CS5112: Algorithms and Data Structures for Applications

Lecture 20: Image segmentation

Ramin Zabih

Some content from: Wikipedia/Google image search
Lecture Outline

• Segmentation, in grayscale and color images
• Image representations
• Edge detection
• Segmentation with shortest paths
• Clustering based segmentation
• Mean shift segmentation
• Non-parametric density estimation (Parzen)
• Multi-modal distributions
Images and segmentation

• Images are 2D arrays, typically 512-by-512 or bigger
  – Video is images at 30 hz
• Entries are 8 bit (grayscale) or 24 bit (R/G/B color)
  – Black is 0 or 0/0/0, white is 255 or 255/255/255
• Segmentation: produce a meaningful partition of the image
  – Allegedly task independent
• Warning: color does not work the way you think it does
  – “Color constancy”
Image and feature space
Image segmentation example

Input Image: cameraman

segmented Image: cameraman
Color constancy famous example
Comparison of segmentation algorithms
Three common image representations used for most algorithms and image processing:

1. Array representation. This is what a camera produces.
2. Feature representation, where each pixel is mapped into a feature space. Simplest example: feature = intensity or color.
3. Graph representation. Pixels are nodes, edges connect adjacent pixels, usually 4-connected (grid)
Image as a graph

- Nodes are pixels, edges connect adjacent pixels
- Question: what weights do we put on the edges?
- This depends on what we are trying to compute
- Sometimes we want to know how similar two pixels are

Usual definition: affinity = \( \exp \left( - \frac{ (I(p) - I(q))^2 }{ 2\sigma^2 } \right) \)

- This is often computed with a larger neighborhood system
  - Neighbors of a pixel are all pixels within some radius
Normalized cuts

- Famous computer vision paper (Shi and Malik, 2000)
- Used spectral methods, i.e. eigenvectors of the affinity matrix
- Well beyond the scope of this course
Edge detection and gradients

• For some applications you need a small value for similar pixels
  – An edge that you want to follow around an object
• You can invert affinity, but in practice it works much better to use something based on edge detection
Edge detection ideas

- Two basic notions, easily seen in 1D
  - Large first derivative (gradient, Canny-style)
  - Zero second derivative (Laplacian)
Intelligent scissors

• Idea: shortest paths
  – E.N. Mortensen and W.A. Barrett, Interactive Segmentation with Intelligent Scissors, SIGGRAPH 1995
• Adobe calls this the “Magnetic Lasso”
  – Video here
• Basic idea: image is a graph, connectivity is how much perpendicular contrast there is between adjacent pixels
  – Computed based on edge detection
Video demonstration
Mean shift algorithm

- Hill climbing algorithm based on local density of data
  - Density increases as we get near “center”
Computing mean shift

• About as simple and effective an algorithm as anything
• Only one parameter!
• Given a box of radius $r$
  – Compute the centroid of the data within the box
  – Subtract the center of the box
  – This is the mean shift vector
• Take a (scaled) step in that direction until you are ‘done’
• Computes a local mode
Local modes
Mean shift segmentations
Multi-modal distributions

• Statistics has historically focused on unimodal distributions
  – Normal distribution, justified by the Central Limit Theorem (Gauss)
• Multi-modal distributions are:
  – Problematic
  – All over the place in computer vision
    • Even compared to machine learning
• Why are they hard? How do we handle them?
Multi-modal distributions are hard

• Standard statistical measures are not meaningful
  – They often implicitly assume normal distribution
    • Or something very close to it

• This is sometimes described in terms of outliers

• What is the average weight of the humans in this picture?
How to handle?

• Easiest cases:
  – Very small number of outliers
    • Statistics textbook historically suggest you plot your data and filter it (!)
    • But you can sometimes pre-filter the data
  – Two gaussians (mixture of gaussians, aka MOG)
    • Example: mischievous lab partner in chemistry
Two gaussians
Standard algorithm: expectation maximization

• We have a chicken and egg problem
• If we knew which data is water and which is beer, we could compute the mean and variance separately
• If we know the mean and variance were for beer and water, we could figure out which data is water and which is beer
• But we don’t know anything!
• So, just like in k-means, we guess and iterate
EM ITERATION 1

From P. Smyth
ICML 2001
Red Blood Cell Volume
Red Blood Cell Hemoglobin Concentration

EM ITERATION 3

From P. Smyth
ICML 2001
From P. Smyth
ICML 2001
EM ITERATION 15

Red Blood Cell Volume
Red Blood Cell Hemoglobin Concentration

From P. Smyth
ICML 2001
EM ITERATION 25

From P. Smyth
ICML 2001