- An image is a 2D function $f(x, y)$ where $x$ and $y$ are spatial co-ordinates and the amplitude $f$ at any point $(x, y)$ is called the 'Intensity' or 'gray level' of the image at that point.

- When $x$, $y$ and amplitude of $f$ are all finite and discrete, then the image is called the 'Digital Image'.

- Processing a digital image using a digital computer is called DIP.

- Analog Image has infinite values of $x$ and $y$.
  $x$ and $y$ being real means there are infinite divisions in the number line. But digital images only allow discrete pixels.

If we use a digital computer to process an Analog Image, huge memory and computations are required.

- We do we need it?
  (i) Improvement of pictorial information
  (ii) Autonomous machine vision
  (iii) Efficient storage and transmission

- Applications:
  Most fields have some use of DIP.

Spectra

γ-ray: nuclear medecine & astronomy
X-ray: medical diagnostics, industry etc.
UV-ray: Lithography, microscopy, lasers
Visible and Infra-red: Remote sensing, security
Microwave: Radar imaging
Radio-wave: MRI

- Human Perception

Improve the quality for better understanding.

(i) Noise filtering

(ii) Contrast Enhancement

(iii) De-blurring

- Images and using them with Machine Learning can help us to predict many things.

- Machine Vision

Interest is on procedures for extraction of image information suitable for computer processing.

(i) Automated inspection
(ii) Biometric security
(iii) Automatic target detection and tracking

- Video and Image compression and transmission

Can be achieved by removing redundancy.

## Representation

$$f(x, y) = r(x, y) * i(x, y)$$

$r \rightarrow$ Reflectance of surface

$i \rightarrow$ Intensity of incident radiation

To convert digitally,

Spatial discretization by grids
Intensity discretization by quantization

Image is now represented as a 2D matrix.
If $(m \times n)$ is the size, thats the resolution.

$$\begin{matrix} f(0,0) & f(0,1) & \cdots & \cdots & f(0,N-1) \\ f(1,0) & f(1,1) & \cdots & \cdots & f(1,N-1) \\ \vdots & \vdots & & & \vdots \\ f(r,0) & f(r,1) & \cdots & \cdots & f(r,N-1) \\ \vdots & \vdots & & & \vdots \\ f(M-1,0) & f(M-1,1) & \cdots & \cdots & f(M-1,N-1) \end{matrix}$$

$0 \leq f(x,y) \leq 255$ if it is 8-bit image

- Steps in DIP :
  - (i) Image Acquisition (done by sensor)
  - (ii) Pre-processing (Enhance the quality by filtering and contrast enhancement)
  - (iii) Segmentation
  - (iv) Feature Extraction (Identifying features and objects)
  - (v) Recognition and Interpretation (understanding the labels)

31/07/2023

Each $f(x_1, x_2)$ is known as 'Image Element' or 'Picture Element' or 'Pixel'.

- Pixels and neighbours

1. 4-Neighbour

In this system, $P(x, y)$ has 4 neighbours which are $(x, y-1)$, $(x, y+1)$, $(x-1, y)$ and $(x+1, y)$.

2. Diagonal Neighbour

In this system, $P(x, y)$ has 4 diagonal neighbours which are $(x-1, y-1)$, $(x-1, y+1)$, $(x+1, y-1)$ and $(x+1, y+1)$.

3. 8-Neighbour

In this system, all 8 previous mentioned neighbours are considered.

- Adjacency

1. 4-Adjacency
2. 8-Adjacency
3. m-Adjacency

Let V be the set of gray-level values used to define adjacency.

For binary image, $V = \{1\}$

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

m-adjacency : Two pixels p and q with values from V are m-adjacent if –

(i) q is in the set $N_4(p)$

(ii) q is in the set $N_D(p)$ and $N_4(p) \cap N_4(q)$ ~~whose~~ no pixel whose values are from V

- Why is m-adjacency required?

⇒ To remove the ambiguity of travelling b/w pixels in case of 8-adjancy.

- $S_1$ and $S_2$ are adjacent if some pixel in $S_1$ is adjacent to some pixel in $S_2$.

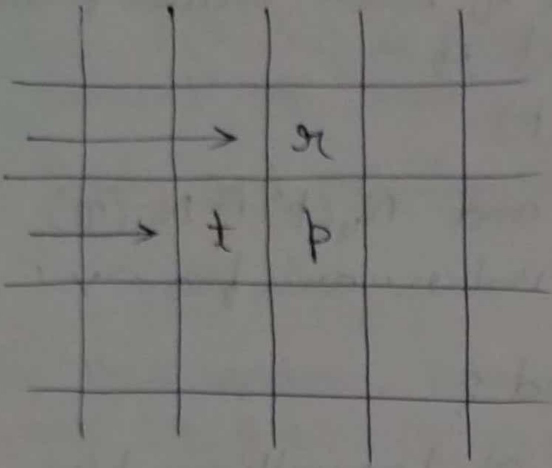
- A path from pixel 'p' with co-ordinate $(x, y)$ to pixel 'q' with co-ordinate $(s, t)$ is a sequence of distict pixels with co-ordinates $(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)$ where $(x, y) \equiv (x_0, y_0)$ and $(s, t) \equiv (x_n, y_n)$ and given that $(x_i, y_i)$ and $(x_{i-1}, y_{i-1})$ are adjcent for all $i$, $1 \leq i \leq n$.

- Connected components

S represents a subset of pixels in an image. Two pixels 'p' and 'q' are said to be connected in S if there exist a digital path between them consisting entirely of pixels in S. For any pixel 'p' in S, the set of pixels that are connected to it is called a 'connected component' of S.

A CAT IS ON THE TABLE

There are 16 connected components in the above text which is part of an image.




$I(p) \rightarrow$ Pixel value at position $p$

$L(p) \rightarrow$ label assignment to pixel location $p$

if $(I(P) == 0)$
    move to next scanning position
else if $(I(P) == 1)$ && $(I(r) = I(t))$ && $(I(r) == 0)$
    assign a new label to $p$
else if $(I(P) == 1)$ and only one neighbour is 1
    assign its label to $p$.
else if $(I(P) == 1)$ && $(I(r) = I(t))$ && $(I(r) == 1)$
    if $L(r) == L(t)$
        $L(P) = L(r)$
    else if $L(r) \,!= L(t)$
        assign one of the label to P
        make a note that $L(r) \eqsim L(p)$

This algorithm gives use how many connected components are there in an image.

- Book

Digital Image Processing
Gonzalez
Woods

stack may overflow if image is very large and we consider 8-adjacency, because this algorithm is recursive. (Made by Sir)
we check a pixel, then check all of its 8-neighbours and so on. So, the stack keeps growing.

- Distance Measure

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

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

and $D(p, q) = 0$ iff $p \equiv q$

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

(c) $D(p, z) \leq D(p, q) + D(q, z)$

-> Euclidean Distance

$$D(p, q) = \sqrt{(x-s)^2 + (y-t)^2}$$

-> City Block Distance ($D_4$)

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

It forms a diamond of equidistant pixels.

| | | | | |
|---|---|---|---|---|
| | | 2 | | |
| | 2 | 1 | 2 | |
| 2 | 1 | 0 | 1 | 2 |
| | 2 | 1 | 2 | |
| | | 2 | | |

→ Chess - Board Distance ($D_8$)

$D_8$ between two pixels $p$ and $q$ is defined as

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

It forms squares of equidistant pixels.

| | | | | |
|---|---|---|---|---|
| 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 |

⊙ Arithmetic and Logical operations

Logical → only on binary images

Arithmetic → Applicable on both gray-level and binary images.

## Image Enhancement in Spatial Domain

Objective : Processed image will be more suitable for a specific application

Image enhancement techniques that are available can be classified broadly into two categories -

(a) Spatial Domain

(b) Frequency Domain

In an image, the variation of intensity gives rise to the frequency.

Here, we use spatial domain techniques i.e. we will directly manipulate pixel values.

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

$f(x,y)$ is the input image

$g(x,y)$ is the processed image

T is an operator on f defined over some neighborhood of $(x,y)$.

- we can write,

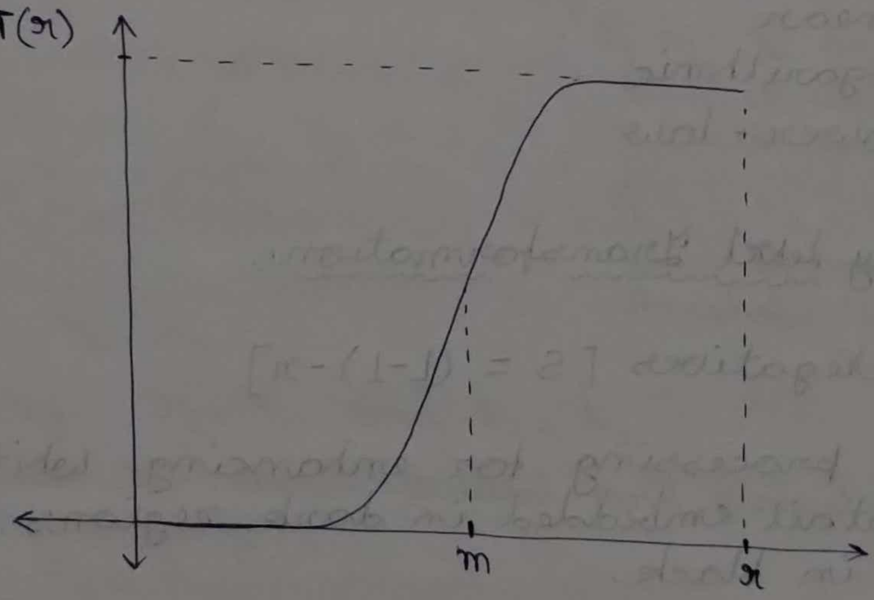
$$s = T(r)$$

where $r = f(x,y)$

$s = g(x,y)$

- Neighbourhood can be 1X1 (pixel itself), 3X3, 5X5 or 7X7.

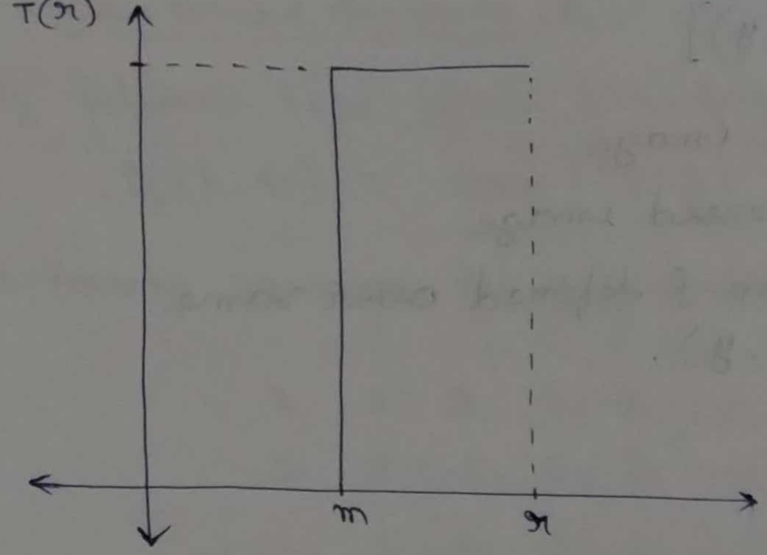
- example :




This function will increase the contrast of the image.

This is not dependent on neighbourhood and is hence called 'Point Processing'.

The enhancement depends only only on the pixel value.




This will convert a gray level image to a binary image.

- Transformations which are dependent on neighborhood, are called 'Mask Processing', 'Kernel Processing' or 'Window Processing'.
- Image enhancement can be done in -
  - (i) Linear
  - (ii) Logarithmic
  - (iii) Power - law

→ Binary Gray Level Transformations

1. Image Negatives $[S = (L-1) - r]$

We need the processing for enhancing white or gray detail embedded, in dark regions dominating in black.

2. Log Transformation

$$[S = c \log(1+r)]$$

This transformation maps a narrow range of low gray-level values input image into a wider range of output values and vice versa.

It also reduces the dynamic range of an image.

3. Power-Law Transformation

$$[s = c\, r^{\gamma}]$$

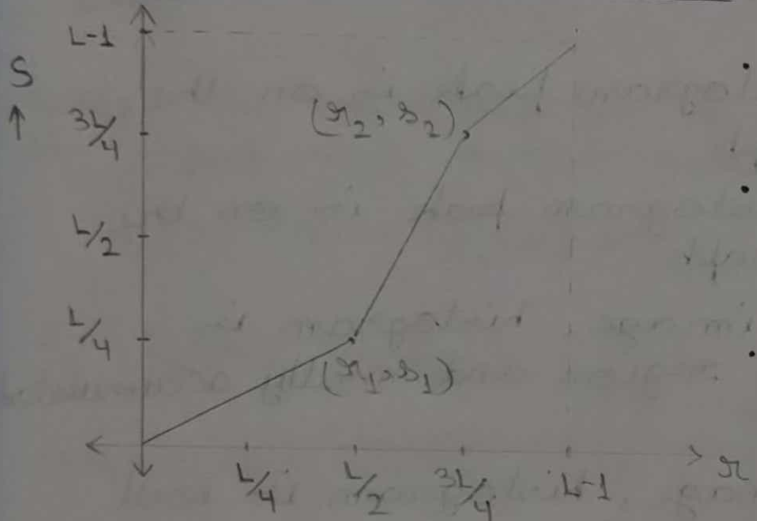
This is used to correct display outputs, print settings etc.
Enhancing water-bodies in ariel image.

07/08/2023

Contrast Stretching : Low contrast image is due to poor illumination, lack of dynamic range sensing device and also due to the wrong aperture of lens.

'Contrast Stretching' is the method of increasing the dynamic range of an image.

- Piece-Wise Linear Function




- Shape of this curve depends on $(r_1, s_1)$ and $(r_2, s_2)$
- If $r_1 = s_1$ and $r_2 = s_2$, then it will be a straight line
- If $r_1 = r_2$, $s_1 = 0$, $s_2 = L-1$, then it will be a Threshold Function

Generally, $r_1 \leq r_2$ and $s_1 \leq s_2$.
Function is single-valued and monotonically increasing.

$r_{min}$ = Minimum graylevel present in the actual image
$r_{max}$ = Maximum graylevel present in the actual image

- Gray-Level Slicing

This is used to enhance a particular range of gray values and suppress others / without disturbing others.
These are the two ways it can be done.

→ Histogram Processing

Histogram of a digital image with a graylevel is the range (0, L-1), is a discrete function $h(r_k) = n_k$, where, $r_k$ is the $k^{th}$ graylevel and $n_k$ is the number of pixels the image having graylevel $r_k$.

In other words, it is an array of size L, where the ith index represent the graylevel i and stores the number of pixels with graylevel i.

$$p(r_k) = \frac{n_k}{n}$$

where, n is the total number of pixels in the image

$n = MN$ (has M rows and N columns)

Histogram manipulation can be used to enhance the image.

- In a dark image, histogram peak is on the lower side of the graph
- In a bright imge, histogram peak is on the higher side of the graph
- In a low-constrast image, histogram is located on a narrow region and mostly accumulated in the centre.
- In a high-contrast image, histogram is well distributed and almost uniform

Our task is to develop a function which covers the entire range of graylevel in the histogram. This can be done by 'Histogram equalization'.

⊙ Histogram Equalization

$r \rightarrow$ Input graylevel (0, L-1)

After normalization : (0, 1)

$S = T(r)$ is the output graylevel

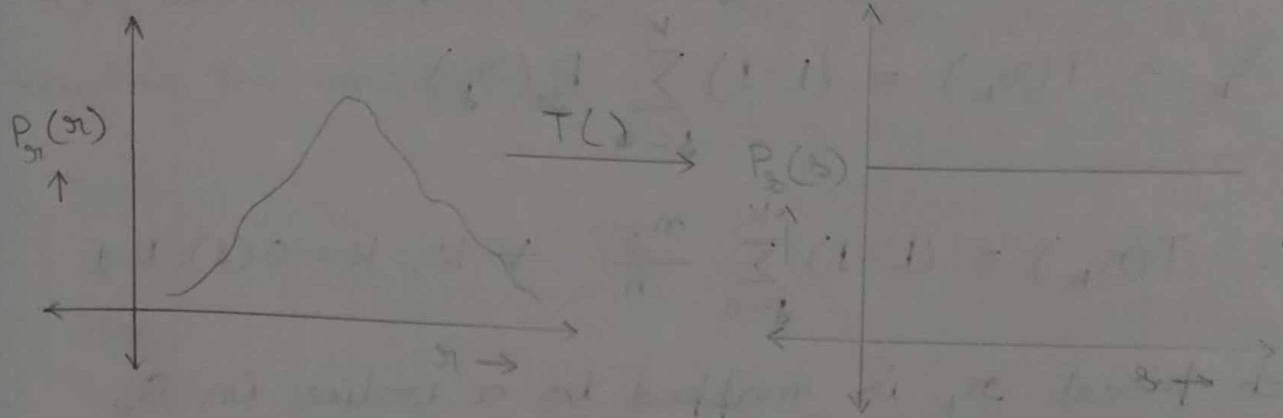
$T(r)$ produces a level 's' for every pixel 'r' in the original image

The transformation function T must satisfy two conditions –

(a) $T(r)$ is a single valued and monotonically increasing function in the domain $[0, L-1]$

(b) T is a mapping from $[0, L-1] \rightarrow [0, L-1]$

If probability density functions $P_r(r)$ and $P_s(s)$ represent the pixel distribution in input and output image respectively, we need a function such that,




Since, the total number of pixels is same in both the input and output images, so,

$$P_s(s)\, ds = P_r(r)\, dr$$

$$\Rightarrow P_s(s) = P_r(r) \left| \frac{dr}{ds} \right| \quad \ldots (1)$$

Now, $$s = T(r) = (L-1) \int_0^r P_r(w)\, dw$$

[cumulative distribution function]

We have to prove that $P_s(s)$ is uniformly distibuted.

$$\frac{ds}{dr} = \frac{L-1}{dr} \left[ \int_0^r P_r(w)\, dw \right]$$

$$= (L-1)\, P_r(r)$$

$$\therefore \frac{dr}{ds} = \frac{1}{(L-1)\, P_r(r)} \quad \ldots (2)$$

From (1) and (2), we have, $P_s(s) = \frac{1}{L-1}$

So, for a continuous curve, $P_s(s)$ covers the whole range, has a uniform distribution.

But, our image is digital,

In this domain, we will use probability instead of pdf and summation instead of integration.

$$P_r(r_k) = \frac{n_k}{n} \; ; \; k = 0(1)L-1$$

The discrete version of the transformation is:

$$S_k = T(r_k) = (L-1)\sum_{j=0}^{k} P_r(r_j)$$

$$\therefore \; T(r_k) = (L-1)\sum_{j=0}^{k} \frac{n_j}{n} \quad \forall k,\; k = 0(1)\,L-1$$

Each pixel $r_k$ is mapped to a value in $S_k$.

09/08/2023

Q. Suppose that the intensity value in an image have the p.d.f.

$$P_r(r) = \begin{cases} \frac{2r}{(L-1)^2}, & 0 \leq r \leq L-1 \\ 0, & \text{otherwise} \end{cases}$$

Find the equalised output.

=> we have,

$$S = T(r) = (L-1)\int_0^r P_r(w)\,dw$$

$$= \frac{2(L-1)}{(L-1)^2}\int_0^r w\,dw$$

$$= \frac{r^2}{L-1}$$

$$\therefore \; \frac{ds}{dr} = \frac{2r}{L-1}$$

$$\therefore P_s(s) = P_r(r) \cdot \frac{dr}{ds}$$

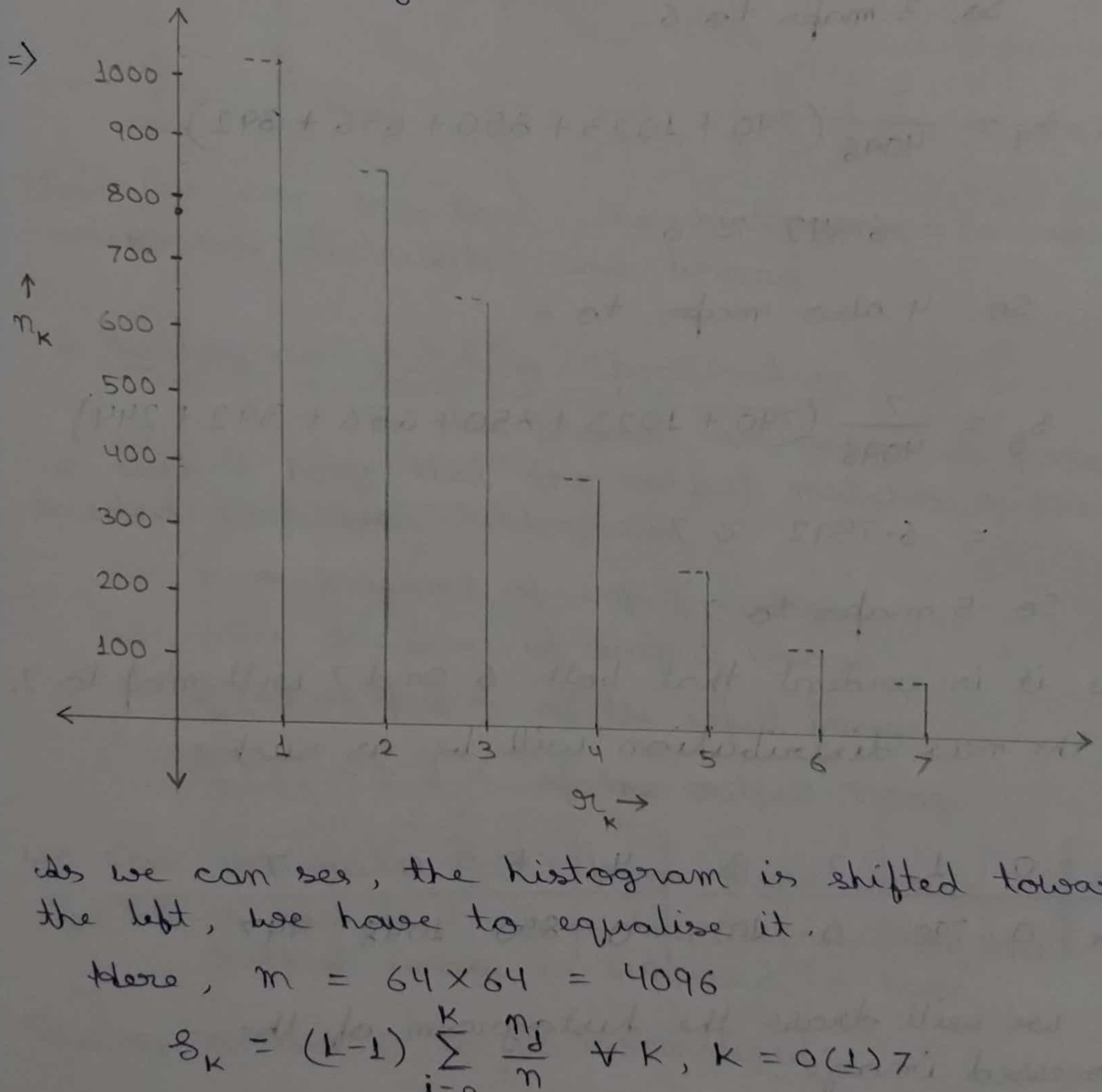
$$= \frac{2r}{(L-1)^2} \cdot \frac{L-1}{2r} = \frac{1}{L-1}$$

Q. Suppose that, a 3-bit image ($L=8$) of size $64 \times 64$ has the intensity distribution

| $r_k$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| $n_k$ | 790 | 1023 | 850 | 656 | 392 | 244 | 122 | 81 |

equalise the image.

=>




As we can see, the histogram is shifted towards the left, we have to equalise it.

Here, $n = 64 \times 64 = 4096$

$$s_k = (L-1) \sum_{j=0}^{k} \frac{n_j}{n} \quad \forall k,\ k = 0(1)7$$

$$\therefore s_0 = 7 \times \frac{790}{4096} = 1.3503 \approx 1$$

So, 0 maps to 1.

$\therefore s_1 = \frac{7}{4096}(790 + 1023) = 3.0982 \approx 3$

So, 1 maps to 3.

$\therefore s_2 = \frac{7}{4096}(790 + 1023 + 850) = 4.551 \approx 5$

So, 2 maps to 5.

$\therefore s_3 = \frac{7}{4096}(790 + 1023 + 850 + 656)$

$= 5.6721 \approx 6$

So, 3 maps to 6.

$\therefore s_4 = \frac{7}{4096}(790 + 1023 + 850 + 656 + 392)$

$= 6.342 \approx 6$

So, 4 also maps to 6.

$\therefore s_5 = \frac{7}{4096}(790 + 1023 + 850 + 656 + 392 + 244)$

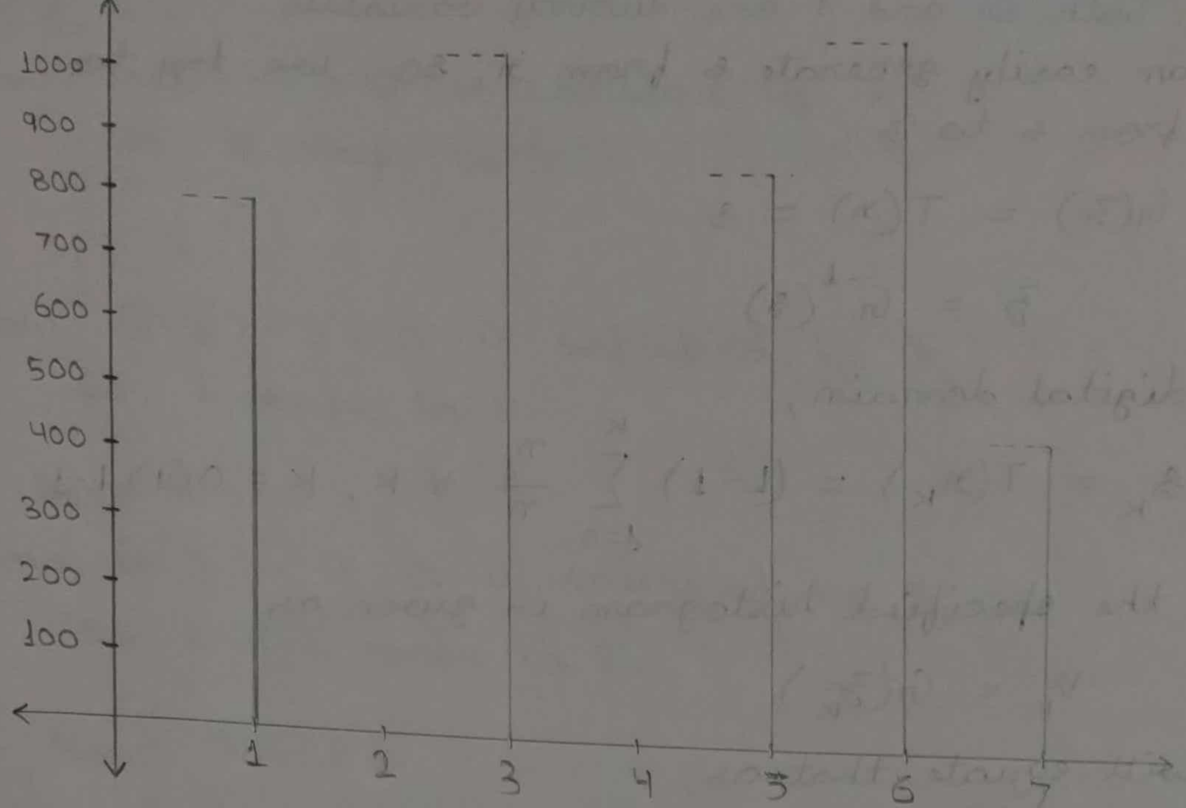
$= 6.7592 \approx 7$

So, 5 maps to 7.

Now, it is evident that both 6 and 7 will map to 7.
So, the new distribution will be as such:

| $s_k$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| $n_k$ | 0 | 790 | 0 | 1023 | 0 | 850 | 1048 | 447 |

Now, we will draw the histogram of the processed image.




Now, we can see that, the histogram is more uniformly distributed and broad.

## ⊙ Histogram Matching / Specification

In this process, an image is processed in such a way that the output matches a pre-decided particular histogram.

$r \rightarrow$ graylevel of input image

$z \rightarrow$ graylevel of output image

$P_r(r) \rightarrow$ p.d.f. of the input image

$P_z(z) \rightarrow$ p.d.f. of the output image

We can estimate $P_r(r)$ from given image.
$P_z(z)$ is specified, it is the distribution the output image is supposed to have

Taking $r$ and $z$ to be normalized,

$$s = T(r) = \int_0^r P_r(w)\, dw$$

We consider another transformation function,

$$G(z) = \int_0^z P_z(t)\, dt$$

Here, both w and t are dummy variables.
We can easily generate s from r, so, we try to go from s to z.

$$G(z) = T(r) = s$$

$$\therefore z = G^{-1}(s)$$

In digital domain,

$$s_K = T(r_K) = (L-1) \sum_{j=0}^{K} \frac{n_j}{n} \quad \forall K, K = 0(1) L-1$$

Say, the specified histogram is given as

$$v_K = G(z_K)$$

We will equate that as,

$$G(z_K) = \sum_{j=0}^{K} (L-1) \cdot P_z(z_j) = s_K \quad \forall K, K = 0(1) L-1$$

Both $v_K$ and $s_K$ are 1D arrays.

$$z_K = G^{-1}(s_K)$$

So, we have to consider the smallest z for which

$$G(\hat{z}) - s_K \geq 0$$

Q. From the previous image, reach to

| $z_K$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| Specified $P_z(z_K)$ | 0.00 | 0.00 | 0.00 | 0.15 | 0.20 | 0.30 | 0.20 | 0.15 |

=> We have,

$$G(z_K) = 7 \sum_{j=0}^{K} P_z(z_j)$$

$\therefore G(z_0) = 0.0$

$G(z_1) = 0.0$

$G(z_2) = 0.0$

$G(z_3) = 1.05$

$G(z_4) = 2.45$

$G(z_5) = 4.55$

$G(z_6) = 5.95$

$G(z_7) = 7$

$\because s_0 = 1$

Now, $G(\hat{z}) - 1 \geq 0$ is satisfied by $z_3$.

So, 0 maps to 3.

$\because s_1 = 3$

Now, $G(\hat{z}) - 3 \geq 0$ is satisfied by $z_5$.

So, 1 maps to 5.

$\because s_2 \approx 5$

Now, $G(\hat{z}) - 4.55 \geq 0$ is satisfied by $z_5$.

So, 2 also maps to 5.

$\because s_3 = 5.6721$

Now, $G(\hat{z}) - 5.6721 \geq 0$ is satisfied by $z_6$.

So, 3 maps to 6.

$\because s_4 = 6.342$

Now, $G(\hat{z}) - 6.342 \geq 0$ is satisfied by $z_7$.

So, 4 maps to 7.

$\because s_5 = 6.7592$

Now, $G(\hat{z}) - 6.7592 \geq 0$ is satisfied by $z_7$.

So, 5 also maps to 7.

We can predict that both 6 and 7 will also map to 7.

So, the new distribution is:

| $z_k$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| $n_k$ | 0 | 0 | 0 | 790 | 0 | 1873 | 656 | 839 |

14/08/2023

→ Mask Processing

Let us consider a 3×3 mask,

| | | |
|---|---|---|
| (-1,-1) | (-1,0) | (-1,1) |
| (0,-1) | (0,0) $W_{ij}$ | (0,1) |
| (1,-1) | (1,0) | (1,1) |

$$R = \sum_{a=-1}^{1} \sum_{b=-1}^{1} W_{ij} \cdot f(a+x, b+y)$$ is the response.

⊙ Smoothing Special Filters

This is used to reduce irrelevant detail, noise, smoothening a false contour.

$$R = \frac{1}{9} \sum_{i=1}^{9} z_i$$ is called 'Averaging Filter'.

Side-effect : This will blur the image. Sharp transitions in the edges will get blurred because of averaging.

To reduce blurring, we use,

$$\frac{1}{16} \begin{vmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{vmatrix}$$

Here, we are giving more prominence to the centre pixel, so blurring will be reduced. This is like weighted average.

Generally,

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

If mask is m×n, then,

$m = 2a+1$ and $n = 2b+1$

## Order - Statistic Filters

This might be a non-linear filter. Here, the response depends on rank of pixels within the mask.

Eg. - Median Filter
we take the median of all pixel values in the mask and then replace the centre one with it.
This reduces the impulse and salt-pepper noise.

## Sharpening Spatial Filters

To highlight the fine detail in an image which has blurred because of some error or during capture, this filter is used.
We will use differentiation to enhance the fine detail.

$$\frac{d}{dx} f(x) = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}$$

In our case (digital image),

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

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

Rules :

When medium is homogenous i.e. there is a constant graylevel, we call it a 'Flat Segment'. When there is controlled linear change in graylevel, we call it 'Step' and when there is a constant slope, we call it a 'Ramp'.

(A) First Order Derivative -

(i) must be zero in the flat segment
(ii) must be non-zero at the onset of graylevel step or ramp
(iii) must be non-zero along ramps

(B) The Second order Derivative -

(i) must be zero in the flat segment

(ii) must be non-zero at the onset and the end of a graylevel step or ramp

(iii) must be zero along the ramp of constant slope

* Step-Transitions and Ramp-Transitions generally represent edges in an image.

First order derivative does not give single pixels when detecting an edge → 'Thick edge'

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

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

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

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

* Isotropic : A transformation is said to be Isotropic if it does not change the output when the the input image is rotated by a certain degree.

$$\begin{vmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{vmatrix}$$ is $90^0$ isotropic.

$$\begin{vmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{vmatrix}$$ is $45^0$ isotropic.

- Our motive is to get a response of 0 in flat segments. So, all features will get vanished. That is why, the sharpened image has to be superimposed on the original one.

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

[If central element is negative]

16/08/2023

## Mathematical Morphology

Morphology is a branch in Biology that deals with size, shape and structure of animals and plants.

Now, Mathematical Morphology is used to extract an image component and describe its features such as boundary, skeleton or convex arc.

Morphological operation consists of two sets –

(i) Bigger set or Image

(ii) Smaller set or Structuring element

→ **Reflection of a set**

$$\hat{B} = \{ w \mid w = -b,\ b \in B \}$$

$\hat{B}$ is the reflection of B.

→ **Translation of a set**

$$(B)_z = \{ w \mid w = b + z,\ b \in B \}$$

where $z(z_1, z_2)$ is a vector.

$(B)_z$ is the translation of B by z.

⊙ **Dilation**

If B is a structuring element and X is an image, then,

Dilation of X by B denoted as '$X \oplus B$' is defined as,

$$X \oplus B = \{ p \mid p = x + b,\ x \in X,\ b \in B \}$$

Alternatively,

$$X \oplus B = \{p \mid (\hat{B})_p \cap X \neq \phi\}$$

<u>Example</u> : Let $X = \{3,2\}$ and $B = \{(-1,0), (2,0)\}$.

By the first definition,

$$X \oplus B = \{(2,2), (5,2)\}$$

By second definition,

$$\hat{B} = \{(1,0), (-2,0)\}$$

Now, $(\hat{B})_p$ will be done and we have to test for every $p$ in $X$.

For $p \equiv (2,2)$

$$(\hat{B})_{(2,2)} = \{(0,2), (3,2)\}$$

Here, $(\hat{B})_{(2,2)} \cap X \neq \phi$

So, $(2,2)$ is accepted.

For $p \equiv (3,4)$

$$(\hat{B})_{(3,4)} = \{(4,4), (1,4)\}$$

Here, $(\hat{B})_{(3,4)} \cap X = \phi$

So, $(3,4)$ is not accepted.

For $p \equiv (5,2)$

$$(\hat{B})_{(5,2)} = \{(3,2), (6,2)\}$$

Here, $(\hat{B})_{(5,2)} \cap X \neq \phi$

So, $(5,2)$ is also accepted.

It is same by both definitions.

- Practical application :

X

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | | | | |
| | 1 | 1 | 1 | 1 | | |
| | 1 | | | 1 | | |
| | 1 | 1 | 1 | 1 | | |
| | 1 | 1 | 1 | 1 | | |
| X | | | | | | |

origin

B

| | | |
|---|---|---|
| 1 | 1 | 1 |
| 1 | 1 | 1 |
| 1 | 1 | 1 |

origin

Given X and B, what is the result of dilation.

=> Here,

$$X = \{(1,1), (1,2), (1,3), (1,4), (2,1), (2,2), (2,4), (3,1), (3,2), (3,4), (4,1), (4,2), (4,3), (4,4)\}$$

$$B = \{(-1,-1), (-1,0), (-1,-1), (0,1), (0,0), (0,-1), (1,1), (1,0), (1,-1)\}$$

$$\therefore\ X \oplus B = \{(i,j) \mid 0 \leq i,j \leq 5\}$$

Advantage - The noise is eliminated

Side-effect - The size of the object is extended

- Properties of 'Dilation' :

(i) It is commutative

$$X \oplus B = B \oplus X$$

(ii) It is associative

$$X \oplus (B \oplus D) = (X \oplus B) \oplus D$$

(iii) $X \oplus B = \bigcup X_b \quad \forall b \in B$

(iv) Translation Invariance

$$X_h \oplus B = (X \oplus B)_h$$

(v) Dilation is an increasing transform
If we had two sets X and Y such that $X \subseteq Y$, then,

$$X \oplus B \subseteq Y \oplus B$$

- Erosion

Erosion, denoted as '$X \ominus B$' is defined as

$$X \ominus B = \{p \mid p + b \in X \ \forall b \in B\}$$

Alternatively,

$$X \ominus B = \{p \mid (B)_p \subseteq X\}$$

Example : If X is like
$$\begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix}$$

and B is $\{(-1,0), (1,0)\}$.

Then $X \ominus B = \begin{vmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{vmatrix}$

- Properties of 'erosion':

(i) $X \ominus B = \bigcap X_{-b} \quad \forall b \in B$

(ii) Translation invariance

$$X_h \ominus B = (X \ominus B)_h$$

$$X \ominus B_h = (X \ominus B)_{-h}$$

(iii) Increasing transform

If $X \subseteq Y$, then

$$X \ominus B \subseteq Y \ominus B$$

(iv) B and D are structuring elements such that $D \subseteq B$, then,

$$X \ominus B \subseteq X \ominus D$$

→ Duality

Erosion and Dilation are duals of each other w.r.t set complement and reflection.

$$(A \ominus B)^c = A^c \oplus \hat{B}$$

$$(A \oplus B)^c = A^c \ominus \hat{B}$$

It is useful when structuring element is symmetric about origin.

Proof : $A \ominus B = \{z \mid (B)_z \subseteq A\}$

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

If $(B)_z$ is contained in A, then $(B)_z \cap A^c = \phi$

$$\therefore (A \ominus B)^c = \{z \mid (B)_z \cap A^c = \phi\}^c$$

The complement of the set of z's that satisfy $(B)_z \cap A^c = \phi$ is the set of z's such that $(B)_z \cap A^c \neq \phi$.

$$\therefore (A \ominus B)^c = \{z \mid (B)_z \cap A^c \neq \phi\}$$

$$\therefore (A \ominus B)^c = A^c \oplus B$$

[Proved]

21/08/2023

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | | | | | | |
| 1 | 1 | 1 | 1 | | | | 1 | 1 | 1 |
| | | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 1 | 1 | 1 | 1 | | | | 1 | 1 | 1 |
| 1 | 1 | 1 | 1 | | | | | | |

There are two objects and left one has Internal noise. Now there is a thin line joining the objects, which is external noise.

We take, $B = \begin{matrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{matrix}$

$A \bullet B = (A \oplus B) \ominus B$

This is the 'closing operation'.

It removes internal noise.

$A \circ B = (A \ominus B) \oplus B$

This is the 'opening operation'

It removes external noise.

- Bengali words are connected components. So, we can find the length of words. The median length of words will be taken as the size of a line-like structuring element.
  The performing opening operation, we can extract the 'শিরোরেখা' of Bengali script.

-> Hit-or-Miss Transform

Denoted as 'A by B'

$A \circledast B = (A \ominus B_1) \cap (A^c \ominus B_2)$

$B_1 \rightarrow$ Set of elements in B associated with object

$B_2 \rightarrow$ Set of elements in B associated with background

This transformation is used for shape detection.

→ Boundary Extraction

The boundary $\beta(A) = A - (A \ominus B)$, where B is a 3×3 structuring element of all 1's.

→ Region - Filling

We take $B = \begin{vmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{vmatrix}$ as structuring element.

Let,

$$\begin{array}{ccccccc}
 & & & 1 & 1 & 1 & \\
 & 1 & & & & 1 & \\
1 & & & & & 1 & \\
 & 1 & & & & 1 & \\
 & 1 & & & & 1 & \\
 & 1 & & & & 1 & \\
 & 1 & & & & 1 & 1 \\
1 & & & & & & 1 \\
1 & 1 & 1 & & 1 & 1 & 1
\end{array}$$

be the object.

We are considering 8- connectivity here.

Let, $p = 1$, we are setting a pixel inside the boundary of the object.

Now, $X_0 = \{P\}$

The region will be filled by successive dilations.

$$X_K = (X_{K-1} \oplus B) \cap A^c$$

until $X_K = X_{K-1}$

23/08/2023

→ Connected Component Extraction

Given a set A, we have y as a connected component in the set A.

Here, we have to,

$$X_K = (X_{K-1} \oplus B) \cap A$$

until $X_K = X_{K-1}$

until $y = X_K$

Here, choose any object point as p.

- A set is said to be convex if all points contained in a straight line b/w two points belonging to the set, is also belonging to that set.

- If S is any arbitrary set,
  H is said to be the 'convex Hull' of S if H is the smallest set that contain S.

H-S is called the 'convex deficiency'.

- We have to take 4 structuring elements here.

$$B_1 = \begin{vmatrix} 1 & X & X \\ 1 & 0 & X \\ 1 & X & X \end{vmatrix}$$

We get $B_i$ by rotating $B_{i-1}$ by $90^\circ$ to the right.
We obtain upto $B_4$.

$$X_K^i = X_{K-1}^i \circledast B^i$$

when $X_K^i = X_{K-1}^i$,

$$D^i = X_K^i$$

$$\therefore \text{Convex hull } C = \bigcup_{i=1}^{4} D^i$$

- The above operation does not give the smallest set. This problem can be manged by not setting pixels which are beyond the boundary.