图像处理与分析

— 图像分割

授课教师：孙剑

jiansun@mail.xjtu.edu.cn

http://jiansun.gr.xjtu.edu.cn

西安交通大学 数学与统计学院

目录

- 基于聚类的图像分割算法
  - k-means算法
  - mean-shift算法
- 基于图的图像分割算法

图像分割问题

- 图像分割：将图像区域分割为颜色、纹理或语义一致的区域。
  - 底层分割：将图像分割为颜色或纹理一致的区域，但每个分割区域内没有确定的语义信息；
  - 中层分割：将图像分割为具有中层语义信息的图像块。例如前景／背景分割；
  - 语义分割：将图像逐点分割为不同语义信息的图像分割块。

分割和聚类

- **图像分割任务：**通过将图像或者视频分割为多个分割块，获得图像或视频的一个更为精炼的表达。

$$\Omega = \bigcup_i \Omega_i \qquad \Omega_i \cap \Omega_j = \varnothing,\ i \neq j$$

- **基于聚类的分割算法思想：**将相似的图像特征进行聚类，每一类即为图像的一个分割块。
  - 图像特征：颜色、纹理等
  - 如何判断特征的相似性？
  - 如何建模图像聚类问题？

分割和聚类

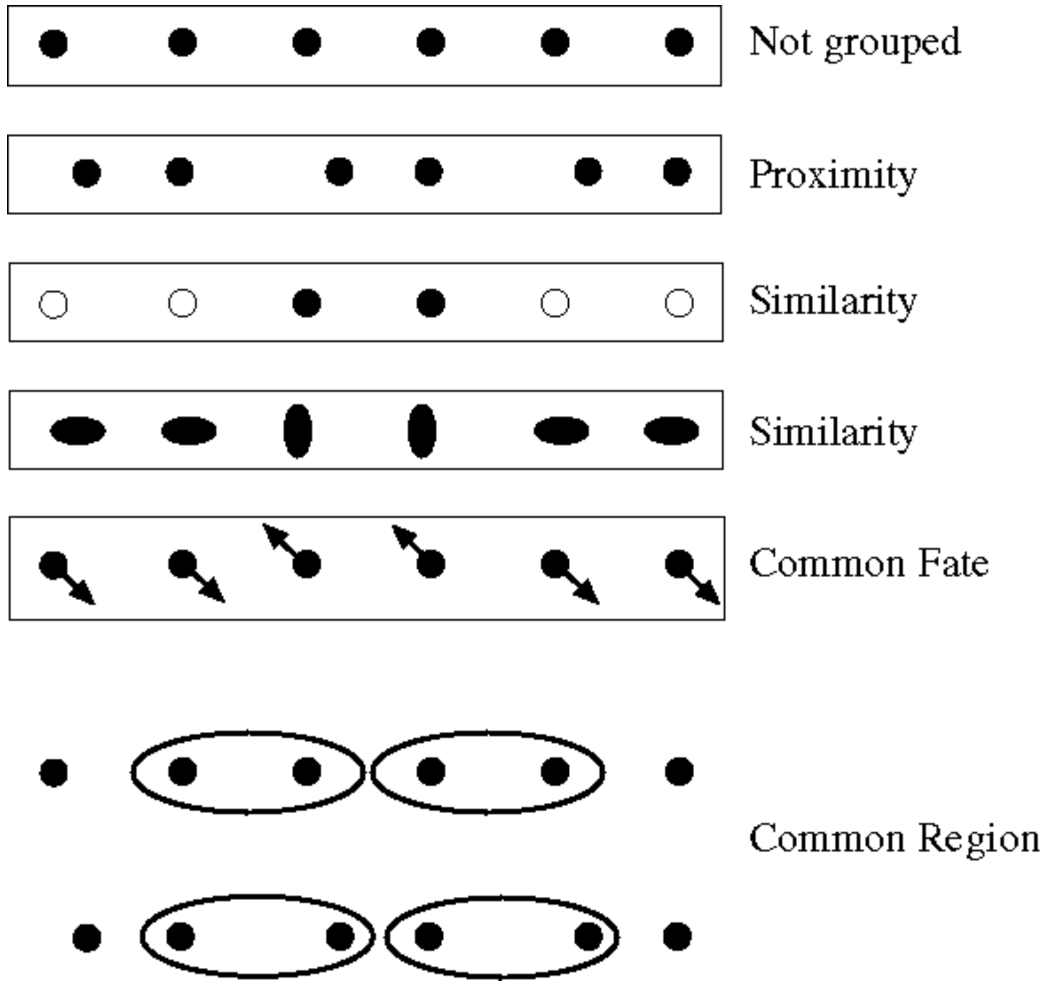
- 分割算法的基本思想
  - 自下而上的分割方法：基于特征的局部相似性进行像素点的聚类
  - 自上而下的分割方法：像素位于同一个物体之上，则认为像素属于一类。

自下而上的分割方法

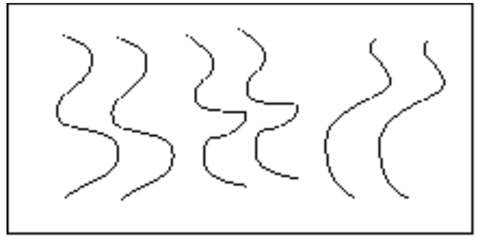
自上而下的分割方法

分割和聚类

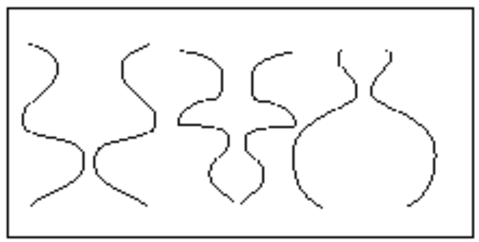
- 视觉聚类原理：Gestalt心理学原理，即描述人是如何将感知到的事物聚合到一起。



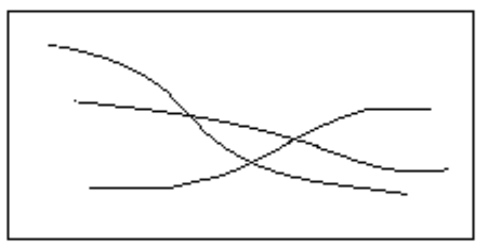
分割和聚类



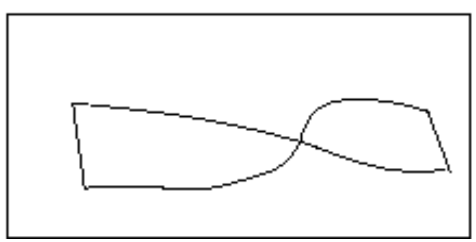
Parallelism



Symmetry



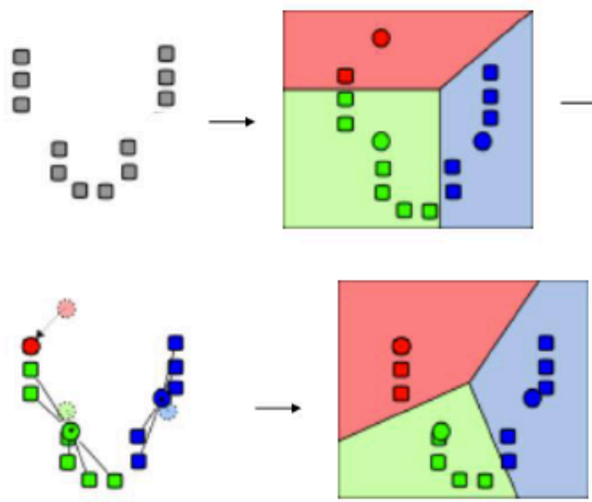
Continuity



Closure

基于k-means聚类的图像分割算法

- K-Means聚类算法：是典型的基于特征相似性的聚类算法



算法流程：

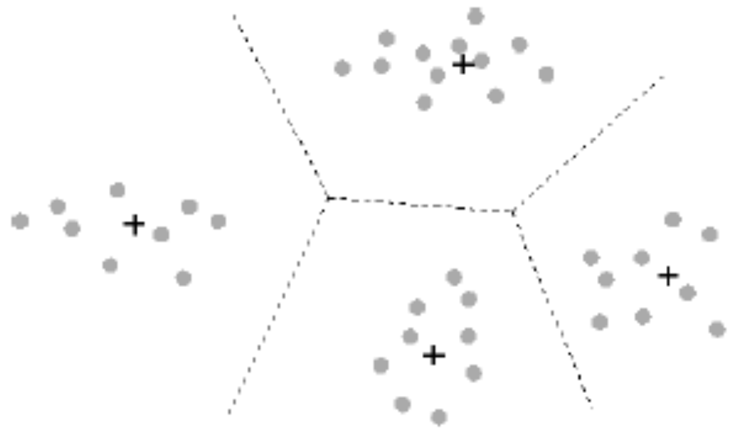
初始化K个中心点: $\{\mu_1,\cdots,\mu_K\}$

迭代进行如下步骤：

Step 1: 将每个点 x_i 赋给最近邻中心点所在的类，即：$l(x)=\arg\min_i \| x_j-\mu_i \|$ ；

Step 2: 将重新计算每个类别的中心点

$$\mu_c=\frac{1}{N}\sum_{l(x_j)=c} x_j,$$



Source:
http://www.heikohoffmann.de/htmlthesis/node28.html

- Solving the optimization problem

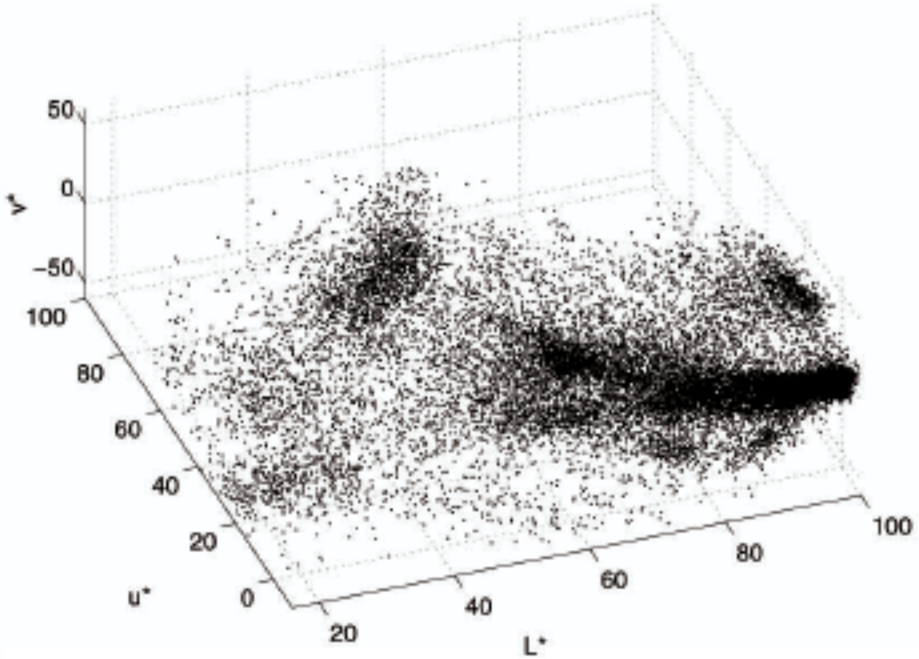
$$\arg\min_{\mathbf{S}} \underbrace{\sum_{i=1}^{k} \sum_{\mathbf{x}_j \in S_i} \|\mathbf{x}_j - \boldsymbol{\mu}_i\|^2}_{e(\boldsymbol{\mu}_i)}$$

- Every iteration is a step of gradient descent

$$\frac{\partial e}{\partial \boldsymbol{\mu}_i} = 0 \longrightarrow \boxed{\boldsymbol{\mu}_i^{t+1} = \frac{1}{|S_i^{(t)}|} \sum_{\mathbf{x}_j \in S_i^{(t)}} \mathbf{x}_j}$$

基于k-means聚类的图像分割算法

输入图像



颜色(LUV)特征

基于k-means聚类的图像分割算法

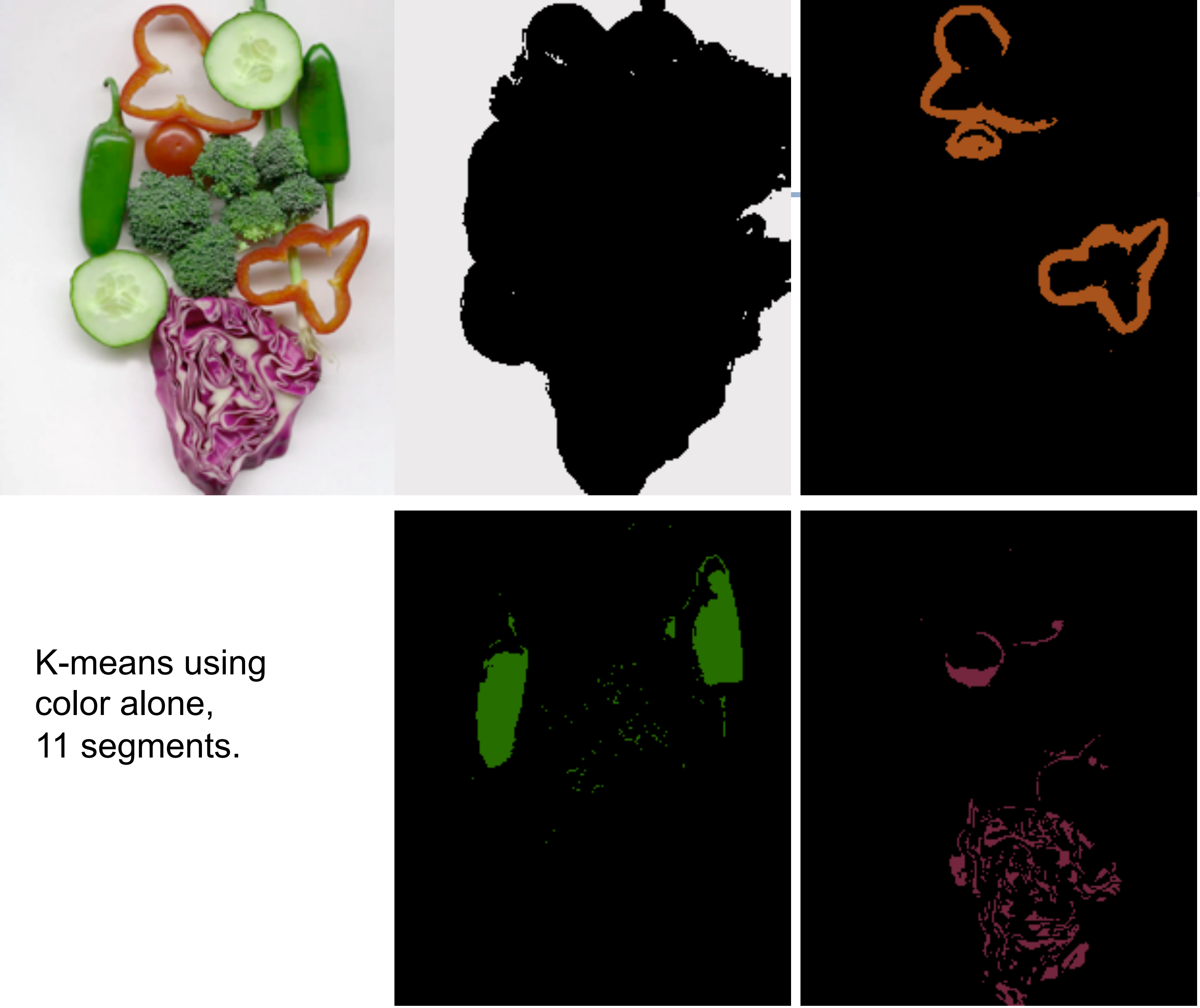
输入图像

灰度值聚类结果

颜色值聚类结果

K-means 聚类结果

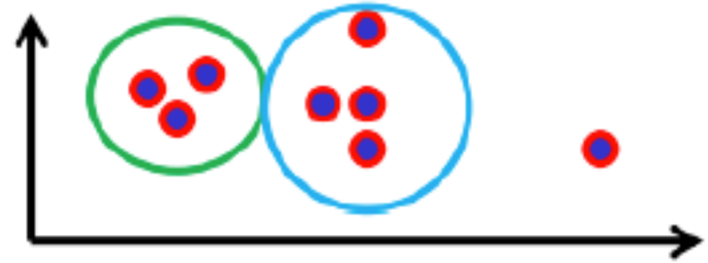
From: Computer Vision - A Modern Approach. Slides by D.A. Forsyth

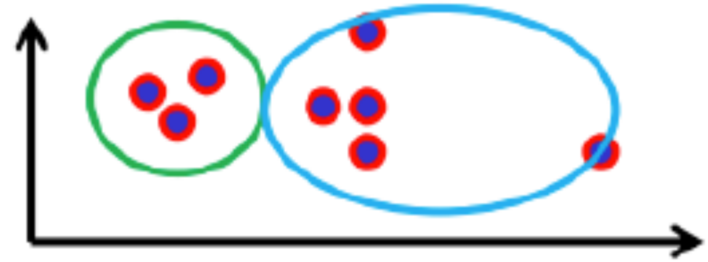


K-means using
color alone,
11 segments.

基于k-means聚类的图像分割算法

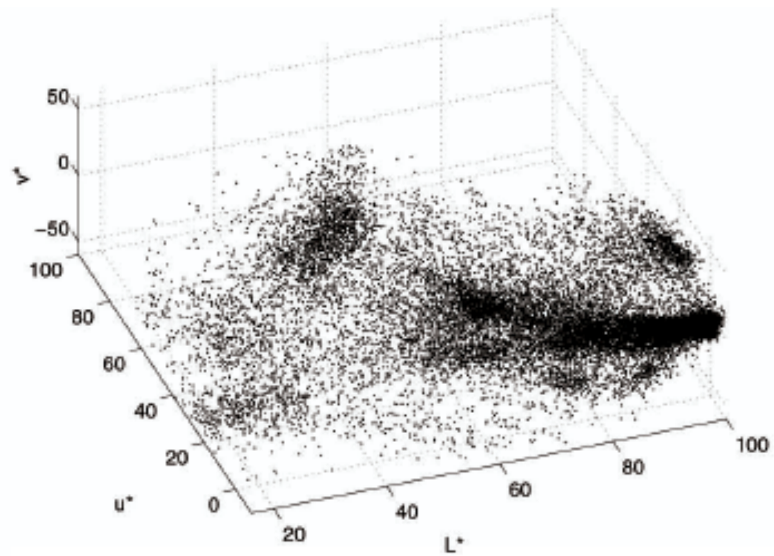
- K-Means聚类算法的优点
  - 简单、快速
  - 收敛到目标函数的局部极小值点
- K-means聚类算法的缺点
  - 需要选择合适的K
  - 对初始化敏感
  - 对数据点分布要求：对球形分布数据聚类结果好
  - 对误差敏感





基于mean-sift聚类的图像分割算法

Mean-sift聚类基本思想：在特征空间中找到特征密度分布的局部极大值点。



IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 24, NO. 5, MAY 2002

［参考文献］：

Mean Shift: A Robust Approach Toward Feature Space Analysis

Dorin Comaniciu, *Member, IEEE*, and Peter Meer, *Senior Member, IEEE*

基于mean-shift聚类的图像分割算法

特征空间的密度估计（核密度估计算法）：

$$\hat{f}_{h,K}(\mathbf{x}) = \frac{c_{k,d}}{nh^d} \sum_{i=1}^{n} k\left(\left\|\frac{\mathbf{x}-\mathbf{x}_i}{h}\right\|^2\right).$$

Mean-shift算法的聚类目标：寻找密度的局部极值点$\nabla f(\mathbf{x}) = \mathbf{0}$

$$\begin{aligned}
&\hat{\nabla} f_{h,K}(\mathbf{x}) \\
&= \frac{2c_{k,d}}{nh^{d+2}} \sum_{i=1}^{n} (\mathbf{x}_i - \mathbf{x}) g\left(\left\|\frac{\mathbf{x}-\mathbf{x}_i}{h}\right\|^2\right) \\
&= \frac{2c_{k,d}}{nh^{d+2}} \left[\sum_{i=1}^{n} g\left(\left\|\frac{\mathbf{x}-\mathbf{x}_i}{h}\right\|^2\right)\right] \left[\frac{\sum_{i=1}^{n} \mathbf{x}_i g\left(\left\|\frac{\mathbf{x}-\mathbf{x}_i}{h}\right\|^2\right)}{\sum_{i=1}^{n} g\left(\left\|\frac{\mathbf{x}-\mathbf{x}_i}{h}\right\|^2\right)} - \mathbf{x}\right],
\end{aligned}$$

Mean-shift

$$\mathbf{m}_{h,G}(\mathbf{x}) = \frac{\sum_{i=1}^{n} \mathbf{x}_i g\left(\left\|\frac{\mathbf{x}-\mathbf{x}_i}{h}\right\|^2\right)}{\sum_{i=1}^{n} g\left(\left\|\frac{\mathbf{x}-\mathbf{x}_i}{h}\right\|^2\right)} - \mathbf{x},$$

基于mean-shift聚类的图像分割算法

Mean-Shift迭代公式：

$$\mathbf{y}_{j+1} = \frac{\sum_{i=1}^{n} \mathbf{x}_i g\left(\left\|\frac{\mathbf{x}-\mathbf{x}_i}{h}\right\|^2\right)}{\sum_{i=1}^{n} g\left(\left\|\frac{\mathbf{x}-\mathbf{x}_i}{h}\right\|^2\right)} \qquad j = 1, 2, \ldots$$

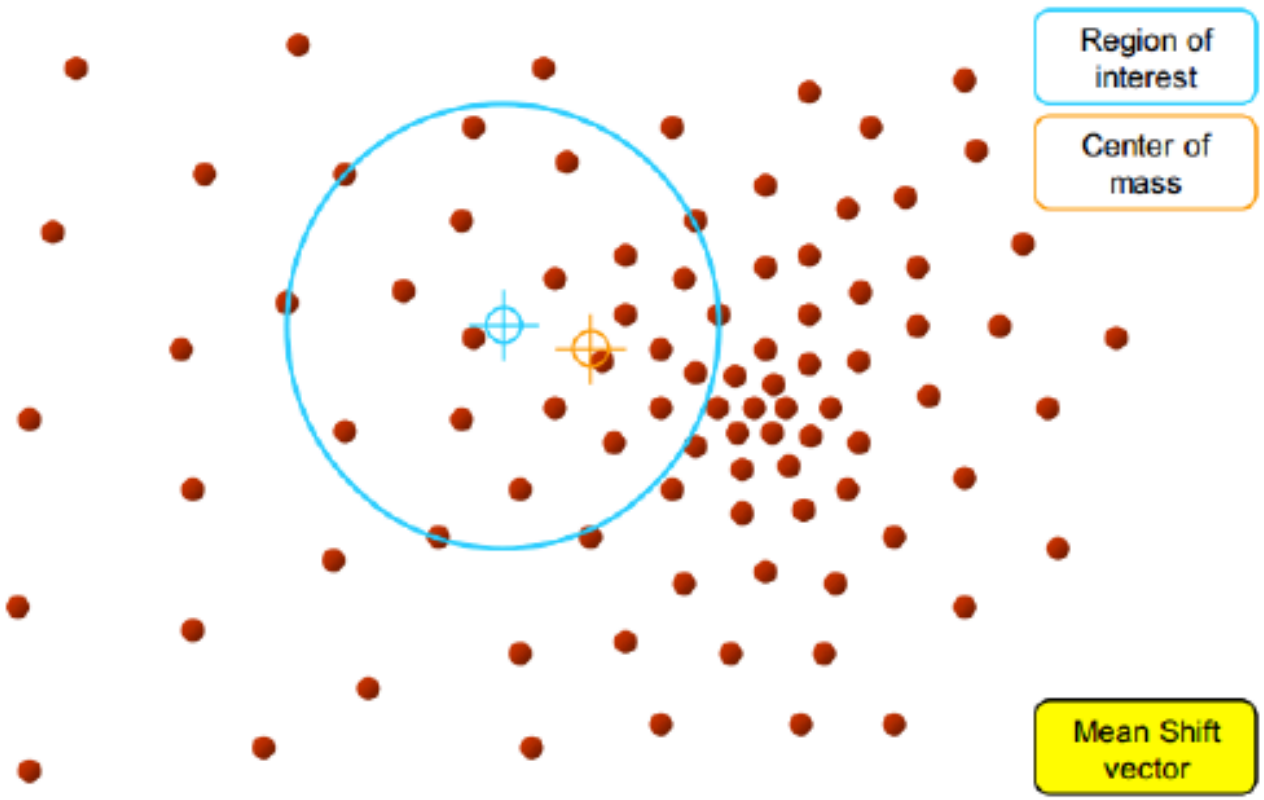
$$x = y_i$$

Theorem 1. *If the kernel K has a convex and monotonically decreasing profile, the sequences $\{\mathbf{y}_j\}_{j=1,2...}$ and $\{\hat{f}_{h,K}(j)\}_{j=1,2...}$ converge and $\{\hat{f}_{h,K}(j)\}_{j=1,2...}$ is monotonically increasing.*

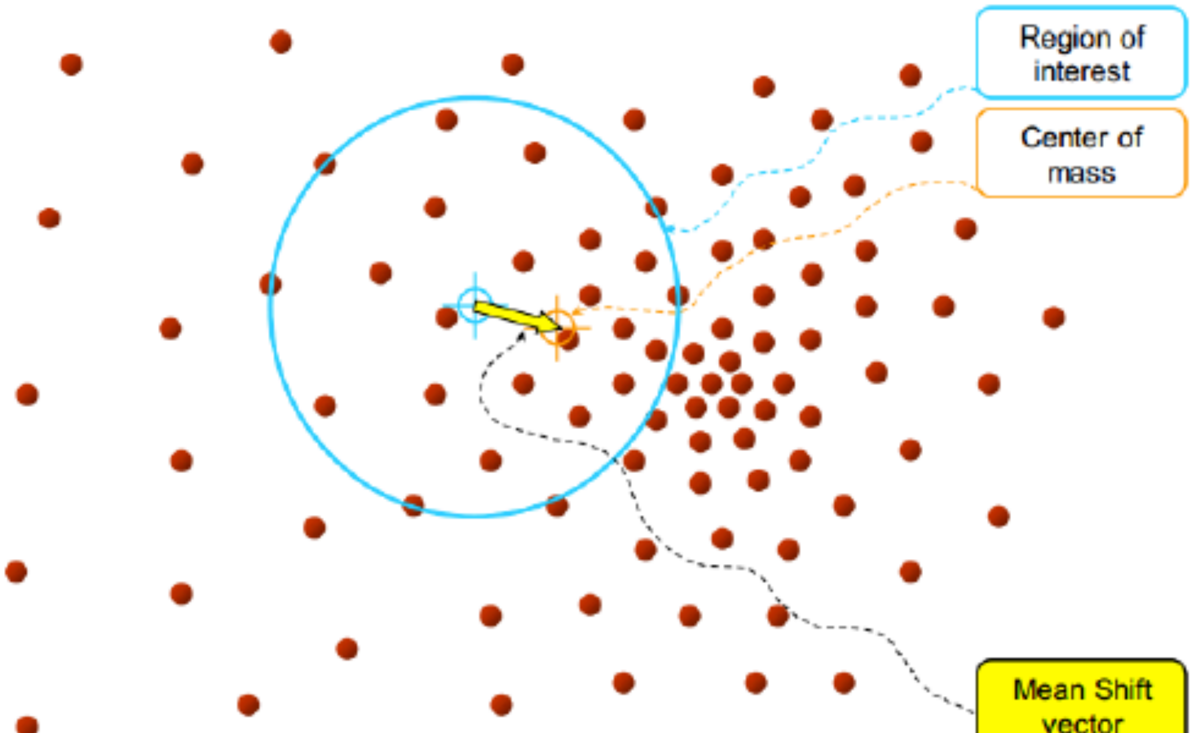
Theorem 2. *The cosine of the angle between two consecutive mean shift vectors is strictly positive when a normal kernel is employed, i.e.,*

$$\frac{\mathbf{m}_{h,N}(\mathbf{y}_j)^\top \mathbf{m}_{h,N}(\mathbf{y}_{j+1})}{\|\mathbf{m}_{h,N}(\mathbf{y}_j)\|\|\mathbf{m}_{h,N}(\mathbf{y}_{j+1})\|} > 0. \qquad (25)$$

Mean-shift

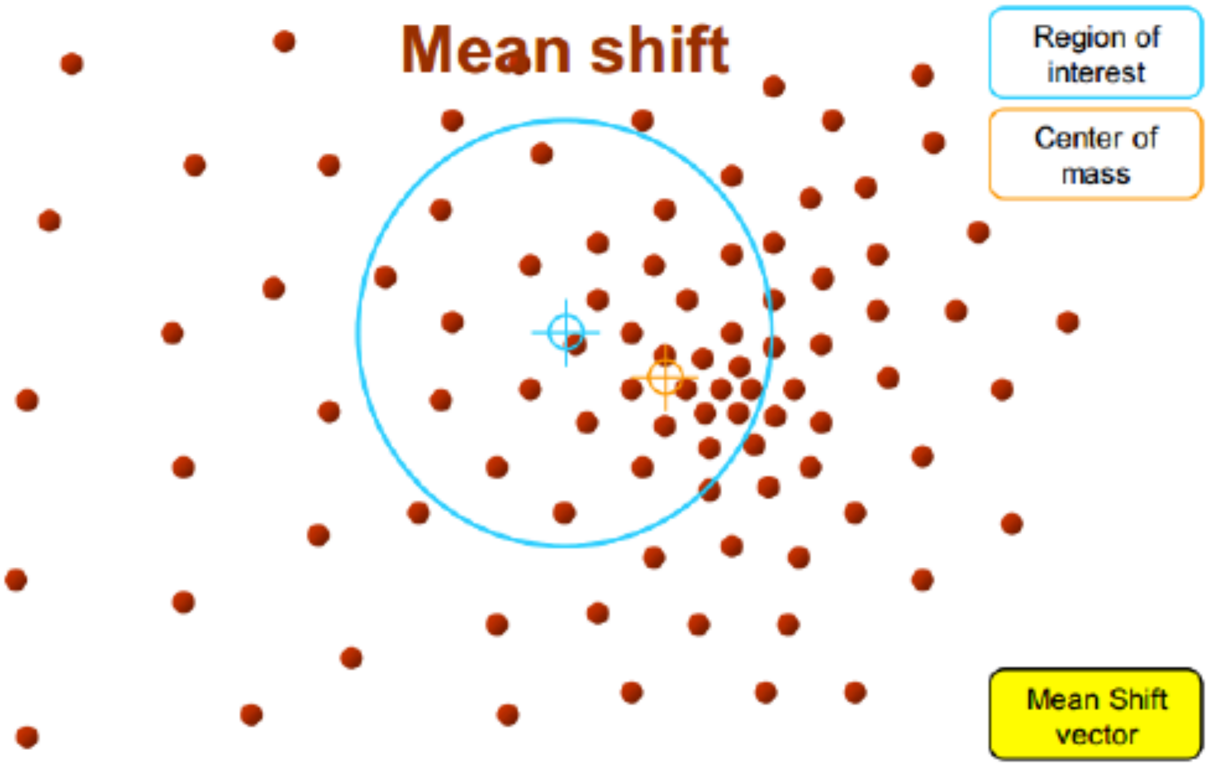


Mean-shift



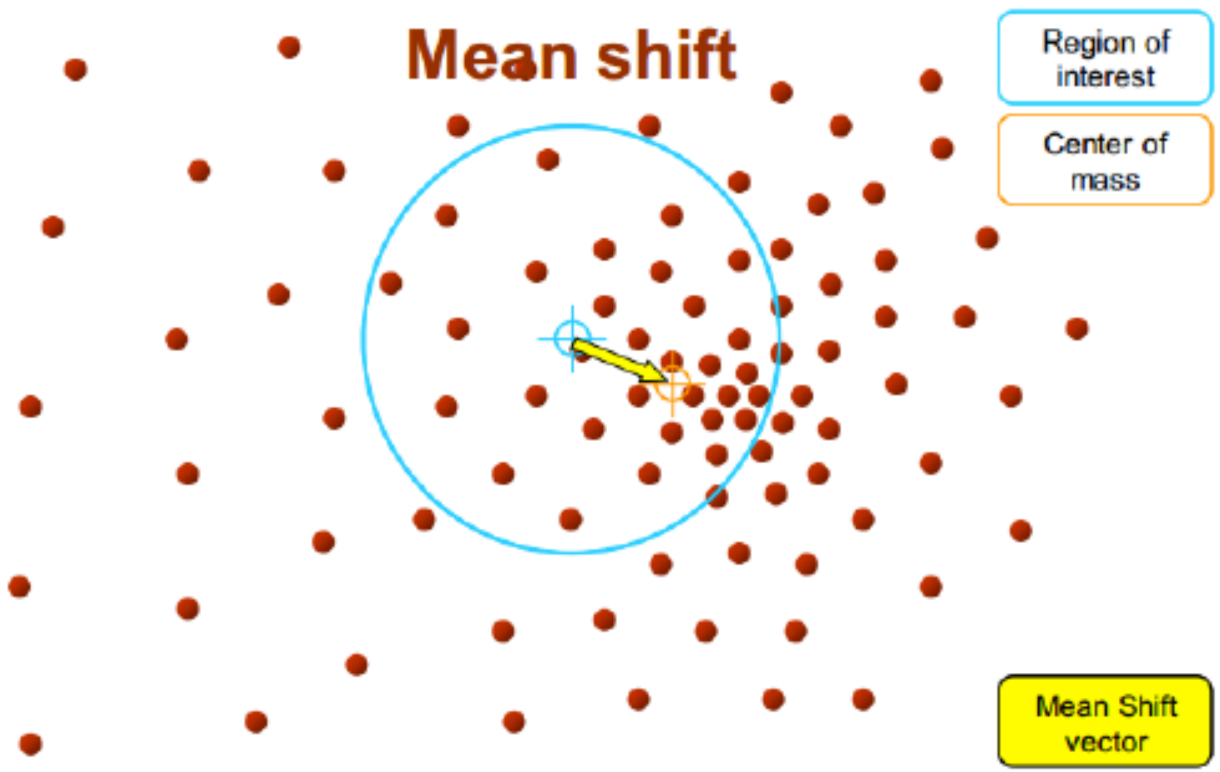
Slide by Y. Ukrainitz & B. Sarel

Mean-shift



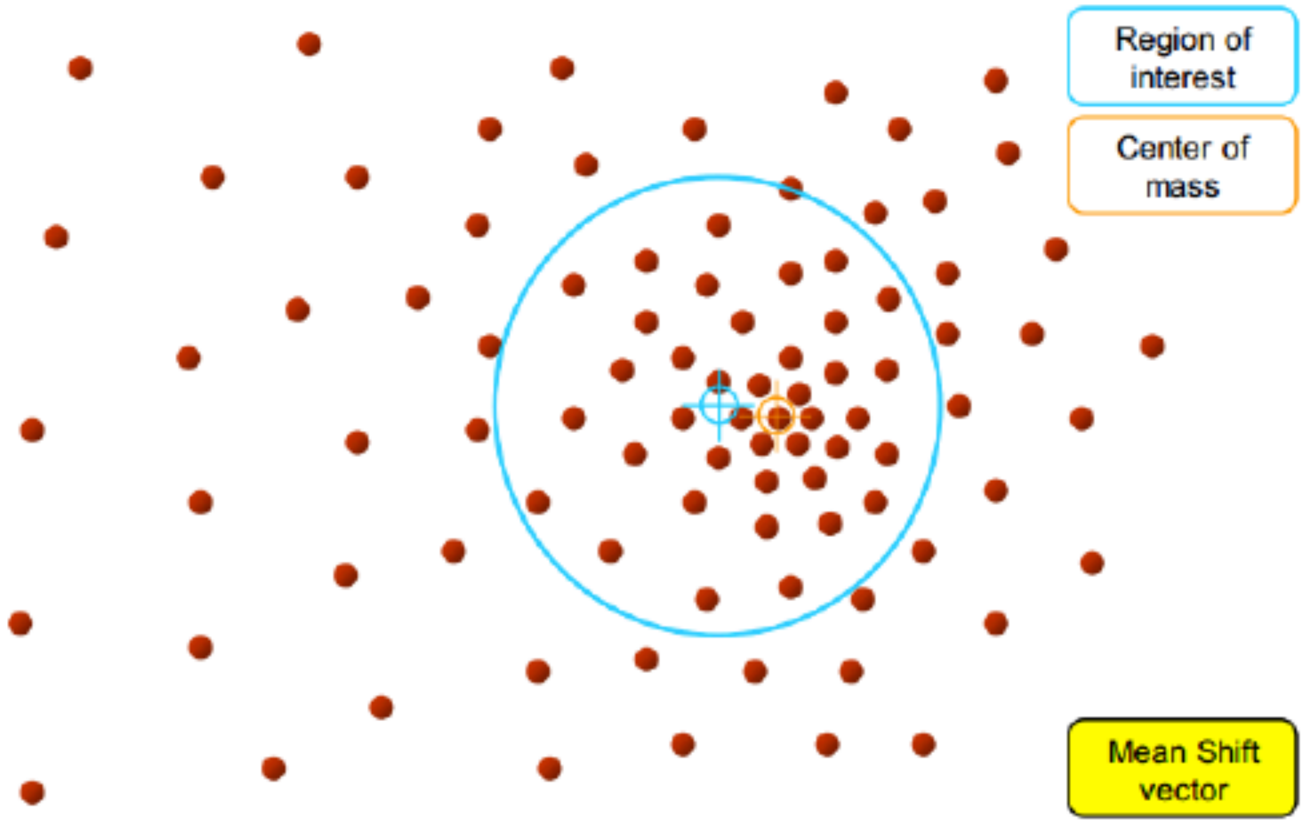
Slide by Y. Ukrainitz & B. Sarel

Mean-shift

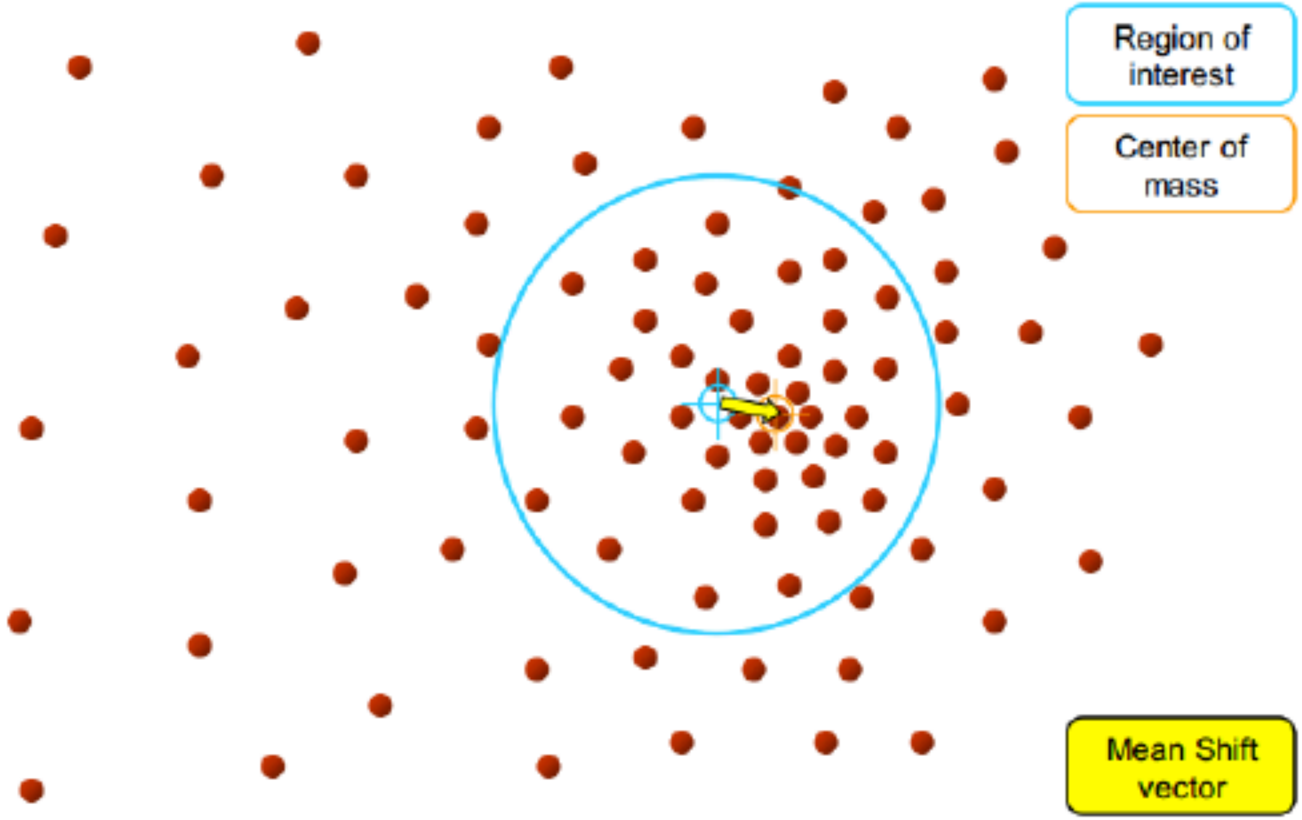


Slide by Y. Ukrainitz & B. Sarel

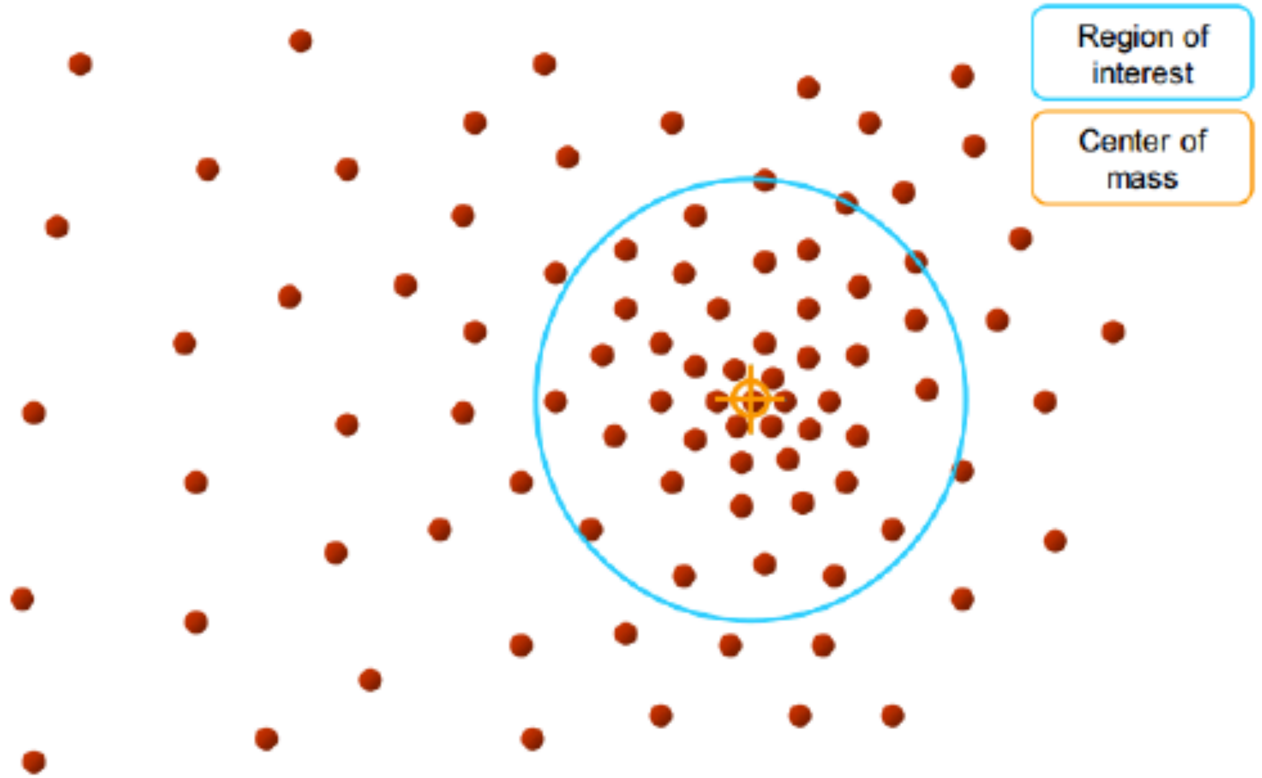
Mean-shift



Mean-shift



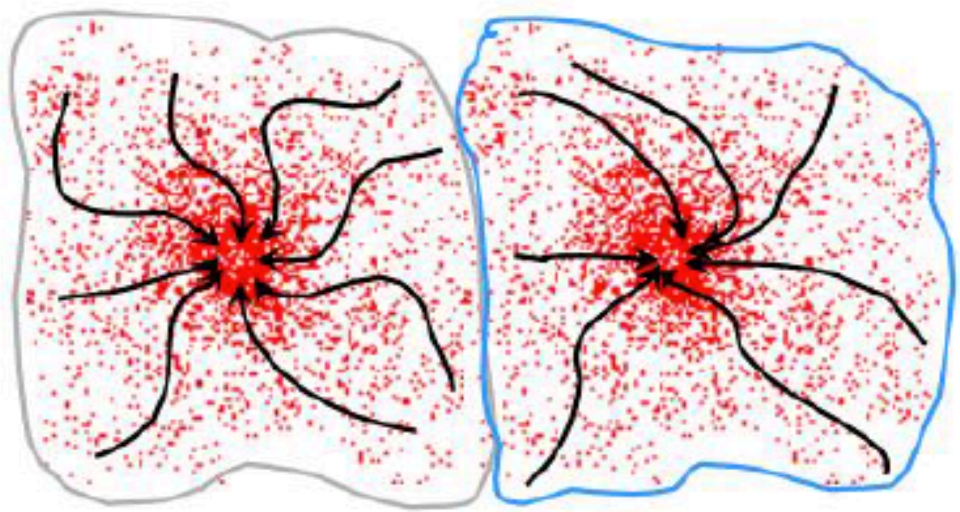
Mean-shift



Slide by Y. Ukrainitz & B. Sarel

Mean-shift

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

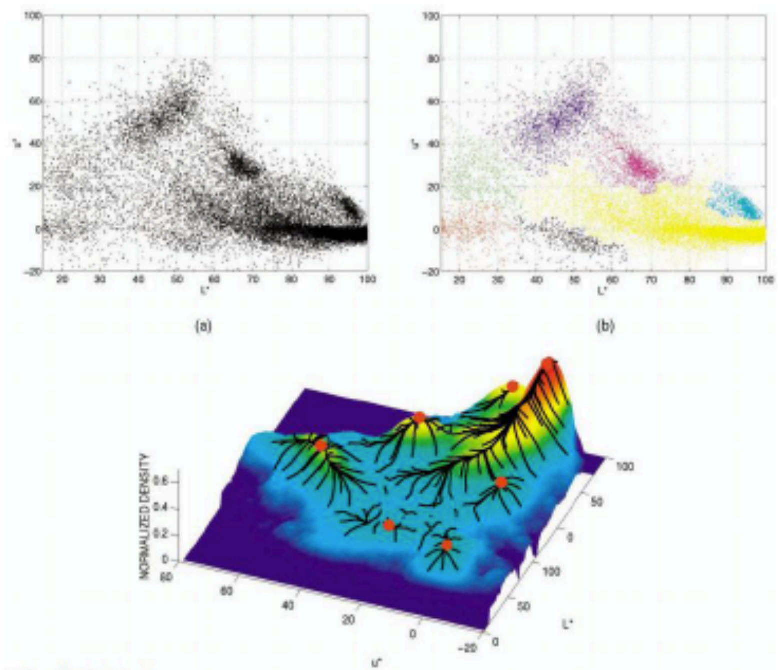


Slide by Y. Ukrainitz & B. Sarel

基于mean-shift聚类的图像分割算法

Mean-Shift算法流程：

- 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



Slide by Y. Ukrainitz & B. Sarel

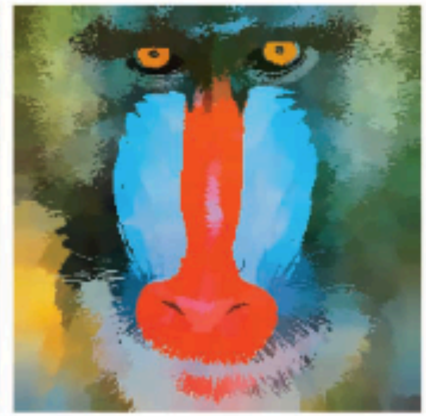
Original

$(h_s, h_r) = (8, 8)$

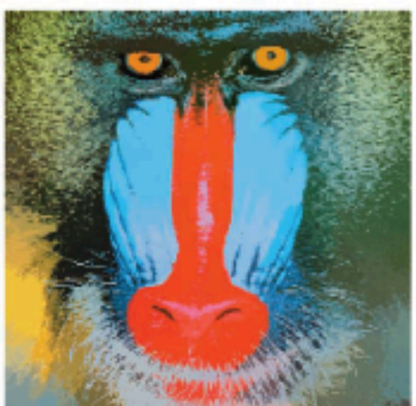
$(h_s, h_r) = (8, 16)$

$(h_s, h_r) = (16, 4)$

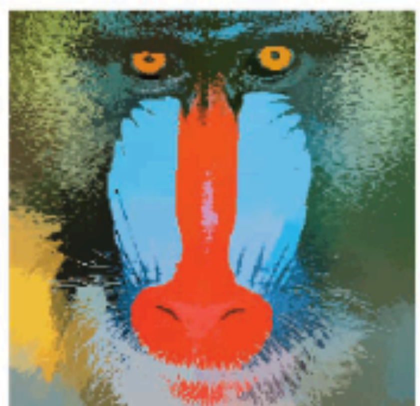
$(h_s, h_r) = (16, 8)$



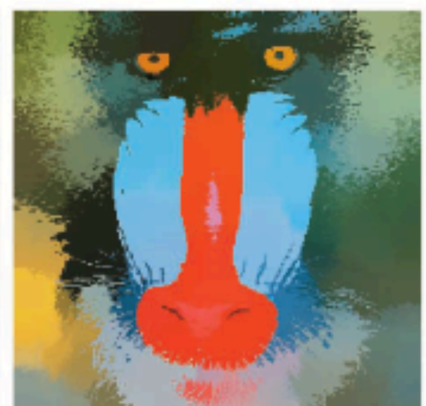
$(h_s, h_r) = (16, 16)$



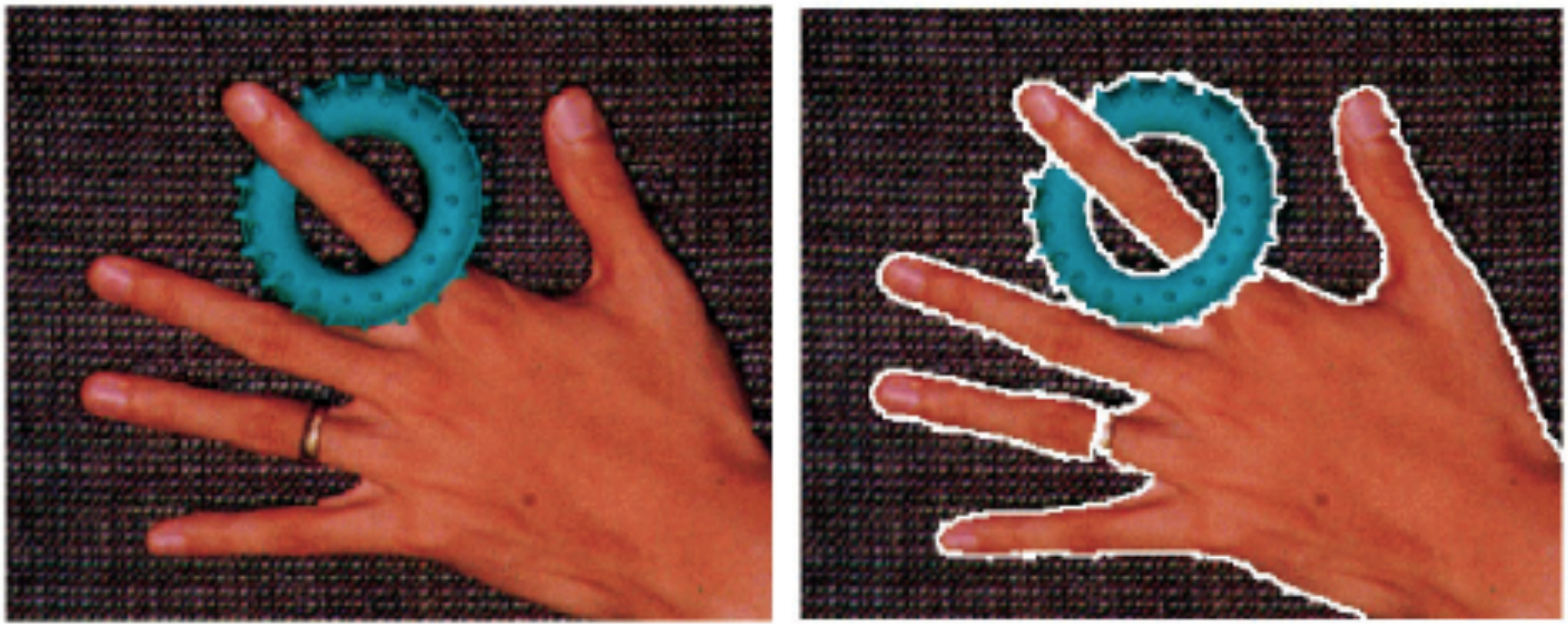
$(h_s, h_r) = (32, 4)$



$(h_s, h_r) = (32, 8)$



$(h_s, h_r) = (32, 16)$



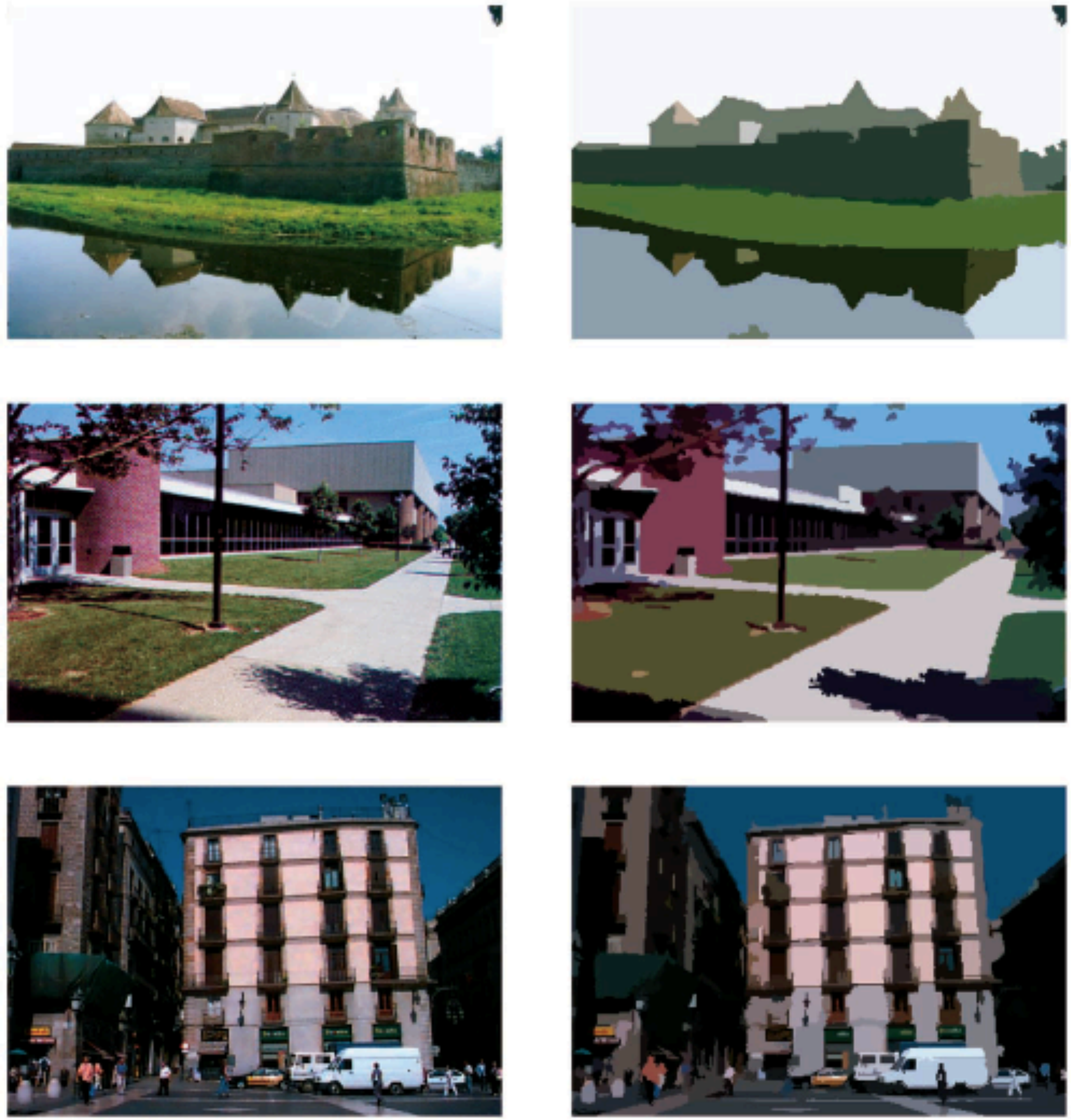
(a) (b)

Mean-Shift算法程序：

http://coewww.rutgers.edu/riul/research/code.html

Edge Detection and Image Segmentation System

Mean-Shift 滤波：



基于mean-shift聚类的图像分割算法

Mean-Shift算法优点：

- 可以聚类非球形分布的数据类
- 仅依赖于一个参数（窗口参数）
- 可找到多个局部极大值
- 对异常点具有鲁棒性

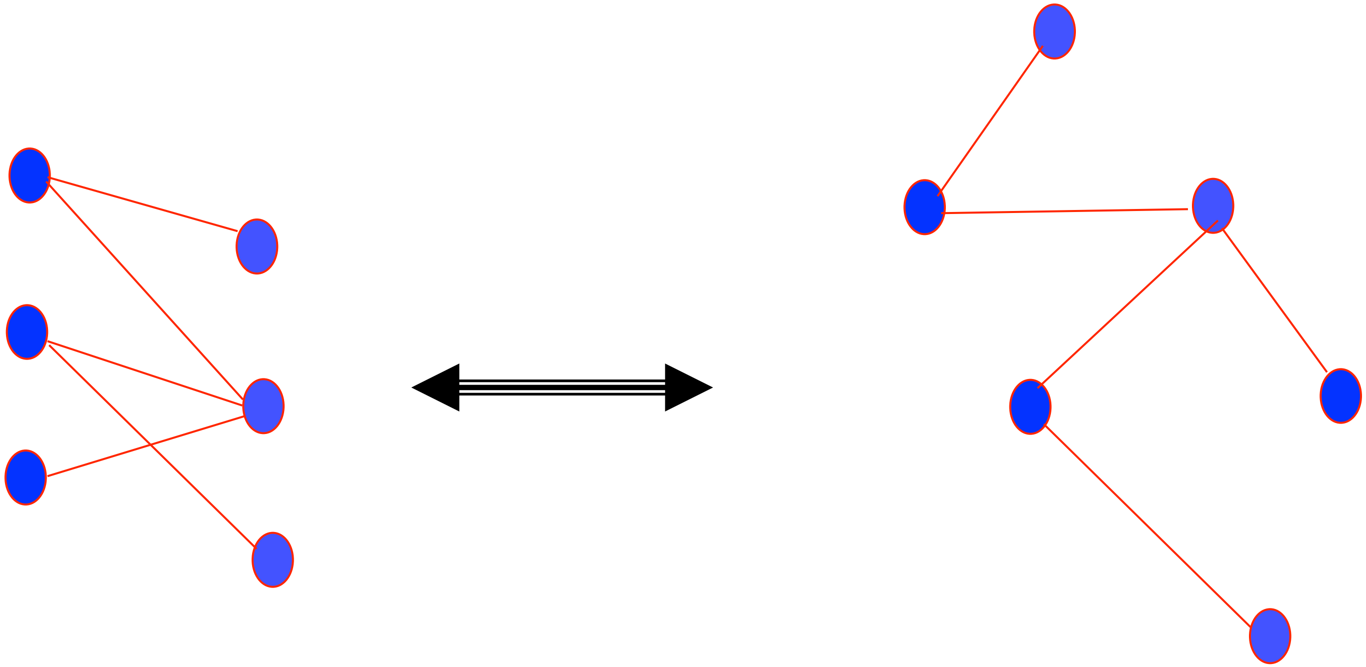
Mean-Shift算法缺点：

- 输出的聚类结果依赖于窗口函数
- 计算量很高
- 对高维度数据的可扩展性不足

基于mean-shift聚类的图像分割算法

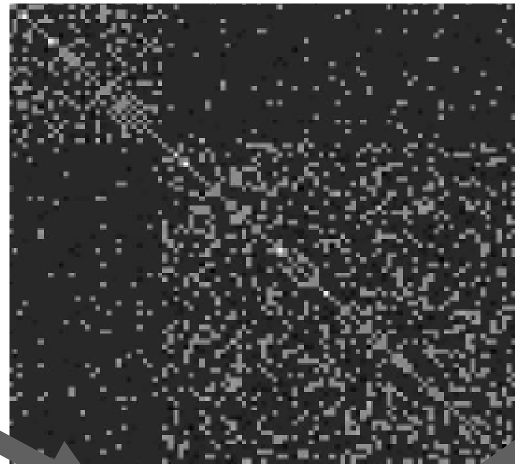
- 图(Graph)

G(V,E)：图G是由节点集合V(G)，和边集合 E(G) 构成，图上的边建模了节点之间的连接关系。



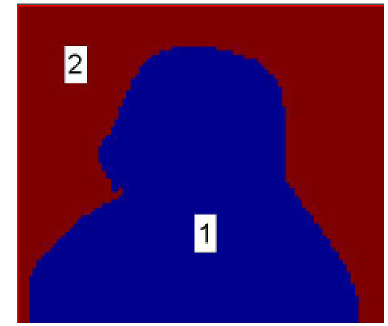
基于图的图像分割算法

Image (I)



Graph Affinities (W)

Eigenvector X(W)



Discretization

Intensity
Color
Edges
Texture

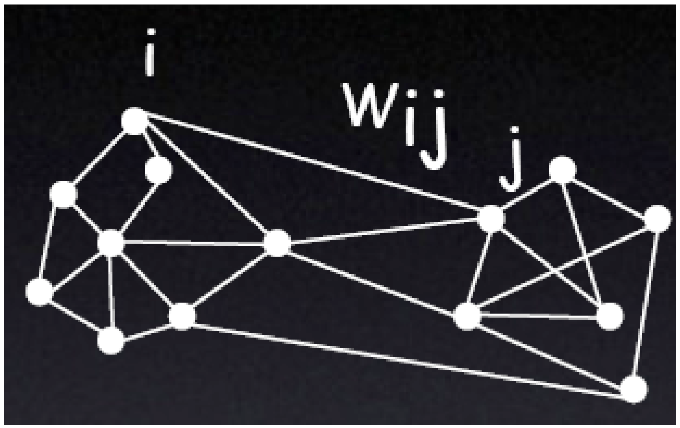
$$Ncut(A,B) = cut(A,B)\left(\frac{1}{vol(A)} + \frac{1}{vol(B)}\right)$$

$$(D-W)X = \lambda DX$$

$$X_A(i) = \begin{cases} 1 & if \quad i \in A \\ 0 & if \quad i \notin A \end{cases}$$

Slide from Timothee Cour (http://www.seas.upenn.edu/~timothee)

基于图的图像分割算法—图的定义



G = {V,E}

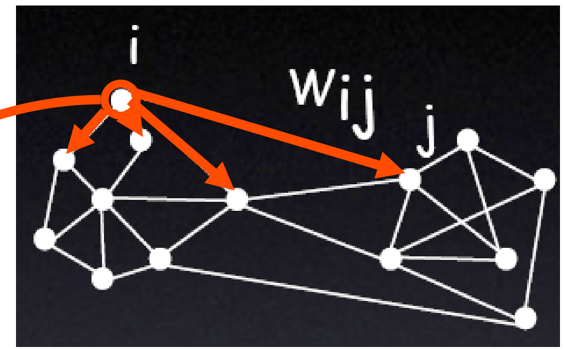
V: 图节点
E: 节点之间的连接边

V: 像素
E: 像素之间的相似形

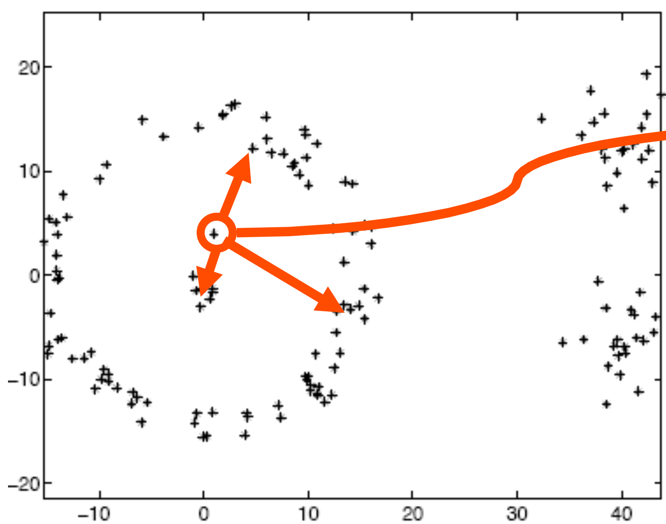
Slides from Jianbo Shi

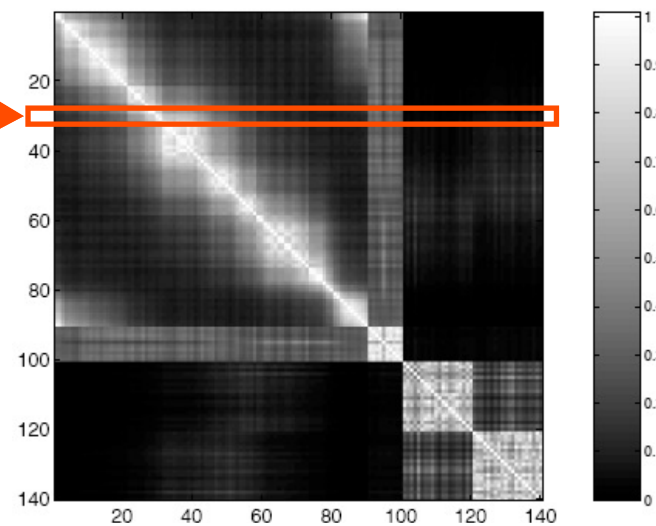
基于图的图像分割算法—图的定义

- 相似性矩阵: $W = \left[w_{i,j} \right]$



$$w_{i,j} = e^{\frac{-\left\| X_{(i)} - X_{(j)} \right\|_2^2}{\sigma_X^2}}$$



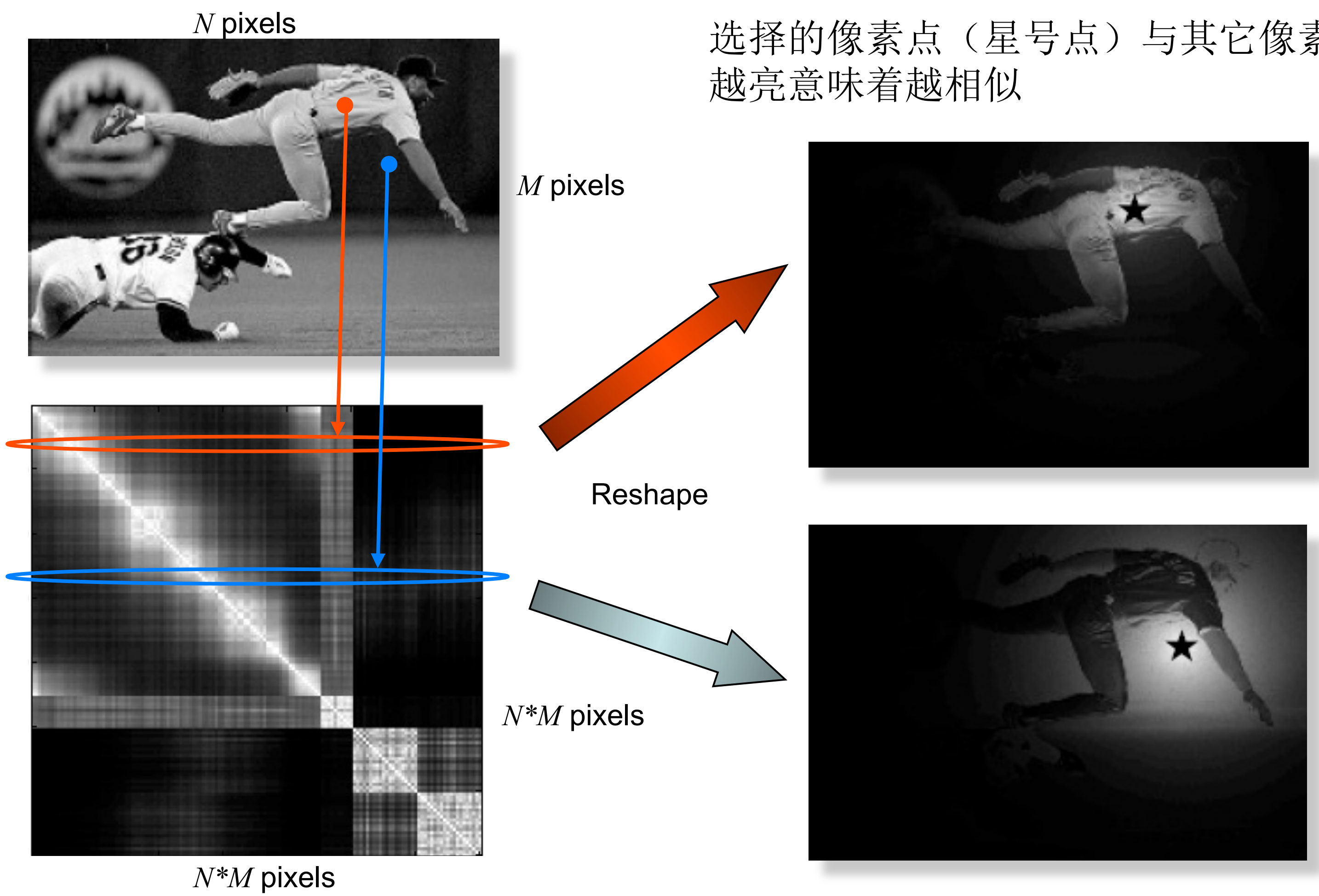


Slides from Jianbo Shi

基于图的图像分割算法—图的定义

- 相似性矩阵:

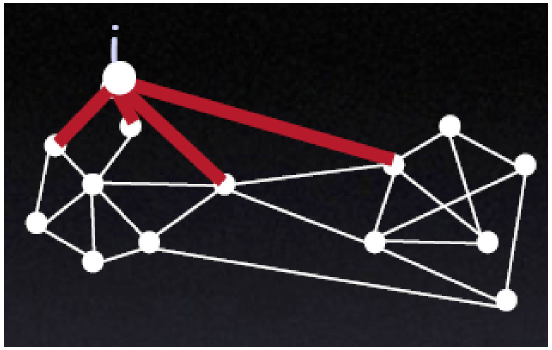
选择的像素点（星号点）与其它像素点的相似性越亮意味着越相似

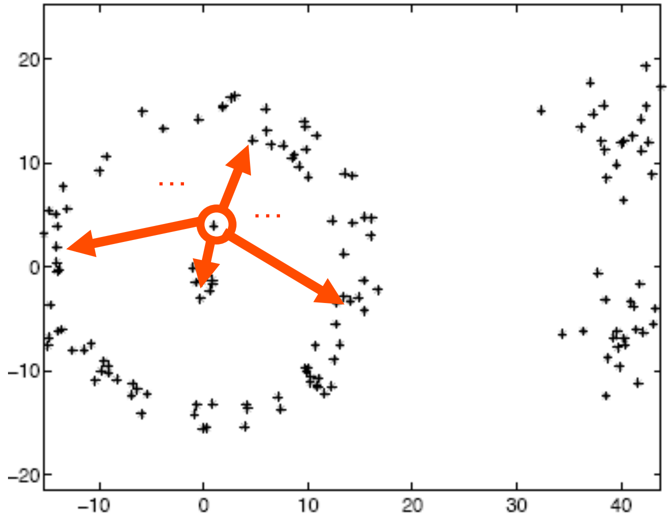


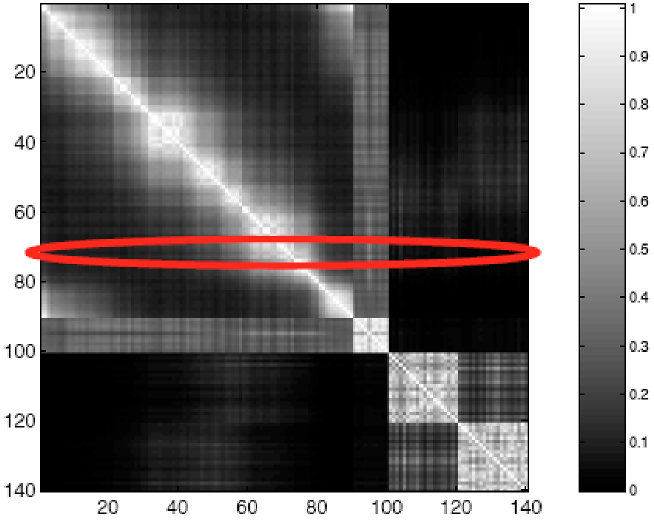
基于图的图像分割算法—图的定义

- 图节点的度数:

$$d_i = \sum_j w_{i,j}$$





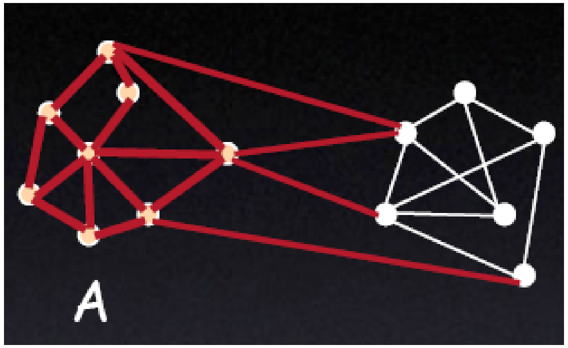


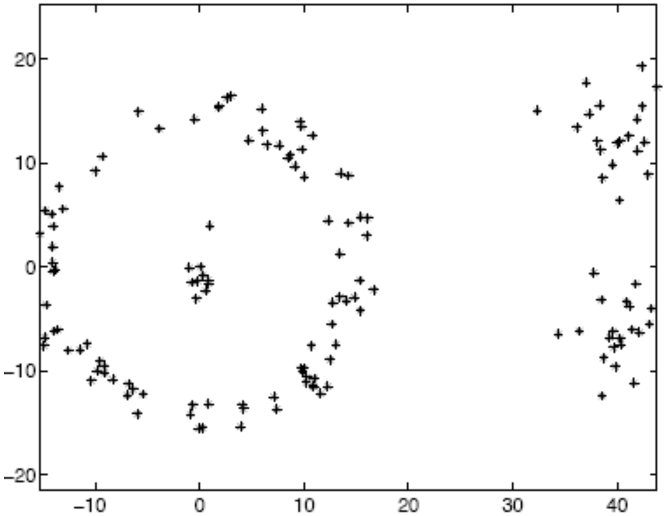
Slides from Jianbo Shi

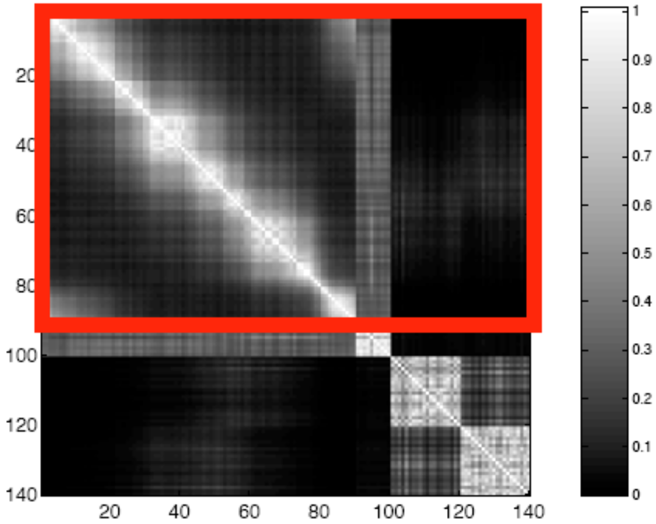
基于图的图像分割算法—图的定义

- 集合的体积:

$$vol(A) = \sum_{i \in A} d_i, A \subseteq V$$





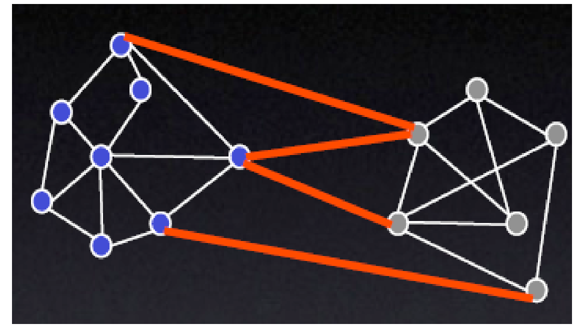


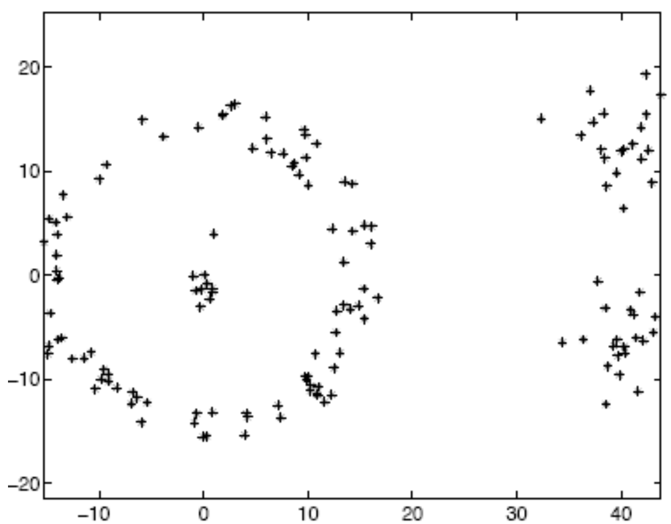
Slides from Jianbo Shi

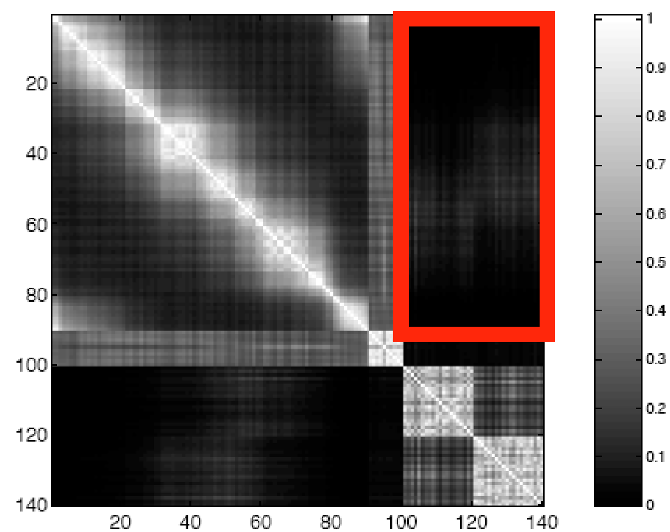
基于图的图像分割算法—图的定义

- 图上的切(Cut):

$$cut(A,\overline{A}) = \sum_{i\in A, j\in \overline{A}} w_{i,j}$$







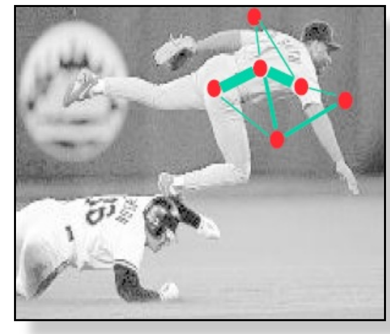
Slides from Jianbo Shi

基于图的图像分割算法一图的定义

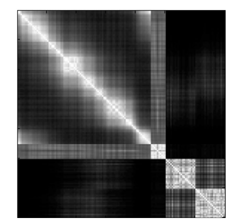
分割矩阵X:

$$X = [X_1, \ldots, X_K]$$

$$X = \begin{array}{cc} & \text{segments} \\ & \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{pixels} \end{array}$$



逐点相似矩阵(Affinity Matrix)W: $W(i,j) = aff(i,j)$



度数矩阵(Degree Matrix)D: $D(i,i) = \sum_j w_{i,j}$

拉普拉斯矩阵L: $L = D - W$

基于图的图像分割算法—相似性度量

亮度

$$W(i,j)=e^{\frac{-\left\|I_{(i)}-I_{(j)}\right\|_2^2}{\sigma_I^2}}$$

距离

$$W(i,j)=e^{\frac{-\left\|X_{(i)}-X_{(j)}\right\|_2^2}{\sigma_X^2}}$$

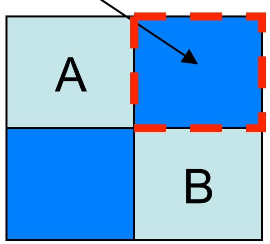
纹理

$$W(i,j)=e^{\frac{-\left\|c_{(i)}-c_{(j)}\right\|_2^2}{\sigma_c^2}}$$

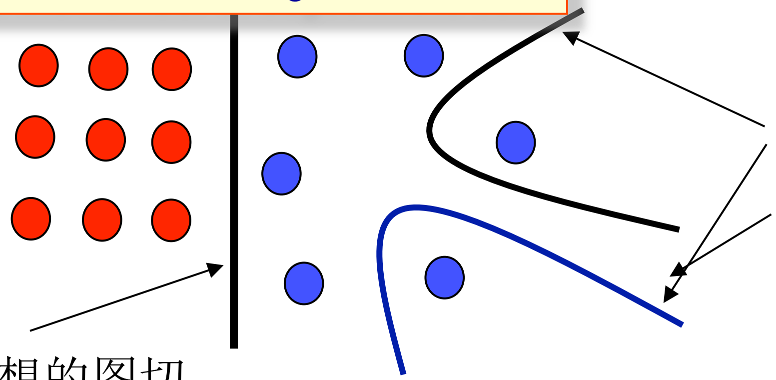
基于图的图像分割算法—最小切

- 分割准则（最小切）：

$$\min cut(A,B) = \min_{A,B} \sum_{u \in A, v \in B} w(u,v)$$



Problem!
Weight of cut is directly proportional to the number of edges in the cut.



这两种图切的代价低于理想的图切代价

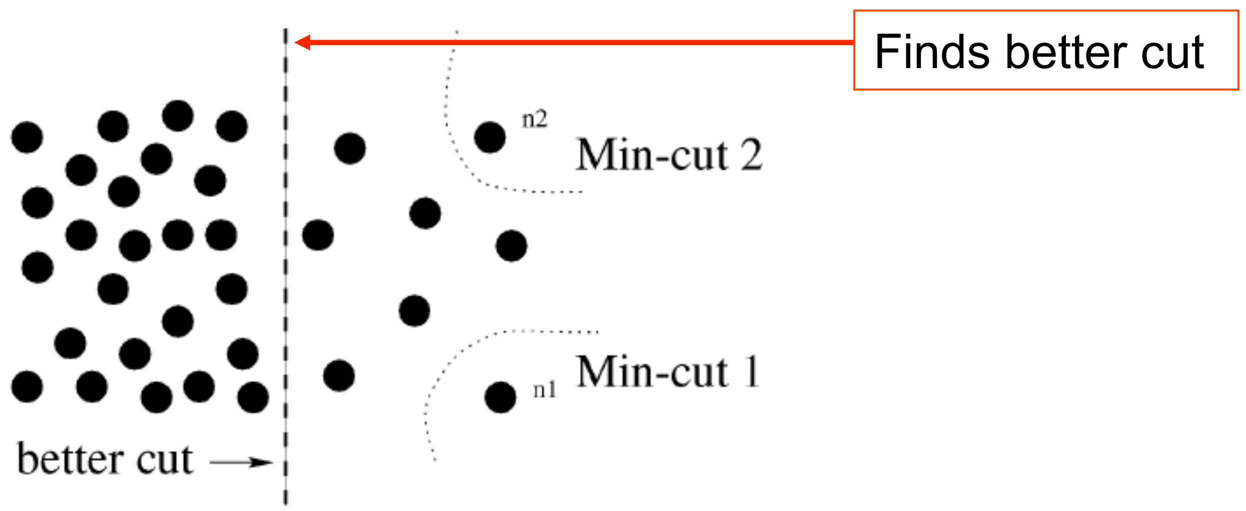
理想的图切

First proposed by Wu and Leahy

基于图的图像分割算法—正规切

正规切或平衡切（Normalized cut or balanced cut）：

$$Ncut(A,B)=cut(A,B)\left(\frac{1}{vol(A)}+\frac{1}{vol(B)}\right)$$



参考文献：

Jianbo Shi, Jitendra Malik, “Normalized Cuts and Image Segmentation,” IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997

基于图的图像分割算法－正规切

- 如何极小化*Ncut*?
  - 将图切目标函数转化为矩阵形式，可以得到简化形式：

$$\min_x Ncut(x) = \min_y \frac{y^T(D-W)y}{y^T D y}$$

NP-Hard!

Subject to: $y^T D 1 = 0$

Rayleigh熵

基于图的图像分割算法—正规切

- 将目标函数进行松弛化，得到如下连续的推广特征值系统：

$$\max_y \left(y^T (D-W) y \right) \text{ subject to } \left(y^T D y = 1 \right)$$

即求解： $(D-W)y = \lambda D y$

- 最小特征值为0：$(D-W)1 = 0$ ，相应的特征向量为y_0=1.
- 第2个最小特征向量是问题的真正解

基于图的图像分割算法—正规切

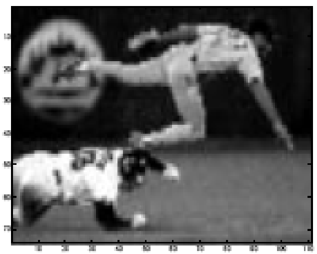
1. 定义两个图节点之间的相似性:

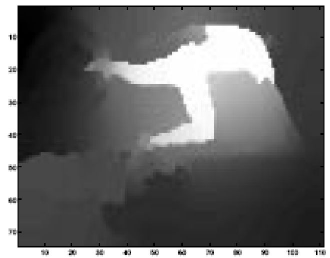
$$w_{i,j} = e^{\frac{-\|F_{(i)} - F_{(j)}\|_2^2}{\sigma_I^2} + \frac{-\|X_{(i)} - X_{(j)}\|_2^2}{\sigma_X^2}}$$

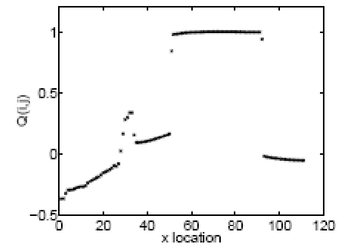
2. 计算相似性矩阵(W)和度数矩阵(D).
3. 求解 $(D-W)y = \lambda D y$
4. 利用第二小特征值对应的特征向量进行图分割.

基于图的图像分割算法－正规切

- 二值化分割：将图像分割为两个区域







分割阈值的选择：

a) 选择为常数(0, or 0.5).

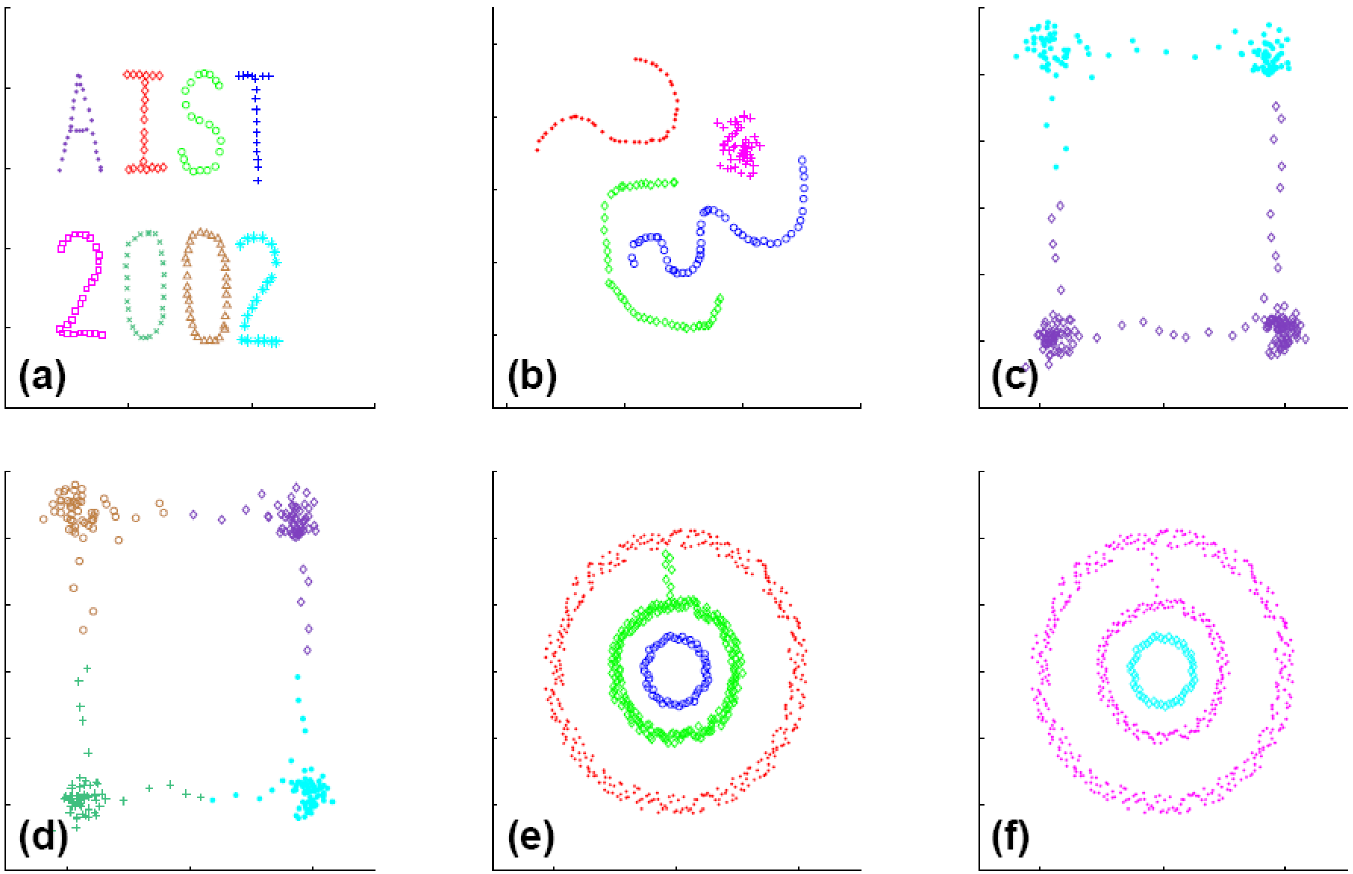
b) 选择为种值.

c) 选择一个阈值可以极小化Ncuts目标函数

基于图的图像分割算法－正规切

- 多值分割：将图像分割为多个图像块
- 算法流程: 使用多个特征向量作为特征，采用k-means 聚类获得初始分割，然后进行如下处理:
  a) 融合分割块以极小化k-way正规切目标函数.
  b) 使用k个分割块，并用穷举搜索算法找到最优分割；

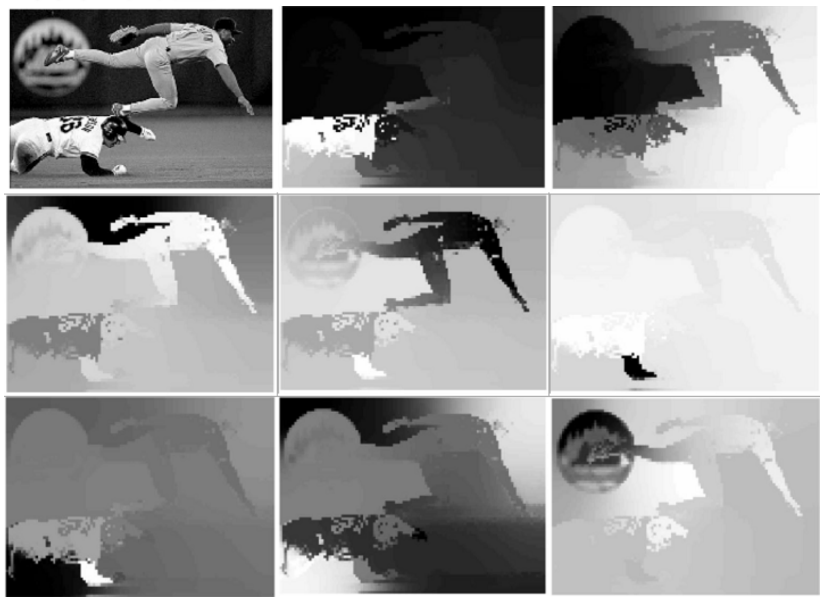
基于图的图像分割算法—正规切



Images from Matthew Brand (TR-2002-42)

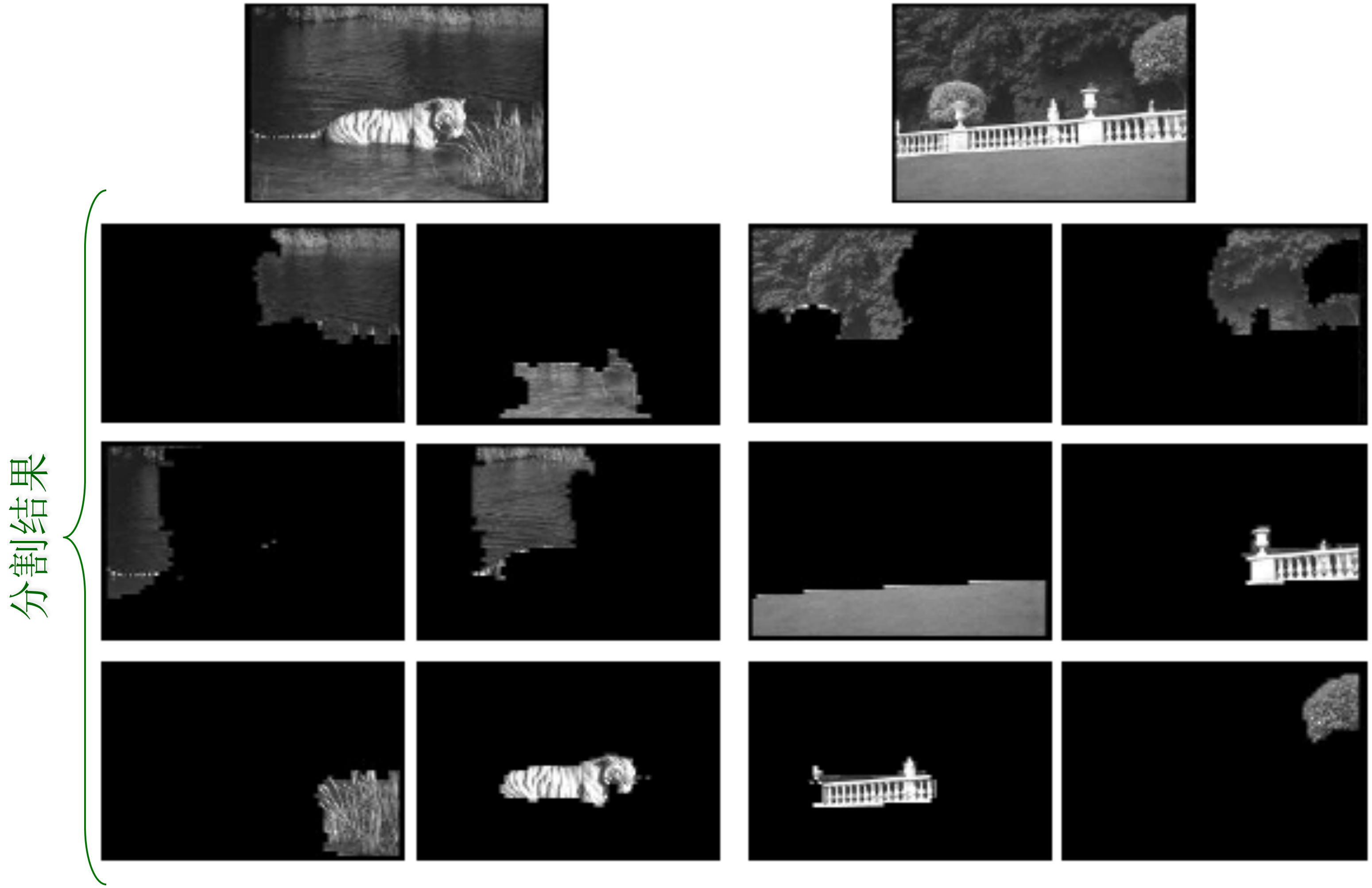
基于图的图像分割算法—正规切

特征向量



分割块

基于图的图像分割算法—正规切



* Slide from Khurram Hassan-Shafique CAP5415 Computer Vision 2003

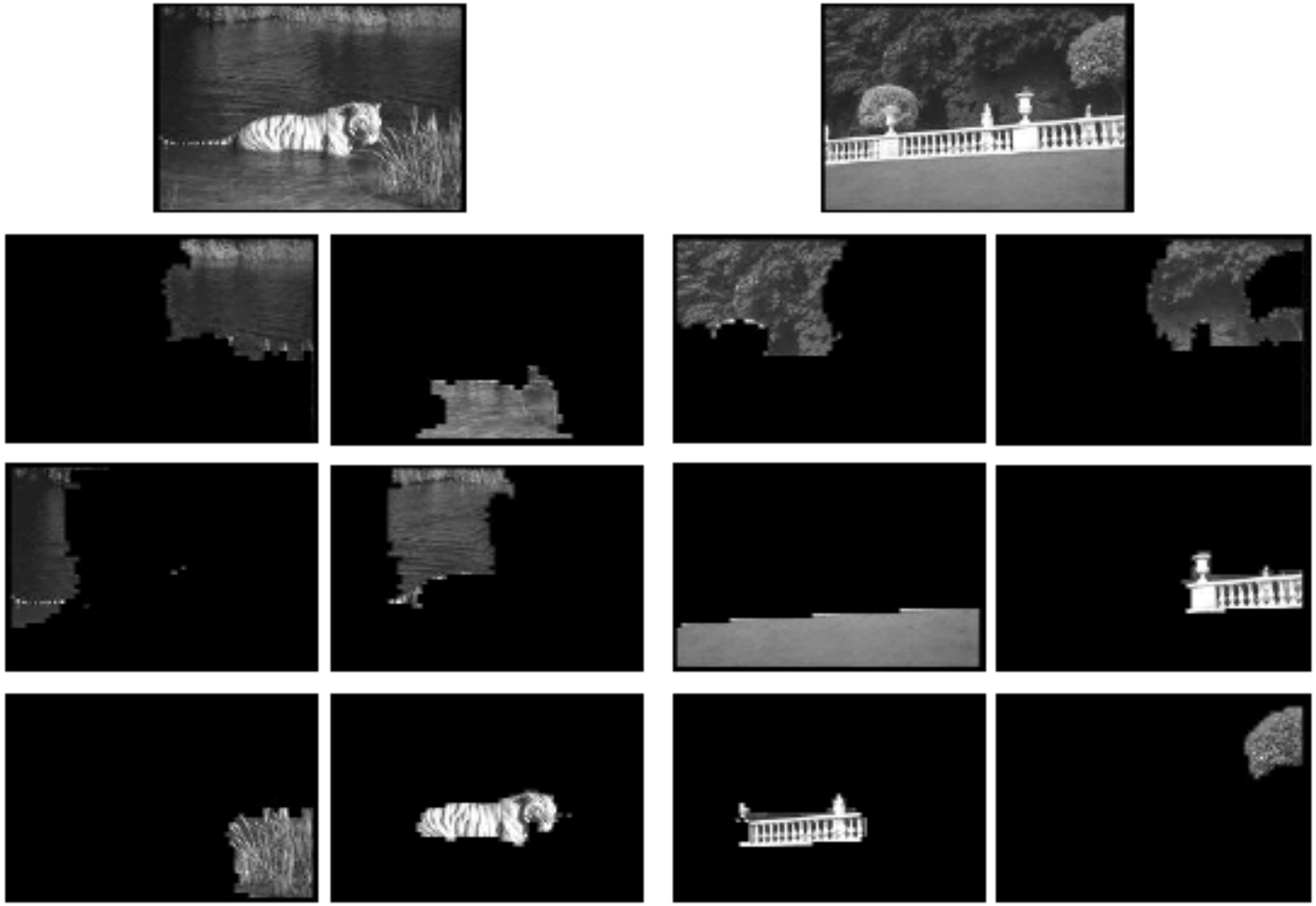
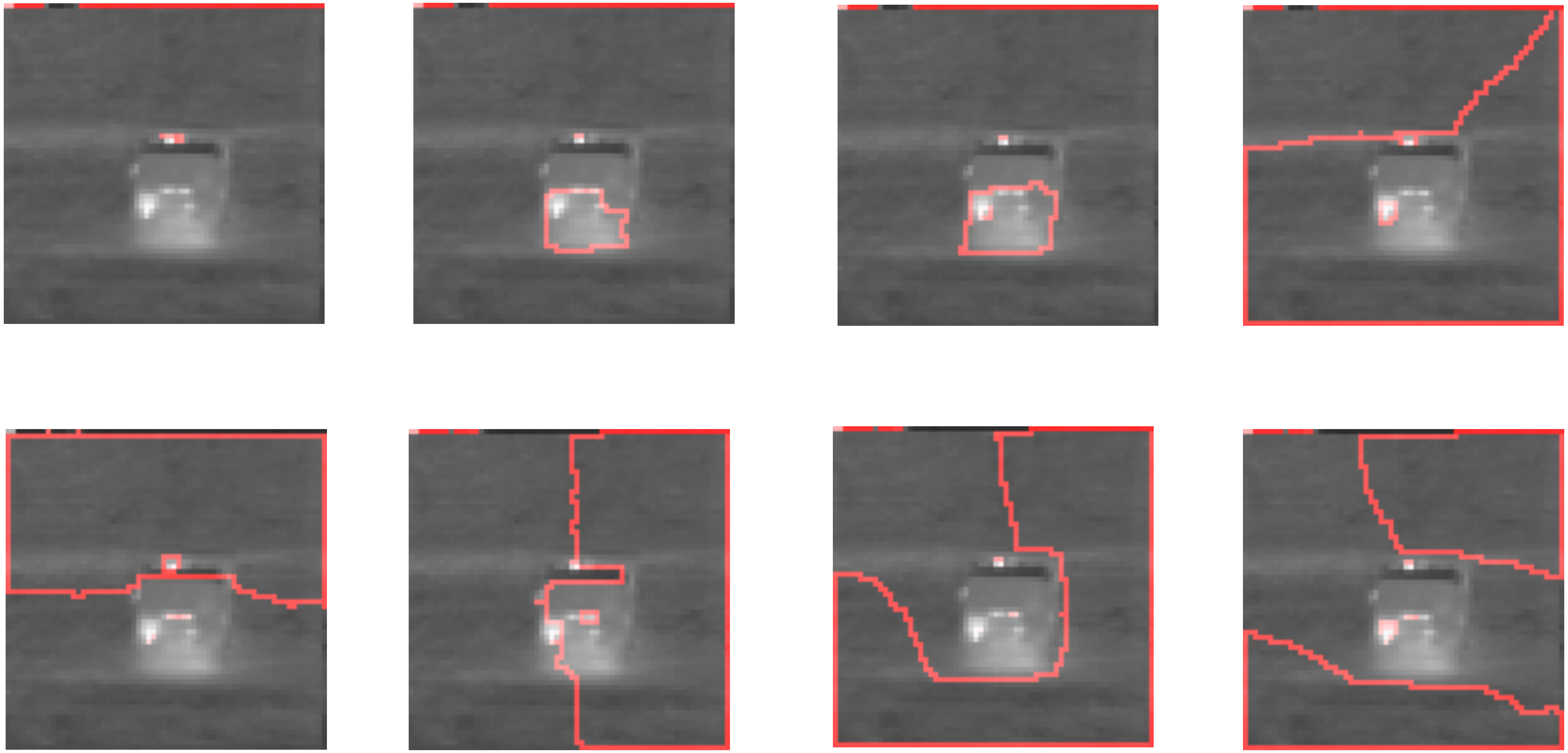
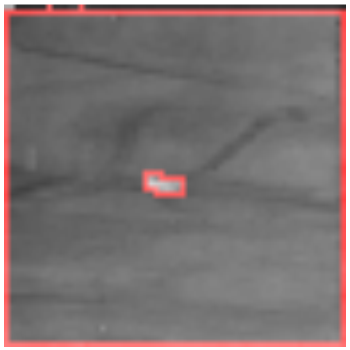
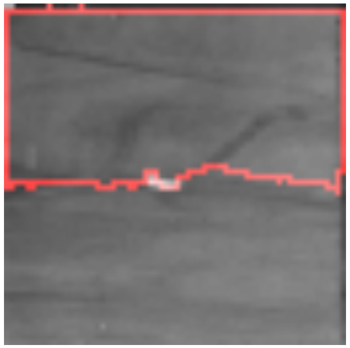
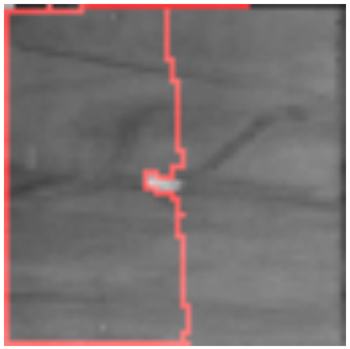
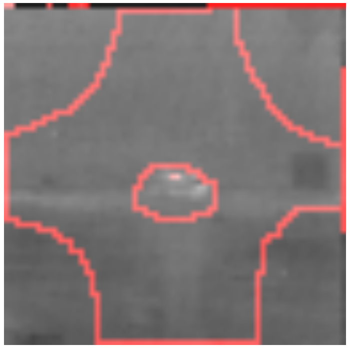
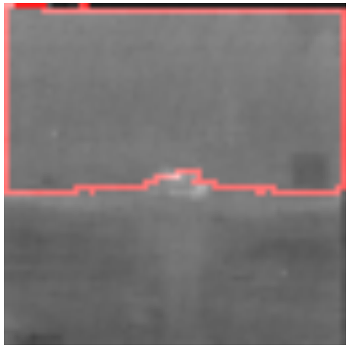
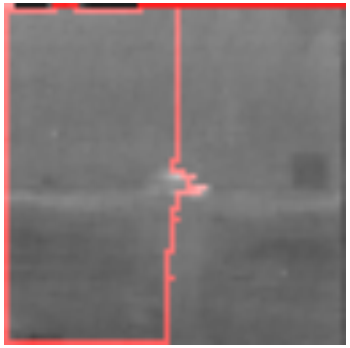


Figure from “Image and video segmentation: the normalised cut framework”, by Shi and Malik, 1998

Figure from “Normalized cuts and image segmentation,” Shi and Malik, 2000



作业

- 复习Mean-shift和Normalized Cut算法思想。参考文献：

  Jianbo Shi, Jitendra Malik, “Normalized Cuts and Image Segmentation,” IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997

  D. Comaniciu, et.al., “Mean Shift: A Robust Approach Toward Feature Space Analysis” IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002

  尝试运行两种算法程序：

  http://www.cis.upenn.edu/~jshi/software/

  http://coewww.rutgers.edu/riul/research/code.html

  Edge Detection and Image Segmentation System