Digital Image Processing By Gonzalez and Wood

Image Processing & Pattern Recognition (Digital)

For numerical data processing, digital processing computer is more preferred than a human being.

For non-numerical data processing, the birth of the artifical intelligence took place.

visual information which we capture from society and environment.

Image Processing (one end) ←·····················→ Computer vision (other end)

Image: a multidimensional function of spatial coordinates, (x, y) for 2D (photograph)

$f(x, y) \rightarrow$ intensity

(x, y, z) for 3D (CT scan image)

(x, y, t) for movies.

Digital Image: discretized both in spatial coordinates and amplitude levels.

$$I = \{(x, f(x)) : x \in X, f(x) \in F\}$$

$X \rightarrow$ point set

$F \rightarrow$ amplitude set.

$(x, f(x))$ is a pixel

Types of images-

1) Digital
2) Optical (eg. lenses)
3) Electrical-analog (eg. old TVs)
   ↳ x, y and f(x, y) is continuous

Advantages of digital :-

① Precision :-
No quality degradation in digital image. (thousand of copies)

② Flexibility :-
Dozens of variations are possible (brightness, contrast)

a digital image transmission needs high frequency bandwidth.

Analog image → Spatial sampling → Amplitude quantization → Digital image

Objective of Digital Image Processing →

i) improvement of pictorial information for human perception.

ii) processing of image data for storage, transmission and representation for autonomous machine perception.

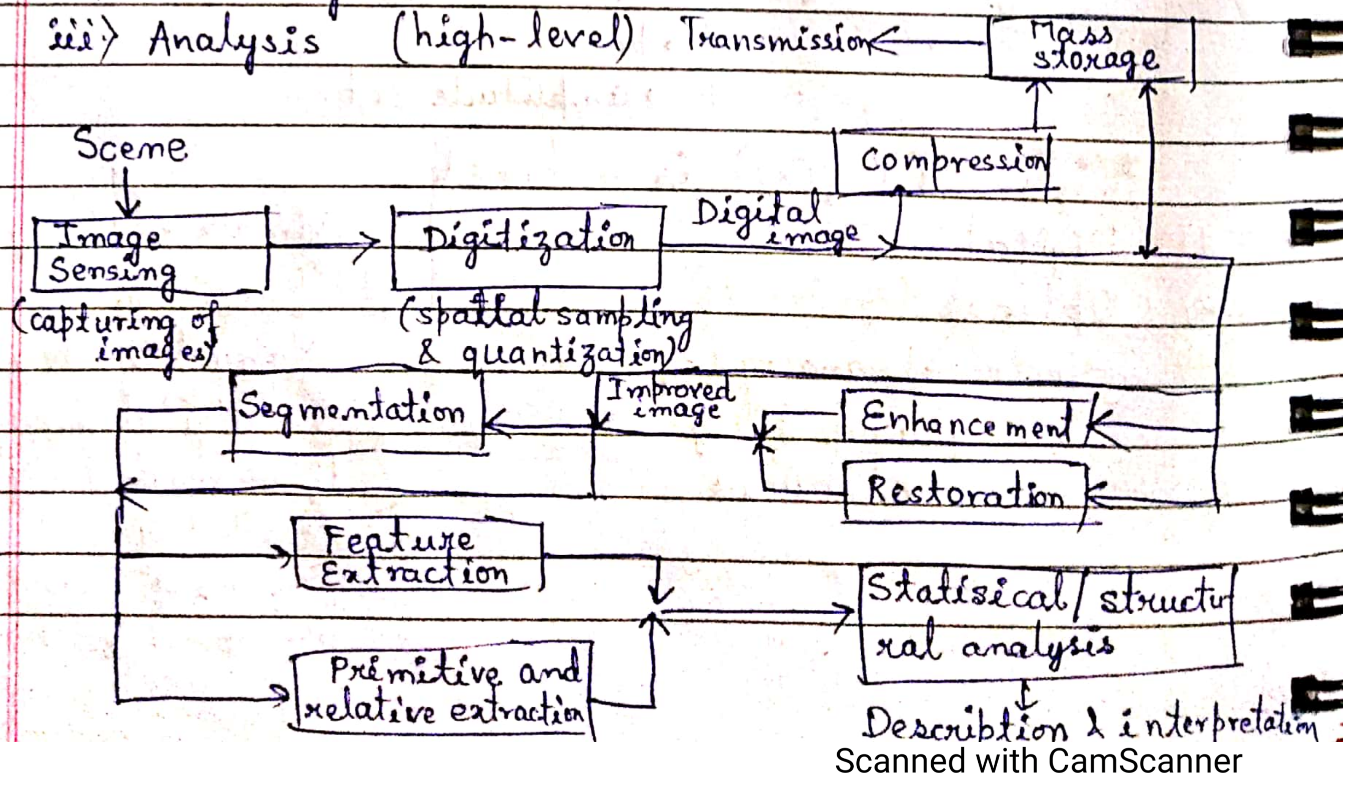
1) Image Acquisition, Discretization, Representation (for efficient storage and transmission)

2) Denoising (improving the visual quality)

3) partition background from objects (segmentation)

4) Translation, Rotation, Scaling.

i) Discretize and representation (low-level)
ii) Processing (mid-level)
iii) Analysis (high-level)



(heuristic approach) Enhancement → ~~don't~~ no need to understand the nature of degradation

Restoration (objective approach) → need to understand the nature of degradation (there is some capturing degradation)

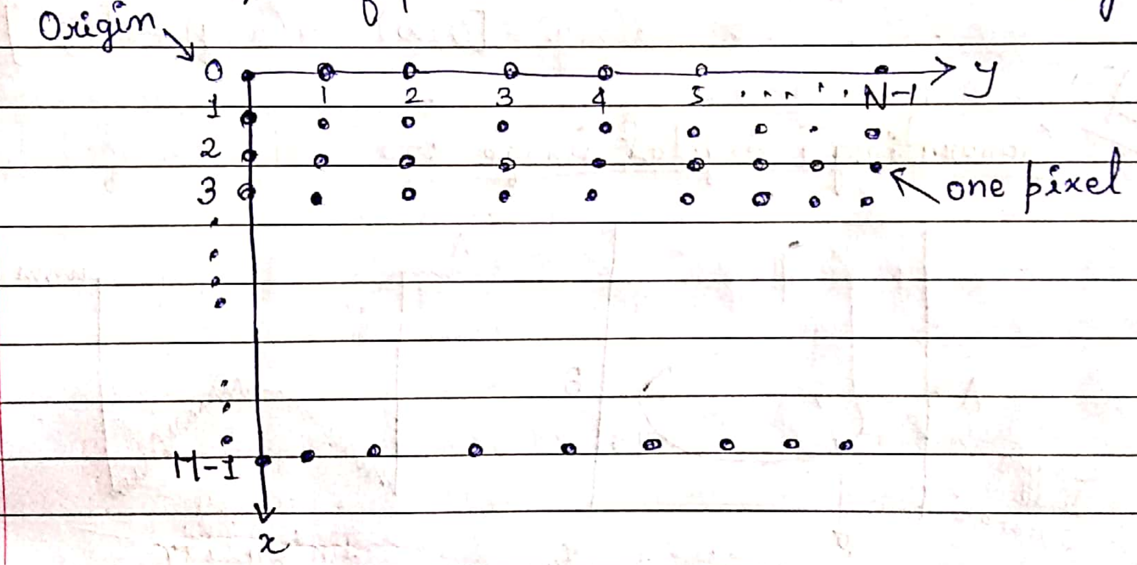
$f(x,y)$ convolutes with $h(x,y)$ to result $g(x,y)$

$f(x,y)$ ↑ original image; $h(x,y)$ ↑ some degradation; $g(x,y)$ ↑ resultant image (degraded)

∴ deconvolution is restoration.

Segmentation → # image is a collection of objects and objects have to be partitioned from background.

object is a set of pixels with certain connectivity.



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

$$A = \begin{bmatrix} a_{00} & a_{01} & \cdots & a_{0,N-1} \\ \vdots & & & \\ a_{M-1,0} & \cdots & \cdots & a_{M-1,N-1} \end{bmatrix} ; \; a_{ij} = f(x=i, y=j)$$

Discretization in x & y → Sampling
Discretization in f → Quantization

Color image or

1) <u>RGB image</u> :-

\# each pixel contains a vector representing red, green and blue components.

2) <u>Index image</u> :-

\# each pixel contains index number pointing to a color in a color table.

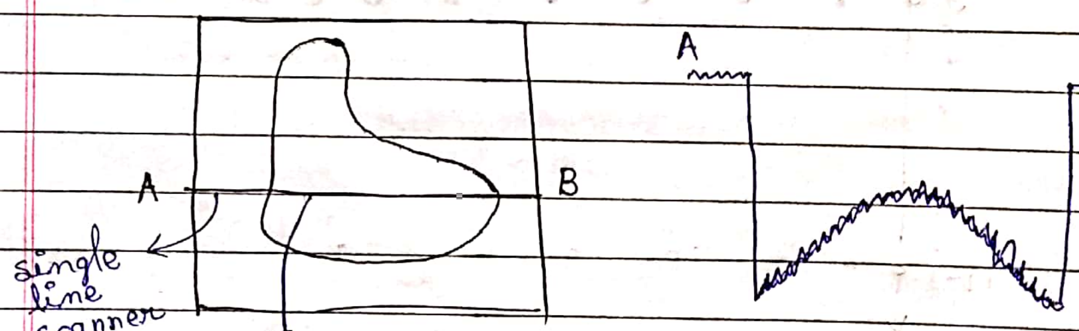
$$f(x,y) = i(x,y) \times r(x,y)$$

(illumination) (reflectance)

$$i(x,y) \in [0,\infty) \qquad r(x,y) \in [0,1]$$

15/02/'21

<u>Generating a digital image</u> :-



change in intensity value as we move across the line

Intensity profile (continuous like function)

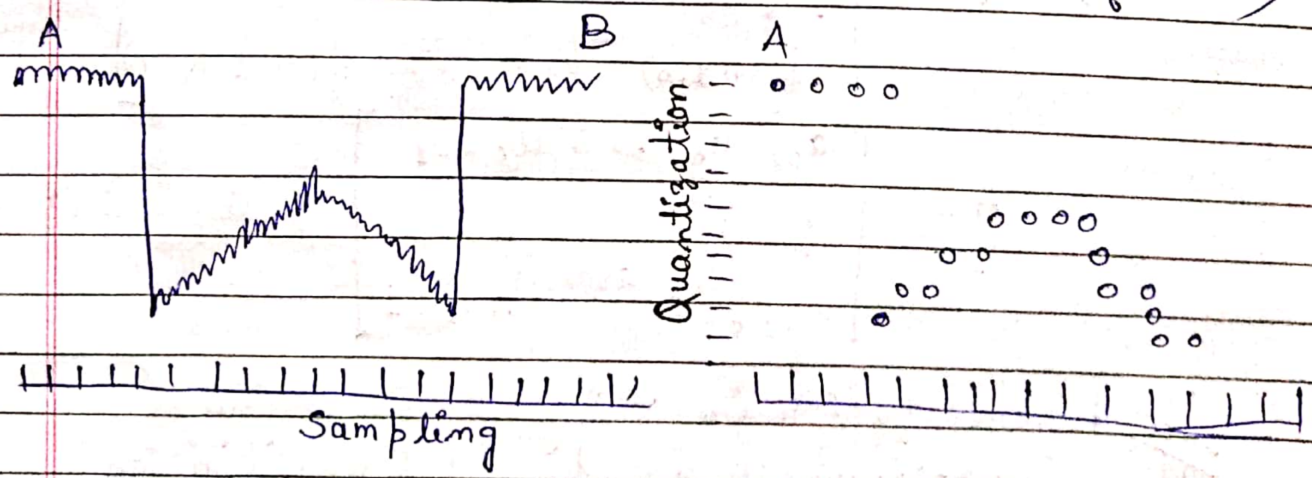
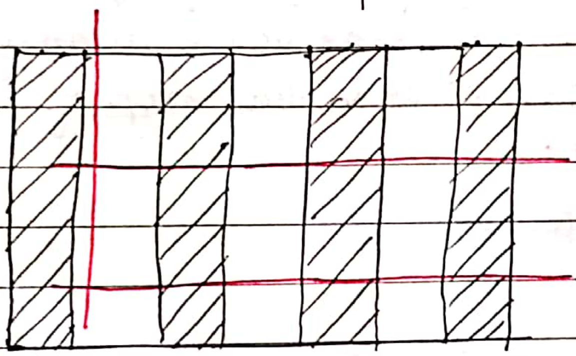


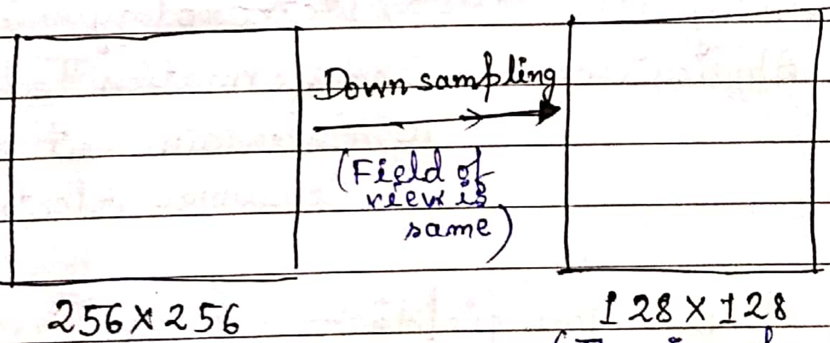
Image sampling: discretize an image in the spatial domain.

Spatial resolution / Image resolution: pixel size or number of pixels.

Given the area or field of view, a high resolution image means no. of pixels are more and they have a high information content.

How to choose the spatial resolution?





(A single pixel carries more information than the next one)

(The size of a pixel is 4 times more than the previous one)

In a case of a 64 X 64 image, a small rectangle has the same intensity value, but in case of 256 X 256 image, it contains 16 diff. pixels having diff. intensity.

Can we increase spatial resolution by interpolation?

In case of interpolation, in b/w two rows, we can add another row which can contain the average intensity of the two rows.

Here, we are increasing the number of pixels but still the visual quality is not enhanced.

We need some intelligent interpolation.

Information content is related with the variation in the data.

Typical operations are:-

1) Remove the blur from an image
2) Smooth out graininess, speckle or noise in an image
3) Improve the contrast or other visual prop. of image prior to displaying it.
4) Magnify, minify.

Applications of Image Processing -

1) Office Automation.
2) Industrial Automation
3) Bio-medical
4) Remote Sensing
5) Scientific Applications.
6) Criminology
7) Astronomy and Space Applications.
8) Meteorology
9) Information Technology
10) Entertainment and Consumer Electronics.

Relation of IP with other fields:-

Input \ Output	Image	Description
Image	IP	IPR, CV
Description	CG	Other Data Proc.

Adv. of IP over - a) Optical and Photographic
b) Electric analog ⓐ Precisions ⓑ Flexibility

Disadv: i) Bandwidth
ii) Speed and expense

08|03|21

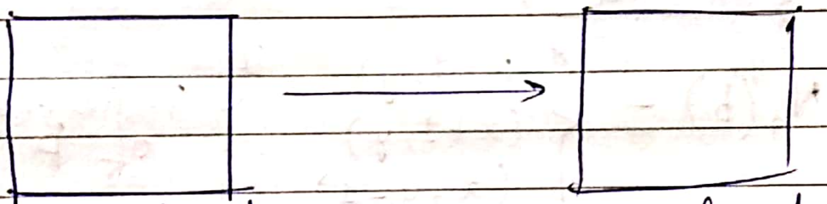
Image quantization:-

COLOR DEPTH/COLOR RESOLUTION

1) No. of colors or gray levels or
2) No. of bits representing each pixel value.
3) No. of colors or gray levels N_c is given by:-

$$N_c = 2^b$$

($b \leftarrow$ no. of bits)

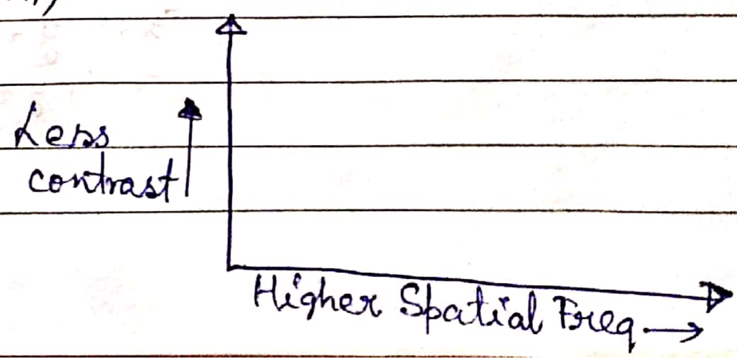


256 levels (For representing each intensity value, we need 8 bits) $\Rightarrow$ 128 levels (. . . . ., 4 bits)

Visual quality of an image depends → (or can be changed)

1. Spatial resolution
2. Changing Color depth.

Low detail image
(need more ~~pixel~~ color depth)



(0,0) $\rightarrow x$, $\downarrow y$

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

Conventional Indexing method

Connected :- Intensity values are similar and they are neighbors (edge, diagonal)

Neighbors of a Pixel :-

$$N_4(p) = \left\{ \begin{array}{l} (x-1, y) \\ (x+1, y) \\ (x, y-1) \\ (x, y+1) \end{array} \right\} \text{4-neighbors of } p$$

$q \in N_4(p)$ implies $p \in N_4(q)$

$$N_8(p) = \left\{ \begin{array}{l} (x-1, y-1) \\ (x, y-1) \\ (x+1, y-1) \\ (x-1, y) \\ (x+1, y) \\ \vdots \end{array} \right.$$

$$N_D(p) = \left\{ \begin{array}{l} (x-1, y-1) \\ (x+1, y-1) \\ (x-1, y+1) \\ (x+1, y+1) \end{array} \right\} \text{Diagonal neighbors of } P.$$

Connectivity:– Two pixels are connected if they are in the same class (same color or same range of intensity) and they are neighbors of one another.

A pixel p is adjacent to pixel q if they are connected.

Path:– path from pixel p at (x,y) to pixel q at (s,t) is a sequence of

Distance:–
For pixel $p, q,$ and z with coordinates (x,y), (s,t) and (u,v), D is a distance function or metric if:
1) $D(p,q) \geq 0$ ($D(p,q) = 0$ if and only if $p = q$)
2) $D(p,q) = D(q,p)$ (Symmetricity)
3) $D(p,z) \leq D(p,q) + D(q,z)$ (Triangle Inequality)

$$D_e(p,q) = \sqrt{(x-s)^2 + (y-t)^2} \quad \{\text{Euclidean Distance}\}$$

$$D_4(p,q) = |x-s| + |y-t| \quad \{\text{City-Block Distance}\}$$

$$D_8(p,q) = \max(|x-s|, |y-t|) \quad \{\text{Chessboard distance}\}$$

Spatial Domain:– (Flexible, easy to manipulate)
Space where all pixels form an image.
(we represent an image by $f(x,y)$)
Images in spatial domain are picture in xy plane where "distance" is meaningful)

Frequency Domain:– (Robust, Computationally expensive)
Using Fourier Transform, word "distance" is lost but the word "Frequency" becomes alive.

12/3/21

Imaging

Image acquisition

the process of sensing our surroundings and then representing the measurements in the form of an image.

different from that of image creation, the process of sensing involves sensor like camara.

Passive Imaging	Active imaging
# makes use of already existing light sources (Ex- Sunlight)	# makes use of artificial lights.
# restricted to day time	# no restriction
	# complicated and expensive
	# takes training to perform.
	# Ex:- NMR, MRI, PET etc.
	# extensively used in medical fields
	# needs control in radiation for safeguard and accurate diagnosis.
	# also used in remote sensing. ↳ SAR (Synthetic Aperture Radar) enables capturing of images day and night

Energy Sources :-

light is familiar and inherently safe
light can be generated reliably

Advantages of Silicon as a sensor →

1) # Reflected lights are linear in nature (visible lights)
2) # other integrated devices can be integrated on a chip easily.
3) # Silicon as a CCD (Charge Coupled Device) and 'Si' being a semiconductor device, all functions can be integrated in a chip.

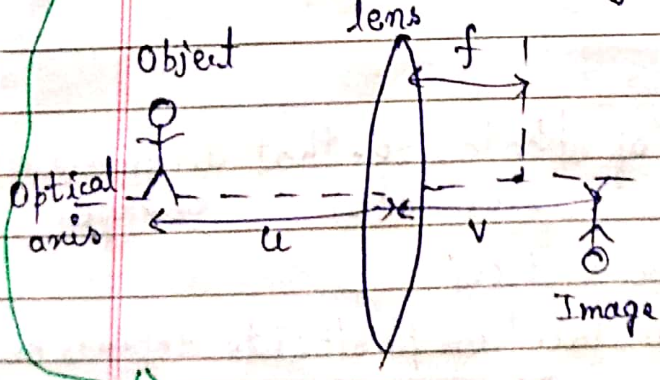
Attributes of a Camera :-

1) Magnification factor
2) Light gathering capacity.

$$m = \frac{\text{image size}}{\text{object size}}.$$

$$\frac{v}{u} = \frac{\text{image size}}{\text{object size}} \qquad m = \frac{v}{u} \qquad \frac{1}{f} = \frac{1}{u} + \frac{1}{v}$$

$$f = \frac{um}{m+1} \qquad m = \frac{\text{image size}}{\text{object size}} = \frac{10}{100} = 0.1$$



Aperture is the light gathering capacity of a lens.
↳ denoted by f number (dimensionless)
↳ obtained by dividing focal length by diameter of the aperture.

Values of aperture are denoted by a sequence of numbers as f 2·8, f 4, f 5·6, f 8, f 11
(incident power becomes half ⟶)

Spherical Aberration
Central and off-central light rays are bought to focus at different distances

Coma
occurs for obliquely incident light when the off-centre rays come to focus on one side of central rays, producing a comet shaped image of point object

Pincushion Distortion
tendency of st. lines to be bowed inwards towards centre of image.

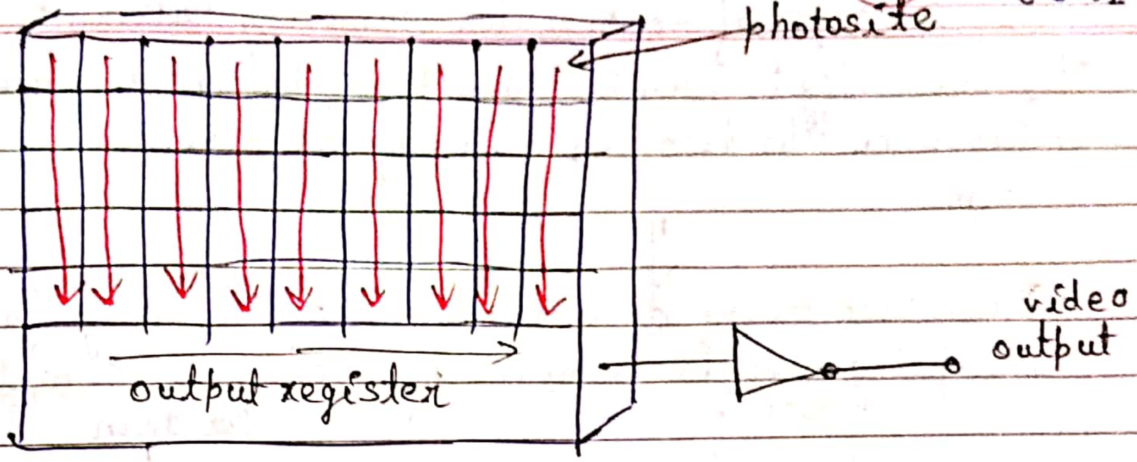
Barrel distortion
tendency of st. lines to be bowed outwards.

* Group of lenses can overcome this problem.
# Aperture needs to be small ⇒ depth of image increases and should have highest resolving power.

CCD Sensor :-
comprises of 2D-array of photosites (elements) that accumulates charge.

Size of individual square $\simeq 1\ cm^2$
Amount of charge accumulated on photosite depends on-
1) amount of light falling on it.
2.) duration of illumination
→ can be max^m = 10^6

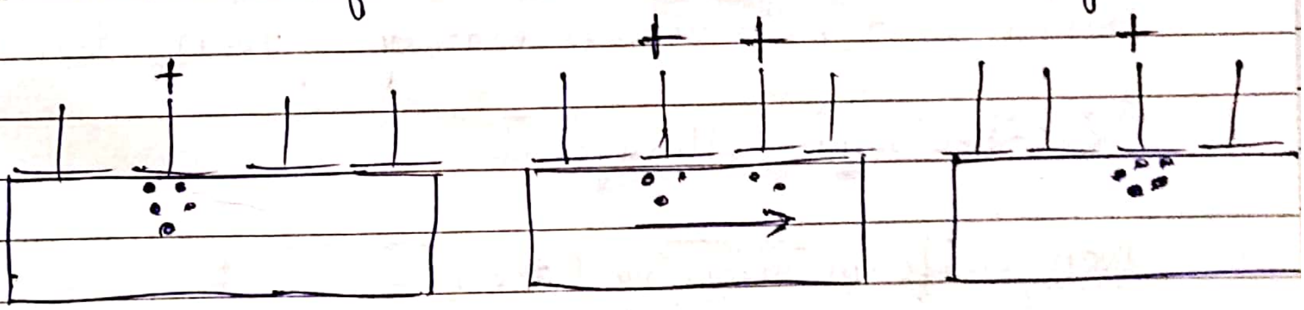
Thermally generated carriers are also there.
(Due to thermal agitation, there will movement of e^-)



Even if there is no light falling, there will still be some $\underbrace{\text{thermally generated carriers}}_{\text{gives rise to dark current.}}$

Electrodes shifts the accumulated charge to the register. There is a timing clock that activates the electrodes.

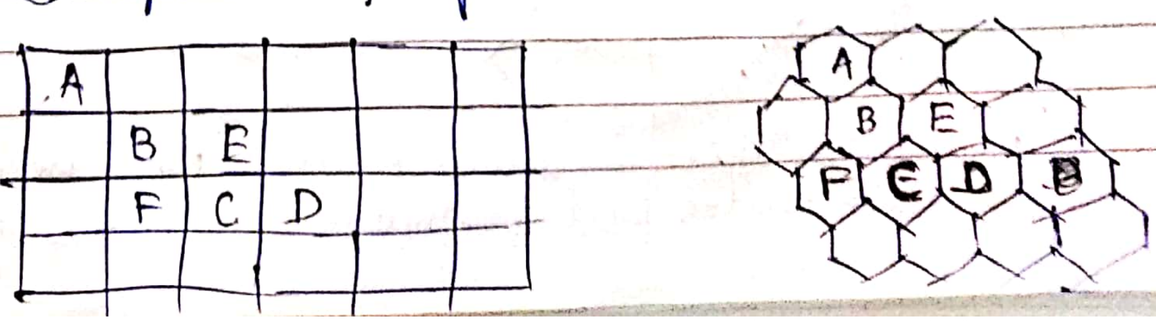
There is no defect unless the electrons are defected.



Sampling Rate and Pattern

When doing the digitization, if we don't maintain the freq. as Nyquist's Sampling rate ⇒ Aliasing.

Rectangular Sampling Pattern



Disadvantages of Rectangular Sampling Pattern :-

1) Connectivity :- Even if E and F are disjoint, they are connected
2) Distance b/w two centre along all directions are same

⇓

These two ~~dista~~ disadvantages of Rectangular sampling ~~rate~~ pattern is improved by Hexagonal sampling pattern.

Energies of signal are mostly distributed /oriented around vertical or horizontal direction in most of the natural images

↳ Reasons for still using Rectangular Sampling Pattern.

15/3/'21

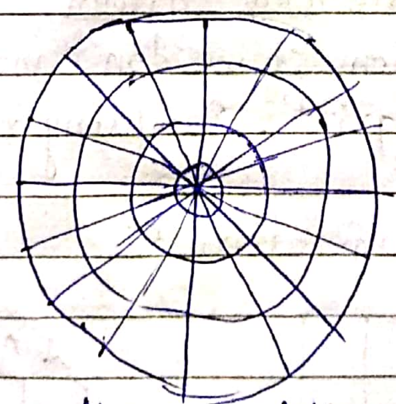
Uniform Sampling Pattern :-

Before perceiving the scene, it is difficult to tell which part has more visual ~~clarity~~ quality than other.

Ex :- Rectangular, Hexagonal.

Non-uniform Sampling Pattern :-

pixel is represented by r and θ.



(Log-polar Sampling)

(we can use this sampling pattern when we know that the centre part has more visual quality)

Rotation invariance :- if we are first given an image and then we perform rotation operations (or vice-versa), then it is rotation invariance if both are same or produce same effects.

This is called Attentive Vision Strategy

Image Enhancement in Spatial Domain

to improve the visual appearance of an image.

Spatial Domain :- Directly manipulating the pixel value.

Frequency Domain :- doing the fourier transformation, applying some manipulation, then doing inverse Fourier transformation.

Operate directly on pixels composing an image

- Pixel gray level info. is used.
- Gray level of pixel $f(x,y)$ is denoted by $g(x,y) = T[f(x,y)]$.
- Usually, a neighborhood about image position (x,y) is considered by using a square or rectangular pixel area centered at (x,y).

Generally, $f(x,y)$ is random. So, it can be represented by a random variable. Similarly, for $g(x,y)$.

To improve the visual quality →

a) Contrast stretching :-

Black to be pushed towards black and white towards white

T can be working as:—

a) Point Processing
b) Neighborhood Processing
c) Ensemble of Images.

When the neighborhood is 1 X 1, then g depends only on the value of f at that (x, y) and T becomes a gray-level transformation (or mapping) function

$s = T(r)$

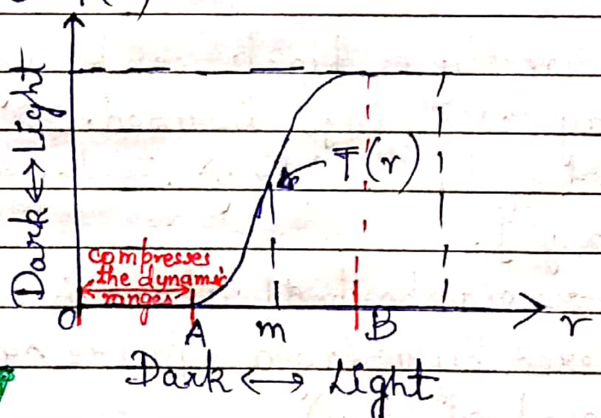
r, s: gray levels of f(x, y) and g(x, y) at (x, y)

Point Processing Techniques (eg → contrast stretching, thresholding, etc.)

Gray Level Transformation:—

$s = T(r)$

output intensity ← ↳ input intensity



Contrast Enhancement.

(the diff. b/w dark and light is ⇒ A–B enhanced).

Contrast Stretching:—

Any value below 'm' is considered as Dark and any value above 'm' is considered as Light.

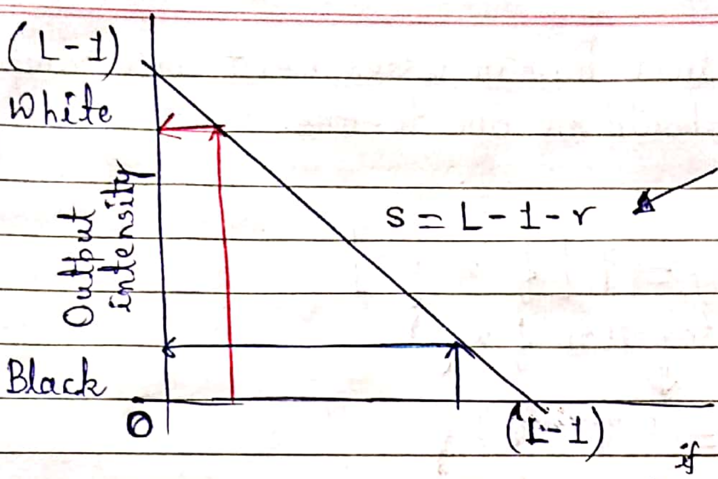


Image negation operation
When $r=0$, $S=L-1$
When $r=L-1$, $S=0$

L = no. of gray levels.

(if a small light region is hugely dominated by a dark region or vice-versa, we can go for negation)

19/3/21

(Image Enhancement)

* Approaches are application specific.
* Some subjective scores are assignment. In case of computers, they calculate objective scores.

PSNR (Peak Signal to Noise Ratio)
(quantitative measure)

* Image quality may not be right because of the wrong image acquisition process.

An excellent image should have higher numerical values than a poor quality image. But this is not always the case in real scenario.

Assessment we do is w.r.t to the poor quality image, not with w.r.t. to the undistored image as we don't have undistorted image.

SSIM (Structural Similarity index measure)
(to assess how good an algorithm is)

* Objective is that human assessment and computer assessment should be one-to-one.

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

↳ better than $f(x\ y)$

⇓

$$s = T(r)$$

s → random variable that determines enhanced image

r → random variable that determines the input image

Point processing :-

Each and every point should be process without considering their neighborhood or connectivity

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

Expansion in dynamic range of output image by the lower part of input range.

Compression in dynamic range of output image by the higher part of input range.

For an inverse log, it is just the opposite.

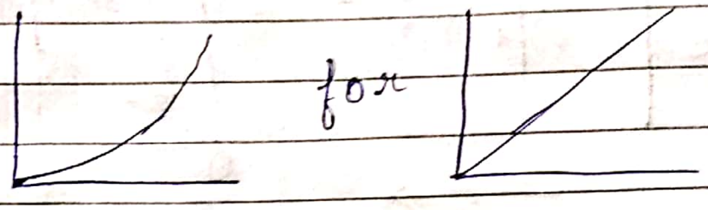
Fourier spectrum scales 0 - 255 to $0 - 10^6$.

But, our computer system can't store 10^6 so it maps it back to 255, so the values 180 - 200 will be mapped ~~to~~ very close to 0.

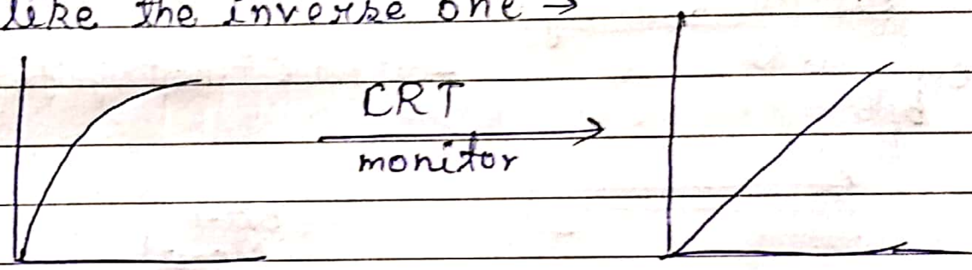
Power Law Transformations :-

$$s = c r^{\gamma}$$

Voltage to intensity mapping for a CRT monitor is :-

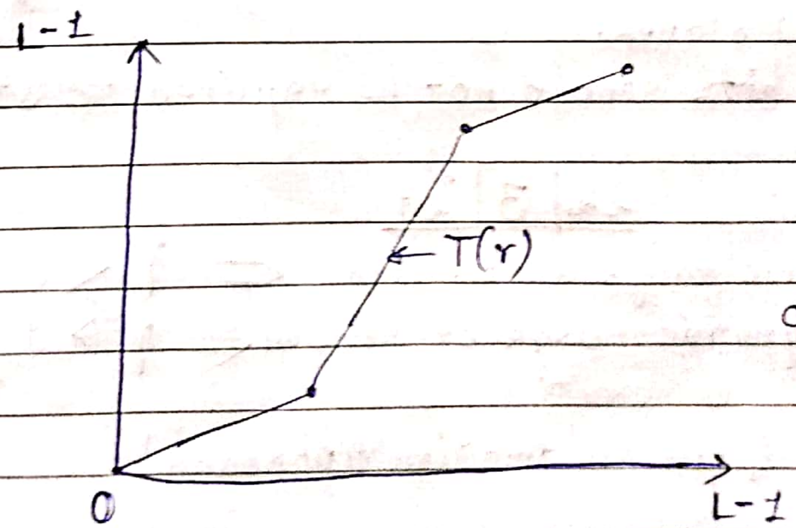


So, to avoid this, we use Gamma correction which first sets the input image to be fed to the CRT monitors like the inverse one →



Contrast Stretching :-

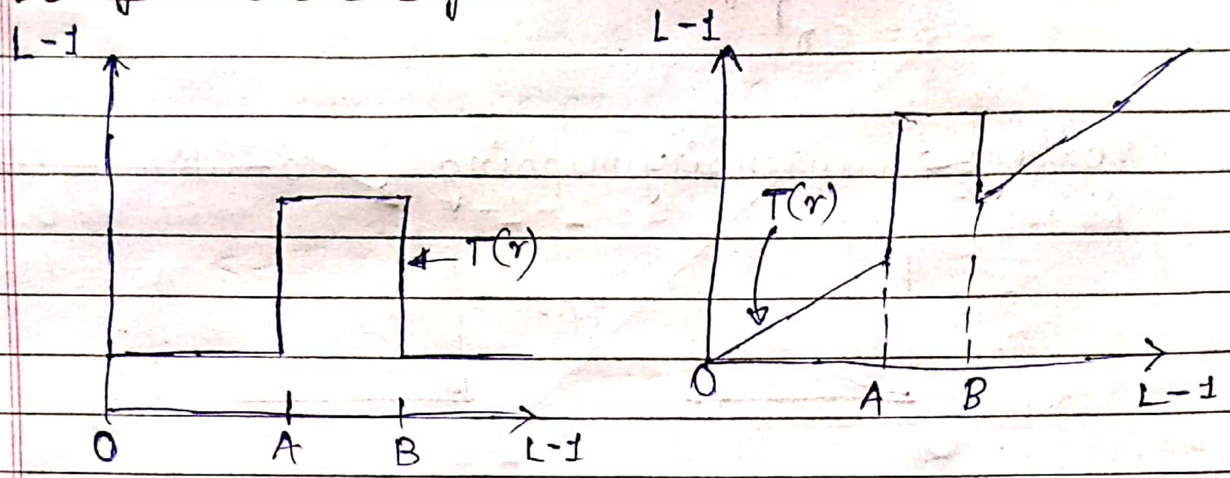
Contrast means the difference between brightest and darkest intensities



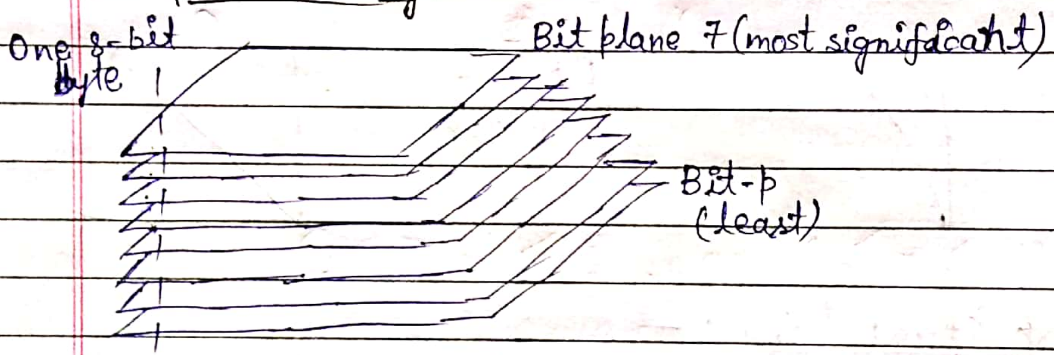
Slope of T(r)

- >1 contrast increases
- $=1$ no change
- <1 contrast decreases

Gray Level Slicing:-



Bit-plane Slicing:-



Relatively different redundancy we can employ for different bit planes.
Lower order bits should not be required to store.

22/3/'21

log ⇐ dynamic range compression ⇐ $\gamma < 1$
inverse log ⇐ dynamic range expression ⇐ $\gamma > 1$

Histogram:- (for automation approach)

→ no. of times a gray value occurs
When it is normalized, it gives as a probability.

also used in
→ Enhancement, Segmentation, Compression

$$h(r_k) = n_k$$

Incase of a dark image, histogram is concentrated around the left side and in case of a bright image, it is concentrated around the right side

∴ We need an o/p image whose histogram will be stretched towards the whole dynamic range.

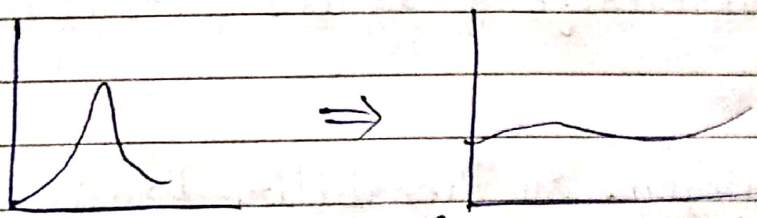
In case of a low contrast image, histogram is located in the middle of intensity value and in case of a high contrast image, it is distributed towards the whole dynamic range.

Requirements:-

1) histogram to be stretched across the dynamic range.
2) no. of intensity values should be also equal.
(as close to an uniform distribution)

Approaches :-

1) Histogram equalization :-



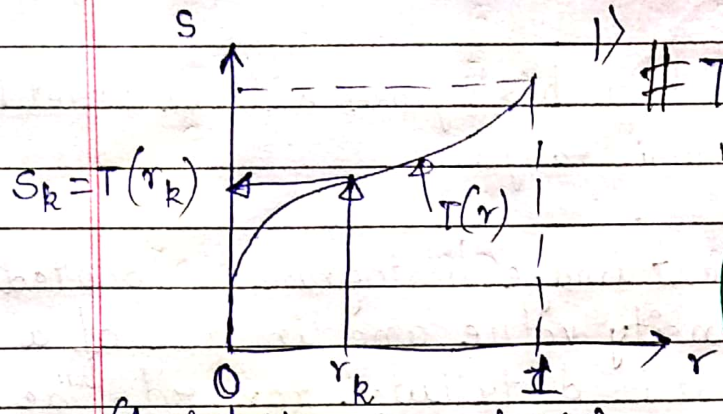
(output should have a large dynamic range)

2) Histogram matching:-
to make the histogram as the desire of the customer

HISTOGRAM EQUALIZATION :-

Assumptions
- # input image is a continuous/analog image
- # normalized dynamic range of [0, 1) — 0: dark, 1: bright

$$s = T(r)$$



(An intuitive form of T(r))

1) # T is single-valued and monotonically increasing.

inversion is possible for each i/p value

(the dark is dark in O/P image and the bright is bright in O/P Image (relative gradation of dark and bright image is preserved))

2) $0 \leq T(r) \leq 1$ for $0 \leq r \leq 1$

(normalized range of the O/P is preserved after applying the operation)

All intensity should have same probability.

Best way to characterize 'r' is its probability density function.

Histogram is analogous to Probability Density Function (PDF) which represent density of population.

Let s and r be Random variables with PDF $p_s(s)$ and $p_r(r)$ respectively and relation b/w s and r is :-

$$s = T(r)$$

$$\boxed{p_s(s) = p_r(r)\left|\frac{dr}{ds}\right|}$$

If we have $p_r(r)$, we can go for Cumulative Density Function. We know that PDF is always +ve and if we go on adding it cumulatively (integration), it will be monotonically increasing.

$$\text{Let } s = T(r) = \underbrace{\int_0^r p_r(w)\,dw}_{\text{cumulative Density function (CDF)}}$$

$$p_s(s) = p_r(r)\left|\frac{dr}{ds}\right|$$

$$= p_r(r)\cdot\left|\frac{1}{\frac{ds}{dr}}\right| = p_r(r)\cdot\left|\frac{1}{\frac{d\left(\int_0^r p_r(w)\,dw\right)}{dr}}\right|$$

$$= p_r(r)\cdot\left|\frac{1}{p_r(r)}\right| = 1 \quad \text{(UNIFORM DISTRIBUTION)}$$

LEIBINITZ'S RULE

Differentiating a definite integral w.r.t to its upper limit value is basically the integrand value at the upper limit value

$$0 \le p_s(s) \le 1$$

∴ So, for discrete/digital image:

$$S_k = T(r_k) = \underbrace{\sum_{j=0}^{k} p_r(r_j)}_{CDF} \quad (k \leftarrow k^{th} \text{ gray value}) \quad k = \{0,1,2,3,\dots,L-1\}$$

$$= \sum_{j=0}^{k} \frac{n_j}{N}$$

n_j = no. of pixels with intensity = j

N = No. of total pixels

Intensity	# pixels	Accumulative Sum
0	20	20/100 = 0.2
1	5	(20+5)/100 = 0.25
2	25	(20+5+25)/100 = 0.5
3	10	(20+5+25+10)/100 = 0.6
4	15	(20+5+25+10+15)/100 = 0.75
5	5	(20+5+25+10+15+5)/100 = 0.8
6	10	(20+5+25+10+15+5+10)/100 = 0.9
7	10	(20+5+25+10+15+5+10+10)/100 = 1.0
Total	100	1.0

O/P Value	Quantized output(s)
0.2 X 7 = 1.4	1
0.25 X 7 = 1.75	2
0.5 * 7 = 3.5	3
0.6 * 7 = 4.2	4
0.75 * 7 = 5.25	5
0.8 * 7 = 5.6	6
0.9 X 7 = 6.3	6
1.0 * 7 = 7	7

↳ Maximum intensity value
(can be 255 depending upon question)

26/03/'21

Histogram Equalization:-

Given an image, the processed image will be stretched across whole dynamic range and the probability of ~~da~~ intensity values are the same.

Histogram matching :-

When the customer specifies the histogram of the output image, we can develop CDF and then normalize it and do the histogram equalization to get an histogram equalized image.

Histogram equalization is fully automatic, but histogram matching is user driven, so why does we need histogram matching?

Original image $\xrightarrow{\text{histogram equalization}}$ the image become a low contrast image. (not a visually pleasing image).

From Histogram equalization,

$$s = T(r) = \int^{r} p_r(w)\,dw \qquad \text{We get } p_s(s) = 1$$

We want an output image to have PDF $p_z(z)$

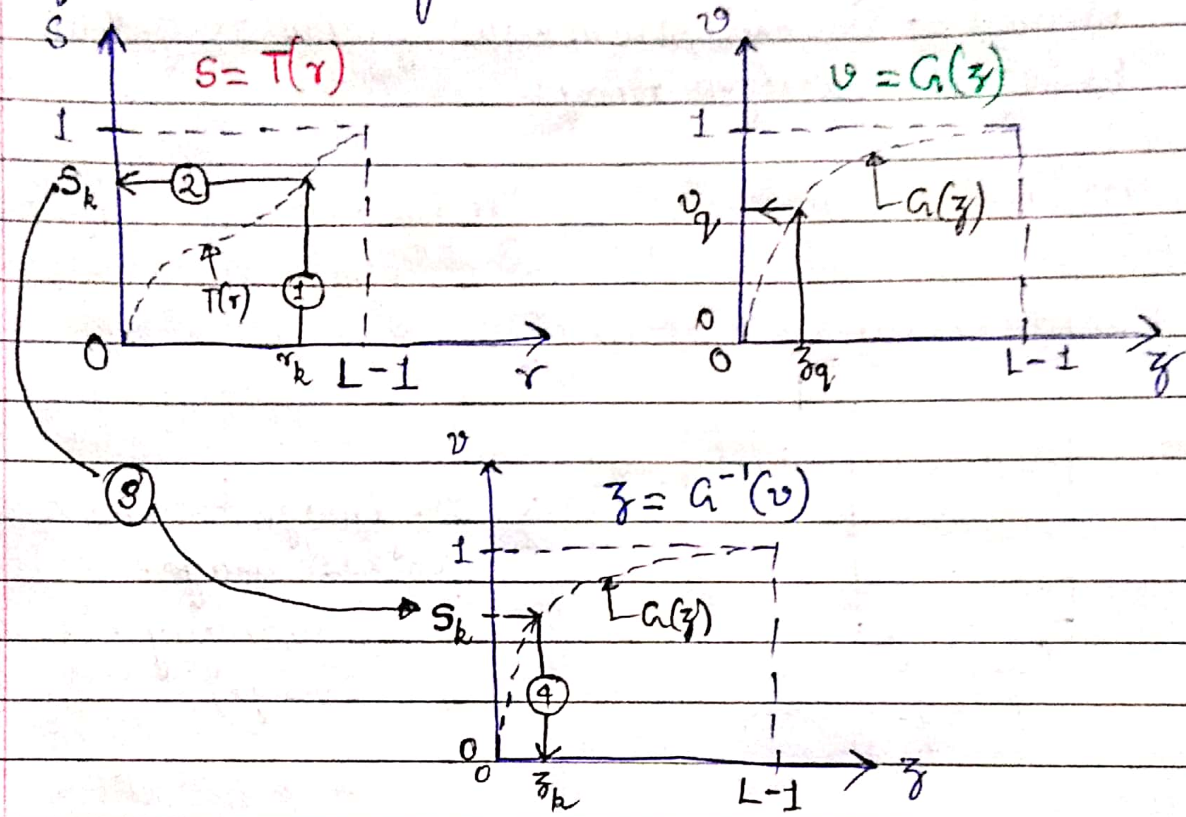
Apply histogram equalization to $p_z(z)$, we get

$$v = G(z) = \int_0^{z} p_z(u)\,du \qquad \text{We get } p_v(v) = 1$$

Since, $p_s(s) = p_v(v) = 1$ therefore, s and v are equivalent.

$$r \xrightarrow{T()} s \xrightarrow{G^{-1}()} z$$

We have been given the input image and the histogram of the output image.



Ex:— Input image histogram — Desired Histogram

Original data →

Intensity (s)	# pixels
0	20
1	5
2	25
3	10
4	15
5	5
6	10
7	10
Total	100

Intensity (z)	# pixels
0	5
1	10
2	15
3	20
4	20
5	15
6	10
7	5
Total	100

1.) Applying Histogram equalization to both tables

r	n_j	$\sum P_r$	s
0	20	0.2	1
1	5	0.25	2
2	25	0.50	3
3	10	0.60	4
4	15	0.75	5
5	5	0.8	6
6	10	0.9	6
7	10	1.0	7

$s_k = T(r_k)$

z	n_j	$\sum P_z$	v
0	5	0.05	0
1	10	0.15	1
2	15	0.30	2
3	20	0.50	4
4	20	0.70	5
5	15	0.85	6
6	10	0.95	7
7	5	1.0	7

$v_k = G(z_k)$

2.) Get a map :

$r \rightarrow s$ $\quad$ $s \rightarrow v$ $\quad$ $v \rightarrow z$

r	s
0	1
1	2
2	3
3	4
4	5
5	6
6	6
7	7

$s_k = T(r_k)$

v	z
0	0
1	1
2	2
4	3
5	4
6	5
7	6
7	7

$z_k = G^{-1}(v_k)$

We get

r	z
0	1
1	2
2	2
3	3
4	4
5	5
6	5
7	6

Actual o/p histogram

z	# pixels
0	0
1	20
2	30
3	10
4	15
5	15
6	10
7	0

Local Enhancement : Local histogram Equalization

Perform histogram equalization on a small neighbourhood.

We can use statistic ~~measure~~ parameters such as Mean, Variance of local area for image enhancement.

$$g(x,y) = \begin{cases} E \cdot f(x,y) & \text{when } \underset{\text{local mean}}{m_{S_{xy}}} \le k_0 \underset{\text{global mean}}{M_G} \text{ and } k_1 D_G \le \underset{\text{local deviation}}{\sigma_{S_{xy}}} \le k_2 \underset{\text{global deviation}}{D_G} \\ f(x,y) & \text{, otherwise} \end{cases}$$

M_G = global ~~deviation~~ mean $\qquad k_2 < 1$

05/04/'21

Local Enhancement : Histogram Statistic for Image Enhancement :-

Global Histogram treats the whole image as an identity
Local Histogram treats the image as composed of several sub-images. It then applies histogram equalization on these sub-images

Low contrast region needs change.

$$g(x,y) = \begin{cases} E \cdot f(x,y) & \text{when } m_{S_{xy}} \le k_0 M_G \text{ and } k_1 D_G \le \sigma_{S_{xy}} \le k_2 D_G \\ f(x,y), & \text{otherwise.} \end{cases}$$

Logic Operations :-

pixel values are converted into bit plane representation

$\underbrace{\text{AND, OR}}_{\text{atleast two images}}$, NOT ↳ a single image

Logical mask of OR/AND is used to take out the region of interest.

Arithmetic Operations :-

Subtraction :-

(a)	Original image	After setting four lower-bit planes to zero	(b)
(c)	Diff. b/w (a) and (b)	Histogram equalized difference image	(d) Error image

Subtraction of an image with an entirely different image ⇒ we have to check which image is totally leading the other image.

Subtraction of two sequence of frames in a video can lead to the tracking of the moving object.

Image Averaging :-
Noise reduction

Degraded image :- $g(x,y) = f(x,y) + \eta(x,y)$

(noise)
(may be White Gaussian ~~image~~ noise)

Image averaging :-

Zero Mean

$$\bar{g}(x,y) = \frac{1}{K}\sum_{i=1}^{K} g_i(x,y)$$

Goal should be to eliminate the effect of $\eta(x,y)$ so that $f(x,y) \approx g(x,y)$

Averaging results in reducing of Noise variance

$$\sigma_{\bar{g}(x,y)} = \frac{1}{\sqrt{K}}\sigma_{\eta(x,y)}$$

We can use subtraction to determine which image is closer to the original image without noise as it will be more black.

06|04|'21

Basics of Spatial Filtering :-

Multiple copies of noise image can lead to transmissi and storage problem

When we are going for cleaning noise, we can loose signal points.

Noise is high frequency content.
As we have a single copy of noise image, pixel to pixel addition is not possible. Since, there is no multiple copies, averaging will be done w.r.t the local neighbourhood.

$$\begin{bmatrix} f(x_1, y_1) & f(x_2, y_2) & \cdots \\ f(x_{n+1}, y_{n+1}) & f(x_{n+2}, y_{n+2}) & \cdots \\ & & \end{bmatrix} \times \underbrace{\begin{bmatrix} w_1 & w_2 & \cdots \\ w_{n+1} & w_{n+2} & \cdots \\ & & \end{bmatrix}}_{\text{Filter}}$$

~~$w_1 f(x_1, y_1) + w_2$~~

Basically, we are doing a convolution operation (window with filter) and then taking an average

3	4	4
9	7	6
3	6	1

$\times$

$\frac{1}{9}$

1	1	1
1	1	1
1	1	1

Mask/ Window/ Template

$$\Rightarrow y = \frac{1}{9}\cdot 3 + \frac{1}{9}\cdot 4 + \frac{1}{9}\cdot 4 + \frac{1}{9}\cdot 9 + \frac{1}{9}\cdot 7 + \frac{1}{9}\cdot 6 + \frac{1}{9}\cdot 3 + \frac{1}{9}\cdot 6 + \frac{1}{9}\cdot 1$$

Imagine that we have a 3 X 3 window that can be placed everywhere on the image.

Step-1:— Move the window to the first location where we want to compute the average value and then select only pixels inside the window.

Step-2:— Compute the average value.

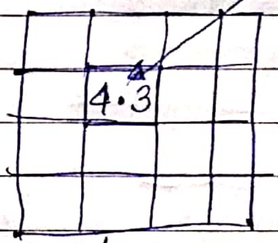
$$y = \sum_{i=1}^{3} \sum_{j=1}^{3} \frac{1}{9} \cdot p(i,j)$$

2	4	1	2	6	2
9	2	3	4	4	4
7	2	9	7	6	7
5	2	3	6	1	5
7	4	2	5	1	2

$\Rightarrow$

2	4	1
9	2	3
7	2	9

$$y = \sum_{i=1}^{3} \sum_{j=1}^{3} \frac{1}{9} \cdot p(i,j)$$

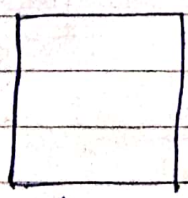


O/P image

Step-3 :- Place the result at the pixel in O/P image.

Step-4 :- Move the window to the next location and go to step 2.

#



A masking window

* 3x3 averaging method is an example of mask operation/ Spatial filtering.

* Mask operation has corresponding mask (sometimes called window or template).

* Mask contains coefficients to be multiplied with pixel values

w(1,1)	w(2,1)	w(3,1)
w(1,2)	w(2,2)	w(3,2)
w(3,1)	w(3,2)	w(3,3)

Mask coefficients

moving average

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

[mask of 3x3 moving average filter has all coefficients $= \frac{1}{9}$]

* Mask operation performed as :-

a) Move the reference point (center) of mask to location to be computed.

b) Compute sum of products b/w mask coefficients and pixels in subimage under the mask.

mask frame

p(1,1)	p(2,1)	p(3,1)
p(1,2)	p(2,2)	p(3,2)
p(1,3)	p(2,3)	p(3,3)

Subimage

reference point of mask

w(1,1)	w(2,1)	w(2,3)
w(1,2)	w(2,2)	w(3,2)
w(1,3)	w(2,3)	w(3,3)

Mask coefficients

$$y = \sum_{i=1}^{N} \sum_{j=1}^{M} w(i,j) \cdot p(i,j)$$

* Spatial filtering →

1) Move the mask over the image at each location

2)

3) Store results at corresponding pixels of o/p image.

4) Move mask to next location and go to step 2

Smoothing Filters:-

1) Mean:- size and shape of window over which mean is computed.

Square window

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

plus shaped window

	1/5	
1/5	1/5	1/5
	1/5	

$$M = m_k m_k^T = \begin{pmatrix} 1/3 \\ 1/3 \\ 1/3 \end{pmatrix} \begin{pmatrix} 1/3 & 1/3 & 1/3 \end{pmatrix}$$

Square mean filter is separable:-

Vertical and horizontal transpose product gives rise to the square mean filter
can be derived by the product of a row and column filter

2) Weighted Mean: A weighted mean is often used in which the weight for a pixel is related to its distance from the center point. The square weighted mean window is also separable.

Square window

1/16	1/8	1/16
1/8	1/4	1/8
1/16	1/8	1/16

plus shaped window

	1/6	
1/6	1/3	1/6
	1/6	

$$w_k = \begin{pmatrix} \frac{1}{4} \\ \frac{1}{2} \\ \frac{1}{4} \end{pmatrix} \begin{pmatrix} 1/4 & 1/2 & 1/4 \end{pmatrix}^T$$

3.) Mode:-
A pixel value is replaced by its most common neighbor.

4.) Median:-
A pixel value is replaced by the median of its neighbours. The median of a set of numbers is the value such that 50% are above and 50% are below. Conceptually, simple but implementation involves sorting operation.

5.) K-nearest neighbour:-
Set v(i,j) to the average of the k pixels in n(i,j) whose values are closest to that of v(i,j). A typical value of k is 6 for (3x3) window

6.) Sigma filter:-
Set v(i,j) equal to the average of all pixels in its neighbourhood whose value is within t counts of the value is within t counts of the value of v(i,j), t is an adjustable parameter. This is called as sigma filter because the parameter t may be derived from sigma or standard deviation of pixel distribution

Calculate the standard deviation of all pixels in the neighbourhood.
Then, include only those points/pixels to calculate average for which the difference is within this standard deviation.

7.) Filters based on pixel values and gradients:-

$$v_{out}(i,j) = \frac{1}{2} v_{in}(i,j) + \frac{1}{2} \sum_{k,l \in n8} w(k,l) v_{in}(k,l)$$

where, $w(k,l)$ is inversely proportional to $|v_{in}(i,j) - v_{in}(k,l)|$

8.) Maximum Homogeneity Filter:-

a) Determine neighborhood of (i,j) most similar to v_{in}

b) Apply some smoothing using only points from this neighborhood.

→ numerical value of pixel needs to be considered regardless of its location.

9.) Filters of the form:

```
If (some condition)
  Apply filter method 1.
else
  Apply filter method 2
End if
```

Ex:-

```
if |v(i,j) − v̄(i,j)| > threshold
   v(i,j) = v̄(i,j)
Else
   Leave v(i,j) unchanged
End if
```

After
(Blurring comes)

12/04/'21

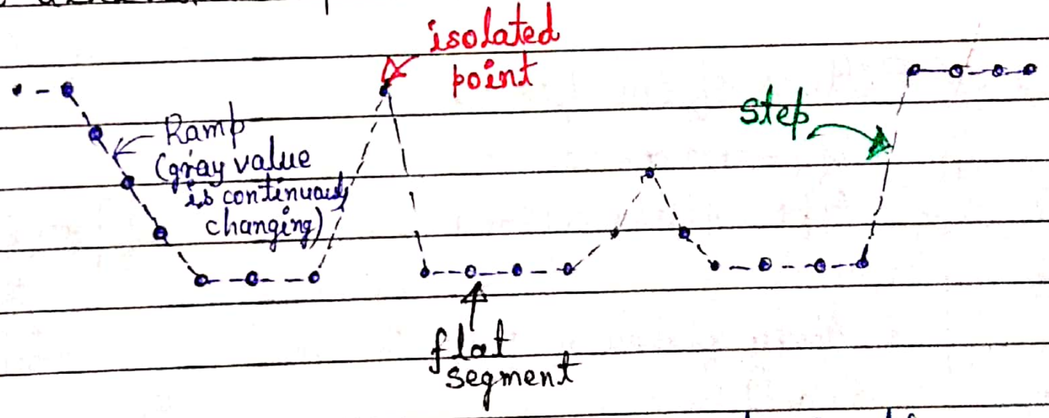
Page No: 110
Date: 12/4/21

Sharpening Spatial Filters :-

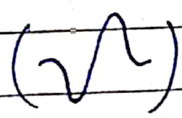
a) Differentiation operation.
b) Includes applications, electronic printing, medical imaging to industrial ~~application~~ inspection and autonomous guidance in military systems.
c) 1^{st} order derivative → where gray values are same in an image
   i) must be zero at flat segments
   ii) must be non-zero at onset of a gray level step or ramp.
   iii) must be non-zero along the ramps.
d) 2^{nd} order derivative → i) must be zero at flag segments
   ii) must be non-zero at the onset and at the end of a ramp.
   iii) must be zero along the ramps.

Averaging :- clean the noise (edge preserving smoothing)
Sharpening :- suppress the noise (enhancing the edge details)

1. To define derivative operator.



* For step like function, second order derivatives have double line behaviour



$$\begin{bmatrix} 1 & -1 \end{bmatrix}_{w_1 \; w_2} \begin{bmatrix} f(x+1) \\ f(x) \end{bmatrix} \qquad \frac{\partial f}{\partial x} = f(x+1) - f(x)$$

$$\begin{array}{ccc} f(x+1) & f(x) & f(x-1) \\ (1 & -2 & 1) \\ w_1 & w_2 & w_3 \end{array} \qquad \frac{\partial^2 f}{\partial x^2} = f(x+1) + f(x-1) - 2f(x)$$

(2) Developing filter coefficients :—

\# Sum of the filter coefficients must be zero.

→ As for a flat segment, first order derivative must be zero. For it to be zero the relation must be like this b/w filter coefficients.

\# 1st order derivative i) thicker edges in an image while 2nd order derivative enhances fine details such as thin lines and isolated points.

\# 1st derivative generally have a stronger response to a gray-level step and while 2nd order derivative produces a double response at step change in gray level.

\# 2nd order derivative (scalar operation) is easy to implement. ↳ Laplacian operator. (rotation invariant)

Isotropic filters (rotation invariant).

$$\nabla = \underbrace{\hat{i} \frac{d}{dx} + \hat{j}}_{\text{unit vectors}} \frac{d}{dy}$$

$$\nabla^2 = \left(\hat{i}\frac{d}{dx} + \hat{j}\frac{d}{dy}\right) \cdot \left(\hat{i}\frac{d}{dx} + \hat{j}\frac{d}{dy}\right)$$

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

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

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

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

Image in 2-D matrix (3X3 neighbourhood)

$$\begin{Bmatrix} f(x-1,y-1) & f(x-1,y) & f(x-1,y+1) \\ f(x,y-1) & f(x,y) & f(x,y+1) \\ f(x+1,y-1) & f(x+1,y) & f(x+1,y+1) \end{Bmatrix}$$

$$\begin{matrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{matrix}$$

if we choose a kernel like this, then convolving this with 3X3 neighbourhood will give the ∇^2.

change in both x and y direction

$$\begin{matrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{matrix}$$

if we also want to take care of both $\frac{\partial^2 f}{\partial x \partial y}$ and $\frac{\partial^2 f}{\partial y \partial x}$ terms, then this filter can be used.

23/04/21

Smoothing operations looses the edge preservation, thereby sharpening is used.

* $\underbrace{\text{First order derivative}}_{\text{vector}}$ enhance the thick edges
* $\underbrace{\text{Second order "}}_{\text{scalar quantity}}$ enhances the thin edges.

for other cases, 1st order derivatives and 2nd order derivatives for flat segments will be non-zero

* Sum of the filter coefficients must be zero.
* Laplacian operator is rotation invariant.

-1	-1	-1
-1	8	-1
-1	-1	-1

0	-1	0
-1	4	-1
0	-1	0

⇒ center of mask is positive

1	1	1
1	-8	1
1	1	1

0	1	0
1	-4	1
0	1	0

⇒ center of mask is negative

$$g(x,y) = \begin{cases} f(x,y) - \nabla^2 f(x,y) & \text{[if center coefficient of Laplacian mask is -ve]} \\ \underbrace{f(x,y)}_{\text{added}} + \nabla^2 f(x,y) & \text{[— do +ve —]} \end{cases}$$

~~In order~~ to preserve the background tonality [to avoid ghost like appearance]

<u>Application</u> :— Enhance edge, line, point.

<u>Disadvantage</u> :— Enhances no flag segments (ghost like) appearance

0	-1	0
-1	5	-1
0	-1	0

$$\underbrace{f(x,y)}_{\text{preserve the background tonality}} - \underbrace{\nabla^2 f(x,y)}_{\text{enhances the edge}}$$

(for directly implementing the $g(x,y)$)

$P - \nabla^2 P$

(Mask for ↓ is Composite Laplacian Mask)

Or

-1	-1	-1
-1	9	-1
-1	-1	-1

Unsharp masking and High-Boost Filtering :-

Unsharp masking → $f_s(x,y) = f(x,y) - \bar{f}(x,y)$

[sharpened image] [smoothing operation (blurred image)]

* Here, we are bringing another function, $\bar{f}(x,y)$ which is averaging (unsharpening image) so it is called unsharp masking.

Blurring → suppresses the high frequency
[$\bar{f}(x,y)$] (rich with low frequency)

$f(x,y)$ contains both low and high frequency

$\therefore f(x,y) - \bar{f}(x,y)$ contains high freq.

High-boost filtering :-

$$f_{hb}(x,y) = A f(x,y) - \bar{f}(x,y)$$

$$f_{hb}(x,y) = (A-1) f(x,y) + f(x,y) - \bar{f}(x,y)$$

$$= (A-1) f(x,y) + f_s(x,y)$$

More higher ~~than~~ the A is, more is high-boosting

-1	-1	-1
-1	k+8	-1
-1	-1	-1

0	-1	0
-1	k+4	-1
0	-1	0

$$f_{hb} = \begin{cases} A f(x,y) - \nabla^2 f(x,y) & \text{[center of the mask is negative]} \\ A f(x,y) + \nabla^2 f(x,y) & \text{[center of the mask is positive]} \end{cases}$$

First order Partial Derivative:-

$$\nabla f = \hat{i}\frac{\partial f}{\partial x} + \hat{j}\frac{\partial f}{\partial y} \Big\} \text{vector}$$

$$|\nabla f| = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2}$$

individually $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ are linear

But, $|\nabla f|$ is non-linear as it contains square root terms.

$\frac{\partial P}{\partial x}, \frac{\partial P}{\partial y} \Rightarrow$ not rotational invariant.

$|\nabla P|$ → rotational invariant w.r.t $90°$

* $|\nabla P| = \left|\frac{\partial P}{\partial x}\right| + \left|\frac{\partial P}{\partial y}\right|$ [To save the enormous computation, we can define $|\nabla P|$ as like this]

~~Pre-witt's filter~~

$$\begin{matrix} z_1 & z_2 & z_3 \\ z_4 & z_5 & z_6 \\ z_7 & z_8 & z_9 \end{matrix}$$

horizontal edge (horizontal change) $z_6 - z_5$ $\begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}$

vertical edge (vertical change) $z_8 - z_5$ $\begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix}$

Robert Cross Gradient filter

$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ diagonal edge $(z_5 - z_9)$

$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ diagonal edge $(z_5 - z_7)$

Pixel values are concentrated b/w two values.

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

Pre-witts filter

* $z_4 - z_6$ is applied and is saved at z_5

Sobel Operators

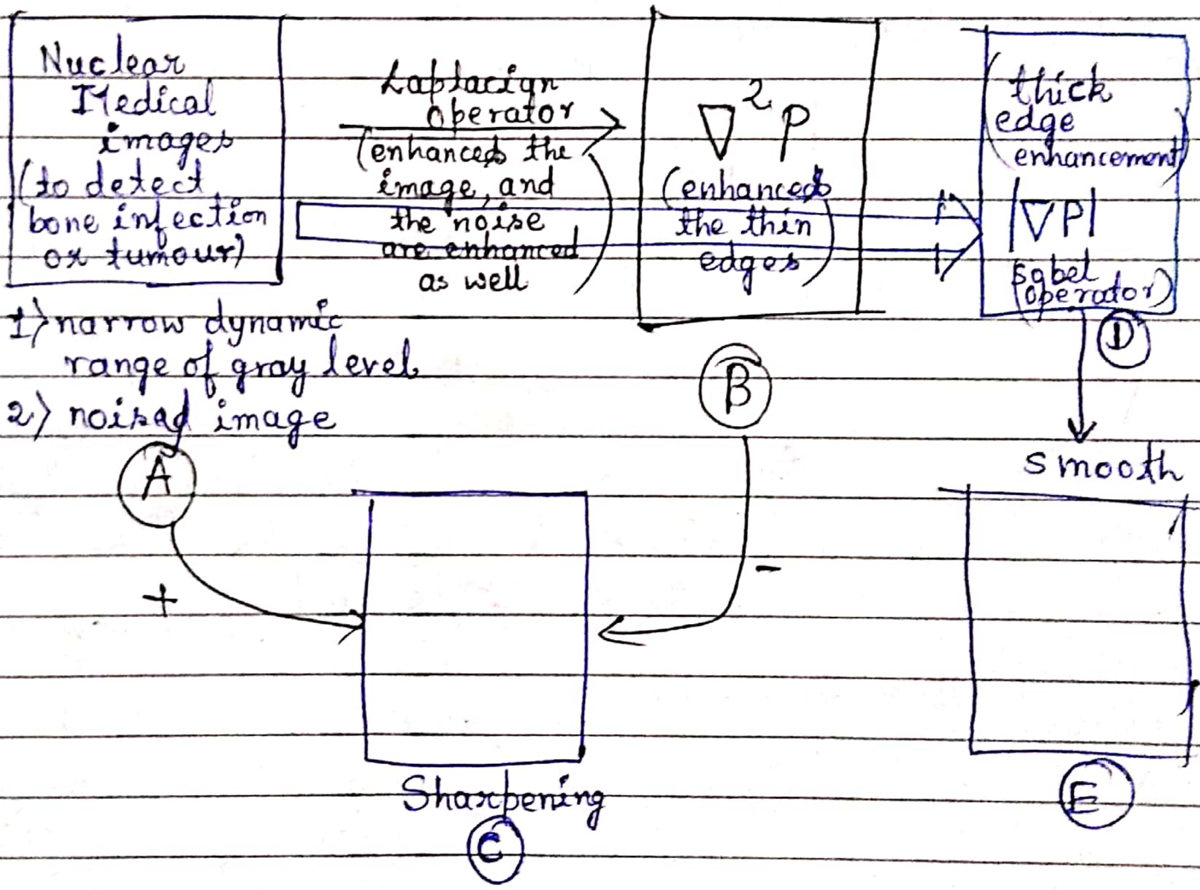
-1	0	1
-2	0	2
-1	0	1

to compute $\frac{\partial P}{\partial x}$.

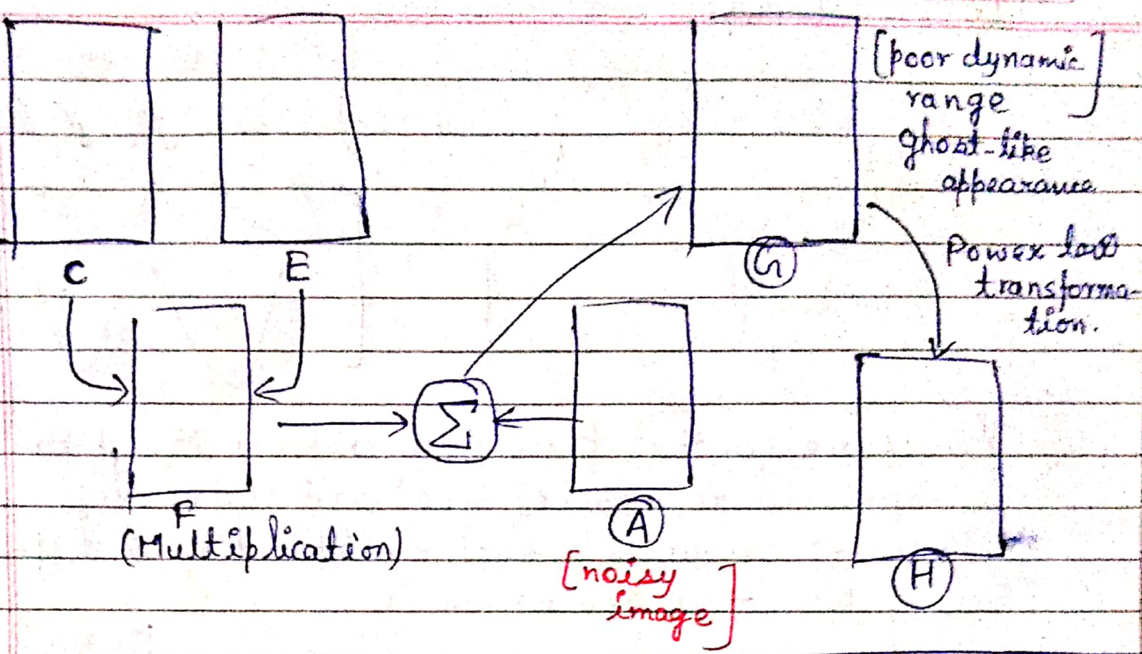
-1	-2	-1
0	0	0
1	2	1

to compute $\frac{\partial P}{\partial y}$.

Image Enhancement in the Spatial Domain :-



* Isolated details of the image is lost in case of Median filter, so it is not suggested in medical fields, even though it is better than mean filtering.



26/04/'21

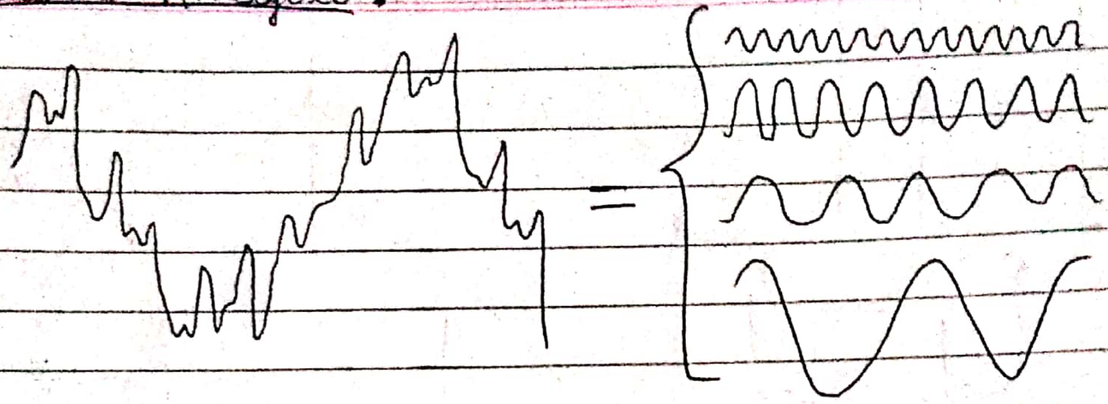
Image Enhancement in Frequency domain:-

Convolution in spatial domain ≡ Multiplication in frequency domain.

$$f(x,y) * h(x,y) \Longleftrightarrow F(u,v) \cdot \underbrace{H(u,v)}_{\text{Filter Transfer function}}$$

* $F(u,v)$ and $H(u,v)$ are the DFT of f and h respectively
* Multiplication on right hand side is component wise i.e., $|F(u,v)| \times |H(u,v)|$
* In some cases, we need to perform preprocessing operation to locate the filter at the centre of the plane.

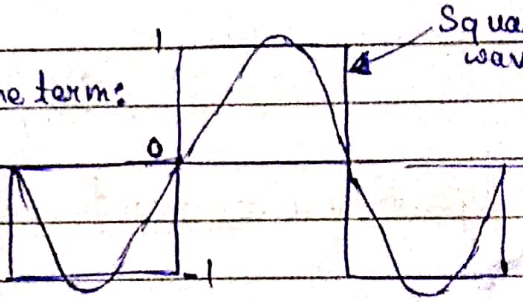
Fourier Analysis :-



Any function that periodically repeats itself can be expressed as a sum of sines and cosines of different frequencies each multiplied by a different coefficient — a Fourier Series

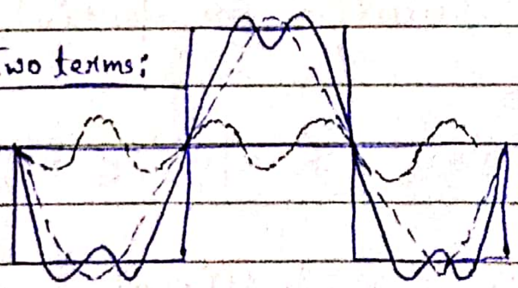
Fourier transform basic functions

One term:

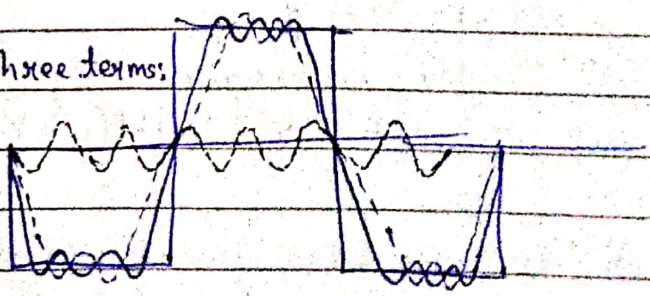


Approximating a square wave as the sum of sine waves

Two terms:



Three terms:



Fourier Transform:-

1-D, Discrete Case :-

Fourier Transform: $F(u) = \frac{1}{M} \sum_{x=0}^{M-1} f(x) e^{\frac{-j2\pi ux}{M}}$, $u = 0, \ldots, M-1$

Inverse Fourier Transform : $f(x) = \sum_{u=0}^{M-1} F(u) e^{\frac{j2\pi ux}{M}}$, $x = 0, \ldots, M-1$

F(u) can be written as:

$$F(u) = R(u) + j\,I(u) \Rightarrow$$

where $|F(u)| = \sqrt{R(u)^2 + I(u)^2}$ ⇒ magnitude

Polar coordinate:

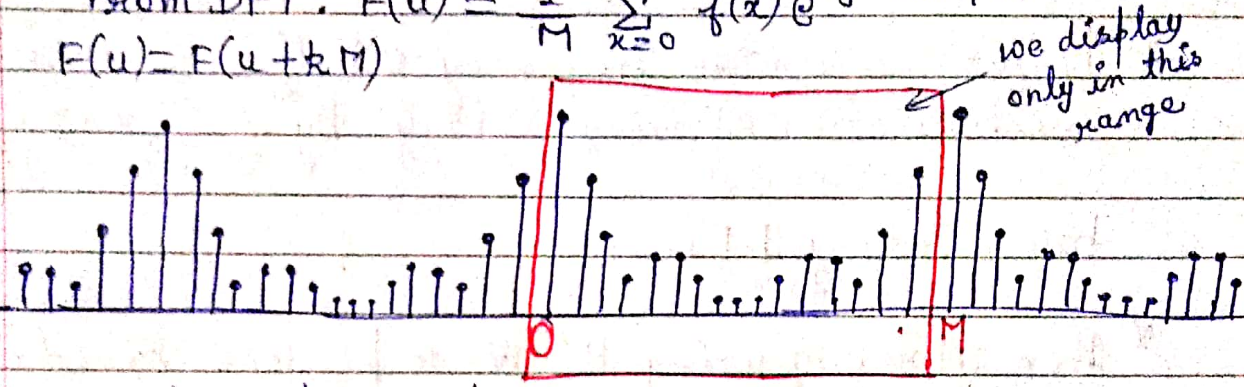
$$F(u) = |F(u)|\, e^{-j\phi(u)}$$

$\phi(u) = \tan^{-1}\left(\frac{I(u)}{R(u)}\right)$ ⇒ phase

Discrete Fourier Transform: Periodicity :—

From DFT: $F(u) = \frac{1}{M} \sum_{x=0}^{M-1} f(x)\, e^{-j 2\pi u x / M}$

$F(u) = F(u + kM)$



u = no. of complete cycles of the sinusoid that fits into the width M of the image. These form the basis func^n of the frequency domain representation and the weights for each sine and cosine function are known as Fourier coefficients.

* For an image of size M×N pixels:—

2-D DFT

$$F(u,v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y)\, e^{-j 2\pi (ux/M + vy/N)}$$

u = frequency in x direction, u = 0, ..., M−1

v = frequency in y direction, v = 0, ..., N−1

→ Same for F(u,v)

2-D IDFT

$$f(x,y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u,v) \times e^{j 2\pi (ux/M + vy/N)}$$

x = 0, ..., M−1
y = 0, ..., N−1

Inverse transform:— If we attempt to reconstruct the image with an inverse Fourier transform after destroying either the phase information or the amplitude information, reconstruction will fail.

* Phase carries more information

28|04|21

Butterworth Low Pass Filter:-

* Image to be filtered is a M×N pixeled image.
* D(u,v) is the distance from the centre to the point (u,v), given by:-

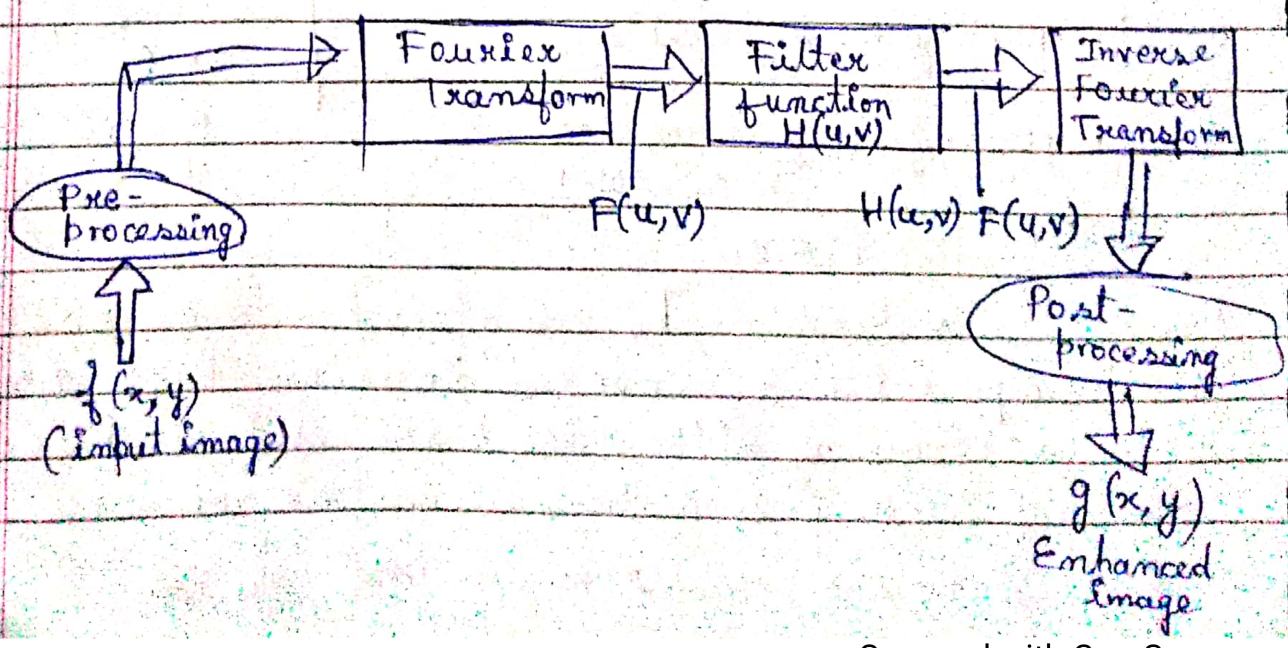
$$D(u,v) = \left[\left(u-\frac{M}{2}\right)^2 + \left(v-\frac{N}{2}\right)^2\right]^{1/2}$$

* To get the centre frequency at $\left(\frac{M}{2}, \frac{N}{2}\right)$, we need multiplication of image $f(x,y)$ by $(-1)^{x+y}$ [pre-processing] before taking DFT.

* Then filtering using H(u,v) to be done based on the application i.e., low-pass filtering for smoothing, noise cleaning, high-pass filtering for edge enhancement.

* Take the inverse DFT of filtered image.
* Then, again post processing is done multiplying again by $(-1)^{x+y}$.

Different steps in Frequency domain enhancement:-



* Multiplication in freq-domain is easier than convolution in the spatial domain

$$g(x,y) = f(x,y) * h(x,y) \Leftrightarrow F(u,v) \cdot H(u,v) = G(u,v)$$

Homomorphic Filtering:-

\# $f(x,y)$ suffers from range compression and poor contrast.

$$f(x,y) = \underbrace{i(x,y)}_{\text{illumination}} \cdot \underbrace{r(x,y)}_{\text{reflectance}}$$

$$\mathcal{F}\{f(x,y)\} \neq \mathcal{F}\{i(x,y)\}\,\mathcal{F}\{r(x,y)\}$$

$\mathcal{F}$ → Fourier transform.

$$z(x,y) = \ln f(x,y) = \ln i(x,y) + \ln r(x,y)$$

$$\mathcal{F}\{z(x,y)\} = \mathcal{F}\{\ln i(x,y)\} + \mathcal{F}\{\ln r(x,y)\}$$

~~Fourier~~

$$Z(u,v) = F_i(u,v) + F_r(u,v)$$

$$f(x,y) \Rightarrow \underbrace{\boxed{\ln}}_{\text{(pre-processing)}} \Rightarrow \boxed{DFT} \Rightarrow \boxed{H(u,v)} \Rightarrow \boxed{(DFT)^{-1}} \Rightarrow \underbrace{\boxed{\exp}}_{\text{(post-processing)}} \Rightarrow g(x,y)$$

$F_i(u,v)$ and $F_r(u,v)$ are fourier transform of $\ln i(x,y)$ and $\ln r(x,y)$ respectively.

If we process $Z(u,v)$ by means of a filter function $H(u,v)$

$$S(u,v) = H(u,v)\cdot Z(u,v)$$
$$= H(u,v)\cdot F_i(u,v) + H(u,v)\cdot F_r(u,v)$$

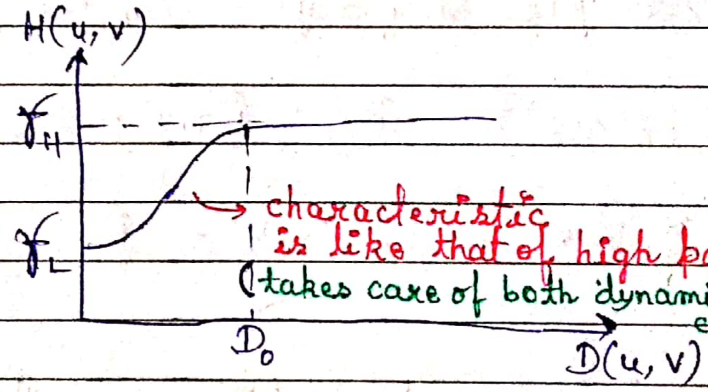
Where, $S(u,v)$ is Fourier transform of the result.
In spatial domain,

$$s(x,y) = \mathcal{F}^{-1}\{S(u,v)\}$$
$$= \mathcal{F}^{-1}\{H(u,v)\, F_i(u,v)\} + \mathcal{F}^{-1}\{H(u,v)\cdot F_r(u,v)\}$$

$$i'(x,y) = \mathcal{F}^{-1}\{H(u,v)\, F_i(u,v)\}$$

$$r'(x,y) = \mathcal{F}^{-1}\{H(u,v)\cdot F_r(u,v)\}$$

$$s(x,y) = i'(x,y) + r'(x,y)$$



characteristic is like that of high pass filter.
(takes care of both dynamic range expansion and contrast stretching)

$$g(x,y) = e^{s(x,y)} = e^{i'(x,y)}\cdot e^{r'(x,y)}$$
$$= i_0(x,y)\cdot r_0(x,y)$$

* A bell ~~is int~~ in its actual form is an integration while in inverted form, it is a differentiation

$$H(u,v) = (\gamma_H - \gamma_L)\cdot\left(1 - e^{-(u^2+v^2)}\right)$$

Spatial Domain		Frequency domain		
Window size is larger	⇒	smaller the radius of cylinder	⇒	Higher the Blurred effect

Brightness Image — Fourier Transform

- → Low-pass filtered (solid cylinder) — Inverse Transform (Blurred image as high frequency is lost)
- → High-pass filtered (hollow cylinder) — Inverse Transform (edge will be preserved) (texture)
- → Band-pass filtered (annular cylinder) — Inverse Transform

Gaussian filter < Butter worth filter < Ideal filter

Butterworth Low Pass Filter

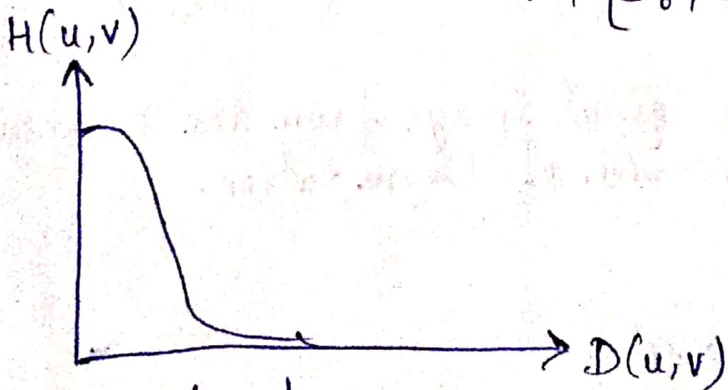
$$D(u,v) = \left[\left(u-\frac{M}{2}\right)^2 + \left(v-\frac{N}{2}\right)^2\right]^{1/2}$$

- For a low pass filter, this means that as $D(u,v)$ increases, $H(u,v)$ decreases. This effect is to dampen the higher frequencies which are represented as being a distance far from the centre and emphasize the lower frequencies which are represented by points close to the centre.
- D_0 is known as the cut-off frequency. In an ideal low pass filter, this is the point above which all frequencies would be eliminated.
- Increasing D_0 increases the number of frequencies that are passed (or lessen the damping effects of higher frequenc
- Decreasing D_0 means a smaller number of frequencies are allowed to pass (or the dampening effects of higher frequenc are increased). This would result in a more blurred image

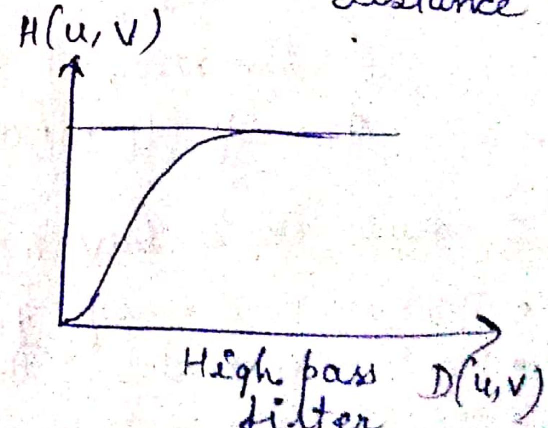
Butterworth High Pass filter is :-

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

$n \rightarrow$ order
$D_0 \rightarrow$ cut off distance



Low pass filter



High pass filter

application, i.e. low-pass filtering for smoothing, noise cleaning, high-pass filtering for edge enhancement.

For a high pass filter,

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

Inverse DFT of the filtered image is then done.

Then again post processing is done by multiplying again by $(-1)^{x+y}$.

30/4/'21

Laplacian in Frequency :- (for thin edge smoothing)

$$\mathcal{F}\left[\frac{d^n f(x)}{dx^n}\right] = (ju)^n F(u)$$

$$\mathcal{F}\left[\frac{\partial^2 f(x,y)}{\partial x^2} + \frac{\partial^2 f(x,y)}{\partial y^2}\right] = (ju)^2 F(u,v) + (jv)^2 F(u,v)$$

$$= -(u^2 + v^2) F(u,v)$$

$$\therefore \mathcal{F}[\nabla^2 f(x,y)] = -(u^2+v^2) F(u,v)$$

$$H(u,v) = -(u^2+v^2)$$

$$H(u,v) = -\left[\left(u-\frac{M}{2}\right)^2 + \left(v-\frac{N}{2}\right)^2\right]$$

$$\nabla^2 f(x,y) = \mathcal{F}^{-1}\left\{-\left[\left(u-\frac{M}{2}\right)^2 + \left(v-\frac{N}{2}\right)^2\right] F(u,v)\right\}$$

$$\nabla^2 f(x,y) \Leftrightarrow -\left[\left(u-\frac{M}{2}\right)^2 + \left(v-\frac{N}{2}\right)^2\right] F(u,v).$$

Unsharp masking, high-boost filtering and High frequency Emphasis filtering :-

$$f_{hp}(x,y) = f(x,y) - f_{lp}(x,y)$$

[ghost like image (high freq.)] [low freq components]

$$f_{hb}(x,y) = A\, f(x,y) - f_{lp}(x,y)$$

[high boost]

$$= (A-1)\, f(x,y) + f(x,y) - f_{lp}(x,y)$$

$$= (A-1)\, f(x,y) + f_{hp}(x,y)$$

$$F_{hp}(u,v) = F(u,v) - F_{lp}(u,v)$$

$$F_{lp}(u,v) = H_{lp}(u,v) \cdot F(u,v)$$

$$\boxed{H_{hp}(u,v) = 1 - H_{lp}(u,v)}$$

$$H_{hb}(u,v) = (A-1) + H_{hp}(u,v)$$

Digital Image Restoration:

* Here, we have been given a degraded image

Original image × Degradation function

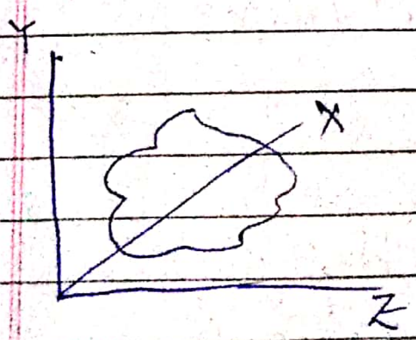
* In enhancement, approaches are heuristic to defeat short comings of the human visual system (subjective)

* In restoration, estimates of the degradation function is done to undo the degradation. (objective)

* We have to meet a goodness criteria.
* Restoration is also done for both spatial domain and frequency domain.
* It is mathematically richer process

* By supposing that the degradation is linear, we can exploit the characteristics of linear system (like convolution).

* If our degradation consideration is non-linear, mathematical processes becomes complex and tough.



▷ Sensor → $g(x,y)$ → A/D converter → $g(r,c)$

Ex:- Camera, Scanner, Ultrasound sensor, Infra-red sensor, MRI, CT, etc.

1. Geometric transformation
2. Photometric transformation

Improvement of Image quality :—

1) Image enhancement :— Adhoc processes improve the quality of the image using some heuristics designed on the basis of users' experience and application in hand.

2) Image Restoration :— Estimates the original (undergraded) image from the observed image based on the knowledge in terms of model of the degradation process.

Degradation Process :—

- Various sources of degradation.
- Nature of degradation.
- Examples of Spatial degradation
- Estimation of Degradation.
- Knowledge about noise.
- Restoration problem ill-conditioned to singular problem.

$$g = h * f + n$$

g = degraded image
f = original un-degraded image
n = noise

$$G = HF$$
$$G/H = F$$

Image Degradation Model

* Photometric transformation function $h(x,y)$ is linear and shift invariant.
* Noise is signal-independent and additive.
* Original and observed image intensity is non-negative.

Assumptions

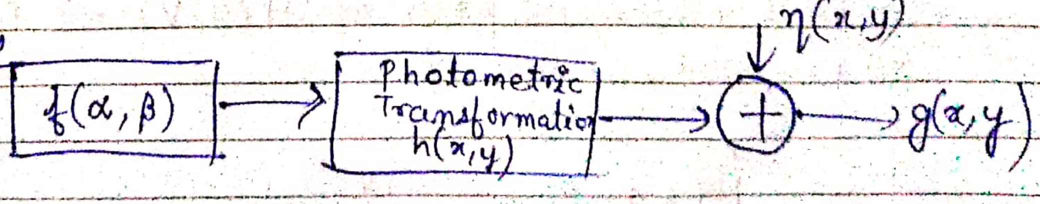


Image Formation Equation

* Suppose
  - $f(x,y)$ is the original image.
  - $h(x,y)$ is linear shift invariant transformation.
  - $\eta(x,y)$ is signal independent additive noise
  - $g(x,y)$ is the observed and recorded image.

* Image formation may be described as :—

$$g(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\alpha,\beta)\cdot h(x-\alpha, y-\beta)\, d\alpha\, d\beta + \eta(x,y)$$

* In discrete domain, for $M \times N$ digital image

$$g(r,c) = \sum_{j=0}^{N-1} \sum_{i=0}^{M-1} f(i,j)\, h(r-i, c-j) + \eta(r,c)$$

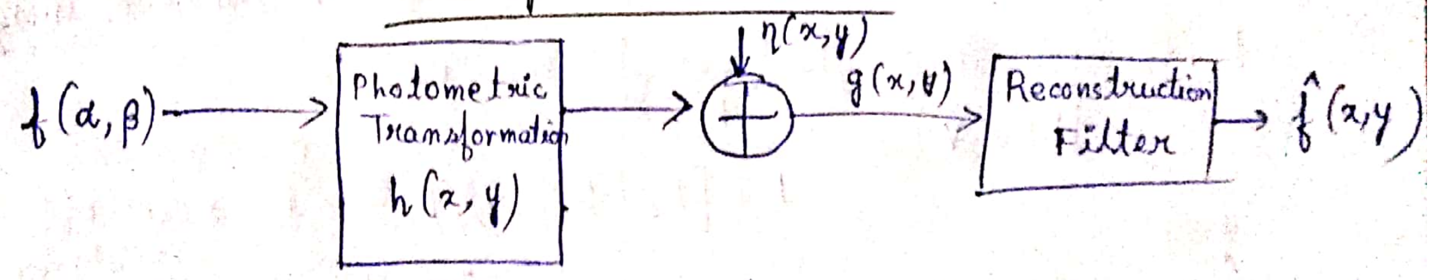
(2-dimensional convolution)

Using matrix vector notation,

$g = Hf + \eta$ (representing MN linear equation)

Degradation matrix is of size $MN \times MN$.

Image Restoration



Estimation of the original image from the observed one by effective inversion of Degradation.

Formulation :- Given →

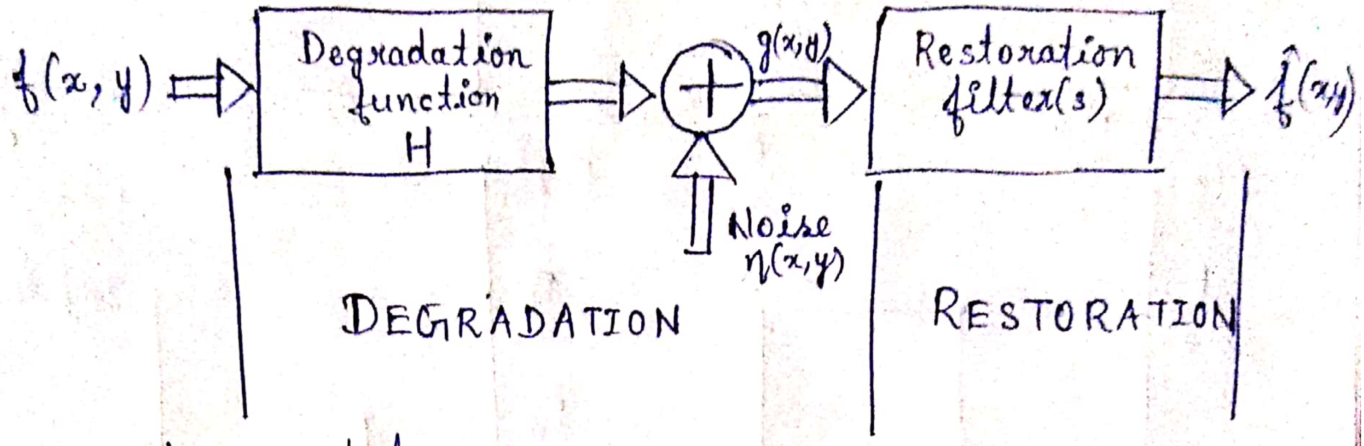
- the observed image $g(x,y)$ and
- the precise model of the PSF in form of $h(x,y)$ and
- the statistical knowledge about the noise $\eta(x,y)$

We need to estimate →

- the original image $f(x,y)$
- based on the relation : $g = Hf + \eta$

- Satisfying some criteria

* Image restoration is to restore a degraded image back to the original image while image enhancement is to manipulate the image so that it is suitable for a specific application.



Degradation model:

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

where, $h(x,y)$ is a system that causes image distortion and $\eta(x,y)$ is noise.

Noise Models:-

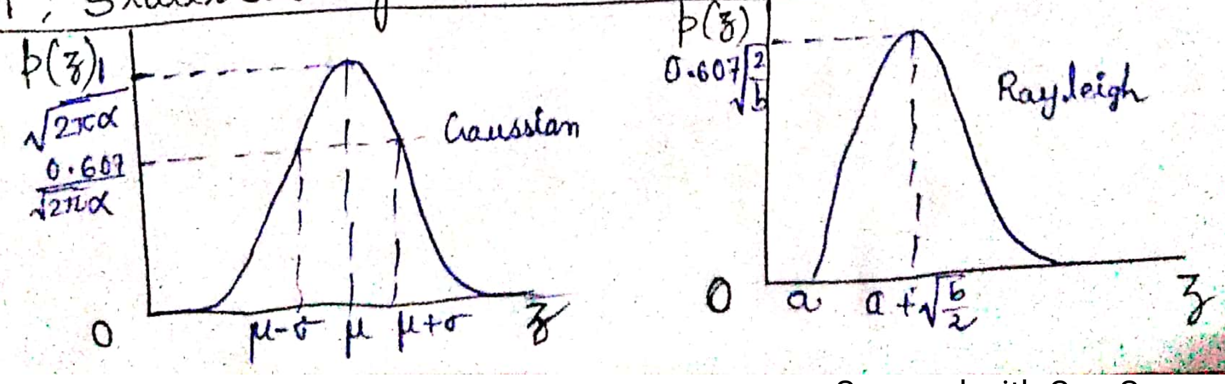
Noise cannot be predicted but can be approximately described in statiscal way using the PDF:-

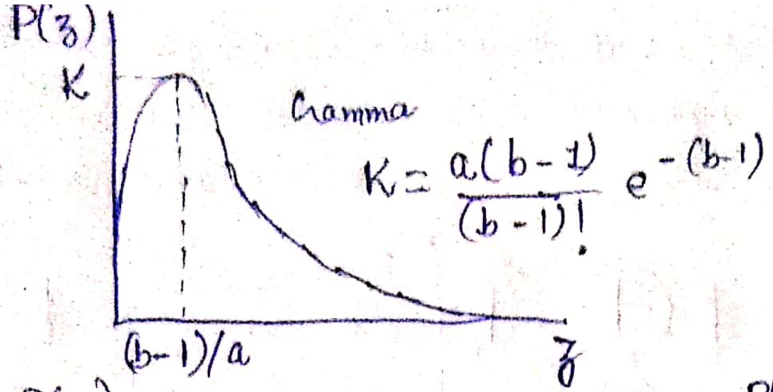
Gaussian noise: $p(z) = \frac{1}{2\pi\sigma} e^{-(z-\mu)^2/2\sigma^2}$

Rayleigh noise:
$$p(z) = \begin{cases} \frac{2}{b}(z-a)e^{-(z-a)^2/b}, & z \geq a \\ 0, & z < a \end{cases}$$

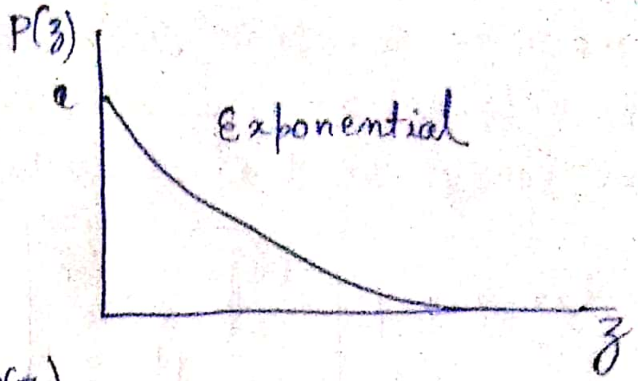
Erlang (Gamma) noise:
$$p(z) = \begin{cases} \frac{a^b z^{b-1}}{(b-1)!}(z-a)e^{-az}, & z \geq 0 \\ 0, & z < 0 \end{cases}$$

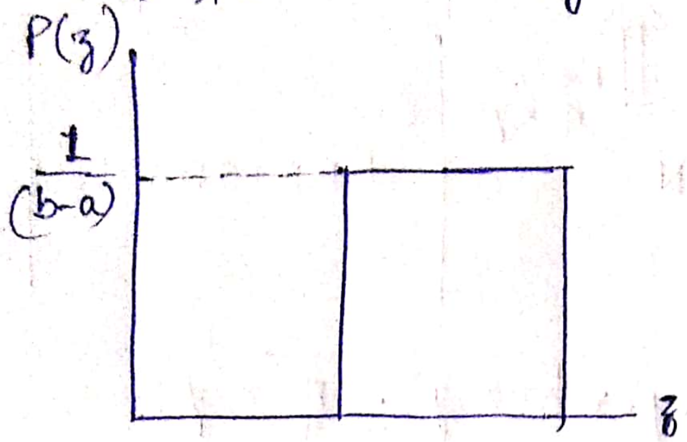
PDF: Statiscal way to describe Noise:-





$$K = \frac{a(b-1)}{(b-1)!} e^{-(b-1)}$$





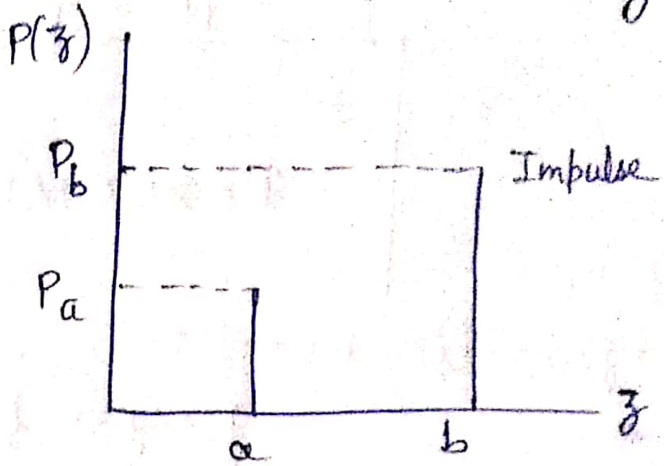
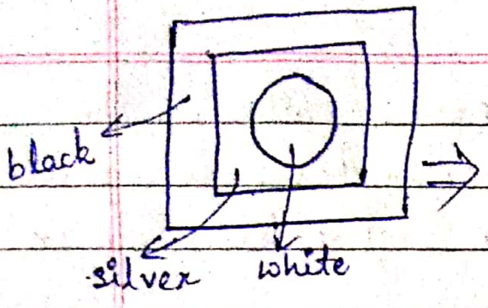
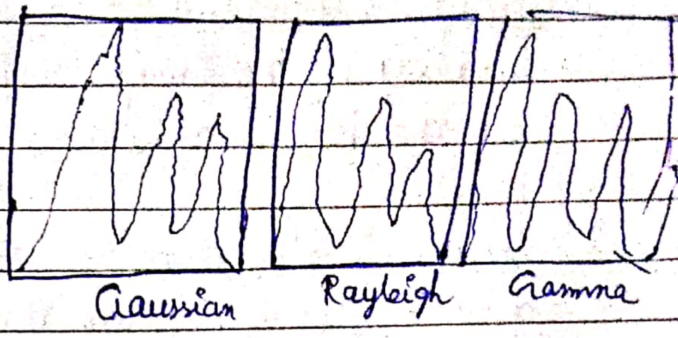


Image Degradation with Additive Noise:

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





Periodic noise:- looks like dots in frequency domains.

3/5/'21

Degradation model:-

degradation function, additive noise

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

or

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

Purpose: to estimate $h(x,y)$ or $H(u,v)$

Why? [If we know exactly $h(x,y)$, regardless of noise, we can do deconvolution to get $f(x,y)$ back from $g(x,y)$.]

Methods:

1. Estimation by Image Observation.
2. Estimation by Experiment.
3. Estimation by Modeling

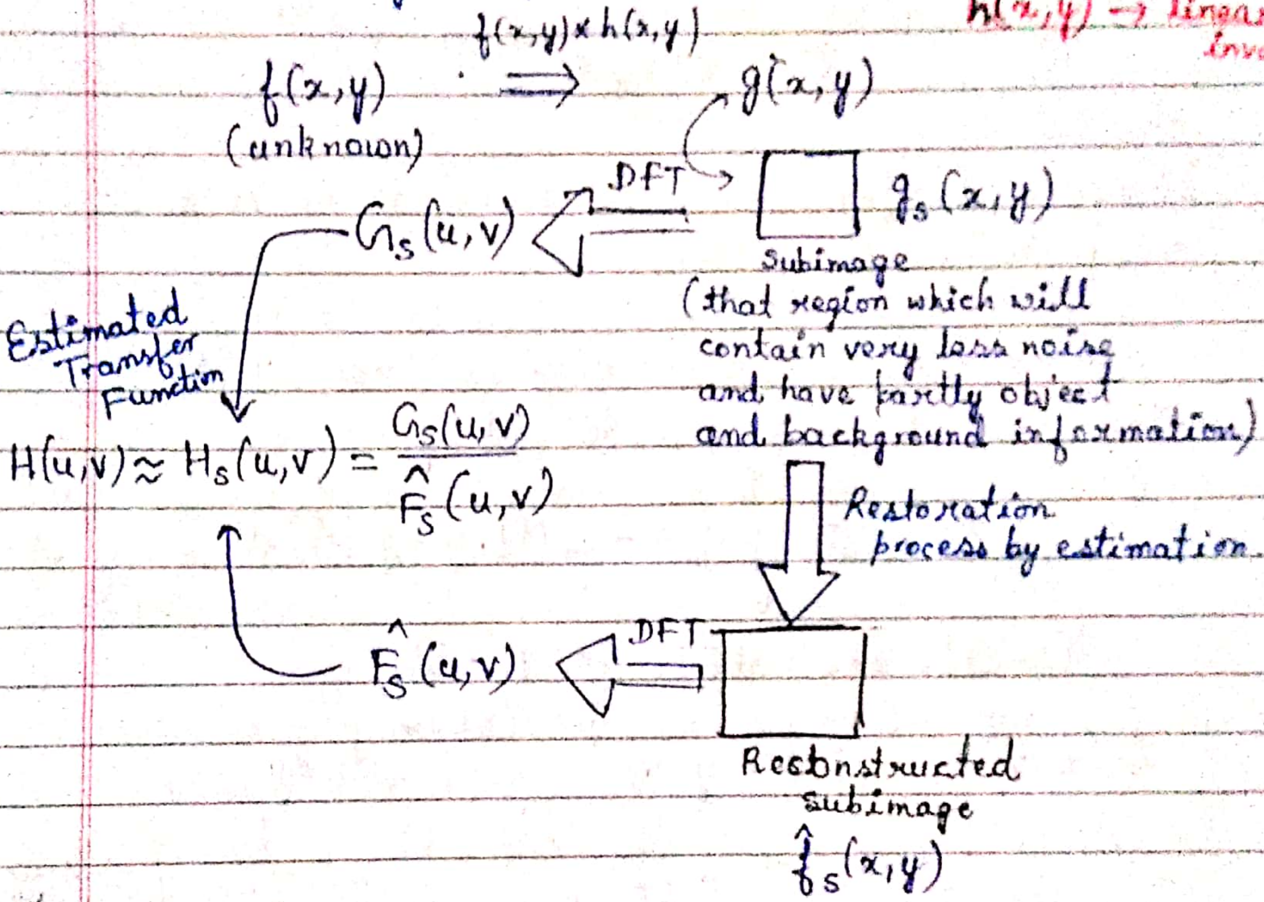
* We will never be able to get the knowledge of complete description of $H(u,v)$, we can only estimate it. $\Rightarrow$ Blind deconvolution

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

* Degraded image is basically a convolution of original image with the degradation function. To get back the original image, we have to use the deconvolution operation

Estimation by Image Observation :-

$h(x,y) \rightarrow$ linear-space invariant



This case is used when we know only $g(x,y)$ and cannot repeat the experiment!

Estimation by Experiment :-

* Use $f(x,y)$ and change the setting of the device and check for which settings, the degraded images being produced are similar

Input impulse image $\rightarrow$ System $H()$ $\rightarrow$ Response image from the system $g(x,y)$

$A\delta(x,y)$

$\downarrow$ DFT

$$DFT\{A\delta(x,y)\} = A \rightarrow H(u,v) = \frac{G(u,v)}{A} \leftarrow G(u,v)$$

$g(x,y) \xrightarrow{DFT} G(u,v)$

Estimation by Modeling :-

(used when we know physical mechanism underlying the image formation process that can be expressed mathematically).

Ex:- Atmospheric Turbulence model.

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

B/w the Imaging platform and the scene, there is a uniform planar motion.

Assume, that camera velocity is $(x_0(t), y_0(t))$

The blurred image is obtained by -

$$g(x,y) = \int_0^T f(x + x_0(t), y + y_0(t))\, dt$$

T= exposure time

$$G(u,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)\, e^{-j2\pi(ux+vy)}\, dx\, dy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left[\int_0^T f(x + x_0(t), y + y_0(t))\, dt \right] e^{-j2\pi(ux+vy)}\, dx\, dy$$

$$= \int_0^T \left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x+x_0(t), y+y_0(t))\, e^{-j2\pi(ux+vy)}\, dx\, dy \right] dt$$

$$= \int_0^T \left[F(u,v)\, e^{-j2\pi(ux_0(t) + vy_0(t))} \right] dt$$

$$= F(u,v) \int_0^T e^{-j2\pi(ux_0(t) + vy_0(t))}\, dt$$

Then, we get the motion blurring transfer function :

$$H(u,v) = \int_0^T e^{-j2\pi(ux_0(t) + vy_0(t))}\, dt.$$

For constant motion, $(x_0(t), y_0(t)) = (at, bt)$

$$H(u,v) = \int_0^T e^{-j2\pi(ua+vb)t} dt = \frac{T}{\pi(ua+vb)} \sin(\pi(ua+vb)) e^{-j\pi(ua+vb)}$$

$$\therefore H(u,v) = \frac{T}{\pi(ua+vb)} \cdot \sin(\pi(ua+vb)) e^{-j\pi(ua+vb)}.$$

05/05/'21 **Inverse Filter** (practically not used)

From degradation model:

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

After we obtain $H(u,v)$, we can estimate $F(u,v)$ by the inverse filter:

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

Problems

1. we don't know the frequency domain description of $N(u,v)$
2. Noise is enhanced, when $H(u,v)$ is small.

To avoid the side effect of enhancing noise, we can apply this formulation to freq. component (u,v) with in a radius D_0 from the center of $H(u,v)$.

Restoration: Formulation :-

Given
1) Observed image : $g(x,y)$
(degraded)
2) Precise model of the PSF in form of $h(x,y)$ and
3) statistical knowledge about the noise $\eta(x,y)$

We need to estimate
- the original image $f(x,y)$.
- based on the relation : $g = Hf + \eta$.

satisfying some criteria.

① Algebraic approach ② Probabilistic approach

Minimum Mean Square Estimation :- (a single, a single equipment, noise is coming from same source)

Suppose there are n images $g_1, g_2, g_3, \ldots, g_n$
where, $g_i = Hf_i + \eta_i$ $\quad i = 1, 2, 3, \ldots, n$

We need to design a filter '$P(\cdot)$' such that
$$\hat{f}_i = Pg_i$$
→ kind of like inverse degradation func^n

and average error of $e_i = \|f_i - \hat{f}_i\|^2$ is minimum

$\|\ \|$ → norm $\quad \|\ \|^2$ → square norm

$e_i = f_i - \hat{f}_i$, e_i can be positive or negative
To make it positive, we do square norm.

$$\|e_i\|^2 = e_i^T \cdot e_i \quad \text{(in matrix form)}$$

$$Exp(e) = Exp\{\|f_i - \hat{f}_i\|^2\}$$

$$= Exp\{(f_i - \hat{f}_i)^T (f_i - \hat{f}_i)\}$$

$$= Exp[\,tr\{(f_i - \hat{f}_i)(f_i - \hat{f}_i)^T\}] \text{ is minimum}$$

Minimization of expectation of $E^T \cdot E$ is same as minimization of expectation of trace of $E \cdot E^T$

* Error in estimation depends on the filter 'P'.
* So, error must be a function of 'P'

$$e(P) = Exp[tr\{(f_i - P(g_i))(f_i - P(g_i))^T\}]$$

$$= Exp[tr\{(f_i - P(Hf_i + n_i))(f_i - P(Hf_i + n_i))^T\}]$$

$$= Exp[tr\{(f_i - PHf_i - Pn_i)(f_i^T - f_i^T H^T P^T - n_i^T P^T)\}]$$

Since both Exp (expectation) and tr (trace) are both the linear operations, they can be exchanged.

$$e(P) = tr\Big[Exp\Big\{f_i f_i^T - f_i f_i^T H^T P^T - f_i n_i^T P^T - PHf_i f_i^T + PHf_i f_i^T H^T P^T + PHf_i n_i^T \cdot P^T - Pn_i f_i^T + Pn_i f_i^T H^T P^T + Pn_i n_i^T P^T\Big\}\Big]$$

$Exp(f_i f_i^T) = R_f$ = autocorrelation matrix

$Exp(f_i n_i^T) = 0$ [Since signal and noise are independent]
$= Exp(n_i f_i^T)$

$Exp(n_i n_i^T) = R_n$ = autocorrelation noise matrix

$$\therefore e(P) = tr[R_f - R_f H^T P^T - PHR_f + PHR_f H^T P^T + PR_n P^T]$$

$$g_i = H f_i + n_i$$

$(mn \times 1)$ $(mn \times mn)$ $(mn \times 1)$ $(mn \times 1)$

$$e(P) = tr[R_f - R_f H^T P^T - P H R_f + P H R_f H^T P^T + P R_\eta P^T]$$

Taking derivative of e w.r.t P and making it 0 (column vector)

$$\frac{de}{dp} = 0 = -R_f H^T - R_f H^T + 2PHR_f H^T + 2PR_\eta$$

Solving for 'p',

$$P^* = R_f H^T (H R_f H^T + R_\eta)^{-1}$$

And, in Fourier domains,

$$P^*(u,v) = \frac{R_f(u,v) H^*(u,v)}{R_f(u,v)|H(u,v)|^2 + R_\eta(u,v)}$$

(optimal value of P)

i-th image is estimated in Fourier domain as :-

$$\hat{F}_i(u,v) = P(u,v) \cdot G(u,v) = \frac{H^*(u,v)}{|H(u,v)|^2 + \frac{R_\eta(u,v)}{R_f(u,v)}} G(u,v)$$

$P(u,v) \rightarrow$ Weiner filter

If we take ~~Inverse~~ fourier transform of autocorrelation function is Power spectral Density.

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

$$= \left[\frac{1}{H(u,v)} \frac{|H(u,v)|^2}{|H(u,v)|^2 + S_\eta(u,v)/S_f(u,v)}\right] G(u,v)$$

I need to know that PSD of noise ($S_\eta(u,v)$) and PSD of f ($S_f(u,v)$). But here, we need the ratio of $S_\eta(u,v)$ and $S_f(u,v)$, so as an engineer we can substitute it as a constant and tweak it as per our control.

* The intuition that we could have also substituted

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

to be a constant,

but here we have eliminated the fact that for low $H(u,v)$, it can be difficult. As now it depends on both $|H(u,v)|^2$ and $\dfrac{S_\eta(u,v)}{S_f(u,v)}$.

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

* When there is a high domination of noise, Weiner filter gives better results than Inverse filter.

* Weiner's filter gives optimal restoration of the image after taking an ensemble of the degraded image. It does not ascertain which degraded image will produce the optimal restoration.

$\therefore$ Our design variable instead of 'P' should be $\hat{f}$.

Least-square Estimation :-

$$e(\hat{f}) = \|g - H\hat{f}\|^2 - \|\eta\|^2$$

Diff. b/w the observed image and re-degraded estimated image

$$= (g - H\hat{f})^T (g - H\hat{f}) - \eta^T \eta$$

$$= (g^T - \hat{f}^T H^T)(g - H\hat{f}) - \eta^T \eta$$

$$= g^T g - g^T H \hat{f} - \hat{f}^T H^T g + \hat{f}^T H^T H \hat{f} - \eta^T \eta$$

Differentiating w.r.t $\hat{f}$ and equating it to $\vec{0}$ (column vector)

$$\frac{\partial e(\hat{f})}{\partial \hat{f}} = \vec{0} = 0 - H^T g - H^T g + 2 H^T H \hat{f} - 0$$

$$\Rightarrow \hat{f} = (H^T H)^{-1} H^T g = H^{-1} g$$

And, in the Fourier domain, the estimated image

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

Singularity Problem $(g_i = H f_i + \eta_i)$

(I have many solutions for estimation)

- no. of variables (unknown) is more than no. of equations.
- image may not be estimated if for some (u,v), $H(u,v)$ becomes $\hat{0}$.

06/05/21

Constrained Least-square Estimation:—

* Reason for singularity → non-uniqueness of solution (we may have large number of estimated images for a given observed image)

* Suppose, the estimated image $\hat{f}$ is obtained by some means which satisfies a criterion $Q(\hat{f})$

* Error ($e(\hat{f})$) in estimated is defined as how good it satisfies $Q(\hat{f})$ given image model as constraint, i.e., $g = Hf + \eta$

I don't have any knowledge of η. From least-square, we can get $\hat{f}$. Out of many possible $\hat{f}$, we will take that for which $g - H\hat{f} = \eta$ ⇒ Equality constraint

$$e(\hat{f}) = \|Q\hat{f}\|^2 + \lambda \underbrace{\left(\|g - H\hat{f}\|^2 - \|\eta\|^2\right)}_{\text{measurement error}}$$

$$= (Q\hat{f})^T(Q\hat{f}) + \lambda\left[(g - H\hat{f})^T(g - H\hat{f}) - \eta^T\eta\right]$$

$$= (\hat{f}^T Q^T) Q\hat{f} + \lambda\left[(g^T - \hat{f}^T H^T)(g - H\hat{f}) - \eta^T\eta\right]$$

$$= \hat{f}^T Q^T Q\hat{f} + \lambda\left[g^T g - g^T H\hat{f} - \hat{f}^T H^T g + \hat{f}^T H^T H\hat{f} - \eta^T\eta\right].$$

Differentiating w.r.t $\hat{f}$ and equating to zero:-

$$\frac{\delta e(\hat{f})}{\delta \hat{f}} = 0 = 2Q^T Q\hat{f} + \lambda\left[0 - H^T g - H^T g + 2H^T H\hat{f} - 0\right]$$

Solving for $\hat{f}$, we get:-

$$\hat{f} = (H^T H + \gamma Q^T Q)^{-1} H^T g, \text{ where } \gamma = \frac{1}{\lambda}$$

And, in fourier domain, estimated image is:-

$$\hat{F}_i(u,v) = \frac{H^*(u,v)}{H^*(u,v) \cdot H(u,v) + \gamma Q^*(u,v)\, Q(u,v)} G(u,v)$$

If $Q^*(u,v)Q(u,v)$ is substituted with $\frac{R_\eta(u,v)}{R_f(u,v)}$, then

$$\hat{F}_i(u,v) = \frac{H^*(u,v)}{H^*(u,v)\, H(u,v) + \gamma \frac{R_\eta(u,v)}{R_f(u,v)}} G(u,v)$$

This is known as Parametric Weiner filter.

Chandra and Dutta, Majumder
(Digital Image Processing
and Analysis)

Degradation model:

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

Written in a matrix form

$g = Hf + \eta$

Now, We will ~~look~~ look for smoothness criteria.

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

Objective:- find the min^m of criterion func^n

$$C = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \left[\nabla^2 f(x,y)\right]^2$$

Subject to the constraint

$$\|g - H\hat{f}\|^2 = \|\eta\|^2 \quad \text{where } \|w\|^2 = w^T w.$$

We get a constrained least square filter.

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

where, $P(u,v)$ = Fourier transform of $p(x,y)$

$$= \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$

* γ is adaptively adjusted to achieve the best result.

Adjusting γ:-

Define $r = g - H\hat{f}$. It can be shown that $\phi(\gamma) = r^T r = \|r\|^2$

We want to adjust γ so that $\|r\|^2 = \|\eta\|^2 \pm a$ → ①
(user defined)

① Specify an initial value of γ.
② compute $\|r\|^2$
③ Stop if ① is satisfied
Otherwise return step 2 after increasing γ if $\|r\|^2 < \|\eta\|^2 - a$
Or decreasing γ if $\|r\|^2 > \|\eta\|^2 + a$

Use the new value of γ to recompute.

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

$$\left.\begin{aligned}
\hat{F}(u,v) &= \left[\frac{H^*(u,v)}{|H(u,v)|^2 + \gamma |P(u,v)|^2}\right] G(u,v) \\
R(u,v) &= G(u,v) - H(u,v)\hat{F}(u,v) \\
\|r\|^2 &= \frac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} r^2(x,y)
\end{aligned}\right\} \text{For computing } \|r\|^2$$

$$\left.\begin{aligned}
m_\eta &= \frac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} \eta(x,y) \\
\sigma_\eta^2 &= \frac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} [\eta(x,y) - m_\eta]^2 \\
\|\eta\|^2 &= MN[\sigma_\eta^2 - m_\eta]
\end{aligned}\right\} \text{For computing } \|\eta\|^2$$

Geometric Operations :-

* Task is to give the image a geometric operation.
* it enables change in image geometry by moving pixels around in a carefully constrained manner.
* Every image has its own geometry that is visible in the form of spatial relationship among the group of pixels representing image features.
* Two images of same scene and same image features may have different geometries.
* For example, the distance between a pair of features in one image is greater than the other or a group of features lie in a straight line in one image may not be in other image
* Two needs for geometric operations —

a) to remove distortion inherent in imaging. i.e, images that suffer from geometric distortion.
In modern camera (lenses and CCD sensor), this is not ~~the~~ much needed, however, geometric operation is required for precise measurement in sizes and shapes of image features.

Typical use may include fish eye lens or log polar type where geometric correction is essential.
It can also be seen in imaging in remote sensing from spacecraft or aircraft.

b) to introduce deliberate distortion for matching one image with another

Ex:- i) Cartographer deliberately introduces distortion on terrestrial or planetary surfaces to conform a particular map projection

ii) image registration for two or more images with same scene to match up same image features

iii) object detection by subtracting successive frames, medical imaging for improving diagnosis through combining different imaging (x ray CT, NMRI)

* Three elements common for geometric operations —
  a) transformation equation that moves a pixel in new location.
  b) a procedure for applying these equations to an image
  c) some way for computation of gray value/intensity in a new location.

Simple technique for enlarging an image by an factor n is to copy each pixel value by a block of $n \times n$ pixels in output image. Two drawbacks →
  a) n can't be any arbitrary non-integer value
  b) greatly enhanced images have blocky appearance

Simpler problems are seen for image shrinking by integer value of "n"
  a) — do —
  b) information loss due to loss can't be restored, however, $(n \times n)$ block of gray values may be replaced by mean or median value - though not much improvement in visuality.

$$x' = T_x(x, y)$$
$$y' = T_y(x, y)$$

T_x, T_y → polynomials of x and y.

In simplest form, T_x and T_y are linear function of x and y →

$$x' = a_0 x + a_1 y + a_2$$
$$y' = b_0 x + b_1 y + b_2$$
(AFFINE TRANSFORMATION)

i) Typical affine operations may include any combination of translation, rotation, scaling and shearing. Coefficients a's, b's, c's are represented in matrix form.

ii) Complex operations like warping and morphing.

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} a_0 & a_1 & a_2 \\ b_0 & b_1 & b_2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Homogeneous coordinates
(representation of 2D points as three dimensional vectors. [using third dimension = 1])

Another option which is less compact and difficult to manipulate →

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a_0 & a_1 \\ b_0 & b_1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} a_2 \\ b_2 \end{bmatrix}$$

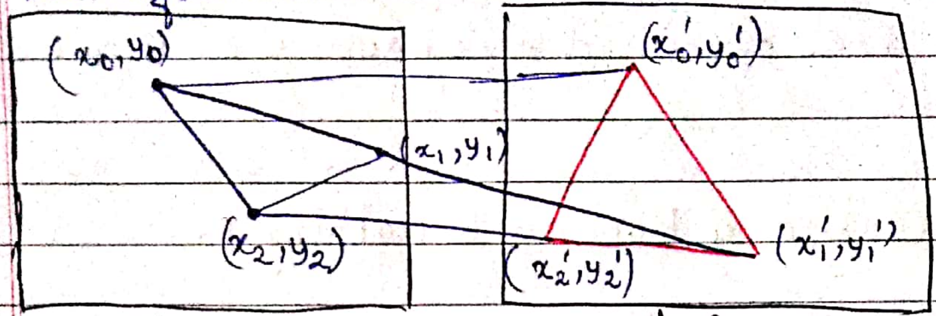
Under affine transformation, straight lines remain straight line and parallel lines remain parallel.
Ex:- Translation of pixel by 3 pixel down and 5 pixel up

$$x' = x + 5$$
$$y' = y + 3$$
$$\begin{bmatrix} 1 & 0 & 5 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix}$$

Transformation	a_0	a_1	a_2	b_0	b_1	b_2
Translation by $\Delta x, \Delta y$	1	0	Δx	0	1	Δy
Scaling by a factor s	s	0	0	0	s	0
Clockwise rotation through θ	$\cos\theta$	$-\sin\theta$	0	$\sin\theta$	$\cos\theta$	0
Horizontal shear by factor s	1	s	0	0	1	0

Alternative implementation (other than coefficient matrix)

On xy plane, three coordinate points before and after transformation →



I/P image   O/P image

Geometric transformation

Spatial transformation + Gray level interpolation

(Defines how pixels are to be rearranged in the spatially transformed image)

(assigns gray level values to pixels in spatially transformed image)

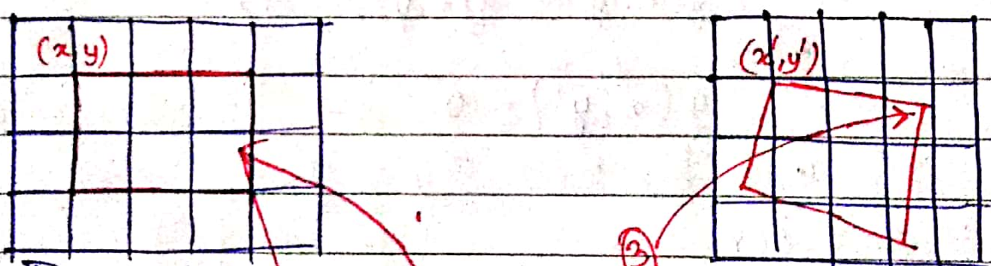


Image f to be restored   Distorted image g

1) Select coordinate (x, y) in f to be restored.

2) Compute :-

$$x' = r(x, y)$$
$$y' = s(x, y)$$

3) Go to pixel (x', y') in a distorted image g

4) get pixel value at $g(x', y')$ By gray level interpolation

5) store that value in pixel $f(x, y)$

Image Rotation by Forward mapping :-

* It is wasteful, not guaranteed lie within the image. Each o/P pixel addressed many times or worse still, not at all.

Image Rotation by Backward mapping :-

Create an o/p image, g, of dimensions M×N.

$a_0 = \cos\theta$

$a_1 = \sin\theta$

$b_0 = -a1$

$b_1 = a0$

for all pixel coordinates x', y' in g do

  $x = \text{round}(a_0 x' + a_1 y')$

  $y = \text{round}(b_0 x' + b_1 y')$

  if (x, y) is inside f, then

    $g(x', y') = f(x, y)$

  else

    $g(x', y') = 0$

  end if

end for

Still two problems remain unsolved. —

i) calculated coordinates are real number.

ii) pixels are outside the bounds of image (this time in i/P image)

Spatial Transformations :-

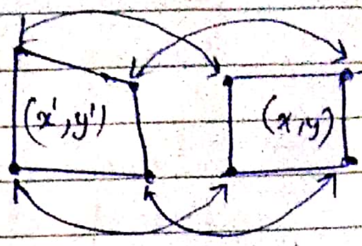
To map b/w pixel coordinate (x, y) of f and pixel coordinate (x', y') of g

$$x' = r(x, y) \qquad y' = s(x, y)$$

For a bilinear transformation mapping b/w a pair of quadrilateral regions,

$$x' = r(x,y) = c_1 x + c_2 y + c_3 xy + c_4$$
$$y' = s(x,y) = c_5 x + c_6 y + c_7 xy + c_8$$

To obtain $r(x,y)$ and $s(x,y)$, we need to know 4 pairs of coordinates (x,y) and its corresponding (x',y') which are called tiepoints.

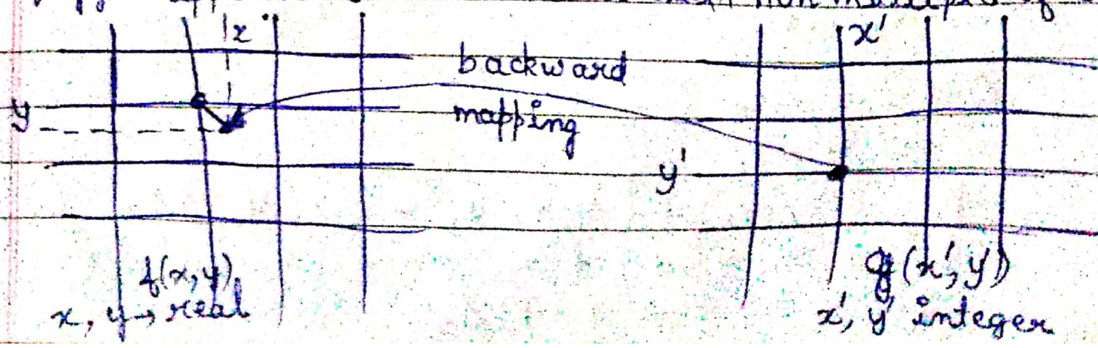


Gray level Interpolation : Nearest Neighbour :-
Since (x',y') may not be at an integer coordinate, we need to interpolate the value of $g(x',y')$

1. Nearest neighbour selection/zero order interpolation
2. Bilinear interpolation
3. Bi-cubic interpolation.

Zero order interpolation :-
The rounding off (x',y') to their nearest integers is known as zero-order or nearest neighbour interpolation.

Advantages :- 1) Computationally simple.
Disadvantages :- 1) Look blocky when scaled up in size by a large factor
2) horizontal or vertical lines may have jagged appearance when rotated with non multiple of 90°



First order interpolation :—

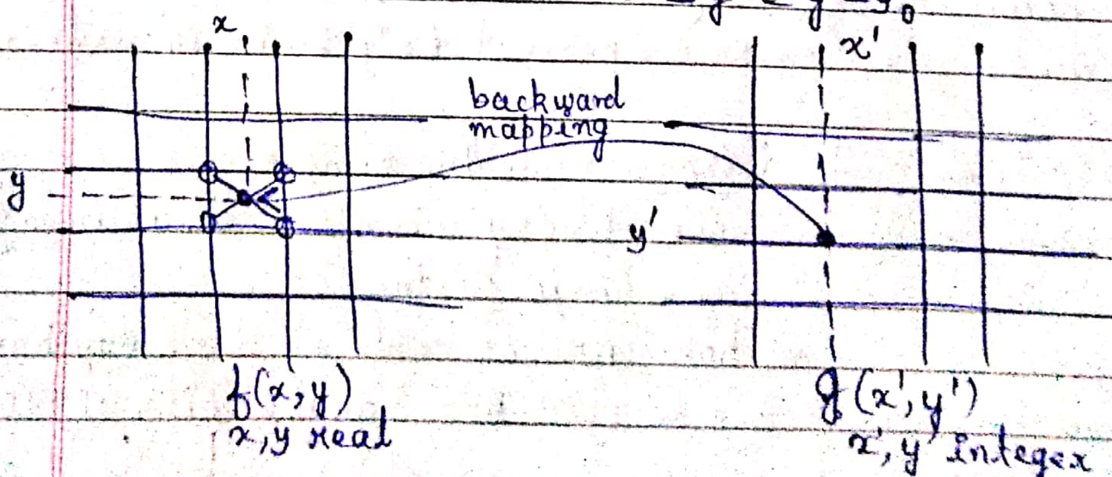
* requires ten additions and or subtractions and four multiplications including calculation of Δx and Δy
* slightly diff. approach requires eight additions or subtractions and three multiplications - looks not much benefit but for real images and for each pixel, this savings is a lot.

* It demands much more computation than zero-order interpolation but produces smoother and pleasant appearance.

* computes the o/p pixel gray value as the distance weighted function of four neighbouring pixel's gray values.

$$f(x',y') = f(x_0,y_0) + [f(x_1,y_0) - f(x_0,y_0)]\Delta x + [f(x_0,y_1) - f(x_0,y_0)]\Delta y + [f(x_1,y_1) + f(x_0,y_0) - f(x_0,y_1) - f(x_1,y_0)]\Delta x \Delta y$$

$$\Delta x = x' - x_0 \text{ and } \Delta y = y' - y_0$$

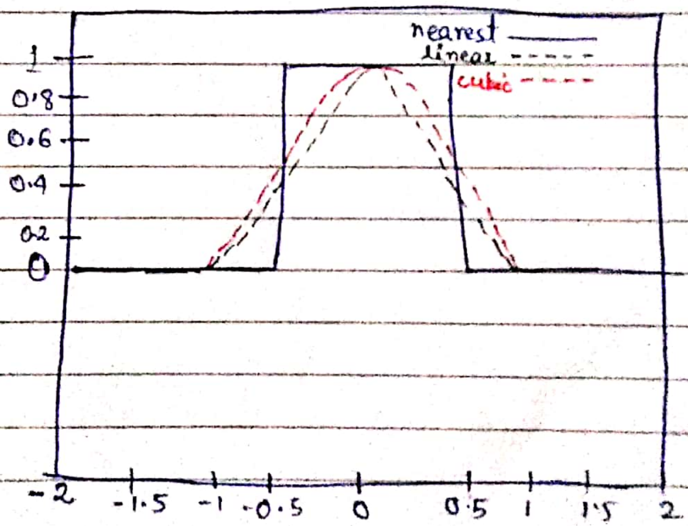


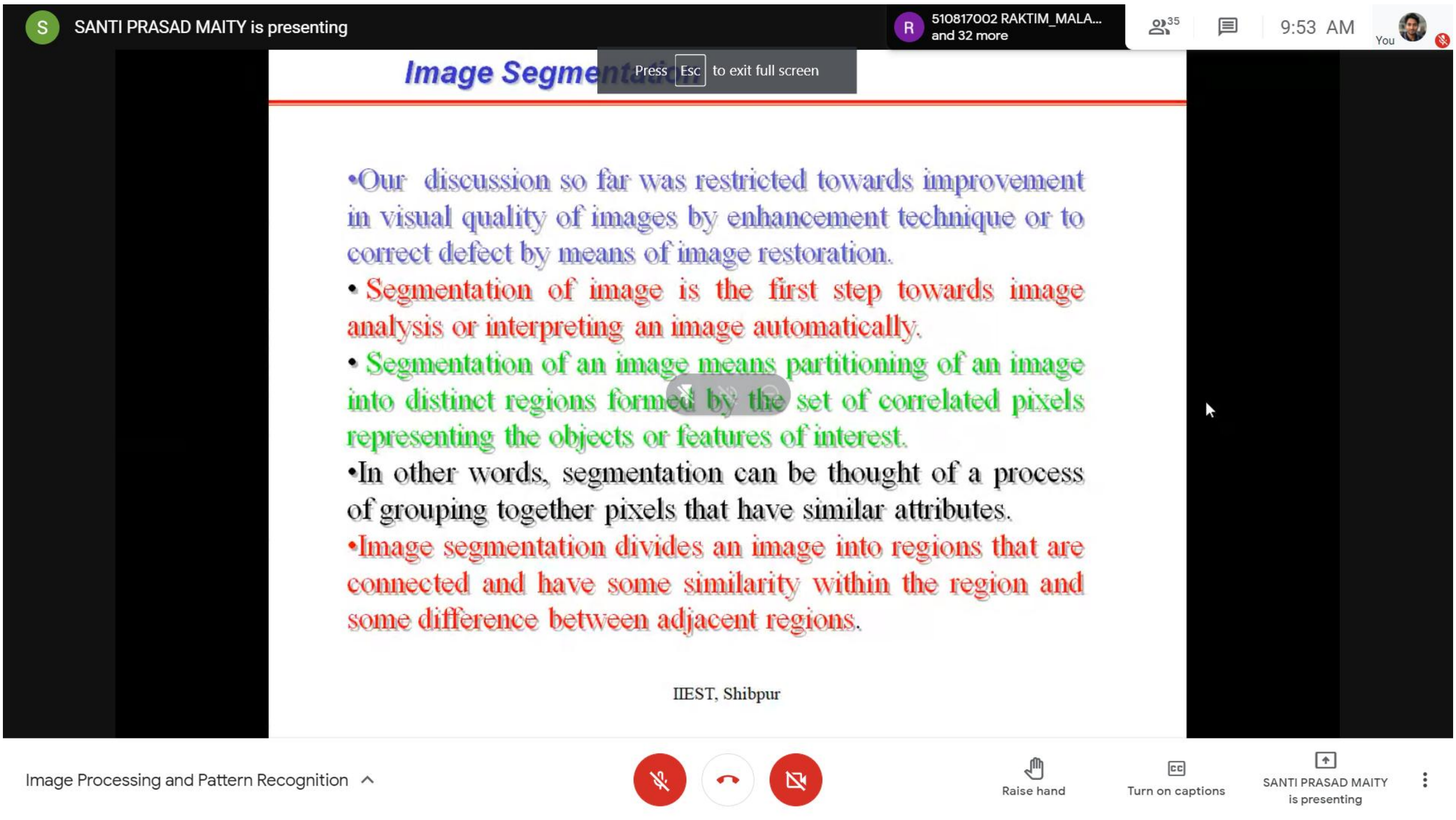
$$\Delta x = x' - x_0$$
$$\Delta y = y' - y_0$$
$$p = f(x_0,y_0) + [f(x_1,y_0) - f(x_0,y_0)]\Delta x$$
$$q = f(x_0,y_1) + [f(x_1,y_1) - f(x_0,y_1)]\Delta x$$
$$f(x',y') = p + (q-p)\Delta y$$

Higher order (bi-cubic) interpolation :-

* uses 16X16 neighbourhood pixels.
* convolution of (16X16) neighbors using cubic fn. $f(x)$
* jump discontinuity in zero-order causes blockiness/ jaggedness.
* Linear and bi-cubic are like low pass filter, difference ~~in~~ continuous slope in the latter, while the former shows sudden change in slopes.

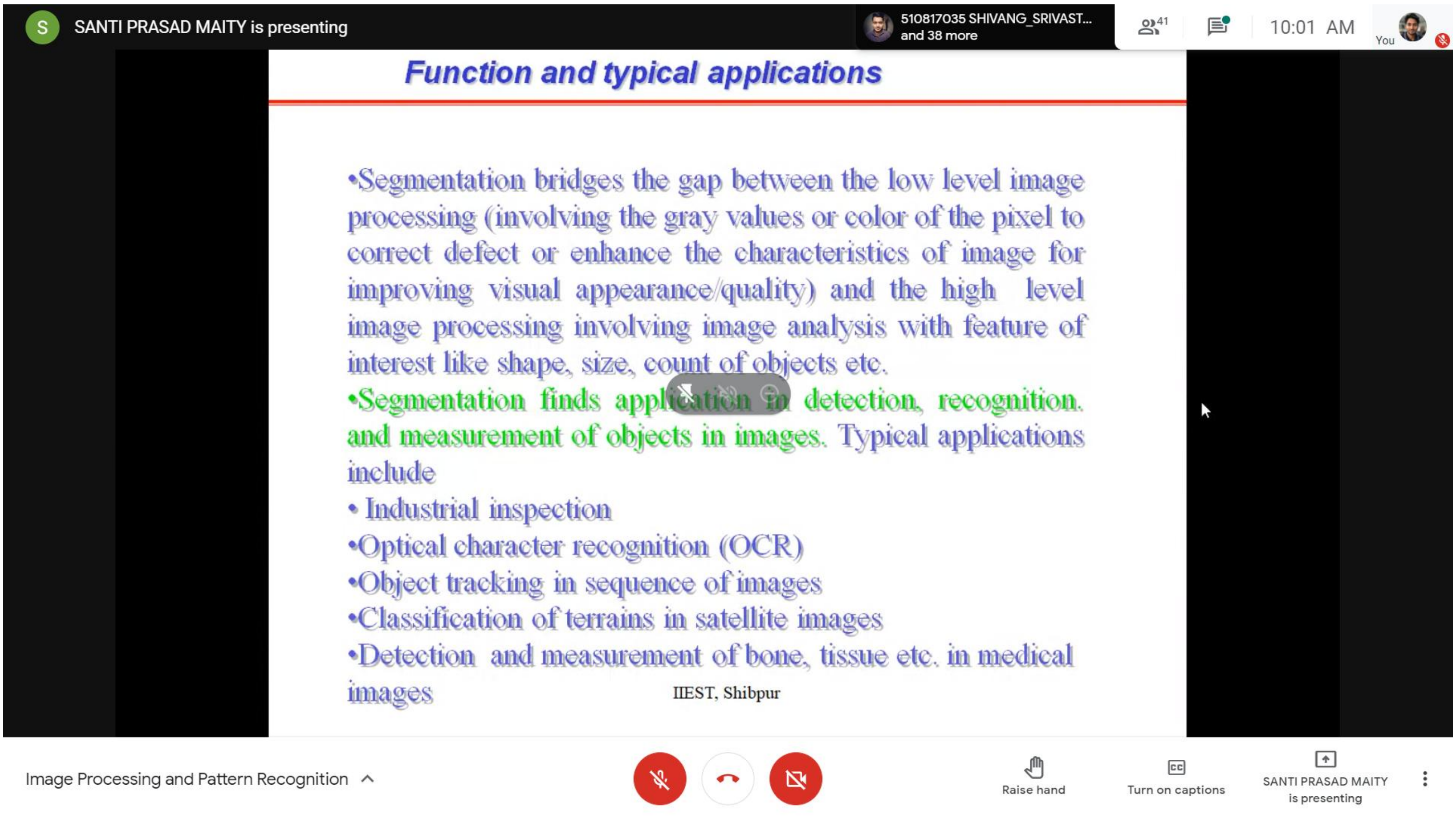
* cubic function has -ve slope, meaning that weighted sum of neighbouring pixels minus the contribution of slightly farther pixels.

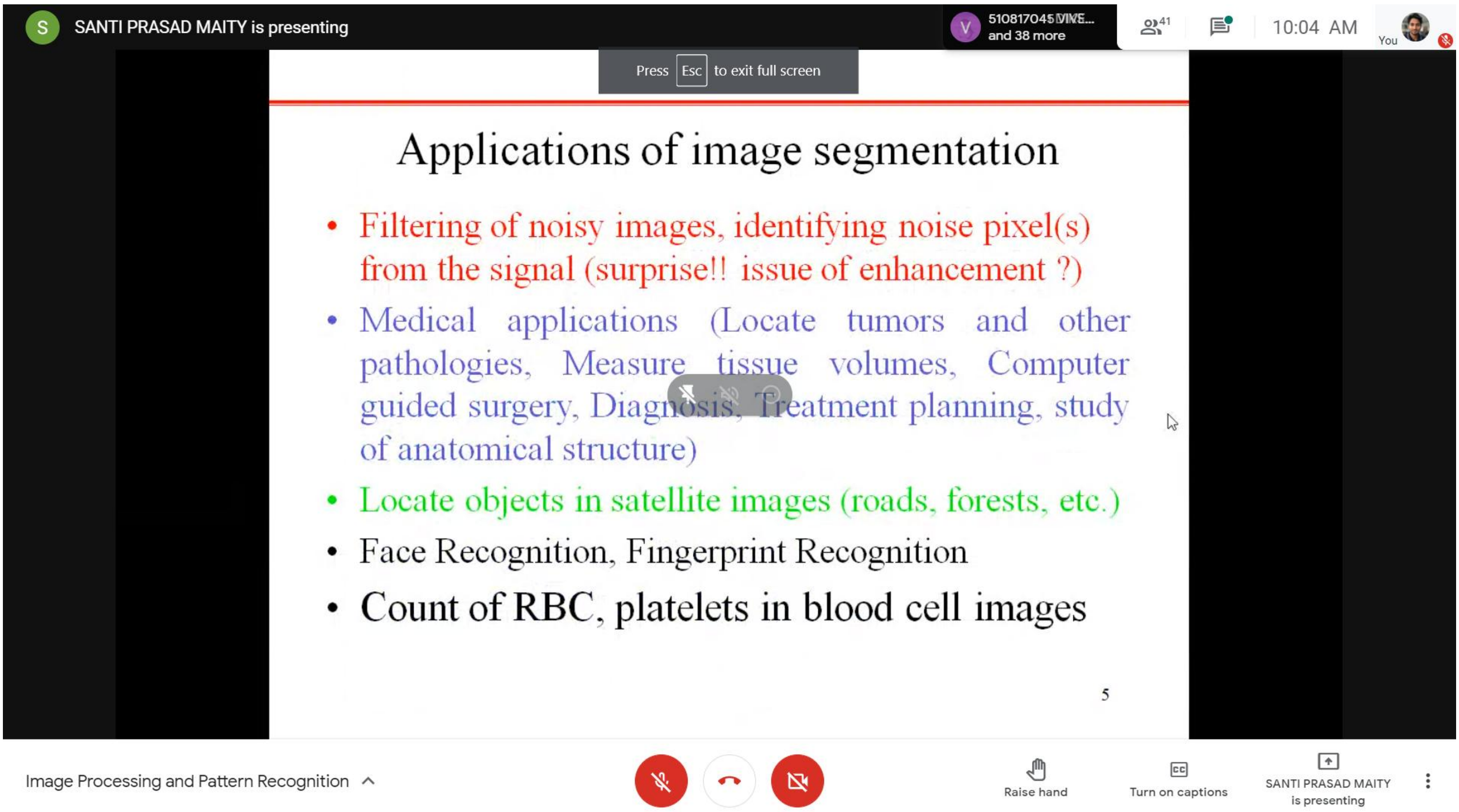


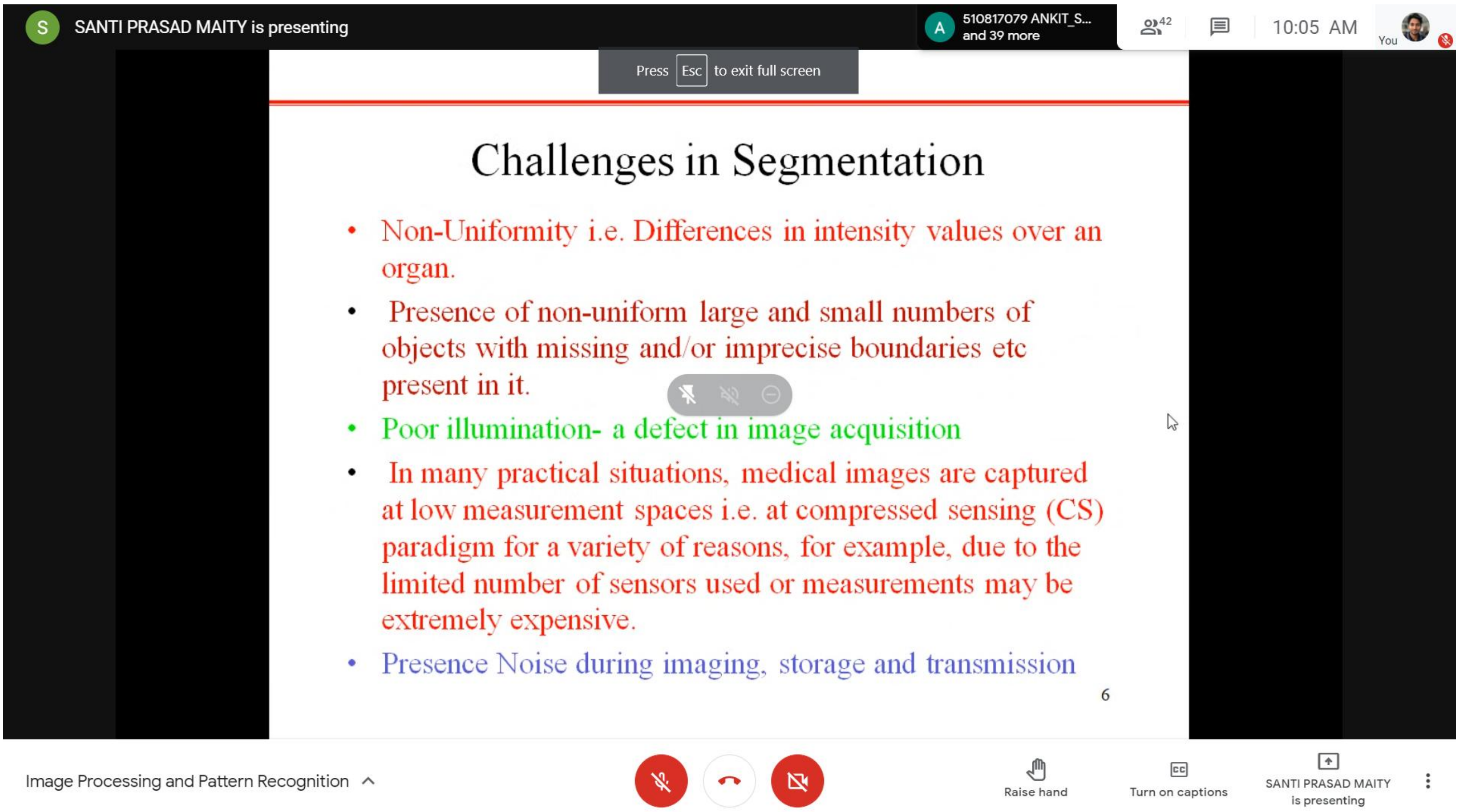

SANTI PRASAD MAITY is presenting
510817002 RAKTIM_MALA...
and 32 more
9:53 AM
You
Image Segme
Press Esc to exit full screen
•Our discussion so far was restricted towards improvement in visual quality of images by enhancement technique or to correct defect by means of image restoration.
• Segmentation of image is the first step towards image analysis or interpreting an image automatically.
• Segmentation of an image means partitioning of an image into distinct regions formed by the set of correlated pixels representing the objects or features of interest.
•In other words, segmentation can be thought of a process of grouping together pixels that have similar attributes.
•Image segmentation divides an image into regions that are connected and have some similarity within the region and some difference between adjacent regions.
IIEST, Shibpur
Image Processing and Pattern Recognition
Raise hand
Turn on captions
SANTI PRASAD MAITY
is presenting

Goals of image segmentation

- The purpose of image segmentation is to partition an image into *meaningful* regions with respect to a particular application
- The goal is usually to find individual objects in an image.
- Thus segmentation is based on measurements taken from the image and might be *grey level*, *colour*, *texture*, *depth* or *motion*


SANTI PRASAD MAITY is presenting
510817035 SHIVANG_SRIVAST...
and 38 more
10:01 AM
You
Function and typical applications
•Segmentation bridges the gap between the low level image processing (involving the gray values or color of the pixel to correct defect or enhance the characteristics of image for improving visual appearance/quality) and the high level image processing involving image analysis with feature of interest like shape, size, count of objects etc.
•Segmentation finds application in detection, recognition, and measurement of objects in images. Typical applications include
• Industrial inspection
•Optical character recognition (OCR)
•Object tracking in sequence of images
•Classification of terrains in satellite images
•Detection and measurement of bone, tissue etc. in medical images
IIEST, Shibpur
Image Processing and Pattern Recognition
Raise hand
Turn on captions
SANTI PRASAD MAITY
is presenting


SANTI PRASAD MAITY is presenting
Press Esc to exit full screen
Applications of image segmentation
• Filtering of noisy images, identifying noise pixel(s) from the signal (surprise!! issue of enhancement ?)
• Medical applications (Locate tumors and other pathologies, Measure tissue volumes, Computer guided surgery, Diagnosis, Treatment planning, study of anatomical structure)
• Locate objects in satellite images (roads, forests, etc.)
• Face Recognition, Fingerprint Recognition
• Count of RBC, platelets in blood cell images
5
10:04 AM
Image Processing and Pattern Recognition
Raise hand
Turn on captions
SANTI PRASAD MAITY is presenting


SANTI PRASAD MAITY is presenting
510817079 ANKIT_S...
and 39 more
42
10:05 AM
You
Press Esc to exit full screen
Challenges in Segmentation
• Non-Uniformity i.e. Differences in intensity values over an organ.
• Presence of non-uniform large and small numbers of objects with missing and/or imprecise boundaries etc present in it.
• Poor illumination- a defect in image acquisition
• In many practical situations, medical images are captured at low measurement spaces i.e. at compressed sensing (CS) paradigm for a variety of reasons, for example, due to the limited number of sensors used or measurements may be extremely expensive.
• Presence Noise during imaging, storage and transmission
6
Image Processing and Pattern Recognition
Raise hand
Turn on captions
SANTI PRASAD MAITY is presenting

SANTI PRASAD MAITY is presenting

510817020 ANINDYA_KU...
and 39 more

10:07 AM

You

Role of segmentation and challenges

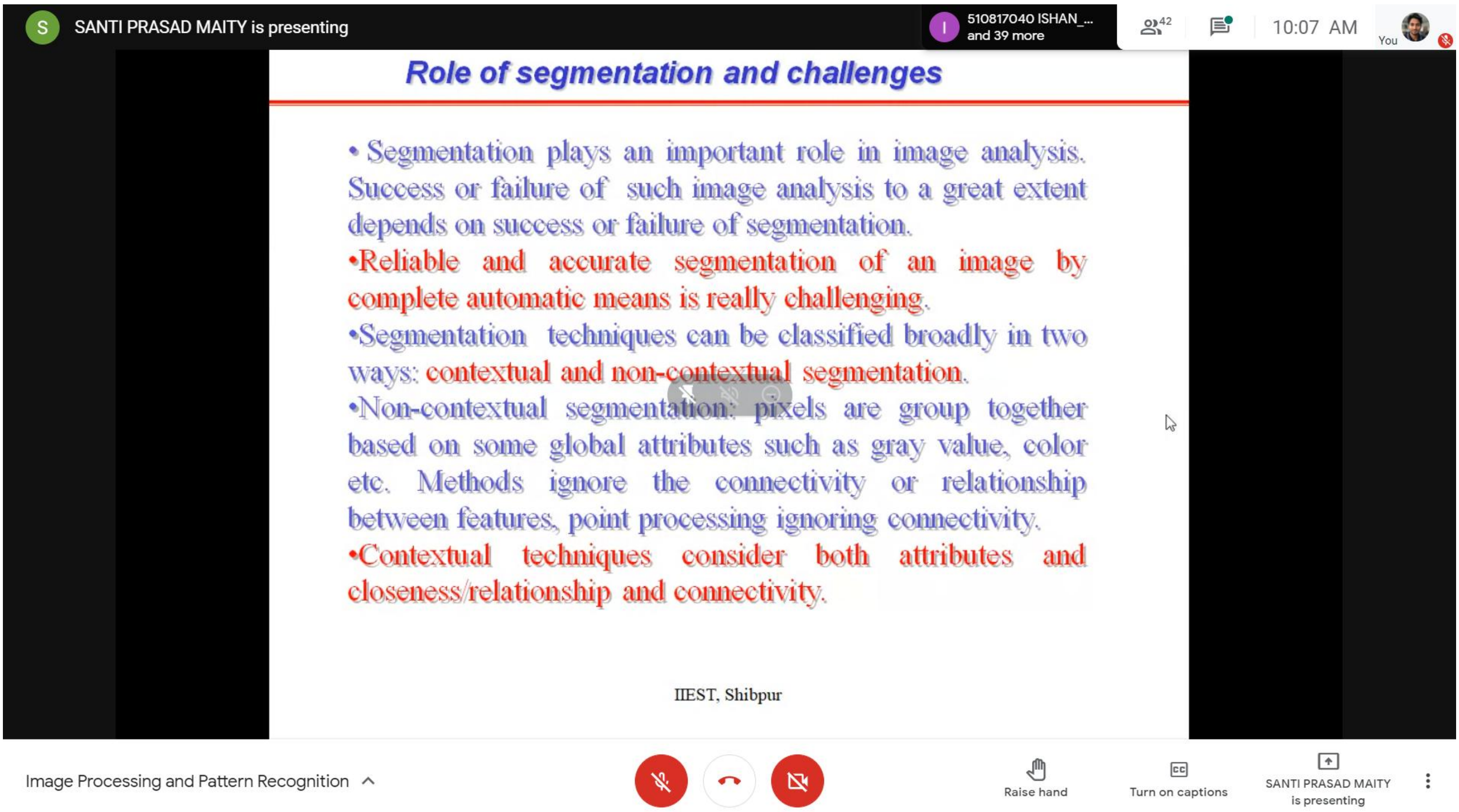
- Segmentation plays an important role in image analysis. Success or failure of such image analysis to a great extent depends on success or failure of segmentation.
- Reliable and accurate segmentation of an image by complete automatic means is really challenging.
- Segmentation techniques can be classified broadly in two ways: contextual and non-contextual segmentation.
- Non-contextual segmentation: pixels are group together based on some global attributes such as gray value, color etc. Methods ignore the connectivity or relationship between features, point processing ignoring connectivity.
- Contextual techniques consider both attributes and closeness/relationship and connectivity.

Image Processing and Pattern Recognition

Raise hand

Turn on captions

SANTI PRASAD MAITY
is presenting

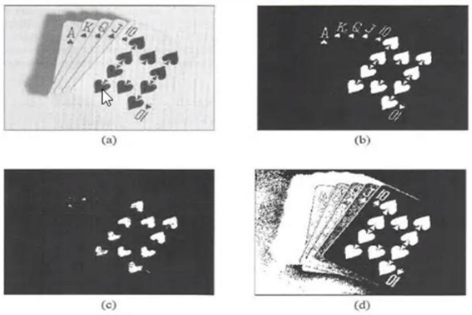

SANTI PRASAD MAITY is presenting
510817040 ISHAN_...
and 39 more
10:07 AM
You
Role of segmentation and challenges
• Segmentation plays an important role in image analysis. Success or failure of such image analysis to a great extent depends on success or failure of segmentation.
•Reliable and accurate segmentation of an image by complete automatic means is really challenging.
•Segmentation techniques can be classified broadly in two ways: contextual and non-contextual segmentation.
•Non-contextual segmentation: pixels are group together based on some global attributes such as gray value, color etc. Methods ignore the connectivity or relationship between features, point processing ignoring connectivity.
•Contextual techniques consider both attributes and closeness/relationship and connectivity.
IIEST, Shibpur
Image Processing and Pattern Recognition
Raise hand
Turn on captions
SANTI PRASAD MAITY is presenting

Importance of proper T value?

In thresholded image, non-zero value is assumed to be "interesting" and zero (0) value indicates no significance. However, if we like to highlight dark feature, then it can be written as follows:

$$g(x, y) = \begin{cases} 1, & f(x, y) \leqslant T, \\ 0, & f(x, y) > T, \end{cases}$$

The success or otherwise of thresholding depends on the selection of the proper "threshold" value. This can be seen from the images



Importance of accurate threshold selection. (a) Input image. (b) Correct choice of threshold ($T = 90$). (c) Threshold too low ($T = 40$). (d) Threshold too high ($T = 215$).

SANTI PRASAD MAITY is presenting

510817031 RAJESH_PAT... and 37 more

10:10 AM

Role of segm Press Esc to exit full screen allenges

• Segmentation plays an important role in image analysis. Success or failure of such image analysis to a great extent depends on success or failure of segmentation.

•Reliable and accurate segmentation of an image by complete automatic means is really challenging.

•Segmentation techniques can be classified broadly in two ways: contextual and non-contextual segmentation.

•Non-contextual segmentation: pixels are group together based on some global attributes such as gray value, color etc. Methods ignore the connectivity or relationship between features, point processing ignoring connectivity.

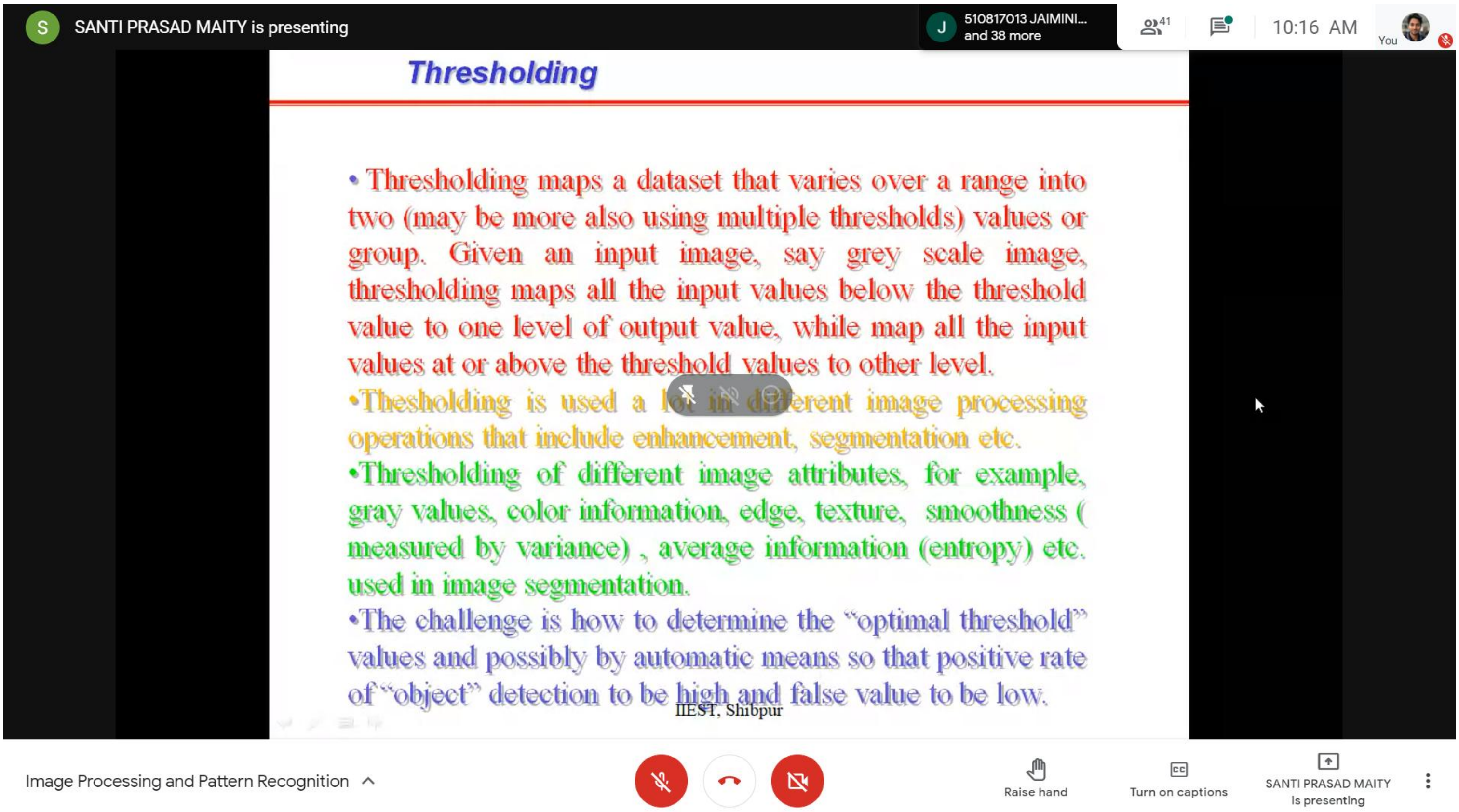
•Contextual techniques consider both attributes and closeness/relationship and connectivity.

Image Processing and Pattern Recognition

Raise hand

Turn on captions

SANTI PRASAD MAITY is presenting



S
SANTI PRASAD MAITY is presenting
J
510817013 JAIMINI...
and 38 more
41
10:16 AM
You
Thresholding
• Thresholding maps a dataset that varies over a range into two (may be more also using multiple thresholds) values or group. Given an input image, say grey scale image, thresholding maps all the input values below the threshold value to one level of output value, while map all the input values at or above the threshold values to other level.
•Thesholding is used a lot in different image processing operations that include enhancement, segmentation etc.
•Thresholding of different image attributes, for example, gray values, color information, edge, texture, smoothness (measured by variance) , average information (entropy) etc. used in image segmentation.
•The challenge is how to determine the "optimal threshold" values and possibly by automatic means so that positive rate of "object" detection to be high and false value to be low.
IIEST, Shibpur
Image Processing and Pattern Recognition
Raise hand
Turn on captions
SANTI PRASAD MAITY
is presenting


SANTI PRASAD MAITY is presenting
510817027 SHIVAM_BHASK...
and 39 more
10:20 AM
You
Non-contextual
Press Esc to exit full screen
Thresholding
•Given an input image, say grey scale image, image thresholding yields two values depending on whether some property measured from the image falls below (assigned to one level) or the property equals or exceeds (assigned to other level) the threshold. Thus thresholding creats a binary image, the generated output image depends on the property being thresholded.
•Revert back to our discussion on edge detection (image enhancement), thresholding implies whether the edge value i.e. the strength of the gradient value is below the threshold value (set the output value 0, considered not a proper' edge) otherwise set the value 1 or 255 (for display purpose) to indicate the proper edge.
IIEST, Shibpur
Image Processing and Pattern Recognition
Raise hand
Turn on captions
SANTI PRASAD MAITY
is presenting

Thresholding of pixel grey level

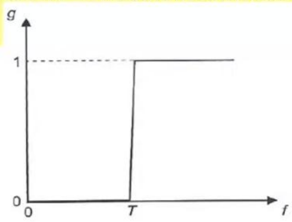
The most common form of image thresholding uses pixel grey value

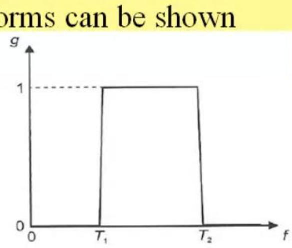
$$g(x, y) = \begin{cases} 0, & f(x, y) < T, \\ 1, & f(x, y) \geqslant T, \end{cases}$$

Here , T is the threshold, f(x,y) and g(x,y) are the grey value of input and the output image, respectively. A variation of above form is as follows

$$g(x, y) = \begin{cases} 0, & f(x, y) < T_1, \\ 1, & T_1 \leqslant f(x, y) \leqslant T_2 \\ 0, & f(x, y) > T_2. \end{cases}$$

Two thresholds T_1 and T_2 are used to define acceptable range of grey values. The above two mathematical forms can be shown pictorially as



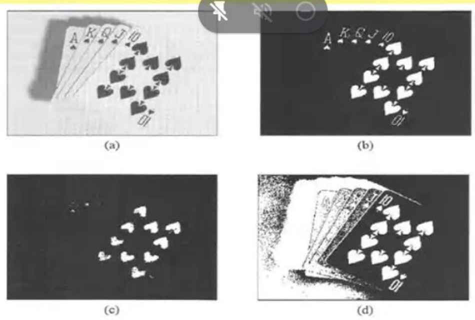


Importance of proper T value?

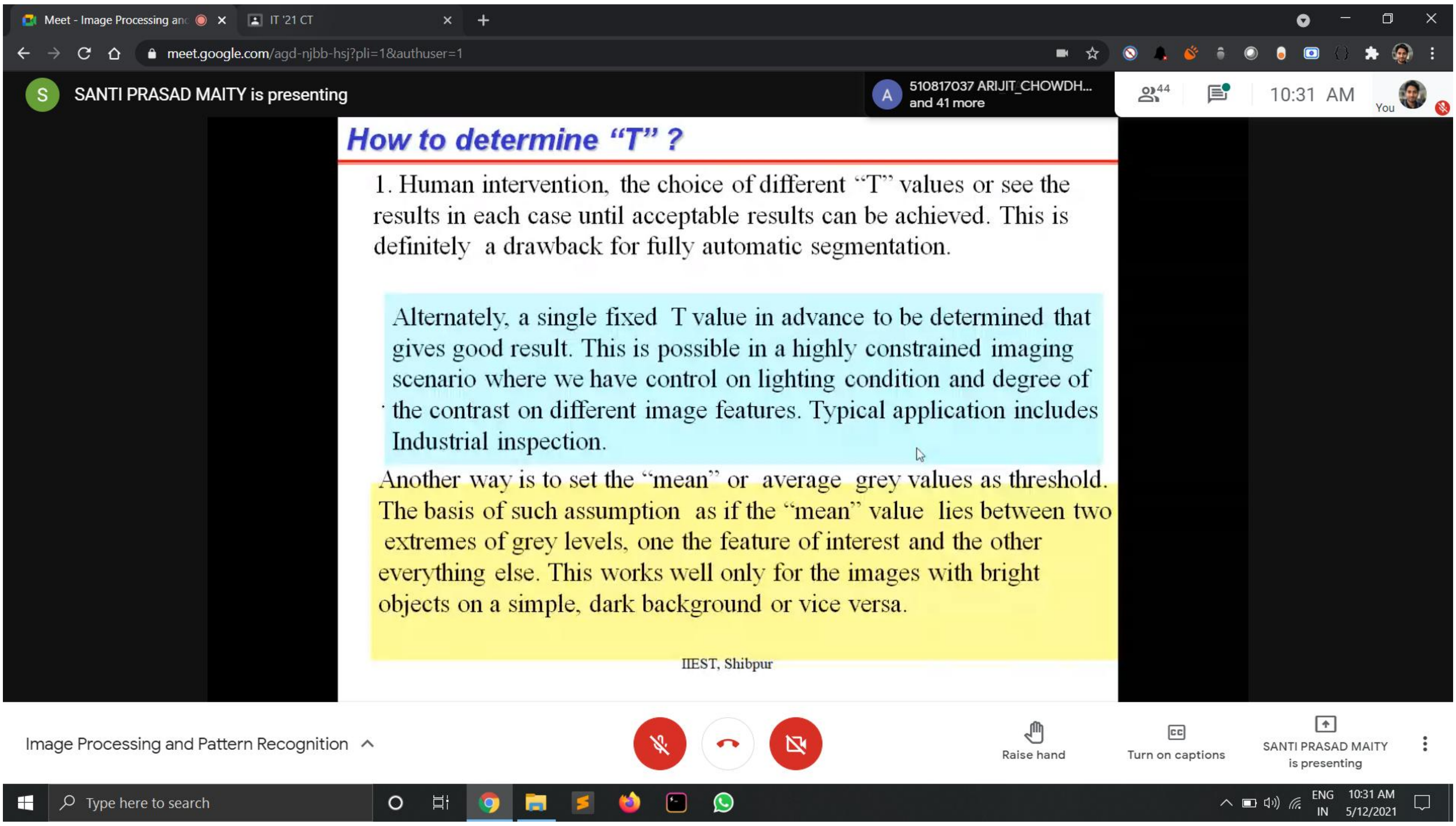
In thresholded image, non-zero value is assumed to be "interesting" and zero (0) value indicates no significance. However, if we like to highlight dark feature, then it can be written as follows:

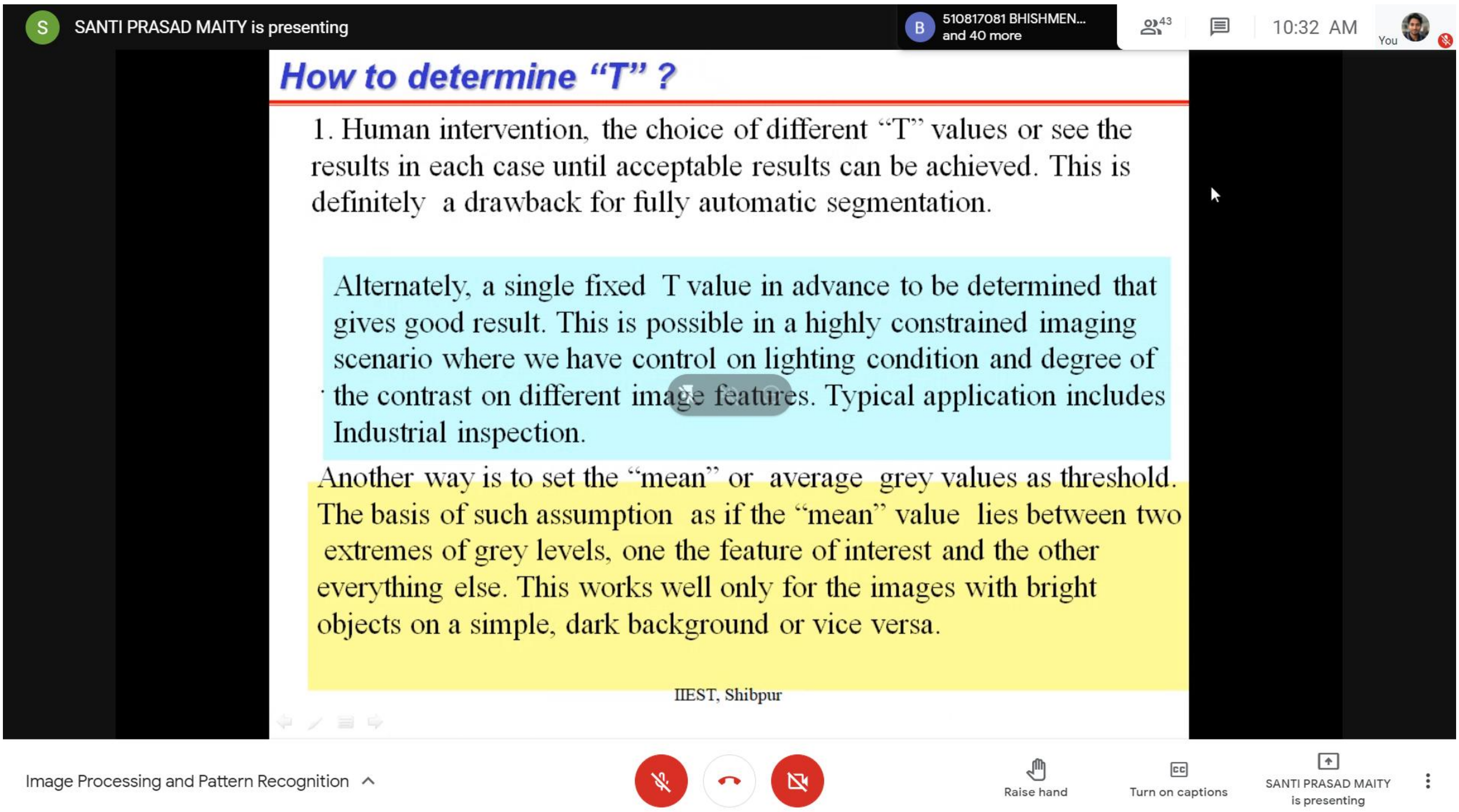
$$g(x, y) = \begin{cases} 1, & f(x, y) \leqslant T, \\ 0, & f(x, y) > T. \end{cases}$$

The success or otherwise of thresholding depends on the selection of the proper "threshold" value. This can be seen from the images



Importance of accurate threshold selection. (a) Input image. (b) Correct choice of threshold ($T = 90$). (c) Threshold too low ($T = 40$). (d) Threshold too high ($T = 215$).

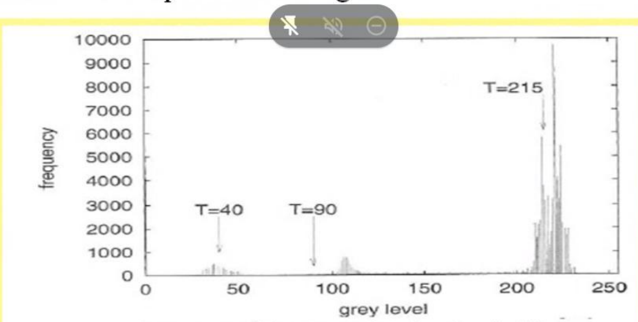

Meet - Image Processing and
IT '21 CT
meet.google.com/agd-njbb-hsj?pli=1&authuser=1
SANTI PRASAD MAITY is presenting
510817037 ARIJIT_CHOWDH...
and 41 more
10:31 AM
You
How to determine "T" ?
1. Human intervention, the choice of different "T" values or see the results in each case until acceptable results can be achieved. This is definitely a drawback for fully automatic segmentation.
Alternately, a single fixed T value in advance to be determined that gives good result. This is possible in a highly constrained imaging scenario where we have control on lighting condition and degree of the contrast on different image features. Typical application includes Industrial inspection.
Another way is to set the "mean" or average grey values as threshold. The basis of such assumption as if the "mean" value lies between two extremes of grey levels, one the feature of interest and the other everything else. This works well only for the images with bright objects on a simple, dark background or vice versa.
IIEST, Shibpur
Image Processing and Pattern Recognition
Raise hand
Turn on captions
SANTI PRASAD MAITY
is presenting
Type here to search
ENG
IN
10:31 AM
5/12/2021


SANTI PRASAD MAITY is presenting
510817081 BHISHMEN...
and 40 more
10:32 AM
You
How to determine "T" ?
1. Human intervention, the choice of different "T" values or see the results in each case until acceptable results can be achieved. This is definitely a drawback for fully automatic segmentation.
Alternately, a single fixed T value in advance to be determined that gives good result. This is possible in a highly constrained imaging scenario where we have control on lighting condition and degree of the contrast on different image features. Typical application includes Industrial inspection.
Another way is to set the "mean" or average grey values as threshold. The basis of such assumption as if the "mean" value lies between two extremes of grey levels, one the feature of interest and the other everything else. This works well only for the images with bright objects on a simple, dark background or vice versa.
IIEST, Shibpur
Image Processing and Pattern Recognition
Raise hand
Turn on captions
SANTI PRASAD MAITY
is presenting

Geometric Transformation

* Choosing T value so that a fixed proportion of pixel values are detected. This works well if we know in advance/a priori that a certain proportion of pixels associated with feature of interest. Typical applications include OCR or industrial inspection.

Histogram of the images may also be used for "threshold" selection. This is based on the assumption that different features give rise to different peaks in histogram.



Histogram of the image showing thresholds that are too low (40), correct (90) and too high (215).

SANTI PRASAD MAITY is presenting

510817072 KUMAR_SOU... and 33 more

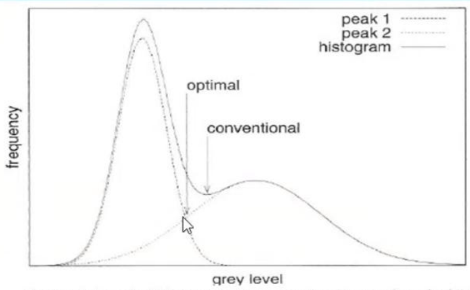
10:41 AM

Selection of T value

Press Esc to exit full screen

* Often histogram peaks corresponding to two features overlap. The degree of overlap depends on peak separation and peak width. In such case threshold may be chosen in the valley between the two peaks. Such threshold (not the optimal) leads to false detection or rejection of some pixels. The optimal threshold minimizes false detection or rejection of pixels.

The optimal threshold, in general, does not occur at the lowest point in the valley between the two overlapping peaks.



Optimal threshold between two overlapping peaks of a histogram.

Image Processing and Pattern Recognition

Raise hand

Turn on captions

SANTI PRASAD MAITY is presenting

Iterative method of automatic threshold selection

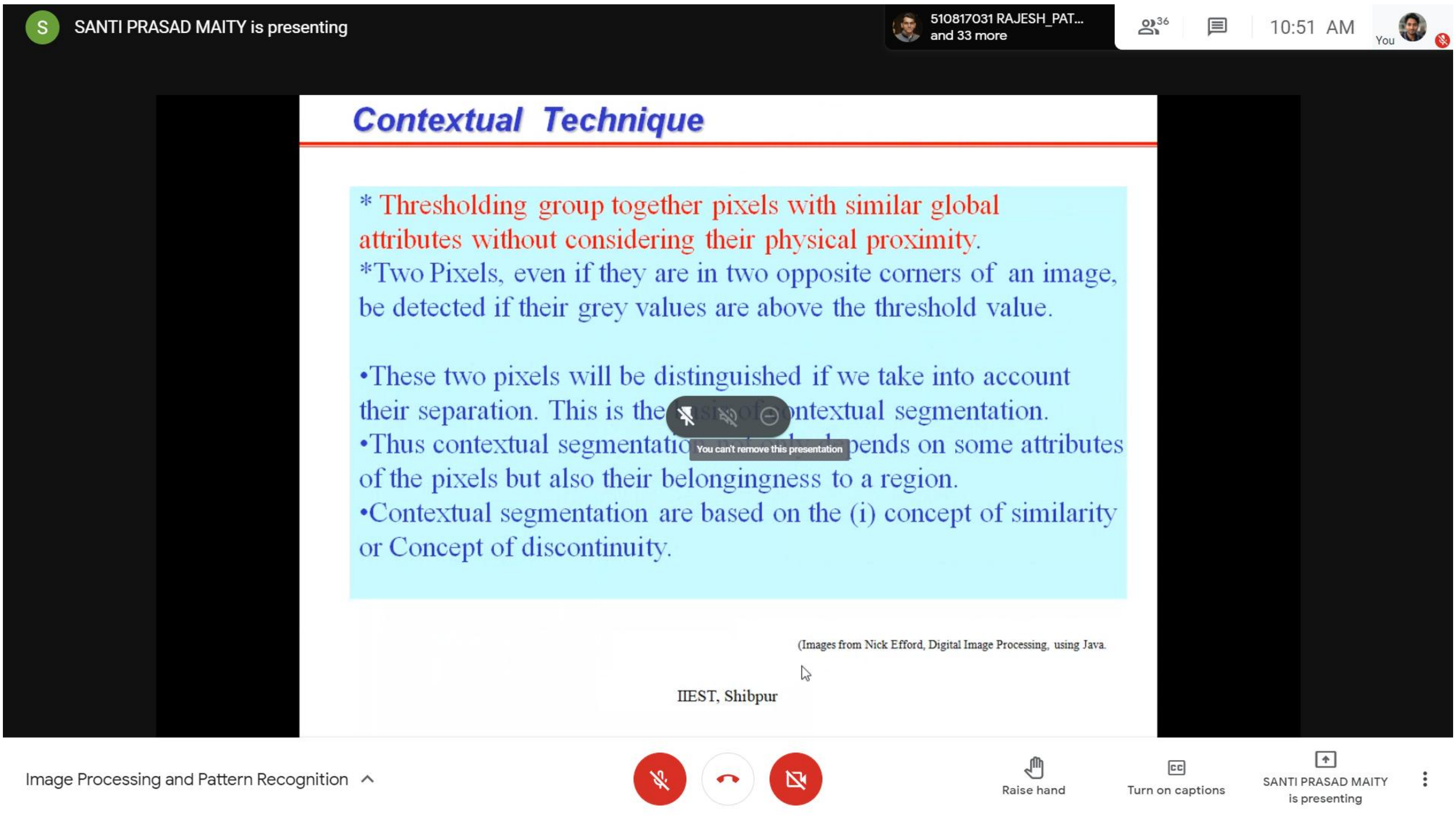
One way to determine is iterative method that starts with an initial guess as threshold and then refines the estimate by successive passes through the images.
* Initial guess may be the mean grey level of the image or the average value of the mean of the corner pixels and the mean of the all other pixel grey values. Assumption considers that cornel pixels represent the background rather than objects.

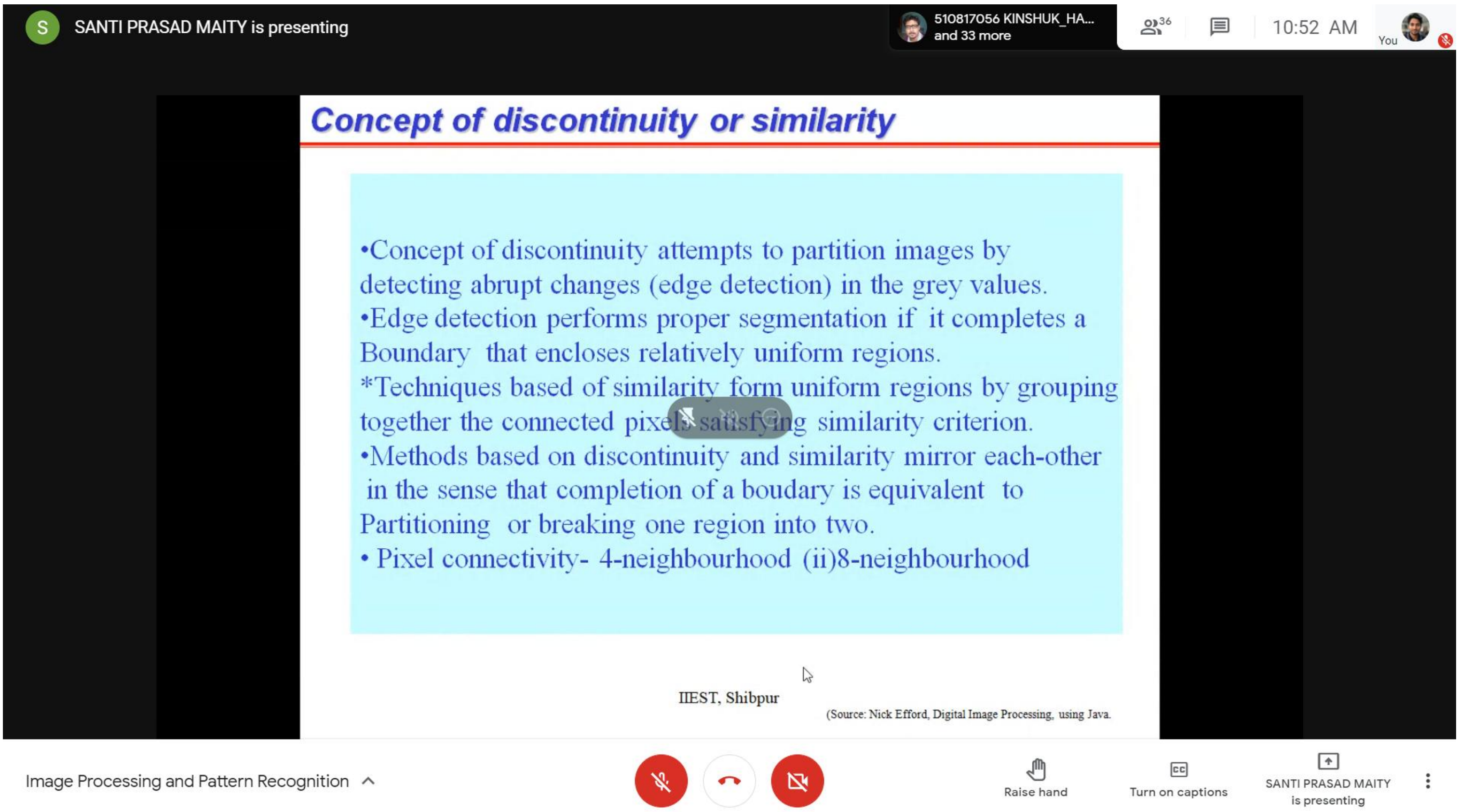
Four to ten iterations converge (algo. For IP lab/assignm ent)

Iterative threshold determination.

Compute μ_1, the mean grey level of the corner pixels
Compute μ_2, the mean grey level of all other pixels
$T_{\text{old}} = 0$
$T_{\text{new}} = (\mu_1 + \mu_2)/2$
while $T_{\text{new}} \neq T_{\text{old}}$ **do**
  μ_1 = mean grey level of pixels for which $f(x, y) < T_{\text{new}}$
  μ_2 = mean grey level of pixels for which $f(x, y) \geqslant T_{\text{new}}$
  $T_{\text{old}} = T_{\text{new}}$
  $T_{\text{new}} = (\mu_1 + \mu_2)/2$
end while

(Source: Nick Efford, Digital Image Processing, using Java.


SANTI PRASAD MAITY is presenting
510817031 RAJESH_PAT...
and 33 more
10:51 AM
You
Contextual Technique
* Thresholding group together pixels with similar global attributes without considering their physical proximity.
*Two Pixels, even if they are in two opposite corners of an image, be detected if their grey values are above the threshold value.
•These two pixels will be distinguished if we take into account their separation. This is the ntextual segmentation.
•Thus contextual segmentatio pends on some attributes of the pixels but also their belongingness to a region.
•Contextual segmentation are based on the (i) concept of similarity or Concept of discontinuity.
You can't remove this presentation
(Images from Nick Efford, Digital Image Processing, using Java.
IIEST, Shibpur
Image Processing and Pattern Recognition
Raise hand
Turn on captions
SANTI PRASAD MAITY
is presenting


SANTI PRASAD MAITY is presenting
510817056 KINSHUK_HA...
and 33 more
10:52 AM
You
Concept of discontinuity or similarity
•Concept of discontinuity attempts to partition images by detecting abrupt changes (edge detection) in the grey values.
•Edge detection performs proper segmentation if it completes a Boundary that encloses relatively uniform regions.
*Techniques based of similarity form uniform regions by grouping together the connected pixels satisfying similarity criterion.
•Methods based on discontinuity and similarity mirror each-other in the sense that completion of a boudary is equivalent to Partitioning or breaking one region into two.
• Pixel connectivity- 4-neighbourhood (ii)8-neighbourhood
IIEST, Shibpur
(Source: Nick Efford, Digital Image Processing, using Java.
Image Processing and Pattern Recognition
Raise hand
Turn on captions
SANTI PRASAD MAITY
is presenting

Region Similarity

The similarity of a region of pixels may be characterized by uniformity predicate, a logical statement that is true if neighboring pixels are very similar in some attribute, say, grey values, color. etc.

Assign-
ment of IP lab

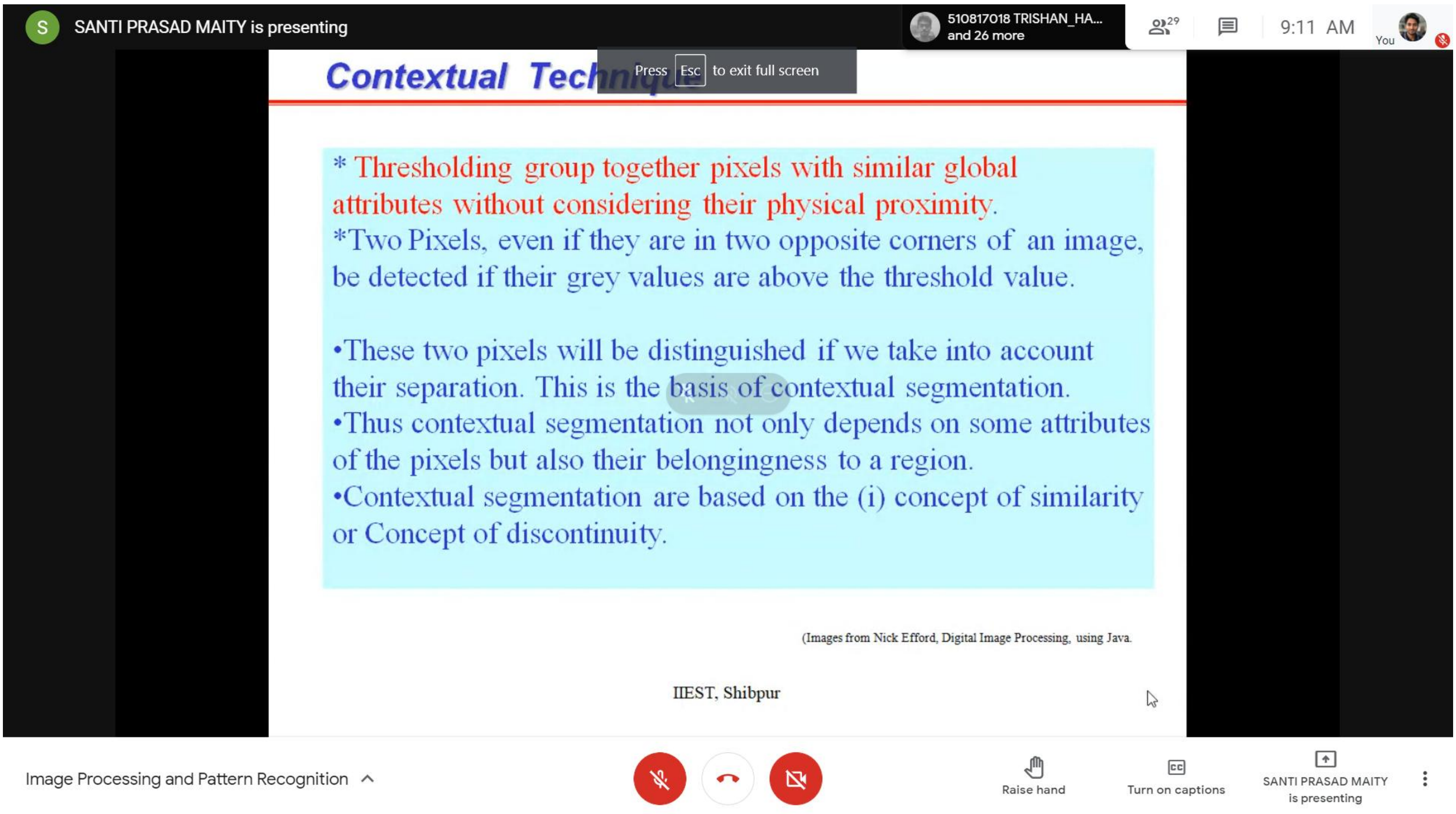
$$P(R) = \begin{cases} \text{TRUE} & \text{if } |f(j,k) - f(m,n)| \leqslant \Delta \\ \text{FALSE} & \text{otherwise,} \end{cases}$$

Where (j, k) and (m, n) are the coordinates of the neighboring pixels and f is their grey values. Region R to be uniform if and only if the two neighboring pixels differ in grey values by no more than Δ

A misconception may be that the grey level variation within the regions may be within the bound specified. However, small change from neighbor to neighbor can accumulate and exceeds the bound between the pixels in two opposite sides of region.

IIEST, Shibpur

(Images from Rafael C. Gonzalez and Richard E. Wood, Digital Image Processing, 2nd Edition.


SANTI PRASAD MAITY is presenting
510817018 TRISHAN_HA...
and 26 more
9:11 AM
You
Contextual Technique
Press Esc to exit full screen
* Thresholding group together pixels with similar global attributes without considering their physical proximity.
*Two Pixels, even if they are in two opposite corners of an image, be detected if their grey values are above the threshold value.
•These two pixels will be distinguished if we take into account their separation. This is the basis of contextual segmentation.
•Thus contextual segmentation not only depends on some attributes of the pixels but also their belongingness to a region.
•Contextual segmentation are based on the (i) concept of similarity or Concept of discontinuity.
(Images from Nick Efford, Digital Image Processing, using Java.
IIEST, Shibpur
Image Processing and Pattern Recognition
Raise hand
Turn on captions
SANTI PRASAD MAITY
is presenting

Region Similarity

A similar predicate may be as follows

$$P(R) = \begin{cases} \text{TRUE} & \text{if } |f(j,k) - \mu_R| \leqslant \Delta \\ \text{FALSE} & \text{otherwise,} \end{cases}$$

Assignment of IP lab

where f(j, k) is the grey value at pixel (j,k), μ_R is the mean grey value of all pixels in the region R except the pixel at (j,k).

* The value of Δ in both the uniformity predicates has an important role in segmentation results. Hence, its proper value to be selected. Its optimal value can be determined by solving some criterion of P(R), may be characterized by variance of R or by average information content which is a measure of entropy for the region R.** (a research or project problem)

(Images from Rafael C. Gonzalez and Richard E. Wood, Digital Image Processing, 2nd Edition.

Region Growing: Typical Example

A typical example of region grown up with 8-connectivity and Δ =3 is shown below

* Region developed is sensitive to connectivity (4 or 8) and value of uniformity threshold (see figure of dice, two results with Δ value 41, 42)

(a)

0	0	5	6	7
1	1	5	8	7
0	1	6	7	7
2	0	7	6	6
0	1	5	6	5

(b)

0	0	5	6	7
1	1	5	8	7
0	1	6	7	7
2	0	[illegible]	6	6
0	1	5	6	5

(c)

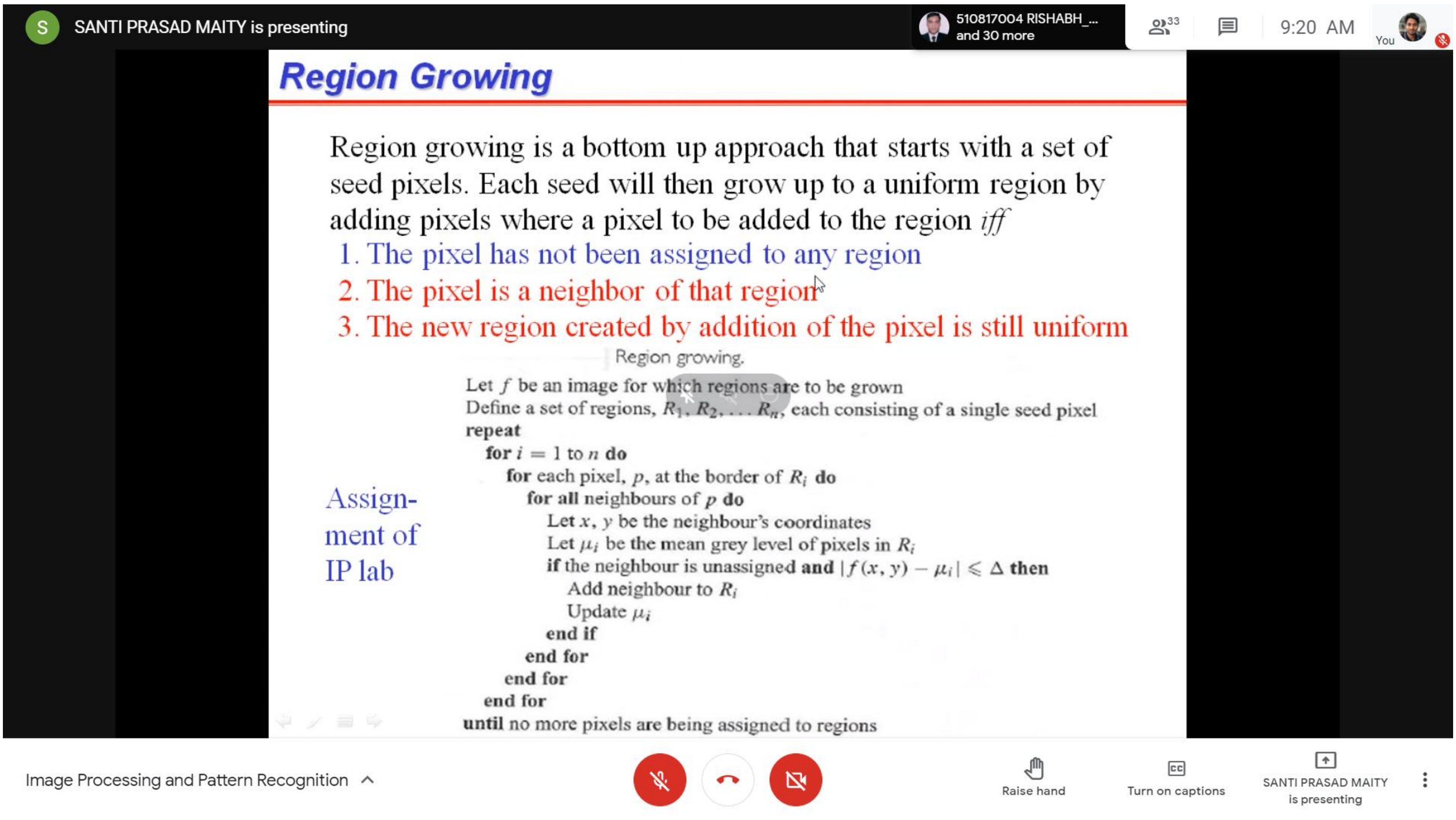
0	0	5	6	7
1	1	5	8	7
0	1	6	7	7
2	0	7	6	6
0	1	5	6	5

Region growing. (a) Seed pixels. (b) First iteration. (c) Final iteration.

(a) (b)

Sensitivity of Δ. (a) Region grown with $\Delta = 41$. (b) Region grown with $\Delta = 42$.

(Images from Nick Efford, Digital Image Processing, using Java.


S
SANTI PRASAD MAITY is presenting
510817004 RISHABH_...
and 30 more
33
9:20 AM
You
Region Growing
Region growing is a bottom up approach that starts with a set of seed pixels. Each seed will then grow up to a uniform region by adding pixels where a pixel to be added to the region iff
1. The pixel has not been assigned to any region
2. The pixel is a neighbor of that region
3. The new region created by addition of the pixel is still uniform
Assign-ment of IP lab
Region growing.
Let f be an image for which regions are to be grown
Define a set of regions, R1, R2, ... Rn, each consisting of a single seed pixel
repeat
for i = 1 to n do
for each pixel, p, at the border of Ri do
for all neighbours of p do
Let x, y be the neighbour's coordinates
Let μi be the mean grey level of pixels in Ri
if the neighbour is unassigned and |f(x, y) − μi| ⩽ Δ then
Add neighbour to Ri
Update μi
end if
end for
end for
end for
until no more pixels are being assigned to regions
Image Processing and Pattern Recognition
Raise hand
Turn on captions
SANTI PRASAD MAITY
is presenting

Limitations of Region Growing Techniques

•Region growing (RG)technique is not a stable operation. Results differ with the (i) choice of connectivity (4 or 8-connectivity) and (ii) Choice of value of Δ

•See (previous slide) difference in segmentation results for dice image wit Δ value 41 and 42, respectively.

•Complete segmentation must satisfy a number of criteria

(i) All pixels must be assigned to regions
(ii) Each pixel must belong to a single region
(iii) Each region must be a connected set of pixels
(iv) Each region must be uniform
(v) Any merged pairs of adjacent regions must be non-uniform

Limitation of RG: RG satisfies third and fourth criteria.

RG does not satisfy first and second one, as number of seeds specified may not sufficient to create region for every pixel.

RG also does not satisfy the fifth criterion

IIEST, Shibpur

(Images from Nick Efford, Digital Image Processing, using Java.

Spilt and merge method

* A complete segmentation is possible through top-down (remember region growing was a bottom-up approach) approach in which the whole image is considered as a single region.
* Hence, uniformity predicate will not be met for the whole image which needs division of the image into sub-regions.

- The Sub-regions are split or merged iteratively until uniformity predicate is met or until desired number of regions are formed.
- If image be a square image, then it needs division in quadrant, sub-quadrant so on until, for any region R, P(R) is True
- The use of splitting method alone may lead to many small, uniform adjacent regions with identical properties. Hence, merging of regions, say Ri and Rj, is needed so that P(Ri U Rj) is true (U indicates union operation).

IIEST, Shibpur

(Images from Nick Efford, Digital Image Processing, using Java.

SANTI PRASAD MAITY is presenting

510817009 SUDIPTI_MAN... and 33 more

9:33 AM

Segmentation using other image properties

- Apart from grey value or color value, edge, texture, depth, motion, entropy etc are used for thresholding to accomplish image segmentation. Even there are variations in various thresholdings

Global thresholding- When 'Thr' depends only on the grey level value of image and 'Thr' is solely related to the properties of pixel in the image.
Otsu method, entropy based thresholding (see literature)
Local thresholding- If threshold *T depends on both f (x, y) and p(x, y).*
Statistical thresholding, 2-D entropy-based thresholding, histogram transformation thresholding (see literature)

Dynamic thresholding- If, in an image, there are several objects taking up different gray level regions, the image should be partitioned with various dynamic thresholds(*T1,T2 …Tn), depending on f (x, y) , p(x, y)and the spatial coordinates x and y,*
thresholding image, Watershed, interpolatory thresholding (see literature)

Image Processing and Pattern Recognition

Raise hand

Turn on captions

SANTI PRASAD MAITY is presenting

Segmentation using other image properties

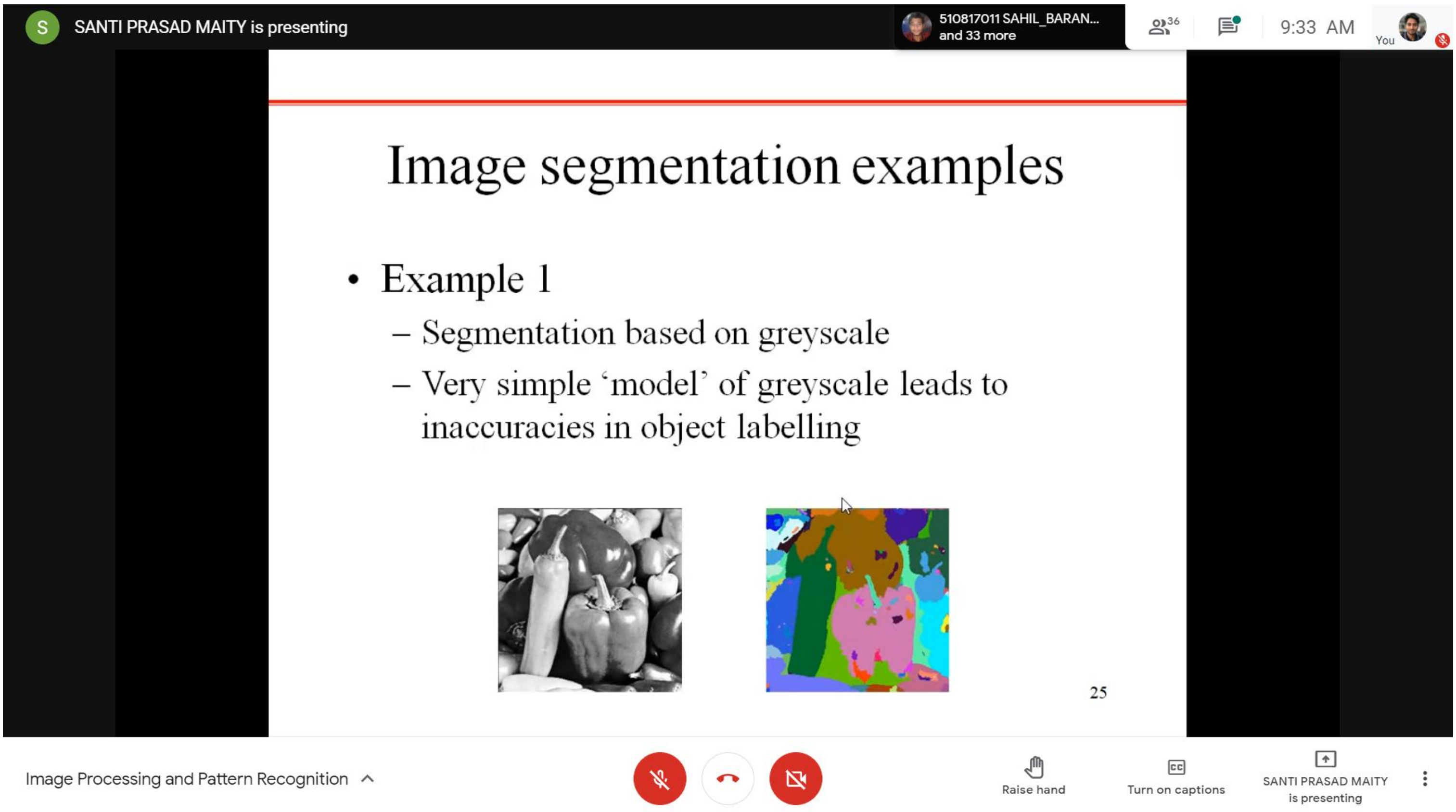
•Apart from grey value or color value, edge, texture, depth, motion, entropy etc are used for thresholding to accomplish image segmentation. Even there are variations in various thresholdings

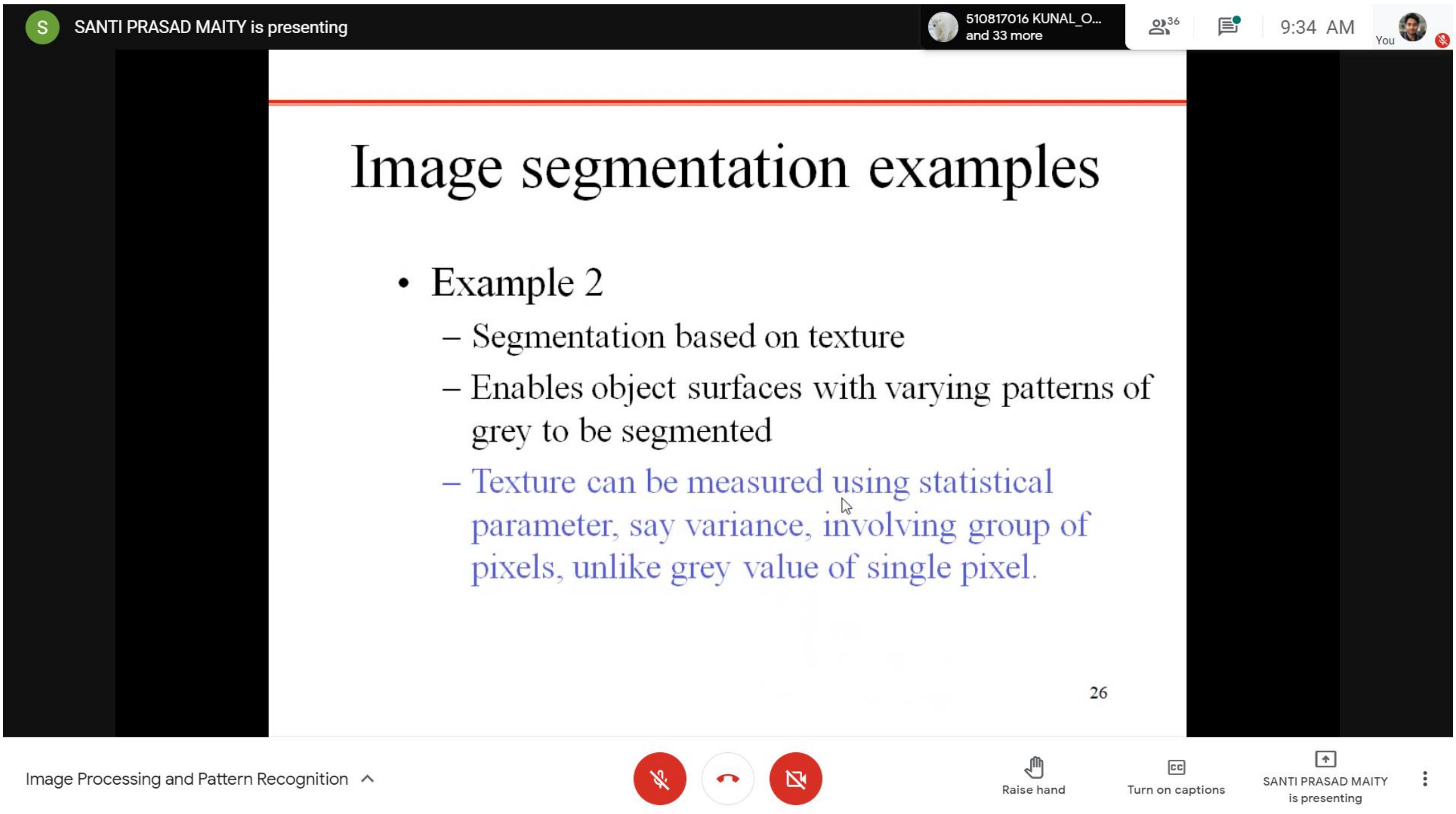
Global thresholding- When 'Thr' depends only on the grey level value of image and 'Thr' is solely related to the properties of pixel in the image.

Otsu method, entropy based thresholding (see literature)

Local thresholding- If threshold *T depends on both f (x, y) and p(x, y).*
Statistical thresholding, 2-D entropy-based thresholding, histogram transformation thresholding (see literature)

Dynamic thresholding- If, in an image, there are several objects taking up different gray level regions, the image should be partitioned with various dynamic thresholds(*T1,T2 ...Tn), depending on f (x, y) , p(x, y)and the spatial coordinates x and y,*
thresholding image, Watershed, interpolatory thresholding (see literature)


S
SANTI PRASAD MAITY is presenting
510817011 SAHIL_BARAN...
and 33 more
36
9:33 AM
You
Image segmentation examples
• Example 1
– Segmentation based on greyscale
– Very simple 'model' of greyscale leads to inaccuracies in object labelling
25
Image Processing and Pattern Recognition
Raise hand
Turn on captions
SANTI PRASAD MAITY is presenting


SANTI PRASAD MAITY is presenting
510817016 KUNAL_O...
and 33 more
9:34 AM
You
Image segmentation examples
• Example 2
– Segmentation based on texture
– Enables object surfaces with varying patterns of grey to be segmented
– Texture can be measured using statistical parameter, say variance, involving group of pixels, unlike grey value of single pixel.
26
Image Processing and Pattern Recognition
Raise hand
Turn on captions
SANTI PRASAD MAITY
is presenting

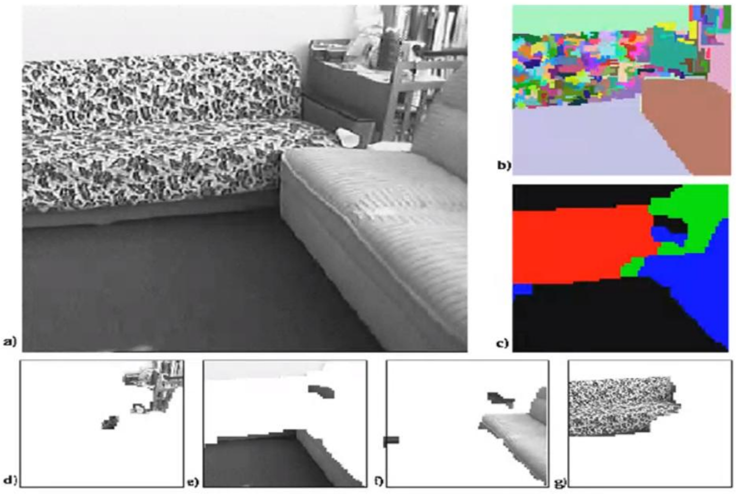
SANTI PRASAD MAITY is presenting

510817020 ANINDYA_KU...
and 33 more

9:34 AM

You

Image segmentation examples



27

Image Processing and Pattern Recognition

Raise hand

Turn on captions

SANTI PRASAD MAITY
is presenting

S SANTI PRASAD MAITY is presenting

510817031 RAJESH_PAT... and 33 more

36

9:34 AM

You

Image segmentation examples

- Example 3
  - Segmentation based on motion
  - The main difficulty of motion segmentation is that an intermediate step is required to (either implicitly or explicitly) estimate an *optical flow field*
  - The segmentation must be based on this estimate and not, in general, the true flow

Image Processing and Pattern Recognition

Raise hand

Turn on captions

SANTI PRASAD MAITY is presenting

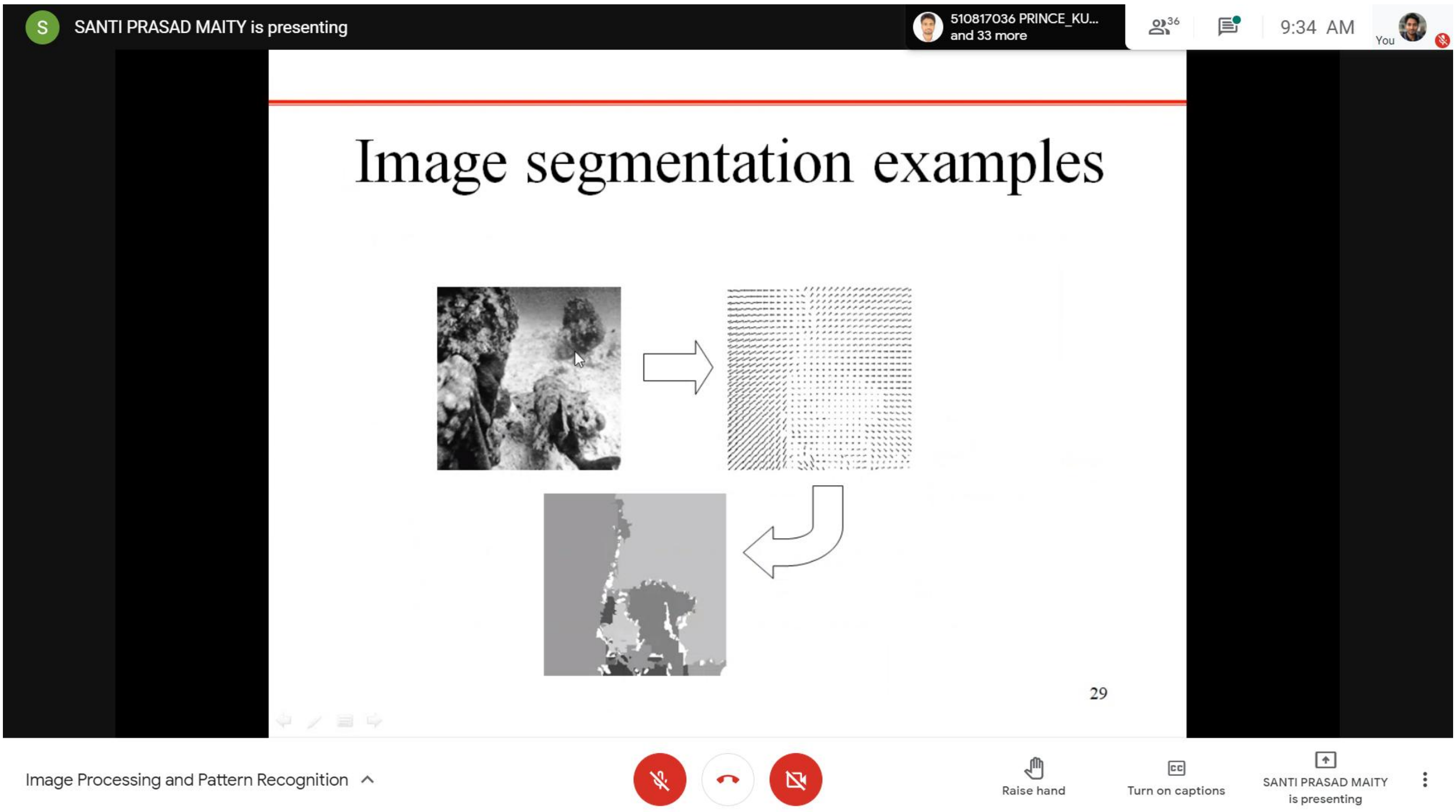

SANTI PRASAD MAITY is presenting
510817036 PRINCE_KU...
and 33 more
36
9:34 AM
You
Image segmentation examples
29
Image Processing and Pattern Recognition
Raise hand
Turn on captions
SANTI PRASAD MAITY
is presenting

Image segmentation examples

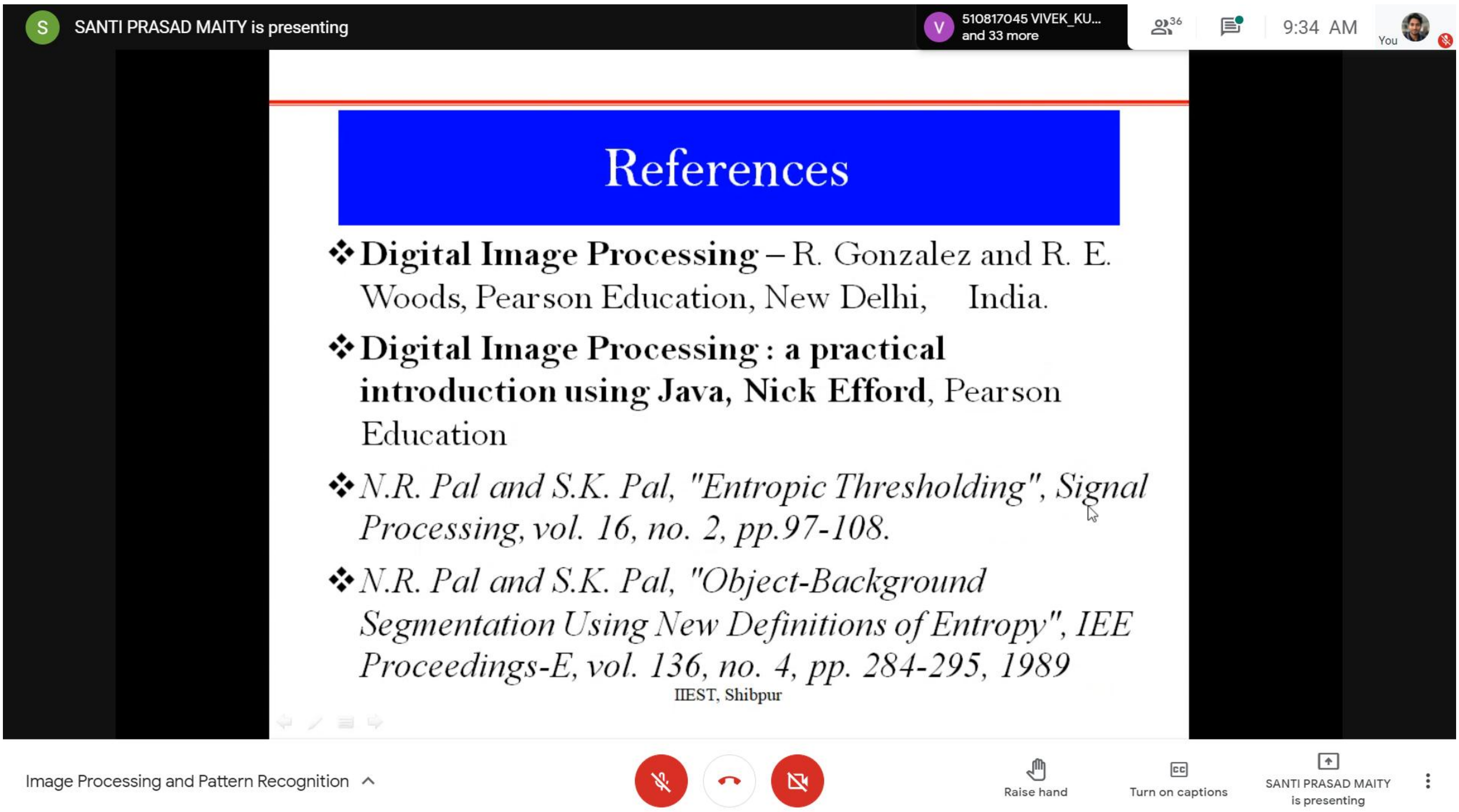
- Example 4
  - Segmentation based on depth
  - This example shows a range image, obtained with a laser range finder
  - A segmentation based on the range (the object distance from the sensor) is useful in guiding mobile robots

Image segmentation examples

Original image

Range image

Segmented image


SANTI PRASAD MAITY is presenting
510817045 VIVEK_KU...
and 33 more
9:34 AM
You
References
❖Digital Image Processing – R. Gonzalez and R. E. Woods, Pearson Education, New Delhi, India.
❖Digital Image Processing : a practical introduction using Java, Nick Efford, Pearson Education
❖N.R. Pal and S.K. Pal, "Entropic Thresholding", Signal Processing, vol. 16, no. 2, pp.97-108.
❖N.R. Pal and S.K. Pal, "Object-Background Segmentation Using New Definitions of Entropy", IEE Proceedings-E, vol. 136, no. 4, pp. 284-295, 1989
IIEST, Shibpur
Image Processing and Pattern Recognition
Raise hand
Turn on captions
SANTI PRASAD MAITY
is presenting