Segmentation and Clustering

COS 429: Computer Vision

Acknowledgments: T. Funkhouser, K. Grauman, S. Lazebnik, S. Seitz, X. Ren, S. Rusinkiewicz
Many announcements

- HW2 graded and will be returned today (a big thank you to Kyle and Riley!!)
- HW3 released after fall break
- Remember: work with the same partner on at most 2 assignments
- No office hours during fall break
- Final project: fall break is a great time to start!
  - Please start coming to office hours to brainstorm, narrow down a topic, ask for papers to read, discuss the scope, …
  - [http://www.cs.princeton.edu/courses/archive/fall17/cos429/project.html](http://www.cs.princeton.edu/courses/archive/fall17/cos429/project.html)
- Attendance after fall break: there will be questions on hw4 that *inspire* attendance (details later)
- Midterm
Segmentation and Clustering

- **Segmentation:**
  Divide image into regions of similar contents

- **Clustering:**
  Aggregate pixels into regions of similar contents
Goal

- Separate image into coherent “regions”

Berkeley segmentation database:
http://www.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/segbench/
But Wait!

- We speak of “segmenting” foreground from background
- Segmenting out skin colors
- Segmenting out the moving person
- How do these relate to “similar regions”?
Segmentation and Clustering Applications

“Intelligent scissors”

Foreground / background segmentation

Finding skin-colored regions

Finding the moving objects

Finding the cars in a video sequence
Segmentation and Clustering Applications

“Intelligent scissors”

Foreground / background segmentation

Finding skin-colored regions

Finding the moving objects

Finding the cars in a video sequence

Statistics

Templates
Questions

• What is coherent?
  – Similar color?
  – Similar texture?
  – Spatial proximity?

• What kinds of regions?
  – Nearly convex?
  – Smooth boundaries?
  – Nearly equal sizes?
  – What granularity?
Gestalt Grouping Cues

- Not grouped
- Proximity
- Similarity
- Similarity
- Common Fate
- Common Region
- Parallelism
- Symmetry
- Continuity
- Closure
Segmentation and Clustering

• Defining regions
  – Should they be compact? Smooth boundary?

• Defining similarity
  – Color, texture, motion, …

• Defining similarity of regions
  – Minimum distance, mean, maximum
Let’s start simple
Clustering Based on Color

• Let’s make a few concrete choices:
  – Arbitrary regions
  – Similarity based on color only
  – Similarity of regions = distance between mean colors
Divisive Clustering

- Start with whole image in one cluster
- Iterate:
  - Find cluster with largest intra-cluster variation
  - Split into two pieces that yield largest inter-cluster distance
- Stopping criteria?
Divisive Clustering

• Start with whole image in one cluster

• Iterate:
  – Find cluster with largest intra-cluster variation
  – Split into two pieces that yield largest inter-cluster distance

• Stopping criteria?
Divisive Clustering

- Start with whole image in one cluster
- Iterate:
  - Find cluster with largest intra-cluster variation
  - Split into two pieces that yield largest inter-cluster distance
- Stopping criteria?
Hierarchical Clustering

• Start with each pixel in its own cluster
• Iterate:
  – Find pair of clusters with smallest inter-cluster distance
  – Merge
• Stopping criteria?
Hierarchical Clustering

- Start with each pixel in its own cluster
- Iterate:
  - Find pair of clusters with smallest inter-cluster distance
  - Merge
- Stopping criteria?
Hierarchical Clustering

- Start with each pixel in its own cluster
- Iterate:
  - Find pair of clusters with smallest inter-cluster distance
  - Merge
- Stopping criteria?
Hierarchical Clustering

- Start with each pixel in its own cluster
- Iterate:
  - Find pair of clusters with smallest inter-cluster distance
  - Merge
- Stopping criteria?
Hierarchical Clustering

- Start with each pixel in its own cluster
- Iterate:
  - Find pair of clusters with smallest inter-cluster distance
  - Merge
- Stopping criteria?
Hierarchical Clustering

- Start with each pixel in its own cluster
- Iterate:
  - Find pair of clusters with smallest inter-cluster distance
  - Merge
- Stopping criteria?
Hierarchical Clustering

• Start with each pixel in its own cluster
• Iterate:
  – Find pair of clusters with smallest inter-cluster distance
  – Merge
• Stopping criteria?
Hierarchical Clustering

- Start with each pixel in its own cluster
- Iterate:
  - Find pair of clusters with smallest inter-cluster distance
  - Merge
- Stopping criteria?
Hierarchical Clustering

• Conservative stopping criteria yields “superpixels”, which can be used as starting point for more complex algorithms.
Problems with These Algorithms

- **Greedy**
  - Decisions made early in process dictate final result
- **Making “good” early decisions is hard/expensive**
  - Many possibilities at each iteration
  - Computing “good” merge or split is expensive
- **Heuristics to speed things up:**
  - For agglomerative clustering, approximate each cluster by average for distance computations
  - For divisive clustering, use summary (histogram) of a region to compute split
More advanced: iterative, better metrics, …
Instead of merging or splitting, start out with the clusters and move them around

1. Pick number of clusters $k$
2. Randomly scatter $k$ “cluster centers” in color space
3. Repeat:
   a. Assign each data point to its closest cluster center
   b. Move each cluster center to the mean of the points assigned to it
Results of Clustering

Original Image  \hspace{1cm} k\text{-means, } k=5 \hspace{1cm} k\text{-means, } k=11
Results of Clustering

Sample clusters with $k$-means clustering based on color
Other Distance Measures

• Suppose we want to have compact regions
• New feature space: 5D
  (2 spatial coordinates, 3 color components)
• Points close in this space are close both in color and in actual proximity
Results of Clustering

Sample clusters with $k$-means clustering based on color and distance
Other Distance Measures

• Problem with simple Euclidean distance: what if coordinates range from 0-1000 but colors only range from 0-255?
  – Depending on how things are scaled, gives different weight to different kinds of data

• Weighted Euclidean distance: adjust weights to emphasize different dimensions

\[ \text{dist}(x, y)^2 = \sum_i c_i (x_i - y_i)^2 \]
Mahalanobis Distance

- Automatically assign weights based on actual variation in the data

\[ \text{dist}(x, y)^2 = (x - y)^T C (x - y) \]

where \( C \) is covariance matrix of all points

- Gives each dimension “equal” weight

- Also accounts for correlations between different dimensions
$k$-means Pros and Cons?
$k$-means Pros and Cons

- **Pros**
  - Very simple method

- **Cons**
  - Need to pick $k$
  - Converges to a local minimum
  - Sensitive to initialization
  - Sensitive to outliers
  - Only finds “spherical” clusters

Sensitive to outliers

Spherical clusters
Mean shift
Mean Shift Clustering

- Seek *modes* (peaks) of density in feature space

Image

Feature space (color values)
Mean Shift Clustering

• Algorithm:
  – Initialize windows at individual feature points
  – Perform mean shift for each window until convergence
  – Merge windows that end up near the same “peak” or mode
Mean Shift Clustering
Mean Shift Clustering
Mean Shift Clustering

- Search window
- Center of mass
- Mean Shift vector

Ukrainitz & Sarel
Mean Shift Clustering
Mean Shift Clustering
Mean Shift Clustering
Mean Shift Clustering

Search window

Center of mass

Ukrainitz & Sarel
Mean Shift Clustering

- Cluster data points in the attraction basin of a mode
  - Separate segment for each mode
  - Assign points to segments based on which mode is at the end of their mean shift trajectories
Mean Shift Clustering

Density

Segments

Modes

Ukrainitz & Sarel
Mean Shift Results

http://www.caip.rutgers.edu/~comanici/MSPAMI/msPamiResults.html
Mean Shift Results
Mean Shift Pros and Cons

• Pros
  – Finds variable number of modes
  – Does not assume spherical clusters
  – Just a single parameter (window size)
  – Robust to outliers

• Cons
  – Output depends on window size
  – Computationally expensive
  – Does not scale well with dimension of feature space
Graph-based algorithms
Segmentation Based on Graph Cuts

- Create weighted graph:
  - Nodes = pixels in image
  - Edge between each pair of nodes
  - Edge weight = similarity (intensity, color, texture, etc.)
Segmentation Based on Graph Cuts

- Partition into disconnected segments
- Easiest to break links that have low cost (low similarity)
  - similar pixels should be in the same segments
  - dissimilar pixels should be in different segments
Cuts in a Graph

- **Link Cut**
  - set of links whose removal makes a graph disconnected
  - cost = sum of costs of all edges

- **Min-cut**
  - fast (polynomial-time) algorithm
  - gives segmentation
But Min Cut Is Not Always the Best Cut...
Cuts in a Graph

- Normalized Cut
  - removes penalty for large segments
  
  \[ N\text{cut}(A, B) = \frac{\text{cut}(A, B)}{\text{volume}(A)} + \frac{\text{cut}(A, B)}{\text{volume}(B)} \]

  - \( \text{volume}(A) = \text{sum of costs of all edges that touch } A \)
  - no fast exact algorithms…
Cuts in a Graph

\[ N_{\text{cut}}(A, B) = \frac{\text{cut}(A, B)}{\text{volume}(A)} + \frac{\text{cut}(A, B)}{\text{volume}(B)} \]
Interpretation as a Dynamical System

Treat the links as springs and shake the system

- elasticity proportional to cost
- vibration “modes” correspond to segments
  - can compute these by solving a generalized eigenvector problem
  - for more details, see

Efficient graph-based segmentation

http://cs.brown.edu/~pff/segment/

Beautiful code available
Boundary detection
Designing Grouping Features

Low-level cues
- Brightness similarity
- Color similarity
- Texture similarity

Mid-level cues
- Contour continuity
- Convexity
- Parallelism
- Symmetry

High-level cues
- Object knowledge
- Scene structure
Brightness and Color Contrast

- Color (e.g., 1976 CIE L*a*b* colorspace)
- Brightness Gradient $BG(x,y,r,\theta)$
  $\chi^2$ difference in $L^*$ distribution
- Color Gradient $CG(x,y,r,\theta)$
  $\chi^2$ difference in $a^*$ and $b^*$ distributions

\[
\chi^2(g, h) = \frac{1}{2} \sum_i \frac{(g_i - h_i)^2}{g_i + h_i}
\]
Texture Contrast

- Texture Gradient $TG(x,y,r,\theta)$
  - $\chi^2$ difference of texton histograms
  - Textons are vector-quantized filter outputs (through k-means)
Boundary Classification

non-boundaries

boundaries

I

BG

CG

TG
W(p1,p2) >> W(p1,p3) as p1 and p2 are more likely to belong to the same region than are p1 and p3, which are separated by a strong boundary.
Combining Cues

Evaluation
Option 1: human agreement

Berkeley segmentation dataset

A Measure for Objective Evaluation of Image Segmentation Algorithms
R. Unnikrishnan C. Pantofaru M. Hebert
CVPR 2005
Option 2: usefulness for end goal

https://pdollar.wordpress.com/2013/12/22/generating-object-proposals/
Selective search for object recognition

Segmentation as Selective Search for Object Recognition.
Koen E. A. van de Sande, Jasper R. R. Uijlings, Theo Gevers, Arnold W. M. Smeulders
ICCV 2011

Figure 2: Two examples of our selective search showing the necessity of different scales. On the left we find many objects at different scales. On the right we necessarily find the objects at different scales as the girl is contained by the tv.
However, worth noting:

Object-Proposal Evaluation Protocol is 'Gameable'
Neelima Chavali, Harsh Agrawal, Aroma Mahendru, Dhruv Batra
CVPR 2016

Figure 1: Method 1 visually seems to recall more categories such as plates, glasses that Method 2 missed. Despite that, the computed recall for Method 2 is higher because it recalled all instances of PASCAL categories that were present in the ground truth. Note that the number of proposals generated by both methods is equal in this figure.

https://filebox.ece.vt.edu/~aroma/web/object-proposals.html
Summary

• **Segmentation:**
  – Partitioning image into coherent regions

• **Algorithms:**
  – Divisive and hierarchical clustering
  – $k$-means clustering
  – Mean shift clustering
  – Graph cuts

• **Applications**
  – Image processing, object recognition, interactive image editing, etc.
Next week

Midterms:

The delicious ice cream center of your semester.

http://viterbivoices.usc.edu/wp-content/uploads/2012/10/midterms-spring-08-300x300.jpeg