Image Processing and Analysis

Unit I

Digital Image Fundamentals

* **Image:** A two dimensional function $f(x,y)$ where x and y are spatial (plane) coordinates.

The amplitude of f at any pair of coordinates (x,y) is called the intensity or gray level of the image at that point.

* **Digital Image:** When x, y and the intensity values of f are all finite, discrete quantities, the image is called a digital image. A digital image is composed of a finite number of elements, each of which has a particular location and value. These elements are called pixels.

* **Image Types**

(i) binary images – black and white – pixels have either 0 (black) or 1 (white) values

(4B) eg. negative of images / outlines of images

(ii) grayscale images - have black, white, shades of gray – pixels range from 0 – 255

0 → black, 255 → white

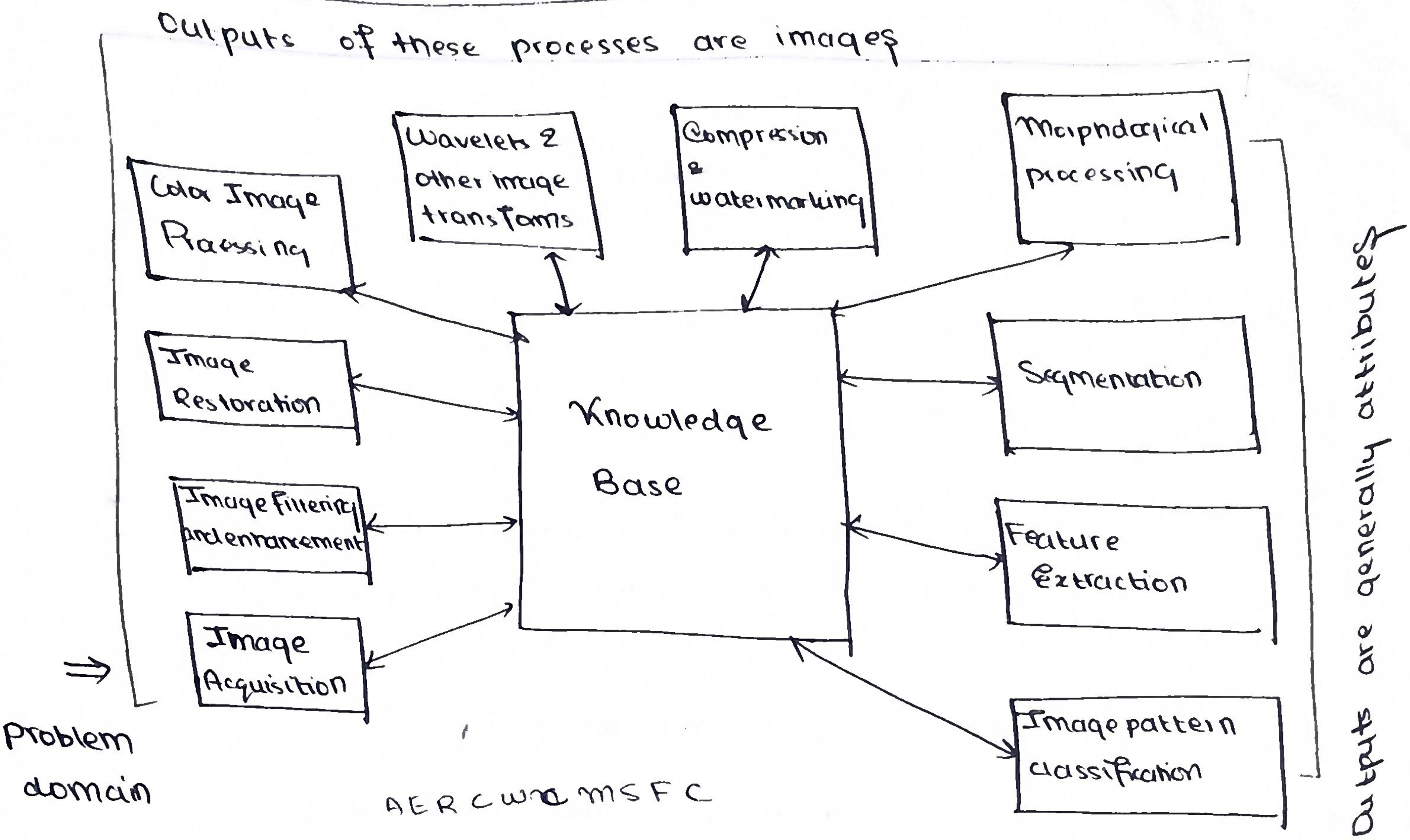
eg. black & white films

grayscale images have scalar values

(iii) RGB images - color images – r, g, b have vector values

r, g, b separately range from 0 – 255.

* Fundamental Steps in Digital Image Processing



① **Image Acquisition** : → capturing an image using sensors / a camera

→ sometimes i/p may be an already digitized image

→ preprocessing steps such as scaling are done

② **Image Enhancement** : → the process of manipulating an image so the result is more suitable than the original for a specific application. Enhancement techniques are problem oriented.

→ Enhancement is highly subjective, based on human preferences

③ **Image Restoration** : → deals with improving the appearance of an image.

→ Image restoration is objective, as they tend to be based on mathematical or probabalistic models of image degradation.

(4) Color Image Processing : involves color models and processing of colors in a digital domain

(5) Wavelets - used to represent images in various levels of resolution

(6) Compression - deals with techniques for reducing the storage required to save an image, or the bandwidth required to transmit it.

→ compression is enforced in the form of image file extensions, like the jpg file extension used in the JPEG (Joint Photographic Experts Group) image compression standard.

(7) Morphological Processing : deals with tools for extracting image components that are useful in the representation & description of shape.

erosion = opening up = thinning → used when 2 objects are clumped together

dilation = to thicken edges | boundaries of objects

(8) Segmentation : partitions an image into its constituent parts or objects, separates the foreground and background.

(9) Feature extraction : follows the output of a segmentation stage

→ Feature extraction consists of feature detection and description

→ Feature description assigns quantitative attributes to the detected features, eg corners & their orientation & location in the image.

(10) Image Pattern Classification → the process of assigning a label to an object based on its feature descriptors.

→ can use classical approaches like minimum distance, correlation & Bayes classifier

→ can also use deep neural networks / deep CNNs.

(11) Knowledge Base → knowledge about a problem domain is coded into an image processing system in the form of a knowledge base.

→ knowledge may be in the form of detailing regions of an image where the information of interest is located, limiting the search

→ can also be complex, with high-resolution images in connection w/ change-detection applications (satellite images)

* Image Sensing and Acquisition

* Brightness Adaptation and Discrimination

→ Digital images are displayed as a discrete set of intensities, and the eye's ability to discriminate these variations matter.

→ The range of intensity levels for which the human eye can adapt is around 10^{10} from scotopic to glare limit

<u>Brightness Adaptation Level</u> : The sensitivity level for a given lighting condition is called as the brightness adaptation level.

→ The response of the visual system can be characterized w.r.t a particular brightness adaptation level.

<u>Brightness Discrimination</u> : The ability of the eye to discriminate between changes in light intensity at any specific adaptation level.

→ If the background illumination I is less, the ability of the eye to discriminate a small change in intencity is less.

→ There must be a large increment in intensity ΔI for the eyes to be able to discern an object if ambiance illumination I is low.

→ The ratio $\frac{\Delta I}{I}$ = Weber ratio

Small weber ratio ⇒ better discrimination

Large weber ratio ⇒ poor discrimination

→ leads to phenomena like the Mach band effect and simultaneous contrast.

* ## Mach band effect

→ exaggerates contrast between edges of slightly different shades of gray

→ As soon as they contact one another, they trigger edge detection in the human visual system.

→ Although the intensity ~~of~~ the stripes is constant, we perceive a brightness pattern that is strong near the boundaries.

→ The visual system tends to undershoot or overshoot around the boundaries ~~of~~ regions of different intensities.

* ## Simultaneous Contrast

→ A region's perceived brightness does not simply depend on its intensity

→ They appear darker to the eye as the background gets lighter.

* ## Sampling and Quantization

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

→ A frame grabber or digitizer is used to sample and quantize the analogue video signal.

→ The continuous data must be converted into digital form. There are two steps in image digitization:

(1) Sampling - related to coordinate values - digitizing the coordinate values.

<u>Quantization</u> : related to intensity values - digitizing the amplitude values. The transition between the continuous values of the image function and its digital equivalent is called quantization.

→ The number of quantization levels should be high enough for human perception of fine shading in the image.

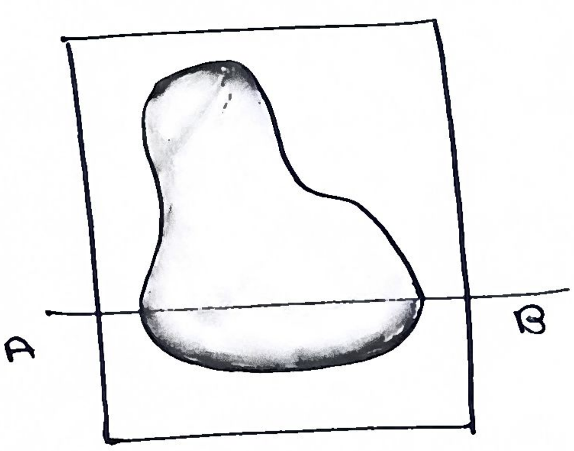
→ If quantization is done with insufficient brightness levels, then there may be false contours in the image.

<u>Note</u>: Sampling → determines spatial resolution of the digitized image

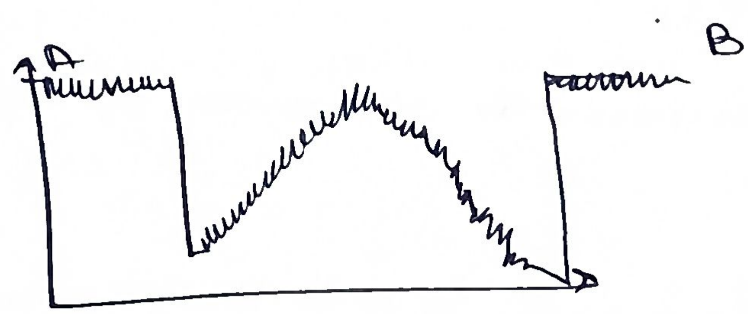
Quantization → determines the no. of gray levels in the digitized image.

<u>Example</u>

Input Image

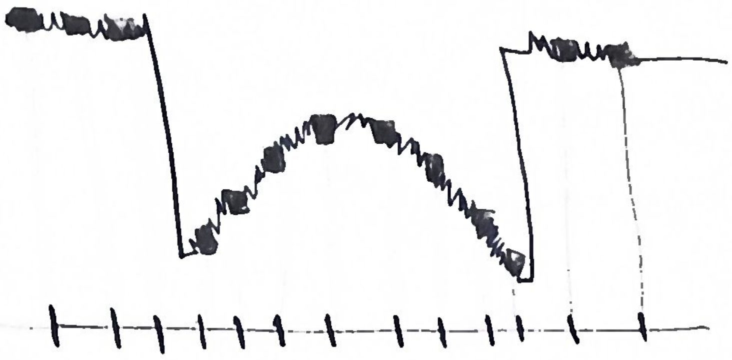


<u>Step1</u> :→ Draw a plot of amplitude (intensity) values of the continuous image along the line segment AB.



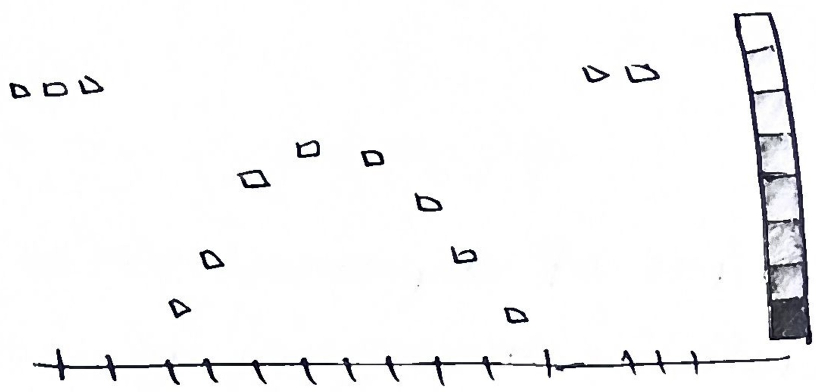
→ The random variations are due to image noise.

<u>Step2</u>: Take equally spaced samples along AB. The discrete spatial values are indicated by the tick marks at the bottom.



Step3: The values of the sample still vertically span a continuous range of intensity values. The intensity values must be quantized into discrete values.

→ The vertical gray bar depicts the intensity scale divided into each of the eight intensity intervals. The continuous intensity levels are quantized by assigning each of the eight values to the sample points.



* Resizing Images

→ Zooming increases the dimensions of an image, and creates new pixel locations.

→ New gray-level values have to be assigned to these locations.

→ This is done using interpolation techniques.

* Representing Digital Images

→ Let $f(s,t)$ represent a continuous image function of 2 continuous variable s and t.

→ Suppose that the continuous image is sampled into a digital image $f(x,y)$ containing M rows and N columns, where (x,y) are discrete coordinates

→ Then x & y range from $x = 0,1,2 \dots M-1$ and $y = 0,1,2 \dots N-1$

→ M×N is represented as

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

* Image Interpolation

→ Interpolation is the process of using known data to estimate values at unknown locations.

→ Interpolation is used in tasks such as zooming, shrinking, rotating and geometrically correcting digital images.

Types

(i) Nearest Neighbor Interpolation: Assign each new location the intensity of its nearest neighbor in the original image

→ Easy to implement, but has the tendency to produce undesirable artifacts, such as severe distortion of straight edges.

(ii) <u>Bilinear Interpolation</u> : The four nearest neighbors are used to estimate the intensity at a given location.

→ Let (x,y) denote the coordinates of the location to which one wishes to assign an intensity value.

→ Let $v(x,y)$ denote the intensity value at that point.

→ For bilinear interpolation, the assigned value is obtained using the equation :

$$v(x,y) = ax + by + cxy + d$$

a, b, c, d are the 4 nearest neighbors of the point (x,y).

→ Bilinear interpolation gives much better results than nearest neighbor interpolation, with a modest increase in computational burden.

(iii) <u>Bicubic Interpolation</u>

→ Involves the 16 nearest neighbors of a point

→ The intensity value assigned to (x,y) is obtained using the equation

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

→ The 16 coefficients are determined from the 16 equations w/ 16 unknowns that are written w/ the 16 nearest neighbors of the point (x,y)

→ Bicubic interpolation does a better job preserving fine details than bilinear interpolation

→ It is the standard used in commercial image editing applications like Adobe Photoshop

* Relationships between pixels

① Neighbors of a pixel

4-neighbors : Any pixel $p(x,y)$ has 2 vertical and 2 horizontal neighbors, given by:

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

→ This set of pixels is called the 4-neighbors of P and is denoted by $N_4(P)$.

→ Each of them are at a unit distance from P.

→ The 4 diagonal neighbors have coordinates
$(x+1, y+1)$, $(x+1, y-1)$, $(x-1, y+1)$, $(x-1, y-1)$ denoted by $N_D(P)$.

→ These neighbors, along with the 4-neighbors, are called the 8-neighbors of p, denoted by $N_8(p)$.

Neighborhood : The set of image locations of the neighbors of a point p is called the neighborhood of p.

→ The neighborhood is said to be closed if it contains p. Otherwise, the neighborhood is said to be open.

② Adjacency

→ relationship between pixels in an image based on their spatial positions

→ Let V be the set of intensity values used to define adjacency.

There are 3 types of adjacency:

(i) <u>4-adjacency</u> : Two pixels p and q with values from V are 4-adjacent if q is in the set $N_4(p)$.

(ii) <u>8-adjacency</u> : Two pixels p and q with values from V are 8-adjacent if q is in the set $N_8(p)$.

(iii) <u>m-adjacency</u> : Two pixels p and q with values from V are m-adjacent, if

- q is in $N_4(p)$
- q is in $N_D(p)$ and the set $N_4(p) \cap N_4(q)$ is empty

has no pixels whose values are from V.

<u>Example</u>

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

8-adjacent   m-adjacent

The two pixels should not have a common 4 connected neighbor.

③ Paths

→ A digital path from pixel p with coordinates (x_0, y_0) to pixel q with coordinates (x_n, y_n) is a sequence of distinct pixels with coordinates

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

where points (x_i, y_i) and (x_{i-1}, y_{i-1}) are adjacent for $1 \le i \le n$

→ n is the <u>length</u> of the path

→ If $(x_0, y_0) = (x_n, y_n)$, the path is closed.

* Connected Components

→ If p and q are ~~subsets~~ pixels of an image subset S, then p is connected to q in S, if there is a path from p to q consisting entirely of pixels in S.

→ For every pixel p in S, the set of pixels in S that are connected to p is called a connected component of S.

→ If S has only one connected component, then S is called a connected set

* Regions and Boundaries

→ A subset R of pixels in an image is called a region of the image if R is a connected set.

→ The boundary of the region R is the set of pixels in the first & last rows & columns of the image, that have one or more neighbors that are not in R.

→ All points in R = foreground, the others are in the background.

* Distance Measures

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

(i) $D(p,q) \geq 0$

(ii) $D(p,q) = D(q,p)$

(iii) $D(p,s) \leq D(p,q) + D(q,s)$

Euclidean distance

$$D(p,q) = \left[(x-u)^2 + (y-v)^2\right]^{1/2}$$

City Block distance

$$D_4(p,q) = |x-u| + |y-v|$$

→ In this case, pixels having a D_4 distance from (x,y) that is less than or equal to some value d form a diamond centered at (x,y).

For d=2

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

value = 1 ⇒ 4 no pixels

Chessboard distance

$$D_8(p,q) = \max(|x-u|, |y-v|)$$

→ In this case, the pixels with D_8 distance from (x,y) with some value ≤ 2 form a square centered at (x,y).

for D = 2

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

value = 1 = 8 no pixels

Numericals

① What is the storage requirement for a 1024 x 1024 binary image?
(i) (assume that 1 bit for each pixel)

Ans = 1024 x 1024
= ~~4.27,072~~

(ii) For a grayscale image?

=> 1024 x 1024 x 8 bits

(iii) For a color image?

=> 1024 x 1024 x 8 x 3

② Assume that $F(x,y) = \begin{bmatrix} 0 & 7 \\ 3 & 15 \end{bmatrix}$. Apply bilinear interpolation
(i)

$$\begin{bmatrix} 0 & 0 & 7 & 0 \\ 0 & 0 & 0 & 0 \\ 3 & 0 & 15 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 0 & 3.5 & 7 & 3.5 \\ 0 & 0 & 0 & 0 \\ 3 & 9 & 15 & 7.5 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 0 & 3.5 & 7 & 3.5 \\ 1.5 & 6.25 & 11 & 5.5 \\ 3 & 9 & 15 & 7.5 \\ 1.5 & 4.5 & 7.5 & 3.25 \end{bmatrix} = \begin{bmatrix} 0 & 4 & 7 & 4 \\ 2 & 6 & 11 & 6 \\ 3 & 9 & 15 & 8 \\ 2 & 5 & 8 & 3 \end{bmatrix}$$

(ii) Apply ordered dithering technique to enlarge the image

Use the threshold mask for thresholding

$$T = \begin{bmatrix} 0 & 8 & 2 & 10 \\ 12 & 4 & 14 & 6 \\ 3 & 11 & 1 & 9 \\ 15 & 7 & 13 & 5 \end{bmatrix}$$

Steps : Compare F and T

If $F < T \Rightarrow 0$

$F \geq T = 1$ (0, when $F = T = 0$)

Ans

$$F = \begin{bmatrix} 0 & 0 & 7 & 7 \\ 0 & 0 & 7 & 7 \\ 3 & 3 & 15 & 15 \\ 3 & 3 & 15 & 15 \end{bmatrix}$$

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

* Thresholding - divide by 255 - round to 0 or 1

* Adaptive Thresholding - based on the intensity values in a image, choose the appropriate threshold.

CAT - 1

Image Processing and Analysis

(1) Unit 1 notes

Numericals

1. An image segment is shown below. Let V be the set of gray level values used to define connectivity in the image. Compute the D_4, D_8 and D_m between p & q.

V is defined as (i) $V = \{2, 3\}$ (ii) $V = \{2, 6\}$

(p) 2	3	2	6	1
6	2	3	6	2
5	3	2	3	5
2	4	3	5	2
4	5	2	3	6 (q)

Ans: D_4 and D_8 depend only on the coordinates of the pixels & not on the gray levels. $p(x,y) = (0,0)$ $q(s,t) = (5,5)$

$D_4(p,q) = |x-s| + |y-t| = |0-5| + |0-5|$

$= 10$ units

$D_8(p,q) = \max(|x-s|, |y-t|) = \max(|0-5|, |0-5|)$

$= 5$

Dm: shortest m-path between points. depends on ~~along~~ the pixels along the path & values of neighbors

$V = \{2, 3\}$ $\quad 6 \notin V \quad$ no Dm

$V = \{2, 6\}$

(p) 2	3	2	6	1
6 — 2		3	6 — 2	
5	3	2	3	5
2	4	3	5	2
4	5	2	3	6 (q)

no path - no Dm

2. Find D_4, D_8, D_m for the image segment below

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

$V = \{0, 1\}$

$V = \{1, 2\}$

$p(x, y) = (4, 1) \quad q(s, t) = (1, 4)$

$D_4 = |x-u| + |y-t|$

$= 3 + 3 = 6$

$D_8 = \max(|x-u|, |y-t|)$

$= \max(3, 3) = 3$

D_m: (i) $V = \{0, 1\}$ $\Rightarrow$

3	1	2	1
0	2	0	2
1	2	1	1
1 — 0		1	2

path = 4

(ii) $V = \{1, 2\}$

$$\begin{matrix} 3 & 1 & 2 & 1 \\ 0 & 2 & 0 & 2 \\ 1 & 2 & 1 & 1 \\ 1 & 0 & 1 & 2 \end{matrix}$$

path = 4

3. Consider a 1024 x 1024 image. What is the storage requirement If the image is:

(i) binary $\Rightarrow$ 1 bit for a pixel

$= 1024 \times 1024$

(ii) grayscale $= 1024 \times 1024 \times 8$ (8 bits for a pixel)

(iii) color $= 1024 \times 1024 \times 8 \times 3$

4. Consider the following image

$$F(x,y) = \begin{bmatrix} 0 & 7 \\ 3 & 13 \end{bmatrix}$$

(i) Apply linear interpolation

$$\begin{bmatrix} 0 & 0 & 7 & 7 \\ 0 & 0 & 7 & 7 \\ 3 & 3 & 15 & 15 \\ 3 & 3 & 15 & 13 \end{bmatrix} \qquad \begin{bmatrix} 0 & 0 & 7 & 0 \\ 0 & 0 & 0 & 0 \\ 3 & 0 & 15 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 0 & 3.5 & 7 & 3.5 \\ 0 & 0 & 0 & 0 \\ 3 & 9 & 13 & 7.5 \\ 0 & 0 & 0 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 0 & 3.5 & 7 & 3.5 \\ 1.5 & 6.25 & 11 & 5.5 \\ 3 & 9 & 13 & 7.5 \\ 1.5 & 4.5 & 7.5 & 3.25 \end{bmatrix} \Rightarrow \begin{bmatrix} 0 & 4 & 7 & 0 \\ 2 & 6 & 11 & 6 \\ 3 & 9 & 13 & 8 \\ 2 & 5 & 8 & 3 \end{bmatrix}$$

(ii) Apply ordered dithering with the given mask

$$F = \begin{bmatrix} 0 & 0 & 7 & 7 \\ 0 & 0 & 7 & 7 \\ 3 & 3 & 15 & 15 \\ 3 & 3 & 15 & 15 \end{bmatrix} \qquad T = \begin{bmatrix} 0 & 8 & 2 & 10 \\ 12 & 4 & 14 & 6 \\ 3 & 11 & 1 & 9 \\ 15 & 7 & 13 & 5 \end{bmatrix}$$

~~F = T → 0~~

$F < T = 0$

$F \geq T \Rightarrow 1$

(when $F = T = 0 \Rightarrow 0$)

$$\Rightarrow \begin{bmatrix} 0 & 0 & \cancel{0}\,1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

5. Obtain the digital negative of the following 8-bit sub image.

$$\begin{bmatrix} 139 & 205 & 105 \\ 141 & 252 & 99 \\ 201 & 15 & 76 \end{bmatrix}$$

Ans

8 bits $\Rightarrow L = 256$

$s = L - 1 - r$

$$= \begin{bmatrix} 116 & 50 & 150 \\ 114 & 3 & 156 \\ 54 & 240 & 179 \end{bmatrix}$$

Question Paper Qs

① Calculate the memory requirement of in bytes of a 7 bit RGB color image with a size of 512 × 512

$= 512 \times 512 \times 7 \times 3$

$= 5,505,024$

then div by 8 ?

$\boxed{688,128}$

$\boxed{\text{no. of channels} \times \text{no. of bits} \times \text{image size}}$

② Choose the parameter to determine the spatial resolution of an image.

- sampling

☆☆☆

③ Consider the 2 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) m - adjacent

(c) 8 - adjacent

$\boxed{\text{m-adj}}$

check

(i) if 4-adj (or)

(ii) if D-adj **and** $N_4(p) \cap N_4(q) = \phi$

	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	1	1	1	0	0	1	1	

→ not 4-adjacent

→ 8 - adjacent

For m-adjacency:

(i) not 4 adj ✗

(ii) is D adjacent

$N_4(p) \cap N_4(q) \neq \phi$

but the common value is 0

here $V = \{1\}$

$\Rightarrow$ m-adjacent

(4) Apply the zooming methods - nearest neighbor and bilinear interpolation to zoom the given image to the size 2x. Is the result same in both methods?

69	50	80
45	60	66
30	55	80

Ans Nearest neighbor interpolation

horizontally

69	69	50	50	80	80
45	45	60	60	66	66
30	30	55	55	80	80

vertically

69	69	50	50	80	80
69	69	50	50	80	80
45	45	60	60	66	66
45	45	60	60	66	66
30	30	55	55	80	80
30	30	55	55	80	80

Bilinear Interpolation

$$\begin{bmatrix} 69 & 50 & 80 \\ 45 & 60 & 66 \\ 30 & 55 & 80 \end{bmatrix} \xrightarrow[\text{padding}]{\text{horizontal}} \begin{bmatrix} 69 & 0 & 50 & 0 & 80 \\ 45 & 0 & 60 & 0 & 66 \\ 30 & 0 & 55 & 0 & 80 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 69 & 60 & 50 & 65 & 80 \\ 45 & 53 & 60 & 63 & 66 \\ 30 & 43 & 55 & 68 & 80 \end{bmatrix}$$

vertical padding

$$\begin{bmatrix} 69 & 60 & 50 & 65 & 80 \\ 0 & 0 & 0 & 0 & 0 \\ 45 & 53 & 60 & 63 & 66 \\ 0 & 0 & 0 & 0 & 0 \\ 30 & 43 & 55 & 68 & 80 \end{bmatrix}$$

Find avgs.

(5) CAT I

A grey scale image of dimension 2.5 inch × 2 inch is scanned at the rate of 150 dpi. Calculate the following:

(a) How many bits are req. to represent the image?

(b) How much time is required to transmit the image if the speed is 28 kbps?

$$DPI = \frac{\text{img. resolution}}{\text{size in inches}}$$

img. resolution = 150 × 2.5 × 2 ×

Total bits = 150 × 8 × (2.5 × 300) × (2 × 300)

540,000,000

time = 19285.71