IMAGE REPRESENTATION & FEATURE EXTRACTION

Dr. Sandeep Kumar Jain

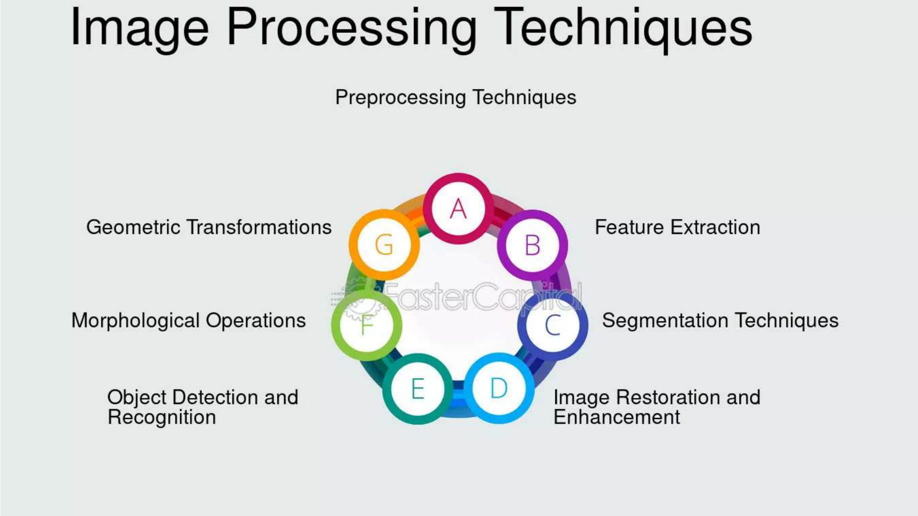


Image Processing Techniques

Preprocessing Techniques

Feature Extraction

Segmentation Techniques

Image Restoration and Enhancement

Object Detection and Recognition

Morphological Operations

Geometric Transformations

Motivation

- One of the major concern of image processing is image (object) recognition
    - Objects are represented as a collection of pixels in an image
- Our Task: To describe the region based on the chosen representation

Image acquisition => digital image

↓

Preprocessing => better image

↓

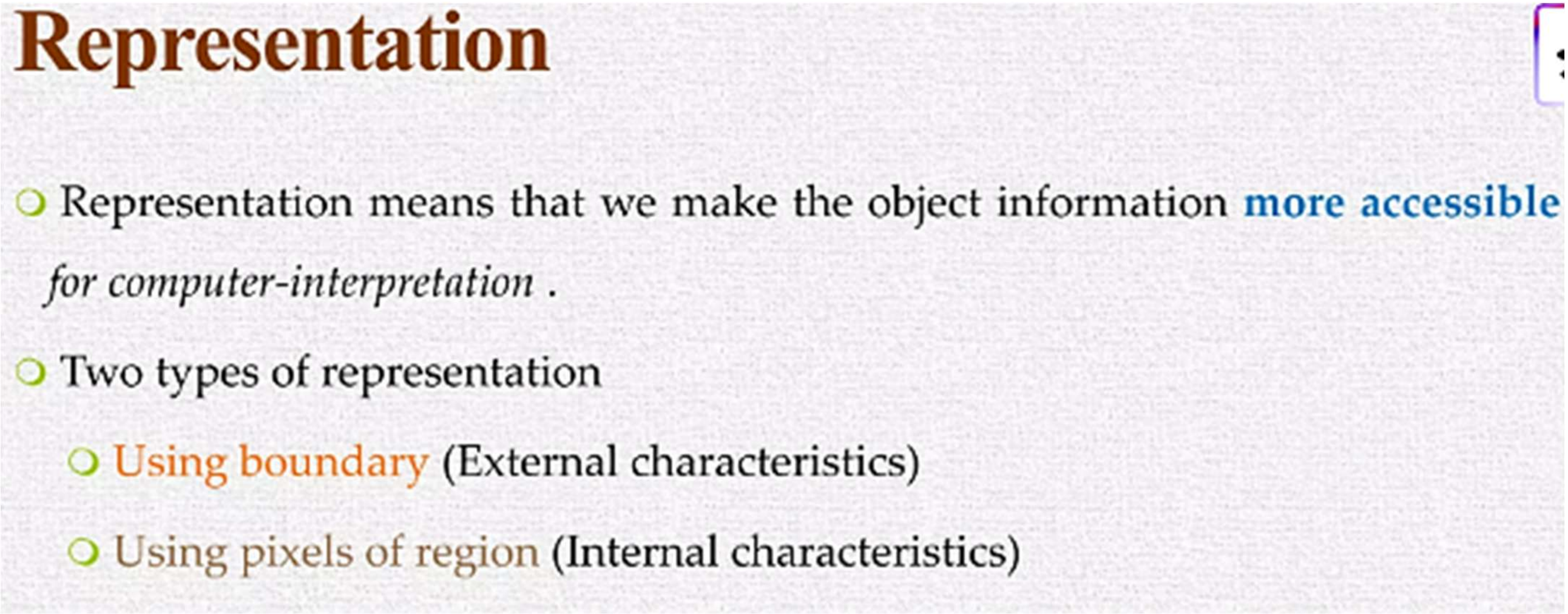
Segmentation => basic features

↓

Representation and description => advanced features

↓

Object recognition

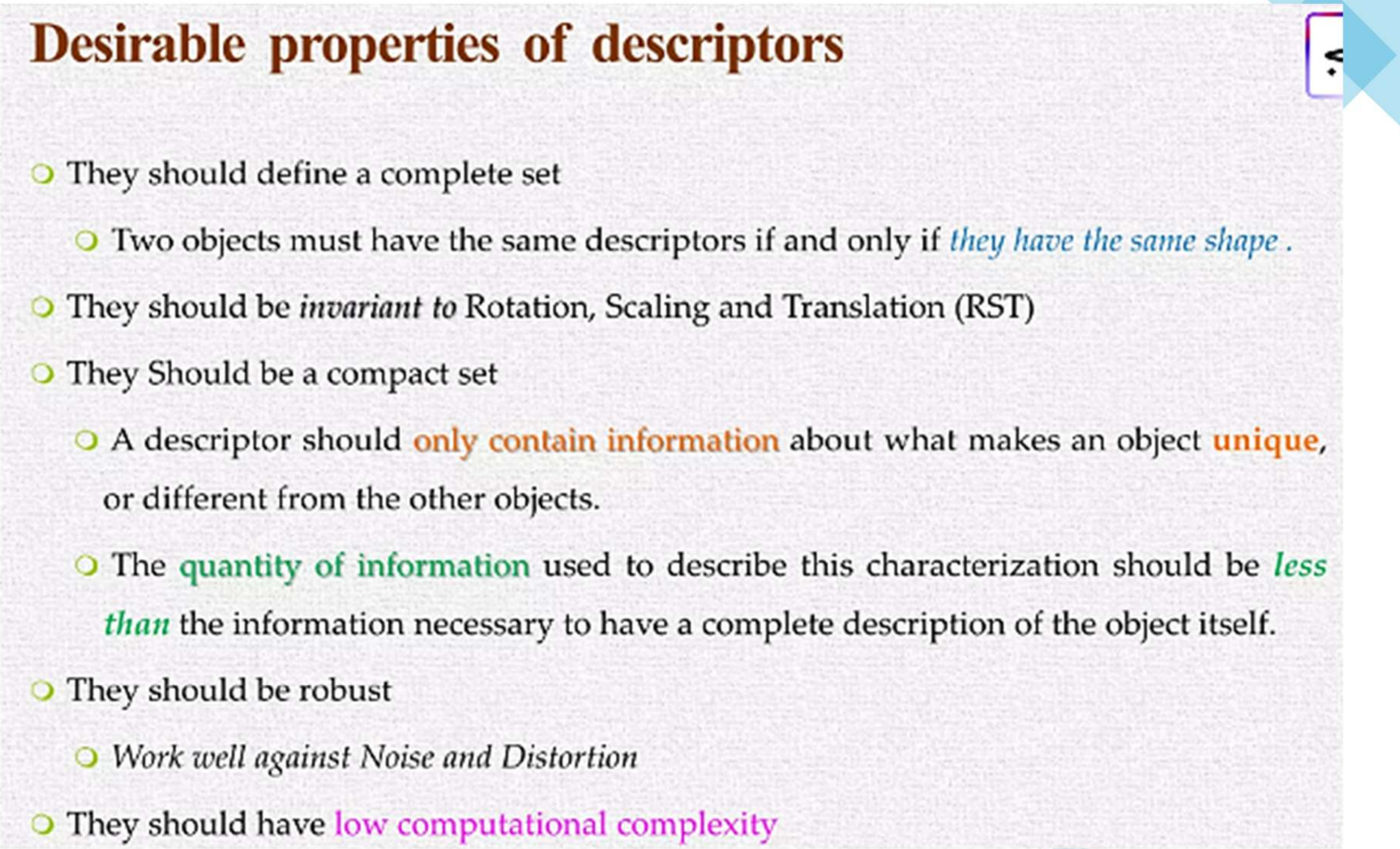


Representation

- Representation means that we make the object information **more accessible** *for computer-interpretation* .
- Two types of representation
  - Using boundary (External characteristics)
  - Using pixels of region (Internal characteristics)

Description

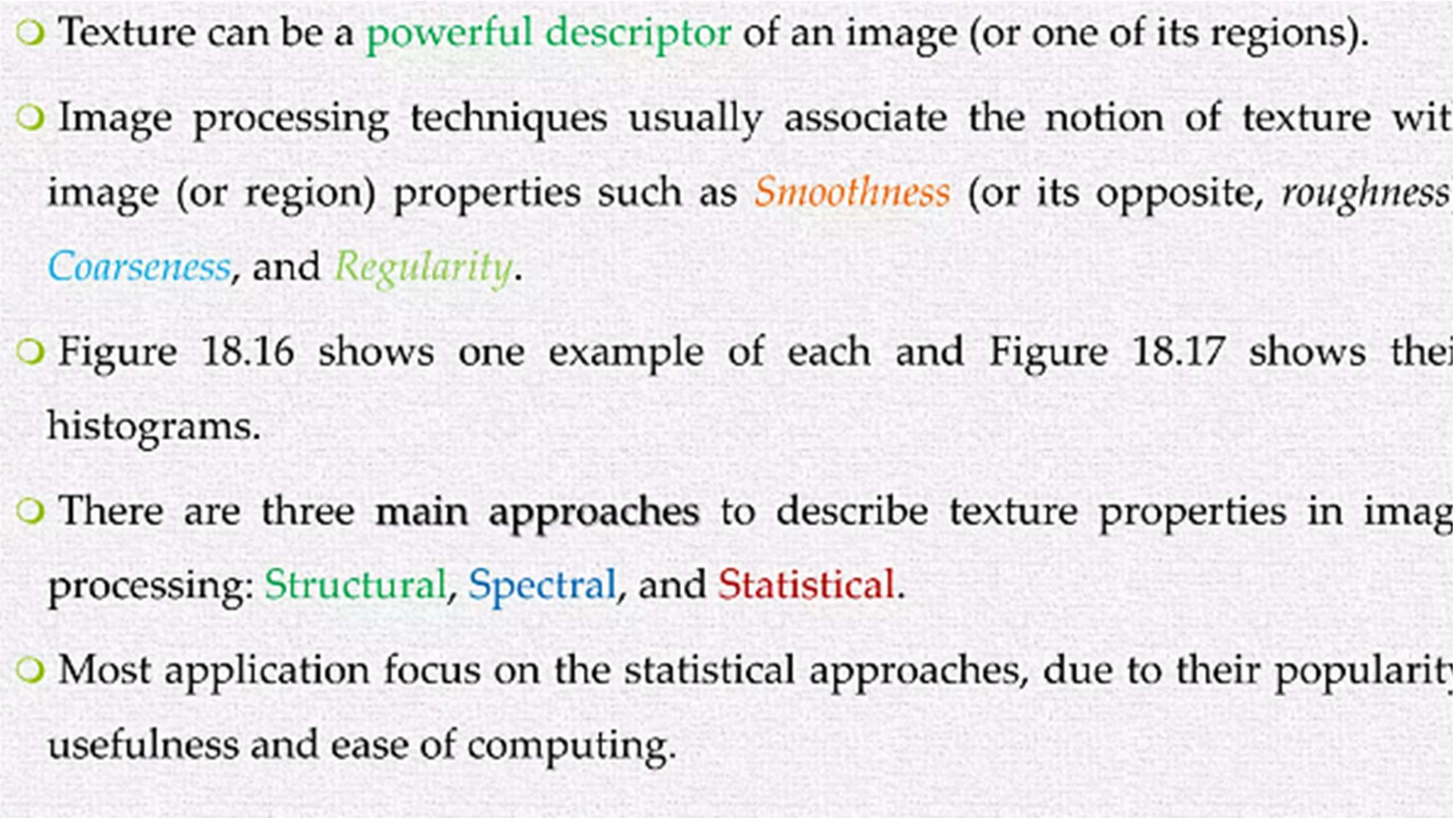
- Description means that *we quantify* our representation of the object
- **Boundary Descriptors**
  - Geometrical descriptors : Diameter, perimeter, eccentricity, curvature
  - Shape Numbers
  - Fourier Descriptors
  - Statistical Moments
- **Regional Descriptors**



Desirable properties of descriptors

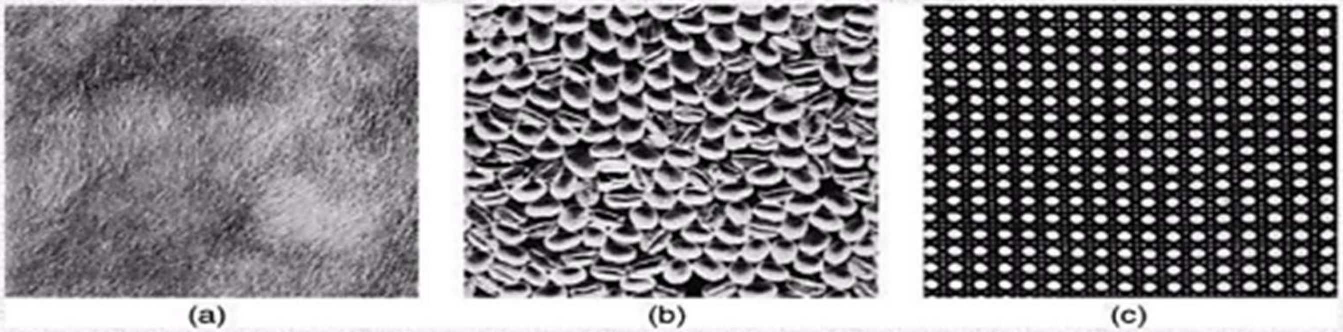
- They should define a complete set
  - Two objects must have the same descriptors if and only if *they have the same shape .*
- They should be ***invariant to*** Rotation, Scaling and Translation (RST)
- They Should be a compact set
  - A descriptor should only contain information about what makes an object **unique**, or different from the other objects.
  - The quantity of information used to describe this characterization should be *less than* the information necessary to have a complete description of the object itself.
- They should be robust
  - *Work well against Noise and Distortion*
- They should have low computational complexity

Texture Features

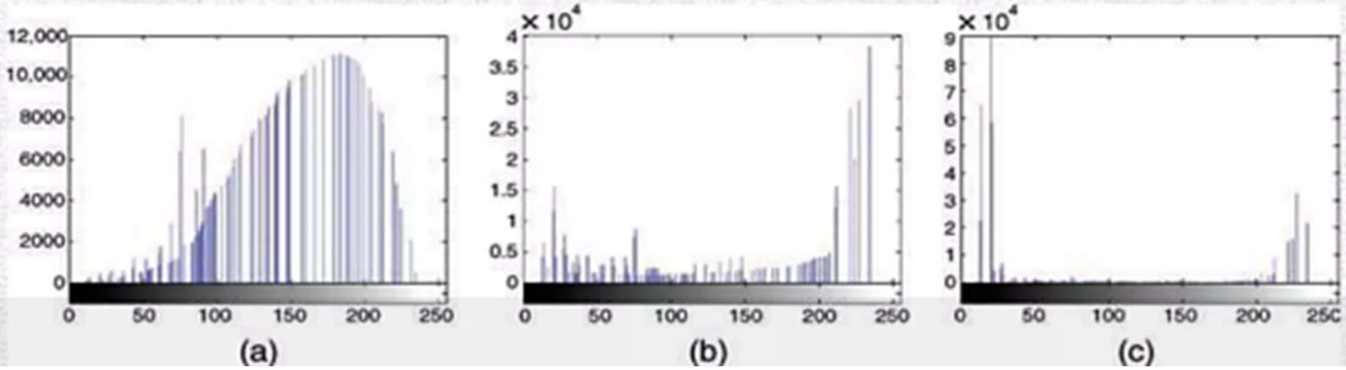


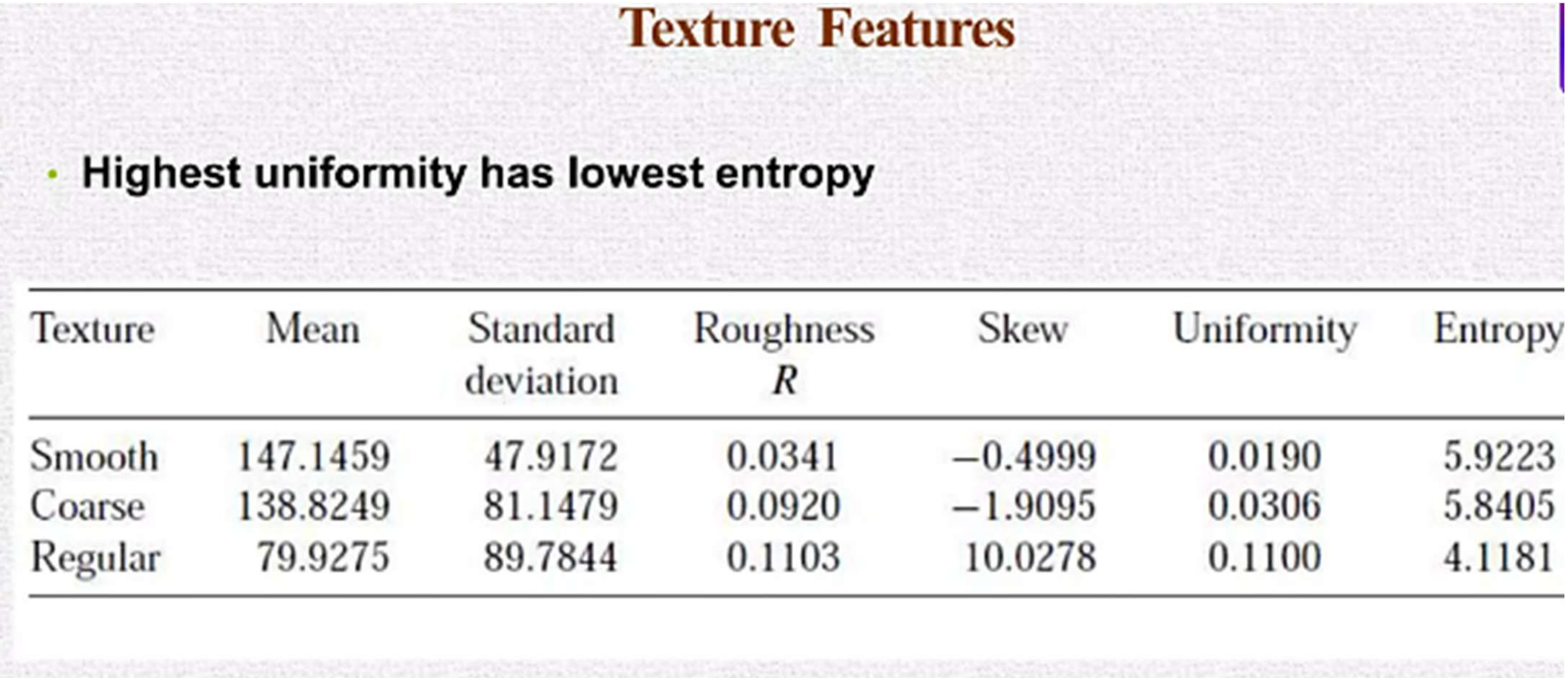
- Texture can be a powerful descriptor of an image (or one of its regions).
- Image processing techniques usually associate the notion of texture wit image (or region) properties such as *Smoothness* (or its opposite, *roughness* *Coarseness*, and *Regularity*.
- Figure 18.16 shows one example of each and Figure 18.17 shows thei histograms.
- There are three **main approaches** to describe texture properties in imag processing: Structural, Spectral, and Statistical.
- Most application focus on the statistical approaches, due to their popularit usefulness and ease of computing.

- FIGURE 18.16 Example of images with smooth (a), coarse (b), and regular (c) texture. Images from the Brodatz textures data set.



- FIGURE 18.17 Histograms of images in Figure 18.16.





Texture Features

- Highest uniformity has lowest entropy

Texture	Mean	Standard deviation	Roughness *R*	Skew	Uniformity	Entropy
Smooth	147.1459	47.9172	0.0341	−0.4999	0.0190	5.9223
Coarse	138.8249	81.1479	0.0920	−1.9095	0.0306	5.8405
Regular	79.9275	89.7844	0.1103	10.0278	0.1100	4.1181

1. Image/Object Feature Extraction

Feature extraction means computing **quantitative descriptors** that capture an image’s texture, shape, or structure so that objects can be recognized or compared.

a. Textural Features – Gray Level Co-occurrence Matrix (GLCM)

- **Idea:** Texture describes spatial repetition of intensity patterns.
- **GLCM:** A matrix where each entry (i, j)counts how often a pixel of gray level *i* occurs adjacent to a pixel of gray level *j* at a specific distance and direction.
- **Common Features Derived:**
    - **Contrast:** Measures local intensity variation.
    - **Correlation:** How correlated a pixel is to its neighbor.
    - **Energy (Angular Second Moment):** Uniformity of texture.
    - **Homogeneity:** Closeness of distribution to the diagonal (smoothness).

1. What Is GLCM?

GLCM is a **second-order statistical method** capturing the *spatial relationship* of gray levels, providing rich descriptors (contrast, energy, homogeneity, etc.) that outperform simple first-order histograms for texture classification and segmentation.

- **Texture** describes how pixel intensities vary spatially (smooth, coarse, striped, etc.).
- The **Gray Level Co-occurrence Matrix (GLCM)** captures how often pairs of pixels with specific gray levels occur next to each other in a given spatial relationship.

In short, GLCM is a 2-D matrix **P(i, j)** where:

- *i* = gray level of a reference pixel.
- *j* = gray level of its neighbor.
- Each cell counts how frequently that pair appears in the image with a specified **distance (d)** and **direction (θ)**.

2. Steps to Construct a GLCM

1. **Convert to Gray Scale**

Reduce color images to grayscale. Often quantized to fewer levels (e.g., 8 or 16) for manageable matrix size.

2. **Choose Parameters**
    1. **Distance (d):** pixel offset (commonly 1).
    2. **Direction (θ):** 0°, 45°, 90°, 135° (or all four).

3. **Pair Counting**

For every pixel, find its neighbor at (d, θ).

If reference pixel has value *i* and neighbor *j*, increment P(i, j).

4. **Normalize (optional)**

Divide by total pairs so that the matrix sums to 1. This turns counts into joint probabilities.

4. Practical Considerations

- **Multiple Directions:**

Compute GLCM for θ = 0°, 45°, 90°, 135° and average the features for rotation invariance.

- **Quantization:**

Fewer gray levels (e.g., 8) reduce computation but may lose subtle details.

- **Windowed Analysis:**

Compute GLCM features in sliding windows for local texture maps.

5. Applications

- **Medical Imaging:** Tumor texture analysis in MRI/CT.
- **Remote Sensing:** Land-cover classification (forest vs. urban).
- **Quality Inspection:** Detecting surface defects on fabrics/metals.
- **Biometrics:** Palmprint and fingerprint texture recognition.

3. Common Statistical Features Derived from GLCM

Once you have the normalized matrix P(i,j), several second-order statistics are computed:
Other derived metrics: **Cluster Shade/Prominence, Dissimilarity, Maximum Probability**, etc.

Feature	Formula	Interpretation
Contrast		Measures local intensity variation; higher for edges/rough textures.
Energy (Angular Second Moment)	$\sum_{i,j} P\,(ij)^2,$	Uniformity. High for homogeneous textures.
Homogeneity (Inverse Difference Moment)	(\sum_{i,j} \frac{P(i,j)}{1 +	i - j
Correlation	$\sum_{i,j} \frac{(i-\mu_i)(j-\mu_j)P(i,j)}{\sigma_i\sigma_j}$	Measures how correlated a pixel is to its neighbor.
Entropy	$-\sum_{i,j} P\,(i,j)\log P\,(i,j)$	Randomness of texture.

Mini-Example

Consider a 4×4 image with gray levels 0 and 1:

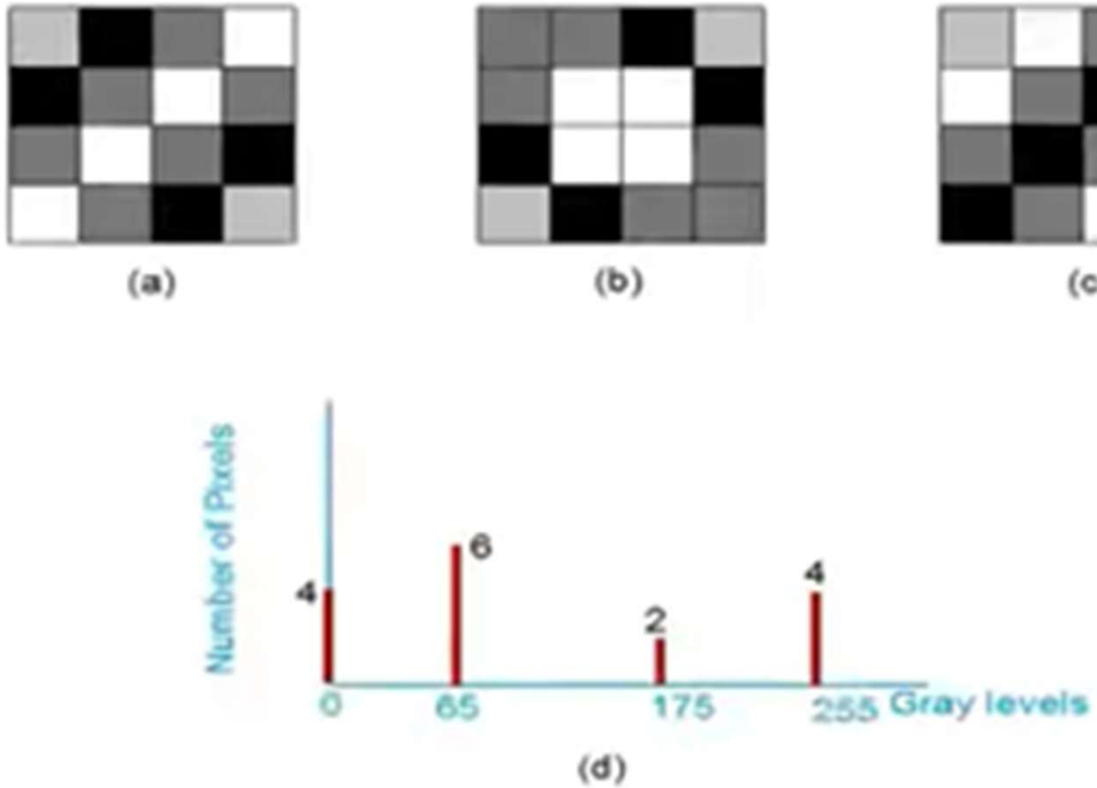
```
0 0 1 1
0 0 1 1
1 1 0 0
1 1 0 0
```

- Distance = 1, Direction = 0° (horizontal neighbors).
- Count pairs:
    - (0,0): 4 times
    - (0,1): 2 times
    - (1,0): 2 times
    - (1,1): 4 times

GLCM =

$$\begin{bmatrix} 4 & 2 \\ 2 & 4 \end{bmatrix}$$

Normalized = each entry /12



(a)

(b)

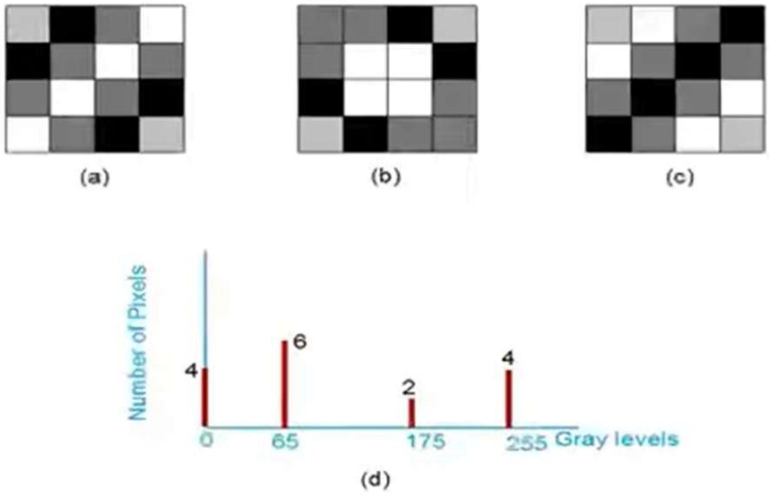
(c)

(d)

Image texture gives us information about the spatial arrangement of color or intensities in an image

- A simple one-dimensional histogram is not useful in characterizing texture

for example, All three images have the Same histogram.



(a) (b) (c)

(d)

- Hence a two-dimensional dependence matrix known as a gray-level co-occurrence matrix is extensively used in texture analysis.
- The co-occurrence matrix captures numerical features of a texture.
- Numerical features calculated from the co-occurrence matrix can be used to represent, classify, and compare textures.

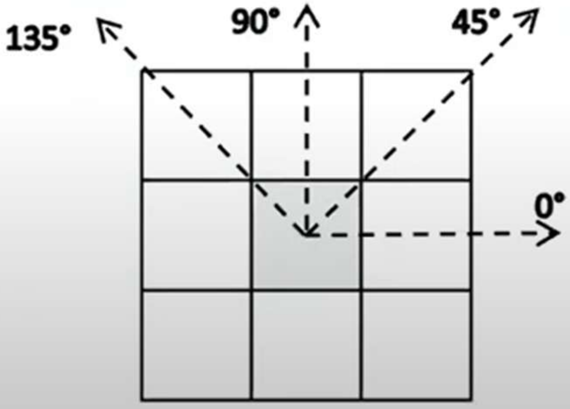
Computation of Co-Occurrence Matrix

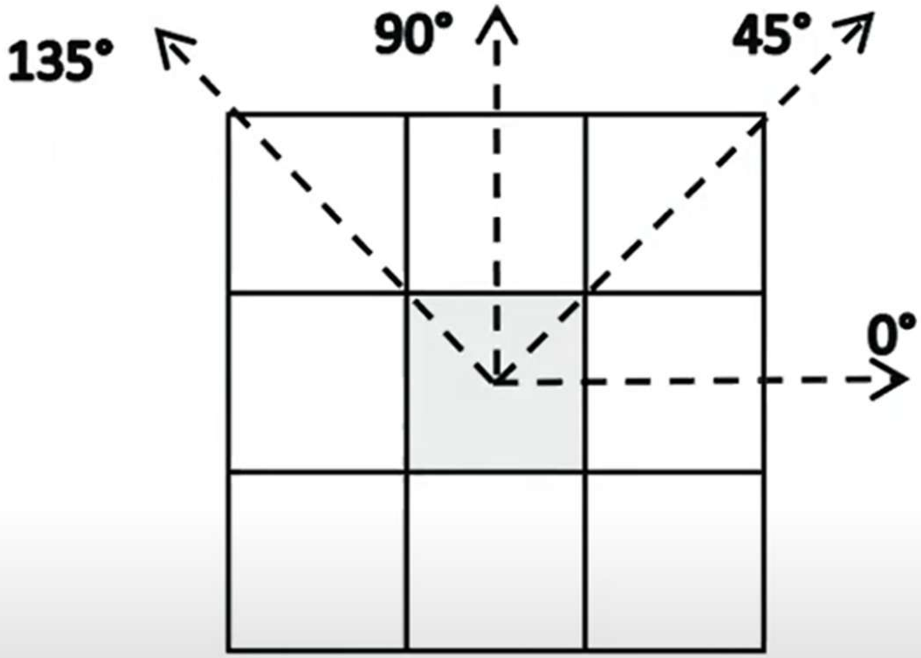
- It has size $N \times N$ (N = Number of gray-values) i.e., the rows & columns represent the set of possible pixel values.

- It is computed based on two parameters:

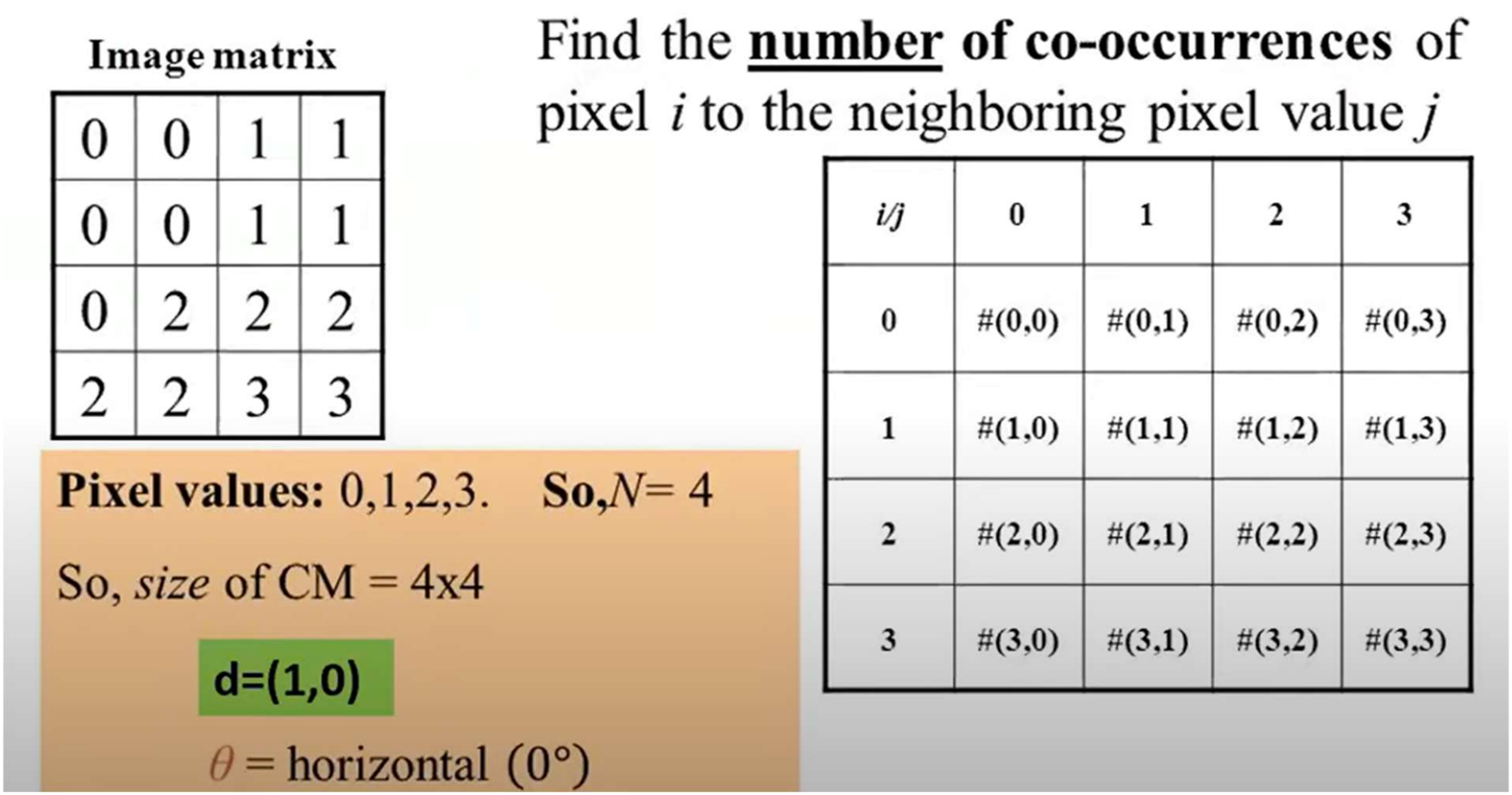
d → Relative **distance** between the pixel pair (measured in pixel number. e.g., 1, 2, …)

θ → Relative **orientation** / rotational angle.





we consider θas horizontal (0°), front diagonal (45°), vertical (90°) and back diagonal (135°)



Find the **number** **of co-occurrences** of pixel *i* to the neighboring pixel value *j*

Image matrix

0	0	1	1
0	0	1	1
0	2	2	2
2	2	3	3

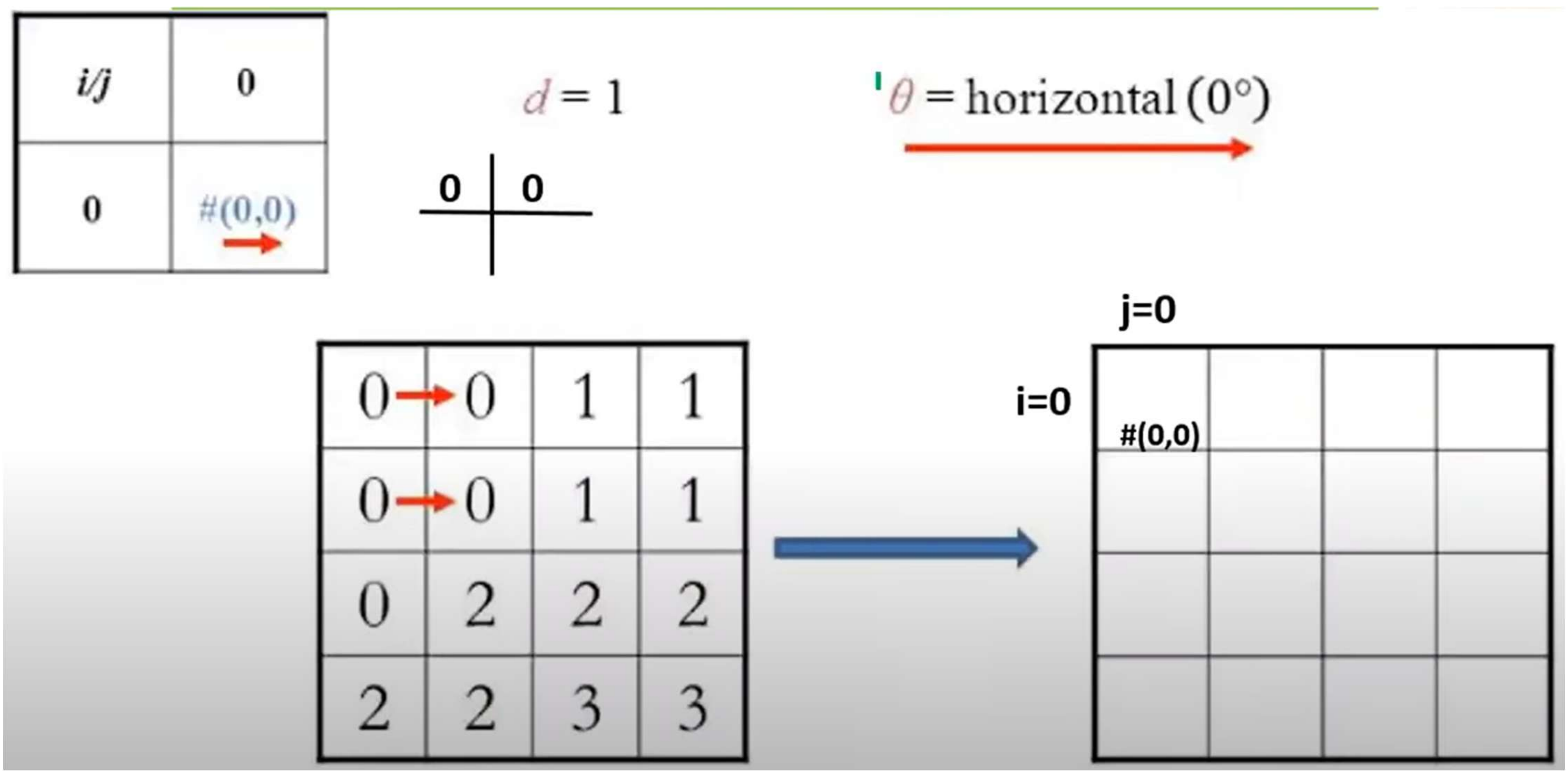
i/j	0	1	2	3
0	#(0,0)	#(0,1)	#(0,2)	#(0,3)
1	#(1,0)	#(1,1)	#(1,2)	#(1,3)
2	#(2,0)	#(2,1)	#(2,2)	#(2,3)
3	#(3,0)	#(3,1)	#(3,2)	#(3,3)

Pixel values: 0,1,2,3. **So,**N= 4

So, *size* of CM = 4x4

d=(1,0)

θ = horizontal (0°)



i/j	0
0	#(0,0) →

$d = 1$

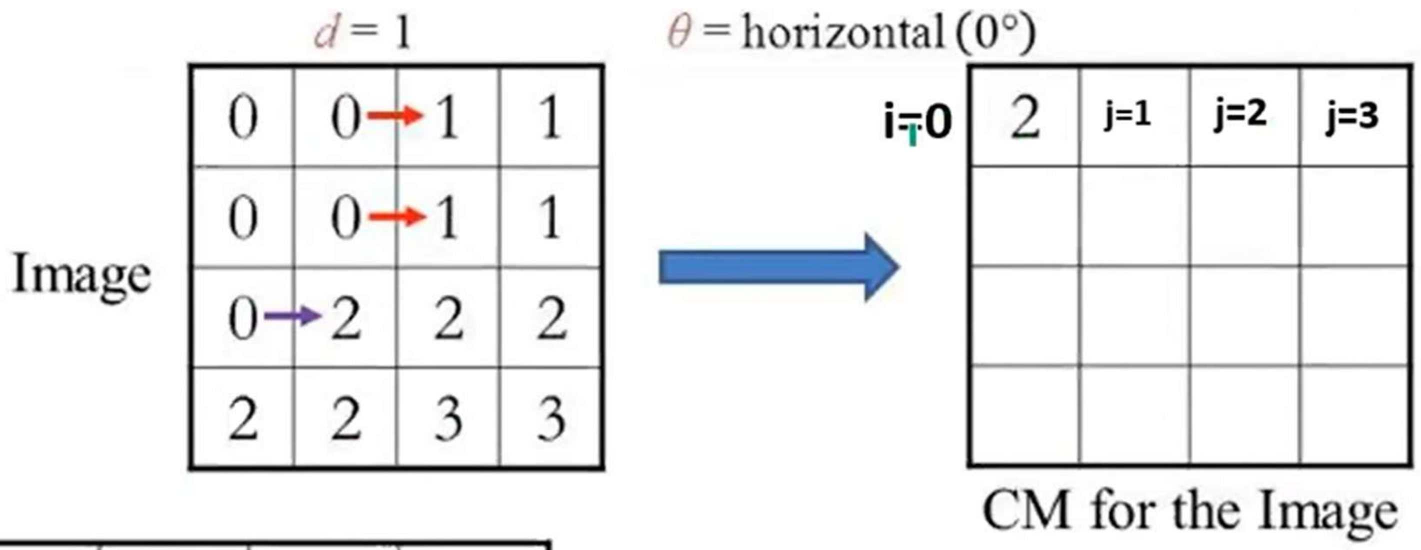
0	0

θ = horizontal (0°) →

0 →	0	1	1
0 →	0	1	1
0	2	2	2
2	2	3	3

→

	j=0			
i=0	#(0,0)			



$d = 1$ θ = horizontal (0°)

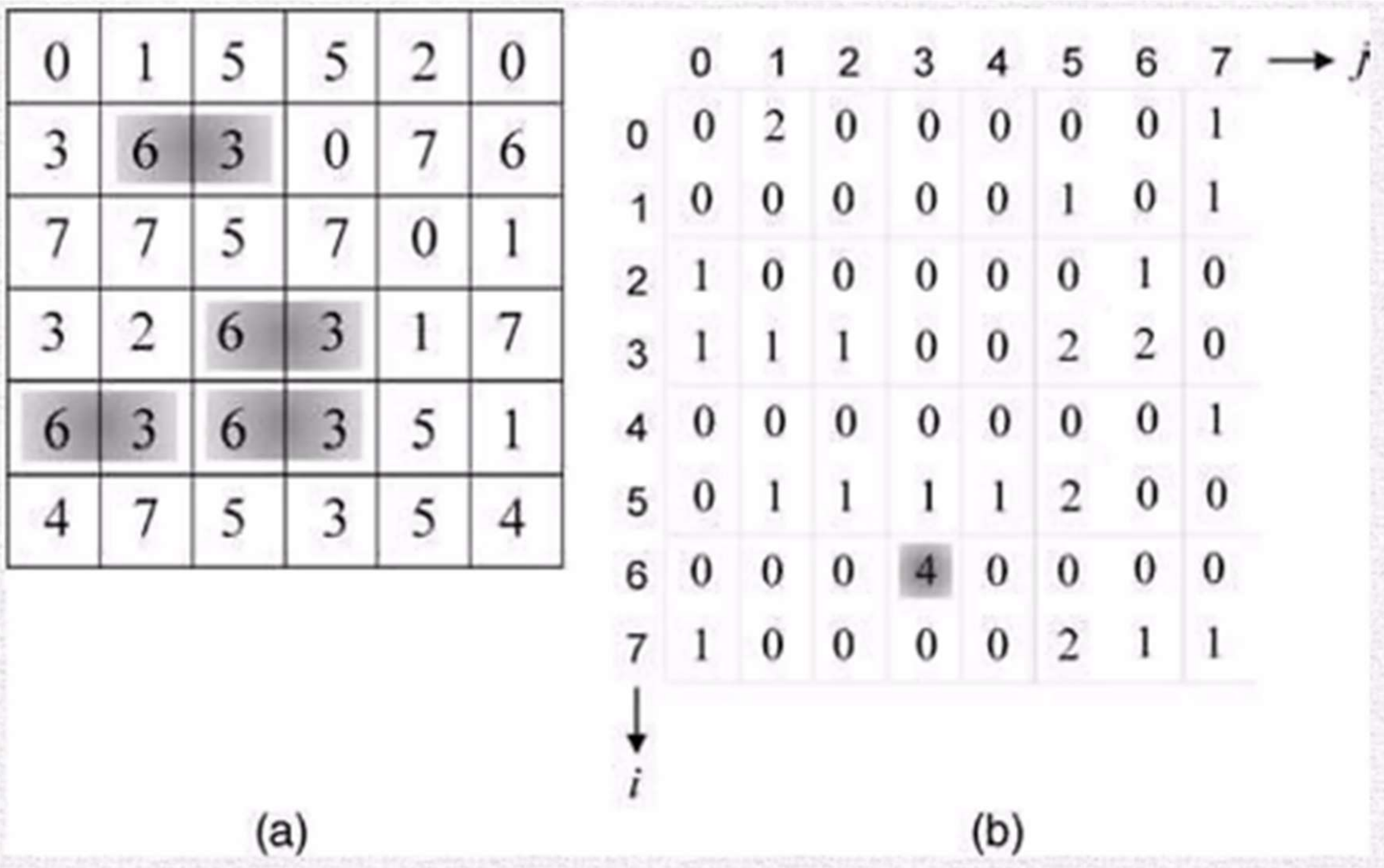
Image

0	0 →	1	1
0	0 →	1	1
0 →	2	2	2
2	2	3	3

i=0	2	j=1	j=2	j=3

CM for the Image

i/j	0	1	2	3
0		#(0,1)	#(0,2)	#(0,3)
1				
2				



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

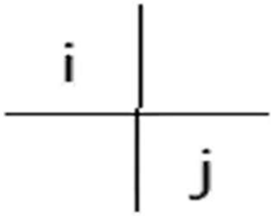
(a)

	0	1	2	3	4	5	6	7	→ *j*
0	0	2	0	0	0	0	0	1	
1	0	0	0	0	0	1	0	1	
2	1	0	0	0	0	0	1	0	
3	1	1	1	0	0	2	2	0	
4	0	0	0	0	0	0	0	1	
5	0	1	1	1	1	2	0	0	
6	0	0	0	4	0	0	0	0	
7	1	0	0	0	0	2	1	1	
↓ *i*									

(b)

Example of Computation

2	1	2	0	1
0	2	1	1	2
0	1	2	2	0
1	2	2	0	1
2	0	1	0	1



$$\frac{1}{16} \times \begin{array}{c|ccc} i \backslash j & 0 & 1 & 2 \\ \hline 0 & 0 & 2 & 2 \\ 1 & 2 & 1 & 2 \\ 2 & 2 & 3 & 2 \end{array}$$

Co-Occurrence Matrix

Consider the simple 5 x 5 image

Gray levels are 0, 1, and 2

N=3

$\boldsymbol{\theta}$ = 135

2	1	2	0	1
0	2	1	1	2
0	1	2	2	0
1	2	2	0	1
2	0	1	0	1

i 0

j 1

$$\frac{1}{16} \times \begin{array}{c|ccc} i \backslash j & 0 & 1 & 2 \\ \hline 0 & 0 & 2 & 2 \\ 1 & 2 & 1 & 2 \\ 2 & 2 & 3 & 2 \end{array}$$

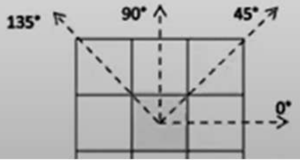
Co-Occurrence Matrix

Consider the simple 5 x 5 image

Gray levels are 0, 1, and 2

N=3

$\boldsymbol{\theta}$ = 135



In a 5 x 5 image there are 16 pairs of pixels which satisfy this spatial separation.
We divide **Co-Occurrence Matrix by 16 to normalize values**

Since there are only three gray levels, P[i, j] is a 3 x 3 matrix

2	1	2	0	1
0	2	1	1	2
0	1	2	2	0
1	2	2	0	1
2	0	1	0	1

i 2

j 1

$$\frac{1}{16} \times \begin{array}{c|ccc} i \backslash j & 0 & 1 & 2 \\ \hline 0 & 0 & 2 & 2 \\ 1 & 2 & 1 & 2 \\ 2 & 2 & 3 & 2 \end{array}$$

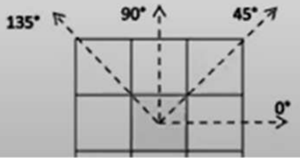
Co-Occurrence Matrix

Consider the simple 5 x 5 image

Gray levels are 0, 1, and 2

N=3

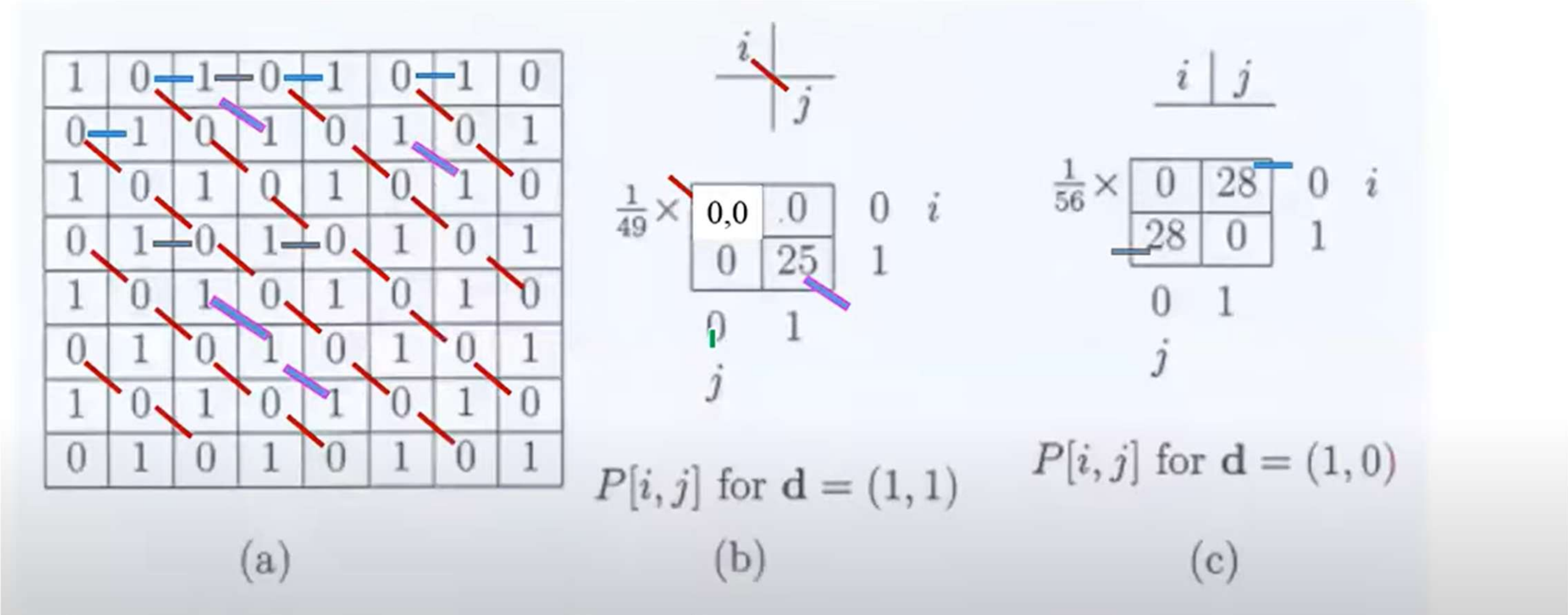
θ = 135



In a 5 x 5 image there are 16 pairs of pixels which satisfy this spatial separation.

We divide **Co-Occurrence Matrix by 16 to normalize values**

Since there are only three gray levels, P[i, j] is a 3 x 3 matrix



1	0	1	0	1	0	1	0
0	1	0	1	0	1	0	1
1	0	1	0	1	0	1	0
0	1	0	1	0	1	0	1
1	0	1	0	1	0	1	0
0	1	0	1	0	1	0	1
1	0	1	0	1	0	1	0
0	1	0	1	0	1	0	1

(a)

A 8 x 8 checkboard

$$\begin{array}{c|c} i & \\ \hline & j \end{array}$$

$$\frac{1}{49}\times \begin{array}{|c|c|cc} \hline 0,0 & 0 & 0 & i \\ \hline 0 & 25 & 1 & \\ \hline 0 & 1 & & \\ j & & & \end{array}$$

$P[i,j]$ for $\mathbf{d}=(1,1)$

(b)

The gray-level co-occurrence

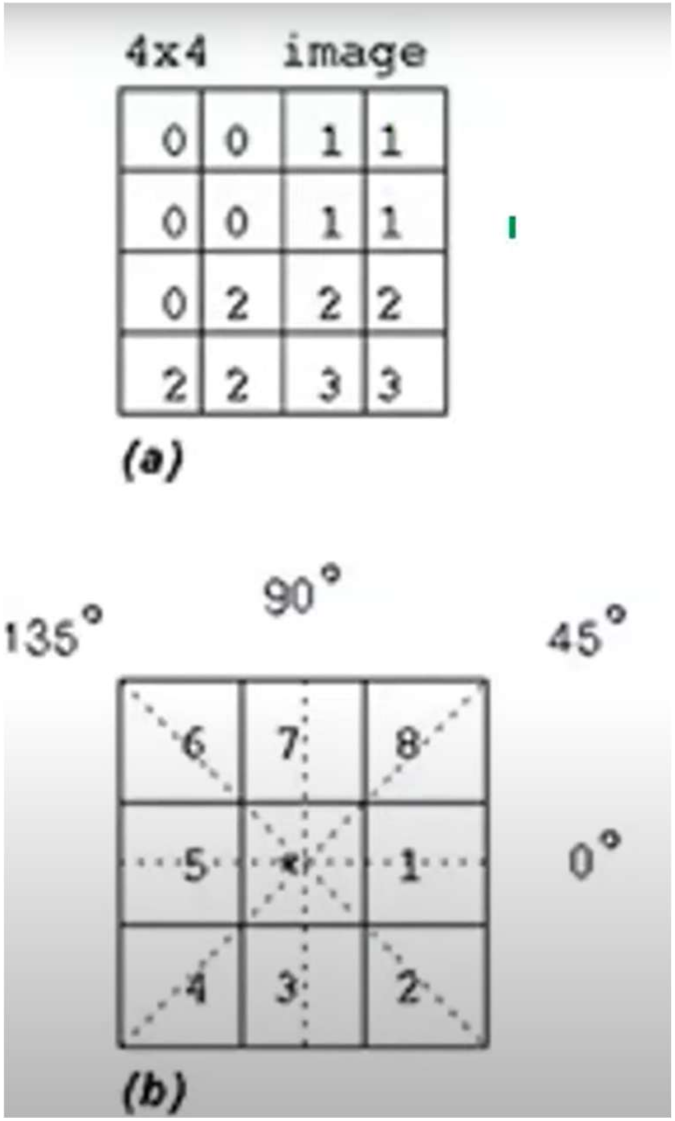
$$\begin{array}{c|c} i & j \\ \hline \end{array}$$

$$\frac{1}{56}\times \begin{array}{|c|c|cc} \hline 0 & 28 & 0 & i \\ \hline 28 & 0 & 1 & \\ \hline 0 & 1 & & \\ j & & & \end{array}$$

$P[i,j]$ for $\mathbf{d}=(1,0)$

(c)

The gray-level co-occurrence

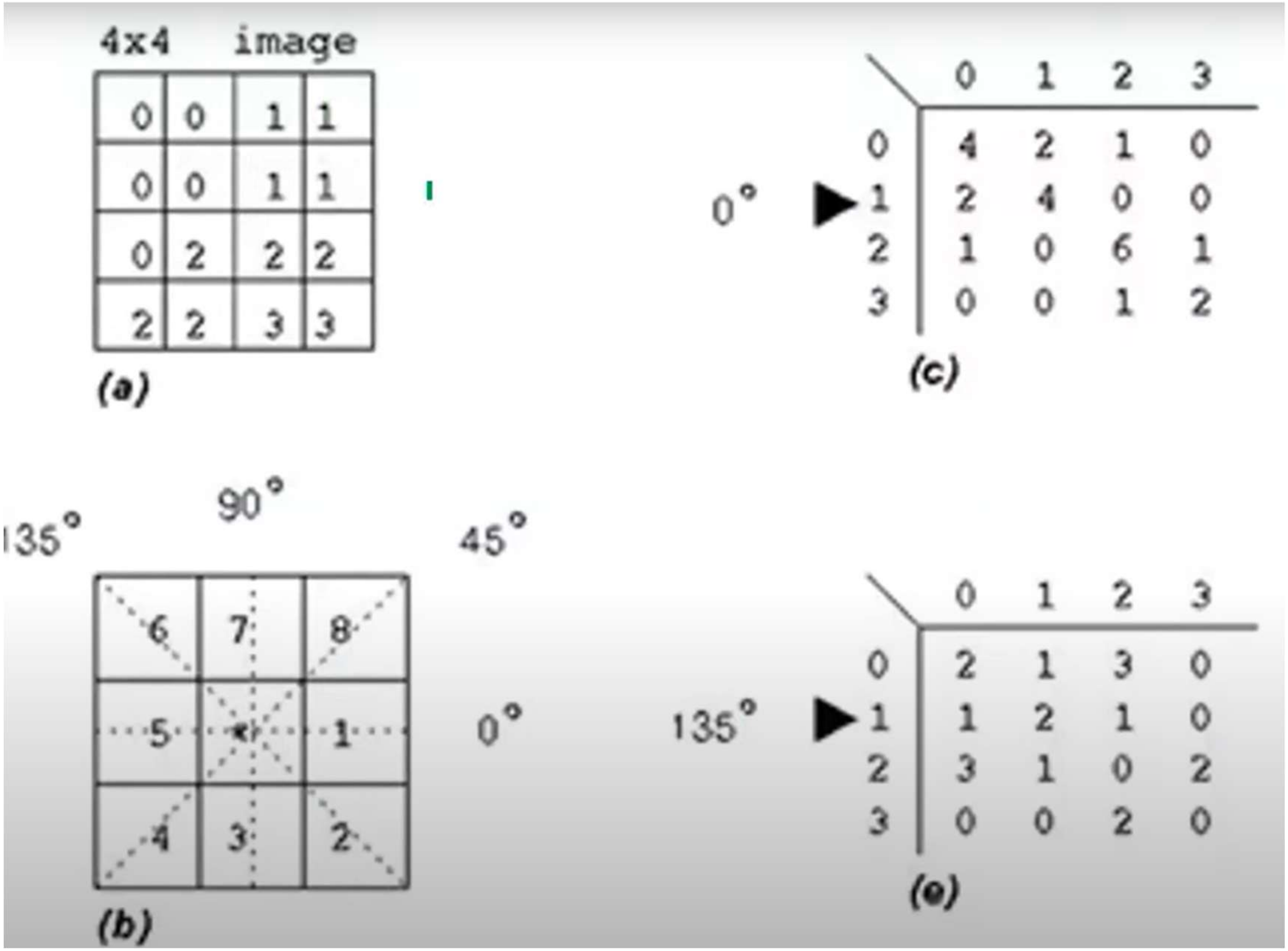


4x4 image

0	0	1	1
0	0	1	1
0	2	2	2
2	2	3	3

(a)

(b)



4x4 image

0	0	1	1
0	0	1	1
0	2	2	2
2	2	3	3

(a)

0° ▶

	0	1	2	3
0	4	2	1	0
1	2	4	0	0
2	1	0	6	1
3	0	0	1	2

(c)

(b)

135° ▶

	0	1	2	3
0	2	1	3	0
1	1	2	1	0
2	3	1	0	2
3	0	0	2	0

(e)

Example: For the 4×4 image below, compute the GLCM for the horizontal right neighbor (offset = (1,0)), normalize it, and calculate **Contrast**, **Energy (and ASM)**, **Homogeneity**, and **Correlation**. Show steps.

Image (gray levels 0–3):

```
I = [[0, 1, 1, 0],
     [2, 1, 0, 0],
     [2, 2, 1, 0],
     [3, 2, 1, 1]]
```

Step 1 — Build (non-symmetric) GLCM for offset (1,0) (right neighbor)

Scan each pixel and its right neighbour (ignore last column). Count occurrences (rows = reference gray level i, columns = neighbour gray level j). Counting gives:

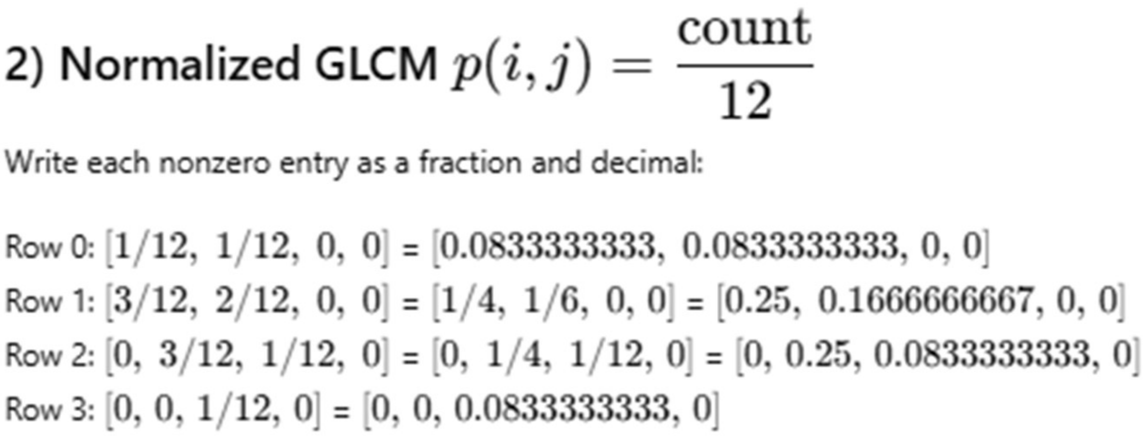
```
GLCM (counts) =
[[1, 1, 0, 0],
 [3, 2, 0, 0],
 [0, 3, 1, 0],
 [0, 0, 1, 0]]
```

Add all counts:

$$1 + 1 + 0 + 0 + 3 + 2 + 0 + 0 + 0$$
$$+ 3 + 1 + 0 + 0 + 0 + 1 + 0 = 12.$$

So total = **12**.

Total pairs = 12.



2) Normalized GLCM $p(i, j) = \frac{\text{count}}{12}$

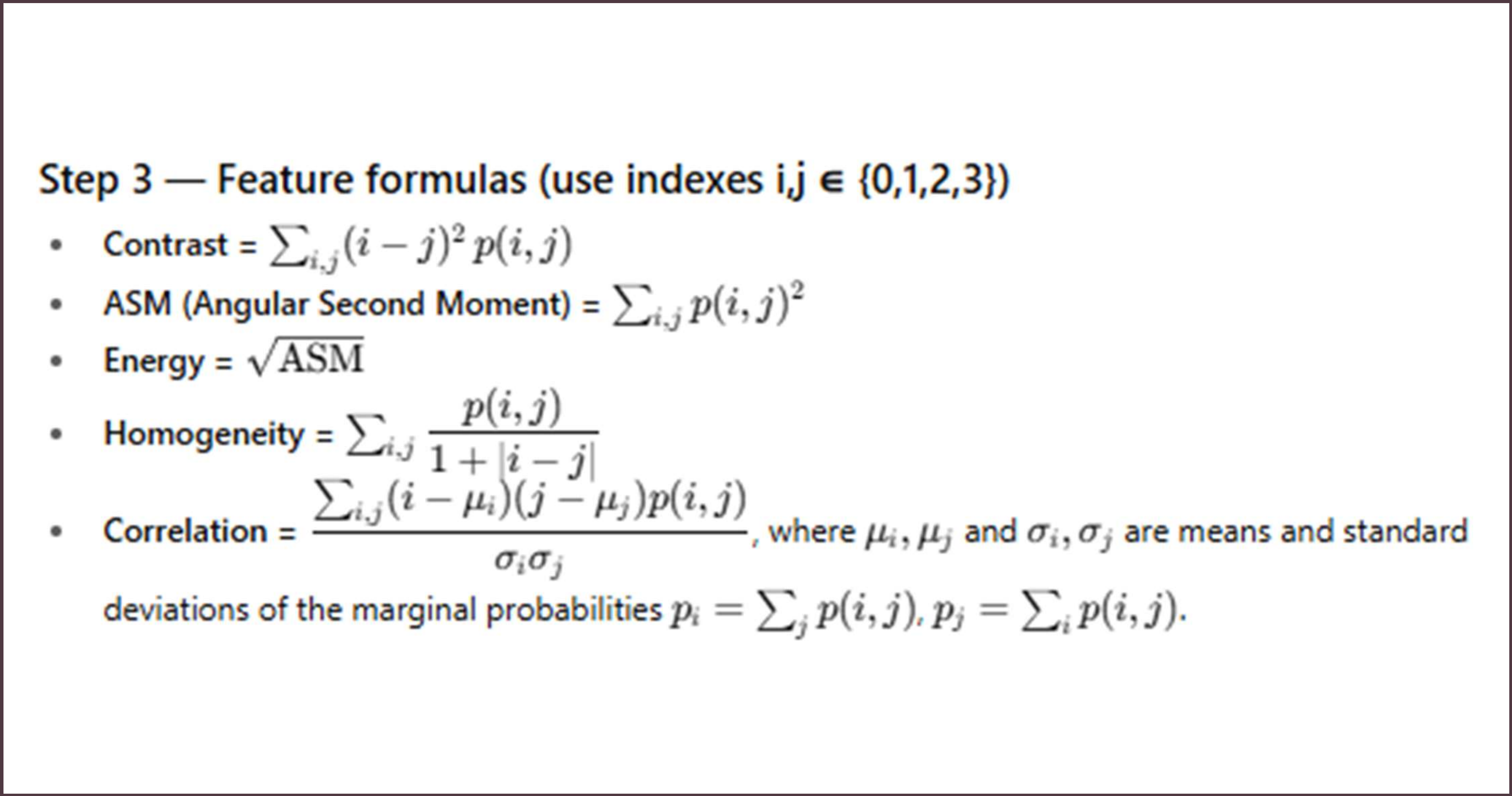
Write each nonzero entry as a fraction and decimal:

Row 0: $[1/12,\ 1/12,\ 0,\ 0]$ = $[0.0833333333,\ 0.0833333333,\ 0,\ 0]$
Row 1: $[3/12,\ 2/12,\ 0,\ 0]$ = $[1/4,\ 1/6,\ 0,\ 0]$ = $[0.25,\ 0.1666666667,\ 0,\ 0]$
Row 2: $[0,\ 3/12,\ 1/12,\ 0]$ = $[0,\ 1/4,\ 1/12,\ 0]$ = $[0,\ 0.25,\ 0.0833333333,\ 0]$
Row 3: $[0,\ 0,\ 1/12,\ 0]$ = $[0,\ 0,\ 0.0833333333,\ 0]$

Step 2 — Normalized to get probability matrix P = G / sum(G)

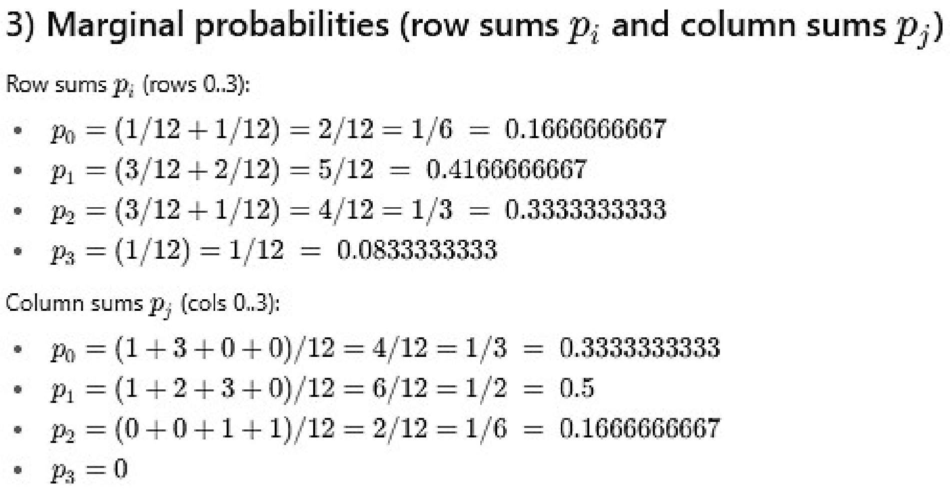
Divide each entry by 12:

```
P =
[[0.08333333, 0.08333333, 0.        , 0.        ],
 [0.25      , 0.16666667, 0.        , 0.        ],
 [0.        , 0.25      , 0.08333333, 0.        ],
 [0.        , 0.        , 0.08333333, 0.        ]]
```



Step 3 — Feature formulas (use indexes i,j ∈ {0,1,2,3})

- Contrast = $\sum_{i,j} (i-j)^2\, p(i,j)$
- ASM (Angular Second Moment) = $\sum_{i,j} p(i,j)^2$
- Energy = $\sqrt{\text{ASM}}$
- Homogeneity = $\sum_{i,j} \dfrac{p(i,j)}{1+|i-j|}$
- Correlation = $\dfrac{\sum_{i,j}(i-\mu_i)(j-\mu_j)p(i,j)}{\sigma_i\sigma_j}$, where μ_i, μ_j and σ_i, σ_j are means and standard deviations of the marginal probabilities $p_i = \sum_j p(i,j)$, $p_j = \sum_i p(i,j)$.



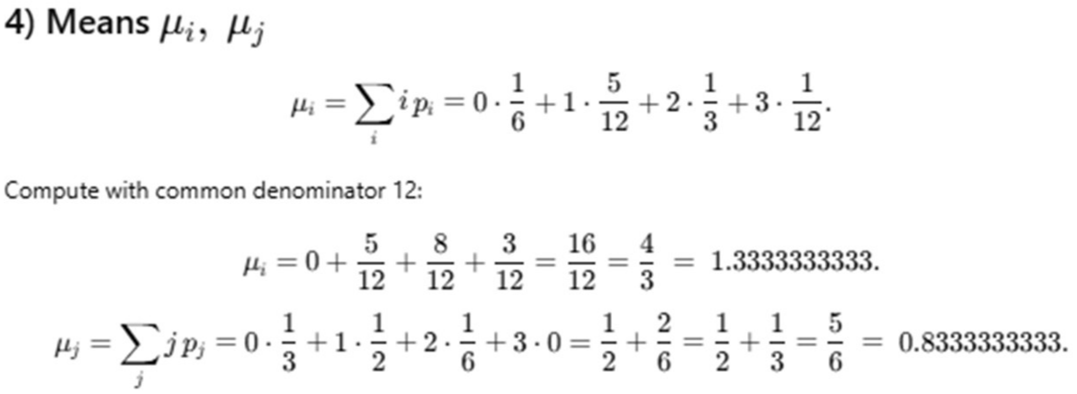
3) Marginal probabilities (row sums p_i and column sums p_j)

Row sums p_i (rows 0..3):

- $p_0 = (1/12 + 1/12) = 2/12 = 1/6 \;=\; 0.1666666667$
- $p_1 = (3/12 + 2/12) = 5/12 \;=\; 0.4166666667$
- $p_2 = (3/12 + 1/12) = 4/12 = 1/3 \;=\; 0.3333333333$
- $p_3 = (1/12) = 1/12 \;=\; 0.0833333333$

Column sums p_j (cols 0..3):

- $p_0 = (1 + 3 + 0 + 0)/12 = 4/12 = 1/3 \;=\; 0.3333333333$
- $p_1 = (1 + 2 + 3 + 0)/12 = 6/12 = 1/2 \;=\; 0.5$
- $p_2 = (0 + 0 + 1 + 1)/12 = 2/12 = 1/6 \;=\; 0.1666666667$
- $p_3 = 0$



4) Means $\mu_i,\ \mu_j$

$$\mu_i = \sum_i i\,p_i = 0 \cdot \frac{1}{6} + 1 \cdot \frac{5}{12} + 2 \cdot \frac{1}{3} + 3 \cdot \frac{1}{12}.$$

Compute with common denominator 12:

$$\mu_i = 0 + \frac{5}{12} + \frac{8}{12} + \frac{3}{12} = \frac{16}{12} = \frac{4}{3} \;=\; 1.3333333333.$$

$$\mu_j = \sum_j j\,p_j = 0 \cdot \frac{1}{3} + 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{6} + 3 \cdot 0 = \frac{1}{2} + \frac{2}{6} = \frac{1}{2} + \frac{1}{3} = \frac{5}{6} \;=\; 0.8333333333.$$

Step 4 — Compute intermediate marginals and stats

Marginal probabilities:

•p_i =row sums = $[0.1666667$，，0.4166667，$0.3333333\ 0.0833333]$

•p_j =column sums = $[0.3333333$，，0.5，$0.1666667\ 0]$

Means:

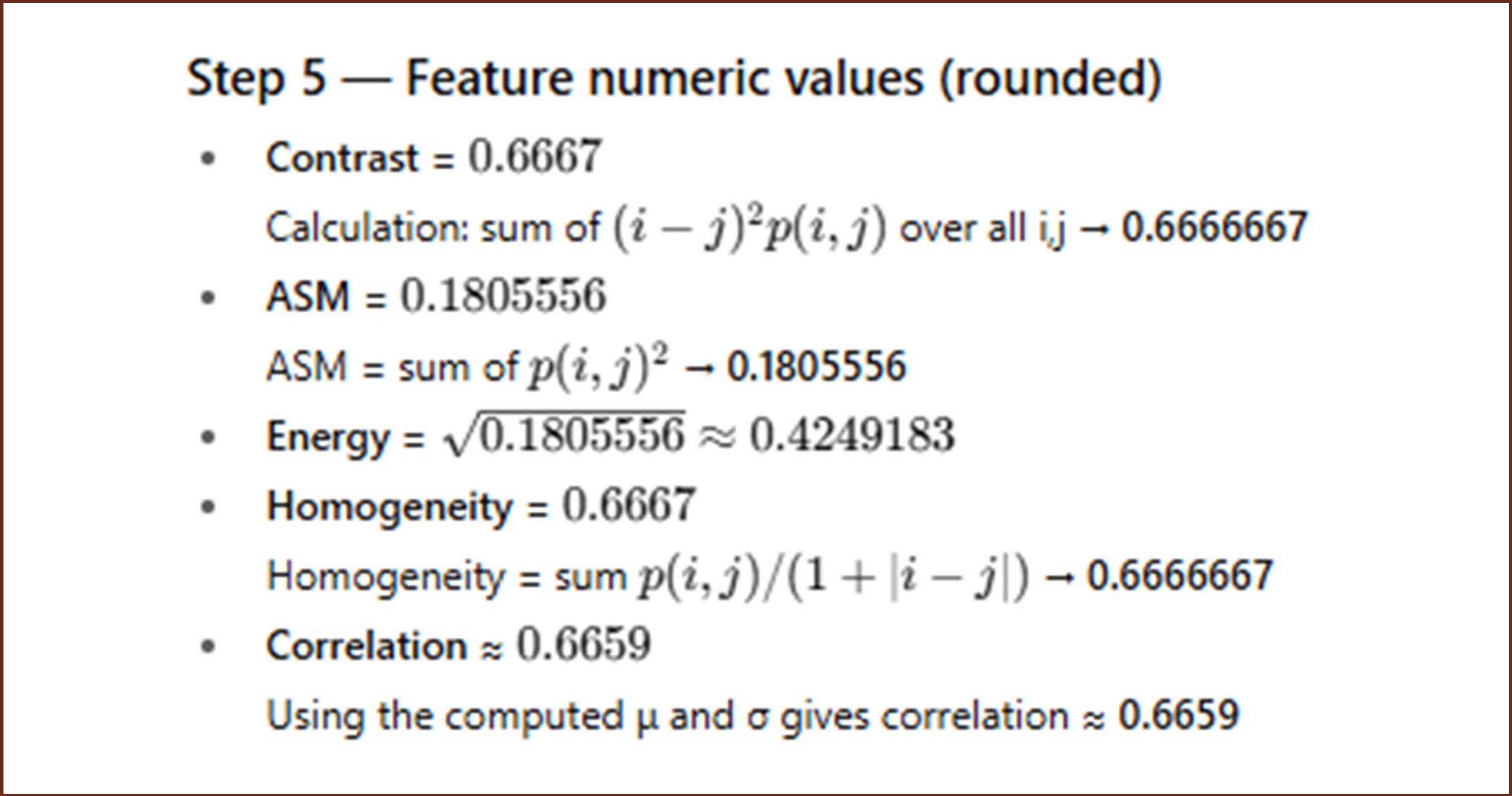
•$\mu_i = \sum i\ p_i = 1.3333333$

•$\mu_j = \sum j\ p_j = 0.8333333$

Standard deviations:

•$\sigma_i \approx 0.8498366$

•$\sigma_j \approx 0.6871843$



Step 5 — Feature numeric values (rounded)

- **Contrast =** 0.6667

  Calculation: sum of $(i-j)^2p(i,j)$ over all i,j → **0.6666667**

- **ASM =** 0.1805556

  ASM = sum of $p(i,j)^2$ → **0.1805556**

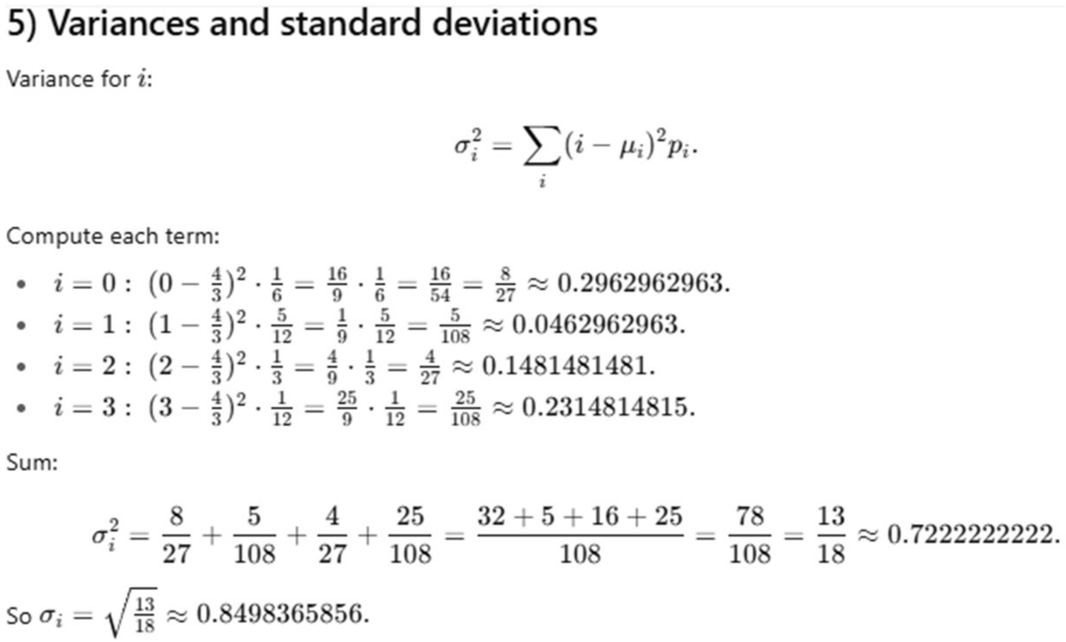
- **Energy =** $\sqrt{0.1805556} \approx 0.4249183$

- **Homogeneity =** 0.6667

  Homogeneity = sum $p(i,j)/(1+|i-j|)$ → **0.6666667**

- **Correlation ≈** 0.6659

  Using the computed μ and σ gives correlation ≈ **0.6659**



5) Variances and standard deviations

Variance for i:

$$\sigma_i^2 = \sum_i (i-\mu_i)^2 p_i.$$

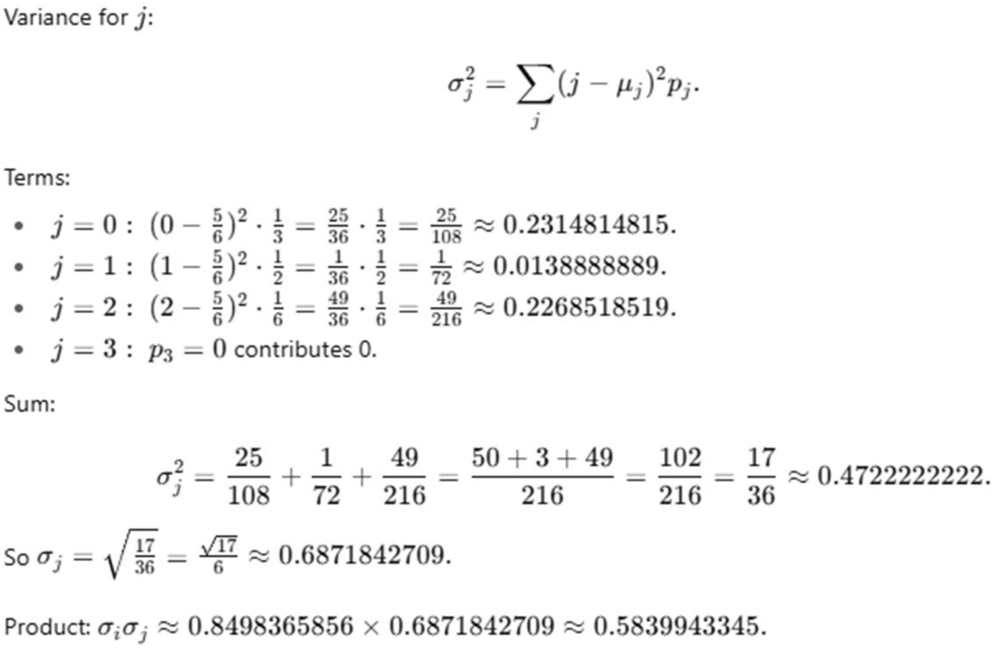
Compute each term:

- $i=0:\ (0-\frac{4}{3})^2\cdot\frac{1}{6}=\frac{16}{9}\cdot\frac{1}{6}=\frac{16}{54}=\frac{8}{27}\approx 0.2962962963.$
- $i=1:\ (1-\frac{4}{3})^2\cdot\frac{5}{12}=\frac{1}{9}\cdot\frac{5}{12}=\frac{5}{108}\approx 0.0462962963.$
- $i=2:\ (2-\frac{4}{3})^2\cdot\frac{1}{3}=\frac{4}{9}\cdot\frac{1}{3}=\frac{4}{27}\approx 0.1481481481.$
- $i=3:\ (3-\frac{4}{3})^2\cdot\frac{1}{12}=\frac{25}{9}\cdot\frac{1}{12}=\frac{25}{108}\approx 0.2314814815.$

Sum:

$$\sigma_i^2=\frac{8}{27}+\frac{5}{108}+\frac{4}{27}+\frac{25}{108}=\frac{32+5+16+25}{108}=\frac{78}{108}=\frac{13}{18}\approx 0.7222222222.$$

So $\sigma_i=\sqrt{\frac{13}{18}}\approx 0.8498365856.$



Variance for j:

$$\sigma_j^2=\sum_j (j-\mu_j)^2 p_j.$$

Terms:

- $j=0:\ (0-\frac{5}{6})^2\cdot\frac{1}{3}=\frac{25}{36}\cdot\frac{1}{3}=\frac{25}{108}\approx 0.2314814815.$
- $j=1:\ (1-\frac{5}{6})^2\cdot\frac{1}{2}=\frac{1}{36}\cdot\frac{1}{2}=\frac{1}{72}\approx 0.0138888889.$
- $j=2:\ (2-\frac{5}{6})^2\cdot\frac{1}{6}=\frac{49}{36}\cdot\frac{1}{6}=\frac{49}{216}\approx 0.2268518519.$
- $j=3:\ p_3=0$ contributes 0.

Sum:

$$\sigma_j^2=\frac{25}{108}+\frac{1}{72}+\frac{49}{216}=\frac{50+3+49}{216}=\frac{102}{216}=\frac{17}{36}\approx 0.4722222222.$$

So $\sigma_j=\sqrt{\frac{17}{36}}=\frac{\sqrt{17}}{6}\approx 0.6871842709.$

Product: $\sigma_i\sigma_j\approx 0.8498365856\times 0.6871842709\approx 0.5839943345.$

6) Contrast

Contrast $= \sum_{i,j}(i-j)^2 p(i,j)$. Only nonzero $p(i,j)$ contribute. Compute each:

Nonzero positions and contributions:

- (0,0): $p = \frac{1}{12}$, $(0-0)^2 = 0 \Rightarrow 0.$
- (0,1): $p = \frac{1}{12}$, $(0-1)^2 = 1 \Rightarrow \frac{1}{12}.$
- (1,0): $p = \frac{1}{4}$, $(1-0)^2 = 1 \Rightarrow \frac{1}{4}.$
- (1,1): $p = \frac{1}{6}$, $(1-1)^2 = 0 \Rightarrow 0.$
- (2,1): $p = \frac{1}{4}$, $(2-1)^2 = 1 \Rightarrow \frac{1}{4}.$
- (2,2): $p = \frac{1}{12}$, $(2-2)^2 = 0 \Rightarrow 0.$
- (3,2): $p = \frac{1}{12}$, $(3-2)^2 = 1 \Rightarrow \frac{1}{12}.$

Sum:

$$\text{Contrast} = \frac{1}{12} + \frac{1}{4} + \frac{1}{4} + \frac{1}{12} = \frac{1+3+3+1}{12} = \frac{8}{12} = \frac{2}{3} \approx 0.6666666667.$$

7) ASM and Energy

ASM $= \sum p(i,j)^2$. Nonzero p's: four times $1/12$, two times $1/4$, one time $1/6$.

$$\text{ASM} = 4 \cdot \left(\frac{1}{12}\right)^2 + 2 \cdot \left(\frac{1}{4}\right)^2 + 1 \cdot \left(\frac{1}{6}\right)^2 = 4 \cdot \frac{1}{144} + 2 \cdot \frac{1}{16} + \frac{1}{36}.$$

Compute:

$$= \frac{4}{144} + \frac{2}{16} + \frac{1}{36} = \frac{1}{36} + \frac{1}{8} + \frac{1}{36} = \frac{1}{8} + \frac{2}{36} = \frac{1}{8} + \frac{1}{18} = \frac{9+4}{72} = \frac{13}{72} \approx 0.1805555556.$$

Energy $= \sqrt{\text{ASM}} = \sqrt{\frac{13}{72}} \approx 0.4249183480.$

8) Homogeneity

Homogeneity $= \sum \dfrac{p(i,j)}{1+|i-j|}$.

Compute each nonzero:

- (0,0): $p = \frac{1}{12},\ |0-0| = 0 \Rightarrow$ contrib $\frac{1}{12}$.
- (0,1): $p = \frac{1}{12},\ |0-1| = 1 \Rightarrow$ contrib $\frac{1}{12} \cdot \frac{1}{2} = \frac{1}{24}$.
- (1,0): $p = \frac{1}{4},\ |1-0| = 1 \Rightarrow$ contrib $\frac{1}{4} \cdot \frac{1}{2} = \frac{1}{8}$.
- (1,1): $p = \frac{1}{6},\ |1-1| = 0 \Rightarrow$ contrib $\frac{1}{6}$.
- (2,1): $p = \frac{1}{4},\ |2-1| = 1 \Rightarrow$ contrib $\frac{1}{8}$.
- (2,2): $p = \frac{1}{12},\ |2-2| = 0 \Rightarrow$ contrib $\frac{1}{12}$.
- (3,2): $p = \frac{1}{12},\ |3-2| = 1 \Rightarrow$ contrib $\frac{1}{24}$.

Sum groups:

- denom =1 terms: $\frac{1}{12} + \frac{1}{6} + \frac{1}{12} = \frac{2}{12} + \frac{2}{12} = \frac{4}{12}$
- denom =2 terms: $\frac{1}{24} + \frac{1}{8} + \frac{1}{8} + \frac{1}{24} = \frac{1}{12} + \frac{1}{4} =$

Total homogeneity $= \frac{1}{3} + \frac{1}{3} = \frac{2}{3} \approx 0.6666666667.$

9) Correlation

Correlation numerator:

$$N = \sum_{i,j} (i-\mu_i)(j-\mu_j)\, p(i,j).$$

We compute only nonzero p's. Recall $\mu_i = 4/3,\ \mu_j = 5/6.$

Compute each term exactly:

1. (0,0): $p = \frac{1}{12}.\ (0-\frac{4}{3})(0-\frac{5}{6}) = (-\frac{4}{3})(-\frac{5}{6}) = \frac{20}{18} = \frac{10}{9}.$
   Contribution = $\frac{10}{9} \cdot \frac{1}{12} = \frac{10}{108} = \frac{5}{54}.$
2. (0,1): $p = \frac{1}{12}.\ (0-\frac{4}{3})(1-\frac{5}{6}) = (-\frac{4}{3})(\frac{1}{6}) = -\frac{4}{18} = -\frac{2}{9}.$
   Contribution = $-\frac{2}{9} \cdot \frac{1}{12} = -\frac{2}{108} = -\frac{1}{54}.$

Denominator for correlation is $\sigma_i\sigma_j$ (computed above) $\approx 0.5839943345.$

Thus

$$\text{Correlation} = \frac{N}{\sigma_i\sigma_j} = \frac{7/18}{0.5839943345} \approx 0.6659120918.$$

(Exact symbolic form: $\dfrac{7/18}{\sqrt{221/648}}$, since $\sigma_i\sigma_j = \sqrt{\frac{13}{18} \cdot \frac{17}{36}} = \sqrt{\frac{221}{648}}$.)

5. (2,1): $p = \frac{1}{4}.\ (2-\frac{4}{3})(1-\frac{5}{6}) = (\frac{2}{3})(\frac{1}{6}) = \frac{1}{9}.$
   Contribution = $\frac{1}{9} \cdot \frac{1}{4} = \frac{1}{36}.$
6. (2,2): $p = \frac{1}{12}.\ (2-\frac{4}{3})(2-\frac{5}{6}) = (\frac{2}{3})(\frac{7}{6}) = \frac{14}{18} = \frac{7}{9}.$
   Contribution = $\frac{7}{9} \cdot \frac{1}{12} = \frac{7}{108}.$
7. (3,2): $p = \frac{1}{12}.\ (3-\frac{4}{3})(2-\frac{5}{6}) = (\frac{5}{3})(\frac{7}{6}) = \frac{35}{18}.$
   Contribution = $\frac{35}{18} \cdot \frac{1}{12} = \frac{35}{216}.$

Now sum these contributions (use denominator 216 to combine):

- $\frac{5}{54} = \frac{20}{216}$
- $-\frac{1}{54} = -\frac{4}{216}$
- $\frac{5}{72} = \frac{15}{216}$
- $-\frac{1}{108} = -\frac{2}{216}$
- $\frac{6}{216}$
- $= \frac{14}{216}$
- $= \frac{35}{216}$

erators: $20 - 4 + 15 - 2 + 6 + 14 + 35 = 84.$

$\frac{84}{216} = \frac{7}{18} \approx 0.3888888889.$

Final compact results (fraction & decimal)

- **Contrast = $\frac{2}{3}$ = 0.6666666667**
- **ASM = $\frac{13}{72}$ = 0.1805555556**
- **Energy = $\sqrt{\frac{13}{72}}$ ≈ 0.4249183480**
- **Homogeneity = $\frac{2}{3}$ = 0.6666666667**
- **Correlation ≈ 0.6659120918**

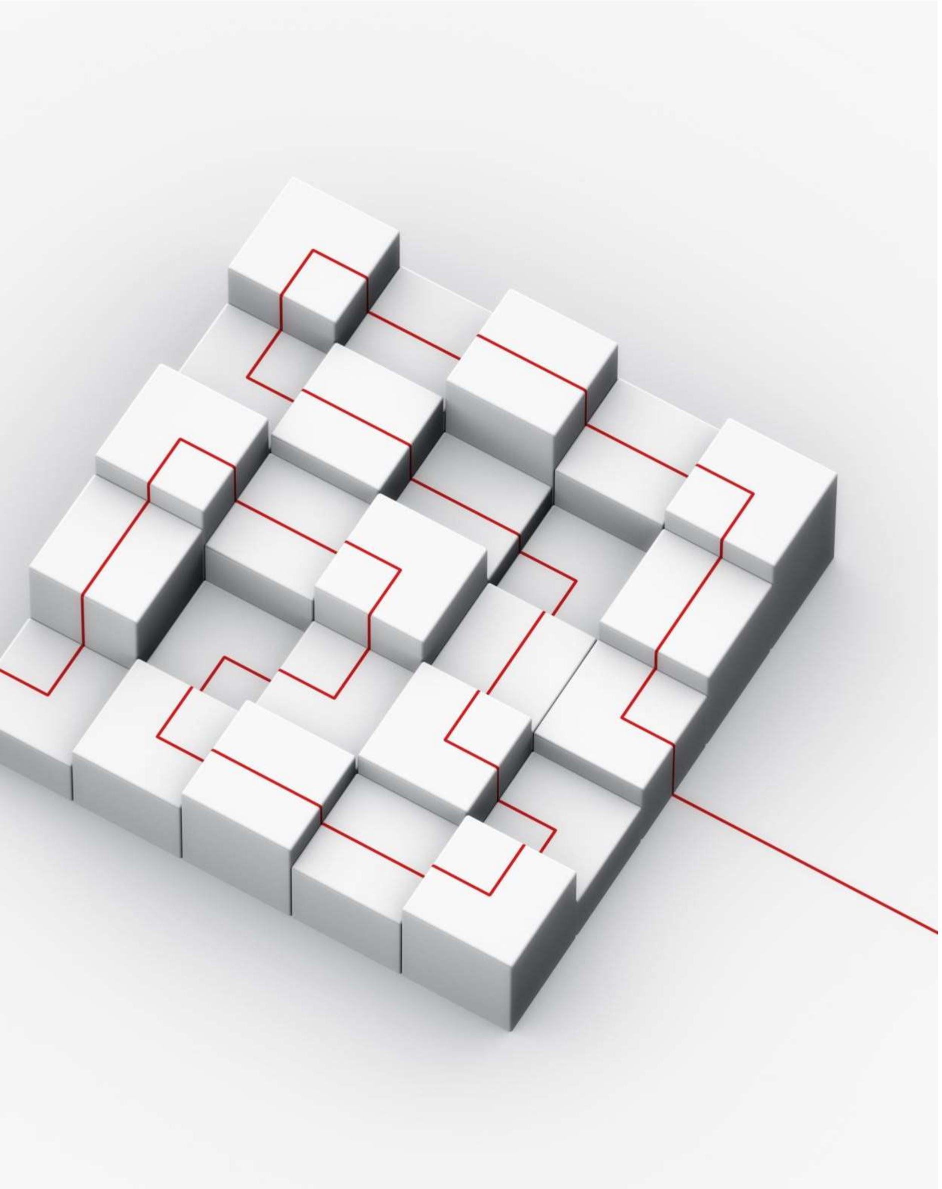
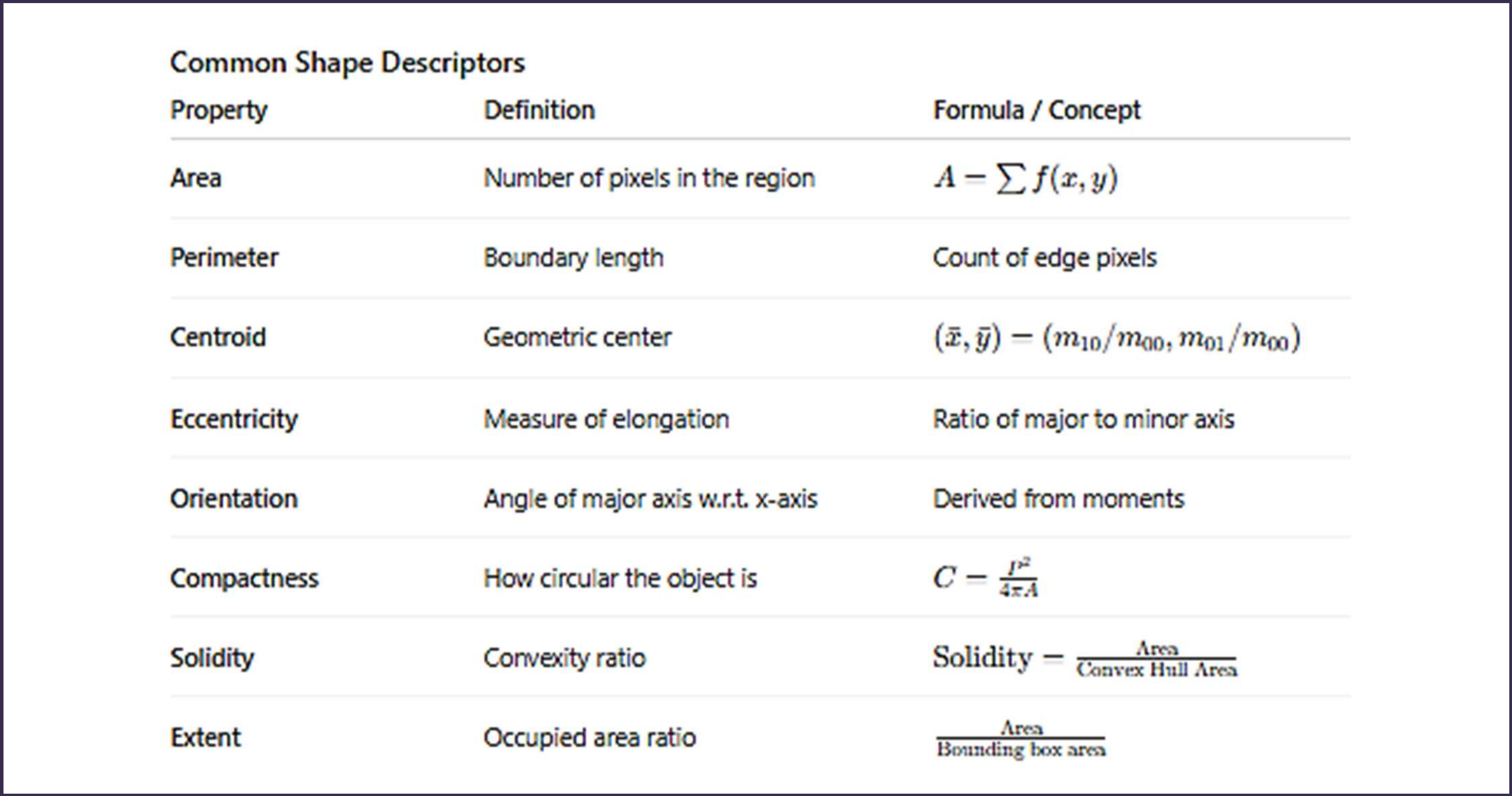


Image Feature Extraction (Shape-Based Features)

Introduction

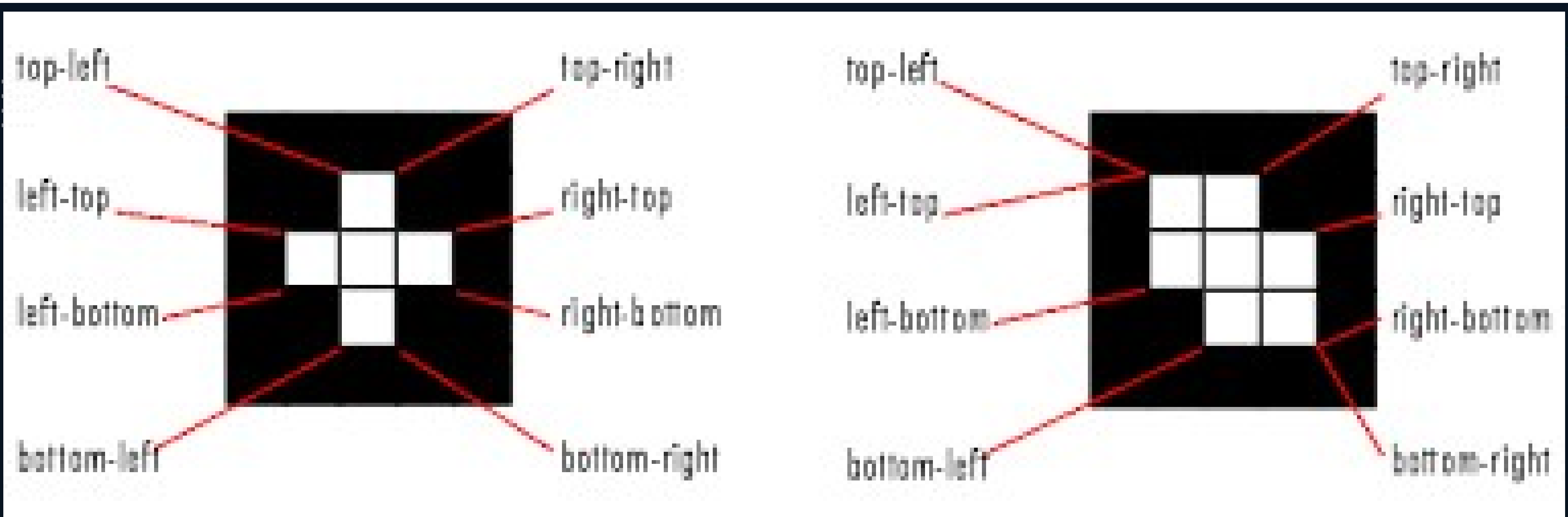
Feature extraction is a process of transforming image data into a set of measurable attributes that can be used for analysis, classification, and recognition.

Shape-based feature extraction focuses on **geometric and structural properties** of objects within an image.



Common Shape Descriptors

Property	Definition	Formula / Concept
Area	Number of pixels in the region	$A = \sum f(x, y)$
Perimeter	Boundary length	Count of edge pixels
Centroid	Geometric center	$(\bar{x}, \bar{y}) = (m_{10}/m_{00}, m_{01}/m_{00})$
Eccentricity	Measure of elongation	Ratio of major to minor axis
Orientation	Angle of major axis w.r.t. x-axis	Derived from moments
Compactness	How circular the object is	$C = \frac{P^2}{4\pi A}$
Solidity	Convexity ratio	$\text{Solidity} = \frac{\text{Area}}{\text{Convex Hull Area}}$
Extent	Occupied area ratio	$\frac{\text{Area}}{\text{Bounding box area}}$



1. Area

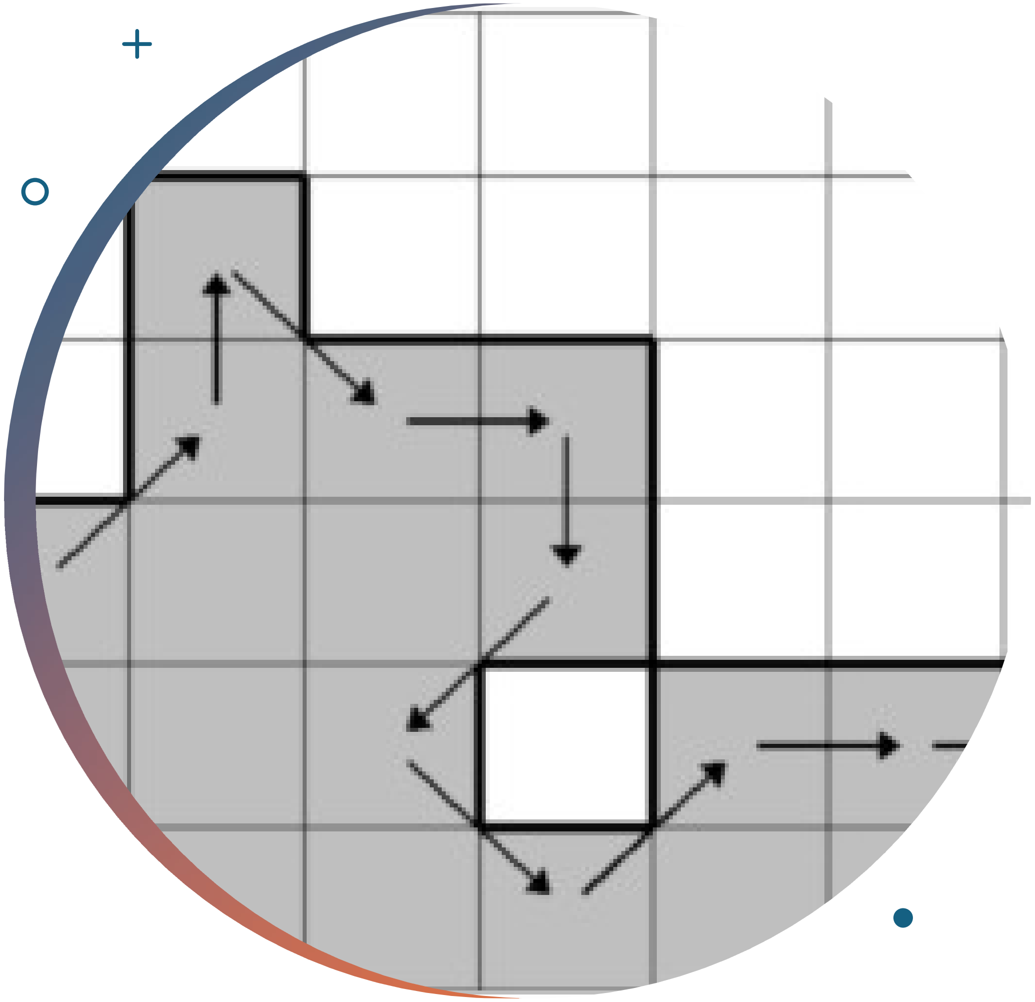
Definition: Number of pixels inside the object or region.

Formula:

$$A = \sum f\,(x, y)$$

Meaning: It tells how large the object is in the image.

Real-life example: If you have an image of a leaf, the *Area* represents how many pixels belong to the leaf — i.e., how big the leaf is.

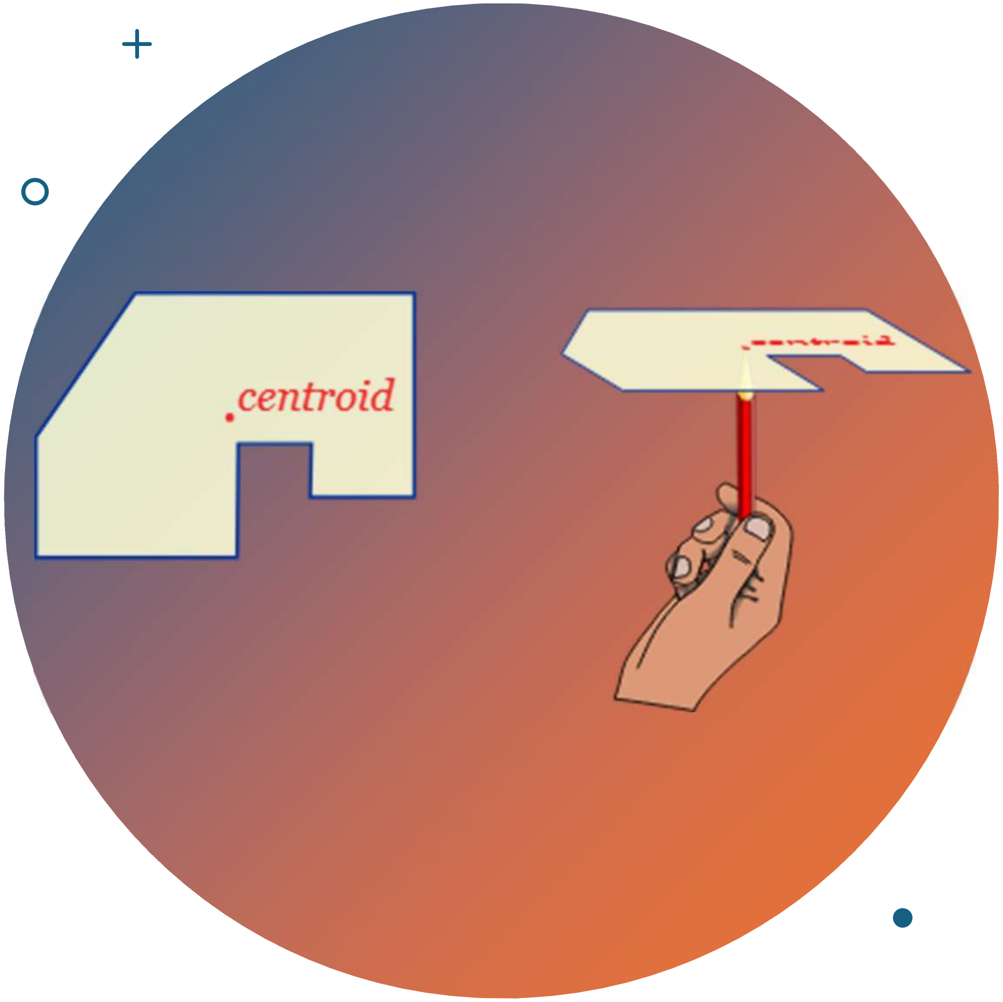


2. Perimeter

Definition: Total length of the boundary pixels of the object.

Meaning: It measures how long the outer edge of the shape is.

Real-life example: The circular edge length of a coin — a larger coin will have a longer perimeter.



3. Centroid

Definition: The geometric center of the object.

Formula:

$$(\bar{x}, \bar{y}) = \left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}}\right)$$

Meaning: It is the balance point or the average location of all pixels in the object.

Real-life example: If you balance a leaf on a needle, the point where it perfectly balances is its centroid.

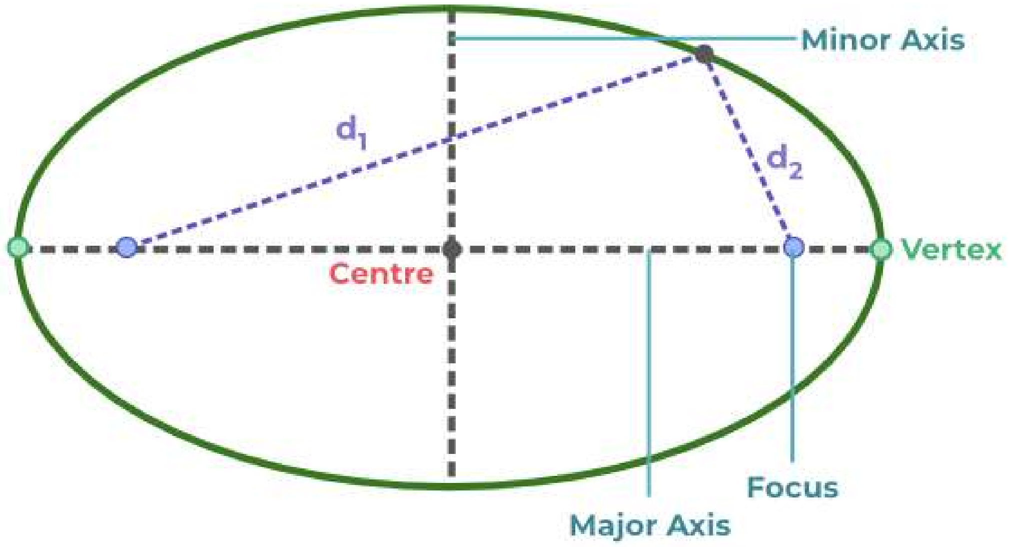
4. Eccentricity

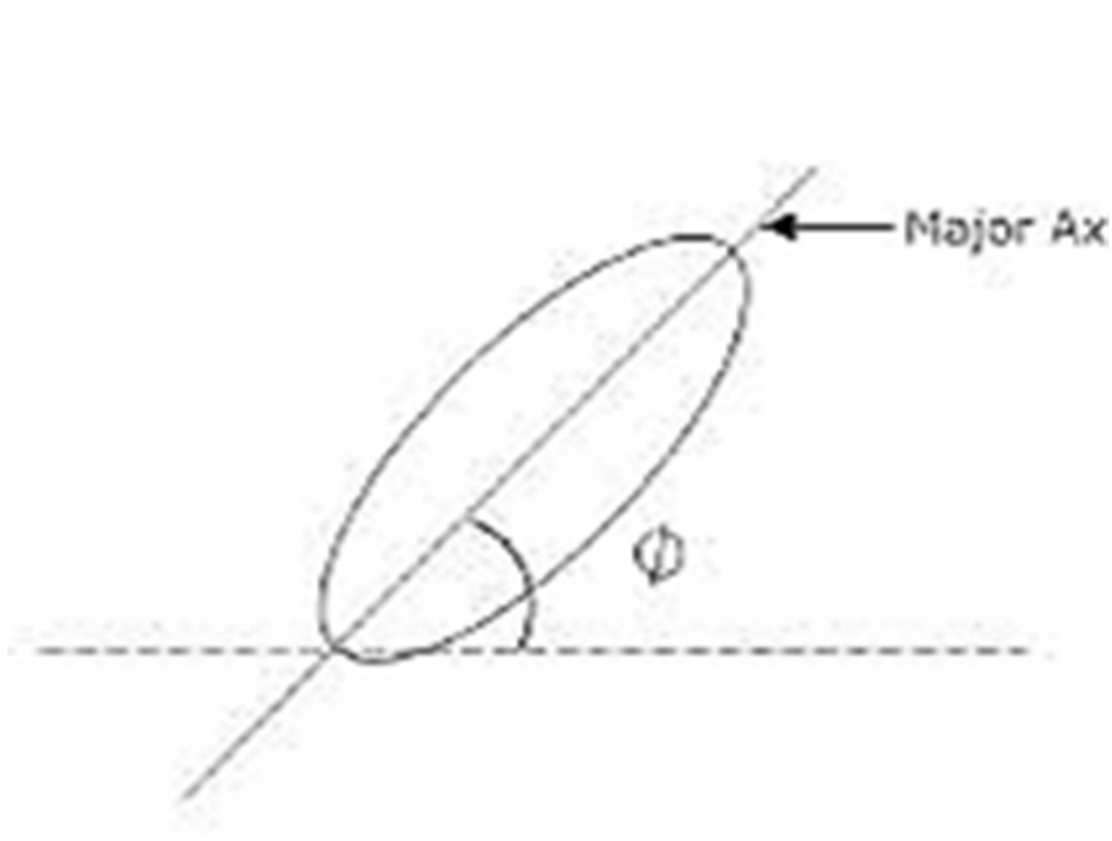
Definition: A measure of elongation — the ratio of the major axis to the minor axis.

Meaning:

- Perfectly circular shape → Eccentricity ≈ 0
- Very long or stretched shape → Eccentricity ≈ 1

Real-life example: A tennis ball (round) has low eccentricity, while a badminton racket (elongated) has high eccentricity.





5. Orientation

Definition: The angle between the object's major axis and the x-axis.

Meaning: It indicates how much the object is tilted or rotated in the image.

Real-life example: If a leaf in an image is leaning toward the right, its orientation gives the tilt angle of that leaf.

6. Compactness

Definition: Describes how circular or compact the object is.

Formula:

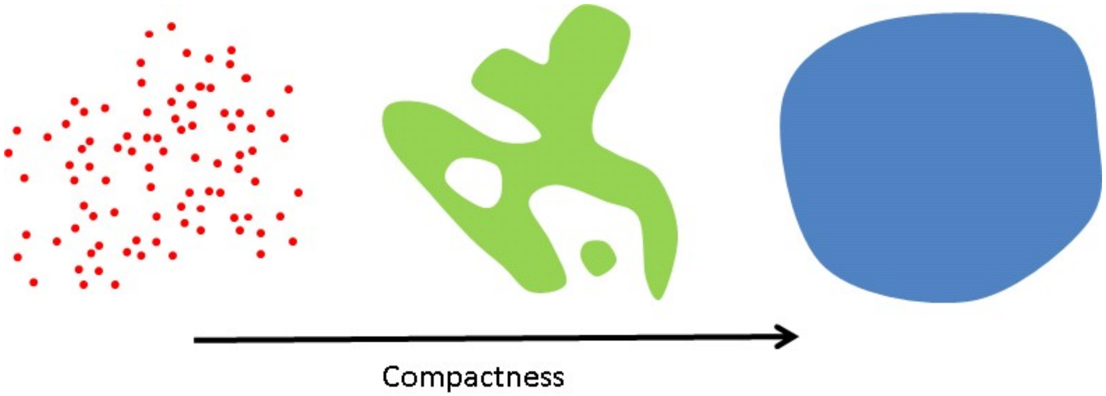
$$C = \frac{P^2}{4\pi A}$$

Meaning:

- For a perfect circle → Compactness ≈ 1
- For irregular or elongated shapes → Compactness > 1

Real-life example:

A ball has compactness close to 1, while a star-shaped object has a higher value.



7. Solidity

Definition: Ratio of the object’s area to the area of its convex hull.

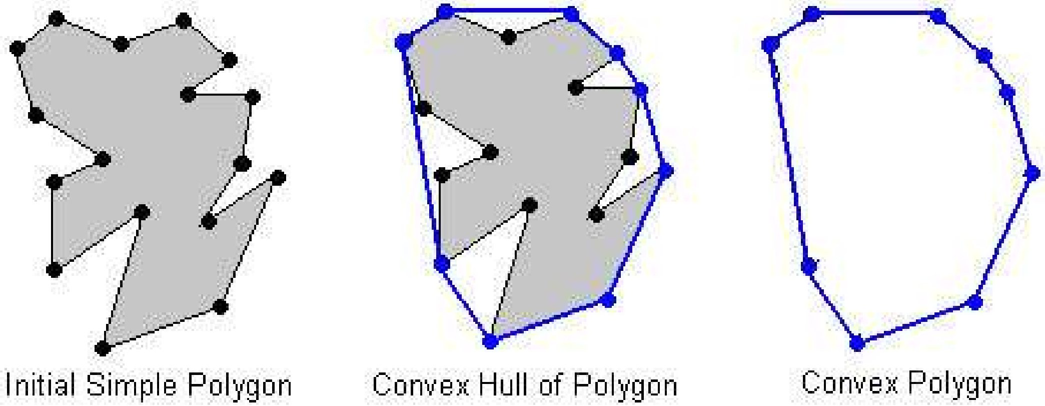
Formula:

$$\text{Solidity} = \frac{Area}{Convex\ Hull\ Area}$$

Meaning: It measures how “solid” or filled the object is.

Real-life example:

A completely filled circle → Solidity ≈ 1,
A ring shape with a hole inside → Solidity < 1.



8. Extent

Definition: Ratio of the object’s area to the area of its bounding box.

Formula:

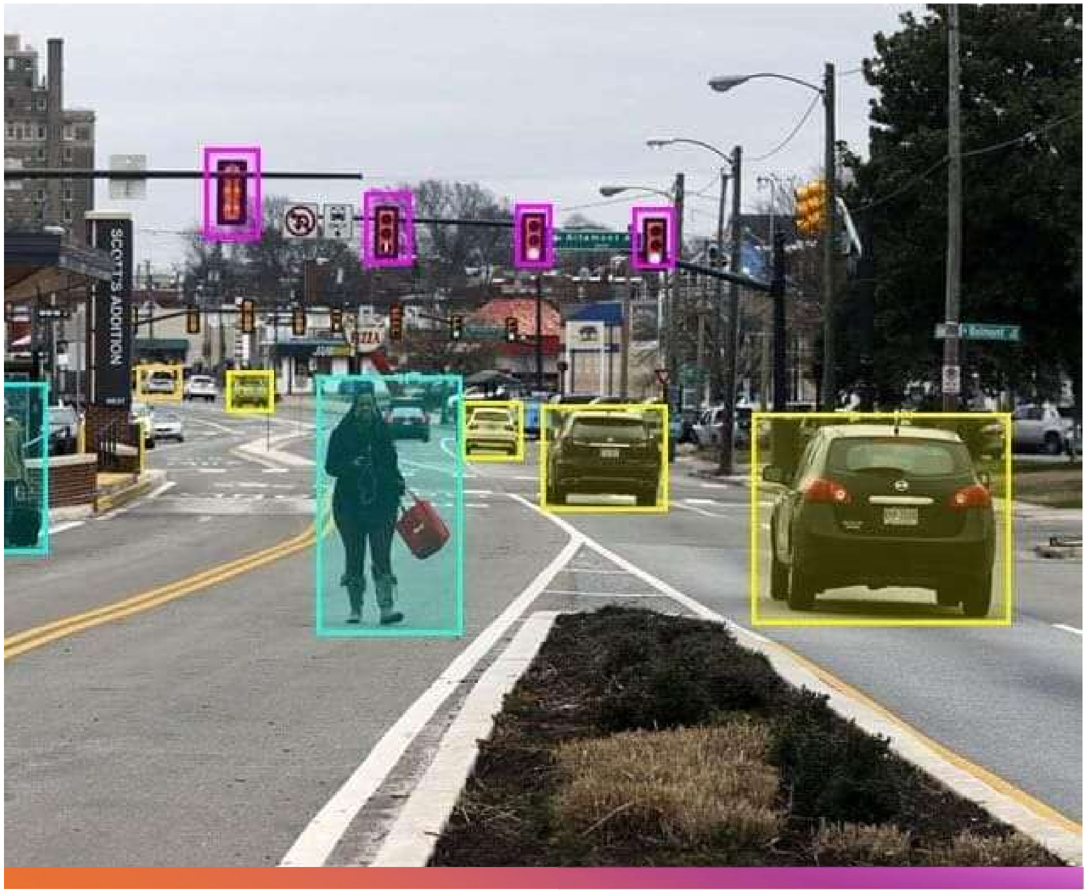
- $\text{Extent} = \dfrac{Area}{Bounding\ Box\ Area}$

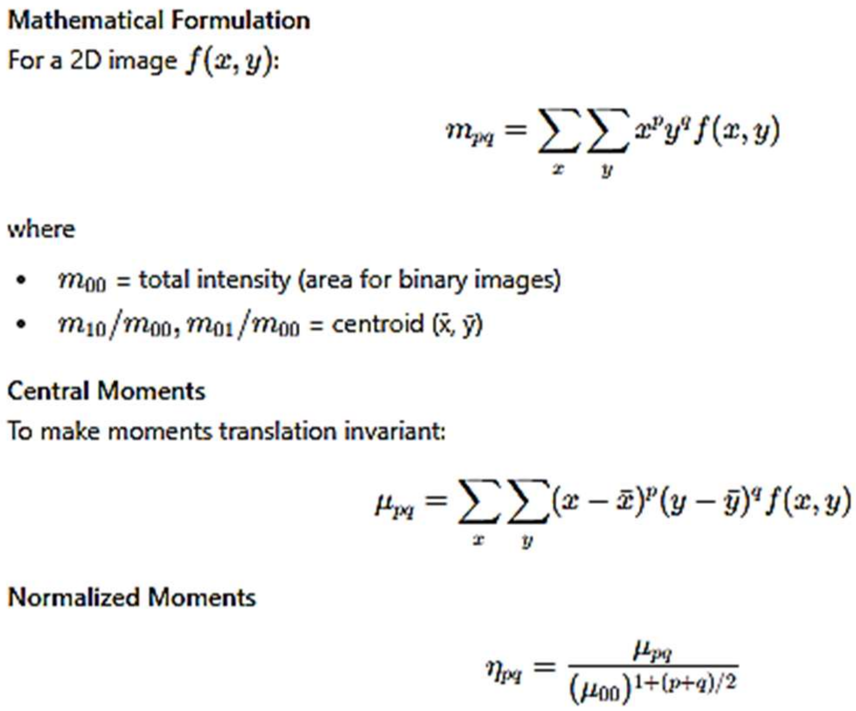
Meaning: It shows how much area the object occupies inside the bounding rectangle that encloses it.

Real-life example:

If a leaf fits perfectly inside a rectangle → Extent ≈ 1,

If the leaf is slanted and leaves empty space → Extent < 1.





Mathematical Formulation

For a 2D image $f(x, y)$:

$$m_{pq} = \sum_x \sum_y x^p y^q f(x, y)$$

where

- m_{00} = total intensity (area for binary images)
- $m_{10}/m_{00}, m_{01}/m_{00}$ = centroid (x̄, ȳ)

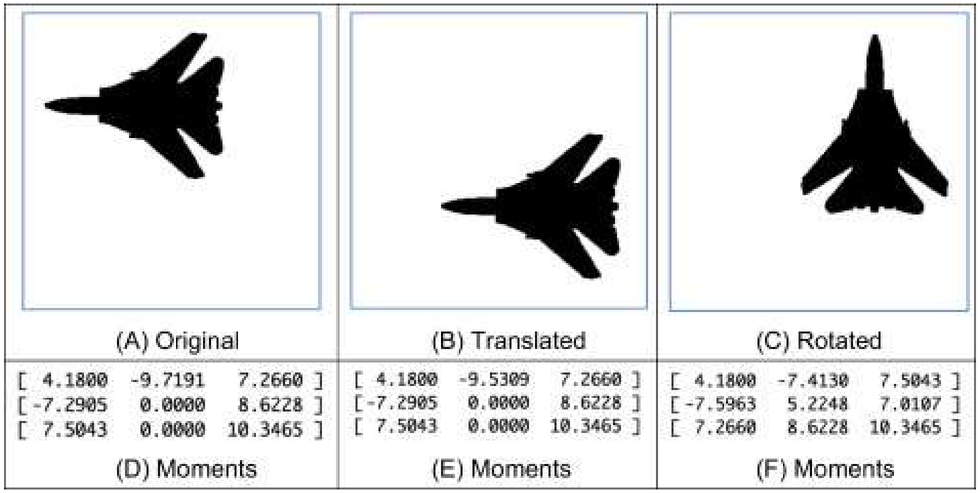
Central Moments

To make moments translation invariant:

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y)$$

Normalized Moments

$$\eta_{pq} = \frac{\mu_{pq}}{(\mu_{00})^{1+(p+q)/2}}$$



(A) Original

(B) Translated

(C) Rotated

```
[ 4.1800  -9.7191   7.2660 ]
[-7.2905   0.0000   8.6228 ]
[ 7.5043   0.0000  10.3465 ]
```

(D) Moments

```
[ 4.1800  -9.5309   7.2660 ]
[-7.2905   0.0000   8.6228 ]
[ 7.5043   0.0000  10.3465 ]
```

(E) Moments

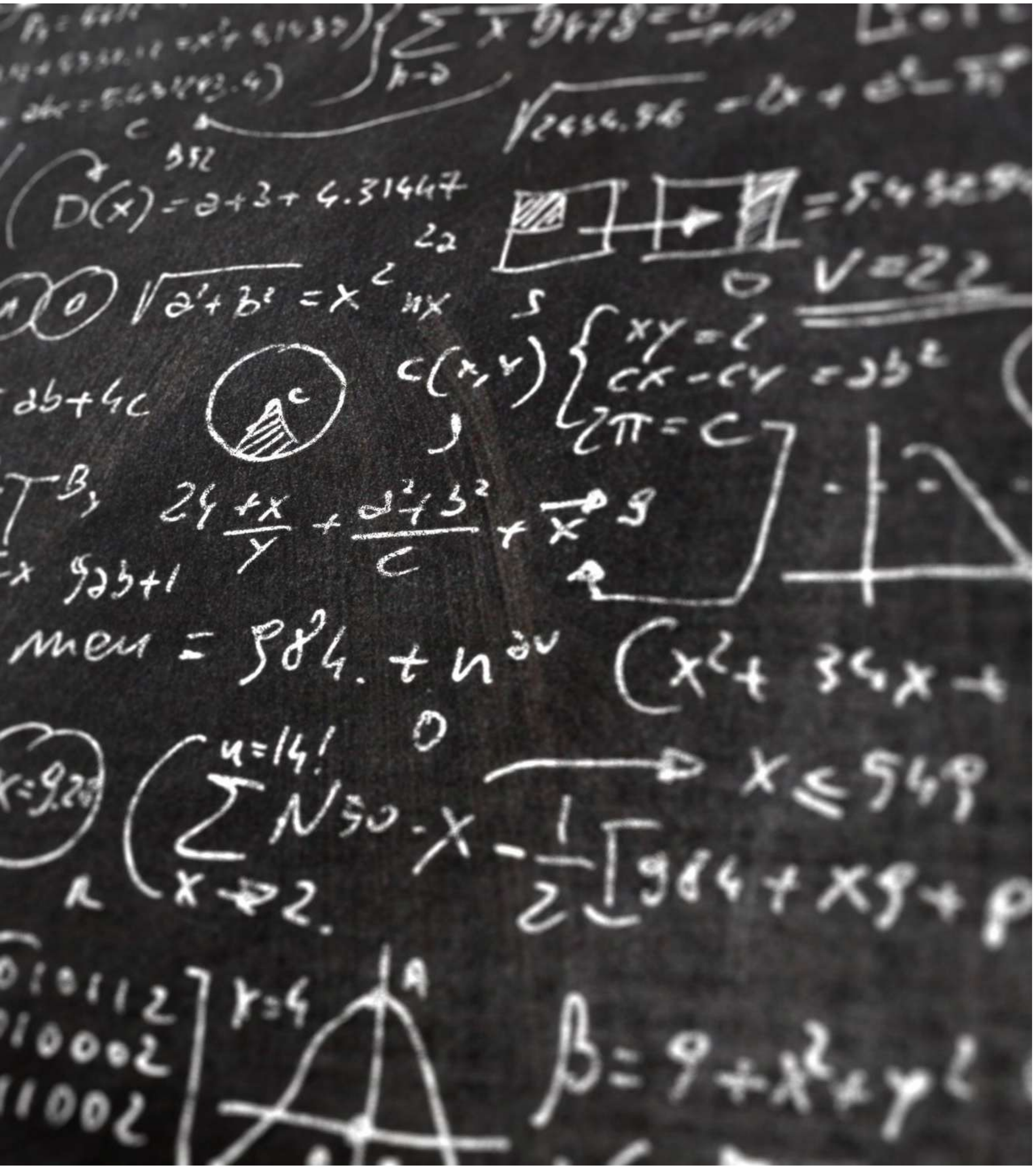
```
[ 4.1800  -7.4130   7.5043 ]
[-7.5963   5.2248   7.0107 ]
[ 7.2660   8.6228  10.3465 ]
```

(F) Moments

Moments

Definition

- Moments are statistical measures that describe the **shape** and **spatial distribution** of pixel intensities in an image.
- They provide compact representations of **area, centroid, orientation, and other geometric features**.



Applications

- Handwritten digit recognition
- Object identification
- Shape similarity measurement

- **Connected Component Analysis (CCA)**

Definition

- CCA is used to **label** and **count distinct objects** (connected regions) in a binary image.

Process

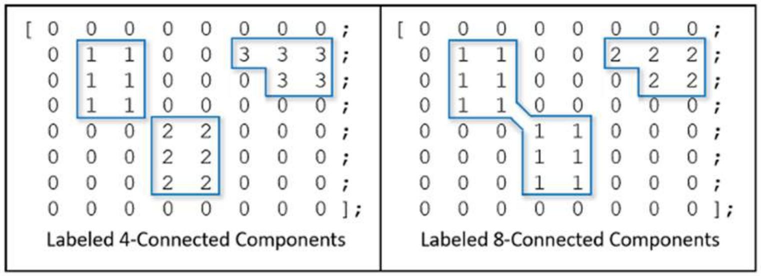
- Convert image to binary (foreground = 1, background = 0).
- Scan the image pixel by pixel.
- Assign a unique label to each group of connected 1’s.
- Connectivity types:
    - **4-connectivity:** Pixels share edge.
    - **8-connectivity:** Pixels share edge or corner.

Output

- Each connected region is given a unique label → used for:
- Object counting
- Blob detection
- Region-based feature extraction

Applications

- Counting cells in microscopy images
- Segmentation for OCR

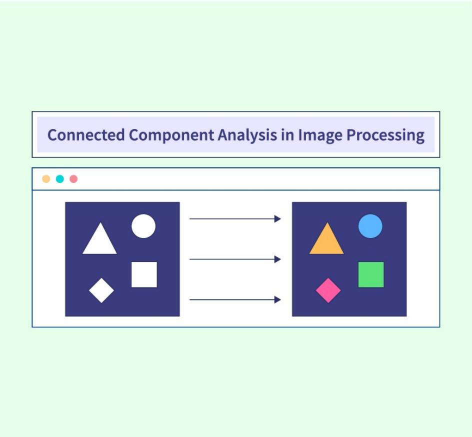


```
[ 0  0  0  0  0  0  0  0 ;
  0  1  1  0  0  3  3  3 ;
  0  1  1  0  0  0  3  3 ;
  0  1  1  0  0  0  0  0 ;
  0  0  0  2  2  0  0  0 ;
  0  0  0  2  2  0  0  0 ;
  0  0  0  2  2  0  0  0 ;
  0  0  0  0  0  0  0  0 ];
```

Labeled 4-Connected Components

```
[ 0  0  0  0  0  0  0  0 ;
  0  1  1  0  0  2  2  2 ;
  0  1  1  0  0  0  2  2 ;
  0  1  1  0  0  0  0  0 ;
  0  0  0  1  1  0  0  0 ;
  0  0  0  1  1  0  0  0 ;
  0  0  0  1  1  0  0  0 ;
  0  0  0  0  0  0  0  0 ];
```

Labeled 8-Connected Components



Convex Hull

Definition

- The convex hull of a shape is the **smallest convex polygon** that completely encloses the object.

Intuition

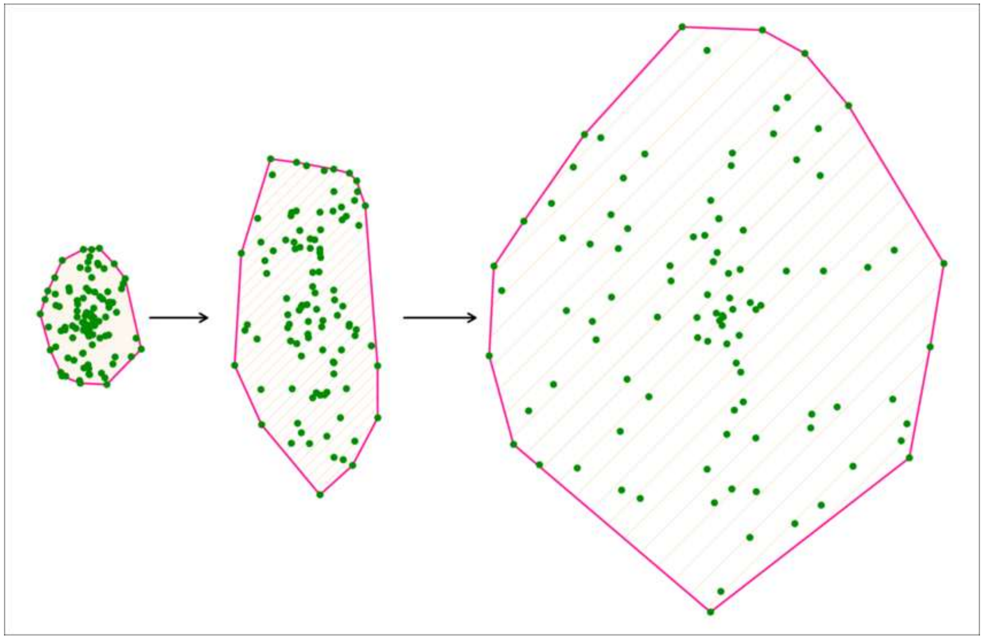
- If you imagine stretching a rubber band around the object — the band takes the convex hull shape.

Properties

- Always convex (no inward dent).
- Can be used to measure **convexity** and **shape irregularity**.

Applications

- Shape comparison
- Object boundary analysis
- Defect detection (difference between object area and convex hull area)



Distance Transform

Definition

- For a **binary image**, the distance transform replaces each foreground pixel with the **distance to the nearest background pixel**.

Mathematical Form

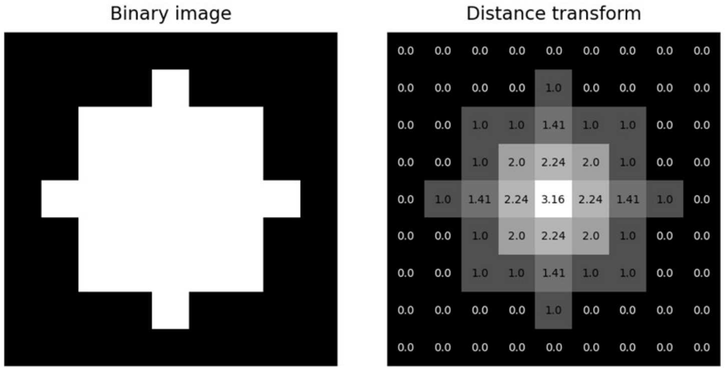
- $D(x, y) = \min_{(i, j) \in background} \sqrt{\left(x - i\right)^2 + \left(y - j\right)^2}$

Distance Metrics

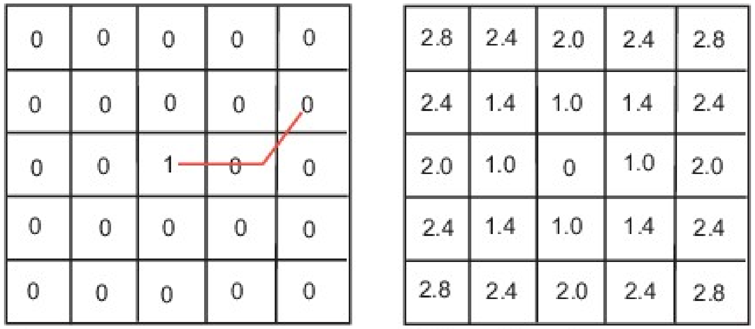
- **Euclidean Distance:** exact geometric distance.
- **City Block (Manhattan):** |x1 − x2| + |y1 − y2|
- **Chessboard:** max(|x1 − x2|, |y1 − y2|)

Applications

- Medial axis extraction
- Path planning
- Object thickness analysis



0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0
0.0	0.0	1.0	1.0	1.41	1.0	1.0	0.0	0.0
0.0	0.0	1.0	2.0	2.24	2.0	1.0	0.0	0.0
0.0	1.0	1.41	2.24	3.16	2.24	1.41	1.0	0.0
0.0	0.0	1.0	2.0	2.24	2.0	1.0	0.0	0.0
0.0	0.0	1.0	1.0	1.41	1.0	1.0	0.0	0.0
0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

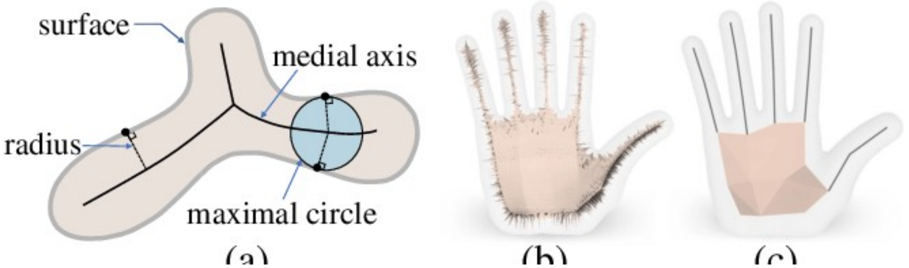


0	0	0	0	0
0	0	0	0	0
0	0	1	0	0
0	0	0	0	0
0	0	0	0	0

Image

2.8	2.4	2.0	2.4	2.8
2.4	1.4	1.0	1.4	2.4
2.0	1.0	0	1.0	2.0
2.4	1.4	1.0	1.4	2.4
2.8	2.4	2.0	2.4	2.8

Distance Transform



Medial Axis Transform (MAT)

Definition

- The **medial axis** is the set of all points having **more than one closest point** on the object boundary.
- It represents the **symmetry axis** or “skeleton” of a shape.

Computation

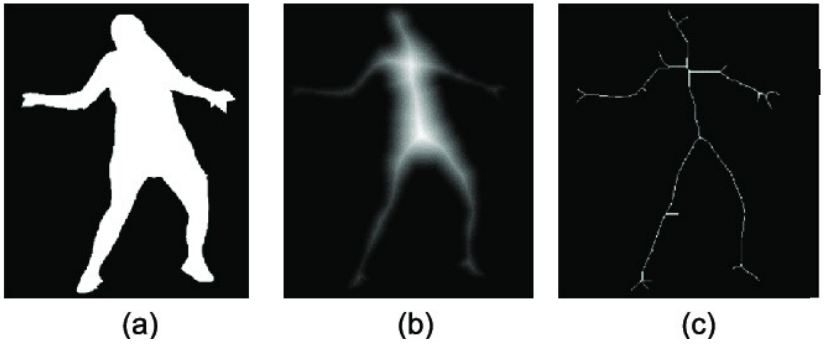
- Apply **distance transform** first.
- Points that are local maxima of the distance map form the **medial axis**.

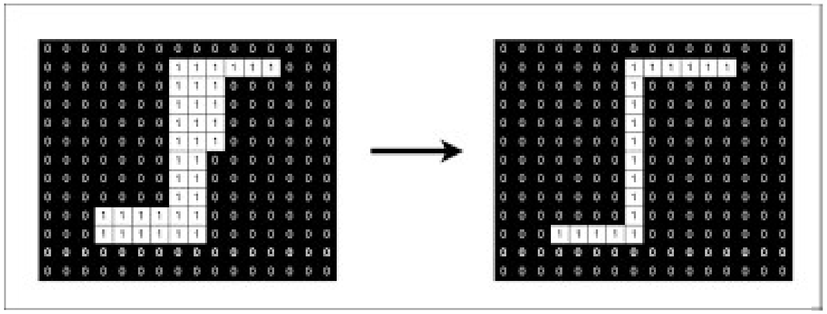
Properties

- Captures the **structural shape**.
- Sensitive to noise, so **pruning** is often applied.

Applications

- Shape matching and recognition
- Object representation in compact form





0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	1	1	1	1	1	1	0	0	0
0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0
0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0
0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0
0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0
0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0
0	0	0	1	1	1	1	1	1	0	0	0	0	0	0	0
0	0	0	1	1	1	1	1	1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

→

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	1	1	1	1	1	1	0	0	0
0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
0	0	0	1	1	1	1	1	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Skeletonization / Thinning

Definition

- Skeletonization reduces an object to a **one-pixel-wide representation** while preserving its **topology**.

Process

- Iteratively remove boundary pixels:
- Preserve connectivity.
- Do not break apart the object.

Difference

- **Thinning:** reduces object gradually from boundary inward.
- **Skeletonization:** emphasizes central lines using medial axis concepts.

Applications

- Character recognition (OCR)
- Fingerprint and vein pattern analysis
- Shape analysis and matching

Image Registration

- Registration aligns two or more images of the same scene so that corresponding points match.

Meaning of Image Registration

- **Image Registration** is a process of **aligning two or more images** of the same scene or object, so that corresponding points (features) in the images match perfectly.

👉 In simple words:

- “Image registration means matching one image with another so that both show the same objects in the same position, size, and orientation.”

Why It Is Needed

Image registration is used when:

- The **same object** is captured at **different times** (e.g., satellite before & after flood)
- Images are taken from **different sensors** (e.g., thermal + optical)
- The object is viewed from **different angles or positions**

This helps us understand **what has changed**, **how much it has changed**, or **how two images relate spatially**.

Steps in Image Registration

- **Feature Detection:**

Identify key points (like corners, edges, or textures) in both images.

- **Feature Matching:**

Find corresponding points between the two images (e.g., same corner or shape).

- **Transformation:**

Adjust one image — by rotation, scaling, or shifting — so that it aligns with the other.

- **Resampling and Output:**

Reconstruct the adjusted image so both are now on the same coordinate system (aligned properly).

Simple Example

Imagine you have two satellite images:

- One taken in the **morning**
- One taken in the **evening**

You want to **overlay** them to see how lighting or shadows have changed.

But they won’t match perfectly — maybe the satellite’s angle changed a little.

So, before comparing, you must **register (align)** them first.

That process is **Image Registration**.

Example — “Doctor and MRI Scans” 🏥

At a hospital, **Dr.** has a patient. Patient had a **brain MRI scan** last month, and another MRI this month.

- The doctor wants to check if the **tumor has grown or reduced** in size.

But there’s a problem:

- In the first MRI, patient’s head was slightly tilted.
- In the second, his head was perfectly straight.

If the doctor compares them directly, the brain structures won’t line up.

👉 So, the doctor uses an **Image Registration software**.

The software:

- Finds common features (like brain boundaries, ventricles) in both MRIs
- Slightly **rotates and shifts** one image
- Aligns it exactly with the other

Now the doctor can overlay both scans and clearly see how much the tumor has changed.

➡️ This whole process is called **Image Registration**.

Real-Life Applications

Domain	Application
🛰️ **Remote Sensing**	Aligning images from different satellites or times
🧬 **Medical Imaging**	Comparing MRI, CT, or PET scans over time
📷 **Computer Vision**	Image stitching for panoramas or augmented reality
🕵️ **Change Detection**	Detecting environmental or structural changes over time

Mono-modal Image Registration

Meaning:

- Mono-modal registration means **aligning images taken using the same imaging modality (type of sensor)** — i.e., the same kind of camera or scanner. In this case, the **image intensity characteristics are similar** in both images.

Example:

- Two **MRI brain scans** taken at different times (say, before and after treatment). Even though the brain may slightly shift or deform, both images come from the **same MRI machine**, so their contrast and brightness are similar.
- The goal is to align them to detect **changes or growth** in a tumor.

Example:

- Imagine a photographer takes **two pictures of a tree** using the **same camera** — one in the morning and one in the evening.

Because the light or camera angle may vary slightly, the pictures don’t overlap perfectly.

Mono-modal registration helps align these two photos so that **every leaf of the tree matches perfectly** in both images.

Mono-modal Registration

Aspect	Details
Definition	Alignment of images acquired **from the same imaging modality/sensor type**—for example, MRI-to-MRI, CT-to-CT, or two visible-light photographs.
Intensity Relationship	Pixel intensities are expected to have a **direct and consistent relationship** across images (linear or near-linear).
Typical Scenarios	• Tracking changes over time (longitudinal patient scans) • Multi-view photography with the same camera • Remote sensing using the same satellite sensor on different dates
Common Transform Models	Rigid (translation + rotation), similarity (add uniform scale), affine, or projective. Deformable models used when anatomy changes (e.g., breathing motion).
Similarity/Cost Measures	Since intensity patterns match, simple **intensity-based metrics** work well: **• Mean Squared Error (MSE)** **• Cross-Correlation / Normalized Cross-Correlation** **• Sum of Absolute Differences (SAD)**
Key Challenge	Mainly geometric distortions—camera motion, patient movement, or slight deformations—rather than differences in intensity mapping.

Multi-modal Image Registration

Meaning:

- Multi-modal registration means **aligning images taken from different imaging modalities** — i.e., using **different sensors or technologies**.

Here, **intensity values and contrast differ** because each modality captures different physical properties.

Example:

- Aligning **CT (Computed Tomography)** and **MRI** images of the same patient's brain.
  - CT shows **bone structure** clearly.
  - MRI shows **soft tissues** clearly.
    By aligning both, doctors can get a **comprehensive view** for accurate diagnosis or surgery planning.

Example:

Suppose a **security team** has two types of cameras at a gate —

- one **thermal camera** (shows heat), and
- one **normal color camera** (shows appearance).

If a person passes by, both cameras capture different information.

- Multi-modal registration helps combine both views so the person's **shape (color image)** and **body temperature (thermal image)** match perfectly.

Multi-modal Registration

Aspect	Details
Definition	Alignment of images captured **by different imaging modalities or sensors**—e.g., **CT–MRI**, **PET–CT**, **Infrared–Visible**, **Ultrasound–MRI**, **SAR–Optical satellite data**.
Intensity Relationship	Pixel intensity correspondence is **non-linear and often unknown**. A bright region in one modality might not be bright in the other (e.g., bones are bright in CT but dark in MRI).
Typical Scenarios	• Medical diagnostics combining structural (CT) and functional (PET) data. • Fusing radar and optical satellite images for environmental monitoring. • Security/defense: night-vision (IR) with daylight images.
Transform Models	Same geometric models as mono-modal (rigid, affine, deformable) but may need more flexible local warping when sensors distort differently.
Similarity/Cost Measures	Must be **intensity-invariant**. Popular choices: • **Mutual Information (MI)** and Normalized MI • Entropy-based measures • Feature-based metrics (SIFT/SURF keypoints with geometric matching)
Key Challenges	• Different noise characteristics and resolutions. • Non-linear intensity mapping. • Artifacts unique to each modality (e.g., CT beam-hardening vs. MRI distortions).

Key Difference Summary Table

Feature	Mono-modal Registration	Multi-modal Registration
Image Type	Same modality	Different modalities
Intensity Similarity	Similar	Different
Example	Two MRI scans	MRI and CT scan
Use Case	Motion correction, temporal change analysis	Data fusion, multi-sensor analysis
Challenge	Small intensity variation	Large contrast and texture variation

Real-World Examples

- **Medical Imaging**
    - *Mono*: Registering two MRI scans of a brain before and after treatment to detect tumor growth.
    - *Multi*: Aligning PET (functional metabolic data) with CT (anatomical structure) for accurate tumor localization.
- **Remote Sensing**
    - *Mono*: Aligning two Landsat optical images from different dates for vegetation change analysis.
    - *Multi*: Registering Synthetic Aperture Radar (SAR) with optical imagery to combine surface texture and visual information.
- **Industrial/Surveillance**
    - *Mono*: Matching consecutive frames from a production line camera.
    - *Multi*: Aligning infrared thermal images with standard RGB images for fault detection.

Summary

Understanding these distinctions guides the **choice of algorithms**, **similarity measures**, and **preprocessing steps** in image processing and computer vision tasks.

Key Insight

Mono-modal registration is about **geometry** (aligning similar images).

Multi-modal registration is about **both geometry and information fusion**, requiring similarity measures that handle **different intensity mappings**.

b. Global vs. Local

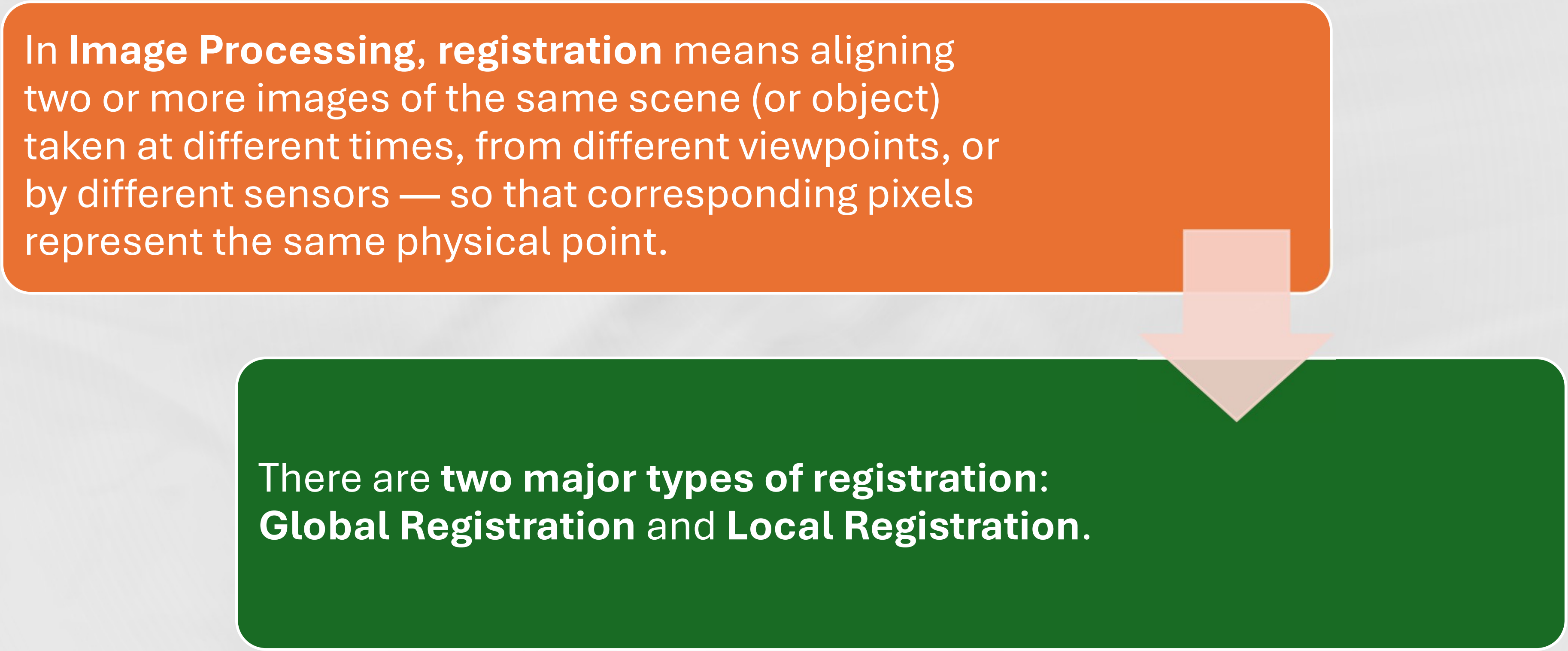
- **Global Registration:** One transformation (rigid, affine, projective) aligns the entire image.
- **Local (Non-rigid / Elastic):** Allows spatially varying transformations to handle deformations (e.g., organ motion).

c. Transforms

- **Rigid:** Translation + Rotation.
- **Similarity:** Rigid + Uniform scaling.
- **Affine:** Adds shear and non-uniform scaling.
- **Projective / Homography:** Handles perspective changes.
- **Deformable / B-spline:** Complex warps for local alignment.

d. Similarity Measures

- **Mono-modal:** Mean Squared Error (MSE), Cross-Correlation.
- **Multi-modal:** Mutual Information (MI) is standard.
- **Feature-based:** Matching keypoints (SIFT, SURF) with geometric consistency.



In **Image Processing**, **registration** means aligning two or more images of the same scene (or object) taken at different times, from different viewpoints, or by different sensors — so that corresponding pixels represent the same physical point.

There are **two major types of registration**: **Global Registration** and **Local Registration**.

1. Global Registration

Definition:

Global registration applies a **single geometric transformation** to the **entire image**.
It assumes that the relationship between the reference and target images is **uniform across the whole image**.

Examples of transformations used:

- Translation (shift)
- Rotation
- Scaling
- Affine transformation (combination of translation, rotation, scaling, and shear)

When it is used:

- When images differ only by **simple global changes**, e.g., camera movement, zoom, or rotation.
- When the entire scene moves as a **rigid body** (no local deformation).

Example (Real-life):

If two satellite images of a city are taken from different angles or on different days, global registration can align them using rotation and translation so that roads and buildings overlap properly.

- **Mathematical form:**
- $x' = a_1 x + a_2 y + a_3$
- $y' = b_1 x + b_2 y + b_3$
- Where $(x, y)$are original coordinates and $(x', y')$are transformed coordinates.

2. Local Registration

Definition:
Local registration handles **spatially varying deformations** by applying **different transformations** to different parts (regions or pixels) of the image.

When it is used:

- When there are **non-rigid or local deformations** such as tissue movement in medical images, or facial expression changes in human faces.
- When the object shape changes locally but not globally.

Techniques used:

- Thin Plate Splines
- B-spline deformation
- Optical Flow methods
- Non-rigid (elastic) registration

Example (Real-life):
In medical imaging, if MRI scans of a patient's brain are taken at different times, local registration can align them accurately even if soft tissues have slightly shifted or deformed.

Mathematical form:

$$(x', y') = T(x, y) = (x + u(x, y),\ y + v(x, y))$$

where $u(x, y)$ and $v(x, y)$ are local displacement fields.

Summary Table

Feature	**Global Registration**	**Local Registration**
Transformation	Single for entire image	Varies spatially
Handles	Rigid or affine transformations	Non-rigid, elastic deformations
Complexity	Low	High
Computation	Fast	Computationally intensive
Examples	Satellite, aerial, panoramic images	Medical, facial, biological images