Segmentation, Grouping, and Fitting

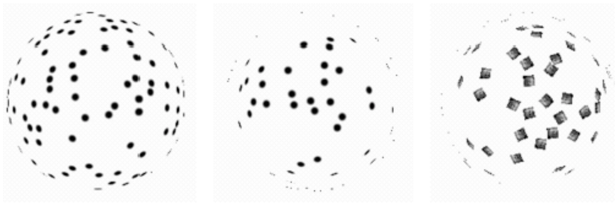
- Collect together tokens that belong together
- Gives a compact representation from an image/motion sequence/set of tokens that can be significantly easier to deal with
- What is the “right” group is often dependent on the application
- Broad theory is not known at present (and may not exist)
- These are general concepts--apply to many things, not just breaking images into regions of the same color.

Segmentation, Grouping, and Fitting

- Terminology varies and the usage and the meaning of segmentation, grouping, and fitting overlap. Somewhat common usage:
  - Grouping (or clustering) is quite general sometimes suggest a relatively high level (group the black and white halves of a penguin together).
  - Segmentation is suggestive of the grouping is done at a low level and is quite spatially (or temporally coherent) given regions in time or space.
  - Fitting suggests the focus is on a model associated with tokens. Issues:
    - which model?
    - which token goes to which element in the model (correspondence)?
    - how many elements in the model (how complex should it be)?

General ideas

- Tokens
  - whatever we need to group (e.g. pixels, points, surface elements)
- Top down segmentation
  - tokens belong together because they lie on the same object
- Bottom up segmentation
  - tokens belong together because they are locally coherent
- These two are not mutually exclusive

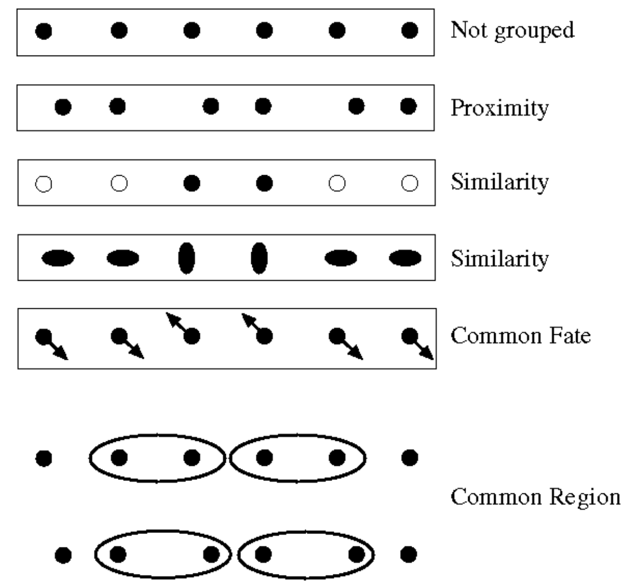


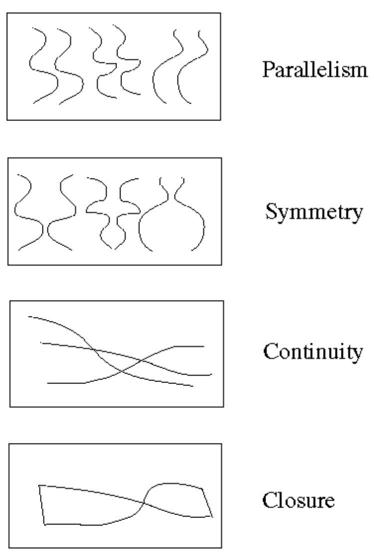
Why do these tokens belong together?

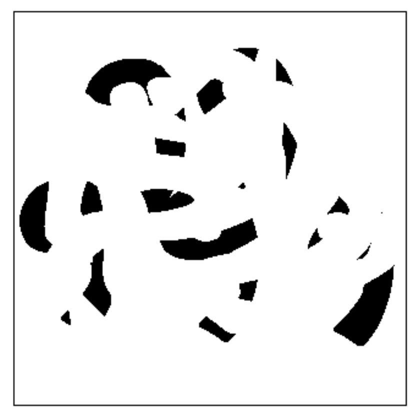
Basic ideas of grouping in humans

- Figure-ground discrimination
  - grouping can be seen in terms of allocating some elements to a figure, some to ground (impoverished theory)

- Gestalt factors
  - Elements in a collection of elements can have properties that result from relationships
  - A series of factors affect whether elements should be grouped together







*

Segmentation as clustering

- Cluster together (pixels, tokens, etc.) that belong together
- We assume that we can compute how close tokens are, or how close a token is to cluster.

Why is clustering hard?

Main reason

- The number of possible clusterings is exponential in the number of data points

Other important issues

- The number of clusters is usually **not** known
- A good distance function between points may not be known
- A good model explaining the existence of clusters is usually not available.
- High dimensionality (discussed more later)

Data Representation

- Most common is an N dimensional “feature” vector.
- Most common distance is Euclidian distance.
- Be careful with scaling and units!
  - Many algorithms are **not** invariant to changing the scale of one vector component relative to the others

K-Means

- Choose a fixed number of clusters (“K”)
- Choose cluster centers (**means**) and point-cluster allocations (membership) to minimize the error

$$\sum_{i \in \text{clusters}} \left\{ \sum_{j \in \text{elements of i'th cluster}} \left\| x_j - \mu_i \right\|^2 \right\}$$

- **x**’s could be any set of features for which we can compute a distance (careful with scaling)

K-Means

- Want to minimize

$$\sum_{i \in \text{clusters}} \left\{ \sum_{j \in \text{elements of i'th cluster}} \left\| x_j - \mu_i \right\|^2 \right\}$$

- **Cannot** do this optimization by search, because there are too many possible allocations.
- Standard difficulty which we handle with an iterative process (chicken and egg)

K-Means algorithm (intuition)

- If we know the cluster centers, the best cluster for each point is easy to compute
  - Just compute the distance to each to find the closest
- If we know the best cluster for each point, the cluster centers are also easy to compute
  - Just average the points in each cluster
- Algorithm
  - 1) Guess one of the two.
  - 2) Alternatively re-compute the values for each

K-means flow chart

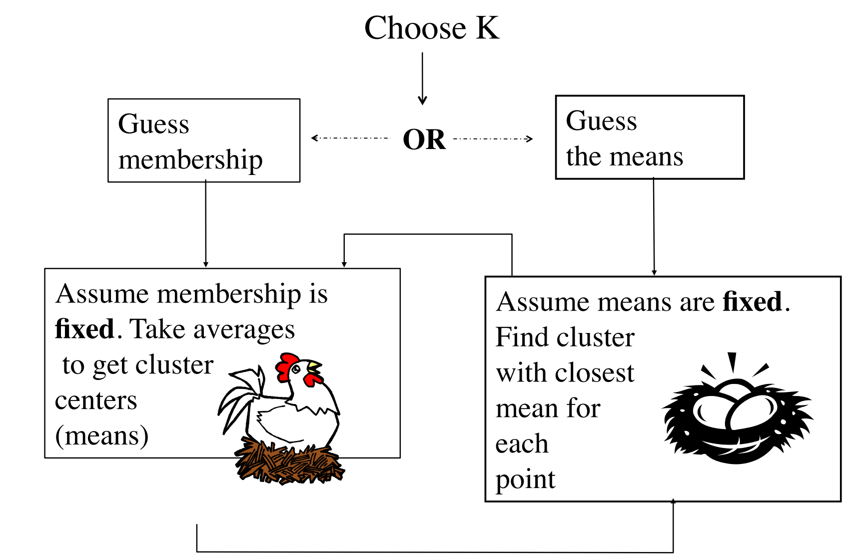
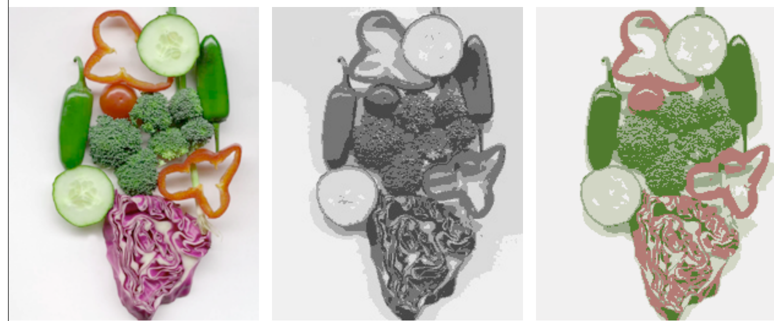


Image | Clusters on intensity | Clusters on color

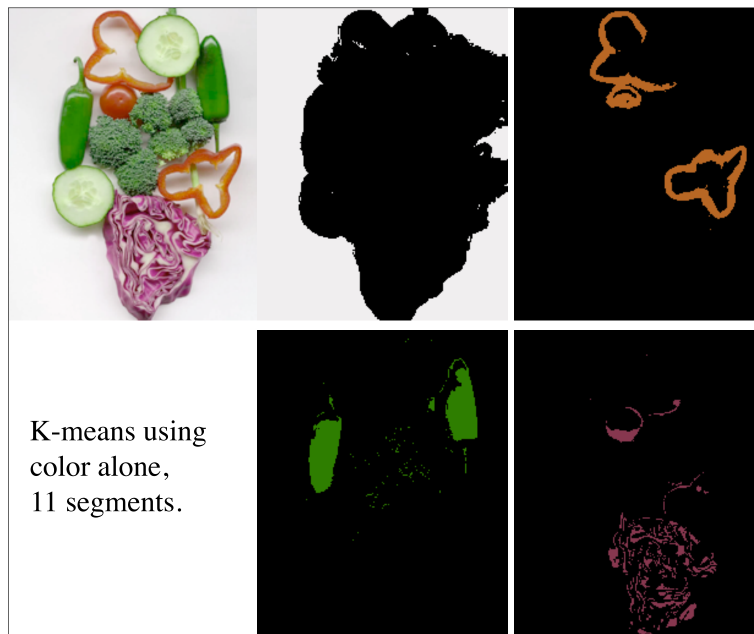


K-means clustering using intensity alone and color alone
(Assuming 5 segments, i.e. k=5)

Image

Clusters on color

K-means using color alone, 11 segments



K-means using color alone, 11 segments.