UNIT: 1

Computer Vision

Dashrath Nandan

★ Computer Vision

Computer Vision is a field of AI and computer science that focuses on enabling computers to interpret and understand the visual world.

Key Concepts of CV : Image Acquisition, Image Processing, Feature Extraction, Object Recognition.

Applications : Autonomous vechiles, Facial Recognition, Healthcare, Surveillance and Security, AR, Robotics, Scientific Research.

★ Digital Image Representation

Digital image representation involues the encoding of visual information in a form that computers and digital devices can process, store and display. This representation is based on pixels, Color Models and file formats.

* **Pixels**: The smallest unit of a digital image. Images are composed of a grid of pixels, each with a specific color value.

* **Color Models**: Color range:- (0-255)

(i) **Greyscale**: This black-and-white looking mode comprises various shades of Grey within an image. Each image pixel has a value ranging from 0 (Black) - 255 (White)

$$\text{Grayscale} = 0.299R + 0.587G + 0.114B$$

(ii) RGB (Red, Green, Blue) : The RGB color model is an additive system, because it adds wavelengths of the primary colours Red, Green and Blue togther to create a broad range of colours.

(iii) HSV (Hue, Saturation, Value):
- Hue → 0° to 360°, representing the type of color.
- Saturation → 0% - 100%, representing intensity of color.
- Value → Brightness, 0 → total black darkness to 100 → full brightness.

(iv) CMYK (Cyan, Magenta, Yellow, key/Black) → for printing.

* **Bit Depth**: The number of bits used to represent the color of each pixel.

1-bit: Black and White images ; 8-bit: Grayscale (256 shades of Grey)
24-bit: 16.7 million colors (RGB) ; 32-bit: RGB with alpha channel.

* **Image File Formats**:

Bitmap (BMP), JPEG, PNG, GIF, TIFF.
JPEG → Lossy Compression; PNG, GIF → Lossless Compression.

* **Image Representation Method / Format**:

Raster	Vector
• Pixel based, raster images are constructed through pixels.	• These are constructed through lines, curves and fills.
• They are not scalable.	• Scalable to any size.
• Prefers graphic formats like GIF, JPEG, PNG, etc.	• Prefers EPS, WMF, PICT, etc.
• Best for editing photos,	• Best when it comes to drawing, illustration and logos.
• We can easily convert a raster file to any file format.	• We cannot change vector file.

★ To create a digital image, we need to convert the continuous sensed data into digital form. This process includes:

1. Sampling :- Digitizing the co-ordinate value.
2. Quantization :- Digitizing the amplitude value.

	Sampling	Quantization
i)	Digitization of Co-ordinate Values.	Digitization of amplitude Values.
ii)	Sampling is done prior to the quantization process.	It is done after the Sampling process.
iii)	X-axis (time) – discretized. Y-axis (amplitude) Continous	X-axis (time) – Continous Y-axis (amplitude) – discretized
iv)	It determines the spatial (loc.) resolution of digitized image.	It determines the no. of grey levels in digitized images.
Adv :- v)	Data reduction, compatibility and Resolution Control.	Data Compression, Noise reduction, Simplified Processing.
vi)	Measured in PPI or DPI	bits per pixel (bpp).
vii)	Storage increases with higher sampling rates.	Increases with more quantization levels.
viii)	Affects the sharpness and clarity of image.	Affects the color richness of the image.

★ **Image Processing**

Image processing involves manipulating and analyzing images to enhance their quality, extract meaningful information or prepare them for further tasks.

* **Steps of Image processing:**

1. **Image Acquisition :** Capturing the image using Sensor or camera.

Steps: Selection of Sensor, Image Capture.

2. Pre-Processing: Removing noise and correcting distortions.
   Steps: Noise Reduction, Image Enhancement, Geometric Correction.

3. Image Restoration: Removing degradation from an image such as blurring, noise & distortions

4. Image Segmentation: Dividing the image ~~to~~ into meaningful region for analysis.

5. Feature Extraction: Identifying and extracting significant features from the image.

6. Image Representation and Description: Representing an image in a way that can be analyzed and manipulated by computer, and describing features in a compact way.

7. Image Recognition and Interpretation: Identifying objects or patterns within the image and making decision based on the analysis.

8. Image Compression: Reducing the size of the image for storage and transmission.

9. Image Display and Visualization: Presenting the processed image in a visual form.

10. Post-Processing: Refining and fine-tuning the processed image.
    Step: filtering, etc.

11. Output: Storing, sharing or using processed image.

★ Image Acquisition

Image Acquisition is the first step in the image processing workflow, involving the capture of a digital image from a physical scene using Sensor or camera.

Components: Image Sensor, optics, Lighting, Analog-to-Digital Conversion (ADC).

***Steps**:- 1. Selection of Sensor/Capture Device,
2. Setting up the Scene, 3. Capturing the image
4. Digitization, 5. Image Storage.

* Types of Image Acquisition Systems:

1. 2D Imaging System – Digital cameras, scanners.
2. 3D Imaging System – LIDAR, Structured Light Scanner.
3. Thermal (Infrared Cam.) 4. Medical I.S. (X-ray, MRI)
5. Remote Sensing System – Satellite and Aerial Sensor.

***Applications**: Medical Imaging, Remote Sensing, Industrial Inspection, Security, etc.

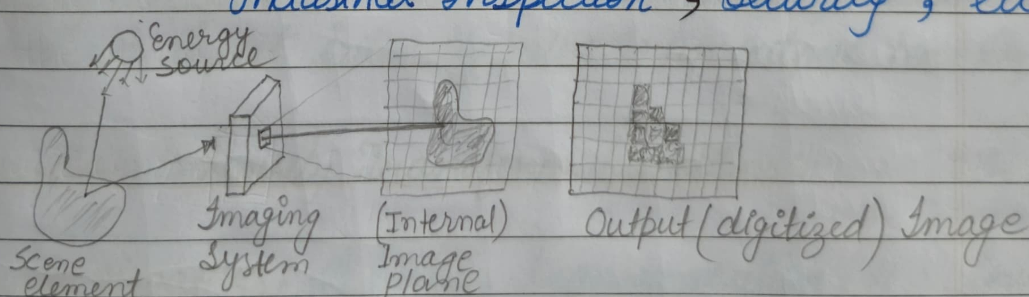


Image Acquisition Model

* Color Image Representation:

Color image representation refers to the method and formats used to encode and store color information in digital image.

Key Concept:—

- Pixels: The smallest unit of a digital image.
- Color Models: RGB, HSV, CMYK, etc
- Image Representation Formats:
  - Bitmap (Pixel-Based) Representation
  - Vector Representation.
- Compression:

i) Lossy Compression → Reduce the file size by compressing the image data, may result in some loss of details. (JPEG)

ii) Lossless Compression → No loss of details (eg. PNG).

★ Intensity Transforms Functions (Experiment – 1.3)

Intensity transformations functions or point processing operations, techniques used in image processing to modify the intensity values of individual pixels in an image.

- Intensity Value (Gray Level): The value representing the brightness of a pixels.
- Intensity Transform Function: A mathematical function that maps an input intensity value to an output intensity value.

* Common Intensity Transform Functions:

1. Image Negatives: A linear transformation, that involves inverting the pixel values of an image.

$$S = L - 1 - r$$

S: Output pixel value, L: Maximum pixel value, r: input P.V.

2. Log-Transformation: This technique used in image processing and CV to enhance the

details in images, particulary in low-intensity regions.

$$S/I_{log} = c \cdot \log(1+r)$$ → input p.v.

S ↓ Output p.v. c ↓ Constant Scaling factor.

3. <u>Gamma Transformation</u>: Also known as gamma correction or gamma adjustment, is a non-linear image enhancement used to correct or adjust the brightness and contrast of an image.

$$S = C \cdot r^{\gamma}$$ γ → gamma value

$\gamma = 1$: No change to the image (Linear transformation).
$\gamma > 1$: Darkens the image, enhances bright areas.
$\gamma < 1$: Lightens the image, enhances dark areas.

4. <u>Histogram Equilization</u>: To improve the contrast of an image by redistributing the intensity values so that histogram of the output image is more uniform.

5. <u>Piecewise Linear Transformation</u>: It involves applying different transformations to different regions or Segments of an image based on their intensity.

* <u>Application</u>:
- Image Enhancement.
- Medical Imaging.
- Remote Sensing.
- Photography.

Ex: Histogram Equalization:

Perform histogram equalization for the following image -

1	2	1	1	1
2	5	3	5	2
2	5	5	5	2
2	5	3	5	2
1	1	1	2	1

Input Image.

max. value = 5

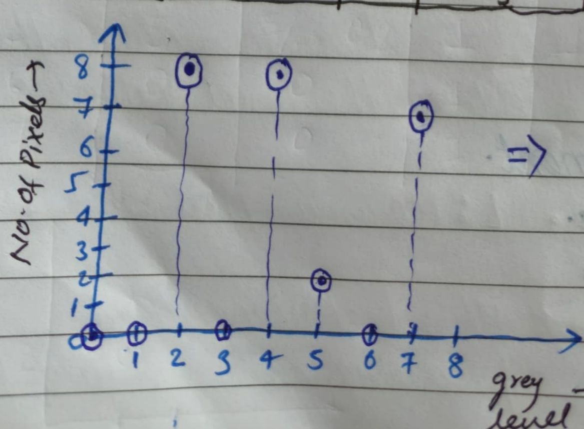
$2^0 = 1,\ 2^1 = 2,\ 2^2 = 4,\ 2^3 = 8$ $\quad 5 < 8 \quad \therefore L = 8$

$\therefore (0, L-1)$ will be interval.

Greylevel	n_k No. of Pixels	PDF (n_k/n)	CDF	f_k cdf $\times$ 7	Histogram Equalization level
0	0	0	0	$0 \times 7 = 0$	0
1	8	$8/25 = 0.32$	0.32	$.32 \times 7 = 2.24$	2
2	8	0.32	0.64	4.48	4
3	2	0.08	0.72	5.04	5
4	0	0	0.72	5.04	5
5	7	0.28	1	7	7
6	0	0	1	7	7
7	0	0	1	7	7

$n = 25$

New gray levels	0	2	4	5	7
No. of Pixels	0	8	8	2	7



=>

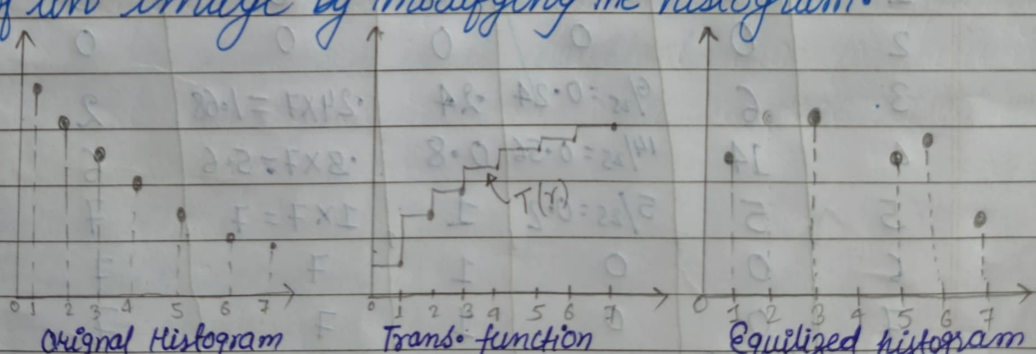
2	4	2	2	2
4	7	5	7	4
4	7	7	7	4
4	7	5	7	4
2	2	2	4	2

Output Image.

★ Histogram Processing

Histogram processing is a fundamental technique in image processing that involves the manipulation of an image's histogram to enhance or modify its visual appearance.

- **Histogram:** A histogram displays the frequency distribution of intensity values in an image.
- **Histogram Equilization:** It involves – Compute histogram, Compute CDF, Normalize CDF, Map Original Intensities.
- **Histogram Matching (Specification):** To modify the histogram of an image to match a specified histogram. It involves – Target histogram, Mapping Func, Apply Matching.
- **Histogram Stretching:** To increase the dynamic range of an image by streching the range of intensity values.
- **Clipping and Contrast Adjustment:** To control the brightness and contrast of an image by modifying the histogram.



*** Applications:** Medical Imaging, Remote Sensing, Photography.

*** Advantages:** Improved Visual Quality, Simple and Effective, Data Normalization.

Spatial Filtering

Spatial filtering technique used to enhance or extract features from an image by manipulating its spatial domain. This process involves applying a filter (or kernal) to each pixel in the image to produce a new pixel value based on the values of neighboring pixels.

- Filter (kernel): A small matrix of numbers used to modify the intensity values of the pixels in the image.
- Convolution: The primary operation in spatial filtering where the filter is applied to an image.

Types of Spatial Filters:

1. Linear Filters:
   (i) Smoothing Filter: Reduce noise and detail by averaging pixel values. Mean Filter, Gaussian Filter.
   (ii) Edge Detection filter: Detect edges by highlighting areas with rapid intensity changes. Sobel, Prewitt and Laplacian Filter.
2. Non-linear Filters:
   Median Filter: Replace each pixel value with the median of the pixel values in its neighborhood.

Application:

- Noise Reduction.
- Edge Enhancement.
- Reduce Blurring.
- Texture Analysis.
- Image Sharpening.

Fourier Transform and its Properties

It is a mathematical tool used in image processing, signal processing to analyze and represent signal and image in terms of their frequency components.

It transform a function (or Signal) from its original domain (often time or Space) into the frequency domain.

$\underbrace{f(x,y)}_{\text{Spatial domain}} \rightarrow$ $\underbrace{\boxed{\text{Transform}} \xrightarrow{T(u,v)} \boxed{\text{Operation } R} \xrightarrow{R[T(u,v)]} \boxed{\text{Inverse transform}}}_{\text{Transform domain}}$ $\rightarrow \underbrace{g(x,y)}_{\text{Spatial domain}}$

Key steps:
i) Transform the image
ii) Carry the task(s) in the transform domain.
iii) Apply inverse transform to return to spatial domain.

Properties:

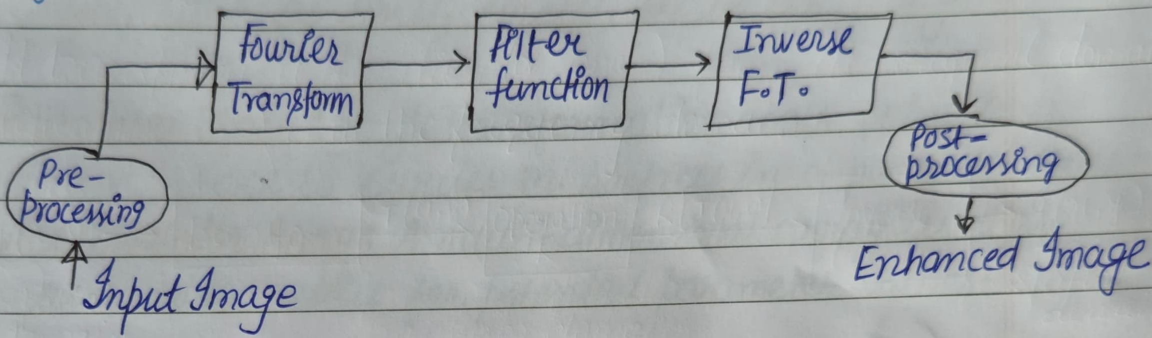
i) Linearity: F.T. is a linear operation.
ii) Shift (Translation): Shifting results in a phase shift in freq. domain
iii) Scaling: Scaling affects the spread in freq. domain.
iv) Symmetry: For real-valued func^n, F.T. has Symmetry.
v) Parseval's Theorem: The energy of a signal is preserved.
vi) Convolution Theorem: Convolution in time domain corresponds to multiplication in freq. domain.
vii) Duality: $f(x) \longleftrightarrow F(u)$
$F(x) \longleftrightarrow f(-u)$

Application:

Image Filtering, Image Compression, Image Analysis, Noise Reduction.

★ Frequency domain Filters

These filters operate on the Fourier Transform of a signal to modify its frequency components, effecitively altering the signal's characteristics in the time domain.



* Types:

1. Low-Pass Filters (LPF): To remove high-frequency components (noise, fine details) and retain low-freq. Compo
2. High-Pass Filters (HPF): To remove low-frequency components (Smooth regions) and retain high-freq. Components.
3. Band-Pass and Band-Stop Filters: Retain a specific range of frequencies outside this range.
4. Directional Filters: To emphasize or attenuate features in specific direction.

* Steps:

i) Transform Image.
ii) Apply Filter.
iii) Inverse Transform.
iv) Result.

Application: Noise Reduction, Edge detection, Feature extraction.

Advantages: Effective Noise Reduction, Edge Enhancement, Flexibility.

★ Hough Transformation

Hough Transformation is a feature extraction techingue used in image analysis, CV and digital image processing.
↳ It's primary purpose is to detect simple geometric shapes, such as lines, circles, etc in an image.

- **Parameter Space:** Hough Transformation maps point in the image space to curves or surfaces in a parameter space.
- **Accumulator Array:** A multi-dimensional array used to accumulate votes for potential parameter values.

* Steps involved —

	Line Detection using H.T.	Circle detection.
1.	Edge detection	→ Edge detection.
2.	Mapping edge points	→ Mapping edge points
3.	Identifying the local maxima in accumulator space.	→ Accumulating votes in a 3D accumulator array.
4.	Mapping these maxima back to the image space.	→ Detecting peaks in the accu. arrays to find circle.

* Properties

- Robustness to noise.
- Computational Costs.
- Shape Detection.

* Application:

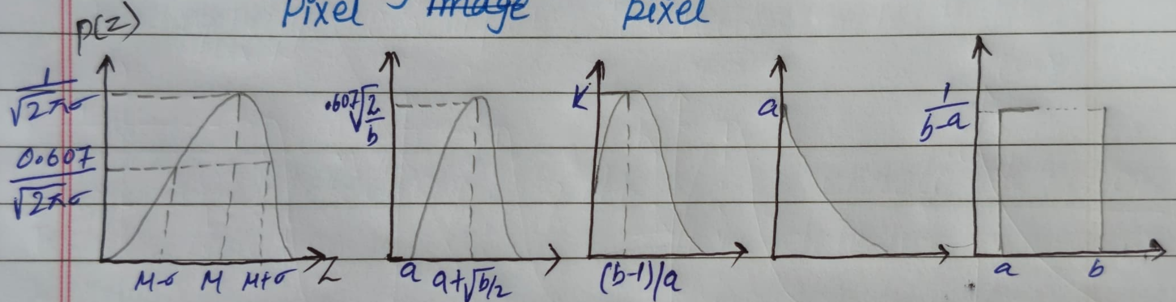
- Edge detection.
- Shape Recognition.
- Object Detection in Robotics.
- Medical Imaging.

Image Noise and Restoration

- The source of noise in digital images arieses during image acquisition (digitization) and transformation.
- Image restoration attempts to restore image that have been damaged.

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

resulting ~~image~~ noisy Pixel — original pixel — noise term



Gaussian (i) Rayleigh (ii) Erlang (iii) Exponential (iv) Uniform (v)

Image Restoration Techniques:

- Spatial Filter :-
  - ↳ Median Filter, excellent at noise removal.
  - ↳ Mean Filter, ↳ Adaptive Filters, remove impulse noise, Reduce distortion, provide smoothing.
  - ↳ Band Reject Filter, removing periodic noise.
- Frequency Domain Methods – Fourier Transformation.
- Statistical Methods – Bayesian Method, Regularization.

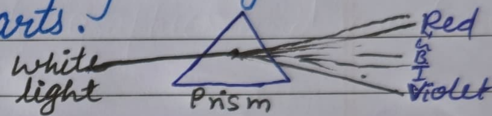
UNIT: 2

Dashrath Nandan

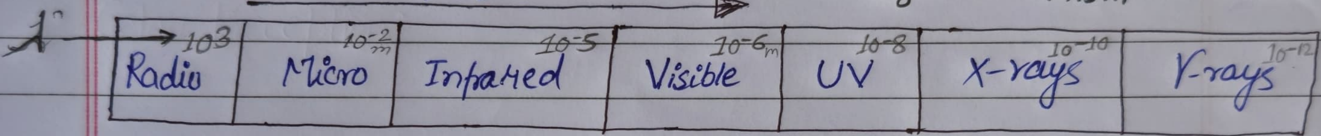
Computer Vision

★ Colour Fundamental

Color fundamental refers to the basic principles and concepts that describe how colors are preceived, represented and manipulated in various fields, including image processing, computer graphics, and visual arts.



λ↓ freq.↑



Electromagnetic Spectrum

1. Color Perception: Human Vision, Visible spectrum, Color attributes — Hue, Saturation, Brightness (Unit 1 notes)

2. Color Models: RGB, CMYK, HSV/HSI, Greyscale.

3. Color space: A color space is a specific organization of colors, often defined by a color model and a range of allowable Value.

4. Color Mixing:
   - Additive Color Mixing: Involves combining different colors of light to create new colors. Used in Screens.
   - Subtractive Color Mixing: Involves combining different pigments or dyes that absorb (subtract) light. Used in printing.

5. Color Balance and Correction: White balance, color correction and Gamma correction.

6. Color Vision deficiency: Color perceive differently due to absence or malfunction of cone cells in eye.

* **Chromatic Color**: It is a color that has a hue, saturation and lightness.

3 basic qualities:-

i) Radiance : The total amount of energy that flow from light source.

ii) Luminance : Amount of energy an observer perceives from light source.

iii) Brightness : Intensity of light.

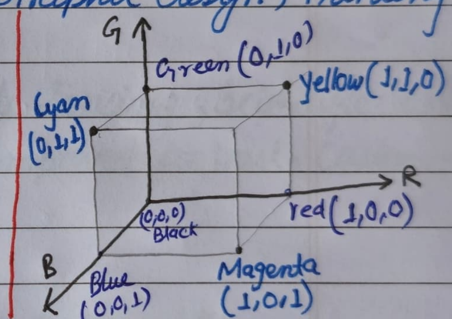
The proportion of blue used in colour is calculated as—

$$Z = 1 - (x + y)$$

* **Application of Color Fundamentals**—

Digital imaging, Computer Vision, Graphic design, Printing.

$R \rightarrow$ Color Monitor $\Rightarrow$ RGB
$G \rightarrow$
$B \rightarrow$



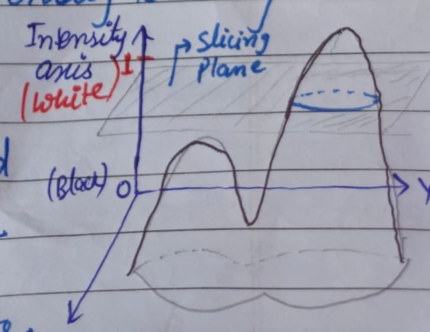
Pseudo Color Image Processing

Pseudo-color image processing is a technique used to enhance visual interpretation of grayscale images by mapping intensity levels to specific colors.

* **Methods:**

1. **Intensity Slicing**: It is also called density slicing or color coding.

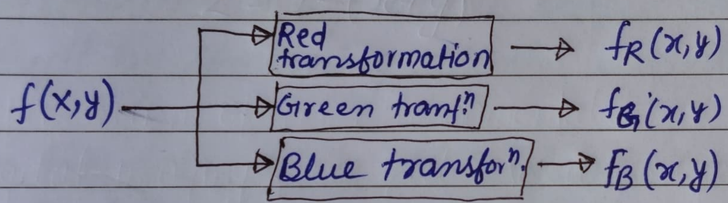
Imagine a greyscale image, as a 3D function with the intensity being the third dimension. Placing a plane parallel to the horizontal plane for the position coordinate would "Slice" the image into two parts.



After that we can assign different colors to different levels.

II. Color Transformation Functions:

This technique performs three independent tarnsformations on the intensity of the image, resulting in three images which are the R, G, B component images used to produce a color image.

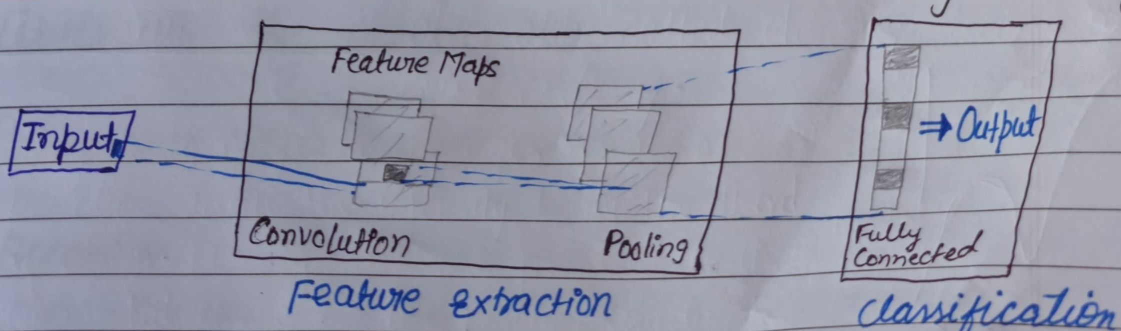


* **Application**: Medical imaging, Remote Sensing, Scientific Visul^{n}.
* **Advantages**: Enhanced Visualization, Improved contrast, Customizability

★ Image Processing using Neural Networks

Neural networks, particularly CNN are widely used for image processing tasks such as classification, Segmentation, Object detection and enhancement.

* Convolutional Neural Networks (CNN) in Image Processing



CNN is specialized for application in image & video recognition.

* CNN Layers:

1.) Convolutional Layers – Apply convolution operation (filter or kernel) to the input image to extract local features.

2.) Pooling layer – These layers reduce the spatial dimensions of feature maps, making the network more efficient. Max pooling and Average pooling are common pooling techniques.

3.) Fully Connected layer – Used to output the final classification result by combining all the learned features.

- The Full dataset is divided into two parts – Training Set and Test Set.

* Performance Metrics for image processing using NN.

1. Confusion Matrix:

	Predicted (+)	Pred. (−)
Actual Positive	TP	FN
Actual Negative	FP	TN

TP: True Positive
TN: True Negative
FP: False Positive
FN: False Negative

A confusion is a table used to describe the performance of a classification model by comparing the predicted lebels with the actual labels.

Performance Metrices derived from Confusion Matrix –

2.) Accuracy:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} = \frac{\text{No. of Correct Predc}^n}{\text{Total no. of Predic}^n}$$

3. **Precision**: It measures how many of the predicted positive instance are actually positive. It overcomes limitation of accuracy.

$$\text{Precision} = \frac{TP}{TP+FP}$$

4. **Recall or Sensitivity**: It aims to calculate the proportion of actual positive that was identified correctly by the model.

$$\text{Recall} = \frac{TP}{(TP+FN)}$$

5. **F1-Score**: Checks productivity or to maintain stability.

$$\text{F1-Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

6. **Specificity** (True Negative Rate): It measures how many of the actual negative instances were correctly identified.

$$\text{Specificity} = \frac{TN}{TN+FP}$$

Example:

	Pred. +	Pred. −
Act. +	50 TP	10 FN
Act. −	5 FP	35 TN

$$\text{Acc.} = \frac{50+35}{50+10+5+35} \approx 85\%$$

$$\text{Precision} = \frac{50}{50+5} \cong 0.91$$

$$\text{Recall} = \frac{50}{50+10} \approx 0.83$$

$$\text{F1-score} = \frac{2 \times 0.91 \times 0.83}{0.91+0.83} \approx 0.87$$

✯ Basic Operation on Binary and Greyscal Images

Binary Image	Greyscale Image
It consists of only two colors, typically black and white.	Consist of various shades of gray, ranging from B to W.
Each pixel represented by a single bit (0 (B) or 1 (W)).	Each pixel represented by 8 bits, allowing 256 diff. shades of gray.
Used in document scanning, barcode reading, etc.	Used in photography, medical imaging and Comp. Vision applcⁿ.
Smaller size file.	Large file size as compared to binary image.

Structuring Element: Also called a Kernel or Mask, is a small binary or grayscale matrix used to probe or scan the input image. It defines the neighborhood around each pixel that is considered during operations like dilation, erosion or opening.

* Morphological Operations:

1.) **Dilation:** Dilation is a morphological operation. that expands or "grow" the boundaries of foreground (white) object in an image.

- In binary images, dilation add pixels to the boundaries of object. If at least one neighborhood is 1, central pixel set to 1.
- In grayscale image, dilation replace the pixel value with the maximum value in neighborhood.

Input (3X3 Kernel) → Structuring element

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

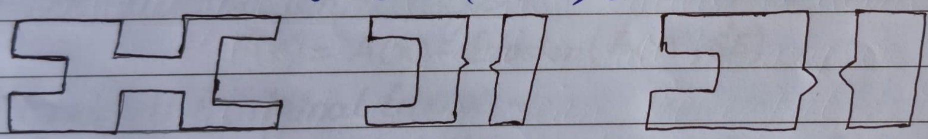
Binary Image

2.) **Erosion**: Erosion is the opposite of dilation. It "Shrinks" the boundaries of foreground objects, reducing their size. In binary, it remove pixels from boundaries and in Grayscale, it darkens region by reducing elements.

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

3.) **Opening**: Opening is a combination of erosion followed by dilation. It is used to remove small object or noise while preserving the shape & size of larger object, that are close to each other.

$$f \circ S = (f \ominus S) \oplus S$$



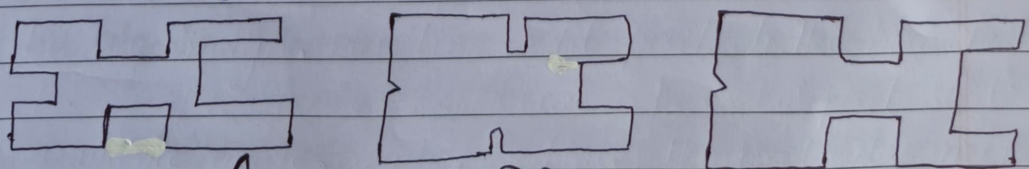
A — original shape

$A \ominus B$ — After Erosion

$A \circ B = (A \ominus B) \oplus B$ — After dilation (Opening)

4.) **Closing**: Closing is defined as dilation followed by erosion. Used to fill small holes or gaps in objects and smooth boundaries.

$$f \bullet S = (f \oplus S) \ominus S$$



A — Original Shape

$A \oplus B$ — After dilation

$A \bullet B = (A \oplus B) \ominus B$ — After erosion (Closing)

★ Morphological Algorithms

Morphological algorithm are set of image processing techniques that use mathematical functions to analyze and manipulate the structure of objects in an image.

- These are used for object recognition, shape analysis, noise removal, and image enhancement.

1. Opening & Closing.

2. Boundary Extraction: Boundary extraction is the process of obtaining the boundaries or edges of object within an image.

Structuring element

$$B(x) = A(x) - Erosion(A(x), SE)$$

1) Erode the original image.
2) Subtract the eroded image from the original image.

The remaining will be highlighting the boundaries of objects.

3. Region Extraction: It involves identifying and extracting connected components from an image. This process often uses Connected-Component labelling.

Methods—

i) Flood fill algorithm: Group pixels with similar properties into a single region.

ii) Morphological operations

Application — Brain tumor detection, medical imaging, etc.

4. Convex Hull: The convex hull of a shape is the smallest convex shape that entirely encloses the objects.

The convex hull can be computed by identifying the extreme points of the object and connecting them in such a way that the resulting shape is convex.

Application — Face recognition, shape analysis, object recognition, etc.

5. **Thinning**: Thinning is a morphological operation that reduces the thickness of object boundaries or lines to a single-pixel-wide skeleton preserving the original structure. $\boxed{A \otimes B = A - (A * B) = A \cap (A * B)^c}$

Thinning is done by repeatedly applying erosion operation until the objects are reduced to a skeleton.

Application :- OCR (Optical Character Recognition), fingerprint recognition, etc.

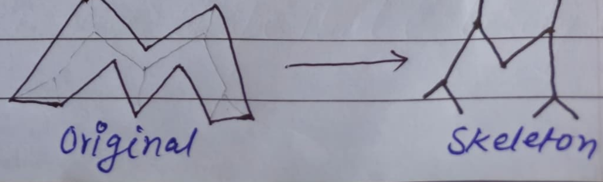
6. **Thickening**: Thickening is opposite of thinning. It is typically performed by a series of dilation operations, gradually expanding the object until it reaches a desired thickness.

$\boxed{A \odot B = A \cup (A * B)}$ → B is a structuring element.

Application— road detection or Vessel enhancement in medical imaging.

7. **Skeletonization**: It is useful in reducing the shape to its "Skeleton", a thin version of the object that retains its overall structure.

Method→ Skeletonization can be seen as more extreme form of thinning.

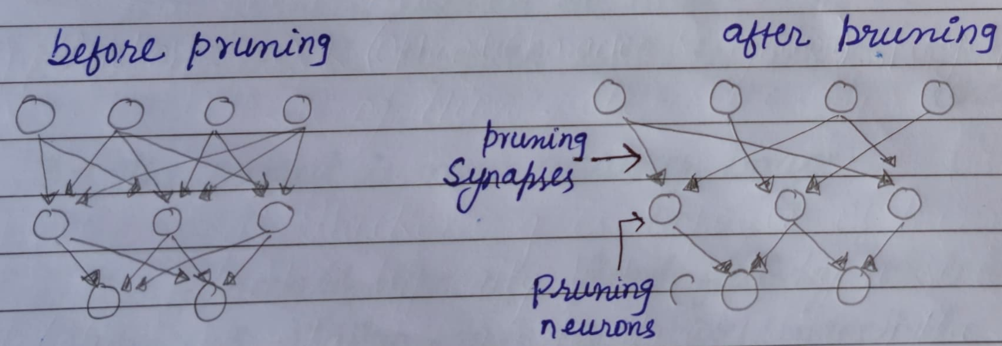


Application – Pattern recognition and shape analysis.

8. Pruning: Pruning is a process used to remove small, spurious branches from the skeleton of an object.

Method – Pruning is achieved by detecting and removing small, unnecessary branches or endpoints from the skeleton. This process typically follows thinning or skeletonization.

Application – Fingerprint analysis, Road network analysis.



* Morphological Reconstruction

Morphology in CV, is a set of processing techniques that use mathematical functions to modify the shapes of object in an image.

Morphological reconstruction is an image processing techniques that uses two images (Marker & Mask) to extract info. about shapes in an image.

- Marker Image: A binary or grayscale image that serves as the starting point for the reconstruction. It defines the area that should be preserved during reconstⁿ process.
- Mask Image: The Original Image (or another image) provides the contextual info. used to guide the reconstruction. It usually contains the desired shape or feature to be preserved
- Morphological Operations: Dilation and Erosion, these operation continues until the marker image can't be further dilated without exceeding the mask.

*Steps: 1→ The algorithm dilates the marker image, while constraining it to remain within the mask.

2→ The dilation is repeated until the image value stop changing.

3→ The output is new marker image.

⇒ It is the generalization of flood-filling. It is based on morphological dilation, but uses pixel connectivity insted of a structuring element with a specific shape and size.

*Application: Segmenting MRI images, Removing shadow from images, Noise removal, Shape Analysis.

Advantages: • Preservation of Structure.
• Robustness.
• Flexibility.

UNIT: 3

Computer Vision

Dashrath Nandan.

classmate Date ____ Page ____

Point Detection

It is a technique in image processing and CV used to identify and locate specific, small regions in an image that exhibit a distinctive feature, often referred as 'interest points' or 'keypoints'.

- Points detection can be achieved simply using mask—

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

Points are detected at those pixels in the subsequent filtered image that are above set threshold.

Algorithms:

i) Harris Corner Detector:- Identifies corners in an image by analyzing changes in intensity in Multiple dirⁿ.

ii) Difference of Gaussian (DOG): Uses Gaussian filters to blur and particularly used in blob detection (Circular Region). using SIFT (Scale-Invarient Feature Transform)

Application:

Feature detection and matching, Image alignment, Motion detection, Medical Imaging.

Line Detection

Line detection algorithms aim to find Straight lines in an image.

Horizontal

-1	-1	-1
2	2	2
-1	-1	-1

+45°

-1	-1	2
-1	2	-1
2	-1	-1

Verticle

-1	2	-1
-1	2	-1
-1	2	-1

−45°

2	-1	-1
-1	2	-1
-1	-1	2

* Algorithm:

i) Hough Transformation :- A voting based algorithm that detects line by transforming edge points into parameter space and finding peaks correspoinding to potential lines.

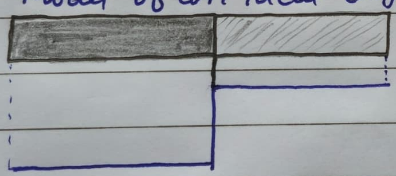
* Applications :

lane detection in autonomous driving, detecting text in OCR, Architectural analysis.

★ Edge Detection

Edge detection identifies the boundaries or contours within an image, where there is a significant change in intensity.

Model of an ideal digital image



: Grey-Level profile of a horizontal line through the image.

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

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

Sobel

* Algorithms:

i) Canny Edge Dector :- A widely used multi-step edge detector. It uses Gaussian filtering, gradient calculation, non-maximum suppression and hysteresis thresholding to detect strong and weak edges.

• Steps:-

1. Noise Reduction.
2. Gradient Calculation.
3. Non-maximum suppression.
4. Double thresholding.
5. Edge tracking by hysteresis.

• Advantages:

Produce highly accurate, thin edge with minimal noise

ii) Laplacian of Gaussian (LoG) :- The LOG combines a Gaussian filter for smoothing and a Laplacian operator to find edges.

- It is used for detecting thincker boundaries.

Steps:	Advantage:
1. Gaussian smoothing 2. Laplacian filtering 3. Zero-crossing detection	Ideal for detecting edges in image with complex textures.

iii) Sobel Operator :- A gradient-based method that detects edges using convolution with Sobel kernels in horizontal and Vertical directions.

★ Corner Detection.

Corner dection algorithms identify points where two-or-more edges meet. Corners are useful for tracking features in image.

* Algorithm

i) Harris Corner Detector :- It uses the gradient of image intensity and looks for regions with Significant variation in Intensity in two directions.

* Application :- Provide robust features for object tracking, image matching, 3-D reconstruction, etc.

→ Point, line, edge and Corner detections are fundamen-tal techniques in CV, used to identify features in an image.

★ Edge and Boundary Linking

These processes in image processing follows edge detection to form continuous, meaningful boundaries or contours of objects in an image.
These techniques helps in connecting these disjointed edges to form complete object boundaries.

★ Edge Linking

Edge linking is the process of connecting isolated edge pixels to form continuous edge based on certain criteria such as proximity, gradient direction, and intensity.

* Techniques:

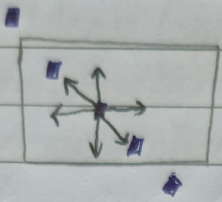
1.) Local Edge Linking (Pixel-based) :- It connects edge points based on local info. such as their proximity and edge gradient orientation.

2.) Hough Transform :- Used for detecting geometric shapes.
- Line Detection.
- Circle Detection

3.) Edge Relaxation :- Nearby edge pixels are examined iteratively, and the decision to link them is refined based on the analysis of nearby pixels relationships.

4.) Graph-based Edge Linking :- Treat edge points as node in a graph and links them based on their proximity and gradient orientation.

* Applications :- Object Recognition, Shape Analysis, Face detection & Recognition.



■ : Edge points
↑ : Direction of search
□ : Neighbourhood

* Boundary Linking

It focuses on connecting edges to form complete object boundaries or Contours.
The goal is to connect fragmented or incomplete edges to create meaningful boundaries encloses the regions.

* Techniques

1.) Contour Tracing :- In this method, once an edge pixel is found, the algorithm "follows" along the edge to trace out the full boundary of the object.

2.) Region-based Segmentation for Boundary detection :- After, initial edge detection, the regions are analyzed to group pixels based on their intensity or texture properties.

3.) Active Contours (Snakes) :- Snakes, are energy-minimizing curves that evolve to fit the boundary of objects.

4.) Graph-Cut Method :- Graph partitioning problem, where the goal is to cut the graph in such a way that the boundary of an object is identified.

* Challenges in Edge and Boundary Linking

- Noise and Weak Edges.
- Texture Complexity.
- Broken or Disconnected Edges.

* Segmentation

Image segmentation is the process of partitioning an image into multiple segments (set of pixels) to simplify or change the representation of an image, making it more meaningful and easier to analyze.

* Types of Segmentation

1.) Thresholding: Converts an image to binary form by setting a threshold value, separating the foreground from the background.

- Basic types of thresholding:-

i) Global Thresholding:- A single threshold value T is selected for the entire image.

ii) Adaptive Thresholding:- Threshold value changes across different regions of the image to handle varying illuminations.

- Mean Adaptive Thresholding
- Gaussian Adaptive Thresholding:- uses a weighted sum of the pixel intensities in local region.

iii) Band Thresholding:- A range of intensities is used to define a foreground object.

iv) Multilevel Thresholding :- Instead of converting the image to two classes (Foreground and BG), the image is divided into several classes based on multiple threshold value.

2.) Region-based Segmentation: It divides the image into segments by grouping pixels with similar properties such as color, Intensity, etc.

Techniques:-

i) Region Growing : The procedure in which pixel are grouped into larger regions based on some predefined condition is known as region growing. The basic approach is to pick a seed point (Starting Point) and growing regions from this seed pixel.

ii) Region Splitting : In this, we try to satisfy the homogenity property where pixels that are similar are grouped together.

iii) Region Merging : Once the splitting is done, adjacent regions are examined. If two adjacent regions are found to be similar, they are merged into a larger region.

iv) Split and Merge (Quadtrees): This method starts with the entire image and recursively splits it into quadrants (4 region). After splitting, adjacent regions are merged if they are sufficently similar.

* The regions that are formed using region-based segmentation must have following properties :-

i) The sum of all regions is equal to the whole image.
ii) Each region is contiguous & connected.
iii) There is no overlap of pixels.
iv) Each region must satisfy uniformity condition.
v) Two adjacent regions don't have anything in common.

3.) Graph-based Region Segmentation : This method treats image pixels as node in a graph. Edges between nodes represent the similarity between neighbouring pixels. The goal is to partition the graph such that regions with high similarity are grouped together and dissimilar regions are seperated.

* Application :- Medical imaging, Autonomous driving, Object detection.

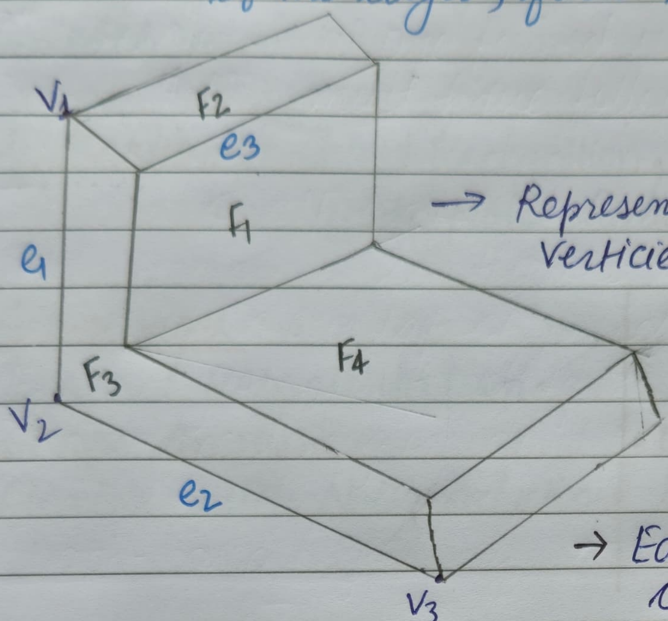
Boundary Representation

B-Rep is a method used in image processing to define the shape of an object by representing its boundaries rather than its interior.

* The description of the object can be into 2 parts :-

i) Topology :- It records the connectivity of the faces, edges & vertices by means of pointers in the data structure.

ii) Geometry :- It describes the exact shape & position of each of the edges, faces & vertices.



→ Representation of faces, edges & verticies.

→ The geometry of a vertex is just its position in space as given by its (x, y, z) coordinates.

→ Edges may be straight line, circular arcs, etc.

* Techniques :

i) Chain Code Representation :- Chain code represents the boundary as a sequence of directions relative to the previous pixel along the boundary. Each pixel on the boundary is assigned a directional code (e.g. 0-7 for an 8-direction code in a grid).

- A chain code is a lossless image representation algorithm for binary images.
- The incoming and outgoing direction are estimated by using the three directly connected neighbouring boundary points before and after the current point.

ii) Boundary Descriptors :- It represent boundaries by analyzing their geometric properties rather than focusing on individual pixels or segments.

iii) Run-Length Encoding for Boundaries (RLE) :- This method represents boundaries by storing only the starting and ending coordinates of sequences of boundary pixels in each row.

* Application :- Object Recognition, Shape Analysis, Flexible with Noise Reduction.

Advantages	Disadvantages
i) Compact Representation.	• Sensitive to Noise.
ii) Shape Analysis.	• Dependence on Accurate Edge Detection.
iii) Flexible with Noise Reduction.	• Complex shape.

★ Region Representation

It is a technique in image processing used to represent objects or areas in an image based on their interior region rather than their boundaries.

- Key Concepts :- Homogenity, Connectivity, Compactness.
- Region representation involves grouping pixels that share similar properties, such as color, intensity, etc.

* Type of Region Representation

i) Region-Based Segmentation Representation.
ii) Run-Length Encode (RLE).
iii) Pixel-Based Representation :- Each pixel within a region is represented individually, with properties such as intensity, color, or texture recorded for each pixel.

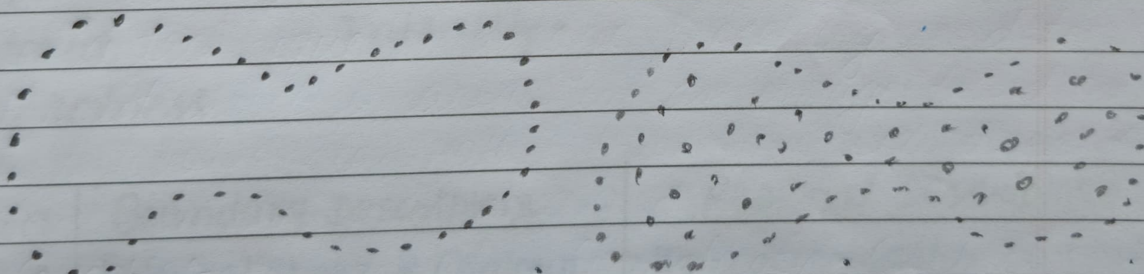
iv) Binary Spatial Array Representation :- Also known as Bitmask, represent each region by setting specific bits to 1 for pixels that belong to the region and 0 for those that do not.

v) Skeletonization.

vi) Texture-Based Region Representation.

* Application :- Object detection, Pattern Recognition, Content-Based Image Retrieval.

Advantages	Disadvantages
i) Preserve object area.	→ High memory Requirement.
ii) Support Complex Shape Analysis.	→ Computational Complexity.
iii) Flexibility in Property Analysis.	→ Sensitive to Initial Segmentation.



Boundary-based   Region-based
Representations

★ Boundary Descriptors

Boundary descriptors focus on the shape and characteristics of the outline or contour of an object. These descriptor provide information about the form, structure, and spatial arrangement of objects boundary.

classmate

Date
Page

* Common Boundary Descriptors:

i) Chain code Representation.

ii) Fourier Descriptors :- This is computed by a applying a Fourier transform to the boundary coordinates (x,y) of object.

iii) Curvature : Helps in identifying significant shapes.

iv) Boundary Length and Perimeter :- Measure the total length of the boundary.

* Regional Descriptors

Region-based descriptors, focus on the properties of the pixels within the object, rather than its boundary.

* Common Regional Descriptors:

i) Area

ii) Mean and Standard deviation of Intensity.

iii) Centroid

iv) Compactness

*

Aspect	Boundary Descriptors	Regional Descriptors
• Focus	External Shape & Contour	- Interior properties of region
• Applicⁿ	Shape analysis, Object recognition.	- Texture analysis, Classification, image reterival.
• Key Metrices:	Curvature, Fourier descriptor, Shape Signature.	- Area, centroid, texture, moments.
• Complexity	Dependent on boundary Precision	- Often computationally intensive for large region.
• Adv :	Efficient shape matching and object detection.	- Provide detailed internal feature representation.
• Disadv:	Sensitive to boundary Noise.	- May require large memory for high-resolution image.

★ Image Wrapping

Image Wrapping is a transformation technique used in image processing to change the spatial configuration of an image by manipulating the position, shape or orientation within an image.

- Transformation function, that dictate how each pixel in original image maps to a new position.
- Mapping:
  Forward mapping: Each pixel in the source image is mapped directly to the destination image.
  Backward mapping: Each pixel in destination image is mapped back to a location in the source image.
- Interpolation: When pixels are mapped to non-Integer coordinates, interpolation is used to estimate new positions.

✱ Types of Image Wrapping

I) Geometric Transformation:- Rotates the entire image around a specific point, often image centre.

II) Perspective Transformation:- Used to change the viewpoint of an image, useful for simulating 3D perspective.

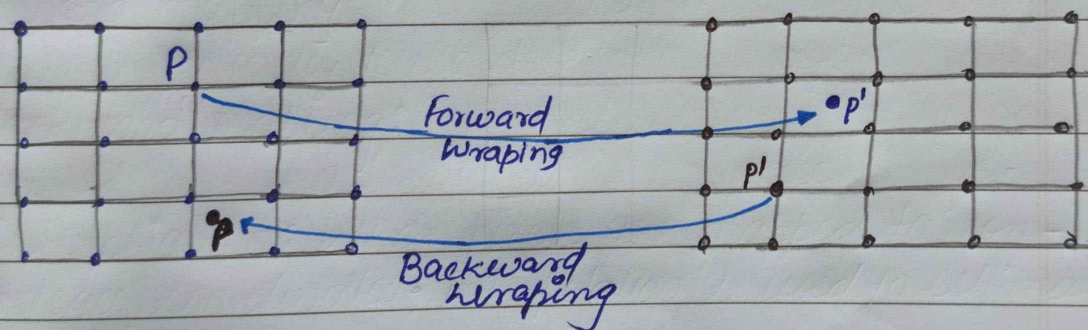
III) Non-Linear Wrapping:- Based on function that wrap the image differently in different area.

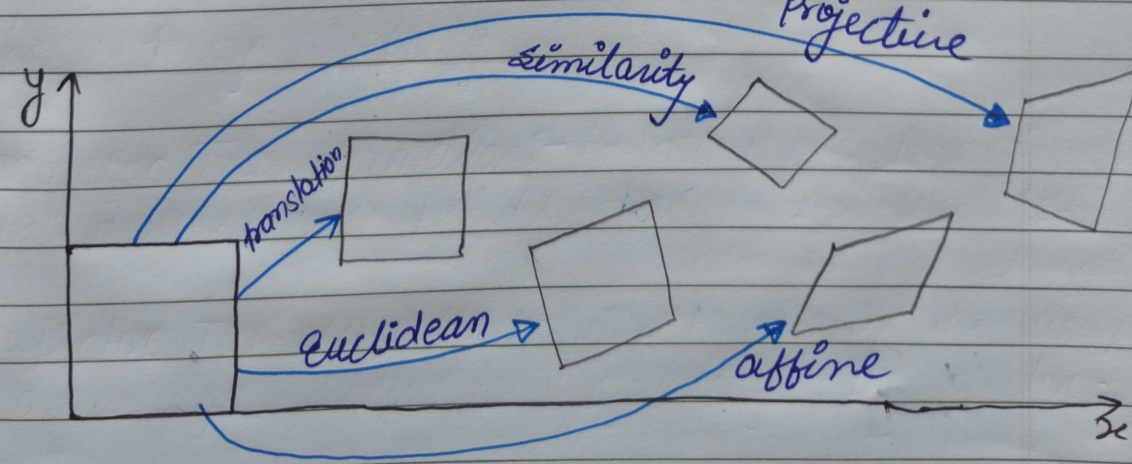
IV) Morphing:- Gradually transform one image into another by wraping the source image and blending it with target image.

v) Elastic Deformation :- It stratches or contracts regions within an image.

* Application :-

i) Image Registration and Alignment.
ii) Perspective Correction.
iii) Augmented Reality.
iv) Panorama Stitching.





Transformation and Wraping