DIGITAL IMAGE PROCESSING

COURSE FILE

IV/IV B.Tech I SEM

FACULTY-INCHARGE
Mr. T. PAVAN VINAYAK

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

AY: 2022-2023

COURSE FILE

SUBJECT: Digital Image Proessing

ACADEMIC YEAR: 2022-2023

Regulation: R18

NAME OF THE FACULTY: Mr. T. Pavan Vinayak
DEPARTMENT: ECE
YEAR & SEMESTER: IV-I
SUBJECT CODE: EC713PE

INDEX		
COURSE FILE		
S.NO	TOPIC	PAGE NO

Mission

- To impart quality technical education to produce industry-ready engineers.
- To motivate the students towards research leading to multidisciplinary innovative projects.
- To produce future leaders with cohesive teamwork.
- Maintain high academic standards and teaching quality that encourages critical thinking and self-evaluation.

Vision

- Imparting quality technical education through research, innovation and teamwork for a lasting technology development in the area of Electronics and Communication Engineering

Program Educational Objectives (PEOs) :

- **PEO1:** To excel in different fields of electronics and communication as well as in multidisciplinary areas. This can lead to a new era in developing a good electronic product.
- **PEO2:** To increase the ability and confidence among the students to solve any problem in their profession by applying mathematical, scientific and engineering methods in a better and efficient way.
- **PEO3:** To provide a good academic environment to the students which can lead to excellence, and stress upon the importance of teamwork and good leadership qualities, written ethical codes and guide lines for lifelong learning needed for a successful professional career.
- **PEO4:** To provide student with a solid foundation to students in all areas like mathematics, science and engineering fundamentals required to solve engineering problems, and also to pursue higher studies.
- **PEO5:** To expose the student to the state of art technology so that the student would be in a position to take up any assignment after his graduation.

Program Specific Outcomes (PSOs) of the E.C.E. Department

The graduates of the department will attain:

PSO1: The ability to absorb and apply fundamental knowledge of core Electronics and Communication Engineering subjects in the analysis, design, and development of various types of integrated electronic systems as well as to interpret and synthesize the experimental data leading to valid conclusions.

PSO2: Competence in using electronic modern IT tools (both software and hardware) for the design and analysis of complex electronic systems in furtherance to research activities.

PSO3: Excellent adaptability to changing work environment, good interpersonal skills as a leader in a team in appreciation of professional ethics and societal responsibilities.

Program Outcomes(PO'S):

a) Graduates will demonstrate ability to apply knowledge of Mathematics, Science and Engineering.

b) Graduate will demonstrate an ability to visualize and work on laboratory and multidisciplinary tasks.

c) Graduates should have the ability to design and synthesize a system, a component or process as per needs and specifications.

d) Graduates will have ability to function on Multidisciplinary teams.

e) Graduates will demonstrate an ability to solve the practical problem that arises by identifying and relating them to the basic subjects they have studied.

f) Graduates will have knowledge of professional and ethical responsibilities.

g) Graduates are expected to communicate effectively.

h) Graduates will show the understanding of impact of engineering solutions in a global, economic, environmental and societal context.

i) Graduates will develop confidence for self study and ability for lifelong learning.

j) Graduates will have knowledge of contemporary issues.

k) Graduates will have knowledge on modern engineering tools, software and equipment to analyze problems.

l) Graduates will be in a position to serve the mankind by contributing directly or indirectly in a way of development.

SYLLABUS

UNIT – I

Digital image fundamentals & Image Transform: Digital image fundamentals, Sampling and Quantization, Relationship between Pixels.

Image Transforms: 2-D FFT, Properties, Walsh Transform, Hadamard Transform, Discrete Cosine Transform, Haar Transform, Slant Transform, Hotelling Transform

UNIT – II

Image Enhancement(Spatial Domain): Introduction, Image enhancement in Spatial Domain, Enhancement through point Operation, Types of point operation, Histogram Manipulation, Linear and Non-Linear gray level Transformation, Local or Neighborhood Operation, Median Filter, Spatial Domain High - Pass Filtering.

Image Enhancement (Frequency Domain): Filtering in Frequency Domain, Low pass (Smoothing), High pass (Sharpening) filters in frequency Domain

UNIT – III

Image Restoration: Introduction, Degradation Model, Algebraic Approach to Restoration, Inverse Filtering, Least Mean Square Filters, Constrained Least Squares Restoration, Iterative Restoration.

UNIT – IV

Image Segmentation: Detection of Discontinuities, Edge Linking, Boundary Detection, Thresholding, Region Oriented Segmentation, Morphological Image processing, Dilation and Erosion, Dilation, Structuring Element Decomposition, Erosion, Combining Dilation and Erosion, Opening and Closing, The Hit or Miss Transformation.

UNIT – V

Image Compression: Redundancies, Redundancies and removal Methods, Fidelity Criteria, Image Compression Models, Huffman coding, Arithmetic coding, Error Free Compression, Lossy Compression, Lossy Predictive Coding, Lossless Predictive Coding, Transform Based Compression, JPEG 2000 Standards.

TEXT BOOKS:

1. Digital Image processing – Rafael C. Gonzalez, Richard E Woods, 3rd Edition, pearson, 2008
2. Digital Image processing- S Jayaraman, S Esakkirajan, T Veerakumar-TMH, 2010

REFERENCE BOOKS:

1. Digital Image processing and analysis-Human and computer vision Application with using CVIP Tools – Scotte Umbaugh, 2nd Ed, CRC press, 2011
2. Digital Image processing using MATLAB - Rafael C. Gonzalez, Richard E Woods and steven L. Eddings, 2nd Edition, TMH, 2010
3. Fundamentals of Digital image processing- A K Jain, PHI, 1989

INDIVIDUAL TIME TABLE

NAME OF THE FACULTY: Mr. T. Pavan Vinayak

Day/ Time	9:15 am to 10:15 am	10:15 am to 11:15 am	11:15 am to 12:15 am	12:15 am to 1:15 pm	1:15pm to 2:00 pm	2:00 pm to 3:00 pm	3:00 pm to 4:00 pm
Monday		DIP			LUNCH		
Tuesday							
Wednesday	DIP	DIP					
Thursday				DIP			
Friday		DIP					
Saturday							

CLASS TIME TABLE:

Day/ Time	9:15 am to 10:15 am	10:15 am to 11:15 am	11:15 am to 12:15 pm	12:15 pm to 1:15 pm	1:15 pm to 2:00 pm	2:00 pm to 3:00 pm	3:00 pm to 4:00 pm
Monday	DBMS	DIP	RSGIS	MWOC	LUNCH	MAJOR STAGE-I	
Tuesday	RSGIS	MAJOR STAGE-I		PPLE		MWOC LAB/ MINI PROJECT	
Wednesday	DIP	DIP	DBMS	RSGIS		MAJOR STAGE-I	
Thursday	MWOC	DBMS	PPLE	DIP		SEMINAR	
Friday	PPLE	DIP	DBMS	MWOC		MINI PROJECT/ MWOC LAB	
Saturday	MWOC	DBMS	MINI PROJECT			PPLE	LIBRARY

STUDENT LIST:

SNo	H.T.No	Name
1	19S11A0461	ABHHINANDU K G
2	19S11A0462	AKHIL MARYADA
3	19S11A0463	AKHIL VIKKURTHI
4	19S11A0464	AMULYA K
5	19S11A0465	ARAVIND REDDY SINGIDI
6	19S11A0466	BHARGAV REDDY M
7	19S11A0467	BHASKAR NAGA SAI N
8	19S11A0468	DHEERAJ KUMAR S
9	19S11A0469	ESHWAR TEJA VALABOJU
10	19S11A0470	HARIKRISHNA ANUMULA
11	19S11A0471	HEMA SRI YERRA
12	19S11A0472	KARTHIK KUMAR KASULAVADHA
13	19S11A0473	KARTHIK MUMMADI
14	19S11A0474	KETHAN PATEL
15	19S11A0475	KOWSHIK C
16	19S11A0476	KRISHNA MUSTIPALLY
17	19S11A0477	LIKHITH REDDY UPPELA
18	19S11A0478	MADHURYA AKKINAGUNTA
19	19S11A0479	MANIDEEP THATIPELLY
20	19S11A0480	MANVITH REDDY NAREDLA
21	19S11A0481	MD RAASHID ALI
22	19S11A0482	NARASIMHA RAJU B
23	19S11A0483	NARENDRA DIKONDA
24	19S11A0484	NAVYASRI RAYUDU
25	19S11A0485	NEERAJ REDDY M
26	19S11A0486	NIREEKSHAN H
27	19S11A0487	NITHIN CHOWDARY
28	19S11A0488	PAVAN S
29	19S11A0489	PRAVEEN KOLAGANI
30	19S11A0491	RAMYA PATLOLLA
31	19S11A0492	REDDY PAVAN KALYAN CH
32	19S11A0493	RUCHITHA ASKANI
33	19S11A0494	RUCHITHA THALLA
34	19S11A0495	SAHITHI KURA

35	19S11A0497	SAI SANGAMESH GUPTA VATTAMVAR
36	19S11A0498	SAI SUMANTH CHARY VADLA
37	19S11A0499	SAMAD REDDY GADILA
38	19S11A04A0	SANMAPRIYA SURAKASULA
39	19S11A04A1	SHAIK ABUTALHA
40	19S11A04A2	SHAILESH PALA
41	19S11A04A3	SHEKAR BAGANNAGARI
42	19S11A04A4	SHIVA SAI KUMAR B
43	19S11A04A5	SOWMYA NALLAGORLA
44	19S11A04A6	SRAVAN KUMAR JAJAM
45	19S11A04A7	SRISHNA GONE
46	19S11A04A8	SRIVIDHYA MAMIDI
47	19S11A04A9	TARUN CHANDRA T
48	19S11A04B0	TARUN KUMAR GUMMALA
49	19S11A04B1	THARUN MIDDE
50	19S11A04B3	UTHAM KUMAR REDDY BADDAM
51	19S11A04B4	UTTEJ KYATHAM
52	19S11A04B5	VAMSHIKRISHNA KOTHI
53	19S11A04B6	VARSHINI PAPANKA
54	19S11A04B7	VASAVI MANDEPUDI
55	19S11A04B8	VIJAYA BADETI
56	19S11A04B9	VIVEK THALLA
57	19S11A04C0	YASHWANTH KOTLA
58	20S15A0421	RAMYA MADUPU
59	20S15A0422	SAI ADITHYA CHATLA
60	20S15A0423	SAI DHANUSH GUGULOTHU
61	20S15A0424	SAI PRAKASH GAJULA
62	20S15A0425	SAI TEJA KARNE
63	20S15A0426	SARA SUSHANK
64	20S15A0427	SHIVA KUMAR KALPAGURI
65	20S15A0428	SRI RAM MANIKANTA PALLA
66	20S15A0429	SUJITH YAMSANI
67	20S15A0430	VAMSHI GOPAGANI

LECTURE PLAN

Unit No.	Lesson No.	Date	No. of Periods	Topic/Sub Topic	Mode of Teaching	Course Outcome (COS)	Program Outcome (POS)	Reference Text Books
I	1.1	01/09/2022	3	Digital Image Fundamentals	PPT	CO 1	1	T1, R1
	1.2	05/09/2022	1	Sampling and Quantization	PPT		2	T1
	1.3	07/09/2022	2	Relationship between pixels	PPT		3	T1
	1.4	08/09/2022	2	2-D FFT, Properties	Chalk & Black board		2	T2, R3
	1.5	14/09/2022	1	Walsh Transform	Chalk & Black board		2	T2, R3
	1.6	15/09/2022	1	Hadamard Transform	Chalk & Black board		2	T2, R3
	1.7	16/09/2022	1	Discrete Cosine Transform	Chalk & Black board		2	T2, R3
	1.8	19/09/2022	1	Haar Transform	Chalk & Black board		2	T2, R3
	1.9	21/09/2022	1	Slant Transform	Chalk & Black board		2	T2, R3
	1.10	21/09/2022	1	Hotelling Transform	Chalk & Black board		2	T2, R3

		22/09/2022	ASSIGNMENT TEST – I					
II	2.1	23/09/2022	1	Image Enhancement in Spatial Domain	PPT	CO 2	1,3	T2
	2.2	26/09/2022	0.5	Enhancement Through point operation	PPT		1,3	T2
	2.3	26/09/2022	0.5	Types of point operation	PPT		1,3	T2
	2.4	28/09/2022	1	Histogram manipulation	PPT		4	T2
	2.5	10/10/2022	1	Linear and non linear grey level transformation	PPT		3	T2
	2.6	12/10/2022	0.5	Local or neighborhood operation	PPT		3	T2
	2.7	12/10/2022	0.5	Median filter	PPT		3	T2
	2.8	13/10/2022	1	Spatial domain high pass filtering	PPT		3	T2
	2.9	14/10/2022	1	Filtering in Frequency Domain	PPT		3	R2
	2.10	17/10/2022	5	Low pass(Smoothing) and High pass(Sharpening) filters in Frequency domain	NPTEL		3	T1
		20/10/2022	ASSIGNMENT TEST – II					

III	3.1	21/10/2022	3	Degradation model	Chalk & Black board/ PPT	CO 3	3	T1
	3.2	27/10/2022	3	Algebraic approach to restoration	Chalk & Black board/ PPT		4	T1
	3.3	31/10/2022	1	Inverse filtering	Chalk & Black board/ PPT		4	T1
	3.4	09/11/2022	3	Least mean square filters	Chalk & Black board/ PPT		4	T1
	3.5	11/11/2022	3	Constrained least squares restoration	NPTEL		4	T2
	3.6	17/11/2022	2	Interactive restoration	PPT		4	T2
		21/11/2022	ASSIGNMENT TEST-III					
	4.1	23/11/2022	2	Introduction to Image Segmentation, Detection of discontinuities	Chalk & Black board/ PPT		3	T1
	4.2	24/11/2022	2	Edge linking and boundary detection	Chalk & Black board/ PPT		4	T1
	4.3	28/11/2022	2	Thresholding	Chalk & Black board/ PPT		4	T1
	4.4	30/11/2022	2	Region oriented Segmentation	NPTEL		4	T1
	4.5	02/12/2022	1	Introduction to Morphological	Chalk & Black		3	T1

IV				Image Processing	board/ PPT	CO 4		
	4.6	05/12/2022	1	Structuring element decomposition	Chalk & Black board/ PPT		4	T1
	4.7	07/12/2022	1	Dilation	Chalk & Black board/ PPT		4	T1
	4.8	08/12/2022	1	Erosion	Chalk & Black board/ PPT		4	T1
	4.9	09/12/2022	1	Combing Erosion and Dilation	Chalk & Black board/ PPT		4	T1
	4.10	10/12/2022	1	Opening and Closing	Chalk & Black board/ PPT		4	T1
	4.11	12/12/2022	1	Hit or Miss Transformation	NTPEL		4	T1
		13/12/2022	ASSIGNMENT TEST-IV					
	5.1	14/12/2022	3	Redundancies and their removal methods	Chalk & Black board		3	T1
	5.2	16/12/2022	1	Fidelity criteria	Chalk & Black board		3	T1
	5.3	19/12/2022	3	Image Compression Models, Huffman Arithmetic Coding	Chalk & Black board		3	T1
	5.4	22/12/2022	1	Error free Compression	Chalk & Black		3	T1

V					board	CO 5		
	5.5	23/12/2022	1	Lossy Compression	PPT		3	T1
	5.6	28/12/2022	2	Lossy and lossless predictive coding	PPT		3	T1
	5.7	29/12/2022	1	Transform based compression	PPT		3	T1
	5.8	30/12/2022	1	JPEG 2000 standards	NPTEL		4	T1
			ASSIGNMENT TEST –V					

Mapping of CO's with PO'S and PSO's

Subject Name: Digital Image Processing **Branch/Year/Sem:** ECE IV-I

Course Code: EC713PE **Regulation:** R18

S.No	COURSE OBJECTIVES
COb 1	To provide a approach towards image processing and introduction about 2D transforms
COb 2	To expertise about enhancement methods in time and frequency domain
COb 3	To expertise about segmentation and compression techniques
COb 4	To understand the Morphological operations on an image.

S.No	COURSE OUTCOMES After completing this course The student will be able to:
CO 1	Explore the fundamental relations between pixels and utility of 2-D transforms in image Processer.
CO 2	Understand the enhancement processes on an image.
CO 3	Understand the segmentation and restoration processes on an image.
CO 4	Implement the various Morphological operations on an image.
CO 5	Understand the need of compression and evaluation of basic compression algorithms.

CO's-PO's Relationship Matrix

Course Outcomes	PO'S											
	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7	PO 8	PO 9	PO 10	PO 11	PO 12
CO 1	3	3	1	1	1	-	-	-	-	-	-	1
CO 2	3	3	1	2	3	-	-	-	-	-	-	1
CO 3	3	3	2	2	3	-	-	-	-	-	-	1
CO 4	3	3	2	2	3	-	-	-	-	-	-	1
CO 5	3	3	2	2	3	-	-	-	-	-	-	1

Note: Indicate the relationship by rating from 1 to 3**(1-Low, 2-Medium, 3-High)**

CO's-PSO's Relationship Matrix

Course Outcomes	PSO 1	PSO 2	PSO 3
CO 1	3	3	-
CO 2	3	3	-
CO 3	3	3	-
CO 4	3	3	-
CO 5	3	3	-

Note: Indicate the relationship by rating from 1 to 3**(1-Low, 2-Medium, 3-High)**

UNIT-I
DIGITAL IMAGE FUNDAMENTALS AND IMAGE TRANSFORMS

1.1 INTRODUCTION:

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 amplitude 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 referred to as picture elements, image elements, pels, and pixels. Pixel is the term most widely used 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 ultra-sound, electron microscopy, and computer-generated images. Thus, digital image processing encompasses a wide and varied field of applications. There is no general agreement among authors regarding where image processing stops and other related areas, such as image analysis and computer vision, start. Sometimes a distinction is made by defining image processing as a discipline in which both the input and output of a process are images. We believe this to be a limiting and somewhat artificial boundary. For example, under this definition, even the trivial task of computing the average intensity of an image (which yields a single number) would not be considered an image processing operation. On the other hand, there are fields such as computer vision whose ultimate goal is to use computers to emulate human vision, including learning and being able to make inferences and take actions based on visual inputs. This area itself is a branch of artificial intelligence (AI) whose objective is to emulate human intelligence. The field of AI is in its earliest stages of infancy in terms of development, with progress having been much slower than originally anticipated. The area of image analysis (also called image understanding) is in between image processing and computer vision.

There are no clear-cut boundaries in the continuum from image processing at one end to computer vision at the other. However, one useful paradigm is to consider three types of computerized processes in this continuum: low-, mid-, and high-level processes. Low-level processes involve primitive operations such as image preprocessing to reduce noise, contrast enhancement, and image sharpening. A 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). Finally, 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 and, in addition, encompasses processes that extract attributes from images, up to and including the recognition of individual objects. As a simple illustration to clarify these concepts, consider the area of automated analysis of text. The processes of acquiring an image of the area containing the text, preprocessing that image, extracting (segmenting) the individual characters, describing the characters in a form suitable for computer processing, and recognizing those individual characters are in the scope of what we call digital image processing.

1.1.1 Representing Digital Images:

We will use two principal ways to represent digital images. Assume that an image f(x, y) is sampled so that the resulting digital image has M rows and N columns. The values of the coordinates (x, y) now become discrete quantities. For notational clarity and convenience, we shall use integer values for these discrete coordinates. Thus, the values of the coordinates at the origin are (x, y) = (0, 0). The next coordinate values along the first row of the image are represented as (x, y) = (0, 1). It is important to keep in mind that the notation (0, 1) is used to signify the second sample along the first row. It does not mean that these are the actual values of physical coordinates when the image was sampled. Figure 1.1 shows the coordinate convention used.

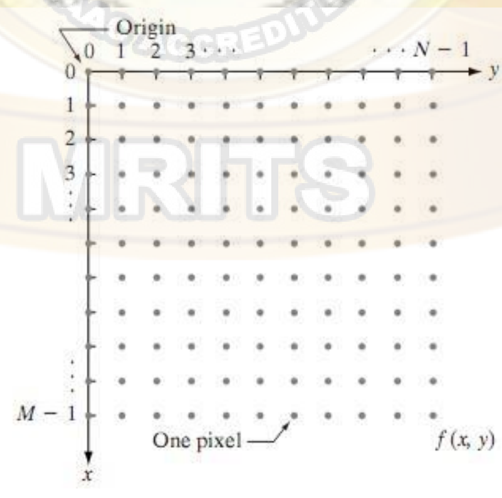


Fig 1.1 Coordinate convention used to represent digital images

The notation introduced in the preceding paragraph allows us to write the complete M*N digital image in the following compact matrix form:

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

The right side of this equation is by definition a digital image. Each element of this matrix array is called an image element, picture element, pixel, or pel.

1.2 FUNDAMENTAL STEPS IN DIGITAL IMAGE PROCESSING:

Image acquisition is the first process shown in Fig1.2. Note that acquisition could be as simple as being given an image that is already in digital form. Generally, the image acquisition stage involves preprocessing, such as scaling.

Image enhancement is among the simplest and most appealing areas of digital image processing. Basically, the idea behind enhancement techniques is to bring out detail that is obscured, or simply to highlight certain features of interest in an image. A familiar example of enhancement is when we increase the contrast of an image because “it looks better.” It is important to keep in mind that enhancement is a very subjective area of image processing.

Image restoration is an area that also deals with improving the appearance of an image. However, unlike enhancement, which is subjective, image restoration is objective, in the sense that restoration techniques tend to be based on mathematical or probabilistic models of image degradation. Enhancement, on the other hand, is based on human subjective preferences regarding what constitutes a “good” enhancement result.

Color image processing is an area that has been gaining in importance because of the significant increase in the use of digital images over the Internet.

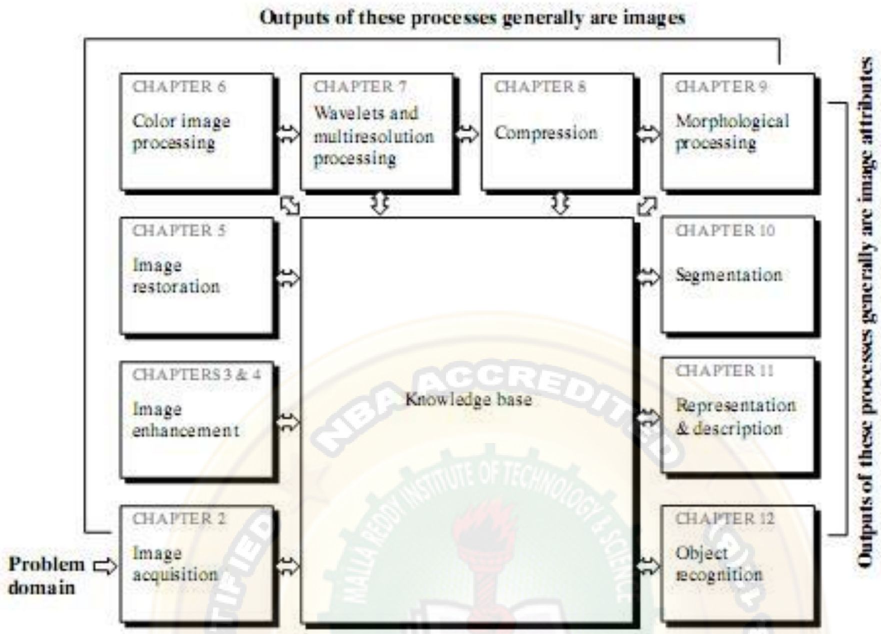


Fig.1.2. Fundamental steps in Digital Image Processing

Wavelets are the foundation for representing images in various degrees of resolution. Compression, as the name implies, deals with techniques for reducing the storage required to save an image, or the bandwidth required to transmit it. Although storage technology has improved significantly over the past decade, the same cannot be said for transmission capacity. This is true particularly in uses of the Internet, which are characterized by significant pictorial content. Image compression is familiar (perhaps inadvertently) to most users of computers in the form of image file extensions, such as the jpg file extension used in the JPEG (Joint Photographic Experts Group) image compression standard.

Morphological processing deals with tools for extracting image components that are useful in the representation and description of shape.

Segmentation procedures partition an image into its constituent parts or objects. In general, autonomous segmentation is one of the most difficult tasks in digital image processing. A rugged segmentation procedure brings the process a long way toward successful solution of imaging problems that require objects to be identified individually. On the other hand, weak or erratic segmentation algorithms almost always guarantee eventual failure. In general, the more accurate the segmentation, the more likely recognition is to succeed.

Representation and description almost always follow the output of a segmentation stage, which usually is raw pixel data, constituting either the boundary of a region (i.e., the set of pixels separating one image region from another) or all the points in the region itself. In either case, converting the data to a form suitable for computer processing is necessary. The first decision that must be made is whether the data should be represented as a boundary or as a complete region. Boundary representation is appropriate when the focus is on external shape characteristics, such as corners and inflections. Regional representation is appropriate when the focus is on internal properties, such as texture or skeletal shape. In some applications, these representations complement each other. Choosing a representation is only part of the solution for transforming raw data into a form suitable for subsequent computer processing. A method must also be specified for describing the data so that features of interest are highlighted. Description, also called feature selection, deals with extracting attributes that result in some quantitative information of interest or are basic for differentiating one class of objects from another.

Recognition is the process that assigns a label (e.g., “vehicle”) to an object based on its descriptors. We conclude our coverage of digital image processing with the development of methods for recognition of individual objects.

1.3 COMPONENTS OF AN IMAGE PROCESSING SYSTEM:

As recently as the mid-1980s, numerous models of image processing systems being sold throughout the world were rather substantial peripheral devices that attached to equally substantial host computers. Late in the 1980s and early in the 1990s, the market shifted to image processing hardware in the form of single boards designed to be compatible with industry standard buses and to fit into engineering workstation cabinets and personal computers. In addition to lowering costs, this market shift also served as a catalyst for a significant number of new companies whose specialty is the development of software written specifically for image processing.

Although large-scale image processing systems still are being sold for massive imaging applications, such as processing of satellite images, the trend continues toward miniaturizing and blending of general-purpose small computers with specialized image processing hardware. Figure 3 shows the basic components comprising a typical general-purpose system used for digital image processing. The function of each component is discussed in the following paragraphs, starting with image sensing.

With reference to sensing, two elements are required to acquire digital images. The first is a physical device that is sensitive to the energy radiated by the object we wish to image. The second, called a digitizer, is a device for converting the output of the physical sensing device into digital form. For instance, in a digital video camera, the sensors produce an electrical output proportional to light intensity. The digitizer converts these outputs to digital data.

Specialized image processing hardware usually consists of the digitizer just mentioned, plus hardware that performs other primitive operations, such as an arithmetic logic unit (ALU), which performs arithmetic and logical operations in parallel on entire images. One example of how an ALU is used is in averaging images as quickly as they are digitized, for the purpose of noise reduction. This type of hardware sometimes is called a front-end subsystem, and its most distinguishing characteristic is speed. In other words, this unit performs functions that require fast data throughputs (e.g., digitizing and averaging video images at 30 framess) that the typical main computer cannot handle.

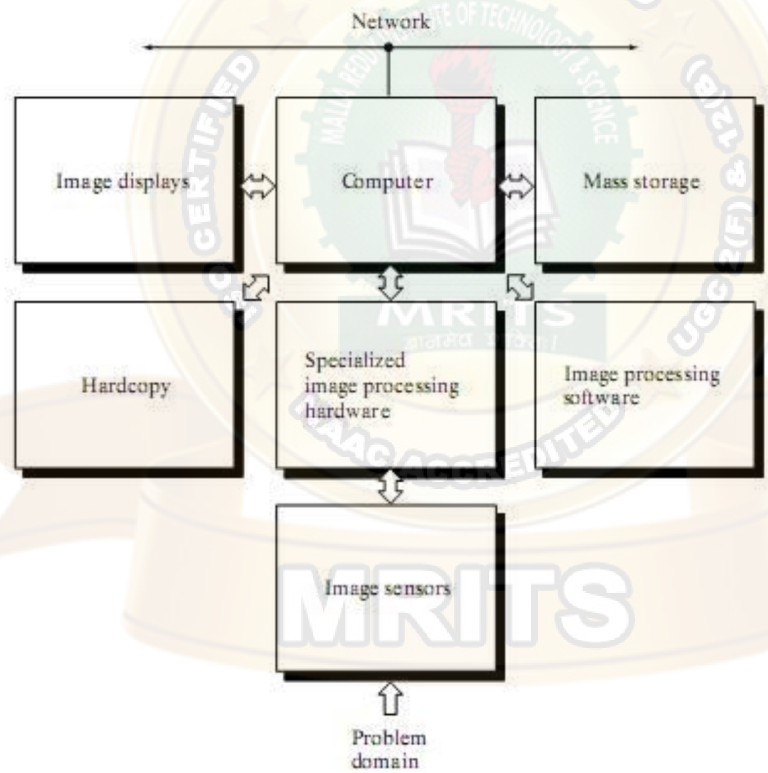


Fig.1.3. Components of a general purpose Image Processing System

The computer in an image processing system is a general-purpose computer and can range from a PC to a supercomputer. In dedicated applications, some times specially designed computers are used to achieve a required level of performance, but our interest here is on general-purpose image processing systems. In these systems, almost any well-equipped PC-type machine is suitable for offline image processing tasks.

Software for image processing consists of specialized modules that perform specific tasks. A well-designed package also includes the capability for the user to write code that, as a minimum, utilizes the specialized modules. More sophisticated software packages allow the integration of those modules and general-purpose software commands from at least one computer language.

Mass storage capability is a must in image processing applications. An image of size 1024*1024 pixels, in which the intensity of each pixel is an 8-bit quantity, requires one megabyte of storage space if the image is not compressed. When dealing with thousands, or even millions, of images, providing adequate storage in an image processing system can be a challenge. Digital storage for image processing applications falls into three principal categories: (1) short-term storage for use during processing, (2) on-line storage for relatively fast re-call, and (3) archival storage, characterized by infrequent access. Storage is measured in bytes (eight bits), Kbytes (one thousand bytes), Mbytes (one million bytes), Gbytes (meaning giga, or one billion, bytes), and Tbytes (meaning tera, or one trillion, bytes). One method of providing short-term storage is computer memory. Another is by specialized boards, called frame buffers, that store one or more images and can be accessed rapidly, usually at video rates (e.g., at 30 complete images per second).The latter method allows virtually instantaneous image zoom, as well as scroll (vertical shifts) and pan (horizontal shifts). Frame buffers usually are housed in the specialized image processing hardware unit shown in Fig.3.Online storage generally takes the form of magnetic disks or optical-media storage. The key factor characterizing on-line storage is frequent access to the stored data. Finally, archival storage is characterized by massive storage requirements but infrequent need for access. Magnetic tapes and optical disks housed in “jukeboxes” are the usual media for archival applications.

Image displays in use today are mainly color (preferably flat screen) TV monitors. Monitors are driven by the outputs of image and graphics display cards that are an integral part of the computer system. Seldom are there requirements for image display applications that cannot be met by display cards available commercially as part of the computer system. In some cases, it is necessary to have stereo displays, and these are implemented in the form of headgear containing two small displays embedded in goggles worn by the user.

Hardcopy devices for recording images include laser printers, film cameras, heat-sensitive devices, inkjet units, and digital units, such as optical and CD-ROM disks. Film provides the highest possible resolution, but paper is the obvious medium of choice for written material. For presentations, images are displayed on film transparencies or in a digital medium if image projection equipment is used. The latter approach is gaining acceptance as the standard for image presentations.

Networking is almost a default function in any computer system in use today. Because of the large amount of data inherent in image processing applications, the key consideration in image transmission is bandwidth. In dedicated networks, this typically is not a problem, but communications with remote sites via the Internet are not always as efficient. Fortunately, this situation is improving quickly as a result of optical fiber and other broadband technologies.

1.4 ELEMENTS OF VISUAL PERCEPTION:

Although the digital image processing field is built on a foundation of mathematical and probabilistic formulations, human intuition and analysis play a central role in the choice of one technique versus another, and this choice often is made based on subjective, visual judgments.

1.4.1 Structure of the Human Eye:

Figure 4 shows a simplified horizontal cross section of the human eye. The eye is nearly a sphere, with an average diameter of approximately 20 mm. Three membranes enclose the eye: the cornea and sclera outer cover; the choroid; and the retina. The cornea is a tough, transparent tissue that covers the anterior surface of the eye. Continuous with the cornea, the sclera is an opaque membrane that encloses the remainder of the optic globe. The choroid lies directly below the sclera. This membrane contains a network of blood vessels that serve as the major source of nutrition to the eye. Even superficial injury to the choroid, often not deemed serious, can lead to severe eye damage as a result of inflammation that restricts blood flow. The choroid coat is heavily pigmented and hence helps to reduce the amount of extraneous light entering the eye and the backscatter within the optical globe. At its anterior extreme, the choroid is divided into the ciliary body and the iris diaphragm. The latter contracts or expands to control the amount of light that enters the eye. The central opening of the iris (the pupil) varies in diameter from approximately 2 to 8 mm. The front of the iris contains the visible pigment of the eye, whereas the back contains a black pigment.

The lens is made up of concentric layers of fibrous cells and is suspended by fibers that attach to the ciliary body. It contains 60 to 70%water, about 6%fat, and more protein than any other tissue in the eye. The lens is colored by a slightly yellow pigmentation that increases with age. In extreme cases, excessive clouding of the lens, caused by the affliction commonly referred to as cataracts, can lead to poor color discrimination and loss of clear vision. The lens absorbs approximately 8% of the visible light spectrum, with relatively higher absorption at shorter wavelengths. Both infrared and ultraviolet light are absorbed appreciably by proteins within the lens structure and, in excessive amounts, can damage the eye.

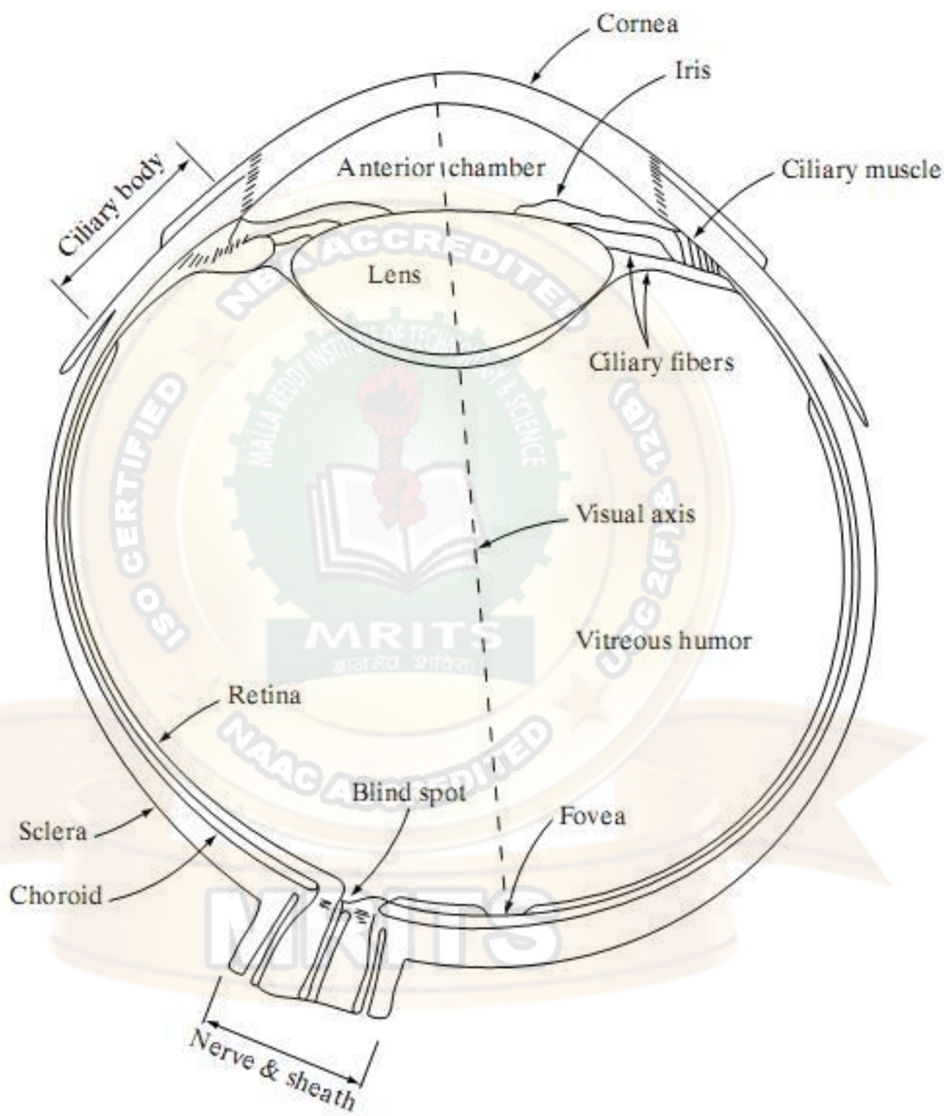


Fig.1.4.1 Simplified diagram of a cross section of the human eye.

The innermost membrane of the eye is the retina, which lines the inside of the wall's entire posterior portion. When the eye is properly focused, light from an object outside the eye is imaged on the retina. Pattern vision is afforded by the distribution of discrete light receptors over the surface of the retina. There are two classes of receptors: cones and rods. The cones in each eye number between 6 and 7 million. They are located primarily in the central portion of the retina, called the fovea, and are highly sensitive to color. Humans can resolve fine details with these cones largely because each one is connected to its own nerve end. Muscles controlling the eye rotate the eyeball until the image of an object of interest falls on the fovea. Cone vision is called photopic or bright-light vision. The number of rods is much larger: Some 75 to 150 million are distributed over the retinal surface. The larger area of distribution and the fact that several rods are connected to a single nerve end reduce the amount of detail discernible by these receptors. Rods serve to give a general, overall picture of the field of view. They are not involved in color vision and are sensitive to low levels of illumination. For example, objects that appear brightly colored in daylight when seen by moonlight appear as colorless forms because only the rods are stimulated. This phenomenon is known as scotopic or dim-light vision.

1.4.2 Image Formation in the Eye:

The principal difference between the lens of the eye and an ordinary optical lens is that the former is flexible. As illustrated in Fig. 1.4.1, the radius of curvature of the anterior surface of the lens is greater than the radius of its posterior surface. The shape of the lens is controlled by tension in the fibers of the ciliary body. To focus on distant objects, the controlling muscles cause the lens to be relatively flattened. Similarly, these muscles allow the lens to become thicker in order to focus on objects near the eye. The distance between the center of the lens and the retina (called the focal length) varies from approximately 17 mm to about 14 mm, as the refractive power of the lens increases from its minimum to its maximum. When the eye

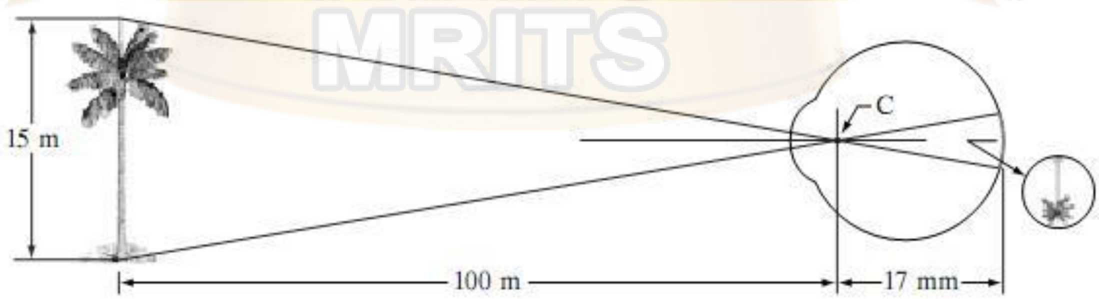


Fig.1.4.2. Graphical representation of the eye looking at a palm tree Point C is the optical center of the lens.

focuses on an object farther away than about 3 m, the lens exhibits its lowest refractive power. When the eye focuses on a nearby object, the lens is most strongly refractive. This information makes it easy to calculate the size of the retinal image of any object. In Fig. 1.4.2, for example, the observer is looking at a tree 15 m high at a distance of 100 m. If h is the height in mm of that object in the retinal image, the geometry of Fig.4.2 yields 15/100=h/17 or h=2.55mm. The retinal image is reflected primarily in the area of the fovea. Perception then takes place by the relative excitation of light receptors, which transform radiant energy into electrical impulses that are ultimately decoded by the brain.

1.4.3. 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 range of light intensity levels to which the human visual system can adapt is enormous—on the order of 10^{10}—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. Figure 1.4.3, 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 10^6. The transition from scotopic to photopic vision is gradual over the approximate range from 0.001 to 0.1 millilambert (–3 to –1 mL in the log scale), as the double branches of the adaptation curve in this range show.

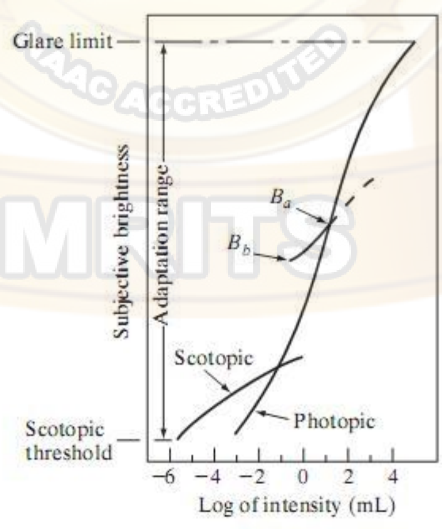


Fig.1.4.3. Range of Subjective brightness sensations showing a particular adaptation level.

The essential point in interpreting the impressive dynamic range depicted in Fig.4.3 is that the visual system cannot operate over such a range simultaneously. Rather, it accomplishes this large variation by changes in its overall sensitivity, a phenomenon known as brightness adaptation. The total range of distinct intensity levels it 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 Ba in Fig. 4.3. 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 Bb at and below which all stimuli are perceived as indistinguishable blacks. The upper (dashed) 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 Ba.

1.5. IMAGE ACQUISITION:

1.5.1 Image Sensing and Acquisition:

The types of images in which we are interested are generated by the combination of an “illumination” source and the reflection or absorption of energy from that source by the elements of the “scene” being imaged. We enclose illumination and scene in quotes to emphasize the fact that they are considerably more general than the familiar situation in which a visible light source illuminates a common everyday 3-D (three-dimensional) scene. For example, the illumination may originate from a source of electromagnetic energy such as radar, infrared, or X-ray energy. But, as noted earlier, it could originate from less traditional sources, such as ultrasound or even a computer-generated illumination pattern.

Similarly, the scene elements could be familiar objects, but they can just as easily be molecules, buried rock formations, or a human brain. We could even image a source, such as acquiring images of the sun. Depending on the nature of the source, illumination energy is reflected from, or transmitted through, objects. An example in the first category is light reflected from a planar surface. An example in the second category is when X-rays pass through a patient’s body for the purpose of generating a diagnostic X-ray film. In some applications, the reflected or transmitted energy is focused onto a photo converter (e.g., a phosphor screen), which converts the energy into visible light. Electron microscopy and some applications of gamma imaging use this approach.

Figure 1.5.1 shows the three principal sensor arrangements used to transform illumination energy into digital images. The idea is simple: Incoming energy is transformed into a voltage by the combination of input electrical power and sensor material that is responsive to the particular type of energy being detected. The output voltage waveform is the response of the sensor(s), and a digital quantity is obtained from each sensor by digitizing its response.

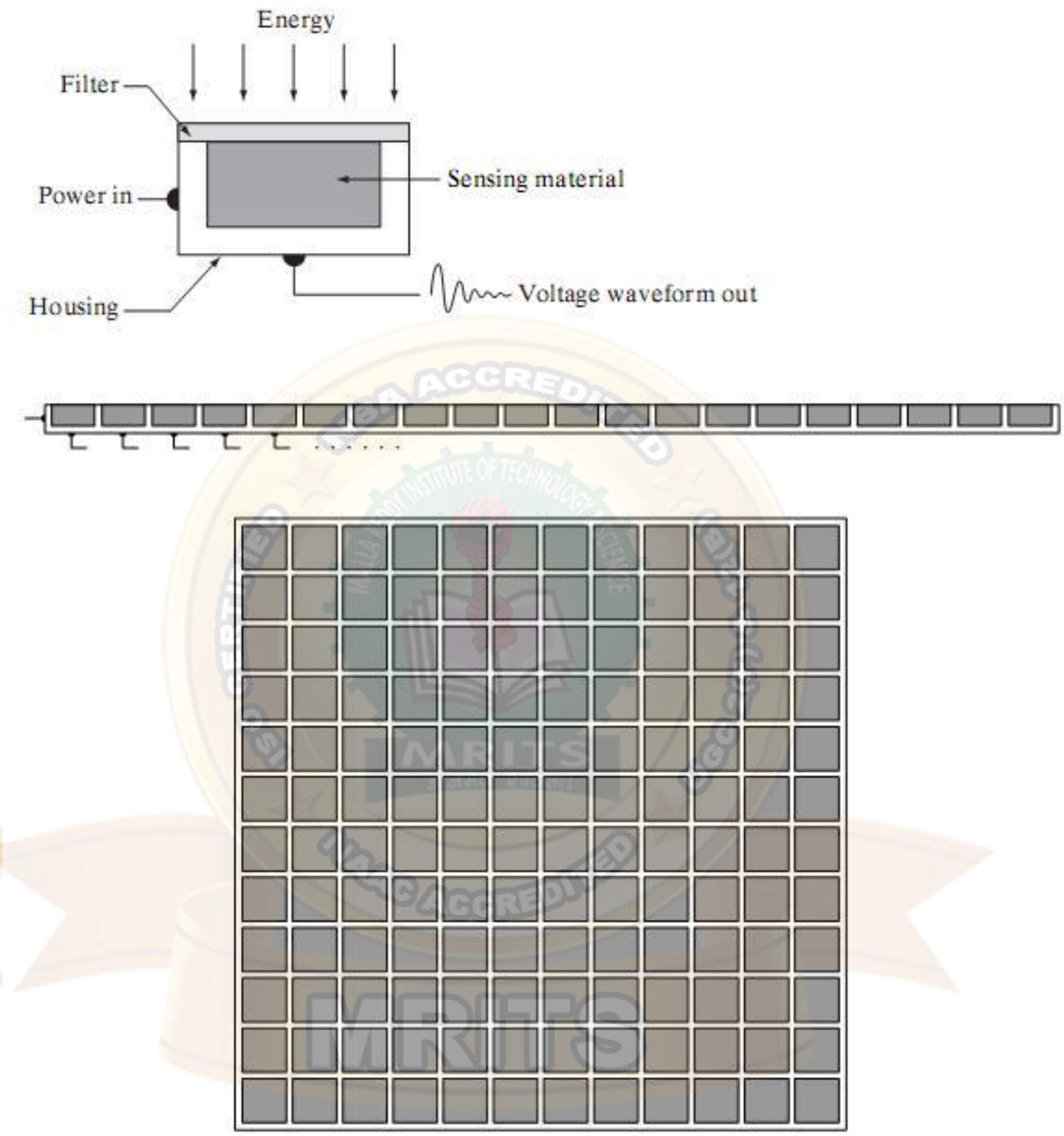


Fig.1.5.1 (a) Single imaging Sensor (b) Line sensor (c) Array sensor

1.5.2 Image Acquisition Using a Single Sensor:

Figure 1.5.1 (a) shows the components of a single sensor. Perhaps the most familiar sensor of this type is the photodiode, which is constructed of silicon materials and whose output voltage waveform is proportional to light. The use of a filter in front of a sensor improves selectivity. For example, a green (pass) filter in front of a light sensor favors light in the green band of the color spectrum. As a consequence, the sensor output will be stronger for green light than for other components in the visible spectrum.

In order to generate a 2-D image using a single sensor, there has to be relative displacements in both the x- and y-directions between the sensor and the area to be imaged. Figure 1.5.2 shows an arrangement used in high-precision scanning, where a film negative is mounted onto a drum whose mechanical rotation provides displacement in one dimension. The single sensor is mounted on a lead screw that provides motion in the perpendicular direction. Since mechanical motion can be controlled with high precision, this method is an inexpensive (but slow) way to obtain high-resolution images. Other similar mechanical arrangements use a flat bed, with the sensor moving in two linear directions. These types of mechanical digitizers sometimes are referred to as microdensitometers.

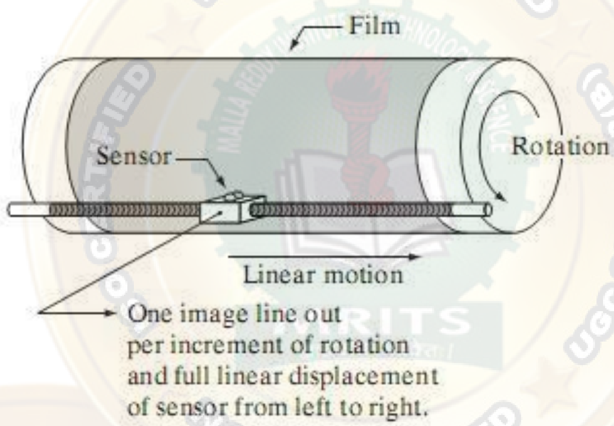


Fig.1.5.2. Combining a single sensor with motion to generate a 2-D image

1.5.3 Image Acquisition Using Sensor Strips:

A geometry that is used much more frequently than single sensors consists of an in-line arrangement of sensors in the form of a sensor strip, as Fig. 1.5.1 (b) shows. The strip provides imaging elements in one direction. Motion perpendicular to the strip provides imaging in the other direction, as shown in Fig. 1.5.3 (a).This is the type of arrangement used in most flat bed scanners. Sensing devices with 4000 or more in-line sensors are possible. In-line sensors are used routinely in airborne imaging applications, in which the imaging system is mounted on an aircraft that flies at a constant altitude and speed over the geographical area to be imaged. One-dimensional imaging sensor strips that respond to various bands of the electromagnetic spectrum are mounted perpendicular to the direction of flight. The imaging strip gives one line of an image at a time, and the motion of the strip completes the other dimension of a two-dimensional image. Lenses or other focusing schemes are used to project the area to be scanned onto the sensors.

Sensor strips mounted in a ring configuration are used in medical and industrial imaging to obtain cross-sectional ("slice") images of 3-D objects, as Fig. 1.5.3 (b) shows. A rotating X-ray source provides illumination and the portion of the sensors opposite the source collect the X-ray energy that pass through the object (the sensors obviously have to be sensitive to X-ray energy).This is the basis for medical and industrial computerized axial tomography (CAT). It is important to note that the output of the sensors must be processed by reconstruction algorithms whose objective is to transform the sensed data into meaningful cross-sectional images.

In other words, images are not obtained directly from the sensors by motion alone; they require extensive processing. A 3-D digital volume consisting of stacked images is generated as the object is moved in a direction perpendicular to the sensor ring. Other modalities of imaging based on the CAT principle include magnetic resonance imaging (MRI) and positron emission tomography (PET).The illumination sources, sensors, and types of images are different, but conceptually they are very similar to the basic imaging approach shown in Fig. 5.3 (b).

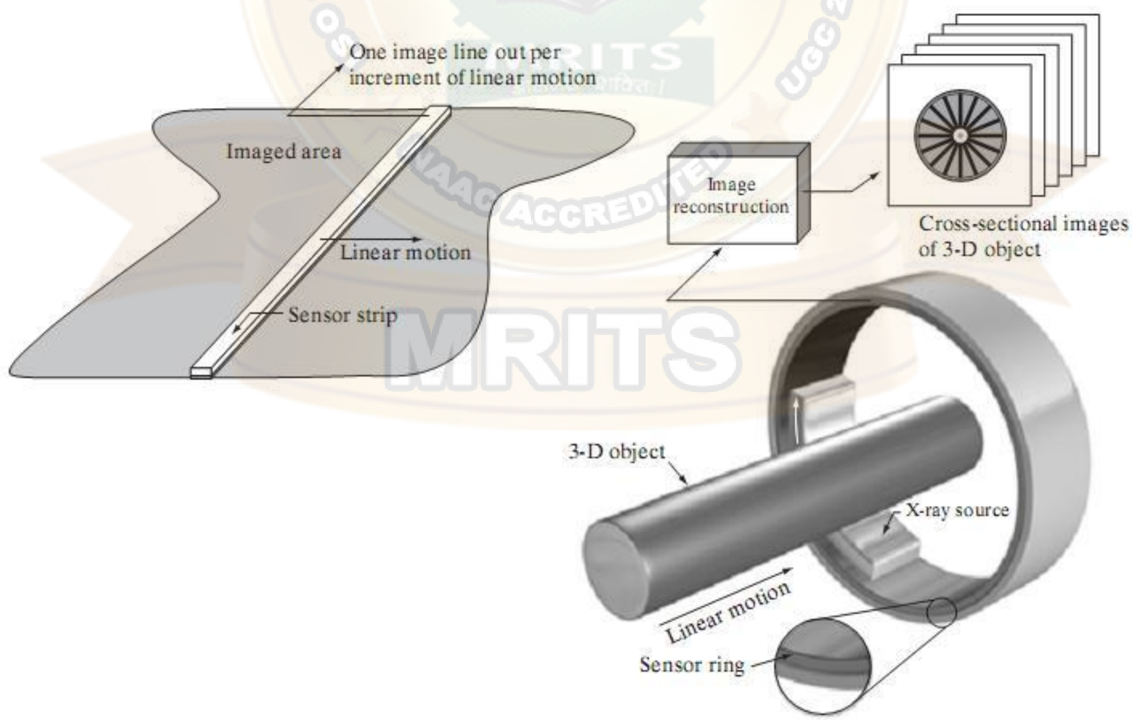


Fig.1.5.3 (a) Image acquisition using a linear sensor strip (b) Image acquisition using a circular sensor strip.

1.5.4 Image Acquisition Using Sensor Arrays:

Figure 1.5.1 (c) shows individual sensors arranged in the form of a 2-D array. Numerous electromagnetic and some ultrasonic sensing devices frequently are arranged in an array format. This is also the predominant arrangement found in digital cameras. A typical sensor for these cameras is a CCD array, which can be manufactured with a broad range of sensing properties and can be packaged in rugged arrays of 4000 * 4000 elements or more. CCD sensors are used widely in digital cameras and other light sensing instruments. The response of each sensor is proportional to the integral of the light energy projected onto the surface of the sensor, a property that is used in astronomical and other applications requiring low noise images. Noise reduction is achieved by letting the sensor integrate the input light signal over minutes or even hours. Since the sensor array shown in Fig. 1.5.4 (c) is two dimensional, its key advantage is that a complete image can be obtained by focusing the energy pattern onto the surface of the array. The principal manner in which array sensors are used is shown in Fig.1.5.4. This figure shows the energy from an illumination source being reflected from a scene element, but, as mentioned at the beginning of this section, the energy also could be transmitted through the scene elements. The first function performed by the imaging system shown in Fig.1.5.4 (c) is to collect the incoming energy and focus it onto an image plane. If the illumination is light, the front end of the imaging system is a lens, which projects the viewed scene onto the lens focal plane, as Fig. 1.5.4(d) shows. The sensor array, which is coincident with the focal plane, produces outputs proportional to the integral of the light received at each sensor. Digital and analog circuitry sweep these outputs and converts them to a video signal, which is then digitized by another section of the imaging system. The output is a digital image, as shown diagrammatically in Fig. 1.5.4 (e).

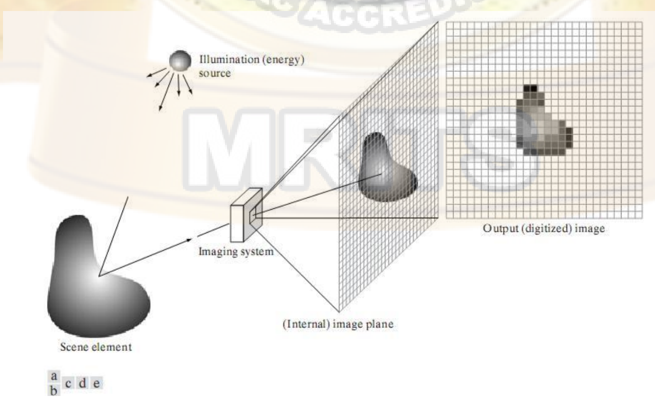


Fig.1.5.4 An example of the digital image acquisition process (a) Energy (“illumination”) source (b) An element of a scene (c) Imaging system (d) Projection of the scene onto the image plane (e) Digitized image

1.6. IMAGE SAMPLING AND QUANTIZATION:

1.6.1Image Sampling and Quantization:

The output of most sensors is a continuous voltage waveform whose amplitude and spatial behavior are related to the physical phenomenon being sensed. To create a digital image, we need to convert the continuous sensed data into digital form. This involves two processes: sampling and quantization.

1.6.2 Basic Concepts in Sampling and Quantization:

The basic idea behind sampling and quantization is illustrated in Fig.1.6.1. Figure 1.6.1(a) shows a continuous image, f(x, y), that we want to convert to digital form. An image may be continuous with respect to the x- and y-coordinates, and also in amplitude. To convert it to digital form, we have to sample the function in both coordinates and in amplitude. Digitizing the coordinate values is called sampling. Digitizing the amplitude values is called quantization.

The one-dimensional function shown in Fig.1.6.1 (b) is a plot of amplitude (gray level) values of the continuous image along the line segment AB in Fig. 1.6.1(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.1.6.1 (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. 6.1 (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.6.1 (d). Starting at the top of the image and carrying out this procedure line by line produces a two-dimensional digital image.

Sampling in the manner just described assumes that we have a continuous image in both coordinate directions as well as in amplitude. In practice, the method of sampling is determined by the sensor arrangement used to generate the image. When an image is generated by a single sensing element combined with mechanical motion, as in Fig. 2.13, the output of the sensor is quantized in the manner described above. However, sampling is accomplished by selecting the number of individual mechanical increments at which we activate the sensor to collect data. Mechanical motion can be made very exact so, in principle; there is almost no limit as to how fine we can sample an image.

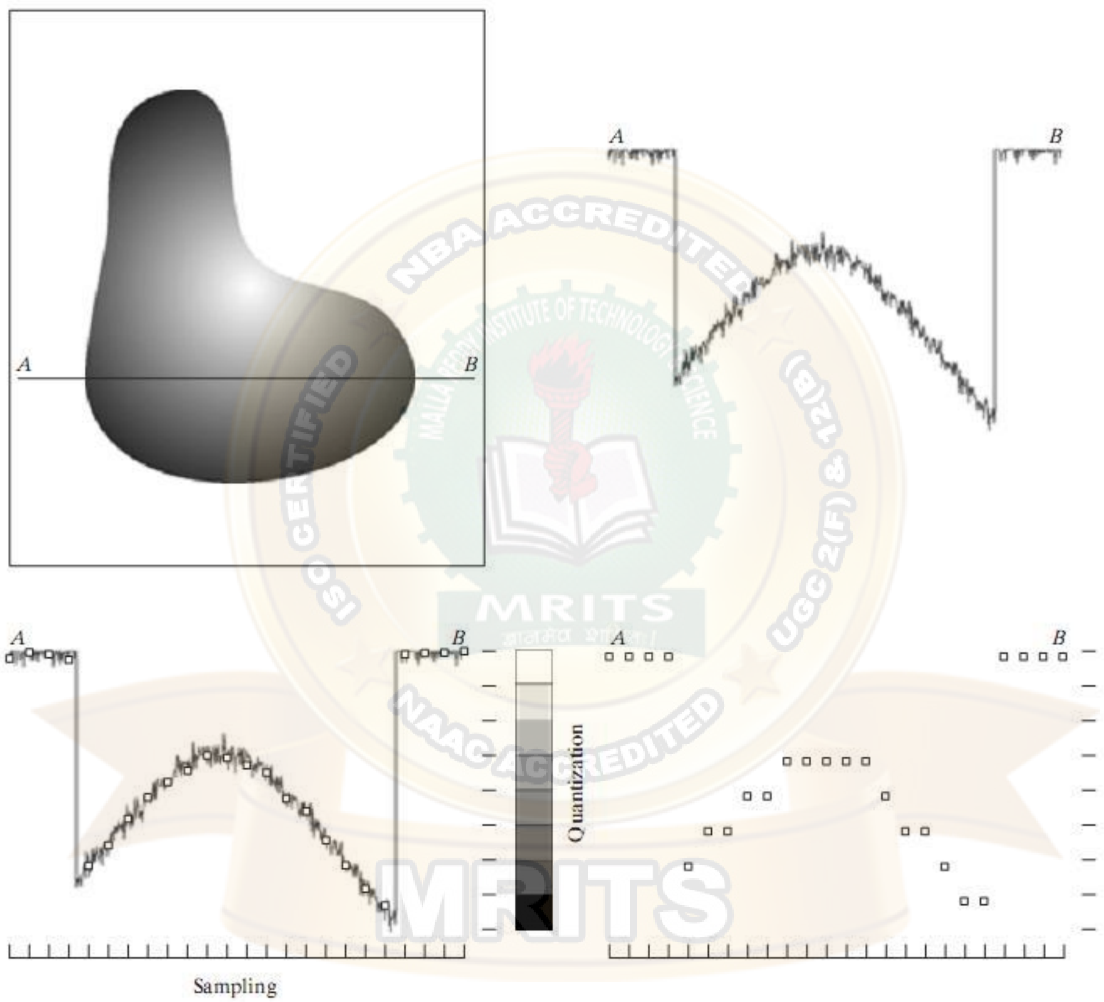


Fig.1.6.1. Generating a digital image (a) Continuous image (b) A scan line from A to Bin the continuous image, used to illustrate the concepts of sampling and quantization (c) Sampling and quantization. (d) Digital scan line

However, practical limits are established by imperfections in the optics used to focus on the sensor an illumination spot that is inconsistent with the fine resolution achievable with mechanical displacements. When a sensing strip is used for image acquisition, the number of sensors in the strip establishes the sampling limitations in one image direction. Mechanical motion in the other direction can be controlled more accurately, but it makes little sense to try to achieve sampling density in one direction that exceeds the sampling limits established by the number of sensors in the other. Quantization of the sensor outputs completes the process of generating a digital image.

When a sensing array is used for image acquisition, there is no motion and the number of sensors in the array establishes the limits of sampling in both directions. Figure 1.6.2 illustrates this concept. Figure 1.6.2 (a) shows a continuous image projected onto the plane of an array sensor. Figure 1.6.2 (b) shows the image after sampling and quantization. Clearly, the quality of a digital image is determined to a large degree by the number of samples and discrete gray levels used in sampling and quantization.

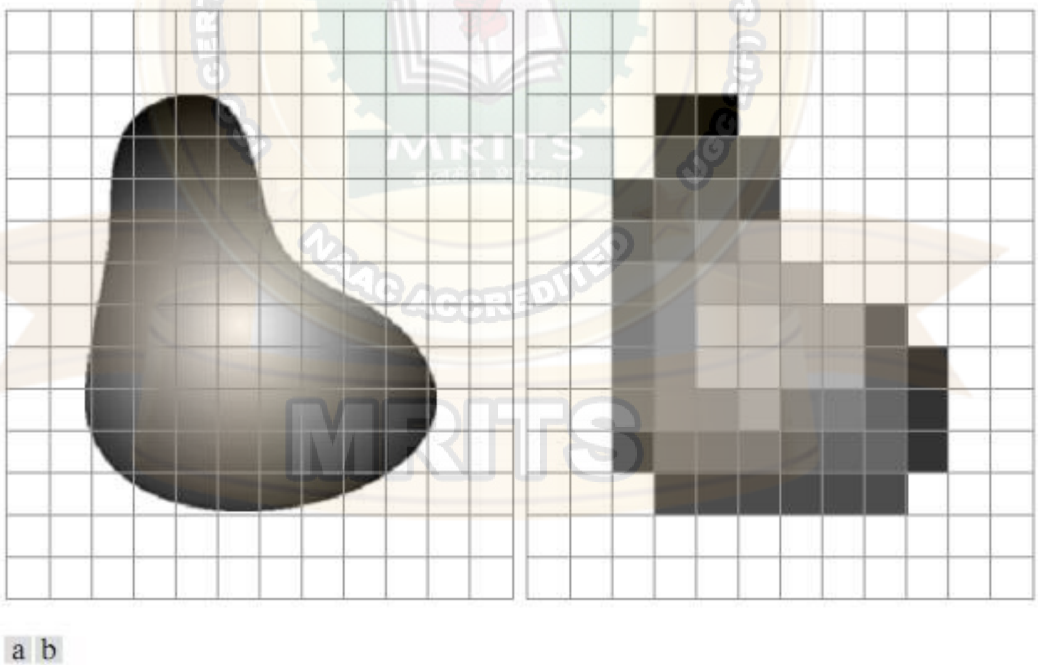


Fig.1.6.2. (a) Continuos image projected onto a sensor array (b) Result of image sampling and quantization.

1.6.3 Spatial and Gray-Level Resolution:

Sampling is the principal factor determining the spatial resolution of an image. Basically, spatial resolution is the smallest discernible detail in an image. Suppose that we construct a chart with vertical lines of width W, with the space between the lines also having width W.A line pair consists of one such line and its adjacent space. Thus, the width of a line pair is 2W, and there are 1/2Wline pairs per unit distance. A widely used definition of resolution is simply the smallest number of discernible line pairs per unit distance; for example, 100 line pairs per millimeter. Gray-level resolution similarly refers to the smallest discernible change in gray level. We have considerable discretion regarding the number of samples used to generate a digital image, but this is not true for the number of gray levels. Due to hardware considerations, the number of gray levels is usually an integer power of 2.

The most common number is 8 bits, with 16 bits being used in some applications where enhancement of specific gray-level ranges is necessary. Sometimes we find systems that can digitize the gray levels of an image with 10 or 12 bit of accuracy, but these are the exception rather than the rule. When an actual measure of physical resolution relating pixels and the level of detail they resolve in the original scene are not necessary, it is not uncommon to refer to an L-level digital image of size M*N as having a spatial resolution of M*N pixels and a gray-level resolution of L levels.

Fig.1.7.1. A 1024*1024, 8-bit image subsampled down to size 32*32 pixels The number of allowable gray levels was kept at 256.

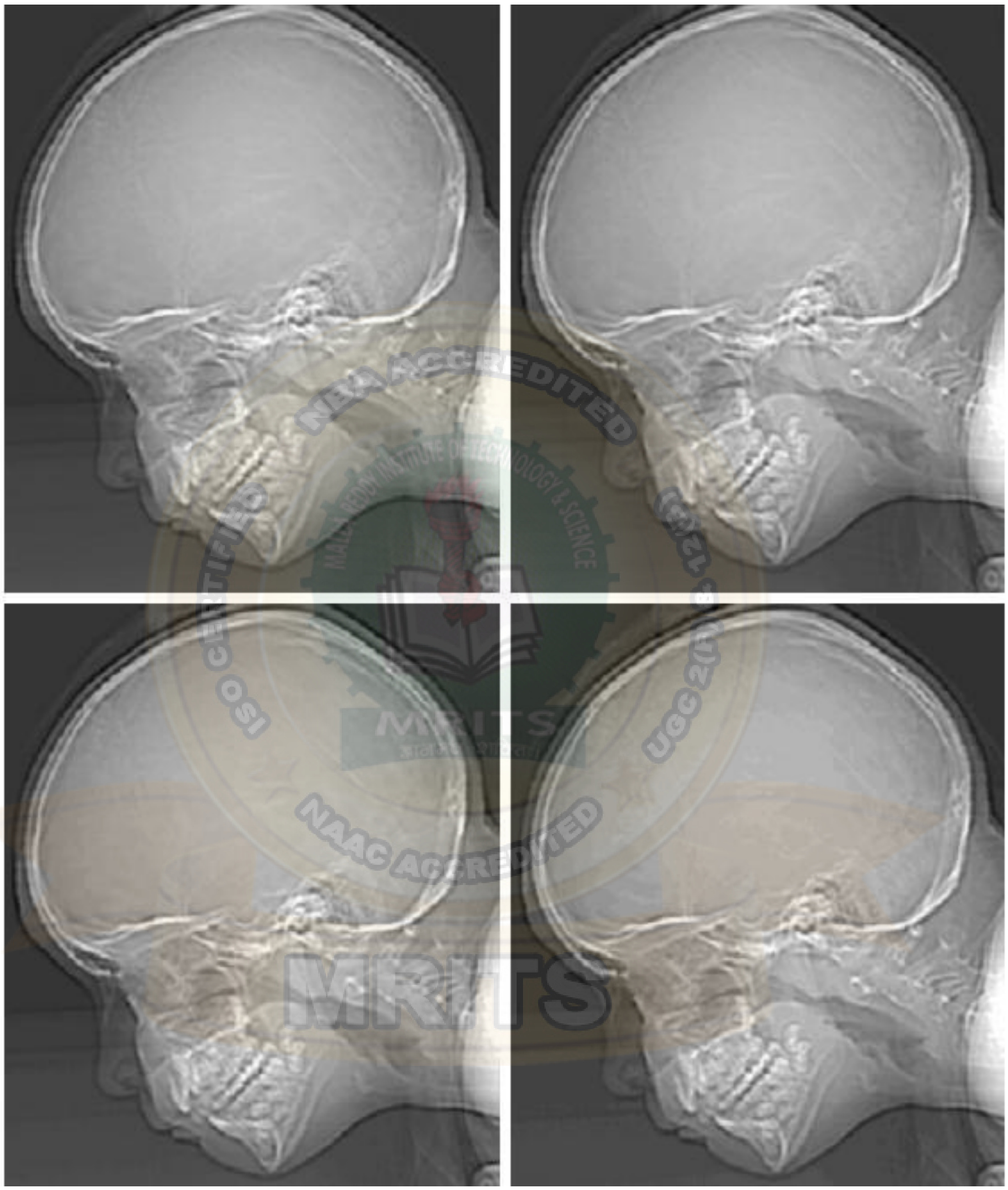
The subsampling was accomplished by deleting the appropriate number of rows and columns from the original image. For example, the 512*512 image was obtained by deleting every other row and column from the 1024*1024 image. The 256*256 image was generated by deleting every other row and column in the 512*512 image, and so on. The number of allowed gray levels was kept at 256. These images show the dimensional proportions between various sampling densities, but their size differences make it difficult to see the effects resulting from a reduction in the number of samples. The simplest way to compare these effects is to bring all the subsampled images up to size 1024*1024 by row and column pixel replication. The results are shown in Figs. 1.7.2 (b) through (f). Figure1.7.2 (a) is the same 1024*1024, 256-level image shown in Fig.1.7.1; it is repeated to facilitate comparisons.

a b c
d e f

Fig. 1.7.2 (a) 1024*1024, 8-bit image (b) 512*512 image resampled into 1024*1024 pixels by row and column duplication (c) through (f) 256*256, 128*128, 64*64, and 32*32 images resampled into 1024*1024 pixels

Compare Fig. 1.7.2(a) with the 512*512 image in Fig. 1.7.2(b) and note that it is virtually impossible to tell these two images apart. The level of detail lost is simply too fine to be seen on the printed page at the scale in which these images are shown. Next, the 256*256 image in Fig. 1.7.2(c) shows a very slight fine checkerboard pattern in the borders between flower petals and the black back-ground. A slightly more pronounced graininess throughout the image also is beginning to appear. These effects are much more visible in the 128*128 image in Fig. 1.7.2(d), and they become pronounced in the 64*64 and 32*32 images in Figs. 1.7.2 (e) and (f), respectively.

In the next example, we keep the number of samples constant and reduce the number of gray levels from 256 to 2, in integer powers of 2.Figure 1.7.3(a) is a 452*374 CAT projection image, displayed with k=8 (256 gray levels). Images such as this are obtained by fixing the X-ray source in one position, thus producing a 2-D image in any desired direction. Projection images are used as guides to set up the parameters for a CAT scanner, including tilt, number of slices, and range. Figures 1.7.3(b) through (h) were obtained by reducing the number of bits from k=7 to k=1 while keeping the spatial resolution constant at 452*374 pixels. The 256-, 128-, and 64-level images are visually identical for all practical purposes. The 32-level image shown in Fig. 1.7.3 (d), however, has an almost imperceptible set of very fine ridge like structures in areas of smooth gray levels (particularly in the skull).This effect, caused by the use of an insufficient number of gray levels in smooth areas of a digital image, is called false contouring, so called because the ridges resemble topographic contours in a map. False contouring generally is quite visible in images displayed using 16 or less uniformly spaced gray levels, as the images in Figs. 1.7.3(e) through (h) show.



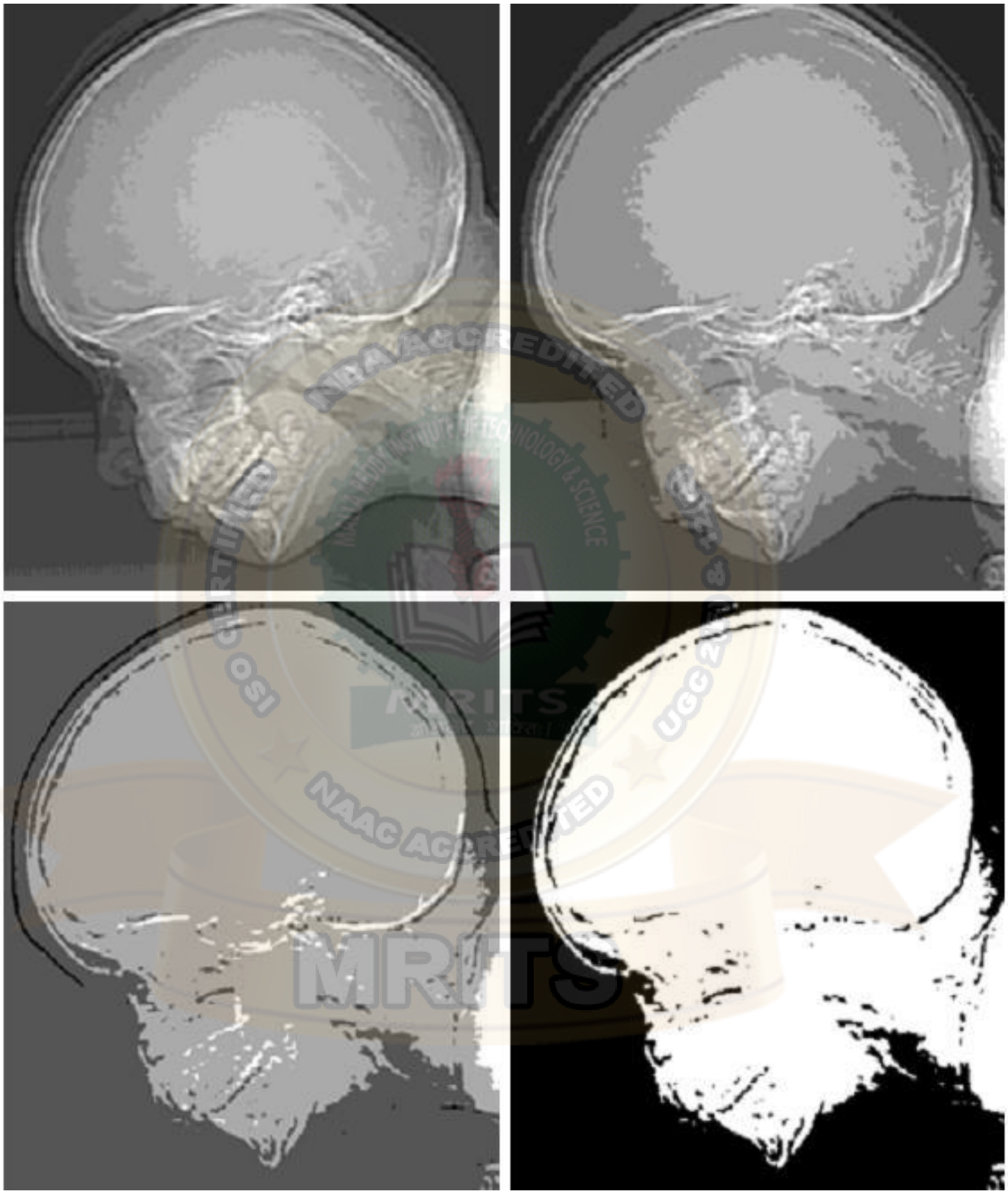


Fig. 1.7.3 (a) 452*374, 256-level image (b)–(d) Image displayed in 128, 64, and 32 gray levels, while keeping the spatial resolution constant (e)–(g) Image displayed in 16, 8, 4, and 2 gray levels.

As a very rough rule of thumb, and assuming powers of 2 for convenience, images of size 256*256 pixels and 64 gray levels are about the smallest images that can be expected to be reasonably free of objectionable sampling checker-boards and false contouring.

The results in Examples 1.7.2 and 1.7.3 illustrate the effects produced on image quality by varying N and k independently. However, these results only partially answer the question of how varying N and k affect images because we have not considered yet any relationships that might exist between these two parameters.

An early study by Huang [1965] attempted to quantify experimentally the effects on image quality produced by varying N and k simultaneously. The experiment consisted of a set of subjective tests. Images similar to those shown in Fig.1.7.4 were used. The woman's face is representative of an image with relatively little detail; the picture of the cameraman contains an intermediate amount of detail; and the crowd picture contains, by comparison, a large amount of detail. Sets of these three types of images were generated by varying N and k, and observers were then asked to rank them according to their subjective quality. Results were summarized in the form of so-called isopreference curves in the Nk-plane (Fig.1.7.5 shows average isopreference curves representative of curves corresponding to the images shown in Fig. 1.7.4).Each point in the Nk-plane represents an image having values of N and k equal to the coordinates of that point.

Fig.1.7.4 (a) Image with a low level of detail (b) Image with a medium level of detail (c) Image with a relatively large amount of detail

Points lying on an isopreference curve correspond to images of equal subjective quality. It was found in the course of the experiments that the isopreference curves tended to shift right and upward, but their shapes in each of the three image categories were similar to those shown in

Fig. 1.7.5. This is not unexpected, since a shift up and right in the curves simply means larger values for N and k, which implies better picture quality.

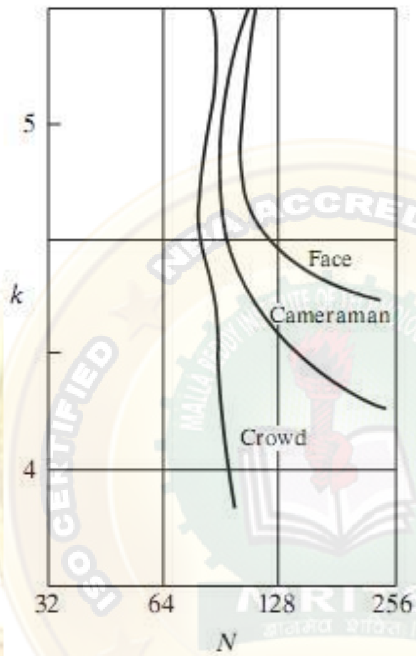


Fig.1.7.5. Representative isopreference curves for the three types of images in Fig.1.7.4

The key point of interest in the context of the present discussion is that isopreference curves tend to become more vertical as the detail in the image increases. This result suggests that for images with a large amount of detail only a few gray levels may be needed. For example, the isopreference curve in Fig.1.7.5 corresponding to the crowd is nearly vertical. This indicates that, for a fixed value of N, the perceived quality for this type of image is nearly independent of the number of gray levels used. It is also of interest to note that perceived quality in the other two image categories remained the same in some intervals in which the spatial resolution was increased, but the number of gray levels actually decreased. The most likely reason for this result is that a decrease in k tends to increase the apparent contrast of an image, a visual effect that humans often perceive as improved quality in an image.

1.9 BASIC RELATIONSHIPS AND DISTANCE MEASURES BETWEEN PIXELS IN A DIGITAL IMAGE:

1.9.1 Neighbors of a Pixel:

A pixel p at coordinates (x, y) 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 N_4 (p). Each pixel is a unit distance from (x, y), and some of the neighbors of p lie outside the digital image if (x, y) 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 points in N_D (p) and N_8 (p) fall outside the image if (x, y) is on the border of the image.

1.9.2 Connectivity:

Connectivity between pixels is a fundamental concept that simplifies the definition of numerous digital image concepts, such as regions and boundaries. To establish if two pixels are connected, it must be determined if they are neighbors and if their gray levels satisfy a specified criterion of similarity (say, if their gray levels are equal). For instance, in a binary image with values 0 and 1, two pixels may be 4-neighbors, but they are said to be connected only if they have the same value.

Let V be the set of gray-level values used to define adjacency. In a binary image, V={1} if we are referring to adjacency of pixels with value 1. In a grayscale image, the idea is the same, but set V typically contains more elements. For example, in the adjacency of pixels with a range of possible gray-level values 0 to 255, set V could be any subset of these 256 values. We consider three types of adjacency:

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

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

(c) m-adjacency (mixed adjacency).Two pixels p and q with values from V are m-adjacent if

(i) q is in N_4 (p), or

(ii) q is in N_D (p) and the set has no pixels whose values are from V.

Mixed adjacency is a modification of 8-adjacency. It is introduced to eliminate the ambiguities that often arise when 8-adjacency is used. For example, consider the pixel arrangement shown in Fig.1.9

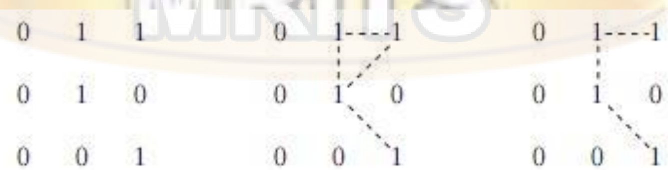
(a) for V= {1}.The three pixels at the top of Fig.9 (b) show multiple (ambiguous) 8-adjacency, as indicated by the dashed lines. This ambiguity is removed by using m-adjacency, as shown in Fig. 1.9 (c).Two image subsets S1 and S2 are adjacent if some pixel in S1 is adjacent to some pixel in S2. It is understood here and in the following definitions that adjacent means 4-, 8-
, or m-adjacent. A (digital) path (or curve) from pixel p with coordinates (x, y) to pixel q with coordinates (s, t) is a sequence of distinct pixels with coordinates

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

where $(x_0, y_0) = (x, y), (x_n, y_n) = (s, t)$, and pixels (x_i, y_i) and (x_{i-1}, y_{i-1}) are adjacent for $1 \leq i \leq n$. In this case, n is the length of the path. If $(x_0, y_0) = (x_n, y_n)$, the path is a closed path. We can define 4-, 8-, or m-paths depending on the type of adjacency specified. For example, the paths shown in Fig. 9 (b) between the northeast and southeast points are 8-paths, and the path in Fig. 9 (c) is an m-path. Note the absence of ambiguity in the m-path. Let S represent a subset of pixels in an image. Two pixels p and q are said to be connected in S if there

exists a path between them consisting entirely of pixels in S. For any pixel p in S, the set of pixels that are connected to it in S is called a connected component of S. If it only has one connected component, then set S is called a connected set.

Let R be a subset of pixels in an image. We call R a region of the image if R is a connected set. The boundary (also called border or contour) of a region R is the set of pixels in the region that have one or more neighbors that are not in R. If R happens to be an entire image (which we recall is a rectangular set of pixels), then its boundary is defined as the set of pixels in the first and last rows and columns of the image. This extra definition is required because an image has no neighbors beyond its border. Normally, when we refer to a region, we are referring to a subset



a b c

Fig1.9 (a) Arrangement of pixels; (b) pixels that are 8-adjacent (shown dashed) to the center pixel; (c) m-adjacency

of an image, and any pixels in the boundary of the region that happen to coincide with the border of the image are included implicitly as part of the region boundary.

1.9.3 Distance Measures:

For pixels p, q, and z, with coordinates (x, y), (s, t), and (v, w), respectively, D is a distance function or metric if

(a) $D(p, q) \geq 0 \quad (D(p, q) = 0 \quad \text{iff} \quad p = q)$,
(b) $D(p, q) = D(q, p)$, and
(c) $D(p, z) \leq D(p, q) + D(q, z)$.

The **Euclidean distance** between p and q is defined as

$$D_e(p, q) = \left[(x - s)^2 + (y - t)^2\right]^{\frac{1}{2}}.$$

For this distance measure, the pixels having a distance less than or equal to some value r from(x, y) are the points contained in a disk of radius r centered at (x, y).

The **D_4 distance (also called city-block distance)** between p and q is defined as

$$D_4(p, q) = |x - s| + |y - t|.$$

In this case, the pixels having a D_4 distance from (x, y) less than or equal to some value r form a diamond centered at (x, y). For example, the pixels with D_4 distance ≤ 2 from (x, y) (the center point) form the following contours of constant distance:

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

The pixels with D_4 =1 are the 4-neighbors of (x, y).

The **D_8 distance (also called chessboard distance)** between p and q is defined as

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

In this case, the pixels with D_8 distance from(x, y) less than or equal to some value r form a square centered at (x, y). For example, the pixels with D_8 distance ≤ 2 from(x, y) (the center point) form the following contours of constant distance:

$$\begin{matrix} 2 & 2 & 2 & 2 & 2 \\ 2 & 1 & 1 & 1 & 2 \\ 2 & 1 & 0 & 1 & 2 \\ 2 & 1 & 1 & 1 & 2 \\ 2 & 2 & 2 & 2 & 2 \end{matrix}$$

The pixels with D_8=1 are the 8-neighbors of (x, y). Note that the D4 and D8 distances between p and q are independent of any paths that might exist between the points because these distances involve only the coordinates of the points. If we elect to consider m-adjacency, however, the D_m distance between two points is defined as the shortest m-path between the points. In this case, the distance between two pixels will depend on the values of the pixels along the path, as well as the values of their neighbors. For instance, consider the following arrangement of pixels and assume that p, p_2 , and p_4 have value 1 and that p_1 and p_3 can have a value of 0 or 1:

$$\begin{matrix} & p_3 & p_4 \\ p_1 & p_2 & \\ p & & \end{matrix}$$

Suppose that we consider adjacency of pixels valued 1 (i.e. = {1}). If p_1 and p_3 are 0, the length of the shortest m-path (the Dm distance) between p and p_4 is 2. If p1 is 1, then p_2 and p will no longer be m-adjacent (see the definition of m-adjacency) and the length of the shortest m-path becomes 3 (the path goes through the points $pp_1p_2p_4$). Similar comments apply if p_3 is 1 (and p_1 is 0); in this case, the length of the shortest m-path also is 3. Finally, if both p_1 and p_3 are 1 the length of the shortest m-path between p and p4 is 4. In this case, the path goes through the sequence of points pp1p2p3p4.

IMAGE TRANSFORMS

Image transform is basically a representation of an image. There are two reasons for transforming an image from one representation to another. First, the transformation may isolate critical components of the image pattern so that they are directly accessible for analysis. Second, the transformation may place the image data in a more compact form so that they can be stored and transmitted efficiently.

1. 2-D FFT and its Properties:

Fourier Transform and its inverse:

Let f(x) be a continuous function of a real variable x. The Fourier transform of f(x) is defined by the equation

$$\Im\{f(x)\} = F(u) = \int_{-\infty}^{\infty} f(x)\exp[-j2\pi ux]\,dx$$

where j = √-1

Given F(u), f(x) can be obtained by using the inverse Fourier transform

$$\Im^{-1}\{F(u)\} = f(x) = \int_{-\infty}^{\infty} F(u)\exp[j2\pi ux]\,du.$$

The Fourier transform exists if f(x) is continuous and integrable and F(u) is integrable. The Fourier transform of a real function, is generally complex,

$$F(u) = R(u) + jI(u)$$

where R(u) and I(u) are the real and imiginary components of F(u). F(u) can be expressed in exponential form as

$$F(u) = |F(u)|\, e^{j\text{Ø}(u)}$$

where

$$|F(u)| = [R^2(u) + I^2(u)]^{1/2}$$

and

$$\text{Ø}(u, v) = \tan^{-1}[\, I(u, v)/R(u, v)\,]$$

The magnitude function |F (u)| is called the Fourier Spectrum of f(x) and Φ(u) its phase angle.

The variable u appearing in the Fourier transform is called the frequency variable.

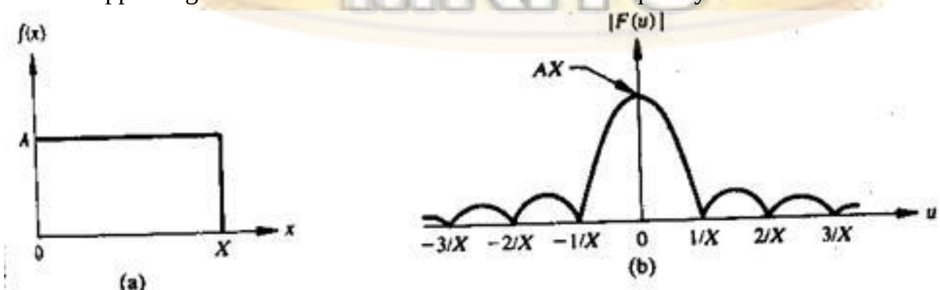


Fig.6. A simple function and its Fourier spectrum

The Fourier transform can be easily extended to a function f(x, y) of two variables. If f(x, y) is continuous and integrable and F(u,v) is integrable, following Fourier transform pair exists

and

$$\Im\{f(x, y)\} = F(u, v) = \iint_{-\infty}^{\infty} f(x, y)\exp[-j2\pi(ux + vy)]\,dx\,dy$$

$$\mathfrak{F}^{-1}\{F(u,v)\} = f(x,y) = \int\int_{-\infty}^{\infty} F(u,v)\exp[j2\pi(ux+vy)]\,du\,dv$$

where u, v are the frequency variables.
The Fourier spectrum, phase, are

$$|F(u,v)| = [R^2(u,v) + I^2(u,v)]^{1/2}$$

$$\text{Ø}(u,v) = \tan^{-1}[\,I(u,v)/R(u,v)\,]$$

Discrete Fourier Transform and its Inverse.

The discrete Fourier transform pair that applies to sampled function is given by,

$$F(u) = \frac{1}{N}\sum_{x=0}^{N-1} f(x)\exp[-j2\pi ux/N]$$

for u = 0, 1, 2 . . . . , N-1, and

$$f(x) = \sum_{u=0}^{N-1} F(u)\exp[j2\pi ux/N]$$

for x = 0, 1, 2 . . . ., N-1.

In the two variable case the discrete Fourier transform pair is

$$F(u,v) = \frac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)\exp[-j2\pi(ux/M + vy/N)]$$

for u = 0, 1, 2 . . . , M-1, v = 0, 1, 2 . . . , N - 1, and

$$f(x,y) = \sum_{u=0}^{M-1}\sum_{v=0}^{N-1} F(u,v)\exp[j2\pi(ux/M + vy/N)]$$

for x = 0, 1, 2 . . . , M-1, y = 0, 1, 2 . . . , N-1.

If M = N, then discrete Fourier transform pair is

$$F(u,v) = \frac{1}{N}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y)\exp[-j2\pi(ux + vy)/N]$$

For u, v = 0, 1, 2 . . . , N – 1, and

$$f(x,y) = \frac{1}{N}\sum_{u=0}^{N-1}\sum_{v=0}^{N-1} F(u,v)\exp[j2\pi(ux + vy)/N]$$

For x, y = 0, 1, 2 . . . , N – 1

Separability Property:

The separability property of 2D-DFT states that, the discrete Fourier transform pair can be expressed in the separable forms. i.e.

$$F(u,v) = \frac{1}{N}\sum_{x=0}^{N-1}\exp[-j2\pi ux/N]\sum_{y=0}^{N-1} f(x,y)\exp[-j2\pi vy/N] \qquad \text{——— (1)}$$

For u, v = 0, 1, 2 . . . , N – 1, and

$$f(x, y) = \frac{1}{N}\sum_{u=0}^{N-1}\exp[j2\pi ux/N]\sum_{v=0}^{N-1}F(u, v)\exp[j2\pi vy/N] \quad \text{———— (2)}$$

For x, y = 0, 1, 2 . . . , N – 1

The principal advantage of the separability property is that F(u,v) or f(x,y) can be obtained in two steps by successive applications of the 1-D Fourier transform or its inverse. This advantage becomes evident if equation (1) is expressed in the form

$$F(u, v) = \frac{1}{N}\sum_{x=0}^{N-1}F(x, v)\exp[-j2\pi ux/N] \quad \text{———— (3)}$$

Where,

$$F(x, v) = N\left[\frac{1}{N}\sum_{y=0}^{N-1}f(x, y)\exp[-j2\pi vy/N]\right]. \quad \text{———— (4)}$$

For each value of x, the expression inside the brackets in eq(4) is a 1-D transform, with frequency values v = 0, 1, . . . , N-1. Therefore the 2-D function f(x, v) is obtained by taking a transform along each row of f(x, y) and multiplying the result by N. The desired result, F(u, v), is then obtained by taking a transform along each column of F(x, v), as indicated by eq(3).

Translation Property:

The translation properties of the Fourier transform pair are

$$f(x, y)\exp[j2\pi(u_0x + v_0y)/N] \Leftrightarrow F(u - u_0, v - v_0) \quad (1)$$

and

$$f(x - x_0, y - y_0) \Leftrightarrow F(u, v)\exp[-j2\pi(ux_0 + vy_0)/N] \quad (2)$$

Where the double arrow indicates the correspondence between a function and its Fourier Transform,

Equation (1) shows that multiplying f(x, y) by the indicated exponential term and taking the transform of the product results in a shift of the origin of the frequency plane to the point (u_o, v_o).

Consider the equation (1) with $u_o = v_o = N/2$ or

$$\exp[j2\Pi(u_ox + v_oy)/N] = e^{j\Pi(x+y)} = (-1)^{(x+y)}$$

and

$$f(x, y)(-1)^{x+y} \Leftrightarrow F(u - N/2, v - N/2)$$

Thus the origin of the Fourier transform of f(x, y) can be moved to the center of its corresponding N x N frequency square simply by multiplying f(x, y) by $(-1)^{x+y}$. In the one variable case this shift reduces to multiplication of f(x) by the term $(-1)^x$. Note from equation (2) that a shift in f(x, y) does not affect the magnitude of its Fourier transform as,

$$|F(u, v)\exp[-j2\pi(ux_0 + vy_0)/N]| = |F(u, v)|.$$

Distributivity Property: From the definition of the continuous or discrete transform pair,

$$\Im\{f_1(x, y) + f_2(x, y)\} = \Im\{f_1(x, y)\} + \Im\{f_2(x, y)\}$$

and, in general,

$$\Im\{f_1(x, y) \cdot f_2(x, y)\} \neq \Im\{f_1(x, y)\} \cdot \Im\{f_2(x, y)\}.$$

In other words, the Fourier transform and its inverse are distributive over addition but not over multiplication.

Scaling Property:

For two scalars a and b,

$$af(x, y) \Leftrightarrow aF(u, v)$$

$$f(ax, by) \Leftrightarrow \frac{1}{|ab|} F(u/a, v/b).$$

2. Walsh Transform:

The discrete Walsh transform of a function *f* (x), denoted *W(u),* is given by

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

Walsh transform kernel is symmetric matrix having orthogonal rows and columns. These properties, which hold in general, lead to an inverse kernel given by

$$h(x, u) = \prod_{i=0}^{n-1} (-1)^{b_i(x) b_{n-1-i}(u)}.$$

Thus the inverse Walsh transform is given by

$$f(x) = \sum_{u=0}^{N-1} W(u) \prod_{i=0}^{n-1} (-1)^{b_i(x) b_{n-1-i}(u)}.$$

The 2-D forward and inverse Walsh kernels are given by

$$g(x, y, u, v) = \frac{1}{N} \prod_{i=0}^{n-1} (-1)^{[b_i(x) b_{n-1-i}(u) + b_i(y) b_{n-1-i}(v)]}$$

and

$$h(x, y, u, v) = \frac{1}{N} \prod_{i=0}^{n-1} (-1)^{[b_i(x) b_{n-1-i}(u) + b_i(y) b_{n-1-i}(v)]}.$$

Thus the forward and inverse Walsh transforms for 2-D are given by

$$W(u, v) = \frac{1}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \prod_{i=0}^{n-1} (-1)^{[b_i(x) b_{n-1-i}(u) + b_i(y) b_{n-1-i}(v)]}$$

and

$$f(x, y) = \frac{1}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} W(u, v) \prod_{i=0}^{n-1} (-1)^{[b_i(x) b_{n-1-i}(u) + b_i(y) b_{n-1-i}(v)]}.$$

The Walsh Transform kernels are seperable and symmetric, because

$$\begin{aligned} g(x, y, u, v) &= g_1(x, u) g_1(y, v) \\ &= h_1(x, u) h_1(y, v) \\ &= \left[\frac{1}{\sqrt{N}} \prod_{i=0}^{n-1} (-1)^{b_i(x) b_{n-1-i}(u)}\right]\left[\frac{1}{\sqrt{N}} \prod_{i=0}^{n-1} (-1)^{b_i(y) b_{n-1-i}(v)}\right]. \end{aligned}$$

Values of the 1-D walsh transform kernel for N = 8 is

$u \backslash x$	0	1	2	3	4	5	6	7
0	+	+	+	+	+	+	+	+
1	+	+	+	+	−	−	−	−
2	+	+	−	−	+	+	−	−
3	+	+	−	−	−	−	+	+
4	+	−	+	−	+	−	+	−
5	+	−	+	−	−	+	−	+
6	+	−	−	+	+	−	−	+
7	+	−	−	+	−	+	+	−

3. Hadamard Transform:

1-D forward kernel for hadamard transform is

$$g(x, u) = \frac{1}{N}(-1)^{\sum_{i=0}^{n-1} b_i(x)b_i(u)}$$

Expression for the 1-D forward Hadamard transform is

$$H(u) = \frac{1}{N}\sum_{x=0}^{N-1} f(x)(-1)^{\sum_{i=0}^{n-1} b_i(x)b_i(u)}$$

Where $N = 2^n$ and u has values in the range 0, 1, . . . , N-1.

1-D inverse kernel for hadamard transform is

$$h(x, u) = (-1)^{\sum_{i=0}^{n-1} b_i(x)b_i(u)}$$

Expression for the 1-D inverse Hadamard transform is

$$f(x) = \sum_{u=0}^{N-1} H(u)(-1)^{\sum_{i=0}^{n-1} b_i(x)b_i(u)}$$

The 2-D kernels are given by the relations

and

$$h(x, y, u, v) = \frac{1}{N}(-1)^{\sum_{i=0}^{n-1} [b_i(x)b_i(u) + b_i(y)b_i(v)]}$$

2-D Hadamard transform pair is given by following equations

$$H(u, v) = \frac{1}{N}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x, y)(-1)^{\sum_{i=0}^{n-1} [b_i(x)b_i(u) + b_i(y)b_i(v)]}$$

$$f(x, y) = \frac{1}{N}\sum_{u=0}^{N-1}\sum_{v=0}^{N-1} H(u, v)(-1)^{\sum_{i=0}^{n-1} [b_i(x)b_i(u) + b_i(y)b_i(v)]}$$

Values of the 1-D hadamard transform kernel for N = 8 is

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

The Hadamard matrix of lowest order N = 2 is

$$\mathbf{H}_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

If $\mathbf{H}_N$ represents the matrix of order N, the recursive relationship is given by

$$\mathbf{H}_{2N} = \begin{bmatrix} \mathbf{H}_N & \mathbf{H}_N \\ \mathbf{H}_N & -\mathbf{H}_N \end{bmatrix}$$

Where $\mathbf{H}_{2N}$ is the Hadamard matrix of order 2N and $N = 2^n$

4. Discrete Cosine Transform:

The 1-D discrete cosine transform is defined as

$$C(u) = \alpha(u) \sum_{x=0}^{N-1} f(x) \cos\left[\frac{(2x+1)u\pi}{2N}\right]$$

For u = 0, 1, 2, . . , N-1. Similarly the inverse DCT is defined as

$$f(x) = \sum_{u=0}^{N-1} \alpha(u) C(u) \cos\left[\frac{(2x+1)u\pi}{2N}\right]$$

For u = 0, 1, 2, . . , N-1

Where α is

$$\alpha(u) = \begin{cases} \sqrt{\dfrac{1}{N}} & \text{for } u = 0 \\ \sqrt{\dfrac{2}{N}} & \text{for } u = 1, 2, \ldots, N-1. \end{cases}$$

The corresponding 2-D DCT pair is

$$C(u, v) = \alpha(u)\alpha(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \cos\left[\frac{(2x+1)u\pi}{2N}\right] \cos\left[\frac{(2y+1)v\pi}{2N}\right]$$

For u, v = 0, 1, 2, . . , N-1, and

$$f(x, y) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} \alpha(u)\alpha(v) C(u, v) \cos\left[\frac{(2x+1)u\pi}{2N}\right] \cos\left[\frac{(2y+1)v\pi}{2N}\right]$$

For x, y= 0, 1, 2, . . , N-1

5. Haar Transform.

The Haar transform is based on the Haar functions, $h_k(z)$, which are defined over the continuous, closed interval z ε [0, 1], and for k = 0, 1, 2 . . . , N-1, where $N = 2^n$. The first step in generating the Haar transform is to note that the integer k can be decomposed uniquely as

$$k = 2^p + q - 1$$

where $0 \le p \le n-1$, q = 0 or 1 for p = 0, and $1 \le q \le 2^p$ for $p \ne 0$. For example, if N = 4, k, q, p have following values

k	p	q
0	0	0
1	0	1
2	1	1
3	1	2

The Haar functions are defined as

$$h_0(z) \triangleq h_{00}(z) = \frac{1}{\sqrt{N}} \quad \text{for } z\ \varepsilon\ [0, 1] \ \ldots\ldots\ (1)$$

and

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

These results allow derivation of Haar transformation matrices of order N x N by formation of the *i*th row of a Haar matrix from elements oh $h_i(z)$ for z = 0/N, 1/N, . . . , (N-1)/N. For instance, when N = 2, the first row of the 2 x 2 Haar matrix is computed by using $h_o(z)$ with z = 0/2, 1/2. From equation (1) , $h_o(z)$ is equal to $1/\sqrt{2}$, independent of z, so the first row of the matrix has two identical $1/\sqrt{2}$ elements. Similarly row is computed. The 2 x 2 Haar matrix is

$$\mathbf{A}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Similarly matrix for N = 4 is

$$\mathbf{A}_4 = \frac{1}{\sqrt{4}} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ \sqrt{2} & -\sqrt{2} & 0 & 0 \\ 0 & 0 & \sqrt{2} & -\sqrt{2} \end{bmatrix}$$

Properties of Haar transform:

1. The Haar transform is real and orthogonal.
2. The Haar transform is very fast. It can implement O(n) operations on an N x 1 vector.
3. The mean vectors of the Haar matrix are sequentially ordered.
4. It has a poor energy deal for image.

6. Slant transform:

The Slant transform matrix of order N x N is the recursive expression S_n is given by

$$\frac{1}{\sqrt{2}}\left[\begin{array}{cc|c|cc|c} 1 & 0 & \mathbf{0} & 1 & 0 & \mathbf{0} \\ a_N & b_N & & -a_N & b_N & \\ \hline \mathbf{0} & & \mathbf{I}_{(N/2)-2} & \mathbf{0} & & \mathbf{I}_{(N/2)-2} \\ \hline 0 & 1 & \mathbf{0} & 0 & -1 & \mathbf{0} \\ -b_N & a_N & & b_N & a_N & \\ \hline \mathbf{0} & & \mathbf{I}_{(N/2)-2} & \mathbf{0} & & -\mathbf{I}_{(N/2)-2} \end{array}\right]\left[\begin{array}{c|c} \mathbf{S}_{N/2} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{S}_{N/2} \end{array}\right]$$

Where I_m is the identity matrix of order M x M, and

$$\mathbf{S}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

The coefficients are

$$a_N = \left[\frac{3N^2}{4(N^2-1)}\right]^{1/2}$$

and

$$b_N = \left[\frac{N^2 - 4}{4(N^2 - 1)}\right]^{1/2}$$

The slant transform for N = 4 will be

$$S_4 = \frac{1}{\sqrt{4}}\begin{bmatrix} 1 & 1 & 1 & 1 \\ \frac{3}{\sqrt{5}} & \frac{1}{\sqrt{5}} & \frac{-1}{\sqrt{5}} & \frac{-3}{\sqrt{5}} \\ 1 & -1 & -1 & 1 \\ \frac{1}{\sqrt{5}} & \frac{-3}{\sqrt{5}} & \frac{3}{\sqrt{5}} & \frac{-1}{\sqrt{5}} \end{bmatrix}$$

Properties of Slant transform:

(i) The slant transform is real and orthogonal.

$$S = S^*, \; S^{-1} = S^T$$

(ii) The slant transform is fast, it can be implemented in ($N \log_2 N$) operations on an N x 1 vector.
(iii) The energy deal for images in this transform is rated in very good to excellent range.
(iv) The mean vectors for slant transform matrix S are not sequentially ordered for $n \geq 3$.

7. Hotelling Transform:

The basic principle of hotelling transform is the statistical properties of vector representation. Consider a population of random vectors of the form,

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

And the mean vector of the population is defined as the expected value of x i.e.,

$$m_x = E\{x\}$$

The suffix m represents that the mean is associated with the population of x vectors. The expected value of a vector or matrix is obtained by taking the expected value of each elememt. The covariance matrix C_x in terms of x and m_x is given as

$$C_x = E\{(x-m_x)(x-m_x)^T\}$$

T denotes the transpose operation. Since, x is n dimensional, $\{(x-m_x)(x-m_x)^T\}$ will be of n x n dimension. The covariance matrix is real and symmetric. If elements x_i and x_j are uncorrelated, their covariance is zero and, therefore, $c_{ij} = c_{ji} = 0$.

For M vector samples from a random population, the mean vector and covariance matrix can be approximated from the samples by

$$\mathbf{m}_x = \frac{1}{M}\sum_{k=1}^{M} \mathbf{x}_k$$

and

$$\mathbf{C}_x = \frac{1}{M}\sum_{k=1}^{M} \mathbf{x}_k\mathbf{x}_k^T - \mathbf{m}_x\mathbf{m}_x^T.$$

Applications of digital image processing:

Some of the major fields in which digital image processing is widely used are mentioned below

1. Gamma Ray Imaging- Nuclear medicine and astronomical observations.
2. X-Ray imaging – X-rays of body.
3. Ultraviolet Band –Lithography, industrial inspection, microscopy, lasers.
4. Visual And Infrared Band – Remote sensing.

Microwave Band – Radar imaging.

UNIT-III
IMAGE RESTORATION

3.1 Model of the Image Degradation/Restoration Process:

The Fig. 3.1 shows, the degradation process is modeled as a degradation function that, together with an additive noise term, operates on an input image f(x, y) to produce a degraded image g(x, y). Given g(x, y), some knowledge about the degradation function H, and some knowledge about the additive noise term η(x, y), the objective of restoration is to obtain an estimate *f*(x, y) of the original image. the estimate should be as close as possible to the original input image and, in general, the more we know about H and η, the closer *f*(x, y) will be to f(x, y).

The degraded image is given in the spatial domain by

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

where h (x, y) is the spatial representation of the degradation function and, the symbol * indicates convolution. Convolution in the spatial domain is equal to multiplication in the frequency domain, hence

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

where the terms in capital letters are the Fourier transforms of the corresponding terms in above equation.

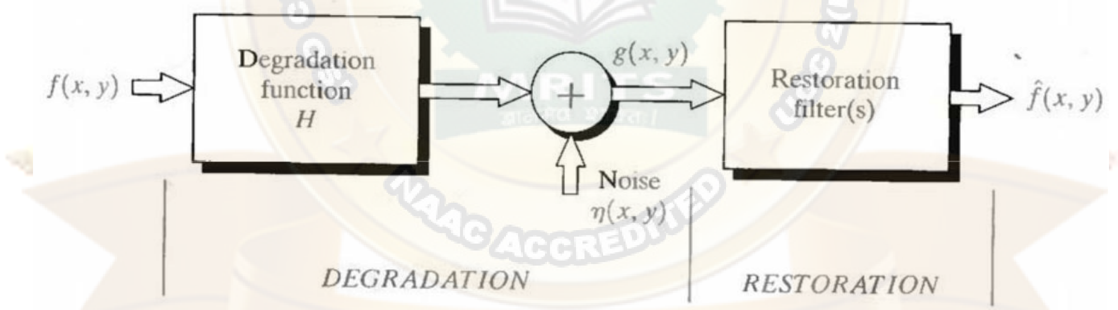


Fig. 3.1 model of the image degradation/restoration process.

3.2 Noise:

The following are among the most common PDFs found in image processing applications.

Gaussian noise

Because of its mathematical tractability in both the spatial and frequency domains, Gaussian (also called normal) noise models are used frequently in practice. In fact, this tractability is so convenient that it often results in Gaussian models being used in situations in which they are marginally applicable at best.

The PDF of a Gaussian random variable, z, is given by

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

… (1)

where z represents gray level, μ is the mean of average value of z, and a σ is its standard deviation. The standard deviation squared, σ^2, is called the variance of z. A plot of this function is shown in Fig. 5.10. When z is described by Eq. (1), approximately 70% of its values will be in the range [(μ - σ), (μ +σ)], and about 95% will be in the range [(μ - 2σ), (μ + 2σ)].

Rayleigh noise

The PDF of Rayleigh noise is given by

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

Erlang (Gamma) noise

The PDF of Erlang noise is given by

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

where the parameters are such that a > 0, b is a positive integer, and "!" indicates factorial.

Exponential noise

The PDF of exponential noise is given by

$$p(z) = \begin{cases} ae^{-az} & \text{for } z \geq 0 \\ 0 & \text{for } z < 0 \end{cases}$$

The mean of this density function is given by

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

$$= 1 / a^2$$

This PDF is a special case of the Erlang PDF, with b = 1.

Uniform noise

The PDF of uniform noise is given by

$$p(z) = \begin{cases} \frac{1}{b-a} & \text{if } a \le z \le b \\ 0 & \text{otherwise.} \end{cases}$$

The mean of this density function is given by

$$\mu = a + b/2$$

$$\sigma^2 = (b - a)^2 / 12$$

Impulse (salt-and-pepper) noise

The PDF of (bipolar) impulse noise is given by

$$p(z) = \begin{cases} P_a & \text{for } z = a \\ P_b & \text{for } z = b \\ 0 & \text{otherwise} \end{cases}$$

If b > a, gray-level b will appear as a light dot in the image. Conversely, level a will appear like a dark dot. If either P_a or P_b is zero, the impulse noise is called unipolar. If neither probability is zero, and especially if they are approximately equal, impulse noise values will resemble salt-and-pepper granules randomly distributed over the image. For this reason, bipolar impulse noise also is called salt-and-pepper noise. Shot and spike noise also are terms used to refer to this type of noise.

3.3 Inverse Filtering

The simplest approach to restoration is direct inverse filtering, where F (u, v), the transform of the original image is computed simply by dividing the transform of the degraded image, G (u, v), by the degradation function

$$\hat{F}(u, v) = \frac{G(u, v)}{H(u, v)}.$$

The divisions are between individual elements of the functions.

But G (u, v) is given by

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

Hence

$$\hat{F}(u, v) = F(u, v) + \frac{N(u, v)}{H(u, v)}.$$

It tells that even if the degradation function is known the undegraded image cannot be recovered [the inverse Fourier transform of F(u, v)] exactly because N(u, v) is a random function whose Fourier transform is not known.

If the degradation has zero or very small values, then the ratio N(u, v)/H(u, v) could easily dominate the estimate F(u, v).

One approach to get around the zero or small-value problem is to limit the filter frequencies to values near the origin. H (0, 0) is equal to the average value of h(x, y) and that this is usually the highest value of H (u, v) in the frequency domain. Thus, by limiting the analysis to frequencies near the origin, the probability of encountering zero values is reduced.

3.4 Wiener Filtering

The inverse filtering approach makes no explicit provision for handling noise. This approach incorporates both the degradation function and statistical characteristics of noise into the restoration process. The method is founded on considering images and noise as random processes, and the objective is to find an estimate *f* of the uncorrupted image f such that the mean square error between them is minimized. This error measure is given by

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

where E{•} is the expected value of the argument. It is assumed that the noise and the image are uncorrelated; that one or the other has zero mean; and that the gray levels in the estimate are a linear function of the levels in the degraded image. Based on these conditions, the minimum of the error function is given in the frequency domain by the expression

$$\begin{aligned}
\hat{F}(u, v) &= \left[\frac{H^*(u, v)S_f(u, v)}{S_f(u, v)|H(u, v)|^2 + S_\eta(u, v)}\right]G(u, v)\\
&= \left[\frac{H^*(u, v)}{|H(u, v)|^2 + S_\eta(u, v)/S_f(u, v)}\right]G(u, v)\\
&= \left[\frac{1}{H(u, v)}\frac{|H(u, v)|^2}{|H(u, v)|^2 + S_\eta(u, v)/S_f(u, v)}\right]G(u, v)
\end{aligned}$$

where we used the fact that the product of a complex quantity with its conjugate is equal to the magnitude of the complex quantity squared. This result is known as the Wiener filter, after N. Wiener [1942], who first proposed the concept in the year shown. The filter, which consists of the terms inside the brackets, also is commonly referred to as the minimum mean square error filter or the least square error filter. The Wiener filter does not have the same problem as the inverse filter with zeros in the degradation function, unless both H(u, v) and $S_\eta(u, v)$ are zero for the same value(s) of u and v.

The terms in above equation are as follows:

H (u, v) = degradation function

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

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

$S_\eta(u, v) = |N(u, v)|^2$ = power spectrum of the noise

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

As before, H (u, v) is the transform of the degradation function and G (u, v) is the transform of the degraded image. The restored image in the spatial domain is given by the inverse Fourier transform of the frequency-domain estimate F (u, v). Note that if the noise is zero, then the noise power spectrum vanishes and the Wiener filter reduces to the inverse filter.

When we are dealing with spectrally white noise, the spectrum $|N(u, v)|^2$ is a constant, which simplifies things considerably. However, the power spectrum of the undegraded image seldom is known. An approach used frequently when these quantities are not known or cannot be estimated is to approximate the equation as

$$\hat{F}(u, v) = \left[\frac{1}{H(u, v)} \frac{|H(u, v)|^2}{|H(u, v)|^2 + K}\right] G(u, v)$$

where K is a specified constant.

3.5 Iterative Restoration

Lucy-Richardson algorithm is a nonlinear restoration method used to recover a latent image which is blurred by a Point Spread Function (psf). It is also known as Richardson-Lucy deconvolution.

With as the point spread function, the pixels in observed image are expressed as,

$$c_i = \sum_j p_{ij} u_j$$

Here,

u_j = Pixel value at location j in the image
c_i = Observed value at i^{th} pixel locacation

The L-R algorithm cannot be used in application in which the psf ***(P_{ij})*** is dependent on one or more unknown variables.

The L-R algorithm is based on maximum- likelihood formulation, in this formulation Poisson statistics are used to model the image. If the likelihood of model is increased, then the result is an equation which satisfies when the following iteration converges.

$$\hat{f}_{k+1}(x, y) = \hat{f}_k(x, y)\left[h(-x, -y)\times\frac{g(x, y)}{h(x, y)\times\hat{f}_k(x, y)}\right]$$

Here,

f = Estimation of undegraded image.

The factor *f* which is present in the right side denominator leads to non-linearity. Since, the algorithm is a type of nonlinear restorations; hence it is stopped when satisfactory result is obtained.

The basic syntax of function deconvlucy with the L-R algorithm is implemented is given below.

fr = Deconvlucy (g, psf, NUMIT, DAMPAR, WEIGHT)

Here the parameters are,

g = Degraded image

f_r = Restored image

psf = Point spread function

NUMIT = Total number of iterations.

The remaining two parameters are,

DAMPAR

The DAMPAR parameter is a scalar parameter which is used to determine the deviation of resultant image with the degraded image (g). The pixels which gel deviated from their original value within the DAMPAR, for these pixels iterations are cancelled so as to reduce noise generation and present essential image information.

WEIGHT

WEIGHT parameter gives a weight to each and every pixel. It is array of size similar to that of degraded image (g). In applications where a pixel leads to improper image is removed by assigning it to a weight as 0'. The pixels may also be given weights depending upon the flat-field correction, which is essential according to image array. Weights are used in applications such as blurring with specified psf. They are used to remove the pixels which are pre9ent at the boundary of the image and are blurred separately by psf.

If the array size of psf is *n x n* then the width of weight of border of zeroes being used is ceil (n / 2)

UNIT-IV
IMAGE SEGMENTATION

4.1 INTRODUCTION:

Gradient operators:

First-order derivatives of a digital image are based on various approximations of the 2-D gradient. The gradient of an image f (x, y) at location (x, y) is defined as the vector

$$\nabla \mathbf{f} = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}.$$

It is well known from vector analysis that the gradient vector points in the direction of maximum rate of change of f at coordinates (x, y). An important quantity in edge detection is the magnitude of this vector, denoted by f, where

$$\nabla f = \text{mag}(\nabla \mathbf{f}) = \left[G_x^2 + G_y^2\right]^{1/2}.$$

This quantity gives the maximum rate of increase of f (x, y) per unit distance in the direction of f. It is a common (although not strictly correct) practice to refer to f also as the gradient. The direction of the gradient vector also is an important quantity. Let α (x, y) represent the direction angle of the vector f at (x, y). Then, from vector analysis,

$$\alpha(x, y) = \tan^{-1}\left(\frac{G_y}{G_x}\right)$$

where the angle is measured with respect to the x-axis. The direction of an edge at (x, y) is perpendicular to the direction of the gradient vector at that point. Computation of the gradient of an image is based on obtaining the partial derivatives f/ x and f/ y at every pixel location. Let the 3x3 area shown in Fig. 1.1 (a) represent the gray levels in a neighborhood of an image. One of the simplest ways to implement a first-order partial derivative at point z_5 is to use the following Roberts cross-gradient operators:

$$G_x = (z_9 - z_5)$$

and

$$G_y = (z_8 - z_6).$$

These derivatives can be implemented for an entire image by using the masks shown in Fig. 4.1(b). Masks of size 2 X 2 are awkward to implement because they do not have a clear center. An approach using masks of size 3 X 3 is given by

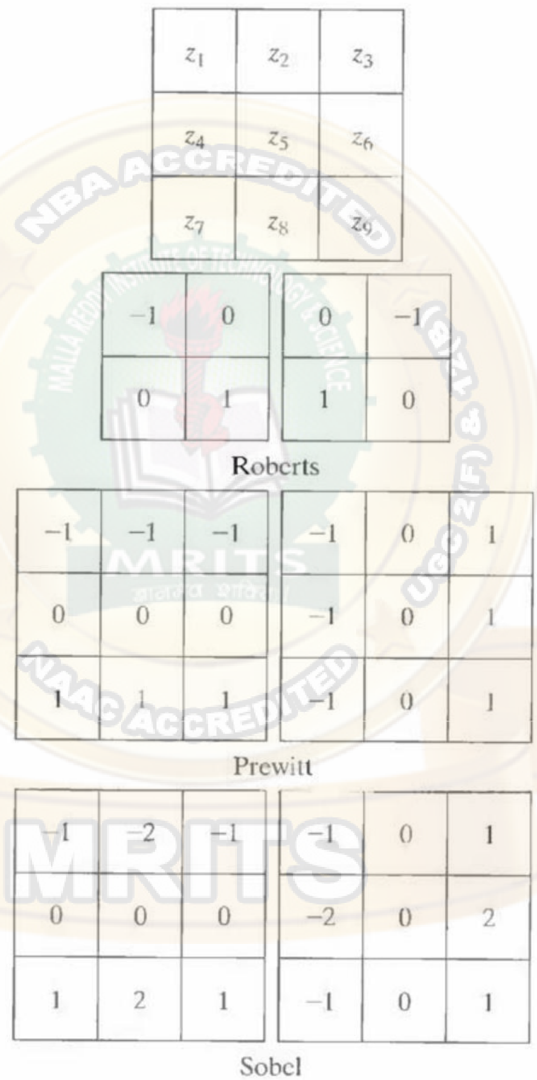


Fig.4.1 A 3 X 3 region of an image (the z's are gray-level values) and various masks used to compute the gradient at point labeled z_5.

$$G_x = (z_7 + z_8 + z_9) - (z_1 + z_2 + z_3)$$

and

$$G_y = (z_3 + z_6 + z_9) - (z_1 + z_4 + z_7).$$

A weight value of 2 is used to achieve some smoothing by giving more importance to the center point. Figures 4.1(f) and (g), called the Sobel operators, and are used to implement these two equations. The Prewitt and Sobel operators are among the most used in practice for computing digital gradients. The Prewitt masks are simpler to implement than the Sobel masks, but the latter have slightly superior noise-suppression characteristics, an important issue when dealing with derivatives. Note that the coefficients in all the masks shown in Fig. 1.1 sum to 0, indicating that they give a response of 0 in areas of constant gray level, as expected of a derivative operator.

The masks just discussed are used to obtain the gradient components G_x and G_y. Computation of the gradient requires that these two components be combined. However, this implementation is not always desirable because of the computational burden required by squares and square roots. An approach used frequently is to approximate the gradient by absolute values:

$$\nabla f \approx |G_x| + |G_y|.$$

This equation is much more attractive computationally, and it still preserves relative changes in gray levels. However, this is not an issue when masks such as the Prewitt and Sobel masks are used to compute Gx and Gy.

It is possible to modify the 3 X 3 masks in Fig. 4.1 so that they have their strongest responses along the diagonal directions. The two additional Prewitt and Sobel masks for detecting discontinuities in the diagonal directions are shown in Fig. 4.2.

4

0	1	1		−1	−1	0
−1	0	1		−1	0	1
−1	−1	0		0	1	1

Prewitt

0	1	2		−2	−1	0
−1	0	1		−1	0	1
−2	−1	0		0	1	2

Sobel

Fig.4.2 Prewitt and Sobel masks for detecting diagonal edges

The Laplacian:

The Laplacian of a 2-D function f(x, y) is a second-order derivative defined as

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

For a 3 X 3 region, one of the two forms encountered most frequently in practice is

$$\nabla^2 f = 4z_5 - (z_2 + z_4 + z_6 + z_8)$$

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

Fig.4.3 Laplacian masks used to implement Eqns. above.

where the z's are defined in Fig. 1.1(a). A digital approximation including the diagonal neighbors is given by

$$\nabla^2 f = 8z_5 - (z_1 + z_2 + z_3 + z_4 + z_6 + z_7 + z_8 + z_9).$$

Masks for implementing these two equations are shown in Fig. 1.3. We note from these masks that the implementations of Eqns. are isotropic for rotation increments of 90° and 45°, respectively.

4.2 EDGE DETECTION:

Intuitively, an edge is a set of connected pixels that lie on the boundary between two regions. Fundamentally, an edge is a "local" concept whereas a region boundary, owing to the way it is defined, is a more global idea. A reasonable definition of "edge" requires the ability to measure gray-level transitions in a meaningful way. We start by modeling an edge intuitively. This will lead us to formalism in which "meaningful" transitions in gray levels can be measured. Intuitively, an ideal edge has the properties of the model shown in Fig. 4.3(a). An ideal edge according to this model is a set of connected pixels (in the vertical direction here), each of which is located at an orthogonal step transition in gray level (as shown by the horizontal profile in the figure).

In practice, optics, sampling, and other image acquisition imperfections yield edges that are blurred, with the degree of blurring being determined by factors such as the quality of the image acquisition system, the sampling rate, and illumination conditions under which the image is acquired. As a result, edges are more closely modeled as having a "ramp like" profile, such as the one shown in Fig.4.3 (b).

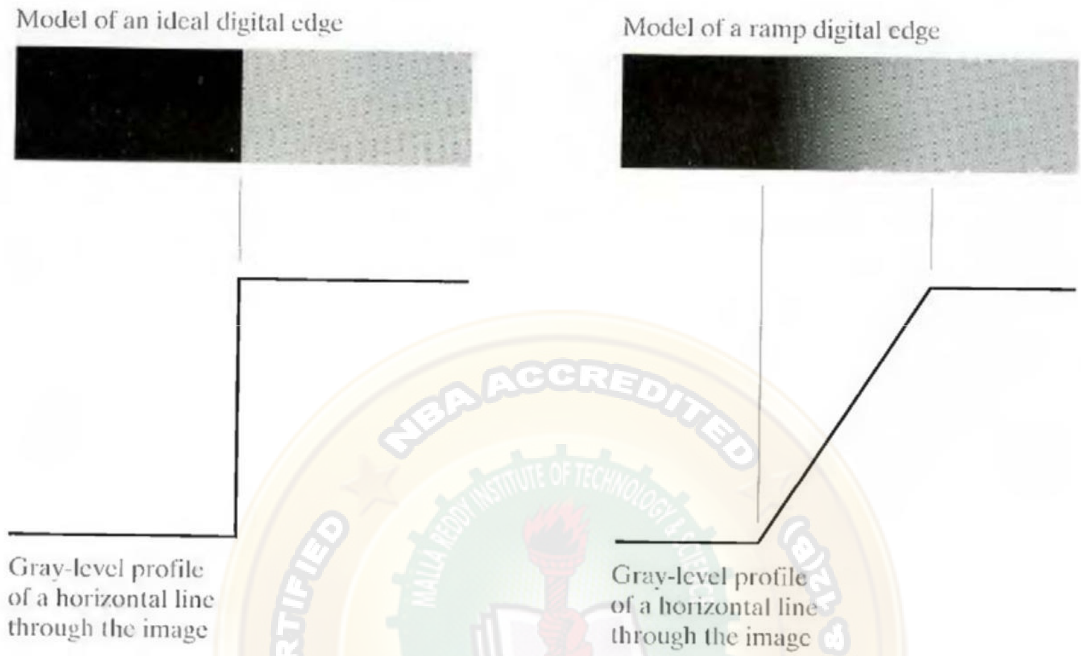


Fig.4.3 (a) Model of an ideal digital edge (b) Model of a ramp edge. The slope of the ramp is proportional to the degree of blurring in the edge.

The slope of the ramp is inversely proportional to the degree of blurring in the edge. In this model, we no longer have a thin (one pixel thick) path. Instead, an edge point now is any point contained in the ramp, and an edge would then be a set of such points that are connected. The "thickness" of the edge is determined by the length of the ramp, as it transitions from an initial to a final gray level. This length is determined by the slope, which, in turn, is determined by the degree of blurring. This makes sense: Blurred edges lend to be thick and sharp edges tend to be thin. Figure 4.4(a) shows the image from which the close-up in Fig. 4.3(b) was extracted. Figure 4.4(b) shows a horizontal gray-level profile of the edge between the two regions. This figure also shows the first and second derivatives of the gray-level profile. The first derivative is positive at the points of transition into and out of the ramp as we move from left to right along the profile; it is constant for points in the ramp; and is zero in areas of constant gray level. The second derivative is positive at the transition associated with the dark side of the edge, negative at the transition associated with the light side of the edge, and zero along the ramp and in areas of constant gray level. The signs of the derivatives in Fig. 4.4(b) would be reversed for an edge that transitions from light to dark.

We conclude from these observations that the magnitude of the first derivative can be used to detect the presence of an edge at a point in an image (i.e. to determine if a point is on a ramp). Similarly, the sign of the second derivative can be used to determine whether an edge pixel lies on the dark or light side of an edge. We note two additional properties of the second derivative around an edge: A) It produces two values for every edge in an image (an undesirable feature); and B) an imaginary straight line joining the extreme positive and negative values of the second derivative would cross zero near the midpoint of the edge. This zero-crossing property of the second derivative is quite useful for locating the centers of thick edges.

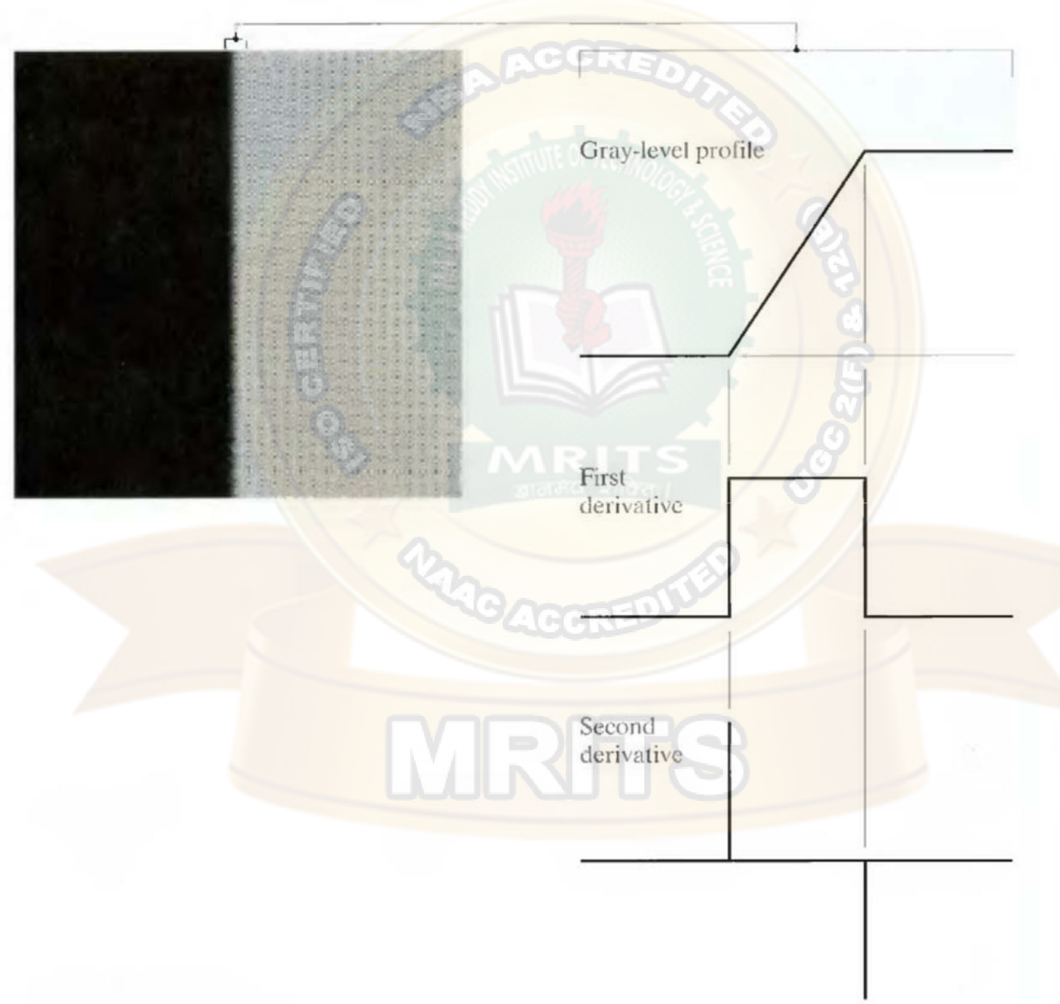


Fig.4.4 (a) Two regions separated by a vertical edge (b) Detail near the edge, showing a gray-level profile, and the first and second derivatives of the profile.

4.3 EDGE LINKING PROCEDURES:

The different methods for edge linking are as follows

(i) Local processing

(ii) Global processing via the Hough Transform

(iii) Global processing via graph-theoretic techniques.

(i) Local Processing:

One of the simplest approaches for linking edge points is to analyze the characteristics of pixels in a small neighborhood (say, 3 X 3 or 5 X 5) about every point (x, y) in an image that has been labeled an edge point. All points that are similar according to a set of predefined criteria are linked, forming an edge of pixels that share those criteria.

The two principal properties used for establishing similarity of edge pixels in this kind of analysis are (1) the strength of the response of the gradient operator used to produce the edge pixel; and (2) the direction of the gradient vector. The first property is given by the value of f.

Thus an edge pixel with coordinates (x_0, y_0) in a predefined neighborhood of (x, y), is similar in magnitude to the pixel at (x, y) if

$$|\nabla f(x, y) - \nabla f(x_0, y_0)| \leq E$$

The direction (angle) of the gradient vector is given by Eq. An edge pixel at (x_0, y_0) in the predefined neighborhood of (x, y) has an angle similar to the pixel at (x, y) if

$$|\alpha(x, y) - \alpha(x_0, y_0)| < A$$

where A is a nonnegative angle threshold. The direction of the edge at (x, y) is perpendicular to the direction of the gradient vector at that point.

A point in the predefined neighborhood of (x, y) is linked to the pixel at (x, y) if both magnitude and direction criteria are satisfied. This process is repeated at every location in the image. A record must be kept of linked points as the center of the neighborhood is moved from pixel to pixel. A simple bookkeeping procedure is to assign a different gray level to each set of linked edge pixels.

(ii) Global processing via the Hough Transform:

In this process, points are linked by determining first if they lie on a curve of specified shape. We now consider global relationships between pixels. Given n points in an image, suppose that we want to find subsets of these points that lie on straight lines. One possible solution is to first find all lines determined by every pair of points and then find all subsets of points that are close to particular lines. The problem with this procedure is that it involves finding $n(n - 1)/2 \sim n^2$ lines and then performing $(n)(n(n - l))/2 \sim n^3$ comparisons of every point to all lines. This approach is computationally prohibitive in all but the most trivial applications.

Hough [1962] proposed an alternative approach, commonly referred to as the Hough transform. Consider a point (x_i, y_i) and the general equation of a straight line in slope-intercept form, $y_i = a.x_i + b$. Infinitely many lines pass through (x_i, y_i) but they all satisfy the equation $y_i = a.x_i + b$ for varying values of a and b. However, writing this equation as $b = -a.x_i + y_i$, and considering the ab-plane (also called parameter space) yields the equation of a single line for a fixed pair (x_i, y_i). Furthermore, a second point (x_j, y_j) also has a line in parameter space associated with it, and this line intersects the line associated with (x_i, y_i) at (a', b'), where a' is the slope and b' the intercept of the line containing both (x_i, y_i) and (x_j, y_j) in the xy-plane. In fact, all points contained on this line have lines in parameter space that intersect at (a', b'). Figure 4.5 illustrates these concepts.

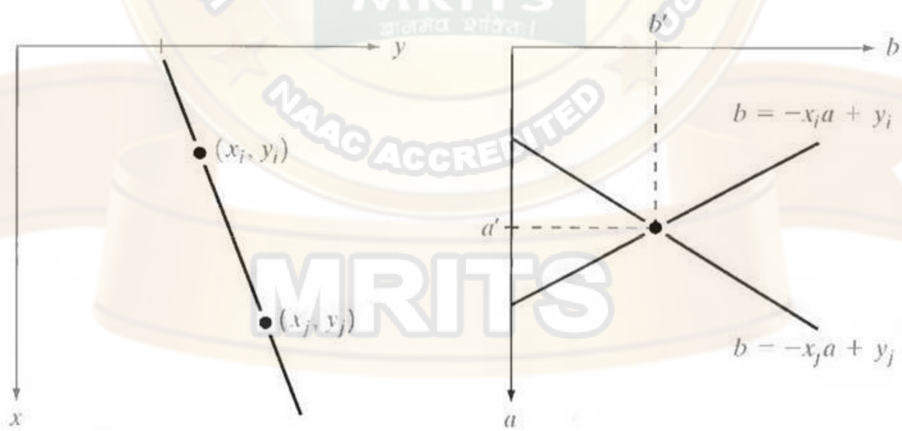


Fig.4.5 (a) xy-plane (b) Parameter space

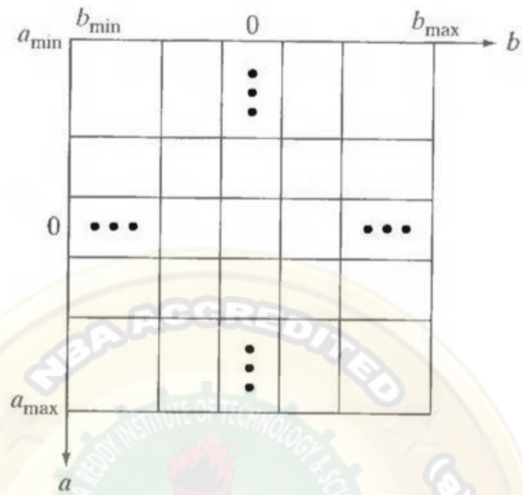


Fig.4.6 Subdivision of the parameter plane for use in the Hough transform

The computational attractiveness of the Hough transform arises from subdividing the parameter space into so-called accumulator cells, as illustrated in Fig. 4.6, where (a_{max} , a_{min}) and (b_{max} , b_{min}), are the expected ranges of slope and intercept values. The cell at coordinates (i, j), with accumulator value A(i, j), corresponds to the square associated with parameter space coordinates (a_i , b_i).

Initially, these cells are set to zero. Then, for every point (x_k, y_k) in the image plane, we let the parameter a equal each of the allowed subdivision values on the fl-axis and solve for the corresponding b using the equation $b = - x_k a + y_k$ .The resulting b's are then rounded off to the nearest allowed value in the b-axis. If a choice of a_p results in solution b_q, we let A (p, q) = A (p, q) + 1. At the end of this procedure, a value of Q in A (i, j) corresponds to Q points in the xy-plane lying on the line $y = a_i x + b_j$. The number of subdivisions in the ab-plane determines the accuracy of the co linearity of these points. Note that subdividing the a axis into K increments gives, for every point (x_k, y_k), K values of b corresponding to the K possible values of a. With n image points, this method involves nK computations. Thus the procedure just discussed is linear in n, and the product nK does not approach the number of computations discussed at the beginning unless K approaches or exceeds n.

A problem with using the equation y = ax + b to represent a line is that the slope approaches infinity as the line approaches the vertical. One way around this difficulty is to use the normal representation of a line:

$$x \cos\theta + y \sin\theta = \rho$$

Figure 4.7(a) illustrates the geometrical interpretation of the parameters used. The use of this representation in constructing a table of accumulators is identical to the method discussed for the slope-intercept representation. Instead of straight lines, however, the loci are sinusoidal curves in the ρθ -plane. As before, Q collinear points lying on a line $x \cos\theta_j + y \sin\theta_j = \rho$, yield Q sinusoidal curves that intersect at (p_i, θ_j) in the parameter space. Incrementing θ and solving for the corresponding p gives Q entries in accumulator A (i, j) associated with the cell determined by (p_i, θ_j). Figure 4.7 (b) illustrates the subdivision of the parameter space.

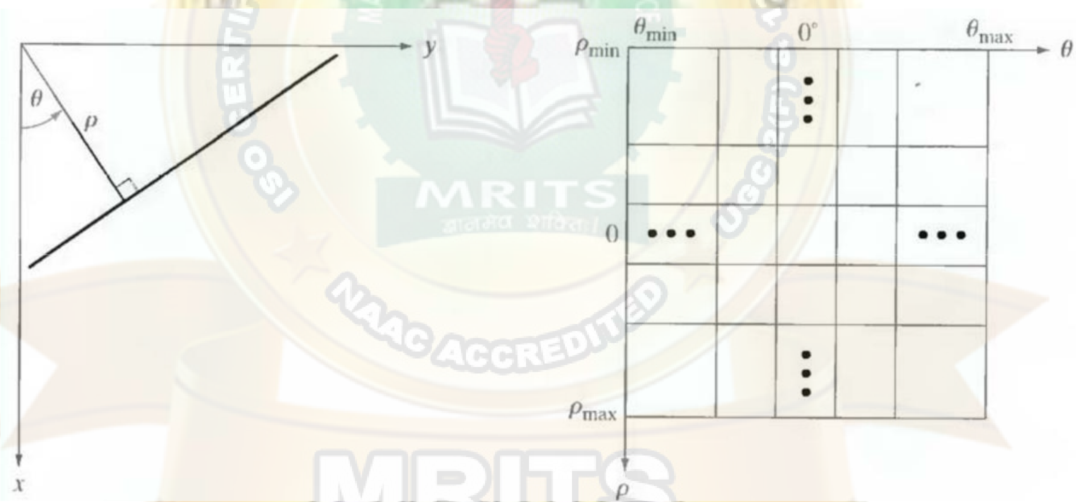


Fig.4.7 (a) Normal representation of a line (b) Subdivision of the ρθ-plane into cells

The range of angle θ is ±90°, measured with respect to the x-axis. Thus with reference to Fig. 4.7 (a), a horizontal line has θ = 0°, with ρ being equal to the positive x-intercept. Similarly, a vertical line has θ = 90°, with p being equal to the positive y-intercept, or θ = - 90°, with ρ being equal to the negative y-intercept.

(iii) Global processing via graph-theoretic techniques

In this process we have a global approach for edge detection and linking based on representing edge segments in the form of a graph and searching the graph for low-cost paths that correspond to significant edges. This representation provides a rugged approach that performs well in the presence of noise.

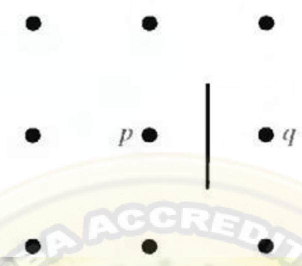


Fig.4.8 Edge clement between pixels p and q

We begin the development with some basic definitions. A graph G = (N,U) is a finite, nonempty set of nodes N, together with a set U of unordered pairs of distinct elements of N. Each pair (n_i, n_j) of U is called an arc. A graph in which the arcs are directed is called a directed graph. If an arc is directed from node n_i to node n_j, then n_j is said to be a successor of the parent node n_i. The process of identifying the successors of a node is called expansion of the node. In each graph we define levels, such that level 0 consists of a single node, called the start or root node, and the nodes in the last level are called goal nodes. A cost $c(n_i, n_j)$ can be associated with every arc (n_i, n_j). A sequence of nodes $n_1, n_2 \ldots n_k$, with each node n_i being a successor of node n_{i-1} is called a path from n_1 to n_k. The cost of the entire path is

$$c = \sum_{i=2}^{k} c(n_{i-1}, n_i).$$

The following discussion is simplified if we define an edge element as the boundary between two pixels p and q, such that p and q are 4-neighbors, as Fig.4.8 illustrates. Edge elements are identified by the xy-coordinates of points p and q. In other words, the edge element in Fig. 3.4 is defined by the pairs (x_p, y_p) (x_q, y_q). Consistent with the definition an edge is a sequence of connected edge elements.

We can illustrate how the concepts just discussed apply to edge detection using the 3 X 3 image shown in Fig. 4.9 (a). The outer numbers are pixel

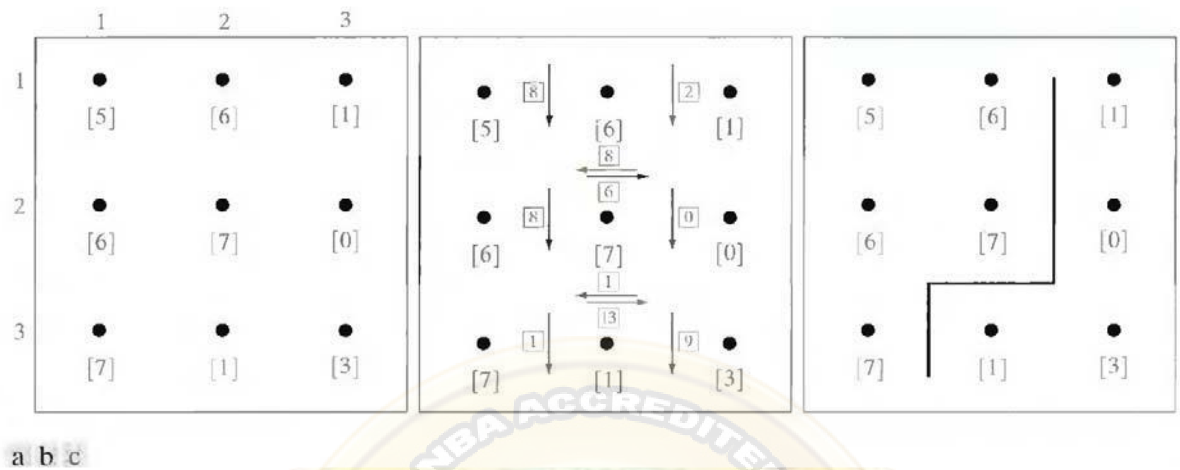


Fig.4.9 (a) A 3 X 3 image region, (b) Edge segments and their costs, (c) Edge corresponding to the lowest-cost path in the graph shown in Fig. 3.6

coordinates and the numbers in brackets represent gray-level values. Each edge element, defined by pixels p and q, has an associated cost, defined as

$$c(p,q) = H - [f(p) - f(q)]$$

where H is the highest gray-level value in the image (7 in this case), and f(p) and f(q) are the gray-level values of p and q, respectively. By convention, the point p is on the right-hand side of the direction of travel along edge elements. For example, the edge segment (1, 2) (2, 2) is between points (1, 2) and (2, 2) in Fig. 3.5 (b). If the direction of travel is to the right, then p is the point with coordinates (2, 2) and q is point with coordinates (1, 2); therefore, c (p, q) = 7 - [7 - 6] = 6. This cost is shown in the box below the edge segment. If, on the other hand, we are traveling to the left between the same two points, then p is point (1, 2) and q is (2, 2). In this case the cost is 8, as shown above the edge segment in Fig. 4.9(b). To simplify the discussion, we assume that edges start in the top row and terminate in the last row, so that the first element of an edge can be only between points (1, 1), (1, 2) or (1, 2), (1, 3). Similarly, the last edge element has to be between points (3, 1), (3, 2) or (3, 2), (3, 3). Keep in mind that p and q are 4-neighbors, as noted earlier. Figure 3.6 shows the graph for this problem. Each node (rectangle) in the graph corresponds to an edge element from Fig. 4.9. An arc exists between two nodes if the two corresponding edge elements taken in succession can be part of an edge.

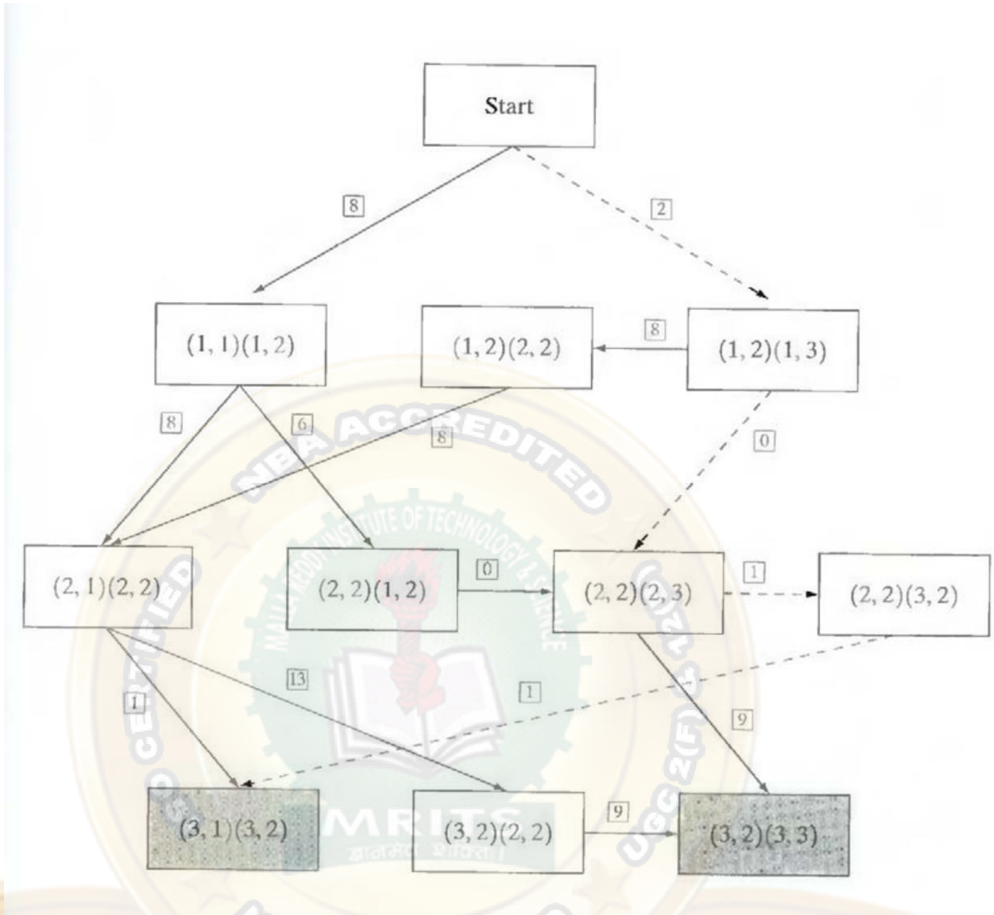


Fig. 4.10 Graph for the image in Fig.3.5 (a). The lowest-cost path is shown dashed.

As in Fig. 4.9 (b), the cost of each edge segment, is shown in a box on the side of the arc leading into the corresponding node. Goal nodes are shown shaded. The minimum cost path is shown dashed, and the edge corresponding to this path is shown in Fig. 4.9 (c).

4.4 THRESHOLDING:

Because of its intuitive properties and simplicity of implementation, image thresholding enjoys a central position in applications of image segmentation.

Global Thresholding:

The simplest of all thresholding techniques is to partition the image histogram by using a single global threshold, T. Segmentation is then accomplished by scanning the image pixel by pixel and labeling each pixel as object or back-ground, depending on whether the gray level of that pixel is greater or less than the value of T. As indicated earlier, the success of this method depends entirely on how well the histogram can be partitioned.

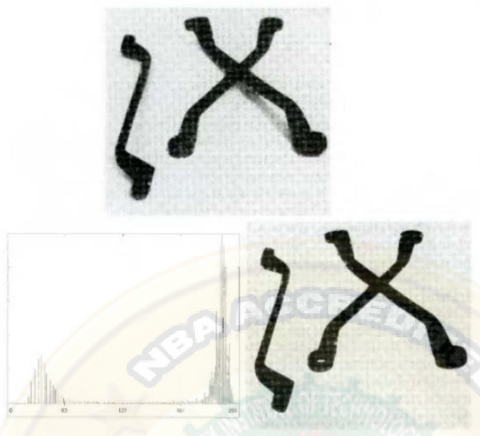


Fig.4.1 FIGURE 10.28 (a) Original image, (b) Image histogram, (c) Result of global thresholding with T midway between the maximum and minimum gray levels.

Figure 4.1(a) shows a simple image, and Fig. 4.1(b) shows its histogram. Figure 4.1(c) shows the result of segmenting Fig. 4.1(a) by using a threshold T midway between the maximum and minimum gray levels. This threshold achieved a "clean" segmentation by eliminating the shadows and leaving only the objects themselves. The objects of interest in this case are darker than the background, so any pixel with a gray level ≤ T was labeled black (0), and any pixel with a gray level ≥ T was labeled white (255).The key objective is merely to generate a binary image, so the black-white relationship could be reversed. The type of global thresholding just described can be expected to be successful in highly controlled environments. One of the areas in which this often is possible is in industrial inspection applications, where control of the illumination usually is feasible.

The threshold in the preceding example was specified by using a heuristic approach, based on visual inspection of the histogram. The following algorithm can be used to obtain T automatically:

1. Select an initial estimate for T.

2. Segment the image using T. This will produce two groups of pixels: G_1 consisting of all pixels with gray level values >T and G_2 consisting of pixels with values < T.

3. Compute the average gray level values μ_1 and μ_2 for the pixels in regions G_1 and G_2.

4. Compute a new threshold value:

$$T = \frac{1}{2}(\mu_1 + \mu_2).$$

5. Repeat steps 2 through 4 until the difference in T in successive iterations is smaller than a predefined parameter T_o.

When there is reason to believe that the background and object occupy comparable areas in the image, a good initial value for T is the average gray level of the image. When objects are small compared to the area occupied by the background (or vice versa), then one group of pixels will dominate the histogram and the average gray level is not as good an initial choice. A more appropriate initial value for T in cases such as this is a value midway between the maximum and minimum gray levels. The parameter T_o is used to stop the algorithm after changes become small in terms of this parameter. This is used when speed of iteration is an important issue.

4.5. REGION BASED SEGMENTATION:

Region-Based Segmentation:

The objective of segmentation is to partition an image into regions. We approached this problem by finding boundaries between regions based on discontinuities in gray levels, whereas segmentation was accomplished via thresholds based on the distribution of pixel properties, such as gray-level values or color.

Basic Formulation:

Let R represent the entire image region. We may view segmentation as a process that partitions R into n subregions, R_1, R_2..., R_n, such that

(a) $\bigcup_{i=1}^{n} R_i = R$.

(b) R_i is a connected region, $i = 1, 2, \ldots, n$.

(c) $R_i \cap R_j = \varnothing$ for all i and j, $i \neq j$.

(d) $P(R_i) = \text{TRUE}$ for $i = 1, 2, \ldots, n$.

(e) $P(R_i \cup R_j) = \text{FALSE}$ for $i \neq j$.

Here, P (R_i) is a logical predicate defined over the points in set R_i and Ǿ` is the null set. Condition (a) indicates that the segmentation must be complete; that is, every pixel must be in a region. Condition (b) requires that points in a region must be connected in some predefined sense. Condition (c) indicates that the regions must be disjoint. Condition (d) deals with the properties that must be satisfied by the pixels in a segmented region—for example P (R_i) = TRUE if all pixels in R_i, have the same gray level. Finally, condition (c) indicates that regions R_i and R_j are different in the sense of predicate P.

Region Growing:

As its name implies, region growing is a procedure that groups pixels or subregions into larger regions based on predefined criteria. The basic approach is to start with a set of "seed" points and from these grow regions by appending to each seed those neighboring pixels that have properties similar to the seed (such as specific ranges of gray level or color). When a priori information is not available, the procedure is to compute at every pixel the same set of properties that ultimately will be used to assign pixels to regions during the growing process. If the result of these computations shows clusters of values, the pixels whose properties place them near the centroid of these clusters can be used as seeds.

The selection of similarity criteria depends not only on the problem under consideration, but also on the type of image data available. For example, the analysis of land-use satellite imagery depends heavily on the use of color. This problem would be significantly more difficult, or even impossible, to handle without the inherent information available in color images. When the images are monochrome, region analysis must be carried out with a set of descriptors based on gray levels and spatial properties (such as moments or texture).

Basically, growing a region should stop when no more pixels satisfy the criteria for inclusion in that region. Criteria such as gray level, texture, and color, are local in nature and do not take into account the "history" of region growth. Additional criteria that increase the power of a region-growing algorithm utilize the concept of size, likeness between a candidate pixel and the pixels grown so far (such as a comparison of the gray level of a candidate and the average gray level of the grown region), and the shape of the region being grown. The use of these types of descriptors is based on the assumption that a model of expected results is at least partially available.

Figure 7.1 (a) shows an X-ray image of a weld (the horizontal dark region) containing several cracks and porosities (the bright, white streaks running horizontally through the middle of the image). We wish to use region growing to segment the regions of the weld failures. These segmented features could be used for inspection, for inclusion in a database of historical studies, for controlling an automated welding system, and for other numerous applications.

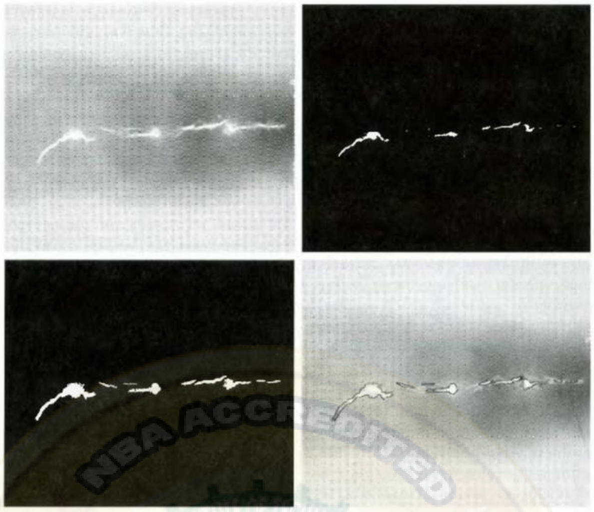


Fig.7.1 (a) Image showing defective welds, (b) Seed points, (c) Result of region growing, (d) Boundaries of segmented ; defective welds (in black).

The first order of business is to determine the initial seed points. In this application, it is known that pixels of defective welds tend to have the maximum allowable digital value B55 in this case). Based on this information, we selected as starting points all pixels having values of 255. The points thus extracted from the original image are shown in Fig. 10.40(b). Note that many of the points are clustered into seed regions.

The next step is to choose criteria for region growing. In this particular example we chose two criteria for a pixel to be annexed to a region: (1) The absolute gray-level difference between any pixel and the seed had to be less than 65. This number is based on the histogram shown in Fig. 7.2 and represents the difference between 255 and the location of the first major valley to the left, which is representative of the highest gray level value in the dark weld region. (2) To be included in one of the regions, the pixel had to be 8-connected to at least one pixel in that region.

If a pixel was found to be connected to more than one region, the regions were merged. Figure 7.1 (c) shows the regions that resulted by starting with the seeds in Fig. 7.2 (b) and utilizing the criteria defined in the previous paragraph. Superimposing the boundaries of these regions on the original image [Fig. 7.1(d)] reveals that the region-growing procedure did indeed segment the defective welds with an acceptable degree of accuracy. It is of interest to note that it was not necessary to specify any stopping rules in this case because the criteria for region growing were sufficient to isolate the features of interest.

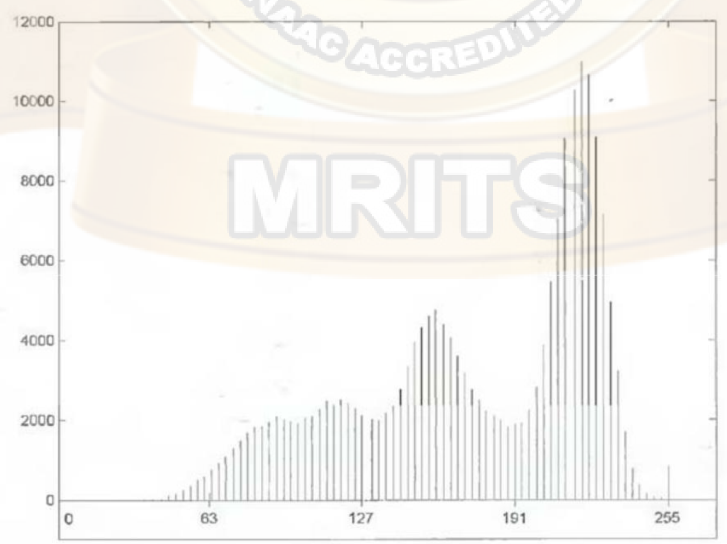


Fig.7.2 Histogram of Fig. 7.1 (a)

Region Splitting and Merging:

The procedure just discussed grows regions from a set of seed points. An alternative is to subdivide an image initially into a set of arbitrary, disjointed regions and then merge and/or split the regions in an attempt to satisfy the conditions. A split and merge algorithm that iteratively works toward satisfying these constraints is developed.

Let R represent the entire image region and select a predicate P. One approach for segmenting R is to subdivide it successively into smaller and smaller quadrant regions so that, for any region R_i, $P(R_i)$ = TRUE. We start with the entire region. If P(R) = FALSE, we divide the image into quadrants. If P is FALSE for any quadrant, we subdivide that quadrant into subquadrants, and so on. This particular splitting technique has a convenient representation in the form of a so-called quadtree (that is, a tree in which nodes have exactly four descendants), as illustrated in Fig. 7.3. Note that the root of the tree corresponds to the entire image and that each node corresponds to a subdivision. In this case, only R_4 was subdivided further.

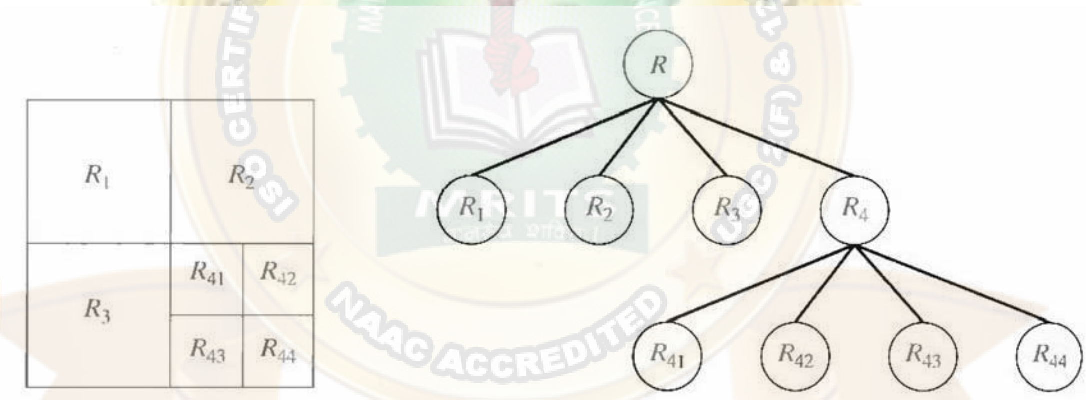


Fig. 7.3 (a) Partitioned image (b) Corresponding quadtree.

If only splitting were used, the final partition likely would contain adjacent regions with identical properties. This drawback may be remedied by allowing merging, as well as splitting. Satisfying the constraints, requires merging only adjacent regions whose combined pixels satisfy the predicate P. That is, two adjacent regions R_j and R_k are merged only if $P(R_j \cup R_k)$ = TRUE.

The preceding discussion may be summarized by the following procedure, in which, at any step we

1. Split into four disjoint quadrants any region R_i, for which $P(R_i)$ = FALSE.

2. Merge any adjacent regions R_j and R_k for which $P(R_j \cup R_k)$ = TRUE.

3. Stop when no further merging or splitting is possible.

Several variations of the preceding basic theme are possible. For example, one possibility is to split the image initially into a set of blocks. Further splitting is carried out as described previously, but merging is initially limited to groups of four blocks that are descendants in the quadtree representation and that satisfy the predicate P. When no further mergings of this type are possible, the procedure is terminated by one final merging of regions satisfying step 2. At this point, the merged regions may be of different sizes. The principal advantage of this approach is that it uses the same quadtree for splitting and merging, until the final merging step.

MORPHOLOGICAL IMAGE PROCESSING

Morphology

- Identification, analysis, and description of the structure of the smallest unit of words
- Theory and technique for the analysis and processing of geometric structures
  - Based on set theory, lattice theory, topology, and random functions
  - Extract image components useful in the representation and description of region shape such as boundaries, skeletons, and convex hull
  - Input in the form of images, output in the form of attributes extracted from those images
  - Attempt to extract the *meaning* of the images

Preliminaries

- Set theory in the context of image processing
  - Sets of pixels represent objects in the image
  - Set of all white pixels in a binary image is a complete morphological description of the image
- Sets in binary images
  - Members of the 2D integer space Z^2
  - Each element of the set is a 2-tuple whose coordinates are the (x, y) coordinates of a white pixel in the image
  - Gray scale images can be represented as a set of 3-tuples in Z^3
  - Set reflection $\hat{B}$

$$\hat{B} = \{ w | w = -b, \text{ for } b \in B \}$$

    - In binary image, $\hat{B}$ is the set of points in B whose (x,y) coordinates have been replaced by $(-x,-y)$
    - Figure 9.1a
  - Set translation
    - Translation of a set B by point $z = (z_1, z_2)$ is denoted by $(B)_z$

$$(B)_z = \{ c | c = b + z, \text{ for } b \in B \}$$

  * In binary image, $(B)_z$ is the set of points in B whose (x, y) coordinates have been replaced by $(x + z_1, y + z_2)$
  * Figure 9.1c
- Set reflection and set translation are used to formulate operations based on so-called *structuring elements*
  * Small sets or subimages used to probe an image for properties of interest
  * Figure 9.2
  * Preference for ses to be rectangular arrays
  * Some locations are such that it does not matter whether they are part of the se
    · Such locations are flagged by × in the se
  * The origin of the se must also be specified
    · Indicated by • in Figure 9.2
    · If se is symmetric and no • is shown, the origin is assumed to be at the center of se
  * Figure 9.3 – A simple set A and an se B
  * Convert A to a rectangular array by adding background elements
  * Make background border large enough to accommodate the entire se when the origin is on the border of original A
  * Fill in the se with the smallest number of background elements to make it a rectangular array
  * Operation of set A using se B
    · Create a new set by running B over A
    · Origin of B visits every element of A
    · If B is completely contained in A, mark that location as a member of the new set; else it is not a member of the new set
    · Results in *eroding* the boundary of A

Erosion and dilation

- Erosion
  - With A and B as sets in Z^2, erosion of A by B, denoted by $A \text{ g } B$ is defined as

$$A \text{ g } B = \{ z \mid (B)_z \subseteq A \}$$

  - Set of all points z such that B, translated by z, is contained in A
  - B does not share any common elements with the background

$$A \text{ g } B = \{ z \mid (B)_z \cap A^c = \varnothing$$

  - Figure 9.4
  - Example: Figure 9.5
    * Erosion shrinks or thins objects in a binary image
    * Morphological filter in which image details smaller than the se are filtered/removed from the image
- Dilation
  - With A and B as sets in Z^2, dilation of A by B, denoted by $A \oplus B$ is defined as

$$A \oplus B = \{ z \mid (\hat{B})_z \cap A \neq \varnothing \}$$

- Reflect B about the origin, and shift the reflection by z
- Dilation is the set of all displacements z such that B and A overlap by at least one element
- An equivalent formulation is

$$A \oplus B = \{ z \mid [(\hat{B})_z \cap A] \subseteq A \}$$

- Grows or thickens objects in a binary image
- Figure 9.6
- Example: Figure 9.7
  - ∗ Bridging gaps in broken characters
  - ∗ Lowpass filtering produces a grayscale image; morphological operation produces a binary image

- Erosion and dilation are based on set operations and therefore, are nonlinear
- Duality

- Erosion and dilation are duals of each other with respect to set complementation and reflection

$$(A \,\mathrm{g}\, B)^c \;=\; A^c \oplus \hat{B}$$
$$(A \oplus B)^c \;=\; A^c \,\mathrm{g}\, \hat{B}$$

- Duality property is especially useful when se is symmetric with respect to its origin so that $\hat{B} = B$
  - Allows for erosion of an image by dilating its background (A^c) using the same se and complementing the results
- Proving duality
  - Definition for erosion can be written as

$$(A \,\mathrm{g}\, B)^c = \{z \mid (B)_z \subseteq A\}^c$$

  - $(B)_z \subseteq A \Rightarrow (B)_z \cap A^c = \emptyset$
  - So, the previous expression yields

$$(A \,\mathrm{g}\, B)^c = \{z \mid (B)_z \cap A^c = \emptyset\}^c$$

  - The complement of the set of z's that satisfy $(B)_z \cap A^c = \emptyset$ is the set of z's such that $(B)_z \cap A^c \neq \emptyset$
  - This leads to

$$(A \,\mathrm{g}\, B)^c \;=\; \{z \mid (B)_z \cap A^c \neq \emptyset\}$$
$$=\; A^c \oplus \hat{B}$$

Opening and closing

- Opening smoothes the contours of an object, breaks narrow isthmuses, and eliminates thin protrusions
- Closing smoothes sections of contours, fusing narrow breaks and long thin gulfs, eliminates small holes, and fills gaps in the contour
- Opening of a set A by se B, denoted by $A \circ B$, is defined by

$$A \circ B = (A \,\mathrm{g}\, B) \oplus B$$

- Closing of a set A by se B, denoted by $A \bullet B$, is defined by

$$A \bullet B = (A \oplus B) \,\mathrm{g}\, B$$

- Geometric interpretation of opening expressed as a fitting process such that

$$A \circ B = \bigcup \{(B)_z \mid (B)_z \subseteq A\}$$

  - Union of all translates of B that fit into A
  - Figure 9.8
- Similar interpretation of closing in Figure 9.9
- Example – Figure 9.10
- Duality property

$$(A \bullet B)^c = (A^c \circ \hat{B})$$

$$(A \circ B)^c = (A^c \bullet \hat{B})$$

- Opening operation satisfies the following properties
  1. $A \circ B \subseteq A$
  2. $C \subseteq D \Rightarrow C \circ B \subseteq D \circ B$
  3. $(A \circ B) \circ B = A \circ B$
- Similarly, closing operation satisfies
  1. $A \subseteq A \bullet B$
  2. $C \subseteq D \Rightarrow C \bullet B \subseteq D \bullet B$
  3. $(A \bullet B) \bullet B = A \bullet B$
  - In both the above cases, multiple application of opening and closing has no effect after the first application
- Example: Removing noise from fingerprints
  - Figure 9.11
  - Noise as random light elements on a dark background

Hit-or-miss transformation

- Basic tool for shape detection in a binary image
  - Uses the morphological erosion operator and a pair of disjoint ses
  - First se fits in the foreground of input image; second se misses it completely
  - The pair of two ses is called *composite structuring element*
- Figure 9.12
  - Three disjoint shapes denoted C, D, and E
    * $A = C \cup D \cup E$
  - Objective: To find the location of one of the shapes, say D
  - Origin/location of each shape given by its center of gravity
  - Let D be enclosed by a small window W
  - *Local background* of D defined by the set difference $(W - D)$
    * Note that D and $W - D$ provide us with the two disjoint ses

$$D \cap (W - D) = \emptyset$$

  - Compute A^c
  - Compute $A \, g \, D$
  - Compute $A^c \, g \, (W - D)$
  - Set of locations where D exactly fits inside A is $(A \, g \, D) \cap (A^c \, g \, (W - D))$
    * The exact location of D
  - If B is the set composed of D and its background, the match of B in A is given by

$$A \sim B = (A \, g \, D) \cap [A^c \, g \, (W - D)]$$

- The above can be generalized to the composite se being defined by $B = (B_1, B_2)$ leading to

$$A \sim B = (A \, g \, B_1) \cap (A^c \, g \, B_2)$$

- B_1 is the set formed from elements of B associated with the object; $B_1 = D$
- $B_2 = (W - D)$

- A point z in universe A belongs to the output if $(B_1)_z$ fits in A (hit) and $(B_2)_z$ misses A

Some basic morphological algorithms

- Useful in extracting image components for representation and description of shape
- Boundary extraction
  - Boundary of a set A
    - ∗ Denoted by $\beta(A)$
    - ∗ Extracted by eroding A by a suitable se B and computing set difference between A and its erosion

$$\beta(A) = A - (A \text{ g } B)$$

  - Figure 9.13
    - ∗ Using a larger se will yield a thicker boundary
  - Figure 9.14
- Hole filling
  - Hole
    - ∗ Background region surrounded by a connected border of foreground pixels
  - Algorithm based on set dilation, complementation, and intersection
  - Let A be a set whose elements are 8-connected boundaries, each boundary enclosing a background (hole)
  - Given a point in each hole, we want to fill all holes
  - Start by forming an array X_0 of 0s of the same size as A
    - ∗ The locations in X_0 corresponding to the given point in each hole are set to 1
  - Let B be a symmetric se with 4-connected neighbors to the origin

0	1	0
1	1	1
0	1	0

  - Compute $X_k = (X_{k-1} \oplus B) \cap A^c \quad k = 1, 2, 3, \ldots$
  - Algorithm terminates at iteration step k if $X_k = X_{k-1}$
  - X_k contains all the filled holes
  - $X_k \cup A$ contains all the filled holes and their boundaries
  - The intersection with A^c at each step limits the result to inside the roi
    - ∗ Also called *conditioned dilation*
  - Figure 9.15
  - Example: Figure 9.16
    - ∗ Thresholded image of polished spheres (ball bearings)
    - ∗ Eliminate reflection by hole filling
    - ∗ Points inside the background selected manually
- Extraction of connected components
  - Let A be a set containing one or more connected components

- Form an array X_0 of the same size as A
  * All elements of X_0 are 0 except for one point in each connected component set to 1
- Select a suitable se B, possibly an 8-connected neighborhood as

1	1	1
1	1	1
1	1	1

- Start with X_0 and find all connected components using the iterative procedure

$$X_k = (X_{k-1} \oplus B) \cap A \quad k = 1, 2, 3, \ldots$$

- Procedure terminates when $X_k = X_{k-1}$; X_k contains all the connected components in the input image
- The only difference from the hole-filling algorithm is the intersection with A instead of A^c
  * This is because here, we are searching for foreground points while in hole filling, we looked for background points (holes)
- Figure 9.17
- Example: Figure 9.18
  * X-ray image of chicken breast with bone fragments
  * Objects of "significant size" can be selected by applying erosion to the thresholded image
  * We may apply labels to the extracted components (region labeling)

- Convex hull
  - Convex set A
    * Straight line segment joining any two points in A lies entirely within A
  - Convex hull H of an arbitrary set of points S is the smallest convex set containing S
  - Set difference $H - S$ is called the *convex deficiency* of S
  - Convex hull and convex deficiency are useful to describe objects
  - Algorithm to compute convex hull $C(A)$ of a set A
    * Figure 9.19
    * Let $B^i, i = 1, 2, 3, 4$ represent the four structuring elements in the figure
      · B^i is a clockwise rotation of B^{i-1} by 90°
    * Implement the equation

$$X_k^i = (X_{k-1} \sim B^i) \cup A \quad i = 1, 2, 3, 4 \text{ and } k = 1, 2, 3, \ldots$$

with $X_0^i = A$

    * Apply hit-or-miss with B^1 till $X_k == X_{k-1}$, then, with B^2 over original A, B^3, and B_4
    * Procedure converges when $X_k^i = X_{k-1}^i$ and we let $D^i = X_k^i$
    * Convex hull of A is given by

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

- Shortcoming of the above procedure
  - Convex hull can grow beyond the minimum dimensions required to guarantee convexity
  - May be fixed by limiting growth to not extend past the bounding box for the original set of points
  - Figure 9.20

- Thinning
  - Transformation of a digital image into a simple topologically equivalent image
    - Remove selected foreground pixels from binary images
    - Used to tidy up the output of edge detectors by reducing all lines to single pixel thickness
  - Thinning of a set A by se B is denoted by $A \otimes B$
  - Defined in terms of hit-or-miss transform as

$$A \otimes B = A - (A \sim B) \\ = A \cap (A \sim B)^c$$

  - Only need to do pattern matching with se; no background operation required in hit-or-miss transform
  - A more useful expression for thinning A symmetrically based on a sequence of ses

$$\{B\} = \{B^1, B^2, \ldots, B^n\}$$

    where B^i is a rotated version of B^{i-1}
  - Define thinning by a sequence of ses as

$$A \otimes \{B\} = ((\ldots((A \otimes B^1) \otimes B^2)\ldots) \otimes B^n)$$

  - Figure 9.21
    - Iterate over the procedure till convergence

- Thickening
  - Morphological dual of thinning defined by

$$A \circledS B = A \cup (A \sim B)$$

  - ses complements of those used for thinning
  - Thickening can also be defined as a sequential operation

$$A \circledS \{B\} = ((\ldots((A \circledS B^1) \circledS B^2)\ldots) \circledS B^n)$$

  - Figure 9.22
  - Usual practice to thin the background and take the complement
    - May result in disconnected points
    - Post-process to remove the disconnected points
    - A g kB indicates k successive erosions of A

$$(A \ominus kB) = ((\ldots((A \ominus B) \ominus B) \ominus \ldots) \ominus B)$$

* K is the last iterative step before A erodes to an empty set

$$K = \max\{k \mid (A \ominus kB) \neq \emptyset\}$$

* $S(A)$ can be obtained as the union of skeleton subsets $S_k(A)$
* A can be reconstructed from the subsets using the equation

UNIT-V
IMAGE COMPRESSION

5.1 INTRODUCTION:

The term data compression refers to the process of reducing the amount of data required to represent a given quantity of information. A clear distinction must be made between data and information. They are not synonymous. In fact, data are the means by which information is conveyed. Various amounts of data may be used to represent the same amount of information. Such might be the case, for example, if a long-winded individual and someone who is short and to the point were to relate the same story. Here, the information of interest is the story; words are the data used to relate the information. If the two individuals use a different number of words to tell the same basic story, two different versions of the story are created, and at least one includes nonessential data. That is, it contains data (or words) that either provide no relevant information or simply restate that which is already known. It is thus said to contain data redundancy.

Data redundancy is a central issue in digital image compression. It is not an abstract concept but a mathematically quantifiable entity. If n_1 and n_2 denote the number of information-carrying units in two data sets that represent the same information, the relative data redundancy R_D of the first data set (the one characterized by n_1) can be defined as

$$R_D = 1 - \frac{1}{C_R}$$

where C_R , commonly called the compression ratio, is

$$C_R = \frac{n_1}{n_2}.$$

For the case $n_2 = n_1$, $C_R = 1$ and $R_D = 0$, indicating that (relative to the second data set) the first representation of the information contains no redundant data. When $n_2 << n_1$, $C_R \rightarrow \infty$ and $R_D \rightarrow 1$, implying significant compression and highly redundant data. Finally, when $n_2 >> n_1$, $C_R \rightarrow 0$ and $R_D \rightarrow \infty$, indicating that the second data set contains much more data than the original representation. This, of course, is the normally undesirable case of data expansion. In general, C_R and R_D lie in the open intervals $(0,\infty)$ and $(-\infty, 1)$, respectively. A practical compression ratio, such as 10 (or 10:1), means that the first data set has 10 information carrying units (say, bits) for every 1 unit in the second or compressed data set. The corresponding redundancy of 0.9 implies that 90% of the data in the first data set is redundant.

In digital image compression, three basic data redundancies can be identified and exploited: **coding redundancy**, **interpixel redundancy**, and **psychovisual redundancy**. Data compression is achieved when one or more of these redundancies are reduced or eliminated.

Coding Redundancy:

In this, we utilize formulation to show how the gray-level histogram of an image also can provide a great deal of insight into the construction of codes to reduce the amount of data used to represent it.

Let us assume, once again, that a discrete random variable r_k in the interval [0, 1] represents the gray levels of an image and that each r_k occurs with probability $p_r(r_k)$.

$$p_r(r_k) = \frac{n_k}{n} \qquad k = 0, 1, 2, \dots, L-1$$

where L is the number of gray levels, n_k is the number of times that the kth gray level appears in the image, and n is the total number of pixels in the image. If the number of bits used to represent each value of r_k is $l(r_k)$, then the average number of bits required to represent each pixel is

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

That is, the average length of the code words assigned to the various gray-level values is found by summing the product of the number of bits used to represent each gray level and the probability that the gray level occurs. Thus the total number of bits required to code an M X N image is MNL_{avg}.

Spatial and Temporal Redundancy:

Consider the images shown in Figs. 5.1(a) and (b). As Figs. 5.1(c) and (d) show, these images have virtually identical histograms. Note also that both histograms are trimodal, indicating the presence of three dominant ranges of gray-level values. Because the gray levels in these images are not equally probable, variable-length coding can be used to reduce the coding redundancy that would result from a straight or natural binary encoding of their pixels. The coding process, however, would not alter the level of correlation between the pixels within the images. In other words, the codes used to represent the gray levels of each image have nothing to do with the correlation between pixels. These correlations result from the structural or geometric relationships between the objects in the image.

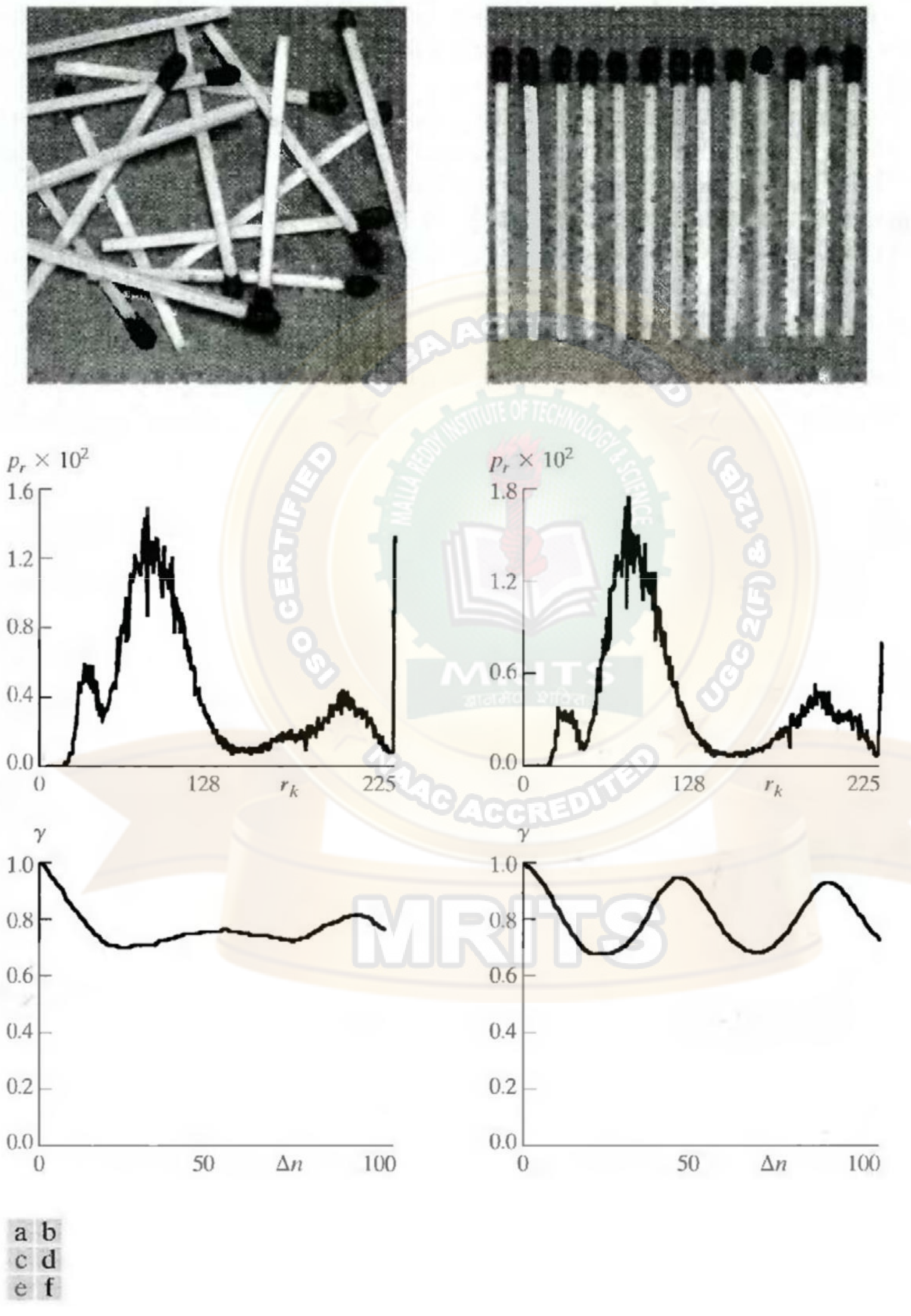


Fig.5.1 Two images and their gray-level histograms and normalized autocorrelation coefficients along one line.

Figures 5.1(e) and (f) show the respective autocorrelation coefficients computed along one line of each image.

$$\gamma(\Delta n) = \frac{A(\Delta n)}{A(0)}$$

where

$$A(\Delta n) = \frac{1}{N - \Delta n} \sum_{y=0}^{N-1-\Delta n} f(x, y) f(x, y + \Delta n).$$

The scaling factor in Eq. above accounts for the varying number of sum terms that arise for each integer value of n. Of course, n must be strictly less than N, the number of pixels on a line. The variable x is the coordinate of the line used in the computation. Note the dramatic difference between the shape of the functions shown in Figs. 5.1(e) and (f). Their shapes can be qualitatively related to the structure in the images in Figs. 5.1(a) and (b).This relationship is particularly noticeable in Fig. 5.1 (f), where the high correlation between pixels separated by 45 and 90 samples can be directly related to the spacing between the vertically oriented matches of Fig. 5.1(b). In addition, the adjacent pixels of both images are highly correlated. When n is 1, γ is 0.9922 and 0.9928 for the images of Figs. 5.1 (a) and (b), respectively. These values are typical of most properly sampled television images.

These illustrations reflect another important form of data redundancy—one directly related to the interpixel correlations within an image. Because the value of any given pixel can be reasonably predicted from the value of its neighbors, the information carried by individual pixels is relatively small. Much of the visual contribution of a single pixel to an image is redundant; it could have been guessed on the basis of the values of its neighbors. A variety of names, including spatial redundancy, geometric redundancy, and interframe redundancy, have been coined to refer to these interpixel dependencies. We use the term interpixel redundancy to encompass them all.

In order to reduce the interpixel redundancies in an image, the 2-D pixel array normally used for human viewing and interpretation must be transformed into a more efficient (but usually "nonvisual") format. For example, the differences between adjacent pixels can be used to represent an image. Transformations of this type (that is, those that remove interpixel redundancy) are referred to as mappings. They are called reversible mappings if the original image elements can be reconstructed from the transformed data set.

Irrelevant Information:

The brightness of a region, as perceived by the eye, depends on factors other than simply the light reflected by the region. For example, intensity variations (Mach bands) can be perceived in an area of constant intensity. Such phenomena result from the fact that the eye does not respond with equal sensitivity to all visual information. Certain information simply has less relative importance than other information in normal visual processing. This information is said to be psychovisually redundant. It can be eliminated without significantly impairing the quality of image perception.

That psychovisual redundancies exist should not come as a surprise, because human perception of the information in an image normally does not involve quantitative analysis of every pixel value in the image. In general, an observer searches for distinguishing features such as edges or textural regions and mentally combines them into recognizable groupings. The brain then correlates these groupings with prior knowledge in order to complete the image interpretation process. Psychovisual redundancy is fundamentally different from the redundancies discussed earlier. Unlike coding and interpixel redundancy, psychovisual redundancy is associated with real or quantifiable visual information. Its elimination is possible only because the information itself is not essential for normal visual processing. Since the elimination of psychovisually redundant data results in a loss of quantitative information, it is commonly referred to as quantization.

This terminology is consistent with normal usage of the word, which generally means the mapping of a broad range of input values to a limited number of output values. As it is an irreversible operation (visual information is lost), quantization results in lossy data compression.

5.2 FIDELITY CRITERION:

The removal of psychovisually redundant data results in a loss of real or quantitative visual information. Because information of interest may be lost, a repeatable or reproducible means of quantifying the nature and extent of information loss is highly desirable. Two general classes of criteria are used as the basis for such an assessment:

A) Objective fidelity criteria and

B) Subjective fidelity criteria.

When the level of information loss can be expressed as a function of the original or input image and the compressed and subsequently decompressed output image, it is said to be based on an objective fidelity criterion. A good example is the root-mean-square (rms) error between an input and output image. Let f(x, y) represent an input image and let f(x, y) denote an estimate or approximation of f(x, y) that results from compressing and subsequently decompressing the input. For any value of x and y, the error e(x, y) between f (x, y) and f^ (x, y) can be defined as

$$e(x, y) = \hat{f}(x, y) - f(x, y)$$

so that the total error between the two images is

$$\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}\left[\hat{f}(x, y) - f(x, y)\right]$$

where the images are of size M X N. The root-mean-square error, e_{rms}, between f(x, y) and f^(x, y) then is the square root of the squared error averaged over the M X N array, or

$$e_{\text{rms}} = \left[\frac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}\left[\hat{f}(x, y) - f(x, y)\right]^2\right]^{1/2}$$

A closely related objective fidelity criterion is the mean-square signal-to-noise ratio of the compressed-decompressed image. If f^ (x, y) is considered to be the sum of the original image f(x, y) and a noise signal e(x, y), the mean-square signal-to-noise ratio of the output image, denoted SNR_{rms}, is

$$SNR_{\text{ms}} = \frac{\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}\hat{f}(x, y)^2}{\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}\left[\hat{f}(x, y) - f(x, y)\right]^2}.$$

The rms value of the signal-to-noise ratio, denoted SNR_{rms}, is obtained by taking the square root of Eq. above.

Although objective fidelity criteria offer a simple and convenient mechanism for evaluating information loss, most decompressed images ultimately are viewed by humans. Consequently, measuring image quality by the subjective evaluations of a human observer often is more appropriate. This can be accomplished by showing a "typical" decompressed image to an appropriate cross section of viewers and averaging their evaluations. The evaluations may be made using an absolute rating scale or by means of side-by-side comparisons of f(x, y) and f^(x, y).

5.3 IMAGE COMPRESSION MODEL:

Fig. 5.2 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.

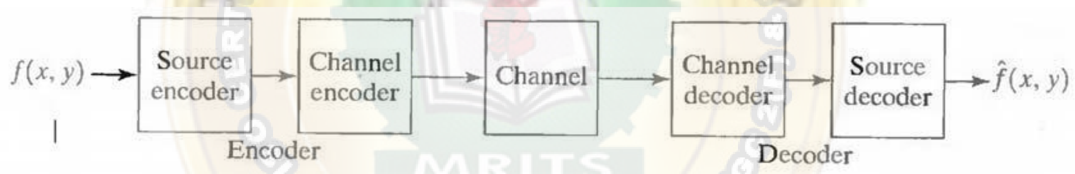


Fig.3.2 A general compression system model

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. 5.3 (a) shows, each operation is designed to reduce one of the three redundancies. Figure 5.3 (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.

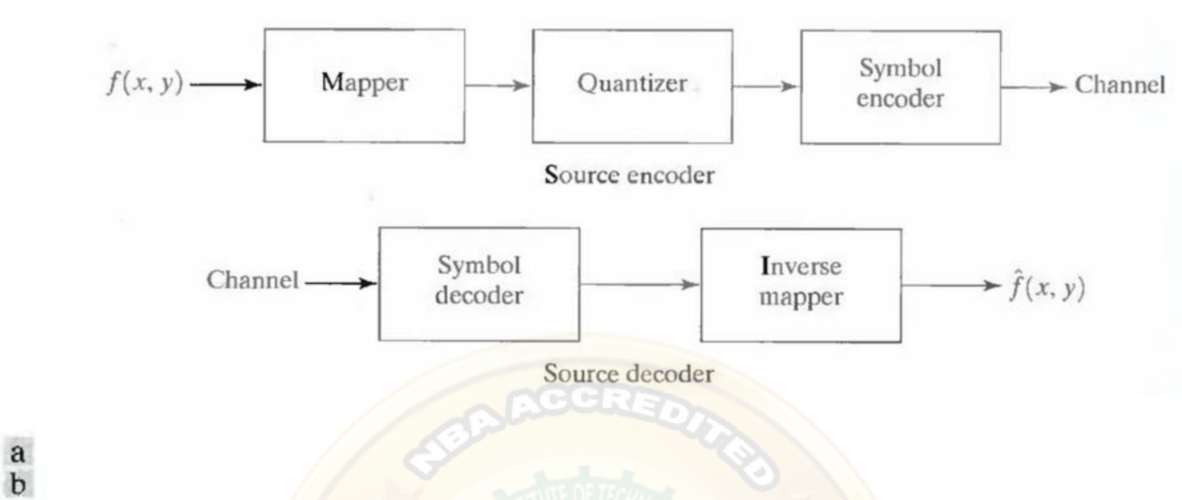


Fig.5.3 (a) Source encoder and (b) source decoder model

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. 5.3 (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 5.3(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. 5.3(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. 5.3(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. 5.3(b).

The Channel Encoder and Decoder:

The channel encoder and decoder play an important role in the overall encoding-decoding process when the channel of Fig. 3.1 is noisy or prone to error. They are designed to reduce the impact of channel noise by inserting a controlled form of redundancy into the source encoded data. As the output of the source encoder contains little redundancy, it would be highly sensitive to transmission noise without the addition of this "controlled redundancy." One of the most useful channel encoding techniques was devised by R. W. Hamming (Hamming [1950]). It is based on appending enough bits to the data being encoded to ensure that some minimum number of bits must change between valid code words. Hamming showed, for example, that if 3 bits of redundancy are added to a 4-bit word, so that the distance between any two valid code words is 3, all single-bit errors can be detected and corrected. (By appending additional bits of redundancy, multiple-bit errors can be detected and corrected.) The 7-bit Hamming (7, 4) code word $h_1, h_2, h_3 \ldots, h_6, h_7$ associated with a 4-bit binary number $b_3b_2b_1b_0$ is

$$\begin{aligned} h_1 &= b_3 \oplus b_2 \oplus b_0 & h_3 &= b_3 \\ h_2 &= b_3 \oplus b_1 \oplus b_0 & h_5 &= b_2 \\ h_4 &= b_2 \oplus b_1 \oplus b_0 & h_6 &= b_1 \\ & & h_7 &= b_0 \end{aligned}$$

where denotes the exclusive OR operation. Note that bits h_1, h_2, and h_4 are even- parity bits for the bit fields b_3 b_2 b_0, $b_3b_1b_0$, and $b_2b_1b_0$, respectively. (Recall that a string of binary bits has even parity if the number of bits with a value of 1 is even.) To decode a Hamming encoded result, the channel decoder must check the encoded value for odd parity over the bit fields in which even parity was previously established. A single-bit error is indicated by a nonzero parity word $c_4c_2c_1$, where

$$c_1 = h_1 \oplus h_3 \oplus h_5 \oplus h_7$$
$$c_2 = h_2 \oplus h_3 \oplus h_6 \oplus h_7$$
$$c_4 = h_4 \oplus h_5 \oplus h_6 \oplus h_7.$$

If a nonzero value is found, the decoder simply complements the code word bit position indicated by the parity word. The decoded binary value is then extracted from the corrected code word as $h_3h_5h_6h_7$.

5.4 HUFFMAN CODING:

The most popular technique for removing coding redundancy is due to Huffman (Huffman [1952]). When coding the symbols of an information source individually, Huffman coding yields the smallest possible number of code symbols per source symbol. In terms of the noiseless coding theorem, the resulting code is optimal for a fixed value of n, subject to the constraint that the source symbols be coded one at a time.

The first step in Huffman's approach is to create a series of source reductions by ordering the probabilities of the symbols under consideration and combining the lowest probability symbols into a single symbol that replaces them in the next source reduction. Figure 5.4 illustrates this process for binary coding (K-ary Huffman codes can also be constructed). At the far left, a hypothetical set of source symbols and their probabilities are ordered from top to bottom in terms of decreasing probability values. To form the first source reduction, the bottom two probabilities, 0.06 and 0.04, are combined to form a "compound symbol" with probability 0.1. This compound symbol and its associated probability are placed in the first source reduction column so that the probabilities of the reduced source are also ordered from the most to the least probable. This process is then repeated until a reduced source with two symbols (at the far right) is reached.

The second step in Huffman's procedure is to code each reduced source, starting with the smallest source and working back to the original source. The minimal length binary code for a two-symbol source, of course, is the symbols 0 and 1. As Fig. 5.5 shows, these symbols are assigned to the two symbols on the right (the assignment is arbitrary; reversing the order of the 0 and 1 would work just as well). As the reduced source symbol with probability 0.6 was generated by combining two symbols in the reduced source to its left, the 0 used to code it is now assigned to both of these symbols, and a 0 and 1 are arbitrarily

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

Fig.5.4 Huffman source reductions.

Original source			Source reduction							
Sym.	Prob.	Code	1		2		3		4	
a_2	0.4	1	0.4	1	0.4	1	0.4	1	0.6	0
a_6	0.3	00	0.3	00	0.3	00	0.3	00	0.4	1
a_1	0.1	011	0.1	011	0.2	010	0.3	01		
a_4	0.1	0100	0.1	0100	0.1	011				
a_3	0.06	01010	0.1	0101						
a_5	0.04	01011								

Fig.5.5 Huffman code assignment procedure.

appended to each to distinguish them from each other. This operation is then repeated for each reduced source until the original source is reached. The final code appears at the far left in Fig. 5.5. The average length of this code is

$$\begin{aligned} L_{avg} &= (0.4)(1) + (0.3)(2) + (0.1)(3) + (0.1)(4) + (0.06)(5) + (0.04)(5) \\ &= 2.2 \text{ bits/symbol} \end{aligned}$$

and the entropy of the source is 2.14 bits/symbol. The resulting Huffman code efficiency is 0.973.

Huffman's procedure creates the optimal code for a set of symbols and probabilities subject to the constraint that the symbols be coded one at a time. After the code has been created, coding and/or decoding is accomplished in a simple lookup table manner. The code itself is an instantaneous uniquely decodable block code. It is called a block code because each source symbol is mapped into a fixed sequence of code symbols. It is instantaneous, because each code word in a string of code symbols can be decoded without referencing succeeding symbols. It is uniquely decodable, because any string of code symbols can be decoded in only one way. Thus, any string of Huffman encoded symbols can be decoded by examining the individual symbols of the string in a left to right manner. For the binary code of Fig. 5.5, a left-to-right scan of the encoded string 010100111100 reveals that the first valid code word is 01010, which is the code for symbol a_3. The next valid code is 011, which corresponds to symbol a_1. Continuing in this manner reveals the completely decoded message to be $a_3a_1a_2a_2a_6$.

5.6 ARITHMETIC CODING:

Unlike the variable-length codes described previously, arithmetic coding generates nonblock codes. In arithmetic coding, which can be traced to the work of Elias, a one-to-one correspondence between source symbols and code words does not exist. Instead, an entire sequence of source symbols (or message) is assigned a single arithmetic code word. The code word itself defines an interval of real numbers between 0 and 1. As the number of symbols in the message increases, the interval used to represent it becomes smaller and the number of information units (say, bits) required to represent the interval becomes larger. Each symbol of the message reduces the size of the interval in accordance with its probability of occurrence. Because the technique does not require, as does Huffman's approach, that each source symbol translate into an integral number of code symbols (that is, that the symbols be coded one at a time), it achieves (but only in theory) the bound established by the noiseless coding theorem.

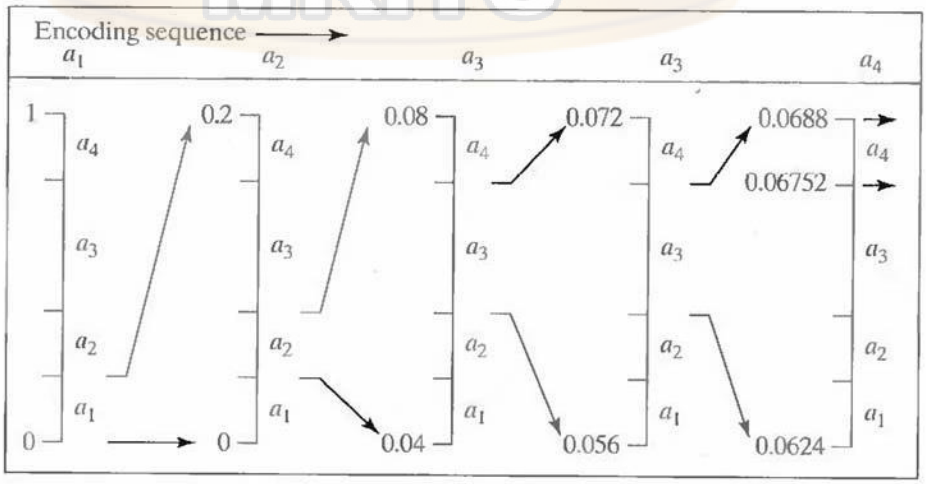


Fig.5.6 Arithmetic coding procedure

Figure 5.6 illustrates the basic arithmetic coding process. Here, a five-symbol sequence or message, $a_1a_2a_3a_3a_4$, from a four-symbol source is coded. At the start of the coding process, the message is assumed to occupy the entire half-open interval [0, 1). As Table 5.1 shows, this interval is initially subdivided into four regions based on the probabilities of each source symbol. Symbol ax, for example, is associated with subinterval [0, 0.2). Because it is the first symbol of the message being coded, the message interval is initially narrowed to [0, 0.2). Thus in Fig. 5.1 [0, 0.2) is expanded to the full height of the figure and its end points labeled by the values of the narrowed range. The narrowed range is then subdivided in accordance with the original source symbol probabilities and the process continues with the next message symbol.

Source Symbol	Probability	Initial Subinterval
a_1	0.2	[0.0, 0.2)
a_2	0.2	[0.2, 0.4)
a_3	0.4	[0.4, 0.8)
a_4	0.2	[0.8, 1.0)

Table 5.1 Arithmetic coding example

In this manner, symbol a_2 narrows the subinterval to [0.04, 0.08), a_3 further narrows it to [0.056, 0.072), and so on. The final message symbol, which must be reserved as a special end-of- message indicator, narrows the range to [0.06752, 0.0688). Of course, any number within this subinterval—for example, 0.068—can be used to represent the message.

In the arithmetically coded message of Fig. 5.1, three decimal digits are used to represent the five-symbol message. This translates into 3/5 or 0.6 decimal digits per source symbol and compares favorably with the entropy of the source, which is 0.58 decimal digits or 10-ary units/symbol. As the length of the sequence being coded increases, the resulting arithmetic code approaches the bound established by the noiseless coding theorem.

In practice, two factors cause coding performance to fall short of the bound: (1) the addition of the end-of-message indicator that is needed to separate one message from an-other; and

(2) the use of finite precision arithmetic. Practical implementations of arithmetic coding address the latter problem by introducing a scaling strategy and a rounding strategy (Langdon and Rissanen [1981]). The scaling strategy renormalizes each subinterval to the [0, 1) range before subdividing it in accordance with the symbol probabilities. The rounding strategy guarantees that the truncations associated with finite precision arithmetic do not prevent the coding subintervals from being represented accurately.

5.7 LOSSLESS PREDICTIVE CODING:

The error-free compression approach 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 5.7 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_n, 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 $f^\wedge{}_n$ and used to form the difference or prediction error which is coded using a variable-length code (by the symbol encoder) to generate the next element of the compressed data stream.

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

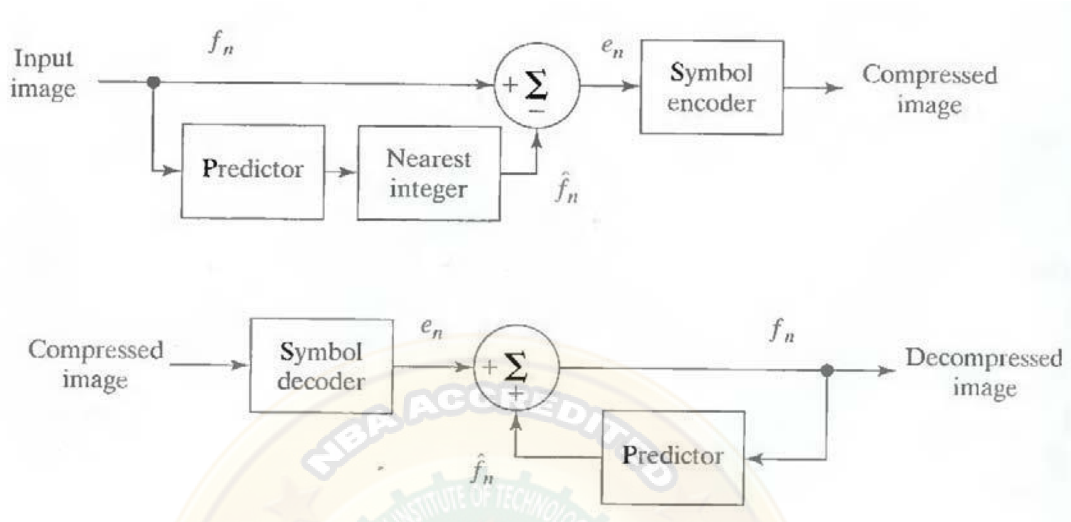


Fig.5.7 A lossless predictive coding model: (a) encoder; (b) decoder

The decoder of Fig. 5.7 (b) reconstructs en from the received variable-length code words and performs the inverse operation

$$f_n = e_n + \hat{f}_n.$$

Various local, global, and adaptive methods can be used to generate f^n. In most cases, however, the prediction is formed by a linear combination of m previous pixels. That is,

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

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,..., m are prediction coefficients. In raster scan applications, the subscript n indexes the predictor outputs in accordance with their time of occurrence. That is, f_n, $\hat{f}_n$ and e_n in Eqns. above could be replaced with the more explicit notation f (t), f^(t), and e (t), where t represents time. In other cases, n is used as an index on the spatial coordinates and/or frame number (in a time sequence of images) of an image. In 1-D linear predictive coding, for example, Eq. above can be written as

$$\hat{f}_n(x, y) = \text{round}\left[\sum_{i=1}^{m} \alpha_i f(x, y - i)\right]$$

where each subscripted variable is now expressed explicitly as a function of spatial coordinates x and y. The Eq. indicates that the 1-D linear prediction f(x, y) is a function of the previous pixels on the current line alone. In 2-D predictive coding, the prediction is a function of the previous pixels in a left-to-right, top-to-bottom scan of an image. In the 3-D case, it is based on these pixels and the previous pixels of preceding frames. Equation above cannot be evaluated for the first m pixels of each line, so these pixels must be coded by using other means (such as a Huffman code) and considered as an overhead of the predictive coding process. A similar comment applies to the higher-dimensional cases.

5.8 LOSSY PREDICTIVE CODING:

In this type of coding, we add a quantizer to the lossless predictive model and examine the resulting trade-off between reconstruction accuracy and compression performance. As Fig.5.8 shows, the quantizer, which 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. It maps the prediction error into a limited range of outputs, denoted e^n which establish the amount of compression and distortion associated with lossy predictive coding.

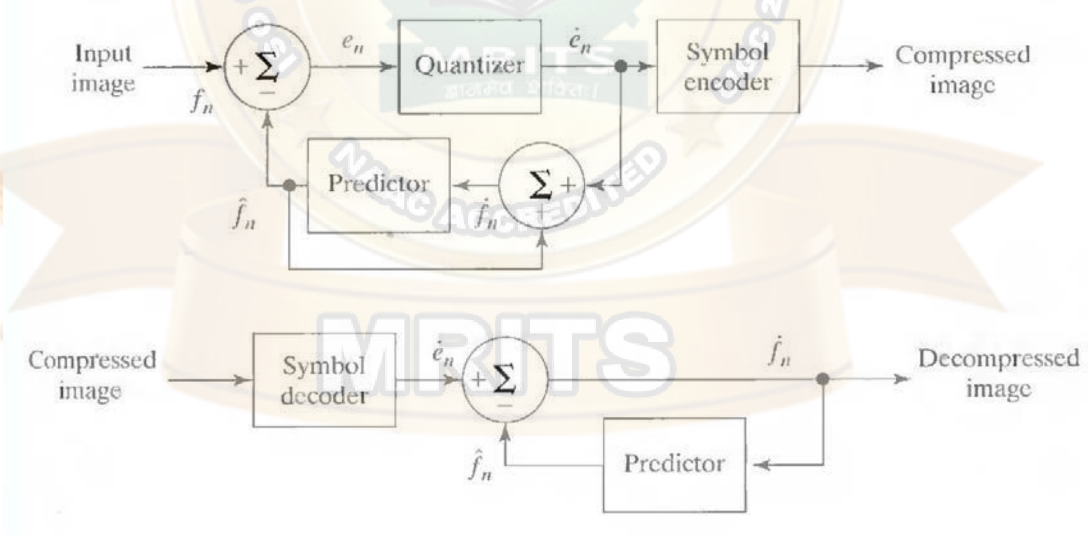


Fig.5.8 A lossy predictive coding model: (a) encoder and (b) decoder.

In order to accommodate the insertion of the quantization step, the error-free encoder of figure must be altered so that the predictions generated by the encoder and decoder are equivalent. As Fig.5.8 (a) shows, this is accomplished by placing the lossy encoder's predictor within a feedback loop, where its input, denoted $\dot{f}_n$, is generated as a function of past predictions and the corresponding quantized errors. That is,

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

This closed loop configuration prevents error buildup at the decoder's output. Note from Fig.5.8 (b) that the output of the decoder also is given by the above Eqn.

5.9 BLOCK TRANSFORM CODING:

All the predictive coding techniques operate directly on the pixels of an image and thus are spatial domain methods. In this coding, we consider compression techniques that are based on modifying the transform of an image. In transform coding, 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. For most natural images, a significant number of the coefficients have small magnitudes and can be coarsely quantized (or discarded entirely) with little image distortion. A variety of transformations, including the discrete Fourier transform (DFT), can be used to transform the image data.

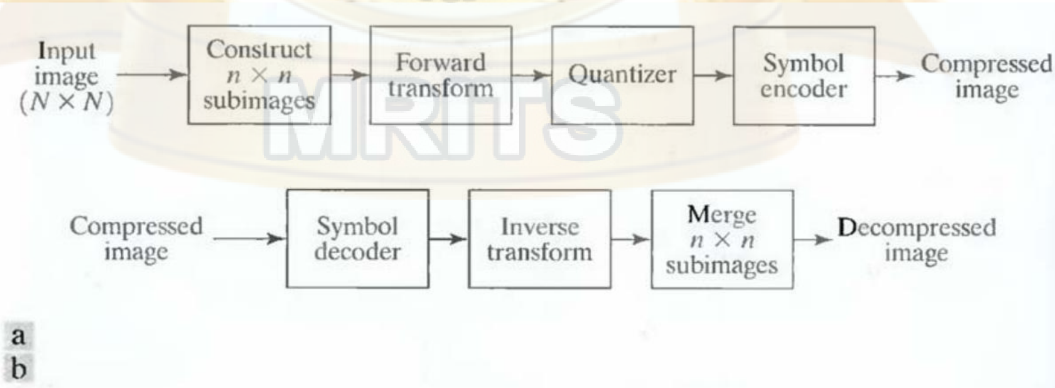


Fig. 5.9 A transform coding system: (a) encoder; (b) decoder.

Figure 5.9 shows a typical transform coding system. The decoder implements the inverse sequence of steps (with the exception of the quantization function) of the encoder, which performs four relatively straightforward operations: subimage decomposition, transformation, quantization, and coding. An N X N input image first is subdivided into subimages of size n X n, which are then

transformed to generate $(N/n)^2$ subimage transform arrays, each of size n X n. The goal of the transformation process is to decorrelate the pixels of each subimage, or to pack as much information as possible into the smallest number of transform coefficients. The quantization stage then selectively eliminates or more coarsely quantizes the coefficients that carry the least information. These coefficients have the smallest impact on reconstructed subimage quality. The encoding process terminates by coding (normally using a variable-length code) the quantized coefficients. Any or all of the transform encoding steps can be adapted to local image content, called adaptive transform coding, or fixed for all subimages, called nonadaptive transform coding.

5.10 JPEG 2000 STANDARD:

JPEG-2000 extends the popular 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 just described. Coefficient quantization is adapted to individual scales and subbands and the quantized coefficients are arithmetically coded on a bit-plane basis Using the notation of the standard, an image is encoded as follows (ISO/IEC [2000]).

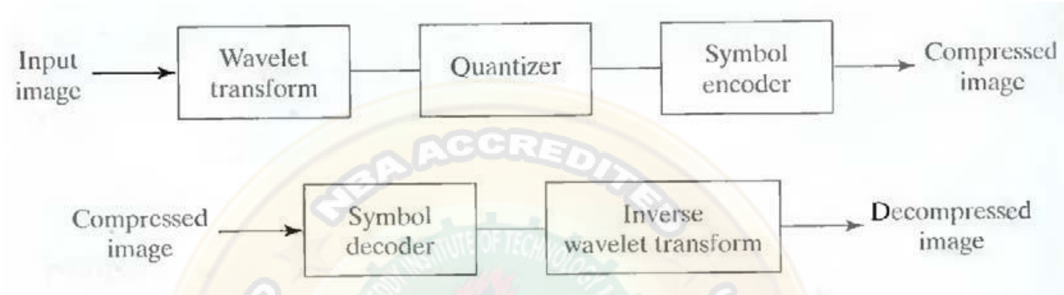
The first step of the encoding process is to DC level shift the samples of the *Ssiz*-bit unsigned image to be coded by subtracting If the image has more than one *component*—like the red, green, and blue planes of a color image—each component is shifted individually. If there are exactly three components,

they may be optionally decorrelated using a reversible or nonreversible linear combination of the components.

Wavelet Coding:

The wavelet coding is based on the idea that the coefficients of a transform that decorrelates the pixels of an image can be coded more efficiently than the original pixels themselves. If the transform's basis functions—in this case wavelets—pack most of the important visual information into a small number of coefficients, the remaining coefficients can be quantized coarsely or truncated to zero with little image distortion.

Figure 5.10 shows a typical wavelet coding system. To encode a 2^J X 2^J image, an analyzing wavelet, Ψ, and minimum decomposition level, J - P, are selected and used to compute the image's discrete wavelet transform. If the wavelet has a complimentary scaling function φ, the fast wavelet transform can be used. In either case, the computed transform converts a large portion of the original image to horizontal, vertical, and diagonal decomposition coefficients with zero mean and Laplacian-like distributions.



ig.5.10 A wavelet coding system: (a) encoder; (b) decoder.

Since many of the computed coefficients carry little visual information, they can be quantized and coded to minimize intercoefficient and coding redundancy. Moreover, the quantization can be adapted to exploit any positional correlation across the P decomposition levels. One or more of the lossless coding methods, including run-length, Huffman, arithmetic, and bit-plane coding, can be incorporated into the final symbol coding step. Decoding is accomplished by inverting the encoding operations—with the exception of quantization, which cannot be reversed exactly.

The principal difference between the wavelet-based system and the transform coding system is the omission of the transform coder's subimage processing stages.

Because wavelet transforms are both computationally efficient and inherently local (i.e., their basis functions are limited in duration), subdivision of the original image is unnecessary.

BEYOND SYLLABUS

IMAGE PROCESSING USING OPENCV

1. ***Open Cv Tutorial-I***

```
# USAGE

# python opencv_tutorial_01.py

# import the necessary packages

import imutils

import cv2


# load the input image and show its dimensions, keeping in mind that

# images are represented as a multi-dimensional NumPy array with

# shape no. rows (height) x no. columns (width) x no. channels (depth)

image = cv2.imread("jp.png")

(h, w, d) = image.shape

print("width={}, height={}, depth={}".format(w, h, d))


# display the image to our screen -- we will need to click the window

# open by OpenCV and press a key on our keyboard to continue execution

cv2.imshow("Image", image)

cv2.waitKey(0)
```

```
# access the RGB pixel located at x=50, y=100, keepind in mind that

# OpenCV stores images in BGR order rather than RGB

(B, G, R) = image[100, 50]

print("R={}, G={}, B={}".format(R, G, B))


# extract a 100x100 pixel square ROI (Region of Interest) from the

# input image starting at x=320,y=60 at ending at x=420,y=160

roi = image[60:160, 320:420]

cv2.imshow("ROI", roi)

cv2.waitKey(0)


# resize the image to 200x200px, ignoring aspect ratio

resized = cv2.resize(image, (200, 200))

cv2.imshow("Fixed Resizing", resized)

cv2.waitKey(0)


# fixed resizing and distort aspect ratio so let's resize the width

# to be 300px but compute the new height based on the aspect ratio

r = 300.0 / w

dim = (300, int(h * r))

resized = cv2.resize(image, dim)
```

```
cv2.imshow("Aspect Ratio Resize", resized)
cv2.waitKey(0)

# manually computing the aspect ratio can be a pain so let's use the
# imutils library instead
resized = imutils.resize(image, width=300)
cv2.imshow("Imutils Resize", resized)
cv2.waitKey(0)

# let's rotate an image 45 degrees clockwise using OpenCV by first
# computing the image center, then constructing the rotation matrix,
# and then finally applying the affine warp
center = (w // 2, h // 2)
M = cv2.getRotationMatrix2D(center, -45, 1.0)
rotated = cv2.warpAffine(image, M, (w, h))
cv2.imshow("OpenCV Rotation", rotated)
cv2.waitKey(0)

# rotation can also be easily accomplished via imutils with less code
rotated = imutils.rotate(image, -45)
cv2.imshow("Imutils Rotation", rotated)
```

```
cv2.waitKey(0)


# OpenCV doesn't "care" if our rotated image is clipped after rotation
# so we can instead use another imutils convenience function to help
# us out
rotated = imutils.rotate_bound(image, 45)
cv2.imshow("Imutils Bound Rotation", rotated)
cv2.waitKey(0)
# apply a Gaussian blur with a 11x11 kernel to the image to smooth it,
# useful when reducing high frequency noise
blurred = cv2.GaussianBlur(image, (11, 11), 0)
cv2.imshow("Blurred", blurred)
cv2.waitKey(0)
# draw a 2px thick red rectangle surrounding the face
output = image.copy()
cv2.rectangle(output, (320, 60), (420, 160), (0, 0, 255), 2)
cv2.imshow("Rectangle", output)
cv2.waitKey(0)


# draw a blue 20px (filled in) circle on the image centered at
# x=300,y=150
```

```
output = image.copy()

cv2.circle(output, (300, 150), 20, (255, 0, 0), -1)

cv2.imshow("Circle", output)

cv2.waitKey(0)


# draw a 5px thick red line from x=60,y=20 to x=400,y=200

output = image.copy()

cv2.line(output, (60, 20), (400, 200), (0, 0, 255), 5)

cv2.imshow("Line", output)

cv2.waitKey(0)


# draw green text on the image

output = image.copy()

cv2.putText(output, "OpenCV + Jurassic Park!!!", (10, 25),

        cv2.FONT_HERSHEY_SIMPLEX, 0.7, (0, 255, 0), 2)

cv2.imshow("Text", output)

cv2.waitKey(0)
```

2. ***Open Cv Tutorial-II***

```
#USAGE

#python opencv_tutorial_02.py --image tetris_blocks.png

# import the necessary packages
```

```
import argparse

import imutils

import cv2


# construct the argument parser and parse the arguments

ap = argparse.ArgumentParser()

ap.add_argument("-i", "--image", required=True,

        help="path to input image")

args = vars(ap.parse_args())


# load the input image (whose path was supplied via command line

# argument) and display the image to our screen

image = cv2.imread(args["image"])

cv2.imshow("Image", image)

cv2.waitKey(0)


# convert the image to grayscale

gray = cv2.cvtColor(image, cv2.COLOR_BGR2GRAY)

cv2.imshow("Gray", gray)

cv2.waitKey(0)
```

```
# applying edge detection we can find the outlines of objects in
# images
edged = cv2.Canny(gray, 30, 150)
cv2.imshow("Edged", edged)
cv2.waitKey(0)

# threshold the image by setting all pixel values less than 225
# to 255 (white; foreground) and all pixel values >= 225 to 255
# (black; background), thereby segmenting the image
thresh = cv2.threshold(gray, 225, 255, cv2.THRESH_BINARY_INV)[1]
cv2.imshow("Thresh", thresh)
cv2.waitKey(0)

# find contours (i.e., outlines) of the foreground objects in the
# thresholded image
cnts = cv2.findContours(thresh.copy(), cv2.RETR_EXTERNAL,
        cv2.CHAIN_APPROX_SIMPLE)
cnts = imutils.grab_contours(cnts)
output = image.copy()

# loop over the contours
```

```
for c in cnts:
	# draw each contour on the output image with a 3px thick purple
	# outline, then display the output contours one at a time
	cv2.drawContours(output, [c], -1, (240, 0, 159), 3)
	cv2.imshow("Contours", output)
	cv2.waitKey(0)

# draw the total number of contours found in purple
text = "I found {} objects!".format(len(cnts))
cv2.putText(output, text, (10, 25),  cv2.FONT_HERSHEY_SIMPLEX, 0.7,
	(240, 0, 159), 2)
cv2.imshow("Contours", output)
cv2.waitKey(0)

# we apply erosions to reduce the size of foreground objects
mask = thresh.copy()
mask = cv2.erode(mask, None, iterations=5)
cv2.imshow("Eroded", mask)
cv2.waitKey(0)

# similarly, dilations can increase the size of the ground objects
```

```
mask = thresh.copy()
mask = cv2.dilate(mask, None, iterations=5)
cv2.imshow("Dilated", mask)
cv2.waitKey(0)

# a typical operation we may want to apply is to take our mask and
# apply a bitwise AND to our input image, keeping only the masked
# regions
mask = thresh.copy()
output = cv2.bitwise_and(image, image, mask=mask)
cv2.imshow("Output", output)
cv2.waitKey(0)
```

3. ***Open Cv Histogram Equalization***

```
# USAGE
# python grayscale_histogram.py --image beach.png
# import the necessary packages
```

```
from matplotlib import pyplot as plt

import argparse

import cv2


# construct the argument parser and parse the arguments

ap = argparse.ArgumentParser()

ap.add_argument("-i", "--image", required=True,

        help="path to the image")

args = vars(ap.parse_args())


# load the input image and convert it to grayscale

image = cv2.imread(args["image"])

image = cv2.cvtColor(image, cv2.COLOR_BGR2GRAY)


# compute a grayscale histogram

hist = cv2.calcHist([image], [0], None, [256], [0, 256])


# matplotlib expects RGB images so convert and then display the image

# with matplotlib

plt.figure()

plt.axis("off")
```

```
plt.imshow(cv2.cvtColor(image, cv2.COLOR_GRAY2RGB))


# plot the histogram
plt.figure()
plt.title("Grayscale Histogram")
plt.xlabel("Bins")
plt.ylabel("# of Pixels")
plt.plot(hist)
plt.xlim([0, 256])


# normalize the histogram
hist /= hist.sum()
# plot the normalized histogram
plt.figure()
plt.title("Grayscale Histogram (Normalized)")
plt.xlabel("Bins")
plt.ylabel("% of Pixels")
plt.plot(hist)
plt.xlim([0, 256])
plt.show()
```

4. ***Open Cv Image Convolution***

```
# USAGE
# python convolutions.py --image 3d_pokemon.png

# import the necessary packages
from skimage.exposure import rescale_intensity
import numpy as np
import argparse
import cv2

def convolve(image, kernel):
	# grab the spatial dimensions of the image, along with
	# the spatial dimensions of the kernel
	(iH, iW) = image.shape[:2]
	(kH, kW) = kernel.shape[:2]

	# allocate memory for the output image, taking care to
	# "pad" the borders of the input image so the spatial
	# size (i.e., width and height) are not reduced
	pad = (kW - 1) // 2
	image = cv2.copyMakeBorder(image, pad, pad, pad, pad,
		cv2.BORDER_REPLICATE)
```

```
output = np.zeros((iH, iW), dtype="float32")


# loop over the input image, "sliding" the kernel across
# each (x, y)-coordinate from left-to-right and top to
# bottom
for y in np.arange(pad, iH + pad):
    for x in np.arange(pad, iW + pad):
        # extract the ROI of the image by extracting the
        # *center* region of the current (x, y)-coordinates
        # dimensions
        roi = image[y - pad:y + pad + 1, x - pad:x + pad + 1]

        # perform the actual convolution by taking the
        # element-wise multiplicate between the ROI and
        # the kernel, then summing the matrix
        k = (roi * kernel).sum()

        # store the convolved value in the output (x,y)-
        # coordinate of the output image
        output[y - pad, x - pad] = k
```

```
	# rescale the output image to be in the range [0, 255]
	output = rescale_intensity(output, in_range=(0, 255))
	output = (output * 255).astype("uint8")

	# return the output image
	return output

# construct the argument parse and parse the arguments
ap = argparse.ArgumentParser()
ap.add_argument("-i", "--image", required=True,
	help="path to the input image")
args = vars(ap.parse_args())

# construct average blurring kernels used to smooth an image
smallBlur = np.ones((7, 7), dtype="float") * (1.0 / (7 * 7))
largeBlur = np.ones((21, 21), dtype="float") * (1.0 / (21 * 21))

# construct a sharpening filter
sharpen = np.array((
	[0, -1, 0],
	[-1, 5, -1],
```

```
        [0, -1, 0]), dtype="int")


# construct the Laplacian kernel used to detect edge-like
# regions of an image
laplacian = np.array((
        [0, 1, 0],
        [1, -4, 1],
        [0, 1, 0]), dtype="int")


# construct the Sobel x-axis kernel
sobelX = np.array((
        [-1, 0, 1],
        [-2, 0, 2],
        [-1, 0, 1]), dtype="int")


# construct the Sobel y-axis kernel
sobelY = np.array((
        [-1, -2, -1],
        [0, 0, 0],
        [1, 2, 1]), dtype="int")
```

```
# construct the kernel bank, a list of kernels we're going
# to apply using both our custom `convole` function and
# OpenCV's `filter2D` function
kernelBank = (
	("small_blur", smallBlur),
	("large_blur", largeBlur),
	("sharpen", sharpen),
	("laplacian", laplacian),
	("sobel_x", sobelX),
	("sobel_y", sobelY)
)

# load the input image and convert it to grayscale
image = cv2.imread(args["image"])
gray = cv2.cvtColor(image, cv2.COLOR_BGR2GRAY)

# loop over the kernels
for (kernelName, kernel) in kernelBank:
	# apply the kernel to the grayscale image using both
	# our custom `convole` function and OpenCV's `filter2D`
	# function
```

```
    print("[INFO] applying {} kernel".format(kernelName))
    convoleOutput = convolve(gray, kernel)
    opencvOutput = cv2.filter2D(gray, -1, kernel)

    # show the output images
    cv2.imshow("original", gray)
    cv2.imshow("{} - convole".format(kernelName), convoleOutput)
    cv2.imshow("{} - opencv".format(kernelName), opencvOutput)
    cv2.waitKey(0)
    cv2.destroyAllWindows()
```

5. ***Open Cv Image Blurring***

```
# USAGE
# python blurring.py

# import the necessary packages
import argparse
import cv2

# construct the argument parser and parse the arguments
ap = argparse.ArgumentParser()
ap.add_argument("-i", "--image", type=str, default="mrits.png",
```

```
        help="path to input image")
args = vars(ap.parse_args())


# load the image, display it to our screen, and initialize a list of
# kernel sizes (so we can evaluate the relationship between kernel
# size and amount of blurring)
image = cv2.imread(args["image"])
cv2.imshow("Original", image)
kernelSizes = [(3, 3), (9, 9), (15, 15)]


# loop over the kernel sizes
for (kX, kY) in kernelSizes:
        # apply an "average" blur to the image using the current kernel
        # size
        blurred = cv2.blur(image, (kX, kY))
        cv2.imshow("Average ({}, {})".format(kX, kY), blurred)
        cv2.waitKey(0)


# close all windows to cleanup the screen
cv2.destroyAllWindows()
cv2.imshow("Original", image)
```

```
# loop over the kernel sizes again
for (kX, kY) in kernelSizes:
        # apply a "Gaussian" blur to the image
        blurred = cv2.GaussianBlur(image, (kX, kY), 0)
        cv2.imshow("Gaussian ({}, {})".format(kX, kY), blurred)
        cv2.waitKey(0)

# close all windows to cleanup the screen
cv2.destroyAllWindows()
cv2.imshow("Original", image)

# loop over the kernel sizes a final time
for k in (3, 9, 15):
        # apply a "median" blur to the image
        blurred = cv2.medianBlur(image, k)
        cv2.imshow("Median {}".format(k), blurred)
        cv2.waitKey(0)
```

6. ***Open Cv Morphological Processing***

```
# USAGE
# python morphological_hats.py --image car.png
```

```
# import the necessary packages

import argparse

import cv2

# construct the argument parser and parse the arguments

ap = argparse.ArgumentParser()

ap.add_argument("-i", "--image", required=True,

	help="path to input image")

args = vars(ap.parse_args())


# load the image and convert it to grayscale

image = cv2.imread(args["image"])

gray = cv2.cvtColor(image, cv2.COLOR_BGR2GRAY)


# construct a rectangular kernel (13x5) and apply a blackhat

# operation which enables us to find dark regions on a light

# background

rectKernel = cv2.getStructuringElement(cv2.MORPH_RECT, (13, 5))

blackhat = cv2.morphologyEx(gray, cv2.MORPH_BLACKHAT, rectKernel)


# similarly, a tophat (also called a "whitehat") operation will

# enable us to find light regions on a dark background
```

```
tophat = cv2.morphologyEx(gray, cv2.MORPH_TOPHAT, rectKernel)


# show the output images

cv2.imshow("Original", image)

cv2.imshow("Blackhat", blackhat)

cv2.imshow("Tophat", tophat)

cv2.waitKey(0)
```

ASSIGNMENT-I

1. Consider the two image subsets, S_1 and S_2 shown in the following figure. For V = {1} determine whether these two subsets are (a) 4-adjacent, (b) 8-adjacent, or (c) *m*-adjacent.

	S_1				S_2				
0	0	0	0	0	0	0	1	1	0
1	0	0	1	0	0	1	0	0	1
1	0	0	1	0	1	1	0	0	0
0	0	1	1	1	0	0	0	0	0
0	0	1	1	1	0	0	1	1	1

2. Calculate DCT Matrix for N=4.
3. Perform histogram equalization of an image whose pixel intensity distribution is given in table:

Gray Level	0	1	2	3	4	5	6	7
Number of Pixels	790	1023	850	656	329	245	122	81

4. List the Formulae and Sketch the perceptive plot, Cross sectional view and Graphical Representation of Image Sharpening Filters.
5. Explain In detail about Image Restoration model.

ASSIGNMENT-II

1. a) Decode the encoded string "010100000101011111 0100" by generating Huffman code for the following data.

Symbol	A1	A2	A3	A4	A5	A6
Probability	0.1	0.4	0.06	0.1	0.04	0.3

b) The arithmetic decoding process is the reverse of the encoding procedure. Decode the message 0.23355 given the coding model.

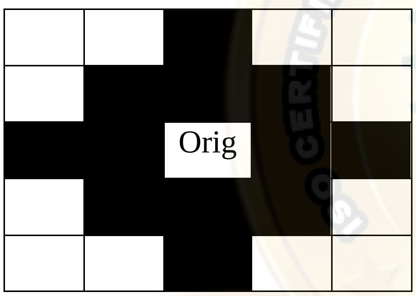
Symbol	A	E	I	O	U	!
Probability	0.2	0.3	0.1	0.2	0.1	0.1

2. a) With a neat block diagram explain about transform based image compression technique?

b) Explain different types of Redundancies and their removal methods?

3. a) For the Images A, B perform Erosion and Dilation.

A =



B =

b) Explain Opening, Closing and Hit or Miss transformation Algorithms with an example.

4. a) Explain how Segmentation can be performed using Derivatives and using Filter Masks.

b) Explain in detail about Similarity based segmentation.

5. a) Explain the Concept of Inverse filtering for Image restoration and derive the Transfer function of Wiener Filter or Minimum Mean Square Error Filter.

b) Explain Different Types of Restoration Filters which removes Noise and Periodic Noise.

LONG ANSWERS TYPE QUESTIONS

Bloom's Taxonomy

I. **R**emembering II. **U**nderstanding III. **Ap**plying IV.**An**alyzing
V. **E**valuating VI. **C**reating

UNIT-I (CO-1)

S.No	Question	Marks	Bloom's Level
1.	With a neat block diagram, Explain the fundamental steps in digital image processing.	5M	U
2.	Explain in detail about Image Sampling and Quantization?	5M	U
3.	Let V = {0,1} be the set of intensity values used to define adjacency. Compute the lengths of the shortest 4-path, 8-path, and *m*-path between *p* and *q* in the following image. If a particular path does not exist between these two points, explain why. $\begin{array}{ccccc} & 3 & 1 & 2 & 1\,(q) \\ & 2 & 2 & 0 & 2 \\ & 1 & 2 & 1 & 1 \\ (p) & 1 & 0 & 1 & 2 \end{array}$	5M	An
4.	With mathematical expressions explain the Haar transform for N=4.	5M	A
5.	With mathematical expressions explain DCT for N=4.	5M	A
6.	Find the Slant transform matrix for N = 4.	5 M	A
7.	Find the Walsh Transform matrix for N = 4.	5 M	A
8.	Explain how KL Transform Basis obtained?	5M	A

UNIT-II (CO-2)

S.No	Question	Marks	Bloom's Level
1.	Define Histogram and Sketch the histogram of basic image types? Explain with an example about histogram Equalization technique?	5M	U,An
2.	Explain Image Enhancement by linear and nonlinear Gray level transformation.	5M	U
3.	What is spatial filter? List out different types of spatial filters?	5M	U
4.	Explain in detail about image smoothing filters in frequency domain?	5M	U
5.	With a neat diagram explain the steps involved in enhancing the images in frequency domain?	5M	U
6.	Explain in detail about image Sharpening filters in frequency domain?	5M	U

UNIT-III (CO-3)

S.No	Question	Marks	Bloom's Level
1.	What is meant by image restoration? Explain the image Restoration model.	5M	R, An
2.	Compare and contrast image enhancement and image restoration techniques?	5M	An
3.	Derive the Transfer function of the wiener filter?	5M	An
4.	Explain about Interactive Restoration?	5M	R
5.	Explain about Inverse Filtering?	5M	An
6.	Explain about Constrained Least Squares Filters?	5M	An

UNIT-IV (CO-4)

S.No	Question	Marks	Bloom's Level (R,U,Ap, An,E,C)
1.	What is meant by the Gradient and the Laplacian? Discuss the role of Filter masks in image segmentation.	5M	U, Ap
2.	Discuss about Region based Segmentation?	5M	U
3.	Discuss about Thresholding techniques for Image Segmentation?	5M	U
4.	Discuss how Edges are linked for detecting Boundaries using a) Local Processing b) Regional Processing c) Global Processing	5M	U
5.	With an example explain about a) Erosion b) Dilation.	5 M	Ap, E
6.	With an example explain about a) Opening b) Closing c) Hit or Miss Transformation.	5M	Ap, E
7.	With an example explain different Morphological Algorithms.	5M	Ap,E

UNIT-V (CO-5)

S.No	Question	Marks	Bloom's Level (R,U,Ap, An,E,C)
1.	Explain different types of Redundancies and their removal methods?	5M	U
2.	With an example explain Huffman coding.	5M	Ap, An
3.	With an example explain Arithmetic coding.	5M	Ap, An
4.	With a neat block diagram explain about transform based image compression technique?	5M	U
5.	With a neat block diagram explain basic compression Model?	5M	U
6.	Compare Lossess and Lossy Predictive Coding?	5M	U
7.	How images are compressed using JPEG 2000 Standard?	5M	Ap

Digital Image Processing

SHORT ANSWERS TYPE QUESTIONS

Bloom's Taxonomy

I. **R**emembering II. **U**nderstanding III. **Ap**plying IV.**An**alyzing
V. **E**valuating VI. **C**reating

UNIT-I (CO-1)

S.No	Question	Marks	Bloom's Level
1.	Define Weber Ratio?	2M	R
2.	What is city block distance?	2M	R
3.	Define image resolution?	2M	R
4.	What are the steps involved in DIP?	2M	R
5.	Define Sampling and Quantization?	2M	R
6.	List the properties of Walsh Transform?	2M	R
7.	Define Digital Image?	2M	R
8.	What is Euclidean Distance?	2M	R
9.	Define Slant Transform?	2M	R
10.	Define Connectivity of pixels?	2M	R

UNIT-II (CO-2)

Digital Image Processing

S.No	Question	Marks	Bloom's Level
1.	What is mean by Image Subtraction?	2M	R
2.	What are Piecewise-Linear Transformations?	2M	R
3.	Specify the objective of image enhancement techniques.	2M	R
4.	Differentiate between linear spatial filter and non-linear spatial filter.	2M	R
5.	Define histogram.	2M	R
6.	What is the need of image enhancement?	2M	R
7.	Define Gaussian smoothing filter?	2M	R
8.	What is Image Negative?	2M	R
9.	What is arithmetic mean filter?	2M	R
10.	Sketch the block diagram of Filtering in Frequency Domain?	2M	R

UNIT-III (CO-3)

S.No	Question	Marks	Bloom's Level
1.	What is degradation function?	2M	R
2.	What is Gray-level interpolation?	2M	R
3.	What is meant by image restoration?	2M	R
4.	What is inverse filtering?	2M	R
5.	What is the difference between image restoration and image enhancement?	2M	R
6.	Draw the model of Image Restoration process.	2M	R
7.	Classify different types of Image Restoration methods?	2M	R
8.	How to estimate Image degradation function by experimentation?	2M	R
9.	Classify the types of Noise?	2M	R
10.	Differentiate Constrained and Unconstrained Restoration filters?	2M	R

UNIT-IV (CO-4)

S.No	Question	Marks	Bloom's Level
1.	What are the logic operations involving binary images?	2M	R
2.	What is convex hull?	2M	R
3.	Define region growing.	2M	R
4.	What are the three types of discontinuity in digital image?	2M	R
5.	List different types of discontinuities in digital image.	2M	R
6.	What is global, Local and dynamic threshold?	2M	R
7.	Define Duality of Erosion and Dilation?	2M	R
8.	Define different types of MASKS for Line Detection?	2M	R
9.	Explain how point can be detected in an image?	2M	R
10.	Define different types of MASKS for Edge Detection?	2M	R

UNIT-V (CO-5)

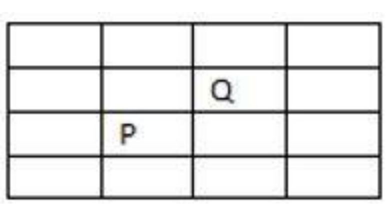
S.No	Question	Marks	Bloom's Level
1.	Define Compression Ratio.	2M	R
2.	What is Arithmetic Coding?	2M	R
3.	Define Huffman coding.	2M	R
4.	What are different compression methods?	2M	R
5.	What is the need of image compression?	2M	R
6.	Give the characteristics of lossless compression.	2M	R
7.	Discuss about irrelevant information?	2M	R
8.	Differentiate lossless and lossy compression model?	2M	R
9.	What is Subjective Fidelity Criteria?	2M	R
10.	Explain about Mapper?	2M	R

UNIT-I

DIGITAL IMAGE FUNDAMENTALS AND IMAGE TRANSFORMS

Objective Type Bits

1. Number of bits required to store a 128 X 128 image with 256 gray levels_________
a) 131072 b)327680 c) 294912 d)1048576
2. Number of bits required to store a 256 X 256 image with 32 gray levels _________
a) 131072 b)327680 c)294912 d)1048576
3. An image is considered to be a function of f(x,y) where f represents ______
a) Amplitude of image b) Resolution c) Width d) Height
4. An image is considered to be a function of f(x,y) where x,y represents ______
a) Amplitude of image b) Resolution c)Spatial co-ordinates d) Height
5. The process of extracting information from image is called _______
a) Image compression b) Enhancement c) Segmentation d) Image analysis
6. Sampling frequency greater than Nyquist rate is called______
a) Nyquist sampling b) over sampling c) Under sampling d) Critical sampling
7. Effect caused by under sampling is called______
a) Smoothing b) sharpening c) zooming d) Aliasing
8. What is the tool used in tasks such as zooming, shrinking etc?
a) Transforms b) filters c) Interpolation d) All the above
9. An image is 2400 pixels wide and 2400 pixels high. The image has resolution of 600 dpi. What will be the physical size of the image?
a) 1.7 inches X 1.7 inches b) 4 cm X 4 cm c) 4 inches X 4 inches d) 1.7 cm X 1.7 cm
10. In the following figure which of the option is true?



i) $Q \in N_4(P)$ ii) $Q \in N_8(P)$ iii) $Q \in N_D(P)$
a) Only i b) only ii c) both i and ii d) both ii and iii
11. Image transforms represents a given image as a series summation of a set of _____
a) Unitary matrices b) skew symmetric c) symmetric d) None of the above

12. DFT is applied to

a) Infinite Discrete Sequences b) finite Discrete Sequence c) Continuous Finite signal d) Continuous infinite signal

13. Condition for unitary matrix ______

a) $A^{-1} = A^{*T}$ b) $A = A^{*T}$ c) $A^{-1} = A^{*}$ d) $A^{-1} = A^{T}$

14. In Haar transform $H_K(Z)$ k is given as_______

a) $K= 2^P + q-1$ b) $K= 2^P + q+1$ c) $K= 2^P + q$ d) $K= p + q-1$

15. For K=0 Haar transform $H_K(Z)$ =________

a) $(1/\sqrt{N})$ b) $(2/\sqrt{N})$ c) $\sqrt{N}$ d) $(2\sqrt{N})$

16. Hotelling transform is also referred as ______

a) Haar transform b) Hadamard transform c) KL transform d) None of the above

17. In slant transform a_1 is given as _____

a) 1 b) -1 c) $(1/\sqrt{N})$ d) $\sqrt{N}$

18. DCT transform widely used in________

a) Image acquisition b)Image Segmentation c)Image Filtering d) Image Compression

19. Walsh transform is widely used in

a) Speech recognition b) Medical image processing c) Digital holography d) All the above

20. Hadamard transform is defined as _______

a) $H_n = H_{n-1} \otimes H_1$ b) $H_n = H_1 \otimes H_{n-1}$ c) $H_n = (\frac{1}{\sqrt{2}})\begin{bmatrix} H_{n-1} & H_{n-1} \\ H_{n-1} & -H_{n-1} \end{bmatrix}$ d) All the above

Fill in the Blanks

1. Gray scale Interval is given as ______
2. Digitizing spatial co-ordinates of the image is called ______
3. Digitizing Amplitude of the image is called ______
4. Intensity levels also known as______
5. Line pairs per unit distance is known as ______
6. Light with color is known as ______
7. Absence of receptor is called ________
8. Assume that a 10m high structure is observed from a distance of 20m. what is size of retinal image ? assume that the distance between the lens and retina is 14mm___________

9. Consider two images of 8 bit integer type F1=$\begin{bmatrix} 3 & 150 \\ 15 & 175 \end{bmatrix}$ and F2= $\begin{bmatrix} 150 & 125 \\ 55 & 155 \end{bmatrix}$ perform F1+F2 and F1-F2______________
10. Consider an image point [2 2]. Now rotate an image point around origin by 45^0 angle in anti clockwise direction around origin then modified image point is given by _______
11. Find the Kronecker product A⊗B of the images A = $\begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}$ and B = $\begin{bmatrix} 2 & 1 \\ 2 & 3 \end{bmatrix}$ _____________________
12. Hotelling transform is defined as _______
13. Haar transform for K≥1 is given as _______________
14. Slant transform is defined as___________
15. Walsh transform is defined as__________
16. DCT transform is defined as ________
17. Simplest form of the wavelet transform is ______
18. 2D DFT of image $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ is given as __________
19. Image is expanded in terms of discrete set of basis arrays called ___________
20. The basis image can be generated by _____________

KEY

I Objective Type

1. A
2. B
3. A
4. C
5. D
6. B
7. D
8. C
9. C
10. D

11. A
12. B
13. A
14. A
15. A
16. C
17. A
18. D
19. D
20. D

II Fill in the Blanks

1. $[L_{min}, L_{max}]$ or [0, L-1] or $[0, 2^k-1]$
2. Sampling
3. Quantization
4. Gray levels
5. Spatial resolution
6. Chromatic Light
7. Blind spot
8. 7mm
9. F1+F2= $\begin{bmatrix} 153 & 255 \\ 70 & 255 \end{bmatrix}$ F1-F2 = $\begin{bmatrix} 0 & 25 \\ 0 & 25 \end{bmatrix}$
10. [0 2.8]
11. A$\otimes$B = $\begin{pmatrix} 2 & 1 & 4 & 2 \\ 2 & 3 & 4 & 6 \\ 6 & 3 & 4 & 2 \\ 6 & 9 & 4 & 6 \end{pmatrix}$
12. KL transform $V = \emptyset^{*T} U$
13. Haar transform

$$h_0(x) \triangleq h_{0,0}(x) = \frac{1}{\sqrt{N}}, \qquad x \in [0, 1].$$

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

14. Slant transform

$$\mathbf{S}_n = \frac{1}{\sqrt{2}} \left[\begin{array}{cc|c|cc|c} 1 & 0 & \mathbf{0} & 1 & 0 & \mathbf{0} \\ a_n & b_n & & -a_n & b_n & \\ \hline \multicolumn{2}{c|}{\mathbf{0}} & \mathbf{I}_{(N/2)-2} & \multicolumn{2}{c|}{\mathbf{0}} & I_{(N/2)-2} \\ \hline 0 & 1 & \mathbf{0} & 0 & -1 & \mathbf{0} \\ -b_n & a_n & & b_n & a_n & \\ \hline \multicolumn{2}{c|}{\mathbf{0}} & I_{(N/2)-2} & \multicolumn{2}{c|}{\mathbf{0}} & -\mathbf{I}_{(N/2)-2} \end{array}\right] \left[\begin{array}{c|c} \mathbf{S}_{n-1} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{S}_{n-1} \end{array}\right]$$

15. Walsh transform

$$g(n,k) = \frac{1}{N}\prod_{i=0}^{m-1}(-1)^{b_i(n)b_{m-i-i}(k)}$$

16. DCT

$$c(k,n) = \begin{cases} \dfrac{1}{\sqrt{N}} & k = 0,\ 0 \le n \le N-1 \\ \sqrt{\dfrac{2}{N}} \cos \dfrac{\pi(2n+1)k}{2N}, & 1 \le k \le N-1,\ 0 \le n \le N-1 \end{cases}$$

17. Haar

18. $\begin{matrix} 4 & 0 \\ 0 & 0 \end{matrix}$

19. Basis Image

20. Unitary matrix

UNIT-II

IMAGE ENHANCEMENT (SPATIAL AND FREQUENCY DOMAIN)

Objective Type Bits

1. In formula S = T[r] T is the

a) Transformed image b) Transformation vector c) Transformation theorem d) Transformation function

2. The technique which improves the quality of the image for human perception is

a) image compression b) image restoration c) image segmentation d) image enhancement

3. Process used to correct the power law response is called____

a) Contrast modification b) brightness modification c) Gamma correction d) None

4. Process that increases the dynamic range of gray levels in an image is called ______

a) Linear stretching b) Thresholding c) Contrast stretching d) Color stretching

5. Histogram equalization is also referred as____

a) Image enhancement b) Histogram specification c) Histogram linearization d) None

6. Contrast adjustment is done by _________

a) Rotating b) Transforming c) Scaling d) all the above

7. Sum of all the components in normalized histogram is equal to

a) 1 b) 0 c) infinite d) None

8. Image enhancement is done in

a) Spatial domain b) Frequency domain c)Both a and b d) None of the above

9. Image enhancement in spatial domain is done by

a) Point operation b) Mask operation c) Global Operation d) All the above

10. Which of the following Technique is point operation

a) Brightness modification b) Contrast adjustment c) Histogram Manipulation d) All

11. Which of the following technique is used to obtain frequency domain filters using spatial domain filters

a) Convolution b) Impulse function c) both a and b d) None of the above

12. If size of the input image is M X N then padding image size if given as

a) 2M X 2N b) M X N c) M X M d) N X N

13. Accepting or rejecting of certain frequency components in the image called as___

a) Filter b) slicer c) Eliminator d) All the above

14. Sharpening of image is achieved by _______
a) Low pass filter b) High pass filter c) Band pass filter d) Band reject filter
15. Smoothing of image is achieved by______
a) Low pass filter b) High pass filter c) Band pass filter d) Band reject filter
16. Which of the following filter does not produce ringing effect?
a) Ideal low pass filter b) Butterworth low pass filter c) Gaussian low pass filter d) None
17. ______ filter is used to emphasize high frequency components representing image details without eliminating low frequency components representing basic form signal.
a) Gaussian filter b) Butterworth filter c) High boost filter d) Ideal filter
18. High pass filters are used for
a) Smoothing b) blurring c) Sharpening d) None
19. Smoothing an image results
a) Blur b) sharpen c) Segment d) None
20. In frequency domain, which is the equivalent operation of two images in spatial domain___
a) Correlation b) Enhancement c) Restoration d) Convolution

Fill in the blanks

1. Image plane itself is referred as _______
2. Process of manipulating an image, so that the result is more suitable than original image for specific application is referred as _____
3. To increase the brightness of the image the transformation function is given as______
4. Histogram of an image is given as ________
5. The transformation function of image negative is given as _______
6. The Log transformation function is given as _______
7. 3x3 Spatial Low pass filter mask is given as _______
8. 3X3 spatial high pass filter mask is given as ___________
9. The spatial filter mask is also known as___________
10. Spatial Gaussian filters are generated by using _______________
11. Ideal low pass filter is defined as _________
12. Butterworth low pass filter is defined as___________
13. Gaussian low pass filter is defined as _____________

14. Ideal high pass filter is defined as ________
15. Gaussian high pass filter is defined as ____________
16. Butterworth high pass filter is defined as _______________
17. The relationship between high pass and low pass filter in frequency domain is given as______________
18. Filtering in frequency domain consists of Modifying ___________ of an image.
19. The basic filtering equation is given as ______
20. For higher order values of ‘n’ in butterworth filter approaches to __________

KEY

I Objective Type

1. D
2. D
3. C
4. C
5. C
6. C
7. A
8. C
9. D
10. D
11. C
12. A
13. A
14. B
15. A
16. C
17. C
18. C
19. A
20. D

II Fill in the Blanks

1. Spatial domain
2. Image Enhancement
3. S = r+k
4. Histogram $h(r_k) = n_k$
5. S=L-1-r
6. S=c log(1+r)
7. 3x3 LPF

$$\frac{1}{9}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

8. 3x3HPF

$$\frac{1}{9}\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

9. Kernel or window
10. Pascal triangle
11. Ideal LPF

$$H(u, v) = \begin{cases} 1 & \text{if } D(u, v) \leq D_0 \\ 0 & \text{if } D(u, v) > D_0 \end{cases}$$

12. Butterworth LPF

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

13. Gaussian LPF

$$H(u, v) = e^{-D^2(u, v)/2D_0^2}$$

14. Ideal HPF

$$H(u, v) = \begin{cases} 0 & \text{if } D(u, v) \leq D_0 \\ 1 & \text{if } D(u, v) > D_0 \end{cases}$$

15. Butterworth HPF

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

16. Gaussian HPF

$$H(u,v) = 1 - e^{-D^2(u,v)/2D_0^2}$$

17. Relation is

$$H_{\text{HP}}(u,v) = 1 - H_{\text{LP}}(u,v)$$

18. Fourier transform
19. g(x,y) = IDFT[H(u,v).F(u,v)]
20. Ideal filter

UNIT-III

IMAGE RESTORATION

Objective Type Bits

1. The process of reconstructing the image that was degraded by a known degradation function is known as ___________

a) Image enhancement b) Image compression c) Image segmentation d) Image restoration

2. Degradation of image is mainly due to _________

a) Sensor distortions b) Optical distortions c) Low Illumination d) All the above

3. Purpose of restoration of an image is to gain ____________

a) Degraded image b) Original image c) Pixels d) co-ordinates

4. principal ways to estimate the degraded function are

a) By observation b) By experimentation c) By mathematical modeling d) All the Above

5. Process of restoration without knowing degradation function is known as ______

a) Iterative restoration b) Stochastic restoration c) Unconstrained restoration d) Blind Convolution

6. Process of restoration by minimizing the parameter of degradation is known as______

a) Iterative restoration b) Constrained restoration c) Unconstrained restoration d) Blind Convolution

7. Which filter used is used for removing salt and pepper noise in the image?

a) High boost filter b) sharpening filter c) gaussian filter d) Median filter

8. Restoration is done in

a) Spatial domain b) Frequency domain c) both a an b d) None of the above

9. Which of the following are restoration filters?

a) Min and Max Filters b) adaptive filters c) Order static filter d) All the above

10. Impulse noise is also referred as

a) Uniform noise b) Exponential noise c) Gaussian noise d) Salt and pepper noise

11. In Minimum mean square error filtering,

S1: Both the degradation function and statistical characteristics of noise considered for restoration.

S2: Both the image and noise are considered as random processes Objective find an estimate

a) Both S1 and S2 True b) Both S1 and S2 False c) S1 True and S2 false d) S1 false and S2 true

12. Purpose of restoration is to gain_______

a) Degraded image b) pixels c) Co-ordinates d) Original image

13. Which of the following is drawback of Wiener filter

a) Gamma doesn't require basic knowledge
b) Constant approximation of power spectrum is not always suitable
c) Constant approximation of power spectrum is always suitable
d) None of these

14. Which of the following is algebraic approach of restoration?

a) Constrained Restoration
b) Unconstrained Restoration
c) Both a & b
d) None

15. Which of the following are not used for restoration?

a) Inverse filter
b) Wiener filter
c) Iterative restoration
d) Masking

16. Which of the following statement(s) is(are) true?

Statement 1: The drawback of Wiener filter is that constant approximation of a ratio of power spectra is not always suitable

Statement 2: Optimum restoration using Iterative Selection of Gamma does not require knowledge of mean and variance of noise

a) Statement 1 is True; Statement 2 is False

b) Statement 1 is True; Statement 2 is True

c) Statement 1 is False; Statement 2 is True

d) Statement 1 is False; Statement 2 is False

17. Which of the following degradation model estimation method corresponds to blind deconvolution?

a) By observation

b) By experimentation

c) Mathematical modelling

d) All of these

18. Assume the degradation model for atmospheric turbulence is defined in frequency domain.If the value of K is increased, then atmospheric turbulence

$H(U,V) = e^{-k(U^2+V^2)^{5/6}}$

a) Decreases

b) Increases

c) Remains constant

d) None of these

19. Which of the following statement(s) is(are) true?

Statement 1: Actual degradation function is seldom known completely during image restoration.

Statement 2: Mathematical modeling for degradation cannot take into account environmental conditions that cause degradation.

a) Statement 1 is False; Statement 2 is True

b) Statement 1 is True; Statement 2 is True

c) Statement 1 is True; Statement 2 is False

d) Statement 1 is False; Statement 2 is False

20. In image restoration techniques, the degradation function satisfy the property:
    a) Linear
    b) Invariant
    c) Both a and b
    d) None of the above

Fill in the blanks

1. Image restoration process is ________ type compared with image enhancement.
2. Image Restoration model is given as__________________
3. Periodic noise in an image arises typically from ____________ during image acquisition.
4. Estimation function of image degradation using Observation is given by______
5. Estimation function of image degradation using Experimentation is given by______
6. Estimation function of image degradation using mathematical modeling is given by______
7. The restored image in spatial domain is given as _____________
8. The restored image in Frequency domain is given as _____________
9. Salt and pepper noise is defined as_____________
10. The process of reducing noise in image is referred as ___________
11. The wiener filter is equal to inverse filter when ________
12. The advantage of Wiener filter over inverse filter is________
13. Inverse filtering function is given by__________
14. Wiener filtering function is given by____________
15. Lucy-richard algorithm is _______________ restoration
16. Restoration technique if degradation function is known is referred as_______
17. Restoration technique if noise function is estimated is referred as______
18. Restoration technique if noise is zero it is referred as________
19. f(x, y) cannot be reconstructed by using $f(x,y)=H^{-1}(g(x,y)-\eta(x,y))$ in degradation model because it requires a_________
20. The drawback of the wiener filter is___________

KEY

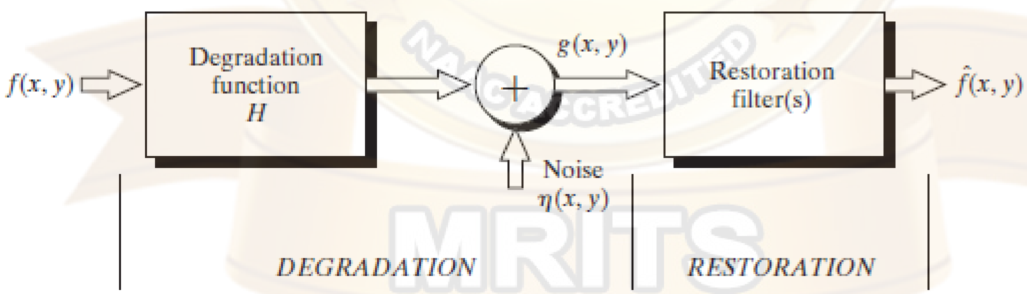
I Objective Type

1. D
2. D

3. B
4. D
5. B
6. A
7. D
8. C
9. D
10. D
11. A
12. D
13. B
14. C
15. D
16. A
17. D
18. B
19. C
20. C

II Fill in the Blanks

1. Objective
2. Model



3. Electrical or electro mechanical interference

4. Observation

$$H_s(u, v) = \frac{G_s(u, v)}{\hat{F}_s(u, v)}$$

5. Experimentation

$$H(u, v) = \frac{G(u, v)}{A}$$

6. Mathematical model

$$H(u, v) = e^{-k(u^2+v^2)^{5/6}}$$

7. Spatial domain

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

8. Frequency domain

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

9. Salt and pepper noise

$$p(z) = \begin{cases} P_a & \text{for } z = a \\ P_b & \text{for } z = b \\ 0 & \text{otherwise} \end{cases}$$

10. Denoising
11. Noise power spectrum is equal to zero.
12. Transfer function of wiener filter is chosen to minimize the mean square error using statistical information of both image and noise field.
13. Inverse filtering

$$\hat{F}(u, v) = F(u, v) + \frac{N(u, v)}{H(u, v)}$$

14. Wiener filtering

$$W(f_1, f_2) = \frac{H^*(f_1, f_2)S_{xx}(f_1, f_2)}{|H(f_1, f_2)|^2 S_{xx}(f_1, f_2) + S_{\eta\eta}(f_1, f_2)},$$

15. Non linear restoration
16. Deterministic Restoration
17. Constrained restoration
18. Unconstrained restoration
19. solution to a large number of simultaneous linear equations
20. In Wiener filter, constant approximation of a ratio of power spectra is not always suitable

UNIT-IV

Objective Type Bits

1. The application of image segmentation is(are) _____
   a) To detect isolated points b) To detect lines and edges c) Both a and b d) none
2. Two regions are said to be adjacent if they are
   a) Region b) Boundary c) edges d) connected set
3. Derivatives of an image can be modeled as _______
   a) Multiplication b) Addition c) division d) Difference
4. Following operator is used to detect edges?
   a) Gaussian b) Laplacian c) Ideal d) Butterworth
5. _______ is connected pixel that lie on boundary between two regions
   a) Edge b) Point c) Boundary d) None
6. Which of the following edge defines a perfect transition from one segment to another?
   a) Roof edge b) Ramp edge c) Line edge d) step edge
7. Segmentation using variable threshold in which the threshold value depends on characteristics of neighborhood of image
   a) Local threshold b) optimal threshold c) global threshold d) adaptive threshold
8. ____ is the position of sign change of first derivative among neighbouring points
   a) Edge crossing b) point crossing c) line crossing d) zero crossing
9. _______ is a procedure that groups pixels into larger region
   a) Region merging b) Region shrinking c) Region growing d) All the above
10. When threshold T only depends on Spatial Coordinates it is called ______ threshold
    a) Dynamic b) variable c) global d) binary
11. Probing of image is referred as
    a) Morphological processing b) Enhancement c) Segmentation d) Restoration
12. Window function of morphing operation is referred as
    a) Masking b) kernel c) Structuring element d) none
13. Morphological processing mainly follows
    a) SET theory b) Enhancement c) Restoration d) Compression
14. Which of the following is morphological processing

a) Erosion b) Dilation c) Both a and b d) None

15. Which of the following is not morphological processing

a) Opening b) closing c) both a and b d) None

16. ________ operation eliminates small holes

a) Closing b) opening c) dilation d) none

17. The following is the basic tool for shape detection

a) Dilation b) hit or miss transformation c) strel function d) thickening

18. The following operation fills gaps in the contour

a) Dilation b) opening c) closing d) erosion

19. The operation of erosion followed by dilation is called as

a) Dilation b) opening c) closing d) erosion

20. The operation of dilation followed by erosion is called as

a) Dilation b) opening c) closing d) erosion

Fill in the blanks

1. Image segmentation is classified into____________
2. Region oriented segmentation is classified into______________
3. ___________is similarity based segmentation
4. ___________ is discontinuities based segmentation
5. Sobel operator is given by________
6. _______ pixels at which the intensity of an image function changes abruptly.
7. When pixel value is greater than Threshold T, then pixel is referred as___
8. When pixel value is less than threshold T, then pixel is referred as____
9. Prewitt operator is given by______
10. If the threshold T is depends on spatial co-ordinates then thresholding is referred as__
11. Erosion is defined as_______
12. Dilation is defined as________
13. Opening is defined as________
14. Closing is defined as______
15. Hit and Miss transformation is given as____________
16. Dilation ________ the boundary.

17. Erosion ____________ the boundary.
18. Pruning method is complement method of _____
19. Closing operation removes_____
20. Opening operation removes_______

UNIT-IV

I Objective Type

1. C
2. D
3. D
4. A
5. A
6. D
7. A
8. D
9. C
10. A
11. A
12. C
13. A
14. C
15. D
16. A
17. B
18. C
19. B
20. C

II Fill in the Blanks

21. Discontinuities and similarity
22. Region growing, merging and splitting
23. Thresholding, region based
24. Isolated points, lines and edges
25. Sobel

–1	–2	–1		–1	0	1
0	0	0		–2	0	2
1	2	1		–1	0	1

Sobel

26. Edge
27. Object point
28. Background point

29. Prewitt

0	1	1		–1	–1	0
–1	0	1		–1	0	1
–1	–1	0		0	1	1

Prewitt

30. Dynamic Thresholding
31. Erosion

$$A \ominus B = \{z | (B)_z \subseteq A\}$$

32. Dilation

$$A \oplus B = \left\{z | [(\hat{B})_z \cap A] \subseteq A\right\}$$

33. Opening

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

34. Closing

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

35. Hit or Miss

$$\begin{aligned} A \circledast B &= (A \ominus B_1) \cap (A^c \ominus B_2) \\ &= (A \ominus B_1) - (A \oplus \hat{B}_2) \end{aligned}$$

36. Expands
37. Shrinks
38. Thinning
39. Holes
40. Sharp peaks

UNIT-V

IMAGE COMPRESSION

Objective Type Bits

1. Which of the following are types redundancy occur in images_____
   a) Coding b) inter pixel c) Psycho visual d) All
2. Coding redundancy can be removed by_____
   a) Huffmann coding b) Variable length coding c) Arithmetic coding d) All
3. Mapping is a_______ Process
   a) Reversible b) irreversible c) Noiseless d) Noisy
4. Quantizer is a_____ Process
   a) Reversible b) irreversible c) Noiseless d) Noisy
5. Which of the following are error free compression technique_____
   a) Huffmann coding b) Variable length coding c) Arithmetic coding d) All
6. Which of the following is/are Noisy compression technique __________
   a) Block transform b) length coding c) Arithmetic coding d) Huffmann coding
7. Which of the following block doesn’t exist in error free compression technique
   a) Mapper b) Symbol encoder c) Quantizer d) Symbol decoder
8. JPEG 2000 Standard is _________
   a) Lossy compression b) Lossless compression c) Both a and b d) None
9. Which of the following blocks consist in JPEG 2000 Standard ______
   a) Arithmetic coding b) DWT c) Both a and b d) None
10. Spatial redundancy is removed by_________
    a) LZW b) arithmetic coding c) Variable run length d) ALL
11. Reducing the amount of the data required to represent an image is known as_____
    a) Image segmentation b) Image restoration c) Image compression d) Morphological processing.
12. In transform based image compression, for efficient compression which of the following parameters is/are considered
    a) Sub image selection b) transform selection c) both a and b d) None of the above
13. For transform selection in transform based image compression which of the following parameters is/are considered

a) Non linearity b) De correlation c) Non orthogonality d) All the above

14. Preprocessing in JPEG 2000 standard involves

a) DWT b) Rate control c) Tiling d) Quantization

15. Preprocessing in JPEG 2000 standard involves

a) ICT b) level off set c) Tiling d) All the above

16. The basic component in lossless predictive coding______

a) Quantizer b) Predictor c) sampler d) DWT

17. The basic components in Lossy predictive coding______

a) Quantizer b) Predictor c) symbol encoder d) All

18. Arithmetic coding generates______

a) Linear block codes b) Non linear block codes c) Binary codes d) Gray codes

19. Which of the following is continuous tone image compression standard

a) JBIG b) PNG c) TIFF d) CCITT

20. Which of the following is binary image compression standard

a) JBIG b) BMP c) GIF d) JPEG

Fill in the blanks

1. Spatial redundancy also known as_______
2. Psycho visual redundancy also known as_________
3. JPEG stands for________
4. BPP stands for__________
5. Information is assigned a sequence of code symbols called a____
6. MPEG stands for________
7. TIFF stands for_______
8. AVS stands for________
9. Repeated information is referred as_________
10. Irrelevant information is referred as____________
11. Information loss using objective fidelity criterion is given by_______
12. As per TASO rating scale rating for unusable image is___
13. Mapper in image compression is used for _______

14. Quantizer in image compression model removes__________
15. Symbol encoder in image compression model used for______
16. For the 10:1 compression ratio how much of redundant data is removed_______
17. Compression ratio is given by _________
18. In arithmetic coding the final code is given by________
19. In JPEG 2000 compression standard ICT performs___________
20. In transform based compression the sub image size n is given by_______

KEY

I Objective Type

1. D
2. D
3. A
4. B
5. D
6. A
7. C
8. C
9. B
10. D
11. C
12. C
13. B
14. C
15. D
16. B
17. D
18. B
19. B or C
20. A

II Fill in the Blanks

21. Inter pixel Redundancy
22. Irrelevant Information
23. Joint Photography Expert Group
24. Bits Per Pixel

25. Codeword
26. Moving Picture Expert Group
27. Tagged Image File Format
28. Audio Video Standard
29. Redundancy
30. Redundancy
31. $e(x,y) = \hat{f}(x,y) - f(x,y)$
32. 6
33. Removing Spatial redundancy
34. Irrelevant information or psycho visual redundancy
35. Removing coding redundancy
36. 90%
37. b/b'
38. (Final lower interval + Final upper interval)/2
39. Irreversible color transform
40. Powers of 2

NPTEL VIDEOS:

https://nptel.ac.in/courses/117105135

STUDENT SEMINAR TOPICS

1. Digital Image Watermarking
2. Wavelets
3. Segmentation using clustering and super pixels
4. Feature Extraction
5. Image Pattern Classification
6. Noise models in Color Images
7. Convolutional Neural Networks
8. Deep Learning

UNIT-I

DIGITAL IMAGE FUNDAMENTALS AND IMAGE TRANSFORMS

Objective Type Bits

1. Number of bits required to store a 128 X 128 image with 256 gray levels_________

a) 131072 b)327680 c) 294912 d)1048576

2. Number of bits required to store a 256 X 256 image with 32 gray levels _________

a) 131072 b)327680 c)294912 d)1048576

3. An image is considered to be a function of f(x,y) where f represents ______

a) Amplitude of image b) Resolution c) Width d) Height

4. An image is considered to be a function of f(x,y) where x,y represents ______

a) Amplitude of image b) Resolution c)Spatial co-ordinates d) Height

5. The process of extracting information from image is called _______

a) Image compression b) Enhancement c) Segmentation d) Image analysis

6. Sampling frequency greater than Nyquist rate is called______

a) Nyquist sampling b) over sampling c) Under sampling d) Critical sampling

7. Effect caused by under sampling is called______

a) Smoothing b) sharpening c) zooming d) Aliasing

8. What is the tool used in tasks such as zooming, shrinking etc?

a) Transforms b) filters c) Interpolation d) All the above

9. An image is 2400 pixels wide and 2400 pixels high. The image has resolution of 600 dpi. What will be the physical size of the image?

a) 1.7 inches X 1.7 inches b) 4 cm X 4 cm c) 4 inches X 4 inches d) 1.7 cm X 1.7 cm

10. In the following figure which of the option is true?

		Q	
	P		

i) $Q \in N_4(P)$ ii) $Q \in N_8(P)$ iii) $Q \in N_D(P)$

a) Only i b) only ii c) both i and ii d) both ii and iii

11. Image transforms represents a given image as a series summation of a set of _____

a) Unitary matrices b) skew symmetric c) symmetric d) None of the above

12. DFT is applied to

a) Infinite Discrete Sequences b) finite Discrete Sequence c) Continuous Finite signal d) Continuous infinite signal

13. Condition for unitary matrix ______

a) $A^{-1} = A^{*T}$ b) $A = A^{*T}$ c) $A^{-1} = A^{*}$ d) $A^{-1} = A^{T}$

14. In Haar transform $H_K(Z)$ k is given as_______

a) $K= 2^P + q-1$ b) $K= 2^P + q+1$ c) $K= 2^P + q$ d) $K= p + q-1$

15. For K=0 Haar transform $H_K(Z)$ =________

a) $(1/\sqrt{N})$ b) $(2/\sqrt{N})$ c) $\sqrt{N}$ d) $(2\sqrt{N})$

16. Hotelling transform is also referred as ______

a) Haar transform b) Hadamard transform c) KL transform d) None of the above

17. In slant transform a_1 is given as _____

a) 1 b) -1 c) $(1/\sqrt{N})$ d) $\sqrt{N}$

18. DCT transform widely used in________

a) Image acquisition b)Image Segmentation c)Image Filtering d) Image Compression

19. Walsh transform is widely used in

a) Speech recognition b) Medical image processing c) Digital holography d) All the above

20. Hadamard transform is defined as _______

a) $H_n = H_{n-1} \otimes H_1$ b) $H_n = H_1 \otimes H_{n-1}$ c) $H_n = (\frac{1}{\sqrt{2}}) \begin{bmatrix} H_{n-1} & H_{n-1} \\ H_{n-1} & -H_{n-1} \end{bmatrix}$ d) All the above

Fill in the Blanks

1. Gray scale Interval is given as ______
2. Digitizing spatial co-ordinates of the image is called ______
3. Digitizing Amplitude of the image is called ______
4. Intensity levels also known as______
5. Line pairs per unit distance is known as ______
6. Light with color is known as ______
7. Absence of receptor is called ________
8. Assume that a 10m high structure is observed from a distance of 20m. what is size of retinal image ? assume that the distance between the lens and retina is 14mm___________
9. Consider two images of 8 bit integer type F1=$\begin{bmatrix} 3 & 150 \\ 15 & 175 \end{bmatrix}$ and F2= $\begin{bmatrix} 150 & 125 \\ 55 & 155 \end{bmatrix}$ perform F1+F2 and F1-F2______________

10. Consider an image point [2 2]. Now rotate an image point around origin by 45^0 angle in anti clockwise direction around origin then modified image point is given by _______

11. Find the Kronecker product A⊗B of the images A = $\begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}$ and B = $\begin{bmatrix} 2 & 1 \\ 2 & 3 \end{bmatrix}$ ____________________

12. Hotelling transform is defined as _______

13. Haar transform for K≥1 is given as _______________

14. Slant transform is defined as___________

15. Walsh transform is defined as__________

16. DCT transform is defined as ________

17. Simplest form of the wavelet transform is ______

18. 2D DFT of image $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ is given as __________

19. Image is expanded in terms of discrete set of basis arrays called ___________

20. The basis image can be generated by _____________

KEY

I Objective Type

1. A
2. B
3. A
4. C
5. D
6. B
7. D
8. C
9. C
10. D
11. A
12. B
13. A
14. A
15. A

16. C

17. A

18. D

19. D

20. D

II Fill in the Blanks

1. $[L_{min}, L_{max}]$ or [0, L-1] or $[0, 2^k-1]$
2. Sampling
3. Quantization
4. Gray levels
5. Spatial resolution
6. Chromatic Light
7. Blind spot
8. 7mm
9. F1+F2= $\begin{bmatrix} 153 & 255 \\ 70 & 255 \end{bmatrix}$ F1-F2 = $\begin{bmatrix} 0 & 25 \\ 0 & 25 \end{bmatrix}$
10. [0 2.8]
11. A⊗B = $\begin{pmatrix} 2 & 1 & 4 & 2 \\ 2 & 3 & 4 & 6 \\ 6 & 3 & 4 & 2 \\ 6 & 9 & 4 & 6 \end{pmatrix}$
12. KL transform $V = \emptyset^{*T} U$
13. Haar transform

$$h_0(x) \triangleq h_{0,0}(x) = \frac{1}{\sqrt{N}}, \qquad x \in [0, 1].$$

$$h_k(x) \triangleq h_{p,q}(x) = \frac{1}{\sqrt{N}} \begin{cases} 2^{p/2}, & \frac{q-1}{2^p} \leq x < \frac{q-\frac{1}{2}}{2^p} \\ -2^{p/2}, & \frac{q-\frac{1}{2}}{2^p} \leq x < \frac{q}{2^p} \\ 0, & \text{otherwise for } x \in [0,1] \end{cases}$$

14. Slant transform

$$\mathbf{S}_n = \frac{1}{\sqrt{2}} \left[\begin{array}{cc|c|cc|c} 1 & 0 & \mathbf{0} & 1 & 0 & \mathbf{0} \\ a_n & b_n & & -a_n & b_n & \\ \hline \multicolumn{2}{c|}{\mathbf{0}} & \mathbf{I}_{(N/2)-2} & \multicolumn{2}{c|}{\mathbf{0}} & I_{(N/2)-2} \\ \hline 0 & 1 & \mathbf{0} & 0 & -1 & \mathbf{0} \\ -b_n & a_n & & b_n & a_n & \\ \hline \multicolumn{2}{c|}{\mathbf{0}} & I_{(N/2)-2} & \multicolumn{2}{c|}{\mathbf{0}} & -\mathbf{I}_{(N/2)-2} \end{array}\right] \left[\begin{array}{c|c} \mathbf{S}_{n-1} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{S}_{n-1} \end{array}\right]$$

15. Walsh transform

$$g(n,k) = \frac{1}{N} \prod_{i=0}^{m-1} (-1)^{b_i(n) b_{m-i-i}(k)}$$

16. DCT

$$c(k,n) = \begin{cases} \frac{1}{\sqrt{N}} & k = 0,\ 0 \le n \le N-1 \\ \sqrt{\frac{2}{N}} \cos \frac{\pi(2n+1)k}{2N}, & 1 \le k \le N-1,\ 0 \le n \le N-1 \end{cases}$$

17. Haar

18. $\begin{matrix} 4 & 0 \\ 0 & 0 \end{matrix}$

19. Basis Image

20. Unitary matrix

UNIT-II

IMAGE ENHANCEMENT (SPATIAL AND FREQUENCY DOMAIN)

Objective Type Bits

1. In formula S = T[r] T is the

a) Transformed image b) Transformation vector c) Transformation theorem d) Transformation function

2. The technique which improves the quality of the image for human perception is

a) image compression b) image restoration c) image segmentation d) image enhancement

3. Process used to correct the power law response is called____

a) Contrast modification b) brightness modification c) Gamma correction d) None

4. Process that increases the dynamic range of gray levels in an image is called ______

a) Linear stretching b) Thresholding c) Contrast stretching d) Color stretching

5. Histogram equalization is also referred as____

a) Image enhancement b) Histogram specification c) Histogram linearization d) None

6. Contrast adjustment is done by _________

a) Rotating b) Transforming c) Scaling d) all the above

7. Sum of all the components in normalized histogram is equal to

a) 1 b) 0 c) infinite d) None

8. Image enhancement is done in

a) Spatial domain b) Frequency domain c)Both a and b d) None of the above

9. Image enhancement in spatial domain is done by

a) Point operation b) Mask operation c) Global Operation d) All the above

10. Which of the following Technique is point operation

a) Brightness modification b) Contrast adjustment c) Histogram Manipulation d) All

11. Which of the following technique is used to obtain frequency domain filters using spatial domain filters

a) Convolution b) Impulse function c) both a and b d) None of the above

12. If size of the input image is M X N then padding image size if given as

a) 2M X 2N b) M X N c) M X M d) N X N

13. Accepting or rejecting of certain frequency components in the image called as___

a) Filter b) slicer c) Eliminator d) All the above

14. Sharpening of image is achieved by _______

a) Low pass filter b) High pass filter c) Band pass filter d) Band reject filter

15. Smoothing of image is achieved by______

a) Low pass filter b) High pass filter c) Band pass filter d) Band reject filter

16. Which of the following filter does not produce ringing effect?

a) Ideal low pass filter b) Butterworth low pass filter c) Gaussian low pass filter d) None

17. ______ filter is used to emphasize high frequency components representing image details without eliminating low frequency components representing basic form signal.

a) Gaussian filter b) Butterworth filter c) High boost filter d) Ideal filter

18. High pass filters are used for

a) Smoothing b) blurring c) Sharpening d) None

19. Smoothing an image results

a) Blur b) sharpen c) Segment d) None

20. In frequency domain, which is the equivalent operation of two images in spatial domain___

a) Correlation b) Enhancement c) Restoration d) Convolution

Fill in the blanks

1. Image plane itself is referred as _______
2. Process of manipulating an image, so that the result is more suitable than original image for specific application is referred as _____
3. To increase the brightness of the image the transformation function is given as______
4. Histogram of an image is given as ________
5. The transformation function of image negative is given as _______
6. The Log transformation function is given as _______
7. 3x3 Spatial Low pass filter mask is given as _______
8. 3X3 spatial high pass filter mask is given as ___________
9. The spatial filter mask is also known as___________
10. Spatial Gaussian filters are generated by using _______________
11. Ideal low pass filter is defined as _________
12. Butterworth low pass filter is defined as___________
13. Gaussian low pass filter is defined as _____________
14. Ideal high pass filter is defined as ________
15. Gaussian high pass filter is defined as ____________

16. Butterworth high pass filter is defined as _______________
17. The relationship between high pass and low pass filter in frequency domain is given as_____________
18. Filtering in frequency domain consists of Modifying ___________ of an image.
19. The basic filtering equation is given as ______
20. For higher order values of ‘n’ in butterworth filter approaches to __________

KEY

I Objective Type

1. D
2. D
3. C
4. C
5. C
6. C
7. A
8. C
9. D
10. D
11. C
12. A
13. A
14. B
15. A
16. C
17. C
18. C
19. A
20. D

II Fill in the Blanks

1. Spatial domain
2. Image Enhancement
3. S = r+k
4. Histogram $h(r_k) = n_k$
5. S=L-1-r

6. S=c log(1+r)
7. 3x3 LPF

$$\frac{1}{9}\begin{bmatrix}1 & 1 & 1\\1 & 1 & 1\\1 & 1 & 1\end{bmatrix}$$

8. 3x3HPF

$$\frac{1}{9}\begin{bmatrix}-1 & -1 & -1\\-1 & 8 & -1\\-1 & -1 & -1\end{bmatrix}$$

9. Kernel or window
10. Pascal triangle
11. Ideal LPF

$$H(u, v) = \begin{cases}1 & \text{if } D(u, v) \le D_0\\0 & \text{if } D(u, v) > D_0\end{cases}$$

12. Butterworth LPF

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

13. Gaussian LPF

$$H(u, v) = e^{-D^2(u, v)/2D_0^2}$$

14. Ideal HPF

$$H(u, v) = \begin{cases}0 & \text{if } D(u, v) \le D_0\\1 & \text{if } D(u, v) > D_0\end{cases}$$

15. Butterworth HPF

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

16. Gaussian HPF

$$H(u, v) = 1 - e^{-D^2(u,v)/2D_0^2}$$

17. Relation is

$$H_{\mathrm{HP}}(u, v) = 1 - H_{\mathrm{LP}}(u, v)$$

18. Fourier transform
19. g(x,y) = IDFT[H(u,v).F(u,v)]
20. Ideal filter

UNIT-III

IMAGE RESTORATION

Objective Type Bits

1. The process of reconstructing the image that was degraded by a known degradation function is known as ___________

a) Image enhancement b) Image compression c) Image segmentation d) Image restoration

2. Degradation of image is mainly due to _________

a) Sensor distortions b) Optical distortions c) Low Illumination d) All the above

3. Purpose of restoration of an image is to gain ____________

a) Degraded image b) Original image c) Pixels d) co-ordinates

4. principal ways to estimate the degraded function are

a) By observation b) By experimentation c) By mathematical modeling d) All the Above

5. Process of restoration without knowing degradation function is known as ______

a) Iterative restoration b) Stochastic restoration c) Unconstrained restoration d) Blind Convolution

6. Process of restoration by minimizing the parameter of degradation is known as______

a) Iterative restoration b) Constrained restoration c) Unconstrained restoration d) Blind Convolution

7. Which filter used is used for removing salt and pepper noise in the image?

a) High boost filter b) sharpening filter c) gaussian filter d) Median filter

8. Restoration is done in

a) Spatial domain b) Frequency domain c) both a an b d) None of the above

9. Which of the following are restoration filters?

a) Min and Max Filters b) adaptive filters c) Order static filter d) All the above

10. Impulse noise is also referred as

a) Uniform noise b) Exponential noise c) Gaussian noise d) Salt and pepper noise

11. In Minimum mean square error filtering,

S1: Both the degradation function and statistical characteristics of noise considered for restoration.

S2: Both the image and noise are considered as random processes Objective find an estimate

a) Both S1 and S2 True b) Both S1 and S2 False c) S1 True and S2 false d) S1 false and S2 true

12. Purpose of restoration is to gain_______

a) Degraded image b) pixels c) Co-ordinates d) Original image

13. Which of the following is drawback of Wiener filter

a) Gamma doesn't require basic knowledge
b) Constant approximation of power spectrum is not always suitable
c) Constant approximation of power spectrum is always suitable
d) None of these

14. Which of the following is algebraic approach of restoration?

a) Constrained Restoration
b) Unconstrained Restoration
c) Both a & b
d) None

15. Which of the following are not used for restoration?

a) Inverse filter
b) Wiener filter
c) Iterative restoration
d) Masking

16. Which of the following statement(s) is(are) true?

Statement 1: The drawback of Wiener filter is that constant approximation of a ratio of power spectra is not always suitable

Statement 2: Optimum restoration using Iterative Selection of Gamma does not require knowledge of mean and variance of noise

a) Statement 1 is True; Statement 2 is False
b) Statement 1 is True; Statement 2 is True
c) Statement 1 is False; Statement 2 is True
d) Statement 1 is False; Statement 2 is False

17. Which of the following degradation model estimation method corresponds to blind deconvolution?

a) By observation

b) By experimentation

c) Mathematical modelling

d) All of these

18. Assume the degradation model for atmospheric turbulence is defined in frequency domain.If the value of K is increased, then atmospheric turbulence

$H(U,V) = e^{-k(U^2+V^2)^{5/6}}$

a) Decreases

b) Increases

c) Remains constant

d) None of these

19. Which of the following statement(s) is(are) true?

Statement 1: Actual degradation function is seldom known completely during image restoration.

Statement 2: Mathematical modeling for degradation cannot take into account environmental conditions that cause degradation.

a) Statement 1 is False; Statement 2 is True

b) Statement 1 is True; Statement 2 is True

c) Statement 1 is True; Statement 2 is False

d) Statement 1 is False; Statement 2 is False

20. In image restoration techniques, the degradation function satisfy the property:

a) Linear

b) Invariant

c) Both a and b

d) None of the above

Fill in the blanks

1. Image restoration process is ________ type compared with image enhancement.
2. Image Restoration model is given as___________________
3. Periodic noise in an image arises typically from _____________ during image acquisition.
4. Estimation function of image degradation using Observation is given by______
5. Estimation function of image degradation using Experimentation is given by______
6. Estimation function of image degradation using mathematical modeling is given by______
7. The restored image in spatial domain is given as ______________
8. The restored image in Frequency domain is given as ______________

9. Salt and pepper noise is defined as______________
10. The process of reducing noise in image is referred as ____________
11. The wiener filter is equal to inverse filter when _________
12. The advantage of Wiener filter over inverse filter is_________
13. Inverse filtering function is given by___________
14. Wiener filtering function is given by_____________
15. Lucy-richard algorithm is ________________ restoration
16. Restoration technique if degradation function is known is referred as_______
17. Restoration technique if noise function is estimated is referred as______
18. Restoration technique if noise is zero it is referred as_________
19. f(x, y) cannot be reconstructed by using $f(x,y)=H^{-1}(g(x,y)-\eta(x,y))$ in degradation model because it requires a__________
20. The drawback of the wiener filter is____________

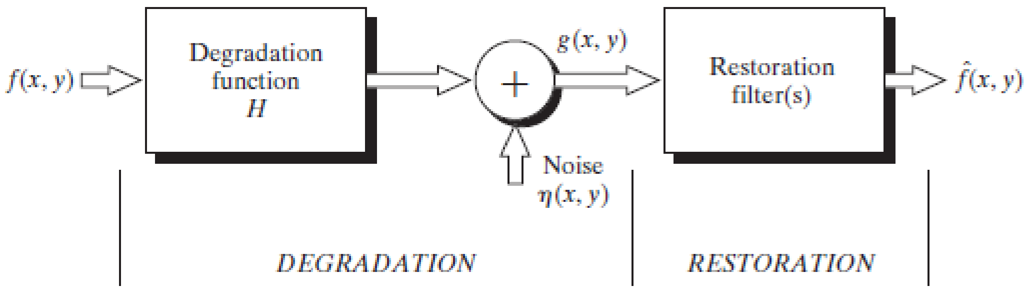
KEY

I Objective Type

1. D
2. D
3. B
4. D
5. B
6. A
7. D
8. C
9. D
10. D
11. A
12. D
13. B
14. C
15. D
16. A
17. D
18. B
19. C
20. C

II Fill in the Blanks

1. Objective
2. Model



3. Electrical or electro mechanical interference

4. Observation

$$H_s(u, v) = \frac{G_s(u, v)}{\hat{F}_s(u, v)}$$

5. Experimentation

$$H(u, v) = \frac{G(u, v)}{A}$$

6. Mathematical model

$$H(u, v) = e^{-k(u^2+v^2)^{5/6}}$$

7. Spatial domain

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

8. Frequency domain

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

9. Salt and pepper noise

$$p(z) = \begin{cases} P_a & \text{for } z = a \\ P_b & \text{for } z = b \\ 0 & \text{otherwise} \end{cases}$$

10. Denoising
11. Noise power spectrum is equal to zero.
12. Transfer function of wiener filter is chosen to minimize the mean square error using statistical information of both image and noise field.
13. Inverse filtering

$$\hat{F}(u, v) = F(u, v) + \frac{N(u, v)}{H(u, v)}$$

14. Wiener filtering

$$W(f_1, f_2) = \frac{H^*(f_1, f_2) S_{xx}(f_1, f_2)}{|H(f_1, f_2)|^2 S_{xx}(f_1, f_2) + S_{\eta\eta}(f_1, f_2)},$$

15. Non linear restoration
16. Deterministic Restoration
17. Constrained restoration
18. Unconstrained restoration
19. solution to a large number of simultaneous linear equations
20. In Wiener filter, constant approximation of a ratio of power spectra is not always suitable

UNIT-IV

Objective Type Bits

1. The application of image segmentation is(are) _____
   a) To detect isolated points b) To detect lines and edges c) Both a and b d) none
2. Two regions are said to be adjacent if they are
   a) Region b) Boundary c) edges d) connected set
3. Derivatives of an image can be modeled as _______
   a) Multiplication b) Addition c) division d) Difference
4. Following operator is used to detect edges?
   a) Gaussian b) Laplacian c) Ideal d) Butterworth
5. _______ is connected pixel that lie on boundary between two regions
   a) Edge b) Point c) Boundary d) None
6. Which of the following edge defines a perfect transition from one segment to another?
   a) Roof edge b) Ramp edge c) Line edge d) step edge
7. Segmentation using variable threshold in which the threshold value depends on characteristics of neighborhood of image
   a) Local threshold b) optimal threshold c) global threshold d) adaptive threshold
8. ____ is the position of sign change of first derivative among neighbouring points
   a) Edge crossing b) point crossing c) line crossing d) zero crossing
9. _______ is a procedure that groups pixels into larger region
   a) Region merging b) Region shrinking c) Region growing d) All the above

10. When threshold T only depends on Spatial Coordinates it is called ______ threshold
    a) Dynamic b) variable c) global d) binary
11. Probing of image is referred as
    a) Morphological processing b) Enhancement c) Segmentation d) Restoration
12. Window function of morphing operation is referred as
    a) Masking b) kernel c) Structuring element d) none
13. Morphological processing mainly follows
    a) SET theory b) Enhancement c) Restoration d) Compression
14. Which of the following is morphological processing
    a) Erosion b) Dilation c) Both a and b d) None
15. Which of the following is not morphological processing
    a) Opening b) closing c) both a and b d) None
16. ________ operation eliminates small holes
    a) Closing b) opening c) dilation d) none
17. The following is the basic tool for shape detection
    a) Dilation b) hit or miss transformation c) strel function d) thickening
18. The following operation fills gaps in the contour
    a) Dilation b) opening c) closing d) erosion
19. The operation of erosion followed by dilation is called as
    a) Dilation b) opening c) closing d) erosion
20. The operation of dilation followed by erosion is called as
    a) Dilation b) opening c) closing d) erosion

Fill in the blanks

1. Image segmentation is classified into______________
2. Region oriented segmentation is classified into________________
3. __________is similarity based segmentation
4. __________ is discontinuities based segmentation
5. Sobel operator is given by________
6. _______ pixels at which the intensity of an image function changes abruptly.
7. When pixel value is greater than Threshold T, then pixel is referred as___
8. When pixel value is less than threshold T, then pixel is referred as____
9. Prewitt operator is given by______

10. If the threshold T is depends on spatial co-ordinates then thresholding is referred as__
11. Erosion is defined as_______
12. Dilation is defined as________
13. Opening is defined as________
14. Closing is defined as______
15. Hit and Miss transformation is given as____________
16. Dilation ________ the boundary.
17. Erosion ___________ the boundary.
18. Pruning method is complement method of _____
19. Closing operation removes_____
20. Opening operation removes_______

UNIT-IV

I Objective Type

1. C
2. D
3. D
4. A
5. A
6. D
7. A
8. D
9. C
10. A
11. A
12. C
13. A
14. C
15. D
16. A
17. B
18. C
19. B
20. C

II Fill in the Blanks

21. Discontinuities and similarity
22. Region growing, merging and splitting
23. Thresholding, region based
24. Isolated points, lines and edges
25. Sobel

−1	−2	−1
0	0	0
1	2	1

−1	0	1
−2	0	2
−1	0	1

Sobel

26. Edge
27. Object point
28. Background point

29. Prewitt

0	1	1
−1	0	1
−1	−1	0

−1	−1	0
−1	0	1
0	1	1

Prewitt

30. Dynamic Thresholding
31. Erosion

$$A \ominus B = \{z | (B)_z \subseteq A\}$$

32. Dilation

$$A \oplus B = \left\{z | [(\hat{B})_z \cap A] \subseteq A\right\}$$

33. Opening

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

34. Closing

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

35. Hit or Miss

$$\begin{aligned} A \circledast B &= (A \ominus B_1) \cap (A^c \ominus B_2) \\ &= (A \ominus B_1) - (A \oplus \hat{B}_2) \end{aligned}$$

36. Expands
37. Shrinks
38. Thinning

39. Holes
40. Sharp peaks

UNIT-V

IMAGE COMPRESSION

Objective Type Bits

1. Which of the following are types redundancy occur in images_____
a) Coding b) inter pixel c) Psycho visual d) All
2. Coding redundancy can be removed by_____
a) Huffmann coding b) Variable length coding c) Arithmetic coding d) All
3. Mapping is a_______ Process
a) Reversible b) irreversible c) Noiseless d) Noisy
4. Quantizer is a_____ Process
a) Reversible b) irreversible c) Noiseless d) Noisy
5. Which of the following are error free compression technique_____
a) Huffmann coding b) Variable length coding c) Arithmetic coding d) All
6. Which of the following is/are Noisy compression technique __________
a) Block transform b) length coding c) Arithmetic coding d) Huffmann coding
7. Which of the following block doesn’t exist in error free compression technique
a) Mapper b) Symbol encoder c) Quantizer d) Symbol decoder
8. JPEG 2000 Standard is _________

a) Lossy compression b) Lossless compression c) Both a and b d) None

9. Which of the following blocks consist in JPEG 2000 Standard ______

a) Arithmetic coding b) DWT c) Both a and b d) None

10. Spatial redundancy is removed by_________

a) LZW b) arithmetic coding c) Variable run length d) ALL

11. Reducing the amount of the data required to represent an image is known as_____

a) Image segmentation b) Image restoration c) Image compression d) Morphological processing.

12. In transform based image compression, for efficient compression which of the following parameters is/are considered

a) Sub image selection b) transform selection c) both a and b d) None of the above

13. For transform selection in transform based image compression which of the following parameters is/are considered

a) Non linearity b) De correlation c) Non orthogonality d) All the above

14. Preprocessing in JPEG 2000 standard involves

a) DWT b) Rate control c) Tiling d) Quantization

15. Preprocessing in JPEG 2000 standard involves

a) ICT b) level off set c) Tiling d) All the above

16. The basic component in lossless predictive coding______

a) Quantizer b) Predictor c) sampler d) DWT

17. The basic components in Lossy predictive coding______

a) Quantizer b) Predictor c) symbol encoder d) All

18. Arithmetic coding generates______

a) Linear block codes b) Non linear block codes c) Binary codes d) Gray codes

19. Which of the following is continuous tone image compression standard

a) JBIG b) PNG c) TIFF d) CCITT

20. Which of the following is binary image compression standard

a) JBIG b) BMP c) GIF d) JPEG

Fill in the blanks

1. Spatial redundancy also known as_______
2. Psycho visual redundancy also known as_________
3. JPEG stands for________

4. BPP stands for__________
5. Information is assigned a sequence of code symbols called a____
6. MPEG stands for________
7. TIFF stands for_______
8. AVS stands for________
9. Repeated information is referred as_________
10. Irrelevant information is referred as____________
11. Information loss using objective fidelity criterion is given by_______
12. As per TASO rating scale rating for unusable image is___
13. Mapper in image compression is used for _______
14. Quantizer in image compression model removes__________
15. Symbol encoder in image compression model used for______
16. For the 10:1 compression ratio how much of redundant data is removed_______
17. Compression ratio is given by _________
18. In arithmetic coding the final code is given by________
19. In JPEG 2000 compression standard ICT performs___________
20. In transform based compression the sub image size n is given by_______

KEY

I Objective Type

1. D
2. D
3. A
4. B
5. D
6. A
7. C
8. C
9. B
10. D
11. C
12. C
13. B
14. C
15. D
16. B
17. D

18. B
19. B or C
20. A

II Fill in the Blanks

21. Inter pixel Redundancy
22. Irrelevant Information
23. Joint Photography Expert Group
24. Bits Per Pixel
25. Codeword
26. Moving Picture Expert Group
27. Tagged Image File Format
28. Audio Video Standard
29. Redundancy
30. Redundancy
31. $e(x,y) = \hat{f}(x,y) - f(x,y)$
32. 6
33. Removing Spatial redundancy
34. Irrelevant information or psycho visual redundancy
35. Removing coding redundancy
36. 90%
37. b/b’
38. (Final lower interval + Final upper interval)/2
39. Irreversible color transform
40. Powers of 2

SHORT ANSWERS TYPE QUESTIONS

Bloom's Taxonomy

I. **R**emembering II. **U**nderstanding III. **Ap**plying IV.**An**alyzing
V. **E**valuating VI. **C**reating

UNIT-I (CO-1)

S.No	Question	Marks	Bloom's Level
1.	Define Weber Ratio?	2M	R
2.	What is city block distance?	2M	R
3.	Define image resolution?	2M	R
4.	What are the steps involved in DIP?	2M	R
5.	Define Sampling and Quantization?	2M	R
6.	List the properties of Walsh Transform?	2M	R
7.	Define Digital Image?	2M	R
8.	What is Euclidean Distance?	2M	R
9.	Define Slant Transform?	2M	R
10.	Define Connectivity of pixels?	2M	R

UNIT-II (CO-2)

S.No	Question	Marks	Bloom's Level
1.	What is mean by Image Subtraction?	2M	R
2.	What are Piecewise-Linear Transformations?	2M	R
3.	Specify the objective of image enhancement techniques.	2M	R
4.	Differentiate between linear spatial filter and non-linear spatial filter.	2M	R
5.	Define histogram.	2M	R
6.	What is the need of image enhancement?	2M	R
7.	Define Gaussian smoothing filter?	2M	R
8.	What is Image Negative?	2M	R
9.	What is arithmetic mean filter?	2M	R
10.	Sketch the block diagram of Filtering in Frequency Domain?	2M	R

UNIT-III (CO-3)

S.No	Question	Marks	Bloom's Level
1.	What is degradation function?	2M	R
2.	What is Gray-level interpolation?	2M	R
3.	What is meant by image restoration?	2M	R
4.	What is inverse filtering?	2M	R
5.	What is the difference between image restoration and image enhancement?	2M	R
6.	Draw the model of Image Restoration process.	2M	R
7.	Classify different types of Image Restoration methods?	2M	R
8.	How to estimate Image degradation function by experimentation?	2M	R
9.	Classify the types of Noise?	2M	R
10.	Differentiate Constrained and Unconstrained Restoration filters?	2M	R

UNIT-IV (CO-4)

S.No	Question	Marks	Bloom's Level
1.	What are the logic operations involving binary images?	2M	R
2.	What is convex hull?	2M	R
3.	Define region growing.	2M	R
4.	What are the three types of discontinuity in digital image?	2M	R
5.	List different types of discontinuities in digital image.	2M	R
6.	What is global, Local and dynamic threshold?	2M	R
7.	Define Duality of Erosion and Dilation?	2M	R
8.	Define different types of MASKS for Line Detection?	2M	R
9.	Explain how point can be detected in an image?	2M	R
10.	Define different types of MASKS for Edge Detection?	2M	R

UNIT-V (CO-5)

S.No	Question	Marks	Bloom's Level
1.	Define Compression Ratio.	2M	R
2.	What is Arithmetic Coding?	2M	R
3.	Define Huffman coding.	2M	R
4.	What are different compression methods?	2M	R
5.	What is the need of image compression?	2M	R
6.	Give the characteristics of lossless compression.	2M	R
7.	Discuss about irrelevant information?	2M	R
8.	Differentiate lossless and lossy compression model?	2M	R
9.	What is Subjective Fidelity Criteria?	2M	R
10.	Explain about Mapper?	2M	R

LONG ANSWERS TYPE QUESTIONS

Bloom's Taxonomy

I. **R**emembering II. **U**nderstanding III. **Ap**plying IV.**An**alyzing
V. **E**valuating VI. **C**reating

UNIT-I (CO-1)

S.No	Question	Marks	Bloom's Level
1.	With a neat block diagram, Explain the fundamental steps in digital image processing.	5M	U
2.	Explain in detail about Image Sampling and Quantization?	5M	U
3.	Let $V = \{0,1\}$ be the set of intensity values used to define adjacency. Compute the lengths of the shortest 4-path, 8-path, and *m*-path between *p* and *q* in the following image. If a particular path does not exist between these two points, explain why. $\begin{array}{ccccc} & 3 & 1 & 2 & 1\,(q) \\ & 2 & 2 & 0 & 2 \\ & 1 & 2 & 1 & 1 \\ (p) & 1 & 0 & 1 & 2 \end{array}$	5M	An
4.	With mathematical expressions explain the Haar transform for N=4.	5M	A
5.	With mathematical expressions explain DCT for N=4.	5M	A
6.	Find the Slant transform matrix for N = 4.	5 M	A
7.	Find the Walsh Transform matrix for N = 4.	5 M	A
8.	Explain how KL Transform Basis obtained?	5M	A

UNIT-II (CO-2)

S.No	Question	Marks	Bloom's Level
1.	Define Histogram and Sketch the histogram of basic image types? Explain with an example about histogram Equalization technique?	5M	U,An
2.	Explain Image Enhancement by linear and nonlinear Gray level transformation.	5M	U
3.	What is spatial filter? List out different types of spatial filters?	5M	U
4.	Explain in detail about image smoothing filters in frequency domain?	5M	U
5.	With a neat diagram explain the steps involved in enhancing the images in frequency domain?	5M	U
6.	Explain in detail about image Sharpening filters in frequency domain?	5M	U

UNIT-III (CO-3)

S.No	Question	Marks	Bloom's Level
1.	What is meant by image restoration? Explain the image Restoration model.	5M	R, An
2.	Compare and contrast image enhancement and image restoration techniques?	5M	An
3.	Derive the Transfer function of the wiener filter?	5M	An
4.	Explain about Interactive Restoration?	5M	R
5.	Explain about Inverse Filtering?	5M	An
6.	Explain about Constrained Least Squares Filters?	5M	An

UNIT-IV (CO-4)

S.No	Question	Marks	Bloom's Level (R,U,Ap, An,E,C)
1.	What is meant by the Gradient and the Laplacian? Discuss the role of Filter masks in image segmentation.	5M	U, Ap
2.	Discuss about Region based Segmentation?	5M	U
3.	Discuss about Thresholding techniques for Image Segmentation?	5M	U
4.	Discuss how Edges are linked for detecting Boundaries using a) Local Processing b) Regional Processing c) Global Processing	5M	U
5.	With an example explain about a) Erosion b) Dilation.	5 M	Ap, E
6.	With an example explain about a) Opening b) Closing c) Hit or Miss Transformation.	5M	Ap, E
7.	With an example explain different Morphological Algorithms.	5M	Ap,E

UNIT-V (CO-5)

S.No	Question	Marks	Bloom's Level (R,U,Ap, An,E,C)
1.	Explain different types of Redundancies and their removal methods?	5M	U
2.	With an example explain Huffman coding.	5M	Ap, An
3.	With an example explain Arithmetic coding.	5M	Ap, An
4.	With a neat block diagram explain about transform based image compression technique?	5M	U
5.	With a neat block diagram explain basic compression Model?	5M	U
6.	Compare Lossess and Lossy Predictive Coding?	5M	U
7.	How images are compressed using JPEG 2000 Standard?	5M	Ap

ASSIGNMENT-I

1. Consider the two image subsets, S_1 and S_2 shown in the following figure. For V = {1} determine whether these two subsets are (a) 4-adjacent, (b) 8-adjacent, or (c) *m*-adjacent.

$$
\begin{array}{c|cccc|cccc|c}
 & & S_1 & & & & S_2 & & & \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\
1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1
\end{array}
$$

2. Calculate DCT Matrix for N=4.
3. Perform histogram equalization of an image whose pixel intensity distribution is given in table:

Gray Level	0	1	2	3	4	5	6	7
Number of Pixels	790	1023	850	656	329	245	122	81

4. List the Formulae and Sketch the perceptive plot, Cross sectional view and Graphical Representation of Image Sharpening Filters.
5. Explain In detail about Image Restoration model.

ASSIGNMENT-II

1. a) Decode the encoded string “0101000001010111110100” by generating Huffman code for the following data.

Symbol	A1	A2	A3	A4	A5	A6
Probability	0.1	0.4	0.06	0.1	0.04	0.3

b) The arithmetic decoding process is the reverse of the encoding procedure. Decode the message 0.23355 given the coding model.

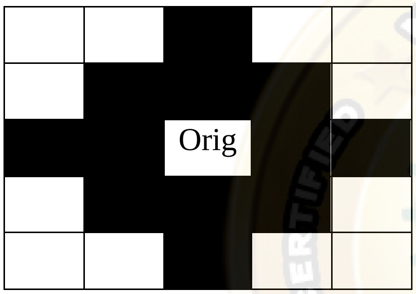
Symbol	A	E	I	O	U	!
Probability	0.2	0.3	0.1	0.2	0.1	0.1

2. a) With a neat block diagram explain about transform based image compression technique?

b) Explain different types of Redundancies and their removal methods?

3. a) For the Images A, B perform Erosion and Dilation.

A =



B =

b) Explain Opening, Closing and Hit or Miss transformation Algorithms with an example.

4. a) Explain how Segmentation can be performed using Derivatives and using Filter Masks.

b) Explain in detail about Similarity based segmentation.

5. a) Explain the Concept of Inverse filtering for Image restoration and derive the Transfer function of Wiener Filter or Minimum Mean Square Error Filter.

b) Explain Different Types of Restoration Filters which removes Noise and Periodic Noise.

STUDENT SEMINAR TOPICS

1. Digital Image Watermarking
2. Wavelets
3. Segmentation using clustering and super pixels
4. Feature Extraction
5. Image Pattern Classification
6. Noise models in Color Images
7. Convolutional Neural Networks
8. Deep Learning

R15

Code No: 127CJ

JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY HYDERABAD

B. Tech IV Year I Semester Examinations, November/December - 2018

DIGITAL IMAGE PROCESSING

(Common to ECE, ETM)

Time: 3 Hours **Max. Marks: 75**

Note: This question paper contains two parts A and B.
Part A is compulsory which carries 25 marks. Answer all questions in Part A. Part B consists of 5 Units. Answer any one full question from each unit. Each question carries 10 marks and may have a, b, c as sub questions.

PART- A

(25 Marks)

1.a) What is Digital Image Processing? [2]
b) Define Walsh Transform. [3]
c) What is the objective of image enhancement technique? [2]
d) List the steps involved in frequency domain filtering. [3]
e) Compare Image enhancement and Restoration techniques. [2]
f) Write the drawbacks of image restoration using inverse filtering. [3]
g) List the applications of segmentation. [2]
h) What is global, Local and dynamic or adaptive threshold? [3]
i) What is image compression? [2]
j) List out the JPEG 2000 standards. [3]

PART-B

(50 Marks)

2.a) Explain the basic concepts of sampling and quantization in the generation of digital image.
b) Explain the following terms:
i) Adjacency ii) Connectivity iii) Regions iv) Boundaries. [5+5]

OR

3.a) Compare and contrast different image transform techniques.
b) Find out the Slant transform matrix for N=8. [5+5]

4.a) Illustrate the histograms of basic Image types.
b) Discuss any one method of an image enhancement through point operation. [4+6]

OR

5.a) Explain image smoothing using ideal lowpass filters.
b) List various approaches used in Image enhancement and then discuss any one method of it. [5+5]

6. Discuss in detail the image restoration using minimum mean square error filtering. [10]

OR

7.a) How degradation function is estimated? Explain.
b) Briefly explain the interactive image restoration. [5+5]

8.a) Explain briefly the segmentation based on thresholding.
b) Discuss briefly the region based segmentation. [5+5]

OR

9. Discuss in detail the following morphological operations:
a) Erosion
b) Dilation [5+5]

10.a) What is Error Free Compression? Explain.
b) Discuss briefly the Image compression using Arithmetic coding. [5+5]

OR

11. Draw the functional block diagram of image compression system and explain the purpose of each block. [10]

--ooOoo--

R15

Code No:127CJ

JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY HYDERABAD
B.Tech IV Year I Semester Examinations, May/June - 2019
DIGITAL IMAGE PROCESSING
(Electronics and Communication Engineering)

Time: 3 Hours **Max.Marks: 75**

Note: This question paper contains two parts A and B.
Part A is compulsory which carries 25 marks. Answer all questions in Part A. Part B consists of 5 Units. Answer any one full question from each unit. Each question carries 10 marks and may have a, b, c as sub questions.

PART- A

(25 Marks)

1.a) How to represent the image? [2]
b) What is 4-, 8-, m- connectivity? [3]
c) What is High boost High pass filter? [2]
d) Compare linear and nonlinear gray level transformations. [3]
e) What are the advantages of Restoration? [2]
f) What are the different sources of degradation? [3]
g) What is erosion? [2]
h) How discontinuity property is used in image segmentation? [3]
i) What is mean by redundancy? [2]
j) What is fidelity? How it is used in image processing? [3]

PART-B

(50 Marks)

2.a) How to sample the image and how it differ from signal sampling?
b) Explore the relationship between pixels. [4+6]

OR

3.a) Define 2-D DFT and prove its convolution property and also write its applications.
b) Derive the 8×8 Slant transform matrix and write its order of sequence. [5+5]

4.a) Explain local enhancement techniques and compare it with global enhancement techniques.
b) Explain Histogram equalization method with example. [5+5]

OR

5.a) Consider the following image segment x and enhance it using the equation $y = k x$ where k is constant and y is output image.

0	1	2	3	4	5	6	7
54	35	64	53	123	43	56	45

b) Explain how low pass filter is used to enhance the image in frequency domain? [5+5]

6.a) Explain how image restoration improves the quality of image.
b) What is inverse filter? How it is used for image restoration? [5+5]

OR

7.a) How wiener filter is used for image restoration? What are the limitations of it?
b) What are the applications of restoration? [6+4]

8.a) How edge linking process is used to segment the image?
b) How to choose the threshold value while segmenting the image? [5+5]

OR

9.a) What are necessary condition to apply region based segmentation?
b) What is mean by Hit and Miss morphological operation? Write some example. [5+5]

10. Suppose the alphabet is *[A, B,C]*, and the known probability distribution is $P_A = 0.5$, $P_B = 0.4$, $P_c = 0.1$. For simplicity, let's also assume that bothencoder and decoder know that the length of the messages is always 3, so thereis no need for a terminator.
a) How many bits are needed to encode the message BBB by Huffmancoding?
b) How many bits are needed to encode the message BBB by arithmetic coding?
c) Analyze and compare the results of (a) and (b). [10]

OR

11.a) Draw the general block diagram of compression modal and explain the significance of each block.
b) Explain the loss-less prediction code for image compression with neat diagrams and equations. [5+5]

--ooOoo--

Code No: 127CJ **R15**

JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY HYDERABAD

B. Tech IV Year I Semester Examinations, December - 2019

DIGITAL IMAGE PROCESSING

(Electronics and Communication Engineering)

Time: 3 Hours **Max. Marks: 75**

Note: This question paper contains two parts A and B.
Part A is compulsory which carries 25 marks. Answer all questions in Part A. Part B consists of 5 Units. Answer any one full question from each unit. Each question carries 10 marks and may have a, b, c as sub questions.

PART- A

(25 Marks)

1.a) Define Weber Ratio. [2]
b) Explain three different level processes in image processing. [3]
c) What is mean by Image Subtraction? [2]
d) What is the need of image enhancement? [3]
e) What is the difference between image restoration and image enhancement? [2]
f) Give the relation for degradation model for Continuous function. [3]
g) List different types of discontinuities in digital image. [2]
h) Write short notes on morphological gradient. [3]
i) What is the need of image compression? [2]
j) Give the characteristics of lossless compression. [3]

PART-B

(50 Marks)

2. With mathematical expressions explain the Slant transform and explain how it is useful in Image processing. [10]

OR

3.a) Compute Haar Transform for N = 8 Value.
b) Explain how Fourier transforms are useful in digital image processing and explain the properties of Fourier transform. [5+5]

4.a) How Gray level transformation helps in contrast enhancement? Discuss.
b) Explain Spatial filtering in Image enhancement. [5+5]

OR

5.a) Discuss how the various filter masks are generated to sharpen images in spatial filters.
b) Illustrate homomorphic filtering approach for image enhancement. [5+5]

6.a) With relevant mathematical expressions, explain how a Wiener filter achieves restoration of a given degraded image.

b) Discuss the minimum mean square error filtering. [6+4]

OR

7.a) Explain linear position invariant degradation employed for image restoration.

b) Write about Constrained Least Squares Restoration in detail. [5+5]

8.a) What is meant by edge linking? Explain edge linking using local processing.

b) Explain the basics of intensity thresholding in image segmentation. [5+5]

OR

9.a) Explain with examples morphological operations dilation and erosion.

b) Explain the significance of thresholding in image segmentation. [5+5]

10.a) Describe arithmetic coding with an example for compression of image.

b) List and explain the steps involved in JPEG compression. [5+5]

OR

11.a) Draw and explain the general image compression system model.

b) Explain briefly the transform based compression. [5+5]

--ooOoo--

Code No: 117CJ **R13**

JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY HYDERABAD

B. Tech IV Year I Semester Examinations, November/December - 2016

DIGITAL IMAGE PROCESSING

(Electronics and Communication Engineering)

Time: 3 Hours **Max. Marks: 75**

Note: This question paper contains two parts A and B.
Part A is compulsory which carries 25 marks. Answer all questions in Part A. Part B consists of 5 Units. Answer any one full question from each unit. Each question carries 10 marks and may have a, b, c as sub questions.

PART- A

(25 Marks)

1.a) Define Weber Ratio [2]
b) What is city block distance [3]
c) What is mean by Image Subtraction? [2]
d) What are Piecewise-Linear Transformations [3]
e) What is degradation function? [2]
f) What is Gray-level interpolation? [3]
g) What are the logic operations involving binary images [2]
h) What is convex hull? [3]
i) Define Compression Ratio [2]
j) What is Arithmetic Coding? [3]

PART-B

(50 Marks)

2.a) Discuss the role of sampling and quantization with an example.
b) With a neat block diagram, explain the fundamental steps in digital image processing.[5+5]

OR

3.a) Discuss the Relationship between Pixels in detail.
b) Discuss optical illusions with examples. [5+5]

4.a) State different types of processing used for image enhancement.
b) Explain in detail smoothing frequency-domain filters related to images. [5+5]

OR

5.a) Explain any two methods used for digital image zooming and shrinking.
b) Discuss two dimensional orthogonal unitary transforms. [5+5]

6.a) Discuss the minimum mean square error filtering.
b) Explain the model of image degradation process. [5+5]

OR

7.a) Discuss in detail the Inverse Filtering.
b) Write about Constrained Least Squares Restoration in detail. [5+5]

8.a) Write Edge Linking And Boundary Detection.
b) Write about detection of discontinuities. [5+5]

OR

9.a) Discuss the Region Oriented Segmentation.
b) Explain about Hit or Miss Transformation. [5+5]

10.a) Explain about Lossy and Lossless Predictive Coding
b) Explain about the methods of removal of redundancy. [5+5]

OR

11.a) Discuss the Transform Based Compression.
b) Write a short note on JPEG 2000 Standards. [5+5]

--ooOoo--

R13

Code No: 117CJ

JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY HYDERABAD
B. Tech IV Year I Semester Examinations, March - 2017
DIGITAL IMAGE PROCESSING
(Electronics and Communication Engineering)

Time: 3 Hours **Max. Marks: 75**

Note: This question paper contains two parts A and B.
Part A is compulsory which carries 25 marks. Answer all questions in Part A. Part B consists of 5 Units. Answer any one full question from each unit. Each question carries 10 marks and may have a, b, c as sub questions.

Part- A (25 Marks)

1.a) Define image resolution. [2]
b) What are the steps involved in DIP? [3]
c) Specify the objective of image enhancement techniques. [2]
d) Differentiate between linear spatial filter and non-linear spatial filter. [3]
e) What is meant by image restoration? [2]
f) What is inverse filtering? [3]
g) Define region growing. [2]
h) What are the three types of discontinuity in digital image? [3]
i) Define huffman coding. [2]
j) What are different compression methods?

Part-B (50 Marks)

2.a) What is meant by digital image processing? What are the applications of it? How an image is represented digitally?
b) Non uniform sampling is useful for what type of images. Give reasons. [5+5]

OR

3.a) Is fast algorithm applicable for computation of Hadamard transform, if so what are the problems encountered in implementation.
b) Explain Discrete Cosine Transform and specify its properties. [5+5]

4.a) What is a histogram of an image? Sketch histograms of basic image types.
b) Discuss how histogram is useful for image enhancement. [5+5]

OR

5. What are the techniques used for image smoothing? Explain any one spatial and one frequency techniques used for image smoothing. [10]

6. Describe constrained least square filtering technique for image restoration and derive its transfer function. [10]

OR

7. Describe with mathematical model, both constrained and unconstrained restoration. [10]

8.a) Explain the segmentation techniques that are based on finding the regions.
b) Write the applications of segmentation. [7+3]

OR

9.a) Explain any two methods for linking the edge pixels to form a boundary of an object.
b) Explain with examples morphological operations dilation and erosion. [7+3]

10.a) Explain the schematics of image compression standard JPEG.
b) Draw and explain a general compression system model. [5+5]

OR

11.a) Describe in detail the lossless predictive coding error free compression.
b) Explain briefly the transform based compression. [5+5]

---ooOoo---

Code No: 117CJ **R13**

JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY HYDERABAD
B. Tech IV Year I Semester Examinations, November/December - 2017
DIGITAL IMAGE PROCESSING
(Common to ECE, ETM)

Time: 3 Hours **Max. Marks: 75**

Note: This question paper contains two parts A and B.
Part A is compulsory which carries 25 marks. Answer all questions in Part A. Part B consists of 5 Units. Answer any one full question from each unit. Each question carries 10 marks and may have a, b, c as sub questions.

Part- A

(25 Marks)

1.a) Define Sampling and Quantization. [2]
b) List the properties of Walsh Transform. [3]
c) Define histogram. [2]
d) What is the need of image enhancement? [3]
e) What is the difference between image restoration and image enhancement? [2]
f) Draw the model of Image Restoration process. [3]
g) List different types of discontinuities in digital image. [2]
h) What is global, Local and dynamic threshold? [3]
i) What is the need of image compression? [2]
j) Give the characteristics of lossless compression. [3]

Part-B

(50 Marks)

2. With mathematical expressions explain the Slant transform and explain how it is useful in Image processing. [10]

OR

3.a) List and explain the fundamental steps in digital image processing.
b) Discuss briefly the following:
i) Neighbours of pixels ii) connectivity. [5+5]

4.a) Explain the use of histogram statistics for image enhancement.
b) How Gray level transformation helps in contrast enhancement? Discuss. [5+5]

OR

5.a) Compare and contrast spatial domain and frequency domain techniques of Image enhancement.
b) Discuss any one frequency domain technique of Image smoothing. [5+5]

6. What is meant by image restoration? Explain the image degradation model. [10]

OR

7. Discuss in detail the image restoration using inverse filtering. [10]

8.a) Explain the basics of intensity thresholding in image segmentation.
b) Explain about morphological hit-or-miss transform. [5+5]

OR

9.a) Discuss in detail the edge linking using local processing.
b) Discuss briefly the region based segmentation. [6+4]

10.a) Discuss briefly the Image compression using Huffman coding.
b) What is the importance of compression in Image processing? [7+3]

OR

11.a) Draw and explain the image compression model.
b) List and explain the steps involved in JPEG compression. [6+4]

--ooOoo--

R13

Code No: 117CJ

JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY HYDERABAD
B. Tech IV Year I Semester Examinations, April/May - 2018
DIGITAL IMAGE PROCESSING
(Common to ECE, ETM)

Time: 3 Hours **Max. Marks: 75**

Note: This question paper contains two parts A and B. Part A is compulsory which carries 25 marks. Answer all questions in Part A. Part B consists of 5 Units. Answer any one full question from each unit. Each question carries 10 marks and may have a, b, c as sub questions.

PART- A

(25 Marks)

1.a) Define a digital image. [2]
b) Draw an image for image processing system. [3]
c) Present a note on smoothing linear filters. [2]
d) What are the applications of gray level slicing? [3]
e) Present a note on WEIGHT parameter. [2]
f) What are the spatial and frequency properties of noise? [3]
g) What are the applications of image segmentation? [2]
h) What is meant by watermarking? [3]
i) Define image compression. [2]
j) What is meant by error free compression? [3]

PART-B

(50 Marks)

2.a) Distinguish between digital image and binary image.
b) Explain a simple image model. [5+5]

OR

3.a) Explain the properties of slant transforrm.
b) Write short notes on hadamard transform. [5+5]

4. Explain image enhancement by point processing. [10]

OR

5.a) Explain about Ideal Low Pass Filter(ILPF) in frequency domain.
b) What is high frequency filtering? [5+5]

6.a) Write about component image observation model.
b) Discuss about Erlang noise. [5+5]

OR

7. Discuss about constrained and unconstrained restorations. [10]

8.a) Explain about Hough transform with an example.
b) What is the role of thresholding in segmentation? [5+5]

OR

9.a) Write short notes on dilation and erosion.
b) Give an overview of digital image watermarking methods. [5+5]

10. Discuss various image compression models. [10]

OR

11.a) Write a short note on fidelity criterion.
b) Explain Huffman coding technique. [5+5]

--ooOoo--

Code No: 137SE

RA

JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY HYDERABAD
B. Tech IV Year I Semester Examinations, December - 2019
DIGITAL IMAGE PROCESSING
(Electronics and Communication Engineering)

Time: 3 Hours **Max. Marks: 75**

Note: This question paper contains two parts A and B.
Part A is compulsory which carries 25 marks. Answer all questions in Part A. Part B consists of 5 Units. Answer any one full question from each unit. Each question carries 10 marks and may have a, b as sub questions.

PART – A

(25 Marks)

1.a) How hotelling transform differ from other? [2]
b) What is gray level in image processing? [3]
c) Write the application of sharpening filters. [2]
d) Name the categories of Image Enhancement. [3]
e) Give one example for the principal source of noise. [2]
f) How a degradation process is modeled? [3]
g) Write about linking edge points. [2]
h) Write the applications of segmentation. [3]
i) What is JPEG? [2]
j) What is the Need for Compression? [3]

PART – B

(50 Mark

2.a) Define Haar transform and Write the properties of Haar transform.
b) Discuss the role of sampling and quantization with an example. [5+5]

OR

3.a) Discuss the Relationship between Pixels in detail.
b) Find out the Slant transform matrix for N=8. [5+5]

4.a) Explain any two methods used for digital image zooming and shrinking.
b) Explain image enhancement by linear and non-linear gray level transformations. [5+5

OR

5.a) Contrast spatial domain and frequency domain techniques of Image enhancement.
b) Explain the use of histogram statistics for image enhancement. [5+5

6.a) How degradation function is estimated? Explain.
b) Explain about Constrained Least Squares Restoration in detail. [5+

OR

7.a) Briefly explain the interactive image restoration.
b) Discuss in detail the image restoration using inverse filtering. [5

8.a) How the discontinuity is detected in an image using segmentation?
b) Define the Hit or Miss Transformation and write its applications. [5+5]

OR

9. Explain region band segmentation methods in detail and compare them with other methods. [10]

10.a) Discuss the Transform Based Compression.
b) Explain about Lossy and Lossless Predictive Coding. [5+5]

OR

11.a) Explain about redundancies and their removal in compression.
b) Discuss briefly the Image compression using Huffman coding. [5+5]

---ooOoo---

Code No: 136BD **R16**

JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY HYDERABAD
B. Tech III Year II Semester Examinations, May - 2019
DIGITAL IMAGE PROCESSING
(Electronics and Communication Engineering)

Time: 3 hours **Max. Marks: 75**

Note: This question paper contains two parts A and B.
Part A is compulsory which carries 25 marks. Answer all questions in Part A. Part B consists of 5 Units. Answer any one full question from each unit. Each question carries 10 marks and may have a, b, c as sub questions.

PART - A

(25 Marks)

1.a) Compare and contrast digital image and binary image. [2]
b) Define spatial and gray level resolution. [3]
c) List the various areas of application of image subtraction [2]
d) Explain about median filtering. [3]
e) Explain about alpha-trimmed mean filter? [2]
f) Write short notes on Max and Min filters. [3]
g) What is meant by edge in a digital image? [2]
h) What is meant by optimal Thresholding? [3]
i) Write short notes on spatial redundancy. [2]
j) Explain the Fidelity criteria. [3]

PART - B

(50 Marks)

2.a) State and prove separability property of 2D-DFT.
b) Explain the role of Discrete Cosine transform in image processing. [5+5]

OR

3.a) Obtain the slant transforms matrix For N=8.
b) Develop an FFT algorithm using successive doubling method. [5+5]

4. Sketch perspective plot of an 2-D Ideal Low pass filter transfer function and filter cross section and explain its usefulness in Image enhancement. [10]

OR

5.a) What is meant by Histogram of an image. Write and explain with an example an algorithm for histogram equalization.
b) What is meant by the Gradient and the Laplacian? Discuss their role in image enhancement. [5+5]

6.a) Illustrate the use of adaptive median filter for noise reduction in an image.
b) Outline the different approaches to estimate the noise parameters in an image. [5+5]

OR

7.a) What are the different ways to estimate the degradation function? Explain.
b) Explain about noise reduction in an image using band reject and band pass filters. [5+5]

8.a) Explain about Global Processing by making use of Hough Transform?
b) Explain the following morphological algorithms [4+6]
i) Boundary extraction ii) Hole filling

OR

9.a) With necessary figures, explain the opening and closing operations.
b) Describe the procedure for image segmentation based on region growing with relevant examples. [5+5]

10.a) Consider an 8- pixel line of gray-scale data, {12,12,13,13,10,13,57,54}, which has been uniformly quantized with 6-bit accuracy. Construct its 3-bit IGS code.
b) What is bit-plane slicing? How it is used for achieving compression? [5+5]

OR

11.a) With the help of a block diagram explain about transform coding system.
b) Summarize the various types of data redundancies? [5+5]

---ooOoo---

R16

Code No: 136BD

JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY HYDERABAD

B. Tech III Year II Semester Examinations, December - 2019

DIGITAL IMAGE PROCESSING

(Electronics and Communication Engineering)

Time: 3 hours **Max. Marks: 75**

Note: This question paper contains two parts A and B.
Part A is compulsory which carries 25 marks. Answer all questions in Part A. Part B consists of 5 Units. Answer any one full question from each unit. Each question carries 10 marks and may have a, b, c as sub questions.

PART - A **(25 Marks)**

1.a) Define Image Sampling. [2]
b) What is Image Transform? List the applications of Transform. [3]
c) What is histogram equalization? [2]
d) Write the application of sharpening filters. [3]
e) What is inverse filtering? [2]
f) Draw the model of image degradation process. [3]
g) What is the advantage of using sobel operator? [2]
h) How the discontinuity is detected in an image using segmentation? [3]
i) What is lossy compression? [2]
j) What is the Need for Compression? [3]

PART - B **(50 Marks)**

2. Define 2-D DCT for N=4 and obtain N×N cosine transform matrix and write its properties and also its applications. [10]

OR

3.a) Define the following terms with respect to image
i) Distance measure ii) Connectivity iii) Neighborhood
b) Derive 4×4 Haar transform. [6+4]

4. Perform Histogram equalization of an image whose pixel intensity distribution is given in table. [10]

Gray Level	0	1	2	3	4	5	6	7
Number of pixels	790	1023	850	656	329	245	122	81

OR

5. What is meant by frequency filtering? Discuss in detail about smoothing and sharpening frequency filtering. [10]

6. Explain image degradation model /restoration process in detail. [10]

OR

7. Discuss the interactive restoration methods. [10]

8. What is segmentation? Explain the concept of region based segmentation techniques. [10]

OR

9. Explain the Morphological Erosion and Dilation combination process with one example. [10]

10. Explain how compression is achieved in transform coding with one example. [10]

OR

11. Explain the parts of the JPEG image compression block diagram. [10]

---ooOoo---