

Topological Isomorphism and Target Tracking in Video

拓扑同构 与 视频目标跟踪

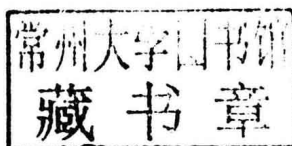
付维娜 于 萍 著

西安电子科技大学出版社
<http://www.xduph.com>

拓扑同构与视频目标跟踪

Topological Isomorphism and Target Tracking in Video

付维娜 于萍 著



西安电子科技大学出版社

内 容 简 介

本书是依据作者在博士学习期间的研究成果,并结合计算机视觉领域中视频目标跟踪问题的研究现状以及信息时代对视频信息处理的具体需要编写而成的。

全书共分五章:第1章介绍了视频目标跟踪的相关基础理论知识,论述了该研究对国家和人民生活的重要意义,同时对视频目标检测、视频目标跟踪的历史与研究现状进行了回顾与分析;第2章~第4章系统地介绍了视频监控中的移动目标跟踪方法,主要包括一种基于背景动态重建的视频移动目标检测方法和另一种结合目标颜色信息拓扑关系的目标跟踪方法,并进行实时性的具体实现;第5章是总结与展望。

本书可供计算机相关领域研究者学习和参考。

图书在版编目(CIP)数据

拓扑同构与视频目标跟踪/付维娜,于萍著. —西安:西安电子科技大学出版社,2018.5
ISBN 978-7-5606-4910-8

I. ① 拓… II. ① 付… ② 于… III. ① 目标跟踪—研究 IV. ① TN953

中国版本图书馆 CIP 数据核字(2018)第 070077 号

策 划 刘小莉

责任编辑 师 彬 阎 彬

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 陕西天意印务有限责任公司

版 次 2018年5月第1版 2018年5月第1次印刷

开 本 787毫米×960毫米 1/16 印 张 5

字 数 95千字

印 数 1~1000册

定 价 23.00元

ISBN 978-7-5606-4910-8/TN

XDUP 5212001-1

如有印装问题可调换

前 言

视频序列中的移动目标跟踪是计算机视觉领域的热点研究课题之一,在军事和民用领域都有广泛的应用,吸引了众多国内外学者的普遍关注。研究者提出了许多有效的跟踪算法,跟踪的性能也不断得到提高。然而,仍存在如非线性形变,目标姿态变化,前景与背景间的遮挡、交错,复杂背景中噪声干扰等各种影响跟踪的问题。因此,提高视频跟踪算法的鲁棒性、准确性仍然是一项充满挑战性的工作。此外,跟踪算法是否能够保证实时性也是需要研究的方面。

本书针对视频中移动目标检测和跟踪算法进行研究,主要完成以下三个方面的创新性工作:

(1) 针对目标检测和跟踪过程中的背景未知、背景复杂问题,提出一种基于背景动态重建的视频移动目标检测方法。该方法结合目标方向动态重建背景,并在新建立的背景基础上进行目标检测和跟踪,解决了现有方法中前景检测依赖已知背景的问题。并且,对背景进行形态学运算,加强了针对轻微抖动、亮度变化等复杂背景的抗干扰能力,提高了检测和跟踪的准确率和效率。

(2) 针对视频目标跟踪过程中目标的遮挡问题,提出一种结合目标颜色信息拓扑关系的目标跟踪方法。该方法将目标颜色位置拓扑关系作为新特征与其他特征进行融合以实现目标跟踪,解决了传统目标跟踪方法将颜色成分相同而位置不同的其他背景识别为目标的问题;通过对拓扑结构矩阵进行近似同构性的判断,解决了由于部分颜色信息被掩盖导致的识别错误问题,为此类特征融合的目标跟踪算法提供了新的思路;通过将多个特征弱分类器组成级联强分类器来建立目标判决模型,有效地解决了目标在运动中经常出现的单一特征缺失导致跟踪无法继续的问题,提高了视频中目标跟踪算法的有效性和准确性。

(3) 针对传统多移动目标跟踪方法计算量大,不能保证计算的实时性问题,提出一种视频实时多移动目标跟踪的分布式方法,并设计了合理的分布式调度算法。该分布式方法将前景按照其连通性分解成若干子目标进行跟踪,有效地降低了因为被跟踪目标过大、过多导致的时耗问题,提高了多移动目标跟踪的效率,达到了实时性标准。

综上，本书对视频目标检测和跟踪过程中的背景重建、多特征融合的目标跟踪方法及多目标跟踪的分布式方法等问题进行了研究，解决了目前大多数跟踪都需要依赖已知背景的局限，消除了背景复杂、成像设备抖动、亮度变化等情况对跟踪的不利影响；找到了一组能够较好地反映移动目标特点的特征，有效地避免了由于遮挡和其他原因导致的特征损耗和淹没；解决了多目标跟踪计算量过大的问题，取得了较好的时效性并实现了多移动目标的实时跟踪。

编 者
2018 年 1 月

目 录

第1章 相关背景与知识基础	1
1.1 视频目标跟踪研究的意义及背景	1
1.2 视频目标跟踪的主要应用领域	3
1.3 相关领域研究现状	4
1.3.1 视频目标检测技术	4
1.3.2 视频目标跟踪技术	10
1.4 视频移动跟踪过程中存在的难点问题	13
1.5 本书的主要工作与结构	15
第2章 基于背景重建的移动目标检测	17
2.1 基于背景差法的移动目标检测技术	17
2.2 结合目标方向的背景重建方法	18
2.2.1 视频图像预处理	18
2.2.2 基于三差帧法的目标移动模糊方向分析	19
2.2.3 动态背景重建方法	21
2.3 基于背景重建和移动特征的目标轨迹跟踪	23
2.4 算法实验与分析	23
2.4.1 室内场景的背景重建实验与检测跟踪实验	24
2.4.2 室外场景的背景重建实验与检测跟踪实验	25
2.5 本章小结	27
第3章 基于近似拓扑同构的多特征融合视频移动目标跟踪方法	28
3.1 视频移动目标跟踪技术	28
3.2 移动目标分割与目标拓扑特征的近似同构	31
3.2.1 色块分割	32
3.2.2 色块拓扑信息提取和近似拓扑同构	33
3.3 基于近似拓扑同构的多移动目标跟踪	37

3.3.1 特征提取	37
3.3.2 特征匹配	38
3.4 算法实验与分析	41
3.4.1 颜色块分割实验与分析.....	41
3.4.2 特征信息提取实验与分析.....	42
3.4.3 特征匹配和移动跟踪实验与分析.....	48
3.4.4 与经典目标跟踪算法比较实验与分析.....	50
3.5 本章小结	54
第4章 分布式视频移动目标跟踪	56
4.1 分布式视频目标跟踪技术	56
4.2 使用拓扑矩阵分割目标	56
4.2.1 前景目标提取	57
4.2.2 多目标分割	58
4.3 基于目标分割的动态分布式跟踪方法.....	60
4.4 算法实验与分析	62
4.4.1 移动特征实验与分析	62
4.4.2 颜色块分割实验与分析.....	64
4.4.3 动态分布式跟踪方法实验与分析.....	66
4.5 本章小结	70
第5章 总结与展望	71
5.1 总结	71
5.2 展望	72
参考文献	73

第1章 相关背景与知识基础

人类感知世界有多种途径,如视觉、听觉、触觉、嗅觉等,视觉是其中最重要的一种方式。俗话说:“百闻不如一见”,外界事物的绝大多数信息都是通过视觉获得的,而且视觉感知环境信息的效率很高。相对于其他类型信息(文字、声音等),视觉信息能够更加形象、逼真地反映这个世界。因此,人们希望计算机能够具有人类的部分视觉功能,帮助甚至代替人眼和大脑对外界事物进行观察和感知,赋予机器视觉是人类多年以来的梦想。随着微电子技术、通信技术和计算机技术的发展,人类可以通过计算机和成像设备对外部图像信息进行获取、处理、分析和模拟。由此,形成了一门新兴的学科——计算机视觉。计算机视觉领域已成为计算机相关研究中的一个活跃的、有潜力的研究领域。在计算机视觉的主要研究中,对视频图像中移动目标的检测、跟踪是一个重要的研究课题,尤其在大型公共场所的安全监控和危险预警方面具有广泛的应用背景。

本章首先介绍视频移动目标跟踪的研究背景及意义;随后给出视频跟踪的研究现状,并介绍当前视频目标跟踪算法的分类和主要技术手段,包括与视频跟踪领域相关的基本概念和方法;最后,在此基础上,总结视频目标跟踪研究中存在的主要问题以及目前的研究发展趋势。

1.1 视频目标跟踪研究的意义及背景

计算机视觉是指使用计算机代替人脑来观察和理解世界,是研究如何使人工系统从图像或多维数据中“感知”的科学,融合了包括视频图像处理、机器学习、人工智能、模式识别等多个领域的研究,吸引了众多国内外学者和研究机构的普遍关注和积极投入。计算机视觉最早是在20世纪50年代由统计模式识别理论发展起来的,最先主要研究二维图像的分析 and 识别,如光学字符识别,工件表面、显微图片和航空图片的分析及解释等。1965年,麻省理工学院的Roberts开创了以理解三维场景为目的的三维计算机视觉的研究。20世纪70年代中期,麻省理工学院的人工智能实验室开设了“机器视觉”(Machine Vision)

课程。从 20 世纪 80 年代开始,对计算机视觉的研究进入高潮,新的方法和理论不断出现,并进入到实际应用。而计算机工业水平的飞速提高以及人工智能、并行处理和神经网络等学科的发展,更加促进了计算机视觉理论的实用化。目前,计算机视觉技术正广泛地应用于图像处理、计算几何、机器人学等多个领域中。

典型的计算机视觉系统(Typical Computer Vision System)分为目标检测、目标跟踪、目标识别、目标行为理解与预测四个部分,如图 1.1 所示。目标检测是目标跟踪的基础,目标检测与目标跟踪又是目标识别和目标行为理解与预测的基础。一个成熟的计算机视觉模型至少具有检测、跟踪两个模块,识别模块或行为理解与预测模块并不是必需的。

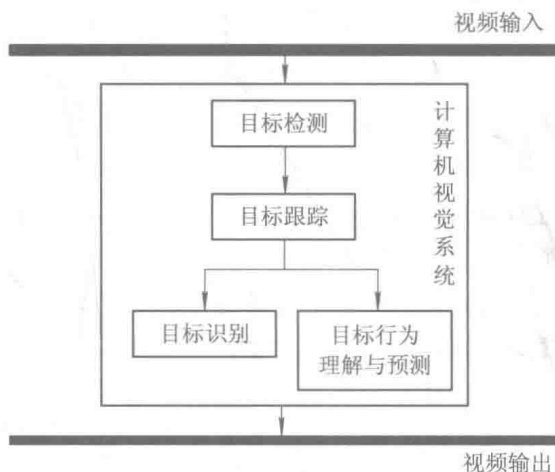


图 1.1 典型的计算机视觉系统

目标检测是计算机视觉领域的关键技术之一,是目标跟踪和行为理解等更高层次研究的基础。目标检测的目的是为了将用户感兴趣的区域准确地从视频序列中抽象出来,并清晰完整地进行描述,以便进行更深层的应用。

目标跟踪是在目标检测基础上进行的,甚至有时将目标检测作为目标跟踪的一部分。一个成熟的目标跟踪系统一定包含目标检测模块和目标跟踪模块。与目标检测一样,目标跟踪同样是计算机视觉领域研究的核心问题,是该领域内诸多其他研究,如目标识别、目标行为分析等的基础。只有准确、稳健、有效地跟踪目标,其他的技术,如行为分析、视频检索、危险预测等才能够进行下去。所谓视频移动目标跟踪,是指对视频图像序列中的特定目标进行检测、提取、识别和跟踪,获得目标的位置参数,如目标质心的位置、速度、加速度,或者目标整体所占的图像区域,或是目标的运动轨迹,等等,从而进行后续深入的处理与分析,以实现特定目标的行为理解,完成更高级的任务。

目标识别是指将一个特殊目标(或一种类型的目标)从其他目标(或其他类型的目标)中

区分出来的过程。它既包括两个非常相似目标的识别,也包括一种类型的目标同其他类型目标的识别。例如,从交通视频监控场景中寻找指定的车辆、人脸识别、指纹识别、手势识别等。

目标行为理解与预测是指在目标检测和跟踪的基础上对目标行为进行分析,并预测该目标的后续行为。目标行为理解与预测常常用于安保场合,如在地铁站、汽车站等公众场所某时段内忽然聚集了大量人群的危险报警、对手势含义的行为理解等。

1.2 视频目标跟踪的主要应用领域

作为计算机视觉领域里最重要的组成部分,视频移动目标检测与跟踪算法得到越来越多的关注,在军事和民用领域都有非常广泛的应用。在军事上,视频移动目标跟踪技术主要应用于精确制导系统(Precision Guide System, PGS)、战场机器人自主导航、无人机着陆引导、靶场光电跟踪等领域。使用目标跟踪技术可以在线、实时地获得目标的精确位置、面积、运动速度等有效信息,它是军事领域智能化环节中最重要的一环。目标检测与跟踪系统在民用领域的应用更加广泛,典型应用包括以下几个方面。

1. 智能视频监控(Intelligent Monitor System, IMS)

智能视频监控是一种全自动、全天候、实时监控的智能系统。通过在监控系统中增加智能视频分析模块,可自动识别不同物体,用来定位事故现场、判断异常情况,从而发出警报或触发其他动作,有效地进行预警、处理和取证。智能监控系统应用广泛,如银行、商场、工厂企业、大型赛事等对安保产品的需求日益增多。在视频监控领域,固定场景的多目标跟踪是研究的重点,尤其是目标之间的相互遮挡或目标的自遮挡对目标检测和跟踪的准确率造成了严重的负面影响。因此,对解决遮挡问题的研究已成为多目标跟踪的一个非常重要的子领域研究^[11]。

2. 智能交通控制(Intelligent Transportation System, ITS)

智能交通控制是一个基于现代电子信息技术的面向交通运输、车辆控制的服务系统。当今社会,随着汽车等交通工具的广泛普及,车辆数量越来越多,从而带来了交通拥堵等新问题。为了有效管理交通流量,智能交通系统越来越受到关注。智能交通控制可以收集、处理、发布、交换、分析、利用交通信息,为交通参与者提供多样性的服务。如可以通过分析摄像头获取的视频信息对交通流量进行控制;可以通过对车辆进行在线、实时的监测和跟踪,获得车辆的速度、车流的密度、道路的堵塞状况、交通事故记录等信息,并可进行违章逃逸车辆追踪、车牌识别、车辆的异常行为分析等。

3. 人机交互(Intelligent Human-Computer Interaction, HCI 或 Human-Machine Interaction, HMI)

人机交互是一门研究系统与用户之间的交互关系的学问,具体的应用有手势识别、唇语识别、人体姿势识别、人眼识别等。这些应用需要以视频目标跟踪技术为基础,通过跟踪移动目标并识别它的行为,才能构建语义和计算机指令之间的关系。

4. 智能医学诊断(Intelligent Medical Diagnosis, IMD)

智能医学诊断是指目标跟踪技术利用目标在几何上的连续性和时间上的相关性,去掉图像中的噪声,使诊断结果更加准确可靠。目标检测和跟踪技术在超声波和核磁共振序列图像的自动分析中有着重要的应用。

5. 智能视频检索(Intelligent Video Retrieval, IVR)

智能视频检索利用人工智能代替人工搜索,为视频分析提供了很大的帮助。它通过选取待搜索目标并提取其特征,能够在图像序列中跟踪选定目标,找到其移动轨迹。

6. 视频语义分析(Video Semantic Analysis, VSA)

在视频序列里,同一个移动物体在若干帧之间的关联有自己专属的移动轨迹,该轨迹使其区别于其他移动物体。移动跟踪在计算机图像语义分析的过程中非常重要,如何保持跟踪的正确率,是目前的一个重要研究课题。

此外,视频目标跟踪技术还在视频会议、三维重构等领域具有重要的应用价值。因此,对视频目标检测与跟踪技术的研究具有非常重要的意义。

1.3 相关领域研究现状

视频目标检测是视频目标跟踪的基础,本节主要从视频目标检测和视频目标跟踪这两个方面对目前的研究状况与相关技术进行介绍。

1.3.1 视频目标检测技术

目标检测的目的是将前景(运动目标)从背景中准确地分离出来,背景用来描述一个不包含感兴趣的运动物体的场景。目标检测可以被看做是视频中运动物体研究的基础,对于诸如智能视觉监控系统及其他功能的有效应用发挥着支撑和决定性作用。移动目标检测结果的准确性往往直接影响后续研究的效果。因此,准确的视频移动目标检测技术一直是研究者研究的重要课题之一。通常情况下,移动目标检测并不单独地构成应用,而是作为一个组成部分出现在实际应用中。所以针对移动目标检测结果的具体要求随着应用的改变而

有很大的区别。例如,对交通监控场景中车辆的检测与跟踪就和对视频中用户手势或表情的识别有所不同,前者的关键在于如何在复杂的光照变化下有效提取运动物体,而后的难点在于怎样从躯体大范围的运动背景中将手势或表情的变化进行识别并提取出来。

视频中的移动目标检测是一件非常有挑战性的工作,不同应用场景的需求差异很大。在一个应用中性能很好的算法可能在另一个场景中完全不适用。如在交通视频监控中,由于太阳位置变化和风吹动的云朵导致的阴影变化,能够使道路的亮度不断变化,可能导致将道路检测成移动目标。类似的干扰因素还有很多,主要有移动物体产生的阴影、风吹动的树叶、水面的波纹、摄像机的抖动等。此外,还有其他的干扰因素,如在道路监控中,车辆之间互相遮挡等现象。当然这些困难也并不是总是存在,如室内视频监视就没有大尺度光照变化的问题;也有些应用中不需要考虑这些问题,如在视频编码中阴影、水波等需要一同视为运动物体,并不能被忽略。由此可见,使用单一的目标检测技术应对所有的应用场合是很困难的。学者们在该领域做了大量的研究工作,相关的算法技术不断涌现。目前,常见的视频移动目标检测算法主要分为帧差法、光流法和背景差法三类。

1. 帧差法(帧间差分法)

摄像机采集的视频序列具有连续性的特点,如果场景内没有运动目标,则连续帧的变化很微弱;如果存在运动目标,则连续的帧和帧之间会有明显的变化。帧间差分法(Temporal Difference)就是借鉴了上述思想。由于场景中的目标在运动,目标的影像在不同图像帧中的位置就相应不同。该类算法对时间上连续的两帧或三帧图像进行差分运算,不同帧对应的像素点相减,判断灰度差的绝对值,当绝对值超过一定阈值时,即可判断为运动目标,从而实现目标的检测功能。

1) 两帧差分法(Two Frames Difference Method)

两帧差分法的运算过程如图 1.2 所示。

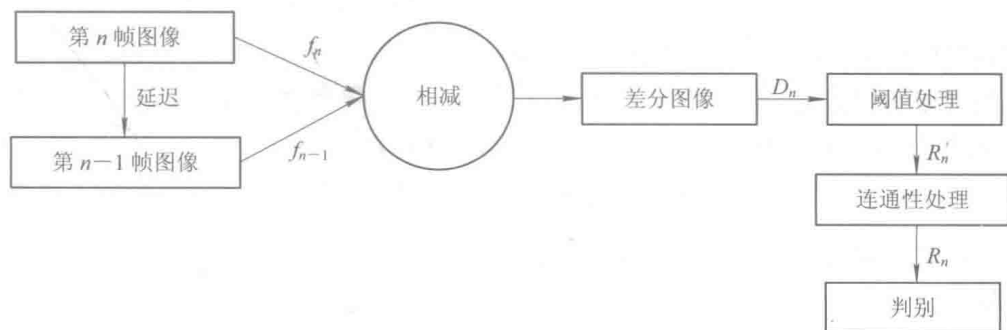


图 1.2 两帧差分法示意图

记视频序列中第 n 帧和第 $n-1$ 帧图像为 f_n 和 f_{n-1} , 两帧对应像素点的灰度值记为 $f_n(x, y)$ 和 $f_{n-1}(x, y)$, 将两帧图像对应像素点的灰度值进行相减, 并取其绝对值, 得到差分图像 D_n :

$$D_n(x, y) = |f_n(x, y) - f_{n-1}(x, y)| \quad (1.1)$$

设定阈值 T , 对像素点逐个进行二值化处理, 得到二值化图像 R'_n :

$$R'_n(x, y) = \begin{cases} 255, & D_n(x, y) > T \\ 0, & \text{其他} \end{cases} \quad (1.2)$$

其中, 灰度值为 255 的点即为前景(运动目标)点, 灰度值为 0 的点即为背景点; 对图像 R'_n 进行连通性分析, 最终可得到含有完整运动目标的图像 R_n 。

2) 三帧差分法(Three Frames Difference Method)

两帧差分法适用于目标运动较为缓慢的场景。当目标运动较快时, 由于目标在相邻帧图像上的位置相差较大, 两帧图像相减后并不能得到完整的运动目标, 因此研究者在两帧差分法的基础上提出了三帧差分法, 其运算过程如图 1.3 所示。

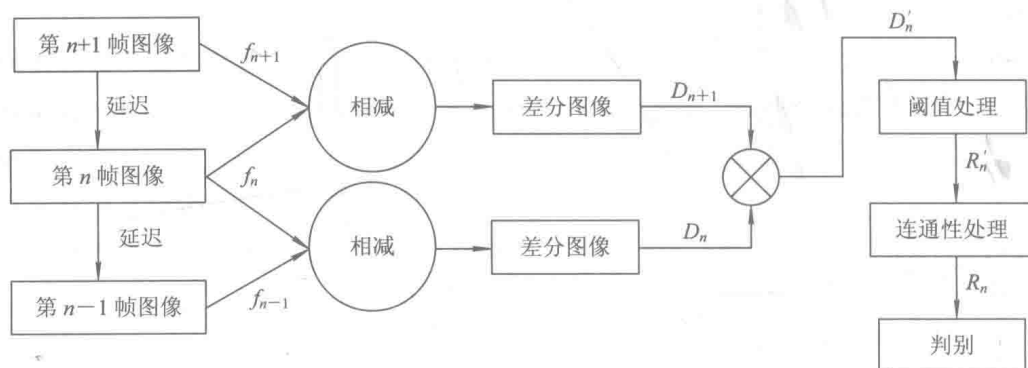


图 1.3 三帧差分法示意图

记视频序列中第 $n+1$ 帧、第 n 帧和第 $n-1$ 帧的图像分别为 f_{n+1} 、 f_n 和 f_{n-1} , 三帧对应像素点的灰度值记为 $f_{n+1}(x, y)$ 、 $f_n(x, y)$ 和 $f_{n-1}(x, y)$, 按照公式(1.1)分别得到差分图像 D_{n+1} 和 D_n 。对差分图像 D_{n+1} 和 D_n 进行与操作, 得到图像 D'_n :

$$D'_n(x, y) = |f_{n+1}(x, y) - f_n(x, y)| \cap |-f_n(x, y) - f_{n-1}(x, y)| \quad (1.3)$$

然后再进行阈值处理、连通性分析, 最终提取出运动目标。

在帧差分法中, 阈值 T 的选择非常重要。如果阈值 T 选取的值太小, 则无法抑制差分图像中的噪声; 如果阈值 T 选取的值太大, 又有可能掩盖差分图像中目标的部分信息, 而且固定的阈值 T 无法适应场景中光线变化等情况。为此, 有人提出了在判决条件中加入对整

体光照敏感的添加项的方法, 将判决条件修改为

$$\text{Max}_{(x,y) \in A} |f_n(x,y) - f_{n-1}(x,y)| > T + \tau \frac{1}{N_A} \sum_{(x,y) \in A} |f_n(x,y) - f_{n-1}(x,y)| \quad (1.4)$$

其中, N_A 为待检测区域中像素的总数目; τ 为光照的抑制系数; A 为一帧完整图像。添加项 $\sum_{(x,y) \in A} |f_n(x,y) - f_{n-1}(x,y)|$ 表示 A 中光照的变化情况。如果场景中的光照变化较小, 则该项的值趋向于零; 如果场景中的光照变化明显, 则该项数值明显增大, 导致式(1.4)右侧判决条件自适应地增大, 最终的判决结果为没有运动目标, 这样就有效地抑制了光线变化对运动目标检测结果的影响。

帧差法在视频移动目标检测中应用广泛, 适用于背景固定不变的情况。帧差法更新速度快、计算量小, 能够快速检测出场景中的运动目标, 但是对环境噪声较为敏感的同时, 此算法对于比较大的、颜色一致的运动目标, 有可能在目标内部产生空洞, 无法完整地提取运动目标。阈值的选择也很关键, 若选择过低不足以抑制图像中的噪声, 若过高则丢失图像中有用的变化。故帧间差分法通常不单独用在目标检测中, 往往与其他的检测算法结合使用。

2. 光流法

光流法是由 Gibson 在 1950 年提出的, 它利用光流方程计算出每个像素点的运动状态矢量, 将具有相同光流向量的区域看做一个运动目标, 进而对这些像素点进行跟踪。通俗地说, 光流就是我们在这个运动的世界里感觉到的明显的视觉运动。例如, 当我们坐在火车上往窗外看时, 可以看到树、地面、建筑, 等等, 都在往后移动。这个运动就是光流。而且, 我们可以通过不同目标的运动速度判断它们与我们的距离。一些比较远的目标, 例如云和山, 它们移动得很慢, 感觉就像静止一样。但一些离得比较近的物体, 例如建筑和树, 就比较快地往后移动, 距离越近, 速度越快。一些离得非常近的物体, 例如路面的标记、草地等, 快到好像在我们耳旁发出“嗖嗖”的声音。光流除了提供远近信息外, 还可以提供角度信息。与人类视线成 90° 方向运动的物体速度要比其他角度的快, 与人类视线成 0° 运动的物体, 感觉上是物体朝着我们的方向直接撞过来, 我们就感受不到它的运动(光流)了, 看起来好像是静止的。当它离我们越近, 就越来越大。

光流法是空间运动物体在观察成像平面上的像素运动的瞬时速度, 是利用图像序列中像素在时间域上的变化以及相邻帧之间的相关性来找到上一帧跟当前帧之间存在的对应关系, 从而计算出相邻帧之间物体的运动信息的一种方法。一般而言, 光流是由于场景中前景目标本身的移动、相机的运动, 或者两者的共同运动所产生的。在一个图片序列

中,把每张图像中每个像素的运动速度和运动方向找出来就是光流场。第 t 帧的时候 A 点的位置是 (x_1, y_1) , 在第 $t+1$ 帧的时候再找到 A 点, 假如它的位置是 (x_2, y_2) , 则 A 点的运动是:

$$(u_x, v_y) = (x_2, y_2) - (x_1, y_1)$$

那么如何计算第 $t+1$ 帧的时候 A 点的位置呢? 这就存在很多的光流计算方法了。1981 年, Horn 和 Schunck 创造性地将二维速度场与灰度相联系, 引入光流约束方程, 得到光流计算的基本算法。人们基于不同的理论基础提出各种光流计算方法, 算法性能各有不同。Barron 等人对多种光流计算技术进行了总结, 按照理论基础与数学方法的区别把它们分成四种: 基于梯度的方法、基于匹配的方法、基于能量的方法和基于相位的方法。近年来, 神经动力学方法也颇受学者重视。

因此, 光流法在适当的平滑性约束条件下, 主要根据图像序列的时空梯度估算运动场, 通过分析运动场的变化对运动目标和场景进行检测与分割。我们也可以把光流法分为基于全局光流场和基于特征点光流场两种方法。全局光流场通过比较运动目标与背景之间的运动差异对运动目标进行光流分割, 缺点是计算量大。特征点光流场通过特征来匹配计算特征点处的流速, 其比全局光流场算法计算量小, 但是却很难精确地提取运动目标的形状。光流法的优点是可以有效地在摄像机运动的情况下检测出运动目标, 并且能够同时完成运动目标检测和跟踪。而且, 光流法不需要预先知道场景的任何信息, 就能够检测到运动。光流法的缺点是计算复杂度高, 很难达到实时检测; 容易受到噪声、光照变化和背景扰动的影响, 抗噪性低; 很难将运动目标的轮廓完整地提取出来。

3. 背景差法(Three Stages of the Background Difference Method)

背景差法首先对视频流中的图像序列进行背景建模, 得到场景中的背景图, 然后利用当前帧与背景图像的差获得前景图(移动目标)。这种基于背景重建的目标检测方法一般分为三个阶段: 背景模型建立、运动检测和背景模型更新, 如图 1.4 所示。



图 1.4 背景差法的三个阶段

背景模型建立是首先对若干视频图像序列进行训练, 然后提取当前视频场景背景的特征, 并将背景的特征用数学模型表达出来。运动检测是将当前帧与背景模型进行比较, 差异较大的点视为运动的点。背景模型更新是根据视频图像变化情况对背景模型进行不断调整, 使建立的背景模型能够适应光照变化、图像噪声等带来的干扰。背景差法的性能依赖

于所使用的背景建模技术。传统的背景建模方法(Traditional Background Modeling Method)有均值背景建模方法、中值背景建模方法、卡尔曼滤波建模方法、单高斯背景建模方法、混合高斯背景建模方法、基于核密度估计的背景建模方法和码本背景建模方法等。传统的背景建模方法及其使用的主要技术手段见表 1.1。

表 1.1 传统的背景建模方法及主要技术手段

传统的背景建模方法	主要技术手段
均值背景建模方法	对一些连续帧取像素平均值
中值背景建模方法	在一段时间内, 将连续 N 帧图像序列中对应位置的像素点灰度值按从小到大排列, 然后取中间值作为背景图像中对应像素点的灰度值
卡尔曼滤波建模方法	该算法把背景认为是一种稳态的系统, 把前景图像认为是一种噪声, 用基于卡尔曼滤波理论的时域递归低通滤波来预测变化缓慢的背景图像
单高斯背景建模方法	将图像中每一个像素点的灰度值看成是一个随机过程 X , 并假设该点的某一像素灰度值出现的概率服从高斯分布
混合高斯背景建模方法	将背景图像的每一个像素点按多个高斯分布的叠加来建模, 每种高斯分布可以表示一种背景场景, 多个高斯模型混合使用来模拟出复杂场景中的多模态情形
基于核密度估计的背景建模方法	基于概率论中的核密度估计理论
码本背景建模方法	根据像素点的连续采样值的颜色失真程度及其亮度范围, 将背景像素用码本表示, 然后利用背景差分法思想对新输入像素值与其对应码本进行比较判断, 从而提取出前景目标像素

一些学者对上述常见的一些背景建模方法进行了对比实验。通过实验发现: 复杂的背景建模算法的移动目标检测效果较好, 但其时间复杂性和空间复杂性较高; 简单的背景建模方法时间复杂性和空间复杂性较低, 但其检测效果却不能令人满意。背景差法的优点是能有效完整地分割出运动目标; 缺点是对变化速度较快的背景, 使用起来效果不够理想。本书第 2 章研究的目标检测方法就属于背景差法范畴。

表 1.2 对常见的视频移动目标检测方法的优缺点进行了总结。此外, 除了上述较为常

用的运动目标检测方法之外,还有一些其他的视频移动目标检测方法,主要有基于人工神经网络的方法、基于主动轮廓模型的方法和基于小波变换的方法。

表 1.2 常见的视频移动目标检测方法优缺点

	优 点	缺 点
帧差法	计算复杂性低,适用于动态背景	容易出现空洞现象 受目标速度快慢影响较大 难以完整地提取目标轮廓
光流法	适用于摄像机运动的情况 能够同时进行目标检测和跟踪	计算复杂性高 抗噪性弱,容易受到光照变化、背景扰动等干扰 很难完整地检测目标
背景差法	抗噪性强 能有效完整地分割出移动目标	不适用于背景变化较快的情况 计算复杂性与背景建模方法有关

1.3.2 视频目标跟踪技术

目标跟踪的概念最早由 Wax 在 1955 年提出。1960 年, Kalman 提出了著名的卡尔曼滤波算法。1964 年, Sittler 提出贝叶斯理论,即利用目标点轨迹和目标运动路径最优数据关联改进目标跟踪算法。视频目标跟踪问题随着计算机视觉技术的发展而逐渐成为研究热点。在 20 世纪 80 年代以前,由于计算机技术与视觉技术发展有限,计算机图像的处理主要以静态图像为主,在动态图像列的分析中,对运动目标的跟踪带有很强的静态图像分析的特点。20 世纪 80 年代初,光流法被提出之后,动态图像序列分析进入了一个研究的高潮。其中对光流法的研究热潮从其产生一直持续到了 20 世纪 90 年代中期。除了光流法之外,还出现了其他众多的视觉跟踪算法。

根据研究角度的不同,目标跟踪问题有不同的分类方法。

(1) 从目标数量的角度分类,可以将目标跟踪分为单目标跟踪和多目标跟踪。单目标跟踪是指对单个移动物体的跟踪。单目标跟踪并不是一个简单的问题,首先从场景中准确地检测和提取运动目标就比较困难,会受到背景变换、光线、噪声的多重影响。多目标跟踪要比单目标跟踪更加困难,在多目标跟踪过程中,必须考虑到多个目标在场景中会互相遮挡、合并、分离等情况,这也是多目标跟踪问题的难点。

(2) 从目标性质的角度分类,可以将目标跟踪分为刚体跟踪和非刚体跟踪。刚体跟踪是指在跟踪过程中目标形状基本不变,或者只有尺度、旋转等简单变化,如交通监控