 2011, 2018]**

Answer:

1st Part:

The output (response) of a smoothing linear spatial filter is simply the average of pixels contained in the neighborhood of the filter mask. These filters sometimes are called averaging filters. The idea behind the smoothing filters is straight forward. By replacing the value of every pixel in an image by the average of the intensity level in the neighborhood defined by the filter mask, this process results in an image with reduced "sharp" transitions in intensities. Because random noise typically consists of sharp transitions in intensity levels, the most obvious application of smoothing is noise reduction. However edges also are characterized by sharp intensity transition, so averaging filters have the undesirable side effect that they blur edges. Another application of this type of process includes the smoothing of false contour that result from using an insufficient number of intensity levels. Fig shows the two 3×3 smoothing filters.

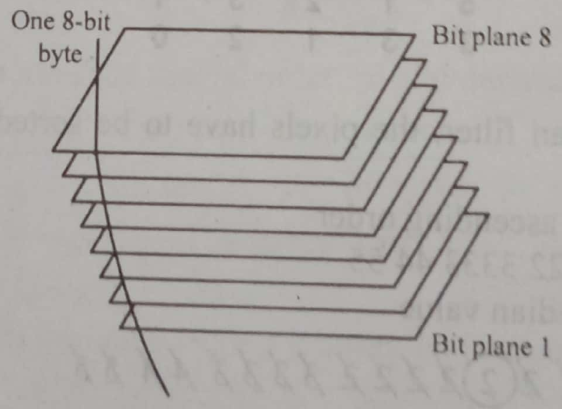
$$\frac{1}{9}\times\begin{array}{|c|c|c|}\hline 1 & 1 & 1\\\hline 1 & 1 & 1\\\hline 1 & 1 & 1\\\hline\end{array} \qquad \frac{1}{16}\times\begin{array}{|c|c|c|}\hline 1 & 2 & 1\\\hline 2 & 4 & 2\\\hline 1 & 2 & 1\\\hline\end{array}$$

The constant multipliers in front of each mask is equal to 1diveded by the sum of the values of its coefficients, as is required to compute an average.

Last Part:

Pixels are digital number composed of bits. For example, the intensity of each pixel in a 256-level gray-scale image is composed of 8 bit. Instead of highlighting intensity-level ranges, we could highlight the contribution made to total image appearance by specific bits.

As illustration, an 8-bit image may be considered as being composed of eight 1-bit planes, with plane of containing the lowest order bit of all pixels in the image and plane 8 all the highest order bits.



11. If all the pixels of an image are shuffled, will there be any change in the histogram of that image? Justify your answer. **[WBUT 2011]**

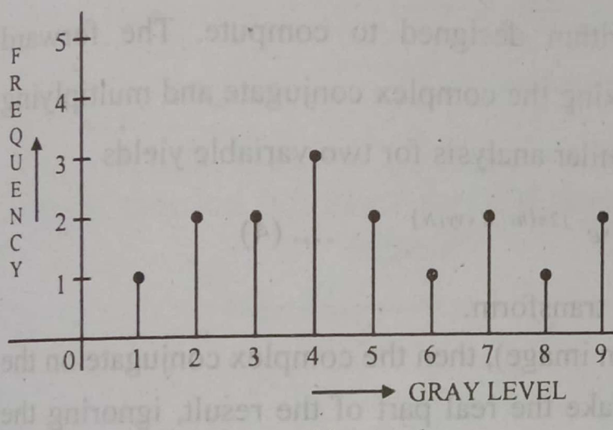
Answer:

If all the pixels in an image are shuffled there will not be any change in the histogram. A histogram gives only the frequency of occurrence of the gray level. Consider the following two images.

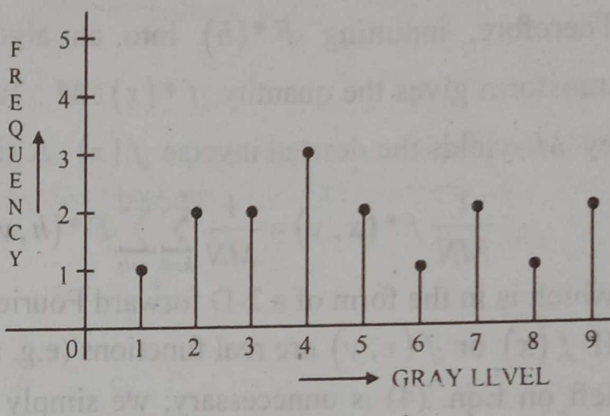
$$\begin{vmatrix} 1 & 2 & 4 & 8 \\ 3 & 5 & 7 & 9 \\ 4 & 2 & 4 & 6 \\ 9 & 7 & 3 & 5 \end{vmatrix} \qquad \begin{vmatrix} 9 & 7 & 3 & 5 \\ 3 & 5 & 7 & 9 \\ 1 & 2 & 4 & 8 \\ 4 & 2 & 4 & 6 \end{vmatrix}$$

Image 1 Image 2

Image 2 is obtained by shuffling the rows of image 1. Their corresponding histogram are



Histogram of Image 1



Histogram of Image 2

From the two histogram it is clear that even if all the pixels in an image are shuffled there will not be any change in the histogram of the image.

12. What is understood by spatial domain representation? **[WBUT 2012, 2019]**
Compute the Inverse Fourier Transform using Forward Transform Algorithm. **[WBUT 2012]**

Answer:

1st Part:

The spatial domain can be denoted by the expression $g(x, y) = T[f(x, y)]$ where $f(x, y)$ is the input image, $g(x, y)$ is the output image and T is an operator on f defined over a neighbourhood of point (x, y). The operator can apply to a single image or to a set of images, such as performing the pixel-by-pixel sum of a sequence of images for noise reduction.

2nd Part:

2-D Fourier transform can be computed via the application of 1-D transforms. The two equations of above transform are

$$F(h)=\frac{1}{M}\sum f(x)\cdot e^{-j2\pi hx/M} \quad \text{.... (1)}$$

for $h=0,1,2,\ldots\ldots,M-1$ and

$$f(x)=\sum_{h=0}^{M-1} F(h)\cdot e^{j2\pi hx/M} \quad \text{.... (2)}$$

for $x=0,1,2,\ldots\ldots,M-1$.

Taking the complex conjugate of Eqn. (2) and dividing both sides by M yields

$$\frac{1}{M}f^*(x)=\frac{1}{M}\sum_{h=0}^{M-1} F^*(h)\cdot e^{-j2\pi hx/M} \quad \text{.... (3)}$$

Comparing this result (3) with (1), it shows the right side of (3) is in the form of forward Fourier transform.

Therefore, inputting $F^*(h)$ into an algorithm designed to compute. The forward transform gives the quantity $f^*(x)/M$. Taking the complex conjugate and multiplying by M -yields the desired inverse $f(x)$. A similar analysis for two variable yields

$$\frac{1}{MN}f^*(x,y)=\frac{1}{MN}\sum_{h=0}^{M-1}\sum_{v=0}^{N-1} F^*(h,v)\cdot e^{-j2\pi(hx/M+vy/N)} \quad \text{.... (4)}$$

which is in the form of a 2-D forward Fourier transform.

If $f(x)$ or $f(x,y)$ are real functions (e.g. an image), then the complex conjugate on the left on Eqn. (4) is unnecessary; we simply take the real part of the result, ignoring the complex term.

13. What is Quad-tree? **[WBUT 2012]**

Answer:

The quad-tree data structure is widely used in digital image processing and computer graphics for modelling spatial segmentation of images and surfaces. It is a free in which each node has four descendants. Since most algorithms based on quad-tree require complex navigation between nodes, efficient traversal methods as well as efficient storage technique are of great interest. In computer graphics quad-trees and octrees are used as spatial partitioning representations. More recently quad-trees have become popular to store triangulated surfaces as shown below.

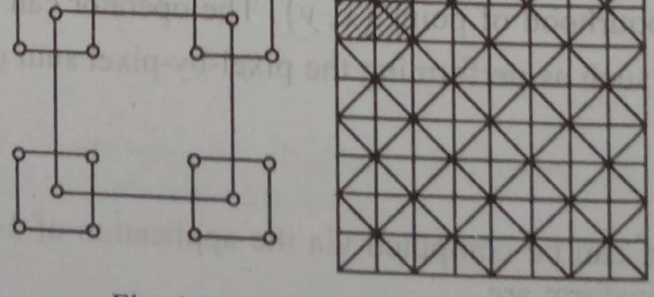


Fig: (a)
Quad-tree

Fig: (b)
Triangulation

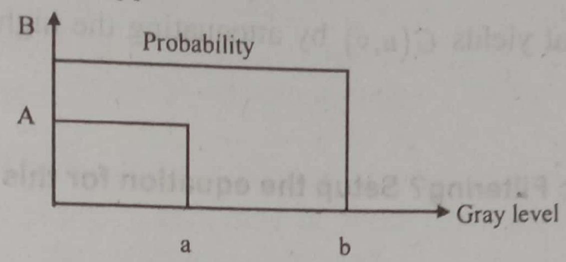
14. What is salt and pepper noise? **[WBUT 2012, 2013, 2014]**
What is Gaussian noise? **[WBUT 2013, 2014]**

Answer:

1st Part:

Salt and pepper noise is also known as impulse noise. This noise can be caused by sharp and sudden disturbances in the image signal. Its appearance is randomly scattered white or black (or both) pixel over the image.

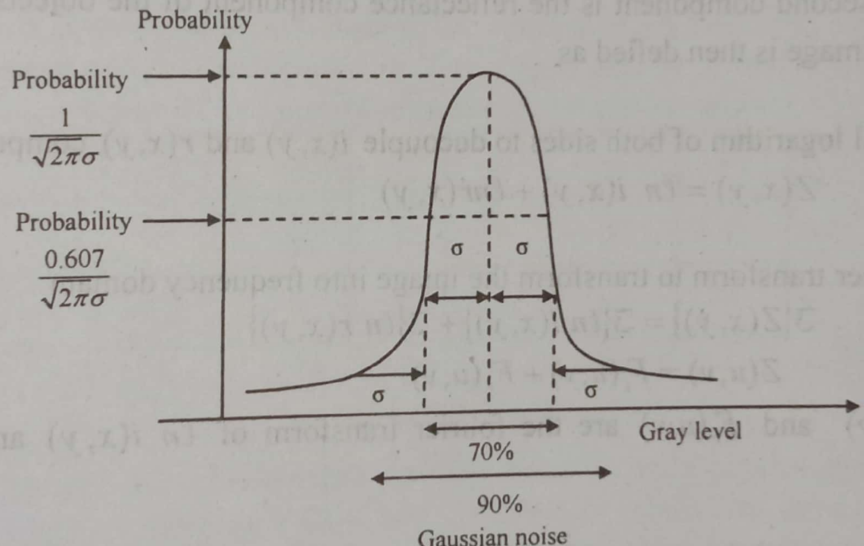
The block diagram of salt & pepper noise



Salt & pepper noise

2nd Part:

Gaussian noise is caused by random fluctuations in the signal. It has a normal probability density function. It is useful for modeling natural processes which introduce noise



Gaussian noise

Here only possible values a and b are possible. For an 8 bit/pixel image, the typical intensity value for pepper noise is close to 0 and for salt noise is close to 255.

It is caused by malfunctioning camera sensor cells, by memory cell failure or by synchronization errors in the image digitizing.

15. What is the "frequency" of an image? Explain the smoothing frequency domain filters. **[WBUT 2013]**

Answer:

1st Part:

Any image in spatial domain cam be represented in a frequency domain. High frequency component corresponds to edges in an image and low frequency components in an image corresponds to smooth regions.

2^{nd} **Part:**
Smoothing frequency domain filter edges and other sharp transition of the gray levels of an image contribute significantly to the high frequency components of its Fourier transformation. Hence smoothing is achieved in the frequency domain by attenuation a specified range of high frequency components in the transform of a given image. Basic model of filtering in the frequency domain is $F(u,v) = H(u,v).F(u,v)$

$F(u,v)$ is the Fourier transform of the image to be smoothed. Objective is to find out a filter function $H(u,v)$ that yields $G(u,v)$ by attenuating the high frequency component of $F(u,v)$.

16. What is Holomorphic Filtering? Setup the equation for this filter. **[WBUT 2014, 2018]**

Answer:
Homomorphic filtering uses illumination – reflectance model in its operation. This model consider the image is been characterized by two primary components. The first component is the amount of source illumination incident on the scene being viewed $i(x, y)$. The second component is the reflectance component of the objects on the scene $r(x, y)$. The image is then defied as

Stage 1
Take a natural logarithm of both sides to decouple $i(x, y)$ and $r(x, y)$ components i.e.

$$Z(x, y) = \ell n\ i(x, y) + \ell n r(x, y)$$

Stage 2
Use the Fourier transform to transform the image into frequency domain

$$\Im\{Z(x, y)\} = \Im\{\ell n\ i(x, y)\} + \Im\{\ell n\ r(x, y)\}$$

or $$Z(u,v) = F_i(u,v) + F_r(u,v)$$

where $F_i(u,v)$ and $F_r(u,v)$ are the fourier transform of $\ell n\ i(x, y)$ and $\ell n\ r(x, y)$ respectively.

Stage 3
High pass the $z(u,v)$ by means of a filter function $H(u,v)$ in frequency domain and get a filtered version $S(u,v)$ as

$$S(u,v) = H(u,v) \cdot Z(u,v) = H(u,v)F_i(u,v) + H(u,v)F_r(u,v)$$

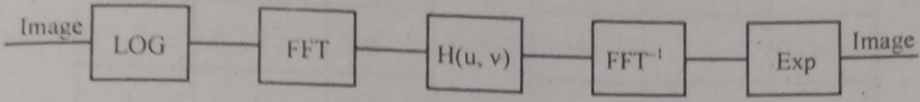
State 4
Take an inverse fourier transform to get the filtered image in the spatial domain

$$S(x, y) = \Im^{-1}\{S(u,v)\} = \Im^{-1}\{H(u,v)F_i(u,v) + H(u,v)F_r(u,v)\}$$

State 5
The filtered enhanced image $g(x, y)$ can be obtained as

$$g(x, y) = \exp\{S(x, y)\}$$

The block diagram is as follows:



17. a) What are the steps to perform histogram equalization? [WBUT 2016]
b) What is the necessity for histogram equalization?

Answer:

a) Histogram equalization can be done in three steps:
i) Compute the histogram of the image
ii) Calculate the normalized sum of histogram.
iii) Transform the input image to an output image

b) Histogram equalization is used to enhance contrast. It is not necessary that contrast will always be increase in this. There may be some cases where histogram equalization can be worse. In that case the contrast is decreased.

18. Show how Prewitt and Sobel operators can be used to smooth the noise. [WBUT 2016]

Answer:

Sobel operator is a simple approximation to the concept of gradient with smoothing. The 3×3 convolution mask is usually used to detect gradients in X and y directions. The operator consists of a pair of 3×3 convolution kernels as shown below:

+1	+2	+1
0	0	0
−1	−2	−1

G_x

+1	+2	+1
0	0	0
−1	−2	−1

G_y

Fig: Masks used by Sobel operator

One kernel is simply the other rotated by 90°. These kernels are designed to respond maximally to edges running vertically and horizontally relative to the pixel grid, one kernel for each of the two perpendicular orientations. The kernels can be applied separately to the input image, to produce separate measurements of the gradient components in each orientation (call these Gx and Gy). These can then be combined together to find the absolute magnitude of the gradient at each point and the orientation of that gradient. The gradient magnitude is given by $|G| = \sqrt{Gx^2 + Gy^2}$. Typically, an approximate magnitude is computed using $|G| = |Gx| + |Gy|$ which is much faster to compute.

The Prewitt filter is very similar to sobel filter. The 3×3 total convolution mask is used to detect gradient in the x and y directions as shown below.

−1	0	+1
−1	0	+1
−1	0	+1

G_x

+1	+1	+1
0	0	0
−1	−1	−1

G_y

Prewitt filter is a fast method for edge detection. The difference with respect to sobel filter is the spectral response. It is well suitable for well-contrasted noiseless image. Effect of noise is reduced in case of Prewitt and sobel operators by inherent averaging of neighbouring pixels. Therefore, to achieve the desired result, gradient operators are usually preceded by noise cleaning. The result of these operator are

$$\theta = \arctan\left(\frac{G_y}{G_x}\right)$$

19. Explain what you understand about image histogram of a digital image. What is the DC component of the following image? **[WBUT 2019]**

$$f = \begin{bmatrix} 1 & 3 & 4 \\ 5 & 6 & 7 \\ 8 & 9 & 11 \end{bmatrix}$$

Suggest one way of generating colour histograms in the RGB colour space from images. What are the number histogram components in such a histogram?

Answer:

1st Part:

The histogram of an image normally refers to a histogram of the pixel intensity values. The histogram is a graph showing the number of pixels in an image at each different intensity values found in that image.

2nd Part:

DC component of an image

$$F(\theta,0) = \sum_{x=0}^{M-1}\sum_{Y=0}^{N-1} f(x,y)$$

i.e., $F(\theta,0)$ is the sum of all terms on the original matrix

i.e. $$f = \begin{bmatrix} 1 & 3 & 4 \\ 5 & 6 & 7 \\ 8 & 9 & 11 \end{bmatrix}$$

So, $$F(\theta,0) = [1+3+4+5+6+7+8+9+11] = 54$$

3rd Part:

A standard way of generating a colour histogram of an image is to concatenate 'N' higher order bits for the red, green and blue values in the RGB space.

4th Part:
The histogram then has 2^{3N} bins which accumulate the count of pixels with similar colour.

20. What are spatial domain and frequency domain technique? Explain Mask or Kernels. What is Median Filter? What is Image Enhancement? Explain gray level slicing. **[WBUT 2019]**

Answer:
1st Part:
Spatial domain technique deals with image plane itself. It works based on direct manipulation of pixels.
Frequency domain technique works based on modifying Fourier transform. It deals with the rate of pixel change.

2nd Part:
In image processing, a kernel, convolution matrix, or mask is a small matrix. It is used for blurring, sharpening, embossing, edge detection, and more. This is accomplished by doing a convolution between a kernel and an image.

3rd Part:
The Median Filter is a non-linear digital filtering technique, often used to remove noise from an image or signal. Such noise reduction is a typical pre-processing step to improve the results of later processing. Median filtering is very widely used in digital image processing because, under certain conditions, it preserves edges while removing noise, also having applications in signal processing.

4th Part:
Image enhancement is the process of adjusting digital images so that the results are more suitable for display or further image analysis. For example, you can remove noise, sharpen, or brighten an image, making it easier to identify key features.

5th Part:
Grey level slicing is equivalent to band pass filtering. It manipulates group of intensity levels in an image up to specific range by diminishing rest or by leaving them alone. This transformation is applicable in medical images and satellite images such as X-ray flaws, CT scan.

Long Answer Type Questions

1. a) What effect would, setting to zero the lower order bit planes, have on the histogram of an image, in general?
b) What would be observed on the histogram if higher order bit planes are set to '0'?
c) Obtain the Haar Transform matrix for $N = 8$. **[WBUT 2010]**

Answer:

a) Removing the low order bit planes would mean the loss of some high frequency details. Furthermore, the image histogram will be more sparse as compared with the all 8-bit plane case. This is because, there will be no component representing intermediate pixel values such as 1, 2, 3, 4, 5, 6, 7 and 9, 10, 11, 12, 13, 14, 15 etc. Instead there will be 0 and 8 and 16 etc. This would cause the height some of the remaining histogram peaks to increase in general. Typically, less variability in gray level values will reduce contrast.

b) Removing the high order bit planes would mean the loss of some very important DC components away from the image.
The meaning of this is that the image is much darker and a lot of the low frequency components will be lost.
The most visible effect would be significant darkening of the image. For example, dropping the highest bit would limit the highest level in an 8-bit image to be 127. Because the no of pixels would remain constant, the height of some of histogram peaks would increase. The general shape of the histogram would now be taller and narrower, with no histogram components being located past 127.

c) The family of N Haar functions $h_k(t)$, $(k = 0, \ldots\ldots, N-1)$ are defined on the interval $0 \le t \le 1$. The shape of the specific function $h_k(t)$ of a given index k depends on two parameters pt and q: $k = 2^p + q - 1$.
For any value of $k \ge 0$, p and q are uniquely determined so that 2^p is the largest power of 2 contained in $k\left(2^p < k\right)$ and $q-1$ is the remainder $q-1 = k-2^p$. For example, when $N = 16$. The index k with the corresponding p and q are shown below:

k	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
p	0	0	1	1	2	2	2	2	3	3	3	3	3	3	3	3
q	0	1	1	2	1	2	3	4	1	2	3	4	5	6	7	8

Now, the Haar functions can be defined recursively as:

- When $k = 0$, the Haar function is defined as a constant $h_0(t) = \frac{1}{\sqrt{n}}$.
- When $k > 0$, the Haar function is defined by

$$h_k(t) = \frac{1}{\sqrt{N}} \begin{cases} 2^{p/2}; & (q-1)/2^p \le t < (q-0.5)/2^p \\ -2^{p/2}; & (q-0.5)/2^p \le t < q/2^p \\ 0; & \text{otherwise} \end{cases}$$

from the definition, it can be seen that p determines the amplitude and width of the non-zero part of the function, while q determines the position of the non-zero part of the function.

The N Haar functions can be sampled at $t = \frac{m}{N}$, where $m = 0, \ldots\ldots, N-1$ to form an N by N matrix.

For discrete Haar transform,

For example when $N = 2$, we have

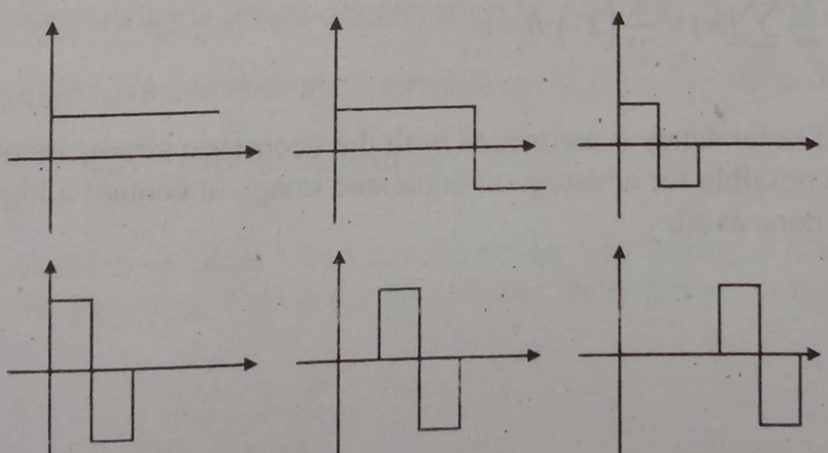
$$H_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

when $N = 4$, we have

$$H_4 = \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ \sqrt{2} & -\sqrt{2} & 0 & 0 \\ 0 & 0 & \sqrt{2} & -\sqrt{2} \end{bmatrix}$$

When $N = 8$,

$$H_8 = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ \sqrt{2} & \sqrt{2} & \sqrt{2} & \sqrt{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sqrt{2} & \sqrt{2} & \sqrt{2} & \sqrt{2} \\ 2 & -2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & -2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & -2 \end{bmatrix}$$



We see that all Haar functions $(h_k(t))$, $k > 0$ contains a single prototype shape composed of a square wave and its negative version, and the parameters (i) p specifies the magnitude and width of the shape (ii) q specifies the position (or shift) of the shape.

2. a) Can variable length coding procedure be used to compress a histogram equalized image with two gray levels?
b) Can such an image contain inter-pixel redundancies that could be exploited for data compression?
c) Find a set of code words and word length using Huffman coding scheme for a set of input gray levels with probabilities as given below: [WBUT 2010]

Symbol:	S1	S2	S3	S4	S5	S6	S7	S8
Probability:	0.02	0.15	0.03	0.15	0.05	0.2	0.1	0.3

Answer:
a) A histogram equalized image has an intensity distribution which is uniform. That is, all intensities are equally probable.

So, $$L_{avg} = \frac{1}{2^n}\sum_{k=0}^{2^n-1}\ell(r_k) \qquad \text{.... (1)}$$

∵ Average no.-of bits required to represent each pixel is

$$L_{avg} = \sum_{R=0}^{L-1}\ell(r_k)P_r(r_k) \qquad \text{.... (2)}$$

$P_r(r_k)$ is the probability of each r_k $\ell(r_k)$ is the no. of bits used to represent each value of r_k.
Eqn. (1) is derived from (2)]

where $\frac{1}{2^n}$ is the probability of occurrence of any intensity. Since all intensities are equally probably there is no advantage to assigning any particular intensity fewer bits than any other. Thus, we assign each the fewest possible bits required to cover the 2^n levels. This, of course is n bits and L_{avg} becomes n bits also

$$L_{avg} = \frac{1}{2^n}\sum_{k=0}^{2^n-1}(n) = \frac{1}{2^n}(2^n)\cdot n = n$$

b) Since spatial redundancy is associated with the geometric arrangement of intensities in the image, it is possible for a histogram equalized image to contain a high level of spatial redundancy or none at all.

c)

Symbols	Prob.	Code	Stage	Stage	Stage	Stage	Stage	Stage
S_8	0.3	00	0.3(00)	0.3(00)	0.3(00)	0.3(00)	0.4(1)	0.6(0)
S_6	0.2	10	0.2(10)	0.2(10)	0.2(10)	0.3(01)	0.3(00)	0.4(1)
S_4	0.15	010	0.15(010)	0.15(010)	0.2(11)	0.2(10)	0.3(01)	
S_2	0.15	011	0.15(011)	0.15(011)	0.15(010)	0.2(11)		
S_7	0.1	110	0.1(110)	0.1(110)	0.15(011)			
S_5	0.05 (1100)	1100	0.05(1100)	0.1(111)				
S_3	0.03 (1110)	1110	0.05(1101)					
S_1	0.02 (1111)	1111						

In the table source symbols are sorted as per their probability from high to low value. The least probabilities 0.03 and 0.02 are merged to give another compound symbol of 0.05 probability and again it is sorted from high to low value. This is repeated till only two probabilities 0.6 and 0.4 remain in the foremost column. The next step is to code each reduced source starting with the shortest source and following its trail back to its original source symbol. The minimum length binary codes for two source symbols are obviously 0 and 1. The Average length of the code is calculated as

$$L_{\text{avg}} = 0.3*2+0.2*2+0.15*3+0.15\times3+0.1*3+0.1*3+0.05\times4$$
$$+0.05*4+0.03*5+0.02*5$$

3. a) Under what condition the does Butterworth low-pass filter become an ideal low-pass filter?
b) Show that a high-pass filtered image in frequency domain can be obtained by using the method of subtracting a low-pass filtered image from the original.
c) An image is blurred by uniform acceleration in x direction. if the image is at rest at time $t=0$ and accelerate with an acceleration $X_0(t)=\frac{at^2}{2}$ for a time T. Find the transfer function $H(u,v)$ to reconstruct the image. **[WBUT 2010]**

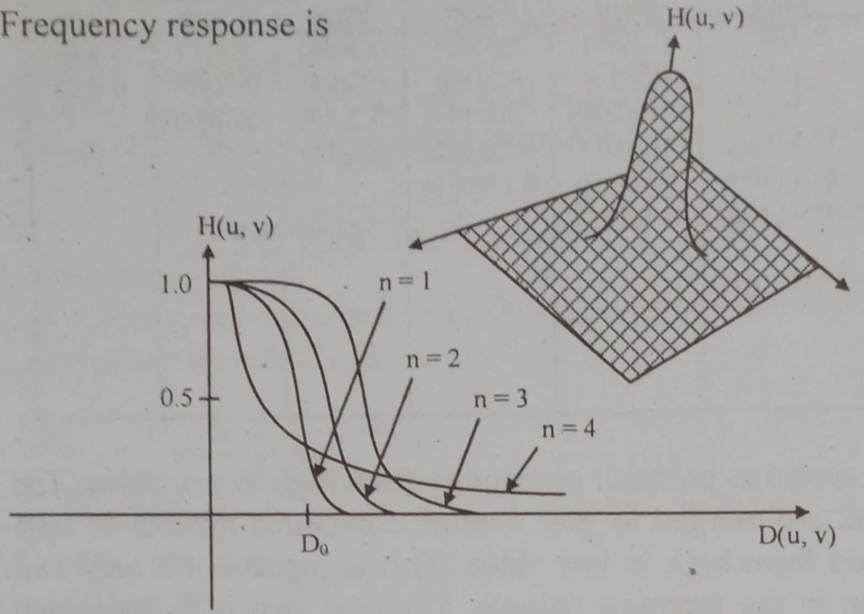
Answer:

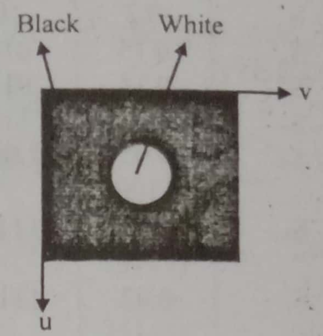
a) The ringing problem of ideal filters can be avoided by having a smoother transition around the cut off frequency. This is the idea behind the Butterworth filter. Its frequency response is given by

$$H(u,v)=\frac{1}{1+\left[D(u,v)/D_0\right]^{2n}}$$

where n is a free parameter, called the order of the filter. The larger the order is, the sharper the transition around D_0, and the more the Butterworth filter resembles the ideal filter.

At the cutoff frequency, $D(u, v) = D_0$, the response is exactly half of its maximal value.
Frequency response is





b) As the opposite of low pass filtering for image smoothing and noise reduction, high pass filtering can sharpen the image, thereby enhancing and emphasizing the detailed information (high special frequency components) in the image.
High pass filtering can be carried out by subtracting low-pass filtered image from its original version, which can be considered as all-pass filtered by a delta function kernel.

$$W_{ap} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

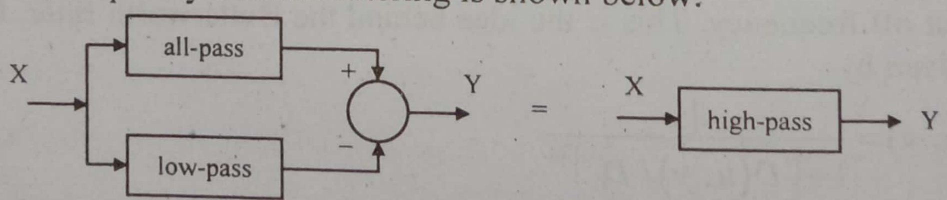
When an image is convolved with this, delta function is not changed

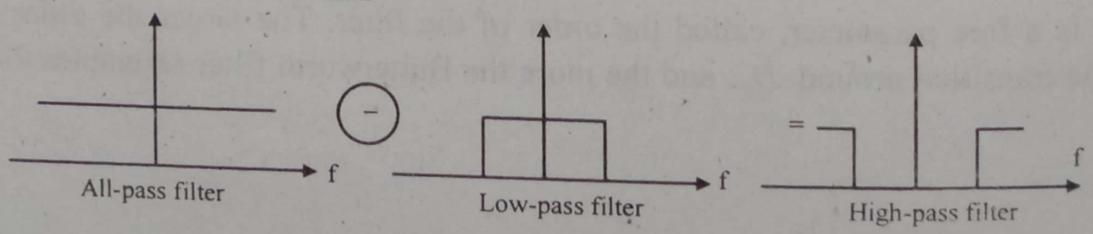
$$W_{ap} * I_0 = I_0$$

As convolution is a linear operator, we have

$$I_{hp} = I_{ap} - I_{\ell p} = W_{ap} * I_0 - W_{\ell p} * I_0 = \left(W_{ap} - W_{\ell p}\right) * I_0 = W_{hp} * I_0$$

where $W_{hp} = W_{ap} - W_{\ell p}$ is the high-pass kernel corresponding to the low-pass kernel $W_{\ell p}$.
Equivalently the frequency domain filtering is shown below:





We can therefore obtain a high-pass filtering kernel corresponding to each of the low-pass filter kernels by subtracting the low-pass kernel from the all-pass kernel. The resulting kernels are various forms of high-pass filtering kernels, also called the Laplace operators.

$$W_{hp} = W_{ap} - W_{lp} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} - \frac{1}{6}\begin{bmatrix} 0 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \frac{1}{6}\begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$

$$W_{hp} = W_{ap} - W_{lp} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} - \frac{1}{9}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} = \frac{1}{9}\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

$$W_{hp} = W_{ap} - W_{lp} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} - \frac{1}{16}\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} = \frac{1}{16}\begin{bmatrix} -1 & -2 & -1 \\ -2 & 12 & -2 \\ -1 & -2 & -1 \end{bmatrix}$$

The sum of the resulting high-pass filter is always zero. When such a high-pass kernel is convolved with a region of an image where all pixels have same gray level, the result is zero i.e., zero spatial frequency component is totally suppressed by the high-pass filter. Similar band-pass filter can be obtained by finding the difference between two low-pass filters of different cut-off frequencies.

c) Suppose the image $f(x, y)$ has a linear motion $x_0(t)$ and $y_0(t)$ are the objects component motion in x and y-direction respectively and the exposure time is T. Without considering other factors, the captured motion-blurred image $g(x, y)$ is

$$g(x, y) = \int_0^T f\left[\left(x - x_0(t), y - y_0(t)\right)\right] dt$$

Transform the above equation with Fourier transform, the above equation can be expressed as

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

$$= \int_0^T \left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f\left[x - x_0(t), y - y_0(t)\right] e^{-j2\pi(ux+vy)} dx \cdot dy\right] dt$$

Exchange integral order and apply properties of Fourier transform, the above equation can be expressed as

$$G(u, v) = F(u, v) \int_0^T e^{-j2\pi\left[ux_0(t)+vy_0(t)\right]} dt \quad \text{.... (1)}$$

Define $H(u, v) = \int_0^T e^{-j2\pi\left[ux_0(t)+vy_0(t)\right]} dt$ .... (2)

Eqn. (1) can be expressed as

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

where $H(u, v)$ is called transfer function of motion blur. If properties of component motions $x(t)$ and $y(t)$ are known, the transfer function can be calculated with Eqn. (2). Now, according to the question, image only has a uniform rest at $t = 0$ and accelerate with an acceleration $x_0(t) = \dfrac{at^2}{2}$ for a time T, then the transfer function $H(u, v)$ will be

$$H(u, v) = \int_0^T e^{-j2\pi x_0(t)} = \int_0^T e^{-j2\pi}\frac{at^2}{2} = a\int_0^T e^{-j\pi t^2}\, dt$$

4. What is Histogram of an image and why is it used for image processing? What do you mean by Histogram equalization? **[WBUT 2011]**

OR,

What do you understand by the terms 'histogram' and 'histogram equalization'? **[WBUT 2016]**

Answer:

The histogram of a digital image with gray levels in the range $[0, L-1]$ is a discrete function $h(r_k) = (n_k)$ where r_k is the k th gray level and n_k is the number of pixels in the image having gray level r_k.

Histogram can be used effectively for image enhancement. It is simple to calculate in software and also level themselves to economic hardware implementations, thus making them a popular tool for real time image processing.

Histogram equalization is a technique for adjusting image intensities to enhance contrast. Let f be a given image represented as a m_r by m_c matrix of integer pixel intensities ranging from 0 to $L-1$. L is the number of possible intensities values, often 256. Let p denote the normalized histogram of f with a bin for each possible intensity.

So, $$p_n = \frac{\text{No. of pixels with intensity } n}{\text{Total no. of pixels } (n = 0, 1, \ldots\ldots, L-1)}$$

The histogram equalized image g will be defined by

$$g_{i,j} = \text{floor}(L-1)\sum_{n=0}^{f_{i,j}} p_n$$

5. a) What do you mean by image negative? Explain. **[WBUT 2012, 2013, 2017]**
b) Explain Intensity slicing with example. **[WBUT 2012, 2013, 2017, 2019]**
c) Why do we need Log Transformation in dynamic range compression? **[WBUT 2012, 2013, 2017, 2019]**

Answer:

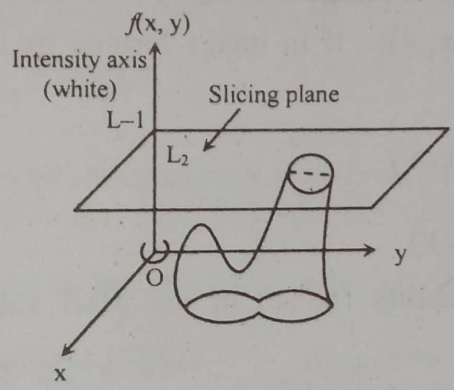
a) The negative of an image with gray level in the range [0, h, –1] is obtained by using the negative transformation. The expression of the transformation is

$$S = L - 1 - r \quad (\because\ S = \text{processed image, } r = \text{Original image})$$

L is the no of levels.

Reverting the intensity levels of an image in this manner produces the equivalent of a photographic negative. This type of processing is practically suited for enhancing white or gray details embedded in dark regions of an image specially when the black area are dominant in size.

b) The technique of intensity slicing is one of the simplest examples of pseudo color image processing. If an image is interpreted as a 3D-function, the method can be viewed as one of the placing planes parallel to the coordinate plane of the image; each plane then slices the function in the area of intersection. Fig. shows an example of using a plane at $f(x, y) = l_2$ to slice the image function into two levels.



If a different color is assigned to each side of the plane in figure, any pixel whose intensity level is above the plane will be coded with one color, and any pixel below the plane will be the coded with the other. Level that lie on the plane itself may be arbitrarily assigned one of the two colors. The result is a two color image whose relative appearance can be controlled by moving the slicing plane up and down the intensity axis.

c) The general form of transformation is $S = c\log(1+r)$; where c is a constant and $r \geq 0$. This transformation maps a narrow range of gray level values in the input image into a wider range of output gray levels. The opposite is true for higher value of input levels. We would use this transformation to expand the values of dark pixels in an image white compressing the higher level values. The opposite is true for inverse log transformation. The log transformation function has an important characteristics that it compresses the dynamic range of images with large variations in pixel values i.e., Fourier spectrum.

6. a) What is image averaging? **[WBUT 2012, 2019]**

b) Discuss Histogram characteristics for dark, bright, low contrast images. **[WBUT 2012]**

c) Equalize the following histogram. Show the histogram before and after equalization. **[WBUT 2012]**

Gray Level	0	1	2	3	4	5	6	7
Number of pixels	10	20	12	8	0	0	0	0

Answer:

a) Consider a noisy image $g(x, y)$ formed by the addition of noise $n(x, y)$ to the original image $f(x, y)$

$$g(x, y) = f(x, y) + n(x, y).$$

Assuming that at every point of coordinate (x, y), the noise is uncorrelated and has zero average value. The objective of **image averaging** is to reduce the noise content by adding a set of noise image $\{g_i(x, y)\}$. If in image formed by image averaging k different noisy images, then

$$\bar{g}(x, y) = \frac{1}{k}\sum_{i=1}^{k} g_i(x, y)$$

$$E\{g(x, y)\} = f(x, y)$$

As k increases, the variability (noise) of the pixel value at each location (x, y) decreases.

$E\{g(x, y)\} = f(x, y)$ means that $g(x, y)$ approaches $f(x, y)$ as the number of noisy image used in the averaging process increases.

Image averaging is important in various applications such as in the field of astronomy where the images are low light levels.

b) We have seen in the dark image that the components of the histogram are concentrated on the low (dark) side of the gray scale. Similarly, the components of the histogram of the bright image are biased forward the high side of the gray scale. An image with low contrast has a histogram that will be narrow and will be centered toward the middle of the gray scale. For a monochrome image this implies a dull, washed-out gray look. Finally, we see that the components of the histogram in the high-contrast image cover a broad range of the gray-scale and further that the distribution of pixels is not too far from uniform with very few vertical lines being much higher than the others. Intuitively, it is reasonable to conclude that an image whose pixels tend to occupy the entire range of possibly gray levels, and in addition, tend to be distributed uniformly, will have an appearance of high contrast and will exhibit a large variety of gray tones.

c)

Gray level	No. of pixel
0	10
1	20
2	12
3	8
4	0
5	0
6	0
7	0

Step 1: Create a running sum of the histogram values. This means that the first value is 10, the second is $10+20=30$, next $10+20+12=42$ and then $10+20+12+8=50$ and so on. Here, we get 10, 30, 42, 50, 50, 50, 50, 50

Step 2: Normalize by dividing the total no. of pixels. The total no. of pixels $=50$ so, we get, $\frac{10}{50}, \frac{20}{50}, \frac{12}{50}, \frac{8}{50}, 0, 0, 0$ and 1.

Step 3: Multiply these values by the maximum gray-level values, in this case 7 and then round the result to the closest integer. After this, we obtain 1, 2, 1, 1, 0, 0, 0 and 0.

Step 4: Map the original values to the result from step 3 by a one-to-one correspondence i.e.,

Original gray level value	Histogram equalized value
0	1
1	2
2	1
3	1
4	0
5	0
6	0
7	0

7. a) What is Histogram of an image and why processing? What do you mean by Histogram specification?
b) What is the difference between spatial domain filtering and frequency domain filtering? Describe edge detection operator and contrast stretching of an image.
c) Briefly describe the smoothing linear spatial filtering. What is bit plane slicing method? **[WBUT 2015]**

Answer:

a) 1st Part:
Histogram of an image, like other histograms also shows frequency. But an image histogram, shows frequency of pixels intensity values. In an image histogram, the x-axis shows the gray level intensities and y-axis shows the frequency of these intensities.

2nd Part:
Histograms has many uses in image processing. The first use as it has also been discussed is the analysis of the image. We can predict about an image by just looking at its histogram. The second use of histogram is for brightness purposes, and also adjusting contrast of an image. Another important use of histogram is to equalize an image and it has wide use in thresholding.

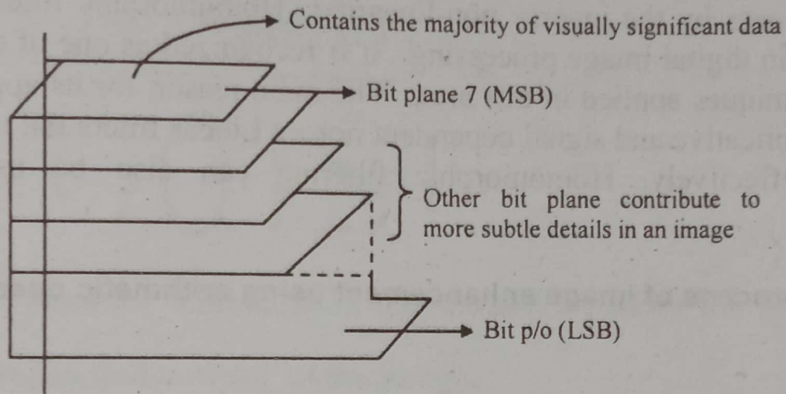
3rd Part:
Histogram equalization automatically determines a transformation function that seeks to produce an output image that has a uniform histogram when automatic enhancement is desired, this is a good approach because the results from this technique are predictable and the method is simple to implement. 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 specification.

b) 1st part:
In spatial domain, we deal with images as it is. The value of the pixels of the image change w.r.t. scene. Whereas in frequency domain, we deal with the rate at which the pixel values are changing in spatial domain.
In simple spatial domain, we directly deal with the image matrix, whereas in frequency domain we deal an image like,
input image → frequency distribution → processing → Inverse transformation → output image.

2nd Part:
Edge detection refers to the boundaries where there is a sharp change in the intensity or brightness of the image. Hence the obtained boundary marks the edge or the contour of the desired object. This way the object can be segmented from the image by the detection of its edges. The final output that is received by applying edge detection algorithm is a binary image. There are many edge detection operators that can be used in gradient based methods such as **sobel operator**, **canny operator**, **Laplace operator** etc. Generally the method of edge detection requires a balance between the proper detection of the edges and the level of the noise present.
Contrast stretching (often called normalization) is a simple image enhancement technique that attempts to improve the contrast in an image by stretching the range of intensity values it contains to span a desired range of values, e.g., the full range of pixel values that the image type concerned allows. It differs from the more sophisticated histogram equalization in that it can only apply a linear scaling function to the image pixel values. As a result the enhancement is less harsh.

c) 1st part:
Smoothing linear spatial filters are used for blurring and for noise reduction. Blurring is used in pre-processing steps such as removal of small details from an image prior to object extraction and bridging of small gaps in lines or curves. Noise reduction can be accomplished by blurring. There are two ways of smoothing spatial filters i.e., smoothing linear filters and order statistics filters.

2nd Part:
Instead of highlighting gray level images, highlighting the contribution made to total images appearance by specific bits might be desired. Suppose that each pixel in an image is represented by 8 bits. Imagine the image is composed of 8, 1-bit planes ranging from bit plane 1–0 (LSB) to bit plane 7 (MSB). In terms of 8bit bytes, plane 0 contains all lowest order bits in the bytes comprising the pixels in the image and plane 7 contains all high order bits.



Separating a digital image into its bit planes is useful for analyzing the relative importance played by each bit of the image, implying, it determines the adequacy of number of bits used to quantize each pixel useful for image compression.
In terms of bit plane extraction for a 8-bit image, it is seen that binary image for bit plane 7 is obtained by proceeding the input image with a thresholding gray-level transformation function that maps all levels between 0 and 127 to one level (i.e., 0) and maps all levels from 129 to 253 to another (i.e., 255).

8. a) Discuss about spatial resolution and gray – level resolution. [WBUT 2015]
b) What is Homomorphic filtering? Mention the domain of application of this filter.

Answer:
a) Sampling is the principal factor determining the spatial resolution of an image. Basically spatial resolution is the smallest discernible detail in an image.
As an example, suppose we construct a chart with vertical lines of width W and with space between the lines also having width W. A line pair consists of one such line and its adjacent space. Thus width of line pair is 2W and there are $\frac{1}{2W}$ line-pairs per unit distance. A widely used definition of resolution is simply the smallest number of discernible lines pairs per unit distance; i.e. 100 lines pairs/mm. Gray level resolution refers to the smallest discernible change in gray level. The measurement of discernible changes in gray level is a highly subjective process. Due to hardware constraints, the

number of gray levels is usually an integer power of two. The most common value is 8 bits. It can vary depending on applications.

b) In many applications are combined in a rather complicated way. Convolved signals are encountered in seismic signal processing, digital speech processing, digital echo removal and digital image restoration. Signals combined in a non-linear way are encountered in digital signal processing for communication systems and in digital image filtering. Classical linear processing techniques are not so useful in those cases because the superposition property does not hold any more. Therefore a special class of filters has been developed for the processing of convolved and non-linearly related signals. They are called homorphic filters. Their basic characteristics is that they use non-linearities (mainly the logarithm) to transform convolved or non-linearly related signals to additive signals and then to process them by linear filters. The output of the linear filter is transformed afterwords by the inverse non-linearity. Homomorphic filtering has found many applications in digital image processing. It is recognized as one of the oldest non-linear filtering techniques applied in this area. The main reason for its application is the need to filter multiplicative and signal dependent noise. Linear filters fail to remove such types of noise effectively. Homomorphic filtering can also be used in image enhancement.

9. a) Explain the process of image enhancement using arithmetic operator. **[WBUT 2016]**

Answer:
Arithmetic operation like subtraction, addition, multiplication and division. It may use multiplication to implement gray-level masks.
Image subtraction of a mask from an image is given as

$$g(x,y)=f(x,y)-h(x,y)$$

Where $f(x,y)$ is the image and $h(x,y)$ is the image mask.
Image subtraction is used in medical imaging e.g. mask mode radiography.
In image averaging, a noisy image can be modeled as

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

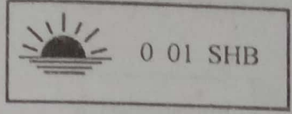
where $f(x,y)$ is the original image and $\eta(x,y)$ is the noise.
Averaging k different noisy images

$$\bar{g}(x,y)=\frac{1}{M}\sum_{i=1}^{M}g_i(x,y)$$

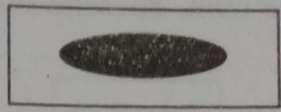
As k increases, the variability of the pixels values at each location decreases. This means that $g(x,y)$ approaches $f(x,y)$ as the number of noisy images used in the averaging process increases. Registering of the images is necessary to avoid blurring in the output image.
Image arithmetic applies one of the standard arithmetic operations or a logical operator to two or more images. The operators are applied in a pixel-by-pixel way. Logical operators

are often used to combine two (mostly binary) images. In the case of integer images, the logical operators is normally applied in a bitwise way.

Example:
Arithmetic and logical operations

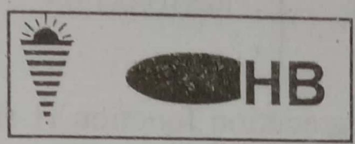


a

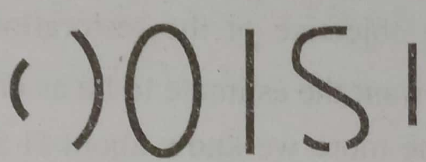


b

NOT (a)



a.b



a + b

b) What is the need for interpolation technique? Consider the image $F = \begin{pmatrix} 2 & 2 \\ 3 & 2 \end{pmatrix}$. Calculate the mean and entropy of the image. **[WBUT 2016]**

OR,

What is the need for interpolation technique? Consider a 2 × 2 image given below: Calculate the mean and entropy of the image. **[WBUT 2018]**

2	2
3	2

Answer:

Consider the image. $F = \begin{pmatrix} 2 & 2 \\ 3 & 2 \end{pmatrix}$ gray level are 2 and 3

Probability of occurrence of gray level $2 = \frac{3}{4}$

and gray level $3 = \frac{1}{4}$

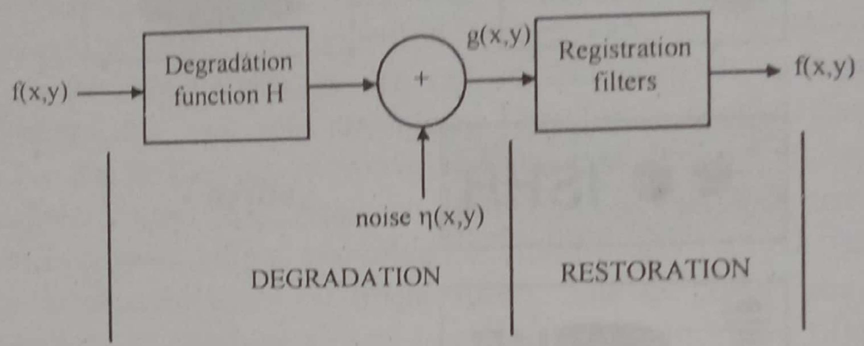
Entropy $H = -\sum P_i \log_2 P_i$ bits/Pixel

$$= -\left[\frac{3}{4}\log_2\left[\frac{3}{4}\right] + \frac{1}{4}\log_2\left[\frac{1}{4}\right]\right] = 2.312$$

Mean $= \frac{1}{4}(2+2+3+2) = \frac{1}{4} \times 9 = 2.25$

c) Explain the model of image restoration process in presence of noise. **[WBUT 2016]**

Answer:
Fig. below shows the degradation process is modeled as a degrade function that, together with an additive noise term, operates on an input image $f(x,y)$ to produce a degraded image $g(x,y)$.



Given $g(x,y)$, some knowledge about the degradation function H and some knowledge about the additive noise term $\eta(x,y)$, the objective of the restoration is to obtain an estimate $f(x,y)$ of the original image. We want the estimate to be as close as possible to the original input image and, in general, the more we know about H and η, the closer $f(x,y)$ will be to $f(x,y)$. The restoration approach used throughout is based on various types of image restoration filters. If H is a linear, position-invariant process, then the degraded image is given in the spatial domain by

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

where $h(x,y)$ is the spatial representation of the degraded function and as it the symbol '*' indicates convolution. We know that convolution in the spatial domain is analogous to the multiplication in the frequency domain, so we write the model in an equivalent frequency domain representation. i.e.

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

where the term in capital letters are the Fourier transforms. The above two equations are the bases for most of the restoration model.

10. a) Consider a 3 × 3 image given below and show that Laplacian mask is an isotropic filter. **[WBUT 2018]**

1	2	1
2	4	2
4	8	4

Answer:
3 × 3 image is given below:

1	2	1
2	4	2
4	8	4

The laplacian is an linear operator and it forms an isotropic filter. The laplacian highlights grey level discontinuities and will output an image with a black background and grey lines where the edges of an object in the image are. In order to get a sharpened image, typically, the resulting laplacian filtered image (or a weighted version of it) is added to the original image. In practice, a mask is typically used that will carry out both steps at once.

The laplacian is given by

$$\nabla^2 f = \frac{\delta^2 f}{\delta x^2} + \frac{\delta^2 f}{\delta y^2}$$

The discrete form is

$$\nabla^2 f = [f(x+1,y) + f(x-1,y) + f(x,y+1) + f(x,y-1)] - 4f(x,y)$$

Combining the additional laplacian into the image if

$$g(x,y) = f(x,y) + \nabla^2 f(x,y)$$

Then

$$f(x,y) = f(x,y) - [f(x+1,y) + f(x-1,y) + f(x,y+1) + f(x,y-1)] + 4f(x,y)$$

$$g(x,y) = 5f(x,y) - [f(x+1,y) + f(x-1,y) + f(x,y+1) + f(x,y-1)]$$

The mask

$$\begin{matrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 4 & 8 & 4 \end{matrix}$$

Consider the original one i.e.

$$\begin{matrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{matrix}$$

By multiplying both

$$1\times1 + 2\times1 + 4\times1 + 2\times1 + 4*(-8) + 8\times1 + 1\times1 + 2\times1 + 4\times1 = -8$$

Similarly, by undercharging the new mask will be

$$\begin{matrix} 1 & 2 & 4 \\ 2 & 4 & 8 \\ 1 & 2 & 4 \end{matrix}$$

Again by multiplying the original one, we get

$$1\times1 + (2\times1) + 1\times1 + 2\times1 + 4\times(-8) + 2\times1 + 4\times1 + 8\times1 + 4\times1$$

$$= 1 + 2 + 1 + 2 - 32 + 2 + 4 + 8 + 4 = -8$$

So, it is isotropic.

b) Given below is a 3 × 3 image. What will be the value of the centre pixel when this image is passed through a
(i) Arithmetic mean filter
(ii) Harmonic mean filter
(iii) Max-filter
(iv) Min-filter **[WBUT 2018]**

1	7	5
6	2	3
1	4	2

Answer:

(i) Arithmetic Mean Filter

The arithmetic mean filter is defined as the average of all pixels within a local region of an image. It is defined as

$$\bar{x} = \frac{1}{n}(x_1 + x_2 + \dots + x_n)$$

i.e. $$\bar{x} = \frac{1}{9}(1+7+5+6+2+3+1+4+2)$$

$$= \frac{31}{9} = 3.44$$

∴ Center pixel will be = 3

(ii) Harmonic Mean Filter:

It is defined as

$$H = \frac{n}{\frac{1}{x_1}+\frac{1}{x_2}+\frac{1}{x_3}+\dots+\frac{1}{x_n}} = \frac{9}{\frac{1}{1}+\frac{1}{7}+\frac{1}{5}+\frac{1}{6}+\frac{1}{2}+\frac{1}{3}+\frac{1}{1}+\frac{1}{4}+\frac{1}{2}} = \frac{9}{4.08} = 2.2$$

Center pixel will be = 2

(iii) Max Filter:

1 7 5 6 2 3 1 4 2

After arranging,

1 1 2 2 <u>3</u> 4 5 6 7

Max value = 7
Center pixel = 3

(iv) Min-Filter:

1 7 5 6 3 1 4 2

After arranging,

1 1 2 2 3 4 5 6 7

Min value = 1
Center pixel = 3

c) Contrast and compare the spatial domain filtering and frequency domain filtering. **[WBUT 2018]**

Answer:

Refer to Question No. 7(a) (1st Part) of Long Answer Type Questions.

11. Describe an algorithm for image sharpening using DFT and IDFT. Identify the parameters affecting the sharpening operation in your algorithm. **[WBUT 2019]**

Answer:

1st Part:

1. Multiply the input image by $(-1)^{x+y}$ to center the transform to $u = \frac{M}{2}$ and $v = \frac{N}{2}$
2. Compute $F(U,V)$ the DFT of the image form (1)
3. Multiply $F(U,V)$ by a filter function $H(U,V)$
4. Compute the inverse DFT (IDFT) of the result obtained in step (3).
5. Obtain the real part of the result obtained in step (4)
6. Multiply the result in (5) by $(-1)^{x+y}$ to cancel the multiplication of the input image.

2nd Part:

Image sharpening can be achieved in the frequency domain by using a high pass filter. The high pass filter attenuates the low frequency component without disturbing the high frequency information in the Fourier transform. High pass filters perform the reverse operation of the low pass filters. Thus, the transfer function of the high pass filters can be obtained

$$H_{np}(U,V) = 1 - H_{\ell p}(U,V)$$

12. Write short notes on the following:

a) Filtering in frequency domain **[WBUT 2012, 2014]**
b) Histogram specification **[WBUT 2013]**
c) Wiener filtering **[WBUT 2013]**
d) Point processing **[WBUT 2013]**
e) Spatial filtering **[WBUT 2013, 2014]**
f) Local enhancement **[WBUT 2015]**
g) Image sharpening filters **[WBUT 2015]**
h) Smoothing linear spatial filtering **[WBUT 2017]**
i) Salt and Pepper noise **[WBUT 2019]**

Answer:

a) Filtering in frequency domain:

The reason for doing the filtering in the frequency domain is generally because it is computationally faster to perform two 2D Fourier transform and a filter multiply then to perform a convolution in the image (spatial) domain. This is particularly so as the filter size increases.

Filtering in the frequency domain consists of modifying the Fourier transform of an image and then computing the inverse transform to obtain the processed result. Thus, given a digital image $f(x,y)$ of size $M \times N$, the basic filtering equation has the form $g(x,y) = \Im^{-1}[H(u,v)F(u,v)]$ where $\Im^{-1}$ is the IDFT, $F(u,v)$ is the DFT of the input image, $f(x,y)$ and $H(u,v)$ is a filter function and $g(x,y)$ is the filtered image.

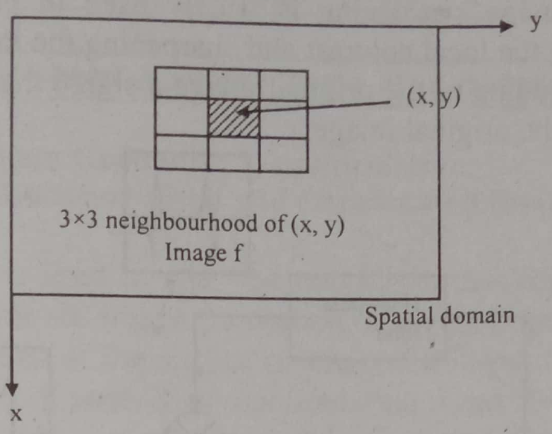
b) Histogram specification:
Histogram specification has been successfully used in digital image processing over the years. Mainly used as an image enhancement technique, methods such as histogram equalization cam yield good contrast with almost no effort in terms of inputs to the algorithm or the computational time required. More elaborate histograms can take on problems faced by histogram equalization at the expense of having to define the final histograms in innovative ways that may require some extra processing time but are nevertheless fast enough to be considered for real time applications.

c) Wiener filtering:
The inverse filtering is a restoration technique for de-convolution i.e., when the image is blurred by a known lowpass filter, it is possible to recover the image by inverse filtering or generalized inverse filtering. However inverse filtering is very sensitive to additive noise. The approach of reducing one degradation at a time allows us to develop a restoration algorithm for each type of degradation and simply combine them. The Wiener filtering executes an optimal tradeoff between inverse filtering and noise smoothing. It removes the additive noise and inverts the blurring simultaneously. The Wiener filtering is optional in terms of the mean square error. In other words, it minimizes the overall mean square error in the process of inverse filtering and noise smoothing. The Wiener filtering is a linear estimation of the original image. The approach is based on a stochastic framework. To implement the Weiner filter in practice we have to estimate the power spectra of the original image and the additive noise.

d) Point processing:
Point processing operations fakes the form $S = T(r)$ where S refers the proposed image pixel value and r refers to the original image pixel value. Point processing are processing that can be done based on a single image pixel. The output of the process is a single value. Usually, this output value replaces the input pixel in the image. The process is carried out pixel-wise. Examples: A number of examples are Gray-level mapping, histograms, segmentation using thresh-oldness.

e) Spatial filtering:
Spatial filter is an image **operation** where each pixel value is changed by a function of the intensities of pixels in a neighbourhood.
So, spatial filter consists of (1) a neighbourhood (typically a small rectangle) and a predefined operation that is performed on the image pixels encompassed by the neighbourhood. Filtering creates a new pixel with coordinates equal to the coordinates of the center of the neighbourhood and whose value is the result of the filtering operation. A processed (filtered) image is generated as the centered of the filter visits each pixel in

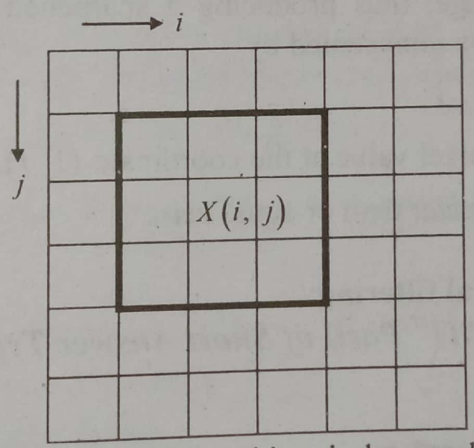
the input image. If the operation performed on the image pixel is linear then the filter is called linear spatial filter, otherwise non-linear.



f) Local enhancement:

Sometimes we may need to enhance details over small areas in an image, which is called local enhancement.

We know that local histogram equalization is the extension of global histogram equalization. This method is based on block processing. A small sized window is selected from the image for enhancement. This window slides through every pixel of the image in order to enhance that image as shown below:



Only those pixels are considered that fall in this window and enhancement is performed only for the center pixel of this window. Unlike global contrast enhancement this method includes local information of the image.

g) Image sharpening filters:

Human perception is highly sensitive to edges and fine details of an image, and since they are composed primarily by high frequency components, the visual quality of an image can be enormously degraded if the high frequencies are attenuated or completely removed. In contrast, enhancing the high frequency components of an image leads to an improvement in the visual quality.

Image sharpening refers to any enhancement technique that highlights edges and fine details in an image. Image sharpening is widely used in printing and photographic industries for increasing the local contrast and sharpening the images. In principle, image sharpening consists of adding to the original image a signal that is proportional to a high pass filtered version of the original image.

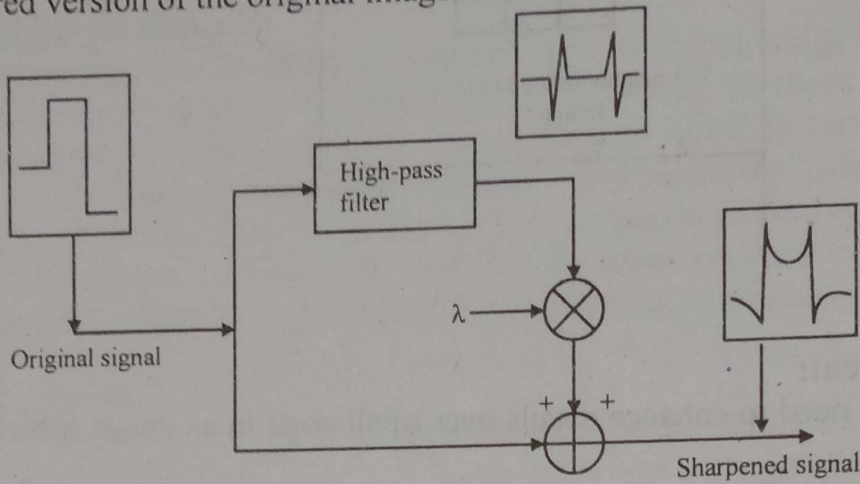


Figure illustrates this procedure, often referred to as unsharp masking on a one-dimensional signal. The original image is first filtered by a high-pass filter that extracts the high frequency components and then a scaled version of the high-pass filter output is added to the original image, thus producing a sharpened image of the original. The sharpening operation can be represented by

$$S_{i,j} = x_{i,j} + \lambda F\left(x_{i,j}\right)$$

where $x_{i,j}$ is the original pixel value at the coordinate (i, j), $F(\cdot)$ is the high-pass filter, λ is a tuning parameter greater than or equal zero.

h) Smoothing linear spatial filtering:
Refer to Question No. 10(1st Part) of Short Answer Type Questions.

i) Salt and Pepper noise:
Refer to Question No. 14(1st Part) of Short Answer Type Questions.

IMAGE RESTORATION

Short Answer Type Questions

1. a) List down the various Geometric Transformation.
b) Difference between Unconstrained and Constrained Restoration. **[WBUT 2009]**

Answer:

a) This transformation is used to alter the coordinate description of image. The basic geometric, transformation are image translation, scaling, image rotation. The need for transformation is that most of the signals or images are time domain i.e. signals can be measured with a function of time. This representation is not always best. For most image processing applications any one of the mathematical transformation are applied to the signal or images to obtain further information from that signal.

b) In unconstrained restoration, no a priori knowledge about the noise is assumed. The restoration is carried out in a least square sense of the estimation error. Inverse filtering is a commonly used restoration approach.
In constrained restoration some a priori knowledge of the noise is used to constrained restoration, some a priori knowledge of the noise is used to constrain the least square computation. The restoration is carried out using the method of Lagrange multipliers. Wiener filtering is a statistical method for constrained restoration. To estimate the original image, noise $n = g - Hf$ has to be maximized, where g is the degraded image, H is the system operator, f is the input image.

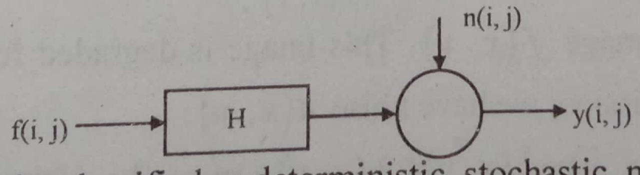
2. Explain the restoration model for continuous function in detail. **[WBUT 2012, 2013, 2014]**

Answer:

The image restoration process model is $y(i,j) = H[f(i,j)] + n(i,j)$

where $y(i,j)$ is the degraded image
$f(i,j)$ is the original image
H is an operator that represents the degradation process.
$n(i,j)$ is the external noise which is assumed to be image independent.

Fig. below is a schematic diagram for a generic degradation process described by the above model.



Restoration method can be classified as deterministic, stochastic, non-blind, blind etc.
Since $y(i,j) = H[f(i,j)] + n(i,j)$
If we ignore the presence of the external noise $n(i,j)$, we get

$$y(i,j) = H[f(i,j)]$$

H is linear of

$$H[k_1 f_1(i,j)+k_2 f_2(i,j)]=k_1 H[f_1(i,j)]+k_2 H[f_2(i,j)]$$

H is position invariant if

$$H[f(i-a,j-b)]=y(i-i,j-b)$$

3. Describe the region growing technique for image segmentation and mention the Problem associate with it. [WBUT 2012, 2017]

OR,

Briefly explain about region growing technique. [WBUT 2014]

Answer:

Region growing segmentation is an approach to examine the neighbouring pixels of the initial "seed points" and determine if the pixels are added to the seed point or not.

Step 1: Selecting a set of one or more starting point (seed) often can be based on the nature of the problem.

Step 2: The region are grown from these seed points to adjacent point depending on a threshold or criteria (8-connected) we make.

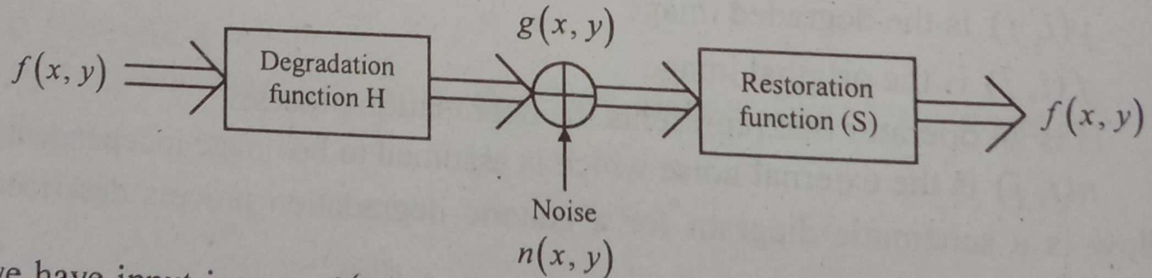
Step 3: Region growth should stop when no more pixels satisfy the criteria for inclusion in that region. Here suitable selection of seed point is important and more information of the image is better. The result of an image after region growing still have points gray-level higher than the threshold but not connected with the object in image.

Long Answer Type Questions

1. Describe the model of image degradation. Give some example of added noise. explain some ways of estimating the degradation function. Describe the estimation of degradation function by mathematical modelling for the degraded image captured while the object is moving in X & Y direction in constant velocity. Briefly state the need of Wiener filter. [WBUT 2009]

Answer:

1st part:



Here, we have input image $f(x, y)$. This image is degraded function $H(x, y)$. Now, during channel transmission, we have noise $h(x, y)$.

Thus, we have $g(x,y)=h(x,y)*f(x,y)+n(x,y)$. Hence, of we have some knowledge about $H(u, v)$ and noise $n(x, y)$, then we get an image $f(x, y)$ that is very close to real image $f(x, y)$.

2nd part:
Some examples of added noise are:
Gaussian noise, Rayleigh noise exponential noise, uniform noise etc.

3rd part:
There are three principal ways to estimate the degraded function for use in image restoration:
1) Observation, 2) Experimentation and 3) Mathematical modelling.
The process of restoring an image by using a degraded function that has been estimated in some ways sometimes called blind convolution, due to the fact that the true degradation function is seldom known completely.

4th part:
Suppose that we assume that image has been blurred by uniform linear motion between the image and the sensor during image acquisition.

a) $f(x,y)$ is a going planner motion $x(t_0)$ and $y(t_0)$ in x and y direction respectively.

So, $$g(x,y)=f\left[x-x_0(t),\ y-y_0(t)\right]$$

b) If shutter opening time is t, then complete image is given by

$$g(x,y)=\int_0^T f\left[x-x_0(t),\ y-y_0(t)\right]dt$$

c) To calculate $G(u,v)$,

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

So, $$G(u,v)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\left[\int_0^T f\left[x-x_0(t),\ y-y_0(t)\right]e^{-j2\pi(ux+vy)}\cdot dx\cdot dy\cdot dt\right]$$

Rearranging the equation

$$G(u,v)=\int_0^T\left[\int_0^{\infty}\int_0^{\infty} f\left[x-x_0(t),\ y-y_0(t)\right]e^{-j2\pi(ux+vy)}\cdot dx\cdot dy\cdot dt\right]$$

$$=\int_0^T F(u,v)e^{-j2\pi}\left[ux_0(t)+vy_0(t)\right]dt$$

$$F(u,v)\int_0^T e^{-j2\pi}\left[ux_0(t)+vy_0(t)\right]dt$$

d) So, we have

$$H(u,v)=\int_0^T e^{-j2\pi}\left[ux_0(t)+vy_0(t)\right]dt$$

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

Thus, we can calculate the value of $H(u,v)$ for an image.

5th part:
The inverse filtering is a restoration technique for de-convolution i.e., when the image is blurred by a known low pass filter, it is possible to recover the image by inverse filtering or generalized inverse filtering. However, inverse filtering is very sensitive to additive noise. the approach of reducing one degradation at a time allows us to develop a restoration algorithm for each type of degradation and simply combine them. The wiener filtering executes an optimal tradeoff between inverse filtering and noise smoothing. The wiener filtering is optimal in terms of the mean square error. In over words, it minimizes the overall mean square error in the process of inverse filtering and noise smoothing. It is a linear estimation of the original image. The approach is based on a stochastic, framework.

2. a) Discuss about threshold coding implementation.
b) Make a comparative study of least square restoration and constrained least square restoration methods. [WBUT 2011]

Answer:
a) Threshold coding is inherently adaptive in the sense that the location of the transform co-efficients retained for each subimage vary from one subimage to another. In fact, threshold coding is the adaptive transform coding approach most often used in practise because of its computational simplicity. The underlying concept is that, for any subimage, the transform co-efficients of largest magnitude make the most significant contribution to reconstructed subimage quality. Because the location of the maximum co-efficients vary from one subimage to another.
There are three basic ways to threshold a transformed subimage: (1) A single global threshold can be applied to all subimages, (2) A different threshold can be used for each subimage, (3) The threshold can be varied as a function of the location of each co-efficient within the subimage.

b) Minimum mean-square error estimation provides a filter that produces optimal restoration of a particular image. This may be obtained by the method of least square. In absence of any knowledge about noise we may treat it as a measurement error. So, to restore a particular image, we minimize the difference between the observed image and the re-degraded estimated image.
The method of constrained least square can be defined as the optimization of some criterion of goodness or quality of image subject to the constraint that residual norm between the image and the re-degraded estimated image be equal to the norm of the noise vector. Constraint least square is a linear image restoration technique but least square restoration is not.
Constraint least square retain the simplicity of a spatially invariant linear filter but require prior knowledge of mean and variance of the noise.
When point-spread function is symmetric, least square restoration does not work properly, to select a solution a constraint least square approach is used.

3. Write short notes the following:
a) Image Restoration model **[WBUT 2016]**
b) Constrained least squares **[WBUT 2011]**
c) Contrast stretching **[WBUT 2012, 2013, 2019]**

Answer:

a) Image Restoration model:

Refer to Question No. 2 of Short Answer Type Questions.

b) Constrained least squares:

The constrained least square filter is another approach for overcoming some of the difficulties of the inverse filter (excessive noise amplification) and of the weiner filter (estimation of the power spectrum of the ideal image), white still retaining the simplicity of a spatially invariant linear filter. But it is required to have a prior knowledge about mean and variance of the noise. The constraint least square algorithm is based on finding a direct solution using a criterion C, which ensures optimal smoothness of the deblurred image. It is a linear image restoration technique in which the smoothness of the restored image is maximized subject to a constraint on the fidelity of the restored image.

c) Contrast stretching:

Contrast stretching is a simple image enhancement technique that attempts to improve the contrast in an image by stretching the range of intensity values it contains to span a desired range of values e.g., The full range of pixel values that the image type concerned allows. It differs from the more sophisticated histogram equalization in that it can only apply a linear scaling function to the image pixel values. As a result the enhancement is less harash.

IMAGE SEGMENTATION

Multiple Choice Type Questions

1. Representation & description almost always follow the output of a [WBUT 2014]
   a) Segmentation stage
   b) Filtering stage
   c) Compression stage
   d) All of these

Answer: (a)

2. Isolated point can be best detected using [WBUT 2014]
   a) 1^{st} derivative
   b) 2^{nd} derivative
   c) 3^{rd} derivative
   d) None of these

Answer: (b)

3. What operation is needed for background removal? [WBUT 2015]
   a) Image multiplication
   b) Image addition
   c) Image subtraction
   d) all of these

Answer: (d)

4. Zero padding of a signal [WBUT 2019]
   a) reduces aliasing
   b) increases frequency
   c) increases time resolution
   d) has no effect

Answer: (c)

Short Answer Type Questions

1. Define Edge detection. [WBUT 2009]

Answer:

Edge detection is the name for a set of mathematical methods which aim at identifying points in a digital image at which the image brightness changes sharply or more formally, has discontinuities. The point at which image brightness changes sharply are typically organised into a set of curved line segments termed edges. Edge detection is a fundamental tool in image processing, computer vision particularly in the areas of feature detection and extraction.

2. What is the difference between local and global shareholding? [WBUT 2013]

Answer:

Refer to Question No. 2(1^{st} Part) of Long Answer Type Questions.

3. Write down the usefulness of segmentation? [WBUT 2014, 2017]

Answer:

Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. More precisely image segmentation is the process of assigning a label to every pixel in an image such that pixel with the same label share certain characteristic. Image segmentation is the process of dividing an image into multiple parts. This is typically used to identify objects or other relevant information in digital images.

4. Consider the image segment **[WBUT 2017]**

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

Let $v = \{0,1\}$. Compute the (i) shortest 4-adjacency distance and (ii) m-adjacency distance between the pixels p and q.

Answer:

When v = {0,1}, 4-path does not exist between p and q because it is impossible to get from p to q by travelling along points that are both 4- adjacent and also have valves from v. Fig, (a) show this condition, it is not possible to get to q . The length of the shortest m-path (shown dashed) is 5. Both of these shortest paths are unique in this case. One possibility for

The shortest 4-path when v= {1, 2} is shown in fig. (b). Its length is 6. It is easily verified that another 4-path of the same length exits between p and q.

$$\begin{array}{cccc} 3 & 1 & 2 & 1(q) \\ 2 & 2 & 0 & 2 \\ & & \uparrow & \\ 1 & 2 & 1 & 1 \\ & & \uparrow & \\ (p)1 \rightarrow & 0 \rightarrow & 1 & 2 \end{array} \qquad \begin{array}{cccc} 3 & 1 & 2 & 1(q) \\ 2 & 2 & 0 & 2 \\ 1 \rightarrow & 2 \rightarrow & 1 \rightarrow & 1 \\ \uparrow & & & \\ (P)1 & 0 & 1 & 2 \end{array}$$

(a) (b)

Long Answer Type Questions

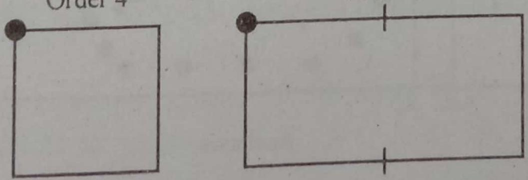
1. a) What is boundary descriptor? **[WBUT 2010]**
b) Explain Fourier descriptor.

Answer:

a) There are several simple geometric measures that can be useful for describing a boundary. The length of a boundary which is the number of pixels along a boundary gives a rough approximation of its length.

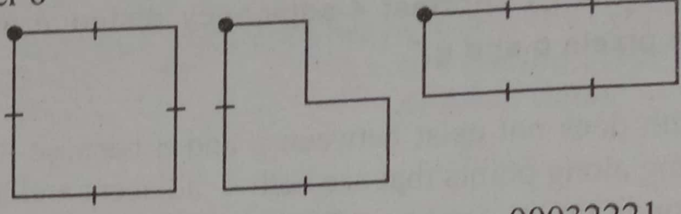
Curvature is the rate of change of slope. To measure a curvature accurately at a point in a digital boundary is difficult. The difference between the slopes of adjacent boundary segments is used as a descriptors of curvature at the point of intersections of segments.

Order 4 Order 6



Chain code	0321	003221
First difference	3331	303303
Shape no.	3333	033033

The shape number of a boundary is defined as the first difference of a smallest magnitude.
The order n of a shape number is defined as the number of digits in its representation.
The first difference is calculated by treating the chain code as a circular sequence.
Shape number of order 8



Chain code:	00332211	03032211	00032221
Difference:	30303030	33133030	30033003
Shape no.:	03030303	03033133	00330033

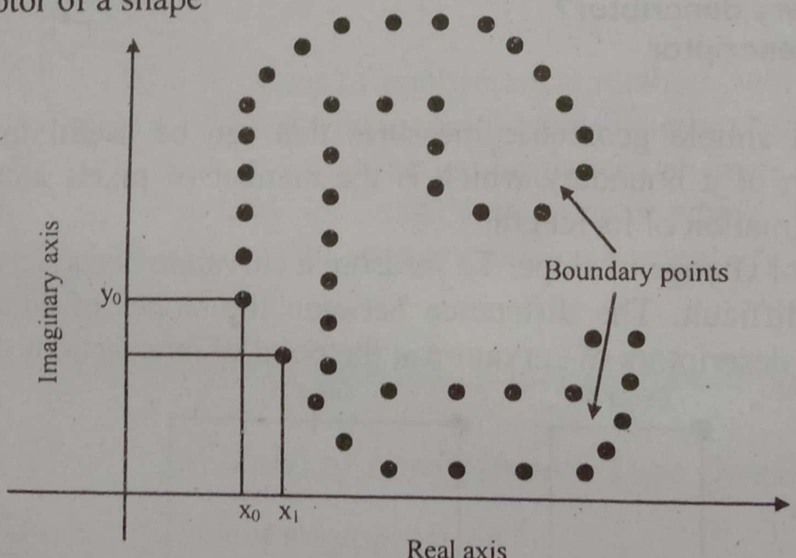
b) Fourier descriptors are a way of encoding the shape of a two-dimensional object by taking the Fourier transform of the boundary, where every (x, y) point on the boundary is mapped to a complex number $x+iy$. The original shape can be recovered from the inverse Fourier transform.
Let $S(k)$ be a co-ordinate of a boundary point k, then

$$S(k)=x(k)+iy(k)$$

Fourier descriptor:

$$a(u)=\frac{1}{k}\sum_{k=0}^{k-1} S(k)e^{\frac{-2\pi uk}{k}}$$

Fourier descriptor of a shape



The fourier descriptors of a shape are calculated as follows:
1) Find the coordinates of the edge pixels of a shape and put them in a list in order, doing clockwise around the shape.
2) Define a complex-valued vector using the coordinates obtained. For example,
3) Take the discrete fourier transform of the complex-valued vector.

Fourier descriptors inherit several properties from the fourier transform.

4) **Translation invariance:** no matter where the shape is located in the image, the fourier descriptor remain the sample.
5) **Scaling:** If the shape is scaled by a factor, the fourier descriptors are scaled by that factor.
6) **Rotation and starting point:** Rotating the shape or selecting a different starting point only affects the phase of the descriptors.

Because the discrete fourier transform is invertible, all the information about the shape is contained in the fourier descriptors.

2. What do you mean by global and local thresholding? What is the basic difference between region growing and split merge technique?

[WBUT 2011, 2013, 2015]

Answer:

1st Part:

In general, the threshold can be chosen by the relation $t=t(r, c, p(r, c))$, where $p(r, c)$ is the feature value at the pixel (r, c). In case of gray level thresholding, $p(r, c)=g(r, c)$ for all (r, c) within the image domain. If t depends on the feature $p(r, c)$ only, it is called local threshold. If t depends on the pixel position (r, c) as well as on the feature $p(r, c)$ at that pixel, it is called global threshold.

2nd Part:

Region-growing methods rely mainly on the assumption that the neighbouring pixels within one region have similar values. The common procedure is to compare one pixel with its neighbours. If a similarity criterion is satisfied, the pixel can be set to belong to the cluster as one or more of its neighbours. Split-and-merge is based on a quad tree partition of an image. It is sometimes called quad tree segmentation. This method starts at the root of the tree that represents the whole image. If it is found non-uniform then it is split into four son squares and so on. If, in contrast, four son squares are homogeneous, they are merged as several connected components. The node in the tree is a segmented node. This process continues recursively until no further splits or merges are possible. Region growing approach is the opposite of the split and merge approach.

Region growing can provide the original images which have clear edges the good segmentation results.

The merge split routine is an optional stage of our region growing based segmentation scheme. It requires a threshold as an input.

3. a) Explain the process of region extraction based on segmentation.

[WBUT 2016]

Answer:

In this technique pixels that are related to an object are grouped for segmentation. The Thresholding techniques is bound with region based segmentation. The area that is detected for segmentation should be closed. Region based segmentation is also termed as similarity based segmentation. These would not be any gap due to missing edge pixels in this in this region based segmentation. The boundaries are identified for segmentation. In each and every step at least one pixel is related to the region and is taken into consideration. After identifying the change in the color and texture, the edge flow is converted into a vector. From this the edges are detected for further segmentation.

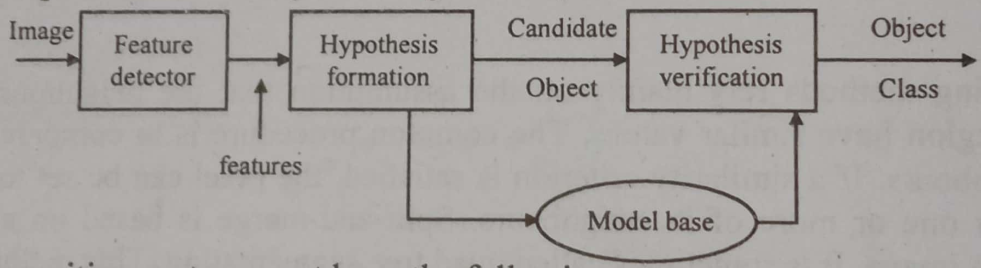
b) With neat sketch, explain the object recognition process. **[WBUT 2016]**

Answer:

An object recognition system finds object in the real world from an image of the world, using object models which are known a priori. This task is surprisingly difficult. Human perform object recognition effortlessly and instantaneously. Algorithmic description of this task for implementation on machines has been very difficult.

The object recognition problem can be defined as a labeling problem based on models of known objects. Formally, given an image containing one or more objects of interest (and background) and a set of labels corresponding to a set of models known to the system, the system should assign correct labels to regions, or a set of regions, in the image. The object recognition problem is closely tied to the segmentation problem without at least a partial recognition of objects segmentation can not be done and without segmentation object recognition is not possible.

Different components of an object recognition system:



An object recognition system must have the following components to perform the task

- Model database (also called model base)
- Feature detector
- Hypothesizer
- Hypothesis verifier

The model database contains all the models known to the system. The information in the model database depends on the approach used for the recognition. It can vary from a qualitative or functional description to precise geometric surface information.

The feature detector applies operations to images and identifies locations of features that help in forming object hypothesis.

c) What is pattern fritting approach? **[WBUT 2016]**

Answer:

Pattern fritting approach

In this approach a gray scale image is considered a topographic surface where altitude at a point is given by its intensity or gray level. The main objective, here, is to fit a pattern over the neighbourhood of the pixel at which edge strength is being calculated. Parameters corresponding to the best fit of the pattern are estimated locally. Finally, properties of edge points are parameters and edges are marked accordingly.

4. a) With appropriate example discuss region split technique for segmentation.
b) How isolated point and line can be identified? **[WBUT 2018]**

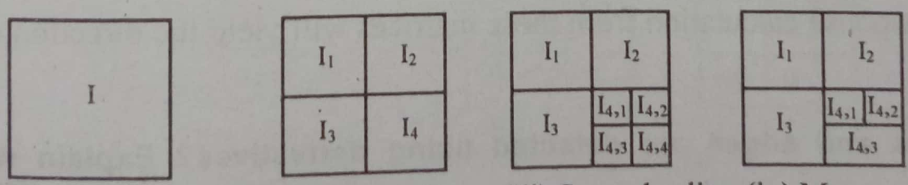
Answer:

a) The basic idea of region splitting is to break the image into a set of disjoint regions which are coherent within themselves:

- Initially take the image as a whole to be the area of interest.
- Look at the area of interest and decide if all pixels contained in the region satisfy some similarity constraint.
- If **true** then the area of interest corresponds to a region in the image.
- If **false** split the area of interest (usually into four equal sub-areas) and consider each of the sub-area as the area of interest in turn.
- This process continues until no further splitting occurs. In the worst case this happened when the areas are just one pixel in size.
- This is a **divide-and-conquer** or **top-down** method.

If only a splitting schedule is used then the final segmentation would probably contain many neighbouring regions that have identical or similar properties.
Thus, a merging process is used after each split which compare adjacent regions and merges then if necessary, algorithms of this nature are called **split and merge** algorithms.
To illustrate the basic principle of these methods let us consider an imaginary image.

- Let I denote the whole image shown in Fig. (i).
- Not all the pixels in I are similar so the region is split as in Fig. (ii).
- Assume that all pixels within regions I_1, I_2 and I_3 respectively are similar but those in I_4 are not.
- Therefore I_4 is split next as in Fig. (iii).
- Now assume that all pixels within each region are similar with respect to that region and that after comparing the split regions, regions $I_{4,3}$ and $I_{4,4}$ are found to be identical.
- These are thus merged together as in Fig. (iv).



(i) Whole image (ii) First split (iii) Second split (iv) Merge

Fig. Example of region splitting and merging

b) A point is the most basic type of discontinuity in a digital image. The most common approach to finding discontinuity is to run an $(n \times n)$ mask over each point in the image. The mask is shown below.

−1	−1	−1
−1	8	−1
−1	−1	−1

(A mask for point detection)

The point is detected at a location (x, y) in an image where the mask is centered. If the corresponding value of R such that $|R| > T$.
Where R is the response of the mask at any point in the image and T is non-negative threshold value. It means that **isolated point** is detected at the corresponding value (x, y). This formulation serves to measures the weighted differences between the center point and its neighbours since the gray level of an isolated point will be very different from that of its neighbors.
Line detection is the next level of complexity in the direction of image discontinuity. For any point in the image, a response can be calculated that will show which direction the point of a line is most associated with. For line detection, we use two mask, ith and jth mask. Then we have,

$$|R_i| > |R_j|, \quad \forall_j \neq i$$

It means that the corresponding points is more likely to be associated with a line in the direction of the mask i. Line detection masks in horizontal direction (Fig. a), 45° direction (fig. (b)), vertical direction (fig. c) and –45° direction (fig. (d)) is shown below.

−1	−1	−1
2	2	2
−1	−1	−1

Fig. (a)

−1	−1	2
−1	2	−1
2	−1	−1

Fig.(b)

−1	2	−1
−1	2	−1
−1	2	−1

Fig. (c)

2	−1	−1
−1	2	−1
−1	−1	2

Fig. (d)

The greatest response calculation from these matrices will yield the direction of the given pixel.

5. How points and edges are detected using derivatives? Explain with proper example. **[WBUT 2019]**

Answer:

An isolated point may be viewed as a line whose length and width are equal to pixel. Point detection in an image is straight forward. Point detection is based on second order derivative filters (Laplacian). Let the spatial filter mast be defined as

W_1	W_2	W_3
W_4	W_5	W_6
W_7	W_8	W_9

The response of the mask at the center point of the origin is

$$R = \sum_{k=1}^{9} W_k Z_k$$

where Z_k is the intensity of pixel and W_k is the mask co-efficient.

A point is said to be detected at the location on which the mask is centered if $|R| \geq T$ where T is a non-negative threshold.

The output $g(x, y)$ is obtained using the expression

$$g(x, y) = \begin{cases} 1 \text{ if } |R(x,y)| \geq T \\ 0 \quad \text{otherwise} \end{cases}$$

Example of point detection mask:

0	1	0
1	−4	1
0	1	0

0	−1	0
−1	4	−1
0	−1	0

−1	−1	−1
−1	8	−1
−1	−1	−1

1	1	1
1	−8	1
1	1	1

NOTE: Sum of mask co-efficient = 0, indicating that the mask response will be zero in areas of constant intensity.

Edge detection:

Edge detection is the most common approach for detecting meaningful discontinuities in gray level. Edge is set of connected pixels that lie on the boundary between two regions. Edge is a local concept whereas a region boundary is a more global idea. Thus, first order derivative can be used to detect the presence of an edge at a point in an image. Second order derivative can be used to determine whether an edge pixel lies on the dark or light side of an edge. Gradient operators, Prewitt's operator, Sobel's operator etc are used to detect edges.

Example:

Z_1	Z_2	Z_3
Z_4	Z_5	Z_6
Z_7	Z_8	Z_9

−1	−1	−1
0	0	0
1	1	1

−1	0	1
−1	0	1
−1	0	1

Prewitt

−1	0
0	1

0	−1
1	0

Roberts

−1	−2	−1
0	0	0
1	2	1

−1	0	1
−2	0	2
−1	0	1

Sobel

Prewitt's operator is

$$G_x = (Z_7 + Z_8 + Z_9) - (Z_1 + Z_2 + Z_3)$$
$$G_Y = (Z_3 + Z_6 + Z_9) - (Z_1 + Z_4 + Z_7)$$

Sobel's Operator is

$$G_x = (Z_7 + 2Z_8 + Z_9) - (Z_1 + 2Z_2 + Z_3)$$
$$G_y = (Z_3 + 2Z_6 + Z_9) - (Z_1 + 2Z_4 + Z_7)$$

Prewitt and Sobel's operator are among the most used in practice for computing digital gradients. Computation of the gradient requires that these two component $(G_x \,\&\, G_y)$ be combined as

$$\Delta f = \sqrt{G_x^2 + G_y^2} \approx |G_x| + |G_y|$$

Δf is the gradient.

6. Write short notes the following:
a) Optimum Thresholding **[WBUT 2009]**
b) Edge detecting operations **[WBUT 2012, 2014, 2017, 2019]**
c) Split and Merge Algorithm **[WBUT 2014]**

Answer:

a) Optimum Thresholding:

Thresholding is a non-linear operation that converts a gray-scale image into a binary image where the two levels are assigned to pixels that are below or above. The specified threshold value we can apply a threshold to data directly from the command line.
e.g., my Binary Image = my Gray Image > Threshold value 255 : 0
It is however for more efficient to use the image threshold operation which also provides several methods for finding the "optimal" threshold value for a given image. Image threshold provides the following methods for determining the threshold value.

1. Automatically calculate a threshold value using an iterative method.
2. Approximately the histogram of the image as a bi-modal distribution and choose a mid-point value as the threshold level.
3. Fuzzy thresholding using as the measure of Fuzziness.
4. Evaluate the threshold based on the last 8 pixel in each row, using alternating rows (Adaptive thresholding).

b) Edge detecting operations:

Edge detection is an image processing technique for finding the boundaries of objects within images. It works by detecting discontinuities in brightness. Edge detection is used for image segmentation and data extracting in areas such as image processing, computer vision and machine vision. Common edge detection algorithms include sobel, canny, prewitt, Roberts and fuzzy logic methods.

c) Split and Merge Algorithm:

Splitting and Merging attempts to divide an image into uniform regions. The basic representational structure is pyramidal i.e. a square region of size m by m at one level of a pyramid has 4 sub-regions of size $m/2$ by $m/2$ below it in the pyramid. Usually the algorithm starts from the initial assumption that the entire image is a single region, then computes the homogeneity criterion to see if it is TRUE. If false, then the square region is split into the four smaller regions. This process is then repeated on each of the sub-regions until no further splitting is necessary. These small square regions are then merged if they are similar to give larger irregular regions. The process terminates when no further merges are possible.

IMAGE DATA COMPRESSION

Multiple Choice Type Questions

1. We have an image in EPS and JPEG formats [WBUT 2009]
   a) the JPEG file will be large in size
   b) the EPS file will be larger in size
   c) both files will be equal in size
   e) none of these are true

Answer: (a)

2. The relation of intensity (I) and R, G & B in RGB colour model is [WBUT 2010]
   a) $I = 0.6R + 0.25G + 0.15B$
   b) $I = \frac{(R+G+B)}{3}$
   c) $I = \frac{(R+2G+B)}{4}$
   d) $I = 0.5R + 0.25G + 0.25B$

Answer: (b)

3. Huffman coding approach reduces [WBUT 2010, 2013, 2018]
   a) coding redundancy only
   b) inter-pixel redundancy only
   c) coding & inter-pixel redundancy
   d) psycho-visual redundancy only

Answer: (a)

4. Which one of the following coding approaches attacks both the coding and inter-pixel redundancy? [WBUT 2010]
   a) Huffman coding
   b) LZW coding
   c) B_2 coding
   d) All of these

Answer: (c)

5. The basic principle of compression matches the principle of [WBUT 2011, 2013, 2014, 2015, 2018]
   a) Channel coding
   b) Line coding
   c) Source coding
   d) All of these

Answer: (c)

6. Which of the following is not used for image compression? [WBUT 2011, 2014]
   a) Block transfer coding
   b) Wavelet coding
   c) LZW coding
   d) Convolution coding

Answer: (a)

7. In 8 distance measurement system, distance between centre pixel and a corner pixel is [WBUT 2012, 2018]
   a) 2 units
   b) $\sqrt{2}$ units
   c) 1 unit
   d) 1.5 units

Answer: (a)

8. How many numbers of colours are present in RGB? [WBUT 2012, 2013]
a) 3
b) 6
c) 216
d) 256

Answer: (d)

9. Compression is achieved by [WBUT 2015]
a) filtering
b) multiplying an array
c) minimizing the redundancies
d) none of these

Answer: (c)

10. Wavelet-based image-compression scheme overcomes the problem of [WBUT 2015]
a) Blur object
b) Blocking artifact
c) Background sharpening
d) all of these

Answer: (b)

11. Wavelets are [WBUT 2014, 2015]
a) of remarkable advantage for image compression
b) foundation for representing images in various degrees of resolution
c) both (a) and (b)
d) none of these

Answer: (b)

12. An example of dictionary based coding technique is [WBUT 2016]
a) Huffman
b) run-length
c) Lempel-Ziv-Welch
d) predictive

Answer: (c)

13. Which one of the following is the lossy image compression? [WBUT 2017]
a) TIFF
b) BMP
c) GIF
d) none of these

Answer: (a)

Short Answer Type Questions

1. For the image shown below, compute the compression ratio by using Huffman coding. Assume that the pixel value is represented by 8 bits. [WBUT 2012]

13	13	13	12
12	33	33	33
33	22	22	22
22	11	11	12

Answer:

Run length coding will be 13(3), 12(2), 33(4), 22(4), 11(2), 12(1)

Therefore, compression ratio will be $= \frac{16}{12} = 4:3$

2. What is the difference between lossy and lossless compression? **[WBUT 2012, 2013, 2014]**

Answer:

Lossless compression is a compression technique that does not lose any data in the compression process. Lossy compression discards some data – the portion of the image that represents the visually insignificant details. Lossy compression will strip a file of some of its redundant data. Because of this data loss, only certain applications are fit for lossy compression, like graphics, audio and video. JPEG uses lossy compression, which is why converting a GIF file to JPEG will reduce it in size.

Lossless compression has advantage and disadvantage. The advantage is that the compressed file will decompress to an exact duplicate of the original file. The disadvantages is that the compression ratio is not all that high, precisely because no data is lost. Lossy compression necessarily reduces the quality of the file to arrive at the resulting highly compressed size, but depending on the need, the loss may acceptable and even unnoticeable in some cases.

3. What is the role of quantization in image compression? **[WBUT 2012]**

OR,

What is the role of quantization in image processing? **[WBUT 2013, 2014, 2019]**

Answer:

Quantization, involved in image processing, is a lossy compression technique achieved by compressing a range of values to a single quantum value. When the number of discrete symbols in a given stream is reduced, the stream becomes more compressible. For example, reducing the number of colors required to represent a digital image makes it possible to reduce its file size. Specific applications include DCT data quantization in JPEG and DWT data quantization in JPEG 2000.

4. An image is 8MB before compression and 2MB after compression. Find the compression ratio and savings percentage. **[WBUT 2018]**

Answer:

$$\text{Compression ratio} = \frac{\text{UnCompressed size}}{\text{Compressed size}} = \frac{8\text{MB}}{2\text{MB}} = 4$$

i.e. 4:1

Saving percentage = 75%

$$(\text{i.e., } 1 - \frac{2}{8} = 1 - \frac{1}{4} = \frac{3}{4} \times 100 = 75\%)$$

5. Discuss briefly Huffman coding. **[WBUT 2019]**

Answer:

Refer to Question No. 8(a) of Long Answer Type Questions.

Long Answer Type Questions

1. Explain the data redundancy and compression ratio of an imaging system. **[WBUT 2009]**

OR,

What do you mean by redundancy and compression ratio? **[WBUT 2012]**

How many types of data redundancy are there in an imaging system? Explain them. Explain lossless and lossy predictive coding. **[WBUT 2009]**

Answer:

1st part:

Data compression is defined as the process of encoding data using a representation that reduces the overall size of data. This reduction is possible when the original dataset contains some type of **redundancy**.

A **compression ratio** is the average number of bits per pixel (bpp) before compression divided by the number of bits per pixel after compression. For example, if an 8-bit image is compressed and each pixel is then represented by 1 bit per pixel, then the compression ratio $=\frac{8}{1}=8$. Equivalently, for a 24 bit image, of the compression ratio $=18$, the compressed image will have $\frac{24}{18}=1.33\,\text{bpp}$.

2nd part:

In general, three basic redundancies exist in digital images.

Psy-visual redundancy:

It is a redundancy corresponding to different sensitivities to all image signals by human eyes. Therefore, eliminating some less relative important information in our visual processing may be acceptable.

Inter-pixel redundancy:

It is a redundancy corresponding to statistical dependencies among pixels, especially between neighbouring pixels.

Coding redundancy:

The uncompressed image usually is coded with each pixel by a fixed length. For example, an image with 256 gray scales is represented by an array of 8-bit integers. Using some variable length code schemes such as Huffman coding and arithmetic coding may produce compression.

3rd part:

Lossy compression is the class of data encoding methods that uses inexact approximations to represent the content. These techniques are used to reduce data size for storage, handling and transmitting content. This is opposed to lossless data compression which does not degrade the image. The amount of data reduction possible using lossy compression is often much higher than through lossless techniques well design lossy compression technology often reduces fold sizes significantly before degradation is noticed by the end user. Even when noticeable by the user, further data reduction may be desirable (e.g., for real-time communication, to reduce transmission times or to reduce storage needs). Lossy compression is most commonly used to compress multimedia data (audio, video and images), especially in applications such as streaming media and internet telephony. Lossless compression typically required for text and data files, such as bank records and text articles. In many cases, it is advantageous to make a master lossless files which is to be used to produce new compressed files.

Predictive coding has proven to be effective for lossless image compression predictive coding estimates a pixel colour value based on the pixel colour values of its neighbouring pixels. To enhance the accuracy of the estimation, prediction scheme can help minimize the upper bound of the residual errors from the prediction. The main component of the predictive coding method is the predictor which exists in both encoder and decoder computes the predictive colour value for a pixel.

2. Describe LZW coding with example. What is transform coding? Give some example of transform coding. **[WBUT 2009, 2010]**

Answer:

1st part:

LZW (Lempel-Ziv-Welch) algorithm used by the UNIX compress command to reduce the size of files e.g., for archival or transmission. This algorithm relives on reoccurrence of byte sequences (strings) in its input. It maintains a table mapping input strings to their associated output codes. The table initially contains mappings for all possible strings of length one. Input is taken one byte at a time to find the longest initial string present in the table. The code for that string is output and then the string is extended with one more input byte b. A new entry is added to the table mapping the extended string to the next unused code. The process repeats, stating from byte b.

LZW compression is the compression of a file into a smaller file using a table-based look up algorithm. Two commonly-used file formats in which LZW compression is used are the GIF and TIFF image format. LZW compression is also suitable for compressing text files.

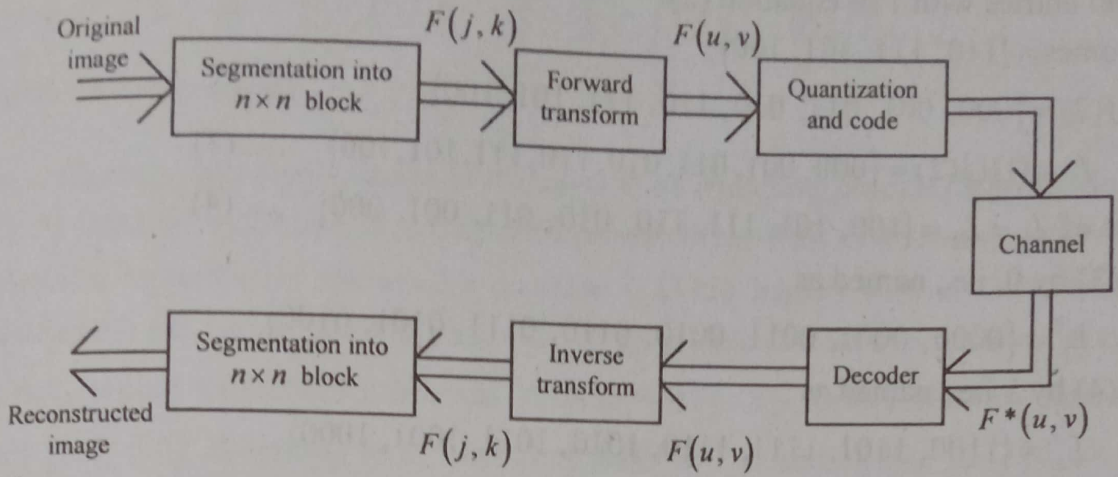
A particular LZW compression algorithm takes each input sequence of bits of a given length and creates an entry in a table for that particular bit pattern, consisting of the pattern itself and a shorter code. As input is head, any pattern that has been read before results in the substitution of the shorter code, effectively compressing the total amount of input to something smaller.

2nd part:

The purpose of transform coding is to convert the data into a form when compression is easier. This transformation will transform the pixels which are correlated into a representation where they are decorrelated. The new values are usually smaller on average than the original values. The net effect is to reduce the redundancy of representation.

For lossy compression, the transform co-efficients can now be quantized according to their statistical properties, producing a much compressed representation of the original image data.

Block diagram of transform coding:



Examples:

Block transform coding:

In order to simplify the computations, block transform coding exploits correlation of the pixels within a number of small blocks that divide the original image. As a result each block is transformed, quantized and coded separately. This technique, using square $8*8$ pixel blocks and be DCT followed by Huffman or arithmetic coding is utilized in the JPEG draft international standard for image compression. The disadvantage of this scheme is the blocking artefacts appear at high compression ratios.

Full frame transform coding:

To avoid the artefacts generated by block transforms, full frame methods, in which the transform is applied to the whole image as a single block, have been investigated in medical imaging research. The tradeoff is the increased computational requirements and the appearance of ringing artefacts. Subband coding is one example among full frame methods.

3. a) Construct the entire 4 bit gray code.
b) Create a general procedure for converting a gray code number to its equivalent binary and use it to decode 0111010100111.
c) The arithmetic decoding process is the reverse of the encoding process. Decode the message 0.23355 using the given coding model. Consider '!' as the terminating symbol. **[WBUT 2010]**

Symbol:	A	E	I	O	U	!
Probability:	0.2	0.3	0.1	0.2	0.1	0.1

Answer:

a) 1 bit Gray Code $=\{0, 1\}$

2 bit Gray Code $=\{00, 01, 11, 10\}$ .... (1)

Reverse is $=\{10,\ 11,\ 01,\ 00\}$ .... (2)

Prefix all entries with 0 of equation (1)

(1) becomes $=\{000,\ 001,\ 011,\ 010\}$

Prefix all entries with 1 of equation (2)

(2) becomes $=\{110,\ 111,\ 101,\ 100\}$

$\therefore\ (1)\cup(2)=\{000,\ 001,\ 011,\ 010,\ 110,\ 111,\ 101,\ 100\}$

Now, $L_1=(1)\cup(2)=\{000, 001, 011, 010, 110, 111, 101, 100\}$ .... (3)

Reverse of $L_1=L_2=\{100,\ 101,\ 111,\ 110,\ 010,\ 011,\ 001,\ 000\}$ .... (4)

Prefix (3) by 0, i.e., named as

$$L_1'=\{0000,\ 0001,\ 0011,\ 0010,\ 0110,\ 0111,\ 0101,\ 0100\}$$

Prefix (4) by 1 i.e., named as

$$L_2'=\{1100,\ 1101,\ 1111,\ 1110,\ 1010,\ 1011,\ 1001,\ 1000\}$$

$$\therefore\quad L_1'\cup L_2'=\{0000,\ 0001,\ 0011,\ 0010,\ 0110,\ 0111,\ 0101,\ 0100,\ 1100,\ 1101,\ 1111,\ 1110,\ 1010,\ 1011,\ 1001,\ 1000\}$$

Hence proved.

b) 1st Part:

For Gray Code to Binary, the general procedure is,

i) Write down the most significant bit of the gray code number on the left because most significant bit of the binary number is same as that of the gray code number.

ii) To find the next binary digit, do the Exclusive-OR ($\oplus$) operation between the next digit of the gray code number and the binary digit just written down.

iii) Continue step (ii) until all the digits of the gray code number Ex-OR ed with binary digits.

Examples:

Let the gray code number be

100110

MSD

Gray Code: 1 0 0 1 1 0

$\oplus$ $\oplus$ $\oplus$ $\oplus$ $\oplus$

Binary Code: 1 1 1 0 1 1

2nd Part:

Gray code number as

0 1 1 1 0 1 0 1 0 0 1 1 1

Binary: 0 1 0 1 1 0 0 1 1 1 0 1 0

Hence the result in Binary

c) The arithmetic decoding process is the reverse of encoding process. Start by dividing the $[0, 1)$ interval according to the symbol probabilities. This is shown in table below. The decoder immediately knows the message 0.23355 begins with an "e", since the coded message lies in the interval $[0.2, 0.5)$. This makes it clear that the second symbol is in "a", which narrows the interval to $[0.2, 0.26)$. To further see this, divide the interval $[0.2, 0.5)$ according to symbol probabilities. Proceeding like this, which is the same procedure used to code the message, we get "$eaii!$".

Table

Symbol	**Probability**	**Range**
a	0.2	$[0.0, 0.2)$
e	0.3	$[0.2, 0.5)$
i	0.1	$[0.5, 0.6)$
o	0.2	$[0.6, 0.8)$
u	0.1	$[0.8, 0.9)$
!	0.1	$[0.9, 1.0)$

4. State the JPEG compression algorithm and draw the schematic diagram of JPEG compressor. **[WBUT 2013, 2019]**

Answer:

1st Part:

The JPEG standard specifies a multilevel coding process:

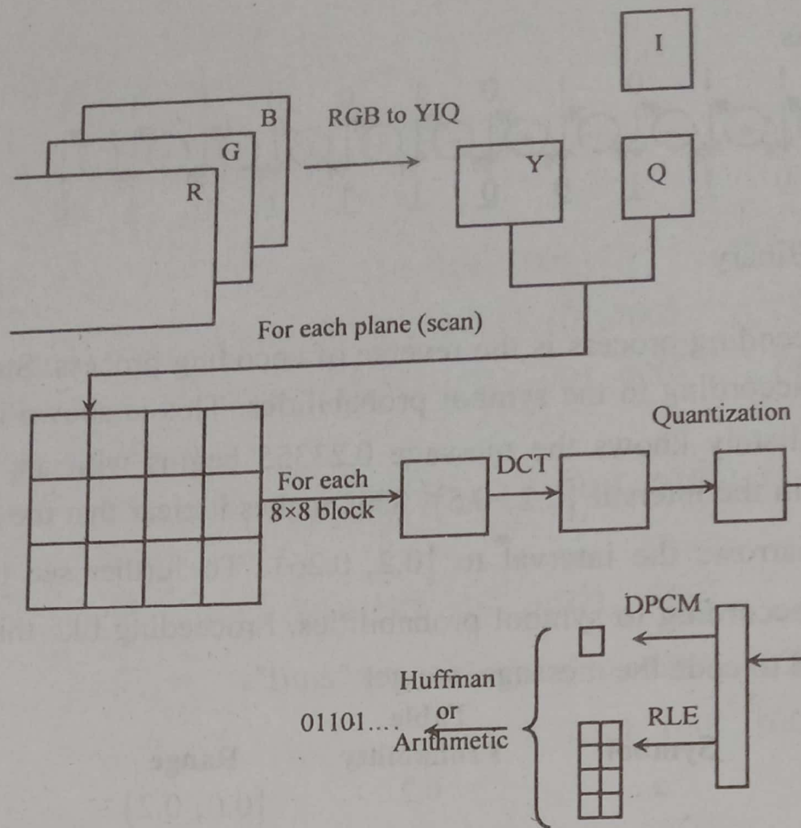
i) Image preparation (Separation into components): The image will be separated into components, which are adapted to the different quality requirements.

ii) Transformation (discrete cosine transform): Each individual component is transferred into a format that allows to make conclusions about the structure of the contents. This offer the option to distinguish between basic and more complex contents.

iii) Quantization (weighting of the contents): The transformed data are weighted according to their meaning for the image contents.

iv) Entropy Coding: Elimination of redundant information.

2nd Part:



5. a) What do you mean by redundancy? What are the types of image Compression? **[WBUT 2014, 2018]**
b) What is compression ratio? **[WBUT 2014, 2018]**
Consider an 8-bit image of 256×256 pixels. After compression, the image size is 6,554 bytes. Find the compression ratio. **[WBUT 2014]**
c) Describe image compression System. **[WBUT 2014]**

Answer:

a) 1st Part:

Representation that contain irrelevant or repeated information are said to contain redundant data. If b and b' denote the number of bits or information carrying units in two representation of the same information, the relative data redundancy R is,

$$R = 1 - \frac{1}{C}$$

Where C is compression ratio and is defined as $C = b/b'$

2nd Part:

Compression techniques are classified into two primary types:

- Lossless compression
- Lossy compression

Lossless allows an image to be compressed and decompressed without losing information. Example of lossless compression is Huffman Coding.

Lossy compression allows a loss in the actual image data.

This is useful in application such as broad cost television and video conferencing. JPEG is the example.

b) 1st Part:
The ratio of the original (uncompressed) image to the compressed image is referred to as compression ratio.

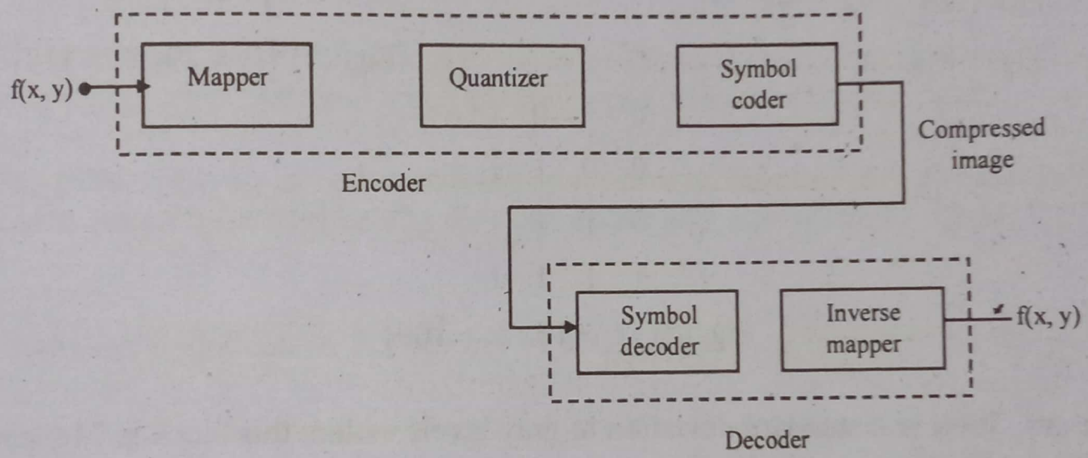
$$C_R = \frac{\text{Uncompressed size}}{\text{Compressed size}}$$

2nd Part:
Uncompressed size $= (256 \times 256 \times 8)/8 = 65{,}536$ bytes

$$\therefore \text{ Compression ratio} = \frac{65{,}536}{6{,}554} = 9.999 \approx 10$$

i.e. $C_R = 10:1$

c) As shown below, the image compression system consists of two distinct structure blocks : an encoder and a decoder.



The encoder is responsible for reducing or eliminating any coding, inter-pixel or psychovisual redundancies in the input image. It consists of,

- **Mapper:** It transforms the input image into a non-visual format designed to reduce interpixel redundancies in the input image. This operation is reversible and may or may not reduce directly the amount of data required.
- **Quantizer:** It reduces the psychovisual redundancies of the input image. This operation is irreversible.
- **Symbol Coder:** It creates a fixed or variable-length code to represent the quantizer output. In a variable length code, the shortest code words are assigned to the most frequently occurring output values, and thus reduce coding redundancy. This operation is reversible.

The decoder contains only two components i.e. symbol decoder and an inverse mapper.

6. a) Why image compression is required? **[WBUT 2015]**

Answer:
Users of digital image processing techniques usually have to handle a large volume of data. Storing image data for future use needs large storage space. Similarly, transmitting image data in reasonable time needs wide channel capacity. To reduce these requirements the technique we use is called data compression. So, compression techniques represent same pictorial information in a more compact form by removing redundancies. In essence, image compression techniques represent image data using fewer bits than what is required for original image.

b) Take any 4×4 gray image of your choice and show how block truncation coding can be applied to compress it. **[WBUT 2015]**

Answer:
Fig. (a) shows a 4×4 block which is a part of a large gray level image.

17	22	20	140
19	26	136	95
25	126	69	98
133	73	92	94

Fig. (a)

0	0	0	1
0	0	1	1
0	1	0	1
1	0	1	1

Fig. (b) $[\ell_1 = 29;\ \ell_2 = 119]$

0	0	0	1
0	0	1	1
0	1	1	1
1	1	1	1

Fig. (d) $[\ell_1 = 16;\ \ell_2 = 109]$

Average gray level and standard deviation of gray levels within, this block is 74 (approx.) and 86.78 (approx.) respectively. Compressed image blocks obtained by using average gray level threshold and threshold corresponding to the largest gray level difference respectively are represented by binary images shown in Fig. (b) and Fig. (d) along with the values of ℓ_1 and ℓ_2.

In these figures, gray levels less than threshold are represented by 0 and those greater than or equal to threshold are by 1. Thus the whole block can be represented by 16 bits and values of ℓ_1 and ℓ_2. Secondly, for Fig. (b), $M_1 = 8$ and $M_2 = 8$ and for Fig. (d) $M_1 = 6$ and $M_2 = 10$. Values of ℓ_1 and ℓ_2 are computed using equations

$$\ell_1 = \bar{g}_i - \sigma\sqrt{\frac{M_2}{M_1}}$$

$$\ell_2 = \bar{g}_i + \sigma\sqrt{\frac{M_1}{M_2}}$$

where $\sigma = \sqrt{\overline{g_i^2} - \bar{g}_i^2}$

ℓ_1 and ℓ_2 gray levels and $\overline{g}_i(r,c)=\begin{cases}\ell_1 & \text{if } g_i(r,c)<\text{threshold}\\ \ell_2 & \text{otherwise}\end{cases}$

and the no. of pixels in the ith block that have approximated value equal to ℓ_1 is M_1 and that equal to ℓ_2 is M_2.

For Fig. (b), $\ell_1=29$ and $\ell_2=119$ and for Fig. (d) $\ell_1=16$ and $\ell_2=109$. So, for each compressed block two extra bytes (i.e., 16 bits) are needed to store these values. Finally, a block is truncated based on these values and the corresponding binary image. Hence, image block reconstructed from Fig. (b) is shown Fig. (c) and that from Fig. (d) in Fig. (e).

29	29	29	119
29	29	119	119
29	119	29	119
119	29	119	119

Fig. (c)

16	16	16	109
16	16	109	109
16	109	109	109
109	109	109	109

Fig. (e)

Fig. (b) is compressed image using average gray as threshold and Fig. (d) is compressed image using the threshold selected based on the largest difference in gray level.

c) Classify different type of redundancies in an image. Mention the different video compression techniques and briefly discuss about any one of them. [WBUT 2015]

Answer:

1st part:

Data compression is defined as the process of encoding data using a representation that reduces the overall size of data. This reduction is possible when the original data set contains some type of redundancy. In general, three basic redundancies exist in digital images that follow:

Psycho-visual redundancy:

It is a redundancy corresponding to different sensitivities to all image signals by human eyes. Therefore, eliminating some less relative important information in one visual processing may be acceptable.

Inter-pixel redundancy:

It is a redundancy corresponding to statistical dependencies among pixels, especially between neighbouring pixels.

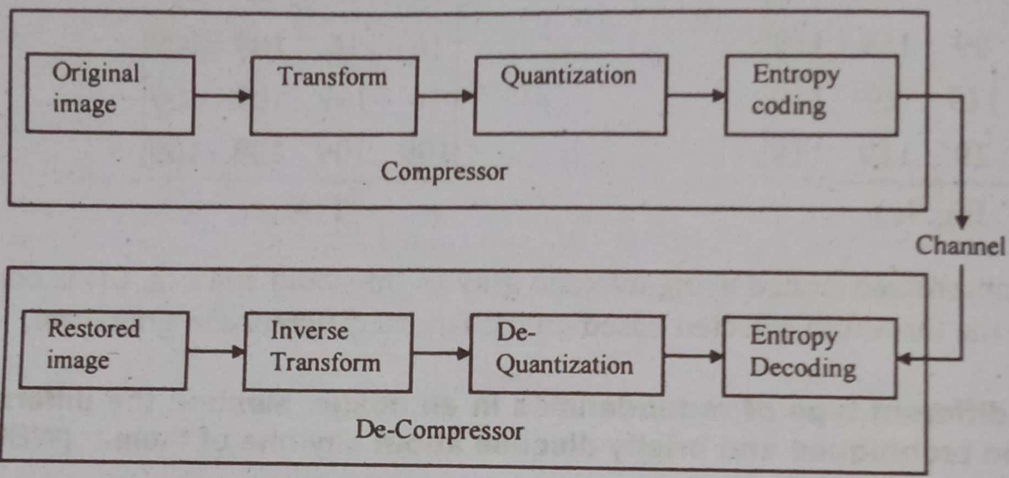
Coding redundancy:

The uncompressed image usually is coded with each pixel by a fixed length. For example, an image with 256 gray scales is represented by an array of 8-bit integer. Using some variable length code schemes such as Huffman coding and arithmetic coding may produce compression.

2nd Part:

During the past two decades, various compression methods have been developed to address major challenges faced by digital imaging. These video compression methods can be classified broadly into **lossy** or **lossless compression**. Lossy compression can achieve a high compression ratio, 50:1 or higher, since it allows some acceptable degradation. Yet it cannot completely recover the original data. On the other hand, lossless compression can completely recover the original data but this reduces the compression ratio to around 2:1.

Lossy compression method:
Generally most lossy compression are three-steps algorithms, each of which is in accordance with three kind of redundancy mentioned above.



The first state is a transform to eliminate the inter-pixel redundancy to pack information efficiently. Then a quantizer is applied to remove psycho-visual redundancy to present the packed information with as few bits as possible. The quantized bits are then efficiently encoded to get more compression from the coding redundancy.

7. a) What is the need for compression? What are different compression methods? **[WBUT 2016]**

Answer:

1[st] **part:** ***Refer to Question No. 6(a) of Long Answer Type Questions.***

2[nd] **Part:** ***Refer to Question No. 2 of Short Answer Type Questions.***

b) Discuss briefly the concept of Huffman coding. **[WBUT 2016, 2018]**

Answer:

Refer to Question No. 8(a) of Long Answer Type Questions.

c) Define interpixel redundancy? **[WBUT 2016]**

Answer:

Interpixel redundancy is defined as failure to identify and utilize data relationships. If a pixel value can be reasonably predicted from its neighbouring pixels, the image is said to contain interpixel redundancy. It depends on the resolution of the image. The higher resolution (spatial) an image, the more probable it is that two neighbouring pixels will depict the same object.

8. Write short notes on the following:

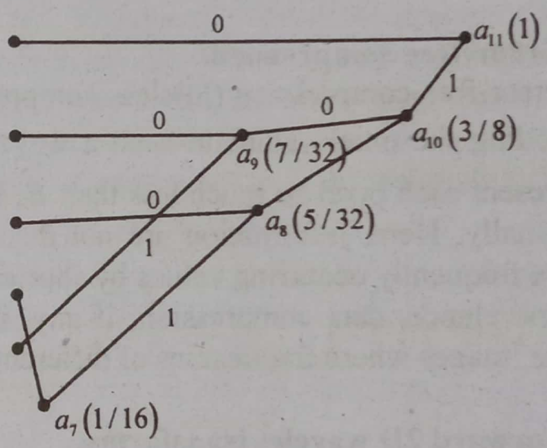
a) Huffman coding **[WBUT 2009]**
b) Lossy compression **[WBUT 2011]**
c) Error-free compression **[WBUT 2011]**
d) Forward 2D wavelet transforms **[WBUT 2012]**
e) Wavelet **[WBUT 2013]**
OR,
Wavelet Transform **[WBUT 2019]**
f) Vector quantization method for image compression **[WBUT 2015]**
g) Run length coding **[WBUT 2016]**
h) Image compression techniques **[WBUT 2017]**
i) LZW compression **[WBUT 2018]**

Answer:

a) Huffman coding:

One optimal codeword design method that is simple to use and which is uniquely decidable and which result in lowest average bit rate is Huffman coding.

Message	Code	Probability
a_1	0	$P_1 = \frac{5}{8}$
a_2	100	$P_2 = \frac{3}{32}$
a_3	110	$P_3 = \frac{3}{32}$
a_4	1110	$P_4 = \frac{1}{32}$
a_5	101	$P_5 = \frac{1}{8}$
a_6	1111	$P_6 = \frac{1}{32}$



An example of Huffman coding is shown in the above figure. In the example $L=6$ with the probability for each message probability noted at each node.

In the 1st step of Huffman coding, we select two message possibilities that have two lowest probabilities. We combine them and form a new node with combined probabilities. We assign '0' to one of the two branches and '1' to other. Reversing this affects the codeword but not the average bit rate. We continue with this process until we are left with one message with probability '1'. To determine the specific codeword assigned to each message possibility. We begin with last node with probability '1', follow the branches that lead to the message possibility of interest and combine the 0's and 1's on the branches.

For example, a_4 has codeword 1110. To compare performance of Huffman coding with the entropy H and uniform length codeword assignment for the above example, we compute average bit rate achieved by uniform length codeword, Huffman coding and the entropy respectively.
For uniform length codeword 3 bit/message.

Huffman Coding

$$\frac{5}{8}\cdot 1+\frac{3}{32}\cdot 3+\frac{3}{32}\cdot 3+\frac{1}{32}-4+\frac{1}{8}\cdot 3+\frac{1}{32}-4=\frac{29}{18}\text{ bits/message}=1.813\text{ bits/message}$$

b) Lossy compression:
Lossy compression is the class of data encoding methods that uses inexact approximations to represent the content. These techniques are used to reduce data size for storage, handling and transmitting content. Well designed lossy compression technology often reduces file sizes significantly before degradation is noticed by the end-user. Even when noticeable by the user, further data reduction may be desirable. Lossy compression is most commonly used to compress multimedia data (audio, video and images), especially in applications such as streaming media and internet telephony.

c) Error-free compression:
In error-free compression (lossless compression) data reduction can also be achieved by encoding the pixel values in such a way that the average number of bits (b') used to represent each pixel be much less than b, the number of bits used to represent each pixel originally. Here, information are not discarding in lossless compression, but represent more frequently occurring values by shorter codes and the less occurring values by longer codes. Hence, data compression, if any, incurs zero error. This method is suitable for those images where frequencies of different gray levels vary widely.

d) Forward 2D wavelet transforms:
A 2-D wavelet transform can be seen as a 1-D wavelet scheme which transform along the rows and then a 1-D wavelet transform along the column. The 2-D DWT operates in a straightforward manner by inserting array transposition between the two 1D-DWT. The rows of the array are processed first with only one level of decomposition. This essentially divides the array into two vertical halves, with the first half storing the average co-efficients, while the second vertical half stores the detail co-efficients. This process is repeated again with the columns, resulting in four sub-bands within the array defined by filter output. One of the prominent features of JPEG 2000 standard, providing it the resolution scalability is the use of 2D-DWT to convert the image samples into a more compressible form.

e) Wavelet:

Wavelets represent another approach to decomposing complex signals into sums of basic functions – in this respect they are similar to Fourier decomposition approaches, but they have an important difference. Wavelets are derived from a basis function $\phi(x)$ called the mother function or analyzing wavelet. The wavelet basis is then provided by the function

$$\phi_{(s,c)}(x) = 2^{-(s/2)}\phi\left(2^{-s}x - l\right)$$

Here, the scale factor s indicates the wavelet's width and the location index l its position. The mother function is obviously key in the application of wavelets; the simplest such is the Haar, which is a simple step.

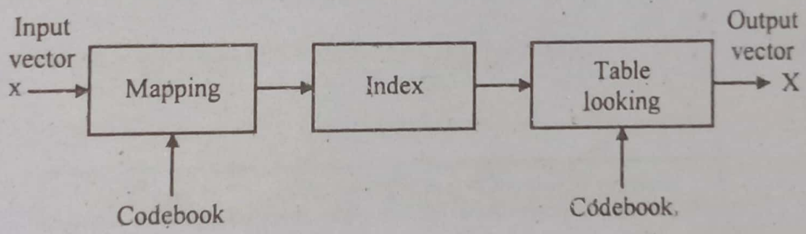
$$\phi(x) = 1 \qquad 0 \le x < 0.5$$

$$\phi(x) = -1 \qquad 0.5 \le x < 1$$

Wavelets have been used with enormous success in data compressive and in image noise suppression.

f) Vector quantization method for image compression:

Vector quantization is a lossy data compression method based on the principle of block coding. It is a fixed-to-fixed length algorithm. In the earlier days, the design of a vector quantizer is considered to be a challenging problem due to the need for multidimensional integration. In 1980, Linde, Buzo and Gray proposed a vector quantization design algorithm based on a training sequence. The use of a training sequence bypasses the need for multidimensional integration. A vector quantization that is designed using this algorithm are referred to LBG-Vector quantization output.



g) Run length coding:

Run length encoding is a data compression algorithm that is supported by most bit map file formats such as TIFF, BMP and PCX.RLE (Run length encoding) is suited for compressing any type of data regardless of its information content, but the content of the data will affect the compression ratio achieved by RLE. For example, if the input string is "wwwwaaadexxxxxx", then the REL function should return

w4a3d1e1×6

RLE may also be used to refer to an early graphics file format supported by "CompuServe" for compressing black and white images, but was widely supported by their later Graphics interchange format.

h) Image compression techniques:
Refer to Question No. 5(a) (2nd Part) & 6(a) of Long Answer Type Questions.

i) LZW compression:
Refer to Question No. 2(1st Part) of Long Answer Type Questions.

MORPHOLOGICAL IMAGE PROCESSING

Multiple Choice Type Questions

1. Erosion [WBUT 2010]
   a) sharps a region b) blurs a region
   c) increases a region d) decreases a region

Answer: (c)

2. .... is not a morphological image processing algorithm. [WBUT 2011]
   a) Thinning b) Skeleton
   c) Both (a) & (b) d) None of these

Answer: (d)

3. Knowledge of which one of the following is not required for morphological image processing? [WBUT 2011, 2012]
   a) Erosion b) Morphological reconstruction
   c) Neural networking d) Duality & dilation

Answer: (d)

4. Which of the following operations is idempotent? [WBUT 2016]
   a) dilation b) erosion c) convolution d) closing

Answer: (d)

5. The morphological operation used to isolate objects which may be just touching one another is [WBUT 2017]
   a) dilation b) erosion c) opening d) closing

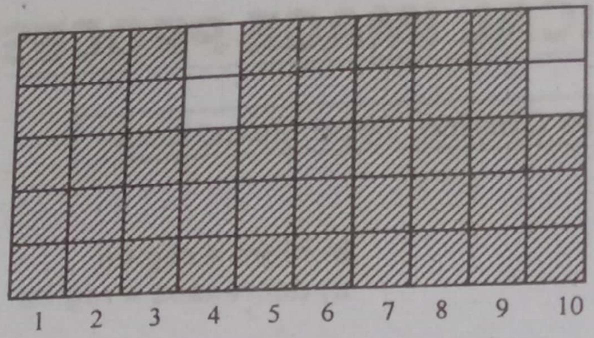
Answer: (b)

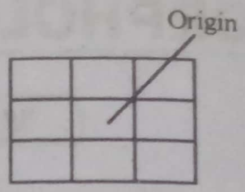
Short Answer Type Questions

1. What is boundary extraction in Morphological image operation? [WBUT 2014, 2017, 2018]

Answer:
When dealing with binary images, the principle application of morphology is extracting image components that are useful in the representation and description of shape. The boundary $\beta(A)$ of a set A, $\beta(A) = A - (A \theta B)$, where B is a suitable structuring element.

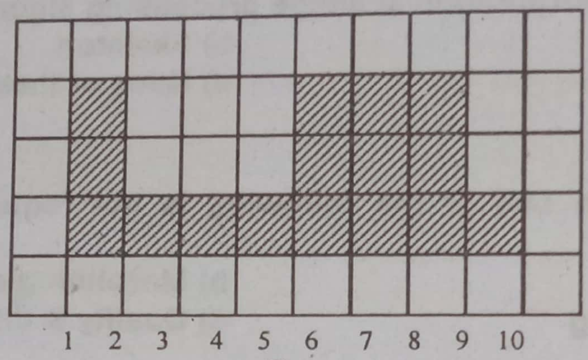
Illustration





Structuring element B

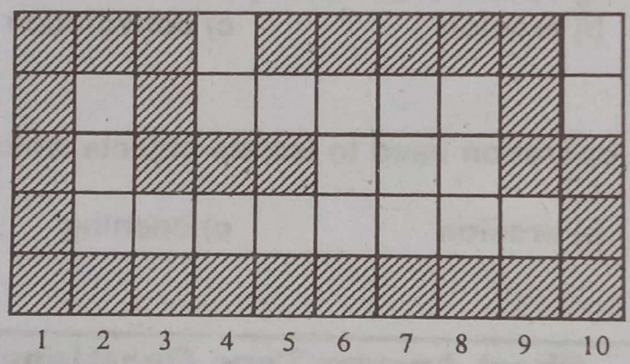
Set A



$A \ominus B$

A encoded by B

Therefore, difference between A and its erosion



$\beta(A)$

2. Explain the operation of an order statistics filter and classify it. **[WBUT 2015]**

Answer:

Order statistics filters are spatial filters whose response is based on ordering the pixel contained in the image area encompassed by the filter. The response of the filter at any point is determined by the ranking result.

Median filter is the best order statistics filter; it replaces the value of a pixel by the median of gray levels in the neighbourhood of the pixel.

i.e., $\hat{f}(x, y) = \text{median}\left(g(s,t)\right)$

The original of the pixel is included in the computation of the median of the filter are quite possible because for certain types of random noise, thus provide excellent noise reduction capabilities with considerably less blurring then smoothing filters of similar size. These are effective for bi-polar and unipolar impulse noise.

3. a) Briefly explain skeletons. **[WBUT 2016]**
b) Write down the effects of the dilation and erosion process.

Answer:

a) Skeleton are the minimal representations of objects in an image while retaining the enter number of the image. The enter number is the number of objects in an image minus the number of holes in those objects. The skeleton on of objects in an image can be found by successive thining until stability is reached.

b) **Effect of dilation Process:**

1. Dilation expands the input image.
2. This process effects both the inside and outside borders of the input image.
3. Dilation fills the holes in the image.
4. The dilation Process smoothens out the image contour.

Effect of erosion process:

1. The erosion process shrinks the given input image.
2. Pixels and holes inside the image are normally removed by the erosion process.
3. Erosion Process removes small objects like noise.

4. Explain the lines and edges detecting process with mathematical models. **[WBUT 2016]**

Answer:

Line masks are horizontal, 45° line, vertical line, –45° line and is given by

$$\text{horizontal line}\begin{bmatrix} -1 & -1 & -1 \\ 2 & 2 & 2 \\ -1 & -1 & -1 \end{bmatrix};\ \text{vertical line}\begin{bmatrix} -1 & 2 & -1 \\ -1 & 2 & -1 \\ -1 & 2 & -1 \end{bmatrix}$$

$$45°\ \text{line}\begin{bmatrix} -1 & -1 & 2 \\ -1 & 2 & -1 \\ 2 & -1 & -1 \end{bmatrix};\ -45°\ \text{line}\begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$$

If at a certain point in the image, $|R_i| > |R_j|$ for all $j \neq i$, that point is said to be more likely associated with a line in the direction of mask i, where R_i & R_j are the response of the mask. Edge detection locates sharp changes in the intensity function. Edges are pixels where brightness changes abruptly. A change of the image function can be described by a gradient that points in the direction of the largest growth of the image function. An edge is a property attached to an individual pixel and is calculated from the image function behavior in a neighbourhood of the pixel. Magnitude of the first derivative detects the presence of the edge. Sign of the second derivative determines whether the edge pixel lies on the dark sign or lightside.

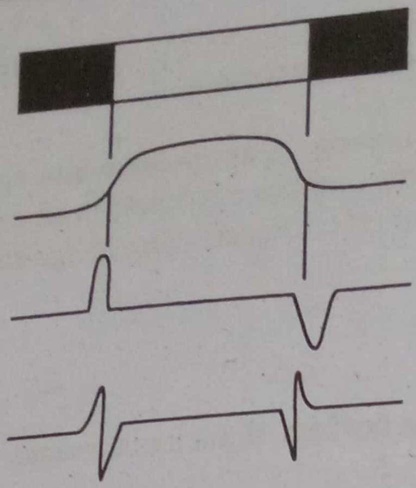


Fig: (a)

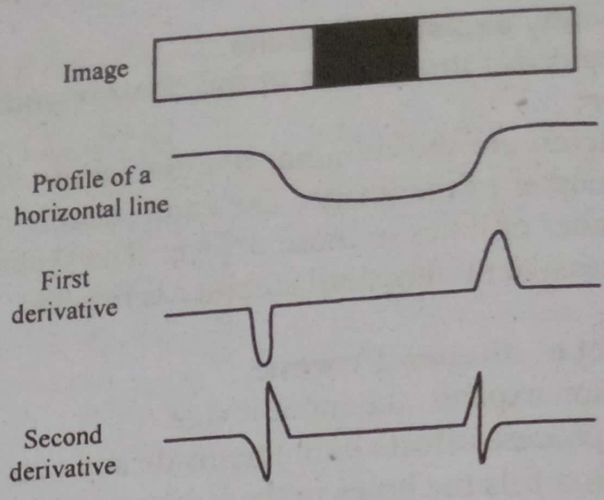


Fig: (b)

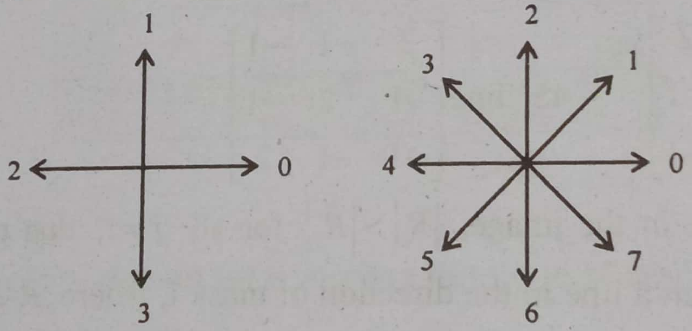
Fig. Edge detection by derivative operators.
a) light stripe on a dark background
b) dark stripe on a light background

Long Answer Type Questions

1. Explain the chain code. **[WBUT 2010]**

Answer:

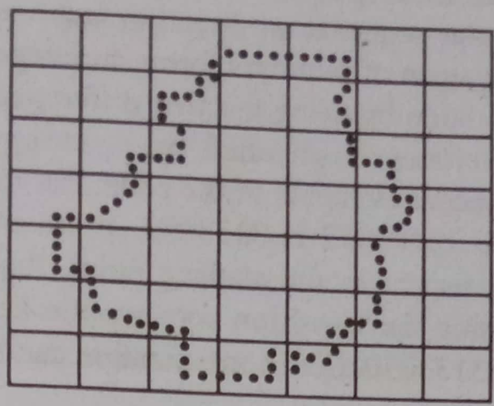
Chain codes are used to represent a boundary for connected sequence of straight line segments of specified length and direction. Typically this representation is based on 4- or 8- connectivity of the segments. The direction of each segment is coded by using a numbering scheme as shown below.



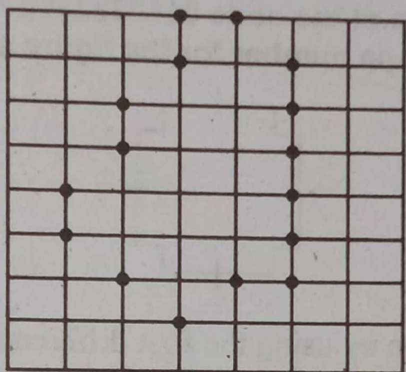
A boundary code formed as a sequence of such directional members is referred to as **Freeman Chain Code**. Digital images usually are required and processed in a grid format with equal spacing in the x- and y-directions, so a chain code can be generated by following a boundary in, say a clockwise direction and assigning a direction to the segments connecting every pair of pixels. This method generally is unacceptable for two principal reasons:

1) The resulting chain tends to be quite long and
2) Any small disturbances along the boundary due to noise or imperfect segmentation cause changes in the code that may not be related to the principal shape features of the boundary.

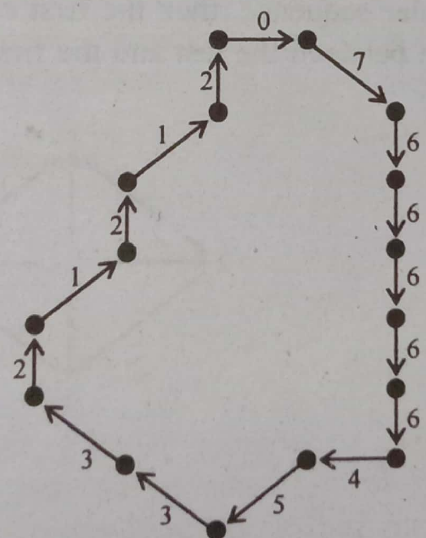
An approach frequently used to circumvent these problems is to resample the boundary by selecting a larger grid spacing as shown below:



Then, as the boundary is traversed, a boundary point is assigned to each node of the large grid, depending on the proximity of the original boundary to that node, as shown in below.



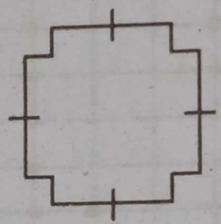
The resampled boundary obtained in this way can be represented by a 4- or 8-code. Fig. below shows the coarser boundary points represented by an 8-directional chain code. It is a simple matter to convert from an 8-code to 4-code and vice-versa.



The starting point in the figure is at the topmost, leftmost point of the boundary which gives the chain code 0766 .... 12. As might be expected, the accuracy of the resulting code representation depends on the spacing of the sampling grid.

The chain code of a boundary depends on the starting point. However, the code can be normalized with respect to the starting point by a straight forward procedure. We simply treat the chain code as a circular sequence of direction numbers and redefine the starting point so that the resulting sequence of numbers forms an integer of minimum magnitude. We can normalize also for rotation by using the first difference of the chain code instead of the code itself. This difference is obtained by counting the number of direction changes that separate two adjacent elements of the code. For instance, the first difference of the 4-direction chain code 10103322 is 3133033. If we treat the code as a circular sequence to normalize with respect to the starting point, then the first element of the difference is computed by using the transition between the last and first components of the chain. Here the result is 33133030. Size normalization can be achieved by altering the size of the resampling grid.

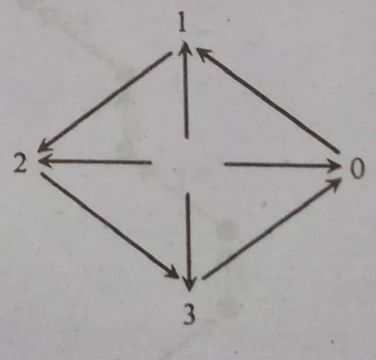
2. a) Show that the first difference of a chain code normalizes it to rotation.
b) Compute the first difference of the code 0110233210332322111.
c) What is the order of the shape number for the figure shown?
d) Obtain the shape number. **[WBUT 2012]**



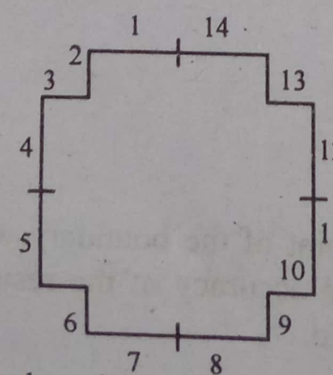
Answer:

a) We can normalize for rotation by using the first difference of the chain code instead of the code itself. The difference is obtained by counting the number of direction changes (in counter clockwise direction) that separate two adjacent elements of the code. For example, the first difference of the 4-direction chain code 10103322 is 3133030. If we elect to treat the code as a circular sequence, then the first element of the difference is computed by using the transition between the last and the first components of the chain. The result is 33133030.

b) 0110233210332322111
The first difference will be
1032103333303130300
by using above diagram.



c)



The order of the shape no. of the above Fig. is 14.

d) Chain code 00303323221211010
Difference 0313031303130313

The shape no. will be smallest magnitude i.e., 0313031303130313

3. a) What do you mean by image recognition? **[WBUT 2017]**
b) Describe any image recognition technique.

Answer:

a) Image recognition also known as pattern recognition, in the context of machine vision, is the ability of software to identify objects, places, people writing and action in images. Computers can use machine vision technologies in combination with a camera and artificial intelligence software to achieve image recognition.

b) Image recognition is the process of identifying and detecting an object or a feature in digital image. This concept is used in many applications. Typical image recognition algorithms include:

- Optical Character recognition
- Pattern matching
- Face recognition
- License plate matching
- Scene change detection

Pattern Matching:

A Pattern is an arrangement of descriptors. The name feature is used often in the pattern recognition literature to denote a descriptor. A pattern, class is a family of patterns that share some common properties. Patterns classes are denoted $w_1, w_2, \ldots\ldots w_n$ where w is the number of classes. Pattern recognition by machine involves techniques for assigning patterns to their respective classes automatically and with a little human intervention as possible. Three common patterns are vectors, strings and trees.

4. Write short notes on the following:
a) Morphology **[WBUT 2009]**
b) Hit-or-miss Transform **[WBUT 2014, 2017]**
c) Erosion and Dilation **[WBUT 2015, 2018]**
d) Chain code **[WBUT 2016, 2018]**

Answer:

a) Morphology:

Morphology is a tool for extracting image components useful in the representation and description of region shape, such as boundaries, skeletons and convex hulls. The language of mathematical morphology is set theory and as such it can apply directly to binary images. Basic operations of mathematical morphology operate on two sets. The first one is the image and the second one is the structuring element sometimes called Kernel. The structuring element used in practise is generally much smaller than the image, often a 3×3 matrix. Erosion and dilation are two basic operators in mathematical morphology.

b) Hit-or-miss Transform:
The hit-and-miss transform is a general binary morphological operation that can be used to look for particular pattern of foreground and background pixels in an image. It is actually the basic operation of binary morphology since almost all the other binary morphological operators can be derived from it. As with other binary morphological operators it takes as input a binary image and a structuring element and produces another binary image as output. The structuring element used in the hit-and-miss is a slight extension to the type that has been introduced for erosion and dilation, in that it contains both foreground and background pixels rather than just foreground pixels i.e. both ones and zeros. The hit-and-miss operation is performed in much the same way as other morphological operators, by translating the origin of the structuring element to all points in the image and then comparing the structuring element with the underlying image pixels.

c) Erosion and Dilation:
Dilation and Erosion are two fundamental morphological operations. Dilation adds pixel to the boundaries of objects in an image, while erosion removes pixel on object boundaries. The number of pixels added to removed from the objects in an image depends on the size and shape of the structuring element used to process the image.

d) Chain code: ***Refer to Question No. 1 of Long Answer Type Questions.***

QUESTION 2015

GROUP – A
(Multiple Choice Type Questions)

1. Answer any *ten* questions:

i) The computation of Walsh coefficient involves
   a) only subtraction
   b) only addition
   ✓c) addition and subtraction
   d) none of the above

ii) Compression is achieved by
   a) filtering
   b) multiplying an array
   ✓c) minimizing the redundancies
   d) none of these

iii) Which of the following is not an image file format?
   a) TIFF
   b) BMP
   c) GIF
   ✓d) none of these

iv) The transform which is widely used in JPEG compression scheme is
   a) FFT
   b) IDFT
   c) Hadamard Transform
   ✓d) Discrete Cosine Transform

v) Wavelet-based image-compression scheme overcomes the problem of
   a) Blur object
   ✓b) Blocking artifact
   c) Background sharpening
   d) all of the above

vi) Usually frequency domain operations are
   a) Global operation
   ✓b) Mask operation
   c) Point operation
   d) none of the above

vii) The smallest units of a digital image is represented by
   a) A one dimensional matrix
   b) Logic 0 or 1
   c) Dot
   ✓d) Pixel

viii) Which of the following statement is true?
   a) The resolution of CMOS image sensor is better than CCD image
   ✓b) The resolution of CCD image sensor is better than CMOS image
   c) CCD image sensor is cost effective than CMOS image
   d) none of the above

ix) What operation is needed for background removal?
a) Image multiplication
b) Image addition
c) Image subtraction
✓d) all of these

x) Wavelets are
a) of remarkable advantage for image compression
✓b) foundation for representing images in various degrees of resolution
c) both (a) and (b)
d) none of these

xi) An example of volume image is
a) A one dimensional image is
b) A two dimensional image
✓c) A three dimensional image
d) all of the above

xii) Basic principle of compression matches the principle of
a) Channel coding
✓b) Source coding
c) Line coding
d) none of these

GROUP – B
(Short Answer Type Questions)

2. Explain Unsharp masking and High boost filtering. Write the expression of Laplacian operator for an image of two variables.
See Topic: IMAGE ENHANCEMENT, Short Answer Type Question No. 7.

3. If an image is rotated by an angle $\frac{\pi}{4}$, will there be any change in the histogram of that image? Justify your answer.
See Topic: DIGITAL IMAGE FORMATION, Short Answer Type Question No. 5.

4. Explain bilinear interpolation method used for image zooming.
See Topic: INTRODUCTION, Short Answer Type Question No. 4.

5. Explain the operation of an order statistics filter and classify it.
See Topic: MORPHOLOGICAL IMAGE PROCESSING, Short Answer Type Question No. 2.

6. Discuss briefly about the usefulness of discrete cosine transform.
See Topic: MATHEMATICAL PRELIMINARIES, Short Answer Type Question No. 4.

GROUP – C
(Long Answer Type Questions)

7. a) What is Histogram of an image and why processing? What do you mean by Histogram specification?

b) What is the difference between spatial domain filtering and frequency domain filtering? Describe edge detection operator and contrast stretching of an image.

c) Briefly describe the smoothing linear spatial filtering. What is bit plane slicing method?

See Topic: IMAGE ENHANCEMENT, Long Answer Type Question No. 7(a), (b) & (c).

8. a) What is m-connectivity among pixels? Give an example.

b) What do you mean by global and local thresholding? What is the basic difference between region growing and split merge technique?

c) What is Hough transform and where is it used?

a) See Topic: MATHEMATICAL PRELIMINARIES, Short Answer Type Question No. 8.

b) See Topic: IMAGE SEGMENTATION, Long Answer Type Question No. 2.

c) See Topic: MATHEMATICAL PRELIMINARIES, Short Answer Type Question No. 5.

9. a) Why image compression is required?

b) Take any 4×4 gray image of your choice and show how block truncation coding can be applied to compress it.

c) Distinguish between image enhancement and image restoration.

a) & b) See Topic: IMAGE DATA COMPRESSION, Long Answer Type Question No. 6(a) & (b).

c) See Topic: IMAGE ENHANCEMENT, Short Answer Type Question No. 2.

10. a) Classify different type of redundancies in an image. Mention the different video compression techniques and briefly discuss about any one of them.

b) Discuss about spatial resolution and gray – level resolution.

c) What is Homomorphic filtering? Mention the domain of application of this filter.

a) See Topic: IMAGE DATA COMPRESSION, Long Answer Type Question No. 6(c).

b) & c) See Topic: IMAGE ENHANCEMENT, Long Answer Type Question No. 8.

11. Write short notes on any *three* of the following:

a) Local enhancement

b) Vector quantization method for image compression

c) Piecewise Linear transformation

d) Erosion and Dilation

e) Image sharpening filters

a) See Topic: IMAGE ENHANCEMENT, Long Answer Type Question No. 12(f).

b) See Topic: IMAGE DATA COMPRESSION, Long Answer Type Question No. 8(f).

c) See Topic: MATHEMATICAL PRELIMINARIES, Long Answer Type Question No. 4(f).

d) See Topic: MORPHOLOGICAL IMAGE PROCESSING, Long Answer Type Question No. 4(c).

e) See Topic: IMAGE ENHANCEMENT, Long Answer Type Question No. 12(g).

QUESTION 2016

GROUP – A

(Multiple Choice Type Questions)

1. Choose the correct alternatives for any *ten* of the following:

i) The spectrum of the visible light is
a) 10-350 nm
✓b) 380-760 nm
c) 760 nm and above
d) none of these

ii) The image acquisition process is used to
a) store image
b) transform image
c) display image
✓d) none of these

iii) Colour image is represented by
a) 2-bit
b) 8-bit
✓c) 24-bit
d) 64-bit

iv) Which of the following operations is idempotent?
a) dilation
b) erosion
c) convolution
✓d) closing

v) If the DFT of $\{f(m,n) \rightarrow F(k,l)\}$ then for the sequence $f(m-m_0,n)$ the DFT becomes
a) $e^{(j\pi m_0 k)/N}.F(k,l)$
b) $e^{-(j\pi m_0 k)/N}.F(k,l)$
✓c) $e^{(j2\pi m_0 k)/N}.F(k,l)$
d) $e^{-(j2\pi m_0 k)/N}.F(k,l)$

vi) The kernel for 2D-DFT with $N=4$ is calculated by the command
a) $dft2(4)$
✓b) $fft2(4)$
c) $dftmtx(4)$
d) none of these

vii) DCT is widely used for
a) image degradation
✓b) image compression
c) image restoration
d) none of these

viii) The photosensitive detector of the human eye is
a) eye lens
b) iris
✓c) retina
d) cornea

ix) Through decimated by 2 operations the sampling rate is
a) increased
✓b) decreased
c) none of these
d) same

x) The histogram equalization process
a) blurs the image
b) fades the image
✓c) improves the brightness of the image
d) none of these

xi) If the size of the mask for averaging is increased, the image will be

a) noise free
✓b) blurred
c) degraded
d) none of these

xii) Salt and pepper noise can be removed by

a) weighted average filter
b) Gaussian filter
✓c) Median filter
d) high boost filter

xiii) An example of dictionary based coding technique is

a) Huffman
b) run-length
✓c) Lempel-Ziv-Welch
d) predictive

xiv) Edge detection of an image broadly means

a) low spatial frequency enhancement
b) high frequency enhancement
c) thresholding low spatial frequency
✓d) detection of intensity variation

xv) The classical Hough transform is concerned with the identification of

✓a) lines in an image
b) zeros in an image
c) poles in an image
d) boundary in an image

xvi) A Wavelet transform is a special case of

a) Laplace transforms
b) Z-transform
✓c) Fourier transform
d) discrete cosine transform

GROUP – B

(Short Answer Type Questions)

2. a) Differentiate between image and scene.
b) What are the image sensors and how are they used?

See Topic: INTRODUCTION, Short Answer Type Question No. 10(a) & (b).

3. a) What do you understand by the terms 'histogram' and 'histogram equalization'?
b) What are the steps to perform histogram equalization?
c) What is the necessity for histogram equalization?

a) See Topic: IMAGE ENHANCEMENT, Long Answer Type Question No. 4.
b) & c) See Topic: IMAGE ENHANCEMENT, Short Answer Type Question No. 17(a) & (b).

4. a) What do you mean by aliasing in the context of image sampling? Explain.
b) What do you mean by the term 'image file format'? Mention some of the frequently used image file format.

See Topic: INTRODUCTION, Short Answer Type Question No. 11(a) & (b).

5. a) Briefly explain skeletons.
b) Write down the effects of the dilation and erosion process.
See Topic: MORPHOLOGICAL IMAGE PROCESSING, Short Answer Type Question No. 3(a) & (b).

6. Show how Prewitt and Sobel operators can be used to smooth the noise.
See Topic: IMAGE ENHANCEMENT, Short Answer Type Question No. 18.

GROUP – C
(Long Answer Type Questions)

7. a) With neat sketch explain briefly about the components of image processing system.
b) Explain the process of image enhancement using arithmetic operator.
c) What is the need for interpolation technique? Consider the image $F = \begin{pmatrix} 2 & 2 \\ 3 & 2 \end{pmatrix}$. Calculate the mean and entropy of the image.
a) See Topic: INTRODUCTION, Short Answer Type Question No. 7.
b) & c) See Topic: IMAGE ENHANCEMENT, Long Answer Type Question No. 9(a) & (b).

8. a) What is 8 bit colour image? For what purpose could it be used? Explain.
b) Explain the model of image restoration process in presence of noise.
c) Explain the process of region extraction based on segmentation.
a) See Topic: INTRODUCTION, Short Answer Type Question No. 8.
b) See Topic: IMAGE ENHANCEMENT, Long Answer Type Question No. 9(c).
c) See Topic: IMAGE SEGMENTATION, Long Answer Type Question No. 3(a).

9. a) With neat sketch, explain the object recognition process.
b) Explain the lines and edges detecting process with mathematical models.
c) What is pattern fritting approach?
a) & c) See Topic: IMAGE SEGMENTATION, Long Answer Type Question No. 3(b) & (c).
b) See Topic: MORPHOLOGICAL IMAGE PROCESSING, Short Answer Type Question No. 4.

10. a) What is the need for compression? What are different compression methods?
b) Discuss briefly the concept of Huffman coding.
c) Define interpixel redundancy?
See Topic: IMAGE DATA COMPRESSION, Long Answer Type Question No. 7(a), (b) & (c).

11. Write short notes on any *three* of the following:
a) Walsh Transformation
b) Image Restoration model
c) Hotelling Transformation
d) Run length coding
e) Chain code

a) See Topic: **MATHEMATICAL PRELIMINARIES, Long Answer Type Question No. 4(d).**
b) See Topic: **IMAGE RESTORATION, Long Answer Type Question No. 3(a).**
c) See Topic: **MATHEMATICAL PRELIMINARIES, Long Answer Type Question No. 4(g).**
d) See Topic: **IMAGE DATA COMPRESSION, Long Answer Type Question No. 8(g).**
e) See Topic: **MORPHOLOGICAL IMAGE PROCESSING, Long Answer Type Question No. 4(d).**

QUESTION 2017

GROUP – A
(Multiple Choice Type Questions)

1. Choose the correct alternatives for any *ten* of the following:

i) What is the range of subjective brightness for human?
a) scotopic threshold to glare limit
✓b) photopic threshold to glare limit
c) scotopic threshold to infinity
d) photopic threshold to infinity

ii) Representation and description almost always follow the output of a
✓a) segmentation stage
b) filtering stage
c) compression stage
d) all of these

iii) An image is a 2D array of
✓a) digital data
b) electrical signals
c) photographic objects
d) light signals

iv) The image function $f(x, y)$ is characterized by two components:
$f(x,y) = i(x,y).r(x,y)$ where
a) $0 < i(x,y) < 1$ & $0 < r(x,y) < \infty$
b) $0 < i(x,y) < 1$ & $0 < r(x,y) < 1$
c) $0 < i(x,y) < \infty$ & $0 < r(x,y) < \infty$
✓d) $0 < i(x,y) < \infty$ & $0 < r(x,y) < 1$

v) The computation of Walsh coefficient involves
a) only subtraction
b) only addition
✓c) addition and subtraction
d) none of these

vi) Which of the following uses 2 × 2 mask for edge detection?
a) Sobel
✓b) Roberts
c) Prewitts
d) Kirsh

vii) The morphological operation used to isolate objects which may be just touching one another is
a) dilation
✓b) erosion
c) opening
d) closing

viii) The transform which is widely used in JPEG compression scheme is

a) FFT
b) IDET
c) Hadamard transform
✓d) discrete cosine transform

ix) Which of the following transforms is used for line detection in image processing?

a) hadamard
✓b) haugh
c) haar
d) slant

x) Which of the following image processing transforms does not satisfy the separability property?

a) Walsh
b) Fourier
c) DCT
✓d) Hotelling

xi) Which of the following is not an image file format?

a) TIFF
b) BMP
c) GIF
✓d) none of these

xii) Which one of the following is the lossy image compression?

✓a) TIFF
b) BMP
c) GIF
d) none of these

GROUP – B

(Short Answer Type Questions)

2. What is Weber ratio? Show the variation of Weber ratio for human eye.

See Topic: INTRODUCTION, Short Answer Type Question No. 9.

3. If an image is rotated by an angle of 60° will there be any change in the histogram of that image? Justify your answer.

See Topic: DIGITAL IMAGE FORMATION, Short Answer Type Question No. 5.

4. Explain the operation of a median filter.

See Topic: IMAGE ENHANCEMENT, Short Answer Type Question No. 8.

5. Distinguish between image enhancement and image restoration.

See Topic: IMAGE ENHANCEMENT, Short Answer Type Question No. 2(1st Part).

6. What is boundary extraction in Morphological image operation?

See Topic: MORPHOLOGICAL IMAGE PROCESSING, Short Answer Type Question No. 1.

GROUP – C

(Long Answer Type Questions)

7. a) Consider the image segment

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

Let $v=\{0,1\}$. Compute the (i) shortest 4-adjacency distance and (ii) m-adjacency distance between the pixels p and q.

b) Discuss one technique of image acquisition.

c) Write down the key stages of digital image processing and explain them.

a) See Topic: IMAGE SEGMENTATION, Short Answer Type Question No. 4.

b) See Topic: INTRODUCTION, Short Answer Type Question No. 12(a).

c) See Topic: INTRODUCTION, Short Answer Type Question No. 7.

8. a) What do you mean by image negative? Explain.

b) Explain Intensity slicing with example.

c) Why do we need Log Transformation in dynamic range compression?

See Topic: IMAGE ENHANCEMENT, Long Answer Type Question No. 5(a), (b) & (c).

9. a) Write down the usefulness of segmentation.

b) Describe the region growing technique for image segmentation and mention the problem associated with it.

c) Discuss the Hough transform method for edge-linking.

a) See Topic: IMAGE SEGMENTATION, Short Answer Type Question No. 3.

b) See Topic: IMAGE RESTORATION, Short Answer Type Question No. 3.

c) See Topic: MATHEMATICAL PRELIMINARIES, Short Answer Type Question No. 5.

10. a) What do you mean by image recognition?

b) What is meant by classification of image?

c) Describe any image recognition technique.

a) & c) See Topic: MORPHOLOGICAL IMAGE PROCESSING, Long Answer Type Question No. 3(a) & (b).

b) See Topic: INTRODUCTION, Short Answer Type Question No. 12(b).

11. Write short notes on any *three* of the following:

   a) Edge detection operation
   b) Hit-or-miss transformation
   c) Smoothing linear spatial filtering
   d) Brightness adaptation
   e) Image compression techniques

a) See Topic: IMAGE SEGMENTATION, Long Answer Type Question No. 6(b).

b) See Topic: MORPHOLOGICAL IMAGE PROCESSING, Long Answer Type Question No. 4(b).

c) See Topic: IMAGE ENHANCEMENT, Long Answer Type Question No. 12(h).

d) See Topic: INTRODUCTION, Long Answer Type Question No. 1(b).

e) See Topic: IMAGE DATA COMPRESSION, Long Answer Type Question No. 8(h).

QUESTION 2018

GROUP – A
(Multiple Choice Type Questions)

1. Choose the correct alternatives for *any ten* of the following:

i) Huffman coding approach reduces
✓a) coding redundancy only
b) inter-pixel redundancy only
c) coding and inter-pixel redundancy
d) psycho-visual redundancy only

ii) The total amount of energy that flows from the light source and it is usually measured in watts (W) is called
a) radiance
✓b) luminance
c) reflectance
d) none of these

iii) The effect caused by the use of an insufficient number of gray levels in smooth areas of a digital image is called
✓a) false contouring
b) gray level slicing
c) bit plane
d) thinning

iv) In 8 distance measurement system distance between centre pixel and a corner pixel is
✓a) 2 units
b) $\sqrt{2}$ units
c) 1 unit
d) 1.5 units

v) Edge detection of an image broadly means
a) low spatial frequency enhancement
b) high spatial frequency enhancement
c) thresholding low spatial frequencies
✓d) none of these

vi) The classical Hough transform is concerned with the identification of
✓a) lines in an image
b) zeros in an image
c) poles in an image
d) none of these

vii) The basic principle of compression matches the principle of
a) channel coding
b) line coding
✓c) source coding
d) all of these

viii) A Wavelet transform is a special case of
a) Laplace transform
b) Z-transform
✓c) Fourier transform
d) none of these

ix) Averaging filter is used for
a) sharpening
b) contrast
c) brightness
✓d) smoothing

x) Wiener filter is used for
a) restoration
b) smoothening
c) sharpening
✓d) none of these

xi) If a function $f(x,y)$ is finite in the space domain, the Fourier transform of $f(x,y)$ will be
✓a) finite
b) infinite
c) undefined
d) zero

xii) A spatial averaging filter in which all the coefficients are equal is called
a) weighted averaging filter
b) median filter
✓c) box filter
d) none of these

GROUP – B
(Short Answer Type Questions)

2. What are the basic steps involved in the image geometrical transformation? Develop the homogeneous form of this transformation.
See Topic: DIGITAL IMAGE FORMATION, Short Answer Type Question No. 3.

3. What is Homomorphic Filtering? Set up the equation for this filter.
See Topic: IMAGE ENHANCEMENT, Short Answer Type Question No. 16.

4. Brightness discrimination is poor at low levels of illumination. Explain.
See Topic: INTRODUCTION, Short Answer Type Question No. 13.

5. What is m-connectivity among pixels? Explain with example.
See Topic: MATHEMATICAL PRELIMINARIES, Short Answer Type Question No. 8.

6. Discuss briefly the Huffman coding.
See Topic: IMAGE DATA COMPRESSION, Long Answer Type Question No. 7(b).

GROUP – C
(Long Answer Type Questions)

7. a) With appropriate example discuss region split technique for segmentation.
b) How isolated point and line can be identified?
c) Consider a 3 × 3 image given below and show that Laplacian mask is an isotropic filter.

1	2	1
2	4	2
4	8	4

a) & b) See Topic: IMAGE SEGMENTATION, Long Answer Type Question No. 4(a) & (b).
c) See Topic: IMAGE ENHANCEMENT, Long Answer Type Question No. 10(a).

8. a) Given below is a 3 × 3 image. What will be the value of the centre pixel when this image is passed through a
(i) Arithmetic mean filter
(ii) Harmonic mean filter
(iii) Max-filter
(iv) Min-filter

1	7	5
6	2	3
1	4	2

b) Contrast and compare the spatial domain filtering and frequency domain filtering.
c) What is unitary transform? Define and compute the equations for unitary transform, fourier transform for both 1-D and 2-D image.

a) & b) See Topic: IMAGE ENHANCEMENT IN THE SPATIAL AND FREQUENCY DOMAIN, Long Answer Type Question No. 10(b) & (c).
c) See Topic: MATHEMATICAL PRELIMINARIES, Long Answer Type Question No. 2.

9. a) What do you mean by redundancy? Name different types of image compression.
b) What is compression ratio? An image is 8MB before compression and 2MB after compression. Find the compression ratio and savings percentage.
c) Briefly describe the smoothing linear spatial filtering. What is bit plane slicing method?

a) See Topic: IMAGE DATA COMPRESSION, Long Answer Type Question No. 5(a).
b) 1st Part: See Topic: IMAGE DATA COMPRESSION, Long Answer Type Question No. 5(b).
2nd Part: See Topic: IMAGE DATA COMPRESSION, Short Answer Type Question No. 4.
c) See Topic: IMAGE ENHANCEMENT, Short Answer Type Question No. 10.

10. a) What is boundary extraction in morphological image operation?
b) What is 8 bit colour image? For what purpose could it be used?
c) What is the need for interpolation technique? Consider a 2 × 2 image given below: Calculate the mean and entropy of the image.

2	2
3	2

a) See Topic: MORPHOLOGICAL IMAGE PROCESSING, Short Answer Type Question No. 1.
b) See Topic: INTRODUCTION, Short Answer Type Question No. 8.
c) See Topic: IMAGE ENHANCEMENT, Long Answer Type Question No. 9(b).

11. Write short notes on *any three* of the following:
a) LZW compression
b) Chain code
c) Dilation and erosion
d) Weber ratio

a) See Topic: **IMAGE DATA COMPRESSION**, Long Answer Type Question No. 8(i).
b) See Topic: **MORPHOLOGICAL IMAGE PROCESSING**, Long Answer Type Question No. 4(d).
c) See Topic: **MORPHOLOGICAL IMAGE PROCESSING**, Long Answer Type Question No. 4(c).
d) See Topic: **INTRODUCTION**, Long Answer Type Question No. 1(c).

QUESTION 2019

GROUP – A
(Multiple Choice Type Questions)

1. Choose the correct alternatives for *any ten* of the following:

i) Each element of image matrix is called
a) dots b) coordinates ✓c) pixel d) value

ii) Histogram is the technique processed in
a) intensity domain b) frequency domain
✓c) spatial domain d) undefined domain

iii) Intensity levels in 8-bit image are
a) 128 b) 255 ✓c) 256 d) 512

iv) What would the time be at 3000K baud, a representative medium speed of a phone DSL (Digital Subscriber Line) connection?
a) 3 sec b) 3 mins ✓c) 3.5 sec d) 3.5 mins

v) What is the basis of numerous spatial domain processing?
a) Transformation b) Scaling ✓c) Histogram d) None of these

vi) Sampling of an image is required for
✓a) Quantization b) Sharpening c) Smoothing d) Digitization

vii) The objective of the sharpening filter is
✓a) highlight the intensity transition b) highlight the low transition
c) highlight the bright transition d) highlight the colour transition

viii) Which is the image processing technique used to improve the quality of image for human viewing?
a) Compression ✓b) Enhancement c) Restoration d) Analysis

ix) Zero padding of a signal
a) reduces aliasing
b) increases frequency
✓c) increases time resolution
d) has no effect

x) The process of extracting information from the image is called as
a) image enhancement
b) image restoration
✓c) image analysis
d) image compression

xi) CAT in imaging stands for
a) Computer Aided Telegraphy
b) Computer Aided Topography
c) Computerized Axial Telegraphy
✓d) Computerized Axial Tomography

xii) A pixel p at coordinate (x, y) has four horizontal and vertical neighbors whose coordinates are given by
a) $(x-1,\ y-1),\ (x+1,\ y+1),\ (x+1,\ y),\ (x,\ y+1)$
✓b) $(x+1,\ y),\ (x,\ y+1),\ (x-1,\ y),\ (x,\ y-1)$
c) $(x+1,\ y+1),\ (x+1,\ y),\ (x,\ y+1),\ (x,\ y)$
d) $(x,\ y-1),\ (x-1,\ y)(x-1,\ y),\ (x-1,\ y-1)$

GROUP – B
(Short Answer Type Questions)

2. Define digital image and explain image and explain image pixel. What are the storage requirements for (500×500) & (1024×1024) binary images?
See Topic: INTRODUCTION, Short Answer Type Question No. 14.

3. Explain what you understand about image histogram of a digital image. What is the DC component of the following image?

$$f = \begin{bmatrix} 1 & 3 & 4 \\ 5 & 6 & 7 \\ 8 & 9 & 11 \end{bmatrix}$$

Suggest one way of generating colour histograms in the RGB colour space from images. What are the number histogram components in such a histogram?
See.Topic: IMAGE ENHANCEMENT, Short Answer Type Question No. 19.

4. Define connectivity. What is the difference between 8-connectivity and m-connectivity? What are the different methods for calculating the distance of pixel? Explain with relevant example.
See Topic: MATHEMATICAL PRELIMINARIES, Short Answer Type Question No. 9.

5. What are spatial domain and frequency domain technique? Explain Mask or Kernels. What is Median Filter? What is Image Enhancement? Explain gray level slicing.
See Topic: IMAGE ENHANCEMENT, Short Answer Type Question No. 20.

6. Distinguish image enhancement and restoration. Explain CMY and CMYK colour model.
1st Part: See Topic: IMAGE ENHANCEMENT, Short Answer Type Question No. 2.
2nd Part: See Topic: DIGITAL IMAGE FORMATION, Short Answer Type Question No. 6.

GROUP – C
(Long Answer Type Questions)

7. a) How is image represented in digital formats?
See Topic: DIGITAL IMAGE FORMATION, Short Answer Type Question No. 7.

b) Explain intensity slicing with example. Explain image negatives. Explain image averaging.
1st Part: See Topic: IMAGE ENHANCEMENT, Long Answer Type Question No. 5(b).
2nd Part: See Topic: IMAGE ENHANCEMENT, Short Answer Type Question No. 3.
3rd Part: See Topic: IMAGE ENHANCEMENT, Long Answer Type Question No. 6(a).

c) What is Log transformation? Why do we need Log transformation in dynamic range compression?
See Topic: IMAGE ENHANCEMENT, Long Answer Type Question No. 5(c).

8. a) What is understood by spatial domain representation? Describe an algorithm for image sharpening using DFT and IDFT. Identify the parameters affecting the sharpening operation in your algorithm.
b) What is Unitary transform? Define and compute the equations for unitary transform for both 1D and 2D image.
a) See Topic: IMAGE ENHANCEMENT, Long Answer Type Question No. 12.
b) See Topic: MATHEMATICAL PRELIMINARIES, Long Answer Type Question No. 2.

9. a) How points and edges are detected using derivatives? Explain with proper example.
b) What is the role of quantization in image processing?
c) Discuss briefly Huffman coding.
a) See Topic: IMAGE SEGMENTATION, Long Answer Type Question No. 5.
b) See Topic: IMAGE DATA COMPRESSION, Short Answer Type Question No. 3.
c) See Topic: IMAGE DATA COMPRESSION, Short Answer Type Question No. 5.

10. a) Draw the schematic of 2D DWT synthesis filter bank for Haar Wavelet transform and explain the components.
b) State the JPEG compression algorithm and draw the schematic diagram of JPEG compressor.
a) See Topic: MATHEMATICAL PRELIMINARIES, Long Answer Type Question No. 3.
b) See Topic: IMAGE DATA COMPRESSION, Long Answer Type Question No. 4.

11. Write short notes on any *three* of the following:
   a) Salt and Pepper noise
   b) Contrast stretching
   c) DCT in Image Compression
   d) Edge detection operators
   e) Wavelet Transform.

a) See Topic: IMAGE ENHANCEMENT, Short Answer Type Question No. 12(i).
b) See Topic: IMAGE RESTORATION, Long Answer Type Question No. 3(c).
c) See Topic: MATHEMATICAL PRELIMINARIES, Long Answer Type Question No. 4(b).
d) See Topic: IMAGE SEGMENTATION, Long Answer Type Question No. 6(b).
e) See Topic: IMAGE DATA COMPRESSION, Long Answer Type Question No. 8(e).