



装备科技译著出版基金



高新科技译丛
通信技术系列



Springer

Video Segmentation and Its Applications

视频分割及其应用

【澳】King Ngi Ngan Hongliang Li 编著
郑丽颖 译 赵琳 审校



国防工业出版社
National Defense Industry Press



装备科技译著出版基金

视频分割及其应用

Video Segmentation and Its Applications

[澳] King Ngi Ngan Hongliang Li 编著

郑丽颖 译 赵琳 审校

国防工业出版社

· 北京 ·

著作权合同登记 图字:军—2013—067 号

图书在版编目(CIP)数据

视频分割及其应用 / (澳)颜庆义,李宏亮编著;郑丽颖译.

—北京:国防工业出版社,2014.4

书名原文:Video segmentation and its applications

ISBN 978-7-118-09216-5

I. ①视... II. ①颜... ②李... ③郑... III. ①视频信号-
图象分割 IV. ①TN941.1

中国版本图书馆CIP数据核字(2014)第044136号

Translation from the English language edition: Video Segmentation and Its Applications By king
Ngi Ngan and Hong Liang Li.

Copy right © 2011 by Springer Science + Business Media, LLC 233 Spring Street, New York
10013, USA

All rights reserved.

本书简体中文版由 Springer Science + Business Media, LLC 授权国防工业出版社独家出版发
行。版权所有,侵权必究。

※

国防工业出版社出版发行

(北京市海淀区紫竹院南路23号 邮政编码100048)

北京嘉恒彩色印刷有限责任公司印刷

新华书店经售

*

开本 710 × 1000 1/16 印张 10½ 字数 168 千字

2014年4月第1版第1次印刷 印数 1—2000册 定价 79.00 元

(本书如有印装错误,我社负责调换)

国防书店:(010)88540777

发行邮购:(010)88540776

发行传真:(010)88540755

发行业务:(010)88540717

前 言

作为视觉信息提取的关键技术,视频分割在数字视频处理、模式识别以及计算机视觉等领域扮演重要角色。视频技术在安全与监控、银行业务监视、视频会议以及个人娱乐等领域的广泛应用得益于视频分割技术的进步。

近 40 年来,视频分割技术有了长足的发展和进步,涌现出大量成果。目前,图像和视频分割仍然是一个热点研究领域,且成绩斐然。因此,亟需将视频分割及其应用的最新研究成果集结成著,本书正是在此需求下编著的。

本书的目的是呈献视频分割及视频分析技术在理论和应用方面的最新成果。本书概括了视频分割的新兴方法以及计算机视觉和视频分析中的一些正在研究的、有发展前景的方法。本书内容不仅包括视频特征提取、语义视频目标分割等图像/视频处理的理论基础和算法,同时也全面描述了视频分割的实际应用技术,填补了市场空白。

本书将为研究人员和相关从业人员全面解读最新的视频分割技术,同时指出可能的应用方法、列举成功的应用案例,可作为从事视频分析、计算机视觉、模式识别、图像和视频处理、人工智能等视频分割领域中各方向研究人员、工程师以及研究生的工具书。

第 1 章首先介绍了基于图论的分割算法、基于密度估计的分割算法以及时域分割算法的发展现状。然后全面介绍了图像/视频分割技术,并讨论了其最新进展。同时本章也指出了图像/视频分割技术所面临的挑战,给出了今后的发展方向。

第 2 章给出了基于本征结构特性的目标分割算法。根据所选取的颜色样本计算协方差矩阵,然后将该协方差矩阵作本征分解,得到本征子空间。利用独立本征模糊 C-均值和耦合本征模糊 C-均值,综合考虑期望颜色的信号和噪声子空间投影,实现了有效的彩色目标分割。根据所提出的算法,利用本征子空间投影,可成功分割彩色目标。

第3章阐述了语义目标分割问题。语义目标分割的目的是将视频帧的每一个像素都标记为具有语义意义的目标类。本章首先讨论了语义目标分割各步骤所涉及的技术及难点问题。分别使用生成方法和判别方法中的代表性模型,即条件随机场和主题模型,完成语义目标分割任务。

第4章总结并介绍了基于学习的视频场景分割技术。根据应用设置以及学习目标,阐述了两大类方法,即泛型方法和基于类型的分析技术。同时,为了实现视频场景分析,本章也指出了视频内容分析与检索中所面临的技术挑战。

第5章描述了有代表性的、最新的多视点图像分割和目标跟踪方法。提出了既适用于无遮挡、也适用于有遮挡人体目标的初始帧深度分割和基于特征的多视点视频跟踪算法。

第6章讨论了视频分割的一些实际应用,包括:医学成像、计算机辅助外科手术、机器视觉、目标识别、监控、基于内容的浏览以及增强现实等。对于给定的某一具体应用,期望的分割质量依赖于粒度水平、目标性状描述精度要求和时间一致性要求。虽然在一般任务的鲁棒和全自动分割方面仍然存在巨大挑战,但是可以通过使用合适的视觉注意模型找到可靠的解决办法。

King Ngai Ngan
中国香港特别行政区

Hongliang Li
中国四川省成都市

2011年1月

目 录

第 1 章 图像/视频分割:现状、趋势、挑战	1
1.1 简介	1
1.2 最新分割技术	4
1.3 发展趋势	11
1.4 技术挑战	18
1.5 小结	19
参考文献	19
第 2 章 基于本征子空间投影的图像分割技术	24
2.1 概述	24
2.2 基于颜色本征结构的目标分割算法	25
2.3 基于自适应本征子空间的彩色目标分割	37
2.4 基于本征子空间投影和模糊 C - 均值的彩色目标分割	43
2.5 小结	53
参考文献	54
第 3 章 语义目标分割	57
3.1 引言	57
3.2 局部视觉线索	61
3.3 基于判别方法的目标分割技术	65
3.4 基于主题模型的目标分割技术	70
3.5 视频目标分割	77
3.6 小结	81
参考文献	81

第4章 视频场景分析：机器学习前瞻	86
4.1 引言	86
4.2 场景语义描述	87
4.3 视频场景分析的通用技术	89
4.4 基于视频类型的内容分析技术	104
4.5 小结	112
参考文献	112
第5章 多视点图像分割和视频跟踪	118
5.1 引言	118
5.2 多视点图像分割	120
5.3 多视点视频跟踪	131
5.4 小结	140
参考文献	140
第6章 视频分割应用	144
6.1 引言	144
6.2 视频分割的发展趋势	145
6.3 基于目标的监控分析	147
6.4 基于目标的可伸缩视频编码	148
6.5 3D 重建	152
6.6 增强现实	153
6.7 小结	154
参考文献	154
附录 缩略语	157

第 1 章

图像/视频分割:现状、趋势、挑战

摘要:分割在数字媒体处理、模式识别、计算机视觉领域中扮演着重要角色。许多实际应用领域,例如:图像解译、视频分析与理解、视频摘要与索引以及数字娱乐,都涉及图像/视频分割问题。最近 20 年,图像/视频分割已成为一个基础问题,对新的模式识别算法及其应用具有重大影响。

本章的主要内容包括:①总结分割技术的研究现状,包括基于图的分割算法、基于密度估计的分割算法以及时域分割算法;②全面介绍图像/视频分割技术,同时讨论该领域的最新进展;③指出图像/视频分割领域面临的挑战以及未来的发展方向。

1.1 简介

古人云:“千言不如一画”。只用一张静态图片就可以传递复杂的语义信息。你是否想过这样一些问题:我们在看一张图片的时候,眼睛是如何从场景中发现感兴趣目标的?大脑是如何理解这一场景的?场景识别过程涉及多少活动?通过语义内容处理可能会找到这些问题的答案,并为场景理解提供有意义的线索^[1]。

通过内容相关服务可以检测语义目标(包含属性集合的有意义实体),并且能够为用户提供灵活的基于内容的访问和操作方式,例如:视频数据库快速索引、高级视频编辑与合成、感兴趣区域的高效编码^[2]。近年来,人们对基于内容的视频数据应用越来越感兴趣,这些应用包括:视频检索与浏览、视频摘要、视频事件分析以及视频编辑。大量多媒体数据的高效访问变得越来越重要。然而,从图像/视频中成功获取语义内容仍是计算机视觉和模式识别中的一项艰巨任务。

为了理解场景内容,需要知道这一内容的基本构成。答案也许就是语义目标,即数据项及其语义内容。语义目标可能由一组显式描述数据项暗含语义的^[4]灵活的元属性构成。每个语义目标都应该明确规定目标和实景之间的联系。因此,在图像理解之前,一个关键的步骤是将图像/视频分割成几个组成部分。

一般地,分割是指将数据划分为具有相似特性的可能子集的过程。分割是语义内容提取的关键技术,在数字多媒体处理、模式识别和计算机视觉方面扮演重要角色。在许多领域中,图像分割是一种面向应用的技术。下面,给出图像/视频分割的一些应用。

(1) 目标识别:在目标识别中,分割作为关键步骤,将图像分成一些具有某种一致视觉特征的小区域,然后将这些区域用于收集并检测目标^[5]。识别中的两个重要任务——特征提取和模型匹配,极大地依赖于图像分割质量。当图像被分割成几个均匀亮度区域之后,由于这些区域是包含丰富信息的描述子,并且对于小的光照和视角变化不敏感,因此可用于推导类模型^[6]。

(2) 视频监视:在这一领域中,可以将目标分割成几部分,通过跟踪运动目标沿时间轴的演变情况提高遮挡目标的跟踪鲁棒性^[7]。分割掩模能够预测并鉴别入侵者或异常情况、展现他们的行为、并快速决定何时向安全中心发出警报。

(3) 视频索引:利用段标注在媒体片段上执行视频索引操作^[8,9]。将与查询目标有关的媒体段排序列表返回到用户端,这一技术已被用于内容分类、表示或者理解。

(4) 数据压缩:数据压缩允许利用合适的编码算法对目标进行独立操作,有助于改善主观质量。采用分割技术将视频序列帧分为任意形状的、有意义的语义目标,通过为这些目标区域分配更多码位的方法,来减少低比特率编码所带来的视觉伪影^[10]。

(5) 计算机视觉:从二维(2D)输入图像或者视频序列分割得到的目标可用于创建三维(3D)场景。例如,基于图像过分割的立体透视图绘制技术,由于匹配的是所有分割区域而不是单个像素,因此,初始匹配值对图像噪声和亮度偏置更具鲁棒性^[11]。

(6) 视频电话:为了获得高感知质量,视频电话以较高的质量对感兴趣区域进行编码。因为前景区域更重要,因此,采用较高的量化等级、更少的码位编码背景区域^[12]。

(7) 数字娱乐:数字娱乐(如:视频抠图和卡通风格转换)利用分割出的目标产生奇幻效果、或者是把这些目标放置于虚拟场景或游戏中。

除此之外,还有一些其他应用,例如,医学诊断、远程教学、工业检测、环境监测或者分割目标的元数据关联等。

在图像/视频处理中,分割已经成为连接原始图像数据和语义内容的有效方式。为了满足未来基于内容的多媒体应用,在过去几十年中,越来越多的研究者探寻从多媒体数据中分割任意目标的有效途径。解决分割问题的方法有很多,可以按照不同标准进行分类。

(1) 基于数据的分类方式:按照原始数据类型,分割可分为图像(例如,自然图像、医学图像或者遥感图像等)分割、视频分割、音频分割以及文本分割。这些分割模式适用于不同的情况。例如,文本分割可用于提取电影字幕,或者将文档分割成感兴趣区域。本章仅关注图像和视频分割。

(2) 基于互动的分类方式:主要有两类——监督模式和无监督模式。监督方法要求用户参与分割过程,允许用户指示跨时空前景区域。由于通过在图像上选取适当的训练数据能够获得目标先验知识,因此,监督方法的性能优于自动方法。无监督方法是指无需关于被分割目标的任何上下文知识,以全自动方式实现目标分割。无监督分割已经成为大量实时应用(例如,视频监视和监控)的关键技术。

(3) 基于特征的分类方式:在图像/视频分割中,特征提取扮演着重要角色。按照所选择的特征空间,分割可以分为基于颜色的分割方法、基于纹理的分割方法、基于亮度的分割方法、基于形状的分割方法、基于运动的分割方法。经常利用这些特征来估计区域性质。例如,对于颜色分割,分组决策依赖于相邻像素的颜色距离;对于运动的分割,关键问题是根据运动线索找到视频中的独立运动目标。

(4) 基于推理的分类方式:分割可分为两种信息传递模式——即自底向上(Bottom-up)的分割和自上向下(Top-down)的分割。第一类方法根据低级视觉特征(如:颜色、纹理、亮度等)而不是感兴趣目标的高级知识实现分割,通常以无监督方式实现。为了得到先验分布,第二类方法通常需要人工标注的图像数据库。通过合并低级分割结果,第二类方法可以完成高级图像识别。

(5) 基于空间的分类方式:基于空间关系,可以分为空域分割和时域分割。第一类方法根据像素空间关系实现分割,第二类方则沿时间轴将帧序列分段。例如,为了分割出具有相似内容的视频帧,可以采用场景分析技术(如:视频剪辑、渐变、

加宽、缩放等)实现场景分割。

(6) 基于类的分类方式:为了从输入图像/视频中提取特定目标(如:人脸、人体目标、汽车或者建筑物),已经提出了许多分割方法。由于事先已经知道了目标类型,因此可以利用目标先验知识改善分割结果。例如,对于人脸分割,根据样本观测到的皮肤颜色分布对于脸部区域检测是非常重要的信息,它有助于获取脸部区域信息。

(7) 基于语义的分类方式:不同于根据纹理或者颜色特征提取一致、均匀区域的无语义分割方法,语义分割是指将图像分割成具有语义意义的区域的过程。

注意:现有的分割方法可能属于以上一个或者多个类别。当然,各个分割模式之间并没有明确的界限,因此,可通过合并不同模式,得到新的分割方法。例如,无监督过分割常常被用于自顶向下分割方法中,其中自顶向下技术可将过分割部分合并为一个语义目标。

由于图像分割是面向应用的,因此,很难按照统一的标准来衡量分割质量。这就意味着“什么是好的分割?”以及“如何区分好的和不好的分割?”等问题极大地依赖于应用背景。因此,许多研究者通过“好分割”假设来回答上述问题。例如,文献[13]指出一个“好分割”应该具有良好的连续状态。

本章的目的是从理论方面和实际应用方面回顾各种图像/视频分割方法。为此,我们仅关注图像/视频分割技术。但是读者有必要知道,最近10年,已经有许多文献从不同角度回顾了图像分割技术,例如,图像分割技术回顾^[14, 15]、超声图像分割概述^[16]、视频分割综述^[17]。本章不仅涉及本领域经典的、或者具有里程碑意义的方法,而且还包括可能会推进未来研究工作的一些发展趋势和技术挑战。

本章内容安排:1.2节回顾已有的图像/视频分割算法;1.3节讨论本领域的新兴方法,探索图像/视频技术的发展趋势;1.4节总结分割方法中的主要技术难点,给出图像/视频分割技术的研究前景。

1.2 最新分割技术

图像分割的历史(如空间域分割)可以追溯到19世纪。此后20年间,图像/视频分割技术发展迅速,涌现出了大量算法。本节主要关注有代表性的几种技术,并

简要描述几个图像分割技术开始发生变革的算法。读者可以在大量的文章和书籍中找到对这些方法的详细描述。

1.2.1 基于图论的分割技术

1.2.1.1 图割算法

1989年, Greig 介绍了用图割(Group Cut)技术求解二值图像最大后验概率估计(Maximum A Posteriori estimation, MAP)问题的方法^[18]。然而,不幸的是该方法直到最近几年才得到重视。Boykov 和 Jolly 做了第一个关于图割技术应用于图像处理方面的报告^[19]。他们将图割技术用于图像复原和交互式图像分割。给定标记目标子集和背景像素子集,该算法基于最小割(Minimum Cut)找到图像的全局最优分割,文献[20—22]也给出了类似算法。

给定一幅图像,基于最小割的分割方法首先创建图 $\mathcal{G} = \langle \mathcal{V}, \mathcal{E} \rangle$, 其中 $\mathcal{V}$ 为顶点集(如:像素集或者区域集), $\mathcal{E}$ 表示边集。特别地,两个终端顶点,即源点 s 和汇点 t , 用于连接这些普通顶点。图像中所有的顶点通过两类边来连接:两个相邻顶点之间的双向 n -连接, 顶点和终端结点之间的 t -连接。

定义割 $\mathcal{C}$ 为将顶点集划分为两个子集的二值划分,这两个子集分别对应于源点(前景)以及汇点(背景)。图割算法的目的是寻找具有全局最小代价的最优切割,即边权重的总和最小,这刚好等于图的最大流^[19]。通常,对于图像 $Z = \{z_i\}$, 图割代价函数可由如下能量函数表示:

$$E(Z) = \sum_{i \in \mathcal{V}^*} E_1(z_i) + \lambda_1 \sum_{\{i,j\} \in \mathcal{E}} E_2(z_i, z_j) \quad (1.1)$$

式中: E_1 和 E_2 分别为数据和平滑度代价函数。第一项 $E_1(z_i)$ 用于设置将像素指定为前景或者背景所对应的惩罚,反映了该像素接近前景或者背景的程度。文献[19]将这一项定义为“目标”和“背景”亮度分布直方图的负对数似然函数。通常,在交互式方法中,可以根据用户标记区域估算这两个分布。也就是说,以弱监督方式根据标记像素估计这两个先验分布。

第二项 $E_2(z_i, z_j)$ 通过设置顶点 z_i 和 z_j 之间的不连续性惩罚来度量它们之间的相似度。如果发现 z_i 和 z_j 之间存在明显边界,则 $E_2(z_i, z_j)$ 趋于零,表明在这两个相邻像素之间出现分割的可能性较大。通过局部亮度梯度或者其他正规化准则可以估算 E_2 。文献[19]采用 *ad-hoc* 函数设置边界惩罚。

利用最大流/最小割算法可以得到最大流问题的精确解^[23]。该算法试图找到一个新的增广路径,使得至少一条边达到饱和,并增大流量,使其接近最大值。当找不到新的增广路径时,就表明找到了最大流,即最小割。

图 1.1 给出了基于图割的分割方法示例。图 1.1(a) 为带有用户输入的原始图像,其中红色和蓝色笔画分别代表背景和目標。图 1.1(b) 为相应的图结构,除了普通顶点,它还包含两个端点顶点(目标(花朵)和背景(叶子))。图 1.1(c) 给出了分割结果,成功地从图像中分割出了花朵目标。

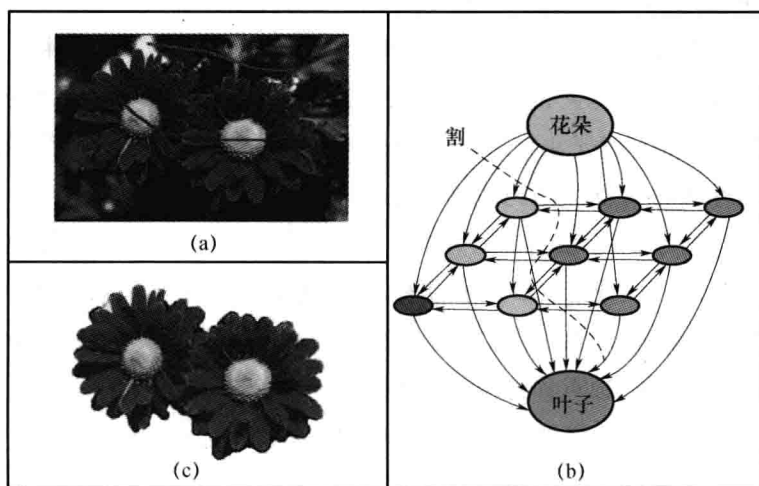


图 1.1 基于分割的分割方法示例

(a) 带有用户标记的原始花朵图像; (b) 包含两个端点顶点的图; (c) 分割结果。

1.2.1.2 随机游走

在 Wechsler 和 Kidode 解决纹理判别和边缘检测问题的早期文献中首次出现了计算机视觉中的随机游走概念^[25]。在早期文献中,随机游走算法将随机游走者置于像素上,检测它们首先到达哪些种子。然而,这样的计算方法并不实用。Grady 采用图论技术解决随机游走问题,首次将其成功应用于图像分割^[26, 27]。为了计算游走者将会最先到达哪个标记种子的期望概率,Grady 通过图建立了随机游走理论和电路理论(或者位势理论)之间的联系。

随机游走分割可以通过具有固定顶点数和边数的图来表示。每条边都被分配了一个实值权重,表示这个随机游走者跨越这条边的似然值。详细算法可概括为

以下四个步骤。

- (1) 初始化:得到具有已知标签的标记像素。
- (2) 映射:利用典型的高斯加权边将图像映射成图。
- (3) 优化:通过解 Dirichlet 问题,计算未标记顶点到达每一个标签的概率。
- (4) 分割:为各个顶点指定对应于最大势的类标签,得到最终分割结果。

图 1.2 给出了随机游走图像分割示例。图 1.2(a) 是标记了背景和目标的原始图像。图 1.2(b) 给出了该原始图像所对应的图,其中的实心 and 空心圆分别表示已知和未知标签。图 1.1(c) 给出了随机游走算法的分割结果。可见,花朵目标被成功分割出来了。

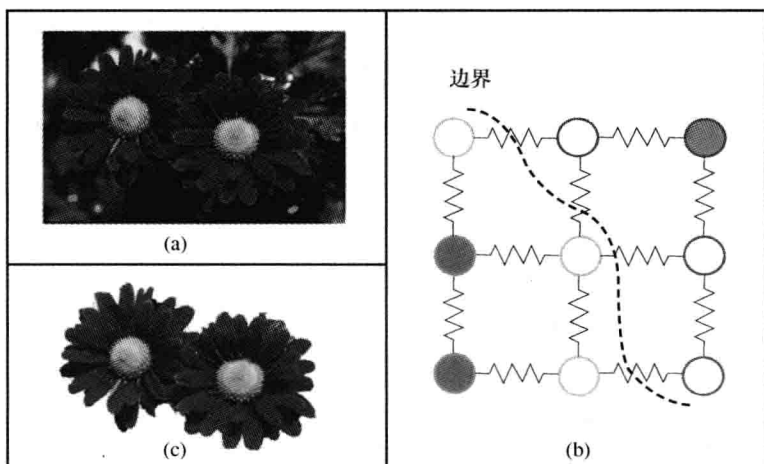


图 1.2 随机游走图像分割示例

(a) 带有用户标记的原始花朵图像; (b) 相应的图模型; (c) 分割结果。

1.2.1.3 谱聚类算法

谱聚类已经成为最流行的现代聚类算法之一。通过标准的线性代数方法可有效地解决谱聚类问题。相比于传统算法(如 k -均值聚类),谱聚类不仅易于实现,而且还有许多优于传统聚类方法的重要特点^[28]。

与图割方法相似,谱聚类也是基于加权图划分思想,可以用相似图 $G=(V,E)$ 表示,其中顶点集 $V=\{v_1,\dots,v_2\}$, E 代表边集,顶点 v_i 表示数据点,它通过边与相邻节点连接。谱聚类的主要工具是基于谱图理论的图拉普拉斯矩阵。最普通的谱聚类算法包括:非正则谱聚类和正则谱聚类^[29, 30]。

利用归一化分割准则, Shi 和 Malik 最先给出了经典的谱聚类分割方法^[29]。假设利用最小割可以将图 G 划分为成两个互斥集合 A 和 B 。定义如下分割:

$$cut(A, B) = \sum_{i \in A, j \in B} w(i, j) \quad (1.2)$$

通常, 通过将数据分成两部分来最小化式(1.2)会相对简单一些。然而, 在实际中这种做法并不能得到满意的划分。为了避免奇异划分, Shi 通过增加额外约束条件(如: 子集体积)重新定义割函数, 提出归一化割(Ncut)。定义如下:

$$Ncut(A, B) = \frac{cut(A, B)}{vol(A)} + \frac{cut(A, B)}{vol(B)} \quad (1.3)$$

其中, $vol(A) = \sum_{i \in A, j \in V} w(i, j)$ 是子集 A 的体积, 表明子集 A 中的节点与整个图顶点之间的关系。式(1.3)的最小化问题可以近似转化为本征值系统求解问题^[29]。

基于 Ncut 算法, 采用如下步骤将图像分割为若干区域:

- (1) 将图像映射为加权图 $G = (V, E)$, 其中顶点对应于像素, 边上的权值对应于像素对相似性。
- (2) 建立仿射矩阵 W 和度矩阵 D 。
- (3) 用第二小本征向量求解广义本征值系统。
- (4) 利用本征向量划分图。
- (5) 完成稳定性检测之后, 如有必要, 重复划分各个已分割部分。

图 1.3 为 Ncut 图像分割示例。图 1.3(a) 是原始图像。因为 Ncut 图像分割

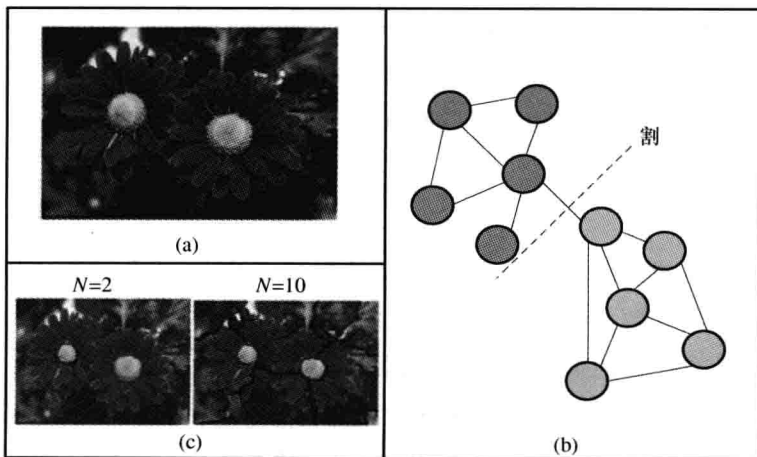


图 1.3 Ncut 分割示例

(a) 原始花朵图像; (b) 相应的图模型; (c) $N=2$ 和 $N=10$ 的分割结果。

方法是在无监督方式下运行的,所以没有用户输入作为已知标签。由该原始图像创建的图如图 1.3(b) 所示,其中虚线代表将该图划分成两部分的一个分割。图 1.3(c) 给出了 $N=2$ 和 $N=10$ 的分割结果。可见,原始图像被分成了不同的区域。

1.2.1.4 高效的图分割算法

文献[24]给出了另一种基于图的图像分割算法。在无监督方式下,该算法基于两两区域比较实现图像分割。利用图像的图表示定义了一个用于度量两个区域之间是否存在边界的谓词,并采用两个区域之间的最小加权边来计算该谓词。尽管算法作出的是贪婪决策,但是其计算时间与图的边数近似成线性关系($O(m \log m)$, m 为图的边数),这在实际应用中是很相当的。通过设置用于平滑输入图像和阈值 θ 的参数 σ 在内的各个参数,可以获得不同的分割结果。关于实验结果和源代码请参见 <http://people.cs.uchicago.edu/pff/segment/>。

1.2.2 基于非参数聚类的分割技术

由 Fukunaga 提出的均值漂移(Mean Shift)分析是一个用于寻找局部样本密度函数模态的非参数循环过程^[31]。Cheng 扩展了原始的均值漂移技术,将其用于图像分析^[33]。具体来说,是通过局部密度峰值估计相似像素的局部密度梯度。已经证明,无论对于静止样本还是对于演化样本,均值漂移过程都是一个二次有界最大化问题^[32]。Comaniciu 和 Meer 将这个算法应用于彩色图像分割^[34]。

给定 d 维空间中的 n 个数据点 x_i , 具有核函数 $K(x)$ 的多变量核密度估计定义如下:

$$\hat{f} = \frac{1}{n} \sum_{i=1}^n K_H(x - x_i) \quad (1.4)$$

对于具有单位阵 $H = h^2 I$ 的径向对称核函数 $\hat{f}$ 可改写为

$$\hat{f} = \frac{1}{nh^d} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right) \quad (1.5)$$

取式(1.5)的梯度并利用一些代数知识,可以得到均值漂移向量:

$$m(x) = C \frac{\nabla \hat{f}(x)}{\hat{f}(x)} \quad (1.6)$$

其中, C 为正的常数, 且

$$m(x) = \frac{\sum_{i=1}^n x_i g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)} - x \quad (1.7)$$

注意: $g(x)$ 是核剖面函数 $k(x)$ 的导数, 即 $g(x) = -k'(x)$ 。

通常情况下, 核函数 $K(x)$ 可分解成两个相异的径向对称核函数(空间域和色域)的乘积。对于静态图像, 均值漂移分割算法包含以下几步^[35]:

- (1) 给定一幅图像, 执行均值漂移滤波过程直至收敛。
- (2) 将所有比空间和颜色带宽更近的点分为一组。
- (3) 为每组分配一个标签。
- (4) 消除具有较少像素点的区域。

图 1.4 为均值漂移图像分割示例。原始图像如图 1.4(a) 所示。图 1.4(b) 给出了不同带宽参数的分割结果, 其中空间和颜色带宽分别为 $h_s = 4, h_r = 10$ 以及 $h_s = 10, h_r = 15$ 。随着核函数带宽的增加, 更多像素的被分为一组, 从而得到了更大的分割区域。该方法也是一种无监督分割技术。

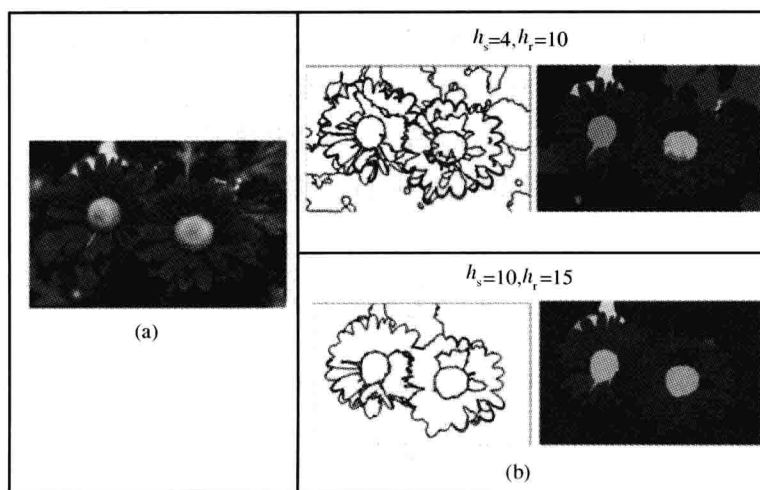


图 1.4 均值漂移分割示例

(a) 原始花朵图像; (b) 不同参数的分割结果。