

信息科学技术学术著作丛书

主动相机运动分割 与目标跟踪理论

崔智高 李艾华 王 涛 苏延召 著



科学出版社

信息科学技术学术著作丛书

主动相机运动分割 与目标跟踪理论

崔智高 李艾华 王 涛 苏延召 著

科 学 出 版 社

北 京

内 容 简 介

主动相机智能监控是当前智能视频监控的主要发展趋势。本书紧紧围绕主动相机智能监控的两大核心技术——运动分割与目标跟踪展开讨论,包括主动相机智能监控、主动相机的运动目标分割技术、主动相机的运动目标跟踪技术,以及主动相机监控系统设计和应用,同时介绍相关方法的研究背景、理论基础和算法描述,并给出相应的实验结果。

本书可作为大专院校及科研院所模式识别、图像处理和机器视觉等领域的高年级本科生、研究生的教材,也可作为相关领域教师、科研人员,以及安防和视频监控行业工程技术人员的参考书。

图书在版编目(CIP)数据

主动相机运动分割与目标跟踪理论/崔智高等著. —北京:科学出版社, 2018. 4

(信息科学技术学术著作丛书)

ISBN 978-7-03-056688-1

I. ①主… II. ①崔… III. ①目标跟踪 IV. ①TN953

中国版本图书馆 CIP 数据核字(2018)第 042032 号

责任编辑:魏英杰 / 责任校对:桂伟利

责任印制:张 伟 / 封面设计:陈 敬

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>

北京中石油彩色印刷有限责任公司印刷

科学出版社发行 各地新华书店经销

*

2018 年 4 月第 一 版 开本:B5 (720×1000)

2018 年 4 月第一次印刷 印张:10 1/4

字数:205 000

定价:90.00 元

(如有印装质量问题,我社负责调换)

《信息科学技术学术著作丛书》序

21 世纪是信息科学技术发生深刻变革的时代,一场以网络科学、高性能计算和仿真、智能科学、计算思维为特征的信息科学革命正在兴起。信息科学技术正在逐步融入各个应用领域,并与生物、纳米、认知等交织在一起,悄然改变着我们的生活方式。信息科学技术已经成为人类社会进步过程中发展最快、交叉渗透性最强、应用面最广的关键技术。

如何进一步推动我国信息科学技术的研究与发展;如何将信息技术发展的新理论、新方法与研究成果转化为社会发展的新动力;如何抓住信息技术深刻发展变革的机遇,提升我国自主创新和可持续发展的能力?这些问题的解答都离不开我国科技工作者和工程技术人员的要求和艰辛付出。为这些科技工作者和工程技术人员提供一个良好的出版环境和平台,将这些科技成就迅速转化为智力成果,将对我国信息科学技术的发展起到重要的推动作用。

《信息科学技术学术著作丛书》是科学出版社在广泛征求专家意见的基础上,经过长期考察、反复论证之后组织出版的。这套丛书旨在传播网络科学和未来网络技术,微电子、光电子和量子信息技术、超级计算机、软件和信息存储技术,数据知识化和基于知识处理的未来信息服务业,低成本信息化和用信息技术提升传统产业,智能与认知科学、生物信息学、社会信息学等前沿交叉科学,信息科学基础理论,信息安全等几个未来信息科学技术重点发展领域的优秀科研成果。丛书力争起点高、内容新、导向性强,具有一定的原创性;体现出科学出版社“高层次、高质量、高水平”的特色和“严肃、严密、严格”的优良作风。

希望这套丛书的出版,能为我国信息科学技术的发展、创新和突破带来一些启迪和帮助。同时,欢迎广大读者提出好的建议,以促进和完善丛书的出版工作。

中国工程院院士

原中国科学院计算技术研究所所长

Handwritten signature in black ink, consisting of three stylized characters: '李' (Li), '小' (Xiao), and '立' (Li).

前 言

随着世界范围内安全意识的不断提高,智能视频监控的研究在军用、民用领域都得到了普遍的重视,因此对智能视频监控的研究具有非常重要的意义。智能视频监控是指以监控摄像机作为前端采集设备,利用计算机视觉的相关技术,对采集的视频序列进行智能分析,从而辅助管理人员获取监控场景中的有用信息,并根据具体需要发出预警信号或实施相关决策。

在智能视频监控的相关研究中,基于单目静止相机的研究已经非常深入,并且广泛应用到很多场合。然而,由于实际监控场景的多样性和复杂性,人们对智能视频监控技术往往有着更高的要求,如场景的多视角监控、目标的多尺度信息等,传统的单目静止相机由于其固有的局限性已经无法满足此类需求。目前,采用主动相机是智能视频监控的主要发展趋势,特别是随着相机制作工艺的完善、机械控制精度的提高,以及生产成本的降低,主动相机的逐渐普及已经成为必然。

本书针对主动相机智能监控的关键技术展开讨论,并重点研究了主动相机智能监控的两大核心技术——运动分割与目标跟踪。全书共7章,从内容看可以分为三个部分。第1章绪论,介绍主动相机智能监控的优势和研究现状;第一部分为主动相机的运动目标分割技术,分别介绍基于多组低秩约束的主动相机运动分割算法(第2章)和基于光流场分析的主动相机运动分割算法(第3章);第二部分为主动相机的运动目标跟踪技术,分别介绍基于地平面约束的静止相机与主动相机目标跟踪算法(第4章)、基于球面坐标和共面约束的双目主动相机目标跟踪算法(第5章),以及基于特征库构造和分层匹配的主动相机参数自修正算法(第6章);第三部分主要介绍基于主动相机的智能监控系统设计及应用(第7章)。

本书由崔智高拟订大纲,撰写第2~5章,并对全书进行统稿,由李艾华撰写第1章,苏延召和王涛撰写第6、7章。在著述过程中得到了火箭军工程大学学校机关、机电教研室的支持和帮助,在此一并表示感谢。另外,感谢研究组的徐斌、姚良、李辉、蔡艳平、金广智、李庆辉、马鹏程、袁梦、鲍振强等同事和同学,为本书提供了很多有价值的素材,并协助完成了纷繁的工作。

主动相机运动分割和目标跟踪是尚在发展中的新技术,限于作者水平,书中不妥之处在所难免,敬请读者指正。

作 者

2017年12月于西安

目 录

《信息科学技术学术著作丛书》序

前言

第 1 章 绪论	1
1.1 智能视频监控技术	1
1.2 主动相机智能监控	4
1.2.1 主动视觉	4
1.2.2 三种常用相机	4
1.2.3 智能视频监控研究的发展趋势	6
1.3 主动相机运动分割和目标跟踪的研究现状	7
1.3.1 主动相机运动分割的研究现状	7
1.3.2 主动相机目标跟踪的研究现状	11
1.3.3 双目视觉系统的研究现状	12
1.4 本书内容安排	16
参考文献	18
第 2 章 基于多组低秩约束的主动相机运动分割算法	24
2.1 引言	24
2.2 基于多组低秩约束的运动轨迹分类	25
2.2.1 背景运动的低秩约束	26
2.2.2 运动轨迹分类方法	26
2.3 视频序列过分割	29
2.4 马尔可夫随机场目标分割模型	29
2.4.1 数据项惩罚函数	30
2.4.2 正则项惩罚函数	32
2.5 实验结果及其分析	33

2.5.1 运动轨迹分类实验结果	33
2.5.2 像素一级目标分割实验结果	35
参考文献	38
第3章 基于光流场分析的主动相机运动分割算法	40
3.1 引言	40
3.2 基于光流梯度幅值和光流矢量方向的目标边界分割	41
3.2.1 光流梯度幅值确定边界	41
3.2.2 光流矢量方向确定边界	42
3.3 基于点在多边形内部原理的目标像素点判断	42
3.4 基于时空马尔可夫随机场模型的前背景像素标记	44
3.4.1 时空马尔可夫随机场模型能量函数设计	44
3.4.2 数据项势能函数设计	44
3.4.3 平滑项势能函数设计	45
3.5 实验结果及其分析	47
参考文献	50
第4章 基于地平面约束的静止相机与主动相机目标跟踪算法 ...	52
4.1 引言	52
4.2 现有协同跟踪方法分析	53
4.2.1 模型拟合方法	54
4.2.2 完全标定方法	56
4.2.3 图像拼接方法	56
4.3 基于地平面约束的协同跟踪算法的总体框架	57
4.4 基于地平面约束的协同跟踪算法的具体实现	59
4.4.1 单应矩阵自标定	59
4.4.2 主点和等效焦距估计	64
4.4.3 静止相机多目标跟踪	66
4.4.4 主动相机参数估计	67
4.5 实验结果及其分析	70
4.5.1 室内场景实验结果	71
4.5.2 室外场景实验结果	72

4.5.3 方法对比	75
参考文献	78
第5章 基于球面坐标和共面约束的双目主动相机目标跟踪算法 ...	
.....	80
5.1 引言	80
5.2 基于球面坐标的目标跟踪算法	81
5.2.1 基于特征匹配的摄像机内部参数获取	81
5.2.2 基于球面经纬坐标的公共坐标系建立	84
5.2.3 结合场景深度范围的协同跟踪	87
5.3 高分辨率全景图生成算法	90
5.3.1 由粗到精的配准模型估计	91
5.3.2 协同运动分离	95
5.3.3 背景区域和前景区域的依次映射	96
5.4 基于共面约束的目标跟踪算法	97
5.4.1 共面约束的引入与估计	97
5.4.2 系统具有的两个性质	98
5.4.3 协同跟踪的实现	100
5.5 实验结果及其分析	101
5.5.1 球面坐标协同跟踪的实验结果	101
5.5.2 高分辨率全景图生成的实验结果	104
5.5.3 共面约束协同跟踪的实验结果	108
5.5.4 两种协同跟踪算法的对比分析	109
参考文献	110
第6章 基于特征库构造和分层匹配的主动相机参数自修正算法 ...	
.....	112
6.1 引言	112
6.2 主动相机参数的误差	113
6.2.1 机械误差	113
6.2.2 累积误差	114
6.3 参数自修正算法的总体框架	115
6.4 参数自修正算法的具体实现	117

6.4.1 特征库构造	117
6.4.2 分层匹配	120
6.4.3 参数修正	122
6.5 实验结果及其分析	124
6.5.1 算法的有效性验证	124
6.5.2 算法的定量比较	128
参考文献	129
第7章 基于主动相机的智能监控系统设计及应用	130
7.1 引言	130
7.2 系统结构设计	130
7.3 系统软件界面	132
7.4 系统应用模块	134
7.4.1 预设位置的自主巡航监视及全景图拼接	134
7.4.2 视频序列的运动目标分割	136
7.4.3 预警区域的多目标依次跟踪	138
7.4.4 重点人员的人脸跟踪与识别	141
7.4.5 深度信息辅助的高分辨率图像获取和遮挡目标跟踪	144
7.4.6 大视场下目标的多分辨率主动跟踪	147
参考文献	150

第 1 章 绪 论

1.1 智能视频监控技术

人类可以利用双眼获取周围场景中的不同颜色和光照,并通过视觉感知系统整合这些信息,同时实现区分物体深度、描述物体特征,以及识别不同物体等智能行为。针对人类对周围物体的这种智能感知能力,许多生理科学、神经科学和认知科学的研究者进行了深入的探索。目前的研究普遍认为,人类的视觉感知能力主要包含三个层次^[1],即神经响应能力、底层视觉能力和中高层感知能力。神经响应能力主要包括颜色区分能力、光照辨别能力等;底层视觉能力包含深度区分能力、运动感知能力、空间敏锐力等;中高层感知能力主要用于提取物体特征、识别不同物体,以及获取语义信息等。

为了使计算机能够模拟人类的上述视觉感知功能,以辅助或代替人类完成一些视觉相关的任务,逐渐形成了计算机视觉这门新兴的学科^[2]。一般认为,计算机视觉起源于 20 世纪 70 年代早期,虽然当时已经存在数字图像处理这个相似的领域,但当时的计算机视觉更关注场景三维结构的恢复,进而实现场景理解的目的,因此早期开展了很多立体视觉、三维结构,以及多视图几何的经典工作。进入 20 世纪 80 年代,更多的研究工作关注视觉分析的复杂数学方法,即将计算机视觉的相关问题纳入数学框架中进行描述,从而可以借助更多的数学工具解决复杂的视觉问题。20 世纪 90 年代以后,计算机视觉的诸多子方向仍然在广泛研究中,但是其中一些问题明显更加活跃,如图像分割、目标跟踪、光流计算等,同时机器学习方法开始逐渐流行,并成功地应用于人脸识别的主分量本征脸分析^[3]。进入 21 世纪,计算机视觉与其他领域的相互影响和交叉融合进一步加深,特别是复杂的机器学习方法在计算机视觉问题中的应用,主导着计算机视觉中很多问题的研究,如物体

识别、物体分割、场景理解等。

经过几十年的不断发展,时至今日,计算机视觉已经成为计算机科学领域最热门的研究方向之一,然而由于计算机视觉涉及信号处理、应用数学、统计学、物理学、计算机科学和认知科学等多个领域,而且其信息来源具有复杂性、多样性和欠定性等特征,因此计算机视觉也被认为是计算机科学领域最具有挑战性的研究方向之一^[4]。

根据信息载体的不同,对计算机视觉的研究可以分为针对静止图片的研究和针对视频序列的研究。针对静止图片的研究更为基本,并且静止图片的采集设备也更为普及,因此目前针对静止图片的研究已经非常广泛。随着视频采集设备的普及,视频源越来越多,针对视频序列的研究也越来越受到重视^[5,6]。相对于静止图片,视频序列由于多了一个时间维度,因此蕴含更多的信息,并且随着视频采集设备的成本越来越低,采集质量越来越高,更多的实际系统采用记录视频序列而非静止图片的方式进行,因此开展针对视频序列的研究具有重要意义。

智能视频监控(intelligent video surveillance, IVS)^[7-9]就是其中的一个重要应用方向。智能视频监控是指以监控摄像机作为前端采集设备,利用计算机视觉的相关技术,对采集的视频序列进行智能分析,如图像预处理、运动目标的分割与跟踪、目标的分类与识别、行为分析和异常检测等,从而辅助管理人员获取监控场景中的有用信息,并根据具体需要发出预警信号或实施相关决策。图 1.1 给出了智能视频监控的典型结构框架。

随着世界范围内安全形势的日益严峻、人类公共安全意识的不断提高,以及军事安全的迫切需要,智能视频监控技术已经渗透到人类日常生活的各个方面。例如,在民用方面,针对银行、广场、停车场等重点场所的监控,可以防止盗窃、暴力、破坏等危害公共安全和公共财产的事件发生,从而保证人民群众的人身安全和财产安全;在军用方面,针对导弹阵地、关键出入口、重点涉密场所的监控,可以防止摧毁、入侵、窃密等危害军事安全的行为发生,从而保证重点军事设施不受损害,保障作战任务的顺利执行。

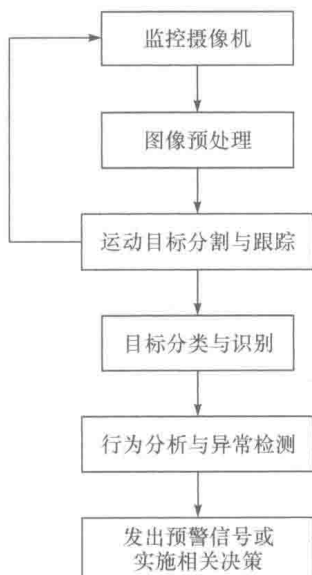


图 1.1 智能视频监控的典型结构框架

同时,智能视频监控也逐渐受到学者的普遍重视,并成为计算机视觉领域的研究热点之一。国内外许多重要期刊,如 *SPM*(*IEEE Transactions on Signal Processing Magazine*)、*IJCV*(*International Journal of Computer Vision*)、*TPAMI*(*IEEE Transactions on Pattern Analysis and Machine Intelligence*)、*CVIU*(*Computer Vision and Image Understanding*)、*MVA*(*Machine Vision and Application*)、*IVC*(*Image and Vision Computing*)、计算机学报、自动化学报等都针对智能视频监控设立了专栏或者是专刊,同时一些有较大影响力的国际计算机视觉会议,如 *ICCV*(*International Conference on Computer Vision*)、*ECCV*(*European Conference on Computer Vision*)、*CVPR*(*Computer Vision and Pattern Recognition*)、*ACCV*(*Asian Conference on Computer Vision*)等都将视频监控纳入议题,有的还为此组织了专题研讨。

综上所述,智能视频监控技术具有重要的研究意义和实用价值,并受到工业界和学术界的广泛重视,因此有必要进一步开展智能视频监控技术的相关研究。

1.2 主动相机智能监控

本书主要针对智能视频监控中的主动相机监控技术。所谓主动相机,指的是一种可以改变水平转动(pan)、垂直转动(tilt)和焦距大小(zoom)的摄像机,也称为 PTZ 镜头。本节首先介绍主动相机智能监控的仿生学基础——主动视觉;然后比较智能监控中使用最多的三种相机,并分析主动相机的优势;最后阐述智能视频监控发展的主要趋势。

1.2.1 主动视觉

源于利用计算机视觉技术模拟人类视觉特性的初衷,研究者逐渐从仿生学的角度研究智能视频监控技术。现有研究发现,人类视觉最显著的特点是具有视觉关注性,即观察者的注意力总是有目的地指向他感兴趣的运动目标,而忽略其他事物,这是人类视觉主动性的一种体现^[10]。也就是说,人类通过图像的主动采集和视觉的关注特性,为感知系统提供目标的多视角和多尺度信息,从而辅助感知系统利用已有知识对目标进行分析和识别,该视觉特性一般称为主动视觉^[11,12]。

一般认为,主动视觉是一种智能的图像采集方式,它根据系统的最终任务、已采集图像数据的分析结果及其误差来动态控制图像的采集过程。这与人眼的转动机制是一致的,即人类一般是有目的地对场景进行观察,当场景提供的信息无法满足要求时,人眼会转动一定的角度,聚焦直至获得目标清晰的图像。从数学的角度,图像采集的主动性能够简化视觉模型的推导,并使很多传统视觉中欠定的问题变得适定,或使传统视觉中的非线性问题线性化^[11,12]。正因为如此,主动视觉逐渐受到智能监控学术界和工业界的重视^[13]。

1.2.2 三种常用相机

监控摄像机是智能视频监控的重要组成部分,目前主要采用的监控摄像机包括静止相机、全向相机和主动相机三种。图 1.2 和表 1.1 分别给出了三种摄像机的示意图及其优缺点对比。



图 1.2 三种常用的监控摄像机

表 1.1 三种常用监控摄像机的优缺点对比

摄像机类型	优点	缺点
静止相机	价格低廉;图像畸变小; 研究相对深入和广泛	视角固定;分辨率单一
全向相机	监控视场大;监控范围广	图像分辨率低;图像畸变程度大; 标定过程复杂
主动相机	视角、分辨率可变;使用灵活; 图像畸变程度小	成本较高;不同厂家的机型 属性差异大

静止相机的价格相对低廉,图像畸变也较小,而且多数视觉分析算法都是针对静止相机采集的视频序列,研究相对深入和广泛,但是静止相机的视角无法改变,而且获得的图像分辨率一般也固定。

在全向相机中,目前使用较多的包括鱼眼相机和折反射式相机,前者实际上是一种短焦距超广角相机,后者则一般由一个传统的电荷耦合器件(charge coupled device, CCD)视觉传感器和一组反光镜片装置构成^[14]。本质上,全向相机仍然属于静止相机,但它可以观察到的视场比静止相机大很多,因此在大场景监控及机器人视觉中已经越来越多地被采用。然而,全向相机仍然具有不可忽视的缺陷,主要表现在图像分辨率低,并且图像畸变程度较大。此外,标定精度容易受图像分辨率限制,标定过程也相对复杂,特别是在标定精度要求较高的场合。

得益于主动视觉研究的深入,以及成像和机械控制技术的发展,主动相机越来越多的应用于智能监控领域,并且逐渐取代传统的静止相机和全向相机^[15]。主动相机可看做一个内外参数均可变化的静止相机,这些参数包括水平转动、竖直转动,以及焦距变化,通过调整这些控制参数,主动相机不但可以改变焦距,从而获得同一目标或区域的不同

分辨率信息,而且可以改变监控视角,从而获得场景中不同目标或区域的监控信息。主动相机的主要缺点包括成本相对较高,特别是控制精度很高的机型;不同厂家的相机属性差异较大,具体表现在相机自由度、控制方式、相机动作范围,以及参数控制精度等。

随着计算机视觉技术的发展,智能视频监控技术已经取得长足的进步,然而目前的绝大多数智能视频监控研究都是基于静止相机,全向相机由于图像畸变严重,研究也相对较少。本质上,静止相机和全向相机的智能监控技术均属于被动视觉的范畴,并不能完全满足智能视频监控的实际需求。例如,在实际应用中,安全值守人员通常希望监控系统能够持续跟踪可疑目标,使其始终处于监控视野内,并获得目标的高分辨率图像,从而为日后事件的回溯提供依据。虽然通过布设多个静止相机,利用预先调整各个相机焦距的方式也可以实现这一目标,但系统配置复杂、灵活性较差,并且即使能够获得目标的高分辨率图像,其获取区域也是固定的。反观主动相机,它的水平转动、竖直转动,以及焦距变化参数具有可变性和可控性,配置简单、视角灵活、分辨率多样,通过设计合适的视觉分析算法,可以方便地获取目标的多视角和多尺度信息。

1.2.3 智能视频监控研究的发展趋势

从上述三种常用相机的分析来看,主动相机的灵活性更强,功能更加完备,因此其实用价值更大。特别是,随着相机制作工艺的完善、机械控制精度的提高,以及生产成本的降低,主动相机的逐渐普及已经成为必然。例如,杭州市在2005年建成了包括10万个摄像机在内的全城监控网络,其中大多数相机是主动相机;武汉市在2013年的平安城市建设中,也安装了大量的主动相机;火箭军部队在信息化建设中,也在阵地、各关键出入口、重点涉密场所安装了热成像仪、主动相机等各种类型的监控设备,从而可以通过统一的信息管理平台监控各个重要区域。可以看出,随着主动相机的广泛应用,智能监控研究的发展趋势之一就是传统静止相机、全向相机到主动相机。