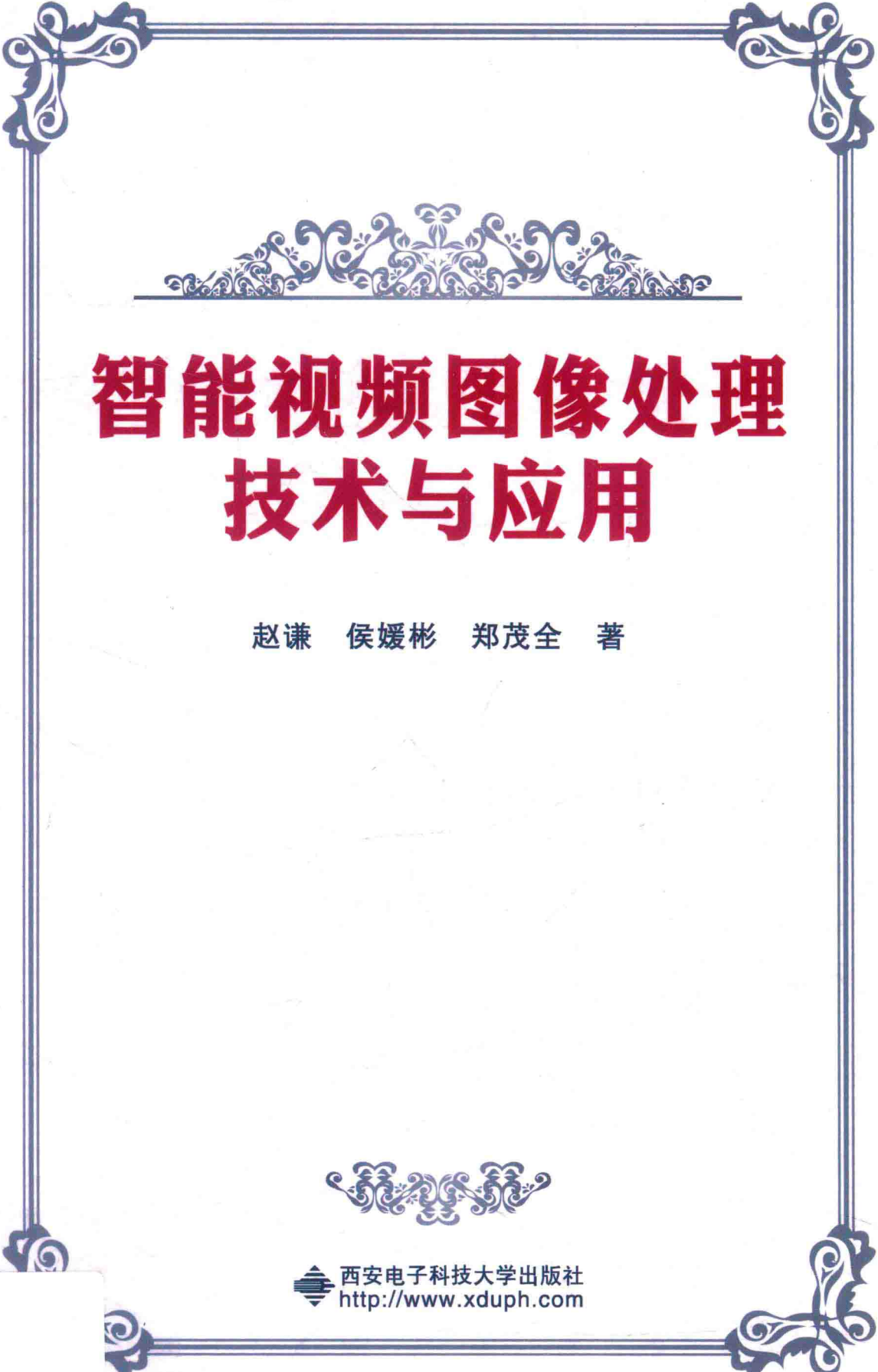




智能视频图像处理 技术与应用

赵谦 侯媛彬 郑茂全 著



西安电子科技大学出版社
<http://www.xduph.com>

智能视频图像处理技术与应用

赵 谦 侯媛彬 郑茂全 著

西安电子科技大学出版社

内 容 简 介

本书在全面综述国内外视频图像的检测、增强、跟踪等技术的基础上,重点介绍了作者在这一领域的研究成果。主要内容包括:分析了矿井图像受噪声影响画面不清等问题,改进了基于模糊熵判别准则合理提取 LFFD 的相似度增强算法;研究了视频监控系统采集点多、历史留存数据量大,不利于后续查找兴趣特征图像等问题,提出了一种基于相关法的欧氏距离配准算法;研究了现实环境视频照度不均、噪声大、极易丢失目标以及工矿企业安全生产对排查前景目标精度要求高等问题,分析了当目标运动信息不足时 CBM 会出现误检或局部漏测等问题,通过联合目标的空间整体性提出了一种基于 CBM 的目标空间整体性背景更新算法;针对公共安全、交通安全行驶中遗留物可能带来的安全隐患等问题,提出了一种基于历史像素稳定度的遗留物检测算法;针对动态目标复杂运动、光照变化及遮挡等因素对目标跟踪性能的影响,分析了现有基于多特征融合的跟踪算法在复杂环境下跟踪准确度不高,且大部分采用单一判定方式来实现多特征融合的情况,提出了一种基于多特征判定准则的目标跟踪融合算法;介绍了基于三频彩色条纹投影轮廓术的微变监测技术。

本书可作为计算机、通信与信息、自动化与控制等专业的高年级本科生与硕士研究生相关课程的辅导教材,也可作为专业技术人员的培训参考书。

图书在版编目(CIP)数据

智能视频图像处理技术与应用/赵谦,侯媛彬,郑茂全著. —西安:西安电子科技大学出版社,2016.11
ISBN 978-7-5606-4304-5

I. ① 智… II. ① 赵… ② 侯… ③ 郑… III. ① 视频信号—数字图像处理—研究 IV. ① TN941.1

中国版本图书馆 CIP 数据核字(2016)第 230109 号

策 划 毛红兵

责任编辑 傅艳霞 马武装

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 虎彩印艺股份有限公司

版 次 2016年9月第1版 2016年11月第1次印刷

开 本 787毫米×1092毫米 1/16 印张 8.5

字 数 195千字

定 价 26.00元

ISBN 978-7-5606-4304-5/TN

XDUP 4596001-1

* * * 如有印装问题可调换 * * *

前言

现实视频作业环境存在动态目标多、噪声大、光照不均、遮挡以及仅依靠人工职守造成智能化低、误检率高等问题，因此有必要研究图像增强、提取特征点快速查询配准、实时动态目标检测、遗留物检测以及目标跟踪等技术。通过理解分析图像画面出现的违规行为、可疑目标和潜在危险，以快速合理的方式发出联动报警，同时为事故后期分析提供第一手资料。本书在对视频图像处理相关技术研究的基础上，重点研究动态目标视频图像处理的相关检测算法。

针对图像受噪声影响导致画面不清等问题，本书在分析研究现有图像增强相关技术的基础上，研究分析了基于模糊熵判别准则合理提取局部模糊分形维数(LFFD)的相似度增强算法。该算法通过模糊熵判别合理 LFFD 融合相似性测度来调整图像对比度，并考虑增强过程中的多参数性在相似度测量理论上的应用。实验结果表明，该算法能较好抑制噪声，提高图像对比度。针对如何获得恰当的图像特征及细节纹理，本书提出一种基于小波分解的 Canny 边缘检测算法。该算法引入小波变换提取灰度图像的高低频分量，以此来获得更多的边缘信息完善特征轮廓，并对 3D 特征点云的精确收集起到关键作用。

目前，视频监控系统采集点多，历史留存数据量大，不利于后续查找特征图像。本书针对这些问题提出一种基于相关法的欧氏距离配准算法。该算法通过利用不同特征点自身信息，在 Harris 算法基础上分别对灰度信息使用梯度相关法，对 SIFT 算法描述子信息使用描述子相关法，并结合特征点间的欧氏距离关系来精确匹配。实验结果表明，该算法降低了误匹配的点数，精确了 3D 点云的收集。

本书研究了基于码书模型(CBM)的运动目标检测算法。针对当目标的运动信息不足时，CBM 可能会出现误检或局部漏测等问题，通过联合目标的空间整体性，提出一种基于 CBM 的目标空间整体性背景更新算法。该算法通过对运动目标空间信息变化的分析，寻找前景中潜在的背景，并联合像素时域统计进行背景模型更新。实验结果表明，该算法可以快速适应背景变化，在处理缓慢移动目标和只有局部运动目标时，能减少由于运动信息不足所造成的误判，同时保证了目标检测的完整性。针对目标检测时受阴影干扰等问题，本书提出一种基于 HSV 空间的码字分量平均算法。该算法通过构建码字加权平均背景模型，并将 RGB 空间转换到 HSV 空间达到更新背景、去除阴影的效果。实验结果表明，该算法对去除阴影有较强的鲁棒性。

在遗留物检测方面，本书研究发现以多层达到背景模型为目的的算法是通过控制不同模型的更新速度，并比较模型之间的差异来判断遗留物的，但此类算法检测速度较慢，对去除“鬼影”效果不理想。因此，本书提出一种基于历史像素稳定度的遗留物检测算法，该算法在运动目标检测的基础上，对不属于背景码书模型的像素点记录其之前若干帧像素信息，构成历史像素集，并通过统计当前像素与历史像素集的匹配程度来判决该像素

点是否稳定,进而判断是否存在遗留物。实验结果表明,该算法检测遗留物准确,而且去除“鬼影”效果较好,实时性强。

动态目标的运动复杂、光照变化以及遮挡等因素对目标跟踪性能影响大,而现有基于多特征融合的跟踪算法在复杂环境下的跟踪准确度不高,且大部分采用单一判定方式来实现多特征融合的问题。因此,本书提出一种基于多特征判定准则的目标跟踪融合算法,该算法首先引入局部背景信息加强对目标的描述,其次在多特征融合过程中利用多种判定准则自适应计算特征权值,最后在 Mean Shift 框架下,结合 Kalman 滤波完成对目标的跟踪。实验结果表明,本文所提算法跟踪准确性高。

在结合视频图像处理应用方面,本书分析研究了基于三频彩色条纹投影轮廓术的微变监测技术。该技术首先利用二维经验模式分解(BEMD)的自适应条纹分析技术解决三频彩色条纹的颜色解耦难题;其次,利用傅立叶变换实现变精度全场包裹相位展开,通过三频变精度得到高精度绝对相位;最后,标定系统,恢复物体高程信息。通过模拟相似实验验证,本书所研究技术具有采集数据精确、处理复杂度低、设备易于安装等优势。

附录中给出了视频图像处理的相关程序,可为读者提供参考。

本书共分为7章,其中第5章及附录由侯媛彬教授(博士生导师)撰写,第3章由郑茂全撰写,其余章节由赵谦撰写。侯媛彬教授负责本书的结构安排。

本书在编写过程中得到了诸多专家的帮助与指导,在此一并表示感谢。

本书也得到朱华伟、赵诚、周勇、王奕婷、寇思玮、任志奇、赵鹏飞等学生的大力支持,作为你们的导师,我很荣幸,感谢你们。

书中难免存在不足之处,欢迎广大读者提出宝贵意见。

编 者
2016.6

目 录

第 1 章 绪论	1
1.1 图像增强现状	2
1.2 特征点提取及配准现状	3
1.3 动态目标检测现状	4
1.4 遗留物检测现状	6
1.5 动态目标跟踪现状	6
1.6 本书研究内容与技术路线	8
1.6.1 研究内容	8
1.6.2 技术路线	9
第 2 章 视频图像的增强	11
2.1 图像增强	11
2.1.1 灰度直方图统计方法	11
2.1.2 图像空域增强	12
2.1.3 二维频域增强的依据	14
2.1.4 图像锐化的处理过程	14
2.2 图像增强的特征值提取	15
2.2.1 梯度测度	15
2.2.2 标准差测量	15
2.2.3 峰度分布测量	15
2.2.4 熵值提取	16
2.2.5 分形维数提取方法分析	16
2.2.6 模糊熵与分形维的关系	17
2.2.7 相似度测量分析	19
2.3 基于模糊熵判别准则合理提取 LFFD 的相似度增强算法	21
2.4 图像边缘检测	22
2.4.1 Canny 边缘检测分析	23
2.4.2 基于小波分解的 Canny 边缘检测算法的提出	24
2.5 煤矿井下图像增强实验与分析	26
2.5.1 基于模糊熵判别准则合理提取 LFFD 的相似度增强实验分析	26
2.5.2 基于小波分解的 Canny 煤矿支护边缘检测实验分析	29
2.6 小结	30
第 3 章 视频图像的特征点提取及配准	31
3.1 特征点提取	31
3.1.1 Harris 特征点提取	32
3.1.2 SIFT 特征点提取	32
3.2 图像特征点配准	34
3.2.1 灰度相关法匹配算法	34

3.2.2	特征描述子的匹配方法	35
3.2.3	基于相关法的欧氏距离配准算法的提出	36
3.3	煤矿井下实验与分析	38
3.3.1	特征点提取实验分析	38
3.3.2	特征点匹配实验分析	38
3.4	小结	40
第4章	动态目标的检测	41
4.1	动态目标检测算法分析	41
4.1.1	光流法	41
4.1.2	帧间差分法	42
4.1.3	背景差分法	43
4.1.4	几种背景模型的建议	44
4.2	基于CBM的目标空间整体性背景更新算法	48
4.2.1	CBM对微动目标的失真分析	48
4.2.2	基于CBM的目标空间整体性背景更新算法的提出	50
4.2.3	基于CBM的目标空间整体性背景更新算法的步骤	50
4.3	动态检测中的阴影去除	51
4.3.1	颜色空间转换	51
4.3.2	HSV空间阴影去除	53
4.3.3	基于HSV空间的码字分量平均算法	54
4.4	动态目标检测实验与分析	54
4.4.1	基于CBM的目标空间整体性背景更新实验分析	54
4.4.2	基于HSV空间的码字分量平均算法实验分析	60
4.5	小结	62
第5章	遗留物的检测	64
5.1	动态目标检测算法在遗留物检测中的不足	64
5.2	基于双背景模型的遗留物检测算法分析	65
5.3	基于历史像素稳定度的遗留物检测算法	66
5.3.1	算法提出的依据	66
5.3.2	算法步骤	67
5.4	遗留物检测实验与分析	69
5.5	小结	70
第6章	动态目标的跟踪	72
6.1	基于Kalman滤波器的跟踪算法分析	72
6.1.1	Kalman滤波器原理	72
6.1.2	Kalman滤波器跟踪算法	73
6.2	基于Mean Shift的目标跟踪算法分析	75
6.3	基于多特征判定准则的目标跟踪融合算法	77
6.3.1	多特征目标跟踪算法分析	77
6.3.2	基于多特征判定准则的目标跟踪融合算法依据与步骤	78
6.4	离散场景同一目标跟踪融合算法	82
6.4.1	CBWH算法分析	82
6.4.2	CBWH改进算法提出	83

6.4.3 离散同一目标跟踪融合算法的步骤与条件 84

6.5 动态目标跟踪实验与分析 86

6.5.1 Kalman 滤波器跟踪算法实验分析 86

6.5.2 Mean Shift 目标跟踪算法实验分析 89

6.5.3 基于多特征判定准则的目标跟踪融合算法实验分析 90

6.5.4 离散场景同一目标跟踪融合算法实验分析 92

6.6 小结 93

第 7 章 基于三频彩色条纹投影轮廓术的微变监测 95

7.1 三维测量相关技术简介 95

7.2 三频彩色条纹投影轮廓术的技术原理 98

7.2.1 经验模式分解 100

7.2.2 背景消减和颜色解耦 102

7.2.3 三频变精度相位展开 103

7.2.4 纹理恢复原理 103

7.2.5 亚像素级匹配 104

7.3 模拟微变监测实验与分析 104

7.4 小结 107

附录 108

附录 A 视频图像采集 108

附录 B 视频图像压缩 JPEG 与 JPEG2000 111

附录 C 移动目标检测中的帧间差分法 116

附录 D 二维运动估计中的三步搜索法 116

参考文献 120

第1章 绪 论

视频监测系统的发展经历了三个阶段。第一阶段为模拟视频监测系统。其缺点是监控数量少,监控区域有限,监控视频只能在监控中心调取查看,不方便扩展。第二阶段为模拟数字混合视频监测系统。其性能显著提高,管理控制功能丰富,远距离传输可靠性高,可实现多媒体信息现场查询,但缺点是智能化因素较低。第三阶段为全数字网络视频监测系统。其具有图像质量稳定可靠、监控方式灵活以及可进行智能化扩展等优点。目前,大部分视频监控系统仍主要依靠人工查看、调阅取证图像信息,然而视频监控范围却不断在扩大,所以依赖人工值守的监控方式必然带来巨大的工作量,具体表现如下:

(1) 监控人员容易产生视觉疲劳。视频采集点多、数据量大,这要求监控人员的身体素质、注意力以及发现异常情况的直观判断能力必须十分出色。

(2) 缺乏有效的智能化分析能力,误报率和漏报率较高。监控人员从海量视频数据中分析提取出有价值信息的局限性较大。

(3) 联动报警机制难以开展。依靠人工做出响应和处理,对突发事件、有征兆事件等难以做到准确判断,更无法保证视频监测系统后续联动报警机制的展开。

在此环境下国内很多高新技术公司相继推出了智能化的数字视频监控系统,从视频监控、信号传输、中心控制、远程监管等各方面提出了全方位的解决办法,可以实现从监控平台到地、市监控指挥中心与更高级别监控指挥中心的联网,提高全社会安全管理水平。

目前,国内监控视频系统功能主要体现在:视频传输到监控中心后使用接收光端机将光信号还原为视频信号,经过视频分配器接入硬盘录像和矩阵窗进行视频监看,也可以输入到画面分割器,使用大屏幕等离子电视监看。中心控制室的视频服务器安装一台8~36路的数字监控主机,对现场状况进行实时视频监控和录像,也可以使用TC-8600视频矩阵切换到电视墙上放大查看。

其他功能还包括:

(1) 超级切换:利用数字视频录像机(Digital Video Recorder, DVR)控制图像在屏幕墙上的任意切换,实现矩阵控制平台与DVR平台的无缝连接,将矩阵系统并入数字网络,提升系统整体性能;

(2) 网络控制:可同时浏览16台DVR上的不同视频,灵活控制任何一台视频服务器上的云台设备;

(3) 完美录像:可定时连续录像、动态报警录像、传感器报警录像,具备先进的压缩技术,可调压缩比、帧率,提高图像质量;

(4) 报警联动:具备可报警输入输出联动摄像机电源、灯光,具有报警前5~999秒预录功能;

(5) 抓拍:单帧及多帧画面可同时抓拍,并保存为JPG格式;

(6) 检索方便:可按照文件、日期、时间、监控点、存储器等进行检索查询;

(7) 放大功能:可动态放大正在回放的视频。

综上所述,目前国内视频监控系统正在经历从基础组网阶段到智能化的转变过程,完全实现了高清晰视频图像采集、录像等功能,但距离智能化、无人值守化阶段还有差距,技术还需提高。因此,研究全数字智能视频监测系统将是视频监控发展的方向,其可以改变以往人工值守的旧模式,依靠智能算法自适应分析视频内容,提高处理的实时性、可靠性,控制事态发展,避免和减少损失。

1.1 图像增强现状

现实场景的视频图像会经常存在混合各种噪声干扰、细节纹理不清、对比度差等问题。因此,研究图像增强技术意义较大^[1]。

图像增强是指原始图像经过增强处理后,在特征值上比原图像更清晰,易于人们视觉观察,也可以理解为对图像处理后感兴趣信息的细节纹理被增强,不感兴趣的成分或干扰噪声被削弱。传统的图像增强技术考虑图像的某一特性,如图像的灰度等级、纹理及边缘轮廓等。Cheng^[2-3]等提出了一种基于测量二维图像相似度的图像增强算法,该方法将灰度、纹理等作为测量二维图像相似度的特征,并取得了一定的增强效果。但该算法对噪声大的图像处理效果并不理想。

美国数学家在1975年提出了“分形”这个概念,却没有对分形进行比较严密的定义,以至于分形理论在很长一段时间里应用相对单一。1986年,B. B. Mandelbrot给出了分形的完整定义,指出分形是“部分和整体之间具有某种相似关系的体系”。此定义为分析自然事物的尺度不变性奠定了基础。后来,关于分形的研究也逐渐从几何形状拓展到时间、空间等领域。传统几何实体的维数一般都是整数的:零维的点、一维的线、二维的面、三维的立体空间以及四维的时间和空间。但在近几十年里,分形不一定是整数的几何维数引起很多研究人员的关注,与此同时,模糊集合理论已经进入研究者的视野。采用基于模糊理论的图像处理技术本质是利用图像的模糊性参数实现对图像的优化处理^[4-7]。基于模糊熵^[8-10]的图像增强算法是基于图像的纹理特征,并以灰度、边缘特征、像素的分布以及像素点的差异等作为参考值,计算特征数学量来判别图像像素点属性的。该方法使每个像素点和其邻域像素点进行比对,根据设定的阈值滤除噪声点保留特征信息,但该方法对于对比度低的图像有较大误差,易造成图像失真。学者们在探索模糊集合理论过程中积累了很多经典的算法,比如Pal. S. K和King. R. A提出的Pal-King算法^[14-17],该算法被广泛应用。

图像增强^[11-12]包含频域增强和空域增强两种。图像频域增强方法有小波变换、contourlet^[13]变换等。文献[2]提出一种基于相似度测量的算法,利用边缘轮廓划分图像,对不同的轮廓进行图像增强。该算法结合灰度、纹理信息以及边缘特征,估算出图像像素的相似度特征量。Fang Shaomei^[18]等进行了区域一致性测度和四阶模型的图像增强研究,这种模型的去噪和边缘检测效果较好。Yang Yuqian^[19]等将自适应图像增强算法结合回归核和局部相似度,能在噪声抑制和边缘增强之间相互取舍,处理结果相对折中。Osinkina^[20]等研究了一种两步走的方法:准六方星形金纳米粒子有序阵列,该方法用以研究增强拉曼散射成像及其光谱。

国内也有众多学者利用相似度测量进行图像增强研究。如文献[3]在基于相似度的图像增强基础上,充分考虑医学超声图像实际情况,加入模糊分形维数特征量,进行图像增强。李沛轩^[21]等对小波变换和模糊熵理论进行了研究,将原始图像通过小波变换后得到高、低频小波系数,并对不同的系数采用不同的增强处理。杨先凤^[22]等提出了基于相似性度量的增强算法,该算法对一幅图像的区域像素进行窗口化处理,并根据相似性基础理论,对提取前后的图像进行相似性检测以达到增强的目的。

在图像增强过程中,准确的边缘提取对于图像增强效果的提升作用明显。边缘^[23]是指每个区域内像素点的灰度值产生阶跃性变化的像素点的集合。1965年,学者提出了边缘检测算子这个概念,经典的边缘检测算子主要是基于一阶微分的边缘检测方法^[24-25],该方法主要利用模板进行局部差分检测提取陡峭的边缘,但其对噪声敏感,使边缘不能获得较好的连接性。这些算法包括 Roberts 算子、Prewitt 算子以及 Sobel 算子等。基于二阶微分的检测方法^[26],包括 Laplacian 算子、LOG 算子^[27]以及由 John Canny 在 1986 年提出的 Canny 边缘检测算法^[28]。Canny 边缘检测算法主要能够滤除噪声、增强图像和检测出多阶段的边缘点,在众多的检测算子中优势明显。近几年,学者们还提出了基于数字形态学的边缘检测方法^[29],该方法主要从数学形态学出发,以形态结构元素去度量提取的边缘点,包括基于分形几何的边缘检测方法^[30]、标记松弛检测方法^[31]以及神经网络的边缘检测算法^[32]等。

1.2 特征点提取及配准现状

图像的特征信息是像素概念下的一种抽象表达,也可以表示为图像各种特征的组合。因此,可利用较少的特征信息来描述图像,这样能在很大程度上减少随机噪声、移动及背景等因素对图像特征信息的影响。目前图像特征主要包括曲面、闭合曲线以及特征点等,在选取所有的这些特征描述前景目标对象时,利用特征信息点进行描述的方法能有效减少信息的数据量,提高获得所有特征信息的速度,以达到实时处理的要求。

特征点主要是一些零交叉点和角点,提取方法主要包括使用 B 样条曲线描述的边界角点^[33]、使用多边形的边线角点^[34]、根据图像梯度方向检测出的角点^[35]等,其中包括 Smith 提出的 SUSAN 角点检测算法^[36]。该算法通过最小化图像像素点的区域,采用统计特性来判定该像素点是角点、边点还是圆点。Harris 和 Stephens 改进了 Plessey 角点法,提出了 Harris 算法^[37]。还有通过检测图像的链码曲率来判定角点的算法^[38],其中包括基于 Gaussian 曲率原则的检测方法,是由 Dreschler 和 Nagel 等^[39]提出的;D. G. Lowe 在 2004 年提出了基于尺度不变特征转换(Scale Invariant Feature Transform, SIFT)算法^[40],该算法首先提取局部特征,其次寻找尺度空间极值点,最后提取位置、尺度以及旋转不变量等。

特征点配准是将两幅图像上的特征点作为对象目标的匹配点,并把提取到的匹配点作为描述对象目标的特征信息。经典的特征配准算法常利用灰度相关性^[41]的思路进行研究,包括 Beardsley 等^[42]提出的通过提取边缘角点作为特征信息点,并根据特征点相关性进行匹配,但这种方法误匹配点数较多。SIFT 算法考虑使用提取的特征点描述子信息来进行匹配,同样也存在误匹配信息。通过匹配特征点获得匹配点集后,由于存在误匹配点,因此有必要进行误匹配点的剔除。传统的 RANSAC 算法^[43]是由 Fishler 和 Bolles 于 1981 年

提出的,该算法能够在误匹配点高达 50%的情况下,取得较多的精确匹配点。融合单应矩阵和极线约束特性作为匹配准则的算法,是由 Pritchett 和 Zisserman 等^[44]提出的,该算法的关键是假设某个特征点及其周边有限区域是空间平面成像,所以认为匹配点之间应近似满足单应矩阵关系。Lhuillier 和 Long Quan 等^[45]提出了一种稠密匹配的方法,该方法将对极几何全局约束、灰度相似性及单应矩阵有限区域约束方法融合起来,对处理纹理复杂的图像特征匹配效果明显,但精度不高。此外,目前还有很多剔除误匹配点的算法,比如最小平方中值法^[46]等。

1.3 动态目标检测现状

动态目标检测可实现对场景中运动目标的自动提取,常用的检测方式可分为以下几种:

(1) 基于运动信息的目标检测。该方法先对场景中运动目标信息进行分析,然后提取运动目标。

(2) 基于图像分割的运动目标检测。该方法将图像分割成不同的相似区域,然后根据相关准则检测分割区域并提取待检目标。

(3) 基于特征信息的目标检测。该方法根据目标的不同选择不同的特征信息来进行检测,首先通过学习,从样本空间中提取目标特征训练分类器,然后根据分类器进行判决,最后从图像中提取出相同特征的目标。

本书根据智能视频监控的要求,选择基于运动信息的目标检测。该方法通过检测视频序列在时间上的连续变化来提取运动目标。常见的检测算法有三种:帧间差分法(Temporal Differencing)^[47]、光流法(Optical Flow)^[48]和背景减法(Background Subtraction)^[49]。

在实际的场景中进行动态目标检测会受到如天气变化、光线突变、摇动的树叶、水面波动等因素的影响,而对于背景的更新过程还要考虑场景中物体的增加和减少情况。针对上述一系列问题,研究者提出了一系列的相关算法,而基于背景建模的背景差分算法,在应对复杂场景时具备较大的优势,因而成为研究的主流。目前该方法大致可分为以下几类^[50]。

1. 基于参数和非参数的背景建模方法

基于参数的建模方法需要事先假设每个像素点的特征值应该符合的分布模型,并准确地估计出该分布模型的参数模型,然后计算得到背景模型。然而,这种假设建立在背景符合一定分布的情况下,但现实的场景并不一定符合这种假设。为了能够处理某种不需要假设的检测场景,文献[51]中提出的非参数化模型,其不需要进行像素特征分布假设,而是通过直接计算特征值概率密度函数来建立概率统计模型。该方法完全依靠训练数据进行概率估计,无需过多先验条件,并可对任意形状目标概率密度进行估计。估计方式需要对场景进行训练,在训练的过程中收集到足够多的样本,再利用核概率密度对样本进行估计,使其收敛到特定的概率密度函数内。核密度预测(Kernel Density Estimation, KDE)方法^[52]是经典的非参数背景建模方法,其无需预先假设概率模型,而是通过足够多的样本估计背景像素的概率密度函数,最终估计出当前像素是否属于背景的概率,其特点是背景更新模型较简单。

由 Stauffer 和 Grimson^[53]联合提出的混合高斯模型(Mixed of Gaussian, MoG)是一种经典的参数化建模方法,该方法为了适应背景的动态变化对每个像素使用多个高斯分布加权进行建模。MoG 使用混合高斯分布来描述图像中每个像素点的特征,通过对背景进行建模以及快速更新可以很好地适应背景的动态变化,通过计算像素的概率密度函数,判断该像素是否是前景目标。MoG 估计的是单个像素在时间序列上的分布,不适用于背景突变的情况。MoG 在室外场景的应用中也遇到一些问题,比如对运动缓慢且纹理相似、面积较大的物体检测时会出现空洞现象,主要是由于整幅图像是一个整体而忽略了图像中像素点的空间关联性,以至于图像受噪声的干扰比较明显。可以通过增加高斯分布的个数来改善检测的效果,但这样又产生了较为复杂的参数调整过程。

上述两种背景建模方法都各有优势,参数化的方式可以减少系统的开销,精简模型的描述在已知分布模型的情况下能达到较好的检测效果;而非参数模型不用指定潜在的分布模型,不需要对参数进行估计,但随着时间的推移需要建立适应多种场景的背景,将占用较大的存储开销。在复杂的计算机视觉应用中,往往存在大量的数据,数据之间相互联系且复杂多变,没有特定的规则形态,这时非参数的方法更能描述场景的变化,从时间和空间复杂度来看,参数化背景建模方法更占有优势。

2. 基于灰度、颜色和纹理的背景建模方法

基于灰度和颜色信息的特征选择是背景模型建立的关键,如前面介绍的 MoG 和 KDE。基于灰度和颜色信息的背景建模算法使用较多,一般能达到很好的检测效果,但是当运动目标和背景颜色相近时,较难检测出前景运动目标,这主要是因为相似物干扰时,检测的目标和背景区分度会降低,产生误检。为了解决这种问题,Marko Heikkila 等^[54]利用图像的纹理特征对背景进行建模,可以很好地解决相似背景的干扰。这种方法通过计算图像的局部二值化处理来描述背景,能减少局部光照变化产生的影响,但是它的检测精度不如 MoG。因此,更多的研究是将这种算法与其他的算法相结合。文献^[55]中结合了像素的灰度和图像的边缘纹理信息,对背景进行建模,实现了特征的多种组合建模方式,很好地利用了不同特征背景建模的优势。

3. 基于像素和区域的背景建模方法

根据每个像素在时间序列上分布的不同来建立背景模型是基于像素的背景建模方法^[56];而基于图像区域的建模方法除了考虑像素本身的特征之外,还充分利用了相邻像素间的空间邻域关系,将背景图像分成众多相同的块,这些块之间或重叠或分离,并从各图像块中提取有效的特征对背景进行建模。在进行运动目标检测时,将图像背景中的区域方块作为整体性特征进行匹配,从而检测出运动目标块。基于像素的背景建模方法假设每个像素在时间序列上的特征值是相互独立的,主要关注自身的特征变化,容易受动态背景和图像噪声干扰。两种建模方法比较,基于图像局部区域的背景建模方法,在考虑图像的局部整体性时,加入了相关像素对区域的影响,这种方式在处理图像的局部突变和较强的背景干扰时效果较好,且能从图像块中提取比单个像素更丰富的特征,能够适应更多的环境变化,处理更多的复杂场景,但是不能得到精确的运动目标,其检测的效果还与所选的特征有关,并且在时间复杂度和空间复杂度上过多依赖所选择的特征块的大小。

随着计算机视觉以及视频分析技术的进一步发展,视频目标检测取得了很多的研究成

果,且在实际应用中能解决部分难题。但是目前还没有被广泛认可的方法能很好地解决所有问题,各种方法都有自己的优势和不足。算法的准确性和鲁棒性越好,算法的实现就越复杂,实时性越差;相反,实时性较强的算法实现起来较简单,但是在抗干扰和背景突变方面能力较差,检测的准确性较低。因此,这些算法在准确性、鲁棒性以及实时性之间的权衡方面还有待细化研究。

1.4 遗留物检测现状

遗留物检测作为一个新兴的研究课题,其相关研究工作也逐步展开。Beynon^[57]等人使用多个摄像机来确定物体的位置,经过场景分割、物体分类、3D跟踪等一系列的步骤来判断是否为遗留物。Chuan-Yu Cho^[58]等人,则利用遗留物在场景中固定不变的特征,通过使用垂直扫描线不断扫描前景目标,如果物体与视频顶部的距离始终不变,则判为遗留物。Fatih Porikli^[59]等人提出了双背景模型的方法检测遗留物或者违规停车,每个背景模型都由 MoG 组成,但是两个背景模型的更新速率不同,通过比较两个模型检测结果的差异,来确定遗留物。Roland Miezancko 和 Dragoljub Pokrajac 等^[60]提出一种在复杂场景下检测和识别遗留物的方法,该方法使用空时纹理特征检测运动目标块,利用支持向量机(Support Vector Machine, SVM)训练分类器对运动块进行分类来判断是否为遗留物。但是由于现实中的遗留物千差万别,该方法检测效果并不理想。文献[61]先使用 MoG 进行检测,再使用均值漂移(Mean Shift)进行跟踪,最后使用不变矩特征训练的分类器来判断是否存在遗留物。

总的来说,这些方法分为两大类:基于跟踪的方法^{[57][60]}和基于检测的方法^[62-63]。基于跟踪的方法受算法复杂度的限制,不适用于背景复杂、人流密集的场所;而基于检测的方法往往使用多层背景模型^[64],通过控制模型的更新速度来检测遗留物,并需要其他辅助算法来排除遗留物被人重新移动后产生的“鬼影”影响。

尽管遗留物检测研究已经取得了一些进展,但是还存在一些问题亟待解决,其中包括:遗留物长时间停留在场景下,进而使其融入背景模型;检测时间较长时若光照发生变化,会产生一定的误检;遗留物被人重新移动后产生“鬼影”现象等。

1.5 动态目标跟踪现状

由于现实监控环境的复杂性,光照变化、目标运动不规则、遮挡、相似物干扰以及各种人为和自然条件下的干扰都会影响目标检测跟踪的准确性、实时性和鲁棒性。因此,如何在各种复杂条件下完成对目标的有效检测跟踪是实现智能视频监控的关键。

动态目标跟踪技术^[65]需要从被检测的场景中自动寻找到目标,然后根据要求来选择被跟踪目标,并通过跟踪算法定位目标的位置,对目标进行持续的跟踪。动态目标的检测是跟踪的前提,只有准确地检测出运动目标才能更好地完成后续的跟踪等智能分析。所以说动态目标跟踪是一座衔接目标检测和目标行为分析与理解判断的桥梁。图 1.1 为动态目标检测及跟踪示意图。

视频图像处理技术是视频序列中动态目标跟踪算法的基础,其根本目的是从视频图像

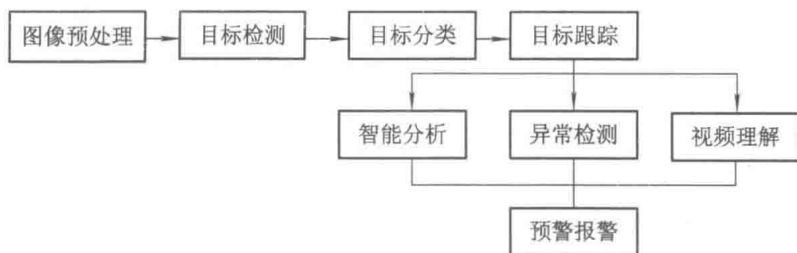


图 1.1 动态目标检测及跟踪示意图

的每一帧中捕获被跟踪目标的真实位置，并结合行为分析及目标识别准确定位目标。根据跟踪准则的不同可分为如下跟踪场景^[66]：单个与多个摄像头；摄像头静止与运动；单个目标跟踪与多个目标同时跟踪；刚性跟踪与非刚性跟踪；可见光与红外跟踪等。本书研究重点是单个摄像头、运动背景相对固定、非刚性、可见光成像条件下的单个目标的跟踪算法。

目前目标跟踪算法分为两类：一类是基于多模式的目标跟踪方法；另一类是基于多特征的目标跟踪方法。

1. 基于多模式的目标跟踪方法

在复杂场景中使用单一方法对目标进行跟踪时经常出现一些问题，主要原因是：

(1) 跟踪过程是复杂多变的，而目标本身又是不断变化的，单一方法不可能处理所有情况；

(2) 很多算法的提出是在一定的条件下实现的，对算法的使用场景有一定的限制，对于不同的跟踪对象跟踪效果存在差异。常用的视频序列中目标跟踪算法主要有自上而下的跟踪过程和自下而上的过程。

在自上而下的方法中常见的是 Kalman 滤波法和粒子滤波法。Kalman 滤波法通过对目标位置参数预测估计，实时调整完善跟踪，具有计算量小等优点，但是面对非线性、非高斯问题时，就会失去跟踪的有效性。为了解决非线性系统带来的跟踪问题，学者们提出了一种扩展的 Kalman 滤波结合随机变量分布的算法来近似非线性，并称之为无味 Kalman 滤波算法。但这种近似方式都会引入误差，从而使有效信息不能正确利用，导致跟踪误差增大。粒子滤波算法可以在非线性、非高斯的条件下完成目标的跟踪，是近年来研究的重点，在跟踪领域得到广泛应用。但是，粒子滤波的算法复杂、计算量大，在粒子经过多次的迭代后会出现严重的粒子退化现象，极大地影响了目标跟踪的持续性。后来的研究者提出了许多的改进算法来解决粒子退化问题。Gordon 等人^[67]提出了重采样机制，成为粒子滤波中不可缺少的一步，可是这种机制会出现样本贫化等问题，解决该问题的有效方法是增加样本数，但是这样就会大大增加计算的复杂度。大量研究表明在环境较复杂情况下，粒子滤波比 Kalman 滤波的跟踪性能更优秀。

在自下而上的跟踪过程中，最常用的是 1975 年由 Fukunage 等人^[68]提出的 Mean Shift 算法及其改进算法。Mean Shift 算法是一种无参估计方法。Comaniciu 等人^[69]在图像分割和目标跟踪中使用的 Mean Shift 算法，没有对目标模板进行必要的更新，而且选择的是固定不变的核函数窗宽，不适用于目标有尺度变化的情况。经过研究者近几年的不断努力，均值平移算法有了较快的发展，这主要是由于 Mean Shift 算法具有优良稳定的特性、较好的实时性、计算量小、易于实现和集成、对环境的变化和目标的形变不敏感以及对复

杂环境具有较强的适应性等优点。但是该算法在实际应用中也存在不足, Mean Shift 是通过不断地迭代寻找最佳的匹配中心的, 当目标运动较快时, 只能通过增加迭代次数寻找目标, 加大了时间复杂度, 影响了跟踪效果, 而且易受到周围相似物的干扰。

通过对上述常用算法的分析可以看出, 目前常用的跟踪算法都是基于某种特定的假设提出的, 要实现对目标的稳定跟踪, 单一的目标跟踪算法略显不足, 所以很多研究者就采用多模式的方式来实现跟踪。一方面, 多模式可以克服单一模型中的不足; 另一方面, 多种模式的合理结合也有效地提高了跟踪的可靠性和鲁棒性。在合理结合的过程中也产生了许多新的思想, 为跟踪技术的进一步发展提供了良好的基础。文献[70]将广泛使用的 Mean Shift 和粒子滤波结合, 该算法对目标分别采用 Mean Shift 和粒子滤波进行跟踪, 以此得到目标位置候选值, 并利用加权混合参考函数来判断当前目标的准确位置。结合 Mean Shift 和粒子滤波在目标跟踪中的各自优势, 可以有效地解决跟踪中经常遇到的遮挡、扭曲以及旋转等情况, 提高跟踪精度, 增强了抗干扰的能力。多模式融合中基于 Kalman 滤波和 Mean Shift 的方法较多, 这种方式既克服了原有算法的不足, 又易于实现, 极大地提高了跟踪性能。但是这种融合并非简单的算法叠加, 而是通过某种机制完成算法的提升。在采用多模式时, 不是模式越多越好, 模式越多越会增加算法实现的复杂度和计算量, 需要在实现融合的过程中综合考虑可靠性和实时性。

2. 基于多特征的目标跟踪方法

在基于特征的目标跟踪方法中, 由于单个特征空间, 在面对复杂环境时不能很好地解决特征相似目标的干扰、光照的变化、特征的突变等问题。在目标跟踪过程中需要对目标进行尽可能多的描述, 如何对跟踪对象有效的描述也是跟踪成功的关键。显然, 单个特征在跟踪时无法提供完整的信息, 需要将单个特征拓展到多特征空间。基于多特征融合的目标跟踪也是近年来使用较多的方式, 这些特征除了常见的颜色、灰度外, 还有纹理、梯度、边缘以及形状等。使用多特征跟踪可以在一种特征失效时, 选择另一种特征继续跟踪。通过多特征中选择区分度较高的特征集合, 可降低算法复杂度, 实现目标的持续跟踪。文献[71]为了解决相似特征的干扰, 利用构建目标与背景图像特征分布方差的比值函数来衡量目标与背景之间的区分度, 并采用各类特征的区分度对特征集进行线性加权, 依据各区分度来进行特征的选择, 使其相互补充, 该算法能有效地对复杂背景下的运动目标进行跟踪。基于多特征的跟踪在特征的选择上并非想象的越多越好, 过多的特征可能会降低跟踪的时效性, 在不同的场景中, 选择最能表示目标的特征作为主要特征, 一般场景中选择 2 到 3 个特征就能达到很好的效果, 继续增加特征数量的效果提升并不明显。

1.6 本书研究内容与技术路线

1.6.1 研究内容

本书针对现实场景中视频图像受噪声干扰大、应用环境复杂以及动态目标智能化安全检测等问题展开研究。主要研究内容如下:

(1) 对相关方法进行了现状分析。

(2) 视频图像的增强。一方面希望提高图像边缘的纹理信息；另一方面又要削减图像的噪声分量。实际上，在除去噪声的同时会削弱图像的边缘信息，增强图像边缘信息的同时也会使一部分噪声得到增强。本书研究相似度和模糊熵判别准则，针对图像不清晰的普遍性问题，利用特征分形维数对比度变换算法合理提取图像信息。为了获得适量而且确切的图像信息，需要对图像通过边缘检测来进行图像分割。本书提出基于小波和 Canny 算法融合的边缘检测算法，使得图像轮廓特征纹理更为清晰。

(3) 视频图像的特征点提取及配准。为了解决经典算法配准产生的误匹配点多的问题，本书提出一种基于相关法的欧氏距离配准算法。该算法通过利用不同特征点自身信息，在 Harris 算法基础上分别对灰度信息使用梯度相关法，对 SIFT 算法描述子信息使用描述子相关法，并结合特征点间的欧氏距离关系来精确匹配，实验结果表明该方法降低了误匹配的点数。

(4) 动态目标的检测。为了解决前景目标检测过程中缓慢和局部运动的问题，本书通过提取运动目标的空间整体信息，提出在处理复杂背景环境时充分利用目标的空间信息，对聚类信息进行有效地描述，实现对视频目标的完整、准确检测，减少运动信息不足时的目标误检率。针对三原色红、绿、蓝(RGB)空间下的码书模型(Code Book Model, CBM)的不足，本书引入色调、饱和度、亮度(Hue, Saturation, Value, HSV)空间下的 CBM，并构建每个像素点的加权平均背景，利用该背景结合阴影检测算法，去除了前景检测中阴影的干扰。

(5) 遗留物的检测。讨论 MoG 和 CBM 在遗留物检测方面存在的不足，介绍了基于双层背景模型的遗留物检测算法。并通过研究遗留物像素点的灰度分布值，提出通过计算每个不属于背景模型的像素点与其之前若干帧组成的历史像素集中的像素值的匹配程度，来判定该像素点是否稳定，是否为遗留物。

(6) 动态目标的跟踪。为解决多特征描述进行匹配跟踪的问题，本书基于 Mean Shift 算法理论，在多特征融合跟踪基础上，提出多准则判定的多特征自适应匹配算法。多特征自适应匹配算法可以在复杂条件下更好地对目标进行描述，实现了特征集合之间准确的匹配跟踪。选择多种判定准则，实现了基于多种途径的特征匹配，并结合 Kalman 滤波算法，实现了对目标的稳定准确跟踪。针对离散场景下的运动目标跟踪问题，本书提出了将帧差法、SIFT 与 CBWH 算法融合的跟踪算法，并且详述了该算法的跟踪过程与步骤，实验结果表明，本书算法能够在离散场景中较好的找到同一目标。

(7) 基于三频彩色条纹投影轮廓术的微变监测。其采用彩色条纹投影与立体视觉融合的三维传感方法，以实时测量动态复杂物体为目标。该技术首先利用 BEMD 的自适应条纹分析技术解决三频彩色条纹的颜色解耦难题；其次，利用傅立叶变换实现变精度全场包裹相位展开，三频变精度得到高精度绝对相位；最后，通过标定系统，恢复物体高程信息，完善立体监测。

(8) 附录中梳理了包括视频的采集、压缩等相关程序，可供参考。

1.6.2 技术路线

图 1.2 为本书研究的动态目标视频监控整体框架，其核心步骤包括通过图像增强获得纹理清晰的视频图像，通过边缘提取及特征点配准获得精确匹配，对图像中的运动目标进