Last Time

- Image segmentation
  - Normalized cut and segmentation
Today

- Segmentation
  - Interactive image segmentation

Images from Rother et al. 2004
Start

- Segmentation
  - Interactive image segmentation

Images from Sun et al. 2004
Segmentation by Graph Cut

- Interactive image segmentation using graph cut
- Binary label: foreground vs. background
- User labels some pixels
  - usually sparser
- Exploit
  - Statistics of known Fg & Bg
  - Smoothness of label
- Turn into discrete graph optimization
  - Graph cut (min cut / max flow)

Slide credit: Y.Y. Chuang
Energy function

- **Segmentation as Labeling**
  - one value per pixel, F or B

- **Energy(labeling) = data + smoothness**
  - Very general situation
  - Will be minimized

- **Data: for each pixel**
  - Probability that this color belongs to F (resp. B)

- **Smoothness (aka regularization):**
  - per neighboring pixel pair
  - Penalty for having different label
  - Penalty is down-weighted if the two pixel colors are very different
  - Similar in spirit to bilateral filter

Slide credit: F. Durand
Data term

- A.k.a regional term (because integrated over full region)
- \( D(L) = \sum_i -\log h[L_i](C_i) \)
- Where \( i \) is a pixel
  - \( L_i \) is the label at \( i \) (F or B),
  - \( C_i \) is the pixel value
  - \( h[L_i] \) is the histogram of the observed Fg (resp Bg)
- Note the minus sign

Slide credit: F. Durand
Hard constraints

- The user has provided some labels
- The quick and dirty way to include constraints into optimization is to replace the data term by a huge penalty if not respected.
- \( D(L_i) = 0 \) if respected
- \( D(L_i) = K \) if not respected
  - e.g. \( K = \#\text{pixels} \)
Smoothness term

- a.k.a boundary term, a.k.a. regularization
- \( S(L) = \sum_{\{i, j\} \in N} B(C_i, C_j) \delta(L_i - L_j) \)

Where \( i, j \) are neighbors
- e.g. 8-neighborhood (but I show 4 for simplicity)

- \( \delta(L_i - L_j) \) is 0 if \( L_i = L_j \), 1 otherwise

- \( B(C_i, C_j) \) is high when \( C_i \) and \( C_j \) are similar, low if there is a discontinuity between those two pixels
- e.g. \( \exp(-||C_i - C_j||^2 / 2\sigma^2) \)
- where \( \sigma \) can be a constant or the local variance

- Note positive sign
Optimization

- $E(L) = D(L) + \lambda S(L)$
- $\lambda$ is a black-magic constant
- Find the labeling that minimizes $E$
- In this case, how many possibilities?
  - $2^9 (512)$
  - We can try them all!
  - What about megapixel images?

Slide credit: F. Durand
Labeling as a graph problem

- Each pixel = node
- Add two nodes F & B
- Labeling: link each pixel to either F or B

Desired result

Slide credit: F. Durand
Data term

- Put one edge between each pixel and F & B
- Weight of edge = minus data term
  - Don’t forget huge weight for hard constraints
  - Careful with sign
Smoothness term

- Add an edge between each neighbor pair
- Weight = smoothness term
Min cut

- Energy optimization equivalent to min cut
- Cut: remove edges to disconnect F from B
- Minimum: minimize sum of cut edge weight
Min cut <=> labeling

- In order to be a cut:
  - For each pixel, either the F or G edge has to be cut

- In order to be minimal:
  - Only one edge label per pixel can be cut (otherwise could be added)
Computing a multiway cut

- With 2 labels: classical min-cut problem
  - Solvable by standard flow algorithms
    - polynomial time in theory, nearly linear in practice
    - Code: C++ from OpenCV
      - Matlab wrapper: http://www.wisdom.weizmann.ac.il/~bagon/matlab.html
  - More than 2 terminals: NP-hard [Dahlhaus et al., STOC ‘92]
    - Code: http://vision.ucla.edu/~brian/gcmex.html

- Efficient approximation algorithms exist
  - Within a factor of 2 of optimal
  - Computes local minimum in a strong sense
    - even very large moves will not improve the energy
GrabCut
Interactive Foreground Extraction using Iterated Graph Cuts

Carsten Rother
Vladimir Kolmogorov
Andrew Blake

Microsoft Research Cambridge-UK
Photomontage

GrabCut – Interactive Foreground Extraction
Framework

- **Input:** Image \( x \in \{R, G, B\}^n \)
- **Output:** Segmentation \( S \in \{0, 1\}^n \)
- **Parameters:** Colour \( \Theta \), Coherence \( \lambda \)
- **Energy:** \( E(\Theta, S, x, \lambda) = E_{Col} + E_{Coh} \)
- **Optimization:** \( \arg\min_{S, \Theta} E(S, \Theta, x, \lambda) \)
Graph Cuts

Boykov and Jolly (2001)

Cut: separating source and sink; Energy: collection of edges

Min Cut: Global minimal energy in polynomial time
Iterated Graph Cut

User Initialisation

\[ \arg \min_{\Theta} E(S, \Theta, x, \lambda) \]

K-means for learning colour distributions

Graph cuts to infer the segmentation
Iterated Graph Cuts

Result

Energy after each Iteration

Guaranteed to converge
Colour Model

Gaussian Mixture Model (typically 5-8 components)

$$E_{Col}(\Theta, S, x) = \sum_n D(S_n, \Theta, x_n)$$

Iterated graph cut
**Coherence Model**

An object is a coherent set of pixels:

\[ E_{coh}(S, x, \lambda) = \lambda \sum_{i,j \text{ adj.}} (S_i \neq S_j) \exp\left\{-\frac{1}{2\sigma^2}||x_i - x_j||^2\right\} \]

Blake et al. (2004): Learn \( \Theta, \lambda \) jointly

\( \lambda = 0 \) \hspace{2cm} \( \lambda = 50 \) \hspace{2cm} \( \lambda = 1000 \)
Moderately straightforward examples

... GrabCut completes automatically
Difficult Examples

Camouflage & Low Contrast

Initial Rectangle

Initial Result

Fine structure

No telepathy

GrabCut – Interactive Foreground Extraction
Evaluation – Labelled Database

Available online: http://research.microsoft.com/vision/cambridge/segmentation/

GrabCut – Interactive Foreground Extraction
Comparison

Boykov and Jolly (2001) vs GrabCut

User Input

Result

Error Rate: 0.72%

Error Rate: 0.72%
Summary

Magic Wand (198?)
Intelligent Scissors
Mortensen and Barrett (1995)
Graph Cuts
Boykov and Jolly (2001)
LazySnapping
Li et al. (2004)
GrabCut
Rother et al. (2004)
Interactive Digital Photomontage

- Combining multiple photos
- Find seams using graph cuts
- Combine gradients and integrate

Aseem Agarwala, Mira Dontcheva, Maneesh Agrawala, Steven Drucker, Alex Colburn, Brian Curless, David Salesin, Michael Cohen, “Interactive Digital Photomontage”, SIGGRAPH 2004
set of originals

photomontage

Slide credit: Y.Y. Chuang
Brush strokes

Computed labeling

Slide credit: Y.Y. Chuang
Next Time

☐ Matting