KARPAGAM ACADEMY OF HIGHER EDUCATION

SUBJECT NAME : DIGITAL IMAGE PROCESSING **CLASS : III B.Sc. (IT)**
SUBJECT CODE :16ITU504A **SYLLABUS** **SEMESTER : V**
BATCH (2016-2019)

16ITU504A **DIGITAL IMAGE PROCESSING** **3H – 3C**

Instruction Hours / week: L: 4 T: 0 P: 0 **Marks:** Int : **40** Ext : **60** Total: **100**

SCOPE

The objectives of this course are tomake the students learn the fundamental theories and techniques of digital image processing, cover the fundamental concepts of visual perception and image acquisition, basic techniques of image manipulation, segmentation and coding, and a preliminary understanding of Computer Vision.

OBJECTIVES

- To perform image manipulations and analysis in many different fields.
- To provide students with the ability to apply knowledge of computing, mathematics, science and engineering to solve problems in multidisciplinary research.

UNIT-I

Introduction: Light, Brightness adaption and discrimination, Pixels, coordinate conventions, Imaging Geometry, Perspective Projection, Spatial Domain Filtering, sampling and quantization. Spatial Domain Filtering: Intensity transformations, contrast stretching, histogram equalization, Correlation and convolution, Smoothing filters, sharpening filters, gradient and Laplacian.

UNIT-II

Hotelling Transform, Fourier Transforms and properties, FFT (Decimation in Frequency and Decimation in Time Techniques), Convolution, Correlation, 2-D sampling, Discrete Cosine Transform, Frequency domain filtering.

UNIT-III:

Image Restoration, Basic Framework, Interactive Restoration, Image deformation and geometric transformations, image morphing, Restoration techniques, Noise characterization, Noise restoration filters, Adaptive filters, Linear, Position invariant degradations, Estimation of Degradation functions, Restoration from projections, Image Compression-Encoder-Decoder model, Types of redundancies, Lossy and Lossless compression, Entropy of an information source, Shannon's 1st Theorem, Huffman Coding, Arithmetic Coding, Golomb

Coding, LZW coding, Transform Coding, Sub-image size selection, blocking artifacts, DCT implementation using FFT, Run length coding.

UNIT – IV

FAX compression (ITUITT Group-3 and Group-4), Symbol-based coding, JBIG-2, Bit-plane encoding, Bit-allocation, Zonal Coding, Threshold Coding, JPEG, Lossless predictive coding, Lossy predictive coding, Motion Compensation

Wavelet based Image Compression: Expansion of functions, Multi-resolution analysis, Scaling functions, MRA refinement equation, Wavelet series expansion, Discrete Wavelet Transform (DWT),Continuous Wavelet Transform, Fast Wavelet Transform, 2-D wavelet Transform, JPEG-2000 encoding, Digital Image Watermarking

UNIT-V

Morphological Image Processing: Basics, SE, Erosion, Dilation, Opening, Closing, Hit-or-Miss Transform, Boundary Detection, Hole filling, Connected components, convex hull, thinning, thickening, skeletons,pruning, Geodesic Dilation, Erosion, Reconstruction by dilation and erosion. Image Segmentation: Boundary detection based techniques, Point, line detection, Edge detection, Edge linking, local processing, regional processing, Hough transform, Thresholding, Iterative thresholding,Otsu's method, Moving averages, Multivariable thresholding, Region-based segmentation, Watershed algorithm, Use of motion in segmentation

Suggested Readings

1. Gonzalez, R. C., & Woods, R. E. (2008). Digital Image Processing(3rd ed.). New Delhi: Pearson Education.
2. Jain, A. K. (1989). Fundamentals of Digital image Processing. New Delhi: Prentice Hall of India.
3. Castleman, K. R. (1996). Digital Image Processing. New Delhi: Pearson Education.
4. Schalkoff. (1989). Digital Image Processing and Computer Vision. New York: John Wiley and Sons.
5. Rafael, C. Gonzalez., Richard, E. Woods.,& Steven Eddins. (2004). Digital Image Processing using MATLAB. New Delhi: Pearson Education.

KARPAGAM ACADEMY OF HIGHER EDUCATION
(Deemed to be University)
(Established Under Section 3 of UGC Act 1956)
Pollachi Main Road, Eachanari Post, Coimbatore - 641021
(For the candidates admitted from 2016 onwards)
DEPARTMENT OF CS, CA & IT

SUBJECT : DIGITAL IMAGE PROCESSING **SEMESTER : V**
SUBJECT CODE: 16ITU504A **CLASS : III B.SC(IT)**

LECTURE PLAN

S.NO	LECTURE DURATION (Hour)	TOPICS TO BE COVERED	SUPPORT MATERIALS
		UNIT I	
1	1	**Introduction:** Light, Brightness adaption and Discrimination, Pixels, Coordinate Conventions, Imaging Geometry	S1: 60-64 S1: 88-91 S2:30
2	1	Perspective Projection, Spatial Domain Filtering, Sampling and Quantization	S1: 138-141 S1:74-76 S2:70
3	1	**Spatial Domain Filtering:** Intensity Transforms, Cotrast Stretching	S2:71-73 S2:74
4	1	Histogram Equalization, Correlation and Convolution	S2:76-78 S2:41-42
5	1	Smoothing Filters, Sharpening Filters	S1:141-145 S1:147-150
6	1	Gradient and Laplacian	S1:156 S2:41
7	1	**Recapitulation of Unit I** **Discussion of Important Questions**	
		Total no. of Hours Planned for Unit – I	**7 Hrs**
		UNIT – II	
1	1	**Transform:** Hotelling Transform, Fourier Transforms and properties	S1:172-176 S1:230-235
2	1	FFT (Decimation in frequency and Decimation in time techniques)	S1:183:189
3	1	Convolution, Correlation	S1:227-230
4	1	2-D Sampling, Discrete Cosine Transform	S1:216,230 S2:51-52

5	1	Frequency domain filtering	S1:178-183
6	**1**	**Recapitulation and Discussion of important Questions**	
		Total no. of Hours Planned for Unit – II	**6 Hrs**
		UNIT III	
1	1	**Image Restoration,** Basic Framework, Interactive Restoration, Image deformation and geometric transformations	S1:243 S2:97 S1:292-294
2	1	Image morphing, Restoration techniques, Noise charecterization	S1:244-249 S1:542-544
3	1	Noise restoration filters, Adaptive filters, Linear, Position invariant degradation	S1:259 S1:276-278
4	1	Estimation of Degradation Function, Restoration from projections, Image Compression-Encoder-Decoder model	S1:278-280 S1:443-445
5	1	Types of Redundancies, Lossy and Lossless Compression, Entropy of an information source, Shannon's 1st Theorem	S1:433-441 S2:177 S1:446
6	1	Huffman Coding, Arithmetic Coding, Golomb Coding, LZW Coding, Transform Coding	S2:177,178 S1:468,489
7	1	Subimage size selection, Blocking artifacts, DCT implementation using FFT, Runlenght Coding	S2:54-57 S2:177
8	1	**Recapitulation and discussion of important Questions**	
		Total no. of Hours Planned for Unit – III	**8 Hrs**
		UNIT IV	
1	1	**Fax Compression,** Symbol-based Coding, JBIG-2, Bit Plane encoding	S2:180-181
2	1	Bit allocation, Zonal Coding, Threshold Coding, JPEG	S1:617-624
3	1	Lossless Predictive coding, Lossy Predictive coding, Motion Compensation	S1:478-489

4	1	**Wavelet based image Compression:** Expansion of functions, Multi-resolution analysis, Scaling functions, MRA Refinement equation	S1:385-391 S1:508
5	1	Wavelet series expansion, Discrete wavelet Transform, Fast wavelet transform, Continuous wavelet Transfrom	S1:394-398 S1:401
6	1	2-D wavelet transform, JPEG 2000 encoding, Digital image watermarking	S1:408 S2:160-162
7	1	**Recapitulation and discussion of important Questions**	
		Total no. of Hours Planned for Unit - IV	**8 Hrs**
		UNIT V	
1	1	**Morphological Lmage Processing,** Basics, SE, Erosion, Dilation, Opening, Closing, Hit or miss transform	S1:542-544 S1:545-547,550 S1:554
2	1	Boundary Detection, Hole Filling, Connected components, Convex hull, Thinning, Thickening, Skeleton, Pruning	S1:556-569
3	1	Geodesic Dilation, Erosion, Recontruction by dilation and erosion **Image Segmentation,** Boundary detection based techniques	S1:572-576 S1:639
4	1	Point, Line Detection, Edge Detection, Edge Linking, Local processing, regional processing, Hough transfrom	S1:590-609
5	1	Thresholding, Iterative thresholding, Otsu's method, Moving averages, Multivariable thersholding	S1:617-633
6	1	Region-based segmentation, watershed algorithm, use of motion in segmentation	S1:634-637 S1:644 S1:648-652
7	1	**Recapitulation and discussion of important Questions**	
8	1	**Previous Year ESE Questions Discussion**	
		Total no. of Hours Planned for Unit - V	**8 Hrs**
		Total Planned Hours	**36 Hrs**

SUPPORTING MATERIAL

1. Gonzalez, R.C., & Woods, R.E. (2008). Digital Image Processing(3rd ed.). New Delhi: Pearson Education.
2. Digital Image Processing, An Algorithmic Approach by Madhuri A.Joshi 2006, Prentice hall of India, New Delhi.
3. Jain, A.K.(1989). Fundamentals of Digital image Processing. New Delhi: Prentice Hall of India.
4. Rafal, C. Gonzalez., Richard, E.Woods & Steven Eddins(2004). Digital Image Processing Using MATLAB. New Delhi: Pearson Education.

Websites

W1: http://www.webopedia.com
W2:www.codecademy.com
W3:www.w3schools.com/
W4:www.javascriptkit.com
W5:http://www.webteacher.com/javascript/
W6: http://www.tizag.com/aspTutorial/
W7: http://www.tutorialspoint.com/xml/

UNIT-I

SYLLABUS

Introduction: Light, Brightness adaption and discrimination, Pixels, coordinate conventions, Imaging Geometry, Perspective Projection, Spatial Domain Filtering, sampling and quantization. Spatial Domain Filtering: Intensity transformations, contrast stretching, histogram equalization, Correlation and convolution, Smoothing filters, sharpening filters, gradient and Laplacian.

INTRODUCTION

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

- improvement of pictorial information for human interpretation;
- processing of image data for storage, transmission, and representation for autonomous machine perception.

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

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

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

Note that a digital image is composed of a finite number of elements, each of which has a particular location and value.

These elements are called *picture elements*, *image elements*, *pels*, and *pixels*. *Pixel* is the term used most widely to denote the elements of a digital image.

Vision is the most advanced of our senses, so it is not surprising that images play the single most important role in human perception.

However, unlike humans, who are limited to the visual band of the electromagnetic (EM) spectrum, imaging machines cover almost the entire EM spectrum, ranging from gamma to radio waves.

They can operate on images generated by sources that humans are not accustomed to associating with images. These include ultrasound, electron microscopy, and computer-generated images. Thus, digital image processing encompasses a wide and varied field of applications.

Low-level process is characterized by the fact that both its inputs and outputs are images.

Mid-level processing on images involves tasks such as segmentation (partitioning an image into regions or objects), description of those objects to reduce them to a form suitable for computer processing, and classification (recognition) of individual objects.

A mid-level process is characterized by the fact that its inputs generally are images, but its outputs are attributes extracted from those images (e.g., edges, contours, and the identity of individual objects).

Higher-level processing involves “making sense” of an ensemble of recognized objects, as in image analysis, and, at the far end of the continuum, performing the cognitive functions normally associated with vision.

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

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

The principal energy source for images in use today is the electromagnetic energy spectrum.

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

LIGHT:

Light is a particular type of electromagnetic radiation that can be sensed by the human eye.

Electromagnetic spectrum:

In 1666, Sir Isaac Newton discovered that when a beam of sunlight is passed through a glass prism, the emerging beam of light is not white but consists instead of a continuous spectrum of colors ranging from violet at one end to red at the other. As in Fig. the range of colors we perceive in visible light represents a very small portion of the electromagnetic spectrum. On one end of the spectrum are radio waves with wavelengths billions of times longer than those of visible light. On the other end of the spectrum are gamma rays with wavelengths millions of times smaller than those of visible light. The electromagnetic spectrum can be expressed in terms of wavelength, frequency, or energy.

Wavelength(λ) and frequency(ν) are related by the expression

$$\lambda = \frac{c}{\nu}$$

where c is the speed of light The energy of the various components of the electromagnetic spectrum is given by the expression (2.2-2) where h is Planck's constant.

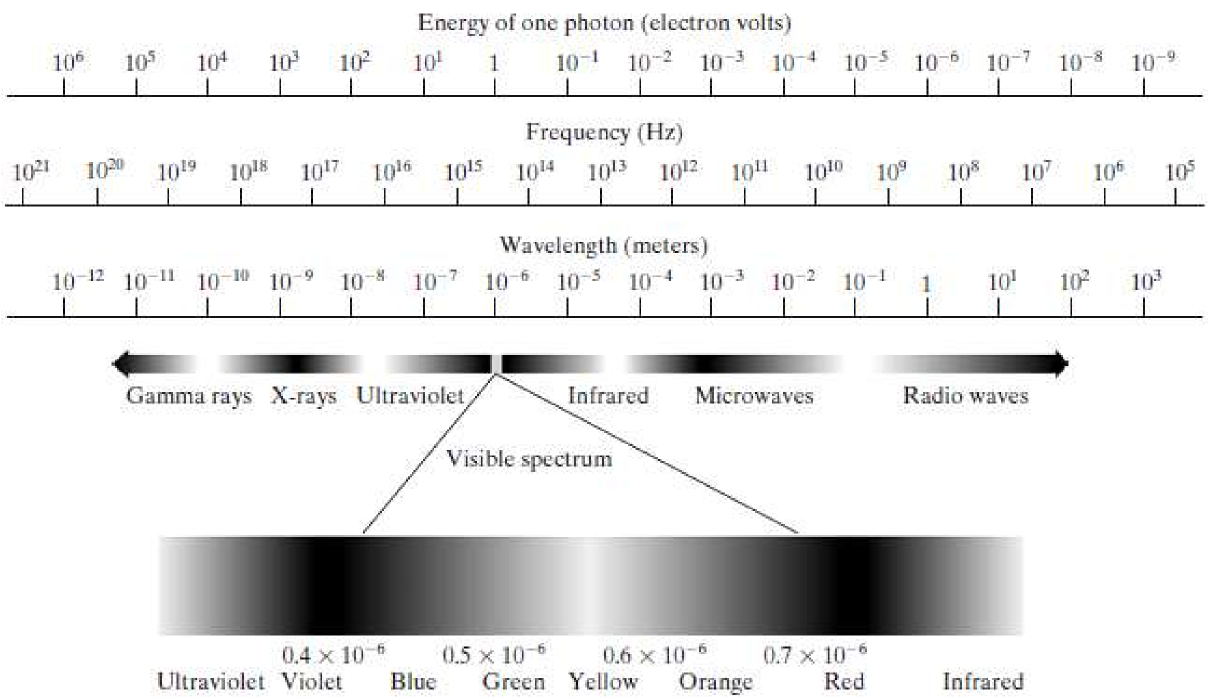


FIGURE 2.10 The electromagnetic spectrum. The visible spectrum is shown zoomed to facilitate explanation, but note that the visible spectrum is a rather narrow portion of the EM spectrum.

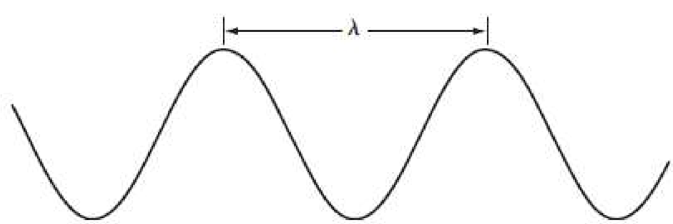
Electromagnetic waves can be conceptualized as propagating sinusoidal waves of varying wavelengths, or they can be thought of as a stream of massless particles, each traveling in a wavelike pattern and moving at the speed of light.

Each massless particle contains a certain amount (or bundle) of energy. Each bundle of energy is called a *photon*.

We see from Eq. (2.2-2) that energy is proportional to frequency, so the higher-frequency (shorter wavelength) electromagnetic phenomena carry more energy per photon. Thus, radio waves have photons with low energies, microwaves have more energy than radio waves, infrared still more, then visible, ultraviolet, X-rays, and finally gamma rays, the most energetic of all. This is the reason why gamma rays are so dangerous to living organisms.

If spectral bands are grouped according to energy per photon, we obtain the spectrum shown in Fig. 1.5, ranging from gamma rays (highest energy) at one end to radio waves (lowest energy) at the other.

FIGURE 2.11 Graphical representation of one wavelength.



Units of measurements

- Frequency is measured in Hertz (Hz)
- Wavelength is measured in meters; also microns (_m or 10□6m) and nanometers (10□9m)
- Energy is measured in electron-volt

- **Photon**
- Massless particles whose stream in a sinusoidal wave pattern forms energy
- Energy is directly proportional to frequency
- Higher frequency energy particles carry more energy

- Radio waves have less energy while gamma rays have more energy, making gamma rays more dangerous to living organisms

The visible band of the electromagnetic spectrum spans the range from approximately (violet) to about (red). For convenience, the color spectrum is divided into six broad regions: violet, blue, green, yellow, orange, and red. No color (or other component of the electromagnetic spectrum) ends abruptly, but rather each range blends smoothly into the next, as shown in Fig. 2.10.

The colors that humans perceive in an object are determined by the nature of the light *reflected* from the object. A body that reflects light relatively balanced in all visible wavelengths appears white to the observer. However, a body that favors reflectance in a limited range of the visible spectrum exhibits some shades of color. For example, green objects reflect light with wavelengths primarily in the 500 to 570 nm range while absorbing most of the energy at other wavelengths.

Light that is void of color is called *monochromatic* (or *achromatic*) *light*.

The only attribute of monochromatic light is its *intensity* or amount.

Because the intensity of monochromatic light is perceived to vary from black to grays and finally to white, the term *gray level* is used commonly to denote monochromatic intensity. We use the terms *intensity* and *gray level* interchangeably in subsequent discussions. The range of measured values of monochromatic light from black to white is usually called the *gray scale*, and monochromatic images are frequently referred to as *gray-scale images*.

Chromatic* (*color*) *light spans the electromagnetic energy spectrum from approximately 0.43 to 0.79 µm as noted previously.

In addition to frequency, three basic quantities are used to describe the quality of a chromatic light source: radiance, luminance, and brightness.

Radiance is the total amount of energy that flows from the light source, and it is usually measured in watts (W).

Luminance**,** measured in lumens (lm), gives a measure of the amount of energy an observer *perceives* from a light source.

For example, light emitted from a source operating in the far infrared region of the spectrum could have significant energy (radiance), but an observer would hardly perceive it; its luminance would be almost zero.

Brightness is a subjective descriptor of light perception that is practically impossible to measure. It embodies the achromatic notion of intensity and is one of the key factors in describing color sensation.

- At the shortwavelength end of the electromagnetic spectrum, we have gamma rays and X-rays. Gamma radiation is important for medical and astronomical imaging, and for imaging radiation in nuclear environments. Hard (high-energy) X-rays are used in industrial applications. Chest and dental X-rays are in the lower energy (soft) end of the X-ray band.

- The soft X-ray band transitions into the far ultraviolet light region, which in turn blends with the visible spectrum at longer wavelengths.

- Moving still higher in wavelength, we encounter the infrared band, which radiates heat, a fact that makes it useful in imaging applications that rely on "heat signatures."

- The part of the infrared band close to the visible spectrum is called the *near-infrared* region.

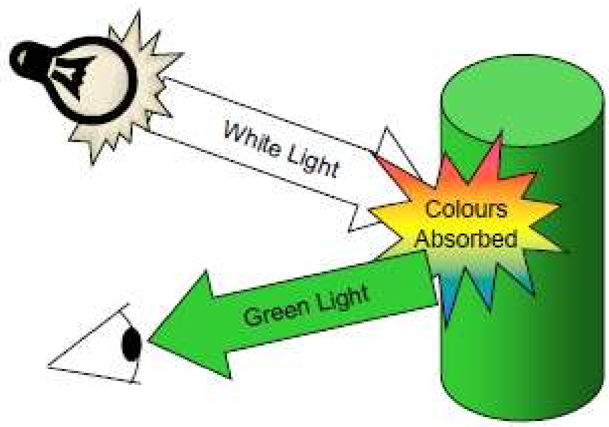
- The opposite end of this band is called the *far-infrared* region.

- This latter region blends with the microwave band. This band is well known as the source of energy in microwave ovens, but it has many other uses, including communication and radar.

- Finally, the radio wave band encompasses television as well as AM and FM radio. In the higher energies, radio signals emanating from certain stellar bodies are useful in astronomical observations.

Light is just a particular part of the electromagnetic spectrum that can be sensed by the human eye

The electromagnetic spectrum is split up according to the wavelengths of different forms of energy. The colours that we perceive are determined by the nature of the light reflected from an object.



For example, if white light is shone onto a green object most wavelengths are absorbed, while green light is reflected from the object

_ Visible spectrum
– 0:43_m (violet) to 0:79_m (red)
– VIBGYOR regions
– Colors are perceived because of light reflected from an object
– Absorption vs reflectance of colors
_ An object appears white because it reflects all colors equally

_ Achromatic or monochromatic light
– No color in light
– Amount of energy describes intensity
_ Quantified by gray level from black through various shades of gray to white
– Monochrome images also called gray scale images

_ Chromatic light

– Spans the energy spectrum from 0.43 to 0.79 _m
– Described by three basic quantities: radiance, luminance, brightness
– Radiance
_ Total amount of energy flowing from a light source
_ Measured in Watts

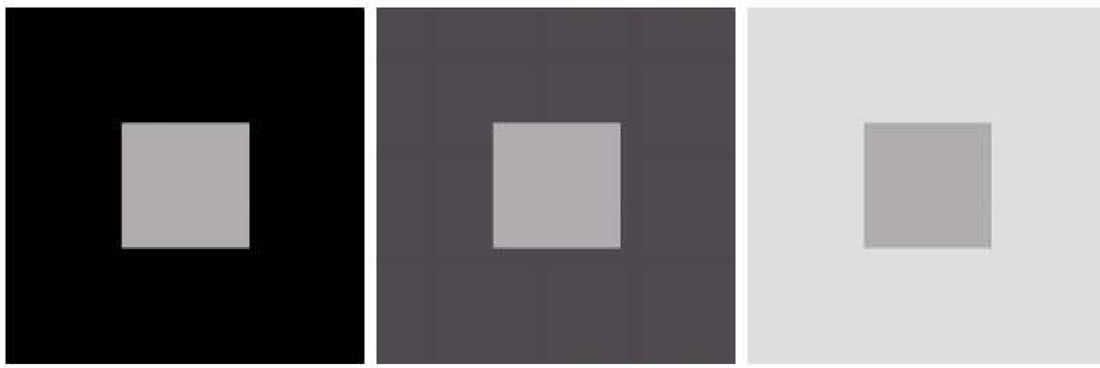
– Luminance
_ Amount of energy perceived by an observer from a light source
_ Measured in lumens (lm)

– Brightness
_ Subjective descriptor of light perception
_ Achromatic notion of intensity
_ Key factor in describing color sensation

BRIGHTNESS ADAPTATION AND DISCRIMINATION

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

The human visual system can perceive approximately 1010 different light intensity levels However, at any one time we can only discriminate between a much smaller number – *brightness adaptation* Similarly, the *perceived intensity* of a region is related to the light intensities of the regions surrounding it



An example of *simultaneous contrast*

Brightness adaptation and discrimination

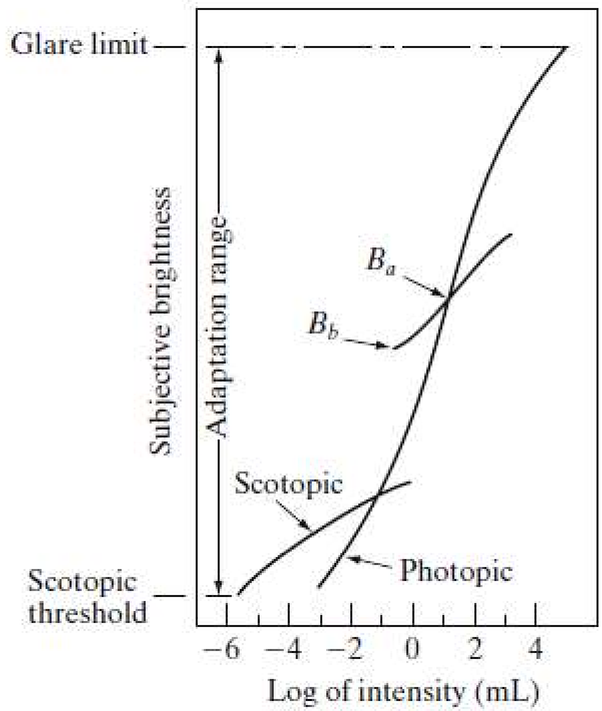
- Digital images displayed as a discrete set of intensities
- Range of human eye is about 10^{10} different light intensity levels, from scotopic threshold to the glare limit
- Subjective brightness (as perceived by humans) is a logarithmic function of light intensity incident on the eye
- Visual system cannot operate over the enormous range simultaneously

Brightness adaptation

- Change in overall sensitivity of perceived brightness
- Number of distinct intensity level that can be perceived simultaneously is small compared to number of levels that can be perceived
- Brightness adaptation level current sensitivity level of the visual system

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

Experimental evidence indicates that *subjective brightness* (intensity as *perceived* by the human visual system) is a logarithmic function of the light intensity incident on the eye.



In fig a plot of light intensity versus subjective brightness, illustrates this characteristic.The long solid curve represents the range of intensities to which the visual system can adapt. In photopic vision alone, the range is about The transition from scotopic to photopic vision is gradual over the approximate range from 0.001 to 0.1 millilambert (to in the log scale), as the double branches of the adaptation curve in this range show.

The essential point in interpreting the impressive dynamic range depicted in Fig. 2.4 is that the visual system cannot operate over such a range *simultaneously.* Rather, it accomplishes this large variation by changing its overall sensitivity, a phenomenon known as *brightness adaptation.*

The total range of distinct intensity levels the eye can discriminate simultaneously is rather small when compared with the total adaptation range. For any given set of conditions, the current sensitivity level of the visual system is called the *brightness adaptation level,* which may correspond, for example, to brightness in Fig. 2.4. The short intersecting curve represents the range of subjective brightness that the eye can perceive when adapted to this level. This range is rather restricted, having a level at and below which all stimuli are perceived as indistinguishable blacks.

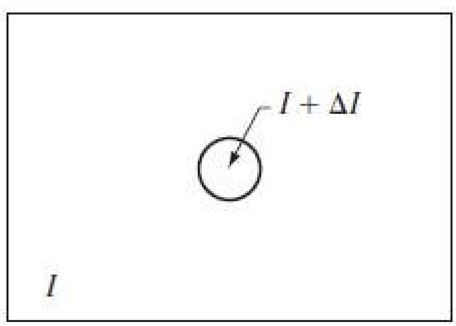
The upper portion of the curve is not actually restricted but, if extended too far, loses its meaning because much higher intensities would simply raise the adaptation level higher than The ability of the eye to discriminate between *changes* in light intensity at any specific adaptation level is also of considerable interest.

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

This area typically is a diffuser, such as opaque glass, that is illuminated from behind by a light source whose intensity, *I*, can be varied. To this field is added an increment of illumination, in the form of a short-duration flash that appears as a circle in the center of the uniformly illuminated field, as Fig. 2.5 shows

Weber ratio

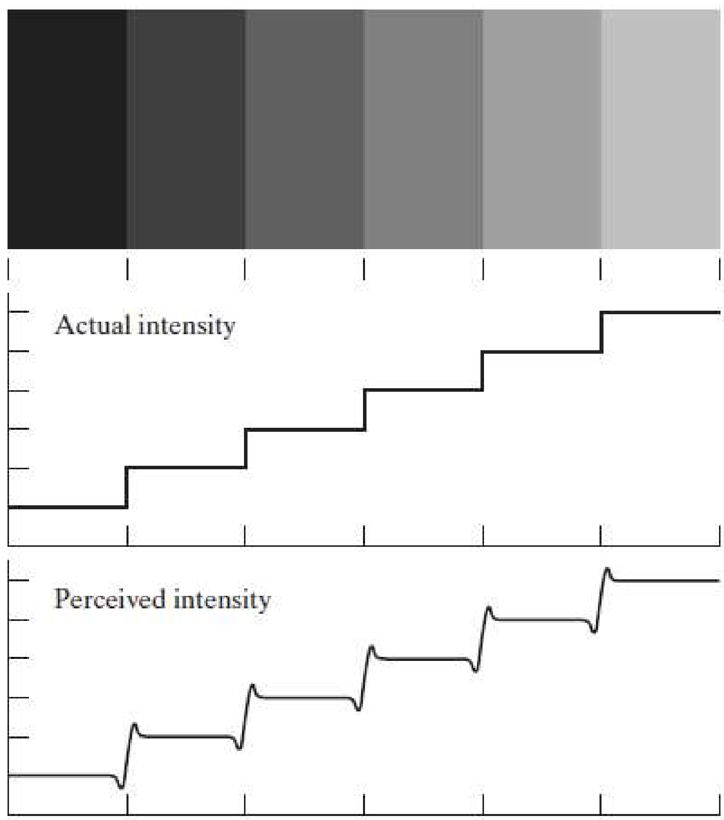
FIGURE 2.5 Basic experimental setup used to characterize brightness discrimination.



- Measure of contrast discrimination ability
- Background intensity given by I
- Increment of illumination for short duration at intensity ΔI (Fig 2.5)
- ΔIc is the increment of illumination when the illumination is visible half the time against background intensity I
- Weber ratio is given by ΔIc/I

- A small value of ΔIc/I implies that a small percentage change in intensity is visible, representing good brightness
- discrimination
- A large value of ΔIc/I implies that a large percentage change is required for discrimination, representing poor
- brightness discrimination
- Typically, brightness discrimination is poor at low levels of illumination and improves at higher levels of background illumination

Mach bands



a
b
c

FIGURE 2.7 Illustration of the Mach band effect. Perceived intensity is not a simple function of actual intensity.

- Brightness pattern near the boundaries shown in stripes of constant intensity (Figure 2.7)
- The bars themselves are useful for calibration of display equipment
  - **Simultaneous contrast**

- A region's perceived brightness does not depend simply on intensity
- Lighter background makes an object appear darker while darker background makes the same object appear brighter (Fig 2.8)
- The second phenomenon, called *simultaneous contrast*, is related to the fact that a region's perceived brightness does not depend simply on its intensity, as Fig. 2.8 demonstrates. All the center squares have exactly the same intensity.
- However, they appear to the eye to become darker as the background gets lighter. A more familiar example is a piece of paper that seems white when lying on a desk, but can appear totally black when used to shield the eyes while looking directly at a bright sky.

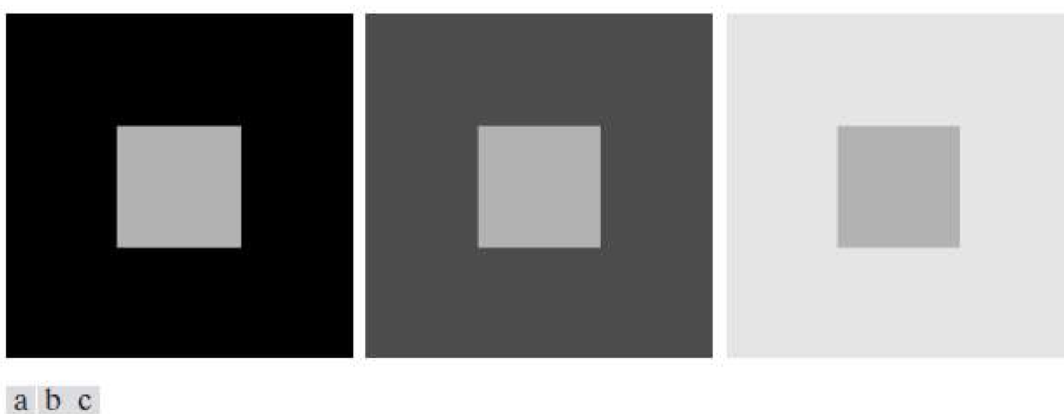


FIGURE 2.8 Examples of simultaneous contrast. All the inner squares have the same intensity, but they appear progressively darker as the background becomes lighter.

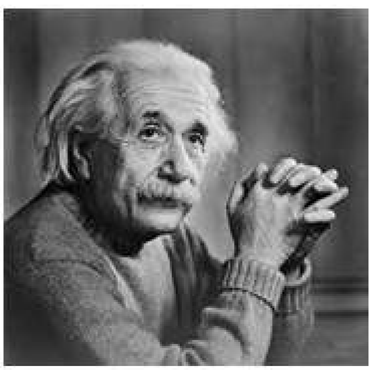
PIXEL

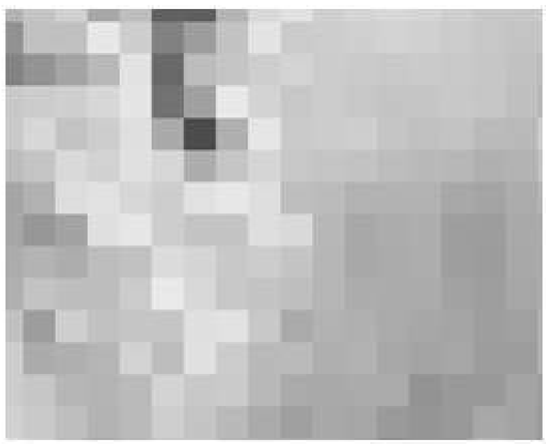
Pixel is the smallest element of an image. Each pixel correspond to any one value. In an 8-bit gray scale image, the value of the pixel between 0 and 255. The value of a pixel at any point correspond to the intensity of the light photons striking at that point. Each pixel store a value proportional to the light intensity at that particular location.

PEL

A pixel is also known as PEL. You can have more understanding of the pixel from the pictures given below.

In the above picture, there may be thousands of pixels, that together make up this image. We will zoom that image to the extent that we are able to see some pixels division. It is shown in the image below.



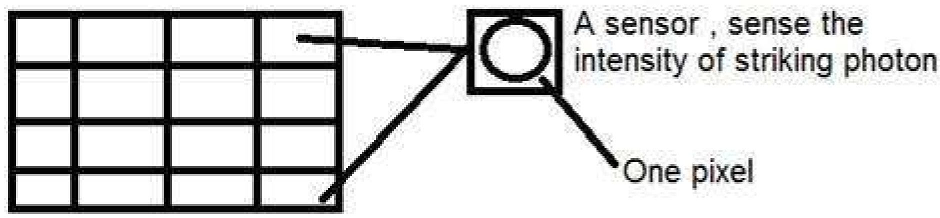


Relationship with CCD array

We have seen that how an image is formed in the CCD array. So a pixel can also be defined as

The smallest division the CCD array is also known as pixel.

Each division of CCD array contains the value against the intensity of the photon striking to it. This value can also be called as a pixel.



Calculation of total number of pixels

We have define an image as a two dimensional signal or matrix. Then in that case the number of PEL would be equal to the number of rows multiply with number of columns.

This can be mathematically represented as below:

Total number of pixels = number of rows (X) number of columns

Or we can say that the number of (x,y) coordinate pairs make up the total number of pixels.

We will look in more detail in the tutorial of image types, that how do we calculate the pixels in a color image.

Gray level

The value of the pixel at any point denotes the intensity of image at that location, and that is also known as gray level.

We will see in more detail about the value of the pixels in the image storage and bits per pixel tutorial, but for now we will just look at the concept of only one pixel value.

Pixel value.(0)

As it has already been define in the beginning of this tutorial, that each pixel can have only one value and each value denotes the intensity of light at that point of the image.

We will now look at a very unique value 0. The value 0 means absence of light. It means that 0 denotes dark, and it further means that when ever a pixel has a value of 0, it means at that point, black color would be formed.

Have a look at this image matrix

0	0	0
0	0	0
0	0	0

Now this image matrix has all filled up with 0. All the pixels have a value of 0. If we were to calculate the total number of pixels form this matrix, this is how we are going to do it.

Total no of pixels = total no. of rows X total no. of columns

= 3 X 3

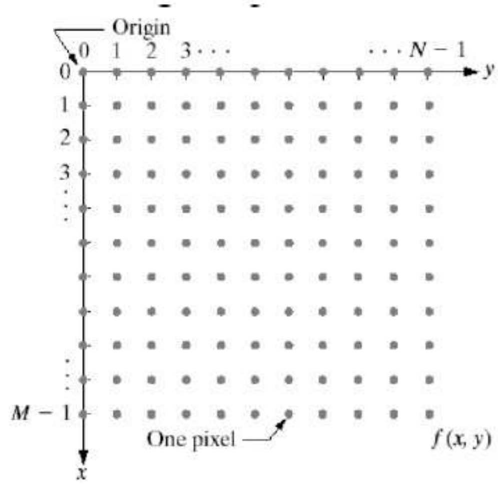
= 9.

It means that an image would be formed with 9 pixels, and that image would have a dimension of 3 rows and 3 column and most importantly that image would be black.

The resulting image that would be made would be something like this

Now why is this image all black. Because all the pixels in the image had a value of 0.

Digital Image Representation



Neighbors of a Pixel

A pixel *p* at coordinates has four *horizontal* and *vertical* neighbors whose coordinates are given by

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

This set of pixels, called the 4-*neighbors* of *p*, is denoted by Each pixel is a unit distance from , and some of the neighbor locations of *p* lie outside the digital image if is on the border of the image.

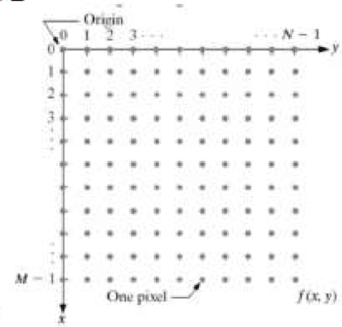
The four *diagonal* neighbors of *p* have coordinates

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

and are denoted by $N_D(p)$. These points, together with the 4-neighbors, are called the 8-*neighbors* of p, denoted by $N_8(p)$. As before, some of the neighbor locations in $N_D(p)$ and $N_8(p)$ fall outside the image if (x, y) is on the border of the image.

Some Basic Relationships Between Pixels

Neighbors of a Pixel

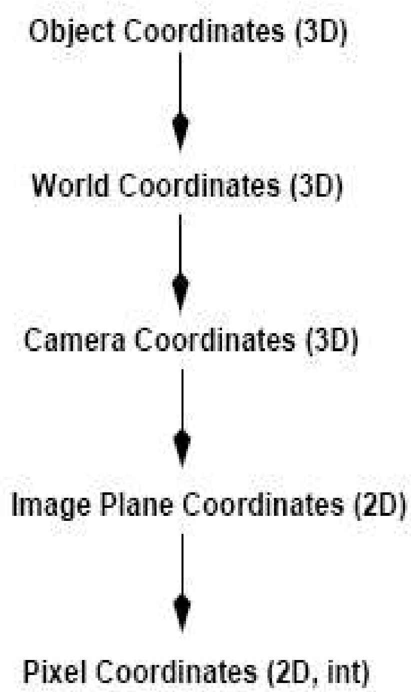


$$N_4 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$ $(x + 1, y), (x - 1, y), (x, y + 1), (x, y - 1).$

$$N_D = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$ $(x + 1, y + 1), (x + 1, y - 1), (x - 1, y + 1), (x - 1, y - 1)$

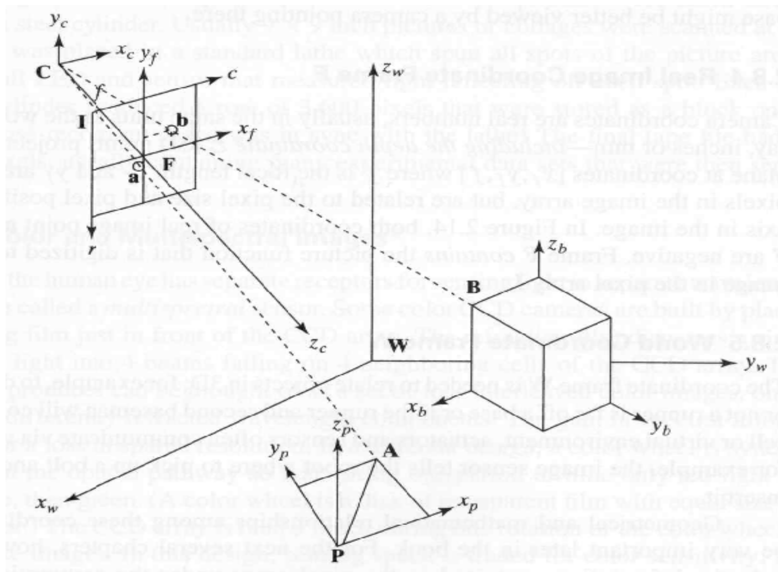
$$N_8 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

COORDINATE CONVENTIONS



(1) Object Coordinate Frame

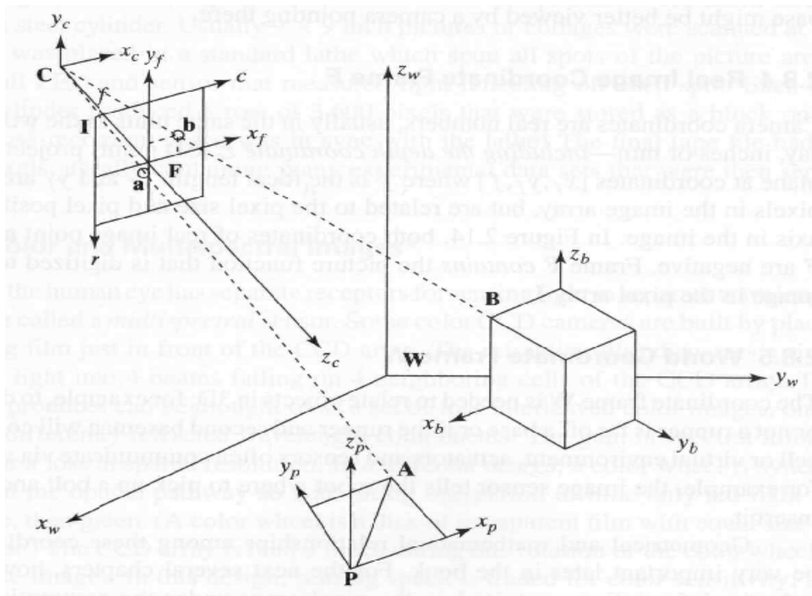
- 3D coordinate system: (x_b, y_b, z_b)
- Useful for modeling objects (i.e., check if a particular hole is in proper position relative to other holes)
- Object coordinates do not change regardless how the object is placed in the scene.

- 

Our notation: $(X_o, Y_o, Z_o)^T$

(2) World Coordinate Frame

- 3D coordinate system: (x_w, y_w, z_w)
- Useful for interrelating objects in 3D



***Our notation:* (X_w, Y_w, Z_w)^T**

(3) Camera Coordinate Frame

- 3D coordinate system: (x_c, y_c, z_c)
- Useful for representing objects with respect to the location of the camera.

Our notation: (X_c, Y_c, Z_c)^T

(4) Image Plane Coordinate Frame (i.e., CCD plane)

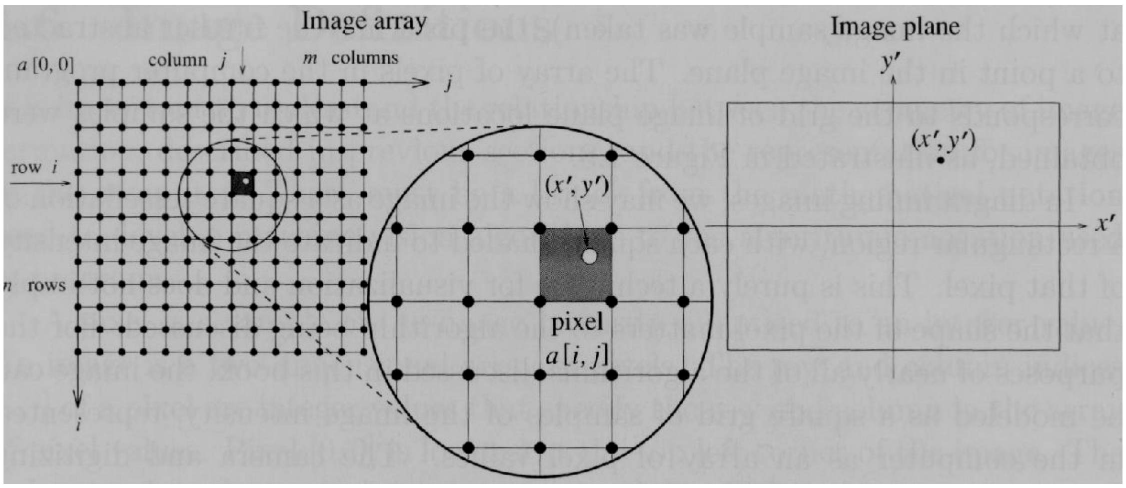
- 2D coordinate system: (x_f, y_f)
- Describes the coordinates of 3D points projected on the image plane.

  Our notation: (x, y)^T

(5) Pixel Coordinate Frame

- 2D coordinate system: (c, r)

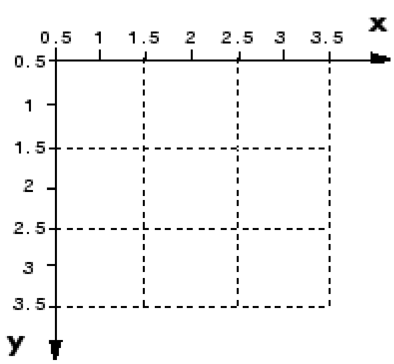
- Each pixel in this frame has integer pixel coordinates.



- **In general, the world and camera coordinate systems are not aligned.**

Spatial Coordinates

Spatial coordinates enable you to specify a location in an image with greater granularity than pixel coordinates. Such as, in the pixel coordinate system, a pixel is treated as a discrete unit, uniquely identified by an integer row and column pair, such as (3,4). In the spatial coordinate system, locations in an image are represented in terms of partial pixels, such as (3.3, 4.7).

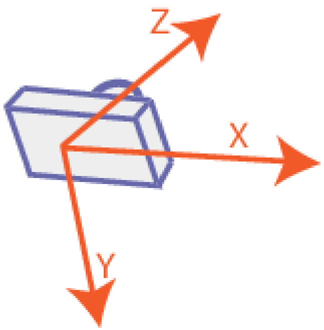


3-D Coordinate Systems

When you reconstruct a 3-D scene, you can define the resulting 3-D points in one of two coordinate systems. In a camera-based coordinate system, the points are defined relative to the center of the camera. In a calibration pattern-based coordinate system, the points are defined relative to a point in the scene.

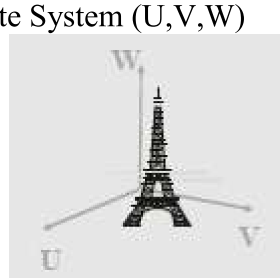
Camera-Based Coordinate System

Points represented in a camera-based coordinate system are described with the origin located at the optical center of the camera.

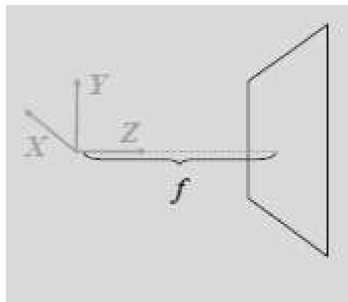


IMAGING GEOMETRY

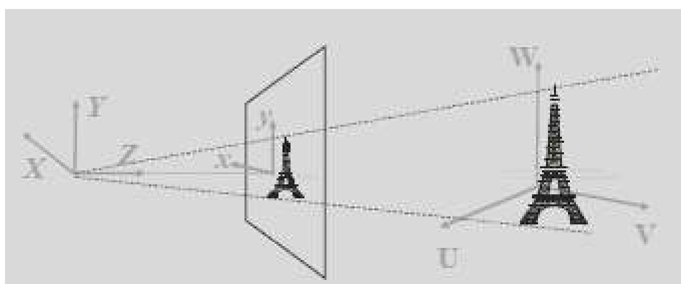
Object of Interest in World Coordinate System (U,V,W)



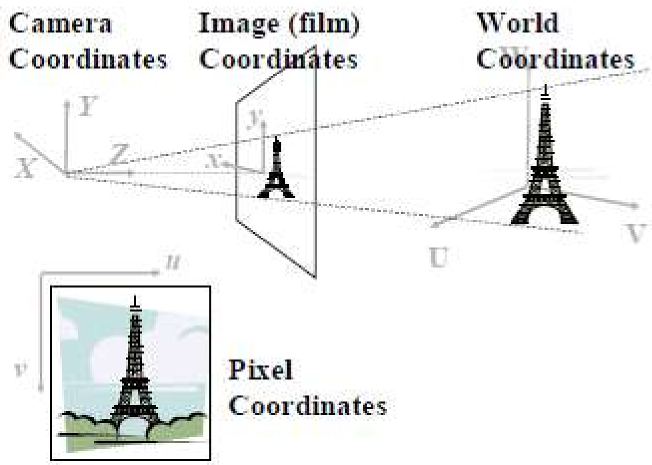
Camera Coordinate System (X,Y,Z).
- Z is optic axis
- Image plane located f units out along optic axis
- f is called focal length



Forward Projection onto image plane.
3D (X,Y,Z) projected to 2D *(x,y)*



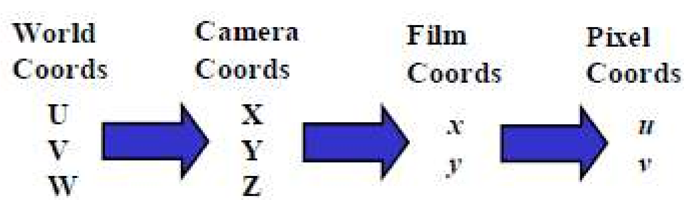
Our image gets digitized into pixel coordinates *(u,v)*



BASIC PERSPECTIVE PROJECTION

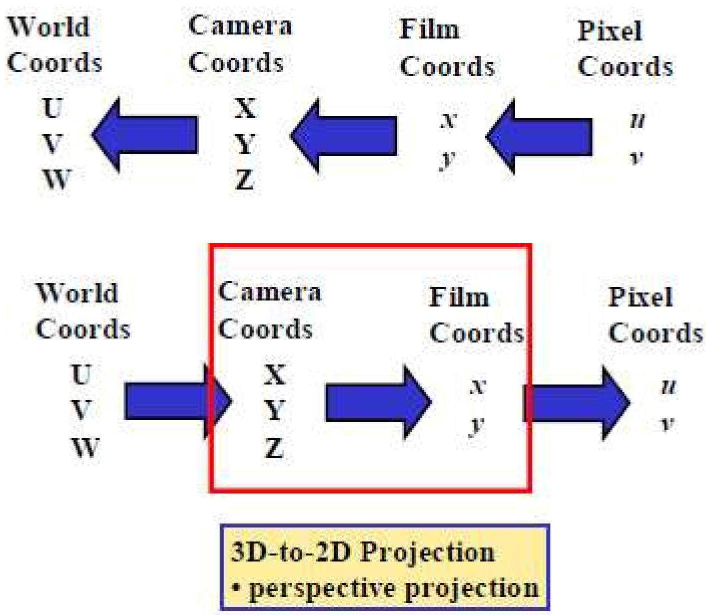
Forward Projection

We want a mathematical model to describe how 3D World points get projected into 2D Pixel coordinates.

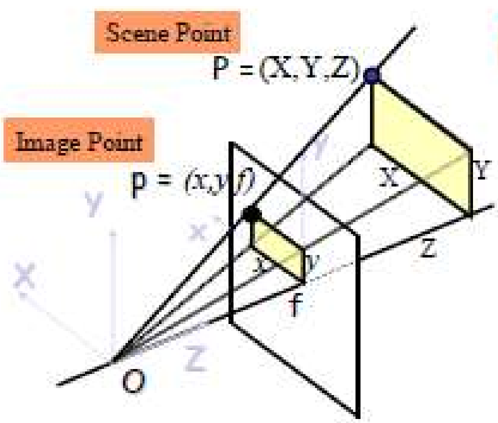


Backward Projection

Note, much of vision concerns trying to derive backward projection equations to recover 3D scene structure from images (via stereo or motion)


World
Coords
U
V
W
Camera
Coords
X
Y
Z
Film
Coords
x
y
Pixel
Coords
u
v
World
Coords
U
V
W
Camera
Coords
X
Y
Z
Film
Coords
x
y
Pixel
Coords
u
v
3D-to-2D Projection
• perspective projection

Basic Perspective Projection



$$x = f\frac{X}{Z}$$

$$y = f\frac{Y}{Z}$$

Homogeneous Coordinates

Represent a 2D point (x,y) by a 3D point (x',y',z') by adding a "fictitious" third coordinate.

By convention, we specify that given (x',y',z') we can recover the 2D point (x,y) as

$$x = \frac{x'}{z'} \quad y = \frac{y'}{z'}$$

Note: (x,y) = (x,y,1) = (2x, 2y, 2) = (k x, ky, k) for any nonzero k (can be negative as well as positive)

Perspective Matrix Equation

$$x = f\frac{X}{Z} \qquad y = f\frac{Y}{Z} \quad \Leftrightarrow \quad \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

SAMPLING AND QUANTIZATION

In order to become suitable for digital processing, an image function f(x,y) must be digitized both spatially and in amplitude.

In order to create an image which is digital, we need to covert continuous data into digital form. There are two steps in which it is done:

Sampling

- **Digitizing the coordinate values is called *sampling.***

Quantization

- **Digitizing the amplitude values is called *quantization***

Sampling : related to coordinates values

Quantization : related to intensity values

The one-dimensional function shown in Fig. 2.16(b) is a plot of amplitude (gray level) values of the continuous image along the line segment *AB* in Fig. 2.16(a).

The random variations are due to image noise. To sample this function, we take equally spaced samples along line *AB,* as shown in Fig. 2.16(c).

The location of each sample is given by a vertical tick mark in the bottom part of the figure the samples are shown as small white squares superimposed on the function.

The set of these discrete locations gives the sampled function.

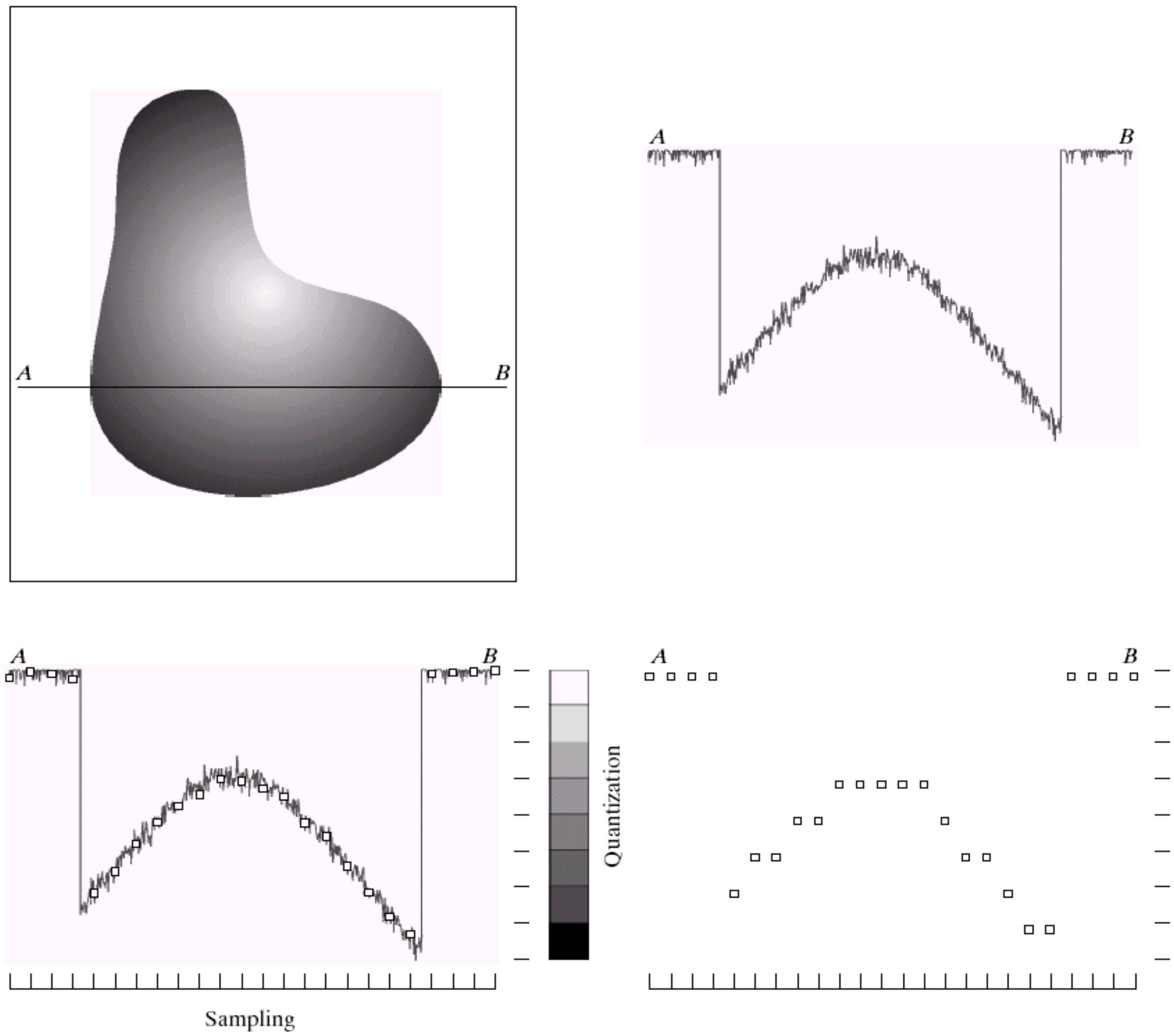
However, the values of the samples still span (vertically) a continuous range of gray-level values.

In order to form a digital function, the gray-level values also must be converted *(quantized)* into discrete quantities.

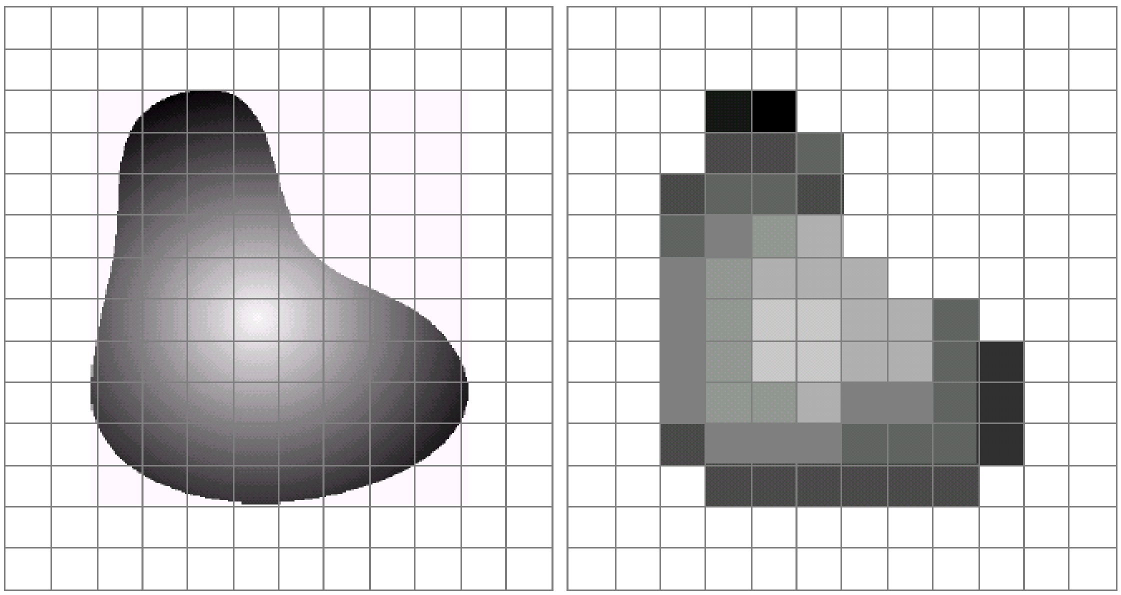
The right side of Fig. 2.16(c) shows the gray-level scale divided into eight discrete levels, ranging from black to white.

The vertical tick marks indicate the specific value assigned to each of the eight gray levels. The continuous gray levels are quantized simply by assigning one of the eight discrete gray levels to each sample. The assignment is made depending on the vertical proximity of a sample to a vertical tick mark. The digital samples resulting from both sampling and quantization are shown in Fig. 2.16(d).

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

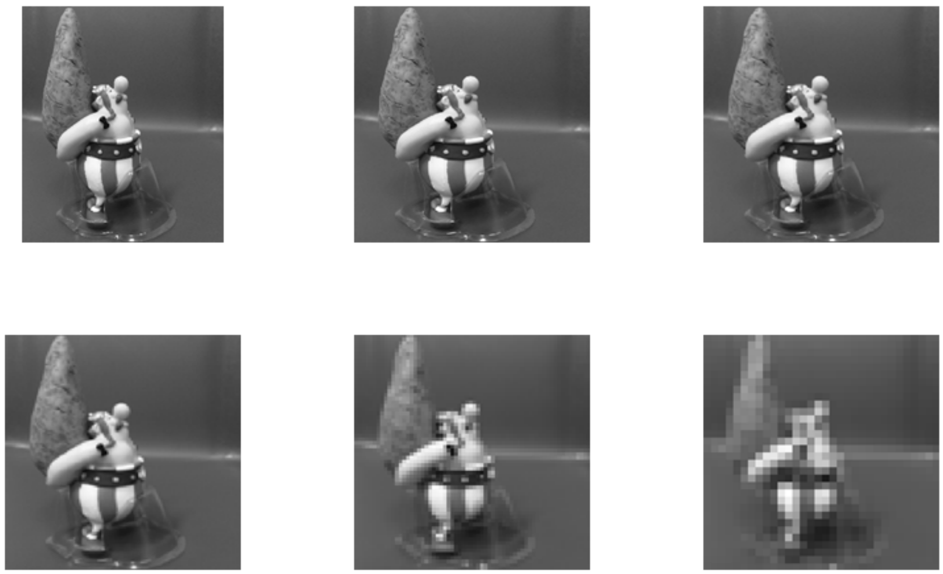


Generating a digital image. (a) Continuous image. (b) A scaling line from A to B in the continuous image, used to illustrate the concepts of sampling and quantization. (c) sampling and quantization. (d) Digital scan line.



a b

FIGURE 2.17 (a) Continuos image projected onto a sensor array. (b) Result of image sampling and quantization.



Spatial Domain Filtering

Filtering is a technique for modifying or enhancing an image. Spatial domain operation or filtering (processed value for the current pixel depends on both itself and surrounding pixels).

Hence Filtering is a neighborhood operation, in which the value of any given pixel in the output image is determined by applying some algorithm to the values of the pixels in the neighborhood of the corresponding input pixel. A pixel's neighborhood is some set of pixels, defined by their locations relative to that pixel.

- **Spatial domain**

image plane itself, directly process the intensity values of the image plane

- **Transform domain**

process the transform coefficients, not directly process the intensity values of the image plane

Spatial Domain Process

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

$f(x,y)$: input image

$g(x,y)$: output image

T : an operator on f defined over
a neighborhood of point (x,y)

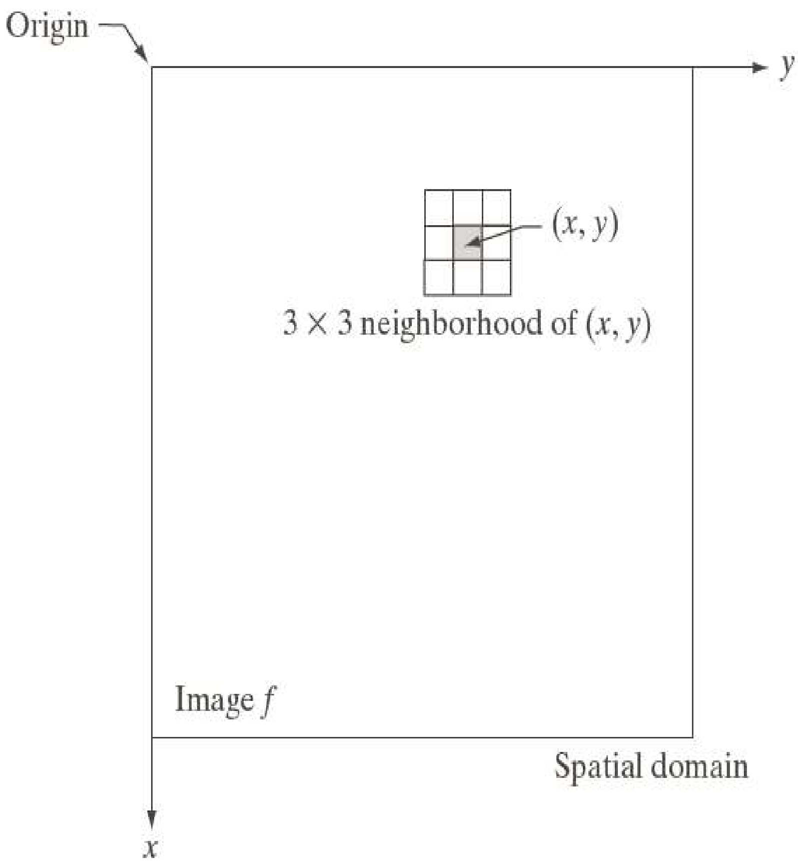


FIGURE 3.1 A 3 × 3 neighborhood about a point (x, y) in an image in the spatial domain. The neighborhood is moved from pixel to pixel in the image to generate an output image.

Some neighborhood operations work with the values of the image pixels in the neighborhood *and* the corresponding values of a subimage that has the same dimensions as the neighborhood. The subimage is called a *filter, mask, kernel, template,* or *window,* with the first three terms being the most prevalent terminology. The values in a filter subimage are referred to as *coefficients,* rather than pixels.

INTENSITY TRANSFORMATIONS

Image enhancement

Enhancing an image provides better contrast and a more detailed image as compare to non enhanced image. Image enhancement has very applications. It is used to enhance medical images, images captured in remote sensing, images from satellite e.t.c

The transformation function has been given below

$$s = T(r)$$

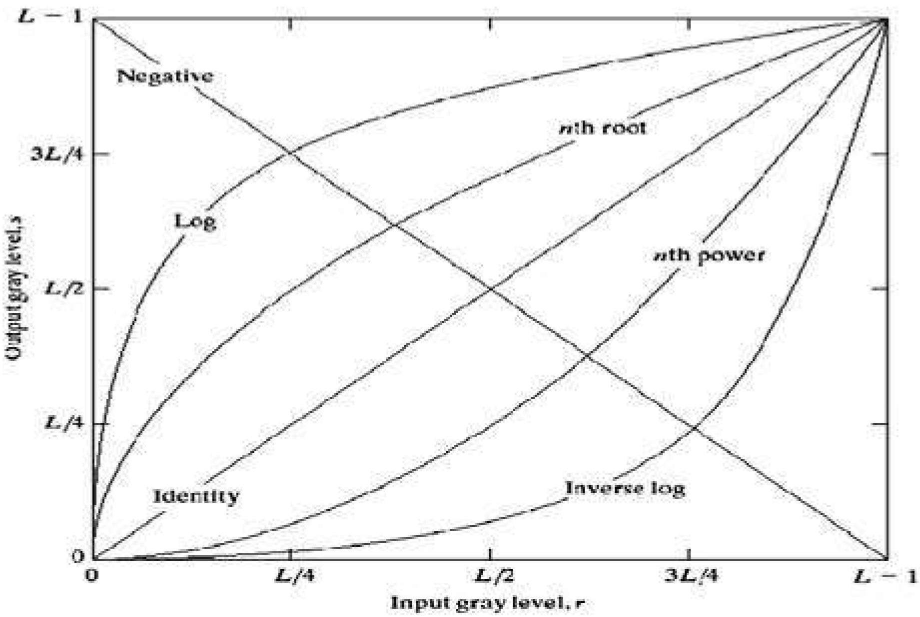
where r is the pixels of the input image and s is the pixels of the output image. T is a transformation function that maps each value of r to each value of s. Image enhancement can be done through gray level transformations which are discussed below.

Gray level transformation

There are three basic gray level transformation.

- **Linear**
- **Logarithmic**
- **Power – law**

The overall graph of these transitions has been shown below.



Linear transformation

First we will look at the linear transformation. Linear transformation includes simple identity and negative transformation.

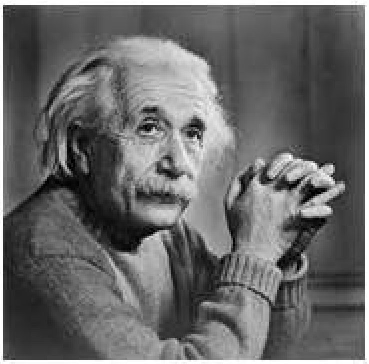
Identity transition is shown by a straight line. In this transition, each value of the input image is directly mapped to each other value of output image. That results in the same input image and output image. And hence is called identity transformation.

Negative transformation

The second linear transformation is negative transformation, which is invert of identity transformation. In negative transformation, each value of the input image is subtracted from the L-1 and mapped onto the output image.

The result is somewhat like this.

Input Image



Output Image

In this case the following transition has been done.

$$s = (L - 1) - r$$

since the input image of Einstein is an 8 bpp image, so the number of levels in this image are 256. Putting 256 in the equation, we get this

$$s = 255 - r$$

So each value is subtracted by 255 and the result image has been shown above. So what happens is that, the lighter pixels become dark and the darker picture becomes light. And it results in image negative.

It has been shown in the graph below.

Logarithmic transformations

Logarithmic transformation further contains two type of transformation. Log transformation and inverse log transformation.

Log transformation

The log transformations can be defined by this formula

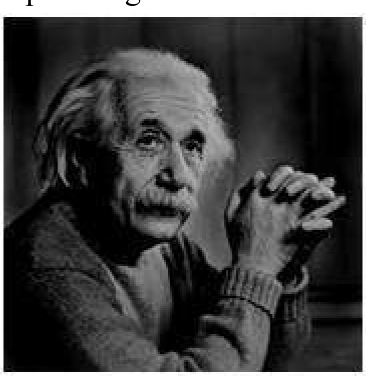
$$s = c \log(r + 1).$$

Where s and r are the pixel values of the output and the input image and c is a constant. The value 1 is added to each of the pixel value of the input image because if there is a pixel intensity of 0 in the image, then log (0) is equal to infinity. So 1 is added, to make the minimum value at least 1.

During log transformation, the dark pixels in an image are expanded as compare to the higher pixel values. The higher pixel values are kind of compressed in log transformation. This result in following image enhancement.

The value of c in the log transform adjust the kind of enhancement you are looking for.

Input Image



Log Tranform Image

The inverse log transform is opposite to log transform.

Power – Law transformations

There are further two transformation is power law transformations, that include nth power and nth root transformation. These transformations can be given by the expression:

$$s=cr^{\gamma}$$

This symbol γ is called gamma, due to which this transformation is also known as gamma transformation.

Variation in the value of γ varies the enhancement of the images. Different display devices / monitors have their own gamma correction, that's why they display their image at different intensity.

This type of transformation is used for enhancing images for different type of display devices. The gamma of different display devices is different. For example Gamma of CRT lies in between of 1.8 to 2.5, that means the image displayed on CRT is dark.

Correcting gamma.

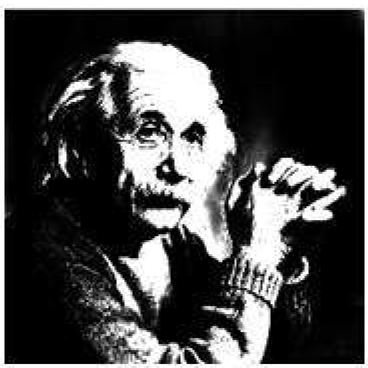
s=cr^γ

s=cr^(1/2.5)

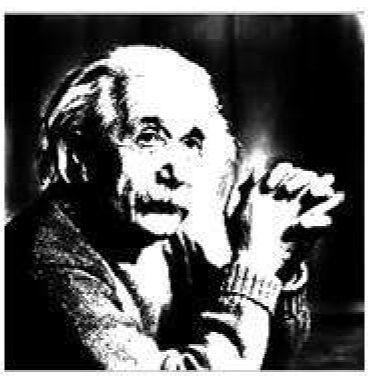
The same image but with different gamma values has been shown here.

For example

Gamma = 10



Gamma = 8



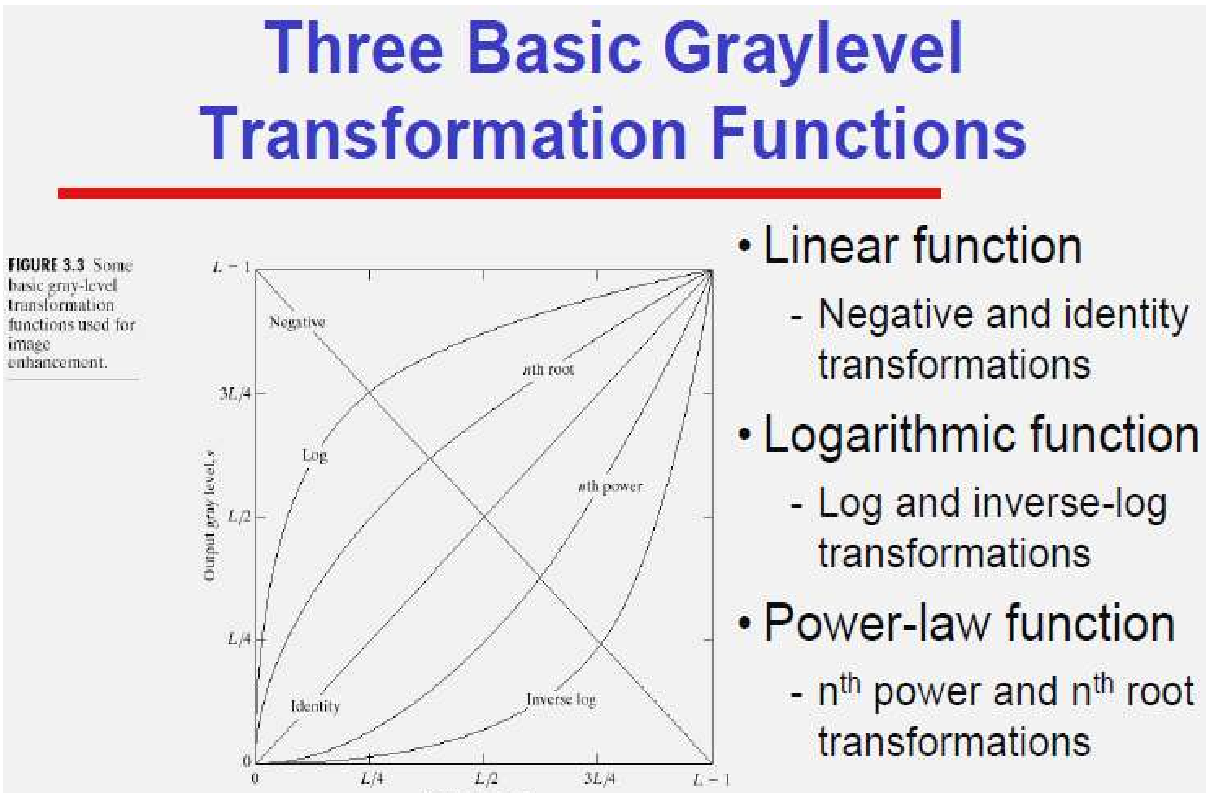

Three Basic Graylevel Transformation Functions
FIGURE 3.3 Some basic gray-level transformation functions used for image enhancement.
L − 1
3L/4
L/2
L/4
0
Output gray level, s
Negative
nth root
Log
nth power
Identity
Inverse log
0
L/4
L/2
3L/4
L − 1
• Linear function
- Negative and identity transformations
• Logarithmic function
- Log and inverse-log transformations
• Power-law function
- n th power and n th root transformations

Image Negatives

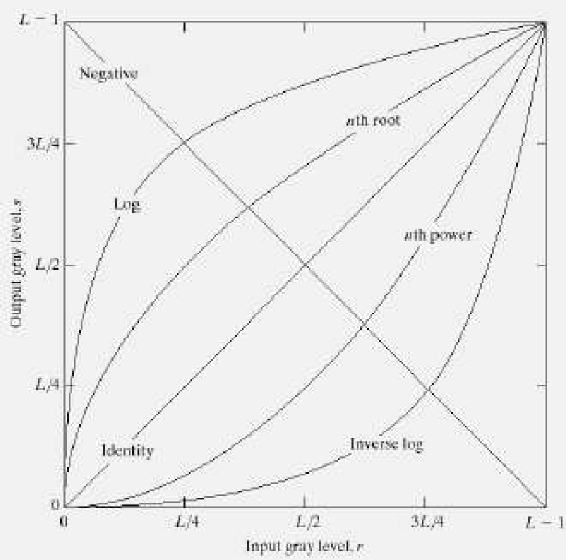
- Negative transformation : $s = (L-1) - r$
- Reverses the intensity levels of an image.
- Suitable for enhancing white or gray detail embedded in dark regions of an image, especially when black area is large.

a b

FIGURE 3.4 (a) Original digital mammogram. (b) Negative image obtained using the negative transformation in Eq. (3.2-1). (Courtesy of G.E. Medical Systems.)

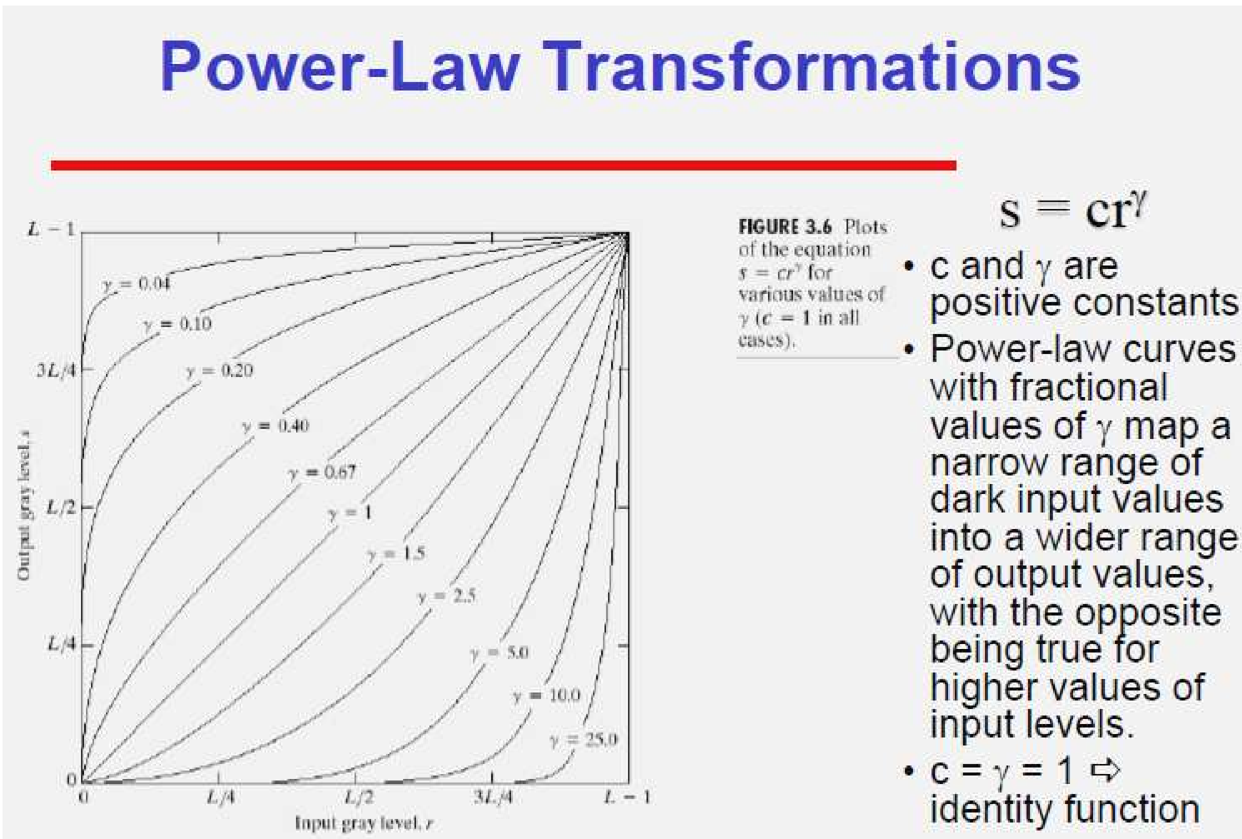
Log Transformations

FIGURE 3.3 Some basic gray-level transformation functions used for image enhancement.



s = c log (1+r)

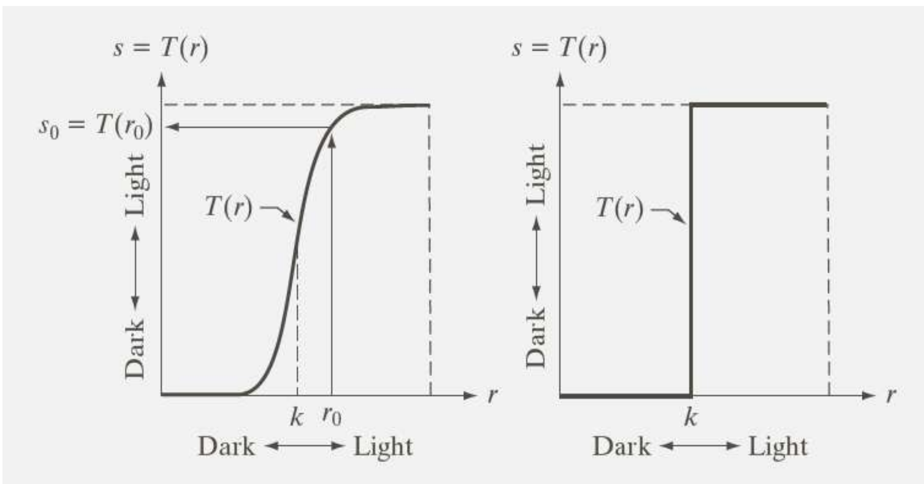
- c is constant and $r \geq 0$
- Log curve maps a narrow range of low graylevels in input image into a wider range of output levels.
- Expands range of dark image pixels while shrinking bright range.
- Inverse log expands range of bright image pixels while shrinking dark range.



Contrast Stretching

- **Contrast Stretching**
  — Expands the range of intensity levels in an image so that it spans the full intensity range of the recording medium or display device.

- **Intensity-level Slicing**
  — Highlighting a specific range of intensities in an image often is of interest.



a b

FIGURE 3.2 Intensity transformation functions. (a) Contrast-stretching function. (b) Thresholding function.

For example, if *T(r)* has the form shown in Fig. 3.2(a), the effect of this transformation would be to produce an image of higher contrast than the original by darkening the levels below *m* and Brightening the levels above *m* in the original image.

In this technique, known as *contrast stretching,* the values of *r* below k are compressed by the transformation function into a narrow range of *s,* toward black. The opposite effect takes place for values of *r* above k.

In the limiting case shown in Fig. 3.2(b), *T(r)* produces a two-level (binary) image. A mapping of this form is called a *thresholding* function. Some fairly simple, yet powerful, processing approaches can be formulated with gray-level transformations. Because enhancement at any point in an image depends only on the gray level at that point.

Contrast stretching

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

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

Figure 3.10(a) shows a typical transformation used for contrast stretching.

The locations of points *(r1,s1)* and *(r2,s2)* control the shape of the transformation in this category often are referred to as *point* function.

If $r1 = s1$ and $r2 = s2$ the transformation is a linear function that produces no changes in gray levels. If $r1 = r2$, $s1 = 0$ and $s2 = L - 1$, the transformation becomes a *thresholding function* that creates a binary image, as illustrated in Fig. 3.2(b).

Intermediate values of (r1,s1) and (r2,s2) produce various degrees of spread in the gray levels of the output image, thus affecting its contrast.

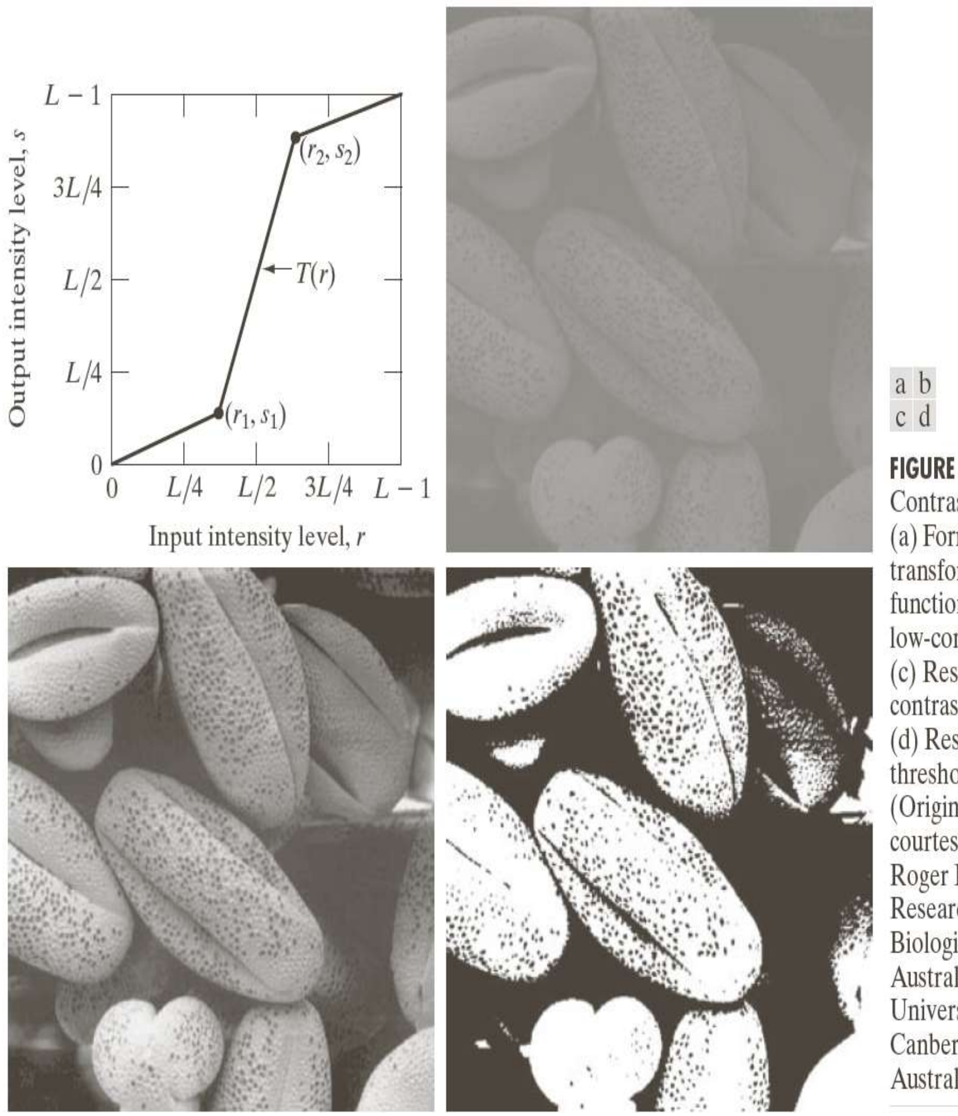


FIGURE 3.10 Contrast stretching. (a) Form of transformation function. (b) A low-contrast image. (c) Result of contrast stretching. (d) Result of thresholding. (Original image courtesy of Dr. Roger Heady, Research School of Biological Sciences, Australian National University, Canberra, Australia.)

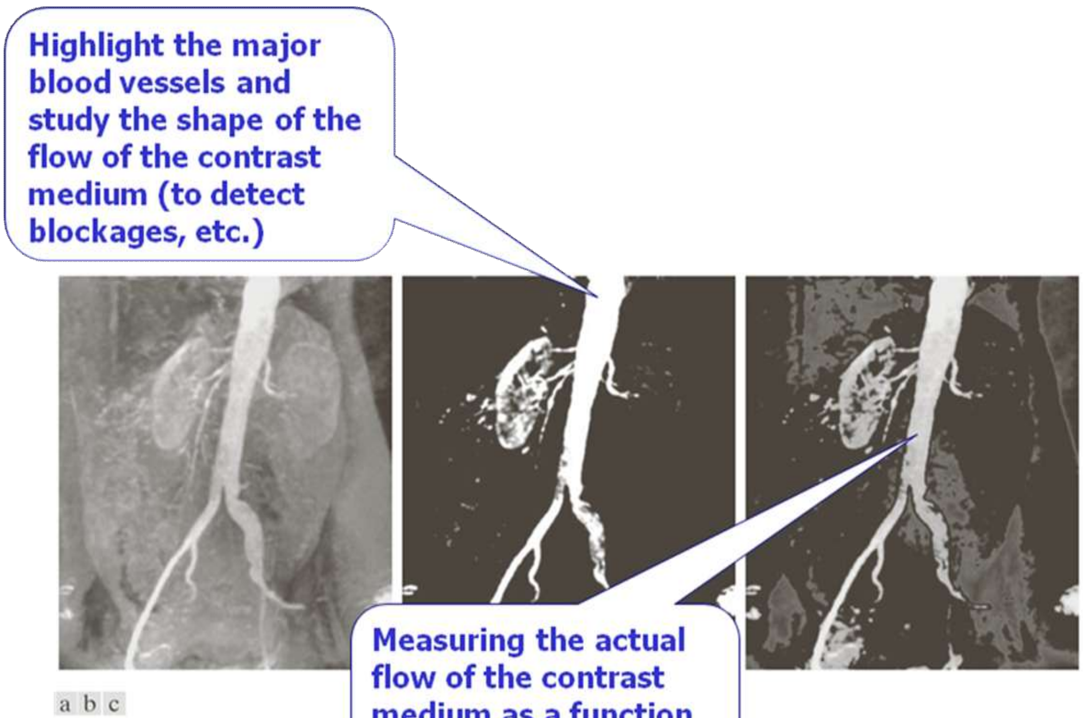


FIGURE 3.12 (a) Aortic angiog... mation of the type illustrated in Fig. 3.11(a), with the range of inte... end of the gray scale. (c) Result of using the transformation in Fig... ack, so that grays in the area of the blood vessels and kidneys were preserved. (Original image courtesy of Dr. Thomas R. Gest, University of Michigan Medical School.)

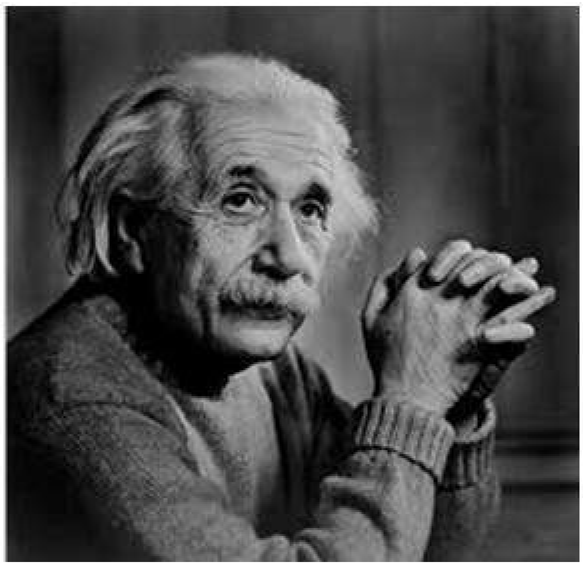
There are two methods of enhancing contrast. The first one is called Histogram stretching that increase contrast. The second one is called Histogram equalization that enhance contrast and it has been discussed in our tutorial of histogram equalization.

Before we will discuss the histogram stretching to increase contrast, we will briefly define contrast.

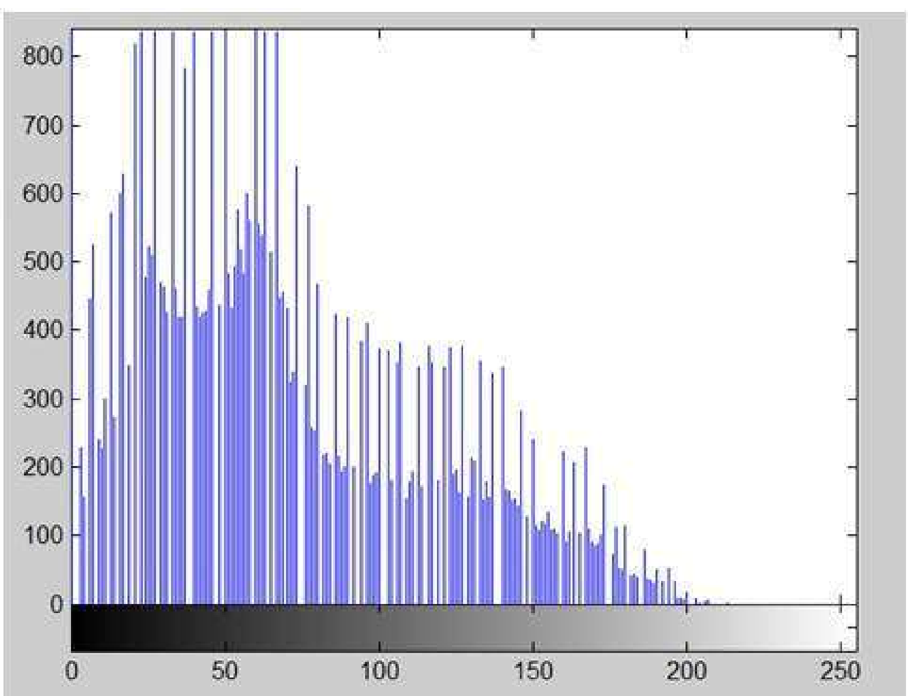
Contrast

Contrast is the difference between maximum and minimum pixel intensity.

Consider this image.



The histogram of this image is shown below.



Now we calculate contrast from this image.

Contrast = 225.

Now we will increase the contrast of the image.

HISTOGRAM PROCESSING

Histogram

Histogram is nothing but a graph that shows frequency of occurrence of data. Histograms has many use in image processing, out of which we are going to discuss one user here which is called histogram sliding.

Brightness

Brightness is a relative term. Brightness can be defined as intensity of light emit by a particular light source.

Contrast

Contrast can be defined as the difference between maximum and minimum pixel intensity in an image.

Using Histogram Statistics for Image Enhancement

A histogram is a graph. A graph that shows frequency of anything. Usually histogram have bars that represent frequency of occurring of data in the whole data set.

A Histogram has two axis the x axis and the y axis.

The x axis contains event whose frequency you have to count.

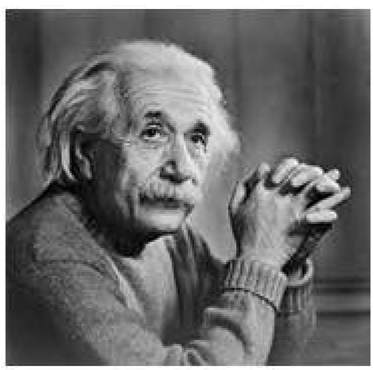
The y axis contains frequency.

HISTOGRAM EQUALIZATION

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

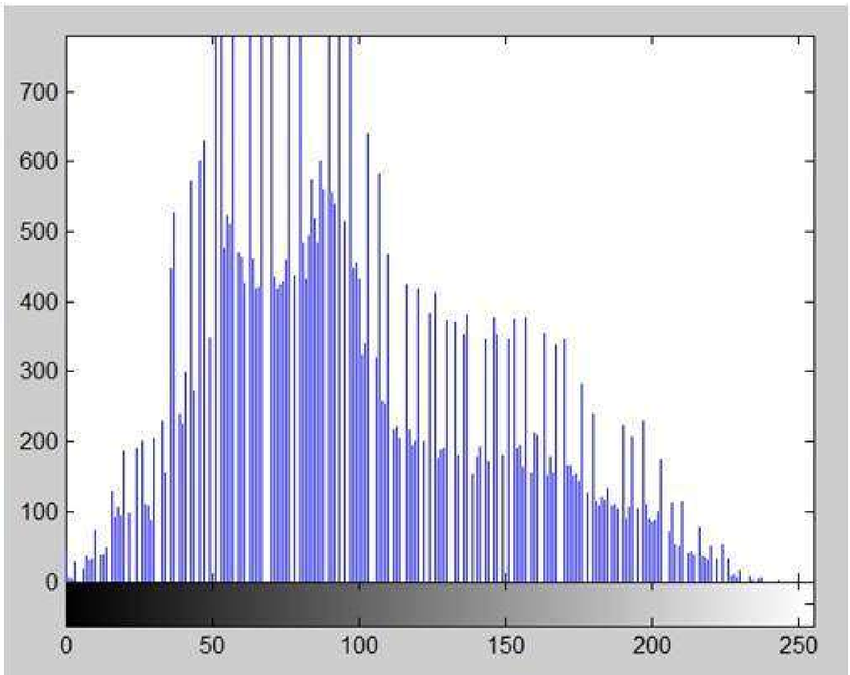
Lets start histogram equalization by taking this image below as a simple image.

Image



Histogram of this image

The histogram of this image has been shown below.



Now we will perform histogram equalization to it.

Histogram $h(r_k) = n_k$

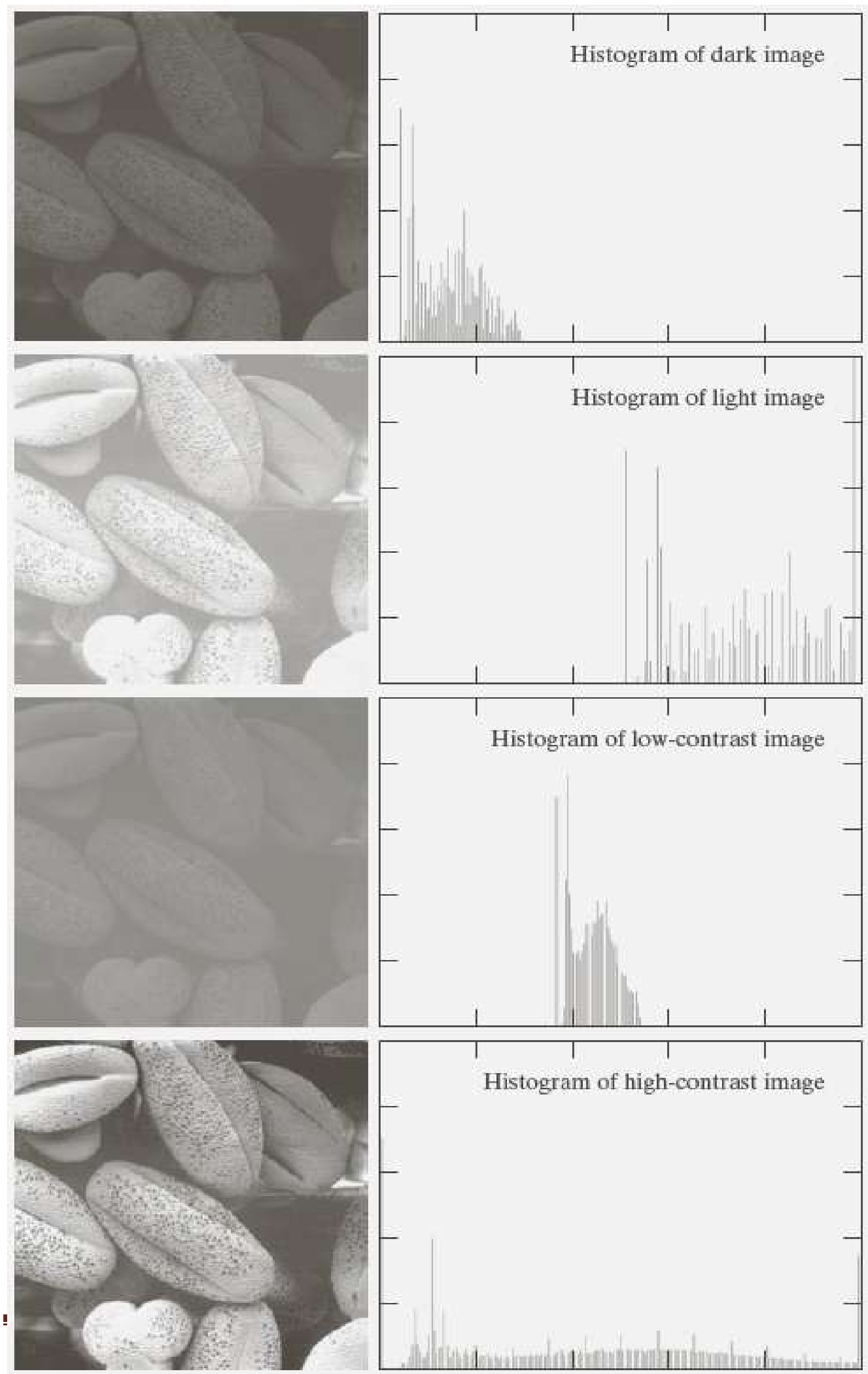
r_k is the k^{th} intensity value

n_k is the number of pixels in the image with intensity r_k

Normalized histogram $p(r_k) = \frac{n_k}{MN}$

n_k: the number of pixels in the image of size $\mathrm{M} \times \mathrm{N}$ with intensity r_k

The horizontal axis of each histogram plot corresponds to gray level values, r_k. The vertical axis corresponds to values of $h(r_k) = n_k$ or $p(r_k) = n_k/n$ if the values are normalized. Thus, as indicated previously, these histogram plots are simply plots of $h(r_k) = n_k$ versus r_k or $p(r_k) = n_k/n$ versus r_k.


Histogram of dark image
Histogram of light image
Histogram of low-contrast image
Histogram of high-contrast image

Histogram Equalization Image

- We note 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 toward 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 possible 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.

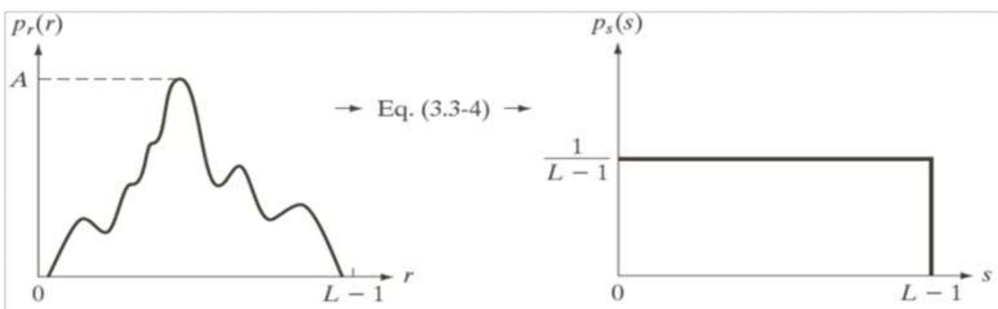
Histogram Equalization

- **Histogram equalization is a technique for adjusting image intensities to enhance contrast.**

Histogram Equalization

The intensity levels in an image may be viewed as random variables in the interval [0, L-1].

Let $p_r(r)$ and $p_s(s)$ denote the probability density function (PDF) of random variables r and s.



a b

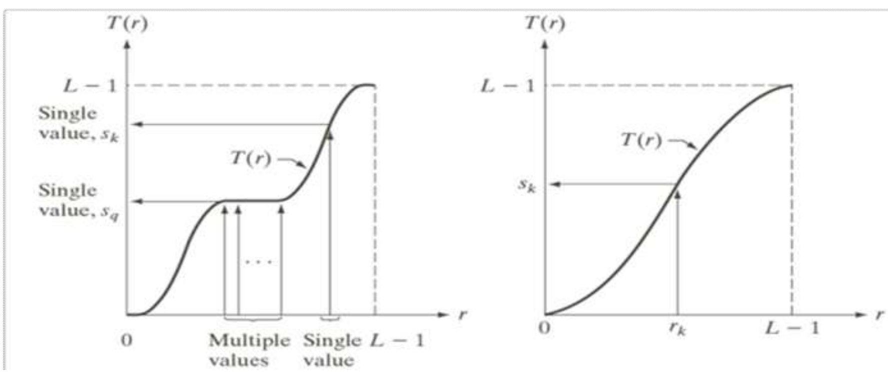
FIGURE 3.18 (a) An arbitrary PDF. (b) Result of applying the transformation in Eq. (3.3-4) to all intensity levels, r. The resulting intensities, s, have a uniform PDF, independently of the form of the PDF of the r's.

Histogram Equalization

$$s = T(r) \qquad 0 \le r \le L-1$$

a. T(r) is a strictly monotonically increasing function in the interval $0 \le r \le L-1$;

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



a b

FIGURE 3.17 (a) Monotonically increasing function, showing how multiple values can map to a single value. (b) Strictly monotonically increasing function. This is a one-to-one mapping, both ways.

Spatial Filtering (cont'd)

- Spatial filtering is defined by:
  - **(1)** A neighborhood
  - **(2)** An operation that is performed on the pixels inside the neighborhood

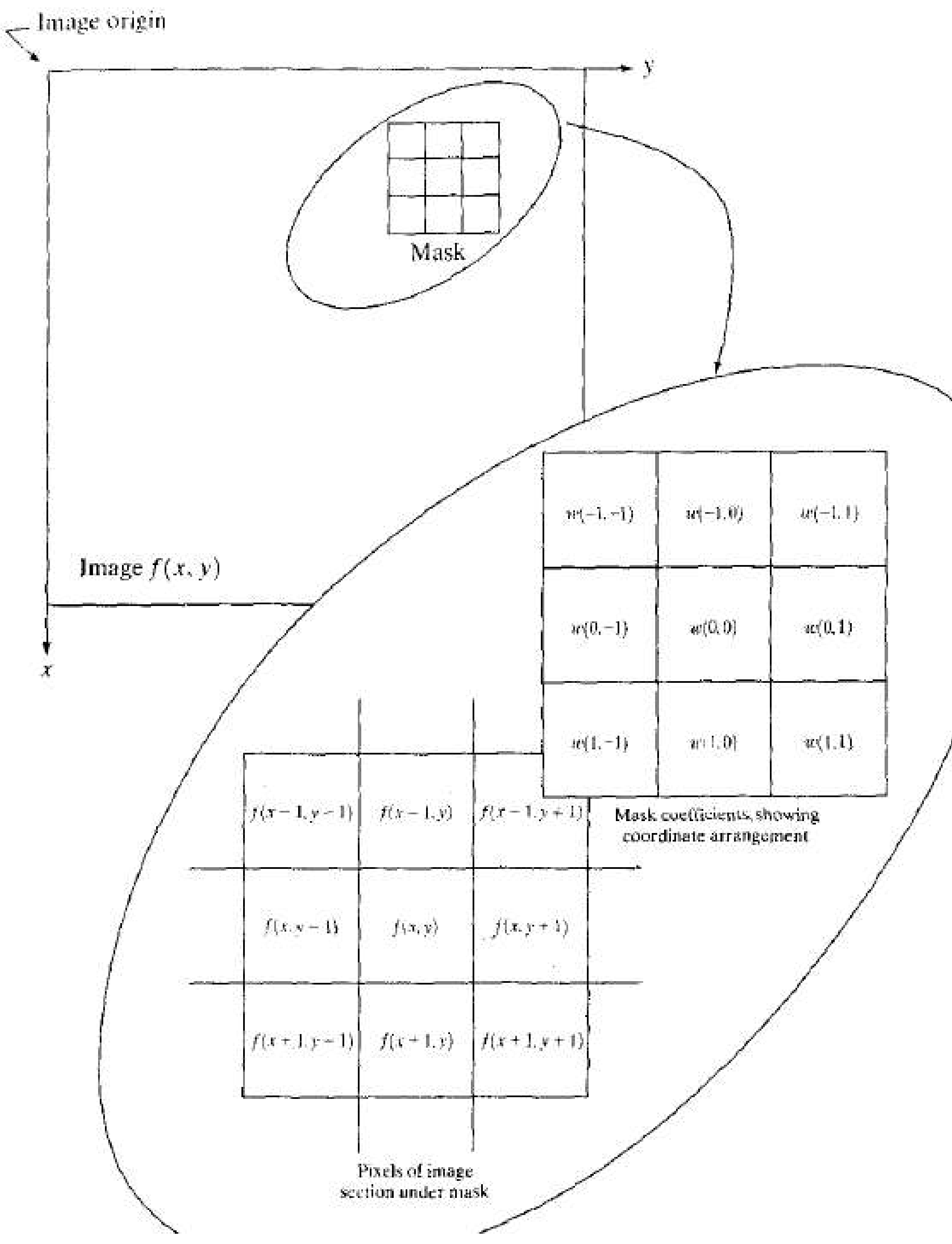


Image origin
y
Mask
Image $f(x, y)$
x
$w(-1,-1)$
$w(-1,0)$
$w(-1,1)$
$w(0,-1)$
$w(0,0)$
$w(0,1)$
$w(1,-1)$
$w(1,0)$
$w(1,1)$
Mask coefficients, showing coordinate arrangement
$f(x-1,y-1)$
$f(x-1,y)$
$f(x-1,y+1)$
$f(x,y-1)$
$f(x,y)$
$f(x,y+1)$
$f(x+1,y-1)$
$f(x+1,y)$
$f(x+1,y+1)$
Pixels of image section under mask

When interest lies on the response, R, of an $m \times n$ mask at any point (x, y), and not on the mechanics of implementing mask convolution, it is common practice to simplify the notation by using the following expression:

$$R = w_1 z_1 + w_2 z_2 + \ldots + w_{mn} z_{mn} = \sum_{i=1}^{mn} w_i z_i \tag{3.5-2}$$

where the w's are mask coefficients, the z's are the values of the image gray levels corresponding to those coefficients, and mn is the total number of coefficients in the mask. For the 3×3 general mask shown in Fig. 3.33 the response at any point (x, y) in the image is given by

$$R = w_1 z_1 + w_2 z_2 + \ldots w_9 z_9 = \sum_{i=1}^{9} w_i z_i. \tag{3.5-3}$$

FIGURE 3.33 Another representation of a general 3×3 spatial filter mask.

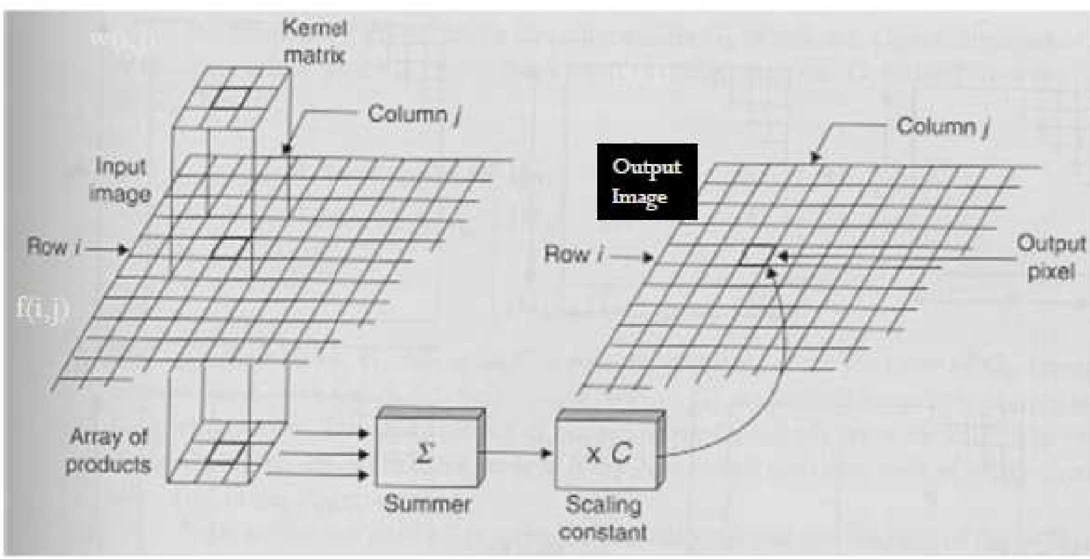
w_1	w_2	w_3
w_4	w_5	w_6
w_7	w_8	w_9

CORRELATION AND CONVOLUTION

Linear Spatial Filtering Methods

- Two main linear spatial filtering methods:
  - Correlation
  - Convolution

Correlation



$$g(x,y) = w(x,y) \bullet f(x,y) = \sum_{s=-K/2}^{K/2} \sum_{t=-K/2}^{K/2} w(s,t) f(x+s, y+t)$$

Convolution

- Similar to correlation except that the mask is first **flipped** both horizontally and vertically.

$$g(x,y) = w(x,y) * f(x,y) = \sum_{s=-K/2}^{K/2} \sum_{t=-K/2}^{K/2} w(s,t) f(x-s, y-t)$$

Convolution can achieve something, that the previous two methods of manipulating images can't achieve. Those include the blurring, sharpening, edge detection, noise reduction e.t.c.

It can be represented as.

It can be mathematically represented as two ways

g(x,y) = h(x,y) * f(x,y)

It can be explained as the "mask convolved with an image".

Or

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

It can be explained as "image convolved with mask".

There are two ways to represent this because the convolution operator(*) is commutative. The h(x,y) is the mask or filter.

What is mask?

Mask is also a signal. It can be represented by a two dimensional matrix. The mask is usually of the order of 1x1, 3x3, 5x5, 7x7 . A mask should always be in odd number, because other wise you cannot find the mid of the mask. Why do we need to find the mid of the mask. The answer lies below, in topic of, how to perform convolution?

How to perform convolution?

In order to perform convolution on an image, following steps should be taken.

- Flip the mask (horizontally and vertically) only once
- Slide the mask onto the image.
- Multiply the corresponding elements and then add them
- Repeat this procedure until all values of the image has been calculated.

Example of convolution

Let's perform some convolution. Step 1 is to flip the mask.

Mask

Let's take our mask to be this.

1	2	3
4	5	6
7	8	9

Flipping the mask horizontally

3	2	1
6	5	4
9	8	7

Flipping the mask vertically

9	8	7
6	5	4
3	2	1

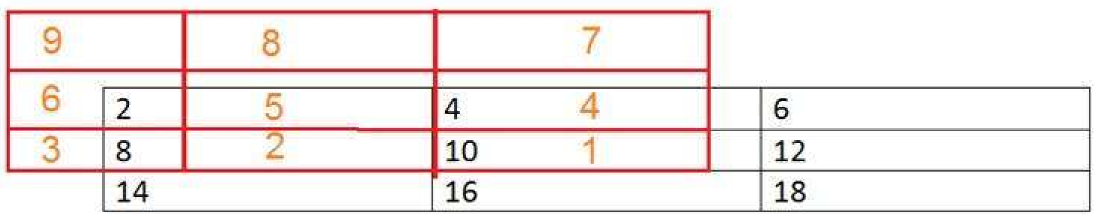
Image

Let's consider an image to be like this

2	4	6
8	10	12
14	16	18

Convolution

Convolving mask over image. It is done in this way. Place the center of the mask at each element of an image. Multiply the corresponding elements and then add them , and paste the result onto the element of the image on which you place the center of mask.



The box in red color is the mask, and the values in the orange are the values of the mask. The black color box and values belong to the image. Now for the first pixel of the image, the value will be calculated as

First pixel = (5*2) + (4*4) + (2*8) + (1*10)

= 10 + 16 + 16 + 10

= 52

Place 52 in the original image at the first index and repeat this procedure for each pixel of the image.

Spatial Convolution

The convolution of a filter $w(x, y)$ of size $m \times n$ with an image $f(x, y)$, denoted as $w(x, y) \quad f(x, y)$

$$w(x, y) \quad f(x, y) = \sum_{s=-a}^{a} \sum_{t=-b}^{b} w(s, t) f(x - s, y - t)$$

Spatial Correlation

The correlation of a filter $w(x, y)$ of size $m \times n$ with an image $f(x, y)$, denoted as $w(x, y) \quad f(x, y)$

$$w(x, y) \quad f(x, y) = \sum_{s=-a}^{a} \sum_{t=-b}^{b} w(s, t) f(x + s, y + t)$$

Spatial Filtering

A spatial filter consists of (a) **a neighborhood**, and (b) **a predefined operation**

Linear spatial filtering of an image of size MxN with a filter of size mxn is given by the expression

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

Linear vs Non-Linear Spatial Filtering Methods

- A filtering method is linear when the output is a weighted sum of the input pixels.

w1	w2	w3
w4	w5	w6
w7	w8	w9

$\boxed{z5' = R = w1z1 + w2z2 + \ldots + z9w9}$

- Methods that do not satisfy the above property are called non-linear.
  - e.g.,

$$z_5' = max(z_k, k = 1, 2, \ldots, 9)$$

What is a mask

A mask is a filter. Concept of masking is also known as spatial filtering. Masking is also known as filtering. In this concept we just deal with the filtering operation that is performed directly on the image.

A sample mask has been shown below

-1	0	1
-1	0	1
-1	0	1

What is filtering

The process of filtering is also known as convolving a mask with an image. As this process is same of convolution so filter masks are also known as convolution masks.

How it is done

The general process of filtering and applying masks is consists of moving the filter mask from point to point in an image. At each point (x,y) of the original image, the response of a filter is calculated by a pre defined relationship. All the filters values are pre defined and are a standard.

Types of filters

Generally there are two types of filters. One is called as linear filters or smoothing filters and others are called as frequency domain filters.

Why filters are used?

Filters are applied on image for multiple purposes. The two most common uses are as following:

- Filters are used for Blurring and noise reduction
- Filters are used for edge detection and sharpness

Blurring and noise reduction

Filters are most commonly 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 large object extraction.

Masks for blurring

The common masks for blurring are.

- **Box filter**
- **Weighted average filter**

In the process of blurring we reduce the edge content in an image and try to make the transitions between different pixel intensities as smooth as possible.

Noise reduction is also possible with the help of blurring.

Edge Detection and sharpness

Masks or filters can also be used for edge detection in an image and to increase sharpness of an image.

Smoothing filters

Smoothing filters are used for blurring and for noise reduction. Blurring is used in preprocessing steps, such as removal of small details from an image prior to (large) object extraction. and bridging of small gaps in lines or curves. Noise reduction can be accomplished by blurring with a linear filter and also by nonlinear filtering.

Smoothing Linear Filters

The output (response) of a smoothing. linear spatial filter is simply the average of the pixels contained in the neighborhood of the filter mask. These filters sometimes are called *averaging filters* also are referred to a *lowpass fillers.*

The idea behind smoothing filters is straightforward. By replacing the value of every pixel in an image by the average of the gray levels in the neighborhood defined by the filter mask. this process results in an image with reduced "sharp" transitions in gray levels.

However. edges (which almost always are desirable features of an image) also are characterized by sharp transitions in gray levels. so averaging filters have the undesirable side effect that they blur edges

Smoothing Spatial Filters

Smoothing filters are used for blurring and for noise reduction

Blurring is used in removal of small details and bridging of small gaps in lines or curves

Smoothing spatial filters include linear filters and nonlinear filters.

Two Smoothing Averaging Filter Masks

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

a b

FIGURE 3.32 Two 3×3 smoothing (averaging) filter masks. The constant multiplier in front of each mask is equal to 1 divided by the sum of the values of its coefficients, as is required to compute an average.

Box filter :
Figure 3.32 shows two 3 X 3 smoothing filters. Use of thefirst filter yields the standard average of the pixels under the mask. This can best be seen by substituting the coefficients of the mask into Eq. (3.5-3):

$$R = \frac{1}{9}\sum_{i=1}^{9} z_i,$$

which is the average of the gray levels of the pixels in the 3 x 3 neighborhood defined by the mask. Note that, instead of being 1/9, the coefficients of the filter are all 1's. The idea here is that it is computationally more efficient to have coefficients valued 1. At the end of the filtering process the entire image is divided by 9. An m X n mask would have a normalizing constant equal to 1/*mn.*

A spatial averaging filter in which all coefficients are equal is sometimes called a *box filter*.

Weighted average :
The second mask yields a so-called *weighted average,* terminology used to indicate that pixels are multiplied by different coefficients, thus giving more importance (weight) to some pixels at the expense of others.

<u>Non-linear Spatial Filters</u>

Order-statistic (Nonlinear) Filters

— Nonlinear

— Based on ordering (ranking) the pixels contained in the filter mask

— Replacing the value of the center pixel with the value determined by the ranking result

E.g., median filter, max filter, min filter

Order-Statistics Filters

- Nonlinear filters whose response is based on ordering (ranking) the pixels contained in the filter support.
- Replace value of the center pixel with value determined by ranking result.
- Order statistic filters applied to *nxn* neighborhoods:
  - median filter: $R = \text{median}\{z_k \mid k = 1,2,\ldots,n^2\}$
  - max filter: $R = \max\{z_k \mid k = 1,2,\ldots, n^2\}$
  - min filter: $R = \min\{z_k \mid k = 1,2,\ldots, n^2\}$

Median Filter

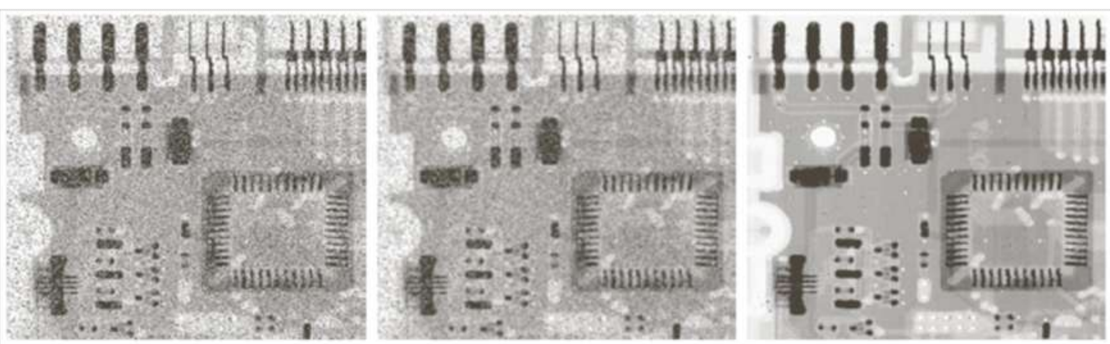
- Sort all neighborhood pixels in increasing order.
- Replace neighborhood center with the median.
- The window shape does not need to be a square.
- Special shapes can preserve line structures.
- Useful in eliminating intensity spikes: salt & pepper noise.

10	20	20
20	200	15
25	20	25

(10,15,20,20,20,20,25,25,200)
Median = 20
Replace 200 with 20

Min Filter	The intensity of the pixel at the center of the mask is replaced by the minimum intensity value of any pixel within the mask. This filter is used to find the dark points in an image.
Max Filter	The intensity of the pixel at the center of the mask is replaced by the maximum intensity value of any pixel within the mask. This filter is used to find the bright points in an image.

Example: Use of Median Filtering for Noise Reduction



a b c

FIGURE 3.35 (a) X-ray image of circuit board corrupted by salt-and-pepper noise. (b) Noise reduction with a 3 × 3 averaging mask. (c) Noise reduction with a 3 × 3 median filter. (Original image courtesy of Mr. Joseph E. Pascente, Lixi, Inc.)

<u>Sharpening Spatial Filters :</u>

The principal objective of sharpening is to highlight fine detail in an image or to enhance detail that has been blurred, either in error or as a natural effect of a particular method of image acquisition.

Uses of image sharpening vary and include applications ranging from electronic printing and medical imaging to industrial inspection and autonomous guidance in military systems.

In the last section, we saw that image blurring could be accomplished in the spatial domain by pixel averaging in a neighborhood.

Since averaging is analogous to integration, it is logical to conclude that sharpening could be accomplished by spatial differentiation.

This, in fact, is the case, and the discussion in this section deals with various ways of defining and implementing operators for sharpening by digital differentiation.

Fundamentally, the strength of the response of a derivative operator is proportional to the degree of discontinuity of the image at the point at which the operator is applied.

Thus, image differentiation enhances edges and other discontinuities (such as noise) and deemphasizes areas with slowly varying gray-level values.

Looking some of the fundamental properties of these derivatives in a digital context. To simplify the explanation; we focus attention On one-dimensional derivatives. In particular, we are interested in the behavior of these derivatives in areas of constant gray level (flat segments), at the onset and end of discontinuities (step and ramp discontinuities), and along gray-level ramps. These types of discontinuities can be used to model noise points, lines, and edges in an image. The behavior of derivatives during transitions into and out of these image features also is of interest.

The derivatives of a digital function are defined in terms of differences.

There are various ways to define these differences. However, we require that any definition we use for a first derivative (1) must be zero in flat segments (areas of constant gray-level values); (2) must be nonzero at the onset of a gray-level step or ramp; and (3) must be nonzero along ramps.

Similarly, any definition of a second derivative (1) must be zero in flat areas; (2) must be nonzero at the onset and end of a gray-level step or ramp; and (3) must be zero along ramps of constant slope.

Since we are dealing with digital quantities whose values are finite, the maximum possible gray-level change also is finite, and the shortest distance over which that change can occur is between adjacent pixels.

A basic definition of the first-order derivative of a one-dimensional function *f(x)* is the difference

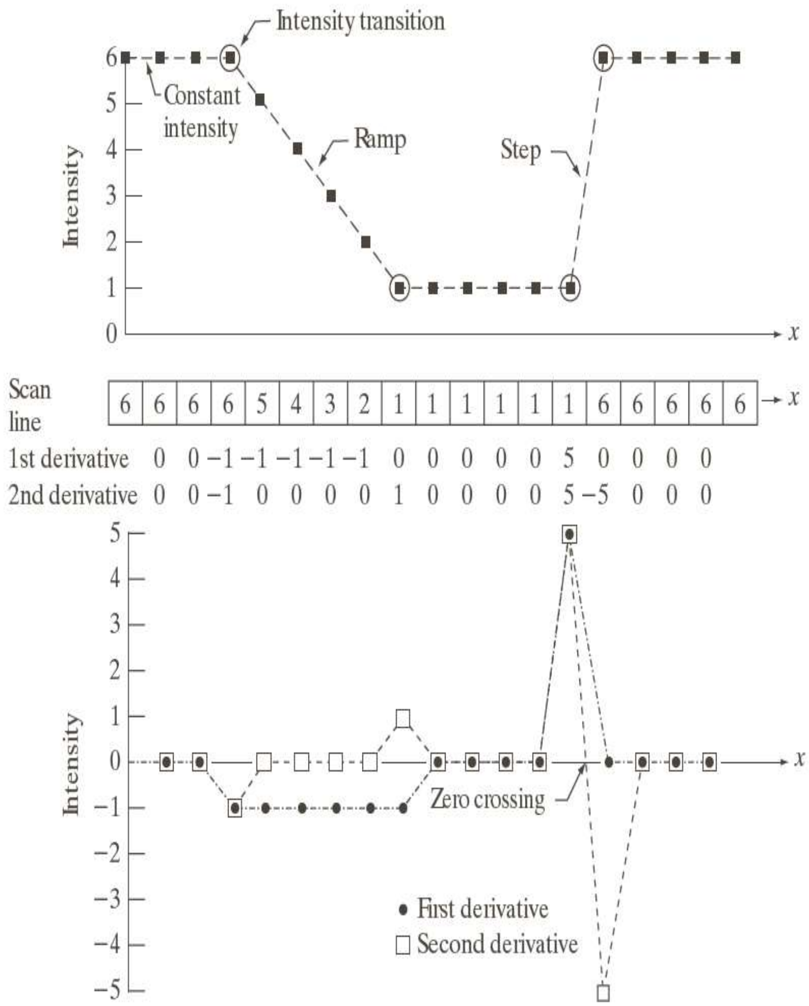
$$\frac{\partial f}{\partial x} = f(x+1) - f(x).$$

Similarly, we define a second-order derivative as the difference

$$\frac{\partial^2 f}{\partial x^2} = f(x+1) + f(x-1) - 2f(x).$$

It is easily verified that these two definitions satisfy the conditions stated previously regarding derivatives of the first and second order.

(I) First -order derivatives generally produce thicker edges in an image and use of first derivatives in image processing. is for edge extraction.
(2) Second-order derivatives have a stronger response to fine detail, such as thin lines and isolated points.
(3) First order derivatives generally have a stronger response to a gray-level step.
(4) Second-order derivatives produce a double response at step changes in gray level.



a
b
c

FIGURE 3.36 Illustration of the first and second derivatives of a 1-D digital function representing a section of a horizontal intensity profile from an image. In (a) and (c) data points are joined by dashed lines as a visualization aid.

Use of Second Derivatives for Enhancement-The Laplacian

The use of two-dimensional, second order derivatives for image enhancement. The approach basically consists of defining a discrete formulation of the second-order derivative and then constructing a filter mask based on that formulation.

We are interested in *isotropic* filters, whose response is independent of the direction of the discontinuities in the image to which the filter is applied. In other words, isotropic filters are *rotation invariant,* in the sense that rotating the image and then applying the filter gives the same result as applying the filter to the image first and then rotating the result.

Development of the method

It can be shown (Rosenfeld and Kak [1982]) that the simplest isotropic derivative operator is the *Laplacian,* which, for a function (image) $f(x, y)$ of two variables, is defined as

Because derivatives of any order are linear operations, the Laplacian is a linear operator.

In order to be useful for digital image processing, this equation needs to be expressed in discrete form. There are several ways to define a digital Laplacian using neighborhoods. The definition of the digital second derivative given in that section is one of the most used. Taking into account that we now have two variables, we use the following notation for the partial second-order derivative in the x-direction:

$$\frac{\partial^2 f}{\partial^2 x^2} = f(x+1, y) + f(x-1, y) - 2f(x, y) \qquad (3.7\text{-}2)$$

and, similarly in the y-direction, as

$$\frac{\partial^2 f}{\partial^2 y^2} = f(x, y+1) + f(x, y-1) - 2f(x, y) \qquad (3.7\text{-}3)$$

The digital implementation of the two-dimensional Laplacian in Eq. (3.7-1) is obtained by summing these two components

$$\nabla^2 f = \left[f(x+1, y) + f(x-1, y) + f(x, y+1) + f(x, y-1)\right] - 4f(x, y). \qquad (3.7\text{-}4)$$

This equation can be implemented using the mask shown in Fig. 3.39(a), which gives an isotropic result for rotations in increments of 90_0

Sharpening Spatial Filters: Foundation

- The first-order derivative of a one-dimensional function f(x) is the difference

$$\frac{\partial f}{\partial x} = f(x+1) - f(x)$$

- The second-order derivative of f(x) as the difference

$$\frac{\partial^2 f}{\partial x^2} = f(x+1) + f(x-1) - 2f(x)$$

Sharpening Spatial Filters: Laplace Operator

The second-order isotropic derivative operator is the Laplacian for a function (image) f(x,y)

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

$$\frac{\partial^2 f}{\partial x^2} = f(x+1, y) + f(x-1, y) - 2f(x, y)$$

$$\frac{\partial^2 f}{\partial y^2} = f(x, y+1) + f(x, y-1) - 2f(x, y)$$

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

Laplacian Mask

Isotropic result for rotations in increments of 90°

0	1	0
1	−4	1
0	1	0

Isotropic result for rotations in increments of 45°

1	1	1
1	−8	1
1	1	1

0	−1	0
−1	4	−1
0	−1	0

−1	−1	−1
−1	8	−1
−1	−1	−1

a b
c d

FIGURE 3.39 (a) Filter mask used to implement the digital Laplacian, as defined in Eq. (3.7-4). (b) Mask used to implement an extension of this equation that includes the diagonal neighbors. (c) and (d) Two other implementations of the Laplacian.

Effect of Laplacian Operator

- Since the Laplacian is a derivative operator
  - it highlights graylevel discontinuities in an image
  - it deemphasizes regions with slowly varying gray levels
- The Laplacian tends to produce images that have
  - grayish edge lines and other discontinuities all superimposed on a dark featureless background

Sharpening Spatial Filters: Laplace Operator

0	1	0	1	1	1
1	−4	1	1	−8	1
0	1	0	1	1	1
0	−1	0	−1	−1	−1
−1	4	−1	−1	8	−1
0	−1	0	−1	−1	−1

a b
c d

FIGURE 3.37 (a) Filter mask used to implement Eq. (3.6-6). (b) Mask used to implement an extension of this equation that includes the diagonal terms. (c) and (d) Two other implementations of the Laplacian found frequently in practice.

Gradient Operator (1)

- The gradient is a vector of directional derivatives.

$$\nabla \mathbf{f} = \begin{bmatrix} f_x \\ f_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$$

- Although not strictly correct, the magnitude of the gradient vector is referred to as the gradient.
- First derivatives are implemented using this magnitude

$$\begin{aligned} \nabla f &= \mathrm{mag}(\nabla \mathbf{f}) \\ &= [f_x^2 + f_y^2]^{1/2} \\ &= \left[\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2\right]^{1/2} \end{aligned}$$

approximation:

$$\nabla f \approx |f_x| + |f_y|$$

3.7.3 Use of First Derivatives for Enhancement—The Gradient

First derivatives in image processing are implemented using the magnitude of the gradient. For a function $f(x, y)$, the gradient of f at coordinates (x, y) is defined as the two-dimensional column *vector*

$$\nabla \mathbf{f} = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \dfrac{\partial f}{\partial x} \\ \dfrac{\partial f}{\partial y} \end{bmatrix} \tag{3.7-12}$$

The magnitude of this vector is given by

$$\begin{aligned} \nabla f &= \text{mag}(\nabla \mathbf{f}) \\ &= \left[G_x^2 + G_y^2\right]^{1/2} \\ &= \left[\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2\right]^{1/2} \end{aligned} \tag{3.7-13}$$

The components of the gradient vector itself are linear operators, but the magnitude of this vector obviously is not because of the squaring and square root operations. On the other hand, the partial derivatives in Eq. (3.7-12) are not rotation invariant (isotropic), but the magnitude of the gradient vector is. Although it is not strictly correct, the magnitude of the gradient vector often is referred to as the *gradient*. In keeping with tradition, we will use this term in the following discussions, explicitly referring to the vector or its magnitude only in cases where confusion is likely.

Sharpening Filters: Derivatives (cont'd)

- The gradient is a **vector** which has magnitude and direction:

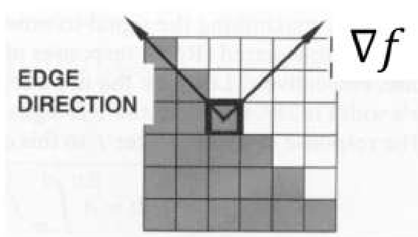
$$magnitude(grad(f)) = \sqrt{\frac{\partial f}{\partial x}^2 + \frac{\partial f}{\partial y}^2} \quad \text{or} \quad \left|\frac{\partial f}{\partial x}\right| + \left|\frac{\partial f}{\partial y}\right|$$

(approximation)

$$direction(grad(f)) = \tan^{-1}\left(\frac{\partial f}{\partial y} / \frac{\partial f}{\partial x}\right)$$

Sharpening Filters: Derivatives (cont'd)

- **Magnitude:** provides information about edge strength.
- **Direction:** perpendicular to the direction of the edge.



Gradient Operator (2)

- The components of the gradient vector are linear operators, but the magnitude is not (square,square root).
- The partial derivatives are not rotation invariant (isotropic), but the magnitude is.
- The Laplacian operator yields a scalar: a single number indicating edge strength at point.
- The gradient is actually a vector from which we can compute edge magnitude and direction.

$$f_{mag}(i,j) = \sqrt{f_x^2 + f_y^2} \quad \text{or} \quad f_{mag}(i,j) = |f_x| + |f_y|$$

$$f_{angle}(i,j) = \tan^{-1}\frac{f_y}{f_x} \qquad \text{where} \qquad \boxed{\begin{array}{l} f_x(i,j) = f(i+1,j) - f(i-1,j) \\ f_y(i,j) = f(i,j+1) - f(i,j-1) \end{array}}$$

The Laplacian is a scalar, giving only the magnitude about the change in pixel values at a point. The gradient gives both magnitude and direction.

Possible Questions

1. Explain about Brightness adaption and discrimination.
2. Describe about Pixel and its coordinates conventions.
3. Illustrate about Spatial domain filtering.
4. Explain about Intensity transformation
5. Elucidate about Contrast stretching.
6. Discuss about Smoothing filters.
7. Explain about Histogram Equalization.
8. Describe about Sharpening filters.

KARPAGAM ACADEMY OF HIGHER EDU

Enable | Enlighten | Enrich
KARPAGAM
ACADEMY OF HIGHER EDUCATION
(Deemed to be University)
(Established Under Section 3 of UGC Act, 1956)

CLASS: III BSC IT COURSE NAME: DIGI
COURSE CODE: 16ITU504A

UNIT 1

	Questions	option1	option2
1	The spatial coordinates of a digital image (	Position	**Brightness**
2	Among the following image processing tec	Optical	**Digital**
3	An image is considered to be a function of	Height of image	Width of image
4	What is pixel?	**Pixel is the elements**	Pixel is the elements of
5	The range of values spanned by the gray s	**Dynamic range**	Band range
6	Which is a colour attribute that describes a	Saturation	**Hue**
7	Which gives a measure of the degree to w	**Saturation**	Hue
8	Which means the assigning meaning to a r	**Interpretation**	Recognition
9	A typical size comparable in quality to mor	256 X 256	**512 X 512**
10	To convert a continuous sensed data into [	Sampling	Quantization
11	To convert a continuous image f(x, y) to dig	Coordinates	Amplitude`
12	For a continuous image f(x, y), how could b	**Digitizing the coordin**	Digitizing the amplitude
13	For a continuous image f(x, y), Quantizatio	Digitizing the coordinat	**Digitizing the amplitu**
14	The quality of a digital image is well deteri	The number of samples	The discrete gray levels
15	Assume that an image f(x, y) is sampled so	**Second sample along**	First sample along seco
16	The resulting image of sampling and quant	Image element or Pictu	Pixel or Pel
17	A continuous image is digitised at _______	random	vertex
18	The transition between continuous values	**Quantisation**	Sampling
19	Images quantised with insufficient brightn	Pixillation	Blurring
20	The smallest discernible change in intensit	**Intensity Resolution**	Contour
21	What is the tool used in tasks such as zoon	Sampling	**Interpolation**
22	The type of Interpolation where for each n	bicubic interpolation	cubic interpolation
23	The type of Interpolation where the intens	cubic interpolation	**nearest neighbour int**
24	Dynamic range of imaging system is a ratic	**Saturation**	Noise
25	For Dynamic range ratio the lower limit is (	Saturation	Brightness
26	Quantitatively, spatial resolution cannot b	line pairs	pixels
27	Of the following, __________ has the maxir	UV Rays	**Gamma Rays**
28	In the Visible spectrum the ______ colour	Violet	Blue
29	Wavelength and frequency are related as :	c = wavelength / freque	frequency = wavelengtl
30	Electromagnetic waves can be visualised a	**sine wave**	cosine wave
31	How is radiance measured?	lumens	**watts**
32	Which of the following is used for chest an	Hard X-Rays	**Soft X-Rays**
33	Which of the following is impractical to me	Frequency	Radiance
34	Massless particle containing a certain amo	**Photon**	Shell
35	What do you mean by achromatic light?	Chromatic light	**Monochromatic light**
36	Which of the following embodies the achr	Luminance	**Brightness**
37	Image processing approaches operating di	Transform domain	**Spatial domain**
38	What is the output of a smoothing, linear s	Median of pixels	Maximum of pixels

39	Which of the following in an image can be	Smooth transitions of g	Smooth transitions of b
40	Which of the following is the disadvantage	**Blur edges**	Blur inner pixels
41	Which of the following comes under the a	Object detection	**Gross representation**
42	Which of the following filters response is b	**Nonlinear smoothing**	Linear smoothing filters
43	Median filter belongs to which category of	Linear spatial filter	Frequency domain filte
44	Which expression is obtained by performin	s=L+1-r	s=L+1+r
45	What is the name of process used to corre	Beta correction	Alpha correction
46	Which of the following transformation fun	Log transformation	Power transformation
47	In which type of slicing, highlighting a spec	**Gray-level slicing**	Bit-plane slicing
48	The process of using known data to estima	Acquisition	**Interpolation**
49	Which of the following is NOT an applicati	Shading Correction	Masking
50	The procedure done on a digital image to a	Neighbourhood Operat	Image Registration
51	______________ is the effect caused by the	Gaussian smooth	Contouring
52	The difference is intensity between the hig	Noise	Saturation
53	The most familiar single sensor used for Im	Microdensitometer	**Photodiode**
54	What is the first and foremost step in Imag	Image restoration	Image enhancement
55	What is the basis for numerous spatial d	Transformations	Scaling
56	In _______ image we notice that the com	bright	**dark**
57	Histogram Equalisation is mainly used fo	**Image enhancement**	Blurring
58	What is the output of a smoothing, linear	Median of pixels	Maximum of pixels
59	Which of the following is the disadvantag	Blur edges	Blur inner pixels
60	Which of the following filters response	**Nonlinear smoothing**	Linear smoothing filte

JCATION

TAL IMAGE PROCESSING
BATCH-2016-2019

option3	option4	Answer
Contrast	Noise	**Brightness**
Electronic	Photographic	**Digital**
Amplitude of image	Resolution of image	**Amplitude of image**
Pixel is the cluster of a	Pixel is the cluster of ar	**Pixel is the elements of a digital image**
Peak range	Resolution range	**Dynamic range**
Brightness	Intensity	**Hue**
Intensity	Brightness	**Saturation**
Acquisition	Segmentation	**Interpretation**
1920 X 1080	1080 X 1080	**512 X 512**
Both Sampling and Q	Neither Sampling nor C	**Both Sampling and Quantization**
All of the mentioned	None of the mentioned	**All of the mentioned**
All of the mentioned	None of the mentioned	**Digitizing the coordinate values**
All of the mentioned	None of the mentioned	**Digitizing the amplitude values**
All of the mentioned	None of the mentioned	**All of the mentioned**
First sample along first	Second sample along s	**Second sample along first row**
All of the mentioned	None of the mentioned	**All of the mentioned**
contour	**sampling**	**sampling**
Rasterisation	None of the Mentioned	**Quantisation**
False Contours	None of the Mentioned	**False Contours**
Saturation	Contrast	**Intensity Resolution**
Filters	None of the Mentioned	**Interpolation**
bilinear interpolation	**nearest neighbour int**	**nearest neighbour interpolation**
bilinear interpolation	bicubic interpolation	**nearest neighbour interpolation**
Brightness	Contrast	**Saturation**
Noise	Contrast	**Noise**
dots	**none of the Mentione**	**none of the Mentioned**
Microwaves	Radio Waves	**Gamma Rays**
Red	Yellow	**Red**
wavelength = c * freque	**c = wavelength * frequ**	**c = wavelength * frequency**
tangential wave	None of the mentioned	**sine wave**
armstrong	hertz	**watts**
Radio waves	Infrared Rays	**Soft X-Rays**
Luminance	**Brightness**	**Brightness**
Electron	None of the mentioned	**Photon**
Infrared light	Invisible light	**Monochromatic light**
Frequency	Radiance	**Brightness**
Inverse transformation	None of the Mentioned	**Spatial domain**
Minimum of pixels	**Average of pixels**	**Average of pixels**

Sharp transitions of g	Sharp transitions of brig	**Sharp transitions of gray levels**
Remove sharp transitic	Sharp edges	**Blur edges**
Object motion	Image segmentation	**Gross representation**
Sharpening filters	Geometric mean filter	**Nonlinear smoothing filters**
Order static filter	Sharpening filter	**Order static filter**
s=L-1-r	s=L-1+r	**s=L-1-r**
Gamma correction	Pie correction	**Gamma correction**
Piece-wise transforma	Linear transformation	**Piece-wise transformation**
Contrast stretching	Byte-level slicing	**Gray-level slicing**
Pixelation	None of the Mentioned	**Interpolation**
Pixelation	Region of Interest opera	**Pixelation**
Geometric Spacial Tran	**Single Pixel Operation**	**Single Pixel Operation**
False Contouring	Interpolation	**False Contouring**
Contrast	Brightness	**Contrast**
CMOS	None of the Mentioned	**Photodiode**
Image acquisition	Segmentation	**Image acquisition**
Histogram	None of the Mentioned	**Histogram**
colourful	All of the Mentioned	**dark**
Contrast adjustment	None of the Mentioned	**Image enhancement**
Minimum of pixels	Average of pixels	**Average of pixels**
Remove sharp transition	Sharp edges	**Blur edges**
Sharpening filters	Geometric mean filter	**Nonlinear smoothing filters**

KARPAGAM ACADEMY OF HIGHER EDUCATI(

CLASS: III BSC IT COURSE NAME: DIGITAL

COURSE CODE: 16ITU504A

UNIT II

	Questions	option1	option2	option3
1	What is the name of the filter that is us	Unsharp mask filt	**Notch filter**	Zero-phase-shift-
2	Which of the following filter(s) attenu	Unsharp mask filt	**Lowpass filter**	Zero-phase-shift f
3	Which of the following filter(s) attenu	Unsharp mask filt	**Highpass filter**	Zero-phase-shift f
4	Which of the following filter have a le	Highpass filter	**Lowpass filter**	Zero-phase-shift f
5	The feature(s) of a highpass filtered in	Have less gray-le	Emphasized trans	An overall sharpe
6	A spatial domain filter of the correspo	Fourier transform	**Inverse Fourier**	None of the ment
7	A frequency domain filter of the corre	**Fourier transfor**	Inverse Fourier tr	None of the ment
8	Which of the following filtering is don	**Gaussian filterin**	Unsharp mask filt	High-boost filteri
9	Which of the following is/are consider	Ideal	Butterworth	Gaussian
10	Which of the following lowpass filters	**Ideal lowpass filt**	Butterworth lowp	Gaussian lowpass
11	Which of the following lowpass filters	**Ideal lowpass filt**	Butterworth lowp	Gaussian lowpass
12	The characteristics of the lowpass filte	Has a dominant c	Has a concentric,	**All of the mentio**
13	In frequency domain terminology, whi	Emphasis filtering	**Unsharp maskin**	Butterworth filter
14	Which of the following is/ are a genera	Lowpass filtering	**High-boost filter**	Emphasis filtering
16	Subtracting Laplacian from an image i	**Unsharp maskin**	Box filter	Median filter
17	A First derivative in image processing	**Magnitude of G**	The Laplacian	All of the mention
18	What is the sum of the coefficient of tl	1	-1	**0**
19	Gradient is used in which of the follow	To aid humans in	As a preprocessin	**All of the mentio**
20	Gradient have some important feature	Enhancing small c	Enhancing promi	**All of the mentio**
21	An image has significant edge details.	**The gradient im**	The gradient imag	Both the gradient
22	Fourier transform of unit impulse at o	undefined	infinity	**1**
23	Continuous functions are sampled to f	Fourier series	Fourier transform	fast Fourier series
24	The common example of 2D interpola	enhancement	sharpening	blurring
25	2D Fourier transform and its inverse a	aperiodic	**periodic**	linear
27	Shrinking of image can be done using	pixel replication	bicubic interpolat	**bilinear interpol**
28	Sum of many infinitely many periodic	aperiodic impulse	periodic impulse	**impulse train**
29	To reduce the effect of aliasing high fr	**attenuated**	accentuated	reduced
30	If f(x,y) is imaginary, then its Fourier	conjugate symme	hermition	**antihermition**
31	The product of two even or two odd f	**even**	odd	prime
32	Unit impulse at every point other than	undefined	infinity	1
33	The sum of cosines and sines coefficie	**Fourier series**	Fourier transform	fast Fourier series
34	A continuous band limited function ca	2D sampling serie	3D sampling theo	1D sampling theo
35	The impulse S(t,v) is function having	one variable	**two variables**	three variables
36	Any function whose Fourier transform	high pass functior	low pass function	**band limited fun**
37	Fourier transform's domain is	**frequency doma**	spatial domain	Fourier domain
38	Giving one period of the periodic conv	periodic convolut	aperiodic convolu	correlation
39	The continuous variables in 2D transfo	continuous freque	spatial variables	**continuous spati**
40	Fourier transform of two continuous f	Fourier series pai	**Fourier transfor**	Fourier series

41	Forward and inverse Fourier transforn	integers	infinite	**finite**
42	The greater, the values of continuous	**contracted**	expanded	discrete
43	Low pass filters are used for image	contrast	sharpening	**blurring**
44	Fourier stated that the periodic functic	sine	cosine	tangent
45	The effect caused by under sampling i	smoothing	sharpening	summation
46	Product of two functions in spatial do	correlation	**convolution**	Fourier transform
47	Digitizing the coordinate values is call	quantization	**sampling**	Fourier transform
48	Shrinking of image is viewed as	**under sampling**	over sampling	critical sampling
49	High pass filters are used for image	contrast	**sharpening**	blurring
50	The Fourier transform is named after l	**joseph Fourier**	john Fourier	sean Fourier
51	Which of the following is the primary	Blurring the imag	**Highlight fine de**	Increase the brigh
52	In spatial domain, which of the follow	Integration	Average	Median
53	In which of the following cases, we w	Flat segments	Step discontinuiti	Ramp discontinui
54	Which of the following is the valid res	Non-zero at flat s	Zero at the onset	**Zero in flat segm**
55	Which of the following is not a valid r	Zero response at	**Nonzero respons**	Zero response at
56	If f(x,y) is an image function of two va	**f(x+1)-f(x)**	f(x)-f(x+1)	f(x-1)-f(x+1)
57	What is the thickness of the edges pro	Finer	Equal	**Thicker**
58	Which of the following derivatives pro	First order deriva	Third order deriv	**Second order de**
59	The domain that refers to image plane	Spatial domain in	Frequency domai	**Spatial domain a**
60	What is accepting or rejecting certain	**Filtering**	Eliminating	Slicing
61	What is the process of moving a filter	Convolution	**Correlation**	Linear spatial filte
62	Convolution and Correlation are funct	Distance	Time	Intensity

ON

IMAGE PROCESSING
BATCH-2016-2019

option4	Answers
None of the mentioned	**Notch filter**
All of the mentioned	**Lowpass filter**
All of the mentioned	**Highpass filter**
None of the mentioned	**Lowpass filter**
All of the mentioned	**All of the mentioned**
All of the mentioned	**Inverse Fourier transform**
All of the mentioned	**Fourier transform**
None of the mentioned	**Gaussian filtering**
All of the mentioned	**All of the mentioned**
All of the mentioned	**Ideal lowpass filters**
All of the mentioned	**Ideal lowpass filters**
None of the mentioned	**All of the mentioned**
None of the mentioned	**Unsharp masking**
All of the mentioned	**High-boost filtering**
None of the mentioned	**Unsharp masking**
None of the mentioned	**Magnitude of Gradient vector**
None of the mentioned	**0**
None of the mentioned	**All of the mentioned**
None of the mentioned	**All of the mentioned**
None of the mentioned	**The gradient image is brighter than the Laplacian image**
0	**1**
digital image	**digital image**
resizing	**resizing**
non linear	**periodic**
row column deletion	**bilinear interpolation**
summation	**impulse train**
removed	**attenuated**
symmetry	**antihermition**
aliasing	**even**
0	**0**
fast Fourier transform	**Fourier series**
2D sampling theorem	**2D sampling theorem**
four variables	**two variables**
band pass function	**band limited function**
time domain	**frequency domain**
circular convolution	**circular convolution**
discrete spatial variables	**continuous spatial variables**
Fourier transform	**Fourier transform pair**

discrete	**finite**
continuous	**contracted**
resizing	**blurring**
Both A and B	**Both A and B**
aliasing	**aliasing**
fast Fourier transform	**convolution**
discrete Fourier transform	**sampling**
nyquist sampling	**under sampling**
resizing	**sharpening**
jay Fourier	**joseph Fourier**
Decrease the brightness of th	**Highlight fine details in the image**
Differentiation	**Differentiation**
Slow varying gray values	**Slow varying gray values**
Zero along ramps	**Zero in flat segments**
Nonzero response along the	**Nonzero response at onset of gray level step**
f(x)+f(x-1)	**f(x+1)-f(x)**
Independent	**Thicker**
First and second order deriva	**Second order derivative**
Frequency domain and Spati	**Spatial domain and Frequency domain respectively**
None of the Mentioned	**Filtering**
Non linear spatial filtering	**Correlation**
Displacement	**Displacement**

KARPAGAM ACADEMY OF HIGHER EDUCATION

CLASS: III BSC CS COURSE NAME: DIGITAL IMAGE PROCESSING

COURSE CODE: 16ITU504A BATCH-2016-2019

UNIT III

SNO	QUESTIONS	CHOICE 1	CHOICE 2	CHOICE 3	CHOICE 4	ANSWER
1	Image restoration is to improve the _______________of the image.	quality	noise	intensity	colour	quality
2	_________ is the operation of taking a corrupt/noisy image and estimating the clean, original image.	Image Restoration	Image Restoration	Image Restoration	Image Restoration	Image Restoration
3	The _________is responsible for reducing or eliminating any coding, interpixel, or psychovisual redundancies in the input image.	source encoder	quantizer	mapper	decoder	source encoder
4	Due to uniform linear motion image is_________________	blurred	sharpened	smoothened	blurred and smoothened	blurred and smoothened
5	Blur is characterized by the _______________ response of the system	filter	noise	impulse	image	impulse
6	Gaussian noise is referred to as __________	red noise	black noise	white noise	normal noise	normal noise
7	Convolution in spatial domain is multiplication in ___________	frequency domain	time domain	spatial domain	plane	frequency domain
8	Linear functions possesses the property of _____________	additivity	homogeneity	multiplication	Both A and B	Both A and B
9	PDF in image processing is called ______________	probability degraded function	probability density function	probabilistic degraded function	probabilistic density function	probability density function
10	Filter that replaces the pixel value with the medians of intensity levels is ___________	arithmetic mean filter	geometric mean filter	median filter	sequence mean filter	median filter
11	The purpose of restoration is to gain ________	degraded image	original image	pixels	coordinates	original image

12	Degraded image is produced using degradation process and ___________	additive noise	destruction	pixels	coordinates	additive noise
13	Degraded image is given in a ___________	frequency domain	time domain	spatial domain	plane	spatial domain
14	Digitizing the image intensity amplitude is called ___________	sampling	quantization	framing	Both A and B	quantization
15	Compressed image can be recovered back by _____________	image enhancement	image decompression	image contrast	image equalization	image decompression
16	Digital video is sequence of ____________	pixels	frames	matrix	coordinates	frames
17	Image compression comprised of ______________	encoder	decoder	frames	encoder and decoder	encoder and decoder
18	Information is the _______________	data	meaningful data	raw data	Both A and B	meaningful data
19	Coding redundancy works on ______________	pixels	matrix	intensity	coordinates	intensity
20	Sequence of code assigned is called ______________	code word	word	byte	nibble	code word
21	Every run length pair introduce new ______________	pixels	matrix	frames	intensity	intensity
22	If the pixels are reconstructed without error mapping is said to be ______________	reversible	irreversible	temporal	facsimile	reversible
23	If the $P_{(E)} = 1$, it means event ______________	does not occur	always occur	no probability	normalization	always occur
24	The basic idea behind Huffman coding is to ______________	expand data by using fewer bits to encode more frequently occuring characters	compress data by using fewer bits to encode fewer frequently occuring characters	compress data by using more bits to encode more frequently occuring characters	compress data by using fewer bits to encode more frequently occuring characters	compress data by using fewer bits to encode more frequently occuring characters
25	Huffman coding is an encoding algorithm used for ______________	lossy data compression	lossless data compression	files greater than 1 Mbit	broadband systems	lossless data compression
26	A Huffman encoder takes a set of characters with fixed length and produces a set of characters of ______________	constant length	fixed length	random length	variable length	variable length
27	A Huffman code: A = 1, B = 000, C = 001, D = 01 ,P(A) = 0.4, P(B) = 0.1, P(C) = 0.2, P(D) = 0.3, The average number of bits per letter is	8.0 bit	2.1 bit	2.0 bit	1.9 bit	1.9 bit

28	The idea with wavelets is to represent a complicated function by _____________	square functions	lines	simple basic functions	sinus functions	simple basic functions
29	Down sampling is to make a digital image file smaller by _____________	removing pixels	removing noise	adding pixels	adding noise	removing pixels
30	Without losing quality, JPEG-2000 can achieve compression ratios of _____________	200:01:00	2000:01:00	20:01	2:01	200:1
31	The best visual compression quality is achieved using _____________	Wavelets	Fourier transform	Dolby	DCT	Wavelets
32	In the coding redundancy technique we use _____________	fixed length code	variable length code	byte	Both A and B	Both A and B
33	Morphology refers to _____________	pixels	matrix	frames	shape	shape
34	FAX is an abbreviation of _____________	fast	female	feminine	facsimile	facsimile
35	Source of information depending on finite no of outputs is called _____________	markov	finite memory source	zero source	Both A and B	Both A and B
36	Types of data redundancy are _____________	1	2	3	4	3
37	Information per source is called _____________	sampling	quantization	entropy	normaliza	entropy
38	Image with very high quality is considered as _____________	good	fair	bad	excellent	excellent
39	Range [0, L-1], where L is the _____________	no of levels	length	no of intensity levels	low quality	no of intensity levels
40	Compression is done for saving _____________	storage	bandwidth	money	Both A and B	Both A and B
41	System of symbols to represent event is called _____________	storage	word	code	nibble	Ccode
42	In the image MxN, M is _____________	rows	column	level	intensity	rows
43	In the image MxN, N is _____________	rows	column	level	intensity	column
44	Inferior image is the image having _____________	low definition	high definition	intensity	coordinates	low definition
45	Histogram equalization refers to image _____________	sampling	quantization	framing	normalization	normalization
46	The simple way to compression is removing	data	**redundant data**	information	meaningful data	**redundant data**
47	Shannons theorem is also called	**noiseless coding theorem**	noisy coding theorem	coding theorem	noiseless theorem	**noiseless coding theorem**

48	Information lost when expressed mathematically is called	markov	finite memory source	**fidelity criteria**	noiseless theorem	**fidelity criteria**
49	Error of the image is referred to as	pixels	matrix	frames	**noise**	**noise**
50	_______ and ________ are two important statistical measures on which the adaptive filtering depends upon.	Mean and variance	Sum and average	Mean and median	Max and Min	Mean and variance
51	In ________ data compression, the integrity of the data is preserved	lossless	lossy	image	text	lossless
52	When coding the symbols of an information source individually, ___________ yields the smallest possible number of code symbols per source symbol.	Huffman coding	Arithmetic coding	LZW coding	Transform coding	Huffman coding
53	___________ is an example of a category of algorithms called dictionary-based encoding.	Huffman coding	Arithmetic coding	Lempel Ziv (LZ) encoding	Transform coding	Lempel Ziv (LZ) encoding
54	_________is associated with the representation of information.	Coding redundancy	Interpixel redundancy	Psychovisual redundancies	Temporal Redundancy	Coding redundancy
55	___________, which deals with the assignment of gray levels to pixels in the spatially transformed image	gray-level interpolation	a spatial transformation	linear transformation	nonlinear transformation	gray-level interpolation
56	__________is due to the correlation between the neighboring pixels in an image.	Interpixel redundancy	Interpixel redundancy	Psychovisual redundancies	Temporal Redundancy	Interpixel redundancy
57	In ____________, a reversible, linear transform (such as the Fourier transform) is used to map the image into a set of transform coefficients, which are then quantized and coded.	Huffman coding	Arithmetic coding	Lempel Ziv (LZ) encoding	Transform coding	Transform coding
58	_________ rejects or pass frequencies in predefined neighborhoods about a center frequency	Adaptive filter	Notch filter	Mean filter	Max filter	Notch filter
59	_________ exist because human perception does not involve quantitative analysis of every pixel or luminance value in the image.	Psychovisual redundancies	Interpixel redundancy	Psychovisual redundancies	Temporal Redundancy	Psychovisual redundancies
60	__________, which defines the "rearrangement" of pixels on the image plane	a spatial transformation	gray-level interpolation	linear transformation	nonlinear transformation	a spatial transformation

UNIT – IV

FAX compression (CSUITT Group-3 and Group-4), Symbol-based coding, JBIG-2, Bit-plane encoding, Bit-allocation, Zonal Coding, Threshold Coding, JPEG, Lossless predictive coding, Lossy predictive coding, Motion Compensation .
Wavelet based Image Compression: Expansion of functions, Multi-resolution analysis, Scaling functions, MRA refinement equation, Wavelet series expansion, Discrete Wavelet Transform (DWT),Continuous Wavelet Transform, Fast Wavelet Transform, 2-D wavelet Transform, JPEG-2000 encoding, Digital Image Watermarking

FAX COMPRESSION (CSUITT GROUP-3 AND GROUP-4)

Fax Group is an encoding format used for fax transmission. There are two types: Fax Group 3, also known as G3, and Fax Group 4, also known as G4. Fax Group 3 and 4 are two of the encoding formats for Tagged Image File Format (TIFF) files. The more commonly used format, Fax Group 3, is Recommendation T.4 of the CCIT , now known as the ITU-T (for Telecommunication Standardization Sector of the International Telecommunications Union).

Fax Group 3 supports one-dimensional image compression of black and white images, On a standard fax machine, Fax Group 3 uses redundancy reduction to enhance speed and is able to transmit a page in one minute or less. Fax Group 3 can achieve compression ratios of 10:1 for office documents and 15:1 for engineering drawings with a resolution of 200 dots per inch (dpi).

Less frequently used, Fax Group 4 (G4) is ITU-T Recommendation T.6 and supports two-dimensional image compression, compressing the line width as well as the line length. Fax Group 4 can achieve compression ratios of 15:1 for office documents and 20:1 for engineering drawings with a resolution of 400 dpi. Unlike Fax Group 3, Fax Group 4 can use Integrated Services Digital Network (ISDN) for transmission.

CCITT Group 4 compression, also referred to as **G4** or **Modified Modified READ** (MMR), is a lossless method of image compression used in Group 4 fax machines defined in the ITU-T T.6 fax standard. It is only used for bitonal (black and white) images. Group 4 compression is based on the Group 3 two-dimensional compression scheme (G3-2D), also known as Modified READ, which is in turn based on the Group 3 one-dimensional compression scheme (G3), also known as Modified Huffman coding. Group 4 compression is available in many proprietary image file formats as well as standardized formats such as TIFF, CALS, CIT (Intergraph Raster Type 24) and the PDF document format.

G4 offers a small improvement over G3-2D by removing the end of line (EOL) codes. G3 and G4 compression both treat an image as a series of horizontal black strips on a white page. Better compression is achieved when there are fewer unique black dots/lines on the page. Both G3-2D and G4 add a two dimensional feature to achieve greater compression by taking advantage of vertical symmetry. A worst-case image would be an alternating pattern of single-pixel black and white dots offset by one pixel on even/odd lines. G4 compression would actually increase the file size on this type of image. G4 typically achieves a 20:1 compression ratio. For an 8.5"×11" page scanned at 200 DPI, this equates to a reduction from 467.5 kB to 23.4 kB (95% compression ratio).

Group 3 & Group 4

GROUP 3

•MODIFIED HUFFMAN METHOD (MHM) – Unidimensional coding method based on the coding of the length of alternate black and white pixel runs using Huffman coding.

GROUP 4 (Also Group 3 Options)

•MODIFIED READ METHOD (MRM) – Bidimensional coding method based on the coding of the variations of the positions of tone transition pixels (black-white or white-black) in relation to the previous line; unidimensional coding may be used every k lines.

•MODIFIED-MODIFIED READ METHOD (MMRM) – Similar to MRM but without periodic unidimensional coding.

JBIG-2

JBIG2 is an image compression standard for bi-level images, developed by the Joint Bi-level Image Experts Group. It is suitable for both lossless and lossy compression. According to a press release[1] from the Group, in its lossless mode JBIG2 typically generates files 3–5 times smaller than Fax Group 4 and 2–4 times smaller than JBIG, the previous bi-level compression standard released by the Group. JBIG2 has been published in 2000 as the international standard ITU T.88,[2] and in 2001 as ISO/IEC 14492.[3]

JBIG • Lossless Compression. • Progressive Coding. • Sequential Coding. • Arithmetic Encoder/Decoder. • Resolution Reduction Algorithm • (optional, can be replaced).

Resolution Reduction Technique • Creates low resolution images. • Combines decimation and filtering in one action. • Uses 9 high resolution and 3 low resolution pixels to determine color of target pixel. • Preserves gray-levels achieved with halftoning.

JBIG Pro's and Con's ☺ Progressive (for binary images). ☺ Better compression for images with up to ☺ 6 bit/pixel (Vs. JPEG) ☺ Better compression than G3 and G4. / Slow and complicated (Vs. JPEG, G.3/4) / Consumes memory resources and needs frame buffers.

Compression Comparison (Cont'd) Typical results for: • Scanned text and line drawings: JBIG ~20-25% better than G4 • Computer generated line-drawing: JBIG ~75% better than G4 • Scanned dither halftones images: JBIG ~85%-90% better than G4

Functionality

Ideally, a JBIG2 encoder will segment the input page into regions of text, regions of halftone images, and regions of other data. Regions that are neither text nor halftones are typically compressed using a context-dependent arithmetic coding algorithm called the MQ coder. Textual regions are compressed as follows: the foreground pixels in the regions are grouped into symbols. A dictionary of symbols is then created and encoded, typically also using context-dependent arithmetic coding, and the regions are encoded by describing which symbols appear where. Typically, a symbol will correspond to a character of text, but this is not required by the compression method.

For lossy compression the difference between similar symbols (e.g., slightly different impressions of the same letter) can be neglected; for lossless compression, this difference is taken into account by compressing one similar symbol using another as a template. Halftone images may be compressed by reconstructing the grayscale image used to generate the halftone and then sending this image together with a dictionary of halftone patterns.[4] Overall, the algorithm used by JBIG2 to compress text is very similar to the JB2 compression scheme used in the DjVu file format for coding binary images.

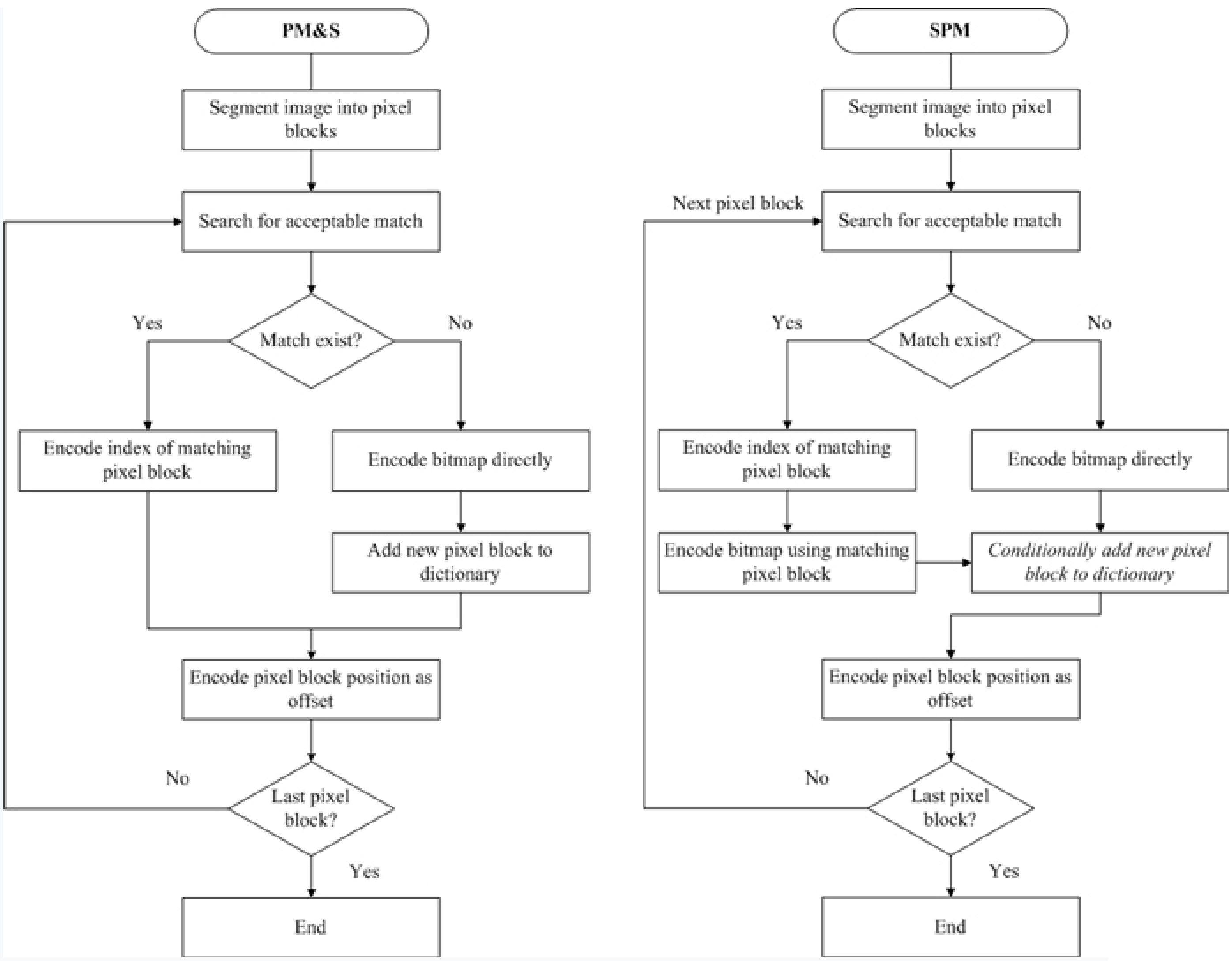
PDF files versions 1.4 and above may contain JBIG2-compressed data. Open-source decoders for JBIG2 are jbig2dec,[5] the java-based jbig2-imageio[6] and the decoder found in versions 2.00 and above of xpdf. An open-source encoder is jbig2enc.[7]

Technical details

Typically, a bi-level image consists mainly of a large amount of textual and halftone data, in which the same shapes appear repeatedly. The bi-level image is segmented into three regions: text, halftone, and generic regions. Each region is coded differently and the coding methodologies are described in the following passage.

Text image data

Text coding is based on the nature of human visual interpretation. A human observer cannot tell the difference between two instances of the same characters in a bi-level image even though they may not exactly match pixel by pixel. Therefore, only the bitmap of one representative character instance needs to be coded instead of coding the bitmaps of each occurrence of the same character individually. For each character instance, the coded instance of the character is then stored into a "symbol dictionary".[8] There are two encoding methods for text image data: pattern matching and substitution (PM&S) and soft pattern matching (SPM). These methods are presented in the following subsections.[9]



Block diagrams of (left) pattern matching and substitution method and (right) soft pattern matching method

Pattern matching and substitution

After performing image segmentation and match searching, and if a match exists, we code an index of the corresponding representative bitmap in the dictionary and the position of the character on the page. The position is usually relative to another previously coded character. If a match is not found, the segmented pixel block is coded directly and added into the dictionary. Typical procedures of pattern matching and substitution algorithm are displayed in the left block diagram of the figure above.

Although the method of PM&S can achieve outstanding compression, substitution errors could be made during the process if the image resolution is low.

Soft pattern matching

In addition to a pointer to the dictionary and position information of the character, refinement data is also required because it is a crucial piece of information used to reconstruct the original character in the image. The deployment of refinement data can make the character-substitution error mentioned earlier highly unlikely. The refinement data contains the current desired character instance, which is coded using the pixels of both the current character and the matching character in the dictionary. Since it is known that the current character instance is highly correlated with the matched character, the prediction of the current pixel is more accurate.

Halftones

Halftone images can be compressed using two methods. One of the methods is similar to the context-based arithmetic coding algorithm, which adaptively positions the template pixels in order to obtain correlations between the adjacent pixels. In the second method, descreening is performed on the halftone image so that the image is converted back to grayscale. The converted grayscale values are then used as indexes of fixed-sized tiny bitmap patterns contained in a halftone bitmap dictionary. This allows decoder to successfully render a halftone image by presenting indexed dictionary bitmap patterns neighboring with each other.

Arithmetic entropy coding

All three region types including text, halftone, and generic regions may all use arithmetic coding. JBIG2 specifically uses the MQ coder.

BIT-PLANE CODING

Another effective technique for reducing an image's interpixel redundancies is to process the image's bit planes individually. The technique, called *bit-plane coding,* IS based on the concept of decomposing a multilevel (monochrome or color) image into a series of binary images and compressing each binary image via one of several well-known binary compression methods. In this section, we describe the most popular decomposition approaches and review several of the more commonly used compression methods. Bit-plane decomposition
The gray levels of an *m-bit* gray-scale Image can be represented m the form of
the base 2 polynomial

$$a_{m-1}2^{m-1} + a_{m-2}2^{m-2} + \quad + a_1 2^1 + a_0 2^0. \qquad (8.4\text{-}2)$$

Based on this property, a simple method of decomposing the image into a collection of binary images is to separate the *m* coefficients of the polynomial into *m* 1-bit *bit planes.*

BIT ALLOCATION

The reconstruction error associated with the truncated series expansion of Eq. (8.5-38) is a function of the number and relative importance of the transform coefficients that are discarded, as well as the precision that is used to represent the retained coefficients. In most transform coding systems, the retained coefficients are selected [that is, the masking function of Eq. (8.5-37) is constructed] on the basis of maximum variance, called *zonal coding,* or on the basis of maximum magnitude, called *threshold coding.* The overall process of truncating, quantizing, and coding the coefficients of a transformed subimage is commonly called *bit allocation.*

Zonal Coding

Zonal coding implementation Zonal coding is based on the information theory concept of viewing information as uncertainty. Therefore the transform coefficients of maximum variance carry the most image information and should be retained in the coding process. The variances themselves can be calculated directly from the ensemble of *(N/ n?* transformed subimage arrays. as in the preceding example, or based on an assumed image model (say. a Markov autocorrelation function). In either case, the zonal sampling process can be viewed, in accordance with Eq. (8.5-38). as multiplying each *T(II, v)* by the corresponding element in a *zonal mask,* which is constructed by placing a 1 in the locations of maximum variance and a 0 in all other locations. Coefficients of maximum variance usually are located around the origin of an image transform. resulting in the typical zonal mask shown in Fig. 8.36(a).

The coefficients retained during the zonal sampling process must be quantized and coded, so zonal masks are sometimes depicted showing the number of bits used to code each coefficient [Fig. 8.36(b)]. In most cases, the coefficients are allocated the same number of bits, or some fixed number of bits is distributed among them unequally. In the first case, the coefficients generally are normalized by their standard deviations and uniformly quantized. In the second case. a quantizer. such as an optimal Lloyd-Max quantizer. is designed for each coefficient. To construct the required quantizers, the zeroth or dc coefficient normally is modeled by a Rayleigh density function, whereas the remaining coefficients are modeled by a Laplacian or Gaussian density.[†] The number of quantization levels (and thus the number of bits) allotted to each quantizer is made proportional to $\log_2 \sigma^2_{T(u,v)}$. This allocation is consistent with rate distortion theory, which indicates that a Gaussian random variable of variance σ^2 cannot be represented by less than $\frac{1}{2}\log_2(\sigma^2/D)$ bits and be reproduced with mean-square error less than D (see Problem 8.11). The intuitive conclusion is that the information content of a Gaussian random variable is proportional to $\log_2(\sigma^2/D)$. Thus the retained coefficients in Eq. (8.5-38)—which (in the context of the current discussion) are selected on the basis of maximum variance—should be assigned bits in proportion to the logarithm of the coefficient variances.

1	1	1	1	1	0	0	0
1	1	1	1	0	0	0	0
1	1	1	0	0	0	0	0
1	1	0	0	0	0	0	0
1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

8	7	6	4	3	2	1	0
7	6	5	4	3	2	1	0
6	5	4	3	3	1	1	0
4	4	3	3	2	1	0	0
3	3	3	2	1	1	0	0
2	2	1	1	1	0	0	0
1	1	1	0	0	0	0	0
0	0	0	0	0	0	0	0

1	1	0	1	1	0	0	0
1	1	1	1	0	0	0	0
1	1	0	0	0	0	0	0
1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

0	1	5	6	14	15	27	28
2	4	7	13	16	26	29	42
3	8	12	17	25	30	41	43
9	11	18	24	31	40	44	53
10	19	23	32	39	45	52	54
20	22	33	38	46	51	55	60
21	34	37	47	50	56	59	61
35	36	48	49	57	58	62	63

a b
c d

FIGURE 8.36 A typical (a) zonal mask, (b) zonal bit allocation, (c) threshold mask, and (d) thresholded coefficient ordering sequence. Shading highlights the coefficients that are retained.

Threshold Coding

Thresholding Coding Implementation Zonal coding usually is implemented by using a single fixed mask for all subimages. Threshold coding. however. is inherently adaptive in the sense that the location of the transform coefficients retained for each subimage vary from one subimage to another. In fact. threshold coding is the adaptive transform coding approach most often used in practice because of its computational simplicity. The underlying concept is that. for an\' subimage. the transform coefficients of largest magnitude make the most significant contribution to reconstructed subimage quality. as demonstrated in the last example.

Because the locations of the maximum coefficients vary from one subimage to another. the elements of Y(u,v)*T(u. v)* normally are reordered (in a predefined manner) to form a I-D. run-length coded sequence. Figure *8.3h(c)* shows a *typical threshold mask* for one subimage of a hypothetical image. This mask provides a convenient way to visualize the threshold coding process for the corresponding subimage. as well as to mathematically describe the process using Eq. (8.5-38).

When the mask is applied [via Eq. (8.5-38)] to the subimage for which it was derived. and the resulting *n* x *n* array is reordered to form an n'-element coefficient sequence in accordance with the zigzag ordering pattern of Fig. 8.34(d). the reordered 1-0 sequence contains several long runs of O's [the zigzag pattern becomes evident by starting at 0 in Fig. 8.36(d) and following the numbers in sequence]. These runs normally are run-length coded. The nonzero or retained coefficients. corresponding to the mask locations that contain a L are represented using one of the variable-length codes of Section 8.4.

There are three basic ways to threshold a transformed subimage or. stated differently. to create a subimage threshold masking function of the form given in Eq. (8.5-37): (l) A single global threshold can be applied to all subimages: (2) a different threshold can be used for each subimage: or (3) the threshold can be varied as a function of the location of each coefficient within the subimage. **In** the first approach. the level of compression differs from image to image. depending on the number of coefficients that exceed the global threshold.

JPEG

Image compression

Image compression is the method of data compression on digital images.

The main objective in the image compression is:

- Store data in an efficient form
- Transmit data in an efficient form

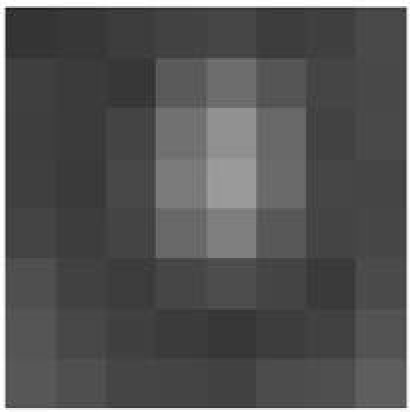
Image compression can be lossy or lossless.

JPEG compression

JPEG stands for Joint photographic experts group. It is the first interanational standard in image compression. It is widely used today. It could be lossy as well as lossless . But the technique we are going to discuss here today is lossy compression technique.

How jpeg compression works

First step is to divide an image into blocks with each having dimensions of 8 x8.



Let's for the record, say that this 8x8 image contains the following values.

$$\begin{bmatrix} 52 & 55 & 61 & 66 & 70 & 61 & 64 & 73 \\ 63 & 59 & 55 & 90 & 109 & 85 & 69 & 72 \\ 62 & 59 & 68 & 113 & 144 & 104 & 66 & 73 \\ 63 & 58 & 71 & 122 & 154 & 106 & 70 & 69 \\ 67 & 61 & 68 & 104 & 126 & 88 & 68 & 70 \\ 79 & 65 & 60 & 70 & 77 & 68 & 58 & 75 \\ 85 & 71 & 64 & 59 & 55 & 61 & 65 & 83 \\ 87 & 79 & 69 & 68 & 65 & 76 & 78 & 94 \end{bmatrix}$$

The range of the pixels intensities now are from 0 to 255. We will change the range from -128 to 127.

Subtracting 128 from each pixel value yields pixel value from -128 to 127. After subtracting 128 from each of the pixel value, we got the following results.

$$\begin{bmatrix} -76 & -73 & -67 & -62 & -58 & -67 & -64 & -55 \\ -65 & -69 & -73 & -38 & -19 & -43 & -59 & -56 \\ -66 & -69 & -60 & -15 & 16 & -24 & -62 & -55 \\ -65 & -70 & -57 & -6 & 26 & -22 & -58 & -59 \\ -61 & -67 & -60 & -24 & -2 & -40 & -60 & -58 \\ -49 & -63 & -68 & -58 & -51 & -60 & -70 & -53 \\ -43 & -57 & -64 & -69 & -73 & -67 & -63 & -45 \\ -41 & -49 & -59 & -60 & -63 & -52 & -50 & -34 \end{bmatrix}$$

Now we will compute using this formula.

$$G_{u,v} = \alpha(u)\alpha(v)\sum_{x=0}^{7}\sum_{y=0}^{7} g_{x,y}\cos\left[\frac{\pi}{8}\left(x+\frac{1}{2}\right)u\right]\cos\left[\frac{\pi}{8}\left(y+\frac{1}{2}\right)v\right]$$

$$\alpha_p(n) = \begin{cases} \sqrt{\frac{1}{8}}, & \text{if } n = 0 \\ \sqrt{\frac{2}{8}}, & \text{otherwise} \end{cases}$$

The result comes from this is stored in let's say A(j,k) matrix.

There is a standard matrix that is used for computing JPEG compression, which is given by a matrix called as Luminance matrix.

This matrix is given below

$$Q_{j,k} = \begin{bmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \\ 72 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \end{bmatrix}$$

Applying the following formula

$$B_{j,k} = \text{round}\left(\frac{A_{j,k}}{Q_{j,k}}\right)$$

We got this result after applying.

$$B_{j,k} = \begin{bmatrix} -26 & -3 & -6 & 2 & 2 & -1 & 0 & 0 \\ 0 & -2 & -4 & 1 & 1 & 0 & 0 & 0 \\ -3 & 1 & 5 & -1 & -1 & 0 & 0 & 0 \\ -4 & 1 & 2 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Now we will perform the real trick which is done in JPEG compression which is ZIG-ZAG movement. The zig zag sequence for the above matrix is shown below. You have to perform zig zag until you find all zeroes ahead. Hence our image is now compressed.

$$B_{j,k} = \begin{bmatrix} -26 & -3 & -6 & 2 & 2 & -1 & 0 & 0 \\ 0 & -2 & -4 & 1 & 1 & 0 & 0 & 0 \\ -3 & 1 & 5 & -1 & -1 & 0 & 0 & 0 \\ -4 & 1 & 2 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Summarizing JPEG compression

The first step is to convert an image to Y'CbCr and just pick the Y' channel and break into 8 x 8 blocks. Then starting from the first block, map the range from -128 to 127. After that you have to find the discrete Fourier transform of the matrix. The result of this should be quantized. The last step is to apply encoding in the zig zag manner and do it till you find all zero.

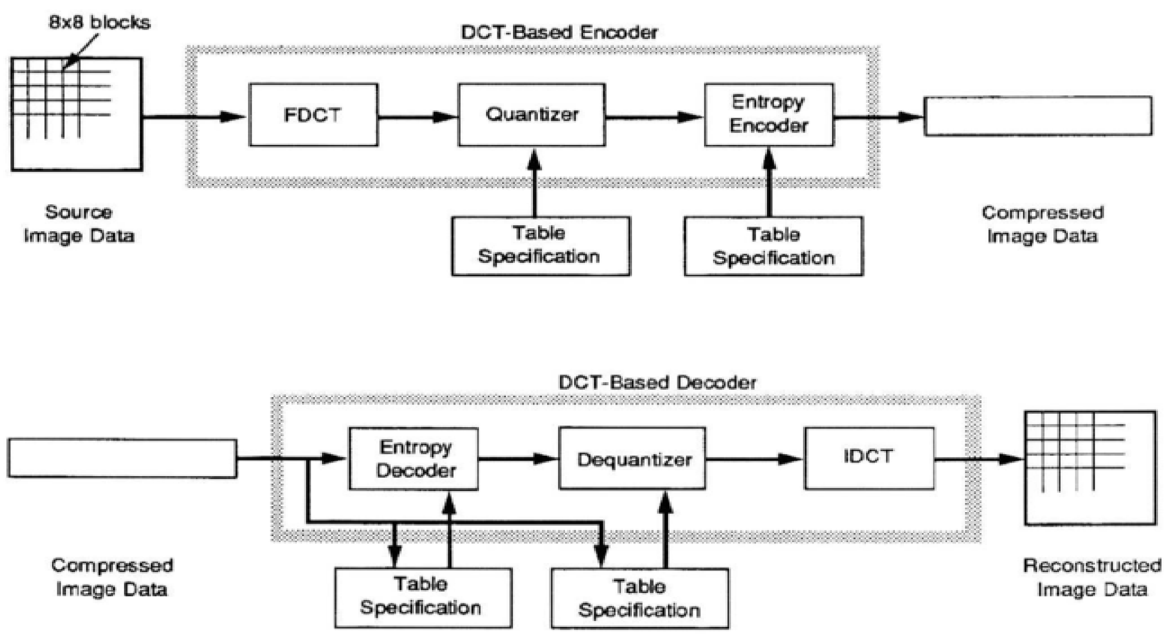
Save this one dimensional array and you are done.

IMAGE COMPRESSION

- **Image compression** is a method to reduce the redundancies in image representation in order to decrease data storage requirements (Gonzalez and Woods, 2013).
- It is a technique used to compress an image without visually reducing the quality of the image itself.
- Data vs. Information.
- The goal of these processes is to represent an image with the same quality level, but in a more solid form.
- The large storage requirement of multimedia data.

The video or image files consume large amount of data and it always required very high bandwidth networks in transmission as well as communication costs.

- Reduce the data storage and maintain the visual image quality (Gonzalez, Woods and Eddins, 2017).
- Increase the speed of transmission by using the repetition property of data.
- The goal of these processes is to represent an image with the same quality level, but in a more solid form.



There are two types of compression methods:

- Lossy image compression
- Lossless image compression

LOSSY COMPRESSION

Unlike the error-free approaches outlined in the previous section, lossy encoding is based on the concept of compromising the accuracy of the reconstructed image m exchange for increased compression. If the resulting distortion (which mayor may not be visually apparent) can be tolerated, the increase m compression can be significant. **In** fact, many lossy encoding techniques are capable of reproducing recognizable monochrome images from data that have been compressed by more than 100: 1 and images that are virtually indistinguishable from the originals at 10: 1 to 50: 1. Error-free encoding of monochrome images, however, seldom results in more than a 3: 1 reduction in data. As indicated in Section 8.2, the principal difference between these two approaches is the presence or absence of the quantizer block of FIg. 8.6.

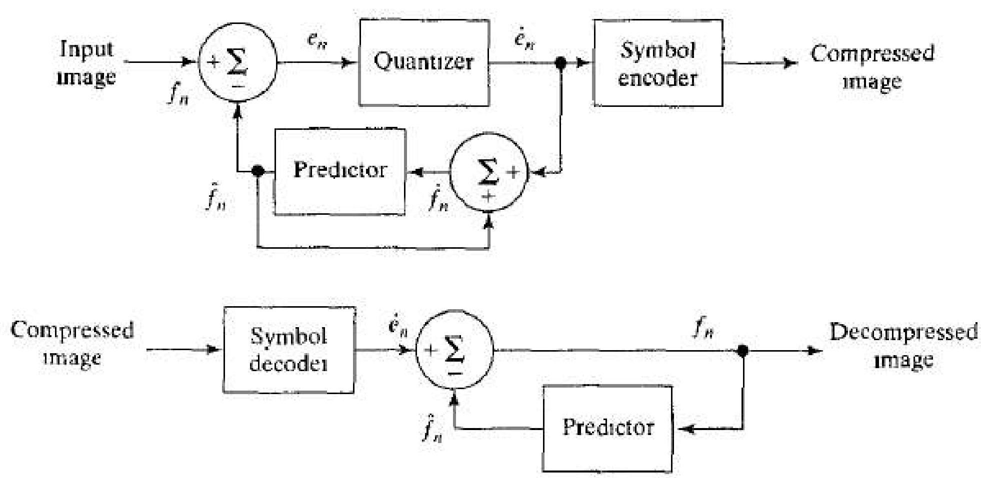
LOSSY PREDICTIVE CODING

- Lossy image compression methods are required to achieve high compression ratios for complex images.
- An image reconstructed Lossy compression can be performed in both spatial or transform domains.
- The process of quantization-dequantization introduces loss in the reconstructed image and is inherently responsible for the "lossy" nature of the compression scheme.
- The quantized transform coefficient is computed by

$$T_{pq} = Round\left(\frac{B_{pq}}{Q_{pq}}\right)$$

Where is the frequency image signals at coordinates (*i*,*j*) in the *k* block.

we add a quantizer to the model introduced in Section 8.4.4 and examine the resulting trade-off between reconstruction accuracy and compression performance. As Fig. 8.21 shows, the quantizer, which absorbs the nearest integer function of the error-free encoder, IS inserted between the symbol



a
b

FIGURE 8.21 A lossy predictive coding model: (a) encoder and (b) decoder.

encoder and the point at which the prediction error is formed. It maps the predictIOn error into a limited range of outputs, denoted *en,* which establish the amount of compression and distortion associated with lossy predictive coding. In order to accommodate the insertion of the quantization step, the error-free encoder of Fig. 8.19(a) must be altered so that the predictions generated by the encoder and decoder are equivalent. As Fig. 8.21(a) shows, this is accomplished by placing the lossy encoder's predictor within a feedback loop, where its input, denoted *f n'* is generated as a function of past predictions and the corresponding quantized errors. That is,

$$\dot{f}_n = \dot{e}_n + \hat{f}_n$$

where f*n* is as defined in Section 8 4.4. This closed loop configuration prevents error buildup at the decoder's output. Note from Fig. 8.21(b) that the output of the decoder also is given by Eq. (8.5-1)

LOSSLESS PREDICTIVE CODING

Lossless compression methods are needed in some digital imaging applications, such as: medical images, x-ray images, etc.

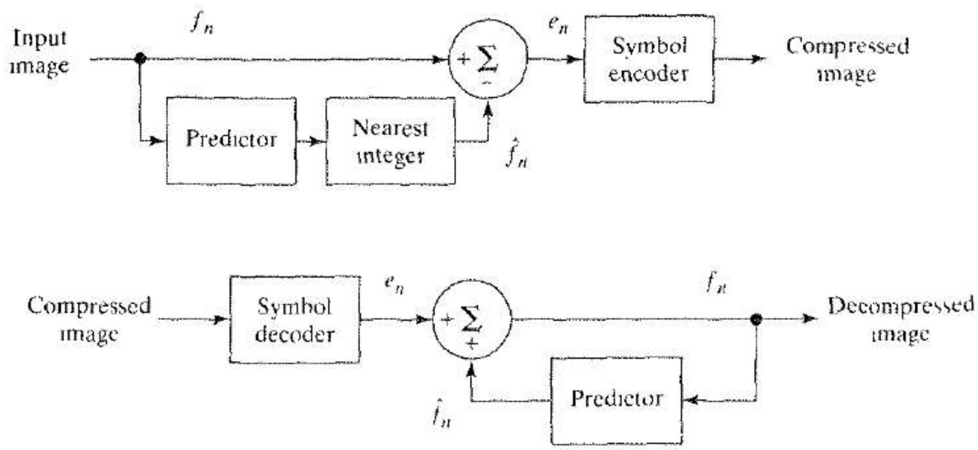
LOSSLESS PREDICTIVE CODING

Let us now turn to an error-free compression approach that does not require decomposition of an image into a collection of bit planes The approach, commonly referred to as *lossless predictive coding,* IS based on eliminating the interpixel redundancies of closely spaced pixels by extracting and coding *only* the new information in each pixel The *new information* of a pixel is defined as the difference between the actual and predicted value of that pixel
Figure 8.19 shows the basic components of a lossless predictive coding system. The system consists of an encoder and a decoder, each containing an identical *predictor.* As each successive pixel of the input image, denoted *f",* is introduced to the encoder, the predictor generates the anticipated value of that pixel based on some number of past inputs. The output of the predictor IS then rounded to the nearest integer, denoted *fn'* and used to form the difference' or *prediction error*

$$e_n = f_n - \hat{f}_n,$$

a
b

FIGURE 8.19 A lossless predictive coding model: (a) encoder, (b) decoder



which is coded using a variable-length code (by the symbol encoder) to generate the next element of the compressed data stream. The decoder of Fig. 8.19(b) reconstructs e_n from the received variable-length code words and performs the inverse operation

$$f_n = e_n + \hat{f}_n \tag{8.4-6}$$

Various local, global, and adaptive (see Section 8.5.1) methods can be used to generate $\hat{f}_n$. In most cases, however, the prediction is formed by a linear combination of m previous pixels. That is,

$$\hat{f}_n = \text{round}\left[\sum_{i=1}^{m} \alpha_i f_{n-i}\right] \tag{8.4-7}$$

where m is the order of the linear predictor, round is a function used to denote the rounding or nearest integer operation, and the α_i for $i = 1, 2, \ldots, m$ are prediction coefficients. In raster scan applications, the subscript n indexes the predictor out-

Lossless compression techniques: run-length coding, Huffman coding, lossless predictive coding etc.

Run-Length Encoding of AC Coefficients

The frequency image signals after quantization process consist of many zeros. Then, a special condition as known end-of-block (EOB) is applied to get an efficient in the entropy code.

A symbol of EOB indicates that the rest of the coefficients in the block are zero. Next, run-length encoding exploits the repeating frequency image signals as the symbols in the sequence a set of the AC coefficients.
The output of run-length encoding represents a sequence value with the consecutive repetition as symbols and the length of occurrence of the symbols

Huffman Coding

- Huffman coding is a famous method that uses variable-length codes (VLC) tables for compressing data (Jayaraman, Veerakumar, and Esakkirajan, 2017).
- Given a set of data symbols and their probabilities, the method creates a set of variable-length codeword with the shortest average length and assigns them to the symbols.

- Achieve minimal redundancy subject to the constraint that the source symbols be coded one at a time.
- Sorting symbols in descending probabilities is the key in the step of source reduction.
- The codeword assignment is not unique. Exchange the labeling of "0" and "1" at any node of binary codeword tree would produce another solution that equally works well.

Only works for a source with finite number of symbols (otherwise, it does not know where to start).

MOTION COMPENSATION

Motion compensation is an algorithmic technique used to predict a frame in a video, given the previous and/or future frames by accounting for motion of the camera and/or objects in the video. It is employed in the encoding of video data for video compression, for example in the generation of MPEG-2 files. Motion compensation describes a picture in terms of the transformation of a reference picture to the current picture. The reference picture may be previous in time or even from the future. When images can be accurately synthesised from previously transmitted/stored images, the compression efficiency can be improved.

Functionality

Motion compensation exploits the fact that, often, for many frames of a movie, the only difference between one frame and another is the result of either the camera moving or an object in the frame moving. In reference to a video file, this means much of the information that represents one frame will be the same as the information used in the next frame.

Using motion compensation, a video stream will contain some full (reference) frames; then the only information stored for the frames in between would be the information needed to transform the previous frame into the next frame.

Illustrated example

The following is a simplistic illustrated explanation of how motion compensation works. Two successive frames were captured from the movie Elephants Dream. As can be seen from the images, the bottom (motion compensated) difference between two frames contains significantly less detail than the prior images, and thus compresses much better than the rest. Thus the information that is required to encode compensated frame will be much smaller than with the difference frame. This also means that it is also possible to encode the information using difference image at a cost of less compression efficiency but by saving coding complexity without motion compensated coding; as a matter of fact that motion compensated coding (together with motion estimation, motion compensation) occupies more than 90% of encoding complexity.

MPEG

In MPEG, images are predicted from previous frames (P frames) or bidirectionally from previous and future frames (B frames). B frames are more complex because the image sequence must be transmitted and stored out of order so that the future frame is available to generate the B frames.[1]

After predicting frames using motion compensation, the coder finds the residual, which is then compressed and transmitted.

Global motion compensation

In global motion compensation, the motion model basically reflects camera motions such as:

- Dolly - moving the camera forward or backward
- Track - moving the camera left or right
- Boom - moving the camera up or down
- Pan - rotating the camera around its Y axis, moving the view left or right
- Tilt - rotating the camera around its X axis, moving the view up or down
- Roll - rotating the camera around the view axis

It works best for still scenes without moving objects.

Block motion compensation

In **block motion compensation** (BMC), the frames are partitioned in blocks of pixels (e.g. macroblocks of 16×16 pixels in MPEG). Each block is predicted from a block of equal size in the reference frame. The blocks are not transformed in any way apart from being shifted to the position of the predicted block. This shift is represented by a *motion vector*.

To exploit the redundancy between neighboring block vectors, (e.g. for a single moving object covered by multiple blocks) it is common to encode only the difference between the current and previous motion vector in the bit-stream. The result of this differencing process is mathematically equivalent to a global motion compensation capable of panning. Further down the encoding pipeline, an entropy coder will take advantage of the resulting statistical distribution of the motion vectors around the zero vector to reduce the output size.

Variable block-size motion compensation

Variable block-size motion compensation (VBSMC) is the use of BMC with the ability for the encoder to dynamically select the size of the blocks. When coding video, the use of larger blocks can reduce the number of bits needed to represent the motion vectors, while the use of smaller blocks can result in a smaller amount of prediction residual information to encode.

Overlapped block motion compensation

Overlapped block motion compensation (OBMC) is a good solution to these problems because it not only increases prediction accuracy but also avoids blocking artifacts. When using OBMC, blocks are typically twice as big in each dimension and overlap quadrant-wise with all 8 neighbouring blocks.

Quarter Pixel (QPel) and Half Pixel motion compensation

In motion compensation, quarter or half samples are actually interpolated sub-samples caused by fractional motion vectors. Based on the vectors and full-samples, the sub-samples can be calculated by using bicubic or bilinear 2-D filtering. See subclause 8.4.2.2 "Fractional sample interpolation process" of the H.264 standard.

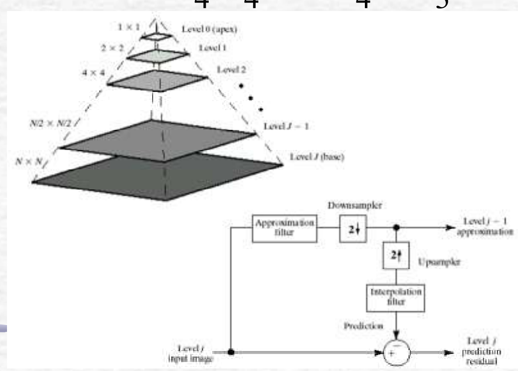
WAVELET BASED IMAGE COMPRESSION:

Unlike the Fourier transform, whose basis functions are sinusoids, wavelet transforms are based on small waves, called *wavelets,* of varying frequency *and limited duration.* This allows them to provide the equivalent of a musical score for an image, revealing not only what notes (or frequencies) to play but also when to play them.

Image Pyramids

Total number of elements in a P+1 level pyramid for P>0 is

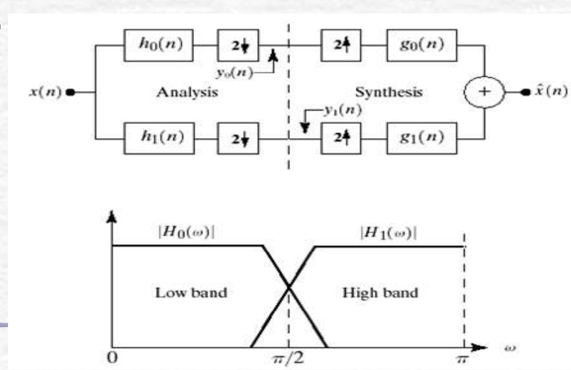
$$N^2(1+\frac{1}{4}+\frac{1}{4^2}+...+\frac{1}{4^P})\leq\frac{4}{3}N^2$$



a
b
FIGURE 7.2 (a) A pyramidal image structure and (b) system block diagram for creating it.

Subband Coding

- An image is decomposed into a set of band-limited components, called subbands, which can be reassembled to reconstruct the original image without error.



The Haar Transform

- Oldest and simplest known orthonormal wavelets.
- T=HFH where
  F: NXN image matrix,
  H: NxN transformation matrix.
- Haar basis functions $h_k(z)$ are defined over the continuous, closed interval [0,1] for k=0,1,..N-1 where $N=2^n$.

Haar Basis Functions

$$k = 2^p + q - 1 \quad \text{where } 0 \le p \le n-1,$$
$$q = 0 \text{ or } 1 \text{ for } p = 0, 1 \le q \le 2^p \text{ for } p \ne 0$$

$$h_0(z) = h_{00}(z) = \frac{1}{\sqrt{N}}, \quad z \in [0,1]$$

$$h_k(z) = h_{pq}(z) = \frac{1}{\sqrt{N}} \begin{cases} 2^{p/2} & (q-1)/2^p \le z < (q-0.5)/2^p \\ -2^{p/2} & (q-0.5)/2^p \le z < q/2^p \\ 0 & \text{otherwise}, z \in [0,1] \end{cases}$$

Multiresolution Expansions

- Multiresolution analysis (MRA)
- A scaling function is used to create a series of approximations of a function or image, each differing by a factor of 2.
- Additional functions, called wavelets, are used to encode the difference in information between adjacent approximations.

Series Expansions

- A signal f(x) can be expressed as a linear combination of expansion functions:

$$f(x) = \sum_k \alpha_k \varphi_k(x)$$

  - Case 1: orthonormal basis: $\langle \varphi_j(x), \varphi_k(x) \rangle = \delta_{jk}$
  - Case 2: orthogonal basis: $\langle \varphi_j(x), \varphi_k(x) \rangle = 0 \quad j \neq k$
  - Case 3: frame: $A\|f(x)\|^2 \leq \sum_k |\langle \varphi_k(x), f(x) \rangle|^2 \leq B\|f(x)\|^2$

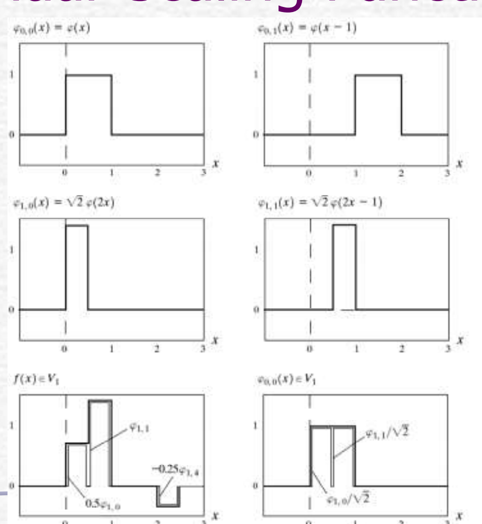
Scaling Functions

- Consider the set of expansion functions composed of integer translations and binary scaling of the real, square-integrable function, $\varphi(x)$,i.e.,

$$\varphi_{j,k}(x) = 2^{j/2}\varphi(2^j x - k)$$

- By choosing φ wisely, $\{\varphi_{j,k}(x)\}$ can be made to span $L^2(\mathbf{R})$

Haar Scaling Function



a b
c d
e f

FIGURE 7.9 Haar scaling functions in V_0 in V_1.

EXPANSION OF FUNCTIONS

MULTIRESOLUTION EXPANSIONS

The previous section introduced three well-known imaging techniques that played an important role in the development of a unique mathematical theory called multiresolution analysis (MRA).ln MRA, a *scaling function* is used to create a series of approximations of a function or image, each differing by a factor of 2 from its nearest neighboring approximations. Additional functions, called *wavelets,* are then used to encode the difference in information between adjacent approximations.

Unlike DCT s and DFT s, which use sinusoidal waves as basis functions, this new variety of transformations use small waves of varying frequency and of limited extent, known as wavelets as basis. The wavelets can be scaled and shifted to analyze the spatial frequency contents of an image at different resolutions and positions. A wavelet can therefore perform analysis of an image at multiple resolutions, making it an effective tool in *multi-resolution analysis* of images. Furthermore, wavelet analysis performs what is known as *space-frequency localization* so that at any specified location in space, one can obtain its details in terms of frequency. It is like placing a magnifying glass above a photograph to explore the details around a specific location. The magnifying glass can be moved up or down to vary the extent of magnification, that is, the level of details and it can be slowly panned over the other locations of the photograph to extract those details.

It is our common observation that the level of details within an image varies from location to location. Some locations contain significant details, where we require finer resolution for analysis and there are other locations, where a coarser resolution representation suffices. A multi-resolution representation of an image gives us a complete idea about the extent of the details existing at different locations from which we can choose our requirements of desired details. Multi-resolution representation facilitates efficient compression by exploiting the redundancies across the resolutions. Wavelet transforms is one of the popular, but not the only approach for multi-resolution image analysis. One can use any of the signal processing approaches to sub-band coding, such as Quadrature Mirror Filters (QMF) in multi-resolution analysis.

MRA Requirements

- Requirement 1: The scaling function is orthogonal to its integer translates.
- Requirement 2:The subspaces spanned by the scaling function at low scales are nested within those spanned at higher resolutions.
- Requirement 3:The only function that is common to all V_j is f(x)=0
- Requirement 4: Any function can be represented with arbitrary precision.

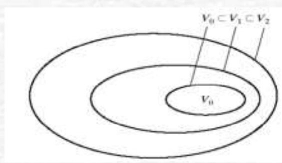
Scaling Functions
Now consider the set of expansion functions composed of integer translations and binary scalings of the real, square-integrable function *'P(x);* that is, the set *{φ.k (x)}* where

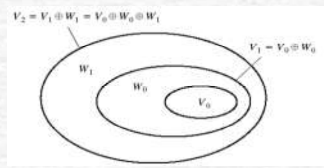
$$\varphi_{j,k}(x) = 2^{j/2}\varphi(2^j x - k) \tag{7.2-10}$$

for all $j, k \in \mathbf{Z}$ and $\varphi(x) \in L^2(\mathbf{R})$.‡ Here, k determines the position of $\varphi_{j,k}(x)$ along the x-axis, j determines $\varphi_{j,k}(x)$'s width—how broad or narrow it is along the x-axis—and $2^{j/2}$ controls its height or amplitude. Because the shape of $\varphi_{j,k}(x)$ changes with j, $\varphi(x)$ is called a *scaling function*. By choosing $\varphi(x)$ wisely, $\{\varphi_{j,k}(x)\}$ can be made to span $L^2(\mathbf{R})$, the set of all measurable, square-integrable functions.

WAVELET TRANSFORMS IN ONE DIMENSION

Wavelet Functions





Wavelet Functions

- A wavelet function, $\psi(x)$, together with its integer translates and binary scalings, spans the difference between any two adjacent scaling subspace, V_j and V_{j+1}.

$$\psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k)$$

$$\psi(x) = \sum_n h_\psi(n)\sqrt{2}\varphi(2x - n) \qquad h_\psi(n) = (-1)^n h_\varphi(1-n)$$

Haar Wavelet Functions

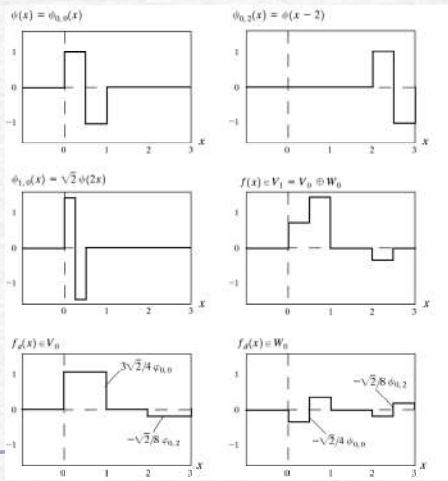


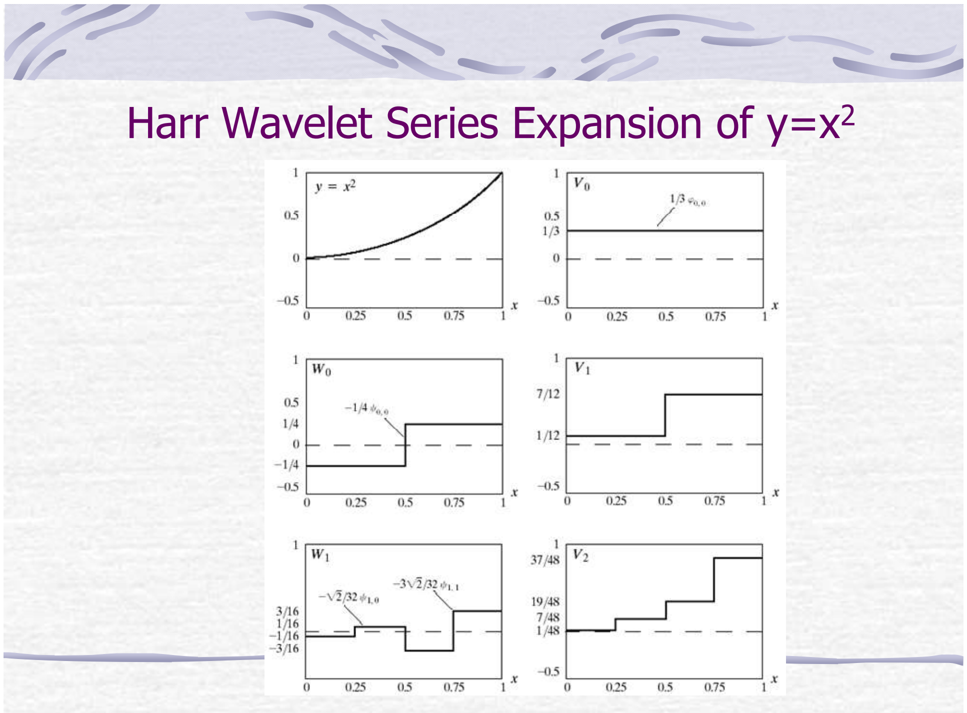
FIGURE 7.12 Haar wavelet functions in W_0 and W_1.

Wavelet Series Expansion

$$f(x) = \sum_k c_{j_0}(k)\varphi_{j_0,k}(x) + \sum_{j=j_0}^{\infty}\sum_k d_j(k)\psi_{j,k}(x)$$

$$c_{j_0}(k) = \int f(x)\varphi_{j_0,k}(x)dx$$

$$d_j(k) = \int f(x)\psi_{j,k}(x)dx$$



We can now formally define several closely related wavelet transformations: the generalized *wavelet series expansion,* the *discrete wavelet transform,* and the *continuous wavelet transform.* Their counterparts in the Fourier domain are the Fourier series expansion, the discrete Fourier transform, and the integral Fourier transform, respectively. **In** Section 7.4, we will define a computationally efficient implementation of the discrete wavelet transform called the *fast wavelet transform.*

THE WAVELET SERIES EXPANSIONS

We begin by defining the *wavelet series expansion* of function $f(x) \in L^2(\mathbf{R})$ relative to wavelet $\psi(x)$ and scaling function $\varphi(x)$. In accordance with Eq. (7.2-27), we can write

$$f(x) = \sum_k c_{j_0}(k)\varphi_{j_0,k}(x) + \sum_{j=j_0}^{\infty} \sum_k d_j(k)\psi_{j,k}(x) \qquad (7.3\text{-}1)$$

where j_0 is an arbitrary starting scale and the $c_{j_0}(k)$'s and $d_j(k)$'s are relabeled α_k's from Eqs. (7.2-12) and (7.2-21), respectively. The $c_{j_0}(k)$'s are normally called the *approximation* or *scaling coefficients*; the $d_j(k)$'s are referred to as the *detail* or *wavelet coefficients*. This is because the first sum in Eq. (7.3-1) uses scaling functions to provide an approximation of $f(x)$ at scale j_0 [unless $f(x) \in V_{j_0}$ and it is exact]. For each higher scale $j \geq j_0$ in the second sum, a finer resolution function—a sum of wavelets—is added to the approximation to provide increasing detail. If the expansion functions form an orthonormal basis or tight frame, which is often the case, the expansion coefficients are calculated—based on Eqs. (7.2-5) and (7.2-9)—as

$$c_{j_0}(k) = \langle f(x), \varphi_{j_0,k}(x)\rangle = \int f(x)\varphi_{j_0,k}(x)\,dx \qquad (7.3\text{-}2)$$

and

$$d_j(k) = \langle f(x), \psi_{j,k}(x)\rangle = \int f(x)\psi_{j,k}(x)\,dx. \qquad (7.3\text{-}3)$$

If the expansion functions are part of a biorthogonal basis, the φ and ψ terms in these equations must be replaced by their dual functions, $\tilde{\varphi}$ and $\tilde{\psi}$, respectively.

SERIES EXPANSIONS

A signal or *functionf(x)* can often be better analyzed as a linear combination of expansion functions

$$f(x) = \sum_k \alpha_k \varphi_k(x)$$

where k is an integer index of the finite or infinite sum, the α_k are real-valued *expansion coefficients*, and the $\varphi_k(x)$ are real-valued *expansion functions*. If the expansion is unique—that is, there is only one set of α_k for any given $f(x)$—the $\varphi_k(x)$ are called *basis functions*, and the *expansion set*, $\{\varphi_k(x)\}$, is called a *basis* for the class of functions that can be so expressed. The expressible functions form a *function space* that is referred to as the *closed span* of the expansion set, denoted

$$V = \overline{\underset{k}{\mathrm{Span}}\{\varphi_k(x)\}}. \tag{7.2-2}$$

To say that $f(x) \in V$ means that $f(x)$ is in the closed span of $\{\varphi_k(x)\}$ and can be written in the form of Eq. (7.2-1).

THE DISCRETE WAVELET TRANSFORM

Like the Fourier series expansion, the wavelet series expansion of the previous section maps a function of a continuous variable into a sequence of coefficients
If the function being expanded is a sequence of numbers. like samples of a continuous function *(x),* the resulting coefficients are called the *discrete wareler*

transform (DWT) of $f(x)$. For this case, the series expansion defined in Eqs (7.3-1) through (7.3-3) becomes the DWT transform pair

$$W_\varphi(j_0, k) = \frac{1}{\sqrt{M}} \sum_x f(x)\varphi_{j_0,k}(x) \tag{7.3-5}$$

$$W_\psi(j, k) = \frac{1}{\sqrt{M}} \sum_x f(x)\psi_{j,k}(x) \tag{7.3-6}$$

for $j \geq j_0$ and

$$f(x) = \frac{1}{\sqrt{M}} \sum_k W_\varphi(j_0, k)\varphi_{j_0,k}(x) + \frac{1}{\sqrt{M}} \sum_{j=j_0}^{\infty} \sum_k W_\psi(j, k)\psi_{j,k}(x). \tag{7.3-7}$$

Here, $f(x)$, $\varphi_{j_0,k}(x)$, and $\psi_{j,k}(x)$ are functions of the discrete variable $x = 0, 1, 2, \ldots, M - 1$. For example, $f(x) = f(x_0 + x\Delta x)$ for some x_0, Δx, and $x = 0, 1, 2, \ldots, M - 1$. Normally, we let $j_0 = 0$ and select M to be a power of 2 (i.e., $M = 2^J$) so that the summations are performed over $x = 0, 1, 2, \ldots, M - 1$, $j = 0, 1, 2, \ldots, J - 1$, and $k = 0, 1, 2, \ldots, 2^j - 1$. For Haar wavelets, the discretized scaling and wavelet functions employed in the transform (i.e., the basis functions) correspond to the rows of the $M \times M$ Haar transformation matrix of Section 7.1.3. The transform itself is composed of M coefficients, the minimum scale is 0, and the maximum scale is $J - 1$. For reasons noted in Section 7.3.1 and illustrated in Example 7.6, the coefficients defined in Eqs. (7.3-5) and (7.3-6) are usually called *approximation* and *detail coefficients*, respectively.

Discrete Wavelet Transform

$$W_\varphi(j_0,k) = \frac{1}{\sqrt{M}} \sum_x f(x)\varphi_{j_0,k}(x)$$

$$W_\psi(j,k) = \frac{1}{\sqrt{M}} \sum_x f(x)\psi_{j,k}(x)$$

$$f(x) = \frac{1}{\sqrt{M}} \sum_k W_\varphi(j_0,k)\varphi_{j_0,k}(x) + \frac{1}{\sqrt{M}} \sum_{j=j_0}^{\infty} \sum_k W_\psi(j,k)\psi_{j,k}(x)$$

THE CONTINUOUS WAVELET TRANSFORM

The natural extension of the discrete wavelet transform is the *continuous wavelet transform* (CWT), which transforms a continuous function into a highly redundant function of two continuous variables-translation and scale. The resulting transform is easy to interpret and valuable for time-frequency analysis. Although our interest is in discrete images, it is covered here for completeness. The continuous wavelet transform of a continuous, square-integrable function, f(x,y), relative to a real-valued wavelet, $\psi(x)$. is

$$W_{\psi}(s, \tau) = \int_{-\infty}^{\infty} f(x)\psi_{s,\tau}(x)\, dx \tag{7.3-8}$$

where

$$\psi_{s,\tau}(x) = \frac{1}{\sqrt{s}}\psi\left(\frac{x-\tau}{s}\right) \tag{7.3-9}$$

and s and τ are called *scale* and *translation* parameters, respectively. Given $W_{\psi}(s, \tau)$, $f(x)$ can be obtained using the *inverse continuous wavelet transform*

$$f(x) = \frac{1}{C_{\psi}} \int_{0}^{\infty} \int_{-\infty}^{\infty} W_{\psi}(s, \tau) \frac{\psi_{s,\tau}(x)}{s^2}\, d\tau\, ds \tag{7.3-10}$$

where

$$C_{\psi} = \int_{-\infty}^{\infty} \frac{|\Psi(u)|^2}{|u|}\, du \tag{7.3-11}$$

and $\Psi(u)$ is the Fourier transform of $\psi(x)$. Equations (7.3-8) through (7.3-11) define a reversible transformation as long as the so-called *admissibility criterion*, $C_{\psi} < \infty$, is satisfied (Grossman and Morlet [1984]). In most cases, this simply means that $\Psi(0) = 0$ and $\Psi(u) \rightarrow 0$ as $u \rightarrow \infty$ fast enough to make $C_{\psi} < \infty$.

The preceding equations are reminiscent of their discrete counterparts—Eqs. (7.2-19), (7.3-1), (7.3-3), (7.3-6), and (7.3-7). The following similarities should be noted:

The Continuous Wavelet Transform

$$W_\psi(s,\tau) = \int_{-\infty}^{\infty} f(x)\psi_{s,\tau}(x)dx \qquad \psi_{s,\tau}(x) = \frac{1}{\sqrt{s}}\psi\left(\frac{x-\tau}{s}\right)$$

$$f(x) = \frac{1}{C_\psi}\int_0^\infty\int_{-\infty}^{\infty} W_\psi(s,\tau)\frac{\psi_{s,\tau}(x)}{s^2}d\tau ds \qquad C_\psi = \int_{-\infty}^{\infty}\frac{|\Psi(u)|^2}{|u|}du$$

<u>THE FAST WAVELET TRANSFORM</u>

The *fast wavelet transform* (FWT) is a computationally efficient implementation of the discrete wavelet transform (DWT) that exploits a surprising but fortunate relationship between the coefficients of the DWT at adjacent scales.

Consider again the multiresolution refinement equation

$$\varphi(x) = \sum_n h_\varphi(n)\sqrt{2}\varphi(2x - n). \tag{7.4-1}$$

Scaling x by 2^j, translating it by k, and letting $m = 2k + n$ gives

$$\begin{aligned}\varphi(2^j x - k) &= \sum_n h_\varphi(n)\sqrt{2}\varphi(2(2^j x - k) - n)\\ &= \sum_m h_\varphi(m - 2k)\sqrt{2}\varphi(2^{j+1}x - m).\end{aligned} \tag{7.4-2}$$

Note that scaling vector h_φ can be thought of as the "weights" used to expand $\varphi(2^j x - k)$ as a sum of scale $j + 1$ scaling functions. A similar sequence of operations—beginning with Eq. (7.2-28)—provides an analogous result for $\psi(2^j x - k)$. That is,

$$\psi(2^j x - k) = \sum_m h_\psi(m - 2k)\sqrt{2}\varphi(2^{j+1}x - m) \tag{7.4-3}$$

where scaling vector $h_\varphi(n)$ in Eq. (7.4-2) is replaced by wavelet vector $h_\psi(n)$ in Eq. (7.4-3).

Wavelet Transforms in Two Dimensions

The one-dimensional transforms of the previous sections are easily extended to two-dimensional functions like images. In two dimensions, a two-dimensional scaling function, $\varphi(x, y)$, and three two-dimensional wavelets, $\psi^H(x, y)$, $\psi^V(x, y)$, and $\psi^D(x, y)$, are required. Each is the product of a one-dimensional scaling function φ and corresponding wavelet ψ. Excluding products that produce one-dimensional results, like $\varphi(x)\psi(x)$, the four remaining products produce the *separable* scaling function

$$\varphi(x, y) = \varphi(x)\varphi(y) \tag{7.5-1}$$

and separable, "directionally sensitive" wavelets

$$\psi^H(x, y) = \psi(x)\varphi(y) \tag{7.5-2}$$

$$\psi^V(x, y) = \varphi(x)\psi(y) \tag{7.5-3}$$

$$\psi^D(x, y) = \psi(x)\psi(y). \tag{7.5-4}$$

These wavelets measure functional variations—intensity or gray-level variations for images—along different directions: ψ^H measures variations along columns (for example, horizontal edges), ψ^V responds to variations along rows (like vertical edges), and ψ^D corresponds to variations along diagonals. The directional sensitivity is a natural consequence of the separability imposed by

Eqs. (7.5-2) to (7.5-4); it does not increase the computational complexity of the two-dimensional transform discussed in this section.

Given separable two-dimensional scaling and wavelet functions, extension of the one-dimensional DWT to two dimensions is straightforward. We first define the scaled and translated basis functions:

$$\varphi_{j,m,n}(x, y) = 2^{j/2}\varphi(2^j x - m, 2^j y - n), \tag{7.5-5}$$

$$\psi^i_{j,m,n}(x, y) = 2^{j/2}\psi^i(2^j x - m, 2^j y - n), \qquad i = \{H, V, D\} \tag{7.5-6}$$

where index i identifies the directional wavelets in Eqs. (7.5-2) to (7.5-4). Rather than an exponent, i is a superscript that assumes the values H, V, and D. The discrete wavelet transform of function $f(x, y)$ of size $M \times N$ is then

$$W_\varphi(j_0, m, n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x, y)\varphi_{j_0,m,n}(x, y) \tag{7.5-7}$$

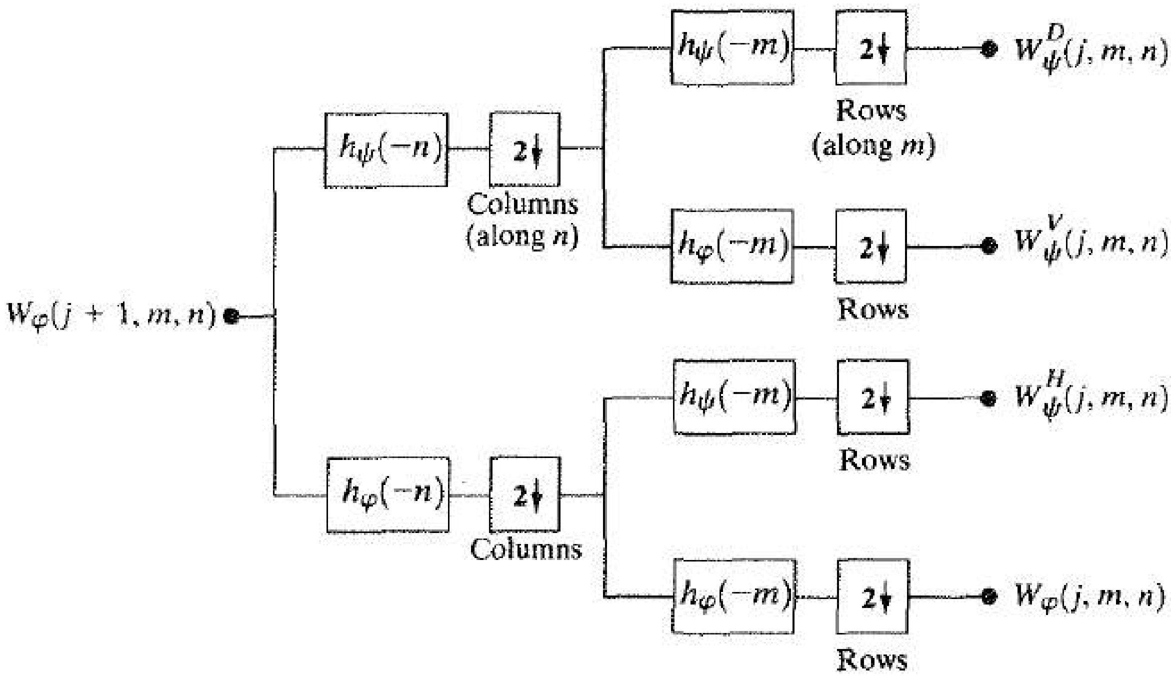
$$W^i_\psi(j, m, n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x, y)\psi^i_{j,m,n}(x, y) \qquad i = \{H, V, D\}. \tag{7.5-8}$$

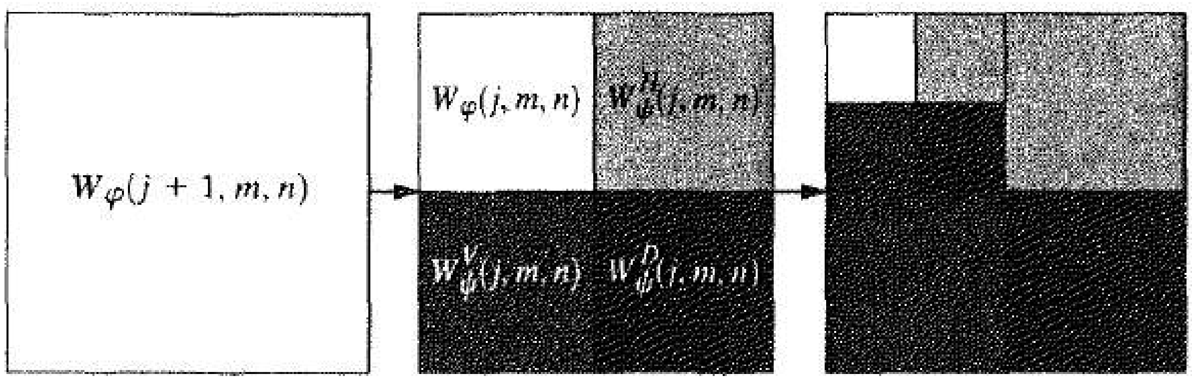
As in the one-dimensional case, j_0 is an arbitrary starting scale and the $W_\varphi(j_0, m, n)$ coefficients define an approximation of $f(x, y)$ at scale j_0. The $W^i_\psi(j, m, n)$ coefficients add horizontal, vertical, and diagonal details for scales $j \geq j_0$. We normally let $j_0 = 0$ and select $N = M = 2^J$ so that $j = 0, 1, 2, \ldots, J - 1$ and $m, n = 0, 1, 2, \ldots, 2^j - 1$. Given the W_φ and W^i_ψ of Eqs. (7.5-7) and (7.5-8), $f(x, y)$ is obtained via the inverse discrete wavelet transform

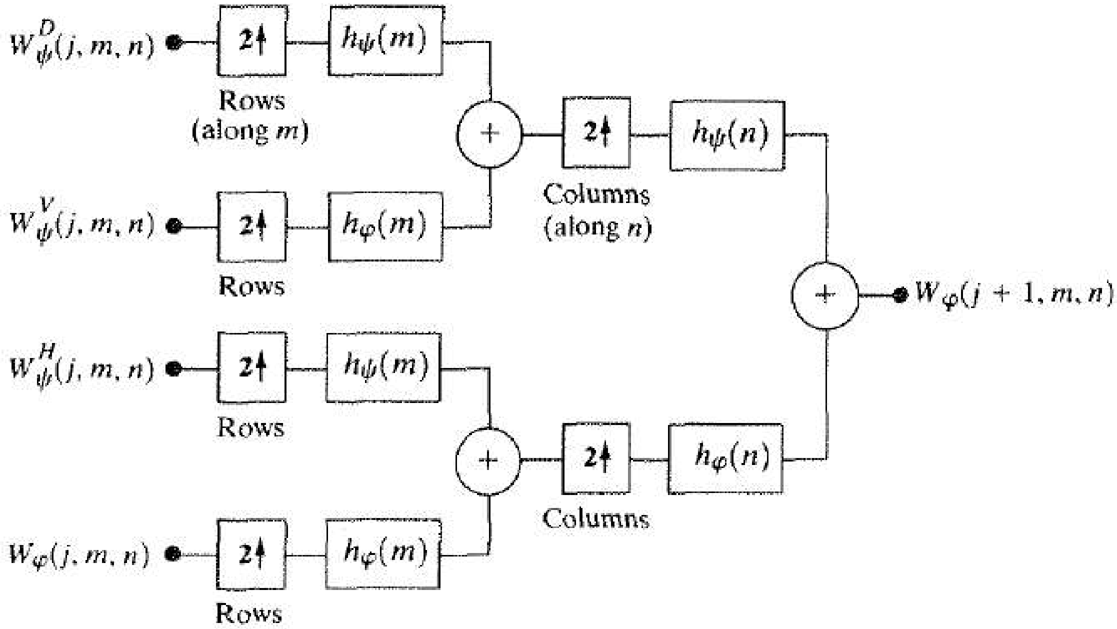
$$f(x, y) = \frac{1}{\sqrt{MN}} \sum_m \sum_n W_\varphi(j_0, m, n)\varphi_{j_0,m,n}(x, y) \tag{7.5-9}$$
$$+ \frac{1}{\sqrt{MN}} \sum_{i=H,V,D} \sum_{j=j_0}^{\infty} \sum_m \sum_n W^i_\psi(j, m, n)\psi^i_{j,m,n}(x, y).$$

Like the one-dimensional discrete wavelet transform, the two-dimensional DWT can be implemented using digital filters and downsamplers. With separable two-dimensional scaling and wavelet functions, we simply take the one-dimensional FWT of the rows of $f(x, y)$, followed by the one-dimensional FWT of the resulting columns. Figure 7.22(a) shows the process in block diagram form. Note that, like its one-dimensional counterpart in Fig. 7.15, the two-dimensional FWT "filters" the scale $j + 1$ approximation coefficients to construct the scale j approximation and detail coefficients. In the two-dimensional case, however, we get three sets of detail coefficients—the horizontal, vertical, and diagonal details.

The single-scale filter bank of Fig. 7.22(a) can be "iterated" (by tying the approximation output to the input of another filter bank) to produce a P scale transform in which scale $j = J - 1, J - 2, \ldots, J - P$. As in the one-dimensional case, image $f(x, y)$ is used as the $W_\varphi(J, m, n)$ input. Convolving its rows with $h_\varphi(-n)$ and $h_\psi(-n)$ and downsampling its columns, we get two subimages whose horizontal resolutions are reduced by a factor of 2. The highpass or detail component characterizes the image's high-frequency information with vertical ori-


$h_\psi(-m)$
$2\downarrow$
$W_\psi^D(j, m, n)$
Rows
(along m)
$h_\psi(-n)$
$2\downarrow$
Columns
(along n)
$h_\varphi(-m)$
$2\downarrow$
$W_\psi^V(j, m, n)$
Rows
$W_\varphi(j + 1, m, n)$
$h_\psi(-m)$
$2\downarrow$
$W_\psi^H(j, m, n)$
Rows
$h_\varphi(-n)$
$2\downarrow$
Columns
$h_\varphi(-m)$
$2\downarrow$
$W_\varphi(j, m, n)$
Rows


$W_\varphi(j + 1, m, n)$
$W_\varphi(j, m, n)$
$W_\psi^H(j, m, n)$
$W_\psi^V(j, m, n)$
$W_\psi^D(j, m, n)$



a
b
c

FIGURE 7.22 The two-dimensional fast wavelet transform: (a) the analysis filter bank; (b) the resulting decomposition; and (c) the synthesis filter bank.

These subimages, which are shown in the middle of Fig. 7.22(b), are the inner products of $f(x, y)$ and the two-dimensional scaling and wavelet functions in Eqs. (7.5-1) through (7.5-4), followed by downsampling by two in each dimension. Two iterations of the filtering process produces the two-scale decomposition at the far right of Fig. 7.22(b).

Figure 7.22(c) shows the synthesis filter bank that reverses the process described above. As would be expected, the reconstruction algorithm is similar to the one-dimensional case. At each iteration, four scale j approximation and detail subimages are upsampled and convolved with two one-dimensional filters—one operating on the subimages' 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.

We conclude the section with two examples that demonstrate the usefulness of wavelets in image processing. As in the Fourier domain, the basic approach is-to

1. Compute the two-dimensional wavelet transform of an image.

2. Alter the transform.
3. Compute the inverse transform.

Because the DWT's scaling and wavelet vectors are used as lowpass and highpass filters. most Fourier based filtering techniques have an equivalent "wavelet domain" counterpart.

The general wavelet-based procedure for *denoising* the image (i.e., suppressing the noise part) is as follows:

1. Choose a wavelet (e.g., Haar, symlet, ...) and number of levels or scales, *P,* for the decomposition. Then compute the FWT of the noisy image.

2. Threshold the detail coefficients-That is, select and apply a threshold to the detail coefficients from scales *J* - 1 to *J* - *P.* This can be accomplished by *hard thresholding,* which means setting to zero the elements whose absolute values are lower than the threshold, or by *soft thresholding,* which involves first setting to zero the elements whose absolute values are lower than the threshold and then scaling the nonzero coefficients toward zero. Soft thresholding eliminates the discontinuity (at the threshold) that is inherent in hard thresholding.

3. Perform a wavelet reconstruction based on the original approximation coefficients at level *J* - *P* and the modified detail coefficients for level.

JPEG 2000

Although not yet formally adopted, JPEG 2000 extends the initial JPEG standard to provide increased flexibility in both the compression of continuous tone still images and access to the compressed data. For example, portions of a JPEG 2000 compressed image can be extracted for retransmission, storage, display, and/or editing. The standard is based on the wavelet coding techniques of Section 8.5.3. Coefficient quantization is adapted to individual scales and subbands and the quantized coefficients are arithmetically coded on a bit-plane basis (see Section *804).* Using the notation of the standard. an image is encoded as follows (ISO/IEC [2000]).

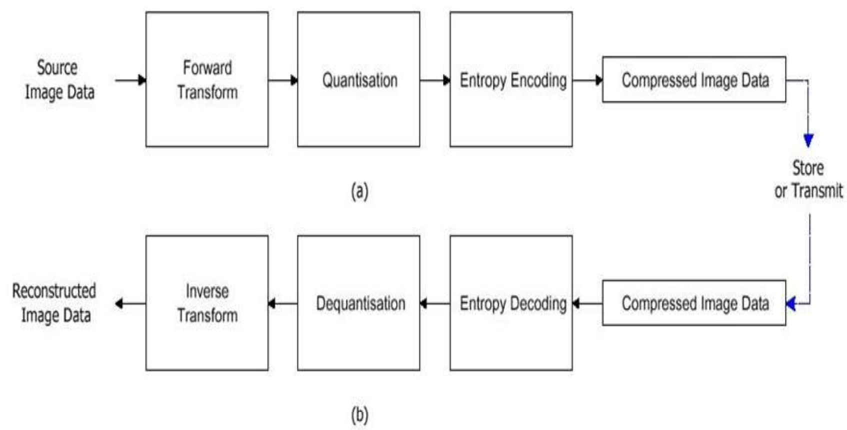
The first step of the encoding process is to DC level shift the samples of the Ssiz-bit unsigned image to be coded by subtracting *28"z-I.If* the image has more than one *component-like* the red, green, and blue planes of a color imageeach component is individually shifted. If there are exactly three components, they may be optionally decorrelated using a reversible or nonreversible linear combination of the components. The *irreversible component transform* of the standard, for example, is

$$Y_0(x, y) = 0.299I_0(x, y) + 0.587I_1(x, y) + 0.114I_2(x, y)$$
$$Y_1(x, y) = -0.16875I_0(x, y) - 0.33126I_1(x, y) + 0.5I_2(x, y) \quad (8.6\text{-}1)$$
$$Y_2(x, y) = 0.5I_0(x, y) - 0.41869I_1(x, y) - 0.08131I_2(x, y)$$

where I_0,.I_1 and I_2 are the level-shifted input components and Yo. YI , and Y2 are the corresponding decorrelated components. If the input components are the red, green, and blue planes of a color image, Eq. (8.6-1) approximates the *R'C'B'* to *Y'ChC,* color video transform (Poynton [1996]).' The goal of the transformation is to improve compression efficiency; transformed components YI and Y2 are difference images whose histograms are highly peaked around zero.

After the image has been level shifted and optionally decorrelated. its components are optionally divided into *tiles.* Tiles are rectangular arrays of pixels that contain the same relative portion of all components. Thus, the tiling process creates *tile components* that can be extracted and reconstructed independently, providing a simple mechanism for accessing and/or manipulating a limited re' gion of a coded image.
The one-dimensional discrete wavelet transform of the rows and columns of each tile component is then computed.

JPEG2000 (encoder-decoder scheme)



Block diagrams of the JPEG2000 (a) encoder and (b) decoder.

6

JPEG2000 - Overview

- The source image is decomposed into components (up to 256).
- The image components are (optionally) decomposed into rectangular tiles. The tile-component is the basic unit of the original or reconstructed image.
- A wavelet transform is applied on each tile. The tile is decomposed into different resolution levels.
- The decomposition levels are made up of subbands of coefficients that describe the frequency characteristics of local areas of the tile components, rather than across the entire image component.
- The sub-bands of coefficients are quantized and collected into rectangular arrays of code blocks.

7

Characteristics:

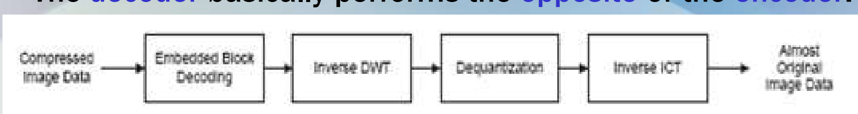
So, what is new in JPEG2000,
comparing to previous encoding protocols???

1. Compress once - decompress many ways
2. Region-Of-Interest encoding
3. Progression
4. Error resilience

32

Decoding

- The decoder basically performs the opposite of the encoder:



- The code-stream is received by the decoder according to the progression order stated in the header. The coefficients in the packets are then decoded and dequantized, and the reverse-ICT is performed:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1.0 & 0.0 & 1.4021 \\ 1.0 & -0.3441 & -0.7142 \\ 1.0 & 1.7718 & 0.0 \end{bmatrix} \begin{bmatrix} Y \\ C_r \\ C_b \end{bmatrix}$$

- In the case of irreversible compression, the decompression results in loss of data. The resulting image is not exactly like the original.

31

DIGITAL IMAGE WATERMARKING

- Watermark--an invisible signature embedded inside an image to show authenticity or proof of ownership
- Discourage unauthorized copying and distribution of images over the internet
- Ensure a digital picture has not been altered
- Software can be used to search for a specific watermark

Types of Watermarking:
The digital watermarking is of four types and they are as follows:

- Visible
- Invisible
- Fragile
- Public

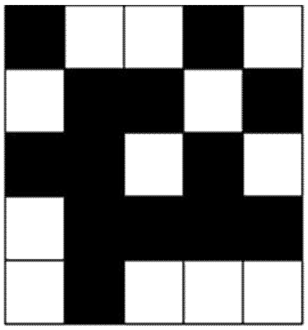
1. **Visible watermarking:** the visible watermarking have the visible message or the logo of the company which represents the ownership of the message, removing of the logo and meddling of the logo breaks the law of copyright.
2. **Invisible watermarking:** the invisible watermarking cannot be seen on the original images and the picture looks like an original image though it has the watermark. The

invisible watermark can be extracted with the detection procedure or with any watermark extraction method.

3. **Fragile watermark:** the fragile watermarks can be demolished by the data manipulation and these are also called as tamper proof watermark. The fragile watermark has the ability to detect the modifications in the signal and also recognizes the place where the modifications have occurred and also the signal before the change.
4. **Public watermark:** The public watermark does not have the protection and these can be read by everyone by availing the unique algorithm.

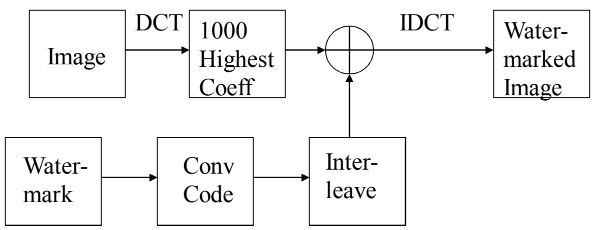
Watermark Properties

- Watermark should appear random, noise-like sequence
- Appear Undetectable
- Good Correlation Properties
  High correlation with signals similar to watermark
  Low correlation with other watermarks or random noise
- Common sequences
  A) Normal distribution
  B) m-sequences



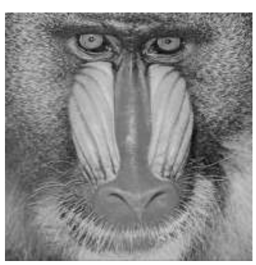
$$W=\begin{bmatrix}1&0&0&1&0\\0&1&1&0&1\\1&1&0&1&0\\0&1&1&1&1\\0&1&0&0&0\end{bmatrix}$$

Setup-Watermark Embedding

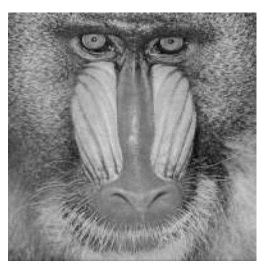


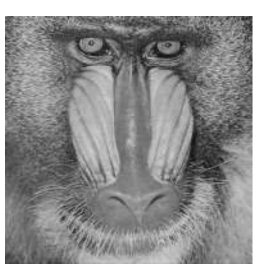
- •DC Component Excluded for 1000 Highest Coefficients
- •Interleaving prevents burst errors
- •Watermarked Image Similar to original image
- •Without coding, ignore Conv Code and Interleave block

Original Image



Watermarked Image, No Coding

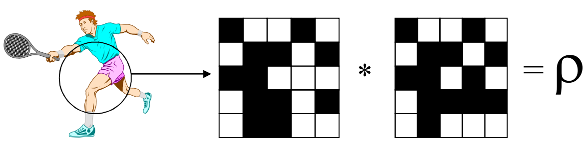




Watermarked Image with Coding

- •512x512 “Mandrill” Image
- •See Handout
- •Both watermarks imperceptible
- •Alterations to original image difficult to notice

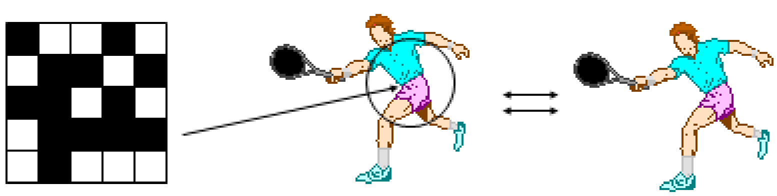
Watermark Detection



Suspected Image | Extracted Watermark | Original Watermark | Correlation

- Watermark Extracted from Suspected Image
- Compute correlation of Extracted and Original Watermark
- Threshold correlation to determine watermark existence

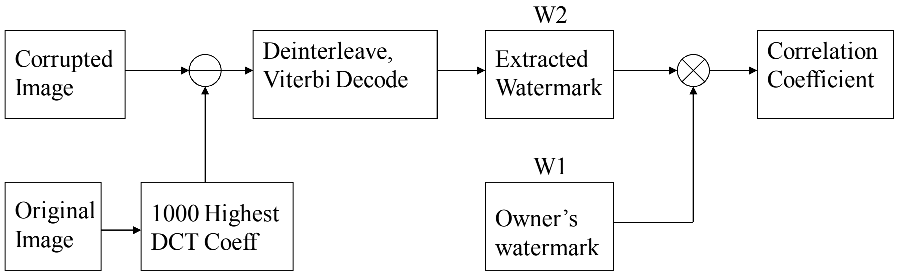
Watermark Embedding



Watermark | Original Image | Watermarked image

- Watermark placed into information content of Original Image to create Watermarked Image
- Image Content
  Spatial Domain (Least Significant Bit)
  FFT - Magnitude and Phase
  Wavelet Transforms
  DCT Coefficients

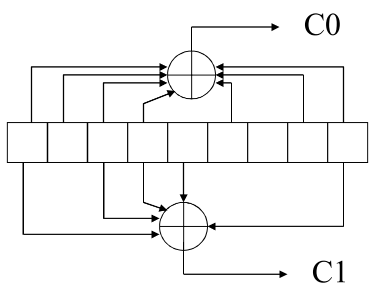
Watermark Detection



- For no coding, deinterleave and decode block ignored
- ρ=E[W1*W2]/{ E[W1^2]E[W2^2]}
- If W1=W2 then ρ=1
- if W1 and W2 are independent, then ρ=0 if E[W1]=0
- Corruptions are additive noise, JPEG Compression Image scaling, and UnZign

Convolutional Codes

Input=[...10110101011100000000]
G0 = [1 1 1 1 0 1 0 1 1]
G1 = [1 0 1 1 1 0 0 0 1]



- Output C0 = conv(G0,Input); Output C1=conv(G1,Input)
- Convolutional code implemented using linear shift registers
- Adds redundancy for error-correction
- Encoding/Decoding well researched
- Good coding performance, very popular

Possible Questions

1. Explain about FAX compression.
2. Explain about symbol based coding.
3. Describe about lossy and lossless predictive coding.
4. Discuss about wavelet based image compression.
5. Illustrate about the concept of scaling function.
6. Ealobarate wavelet series expansion.
7. Explain Discrete wavelet transform.
8. Describe about fast wavelet transform.
9. Explain about 2D wavelet transform.
10. Explain about image watermaking.

KARPAGAM ACADEMY OF HIGHER EDUCATION
CLASS: III BSC IT COURSE NAME: DIGITAL IMAGE PROCESSING
COURSE CODE: 16ITU504A BATCH-2016-2019

UNIT IV

SNO	QUESTIONS	CHOICE 1	CHOICE 2	CHOICE 3	CHOICE 4	ANSWER
1	__________ is an encoding format used for fax transmission.	Fax Group	ITU-T	JPEG	TIFF	Fax Group
2	JBIG stands for ___________	Joint Bi-level Image Experts Group	Joint Bit Image Experts Group	Joint Bi-level Image Encoding Group	Joint Bit Image Encoding Group	Joint Bi-level Image Experts Group
3	___________ is an effective technique for reducing an image's interpixel redundancies is to process the image's bit planes individually	Huffman coding	bit-plane coding	Arithmetic coding	Golomb coding	bit-plane coding
4	The overall process of truncating, quantizing, and coding the coefficients of a transformed subimage is commonly called __________	bit-plane coding	zonal coding	threshold coding	bit allocation	bit allocation
5	_____________ is based on the information th	Zonal coding	bit-plane coding	Arithmetic coding	threshold coding	Zonal coding
6	JPEG stands for ______________	Joint photographic experts group	Joint photographic encoding group	Joint phase experts group	Joint photo encode group	Joint photographic experts group
7	is a method to reduce the redundancies in image representation in order to decrease data storage requirements.	Image enhancement	Image compression	Image segmentation	Image restoration	Image compression
8	_________ consume large amount of data and requires very high bandwidth networks in transmission	video or image files	email	file attachment	text	video or image files
9	__________supports one-dimensional image compression of black and white images, on a standard fax machine.	Fax Group 5	Fax Group 4	Fax Group 3	Fax Group 6	Fax Group 3
10	_______________ supports two-dimensional image compression, compressing the line width as well as the line length.	Fax Group 5	Fax Group 4	Fax Group 3	Fax Group 6	Fax Group 4

11	____________ is an algorithmic technique used to predict a frame in a video, given the previous and/or future frames.	Motion compensation	Huffman coding	bit-plane coding	Arithmetic coding	Motion compensation
12	___________ transforms a continuous function into a highly redundant function of two continuous variables-translation and scale.	Discrete wavelet transform	Continuous wavelet transform	Fast Wavelet transform	Fourier transform	Continuous wavelet transform
13	_____________ is based on the concept of compromising the accuracy of the reconstructed image m exchange for increased compression.	lossy encoding	lossless encoding	lzw encoding	transform coding	lossy encoding
14	______________ is error-free compression approach that does not require decomposition of an image into a collection of bit planes	lossy encoding	lossless encoding	lzw encoding	transform coding	lossless encoding
15	An image is decomposed into a set of band-limited components called ___________	subbands	frequency band	bandwidth	bitplane	subbands
16	____________ absorbs the nearest integer function of the error-free encoder, IS inserted between the symbol encoder and the point at which the prediction error is formed.	quantizer	encoder	decoder	predictor	quantizer
17	Orthonormal filter is developed around filter called ___________	up sampling	filtering	Digital segment filtering	prototype	prototype
18	The base of image pyramid contains ___________	low resolution	high resolution	intensity	blurred portion	high resolution
19	FWT stands for ___________	Fast wavelet transformation	Fast wavelet transform	Fourier wavelet transform	Fourier wavelet transformation	Fast wavelet transform
20	Wavelet series equation is the sum of ___________	scaling coefficient	detail coefficient	span coefficient	Both scaling and detail coefficient	Both scaling and detail coefficient
21	DWT stands for ___________	Discrete wavelet transform	Discrete wavelet transformation	Digital wavelet transform	Digital wavelet transformation	Discrete wavelet transform
22	In MRA, a _______ is used to create a series of approximations of a function or image.	scaling function	signal function	series function	wavelet function	scaling function
23	Scaling vectors are taken as ___________	heights	sharpness	intensity	weights	weights

24	MRA equation is also called ____________	modulating equation	FIR filter	dilation equation	span equation	dilation equation
25	Subspaces spanned are nested at ____________	lower scales	higher scales	mid scales	intense scales	higher scales
26	Prediction residual pyramid is computed in ____________	2 steps	3 steps	4 steps	5 steps	3 steps
27	K multiplication constants in digital filters are called ____________	co efficient	multipliers	subtractors	filter coefficients	filter coefficients
28	The size of the base image will be ____________	N-1 x N-1	N+1 x N-1	N-1 x N	N x N	N x N
29	MRA stands for ____________	Multiresolution analysis	Multiresolution assembly	Multiresemble analysis	Multiresemble assembly	Multiresolution analysis
30	Discarding every sample is called ____________	up sampling	filtering	down sampling	blurring	down sampling
31	Images are ____________	1D arrays	2D arrays	3D arrays	4D arrays	2D arrays
32	Filter banks consists of ____________	1 FIR filter	2 FIR filters	3 FIR filters	4 FIR filters	2 FIR filters
33	Neighborhood averaging produces ____________	histogram	pyramids	mean pyramids	equalized histogram	mean pyramids
34	Decomposing image into band limit components is called ____________	low coding	high coding	intense coding	subband coding	subband coding
35	Narrow wavelets represents ____________	sharp details	finer details	blur details	edge details	finer details
36	High contrast images are considered as ____________	low resolution	high resolution	intense	blurred	low resolution
37	Function having compact support is ____________	histogram	pyramids	mean pyramids	haar function	haar function
38	In multiresolution processing * represents the ____________	complete conjugate operation	complex conjugate operation	complete complex operation	complex complex operation	complex conjugate operation
39	The apex of image pyramid contains ____________	low resolution	high resolution	intensity	blurred portion	low resolution
40	Representing image in more than one resolution is ____________	histogram	image pyramid	local histogram	equalized histogram	image pyramid
41	Moving up in pyramid the size ____________	increases	remain same	decreases	blurred	decreases
42	The scaling function is ____________	pentagonal	square	orthogonal	oval	orthogonal
43	CWT stands for ____________	Complete wavelet transform	Complex wavelet transform	Continuous wavelet transform	Continuous wavelet transformation	Continuous wavelet transform

44	Integer wavelet translates are ___________	pentagonal	square	orthogonal	oval	orthogonal
45	Low contrast images are considered as ___________	low resolution	high resolution	intense	blurred	high resolution
46	Processing the image in small parts is ___________	histogram	pyramids	local histogram	equalized histogram	local histogram
47	Function can be represented with ___________	arbitrary precision	filtering	down sampling	prototype	arbitrary precision
48	One that is not a part of digital filter ___________	unit delay	multiplier	subtractor	adder	subtractor
49	The function changing the shape is called ___________	scaling function	shaping function	down sampling	blurring	scaling function
50	Function space is referred to as ___________	open span	fully span	closed span	span	closed span
51	No filtering produces ___________	Gaussian pyramids	pyramids	mean pyramids	subsampling pyramids	subsampling pyramids
52	Diagonally opposed filters is said to be ___________	modulation	multiplier	cross modulation	subband coding	cross modulation
53	The image pyramid contains ___________	j levels	j-1 levels	j+1 levels	n levels	j levels
54	Subband of input image, showing a(m,n) is called ___________	approximation	vertical detail	horizontal detail	diagonal detail	approximation
55	Decomposition in subband coding is performed to	segment image	reconstruct image	blur image	sharpened image	reconstruct image
56	Lowpass Gaussian filtering produces _______	Gaussian pyramids	pyramids	mean pyramids	equalized histogram	Gaussian pyramids
57	_____________ an invisible signature embedde	Watermark	compression	enhancement	restoration	Watermark
58	___________cannot be seen on the original images and the picture looks like an original image though it has the watermark.	invisible watermarking	visible watermarking	fragile watermarking	public watermarking	invisible watermarking
59	__________watermarks can be demolished by the data manipulation and these are also called as tamper proof watermark.	invisible watermarking	visible watermarking	fragile watermarking	public watermarking	fragile watermarking
60	_____________ does not have the protection and these can be read by everyone by availing the unique algorithm.	invisible watermarking	visible watermarking	fragile watermarking	public watermarking	public watermarking

KARPAGAM ACADEMY OF HI

CLASS: III BSC IT COURSE NAME: I

COURSE CODE: 16ITU5

UNIT V

SNO	QUESTIONS	CHOICE 1
1	Closing is represented by ___________	A .B
2	Closing is used for ___________	separation
3	Erosion is used for object ___________	removing lines
4	Reflection is applied on image's ___________	x coordinate
5	Opening with rolling SE ___________	sharps
6	For line detection we assume that lines are ___________	thin
7	The maximum length of the breaks can gets upto ___________	4pixels
8	Structuring elements runs over image's ___________	rows
9	Structuring elements usually are ___________	square array
10	Duality principle is used when SE is ___________	square
11	Dilation followed by erosion is called ___________	opening
12	Reflection and translation of the image objects are based on ___________	pixels
13	Opening smooths the image's ___________	pixels
14	Two main operations of morphology are ___________	erosion
15	Structuring elements have origins at ___________	top left
16	Structuring element is also called ___________	pixels
17	With dilation process images get ___________	thinner
18	Opening and closing are each others ___________	neighbors
19	Fully containment of the SE in an image is required in ___________	erosion
20	Erosion followed by dilation is called ___________	opening
21	To make the SE rectangular array approach that is used is called ___________	padding
22	Hit-or-miss transformation is used for shape ___________	removal
23	Mathematical morphology is a ___________	set theory
24	Duality principle is valid to involved ___________	one equation
25	Opening is represented by ___________	A o B
26	Subimages used to probe the image is called ___________	pixels
27	Closing produces ___________	narrow breaks
28	Replacing the object from its origin referred to as ___________	reflection
29	Dilation is used for	bridging gaps

31	If the inner region of the object is textured then approach we use is ___________	discontinuity
32	To avoid the negative values taking absolute values in lapacian image doubles ___________	thickness of lines
33	Gradient magnitude images are more useful in ___________	point detection
34	Image having gradient pixels is called ___________	sharp image
35	Diagonal lines are angles at ___________	0
36	Transition between objects and background shows ___________	ramp edges
37	Horizontal lines are angles at ___________	0
38	For edge detection we use ___________	first derivative
39	Sobel gradient is not that good for detection of ___________	horizontal lines
40	Method in which images are input and attributes are output is called ___________	low level processes
41	Image morphology is an important tool in extraction of image ______________	features
42	The difference between the original image and the	higher level gray
43	Top-hat transform is used for _____________	highlighting the right peaks
44	The theory of mathematical morphology is based on_______________	image size
45	well transform is used for _____________	highlighting the bright peaks
46	Thinning operation is used to remove the __________pixels	foreground
47	Morphlogical gradient gives___________	transition from spatial to frequency
48	Structering element is a ______________	mask
49	_____________is a process of removing of the extra tail pixels in an image	erosion
50	Watershed is process of ______________ the object	histogram
51	_________________is process of partition the digital image in to multiple regions	merging
52	__________is set of connected pixel that lie on the boundary between two regions.	point
53	The objective of the sharpening filter is __________	highlight the intensity transitions
54	______________has number of peaks	bimodel histogram

55	_____________is the starting pixel of region growing process.	seed pixel
56	____________is a deformable model that fits a model for segmenting ROI	tiger
57	_______________is the position of sign change of the first derivative among neighboring points	edge
58	__________ has unimodel histogram	one pixel
59	ROI stands for ________________	region of image
60	The hough transform is used to fit points as __________________	line

IGHER EDUCATION

DIGITAL IMAGE PROCESSING
04A BATCH-2016-2019

CHOICE 2	**CHOICE 3**	**CHOICE 4**	**ANSWER**
A+B	A-B	AxB	A .B
compression	decompression	filling holes	filling holes
producing lines	blurring image	sharpening image	removing lines
y coordinate	z coordinate	Both A and B	Both x and y coordinate
shrinks	smooths	deletes	smooths
thick	sharp	blur	thin
3pixels	2pixels	1pixel	2pixels
columns	edges	every element	every element
circular array	triangular array	rectangular array	rectangular array
symmetric	asymmetric	translated	symmetric
closing	blurring	translation	closing
frames	structuring elements	coordinates	structuring elements
lines	contour	boundary	contour
dilation	set theory	Both erosion and dilation	Both erosion and dilation
top right	center	bottom left	center
lines	subimage	noise	subimage
shrinked	thickened	sharpened	thickened
duals	centers	corners	duals
dilation	opening	closing	erosion
closing	blurring	translation	opening
logic diagram	set theory	map	padding
detection	compression	decompression	detection
logic diagram	graph	map	set theory
both equations	any equation	Both A and B	both equations
A+B	A-B	AxB	A o B
frames	structuring elements	coordinates	structuring elements
lines	dots	noise	narrow breaks
compression	decompression	translation	translation
compression	decompression	translation	bridging gaps

similarity	extraction	recognition	similarity
thinness of lines	thickness of edges	thinness of edges	thickness of lines
line detection	area detection	edge detection	edge detection
blur image	gradient image	binary image	gradient image
30	45	90	45
step edges	sharp edges	Both ramp and step edges	Both ramp and step edges
30	45	90	0
second derivative	third derivative	Both A and B	first derivative
vertical lines	Diagonal lines	edges	Diagonal lines
high level processes	mid level processes	edge level processes	mid level processes
colour	intensities	nature	features
low lever gray level	boundary	unfilled regions	boundary
highlighting the dark peaks	highlighting the bright and dark peaks	highlighting the dark and bright peaks	highlighting the right peaks
set theory	probability	correlation	set theory
highlighting the dark peaks	highlighting the bright and dark peaks	highlighting the dark and bright peaks	highlighting the dark peaks
back ground	object	image	foreground
transition from dark to bright	transition from frequency to spatial	none	transition from dark to bright
colour	background	pixel	mask
dilation	hit-miss transform	pruning	pruning
locating	transform	highliting	locating
filling	splitting	transform	splitting
edge	colour	line	edge
highlight the low transitions	highlight the bright transitions	highlight the colour transitions	highlight the intensity transitions
multimodel histogram	histogram	image	multimodel histogram

base pixel	original pixel	image	seed pixel
snake	goat	image	snake
zero-crosing	line	point	zero-crosing
one peak	one valley	one intensity level	one peak
region of interest	region of indicator	restoration of image	region of interest
edge	curve	ROI	curve

KARPAGAM ACADEMY OF HIGHER EDUCATION
(Deemed to be University)
(Established Under Section 3 of UGC Act 1956)
Coimbatore – 641 021

INFORMATION TECHNOLOGY

Fifth Semester
FIRST INTERNAL EXAMINATION – JULY 2018

DIGITAL IMAGE PROCESSING

Class: III B.Sc. IT (A&B) **Duration:** 2 Hours
Date & Session: .7.2017 **Maximum :** 50
Subject Code : 16ITU504A

PART-A (20 X 1 = 20 Marks)
Answer ALL the Questions

1. The value of x,y and amplitude of f are all finite, discrete quantities then we call the image as _____.
   a) Projected image b) Digital image c) Normal image d) Spaital image

2. The amplitude of an image at any pair of coordinate is called as ______.
   a) Frequency b) Light c) Intensity d) Digital

3. The range of light intensity levels to which the human visual system can adapt is _________.
   a) 10^8 b) 10^6 c) 10^{15} d) 10^{10}

4. The process of digitizing the coordinate values is called _______.
   a) Sampling b) Quantization c) Amplitude d) Continuous

5. The set of pixels with four horizontal and vertical neighbours of P, is denoted as _____.
   a) N b) $N_8(P)$ c) $N_4(P)$ d) P_4

6. The process of connectivity between pixels in a digital image is called as ________.
   a) Pixel b) Region c) Line d) Curve

7. The processed image is obtained by mapping each pixel with level r_k in the input image into a corresponding pixel with level s_k in the output image is called ______ equalization.
   a) Histogram b) Linear c) Equalization d) Transformation

8. The subimage of an image is called as __________.
   a) neighbor b) Value c) Filter d) Mask

9. The value in a filter subimage are referred as ________.
   a) Pixel b) Coefficient c) Coordinate d) Value

10. The process of filtering that are performed directly on the pixels of an image is called ________ filtering.
    a) Frequency b) Linear c) Spatial d) Domain

11. The ___________ filter is used for blurring and for noise reduction in an image.
    a) Smoothing b) Sharpening c) Linear d) Continuous

12. The average of the pixels contained in the neighbourhood of the filter mask is called ______ filter.
    a) Low b) Domain c) averaging d) Sharpening

13. A spatial averaging filter in which all coefficients are equal is called _______.
    a) Box filter b) Linear filter c) Average filter d) Pixel

14. The process of separating a function into various components based on frequency content is called as ___________.
    a) Fourier b) Mathematical prism c) Discrete d) Transform

15. The principal use of correlation in an image containing objects or region is called as ________.
    a) Matching b) Mapping c) Coordinate d) Template

16. D distance is also called as ______________ distance.
    a) City block b) Chess board c) Euclidean d) Mean

17. Image restoration is used to improve the ___________ image.
    a) Quantity b) Quality c) Blur d) restore

18. Which of the following make an image difficult to enchance?
    a) Narrow intensity b) Dynamic intensity c) High noise d) All the above

19. Which of the following is a second- order derivative operator?
    a) Histogram b) Laplacian c) Gaussian d) Gradiant

20. Which of the following fails to work on dark intensity distribution?
    a) Laplacian Transform b) Gaussian Transform c) Histogram Equalization d) Transform

PART-B (3 X 2 = 6 Marks)
(Answer ALL the Questions)

21. What is called Light spectrum in digital image processing?
22. Write short notes on contrast stretching of spatial domain.
23. Define Correlation and Convolution.

PART-C (3 X 8 = 24 Marks)
(Answer ALL the Questions)

24. a) Illustrate the concept of Brightness adaption in image processing.

(or)

b) Explain briefly about Pixels and Coordinates conventions.

25. a) Describe about Sampling and Quantization of image processing.

(or)

b) Briefly explain about Smoothing and Sharpening filters.

26. a) Discuss about Fourier transformation and its properities.

(or)

b) Explain in detail about the concept of 2-D Sampling.

KARPAGAM ACADEMY OF HIGHER EDUCATION
(Deemed to be University)
(Established Under Section 3 of UGC Act 1956)
Coimbatore – 641 021

INFORMATION TECHNOLOGY

Fifth Semester
FIRST INTERNAL EXAMINATION – JULY 2018

DATA MINING

Class: III B.Sc. IT (A&B) | **Duration:** 2 Hours
Date & Session: .7.2017 | **Maximum :** 50
Subject Code : 16ITU503B

PART-A (20 X 1 = 20 Marks)
Answer ALL the Questions

1. Data mining is __________________
a) The actual discovery phase of a knowledge discovery process
b) The stage of selecting the right data for a KDD process
c) A collection of data in support of management
d) The analysis of data

2. The removal of noise and inconsistent data in data mining is called as ______________.
a) Data integration **b) Data Cleaning** c) Data modeling d) Data evalution

3. The process of constructing the final dataset from raw data is called as data __________.
a) Understanding b) Evaluation **c) Preparation** d) Deployment

4. Which of the following is not involve in data mining?
a) Knowledge extraction b)Data archaeology
c) Data exploration **d) Data transformation**

5. The study of frequency of items occuring together in transactional databases is called _____________.
a) Association b) Clustering c) Classification d) Regression

6. The principle of maximizing the similarity between objects in a same class is called as ________________ similarity.
a) inter b) data c) class **d) intra**

7. The process of analyzing the current and past states of the attributes and prediction of its future state is called as ____________.

a) Descriptive b) **Predictive** c) Pattern d) Segmentation

8. Classification is _______________ type of learning.
   a) Unsupervised b) Relationship **c) Supervised** d) Domain

9. The process of inspecting, cleaning and modelling data with the objective of discovery useful information is known as ___________.
   a) Data Project b) Data Collection **c) Data Analysis** d) Statistics

10. A statistical methodology that is most often used for numeric prediction is called as _________________.
    a) Statistical **b) Regression** c) Training set d) Class

11. Task of inferring a model from labeled training data is called ________________ learning.
    a) Unsupervised **b) Supervised** c) Reinforcement d) Machine

12. A subdivision of a set of examples into a number of classes is called as _____________.
    a) Time series b) Summarization c) Prediction **d) Classification**

13. The classification is according to the data handled is called as classification under ____________
    a) Type of data b) Database c) Knoweldge d) Techniques

14. Among the following which is not a phase in life cycle of data mining.
    a) Data understanding b) Data preparation **c) Correlation** d) Evaluation

15. The _____________is a summarization of the general characteristics or features of a target class of data.
    a) Characterization b) Classification c) Discrimination d) Selection

16. Strategic value of data mining is ______________.
    a) Cost - sensitive b) Work - sensitive
    c) Time - sensitive d) Technical sensitive

17. The full form of KDD is _________.
    a) Knowledge Database **b) Knowledge Discovery Database**
    c) Knowledge Data House d) Knowledge Data Definition

18. The output of KDD is ____________.
    a) Data **b) Information** c) Query d) Labels

19. The ____________ is an essential process where intelligent methods are applied to extract data patterns.
    a) Data warehousing **b) Data mining** c) Text mining d) Data selection

20. Data mining can also applied to other forms such as ________________.

a) Data streams b) Sequence data c) Networked data d) Spatial data

PART-B (3 X 2 = 6 Marks)
(Answer ALL the Questions)

21. Define Datamining. What are the other names of data mining?

Data Mining refers to extracting or “mining” knowledge from large amounts of data. The overall goals of the data mining process is extraction of information from large data sets and transform it into some understandable structure for further uses. Data mining involves the use of sophisticated data analysis tools to discover previously unknown, valid patterns and relationships in large data sets. Different names of data mining are knowledge extraction, information discovery, information harvesting, data archeology, and data pattern processing.

22. Write a short note on classification techniques.

Classification uses given class labels to order the objects in the data collection. Classification approaches normally use a training set where all objects are already associated with known class labels. The classification algorithm learns from the training set and builds a model. The model is used to classify new objects. For example, after starting a credit policy, the manager of a store could analyze the customers’ behavior vis-à-vis their credit, and label accordingly the customers who received credits with three possible labels “safe”, “risky” and “very risky”. The classification analysis would generate a model that could be used to either accept or reject credit requests in the future.

23. Define superviesd and unsupervised learning in data mining.

Definition for Supervised data mining:

Supervised learning is the Data mining task of inferring a function from **labeled training data**. The training data consist of a set of training examples. In supervised learning, each example is a pair consisting of an input object (typically a vector) and a desired output value (also called the**supervisory signal**). A **supervised learning algorithm** analyzes the training data and produces an inferred function, which can be used for mapping new examples. An optimal scenario will allow for the algorithm to correctly determine the class labels for **unseen instances**. This requires the learning algorithm to generalize from the training data to unseen situations in a “reasonable” way.

Definition for UnSupervised data mining:

In Data mining, the problem of **unsupervised learning** is that of trying to find hidden structure in unlabeled data. Since the examples given to the learner are unlabeled, there is no error or reward signal to evaluate a potential solution. The goal is rather to establish some relationship among all the variables in the data.

PART-C (3 X 8 = 24 Marks)
(Answer ALL the Questions)

24. a) Explain about the classification and life cycle of data mining.

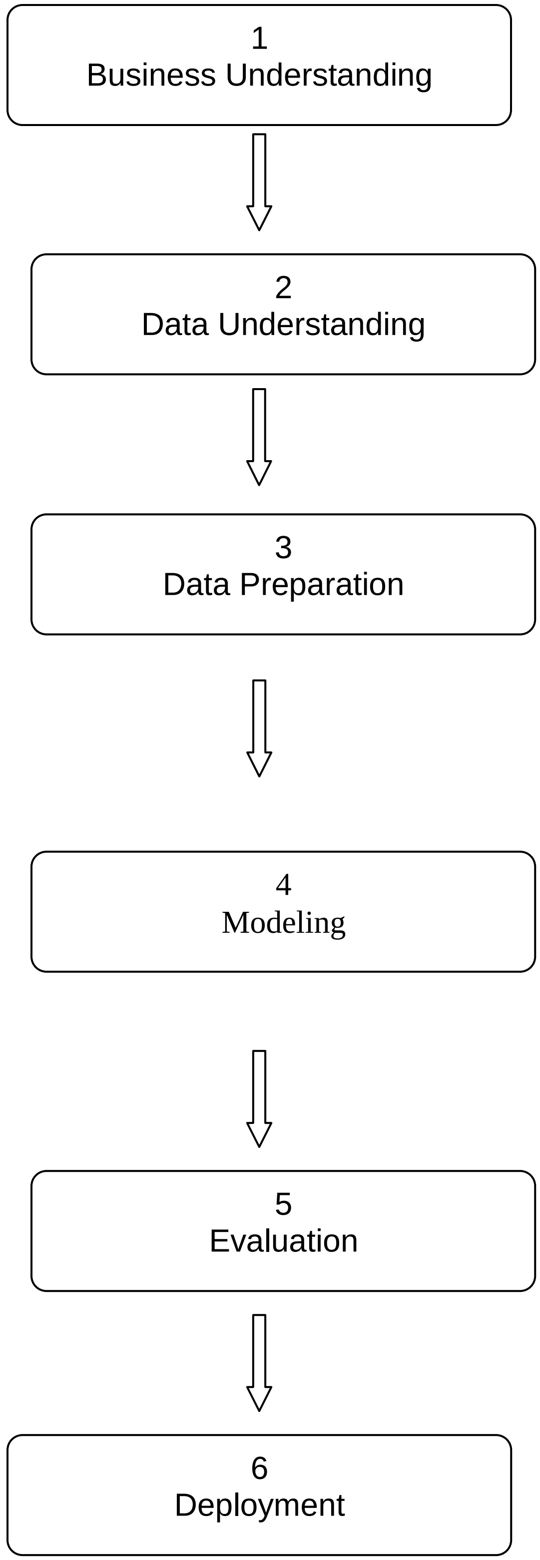
Data Mining Life cycle

The life cycle of a data mining project consists of six phases [91, 26]. The sequence of the phases is not rigid. Moving back and forth between different phases is always required depending upon the outcome of each phase. The main phases are:

Business Understanding:

This phase focuses on understanding the project objectives and requirements from a business perspective, then converting this knowledge into a data mining problem definition and a preliminary plan designed to achieve the objectives.

Figure 3.2: Phases of Data Mining Life Cycle



Data Understanding:

It starts with an initial data collection, to get familiar with the data, to identify data quality problems, to discover first insights into the data or to detect interesting subsets to form hypotheses for hidden information.

Data Preparation:

Covers all activities to construct the final dataset from raw data.

Modeling:

In this phase, various modeling techniques are selected and applied and their parameters are calibrated to optimal values.

Evaluation:

In this stage the model is thoroughly evaluated and reviewed. The steps executed to construct the model to be certain it properly achieves the business

Deployment:

The purpose of the model is to increase knowledge of the data, the knowledge gained will need to be organized and presented in a way that the customer can use it. The deployment phase can be as simple as generating a report or as complex as implementing a repeatable data mining process across the enterprise.

b) Describe briefly about the functionalities of data mining.

Data Mining Functionalities

The kinds of patterns that can be discovered depend upon the data mining tasks employed. By and large, there are two types of data mining tasks: descriptive data mining tasks that describe the general properties of the existing data, and predictive data mining tasks that attempt to do predictions based on inference on available data. The data mining functionalities and the variety of knowledge they discover are briefly presented in the following list:

Characterization:

It is the summarization of general features of objects in a target class, and produces what is called characteristic rules. The data relevant to a userspecified
class are normally retrieved by a database query and run through a summarization module to extract the essence of the data at different levels of abstractions. For example, one may wish to characterize the customers of a store who regularly rent more than 30 movies a year. With concept hierarchies on the attributes describing the target class, the attribute oriented induction method can be used to carry out data summarization. With a data cube containing summarization of data, simple OLAP operations fit the purpose of data characterization.

Discrimination:

Data discrimination produces what are called discriminant rules and is basically the comparison of the general features of objects between two classes referred to as the target class and the contrasting class. For example, one may wish to compare the general characteristics of the customers who rented more than 30 movies in the last year with those whose rental account is lower than 5. The techniques used for data discrimination are similar to the techniques used for data characterization with the exception that data discrimination results include comparative measures.

Association analysis:

Association analysis studies the frequency of items occurring together in transactional databases, and based on a threshold called support, identifies the frequent item sets. Another threshold, confidence, which is the conditional probability than an item appears in a transaction when another item appears, is used to pinpoint association rules. This is commonly used for market basket analysis. For example, it could be useful for the manager to know what movies are often rented together or if there is a relationship between renting a certain type of movies and buying popcorn or pop. The discovered association rules are of the form:

$$P \rightarrow Q \ [s, c],$$

where P and Q are conjunctions of attribute value-pairs, and s (support) is the probability that P and Q appear together in a transaction and c (confidence) is the conditional probability that Q appears in a transaction when P is present. For example, RentType(X,“game”)∧ Age(X,“13-19”)_Buys(X,“pop”)[s=2%, =55%]

The above rule would indicate that 2% of the transactions considered are of customers aged between 13 and 19 who are renting a game and buying a pop, and that there is a certainty of 55% that teenage customers who rent a game also buy pop.

Classification:

It is the organization of data in given classes. Classification uses given class labels to order the objects in the data collection. Classification approaches normally use a training set where all objects are already associated with known class labels. The classification algorithm learns from the training set and builds a model. The model is used to classify new objects. For example, after starting a credit policy, the manager of a store could analyze the customers’ behavior vis-à-vis their credit, and label accordingly the customers who received credits with three possible labels “safe”, “risky” and “very risky”. The classification analysis would generate a model that could be used to either accept or reject credit requests in the future.

Prediction:

Prediction has attracted considerable attention given the potential implications of successful forecasting in a business context. There are two major types of predictions: one can either try to predict some unavailable data values or pending trends, or predict a class label for some data. The latter is tied to classification. Once a classification model is built based on a training set, the class label of an object can be foreseen based on the attribute values of the object and the attribute values of the classes. Prediction is however more often referred to the forecast of missing numerical values, or increase/ decrease trends in time related data. The major idea is to use a large number of past values to consider probable future values.

Clustering:

Similar to classification, clustering is the organization of data in classes. However, unlike classification, in clustering, class labels are unknown and it is up to the clustering algorithm to discover acceptable classes. Clustering is also called unsupervised classification, because the classification is not dictated by given class labels. There are many clustering approaches all based on the principle of maximizing the similarity between objects in a same class (intra-class similarity) and minimizing the similarity between objects of different classes (inter-class similarity).

Outlier analysis:

Outliers are data elements that cannot be grouped in a given class or cluster. Also known as exceptions or surprises, they are often very important to identify. While outliers can be considered noise and discarded in some applications, they can reveal important knowledge in other domains, and thus can be very significant and their analysis valuable.

Evolution and deviation analysis:

Evolution and deviation analysis pertain to the study of time related data that changes in time. Evolution analysis models evolutionary trends in data, which consent to characterizing, comparing, classifying or clustering of time related data. Deviation analysis, on the other hand, considers differences between measured values and expected values, and attempts to find the cause of the deviations from the anticipated values. It is common that users do not have a clear idea of the kind of patterns they can discover or need to discover from the data at hand. It is therefore important to have a versatile and inclusive data mining system that allows the discovery of different kinds of knowledge and at different levels of abstraction. This also makes interactivity an important attribute of a data mining system.

25. a) Briefly explain the techniques followed in Descriptive data mining model.

Descriptive Models

The descriptive model identifies the patterns or relationships in data and explores the properties of the data examined. Ex. Clustering, Summarization, Association rule, Sequence discovery etc. Clustering is similar to classification except that the groups are not predefined, but are defined by the data alone. It is also referred to as unsupervised learning or segmentation. It is the partitioning or segmentation of the data in to groups or clusters. The clusters are defined by studying the behavior of the data by the domain experts. The term segmentation is used in very specific context; it is a process of partitioning of database into disjoint grouping of similar tuples. Summarization is the technique of presenting the summarize information from the data. The association rule finds the association between the different attributes. Association rule mining is a two step process: Finding all frequent item sets, Generating strong association rules from the frequent item sets. Sequence discovery is a process of finding the sequence patterns in data. This sequence can be used to understand the trend.

A. Clustering

Clustering is the most important descriptive data mining technique, in which a set of data items is partitioned into a set of classes also called as groups. Clustering is the process of finding natural groups, called as clusters, in a database. The set of data items in a group have similar characteristics. Hence, it is mainly used for finding groups of similar items. It is the identification process for classes for a set of objects whose classes are not known. These objects ware so clustered that their intra class similarities can be maximized and minimized on the criteria defined by attributes of objects. The object is label with their corresponding clusters once their particular class or clusters are decided. For example, between a data set of customers, that have a similar buying behavior can be identified with subgroups of customers.

B. Summarization

Summarization is called as the abstraction or generalization of data. The summarization technique maps data into subsets with simple descriptions. The summarized data set gives general overview of the data with aggregated information. Summarization can scale up to different levels of abstraction and can be viewed from different angles. It is a key data-mining concept involving techniques for finding a compact description of dataset. Summarization approaches are Basically Mean, Standard Deviation, Variance, tabulating, mode, and median. These approaches are often applied for data analysis, data visualization and automated report generation.

Consider example of long distance call, Costumer data related to call can be summarize as total minutes, total spending, total calls, etc. Such important information can be presented to sales manager for analysis of his costumer and his business. The scaling and viewing from different angles can be done for this example as, calling minutes and spending can be totaled along, its period in weeks, months or years. In the same way long distance, call can be summarized as state call, state-to-state call, national, Asia calls, America calls, etc. Further summarized as Domestic and international calls.

C. Association

The Associations or Link Analysis technique are used to discover relationships between attributes and items. In these techniques, the presence of one pattern implies the presence of another pattern i.e. item is related to another in terms of cause-and-effect. This is common in establishing a form of statistical relationships among different interdependent variables of a model. For market basket analysis, Association Rules is a important technique because all possible combinations of potentially interesting product groupings can be explored [11]. Association rules are as if/then statements that help uncover relationships between seemingly unrelated data in a transaction oriented database, relational database or other information repository. In data mining, association rules are useful for analyzing and predicting customer behavior. They also play an important role in shopping basket data analysis, product clustering, catalog design and store layout. This association rules are also build by programmers use to build programs capable of machine learning.

D. Feature Extraction

Feature Extraction extract a new set of features by decomposing the original set of data. This technique describes the data with a number of features far smaller than the number of original attributes. The word feature in the technique are combination of attributes in the data that have special important characteristics of the data [12]. Feature extraction is mostly applicable to latent semantic analysis, data compression, data decomposition and projection, and pattern recognition, etc. Using feature extraction process the speed and effectiveness of supervised learning can also be improved.

For example, feature extraction is used to extract the themes/features of a document collection, where documents are represented by a set of keywords and their frequencies. Each feature is represented by a combination of keywords. The documents can then be expressed from the collection in terms of the discovered themes.

b) Explain the applications of analytic process in data mining.

Predictive analytics

Predictive analytics is an area of statistics that deals with extracting information from data and using it to predict trends and behavior patterns. Often the unknown event of interest is in the future, but predictive analytics can be applied to any type of unknown whether it be in the past, present or future. For example, identifying suspects after a crime has been committed, or credit card fraud as it occurs.

Predictive Analytics Process

---

1. **Define Project** : Define the project outcomes, deliverable, scope of the effort, business objectives, identify the data sets that are going to be used.
2. **Data Collection** : Data mining for predictive analytics prepares data from multiple sources for analysis. This provides a complete view of customer interactions.

3. **Data Analysis** : Data Analysis is the process of inspecting, cleaning and modelling data with the objective of discovering useful information, arriving at conclusion
4. **Statistics** : Statistical Analysis enables to validate the assumptions, hypothesis and test them using standard statistical models.
5. **Modelling** : Predictive modelling provides the ability to automatically create accurate predictive models about future. There are also options to choose the best solution with multi-modal evaluation.
6. **Deployment** : Predictive model deployment provides the option to deploy the analytical results into everyday decision making process to get results, reports and output by automating the decisions based on the modelling.
7. **Model Monitoring** : Models are managed and monitored to review the model performance to ensure that it is providing the results expected.

In the current literature, concepts ”knowledge discovery”, ”data mining” and ”machine learning” are often used interchangeably. Sometimes the whole KDD process is called data mining or machine learning, or machine learning is considered as a subdiscipline of data mining. To avoid confusion, we follow a systematic division to descriptive and predictive modelling, which matches well with the classical ideas of data mining and machine learning1 . This approach has several advantages:

1. The descriptive and predictive models can often be paired as illustrated in Table 2.1. Thus, descriptive models indicate the suitability of a given predictive model and can guide the search of models (i.e. they help in the selection of the modelling paradigm and the model structure).

2. Descriptive and predictive modelling require different validation techniques. Thus the division gives guidelines, how to validate the results.

3. Descriptive and predictive models have different underlying assumptions (bias) about good models. This is reflected especially by the score functions, which guide the search.

26. a) Write briefly about supervised and unsupervised machine learning algorithms.

Supervised and Unsupervised Machine Learning Algorithms

Supervised Machine Learning

The majority of practical machine learning uses supervised learning.

Supervised learning is where you have input variables (x) and an output variable (Y) and you use an algorithm to learn the mapping function from the input to the output.

Y = f(X)

The goal is to approximate the mapping function so well that when you have new input data (x) that you can predict the output variables (Y) for that data.

It is called supervised learning because the process of an algorithm learning from the training dataset can be thought of as a teacher supervising the learning process. We know the correct answers, the algorithm iteratively makes predictions on the training data and is corrected by the teacher. Learning stops when the algorithm achieves an acceptable level of performance.

Supervised learning problems can be further grouped into regression and classification problems.

- **Classification**: A classification problem is when the output variable is a category, such as "red" or "blue" or "disease" and "no disease".
- **Regression**: A regression problem is when the output variable is a real value, such as "dollars" or "weight".

Some common types of problems built on top of classification and regression include recommendation and time series prediction respectively.

Some popular examples of supervised machine learning algorithms are:

- Linear regression for regression problems.
- Random forest for classification and regression problems.

- Support vector machines for classification problems.

Unsupervised Machine Learning

Unsupervised learning is where you only have input data (X) and no corresponding output variables.

The goal for unsupervised learning is to model the underlying structure or distribution in the data in order to learn more about the data.

These are called unsupervised learning because unlike supervised learning above there is no correct answers and there is no teacher. Algorithms are left to their own devises to discover and present the interesting structure in the data.

Unsupervised learning problems can be further grouped into clustering and association problems.

- **Clustering**: A clustering problem is where you want to discover the inherent groupings in the data, such as grouping customers by purchasing behavior.
- **Association**: An association rule learning problem is where you want to discover rules that describe large portions of your data, such as people that buy X also tend to buy Y.

Some popular examples of unsupervised learning algorithms are:

- k-means for clustering problems.
- Apriori algorithm for association rule learning problems.

Semi-Supervised Machine Learning

Problems where you have a large amount of input data (X) and only some of the data is labeled (Y) are called semi-supervised learning problems.

These problems sit in between both supervised and unsupervised learning.

A good example is a photo archive where only some of the images are labeled, (e.g. dog, cat, person) and the majority are unlabeled.

Many real world machine learning problems fall into this area. This is because it can be expensive or time-consuming to label data as it may require access to domain experts. Whereas unlabeled data is cheap and easy to collect and store.

You can use unsupervised learning techniques to discover and learn the structure in the input variables.

You can also use supervised learning techniques to make best guess predictions for the unlabeled data, feed that data back into the supervised learning algorithm as training data and use the model to make predictions on new unseen data.

Summary

In this post you learned the difference between supervised, unsupervised and semi-supervised learning. You now know that:

Supervised: All data is labeled and the algorithms learn to predict the output from the input data.

Unsupervised: All data is unlabeled and the algorithms learn to inherent structure from the input data.

Semi-supervised: Some data is labeled but most of it is unlabeled and a mixture of supervised and unsupervised techniques can be used.

Do you have any questions about supervised, unsupervised or semi-supervised learning? Leave a comment and ask your question and I will do my best to answer it.

b) Illustrate the major issues in supervised learning methods.

Data mining techniques come in two main forms: supervised (also known as predictive or directed) and unsupervised (also known as descriptive or undirected). Both categories encompass functions capable of finding different hidden patterns in large data sets.

Data analytics tools are placing more emphasis on self service, it's still useful to know which data mining operation is appropriate for your needs before you begin a data mining operation.

Supervised Data Mining

Supervised data mining techniques are appropriate when you have a specific target value you'd like to predict about your data. The targets can have two or more possible outcomes, or even be a continuous numeric value (more on that later).

To use these methods, you ideally have a subset of data points for which this target value is already known. You use that data to build a model of what a typical data point looks like when it has one of the various target values. You then apply that model to data for which that target value is currently unknown. The algorithm identifies the "*new*" data points that match the model of each target value.

Classification

As a supervised data mining method, classification begins with the method.

For example:Imagine you're a credit card company and you want to know which customers are likely to default on their payments in the next few years.

You use the data on customers who have and have not defaulted for extended periods of time as build data (or training data) to generate a classification model. You then run that model on the customers you're curious about. The algorithms will look for customers whose attributes match the attribute patterns of previous defaulters/non-defaulters, and categorize them according to which group they most closely match. You can then use these groupings as indicators of which customers are most likely to default.

Similarly, a classification model can have more than two possible values in the target attribute. The values could be anything from the shirt colors they're most likely to buy, the promotional methods they'll respond to (mail, email, phone), or whether or not they'll use a coupon.

Regression

Regression is similar to classification except that the targeted attribute's values are numeric, rather than categorical. The order or magnitude of the value is significant in some way.

To reuse the credit card example, if you wanted to know what threshold of debt new customers are likely to accumulate on their credit card, you would use a regression model.

Simply supply data from current and past customers with their maximum previous debt level as the target value, and a regression model will be built on that training data. Once run on the new customers, the regression model will match attribute values with predicted maximum debt levels and assign the predictions to each customer accordingly.

This could be used to predict the age of customers with demographic and purchasing data, or to predict the frequency of insurance claims.

Anomaly Detection

Anomaly detection identifies data points atypical of a given distribution. In other words, it finds the outliers. Though simpler data analysis techniques than full-scale data mining can identify outliers, data mining anomaly detection techniques identify much more subtle attribute patterns and the data points that fail to conform to those patterns.

Most examples of anomaly detection uses involve fraud detection, such as for insurance or credit card companies.

Reg.No.________________

[16ITU504A]

KARPAGAM ACADEMY OF HIGHER EDUCATION
(Deemed to be University)
(Established Under Section 3 of UGC Act 1956)
Coimbatore – 641 021

INFORMATION TECHNOLOGY

Fifth Semester
SECOND INTERNAL EXAMINATION – AUGUST 2018

DIGITAL IMAGE PROCESSING

Class: III B.Sc. IT (A&B) **Duration :** 2 Hours
Date & Session : 16.8.2017 **Maximum :**50 Marks

PART-A (20 X 1 = 20 Marks)
Answer ALL the Questions

1. Product of two functions in spatial domain is what, in frequency domain _________.
   a) correlation b) convolution c) Fourier transform d) fast Fourier transform
2. Digitizing the coordinate values is called __________.
   a) quantization b) sampling c) Fourier transform d) discrete Fourier
3. Shrinking of image is viewed as ____________.
   a) under sampling b) over sampling c) critical sampling d) nyquist sampling
4. High pass filters are used for image ____________.
   a) contrast b) sharpening c) blurring d) resizing
5. The Fourier transform is named after French mathematician ______________.
   a) joseph Fourier b) john Fourier c) sean Fourier d) jay Fourier
6. What is the thickness of the edges produced by first order derivatives when compared to that of second order derivatives?
   a) Finer b) Equal c) Thicker d) Independent
7. What is accepting or rejecting certain frequency components called as ____________.
   a) Filtering b) Eliminating c) Slicing d) None of the above
8. What is the process of moving a filter mask over the image and computing the sum of products at each location called as _________.
   a) Convolution b) Correlation c) Linear filtering d) Non linear filtering
9. Convolution and Correlation are functions of _____________.
   Distance b) Time c) Intensity d) Displacement
10. Due to uniform linear motion image is_________________.
    a) blurred b) sharpened c) smoothened d) blurred and smoothened
11. Blur is characterized by the _______________ response of the system.
    a) filter b) noise c) impulse d) image

12. Gaussian noise is referred to as __________.
    a) red noise   b) black noise   c) white noise   d) normal noise
13. Convolution in spatial domain is multiplication in ___________.
    a) frequency domain   b) time domain   c) spatial domain   d) plane
14. Linear functions possess the property of _____________.
    a) additive   b) homogeneity   c) multiplication   d) Both A and B
15. Information lost when expressed mathematically is called _____________.
    a) Markov   b) finite memory source   c) fidelity criteria   d) noiseless theorem
16. Error of the image is referred to as __________.
    a) pixels   b) matrix   c) frames   d) noise
17. The _______ and ________ are two important statistical measures on which the adaptive filtering depends upon.
    a) Mean and variance   b) Sum and average   c) Mean and median   d) Max and Min
18. In ________ data compression, the integrity of the data is preserved.
    a) lossless   b) lossy   c) image   d) text
19. Morphology refers to _____________.
    a) pixels   b) matrix   c) frames   d)shape
20. FAX is an abbreviation of _____________.
    a) fast   b) female   c) feminine   d) facsimile

PART-B (3 X 2 = 6 Marks)
(Answer ALL the Questions)

21. Define low pass and high pass filtering.
22. Write short notes on Image restoration.
23. Define Adaptive filter.

PART-C (3 X 8 = 24 Marks)
(Answer ALL the Questions)

24. a) Explain the basic steps in frequency domain filtering.

    (or)

    b) Illustrate the concept of Geometric Transformation.

25. a) Describe about Image restoration techniques.

    (or)

    b) Briefly explain about Linear, Position invariant degradation.

26. a) Discuss about Degradation functions.

    (or)

    b) Explain in detail about Image Compression.

Reg.No.__________________

[16ITU504A]

KARPAGAM ACADEMY OF HIGHER EDUCATION
(Deemed to be University)
(Established Under Section 3 of UGC Act 1956)
Coimbatore – 641 021

INFORMATION TECHNOLOGY

Fifth Semester
SECOND INTERNAL EXAMINATION – AUGUST 2018

DIGITAL IMAGE PROCESSING

Class: III B.Sc. IT (A&B) **Duration** : 2 Hours
Date & Session : 16.8.2017 **Maximum** :50 Marks

PART-A (20 X 1 = 20 Marks)
Answer ALL the Questions

1. Product of two functions in spatial domain is what, in frequency domain _________.
   a) correlation b) **convolution** c) Fourier transform d) fast Fourier transform
2. Digitizing the coordinate values is called __________.
   a) quantization b) **sampling** c) Fourier transform d) discrete Fourier
3. Shrinking of image is viewed as ____________.
   a) **under sampling** b) over sampling c) critical sampling d) nyquist sampling
4. High pass filters are used for image ____________.
   a) contrast b) **sharpening** c) blurring d) resizing
5. The Fourier transform is named after French mathematician ______________.
   a) **joseph Fourier** b) john Fourier c) sean Fourier d) jay Fourier
6. What is the thickness of the edges produced by first order derivatives when compared to that of second order derivatives?
   a) Finer b) Equal c) **Thicker** d) Independent
7. What is accepting or rejecting certain frequency components called as ____________.
   a) **Filtering** b) Eliminating c) Slicing d) None of the above
8. What is the process of moving a filter mask over the image and computing the sum of products at each location called as _________.
   a) Convolution b) **Correlation** c) Linear filtering d) Non linear filtering
9. Convolution and Correlation are functions of _____________.
   Distance b) Time c) Intensity d) **Displacement**
10. Due to uniform linear motion image is_________________.
    a) blurred b) sharpened c) smoothened d) **blurred and smoothened**
11. Blur is characterized by the _______________ response of the system.
    a) filter b) noise c) **impulse** d) image
12. Gaussian noise is referred to as __________.
    a) red noise b) black noise c) white noise d) **normal noise**
13. Convolution in spatial domain is multiplication in ___________.
    a) **frequency domain** b) time domain c) spatial domain d) plane
14. Linear functions possess the property of _____________.
    a) additive b) homogeneity c) multiplication d) **Both A and B**
15. Information lost when expressed mathematically is called _____________.
    a) Markov b) finite memory source c) **fidelity criteria** d) noiseless theorem

16. Error of the image is referred to as __________.
a) pixels b) matrix c) frames d) **noise**

17. The _______ and ________ are two important statistical measures on which the adaptive filtering depends upon.
a) **Mean and varianc** b) Sum and average c) Mean and median d) Max and Min

18. In ________ data compression, the integrity of the data is preserved.
a) lossless b) lossy c) image d) text

19. Morphology refers to _____________.
a) pixels b) matrix c) frames d)**shape**

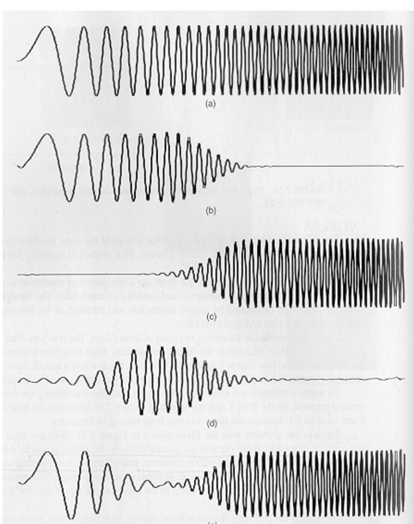
20. FAX is an abbreviation of _____________.
a) fast b) female c) feminine d) **facsimile**

PART-B (3 X 2 = 6 Marks)
(Answer ALL the Questions)

21. Define low pass and high pass filtering.

Converting an image to frequency domain, some filters are applied in filtering process to perform different kind of processing on an image. The processing include blurring an image, sharpening an image e.t.c.

A filter that attenuates high frequencies while “passing” low frequesies is a lowpass filter. A filter that attenuates low frequencies while “passing” high frequecies is a highpass filter.



Original

Low-pass

High-pass

Band-pass

Band-stop

22. Write short notes on Image restoration.

<u>IMAGE RESTORATION:</u>

Image restoration Image restoration techniques aim at modelling a degradation corrupting the image and inverting this degradation to correct the image so that it is as close as possible to the original.

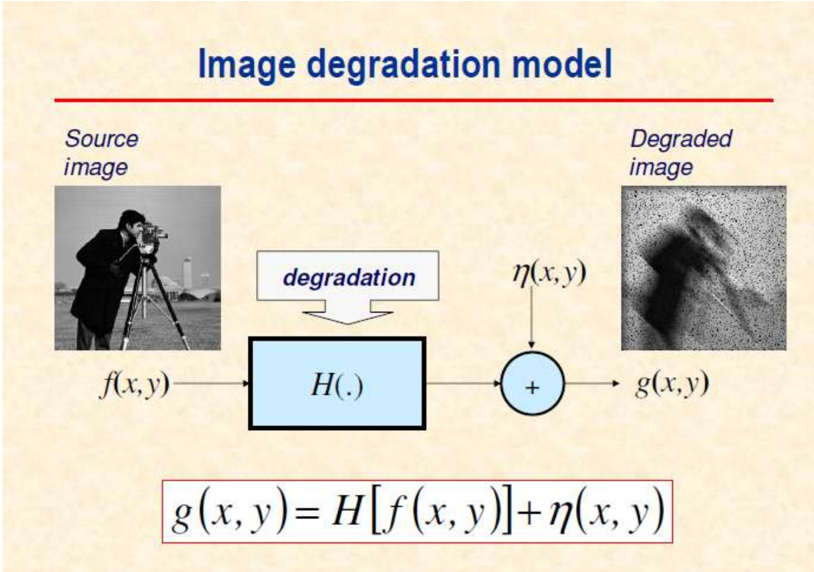


Image Restoration is the operation of taking a corrupt/noisy **image** and estimating the clean, original **image**. Corruption may come in many forms such as motion blur, noise and camera mis-focus. Image restoration is performed by reversing the process that blurred the image and such is performed by imaging a point source and use the point source image, which is called the Point Spread Function (PSF) to restore the image information lost to the blurring process.

23. Define Adaptive filter.

ADAPTIVE FILTERS

An **adaptive filter** is a system with a linear **filter** that has a transfer function controlled by variable parameters and a **means** to adjust those parameters according to an optimization algorithm.

Mean and variance are two important statistical measures on which the adaptive filtering is depends upon. For example if the local variance is high compared to the overall image variance, the filter should return a value close to the present value. Because high variance is usually associated with edges and edges should be preserved.

Noise reduction is the process of removing **noise** from a signal. All signal processing devices, both analog and digital, have traits that make them susceptible to **noise**.

PART-C (3 X 8 = 24 Marks)
(Answer ALL the Questions)

24. a) **Explain the basic steps in frequency domain filtering.**

Frequency domain filtering

Basic Steps

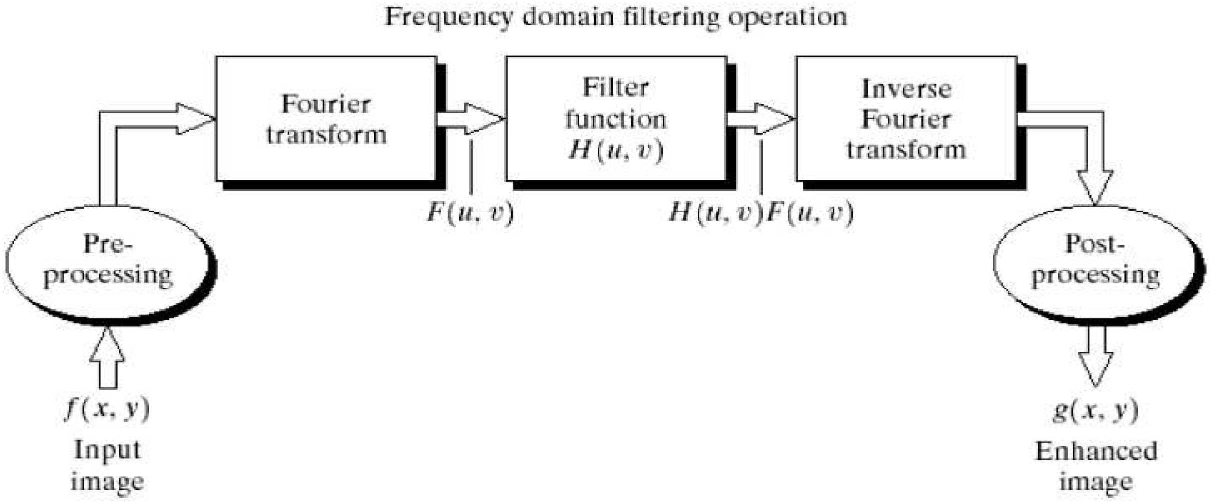
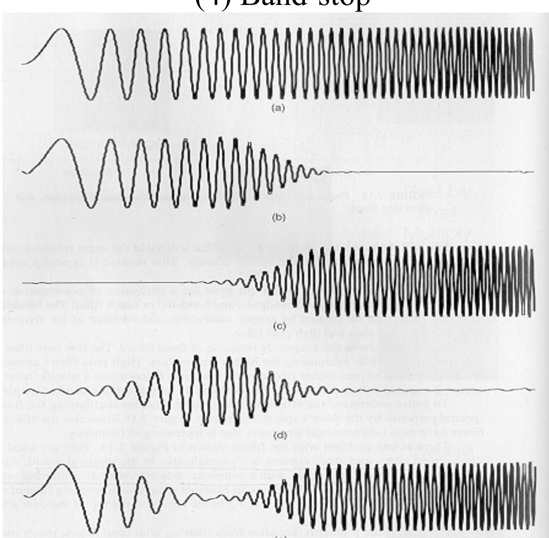


FIGURE 4.5 Basic steps for filtering in the frequency domain.

- Typically, filters are classified by examining their properties in the frequency domain:

    (1) Low-pass
    (2) High-pass
    (3) Band-pass
    (4) Band-stop



Original

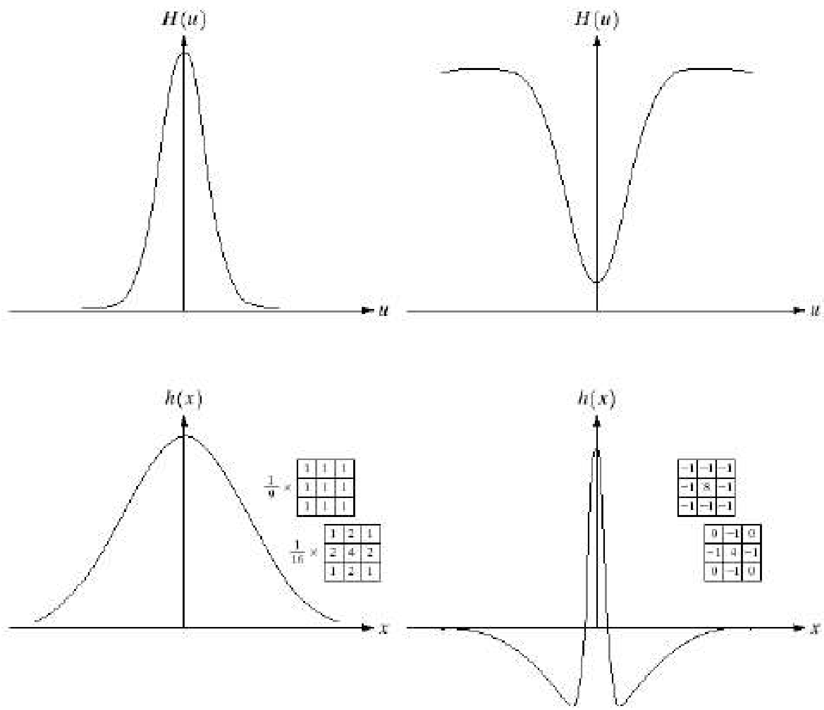
Low-pass

High-pass

Band-pass

Band-stop

Corresponding Filters in the Spatial and Frequency Domains



a b
c d

FIGURE 4.9 (a) Gaussian frequency domain lowpass filter. (b) Gaussian frequency domain highpass filter. (c) Corresponding lowpass spatial filter. (d) Corresponding highpass spatial filter. The masks shown are used in Chapter 3 for lowpass and highpass filtering.

H(u,v) - Ideal Low Pass Filter

b) Illustrate the concept of Geometric Transformation.

<u>GEOMETRIC TRANSFORMATIONS</u>

In terms of digital image processing, a geometric transformation consists of two basic operations: (1) a *spatial transformation,* which defines the "rearrangement" of pixels on the image plane; and (2) *gray-level interpolation,* which deals with the assignment of gray levels to pixels in the spatially transformed image

5.11.1 Spatial Transformations

Suppose that an image f with pixel coordinates (x, y) undergoes geometric distortion to produce an image g with coordinates (x', y'). This transformation may be expressed as

$$x' = r(x, y) \tag{5.11-1}$$

and

$$y' = s(x, y) \tag{5.11-2}$$

where $r(x, y)$ and $s(x, y)$ are the spatial transformations that produced the geometrically distorted image $g(x', y')$. For example, if $r(x, y) = x/2$ and $s(x, y) = y/2$, the "distortion" is simply a shrinking of the size of $f(x, y)$ by one-half in both spatial directions.

If $r(x, y)$ and $s(x, y)$ were known analytically, recovering $f(x, y)$ from the distorted image $g(x', y')$ by applying the transformations in reverse might be possible theoretically. In practice, however, formulating a single set of analytical functions $r(x, y)$ and $s(x, y)$ that describe the geometric distortion process over the entire image plane generally is not possible. The method used most frequently to overcome this difficulty is to formulate the spatial relocation of pixels by the use of *tiepoints*, which are a subset of pixels whose location in the input (distorted) and output (corrected) images is known precisely.

Figure 5.32 shows quadrilateral regions in a distorted and corresponding corrected image. The vertices of the quadrilaterals are corresponding tiepoints.

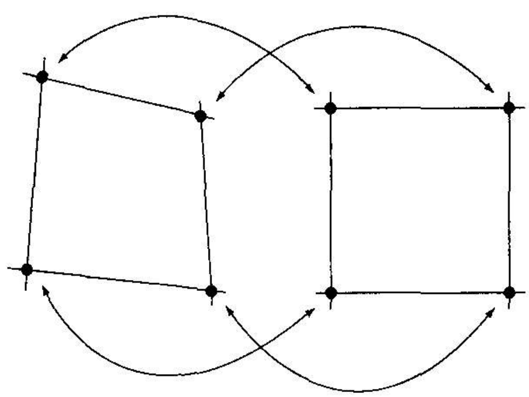


FIGURE 5.32 Corresponding tiepoints in two image segments.

Gray-Level Interpolation

The method discussed in the preceding section steps through integer values of the coordinates (x, y) to yield the restored image $\hat{f}(x, y)$. However, depending on the values of the coefficients c_i, Eqs. (5.11-5) and (5.11-6) can yield noninteger val-

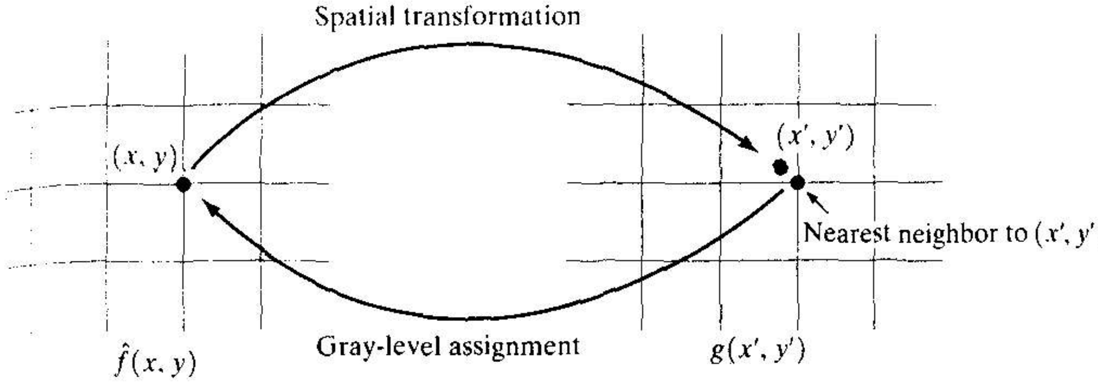


FIGURE 5.33 Gray-level interpolation based on the nearest neighbor concept.

ues for x' and y'. Because the distorted image g is digital, its pixel values are defined only at integer coordinates. Thus using noninteger values for x' and y' causes a mapping into locations of g for which no gray levels are defined. Inferring what the gray-level values at those locations should be, based only on the pixel values at integer coordinate locations, then becomes necessary. The technique used to accomplish this is called *gray-level interpolation*.

- Geometric transforms permit the elimination of geometric distortion that occurs when an image is captured.
- An example is an attempt to match remotely sensed images of the same area taken after one year, when the more recent image was probably not taken from precisely the same position.
- To inspect changes over the year, it is necessary first to execute a geometric transformation, and then subtract one image from the other.

- To inspect changes over the year, it is necessary first to execute a geometric transformation, and then subtract one image from the other.

- A geometric transform is a vector function T that maps the pixel (x,y) to a new position (x',y').

$$X' = T_x(x,y) \text{ , } y' = T_x(x,y)$$

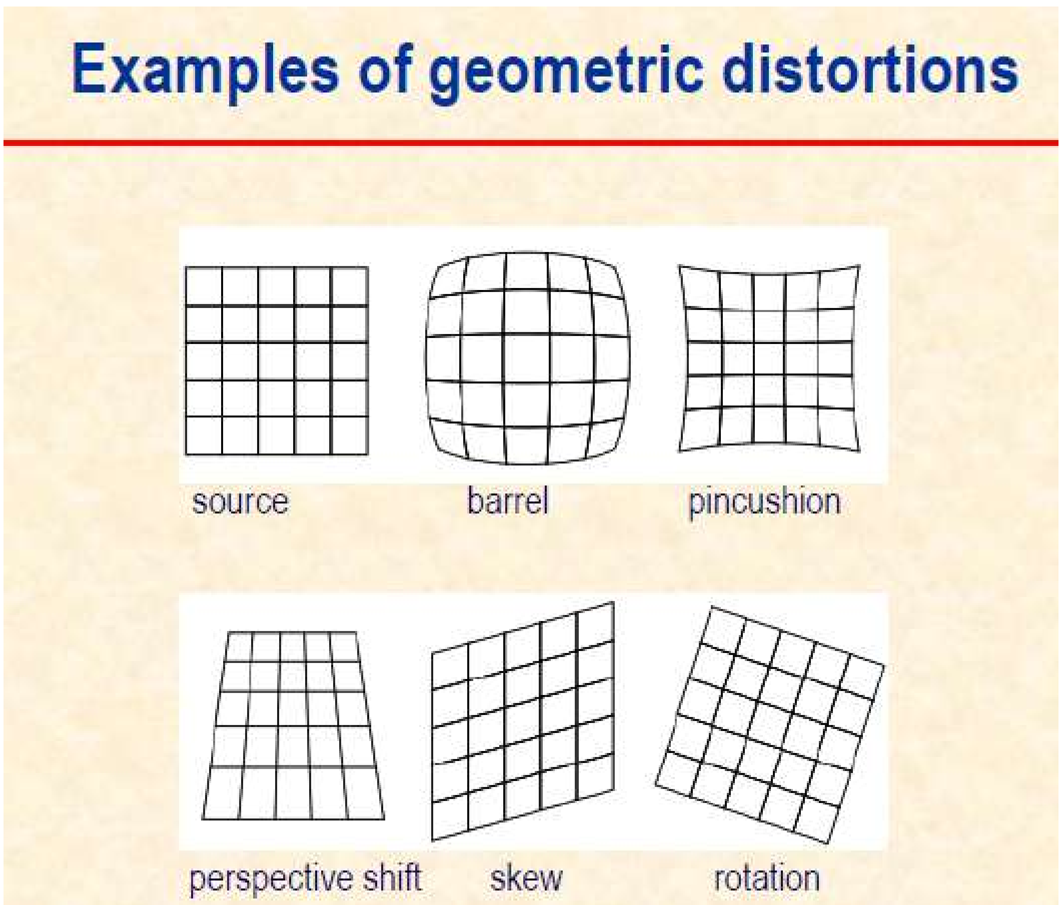
- The transformation equations are either known in advance or can be determined from known original and transformed images.
- Several pixels in both images with known correspondence are used to derive the unknown transformation.

- A geometric transform consists of two basic steps ...

- 1. determining the pixel co-ordinate transformation

- mapping of the co-ordinates of the input image pixel to the point in the output image.
    - the output point co-ordinates should be computed as continuous values (real numbers) as the position does not necessarily match the digital grid after the transform.

- 2. finding the point in the digital raster which matches the transformed point and determining its brightness.
- brightness is usually computed as an interpolation of the brightnesses of several points in the neighborhood.



25. a) Describe about Image restoration techniques.

<u>IMAGE RESTORATION TECHNIQUES</u>

Image restoration is concerned with the reconstruction or estimation of the uncorrupted image from a blurred and noisy one. So, image restoration techniques are mainly categorized for three purposes such as

- restoration from noise,
- linear degradation and
- geometric distortion .

Usage of image restoration techniques has improved enormously since the beginning of the digital image restoration era. Now, different techniques are developed for restoration as shown in figure

NOISE MODELS

In order to restore an image we need to know about the degradation functions. Different models for the noise are described in this section. The set of noise models are defined by specific probability density functions (PDFs). Some commonly found noise models and their corresponding PDFs are given below.

common noise models

11

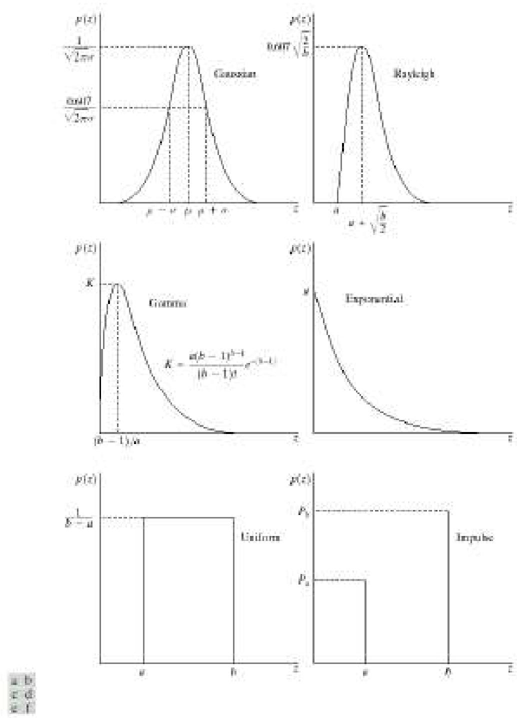


FIGURE 5.2 Some important probability density functions.

Gaussian

$$p(z)=\frac{1}{\sqrt{2\pi}\sigma}e^{-(z-\mu)^2/2\sigma^2}$$

Rayleigh

$$p(z)=\frac{2}{b}(z-a)e^{-(z-a)^2/b}, \text{ for } \quad z\geq a$$

Erlang, Gamma(a, b)

$$p(z)=\frac{a^b z^{b-1}}{(b-a)!}e^{-az}, \text{ for } \quad z\geq 0$$

Exponential

$$p(z)=ae^{-az}, \text{ for } \quad z\geq 0$$

→ additive noise

Salt-and-Pepper:

$$p(z)=P_a\delta(z-a)+P_b\delta(z-b)$$

Speckle noise: $a=a_R+ja_I$

$$|g(x,y)|^2\simeq|f(x,y)|^2|a(x,y)|^2+\eta(x,y)$$

a_R, a_I zero mean, independent Gaussian

→ multiplicative noise on signal magnitude

b) Briefly explain about Linear, Position invariant degradation.

Linear, Position-Invariant Degradations

The input-output relationship in Fig. 5.1 before the restoration stage is expressed as

$$g(x, y) = H[f(x, y)] + \eta(x, y). \tag{5.5-1}$$

For the moment, let us assume that $\eta(x, y) = 0$ so that $g(x, y) = H[f(x, y)]$. Based on the discussion in Section 2.6, H is *linear* if

$$H[af_1(x, y) + bf_2(x, y)] = aH[f_1(x, y)] + bH[f_2(x, y)] \tag{5.5-2}$$

where a and b are scalars and $f_1(x, y)$ and $f_2(x, y)$ are any two input images.

If $a = b = 1$, Eq. (5.5-2) becomes

$$H[f_1(x, y) + f_2(x, y)] = H[f_1(x, y)] + H[f_2(x, y)] \tag{5.5-3}$$

which is called the property of *additivity*. This property simply says that, if H is a linear operator, the response to a sum of two inputs is equal to the sum of the two responses.

With $f_2(x, y) = 0$, Eq. (5.5-2) becomes

$$H[af_1(x, y)] = aH[f_1(x, y)] \tag{5.5-4}$$

which is called the property of *homogeneity*. It says that the response to a constant multiple of any input is equal to the response to that input multiplied by the same constant. Thus a linear operator possesses both the property of additivity and the property of homogeneity.

An operator having the input-output relationship $g(x, y) = H[f(x, y)]$ is said to be *position* (or *space*) *invariant* if

$$H[f(x - \alpha, y - \beta)] = g(x - \alpha, y - \beta) \tag{5.5-5}$$

for any $f(x, y)$ and any α and β. This definition indicates that the response at any point in the image depends only on the *value* of the input at that point, not on its *position*.

With a slight (but equivalent) change in notation in the definition of the discrete impulse function in Eq. (4.2-33), $f(x, y)$ can be expressed in terms of a continuous impulse function:

$$f(x, y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\alpha, \beta)\delta(x - \alpha, y - \beta)\, d\alpha\, d\beta. \tag{5.5-6}$$

This, in fact, is the *definition* using continuous variables of a unit impulse located at coordinates (x, y).

Assume again for a moment that $\eta(x, y) = 0$. Then, substitution of Eq. (5.5-6) into Eq. (5.5-1) results in the expression

$$g(x, y) = H[f(x, y)] = H\left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\alpha, \beta)\delta(x - \alpha, y - \beta)\, d\alpha\, d\beta\right]. \tag{5.5-7}$$

If H is a linear operator and we extend the additivity property to integrals, then

$$g(x, y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} H[f(\alpha, \beta)\delta(x - \alpha, y - \beta)]\, d\alpha\, d\beta. \tag{5.5-8}$$

Because $f(\alpha, \beta)$ is independent of x and y, and using the homogeneity property, it follows that

$$g(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\alpha, \beta) H[\delta(x - \alpha, y - \beta)] \, d\alpha \, d\beta. \tag{5.5-9}$$

The term

$$h(x, \alpha, y, \beta) = H[\delta(x - \alpha, y - \beta)] \tag{5.5-10}$$

is called the *impulse response* of H. In other words, if $\eta(x, y) = 0$ in Eq. (5.5-1), then $h(x, \alpha, y, \beta)$ is the response of H to an impulse of strength 1 at coordinates (x, y). In optics, the impulse becomes a point of light and $h(x, \alpha, y, \beta)$ is commonly referred to as the *point spread function* (PSF). This name arises from the fact that all physical optical systems blur (spread) a point of light to some degree, with the amount of blurring being determined by the quality of the optical components.

26. a) Discuss about Degradation functions.

ESTIMATING THE DEGRADATION FUNCTION

The process of restoring an image by using a degradation function that has been estimated in some way sometimes is called *blind deconvolution,*

- **Estimation by Image observation**
- **Estimation by experimentation**
- **Estimation by modeling**

Estimation by Image Observation

Suppose that we are given a degraded image without any knowledge about the degradation function *H.* One way to estimate this function is to gather information from the image itself. For example, if the image is blurred, we can look at a small section of the image containing simple structures, like part of an object and the background. In order to reduce the effect of noise in our observation, we would look for areas of strong signal content. Using sample gray levels of the object and background, we can construct an unblurred image of the same size and characteristics as the observed subimage. Let the observed subimage be denoted by gJx, *y).* and let the constructed subimage which in reality is our estimate of the original image in that area) be denoted by *I,(x, y).* Then, assuming that the effect of noise is negligible because of our choice of a strong-signal area, it follows from Eq. (5.5-17) that

$$H_s(u, v) = \frac{G_s(u, v)}{\hat{F}_s(u, v)}. \tag{5.6-1}$$

From the characteristics of this function we then deduce the complete function $H(u, v)$ by making use of the fact that we are assuming position invariance. For example, suppose that a radial plot of $H_s(u, v)$ turns out to have the shape of Butterworth lowpass filter. We can use that information to construct a function $H(u, v)$ on a larger scale, but having the same shape.

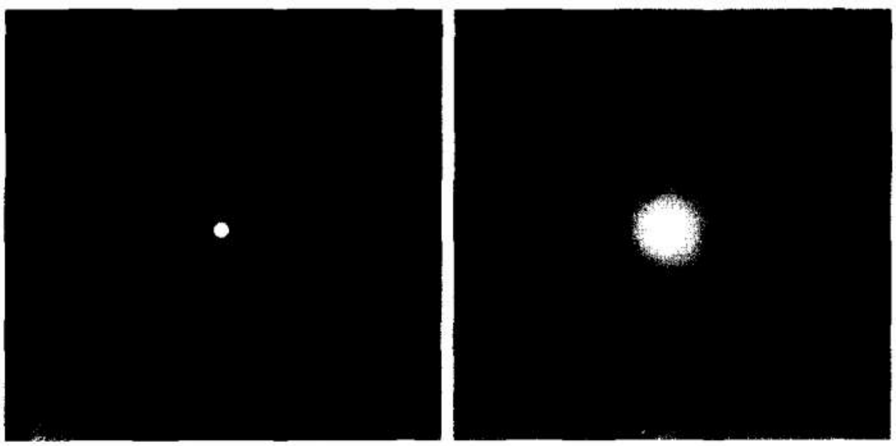
5.6.2 Estimation by Experimentation

If equipment similar to the equipment used to acquire the degraded image is available, it is possible in principle to obtain an accurate estimate of the degradation. Images similar to the degraded image can be acquired with various system settings until they are degraded as closely as possible to the image we wish to restore. Then the idea is to obtain the impulse response of the degradation by imaging an impulse (small dot of light) using the same system settings. As noted in Section 5.5, a linear, space-invariant system is described completely by its impulse response.

An impulse is simulated by a bright dot of light, as bright as possible to reduce the effect of noise. Then, recalling that the Fourier transform of an impulse is a constant, it follows from Eq. (5.5-17) that

$$H(u, v) = \frac{G(u, v)}{A} \tag{5.6-2}$$

where, as before, $G(u, v)$ is the Fourier transform of the observed image and A is a constant describing the strength of the impulse. Figure 5.24 shows an example.



a b

FIGURE 5.24 Degradation estimation by impulse characterization. (a) An impulse of light (shown magnified). (b) Imaged (degraded) impulse.

5.6.3 Estimation by Modeling

Degradation modeling has been used for many years because of the insight it affords into the image restoration problem. In some cases, the model can even take into account environmental conditions that cause degradations. For example, a degradation model proposed by Hufnagel and Stanley [1964] is based on the physical characteristics of atmospheric turbulence. This model has a familiar form:

$$H(u, v) = e^{-k(u^2+v^2)^{5/6}} \tag{5.6-3}$$

where k is a constant that depends on the nature of the turbulence. With the exception of the 5/6 power on the exponent, this equation has the same form as the Gaussian lowpass filter discussed in Section 4.3.3. In fact, the Gaussian LPF is used sometimes to model mild, uniform blurring. Figure 5.25 shows examples obtained by simulating blurring an image using Eq. (5.6-3) with values $k = 0.0025$ (severe turbulence in this case), $k = 0.001$ (mild turbu-

b)Explain in detail about Image Compression.

IMAGE COMPRESSION:

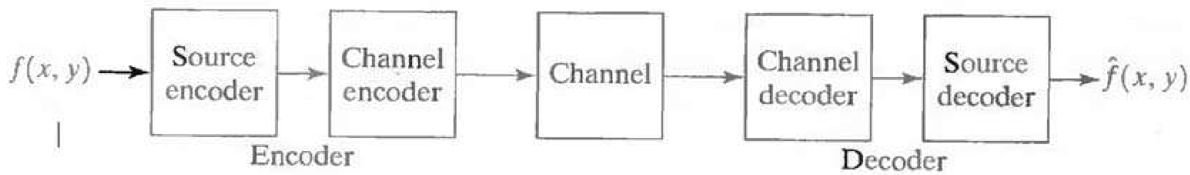
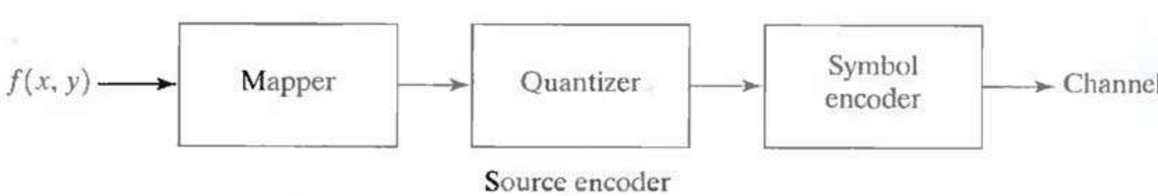


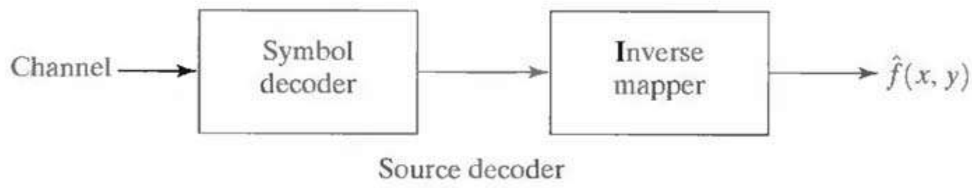
Fig.3.1 A general compression system model

Fig. shows, a compression system consists of two distinct structural blocks: an encoder and a decoder. An input image f(x, y) is fed into the encoder, which creates a set of symbols from the input data. After transmission over the channel, the encoded representation is fed to the decoder, where a reconstructed output image f^(x, y) is generated. In general, f^(x, y) may or may not be an exact replica of f(x, y). If it is, the system is error free or information preserving; if not, some level of distortion is present in the reconstructed image. Both the encoder and decoder shown in Fig. 3.1 consist of two relatively independent functions or subblocks. The encoder is made up of a source encoder, which removes input redundancies, and a channel encoder, which increases the noise immunity of the source encoder's output. As would be expected, the decoder includes a channel decoder followed by a source decoder. If the channel between the encoder and decoder is noise free (not prone to error), the channel encoder and decoder are omitted, and the general encoder and decoder become the source encoder and decoder, respectively.

THE SOURCE ENCODER AND DECODER:

The source encoder is responsible for reducing or eliminating any coding, interpixel, or psychovisual redundancies in the input image. The specific application and associated fidelity requirements dictate the best encoding approach to use in any given situation. Normally, the approach can be modeled by a series of three independent operations. As Fig. 3.2 (a) shows, each operation is designed to reduce one of the three redundancies. Figure 3.2 (b) depicts the corresponding source decoder. In the first stage of the source encoding process, the mapper transforms the input data into a (usually nonvisual) format designed to reduce interpixel redundancies in the input image. This operation generally is reversible and may or may not reduce directly the amount of data required to represent the image.





a
b

Fig.3.2 (a) Source encoder and (b) source decoder model

Run-length coding is an example of a mapping that directly results in data compression in this initial stage of the overall source encoding process. The representation of an image by a set of transform coefficients is an example of the opposite case. Here, the mapper transforms the image into an array of coefficients, making its interpixel redundancies more accessible for compression in later stages of the encoding process.

The second stage, or quantizer block in Fig. 3.2 (a), reduces the accuracy of the mapper's output in accordance with some preestablished fidelity criterion. This stage reduces the psychovisual redundancies of the input image. This operation is irreversible. Thus it must be omitted when error-free compression is desired.

In the third and final stage of the source encoding process, the symbol coder creates a fixed- or variable-length code to represent the quantizer output and maps the output in accordance with the code. The term symbol coder distinguishes this coding operation from the overall source encoding process. In most cases, a variable-length code is used to represent the mapped and quantized data set.

It assigns the shortest code words to the most frequently occurring output values and thus reduces coding redundancy. The operation, of course, is reversible. Upon completion of the symbol coding step, the input image has been processed to remove each of the three redundancies.

Figure 3.2(a) shows the source encoding process as three successive operations, but all three operations are not necessarily included in every compression system. Recall, for example, that the quantizer must be omitted when error-free compression is desired. In addition, some compression techniques normally are modeled by merging blocks that are physically separate in Fig. 3.2(a).
In the predictive compression systems, for instance, the mapper and quantizer are often represented by a single block, which simultaneously performs both operations. The source decoder shown in Fig. 3.2(b) contains only two components: a symbol decoder and an inverse mapper. These blocks perform, in reverse order, the inverse operations of the source encoder's symbol encoder and mapper blocks. Because quantization results in irreversible information loss, an inverse quantizer block is not included in the general source decoder model shown in Fig. 3.2(b).

Reg.No. ______________

[16ITU504A]

KARPAGAM ACADEMY OF HIGHER EDUCATION
(Deemed to be University)
(Established Under Section 3 of UGC Act 1956)
Coimbatore – 641 021

INFORMATION TECHNOLOGY

Fifth Semester
THIRD INTERNAL EXAMINATION – OCTOBER 2018

DIGITAL IMAGE PROCESSING

Class: III B.Sc. IT (A&B) **Duration :** 2 Hours
Date & Session : 9.10.2018 & AN **Maximum :** 50 Marks

PART-A (20 X 1 = 20 Marks)
Answer ALL the Questions

1. The ___________ is an encoding format used for fax transmission.
   a) Fax Group b) ITU-T c) JPEG d) TIFF
2. The overall process of truncating, quantizing, and coding the coefficients of transformed subimage is commonly called ___________ coding.
   a) bit-plane b) zonal c) threshold d) bit allocation
3. The _________ consume large amount of data and requires very high bandwidth networks in transmission.
   a) video files b) email c) file attachment d) text
4. The ___________supports one-dimensional image compression of black and white images, on a standard fax machine.
   a) Fax Group 5 b) Fax Group 4 c) Fax Group 3 d) Fax Group 6
5. The ________________ supports two-dimensional image compression, compressing the line width as well as the line length.
   a) Fax Group 5 b) Fax Group 4 c) Fax Group 3 d) Fax Group 6
6. Subspaces spanned are nested at ____________.
   a) lower scales b) higher scales c) mid scales d) intense scales
7. Prediction residual pyramid is computed in ____________.
   a) 2 steps b) 3 steps c) 4 steps d) 5 steps
8. K multiplication constants in digital filters are called ____________.
   a) co efficient b) multipliers c) subtractors d) filter coefficients
9. The size of the base image will be ____________.
   a) N-1 x N-1 b) N+1 x N-1 c) N-1 x N d) N x N

10. Narrow wavelets represents ___________.
a) sharp details b) finer details c) blur details d) edge details
11. High contrast images are considered as ___________.
a) low resolution b) high resolution c) intense blurred d) low resolution
12. Function having compact support is ___________.
a) histogram b) pyramids c) mean pyramids d) haar function
13. Closing is represented by __________.
a) A .B b) A+B c) A-B d) AxB
14. Closing is used for __________.
a) separation b) compression c) decompression d) filling holes
15. Opening with rolling SE __________.
a) sharps b) shrinks c) smooths d) deletes
16. Structuring element is also called __________.
a) pixels b) lines c) subimage d) noise
17. With dilation process images get __________.
a) thinner b) shrinked c) thickened d) sharpened
18. Opening and closing are each others __________.
a) neighbors b) duals c) centers d) corners
19. Fully containment of the SE in an image is required in __________.
a) erosion b) dilation c) opening d) closing
20. Erosion followed by dilation is called __________.
a) opening b) closing c) blurring d) translation

PART-B (3 X 2 = 6 Marks)
(Answer ALL the Questions)

21. Define G3 and G4 of Fax Compression.
22. Write short notes on Lossless Predictive coding.
23. Define Rules for Dilation and Erosion.

PART-C (3 X 8 = 24 Marks)
(Answer ALL the Questions)

24. a) Discuss about Wavelet based image compression.
(or)
b) Explain about FAX compression.

25. a) Illustrate about the concept of Scaling function.
(or)
b) Briefly explain about Image Watermarking.

26. a) Discuss about Edge Detection.
(or)
b) Explain in detail about Thinning and Thickening of Image Processing.

Reg.No. ______________

[16ITU504A]

KARPAGAM ACADEMY OF HIGHER EDUCATION
(Deemed to be University)
(Established Under Section 3 of UGC Act 1956)
Coimbatore – 641 021

INFORMATION TECHNOLOGY

Fifth Semester
THIRD INTERNAL EXAMINATION – OCTOBER 2018

DIGITAL IMAGE PROCESSING

Class: III B.Sc. IT (A&B) **Duration :** 2 Hours
Date & Session : 9.10.2018 & AN **Maximum :** 50 Marks

PART-A (20 X 1 = 20 Marks)
Answer ALL the Questions

1. The ___________ is an encoding format used for fax transmission.
   a) Fax Group b) ITU-T c) JPEG d) TIFF
2. The overall process of truncating, quantizing, and coding the coefficients of transformed subimage is commonly called ___________ coding.
   a) bit-plane b) zonal c) threshold **d) bit allocation**
3. The __________ consume large amount of data and requires very high bandwidth networks in transmission.
   a) video files b) email c) file attachment d) text
4. The ___________supports one-dimensional image compression of black and white images, on a standard fax machine.
   a) Fax Group 5 b) Fax Group 4 **c) Fax Group 3** d) Fax Group 6
5. The ________________ supports two-dimensional image compression, compressing the line width as well as the line length.
   a) Fax Group 5 **b) Fax Group 4** c) Fax Group 3 d) Fax Group 6
6. Subspaces spanned are nested at ____________.
   a) lower scales **b) higher scales** c) mid scales d) intense scales
7. Prediction residual pyramid is computed in ____________.
   a) 2 steps **b) 3 steps** c) 4 steps d) 5 steps
8. K multiplication constants in digital filters are called ____________.
   a) co efficient b) multipliers c) subtractors **d) filter coefficients**
9. The size of the base image will be ____________.
   a) N-1 x N-1 b) N+1 x N-1 c) N-1 x N **d) N x N**

10. Narrow wavelets represents ___________.
a) sharp details **b) finer details** c) blur details d) edge details

11. High contrast images are considered as ___________.
a) low resolution b) high resolution c) intense blurred d) low resolution

12. Function having compact support is ___________.
a) histogram b) pyramids c) mean pyramids **d) haar function**

13. Closing is represented by __________.
a) A .B b) A+B c) A-B d) AxB

14. Closing is used for __________.
a) separation b) compression c) decompression **d) filling holes**

15. Opening with rolling SE __________.
a) sharps b) shrinks **c) smooths** d) deletes

16. Structuring element is also called __________.
a) pixels b) lines **c) subimage** d) noise

17. With dilation process images get __________.
a) thinner b) shrinked **c) thickened** d) sharpened

18. Opening and closing are each others __________.
a) neighbors **b) duals** c) centers d) corners

19. Fully containment of the SE in an image is required in __________.
a) erosion b) dilation c) opening d) closing

20. Erosion followed by dilation is called __________.
a) opening b) closing c) blurring d) translation

PART-B (3 X 2 = 6 Marks)
(Answer ALL the Questions)

21. Define G3 and G4 of Fax Compression.

Fax Group 3 supports one-dimensional image compression of black and white images, On a standard fax machine, Fax Group 3 uses redundancy reduction to enhance speed and is able to transmit a page in one minute or less. Fax Group 3 can achieve compression ratios of 10:1 for office documents and 15:1 for engineering drawings with a resolution of 200 dots per inch (dpi).

Less frequently used, Fax Group 4 (G4) is ITU-T Recommendation T.6 and supports two-dimensional image compression, compressing the line width as well as the line length. Fax Group 4 can achieve compression ratios of 15:1 for office documents and 20:1 for engineering drawings with a resolution of 400 dpi. Unlike Fax Group 3, Fax Group 4 can use Integrated Services Digital Network (ISDN) for transmission.

22. Write short notes on Lossless Predictive coding.

LOSSLESS PREDICTIVE CODING

Lossless compression methods are needed in some digital imaging applications, such as: medical images, x-ray images, etc. An error-free compression approach that does not require decomposition of an image into a collection of bit planes The approach, commonly referred to as *lossless predictive coding,* IS based on eliminating the interpixel redundancies of closely spaced pixels by extracting and coding *only* the new information in each pixel The *new information* of a pixel is defined as the difference between the actual and predicted value of that pixel .The basic components of a lossless predictive coding system. The system consists of an encoder and a decoder, each containing an identical *predictor.* As each successive pixel of the input image, denoted *f",* is introduced to the encoder, the predictor generates the anticipated value of that pixel based on some number of past inputs. The output of the predictor IS then rounded to the nearest integer, denoted *fn'* and used to form the difference' or *prediction error.*

$$e_n = f_n - \hat{f}_n,$$

23. Define Rules for Dilation and Erosion.

Dilation	The value of the output pixel is the *maximum* value of all the pixels in the input pixel's neighborhood. In a binary image, if any of the pixels is set to the value1, the output pixel is set to 1.
Erosion	The value of the output pixel is the *minimum* value of all the pixels in the input pixel's neighborhood. In a binary image, if any of the pixels is set to 0, the output pixel is set to 0.

PART-C (3 X 8 = 24 Marks)
(Answer ALL the Questions)

24. a) Discuss about Wavelet based image compression.

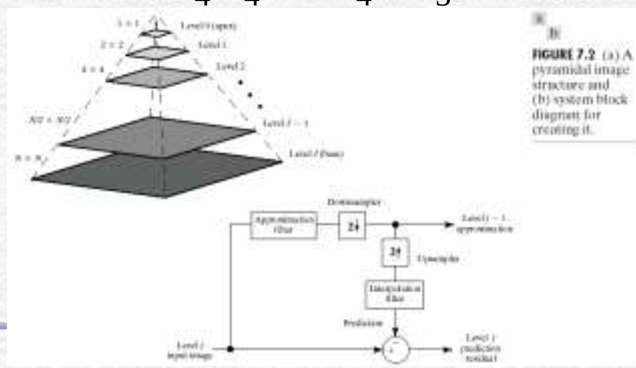
WAVELET BASED IMAGE COMPRESSION:

Unlike the Fourier transform, whose basis functions are sinusoids, wavelet transforms are based on small waves, called *wavelets,* of varying frequency *and limited duration.* This allows them to provide the equivalent of a musical score for an image, revealing not only what notes (or frequencies) to play but also when to play them.

Image Pyramids

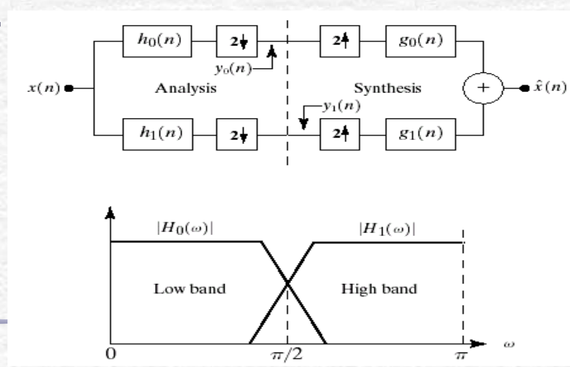
Total number of elements in a P+1 level pyramid for P>0 is

$$N^2\left(1+\frac{1}{4}+\frac{1}{4^2}+\ldots+\frac{1}{4^P}\right) \le \frac{4}{3}N^2$$



Subband Coding

- An image is decomposed into a set of band-limited components, called subbands, which can be reassembled to reconstruct the original image without error.



The Haar Transform

- Oldest and simplest known orthonormal wavelets.
- T=HFH where

  F: NXN image matrix,

  H: NxN transformation matrix.
- Haar basis functions $h_k(z)$ are defined over the continuous, closed interval [0,1] for k=0,1,..N-1 where $N=2^n$.

Haar Basis Functions

$$k = 2^p + q - 1 \quad \text{where } 0 \le p \le n-1,$$
$$q = 0 \text{ or } 1 \text{ for } p = 0, 1 \le q \le 2^p \text{ for } p \ne 0$$

$$h_0(z) = h_{00}(z) = \frac{1}{\sqrt{N}}, \quad z \in [0,1]$$

$$h_k(z) = h_{pq}(z) = \frac{1}{\sqrt{N}} \begin{cases} 2^{p/2} & (q-1)/2^p \le z < (q-0.5)/2^p \\ -2^{p/2} & (q-0.5)/2^p \le z < q/2^p \\ 0 & \text{otherwise}, z \in [0,1] \end{cases}$$

Multiresolution Expansions

- Multiresolution analysis (MRA)
- A scaling function is used to create a series of approximations of a function or image, each differing by a factor of 2.
- Additional functions, called wavelets, are used to encode the difference in information between adjacent approximations.

Series Expansions

- A signal f(x) can be expressed as a linear combination of expansion functions:

$$f(x) = \sum_k \alpha_k \varphi_k(x)$$

  - Case 1: orthonormal basis: $\langle \varphi_j(x), \varphi_k(x) \rangle = \delta_{jk}$
  - Case 2: orthogonal basis: $\langle \varphi_j(x), \varphi_k(x) \rangle = 0 \quad j \neq k$
  - Case 3: frame: $A\|f(x)\|^2 \leq \sum_k |\langle \varphi_k(x), f(x) \rangle|^2 \leq B\|f(x)\|^2$

Scaling Functions

- Consider the set of expansion functions composed of integer translations and binary scaling of the real, square-integrable function, $\varphi(x)$,i.e.,

$$\varphi_{j,k}(x) = 2^{j/2}\varphi(2^j x - k)$$

- By choosing φ wisely, $\{\varphi_{j,k}(x)\}$ can be made to span $L^2(\mathbf{R})$

Haar Scaling Function

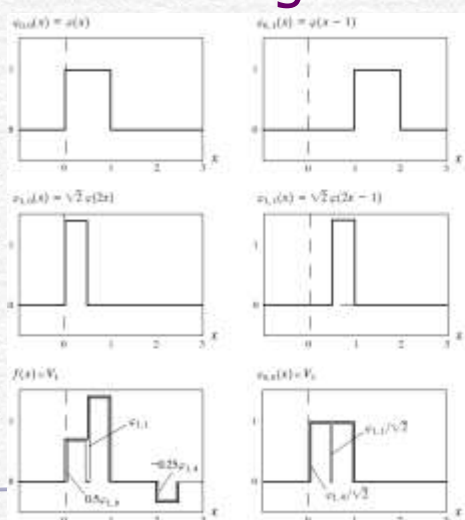


FIGURE 7.9 Haar scaling functions in V_0 in V_1.

(or)

b) Explain about FAX compression.

FAX COMPRESSION (CSUITT GROUP-3 AND GROUP-4)

Fax Group is an encoding format used for fax transmission. There are two types: Fax Group 3, also known as G3, and Fax Group 4, also known as G4. Fax Group 3 and 4 are two of the encoding formats for Tagged Image File Format (TIFF) files. The more commonly used format, Fax Group 3, is Recommendation T.4 of the CCIT , now known as the ITU-T (for Telecommunication Standardization Sector of the International Telecommunications Union).

Fax Group 3 supports one-dimensional image compression of black and white images, On a standard fax machine, Fax Group 3 uses redundancy reduction to enhance speed and is able to transmit a page in one minute or less. Fax Group 3 can achieve compression ratios of 10:1 for office documents and 15:1 for engineering drawings with a resolution of 200 dots per inch (dpi).

Less frequently used, Fax Group 4 (G4) is ITU-T Recommendation T.6 and supports two-dimensional image compression, compressing the line width as well as the line length. Fax Group 4 can achieve compression ratios of 15:1 for office documents and 20:1 for engineering drawings with a resolution of 400 dpi. Unlike Fax Group 3, Fax Group 4 can use Integrated Services Digital Network (ISDN) for transmission.

CCITT Group 4 compression, also referred to as **G4** or **Modified Modified READ** (MMR), is a lossless method of image compression used in Group 4 fax machines defined in the ITU-T T.6 fax standard. It is only used for bitonal (black and white) images. Group 4 compression is based on the Group 3 two-dimensional compression scheme (G3-2D), also known as Modified READ, which is in turn based on the Group 3 one-dimensional compression scheme (G3), also known as Modified Huffman coding. Group 4 compression is available in many proprietary image file formats as well as standardized formats such as TIFF, CALS, CIT (Intergraph Raster Type 24) and the PDF document format.

G4 offers a small improvement over G3-2D by removing the end of line (EOL) codes. G3 and G4 compression both treat an image as a series of horizontal black strips on a white page. Better compression is achieved when there are fewer unique black dots/lines on the page. Both G3-2D and G4 add a two dimensional feature to achieve greater compression by taking advantage of vertical symmetry. A worst-case image would be an alternating pattern of single-pixel black and white dots offset by one pixel on even/odd lines. G4 compression would actually increase the file size on this type of image. G4 typically achieves a 20:1 compression ratio. For an 8.5"×11" page scanned at 200 DPI, this equates to a reduction from 467.5 kB to 23.4 kB (95% compression ratio).

Group 3 & Group 4

GROUP 3

•MODIFIED HUFFMAN METHOD (MHM) – Unidimensional coding method based on the coding of the length of alternate black and white pixel runs using Huffman coding.

GROUP 4 (Also Group 3 Options)

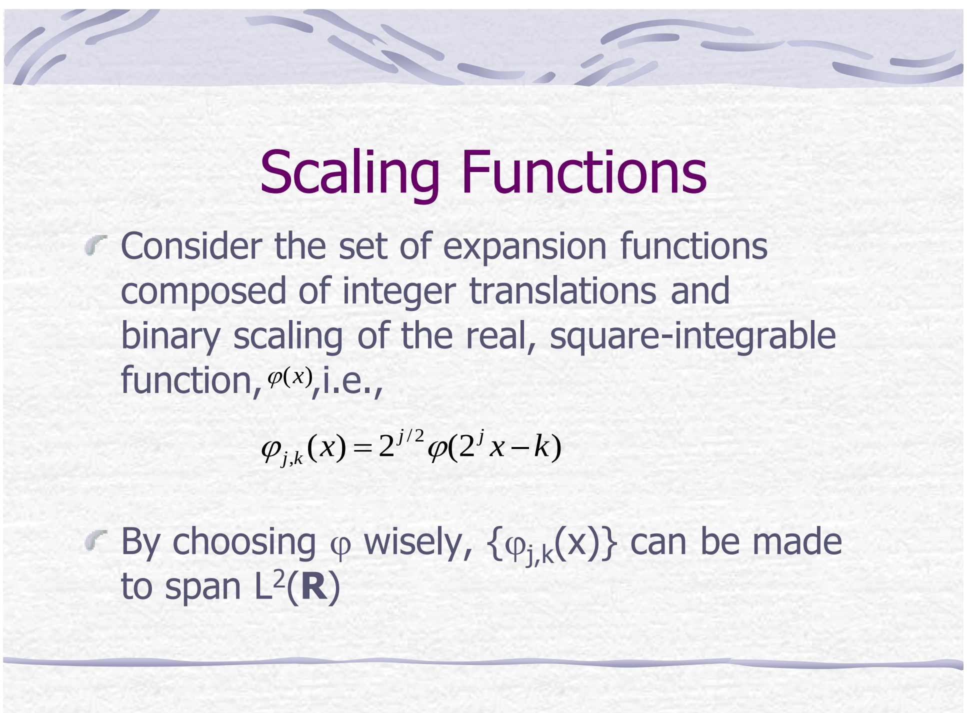
•MODIFIED READ METHOD (MRM) – Bidimensional coding method based on the coding of the variations of the positions of tone transition pixels (black-white or white-black) in relation to the previous line; unidimensional coding may be used every k lines.

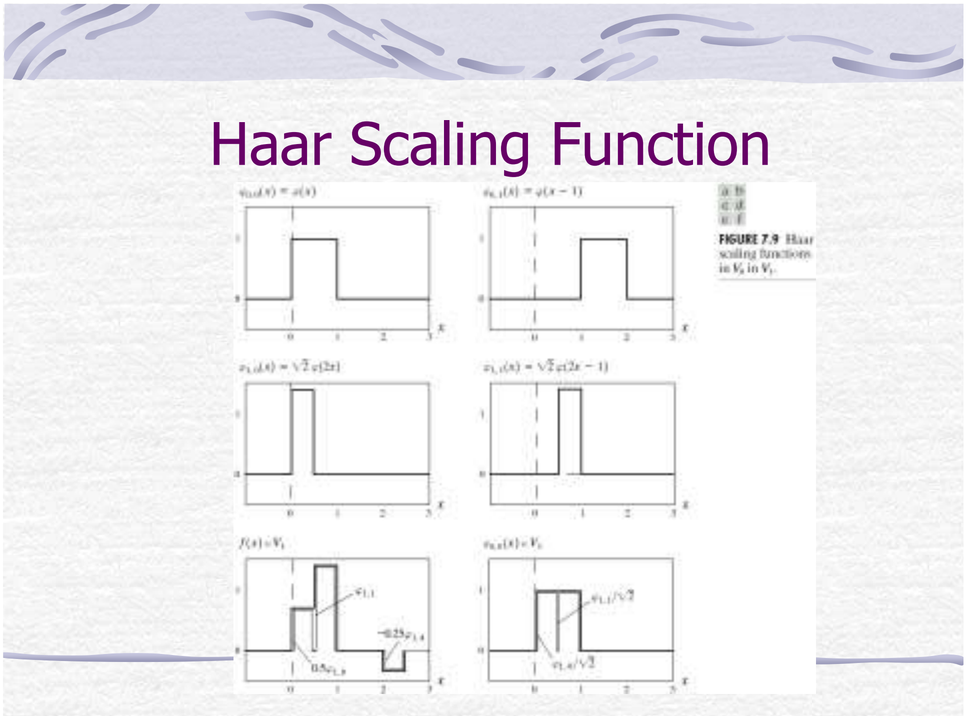
•MODIFIED-MODIFIED READ METHOD (MMRM) – Similar to MRM but without periodic unidimensional coding.

25. a) Illustrate about the concept of Scaling function.

Image pyramids, subband coding and the Haar transform play an important role in a mathematical framework called multiresolution analysis (MRA). In MRA, a **scaling function** is used to create a series of approximations of a signal each differing a factor of 2 in resolution from its nearest neighbour approximation.

Wavelets are **functions** which are form by two resulting coefficients. The detail coefficient and the scale coefficient. From wikipedia **Wavelets** are defined by the**wavelet function** ψ(t) (i.e. the mother **wavelet**) and **scaling function** φ(t) (also called father **wavelet**) in the time domain.





(or)

b) Briefly explain about Image Watermarking.

<u>DIGITAL IMAGE WATERMARKING</u>

- Watermark--an invisible signature embedded inside an image to show authenticity or proof of ownership
- Discourage unauthorized copying and distribution of images over the internet
- Ensure a digital picture has not been altered
- Software can be used to search for a specific watermark

Types of Watermarking:

The digital watermarking is of four types and they are as follows:

- Visible
- Invisible
- Fragile
- Public

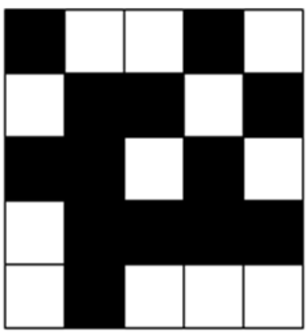
1. **Visible watermarking:** the visible watermarking have the visible message or the logo of the company which represents the ownership of the message, removing of the logo and meddling of the logo breaks the law of copyright.
2. **Invisible watermarking:** the invisible watermarking cannot be seen on the original images and the picture looks like an original image though it has the watermark. The invisible watermark can be extracted with the detection procedure or with any watermark extraction method.
3. **Fragile watermark:** the fragile watermarks can be demolished by the data manipulation and these are also called as tamper proof watermark. The fragile watermark has the

ability to detect the modifications in the signal and also recognizes the place where the modifications have occurred and also the signal before the change.

4. **Public watermark:** The public watermark does not have the protection and these can be read by everyone by availing the unique algorithm.

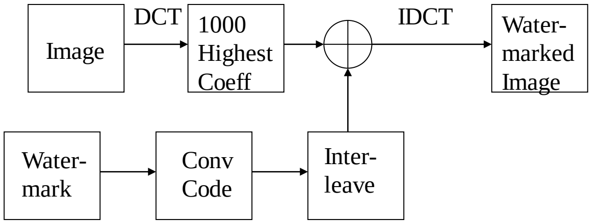
Watermark Properties

- Watermark should appear random, noise-like sequence
- Appear Undetectable
- Good Correlation Properties
  High correlation with signals similar to watermark
  Low correlation with other watermarks or random noise
- Common sequences
  A) Normal distribution
  B) m-sequences



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

Setup-Watermark Embedding



•DC Component Excluded for 1000 Highest Coefficients
•Interleaving prevents burst errors
•Watermarked Image Similar to original image
•Without coding, ignore Conv Code and Interleave block

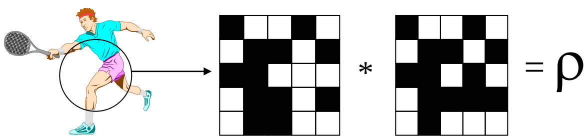
Original Image

Watermarked Image, No Coding

Watermarked Image with Coding

•512x512 “Mandrill” Image
•See Handout
•Both watermarks imperceptible
•Alterations to original image difficult to notice

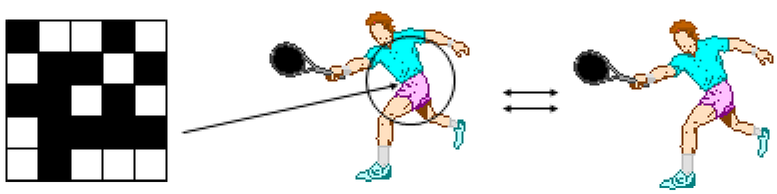
Watermark Detection



Suspected Image | Extracted Watermark | Original Watermark | Correlation

•Watermark Extracted from Suspected Image
•Compute correlation of Extracted and Original Watermark
•Threshold correlation to determine watermark existence

Watermark Embedding



Watermark | Original Image | Watermarked image

•Watermark placed into information content of Original Image to create Watermarked Image
•Image Content
Spatial Domain (Least Significant Bit)
FFT - Magnitude and Phase
Wavelet Transforms
DCT Coefficients

26. a) Discuss about Edge Detection.

Edge Detection

The sudden changes of discontinuities in an image are called as edges. Significant transitions in an image are called as edges.

Types of edges

Generally edges are of three types:

- Horizontal edges
- Vertical Edges
- Diagonal Edges

Why detect edges

Most of the shape information of an image is enclosed in edges. So first we detect these edges in an image and by using these filters and then by enhancing those areas of image which contains edges, sharpness of the image will increase and image will become clearer.

Here are some of the masks for edge detection that we will discuss in the upcoming tutorials.

- Prewitt Operator
- Sobel Operator
- Robinson Compass Masks
- Krisch Compass Masks
- Laplacian Operator.

Above mentioned all the filters are Linear filters or smoothing filters.

Prewitt Operator

Prewitt operator is used for detecting edges horizontally and vertically.

Sobel Operator

The sobel operator is very similar to Prewitt operator. It is also a derivate mask and is used for edge detection. It also calculates edges in both horizontal and vertical direction.

Robinson Compass Masks

This operator is also known as direction mask. In this operator we take one mask and rotate it in all the 8 compass major directions to calculate edges of each direction.

Kirsch Compass Masks

Kirsch Compass Mask is also a derivative mask which is used for finding edges. Kirsch mask is also used for calculating edges in all the directions.

Laplacian Operator

Laplacian Operator is also a derivative operator which is used to find edges in an image. Laplacian is a second order derivative mask. It can be further divided into positive laplacian and negative laplacian.

All these masks find edges. Some find horizontally and vertically, some find in one direction only and some find in all the directions. The next concept that comes after this is sharpening which can be done once the edges are extracted from the image

Sharpening

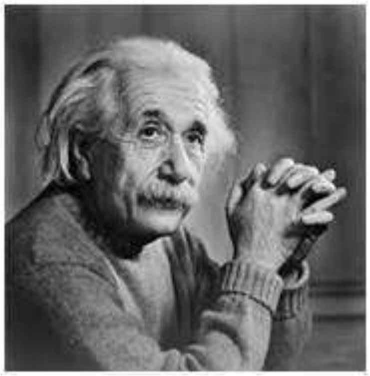
Sharpening is opposite to the blurring. In blurring, we reduce the edge content and in Sharpening, we increase the edge content. So in order to increase the edge content in an image, we have to find edges first.

Edges can be find by one of the any method described above by using any operator. After finding edges, we will add those edges on an image and thus the image would have more edges, and it would look sharpen.

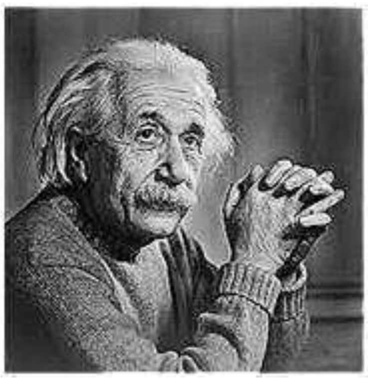
This is one way of sharpening an image.

The sharpen image is shown below.

Original Image



Sharpen Image



(or)

b) Explain in detail about Thinning and Thickening of Image Processing.

Thinning and Thickening

—Thinning is an image-processing operation in which binary valued image regions are reduced to lines

—The purpose of thinning is to reduce the image components to their essential information for further analysis and recognition

—Thickening is changing a pixel from 1 to 0 if any neighbors of the pixel are 1.

—Thickening followed by thinning can be used for filling undesirable holes.

—Thinning followed by thickening is used for determining isolated components and clusters.

Thinning

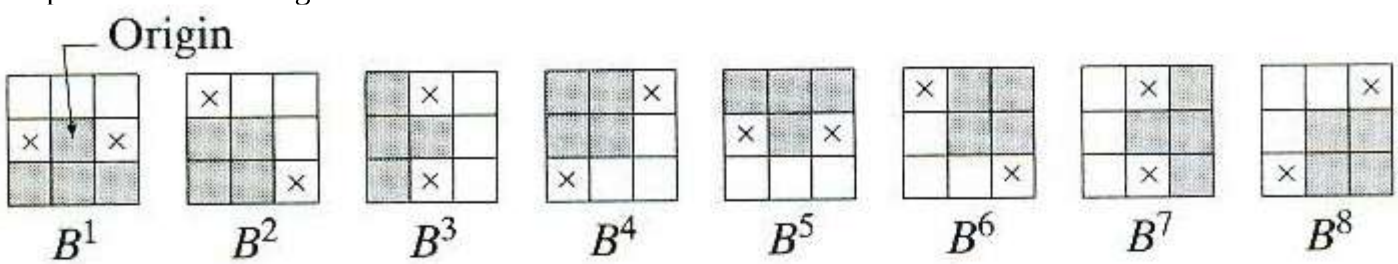
—Thinning is defined in terms of hit or miss as

$$A \otimes B = A - (A \circledast B)$$
$$= A \cap (A \circledast B)^c$$

where B is a sequence of structuring elements like
{B} = {B1, B2, B3, …, Bn} and the operation can be given as

$$A \otimes \{B\} = ((\dots((A \otimes B^1) \otimes B^2) \dots) \otimes B^n)$$

Thinning
—Sample set of structuring elements



Thickening
—Thickening is the morphological dual of thinning and defined as

$$A \odot B = A \cup (A \circledast B)$$

Or

$$A \odot \{B\} = ((\dots((A \odot B^1) \odot B^2) \dots) \odot B^n)$$

Example: