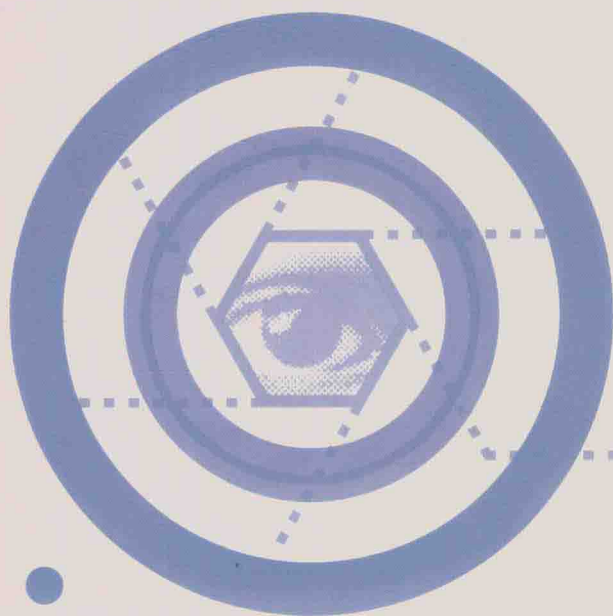


基于视觉显著性的 目标检测方法与应用研究

刘君玲◎著



 科学出版社

基于视觉显著性的目标检测 方法与应用研究

刘君玲 著

科学出版社

北 京

内 容 简 介

本书通过对场景感知的两种加工模式的研究,提出了自底向上与自顶向下相结合的混合加工模型,并将场景图像的局部特征与全局特征并行加工处理,通过特征融合实现了场景图像的显著性目标检测;通过认知心理学方法研究了底层视觉特征对不同类别目标识别的影响,采用神经网络偏置自适应共振理论模型实现了模拟与仿真,并实现了基于视觉显著性特征的目标快速识别;采用视觉显著性目标检测方法实现了火焰与烟雾目标的识别,通过提出的傅里叶频谱时空差异与多步长帧间差累积算法改进了现有的视觉显著性目标检测方法,提高了目标检测的准确性与检测速度。

本书可作为高等院校计算机视觉或图像增强等研究方向的参考教材,也可作为计算机视觉、模式识别等相关行业从业人员的参考用书。

图书在版编目(CIP)数据

基于视觉显著性的目标检测方法与应用研究/刘君玲著. —北京:科学出版社, 2018.3

ISBN 978-7-03-056818-2

I. ①基… II. ①刘… III. ①计算机视觉—目标检测 IV. ①TP391.41

中国版本图书馆 CIP 数据核字 (2018) 第 048796 号

责任编辑:戴薇 王惠 / 责任校对:陶丽荣
责任印制:吕春珉 / 封面设计:东方人华平面设计部

科学出版社 出版

北京东黄城根北街 16 号

邮政编码: 100717

<http://www.sciencep.com>

北京京华虎彩印刷有限公司 印刷

科学出版社发行 各地新华书店经销

*

2018 年 4 月第 一 版 开本: B5 (720×1000)

2018 年 4 月第一次印刷 印张: 6 1/2

字数: 123 000

定价: 48.00 元

(如有印装质量问题,我社负责调换《京华虎彩》)

销售部电话 010-62136230 编辑部电话 010-62135397-2052

版权所有, 侵权必究

举报电话: 010-64030229; 010-64034315; 13501151303



刘君玲 女，1974年

7月出生，副教授，博士，毕业于吉林大学计算机应用专业。主要研究方向：计算机视觉、

场景理解。近年来发表EI检索论文7篇，中文核心论文1篇；获实用新型专利6项，软件著作权2项；主持吉林省科学研究项目3项，教学研究项目3项，教学改革项目1项。获吉林省科技进步奖2项，教学成果奖1项，教学研究成果奖2项，教育技术成果奖3项；获吉林省首届微课大赛一等奖。

前 言

视觉显著性是指场景图像中不同区域的主观感知质量存在差异,使场景中的某些目标从其相邻的背景区域脱颖而出,立刻吸引人们的注意。目前,在计算机视觉领域,显著目标检测引起了研究者的兴趣,将视觉显著性分析方法引入目标检测会大幅提高检测效率。计算机视觉研究领域的图书较多,但涉及视觉显著性技术研究的较少。本书为视觉显著目标检测提供了一个理解现有方法和技术,并为以后的研究做准备的系统框架。

视觉显著性是视觉感知与场景理解的重要研究内容,涉及认知神经科学、认知心理学、计算机视觉和信息处理等多个学科,其研究内容十分广泛,场景认知、目标识别、图像检索等方面都有涉及。视觉显著性源于场景中候选前景目标与背景在某种特定的对比中形成了引起人眼注意的新奇刺激。本书采用计算机视觉与认知心理学的关键技术对场景图像的视觉显著性进行了深入的研究,并将研究成果应用于烟雾与火焰的视觉显著目标检测,实现了场景图像与视频帧中显著目标的自动检测。

本书通过对视觉认知的两种加工模式——自底向上与自顶向下的研究发现,底层视觉刺激可以吸引注意资源的分配,而顶层的视觉感知和先验知识又能很好地指导视觉显著目标的检测,将两者相结合可以提高检测效率。在一幅场景图像中,图像的局部特征对视觉显著目标检测有影响,而全局特征又起到了主导作用,采用知觉混合理论将两种特征并行处理,有利于快速实现显著目标检测。本书采用自底向上与自顶向下相结合的视觉信息加工模式,同时将场景图像的局部特征与全局特征并行加工处理,通过特征融合实现场景图像的显著目标检测。

本书采用认知心理学方法研究快速场景认知与显著特征对目标识别的影响两部分内容,并采用偏置自适应共振理论神经网络模型实现基于显著特征的目标识别模拟与仿真,通过前人的研究发现场景的快速识别受图像的尺寸与颜色等特征的影响。本书通过心理学实验完成了4类目标 bird、ball、butterfly、flower 在彩色图像、灰度图像、边缘过滤及低通过滤下的识别,实验结果表明目标识别发生在快速场景认知之后;在目标识别中,颜色、形状、纹理是具有重要认知作用的视觉特征,不同的目标对不同的特征依赖程度有所不同,影响了识别的速度与识别正确率。本书还采用偏置自适应共振理论神经网络模型对认知心理学实验进行了模拟与仿真,通过对目标特征颜色、形状、纹理特征的提取实现了4类目标的识别,通过两阶段的学习比较体现了偏置自适应共振理论模型自组织、自学习的

能力。

基于视觉显著性的火灾目标检测技术主要是在不需要人为干预的情况下,实现采集图像的自动处理与分析,根据火险视觉目标特征进行检测与匹配,达到动态场景中火险与火情的定位、识别和跟踪。本书在研究谱残差算法基础上,采用多视频帧间频谱差异与单视频帧的窗口间频谱差异完成显著目标检测,利用火焰的颜色视觉特征采用 CIELab 颜色空间聚类方法确定显著对象,并通过块逆概率差方法锁定视觉显著点,根据显著点 RGB 与 CIELab 颜色特征值锁定火焰目标,实现火灾告警。烟雾一般发生在火灾的初期,尤其在室外复杂场景中呈现出蔓延与扩散的特征,并且具有随时间累积的视觉特征。本书采用多步长帧间差累积算法实现显著运动区域的检测,并基于低秩矩阵恢复理论通过多视频帧构建超完备的字典找出前景稀疏部分,实现了烟雾前景对象从复杂背景中的分离。本书同时研究了烟雾颜色的半透明性视觉特征,通过运动区域增长与 HSV 颜色空间显著性检测实现烟雾目标检测。实验表明,本书的方法在检测速度与准确性上取得了很好的效果,且该方法适用于不同的室外场景。

由于著者水平有限,书中的疏漏及不妥之处在所难免,敬请广大读者批评指正。

著 者

目 录

第 1 章 绪论	1
1.1 视觉显著性的研究背景与意义	1
1.2 视觉显著性目标检测概述	3
1.2.1 视觉显著性目标检测算法的分类	3
1.2.2 视觉显著性目标检测算法的比较分析	5
1.2.3 视觉显著性目标检测技术的应用	16
1.3 本书的研究内容与组织框架	16
1.3.1 本书的研究内容	16
1.3.2 本书内容的组织框架	18
第 2 章 基于场景感知的混合处理模型	19
2.1 视觉信息加工模式研究	19
2.1.1 通用视觉信息加工模式	19
2.1.2 主动视觉加工模式	20
2.1.3 场景感知混合加工模式	21
2.2 场景感知特征理论研究	21
2.3 基于混合加工模型的显著目标识别	23
2.3.1 视觉特征的提取	23
2.3.2 基于混合加工模型的显著目标识别过程	30
小结	37
第 3 章 视觉认知的心理学研究与 bART 神经网络模拟	38
3.1 快速场景认知的研究	38
3.1.1 空间尺度在快速场景认知中的作用	40
3.1.2 颜色特征在快速场景认知中的作用	41
3.2 视觉显著特征对目标识别影响的比较研究	42
3.2.1 心理学实验研究	44
3.2.2 实验设计与分析	45
3.3 bART 神经网络模拟	48

3.3.1	自适应共振理论	49
3.3.2	偏置自适应共振理论神经网络模型	50
3.3.3	底层视觉特征的提取	54
3.3.4	实验研究	57
	小结	61
第 4 章	基于视觉显著性的火焰目标检测	63
4.1	视觉显著区域	65
4.1.1	傅里叶变换	65
4.1.2	视觉差异显著区域的生成	68
4.2	显著目标对象的提取	71
4.3	基于显著特征点的火焰检测	73
	小结	76
第 5 章	基于视觉显著性的复杂场景烟雾检测	78
5.1	显著运动区域的提取	79
5.1.1	多步长帧间差累积算法	79
5.1.2	显著运动区域提取步骤	81
5.2	显著运动对象的提取	82
5.2.1	图像的稀疏表示与字典学习	82
5.2.2	稀疏矩阵求解	84
5.2.3	显著运动对象像素点的提取	85
5.3	烟雾目标检测	86
5.3.1	运动增长区域的提取	86
5.3.2	烟雾的颜色显著区域检测	86
5.4	采用多种方法的显著运动目标检测	88
	小结	90
	参考文献	91

第1章 绪 论

1.1 视觉显著性的研究背景与意义

近年来, 监控摄像头已经成为城市景观, 随处可见, 产生的最严重的问题就是信息过载。外围传感器不断产生的传入信号构成了海量数据, 计算机视觉技术面临的任务就是在繁多的图像、视频文件中快速发现感兴趣的目标, 这在今天仍然是具有挑战性的问题。

基于各种任务去理解从视觉场景看到的事物, 称为“视觉感知”^[1]。视觉感知的研究目的是场景理解, 场景理解是在场景图像的各种处理和分析的基础上解释场景图像的内容、场景中包括的对象及它们之间的关系, 即理解场景的高层语义知识。视觉场景是一组特定对象的特定空间关系的抽象, 在场景感知过程中, 视觉组织要解决的基本问题是对刺激的某些区域或部分进行组合, 区分目标与背景, 并判断哪些部分属于同一个客体^[2]。自然场景通过视网膜传递给视觉系统, 由视觉神经传导处理转化为能知觉到的图形, 这不仅需要对视觉刺激进行处理, 还需要把这些信息加工组织形成知觉。格式塔心理学家在 20 世纪早期发现了一些视觉组织的规律, 后来又发现了若干著名的原理及法则, 被称为完形法则 (Gestalt Law), 即接近律、相似律、闭合律、连续律、成员特性律^[3]。心理学家近年来又发现了一些新的规律, 包括: Palmer 等通过研究耳颞神经关联和感官意识的绑定发现位置相邻或相近的视觉像素容易被知觉系统处理为同一个整体^[4]; Engel 等通过对时空结合与神经感觉意识的相关性研究发现, 复杂环境的视觉刺激的感知通常需要对不同特征信息进行平行加工, 通过不同视觉神经元的传导达到不同刺激特征的区分, 经视觉组织整合形成特定的客体, 可以实现与无关信息的分离^[5]。在场景感知信息处理过程中, 关于特征整合问题的研究, Gray 等通过对猫视觉皮层的神经元方向列的研究发现: 脑中不同特征细胞的编码可通过同步发放实现特征与其活动的时间关联^[6]。从场景特征描述的角度来看, Lowe 通过对形状特征尺度不变的关键点的研究给出场景尺度不变特征变换 (Scale-Invariant Feature Transform, SIFT)^[7], 该变换可以使场景图像在发生多尺度变换或发生平移与旋转时也能达到较好的识别效果。Torralba 等通过对全局特征在目标搜索中的作用进行研究发现, 在现实世界场景的观察中, 眼球运动受注意先验文指导^[8]。Oliva

提出了场景 Gist 理论^[9], 该理论指出在一个场景视觉感知过程中涉及 3 级信息处理, 即从底层加工场景各级特征(如颜色)空间频率的 Gist 信息, 到中间层提取场景图像属性信息(如表面、体积), 最后达到高层次的场景语义信息的激活。目标检测(Object Detection)^[10]是视觉认知的重要任务, 主要是将场景对象的分割与识别相结合, 采用对象的几何和统计特征对图像进行分割。目标检测的准确性和实时性是衡量计算机视觉系统的一项重要指标, 尤其是在复杂场景中, 当需要对多个目标进行实时处理时, 目标检测就显得特别重要。目标检测主要是为了检测场景图像中某一类语义对象的实例(如在室外场景中检测行人、建筑物或汽车), 但事实上相同语义的目标在不同场景中可以表现出不同的形状、颜色、尺寸、亮度等, 同时还存在旋转、仿射、遮挡甚至变形等情况, 所以如何提高目标检测的效率仍是计算机视觉系统面临的挑战性问题。

人类视觉可以通过扫视快速地发现场景图像中的目标, 通过分配注意资源可以进一步分析图像的内容与语义, 实现图像理解的过程。视觉场景的所有信息对于观察者来说并不是同等重要。例如, 在完成图像搜索、目标识别及视频语义标注等任务时, 人们所关注的只是图像中较小的一个或多个区域, 大脑只需要对部分重要信息做出响应, 显著性的特征信息更容易引起优先注意。视觉显著性(Visual Saliency)是视觉感知与场景理解的研究热点, 涉及认知神经科学、认知心理学、计算机视觉和信息处理等多个学科, 其研究内容十分广泛, 图像处理、目标认知、图像检索、图像分类等方面的研究都有涉及。视觉显著性是场景图像中不同区域的主观感知的质量^[11], 这使场景中的某些目标从其相邻的区域中脱颖而出, 立刻吸引人们的注意。当场景图像中的前景对象与其背景在某些视觉特征上形成较为明显的视觉差异(或称为新奇刺激)时, 前景对象就突显出来形成视觉显著性。影响显著性的因素主要包括视觉刺激的相似性与差异性。Posner 等^[12]研究发现, 在视觉搜索中, 当目标特征与中央线索一致时反应较快, 显著性的目标获得优先加工; Serences 等^[13]通过对视觉选择性注意和知觉的协同研究发现, 视觉系统用选择性注意神经元的活动来解决链路竞争和显著对象表示, 通过注意力转移这个短暂的控制信号, 可以将视觉系统从一个一致的认知状态“推送”到优先加工特征显著的目标上; Rolls 等^[14]通过神经电生理方法对自然场景视觉注意的研究表明, 颞叶皮层神经元的接受域在处理复杂场景的对象时会变小且集中于中央凹处, 与视觉任务相关的对象特征会被优先处理; Blaser 等^[15]通过特征空间对象的追踪研究, 认为注意的确可以选择与追踪任务相匹配的特定的视觉对象, 这些特定视觉对象的特征空间不仅包括项目定义的空间轨迹, 同时具有明确的特征时间轨迹; Itti 等^[16]通过对图像进行多尺度的方向、颜色和亮度特征提取与差异比较形成了显

著图,实现了前景目标与背景的分离; Bruce 等^[17]采用信息理论方法提出了一个信息最大化关注模型; Avraham 等^[18]提出了估计显著性随机数学模型,采用概率方法找出最有可能的显著目标; Marchesotti 等^[19]提出了一个新颖的显著性视觉特征检测框架,假设图像的全局特征都有相似的描述属性,而每一图片的全局特征又有其独特之处,通过构建一个简单的分类器来检测具有独特之处的视觉特征,从而形成显著图; Liu 等^[20]采用监督学习的方法实现显著性目标检测,完成了基于局部、区域和全局特征的显著对象的分割。

自然场景中的显著性目标检测对于图像分析过程至关重要。注意是人类信息加工过程中的一项重要的心理调节机制,它能够对有限的信息加工进行分配,使感知具备选择能力。如果能够将这种机制引入图像分析领域,将计算资源优先分配给那些容易引起观察者注意的区域,必将极大地提高现有的图像处理分析方法的工作效率。本书将视觉显著性机制引入目标检测,主要通过视觉显著性区域的检测使观察者快速定位场景图像的关键信息,通过制订合理的计算资源分配方案,提高现有目标检测方法的性能。

1.2 视觉显著性目标检测概述

显著性是一个视觉系统具有的对视觉刺激选择注意的能力,这种机制主要是基于行为与任务去过滤掉不相关的信息,而仅对感兴趣的信息进行选择并对其进行处理。近年来,视觉显著性目标检测引起了广泛的关注,主要是因为它能够为目标检测提供快速的解决方案。该机制可以检测到场景图像中最显著、最能够吸引注意的对象,然后对该对象涉及部分进行分割,生成属于显著对象的密度概率图,从而实现前景目标与背景图像的分离。

1.2.1 视觉显著性目标检测算法的分类

1. 根据信息处理层次的不同划分

根据信息处理层次不同,视觉显著性检测算法可分为3类:

第一类是自底向上基于数据驱动的显著性区域检测算法,它依赖于视觉特征本身的刺激,具有代表性的是 Itti 等^[16]提出的模拟生物体视觉注意机制的选择性注意算法(简称 IT 算法)。

第二类是自顶向下任务驱动的显著性目标检测算法,该算法依赖于场景内部的结构、事物本身的先验知识等,如 Goferman 等提出的 CA 算法^[21](Context-Aware Saliency Detection), Oliva^[9]采用 Gist 特征与场景语义关联作为显著性检测的上下

文, 引导注意资源分配。

一方面, 自底向上的显著性检测底层特征一般与背景具有较明显的差异, 能吸引更多的注意资源, 这种注意力检测机制是不自主的、快速的信息处理过程, 当我们关注室外场景时, 路上的行人、行驶的车辆或颜色鲜艳的花朵都会表现出较强的视觉显著性; 另一方面, 自顶向下的显著性检测是基于任务或期望的具有自主性的意愿, 检测需要耗费一定的运算时间。1967 年, Yarbus^[22]完成的著名的注意力依赖于当前任务实验表明, 显著性检测与主观任务有较大的相关性。

第三类是结合自顶向下和自底向上的方法, 目前应用较为广泛, 因为该算法具有两方面的优势: 可保持对象分组假设的一致性, 可避免穷举搜索和误报。例如, Navalpakkam 等^[23]采用贝叶斯方法将显著目标优化整合为一个自顶向下的注意力检测提示, 同时将对比度和方向作为自底向上的线索; Levin 等^[24]同时将自底向上和自顶向下的线索放在条件随机场中。

2. 根据信息处理模式的不同划分

根据信息处理模式的不同, 视觉显著性检测算法可分为 3 类: 一类是基于生物视觉特性的算法, 代表性算法是 IT 算法; 另一类是纯数学计算方法, 如 Achanta 等^[25]提出的全分辨率算法、Hou 等^[26]提出的基于空间频域分析的谱残差 (Spectral Residual, SR) 算法及近几年发展很快的低秩矩阵恢复算法^[27] (Low Rank Matrix Recovery, LRMR); 最后一类是将前两种算法进行融合的方法, 代表性算法是 Harel 等^[28]提出的基于图论的算法 (Graph-Based Visual Saliency, GBVS)。

3. 根据场景图像检测处理单元划分

根据场景图像检测处理单元, 视觉显著性检测算法可分为两类: 基于像素特征的检测方法和基于区域特征的检测方法。基于像素特征的检测方法主要以像素为单位计算不同特征的显著度图, 在场景图像中定位显著对象, 此类方法的特点是计算特征简单快速^[1], 但对于稳健性较差尤其是存在强噪声的图像检测准确性较低, 如亮度对比度 (Luminance Contrast, LC) 算法^[29]采用各像素点颜色统计信息来提取空间显著性图; 基于区域特征的检测方法主要利用视觉分割算法先将图像划分为不同特征的区域块, 然后通过引入竞争机制实现显著对象的区域定位^[1], 该方法基于认知心理学研究, 是较为常用的检测方法, 但由于分割算法造成细小的杂块, 不利于完整的语义对象的标注, 并且计算量较大, 稳健性有待提高。典型的算法有局部对比度的区域对比 (Region Contrast, RC) 算法^[30], 其先用基于图的快速分割算法将图像划分为区域块, 再以块为单位建立整幅图像的稀疏直方图, 通过颜色差异、距离远近、基本元素大小 3 方面度量进行区域显著性分析,

该算法检测结果精度高且全面。

4. 根据对比度划分

根据场景图像特征的对比度,视觉显著性检测算法可分为局部特征对比度检测方法和全局特征对比度检测方法。基于局部特征对比度的显著性检测方法主要会在场景对象的边缘区域产生较高的显著值,致使显著目标区域无法产生很一致的高亮值,如 IT 算法^[16]和 Ma 等提出的一种基于局部对比度分析的 MZ 算法^[31];而基于全局特征对比度的显著性检测方法是在场景图像上整体考查对比度关系,采用像素颜色差异、空间的距离等来考量视觉的显著程度,这类方法适用于将前景对象从背景图像中的分离,典型的有 LC 算法^[29]、RC 算法^[30]。

5. 根据检测对象划分

根据显著性检测对象的不同,将注意分为 3 类:基于空间的显著性检测、基于目标的显著性检测及基于特征的检测。基于空间的显著性检测是指注意单元是图像中感兴趣的区域,注意单元的转移是从一个区域转移到另一个区域,甚至转移到没有明确场景意义的区域。最经典的基于空间的视觉检测模型是由 Itti 依据 Treisman 等^[32]的特征整合理论提出的 Itti 模型。目前存在的另一个视觉选择性注意计算模型是基于目标的注意。视觉显著性检测的单元是场景中的目标、目标的一部分或目标组, Sun 等^[33]首次提出并实现了一个基于目标的显著性检测模型,其主要思想是显著性检测的单元不是空间的某个点或某个区域,而是具有明确物理意义的完整物体。此外,脑成像研究还发现了一种新的显著性检测机制——基于特征的选择性注意, Corbetta 等通过正电子发射断层显像 (Positron Emission Tomography, PET)^[34]及 Treue 等^[35]、Saenz 等^[36]通过功能性磁共振成像 (Functional Magnetic Resonance Imaging, FMRI) 研究发现,视觉刺激特征是由两部分构成的复合体,一部分是由空间特征的形状、大小、方位等特征构成的;另一部分是由非空间特征如密度、颜色和明度等构成的。显著性检测系统选择性地对任何一个维度的信息或某几个维度信息的组合进行加工。基于空间的注意是一种更为基础的显著性检测机制。但是基于特征和基于客体的检测机制又有其独立的脑神经机制,在特定条件下有其独特的功能。

1.2.2 视觉显著性目标检测算法的比较分析

视觉显著性检测算法很多,具有代表性且实验结果具有很高的准确率,同时又有很高引用率的算法有 IT^[16]、CA^[21]、SR^[26]、GBVS^[28] (Graph-Based Visual Saliency, 基于图论的视觉显著性)、LC^[29]、RC^[30]、相位差分析^[37]、FT^[38] (Frequency-

Turned, 频变)、AIM^[17] (Attention-based on Information Maximization, 基于信息最大化的注意力)、MSS^[39] (Maximum Symmetric Surround, 最大对称环绕)、分割^[40]、估算^[41]、CB^[42] (Content-based, 基于内容的)、SWD^[43] (Spatially Weighted Dissimilarity, 空间加权相异性)等。这些算法遵循了不同的视觉显著性规律, 各算法处理的视觉特征不同, 得到的显著图也各有特点。通过以下几种典型算法的比较分析, 可以探索显著性检测方法的基本思路, 为后续的研究与应用提供理论参考。

1. IT 算法

IT 算法是由 Itti 等在 Koch 等提出的生物学结构和模型^[44]的基础上最早发展起来的显著性检测算法, 该算法采用自底向上的信息处理模式将场景图像的各像素点在颜色、亮度、方向方面与领域像素点的对比值作为显著性的度量。图 1.1 描述了 IT 算法的工作过程。

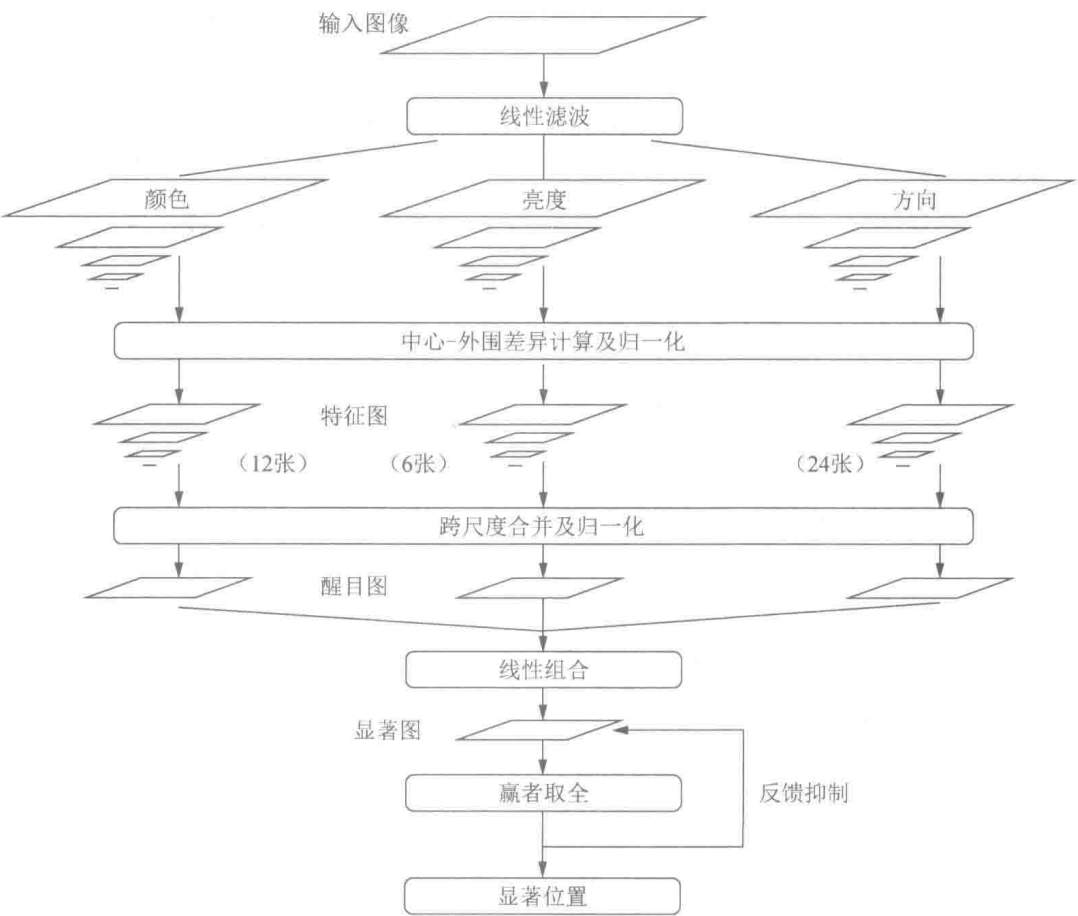


图 1.1 IT 算法的一般结构

算法实现步骤如下:

1) 读取图像, I_0 为初始图像, 进行 $1/2$ 采样得 I_1 , 一直进行 $1/2$ 采样直到得到 I_8 , I_8 为 I_0 的 $1/256$, 如图 1.2 所示。



图 1.2 图像采样

2) 提取特征, 包括 4 种颜色信息 R 、 G 、 B 、 Y 构建的高斯金字塔 $R(\sigma)$ 、 $G(\sigma)$ 、 $B(\sigma)$ 、 $Y(\sigma)$ (其中 $\sigma \in [0,8]$, σ 为尺度因子, $[0,8]$ 表示 9 个尺度), 4 个朝向信息 0° 、 45° 、 90° 、 135° 及亮度 I 共 9 种特征信息。

$$\begin{cases} I = (r + g + b) / 3, \\ R = r - (g + b) / 2, \\ G = g - (r + b) / 2, \\ B = b - (r + g) / 2, \\ Y = (r + g) / 2 - |r - g| / 2 - b. \end{cases} \quad (1-1)$$

式中, R 表示彩色图像红色信息; B 表示彩色图像蓝色信息; G 表示彩色图像绿色信息; r 、 g 、 b 分别表示彩色图像 R 、 G 、 B 三种信息对应元素的取值。

3) 根据 Center-surround 算法, 通过计算小尺度和大尺度特征空间上的特征差别来计算显著性。如图 1.3 所示, 取 $2 \sim 5$ 、 $2 \sim 6$ 、 $3 \sim 6$ 、 $3 \sim 7$ 、 $4 \sim 7$ 、 $4 \sim 8$ 共 6 种跨尺度的差值, 再加上 I 、 RG 、 BY 与 O 的 4 个方向差值即 7 种取差信息, 共计生成 42 幅特征图。其中差运算 $\ominus$ 过程采用先进行插值使其具有相同大小, 然后对应相减取绝对值的方法。

中心是较小尺寸图中的某个像素, 图像的大小 $c \in \{2,3,4\}$; 周围是该像素在较大尺寸图中的对应像素, 较大的图像的大小为 $s = c + \tau$, $\tau \in \{3,4\}$, 灰度特征显著图共 6 张:

$$I(c,s) = |I(c) \ominus I(s)| \quad (1-2)$$

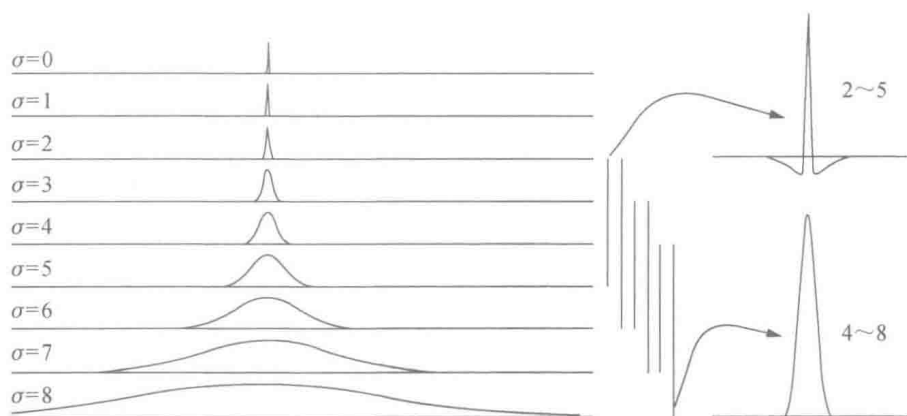
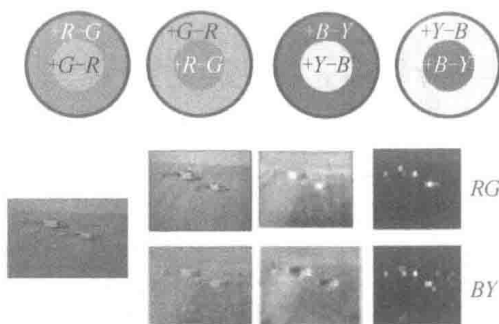


图 1.3 不同尺度特征的差值获取

如图 1.4 所示,采用双色对立原理得到颜色特征显著图, $RG(c,s)$ 和 $BY(c,s)$ 的 6 种差值共有 12 张:

$$\begin{cases} RG(c,s) = |[R(c) - G(c)] \ominus [G(s) - R(s)]|, \\ BY(c,s) = |[B(c) - Y(c)] \ominus [Y(s) - B(s)]|. \end{cases} \quad (1-3)$$

图 1.4 RG 与 BY 颜色差值

Gabor 局部朝向特征显著图为 $O(c,s,\theta)$ (其中 $\theta \in \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$), 获得 24 张:

$$O(c,s,\theta) = |O(c,\theta) \ominus O(s,\theta)| \quad (1-4)$$

4) 特征图融合: 先对特征进行归一化处理 $N(\bullet)$, 然后分别在尺度特地空间求和, 包括 Im (亮度和)、 Cm (RG 、 BY 和 I)、 Om (4 个方向和)。求和与上面取差类似, 先插值, 后求和。

按图 1.5 所示做归一化处理 $N(\bullet)$ 操作后实现了以下 3 个方面的工作:

- ① 各像素的值经归一化处理后, 取值范围为 $[0, M]$;
- ② 找到全局最大值 M , 以及计算出其他局部最大值的均值 $\bar{m}$;
- ③ 显著图内每次归一化后的值都乘以 $(M - \bar{m})^2$ 。

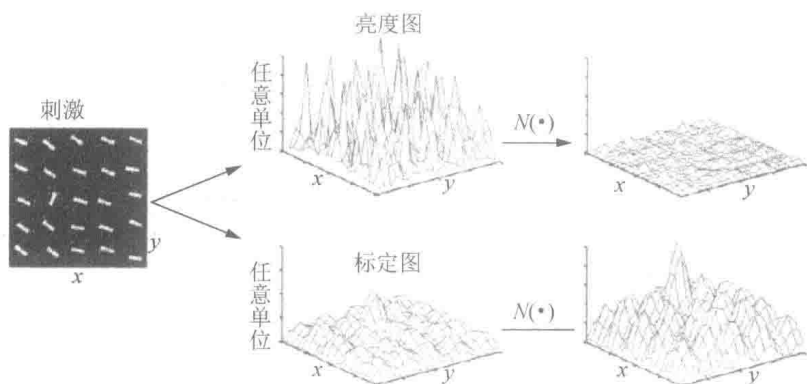


图 1.5 归一化处理因子

通过叠加操作将它们整合到一起，得到图像显著图（图 1.6）：

$$\begin{cases} \bar{I} = \bigoplus_{c=2}^4 \bigoplus_{s=c+3}^{c+4} N[I(c,s)], \\ \bar{C} = \bigoplus_{c=2}^4 \bigoplus_{s=c+3}^{c+4} \{N[RG(c,s)] + N[BY(c,s)]\}, \\ \bar{O} = \sum_{\theta \in \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}} N \left\{ \bigoplus_{c=2}^4 \bigoplus_{s=c+3}^{c+4} N[O(c,s,\theta)] \right\}. \end{cases} \quad (1-5)$$

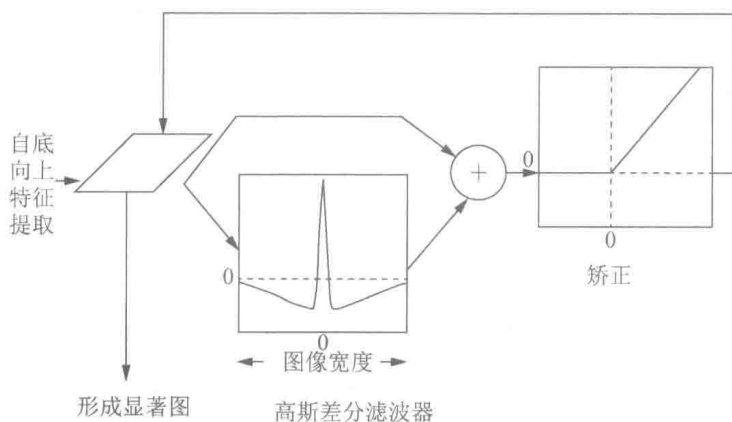


图 1.6 各特征的过滤与叠加

5) 显著图的生成（图 1.7）：

$$S = [N(\bar{I}) + N(\bar{C}) + N(\bar{O})] / 3 \quad (1-6)$$

IT 算法作为最早提出的视觉注意算法具有开创性意义，而且该算法是图像显著性检测技术方面第一个相对较为完善的模型，近年来许多学者在此研究基础上进行了改进优化研究，如图 1.8 所示。