Bapatla Engineering College::Bapatla
Department of EIE
IV/IV B.Tech (Regular) Degree Examinations

Scheme of Evaluation

Sub:DIP

Sub Code:14EI804(B)

(March, 2019)

8^{th} Semester

Prepared by
SrinivasaRao.T
Assistant Professor

....

Bapatla Engineering College::Bapatla
Department of EIE
Spot Evaluation

Class: (4/4) EIE **Date: 28-03-2019**
Sub: DIP **Time: 90 Min**
Sub Code: 14EI804(B) **Maximum Marks: 60**

Answer all of the following questions

1. 12x1=12M

(a) . In digital imaging, a **pixel**, **pel** or **picture element** is a physical point in a raster image, or the smallest addressable element in an all points addressable display devices; so it is the smallest controllable element of a picture represented on the screen.

(b). The "dynamic range" is the difference between the darkest and lightest tones in an image, generally pure black and pure white. It's more often used to talk about the maximum dynamic range of a image.

(c). Image enhancement is the process of adjusting digital **images** so that the results are more suitable for display or further **image** analysis.

(d) . An **image histogram** is a type of **histogram** that acts as a graphical representation of the tonal distribution in a digital **image**. It plots the number of pixels for each tonal value. By looking at the **histogram** for a specific **image** a viewer will be able to judge the entire tonal distribution at a glance.

(e) . Images are often degraded during the data acquisition process. The degradation may involve blurring, information loss due to sampling, quantization effects, and various sources of noise. The purpose of image restoration is to estimate the original image from the degraded data.

(f). Subjective fidelity (Viewed by Human):

- By absolute rating
- By means of side-by-side comparison of and f (x,y) and $\hat{f}\,(x,y)$

Objective fidelity:

Level of information loss can be expressed as a function of the original and the compressed and subsequently decompressed image.

MSE : one of many ways to quantify the difference between values implied by an estimator and the true values of the quantity being estimated

PSNR : The phrase peak signal-to-noise ratio, often abbreviated PSNR, is an engineering term for the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation.

(g) . Compression ratio: $$C_R = \frac{n_1}{n_2}$$

n_1 is numer of bits of original image
n_2 is the number of bits of compressed image

(h). A simple approach to image segmentation is to start from some pixels (seeds) representing distinct image regions and to grow them, until they cover the entire image .For region growing we need a rule describing a growth mechanism and a rule checking the homogeneity of the regions after each growth step.

(i) . Advantage of chain code:

a compressed contour representation

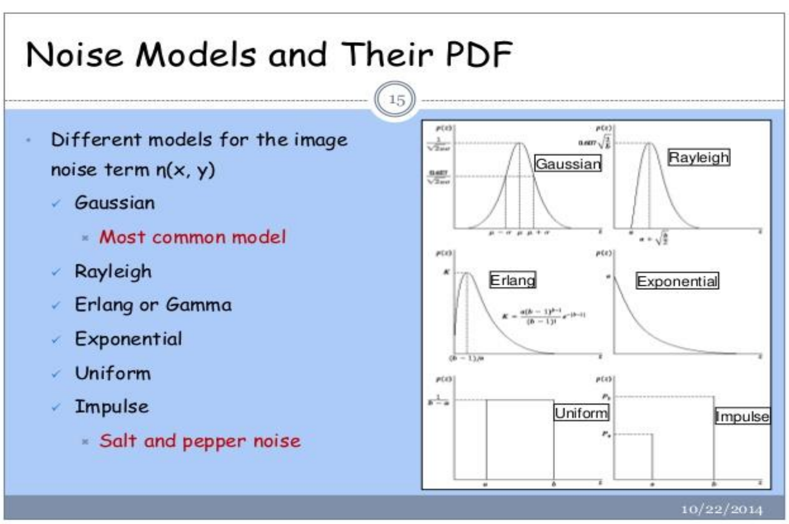
Disadvantages of chain code:

• chain code depends on the starting point Æ can be solved: treat the chain code as a circular sequence and redefining the starting point so that the resulting sequence of numbers is the smallest possible integer

• Operations such as scaling and rotation result in different contours that in practice cannot be normalized (due to a finite grid) and hence in different chain codes. Æ this problem cannot be completely solved but its effect can be reduced by resampling to a coarser grid before chain coding and by a proper orientation of the resampling grid

(j) . The length of a boundary is one of the simple boundary descriptor. The length of the boundary is approximately given by the number of pixels along that boundary.

(k) .



(l) . In image processing, it is often desirable to emphasize high frequency components representing the image details without eliminating low frequency components (such as sharpening). The high-boost filter can be used to enhance high frequency component.

The high-boost filtering technique can be implemented using the masks given below for

$A \geq 1$

0	−1	0		−1	−1	−1
−1	$A+4$	−1		−1	$A+8$	−1
0	−1	0		−1	−1	−1

$$W_{highboost} = AW_{allpass} + W_{highpass} = A\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 & 0 \\ -1 & A+4 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$

$$W_{highboost} = AW_{allpass} + W_{highpass} = A\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & -1 & 0 \\ -1 & 8 & -1 \\ 0 & -1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 & 0 \\ -1 & A+8 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$

UNIT-I

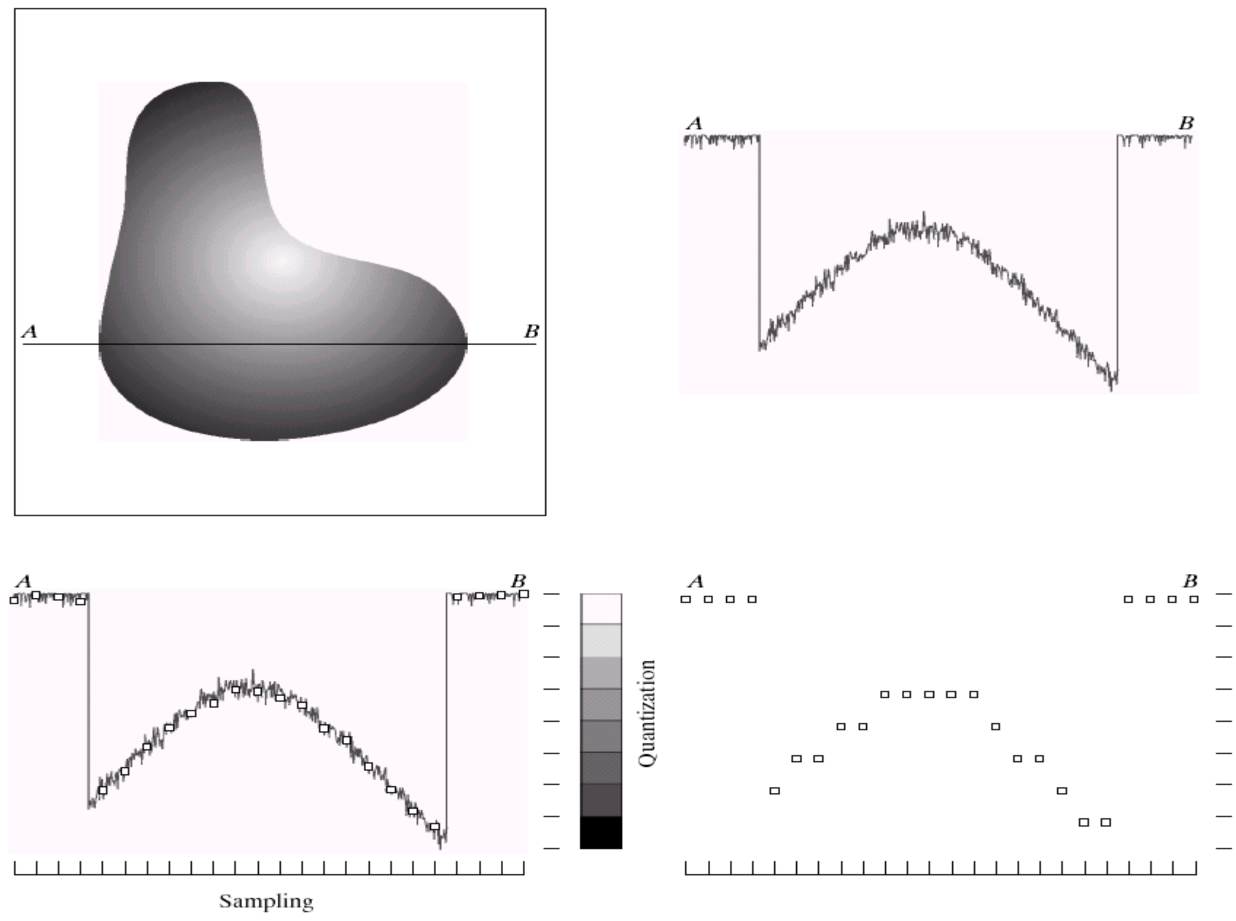
2. (a)

Sampling of digital image	3M
Quantization	3M

To create a digital image, we need to convert the continuous sensed data into digital form . This involves two processes: *sampling* and *quantization.*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.



a b
c d

FIGURE 2.16 Generating a digital image. (a) Continuous image. (b) A scan line from A to B in the continuous image, used to illustrate the concepts of sampling and quantization. (c) Sampling and quantization. (d) Digital scan line.

The result of sampling and quantization is a matrix of real numbers . 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. Sampling is the principal factor determining the *spatial resolution of an image*
subsampling the 1024*1024 image. 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.
This effect, caused by the use of an insufficient number of gray levels in smooth areas of a digital image, is called *false contouring.*
An adaptive sampling scheme can improve the appearance of an image, where the sampling would consider the characteristics of the image.

- i.e. fine sampling in the neighborhood of sharp gray-level transitions (e.g. boundaries)
- Coarse sampling in relatively smooth regions

• Considerations: boundary detection, detail content

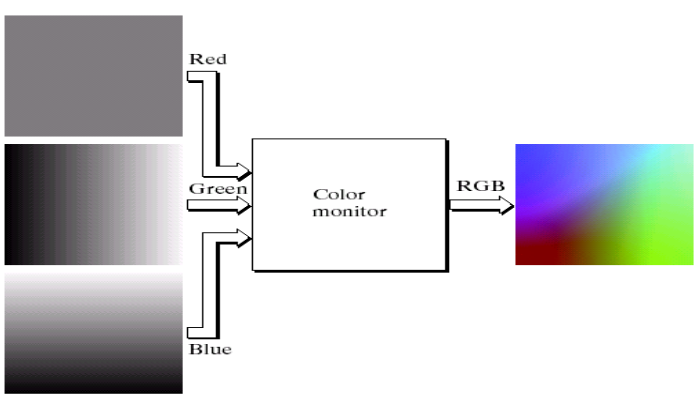
2.(b).RGB to HSI Color Models	4M

Primary color of pigments

Color that subtracts or absorbs a primary color of light and reflects or transmits the other two

Color of light:	R	G	B
Color of pigments:	absorb R	absorb G	absorb B
	Cyan	Magenta	Yellow

Application of additive nature of light colors is TV.



CIE XYZ model

- RGB -> CIE XYZ model

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.431 & 0.342 & 0.178 \\ 0.222 & 0.707 & 0.071 \\ 0.020 & 0.130 & 0.939 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

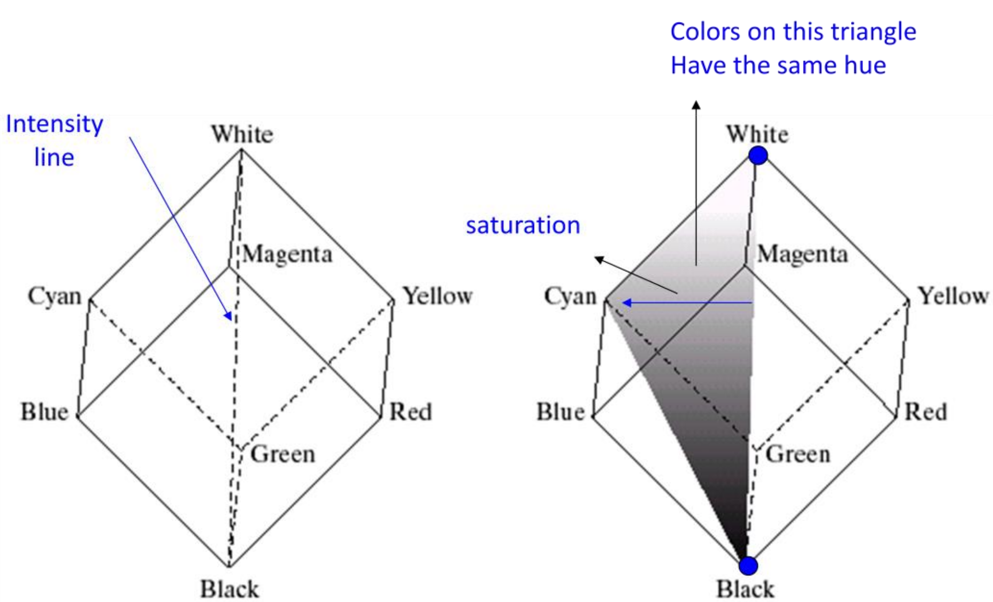
Normalized tristimulus values

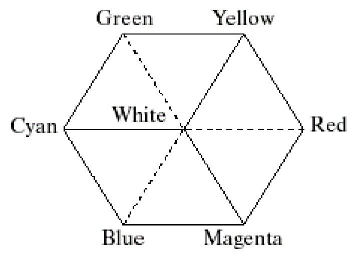
$$x = \frac{X}{X+Y+Z} \qquad y = \frac{Y}{X+Y+Z} \qquad z = \frac{Z}{X+Y+Z}$$

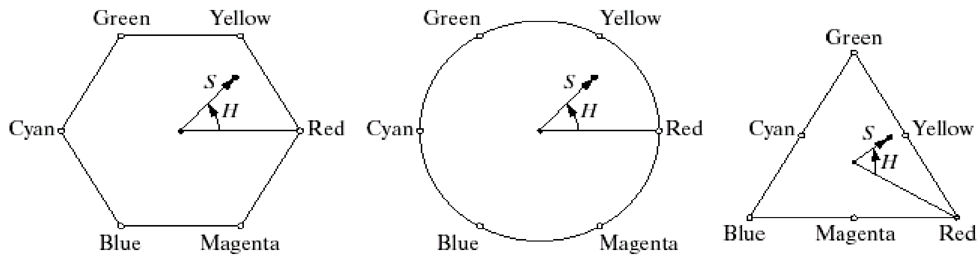
=> x+y+z=1. Thus, x, y (chromaticity coordinate) is enough to describe all colors

HSI color model

- Will you describe a color using its R, G, B components?
- Human describe a color by its hue, saturation, and brightness
  - Hue : color attribute
  - Saturation: purity of color (white->0, primary color->1)
  - Brightness: achromatic notion of intensity
- **RGB -> HSI model**







(OR)

3.(a).

Application of Digital Image Processing	**4M**

Today, there is almost no area of technical endeavor that is not impacted in some way by digital image processing.One of the simplest ways to develop a basic understanding of the extent of image processing applications is to categorize images according to their source (e.g., visual, X-ray, and so on).The principal energy source for images in use today is the electromagnetic energy spectrum. Other important sources of energy include acoustic, ultrasonic, and electronic (in the form of electron beams used in electron microscopy). Synthetic images, used for modelling and visualization, are generated by computer Images based on radiation from the EM spectrum are the most familiar, especially images in the X-ray and visual bands of the spectrum. Electromagnetic waves can be conceptualized as propagating sinusoidal waves of varying wavelengths, or they can be thought of as a stream of massless particles, each traveling in a wavelike pattern and moving at the speed of light. Each massless particle contains a certain amount (or bundle) of energy. Each bundle of energy is called a *photon. If spectral bands are grouped according to energy per* photon, we obtain the spectrum shown in Fig. 1.5, ranging from gamma rays (highest energy) at one end to radio waves (lowest energy) at the other.

Major uses of imaging based on gamma rays include nuclear medicine and astronomical observations. In nuclear medicine, the approach is to inject a patient with a radioactive isotope that emits gamma rays as it decays. Images are produced from the emissions collected by gamma ray detectors.

X-rays are among the oldest sources of EM radiation used for imaging.

The best known use of X-rays is medical diagnostics, but they also are used extensively in industry and other areas, like astronomy. X-rays for medical and industrial imaging are generated using an X-ray tube, which is a vacuum tube with a cathode and anode.

In digital radiography, digital images are obtained by one of two methods: (1) by digitizing X-ray films; or (2) by having the X-rays that pass through the patient fall directly onto devices (such as a phosphor screen) that convert X-rays to light . The light signal in turn is captured by a light-sensitive digitizing system.

Applications of ultraviolet "light" are varied. They include lithography, industrial inspection, microscopy, lasers, biological imaging, and astronomical observations. We illustrate imaging in this band with examples from microscopy and astronomy.

The infrared band often is used in conjunction with visual imaging, so we have grouped the visible and infrared bands in this section for the purpose of illustration. We consider in the following discussion applications in light microscopy, astronomy, remote sensing, industry, and law enforcement.

Another major area of visual processing is remote sensing, which usually includes several bands in the visual and infrared regions of the spectrum.

3. (b).Types of connectivity between the pixels	6M
Distance Measures	2M

- f(x,y): digital image
- Pixels: p,q
- Subset of pixels of f(x,y): S
- A pixel p at (x,y) has 2 horizontal and 2 vertical neighbors:
- (x+1,y), (x-1,y), (x,y+1), (x,y-1)
- This set of pixels is called the 4-neighbors of p: $N_4(p)$

- Each pixel is a unit distance from (x, y), and some of the neighbours of p lie outside the digital image if (x, y) is on the border of the image.
- The 4 diagonal neighbors of p are: ($N_D(p)$)
  - (x+1,y+1), (x+1,y-1), (x-1,y+1), (x-1,y-1)
  - $N_4(p) + N_D(p) \rightarrow N_8(p)$: the 8-neighbors of p
-
-

- **continuity**
- Two pixels are connected if:
  - They are neighbors (i.e. adjacent in some sense -- e.g. $N_4(p)$, $N_8(p)$, …)
  - Their gray levels satisfy a specified criterion of similarity (e.g. equality, …)
- V is the set of gray-level values used to define adjacency (e.g. V={1} for adjacency of pixels of value 1).
- **Adjacency**
- We consider three types of adjacency:
  - 4-adjacency: two pixels p and q with values from V are 4-adjacent if q is in the set $N_4(p)$
  - 8-adjacency : p & q are 8- adjacent if q is in the set $N_8(p)$
  - m-adjacency: p & q with values from V are m-adjacent if
    - q is in $N_4(p)$ or
    - q is in $N_D(p)$ and the set $N_4(p) \cap N_4(q)$ has no pixels with values from V
- Mixed adjacency is a modification of 8-adjacency and is used to eliminate the multiple path connections that often arise when 8-adjacency is used.
- **Path**
- A path (curve) from pixel p with coordinates (x,y) to pixel q with coordinates (s,t) is a sequence of distinct pixels:
  - (x_0,y_0), (x_1,y_1), …, (x_n,y_n)
  - where $(x_0,y_0)=(x,y)$, $(x_n,y_n)=(s,t)$, and
  - (x_i,y_i) is adjacent to (x_{i-1},y_{i-1}), for $1 \le i \le n$; n is the length of the path.
- If (xo, yo) = (x_n, y_n): a closed path
- 4-, 8-, m-paths can be defined depending on the type of adjacency specified.
- If p,q Î S, then q is connected to p in S if there is a path from p to q consisting entirely of pixels in S.

Boundary:

- 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 neighbours 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 .
- **Distance Measures**
- For pixels p,q,z with coordinates (x,y), (s,t), (u,v), D is a distance function or metric if:
  - $D(p,q) \ge 0$ (D(p,q)=0 if p=q)
  - D(p,q) = D(q,p) and
  - $D(p,z) \le D(p,q) + D(q,z)$.

- Euclidean distance:
  - $D_e(p,q) = [(x-s)^2 + (y-t)^2]^{1/2}$
  - Points (pixels) having a distance less than or equal to r from (x,y) are contained in a disk of radius r centered at (x,y).
- **D_4 distance (city-block distance):**
  - $D_4(p,q) = |x-s| + |y-t|$
  - forms a diamond centered at (x,y)
  - e.g. pixels with $D_4 \leq 2$ from p

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

$D_4 = 1$ are the 4-neighbors of p

- **D_8 distance (chessboard distance):**
  - $D_8(p,q) = \max(|x-s|,|y-t|)$
  - Forms a square centered at p
  - e.g. pixels with $D_8 \leq 2$ from p

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

D_8 = 1 are the 8-neighbors of p

Ex:

- assume $p_1, p_2, p_4 = 1$; p_1, p_3 = can have either 0 or 1

	p_3	p_4
p_1	p_2	
p		

If only connectivity of pixels valued 1 is allowed, and p_1 and p_3 are 0, the m-distance between p and p_4 is 2.
If either p_1 or p_3 is 1, the distance is 3.
If both p_1 and p_3 are 1, the distance is 4 (pp1p2p3p4)

UNIT-II

4.(a).Histogram equalization solution	8M

Histogram equalization: discrete solution

Probability (normalized histogram) of gray level r_k

$p(r_k)=n_k/n$, $k=0,1,2,\ldots, L-1$

Original Image								
Gray level(r_k)	0	1	2	3	4	5	6	7
No.Of Pixels(n_k)	8	10	10	2	12	16	4	2
Probability(P(r_k)	0.13	0.16	0.16	0.03	0.18	0.25	0.06	0.03
Cumulative probability	0.13	0.29	0.45	0.48	0.66	0.91	0.97	1.0
CP*20	2.6	5.8	9	9.6	13.2	18.2	19.4	20.0
Floor Rounding	2	5	9	9	13	18	19	20

Total number of pixels=64

We want to change the intensity to 1-20. Let us multiply cumulative probability by 20.

Round of decimal values to lower integer values.

So finally ,

Original Image	0	1	2	3	4	5	6	7
Histogram equalized output	2	5	9	9	13	18	19	20

We can see that the intensity range of the pixels have been increased and hence the

Histogram of the image will look more spread. This in turn is called Histogram equalization.

4.(b).Local enhancement	4M

Contrast enhancement is between the imaging device output and the display voltages. In order to see the dark areas, the light looks over exposed. Contrast Enhancement:

Two General Strategies

Optimize information transmission via intensity mappings

- Histogram as probability
- Information theoretic argument leads to uniform probability distribution, so histogram flattening

Optimize contrast at spatial scales where most important information on object lies

- Smaller scales are where object boundary info is dominant; at yet smaller scales noise is dominant

Global vs. Local

- Global
    - Intensity mappings
        - Intensity Windowing
        - Histogram Equalization
            - Achieving other histograms
    - Optimize contrast at boundary
        - Scale decomposition and then magnification of components at appropriate scales (e.g., MUSICA)
        - Unsharp Masking
- Locally adaptive
    - Intensity mapping
        - Adaptive Histogram Equalization
    - Optimize contrast at boundary
        - Geometry-limited diffusion

Intensity Windowing

- Dedicate the range of display intensities to a limited window of recorded intensities.
- Moves perceived object boundaries.

Unsharp Masking

- $I_{new} = \alpha \cdot (G_{small}*I - G_{big}*I) + G_{big}*I$
- Adds detail to a background image.
- Amplifies Mach bands.

MUSICA – Multiple level-of-detail images, combination of which forms result.

Adaptive Histogram Equalization

Intensities are mapped to their rank in the contextual region (window).

Enhances noise in smooth regions

Correction: limit the slope of the intensity mapping.

(OR)

5. (a). Smoothing	**3M**
shaping techniques	**3M**

Smoothing in the Frequency Domain

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

- Ideal low pass filter
- Butterworth (parameter: *filter order*)
- Gaussian

A 2-D ideal low-pass filter:

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

where D_0 is a specified nonnegative quantity and D(u,v) is the distance from point (u,v) to the center of the frequency rectangle.

- Center of frequency rectangle: (u,v)=(M/2,N/2)
- Distance from any point to the center (origin) of the FT:

$$D(u,v)=(u^2+v^2)^{1/2}$$

Ideal Filter (Lowpass)

- Ideal:
  - all frequencies inside a circle of radius D_0 are passed with no attenuation
  - all frequencies outside this circle are completely attenuated.
- Cutoff-frequency: the point of transition between H(u,v)=1 and H(u,v)=0 (D_0)
- To establish cutoff frequency loci, we typically compute circles that enclose specified amounts of total image power PT.

Butterworth Filter (Lowpass)

- To define a cutoff frequency locus: at points for which H(u,v) is down to a certain fraction of its maximum value.
  - When D(u,v) = D_0, H(u,v) = 0.5
  - i.e. down 50% from its maximum value of 1.

Gaussian Lowpass Filter

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

- D(u,v): distance from the origin of FT
- Parameter: σ=D0 (cutoff frequency)
- The inverse FT of the Gaussian filter is also a Gaussian

Sharpening Spatial Filters

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

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

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

- **we arrive at the following conclusions.**

(1) First-order derivatives generally produce thicker edges in an image.

(2) Second-order derivatives have a stronger response to fine detail, such as thin lines and isolated points.

(3) First order derivatives generally have a stronger response to a gray-level step.

(4) Second- order derivatives produce a double response at step changes in gray level.

We also note of second-order derivatives that, for similar changes in gray-level values in an image, their response is stronger to a line than to a step, and to a point than to a line. In most applications, the second derivative is better suited than the first derivative for image enhancement because of the ability of the former to enhance fine detail. For this, and for reasons of simpler implementation and extensions, we will focus attention initially on uses of the second derivative for enhancement.

5.(b).Histogram specification	6M

- Histogram equalization does not allow interactive image enhancement and generates only one result: an approximation to a uniform histogram. Sometimes though, we need to be able to specify particular histogram shapes capable of highlighting certain gray-level ranges. The procedure for histogram-specification based enhancement is:
- Equalize the levels of the original image using:

$$s = T(r_k) = \sum_{j=0}^{k} \frac{n_j}{n}$$

n: total number of pixels,

nj: number of pixels with gray level rj,

L: number of discrete gray levels

Specify the desired density function and obtain the transformation function G(z):

$$v = G(z) = \sum_{0}^{z} p_z(w) \approx \sum_{i=0}^{z} \frac{n_i}{n}$$

p_z: specified desirable PDF for output

Apply the inverse transformation function ; $z=G^{-1}(s)$ to the levels obtained in step 1.

The new, processed version of the original image consists of gray levels characterized by the specified density $p_z(z)$.

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

Either a particular probability density function (such as a Gaussian density) is specified and then a histogram is formed by digitizing the given function Or a histogram shape is specified on a graphic device and then is fed into the processor executing the histogram specification algorithm

UNIT-III

6.(a).Minimum Mean Square Error(Wiener) filtering for image restoration	6M

Here we discuss an approach that incorporates both the degradation function and statistical characteristics of noise into the restoration process.

Considering images and noise as random variables, the objective is to find an estimate $\hat{f}$ of the uncorrupted image f such that the mean square error between them is minimized.

The error measure is given by

$$e^2 = E\left\{(f - \hat{f})^2\right\} \qquad (5.8\text{-}1)$$

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

By assuming that

1. the noise and the image are uncorrelated;
2. one or the other has zero mean;
3. the intensity levels in the estimate are a linear function of the levels in the degraded image.

Then, the minimum of the error function in (5.8-1) is given in the frequency domain by the expression

$$
\begin{aligned}
\hat{F}(u,v) &= \left[\frac{H^*(u,v)S_f(u,v)}{S_f(u,v)|H(u,v)|^2 + S_\eta(u,v)}\right]G(u,v) \\
&= \left[\frac{H^*(u,v)}{|H(u,v)|^2 + S_\eta(u,v)/S_f(u,v)}\right]G(u,v) \qquad (5.8\text{-}2) \\
&= \left[\frac{1}{H(u,v)}\frac{|H(u,v)|^2}{|H(u,v)|^2 + S_\eta(u,v)/S_f(u,v)}\right]G(u,v)
\end{aligned}
$$

The terms in (5.8-2) are as follows:

$\hat{F}(u,v)$ is the frequency domain estimate

$G(u,v)$ is the transform of the degraded image

$H(u,v)$ is the transform of the degradation function

$H^*(u,v)$ is 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

This result is known as the Wiener filter, which 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 the entire denominator is zero for the same value(s) of u and v.

If the noise is zero, then the Wiener filter reduces to the inverse filter.

One of the most important measures is the signal-to-noise ratio, approximated using frequency domain quantities such as

$$SNR = \frac{\sum_{u=0}^{M-1} \sum_{v=0}^{N-1} |F(u,v)|^2}{\sum_{u=0}^{M-1} \sum_{v=0}^{N-1} |N(u,v)|^2} \tag{5.8-3}$$

The mean square error given in statistical form in (5.8-1) can be approximated also in terms a summation involving the original and restored images:

$$MSE = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \left[f(x,y) - \hat{f}(x,y) \right]^2 \tag{5.8-4}$$

If one considers the restored image to be signal and the difference between this image and the original to be noise, we can define a signal-to-noise ratio in the spatial domain as

$$SNR = \frac{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \hat{f}(x,y)^2}{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \left[f(x,y) - \hat{f}(x,y) \right]^2} \tag{5.8-5}$$

The closer f and $\hat{f}$ are, the larger this ratio will be.

If we are dealing with white noise, the spectrum $|N(u,v)|^2$ is a constant, which simplifies things considerably. However, $|F(u,v)|^2$ is usually unknown.

An approach is used frequently when these quantities are not known or cannot be estimated:

$$\hat{F}(u,v) = \left[\frac{1}{H(u,v)} \frac{|H(u,v)|^2}{|H(u,v)|^2 + K}\right] G(u,v) \qquad (5.8\text{-}6)$$

where K is a specified constant that is added to all terms of $|H(u,v)|^2$.

Note: White noise is a random signal (or process) with a flat power spectral density. In other words, the signal contains equal power within a fixed bandwidth at any center frequency.

6.(b).Transform based image compression system block diagram	**2M**
Explanation of each block	**4M**

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.

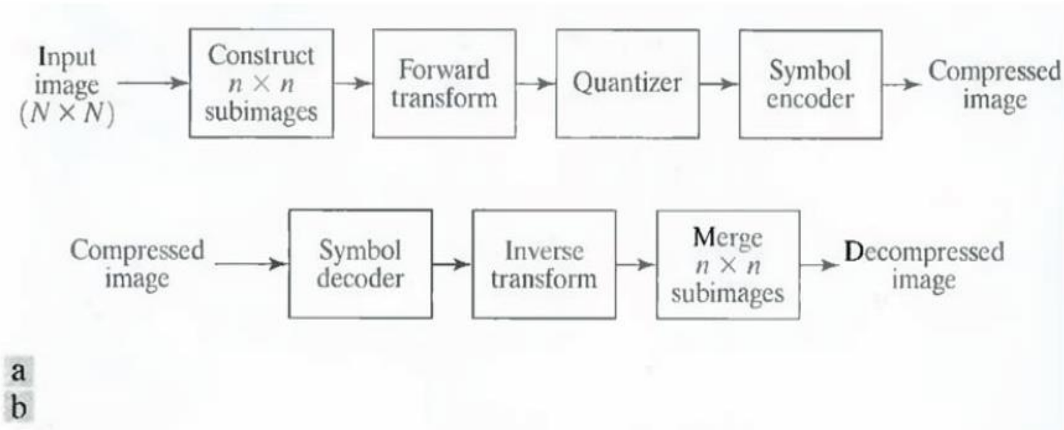


Fig. 10 A transform coding system: (a) encoder; (b) decoder.

Figure 10 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.

(OR)

7. (a).Airthematic coding	3M
Example	3M

Example 4.3.1:

Consider a three-letter alphabet $\mathcal{A} = \{a_1, a_2, a_3\}$ with $P(a_1) = 0.7$, $P(a_2) = 0.1$, and $P(a_3) = 0.2$. Using the mapping of Equation (4.1), $F_X(1) = 0.7$, $F_X(2) = 0.8$, and $F_X(3) = 1$. This partitions the unit interval as shown in Figure 4.1.

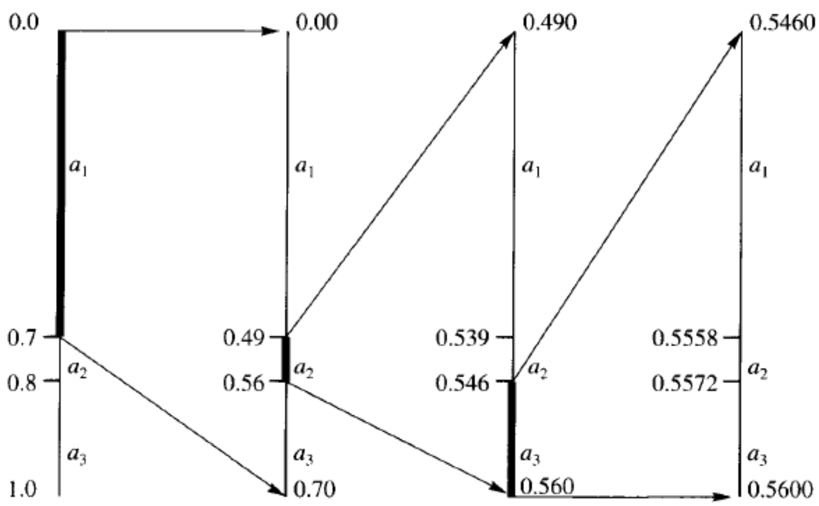


FIGURE 4.1 **Restricting the interval containing the tag for the input sequence $\{a_1, a_2, a_3, \ldots\}$.**

The partition in which the tag resides depends on the first symbol of the sequence being encoded. For example, if the first symbol is a_1, the tag lies in the interval $[0.0, 0.7)$; if the first symbol is a_2, the tag lies in the interval $[0.7, 0.8)$; and if the first symbol is a_3, the tag lies in the interval $[0.8, 1.0)$. Once the interval containing the tag has been determined, the rest of the unit interval is discarded, and this restricted interval is again divided in the same proportions as the original interval. Suppose the first symbol was a_1. The tag would be contained in the subinterval $[0.0, 0.7)$. This subinterval is then subdivided in exactly the same proportions as the original interval, yielding the subintervals $[0.0, 0.49)$, $[0.49, 0.56)$, and $[0.56, 0.7)$. The first partition as before corresponds to the symbol a_1, the second partition corresponds to the symbol a_2, and the third partition $[0.56, 0.7)$ corresponds to the symbol a_3. Suppose the second symbol in the sequence is a_2. The tag value is then restricted to lie in the interval $[0.49, 0.56)$. We now partition this interval in the same proportion as the original interval to obtain the subintervals $[0.49, 0.539)$ corresponding to the symbol a_1, $[0.539, 0.546)$ corresponding to the symbol a_2, and $[0.546, 0.56)$ corresponding to the symbol a_3. If the third symbol is a_3, the tag will be restricted to the interval $[0.546, 0.56)$, which can then be subdivided further. This process is described graphically in Figure 4.1.

Notice that the appearance of each new symbol restricts the tag to a subinterval that is disjoint from any other subinterval that may have been generated using this process. For the sequence beginning with $\{a_1, a_2, a_3, \ldots\}$, by the time the third symbol a_3 is received, the tag has been restricted to the subinterval $[0.546, 0.56)$. If the third symbol had been a_1 instead of a_3, the tag would have resided in the subinterval $[0.49, 0.539)$, which is disjoint from the subinterval $[0.546, 0.56)$. Even if the two sequences are identical from this point on (one starting with a_1, a_2, a_3 and the other beginning with a_1, a_2, a_1), the tag interval for the two sequences will always be disjoint. ♦

7.(b).

7.(b).Median filter	**2M**
Max and min filters	**2M**
Alpha trimmed mean filter	**2M**

Order statistic al filter:

These are spatial filters whose response is based on ordering (ranking)the pixels contained in the image area encompassed by the filter . The response of the filter at any point is determined by the ranking result.

Median filter :

The best known order statistic filter is the median filter which as its name implies, replaces the value of a pixel by the median of the gray levels in the neighbourhood of that pixel

$$\hat{f}(x,y) = \underset{(s,t)\varepsilon S_{xy}}{median}\{g(s,t)\}$$

The original value of the pixel is included in the computation of the median .Median filters are quite popular because ,for certain types of random noise, they provide excellent noise reduction capabilities , with considerably less blurring than linear smoothing filters of same size. Median filter are particularly effective in the presence of both bipolar and unipolar impulse noise.

Max and min filters :

Although the median filter is by far the order statistic filter most used in image processing .It is no means the only one. The median represents the 50th percentile of a ranked set of numbers(ranking lends itself to many other possibilities . For example using 100th percentile results in the so called max filter given by

Max filter: $$\hat{f}(x,y) = \max\{g(s,t)\}_{(s,t)\in S_{xy}}$$

Useful for finding the brightest points in an image

Pepper noise has very low values ,so max filter removes this noise also.

The 0th percentile filter is the min filter

Min filter: $$\hat{f}(x,y) = \min\{g(s,t)\}_{(s,t)\in S_{xy}}$$

This filter is useful for finding the darkest points in the image .Also it reduces salt noise as a result of the min operation.

Mid point filter:

The mid point filter simply computes the mid point between the maximum and minimum values in the area encompassed by the filter.

Midpoint filter $$\hat{f}(x,y) = \frac{1}{2}\left[\max\{g(s,t)\}_{(s,t)\in S_{xy}} + \min\{g(s,t)\}_{(s,t)\in S_{xy}}\right]$$

NB: combines order statistics and averaging.
Works best for randomly distributed noise such as Gaussian or uniform

Alpha-trimmed mean filter $$\hat{f}(x,y) = \frac{1}{mn-d}\sum_{(s,t)\in S_{xy}} g_r(s,t)$$

Where g_r represents the image g in which the $d/2$ lowest and $d/2$ highest intensity values in the neighbourhood S_{xy} were deleted
NB: $d = 0$ => arithmetic mean filter, $d = mn-1$ => median filter
For other values of d, useful when multiple types of noise (e.g. combination of salt-and-pepper and Gaussian Noise)

If we choose $d = \frac{(mn-1)}{2}$ the filter becomes a median filter.

For other values of d ,the alpha trimmed filter is useful in situations involving multiple types of noise ,such as combination of salt and pepper and Gaussian noise.

UNIT-IV

8.(a).Region growing	6M

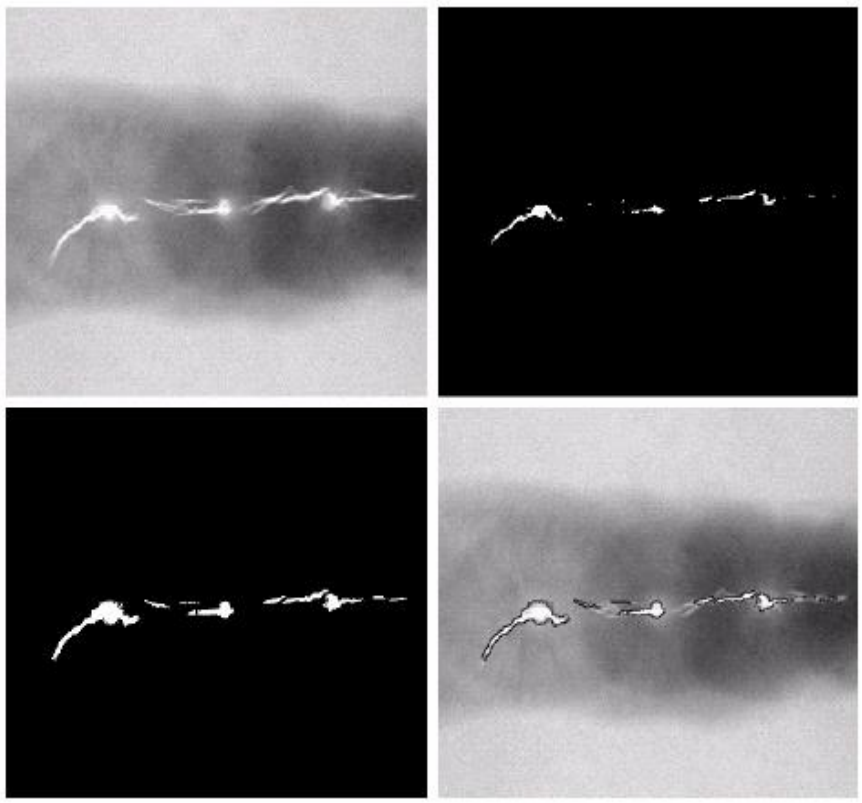
Region Growing:

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.



a b
c d

FIGURE 10.40 (a) Image showing defective welds. (b) Seed points. (c) Result of region growing. (d) Boundaries of segmented defective welds (in black). (Original image courtesy of X-TEK Systems, Ltd.).

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.

(OR)

9.(a).Line detection techniques	**6M**

3 basic types of gray-level discontinuities:

Points

Lines

Edges

- Common method of detection: run a mask through the image.

FIGURE 10.1 A general 3×3 mask.

w_1	w_2	w_3
w_4	w_5	w_6
w_7	w_8	w_9

Point Detection

This procedure involves computing the sum of products of the coefficients with the gray levels contained in the region encompassed by the mask

- R:The response of the mask at any point in the image is given by
- **T: nonnegative threshold:**

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

−1	−1	−1
−1	8	−1
−1	−1	−1

$$|R| \geq T$$

Line Detection

−1	−1	−1
2	2	2
−1	−1	−1

Horizontal

−1	−1	2
−1	2	−1
2	−1	−1

+ 45°

−1	2	−1
−1	2	−1
−1	2	−1

Vertical

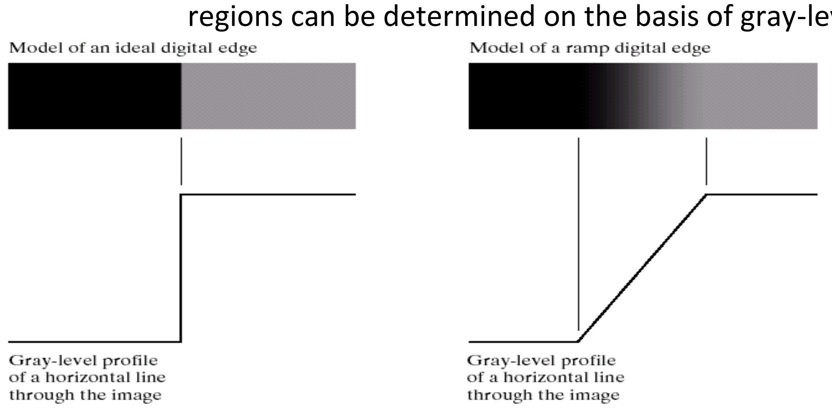
2	−1	−1
−1	2	−1
−1	−1	2

−45°

- **If at a certain point $|R_i|>|R_j|$, this point is more likely associated with a line in the direction of mask i.**

Edge Detection

- edge vs. boundary
- An edge is a set of connected pixels that lie on the boundary between two regions.
- An edge is a “local” concept where as a region boundary ,owing to the way it is defined ,is more global idea.
  - Assumption:
- the regions are sufficiently homogeneous, so that the transition between two regions can be determined on the basis of gray-level discontinuities alone.



a b

FIGURE 10.5 (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.

9.(b).Local thresholding techniques for segmentation	3M
global thresholding techniques for segmentation	3M

- **Local processing:**
- One of the simplest approaches for linking edge points is to analyze the characteristics of pixels in a small neighbourhood (say, 3x3 or 5x5).

• All points that are similar according to a set of predefined criteria are linked.

• 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(angle) of the gradient vector.

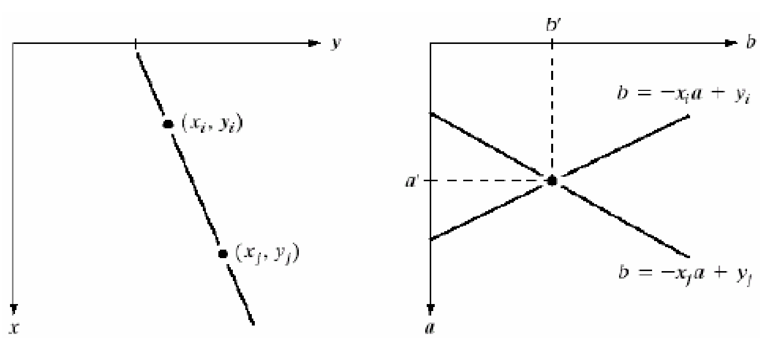
Thus an edge pixel with coordinates (x_0,y_0) *in a* magnitude to the pixel at (*x* , *y) if where E is a non negative threshold*

• An edge pixel at (x_0,y_0) *in the predefined* neighbourhood of *(x,y) has an angle similar to* the pixel at *(x,y) if*

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

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.



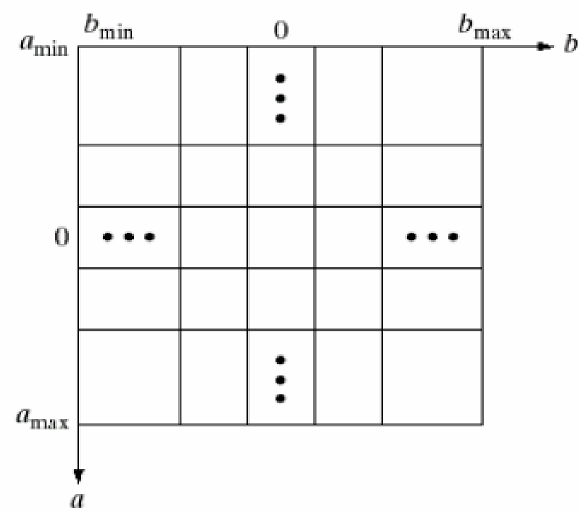
a b

FIGURE 10.17
(a) *xy*-plane.
(b) Parameter space.

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 3.1 illustrates these concepts.

The computational attractiveness of the Hough transform arises from subdividing the parameter space into so-called accumulator cells, as illustrated in Fig. 3.2, 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).

FIGURE 10.18 Subdivision of the parameter plane for use in the Hough transform.



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:

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