Computer Vision – T2.2a Segmentation by Clustering

MAP-I Doctoral Programme

Miguel Tavares Coimbra

MAP i DOCTORAL PROGRAMME IN COMPUTER SCIENCE

Outline

- Introduction
- Applications
- Simple clustering
- K-means clustering
- Graph-theoretic clustering

Acknowledgements: These slides follow Forsyth and Ponce's "Computer Vision: A Modern Approach", Chapter 14.

Topic: Introduction

- Introduction
- Applications
- Simple clustering
- K-means clustering
- Graph-theoretic clustering

What is ‘Segmentation’? (again?)

- Traditional definition:

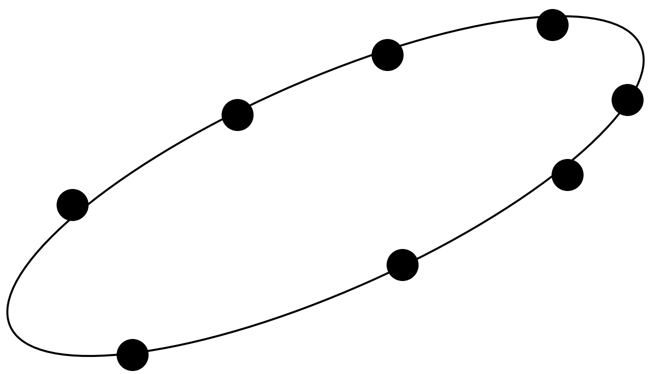
  “Separation of the image in different areas“

  - Decompose an image into “superpixels”.
  - Colour and texture coherence.

- Aren’t there other ways to look at the ‘Segmentation’ concept?

Other '*Segmentation*' problems

- Fitting lines to edge points.



We can't see this as 'separating an image in different areas'!

- Fitting a fundamental matrix to a set of feature points.

This one is complicated!
Check Forsyth and Ponce, chap.14

Segmentation as Clustering

- Tries to answer the question:

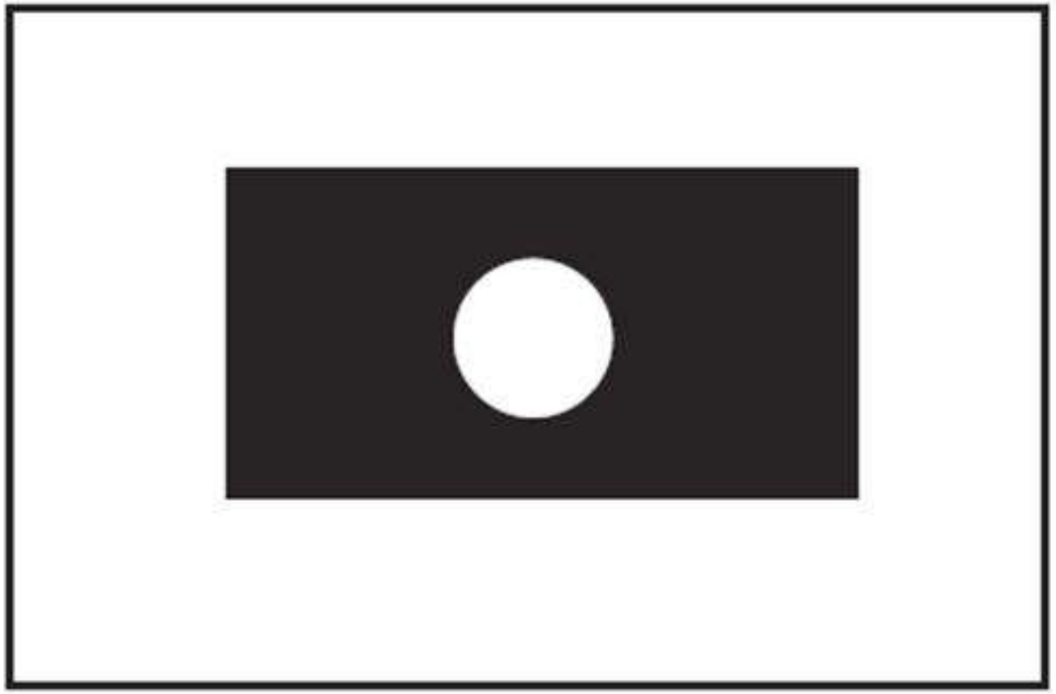
  "Which components of the data set naturally belong together?"
- Two approaches:
  - Partitioning
    - Decompose a large data set into pieces that are 'good' according to our model.
  - Grouping
    - Collect sets of data items that 'make sense' according to our model.

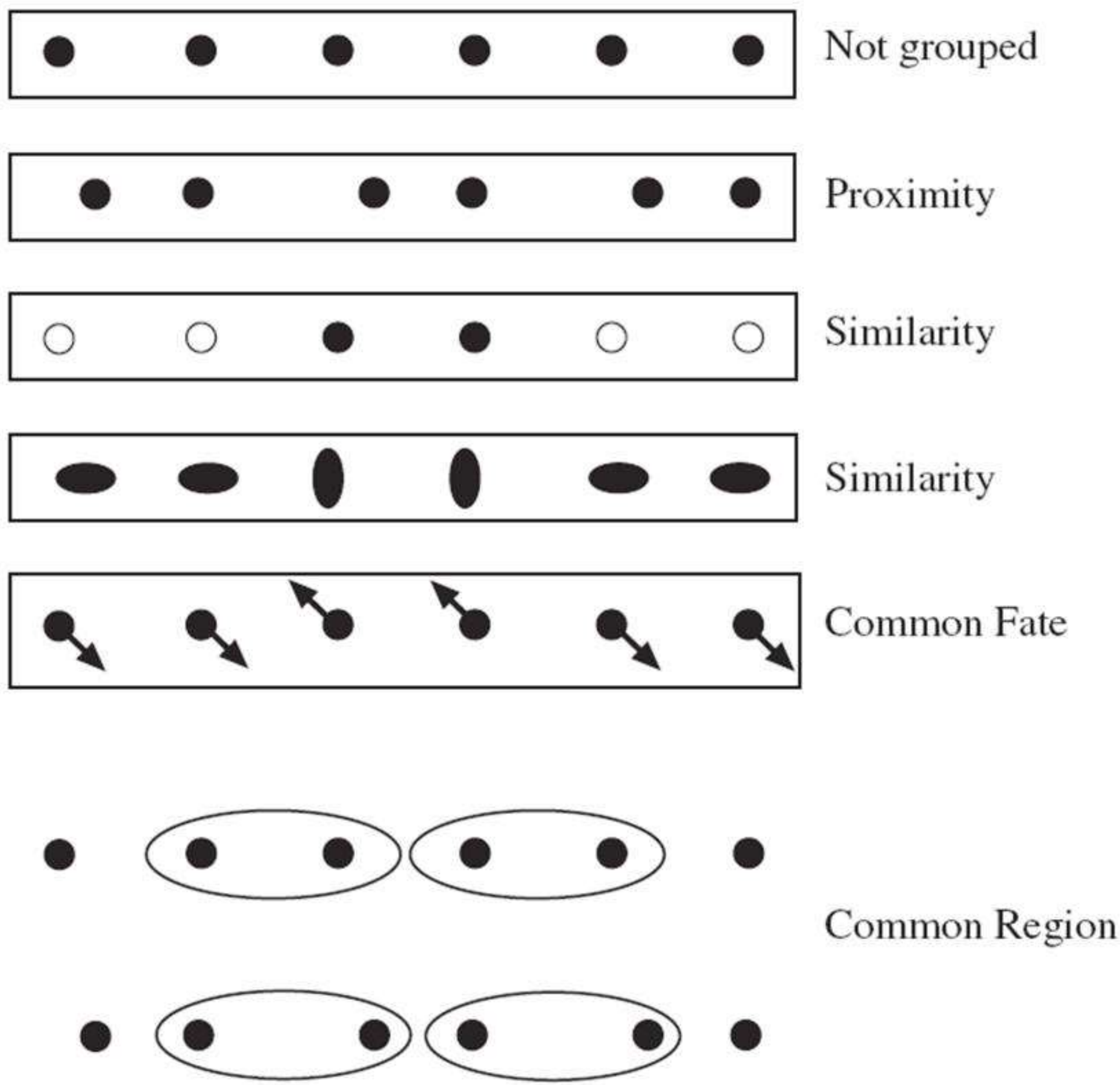
Human Clustering

- How de we humans *cluster* images?
  - Well… we don't really know...
- Gestalt school of psychologists
  - Attempts to study this problem.
  - Key ideas:
    - Context affects perception. So…
    - Responses to stimuli are not important.
    - Grouping is the key to understanding visual perception.

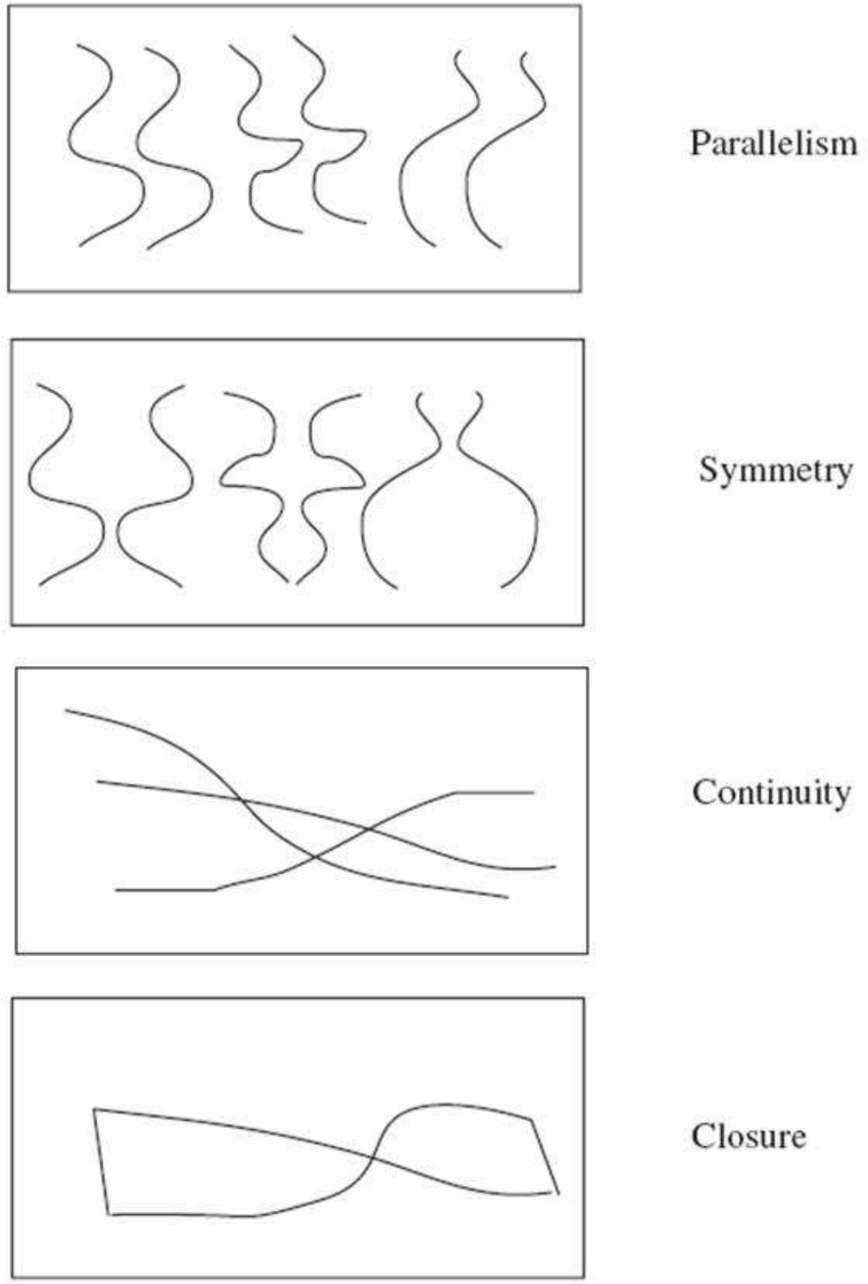
Figure-Ground ambiguity:

- Is this a white circular figure over a black ground?
- Or a black table with a white circular hole in it?





Examples of Gestalt factors that lead to grouping



Examples of Gestalt factors that lead to grouping

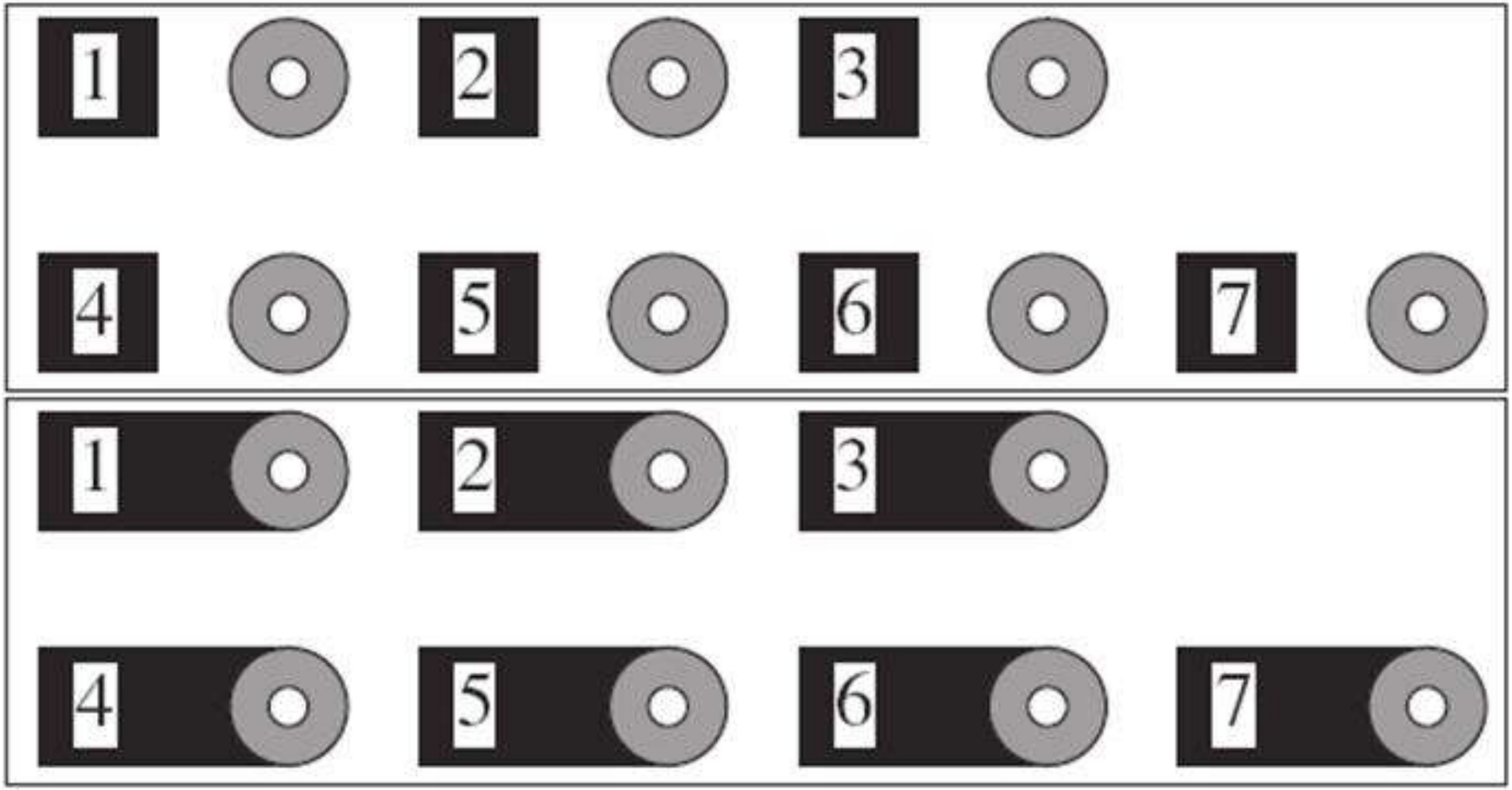


Figure 16.7. An example of grouping phenomena in real life. The buttons on an elevator in the computer science building at U.C. Berkeley used to be laid out as in the **top** figure. It was common to arrive at the wrong floor and discover that this was because you'd pressed the wrong button — the buttons are difficult to group unambiguously with the correct label, and it is easy to get the wrong grouping at a quick glance. A public-spirited individual filled in the gap between the numbers and the buttons, as in the **bottom** figure, and the confusion stopped because the proximity cue had been disambiguated.

Gestalt in Practice

- Rules function fairly well as explanations.
- However, they are insufficient to form an algorithm.
- So, how is Gestalt useful?
  - Gives us ‘hints’ on where to go.
  - Shatters the traditional definition of segmentation, clearly showing us that we need something better.
  - Context is vital! Grouping is vital!

Topic: Applications

- Introduction
- Applications
- Simple clustering
- K-means clustering
- Graph-theoretic clustering

Practical usefulness of *SbC*

- Sometimes simple algorithms are ‘good enough’.
- When?
  - Easy to tell what a useful decomposition is.
- Two examples:
  - Background subtraction
    - Anything that doesn’t look like a known background is interesting.
  - Shot boundary detection
    - Substantial changes in video are interesting.


Is there a person here? Where?

Background Subtraction

- What if I know this?

Background Subtraction

- Subtract!
- Limitations?

Background Subtraction

- Objective:
  - Separate the foreground objects from a static background.
- Large variety of methods:
  - Mean & Threshold [CD04]
  - Normalized Block Correlation [Mats00]
  - Temporal Derivative [Hari98]
  - Single Gaussian [Wren97]
  - Mixture of Gaussians [Grim98]

Where is this useful?

- If we have:
  - Static cameras
  - Large stable backgrounds
  - Low occlusion of foreground objects
- Examples:
  - Detecting parts on a conveyor belt
  - Counting cars in an overhead view of a road
  - Counting people in CCTV cameras

Eric Harris and
Dylan Klebold, in
the Columbine High
School Massacre
via CCTV cameras

The men alleged to be responsible for the 7 July attacks on London, captured on CCTV.

Modelling the background

- Photograph an 'empty' scene.
  - Sometimes it is difficult to obtain an 'empty scene'.
  - Does not handle changes in background (e.g. the *light switch* problem).
- More adaptive solutions:
  - Moving average of each pixel.
  - Many other possibilities:

    Toyama et al., "Wallflower: Principles and Practice of Background Maintenance", International Conference on Computer Vision, Vol. 1 (1999), 255.

Typical Algorithm

- Form a background estimate $B^{(0)}$. At each frame f:
  - Update the background estimate

$$B^{(n+1)} = \frac{w_a f^{(n)} + \sum_i w_i B^{(n-i)}}{w_c}$$

  - Subtract the background from the frame.

$$S^{(n)} = \left| F^{(n)} - B^{(n)} \right|$$

  - Threshold the resulting difference.

http://www.merl.com/projects/pedestrian/

Shot Boundary Detection

- Long sequences of video are composed of *shots*.
- Can we find the boundaries of these *shots* automatically?
  - Find frames in the video that are *significantly different* from the previous frame.
  - Distance measures! (lots of options here…)


Pan Right


HORNBACH
Pan Right

Sample Distance Measures

- Frame differencing
  - Pixel-by-pixel difference.
- Histogram-based
  - Differences in colour histograms.
- Block-based comparison
  - Divide image in blocks and compare blocks directly.
- Edge differencing
  - Compare edge maps.

Many other exist!

Topic: Simple clustering

- Introduction
- Applications
- Simple clustering
- K-means clustering
- Graph-theoretic clustering

What do we mean by ‘clustering’?

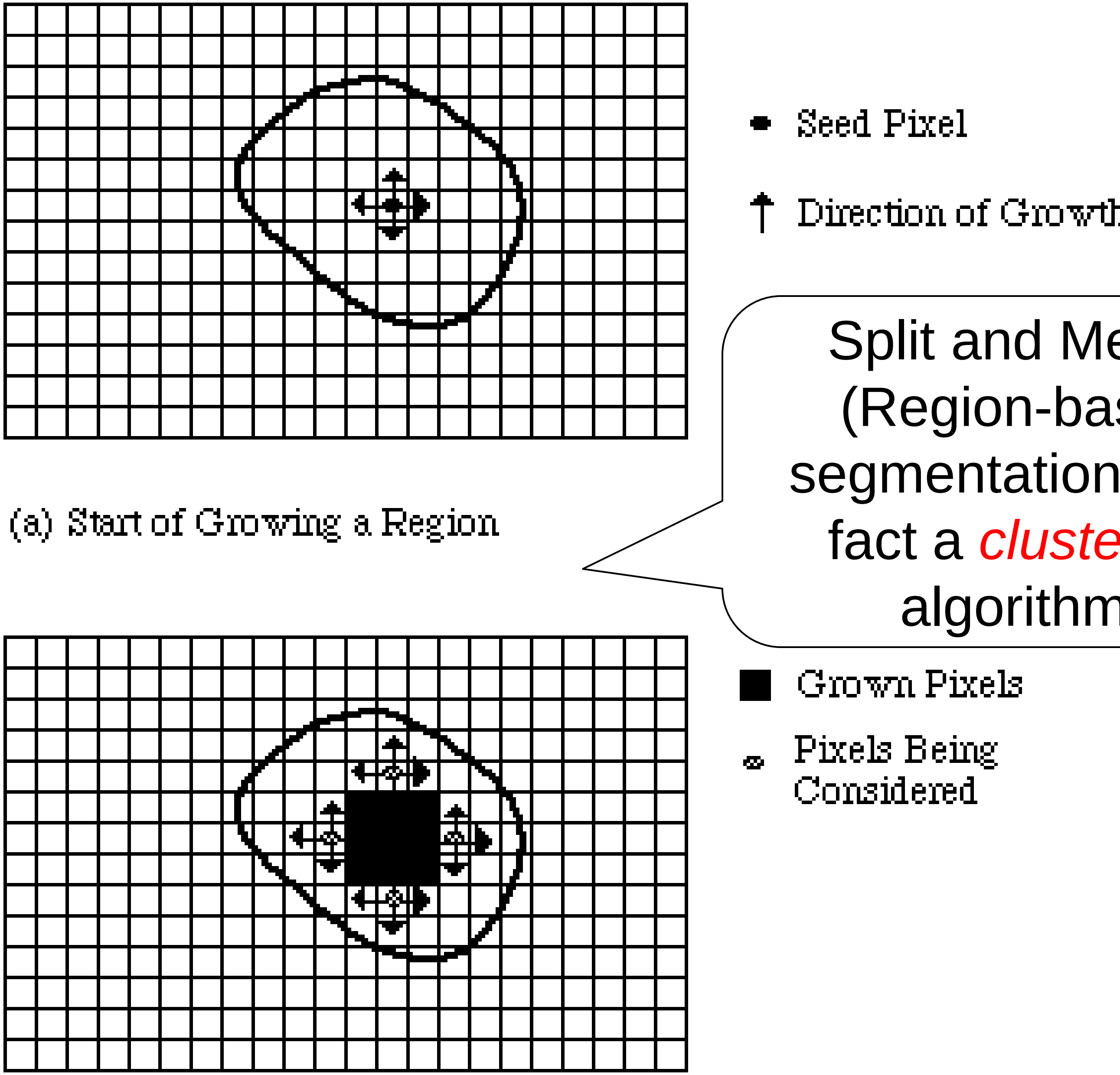
- “Clustering is a process whereby a data set is replaced by **clusters**, which are collections of data points that belong together”

  Forsyth and Ponce, “Computer Vision: A modern approach”

- Why do points “belong together”?
  - Same colour.
  - Same texture.
  - Same… something!

Simple clustering

- Two natural types of clustering:
  - Divisive clustering
    - Entire data set is regarded as a cluster.
    - Clusters are recursively split.
  - Agglomerative clustering
    - Each data item is a cluster.
    - Clusters are recursively merged.
- Where have I seen this before?



(b) Growing Process After a Few Iterations

Split and Merge (Region-based segmentation) is in fact a *clustering* algorithm.

Generic simple clustering algorithms

- Divisive Clustering
  - Construct a single cluster containing all points
  - While the clustering is not satisfactory
    - Split the cluster that yields the two components with the largest inter-cluster distance
  - end
- Agglomerative Clustering
  - Make each point a separate cluster
  - While the clustering is not satisfactory
    - Merge the two clusters with smallest inter-cluster distance
  - end

Which inter-cluster distance?

What does this mean?

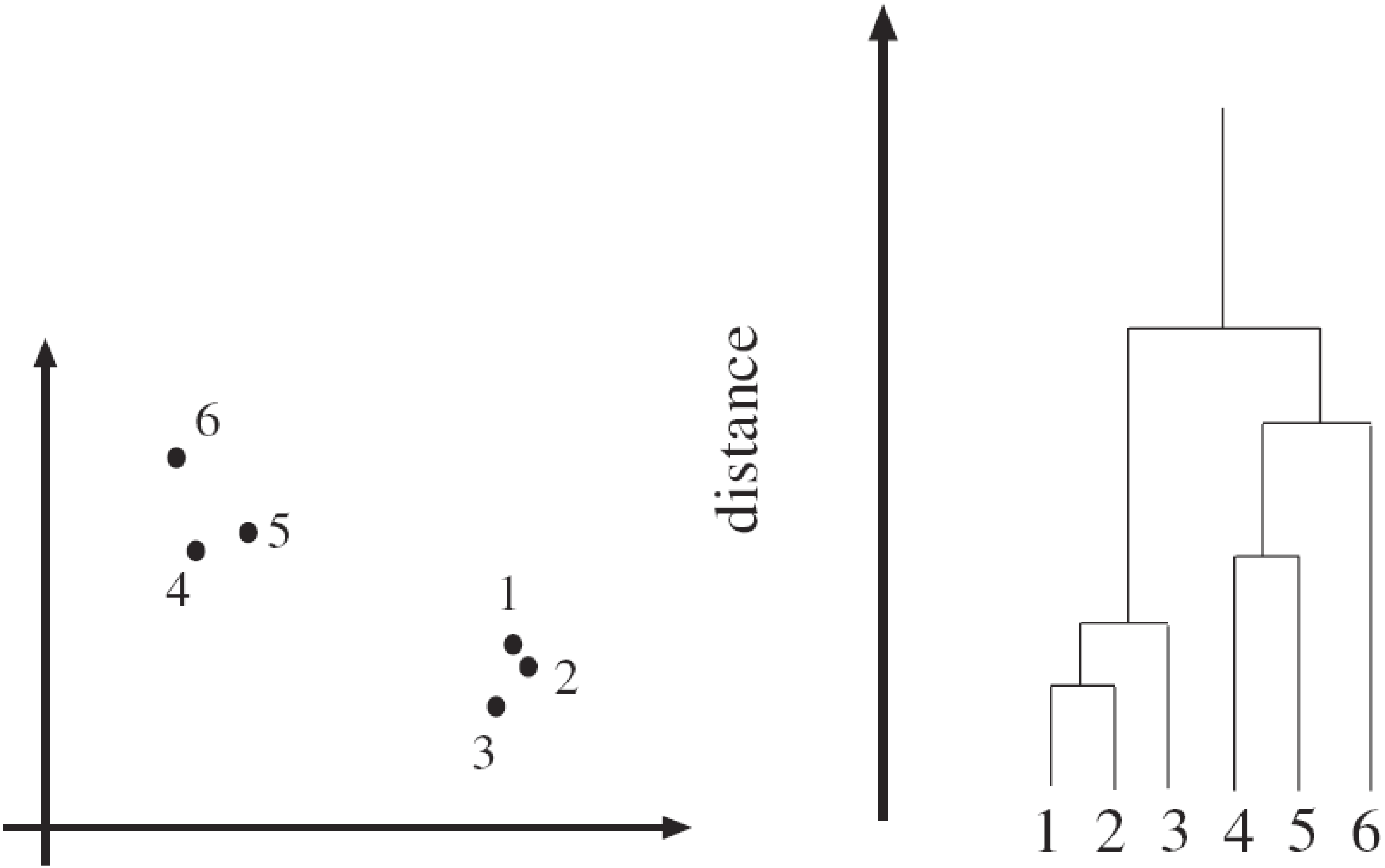


Figure 16.12. Left, a data set; right, a dendrogram obtained by agglomerative clustering using single link clustering. If one selects a particular value of distance, then a horizontal line at that distance will split the dendrogram into clusters. This representation makes it possible to guess how many clusters there are, and to get some insight into how good the clusters are.

Simple clustering with images

- Some specific problems arise:
  - Lots of pixels! Graphical representations are harder to read.
  - Segmentation: It is desirable that certain objects are connected. How to enforce this?
  - When do we stop splitting/merging process?
- Complex situations require more complex clustering solutions!

Topic: K-means clustering

- Introduction
- Applications
- Simple clustering
- K-means clustering
- Graph-theoretic clustering

Objective function

- What if we know that there are *k* clusters in the image?
- We can define an *objective function*!
  - Expresses how good my representation is.
- We can now build an algorithm to obtain the *best* representation.

Caution! “*Best*” given my objective function!

K-means Clustering

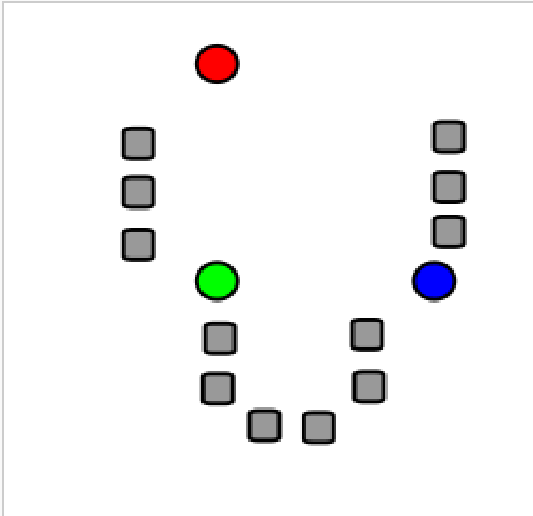
- Assume:
  - We have k clusters.
  - Each cluster i has a centre c_i.
  - Element j to be clustered is described by a feature vector x_j.
- Our objective function is thus:

What does this mean?

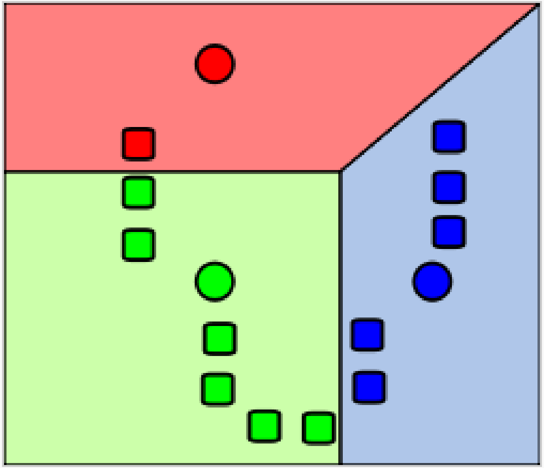
$$\Phi(clusters, data) = \sum_{i \in clusters} \left\{ \sum_{j \in cluster(i)} (x_j - c_i)^T (x_j - c_i) \right\}$$

Iteration step

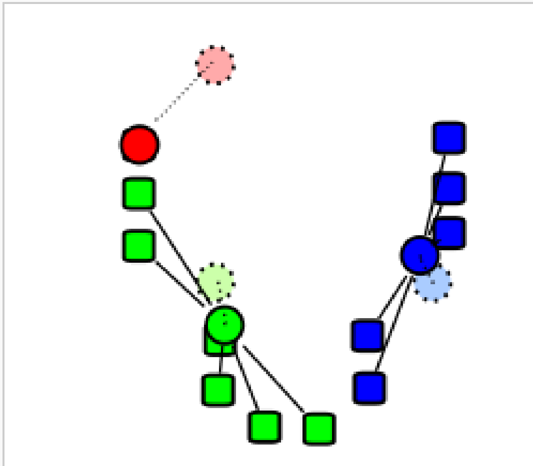
- Too many possible allocations of points to clusters to search this space for a minimum.
- Iterate!
  - Assume cluster centres are known and allocate each point to the closest cluster centre.
  - Assume the allocation is known and choose a new set of cluster centres. Each centre is the mean of the points allocated to that cluster.



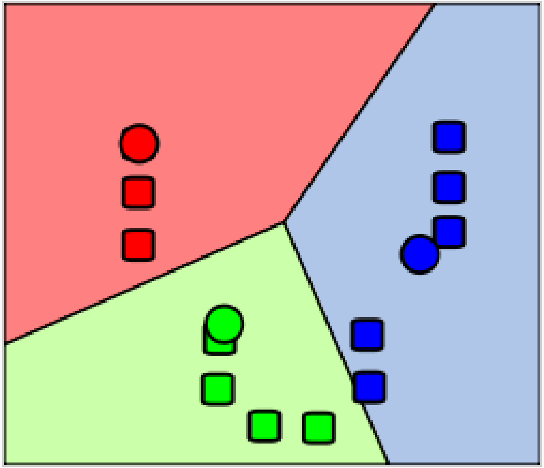
Shows the initial randomized centroids and a number of points.



Points are associated with the nearest centroid.



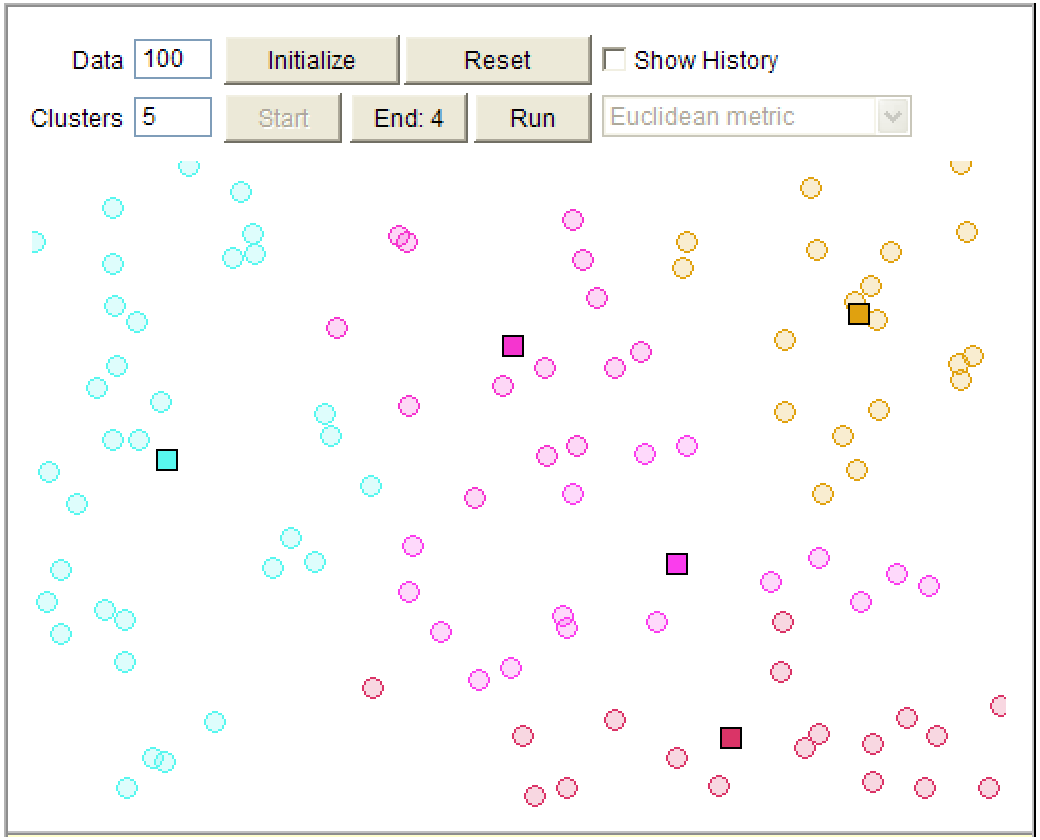
Now the centroids are moved to the center of their respective clusters.



Steps 2 & 3 are repeated until a suitable level of convergence has been reached.

[Wikipedia]

Interactive Java Tutorial



http://home.dei.polimi.it/matteucc/Clustering/tutorial_html/AppletKM.html

Topic: Graph-theoretic clustering

- Introduction
- Applications
- Simple clustering
- K-means clustering
- Graph-theoretic clustering

Using graphs

- Clustering can be seen as a problem of "*cutting graphs into good pieces*".
- Data Items
  - Vertex in a weighted graph.
  - Weights are large if elements are similar.
- Cut edges
  - Cut edges with small weights.
  - Keep connected components with large interior weights.

Regions!

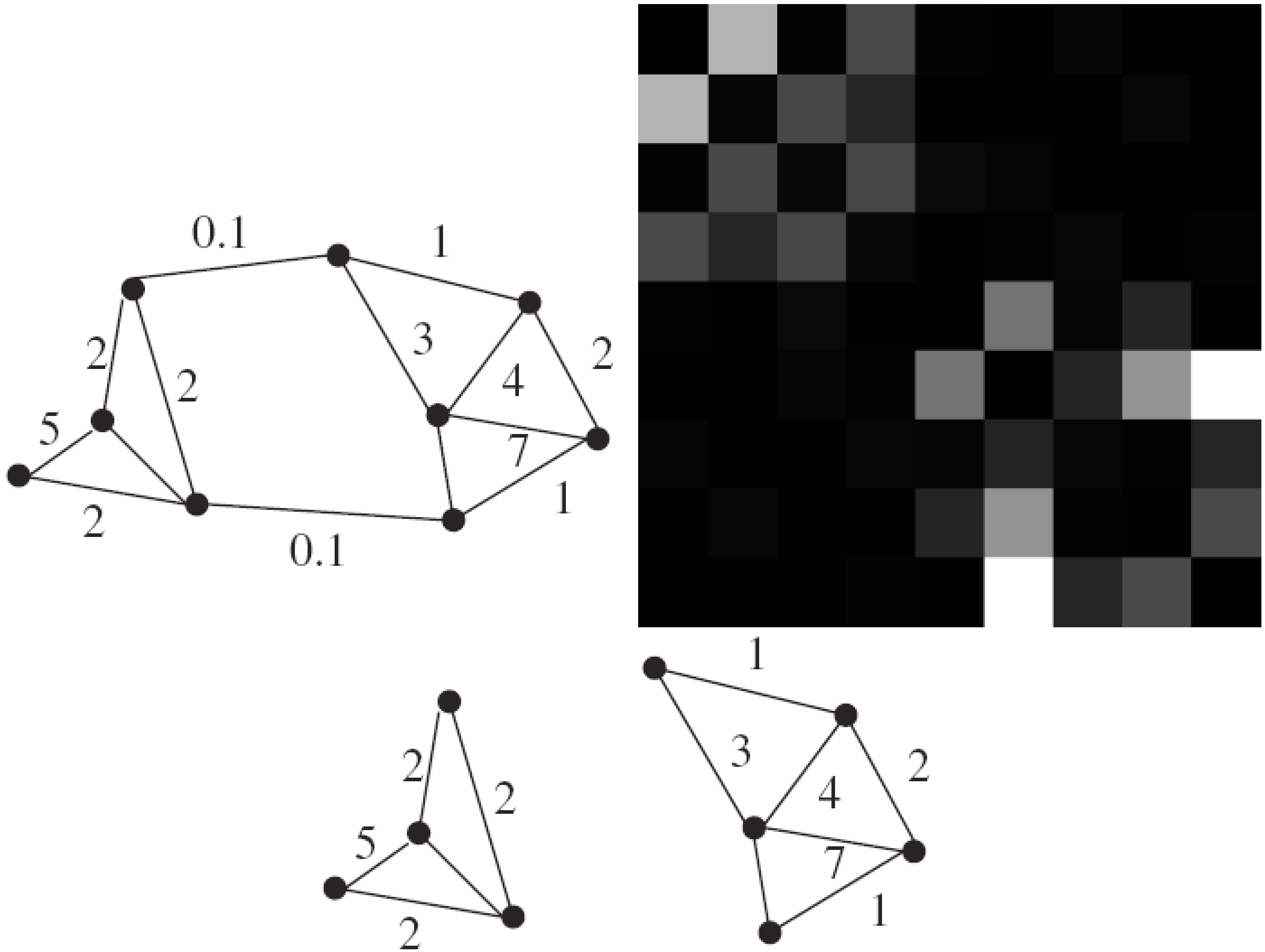


Figure 16.16. On the **top left**, a drawing of an undirected weighted graph; on the **top right**, the weight matrix associated with that graph. Larger values are lighter. By associating the vertices with rows (and columns) in a different order, the matrix can be shuffled. We have chosen the ordering to show the matrix in a form that emphasizes the fact that it is very largely block-diagonal. The figure on the **bottom** shows a cut of that graph that decomposes the graph into two tightly linked components. This cut decomposes the graph's matrix into the two main blocks on the diagonal.

Graphs and Clustering

- Associate each element to be clustered with a vertex on a graph.
- Construct an edge from every element to every other.
- Associate a weight with each edge based on a similarity measure.
- Cut the edges in the graph to form a good set of connected components.

Weight Matrices

- Typically look like block diagonal matrices.
- Why?
  - Interclusters similarities are strong.
  - Intracluster similarities are weak.
- Split a matrix into smaller matrices, each of which is a block.
- Define *Affinity Measures*.

More on this

- Affinity measures
  - Affinity by Distance
  - Affinity by Intensity
  - Affinity by Colour
  - Affinity by Texture
- Popular method: Normalized cuts

Jianbo Shi and Jitendra Malik, “Normalized Cuts and Image Segmentation”, IEEE Transactions on Pattern Analysis And Machine Intelligence, Vol. 22, No. 8, August 2000

Want to know more?
Check out: Forsyth and Ponce, Section 14.5

Resources

- Forsyth and Ponce, Chapter 14
- Jianbo Shi and Jitendra Malik, "Normalized Cuts and Image Segmentation", IEEE Transactions on Pattern Analysis And Machine Intelligence, Vol. 22, No. 8, August 2000