Segmentation

Lu Peng
School of Computer Science,
Beijing University of Posts and Telecommunications

Machine Vision Technology							
Semantic information				Metric 3D information			
Pixels	Segments	Images	Videos	Camera		Multi-view Geometry	
Convolutions Edges & Fitting Local features Texture	Segmentation Clustering	Recognition Detection	Motion Tracking	Camera Model	Camera Calibration	Epipolar Geometry	SFM
10	4	4	2	2	2	2	2

Image segmentation

Source： S. Lazebnik

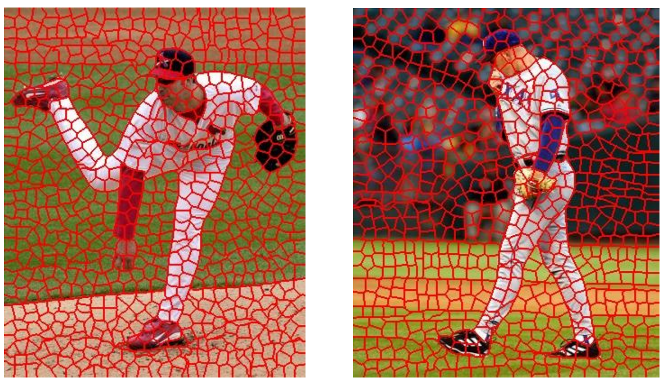
The goals of segmentation

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

先把看起来相似的 pixels 组一块

"superpixels"

目前还是过分割的一个状态

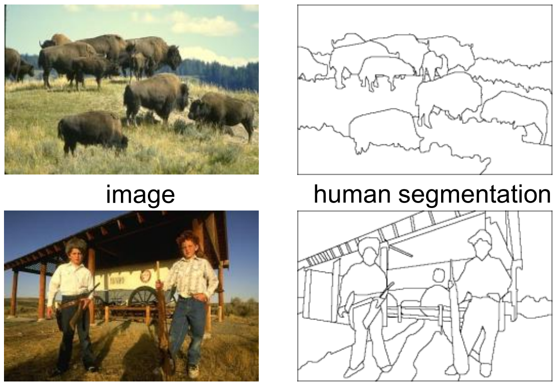


X. Ren and J. Malik. Learning a classification model for segmentation. ICCV 2003.

Source： S. Lazebnik

The goals of segmentation

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

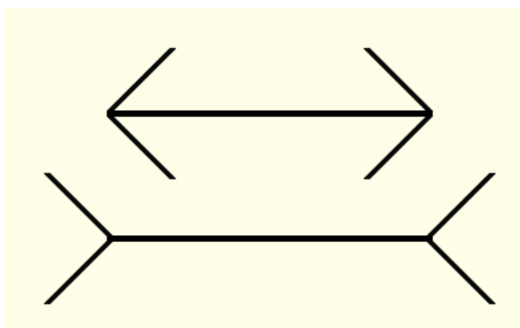


Berkeley segmentation database:
http://www.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/segbench/

Inspiration from psychology

- The Gestalt school: Grouping is key to visual perception

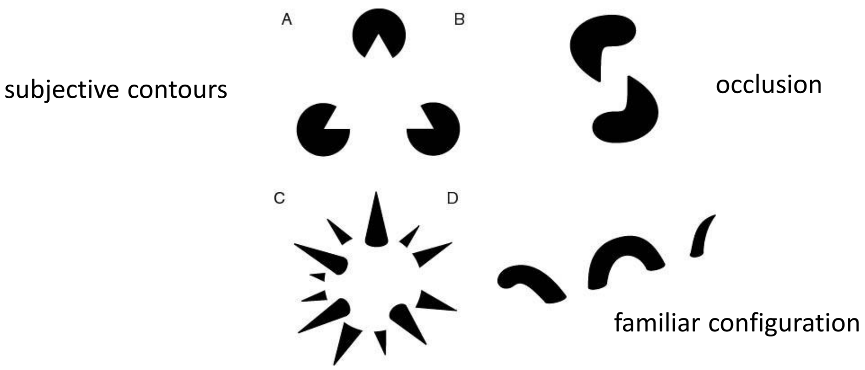
The Muller-Lyer illusion



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

The Gestalt school

- Elements in a collection can have properties that result from relationships
  - “The whole is greater than the sum of its parts”



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

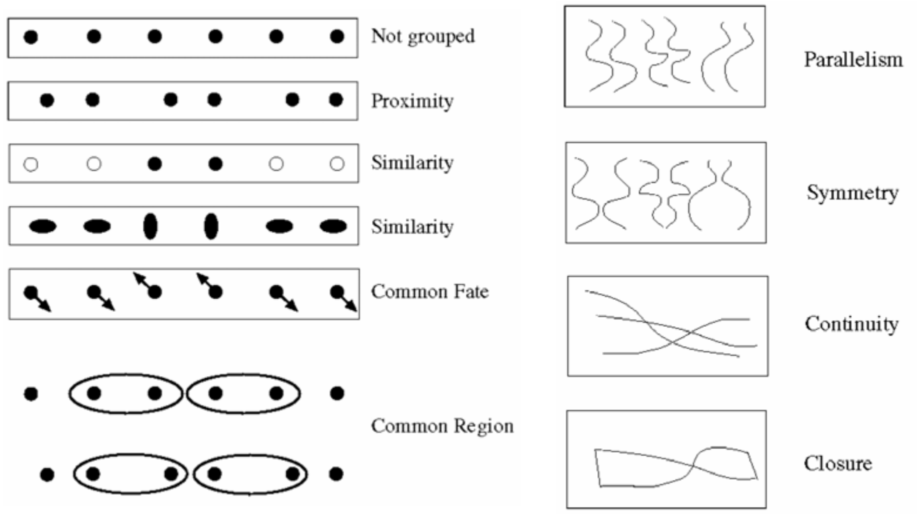
Source：S. Lazebnik

Emergence

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

Gestalt factors

分割依据



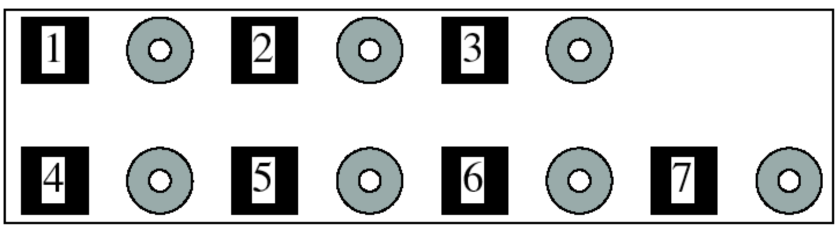
对称

连续

封闭

Source： S. Lazebnik

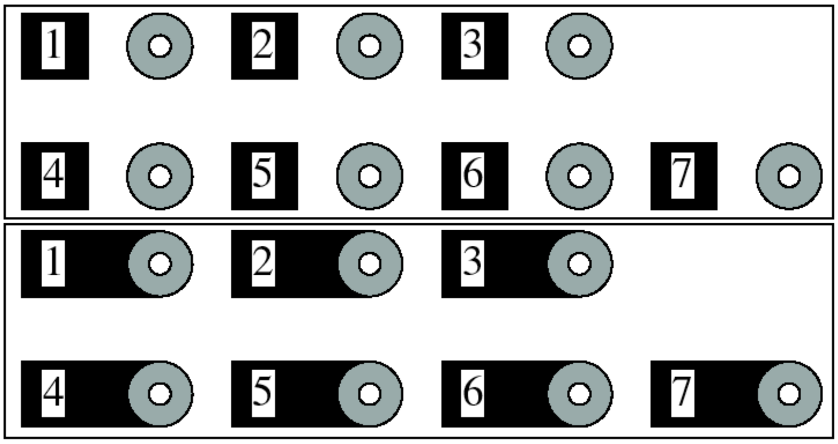
Grouping phenomena in real life



Forsyth & Ponce, Figure 14.7

Source： S. Lazebnik

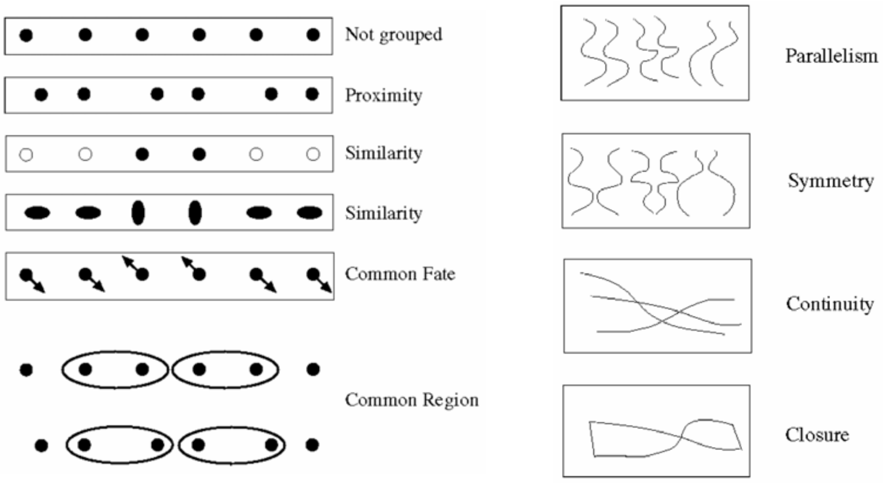
Grouping phenomena in real life



Forsyth & Ponce, Figure 14.7

Source： S. Lazebnik

Gestalt factors

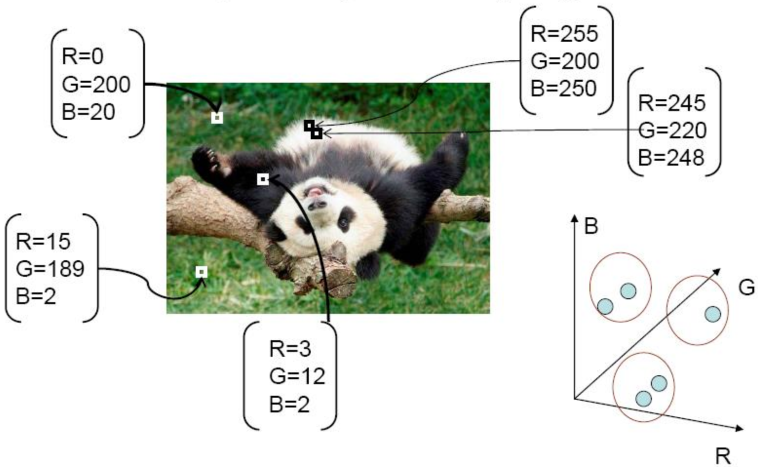


These factors make intuitive sense, but are very difficult to translate into algorithms

Source： S. Lazebnik

Segmentation as clustering

- Cluster similar pixels (features) together



色彩空间

Source: K. Grauman

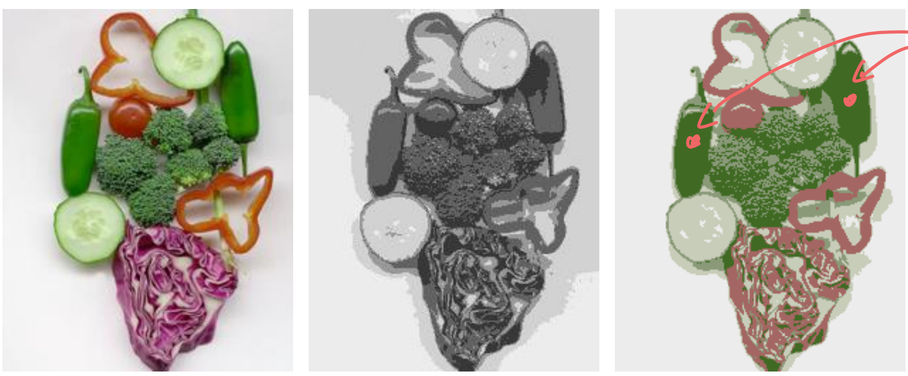
Segmentation as clustering

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

灰度图，一维空间聚类

彩色，三维空间聚类

Image | Intensity-based clusters | Color-based clusters



分类出了辣椒，但无法区分哪个pixel属于哪个辣椒（语义分割）

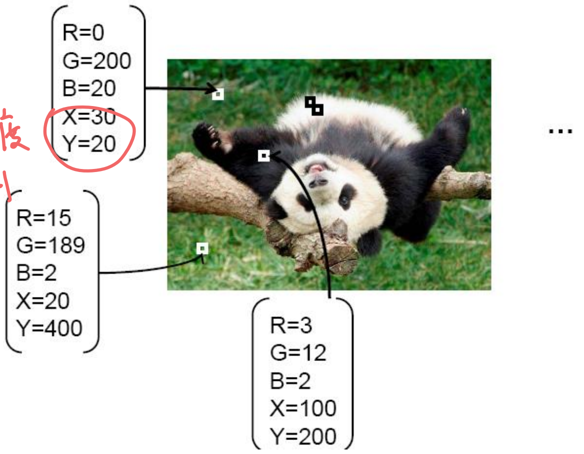
如果能区分，这就是 实例分割

Source： S. Lazebnik

Segmentation as clustering

- Cluster similar pixels (features) together

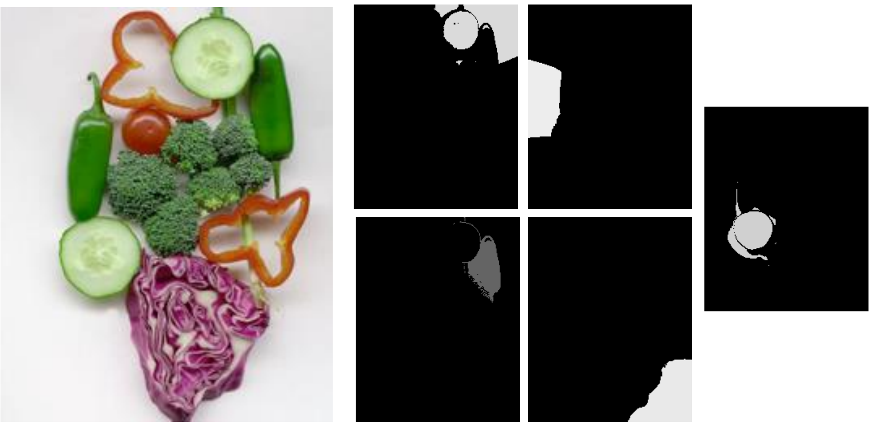
增加维度来区分案例



Source: K. Grauman

Segmentation as clustering

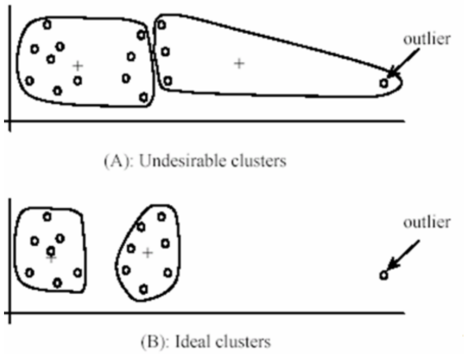
- Clustering based on (r,g,b,x,y) values enforces more spatial coherence



Source： S. Lazebnik

K-Means for segmentation

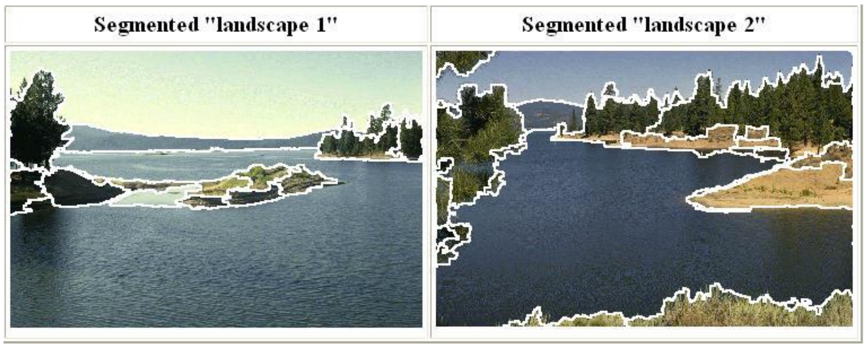
- Pros
  - Very simple method
  - Converges to a local minimum of the error function
- Cons
  - Memory-intensive
  - Need to pick K
  - Sensitive to initialization
  - Sensitive to outliers
  - Only finds "spherical" clusters



Source： S. Lazebnik

Mean shift clustering and segmentation

- An advanced and versatile technique for clustering-based segmentation



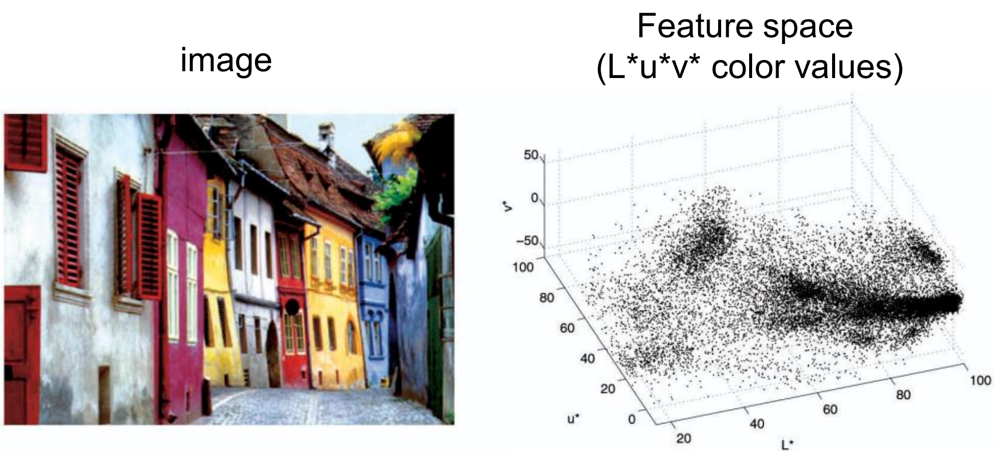
http://www.caip.rutgers.edu/~comanici/MSPAMI/msPamiResults.html

模式分析和机器智能

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

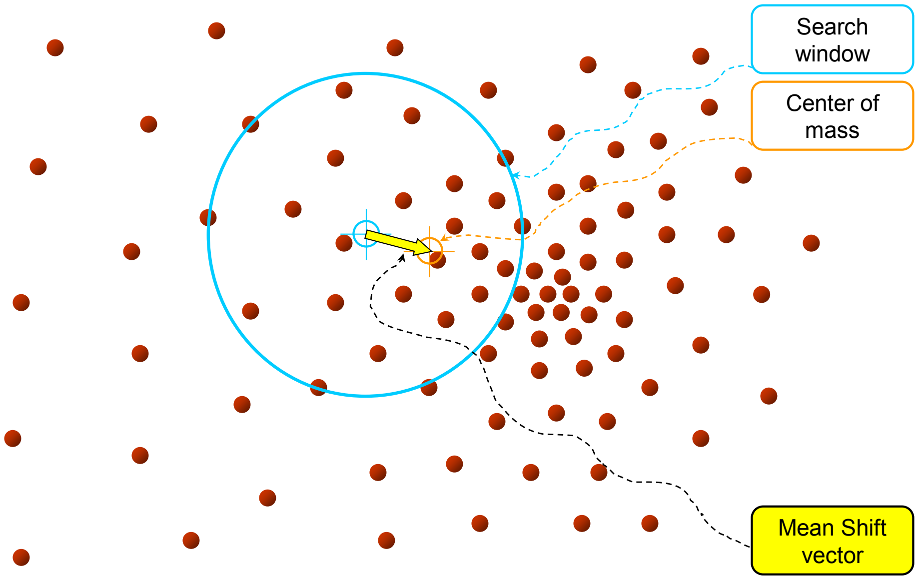
Mean shift algorithm

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



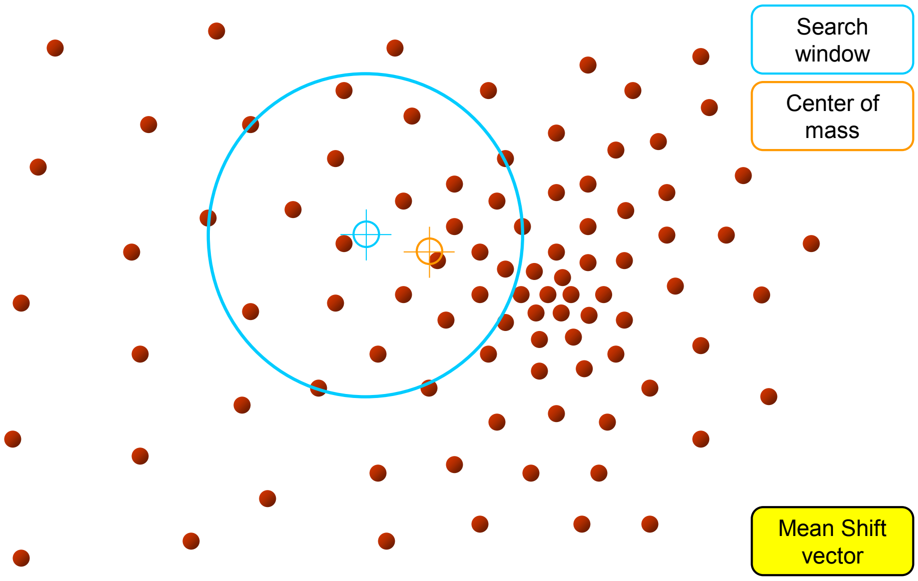
Source: S. Lazebnik

Mean shift



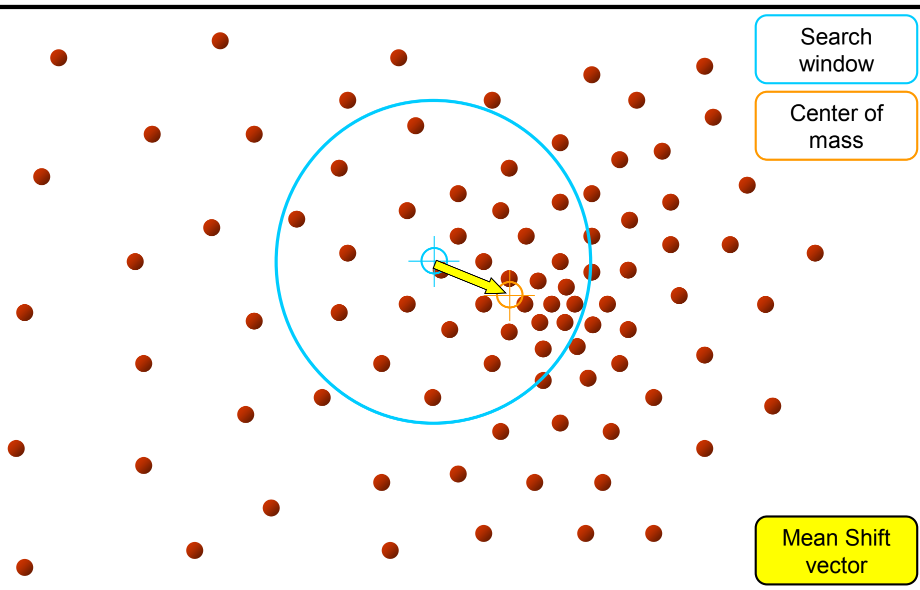
Soruce: Y. Ukrainitz & B. Sarel

Mean shift



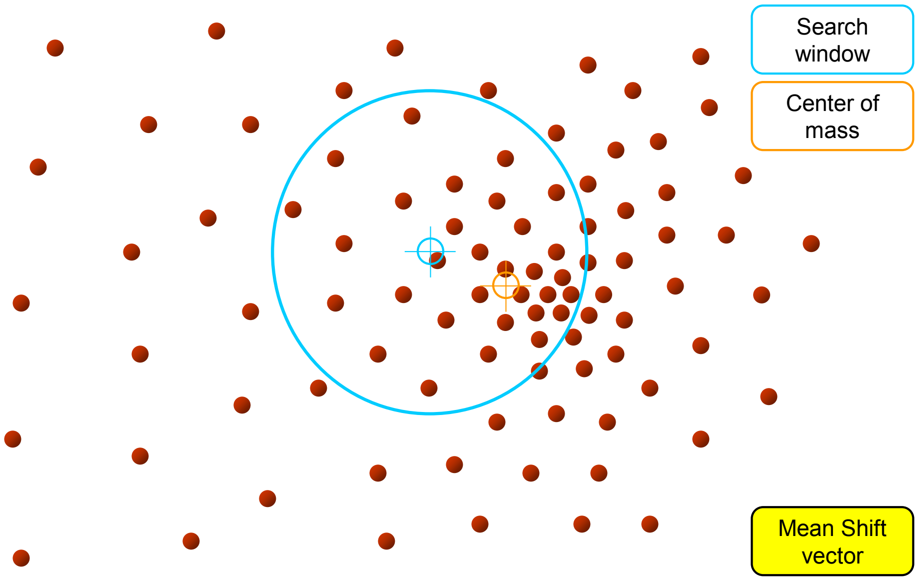
Soruce: Y. Ukrainitz & B. Sarel

Mean shift



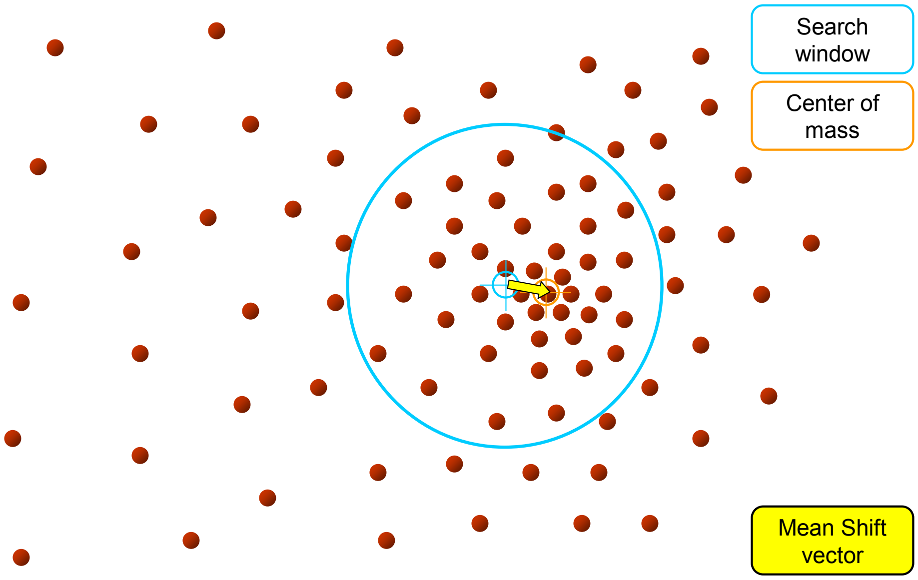
Soruce: Y. Ukrainitz & B. Sarel

Mean shift



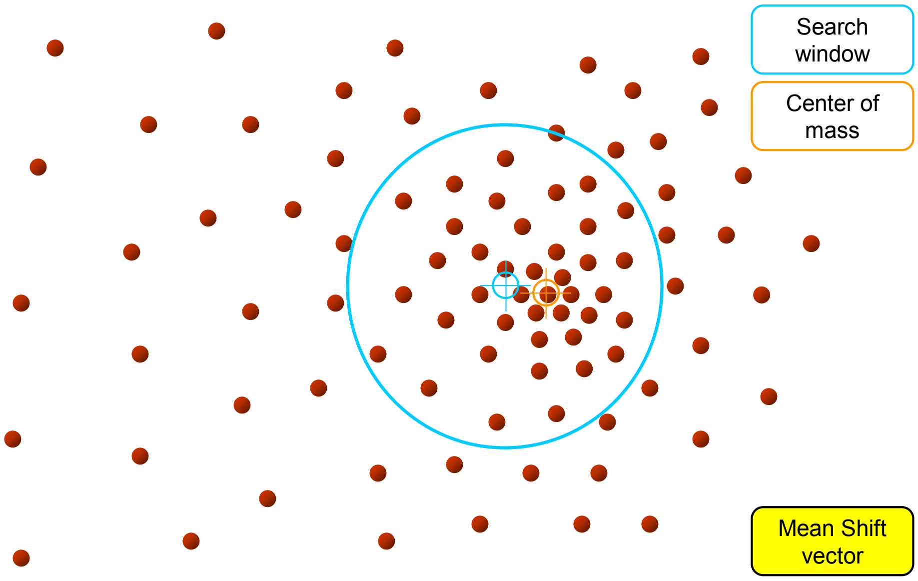
Soruce: Y. Ukrainitz & B. Sarel

Mean shift



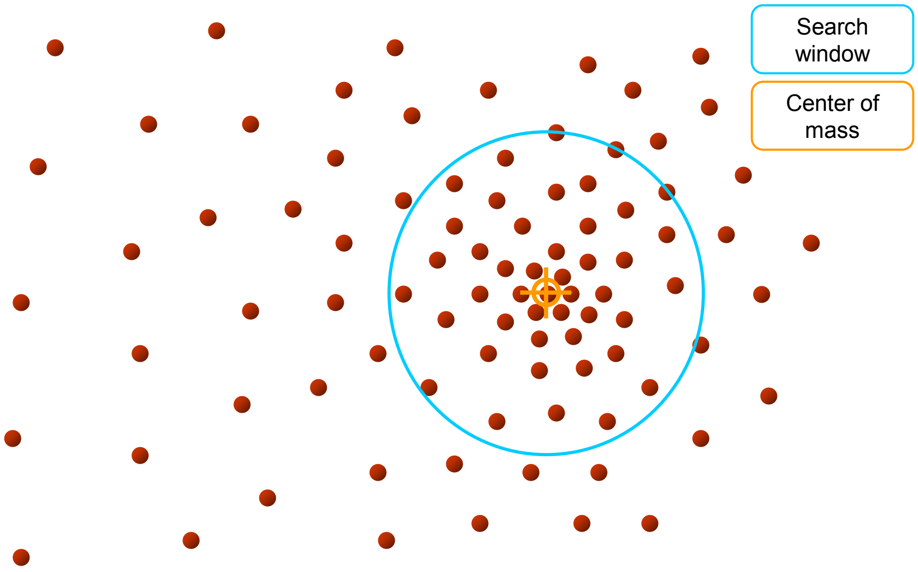
Soruce: Y. Ukrainitz & B. Sarel

Mean shift



Soruce: Y. Ukrainitz & B. Sarel

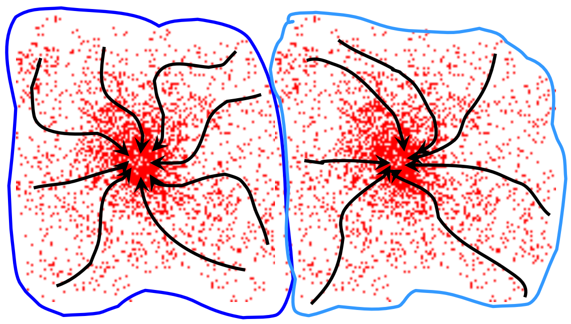
Mean shift



Soruce: Y. Ukrainitz & B. Sarel

Mean shift clustering

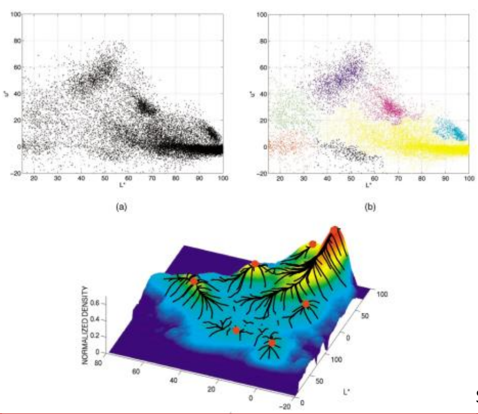
- Cluster: all data points in the attraction basin of a mode
- Attraction basin: the region for which all trajectories lead to the same mode



Soruce: Y. Ukrainitz & B. Sarel

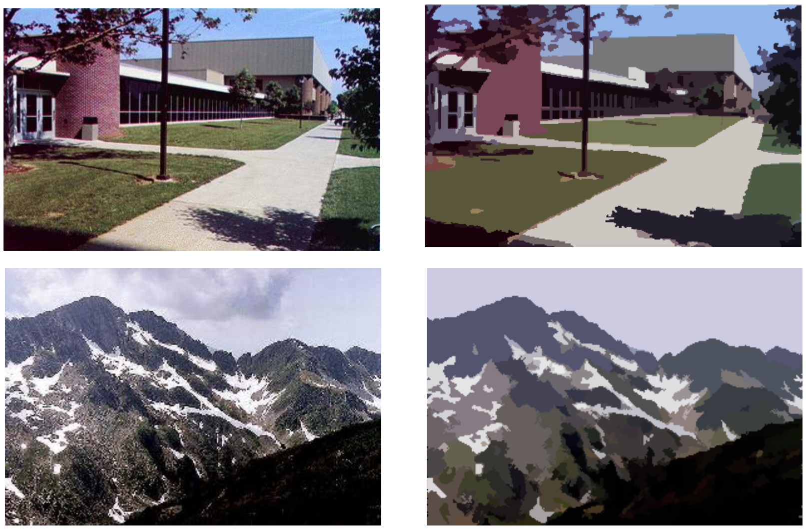
Mean shift clustering/segmentation

- 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



Source： S. Lazebnik

Mean shift segmentation results



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

Source： S. Lazebnik

More results

Source： S. Lazebnik

More results

Source： S. Lazebnik

Mean shift pros and cons

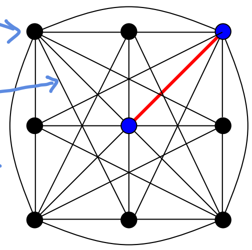
- Pros
  - Does not assume spherical clusters
  - Just a single parameter (window size)
  - Finds variable number of modes
  - Robust to outliers
- Cons
  - Output depends on window size
  - Computationally expensive
  - Does not scale well with dimension of feature space

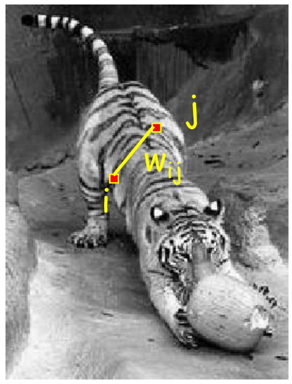
Source： S. Lazebnik

Images as graphs

pixel

边表示2个pixel间的相似度



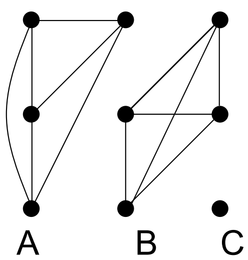


- Node for every pixel
- Edge between every pair of pixels (or every pair of “sufficiently close” pixels)
- Each edge is weighted by the *affinity* or similarity of the two nodes

Source: S. Seitz

Segmentation by graph partitioning

删除相似性最小的边，即完成了分割



- 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:

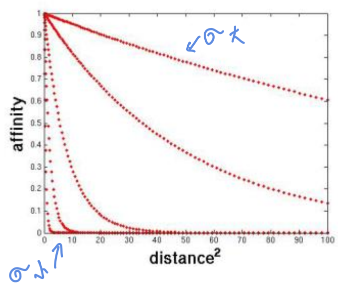
距离

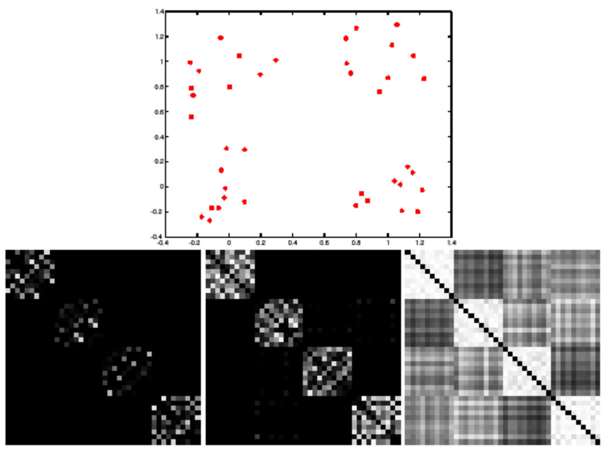
$$\text{similarity} = \exp\left(-\frac{1}{2\sigma^2}\,\mathrm{dist}(\mathbf{x}_i, \mathbf{x}_j)^2\right)$$

Source: S. Lazebnik

Scale affects affinity

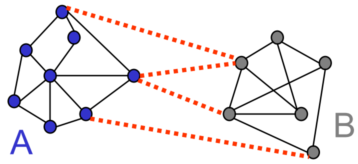
- Small σ: group only nearby points
- Large σ: group far-away points





Source: S. Lazebnik

Graph cut



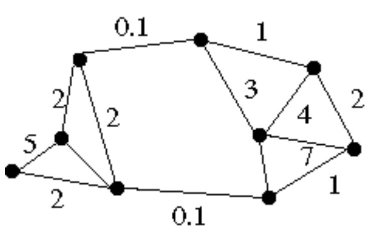
- Set of edges whose removal makes a graph disconnected
- Cost of a cut: sum of weights of cut edges
- A graph cut gives us a segmentation
  - What is a “good” graph cut and how do we find one?

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

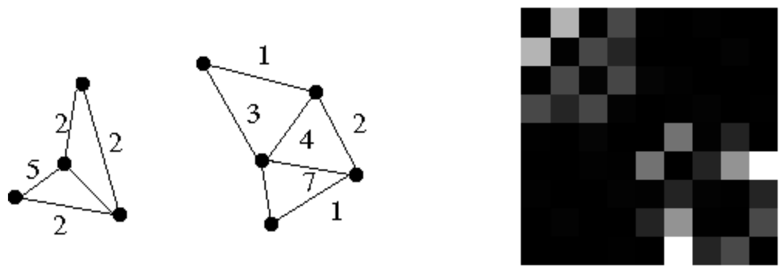


Source： S. Lazebnik

Minimum cut

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

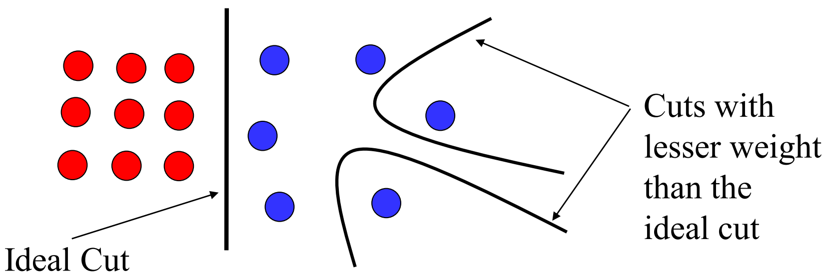
Minimum cut example



Source: S. Lazebnik

Normalized cut

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



单个pixel容易被切，因为它可能只有1个边，即weight自然很小

Source: Khurram Hassan-Shafique

Normalized cut 为了避免单个pixel被切

- 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: V表示所有类别，我们要使下面这个表达式最小，即实现了Normalized Cut

w(A,B)表示切完AB后，A与B的weight和，所以w(A,B)越小，A与B分割就越好

若A是单个pixel，w(A,B)小，w(A,V)会更小

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

所以 $\frac{w(A,B)}{w(A,V)}$ 会变大，$\frac{w(A,B)}{w(B,V)}$ 同理

$w(A, B)$ = sum of weights of all edges between A and B

所以当整个表达式达到最小时，不会出现切割单个pixel的情况，这就是 Normalized Cut

求解方法

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

Source：S. Lazebnik

W是相似度矩阵

$$\begin{array}{c|c} & 1\ 2\ \cdots\ 100 \\ \hline \begin{matrix}1\\2\\\vdots\\100\end{matrix} & \begin{pmatrix} 0 & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{pmatrix} \end{array}$$

自己跟自己计算填0

⇐ 则W是这样

这张图总共有100个pixels

W是对称矩阵

D只有对角线上有元素 分别是W每行的和

$$\begin{pmatrix} x & & & \\ & x & & \\ & & x & \\ & & & \ddots \\ & & & & x \end{pmatrix}$$

Normalized cut

- 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

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

⇐使它最小，求y

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

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

Source：S. Lazebnik

Normalized cut

- Finding the exact minimum of the normalized cut cost is NP-complete, but if we *relax* y to take on arbitrary values, then 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

Source: S. Lazebnik

Normalized cut algorithm

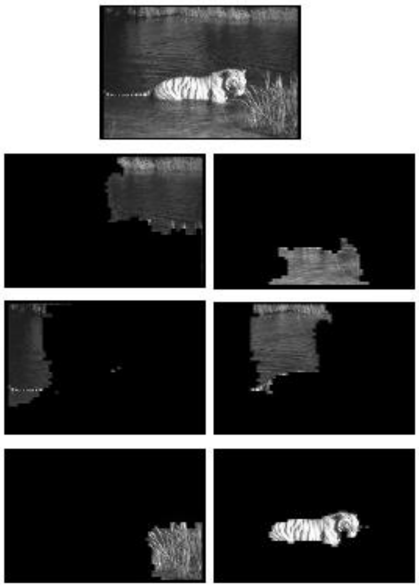
1. Represent the image as a weighted graph $G = (V,E)$, compute the weight of each edge, and summarize the information in D and W
2. Solve $(D - W)y = \lambda Dy$ for the eigenvector with the second smallest eigenvalue
3. Use the entries of the eigenvector to bipartition the graph

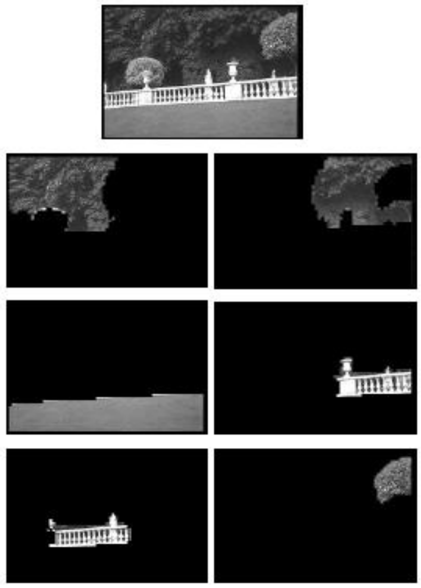
To find more than two clusters:

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

Source: S. Lazebnik

Example result





Source： S. Lazebnik

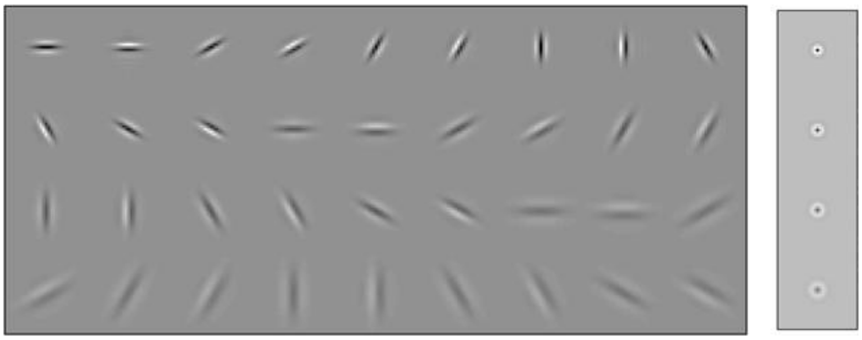
Challenge

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

Source： S. Lazebnik

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

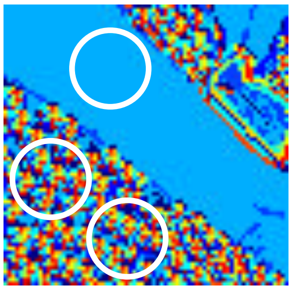
- Convolve image with a bank of filters
- Find *textons* by clustering vectors of filter bank outputs

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.

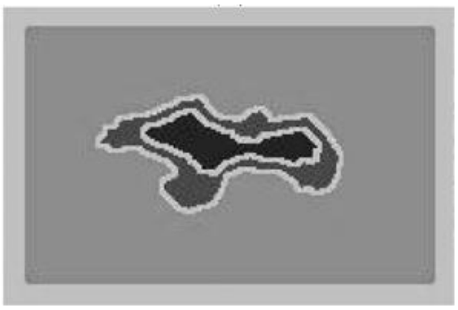
Using texture features for segmentation

- 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.

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

Source： S. Lazebnik

Results: Berkeley Segmentation Engine

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

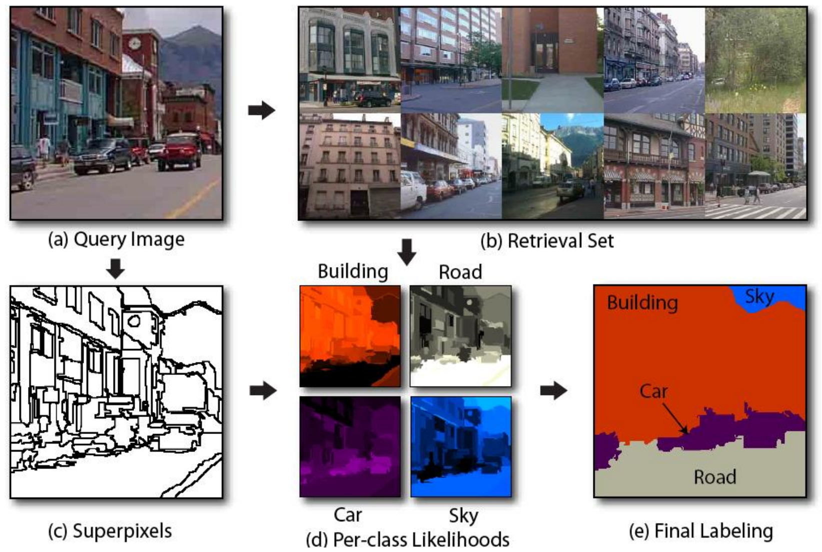
Source： S. Lazebnik

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

Source：S. Lazebnik

Segments as primitives for recognition



J. Tighe and S. Lazebnik, ECCV 2010