Transformation matrices in 3D :- involves the z-axis in addition to x & y axis so the matrices become 4x4 to accommodate homogeneous coordinates.

1. Translation (3D) :- $\begin{bmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$

$t_x, t_y, t_z$ are translations in x, y & z directions

2. Rotation (3D) : can occur about the x, y or z axis.

-> Rotation about the x-axis $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

-> Rotation about the y-axis $\begin{bmatrix} \cos\theta & 0 & \sin\theta & 0 \\ 0 & 1 & 0 & 0 \\ -\sin\theta & 0 & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

-> Rotation about the z-axis $\begin{bmatrix} \cos\theta & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

3. Shear (3D) : can also applied in 3D to distort the image along one or more axes.

-> Shear along x axis $\begin{bmatrix} 1 & \lambda_{xy} & \lambda_{xz} & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

where $\lambda_{xy}$ is the shear in the xy plane, & $\lambda_{xz}$ is the shear in the xz plane.

4. Reflection (3D) flips the image along one or more axes.

→ Reflection accross the xy-plane
$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

→ Reflection across the yz-plane
$$\begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

→ Reflection across the xz-plane
$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Advantages of using 3X3 matrix for transformation.

1. Translation handling - A 3x3 matrix includes an extra dimension to handle translation along with rotation, scaling & shear in one unified framework.
2. Combination of Transformations - 3X3 matrix allows you to easily combine transformations like translation, rotation & scaling into a single matrix.
3. Homogeneous coordinates : By using homogeneous coordinates, the transformation of points at infinity can be handled, which is essential in many CV tasks like projective geometry.

Q15. Define segmentation, give application.

Ans Segmentation is the process of dividing an image into multiple regions or segments. The goal is to simplify or change the representation of an image to make

VISION

it more meaningful & easier to analyze. In segmentation, similar pixels are grouped together based on criteria such as color, intensity, textures, or edges resulting in regions that correspond to objects, boundaries or areas of interest image

Types of segmentation:

-> Thresholding - where pixel intensities are compared to a threshold value to separate foreground from background

-> Edge based segmentation: Detect the boundaries between regions based on edge detection methods like sobel, canny etc

-> Region based segmentation: Divides the image into regions based on the similarity of pixels such as region growing.

-> Clustering based segmentation: Used clustering algo like k-means to group similar pixels.

-> Semantic segmentation: Assigns a label to each pixel of the image

-> Instance segmentation: Identifies different instances of the same object category within an image.

Applications of segmentation:-

1. Object detection & Recognition - segmenting objects like vehicles, pedestrains, or animals for autonomous driving, surveillance, or scene understanding
Benefit:- Allows detection & classification of individual object within an image, helping to track movement & identify objects.

2. Medical Imaging: segmenting tumors, organs or tissues in medical scans (MRI, CT, or X-ray images).
Benefit:- Helps in diagnosing disease, planning treatments & conducting medical research by accurately isolating regions like tumors or organs.

3. Image compression : segmenting images into regions for more efficient compression.
Benefit : By grouping similar regions, segmentation can reduce redundant information & improve compression without significant loss of image quality.

4. Scene understanding : segmenting different parts of a scene for autonomous vehicles or robotic navigation.
Benefit : Enable robots or autonomous system to recognize road, buildings, pedestrians, obstacles etc for decision making.

5. Facial recognition & feature detection : segmenting facial features like eyes, nose & mouth for applications like security authentication or AR filters.
Benefit : Helps in identifying key facial landmarks, enabling facial recognition, emotion detection, & augmented reality effects.

6. Background substraction : segmenting the foreground from the background in video sequences
Benefit - useful in surveillance, motion detection, & video compression by isolating objects of interest in a scene.

7. Optical character Recognition : segmenting individual characters from scanned text documents for recognition.
Benefit - Helps in extracting & recognizing text from images or scanned documents for digital storage or processing.

Q17. Define thresholding. Give difference between global & local thresholding?

Ans Thresholding is a simple image segmentation technique used to convert a grayscale image into a binary image. It works by selecting a threshold value, & every pixel in the image is compared to this threshold.

* Pixel greater than the threshold are assigned to one class (white or 1)
* Pixel less than or equal to the threshold are assigned to another class (e.g. black or 0)

| | Global Thresholding | Local Thresholding |
|---|---|---|
| 1 | Threshold value :- Use a single threshold value for the entire image. | Calculates different threshold values for different regions. |
| 2. | Usage scenario :- Best for images with uniform lighting & contrast | Suitable for images with varying lighting conditions. |
| 3. | Computation :- simple and computationally efficient | More complex & computationally intensive. |
| 4. | Example Algorithm : Otsu's method (automatically determines global threshold). | Adaptive Gaussian Thresholding, Adaptive mean Thresholding |
| 5. | Output Consistency : The entire image is binarized with the same logic, making it less flexible for varying conditions. | Different parts of image are binarized independently, providing flexibility. |
| 6. | Limitation : May fail in images with shadows, uneven lighting, or varying object intensities | Handles uneven lighting but may over-segment in some cases. |

Q18. What is Ostu binarization ?

Ans 18. Ostu's method is a global thresholding technique to automatically find the optimal threshold value for converting a gray scale image into a binary image. Ostu's method works by minimizing the intra-class variance or equivalently, maximizing the inter-class variance of the pixel intensities in an image. This ensures the two resulting pixel classes - foreground & background.

Steps of Ostu's Binarization:

1. Histogram Generation : Compute the histogram of pixel intensity values for the entire image.
2. Search for the optimal Threshold : Ostu's algo test every possible threshold value
   For each threshold, it divides the pixel into two classes:
   -> Foreground : Pixels with intensities greater than threshold.
   -> Background : Pixels with intensities less than or equal to threshold.
3. Calculate intra-class variance :- For each threshold, calculate the intra-class-variance which is the weighted sum of the variances of two classes
   -> Ostu's method seek to minimize this value, as it indicates a clearer separation between foreground & background.
4. Select the optimal threshold : The threshold that minimizes the intra-class variance is selected as optimal threshold.
5. Apply threshold : Finally, the optimal threshold is applied to the image, converting it into a binary image.

Formula for Ostu's Method:

Let,

VISION

→ $w_0(t)$ and $w_1(t)$ represent the probabilities of the two classes separated by a threshold $t$,
→ $\mu_0(t)$ and $\mu_1(t)$ are the mean intensities of two classes,
→ The total variance $\sigma^2(t)$ is defined as:

$$\sigma^2(t) = w_0(t) \cdot w_1(t) \cdot (\mu_0(t) - \mu_1(t))^2$$

Otsu's methods maximizes $\sigma^2(t)$ by varying $t$.

Advantages of Otsu's Binarization:
→ Automatic :- It automatically finds the best threshold, eliminating the need for manual input.
→ Global applicability :- works well when the image has a bimodal histogram, i.e two distinct peaks representing the foreground & background.

Limitations of Otsu's Binarization :
→ Not effective for non-uniform illumination :- Otsu's method assumes the image has uniform lighting & distinct foreground-background separation.
→ Sensitive to noise :- If there is a lot of noise in the image, Otsu's method may not find an optimal threshold.

Applications of Otsu's Binarization :
1) Document scanning :- Used to binarize text from scanned documents (black text on white background)
2) Medical Imaging :- Helps in separating objects like cells, tissues, or tumors in X-ray, MRI, or CT scans.
3) Object detection :- Useful in preprocessing images for detecting objects in scenes with relatively clear separation between the object & background.

Q19. Define adaptive thresholding. With respect to morphological operators. Explain structuring elements & types with ex?

Ans 19. Adaptive thresholding is a technique where the threshold value is calculated dynamically for smaller regions of an image rather than using a single global threshold for the entire image. This method is particularly useful for images with non-uniform lighting, shadows, or varying background intensities.

Common methods for adaptive thresholding :-

1) Mean adaptive thresholding : The threshold value is set to the mean of the pixel intensities in the neighborhood region.

2) Gaussian adaptive thresholding : A weighted sum of the pixel intensities the neighborhood is used to compute the threshold.

Morphological operations : are used to process binary or grayscale images based on the shapes or structures present in the image. These operations rely on the concept of a structuring element to probe the image, modifying pixel values based on the local neighborhood.

Structuring elements in morphological operations

A structuring element is a small binary matrix use to probe an image during morphological operations. It defines the neighborhood over which the operation will be applied. The shape & size of the structuring element determine the extent & type of transformation performed on the image.

Types of structuring elements :

1. Rectangular structuring element :
Shape : A matrix filled with ones, typically used for regular, box-shaped objects

Ex :-
$$\begin{matrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{matrix}$$

2. Elliptical (circular) structuring element :
Shape : A circular kernel, useful for rounded objects in an image . Ex →
$$\begin{matrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 \end{matrix}$$

3. Cross-shaped structuring element :-
Shape : A cross like pattern used for connectivity & specific shapes.
Ex -
$$\begin{matrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{matrix}$$

Q20. Dilation & erosion (definition) effect on image, example with image.

Ans Dilation :- is a morphological operation that adds pixels to the boundaries of object in a binary or grayscale image. It expands the shape of foreground objects by filling gaps or holes, making the object larger & smoother.
Effect on Image :-
→ Increases the size of objects in the image.

-> Fills small holes or gaps in objects.
-> Connects disconnected regions or objects.
-> Useful for bridging small breaks & closing gaps btw pixels.

Mathematical definition.

-> For a binary image, the dilation of set A by structuring element B is defined as:

$$A \oplus B = \{ z \mid (\hat{B})_z \cap A \neq \phi \}$$

-> This means that dilations shifts the structuring element B over the image & adds pixels to the object where the structuring element fits.

Example of dilation:

-> Original image - A binary image with a small object gaps.
-> Dilated image - The object will appear larger, with gaps between pixels or objects filled.

2. Erosion - is a morphological operation that removes pixels from the boundaries of object in a binary or grayscale image. It shrinks the foreground objects by eroding away the outer boundaries, making obj smaller.

Effect on image

-> Decrease the size of objects
-> Removes all noise or speckles from the image.
-> separates objects that are connected
-> Useful for shrinking obj or removing small artifacts & isolating important features

Mathematical definition:

For a binary image the erosion of set A by structuring element B is defined as

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

-> This means erosion shifts the structuring element &

VISION

over the image & remove pixels from the object if the structure element does not completely fit inside obj.

Example of erosion:

- Original image :- A binary image with small objects or connected components
- Eroded image :- The object will appear smaller, with thin feature or small noise removed.

Effect of dilation & erosion on image:

- If you have an image with small white dots (noise) on a black background, dilation would enlarge the dots & fill gaps.

Before (Dilation Example) After dilation

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

- If you have an image with thick white shapes on a black background, erosion would shrink the shape & remove small white specks

Before (Erosion Example) After erosion

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

Q21. Erosion, dilation, opening, closing.

Ans Erosion :- removes pixels from the edges of object in an image. It "erodes" away the boundaries of the foreground object, making the object smaller.

Effect on image

- Reduce size of objects.
- Removes small noise or thin structures.
- Seperates connected object.

Use case - To remove noise or separate touching objects.

Visual Example - If a white object is surrounded by black pixels, erosion will shrink the white area.

| Before erosion: | | | | | | After Erosion: | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 1 | | 0 | 0 | 1 | 0 | 0 |
| 1 | 1 | 1 | 1 | 1 | ---> | 0 | 1 | 1 | 1 | 0 |
| 1 | 1 | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | 1 |
| 1 | 1 | 1 | 1 | 1 | | 0 | 1 | 1 | 1 | 0 |
| 1 | 1 | 1 | 1 | 1 | | 0 | 0 | 1 | 0 | 0 |

Dilation :- add pixels to the boundaries of object in an image. It "grows" the foreground object, making it larger & closing small holes or gaps.

Effect on image:

- Increase the size of objects.
- Fills small holes and gaps.
- Connect broken parts of objects

Use case: To fill small gaps or connect broken components.

Visual Example: If a white object has gaps or holes, dilation will enlarge the white area & fill them.

Before dilation :

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

---->

After dilation

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

3. Opening :- is a combination of erosion followed by dilation. It first removes small noise or unwanted small objects with erosion & then restores the shape of the remaining object using dilation.

Effect on image :

-) Remove small noise

-) Smooths the contours of object without significantly changing their area.

Use case : To remove small objects or isolated noise from the image while preserving the main shapes.

Visual Example : If their is a small noise or small bright spots, opening removes them.

Before opening

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

----->

After Opening

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

4. Closing :- is a combination of dilation followed by erosion. It fills small holes or gaps within object using dilation & then reduces the size of object back to their original shape using erosion.

Effect on image :-

-> close small holes or gaps within objects.

-> Smooth contours.

Use Cases :- To fill small holes or connect broken parts of objects in the image.

Visual Example :- If there are small gaps or holes in an object, closing will fill them:

Before closing

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

---->

After closing

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