

# 计算机视觉基础

## 章毓晋

清华计算机图书·译丛

Foundations of Computer Vision

## 计算机视觉基础

本书介绍计算机视觉的基础内容，比较侧重计算几何和目标检测方面。本书对图像网格的构建和叠加、德劳内三角剖分和沃罗诺伊镶嵌、多边形拼贴、图像拓扑等都有比较全面深入的介绍，并对图像结构给出了直观可视的描述。书中提供了大量相应的Matlab程序，可结合原理学习进行实验，以进一步加深理解并解决实际问题。

本书可作为信号与信息处理、通信与信息系统、电子与通信工程、模式识别与智能系统、计算机科学等学科大学高年级本科生或研究生专业课教材和教学参考书，还可供涉及图像和机器视觉技术的应用行业（如生物医学、电视广播、工业自动化、文档识别、机器人、电子医疗设备、遥感测绘、增强现实、智能交通和军事侦察等）的科技工作者和从业者参考和自学。

课件下载·样书申请



书圈

清华社官方微信信号



扫 我 有 惊 喜



定价：98.00元

清华计算机图书·译丛

计算机视觉基础

清华大学出版社



Springer

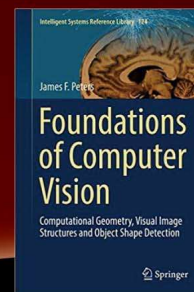
清华计算机图书·译丛

Foundations of Computer Vision

## 计算机视觉基础

[加] 詹姆斯·彼得斯 (James F. Peters) 著

章毓晋 译



清华大学出版社



## 第7章 视频处理，实时和离线 视频分析介绍

在帧图像的沃罗诺伊拼贴中，可以通过对多边形（区域）的形状，位置和分布变化来检测视频帧变化

对视频帧变化的研究可以实时或离线完成

从计算机视觉的角度来看，摄像机拍摄的场景取决于摄像机的光圈角度及其视场的视野，类似于人类的感知角度



## 7.1 视频处理基础

视频中的基本单位是帧。帧是线性序列图像中的单个数字图像

### 视频分析的基本步骤

摄像头 → 图像采集 → 预处理 → 帧结构化 → 分类

帧图像点处理：

伪彩色，像素选择，像素操作，像素（点）处理，滤波和分割



## 7.1 视频处理基础

视频处理从图像采集过程开始

图像采集基本上是一个两步过程，其中将单幅图像（称为帧）添加到图像序列中

为存储视频会消耗大量内存。因此，图像压缩是视频图像采集中的核心问题

量化器由编码器和解码器确定，编码器对输入信号值的集合进行分类，解码器指定输出值集合



## 7.1 视频处理基础

斑块（二值大目标）是二值图像中一组路径连接的像素。连通性的概念使得可以将斑块的概念扩展为灰度中的灰色斑块和彩色图像中的彩色斑块

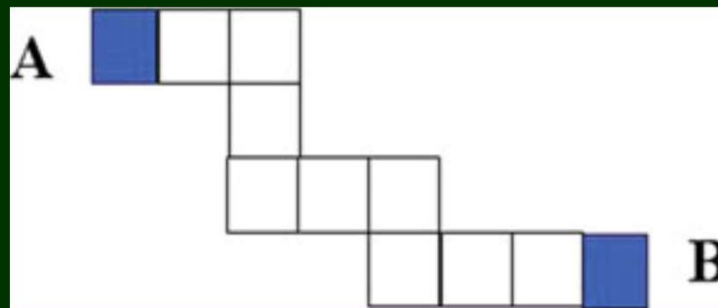
如果多边形共享一个或多个点，则多边形是连接的。具有公共边包含 $n$ 个点的连接沃罗诺伊多边形是 $n$ -相邻的。类似地，具有公共顶点包含 $n$ 个点的德劳内三角形是连接的并且是 $n$ -相邻的



## 7.1 视频处理基础

具有 $n$ 个像素或体素的序列 $p_1, \dots, p_i, p_{i+1}, \dots, p_n$ 是一条路径，如果 $p_i$ 和 $p_{i+1}$ 相邻（ $p_i$ 和 $p_{i+1}$ 之间没有像素）。像素 $p$ 和 $q$ 是路径连接的，前提是存在以 $p$ 和 $q$ 作为端点的路径

类似地，图像形状 $A$ 和 $B$ （任何多边形）是路径连接的，只要 $n$ 个相邻形状构成序列 $S_1, \dots, S_i, S_{i+1}, \dots, S_n$ ，且 $A = S_0, B = S_n$

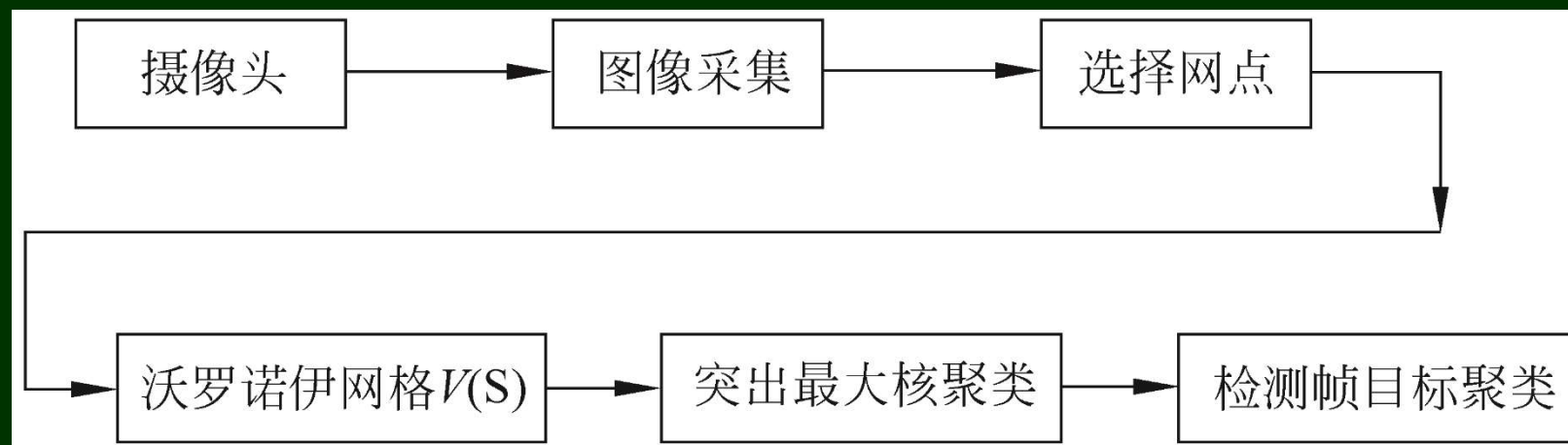




## 7.1 视频处理基础

每个视频帧都可以使用沃罗诺伊图进行拼贴（镶嵌），或使用德劳内的三角测量方法拼贴每个帧

任何形式的框架拼贴的自然结果是网格聚类和目标识别





## 7.2 视频帧的沃罗诺伊拼贴

平面沃罗诺伊拼贴是对具有非重叠多边形表面的覆盖。设 $\times$ 表示沃罗诺伊区域中质心的位置，那么可用连接围绕目标的沃罗诺伊区域质心的线来识别帧目标的轮廓



轮廓1



轮廓2





## 7.3 在视频帧中检测目标形状

通过在每个视频帧上构建沃罗诺伊拼贴（也称为沃罗诺伊图），可辅助检测视频序列中人运动中的个人空间

个人空间由运动中的人之间的舒适距离限定

亲密： $0 \leq d \leq 0.5\text{m}$ （友谊距离）

私人： $0.5 \leq d \leq 1.25\text{m}$ （会话距离）

社交： $1.25 \leq d \leq 3.5\text{m}$ （非个人距离）

公众： $d \geq 3.5\text{m}$ （公开演讲距离）



## 7.3 在视频帧中检测目标形状

基于人与人之间舒适距离的概念，可以研究连续视频帧中人的运动

相对于连续沃罗诺伊视频帧拼贴中人与人之间距离的变化，引入了感知个人空间（PPS）的概念

$$\text{PPS}(V_f) \geq \frac{\alpha\pi}{360^\circ} R_c^2 \text{ (视频帧感知个人距离)}$$

透镜（人眼）的孔径角是从焦点看到的透镜孔径的视角

$$\alpha = 2 \tan^{-1} \left( \frac{D}{2f} \right) \text{ (孔径角)}$$



## 7.4 测量目标的形状相似性和 沃罗诺伊视觉外壳

在视频帧的沃罗诺伊拼贴中，目标的视觉外壳（VH）是围绕目标的多边形的集合

设 $A$ 是目标轮廓上的点集， $B$ 为已知目标的轮廓。  
在点 $x$ 和集合 $A$ （记为 $D(x, A)$ ）之间的豪斯道夫距离

$$D(x, A) = \min \{ \|x - a\| : a \in A \} \text{ (豪斯道夫点 - 集合距离)}$$

两个轮廓 $A$ 和 $B$ 之间的相似距离 $D(A, B)$

$$D(A, B) = \max \left\{ \max_{a \in A} D(a, B), \max_{b \in B} D(b, A) \right\} \text{ (相似距离)}$$



## 7.4 测量目标的形状相似性和 沃罗诺伊视觉外壳

视频帧序列中的视觉外壳的形状之间的相似性是相对于已知视觉外壳形状之间的距离来测量的

设 $A, B_i$ 为距离 $D(A, B_i)$ 的已知形状。令 $S_0$ 表示已知形状，与 $n$ 个视频帧序列中的形状 $S_j, 1 \leq j \leq n$ 进行比较  
将距离 $D(A, B_i)$ 与 $S_0$ 和 $S_i$ 之间的距离总和进行比较

$$D(A, B_i) \approx \sum_{j=1}^n \left( \sum_{\substack{a \in S_0 \\ b \in S_j}} \|a - b\| \right) \quad (\text{相似帧形状})$$



## 7.4 测量目标的形状相似性和 沃罗诺伊视觉外壳

设  $\varepsilon > 0$  是一个小的数，则形状  $A$  和  $S_0$  被认为是接近的，只要

$$\text{shapeDiff}(S_0, S_j) := \sum_{j=1}^n \left( \sum_{\substack{a \in S_0 \\ b \in S_j}} \|a - b\| \right)$$
$$|D(A, B_i) - \text{shapeDiff}(S_0, S_j)| \leq \varepsilon$$

对于形状  $S_1, \dots, S_i, S_{i+1}, \dots, S_n$  和一些小的数  $\varepsilon > 0$ ，  
检查下式是否成立：

$$|D(A, B_i) - \text{shapeDiff}(A, S_j)| \leq \varepsilon \quad \text{对 } 1 \leq j \leq n$$



## 7.5 最大核聚类

曲面的沃罗诺伊镶嵌中的每个多边形都是包含与核相邻的所有多边形的聚类的中心（核）。沃罗诺伊网格核是任何与核相邻的沃罗诺伊区域的中心沃罗诺伊区域

### 最大核聚类（MNC）

如果核 $N$ 在镶嵌表面中具有最高数量的相邻多边形（记为 $\max CN$ ），则具有核 $N$ 的连接多边形聚类是最大的



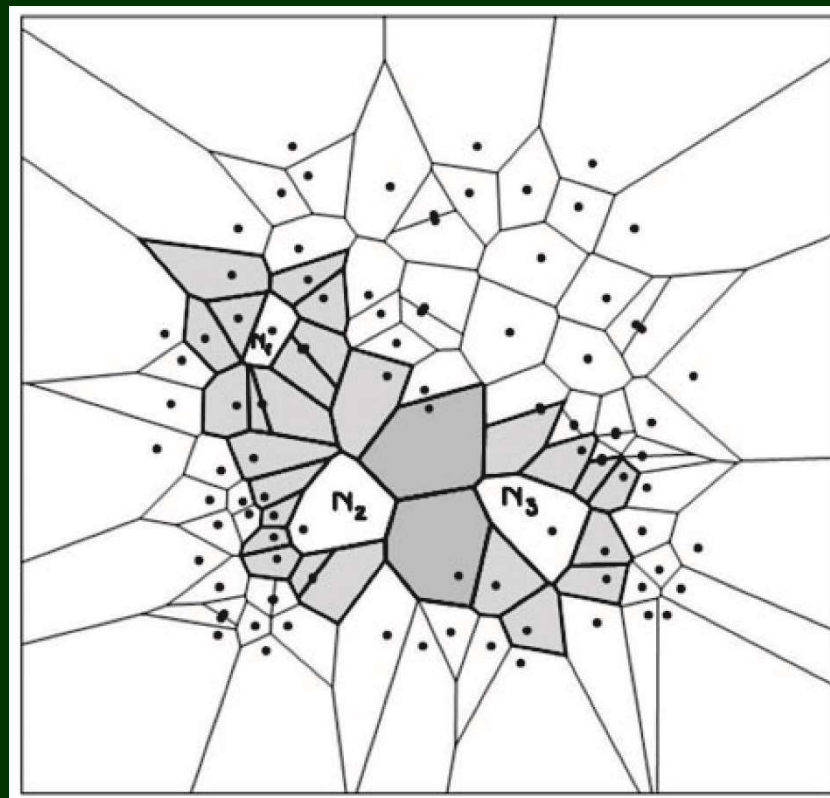
## 7.5 最大核聚类

设 $X$ 是沃罗诺伊区域的集合， $X$ 在曲面镶嵌中包含最大核聚类 $CN_1$ 、 $CN_2$ 、 $CN_3$

$CN_1$ 与 $CN_2$ 相邻

$CN_2$ 与 $CN_3$ 共享沃罗诺伊区域

相邻或重叠的MNC具有由其周边确定的形状，它们围绕镶嵌图像中的感兴趣区域



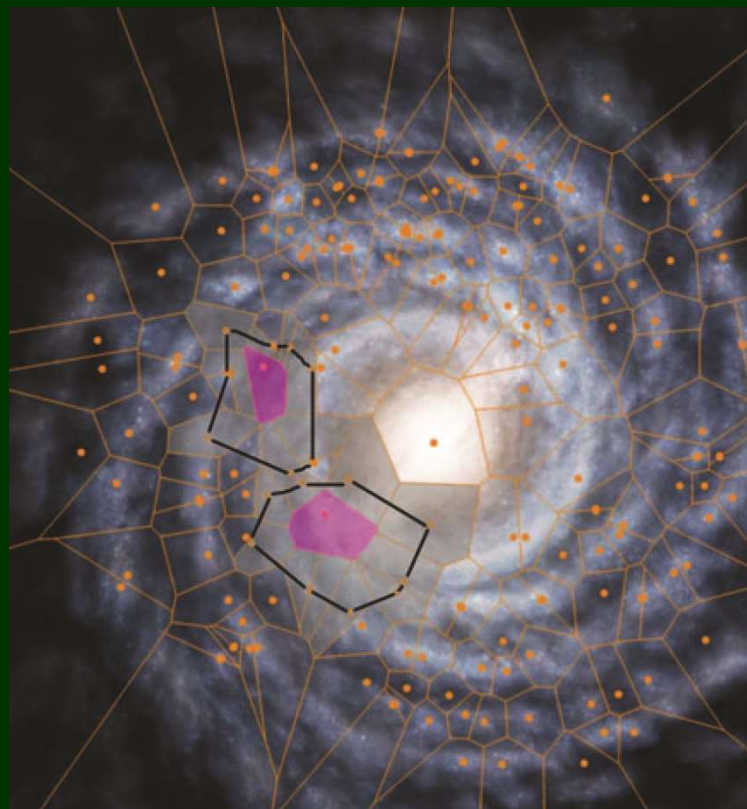


## 7.5 最大核聚类

银河系的地图



两个最大核聚类的轮廓

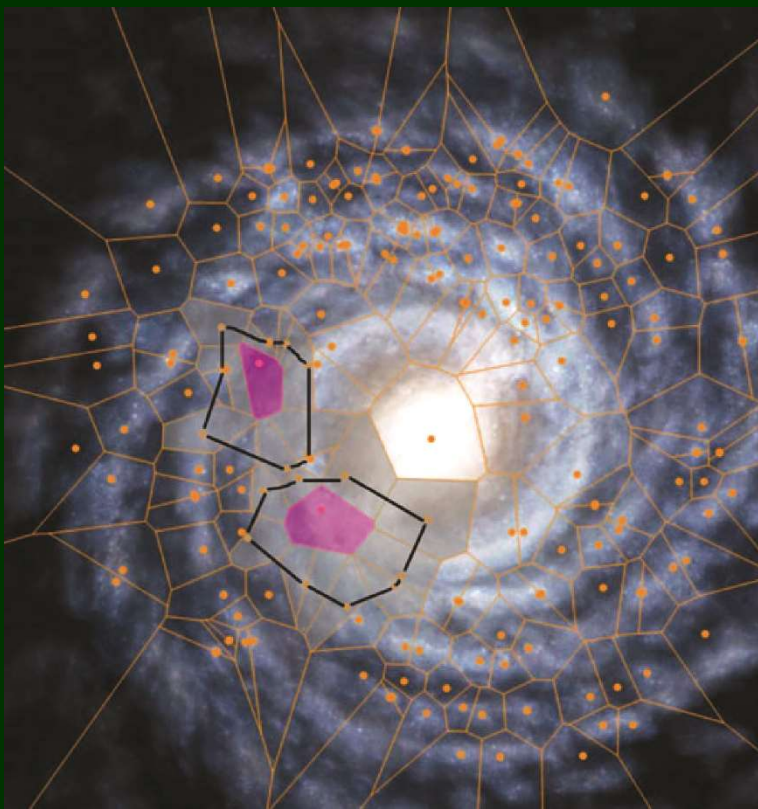




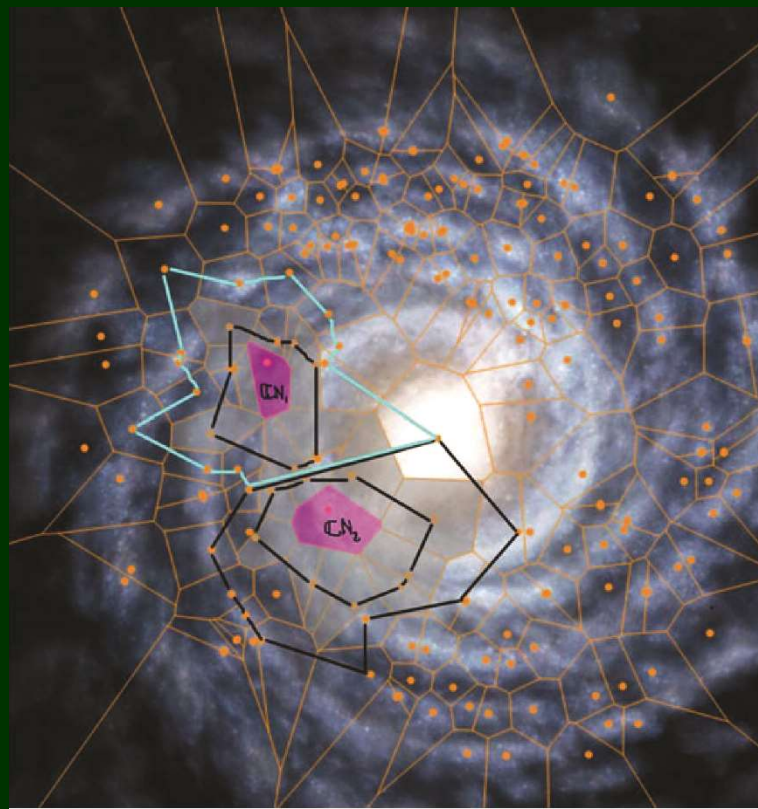


## 7.5 最大核聚类

聚类轮廓



直边通路连接





## 7.5 最大核聚类

### (1) 精细聚类轮廓

在与聚类核相邻的多边形中，精细聚类轮廓是其中每对相邻生成点由直边相连接的通路。换句话说，精细聚类轮廓是包含相邻直边的直边连接通路。如果边缘具有共同的端点，则直边相邻

### (2) 粗糙聚类轮廓

在围绕与聚类核相邻的那些多边形的多边形中，粗糙聚类轮廓是其中每对相邻生成点通过直边连接的通路



## 7.7 形状距离

计算MNC轮廓边缘集之间的距离代价

设 $e_i$ 、 $e_j$ 为边缘集，取 $a, b > 0$ 为用于调整如下定义代价函数 $C_{\text{dist}}(e_i, e_j)$ 的常数：

$$C_{\text{dist}}(e_i, e_j) = \frac{a}{1 + \exp(-b \|e_i - e_j\|)}$$

对 $a$ 和 $b$ 的选择基于达到 $e_i$ 和 $e_j$ 之间距离的最大代价进行。例如，设 $b := 1$ 并且 $a := D(e_i, e_j)$ （最小距离）

$$D(e_i, e_j) = \min \{ \|x - y\| : x \in e_i, y \in e_j \}$$



## 7.8 边缘集的权函数

对于边缘集 $e_i$ 和 $e_j$ ，距离代价权函数 $C(e_i, e_j)$

$$C(e_i, e_j) = C_{\text{dist}}(e_i, e_j) + C_{\text{shape}}(e_i, e_j)$$

包括距离和形状两个因素

如果代价函数针对一对边缘集合 $e_{\text{target}}$ 和 $e_j$ ，即目标MNC轮廓边缘集和视频中的样本边缘集，则得到的 $C(e_{\text{target}}, e_j)$ 为

$$C(e_{\text{target}}, e_j) = C_{\text{dist}}(e_{\text{target}}, e_j) + C_{\text{shape}}(e_{\text{target}}, e_j)$$



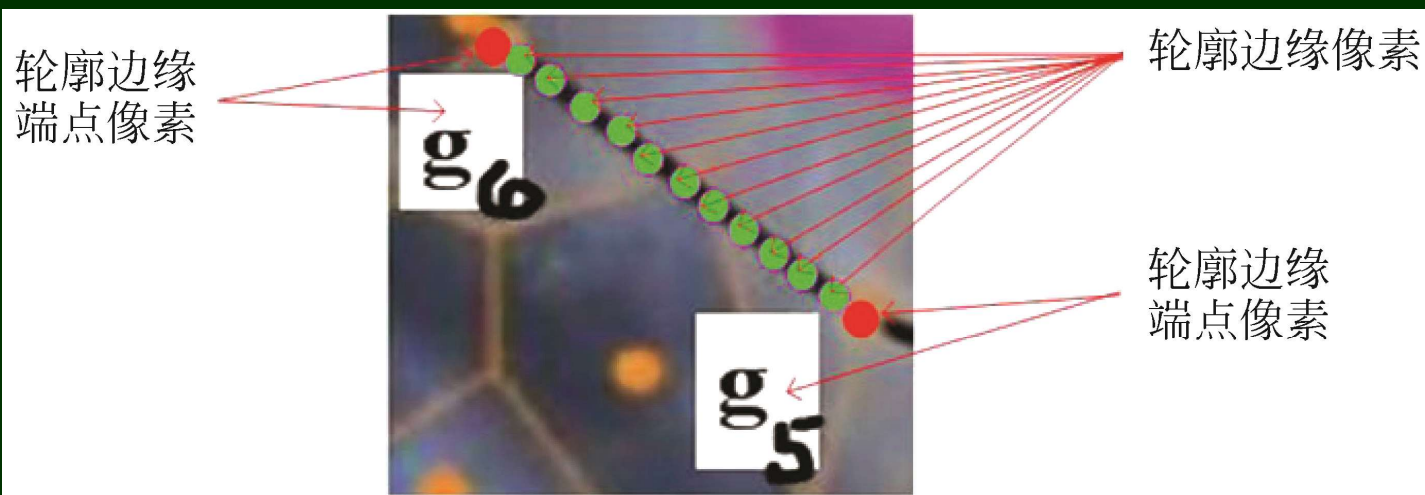
## 7.9 最大边缘集

最大MNC轮廓边缘集（记为 $\max e_i$ ）是包含所有轮廓边缘像素的边缘集

$\max |e_i| = \# \text{轮廓边缘像素}$ ，而不仅是直边缘的端点

$\max e_i$ 将包括端点 $g_5$ 、 $g_6$ 以及轮廓直边 $\overline{g_5 g_6}$ 的内部

像素

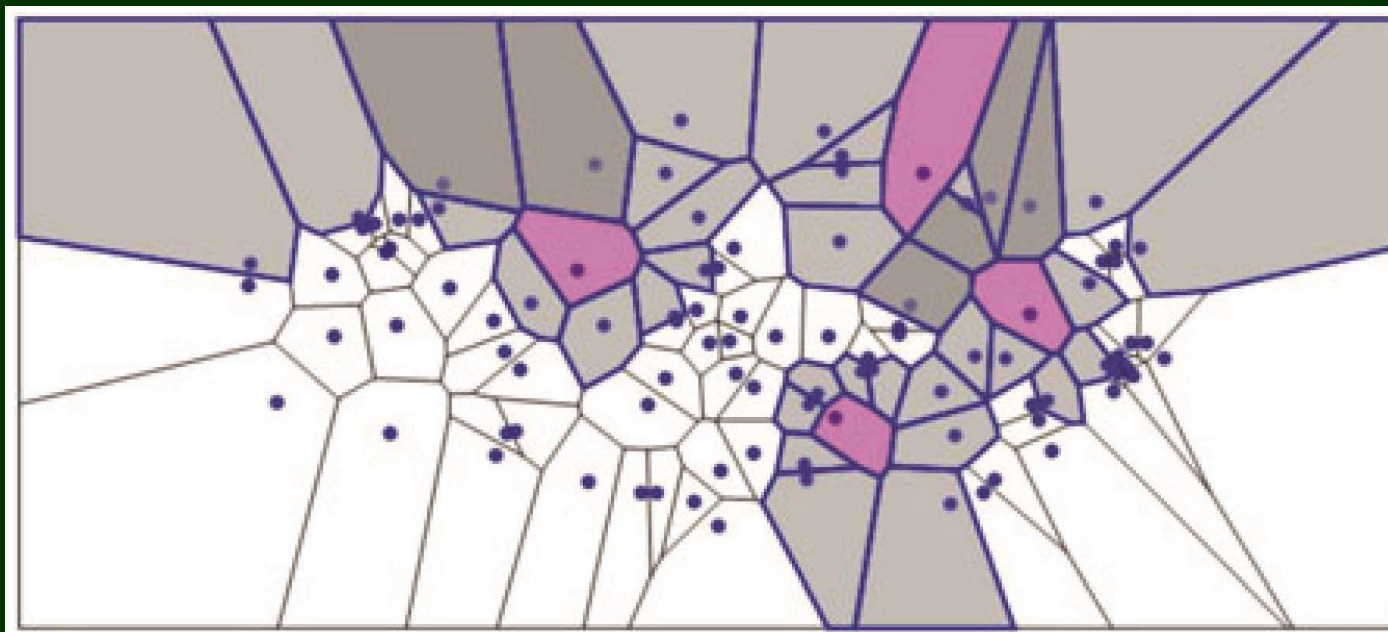




## 7.9 最大边缘集

粗糙轮廓边缘集是由连接围绕精细轮廓多边形生成点的直边所定义的边缘集

考虑沃罗诺伊网格中4个MNC



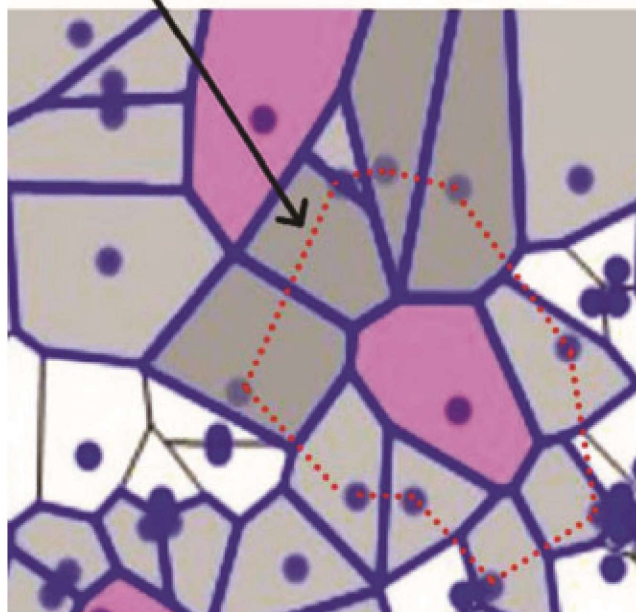




## 7.9 最大边缘集

(红色)  
虚线  
表示精细  
MNC轮  
廓边缘  
 $\max e_{\text{fine}}$   
中的端  
点和内  
部直边  
像素

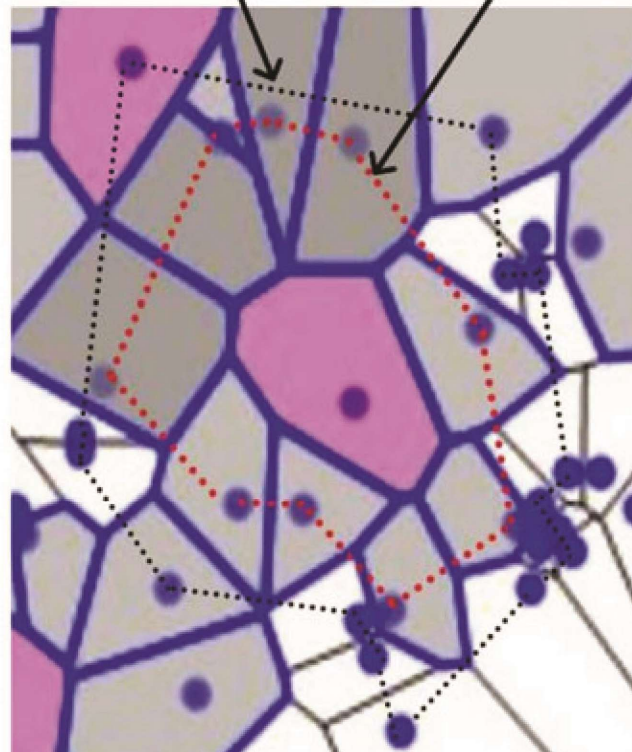
精细轮廓边缘集



(a) 精细边缘集

粗糙轮廓边缘集

精细轮廓边缘集



(b) 粗糙边缘集

(黑色)  
虚线  
表示粗糙  
MNC轮  
廓边缘  
 $\max e_{\text{coarse}}$   
中的端  
点和内  
部直边  
像素



## 7.9 最大边缘集

设MNC1和MNC2是沃罗诺伊曲面镶嵌中的一对最大核聚类。如果它们具有相邻的多边形或者具有至少一个共同的多边形，则MNC1和MNC2是连接的





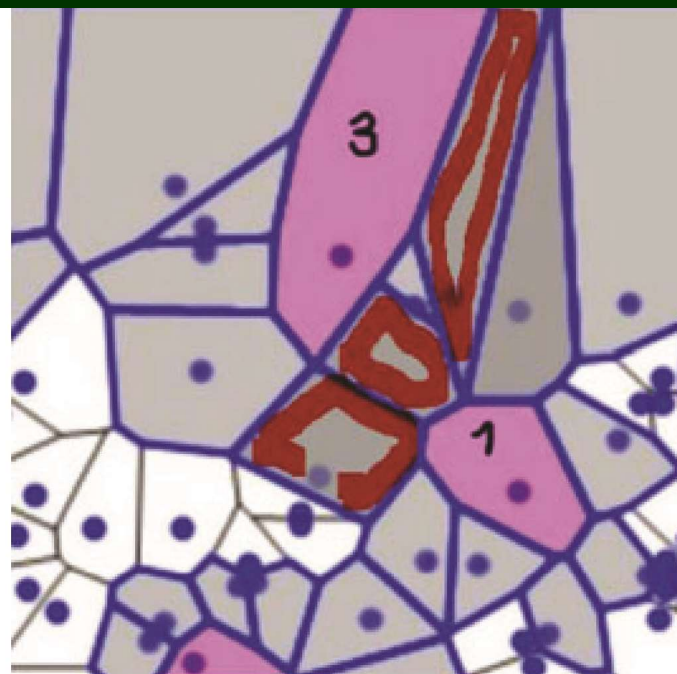


## 7.9 最大边缘集

强连接的MNC，因为它们共享多于一个像素



(a) MNC1和MNC2具有相邻的沃罗诺伊网格多边形



(b) MNC1和MNC3共享3个沃罗诺伊网格多边形



## 译者（章毓晋）联系信息

- ✎ 通信地址：北京清华大学电子工程系
- ✎ 邮政编码：100084
- ✎ 办公地址：清华大学罗姆楼，6层305室
- ✎ 办公电话：(010) 62798540
- ✎ 传真号码：(010) 62770317
- ✎ 电子邮件：[zhang-yj@tsinghua.edu.cn](mailto:zhang-yj@tsinghua.edu.cn)
- ✎ 个人主页：[oa.ee.tsinghua.edu.cn/~zhangyujin/](http://oa.ee.tsinghua.edu.cn/~zhangyujin/)  
(下载更新的讲稿和教材修改表)