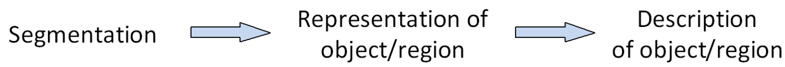
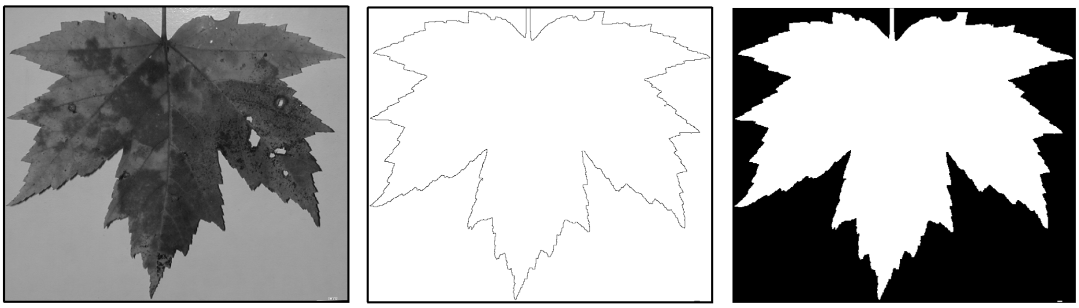
9 – REPRESENTATION AND DESCRIPTION (A)

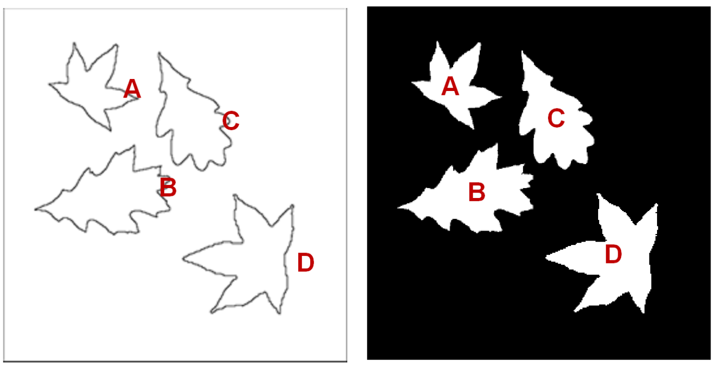
After an image has been segmented into regions, the resulting collection of segmented pixels are usually represented and described in a form suitable for further processing.

The representation of an object may be based on

- the boundary
$$S = \{b_1, b_2, b_3, ...\}$$
where the b_i's are the boundary pixels
- the pixels comprising the object
$$S = \{p_1, p_2, p_3, ...\}$$
where the p_i's are the '1' pixels that comprise the object





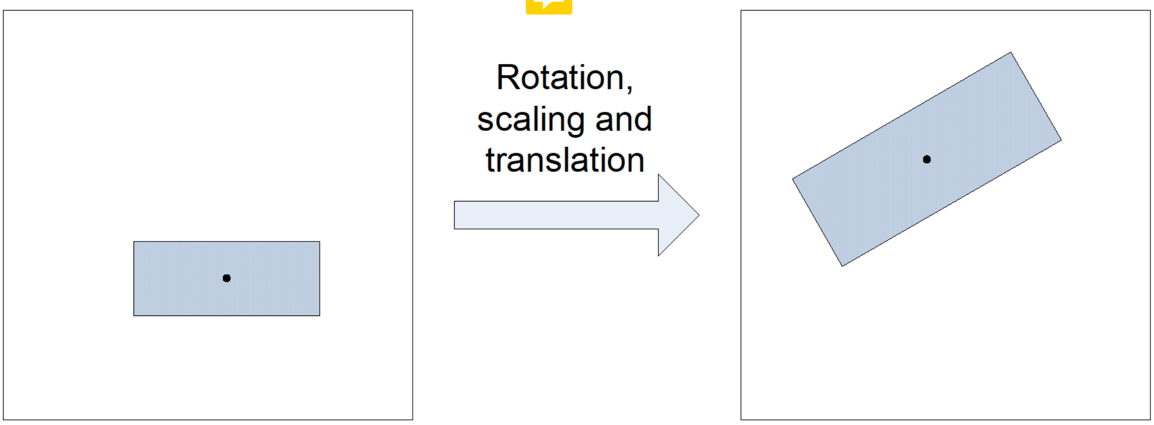


Object	Perimeter	Area
A	643	2900
B	883	4600
C	721	6600
D	901	6300

Features/descriptors

The object then has to be described based on the chosen representation. For example, a region may be represented by its boundary, with the boundary described by its perimeter, the orientation of the straight line joining the extreme points, etc.

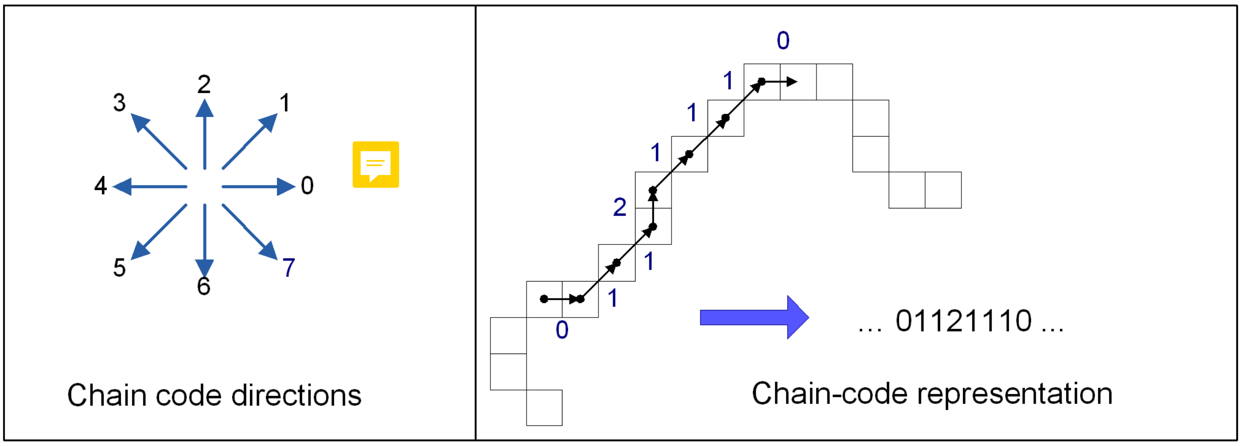
Features have to be extracted from the object for the purpose of recognition. In some machine vision applications, these descriptors should be independent of object orientation, size and location, or RST-invariant (rotation-scale-translation). They should also be sufficiently discriminatory.



BOUNDARY-BASED SCHEMES

Chain Codes

Chain codes are used to represent a boundary by a connected sequence of line segments of specified length and direction. Descriptors such as boundary length (perimeter), enclosed area, and boundary curvature can be easily computed from chain codes.



Chain code directions

Chain-code representation

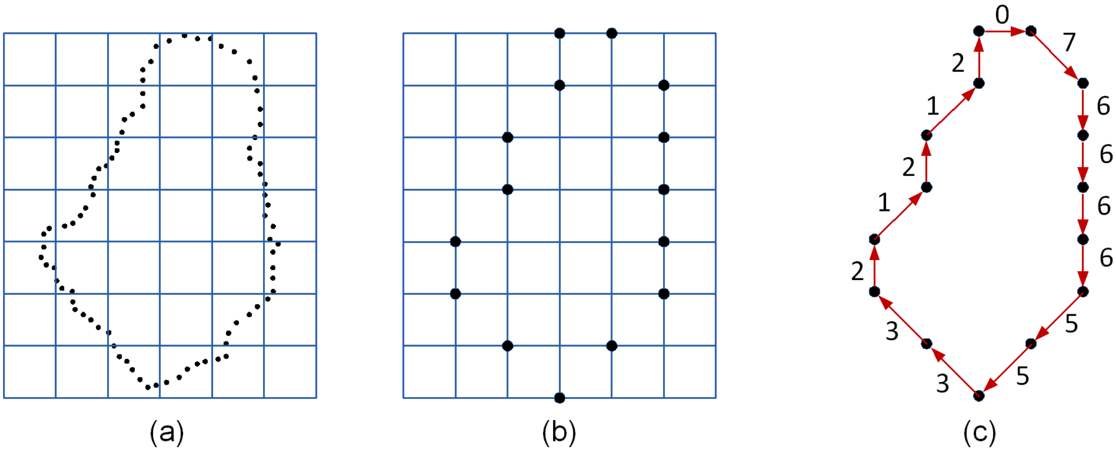
–

Method 1

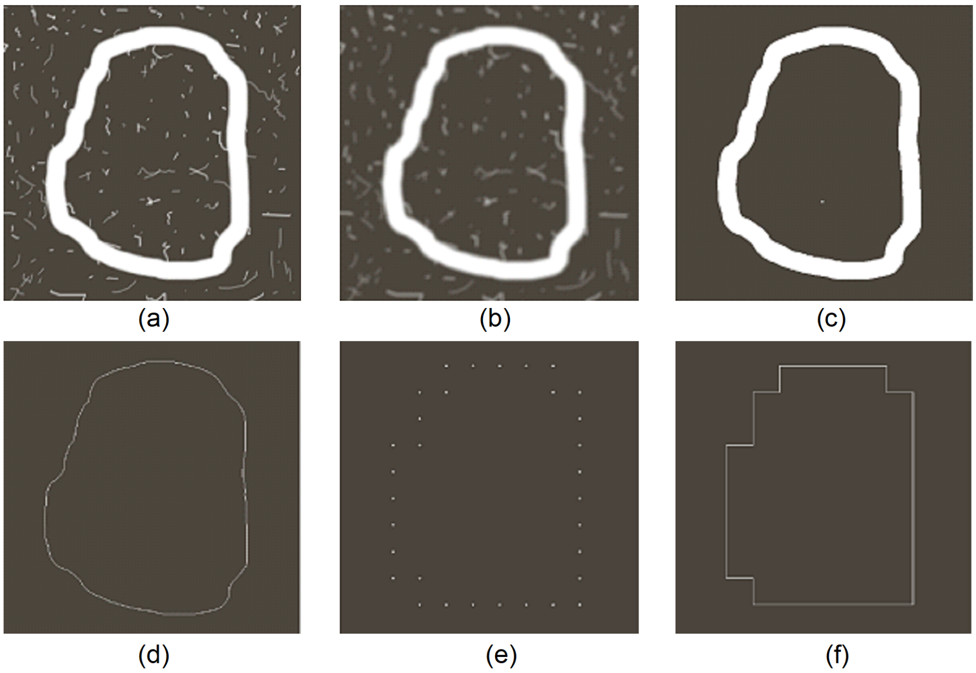
Traverse the boundary in a clockwise direction and assign a direction to the segments connecting every pair of pixels. However, the resulting chain code is usually quite long and sensitive to noise effects.

Method 2

Resample the boundary by selecting a larger grid spacing. As the boundary is traversed, assign a boundary point to the node of the grid to which it is closest.



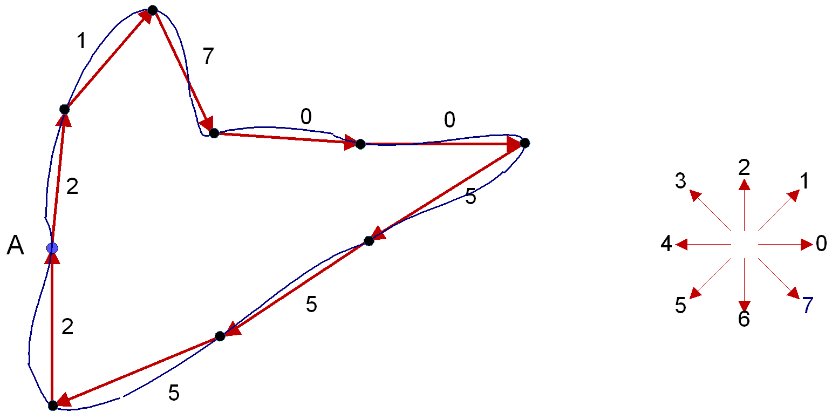
(a) Digital boundary with resampling grid superimposed. (b) Result of resampling. (c) Chain code.



(a) Noisy image. (b) Smoothed image. (c) Smoothed image after thresholding. (d) After boundary tracking. (e) Resampled boundary in a grid with nodes 50 pixels apart. (f) Connected boundary points. (G&W)

Method 3

Divide the boundary into segments of equal length (each segment having the same number of pixels). Join the endpoints of each segment by a straight line. The resulting code is 217005552.



Generation of chain code by boundary subdivision

Normalization for starting point

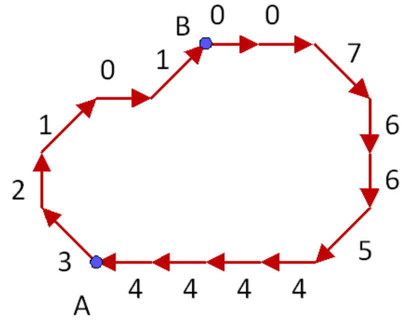
For the same boundary, starting at different points will result in different chain code representations. To ensure that the identical boundaries are represented by the same sequence of numbers, we may employ a normalization procedure. This is achieved by redefining the starting point so that the resulting sequence of numbers forms an integer of minimum magnitude. In the figure below, if we start from an arbitrary point, A, the code is

$$321010076654444$$

The smallest integer that we can derive from this code is

$$007665444432101$$

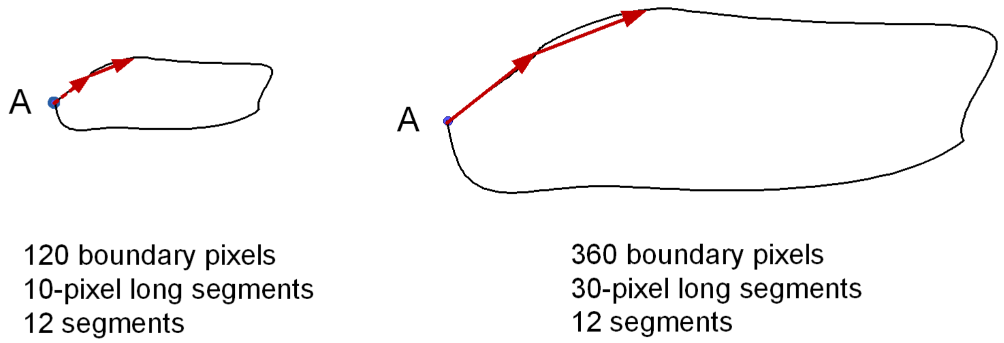
equivalent to starting at B.



Size normalization

A larger version of a boundary will be represented by a longer sequence of numbers compared to the smaller version. By employing a normalization procedure, we can represent a boundary by a code that does not change in length even though there is scaling.

This is achieved by subdividing all object boundaries into the same number of equal segments and adjusting the code segment lengths to fit the subdivisions.

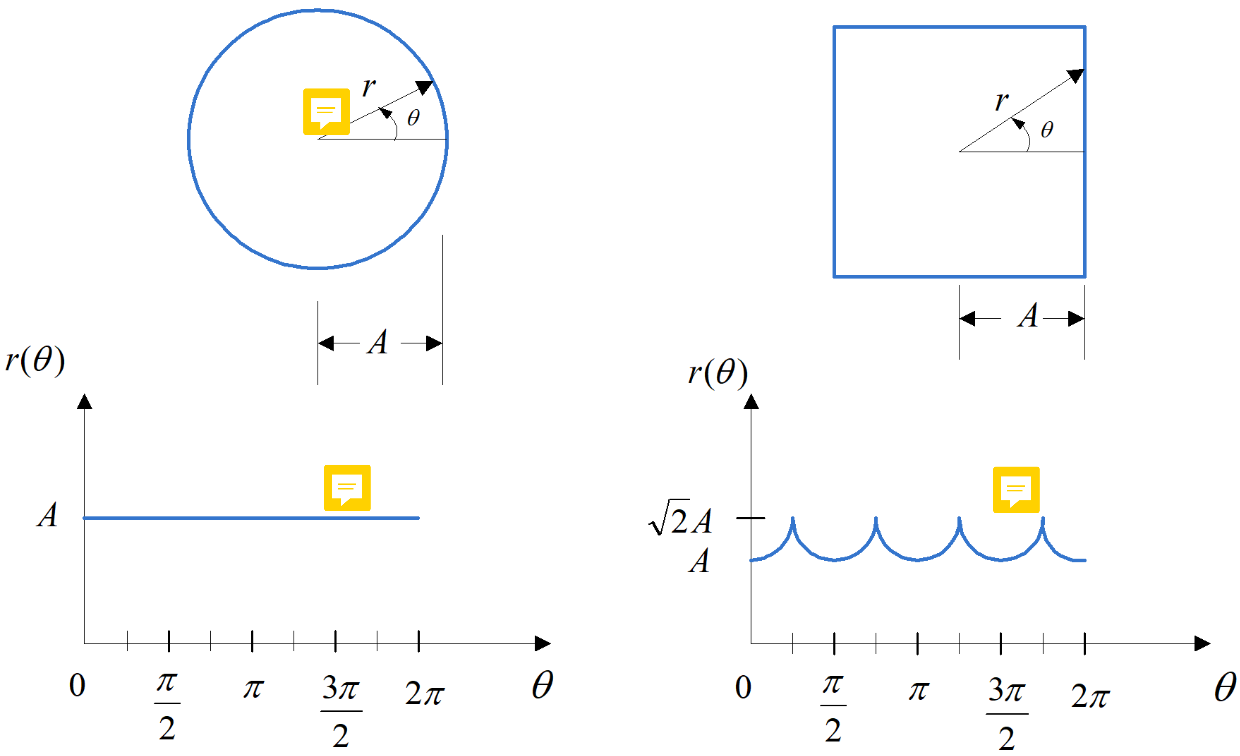


Signatures

A signature is a one-dimensional functional representation of boundary. The basic idea is to reduce the (2-D) boundary representation to a one-dimensional function.

(a) Distance-angle function

It is obtained by plotting the distance of the boundary from the centroid as a function of angle. The result is also known as the $r - \theta$ plot or centroidal profile.



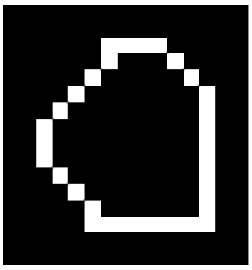
The centroid may be estimated from the boundary pixels

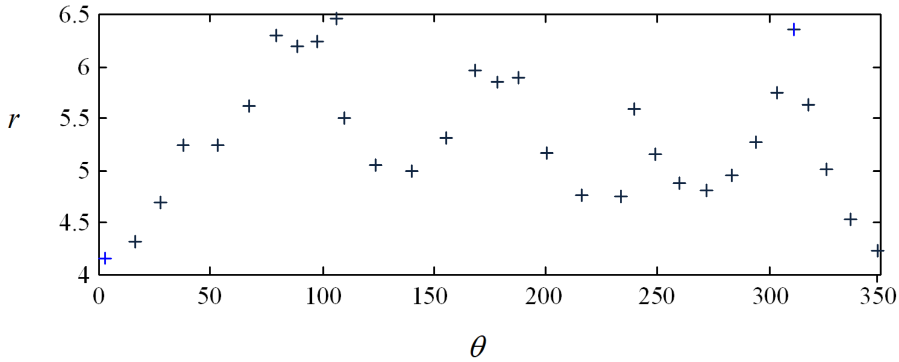
$$B = \{(x_i, y_i), \quad i = 1, 2, \ldots, N\}$$

The coordinates of the centroid are given by

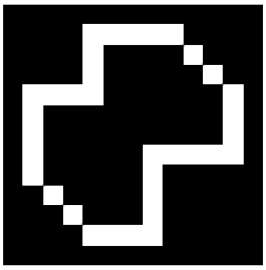
$$\bar{x} = \frac{1}{N}\sum_i x_i \qquad \bar{y} = \frac{1}{N}\sum_i y_i \tag{1}$$

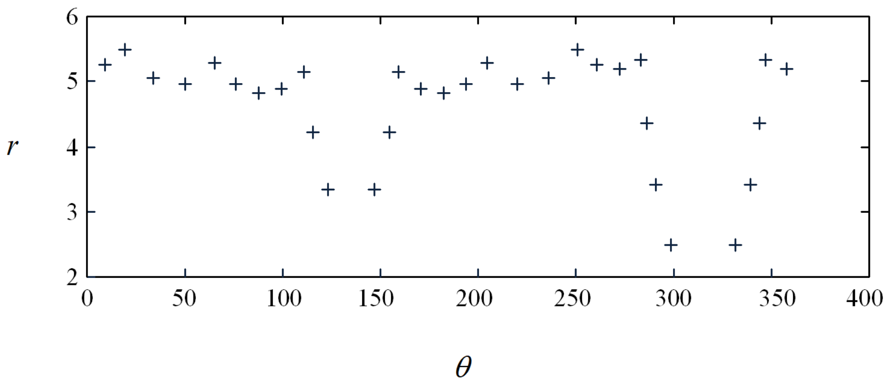
Boundary B1
32 points

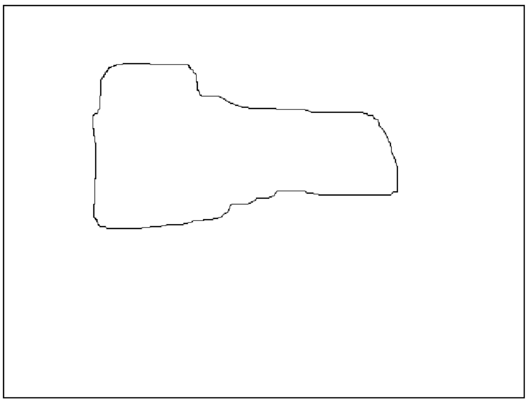




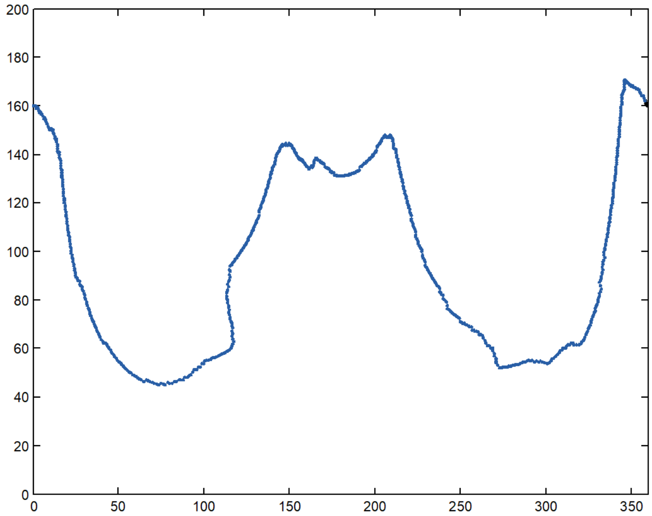
Boundary B2
32 points



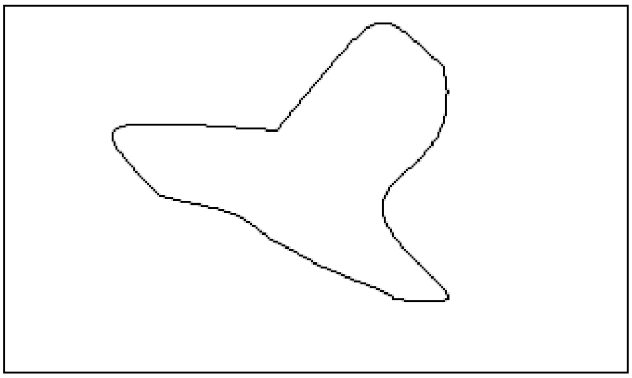




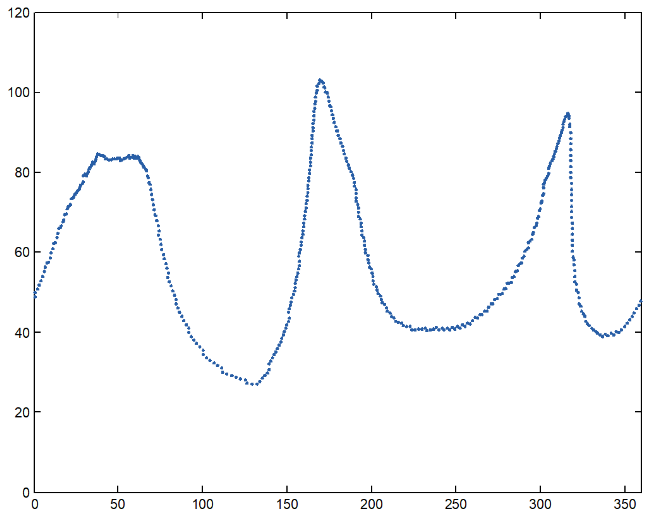
Boundary



Distance-angle function



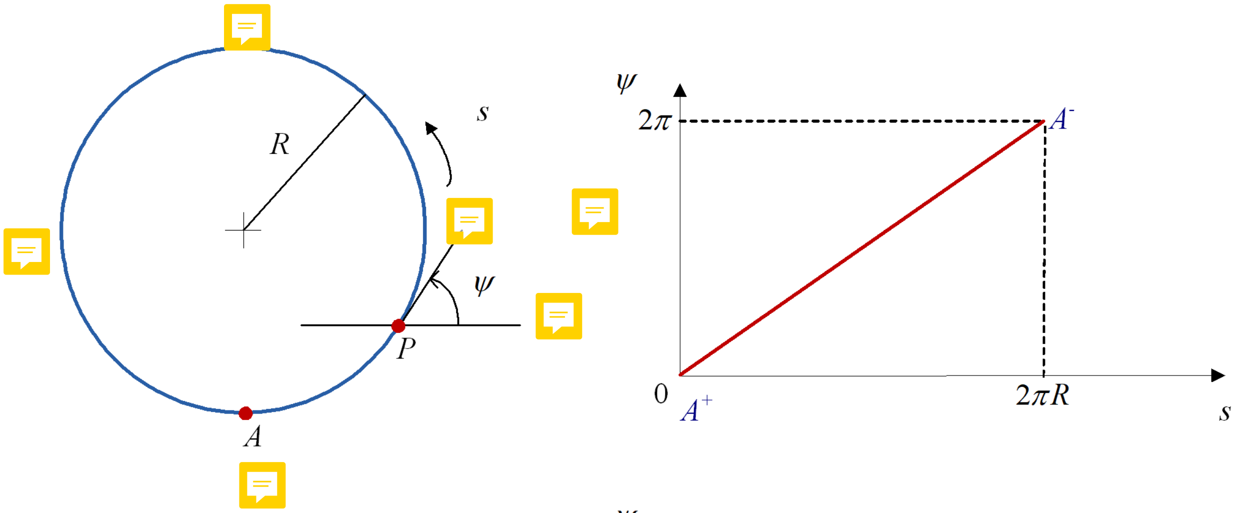
Boundary

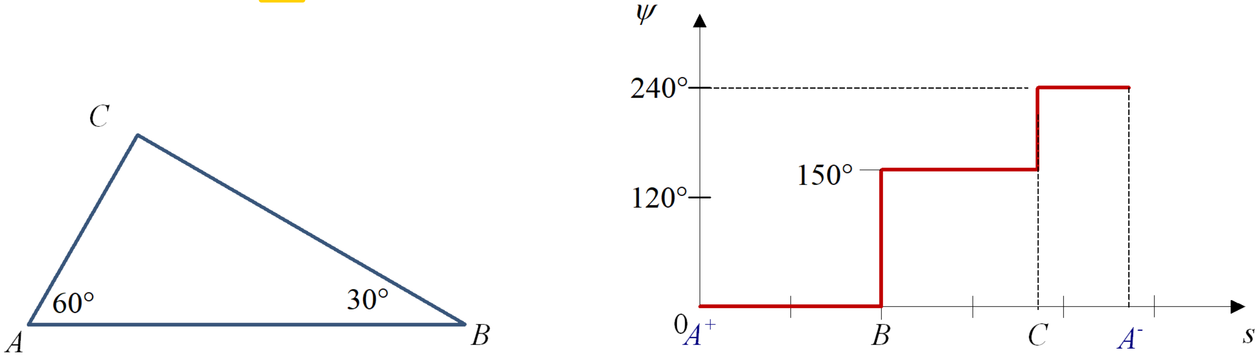


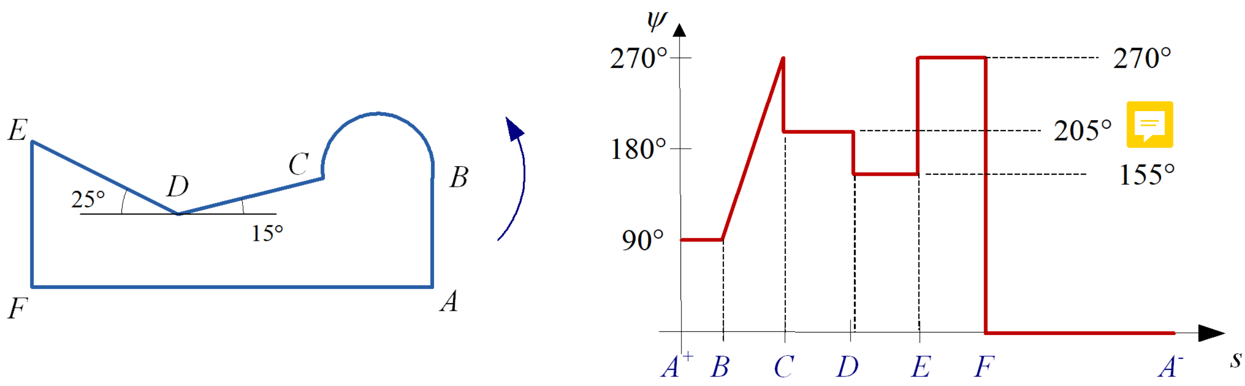
Distance-angle function

(b) Slope-density function

- Also known as the ψ-s plot.
- A plot of the tangential orientation ψ $(0 \leq \psi \leq 2\pi)$, as a function of the boundary distance s.
- The angle ψ is measured with respect to a fixed line, usually the horizontal line, and the tangent to the boundary.
- The starting point (A) is arbitrary.







Boundary Matching with Signatures

How do we differentiate between signatures corresponding to different boundary shapes? Let a be a discrete variable denoting amplitude variations in a signature, and let $p(a_i)$, $i = 1, 2, \ldots, K$, denote the corresponding histogram:

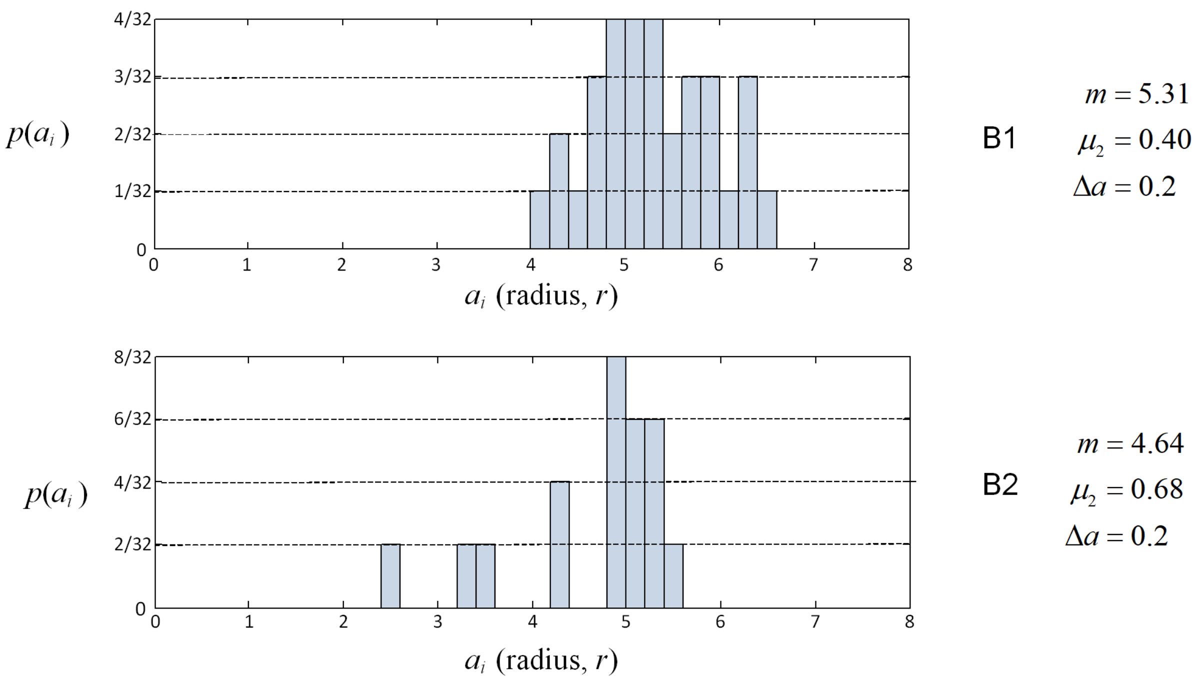
- a represents r in the case of the distance-angle function and ψ in the case of the slope-density function
- K is the number of discrete amplitude increments of a
- $\sum_{i=1}^{K} p(a_i) = 1$.

The nth moment of a about its mean, m is

$$\mu_n(a) = \sum_{i=1}^{K} (a_i - m)^n p(a_i)$$

$$m = \sum_{i=1}^{K} a_i p(a_i).$$

Only the first few moments are required to differentiate between signatures of clearly distinct shapes.



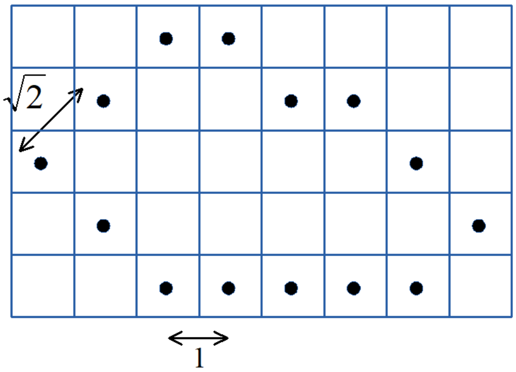
Shape Descriptors

Area

The area of a region is given by the total number of pixels enclosed by the boundary and inclusive of the boundary pixels.

Perimeter

The perimeter can be approximated by the number of boundary pixels. A more accurate method is to multiply the diagonal boundary segments by a factor $\sqrt{2}$.



number of boundary pixels = 14

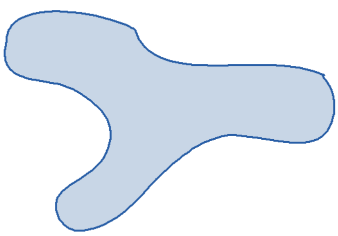
more accurate estimate $= 6 + 8\sqrt{2} = 17.3$

Compactness

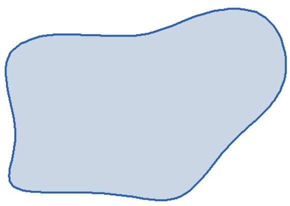
Defined as

$$\gamma = \frac{(\text{perimeter})^2}{4\pi \times (\text{area})} \tag{2}$$

For a disk, γ is minimum and is equal to 1.

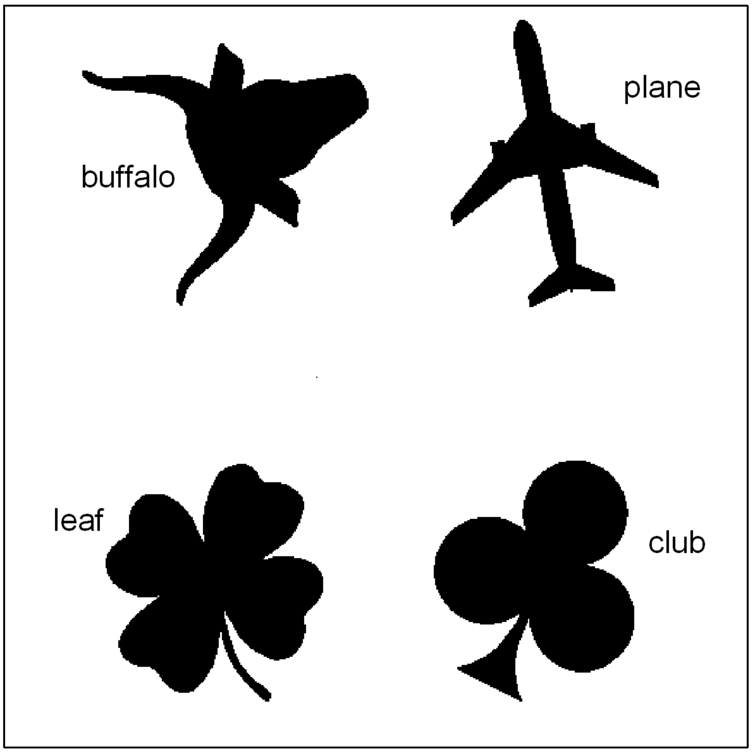


High compactness



Low compactness

Example



	Area		Perimeter		Compactness	
	Raw values	Norm. values	Raw values	Norm. values	Raw values	Norm. values
buffalo	10472	69	747.5	93	4.25	62
plane	6680	44	757.6	95	6.84	100
leaf	15178	100	799.7	100	3.35	49
club	13878	91	683.0	85	2.67	39

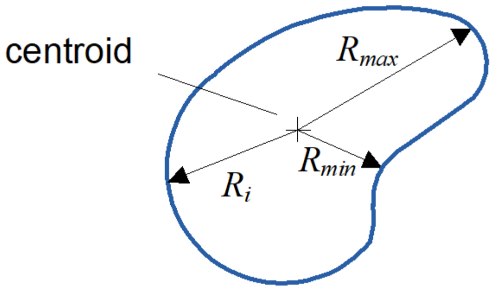
Radial distances

The radial distance R_i is the distance of a boundary pixel from the centroid. For a boundary comprising N pixels:

$$\text{Average radial distance, } R_{av} = \frac{1}{N}\sum_{i=1}^{N} R_i \quad (3)$$

$$\text{Minimum radial distance, } R_{min} = \min(R_i) \quad (4)$$

$$\text{Maximum radial distance, } R_{max} = \max(R_i) \quad (5)$$

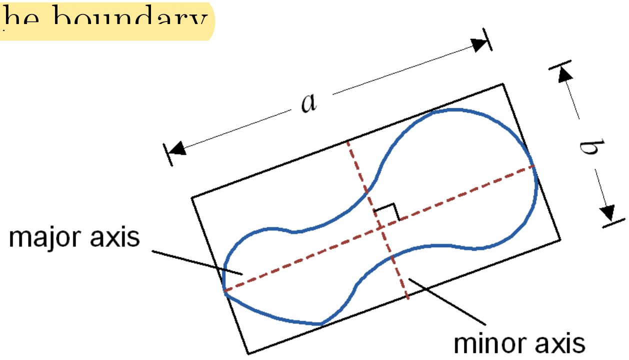


Bounding box

The major axis is the straight-line segment joining the two boundary points furthest away from each other.

The minor axis

- is perpendicular to the major axis
- intersects the mid-point of the major axis
- is of length such that a bounding box can be formed that just encloses the boundary



Eccentricity

Eccentricity is defined as

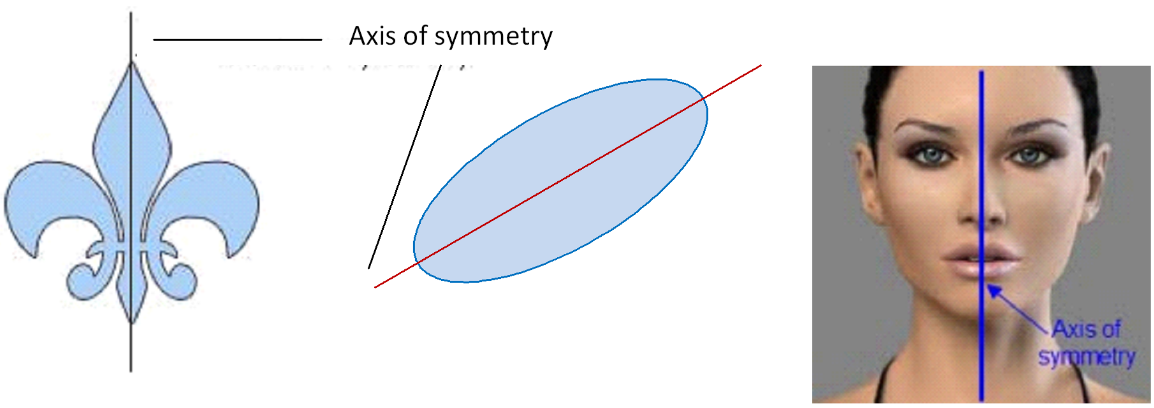
$$\epsilon = \frac{\text{length of major axis}}{\text{length of minor axis}} \quad \text{or} \quad \frac{\text{maximum radial distance}}{\text{minimum radial distance}} \tag{6}$$

(It is given by a/b in the above diagram.)

Axis of symmetry

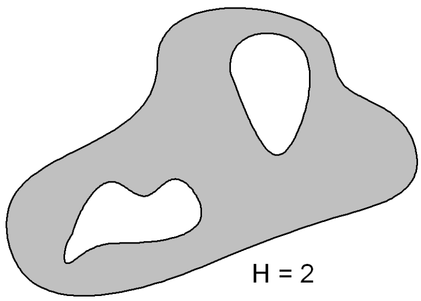
An axis of symmetry is a line that divides the figure into two symmetrical parts in such a way that the figure on one side is the mirror image of the figure on the other side.

When the shape is folded in half along the axis of symmetry, then the two halves are congruent.

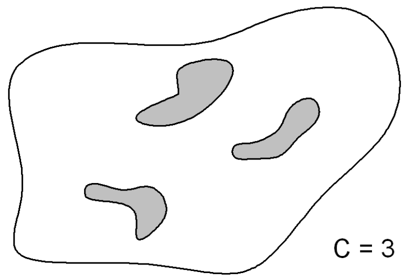


Topological Descriptors

Topology is the study of properties of a figure that are unaffected by any deformation, as long as there is no tearing or joining of the figure.



Number of holes, H

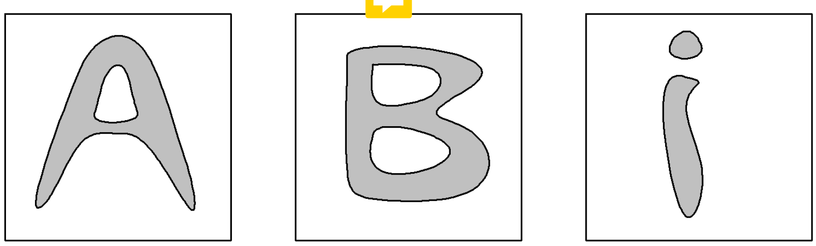


Number of connected components, C

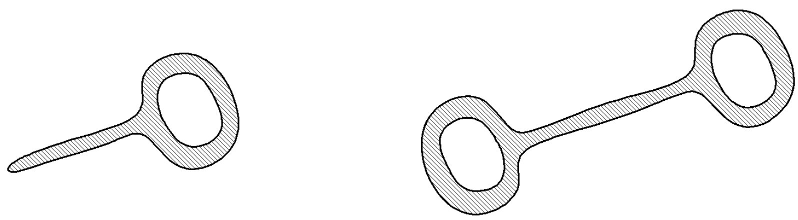
The Euler number is defined as

$$E = C - H. \tag{7}$$

The Euler number is invariant to RST transformations.



E = 0 E = -1 E = 2



E = 0 E = -1