VIT-AP UNIVERSITY

Computer Vision

(Course Code: 4047)

Module-1:Lecture-3: Image Representation and Analysis

Gundimeda Venugopal, Professor of Practice, SCOPE

Key Stages in Digital Image Processing

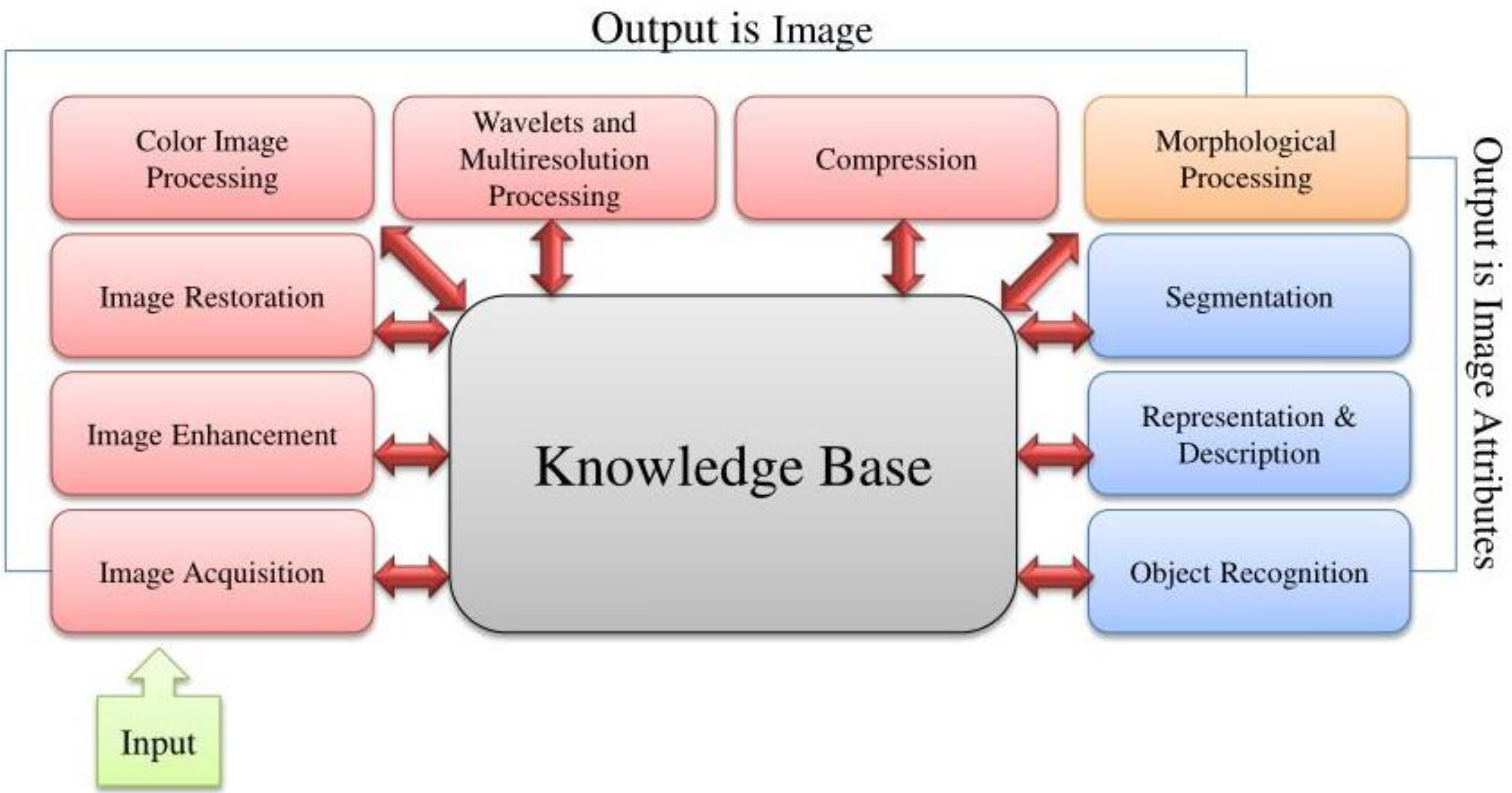


Image Representation and Description

- After an image is segmented into regions, the regions are represented and described in a form suitable for computer processing (descriptors).
- Representing a region:
  1. In terms of its external characteristics (boundary)
  2. In term of its internal characteristics

  Exp: A region might be represented by the length of its boundary.

  – External representations are used when the focus is on shape of the region.

  – Internal representations are used when the focus is on color and texture.

  – Representations should be insensitive to rotation and translation.

Image Description deals extracting attributes that result in some quantitative information of Interest.

Boundary Following Algorithm

Many algorithms require the points on the boundary of region be ordered Clockwise (or anticlockwise)
We assume

a) Object and Background are represented by 1 and 0 respectively
b) Image are padded with 0s to eliminate the possibility of object merging with image border

Let the starting point b_0 be the uppermost, leftmost point in the image. c_0 the west neighbor of b_0. Examine 8 neighbors of b_0 starting at c_0 & proceed in clockwise direction.

Let b1 denote first neighbor encountered with value 1 & c1 be background point immediately preceding b1 in the sequence.

		1	1	1	1	
	1			1		
		1		1		
	1			1		
	1	1	1	1		

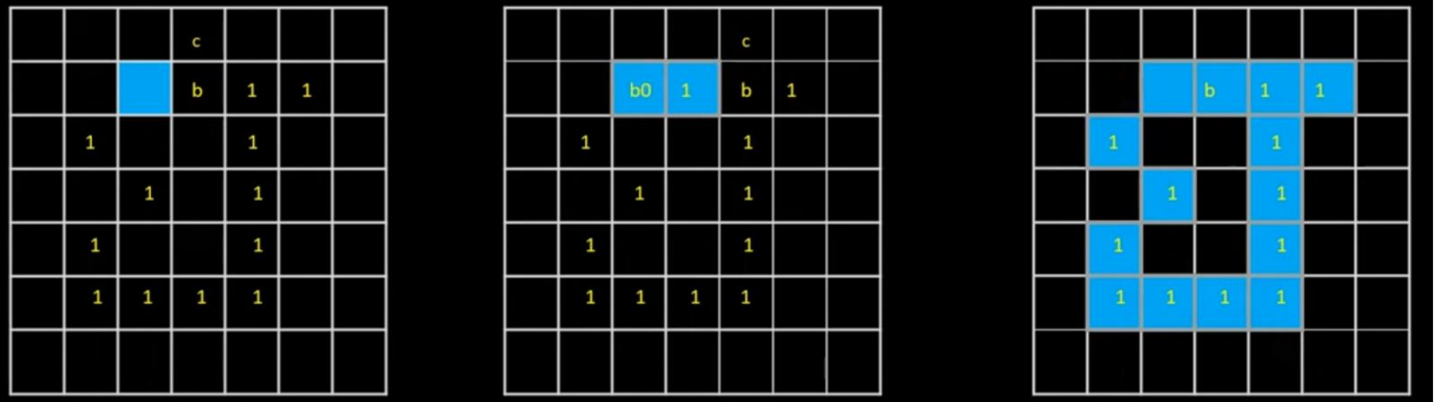
	c0	b0	1	1	1	
	1			1		
		1		1		
	1			1		
	1	1	1	1		

Boundary Following Algorithm

2) Let b = b1 & c = c1

3) Let the 8-neighbors of b, starting at c and proceeding in clockwise directions be denoted by n1, n2, ...n8. Find first n_k labeled 1.

4)Let b = nk & c = nk-1



- 5) Repeat step 3 & 4 until b = b0 & next boundary point found is b1. The sequence of b points found when the algorithm stops constitutes the set of ordered boundary points.

- The algorithm is also called as MOORE BOUNDARY TRACKING ALGORITHM.

Polygonal Approximation: Minimum Parameter Polygon

- The goal of polygon approximation is to represent an object boundary by a polygon.
- The minimum perimeter polygon consists of line segments that minimize distances between boundary pixels.
- Here the boundary is enclosed by a set of concatenated cells. The enclosure has two walls corresponding to the inside and outside boundaries of the strip of cell.
- Think of the object boundary as a rubber band contained within the wall.
- The rubber band shrinks and produces a polygon of minimum perimeter that fit the geometry established by the cell strip.

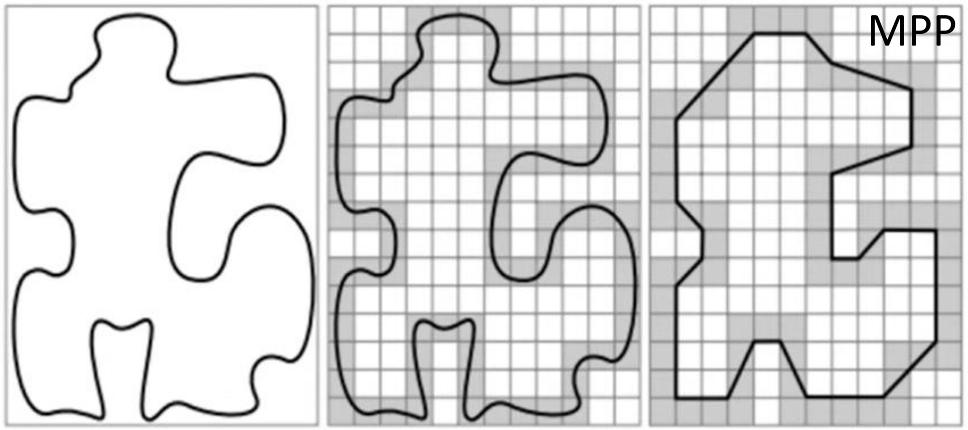


FIGURE 11.6 (a) An object boundary (black curve). (b) Boundary enclosed by cells (in gray). (c) Minimum-perimeter polygon obtained by allowing the boundary to shrink. The vertices of the polygon are created by the corners of the inner and outer walls of the gray region.

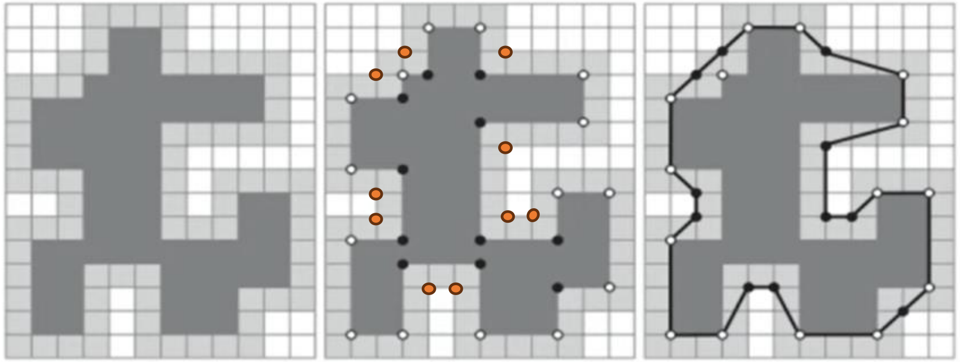


FIGURE 11.7 (a) Region (dark gray) resulting from enclosing the original boundary by cells (see Fig. 11.6). (b) Convex (white dots) and concave (black dots) vertices obtained by following the boundary of the dark gray region in the counterclockwise direction. (c) Concave vertices (black dots) displaced to their diagonal mirror locations in the outer wall of the bounding region; the convex vertices are not changed. The MPP (black boundary) is superimposed for reference.

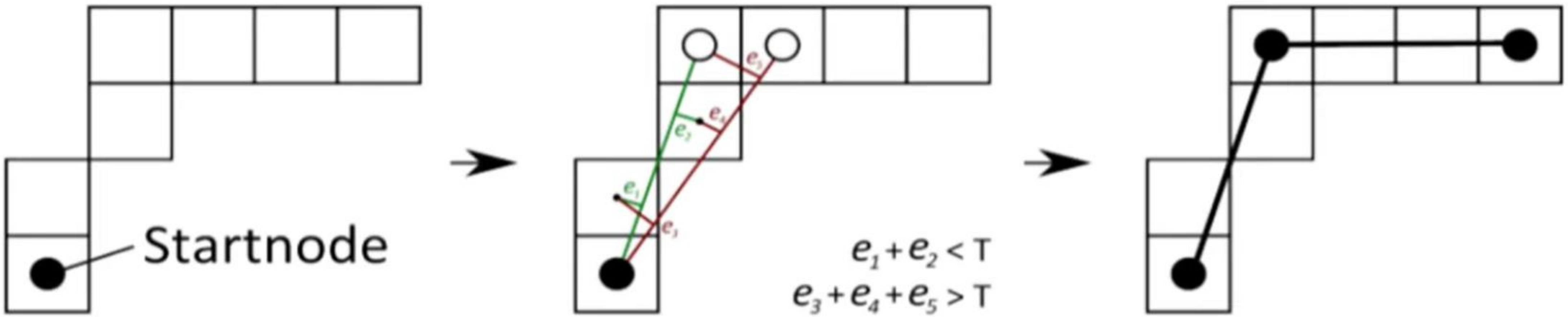
b) Diagonal Mirror locations in the outer wall of bounding region denoted with orange dots

MPP Approximation Advantages

- In essence, polygon approximation with minimum-perimeter polygons helps strike a balance between accurately representing a shape and simplifying it for practical purposes.
- It's like creating a simpler version of a complex shape by connecting a few key points with the shortest line possible, while still keeping the main characteristics of the shape intact.
- This technique is commonly used in various fields of image processing, computer graphics, and geographic information systems.

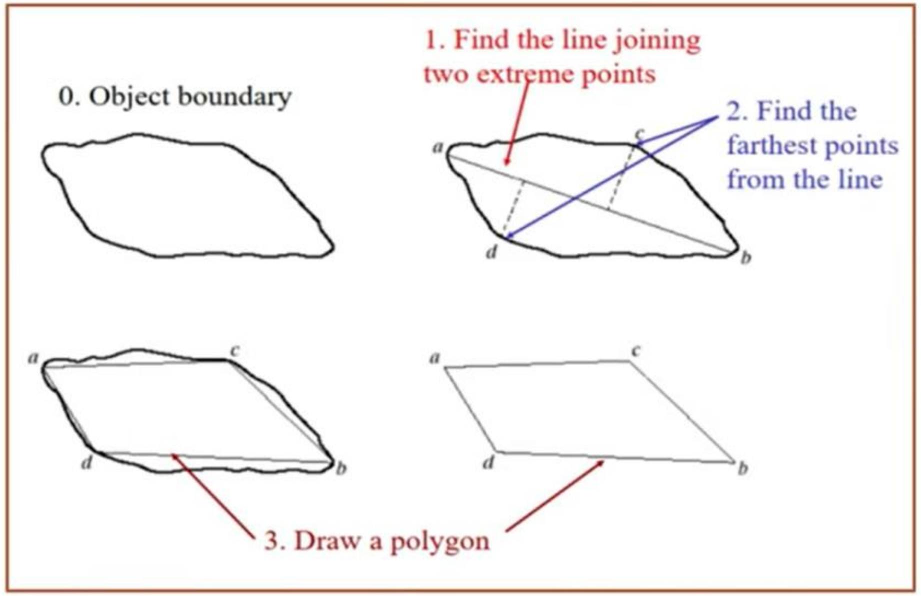
Merging: Polygonal Approximation

- Merge points around the boundary and fit a least-square-error line to the points until an preset threshold is exceeded.
- Next, start a new line and repeat above step.
- When the start point is reached, the intersections of adjacent lines are the vertices of the polygon.
- Here the disadvantage is that the vertices do not always correspond to inflections(corners).



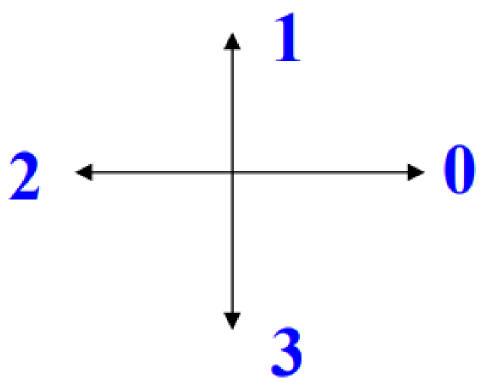
Splitting: Polygonal Approximation

- Find the line joining two extreme points. Choose a threshold(e.g. Half of this line).
- Find the farthest perpendicular points on the boundary from the above mentioned line. If this distance exceed threshold, the point becomes a vertex and again subdivide the segment into two sub-segments, if not no change in original line.
- Repeat above step until the initial point is reached.
- This technique has the advantage in finding prominent inflection points.



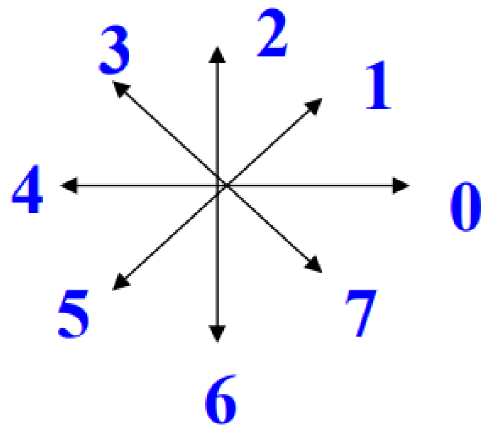
Boundary Representation: Chain Code

- Chain Code: Used to represent a boundary by a connected sequence of straight line segments.
  - 4 or 8 connectivity is used
  - The direction of each segment is coded by a numbering scheme.



Method:

- Follow the boundary in a specific (clockwise) direction.
- Assign a direction to the segment connecting every pair of pixels.

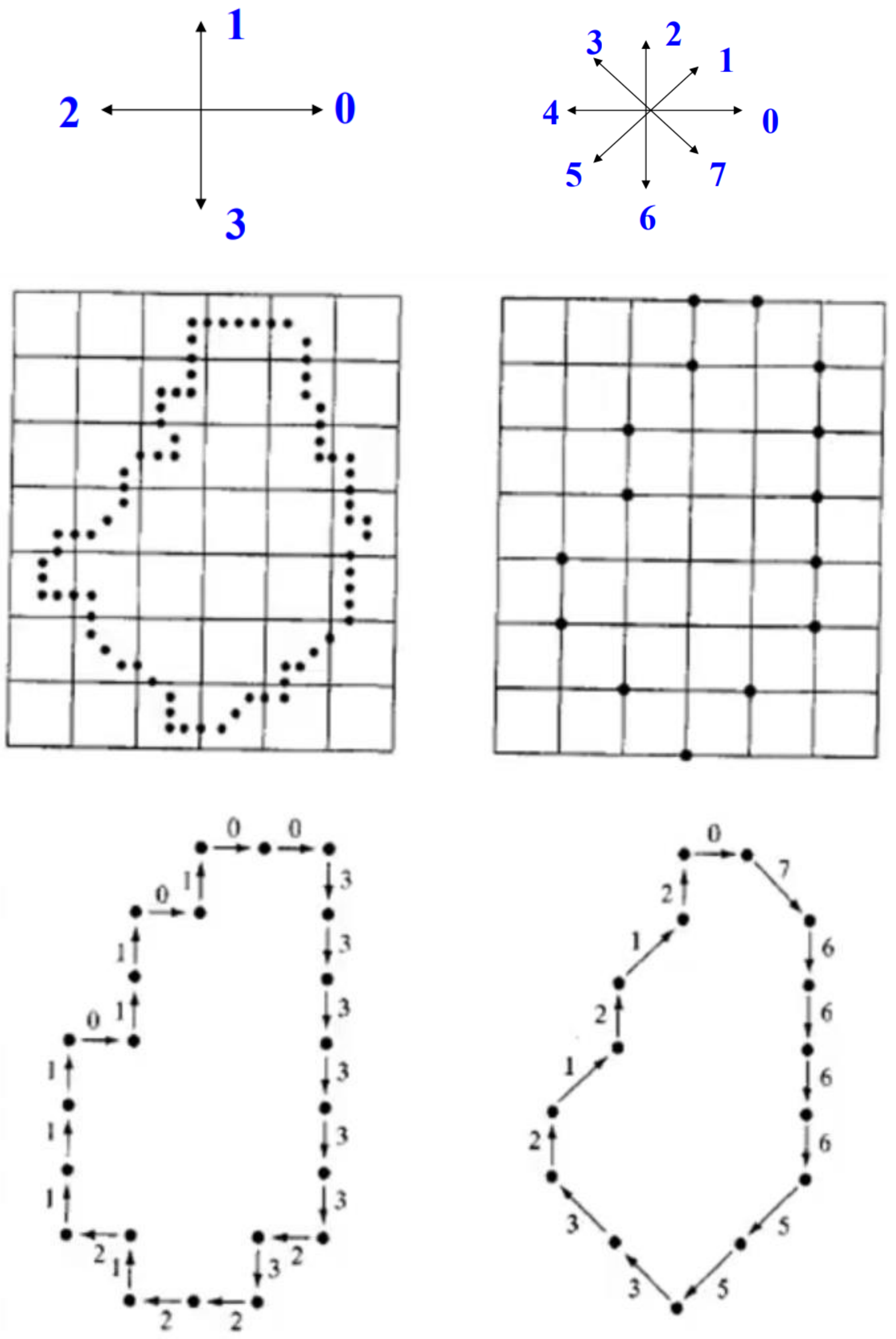


Chain Code

- Solution: Resample the boundary by selecting a larger grid spacing

- Then, as the boundary is traversed, a boundary point is assigned to each node of the large grid, depending upon the proximity of original boundary of the node

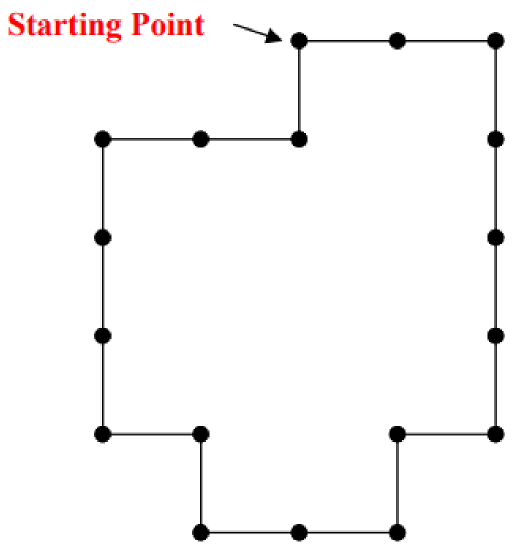
- The resampled boundary can now be represented by 4 or 8 code

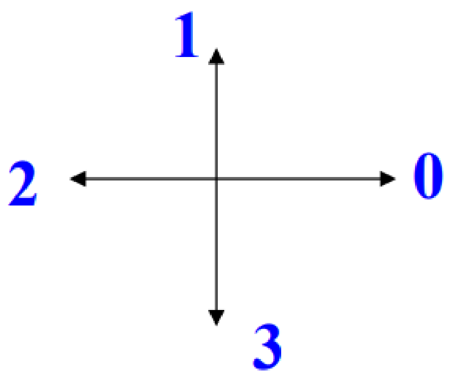


Chain Code

Exp: 003333232221211101

- Problems:
  - The chain code depends on the starting point.
  - It changes with rotation.



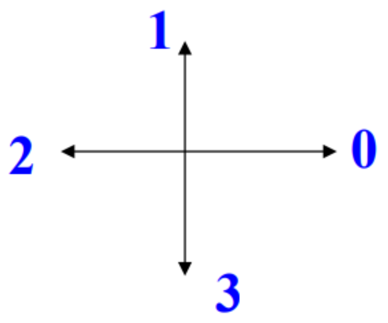


Chain Code

Solutions:

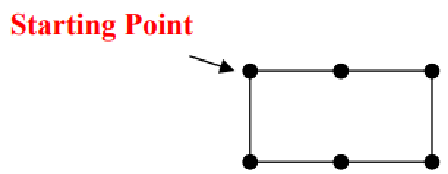
- Treat the chain code as a circular sequence of numbers. Circulate until the number is of minimum magnitude.
- Use the difference of chain code instead of the code itself: count counterclockwise the number of directions that separates two adjacent elements (First difference)

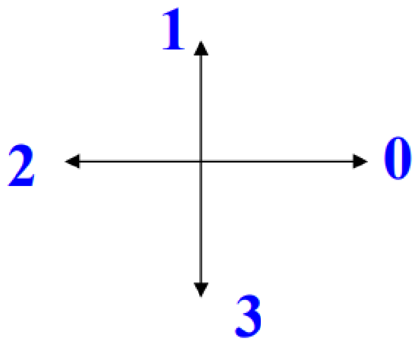
- Exp: 10103322
- First difference code: 33133030



Shape Number

- Shape number: first difference of smallest magnitude (in the chain code)
  - Exp: chain code 003221
  - First difference code: 303303
  - Shape number: 033033





Fourier Descriptor

- Fourier descriptors are a way of encoding the shape of a two-dimensional object by taking the Fourier transform of the boundary, where every point on the boundary is mapped to a complex number.
- Fourier descriptors refer to a set of numbers derived from Fourier coefficients that characterize the frequency content of a shape, allowing for a compact representation of the shape by capturing its main features while ignoring small variations caused by noise.
- The original shape can be recovered from the inverse Fourier transform. However, if only a few terms of the inverse are used, the boundary becomes simplified, providing a way to smooth or filter the boundary.
- Few Fourier descriptors can be used to capture the gross essence of a boundary. They can be used as a basis for differentiating between distinct boundary shapes.
- Local frequency components provide most of the features a shape (boundary).
- High Frequency components account for finer detail (corners). They are easily affected by noise and only represent detail that is of little value to recognition.

N point boundary

$(x_0, y_0), (x_1, y_1), \ldots, (x_{N-1}, y_{N-1})$

$s(k) = x_k + jy_k$

N point DFT of s(k):

$$a(u) = \frac{1}{N} \sum_{k=0}^{N-1} s(k) \exp(-j2\pi uk / N)$$

$a(u)$ are called Fourier descriptors.

To define a curve that passes through the k sampled points, we need to consider only k/2 coefficients.

In practice, Fourier descriptors are computed for fewer coefficients than the limit of k/2.

https://demonstrations.wolfram.com/FourierDescriptors/

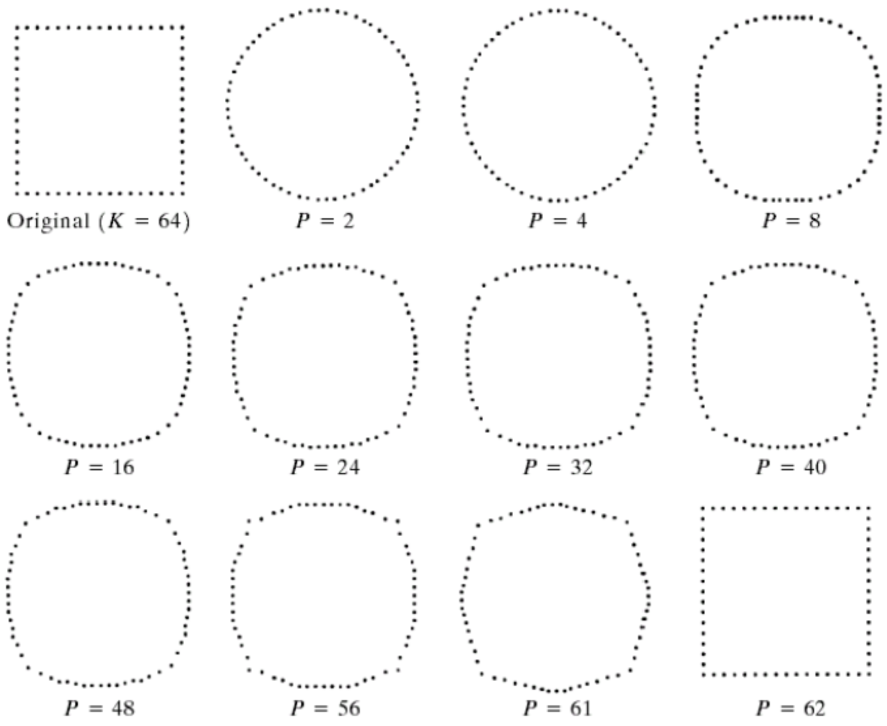
Fourier Descriptor

If P of the Fourier descrptors are used

$$\hat{s}(k) = \sum_{u=0}^{P-1} a(u)\exp(j2\pi uk/N)$$

$P < N \Rightarrow$ High frequency details of the boundary (e.g., corners) are removed.

- Fourier descriptors are not directly insensitive to translation, rotation and scaling.
- Magnitude of the Fourier descriptors is insensitive to rotation.



At P=8, Reconstructed starts looking like a square than a circle
At P=56, Corners start to break out of the sequence
At P=61, the curves begin to straighten

Fourier Descriptors: Effect of Rotation, Translation & Scaling

Transformation	Boundary	Fourier Descriptor
Identity	$s(k)$	$a(u)$
Rotation	$s_r(k) = s(k)e^{j\theta}$	$a_r(u) = a(u)e^{j\theta}$
Translation	$s_t(k) = s(k) + \Delta_{xy}$	$a_t(u) = a(u) + \Delta_{xy}\delta(u)$
Scaling	$s_s(k) = \alpha s(k)$	$a_s(u) = \alpha a(u)$
Starting point	$s_p(k) = s(k - k_0)$	$a_p(u) = a(u)e^{-j2\pi k_0 u/K}$

Reference: https://cis.temple.edu/~lakaemper/courses/cis595_2004/papers/fourierShape.pdf

Regional Descriptors

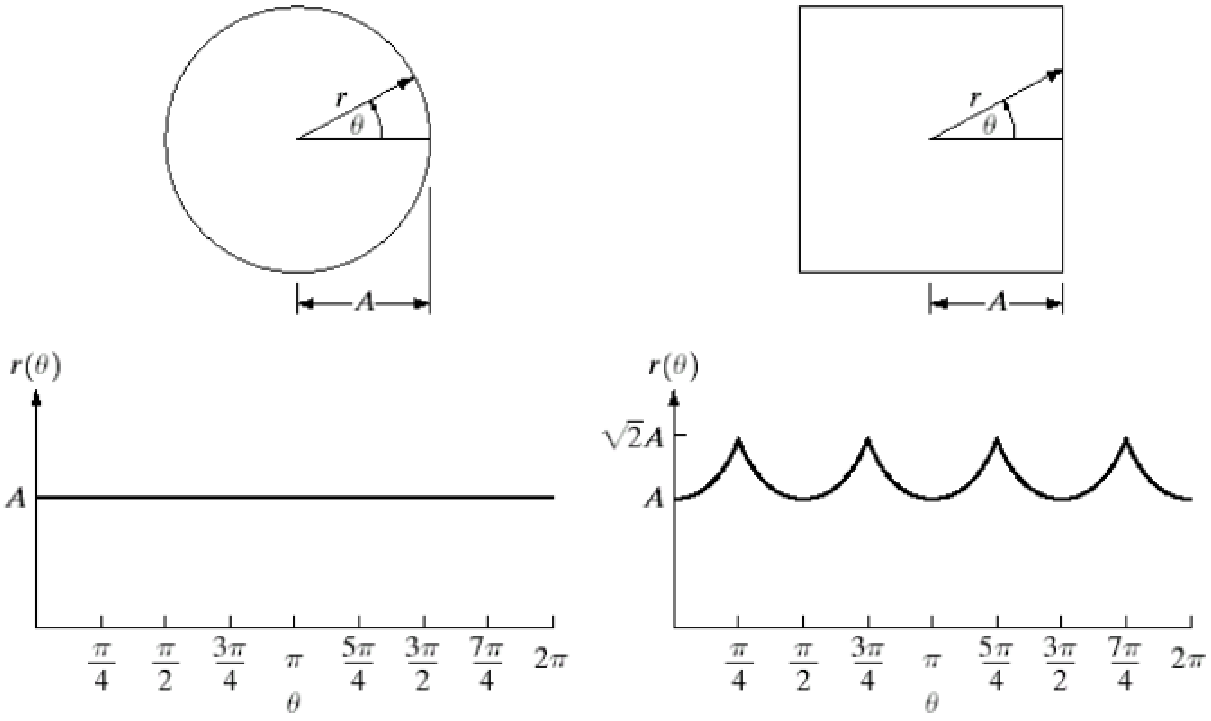
- Area: number of pixels contained within a region
- Compactness: $(\text{perimeter})^2/\text{area}$
- Min and Max of the gray levels in the region
- Mean and median of gray levels

Boundary Representation: Signature

- Signature: a 1-D functional representation of a boundary
- Different ways of generating signature
- Plot distance form centroid to boundary as a function of angle

a b

FIGURE 11.5 Distance-versus-angle signatures. In (a) $r(\theta)$ is constant. In (b), the signature consists of repetitions of the pattern $r(\theta) = A \sec\theta$ for $0 \leq \theta \leq \pi/4$ and $r(\theta) = A \csc\theta$ for $\pi/4 < \theta \leq \pi/2$.



Signature

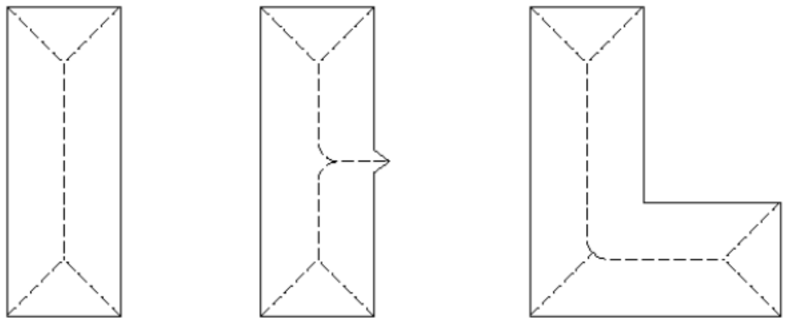
- Invariant to translation, depend on rotation and scaling
- To make it invariant to rotation we should select the same starting point regardless of the orientation
  - Select starting point farthest from centroid (if unique)
- To make it invariant to scaling we can normalize to a particular range
- Other signatures: traverse the boundary, at each point plot the angle between a line tangent to the boundary and a reference line
- Slope-density-function: histogram of tangent-angle values
- Straight segments will form the peaks of histogram

Skeletons

- An important approach to representing structural shape of a plane region is to reduce it to a graph
- This may be accomplished by obtaining the skeleton of the region via a thinning algorithm.
- Applications in automated inspection ….
- Definition of skeleton is based on medial axis transformation (MAT)

Skeletons

- MAT of a region R with border B: for each point p in R, find the closest neighbor in B. If p has more than one such neighbor, it belongs to medial axis (skeleton)
- MAT is based on “prairie fire concept”.
- Direct implementation of MAT is computationally expensive
- Alternative algorithms have been proposed that “thin” the boundary of a region until the skeleton is left



a b c

FIGURE 11.7 Medial axes (dashed) of three simple regions.

Skeletons

- An algorithm for thinning binary regions (assume region points are 1 and background points are 0)
- The algorithm has two steps which are applied to all the pixels on the contour of the region
- A contour point is any pixel with value 1 and having at least one 8-neighbor valued 0.
- In each step the boundary point that satisfy a set of conditions are flagged and then deleted

Skeletons

- Step 1 flags a contour point p1 if the following conditions are satisfied:

  a) $2 \le N(p1) \le 6$ (Not a Isolated or End point)

  b) $T(p1) = 1$ (Not a corner or junction point)

  c) $p2.p4.p6 = 0$

  d) $p4.p6.p8 = 0$

  (This is Not part of a horizontal /vertical line

p_9	p_2	p_3
p_8	p_1	p_4
p_7	p_6	p_5

- N(p1): number of nonzero neighbors of p1
- T(p1) : number of 0 to 1 transitions in ordered sequence p2,p3,….p8,p9,p2.

0	0	1
1	p_1	0
1	0	1

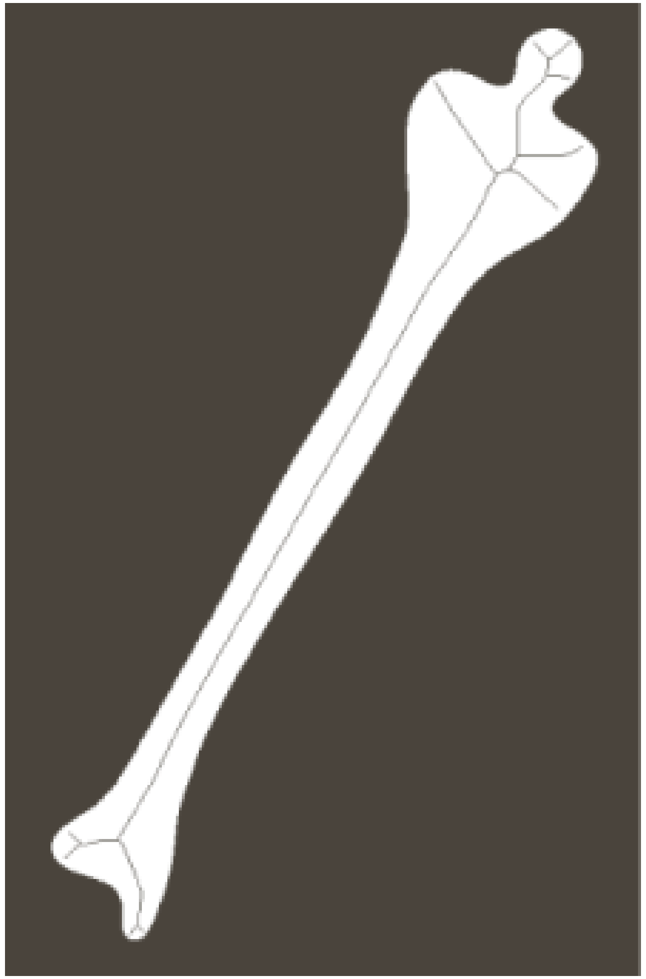
Skeletons

- After step 1 is applied to all border points those that are flagged are deleted (changed to 0)
- In step 2 conditions (a) and (b) remain the same but (c) and (d) are changed to:

$$c') \qquad p2.p4.p8 = 0$$
$$d') \qquad p2.p6.p8 = 0$$

- After step 2 is applied to all border points remaining after step 1, those that are flagged are deleted (changed to 0)
- This procedure is applied iteratively until no further points are deleted.

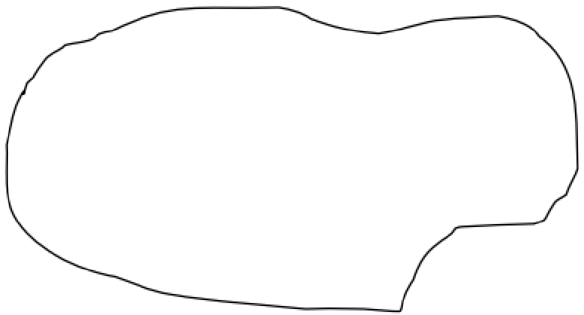
Skeletons



Human leg bone and skeleton of the region shown superimposed.

Simple Boundary Descriptors

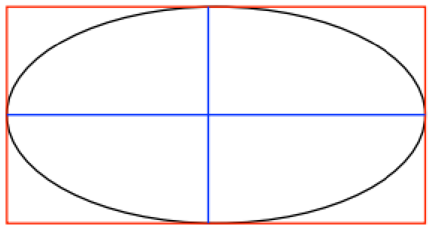
- Length: number of pixels along the contour of a region
- Diameter: $Dima(B)=max[D(p_i,p_j)]$
  - p_i,p_j are points on the boundary.



- Curvature: rate of change of slope.
- For digital images: the difference between the slopes of adjacent boundary segments.

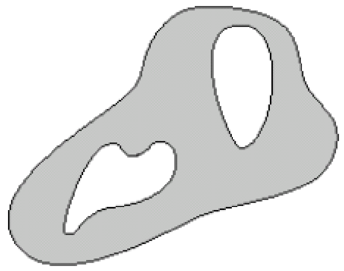
Simple Boundary Descriptors

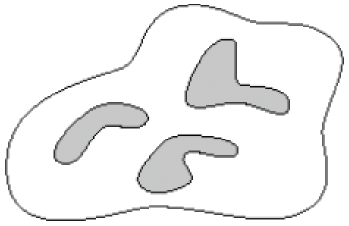
- Major axis: straight line segment joining the two points farthest from each other on the boundary
- Minor axis: Perpendicular to the major axis and of such length that a box could be formed to enclose the boundary.
- Eccentricity=Major axis/Minor axis
- Basic rectangle (bounding box): the rectangle formed by minor and major axes enclosing the boundary.



Topological Descriptors

- Topology: study of properties of a figure that are unaffected by any deformation, as long as there is no tearing or joining of the figure (rubber sheet distortions)
- Since stretching affects distance, topological properties do not depend on the notion of distance
- Number of holes in a region (H)
- Number of connected components (C)





Topological Descriptors

- Euler number: E=C-H
- Sometimes a region is represented by straight-line segments (polygonal network)
- V: number of vertices, Q: number of edges, F: number of faces => V-Q+F=C-H=E
- For the figure below right: 7-11+2=1-3=-2

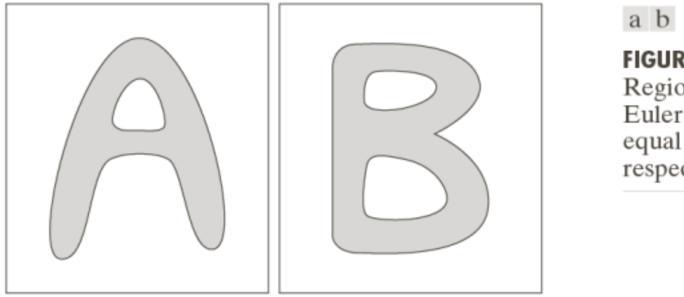


FIGURE 11.25 Regions with Euler number equal to 0 and respectively.

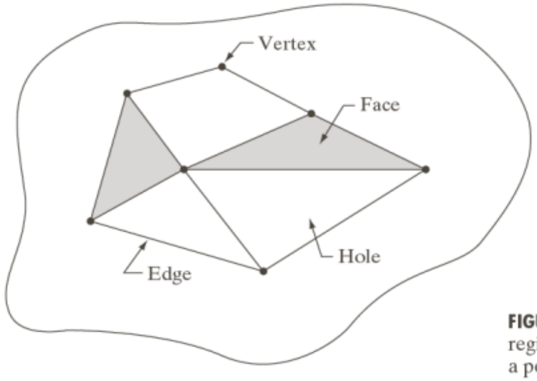


FIGURE 11.26 A region containing a polygonal network.

Color Histogram

Examples of histogram-based descriptors

- Global Color Histogram (GCH): computes a single histogram for the image, using the frequency values as descriptor.
- Border/Interior Classification (BIC): computes two histograms – one for pixels considered as interior (a given pixel is interior if their neighbors have the same color), and another for pixels considered as border (otherwise).

Those methods generate a vector of values that tries to "describe" color in a given image or image region.

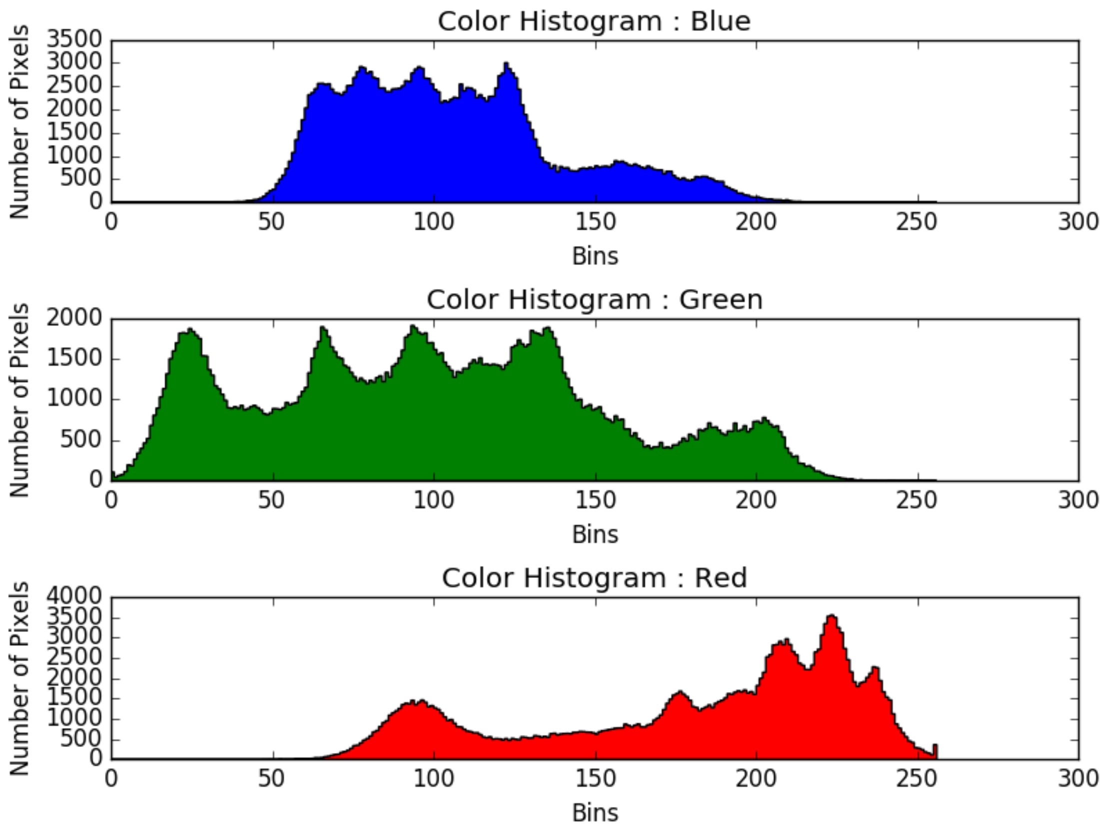
Descriptors are often designed with a distance function that allows the comparison between them.

(a) Input Image

(b) Interior Pixels

(c) Boundary Pixels

Color Histogram



A color histogram is **a representation of the distribution of colours in an image**.

For digital images, a color histogram represents the number of pixels that have colours in each of a fixed list of color ranges, that span the image's color space, the set of all possible colours.

The color histogram can be built for any kind of color space, although the term is more often used for three-dimensional spaces such as RGB or HSV. For monochromatic images, the term **intensity histogram** may be used instead.

For multi-spectral images, where each pixel is represented by an arbitrary number of measurements (for example, beyond the three measurements in RGB), the color histogram is *N*-dimensional, with N being the number of measurements taken. Each measurement has its own wavelength range of the light spectrum, some of which may be outside the visible spectrum.

Global Color Histogram and Local Color Histogram

GCH is the most known color histogram used to detect similar images.

Feature extraction algorithm:

1.Discretize your color-space (images' colors) into n color (You may use just 8*8*8 =512 color instead of 256*256*256=16777216 color).

2.Create a bin for each color.

3.Count number of pixels for each color and store it in histogram's bins.

Matching function:

The most common matching function for this method is Euclidean distance.

- To compare 2 images A, B.

A(R,G,B) : represents number of pixels in color = (R,G,B). (for example A(6,2,4) represents the number of discretized pixels of color R=6,G=2 and B=4).

D: sum Euclidean distances.

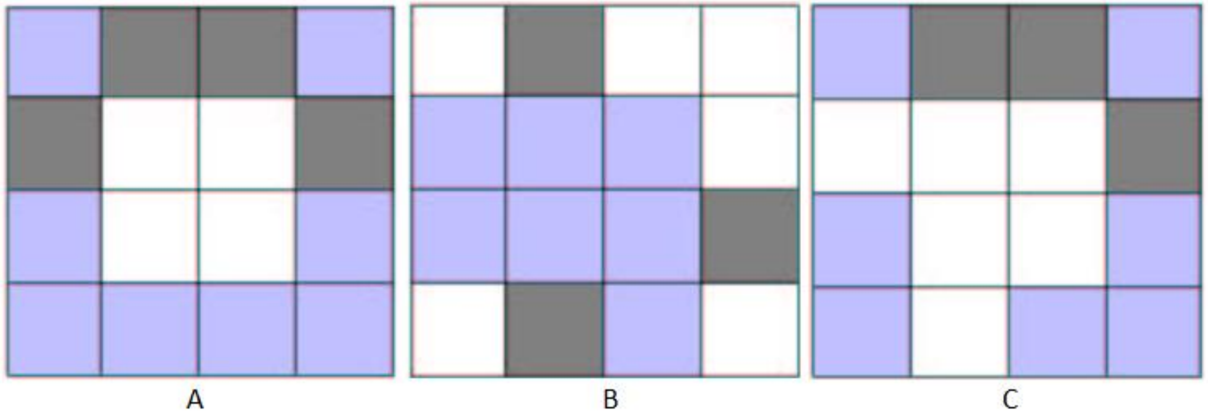
$$D(A,B) = \sum_{i=0}^{n}\sum_{j=0}^{n}\sum_{k=0}^{n}(A(i,j,k) - B(i,j,k))^2$$

Local Color Histogram:

LCH includes information about color's distribution in different regions. It's the same as GCH but at first we divide the image into different block. Where each pair of the blocks (one of them in the first image and the other in the second) will be computed separately using GCH. After that the total distance between the two images will be the sum of all GCH distances between them.

Remember : the larger the distance value, the less similar the images are.

GCH Example



Here C has the same color histogram as B but A is different from them.

Using Euclidian distance for these color histograms we found that D(A,C) = D(A,B) and D(B,C) = 0 but There's a problem here that B, C are not similar at all so D(B,C) shouldn't be zero and D(A,C) should be smaller than D(A,B) because A,C have the same pixels except for only two pixels.

GCH doesn't include information about color spatial distribution.

What is Texture ?

Texture for humans is a concept related to tactile or haptic perception: differences in regions of surfaces: rough or smooth

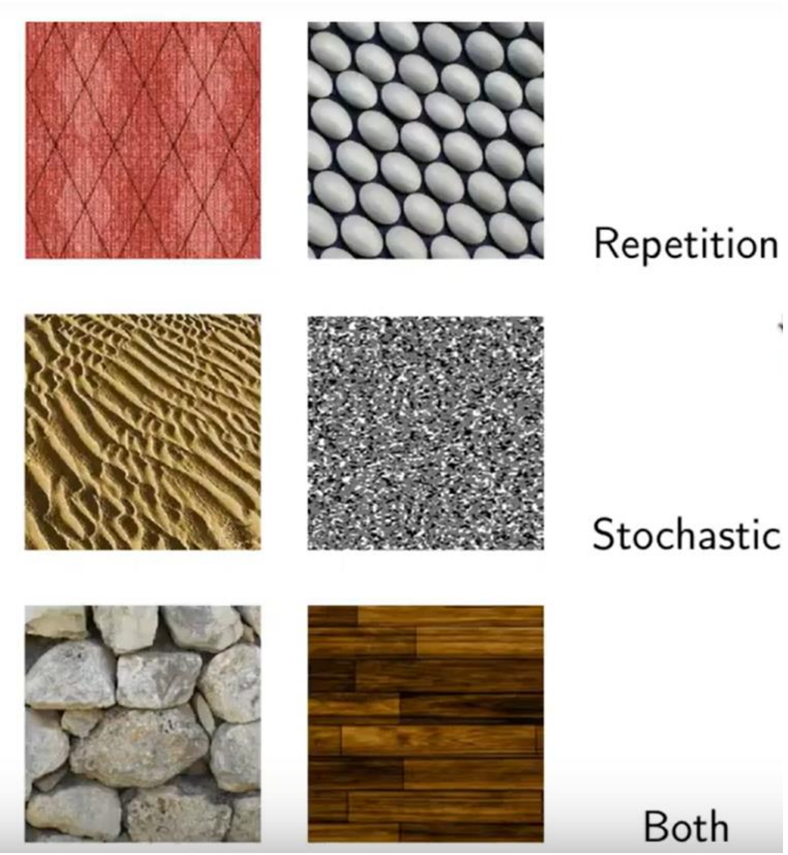
- Image textures is a concept related to local differences in the intensity levels, in which important elements are:
  - Differences in the pixel levels (contrast)
  - Size of regions to be considered (window)
  - Direction (ou lack of direction)

- Represent details in an image

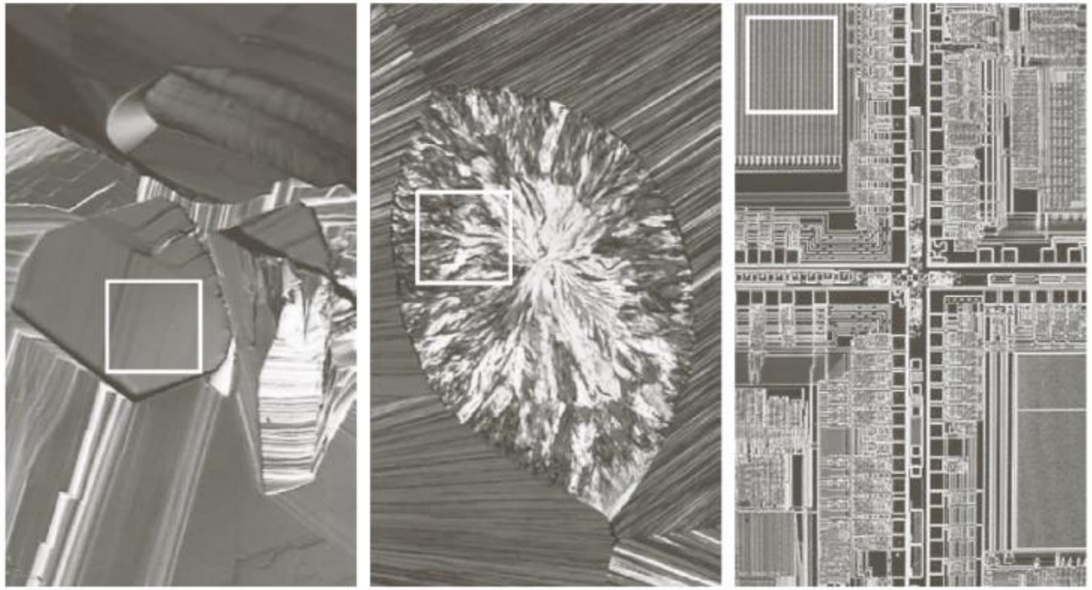
Texture with repeated local patterns

Local pattern

Texture Characteristics



Texture Analysis



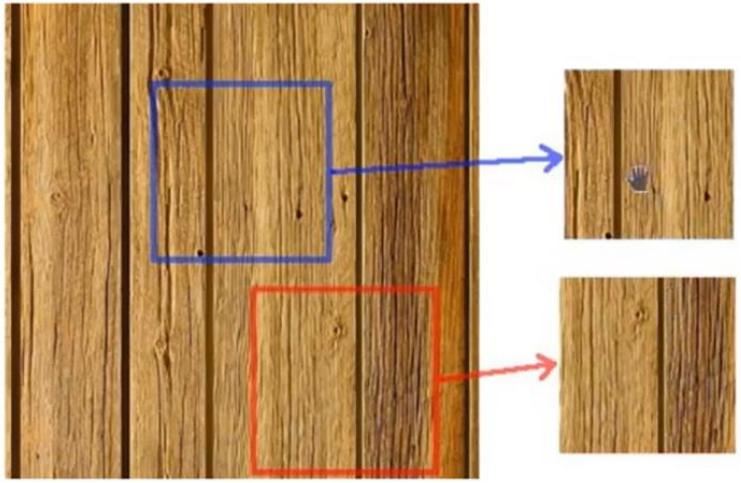
a b c

FIGURE 11.28 The white squares mark, from left to right, smooth, coarse, and regular textures. These are optical microscope images of a superconductor, human cholesterol, and a microprocessor. (Courtesy of Dr. Michael W. Davidson, Florida State University.)

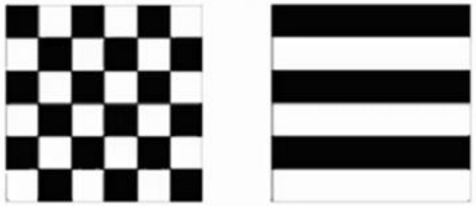
- An important approach to region description is to quantify its texture content
- This descriptor provides measures of smoothness, coarseness and regularity
- 3 approaches in describing the texture of a region: statistical, structural, and spectral.

Texture Analysis

- Compare textures to look for similar or different patterns:



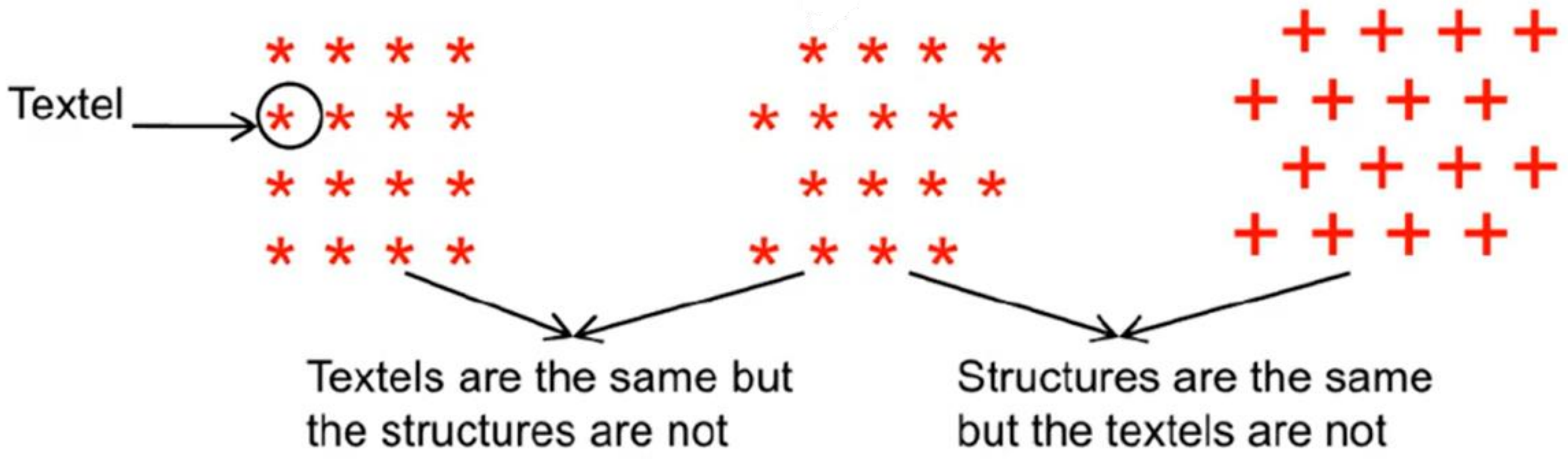
1. Structural (top–down)
   - decompose image in basic elements: texels (*texture elements*) / textons
   - adequate for artificial texture or well-behaved patterns



2. Statistical (bottom–up)
   - characterize texture as a series of statistical properties in a small group of pixels
   - often adequate for natural texture

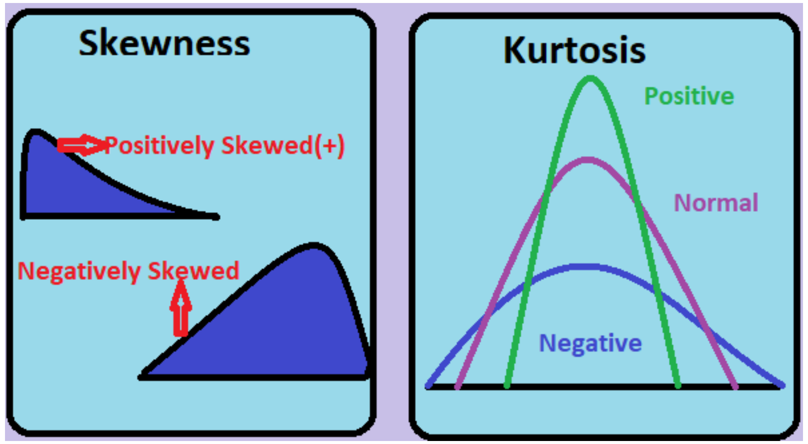
Texture Analysis: Structural Approach

- Texture represented via primitives with regular repeated patterns: textels
  - textel is a group of pixels with similar intensity properties: average value, contrast, flat regions, etc.
  - the granularity of the texture is given by the size of the "primitive".



Texture Analysis: Statistical Approach

- Define and segment regions of textels can be a challenge for natural scenes
  - textures look similar but it is hard to extract their structure.



Low number of athletes running 100m in less than 11sec

Measures computed over the values or the histogram.

- Fixing a window (region) of pixels, compute the mean, standard deviation, skewness and kurtosis
- Other methods such as uniformity and entropy are often computed using the histogram

Texture Analysis: Statistical Approach

- One of the simplest approaches for describing texture is to use statistical moments of the histogram of an image or a region

$$\mu_n(z) = \sum_{i=0}^{L-1} (z_i - m)^n p(z_i)$$

$$m = \sum_{i=0}^{L-1} z_i p(z_i)$$

- $\sigma^2(z) = \mu_2(z)$ is a measure of contrast

$$R = 1 - \frac{1}{1 + \sigma^2(z)}$$

- Third moment is a measure of skewness of histogram

$$\mu_3(z) = \sum_{i=0}^{L-1} (z_i - m)^3 p(z_i)$$

- Measure of uniformity: U is maximum for an image in which all gray levels are equal

$$U = \sum_{i=0}^{L-1} p^2(z_i)$$

- Average entropy: a measure of variability and is zero for constant image

$$e = -\sum_{i=0}^{L-1} p(z_i) \log_2 p(z_i)$$

Texture: Statistical

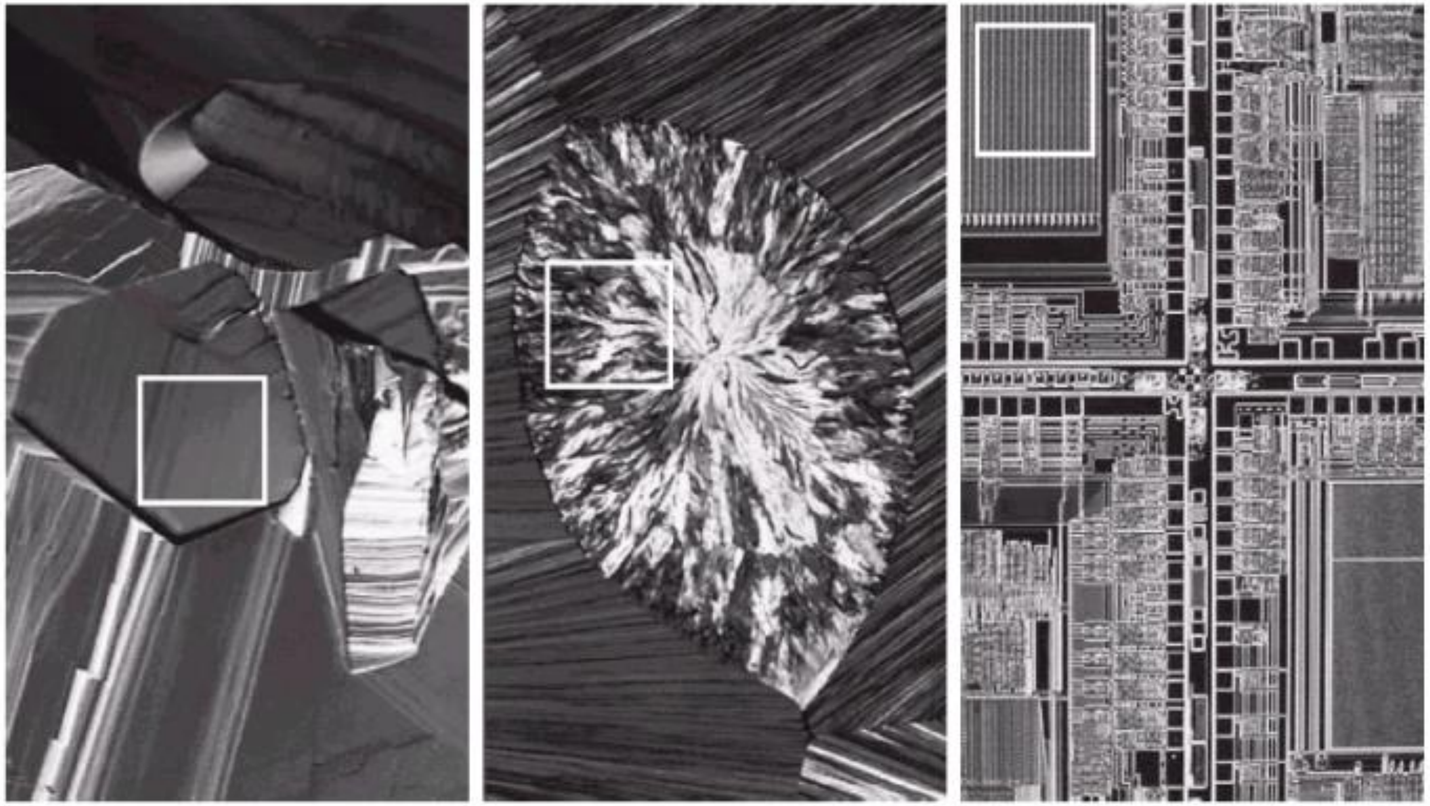


TABLE 11.2 Texture measures for the subimages shown in Fig. 11.22.

Texture	Mean	Standard deviation	R (normalized)	Third moment	Uniformity	Entropy
Smooth	82.64	11.79	0.002	−0.105	0.026	5.434
Coarse	143.56	74.63	0.079	−0.151	0.005	7.783
Regular	99.72	33.73	0.017	0.750	0.013	6.674

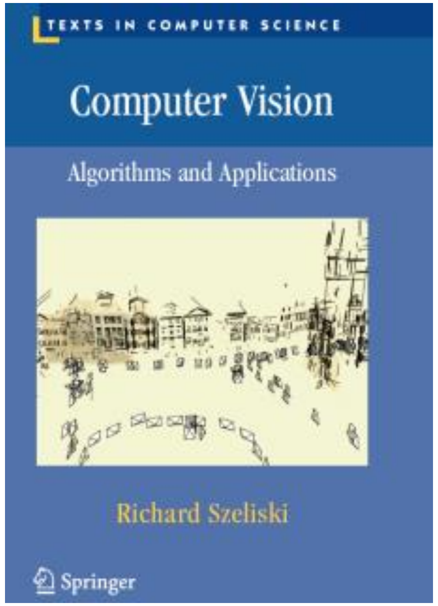
Text Books and References

❖Text Books

- ➢ Richard Szeliski. ”Computer vision: Algorithms and Applications. Springer Nature, Second Edition”, 2022
- ➢ E. R. Davies, Computer Vision Principles, Algorithms, Applications, Learning, Elsevier,5th Edition, 2017

❖ References

- ➢ Rafael C. Gonzales, Richard E. Woods, “Digital Image Processing”, Fourth Edition, Pearson Education, 2018.
- ➢ Richard Szeliski ,”Computer Vision: Algorithms and Applications” , Springer 2015
- ➢ Intro to Digital Image Processing
- ➢ Digital Image Processing 2nd edition



PDF online

❖ Credits: Slides and Video content borrowed from:

- ➢ First Principles of Computer Vision
- ➢ https://robotacademy.net.au/
- ➢ **MIT 6.S094: Computer Vision**
- ➢ **MIT Introduction to Deep Learning | 6.S191**
- ➢ Digital Image Processing 2nd edition