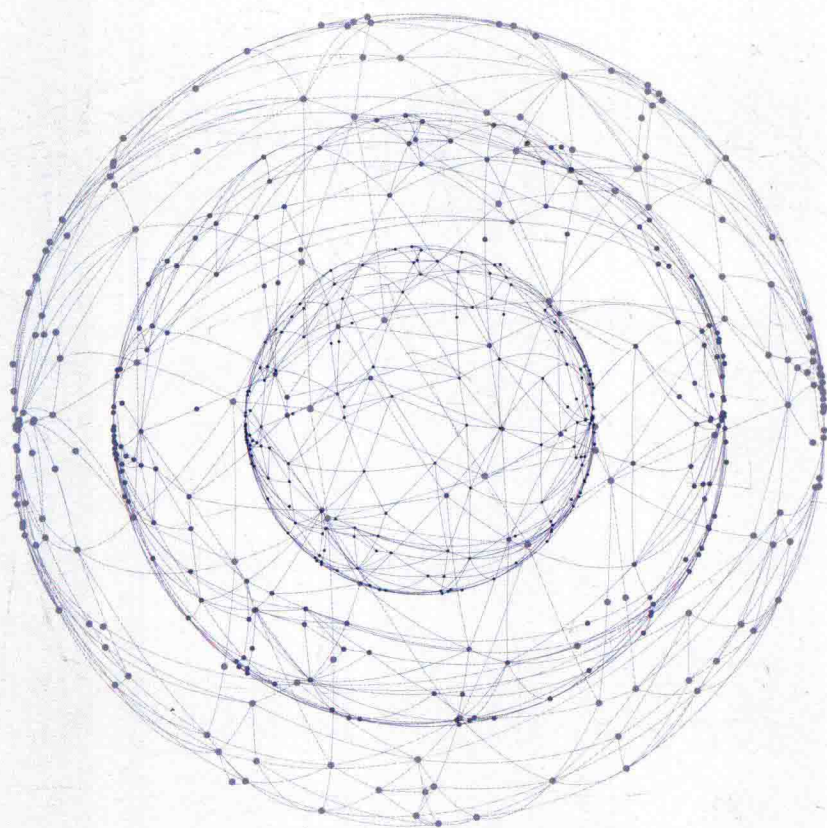


智能科学与技术丛书

复杂场景的 视觉目标跟踪方法

郭文 著



科学出版社

复杂场景的视觉目标跟踪方法

郭 文 著



科学出版社

北 京

内 容 简 介

本书主要介绍了当前国内外与跟踪相关的研究现状,系统总结了作者近些年在跟踪方面所做的工作,介绍了视觉跟踪方面主要的分析思想与相关理论,以及先进的视觉跟踪方法。全书分为8章,具体内容为基于视觉关注度的核跟踪方法、基于黎曼度量的鱼群优化贝叶斯跟踪方法、基于不确定性最大置信度提升算法的视觉跟踪方法、辅助子空间更新的自适应集成随机森林跟踪方法和低秩重检测的多特征时空上下文跟踪方法等在视觉跟踪中的应用。

本书可作为自动化、通信与电子、计算机科学等专业大学本科高年级学生以及计算机视觉与模式识别专业方向的硕士生与博士生的参考书,也可作为高等学校大学生和研究生学习机器视觉、图像处理的参考书。

图书在版编目(CIP)数据

复杂场景的视觉目标跟踪方法/郭文著. —北京:科学出版社,2017
ISBN 978-7-03-055795-7

I. ①复… II. ①郭… III. ①计算机视觉-视觉跟踪 IV. ①TP391.41

中国版本图书馆CIP数据核字(2017)第300853号

责任编辑:赵丽欣 常晓敏 / 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>

三河市骏杰印刷有限公司印刷

科学出版社发行 各地新华书店经销

*

2017年12月第一版 开本:B5(720×1000)

2017年12月第一次印刷 印张:7 1/4

字数:146 000

定价:45.00元

(如有印装质量问题,我社负责调换〈骏杰〉)

销售部电话 010-62136230 编辑部电话 010-62134021

版权所有,侵权必究

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

前 言

进入 21 世纪,随着计算机技术的发展,视频数据已经成为人们记录和保存信息的主要形式之一,这导致视频数据以惊人的速度增长,但是这些海量的视频数据存在着大量的冗余信息,浪费了巨大的存储空间,并且导致很多有价值的信息无法查询。因此如何有效地分析和理解这些视频内容已经成为一个急需解决的问题,这个问题的核心就是如何利用计算机技术对视频内容进行有效的理解与表达。通常情况下,视频中运动的目标是视频内容理解的主要关注对象,人对运动目标的感知程度比对静态信息的感知程度要高,因而基于视觉的运动目标跟踪成了智能视频分析与理解的一个关键技术,它为视频内容的理解提供了重要的保障。

视觉目标跟踪是计算机视觉领域中的研究热点之一,在视频监控、人机交互、智能交通、视频检索等领域具有广阔的应用前景。虽然经过人们在目标跟踪领域几十年的研究,目标跟踪技术已经有了长足的发展,但是实现复杂场景中对任意目标进行稳定、准确的跟踪仍然有很多理论与技术问题亟需解决,因此在复杂场景下特别是在跟踪过程中存在遮挡、背景干扰、目标的复杂运动和目标形状与外观变化等情况下完善快速、鲁棒和准确的目标跟踪方法和理论仍是当前跟踪领域努力的方向。

本书共分 8 章。第 1 章介绍了相关研究的内容及意义;第 2 章介绍了目标跟踪的研究现状;第 3 章提出了基于视觉关注度的核跟踪方法;第 4 章给出了基于黎曼度量的鱼群优化贝叶斯跟踪方法;第 5 章提出了基于不确定性最大置信度提升算法的视觉跟踪方法;第 6 章介绍了辅助子空间更新的自适应集成随机森林跟踪方法;第 7 章介绍了低秩重检测的多特征时空上下文的视觉跟踪方法;第 8 章对本书的内容进行总结,并指出进一步研究的发展方向。

本书的研究工作得到了国家自然科学基金(批准号:61572296、61603372、61472227、61672327、61303086、61328205)、山东省自然科学基金(批准号:ZR2015FL020)和模式识别国家重点实验室开放课题(批准号:201600024)基金的资助。

感谢实验室的研究生游思思对全书资料的整理,并参与了第 2 章内容的编写。参加研究工作的还有庞清乐、丁昕苗、白延成、杨福刚、林忠海、马加庆、慕延华等老师,在此一并致谢!另外要特别致谢徐常胜研究员、华臻教授、朱智林教

授等，感谢他们为本书写作所提供的帮助和支持。此外，信息电子与工程学院控制理论与控制工程学科对本书的出版给予了大力支持，在此也由衷地表示感谢。

由于作者水平有限，全书不妥与疏漏之处在所难免，敬请读者批评指正。

目 录

第 1 章	绪论	1
1.1	视觉目标跟踪的研究意义与发展前景	1
1.2	本书的主要研究内容	2
1.2.1	目标的表观模型	2
1.2.2	运动状态的搜索策略	3
1.3	本书的结构安排	3
第 2 章	目标跟踪研究现状综述	6
2.1	引言	6
2.2	目标跟踪方法的国内外综述	7
2.2.1	常用的目标表观模型	7
2.2.2	常见的目标跟踪方法	9
第 3 章	基于视觉关注度的核跟踪方法	16
3.1	引言	16
3.2	视觉关注度简介	17
3.3	基于视觉关注度的均值漂移跟踪方法	18
3.3.1	均值漂移跟踪方法简介	18
3.3.2	运动目标视觉关注度检测	19
3.3.3	视觉关注度加权的目标定位	21
3.3.4	遮挡情况的处理	22
3.4	实验结果	23
3.5	本章小结	26
第 4 章	基于黎曼度量的鱼群优化贝叶斯跟踪方法	27
4.1	引言	27
4.2	人工鱼群优化算法	28
4.2.1	背景与起源	28
4.2.2	觅食行为	29
4.2.3	聚群行为	29
4.2.4	追尾行为	30
4.3	贝叶斯框架下的跟踪	30
4.3.1	通用贝叶斯框架	30
4.3.2	粒子滤波算法	31

4.4	基于黎曼度量的鱼群优化贝叶斯跟踪方法	34
4.4.1	鱼群优化的粒子滤波算法	34
4.4.2	黎曼度量下的目标表观模型	36
4.4.3	黎曼度量的鱼群优化粒子滤波跟踪方法	38
4.5	实验评估	39
4.5.1	定性的评估	39
4.5.2	定量的评估	44
4.6	本章小结	49
第5章	基于不确定性最大置信度提升算法的视觉跟踪方法	51
5.1	引言	51
5.2	Boosting 算法	52
5.3	最大置信度提升算法	53
5.3.1	非确定性标签提升算法	54
5.3.2	处理无标签数据的 MCB 算法	55
5.3.3	理论分析	56
5.4	基于 MCB 的跟踪方法	58
5.4.1	动机与问题描述	58
5.4.2	跟踪框架	58
5.4.3	分类器特征构建与样本提取	60
5.5	实验评估	61
5.6	本章小结	67
第6章	辅助子空间更新的自适应集成随机森林跟踪方法	69
6.1	引言	69
6.2	基于随机森林的集成跟踪	70
6.2.1	在线的随机森林分类器	70
6.2.2	前景与背景分类	71
6.2.3	基于自组织的多特征融合目标定位	72
6.2.4	辅助增量子空间学习的状态更新	73
6.3	实验评估	75
6.3.1	单特征对比自适应多特征融合	75
6.3.2	定性与定量评估	77
6.4	本章小结	81
第7章	低秩重检测的多特征时空上下文的视觉跟踪方法	82
7.1	引言	82
7.2	时空上下文跟踪	84
7.2.1	上下文先验表达	84

7.2.2 空间上下文表达	85
7.2.3 基于时空上下文的目标定位	85
7.3 低秩重检测的多特征时空上下文跟踪算法	86
7.3.1 多特征融合的时空上下文目标定位	86
7.3.2 基于在线低秩滤波器的重定位	87
7.4 实验结果分析	89
7.4.1 实验细节	89
7.4.2 实验评测	91
7.5 本章小结	97
第 8 章 总结与展望	98
8.1 总结	98
8.2 研究展望	99
参考文献	101

第1章 绪 论

1.1 视觉目标跟踪的研究意义与发展前景

随着计算机视觉技术的发展,视频数据已经成为人们记录和保存信息的主要形式之一。如何在海量的视频数据中挖掘有价值的信息、避免资源浪费就是摆在我们面前的问题。有效地分析和理解这些信息的核心。在于我们采用计算机对存储内容进行有效的理解和表达。通常情况下,视频中运动的目标是视频内容的主要关注对象,由于人对运动目标的感知程度比静态信息的感知程度要高,因而基于视觉的运动目标跟踪成了智能视频分析与理解的一个关键技术,它为视频内容的理解提供了重要的保障。视觉跟踪技术包括计算机视觉、信号处理、图像处理、模式识别、机器学习、人工智能、数据挖掘及多媒体检索等多学科的技术,对促进多学科发展与交叉有重要的意义。复杂场景,如杂乱背景、目标的姿态变化、光照突变及目标遮挡,也为视觉目标的跟踪提出了许多具有挑战性的难题,如何在各种条件下取得鲁棒的目标跟踪结果具有非常重要的理论研究价值。

除了具有重要的理论研究价值外,目标跟踪技术的研究在智能视频监控、人机交互、智能交通、视频压缩与视频检索等领域同样具有广阔的发展前景和重要的实用价值。

1. 智能视频监控

智能视频监控系统的需求主要来自那些对安全要求较高的场所,如银行、商店、停车场等。这些系统的特点是减少对人工的依赖,利用计算机视觉和视频分析的方法自动分析摄像机记录的视频序列,实现对场景中目标的定位、识别和跟踪,并在此基础上分析、判断目标的行为,进而对异常行为做出甄别。智能视频监控系统中采用的技术需要计算机视觉、模式识别和人工智能等学科的支撑,对感兴趣的目标实时检测和连续跟踪是其中的首要环节,是基于视频的运动分析、行为理解等诸多后续处理工作的基础。

2. 人机交互

传统的人机交互是通过计算机键盘和鼠标进行的。在高级用户接口领域中,希望未来的计算机能够与我们进行更加容易和便捷的交流,如手势驱动控制、手语翻译等。计算机无接触式地收集人在计算机前的视频信号,利用计算机视觉的

相关理论分析视频信号, 分辨人的动作, 理解人的意图, 从而得到可靠、准确的连续跟踪结果, 这是这种以人为中心的智能人机交互技术的前提条件。

3. 智能交通

在现代城市交通系统中, 城市的交通管理与调度问题面临严峻的挑战, 主要包括交通流量控制、车辆异常行为检测、行人行为判定等, 这些工作难以通过人工独立完成, 目标跟踪研究也具有非常广泛的应用, 特别是道路上车辆异常行为检测在交通事故的预防和事后处理中具有重要意义, 其核心技术是如何分割和跟踪车辆并对所跟踪车辆的运动轨迹进行判定。

4. 视频压缩与视频检索

随着视频和流媒体技术的发展, 视频数据被广泛采用, 它们比普通的文本信息更直观, 但是数据量也更加巨大。单纯地采用扩大存储容量、增加通信带宽的方法来增加视频存储量是不现实的, 因此迫切需要更有效的视频压缩技术和检索技术。视频检索技术的关键是在对视频数据内容理解的基础上, 对内容进行语义描述, 并通过检索语义特征来完成查询任务。而目标运动分析可以用来获取视频所蕴含的语义信息, 并对视频内容进行语义描述, 鲁棒的运动目标分割和跟踪同样是实现视频压缩去除冗余的重要基础。目前, 新一代的视频图像压缩编码标准 MPEG-4^[1]已经提供了基于运动估计和补偿技术的压缩编码方式。

除了上述民用方面的应用外, 运动目标跟踪技术在军事领域也有广泛的应用前景, 如自动驾驶飞行器、自动目标跟踪、军事机器人, 特别是在制导和导航方面有着十分重要的地位。综上所述, 运动目标跟踪技术的研究有很大的现实意义和应用价值。

1.2 本书的主要研究内容

本书从复杂场景下目标跟踪存在的诸多实际问题出发, 通过对目标进行合理的表现、表达并建立有效的跟踪策略, 以实现精准的目标跟踪。因此, 本书的主要内容将围绕目标跟踪中表现模型的建模、运动状态的估计展开研究。

1.2.1 目标的表现模型

目标的表现模型是指建立有效的目标特征表达以及衡量目标表现模型的相似度, 并要实时予以更新。复杂场景, 如光线变化、遮挡、杂乱背景及自身姿态、形状的变化给目标跟踪带来很大的困难, 容易导致跟踪的失败。一个好的目标的表现模型应该具备如下特性:

- 1) 表达目标的表现可以使目标区域与背景区域具有足够的区分性。

2) 能够在复杂的环境变化下保持足够的准确性。

因此本书将研究如何构建有效的目标表观模型,使其对光照变化、噪声、部分遮挡、形变等复杂因素都比较鲁棒,并且具有一定通用性的目标跟踪算法。

1.2.2 运动状态的搜索策略

运动状态的搜索策略是视觉跟踪中极其重要的问题,这个搜索过程就是目标的定位过程,是寻找候选区域与目标模板最佳匹配的过程。当前,处理跟踪的主要方法大致可以分成自下而上的方法和自上而下的方法。自下而上的方法的跟踪思路是通过优化一个代价函数估计目标所在的位置(如均值漂移算法),这类跟踪算法的实时性比较好,但由于依靠底层的特征数据,容易陷于局部极值;或者通过在线的分类器区分前景与背景所在的区域(如基于分类器的判别式跟踪方法),但这类判别式跟踪方法由于分类器的分类误差极易产生误差积累。自上而下的跟踪方法则将跟踪看成一个贝叶斯的推理问题,通过估计运动状态的分布定位目标(如粒子滤波算法),这类问题存在实时性差和样本退化的问题。因此,本书将从两类方法存在的缺点出发,研究如何避免算法陷于极值,如何提高分类器分类性能和减少样本退化的问题。

1.3 本书的结构安排

本书主要章节之间的逻辑关系如图 1-1 所示,主要章节针对不同问题提出相应的解决办法。第 3 章与第 4 章解决目标表观模型和遮挡问题,第 3 章的方法比较适合监控视频;第 5~7 章针对基于检测/分类方法的漂移问题,第 5 章的方法突出理论,第 6 章的方法注重融合,第 7 章的方法侧重精度与实时性。

第 1 章为绪论,阐述复杂场景的视觉目标跟踪方法的研究意义发展前景与主要研究内容,本书的组织结构及本书的主要贡献。

第 2 章主要对目标跟踪领域的研究现状进行综述。首先详细地介绍目标跟踪中常用的表观模型,然后从基于目标的描述方法和基于目标定位的搜索策略两个不同的角度对现有的目标跟踪方法进行梳理。

第 3 章提出一种基于视觉关注度加权的均值漂移跟踪方法。该章首先简要回顾视觉关注度的相关知识,为了提高均值漂移算法颜色直图表观适应复杂场景下跟踪的能力,通过计算将基于运动信息的动态关注度和基于 Karhunen-Loeve (KL) 变换处理的静态关注度结合起来提取目标的关注区域,采用基于关注度加权的目标区域直方图来提高对目标的表达能力,同时为了提高算法对部分遮挡的鲁棒性,采用交互式的贝叶斯滤波技术,提高跟踪的性能。

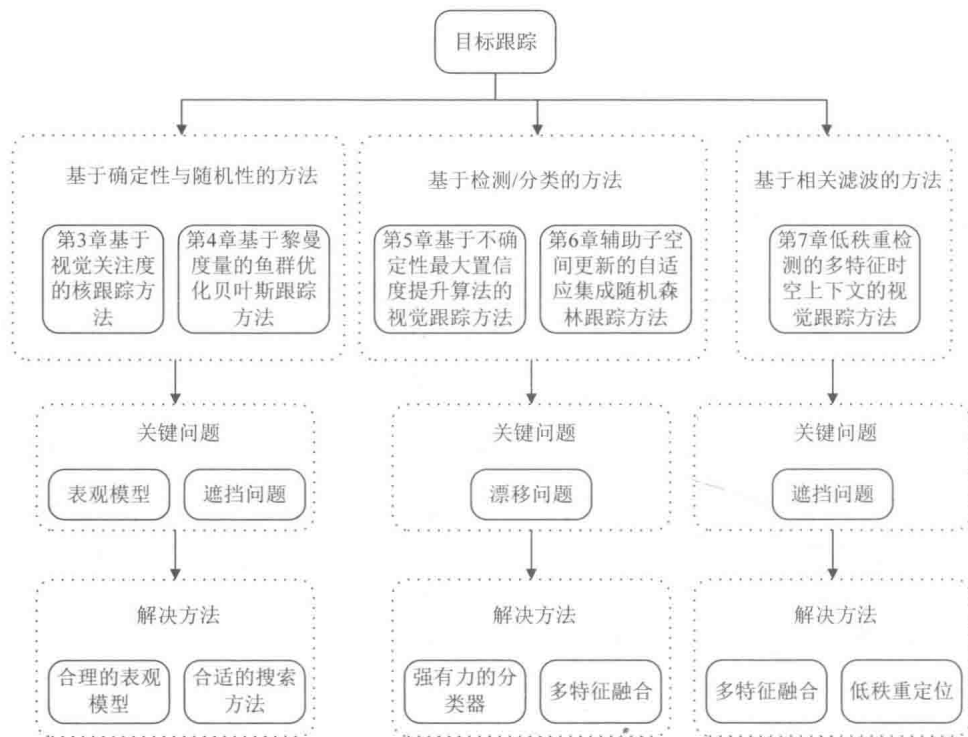


图 1-1 本书主要章节的逻辑关系

第 4 章提出一种基于黎曼度量的鱼群优化贝叶斯目标跟踪方法。该章首先介绍人工鱼群算法、贝叶斯跟踪的一般框架和区域协方差算子的计算，然后通过分析基于协方差贝叶斯跟踪方法存在的问题，提出采用人工鱼群算法提高粒子滤波中粒子的质量，使粒子更多地处于高似然度区域，提高粒子的采样质量，从而能够更好地估计后验概率。目标的相似度计算采用协方差算子黎曼度量，这种表观模型的表达提高了算法对目标跟踪中目标表观的变化以及复杂背景环境的影响。在处理部分遮挡的问题中，采用分块的结构，这提高了算法对部分遮挡的鲁棒性，这几部分合理地融合在一起形成完整的黎曼度量下的鱼群优化贝叶斯目标跟踪框架，最后通过实验验证我们方法的高效性。

第 5 章提出一种基于最大置信度提升算法的视觉跟踪方法。该章首先简要叙述 Boosting 算法；然后从提高分类器精度、扩展算法适用范围的角度出发，将自适应提升算法的确定标签改成非确定标签，提出最大置信度提升算法，并且使其能够处理无标签数据，从理论上证明了这是一种推广的自适应提升算法；最后在集成跟踪的框架下提出基于最大置信度提升算法的目标跟踪方法。实验结果表明，与传统的自适应提升跟踪算法相比，该方法提高了跟踪的精度，在复杂场景下跟踪的结果更加鲁棒。

第 6 章提出一种辅助子空间更新的自适应集成随机森林跟踪方法。该章首先

简要介绍随机森林和子空间学习的理论, 针对集成跟踪方法由于分类器误差积累带来的错误更新导致的跟踪漂移问题, 将随机森林嵌入集成跟踪框架, 并采用自组织的多视角融合的方式得到更准确的置信度图, 从而提高分类器的分类精度, 而子空间学习的更新方法则可以辅助分类器的更新, 减少分类器错误的特征更新, 维持分类器具有较强的判别性的功能。

第7章提出一个低秩重检测的多特征时空上下文的跟踪方法, 通过利用多特征融合的技术改善上下文的信息表达, 完善目标周围的结构信息, 通过有效的低秩矩阵近似分解的方法构建鲁棒的在线重检测器, 该重检测器可以保持历史跟踪信息结构的一致性, 可以在跟踪器跟踪失败后对目标周围进行目标的重定位, 实现跟踪方法的长时有效跟踪。从大量的实验评测中可以得出, 该算法能够有效地处理遮挡等问题引起的跟踪失败的重定位问题, 实验结果也表明该章的算法在跟踪精度、成功率和复杂度上取得了不错的性能, 在总体性能上超过了当前的一些先进的跟踪方法。

第8章对全书进行归纳, 并对未来的工作进行展望。

第2章 目标跟踪研究现状综述

2.1 引言

视觉跟踪技术在视觉监控、运动分析等计算机视觉领域中占有重要的地位。视觉跟踪是指对视频序列中需要跟踪的运动目标进行检测、跟踪,从而获得被跟踪对象的某些参数(如位置、形状、速度等)和运动轨迹的过程,这个过程为进一步的视频处理、分析和实现对运动目标的高级行为理解提供有力的支撑。

目标跟踪问题作为一个计算机视觉领域的热点研究问题,吸引了国内外很多研究机构进行了大量的研究工作。例如,在美国国防高级研究计划局(Defense Advanced Research Projects Agency, DARPA)的资助下,美国卡内基梅隆大学等几家著名的研究机构联合研制的视觉监控重大项目 VSAM (visual surveillance and monitoring, 视觉监控与追踪)^[2], 该项目的主要目的是利用视频理解、网络通信、多种传感器融合技术实现对未来城市、战场等场景的自动监控。美国马里兰大学开发的 W4 (Who? When? Where? What?)^[3] 实时视觉监控系统是 VSAM 的一个子项目, 该系统可以对人及肢体进行实时跟踪, 并且能对人们之间简单的交互行为进行监控。IBM 公司的 3S (Smart Surveillance System) 系统^[4] 为视频监控系统和基于视频的行为分析等应用提供了中间件 (middleware), 现在已经成功应用于美国国防某部门。英国雷丁大学开展了车辆的检测与跟踪^[5], 还与法国的 INRIA (法国国家信息与自动化研究所) 等研究机构联合实施了机场智能监控系统, 能够对停机坪场景下的运动目标进行跟踪并做出异常行为预警。此外, 国内的高校和研究所近年来也在运动目标跟踪领域开展了深入的研究, 并取得可喜的研究成果。

当前目标跟踪领域的学术交流非常活跃, 一些重要的计算机视觉会议上每年都能看到最新的视觉领域的研究成果, 如“国际计算机视觉会议”(IEEE International Conference on Computer Vision, ICCV)、“国际计算机视觉与模式识别会议”(IEEE Conference on Computer Vision and Pattern Recognition, CVPR)、“欧洲计算机视觉会议”(European Conference on Computer Vision, ECCV)、“国际模式识别会议”(International Conference on Pattern Recognition, ICPR)、“国际图像处理会议”(International Conference on Image Processing, ICIP) 等。国际上一些权威的学术期刊, 如《模式分析与机器智能汇刊》(IEEE Transactions on Pattern Analysis and Machine Intelligence, TPAMI)、《计算机视觉国际期刊》(International

Journal of Computer Vision, IJCV)、《图像处理汇刊》(*IEEE Transactions on Image Processing, TIP*)、《视频技术的电路与系统汇刊》(*IEEE Transactions on Circuits and Systems for Video Technology, TCSVT*)、《模式识别》(*Pattern Recognition, PR*)、《计算机视觉与图像理解》(*Computer Vision and Image Understanding, CVIU*)、《图像与视觉计算》(*Image and Vision Computing, IVC*)，以及国内重要的学术期刊，如《自动化学报》《计算机学报》《电子学报》等都大量刊登了代表该领域前沿的视觉跟踪的论文。

2.2 目标跟踪方法的国内外综述

尽管目标跟踪经历了几十年的发展，但该领域仍然在不断发展，目前并没有一个统一的理论框架体系，但是大多数的算法都要围绕两个重要的问题：建立目标表观模型和选择跟踪的方法策略。本章也将从目标的表观建模和跟踪方法进行分析，并对本书所要涉及的内容进行重点阐述。

2.2.1 常用的目标表观模型

建立一个鲁棒的目标表观模型是实现鲁棒的目标跟踪的一个至关重要的条件，应该说在跟踪策略确定的情况下，表观模型决定了跟踪的效果。以下将从表观模型这个角度来分析国内外的研究现状。

颜色直方图^[6-8]因其概念清晰、计算便捷，以及对目标的尺度变化、旋转、形变比较鲁棒的优点，被广泛应用在目标的表观模型中。但是这种基于统计学的概念，忽略了颜色的空间分布，导致其无法将两个颜色组成成分相似但颜色空间分布相差较大的目标区分开来，如图 2-1 所示。

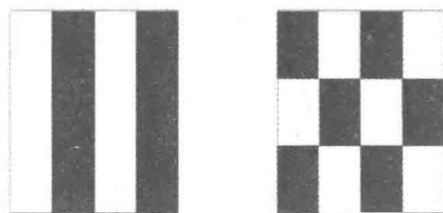


图 2-1 成分相同，分布不同的直方图

除了颜色直方图，Dalal 等^[9]提出的梯度方向直方图 (histograms of oriented gradients, HOG) 非常有影响力。HOG 特征描述了图像局部区域的梯度强度和梯度方向的分布情况，该分布情况能对局部对象的外观和形状进行很好的表征，在行人检测中得到了广泛的应用。当然为了克服颜色直方图丢失空间信息的这种缺陷，在随后的工作^[10-13]中提出了基于核密度估计的表观模型，该模型通过一些空间加权的核来描述目标颜色的空间分布，但是这种模型的主要缺点是计算量大、

复杂度高。在过去的一段时间内,混合高斯表观模型^[14]非常流行并广泛应用在计算机视觉领域和模式识别领域。这类模型利用混合加权的高斯模型来动态地学习像素颜色的时间统计特性。Jepson 等^[15]将目标表示成由三个分量组成的混合高斯分布,包括稳定的表观特征、瞬间特征和噪声过程。其中稳定分量表观特征中变化缓慢的目标区域,瞬间特征代表快速变化的像素点,噪声分量则描述目标表观分布中由噪声引起的外点,将稳定分量赋予更高的权重,可以有效地解决目标的局部表观变化,可以使跟踪更鲁棒,然而这类模型没有考虑像素的空间分布信息。Zhou 等^[16]提出了一个基于混合高斯的融合空间和颜色的自适应表观模型,采用粒子滤波的方法替代了文献[15]中的最大期望(expectation maximization, EM)算法,但是,它对目标表观的全局性变化及噪声非常敏感。为了提高表观模型对目标的刻画性能,目标表观的建模也在不断创新。近来,Lim^[17]等采用子空间学习的方法描述像素的全局统计信息,将图像展开成向量基的集合来表达目标,目标表观的空间信息及像素的邻域关系在这种表达方式中丧失殆尽,从而使模型对目标表观的全局性变化及噪声非常敏感。

这些基于图像区域本身的表观模型构建方法虽然具有操作容易和执行效率高等优点,但由于没有考虑表观像素之间的时空相关信息,不可避免地存在对光照、遮挡及形变等复杂因素敏感的致命缺点,进而导致在实际应用中非常不可靠。为了弥补表观模型中空间领域信息的丢失,文献[18]提出一种利用二维线性判别性分析(2 dimensional linear discriminant analysis, 2DLDA)直接在图像矩阵操作的基础上进行分类,而不是像前述的子空间学习那样把图像矩阵拉伸为向量。文献[19, 20]提出一种叫作张量子空间分析的学习算法,该算法通过学习一个低阶的张量子空间来刻画张量空间的内在局部几何结构,并给出了基于张量子空间学习的一个收敛的解。文献[21]提出将多种底层特征的协方差矩阵作为目标表观模型,协方差矩阵保留了目标的空间、统计及特征的相关信息,使特征之间的融合变得容易。文献[22]在张量与子空间的基础上提出一种用三阶张量表示视频序列段的张量子空间方法,算法保留了图像序列中空间与时间的连贯性,但是其三阶的表示算法泛化能力弱,且运算量较大。同时在处理目标遮挡的问题中,目标分块的表观模型也经常用到,这种表达方式也充分利用了像素的时空信息。文献[23]提出一种分块表达的目标表观模型,它将目标模板分成若干个不基于任何目标模型的图像块(patch),用每个小块对当前帧的目标位置和尺寸进行投票,这些投票融合得到最终的目标位置,因为这种分块方法利用了目标的空间分布,所以对局部遮挡非常有效。文献[24]将目标区域分成几个可以活动的矩形块(rectangle),每个矩形块的位置由目标的预分割结果决定,然后自适应调节这些块的位置,目标的表观模型可以由这些矩形块的直方图更准确地反映出来,对表达形状不规则以及形变较大的目标效果良好。

上述的表观模型大多数仅仅考虑了所关注的目标的特征,很少考虑目标周围

的背景模式及目标与背景之间的关系等，特别是当前景目标与背景之间存在某种相似性时，由此建立的表现模型往往不是最具判别性的。文献[25]提出一种同时考虑前景与背景变化比率的表现模型，该模型首先计算前景与背景条件概率的对数似然比，然后通过排序选择最佳判别性特征作为目标的表现模型，图 2-2 所示是这个构建过程的示意图。

总之，各种表现模型在某一个方面拥有自己的优势，同时也存在自身的缺点，所以面对不同的问题要采用不同的措施。因此，要实现复杂场景下鲁棒的目标跟踪算法，就要构建一种鲁棒的表现模型，以便能够全局地考虑目标的分布信息以及目标在其近邻区域的相对信息。

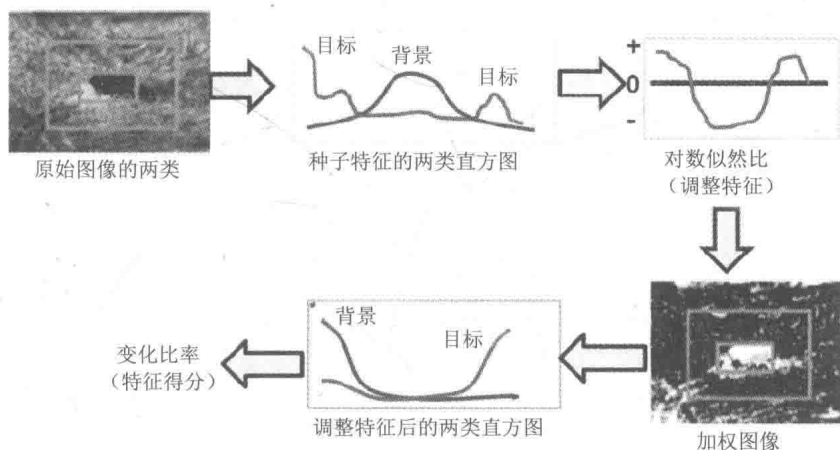


图 2-2 判别性表现特征构建过程示意图

“原始图像的两类”指的是背景的直方图和前景的直方图两类

2.2.2 常见的目标跟踪方法

对于目标跟踪的方法，不同的文献有不同的分类方法^[26-28]。根据跟踪采用的目标描述方式可以大致分为基于模型的跟踪、基于轮廓的跟踪、基于特征的跟踪和基于检测/分类的跟踪四类；如果从跟踪目标的定位策略出发，目标跟踪可以分成确定性搜索策略的跟踪方法和随机性搜索策略的跟踪方法两类。下面将从这两个角度出发对视频跟踪的关键技术的研究现状进行分析。

2.2.2.1 根据目标描述方式的分类

(1) 基于模型的跟踪

该类跟踪方法的基本思想是，首先将目标表现表示为特定的模板，该目标模板一般用由若干个简单几何形状拟合成的几何图形或目标模型函数表示，然后在当前帧采用穷尽搜索寻找与前一帧的目标模板最匹配的区域作为目标区