Last time: Grouping

- Bottom-up segmentation via clustering
  - To find mid-level regions, tokens
  - General choices -- features, affinity functions, and clustering algorithms
  - Example clustering algorithms
    - Mean shift and mode finding: K-means, Mean shift
      - Graph theoretic: Graph cut, normalized cuts
- Grouping also useful for quantization
  - Texton histograms for texture within local region

Recall: Images as graphs

- Fully-connected graph
  - node for every pixel
  - link between every pair of pixels, p,q
  - similarity \( w_{pq} \) for each link
  - similarity is inversely proportional to difference in color and position
Last time: Measuring affinity

40 data points

40 x 40 affinity matrix $A$

$$A(i, j) = \exp\left(-\frac{1}{2\sigma^2}\|x_i - x_j\|^2\right)$$

1. What do the blocks signify?
2. What does the symmetry of the matrix signify?
3. How would the matrix change with larger value of $\sigma$?

Example: weighted graphs

- Suppose we have a 4-pixel image (i.e., a $2 \times 2$ matrix)
- Each pixel described by 2 features

Dimension of data points: $d = 2$
Number of data points: $N = 4$

Example: weighted graphs

Computing the distance matrix:

$$D(i, j) = \|x_i - x_j\|^2$$

for $i = 1:N$
for $j = 1:N$

for $i = 1:N$
for $j = 1:N$

for $i = 1:N$
for $j = 1:N$

for $i = 1:N$
for $j = 1:N$
for \( i = 1: N \)
for \( j = 1: N \)
\[
D(i,j) = \|x_i - x_j\|^2
\]
end
end

Example: weighted graphs
Computing the distance matrix:

\[
D(1,:) = \begin{bmatrix}
0.24 & 0.01 & 0.47 \\
0.15 & 0.24 & 0.29 \\
0.29 & 0.15 & 0.24
\end{bmatrix}
\]

Example: weighted graphs
Computing the distance matrix:

\[
N \times N \text{ matrix}
\]

Example: weighted graphs
Distances \(\rightarrow\) affinities

for \( i = 1: N \)
for \( j = i+1: N \)
\[
A(i,j) = \exp(-1/(2*\sigma^2)*\|x_i - x_j\|^2);
A(j,i) = A(i,j);
end
end
Scale parameter $\sigma$ affects affinity

Distance matrix $D$

Affinity matrix with increasing $\sigma$:

Visualizing a shuffled affinity matrix

If we permute the order of the vertices as they are referred to in the affinity matrix, we see different patterns:

Putting these two aspects together

$$A(i,j) = \exp\left(-\left(\frac{1}{2\sigma^2}\right)\|x_i - x_j\|^2\right)$$
Goal: Segmentation by Graph Cuts

Break graph into segments
- Delete links that cross between segments
  - Easiest to break links that have low similarity
    - similar pixels should be in the same segments
    - dissimilar pixels should be in different segments

Cuts in a graph: Min cut

Link Cut
- set of links whose removal makes a graph disconnected
- cost of a cut: \( \text{cut}(A, B) = \sum_{p \in A, q \in B} w_{pq} \)

Find minimum cut
- gives you a segmentation
- fast algorithms exist

Cuts in a graph: Normalized cut

- Fix bias of Min Cut by normalizing for size of segments:
  \[
  N\text{cut}(A, B) = \frac{\text{cut}(A, B)}{\text{assoc}(A, V)} \cdot \frac{\text{cut}(A, B)}{\text{assoc}(B, V)}
  \]
  \( \text{assoc}(A, V) = \) sum of weights of all edges that touch \( A \)
- \( N \text{cut} \) value is small when we get two clusters with many edges with high weights, and few edges of low weight between them.

J. Shi and J. Malik, Normalized Cuts and Image Segmentation, CVPR, 1997
Normalized cuts: pros and cons

Pros:
- Generic framework, flexible to choice of function that computes weights ("affinities") between nodes
- Does not require model of the data distribution

Cons:
- Time complexity can be high
  - Dense, highly connected graphs \( \rightarrow \) many affinity computations
  - Solving eigenvalue problem
- Preference for balanced partitions

Now: Fitting

- Want to associate a model with multiple observed features
  
  For example, the model could be a line, a circle, or an arbitrary shape.
Fitting: Main idea

- Choose a parametric model that best represent a set of features
- Membership criterion is not local
  - Can’t tell whether a point belongs to a given model just by looking at that point
- Three main questions:
  - What model represents this set of features best?
  - Which of several model instances gets which feature?
  - How many model instances are there?
- Computational complexity is important
  - It is infeasible to examine every possible set of parameters and every possible combination of features

Example: Line fitting

- Why fit lines?
  Many objects characterized by presence of straight lines

- Wait, why aren’t we done just by running edge detection?

Difficulty of line fitting

- Extra edge points (clutter), multiple models:
  - which points go with which line, if any?
- Only some parts of each line detected, and some parts are missing:
  - how to find a line that bridges missing evidence?
- Noise in measured edge points, orientations:
  - how to detect true underlying parameters?
Voting

- It’s not feasible to check all combinations of features by fitting a model to each possible subset.
- **Voting** is a general technique where we let each feature vote for all models that are compatible with it.
  - Cycle through features, cast votes for model parameters.
  - Look for model parameters that receive a lot of votes.
- Noise & clutter features will cast votes too, but typically their votes should be inconsistent with the majority of “good” features.

Fitting lines: Hough transform

- Given points that belong to a line, what is the line?
- How many lines are there?
- Which points belong to which lines?
- **Hough Transform** is a voting technique that can be used to answer all of these questions.
  **Main idea:**
  1. Record vote for each possible line on which each edge point lies.
  2. Look for lines that get many votes.

Finding lines in an image: Hough space

Connection between image (x,y) and Hough (m,b) spaces

- A line in the image corresponds to a point in Hough space
- To go from image space to Hough space:
  - given a set of points (x,y), find all (m,b) such that y = mx + b
Finding lines in an image: Hough space

Connection between image (x,y) and Hough (m,b) spaces
- A line in the image corresponds to a point in Hough space
- To go from image space to Hough space:
  - given a set of points (x,y), find all (m,b) such that y = mx + b
- What does a point (x₀, y₀) in the image space map to?
  - Answer: the solutions of \[ b = -x₀m + y₀ \]
  - this is a line in Hough space

Finding lines in an image: Hough algorithm

How can we use this to find the most likely parameters (m,b) for the most prominent line in the image space?
- Let each edge point in image space vote for a set of possible parameters in Hough space
- Accumulate votes in discrete set of bins; parameters with the most votes indicate line in image space.
Polar representation for lines

Issues with usual \((m, b)\) parameter space: can take on infinite values, undefined for vertical lines.

\[ d : \text{perpendicular distance from line to origin} \]
\[ \theta : \text{angle the perpendicular makes with the x-axis} \]
\[ x \cos \theta + y \sin \theta = d \]

Point in image space \(\rightarrow\) sinusoid segment in Hough space

---

Hough transform algorithm

Using the polar parameterization:

\[ x \cos \theta + y \sin \theta = d \]

Basic Hough transform algorithm

1. Initialize \(H[d, \theta]=0\)
2. For each edge point \((x, y)\) in the image
   1. For \(\theta = \theta_{\text{min}} \text{ to } \theta_{\text{max}}\) // some quantization
      1. \(d = x \cos \theta + y \sin \theta\)
      2. \(H[d, \theta] += 1\)
3. Find the value(s) of \((d, \theta)\) where \(H[d, \theta]\) is maximum
4. The detected line in the image is given by \(d = x \cos \theta + y \sin \theta\)

Time complexity (in terms of number of votes per pt)?

---

Hough line demo


---

Slide credit: Kristen Grauman
1. Image → Canny

2. Canny → Hough votes

3. Hough votes → Edges

Find peaks
Hough transform example

Original image

Canny edges

Vote space and top peaks

Showing longest segments found

Impact of noise on Hough

Image space edge coordinates

Votes

What difficulty does this present for an implementation?
Here, everything appears to be "noise", or random edge points, but we still see peaks in the vote space.

**Extensions**

Recall: when we detect an edge point, we also know its gradient direction.

**Extension 1:** Use the image gradient

1. same
2. for each edge point [x,y] in the image
   \[ \theta = \text{gradient at } (x,y) \]
   \[ d = x \cos \theta + y \sin \theta \]
   \[ |r(d, \theta)| = 1 \]
3. same
4. same
   (Reduces degrees of freedom)

**Extension 2**

• give more votes for stronger edges (use magnitude of gradient)

**Extension 3**

• change the sampling of (d, \theta) to give more/less resolution

**Extension 4**

• The same procedure can be used with circles, squares, or any other shape...

*Source: Steve Seitz*
Hough transform for circles

- Circle: center (a, b) and radius r
  \[(x - a)^2 + (y - b)^2 = r^2\]

- For a fixed radius r

\[r \times b + a = 0\]

Equation of circle?

Equation of set of circles that all pass through a point?

Image space

Hough space

Intersection: most votes for center occur here.

Kristen Grauman
Hough transform for circles

- Circle: center \((a, b)\) and radius \(r\)
  \[ (x - a)^2 + (y - b)^2 = r^2 \]
- For an unknown radius \(r\)

\[
\begin{align*}
\theta & \quad x \\
\rho & \quad y
\end{align*}
\]

Hough space

Image space

Kristen Grauman

---

Hough transform for circles

- Circle: center \((a, b)\) and radius \(r\)
  \[ (x - a)^2 + (y - b)^2 = r^2 \]
- For an unknown radius \(r\), **known** gradient direction

\[
\begin{align*}
\theta & \quad x \\
\rho & \quad y
\end{align*}
\]

Hough space

Image space

Kristen Grauman

---

Hough transform for circles

For every edge pixel \((x, y)\):
- For each possible radius value \(r\):
  - For each possible gradient direction \(\theta\):
    // or use estimated gradient at \((x, y)\)
    \[
    \begin{align*}
    a &= x - \cos(\theta) \quad // \text{column} \\
    b &= y + \sin(\theta) \quad // \text{row} \\
    H[a, b, r] &= +1
    \end{align*}
    \]

\[
\begin{align*}
\theta & \quad x \\
\rho & \quad y
\end{align*}
\]

Hough space

Image space

Kristen Grauman

---

**Time complexity per edge?**

- Check out online demo: [http://www.markschulze.net/java/hough/](http://www.markschulze.net/java/hough/)
Example: detecting circles with Hough

Original   Edges   Votes: Penny

Note: a different Hough transform (with separate accumulators) was used for each circle radius (quarters vs. penny).

Slide credit: Kristen Grauman

Example: detecting circles with Hough

Combined detections   Edges   Votes: Quarter

Coin finding sample images from Vivek Kwatra

Slide credit: Kristen Grauman

Example: iris detection

Gradient+threshold   Hough space (fixed radius)   Max detections

Hemerson Pistori and Eduardo Rocha Costa

Slide credit: Kristen Grauman
Example: iris detection


Voting: practical tips

- Minimize irrelevant tokens first
- Choose a good grid / discretization
  
  Too fine | ? | Too coarse

- Vote for neighbors, also (smoothing in accumulator array)
- Use direction of edge to reduce parameters by 1

Hough transform: pros and cons

Pros
- All points are processed independently, so can cope with occlusion, gaps
- Some robustness to noise: noise points unlikely to contribute consistently to any single bin
- Can detect multiple instances of a model in a single pass

Cons
- Complexity of search time increases exponentially with the number of model parameters
- Non-target shapes can produce spurious peaks in parameter space
- Quantization: can be tricky to pick a good grid size
now suppose those colors encode gradient directions…

• What if we want to detect arbitrary shapes?

**Intuition:**

*What if we want to detect arbitrary shapes?*

**Generalized Hough Transform**

- Define a model shape by its boundary points and a reference point.

**Offline procedure:**

At each boundary point, compute displacement vector: \( r = a - p_i \).

Store these vectors in a table indexed by gradient orientation \( \theta \).

**Detection procedure:**

For each edge point:

- Use its gradient orientation \( \theta \) to index into stored table
- Use retrieved \( r \) vectors to vote for reference point

Assuming translation is the only transformation here, i.e., orientation and scale are fixed.
Generalized Hough for object detection

- Instead of indexing displacements by gradient orientation, index by matched local patterns.

![Training image](image)

"visual codeword" with displacement vectors

B. Leibe, A. Leonardis, and B. Schiele. *Combined Object Categorization and Segmentation with an Implicit Shape Model*, ECCV Workshop on Statistical Learning in Computer Vision 2004

Source: L. Lazebnik

---

Generalized Hough for object detection

- Instead of indexing displacements by gradient orientation, index by "visual codeword"

![Test image](image)

Source: L. Lazebnik

---

Summary

- **Grouping/segmentation** useful to make a compact representation and merge similar features
  - associate features based on defined similarity measure and clustering objective

- **Fitting** problems require finding any supporting evidence for a model, even within clutter and missing features.
  - associate features with an explicit model

- **Voting** approaches, such as the Hough transform, make it possible to find likely model parameters without searching all combinations of features.
  - Hough transform approach for lines, circles, ..., arbitrary shapes defined by a set of boundary points, recognition from patches.
Questions?

See you Tuesday!