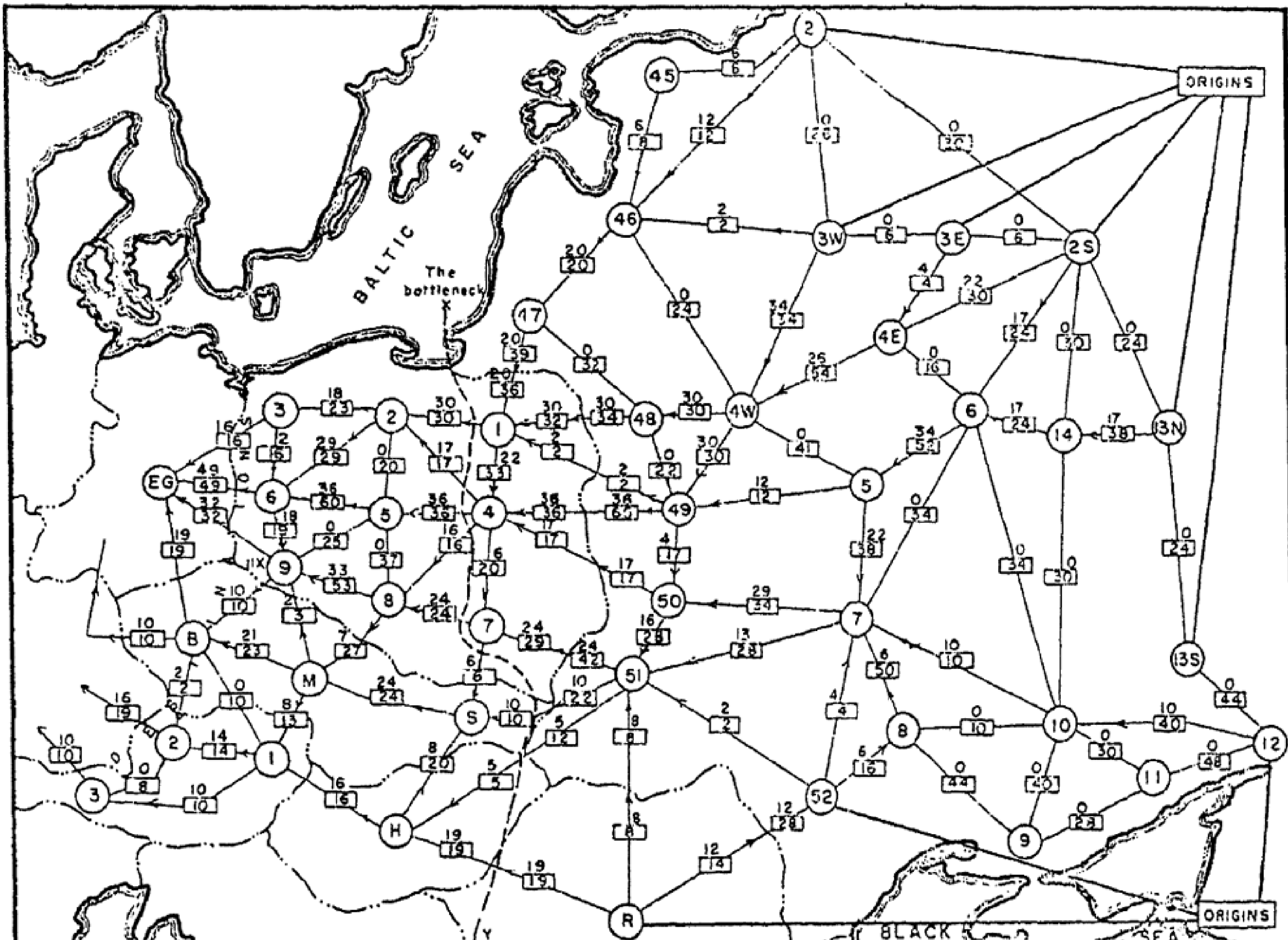

ORIGINS
BALTIC SEA
The bottleneck
X
Y
BLACK SEA
ORIGINS

10/07/11

Segmentation: MRFs and Graph Cuts

Computer Vision

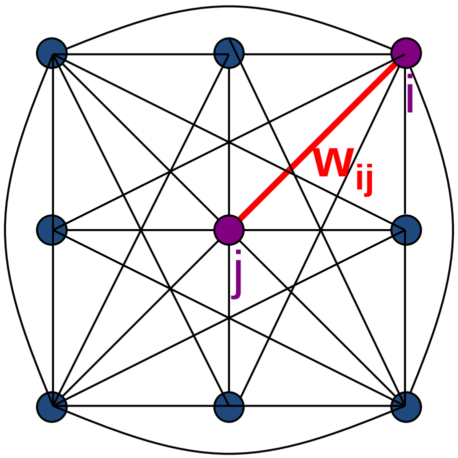
CS 143, Brown

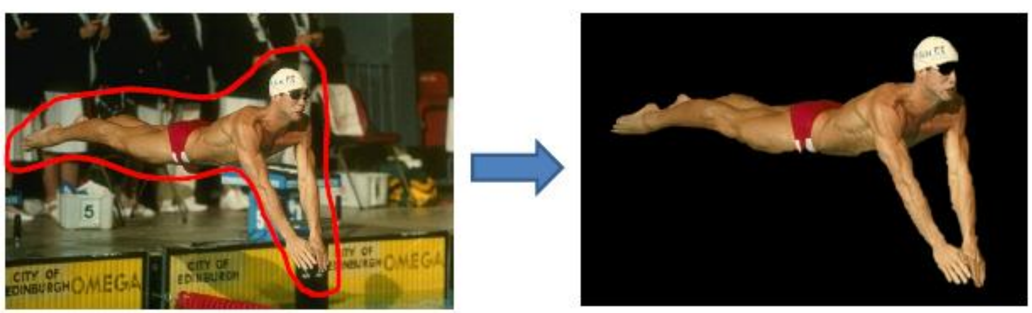
James Hays

Many slides from Kristin Grauman and Derek Hoiem

Today's class

- Segmentation and Grouping
- Inspiration from human perception
  - Gestalt properties
- MRFs
- Segmentation with Graph Cuts



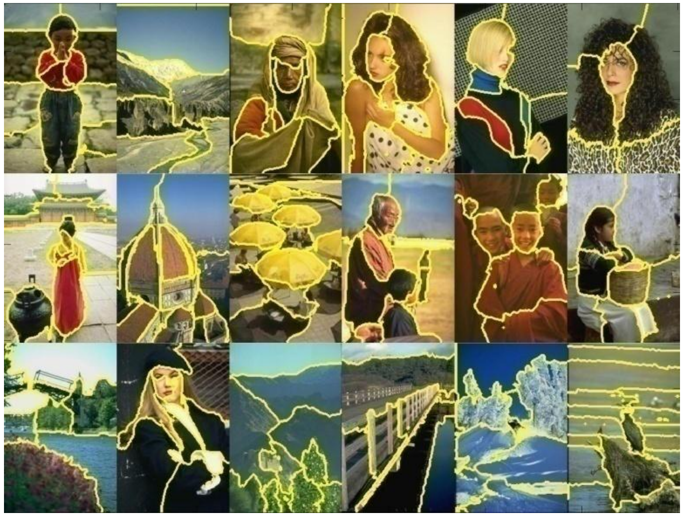


Slide: Derek Hoiem

Grouping in vision

- Goals:
  - Gather features that belong together
  - Obtain an intermediate representation that compactly describes key image or video parts

Examples of grouping in vision



[Figure by J. Shi]

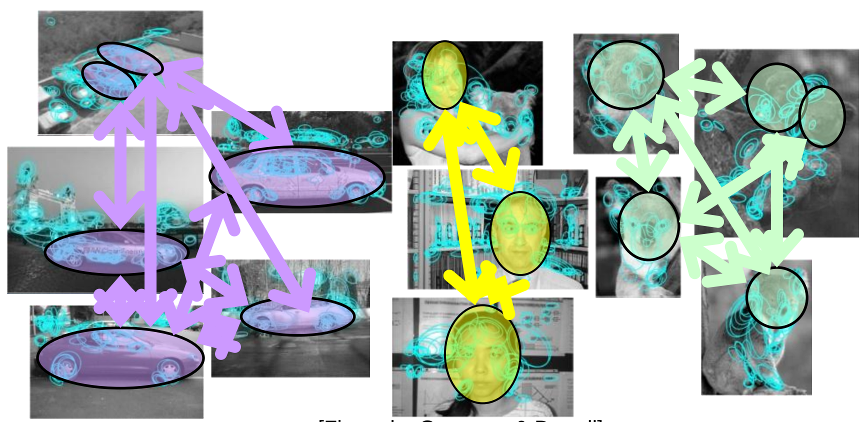
Determine image regions

[http://poseidon.csd.auth.gr/LAB_RESEARCH/Latest/imgs/SpeakDepVidIndex_img2.jpg]

Group video frames into shots

[Figure by Wang & Suter]

Figure-ground



[Figure by Grauman & Darrell]

Object-level grouping

Slide: Kristin Grauman

Grouping in vision

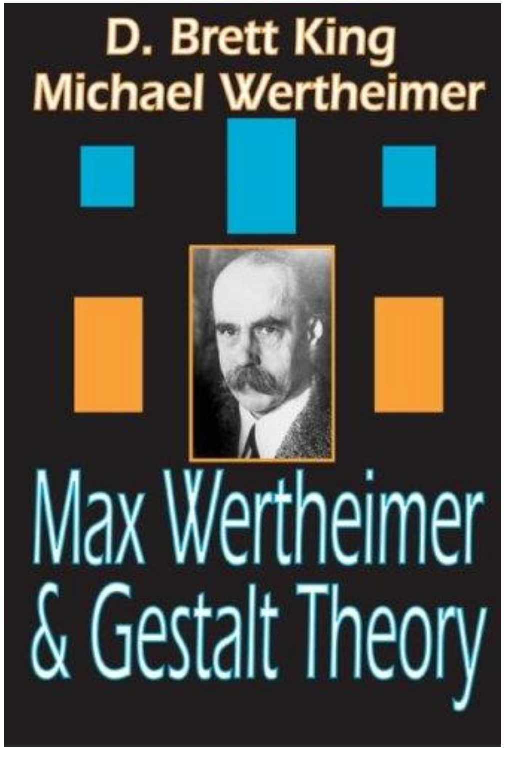
- Goals:
  - Gather features that belong together
  - Obtain an intermediate representation that compactly describes key image (video) parts
- Top down vs. bottom up segmentation
  - Top down: pixels belong together because they are from the same object
  - Bottom up: pixels belong together because they look similar
- Hard to measure success
  - What is interesting depends on the app.

Slide: Kristin Grauman

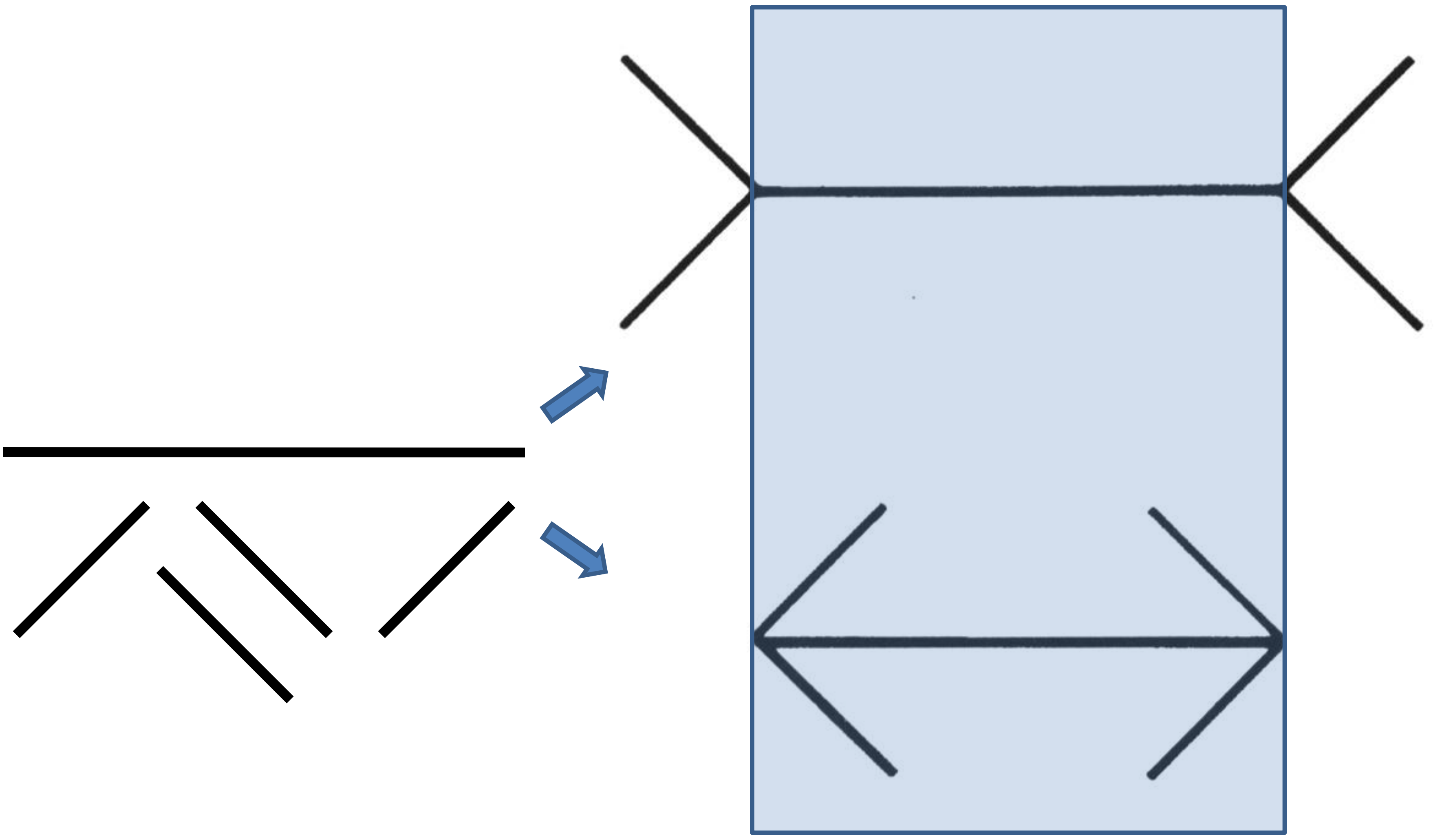
What things should be grouped?
What cues indicate groups?

Gestalt psychology or **Gestaltism**

- German: *Gestalt* - "form" or "whole”
- Berlin School, early 20th century
  - Kurt Koffka, Max Wertheimer, and Wolfgang Köhler
- 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)



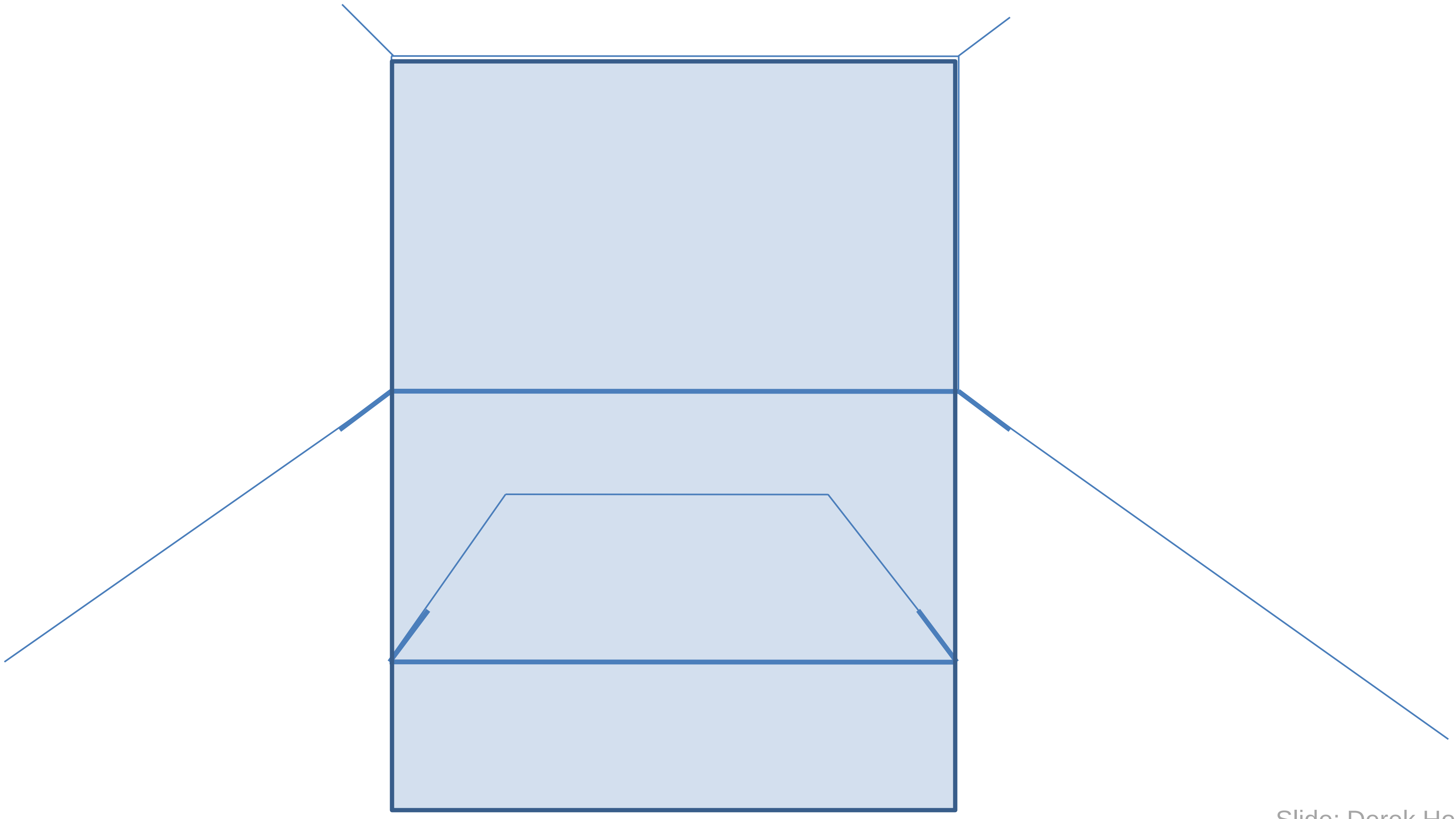
Gestaltism



The Muller-Lyer illusion

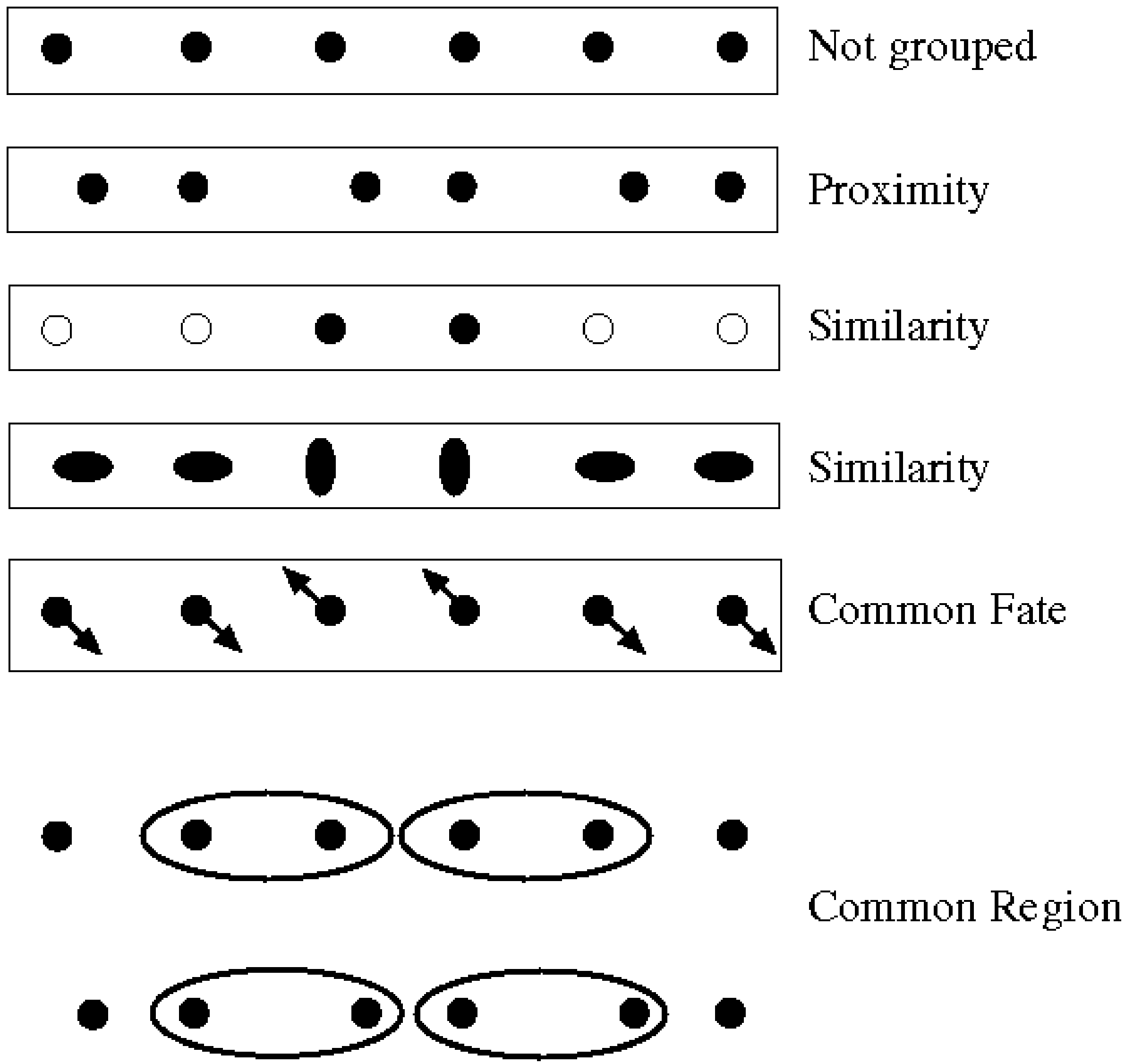
Slide: Derek Hoiem

We perceive the interpretation, not the senses



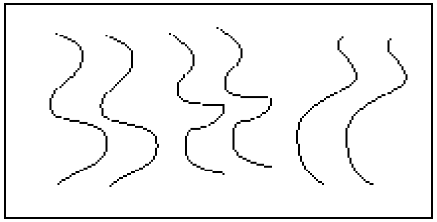
Slide: Derek Hoiem

Principles of perceptual organization

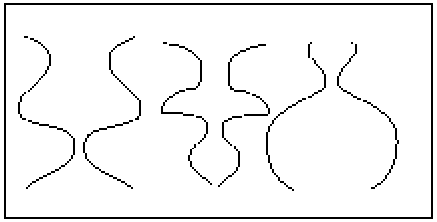


From Steve Lehar: The Constructive Aspect of Visual Perception

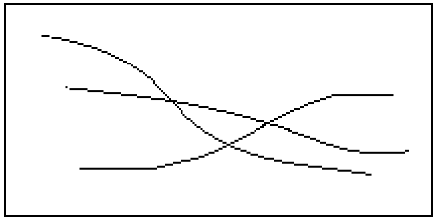
Principles of perceptual organization



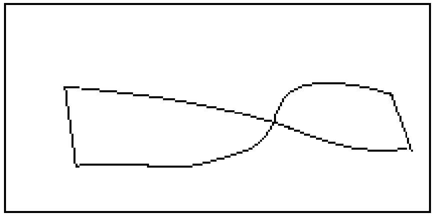
Parallelism



Symmetry



Continuity



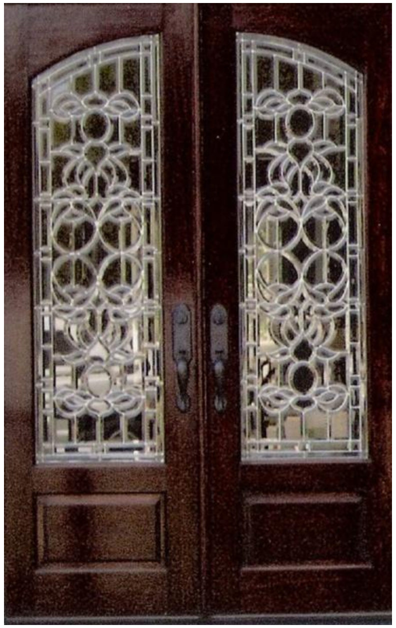
Closure

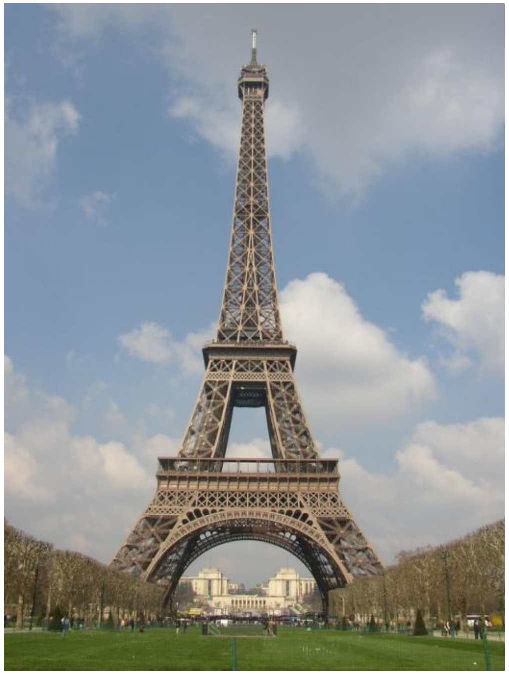
Similarity

http://chicagoist.com/attachments/chicagoist_alicia/GEESE.jpg, http://wwwdelivery.superstock.com/WI/223/1532/PreviewComp/SuperStock_1532R-0831.jpg

Slide: Kristin Grauman

Symmetry





http://seedmagazine.com/news/2006/10/beauty_is_in_the_processingtim.php

Slide: Kristin Grauman

Common fate

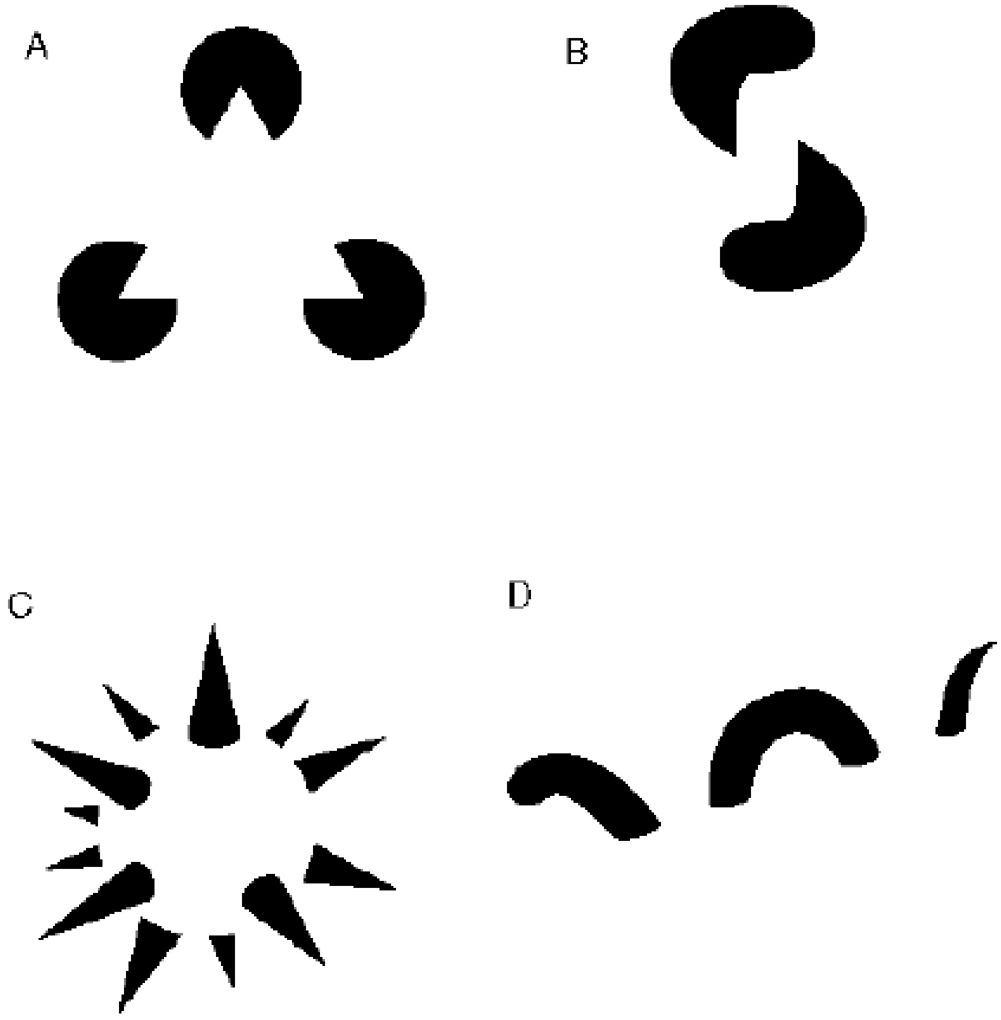
Image credit: Arthus-Bertrand (via F. Durand)

Proximity

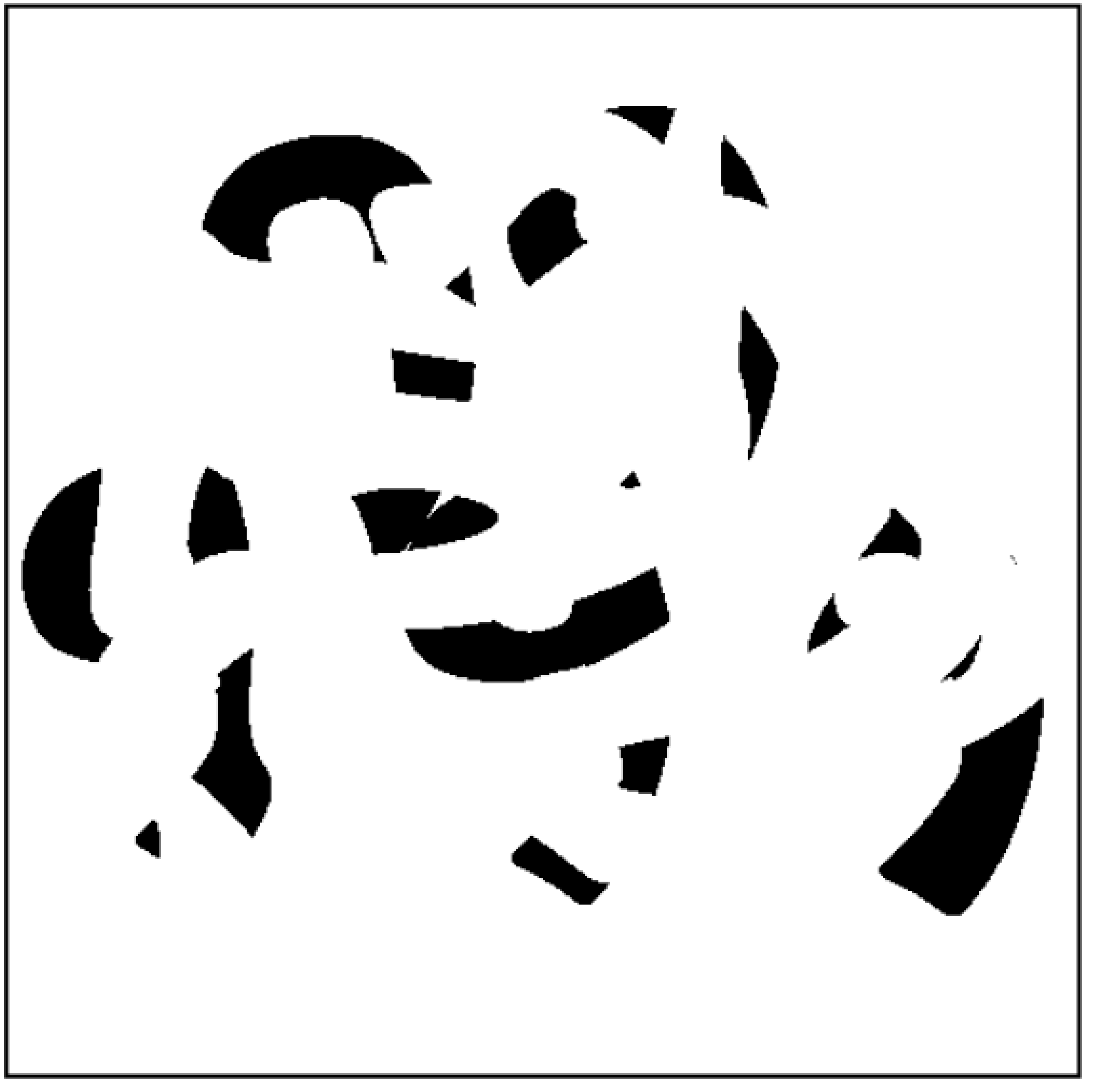
http://www.capital.edu/Resources/Images/outside6_035.jpg

Slide: Kristin Grauman

Grouping by invisible completion



From Steve Lehar: The Constructive Aspect of Visual Perception



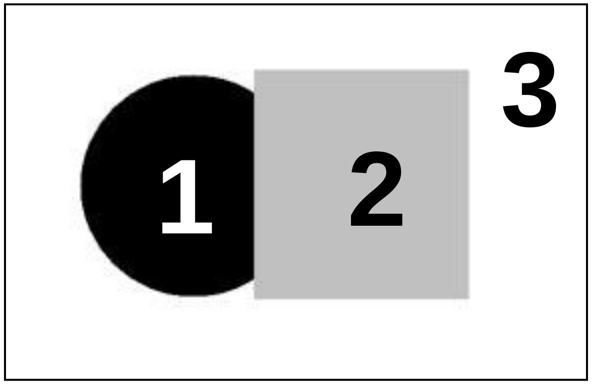
D. Forsyth

Emergence

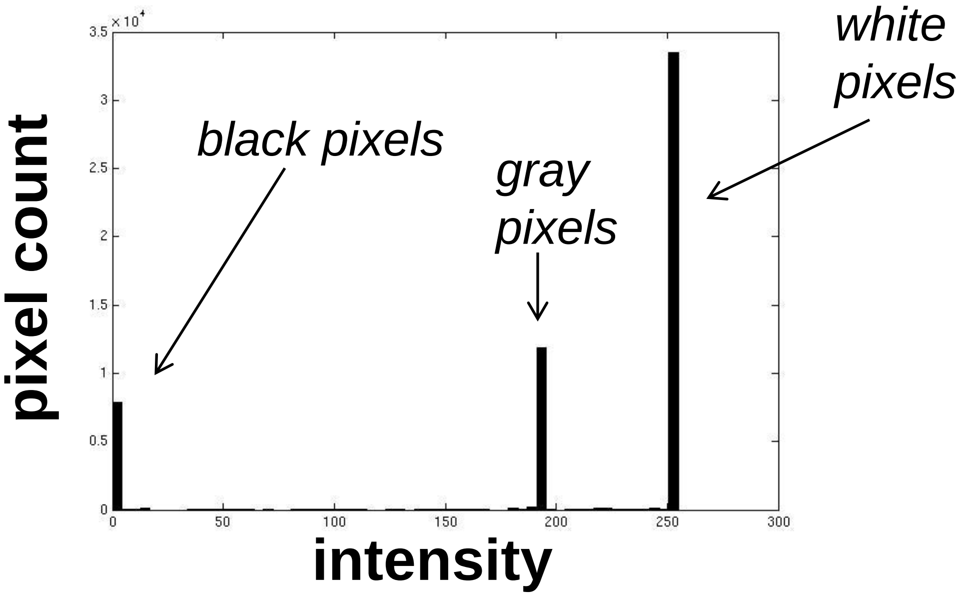
Gestalt cues

- Good intuition and basic principles for grouping
- Basis for many ideas in segmentation and occlusion reasoning
- Some (e.g., symmetry) are difficult to implement in practice

Image segmentation: toy example

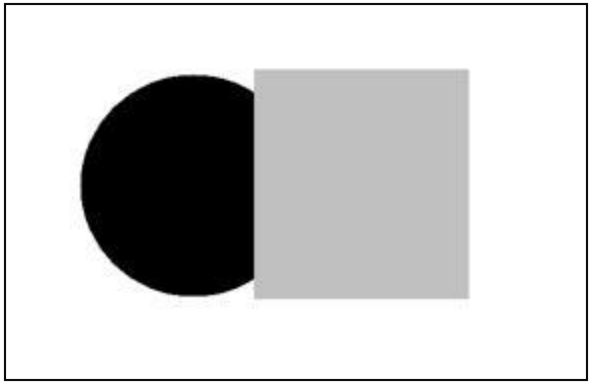


input image

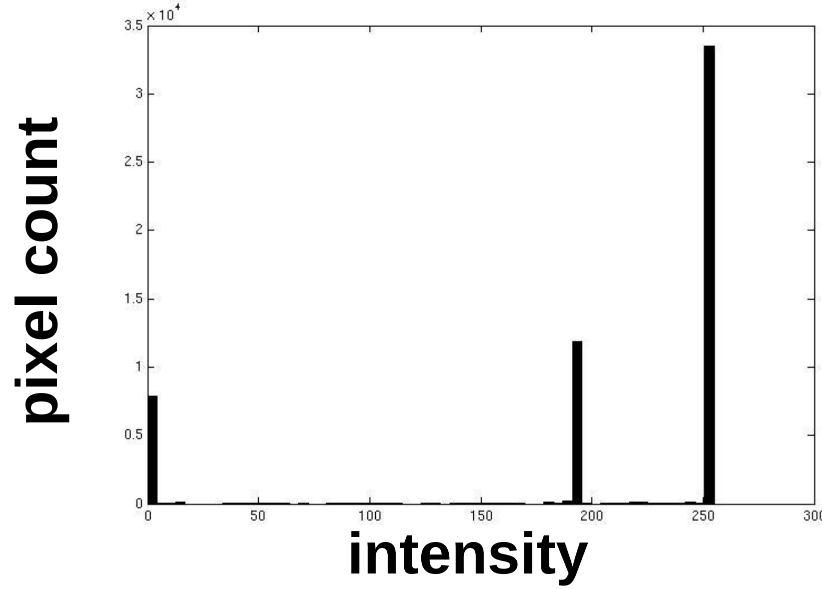


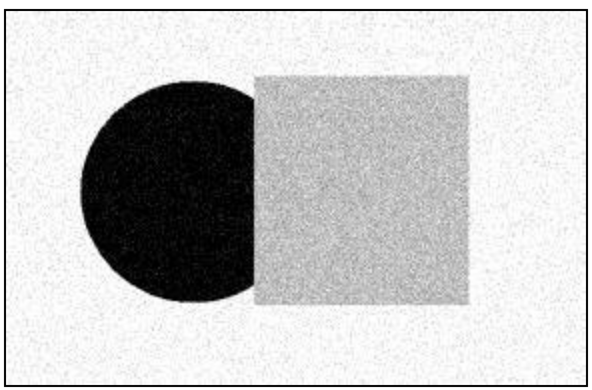
- These intensities define the three groups.
- We could label every pixel in the image according to which of these primary intensities it is.
  - i.e., *segment* the image based on the intensity feature.
- What if the image isn't quite so simple?

Kristen Grauman

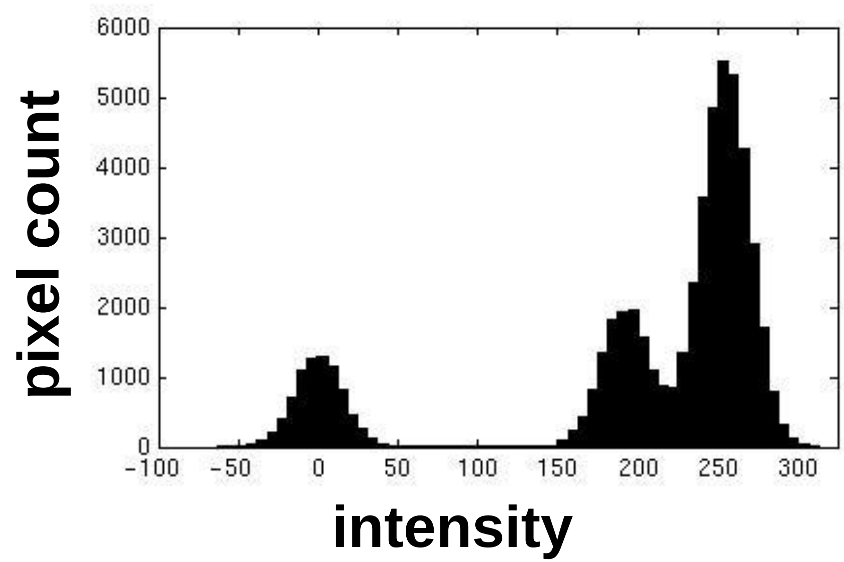


input image

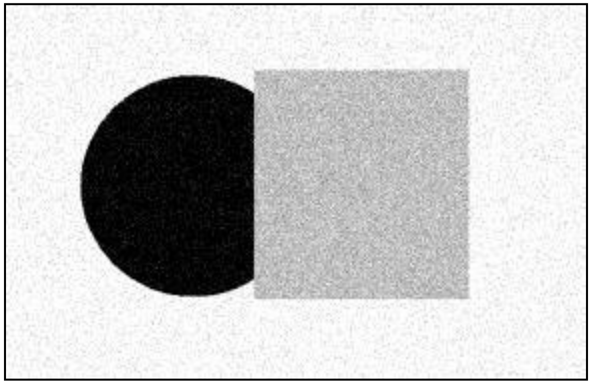




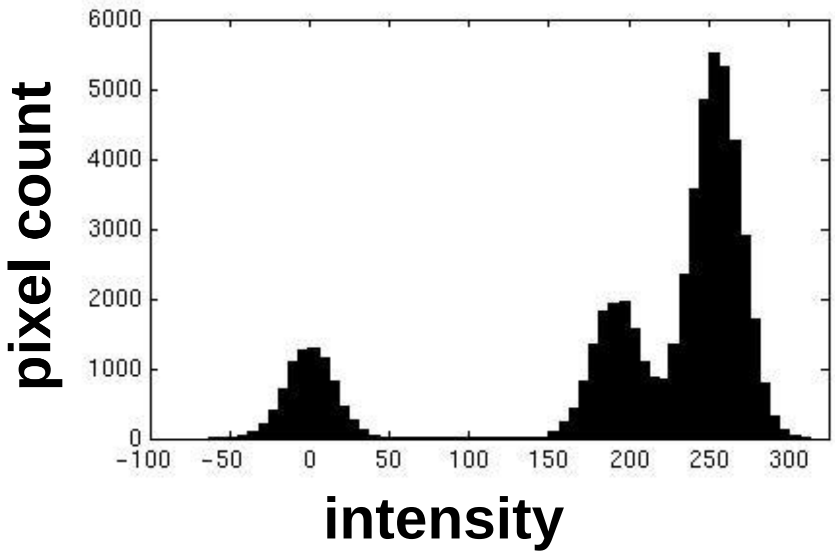
input image



Kristen Grauman



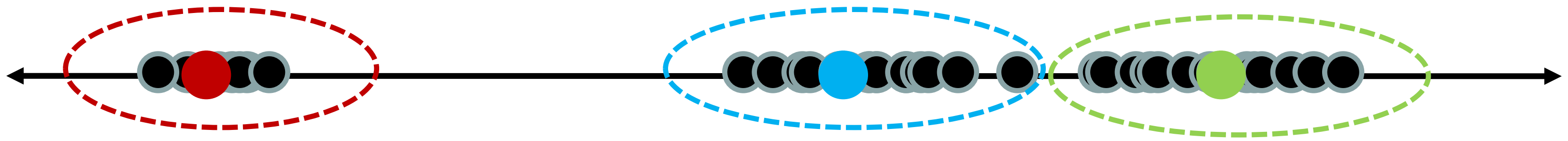
input image



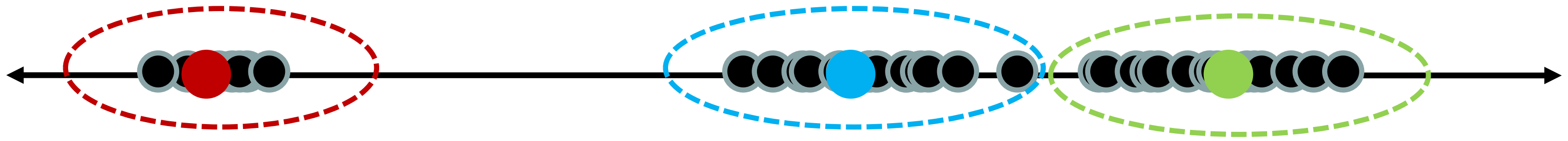
- Now how to determine the three main intensities that define our groups?
- We need to ***cluster.***

Clustering

- With this objective, it is a "chicken and egg" problem:
  - If we knew the **cluster centers**, we could allocate points to groups by assigning each to its closest center.



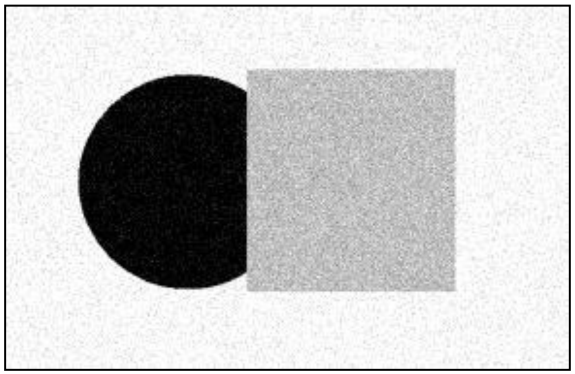
  - If we knew the **group memberships**, we could get the centers by computing the mean per group.



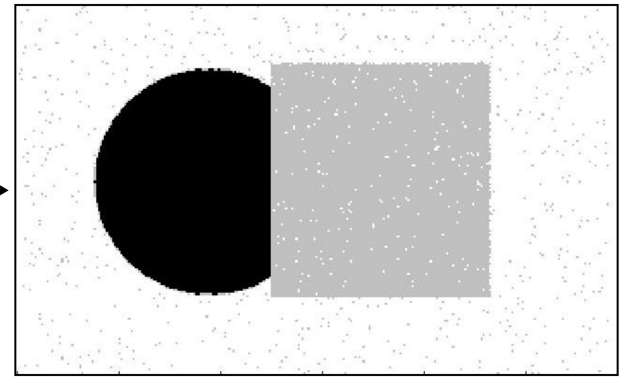
Kristen Grauman

Smoothing out cluster assignments

- Assigning a cluster label per pixel may yield outliers:

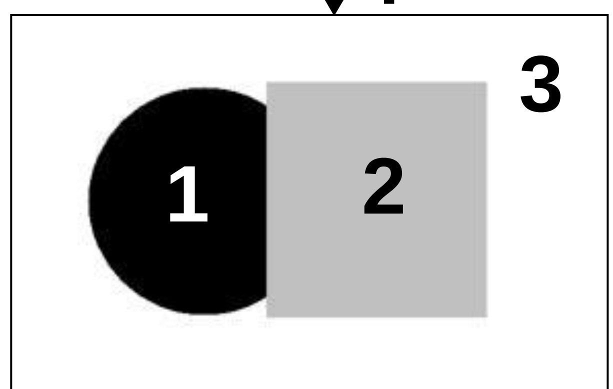


original



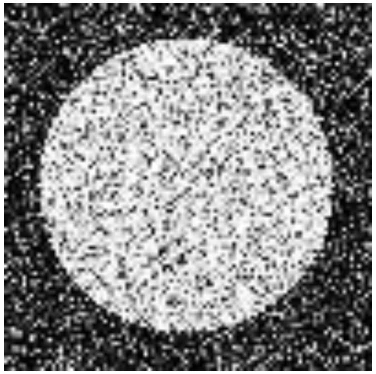
labeled by cluster center's intensity

?



- How to ensure they are spatially smooth?

Solution



P(foreground | image)

Encode dependencies between pixels

Normalizing constant

$$P(\mathbf{y};\theta,data) = \frac{1}{Z}\prod_{i=1..N} f_1(y_i;\theta,data) \prod_{i,j\in edges} f_2(y_i,y_j;\theta,data)$$

Labels to be predicted

Individual predictions

Pairwise predictions

Slide: Derek Hoiem

Writing Likelihood as an "Energy"

$$P(\mathbf{y};\theta,data)=\frac{1}{Z}\prod_{i=1..N}p_1(y_i;\theta,data)\prod_{i,j\in edges}p_2(y_i,y_j;\theta,data)$$

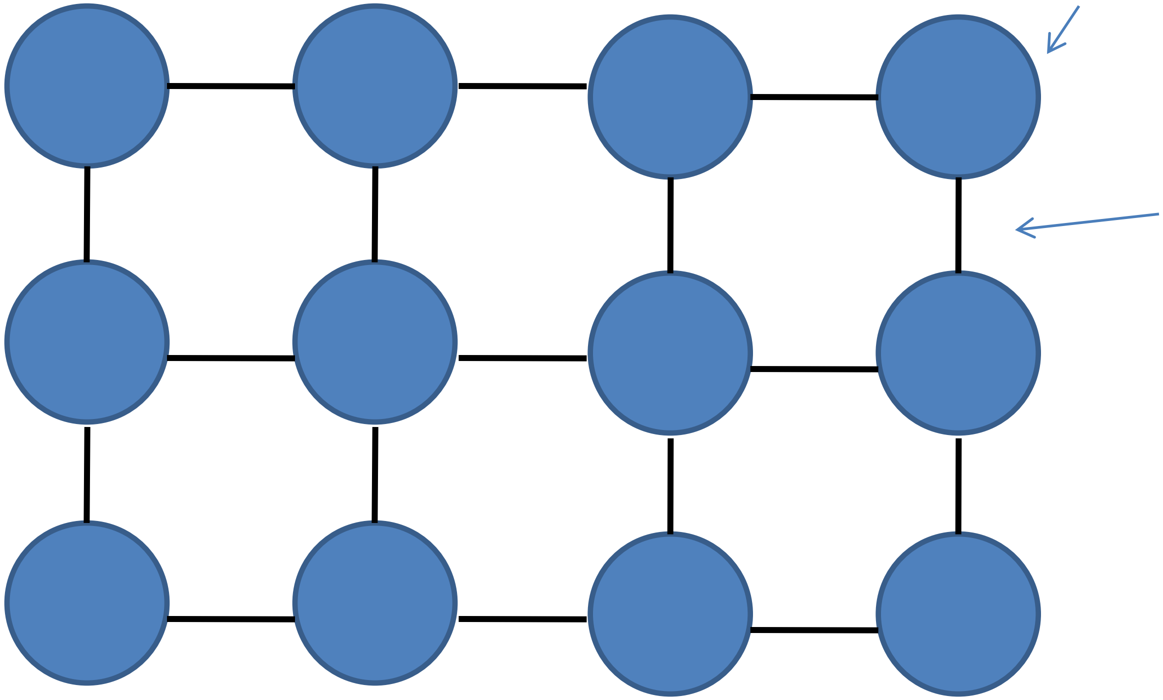
$$Energy(\mathbf{y};\theta,data)=\sum_i\psi_1(y_i;\theta,data)+\sum_{i,j\in edges}\psi_2(y_i,y_j;\theta,data)$$

"Cost" of assignment y_i

"Cost" of pairwise assignment y_i $_,y_j$

Slide: Derek Hoiem

Markov Random Fields



Cost to assign a label to each pixel

Cost to assign a pair of labels to connected pixels

$$Energy(\mathbf{y};\theta,data) = \sum_{i} \psi_1(y_i;\theta,data) + \sum_{i,j \in edges} \psi_2(y_i,y_j;\theta,data)$$

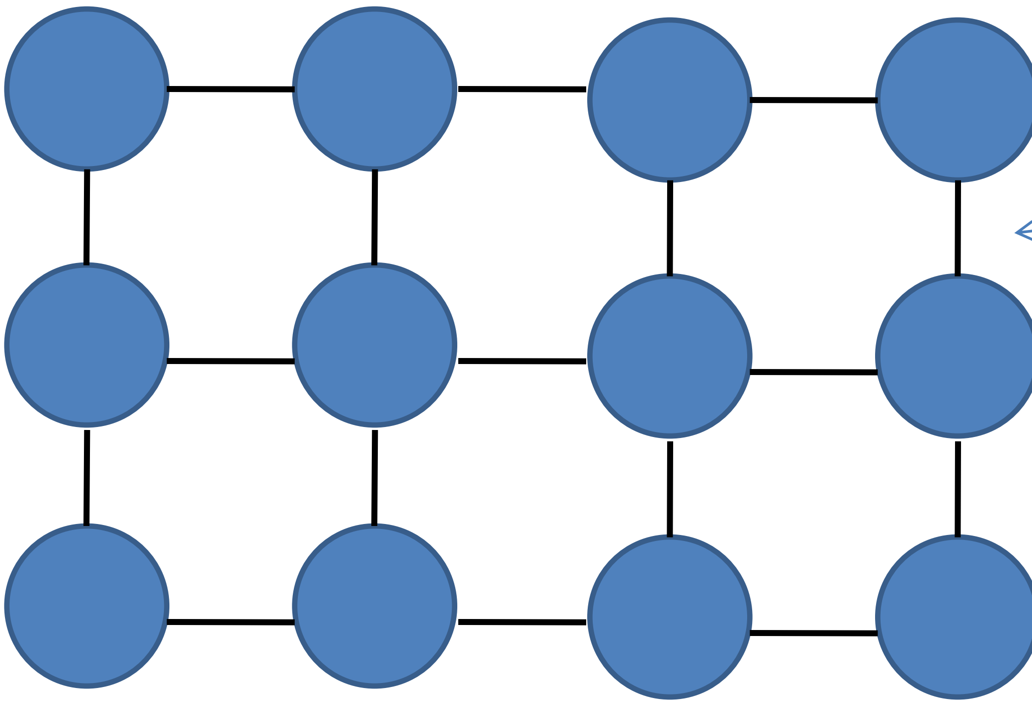
Slide: Derek Hoiem

Markov Random Fields

- Example: “label smoothing” grid

Unary potential

0: -logP(y_i = 0 ; data)
1: -logP(y_i = 1 ; data)



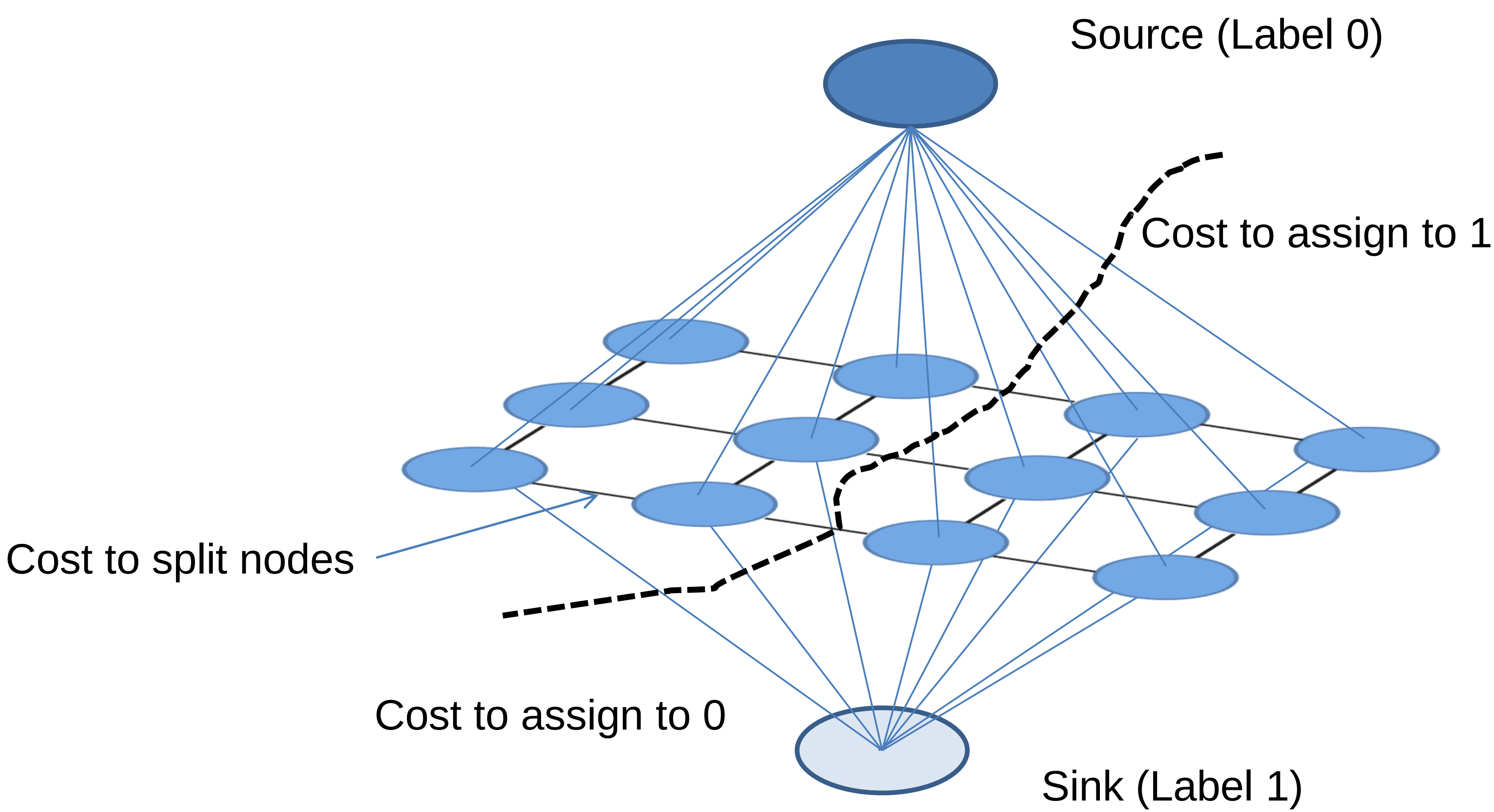
Pairwise Potential

	0	1
0	0	K
1	K	0

$$Energy(\mathbf{y};\theta,data) = \sum_i \psi_1(y_i;\theta,data) + \sum_{i,j \in edges} \psi_2(y_i,y_j;\theta,data)$$

Slide: Derek Hoiem

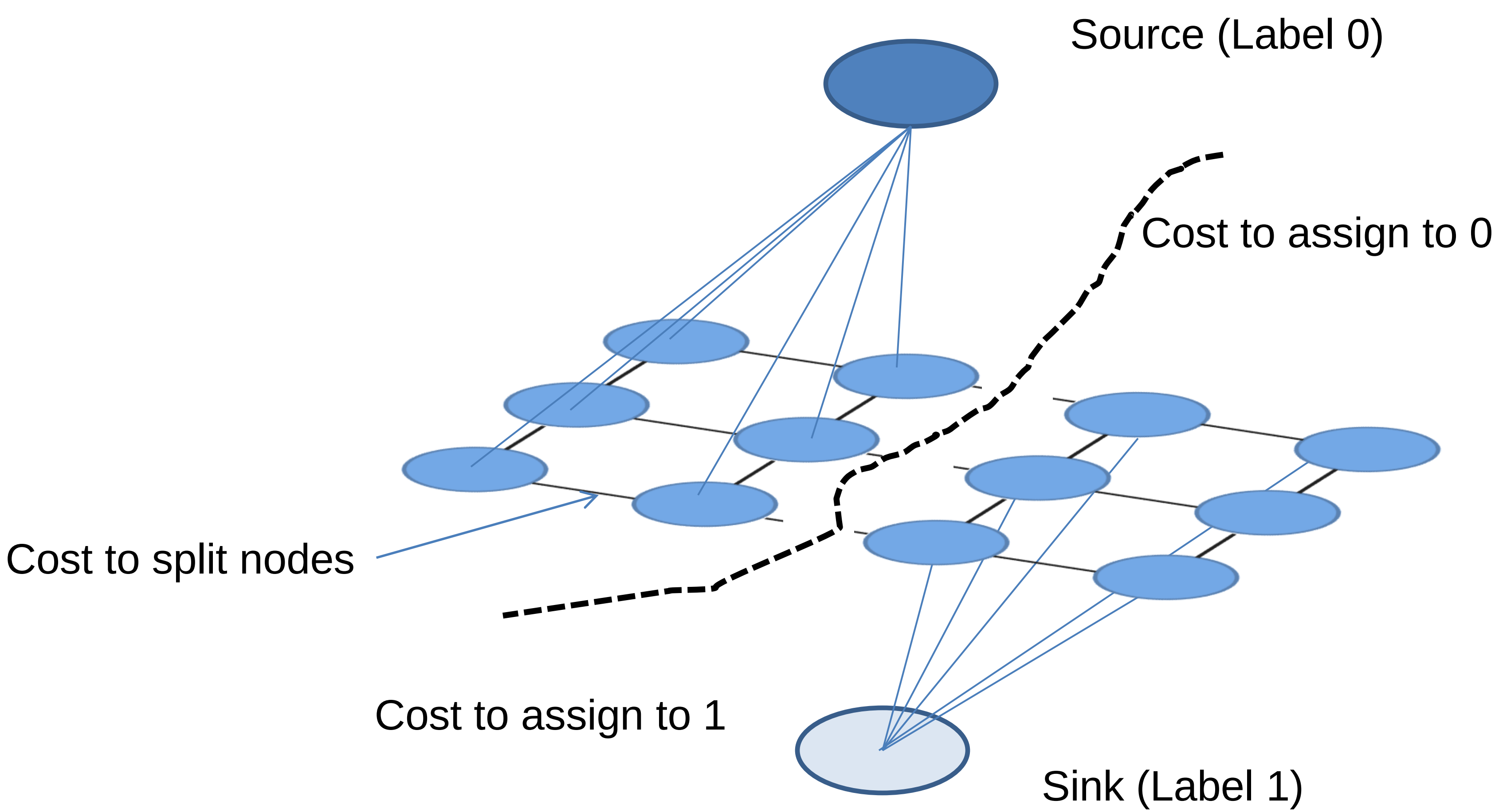
Solving MRFs with graph cuts



$$Energy(\mathbf{y};\theta,data) = \sum_i \psi_1(y_i;\theta,data) + \sum_{i,j\in edges} \psi_2(y_i,y_j;\theta,data)$$

Slide: Derek Hoiem

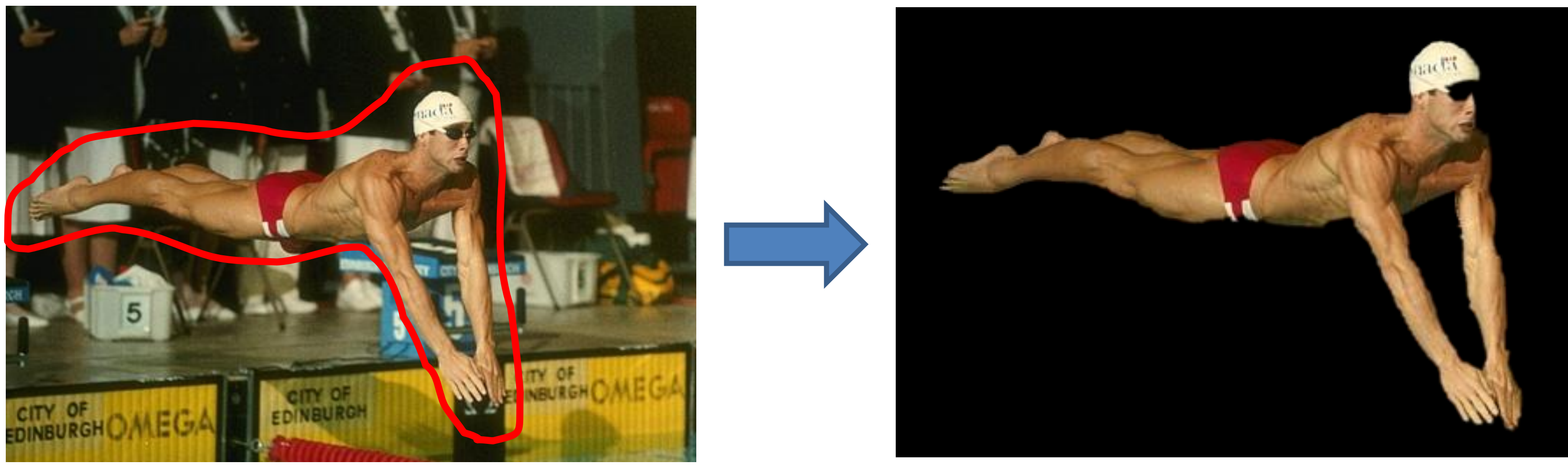
Solving MRFs with graph cuts



$$Energy(\mathbf{y};\theta,data)=\sum_i \psi_1(y_i;\theta,data)+\sum_{i,j\in edges}\psi_2(y_i,y_j;\theta,data)$$

Slide: Derek Hoiem

GrabCut segmentation



User provides rough indication of foreground region.

Goal: Automatically provide a pixel-level segmentation.

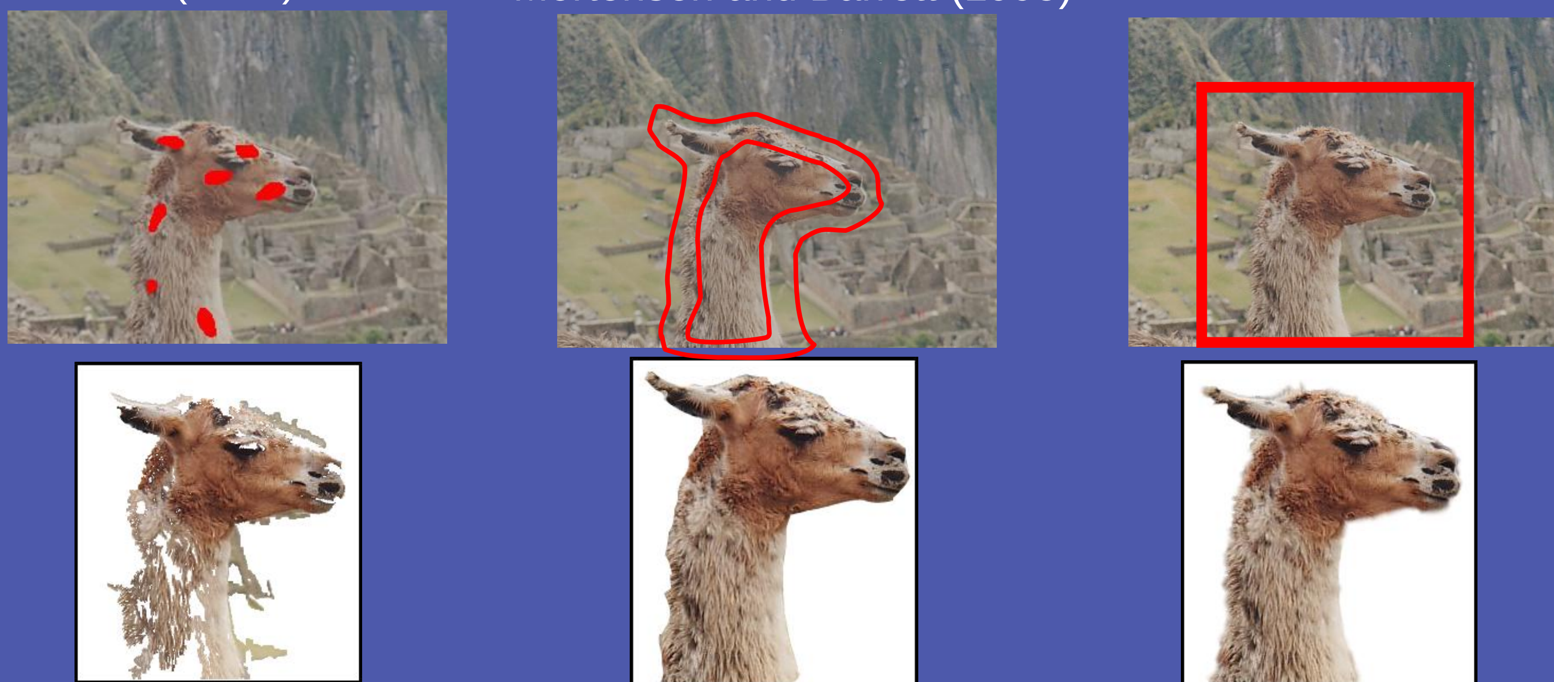
Slide: Derek Hoiem

Grab cuts and graph cuts

	Magic Wand (198?)	Intelligent Scissors Mortensen and Barrett (1995)	GrabCut

User Input

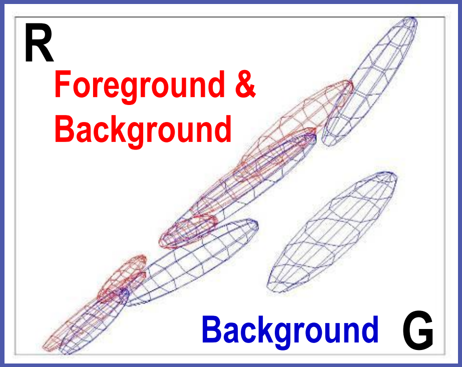
Result



Regions — Boundary — Regions & Boundary

Source: Rother

Colour Model

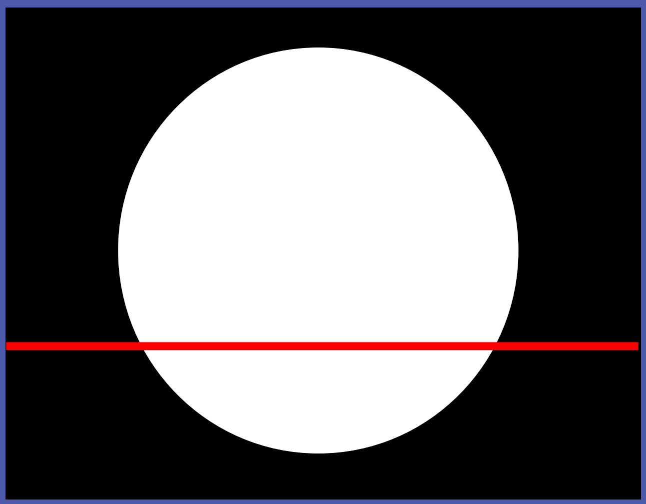


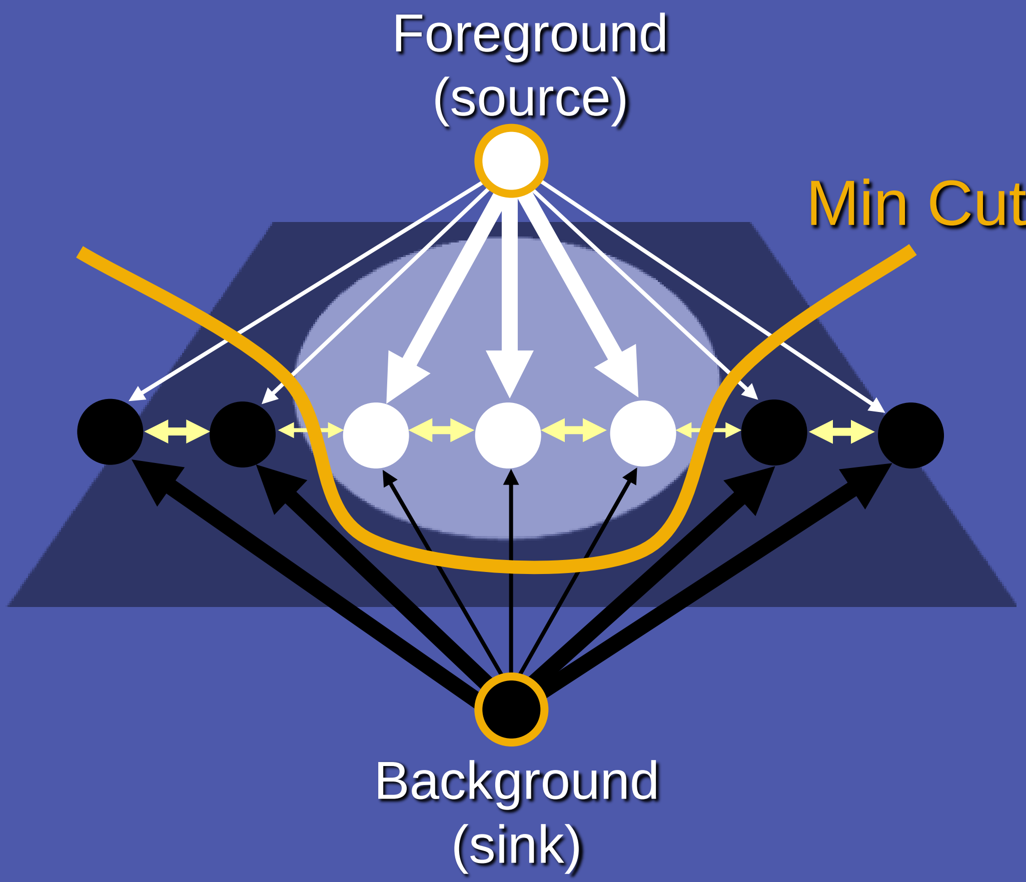
Gaussian Mixture Model (typically 5-8 components)

Source: Rother

Graph cuts

Boykov and Jolly (2001)





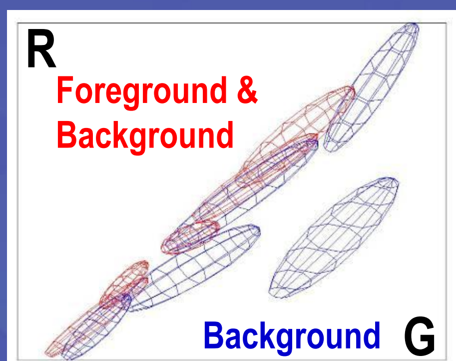
Cut: separating source and sink; Energy: collection of edges

Min Cut: Global minimal enegry in polynomial time

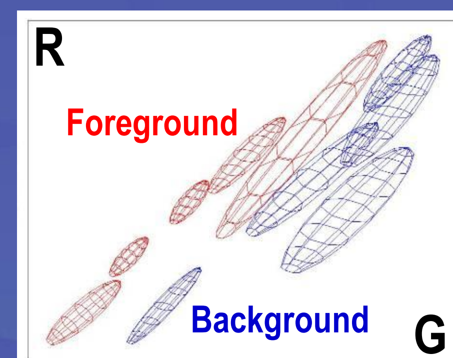
Source: Rother

Colour Model

SIGGRAPH2004



Iterated graph cut



Gaussian Mixture Model (typically 5-8 components)

Source: Rother

GrabCut segmentation

1. Define graph
   – usually 4-connected or 8-connected
2. Define unary potentials
   – Color histogram or mixture of Gaussians for background and foreground

$$unary\_potential(x) = -\log\left(\frac{P(c(x);\theta_{foreground})}{P(c(x);\theta_{background})}\right)$$

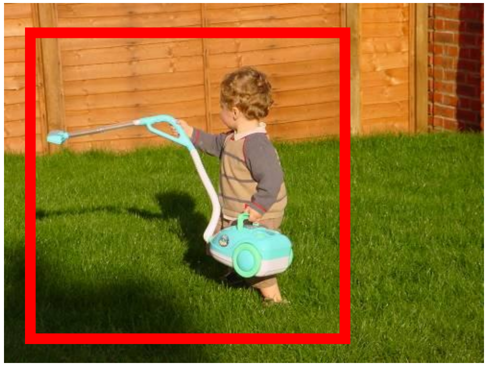
3. Define pairwise potentials

$$edge\_potential(x,y) = k_1 + k_2 \exp\left\{\frac{-\left\|c(x)-c(y)\right\|^2}{2\sigma^2}\right\}$$

4. Apply graph cuts
5. Return to 2, using current labels to compute foreground, background models

Slide: Derek Hoiem

What is easy or hard about these cases for graphcut-based segmentation?



Easier examples

More difficult Examples

Camouflage & Low Contrast

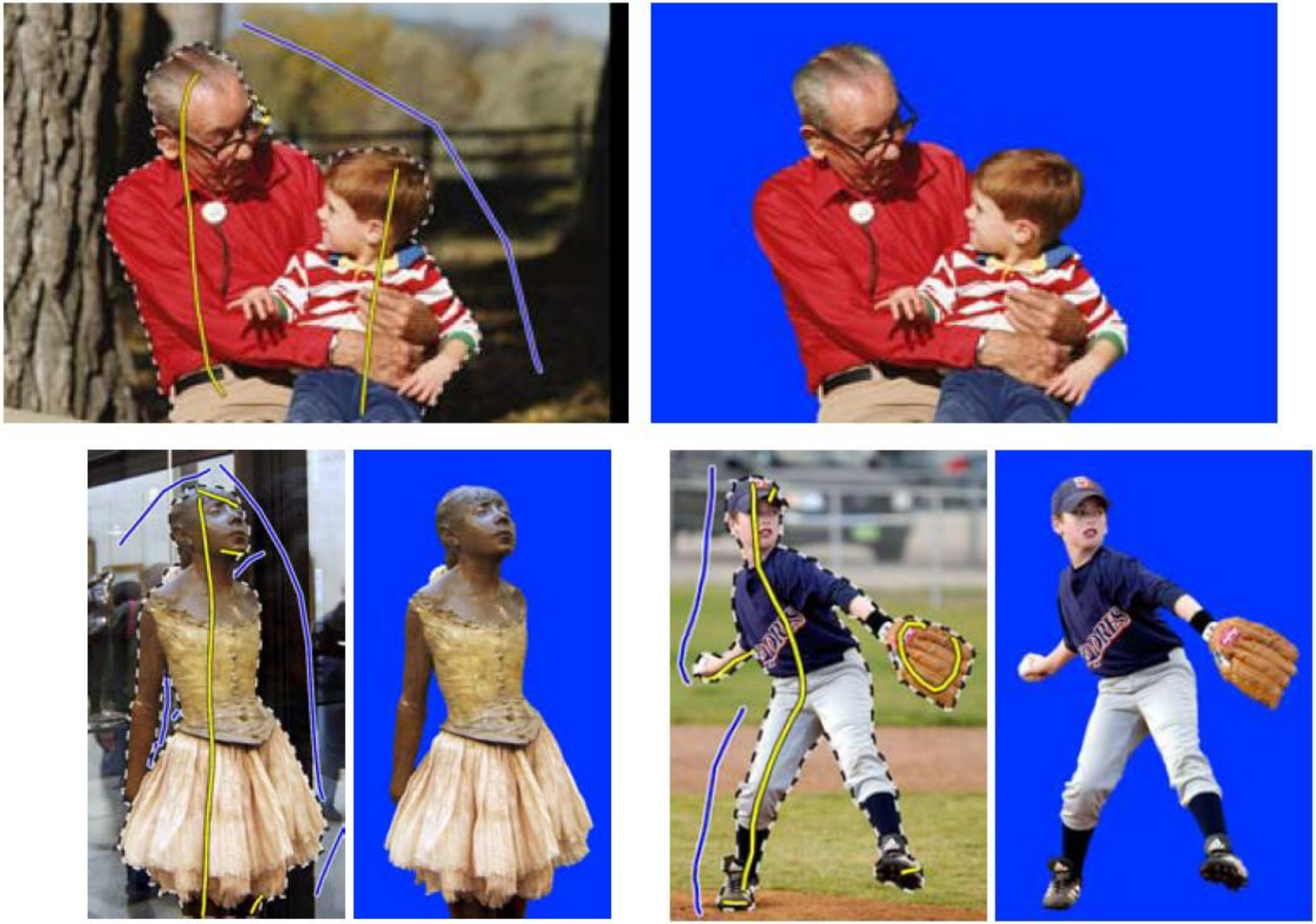
Fine structure

Harder Case

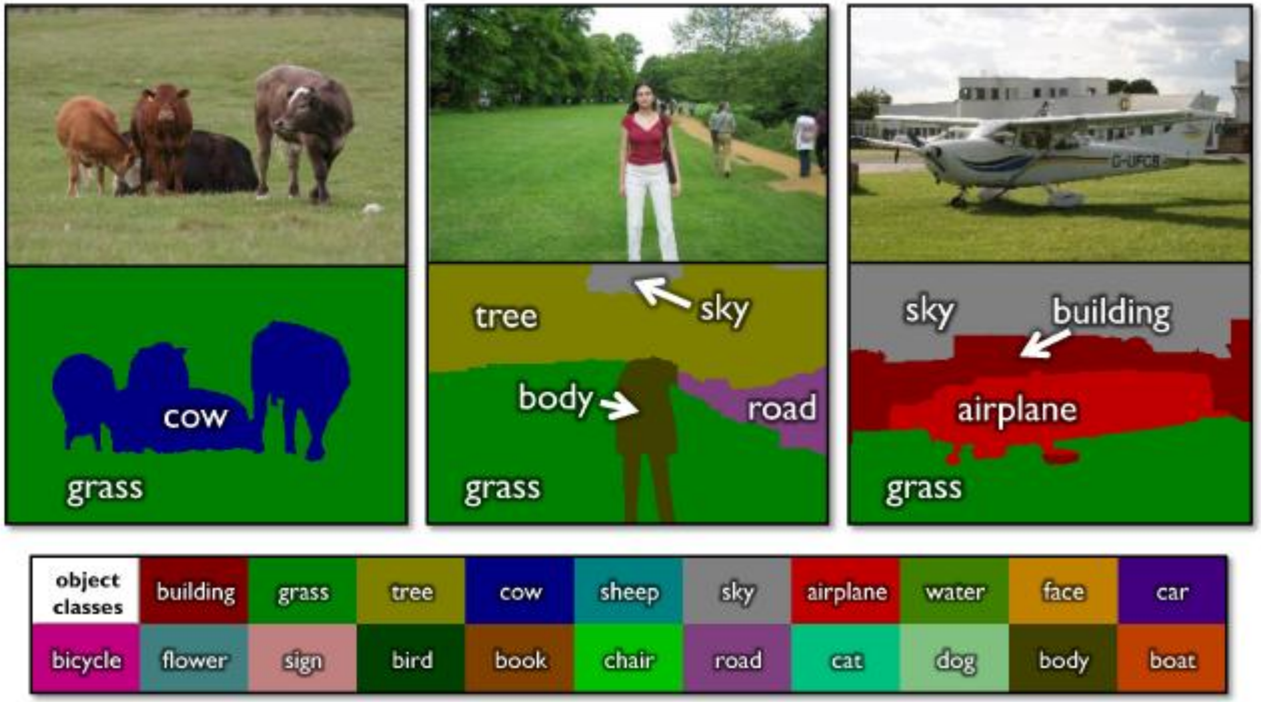
Initial Rectangle

Initial Result

Lazy Snapping (Li et al. SG 2004)

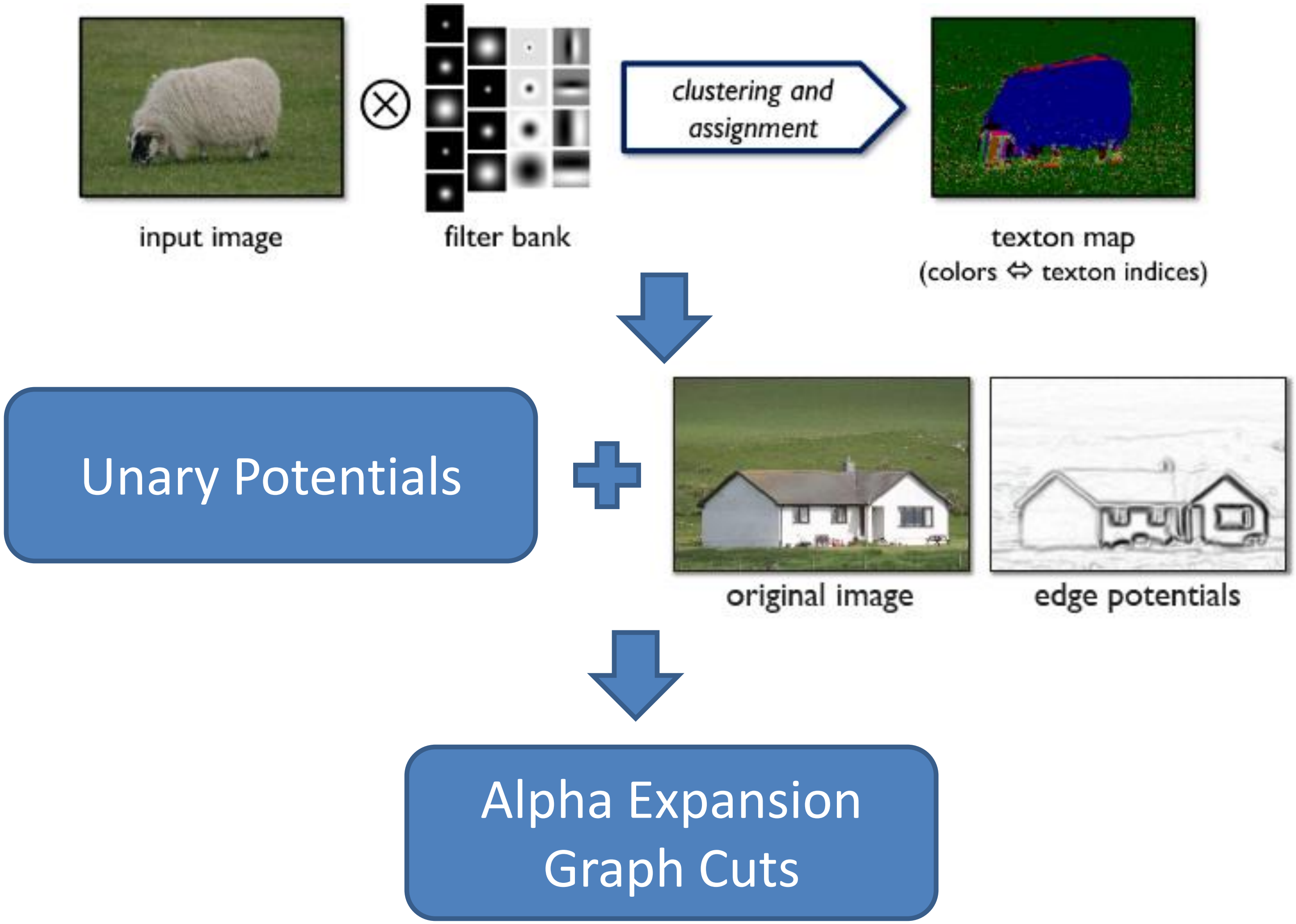


Using graph cuts for recognition



TextonBoost (Shotton et al. 2009 IJCV)

Using graph cuts for recognition



TextonBoost (Shotton et al. 2009 IJCV)

Limitations of graph cuts

- Associative: edge potentials penalize different labels

Must satisfy

$$E^{i,j}(0,0) + E^{i,j}(1,1) \leq E^{i,j}(0,1) + E^{i,j}(1,0)$$

- If not associative, can sometimes clip potentials

- Approximate for multilabel
  - Alpha-expansion or alpha-beta swaps

Slide: Derek Hoiem

Graph cuts: Pros and Cons

- Pros
  - Very fast inference
  - Can incorporate data likelihoods and priors
  - Applies to a wide range of problems (stereo, image labeling, recognition)
- Cons
  - Not always applicable (associative only)
  - Need unary terms (not used for generic segmentation)
- Use whenever applicable

Slide: Derek Hoiem

More about MRFs/CRFs

- Other common uses
  - Graph structure on regions
  - Encoding relations between multiple scene elements
- Inference methods
  - Loopy BP or BP-TRW: approximate, slower, but works for more general graphs

Slide: Derek Hoiem

Further reading and resources

- Graph cuts
  - http://www.cs.cornell.edu/~rdz/graphcuts.html
  - Classic paper: What Energy Functions can be Minimized via Graph Cuts? (Kolmogorov and Zabih, ECCV '02/PAMI '04)

- Belief propagation

  Yedidia, J.S.; Freeman, W.T.; Weiss, Y., "Understanding Belief Propagation and Its Generalizations", Technical Report, 2001: http://www.merl.com/publications/TR2001-022/

- Normalized cuts and image segmentation (Shi and Malik)

  *http://www.cs.berkeley.edu/~malik/papers/SM-n**cut**.pdf*

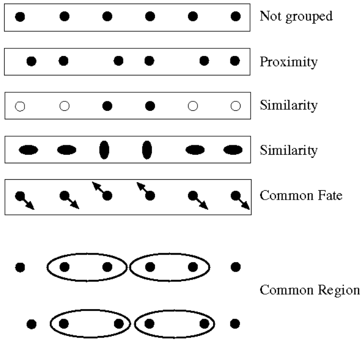
- N-cut implementation

  http://www.seas.upenn.edu/~timothee/software/ncut/ncut.html

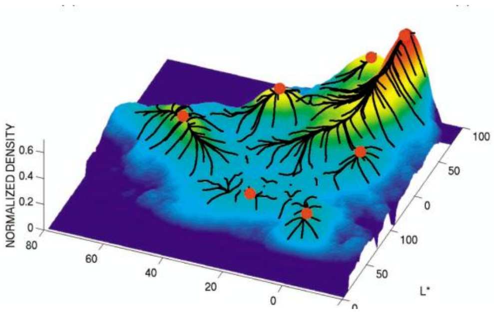
Slide: Derek Hoiem

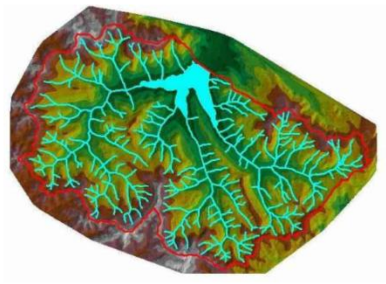
Next Class

- Gestalt grouping



- More segmentation methods



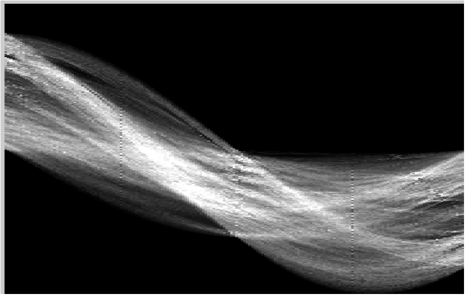


Recap of Grouping and Fitting

Edge and line detection

- Canny edge detector = smooth → derivative → thin → threshold → link

- Generalized Hough transform = points vote for shape parameters



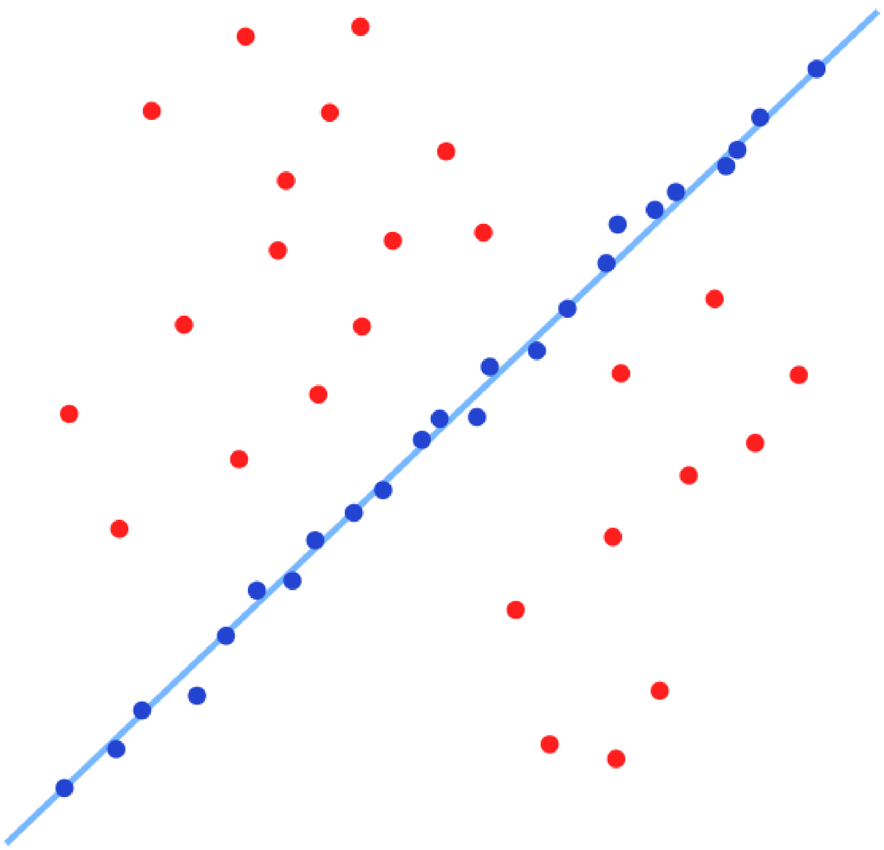
- Straight line detector = canny + gradient orientations → orientation binning → linking → check for straightness

Slide: Derek Hoiem

Robust fitting and registration

Key algorithms

- RANSAC, Hough Transform

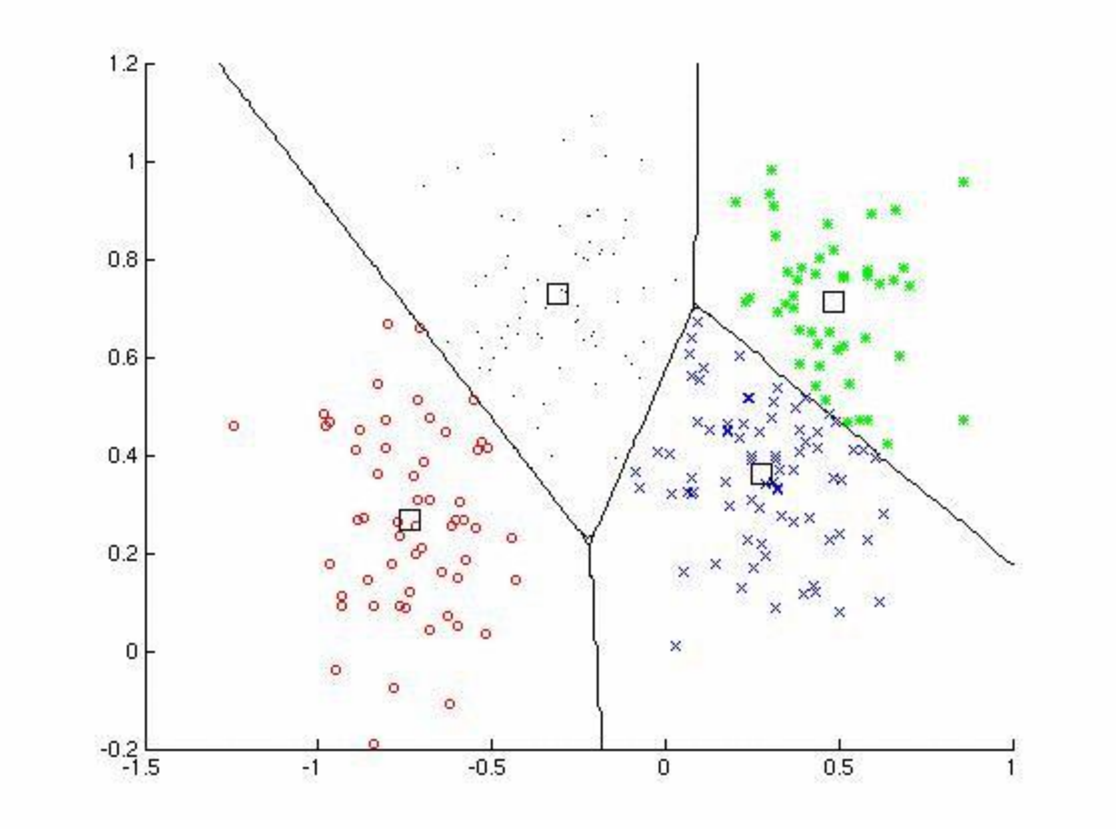


Slide: Derek Hoiem

Clustering

Key algorithm

- K-means



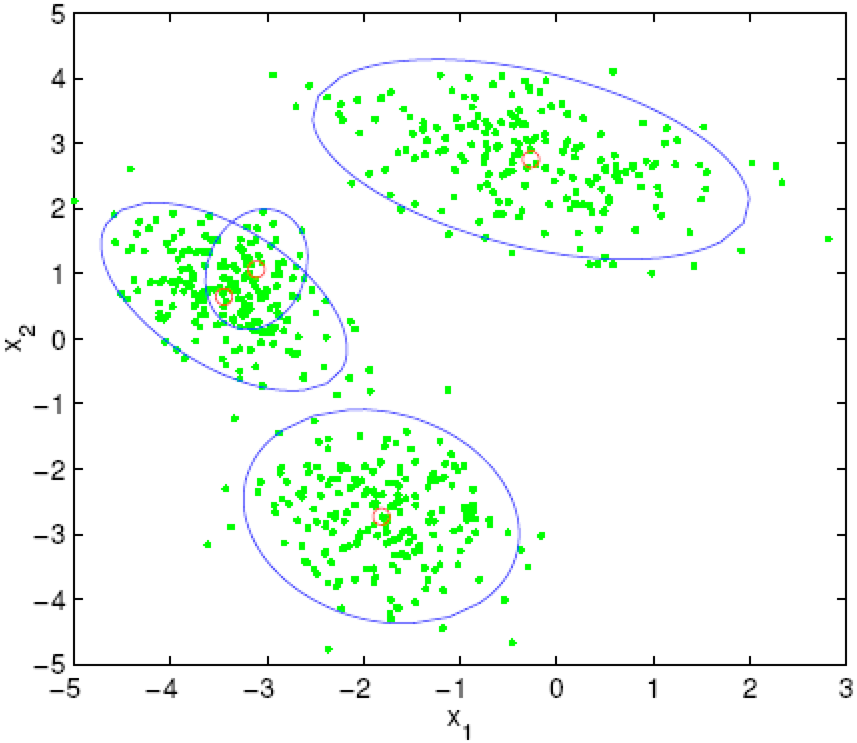
Slide: Derek Hoiem

EM and Mixture of Gaussians

Tutorials:

*http://www.cs.duke.edu/courses/spring04/cps196.1/.../**EM**/tomasiEM.pdf*

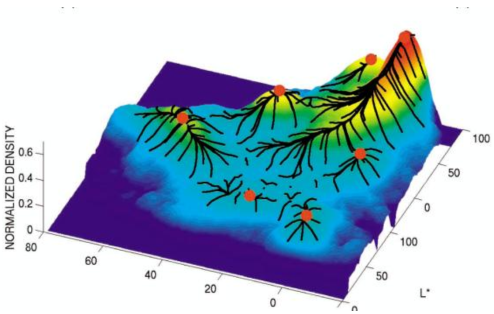
http://www-clmc.usc.edu/~adsouza/notes/mix_gauss.pdf



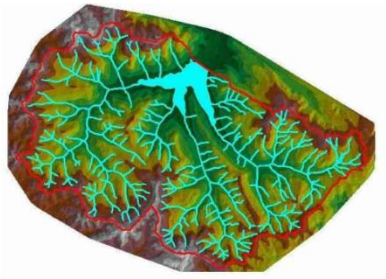
Slide: Derek Hoiem

Segmentation

- Mean-shift segmentation
  - Flexible clustering method, good segmentation



- Watershed segmentation
  - Hierarchical segmentation from soft boundaries



- Normalized cuts
  - Produces regular regions
  - Slow but good for oversegmentation

- MRFs with Graph Cut
  - Incorporates foreground/background/object model and prefers to cut at image boundaries
  - Good for interactive segmentation or recognition

Slide: Derek Hoiem

Next section: Recognition

- How to recognize
  - Specific object instances
  - Faces
  - Scenes
  - Object categories

Slide: Derek Hoiem