



杰出青年学者研究文丛
Distinguished Young Scholars

复杂场景下目标跟踪的 理论与方法

Theories and Methodologies of
Object Tracking under Complex Scenes

◎ 张笑钦 著



ZHEJIANG UNIVERSITY PRESS
浙江大学出版社



杰出青年

复杂场景下目标跟踪的 理论与方法

Theories and Methodologies of
Object Tracking under Complex Scenes

◎ 张笑钦 著



ZHEJIANG UNIVERSITY PRESS
浙江大学出版社

图书在版编目 (CIP) 数据

复杂场景下目标跟踪的理论与方法 / 张笑钦著.

—杭州:浙江大学出版社, 2017. 5

ISBN 978-7-308-16318-7

I. ①复… II. ①张… III. ①目标跟踪—研究 IV.
①TN953

中国版本图书馆 CIP 数据核字 (2016) 第 246344 号

复杂场景下目标跟踪的理论与方法

Fuza Changjing Xia Mubiao Genzong De Lilun Yu Fangfa

张笑钦 著

-
- 责任编辑 张凌静 (zlj@zju.edu.cn)
文字编辑 赵黎丽
责任校对 杨利军 刘 郡
封面设计 刘依群
出版发行 浙江大学出版社
(杭州市天目山路 148 号 邮政编码 310007)
(网址: <http://www.zjupress.com>)
排 版 杭州中大图文设计有限公司
印 刷 杭州钱江彩色印务有限公司
开 本 710mm×1000mm 1/16
印 张 8
彩 插 12
字 数 170 千
版 印 次 2017 年 5 月第 1 版 2017 年 5 月第 1 次印刷
书 号 ISBN 978-7-308-16318-7
定 价 49.00 元
-

版权所有 翻印必究 印装差错 负责调换

浙江大学出版社发行中心联系方式: 0571-88925591; <http://zjdxcs.tmall.com>

浙江省“十三五”重点出版物出版规划项目

前 言

随着计算机和摄像头的普及,计算机视觉技术正在成为研究热点。计算机视觉的终极目标是让计算机代替人的大脑,让摄像头代替人的眼睛,由计算机智能地分析从摄像头中获取的图像序列,对图像序列中的内容进行理解。其中最为关键的步骤是对场景中感兴趣的目标进行跟踪。目标视觉跟踪的任务就是在连续的图像序列中对运动目标进行定位,为后续行为分析提供目标轨迹、运动参数等信息。它在计算机视觉技术中起着至关重要的作用,不仅具有重要的理论研究价值,而且还具有广泛的应用前景和潜在的实用价值。

本书是一本关于复杂场景下目标视觉跟踪理论与方法的专著,书中从目标表观建模、跟踪框架这两个方面对目标视觉跟踪的最新研究成果进行介绍。全书分为两个部分:第1部分主要介绍目标表观建模方面的研究,第2部分主要介绍目标视觉跟踪框架方面的研究。具体来说,第1章介绍了目标视觉跟踪的意义、难点及其在视觉监控、智能交通、人机交互、虚拟现实等方面的应用。随后,分别从目标的表观建模和跟踪框架这两方面展开综述。第2章介绍了一种基于图嵌入判别学习的表观模型,包括基于图嵌入的特征降维、表观模型的构建与更新。第3章介绍了基于张量子空间学习的表观模型,包括张量代数下子空间的计算及在线更新方法。第4章介绍了对数欧氏黎曼空间下目标特征的提取与表示,并在此基础上建立增量量子空间学习的表观模型。第5章介绍了一种将核方法与贝叶斯方法结合的目标跟踪框架。第6章介绍了一种基于奇异值分解的卡尔曼粒子滤波跟踪框架。第7章介绍了一种基于序列粒子群优化的跟踪算法框架,并从理论上证明序列粒子群优化框架实质上是一个基于分步重要性采样的粒子滤波过程。

本书可作为高等学校计算机、自动化、电子信息等专业高年级学生或研究生的教学参考书,对广大从事视觉监控、智能交通、人机交互、虚拟现实等技术的研究、应用和开发的科技人员也有较大的参考价值。

本书由张笑钦制定编写大纲,初稿的第3~4章由李玺撰写,其余章节由张笑钦撰写。全书由张笑钦修改定稿。

该学术专著受到了国家自然科学基金“基于多特征联合稀疏表示和低秩张量恢复的视觉跟踪研究(61472285)”“面向低帧率视频的目标跟踪研究(61100147)”,以及浙江省自然科学基金——杰出青年科学基金“基于 L_p 范数-多任务稀疏表示的目标跟踪研究(LR17F030001)”的资助。本书在编写过程中,得到了中国科学院自动化研究所胡卫明研究员、浙江大学张仲非教授、伦敦大学 Stephen Maybank 等教授的大力支持和指导。李玺、李威、罗文寒、史信楚等参与了相关课题的研究,做出了重要贡献。中国科学院自动化研究所模式识别国家重点实验室为作者开展目标视觉跟踪研究提供了良好的平台,在此一并表示衷心的感谢!

由于作者水平有限,书中错误与疏漏之处在所难免,恳请读者批评指正。

张笑钦

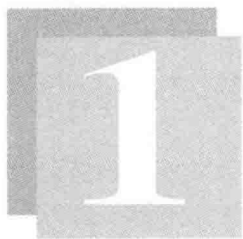
2016年8月

目 录

1 绪 论	1
1.1 意 义	1
1.2 研究现状	3
1.2.1 表观模型	4
1.2.2 跟踪框架	5
1.3 本书的结构	7
2 基于图嵌入判别学习的表观模型	10
2.1 基于图嵌入的学习框架	11
2.1.1 图嵌入	11
2.1.2 图嵌入的降维学习	11
2.2 基于图嵌入学习的目标跟踪	13
2.2.1 问题描述	13
2.2.2 基于样本分布的图嵌入	14
2.2.3 跟踪算法	17
2.3 实验结果	20
3 基于增量张量子空间学习的表观模型	25
3.1 张量代数	25
3.2 张量分解	27
3.3 主成分分析	28
3.4 增量张量子空间学习	33
3.5 基于张量子空间学习的目标跟踪	35
3.6 实验评估	37
4 基于对数欧氏黎曼子空间学习的表观模型	43
4.1 黎曼流形和黎曼度量	43

4.2	协方差矩阵描述子	47
4.3	基于对数欧氏黎曼子空间学习的目标跟踪	47
4.3.1	目标跟踪框架概述	48
4.3.2	目标表达	48
4.3.3	增量对数欧氏黎曼子空间学习	50
4.3.4	贝叶斯目标状态估计	51
4.3.5	实验评估	52
4.4	基于空间对数欧氏表观模型的目标跟踪	57
4.4.1	目标表观分块	58
4.4.2	在线对数欧氏特征空间学习	59
4.4.3	局部空间滤波	60
4.4.4	全局空间滤波	60
4.4.5	相似度匹配	61
4.4.6	贝叶斯目标状态估计	62
4.4.7	实验评估	62
5	基于核贝叶斯的跟踪框架	66
5.1	核贝叶斯框架	67
5.1.1	核方法	67
5.1.2	贝叶斯方法	68
5.1.3	核贝叶斯方法	69
5.1.4	讨论和分析	70
5.2	核贝叶斯跟踪算法	70
5.2.1	基于空间约束混合高斯的表观模型	71
5.2.2	核贝叶斯跟踪框架	72
5.2.3	表观模型的选择性更新	73
5.2.4	实验结果	74
6	基于奇异值分解的卡尔曼粒子滤波框架	77
6.1	无迹卡尔曼粒子滤波算法	77
6.1.1	无迹卡尔曼滤波器	77
6.1.2	无迹卡尔曼粒子滤波算法	78
6.2	基于奇异值分解的卡尔曼粒子滤波算法	79
6.3	基于增量子空间学习的表观模型	80

6.4 实验结果	81
7 基于序列粒子群优化的跟踪框架	83
7.1 基于序列粒子群优化算法的单目标跟踪	83
7.1.1 粒子群优化算法	84
7.1.2 序列粒子群优化算法	86
7.1.3 算法分析	90
7.1.4 跟踪算法	92
7.1.5 实验结果	93
7.2 基于多种群粒子群优化算法的多目标跟踪	96
7.2.1 动机	96
7.2.2 多目标跟踪算法	97
7.2.3 算法理论分析	102
7.2.4 实验结果	105
参考文献	113
附录 彩图	123
索 引	145



绪 论

1.1 意 义

目标视觉跟踪的目的是通过对摄像机所拍摄的视频图像序列进行处理,检测、定位和跟踪其中运动的目标。它是一个模式识别、计算机视觉、图像处理、人工智能、机器学习、概率论及统计分析等多学科交叉的前沿性研究课题。尽管经过十多年的研究,不同研究者提出了很多跟踪算法,然而由于目标表观的变化和环境的干扰,复杂场景下鲁棒和长期的目标跟踪依旧是一个研究难题,这使得目标视觉跟踪具有较大的理论研究价值。另一方面,目标视觉跟踪研究具有极其广泛的应用前景,简单归纳起来主要有如下 5 个方面。

(1)视觉监控。对视觉监控系统的需求主要来自那些对安全要求敏感的场所,如机场、银行、商店等。尽管目前监控摄像机在商业应用中已经普遍存在,但并未充分发挥其实时主动的监督媒介的作用。因为目前的做法通常是将摄像机的输出结果简单地记录下来,当异常现象发生后(如停车场中的车辆被盗),安保人员才通过查看摄像机的监视记录来获取当时的信息,但往往为时已晚。许多场合需要的是每天 24 小时不间断地实时监视,并自动分析摄像机捕捉的图像数据;当异常现象发生时,能够给监视人员以准确的警报,且要求摄像机能够自动运动,使得光心对准异常情况发生的地点,并调整摄像机焦距,使得图像最为清晰,从而让监视人员能够迅速做出反应,避免一些不正常事件的发生;同时,也减少雇用大批监视人员所需要

的人力、物力和财力的投入。要实现上述的智能监控,第一步就是要实现对场景中的运动目标(人或者车辆)进行准确的跟踪。只有在跟踪的基础上才可以进行更高层次的分析,如行为识别、异常检测等。另外,目标视觉跟踪在自动售货机、ATM、公共场所行人的拥挤状态及商店中消费者流量统计等监控方面都有着相应的应用。

(2)智能交通。从2000年开始,我国主要城市的交通状况开始恶化,机动车数量的迅速增长给交通运输和管理带来了巨大挑战;道路阻塞问题日益严重,交通事故频繁发生,这些都使交通建设面临巨大压力。因此,智能交通系统已成为我国城市交通运输的研究重点和发展方向。它通过对交通视频中车辆和人的跟踪和轨迹分析,来确定其行为是否合理。对于检测到的不合法的行为进行报警,并通过实时拍照等方式记录违章行为,为交通监管部门进行道路管理提供依据。同时,它也可以实时获取通行车速、交通流量、道路通行能力等信息,为智能交通系统实施交通调度、诱导及动态规划交通信号配时提供决策参考。

(3)感知接口。在高级用户接口应用领域中,我们希望未来的机器能像人一样与我们更加容易和便捷地交流,如手势驱动控制、手语翻译等。人与人之间的信息交流主要依靠语言,并适当结合手势、身体姿势和面部表情等,故视觉信息可以作为语音和自然语言理解的有效补充来完成更加自然和智能的人机交互。未来的计算机应该具备感知外部环境的能力,即代替传统的键盘输入模式,独立地提取周围环境的重要信息(如检测到人的存在等);能够通过对人的手势和姿态进行跟踪,以及对行为和轨迹的分析来与人进行相应的信息交流。对于机场等高噪声的场合,基于视觉的高级用户接口也能够提供比语音识别更加准确的信息输入。

(4)运动分析。分割图像中的人体部分,在图像序列中跟踪并分析感兴趣的关节运动,对于建立人体的几何模型、解释人体的运动行为机制,从而提高人体的运动机能等都有着积极的作用。研究结论可以应用于指导体育运动、舞蹈训练等。传统的步态运动分析属于肌电图学、关节动力学的研究领域,目前它已经被广泛应用于医疗诊断分析中。医学步态分析是一个旨在提供诊断和治疗支持的研究领域,可以提供人体正常步态建模的线索,有助于开发生物反馈系统来分析病人的步态,判断其腿部受伤情况或者畸形程度,从而采取积极的整形补偿或其他有效的治疗手段。

(5)虚拟现实。目前,许多电脑游戏中目标动画的形体、运动和行为交互设计的逼真性实际得益于物理空间中目标的运动分析,包括目标骨架模型、关节运动机制的获取、姿势的恢复等。基于互联网络的交互式虚拟空间

的开发才刚刚起步,如虚拟聊天室在通过文本交流的同时可以通过二维图标来导航用户,如果增加手势、头的姿势、面部表情等线索,将会给参与者提供更加丰富的交互形式,而这一切都是以跟踪结果为基础的。另外,目标视觉跟踪在其他虚拟现实场合,如视频会议、动画、虚拟工作室等,也有相应的应用前景。

由于目标视觉跟踪有着广泛的应用前景,国内外很多研究机构已经开展了大量相关项目的研究。例如,1997年美国国防部高级研究项目署(Defense Advanced Research Projects Agency)设立了视觉监控重大项目VSAM(Visual Surveillance and Monitoring),主要研究用于战场及普通民用场景监控的自动视频理解技术;英国的雷丁大学(University of Reading)开展了对车辆和行人的跟踪及其交互作用识别的相关研究^[1];英国利兹大学也同时开展了基于模型的行人监控系统的研究^[2]。在德国的法兰克福理工学院,Haag和Nagel^[3]从事交通监控系统的研究。IEEE从1998年起资助了国际视觉监控系统研讨会。在硬件设备方面,国外许多公司设计了各种专用设备,如Action camera^[4], Smart camera^[5], Omnidirectional camera^[6],等等。国内的中科院自动化所模式识别国家重点实验室(National Laboratory of Pattern Recognition)开发的Vstar演示系统,对交通场景中的车辆进行实时跟踪与异常行为检测。当前,国际上一些权威期刊,如IEEE TPAMI、IJCV、CVIU、IEEE TCSVT、IVC和重要的学术会议,如ICCV、CVPR、ECCV、ACCV、ICIP、ICPR、IWVS等,将目标视觉跟踪研究作为主题内容之一,为该领域的研究人员提供了更多的交流机会。复杂场景下的目标视觉跟踪受到了世界上许多重要研究机构的高度重视,这从另一方面说明了它的重要性。

1.2 研究现状

跟踪问题等价于在连续的图像帧间创建基于位置、速度、形状、纹理、色彩等有关特征的对应匹配问题。大体上,目标跟踪的算法都要涉及两个关键性问题:

表观模型——如何对目标进行建模,并且实时更新;

跟踪框架——如何对候选区域与目标模板进行有效的匹配。

以下我们将从这两个方面分析国内外的研究现状。

1.2.1 表观模型

目标区域的颜色直方图^[7-8]通常被用来刻画目标的表观特性。由于颜色直方图提取简便,并且对目标尺度、旋转以及非刚性形变较为鲁棒,从而被研究人员广泛地使用。然而,颜色直方图存在一个潜在的问题,那就是它完全忽略了目标表观颜色的空间分布信息。这样导致的结果就是当两个目标的颜色组成成分非常相似但颜色空间分布相差比较大的时候,颜色直方图无法区分这两个目标。为了克服颜色直方图丢失空间信息的缺陷,基于核密度估计的表观模型应运而生^[9-10]。该模型的优点是通过运用一些空间加权的核来刻画目标的颜色空间分布信息。不幸的是,这种类型的模型有一个共同的缺点,那就是计算和存储复杂度非常高,以至于不能被广泛使用。近年来,混合高斯表观模型^[11-17]非常流行,并且被广泛地应用于计算机视觉和模式识别领域。基于混合高斯的表观模型利用混合加权的高斯来动态地学习像素颜色的时间统计特性。然而,这种模型只是独立地处理每个像素,并没有考虑像素间的内在联系,从而使模型对噪音非常敏感。Wang 等人^[18]提出基于混合高斯的融合空间和颜色的自适应表观模型,该模型能够获取丰富的局部空间和颜色分布信息。但是,该模型也存在一个较大的缺点,那就是像素的全局时空统计特性不能被该模型有效地刻画,从而导致对目标表观的全局性的变化以及噪音非常敏感。另外,基于条件随机场的表观模型^[19]通过马尔可夫随机场来建模邻近像素之间的内在关系,但是这样一来,模型的训练代价非常巨大,同时也不能考虑像素的全局分布信息。近年来,基于在线子空间学习的方法^[20]被用来描述像素的全局统计信息,但该类方法首先要将图像展成向量,使得目标表观的空间信息几乎完全丢失,这样使得模型对目标表观的全局性变化以及噪音非常敏感。

最近,很多人使用基于张量的子空间方法来建模目标表观的时空统计特性,这样能够极大地减少目标的时空分布冗余,从而能构建一个紧密的目标表观模型。但是目前的基于张量的目标表达方法大多数是离线操作的,它们不能在线更新,从而导致很大的运算和存储代价的付出。这里我们简要地回顾一下最近相关文献中基于张量的目标建模方法。Yang 等人^[21]提出用 2DPCA(二维主成分分析)来表征一幅图像。Ye 等人^[22]提出 2DLDA(二维线性判别分析方法),该方法直接在对图像矩阵操作的基础上进行分类,而不需要把图像展成向量。Ye^[23]提出一种 GLRAM 算法,该算法可以构建一个低阶的模型来逼近一个序列矩阵。Ye 等人^[24]提出一种 GPCA(广

义主成分分析)的降维算法,该算法直接对图像集构建一个低阶的矩阵表示。Wang 和 Ahja^[25]提出一种 R 阶张量分解方法,该方法能够较好地减少张量的时空冗余。Yan 等人^[26]开发了一种张量的判别分析方法,该方法将传统的判别分析方法扩展到了张量空间。在 Vasilescu 和 Terzopoulos 的研究^[27-28]中,图像集不同模态的奇异值分解被用来构建人脸图像的紧致表达。He 等人^[29]提出了一种叫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。Wang 等人^[30]给出了针对一般基于张量子空间学习的一个收敛的解。Sun 等人^[31]提出动态和流张量分析来挖掘高阶的数据流。同样在 Sun 等人^[32]的研究中,基于窗的张量分析方法被用来建模实时的数据流。综观以上基于张量的方法,它们存在一个共同的问题,即它们不能在线自适应地更新基于张量的特征空间(均值和特征基),因此无法将这种有效的建模方法应用于时序数据,如视频流数据,从而无法将该方法有效地应用到视觉跟踪领域。

上述的目标跟踪方法所使用的表观模型通常只考虑到正样本(目标物体)模式,如通过模板、直方图、混合高斯模型或者张量表示等方法,很少考虑正样本周围的负样本(背景)的模式以及正、负样本之间的局部模式等。因此产生的表观模型往往不是最具判别性的,进而导致在实际应用中对光照、遮挡以及形变等复杂因素不够鲁棒。如何设计一种表观模型,能够有效地抓住目标样本的分布信息以及目标样本与其近邻区域中背景样本的相对分布信息,是设计鲁棒跟踪器的关键所在。

1.2.2 跟踪框架

对于跟踪的框架而言,从优化的角度来说,目前存在的跟踪算法的理论框架大致可以分为两类:一类称为确定性(deterministic)跟踪框架,另一类称为随机性(stochastic)跟踪框架。下面我们将进一步介绍这两种跟踪框架。

1.2.2.1 确定性跟踪框架

典型的确定性跟踪框架根据事先设定的相似度损失函数(similarity cost function),在当前图像帧内的局部区域中通过梯度优化的迭代方法找到与目标模板最为相似的区域。较为广泛使用的损失函数是误差平方和函数(sum of squared difference, SSD)^[33]及基于核直方图^[34-36]的损失函数。其中,SSD 损失函数是基于模板与当前图像的像素灰度差的平方和,并通过

迭代优化获得目标的运动参数;基于核直方图的损失函数是衡量两个核直方图之间的相似度,并采用均值漂移(mean shift)来实现对帧间目标偏移量的估计。目前,更为鲁棒的损失函数不断被提出,如基于混合高斯的损失函数^[37]、基于子空间学习^[38]的损失函数。同时,诸如模拟退火^[39]、遗传算法^[40]等其他优化方法也被用来搜索与目标模板最为相似的局部最优区域。总的来说,确定性跟踪框架一般计算量较小,算法收敛快,效率较高,但容易陷入局部极小值,导致跟踪的结果不够精确。随着跟踪误差的积累,最后可能导致跟踪失败。

1.2.2.2 随机性跟踪框架

随机性跟踪框架将跟踪问题看作概率推理(probability inference)问题,目标是通过已知物体的先验知识,结合观测结果来估计物体的状态,即转化为最大化后验概率问题。技术上的困难在于如何得到关于目标状态的后验概率的准确描述。

跟踪过程由预测和校正两步组成。

(1) 预测:根据前 $i-1$ 帧的观测值(用 $y_0 \cdots y_{i-1}$ 表示)预测第 i 帧目标的状态(用 x_i 表示),即有如下表达式:

$$\begin{aligned}
 & p(x_i | y_0 \cdots y_{i-1}) \\
 &= \int p(x_i, x_{i-1} | y_0 \cdots y_{i-1}) dx_{i-1} \\
 &= \int p(x_i | x_{i-1}, y_0 \cdots y_{i-1}) p(x_{i-1} | y_0 \cdots y_{i-1}) dx_{i-1} \\
 &= \int p(x_i | x_{i-1}) p(x_{i-1} | y_0 \cdots y_{i-1}) dx_{i-1} \\
 &\propto p(x_i | x_{i-1}) p(x_{i-1} | y_0 \cdots y_{i-1}) \quad (1-1)
 \end{aligned}$$

(2) 校正:结合第 i 帧的观测值修正状态的预测值, i 时刻的后验概率 $p(x_i | y_0 \cdots y_i)$ 由下式表达:

$$\begin{aligned}
 & p(x_i | y_0 \cdots y_i) \\
 &= \frac{p(y_i | x_i, y_0 \cdots y_{i-1}) p(x_i | y_0 \cdots y_{i-1}) p(y_0 \cdots y_{i-1})}{p(y_0 \cdots y_i)} \\
 &= \frac{p(y_i | x_i, y_0 \cdots y_{i-1}) p(x_i | y_0 \cdots y_{i-1})}{\int p(y_i, x_i | y_0 \cdots y_{i-1}) dx_i} \\
 &= \frac{p(y_i | x_i, y_0 \cdots y_{i-1}) p(x_i | y_0 \cdots y_{i-1})}{\int p(y_i | x_i, y_0 \cdots y_{i-1}) p(x_i | y_0 \cdots y_{i-1}) dx_i} \\
 &\propto p(y_i | x_i) p(x_i | y_0 \cdots y_{i-1}) \quad (1-2)
 \end{aligned}$$

上述推导过程中的假设条件请参考 Arulampalam 等人的工作^[41]。其中有两个重要的先验模型:一个是预测过程中的 $p(x_i | x_{i-1})$,它是用来描述系统动态特性的运动模型;另一个是校正过程中的观测模型 $p(y_i | x_i)$,用来描述观测值与模型的相似程度。由于运动模型和观测模型不同,适用的跟踪滤波器也不同。当两者都服从高斯分布且为线性关系时,可以利用卡尔曼滤波(Kalman filter)^[42]来获得后验概率的解析表达。当两者为非线性关系时,可以利用扩展卡尔曼滤波(extended Kalman filter, EKF)^[43]求解后验概率,此时后验概率仍然满足高斯分布。一种替代 EKF 的方法是无损卡尔曼滤波(unscented Kalman filter, UKF)^[44]。如果状态空间是离散的,而且是由有限个状态组成的,则可以使用隐马尔可夫模型^[45]进行跟踪。但是,在实际的视觉跟踪过程中,后验概率的分布往往是非线性、非高斯、多模态的,此时序贯蒙特卡罗方法(sequential Monte Carlo methods, SMC)^[46]被提出用于解决这类情况,其原理是利用重要性采样来逼近后验分布。该方法在刚提出之时并未在视觉领域引起很大关注,直到 1996 年,由 Isard 和 Blake^[47]提出 condensation 算法之后,才引起广泛注意。目前,这一类方法又被称为粒子滤波(particle filter)^[48],是当前视觉跟踪方法中的研究热点。另外,UKF 思想也可被粒子滤波方法借用,这种情况下粒子滤波被称为无损粒子滤波(unscented particle filter, UPF)^[49]。

与确定性跟踪框架相比,随机性跟踪框架更为鲁棒,但是计算量过大,而且随着目标个数的增长呈指数增长。另外,由于没有好的机制选择重要性分布,粒子滤波存在着严重的样本退化问题。确定性跟踪框架和随机性跟踪框架有着各自的优点和缺点,若将确定性跟踪框架和随机性跟踪框架这两种思路结合起来使用,就有可能克服各自的弱点,实现更为有效的跟踪。

1.3 本书的结构

第 1 章大致回顾了运动目标跟踪的研究现状。具体而言,在这一章重点介绍了两个方面的相关工作:常用的表观模型和现有的跟踪框架。

第 2 章提出了一个基于图嵌入判别学习的表观模型,并将其应用到移动摄像机的目标跟踪。该算法将目标跟踪视为一个前景与背景的分类问题。首先通过初始化获得正样本及一系列负样本,并且构建正样本、负样本以及

正、负样本关系的图结构,然后通过图嵌入理论学习正样本的模式空间,以及正、负样本之间的判别空间。基于对这两个模式空间的学习,我们构建鲁棒的表现模型,并通过该表现模型获得样本的置信概率,将样本的最大后验概率作为跟踪结果。最后通过基于子空间的启发式策略选择正样本和负样本,并增量地更新嵌入图的结构。另外,针对粒子滤波缺乏对目标运动有效的估计,本章提出了级联的粒子滤波框架,即在粒子滤波器的前端级联一个多尺度的模板匹配模型来提供对目标运动的有效估计。实验结果表明,与基于增量子空间方法和基于 Fisher 判别的方法相比,该表现模型有效地提高了跟踪的效率和精确度,并且可以有效地处理局部遮挡及高噪声情况下的跟踪。

第 3 章采用了基于张量子空间学习的方法来进行目标表现建模。张量子空间学习不仅能够有效地减少目标区域像素的时空冗余信息,而且也能够融合不同角度的时空信息,从而能够较好地处理噪音干扰、运动模糊、光照变化等复杂情况下的目标跟踪问题。另外,本章还提出了增量张量子空间学习方法,该方法能够在线更新基于张量特征空间的目标表现模型。该章采用粒子滤波器来动态传播样本分布,并利用上述张量子空间模型对样本进行评价,从而对目标状态进行有效的贝叶斯估计。

第 4 章提出了两个基于对数欧氏黎曼子空间学习的表现模型。这两个表现模型能够充分考虑目标表现的时空相关信息,从而有效地解决了在复杂因素影响下目标鲁棒跟踪的问题。具体地说,第一个表现模型首先提取一些图像特征(包括像素灰度值、一阶梯度、二阶梯度等);接着用这些特征进行自相关分析,也就是计算这些特征对应的协方差矩阵;然后将协方差矩阵映射到对数欧氏黎曼空间;最后再对对数欧氏黎曼空间进行线性子空间分析,从而得到基于对数欧氏黎曼特征空间的表现模型。第二个表现模型是在第一个表现模型的基础上进行表现分块和空间加权,进一步考虑目标表现的局部空间相关信息,因此对复杂的外界干扰因素非常鲁棒。另外,增量对数欧氏黎曼子空间学习的方法被用来在线更新这两个表现模型,从而使得它们能够适应外界条件的变化。此外,粒子滤波器被用来动态传播样本分布,从而对目标状态进行有效的贝叶斯估计。

第 5 章讨论了结合确定性跟踪框架和随机性跟踪框架的方法:核贝叶斯跟踪框架。首先将核方法应用于当前帧,获得对运动方向及其偏移量的估计,再将这些信息作为启发式先验信息加入贝叶斯滤波框架中,从而有效结合确定性方法与随机性方法的优点。此外,该框架是一种通用的跟踪框架,任何形式的表现模型都可以在此框架下应用。第 5 章采用了基于空间约束