Segmentation

Lecture 12

Many slides from: S. Lazebnik, K. Grauman and P. Kumar

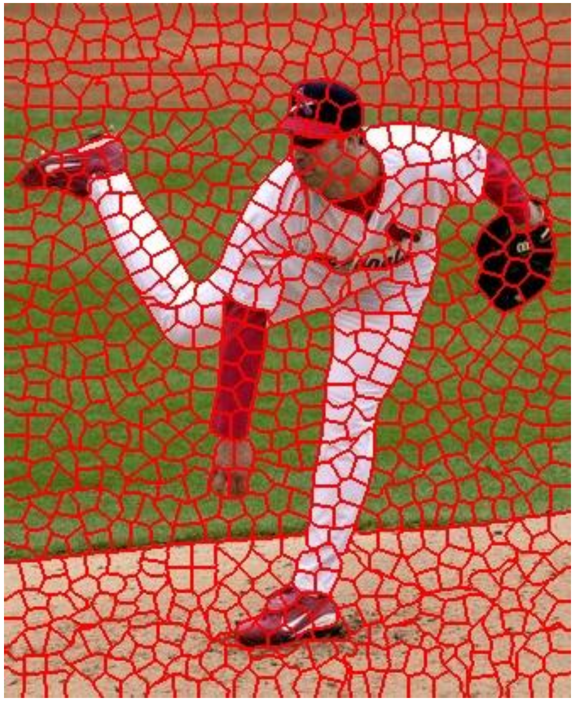
Image Segmentation

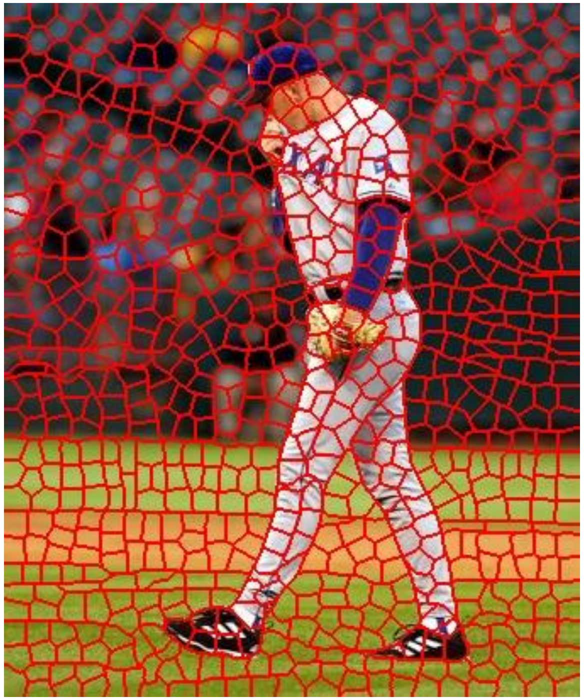
Image segmentation

The goals of segmentation

- Group together similar-looking pixels for efficiency of further processing
  - “Bottom-up” process
  - Unsupervised

“superpixels”





X. Ren and J. Malik. **[Learning a classification model for segmentation.](#)** ICCV 2003.

Slide credit: S. Lazebnik

The goals of segmentation

- Separate image into coherent “objects”
  - “Bottom-up” or “top-down” process?
  - Supervised or unsupervised?

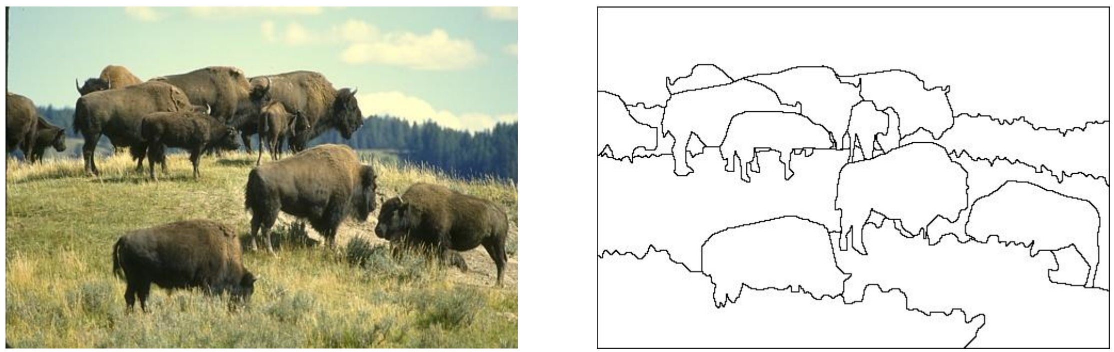
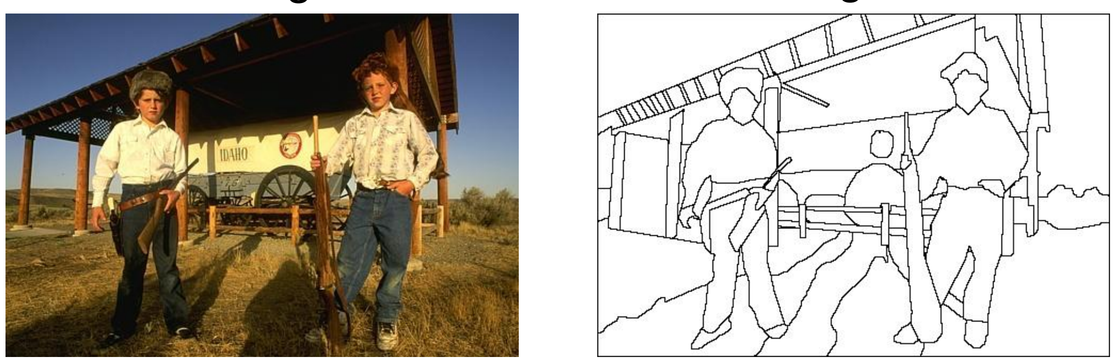


image human segmentation



Berkeley segmentation database:

http://www.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/segbench/

Slide credit: S. Lazebnik

Emergence

http://en.wikipedia.org/wiki/Gestalt_psychology

Overview

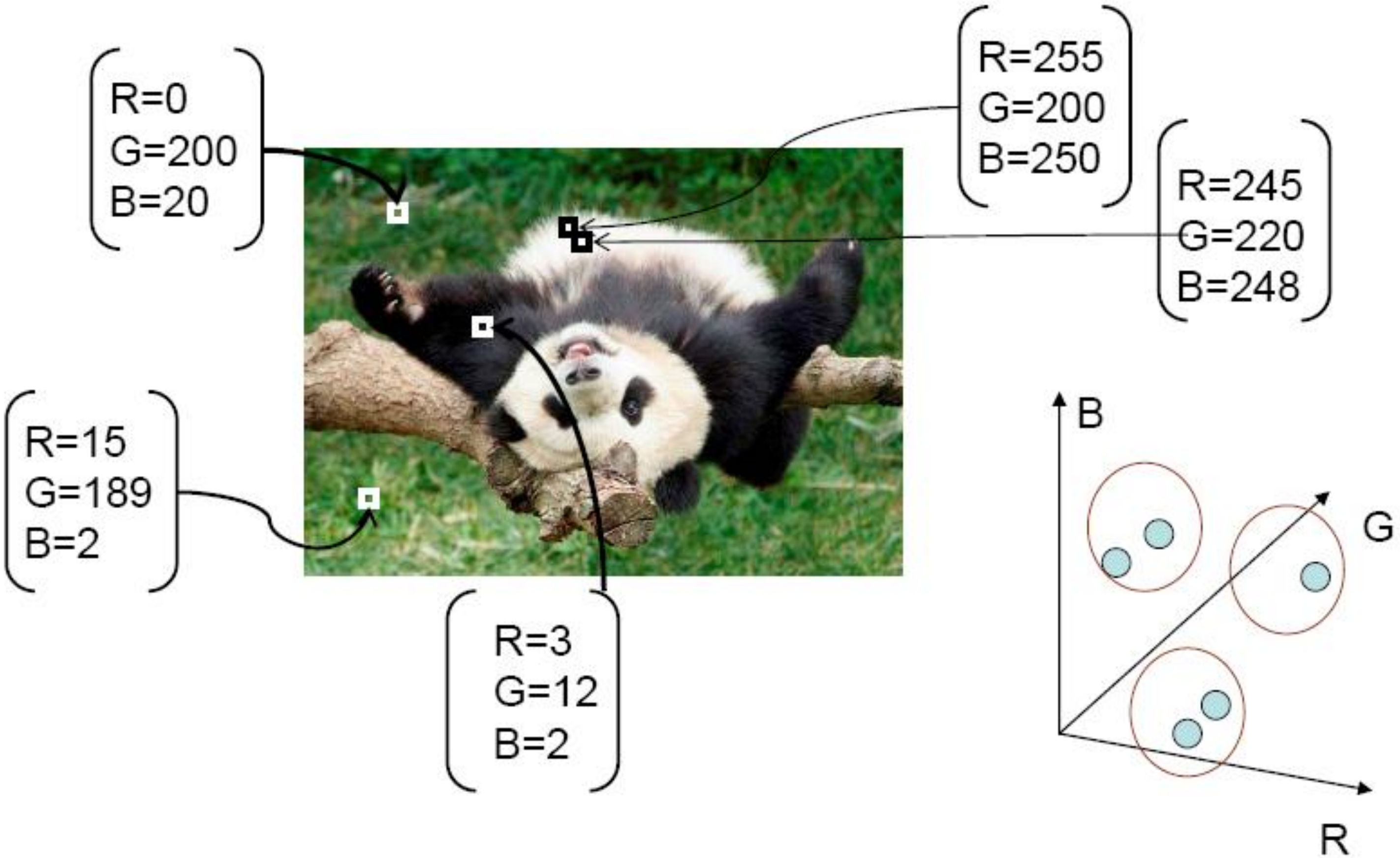
- Bottom-up segmentation
  - Clustering
  - Mean shift
  - Graph-based
- Combining object recognition & segmentation
  - OBJCUT
  - Other methods

Overview

- Bottom-up segmentation
  - Clustering
  - Mean shift
  - Graph-based
- Combining object recognition & segmentation
  - OBJCUT
  - Other methods

Segmentation as clustering

- Cluster similar pixels (features) together



Source: K. Grauman

Segmentation as clustering

- K-means clustering based on intensity or color is essentially vector quantization of the image attributes
  - Clusters don't have to be spatially coherent

Image

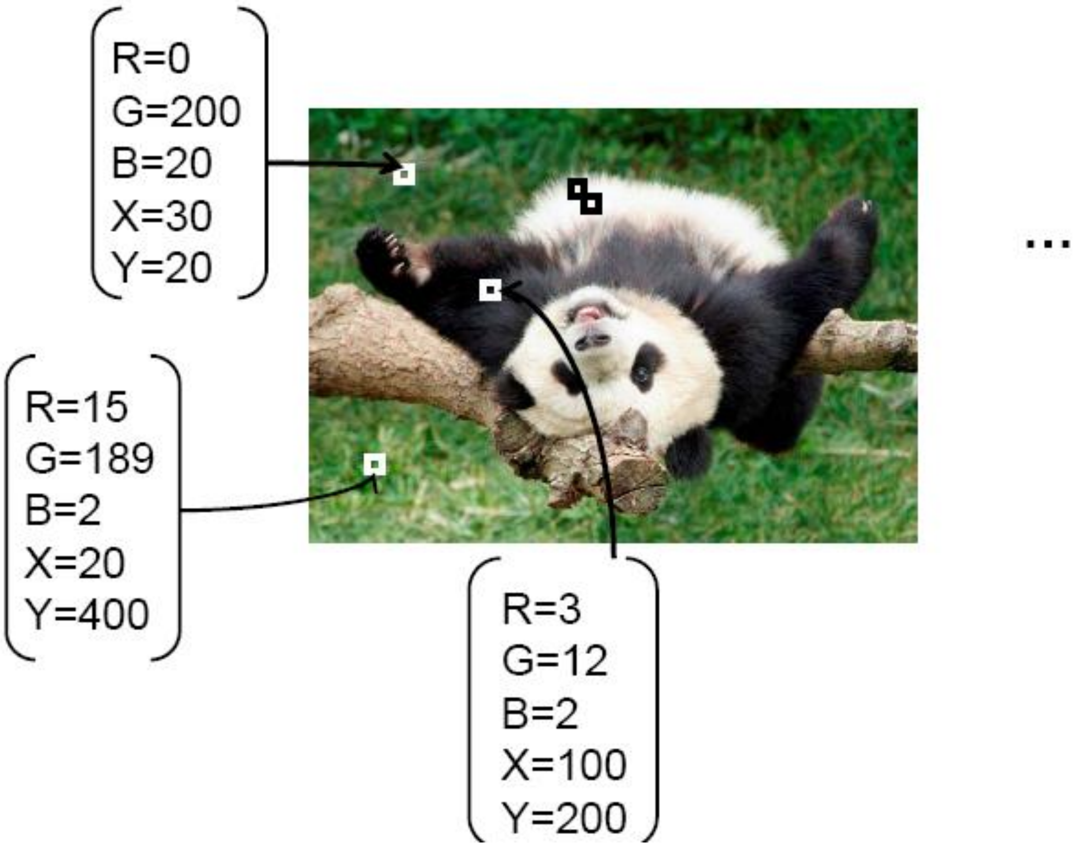
Intensity-based clusters

Color-based clusters

Slide credit: S. Lazebnik

Segmentation as clustering

- Cluster similar pixels (features) together



Source: K. Grauman

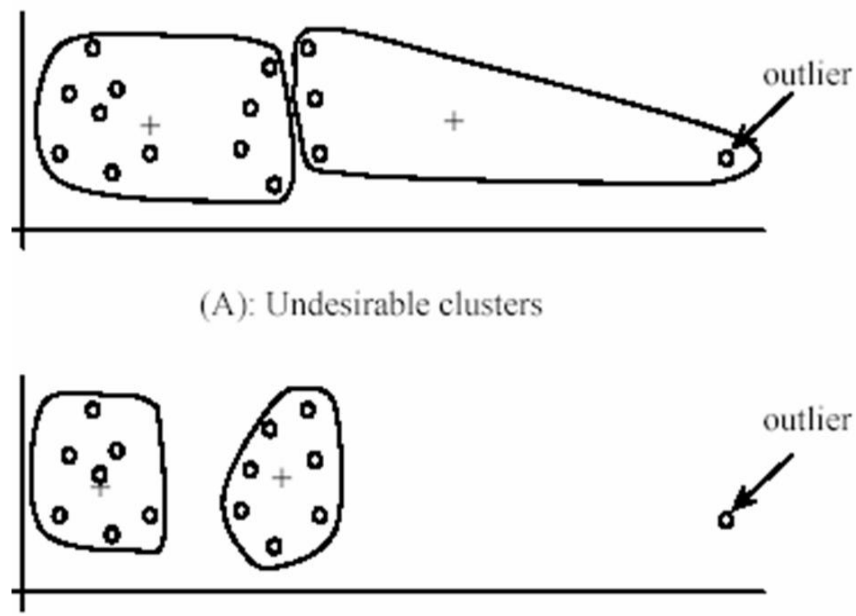
Segmentation as clustering

- Clustering based on (r,g,b,x,y) values enforces more spatial coherence

Slide credit: S. Lazebnik

K-Means for segmentation

- Pros
  - Very simple method
  - Converges to a local minimum of the error function
- Cons
  - Memory-intensive
  - Need to pick K
  - Sensitive to initialization
  - Sensitive to outliers
  - Only finds “spherical” clusters



(A): Undesirable clusters

(B): Ideal clusters

Slide credit: S. Lazebnik

Overview

Bottom-up segmentation

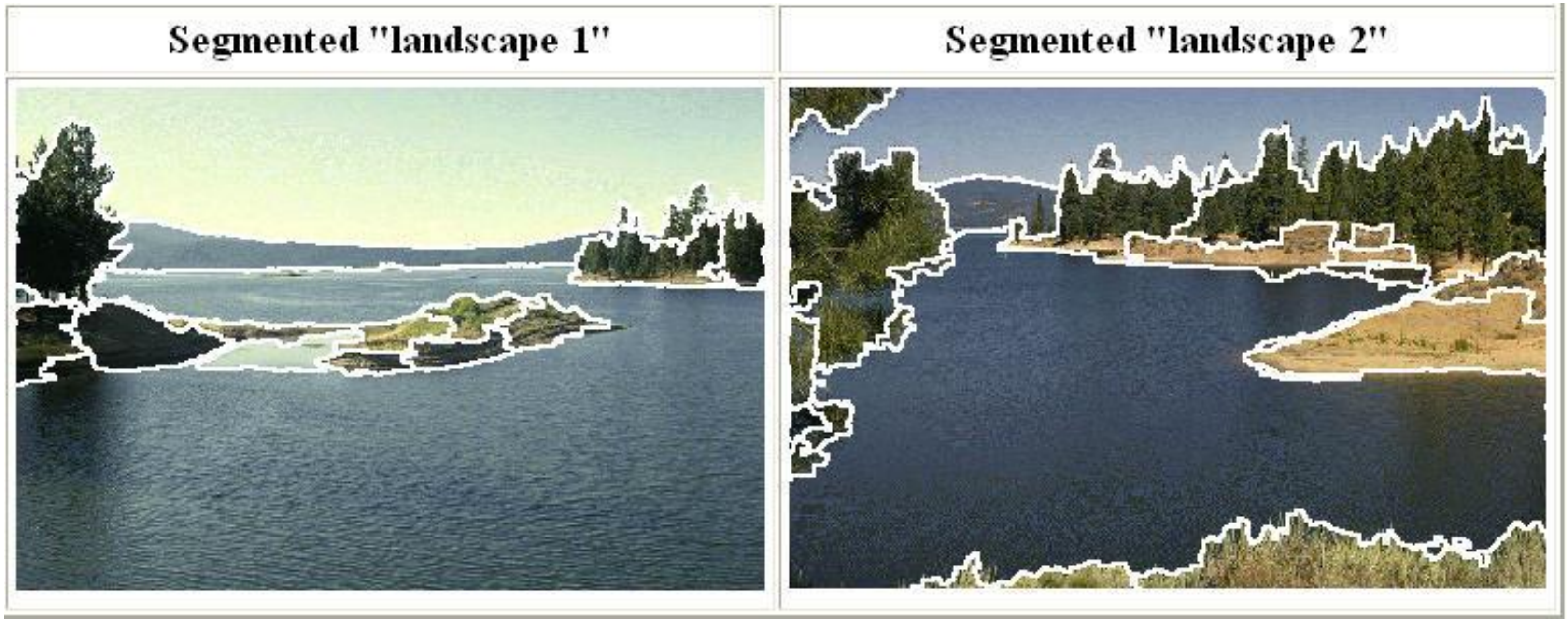
- Clustering
- Mean shift
- Graph-based

Combining object recognition & segmentation

- OBJCUT
- Other methods

Mean shift clustering and segmentation

- An advanced and versatile technique for clustering-based segmentation



http://www.caip.rutgers.edu/~comanici/MSPAMI/msPamiResults.html

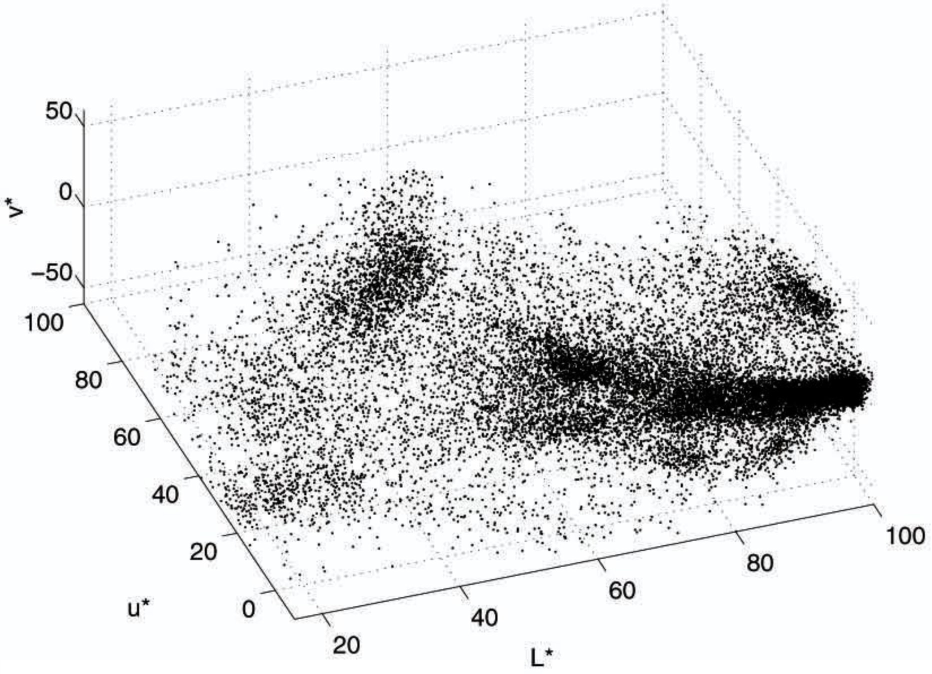
D. Comaniciu and P. Meer, Mean Shift: A Robust Approach toward Feature Space Analysis, PAMI 2002.

Mean shift algorithm

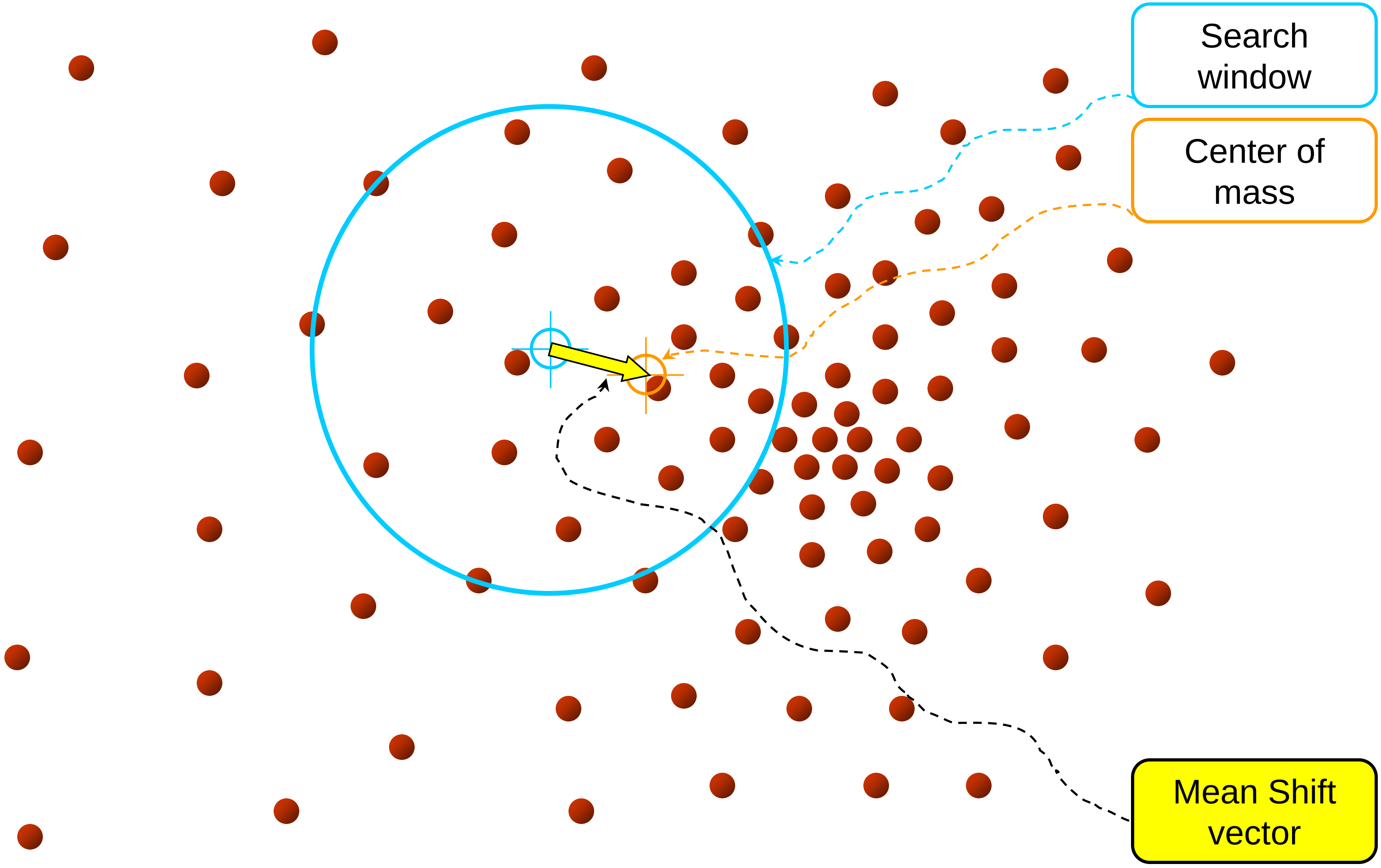
- The mean shift algorithm seeks *modes* or local maxima of density in the feature space

image

Feature space (L*u*v* color values)

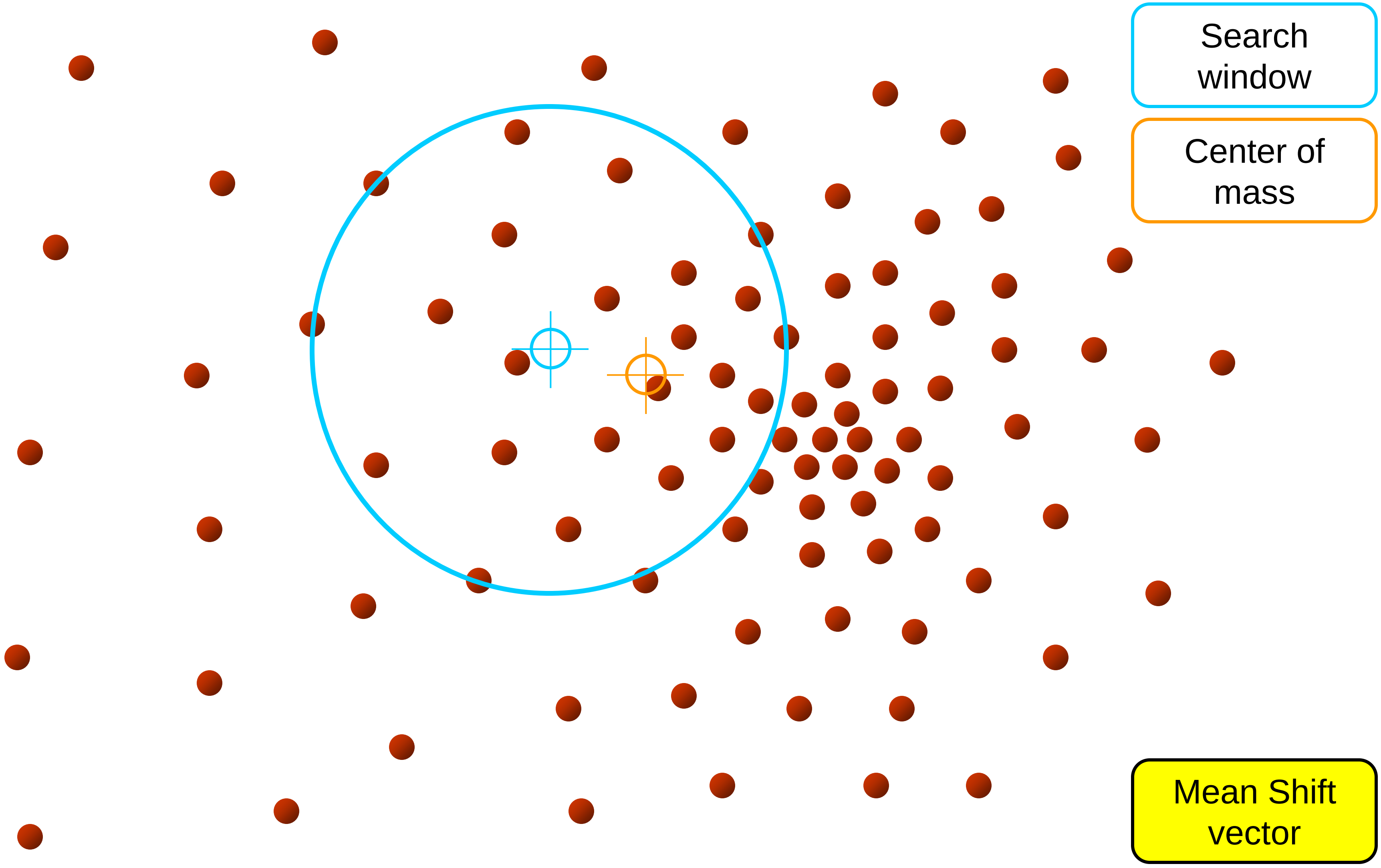


Mean shift



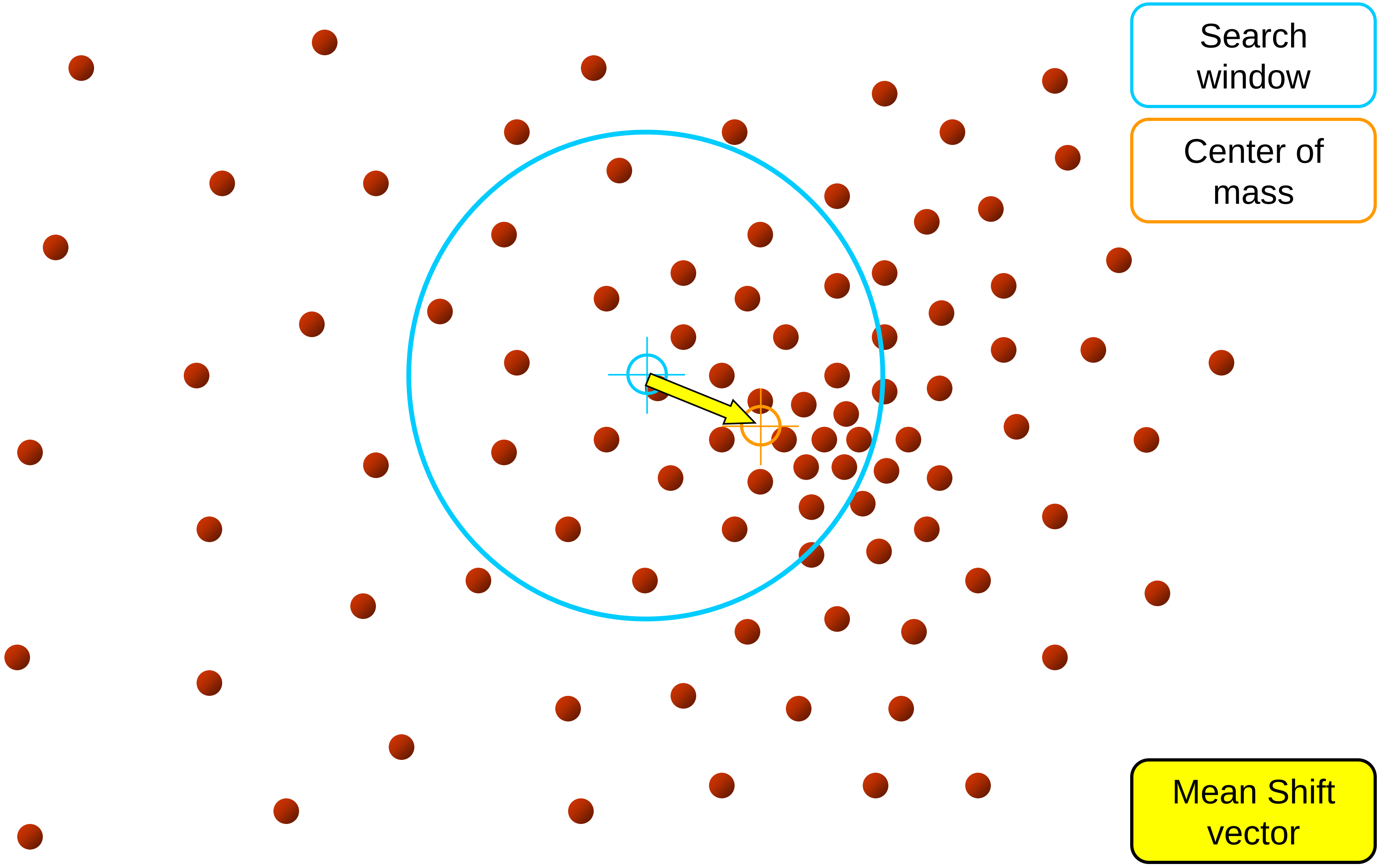
Slide by Y. Ukrainitz & B. Sarel

Mean shift



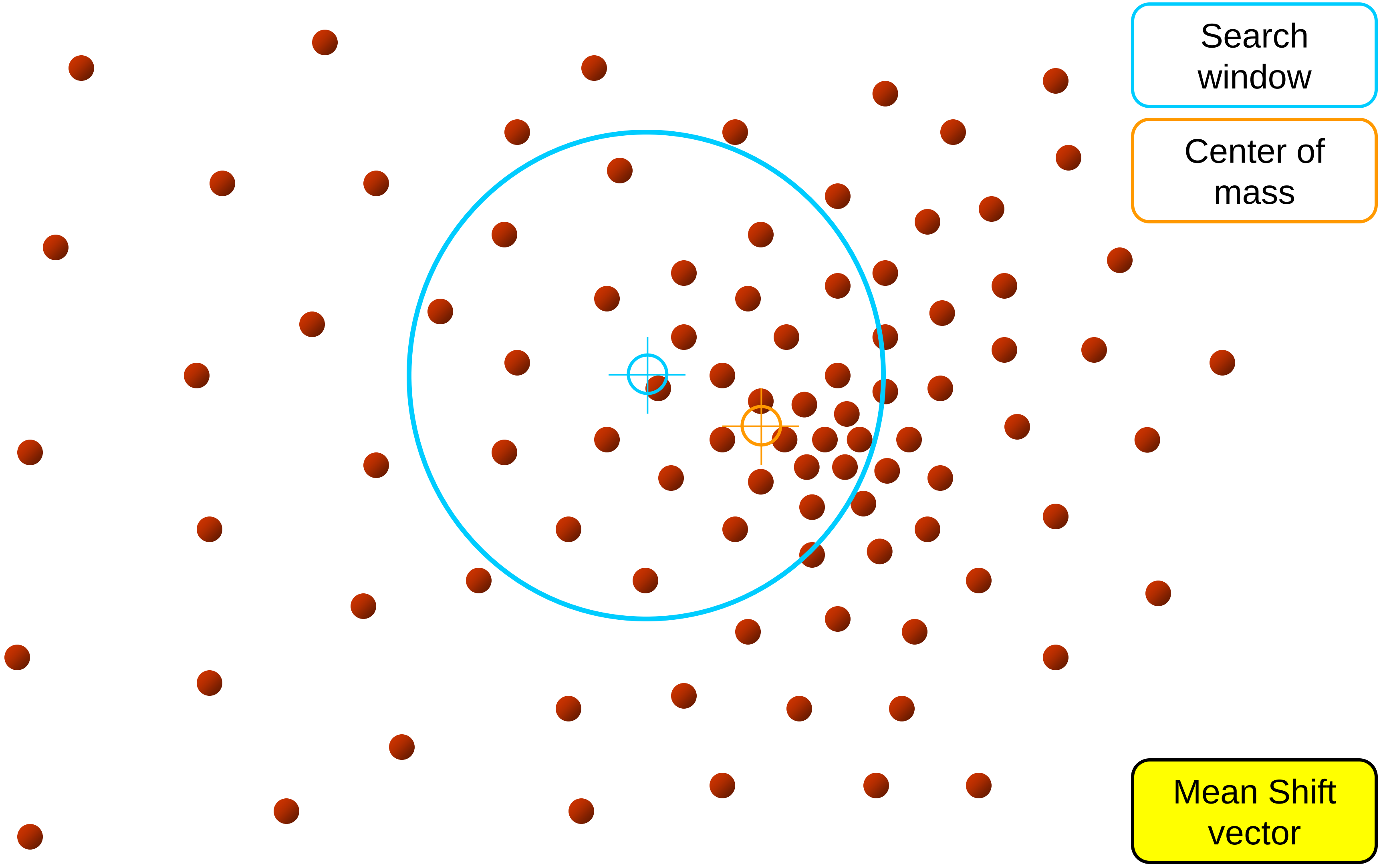
Slide by Y. Ukrainitz & B. Sarel

Mean shift



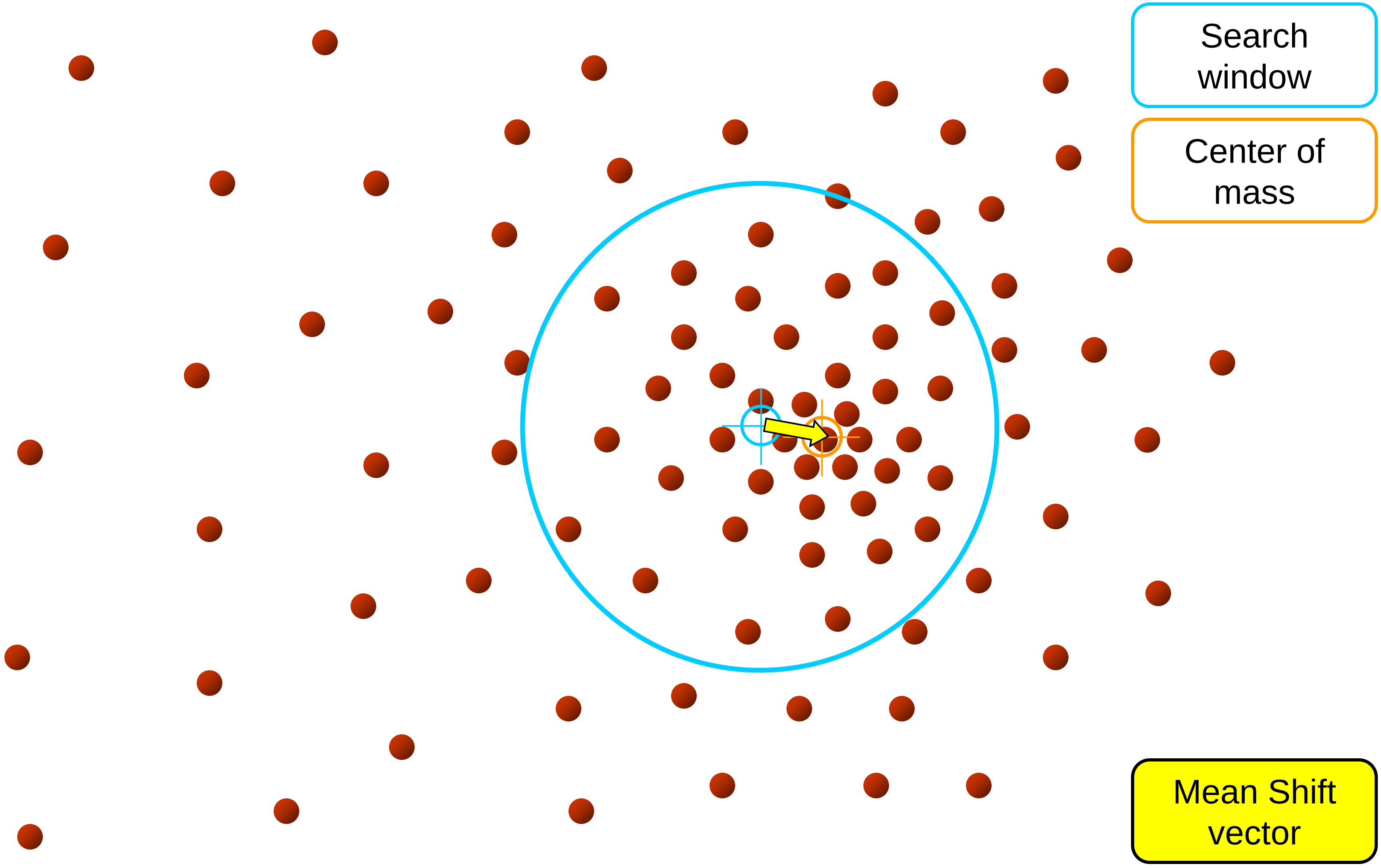
Slide by Y. Ukrainitz & B. Sarel

Mean shift



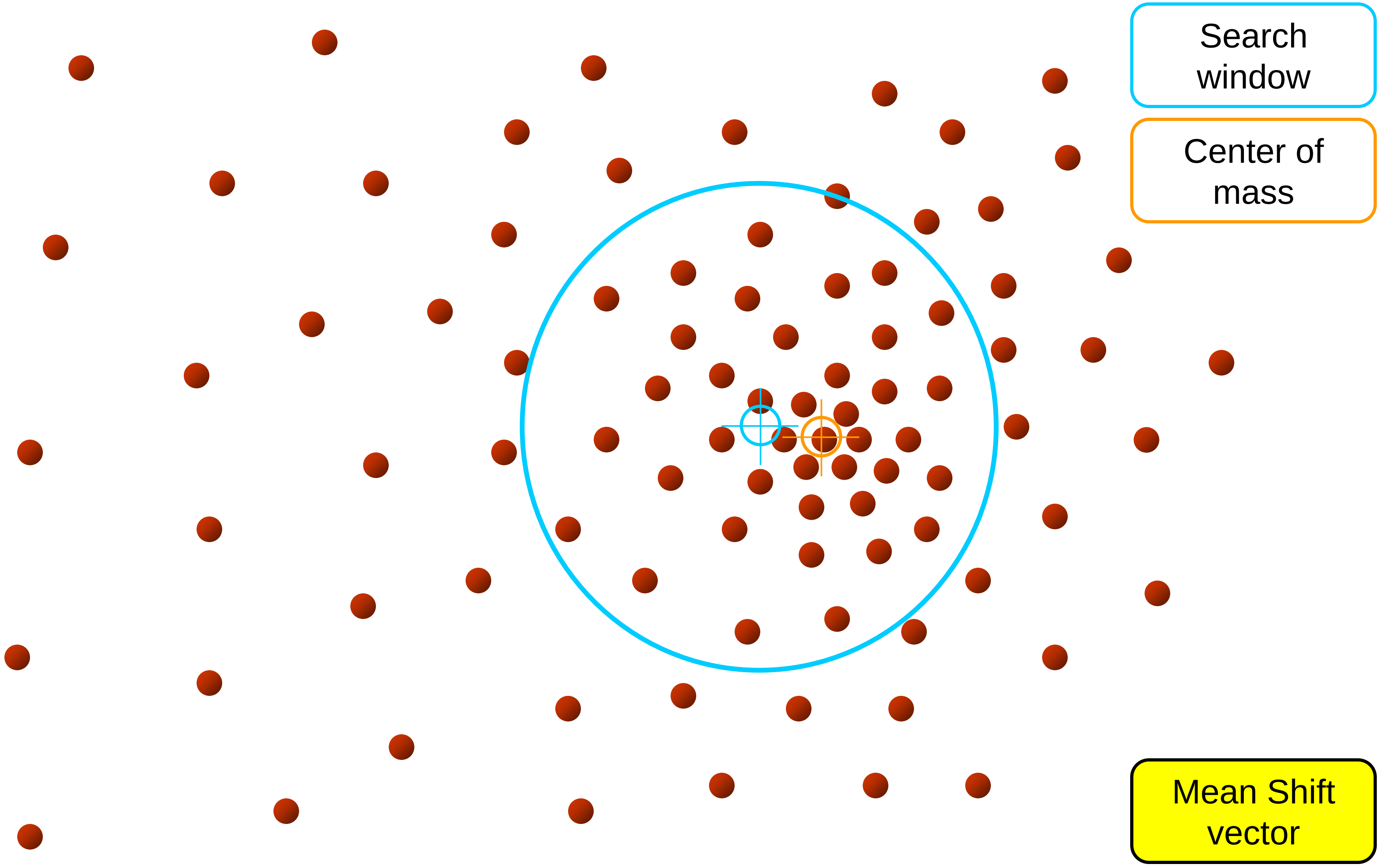
Slide by Y. Ukrainitz & B. Sarel

Mean shift



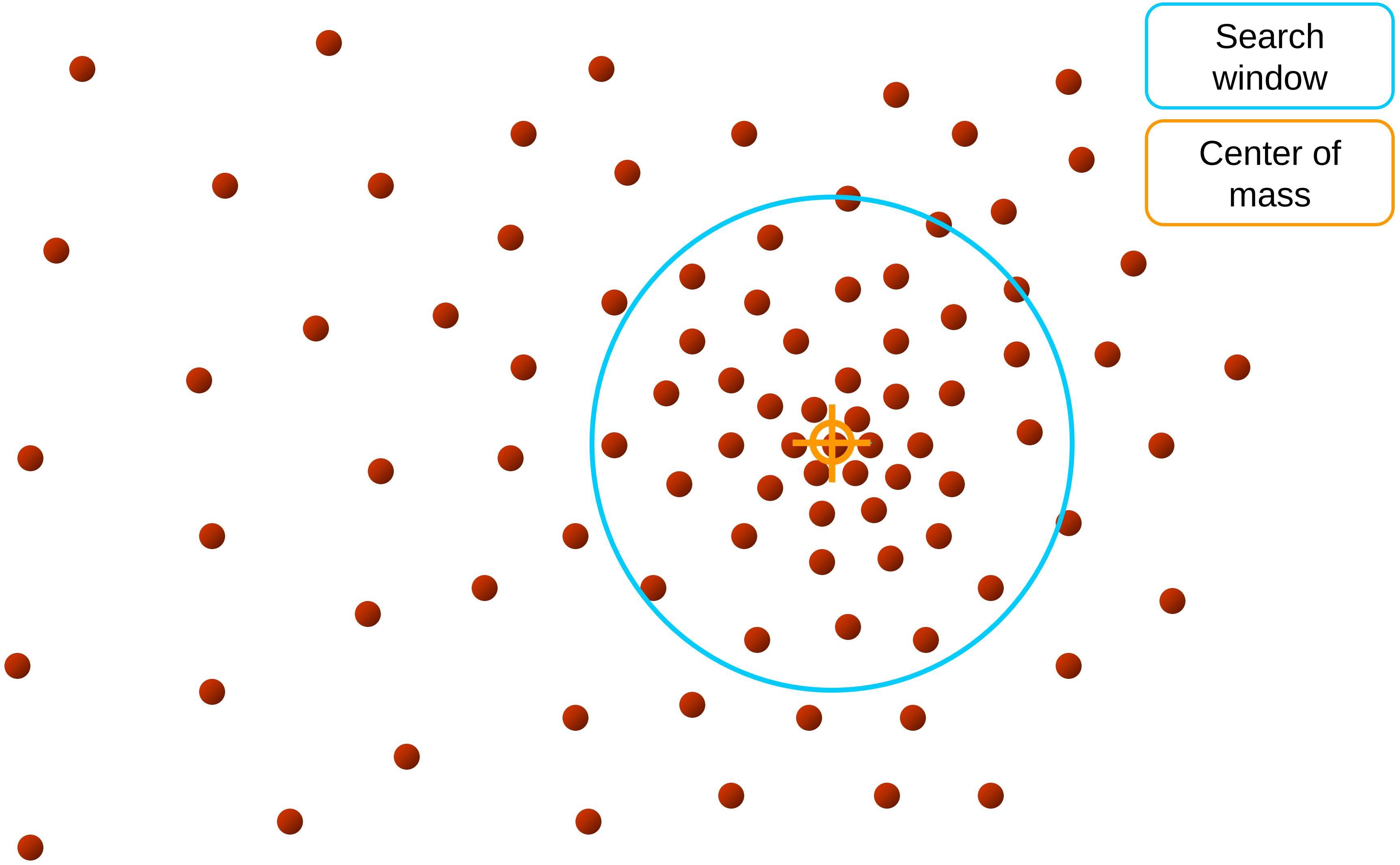
Slide by Y. Ukrainitz & B. Sarel

Mean shift



Slide by Y. Ukrainitz & B. Sarel

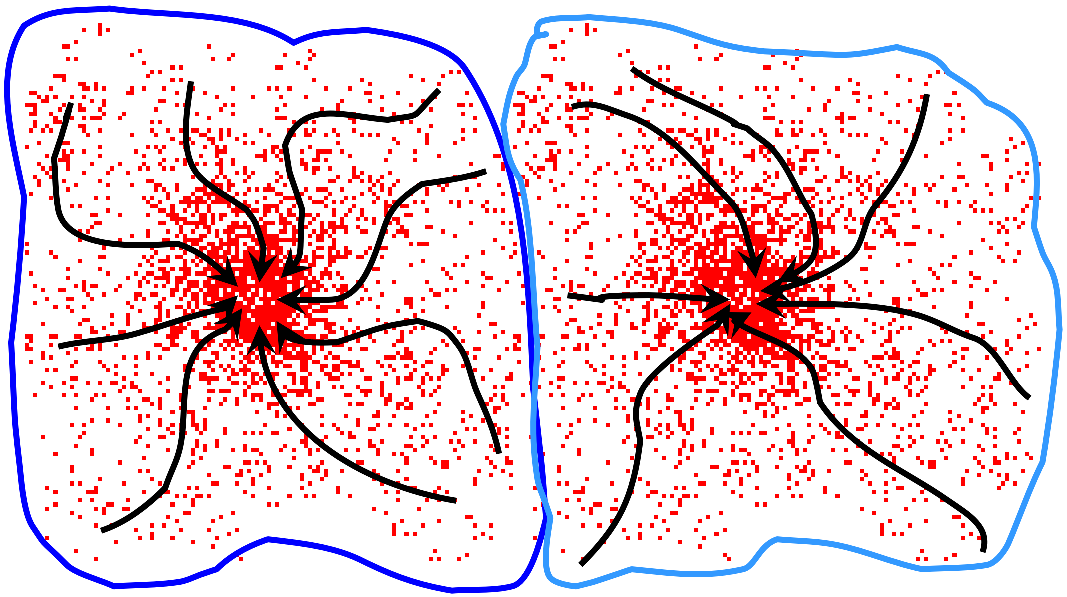
Mean shift



Slide by Y. Ukrainitz & B. Sarel

Mean shift clustering

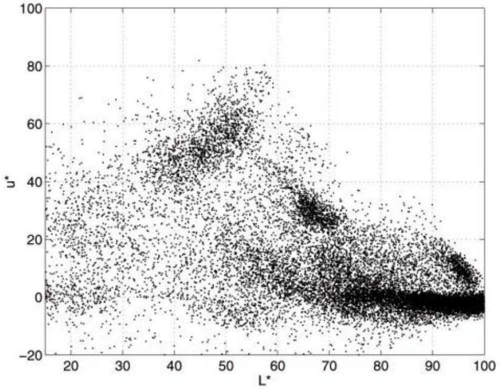
- Cluster: all data points in the attraction basin of a mode
- Attraction basin: the region for which all trajectories lead to the same mode



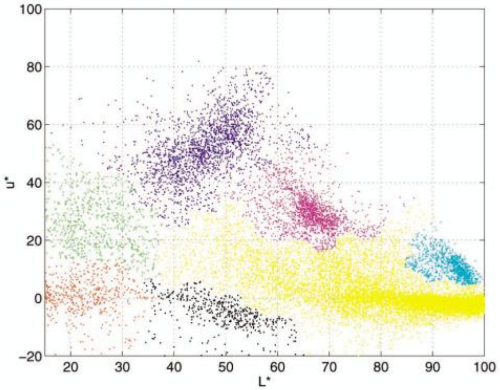
Slide by Y. Ukrainitz & B. Sarel

Mean shift clustering/segmentation

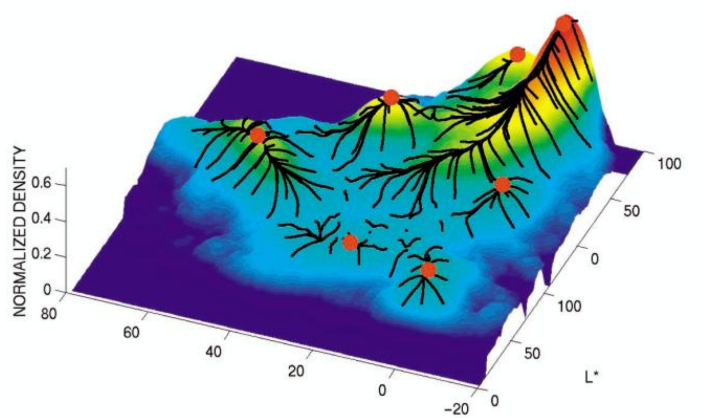
- Find features (color, gradients, texture, etc)
- Initialize windows at individual feature points
- Perform mean shift for each window until convergence
- Merge windows that end up near the same “peak” or mode



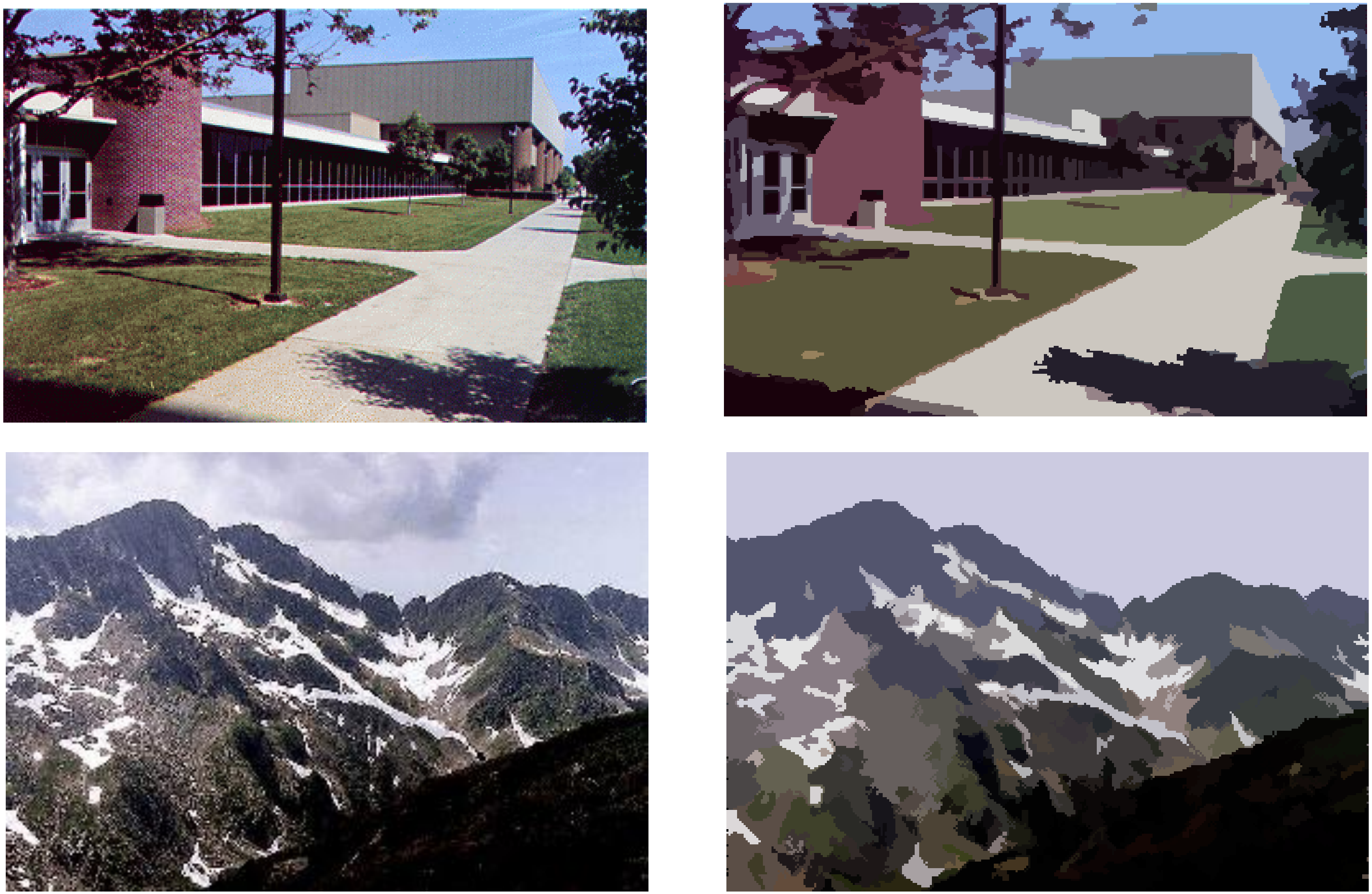
(a)



(b)

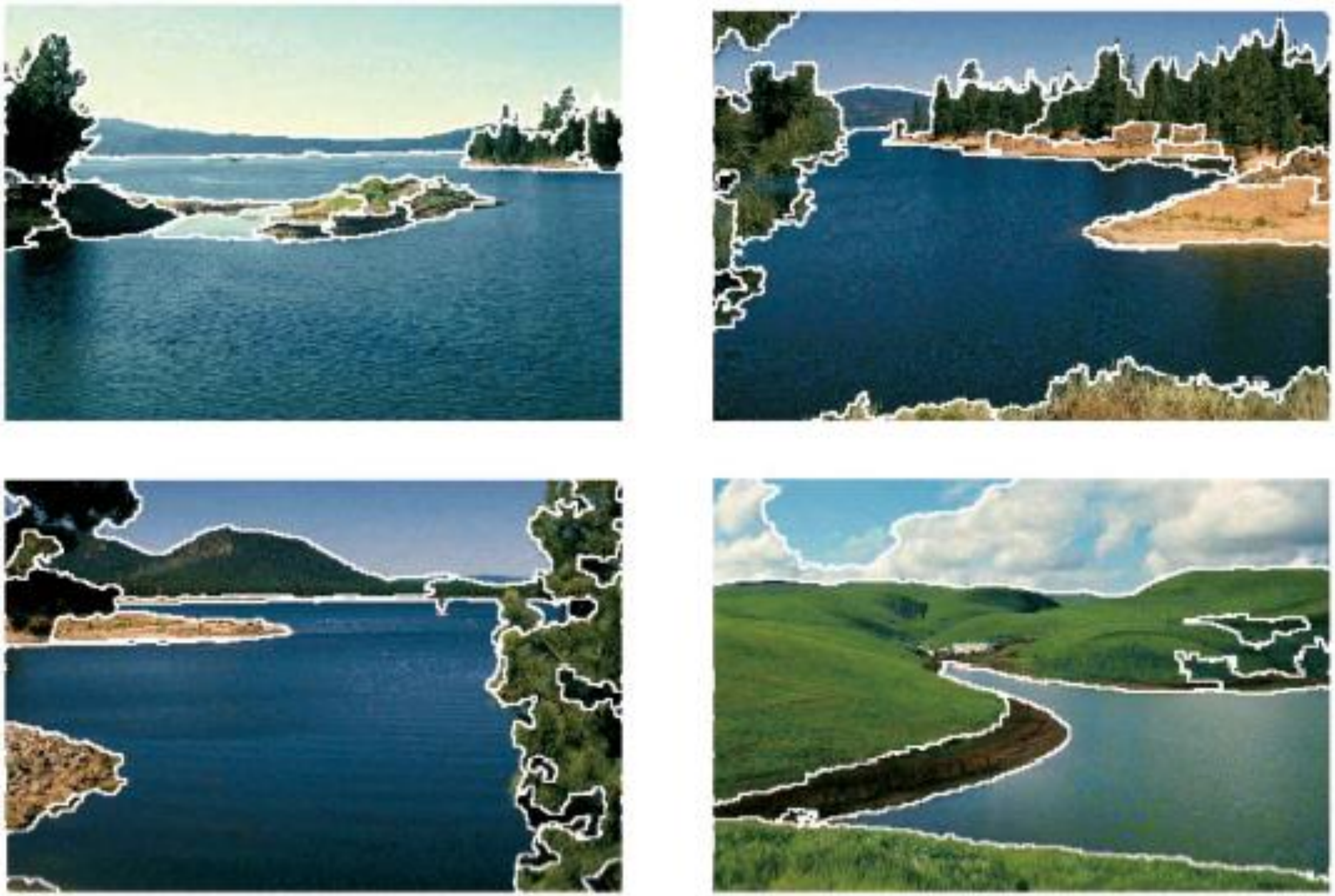


Mean shift segmentation results



http://www.caip.rutgers.edu/~comanici/MSPAMI/msPamiResults.html

More results



More results

Mean shift pros and cons

- Pros
  - Does not assume spherical clusters
  - Just a single parameter (window size)
  - Finds variable number of modes
  - Robust to outliers
- Cons
  - Output depends on window size
  - Computationally expensive
  - Does not scale well with dimension of feature space

Slide credit: S. Lazebnik

Overview

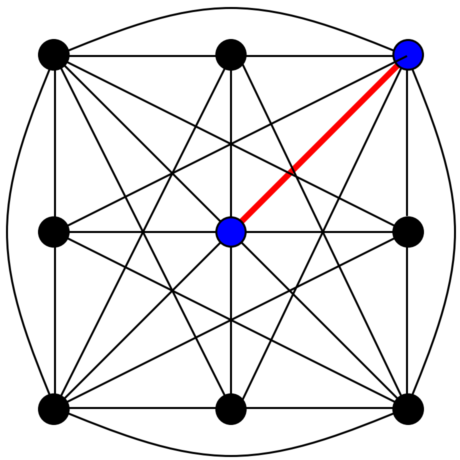
Bottom-up segmentation

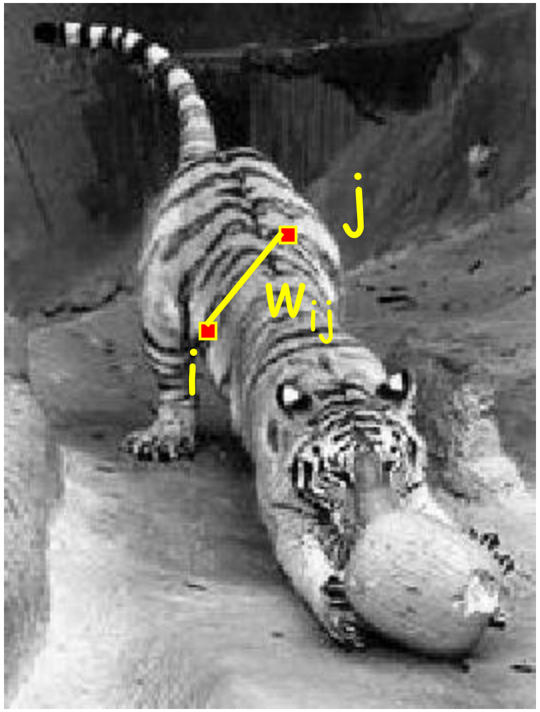
- Clustering
- Mean shift
- Graph-based

Combining object recognition & segmentation

- OBJCUT
- Other methods

Images as graphs

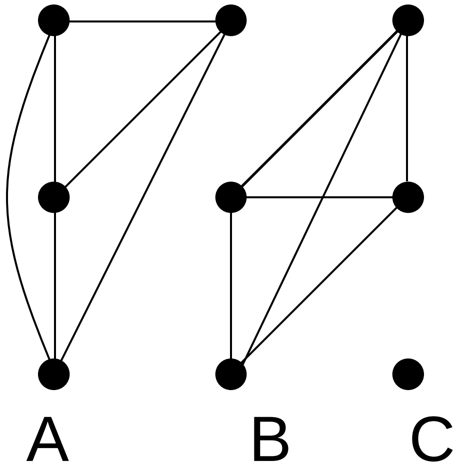




- Node for every pixel
- Edge between every pair of pixels (or every pair of “sufficiently close” pixels)
- Each edge is weighted by the *affinity* or similarity of the two nodes

Source: S. Seitz

Segmentation by graph partitioning



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

Source: S. Seitz

Measuring affinity

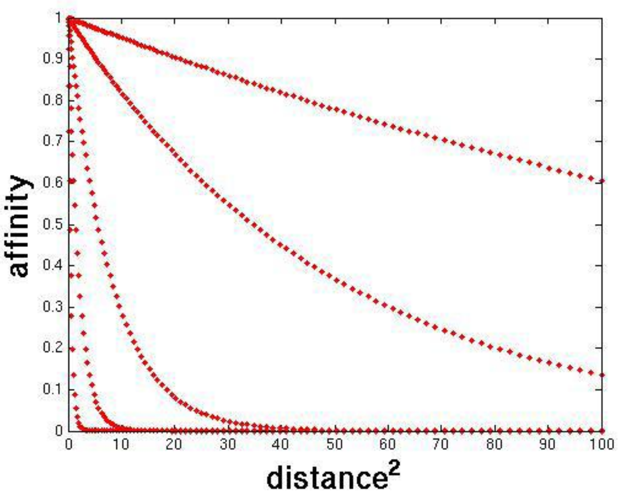
- Suppose we represent each pixel by a feature vector **x**, and define a distance function appropriate for this feature representation
- Then we can convert the distance between two feature vectors into an affinity with the help of a generalized Gaussian kernel:

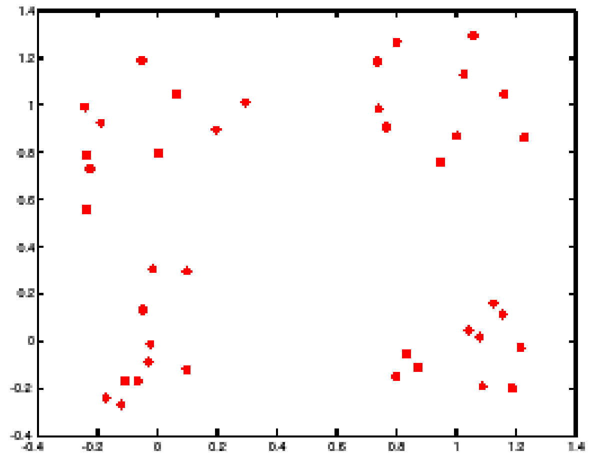
$$\exp\left(-\frac{1}{2\sigma^2}\mathrm{dist}(\mathbf{x}_i, \mathbf{x}_j)^2\right)$$

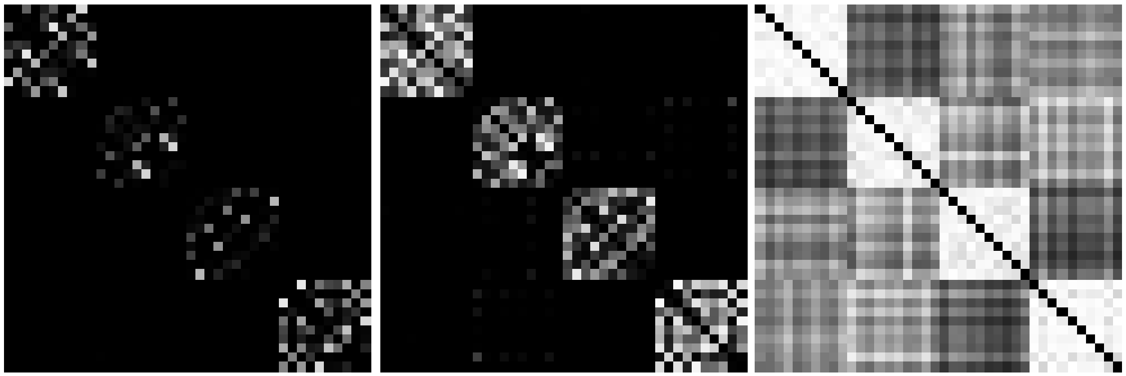
Slide credit: S. Lazebnik

Scale affects affinity

- Small σ: group only nearby points
- Large σ: group far-away points

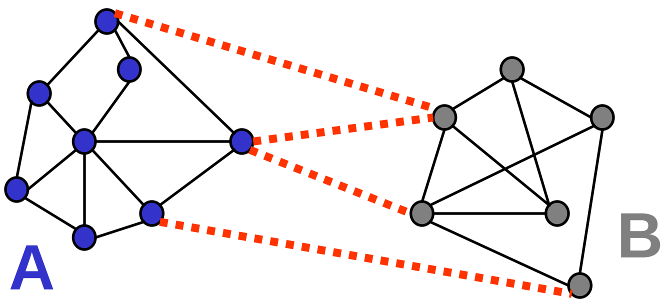






Slide credit: S. Lazebnik

Graph cut



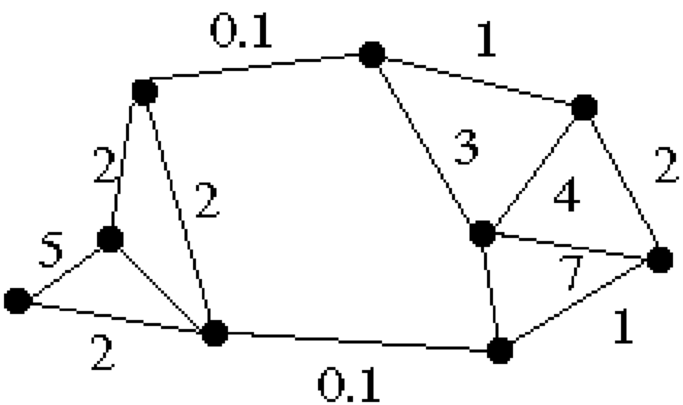
- Set of edges whose removal makes a graph disconnected
- Cost of a cut: sum of weights of cut edges
- A graph cut gives us a segmentation
  - What is a “good” graph cut and how do we find one?

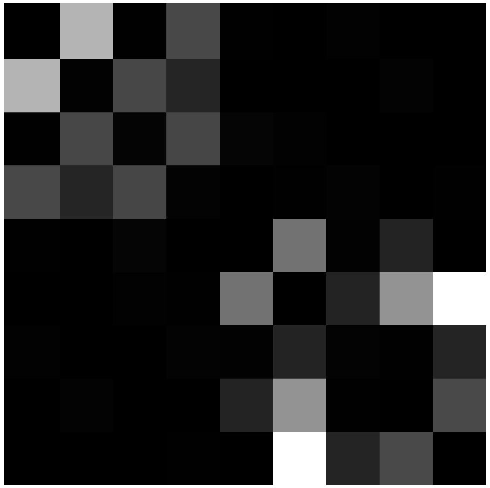
Source: S. Seitz

Minimum cut

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

Minimum cut example



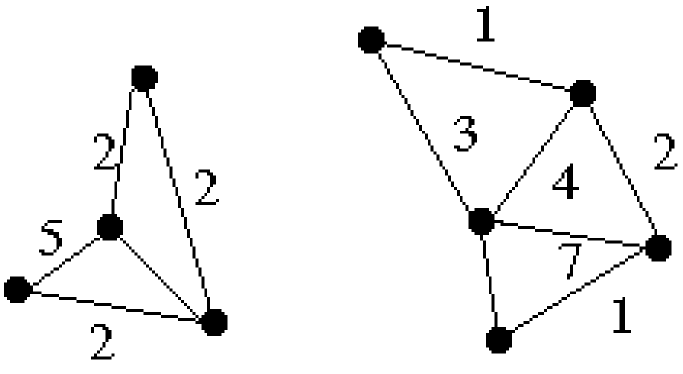


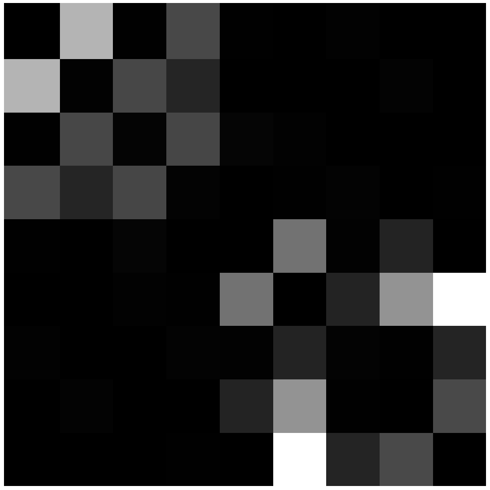
Slide credit: S. Lazebnik

Minimum cut

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

Minimum cut example

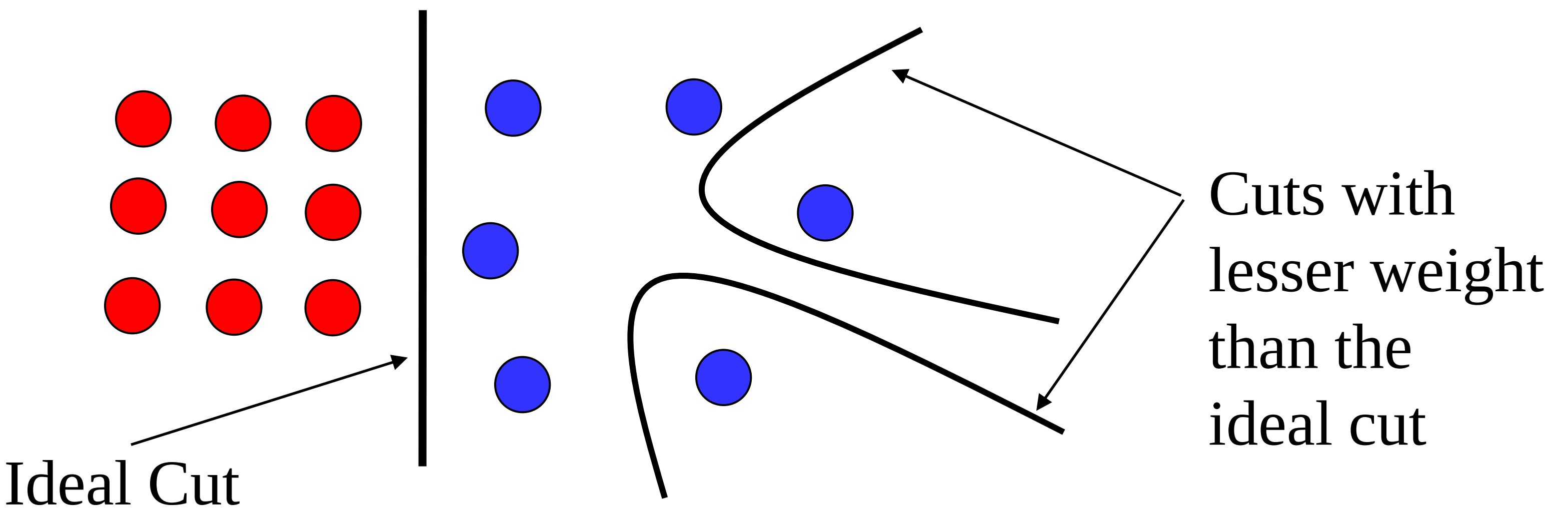




Slide credit: S. Lazebnik

Normalized cut

- Drawback: minimum cut tends to cut off very small, isolated components



* Slide from Khurram Hassan-Shafique CAP5415 Computer Vision 2003

Normalized cut

- Drawback: minimum cut tends to cut off very small, isolated components
- This can be fixed by normalizing the cut by the weight of all the edges incident to the segment
- The *normalized cut* cost is:

$$\frac{w(A,B)}{assoc(A,V)} + \frac{w(A,B)}{assoc(B,V)}$$

w(*A*, *B*) = sum of weights of all edges between *A* and *B*
assoc(A,V) = sum of all weights in cluster A + w(A,B)

J. Shi and J. Malik. [Normalized cuts and image segmentation.] PAMI 2000

Normalized cut

- Finding the exact minimum of the normalized cut cost is NP-complete, but we *relax* to let nodes take on arbitrary values:
- Let *W* be the adjacency matrix of the graph
- Let *D* be the diagonal matrix with diagonal entries $D(i, i) = \Sigma_j W(i, j)$
- Then the normalized cut cost can be written as

$$\frac{y^T(D-W)y}{y^T Dy}$$

where *y* is an indicator vector whose value should be 1 in the *i*th position if the *i*th feature point belongs to A and a negative constant otherwise

J. Shi and J. Malik. Normalized cuts and image segmentation. PAMI 2000

Normalized cut

- We can minimize the relaxed cost by solving the *generalized eigenvalue problem* $(D - W)y = \lambda Dy$
- The solution y is given by the generalized eigenvector corresponding to the second smallest eigenvalue
- Intutitively, the ith entry of y can be viewed as a "soft" indication of the component membership of the ith feature
  - Can use 0 or median value of the entries as the splitting point (threshold), or find threshold that minimizes the Ncut cost

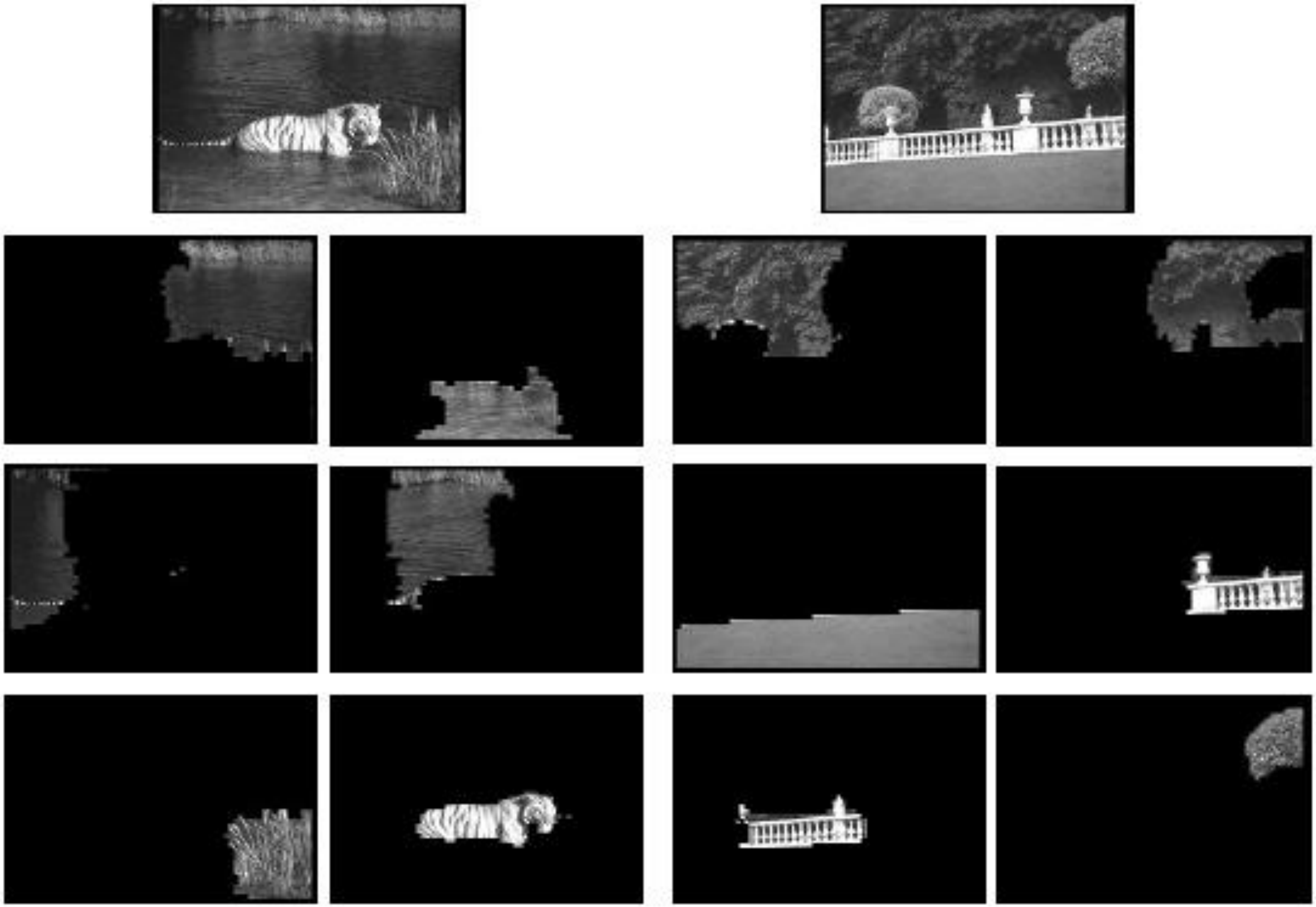
Normalized cut algorithm

1. Represent the image as a weighted graph $G = (V,E)$, compute the weight of each edge, and summarize the information in D and W
2. Solve $(D - W)y = \lambda Dy$ for the eigenvector with the second smallest eigenvalue
3. Use the entries of the eigenvector to bipartition the graph

To find more than two clusters:

- Recursively bipartition the graph
- Run k-means clustering on values of several eigenvectors

Example result

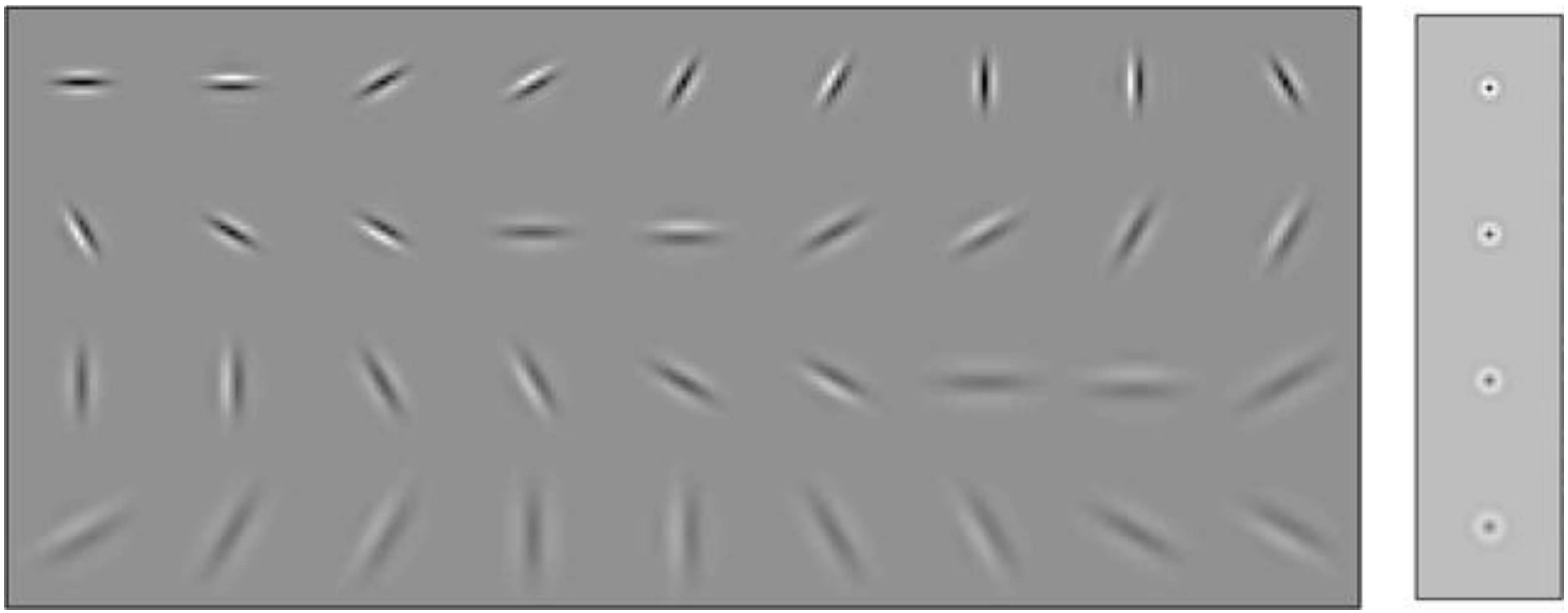


Challenge

- How to segment images that are a “mosaic of textures”?

Using texture features for segmentation

- Convolve image with a bank of filters



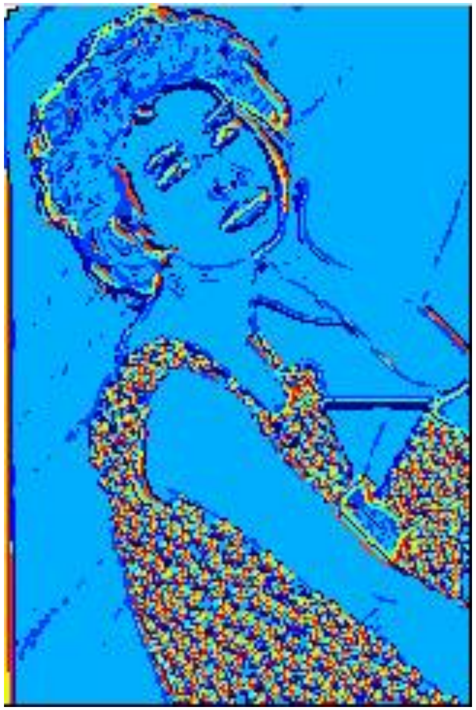
J. Malik, S. Belongie, T. Leung and J. Shi. ["Contour and Texture Analysis for Image Segmentation"](). IJCV 43(1),7-27,2001.

Using texture features for segmentation

- Convolve image with a bank of filters
- Find *textons* by clustering vectors of filter bank outputs

Image

Texton map

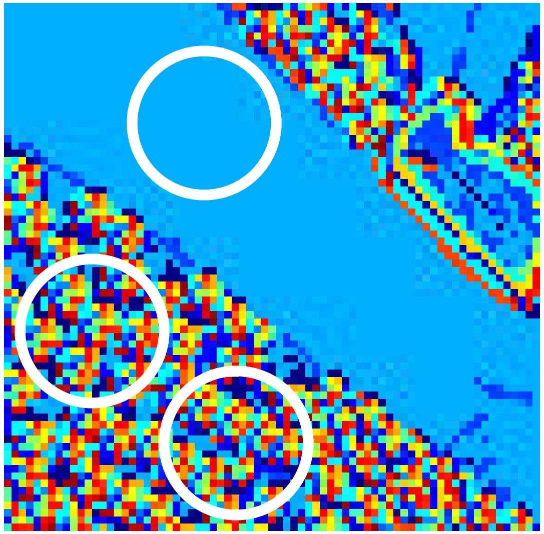


J. Malik, S. Belongie, T. Leung and J. Shi. ["*Contour and Texture Analysis for Image Segmentation*"]. IJCV 43(1),7-27,2001.

Slide credit: S. Lazebnik

Using texture features for segmentation

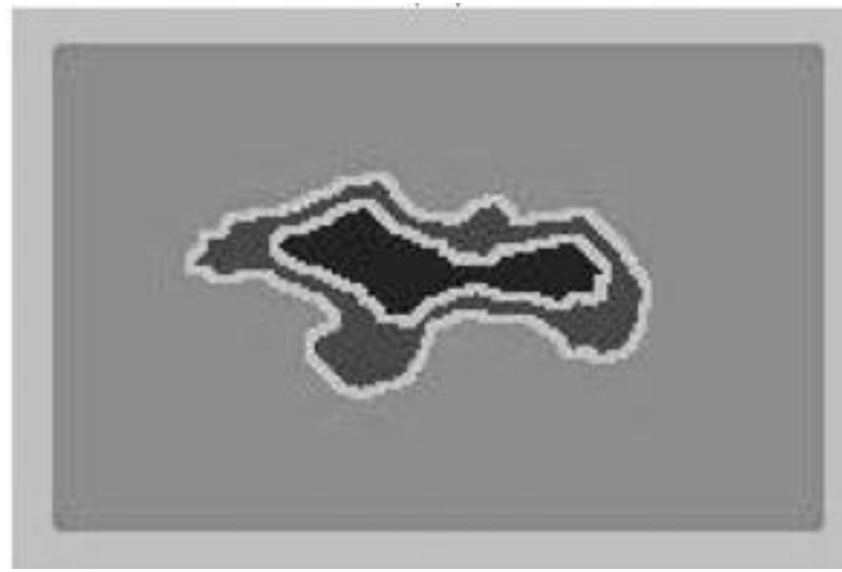
- Convolve image with a bank of filters
- Find *textons* by clustering vectors of filter bank outputs
- The final texture feature is a texton histogram computed over image windows at some "local scale"



J. Malik, S. Belongie, T. Leung and J. Shi. ["*Contour and Texture Analysis for Image Segmentation*"](). IJCV 43(1),7-27,2001.

Slide credit: S. Lazebnik

Pitfall of texture features



- Possible solution: check for “intervening contours” when computing connection weights

J. Malik, S. Belongie, T. Leung and J. Shi. *"Contour and Texture Analysis for Image Segmentation"*. IJCV 43(1),7-27,2001.

Example results

Results: Berkeley Segmentation Engine

http://www.cs.berkeley.edu/~fowlkes/BSE/

Normalized cuts: Pro and con

- Pros
  - Generic framework, can be used with many different features and affinity formulations
- Cons
  - High storage requirement and time complexity
  - Bias towards partitioning into equal segments

Slide credit: S. Lazebnik

Overview

Bottom-up segmentation

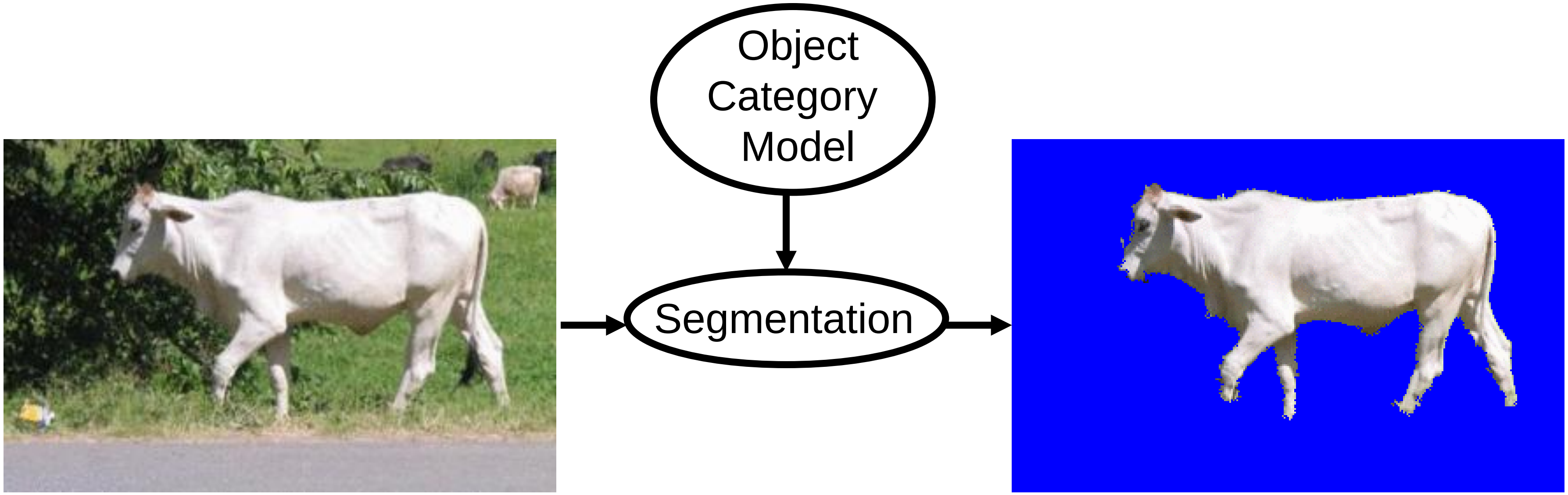
- Clustering
- Mean shift
- Graph-based
- Texton

Combining object recognition & segmentation

- OBJCUT
- Other methods

Aim

- Given an image and object category, to segment the object



Cow Image

Segmented Cow

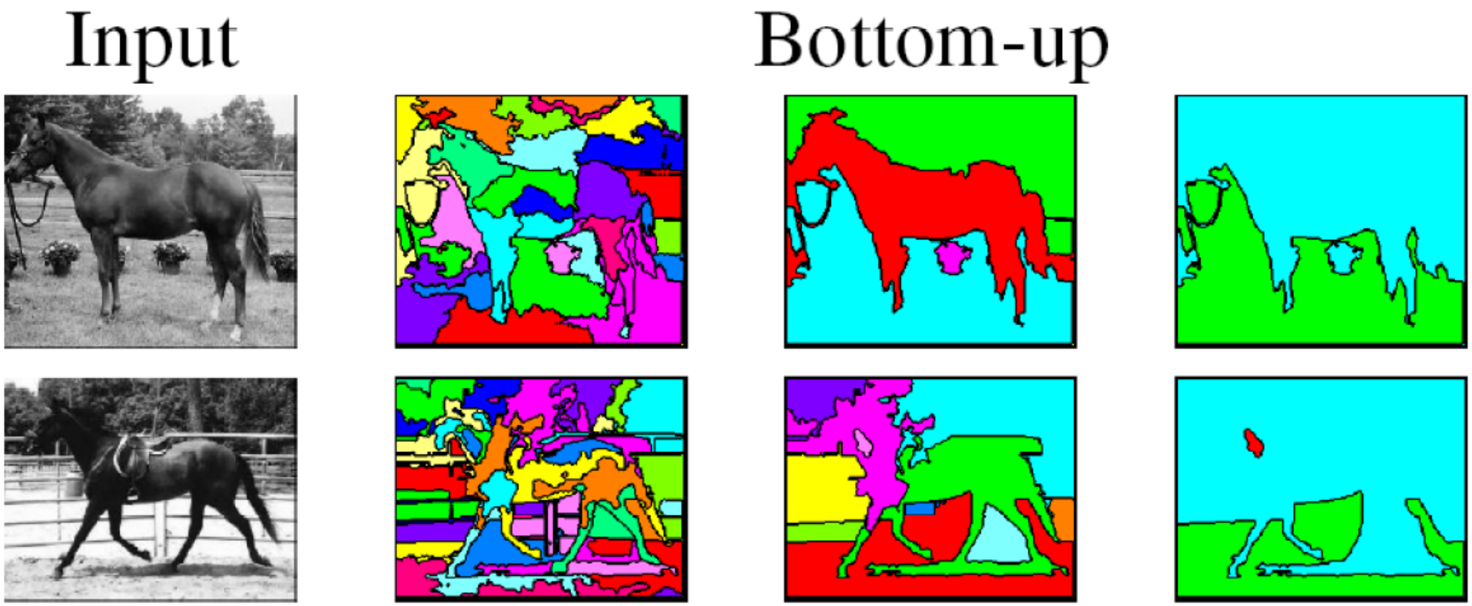
Segmentation should (ideally) be
- shaped like the object e.g. cow-like
- obtained efficiently in an unsupervised manner
- able to handle self-occlusion

Slide from Kumar ‘05

Feature-detector view

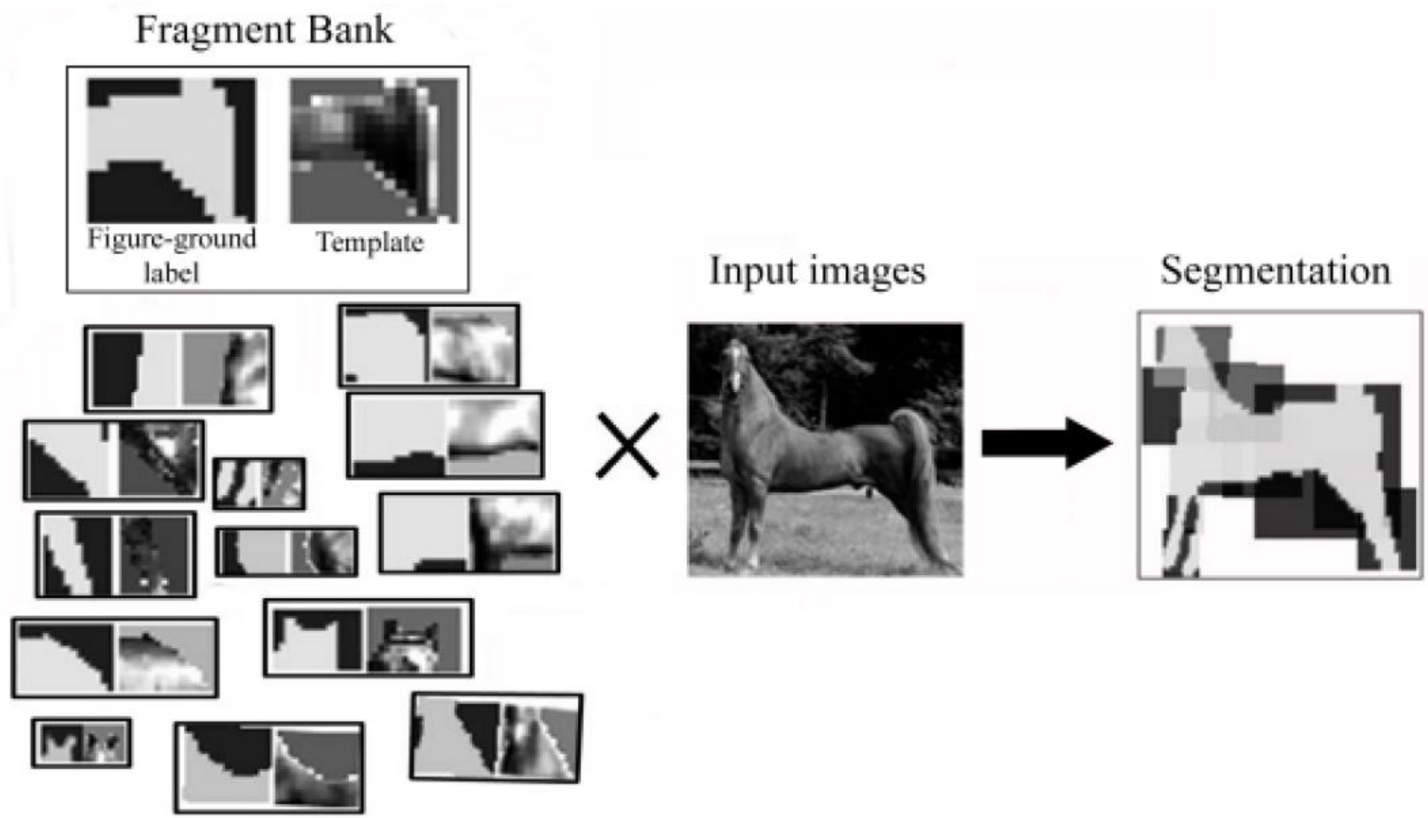
Examples of bottom-up segmentation

- Using Normalized Cuts, Shi & Malik, 1997

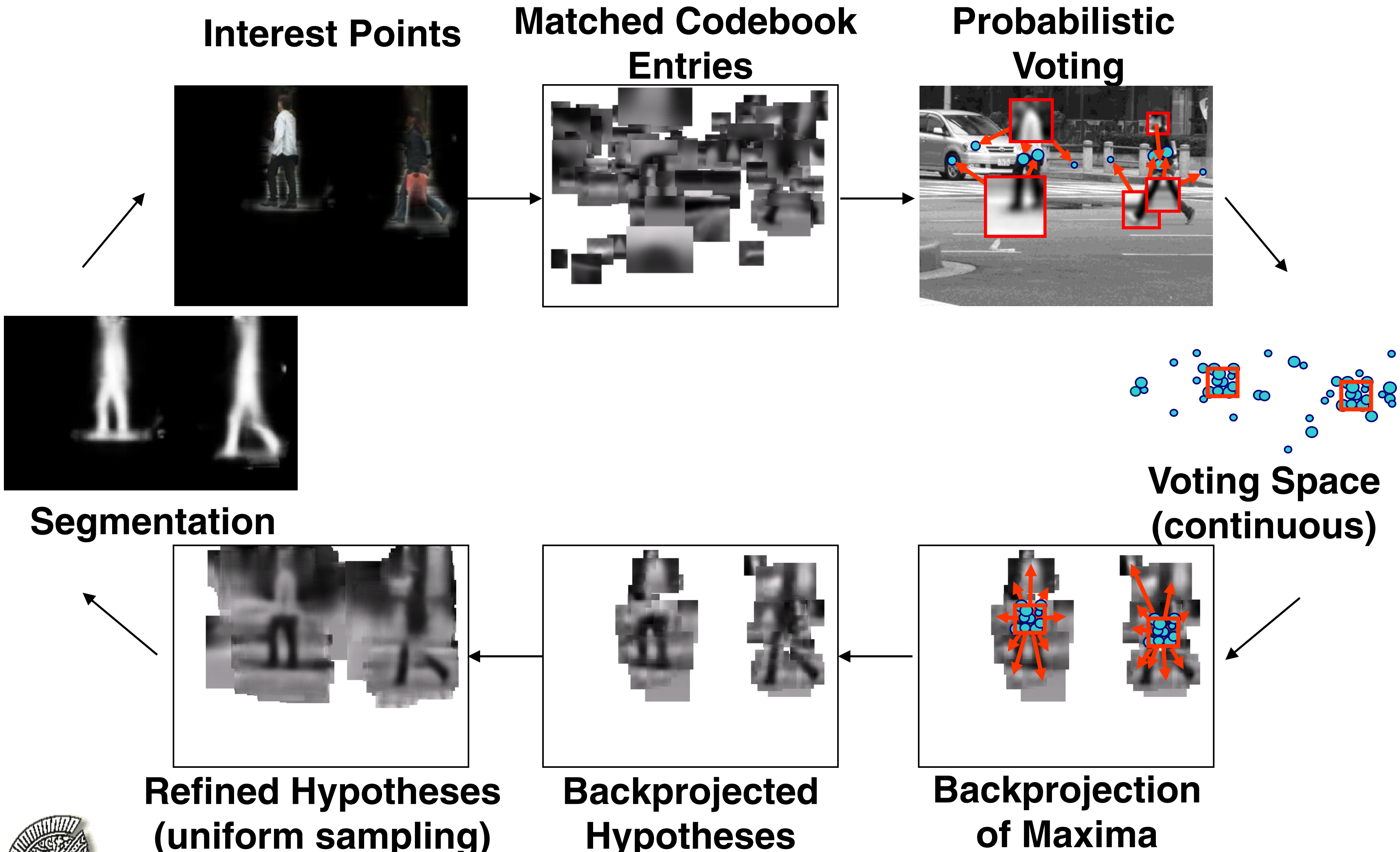


Borenstein and Ullman, ECCV 2002

Jigsaw approach: Borenstein and Ullman, 2002



Implicit Shape Model - Liebe and Schiele, 2003



Liebe and Schiele, 2003, 2005

Overview

- Bottom-up segmentation
  - Clustering
  - Mean shift
  - Graph-based
- Combining object recognition & segmentation
  - OBJCUT
  - Other methods

OXFORD BROOKES UNIVERSITY

UNIVERSITY OF OXFORD

OBJ CUT

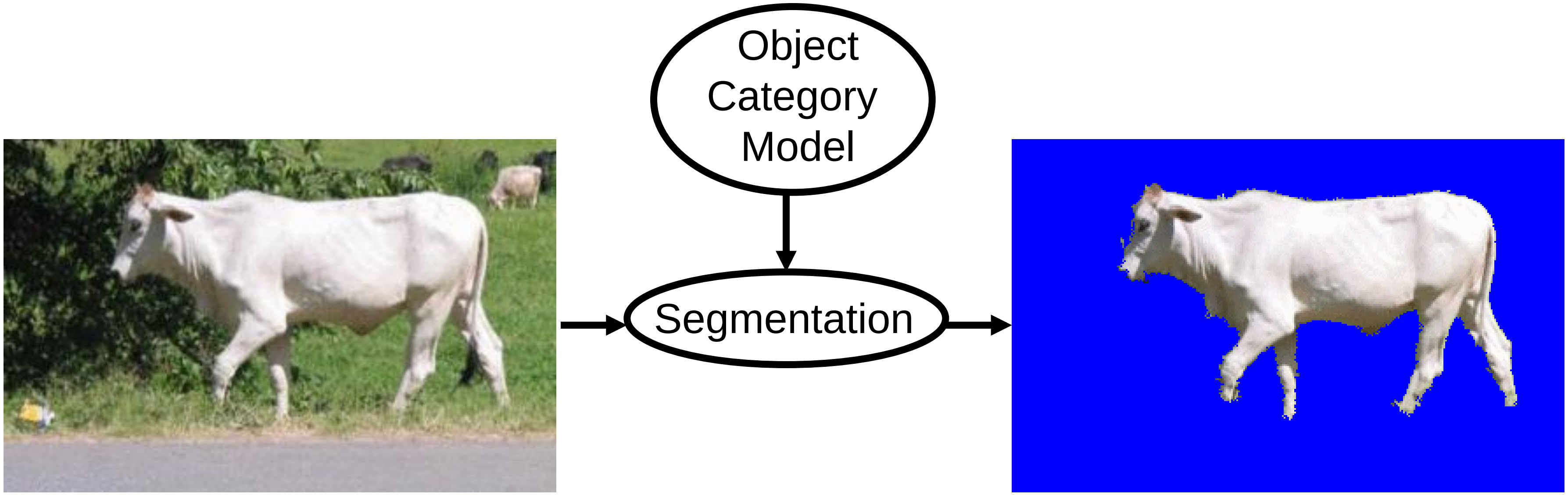
M. Pawan Kumar

Philip Torr

Andrew Zisserman

Aim

- Given an image, to segment the object



Cow Image

Segmented Cow

Segmentation should (ideally) be

- shaped like the object e.g. cow-like
- obtained efficiently in an unsupervised manner
- able to handle self-occlusion

Challenges

Intra-Class Shape Variability



Intra-Class Appearance Variability



Self Occlusion

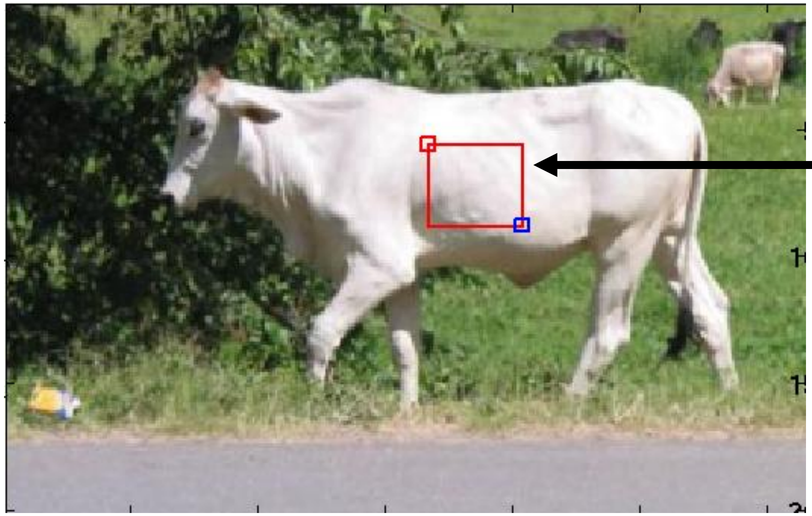


Motivation

Magic Wand

Current methods require user intervention

- Object and background seed pixels (Boykov and Jolly, ICCV 01)
- Bounding Box of object (Rother *et al.* SIGGRAPH 04)



Object Seed Pixels

Cow Image

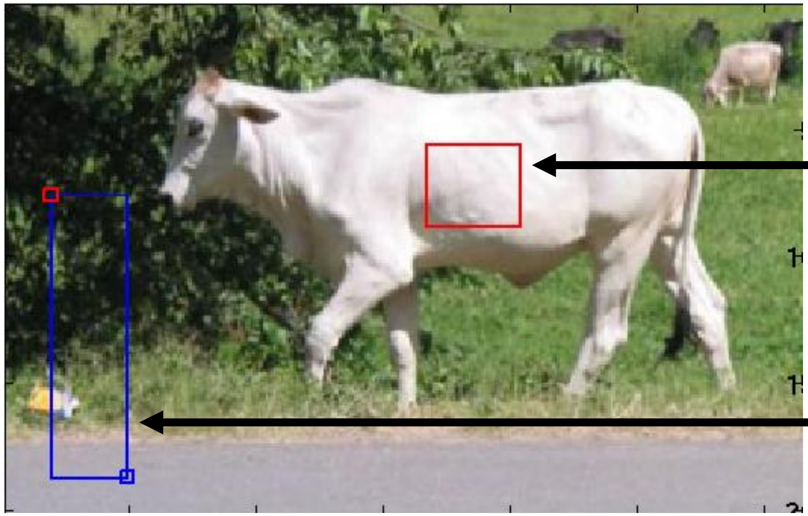
Slide credit: P. Kumar

Motivation

Magic Wand

Current methods require user intervention

- Object and background seed pixels (Boykov and Jolly, ICCV 01)
- Bounding Box of object (Rother *et al.* SIGGRAPH 04)



Object Seed Pixels

Background Seed Pixels

Cow Image

Slide credit: P. Kumar

Motivation

Magic Wand

Current methods require user intervention

- Object and background seed pixels (Boykov and Jolly, ICCV 01)
- Bounding Box of object (Rother *et al.* SIGGRAPH 04)

Segmented Image

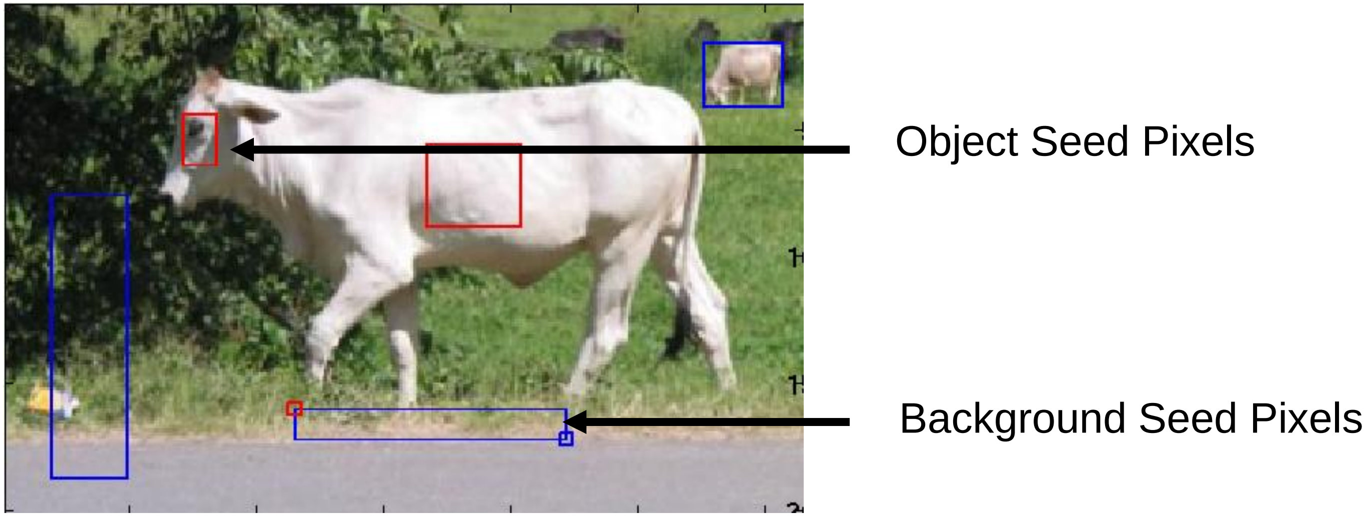
Slide credit: P. Kumar

Motivation

Magic Wand

Current methods require user intervention

- Object and background seed pixels (Boykov and Jolly, ICCV 01)
- Bounding Box of object (Rother *et al.* SIGGRAPH 04)



Cow Image

Slide credit: P. Kumar

Motivation

Magic Wand

Current methods require user intervention

- Object and background seed pixels (Boykov and Jolly, ICCV 01)
- Bounding Box of object (Rother *et al.* SIGGRAPH 04)

Segmented Image

Slide credit: P. Kumar

Motivation

Problem

- Manually intensive
- Segmentation is not guaranteed to be ‘object-like’

Non Object-like Segmentation

Slide credit: P. Kumar

Our Method

- Combine object detection with segmentation
  - Borenstein and Ullman, ECCV ’02
  - Leibe and Schiele, BMVC ’03
- Incorporate global shape priors in MRF
- Detection provides
  - Object Localization
  - Global shape priors
- Automatically segments the object
  - Note our method is completely generic
  - Applicable to any object category model

Slide credit: P. Kumar

Outline

- Problem Formulation
- Form of Shape Prior
- Optimization
- Results

Slide credit: P. Kumar

Problem

- Labelling **m** over the set of pixels **D**

- Shape prior provided by parameter θ

- Energy $E(m, \theta) = \sum \underbrace{\phi_x(\mathbf{D}|m_x) + \phi_x(m_x|\theta)}_{\textbf{Unary terms}} + \sum \underbrace{\phi_{xy}(m_x, m_y) + \phi(\mathbf{D}|m_x, m_y)}_{\textbf{Pairwise terms}}$

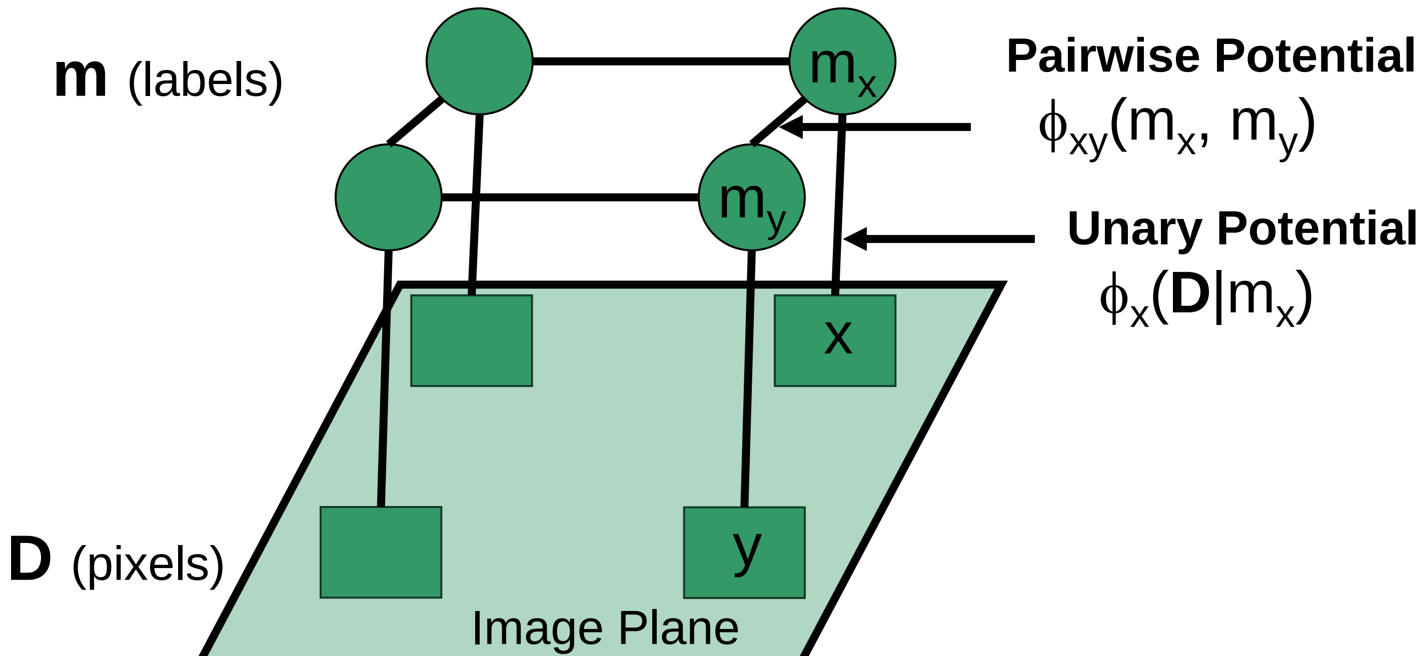
- Unary terms
  - Likelihood based on colour
  - Unary potential based on distance from θ
- Pairwise terms
  - Prior
  - Contrast term

- Find best labelling $\mathbf{m}^* = \arg\min \sum w_i E(\mathbf{m}, \theta_i)$
  - w_i is the weight for sample θ_i

Slide credit: P. Kumar

Markov Random Field (MRF)

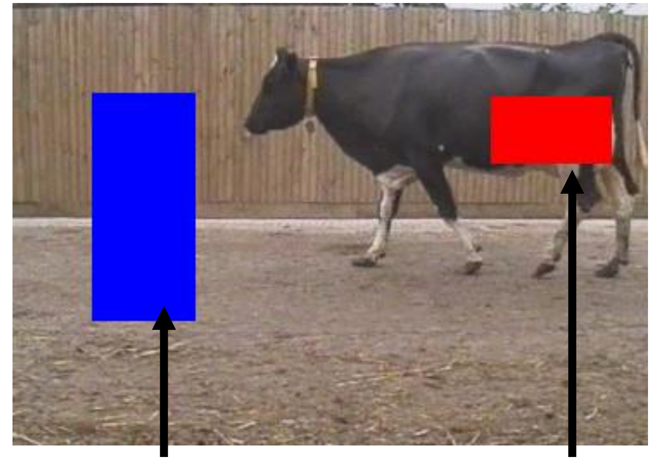
Probability for a labelling consists of
- Likelihood
    - Unary potential based on colour of pixel
- Prior which favours same labels for neighbours (pairwise potentials)



Slide credit: P. Kumar

Example

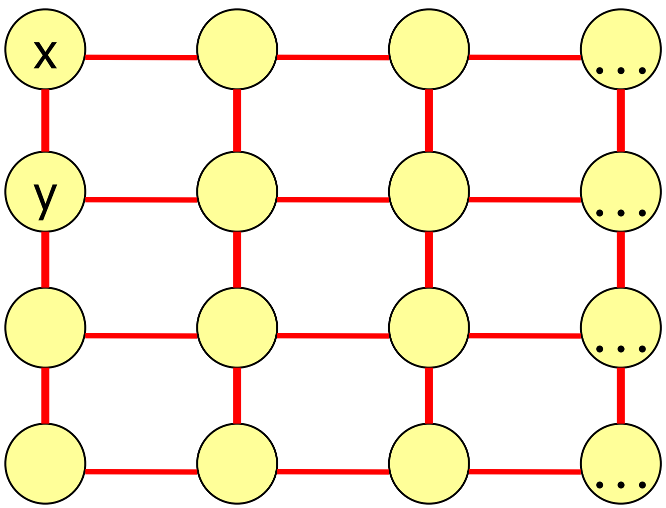
Cow Image



Background Seed Pixels

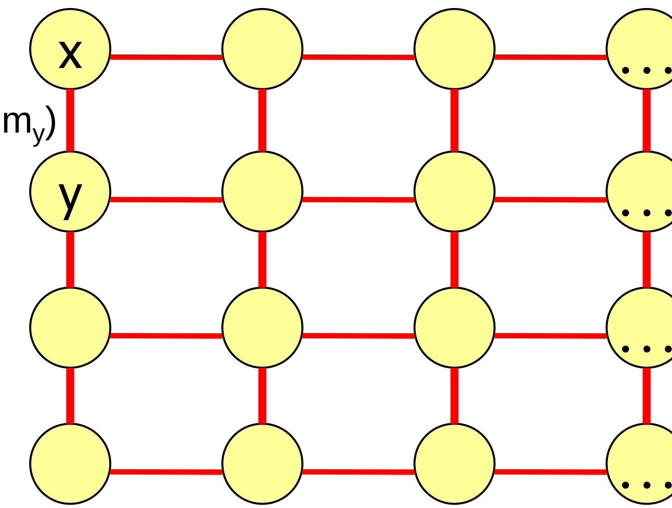
Object Seed Pixels

$$\frac{\phi_x(\mathbf{D}|\text{obj})}{\phi_x(\mathbf{D}|\text{bkg})}$$



Likelihood Ratio (Colour)

$\phi_{xy}(m_x,m_y)$

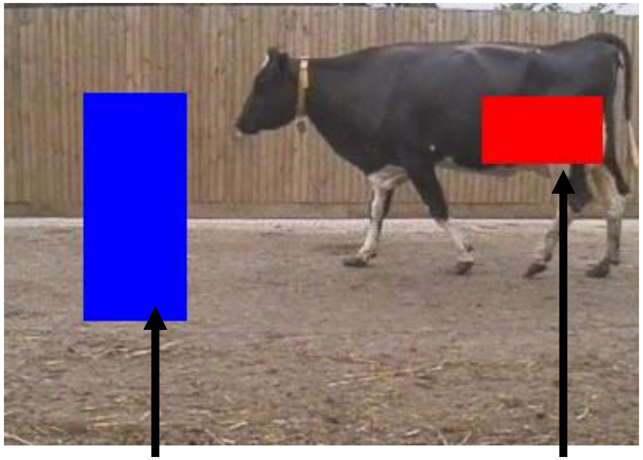


Prior

Slide credit: P. Kumar

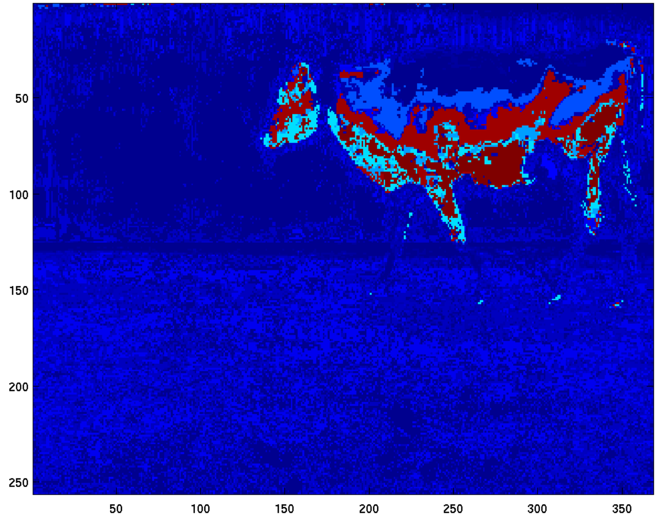
Example

Cow Image

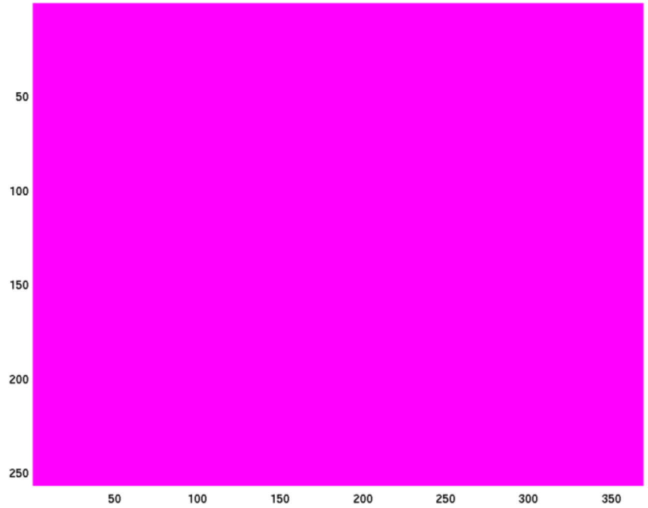


Background Seed Pixels

Object Seed Pixels



Likelihood Ratio (Colour)



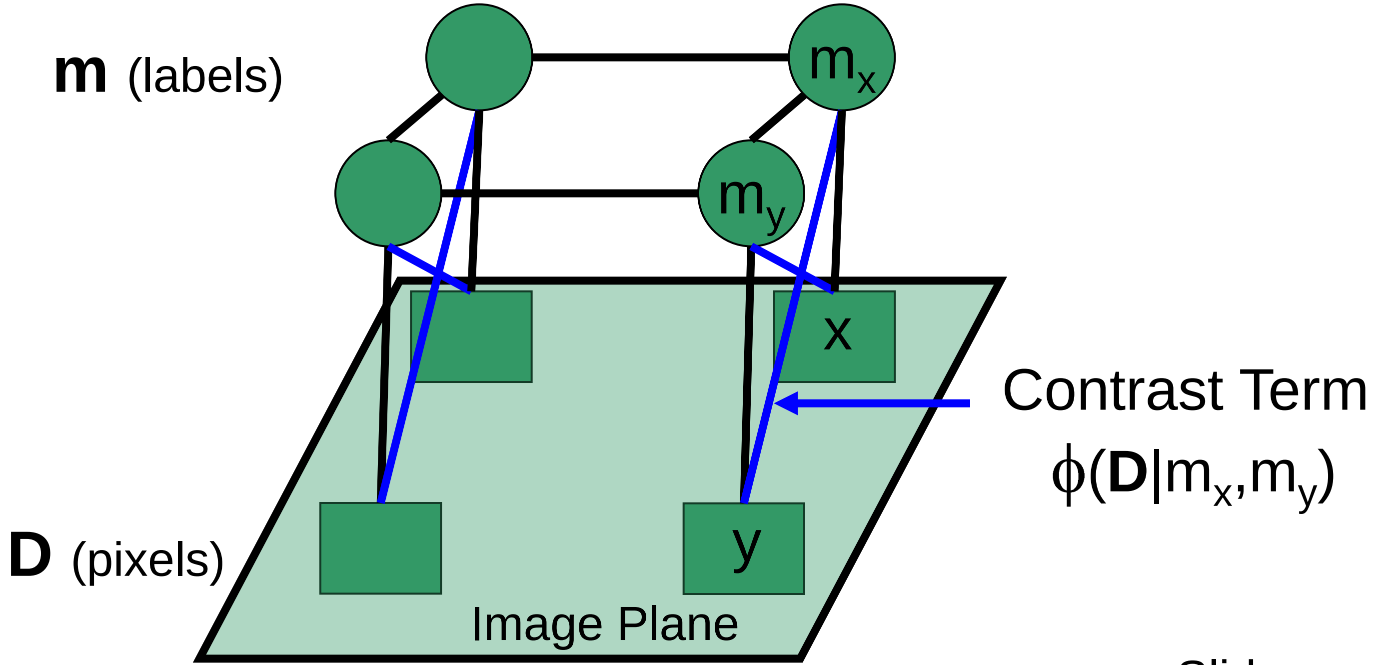
Prior

Slide credit: P. Kumar

Contrast-Dependent MRF

Probability of labelling in addition has

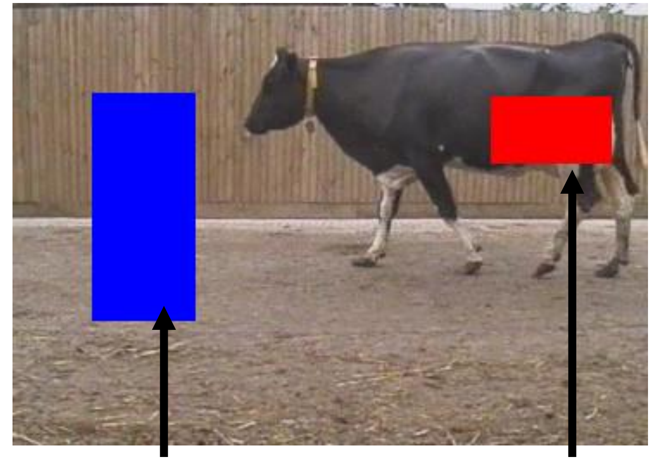
- Contrast term which favours boundaries to lie on image edges



Slide credit: P. Kumar

Example

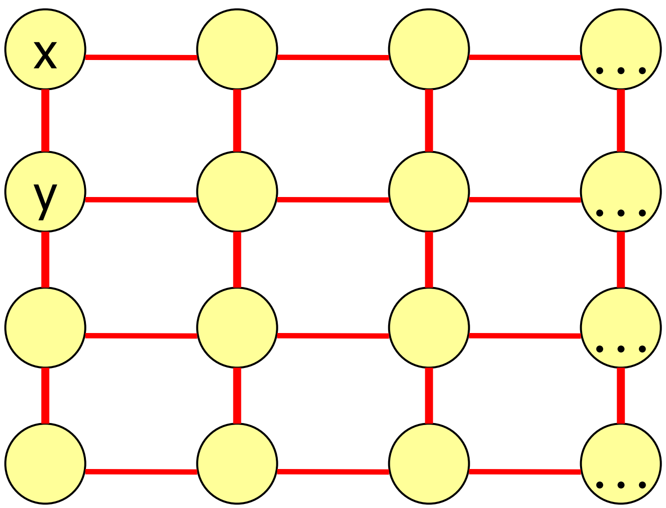
Cow Image



Background Seed Pixels

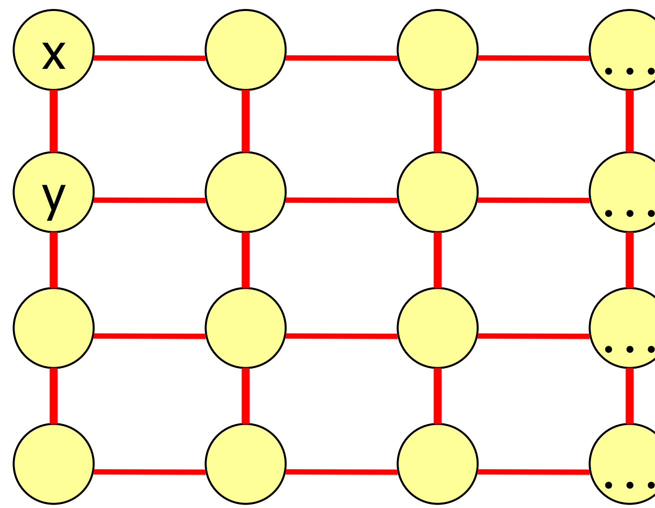
Object Seed Pixels

$$\frac{\phi_x(\mathbf{D}|\text{obj})}{\phi_x(\mathbf{D}|\text{bkg})}$$



Likelihood Ratio (Colour)

$\phi_{xy}(m_x,m_y)+$
$\phi_{xy}(\mathbf{D}|m_x,m_y)$

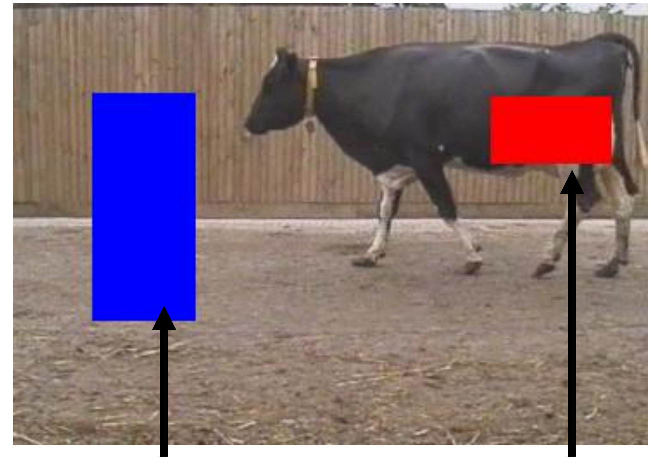


Prior + Contrast

Slide credit: P. Kumar

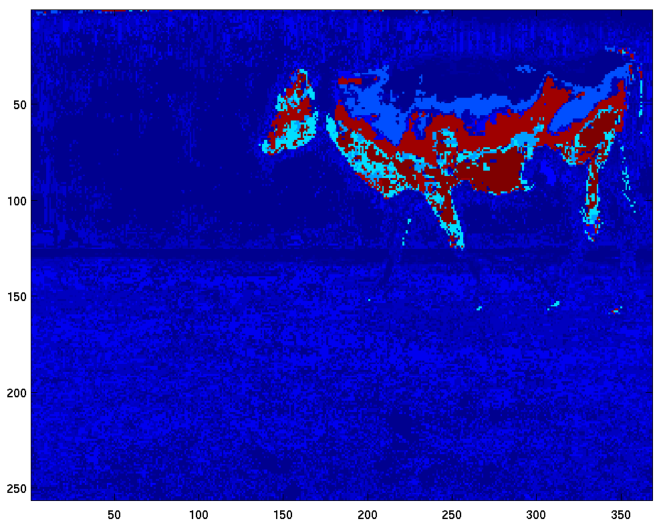
Example

Cow Image

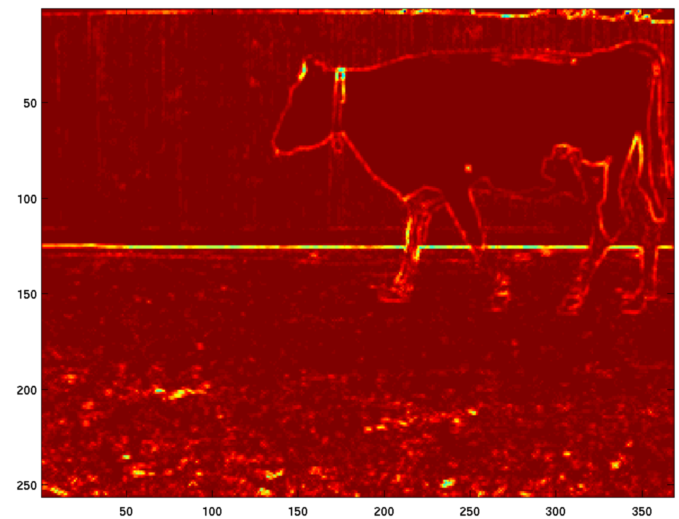


Background Seed Pixels

Object Seed Pixels



Likelihood Ratio (Colour)



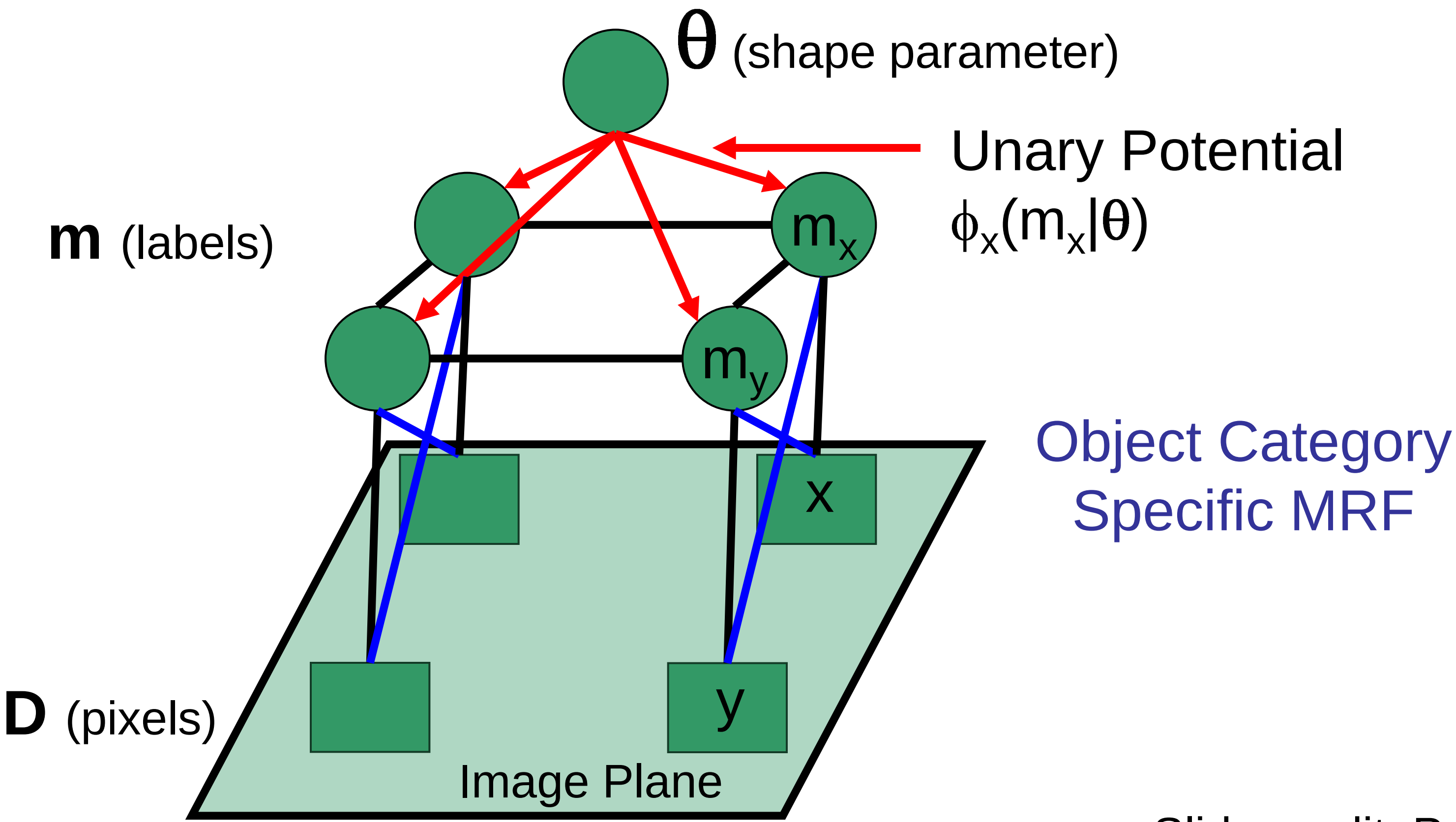
Prior + Contrast

Slide credit: P. Kumar

Our Model

Probability of labelling in addition has

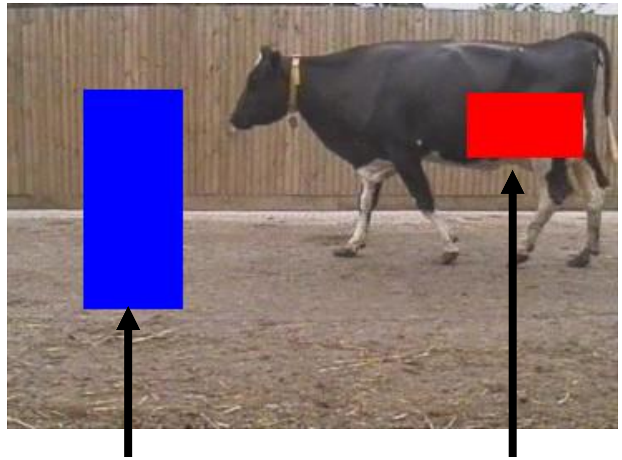
- Unary potential which depend on distance from θ (shape parameter)



Slide credit: P. Kumar

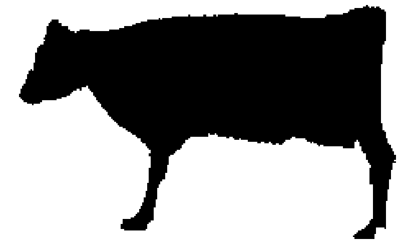
Example

Cow Image

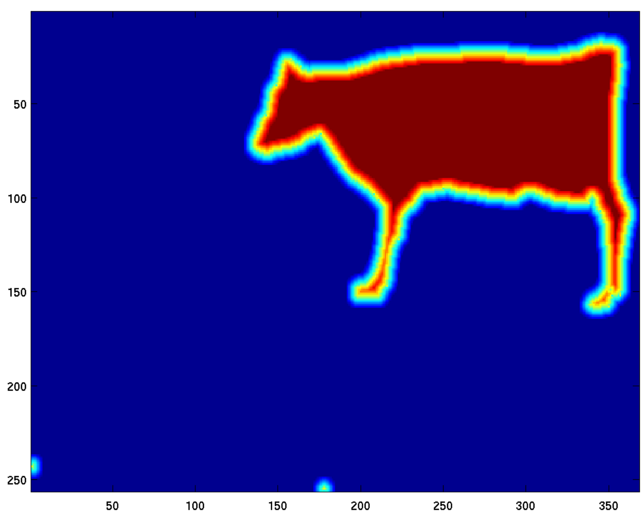


Background Seed Pixels

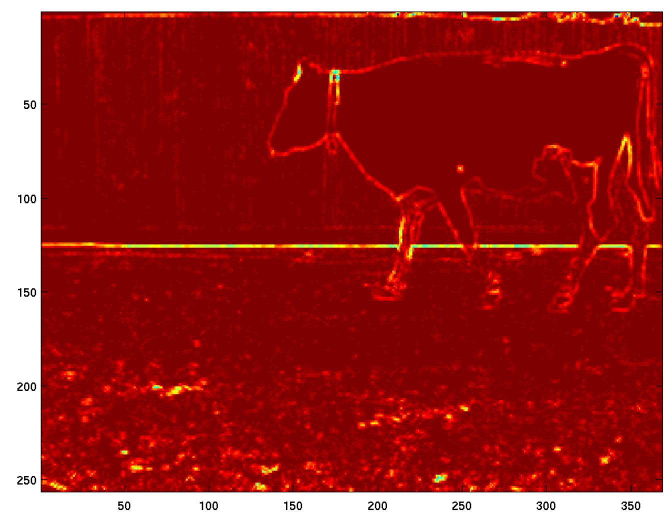
Object Seed Pixels



Shape Prior θ



Distance from θ

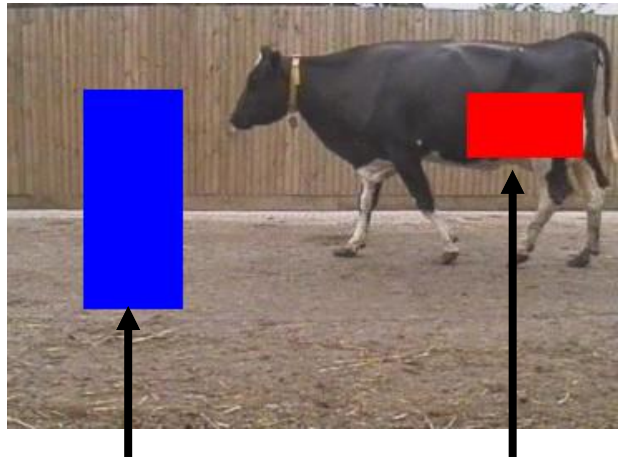


Prior + Contrast

Slide credit: P. Kumar

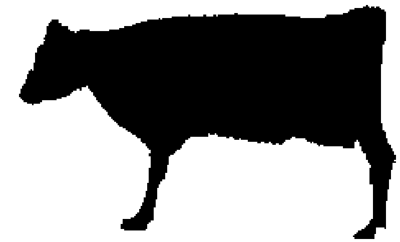
Example

Cow Image

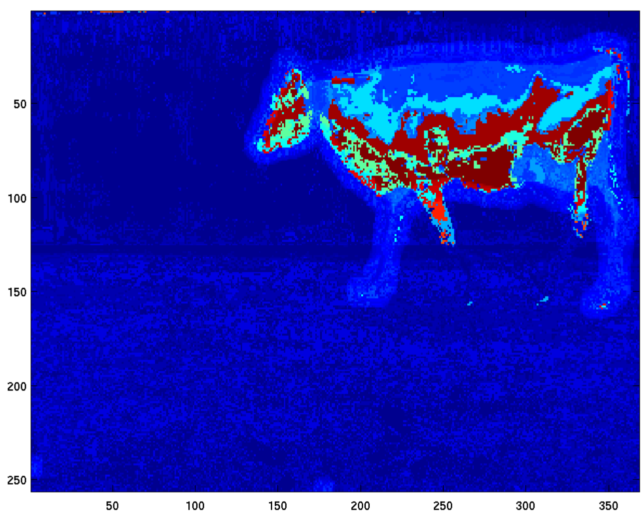


Background Seed Pixels

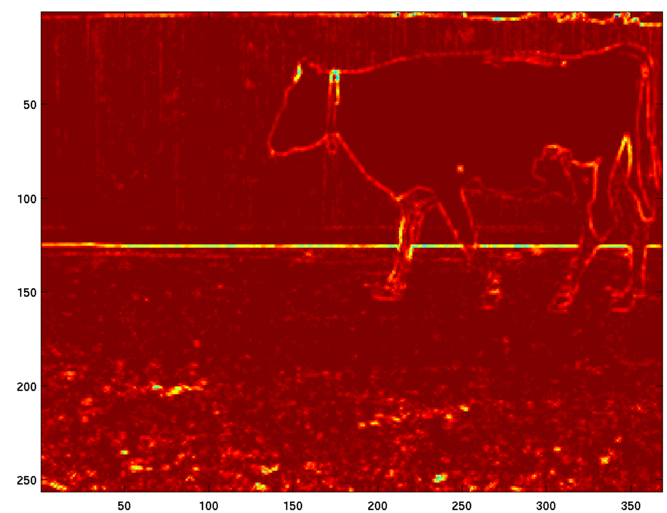
Object Seed Pixels



Shape Prior θ



Likelihood + Distance from θ

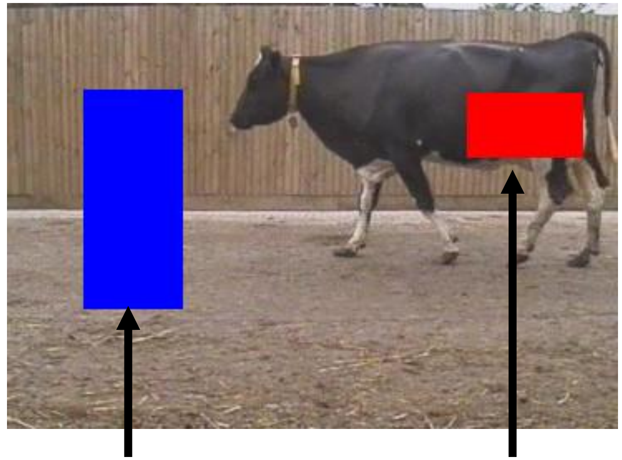


Prior + Contrast

Slide credit: P. Kumar

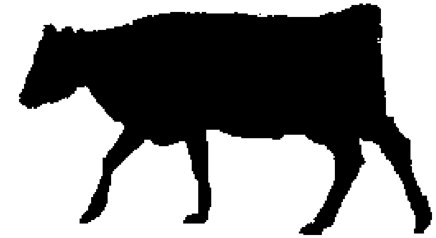
Example

Cow Image

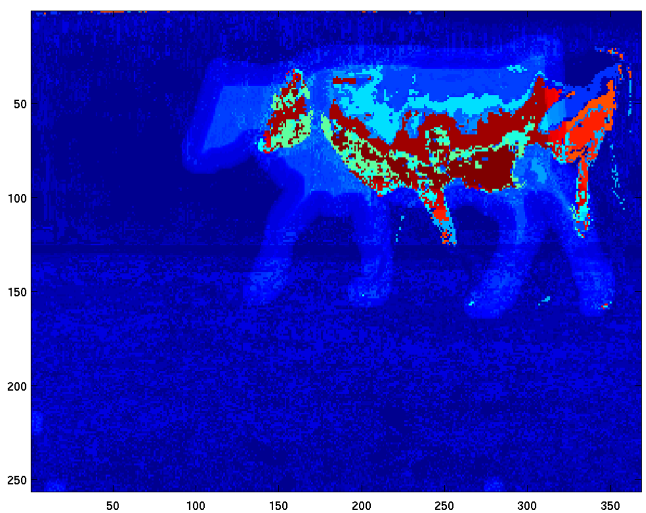


Background Seed Pixels

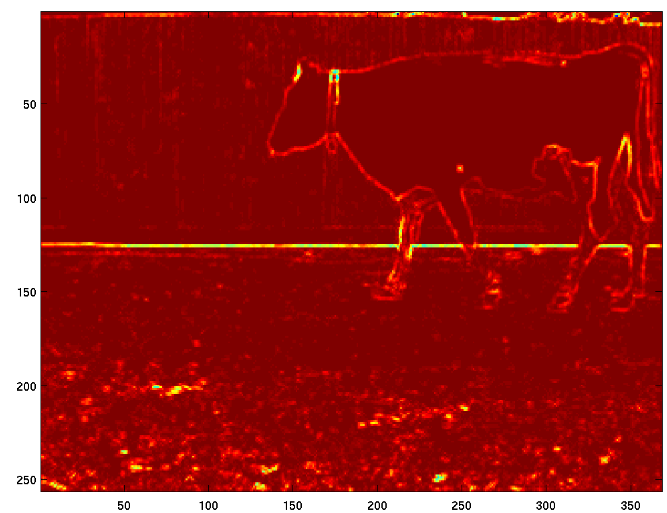
Object Seed Pixels



Shape Prior θ



Likelihood + Distance from θ



Prior + Contrast

Slide credit: P. Kumar

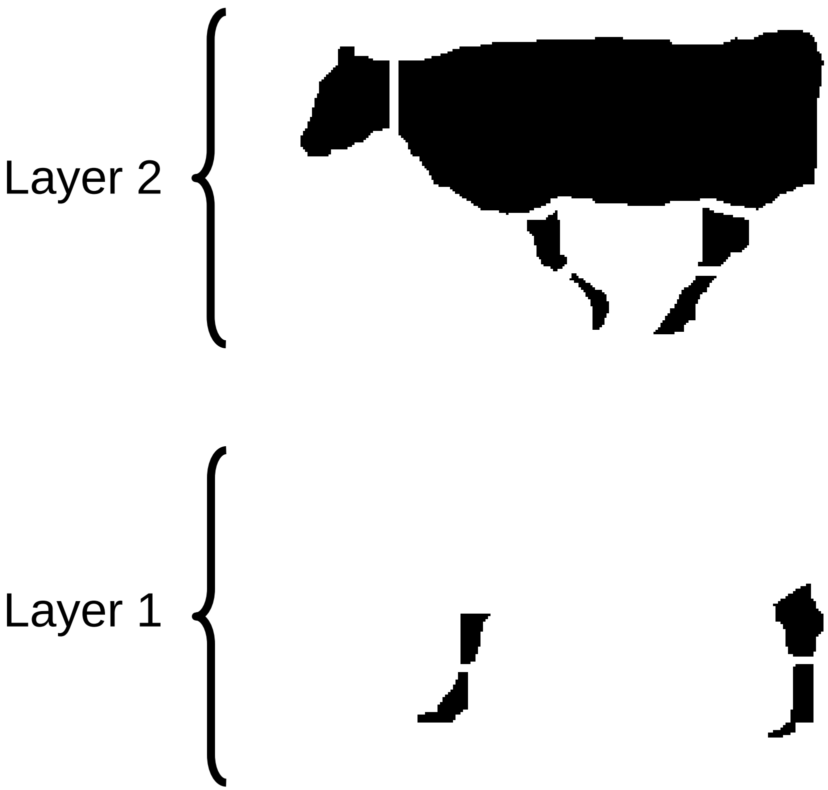
Outline

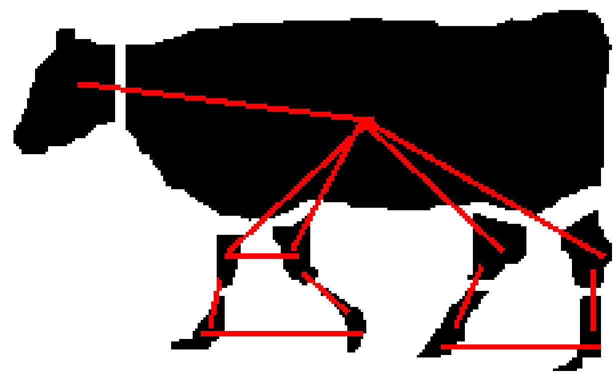
- Problem Formulation
  - Energy E (m, θ) = $\sum\phi_x(\mathbf{D}|m_x)+\phi_x(m_x|\,\theta) + \sum\phi_{xy}(m_x,m_y)+\phi(\mathbf{D}|m_x,m_y)$
- Form of Shape Prior
- Optimization
- Results

Slide credit: P. Kumar

Layered Pictorial Structures (LPS)

- Generative model
- Composition of parts + spatial layout



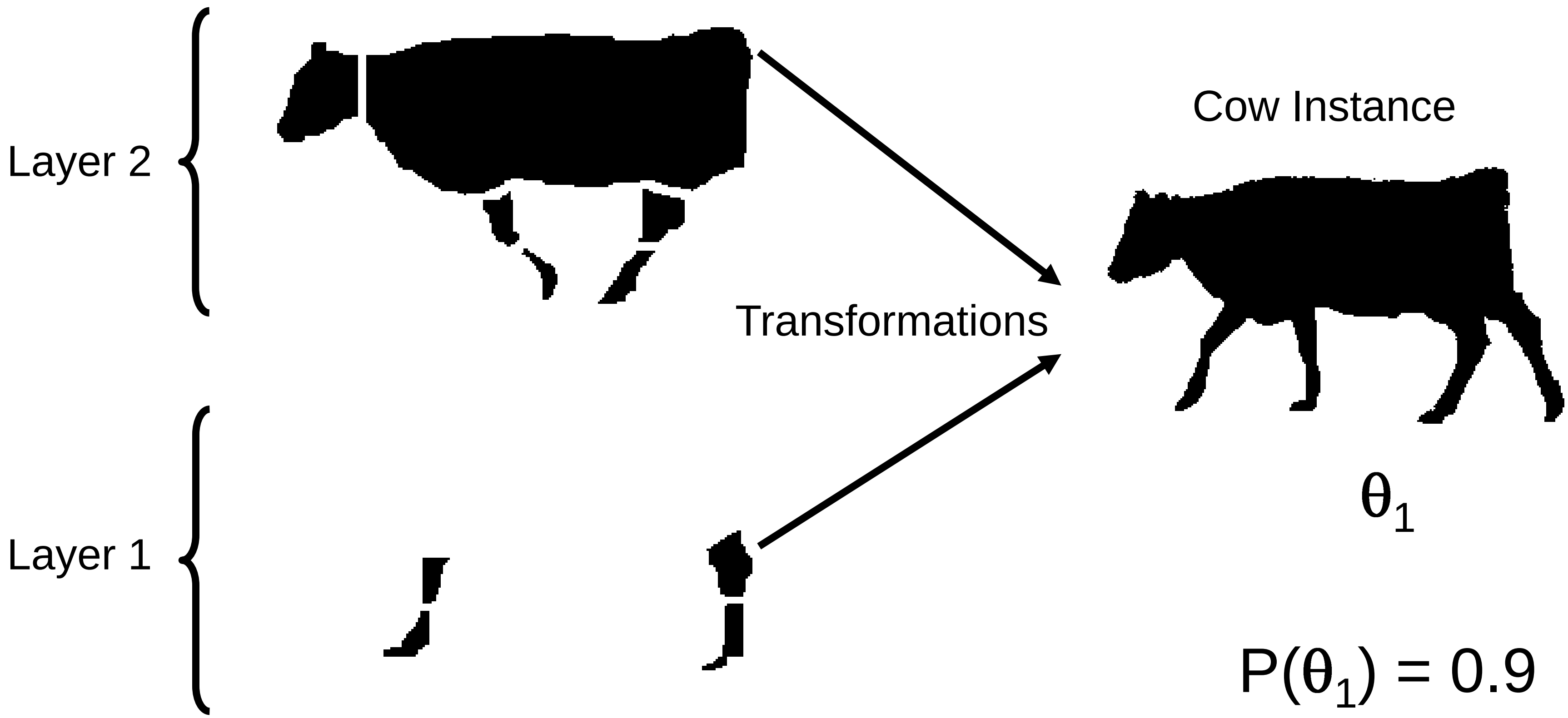


Spatial Layout
(Pairwise Configuration)

Parts in Layer 2 can occlude parts in Layer 1

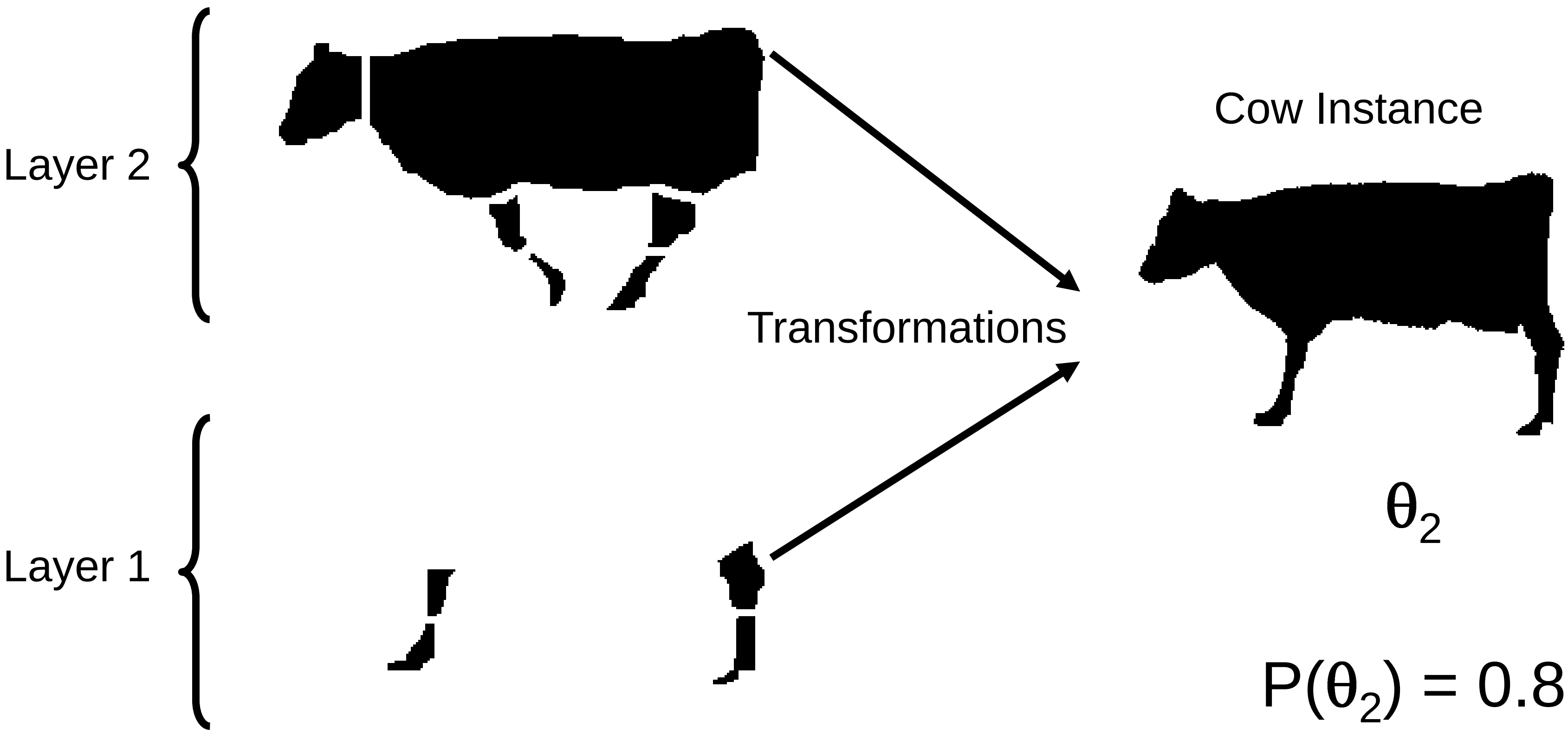
Slide credit: P. Kumar

Layered Pictorial Structures (LPS)



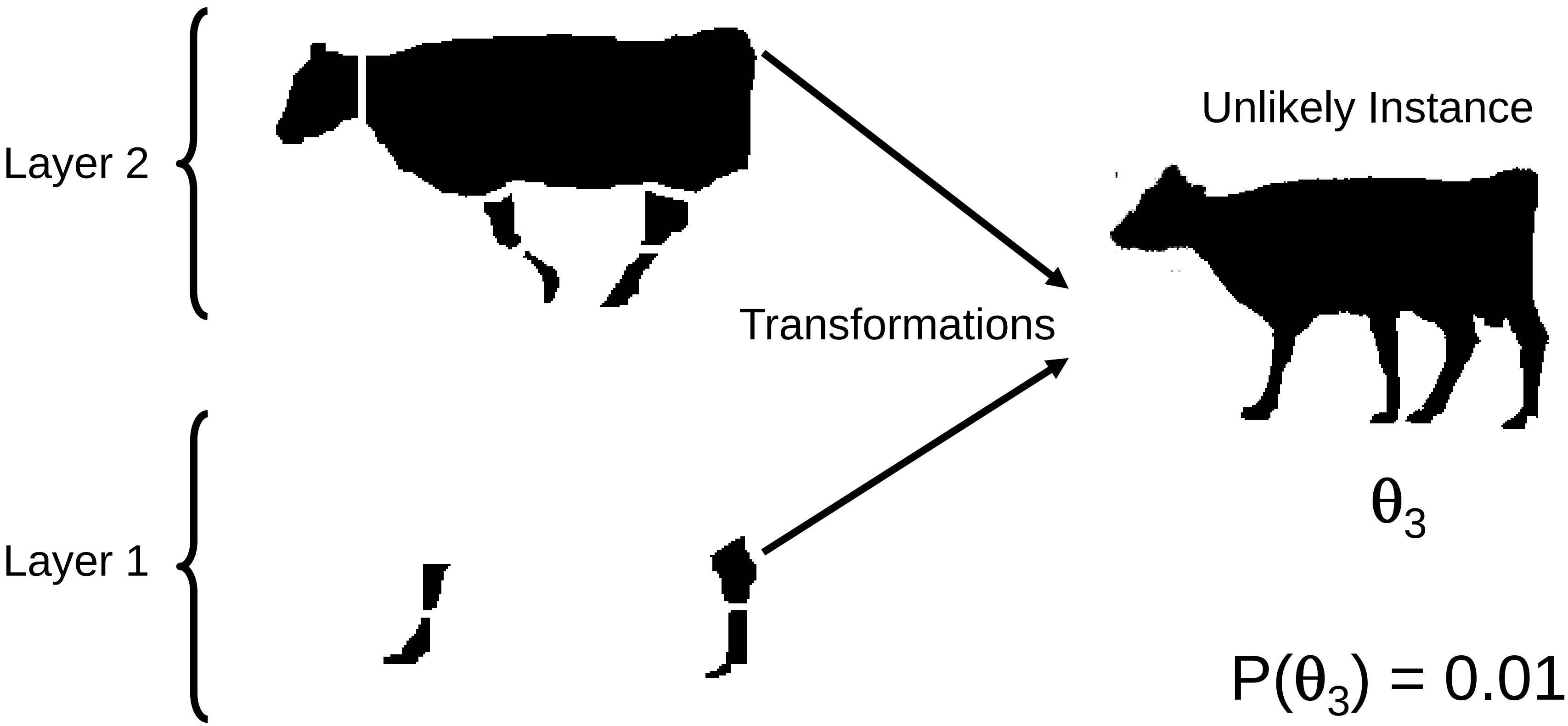
Slide credit: P. Kumar

Layered Pictorial Structures (LPS)



Slide credit: P. Kumar

Layered Pictorial Structures (LPS)



Slide credit: P. Kumar

Outline

- Problem Formulation
- Form of Shape Prior
- Optimization
- Results

Slide credit: P. Kumar

Optimization

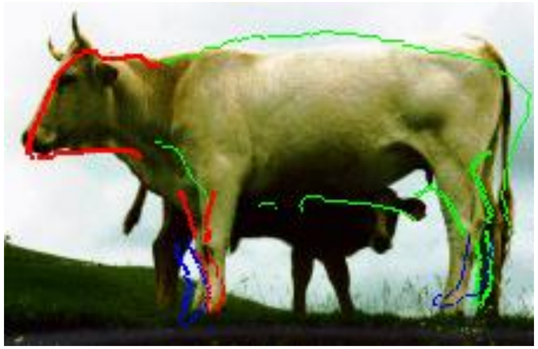
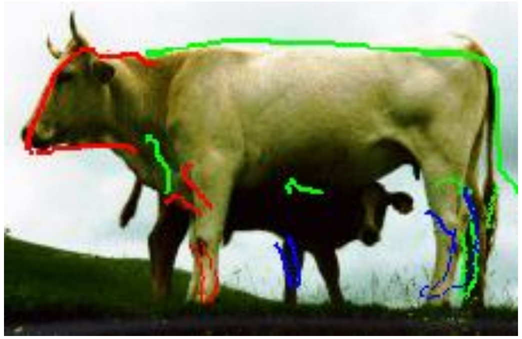
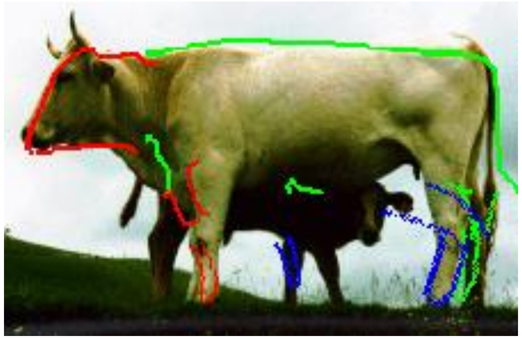
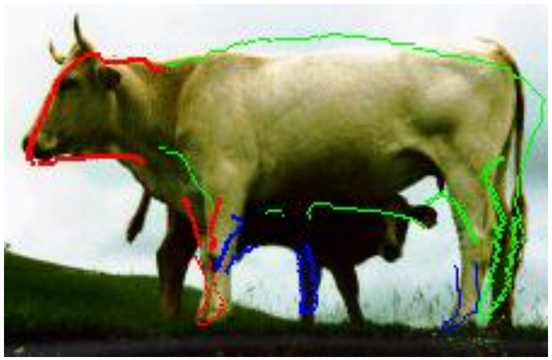
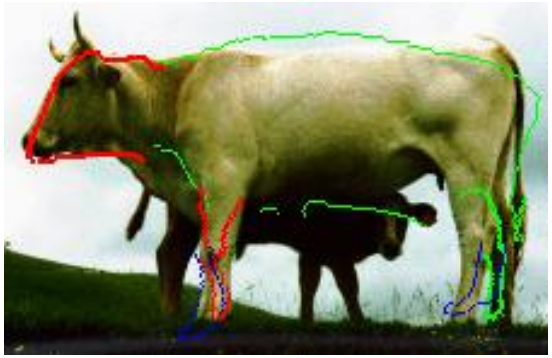
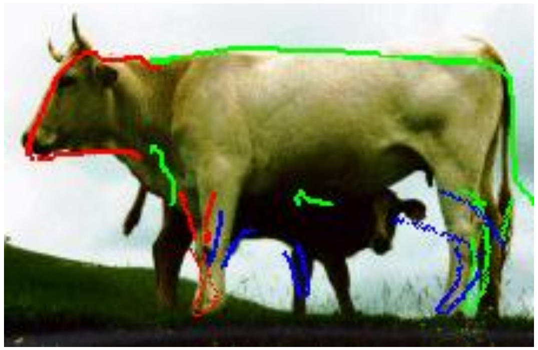
- Given image **D**, find best labelling as
  $\mathbf{m}^* = \arg\max p(\mathbf{m}|\mathbf{D})$

- Treat LPS parameter θ as a latent (hidden) variable

- EM framework
  - E : sample the distribution over θ
  - M : obtain the labelling **m**

Slide credit: P. Kumar

Results of E-Step

- Different samples *localize* different parts well.
- We cannot use only the MAP estimate of the LPS.

Slide credit: P. Kumar

M-Step

- Given samples from p(θ |**m'**,**D**), get new labelling $\mathbf{m}_{new}$

- Sample θ_i provides
  - Object localization to learn RGB distributions of object and background
  - Shape prior for segmentation

- Problem
  - Maximize expected log likelihood using all samples
  - To efficiently obtain the new labelling

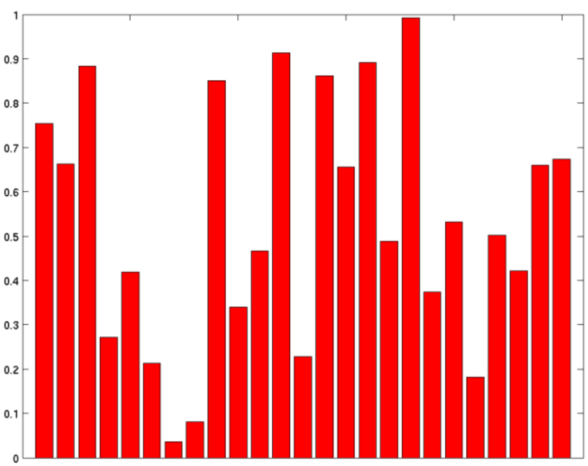
Slide credit: P. Kumar

M-Step

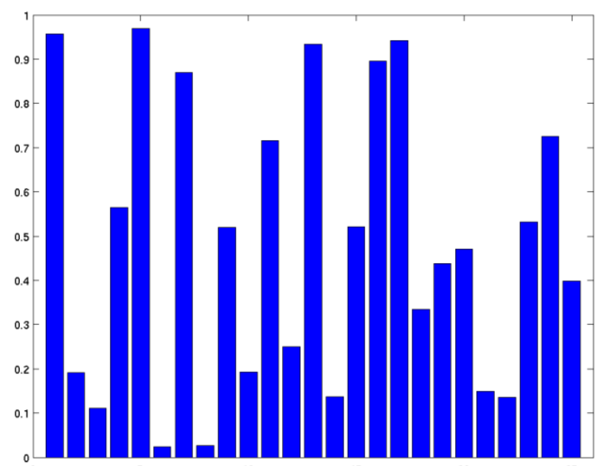
Cow Image

Shape θ_1

$w_1 = P(\theta_1|\mathbf{m}',\mathbf{D})$



RGB Histogram for Object



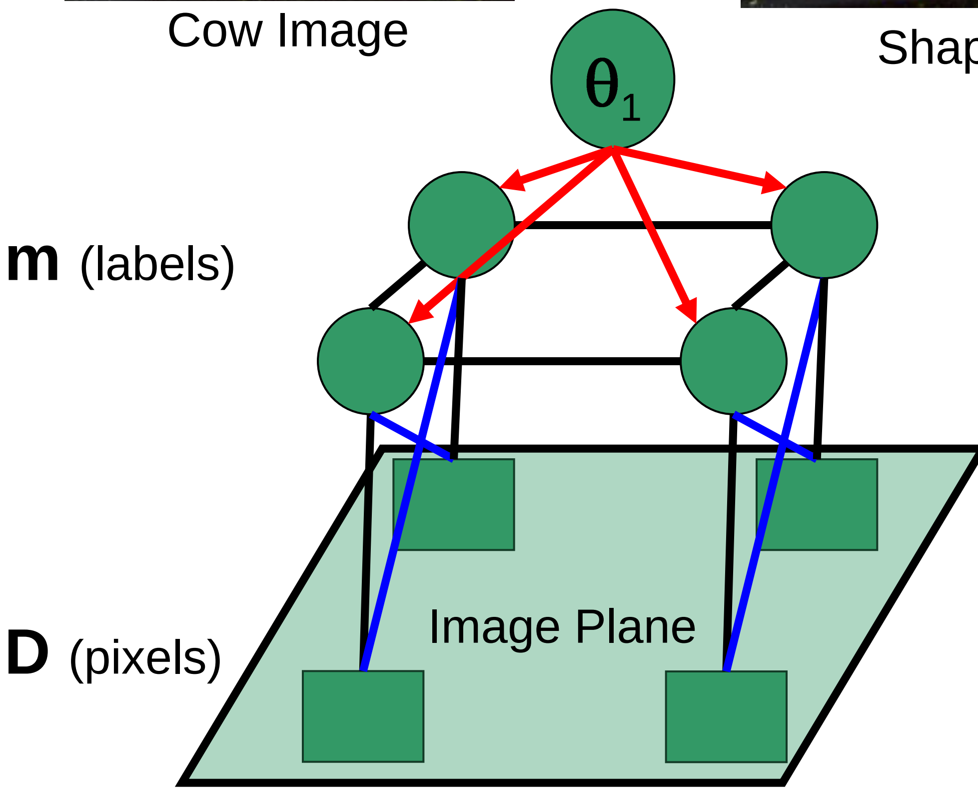
RGB Histogram for Background

M-Step

Cow Image

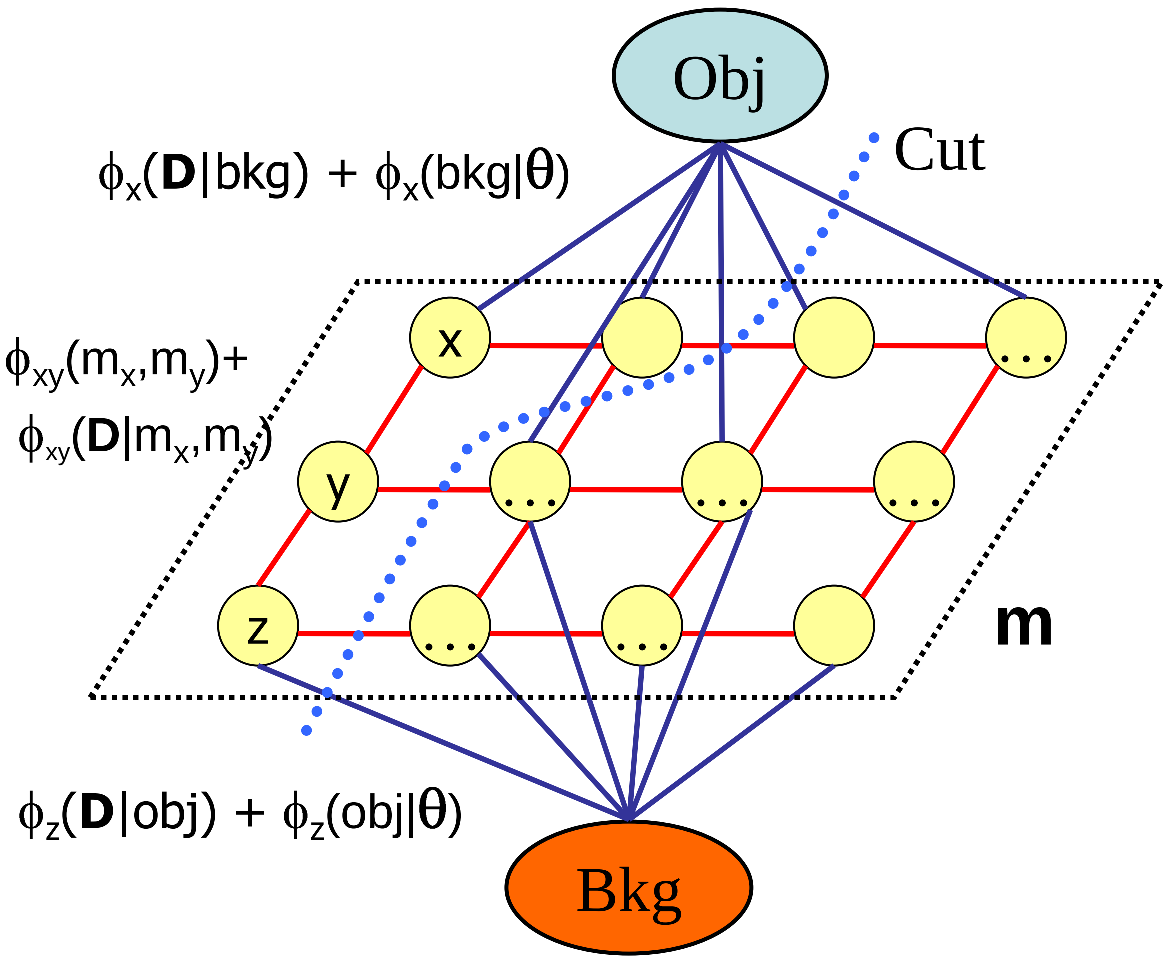
Shape θ_1

$w_1 = P(\theta_1|\mathbf{m}',\mathbf{D})$



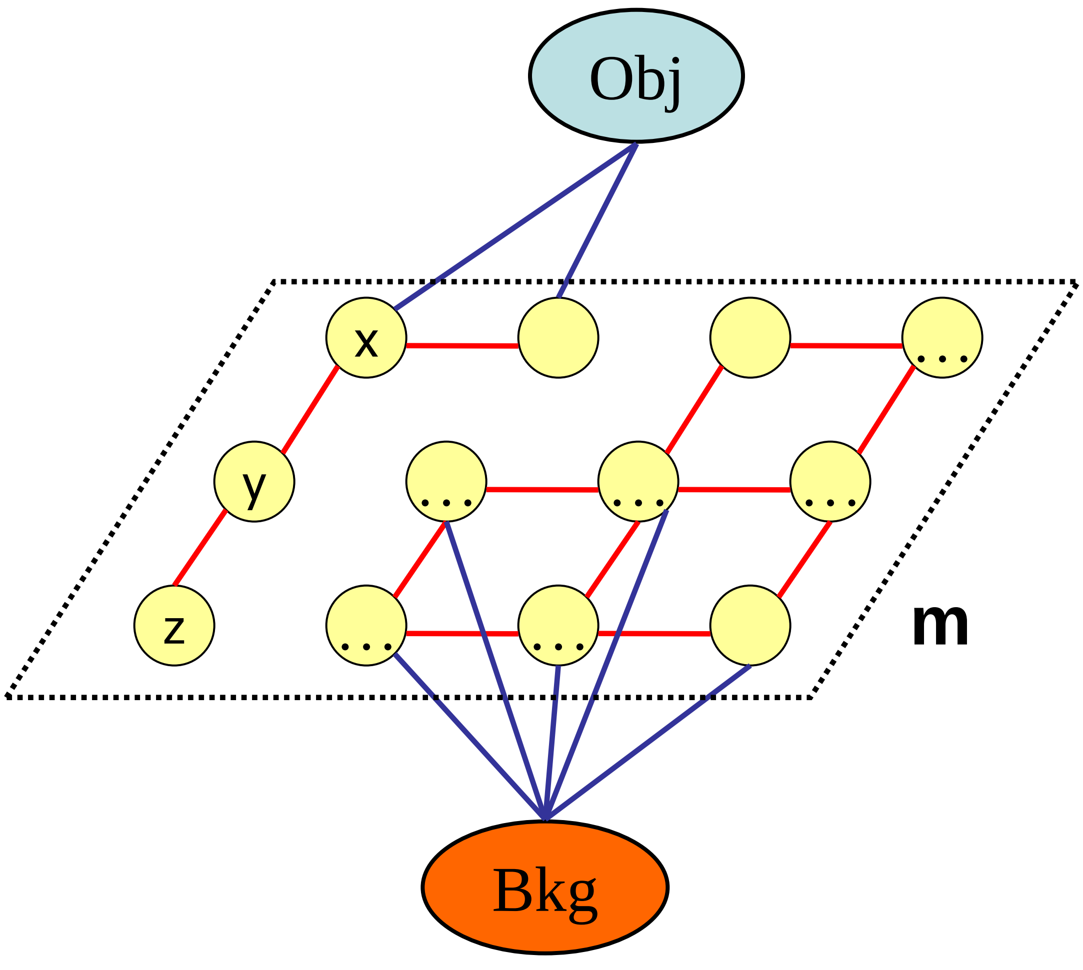
- Best labelling found efficiently using a Single Graph Cut

Segmentation using Graph Cuts



Slide credit: P. Kumar

Segmentation using Graph Cuts



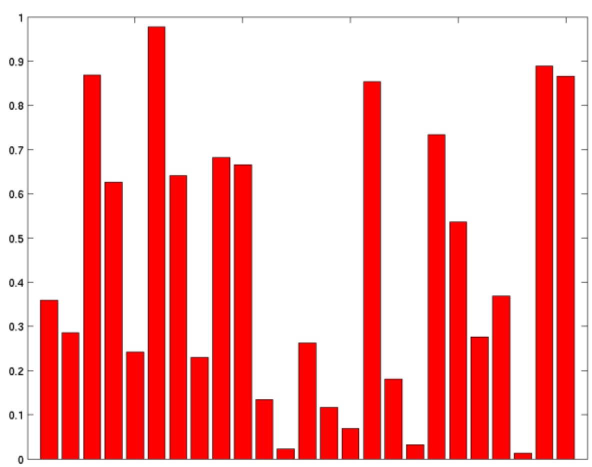
Slide credit: P. Kumar

M-Step

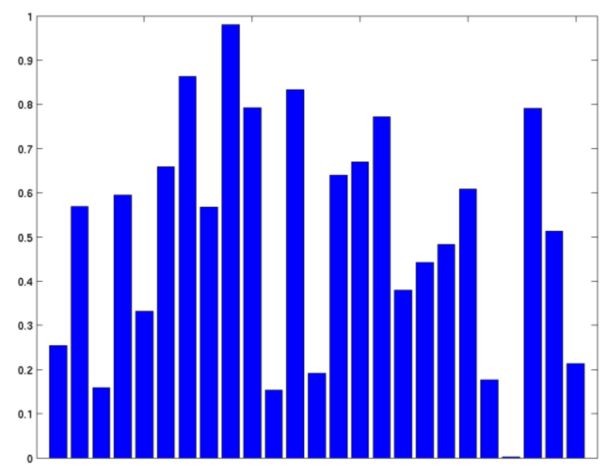
Cow Image

Shape θ_2

$w_2 = P(\theta_2 | \mathbf{m}', \mathbf{D})$



RGB Histogram for Object



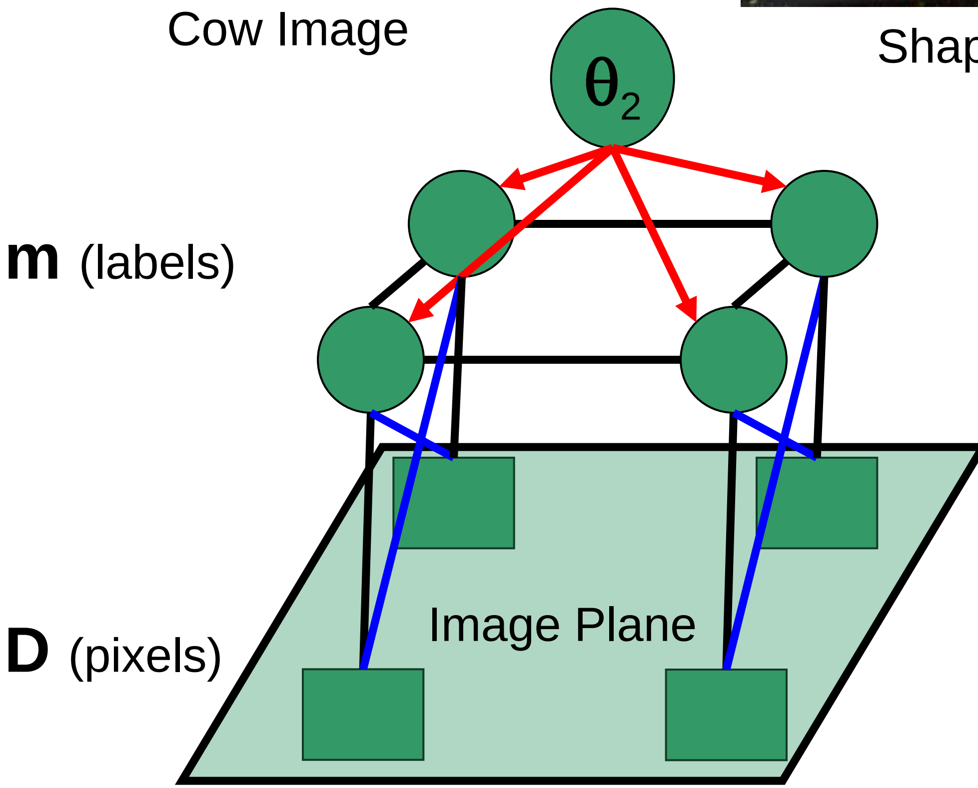
RGB Histogram for Background

M-Step

Cow Image

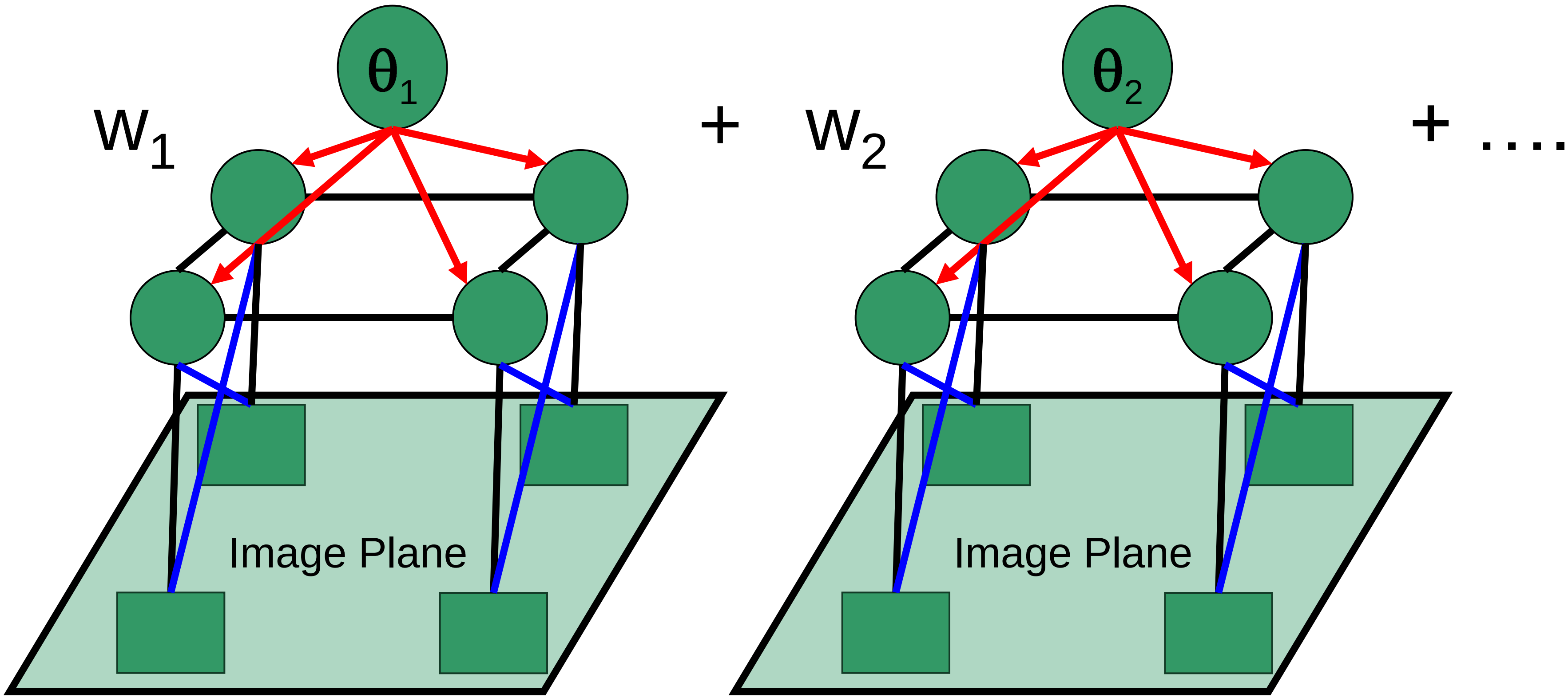
Shape θ_2

$w_2 = P(\theta_2|\mathbf{m}',\mathbf{D})$



- Best labelling found efficiently using a Single Graph Cut

M-Step



$\mathbf{m}^* = \arg\min \sum w_i \, E\,(\mathbf{m}, \theta_i)$

- Best labelling found efficiently using a Single Graph Cut

Outline

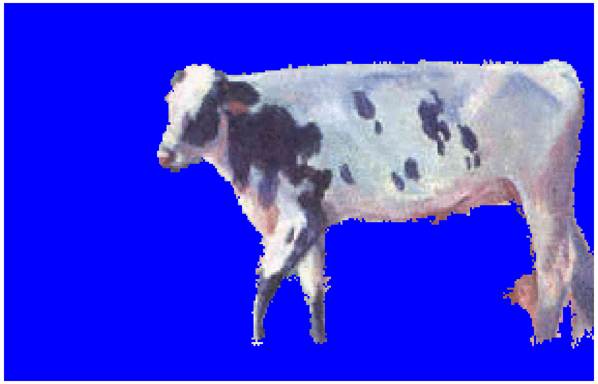
- Problem Formulation
- Form of Shape Prior
- Optimization
- Results

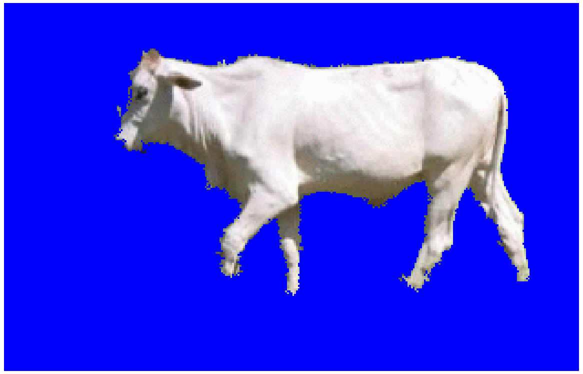
Results

Using LPS Model for Cow

Image

Segmentation





Slide credit: P. Kumar

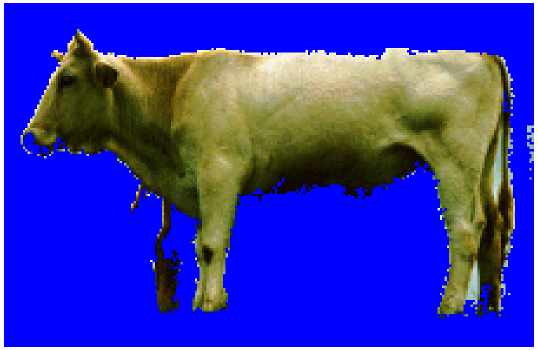
Results

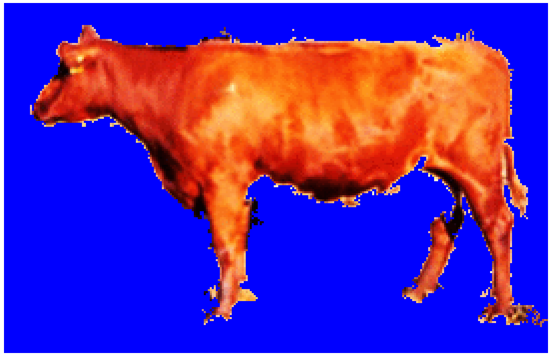
Using LPS Model for Cow

In the absence of a clear boundary between object and background

Image

Segmentation





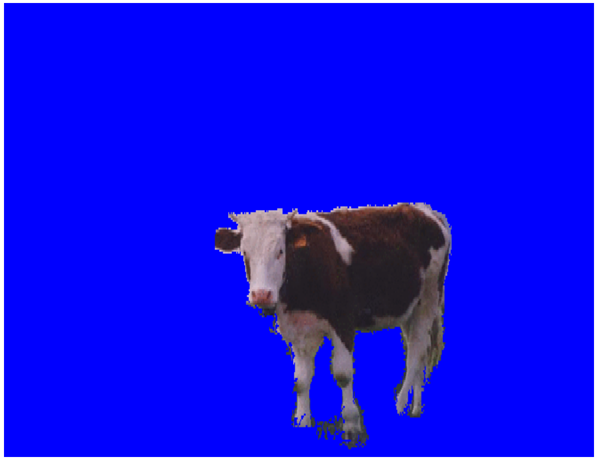
Slide credit: P. Kumar

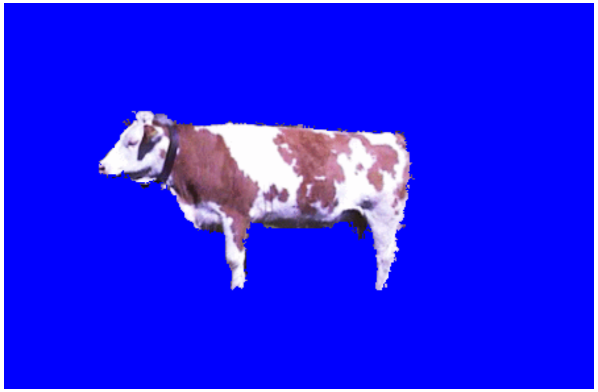
Results

Using LPS Model for Cow

Image

Segmentation

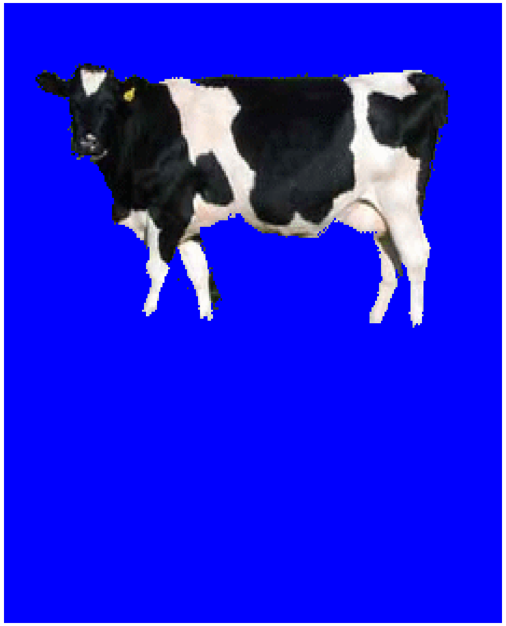




Slide credit: P. Kumar

Results

Using LPS Model for Cow

Image	Segmentation
	
	

Results

Using LPS Model for Horse

Image	Segmentation
	
	

Results

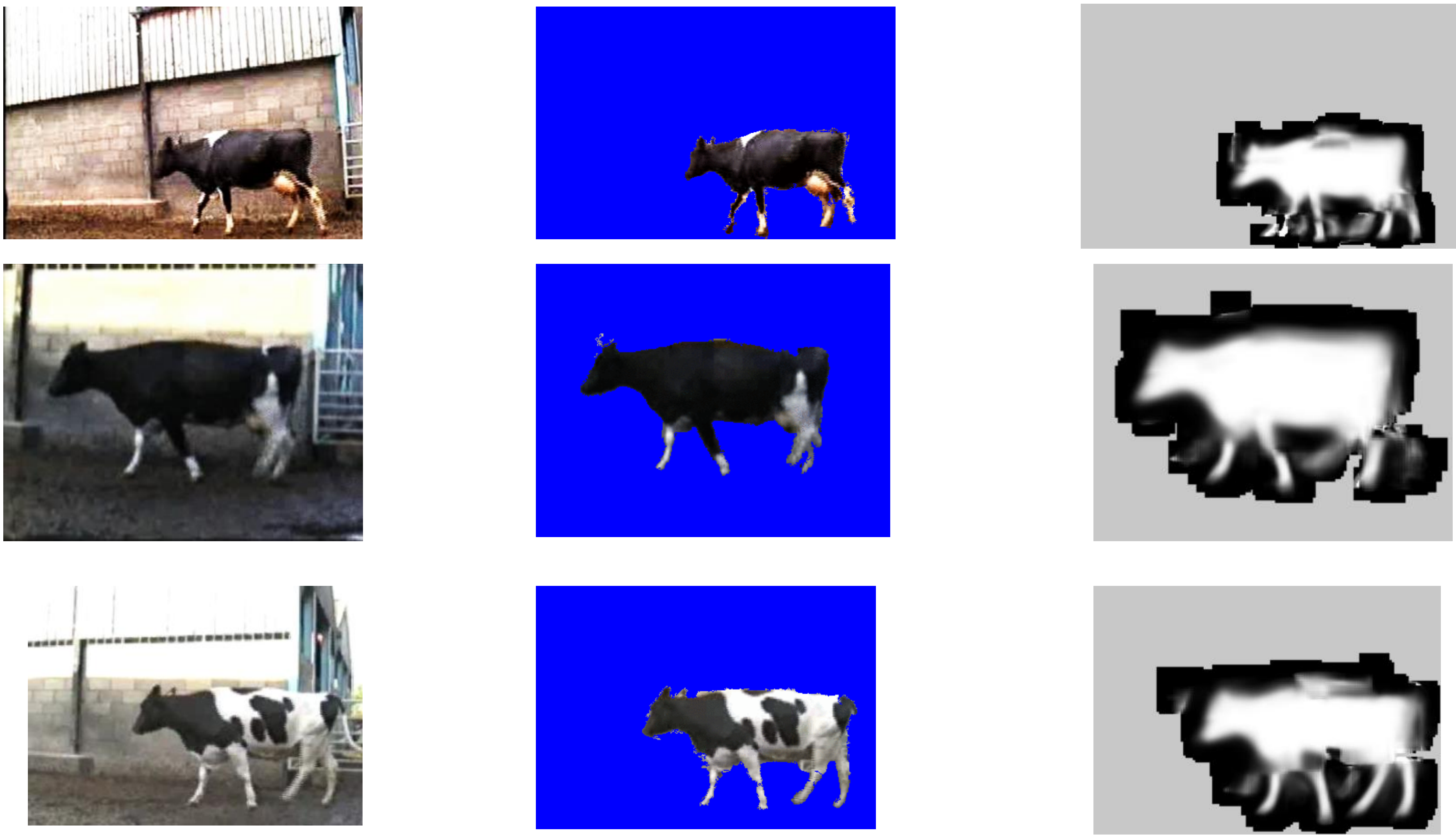
Using LPS Model for Horse

Image

Segmentation

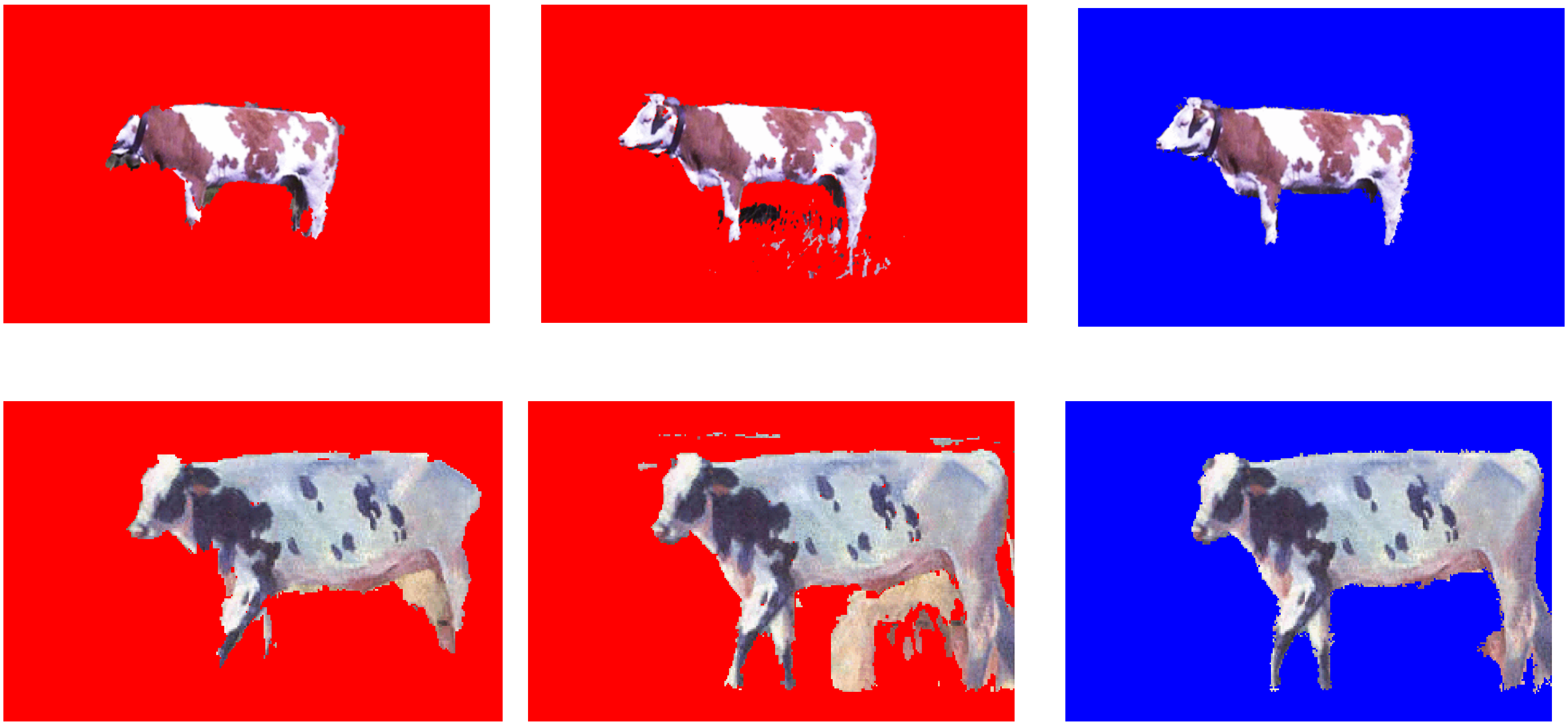
Results

Image | Our Method | Leibe and Schiele



Results

Shape | Appearance | Shape+Appearance



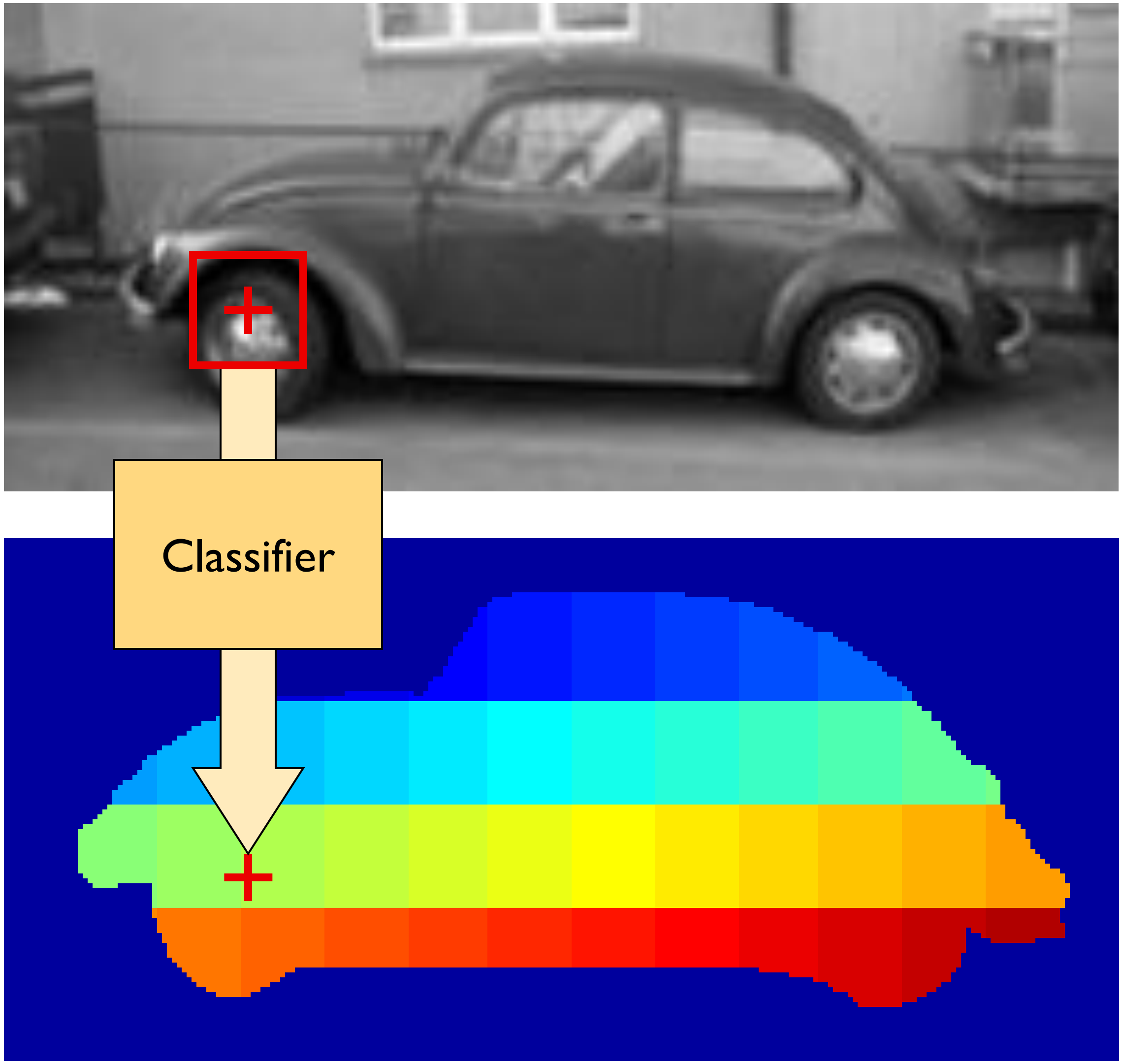
Without $\phi_x(\mathbf{D}|m_x)$ | **Without** $\phi_x(m_x|\theta)$

Overview

- Bottom-up segmentation
  - Clustering
  - Mean shift
  - Graph-based
- Combining object recognition & segmentation
  - OBJCUT
  - Other methods

Layout Consistent Random Field

Winn and Shotton 2006

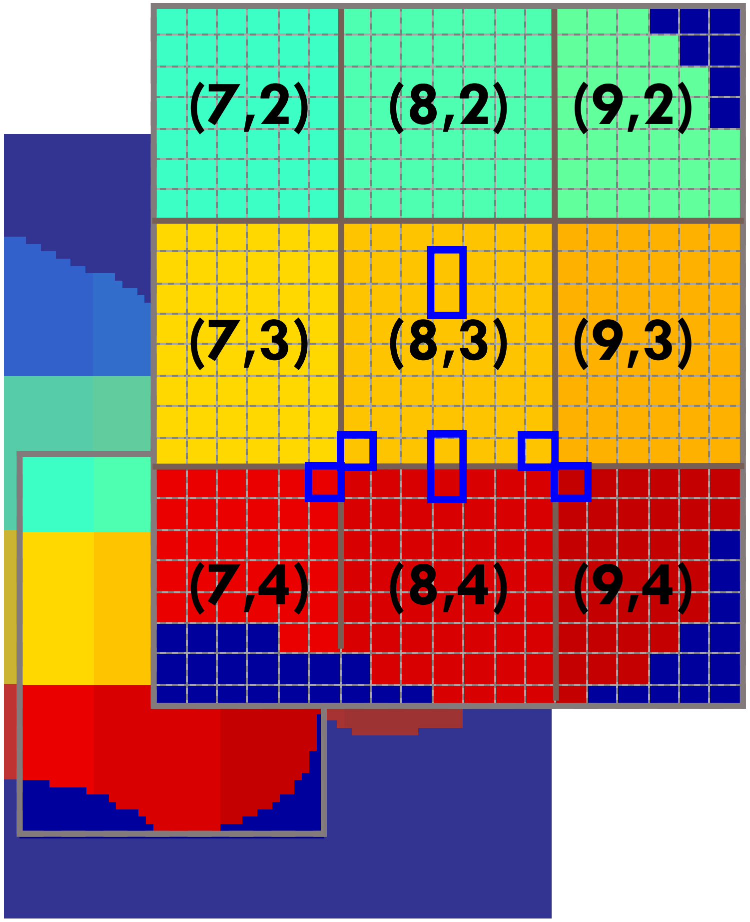


- Decision forest classifier
- Features are differences of pixel intensities

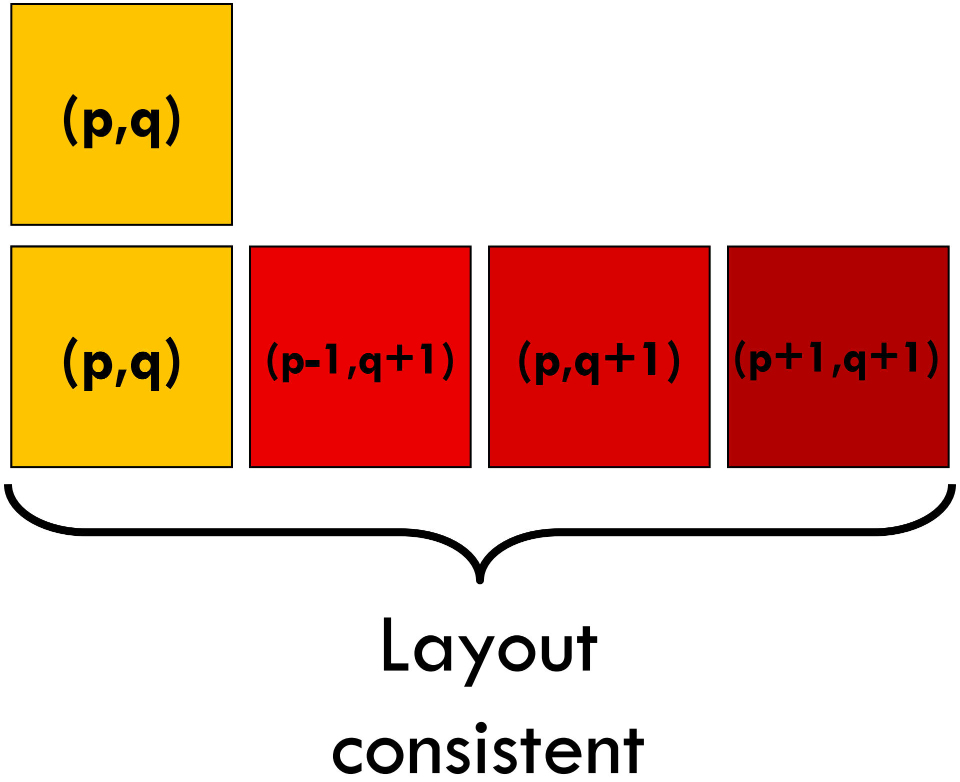
[Lepetit et al. CVPR 2005]

Layout consistency

Winn and Shotton 2006



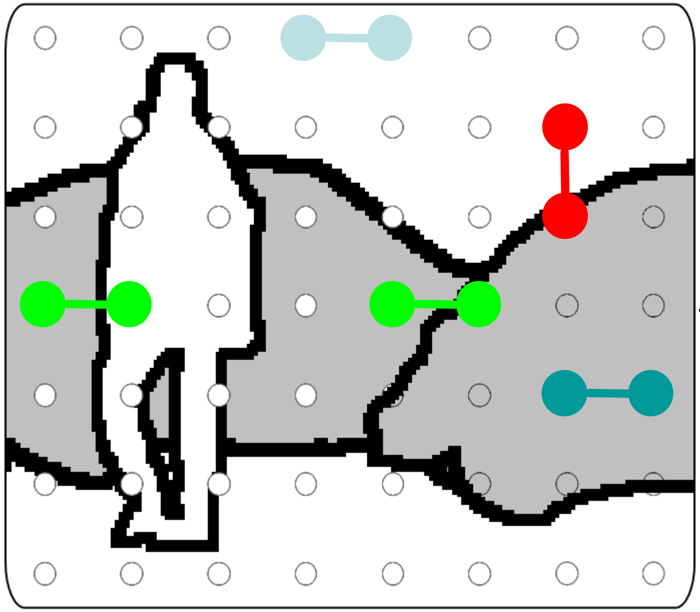
Neighboring pixels



Layout Consistent Random Field

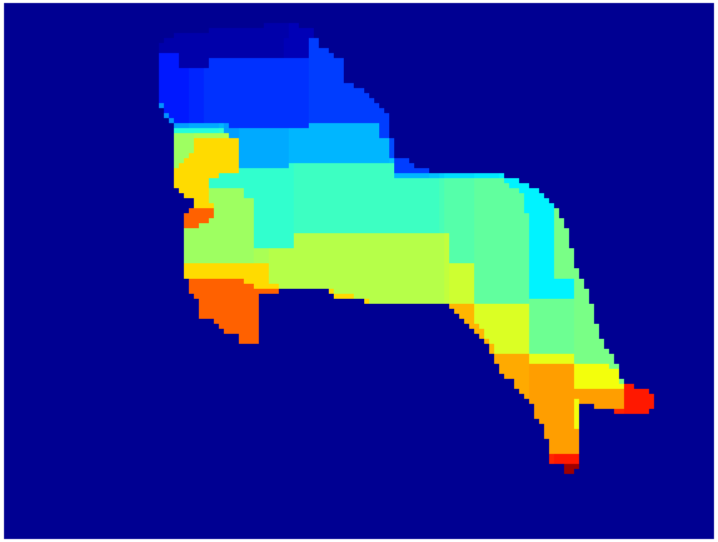
Winn and Shotton 2006

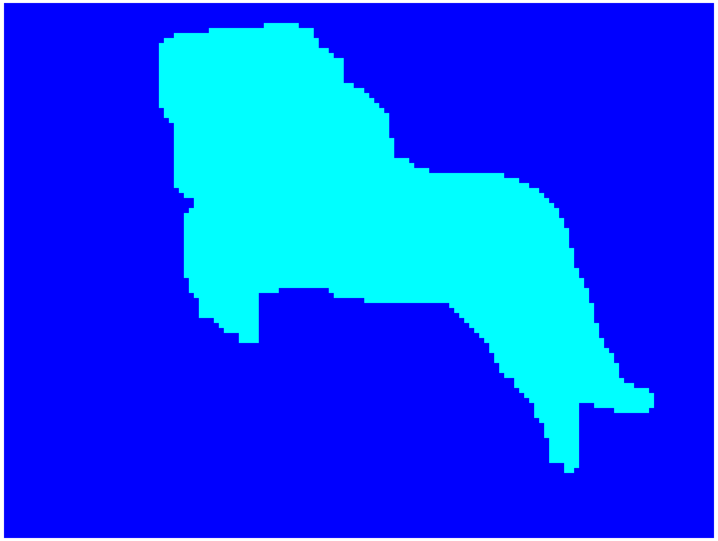
$$P(\mathbf{h} \mid \mathbf{I}; \boldsymbol{\theta}) \propto \underbrace{\prod_i \phi_i(h_i, \mathbf{I}; \theta)}_{\text{Part detector}} \underbrace{\prod_{(i,j)\in E} \psi_{ij}(h_i, h_j, \mathbf{I}; \theta')}_{\text{Layout consistency}}$$



$$-\log \psi_{ij} = \begin{cases} 0 & \text{Consistent foreground} \\ \beta_{\text{bg}} & \text{Background} \\ \beta_{\text{oe}}.e_{ij} & \text{Object edge} \\ \beta_{\text{co}}.e_{ij} & \text{Object occlusion} \\ \beta_{\text{iif}} & \text{Inconsistent} \end{cases}$$

Stability of part labelling





Part color key

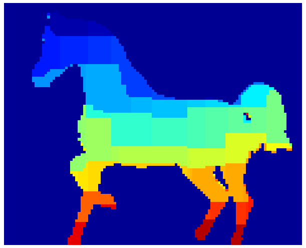


Image parsing: Tu, Zhu and Yuille 2003

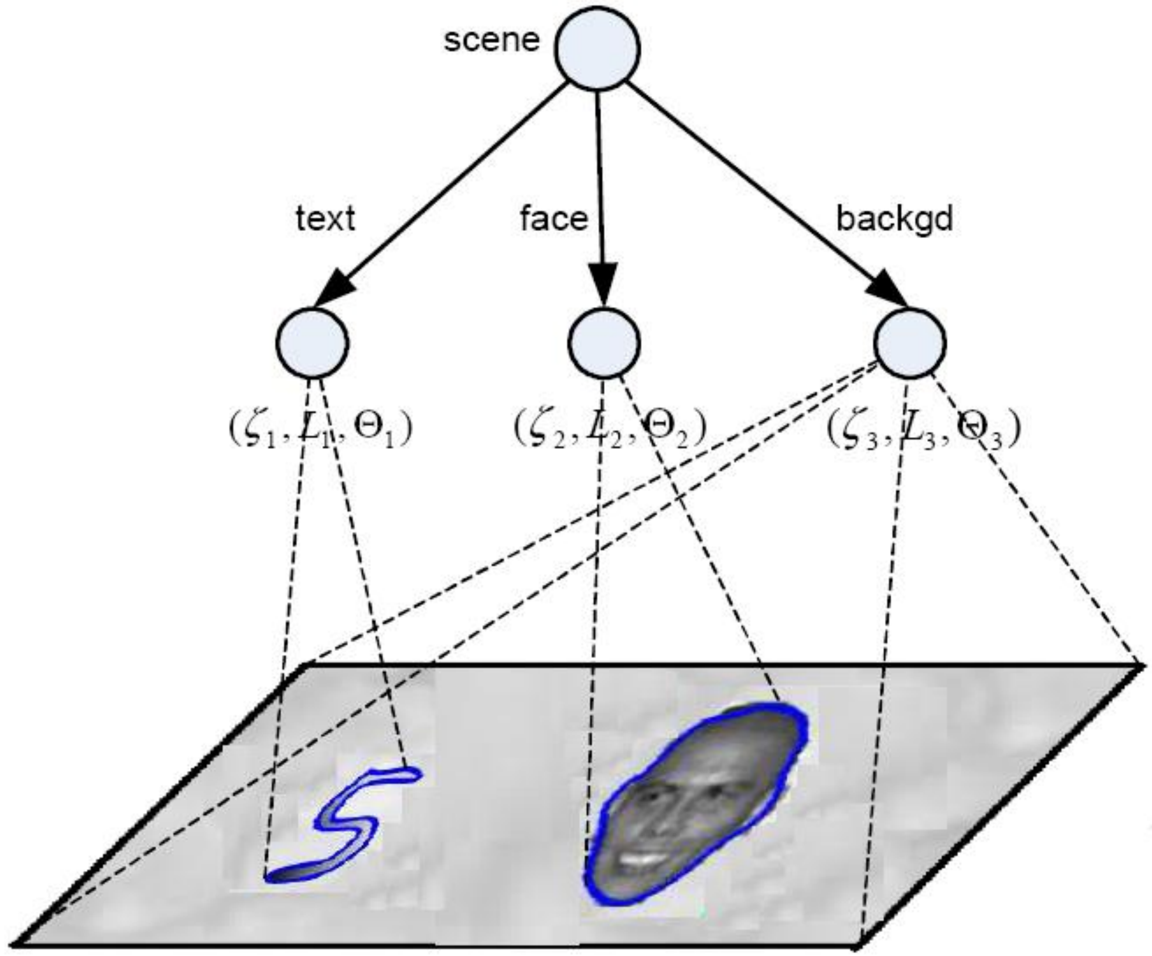
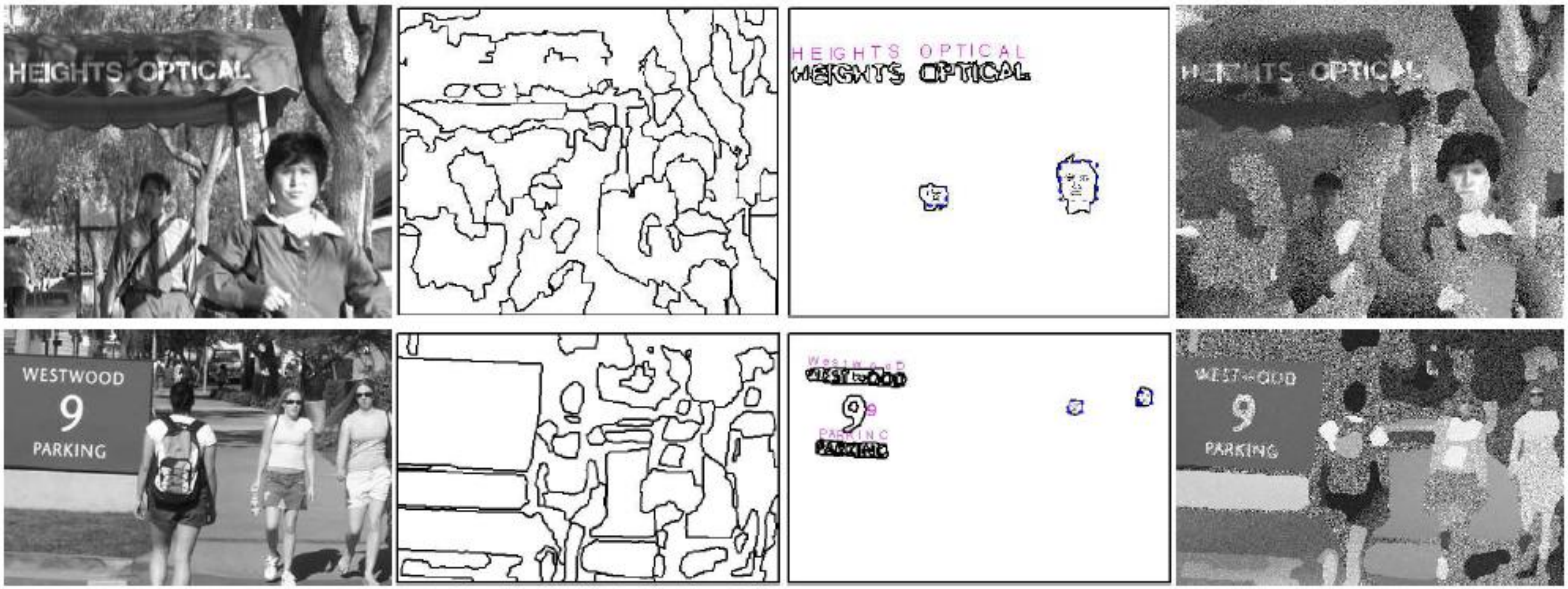


Image parsing: Tu, Zhu and Yuille 2003



a. Input image b. Segmentation c. Object recognition d. Synthesized image

LOCUS model

Kannan, Jojic and Frey 2004
Winn and Jojic, 2005