Image is the Collection of pixel
Video " " " " Image

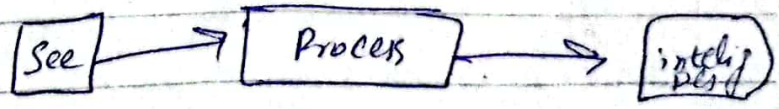
- Image Processing is a method to perform some operations on an image, in order to get an enhanced image or to extract some useful information with the help of algorithm.

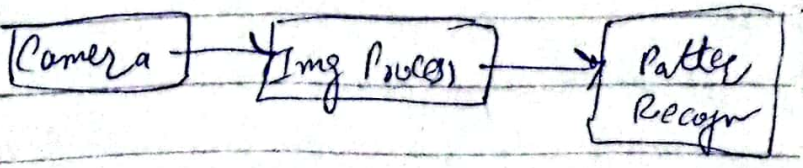
- Computer Vision is the process to learn the AI model.
  - is an interdisciplinary Scientific field that deals with how Computer can gain high-level understanding from digital images or videos.
  - Images to model

- Open CV is an open Source Library for the implementation of image Processing and Computer Vision.
  - numpy library handles the Matrix.

- Gray scale image
  - Single channel

Color Image
3-channel => RGB Color





- Application
  - Cancer detection
  - Still Training
  - Covid-19 dignosis.
  - Mask detection

Imag => is 2D
Video => is 3D

Gray Img: binary (1)

Img Segmentation means the partioning of an image to Connect homogeneous region

Feature extraction (Representation & description)

* Digital Img Processing
  2D array of quantized intensity fellows or may be 2D array of numbers
  ↳ If we want to improve image quality we will change the quantized intensity.
  ↳ def: Its means processing of digital img & manipul on digital hardware usually a Computer.

Object → observe → Digitize → Store → Process → Refresh/stor

Img representation: $f(x,y)$ gives the intensity at posit (x,

Digital Img def:
Value = $\infty(x, y, z, \lambda, t)$
- x: Co-ordinat of Img
- y: Co-ordinat of Img
- z: Depth information
- λ: Color Infor
- t: Frame. → Temporal information

Image Types

- RGB
- GRAYSCALE
- BLACK AND WHITE

Binary Image

- Efficient in terms of Storage (1, 0)

GRAYSCALE Img

- The range is usually a power of 2
  256 levels are Sufficient for most application

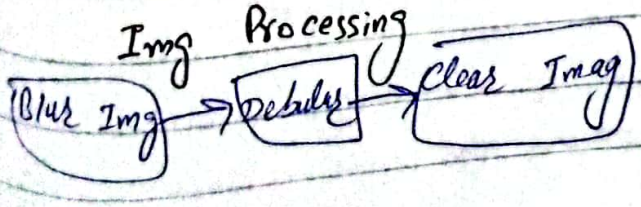
Index Image

- The image has an associated Color map, which is simply a list of all the Colors used in the image.
  Gif, PNG formats, etc.

* Types Of Digital Images.
  - Binary Image - 1 bit / Pixel
  - Grayscal " - 8 bit "
  - True Color or RGB image - 24 bit / Pixel.
  - Index img - 8 bits / Pixel.

Domain means scaling :

Img Processing



* Imag Processing Steps:

1. Acquisition, Sampling / Quantization / Compression.
2. Img enhancement (sub) and restoration → objective
   - Improve the visual quality of Img. like brightness
3. Img Segmentation
   - The Partition of img into Connected homogeneous region
4. Feature Extraction
5. object recognition
6. Img Interpretation

Img Acquisition

- An analog img is obtained by scanning the sensor output.
- Some of the modern scanning devices such as a CCD Camera. Contains an array of Photo detectors.

* Sampling / Quantization / Compression

- Raw vid is very bulky

Ex: The transformation of high-definition uncompressed digital vid at 1024 x 768, 24 bit / pixel 25 frames requires 472 MBps.

- We have to Compress the raw data to store and transmit

Img Enhacent

- Improves the quality of an img.
  - → Enhancing the Contrast, Sharping the edges, removing noise.

Img restoration:

Some as img enhancement, but you have additional information concerning the quality degradation. Ex: removing motion blur of an moving obj.

Reconstruction: reconstructions from projections.

↳ Constructing 3d data from 2D projections in Computer tomography.

↳ Radon transform.

Img representation: using feature.

↳ low level representations using color, texture, shape, motion etc.

high level feature for recognitions.

Feature extraction: like edges.

v. important to detect the boundaries of the object.

Segmentation: Partioning of an image into connected homogeneous regions.

Homogenity defined in terms:

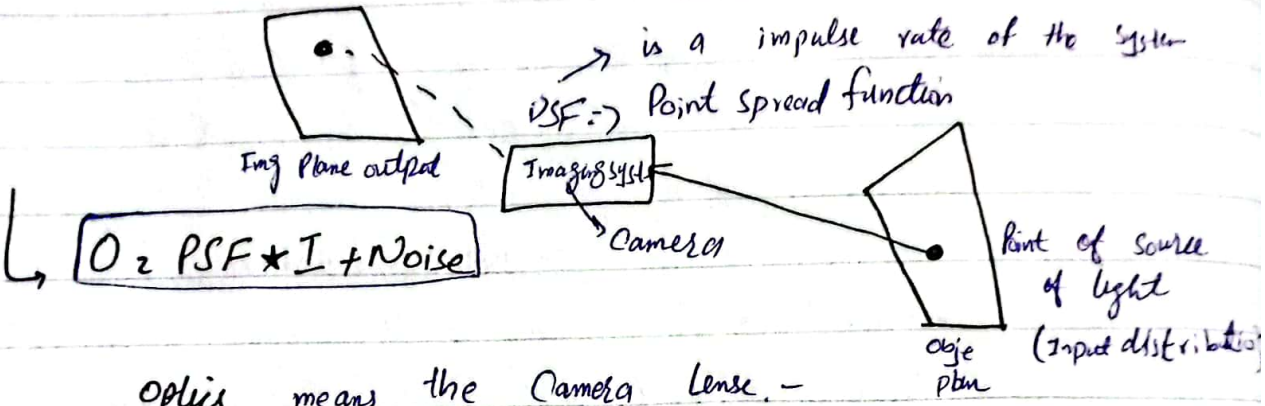
Gray scale.

Color

Texture

Radiometery : means the measurement of light

Img formation : refers to the process of Creating an img from a scene or object in the real world.



optics means the Camera Lense. - and is am getting the img in the Sensor.

Sensors Converts the optical photons into electrical Signal

↳ Radiance
↳ used to measure the distribution of light in space

↳ Irradiance
↳ How much light is arriving at a surface?

To Sum up

Radiance (L) Energy carried by a ray ($W\ m^{-2}\ sr^{-1}$) → unit

Irradiance (E) Energy arriving at a surface.
($W\ m^{-1}$) → unit.

Light at Surfaces.

Many effects when light strikes a surface:
Could be:
↳ Absorbed, transmitted (Skin), reflected (mirror), (Scattered) (milk).

BRDF

Img Formation:
↳ In modeling any img formation process, geometric primitives and transformations are crucial to project 3-D geometric features, ~~img formation~~ into 2-D features.

↳ It needs to know the lightning of the enviremont, Camera optics, sensor properties etc. Therefore while talking about img formation in Computer Vision.

↳ Means ke hmne obj ko clear dekhana he.
↳ Quality of img proper hona Chahiye.

1- Photometric Img Formation
Surface matter
Color "
Light Source matter.

Humain vs Computer Vision System

The humain vision system is complex. Firstly the eye is perceive the env / object different organs & nerves send the signal to the brain. Finally, the brain recognize it & acts accordingly.

↳ For a Computer vision system, the camera takes the image to perceive the env, & the camera input is further fed to the Computer vision models recognize the img & show the output.

Color Models
RGB
CMYK
HSL
HSV

* Basic Models of Reflection
  ↳ Specular: Light bounces off at the incident angel.
    ↳ Eg: Mirror.

* Diffuse: Light scatters in all directions
  ↳ Eg: Cloth, Rough wood

Most Surfaces have both Specular & diffuse

Specularity => Spot where Specular reflection dominates (typically reflects light source)

* Dynamic Range & Camera Response

Typically scenes have a huge dynamic range.

Camera Response is roughly linear in the mid range (15 to 240) but not linear at extremes.

* Conversion from RGB to HSV

* Resolution : is a sampling Parameter, defined in dpi or equivalent measures of spatial pixel density and its standard value for recent screen technologies is 72 dpi

* Color Correction

Multiply R, G & B value by seperate Constant

$$\begin{bmatrix} r \\ g \\ b \end{bmatrix} = \begin{bmatrix} a_r & 0 & 0 \\ 0 & a_g & 0 \\ 0 & 0 & a_b \end{bmatrix} \begin{bmatrix} r \\ g \\ b \end{bmatrix}$$

* Histogram: of an image provides the frequency of the brightness (intensity) value in the img

↳ Histogram captures the distribution of gray levels in the img

Find histogram:
- PDF
- CDF
- Mean
- Variance

* Neighbourhood

Mask is always in rectangle and any nxm.

Shape

→ Neighbourhood

Simply operations on a large neighbourhood of pixels than point

Any size rectangle and any shape filter is possible

Neighbourhood operations

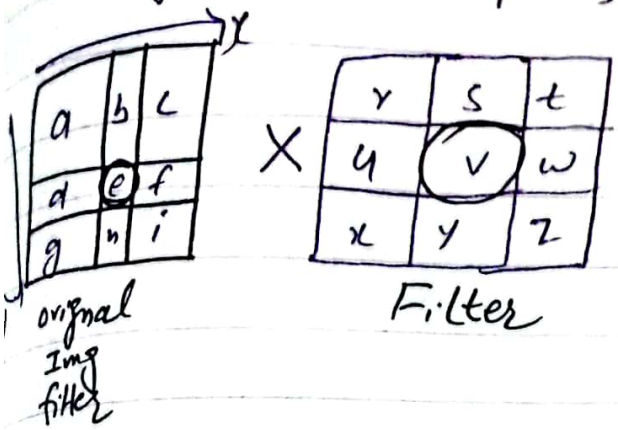
Min value in neighbour hood

Max " " " "

Median " " " "

Spatial Filtering used

↳ Img Smooth, Sharpness, remove noise



$e_{processed} = vxe + vxa + sxb + txc + uxd + wxf + xxg + yxh + zxi$

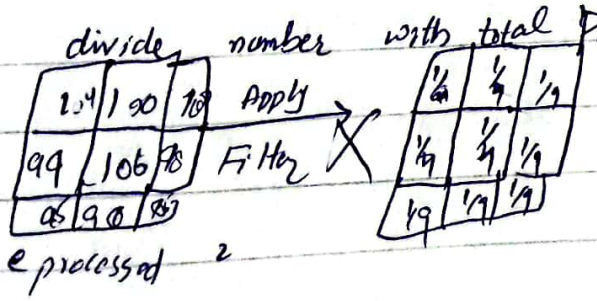
$e_{process}$ = Selected pixel with neighbour + Filter

generated (x,y) = T . $f(x,y)$

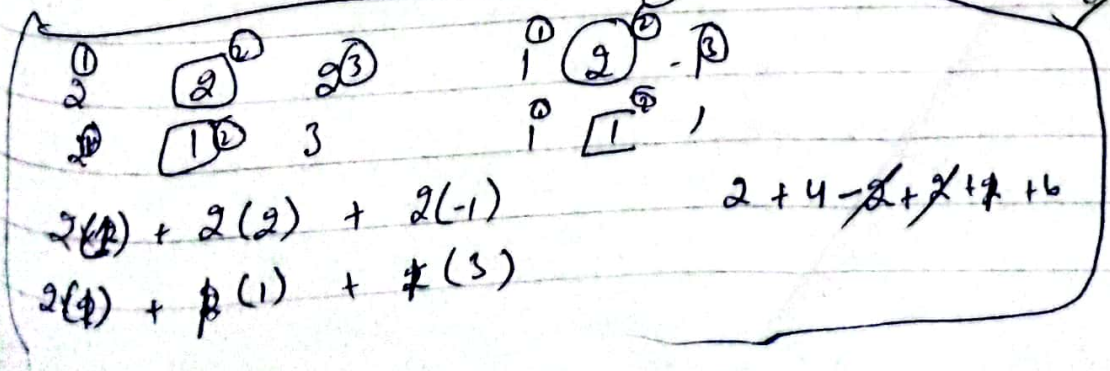
Filtered img — mask — original Img.

Smoothing Filter

↳ Easest method to decreas the pixel value



$1/9 \times 106 + 1/9 \times 104 \times 1/9 \times 99 \times 1/9$



Img Enhacement

FOURIER TRANSFORMATION

This is the second Concept to remove the noise or make the img Shaeper / or Smooth as Compare to orignal Img.

Frequency : Same set of data how many times repeates

Filter Minimize the noize from frequency.
↳ Simply we remove the noise from the img.

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

$$f(x,y) = \frac{1}{MN} \sum_{u} \sum_{v} F(u,v) e^{j2\pi \cdots}$$

↳ Invers DFT

1) Fourier Trans-form is a mathematical tool used to analyzing the freqxency Content of an img or Signal

Hybrid img: Changed on distance

* Template Matching
  - Means predefined object.

Two forms of Template Matching

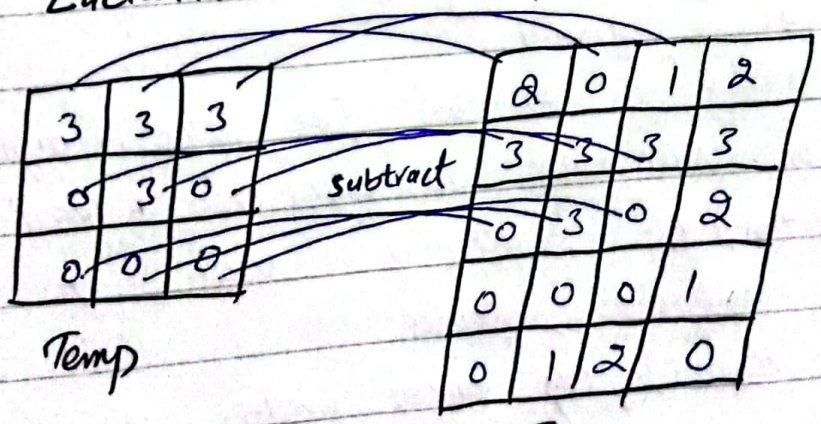
i) Global Template Matching
   - Template represents the whole image.

ii) Local Template Matching
   - we use small object not fully object
   - Simply we use the miner feature of the object
   - e.g: Corner, edges

Global TM

Measure of Similarity in TM

1) Eucledian Distance Measure: $E_2 = \sqrt{\sum_i \sum_j [g(i,j) - t(i-m, j-n)]^2}$



Temp

Img

$$E(1,1) = \sqrt{(2-3)^2 + (3-0)^2 + (3-1)^2 + 9+0+9+0+9+0}$$

$$= \sqrt{1 + 9 + 4 + \quad}$$

$$= \sqrt{41}$$

* Image Pyramids

is a technique used for multi-scale img representation.

An img pyramid is a collection of img that are derived from a single original image at different scale or representation

→ The purpose of img pyramids is to analyze & processing img at multiple scales, allowing for detection and extraction of feature at different level of granularity.

* Two Types of img Pyramids.

i) Gaussian pyramid:

in this type the orignal img is successfully downsamples (i.e: reduced in resolution) to create a series of img at different scales.

→ The downsampling process involves applying a low-pass filter to the orignal img to remove high-freq. components.

ii) Laplacian Pyramid

→ Its derived from the Gaussian pyramid.

→ Each level of Laplacian pyramid represents the difference b/w the corresponding level in the Gaussian pyramid and an upsampled and blured version for the next level Gaussian pyramid.

* Edge detection

An edge is a set of connect pixel that lies on the boundary b/w two regions which differ in grey value.

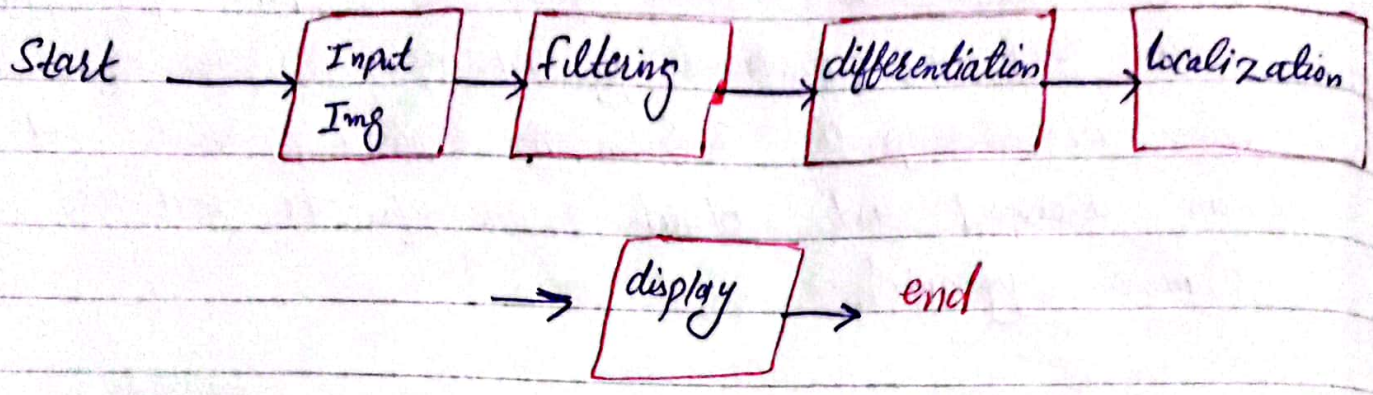
→ Edges provide an outline of the object.

→ An edge can be extracted by:

→ Magnitude of derivate
Direction of the derivate vector.

Step edge	Ramp edge	Spike edge	Roof edge
Suddenly chng in intensity	slow chng in intensity	Quick chng	for a long distance

* Stages in Edge detection:



i) Filtering :- Involves, Smoothing explicity on implicity.

ii) Differentiation :- distignuish edges pixels from other pixels.

iii) Localization :- exact location of edge.

* Hough Transform

It is a technique use to disjoint pixels.

↳ Its used to detect the lines, circles, ecllipses etc.

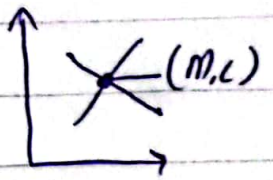
Detecting Lines using Hough Transform

$y = mx + c$

$C = -m * x_i + y_i$

→ we can consider x, y are parameters.

a) m and c are variables.



Q: Using Hough Transform show that the following points are collinear, also find the eqn of line (1,2), (2,3), and (3,4)

Ans Eqn of the line

$y = mx + c$

In order to perform hough transform, we need to convert the line from (x, y) to (m, c) plane.

$c = -mx + y$

i) For $(x,y) = (1,2)$, $c = -m + 2$

if $c = 0$ $\quad 0 = -m + 2$

$m = 2$

if $m = 0$, $\quad c = 2$

Thus $(m, c) = (2, 2)$

ii) For $(x,y) = (2,3)$, $c = -2m + 3$

if $c = 0$ $\quad 0 = -2m + 3$

$2m = 3$

$m = 3/2 = 1.5$

$m = 1.5$

Thus $m, c = 1.5, 3$

iii) For $(x,y) = (3,4)$, $c = -3m + 4$

if $c = 0$, $\quad 0 = -3m + 4$

$3m = 4$

$m = 4/3 = 1.33$

if $m = 0$ $\quad c = 4$

Thus $m, c = 1.33, 4$

* RANSAC (Random Sample Consensus)

is a popular Algorithm used in Computer Vision for robust estimation of parameter from noisy data.

Its useful when dealing with data that contains outliers or is corrupted noise.

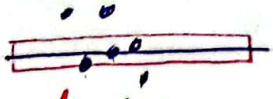
- The RANSAC is used to solve various problem such as fitting models to data, estimating motion parameter or detecting geometric structure in img.
- Its basically desined for handle outliers & noise.

outliers are points that don't "fit" the model

Problem with outliers

Least Squares estimation is sensitive to outliers, so that a few outliers can greatly skew the result.

Count = 3



RANSAC Algorithm in Simple words

i) Randomly select the sample data point.

ii) determine the set of data points which are within a distance threshold t of the model.

iii) If the size of s (number of inliers) is greater than some threshold t, restimating the model using all the points in consensus

iv) After N trials the largest consensus set is selected and the model is reestimated.