



运动目标检测 理论与方法

韩光 才溪 汪晋宽 / 著



中国工信出版集团



电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
<http://www.phei.com.cn>



运动目标检测 理论与方法

韩光 才溪 汪晋宽 / 著

電子工業出版社

Publishing House of Electronics Industry

北京•BEIJING

内 容 简 介

本书系统深入地阐述了运动目标检测的理论与方法,总结了国内外运动目标检测技术的最新成果和最新进展,其中涵盖了作者多年来在该领域所取得的科研成果。

全书共 8 章,内容包括运动目标检测的基本概念和研究现状、运动目标检测的经典方法和评价标准,以及基于单像素特征建模、复杂特征建模、多源信息建模、低维子空间分解、盲源信号分离、三维小波变换的运动目标检测方法。

本书可作为视频分析、图像处理、信号处理等领域高年级本科生、研究生的参考学习用书,也可供从事相关领域研究的高校教师、科研人员以及从事相关行业的工程技术人员阅读参考。

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有,侵权必究。

图书在版编目(CIP)数据

运动目标检测理论与方法/韩光,才溪,汪晋宽著. —北京:电子工业出版社,2018.2

ISBN 978-7-121-33309-5

I. ①运… II. ①韩… ②才… ③汪… III. ①运动目标检测—研究 IV. ①TP72

中国版本图书馆 CIP 数据核字(2017)第 311542 号

策划编辑:张楠

责任编辑:张楠

文字编辑:钱维扬

印刷:三河市华成印务有限公司

装订:三河市华成印务有限公司

出版发行:电子工业出版社

北京市海淀区万寿路 173 信箱 邮编 100036

开本:720×1000 1/16 印张:12 字数:270 千字

版次:2018 年 2 月第 1 版

印次:2018 年 2 月第 1 次印刷

定价:48.00 元

凡所购买电子工业出版社图书有缺损问题,请向购买书店调换。若书店售缺,请与本社发行部联系,联系及邮购电话:(010) 88254888, 88258888。

质量投诉请发邮件至 zlts@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

本书咨询联系方式:(010) 88254579。

Preface 前言

运动目标检测是综合了人工智能、模式识别、图像与信号处理、统计估计等理论技术的交叉型研究领域。作为智能视频分析的核心技术，运动目标检测技术是智能视频分析其他各种后续处理（如目标跟踪、目标识别、行为分析理解等）的前提和基础，它可将人们感兴趣的运动目标从视频数据中智能检测并自动提取分割出来。目前，运动目标检测技术已广泛应用于智能视频监控、智能交通、灾情监控、人机交互、医学运动分析、汽车自动驾驶等领域。

近年来，国内外学者对运动目标检测的研究已提高了视频分析的智能化水平。然而，随着视频分析应用领域的不断扩大以及人们对海量数据分析智能性、准确性要求的不断提高，运动目标检测在面对越来越多复杂的实际应用场景时仍有许多理论和技术上的关键问题有待进一步解决。

作者在总结近年来在运动目标检测领域取得的研究成果的基础上，进一步系统深入地研究了运动目标检测的最新成果和最新方法，分析了运动目标检测技术现存的关键问题，并开展了大量的研究工作力图解决上述问题。作为研究工作的阶段总结，作者将近年来的相关研究成果汇总成册，构成了本书的主要内容，期望为从事视频分析、图像处理、信号处理研究工作的同仁在理论分析方法上提供一些有益的帮助。

全书共分8章，第1章简要介绍运动目标检测的基本概念和研究现状，第2章概述运动目标检测的经典方法和评价标准，第3章～第8章分别对基于单像素特征建模、复杂特征建模、多源信息建模、低维子空间分解、盲源信号分离、三维小波变换的运动目标检测方法展开讨论。

本书由韩光、才溪执笔，汪晋宽统稿。本书在写作过程中参阅和借鉴了大

量国内外文献资料，同时受到了国家自然科学基金（61601108、61701098）资助，在此一并表示衷心的感谢！

由于运动目标检测技术正在飞速发展，加之作者水平有限，书中难免有疏漏和不足，敬请读者批评、指正。

著者
2017年12月

Contents

目录

第 1 章 绪论	1
1.1 运动目标检测研究背景及意义	1
1.2 运动目标检测技术发展及研究现状	6
1.3 运动目标检测的应用难题和研究热点	11
1.3.1 运动目标检测的应用难题	11
1.3.2 运动目标检测的研究热点	12
1.4 本章小结	13
参考文献	13
第 2 章 运动目标检测概述	18
2.1 运动目标检测的经典方法	18
2.1.1 帧间差分法	18
2.1.2 光流法	20
2.1.3 背景减除法	21
2.2 运动目标检测的性能评价	24
2.2.1 主观评价	25
2.2.2 客观评价	25
2.3 运动目标检测的公开数据库	28
2.4 本章小结	42
参考文献	42
第 3 章 基于单像素特征建模的运动目标检测	45
3.1 相关工作与研究现状	45
3.2 基于人类视觉亮度敏感性的运动目标检测算法	46

3.2.1	揭示人类感知规律的韦伯定律	47
3.2.2	适用于复杂图像背景环境的韦伯比的分析与推导	49
3.2.3	基于人类视觉亮度敏感性的自适应匹配判断阈值设置	52
3.2.4	算法描述	53
3.2.5	自适应匹配判断阈值对算法性能的影响分析	55
3.3	实验结果及分析	56
3.3.1	实验环境设置	56
3.3.2	定性分析	59
3.3.3	定量分析	63
3.4	本章小结	65
	参考文献	65
第4章	基于复杂特征建模的运动目标检测	68
4.1	相关工作与研究现状	68
4.2	基于改进脉冲耦合神经网络的运动目标检测算法	69
4.2.1	脉冲耦合神经网络(PCNN)的基础理论	70
4.2.2	基于改进脉冲耦合神经网络的全局特征提取	75
4.2.3	算法描述	79
4.3	实验结果及分析	80
4.3.1	实验环境设置	80
4.3.2	定性分析	81
4.3.3	定量分析	88
4.4	本章小结	90
	参考文献	90
第5章	基于多源信息建模的运动目标检测	93
5.1	相关工作与研究现状	93
5.2	基于红外、可见光多源特征融合建模的运动目标检测算法	94
5.2.1	红外、可见光多源特征融合的理论基础和优势分析	95
5.2.2	算法描述	98
5.3	实验结果及分析	101

5.3.1	实验环境设置	101
5.3.2	定性分析	102
5.3.3	定量分析	106
5.4	本章小结	107
	参考文献	108
第 6 章	基于低维子空间分解的运动目标检测	109
6.1	相关工作与研究现状	109
6.1.1	低维子空间分解理论基础	110
6.1.2	低维子空间分解在运动目标检测中的应用	111
6.2	基于改进在线鲁棒主成分分析的运动目标检测算法	112
6.2.1	自适应稀疏权重的在线鲁棒主成分分析	112
6.2.2	算法描述	116
6.3	实验结果及分析	120
6.3.1	实验环境设置	120
6.3.2	定性分析	122
6.3.3	定量分析	126
6.4	本章小结	128
	参考文献	128
第 7 章	基于盲源信号分离的运动目标检测	130
7.1	相关工作与研究现状	130
7.1.1	盲源信号分离理论基础	131
7.1.2	盲源信号分离在运动目标检测中的应用	132
7.2	基于约束非线性独立成分分析的运动目标检测算法	133
7.2.1	运动目标检测中的非线性盲源分离问题	133
7.2.2	算法描述	135
7.3	实验结果及分析	137
7.3.1	实验环境设置	137
7.3.2	定性分析	138
7.3.3	定量分析	141

7.4 本章小结	142
参考文献	142
第8章 基于三维小波变换的运动目标检测	144
8.1 小波变换基础理论与快速算法	145
8.1.1 一维小波变换	146
8.1.2 三维小波变换	150
8.2 基于三维小波变换的运动目标检测算法	155
8.2.1 三维小波变换用于运动目标检测的理论分析	155
8.2.2 算法描述	157
8.2.3 算法优势分析	159
8.3 实验结果及分析	162
8.3.1 实验环境设置	162
8.3.2 定性分析	166
8.3.3 定量分析	177
8.4 本章小结	182
参考文献	182

第 1 章

绪 论

1.1 运动目标检测研究背景及意义

随着成像设备的日渐普及、信息技术和互联网技术的飞速发展,视频数据的采集、存储、传输方式越来越便捷,由此产生了数据量呈爆炸式增长的海量视频数据。视频分析正步入大数据时代。面对内容丰富、形式多样的海量视频数据,如何从中智能地获取有用信息非常关键,智能视频分析已成为当前的研究热点。

由于海量视频数据中占比例较小的目标区域通常包含有绝大部分信息量,因此,对视频数据中目标信息的自动解译可有效提高海量视频数据的分析利用率。作为提取目标信息的重要手段,运动目标检测是智能视频分析的基础和核心技术之一。

运动目标检测技术就是将人们感兴趣的运动目标从视频数据中智能检测并自动提取分割出来。如图 1.1 和图 1.2 分别给出了运动目标检测的两个实例。图 1.1 的监控场景为一条高速公路,在这样的场景中,人们感兴趣的目标是在路上行驶的车辆;图 1.2 的监控场景是一个地铁站,人们感兴趣的目标主要是出现在该场景中的人。图 1.1 (a) 和图 1.2 (a) 中分别给出了在各自监控场景下拍摄得到视频数据的多帧视频图像,图 1.1 (b) 和图 1.2 (b) 的相应位置则分别给出了对应帧的运动目标检测结果。从这些检测结果中可以看出,对视频数据进行运动目标检测处理能够清晰捕获场景中的运动目标,可获得关于目标

形状、位置及其运动趋势等重要信息。这些检测结果不仅便于人眼观察，而且更为重要的是，它们将为后续对运动目标的进一步智能分析理解提供有关目标的重要数据和信息。



Frame 1508



Frame 1512



Frame 1516



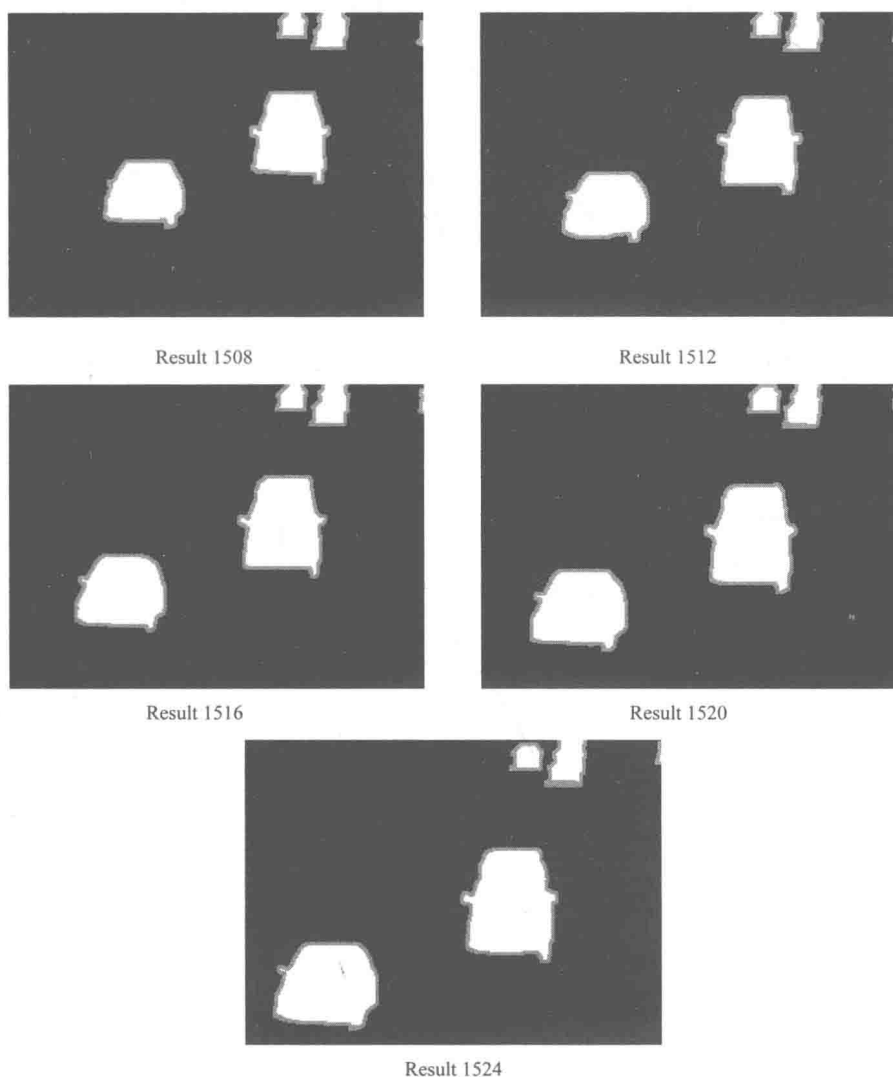
Frame 1520



Frame 1524

(a) 监控视频图像

图 1.1 运动目标检测实例一（高速公路）



Result 1524
(b) 运动目标检测结果

图 1.1 运动目标检测实例一（高速公路）（续）

通过上述实例可以看出，运动目标检测技术是智能视频分析中其他各种后续处理（如目标跟踪、目标识别、行为分析理解等）技术的基础。对视频中出现的运动目标的正确检测及准确分割很大程度上影响着能否对这些运动目标

进行正确的跟踪、识别、分析和理解，同时，运动区域的有效分割还能大大减少后续进一步分析处理的运算量。因此，运动目标检测技术一直都是智能视频分析中的一项重要研究课题。



Frame 400



Frame 410



Frame 420



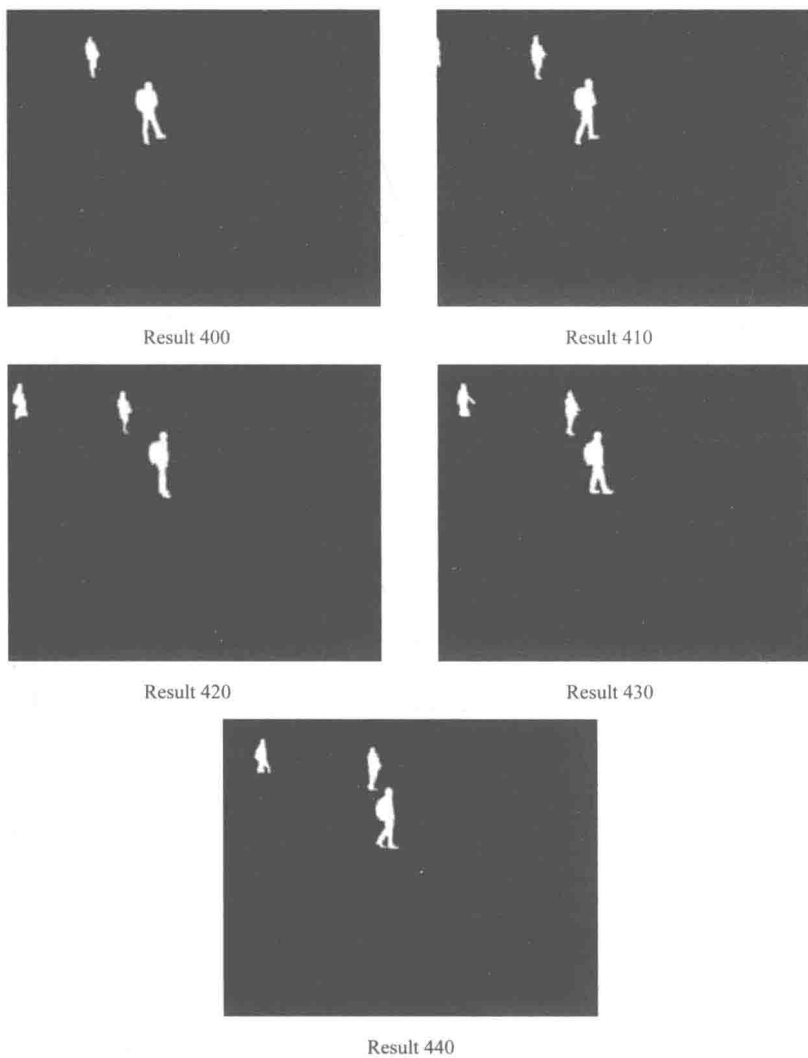
Frame 430



Frame 440

(a) 监控视频图像

图 1.2 运动目标检测实例二（地铁站）



(b) 运动目标检测结果

图 1.2 运动目标检测实例二（地铁站）（续）

运动目标检测技术发展的最初牵引力来自于其在军事监控和安全监控领域的应用，用来提高军事安全监控系统的智能性和可靠性，从而增强战场态势感知、部队安全防护以及国防、民用等场合免受恐怖袭击的保护能力。随着运

运动目标检测技术研究的不断深入和视频采集设备的日渐普及,运动目标检测技术的应用前景也日益广阔,目前已经广泛应用于智能视频监控、智能交通、灾情监测、人机交互、医学运动分析、汽车自动驾驶等领域。

运动目标检测技术应用领域的不断扩大使其在面临越来越多复杂的实际应用场景时将面临更多挑战,同时人们对于安全监控、智能交通等领域中视频分析系统的智能化要求的不断提高也给运动目标检测技术提出了越来越高的性能要求,因此,对运动目标检测技术的深入研究将不仅有利于加强智能视频分析系统在军事安全、社会安全等重要领域的应用能力,促进国家安全及公共安全体系建设,还能带动智能化分析系统在影响人们日常生活的众多领域中的大范围推广和普及,有利于推进我国信息惠民工程建设和智慧城市、智慧社区建设,对促进我国国民经济和社会的和谐发展具有重大意义和深远影响。

1.2 运动目标检测技术发展及研究现状

在智能视频分析系统中,运动目标检测的目的主要是将人们感兴趣的运动目标从视频数据中智能地检测并分割出来。目前,运动目标检测方法主要分为3大类:帧间差分法、光流法和背景减除法。本节分别对3大类运动目标检测方法的研究现状进行分析与总结。

(1) 帧间差分法。

帧间差分法是将视频中时域相邻的图像的特征进行差分运算,而后对所得的差分图像进行二值化阈值处理,以提取运动目标所在位置。帧间差分法不需要监控场景的任何先验知识,也不需要训练过程,且其计算量小,存储数据少,可以迅速输出检测结果,即使应用在嵌入式系统中也可实现实时检测。

帧间差分法目前主要应用于对检测实时性有很高要求的视频分析应用中(如智能交通监控等)。1995年,美国密歇根州立大学 Dubuisson 和 Jain 利用对3个相邻视频帧做两次帧间差分的结果取交集的方式提取车辆的完整轮廓信息^[1]。1998年,美国卡耐基梅隆大学 Lipton 等人利用连续两个视频帧灰度差的绝对值来寻找变化区域^[2]。此外,由于帧间差分法计算量小,也常被作为复杂运动目标检测算法的预处理步骤,用来提取运动目标的初始大致区域。最近,研究人员将帧间差分法与其他运动目标检测方法结合使用,用来提高运动目标检测

的性能。如2015年, Han等人提出将帧间差分法与光流法结合^[3], 以降低光流法的计算复杂度; 2016年, Liu等人提出将帧间差分法与背景减除法结合^[4], 实现水下视频的运动目标检测。

帧间差分法对光照变化、背景扰动、环境噪声等异常敏感, 因此, 将它应用于复杂监控场景时无法准确地获得运动目标的分割区域。特别是当运动目标运动速度较慢时, 帧间差分法易在运动目标检测结果的内部形成空洞, 甚至当运动目标短暂静止时, 无法再检测出整个运动目标。可见, 帧间差分法虽简单易行, 但当智能视频分析系统需要获取并准确分析分割的运动目标(尤其是需对分割出来的运动目标进行后续的识别、理解等高级分析)时则很难满足其需求。

(2) 光流法。

光流法主要通过分析视频的光流场进行运动目标检测。光流是指空间中物体的运动在图像平面上表现出的物体灰度模式的流动。光流法基于场景中物体的灰度守恒假设, 通过分析视频图像平面上物体灰度模式的流动, 发现场景中各物体的运动, 并对运动特性不同的物体进行分割。光流法能够同时捕捉因前景目标运动引起的前景目标的光流和因摄像机运动导致的背景物体的光流, 并能够对它们进行有效区分, 因此, 适用于摄像机存在运动情况下的运动目标检测(如车载摄像机监控等)。

1981年, 美国卡耐基梅隆大学的 Lucas 和 Kanade 共同提出了一个局部光流法的经典算法——Lucas—Kanade 算法^[5]。同年, 美国麻省理工学院人工智能实验室研究人员提出了一个全局光流法的经典算法^[6]。1994年, 加拿大西安大略大学 Barron 等人将多种光流计算方法分类为基于梯度的方法、基于匹配的方法、基于能量的方法和基于相位的方法, 并对它们进行了实验对比分析^[7]。1997年, 德国埃尔朗根纽伦堡大学的研究人员使用光流法进行运动目标的初始分割, 以对基于轮廓的跟踪算法进行初始化^[8]。2000年, 日本德岛大学 Chunle 和 Oe 提出了利用多梯度限制的基于梯度的光流法^[9]。2007年, 美国微软研究院 Baker 等人创建了新的数据库并利用该数据库对经典的多种光流法进行了对比分析^[10]。2012年, 韩国浦项科技大学 Lim 等人提出利用光流和极线约束来区分运动的前景目标和运动的背景, 以实现移动摄像机拍摄视频的运动前景目标的检测^[11]。2015年, 比利时根特大学 Allebosch 等人提出利用光流法检测摄像机运动(如平移、俯仰、抖动), 并对摄像机的运动进行补偿^[12]。

光流法计算量很大,且其基础的灰度守恒假设条件在实际应用中受噪声、光照变化等多因素影响,很难得到满足,因此,光流法一般不适用于对运动目标检测精度和实时性要求较高的监控系统。

(3) 背景减法。

背景减除 (Background Subtraction) 法是目前最流行的一大类运动目标检测方法,其基本思想是从视频中减除背景、保留运动前景目标。文献[13-15]从不同角度对背景减除法进行了较为全面的综述。

早期简单的背景减除法主要是从一些训练数据中分析并估计场景的背景图像,如由各像素在训练数据中的灰度均值或中值形成背景图像,然后,对当前输入帧与背景图像进行差分运算,对于差异较小的区域即判断其为背景,反之则判断其为前景目标。文献[16]详细讨论了使用不同方式估计背景图像、度量背景图像与当前输入图像差异给最终检测结果带来的影响。

随着人们对检测精度的要求不断提高,研究人员逐渐使用背景模型来代替单一的背景图像,以准确描述复杂多变的场景背景。

最初用于描述背景的模型是概率分布模型。1996年,麻省理工学院 Wren 等人提出的 Pfinder 系统^[17]利用了单个高斯分布对场景背景建模,实现了室内场景中的运动目标检测。由于单高斯分布背景模型无法适应场景的动态变化,1999年麻省理工学院 Stauffer 等人提出了使用多个高斯分布对场景背景建模的运动目标检测算法,也被称为高斯混合模型 (Gaussian Mixture Model, GMM) 算法^[18],大量实验证明 GMM 可以有效应对存在缓慢光照变化和重复性运动的动态背景。不过, GMM 算法中描述背景模型的参数很多为经验设定,缺少适应实际情况的灵活性,因此,在此基础上出现了一大批行之有效的 GMM 改进算法。2005年, Lee 等人提出对每个高斯分布使用自适应的学习速率,提高背景模型学习的收敛速率^[19]。2006年, Zivkovic 等人提出对各像素点背景模型中的高斯分布个数进行在线的自适应调整,有效提高了算法的处理速度^[20]。2007年,上海交通大学研究人员在 GMM 基础上加入了空间信息,提出了空时高斯混合模型,用来更好地应对动态背景^[21]。2014年,英国金斯顿大学 Chen 等人提出根据场景全局光照自适应调整 GMM 模型的学习速率,有效增强了算法对快速光照变化的适应能力^[22]。2015年,日本东京大学研究人员提出使用近似 GMM 模型建立背景模型,提高了运动目标检测的准确性^[23]。

除使用高斯分布外,研究人员还尝试着用其他分布来对背景建立模型。