Computer Vision - Lecture 8

Graph-Theoretic Segmentation

26.11.2009

Bastian Leibe
RWTH Aachen
http://www.mmp.rwth-aachen.de

leibe@umic.rwth-aachen.de
Course Outline

- Image Processing Basics
- Recognition I
  - Global Representations
- Segmentation
  - Segmentation and Grouping
  - Graph-Theoretic Segmentation
- Recognition II
  - Subspace representations
- Local Features & Matching
- Object Categorization
- 3D Reconstruction
- Motion and Tracking
Recap: Gestalt Theory

- Gestalt: whole or group
  - Whole is greater than sum of its parts
  - Relationships among parts can yield new properties/features

- Psychologists identified series of factors that predispose set of elements to be grouped (by human visual system)

> "I stand at the window and see a house, trees, sky. Theoretically I might say there were 327 brightnesses and nuances of colour. Do I have "327"? No. I have sky, house, and trees."

Max Wertheimer
(1880-1943)

Untersuchungen zur Lehre von der Gestalt,
http://psy.ed.asu.edu/~classics/Wertheimer/Forms/forms.htm
Recap: Gestalt Factors

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

- These factors make intuitive sense, but are very difficult to translate into algorithms.
Recap: Image Segmentation

- Goal: identify groups of pixels that go together
Recap: K-Means Clustering

• Basic idea: randomly initialize the $k$ cluster centers, and iterate between the two steps we just saw.

1. Randomly initialize the cluster centers, $c_1, \ldots, c_k$
2. Given cluster centers, determine points in each cluster
   - For each point $p$, find the closest $c_i$. Put $p$ into cluster $i$
3. Given points in each cluster, solve for $c_i$
   - Set $c_i$ to be the mean of points in cluster $i$
4. If $c_i$ have changed, repeat Step 2

• Properties
  - Will always converge to some solution
  - Can be a “local minimum”
    - Does not always find the global minimum of objective function:
      $$\sum_{\text{clusters } i} \sum_{\text{points } p \text{ in cluster } i} \|p - c_i\|^2$$
Recap: Expectation Maximization (EM)

- **Goal**
  - Find blob parameters $\theta$ that maximize the likelihood function:
    \[
P(data|\theta) = \prod_x P(x|\theta)
    \]

- **Approach:**
  1. **E-step:** given current guess of blobs, compute ownership of each point
  2. **M-step:** given ownership probabilities, update blobs to maximize likelihood function
  3. Repeat until convergence
Recap: Mean-Shift Algorithm

- **Iterative Mode Search**
  1. Initialize random seed, and window $W$
  2. Calculate center of gravity (the “mean”) of $W$: $\sum_{x \in W} x H(x)$
  3. Shift the search window to the mean
  4. Repeat Step 2 until convergence
Recap: Mean-Shift Clustering

- Cluster: all data points in the attraction basin of a mode
- Attraction basin: the region for which all trajectories lead to the same mode
Recap: Mean-Shift Segmentation

- Find features (color, gradients, texture, etc)
- Initialize windows at individual pixel locations
- Perform mean shift for each window until convergence
- Merge windows that end up near the same “peak” or mode
Recap: Generic Clustering

- We have focused on ways to group pixels into image segments based on their appearance
  - Find groups; “quantize” feature space

- In general, we can use clustering techniques to find groups of similar “tokens”, provided we know how to compare the tokens.
  - E.g., segment an image into the types of motions present
  - E.g., segment a video into the types of scenes (shots) present
Topics of This Lecture

• Graph theoretic segmentation
  ➢ Normalized Cuts
  ➢ Using texture features

• Hierarchical segmentation
  ➢ Case study: segmentation-based recognition

• Segmentation as Energy Minimization
  ➢ Markov Random Fields
  ➢ Graph cuts for image segmentation
  ➢ s-t mincut algorithm
  ➢ Extension to non-binary case
  ➢ Applications
Images as Graphs

- **Fully-connected graph**
  - Node (vertex) for every pixel
  - Link between *every* pair of pixels, \((p,q)\)
  - Affinity weight \(w_{pq}\) for each link (edge)
    - \(w_{pq}\) measures similarity
    - Similarity is *inversely proportional* to difference (in color and position...)

Slide credit: Steve Seitz
Segmentation by Graph Cuts

- Break Graph into Segments
  - Delete links that cross between segments
  - Easiest to break links that have low similarity (low weight)
    - Similar pixels should be in the same segments
    - Dissimilar pixels should be in different segments

Slide credit: Steve Seitz
Measuring Affinity

- **Distance**
  \[ \text{aff}(x, y) = \exp\left\{-\frac{1}{2\sigma_d^2} \|x - y\|^2\right\} \]

- **Intensity**
  \[ \text{aff}(x, y) = \exp\left\{-\frac{1}{2\sigma_d^2} \|I(x) - I(y)\|^2\right\} \]

- **Color**
  \[ \text{aff}(x, y) = \exp\left\{-\frac{1}{2\sigma_d^2} \text{dist}(c(x), c(y))^2\right\} \]
  (some suitable color space distance)

- **Texture**
  \[ \text{aff}(x, y) = \exp\left\{-\frac{1}{2\sigma_d^2} \|f(x) - f(y)\|^2\right\} \]
  (vectors of filter outputs)

Source: Forsyth & Ponce
Scale Affects Affinity

- Small $\sigma$: group only nearby points
- Large $\sigma$: group far-away points

Slide credit: Svetlana Lazebnik
Graph Cut

- Set of edges whose removal makes a graph disconnected
- Cost of a cut
  - Sum of weights of cut edges: \( cut(A, B) = \sum_{p\in A, q\in B} w_{p,q} \)
- A graph cut gives us a segmentation
  - What is a “good” graph cut and how do we find one?

Slide credit: Steve Seitz
Here, the cut is nicely defined by the block-diagonal structure of the affinity matrix.

⇒ How can this be generalized?
Minimum Cut

• We can do segmentation by finding the *minimum cut* in a graph
  - Efficient algorithms exist for doing this

• Drawback:
  - Weight of cut proportional to number of edges in the cut
  - Minimum cut tends to cut off very small, isolated components

Ideal Cut

Cuts with lesser weight than the ideal cut

Slide credit: Khurram Hassan-Shafique
Normalized Cut (NCut)

- A minimum cut penalizes large segments
- This can be fixed by normalizing for size of segments
- The normalized cut cost is:

\[
Ncut(A, B) = \frac{\text{cut}(A, B)}{\text{assoc}(A, V)} + \frac{\text{cut}(A, B)}{\text{assoc}(B, V)}
\]

\[
\text{assoc}(A, V) = \text{sum of weights of all edges in V that touch A}
\]

\[
= \text{cut}(A, B) \left[ \frac{1}{\sum_{p\in A} w_{p,q}} + \frac{1}{\sum_{q\in B} w_{p,q}} \right]
\]

- The exact solution is NP-hard but an approximation can be computed by solving a generalized eigenvalue problem.

• 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

Slide credit: Steve Seitz
NCuts as a Generalized Eigenvector Problem

- Definitions
  
  $W$: the affinity matrix, $W(i, j) = w_{i,j}$;
  
  $D$: the diag. matrix, $D(i,i) = \sum_j W(i, j)$;
  
  $x$: a vector in $\{1, -1\}^N$, $x(i) = 1 \Leftrightarrow i \in A$.

- Rewriting Normalized Cut in matrix form:
  
  $NCut(A,B) = \frac{cut(A,B)}{assoc(A,V)} + \frac{cut(A,B)}{assoc(B,V)}$

  
  $= \frac{(1+x)^T(D-W)(1+x)}{k1^TD1} + \frac{(1-x)^T(D-W)(1-x)}{(1-k)1^TD1}$

  
  $k = \frac{\sum_{x_i > 0} D(i,i)}{\sum_i D(i,i)}$

  
  $= ...$

Slide credit: Jitendra Malik
Some More Math...

We see again this is an unbiased measure, which reflects how tightly on average, nodes within the group are connected to each other.

Another important property of this definition of association and disassociation of a partition is that they are mutually related:

\[
\begin{aligned}
N_{\text{assoc}}(A, B) &= \frac{\text{c}(A, B)}{\text{c}(A) \times \text{c}(B)} + \frac{\text{c}(A, B)}{\text{c}(A) \times \text{c}(B)} \\
N_{\text{disassoc}}(A, B) &= \frac{\text{c}(A) \times \text{c}(B) - \text{c}(A, B)}{\text{c}(A) \times \text{c}(B)} \\
\text{assoc}(A, B) &= -\frac{\text{c}(A) \times \text{c}(B) - \text{c}(A, B)}{\text{c}(A) \times \text{c}(B)} \\
\text{disassoc}(A, B) &= \frac{\text{c}(A) \times \text{c}(B) - \text{c}(A, B)}{\text{c}(A) \times \text{c}(B)}
\end{aligned}
\]

Hence the two partition criteria that we seek in our grouping algorithm, minimizing the disassociation between the groups and maximizing the association within the group, are in fact identical and can be satisfied simultaneously. In our algorithm, we will use this normalized version of the partition criteria.

Once defined the graph partition criteria that we want to optimize, we will show how such an optimal partition can be computed efficiently.

2.4 Computing the optimal partition

Given a partition of nodes of a graph, \( V \), into two sets \( A \) and \( B \), let us be an \( N \times |V| \) dimensional indicator vector, \( \mathbf{e}_i \) if node \( i \) is in \( A \), and \(-1\) otherwise. Let \( \mathbf{d}(i,j) = \sum \mathbf{e}_i \mathbf{e}_j \), be the total connection from node \( i \) to all other nodes. With the definitions above and below, we can rewrite \( \text{N_{eq}}(A, B) \) as:

\[
\begin{aligned}
N_{\text{assoc}}(A, B) &= \frac{\text{c}(A, B)}{\text{c}(A) \times \text{c}(B)} + \frac{\text{c}(A, B)}{\text{c}(A) \times \text{c}(B)} \\
N_{\text{disassoc}}(A, B) &= \frac{\text{c}(A) \times \text{c}(B) - \text{c}(A, B)}{\text{c}(A) \times \text{c}(B)} \\
\text{assoc}(A, B) &= -\frac{\text{c}(A) \times \text{c}(B) - \text{c}(A, B)}{\text{c}(A) \times \text{c}(B)} \\
\text{disassoc}(A, B) &= \frac{\text{c}(A) \times \text{c}(B) - \text{c}(A, B)}{\text{c}(A) \times \text{c}(B)}
\end{aligned}
\]

Let \( D \) be an \( N \times N \) diagonal matrix with \( \alpha \) on its diagonal, \( W \) be an \( N \times N \) symmetrical matrix with \( W(i,j) = \delta_{ij} \), \( \mathbf{b} = \sum \mathbf{d}(i,j), \) and \( \mathbf{s} \) be an \( N \times 1 \) vector of all ones. Using the fact the \( \mathbf{1}\mathbf{1}^T \) and \( \mathbf{1}\mathbf{d}^T \) are indicator vectors for \( s_i \geq 0 \) and \( s_i < 0 \) respectively, we can rewrite \( \text{assoc}(A, B) \) as:

\[
\begin{aligned}
\text{assoc}(A, B) &= \left( \sum_{i \in A} \mathbf{e}_i \right)^T D \left( \sum_{i \in A} \mathbf{e}_i \right) + \left( \sum_{i \in B} \mathbf{e}_i \right)^T D \left( \sum_{i \in B} \mathbf{e}_i \right)
\end{aligned}
\]

Let \( \alpha(n) = \alpha^T (D - \mathbf{W}) \mathbf{e}_n \), \( \beta(n) = \beta^T (D - \mathbf{W}) \mathbf{e}_n \), \( n = \alpha^T (D - \mathbf{W}) \mathbf{e}_n \), and \( M = \alpha^T D \mathbf{e}_n \), we can then further expand the above equation as:

\[
\begin{aligned}
\text{assoc}(A, B) &= \left( \sum_{i \in A} \mathbf{e}_i \right)^T D \left( \sum_{i \in A} \mathbf{e}_i \right) + \left( \sum_{i \in B} \mathbf{e}_i \right)^T D \left( \sum_{i \in B} \mathbf{e}_i \right)
\end{aligned}
\]

\[
\begin{aligned}
\text{assoc}(A, B) &= \frac{\alpha(n) + \beta(n) + 2 \alpha(n) \beta(n)}{M} - \frac{\alpha(n) + \beta(n) + 2 \alpha(n) \beta(n)}{M}
\end{aligned}
\]

\[
\begin{aligned}
\text{assoc}(A, B) &= 1 - \frac{\alpha(n) + \beta(n) + 2 \alpha(n) \beta(n)}{M}
\end{aligned}
\]

\[
\begin{aligned}
\text{assoc}(A, B) &= 1 - \frac{(1 - \alpha(n) + \beta(n) + 2 \alpha(n) \beta(n))}{M}
\end{aligned}
\]

\[
\begin{aligned}
\text{assoc}(A, B) &= 1 - \frac{\alpha(n) + \beta(n) + 2 \alpha(n) \beta(n)}{M}
\end{aligned}
\]

\[
\begin{aligned}
\text{assoc}(A, B) &= 1 - \frac{\alpha(n) + \beta(n) + 2 \alpha(n) \beta(n)}{M}
\end{aligned}
\]
NCuts as a Generalized Eigenvalue Problem

- After simplification, we get
  \[ NCut(A, B) = \frac{y^T (D-W)y}{y^T Dy}, \]  
  with \( y_i \in \{1, -b\} \), \( y^T D1 = 0 \).

- This is a Rayleigh Quotient
  - Solution given by the “generalized” eigenvalue problem
    \[ (D - W)y = \lambda Dy \]
  - Solved by converting to standard eigenvalue problem
    \[ D^{-\frac{1}{2}} (D - W)D^{-\frac{1}{2}} z = \lambda z, \]  
    where \( z = D^{\frac{1}{2}} y \).

- Subtleties
  - Optimal solution is second smallest eigenvector
  - Gives continuous result—must convert into discrete values of \( y \)

This is hard, as \( y \) is discrete!

Relaxation: continuous \( y \).
NCuts Example

Smallest eigenvectors

NCuts segments

B. Leibe

Image source: Shi & Malik
Discretization

- Problem: eigenvectors take on continuous values
  - How to choose the splitting point to binarize the image?

- Possible procedures
  a) Pick a constant value (0, or 0.5).
  b) Pick the median value as splitting point.
  c) Look for the splitting point that has the minimum $NCut$ value:
      1. Choose $n$ possible splitting points.
      2. Compute $NCut$ value.
      3. Pick minimum.
NCuts: Overall Procedure

1. Construct a weighted graph $G= (V,E)$ from an image.
2. Connect each pair of pixels, and assign graph edge weights,
   \[ W(i, j) = \text{Prob. that } i \text{ and } j \text{ belong to the same region.} \]
3. Solve \((D - W)y = \lambda Dy \) for the smallest few eigenvectors. This yields a continuous solution.
4. Threshold eigenvectors to get a discrete cut
   - This is where the approximation is made (we’re not solving NP).
5. Recursively subdivide if NCut value is below a pre-specified value.

NCuts Matlab code available at
http://www.cis.upenn.edu/~jshi/software/
Color Image Segmentation with NCuts
Using Texture Features for Segmentation

- Texture descriptor is vector of filter bank outputs


Slide credit: Svetlana Lazebnik
Using Texture Features for Segmentation

- Texture descriptor is vector of filter bank outputs.
- **Textons** are found by clustering.
Using Texture Features for Segmentation

- Texture descriptor is vector of filter bank outputs.
- **Textons** are found by clustering.
- Affinities are given by similarities of texton histograms over windows given by the “local scale” of the texture.
Results with Color & Texture
Summary: Normalized Cuts

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

• **Cons:**
  - Time and memory complexity can be high
    - Dense, highly connected graphs ⇒ many affinity computations
    - Solving eigenvalue problem for each cut
  - Preference for balanced partitions
    - If a region is uniform, NCuts will find the modes of vibration of the image dimensions
Topics of This Lecture

- **Graph theoretic segmentation**
  - Normalized Cuts
  - Using color and texture features

- **Hierarchical segmentation**
  - Case study: segmentation-based recognition

- **Segmentation as Energy Minimization**
  - Markov Random Fields
  - Graph cuts for image segmentation
  - s-t mincut algorithm
  - Extension to non-binary case
  - Applications

B. Leibe

Image Source: Ronen Basri
Multi-Level Segmentation

Example segmentations for several contrasts

• It is often difficult to extract a single good segmentation
  ➢ Idea: Extract a hierarchy of segmentations instead
Multiscale Segmentation Tree

- Arrange the segmentations in a tree

Example segmentations

Segmentation tree

Sample cutsets

Source: [S. Todorovic, N. Ahuja, CVPR’06]
Multiscale Segmentation Tree

- Now an image corresponds to a segmentation tree.
- An object forms a subtree.

Source: [S. Todorovic, N. Ahuja, CVPR’06]
Case Study: Tree-Based Recognition

Images = Trees

Category present = Many similar subtrees

Extracting similar subtrees = Tree matching

- Main difficulty: structural noise

Category model = Union of similar subtrees

Simultaneous detection, recognition and segmentation of ALL category instances by Matching the model with an image

Source: [S. Todorovic, N. Ahuja, CVPR’06]
Example Results: Cars

10 positive out of 20 training images

Results on test images:

Source: [S. Todorovic, N. Ahuja, CVPR’06]
Example Results: Faces

6 positive out of 12 training images

Results on test images:

- Results invariant to scale and rotation

B. Leibe

Source: [S. Todorovic, N. Ahuja, CVPR’06]
Tree-Based Recognition

• **Pros/Interesting ideas**
  - Instability of segmentation process circumvented by hierarchical segmentation.
  - Unsupervised learning of new category models.
  - Extraction of a semantically meaningful segment hierarchy.
  - Simultaneous recognition and segmentation.
  - Approach invariant to scale and rotation.

• **Cons**
  - Currently, this is still an experimental procedure.
  - Subtree matching is still very expensive: \( O(#\text{nodes}^4) \).
    - Training on 20 images takes \(~2\) hours.
  - Algorithm doesn’t scale well to larger training sets.
Topics of This Lecture

- **Graph theoretic segmentation**
  - Normalized Cuts
  - Using color and texture features

- **Hierarchical segmentation**
  - Case study: segmentation-based recognition

- **Segmentation as Energy Minimization**
  - Markov Random Fields
  - Graph cuts for image segmentation
  - s-t mincut algorithm
  - Extension to non-binary case
  - Applications
Markov Random Fields

- Allow rich probabilistic models for images
- But built in a local, modular way
  - Learn local effects, get global effects out

Observed evidence

Hidden “true states”

Neighborhood relations

Slide credit: William Freeman
MRF Nodes as Pixels

Original image

Degraded image

Reconstruction from MRF modeling pixel neighborhood statistics
MRF Nodes as Patches

\[ \Phi(x_i, y_i) \]
\[ \Psi(x_i, x_j) \]

Image patches

Scene patches

Image

Scene

Slide credit: William Freeman
Network Joint Probability

\[ P(x, y) = \prod_{i} \Phi(x_i, y_i) \prod_{i,j} \Psi(x_i, x_j) \]

- Scene
- Image
- Image-scene compatibility function
- Local observations
- Scene-scene compatibility function
- Neighboring scene nodes
Energy Formulation

- Joint probability

\[ P(x, y) = \prod_i \Phi(x_i, y_i) \prod_{i,j} \Psi(x_i, x_j) \]

- Maximizing the joint probability is the same as minimizing the log

\[ \log P(x, y) = \sum_i \log \Phi(x_i, y_i) + \sum_{i,j} \log \Psi(x_i, x_j) \]

\[ E(x, y) = \sum_i \phi(x_i, y_i) + \sum_{i,j} \psi(x_i, x_j) \]

- This is similar to free-energy problems in statistical mechanics (spin glass theory). We therefore draw the analogy and call \( E \) an energy function.

- \( \phi \) and \( \psi \) are called potentials.

B. Leibe
Energy Formulation

- Energy function

\[ E(x, y) = \sum_i \phi(x_i, y_i) + \sum_{i,j} \psi(x_i, x_j) \]

Single-node potentials \( \phi \)

- Encode local information about the given pixel/patch
- How likely is a pixel/patch to belong to a certain class (e.g. foreground/background)?

Pairwise potentials \( \psi \)

- Encode neighborhood information
- How different is a pixel/patch’s label from that of its neighbor? (e.g. based on intensity/color/texture difference, edges)
Energy Minimization

• Goal:
  - Infer the optimal labeling of the MRF.

• Many inference algorithms are available, e.g.
  - Gibbs sampling, simulated annealing
  - Iterated conditional modes (ICM)
  - Variational methods
  - Belief propagation
  - Graph cuts

• Recently, Graph Cuts have become a popular tool
  - Only suitable for a certain class of energy functions
  - But the solution can be obtained very fast for typical vision problems (~1MPixel/sec).
Topics of This Lecture

- **Graph theoretic segmentation**
  - Normalized Cuts
  - Using color and texture features

- **Hierarchical segmentation**
  - Case study: segmentation-based recognition

- **Segmentation as Energy Minimization**
  - Markov Random Fields
  - Graph cuts for image segmentation
  - s-t mincut algorithm
  - Extension to non-binary case
  - Applications
Graph Cuts for Optimal Boundary Detection

- Idea: convert MRF into source-sink graph

Minimum cost cut can be computed in polynomial time
(max-flow/min-cut algorithms)

\[ w_{pq} = \exp \left( -\frac{\Delta I_{pq}}{2\sigma^2} \right) \]

Slide credit: Yuri Boykov

[Boykov & Jolly, ICCV’01]
Simple Example of Energy

\[ E(L) = \sum_{p} D_p(L_p) + \sum_{pq \in N} w_{pq} \cdot \delta(L_p \neq L_q) \]

- **Regional term**
  \[ D_p(L_p) \]

- **Boundary term**
  \[ w_{pq} = \exp\left\{-\frac{\Delta I_{pq}}{2\sigma^2}\right\} \]

A cut

\[ L_p \in \{s, t\} \]

(binary object segmentation)

Slide credit: Yuri Boykov
Adding Regional Properties

Regional bias example

Suppose $I^s$ and $I^t$ are given “expected” intensities of object and background

\[
D_p(s) \propto \exp \left( -\frac{\|I_p - I^s\|^2}{2\sigma^2} \right)
\]
\[
D_p(t) \propto \exp \left( -\frac{\|I_p - I^t\|^2}{2\sigma^2} \right)
\]

NOTE: hard constrains are not required, in general.

Slide credit: Yuri Boykov

[Boykov & Jolly, ICCV’01]
Adding Regional Properties

“expected” intensities of object and background $I^s$ and $I^t$ can be re-estimated.

$D_p(s) \propto \exp\left(-\|I_p - I^s\|^2 / 2\sigma^2\right)$

$D_p(t) \propto \exp\left(-\|I_p - I^t\|^2 / 2\sigma^2\right)$

EM-style optimization

Slide credit: Yuri Boykov
Adding Regional Properties

• More generally, regional bias can be based on any intensity models of object and background

\[ D_p(L_p) = -\log \Pr(I_p | L_p) \]

given object and background intensity histograms

[Boykov & Jolly, ICCV’01]
How to Set the Potentials? Some Examples

- **Color potentials**
  - e.g. modeled with a Mixture of Gaussians
  \[
  \pi(x_i, y_i; \theta_\pi) = \log \sum_k \theta_\pi(x_i, k) P(k | x_i) N(y_i; \bar{y}_k, \Sigma_k)
  \]

- **Edge potentials**
  - e.g. a “contrast sensitive Potts model”
  \[
  \phi(x_i, x_j, g_{ij}(y); \theta_\phi) = -\theta_\phi^T g_{ij}(y) \delta(x_i \neq x_j)
  \]
  where
  \[
  g_{ij}(y) = e^{-\beta \|y_i - y_j\|^2}
  \]
  \[
  \beta = 2 \cdot \text{avg} \left(\|y_i - y_j\|^2\right)
  \]

- **Parameters** \(\theta_\pi, \theta_\phi\) need to be learned, too! [Shotton & Winn, ECCV’06]
Topics of This Lecture

- **Graph theoretic segmentation**
  - Normalized Cuts
  - Using color and texture features

- **Hierarchical segmentation**
  - Case study: segmentation-based recognition

- **Segmentation as Energy Minimization**
  - Markov Random Fields
  - Graph cuts for image segmentation
  - s-t mincut algorithm
  - Extension to non-binary case
  - Applications
How Does it Work? The s-t-Mincut Problem

Graph \((V, E, C)\)

Vertices \(V = \{v_1, v_2 \ldots v_n\}\)

Edges \(E = \{(v_1, v_2) \ldots\}\)

Costs \(C = \{c_{(1,2)} \ldots\}\)
The s-t-Mincut Problem

What is an st-cut?
An st-cut \((S, T)\) divides the nodes between source and sink.

What is the cost of an st-cut?
Sum of cost of all edges going from \(S\) to \(T\)

5 + 2 + 9 = 16
The s-t-Mincut Problem

What is an st-cut?

An st-cut (S,T) divides the nodes between source and sink.

What is the cost of an st-cut?

Sum of cost of all edges going from S to T

What is the st-mincut?

st-cut with the minimum cost

Source

Sink

2 + 1 + 4 = 7

Slide credit: Pushmeet Kohli
## History of Maxflow Algorithms

### Augmenting Path and Push-Relabel

<table>
<thead>
<tr>
<th>year</th>
<th>discoverer(s)</th>
<th>bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>1951</td>
<td>Dantzig</td>
<td>$O(n^2mU)$</td>
</tr>
<tr>
<td>1955</td>
<td>Ford &amp; Fulkerson</td>
<td>$O(m^2U)$</td>
</tr>
<tr>
<td>1970</td>
<td>Dinitz</td>
<td>$O(n^2m)$</td>
</tr>
<tr>
<td>1972</td>
<td>Edmonds &amp; Karp</td>
<td>$O(m^2 \log U)$</td>
</tr>
<tr>
<td>1973</td>
<td>Dinitz</td>
<td>$O(nm \log U)$</td>
</tr>
<tr>
<td>1974</td>
<td>Karzanov</td>
<td>$O(n^3)$</td>
</tr>
<tr>
<td>1977</td>
<td>Cherkassky</td>
<td>$O(n^2m^{1/2})$</td>
</tr>
<tr>
<td>1980</td>
<td>Galil &amp; Naamad</td>
<td>$O(nm \log^2 n)$</td>
</tr>
<tr>
<td>1983</td>
<td>Sleator &amp; Tarjan</td>
<td>$O(nm \log n)$</td>
</tr>
<tr>
<td>1986</td>
<td>Goldberg &amp; Tarjan</td>
<td>$O(nm \log(n^2/m))$</td>
</tr>
<tr>
<td>1987</td>
<td>Ahuja &amp; Orlin</td>
<td>$O(nm + n^2 \log U)$</td>
</tr>
<tr>
<td>1987</td>
<td>Ahuja et al.</td>
<td>$O(nm \log(n \sqrt{\log U/m}))$</td>
</tr>
<tr>
<td>1989</td>
<td>Cheriyan &amp; Hagerup</td>
<td>$E(nm + n^2 \log^2 n)$</td>
</tr>
<tr>
<td>1990</td>
<td>Cheriyan et al.</td>
<td>$O(n^3/ \log n)$</td>
</tr>
<tr>
<td>1990</td>
<td>Alon</td>
<td>$O(nm + n^{8/3} \log n)$</td>
</tr>
<tr>
<td>1992</td>
<td>King et al.</td>
<td>$O(nm + n^{2+\epsilon})$</td>
</tr>
<tr>
<td>1993</td>
<td>Phillips &amp; Westbrook</td>
<td>$O(nm (\log_{m/n} n + \log^{2+\epsilon} n))$</td>
</tr>
<tr>
<td>1994</td>
<td>King et al.</td>
<td>$O(nm \log_{m/(n\log n)} n)$</td>
</tr>
<tr>
<td>1997</td>
<td>Goldberg &amp; Rao</td>
<td>$O(m^{3/2} \log (n^2/m) \log U)$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$O(n^{2/3} m \log(n^2/m) \log U)$</td>
</tr>
</tbody>
</table>

$n$: #nodes  
$m$: #edges  
$U$: maximum edge weight

Algorithms assume non-negative edge weights.
How to Compute the s-t-Mincut?

Solve the dual maximum flow problem

Compute the maximum flow between Source and Sink

Constraints
- Edges: Flow < Capacity
- Nodes: Flow in = Flow out

Min-cut/Max-flow Theorem
In every network, the maximum flow equals the cost of the st-mincut

Slide credit: Pushmeet Kohli
Maxflow Algorithms

Augmenting Path Based Algorithms

1. Find path from source to sink with positive capacity
2. Push maximum possible flow through this path
3. Repeat until no path can be found

Algorithms assume non-negative capacity

Slide credit: Pushmeet Kohli
Maxflow Algorithms

Augmenting Path Based Algorithms

1. Find path from source to sink with positive capacity
2. Push maximum possible flow through this path
3. Repeat until no path can be found

Algorithms assume non-negative capacity
Maxflow Algorithms

Flow = 0 + 2

Augmenting Path Based Algorithms

1. Find path from source to sink with positive capacity
2. Push maximum possible flow through this path
3. Repeat until no path can be found

Algorithms assume non-negative capacity

Slide credit: Pushmeet Kohli
Maxflow Algorithms

**Augmenting Path Based Algorithms**

1. Find path from source to sink with positive capacity
2. Push maximum possible flow through this path
3. Repeat until no path can be found

**Algorithms assume non-negative capacity**
Maxflow Algorithms

Flow = 2

Augmenting Path Based Algorithms

1. Find path from source to sink with positive capacity
2. Push maximum possible flow through this path
3. Repeat until no path can be found

Algorithms assume non-negative capacity

Slide credit: Pushmeet Kohli
Maxflow Algorithms

Flow = 2

Augmenting Path Based Algorithms

1. Find path from source to sink with positive capacity
2. Push maximum possible flow through this path
3. Repeat until no path can be found

Algorithms assume non-negative capacity
Maxflow Algorithms

Flow = 2 + 4

Augmenting Path Based Algorithms

1. Find path from source to sink with positive capacity
2. Push maximum possible flow through this path
3. Repeat until no path can be found

Algorithms assume non-negative capacity

Slide credit: Pushmeet Kohli
Maxflow Algorithms

Flow = 6

Augmenting Path Based Algorithms

1. Find path from source to sink with positive capacity
2. Push maximum possible flow through this path
3. Repeat until no path can be found

Algorithms assume non-negative capacity

Slide credit: Pushmeet Kohli
Maxflow Algorithms

1. Find path from source to sink with positive capacity
2. Push maximum possible flow through this path
3. Repeat until no path can be found

Algorithms assume non-negative capacity

Augmenting Path Based Algorithms

Flow = 6
Maxflow Algorithms

Flow = 6 + 1

Augmenting Path Based Algorithms

1. Find path from source to sink with positive capacity
2. Push maximum possible flow through this path
3. Repeat until no path can be found

Algorithms assume non-negative capacity

Slide credit: Pushmeet Kohli
B. Leibe
Maxflow Algorithms

Augmenting Path Based Algorithms

1. Find path from source to sink with positive capacity

2. Push maximum possible flow through this path

3. Repeat until no path can be found

Algorithms assume non-negative capacity
Maxflow Algorithms

Augmenting Path Based Algorithms

1. Find path from source to sink with positive capacity
2. Push maximum possible flow through this path
3. Repeat until no path can be found

Algorithms assume non-negative capacity

Slide credit: Pushmeet Kohli
Maxflow in Computer Vision

- Specialized algorithms for vision problems
  - Grid graphs
  - Low connectivity \((m \sim O(n))\)

- Dual search tree augmenting path algorithm
  [Boykov and Kolmogorov PAMI 2004]
  - Finds approximate shortest augmenting paths efficiently
  - High worst-case time complexity
  - Empirically outperforms other algorithms on vision problems
  - Efficient code available on the web

http://www.adastral.ucl.ac.uk/~vladkolm/software.html

Slide credit: Pushmeet Kohli
When Can s-t Graph Cuts Be Applied?

- s-t graph cuts can only globally minimize binary energies that are submodular. [Boros & Hummer, 2002, Kolmogorov & Zabih, 2004]

\[ E(L) = \sum_p E_p(L_p) + \sum_{pq \in N} E(L_p, L_q) \]

- t-links \( L_p \in \{s, t\} \)

- Boundary term

\[ L_p \in \{s, t\} \]

- Regional term

\[ E(L) \text{ can be minimized by s-t graph cuts} \iff E(s,s) + E(t,t) \leq E(s,t) + E(t,s) \]

Submodularity ("convexity")

- Non-submodular cases can still be addressed with some optimality guarantees.
  - Current research topic

B. Leibe
Topics of This Lecture

- **Graph theoretic segmentation**
  - Normalized Cuts
  - Using color and texture features

- **Hierarchical segmentation**
  - Case study: segmentation-based recognition

- **Segmentation as Energy Minimization**
  - Markov Random Fields
  - Graph cuts for image segmentation
  - s-t mincut algorithm
  - Extension to non-binary case
  - Applications
Dealing with Non-Binary Cases

- For image segmentation, the limitation to binary energies is a nuisance.
  \[ \Rightarrow \text{Binary segmentation only} \]

- We would like to solve also multi-label problems.
  - NP-hard problem with 3 or more labels

- There exist some approximation algorithms which extend graph cuts to the multi-label case
  - \( \alpha \)-Expansion
  - \( \alpha\beta \)-Swap

- They are no longer guaranteed to return the globally optimal result.
  - But \( \alpha \)-Expansion has a guaranteed approximation quality (2-approx) and converges in a few iterations.
α-Expansion Move

- Basic idea:
  - Break multi-way cut computation into a sequence of binary s-t cuts.
α-Expansion Algorithm

1. Start with any initial solution
2. For each label “α” in any (e.g. random) order
   1. Compute optimal α-expansion move (s-t graph cuts)
   2. Decline the move if there is no energy decrease

• Stop when no expansion move would decrease energy
\( \alpha \)-Expansion Moves

- In each \( \alpha \)-expansion a given label “\( \alpha \)” grabs space from other labels

For each move we choose the expansion that gives the largest decrease in the energy: binary optimization problem
Topics of This Lecture

- Graph theoretic segmentation
  - Normalized Cuts
  - Using color and texture features

- Hierarchical segmentation
  - Case study: segmentation-based recognition

- Segmentation as Energy Minimization
  - Markov Random Fields
  - Graph cuts for image segmentation
  - s-t mincut algorithm
  - Extension to non-binary case
  - Applications
GraphCut Applications: “GrabCut”

- Interactive Image Segmentation [Boykov & Jolly, ICCV’01]
  - Rough region cues sufficient
  - Segmentation boundary can be extracted from edges

- Procedure
  - User marks foreground and background regions with a brush.
  - This is used to create an initial segmentation which can then be corrected by additional brush strokes.

Slide credit: Matthieu Bray
GrabCut: Data Model

- Obtained from interactive user input
  - User marks foreground and background regions with a brush
  - Alternatively, user can specify a bounding box

Global optimum of the energy
GrabCut: Coherence Model

- An object is a coherent set of pixels:

\[ \psi(x, y) = \gamma \sum_{(m,n) \in C} \delta[x_n \neq x_m] e^{-\beta\|y_m - y_n\|^2} \]

How to choose \( \gamma \)?

Error (%) over training set: 25
Iterated Graph Cuts

Result

Color model (Mixture of Gaussians)

Energy after each iteration

Slide credit: Carsten Rother
This will be included in the next version of MS Office!
Applications: Interactive 3D Segmentation

Slide credit: Yuri Boykov

B. Leibe

[Y. Boykov, V. Kolmogorov, ICCV’03]
Improving Efficiency of Segmentation

• Problem: Images contain many pixels
  ➢ Even with efficient graph cuts, an MRF formulation has too many nodes for interactive results.

• Efficiency trick: Superpixels
  ➢ Group together similar-looking pixels for efficiency of further processing.
  ➢ Cheap, local oversegmentation
  ➢ Important to ensure that superpixels do not cross boundaries

• Several different approaches possible
  ➢ Superpixel code available here
  ➢ http://www.cs.sfu.ca/~mori/research/superpixels/
Superpixels for Pre-Segmentation

Graph structure

<table>
<thead>
<tr>
<th>Image</th>
<th>Dimension</th>
<th>Nodes Ratio</th>
<th>Edges Ratio</th>
<th>Lag with Pre-segmentation</th>
<th>Lag without Pre-segmentation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boy</td>
<td>(408, 600)</td>
<td>10.7</td>
<td>16.8</td>
<td>0.12s</td>
<td>0.57s</td>
</tr>
<tr>
<td>Ballet</td>
<td>(440, 800)</td>
<td>11.4</td>
<td>18.3</td>
<td>0.21s</td>
<td>1.39s</td>
</tr>
<tr>
<td>Twins</td>
<td>(1024, 768)</td>
<td>20.7</td>
<td>32.5</td>
<td>0.25s</td>
<td>1.82s</td>
</tr>
<tr>
<td>Girl</td>
<td>(768, 1147)</td>
<td>23.8</td>
<td>37.6</td>
<td>0.22s</td>
<td>2.49s</td>
</tr>
<tr>
<td>Grandpa</td>
<td>(1147, 768)</td>
<td>19.3</td>
<td>30.5</td>
<td>0.22s</td>
<td>3.56s</td>
</tr>
</tbody>
</table>

Speedup
Summary: Graph Cuts Segmentation

• **Pros**
  - Powerful technique, based on probabilistic model (MRF).
  - Applicable for a wide range of problems.
  - Very efficient algorithms available for vision problems.
  - Becoming a de-facto standard for many segmentation tasks.

• **Cons/Issues**
  - Graph cuts can only solve a limited class of models
    - Submodular energy functions
    - Can capture only part of the expressiveness of MRFs
  - Only approximate algorithms available for multi-label case
Other Applications: Texture Synthesis

Graph-cut textures
(Katra, Schodl, Essa, Bobick 2003)

Similar to “image-quilting” (Efros & Freeman, 2001)
Basic Idea

Input texture

Random placement of blocks

Neighboring blocks constrained by overlap

Minimal error boundary cut

B. Leibe

Slide from Alyosha Efros
Minimal Error Boundary

Overlapping blocks

Vertical boundary

Overlap error

min. error boundary

B. Leibe

Slide from Alyosha Efros
GraphCut Texture Synthesis Results

Original fragments

Results (with perspective correction)

Source: Vivek Kwatra
GraphCut Image Synthesis Results

Source: Vivek Kwatra
Application: Texture Synthesis in the Media

- Currently, still done manually...
Application: Texture Synthesis in the Media

- Currently, still done manually...
How Do we Know?


Slide credit: Kristen Grauman
Another Example


Slide credit: Kristen Grauman
Segmentation: Caveats

- We’ve looked at *bottom-up* ways to segment an image into regions, yet finding meaningful segments is intertwined with the recognition problem.
- Often want to avoid making hard decisions too soon
- Difficult to evaluate; when is a segmentation successful?
  - Often depends on the rest of the system pipeline.
References and Further Reading

- Background information on Normalized Cuts can be found in Chapter 14 of

- Try the NCuts Matlab code at
  - [http://www.cis.upenn.edu/~jshi/software/](http://www.cis.upenn.edu/~jshi/software/)

- Try the GraphCut implementation at
  - [http://www.adastral.ucl.ac.uk/~vladkolm/software.html](http://www.adastral.ucl.ac.uk/~vladkolm/software.html)