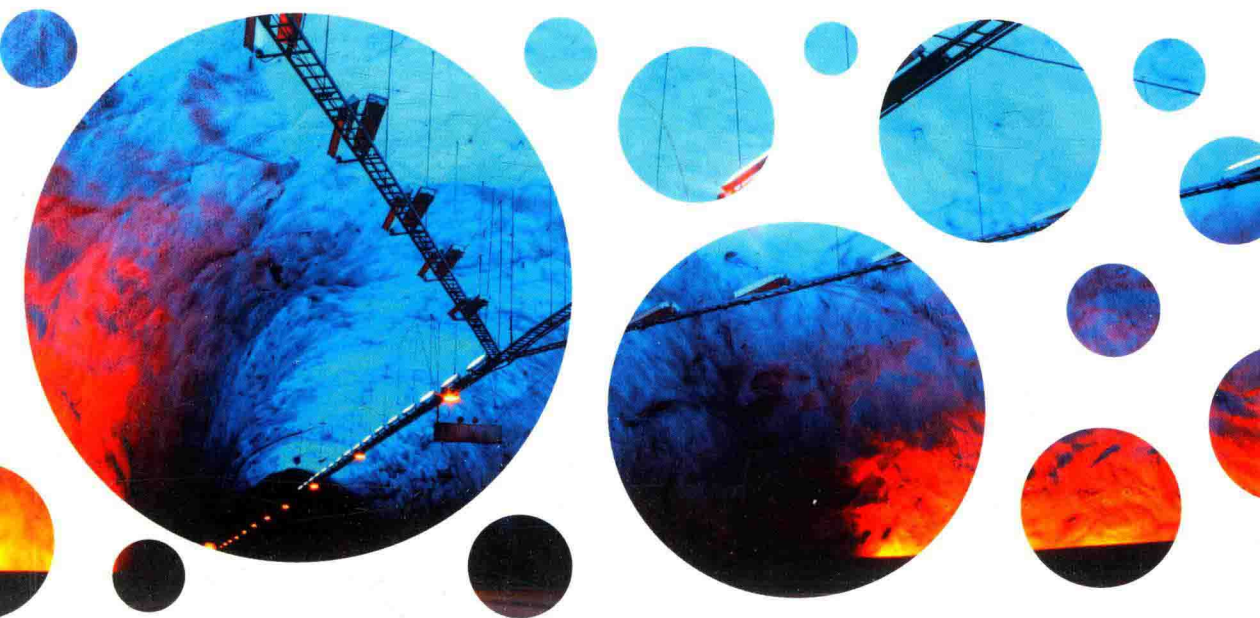




基于视觉显著性的 图像分割

刘占文 著



西安电子科技大学出版社
<http://www.xduph.com>

基于视觉显著性的 图像分割

刘占文 著

西安电子科技大学出版社

内 容 简 介

多示例学习与图割优化是近年来计算机视觉领域广受关注的研究方向。本书系统地论述了多示例学习与图割的基本理论、算法及其在交通视频图像识别中的应用。全书共6章,第一、二章供初学者学习,论述了目标显著性检测方法、基于图论的图像分割方法与多示例学习方法的研究现状,以及图像显著特征的基本定义与相似性度量的常用处理方法。第三、四章论述了显著性目标分割方法的基本原理及具体算法的实现步骤。第五章论述了基于多示例与图割优化的图像分割方法在实际交通视频图像识别中的应用。第六章为结论与展望。

本书适合计算机视觉、图像处理、模式识别等研究方向的读者及开发工程师参考学习。

图书在版编目(CIP)数据

基于视觉显著性的图像分割/刘占文著. —西安:西安电子科技大学出版社, 2019.3

ISBN 978-7-5606-5096-8

I. ①基… II. ①刘… III. ①计算机视觉—检测—图像分割

IV. ①TP391.41 ②TN911.73

中国版本图书馆 CIP 数据核字(2018)第 213562 号

策 划 刘玉芳

责任编辑 刘玉芳 毛红兵

出版发行 西安电子科技大学出版社(西安市太白南路2号)

电 话 (029)88242885 88201467 邮 编 710071

网 址 www.xduph.com

电子邮箱 xdupfxb001@163.com

经 销 新华书店

印刷单位 北京虎彩文化传播有限公司

版 次 2019年3月第1版 2019年3月第1次印刷

开 本 787毫米×960毫米 1/16 印张 9

字 数 124千字

定 价 30.00元

ISBN 978-7-5606-5096-8/TP

XDUP 5398001-1

* * * 如有印装问题可调换 * * *

前 言

图像目标分割是计算机视觉领域的一个重要研究方向，同时也是视觉检测、跟踪与识别等应用的重要基础，其分割质量的好坏在很大程度上影响着整个视觉系统的性能。然而由于缺乏对人类视觉系统的深层认识，图像分割同时也成为了计算机视觉领域的一个经典难题。人类视觉系统能够有选择地注意所观察场景的主要内容，而忽略其他次要内容，视觉的这种选择性注意机制使得高效的信息处理成为可能，同时也启发了计算机视觉的研究者们从注意机制的角度另辟蹊径，因此具有人类视觉特性的图像分割模型将成为图像分割领域一个新的研究热点。

本书主要围绕图像的显著性检测与基于图论的图像分割方法展开研究。针对传统的显著性检测算法所定义的模型缺乏学习能力，以及对显著度的计算不能很好地反映视觉注意机制等问题，提出了一种基于多示例学习的显著性检测算法；并将显著性检测结果应用到基于图割的图像分割算法中，作为图割算法的输入，将图像的显著度引入图割框架，提出了一种基于图割优化的显著目标分割方法，解决了基于图论的图像分割方法计算复杂度高与边界分割不准确的问题。目的是让读者在交通视频图像处理研究领域快速入门，以解决交通视频基本视觉问题为出发点，激发读者对交通视频图像研究的兴趣，为更加深入的学习打好基础，并为解决实际问题提供

研究思路。

我们依托陕西省道路交通智能检测与装备工程技术研究中心、“多源异构交通信息智能检测与融合技术”教育部科技创新团队、高等学校创新引智计划“车路信息感知与智能交通系统创新引智基地”，致力于车路环境感知的嵌入式图像处理与交通视频分析，搭建了多个省市地区交通视频大数据分析平台，并在多示例学习、图割理论、目标分割与跟踪方法，以及智能交通系统中的应用与实现等方面取得了良好的研究成果，本书即是我们在目标分割领域研究工作的初步总结。

本书的完成离不开团队研究生的支持与帮助。特此对王润民博士、沈超博士，徐江、樊星、连心雨、李强、张凡等研究生表示感谢。同时，本书的编写也得到了国家自然科学基金(61703054, 51278058)、装备预研教育部联合基金(6141A02022322)、陕西省重点研发计划工业领域项目(2018ZDXM-GY-044)、长安大学中央高校基本科研业务费高新技术研究培育项目(300102248202)等科研项目的支持，特此感谢。

著 者

2018年9月于长安大学

目 录

第一章 绪论	1
1.1 研究背景与意义	1
1.2 国内外研究现状综述	4
1.2.1 目标显著性检测方法	4
1.2.2 基于图论的图像分割方法	6
1.2.3 多示例学习方法	17
1.3 本文的主要研究内容和章节安排	18
第二章 图像显著特征与相似性度量分析	21
2.1 图像显著特征分析	21
2.1.1 底层特征	22
2.1.2 中高层语义特征	29
2.2 相似性度量分析	31
2.2.1 相似性计算模型	31
2.2.2 相似性度量	39
2.3 本章小结	41
第三章 基于多示例学习的图像目标显著性特征检测	43
3.1 显著性特征的检测	43
3.1.1 图像的预处理	44
3.1.2 亮度梯度特征	45
3.1.3 色彩梯度特征	51
3.1.4 纹理梯度特征	52

3.1.5 图像的边界镜像	54
3.2 基于多示例学习的显著性检测	55
3.2.1 Bag-SVM 算法	56
3.2.2 Ins-SVM 算法	57
3.2.3 APR 算法	58
3.2.4 EMDD 算法	59
3.3 基于多示例学习的显著性检测实验结果分析	60
3.3.1 实验结果对比分析	60
3.3.2 算法性能对比分析	70
3.4 本章小结	72
第四章 基于图割优化的显著性目标分割方法	74
4.1 图的基本概念	75
4.2 基于图割优化的图像显著性目标分割方法	78
4.2.1 自适应图像层次分割方法	78
4.2.2 基于图割优化的图像显著性目标分割方法	80
4.2.3 图像分割质量评价指标分析	86
4.2.4 实验结果对比分析	90
4.3 本章小结	101
第五章 基于多示例学习与图割优化的弱对比度车辆目标分割算法	102
5.1 基于机器视觉的道路交通信息采集与检测系统	103
5.2 基于多示例学习与图割优化的弱对比度车辆目标分割算法	104
5.2.1 车辆目标特征分析	104
5.2.2 算法具体步骤	106
5.3 算法实验与评价	108
5.3.1 算法实验过程	108
5.3.2 实验结果对比	114
5.3.3 算法评价分析	116
5.4 本章小结	118
第六章 结论与展望	119
参考文献	122
致谢	138

第一章 结 论

1.1 研究背景与意义

图像显著性是图像中重要的视觉特征,体现了人类视觉对图像显著区域的重视程度。图像显著性检测是图像重要视觉特征的提取过程,一般以显著度图^[1]的形式表示,它既是视觉注意机制在计算机视觉研究中的主要体现形式,同时对于提高目标识别效率和质量具有极其重要的作用。图像显著性目标分割是按照一定的相似性准则将图像分成各具特性的有意义区域或对象,从而在复杂的背景中把图像特征较为显著的区域或感兴趣的目标对象分离出来的过程。图像显著性检测的结果往往标示了显著目标可能出现的大致位置,使得后续的操作能够有针对性地将处理重点放在显著性区域。因此,将图像显著性检测的结果用于指导图像分割,将资源优先分配给感兴趣的区域,对于解决大尺度图像分割与图像实时分割问题,降低计算量并提高计算机对信息处理的效率等都具有极为重要的研究意义和应用价值。

图像显著性目标分割在图像处理、分析、理解中具有举足轻重的作用,它是最底层图像处理与中间层的图像分析及识别之间的一个关键步骤,其分割质量的优劣、区域界限定位的精度直接影响后续的区域描述以及图像的分析 and 理解。图像分割方法的研究一直是计算机视觉领域的热点问题,在众多研究领域都有非常广泛的应用,包括智能交通检测、生物特征识别、农业环境监测、军事卫星遥感及医学病理影像分析等。例

如,在智能交通领域,从静态复杂交通场景图像或动态交通监控视频序列中分割车辆及车牌,其结果用来识别车型、车牌^[2]或目标跟踪^[3];在生物特征识别领域,首先对生物图像进行分割,从而进一步进行人脸、指纹以及虹膜识别等^[4];在农业环境监测领域,通过图像分割来监测户外农作物的生长状况^[5],及时调整农作物的种植方案,减少经济损失等;在卫星遥感领域,对合成孔径雷达(Synthetic Aperture Radar, SAR)图像的感兴趣区域进行自动提取,可以分析水系、植被、土地、道路、建筑、城市地貌及作物生长状况等^[6],以及多光谱和超光谱遥感的分割分类^[7]等;在工业监控中,通过对精密零件表面的材质纹理进行分割,从而进一步检测零件缺陷^[8];在医学病理影像分析中,医学图像分割是对正常组织和病变组织进行定量分析、区域提取及三维重建等工作的基础。例如对核磁共振图像(Magnetic Resonance Imaging, MRI)进行偏差场纠正和组织分割^[9],以及对X射线断层扫描(Computer Tomography, CT)图像进行病理提取^[10],这些医学图像的分割结果对进一步的计算机辅助诊断、三维重建以及医学图像配准等后续处理都具有重要作用。总之,在各类图像处理应用领域中,都需要通过图像分割来提取特定的、感兴趣的图像目标,以便于进一步分析与识别。

图像分割是图像处理和计算机视觉中一个经典的基本问题,是许多计算机视觉问题不可或缺的步骤,分割输出的质量在很大程度上影响整个视觉系统的性能。在过去几十年,国内外研究学者对图像分割方法进行了深入而广泛的研究,发展到现在已经形成了一个庞大的图像分割方法体系^[11],因此,对图像分割方法的分类从不同的角度具有多种划分途径。从信息学的角度依据图像信息种类不同而分类,大致分为基于区域信息的图像分割方法、基于边缘信息的图像分割方法、结合边缘与区域信息的图像分割方法;从统计学的角度依据图像特征不同而分类,分为基于特征空间的图像分割方法、基于阈值的图像分割方法与基于聚类的图像分割方法;从数学理论的角度依据采用的数学工具不同而分类,分为结合小波与数学形态学的图像分割方法、基于人工神经网络的图像分割方法、基于能量函数优化的图像分割方法以及基于图论的图像分割方

法等；依据分割过程中是否有人工参与而分类，分为自动的图像分割方法和交互式图像分割方法；还有一些其他的图像分割方法，例如基于物理模型的分割方法、基于无监督学习的分割方法和基于有监督学习的分割方法等。

前述的图像分割技术中，有一些分割技术成功地将图像元素映射到图中进行分割以获取感兴趣区域，例如基于图论的图像分割方法与基于能量函数优化的图像分割方法。基于图论的图像分割技术在近 30 年引起了学者们的广泛关注，成为图像分割领域一个较新的研究热点。其主要思想是将图像映射成加权图，将图像像素看做图的顶点，邻接像素之间的关系看做图的边，邻接像素之间的相似性看做边的权值，根据边的权值设计能量函数或图的分割准则，通过最小化能量函数或图割优化完成对图的分割，从而实现图像分割。将传统意义上的图像分割问题转变为图论中的图割最优化问题，图割最优化问题本质上就是对图的顶点分簇进行最优划分。来自图论的有效工具在离散空间中解决了分割问题。在一幅图上指定分割的好处是由于纯粹的组合操作可能不需要离散化，因此不会产生离散化的错误。

本文的研究重点是图像显著性目标分割的问题，对目前图像目标显著性检测方法、基于图论的图像分割方法、图割优化算法以及图像分割质量的评价方法等都进行了深入的总结、分析和研究，提出了一种基于多示例学习的目标显著性检测算法，将显著性检测的结果用于基于图论的图像分割方法中指导图像分割，对图割模型框架等诸多环节进行了优化；采用凝聚层次聚类算法作为图割优化的求解方法，使得分割结果能更好地符合语义感知的输出；深入研究了五种目前公认的著名图像分割评价指标，以 Achanta 等人建立的数据库以及 Berkeley 图像分割数据库的图像为实验对象；对本文算法与其他基于图论的分割方法的分割质量进行了对比分析，选用 PRI、VI、GCE、P-R-F 指数等评价指标对算法进行定量评价，验证了本文算法的有效性。

本文同时受到高等学校学科创新引智计划项目“车-路信息感知与智能交通系统创新引智基地”（“111 计划”基地编号：B14103）、国家物联网

重大示范工程专题研究项目“集成一体化车载通用感知终端设备的研发”(项目编号:2012-364—812-105)和国家自然科学基金项目“车联网环境下融合多源交通信息的车辆行为辨识与安全预警方法”(项目编号:51278058)的资助,针对基于机器视觉的图像目标分割关键技术进行研究。

1.2 国内外研究现状综述

1.2.1 目标显著性检测方法

本文主要关注的是从计算机视觉角度出发的显著性检测,主要分为自底向上和自顶向下的方法。前者由数据驱动,独立于具体任务;后者受意识支配,依赖于具体任务。下面分别介绍这两种方法近些年的研究现状和发展趋势。

1. 自底向上的显著性检测方法

自底向上的显著性检测方法从低层的视觉特征出发,通过颜色、纹理、亮度、尺度等信息的对比度计算,确定出每个视觉单元(像素/区域)的显著性大小。根据计算过程中所采用的支撑信息范围的不同,此类方法又分为局部显著性检测和全局显著性检测。

对于局部显著性检测的方法,每个视觉单元的对比度由待考察单元和周围局部邻域单元的差异来决定。Koch 和 Ullman 最早于 1985 年提出显著度的概念,并研究了视觉系统在观察场景图像时,注意力从一个位置到另外一个位置的转移^[12]。其基本思想是基于颜色、边界方向图等基本特征,用 Winner-Take-All 的方法确定出每个像素和周围邻域的差异度,然后将此差异度作为显著度大小的描述,最后再根据显著度的大小和像素相似度进行注意力的转移建模。受此工作的启发,Itti 等人于 1998 年提出了基于多尺度融合的显著性检测方法^[1]。在一系列不同尺度上的

中心-邻域比较下,将基于不同特征信息所获得的显著度图线性融合起来并进行归一化。在此项工作发表之后,显著性检测如同雨后春笋般发展起来,各种不同的技术也层出不穷。Achanta 等人在显著性检测时更注重体现图像中的高频信息,这使得显著度图中的边缘得到了很好的保护,前景-背景的区分度也更加明显^[13]。同时,实验中所建立的包含 1000 幅图像和手工标注显著性结果的数据库,成为显著性检测中最有影响力的标准数据库之一。在后续工作中,Achanta 等人将显著物体的尺寸与其在图像中所处的位置结合起来,通过在中心-邻域显著性滤波时变化带宽以适应不同大小的可能目标^[14]。Harel 和 Koch 提出了一种基于图的显著性算法^[15]。在提取激励图的基础上,将其归一化为统一的显著度图。Walther 等人将 Itti 的显著性检测模型同层级式的物体识别系统结合起来,用于识别图像中的物体^[16],其原理符合生物视觉机理,实验结果取得了很大的提高。Frintrop 等人将积分图像用于特征计算过程,从而提高了显著性检测的速度,使得该系统可以实时工作^[17]。

对于全局显著性检测的方法,其基本思想是建立在整个图像统计特性的基础上。Wang 和 Li 提出一个两步法则来检测显著性^[18]。首先,通过引入自动通道选择和决策反馈来扩展剩余频谱模型,然后根据格式塔原则补充显著性区域^[19]。Zhang 等人在自然图像统计特性的先验条件下,通过贝叶斯推理和图像底层特征来进行显著性检测^[20]。Cheng 等人根据图像中统计直方图的出现频率和像素在图像中所处的位置来判定显著度与否,并将此结果用于后续的图像语义分割^[21]。Zhai 和 Shah 用时间维和空间维信息从视频的角度检测显著性^[22]。时间维信息主要包含相邻帧之间的运动关系,空间维信息主要包含帧内颜色直方图信息。

2. 自顶向下的显著性检测方法

自顶向下的显著性检测方法一般都是基于统计学习的方法,即根据训练数据学习判定模型后,在测试数据上进行验证。Liu 等人结合多尺度对比度、中心-邻域直方图、颜色分布等特征,用条件随机场(Conditional Random Field, CRF)学习显著度检测模型,并将此模型推广到视频显著性的检测^[23],其构建的图像和视频数据库是此研究领域最早的测试数据

库。Hou 和 Zhang 在对大量自然图像进行统计的基础上发现,显著区域所对应的傅里叶变换的频谱对应于对数频谱曲线上的奇异点,据此可以实现对显著区域的高效检测^[24]。Judd 等人建立了一个人眼追踪的数据库,此数据库包含 15 个被试者在 1003 幅图像上的视觉注意点轨迹记录^[25]。基于此数据库,结合低层、中层、高层图像特征,用 SVM 分类器进一步学习视觉关注点的位置。

除了上述两种类型的显著性检测方法外,还有很多研究者致力于从不同的角度去定义显著性。Chang 等人将一组图像中共同出现的目标物体定义为协同显著性(Co-saliency)^[26]。Goferman 等人对显著性的定义则从上下文的角度出发,其检测的显著目标不仅包含显著物体,还包含支撑其定义的背景^[27]。Wang 等人通过将当前待考察图像与一系列同类图像比较,得出一个外在显著性的定义方式,其体现的是图像区域与其他图像中类似部位的差异大小^[28]。

1.2.2 基于图论的图像分割方法

在众多的图像分割方法中,基于图论的分割方法在实际应用中具有一些很好的特性。它将图像像素明确地组织为合理的数学结构,使分割问题能够公式化表述,更加灵活,计算效率更高。Wu 等人在 1990 年提出,将图的顶点关系作为一种代价函数的数学表述方法来分割图像^[29]。从那以后,大量的研究转移到对图的优化技术的研究上。众所周知,图像分割中的困难之一是其本身所固有的病态性质。由于可能会有对图内容的多种解释,对于一个给定图像的分割很难找到唯一的正确答案。这表明,为了精确地提取出感兴趣的目标,图像分割应该包含中、高层次的信息。20 世纪 90 年代末,一个先进的图像技术出现在使用特定模型相关的线索与上下文信息的组合中,较为有影响力的代表是 S/T 的曲线图切割算法^[30]。它的技术框架与一些按照离散方式的变分方法是密切联系的。到目前为止,S/T 的图割和它的变体已经用于解决许多计算机视觉问题,最终作为这些领域中的优化工具。

本文对基于图论的图像分割方法进行了系统性的研究,一般是将图

分成若干个子图,每一个子图都表示图像中一个感兴趣的对象。当采用统一的表述时,这些方法大致分为五类:(1)基于最小生成树的方法;(2)基于代价函数的图割方法;(3)基于马尔可夫随机场模型的图割方法;(4)基于最短路径的方法;(5)不属于上述任何一类的其他方法。

下面详细描述每一类的具体公式与实现方法,分别讨论了每一类中具有影响力的几种代表性方法的优缺点。尽管将这些方法分成5类,但这5类方法经常一起使用,它们之间的区别主要在于如何定义分割方法所要求的分割质量和使用不同的图特性来实现分割。

1. 基于最小支撑树的方法

最小支撑树(也称为最短路径树)是图理论中一个重要的概念。一幅图 $G=(V, E)$ 的支撑树 T 是一棵树 $T=(V, E')$, 这里 E' 包含于 E 中。一幅图可能有几个不同的支撑树,最小支撑树是所有支撑树中具有最小权值的那个支撑树。计算最小支撑树的算法在文献[31-33]中有详细描述,例如,在 Prim 算法中,最小支撑树由迭代地添加最小边权值的前边而构成,该算法是一种贪婪算法,运行时间呈多项式倍数^[33]。早期的基于最小支撑树的图像分割方法大多基于最小支撑树与聚类结构之间的内部关系^[34],该最小支撑树是由所有支撑树中具有最小权值和的边集构成的,因此保证了顶点之间的连接都满足彼此最大相似性,且所有的顶点在不同类之间跨越最小间隔。现实图像中的复杂场景经常在感知上具有非统一密度的有意义的类,因此更应该考虑类间的相异性和类内的相似性。有研究者用最小支撑树来层次地划分图像,该方法基于相似像素分在一起且不相似像素分离的原则,来获得不同尺度下的分割图像^[35]。通过切割最小支撑树中权值最高的边,得到由相似性最小的相邻子图组成的分割图。在文献[35]中,提出了一些基于最小支撑树的改进算法,例如递归最小支撑树算法。在每一次迭代中,通过分割一个子图来完成一次分割,因此,算法最后能得到一个已知子图数的分割图。显然,这种形式的算法效率较低。文献[36]提出采用一种快速递归最小支撑树算法来加速文献[35]中的算法。文献[37]提出一种基于自适应阈值的最小支撑树算法,该算法充分利用两个子图间的相异性和子图内的相似性,分割的过程伴

随着区域的合并,并且分割结果满足全局属性。与使用常数 K 来设定阈值的单连接聚类相比,自适应阈值是可变的,且根据集合的大小定义。区域的合并准则如式(1.1):

$$|e_i| < \min\left(\text{Int}(C_1) + \frac{K}{|C_1|}, \text{Int}(C_2) + \frac{K}{|C_2|}\right) \quad (1.1)$$

式中, K 是一个常量, $|C_1|$ 和 $|C_2|$ 分别是集合 C_1 和 C_2 的大小, $\text{Int}(C)$ 是 C 的最小支撑树中最大的边权值, $|e_i|$ 是连接集合 C_1 与 C_2 的最小权值边。式(1.1)表明算法对平滑区域的边是敏感的,反之,对高变化区域的边是不敏感的。

由上述讨论可知,以无向权值图作为分割对象,最小支撑树的算法明确定义了聚类的架构,利用亮度、颜色或位置等底层特征来表达像素并很好地实现聚类分割。因此,该类算法经常用于其他高级应用的初级处理^[38, 39]。最小支撑树通常通过削减最高边的权值来形成分割图,对树的进一步切割可得到新的分割图,说明最小支撑树分割方法是一种层次分割,在不损失类特性的条件下可将任何过度分割转化为高级分割。

2. 基于代价函数的图割方法

1) 最小割方法

割的概念早期经常被用于网络流的研究中。1990年初, Wu 和 Leahy 在文献[40]中首次提出使用基于代价函数的图割方法进行图像分割,也称为最小割方法。与最小生成树方法类似,基于代价函数的图割也是建立在具有明确定义的边权值图上的一种概念。相比于上述方法,考虑到图的特性,文献[40]提出一种基于代价函数的通用框架来优化图的分割,针对不同的应用,根据明确定义的分割目标来设计不同的代价函数。文献[40]给出一个明确有意义的图的分割:最小化图割框架将图分割成若干子图并使不同子图之间的割最小,以保证不同子图间的最大相异性,如式(1.2)所示:

$$\min\text{Cut}(\mathbf{A}, \mathbf{B}) = \sum_{u \in \mathbf{A}, v \in \mathbf{B}} w(u, v) \quad (1.2)$$

其中, $\min\text{Cut}(\mathbf{A}, \mathbf{B})$ 表示子图 $\mathbf{A}$ 与子图 $\mathbf{B}$ 的最小割, $w(u, v)$ 表示子图 $\mathbf{A}$ 与 $\mathbf{B}$ 内部所有顶点之间的边的权值和。根据文献[41]Ford-Fulkerson 的网

络流理论体系提出的最大流最小割算法,即最大流等于最小割的容量,可以有效地求解二部图(二分图)最小割。但在文献[40]中,作者研究了一种更通用的 K 分图的情况,这种 K 分图的定义采用了文献[42]中多端网络流理论的思想,得到最小割即获取 K 个顶点集合之间的最大流。最小割的分割方法存在一定的缺陷,根据式(1.2)可以看出,该方法只考虑两个子图之间的割最小,就会趋向于寻找边数较小的割,而单个顶点或小簇顶点集与其补图之间的边数往往最少,因此,偏向于分离出孤立点或者小簇顶点集,如图 1.1 所示。

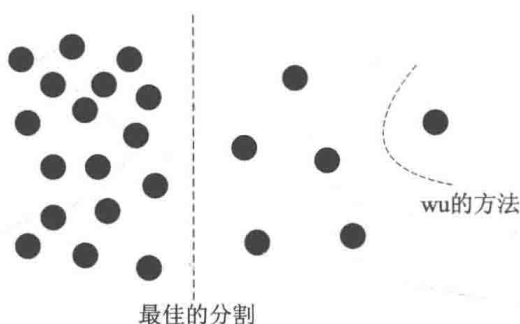


图 1.1 最小割的分割方法缺陷

2) 归一化割方法

为了避免出现图 1.1 所述的情况,同时又考虑到能够将每个顶点集合控制在合理的范围内,学者们开始研究这个问题,相继提出了多种归一化目标代价函数分割框架。其中一个比较著名的代价函数是由 Shi 和 Malik 提出的归一化割(Normalized cut, Ncut)分割准则^[43],还将 Ncut 方法分为 2 分 Ncut 与 K 分 Ncut。该方法的图割不仅与子图之间的顶点的边权值相关,还取决于任一子图与其补图之间的顶点的边权值和。2 分 Ncut 方法的公式原理如下:

$$\begin{aligned} \text{Ncut}(\mathbf{A}, \mathbf{B}) &= \frac{\text{cut}(\mathbf{A}, \mathbf{B})}{\text{assoc}(\mathbf{A}, \mathbf{V})} + \frac{\text{cut}(\mathbf{A}, \mathbf{B})}{\text{assoc}(\mathbf{A}, \mathbf{V})} \\ &= \frac{\sum_{x_i > 0, x_j < 0} -w_{ij}u_iu_j}{\sum_{x_i > 0} d_i} + \frac{\sum_{x_i < 0, x_j > 0} -w_{ij}u_iu_j}{\sum_{x_i < 0} d_i} \end{aligned} \quad (1.3)$$

其中, V 为顶点集, $\text{assoc}(\mathbf{A}, \mathbf{V}) = \sum_{x_i \in A, x_j \in V} w(x_i, x_j) = \sum_{x_i > 0} d_i$, $\sum_{x_i > 0} d_i$ 表示子图 A 与其补图之间的顶点的边权值之和, 同理 $\text{assoc}(\mathbf{A}, \mathbf{V})$ 与此类似。 $U = \{U_1, U_2, \dots, U_i, \dots, U_N\}$ 是一个指示向量, 若 $x_i \in A$, 则 $U_i = 1$; 若 $x_i \in B$, 则 $U_i = -1$ 。 $d_i = \sum_j w(x_i, x_j)$ 表示顶点 x_i 与其余所有顶点的权值和, 令 D 为 N 维对角矩阵, 其对角线上元素 $d_i = \sum_j w(x_i, x_j)$; I 为 $N \times 1$ 维矩阵, 其元素都为 1; X 为 $N \times 1$ 维矩阵, 其元素为 x_i ; $k = \frac{\sum_{x_i > 0} d_i}{\sum_i d_i}$, 令 $b = \frac{k}{1-k}$, 则 $y = (I + X) - b(I - X)$ 。根据矩阵论中瑞利商的理论, 求解式(1.3)的最小化问题就转化为一个求解广义特征系统的问题^[44]。

$$\begin{aligned} \min \text{Ncut}(\mathbf{A}, \mathbf{B}) &= \frac{[(I + x) - b(I - x)]^T (D - W) [(I + x) - b(I - x)]}{bI^T DI} \\ &= \arg. \min_y \frac{y^T (D - W) y}{y^T Dy} \end{aligned} \quad (1.4)$$

如式(1.4)所示的广义特征系统, $(D - W)y = \lambda Dy$, 当 λ 为 $D - W$ 相对于 W 的次小广义特征值时, 所对应的特征向量 y 为式(1.4)求解的最小图割优化值。采用特征值进行图分割的思想可追溯至 20 世纪 70 年代, 次小特征值又被称做 Fielder 值, Fielder 最早在文献[45]中提出使用次小特征值来分割图。Shi 和 Malik 提出的 K 分图法是由 2 分图的关系表达式演变而来的, 如式(1.5)所示:

$$\begin{aligned} \text{Ncut}(\mathbf{A}, \mathbf{B}) &= \frac{\text{cut}(\mathbf{A}_1, \mathbf{V} - \mathbf{A}_1)}{\text{assoc}(\mathbf{A}, \mathbf{V})} + \frac{\text{cut}(\mathbf{A}_2, \mathbf{V} - \mathbf{A}_2)}{\text{assoc}(\mathbf{A}_2, \mathbf{V})} \\ &\quad + \dots + \frac{\text{cut}(\mathbf{A}_k, \mathbf{V} - \mathbf{A}_k)}{\text{assoc}(\mathbf{A}_k, \mathbf{V})} \end{aligned} \quad (1.5)$$

对于上式的求解, 同二分法类似, 将其转化为求解特征系统 $(D - W)y = \lambda Dy$ 的特征值和与其相对应的特征向量, 然后利用 K-means 算法或者其他聚类算法对上一步骤中的特征向量进行聚类。也可以先用聚类算法将原始图像分为 n 块, 然后一次合并 2 块, 遵照最小化 K-way Ncut