

 Springer

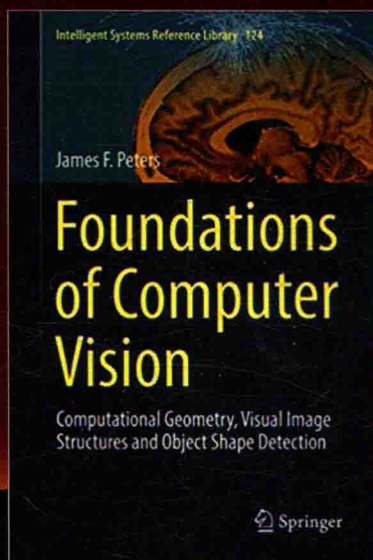
清华计算机图书·译丛

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



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  
北京

北京市版权局著作权合同登记号 图字：01-2018-8799

Translation from English language edition: **Foundations of Computer Vision** by **James F. Peters**  
Copyright © 2017, Springer-Verlag Berlin Heidelberg. Springer Berlin Heidelberg is a part of  
Springer Science+Business Media.

All Rights Reserved.

本书中文简体字翻译版由德国施普林格公司授权清华大学出版社在中华人民共和国境内(不包括中国香港、澳门特别行政区和中国台湾地区)独家出版发行。未经出版者预先书面许可,不得以任何方式复制或抄袭本书的任何部分。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。  
版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

### 图书在版编目(CIP)数据

计算机视觉基础 / (加)詹姆斯·彼得斯(James F. Peters)著;章毓晋译. —北京:清华大学出版社,2019  
(清华计算机图书译丛)

书名原文:Foundations of Computer Vision—Computational Geometry, Visual Image Structures and Object  
Shape Detection

ISBN 978-7-302-52995-8

I. ①计… II. ①詹… ②章… III. ①计算机视觉 IV. ①TP302.7

中国版本图书馆CIP数据核字(2019)第096670号

责任编辑:龙启铭  
封面设计:傅瑞学  
责任校对:徐俊伟  
责任印制:丛怀宇

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦A座 邮 编:100084

社 总 机:010-62770175 邮 购:010-62786544

投稿与读者服务:010-62776969, [c-service@tup.tsinghua.edu.cn](mailto:c-service@tup.tsinghua.edu.cn)

质 量 反 馈:010-62772015, [zhiliang@tup.tsinghua.edu.cn](mailto:zhiliang@tup.tsinghua.edu.cn)

课 件 下 载: <http://www.tup.com.cn>, 010-62795954

印 装 者:小森印刷(北京)有限公司

经 销:全国新华书店

开 本:185mm×260mm

印 张:21

字 数:513千字

版 次:2019年8月第1版

印 次:2019年8月第1次印刷

定 价:98.00元

产品编号:081740-01

# 译者序

本书是一本介绍计算机视觉基础内容，并辅以相关 Matlab 程序（部分还提供了 Mathematica 程序）的书籍。

本书原内容主要总结自作者为本科生讲授计算机视觉课程的讲义和笔记。有些主题是作者（及同事、学生）研究工作的介绍。两者的结合对撰写计算机视觉这种既有较长的历史，又在近期得到广泛关注的领域非常适合。此书也比较适合作为相关课程的教材。

本书除介绍基本的图像处理和计算机视觉的内容外，重点是对计算几何和目标检测的介绍。作者借助德劳内三角剖分和沃罗诺伊镶嵌来研究图像及目标的几何信息，以帮助发现图像中隐藏的模式和嵌入的目标形状。书中对图像网格的构建和叠加、多边形拼贴、拓扑邻域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。

本书具有实用教材的一些特点。书中所介绍的大多数图像处理和计算机视觉技术都提供了相应的 Matlab 程序，读者不仅能了解基本原理，还可以实现相应技术的算法并获得可视的结果。书中给出的问题，既有思考题，也有练习题，其难度也分了不同的档次，方便了选择使用。书中还介绍了一些深入学习或应用的参考文献。

本书从结构上看，共有 9 章正文和两个附录，包括 128 节、70 小节。全书共有编了号的图 354 个、表格 6 个、例 141 个、注释 53 个、引理 4 个、定理 8 个、问题（练习题或思考题）117 个、算法 10 个、列表（Matlab 程序）105 个以及 Mathematica 程序 8 个。另外还有参考文献、作者索引和主题索引。本书可作为相关专业本科生和其他专业研究生学习图像技术的课程教材，也可供从事相关领域科技开发和应用的技术人员自学参考。

本书的翻译基本忠实于原书的描述结构和文字风格。对明显的印刷错误，直接进行了修正。将原书混合编号的例和问题（以及仅在个别章节出现的引理和定理）分别按各自顺序重新进行了编号，将原来由于排版原因而没有依次出现的图，按文中引用的顺序进行了重排，将原来按全书编号的算法改为分章编号以与其他图表等标号一致。另外，对原书附录 B 的词汇表和书后的主题索引，均重新按中文拼音顺序进行了排列，以方便读者查阅。最后，根据中文书籍规范，将矢量和矩阵均改用了粗斜体标注。

感谢清华大学出版社编辑的精心组稿、认真审阅和细心修改。

最后，作者感谢妻子何芸、女儿章荷铭在各方面的理解和支持。

章毓晋

2019 年元旦于书房

通信：北京清华大学电子工程系，100084

电话：(010) 62798540

传真：(010) 62770317

邮箱：zhang-yj@tsinghua.edu.cn

主页：oa.ee.tsinghua.edu.cn/~zhangyujin/



# 前 言

本书介绍计算机视觉基础。计算机视觉（也称为机器视觉）的主要目标是基于由各种相机捕获的图像内容来重建和解释自然场景（参见 R. Szeliski[187]）。计算机视觉系统包括诸如卫星测绘、机器人导航系统、智能扫描仪和遥感系统之类的东西。在计算机视觉研究中，重点是从图像中提取有用的信息（参见 S. Prince [159]）。计算机视觉系统通常模仿人类视觉感知。计算机视觉系统中选择的硬件是某种形式的数码相机，被编程为近似视觉感知。因此，计算机视觉、数字图像处理、光学、光度学和光子学之间存在密切联系（参见 E. Stijns 和 H. Thienpont [184]）。

从计算机视觉的角度来看，光子学是捕捉视觉场景中光的科学。图像处理研究数字图像形成（如从模拟光学传感器信号到数字信号的转换）、操作（如图像滤波、去噪、裁剪）、特征提取（如像素强度、梯度方向、梯度幅度、边缘强度）、描述（如图像边缘和纹理）和可视化（如像素强度直方图）。可以参见 B. Jähne [85]和 S.G. Hoggar [81]的图像处理数学框架，并扩展到许多相关人员对图像处理的看法，如 M. Sonka、V. Hlavac 和 R. Boyle [183]，W. Burger 和 M. J. Burge [21]，R. C. 冈萨雷斯和 R. E. 伍兹[57]，R. C. 冈萨雷斯、R. E. 伍兹和 S. L. 埃丁斯 [58]，V. Hlavac [80]，C. Solomon 和 T. Breckon [181]。这些有用的信息，即可以检测、分析和分类的图像目标形状和模式，为计算机视觉研究者的关注点提供了基石（如[140]）。实际上，计算机视觉是对数字化图像结构和模式的研究，而这是在图像处理和光子学之上的一层的图像分析。计算机视觉将图像处理和光子学包括在其追求图像几何和图像区域模式的技巧集合中。

此外，培养数字图像的智能系统观点可帮助发现隐藏的模式（如图像区域的重复凸轮廓）和嵌入的图像结构（如图像中感兴趣区域的点聚类）。通过量化器可以发现这种结构。量化器将一组（通常是连续的）值限制为离散值。量化器在其计算机视觉里最简单的形式中，观察特定的目标像素强度并在目标邻域中选择最接近的近似值。量化器的输出被 A. Gersho 和 R.M. Gray 称为码本[55, 5.1 节, p.133]（另见 S. Ramakrishnan、K. Rose 和 A. Gersho [161]）。

在图像网格叠加的上下文中，Gersho-Gray 量化器被替换为基于几何的量化器。基于几何的量化器将图像区域限制在其外形轮廓中并在图像中观察特定的目标形状轮廓，与其他具有近似形状的目标轮廓进行比较。在计算机视觉的基础上，基于几何的量化器观察并比较大致相同的图像区域，例如将最大核聚类（MNC）与其他核聚类进行比较。最大核聚类（MNC）是围绕称为核的网格多边形的图像网格多边形的集合（参见 J. F. Peters 和 E. İnan 在 Edelsbrunner 神经中的沃罗诺伊图像镶嵌[147]）。一个图像网格核是一个网格多边形，它是相邻多边形集合的中心。事实上，每个网格多边形都是一个多边形聚类的核。然而，只有一个或多个网格核是最大的。

最大图像网格核是具有最高数量的相邻多边形的网格核。MNC 在计算机视觉中很重

要，因为 MNC 轮廓能近似所考虑图像目标的形状。对图像的沃罗诺伊镶嵌是用多边形对图像的拼贴。图像的沃罗诺伊镶嵌也称为沃罗诺伊网格。对图 0.1 (a) 中音乐家图像的拼贴样本图如图 0.1 (b) 所示。在每个拼贴多边形内的红点 (●) 是沃罗诺伊区域 (多边形) 生成点的示例。更多的信息参见 1.22.1 小节。这个音乐家网格核是如图 0.2 (b) 所示最大核聚类的中心。这是图 0.1 (b) 中音乐家图像网格中唯一的 MNC。该 MNC 也是沃罗诺伊网状神经的一个例子。对图像 MNC 的研究将我们带到图像几何和图像目标形状检测的入门处。更多内容可见 1.22.2 小节。

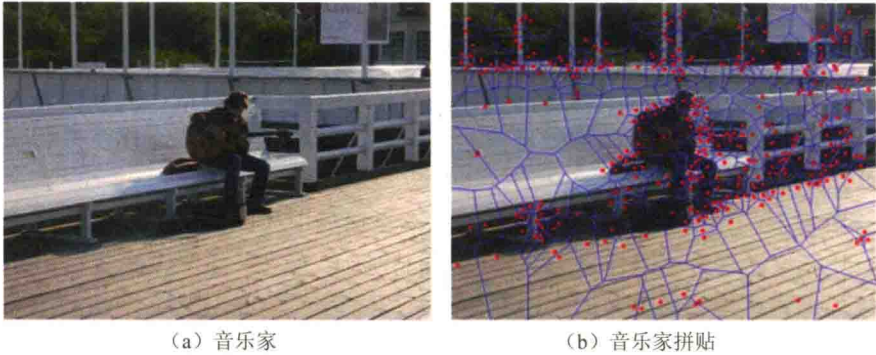


图 0.1 音乐家图像的沃罗诺伊镶嵌

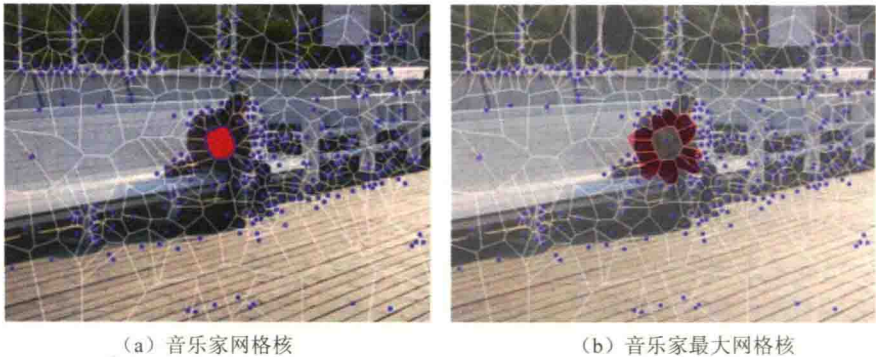


图 0.2 音乐家图像的最大核聚类

每个图像拼贴多边形是内部和顶点像素的凸包。一组图像点的凸包是该组点的最小凸集。一组图像点  $A$  是凸集，只要集合  $A$  中任意两点之间每条直线段上的所有点都包含在集合中。换句话说，知识发现是计算机视觉的核心。知识和对数字图像的理解可用于计算机视觉系统的设计。在视觉系统设计中，需要理解数字图像的组成和结构以及用于分析捕获图像的方法。

本书的重点是对光栅图像的研究。本书的续集将集中在由点 (矢量)，线和曲线组成的矢量图像。每幅光栅图像基本由像素 (典型的像素称为网点或网格生成点)，边缘 (共同的、平行的、交叉的、凸的、凹的、直的、弯曲的、连通的、不连通的)，角度 (矢量角度、矢量之间的角度、像素角度)，图像几何 (沃罗诺伊区域[139]、德劳内三角剖分[138])，颜色，形状和纹理组成。许多计算机视觉和场景分析中的问题可通过找到某些隐藏或未观察到的图像变量核结构的最可能值来解决 (参见 P. Kohi 和 P. H. S. Torr [94])。这样的结构和变量



包括像素的拓扑邻域、像素集合的凸包、图像结构的接近度（和分离度）、像素梯度分布以及描述捕获场景元素的特征矢量。

其他计算机视觉问题包括图像匹配、特征选择、最佳分类器设计、图像区域测量、兴趣点识别、轮廓组合、分割、配准、匹配、识别、图像聚类、模式聚类，见 F. Escolono、P. Suau、B. Bonev [45]和 N. Paragios、Y. Chen、O. Faugeras [136]，地标点形状匹配、图像变形、形状梯度[136]，假着色、像素标记、边缘检测、几何结构检测、拓扑邻域检测、目标识别和图像模式识别。

在计算机视觉中，重点是检测数字图像中常见的基本几何结构和物体形状。这导致了对图像处理 and 图像分析基础知识以及对矢量空间和图像计算几何观点的研究。图像处理的基础包括彩色空间、滤波、边缘检测、空间描述和图像纹理。数字图像是欧几里得空间（2-D 和 3-D）的示例。因此，数字图像的矢量空间观点是其基本特征的自然结果。**数字图像结构**本质上是几何或视觉拓扑结构。图像结构的示例包括图像区域、线段、生成点（例如，洛韦关键点）、像素组、像素的邻域、半空间、像素凸集和图像像素集合的凸包。例如，这种结构可以根据最接近选定点的图像区域或具有指定直径范围的图像区域来查看。一个图像区域是数字图像内部的一组图像点（像素）。任何图像区域的直径是该区域中一对点之间的最大距离。这样的结构也可以在由 2-D 和 3-D 图像中选定点之间连接的线段所形成的三角形区域中找到。

这种结构也常见于 2-D 和 3-D 图像中由封闭的半空间所形成的点集凸包里、或如 G. M. 齐格勒所称的**多面体**[217]里。**图像半空间**是一条线上方或下方的所有点的集合。在所有这三种情况下都能获得对数字化图像的区域视图。更多有关多面体的信息参见附录 B.14。

每个图像区域都有一个形状。一些区域的形状比其他区域的更有趣。有趣的图像区域形状包含感兴趣的目标。这些图像的区域视图导致各种形式的图像分割，它们在识别图像中的目标时具有实用价值。此外，检测感兴趣图像区域的形状可帮助发现图像模式，它们超越了图像处理中对纹理元的研究。纹理元是一个由像素组表示的区域。更多有关形状的信息参见附录 B.18 关于形状和形状边界的介绍。

图像分析侧重于各种数字图像测量（例如，像素大小、像素邻接、像素特征值、像素邻域、像素梯度、图像邻域的接近度）。图像分析中三种基于区域的标准方法是等值线阈值化（二值化图像）、分水岭分割（使用从前景像素到背景区域的距离图计算）和非最大抑制（通过抑制所有比周围像素更不可能的像素来找到局部最大值）[208]。

在图像分析中，前景和背景像素具有不同的邻接（邻域），见 T. Aberra [3]。有三种基本的邻域类型：即罗森菲尔德邻接邻域[101, 169]、豪斯多夫邻域[73, 74]，以及 J. F. Peters [140]和 C. J. Henry [75, 76]的描述性邻域。使用不同的几何形状，像素的邻接邻域由与给定像素相邻的像素限定。一个像素  $p$  的**罗森菲尔德邻接邻域**是一组与  $p$  邻接的像素。邻接邻域常用于数字图像中的边缘检测。

一个像素  $p$  的**豪斯多夫邻域**由与  $p$  的距离小于一个正数  $r$ （称为邻域半径）的所有像素构成。一个像素  $p$  的**描述性邻域**（由  $N(\text{img}(x, y), r)$  表示）由一组具有特征矢量的像素构成，这些特征矢量与描述  $\text{img}(x, y)$  的特征矢量匹配或相似，且在规定的  $r$  半径内。

与邻接邻域不同，描述性邻域可以拥有其中的孔，即具有与邻域中心不匹配和不属于邻域的特征矢量的像素。其他类型的描述性邻域在[140, 1.16 节, pp.29-34]中介绍。

本书的章节源于过去若干年中自己教授过的本科生计算机视觉课程的笔记。本书中的许多主题源于自己与一些研究生和其他人员的讨论和交流,尤其是 S. Ramanna(很多形状,特别是在水晶中)、Anna Di Concilio(邻近性、区域自由几何和如图 0.3 的海景形状)、Clara Guadagni(花神经结构)、Arturo Tozzi(Borsuk-Ulam 定理的见解和 Gibson 形状, Avenarius 形状)、Romy Tozzi(记得  $8, \infty$ )、Zdzisław Pawlak(波兰乡村绘画中的形状)、Lech Polkowski(那些纯粹的,拓扑的和粗糙集结构)、Piotr Artiemjew(蜻蜓翅膀)、Giangiacomo Gerla(那些 UNISA 庭院三角形和空间区域的指示(点))、Gerald Beer(Som Naimpally's 生活中的时刻)、Guiseppe Di Maio(关于邻近性的见解)、Somashekhkar (Som) A. Naimpally(拓扑结构)、Chris Henry(色彩空间,颜色形状集)、Macek Borkowski(3-D 空间观点)、Homa Fashandi、Dan Lockery, Irakli Dochviri, Ebubekir İnan(接近关系和接近组)、Mehmet Ali Öztürk(优美的代数结构)、Mustafa Uçkun、Nick Friesen(住宅的形状)、Özlem Umdu、Doungrat Chitcharoen、Çenker Sandoz(德劳内三角化)、Surabi Tiwari(许多类别)、Kyle Fedoruk(计算机视觉应用:斯巴鲁 EyeSight®)、Amir H. Meghdadi、Shabnam Shahfar、Andrew Skowron(Banacha 的接近性)、Alexander Yurkin、Marcin Wolksi(束或滑轮)、Piotr Wasilewski、Leon Schilmoeler、Jerzy W. Grzymala-Busse(关于粗糙集和 LATEX 提示的见解)、Zbigniew Suraj(Petri 网)、Jarosław Stepaniuk、Witold Pedrycz、Robert Thomas(倾斜的形状)、Marković G. oko(多形体)、Mirosław Pawlak、Pradeepa Yahampath、Gabriel Thomas、Anthony (Tony) Szturm、Sankar K. Pal、Dean McNeill、Guiseppe (Joe) Lo Vetri、Witold Kinsner、Ken Ferens、David Schmidt(集合论)、William Hankley(基于时间的规范)、Jack Lange(黑板拓扑涂鸦)、Irving Sussman(定理和证明中的金块)和 Brian Peters(在墙壁上稍纵即逝的几何形状一瞥)。





图 0.3 沿意大利维耶特里海岸线的海景形状

我们系的技术人员非常乐于助人,尤其是 Mount-First Ng, Ken Biegun, Guy Jonatschick 和 Sinisa Janjic。

我的很多学生提出了有关本书主题的重要建议,特别是 Drew Barclay、Braden Cross、Binglin Li、Randima Hettiarachchi、Enoch A-iyeh、Chidoteremndu(Chido)Chinonyelum Uchime、D. Villar、K. Marcynuk、Muhammad Zubair Ahmad 和 Armina Ebrahimi。



各章的问题已被分了类。以  标记的问题可以迅速地回答。以  标记的问题有可能需要用喝一杯茶或咖啡的时间来完成。解决剩下的问题所需的时间各不相同。

詹姆斯 F.彼得斯  
温尼伯，加拿大

# 目 录

第 1 章 通往机器视觉的基础知识	1
1.1 什么是计算机视觉	1
1.2 分而治之的方法	1
1.3 覆盖在图像上的沃罗诺伊图	4
1.4 计算几何简介	6
1.5 数字图像的框架	8
1.6 数字视觉空间	12
1.7 创建你自己的图像	12
1.8 随机生成图像	14
1.9 显示图像的方法	16
1.10 数字图像的格式	17
1.11 图像数据类型	18
1.12 彩色图像	21
1.12.1 彩色空间	21
1.12.2 彩色通道	21
1.13 彩色查找表	24
1.14 图像几何初步	26
1.15 访问和修改图像像素值	29
1.16 彩色、灰度和二值图像	31
1.17 像素的罗森菲尔德 8-邻域	32
1.18 距离：欧几里得和出租车测度	35
1.19 假彩色：点彩派绘画	37
1.19.1 假彩色 RGB 图像模式	37
1.19.2 假彩色灰度图像模式	39
1.20 数字图像上的矢量空间	41
1.20.1 点积	41
1.20.2 图像梯度	42
1.21 相机看见什么：智能系统视图	43
1.21.1 相机视觉系统中的智能系统方法	43
1.21.2 相机感知的场景彩色	44
1.22 图像几何：图像上的沃罗诺伊和德劳内网格	47
1.22.1 汽车图像上的沃罗诺伊网格	47
1.22.2 沃罗诺伊图像的子网格揭示了什么信息	50
1.23 神经结构	50

1.23.1	汽车图像上的德劳内网格 .....	54
1.23.2	在汽车图像上结合沃罗诺伊和德劳内网格 .....	55
1.24	视频帧网格覆盖 .....	57
1.24.1	离线视频帧处理 .....	57
1.24.2	实时视频处理 .....	60
<b>第 2 章</b>	<b>像素加工 .....</b>	<b>63</b>
2.1	图像元素 .....	63
2.2	分离彩色图像通道 .....	65
2.3	彩色向灰度的转换 .....	66
2.4	对像素强度的代数操作 .....	67
2.5	用边缘像素选择解释像素选择 .....	72
2.6	基于函数修改图像像素值 .....	78
2.7	图像的逻辑操作 .....	81
2.7.1	像素强度的补和逻辑非 .....	81
2.7.2	成对二值图像的 XOR 操作 .....	82
2.8	从背景中提取前景 .....	84
2.9	阈值化彩色通道的合并 .....	85
2.10	增强图像的对比度 .....	87
2.11	伽马变换 .....	88
2.12	伽马校正 .....	89
<b>第 3 章</b>	<b>可视化像素强度分布 .....</b>	<b>91</b>
3.1	直方图和绘制 .....	93
3.1.1	直方图 .....	94
3.1.2	茎干图 .....	95
3.1.3	绘制 .....	96
3.1.4	表面绘制 .....	96
3.1.5	线框表面绘制 .....	98
3.1.6	轮廓绘制 .....	98
3.2	等值线 .....	98
3.3	彩色直方图 .....	100
3.4	自适应阈值化 .....	102
3.5	对比度拉伸 .....	103
3.6	直方图匹配 .....	106
<b>第 4 章</b>	<b>线性滤波 .....</b>	<b>108</b>
4.1	图像滤波的重要性 .....	108
4.2	滤波器核 .....	109
4.3	线性滤波器实验 .....	110
4.4	线性卷积滤波 .....	111
4.5	选取感兴趣区域 .....	112



4.6	给图像加噪声	114
4.7	均值滤波	116
4.8	中值滤波	117
4.9	排序滤波	118
4.10	正态分布滤波	120
<b>第5章</b>	<b>边缘、线、角点、高斯核与沃罗诺伊网格</b>	<b>121</b>
5.1	线性函数	121
5.2	边缘检测	123
5.3	双精度拉普拉斯滤波器	126
5.4	增强数字图像边缘	127
5.5	高斯核	128
5.6	高斯滤波器	130
5.7	高斯滤波器核图像恢复	131
5.8	高斯-拉普拉斯滤波器图像增强	132
5.9	零交叉边缘滤波器图像增强	132
5.10	各向异性与各向同性边缘检测	134
5.11	在数字图像中检测边缘核线	136
5.12	检测图像角点	138
5.13	基于图像角点的沃罗诺伊网格重访	140
5.13.1	沃罗诺伊镶嵌细节	140
5.13.2	沃罗诺伊多边形的位置	140
5.14	构建基于角点的沃罗诺伊网格的步骤	142
5.15	网格生成器集中的极端图像角点	142
5.16	具有极端角点图像上的沃罗诺伊网格	144
5.17	孤立图像边缘的图像梯度方法	145
5.18	角点、边缘和沃罗诺伊网格	147
<b>第6章</b>	<b>德劳内网格分割</b>	<b>149</b>
6.1	德劳内三角化生成三角网格	149
6.2	三角形外接圆	151
6.3	在图像上构建基于角点的德劳内网格	151
6.4	基于质心的德劳内图像网格	152
6.4.1	寻找图像质心	153
6.4.2	寻找图像质心的德劳内网格	154
6.4.3	寻找图像质心的沃罗诺伊网格	154
6.4.4	寻找叠加在德劳内网格上的图像质心沃罗诺伊网格	155
<b>第7章</b>	<b>视频处理、实时和离线视频分析介绍</b>	<b>157</b>
7.1	视频处理基础	158
7.1.1	帧图像点处理	158
7.1.2	图像采集	158

7.1.3	斑块.....	159
7.1.4	帧拼贴和帧几何.....	159
7.2	视频帧的沃罗诺伊拼贴.....	160
7.3	在视频帧中检测目标形状.....	160
7.4	测量目标的形状相似性和沃罗诺伊视觉外壳.....	161
7.5	最大核聚类.....	163
7.6	问题.....	166
7.7	形状距离.....	171
7.8	边缘集的权函数.....	172
7.9	最大边缘集.....	173
7.9.1	粗糙轮廓边缘集.....	173
7.9.2	最大核聚类连通网格区域.....	175
<b>第 8 章</b>	<b>洛韦关键点、最大核聚类、轮廓和形状</b> .....	<b>177</b>
8.1	图像分析.....	178
8.2	场景分析.....	179
8.3	像素边缘增强.....	180
8.4	数字图像的裁剪和稀疏表示.....	183
8.5	形状理论和 2-D 图像目标的形状：面向图像目标形状检测.....	185
8.6	图像像素梯度的朝向和强度.....	186
8.7	高斯差.....	187
8.8	图像关键点：洛韦 SIFT 方法.....	189
8.9	应用：图像网格核的关键点边界.....	191
8.10	超（外）核轮廓.....	193
8.11	最大核聚类轮廓形状的质量.....	194
8.12	粗糙 S2P 和 S3P（2 级和 3 级）最大核聚类轮廓.....	194
8.13	关键点数量的实验.....	196
8.14	双重最大核聚类上的粗糙周边.....	199
8.15	图像最大核聚类区域的莱利熵.....	199
8.16	问题.....	202
<b>第 9 章</b>	<b>后记：形状适合计算机视觉环境的地方</b> .....	<b>204</b>
9.1	自然场景中的目标形状.....	204
9.2	形状估计.....	207
<b>附录 A</b>	<b>Matlab 和 Mathematica 编程</b> .....	<b>210</b>
A.1	第 1 章的程序.....	210
A.1.1	数字图像角点.....	210
A.1.2	沃罗诺伊镶嵌算法的实现.....	211
A.1.3	德劳内镶嵌算法的实现.....	213
A.1.4	沃罗诺伊和德劳内镶嵌结合算法的实现.....	216
A.1.5	第 1 章的离线视频处理程序.....	218

A.1.6	第1章的实时视频处理程序 .....	219
A.2	第2章的程序 .....	221
A.2.1	数字图像像素 .....	221
A.2.2	彩色图像通道 .....	223
A.2.3	彩色向灰度的转换 .....	224
A.2.4	对像素强度的代数操作 .....	225
A.2.5	选择和显示边缘像素的彩色像素强度 .....	228
A.2.6	基于函数的像素值修改 .....	231
A.2.7	对图像的逻辑操作 .....	233
A.3	第3章的程序 .....	234
A.3.1	像素强度直方图(分档) .....	234
A.3.2	像素强度分布 .....	235
A.3.3	像素强度等值线 .....	237
A.4	第4章的程序 .....	238
A.5	第5章的程序 .....	238
A.5.1	1-D 高斯核绘图 .....	238
A.5.2	高斯核实验 .....	240
A.5.3	2-D 高斯核绘图 .....	241
A.5.4	高斯平滑图像 .....	242
A.5.5	图像恢复 .....	243
A.5.6	图像角点 .....	245
A.5.7	具有和没有图像角点的沃罗诺伊网格 .....	247
A.6	第6章的程序 .....	248
A.6.1	查找2-D和3-D图像质心 .....	248
A.6.2	另一种查找图像质心的方法 .....	249
A.6.3	查找图像质心的德劳内网格 .....	250
A.6.4	查找图像质心的沃罗诺伊网格 .....	251
A.6.5	查找覆盖在德劳内网格上的图像质心沃罗诺伊网格 .....	252
A.7	第7章的程序 .....	253
A.7.1	沃罗诺伊镶嵌视频帧中的边缘集测量 .....	253
A.8	第8章的程序 .....	257
A.8.1	高斯金字塔方案 .....	257
A.8.2	小波金字塔方案 .....	258
A.8.3	像素边缘强度 .....	259
A.8.4	绘制反正切值 .....	260
A.8.5	像素几何: 梯度方向和梯度幅度 .....	261
A.8.6	高斯差图像 .....	263
A.8.7	图像关键点和沃罗诺伊网格 .....	263



---

附录 B 词汇表 .....	266
B.1 数字与符号 .....	266
B.2 B .....	268
B.3 C .....	271
B.4 D .....	271
B.5 F .....	273
B.6 G .....	274
B.7 H .....	274
B.8 J .....	274
B.9 K .....	279
B.10 L .....	281
B.11 M .....	284
B.12 N .....	284
B.13 P .....	284
B.14 Q .....	285
B.15 S .....	285
B.16 T .....	287
B.17 W .....	291
B.18 X .....	292
B.19 Z .....	294
参考文献 .....	296
主题索引 .....	306

# 第 1 章 通往机器视觉的基础知识

## 1.1 什么是计算机视觉

计算机视觉的主要目的是基于数码相机拍摄的图像内容重建和解释自然场景[186]。自然场景是视场的一部分，视场是由人类视觉感知或基于光学传感器阵列而获取的。

基于光学传感器阵列的自然场景或者是用相机获取的单幅图像，或者是用摄像机（如网络摄像机）获取的视频帧图像。

每个图像场景的基本内容包括**像素**（相邻的、不相邻的）、**边缘**（共同的、平行的、交叉的、凸的、凹的、直的、弯曲的、连通的、不连通的）、**角度**（矢量、矢量之间、像素）、**图像几何**（沃罗诺伊区域[139]，德劳内三角剖分[138]）、**颜色**、**形状**和**纹理**等。

## 1.2 分而治之的方法

通过用已知几何形状，如三角形（德劳内三角剖分方法）和多边形（沃罗诺伊图方法），来拼贴（镶嵌）场景图像更容易重建和解释自然场景。这是一个分而治之的方法。在计算机视觉中，这类方法的例子包括如下。

**形状检测**：使用德劳内三角剖分的视频帧形状检测见 C. P. Yung、G. P.-T. Choi、K. Chen 和 L.M. Lui [211]（参见图 1.1）。

**剪影**：使用轮廓查找极线以对摄像机网络进行校准是 G. Ben-Artzi、T. Halperin、M. Werman 和 S. Peleg 使用的方法[14]。这里的基本目标是实现双目视觉并确定场景中一个 3-D 物体点在一对 2-D 图像上的位置（单个相机所见）。这里使用**极点**（在成像平面上连接光学中心的直线的交点）以从一对 2-D 图像中提取 3-D 目标。光学中心之间的连线称为**基线**。**极平面**是由 3-D 点  $m$  与光学中心  $C$  和  $C'$  定义的平面。参见图 1.2 中的一对极点和极线。



图 1.1 [211]中三角化的视频帧

**视频点画**：点画使用点集、基本形状和彩色来绘制/渲染图像。视频点画的核心技术是视频帧的**沃罗诺伊镶嵌**。这是 T. Houit 和 F. Nielsen 的方法[83]。该文包含对叠加在视频帧图像上沃罗诺伊图的一个很好的介绍（见[83, 2 节, pp.2-3]）。沃罗诺伊图在图像分割中很有用。它导致所谓的 **Dirichlet 镶嵌图像**，引领一种新形式的  **$k$ -均值图像区域的聚类**（参见图 1.3，了解沃罗诺伊分割方法中的方法）。这种形式的**图像分割**使用聚类中心逼近图像聚类。这是 R. Hettiarachchi 和 J. F. Peters 使用的方法[78]。沃罗诺伊流形由 J. F. Peters 和 C. Guadagni 在[144]中介绍。**流形**是一个局部欧几里得拓扑空间，即

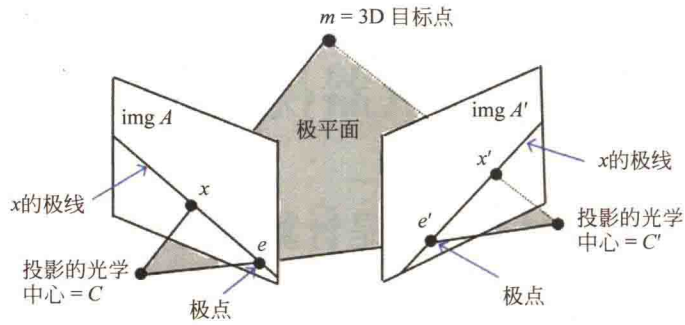


图 1.2 极点和极线

在流形的每个点周围都有一个开放的邻域。具有拓扑结构 $\tau$ 在其上的非空集合 $X$ 就是一个拓扑空间。在一个非空开集 $X$ 上的开集集合 $\tau$ 是 $X$ 上的拓扑，只要它具有某些属性（参见附录 B.9 节对于开集和拓扑结构的定义）。一个开集是空间 $X$ 中点 $A$ 的非空集，包含所有足够接近 $A$ 的点但不包括它的边界点。

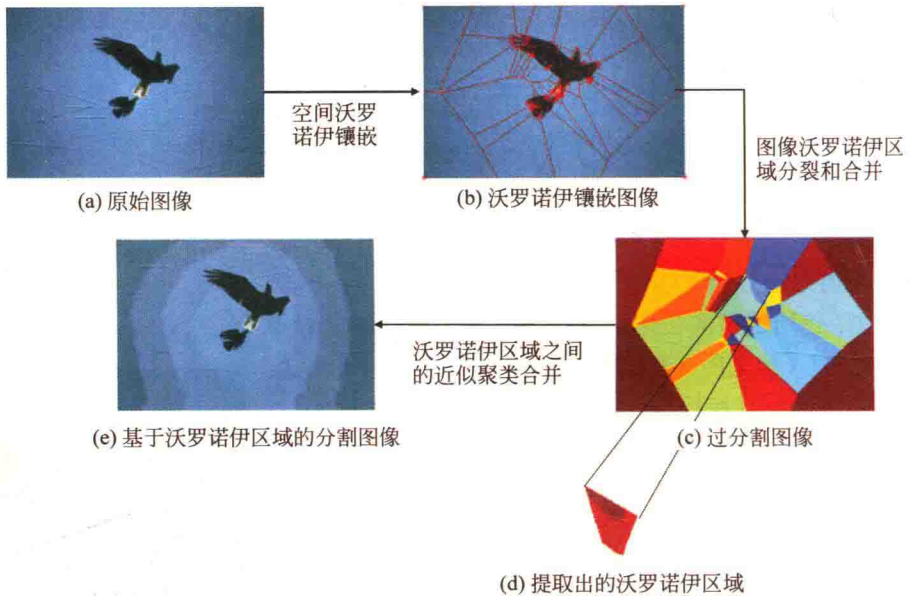


图 1.3 [78]中沃罗诺伊分割方法的步骤

### 例 1.1 开集示例

苹果果肉：没有外皮的苹果。

鸡蛋内部：没有蛋壳的蛋黄。

无墙房间：没有墙壁的房间。

开子图像：不包括其边界点的子图像。图 1.4 所示为一幅处在 $4 \times 5$ 的数字图像 $\text{Img}$ 中的 $2 \times 3$ 子图像 $A$ 。集合 $A$ 是开集，因为它仅包含黑色方块■且不包含沿其灰色像素■表示的边界。