Set 10: Representation and Description

Lecturer Arto Kaarna

Lappeenranta University of Technology (LUT)
School of Engineering Science (LENS)
Machine Vision and Pattern Recognition (MVPR)

Arto.Kaarna@lut.fi
http://www.lut.fi/web/en/school-of-engineering-science/research/machine-vision-and-pattern-recognition

Contents

- Representation schemes
- Boundary descriptors
- Regional descriptors
- Morphology: dilation, erosion, closing, opening
- Principal components
- Relational descriptors

Motivation

Task: 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 (classification, clustering)

Representation Schemes

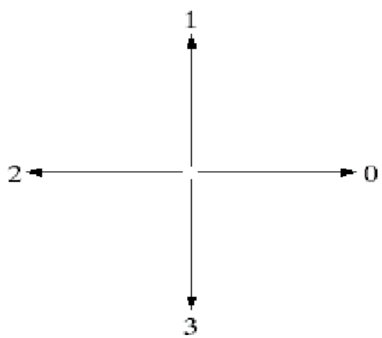
- Schemes that compact data into representations that are more useful in the computation of descriptors (features)
- Schemes:
  - Chain codes
  - Polygonal approximation
  - Signatures
  - Boundary segments
  - Skeletons

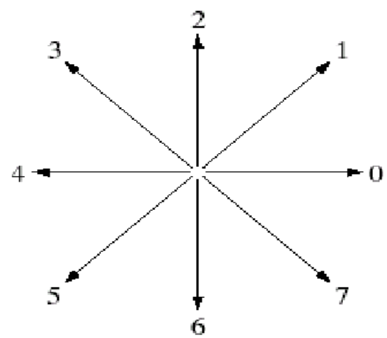
RS: Chain Codes

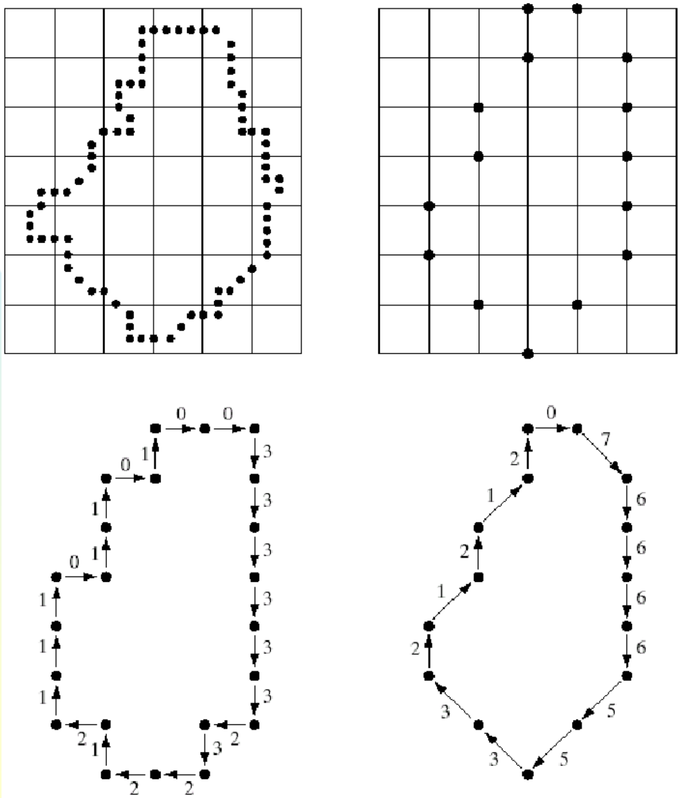
- Chain codes are used to represent a boundary by a connected sequence of straight-line segments of the specified length and direction
- 4-directional chain and 8-directional chain codes
- The chain code of the boundary depends on the starting point
  - Redefine the starting point so that the resulting sequence of numbers forms an integer of minimum magnitude
  - Normalization for rotation: the first difference of the chain code instead of the code itself

RS: Chain Codes

- 4- and 8-directional chain code (make unambiguos by selecting starting point so that the magnitude of the resulting integer is minimal)

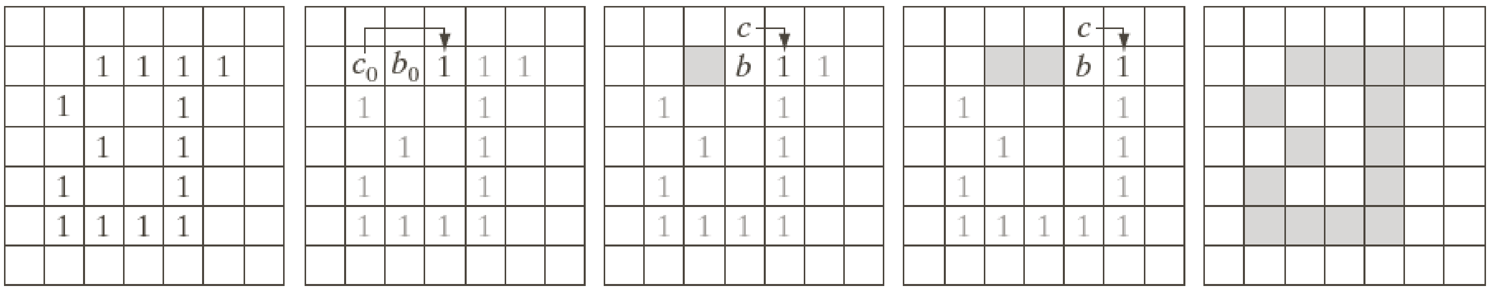






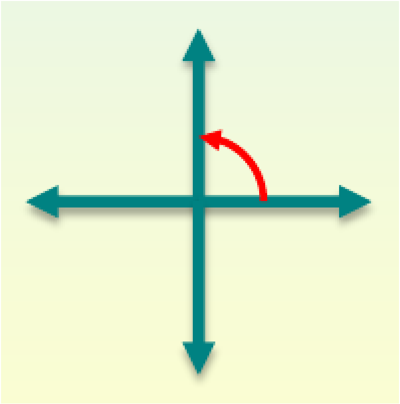
RS: Chain Codes

- Steps in creating the chain code



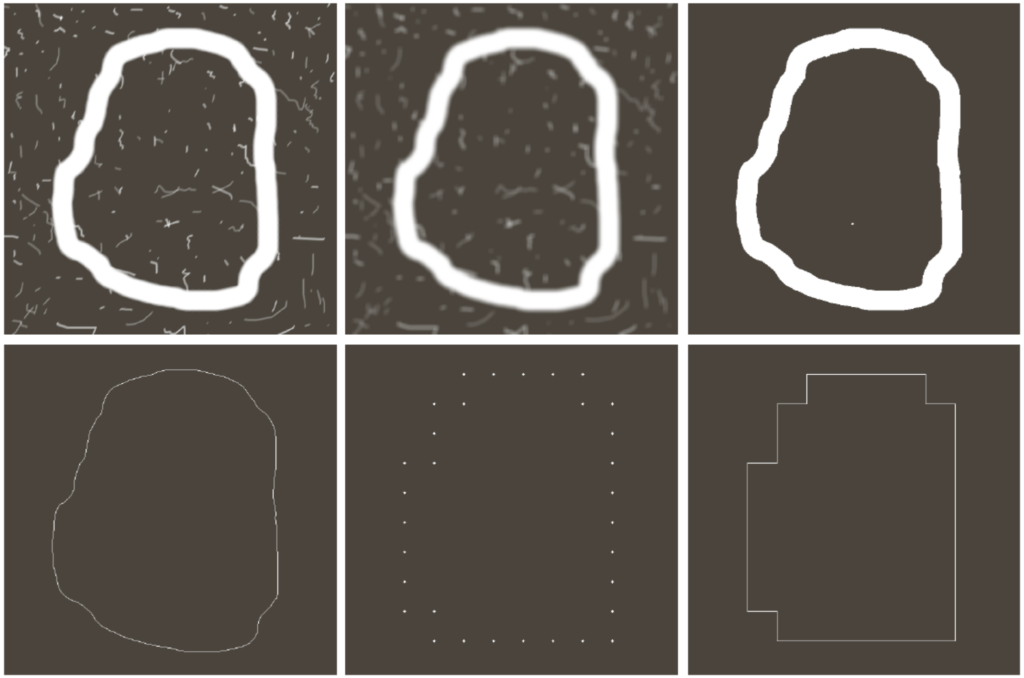
- First difference: count direction changes between consequtive chain codes → rotation invariance, e.g.

CC	FD
01	1
02	2
03	3



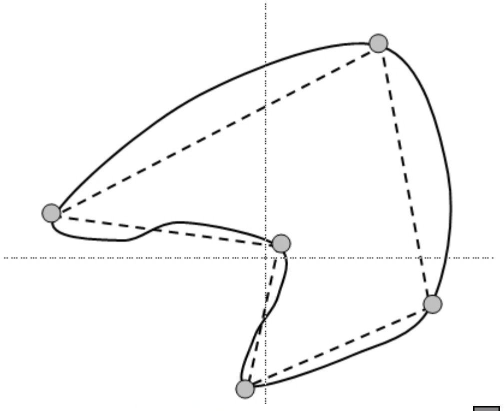
RS: Chain Codes

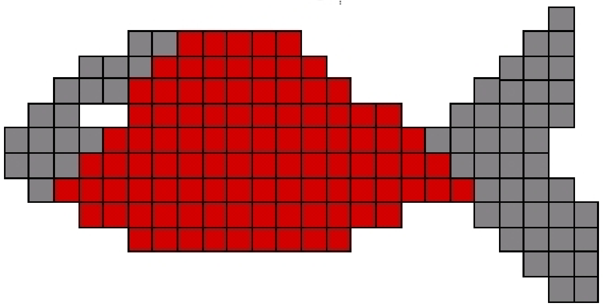
- For a noisy image, the first step is to remove noise, then threshold, find the boundary, sample the boundary, and finally connect the points in the sampled boundary



RS: Polygonal approximation

- A digital boundary can be approximated with arbitrary accuracy by a polygon
- Minimum perimeter polygons
  - A polygon of minimum perimeter fitted to the object boundary enclosed by cells
- Splitting techniques
  - Subdivision of a segment successively into two parts until a given criterion is satisfied

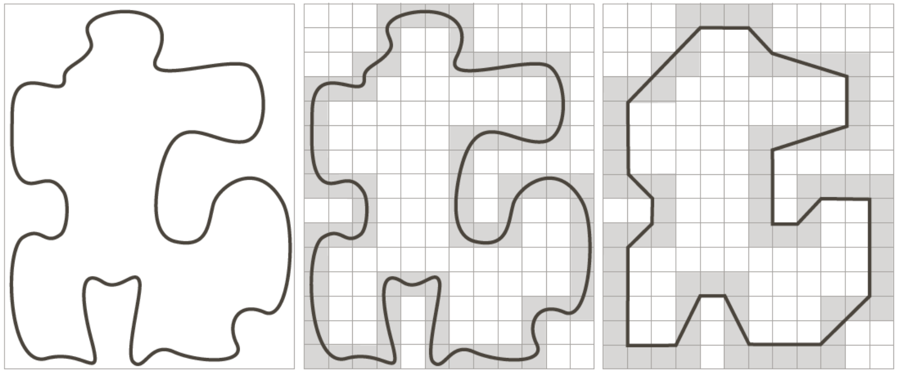




Maximal inscribed polygon

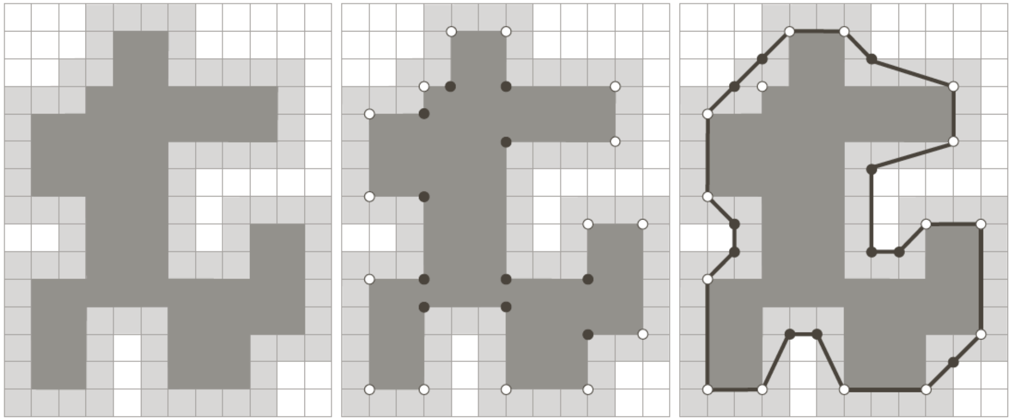
RS: Polygonal approximation

- Finding cells, defining the minimum-perimeter polygon by following the corners/inner/outer corners of the cells



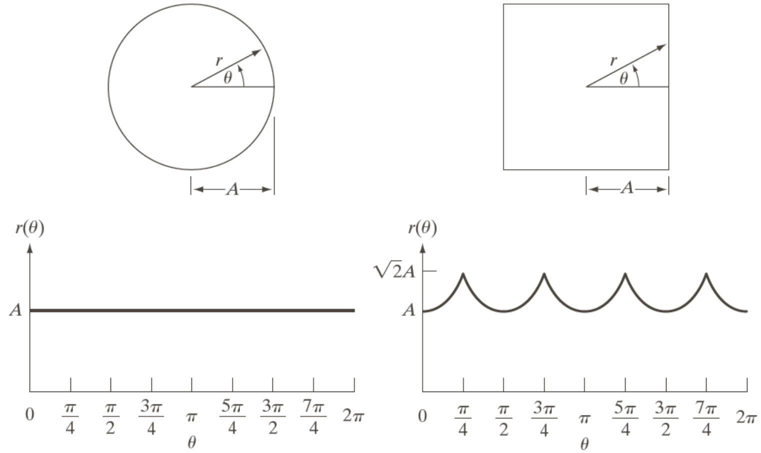
- Minimum perimeter polygon (MPP)

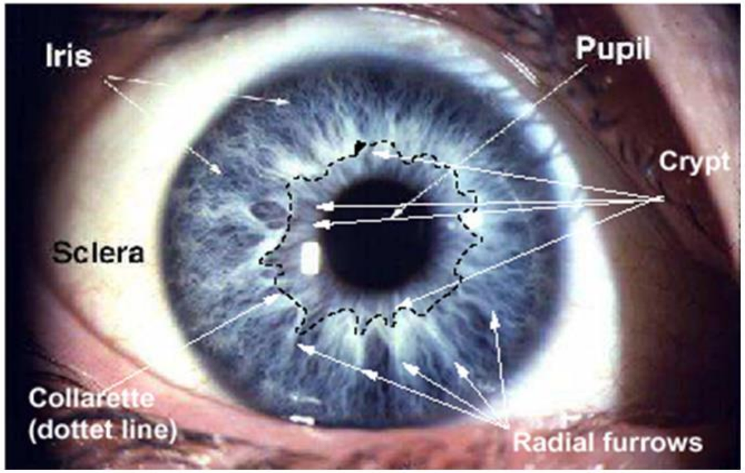
- Convex (W) and concave (B) vertices.

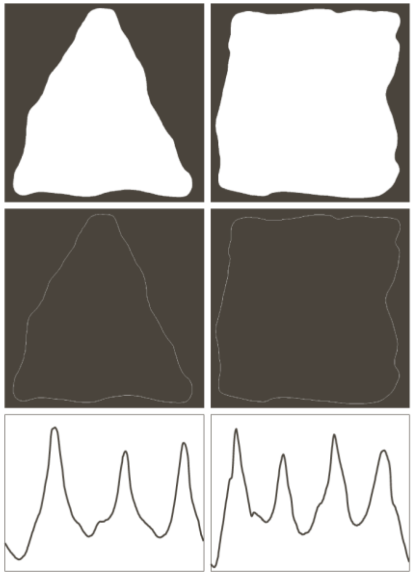


RS: Signatures

- A signature is a 1-D functional representation of a boundary
- To reduce the boundary representation to a 1-D function
- Invariance:
    - Rotation (if starting point is selected suitably)
    - Translation
    - Scaling

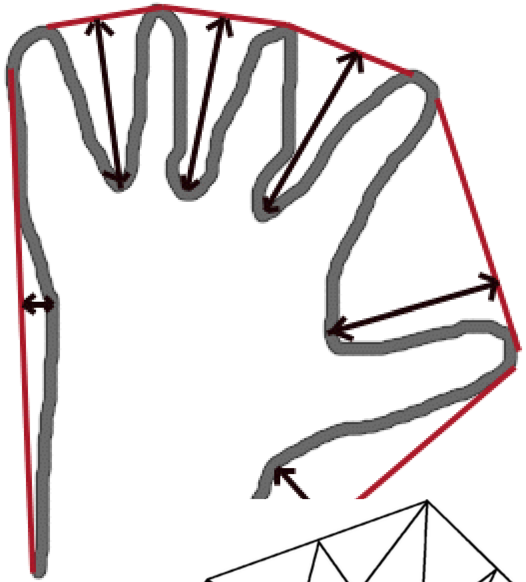


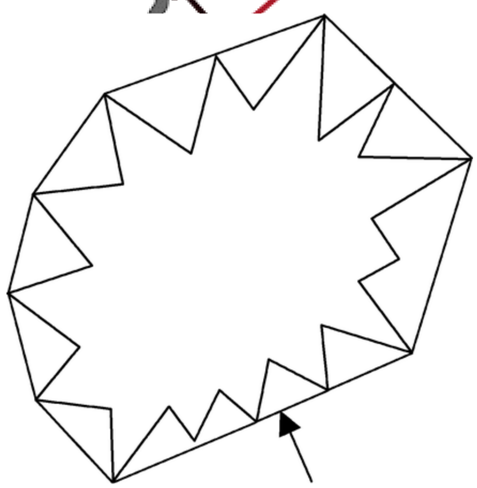


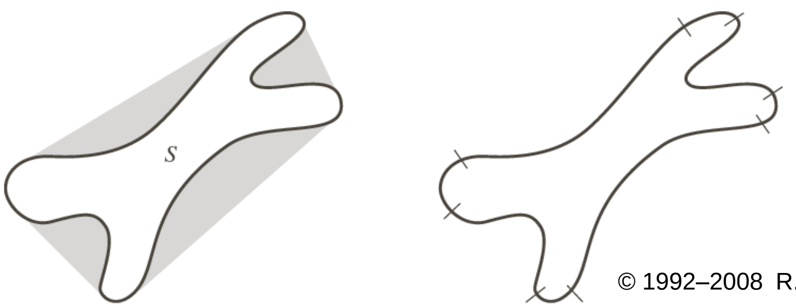


RS: Boundary segments

- Decomposing a boundary into segments
- The convex hull *H* of an arbitrary set *S* is the smallest convex set containing *S*
- The set *H* - *S* is called the convex deficiency *D* of the set *S*





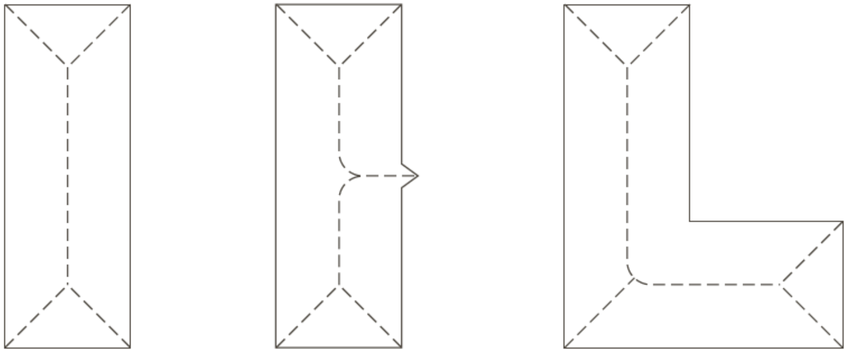


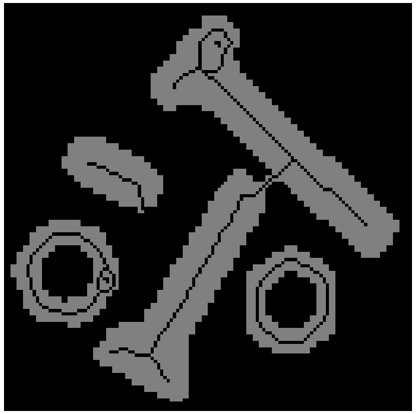
RS: Skeleton of a Region

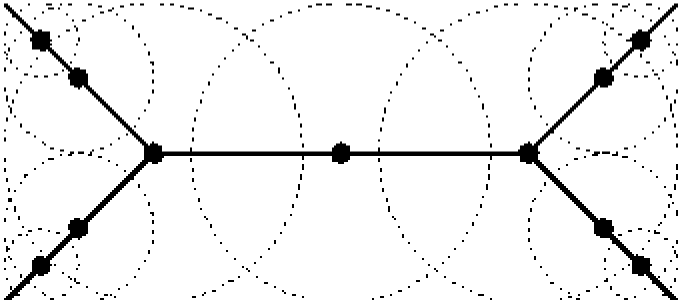
- To reduce the structural shape to a graph
- The skeleton is obtained with a thinning algorithm
- Medial axis transformation (MAT) of a region *R* with border *B* is as follows:
  - For each point *p* in *R* find the closest neighbor in *B.*
  - If *p* has more than one such neighbor it is said to belong to the medial axis (skeleton) of *R.*
- MAT depends on the choice of a distance measure (e.g., Euclidean distance) and it is also computationally expensive => many other thinning algorithms have been proposed

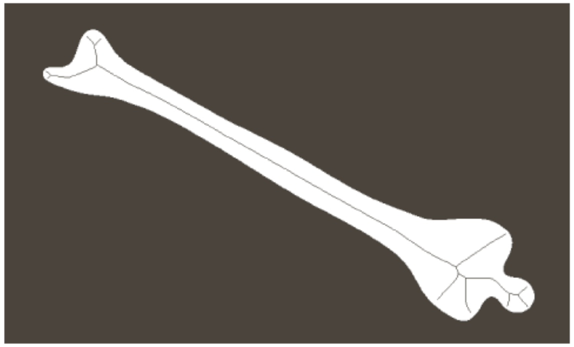
RS: Skeleton of a Region

- Medial axis of some regions









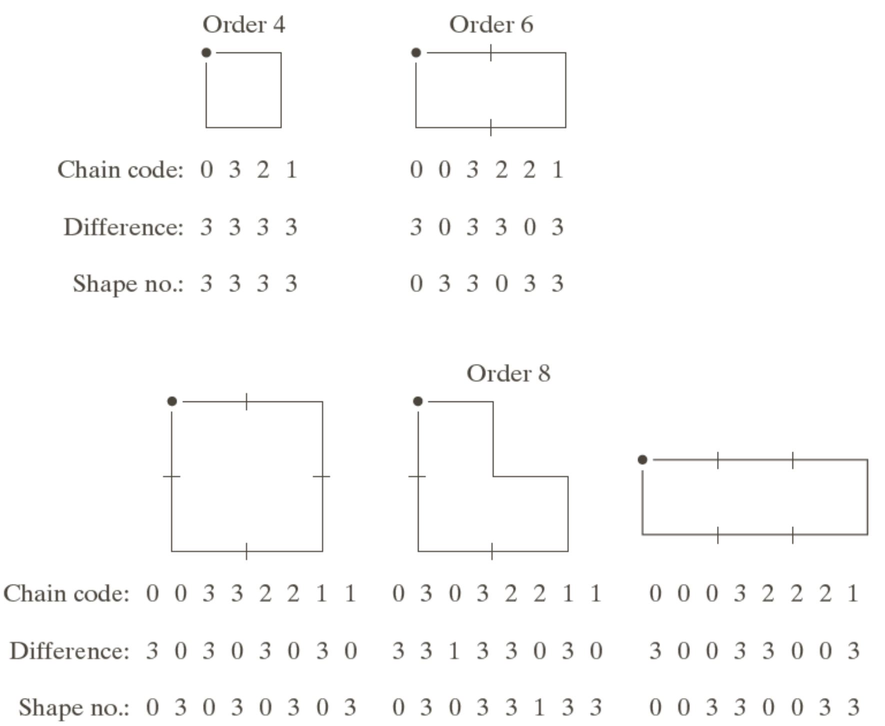
Boundary descriptors

- Simple descriptors
  - Length of a contour: number of pixels along the contour
  - Diameter of a boundary *B* (*D* = distance measure):
    $Diam(B) = \max[D(p_i, p_j)]$, p_i, p_j are points on boundary
  - Curvature: the rate of change of slope
- Shape numbers: the first difference of the smallest magnitude (from the chain code) (order: number of digits in the representation)
- Fourier descriptors
- Moments: the shape of boundary segments described by moments (mean value and variance of a 1-D function)

BS: Shape Numbers

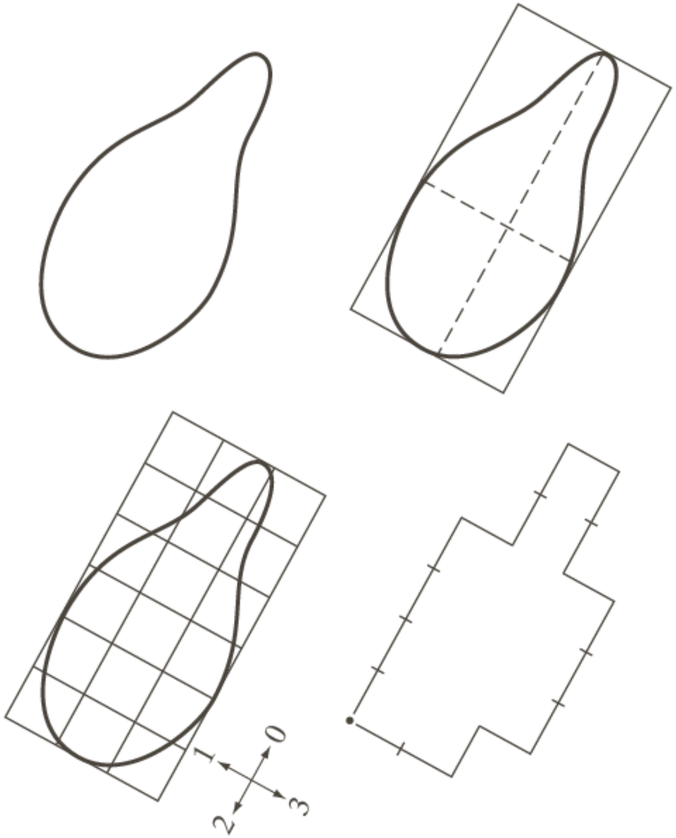
- Shape number of a boundary is defined as the first difference of the smallest magnitude (the code depends on the starting point)

- Shape numbers for
 some simple shapes,
 orders: 4,6, and 8



BS: Shape Numbers

- Steps in finding the shape number
  1. Find the chain code
  2. Find the first differences
  3. Select the first difference with the smallest magnitude.



Chain code:	0	0	0	0	3	0	0	3	2	2	3	2	2	2	1	2	1	1
Difference:	3	0	0	0	3	1	0	3	3	0	1	3	0	0	3	1	3	0
Shape no.:	0	0	0	3	1	0	3	3	0	1	3	0	0	3	1	3	0	3

BS: Fourier Descriptors

- The boundary is described as a complex number

$$s(k) = x(k) + jy(k)$$

- The DFT of $s(k)$ is $a(u), k = 0, K-1$
- An approximation is obtained when we set

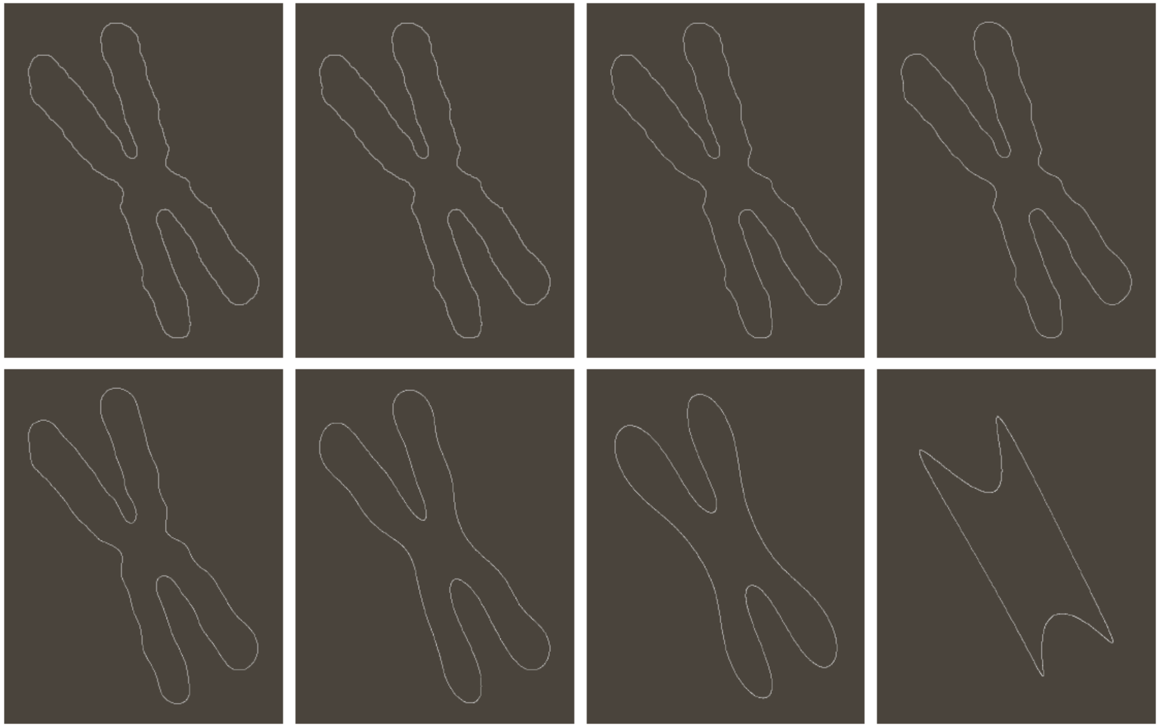
$$a(u) = 0 \text{ for } u > P-1,\ P < K$$

- The descriptors are then

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	$s_p(k) = s(k - k_0)$	$a_p(u) = a(u)e^{-j2\pi k_0 u/K}$

BS: Fourier Descriptors

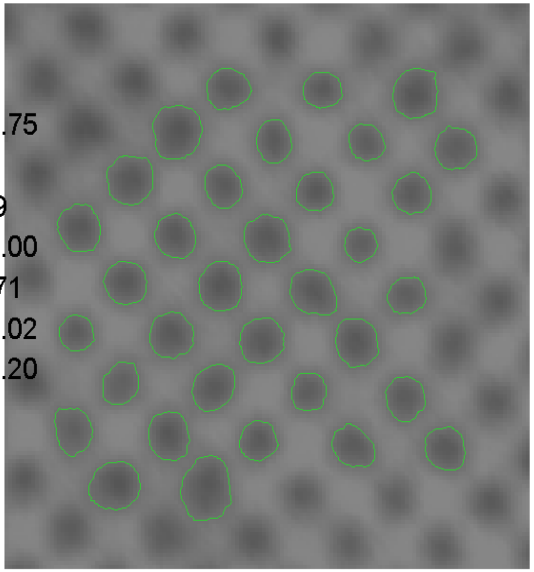
- Approximation, originally 2868 points on the boundary
- 50, 10, 5, 2.5, 1.25, 0,63, 0.28% of the points used



Regional Descriptors

- Simple descriptors
  - Area, perimeter, compactness, principal axes
- Topological descriptors
  - Holes, connected components, Euler number (e.g. the number of objects minus the number of holes in the objects in a binary image)
- Texture
- Moments

Dots	34
Area	94.15 ± 22.75
Stretch	6.93 ± 5.59
Consistency	100.00 ± 0.00
Regularity1	97.78 ± 1.71
Regularity2	102.09 ± 3.02
Coarseness	106.73 ± 2.20



RD: Simple Descriptors of a Region

- Area = the number of pixels contained within its boundary
- Perimeter = the length of its boundary
- Compactness = $perimeter^2/A$
- Circularity = A_R/A_C, two objects with the same perimeter
- Elongatedness = $A_R/(2d)^2$
  - d = erosion steps before the region totally disappears
- Eccentricity (of the ellipse that has the same second moments as the region) = $(\mu_{20} - \mu_{02} + 4\mu_{11})^2/\mu_{00}$
- Principal axes ratio = major axis / minor axis.

RD: Simple Descriptors of a Region

- Eccentricity: ratio of major and minor axes
- Elongatedness: ratio between length and width of bounding rectangle

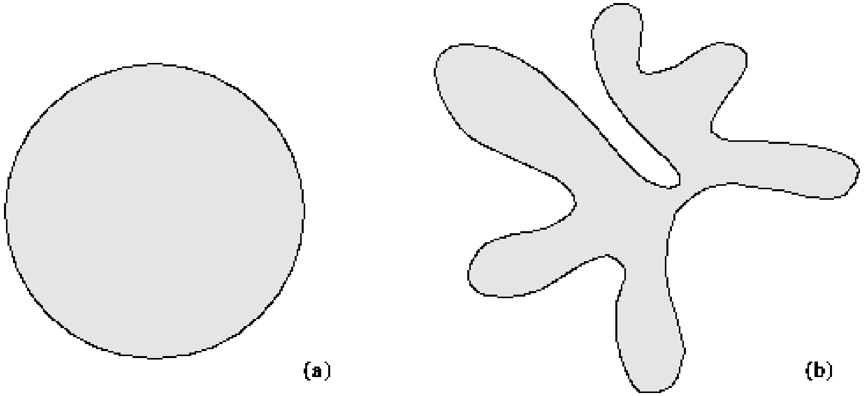


Figure 6.25 *Compactness: (a) Compact, (b) non-compact.*

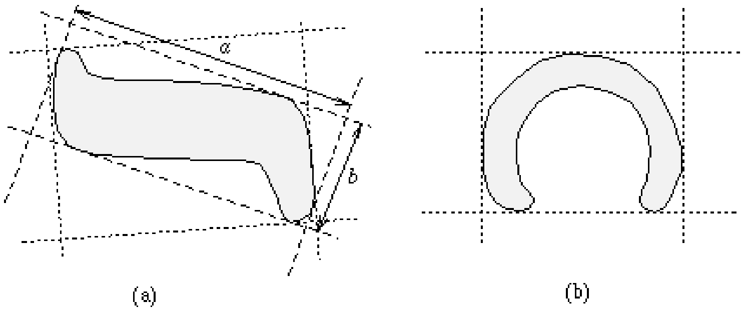
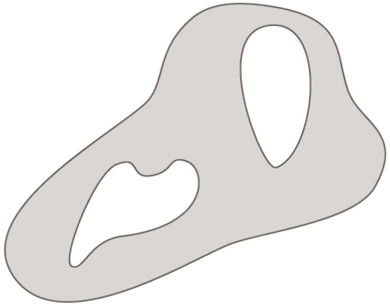
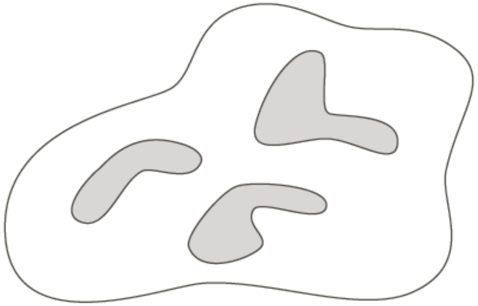


Figure 6.24 *Elongatedness: (a) Bounding rectangle gives acceptable results, (b) bounding rectangle cannot represent elongatedness.*

RD: Topological Descriptors

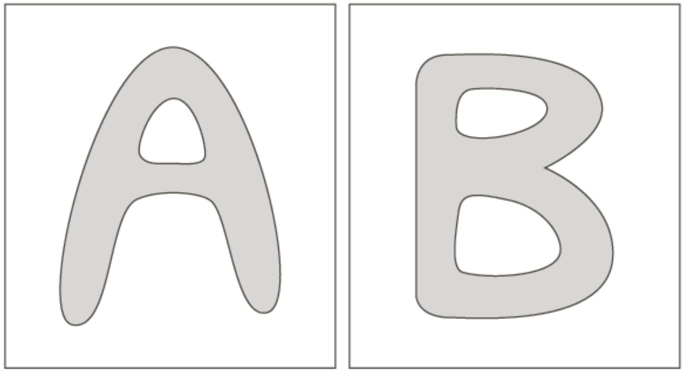
- Topology is the study of properties a figure that are unaffected by any deformation, as long as there is no tearing or joining of the figure (also called rubber-sheet distortions)
- Euler number $E = C - H$.
  - Number of holes in a region H.
  - Number of connected components C.
- 2 holes and 3 connected components

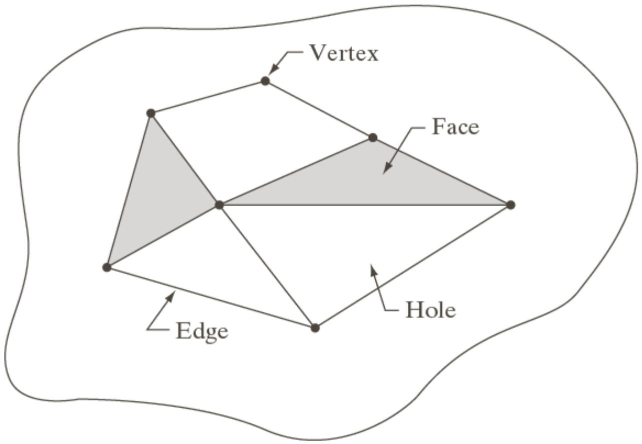




RD: Topological Descriptors

- Euler number 0 and -1





- In more general case (for 3D objects) the vertices *V*, edges *Q* and faces *F*, (and bodies,) are also included

$$V - Q + F = C - H = E$$

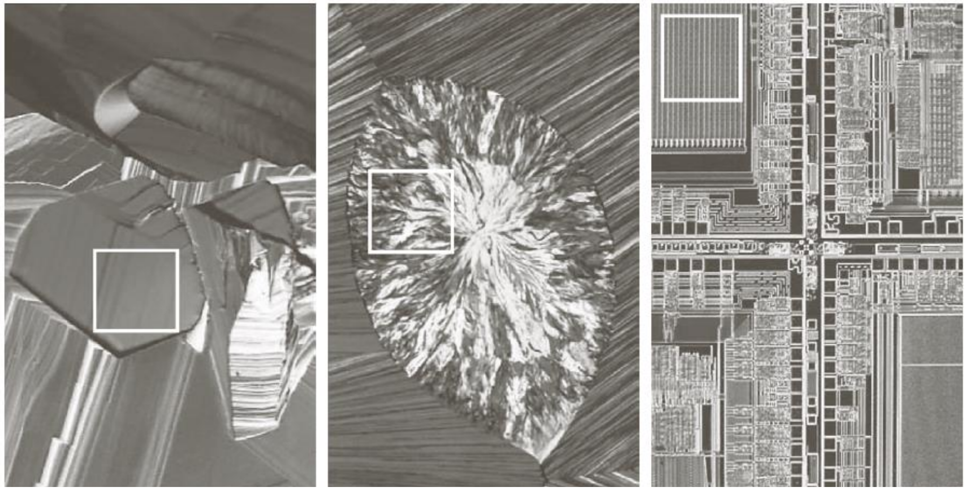
- E.g. object above: 7-11+2=1-3=-2

RD: Texture

- Texture content of a region
- Texture descriptor provides measures of properties such as smoothness, coarseness, and regularity
- Three approaches:
  - Statistical approach: moments, co-occurrence matrix
  - Structural approach: texture primitives
    - Rules similar to productions in parsing, e.g *S->aS*
    - For detecting similar structures in the texture
  - Spectral approach: Fourier spectrum

RD: Texture

- Statistical measures for texture
- E.g. smoothness $R = 1 - 1/(\sigma^2(z))$



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

RD: Texture, Co-occurrence matrix

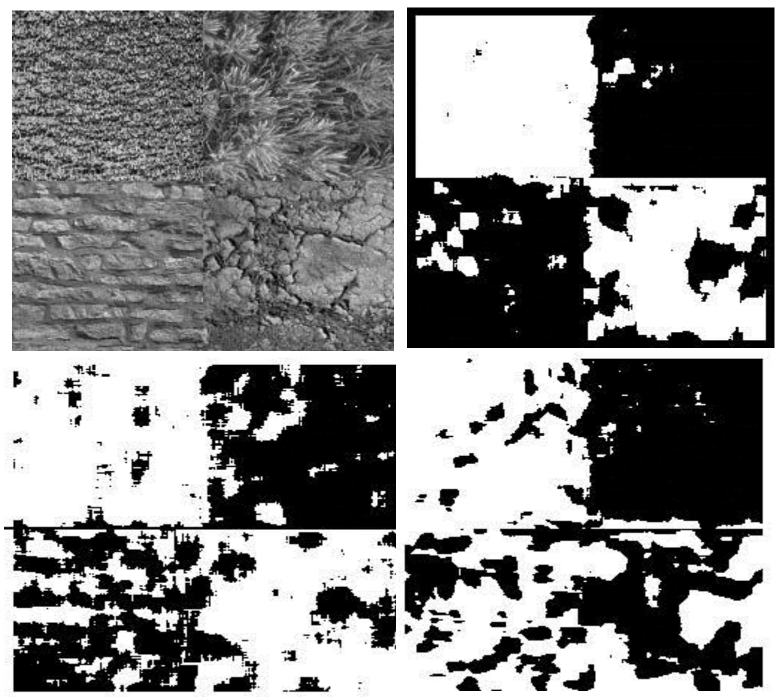
- Original image (upper left): skin, leaves, brick, stone
- Segmentation based on: Contrast (upper right), Homogeneity (lower left), Standard deviation (lower right)

1	1	7	5	3	2
5	1	6	1	2	5
8	8	6	8	1	2
4	3	4	5	5	1
8	7	8	7	6	2
7	8	6	2	6	2

Image f

	1	2	3	4	5	6	7	8
1	1	2	0	0	0	1	1	0
2	0	0	0	0	1	1	0	0
3	0	1	0	1	0	0	0	0
4	0	0	1	0	1	0	0	0
5	2	0	1	0	1	0	0	0
6	1	3	0	0	0	0	0	1
7	0	0	0	0	1	1	0	2
8	1	0	0	0	0	2	2	1

Co-occurrence matrix **G**



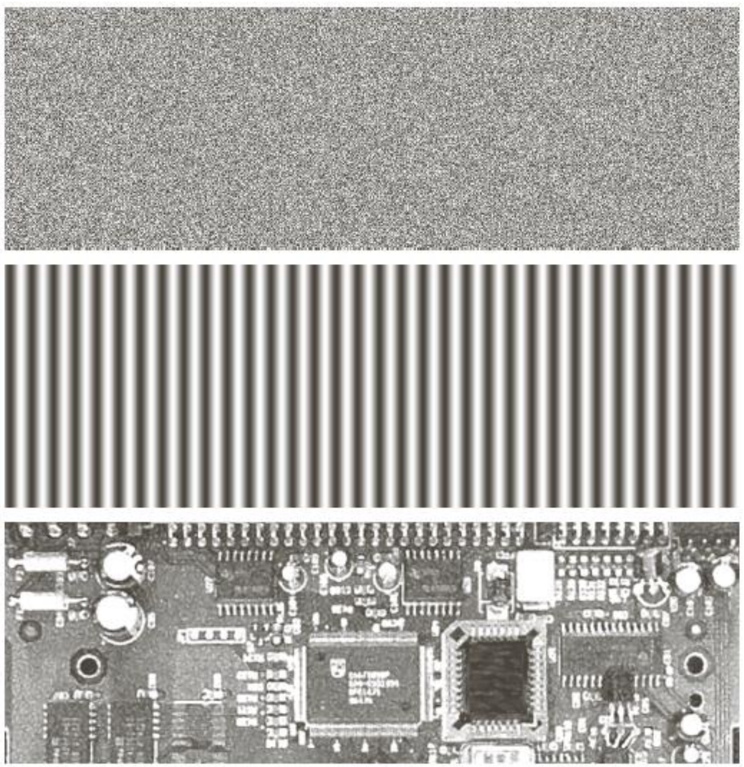
RD: Texture, Co-occurrence matrix

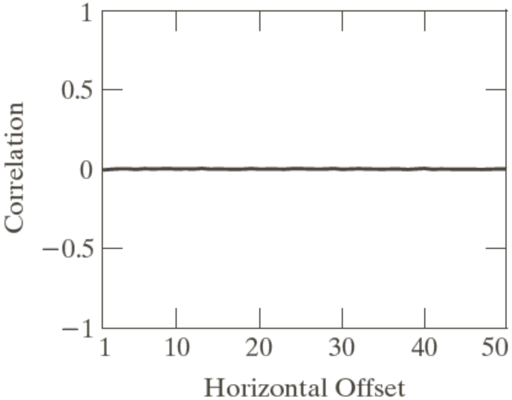
- Some properties found from the co-occurrence matrix

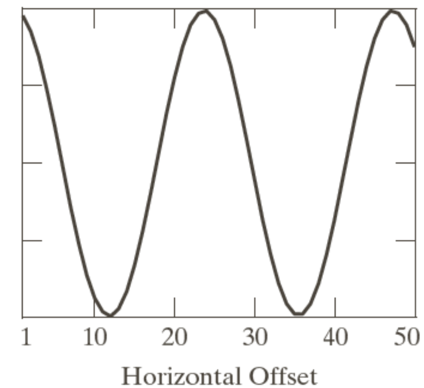
Descriptor	Explanation	Formula
Maximum probability	Measures the strongest response of **G**. The range of values is [0, 1].	$\max_{i,j}(p_{ij})$
Correlation	A measure of how correlated a pixel is to its neighbor over the entire image. Range of values is 1 to −1, corresponding to perfect positive and perfect negative correlations. This measure is not defined if either standard deviation is zero.	$\sum_{i=1}^{K}\sum_{j=1}^{K}\frac{(i-m_r)(j-m_c)p_{ij}}{\sigma_r\sigma_c}$ $\sigma_r \neq 0;\ \sigma_c \neq 0$
Contrast	A measure of intensity contrast between a pixel and its neighbor over the entire image. The range of values is 0 (when **G** is constant) to $(K-1)^2$.	$\sum_{i=1}^{K}\sum_{j=1}^{K}(i-j)^2 p_{ij}$
Uniformity (also called Energy)	A measure of uniformity in the range [0, 1]. Uniformity is 1 for a constant image.	$\sum_{i=1}^{K}\sum_{j=1}^{K}p_{ij}^2$
Homogeneity	Measures the spatial closeness of the distribution of elements in **G** to the diagonal. The range of values is [0, 1], with the maximum being achieved when **G** is a diagonal matrix.	$\sum_{i=1}^{K}\sum_{i=1}^{K}\frac{p_{ij}}{1+\lvert i-j\rvert}$
Entropy	Measures the randomness of the elements of **G**. The entropy is 0 when all p_{ij}'s are 0 and is maximum when all p_{ij}'s are equal. The maximum value is $2\log_2 K$. (See Eq. (11.3-9) regarding entropy).	$-\sum_{i=1}^{K}\sum_{i=1}^{K}p_{ij}\log_2 p_{ij}$

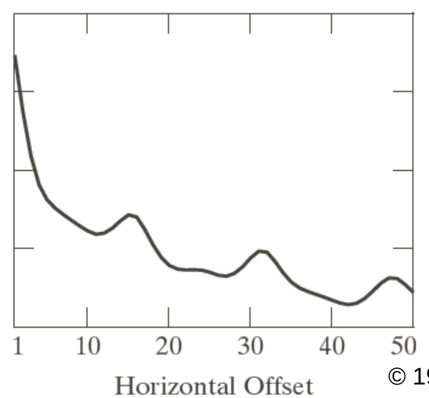
RD: Texture, Co-occurrence matrix

- Effect of the horizontal offset vs. correlation



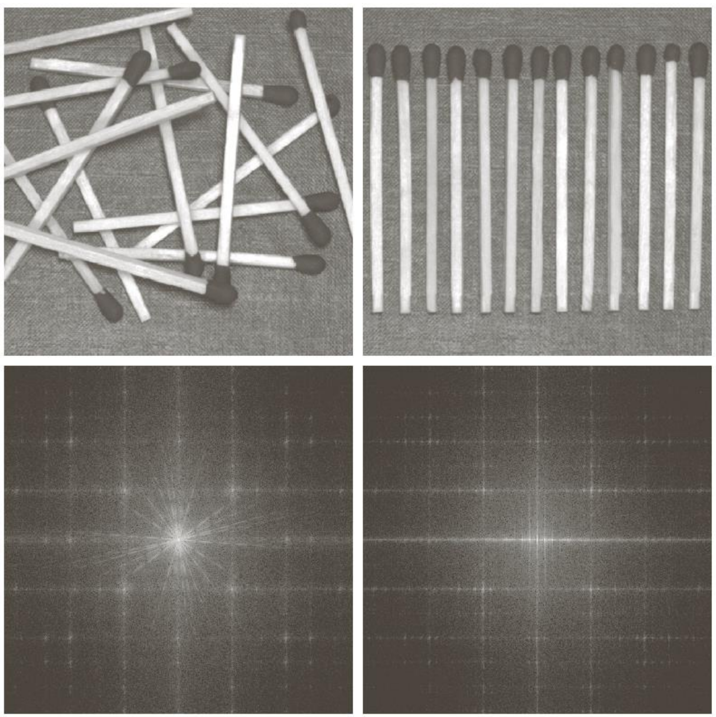


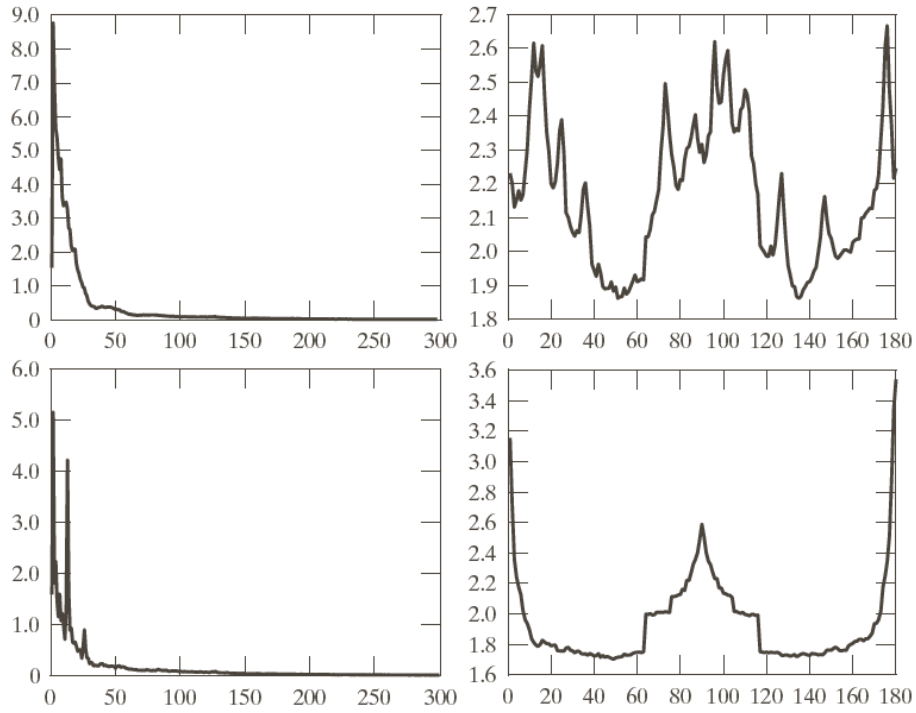




RD: Texture, Fourier signatures

- Fourier signatures $S(R)$ $S(\theta)$
- Original images with DFTs





RD: Moments

- Image moment: a certain weighted average of the image pixel intensities (e.g. centroid can be found with moments)
- For image $f(x,y)$, the central moment of order $(p+q)$ is

$$\mu_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (x-\bar{x})^p (y-\bar{y})^q f(x,y)$$

- Then invariant moments can be derived from the second and third moments
  - Translation, scaling, rotation, mirroring

RD: Moments

- Typical invariant moments, $\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\gamma}}$, $\gamma = \frac{p+q}{2} + 1$

$$\phi_1 = \eta_{20} + \eta_{02}$$

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$$

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2$$

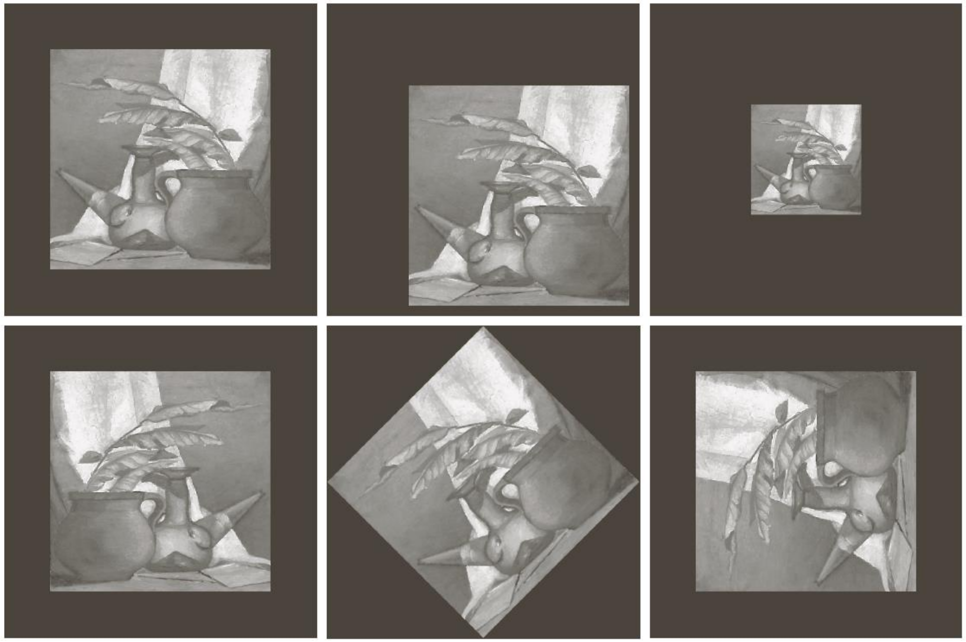
$$\phi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$$

$$\phi_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})$$

$$\phi_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03}) [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$$

RD: Moments

- Examples on moments
- Original image and the processed images

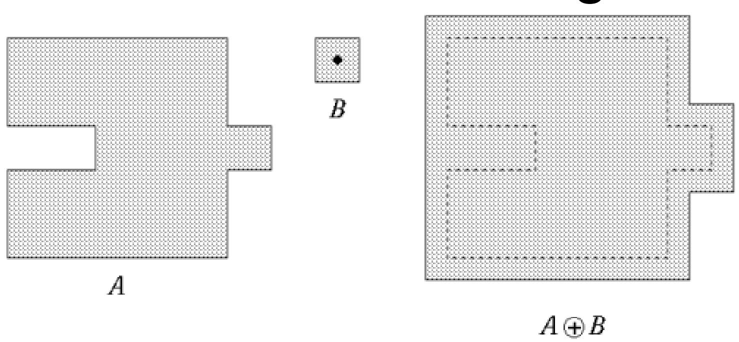


- The corresponding moments

Moment Invariant	Original Image	Translated	Half Size	Mirrored	Rotated 45°	Rotated 90°
ϕ_1	2.8662	2.8662	2.8664	2.8662	2.8661	2.8662
ϕ_2	7.1265	7.1265	7.1257	7.1265	7.1266	7.1265
ϕ_3	10.4109	10.4109	10.4047	10.4109	10.4115	10.4109
ϕ_4	10.3742	10.3742	10.3719	10.3742	10.3742	10.3742
ϕ_5	21.3674	21.3674	21.3924	21.3674	21.3663	21.3674
ϕ_6	13.9417	13.9417	13.9383	13.9417	13.9417	13.9417
ϕ_7	−20.7809	−20.7809	−20.7724	20.7809	−20.7813	−20.7809

Image Morphology

- A tool for extracting image components that are useful in the presentation and description of region such as boundaries, skeletons, the convex hull
- Also useful in pre- and post processing such as morphological filtering, thinning, and pruning
- Dilation: expand the region using an structural element.
- Erosion: shrink the region using an structural element



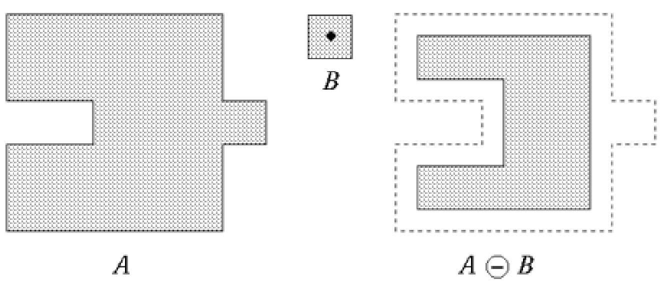


Image Morphology

- Basic definitions

  $A(x) = \{c | c = a + x, for\ a \in A\}$ (translation of set A by point x)

  $\hat{B} = \{x | x = -b, for\ b \in B\}$ (reflection of set B)

- Dilation

$$A + B = \left\{x \middle| \left(\hat{B}\right)_x \cap A \neq \emptyset\right\} = \left\{x | [\left(\hat{B}\right)_x \cap A] \subseteq A\right\}$$

- The dilation of *A* by *B* is the set of all *x* displacements such that $\hat{B}$ and *A* overlap by at least one nonzero element
- Erosion

$$A + B = \left\{x \middle| \left(\hat{B}\right)_x \text{ belongs to the subset of } A\right\}$$

- The erosion of *A* by *B* is the set of all points *x* such that *B*, translated by *x*, is contained in *A*

Image Morphology

- Opening

  $$A \circ B = (A - B) + B$$

- Closing

  $$A \bullet B = (A + B) - B$$

- To cope with distortions (e.g, to fill gaps) and noise (e.g., to remove isolated large noise areas).

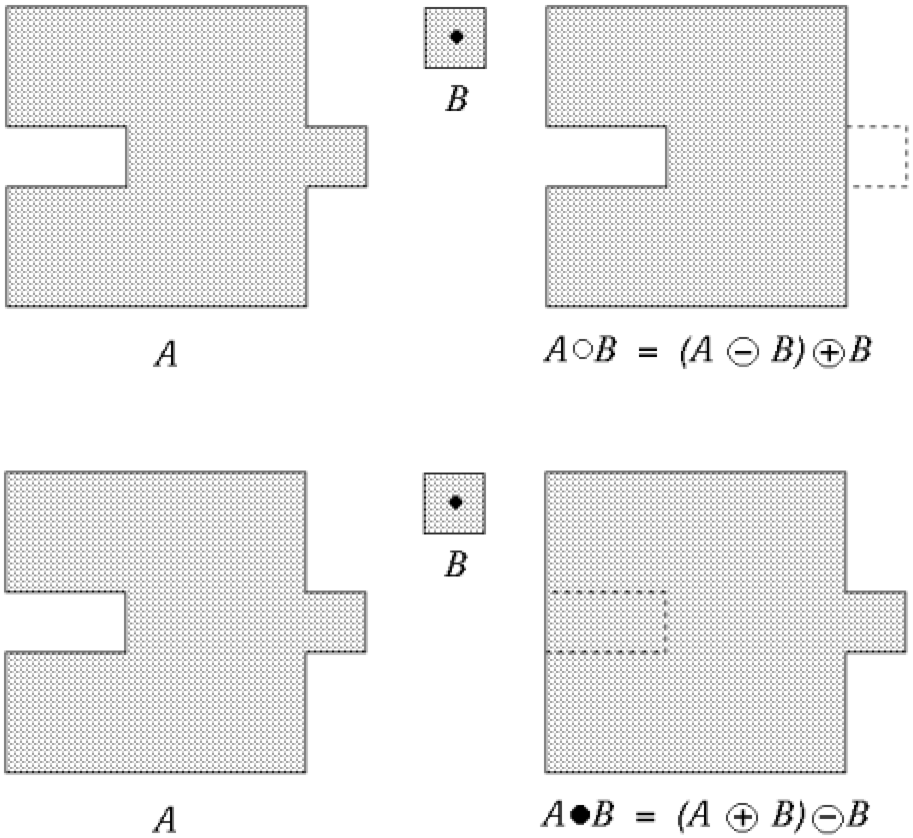
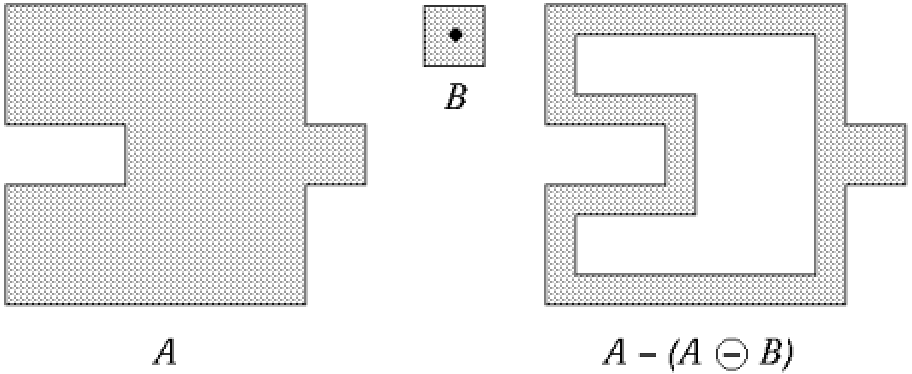


Image Morphology

- Hit or miss (points found in A with B1 and in the complement of A with B2)
- Boundary extraction
- Connected components
- Convex hull
- Thinning
- Thickening
- Skeletons
- Pruning (“sophisticated” thinning and skeletonizing)



Principal Components

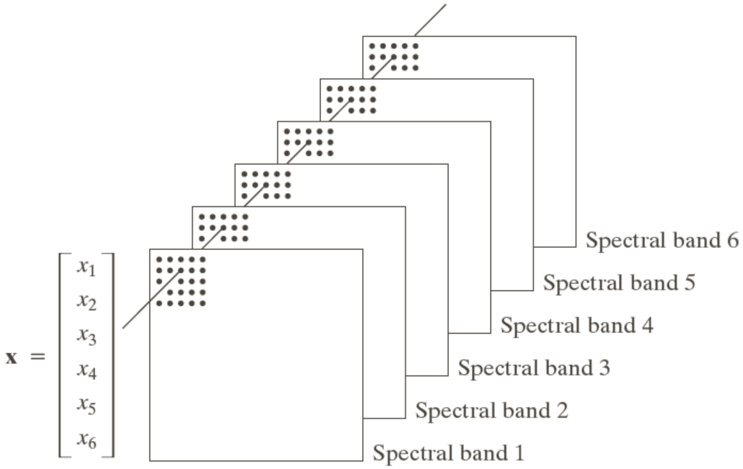
- Principal components of the data are found as eigenvalues and eigenvectors of the covariance matrix of the data values
- The approximation for the covariance matrix is

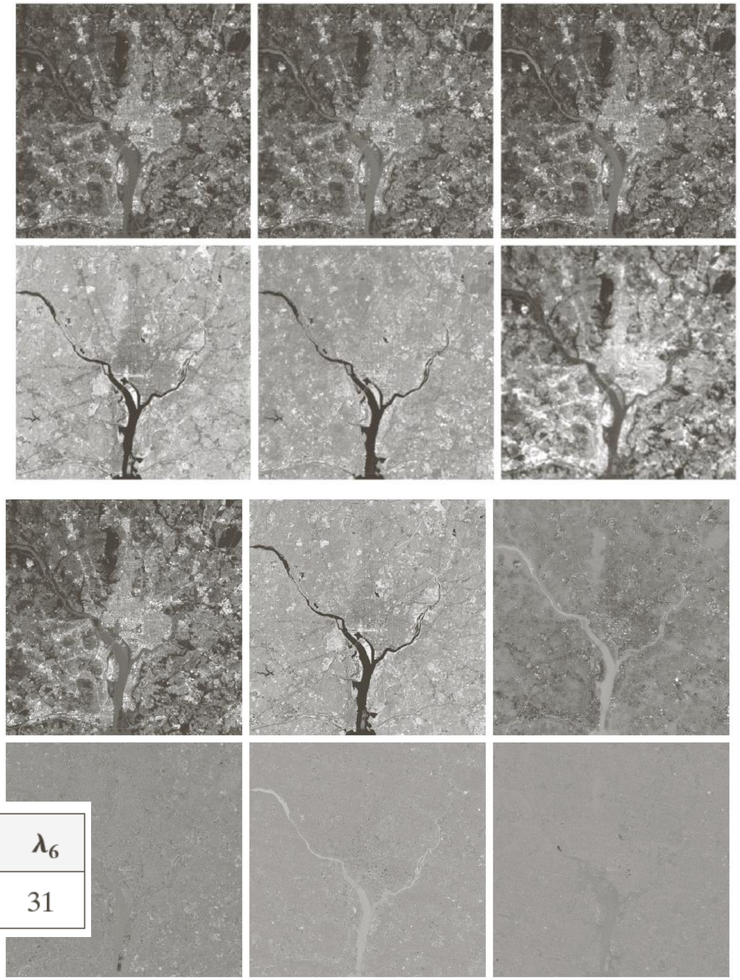
$$C_x = \frac{1}{K}\sum_{k=1}^{K} x_k x_k^T - m_k m_k^T, \quad m_x = \frac{1}{K}\sum_{k=1}^{K} x_k$$

- The principal components are used in
  - Spectral imaging for image compression
  - Color imaging for feature detection
  - etc

Principal Components

- Spectral imaging





λ_1	λ_2	λ_3	λ_4	λ_5	λ_6
10344	2966	1401	203	94	31

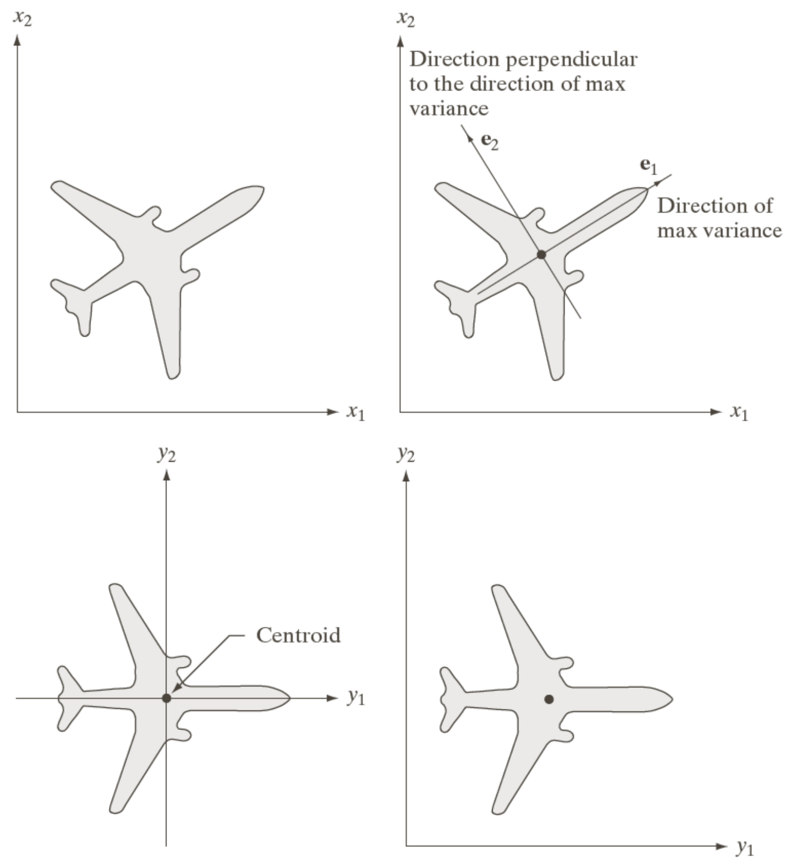
Principal Components

- Gray-scale images
- Points (x_1, x_2) from the region or from the boundary
- If matrix A contains the eigenvectors as rows, then

$$y = A(x - m_x)$$

$$\mathbf{m}_{\mathbf{x}} = \begin{bmatrix} 3 \\ 3 \end{bmatrix} \qquad \mathbf{C}_{\mathbf{x}} = \begin{bmatrix} 3.333 & 2.00 \\ 2.00 & 3.333 \end{bmatrix}$$

$$\mathbf{e}_1 = \begin{bmatrix} 0.707 \\ 0.707 \end{bmatrix} \quad \mathbf{e}_2 = \begin{bmatrix} -0.707 \\ 0.707 \end{bmatrix}$$



Relational descriptors

- To organize components to exploit any structural relationships that may exist between them.
  - Primitive elements.

    $\rightarrow a \quad \downarrow b$
  - Rewriting rules with productions, variables and primitives

    (1) $S \rightarrow aA$ (a followed by A)

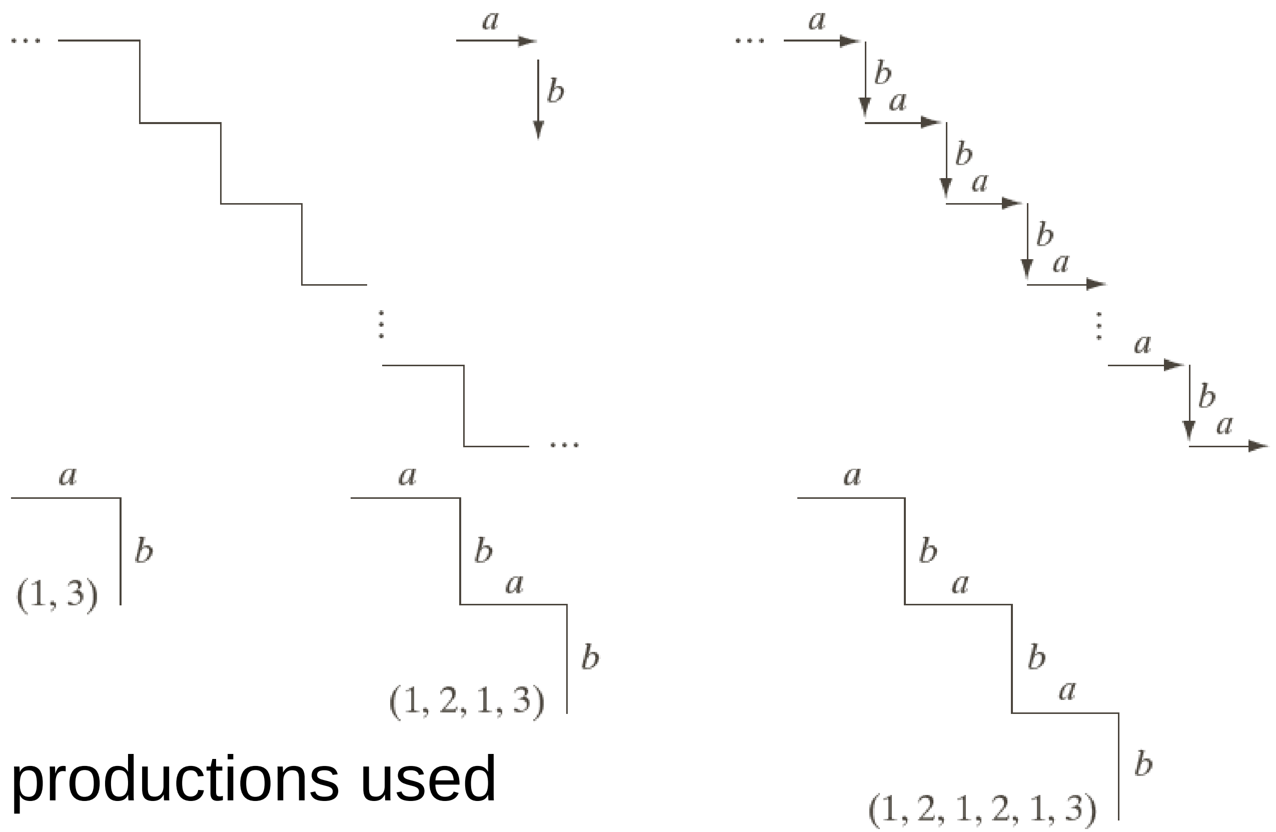
    (2) $A \rightarrow bS$ (b followed by S)

    (3) $A \rightarrow b$ (A derives b)
- In principle the transformation reduces a 2D image to a 1D string

Relational descriptors

- Example of derivations

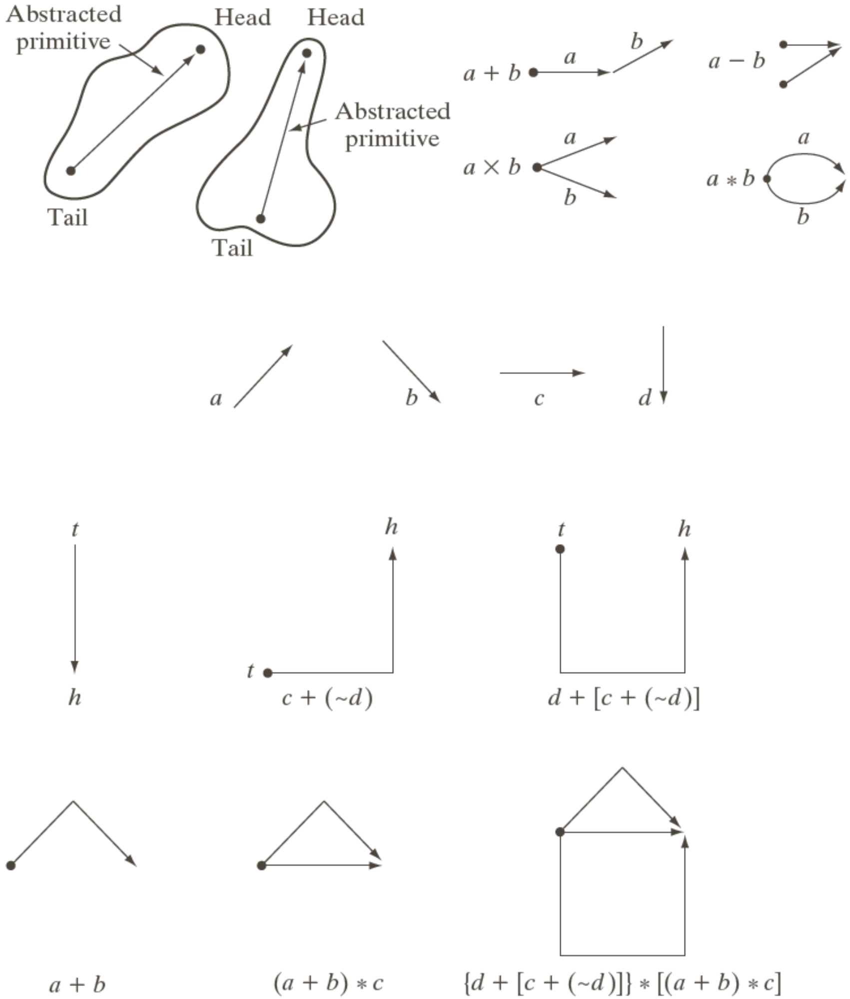
(1) $S \rightarrow aA$,
(2) $A \rightarrow bS$, and
(3) $A \rightarrow b$,



- Output: list of productions used

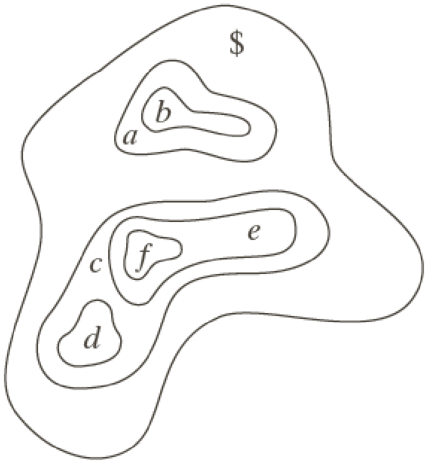
Relational descriptors

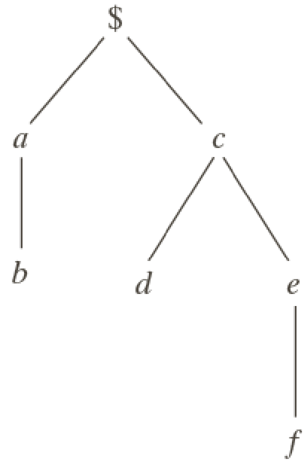
- Head-tail connections for the primitives in the image
- Also operations +,-,*, x defined



Relational descriptors

- A tree-structure is constructed on a specific rule, e.g. the rule ”inside of”
- Information in the tree (root *$*, subtrees *T*)
  - A node contains a description related to that node
  - Information how the nodes are related (connected)





Example applications

- Medical image processing.
  - Diabetes and retinal image analysis using machine vision
  - http://www.it.lut.fi/project/imageret/
- Industrial machine vision.
  - Paper and board printability tests by machine vision in the paper making and printing industry.
  - http://www.it.lut.fi/project/papvision/
- Biometrics.
  - Image-based biometric person authentication.
  - http://www.it.lut.fi/project/facedetect/

Summary

- **Task: 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 (classification, clustering)