Representation and Description

Selim Aksoy
Department of Computer Engineering
Bilkent University
saksoy@cs.bilkent.edu.tr

Representation and description

- Images and segmented regions must be represented and described in a form suitable for further processing.
- The representations and the corresponding descriptions are selected according to the computational and semantic requirements of the image analysis task.
- We will study:
  - Region representations and descriptors.
  - Image representations and descriptors.

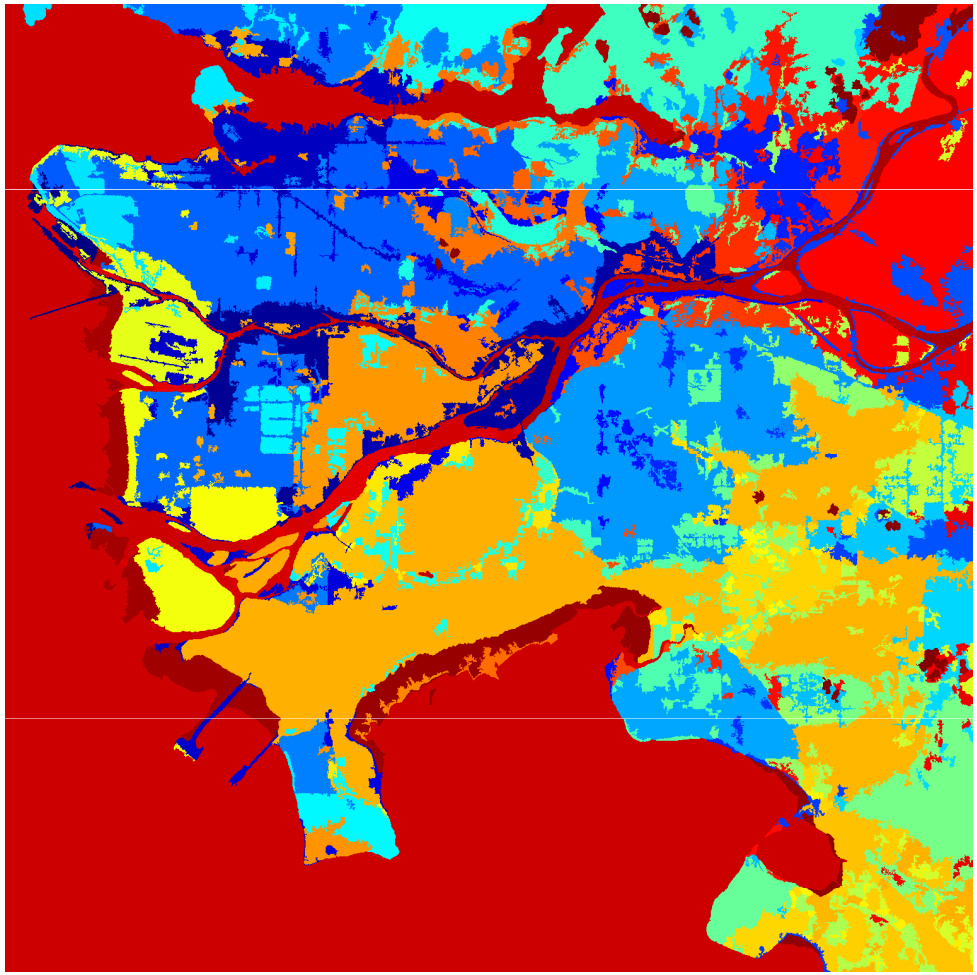
Region representations

- Representing a region involves two choices:
  - External characteristics (boundary).
  - Internal characteristics (pixels comprising the region).
- An external representation is chosen when the primary focus is on shape characteristics.
- An internal representation is selected when the primary focus is on regional properties, such as color and texture.

Labeled images and overlays

- Labeled images are good intermediate representations for regions.
- The idea is to assign each detected region a unique identifier (an integer) and create an image where all pixels of a region will have its unique identifier as their pixel value.
- A labeled image can be used as a kind of mask to identify pixels of a region.
- Region boundaries can be computed from the labeled image and can be overlaid on top of the original image.

Labeled images and overlays

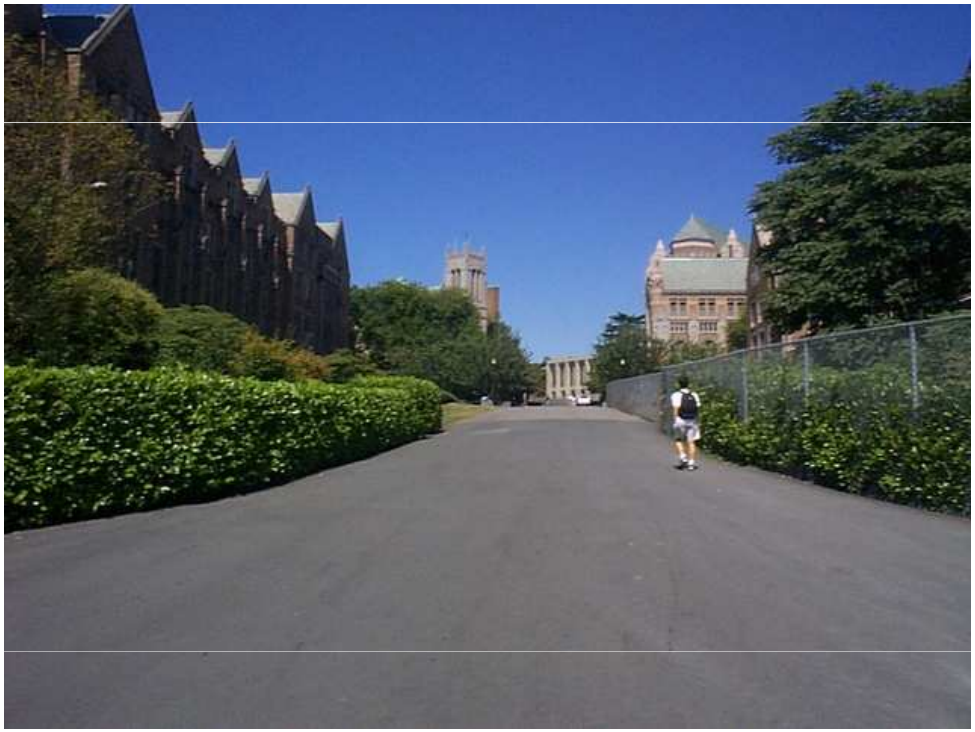


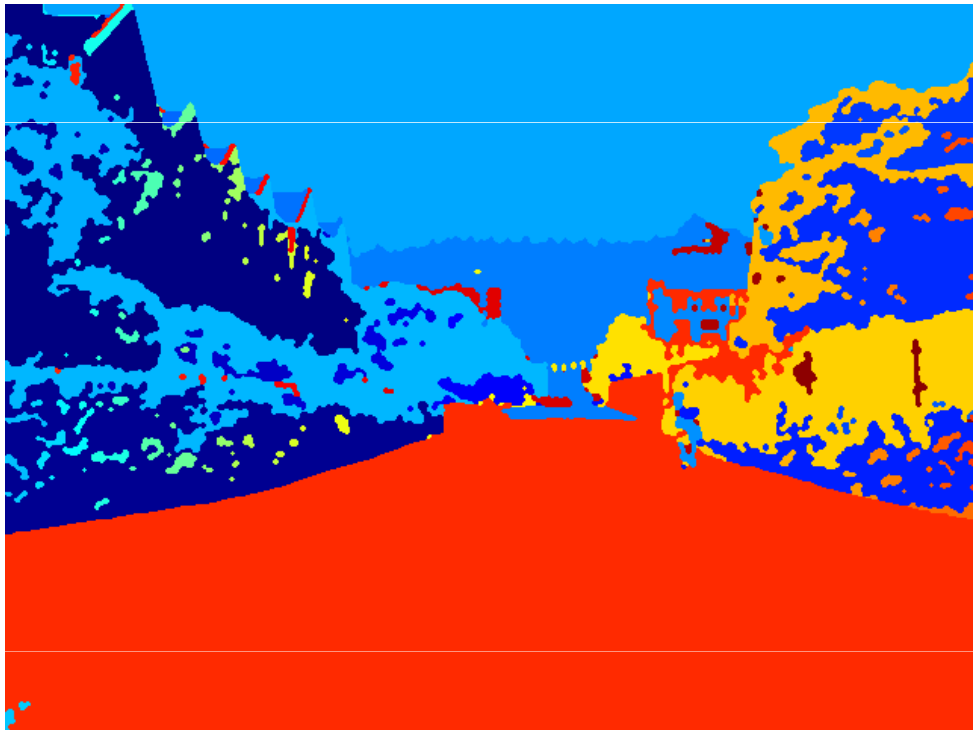
A satellite image and the corresponding labeled image after segmentation (displayed in pseudo-color).

Labeled images and overlays

A satellite image and the corresponding segmentation overlay.

Labeled images and overlays



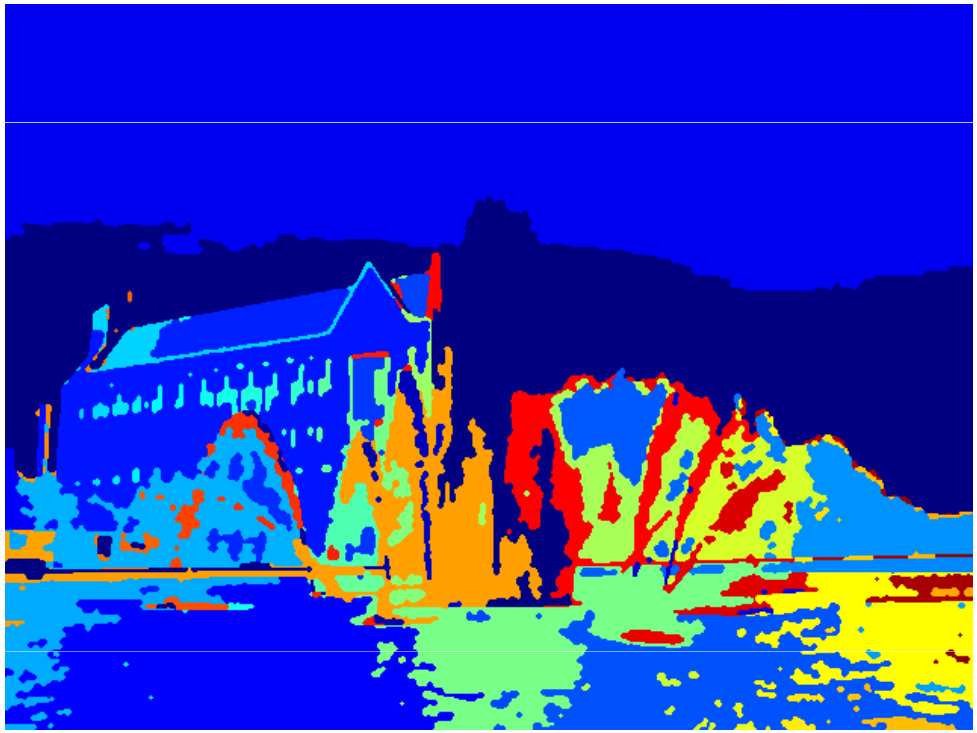


An image and the corresponding labeled image after segmentation (displayed in pseudo-color).

Labeled images and overlays

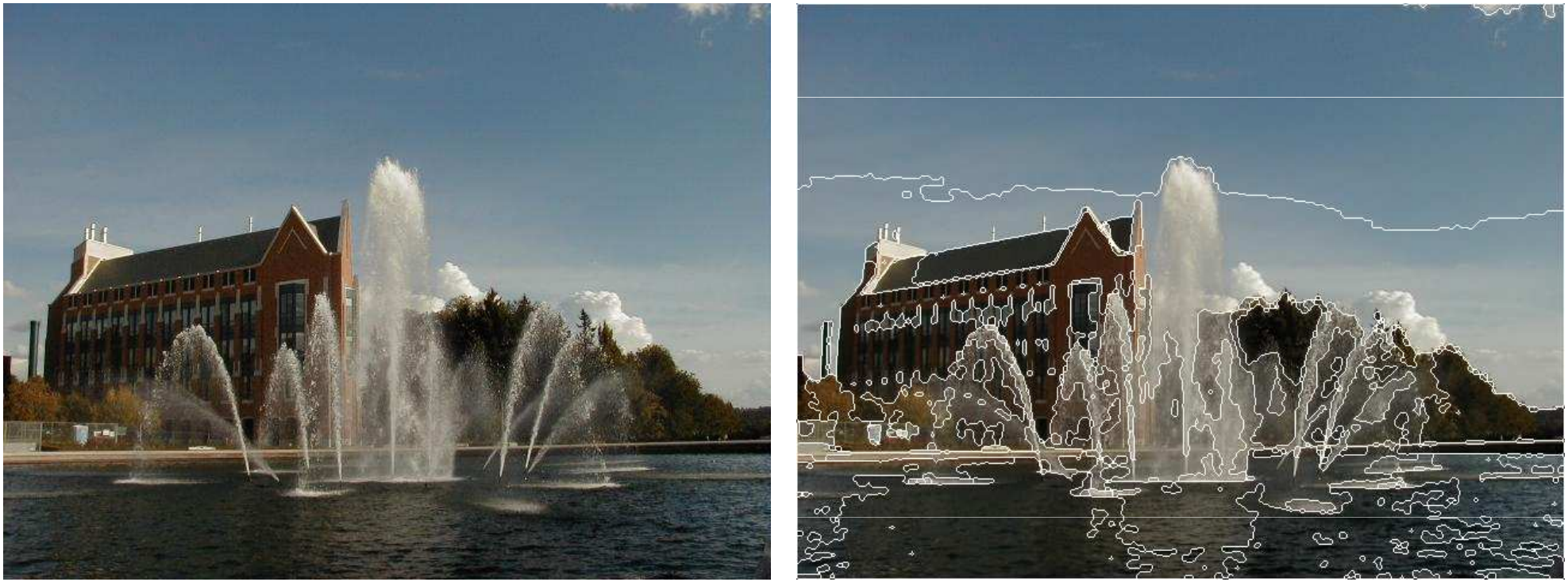
An image and the corresponding segmentation overlay.

Labeled images and overlays



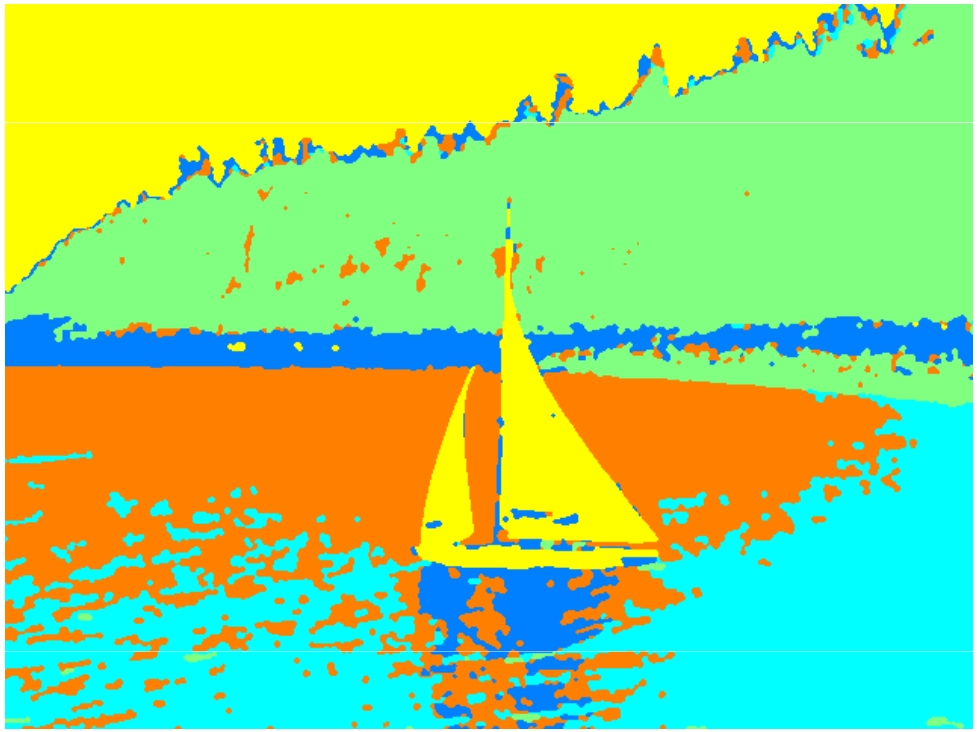
An image and the corresponding labeled image after segmentation (displayed in pseudo-color).

Labeled images and overlays



An image and the corresponding segmentation overlay.

Labeled images and overlays



An image and the corresponding labeled image after segmentation (displayed in pseudo-color).

Labeled images and overlays

An image and the corresponding segmentation overlay.

Chain codes

- Regions can be represented by their boundaries in a data structure instead of an image.
- The simplest form is just a linear list of the boundary points of each region.
- This method generally is unacceptable because:
  - The resulting list tends to be quite long.
  - Any small disturbances along the boundary cause changes in the list that may not be related to the shape of the boundary.
- A variation of the list of points is the chain code, which encodes the information from the list of points at any desired quantization.

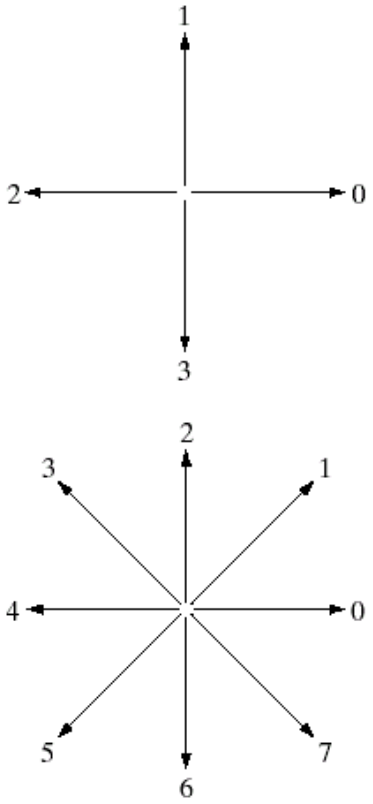
Chain codes

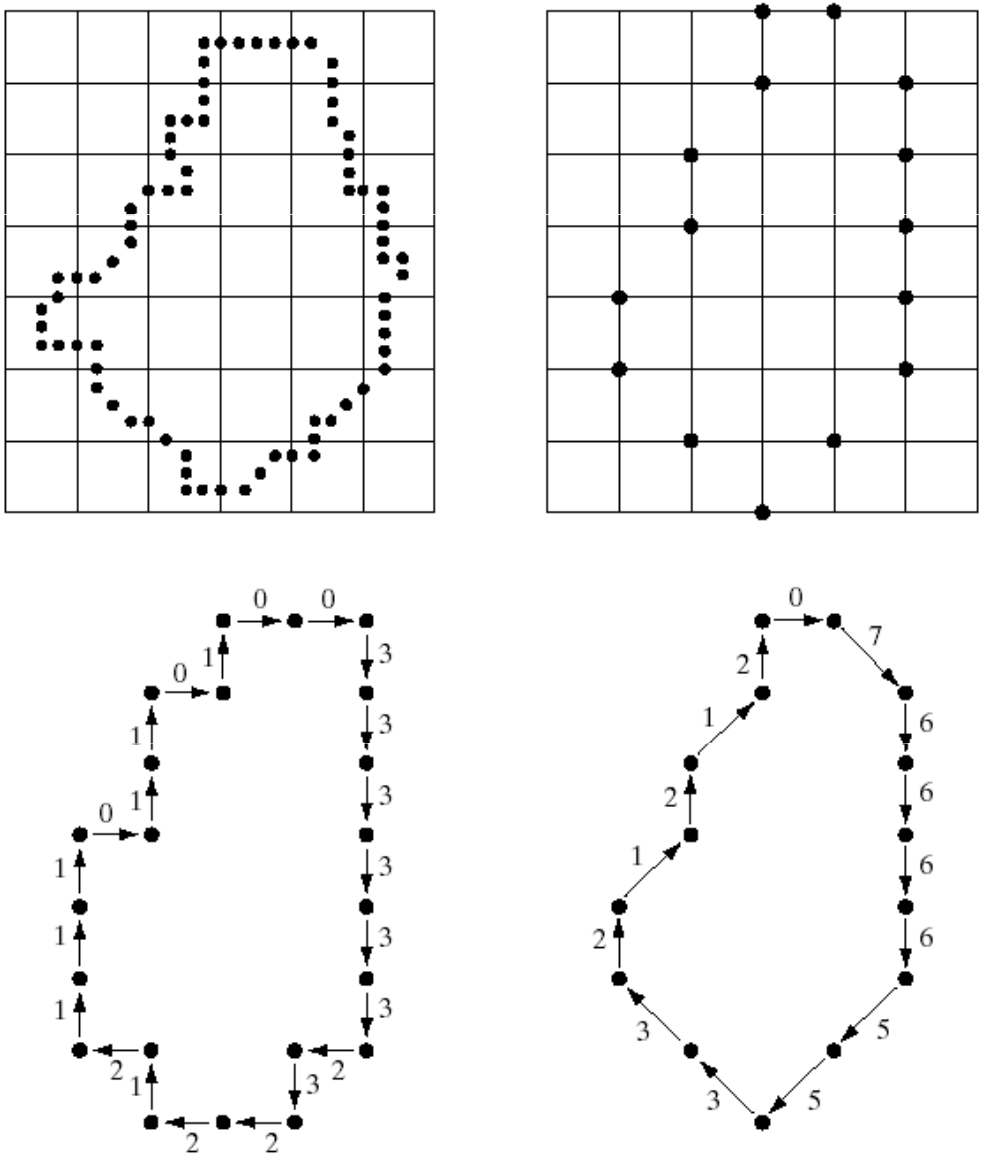
- Conceptually, a boundary to be encoded is overlaid on a square grid whose side length determines the resolution of the encoding.
- Starting at the beginning of the curve, the grid intersection points that come closest to it are used to define small line segments that join each grid point to one of its neighbors.
- The directions of these line segments are then encoded as small integers from zero to the number of neighbors used in encoding.

Chain codes

a b

FIGURE 11.1 Direction numbers for (a) 4-directional chain code, and (b) 8-directional chain code.





a b
c d

FIGURE 11.2 (a) Digital boundary with resampling grid superimposed. (b) Result of resampling. (c) 4-directional chain code. (d) 8-directional chain code.

Chain codes

- The chain code of a boundary depends on the starting point.
- However, the code can be normalized by treating it as a circular sequence.
- Size (scale) normalization can be achieved by altering the size of the sampling grid.
- Rotation normalization can be achieved by using the first difference of the chain code instead of the code itself.
  - Count the number of direction changes (in counter-clockwise direction) that separate two adjacent elements of the code (e.g., 10103322 → 3133030).

Chain codes

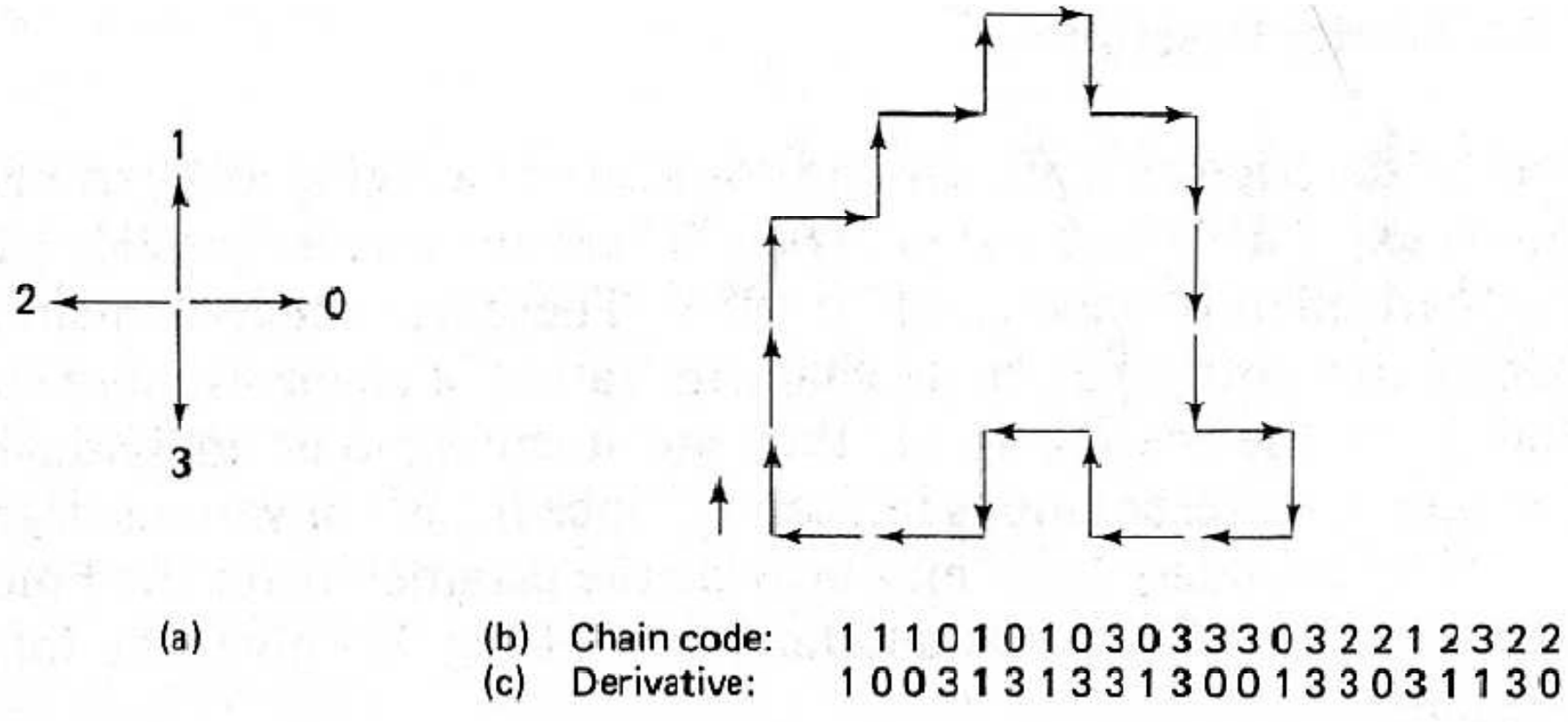


Fig. 8.5 (a) Direction numbers for chain code elements. (b) Chain code for the boundary shown. (c) Derivative of (b).

Polygonal approximations

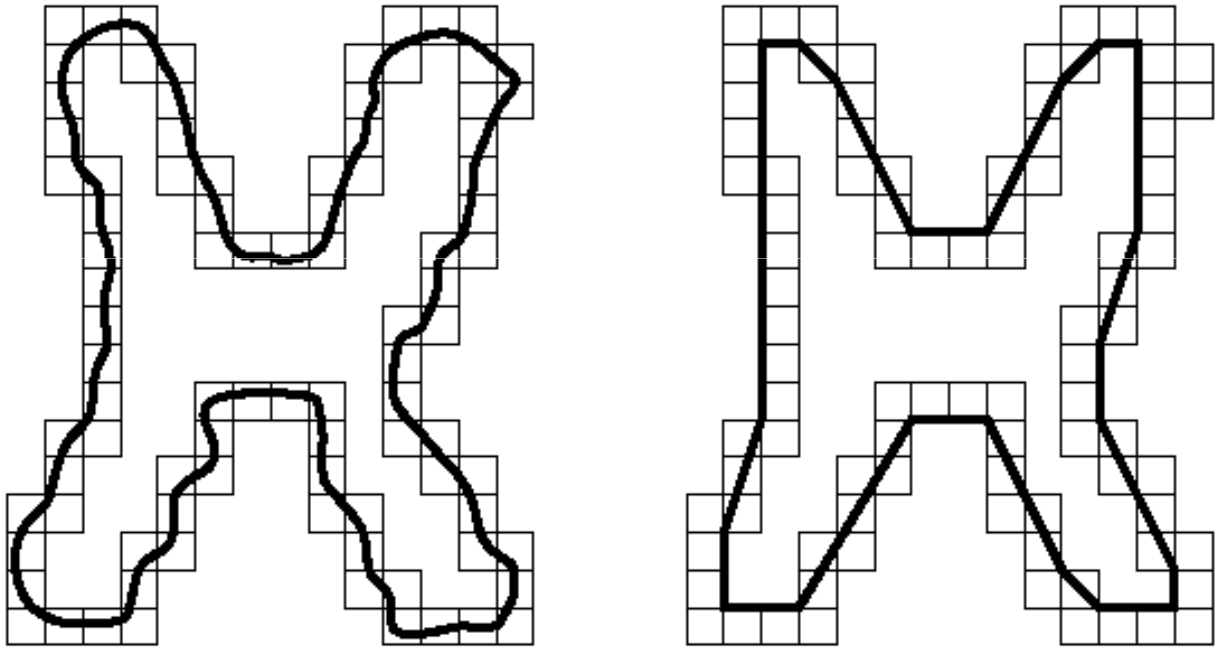
- When the boundary does not have to be exact, the boundary pixels can be approximated by straight line segments, forming a polygonal approximation to the boundary.
- The goal of polygonal approximations is to capture the "essence" of the boundary shape with the fewest possible polygonal segments.
- This representation can save space and simplify algorithms that process the boundary.

Polygonal approximations

- Minimum perimeter polygons:
  1. Enclose the boundary by a set of concatenated cells.
  2. Produce a polygon of minimum perimeter that fits the geometry established by the cell strip.

a b

FIGURE 11.3
(a) Object boundary enclosed by cells. (b) Minimum perimeter polygon.



Polygonal approximations

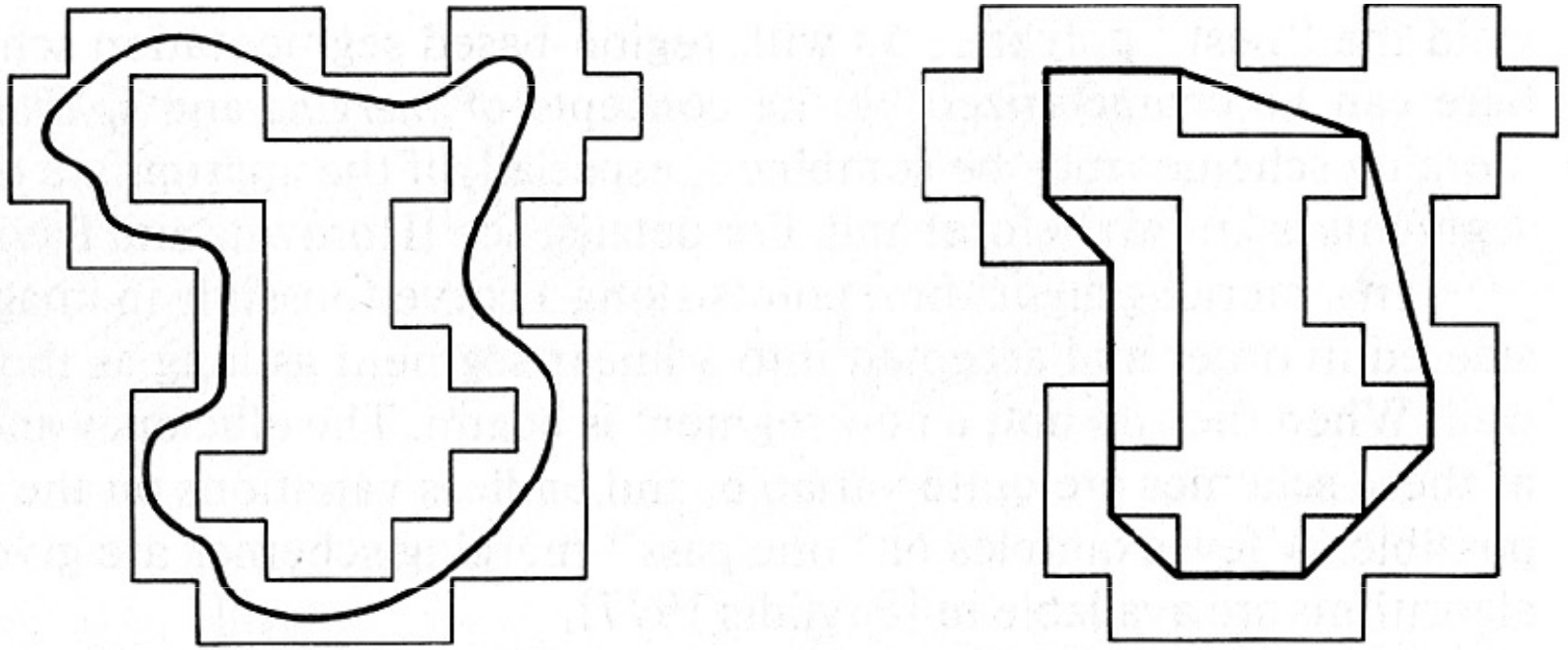
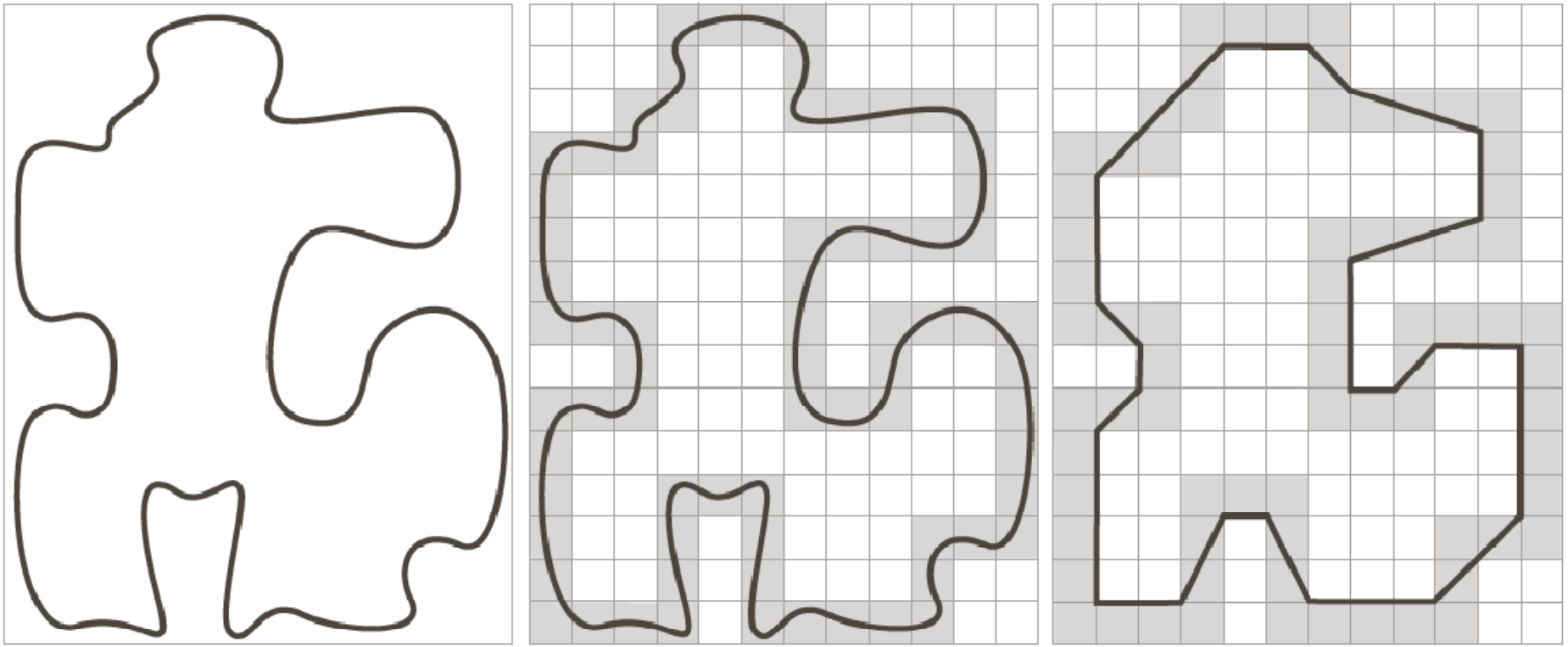


Fig. 8.3 Minimum length polyline.

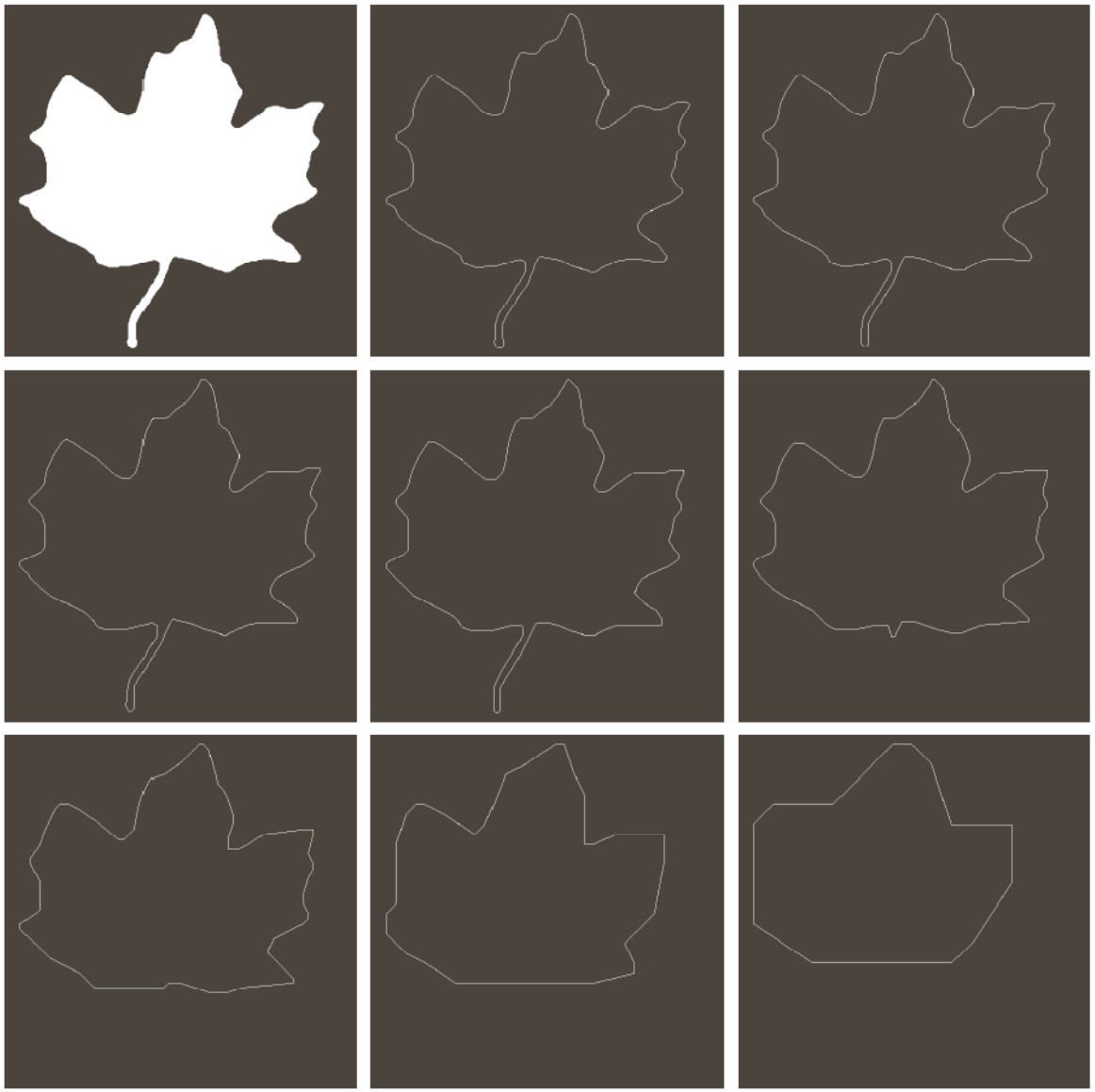
Polygonal approximations



a b c

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.

Polygonal approximations



a b c
d e f
g h i

FIGURE 11.8
(a) 566 × 566 binary image. (b) 8-connected boundary. (c) through (i), MMPs obtained using square cells of sizes 2, 3, 4, 6, 8, 16, and 32, respectively (the vertices were joined by straight lines for display). The number of boundary points in (b) is 1900. The numbers of vertices in (c) through (i) are 206, 160, 127, 92, 66, 32, and 13, respectively.

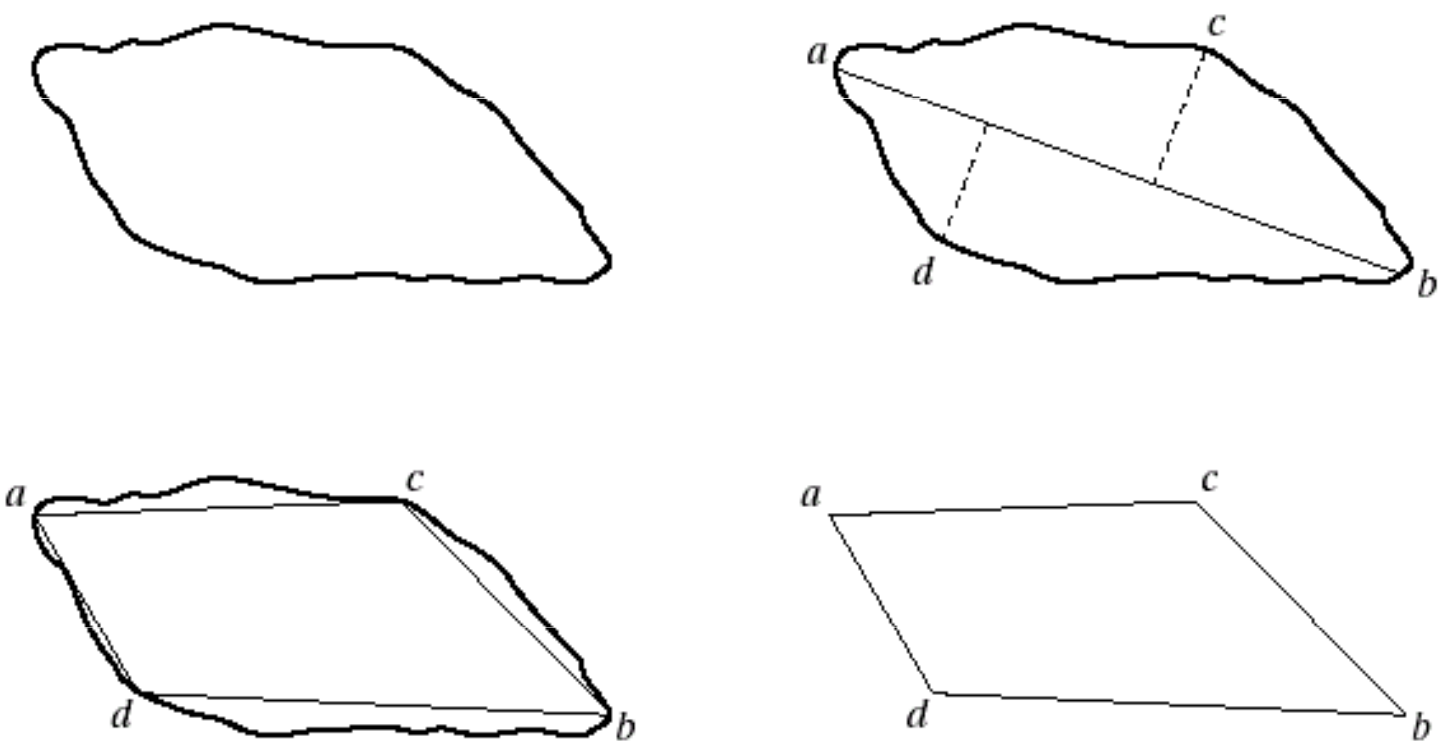
Polygonal approximations

- Merging techniques:
  - Points along a boundary can be merged until the least square error line fit of the points merged so far exceeds a preset threshold.
  - At the end of the procedure, the intersections of adjacent line segments form the vertices of the polygon.

Polygonal approximations

- Splitting techniques:
  - One approach is to subdivide a segment successively into two parts until a specified criterion is satisfied.
  - For instance, a requirement might be that the maximum perpendicular distance from a boundary segment to the line joining its two end points not exceed a preset threshold.
  - If it does, the farthest point from the line becomes a vertex, thus subdividing the initial segment into two.
  - The procedure terminates when no point in the new boundary segments has a perpendicular distances that exceeds the threshold.

Polygonal approximations



a b
c d

FIGURE 11.4
(a) Original boundary. (b) Boundary divided into segments based on extreme points. (c) Joining of vertices. (d) Resulting polygon.

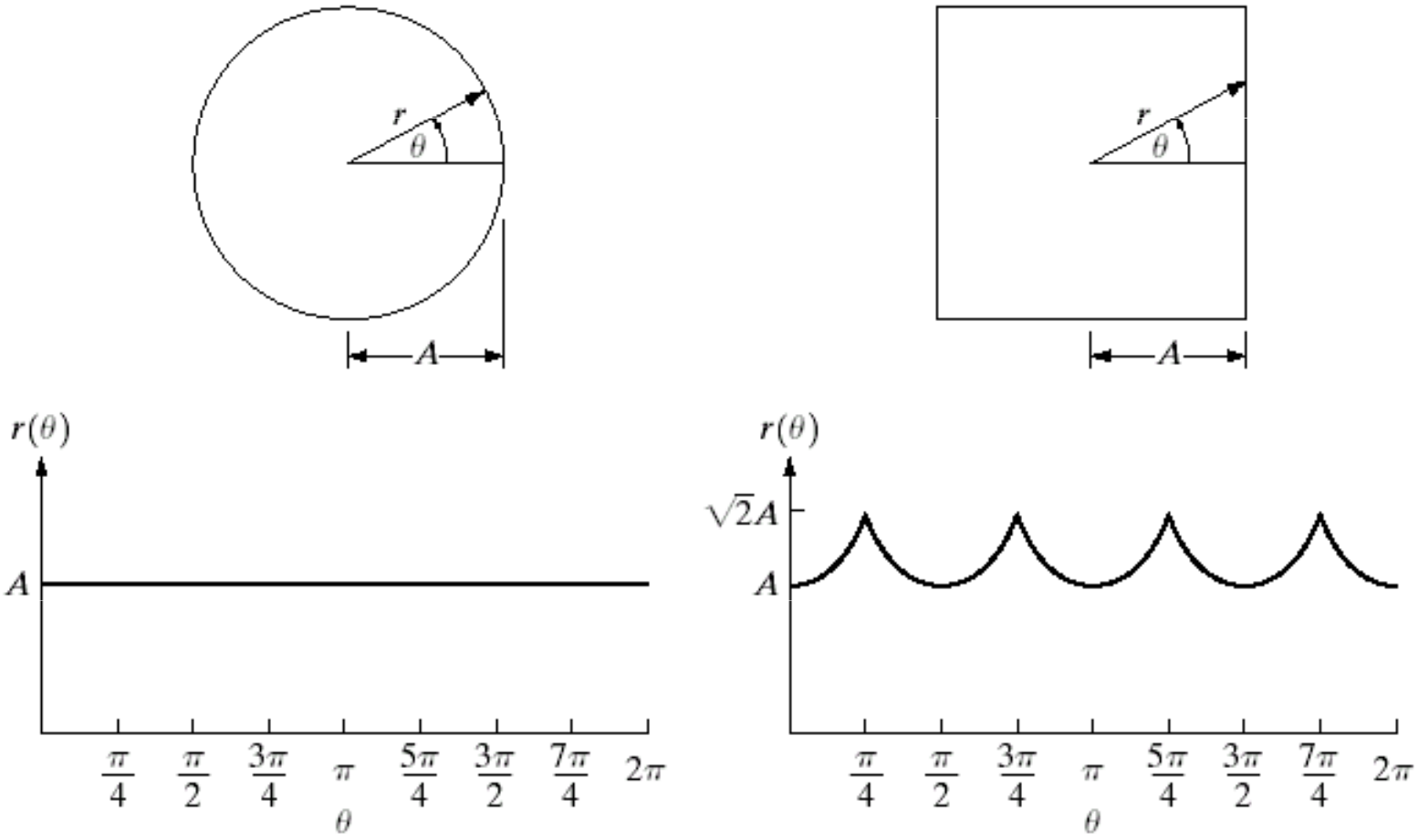
Signatures

- A signature is a 1-D functional representation of a boundary.
- A simple method is to plot the distance from the centroid to the boundary as a function of angle.
  - Normalization with respect to rotation and scaling are needed.
- Another method is to traverse the boundary and, corresponding to each point on the boundary, plot the angle between the line tangent to the boundary at that point and a reference line.
- A variation is to use the histogram of tangent angle values.

Signatures

a b

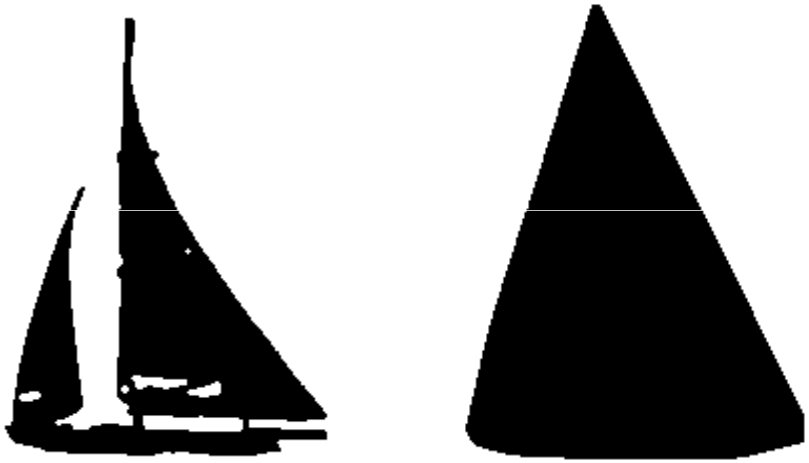
FIGURE 11.5 Distance-versus-angle signatures. In (a) $r(\theta)$ is constant. In (b), the signature consists of repetitions of the pattern $r(\theta) = A\sec\theta$ for $0 \le \theta \le \pi/4$ and $r(\theta) = A\csc\theta$ for $\pi/4 < \theta \le \pi/2$.



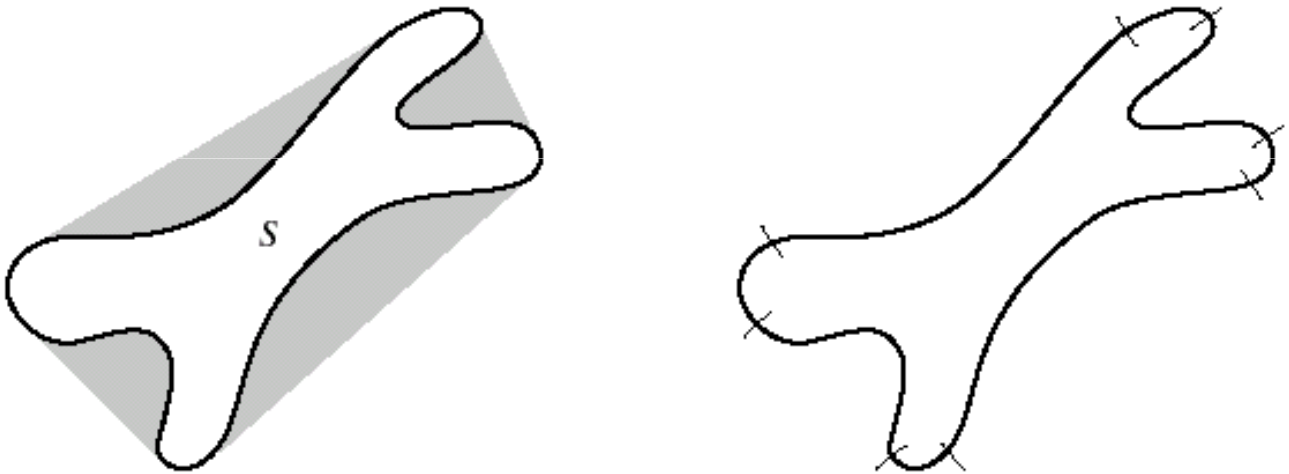
Boundary segments

- Decomposing a boundary into segments can reduce the boundary's complexity and simplify the description process.
- The convex hull H of an arbitrary set S is the smallest convex set containing S.
- The set difference H – S is called the convex deficiency of the set.
- The region boundary can be partitioned by following the contour of S and marking the points at which a transition is made into and out of a component of the convex deficiency.

Boundary segments



A region and its convex hull.



a b

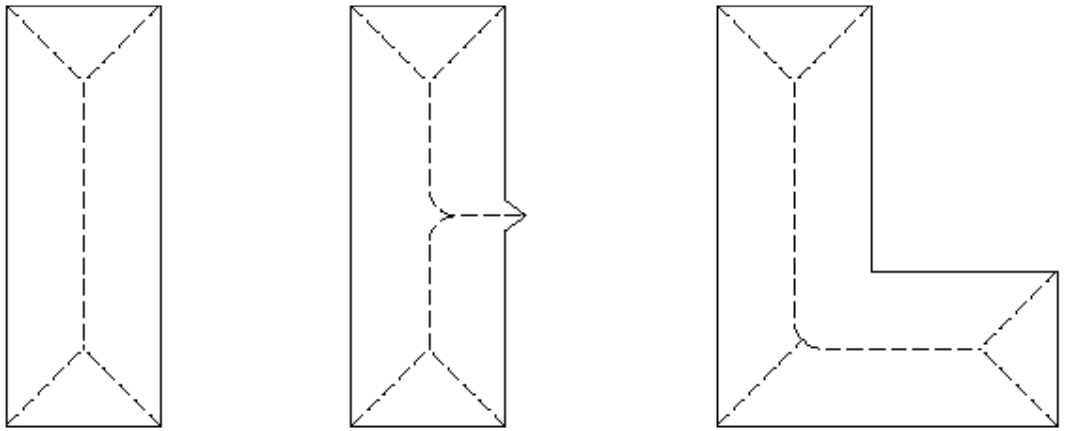
FIGURE 11.6
(a) A region, S, and its convex deficiency (shaded). (b) Partitioned boundary.

Skeletons

- An important approach to representing the structural shape of a plane region is to reduce it to a graph.
- This reduction may be accomplished by obtaining the skeleton of the region.
- The medial axis transformation (MAT) can be used to compute the skeleton.
- The MAT of a region R with border B is as follows:
  - For each point p in R, we find its closest neighbor in B.
  - If p has more than one such neighbor, it is said to belong to the medial axis of R.

Skeletons

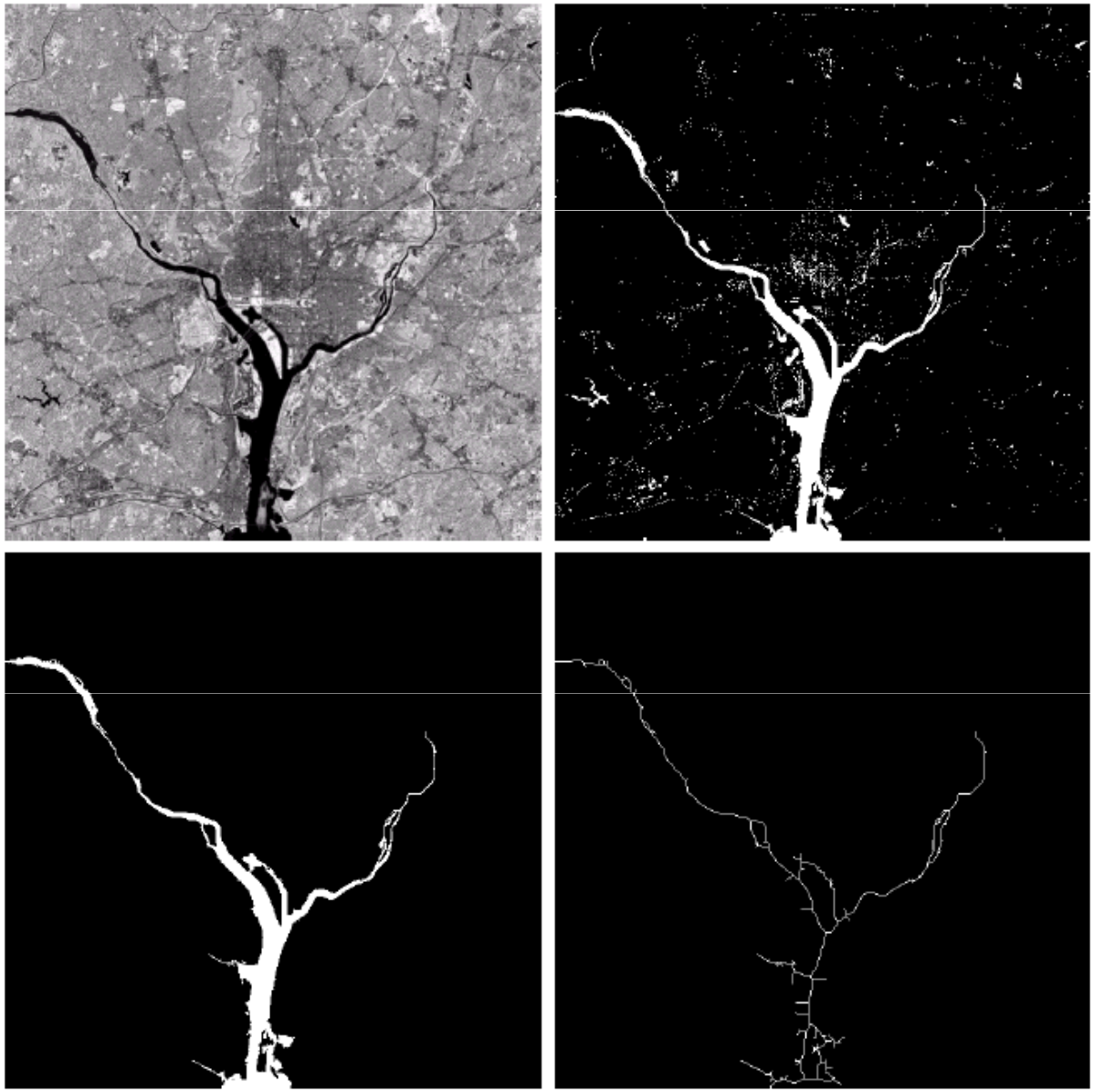
- The MAT can be computed using thinning algorithms that iteratively delete edge points of a region subject to the constraints that deletion of these points
  - does not remove end points,
  - does not break connectivity, and
  - does not cause excessive erosion of the region.



a b c

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

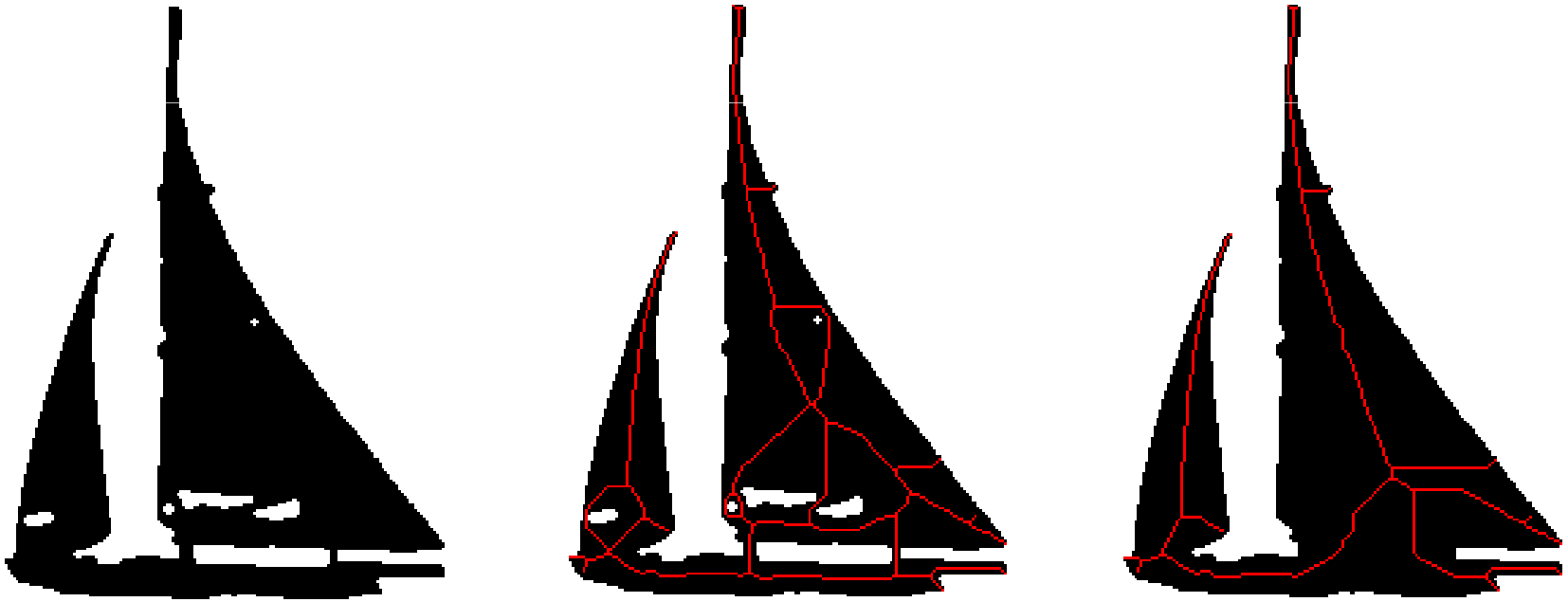
Skeletons



a b
c d

FIGURE 11.21
(a) Infrared image of the Washington, D.C. area. (b) Thresholded image. (c) The largest connected component of (b). Skeleton of (c).

Skeletons



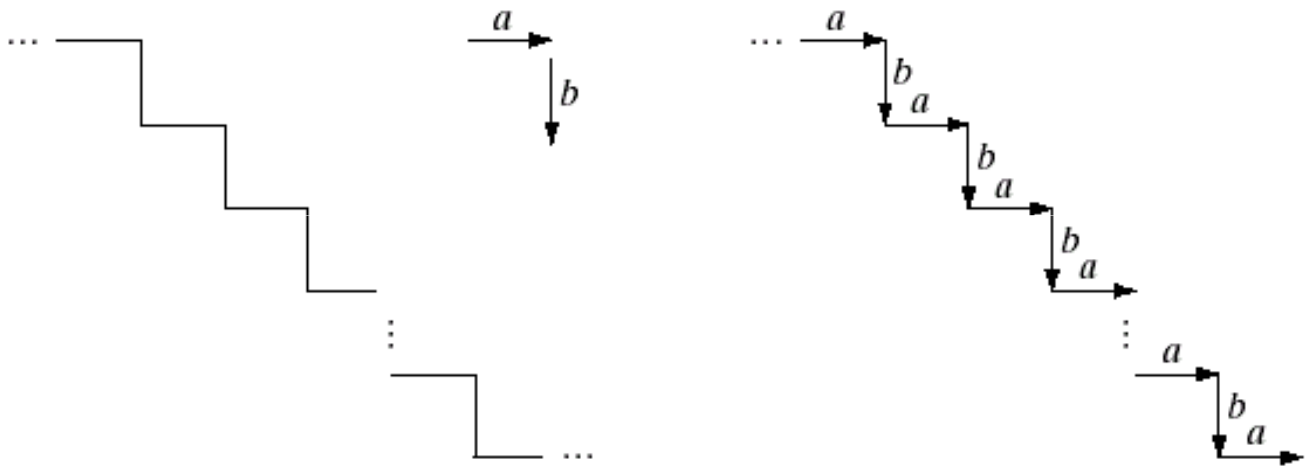
A region, its skeleton, and the skeleton after filling the holes in the region.

Relational representations

- Main purpose is to describe region boundaries using primitive elements.
- Then, string-based representations of region boundaries can be obtained.

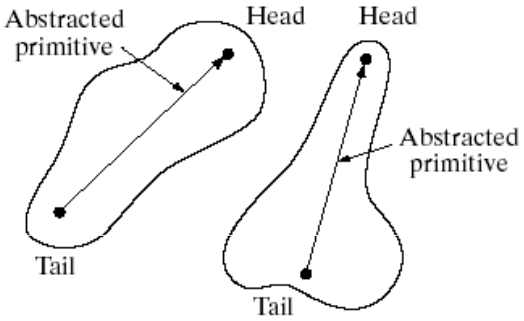
a b

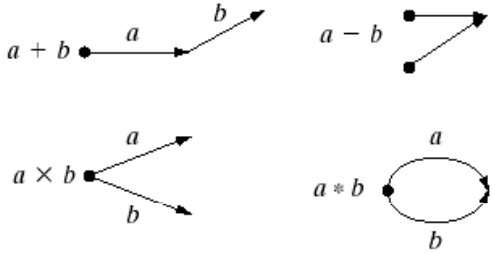
FIGURE 11.30
(a) A simple staircase structure. (b) Coded structure.



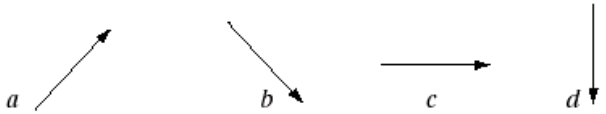
Relational representations

Abstracted primitives



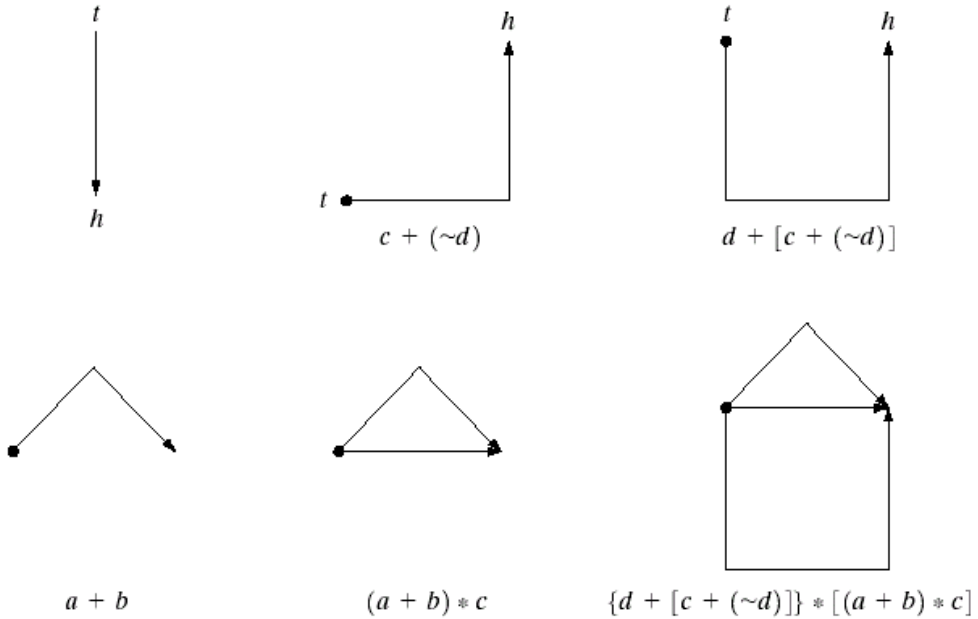


Operations among primitives



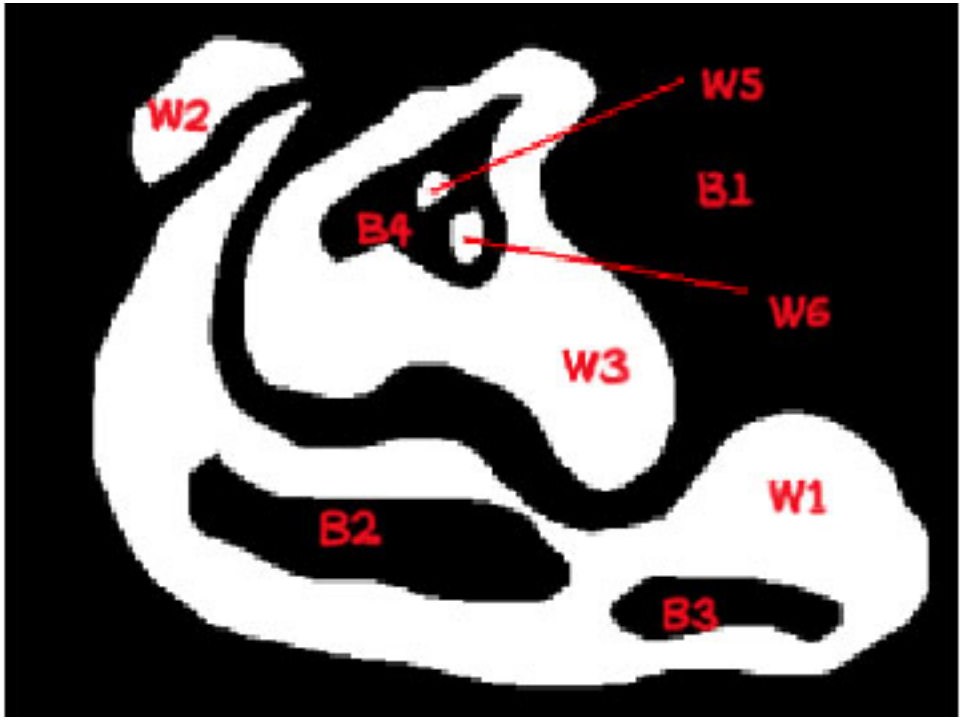
A set of specific primitives

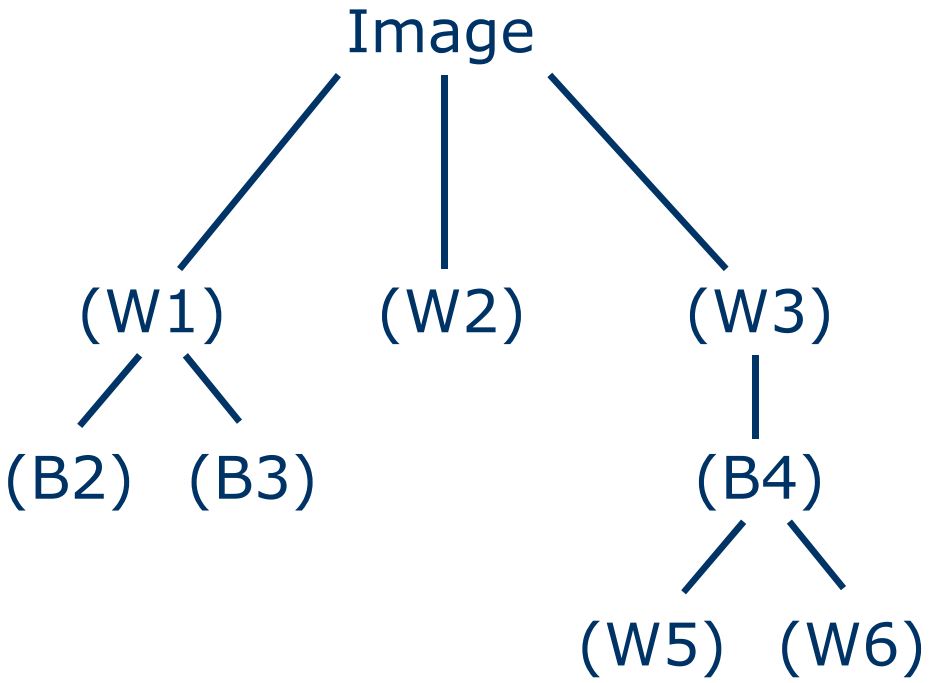
Steps in building a structure



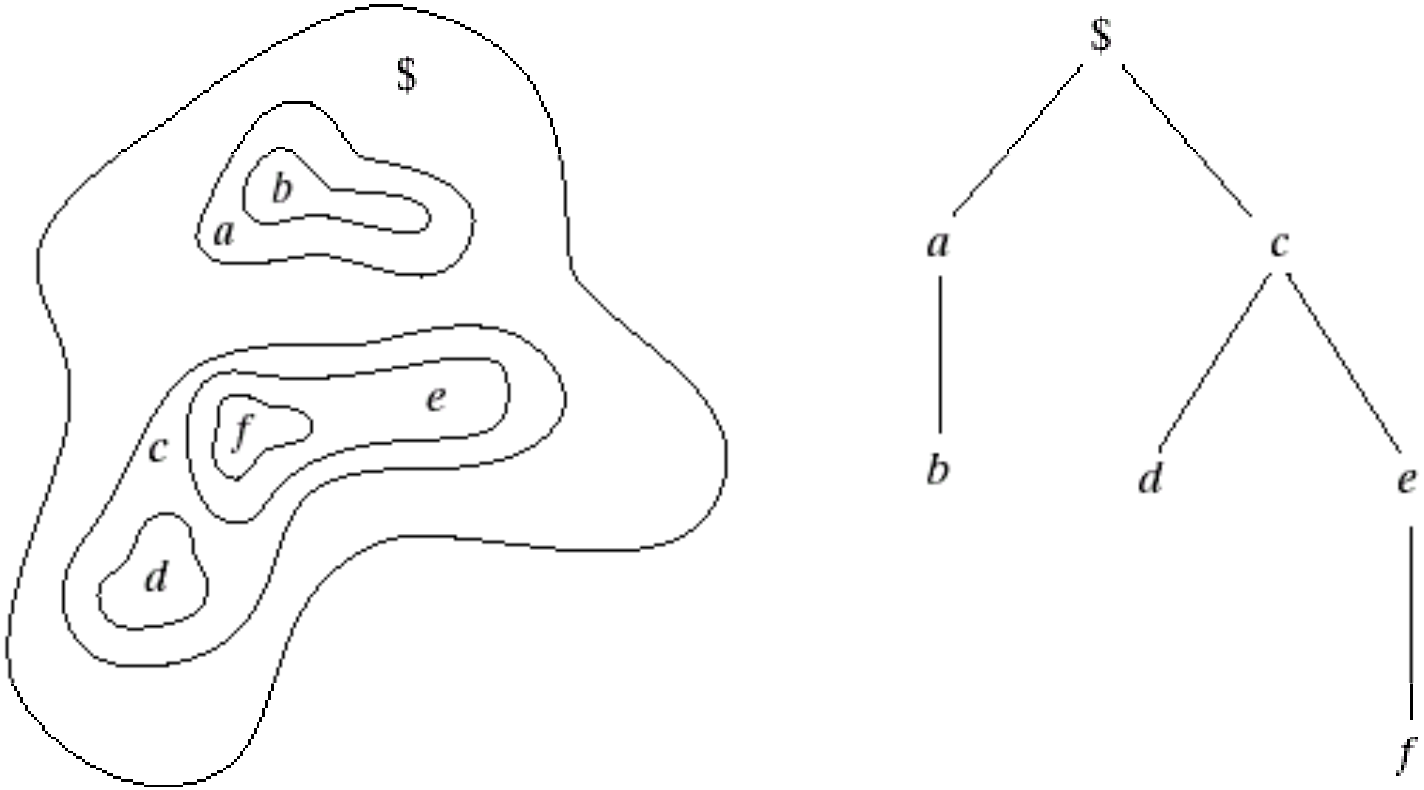
Composite regions

- Some regions can have holes and also multiple parts that are not contiguous.
- A tree structure can be used to represent such regions.





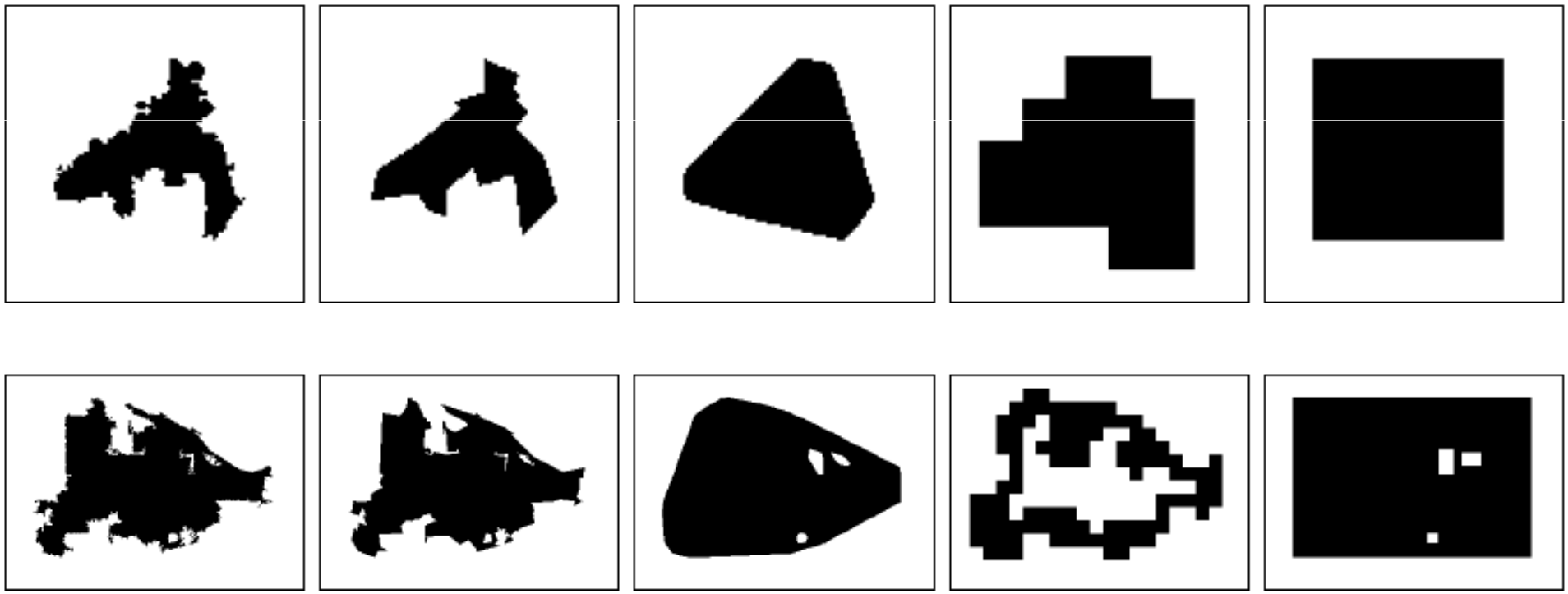
Composite regions



a b

FIGURE 11.35 (a) A simple composite region. (b) Tree representation obtained by using the relationship "inside of."

Examples



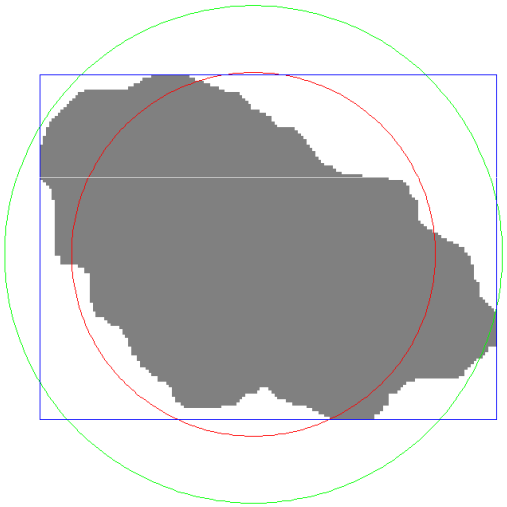
Region representation examples. Rows show representations for two different regions. Columns represent, from left to right: original boundary, smoothed polygon, convex hull, grid representation, and minimum bounding rectangle.

Region descriptors

- Boundary descriptors use external representations to model the shape characteristics of regions.
- Regional descriptors use internal representations to model the internal content of regions.
- Examples:
  - Shape properties (both internal and external)
  - Fourier descriptors (external)
  - Statistics and histograms (internal)

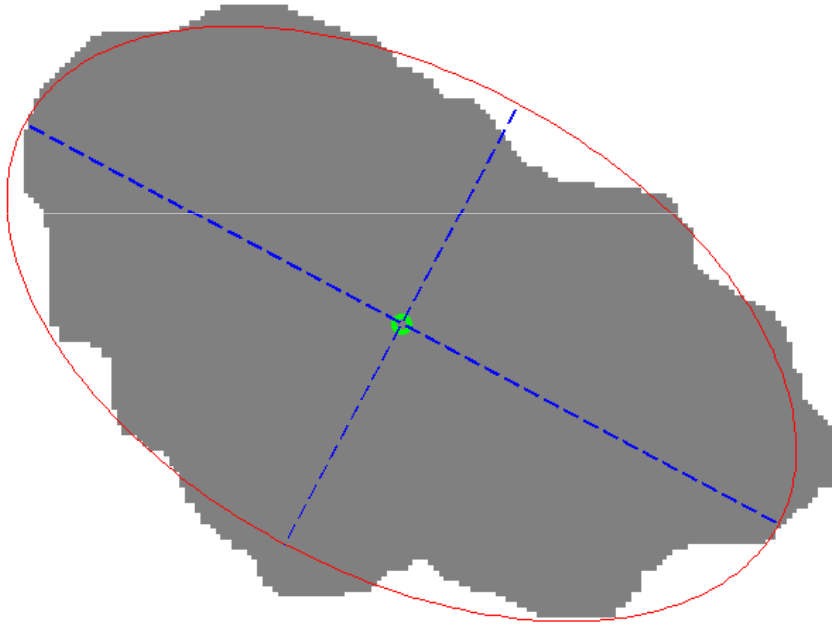
Shape properties

- Area
  - Number of pixels in the region
- Perimeter
  - Number of pixels on the boundary
  - Number of vertical and horizontal components plus $\sqrt{2}$ times the number of diagonal components
- Bounding box
- Diameter
  - Maximum distance between boundary points
- Equivalent diameter
  - Diameter of a circle with the same area as the region



Shape properties

- Principal axes of inertia
  - Compute the mean of pixel coordinates (centroid)
  - Compute the covariance matrix of pixel coordinates
  - Major axis is the eigenvector of the covariance matrix corresponding to the larger eigenvalue
  - Minor axis is the eigenvector of the covariance matrix corresponding to the smaller eigenvalue



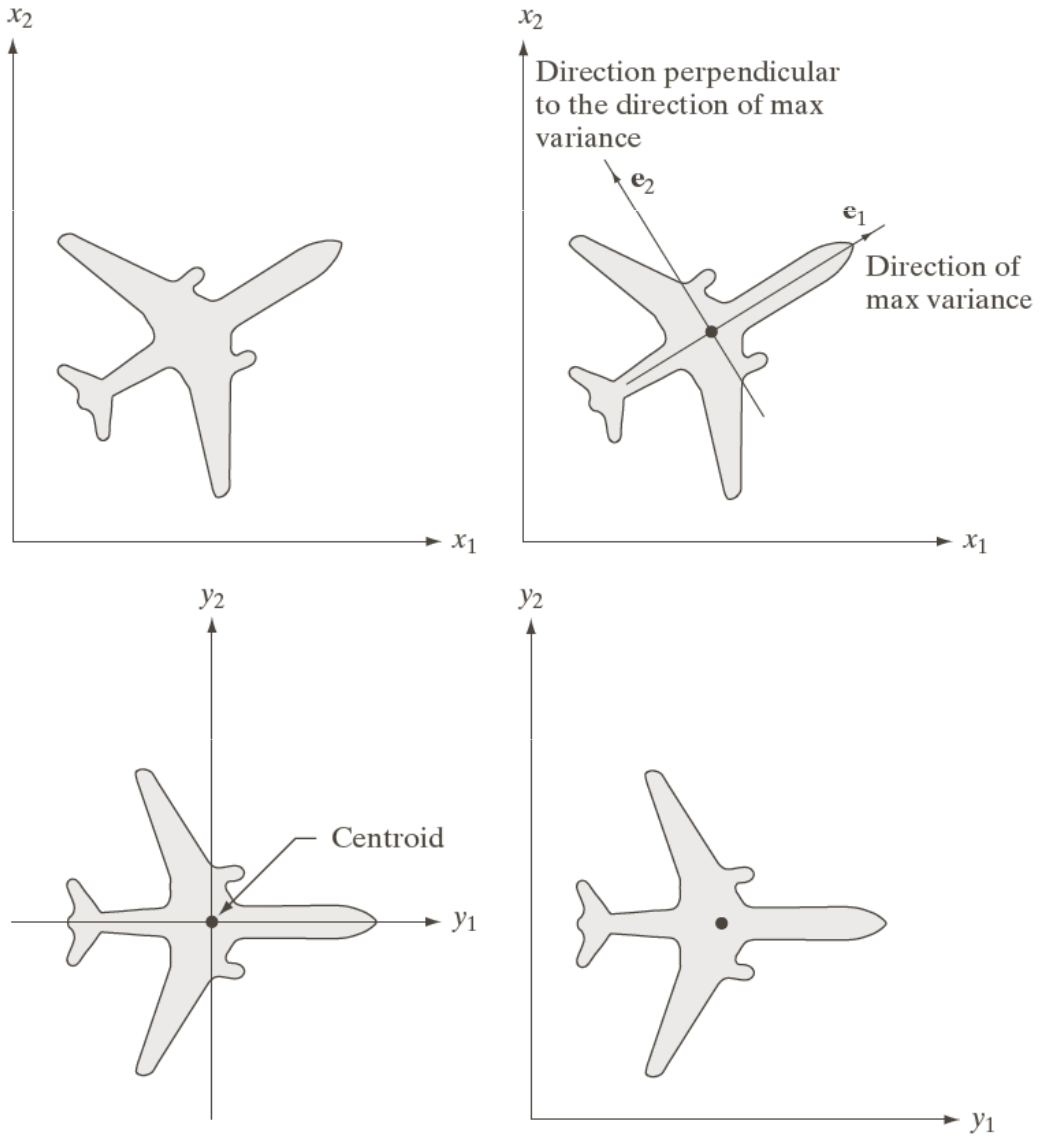
Shape properties

$$\mathbf{m}_x = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$$

$$\mathbf{C}_x = \begin{bmatrix} 3.333 & 2.00 \\ 2.00 & 3.333 \end{bmatrix}$$

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

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

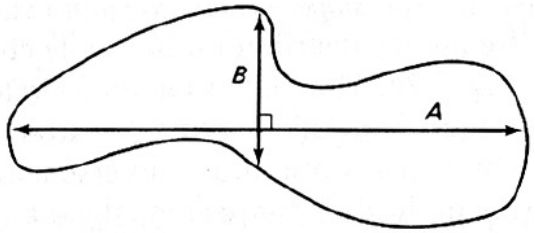


a b
c d

FIGURE 11.43
(a) An object. (b) Object showing eigenvectors of its covariance matrix. (c) Transformed object, obtained using Eq. (11.4-6). (d) Object translated so that all its coordinate values are greater than 0.

Shape properties

- Orientation
  - Orientation of the major axis with respect to the horizontal axis
- Eccentricity (elongation)
  - Ratio of the length of maximum chord A to maximum chord B perpendicular to A
  - Ratio of the principal axes of inertia
- Compactness
  - Perimeter^2 / area (minimized by a disk)
- Extent
  - Area of region / area of its bounding box
- Solidity
  - Area of region / area of its convex hull



Shape properties

- Spatial variances
  - Variance of pixel coordinates along the principal axes of inertia
- Euler number
  - Number of connected regions – number of holes (actually a topological property that describes the connectedness of a region, not its shape)
- Moments

$$m_{ij} = \sum_x (x - \bar{x})^i (y - \bar{y})^j$$

$$\bar{x} = \frac{1}{n} \sum_x x \qquad \text{and} \qquad \bar{y} = \frac{1}{n} \sum_y y.$$

Fourier descriptors

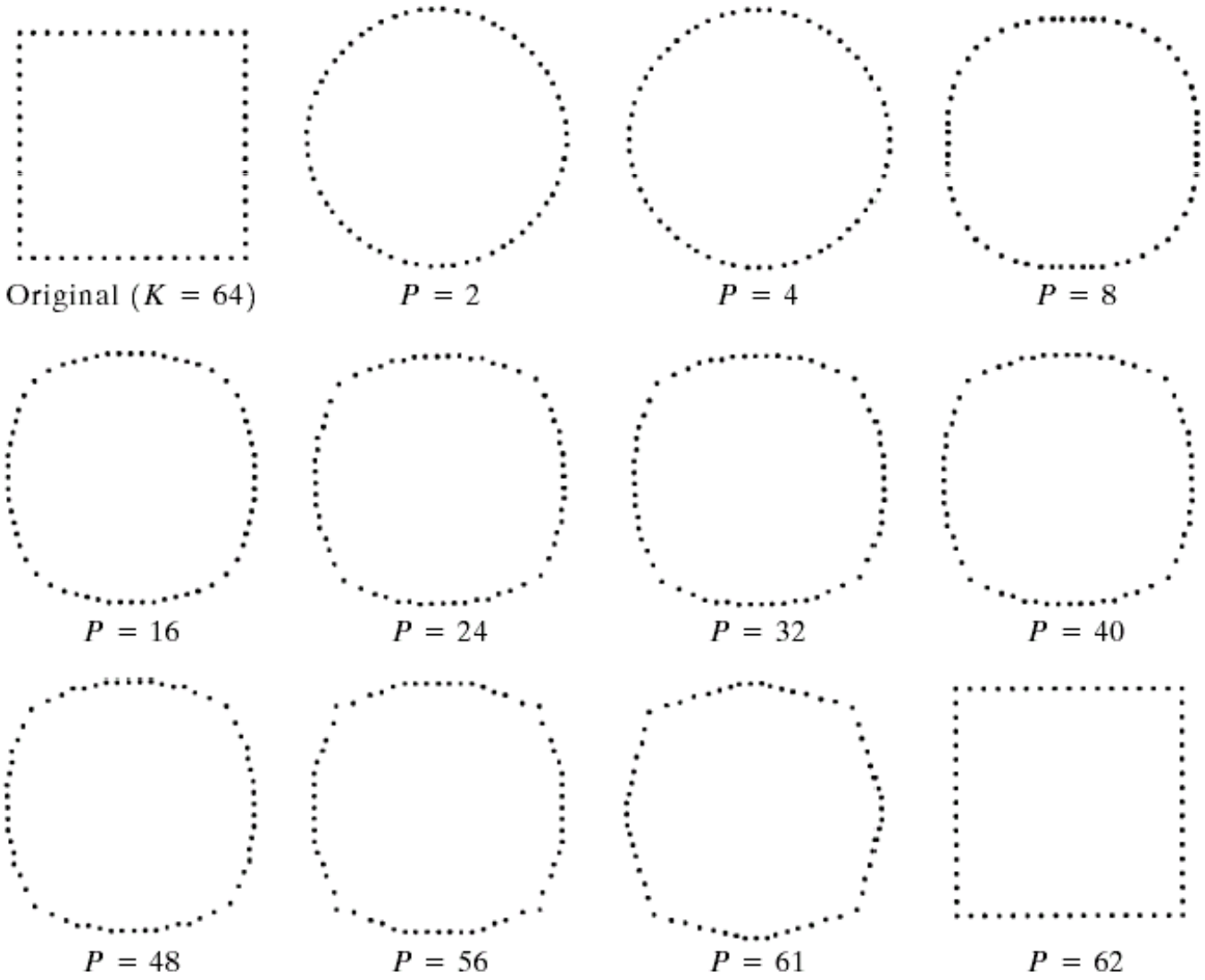
- Given the list of K boundary points, start at an arbitrary point and traverse the boundary in a particular direction.
- The coordinates can be expressed in the form $x(k) = x_k$ and $y(k) = y_k$ where $k = 0, 1, \ldots, K-1$ shows the order of traversal.
- Each coordinate pair can be treated as a complex number $s(k) = x(k) + j\, y(k)$ for $k = 0, 1, \ldots, K-1$.
- The complex coefficients of the discrete Fourier transform of s(k) are called the Fourier descriptors of the boundary.

Fourier descriptors

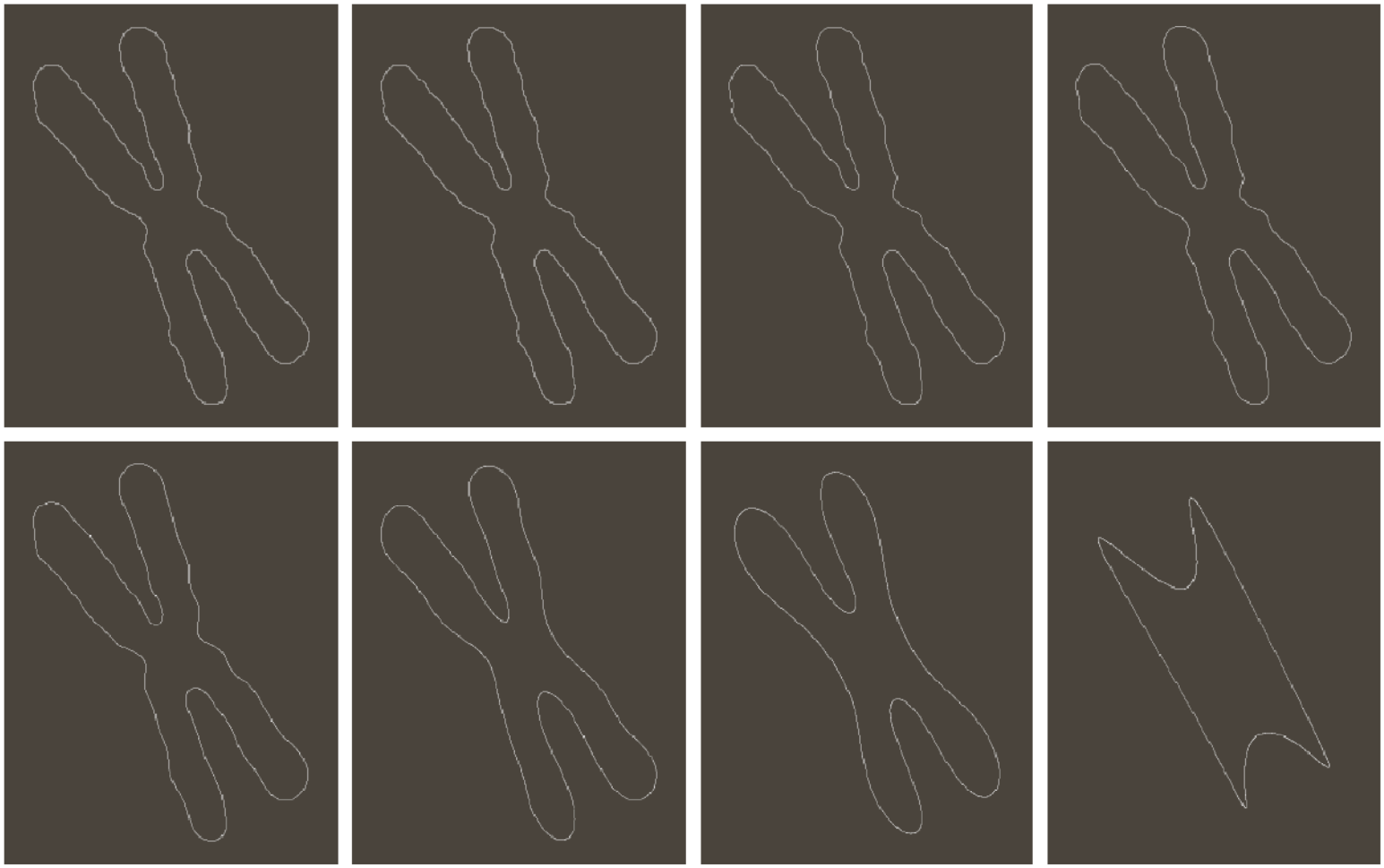
- The inverse Fourier transform reconstructs s(k), i.e., the boundary.
- If only the first P coefficients are used in the reconstruction, we obtain an approximation to the boundary.
- Note that only P terms are used to obtain each component of s(k), but k still ranges from 0 to K-1.
- Since the high-frequency components of the Fourier transform account for fine detail, the smaller P becomes, the more detail is lost on the boundary.

Fourier descriptors

FIGURE 11.14 Examples of reconstruction from Fourier descriptors. P is the number of Fourier coefficients used in the reconstruction of the boundary.



Fourier descriptors



a b c d
e f g h

FIGURE 11.20 (a) Boundary of human chromosome (2868 points). (b)–(h) Boundaries reconstructed using 1434, 286, 144, 72, 36, 18, and 8 Fourier descriptors, respectively. These numbers are approximately 50%, 10%, 5%, 2.5%, 1.25%, 0.63%, and 0.28% of 2868, respectively.

Fourier descriptors

- Fourier descriptors are not directly insensitive to geometrical changes such as translation, rotation and scale, but the changes in these parameters can be related to simple transformations on the descriptors.

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

TABLE 11.1 Some basic properties of Fourier descriptors.

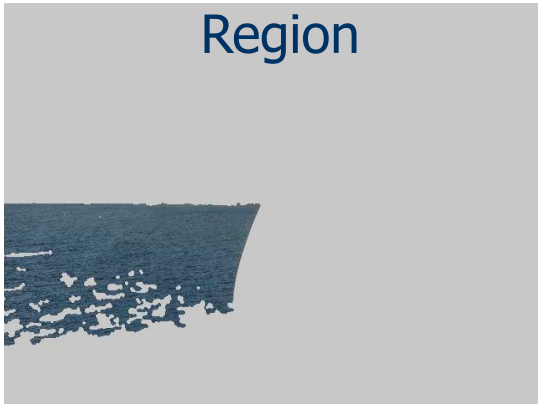
$(\Delta_{xy} = \Delta_x + j\,\Delta_y)$

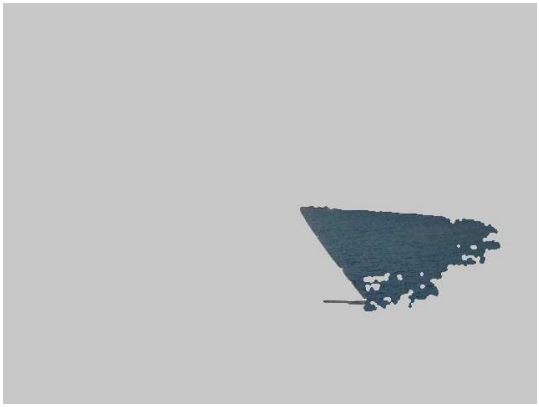
Statistics and histograms

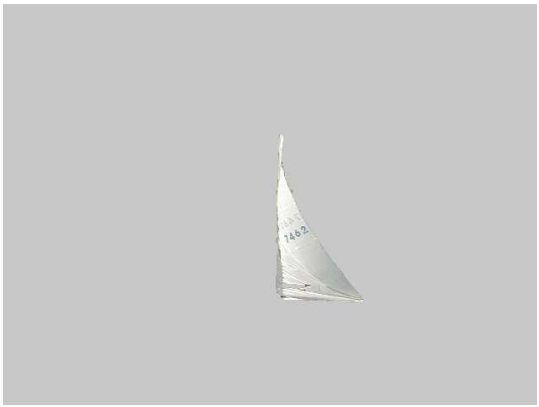
- Contents of regions can be summarized using statistics (e.g., mean, standard deviation) and histograms of pixel features.
- Commonly used pixel features include
  - Gray tone,
  - Color (RGB, HSV, …),
  - Texture,
  - Motion.
- Then, the resulting region level features can be used for clustering, retrieval, classification, etc.

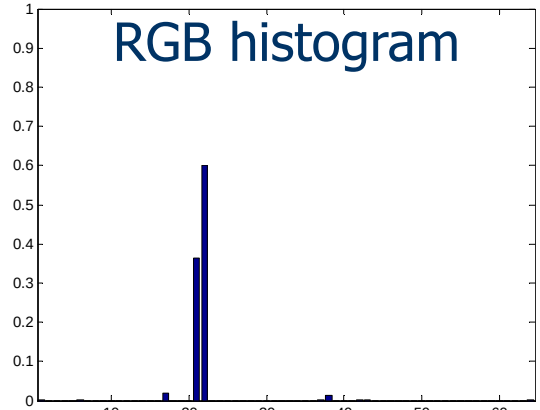
Statistics and histograms

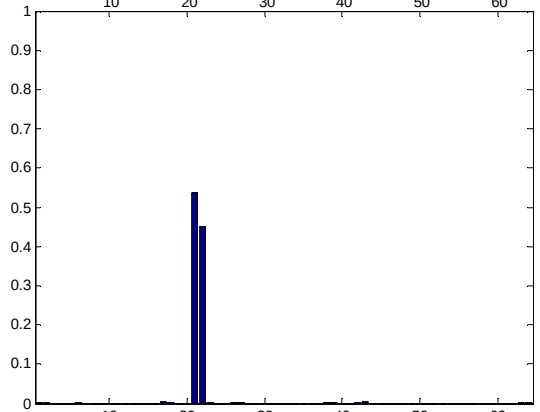
Statistics and histograms

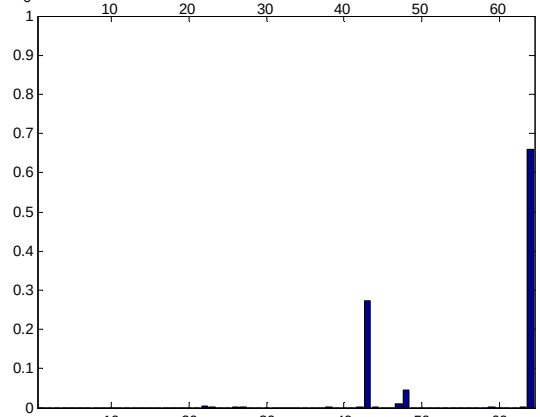


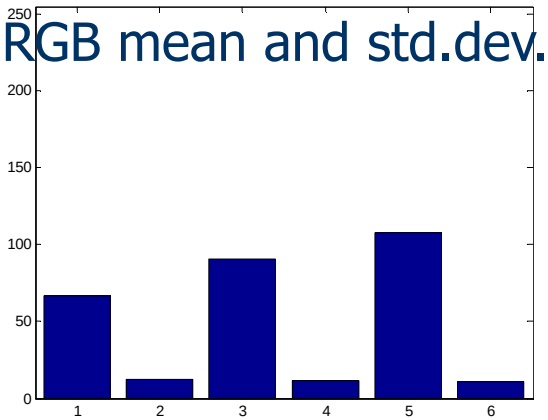


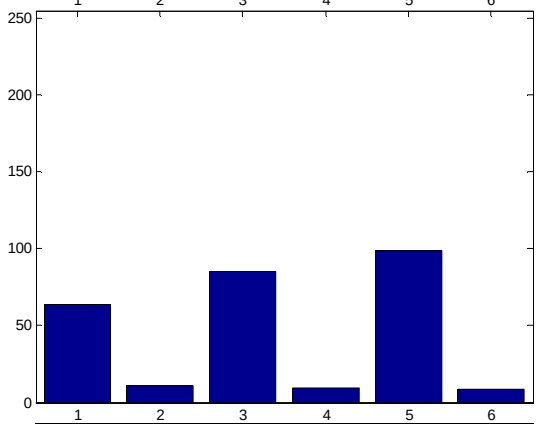


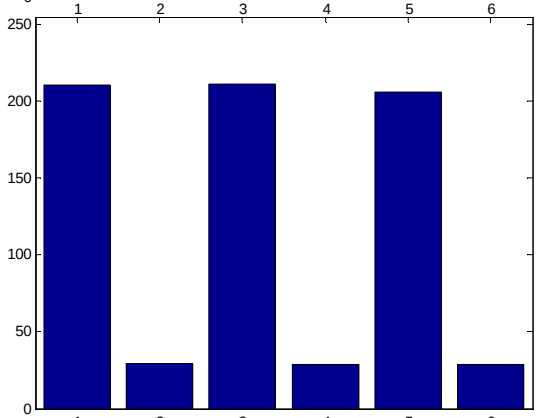




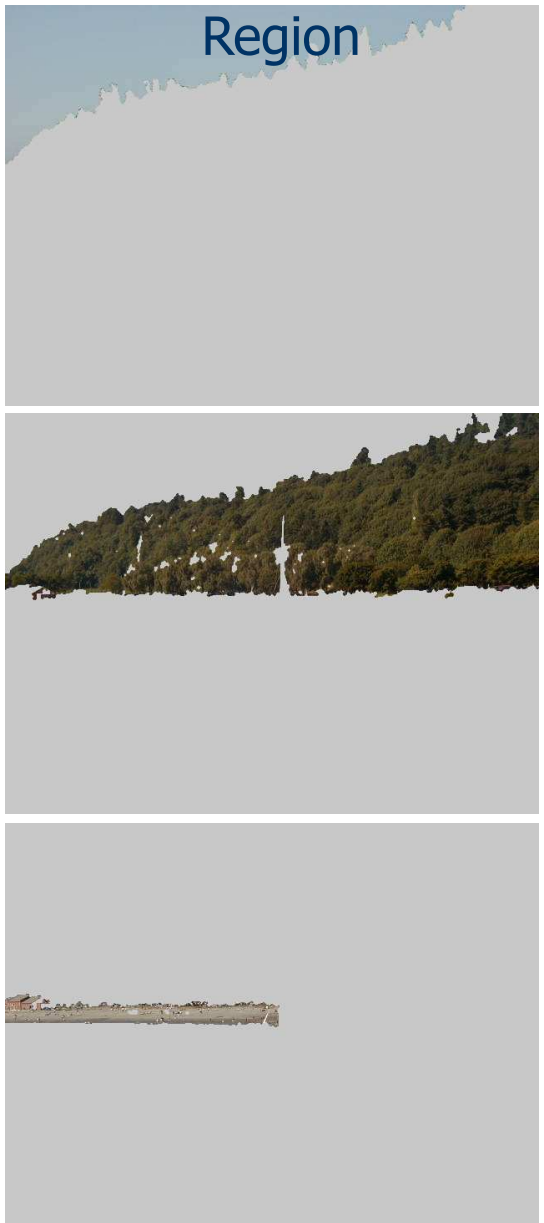








Statistics and histograms



Statistics and histograms

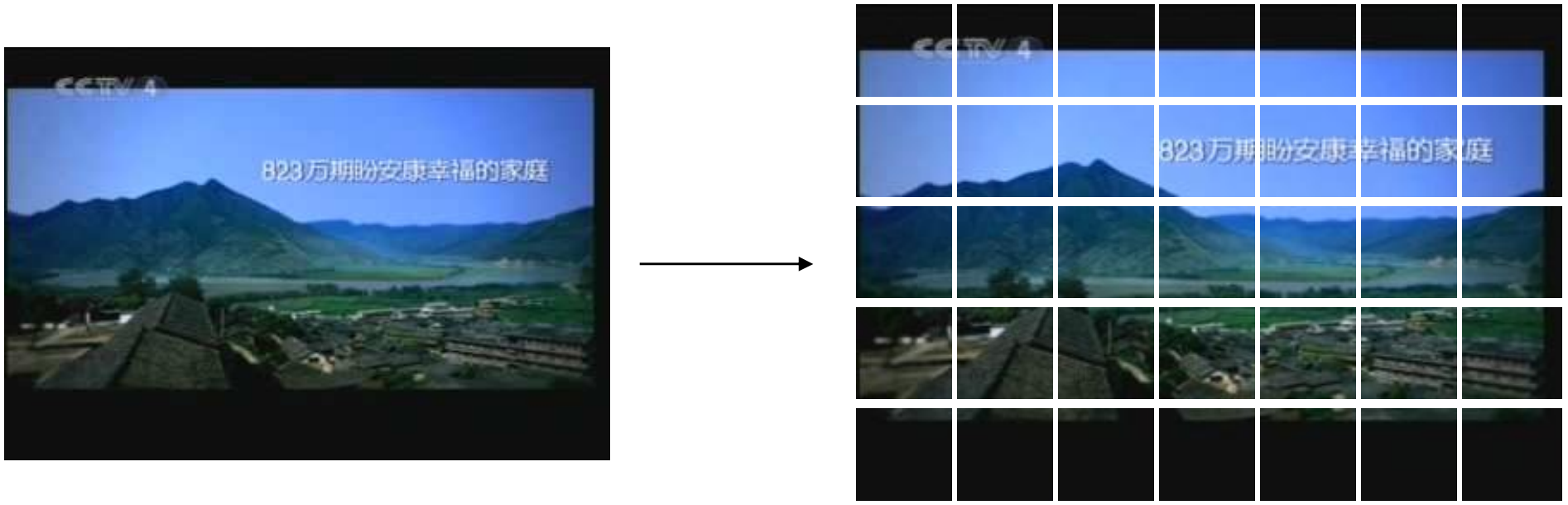
Region clusters obtained using the histograms of the HSV values of their pixels. Each row represents the clusters corresponding to sky, rock, tree, road.

Image representations and descriptors

- Popular image representations (in increasing order of complexity) include:
  - Global representation
  - Tiled representations
  - Quadtrees
  - Region adjacency graphs
  - Attributed relational graphs

Tiled representations

- Images can be divided into fixed size or variable size grids that are overlapping or non-overlapping.



Dividing an image into 5x7 fixed size non-overlapping grid cells.

Tiled representations

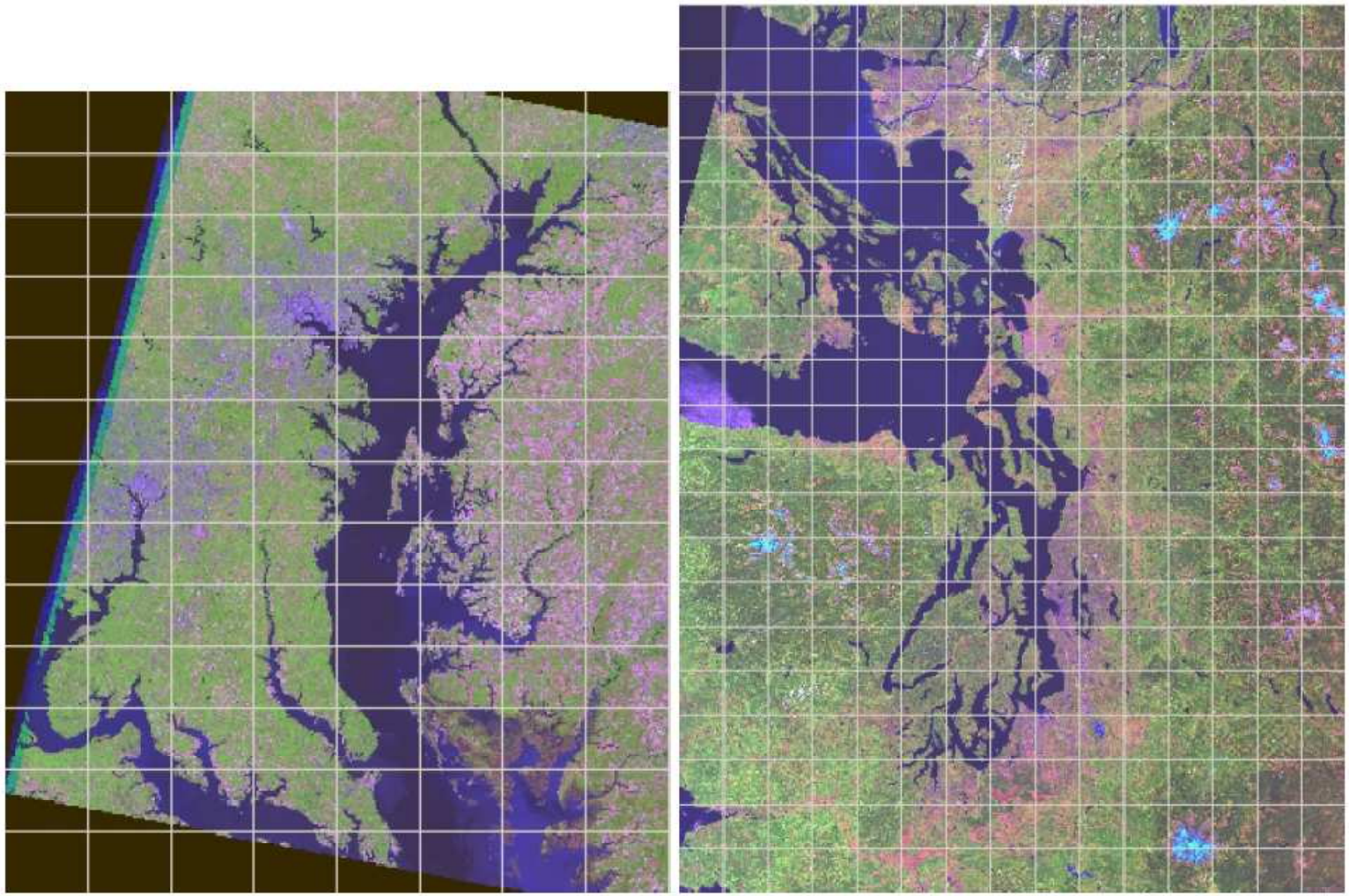
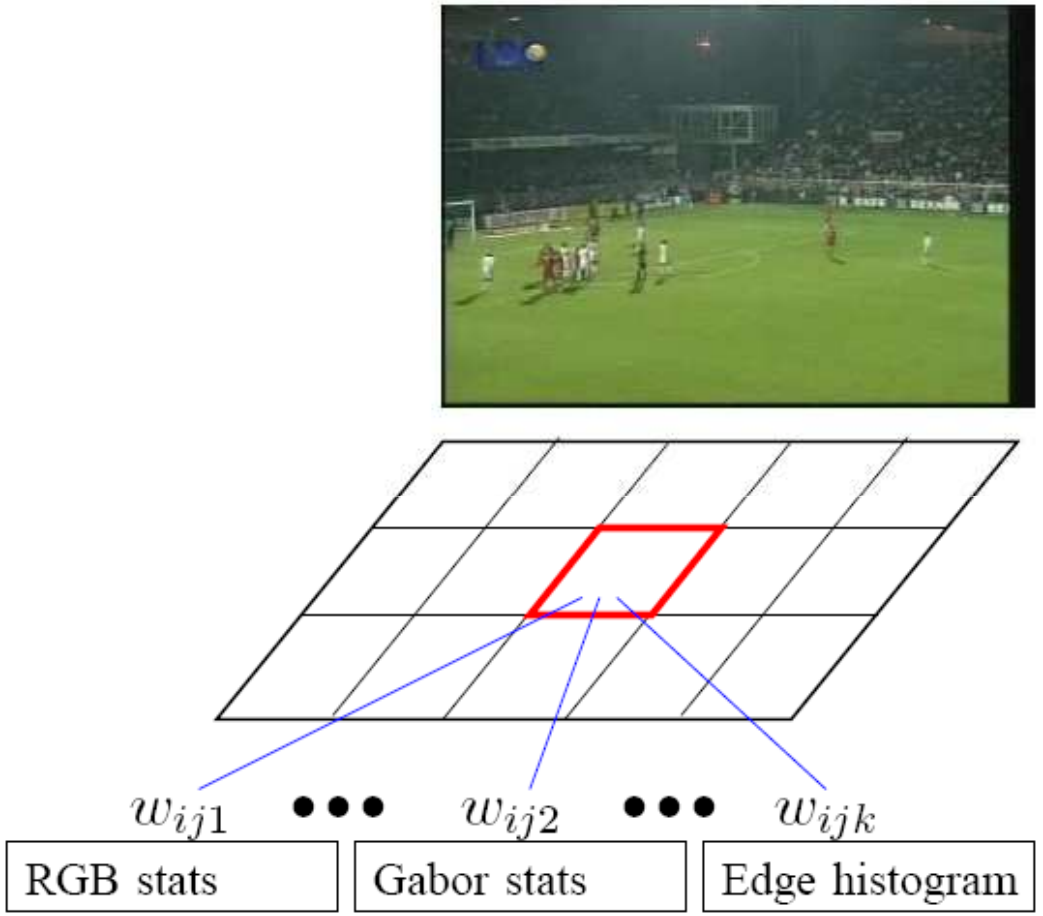


Fig. 2. Multi-spectral LANDSAT scenes of Washington, DC (left), and southern British Columbia and Washington State (right) that are divided into smaller tiles for easier processing. The first scene ($5,376 \times 6,656$; 205 MB) is divided into 104 tiles and the second one ($7,680 \times 10,240$; 450 MB) is divided into 300 tiles. The best tile size for each scene is automatically determined by the system.

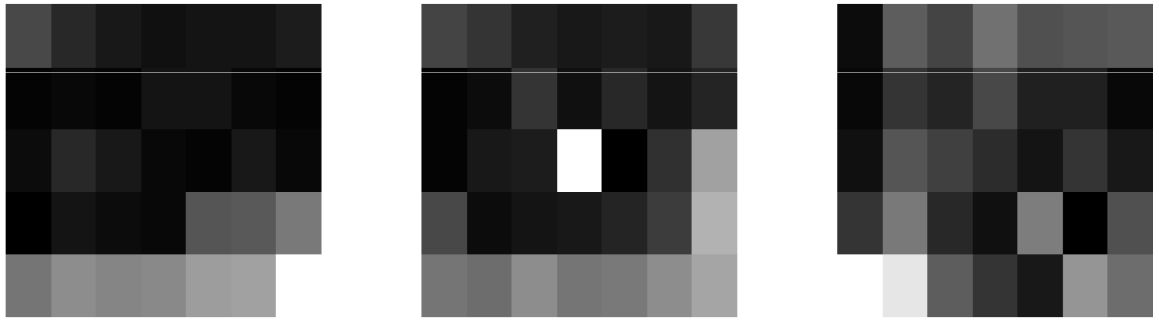
Tiled representations

- Example application: learning the importance of different features in different grid cells using user feedback.



Tiled representations

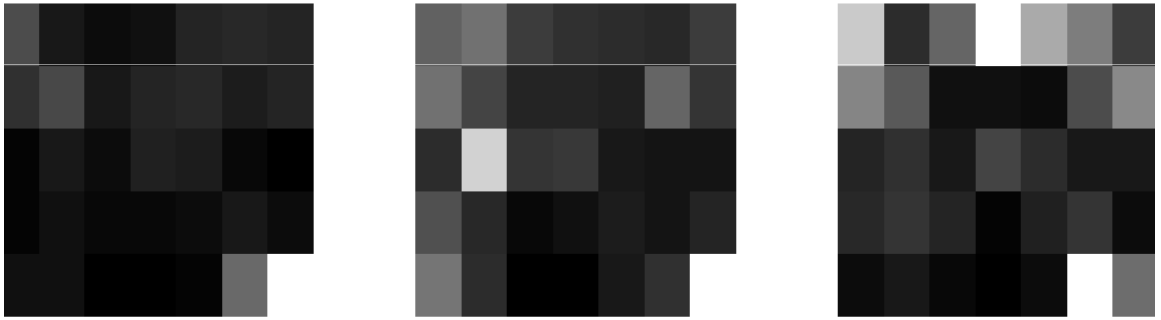
Results of a football query.



Weights for different features (left to right: HSV, LUV, RGB) and grid cells for the football query. Brighter colors represent higher weights.

Tiled representations

Results of a basketball query.

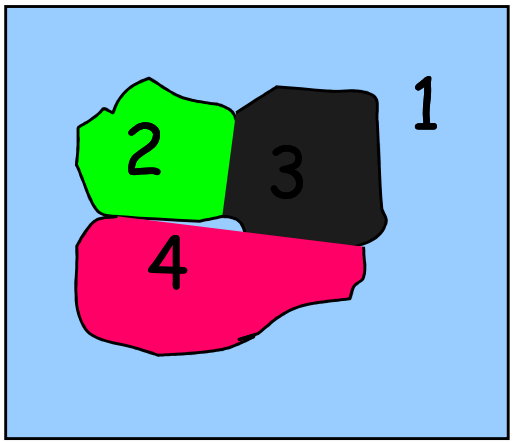


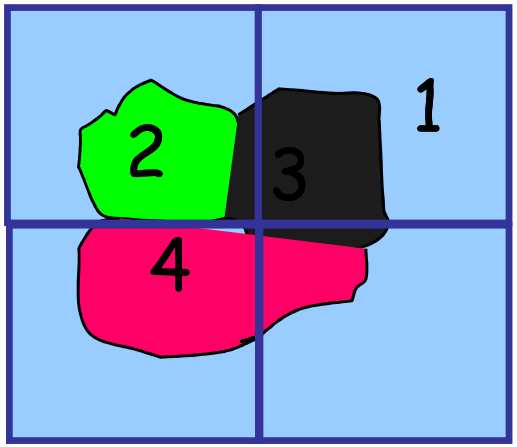
Weights for different features (left to right: HSV, LUV, RGB) and grid cells for the basketball query. Brighter colors represent higher weights.

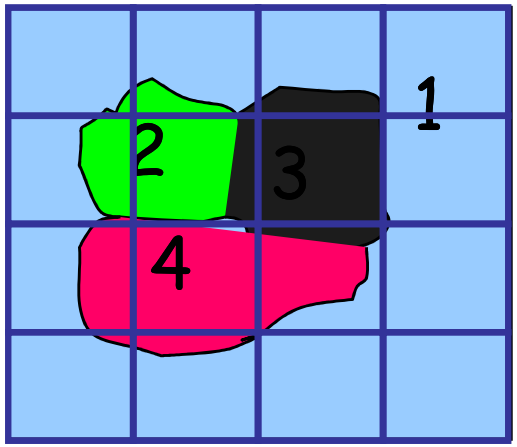
Quadtrees

- Quadtrees:
  - Trees where nodes have 4 children.
- Building a quadtree:
  - Nodes represent regions.
  - Every time a region is split, its node gives birth to 4 children.
  - Leaves are nodes for uniform regions.
- Merging:
  - Siblings that are “similar” can be merged.

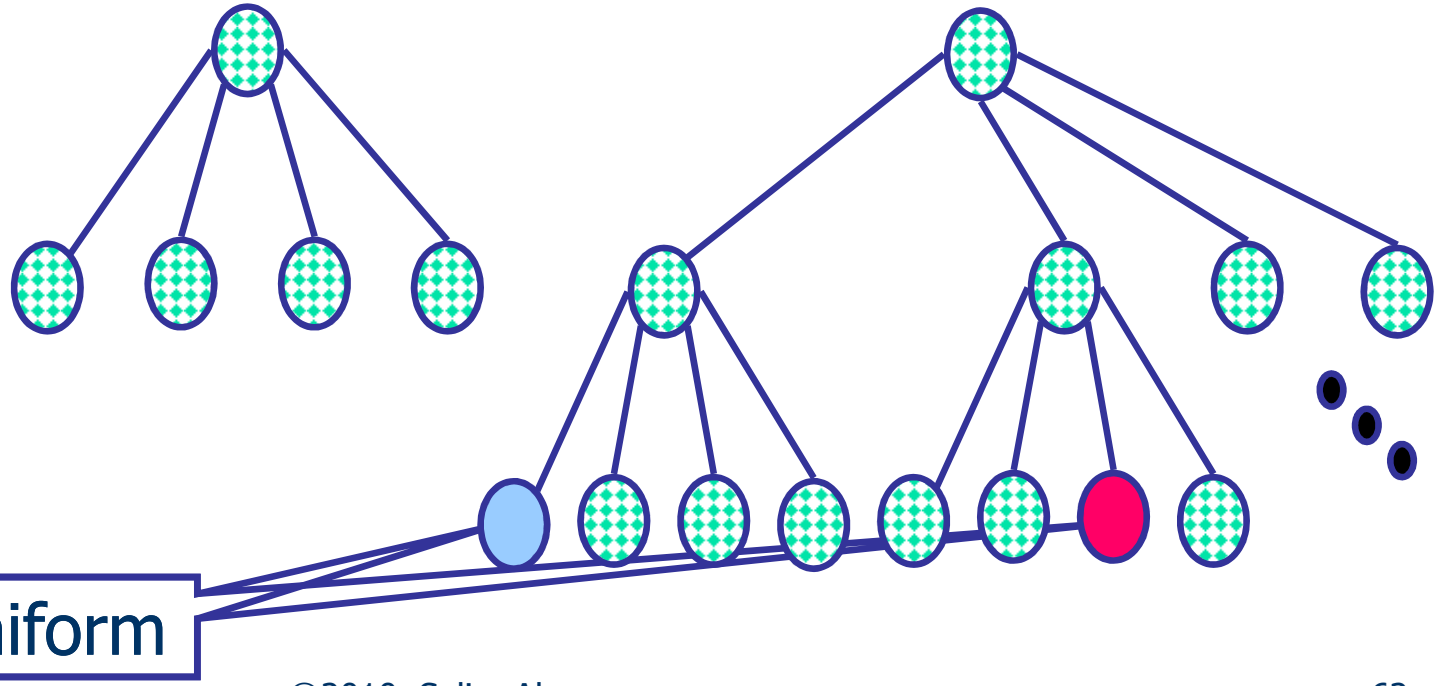
Quadtrees



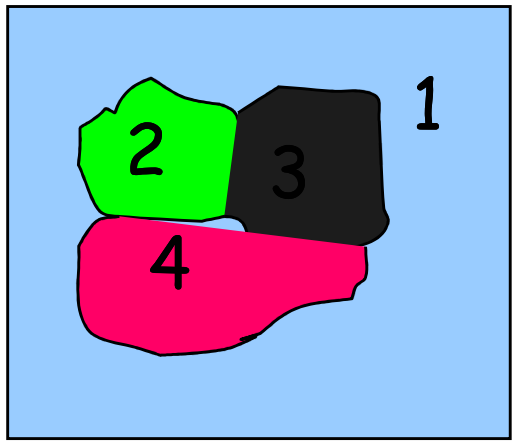


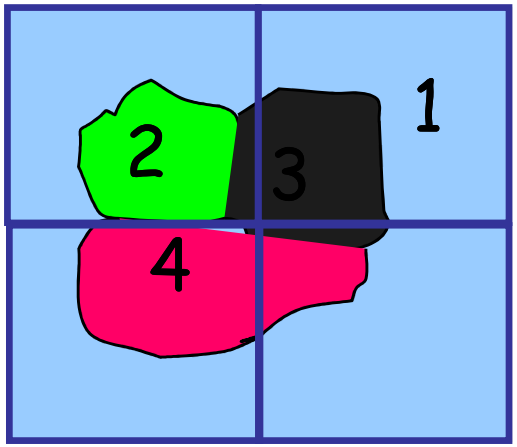


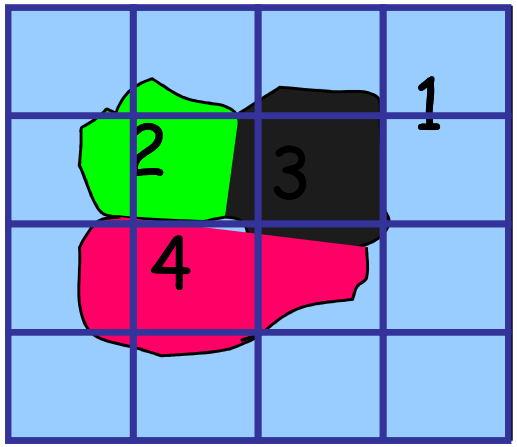
Not uniform

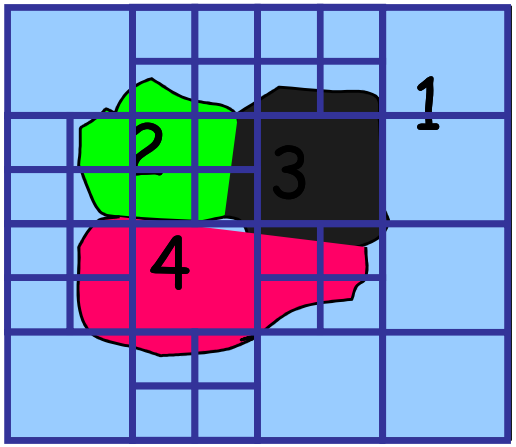


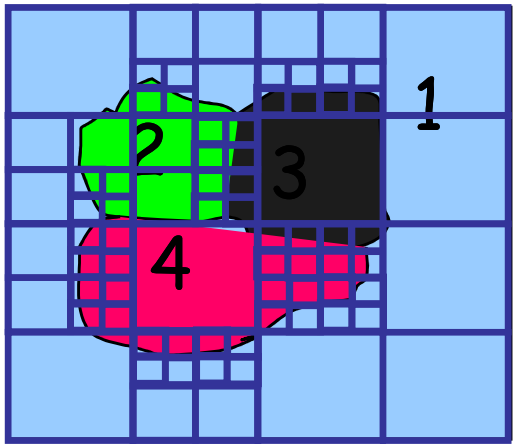
Quadtrees





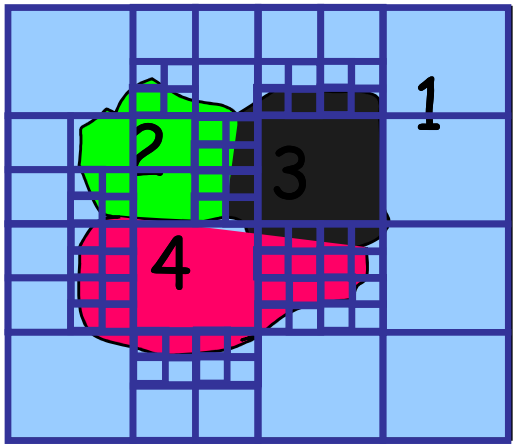


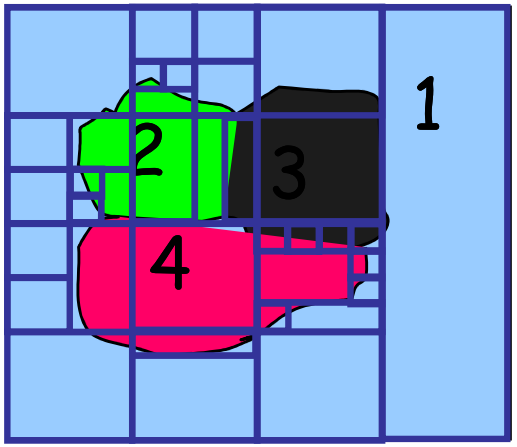


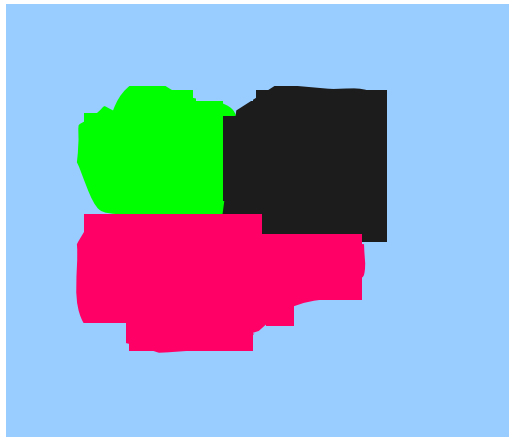


Splitting…

Quadtrees



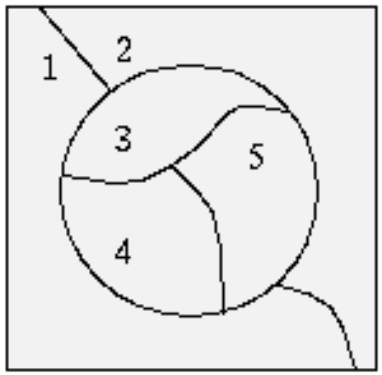




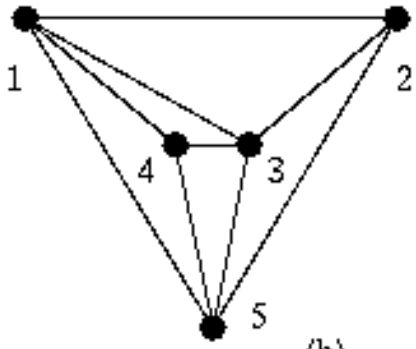
Merging…

Region adjacency graphs

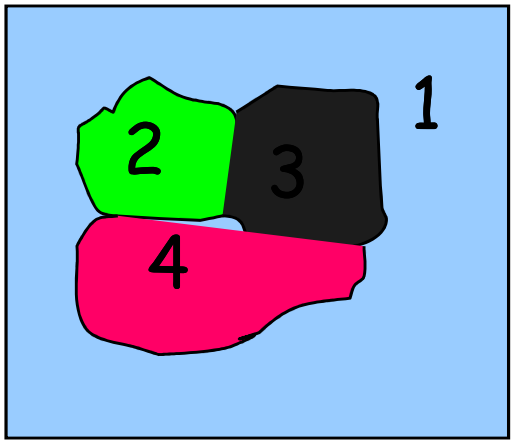
- A region adjacency graph (RAG) is a graph in which each node represents a region of the image and an edge connects two nodes if the regions are adjacent.

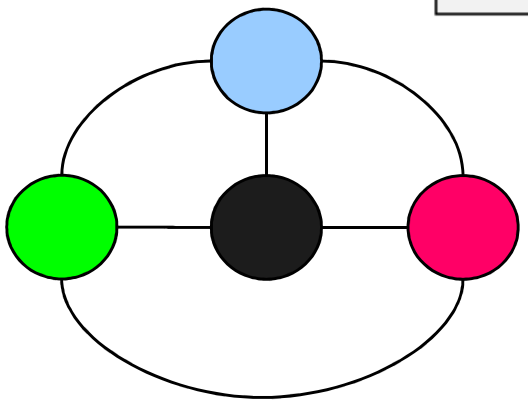


(a)

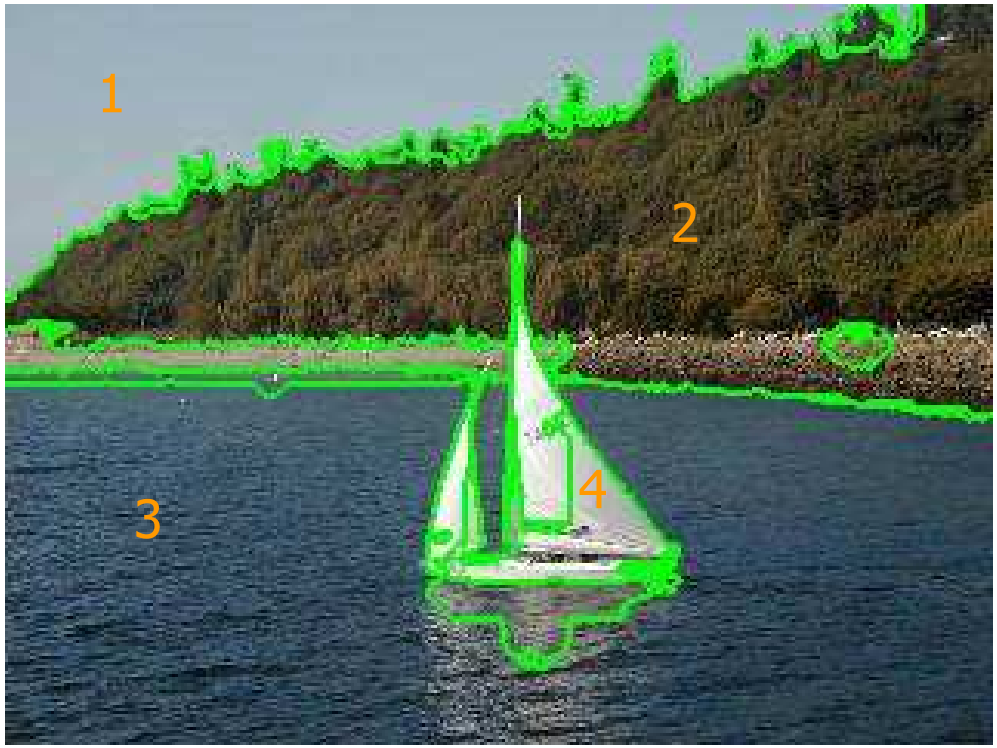


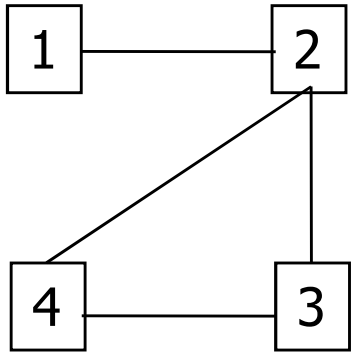
(b)





Region adjacency graphs





Attributed relational graphs

- Attributed relational graphs (ARG) generalize ordinary graphs by attaching discrete or continuous features (attributes) to the vertices and edges.
- Formally, an attributed relational graph $\mathcal{G}$ is a 4-tuple $\mathcal{G}=(\mathcal{N},\mathcal{E},\mu,\nu)$ where
  - $\mathcal{N}$ is the set of nodes,
  - $\mathcal{E} \subseteq \mathcal{N} \times \mathcal{N}$ is the set of edges,
  - $\mu : \mathcal{N} \rightarrow \mathcal{L}_{\mathcal{N}}$ is a function assigning labels to the nodes,
  - $\nu : \mathcal{E} \rightarrow \mathcal{L}_{\mathcal{E}}$ is a function assigning labels to the edges.

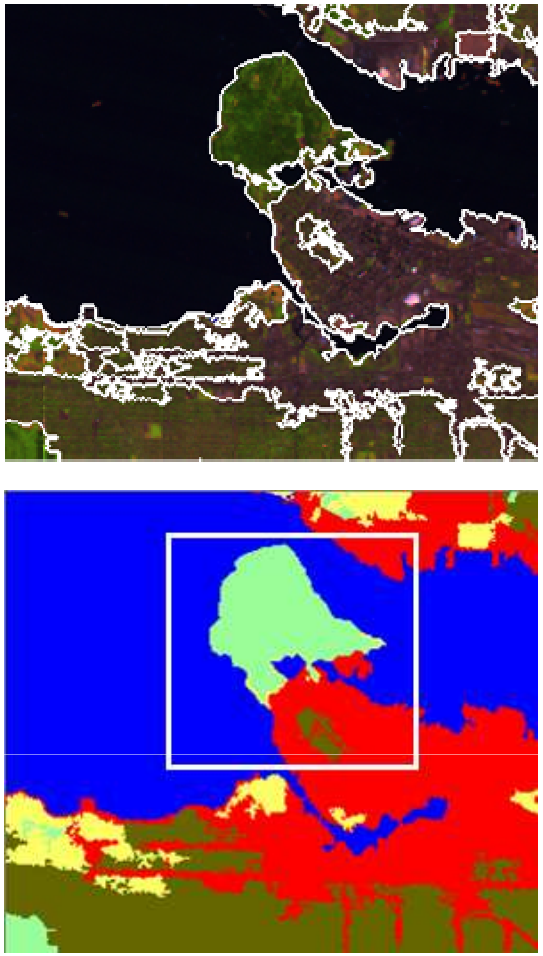
Attributed relational graphs

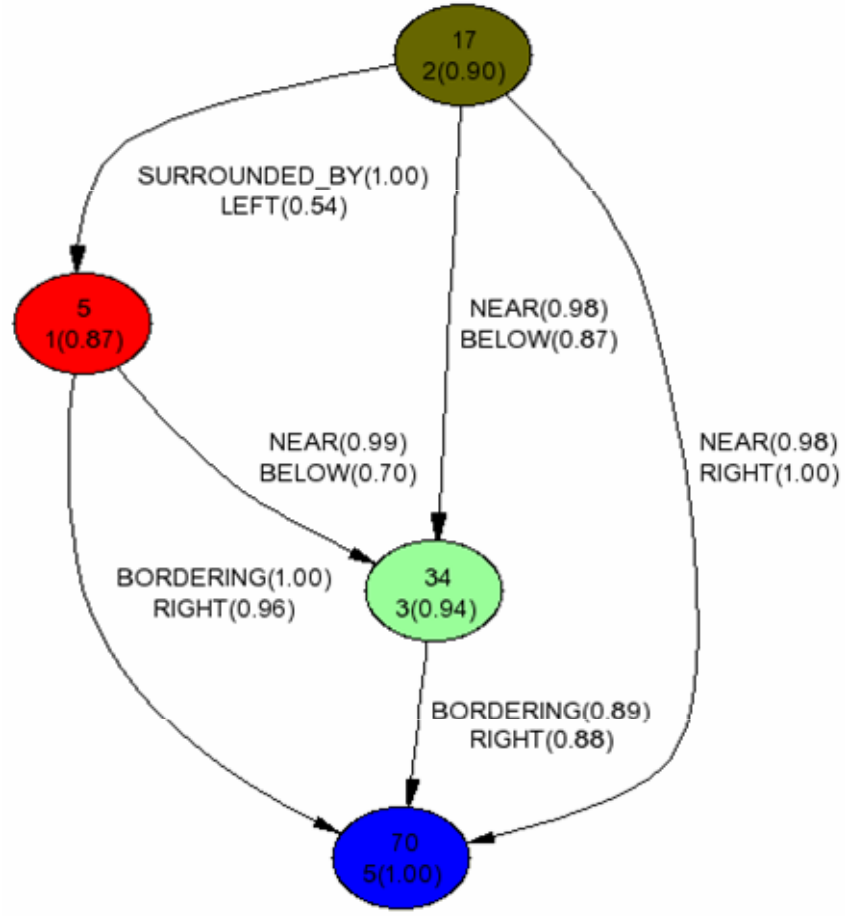
- Many real-world entities can be represented as an ARG.
- For example, visual scenes can be modeled as the composition of the objects with specific spatial/attribute relationships.
- Molecular compounds can be represented as an ARG with atoms as vertices and bonds as edges.
- A web site can be represented as an ARG with web pages as vertices and URL links as edges.

Attributed relational graphs

- Example application: modeling remote sensing image content using ARGs.
  - Nodes in the ARG represent the regions and edges represent the spatial relationships between these regions.
  - Nodes are labeled with the class (land cover/use) names and the corresponding confidence values for these class assignments.
  - Edges are labeled with the spatial relationship classes and the corresponding degrees for these relationships.

Attributed relational graphs

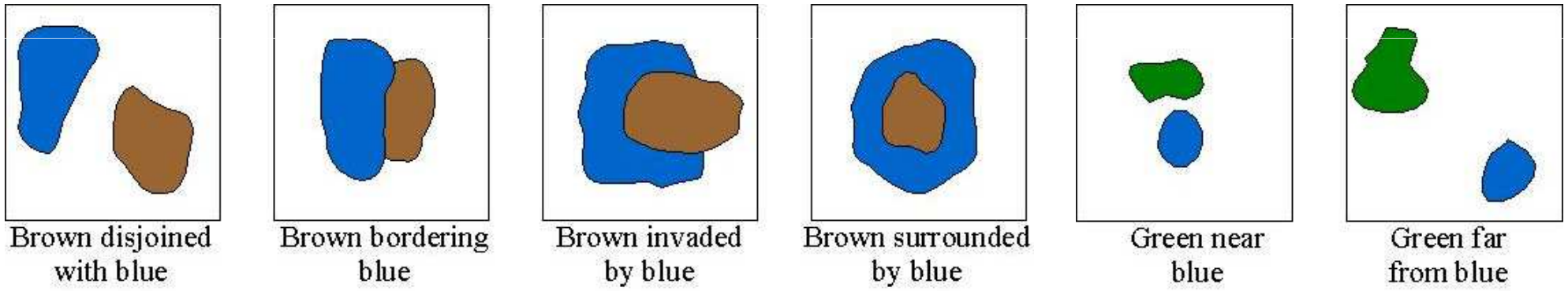




ARG for an example scene (marked with a white rectangle) containing regions classified as water (blue), city center (red), residential area (brown) and park (green). Nodes are labeled using region id, class id, class probability (in parenthesis) and edges are labeled using relationship names and membership values (in parenthesis).

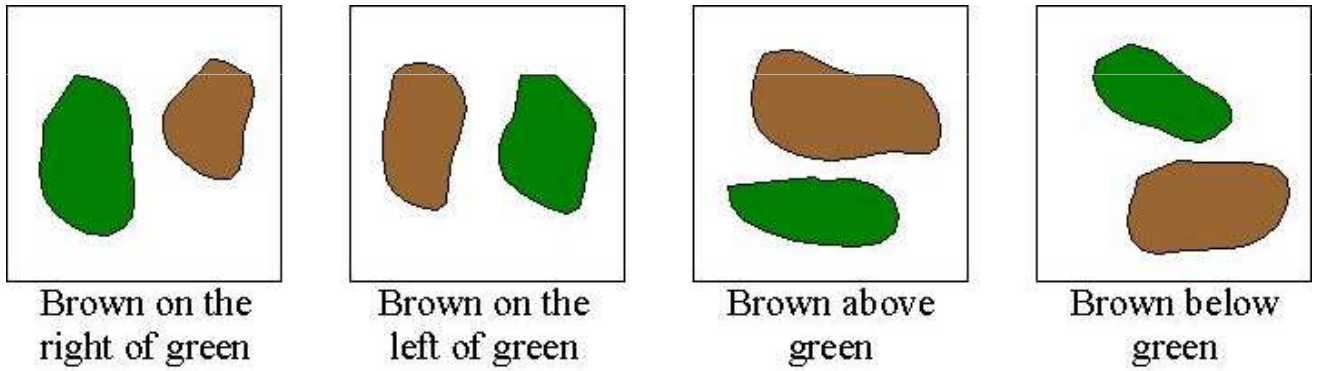
Attributed relational graphs

Pairwise spatial relationships



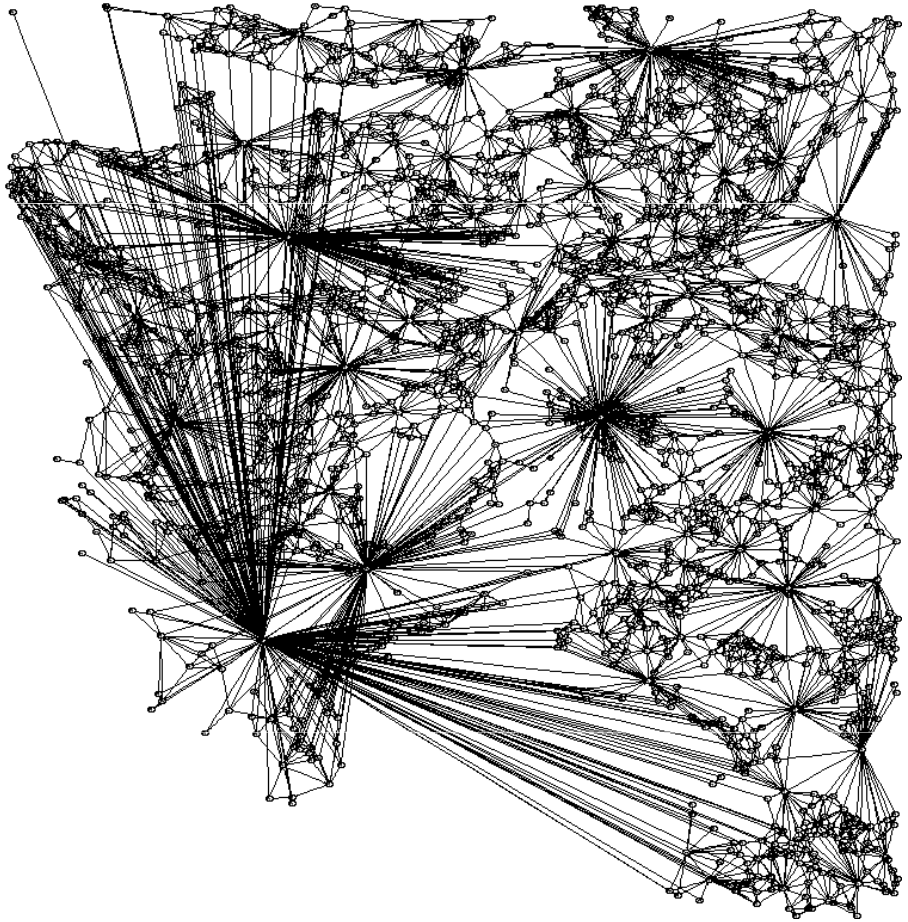
Perimeter-class relationships

Distance-class relationships



Orientation-class relationships

Attributed relational graphs

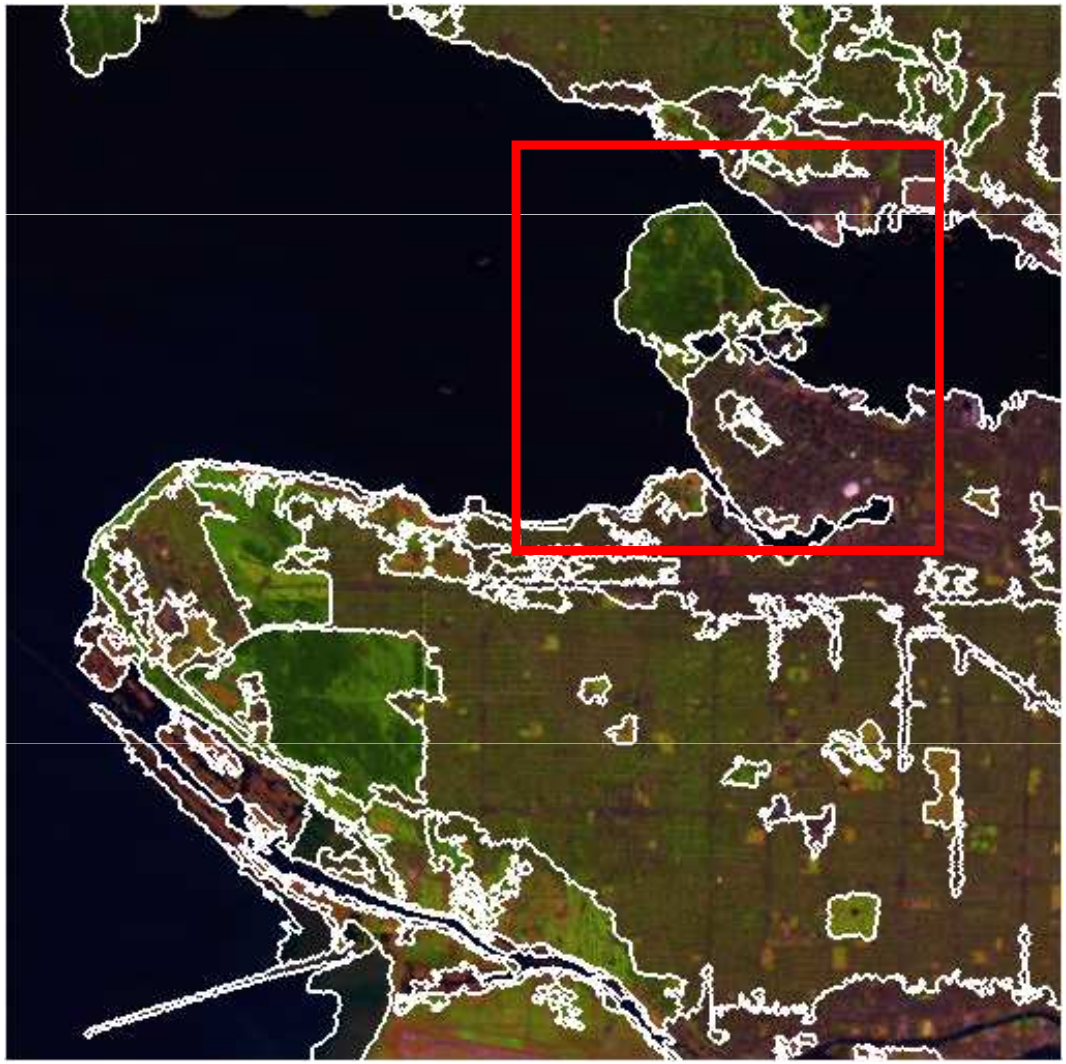


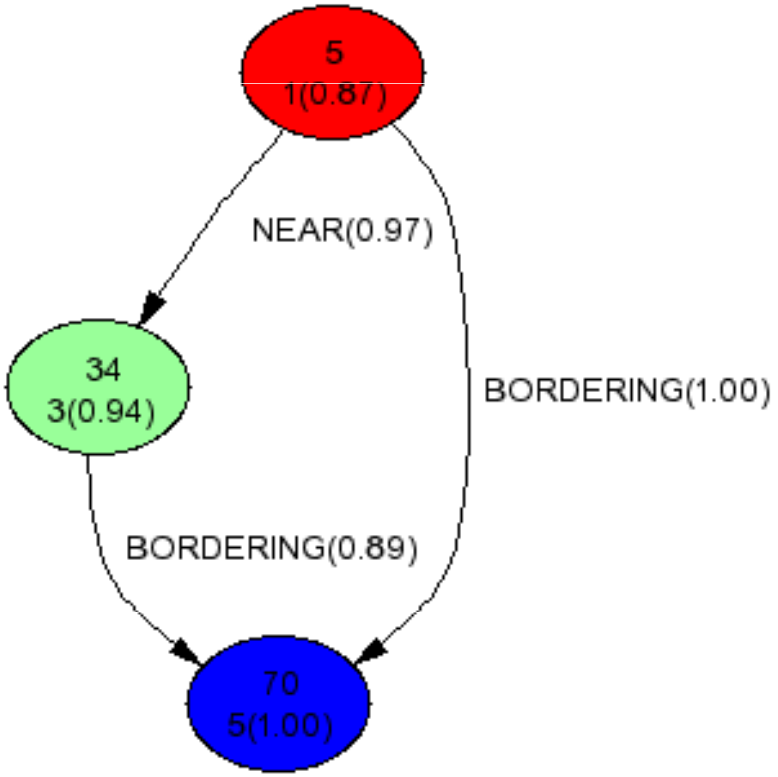
ARG of a LANDSAT scene. Nodes are located at the centroids of the corresponding regions. Edges are drawn only for pairs that are within 10 pixels of each other to keep the graph simple.

Attributed relational graphs

- An important problem is how similarities between two scenes represented using ARGs can be found.
- Relational matching has been extensively studied in structural pattern recognition.
- One possible solution is the "editing distance" between two ARGs that is defined as the minimum cost taken over all sequences of operations (error corrections such as substitution, insertion and deletion) that transform one ARG to the other.

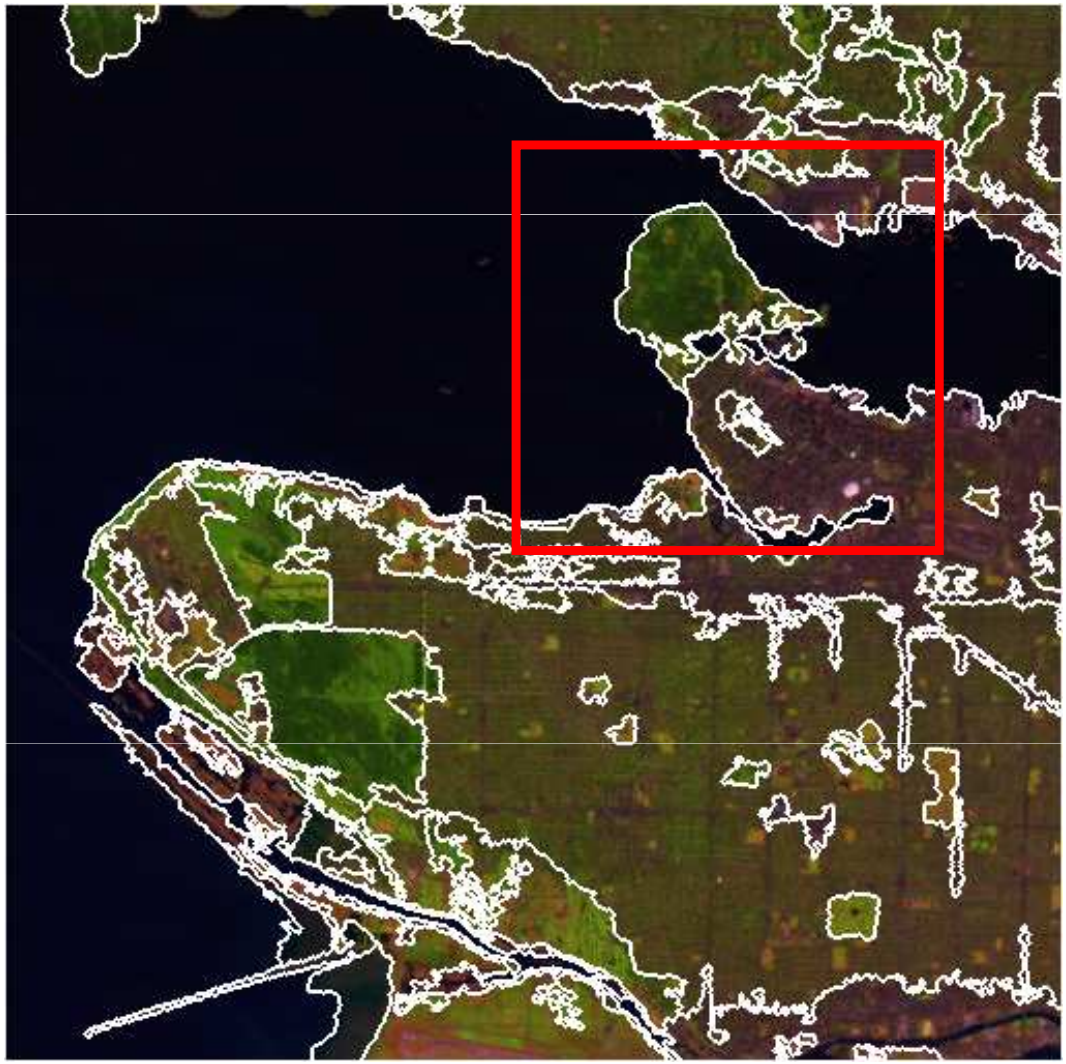
Attributed relational graphs

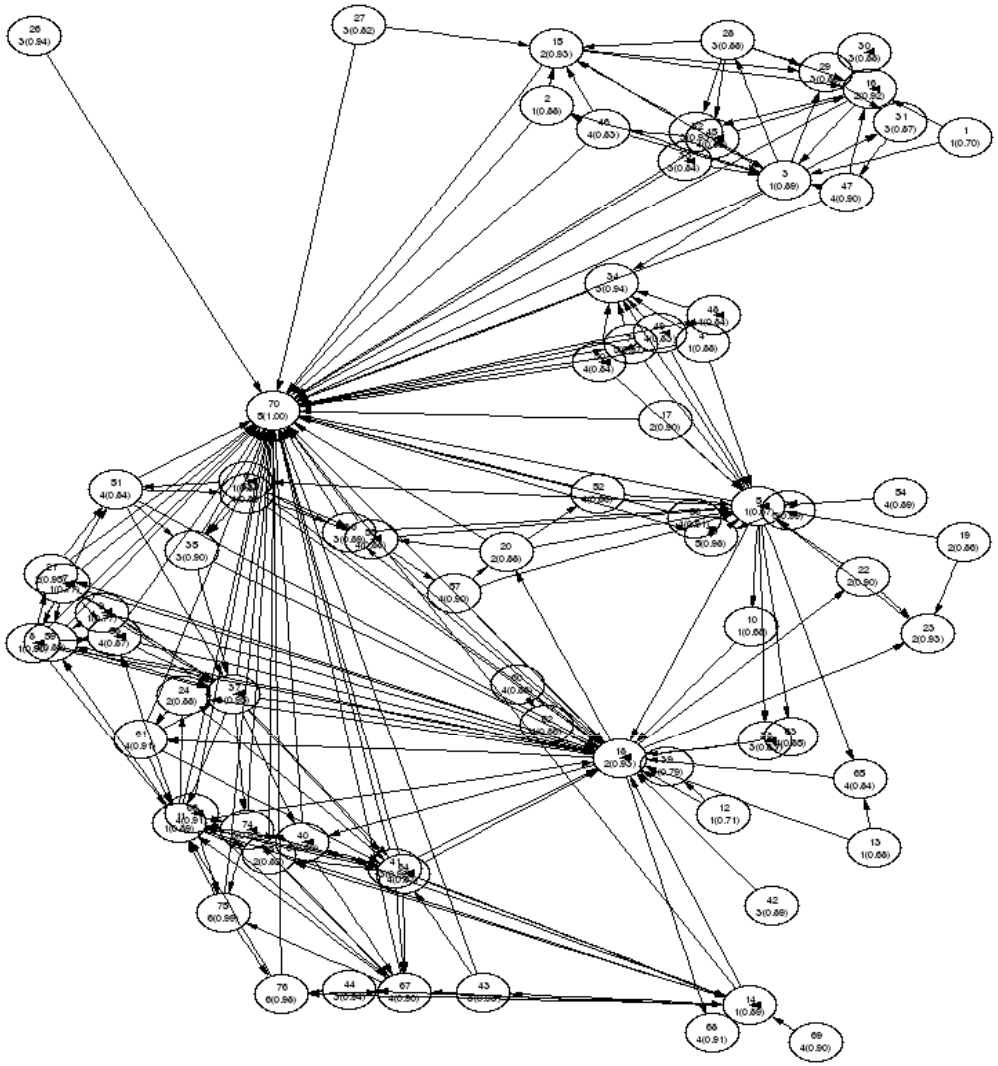




Find the ARG of the area of interest (query).

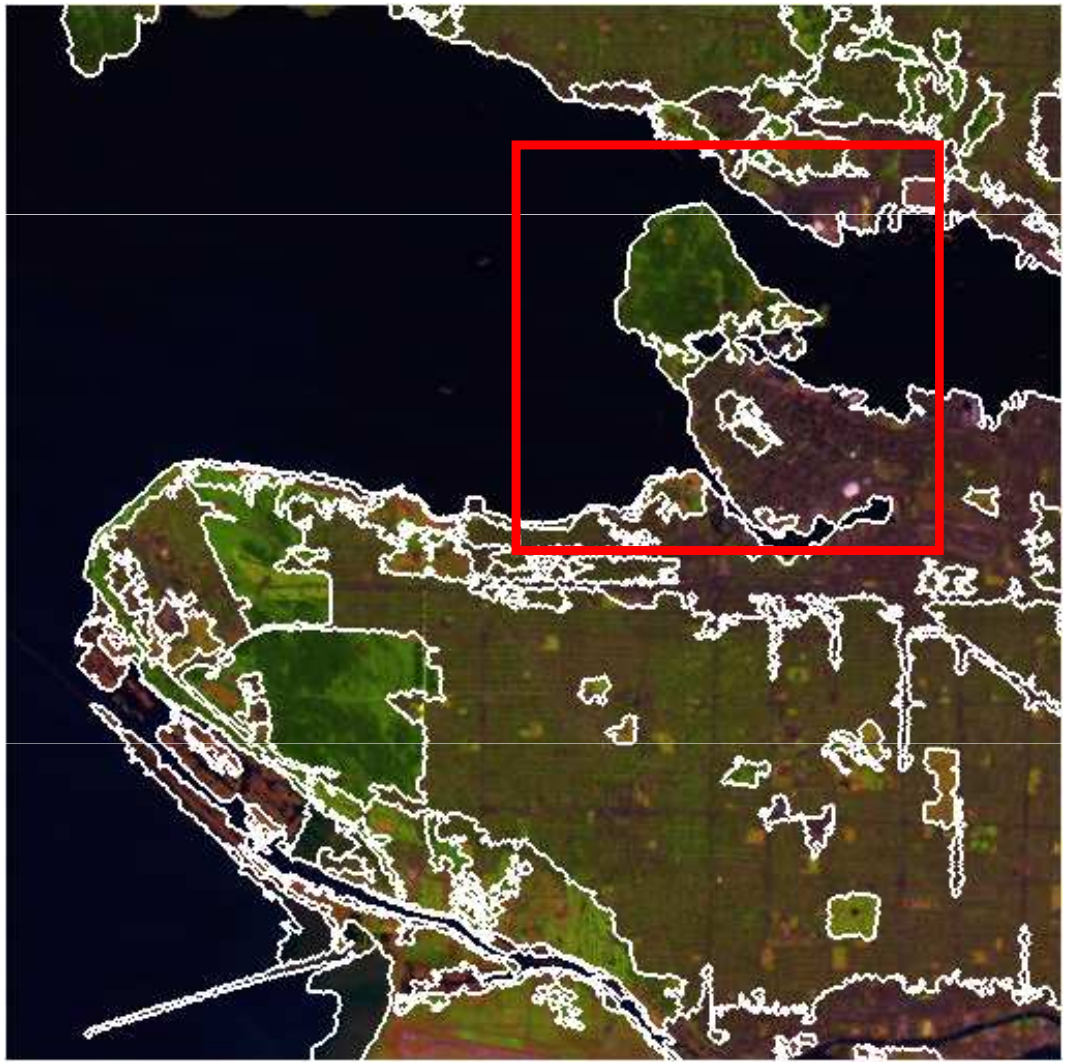
Attributed relational graphs

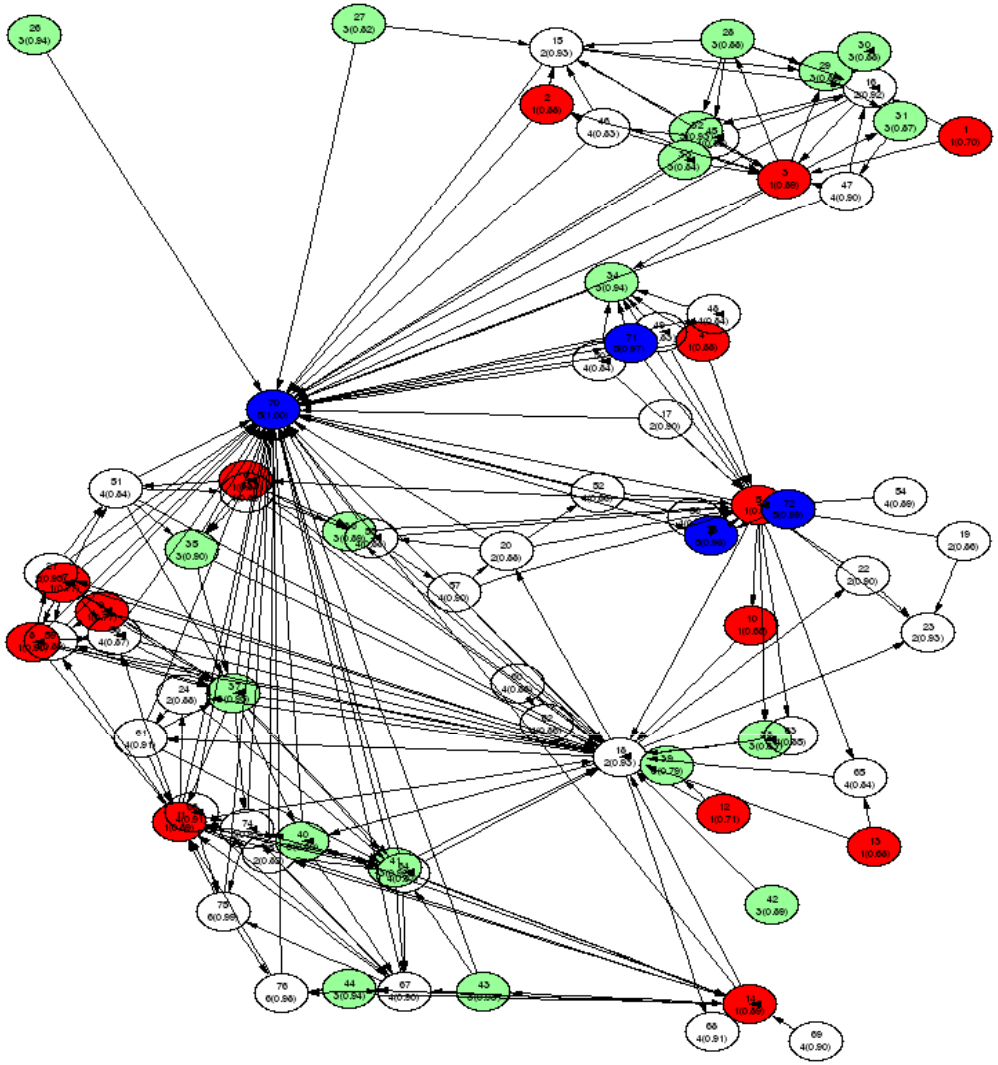




Search the graph for the whole scene …

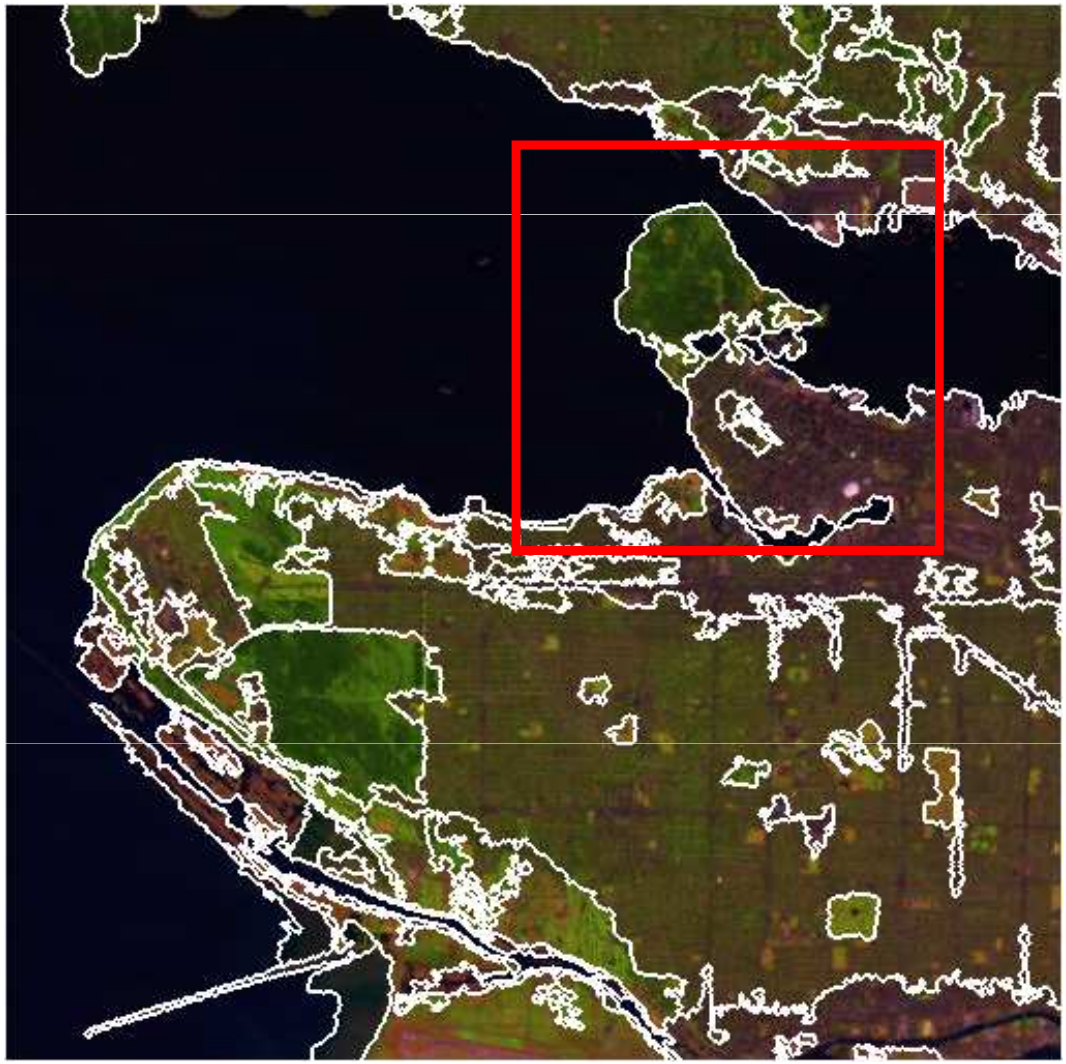
Attributed relational graphs

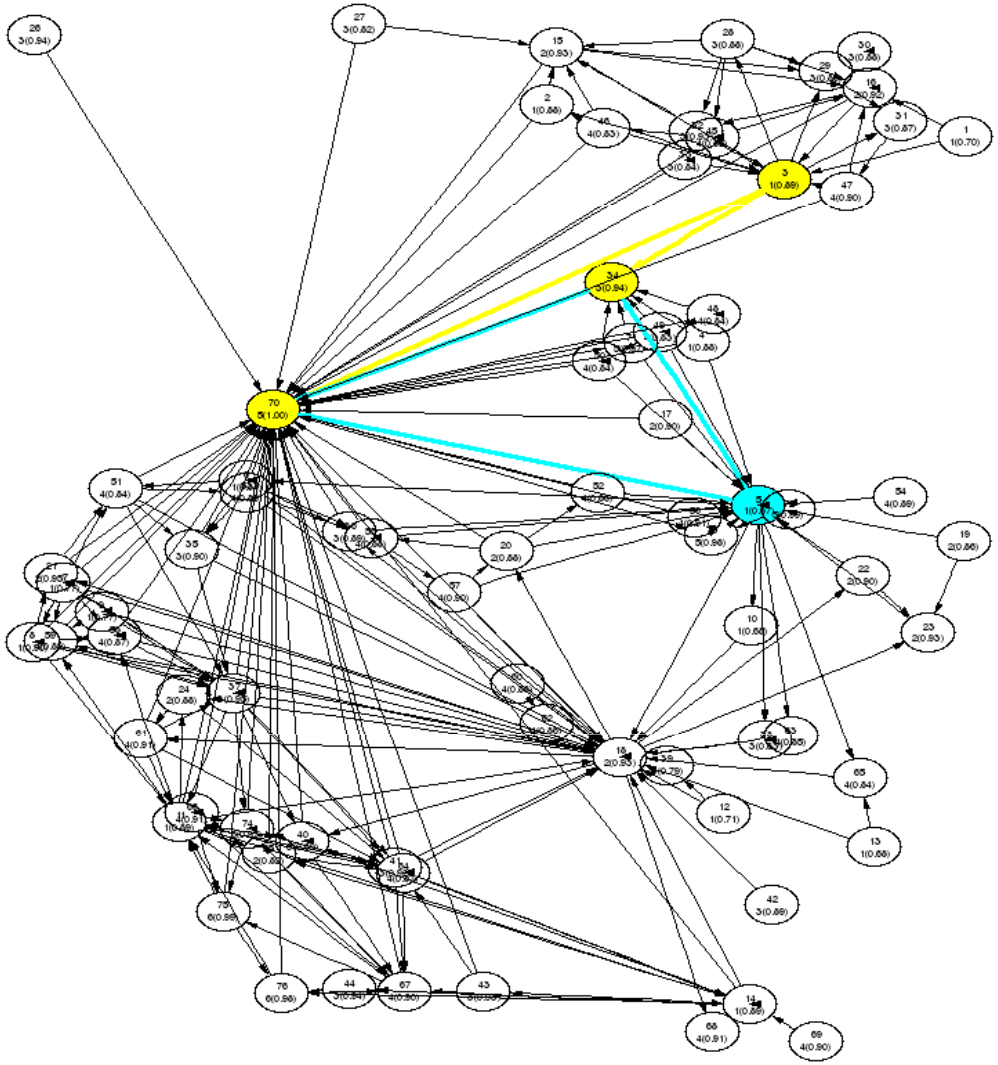




… by first finding the nodes (regions) with similar attributes

Attributed relational graphs





… and then finding the subgraphs with similar edges (relationships).