



Y1523979

分类号 TP391.41

学号 04019026

UDC

密级 公 开

工学博士学位论文

无人机平台运动目标检测与跟踪
及其视觉辅助着陆系统研究

博士生姓名 张 恒

学科专业 航空宇航科学与技术

研究方向 基于图像的精密测量及运动测量

指导教师 于起峰 教授

国防科学技术大学研究生院

二〇〇八年九月

摘 要

无人机因其特有优势而一直备受各国军事专家的青睐。特别是在最近几次局部战争中,无人机更因其赫赫战功,而成为各国军事发展的重要方向之一。论文以无人机平台对运动目标检测跟踪和无人机辅助着陆导航系统为背景,研究了图像角点特征提取、像机自运动消除、复杂背景下运动目标跟踪、目标尺度方向自适应跟踪、小目标实时高精度跟踪和无人机辅助着陆系统等方面的相关算法,系统总结了作者在攻读博士学位期间所做的研究和取得的成果。论文创新点如下:

(1) 传统梯度算子抗噪性不好,而且在梯度方向计算准确性上存在一定误差,论文在分析 Gabor 小波的基础上,提出了一种新梯度算子。新算子在抗噪性和梯度方向计算的准确性上优于传统梯度算子。针对图像角点特征提取问题,本文提出一种基于 Gabor 梯度的角点检测算法。新算法在定位精度、噪声抑制等方面比经典算法有一定的提高。

(2) 对于像机自身存在运动情况下对运动目标进行检测跟踪问题,常用的一类算法是先消除像机自身运动,然后再对目标进行检测跟踪。论文针对像机运动消除问题,提出了一种基于 Gabor 特征描述的像机自运动消除方法。实验结果表明,新算法可以有效地消除像机自身运动,检测出和像机存在相对运动的目标。

(3) 目标在复杂背景中运动时,变化的背景会在一定程度上影响跟踪算法的稳定性和可靠性,所以应尽量抑制背景。对于目标跟踪而言,目标中和背景差异大的区域对于跟踪的有效性、稳定性贡献较大。根据以上原则,论文提出了一种目标区域差异性权值计算方法。实验结果表明,权值计算方法有效地抑制了背景,权值计算结果符合人类直观感受。在获取目标差异性权值的基础上,本文提出了一种基于目标差异性加权的最小二乘影像匹配运动目标跟踪算法。仿真实验分析和自然图像跟踪结果表明,对于复杂背景下运动目标进行跟踪,新方法比传统方法更为有效准确。论文将差异性权值与 Mean Shift 跟踪算法结合起来,提出了一种基于目标差异性加权的 Mean Shift 运动目标跟踪算法,并对算法的收敛性及收敛条件进行了讨论。实验结果表明,改进的 Mean Shift 算法比传统跟踪算法在目标跟踪稳定性、可靠性等方面有一定提高。

(4) 对于目标跟踪问题,目标的初始特征建立或目标的描述区域是否准确是整个跟踪过程中至关重要的一环。针对该问题,本文在分析尺度空间理论的基础上,提出了一种目标最佳椭圆描述区域计算方法,该方法为准确描述目标区域特性奠定了基础。针对图像上同时存在旋转缩放变化的运动目标进行自适应跟踪问题,在分析 Mean Shift 跟踪算法和尺度空间理论的基础上,本文提出了一种尺度

方向自适应 Mean Shift 跟踪算法。实验表明,新算法比现有的 Mean Shift 改进算法更为有效,即可以准确获得目标的位置、大小、方向信息。

(5) 针对小目标实时高精度检测跟踪问题,提出一种基于正负正则化 LOG 算子的尺度自适应小目标实时高精度检测跟踪方法。实验证明,新方法在实时性、尺度自适应性、检测准确性以及抗噪声性上均有较好的表现。

(6) 无人机安全进场着陆过程中,其引导控制尤为关键。目前 GPS 导航因其精度高、使用简单而普遍使用,但信号易受干扰,特别是战争时期不能获得准确的 GPS 信息。论文提出了一种基于摄影测量的无人机辅助着陆导航系统。实验表明,该系统可以实时高精度地获得无人机相对跑道的相对位置,同时可以实时保存无人机着陆过程图像,为事后分析提供有效数据。

关键词: 无人机 运动目标跟踪 角点提取 像机自运动消除 小目标检测与跟踪 无人机辅助着陆

ABSTRACT

Unmanned Aerial Vehicle (UAV) is extra favored by the military experts for its own dominance. Specially, UAV has already been one of the very important military development fields for its success in the last brushfires. Moving target tracking on the platform of UAV and the assistant landing system of UAV are taken as the backgrounds of this dissertation. The image corner detection, camera ego motion compensation, moving target tracking in the cluttering background, the scale and rotation adaptive tracking of the target, puny target real-time and high-accurate detection, the assistant landing system and other relative technologies are investigated in this dissertation. The main contents and achievements of this dissertation are as follows:

(1) There is definite error to calculate the gradient orientation by traditional gradient operators. Based on the analysis of the gabor wavelet, a new gradient calculating operator is proposed. The new operator can obtain the orientation of the gradient more accurately, and restrains the noise more effectively than the traditional gradient operators. And a new corner detecting algorithm based on gabor gradient operator is proposed. Experimental results with some synthetic and real images show that this new algorithm detects the corner more efficiently, locates the corner more accurately, and restrains the noise more effectively than the classical algorithms.

(2) For the tracking of the moving target on the moving platform, one commonly used algorithm is to eliminat the camera ego motion firstly, and then track the moving target. An camera ego motion compensation algorithm based on the gabor feature descriptions is proposed. Experimental results show that the new method can effectively compensate the camera ego motion and detect the moving targets.

(3) When a target moves in the cluttering backgrounds, the variational backgrounds can reduce the stability and reliability of the tracking algorithms. So the backgrounds must be restrained. For the heterogeneous target, the target areas in the image are different for the target's ununiformity. If some areas in the target are more different from the backgrounds, it shows that the diversity of these areas is stronger and these areas are more helpful for target tracking. Based on the above-mentioned principle, a diversity weight calculation algorithm is proposed. Experimental results show that the new algorithm can restrain the backgrounds effectively, and that the weights accord with the human intuitionistic recept. Based on the diversity weights, a diversity weighted least square image matching method is proposed. Simulation experimental analysis and real image tracking results show that the new method can track moving target in the cluttering backgrounds more effectively and accurately than the traditional tracking methods. A new diversity weighted Mean Shift tracking method (DWMS) is proposed. Its convergence condition is discussed. And its convergence is proved. Experimental

results show that DWMS can effectively track moving targets in the cluttering background too.

(4) It is one of the most important problem for target tracking whether the initial features and the describing areas of the target are accurate. Based on the scale-space theory, a new method of the optimal ellipse describing area of the target is proposed. The new method provide a foundation for the feature analysis of the target. It is one of the most difficult research fields of computer vision to track target which takes the change of rotation and scales at the same time. Based on the analysis of the scale-space theory , the current Mean Shift algorithms and the optimal ellipse describing area of the target, a scale and rotation adaptive Mean Shift tracking algorithm is proposed. Experimental results show that the new Mean Shift tracking algorithm can more effectively track the target which takes the change of both rotation and scale than the current mean shift algorithms. At the same time, the new algorithm can supply the size and angle information of the target.

(5) To detect and track puny target high-accurately and in real-time, a new scale adaptive little target detecting and tracking method based on the positive and negative LOG operator is proposed. Experimental results demonstrate that the new method has great capability in real time detection, precise detection and noise restraint.

(6) In the landing process of UAV, navigation control is very important. At present, the GPS navigation is usually used because it is very accurate and easy. But its signal can be easily disturbed during the war. The vision-based navigation system becomes one of the most important research fields for its superiority. A assistant landing system based on the photogrammetry is proposed. Experimental results show that the system can obtain the accurate location of UAV according to the runway, and save the images of the whole landing process of UAV synchronously. which can provide effective information for post analysis.

Key Words : Unmanned Aerial Vehicle(UAV) Moving Target Tracking
Camera ego motion compensation Corner Detection Small Target Tracking
And Detection The Assistant Landing System of UAV

表 目 录

| | |
|--------------------------------------|-----|
| 表 1-1 博士学位论文主要研究成果的发表..... | 23 |
| 表 3-1 5 度边缘方向计算结果 (真实值 0.088) | 46 |
| 表 3-2 10 度边缘方向计算结果 (真实值 0.177) | 47 |
| 表 3-3 15 度边缘方向计算结果 (真实值 0.268) | 47 |
| 表 3-4 20 度边缘方向计算结果 (真实值 0.364) | 47 |
| 表 3-5 25 度边缘方向计算结果 (真实值 0.466) | 47 |
| 表 3-6 30 度边缘方向计算结果 (真实值 0.577) | 47 |
| 表 3-7 35 度边缘方向计算结果 (真实值 0.700) | 48 |
| 表 3-8 40 度边缘方向计算结果 (真实值 0.839) | 48 |
| 表 5-1 纯平移变换下十字丝定位结果比较..... | 90 |
| 表 5-2 平移+旋转变换下十字丝定位结果比较..... | 90 |
| 表 5-3 20%高斯噪声仿射变换下十字丝定位结果比较 | 91 |
| 表 5-4 不同噪声下十字丝定位结果比较..... | 92 |
| 表 5-5 四种算法的平均处理时间比较 单位: 毫秒..... | 95 |
| 表 5-6 四种算法平均处理时间比较 单位: 毫秒..... | 103 |
| 表 7-1 像机标定结果验证 (单位: 米) | 126 |

图 目 录

| | |
|-------------------------------------|----|
| 图 1-1 选择的不同特征图案..... | 21 |
| 图 1-2 论文组织结构..... | 22 |
| 图 2-1 图像被尺度渐增的高斯核平滑的结果..... | 26 |
| 图 2-2 不同尺度的正则化 LOG 算子对指定目标点的响应..... | 27 |
| 图 2-3 Gabor 函数模板..... | 30 |
| 图 2-4 三种常用的核函数..... | 34 |
| 图 3-1 图像角点特征分类..... | 40 |
| 图 3-2 Harris 角点检测原理示意图..... | 42 |
| 图 3.3 四个不同位置的圆模板..... | 43 |
| 图 3.4 USAN 区域的显示..... | 44 |
| 图 3.5 Gabor 模板三维示意图..... | 45 |
| 图 3-6 理想边缘图像..... | 46 |
| 图 3-7 5 度理想边缘..... | 46 |
| 图 3-8 10 度理想边缘..... | 47 |
| 图 3-9 15 度理想边缘..... | 47 |
| 图 3-10 20 度理想边缘..... | 47 |
| 图 3-11 25 度理想边缘..... | 47 |
| 图 3-12 30 度理想边缘..... | 47 |
| 图 3-13 35 度理想边缘..... | 48 |
| 图 3-14 40 度理想边缘..... | 48 |
| 图 3-15 虚假点位置示意图..... | 50 |
| 图 3-16 无噪声的检测结果..... | 52 |
| 图 3-17 加入噪声的检测结果..... | 53 |
| 图 3-18 立方体图像检测结果..... | 54 |
| 图 3-19 房子图像检测结果..... | 55 |
| 图 3-20 实验室图像检测结果..... | 56 |
| 图 4-1 三个不同层次的坐标系统..... | 59 |
| 图 4-2 图像与摄像机坐标系..... | 59 |
| 图 4-3 两个像机从不同视角获取图像的模型..... | 63 |
| 图 4-4 对应点在不同视点图像中的变化关系..... | 65 |
| 图 4-5 基于特征匹配的像机自运动消除算法总体框架..... | 66 |
| 图 4-6 CMCGFD 算法流程图..... | 70 |

| | |
|-----------------------------------|-----|
| 图 4-7 两幅图像对应点及估计点示意图..... | 74 |
| 图 4-8 无人机平台图像兴趣点提取匹配结果..... | 75 |
| 图 4-9 无人机平台像机自运动消除结果对比..... | 76 |
| 图 4-10 丛林背景图像像机自运动消除结果对比..... | 77 |
| 图 4-11 高山背景图像像机自运动消除结果对比..... | 78 |
| 图 4-12 岛屿背景图像像机自运动消除结果对比..... | 79 |
| 图 5-1 内外区域选择示意图..... | 81 |
| 图 5-2 分割后内外区域示意图..... | 83 |
| 图 5-3 复杂背景图像中飞机目标差异性权值计算示意图..... | 84 |
| 图 5-4 不同复杂背景图像目标差异性权值计算结果..... | 85 |
| 图 5-5 DWLSIM 算法步骤..... | 88 |
| 图 5-6 平面仿射变换失真..... | 89 |
| 图 5-7 纯平移变换仿真序列图..... | 89 |
| 图 5-8 纯平移变换十字丝定位误差曲线..... | 90 |
| 图 5-9 平移+旋转变换仿真序列图..... | 90 |
| 图 5-10 平移旋转变换十字丝定位误差曲线..... | 91 |
| 图 5-11 20%高斯噪声仿射变换下仿真序列图..... | 91 |
| 图 5-12 无噪声仿射变换下十字丝定位误差曲线..... | 91 |
| 图 5-13 不同噪声仿真序列图..... | 92 |
| 图 5-14 不同噪声下十字丝定位误差曲线..... | 92 |
| 图 5-15 复杂背景下运动飞机跟踪结果比较（一）..... | 93 |
| 图 5-16 复杂背景下运动飞机跟踪结果比较（二）..... | 94 |
| 图 5-17 DWLSIM 跟踪算法结果局部放大..... | 94 |
| 图 5-18 复杂背景下运动飞机跟踪结果比较（一）..... | 102 |
| 图 5-19 复杂背景下运动飞机跟踪结果比较（二）..... | 103 |
| 图 6-1 正则化 LOG 算子计算得到目标内部描述区域..... | 106 |
| 图 6-2 ENLOG 算子示意图..... | 107 |
| 图 6-3 目标区域加权图像计算结果..... | 108 |
| 图 6-4 目标最佳描述椭圆计算结果..... | 110 |
| 图 6-5 目标最佳描述椭圆计算结果..... | 110 |
| 图 6-6 瓶子尺度方向自适应跟踪结果..... | 116 |
| 图 6-7 Mean shift 加减 10%跟踪结果..... | 116 |
| 图 6-8 Collins 算法跟踪结果..... | 117 |
| 图 6-9 本文算法跟踪结果..... | 118 |

| | |
|---|-----|
| 图 6-10 加减 10%Mean Shift 算法跟踪结果..... | 118 |
| 图 6-11 Collins 算法跟踪结果 | 118 |
| 图 7-1 无飞机沿下滑面着陆示意图..... | 121 |
| 图 7-2 双目交会测量无人机三维位置..... | 121 |
| 图 7-3 无人机辅助着陆系统硬件布局示意图..... | 122 |
| 图 7-4 系统工作流程图..... | 122 |
| 图 7-5 双目交会测量示意图..... | 123 |
| 图 7-6 标定布局示意图..... | 125 |
| 图 7-7 标定点布局示意图(单位:米) | 125 |
| 图 7-8 标志点与标杆的相对位置示意图(单位:厘米) | 126 |
| 图 7-9 实拍标定图像..... | 126 |
| 图 7-10 不同尺度的 PNNLOG 算子对同一高斯点的响应..... | 129 |
| 图 7-11 同一尺度的 PNNLOG 算子对不同尺度高斯点的响应..... | 130 |
| 图 7-12 不同尺度 PNNLOG 算子对同一半径圆台的响应..... | 130 |
| 图 7-13 不同灰度高斯点的不同尺度 LOG 算子响应曲线(一) | 131 |
| 图 7-14 不同高斯点不同尺度 LOG 算子响应曲线(二) | 131 |
| 图 7-15 硬盘存图工作流程图..... | 134 |
| 图 7-16 内存存图模式工作示意图..... | 135 |
| 图 7-17 实时处理与存图模式工作示意图..... | 136 |
| 图 7-18 阳光强烈时拍摄的图像..... | 137 |
| 图 7-19 阳光较弱时拍摄的图像..... | 137 |
| 图 7-20 大雾天气拍摄的图像..... | 137 |
| 图 7-21 850 纳米带通滤光片的光谱透过率曲线..... | 138 |
| 图 7-22 加滤光片前拍摄的图像..... | 138 |
| 图 7-23 加滤光片后拍摄的图像..... | 138 |
| 图 7-24 CCD 摄像机和 DPS 摄像机拍摄结果对比 | 140 |
| 图 7-25 云雾影响消除结果对比..... | 140 |
| 图 7-26 PNNLOG 算子对仿真图检测结果..... | 141 |
| 图 7-27 PNNLOG 算子对真实序列图像检测结果..... | 142 |
| 图 7-28 测量数据与 GPS 数据比对图 | 143 |
| 图 11-1 中心透视投影成像模型..... | 162 |

独创性声明

本人声明所呈交的学位论文是我本人在导师指导下进行的研究工作及取得的科研成果。尽我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表和撰写过的研究成果，也不包含为获得国防科学技术大学或其它教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示谢意。

学位论文题目：无人机平台运动目标检测与跟踪及其视觉辅助着陆系统研究

学位论文作者签名：张恒 日期： 年 月 日

学位论文版权使用授权书

本人完全了解国防科学技术大学有关保留、使用学位论文的规定。本人授权国防科学技术大学可以保留并向国家有关部门或机构送交论文的复印件和电子文档，允许论文被查阅和借阅；可以将学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。

(保密学位论文在解密后适用本授权书。)

学位论文题目：无人机平台运动目标检测与跟踪及其视觉辅助着陆系统研究

学位论文作者签名：张恒 日期： 年 月 日

作者指导教师签名：于起峰 日期： 年 月 日

第一章 绪论

无人机(Unmanned Aerial Vehicle, 简称 UAV)作为无人机驾驶的航空飞行器,走上历史舞台已近百年。1917 年英国人研制出第一架无人机,从此拉开了无人机研究的序幕,而对无人机的广泛应用则始于 20 世纪 50 年代。在军事上,无人机最先作为靶机。之后,随着战争的需要和技术的进步,在无人机上安装了光学摄像设备、红外摄像设备、角雷达反射器、通信侦察接收机和干扰发射机等,从而产生了具有情报侦察、雷达诱饵、电子对抗和对目标实施攻击等功能的无人机。无人机真正投入实战始于 20 世纪 60 年代的越南战争。但取得辉煌战绩是在 1991 年海湾战争。当时,多国部队使用了多种无人机,为战区和各部队指挥官收集和提供了大量的电子情报信息。此外,美军还利用无人机作诱饵,诱使伊拉克防空系统雷达开机并发射导弹,致使伊军防空阵地短短两天内全部瘫痪。1995 年波黑战场上,美军空袭波黑塞族前曾派出 80 架次“捕食者”无人机,对战场进行了实地侦察。空袭时,该型无人机向战斗机提供了准确的目标信息,致使塞军 70% 防空设施和作战指挥系统被摧毁,40%的弹药库被炸毁。在阿富汗战争中,无人机开创了用于攻击作战的先河。2001 年 11 月 15 日,捕食者无人机向塔利班停车场上的车辆发射了两枚海尔法导弹,致使基地组织的二号人物穆罕默德·阿提夫在这次空袭中丧生。2002 年 11 月 3 日,“捕食者”无人机又极其成功地在也门沙漠地带击毁乘有六名“基地”分子的车辆,为美军立下赫赫战功。在民用上,无人机主要用于大地测量、气象观测、城市环境检测、地球资源勘探和森林防火等。

无人机所特有的功能和优越性,注定了从诞生的那一天起,就被各国专家所青睐。随着在战争中的应用,其性能和功能因作战需要而不断提高和扩展。可以预言,无人机将在 21 世纪的战争或局部冲突中显示出更加强大的威力,也是未来实现“零伤亡”战争的首选武器之一。

本文以无人机对运动目标进行跟踪锁定及无人机安全着陆导航为研究背景,围绕运动目标检测与跟踪以及基于视觉无人机辅助着陆系统的研究开展工作。本章将首先介绍论文的研究背景和意义,然后综述相关领域的国内外研究进展,最后介绍论文的内容安排。

1.1 论文研究背景

1.1.1 运动目标检测跟踪技术的研究目的和意义

计算机视觉是研究利用计算机系统实现类似人类视觉系统理解客观世界的新兴学科。它是目前人工智能和生命科学研究的热点领域之一。自上世纪七十年代

后期,以Marr等为代表的研究者提出了一整套视觉计算理论^[1]开始,主动视觉^[2]、基于模型的视觉^[3]以及多传感器技术^[4]的相继兴起和研究,使得Marr理论框架下的病态问题变成了非病态,以前不能解决的问题也有可能得到解决,较大地推动了计算机视觉理论的发展。但是到目前为止,计算机视觉理论仍然存在许多难题有待解决。这些困难严重阻碍着计算机视觉技术从理论研究迈向实用化的进程,作为计算机视觉理论基础之一的“序列图像中运动目标检测与跟踪”就是其中一个富有挑战性的课题。

运动目标检测与跟踪的研究实际上是一个系统工程问题,它跨越了计算机科学、光学、数学以及认知科学等。虽然二十世纪六十年代初就得到了学术界和应用部门的高度重视,但是直到最近几十年随着对人类视觉系统认识的不断深入以及各种复杂算法、微处理器技术、超大规模集成电路技术的高速发展,才使得对运动目标检测和跟踪的研究从理论探索走向实际应用。同时随着现代工业和军事自动化程度的不断提高,基于该技术的各种工程需求不断涌现。

在民用上,运动目标检测与跟踪技术的应用相当广泛,其中最为重要的应用之一是银行、停车场和海岸等对安全要求敏感场合的安全监控。美国自从“911”恐怖袭击之后,在全国范围内安装了视频监测系统,以达到预先报警及时处理的目的。特别当某处出现恐怖袭击时,可以从无人机和地面等多个视频平台对恐怖袭击者和车辆进行实时识别跟踪,为采取进一步措施提供了有利保障。在交通系统中,目标检测与跟踪也具有非常广泛的应用,主要包括交通流量控制、异常行为检测、行人行为判定、智能车辆测量等很多方面。此外该技术在人机交互、人体运动分析、视频压缩等方面发挥至关重要的作用。因此运动目标检测与跟踪技术具有较高的民用价值。

与民用领域相比,军事领域对于视频监视系统的要求要高得多,主要是因为战场环境远比普通的民用环境复杂和苛刻。现代高科技条件下的战争,大量对抗、伪装、欺骗和反辐射技术的应用,使得作战环境日趋复杂,对武器的制导系统提出了更高的要求,不仅要求能够精确击中目标,而且要求进一步击中目标的关键部位。而准确检测、识别和跟踪目标是精确制导的重要环节之一。在各发达国家未来军事技术发展规划中,目标检测与跟踪的地位十分引人注目。美国、法国等已经开始研究在武器上安装高分辨率电视摄像机、微光电视和红外摄像机,以提高检测和跟踪目标的概率。美国在1998年科索沃战争和2003年伊拉克战争中所使用的高技术武器的命中精度之高,尤其是光电夜视和制导技术的应用给人们留下了深刻的印象。因此运动目标检测与跟踪技术具有极其重要的战略意义。

从国内外的研究动态来看,运动目标检测和跟踪研究一直受到广泛的关注。国际上许多著名的期刊和会议都把它作为重要研究领域。国外许多著名高校和研

究所,如Carnegie Mellon University、University of Oxford、MIT和Mitsubishi Electric Research Laboratories等,都专门设立了针对目标检测和跟踪的研究组合和研究实验室。对于运动目标检测和跟踪的许多关键技术,如目标检测、运动分析、像机自运动消除、目标分类与理解以及图像信息融合等,许多学者进行了大量的研究。基于图像分割技术^[5]、基于模板匹配^[6,7]、基于光流场^[8]等目标检测方法已经提出;基于灰度变化^[9]、基于特征分析^[10]、基于变换^[11]等运动分析方法也已日趋成熟;基于马尔可夫随机场的图像融合、加权融合和主成份分析、基于统计的图像融合等图像融合算法,以及大量的目标分类和理解的方法近年来也不断见诸文献。同时也涌现出了一批已进入实用的检测和跟踪系统,如Collins介绍的由CMU和Sarnoff公司合作研究的一种视觉监测系统,该系统使用多个相互协作的摄像机在复杂环境里对人和车辆进行连续跟踪,并对目标类型和行为进行分析。Olson等研究的运动物体检测和事件识别系统,通过检测帧间图像的变化来发现运动物体,在跟踪技术上使用了一阶预测和最近邻匹配技术。但是,当前对运动目标检测和跟踪问题的研究尚处于特定问题设计特定方法的阶段,对场景复杂多变工作环境的自适应能力很差。常常是当背景较简单或者在某种特定的工作场合,设计的方法或系统可以较好的工作。但当背景复杂时,特别是拍摄目标的像机自身存在运动时,该问题常常变得更为困难。这主要是因为:

(1) 现有的检测和跟踪算法或系统都是在一定的假设条件下建立的,它无法适应客观世界的复杂多样性;

(2) 大多数方法只使用目标的二维图像信息,因而无法对目标进行完整的描述,信息损失对目标的检测和识别常常有较大影响,虽然采用多目像机可以在一定程度上弥补这些损失的信息,但是如何很好地融合这些信息尚缺乏有效的方法;

(3) 沉重的计算负担严重制约了该技术的发展。当背景复杂时,目标检测、运动分析、目标分类、目标跟踪变得更加困难,同时对算法的准确性和快速性提出了更高的要求。迄今为止还没有有效的方法可以解决这些问题;

(4) 照明情况、目标表面的物理性质、场景中目标的三维空间关系等因素的综合影响,仅仅反映为图像的灰度信息变化,从而造成对客观世界的多解性,如何综合考虑这些信息常常很困难。

通过以上的分析,可以看出对运动目标检测与跟踪技术进行深入研究,不仅具有深远的理论意义,还具有广泛的实用意义。

1.1.2 无人机辅助着陆技术的研究目的和意义

如何实现无人机安全平稳着陆是目前无人机技术发展迫切需要解决的问题。着陆时,由于操纵复杂、地面干扰因素多,而事故频繁。要解决这个问题,必须

高精度地获取在着陆过程中无人机的方位和姿态信息。

目前常用的无人机着陆导航技术主要包括：惯性导航定位系统(INS)、GPS导航定位、INS/GPS组合导航定位系统和视觉导航定位系统。其中，惯性导航定位是最早研究、最为成熟的导航技术；GPS是近十年兴起的，应用最为广泛，技术也相对成熟的导航技术；伴随着多种导航技术的出现，人们自然地想到将不同的导航技术组合，发挥各自的优势以达到更好的导航效果，INS/GPS组合导航是一种组合较为理想、研究较多的导航技术。但是上述几种导航技术都各有它们的缺点：惯性导航系统是利用加速度计等惯性元件感受运载器在运动过程中的加速度，然后通过积分计算，得到载体的位置与速度等导航参数，它的最大缺点是误差会随着时间的推移而发散；GPS导航是利用导航卫星进行导航定位，它具有精度高、使用简单等优点。在和平时期GPS导航是一种很好的自主导航方法，但是在战争时期，该种导航方式容易受到干扰，因此它只能算是一种半自主导航定位技术。近年来，基于视觉的导航系统因其特有的优势而成为导航系统研究的热点之一，而且已经取得了一些研究成果。传统视觉导航系统的策略是：首先利用无人机实时获取的地物景象与预先得到的基准景象进行匹配；然后在成功匹配的基础上，计算实时景象窗口中心所对应的基准景象的窗口中心坐标，并将该坐标作为无人机的空间平面坐标值，从而实现无人机的空间定位。该系统具有非接触性、精度高、不易受干扰等优点。但是该类系统也存在着一定的问题，如当无人机飞行在两千米高度成像时，姿态角的微小变化和地形的起伏影响都会使成像底点产生较大的偏移；算法较为复杂，实时性难以保证等等。因此，研究有效的辅助着陆系统具有巨大的应用价值。

1.2 国内外相关研究进展

1.2.1 基于视觉的目标跟踪算法研究进展

随着计算机技术以及计算机视觉算法的发展，视觉跟踪问题逐渐成为研究热点之一。二十世纪八十年代以前，由于计算机技术的限制，对图像的处理与分析主要以静态图像为主^[12]。二十世纪八十年代初光流法^[13]提出之后，动态序列图像分析进入了研究高潮。但是由于采用假设的局限性使得光流法对噪声特别敏感，很容易产生错误的结果。从二十世纪八十年代后，国内外学者从不同角度提出了众多的视觉跟踪算法。为了了解跟踪算法的研究进展，首先对视觉跟踪问题进行分类，然后介绍处理视觉跟踪问题的两种思路，最后对具体的视觉跟踪方法进行分类介绍。

1.2.1.1 视觉跟踪问题分类

按不同标准可将视觉跟踪问题分成很多类型，下边给出一些常见的分类：

(1) 可见光图像和红外图像

根据所使用的传感器种类的不同，视觉跟踪问题可分为可见光图像的视觉跟踪和红外图像的视觉跟踪。通常的自然环境条件下，白天常使用可见光图像进行目标跟踪，而晚上则使用红外图像进行目标跟踪。如果将二者所获得的图像进行融合处理，则一般可以得到比单一传感器信息更多的图像，从而极大地提高检测和跟踪运动目标的能力。

(2) 单摄像机和多摄像机

根据所使用摄像机数目的多少，视觉跟踪问题可分为单摄像机视觉跟踪和多摄像机视觉跟踪。目前，绝大多数研究都属于单摄像机的视觉跟踪问题。但是，在实际的视频监视系统中，单摄像机所能监视的区域很有限。因此，在一些大的公共场合，更多的是使用多摄像机组合以扩大监视的范围。此外，单摄像机不能获得运动目标的距离信息，而多摄像机系统却能够得到运动目标的距离信息。但是多摄像机跟踪面临一个单摄像机跟踪所没有的难题，就是如何将不同摄像机所获得的不同视角的运动目标信息正确地对应起来。

(3) 摄像机静止和摄像机运动

根据摄像机是否运动可以将视觉跟踪问题分为摄像机静止时的运动目标跟踪和摄像机运动时的运动目标跟踪。对于大多数的视频监视系统而言，都是在摄像机静止时对某个需要特殊关注的区域进行监视。背景是静止的，而作为前景的运动目标是移动的，这种情况下使用背景差法^[14]能够取得相当不错的结果。但是在很多情况下，摄像机是运动的，其运动的形式可以分为两种：一种是摄像机的支架固定，但摄像机可以偏转，俯仰以及缩放；另一种是摄像机装在某个移动的载体上的，如移动的车辆或飞行的飞机上。在这些情况下，由于背景和前景都是运动的，要准确检测并跟踪运动目标变得更为困难。

(4) 单运动目标和多运动目标

根据视频场景中运动目标数目的多少，视觉跟踪问题可分为单目标跟踪和多目标跟踪两类。单目标跟踪并不是一个简单的问题，由于场景中噪声的影响，往往不能准确地检测出真正的运动目标，即使可以检测出运动目标，也很难将其完整地提取出来，例如由于阴影或光照的影响等造成提取出来的运动目标不完整或中心有空洞。而多目标跟踪比单目标跟踪要困难得多，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。

(5) 刚体和非刚体。

根据视频场景中运动目标类型的不同，视觉跟踪问题可分为对刚体目标跟踪

和对非刚体目标跟踪。所谓刚体是指具有刚性结构的物体，例如车辆、坦克、飞机等，这些物体的共同特点是结构比较规范，能够用三维几何模型进行描述，因此，对这类运动目标的跟踪采用基于三维模型的跟踪方法。而非刚体是指外形能够变化的物体，如细胞、动物、人等，对这类目标通常采用变形模板进行跟踪。

1.2.1.2 处理目标跟踪问题的两种思路

对于目标跟踪问题的处理，总体上讲有两种思路，一种称之为自底向上(Bottom-up)的处理方法；一种称之为自顶向下(Top-down)的处理方法。自底向上的处理方法又称为数据驱动(Data-driven)的方法，这种方法不依赖于先验知识，直接从序列图像中获得目标的运动信息并进行跟踪。自顶向下的处理方法又称为模型驱动(Model-driven)的方法，这种方法依赖于所构建的模型或先验知识，在序列图像中进行匹配运算或求解后验概率，实现目标跟踪。

(1) 自底向上

自底向上实现目标跟踪的思路中以Marr的视觉计算理论^[1]最具代表性。Marr认为视觉过程的主要任务是从二维图像中定量地恢复图像所反映的场景中三维物体的形状和空间位置。为此，Marr将视觉过程分为三个阶段：早期、中期和后期。早期视觉由输入图像获得要素图，包括图像中强度变化剧烈处的位置、几何分布和组织结构；中期视觉实现由输入图像和要素图获得的2.5维图，即获得在以观察者为中心的坐标系里可见表面的深度及不连续轮廓等；后期视觉由输入图像、要素图以及2.5维图获得物体的三维表示。

在自底向上的目标跟踪过程中，跟踪的目的是获得场景中运动目标的位置、速度、加速度、运动轨迹等信息。为得到这些信息，在获得序列图像进行预处理之后，直接从序列图像中检测运动目标，这是早期阶段；检测到运动目标之后，提取识别运动目标，以判定是否开始跟踪，这相当于中期阶段；最后对该目标进行跟踪并获得运动目标的相关运动信息。上述过程按处理顺序共分为四步，第一步为图像预处理，即对获得的序列图像进行消噪或增强，提高图像质量以方便后续处理；第二步为目标检测。在摄像机不动情况下一般采用帧差法^[15]或背景差法^[14]进行变化检测以获取运动目标，此时得到的运动目标是一个团状斑点；当摄像机运动时，需要先消除像机自运动，然后再进一步检测出运动目标。第三步为目标分类。要完成两个任务，首先提取出检测到的运动目标，其次对该目标进行识别；最后在目标跟踪阶段获得运动目标的相关运动信息。这种方法的优点在于能快速检测并跟踪运动目标，便于工程实现，目前已有大量实际系统利用该思路解决目标跟踪问题^[16,17,18]。该方法在摄像机静止情况下效果非常好。如果摄像机是运动的，则实现运动目标检测与跟踪就比较困难，因此在很大程度上限制了该类方法的使用范围。

(2) 自顶向下

自顶向下的跟踪思路中以Bar-Shalom的目标跟踪思想最具代表性[19]。该类算法将目标跟踪问题转换为在贝叶斯(Bayesian)理论框架下, 已知目标状态的先验概率, 在获得新的测量后不断求解目标状态的最大后验概率的过程。也就是说, 在贝叶斯理论框架下, 将视觉跟踪问题看作是“最优猜测”或者“推理”的过程, 通常采用状态空间法实现视觉跟踪。当系统噪声为高斯分布时, 可以利用卡尔曼滤波(Kalman filter)^[19]求解后验概率。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(Hidden Markov Models)^[20]进行跟踪。但是, 经典卡尔曼滤波只能处理线性、高斯、单模态的情况。实际的视觉跟踪过程中, 后验概率分布往往是非线性、非高斯、多模态的。为此, 在1993年左右一种称之为序贯蒙特卡罗(Sequential monte carlo method)的方法^[21,22]被提出用于解决这类情况。该方法在刚提出之时并未受到很大关注, 直到1996年, 由Isard和Blake提出Condensation算法^[23]之后才引起广泛注意。目前这一类方法被称为粒子滤波(Particle filter, 简称PF), 是当前视觉跟踪的研究热点之一。

自顶向下思路利用先验知识对跟踪问题建立模型, 然后利用实际序列图像验证模型的正确性, 这种方法具有坚实的数学理论基础, 有很多数学工具可以使用, 因此一直是理论界研究视觉跟踪问题的主流方法。但是, 先验知识的表述是人工智能中的困难所在, 而且人的很多先验知识也很难用数学形式来表达。此外, 通过模型匹配来实现跟踪往往比较费时, 很难在实际工程中实现实时跟踪。

自底向上跟踪和自顶向下跟踪有着各自的优点和缺点, 若将自底向上和自顶向下这两种思路结合起来使用, 就有可能克服各自弱点, 实现更为有效的跟踪。

1.2.1.3 视觉跟踪算法

从二十世纪八十年代起到目前为止, 出现了众多的视觉跟踪算法。通常, 将视觉跟踪方法分为三类, 分别是基于区域的跟踪算法、基于特征的跟踪算法以及基于模型的跟踪算法。

(1) 基于区域的跟踪算法

基于区域的跟踪算法基本思想是: 首先通过图像分割获得或是预先人为确定区域获得包含目标的模板。然后在序列图像中, 运用相关算法跟踪目标。最常用的相关准则是平方和准则。这种算法的优点在于当目标未被遮挡时, 跟踪精度高。但其缺点首先是费时, 当搜索区域较大时情况尤其严重; 其次, 算法要求目标变形不大, 且不能有较大遮挡, 否则相关精度下降会造成目标跟踪失败。有关基于区域的跟踪算法可参见文献[24,25]。近年来, 对基于区域的跟踪方法关注较多的是如何处理模板变化时的情况, 这种变化是由运动目标姿态变化引起的, 如果能正确预测目标的姿态变化, 则可实现稳定跟踪, 研究这一问题的有关文献可参见[26,27]。

(2) 基于特征的跟踪算法

与基于区域的跟踪算法的不同之处在于, 基于区域的跟踪算法使用目标整体作为相关时的对象, 而基于特征的跟踪算法使用目标的某个或某些局部特征作为相关对象。这种算法的优点在于即使目标的某一部分被遮挡, 只要还有一部分特征可以看到, 就可以完成跟踪任务。这种算法的难点是: 对运动目标, 如何确定目标的特征集。若采用特征过多, 算法计算量将增加, 而采用特征较少, 又容易因描述不够准确唯一而产生错误。Kanoke^[28]和 Mitra^[29]分别对这个问题进行了讨论。通常, 该类算法提取目标的颜色(或灰度)、角点、边缘等信息建立特征描述。如 Koller 采用 Canny 算子获得目标的边缘信息作为特征^[30], 而 Smith 采用 SUSAN 算子获得目标的角点信息作为特征^[31]。在基于特征的跟踪方法中, Mean Shift 算法因其实时性、有效性渐渐成为研究热点之一。Comaniciu^[32]对 Mean Shift 在跟踪中的使用做了论述。Yang Changjiang^[33]对多维图像的 Mean Shift 方法进行了讨论, 使用改进的快速高斯变换提高算法的速度。Collins^[34]将尺度空间理论和 Mean Shift 方法相结合解决了核函数带宽变化时的目标跟踪问题。另外, Nummiaro^[35]研究了粒子滤波器和 Mean Shift 方法相结合的情况, 但是粒子滤波器本身的复杂计算降低了跟踪的实时性。

(3) 基于模型的跟踪算法

根据跟踪目标特性不同, 该类方法又分为刚体目标跟踪和非刚体目标跟踪。

对于刚体目标, 如汽车、飞机和坦克等, 更多的是采用运动目标的三维模型进行跟踪。其基本思想是: 首先由先验知识获得目标的三维结构模型和运动模型, 然后根据实际的序列图像, 确定出目标的三维模型参数, 从而获得目标的瞬时运动参数。Juric^[36]、Drummond^[37]、Gerard^[38]和 Jung^[39]等提出了适应刚体目标的基于模型的跟踪方法。该类方法的优点是, 可以精确地分析目标的三维运动轨迹, 即使在运动目标姿态变化的情况下, 也能够可靠地跟踪。但是其缺点在于, 运动分析的精度取决于几何模型的精度, 在现实生活中要获得所有运动目标的精确几何模型非常困难。这限制了基于模型的跟踪算法的使用, 同时基于 3D 模型的跟踪算法往往需要大量的运算时间, 很难实现实时运动目标跟踪。

对于非刚体目标, 一般采用主动轮廓模型(又称 Snake 模型)进行跟踪。该模型是由 Kass 等人于 1987 年首先提出^[40]。不同于 Marr 提出的分层视觉理论, 他们认为在众多图像理解任务中, 底层事件的正确理解依赖于高层知识。因此, 试图设计一种能量函数: 其局部极值组成了可供高层视觉处理选择的方案, 通过能量项的迭加完成从该组方案中选择最优的一组。这样, 在寻找显著的图像特征时, 高层机制可能通过将图像特征推向一个适当的局部极值点而与模型进行交互。Snake 模型一经提出, 就受到国内外学者的广泛关注。在最近十多年中, 已经被越

来越多研究者成功地应用于计算机视觉的许多领域。该模型具有一些经典方法所无法比拟的优点,如将图像数据、初始估计、目标轮廓及基于知识的约束统一于一个特征提取过程中;经适当地初始化后,它能够自主地收敛于能量极小值状态;尺度空间中由粗到精地极小化能量可以极大地扩展捕获区域和降低计算复杂性。同时,Snake模型也有自身的缺点:对初始外置敏感,需要依赖其他机制将Snake放置在感兴趣的图像特征附近;由于Snake模型的非凸性,它可能收敛到局部极值点,甚至发散。针对Snake模型的缺点,国内外许多研究者提出了许多改进算法。针对在没有图像力的情况下,Kass等人的Snake模型将收缩为一点或一条直线的问题,Cohen等人提出了主动轮廓线的“气球模型”^[41],通过在外力中增加了膨胀力,改善了Snake对初始轮廓的敏感性,并且能跨越图像中的伪边缘点。对于Snake模型无法收敛到轮廓的深度凹陷部分,一些研究者提出了局部自适应法扩大搜索区来解决这一问题,但效果并不理想。Xu等人提出了GVF Snake (gradient vector flow Snake)模型试图解决这一问题^[42]。该模型提供了一种自然的机制用以扩大Snake的捕获区,可以将Snake拖向物体的深度凹陷区。Szloliski等人将Snake的动态模型作为Kalman滤波器的系统模型,同时考虑了系统噪声和观测噪声模型,提出了Kalman Snake跟踪模型^[43]。由于该模型受到弹性动力学模型的制约,因此可以有效地跟踪刚体和非刚体的复杂运动,同时提供被跟踪物体的位置、速度、加速度和轮廓形状等信息。Vieren等提出了一种基于Snake模型的多目标跟踪算法^[44]。但Snake模型比较适合单目标的跟踪,对于多目标的跟踪更多地是采用基于水平集(Level Set)方法的主动轮廓模型^[45,46]。

1.2.2 图像角点检测算法研究进展

图像角点定义为二维图像亮度变化剧烈的斑点或图像边缘曲线上曲率极大值点。角点特征在三维场景重建、运动估计、目标跟踪、目标识别、图像配准与匹配等计算机视觉领域起着非常重要的作用。这些特征点保留了图像重要特征,有利于图像的可靠匹配,同时有效地减少数据量,提高计算速度,使得实时处理成为可能。

目前的角点检测算法多种多样。Kchen和Rosedfeld^[47]于1982年提出了一种基于梯度的角点检测方法,该方法给出了具体的角点检测算子 K ,通过检测 K 的响应值在图像某一领域的极大值来达到提取角点的目的。该算子表现为水平面截线上某点 (x,y) 的曲率与该点的最大梯度的乘积。但田原和梁德群等人^[48]指出 K 在最大梯度方向上并不是极大值点,而是呈现单调变化的,所以在某一个邻域内曲率和该点的最大梯度乘积的极大值并不会出现在角点上,同时指出角点不仅是水平面截线上的曲率极值点,也是该点在最大梯度方向上最大梯度的模达到极大值,是

满足两个条件的点集的交集。考虑到角点作为一种重要的信号特征，属于图像的细节，按照Witkin尺度空间理论^[49]，角点应该在较大的尺度空间存在。基于小波多尺度分析的角点检测^[48]，通过提出不同尺度上角点的对应关系准则由大尺度跟踪到小尺度上精确的角点位置。Rosedfeld^[50]于1999年提出了一种基于模板的角点检测方法。该方法通过模板窗口局部梯度幅值和梯度方向的变化率来计算角点度量值 C ，根据 C 与给定的阈值大小关系来判定该点是否是角点。Harris等人^[51]提出用高斯窗在图像上移动，由模板窗口取得原图像衍生 2×2 的局部结构矩阵 M ，根据矩阵 M 的两个特征值建立一个角点判断算子 R ，由 R 判断该点是否为角点。Tomasi和Kandc^[52]提出的KLT算法也是基于计算窗口模板 $d \times d$ 大小图像的局部结构矩阵，通过矩阵的特征值判定该点是否为角点。Smith和Brady^[53]提出了一种SUSAN角点检测算法。SUSAN算法是基于半径为 r 的圆形模板。该算法对每个像素基于其模板邻域的图像灰度计算角点响应函数，如果该响应值为局部极大值，则认为该点是角点。

由算法的实现和相关结果可以看出，KLT算法比Harris算法检测角点的质量高，但KLT算法适用于角点数目不多且光源简单的情况，Harris适用于角点数目较多且光源复杂的情况。除了对单幅图像能进行角点检测以外，KLT算法比Harris算法对序列图像的角点检测效果更好。但Harris算法的实现公式中有平滑部分，因此具有较强的鲁棒性且对噪声也不太敏感。Kitchen & Rosenfeld 算法和SUSAN算法一般来说不适合序列图像的角点检测，对于单幅图像的角点检测，SUSAN算法要比Kitchen & Rosenfeld算法好。但在实际计算过程中，圆形模板需要离散化，因此带来了较大的量化误差，容易导致边缘点和角点的判断混乱。对于边缘模糊的图像，使用小模板会丢失角点，这就需要动态地判断究竟用哪种模板最优。文献[54]针对此问题提出模糊度的概念，在计算每一个像素的角点响应函数值之前首先测定其模糊度。若达到模糊的标准，就使用大的模板来计算；若清晰，选用小的模板来计算。这使得判定的准确性得到很大的提高。费旭东等人^[55]采用基于知识的查表技术来进行角点的快速提取，其特点是便于用硬件来实现，但必须先得到图像的边界链码表示。文献[56]基于神经细胞感受野数学模型提出双高斯差(Difference Of Gaussian,简称DOG)模型来检测角点，指出高斯二阶微分函数是DOG函数在其两个高斯函数相互逼近时的一个极端形式特例。DOG函数与信号的卷积相当于两个高斯函数与信号的卷积结果之差，因此抗噪声的能力较强。

1.2.3 运动目标检测研究进展

运动目标的检测是计算机视觉中富有挑战性的课题之一，它被广泛应用于靶场测量、目标监视、视频压缩、车辆跟踪和航空航天等方面。许多研究者在此领

域做了大量开创性的工作,取得了较大的进展,不断有高水平、有实用价值的研究成果出现。根据各个算法侧重点不同可以有很多种分类方法。例如,根据运动目标检测需要的图像帧数分为:单帧检测和多帧检测。根据检测利用的信息不同分为:基于颜色(或灰度)差异和基于运动差异两种方法。根据摄像机是否运动可将运动划分为两种模式:摄像机静止物体运动和摄像机运动物体运动,每一种模式需要不同的分析方式和算法。摄像机静止物体运动是一类非常重要的场景分析,包括运动目标检测、目标运动特性估计,主要用于预警、监视和目标跟踪等。摄像机和物体都运动是计算机视觉中最一般的情况,处理起来最为复杂,目前对该问题的研究非常多,是计算机视觉的热点问题之一。实现摄像机运动下运动目标检测的关键在于消除像机自身运动、完整地分割目标、合理地提取特征和准确地识别目标。

由于运动目标检测在军事、国防以及工业等领域具有广泛的应用前景,从而激发了国内外广大科研工作者的浓厚兴趣。近些年,国内外提出了许多有实际意义的算法。下面针对不同算法进行简要综述。

1.2.3.1 差图像法

该类方法是在运动目标检测中使用得最多的一类算法^[57,58,59,60]。它的突出特点是实现简单、运算速度快,检测效果较好。基本原理是将前后两帧图像对应像素点的灰度值相减,在环境亮度变化不大的情况下,如果对应像素灰度相差很小,可以认为此处景物是静止的,如果图像区域某处的灰度变化很大,可以认为这是由于图像中运动物体引起的,将这些区域标记下来,利用这些标记的像素区域,求出运动目标在图像中的位置。由于目标大小、背景亮度的差别,对差图像的分割方法也不尽相同。另外当目标有阴影干扰时也需要进行特殊处理。Rosin 对此进行了详细的讨论^[61]。为了提高差图像法的性能,Jain^[62]在此基础上提出了累积图像差分法,陈朝阳^[63]提出了对称图像差分法,这两种方法都可以在提高检测概率的同时降低虚报概率。Lipton^[64]和 Changick^[65]提出了一种相邻帧差方法,该类方法主要利用时间信息,通过比较序列图像中连续两、三帧中所有的对应位置像素点,并根据规则计算对应像素点的差值,如果差值大于预先选取的阈值,则认为该点对应位置有运动目标存在。

早期的差图像法是以图像背景固定不变为前提的,如果图像背景运动,早期的方法就无能为力了。针对这种情况,杨志华等人提出了基于补偿差分的运动目标检测方法^[66],先对由于摄像机运动引起的背景移动进行补偿再进行差分运算。Araki 等人将两帧图像之间的背景运动用仿射变换来表示,并用最小中值法来求解变换参数,取得了比较理想的补偿效果^[67]。Collins 提出的自适应“背景减”与“相邻帧差”混合算法^[68],能够快速有效地从背景中检测出目标。Paul Viola 等人^[69]

采用了“相邻帧差法”与“Adaboost方法”，实现对步行者的检测。

1.2.3.2 光流法

光流计算技术是 Gibson 于 1950 年提出的。所谓光流是指图像中模式运动的速度，它是一种二维速度场。光流法检测运动目标的基本原理是：给图像中的每一个像素点赋予一个速度矢量，形成一个图像运动场，在特定时刻，图像上的点和三维物体上的点一一对应，这种对应关系可由投影关系得到，根据各个像素点的速度矢量特征，对图像进行动态分析。如果图像中没有运动目标，则光流矢量在整个图像区域是连续变化的，当物体和图像背景存在相对运动时，运动物体所形成的速度矢量必然和邻域背景速度矢量不同，从而检测出运动物体的位置。这种方法优点在于能够检测独立运动目标，不需要预先知道场景的任何信息，并且可应用于摄像机运动的情况。但是由于噪声、多光源、阴影和遮挡等原因，使得光流场计算不是十分可靠和精确，而且多数光流法计算量比较大。Eun^[70]和 Alexei^[71]提出了应用运动目标随时间变化的光流矢量特性在序列图像中检测出运动区域。日本大阪大学的 Okada, Yamamoto 等人对此方法作了深入的研究，并已研制出比较成熟的系统^[72,73,74]。借助于多个数字信号处理器实现了实时目标检测与跟踪，处理速度可以达到 15 帧/秒。

1.2.3.3 基于特征匹配的方法

基于特征匹配的算法首先提取图像某种特征，然后利用特征集合之间的空间关系或属性关系等进行匹配，接着利用匹配好的特征集合鲁棒地估计出两幅图像之间的变换模型参数，最后根据变换模型将两幅图像配准后相减得到运动目标。常用的特征有：点特征，例如角点、线段端点、区域形心等；线特征，例如线段、圆或椭圆曲线等；区域特征，例如区域面积、主轴方向和矩不变量等。这类算法具有受几何变形和灰度变化影响小、可匹配特征类型多等优点，尤其适于图像人工建筑物的匹配。这类方法中，最为典型的是 Lowe 提出的 SIFT 匹配算法^[75]。该方法因其有效性、鲁棒性而受到国内外学者的广泛关注。

1.2.3.4 基于统计模型的方法

基于统计模型的方法，首先利用较简单的算法对运动场进行粗略估计，然后根据马尔可夫随机场理论构造出运动场的间断点分布模型，利用此模型来检测运动场间的断点，从而实现运动目标的检测。实验表明，该类算法比较有效，且运算负担较小，并可以通过硬件实现实时检测^[76]。Rosenberg 提出了另外一种基于统计模型的方法^[77]。该方法在摄像机做剧烈拉伸、旋转运动时仍可以实时检测并跟踪运动目标。Grimson 等人利用跟踪得到的数据不断校正分布式传感器，重新构建场景模型，通过学习对典型的运动模式进行统计描述分类检测运动目标^[78]。Ahmed

等人提出了一种非参数背景模型的实时运动目标检测算法^[79]，它利用统计学习的模式进行背景更新，提取前景目标。Schneiderman 则采用统计模型实现人脸与小汽车的三维目标检测^[80]。

1.2.3.5 基于高阶统计的方法

高阶统计包括高阶累积矩和高阶矩，它具有从高斯信号背景下提取出非高斯信号的独特优点，是研究非高斯过程、非最小相移系统和非线性系统的有力工具^[81]。利用高阶统计解决问题的关键性论文已在 60 年代发表，但直到 70 年代，人们才开始应用高阶统计技术来解决信号处理当中的实际问题。

实际上，对于视频数字图像，区分运动区域与背景是在一个估计场中进行的，由于运动目标的出现，相邻的两帧图像的帧差相当于一个零均值、对称的概率密度函数，其概率分布与高斯信号函数相差较大。在帧差中，由噪声、亮度变化引起的随机变动、慢变的背景纹理都相当于高斯信号，可以利用图像信号的四阶累积或者四阶矩来提取运动区域。要初步地将运动区域提取出来，可先计算局部估计的帧差零延迟四阶矩，自适应地设定与背景变化相适应的阈值，将计算得到的四阶矩和阈值相比较，然后确定运动区域和背景，从而分离出有别于背景波动和噪声的运动区域。

上述处理中如果阈值设定不当，可能导致部分噪声不能被消除；或者由于物体内部纹理信息的不一致，可能导致检测出的目标内部出现空洞现象等问题。所以还要进行边缘提取、区域填充等工作，最终实现运动目标检测。

1.2.3.6 基于小波的方法

小波理论是 20 世纪 80 年代后期发展起来的一种新的信号处理工具，具有良好的时频局部化特征和方向性特征，在运动目标检测和跟踪技术中得到了广泛的应用。这些应用大多是使用传统的 Mallat 小波算法^[82]，但是 Mallat 快速算法是以傅立叶变换为基础，直接在时空域中设计滤波器比较困难，并且计算量较大。为克服这些局限，国内外学者进行深入地研究，提出了许多有效的算法。于等人提出了自适应最大似然比检测算法^[83]，用于检测多光谱图像中的弱目标。通过小波分解，在多尺度上计算由光谱、方向和尺度参数构成的向量，进行目标检测。Casasent 等人将反映方向与空间频率的小波与 Gabor 系数组合在一个框架下^[84]，用于目标的检测。李红艳在分析了小波变换特性的基础上，提出了基于 Haar 小波变换的低信噪比小目标检测算法^[85]。理论分析与仿真结果表明，该方法可有效地提高小目标的信噪比。

1.2.4 小目标检测跟踪研究进展

图像处理中对低信噪比小目标的检测与跟踪问题的研究，起源于远距离搜索与监视，如用 CCD 宽视场望远镜搜索或跟踪天空中的流星、卫星或其它运动目标^[86,87,88,89]，用机载或地面红外（电视）搜索跟踪系统搜索远距离目标等^[90,91]。在这些应用中，由于目标距传感器较远，它们在传感器中的成像为小目标，相对于背景和传感器噪声而言，目标强度较弱，基本上被噪声所淹没。小目标的定义没有固定统一的标准，在不同的历史时期，不同的应用背景中都有着一些差异。目前常用的描述包括“弱”和“小”两个方面，它们分别描述目标属性的两个方面。所谓“弱”是指目标的强度，反映到图像上就是指目标的灰度，在图像中从对比度和信噪比两个方面来描述目标信号的强弱。而“小”指的是目标尺寸，在图像上表现为目标所占的像素个数，一般在图像中目标的几何尺寸小到无法提取任何形状信息，只是一个亮点或亮斑。

在低信噪比情况下检测和跟踪未知位置和速度的运动小目标的主要难点在于：

(1) 由于成像距离远，经过较长距离的大气衰减，传感器接受到的目标能量较弱，噪声和背景杂波在视场中所占比例大，图像信噪比降低，使得目标的检测变得困难；

(2) 小目标可供利用的信息很少，没有明显的形状、尺寸、纹理等信息可以利用，仅仅利用目标本身的灰度信息是很难检测出目标的，必须同时考虑目标的运动信息；

(3) 缺少关于背景的统计先验信息；

(4) 目标的信噪比非常低，以至于很难从单幅图像中检测出来；

(5) 目标可能会在未知时间点出现或消失；

(6) 算法的复杂性和实时性之间矛盾比较突出。

传统的序列图像弱小目标检测与跟踪方法是先根据目标像素强度检测出目标，然后通过序列图像得到目标的运动轨迹，称之为跟踪前检测算法 (Detect Before track, 简称 DBT)；而检测前跟踪算法 (Track Before Detect, 简称 TBD) 则是在每一次扫描后将结果存储起来，然后对假设轨迹包含的点做没有信息损失的相关处理，经过若干帧的累积后，同时得到检测结果与目标的轨迹。

1.2.4.1 TBD方法

TBD 方法是在三维图像中对较多的可能轨迹进行跟踪，但起初并不对这些轨迹是否代表真正目标做出判断，而是对每条跟踪的轨迹计算其后验概率。如果某条轨迹的后验概率函数值超过设定的门限，就认为该条轨迹代表一个目标。近年

来出现的高阶相关、多级假设检验、动态规划、三维搜索匹配滤波等方法都是通过多帧图像的相关处理,利用目标的运动特性,沿目标轨迹积累能量,提高信噪比,从而达到检测识别小目标的目的。总的来说,TBD 算法结构复杂,存储量大,计算量大。

(1) 高阶相关方法

高阶相关法是近几年提出的一种序列图像点目标检测算法。这种算法利用目标轨迹点的时空相关性,通过计算不同帧之间的高阶相关性检测出真实目标。该类算法可以在有噪声的三维图像中检测出作直线或曲线运动的目标^[92,93,94]。该方法和多假设轨迹跟踪类似,当轨迹段达到一定长度时,则确认轨迹,否则作为噪声滤除。Liou^[95]提出了基于神经网络的高阶时空相关算法,它结合目标的运动信息,增加适当的约束,从而提高了背景抑制率,减少了运算量。

将高阶相关法与神经网络技术相结合是目前的一种趋势,它能发挥各自优点,实现高性能的检测与跟踪。但高阶相关算法处理的是二值图像,如果二值化阈值选择的较高,会导致漏检率较高。反之则运算量大。进行相关的阶数对算法的有效性也有很大影响,阶数高时,计算量和存储量都增大;阶数低时,又会造成虚警率增高。

(2) 基于多级假设检验方法

Blostein 和 Huang^[96]提出了多级假设检验的序贯处理算法来达到减少计算量的目的。这一方法将众多可能的目标轨迹以树的结构组织起来,通过对每帧图像的树进行假设检验,以随时去掉没有通过检验的树,达到减少运算量和存储量的目的。随后,Blostein 和 Huang^[97]又提出了多重多级假设检验,该算法实质上是把多阶假设检验判断出的轨迹进行跟踪,进一步滤除虚假轨迹,实现低信噪比序列图像的目标跟踪。李红艳^[98,99]提出了两种快速检测小目标的算法,算法首先采用小波变换对每帧图像进行滤波,然后分别利用基于假设检验算法形成起始轨迹搜索方法、固定采样假设检验算法形成起始轨迹,再对这些起始轨迹进行截断序贯似然比检验,确认轨迹。

该类方法主要有三大优点:第一,传统的具有固定采样长度的假设检测方法在对每一条候选轨迹的计算上效率不高,而多级假设检测方法克服了这一缺点;第二,由于许多候选轨迹在初始几帧中具有相同的观测结果,多级假设检测方法利用了这一特点来减少算法的计算量和存储量。第三,多级假设方法易于与多目标跟踪算法相结合构成融合算法。虽然这种方法具有同时检测出多个做不同方向直线运动目标的能力,但在低信噪比情况下,点目标的各种特性与噪声点更加接近,候选轨迹起始点非常多,可能导致后面的树杈急剧增多,出现组合爆炸,使得计算量迅速增加。

(3) 动态规划方法

动态规划算法是分级决策方法和最优化原理的综合应用,通常又称为 Viterbi 算法, Barniv^[100]试图用动态规划方法代替穷尽搜索。在该方法中,对所有可能的目标轨迹进行穷举搜索,返回单个最好的轨迹作为目标的运动轨迹。随后,Arnold^[101]进一步发展了这类算法,但该算法对于起伏目标模型处理效果较差。Tonissen 等人^[102,103]直接利用假设目标的幅度信息构造值函数,可以检测起伏模型的运动目标,此算法虽然具有良好的检测性能,但跟踪性能较差。Johnston^[104]在 Tonissen 研究基础上,利用极限理论分析了动态规划算法的性能,得出虚警概率和检测概率的近似表达式。李斌和彭嘉雄^[105]根据定义的一种距离,运用动态规划算法找出目标及其运动轨迹。强勇等人^[106]深入研究了二类动态规划算法的机理,将其统一到一个基本递推方程中,并利用目标的一些先验信息,改进了基本的动态规划方程。卢焕章等人^[107,108,109]对动态规划算法进行了大量研究与应用,先后提出了基于动态规划与轨迹置信度检验的弱小目标检测方法、基于方向加权的动态规划算法和二值轨迹关联算法,有效地实现了弱小运动目标的检测。黄林梅^[110]提出了一种改进的实用动态规划算法,易于硬件实现。

动态规划算法将小目标的轨迹搜索问题转换为最优寻迹问题,沿目标轨迹积累能量。该类方法是基于像素级的操作运算,便于硬件实现。但当目标运动速度未知时,计算中所需的速度窗数目无法确定,一旦将速度窗的参数范围放宽,计算量将迅速增大。

(4) 三维匹配滤波方法

Reed 等人在低信噪比序列图像运动小目标检测方面做了系统的研究工作。他们把匹配滤波器理论推广到三维序列图像,提出了三维匹配滤波器理论^[111]。这种方法的基本思想是针对目标所有可能的运动情况,设计相应的多个三维匹配滤波器,将每个滤波器对序列图像的滤波输出结果进行统计,选出使输出信噪比最高的滤波器,根据该滤波器所对应的运动状态确定点目标在图像中的位置和运动轨迹。由于每个匹配滤波器对应着一条完整的轨迹,这种方法可实现对多条轨迹的同时检测。

当已知目标轨迹的速度和方向时,匹配滤波器能从噪声中很好的检测出轨迹。然而,实际情况中目标速度与方向往往未知。为此,Stocker^[112]和 Chen^[113]利用速度滤波器组实现目标检测。一组速度滤波器就是一种速度参数不同的匹配滤波器,它通过划分速度空间实现对不同速度的运动目标检测。然而,对运动参数取值范围较宽的小目标,需要设计较多的速度滤波器,在实际中难于实现。Porat^[114]提出了在频域利用三维匹配滤波器组检测目标的方法。由于三维傅立叶变换计算量很大,因此 Reed 利用在前后各帧图像中匀速直线运动的目标在二维傅立叶变换域上

只相差一个相位的特性,将三维匹配滤波算法简化为不同帧图像的二维傅立叶变换的递推加权叠加形式,称之为递推运动目标指示算法^[115],以实现多帧目标能量的相干累加。熊艳等学者^[116]在递推运动目标指示算法的基础上,将目标的运动模型扩展到加速度恒定情况,提出了一种基于线性变系数差分方程的运动目标检测算法。Bar-Shalom^[117]提出的数值优化算法估计出目标运动参数,以避免速度失配造成的性能损失。宋柳平等^[118]以直线的三维处理为手段实现了匀速运动和机动运动点目标的快速检测。

三维匹配滤波算法是一种应用较为广泛的方法,它把运动小目标检测问题转换为三维变换域中寻找匹配滤波器的问题,使得该类算法具有较好的低信噪比下检测运动小目标能力。但这类方法通常不可实现穷尽搜索,所以只能用于某些特定的场合,如目标做匀速直线运动且目标大小固定的场合。

(5) 基于投影变换和三维搜索相结合的方法

这种检测技术的基本思路是:通过某种形式的投影变换或逻辑运算,将三维序列图像中的目标轨迹检测问题转化为二维投影平面上的轨迹检测问题,使搜索轨迹的运算量减少。序列中每一帧图像经过门限处理后,将检测出点投影在同一投影平面上。在对投影平面进行门限处理后,可采取一些计算量较小的后处理方法检测出轨迹^[119]。但这样检测出的轨迹还无法与三维空间中的轨迹进行对应。更为一般的方法是由投影法确定目标轨迹在三维空间中的大致范围后,再回到三维空间中匹配滤波。

廖斌^[120]提出了综合运用自适应门限检测方法和 Hough 变换的低信噪比红外小目标检测算法,有效地检测出作线性运动的红外小目标。艾斯卡尔^[121]提出了一种基于准优分布变换的微弱点状运动目标检测技术。该方法首先经过变换使未知统计分布变为准高斯分布,再将时空三维搜索简化为沿时间轴投影与二维空域投影进行弱小目标检测。

投影方法减少了检测过程的数据量和存储量,向实时处理迈进了一步,但这种进步以性能下降为代价,特别在噪声较强和目标帧间位移较大时,投影造成的性能下降往往难以接受。

(6) 递推贝叶斯滤波

从贝叶斯学派的观点看,解决 TBD 问题就是在给出 $k-1$ 时刻目标联合状态 (s_{k-1}, δ_{k-1}) 的后验概率密度 $p(s_{k-1}, \delta_{k-1} | z_{1:k-1})$ 及最新测量 z_k 的条件下,估算 k 时刻目标联合状态 (s_k, δ_k) 的后验概率密度 $p(s_k, \delta_k | z_{1:k})$ 。与其他 TBD 算法最大的不同之处在于,基于递推贝叶斯滤波的 TBD 算法通过目标运动模型和测量模型,真正引入了跟踪的思想和算法。由于贝叶斯滤波在工程上难以实现,可通过粒子滤波^[122,123]对其进行实现。基于粒子滤波检测小目标的思想首先是由 Samond^[124]提出的,他通

过运动方程预测出下一时刻目标的运动状态，再利用传感器的测量数据对其进行更新，得到该时刻状态的后验概率分布，最后以目标出现概率作为目标检测的判断准则，检测出真实目标，并估计目标在空间平面内的位置。Ristic^[125]完善了Samond的工作，给出了详细的实现步骤，并且评价了该算法的跟踪误差及检测性能。在Ristic的基础上，Rutten等^[126]对粒子滤波TBD算法作了深入研究，在研究中所采用的测量噪声不是传统的高斯噪声，而是Ricean噪声和Rayleigh噪声，这样使测量更加接近真实数据，并且通过推导得出目标出现概率的计算公式。针对低信噪比下幅值波动的弱小目标跟踪问题，杨小军^[127]提出了一种基于粒子滤波和Bayes似然比方法的联合检测和跟踪算法。

基于递推贝叶斯滤波的TBD算法得到的是目标状态的概率分布，这样可以充分反映出目标的状态信息，能够实现低信噪比下小目标检测，而且运动模型和测量模型均可以是非线性、非高斯，对目标的要求也不需要限定在匀速条件。

1.2.4.2 DBT方法

最简单的DBT方法是通过每个单帧图像进行阈值化处理，将大于阈值的像素判定为可能的目标，然后通过关联多帧图像数据决定其是否为真正的目标。近年来出现了基于形态学的方法、基于像素分析的方法、小波分析方法、基于变换的方法等。该类方法需要被处理的图像中的目标能量足够大，否则不能很好地检测出目标。

(1) 基于形态学的方法

数学形态学是一种用于数字图像处理的方法^[128]。该方法能够去除高频噪声的干扰，是一种较为有效的小目标检测方法。它的主要内容是设计算子用以描述图像各个像素之间的关系，常用的运算包括开、闭、腐蚀、膨胀等。

近些年，国内外学者提出了许多有实际意义的算法。如Wang等人^[129]设计了一种基于数学形态学的红外灰度图像小目标检测算法，并用FPGA硬件实现。张彦^[130]构造了由几种形态学运算组成的非线性算子，结合刚性小平面的三维分析技术快速检测出红外图像中的小目标。

通常，点目标的检测采用多帧累加方法来提高信噪比，该方法对运动速度较快的目标不能实现有效的能量累加，而数学形态法则很好地克服了这一点，并能很容易地检测出低信噪比快速运动的点目标。

(2) 基于像素分析的方法

基于像素分析的方法是根据图像中像素灰度模式的变化来实现运动目标检测、识别与跟踪。该类方法中最直接的方法是对前后两帧图像进行差分运算，这种方法的前提条件是传感器与目标之间的运动是平行移动，在较短的时间间隔里背景灰度基本保持不变，利用前后两帧图像的差分图像信息检测出运动目标。

Cassasent 等人^[131]分析评估了移动背景下探测和跟踪小目标的六种算法(单差、双差、线性内插、抛物线内插、空间差及空间滤波)的性能; Rauch 等人^[132]采用一组时间上的高阶差分来抑制背景干扰, 并得到目标的运动轨迹。

另一种基于像素分析的方法是光流法, 光流法利用图像中目标的运动是由目标像素亮度运动确定, 推导出瞬时光流场, 然后根据光流场进行目标运动检测。文献^[133]给出了运动约束方程:

$$I_x \frac{dx}{dt} + I_y \frac{dy}{dt} + I_t = 0 \quad (1.1)$$

其后光流法都是基于上述约束方程。Horn 和 Schunck^[134]指出光流计算存在病态解问题, 必须引入附加条件才能确定光流的唯一解; Uras^[135]、Reichardt^[136]、Markandey^[137]、李晓春^[138]提出了几种应用光流法检测运动目标的方法; 光流法的理论和算法存在以下几方面的问题: 约束方程只有在梯度很大的点上才严格成立^[139], 计算不稳定, 运算量大等, 上述问题限制了光流法的理论和算法的应用。

(3) 小波分析法

红外图像中背景部分往往呈大面积的连续分布状态, 在红外辐射的强度上也呈渐变过渡状态, 从而使得它们的红外图像在图像灰度空间分布上具有较大的相关性。目标的红外辐射强度与其周围自然背景的辐射强度无关, 且一般都高于背景的辐射强度, 可以认为背景处在图像信号的低频部分, 而目标处在图像信号的高频部分, 通过小波变换将红外图像的低频部分和高频部分进行分离, 然后对各高频分量进行分析, 检测出真实小目标。

Li^[140]和 Zuo^[141]先采用小波变换增强图像抑制噪声, 然后检测目标。Sheng^[142]通过对图像信号作多尺度小波变换, 检测弱小目标和提取不同分辨率下背景纹理的能量特征。Li^[143]通过一种鲁棒的小波多尺度图像配准过程消除主场或背景运动的影响, 然后检测出小目标; Giuseppe^[144]指出, 在经过小波变换滤波后, 可以利用不同分解尺度上的评估信息将目标从背景中分离出来。Li^[145]采用小波滤波器对每帧图像滤波, 提高目标的信噪比, 剔除部分噪声点, 降低了算法的计算量。

利用小波变换实现频率选择和多尺度分解, 可以起到抑制背景噪声和增强目标的作用, 极大地提高信噪比, 从而提高单帧图像中目标检测概率, 降低误检测。当时间分辨率高时, 适合于分析高频信号, 可以更加精确估计信号所在的位置; 当频率分辨率低时, 适合于分析低频信号, 可以观察过程的全貌, 这使得基于小波变换的检测算法具有较好的鲁棒性。

(4) 基于变换的方法

基于变换的方法是利用某一种形式的变换来实现目标的检测与跟踪, 对数学理论的依赖性较强。基于 Fourier 变换的方法是其中比较典型的方法之一, 它利用

对序列图像进行 Fourier 变换时,其结果的频谱幅度不变而相位谱的变化是一个常数的性质检测运动目标。Mahmoud^[146,147]分别用 Hartley 变换和余弦面积变换(CAT)来实现目标检测与跟踪。Falconer^[148]和 Snyder^[149]用 Hough 变换实现目标检测与跟踪。Dong^[150]将 Wigner Ville 分布和 Hough 变换技术相结合,实现小目标检测。这类算法的特点是结构简单,运算速度快,具有较低信噪比下的目标检测能力,是工程应用较多的方法。

1.2.4.3 性能分析

综合分析以上算法,可以得到以下结论:

(1) DBT 算法得到一个观测结果列表,根据目标的运动假设,将它们和已经存在的轨迹相关联,由此决定哪些轨迹是目标,即通过二元决策得到目标,因而误检率较高;

(2) TBD 算法考虑到时间信息与空间信息的相关性,计算量要大于 DBT 算法,但在较高信噪比条件下,DBT 算法的性能要相对优越一些;

(3) 在较高信噪比条件下,检测帧数的增加使得检测概率迅速增大,DBT 算法和 TBD 算法都能在几帧内有效地检测出小目标;而在低信噪比条件下,TBD 算法能在帧数增加的情况下具有较好的检测性能。

1.2.5 无人机着陆导航技术发展现状

无人机进场着陆是执行飞行任务的重要阶段之一。装备先进的着陆导引设备实现精密进场和着陆,对于完成作战训练任务,尤其是对于保障飞行安全极为重要。要实现安全着陆,无人机必须具有较强的导航定位的能力,因此高精度的导航定位技术是实现无人机着陆的关键。

目前用于无人机着陆的导航技术主要包括:惯性导航、GPS导航、INS/GPS组合导航和视觉导航。

惯性导航技术是最早研究、最为成熟的导航技术,也是当前飞行器导航中应用最广泛的一项技术。惯性导航利用惯性敏感元件测量飞行器对于惯性空间的线运动和角运动参数,在给定的运动初始条件下,由计算机推算出飞行器的姿态、方位、速度和位置等参数,从而引导飞行器完成预定的航行任务,是一种自主式的导航设备^[151]。惯性导航系统的优点在于:(1)完全自主式,保密性强,不易受外界条件和人为因素的干扰;(2)既不发射信号,也不接收信号,不存在电磁波传播问题,因此没有无线电干扰;(3)全天候,不受天气限制。但是其缺点在于定位误差随时间而迅速累积增长,在长时间工作后,会产生不同程度的累积误差。因此惯性导航系统不能单独使用,必须与其他导航方法组合使用。

GPS导航技术兴起于70年代,是应用最为广泛、相对成熟的导航技术。该技术

采用多星、高轨、测距体制，以距离作为基本的观测量，通过对四颗卫星同时进行距离测量计算出用户接收机的位置^[152,153]。相对于惯性导航系统而言，GPS定位和测速精度高，并且其误差不随时间的推移而累积，另外它基本上不受地区的限制。GPS导航在和平时期是一种很好的自主导航方法，但是在战争时期，该种导航方式将完全受制于美国军方，而我国的双星定位还不成熟，因此它只能算是一种半自主导航定位技术。

20世纪80年代开始，美、英、法等国的军方和一些民用部门开始对INS/GPS组合导航系统进行研制，并逐渐得到了广泛的应用，它的优点是可以充分利用GPS和INS的互补特点，改善系统的精度，加强系统的抗干扰能力，而且INS可以辅助GPS对信号捕获，从而提高GPS接收机的跟踪能力^[153]。

基于计算机视觉的自主着陆导航技术一出现就备受关注。该类方法主要通过安装在无人机上的摄像机，获得着陆点附近的图像，使用计算机视觉算法，估计出无人机的飞行状态和相对于着陆点的位置，结合其他机载传感器，实现无人机的自主着陆控制和准确着陆。该类导航技术具有不依赖于地面和空中导航设备的优点。最近一些年来，计算机视觉在无人机自主着陆^[154,155,156]等方面得到了广泛的应用并获得了一定的成功。该类导航技术的核心工作之一就是选择特殊形状的合作标志加载在跑道上，使得无人机姿态和方位参数的解算更为简单有效，其中特殊形状的合作标志如图1-1所示，图1-1.a为美国加州大学伯克利分校设计的图案^[154]；图1-1.b为美国南加州大学设计的图案^[155]；图1-1.c为北京航空航天大学设计的图案^[157]。

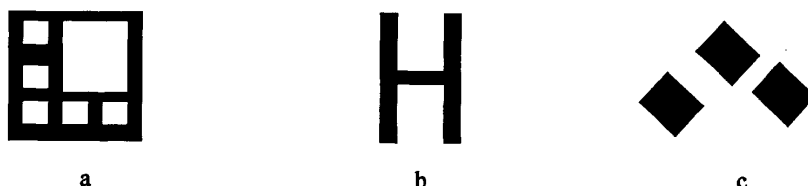


图1-1 选择的不同特征图案

1.3 论文主要内容及创新点

1.3.1 论文主要内容

论文以无人机侦察为应用背景，以实现无人机平台对运动目标检测、跟踪以及无人机辅助着陆系统为目的。论文共分八章，前两章是全文的基础。第一章介绍了论文的研究背景；第二章介绍论文涉及的基本理论；第三章研究了图像角点特征提取定位算法；第四章围绕无人机平台像机自运动消除进行了研究；第五章针对复杂背景下运动目标跟踪问题进行了深入研究；第六章研究了目标尺度方向

自适应跟踪问题；第七章研究了无人机辅助着陆系统及弱小目标实时跟踪问题；第八章对全文做出了总结，并进行了展望。各章的组织结构关系如图 1-2。

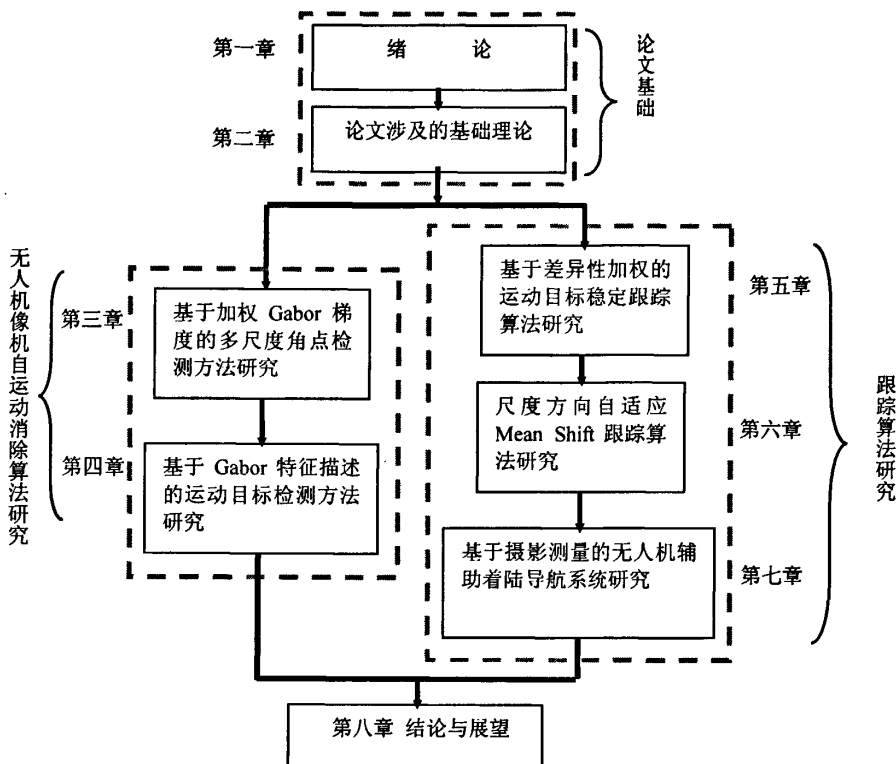


图 1-2 论文组织结构

本文各章节内容安排如下：

第一章介绍了论文的研究背景，对相关研究进行了综述。

第二章主要介绍论文涉及的基本理论。本章首先介绍了尺度空间的基本理论和性质，然后详细介绍了 Gabor 小波及其生物学背景，最后对 Mean Shift 算法进行了详细介绍。本章主要是为后续章节奠定理论基础。

第三章对图像角点特征检测问题进行了研究。本章简要介绍了角点特征分类及检测准则，给出了角点检测算法分类，详细介绍了四种经典的角点检测算法。在分析传统梯度算子不足的基础上，提出了并实现了基于 Gabor 梯度算子的角点检测算法。该方法在角点定位准确性和抗噪性等方面，比经典算法有所提高。新的角点检测方法为基于特征点匹配的运动目标检测算法研究奠定了基础。

第四章对像机自运动消除问题进行了研究。本章提出并实现了一种利用 Gabor 特征描述的像机自运动消除方法。该方法利用了尺度不变性理论在图像上提取兴趣点，然后对每个兴趣点建立 Gabor 特征描述矢量，最后进行特征匹配，消除了像机自身运动。该方法具有自动化程度高、抗背景噪声能力强和检测效果好等优

点。

第五章对复杂背景下运动目标跟踪问题进行研究。本章首先提出一种目标差异性权值计算方法,该权值较好地描述目标区域特性,对背景有较好的抑制作用。为了适应复杂背景下目标在时间序列上发生的平移、旋转和缩放变化情况,将最小二乘影像匹配算法(LSIM)应用到运动目标跟踪问题中,提出并实现了一种基于目标差异性加权的最小二乘影像匹配运动目标跟踪算法(DWLSIM)。本章详细比较了LSIM算法和DWLSIM算法在目标发生各种变换及噪声影响下的跟踪效果。跟踪实验结果表明DWLSIM算法能够降低跟踪窗口中背景点对最小二乘匹配的影响,较LSIM有更高的跟踪精度和抗噪声能力。本章将差异性权值与Mean Shift跟踪算法相结合,提出并实现了一种基于差异性权值的Mean Shift跟踪算法(DWMS),并对算法的收敛性及收敛条件进行了讨论。对复杂背景下运动目标进行跟踪,DWMS算法比经典算法更为有效稳定。

第六章对同时存在尺度方向变化的运动目标跟踪进行了研究。本章在分析尺度空间理论的基础上,提出并实现了一种目标最佳椭圆描述区域确定方法,并在此基础上提出并实现了一种尺度方向自适应Mean Shift跟踪算法。实验结果证明目标最佳椭圆描述区域可以较好地描述目标区域的大小和方向,新的跟踪算法可以根据目标位置、大小及方向的变化自适应地做出调整,跟踪效果较为理想。

第七章对无人机着陆导航问题进行探讨。本章针对弱小目标实时高精度检测跟踪问题,提出并实现了一种正负正则化LOG算子的弱小目标实时高精度检测跟踪算法,同时设计并实现了一套基于摄影测量的无人机辅助着陆导航系统。实验证明,新的跟踪算法在实时性、检测准确性以及抗噪声性上均有较好表现。无人机辅助着陆系统可以实时高精度地获得无人机相对于跑道的位置信息,同时可以实时保存无人机整个着陆过程的序列图像。

第八章是结束语,对本文的研究工作进行总结,归纳了本文的创新点,分析了本文应用前景及不足,并对继续开展的工作进行了讨论。

1.3.2 主要研究成果的发表情况

表 1-1 博士学位论文主要研究成果的发表

| 序号 | 发表论文题目 | 刊物名称 | 发表时间 及排名 | 对应博士学位论 文章节 | 备注 |
|----|-------------------------|-----------------|------------------|------------------|----|
| 1 | 基于加权Gabor梯度的新型多尺度角点检测方法 | 中国图象图形学报第12卷第8期 | 2007.08, 排名第1 | 3.3, 3.4, 3.5 | |

| | | | | | |
|---|--|--------------------------------|----------------|---------------|--|
| 2 | 一种运动平台下运动目标检测方法 | 光电工程第 35 卷第 9 期 | 2008.09, 排名第 1 | 4.3, 4.4, 4.5 | |
| 3 | 基于差异性加权的最小二乘图像匹配的跟踪方法 | 光电工程第 35 卷第 4 期 | 2008.04, 排名第 1 | 5.2, 5.3 | |
| 4 | 基于差异性加权的 Mean Shift 跟踪方法 | 光学技术第 34 卷第 3 期 | 2008.03, 排名第 1 | 5.4, 5.5, | |
| 6 | 一种目标区域最佳椭圆描述确定方法 | 应用光学第 29 卷第 4 期 | 2008.04, 排名第 1 | 6.2 | |
| 6 | A scale rotation adaptive new Mean Shift tracking method | 2007 Photonics Asia Conference | 排名第 1 | 6.3, 6.4, 6.5 | |
| 7 | 一种尺度自适应小目标实时检测方法 | 应用光学第 29 卷第 1 期 | 2008.01, 排名第 1 | 7.2.3 | |

第二章 基本理论介绍

为了便于读者更好的理解后续章节内容，本章对论文涉及的基本理论进行简要介绍，具体安排如下：2.1 节介绍了尺度空间理论和 Lindeberg 特征尺度选择理论；2.2 节简要介绍了 Gabor 小波，给出了 Gabor 小波的生物学背景以及 Gabor 小波变换的数学形式；2.3 节详细介绍了 Mean Shift 理论的基本原理，包括概率密度的基本估计形式、Parzen 窗方法、Mean Shift 方法以及核密度估计的带宽选择问题；2.4 节对本章内容进行总结。

2.1 尺度空间理论

计算机视觉的根本任务是利用计算机系统解释图像，以得到对客观世界图像的正确描述，实现类似人类视觉系统的功能。世界上的万物只有在特定尺度下才能成为有意义的实体，并在不同的尺度下表现出不同特性的成像，所以客观世界图像具有多尺度特性。人的视觉系统也是一个多尺度系统。例如，一片麦田，近看时，可以看到各个麦苗的细节信息，但是远看时，不足以把麦苗区分清楚，而只能将一大片麦苗看作一个整体。从上面的分析可以看出，仅仅在某一固定尺度下分析图像不能表现出图像本身固有的多尺度特性，从而影响了分析结果的准确性。然而，图像分析与理解往往是在某一固定尺度和若干人为设定的先验尺度上进行的。为解决这个问题，人们试图在一组连续尺度上分析图像，而且使不同尺度层上图像之间满足某种特定的关系，以达到综合各尺度层上信息的目的。本节对尺度空间理论中的基本概念和理论进行简要介绍。

2.1.1 尺度空间基本概念

尺度空间理论是有一个形成过程的，早在尺度空间理论提出之前，Marr-Hildreth 就在边缘检测理论中采用了多尺度方法^[158]，他们考虑图像的拉普拉斯变换，并和不同方差的高斯滤波器进行卷积，以检测不同灰度范围内的边缘。

1983 年，Witkin 扩展了 Marr-Hildreth 理论，第一次提出了尺度空间概念^[49]，并给出了一维连续信号的尺度空间定义，这使得尺度空间方法逐渐得到关注和发展。随后，尺度空间理论又诞生出多个分支：高斯尺度空间、小波尺度空间、形态学尺度空间和 Schwarz 尺度空间等。

尺度空间方法的基本思想是：在视觉信息（图像信息）处理模型中引入一个被称为尺度的参数，通过连续变化尺度参数获得不同尺度下的视觉处理信息，然后综合这些信息用以深入挖掘图像的本质特征。尺度空间方法将传统的单尺度视

觉信息处理技术纳入尺度不断变化的动态分析框架中, 因此更容易获得图像的本质特征。

设 $f: R^N \rightarrow R$ 表示任意给定的信号, $f(\mathbf{x})$ ($\mathbf{x} \in R^N$) 的尺度空间 $s(\mathbf{x}, \sigma)$ 是一族信号的集合 $\{L(\mathbf{x}, \sigma), \sigma \geq 0\}$, 这一族信号是原信号在不同尺度下的尺度空间表示。

$$s(\mathbf{x}, \sigma) = \{L(\mathbf{x}, \sigma), \sigma \geq 0\} \quad (2.1)$$

其中

$$L(\mathbf{x}, \sigma) = G(\mathbf{x}, \sigma) * f(\mathbf{x}) \quad (2.2)$$

σ 为尺度参数, $G: R^N \times R \rightarrow R$ 是高斯核:

$$G(\mathbf{x}, \sigma) = \frac{1}{(2\pi\sigma)^{N/2}} e^{-\frac{\mathbf{x}^T \mathbf{x}}{2\sigma^2}} \quad (2.3)$$

Witkin 指出, 当高斯核方差增加时, 高斯核和信号卷积所得结果中的精尺度信息会被压缩, 信号会随着平滑尺度的增大越来越平滑。如图 2-1 所示, 其中图 2-1.a 为原始图像, b 到 f 依次为高斯核方差为 1、2、3、4、5 时原始图像的尺度空间表示, 当尺度参数 σ 逐渐增大时, 图像细节信息渐渐被平滑。

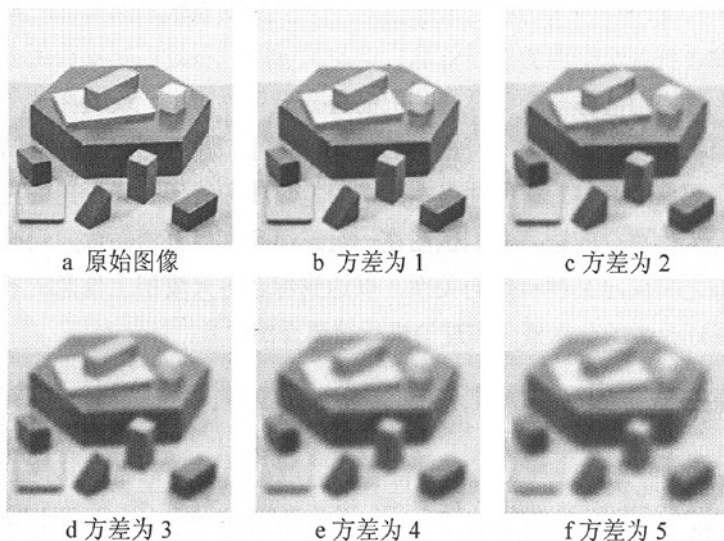


图 2-1 图像被尺度渐增的高斯核平滑的结果

2.1.2 Lindeberg 的特征尺度选择理论

Lindeberg 将连续尺度空间理论拓展到离散空间, 对于图像, 提出了目标特征尺度选择理论和尺度不变性理论。他提出由 σ^2 正则化的二维高斯拉普拉斯算子 (Laplacian Of Gaussian) 具有尺度不变性检测能力, 而且认为图像局部结构的最佳描述区域为不同尺度的正则化 LOG 算子响应的极大值所对应的尺度参数 σ_c 的三倍大小所描述的区域, 其中特征尺度参数 σ_c 描述了目标内部特征大小。

由 σ^2 正则化的 LOG 算子 (Normalized LOG, 简称 NLOG) 表示为:

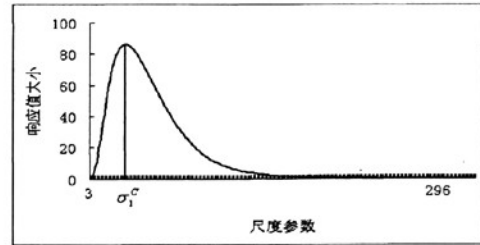
$$NLOG = \sigma^2 \nabla^2 G = \sigma^2 [L_{xx}(\mathbf{x}, \sigma) + L_{yy}(\mathbf{x}, \sigma)] = \left(\frac{x^2 + y^2 - 2\sigma^2}{2\pi\sigma^4} \right) e^{-\frac{x^2 + y^2}{2\sigma^2}} \quad (2.4)$$

其中 $L_{xx}(\mathbf{x}, \sigma)$ 和 $L_{yy}(\mathbf{x}, \sigma)$ 分别 $L(\mathbf{x}, \sigma)$ 关于 x 方向和 y 方向的二阶导数。

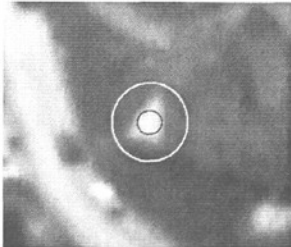
图 2-2.a 为无人机平台上的摄像机拍摄的原始图像, 其中白色十字表示选择的目标点。白色十字为中心、三倍特征尺度 σ_1^c 为半径的白色圆圈区域描述了三角形图像局部结构的最佳描述区域大小。白色十字为中心、特征尺度 σ_1^c 为半径的黑色圆圈描述了三角形结构的内部特征大小。图 2-2.b 为该点的不同尺度参数正则化 LOG 算子的响应曲线, 尺度参数选择为从 3.0 开始间隔 0.5 递增到 296.0。特征尺度 $\sigma_1^c = 29.0$ 为响应曲线局部最大值所对应的尺度参数, 对应的响应值为 85.88。图 2-2.c 为原始图像放大两倍后局部图像, 其中白色十字为原始图像选择的目标点的对应点。白色十字为中心、三倍特征尺度 σ_2^c 为半径的白色圆圈区域描述了放大后三角形图像局部结构的最佳描述区域大小。白色十字为中心、特征尺度 σ_1^c 为半径的黑色圆圈描述了三角形结构的内部特征大小。图 2-2.d 为该点的不同尺度参数正则化 LOG 算子的响应曲线, 尺度参数选择和图 2-2.b 相同, 响应曲线局部极大值所对应的特征尺度 σ_2^c 为 58.5, 对应的响应值为 85.82。从图 2-2.a 和图 2-2.c 可以看出, 三倍特征尺度大小的区域较好地定义了描述该目标区域的大小, 且特征尺度之比反映了图像的缩放变化。对同一空间目标在不同距离拍摄进行拍摄时, 目标在图像上表现出一定的缩放变化, 而特征尺度会根据目标在图像上的大小变化作出相应的调整, 而特征尺度所对应的响应值基本保持不变。



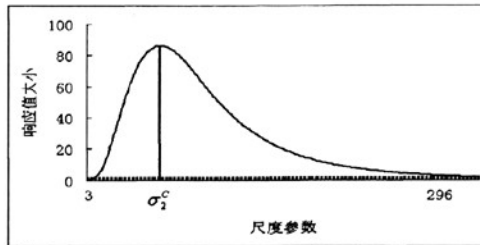
a 原始图像目标最佳描述区域



b 不同尺度参数响应曲线



c 原始图像放大两倍后目标最佳描述区域



d 放大两倍后不同尺度参数响应曲线

图 2-2 不同尺度的正则化 LOG 算子对指定目标点的响应

由于正则化 LOG 算子计算量较大, Lindeberg^[159]于 1994 年提出 DOG 算子 (Difference of Gaussian) 可以作为正则化 LOG 算子的近似表示, 即

$$DOG = G(\mathbf{x}, k\sigma) - G(\mathbf{x}, \sigma) \approx (k-1)\sigma^2 \nabla^2 G = LOG \quad (2.5)$$

其中(2.5)式中的比例因子 k 在尺度空间构造时为一个常数, 该因子并不影响特征尺度的大小, 通常 k 取为 $\sqrt{2}$ 。图像上任意一点不同参数 DOG 算子响应曲线类似于图 2-2.b。

2.2 Gabor 小波

2.2.1 Gabor 函数的提出

众所周知, 在十九世纪初, 法国数学家傅立叶提出的 Fourier 级数具有非比寻常的意义。Fourier 分析方法已经作为信号处理领域最基本的数学工具而被广泛应用。Fourier 变换的基本定义如下:

若在 $L^2(R)$ 中, 模拟信号函数 $f(t)$ 具有有限的能量, 则其 Fourier 变换定义为:

$$\hat{f}(\omega) = \int_{-\infty}^{\infty} e^{-i\omega t} f(t) dt \quad (2.6)$$

由于 Fourier 变换的基函数在时间轴上是无限扩展的, 任何时间局部的信息可以被扩展到整个频率轴上。因而信号的频谱 $\hat{f}(\omega)$ 只能刻画 $f(t)$ 在整个时间轴上的频谱特征, 而不能反映信号在时间轴上的局部区域的频率特征。因此 Fourier 变换适于分析稳定的缓慢变化信号。但是, 在很多实际问题中, 我们所关心的恰恰是信号的局部特征。

1946 年, Gabor^[160]注意到 Fourier 变换在表示非平稳信号方面的不足, 通过与量子力学中的 Heisenberg 不确定性原理的类比, 发现并证明了一维信号的不确定性原理, 即一个同时用时间和频率来刻画信号特征受它的带宽和持续时间乘积的下限所限制。他进一步发现对于一般的信号来说, 都存在最优化的折中, 即任意可以用高斯函数调制的复正弦形式表示的信号都可以达到时域和频域联合不确定关系的下限, 可以同时进行时域和频域获得最佳的分辨率, 这种表示就是 Gabor 函数的最初形式。

2.2.2 Gabor 小波的生物学背景

视觉皮层细胞按其感受野特征可以分为: 简单细胞、复杂细胞和超复杂细胞。简单细胞直接从外侧膝体接受视觉信号, 这一过程构成视觉信息处理过程的关键一步。1980 年, Marcelja^[161]研究发现简单细胞感受野的剖面线可以用一维 Gabor

函数来很好的描述, 并提出了用于描述简单细胞感受野的参数化一维 Gabor 滤波器模型; 与此同时, Daugman 提出了相应的二维 Gabor 滤波器模型^[162]。Pollen 和 Ronner^[163]的研究更进一步说明了这个问题, 他们实验测量了猫的视觉皮层相邻细胞的相位关系, 发现相邻的细胞对具有某种对称关系, 其中一个是偶函数, 另一个是奇函数, 这表明这样的细胞对可以用复 Gabor 函数来模拟。1985 年, Daugman 将 Gabor 不确定性原理从一维推广到二维, 并从理论上证明二维 Gabor 函数达到空间域和频率域的下限^[164]。Jones 和 Palmer^[165]用实验测量了视皮层 V1 区简单细胞空间及空间频谱的响应特性, 在统计意义下, 这些相应特性可以用二维 Gabor 函数很好的拟合。他们指出, 二维 Gabor 函数不仅可以对简单细胞感受野的结构做出统一的数学描述, 而且它们在两个互异空间上可以达到最佳视觉滤波效果。

2.2.3 二维 Gabor 小波

二维 Gabor 滤波器的函数形式为椭圆高斯函数与复平面波的乘积, 具体定义如下:

$$g(x, y, \sigma_x, \sigma_y, \omega_0, \theta) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)\right] \exp[i\omega_0(x\cos\theta + y\sin\theta)] \quad (2.7)$$

式中 σ_x 、 σ_y 为椭圆高斯函数的标准差, ω_0 为复平面波的空间频率。Gabor 函数可以分解为奇函数和偶函数两部分。分别利用特定参数的 Gabor 奇函数和偶函数作为母小波, 通过尺度、平移、旋转参数的变化可以将它扩展为一组小波函数:

$$\begin{aligned} \psi_n(x, y | c_x, c_y, \theta, s_x, s_y) = & \frac{1}{2\pi s_x s_y} \exp\left\{-\frac{1}{2}\left[\frac{1}{s_x^2}((x-c_x)\cos\theta - (y-c_y)\sin\theta)^2\right.\right. \\ & \left.\left. + \frac{1}{s_y^2}((x-c_x)\sin\theta + (y-c_y)\cos\theta)^2\right]\right\} \times \sin\{\omega_0[(x-c_x)\cos\theta + (y-c_y)\sin\theta]\} \end{aligned} \quad (2.8)$$

$$\begin{aligned} \bar{\psi}_n(x, y | c_x, c_y, \theta, s_x, s_y) = & \frac{1}{2\pi s_x s_y} \exp\left\{-\frac{1}{2}\left[\frac{1}{s_x^2}((x-c_x)\cos\theta - (y-c_y)\sin\theta)^2\right.\right. \\ & \left.\left. + \frac{1}{s_y^2}((x-c_x)\sin\theta + (y-c_y)\cos\theta)^2\right]\right\} \times \cos\{\omega_0[(x-c_x)\cos\theta + (y-c_y)\sin\theta]\} \end{aligned} \quad (2.9)$$

其中 $(c_x, c_y, \theta, s_x, s_y)^T$ 为小波参数向量, c_x, c_y 为平移参数, θ 为旋转参数, s_x, s_y 为尺度参数。图 2-3 为一组 Gabor 小波模板, 上一行为奇模板, 下一行为偶模板, 其中从左到右 θ 从 0° 开始间隔 22.5° 增加到 157.5° , ω_0 为 5.0, $s_x = s_y = 18.0$ 。

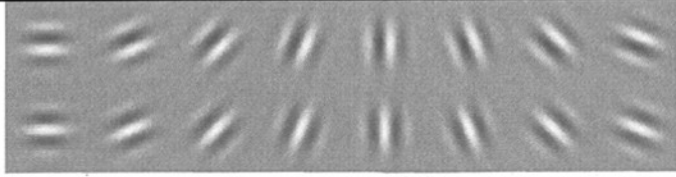


图 2-3 Gabor 函数模板

2.3 Mean Shift 理论

设 $\{\mathbf{x}_i\}_{i=1, \dots, n}$ 是 d 维欧氏空间 R^d 中任意一个点集, 其概率密度函数 p 未知, 那么根据 p 的先验知识, 对概率密度函数的估计 \hat{p} 有两种方法: 参数估计方法和非参数估计方法。

如果假定已知数据点服从某种分布, 如正态分布, 通过样本来估计分布的参数, 如正态分布的均值和方差, 那么这种估计方法为参数估计方法。参数估计方法通常比较简单且容易实现, 但该方法也存在一些致命的缺点, 如要求密度函数的形式已知, 这种假定有时并不成立, 常见的一些函数形式很难拟合实际的概率密度, 经典的密度函数都是单峰的, 而在许多实际情况中却是多峰的, 所以参数估计适用范围有限。

客观世界中, 真实数据很难只服从一种分布, 非参数估计因其不需要假定严格的分布模型而得到广泛应用, 其中最近邻估计和核密度估计是两种典型的非参数估计方法。

Mean Shift 方法也是一种非常优秀的非参数特征空间分析方法, 最初由 Fukunaga 和 Hostetler 于 1975 年提出^[166], 其本质属于核密度估计。该方法一度被人遗忘, 直到 1995 年 Cheng 的相关论文^[167]的发表, 才又一次引起人们的关注。Cheng 对基本的 Mean Shift 方法做了两方面的推广, 首先定义了一族核函数, 使得随着样本与被偏移点的距离不同, 其偏移量对均值偏移量的贡献也不同, 其次还设定了一个权重系数, 使得不同的样本点重要性不一样, 这大大扩大了 Mean Shift 方法的适用范围。另外 Cheng 指出了 Mean Shift 可能应用的领域, 并给出了在图像分析领域具体的例子。此后, 美国 Rutgers 大学的 Comaniciu 在其博士论文^[168]中将 Mean Shift 方法作为主要研究内容, 并在运动目标跟踪和图像分割中进行很好地应用, 验证了 Mean Shift 方法是一种优秀的非参数聚类方法。

核函数的带宽参数选择是 Mean Shift 方法的一个非常重要的因素。2002 年 Maneesh^[169]提出了一种基于小波的自适应带宽选择的 Mean Shift 方法, 2003 年 Dorin Comaniciu^[170]提出了一种基于数据驱动自适应带宽选择的 Mean Shift 方法。同年, Collins 提出了一种基于尺度空间理论的带宽选择机制, 并成功地应用于目标多尺度跟踪领域^[171]。

Mean Shift 方法中,除了带宽参数的选择,对于高维特征空间应用也一直是重要的研究问题。Georgescu 等提出了一种基于 LSH (locality-sensitive hashing) 的快速 Mean Shift 算法^[172],从而使 Mean Shift 方法能够应用于高维特征空间。同年,美国 Maryland 大学的杨长江也对 Mean Shift 方法做了两种改进,一种是采用伪牛顿法代替最陡上升法实现收敛过程^[173],另一种是采用改进的高斯变换加速 Mean Shift 收敛过程^[174],将高斯变换离散分解,并用于 Mean Shift 的迭代收敛过程,使时间复杂度由 $O(M \times N)$ 下降为 $O(M+N)$ 。

论文部分章节针对 Mean Shift 算法提出了改进算法,为了便于对后续章节的理解,本节对 Mean Shift 算法的相关理论作简要介绍。

2.3.1 概率密度的基本估计形式

一个随机样本 x 落在特征空间中某一区域 R 中的概率为

$$P = \int_R p(x) dx \quad (2.10)$$

设 N 个样本 x_1, x_2, \dots, x_N 是从总体中独立抽取的, N 个样本中有 k 个样本落入区域 R 中的概率 P_k 服从离散随机变量的二项分布

$$P_k = C_N^k P^k (1-P)^{N-k} \quad (2.11)$$

使 P_k 取最大的 k 值称为众数,这里记为 m 。如果 $(N+1)P$ 不是整数,则 m 等于 $(N+1)P$ 的整数部分;如果 $(N+1)P$ 是整数,则 $m = (N+1)P - 1$ 和 $m = (N+1)P$ 。众数定义表明有 m 个样本落入区域 R 中的概率最大。若实际落入 R 域的样本有 k 个,这时认为 $k \approx m$ 。由于 $(N+1)P - 1 \leq m \leq (N+1)P$, 所以

$$k \approx NP \approx N\hat{P} \quad (2.12)$$

这里 \hat{P} 是 P 的估计,当 N 较大 P 较小时式(2.12)的近似程度是足够的。当固定 k 时,对 P 的最大似然估计 $\hat{p}_{ML} = k/N$ 。由概率论知, k 的数学期望 $E[k] = NP$ 。综上所述并根据极限情况下的贝努利定理,可以认为 P 的估计 \hat{p} 为

$$\hat{p} = k/N \quad (2.13)$$

设区域 R 的体积为 V , 取 R 足够小, 使

$$p = \int_R p(x) dx \approx p(x)V \quad (2.14)$$

设 $\hat{p}(x)$ 是 $p(x)$ 的估计, 由(2.13)和(2.14)可以得到

$$k/N = \hat{P} = \int_R \hat{p}(x) dx = \hat{p}(x)V \quad (2.15)$$

于是可以得到

$$\hat{p}(x) = \frac{k/N}{V} \quad (2.16)$$

(2.16)式就是 $p(x)$ 的基本估计式，它与 N 、 V 、 k 有关，显然 $\hat{p}(x)$ 和 $p(x)$ 有一定的误差。

从理论上讲，要使 $\hat{p}(x)$ 趋于 $p(x)$ ，就必须让积分区域 R 无限小，即让其体积 V 接近于零，同时让 N 、 k 无穷大。在实际估计时体积 V 不是任意的小，而且样本总数也是有限的，所以 $\hat{p}(x)$ 总是存在误差。为了提高 x 处的概率 $p(x)$ 的估计精度，根据极限理论，可以采用如下步骤一尽量满足理论要求：

(1) 构造包含 x 的区域序列 R_1, R_2, \dots ，各区域 $R_N (N=1, 2, \dots)$ 的体积 V_N 满足

$$\lim_{N \rightarrow \infty} V_N = 0 \quad (2.17)$$

(2) 相对区域 $R_N (N=1, 2, \dots)$ 取 N 个样本进行估计实验，设有 k_N 个样本落入 R_N 中，样本数目应满足：

$$\lim_{N \rightarrow \infty} k_N = \infty \quad (2.18)$$

$$\lim_{N \rightarrow \infty} \frac{k_N}{N} = 0 \quad (2.19)$$

则此估计序列处处收敛于 $p(x)$ 。

$$\hat{p}_N(x) = \frac{k_N/N}{V_N} \quad N=1, 2, \dots \quad (2.20)$$

条件(2.17)保证在给定区域均匀收缩且密度函数基于 x 连续的情况下，空间平均 P/V 能够收敛到 $p(x)$ 。条件(2.18)仅对 $p(x) \neq 0$ 的点才有意义，即当 $p(x) \neq 0$ 时，使 $\hat{p}(x) \neq 0$ ，可使频率在概率意义上收敛于概率 P 。条件(2.19)是式(2.20)收敛的必要条件，它描述了 N 的增长速度要大于 k_N 的增长速度，使 k_N/N 为无穷小，而 k_N/N 和 V_N 为同阶的无穷小，使得 $k_N/(NV_N)$ 为非无穷大的有界数，避免了 $\hat{p}(x) \rightarrow \infty$ 。

通常有两种办法获取一个区域序列满足条件(2.17)、(2.18)和(2.19)，一种是通过定义区域的体积为样本数的函数来收缩一个区域，比如定义 $V_N = 1/\sqrt{N}$ ，从而使 $p_N(x)$ 能够收敛到 $p(x)$ ，这就是 Parzen 窗法。另一种方法是定义 k_N 为样本 N 的函数，比如令 $k_N = \sqrt{N}$ ，这样体积 V_N 会逐步增长直到能包围 x 附近的 k_N 个样本，这就是 K 近邻估计方法。

2.3.2 Parzen 窗法

在 d 维特征空间中，取区域 R_N 是一个 d 维超立方体，其边长为 h_N ，则其体积

为 $V_N = h_N^d$ 。为了能用函数描述区域 R_N 和对落入 R_N 的样本计数，定义基本窗函数 $\varphi(\mathbf{x})$

$$\varphi(\mathbf{x}) = \begin{cases} 1 & \text{当 } |\mathbf{x}_i| \leq 0.5, i=1,2,\dots,d \\ 0 & \text{其他} \end{cases} \quad (2.21)$$

这样， $\varphi(\mathbf{x})$ 定义了一个基于起始点的单位超立方体，通过平移和尺度缩放可以改变超立方体的位置和大小。如果一个样本 \mathbf{x}_j 落入以 \mathbf{x} 为中心、 h_N 为边长的超立方体 R_N 内时，则 $\varphi((\mathbf{x}-\mathbf{x}_j)/h_N)$ 的值为 1，而其他的则为 0。由此，落入这个立方体内的样本数为：

$$k_N = \sum_{j=1}^N \varphi\left(\frac{\mathbf{x}-\mathbf{x}_j}{h_N}\right) \quad (2.22)$$

将上式带入(2.20)式，可得到 \mathbf{x} 点处概率密度估计

$$\hat{P}_N(\mathbf{x}) = \frac{1}{NV_N} \sum_{j=1}^N \varphi\left(\frac{\mathbf{x}-\mathbf{x}_j}{h_N}\right) = \frac{1}{Nh_N^d} \sum_{j=1}^N \varphi\left(\frac{\mathbf{x}-\mathbf{x}_j}{h_N}\right) \quad (2.23)$$

(2.23)式实际上是一个叠加函数。使用 k_N 个以样本 \mathbf{x}_j 为中心的窗函数叠加对 \mathbf{x} 处的概率密度进行估计。易知，样本较密集的区域上概率密度估计值较大。这种方法称为 Parzen 窗法。

(2.23)式给出了一个比较通用的概率密度估计方法。假定我们应用一个更为常用的窗口函数，那么式(2.23)暗含着密度函数的估计是一种对 \mathbf{x} 的函数平均。本质上，窗口函数可以认为是权值函数，即每个样本点根据它到 \mathbf{x} 的距离来对整个估计作出相应的贡献。函数 $\varphi(\mathbf{x})$ 常称为核函数，必须满足下列条件^[175]：

$$\begin{aligned} \sup_{\mathbf{x} \in R^d} |\varphi(\mathbf{x})| &< \infty \\ \int_{R^d} |\varphi(\mathbf{x})| d\mathbf{x} &< \infty \\ \lim_{\|\mathbf{x}\| \rightarrow \infty} \|\mathbf{x}\| \varphi(\mathbf{x}) &= 0 \\ \int_{R^d} \varphi(\mathbf{x}) d\mathbf{x} &= 1 \end{aligned} \quad (2.24)$$

其中 $\|\bullet\|$ 为欧氏范数。

2.3.3 Mean Shift 过程原理

给定点 $\mathbf{x}_i \in R^d, i=1,\dots,n$ ，点 \mathbf{x} 的多元核密度估计函数可以表示为^[168]：

$$\hat{f}(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^n K_H(\mathbf{x}-\mathbf{x}_i) \quad (2.25)$$

其中

$$K_H(\mathbf{x}) = |\mathbf{H}|^{-1/2} K(\mathbf{H}^{-1/2} \mathbf{x}) \quad (2.26)$$

\mathbf{H} 是 $d \times d$ 的正定对称矩阵, d 维的核函数 $K(\mathbf{x})$ 应满足如下条件:

$$\begin{aligned} \int_{\mathbb{R}^d} K(\mathbf{x}) d\mathbf{x} &= 1 \\ \int_{\mathbb{R}^d} \mathbf{x} K(\mathbf{x}) d\mathbf{x} &= \mathbf{0} \\ \lim_{\|\mathbf{x}\| \rightarrow \infty} \|\mathbf{x}\|^d K(\mathbf{x}) &= 0 \\ \int_{\mathbb{R}^d} \mathbf{x} \mathbf{x}^T K(\mathbf{x}) d\mathbf{x} &= c_K I \end{aligned} \quad (2.27)$$

其中 c_K 是常量。通常核函数有以下两种取法:

$$K(\mathbf{x}) = c_{k,d} \prod_{i=1}^d k(x_i) \quad (2.28)$$

$$K(\mathbf{x}) = c_{k,d} k(\|\mathbf{x}\|^2) \quad (2.29)$$

$c_{k,d}$ 是归一化常量。(2.28) 式中函数 $k(x_i)$ 为矢量 \mathbf{x} 的每一个分量 x_i 的函数, 而 (2.29) 式是矢量 \mathbf{x} 长度的函数。常用的核函数有:

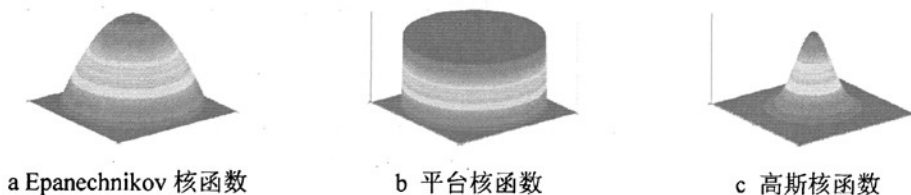


图 2-4 三种常用的核函数

其中:

Epanechnikov

核函数:

$$K_E(\mathbf{x}) = \begin{cases} c(1 - \|\mathbf{x}\|^2) & \|\mathbf{x}\| \leq 1 \\ 0 & \text{其他} \end{cases}$$

平台核函数:

$$K_U(\mathbf{x}) = \begin{cases} c & \|\mathbf{x}\| \leq 1 \\ 0 & \text{其他} \end{cases}$$

高斯核函数:

$$K_N(\mathbf{x}) = ce^{-\|\mathbf{x}\|^2/2}$$

为降低计算复杂度, 带宽矩阵 \mathbf{H} 选择只有一个参数 $h > 0$ 的 H_1 , 由此(2.25)式变为:

$$\hat{f}_{h,k}(\mathbf{x}) = \frac{c_{k,d}}{nh^d} \sum_{i=1}^n k\left(\left\|\frac{\mathbf{x} - \mathbf{x}_i}{h}\right\|^2\right) \quad (2.30)$$

特征空间分析实际上是找满足某种密度函数的模式，它们通常位于密度函数的梯度零值点，所以

$$\nabla f_{h,k}(\mathbf{x}) \equiv \nabla \hat{f}_{h,k}(\mathbf{x}) = \frac{2c_{k,d}}{nh^d} \sum_{i=1}^n (\mathbf{x} - \mathbf{x}_i) k' \left(\left\| \frac{\mathbf{x} - \mathbf{x}_i}{h} \right\|^2 \right) \quad (2.31)$$

定义核函数

$$G(\mathbf{x}) = c_{g,d} g \left(\left\| \frac{\mathbf{x} - \mathbf{x}_i}{h} \right\|^2 \right) \quad (2.32)$$

其中 $g(x) = -k'(x)$

将 $g(\mathbf{x})$ 代入(2.31)式，经整理得

$$\nabla \hat{f}_{h,k}(\mathbf{x}) = \frac{2c_{k,d}}{nh^d} \sum_{i=1}^n g \left(\left\| \frac{\mathbf{x} - \mathbf{x}_i}{h} \right\|^2 \right) \left[\frac{\sum_{i=1}^n \mathbf{x}_i g \left(\left\| \frac{\mathbf{x} - \mathbf{x}_i}{h} \right\|^2 \right)}{\sum_{i=1}^n g \left(\left\| \frac{\mathbf{x} - \mathbf{x}_i}{h} \right\|^2 \right)} - \mathbf{x} \right] \quad (2.33)$$

(2.33)式中等号右边的第一个求和部分正比于以 G 为核的密度估计函数：

$$\hat{f}_{h,G}(\mathbf{x}) = \frac{2c_{k,d}}{nh^d} \sum_{i=1}^n g \left(\left\| \frac{\mathbf{x} - \mathbf{x}_i}{h} \right\|^2 \right) \quad (2.34)$$

而第二部分就是 Mean Shift 矢量

$$m_{h,G}(\mathbf{x}) = \left[\frac{\sum_{i=1}^n \mathbf{x}_i g \left(\left\| \frac{\mathbf{x} - \mathbf{x}_i}{h} \right\|^2 \right)}{\sum_{i=1}^n g \left(\left\| \frac{\mathbf{x} - \mathbf{x}_i}{h} \right\|^2 \right)} - \mathbf{x} \right] \quad (2.35)$$

由(2.33)式、(2.34)式和(2.35)式，可以得到：

$$\nabla \hat{f}_{h,k}(\mathbf{x}) = \frac{2c_{k,d}}{nh^d} \cdot \hat{f}_{h,G}(\mathbf{x}) \cdot m_{h,G}(\mathbf{x}) \quad (2.36)$$

从而

$$m_{h,G}(\mathbf{x}) = \frac{nh^d}{2c_{k,d}} \cdot \frac{\nabla \hat{f}_{h,k}(\mathbf{x})}{\hat{f}_{h,G}(\mathbf{x})} \quad (2.37)$$

由(2.37)式可见，以 K 为核的密度函数梯度估计的零值点可以通过以 G 为核的 Mean Shift 矢量迭代得到。可以看出，Mean Shift 矢量始终指向概率密度函数最大的方向（即梯度方向），因为 Mean Shift 矢量最终要收敛为零，以 G 为核的概率密度函数估计 $\hat{f}_{h,G}(\mathbf{x})$ 在数据点稀疏的地方值比较小，在数据点密集的地方值比较大，因此当(2.37)式左端的 Mean Shift 矢量为零时，可以理解为右端的分母 $\hat{f}_{h,G}(\mathbf{x})$ 最大。

文献[168]给出了 Mean Shift 迭代过程的收敛性详细证明, 本文不再论述。

2.3.4 经典 Mean shift 跟踪算法

由 1.2.1.3 节可知, 视觉跟踪方法分为三类, 分别是基于区域的跟踪算法、基于特征的跟踪算法以及基于模型的跟踪算法。Mean Shift 跟踪算法本质上属于第二类, 即基于特征的跟踪算法。该类算法通常包含以下几步: 一是初始帧中目标特征建立, 在经典 Mean Shift 算法中用颜色(或灰度)直方图描述目标区域特征; 二是后续帧中目标候选特征提取; 三是通过相似性函数判断当前帧中候选区域是否为真实目标。下面结合以上三步, 对经典 Mean Shift 跟踪算法进行简要介绍。

2.3.4.1 计算初始帧的目标模板

$\{\mathbf{x}_i\}_{i=1,\dots,N}$ 为目标窗口中第 i 个像素的坐标, N 为目标窗口的像素点总数, 其中 \mathbf{x}_c 为目标窗口中心, m 为颜色直方图栅格总数。因此目标颜色直方图栅格特征向量中第 u 个分量为:

$$q_u = C \sum_{i=1}^N \left\{ k(\|\mathbf{x}_i - \mathbf{x}_c\|^2) \cdot \delta[b(\mathbf{x}_i) - u] \right\} \quad u \in \{1, \dots, m\} \quad (2.38)$$

式中 $k(\|\mathbf{x}\|^2)$ 是核函数, 而 C 为归一化常数, 定义为:

$$C = 1 / \sum_{i=1}^N [k(\|\mathbf{x}_i - \mathbf{x}_c\|^2)] \quad (2.39)$$

2.3.4.2 计算当前帧候选特征向量

$\{\mathbf{x}_i\}_{i=1,\dots,N_h}$ 为当前帧中搜索窗口中第 i 个像素的坐标, N_h 为搜索窗口中像素总数, 其中 \mathbf{y}_0 为搜索窗口中心, 类似(2.38)式, 计算当前帧中搜索窗口的颜色直方图栅格特征向量的第 u 个分量为:

$$p_u(\mathbf{y}_0) = C_h \sum_{i=1}^{N_h} \left\{ k\left(\left\|\frac{\mathbf{x}_i - \mathbf{y}_0}{h}\right\|^2\right) \cdot \delta[b(\mathbf{x}_i) - u] \right\} \quad u \in \{1, \dots, m\} \quad (2.40)$$

式中 \mathbf{x}_i 为当前帧搜索窗口中第 i 个像素, 归一化常数 C_h 的计算方法类似于 C , 而 h 为核函数的带宽。

$$C_h = 1 / \sum_{i=1}^{N_h} \left[k\left(\left\|\frac{\mathbf{x}_i - \mathbf{y}_0}{h}\right\|^2\right) \right] \quad (2.41)$$

2.3.4.3 相似性函数

用 Bhattacharyya 系数描述初始帧目标模板和当前帧候选向量的相似程度, 其定义为:

$$\rho(\mathbf{y}) \equiv \rho[\mathbf{p}(\mathbf{y}), \mathbf{q}] = \sum_{u=1}^m \sqrt{q_u \cdot p_u(\mathbf{y})} \quad (2.42)$$

式中 \mathbf{y} 表示当前帧候选模板的中心。(2.42)式的几何意义为 m 维单位向量 $(\sqrt{p_1}, \dots, \sqrt{p_m})$ 和 $(\sqrt{q_1}, \dots, \sqrt{q_m})$ 之间夹角的余弦值。所以当两个向量夹角为零时, 两个向量最为相似。

2.3.4.4 当前帧跟踪结果求解

为了使(2.42)式最大, 在当前帧中, 以前一帧跟踪结果作为当前帧搜索窗口的位置, 设 \mathbf{y}_0 为窗口中心为, 在 \mathbf{y}_0 邻域内寻找最优目标位置 \mathbf{y}_1 , 反复迭代, 最后得到在当前帧目标的最优位置。对(2.42)式在 $\mathbf{p}(\mathbf{y}_0)$ 处进行泰勒展开, 相似性函数可以近似表示为:

$$\rho[\mathbf{p}(\mathbf{y}), \mathbf{q}] \approx \frac{1}{2} \sum_{u=1}^m \sqrt{q_u \cdot p_u(\mathbf{y}_0)} + \frac{C_h}{2} \sum_{u=1}^{N_h} \left[\omega_i k \left(\left\| \frac{\mathbf{y} - \mathbf{x}_i}{h} \right\|^2 \right) \right] \quad (2.43)$$

$$\text{其中} \quad \omega_i = \sum_{u=1}^m \left\{ \delta[b(\mathbf{x}) - u] \sqrt{q_u / p_u(\mathbf{y}_0)} \right\} \quad (2.44)$$

通过对相似性函数求最大值, 根据 Mean shift 向量迭代得到当前帧目标的新位置:

$$\mathbf{y}_{j+1} = \frac{\sum_{u=1}^{N_h} \left[\mathbf{x}_i \omega_i g \left(\left\| \frac{\mathbf{y}_j - \mathbf{x}_i}{h} \right\|^2 \right) \right]}{\sum_{u=1}^{N_h} \left[\omega_i g \left(\left\| \frac{\mathbf{y}_j - \mathbf{x}_i}{h} \right\|^2 \right) \right]} \quad (2.45)$$

2.3.5 经典 Mean Shift 跟踪算法实现步骤

(1) 对初始帧图像, 利用(2.38)式统计目标特征描述向量, 将初始帧目标位置作为下一帧中目标初始位置 \mathbf{y}_0 ;

(2) 对于当前帧, 以 \mathbf{y}_j 为中心、带宽参数 h 为半径选择一个圆形区域, 利用(2.40)式统计当前帧的候选特征向量;

(3) 根据(2.44)式计算 Mean Shift 采用权值 ω_i ;

(4) 利用(2.45)式得到当前帧目标的位置 \mathbf{y}_{j+1} ;

(5) 如果 $\|\mathbf{y}_{j+1} - \mathbf{y}_j\| < \varepsilon$ (ε 为预先选择的阈值), 则停止迭代, 输出当前帧目标位置 \mathbf{y}_{j+1} ; 否则令 $\mathbf{y}_j = \mathbf{y}_{j+1}$, 返回第二步。

2.3.6 核密度估计的带宽选择

带宽矩阵 \mathbf{H} 是核密度估计的关键参数。选择带宽矩阵要根据所要估计概率密度的复杂度以及参数的数量综合考虑，通常带宽矩阵有 $p(p+1)/2$ 个值。典型的带宽矩阵选择方法有三种：

1. $\mathbf{H}_1 = h^2 I_p$ 。对每一维设定同样的平滑参数 h ，这意味着在每个方向上平滑程度一样，采用 \mathbf{H}_1 作为带宽矩阵的核具有椭圆对称性，只有一个平滑参数，这种情况下要对每一维的矢量作归一化处理，使每一维具有相同的尺度。

2. $\mathbf{H}_2 = \text{diag}(h_1^2, \dots, h_p^2)$ 。对每一坐标维上设置不同的平衡度，如果对每一维作了尺度处理，令 s_j 为尺度常量，则该带宽矩阵 \mathbf{H}_2 也可以表示为 $h^2 \text{diag}(s_1, \dots, s_p)$ ，因为每一维可具有不同的平衡度，因而有一定的改进和灵活性。

3. $\mathbf{H}_3 = h^2 S$ 。这里 S 是 \mathbf{x} 的协方差矩阵估计，其思想是采用与概率密度具有相同形状的核函数。该形式是多元核密度估计一般意义上的带宽矩阵形式，相当于把数据点通过采用 \mathbf{H}_1 作为带宽矩阵变换到单位化的协方差矩阵估计，然后再变换回原来的尺度。

2.4 本章小结

本章主要介绍了论文涉及到的基本理论，其中首先介绍了尺度空间基本理论，阐述了 Linderberg 特征尺度选择理论以及尺度不变性理论。然后介绍了 Gabor 小波理论的生物学背景及其具体数学表达形式。本章着重介绍了 Mean Shift 理论，详细给出了非参数估计的具体推导，引出了 Parzen 窗密度估计方法，具体讲解了 Mean Shift 过程的基本原理和过程，详细介绍了经典 Mean Shift 跟踪算法，给出了核函数带宽选择方法。

第三章 基于 Gabor 梯度的角点检测算法研究

战场三维重建是未来战场态势分析的重要辅助手段之一。而图像配准是实现场景三维重建的关键技术之一。常用的一类图像配准算法需要提取图像特征（如点特征、线特征或面特征等），对每个特征建立特征描述，通过分析特征描述的相似性建立特征之间的对应关系，由对应的特征集解算出图像对之间的变换模型参数，从而达到配准图像的目的。该类配准算法的性能主要取决于特征提取和特征描述建立两个环节。因此，研究图像特征提取对于该类算法具有重要的意义。

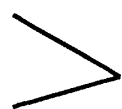
无人机图像是指无人机上的摄像机对其它景物拍摄的图像，按其物理属性可分为自然地物特征图像和人工地物特征图像两类。其中自然地物特征主要包括山脉、河流和海岸线等，而人工地物特征主要包括道路、田地、桥梁和各类人工建筑物等。两类地物特征图像中蕴含着大量角点特征，为无人机图像配准及场景三维重建提供了有利的条件。

本章围绕图像角点特征提取算法进行研究，具体内容安排如下：3.1 节给出了图像特征分类及角点特征的检测准则；3.2 节介绍了常用角点检测算法，分析各个算法的优缺点；3.3 节简要介绍了 Gabor 梯度算子原理并对其性能进行了分析；3.4 节提出了基于 Gabor 梯度的角点检测算法；3.5 节给出了新方法 with 经典角点提取算法的检测结果及分析；3.6 节对本章进行了小结。

3.1 角点特征分类及检测准则

图像特征是物体的物理和几何特性经过光学成像后的反映，可能对应着三维连续景象中的空间不连续点、不同反射系数（与物体的材料和颜色、光线入射和反射的角度有关）物体之间的分界线、物体与背景的分界线以及阴影引起的边界线等。图像特征通常表现为在某邻域内的灰度突变，而不是一个像素点上的灰度突变。

图像特征按其几何特性可分为角点特征、线特征和区域特征三类，其中角点特征是图像的重要特征之一。该类特征为目标识别、立体匹配等应用提供了重要的信息。为了实现角点特征的可靠提取，需要了解角点特征的分类，针对不同类型的角点设计算法进行检测。Guzman^[176]利用形成角点的线段相交特性来对角点特征进行系统的分类，如图 3-1 所示。



(1) “L”型



(2) “T”型



(3) “Y”型

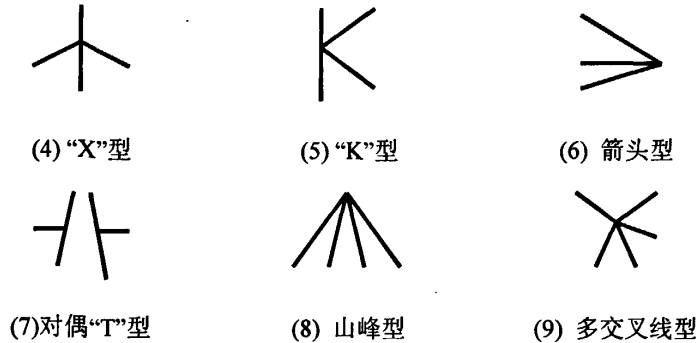


图 3-1 图像角点特征分类

近些年来，各种各样的角点检测算法层出不穷，但是不论算法的原理差异有多大，均应该满足以下四个准则：

- 1、角点检测性：可以检测到微弱的角点；
- 2、定位准确性：和真实位置接近；
- 3、稳定性：在不同的几何变形或者不同光照条件下能检测出同一角点；
- 4、鲁棒性：对于噪声不敏感。

3.2 常用角点特征检测算法

国内外学者很早就对角点检测技术进行了系统研究，提出了许多具有实际意义的算法。按算法原理不同，角点检测算法可以分为基于边缘链码分析、基于模板匹配和基于区域特征属性分析三类。

第一类算法首先提取边缘特征，然后将边缘用链码表示，通过找寻高曲率点以达到提取角点特征的目的。这类算法一般分为三个步骤：边缘提取、链码分析和角点检测。该类算法很大程度上依赖于边缘特征的提取，而边缘提取本身又是图像处理的一个难题，因此这类算法在摄影测量中应用较少。

第二类算法通过定义多种角点模板，计算模板与每个图像子窗口的相似性，从而判断出图像中是否存在与模板相对应的角点。这类方法虽然检测精度比较高，但需要进行前期预测分析，且计算量较大。

第三类算法利用角点邻域的特征属性（例如灰度和梯度信息）提取角点。该类算法中应用最为广泛的是 Moravec 算法^[177]、Harris 算法^[51]、Forstner 算法^[178]和 SUSAN 算法^[53]。下面对这四种经典角点检测算法进行简要介绍。

3.2.1 Moravec 角点检测算法

Moravec 于 1977 年提出利用灰度方差提取角点特征算法，其步骤为：

(1) 计算各像素的兴趣值 (Interest Value, 简称 IV)。在以像素 (c, r) 为中心的 $w \times w$ 的局部邻域内 (如 5×5 的窗口), 按(3.1)式计算四个方向相邻像素灰度差的平方和;

$$\begin{aligned} V_1 &= \sum_{i=-k}^k (I_{c+i,r} - I_{c+i+1,r})^2 \\ V_2 &= \sum_{i=-k}^k (I_{c+i,r+i} - I_{c+i+1,r+i+1})^2 \\ V_3 &= \sum_{i=-k}^k (I_{c,r+i} - I_{c,r+i+1})^2 \\ V_4 &= \sum_{i=-k}^k (I_{c+i,r-i} - I_{c+i+1,r-i-1})^2 \end{aligned} \quad (3.1)$$

其中 $k = INT(w/2)$ 。取其中最小者作为该像素 (c, r) 的兴趣值:

$$IV = \min\{V_1, V_2, V_3, V_4\} \quad (3.2)$$

(2) 给定经验阈值, 将兴趣值大于该阈值的点 (即兴趣值计算窗口的中心) 作为候选点。阈值的选择应以候选点中包括所需要的角点, 而又不含过多的非角点为原则;

(3) 选取候选点中的极值点作为角点。在一定大小的窗口内 (可不同于兴趣值计算窗口, 例如 5×5 , 7×7 , 9×9), 将候选点中兴趣值不是最大者均去掉, 仅留下一个兴趣值最大者, 该像素即为一个角点, 此步骤常称为“非最大化抑制”。

Moravec 算法的性能评价如下:

- (1) 由于利用了四个方向实现局部相关, 所以 IV 是各向异性的;
- (2) 算子实现简单快速。

3.2.2 Harris 角点检测算法

Harris 和 Stephen 于 1988 年提出了著名的 Harris 角点检测算法。其基本思想与 Moravec 兴趣点检测算子相似, 但 Harris 算法是由图像的一阶导数估算出自相关度量, 利用自相关度量建立角点响应函数, 从而达到检测角点的目的。

Harris 算法检测思想为: 对于图像上任意一点, 以该点为中心选取一个小窗口, 在任意方向移动该小窗口, 观察该窗口内灰度变化。图 3-2 中浅灰色三角形表示平坦平面区域, 深灰色正方形表示选取的小窗口, 箭头表示移动方向。对于平坦区域内的像素点, 在任意方向移动窗口, 窗口内灰度变化均很小, 如图 3-2.a; 对于边界上的像素点, 沿着边界方向移动窗口, 窗口内灰度变化很小, 如图 3-2.b; 对于角点, 在任意方向移动窗口, 窗口内灰度均会有显著变化, 如图 3-2.c。

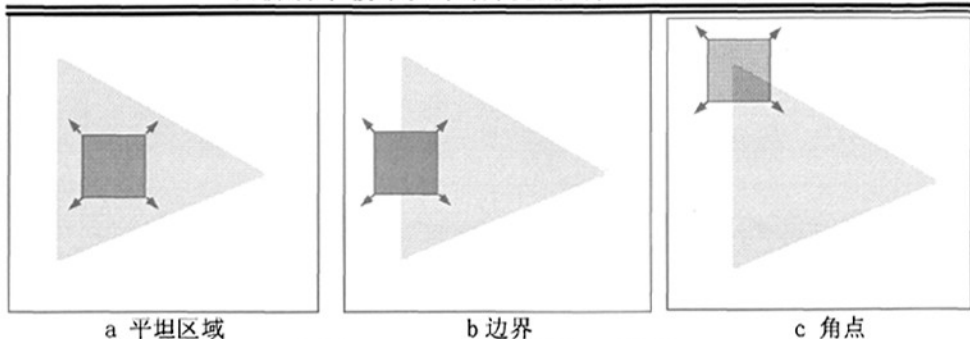


图 3-2 Harris 角点检测原理示意图

移动窗口产生的灰度变化表示为

$$E(u, v) = \sum_{x, y} w(x, y) [I(x+u, y+v) - I(x, y)]^2 \quad (3.3)$$

其中 $w(x, y)$ 为权值函数 (如高斯函数), $I(x+u, y+v)$ 为窗口移动 (u, v) 后的灰度, $I(x, y)$ 为窗口移动前的灰度。对(3.3)式进行线性展开可以得到:

$$E(u, v) = [u, v] \mathbf{M} \begin{bmatrix} u \\ v \end{bmatrix} \quad (3.4)$$

\mathbf{M} 是一个 2×2 的矩阵, 计算方法如下式:

$$\mathbf{M} = \sum_{x, y} w(x, y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \quad (3.5)$$

其中 I_x 与 I_y 分别为图像 $I(x, y)$ 关于 x 方向和 y 方向的梯度。

设 λ_1 、 λ_2 为矩阵 \mathbf{M} 的特征值, 则 λ_1 、 λ_2 与主曲率成比例。为了使 Harris 角点具有旋转不变性, 利用 λ_1 、 λ_2 构造与旋转无关的角点响应函数 (Corner Response Function, 简称 CRF), 其定义为:

$$CRF = \det \mathbf{M} - k(\text{trace} \mathbf{M})^2 \quad (3.6)$$

其中 $\det \mathbf{M}$ 是矩阵 \mathbf{M} 的行列式, $\text{trace} \mathbf{M}$ 为 \mathbf{M} 的迹。 k 是经验常数, 通常取为 $0.04 \sim 0.06$ 。当且仅当所有方向的曲率都大于某一设定阈值时, 才表明角点存在, 此时任一方向的偏移都会对 $E(u, v)$ 产生极大的影响。

Harris 角点检测算子具有如下特性:

- (1) 角点检测的可靠性高;
- (2) 需要确定阈值、常量 k 和高斯函数的方差值。

3.2.3 Forstner 角点检测算子

Forstner 算子是一个高精度角点定位算子, 它通过计算一个窗口 (如 5×5) 内每一个像素灰度的 Roberts 梯度, 将得到的灰度梯度构成协方差矩阵 \mathbf{N} , 找出灰度

变化较大、且灰度的误差椭圆尽可能接近圆的点作为角点。协方差矩阵 \mathbf{N} 由(3.7)式获得。

$$\mathbf{N} = \begin{bmatrix} \sum I_u^2 & \sum I_u I_v \\ \sum I_u I_v & \sum I_v^2 \end{bmatrix} \quad (3.7)$$

其中 $I_u = I_{i+1,j+1} - I_{i,j}$, $I_v = I_{i+1,j} - I_{i,j+1}$ 为 Roberts 梯度。

Forstner 角点检测算法分为最佳窗口选择与最佳窗口内加权重心化两步。对每一个图像窗口计算其兴趣值 q 和 w ：

$$q = \frac{4 \det \mathbf{N}}{(\text{trace} \mathbf{N})^2} \quad (3.8)$$

$$w = \frac{\det \mathbf{N}}{\text{trace} \mathbf{N}} \quad (3.9)$$

其中 $\det \mathbf{N}$ 是协方差矩阵 \mathbf{N} 的行列式, $\text{trace} \mathbf{N}$ 为 \mathbf{N} 的迹。当 q 和 w 均大于给定的阈值且为局部区域内的极值时, 该窗口为最佳窗口。窗口内像素的加权灰度重心为角点位置。

Forstner 算法将灰度变化较大、灰度误差椭圆尽可能圆且为窗口内像素的加权灰度重心作为角点, 所以该算法对外界光照变化不是十分敏感, 稳定性较好。

3.2.4 SUSAN 角点算子

Smith 和 Brady 提出了一种 SUSAN 角点检测算法。图 3.3 为一个灰色矩形位于白色背景上, 选取圆模板 e 置于图上四个不同的位置。

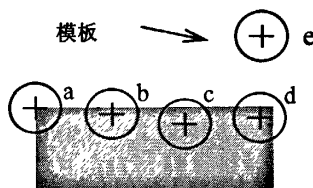


图 3.3 四个不同位置的圆模板

将模板中的各点与核心点（当前点）的灰度值用下面的相似比较函数来进行比较。

$$c(x,y) = \begin{cases} 1 & |I(x,y) - I(x_0,y_0)| \leq t \\ 0 & |I(x,y) - I(x_0,y_0)| > t \end{cases} \quad (3.10)$$

其中 $I(x_0,y_0)$ 和 $I(x,y)$ 分别是核心点和模板中其它点的灰度值, t 是区分特征目标与背景的阈值。阈值 t 的选取要根据图像中目标与背景的对比程度来确定。

由于图像中同一特征区域的内部特征（通常指灰度值）是一致的或相近的, 因此对单个像素点来讲, 其特征只与其周围局部区域灰度相近的点有关。

用相似比较函数对模板中的各点与核心点的灰度值进行比较,得到的与核心点灰度相近的点的集合称为核值相似区 (Univalue Segment Assimilating Nucleus, 简称 USAN 区域)。图 3.4 中模板中的白色区域即为 USAN 区域。由图 3.4 可以看出,在平坦区域 USAN 区最大,如模板 e 所示;在边缘处 USAN 区大小降为一半,如模板 b 所示;而在角点附近 USAN 区会变得更小,如模板 a 所示。由此可以得到 SUSAN 检测角点算法的基本原理:即在角点处的 USAN 区最小。因此可以根据 USAN 区的大小和矩特性来检测角点特征。SUSAN 角点检测算法与基于梯度信息的角点检测算法的不同之处在于不需要对图像求导,所以抗噪声能力较强。

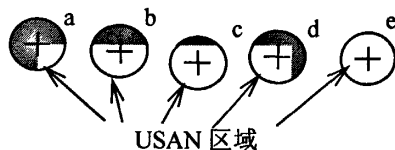


图 3.4 USAN 区域的显示

3.3 Gabor 梯度算子原理及分析

近些年,国内外学者提出了许多角点检测算法,其中非常重要的一类算法以图像 x 方向和 y 方向梯度信息为基础,如 Harris 算法和 Forstner 算法等。梯度信息通常由 Sobel 算子、Roberts 算子或 Prewitt 算子等计算得到。当图像噪声严重时,由这些算子计算得到的梯度信息存在较大误差,从而影响角点提取算法的检测性能。因此,需要找到一种抗噪性更好的梯度算子。

边界方向角的正切值定义为 y 方向梯度与 x 方向梯度的比值。待处理的图像为数字化图像,对于水平、垂直以及正负 45° 方向的边界, Sobel、Roberts 和 Prewitt 等传统梯度算子在求取直线边界方向角的正切值上误差较小。而对于这四个方向以外边界的 y 方向梯度和 x 方向梯度计算存在较大误差。因此,需要一种比传统梯度算子更为准确的梯度算子。

通过以上分析,研究抗噪性好、计算准确的梯度算子对于提高基于梯度信息角点提取算法性能具有非常重要的意义。针对该问题,本文提出了一种基于 Gabor 小波的梯度算子,下面简要介绍该算子,并对该算子与传统算子的检测性能进行分析与比较。

3.3.1 Gabor 梯度算子定义

Gabor 奇函数是 Gabor 函数的虚部,其数学表达形式为:

$$g(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)\right] \sin[\omega_0(x\cos\theta + y\sin\theta)] \quad (3.11)$$

由 Gabor 奇函数数学表达式可以看出, Gabor 奇函数是由高斯函数通过 \sin 函数调制得到。在应用 Gabor 奇函数时, 通常令 $\sigma_x = \sigma_y = \sigma$, 则(3.11)式可表示为:

$$g(x, y, \sigma, \omega_0, \theta) = \frac{1}{2\pi\sigma^2} \exp\left[-\frac{1}{2}\left(\frac{x^2 + y^2}{\sigma^2}\right)\right] \sin[\omega_0(x\cos\theta + y\sin\theta)] \quad (3.12)$$

当 $\sigma\omega_0 \cong 1$ 时, Gabor 奇函数的边缘检测性能最好^[179], 其检测性能要优于传统的简单边缘检测算子(如 Sobel、Roberts 和 Prewitt 等), 与被广泛使用的 Canny 边缘检测算子相比, 具有相近的检测性能。在边缘检测中, 为了抑制噪声, Canny 边缘检测算子采用高斯函数来实现低通滤波, 而对于 Gabor 奇函数, 则有

$$g(x, y, \sigma, \omega_0, \theta) * f(x, y, \sigma) = c \cdot g_1\left(x, y, \sqrt{2}\sigma, \frac{\omega_0}{2}, \theta\right) \quad (3.13)$$

其中 c 为一常数, $f(x, y, \sigma)$ 表示参数为 σ 的高斯函数。(3.13)式的证明可参见文献[179]。可以看出, 由参数 $\sqrt{2}\sigma$ 的 Gabor 奇函数可以分解为参数为 σ 的 Gabor 奇函数与参数为 σ 的高斯函数的卷积, 因而 Gabor 奇函数具有抑制噪声的能力。

令 $\sigma\omega_0 \cong 1$, 通过尺度、平移和旋转参数的变化可以将公式(3.12)表示的 Gabor 奇函数扩展为一组小波^[180]。

$$\begin{aligned} \psi\left(x, y, c_x, c_y, \theta, s\sigma, \frac{\omega_0}{s}\right) &= \frac{1}{2\pi(s\sigma)^2} \exp\left\{-\frac{1}{2(s\sigma)^2} \left[((x-c_x)\cos\theta - (y-c_y)\sin\theta)^2 \right. \right. \\ &\quad \left. \left. + ((x-c_x)\sin\theta + (y-c_y)\cos\theta)^2 \right] \right\} \times \sin\left\{ \frac{\omega_0}{s} [(x-c_x)\cos\theta + (y-c_y)\sin\theta] \right\} \end{aligned} \quad (3.14)$$

图 3.5 表示 $\theta = 0$ 、 $s = 2$ 、 $\sigma = 4$ 、 $\omega_0 = 0.125$ 的大小为 80×80 的 Gabor 奇函数三维示意图。

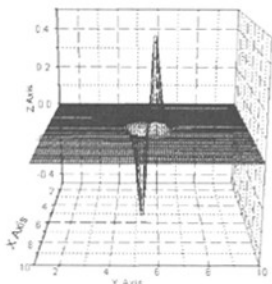


图 3.5 Gabor 模板三维示意图

根据(3.14)式, 令 $\sigma_1 = s\sigma$ 、 $\omega_1 = \omega_0/s$, 可以定义 x 方向和 y 方向 Gabor 梯度算子 (Gabor Gradient Operator, 简称 GGO) 分别为(3.15)式和(3.16)式。

$$G_x = \psi(x, y, c_x, c_y, 90, \sigma_1, \omega_1) \quad (3.15)$$

$$G_y = \psi(x, y, c_x, c_y, 0, \sigma_1, \omega_1) \quad (3.16)$$

3.3.2 Gabor 梯度算子边缘方向检测结果及分析

实验中, 选取 5 度、10 度、15 度、20 度、25 度、30 度、35 度和 40 度八个理想边缘进行计算。对于每个理想边缘图像, 为了消除随机因素的影响, 在每幅图像中心分别选取一个区域, 对区域中每个边缘点利用 Gabor 梯度算子、Sobel 算子和 Prewitt 算子计算其直线边界方向角的正切值, 最后将三种方法每点计算结果的平均值进行比较。图 3-6.a 和图 3-6.b 分别为 15 度与 30 度理想边缘, 其中灰色矩形为选择的计算区域。



图 3-6 理想边缘图像

为了比较三种算子在不同噪声下抗噪性, 对每幅理想边缘图像均分别加入方差为 10%、20%、30%、40%和 50%的高斯噪声。为了消除随机因素的影响, 对于每个级别的噪声分别加载 100 次, 计算 100 幅该级别噪声图像统计结果的平均值作为最终计算结果。具体结果如表 3-1~表 3-8 和图 3-7~图 3-14, 其中表 3-1~表 3-8 为三种算子边缘方向计算结果, 图 3-7~图 3-14 为三种算法理想边缘不同方差高斯噪声下与真实值差的绝对值比较。从计算结果可以看出, 对于八个理想边缘, GGO 算子在计算准确性和抗噪性上要明显优于比 Sobel 算子和 Prewitt 算子, 而 Prewitt 算子要略优于 Sobel 算子。

表 3-1 5 度边缘方向计算结果 (真实值 0.088)

| 高斯噪声方差 | Sobel | Prewitt | GGO |
|--------|-------|---------|-------|
| 10 | 0.180 | 0.173 | 0.089 |
| 20 | 0.195 | 0.183 | 0.093 |
| 30 | 0.220 | 0.198 | 0.101 |
| 40 | 0.248 | 0.212 | 0.116 |
| 50 | 0.256 | 0.238 | 0.135 |

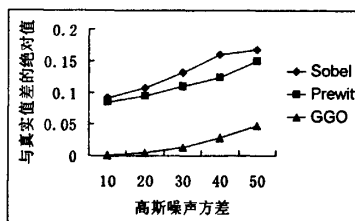


图 3-7 5 度理想边缘

表 3-2 10 度边缘方向计算结果(真实值 0.177)

| 高斯噪声方差 | Sobel | Prewitt | GGO |
|--------|-------|---------|-------|
| 10 | 0.248 | 0.248 | 0.178 |
| 20 | 0.254 | 0.253 | 0.180 |
| 30 | 0.263 | 0.260 | 0.182 |
| 40 | 0.273 | 0.268 | 0.185 |
| 50 | 0.291 | 0.278 | 0.189 |

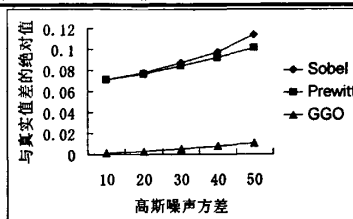


图 3-8 10 度理想边缘

表 3-3 15 度边缘方向计算结果(真实值 0.268)

| 高斯噪声方差 | Sobel | Prewitt | GGO |
|--------|-------|---------|-------|
| 10 | 0.348 | 0.347 | 0.266 |
| 20 | 0.355 | 0.352 | 0.263 |
| 30 | 0.361 | 0.357 | 0.275 |
| 40 | 0.375 | 0.365 | 0.279 |
| 50 | 0.387 | 0.377 | 0.286 |

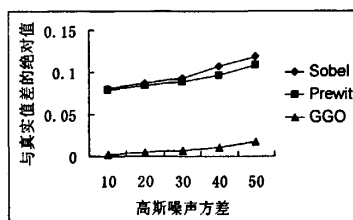


图 3-9 15 度理想边缘

表 3-4 20 度边缘方向计算结果(真实值 0.364)

| 高斯噪声方差 | Sobel | Prewitt | GGO |
|--------|-------|---------|-------|
| 10 | 0.411 | 0.412 | 0.362 |
| 20 | 0.415 | 0.415 | 0.361 |
| 30 | 0.422 | 0.419 | 0.360 |
| 40 | 0.437 | 0.429 | 0.342 |
| 50 | 0.450 | 0.436 | 0.377 |

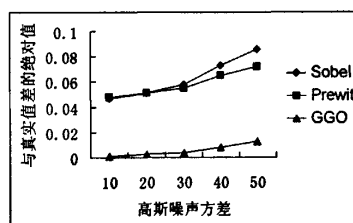


图 3-10 20 度理想边缘

表 3-5 25 度边缘方向计算结果(真实值 0.466)

| 高斯噪声方差 | Sobel | Prewitt | GGO |
|--------|-------|---------|-------|
| 10 | 0.523 | 0.522 | 0.466 |
| 20 | 0.527 | 0.525 | 0.468 |
| 30 | 0.536 | 0.528 | 0.450 |
| 40 | 0.558 | 0.534 | 0.451 |
| 50 | 0.568 | 0.545 | 0.453 |

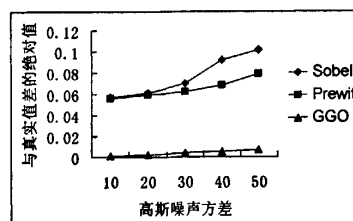


图 3-11 25 度理想边缘

表 3-6 30 度边缘方向计算结果(真实值 0.577)

| 高斯噪声方差 | Sobel | Prewitt | GGO |
|--------|-------|---------|-------|
| 10 | 0.622 | 0.623 | 0.578 |
| 20 | 0.626 | 0.625 | 0.579 |
| 30 | 0.634 | 0.628 | 0.581 |
| 40 | 0.649 | 0.633 | 0.582 |
| 50 | 0.693 | 0.642 | 0.586 |

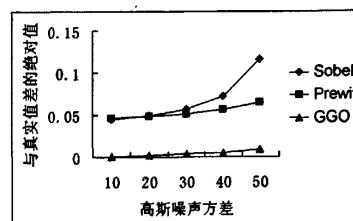


图 3-12 30 度理想边缘

表 3-7 35 度边缘方向计算结果(真实值 0.700)

| 高斯噪声方差 | Sobel | Prewitt | GGO |
|--------|-------|---------|-------|
| 10 | 0.749 | 0.747 | 0.699 |
| 20 | 0.750 | 0.748 | 0.704 |
| 30 | 0.758 | 0.752 | 0.694 |
| 40 | 0.788 | 0.758 | 0.708 |
| 50 | 0.828 | 0.774 | 0.717 |

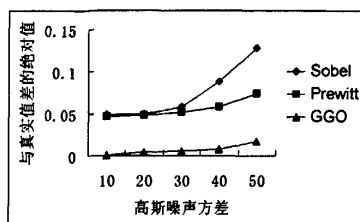


图 3-13 35 度理想边缘

表 3-8 40 度边缘方向计算结果(真实值 0.839)

| 高斯噪声方差 | Sobel | Prewitt | GGO |
|--------|-------|---------|-------|
| 10 | 0.878 | 0.876 | 0.840 |
| 20 | 0.882 | 0.879 | 0.841 |
| 30 | 0.891 | 0.883 | 0.843 |
| 40 | 0.917 | 0.895 | 0.846 |
| 50 | 0.938 | 0.908 | 0.848 |

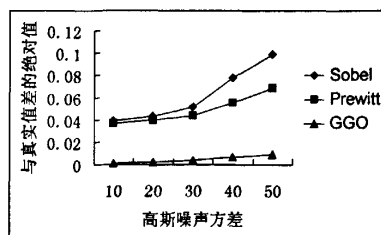


图 3-14 40 度理想边缘

3.4 基于 Gabor 梯度的角点检测算法原理

角点检测中常用的一类算法以梯度信息为基础，而传统梯度算子对噪声较为敏感，并且在边缘方向求取上存在一定的偏差，当图像上存在噪声时，这种偏差更为明显。从上一节可知，Gabor 梯度算子 G_x 与 G_y 在抗噪性和准确性上优于传统梯度算子。在角点检测中，如果将 G_x 和 G_y 代替 I_x 和 I_y ，则可以在一定程度上提高基于梯度信息的角点检测算法的准确性和抗噪性。

本文提出了一种基于 Gabor 梯度算子的角点检测算法（Corner Detecting based on Gabor Gradient Operator，简称 CDGGO）。实验发现 Gabor 梯度算子选取不同参数时，计算精度和噪声抑制上会有不同的结果。当 Gabor 梯度算子选择大尺度参数时（即 σ_1 较大），对于噪声有较好的抑制作用，但是容易将一些真正的角点剔除。而当 Gabor 梯度算子选择小尺度参数时（即 σ_1 较小），定位比较准确，但是容易产生虚假角点。因此，CDGGO 算法综合考虑了这两个特性进行角点提取。

3.4.1 算法原理

3.4.1.1 初始候选点检测

对于图像上任意一点 (i, j) ，以该点为中心选取一个小窗口，在任意方向移动小窗口，观察小窗口内灰度变化。由 Harris 角点检测算法思想可知：对于角点，在任意方向移动窗口，窗口内灰度均会有显著变化。移动窗口产生的灰度变化表示为：

$$E(u, v) = \sum_{x, y} w(x, y) [I(x+u, y+v) - I(x, y)]^2 \quad (3.17)$$

其中 $I(x+u, y+v)$ 为窗口移动 (u, v) 后的灰度, $I(x, y)$ 为窗口移动前的灰度, $w(x, y)$ 为高斯权值。对(3.17)式进行线性展开可以得到:

$$E(u, v) = [u, v] \mathbf{M} \begin{bmatrix} u \\ v \end{bmatrix} \quad (3.18)$$

其中 \mathbf{M} 是一个 2×2 的矩阵, 具体计算方法如下式:

$$\mathbf{M} = \sum_{x, y} w(x, y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \quad (3.19)$$

设 λ_1 、 λ_2 为矩阵 \mathbf{M} 特征值, 则 λ_1 、 λ_2 与主曲率成比例。由 3.3 节分析可知, 传统梯度算子在梯度计算上存在误差, 而 Gabor 梯度算子在梯度计算误差和抗噪性上优于传统梯度算子。所以可以用 x 方向和 y 方向 Gabor 梯度算子, 分别代替(3.19)式中的 I_x 和 I_y , 得到新的矩阵 \mathbf{M}_1

$$\mathbf{M}_1 = \sum_{x, y} w(x, y) \begin{bmatrix} G_x^2 & G_x G_y \\ G_x G_y & G_y^2 \end{bmatrix} \quad (3.20)$$

由(3.13)式可知, 高斯函数 $w(x, y)$ 和 Gabor 奇函数相乘, 等价于一个新的 Gabor 奇函数, 所以可以得到

$$\mathbf{M}_1 = \sum_{x, y} \begin{bmatrix} \hat{G}_x^2 & \hat{G}_x \hat{G}_y \\ \hat{G}_x \hat{G}_y & \hat{G}_y^2 \end{bmatrix} \quad (3.21)$$

其中 \hat{G} 表示一个新的 Gabor 奇函数, \hat{G}_x 和 \hat{G}_y 为 \hat{G} 对应的 Gabor 梯度算子。如果预先定义好 \hat{G} , 则可以不用再计算高斯权值, 从而可以减少计算量。 \mathbf{M}_1 的特征值用 $\lambda_{i,j}^0$ 和 $\lambda_{i,j}^{90}$ 表示。根据 Harris 角点检测算法可知, 对于角点特征, $\lambda_{i,j}^0$ 和 $\lambda_{i,j}^{90}$ 应都比较大, 所以构造点 (i, j) 的响应函数为:

$$Qn_{i,j}^1 = \frac{4.0 \times \lambda_{i,j}^0 \times \lambda_{i,j}^{90}}{(\lambda_{i,j}^0 + \lambda_{i,j}^{90})^2} \quad (3.22)$$

由文献[178]可知, $Qn_{i,j}^1$ 为像素 (i, j) 对应的误差椭圆的圆度:

$$Qn_{i,j}^1 = 1 - \frac{(a^2 - b^2)^2}{(a^2 + b^2)^2} \quad (3.23)$$

其中 a 和 b 分别为椭圆的长、短半轴。如果 a 、 b 中任意一个为零, 则 $Qn_{i,j}^1 = 0$, 表明该点可能位于边缘上, 如果 $a = b$, 则 $Qn_{i,j}^1 = 1$, 表明该点为角点。

对于角点, 由 Harris 角点检测算法思想可知, 在任意方向移动窗口, 窗口内

灰度均会有显著变化。为了提高角点检测算法的准确性，增加两个方向再构造一个响应函数。构造思路为：假设将原始图像旋转 45 度后，以该点为中心选取一个小窗口，在任意方向移动小窗口，窗口内灰度仍然会有显著变化。图像旋转后计算 x 方向和 y 方向梯度，相当于原图上计算 45 度和 135 度的梯度。所以类似 0 度和 90 度的做法，将 45 度与 135 度 Gabor 梯度 $g_{i,j}^{45}$ 与 $g_{i,j}^{135}$ ，分别代替(3.19)式中的 I_x 和 I_y ，得到新的矩阵 M_2 ，其特征值用 $\lambda_{i,j}^{45}$ 和 $\lambda_{i,j}^{135}$ 表示。类似(3.22)式构造点 (i, j) 的响应函数为：

$$Qn_{i,j}^2 = \frac{4.0 \times \lambda_{i,j}^{45} \times \lambda_{i,j}^{135}}{(\lambda_{i,j}^{45} + \lambda_{i,j}^{135})^2} \quad (3.24)$$

综合考虑以上分析，由(3.22)式和(3.24)式定义 CDGGO 算法角点响应函数为：

$$CRF = Qn_{i,j}^1 \times Qn_{i,j}^2 \quad (3.25)$$

保留 CRF 大于指定阈值且为局部最大的点作为角点候选点。

3.4.1.2 虚假点剔除

由于噪声等因素的影响，经上面方法提取的候选点中可能存在虚假点。实验发现，绝大多数虚假点位于图像中较为平坦区域上，而真正的角点应出现在图像梯度变化较为剧烈的地方。对于图像上像素点 (i, j) ，用梯度模表示该点的梯度强度。定义基于 Gabor 梯度算子的梯度模 $Gm_{i,j}$ 为：

$$Gm_{i,j} = \sqrt{g_{i,j}^0 \times g_{i,j}^0 + g_{i,j}^{90} \times g_{i,j}^{90}} \quad (3.26)$$

以像素点 (x, y) 为中心，选取大小为 $m \times m$ 局部区域，该区域内的平均 Gabor 梯度模 Gm_{mean} 定义为：

$$Gm_{mean} = \frac{\sum_{x=0}^m \sum_{y=0}^m Gm_{x,y}}{m^2} \quad (3.27)$$

图 3-15 中黑点表示当前点 (i, j) ，黑色正方形表示以该点为中心选择的局部区域，灰色三角形表示图像上平坦区域。

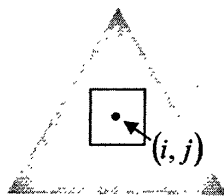


图 3-15 虚假点位置示意图

当正方形区域所有像素 (x, y) 均位于同一平坦区域内时，正方形区域所有像素

的梯度模一致，所以 $Gm_{i,j}$ 近似等于 Gm_{mean} 。利用该信息可以有效地排除位于平坦区域中的虚假点。

3.4.2 算法步骤

新的角点检测算法分为两个步骤：1、在小尺度参数下提取图像角点候选点；2、大尺度参数下，剔除候选点中虚假点。具体步骤如下：

3.4.2.1 小尺度参数下检测角点

1、选取适当的小尺度参数 $(\theta, \sigma_1, \omega_1)$ 构造 Gabor 梯度算子，计算图像上每一像素点 (i, j) 的 0 度、45 度、90 度与 135 度的 Gabor 梯度 $g_{i,j}^0$ 、 $g_{i,j}^{45}$ 、 $g_{i,j}^{90}$ 、 $g_{i,j}^{135}$ ；

2、以图像上每一像素点 (i, j) 为中心，取 $N \times N$ 大小的窗口按(3.25)式计算角点响应值，用非最大化抑制法选取局部 CRF 最大且大于指定阈值的点为角点候选点。

3.4.2.2 大尺度参数下剔除虚假点

1、选取合适的大尺度参数 $(\theta, \sigma'_1, \omega'_1)$ 构造 Gabor 梯度算子，计算角点候选点 (i, j) 的 $m \times m$ 邻域内每一点 (x, y) 的 0 度与 90 度的 Gabor 梯度 $g_{x,y}^0$ 、 $g_{x,y}^{90}$ ；

2、按(3.26)式计算角点候选点 (i, j) 的 Gabor 梯度模 $Gm_{i,j}$ ，按(3.27)式统计该点 $m \times m$ 邻域内所有像素的 Gabor 梯度模的平均值 Gm_{mean} ；

3、若角点候选点的梯度模小于或等于 Gm_{mean} ，则认为该点为虚假点，加以剔除。

3.5 实验结果及分析

为了验证 CDGGO 算法的可靠性、有效性，我们分别对仿真图和自然图像进行了处理，并与 Harris 算法、SUSAN 算法以及 Forstner 算法的检测结果进行比较。四种算法的检测效果均与阈值参数直接相关。调整阈值参数虽然能检测出更多的真实角点，但同时也增加了虚假角点数目。这里，我们的阈值选择原则是所选的阈值应保证检测出尽可能多的真实角点，且虚假点较少。

3.5.1 仿真图角点检测结果

3.5.1.1 无噪声仿真图

图 3-16.a 为 Harris 算法检测结果（参数为 k 为 0.04，高斯加权函数参数为 0.5，阈值参数为全图 CRF 的均值），图 3-16.b 为 SUSAN 算法检测结果（阈值参数为 10），图 3-16.c 为 Forstner 算法检测结果（ q 的阈值为 0.5， w 的阈值为 0.7），图

3-16.d 为 CDGGO 算法提取的结果 (小尺度参数 σ_1 、 ω_1 分别为 1.0 和 1.0, 大尺度参数 σ_1 、 ω_1 分别为 5.0 和 0.2)。实验结果表明, Harris 算法遗漏了一些角点, 在斜线边缘上检测出一些虚假角点。SUSAN 算法遗漏了部分角点, 角点定位较为准确。Forstner 算法虽然检测出大部分角点, 但是也存在部分虚假点。CDGGO 算法提出了图中所有角点, 没有虚假点, 提取效果较为理想, 尤其是 L 型角点的提取效果较好, 但是在 X 型角点定位还存在一定偏差。可以看出, CDGGO 算法检测效果优于其他三种算法。

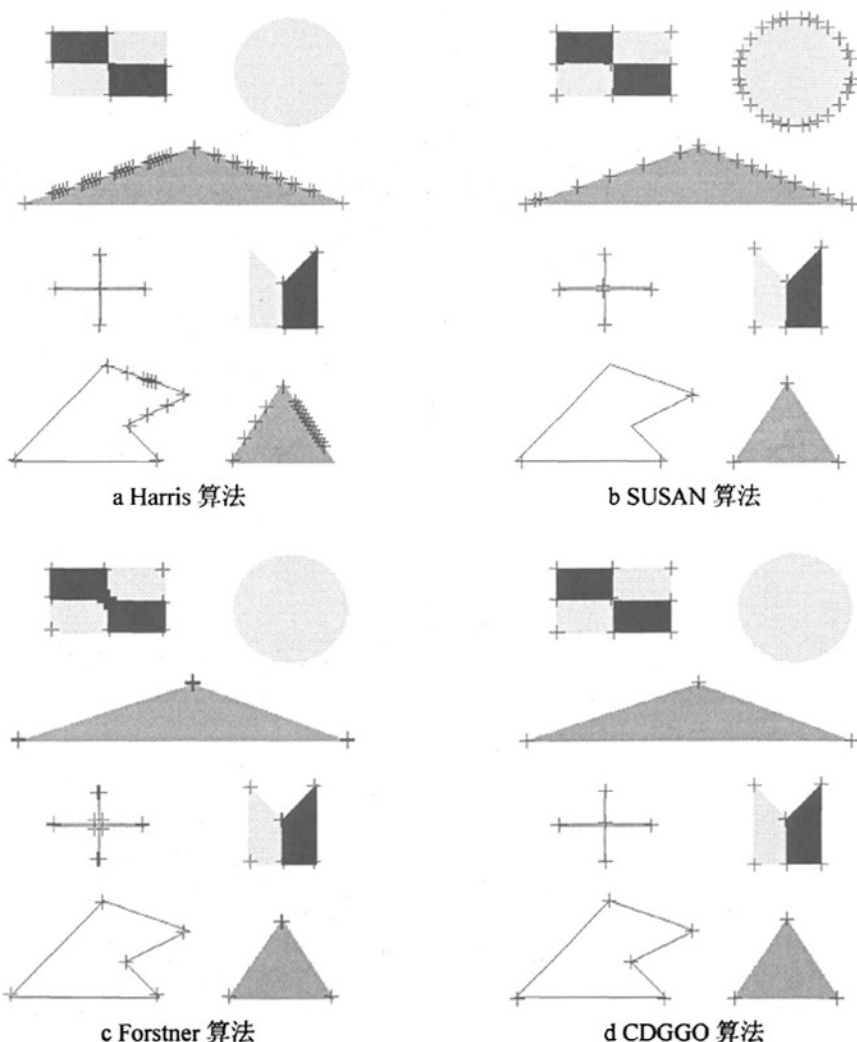


图 3-16 无噪声的检测结果

3.5.1.2 加入噪声的仿真图检测结果比较

仿真图像上先加入了方差为 30% 的高斯噪声, 再加入方差为 30% 的椒盐噪声,

图 3-17.a 为 Harris 算法提取的结果 (参数为 k 为 0.06, 高斯加权函数参数为 0.5, 阈值参数为全图 CRF 的均值), 图 3-17.b 为 SUSAN 算法提取结果 (阈值参数为 30), 图 3-17.c 为 Forstner 算法检测结果 (q 的阈值为 0.75, w 的阈值为 0.8), 图 3-17.d 为 CDGGO 算法提取的结果 (小尺度参数 σ_1 、 ω_1 分别为 4.0 和 0.25, 大尺度参数 σ_1 、 ω_1 分别为 18 和 0.06)。实验结果表明, Harris 算子当选择合适的阈值时, 对噪声有一定的抑制作用。虽然检测出了部分真实角点, 但是也遗漏了一些明显的角点, 同时存在较多虚假点。SUSAN 算法得到的结果不理想, 存在较多虚假点。Forstner 算法检测结果存在较多虚假点。CDGGO 算法对噪声有一定的抑制作用, 检测出大部分真实角点, 真实角点定位较为准确, 但也存在部分虚假角点。可以看出, CDGGO 算法检测效果优于其他三种算法。

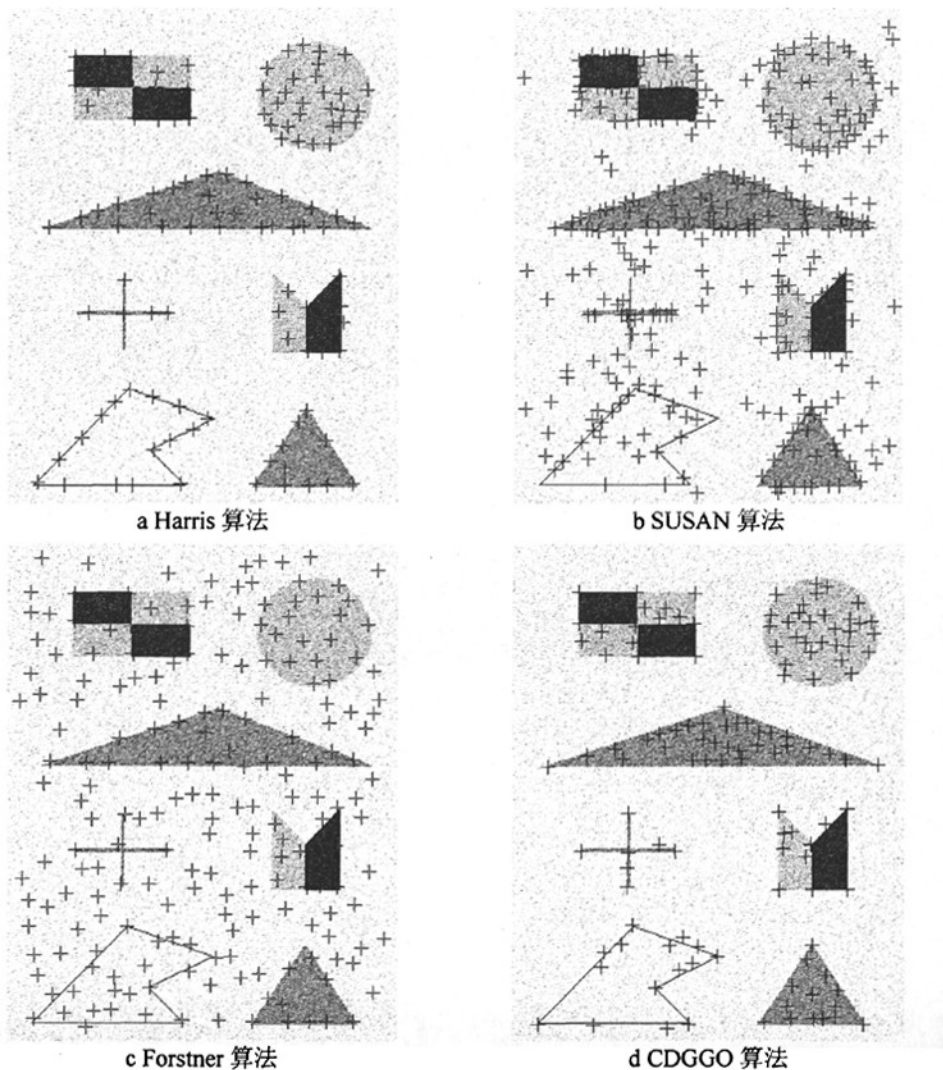


图 3-17 加入噪声的检测结果

3.5.2 真实图像角点检测实验结果

3.5.2.1 立方体图像角点提取结果

图 3-18.a 为 Harris 算法提取的结果 (参数为 k 为 0.04, 高斯加权函数参数为 0.5, 阈值参数为全图 CRF 的均值), 图 3-18.b 为 SUSAN 算法提取结果 (阈值参数为 20), 图 3-18.c 为 Forstner 算法检测结果 (q 的阈值为 0.7, w 的阈值为 0.75), 图 3-18.d 为 CDGGO 算法提取的结果 (小尺度参数 σ_1 、 ω_1 分别为 2.0 和 0.5, 大尺度参数 σ_1 、 ω_1 分别为 6 和 0.17)。实验结果表明, Harris 算法遗漏了部分角点, 角点定位较为准确, 但存在虚假点。SUSAN 算法遗漏了部分角点, 同时检测出一些边缘点, 边缘点的多少和阈值选取直接相关, 在消除边缘点的同时也减少了真正的角点。Forstner 算法遗漏了部分角点, 同时还存在较多虚假角点。而 CDGGO 算法提取效果较好, 虚假点最少, 提取效果优于其他三种算法。

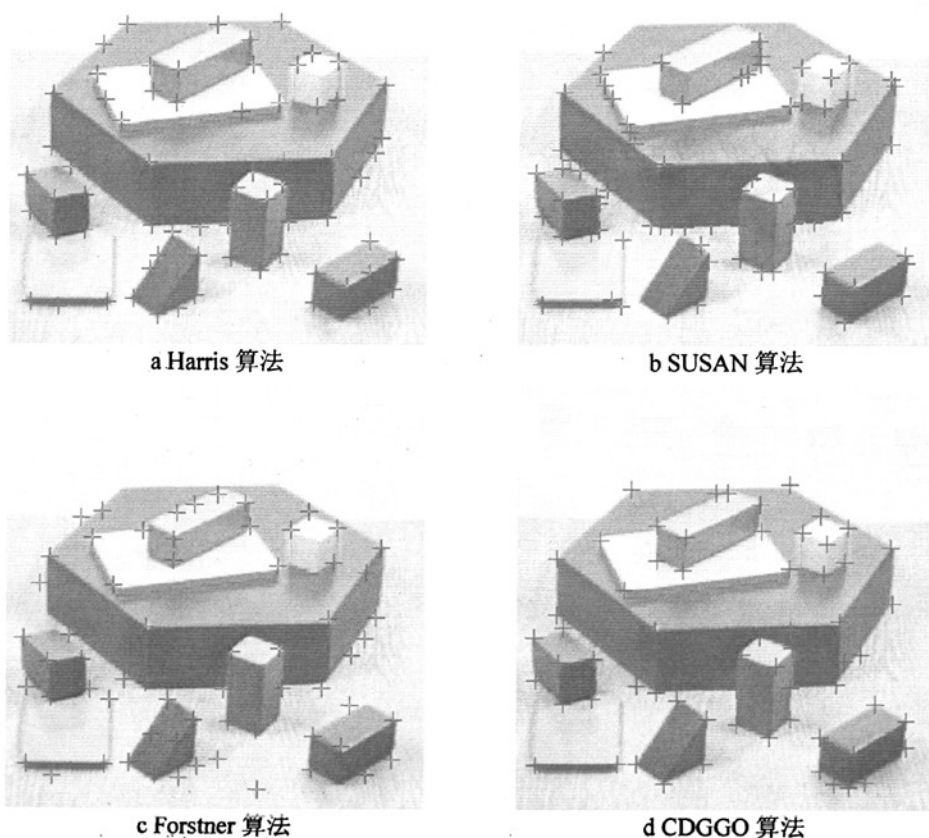


图 3-18 立方体图像检测结果

3.5.2.2 房屋图像角点提取结果

图 3-19.a 为 Harris 算法提取的结果 (参数为 k 为 0.04, 高斯加权函数参数为 0.5, 阈值参数为全图 CRF 的均值), 图 3-19.b 为 SUSAN 算法提取结果 (阈值参数为 30), 图 3-19.c 为 Forstner 算法检测结果 (q 的阈值为 0.75, w 的阈值为 0.75), 图 3-19.d 为 CDGGO 算法提取的结果 (小尺度参数 σ_1 、 ω_1 分别为 3 和 0.34, 大尺度参数 σ_1 、 ω_1 分别为 9 和 0.11)。实验结果表明, Harris 算法遗漏了较多角点, 定位结果有偏差, 存在虚假点。SUSAN 算法遗漏了部分明显特征点, 部分角点定位存在偏差, 同时检测出部分虚假角点。Forstner 算法遗漏了部分角点, 角点定位存在偏差, 同时还存在较多虚假角点。CDGGO 算法检测效果较好, 定位较为准确, 虚假点最少。可以看出, CDGGO 算法检测效果优于其他三种算法。

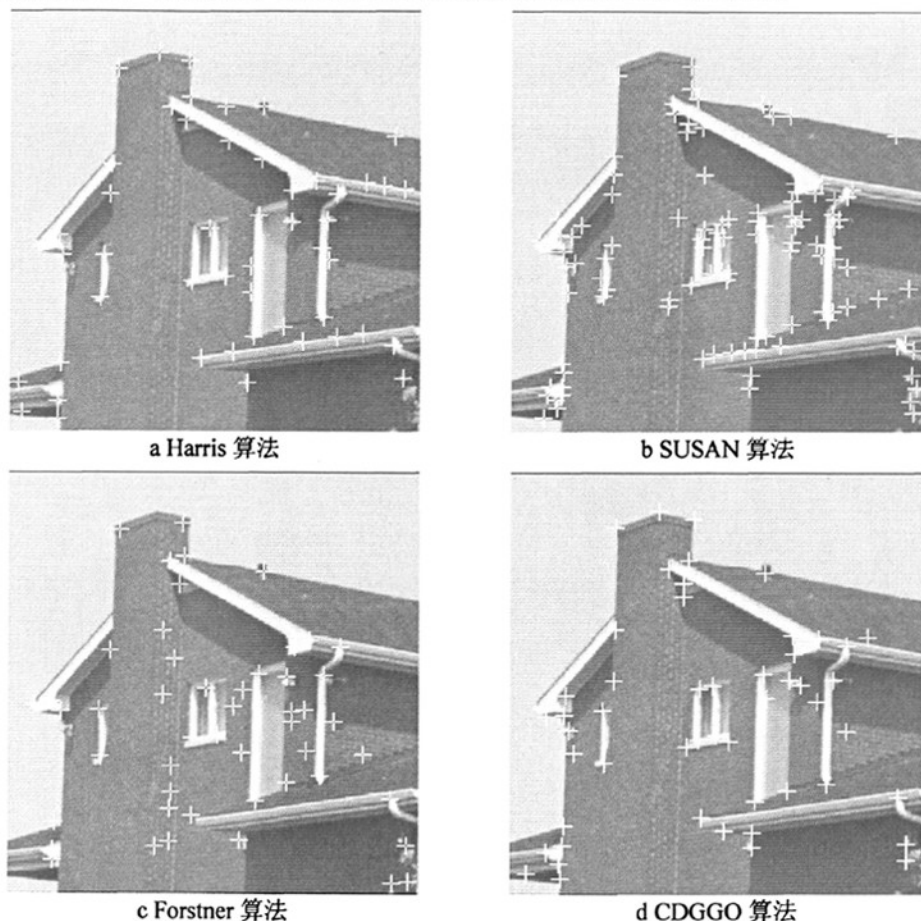


图 3-19 房子图像检测结果

3.5.2.3 实验室图像角点提取结果

图 3-20.a 为 Harris 算法提取的结果 (参数为 k 为 0.04, 高斯加权函数参数为

0.5, 阈值参数为全图 CRF 的均值), 图 3-20.b 为 SUSAN 算法提取结果 (阈值参数为 20), 图 3-20.c 为 Forstner 算法检测结果 (q 的阈值为 0.75, w 的阈值为 0.8), 图 3-20.d 为 CDGGO 算法提取的结果 (小尺度参数 σ_1 、 ω_1 分别为 4 和 0.25, 大尺度参数 σ_1 、 ω_1 分别为 6 和 0.17)。实验结果表明, Harris 算法遗漏了一些较为明显的角点, 角点定位普遍向角点内部偏移。SUSAN 算法能检测出大部分的真实角点, 但存在一些虚假点, 其中大部分虚假点为边缘点。Forstner 算法遗漏了部分角点, 同时还存在较多虚假角点, 角点定位较为准确。而 CDGGO 算法虚假点最少, 且定位准确。可以看出, CDGGO 算法检测效果优于其他三种算法。

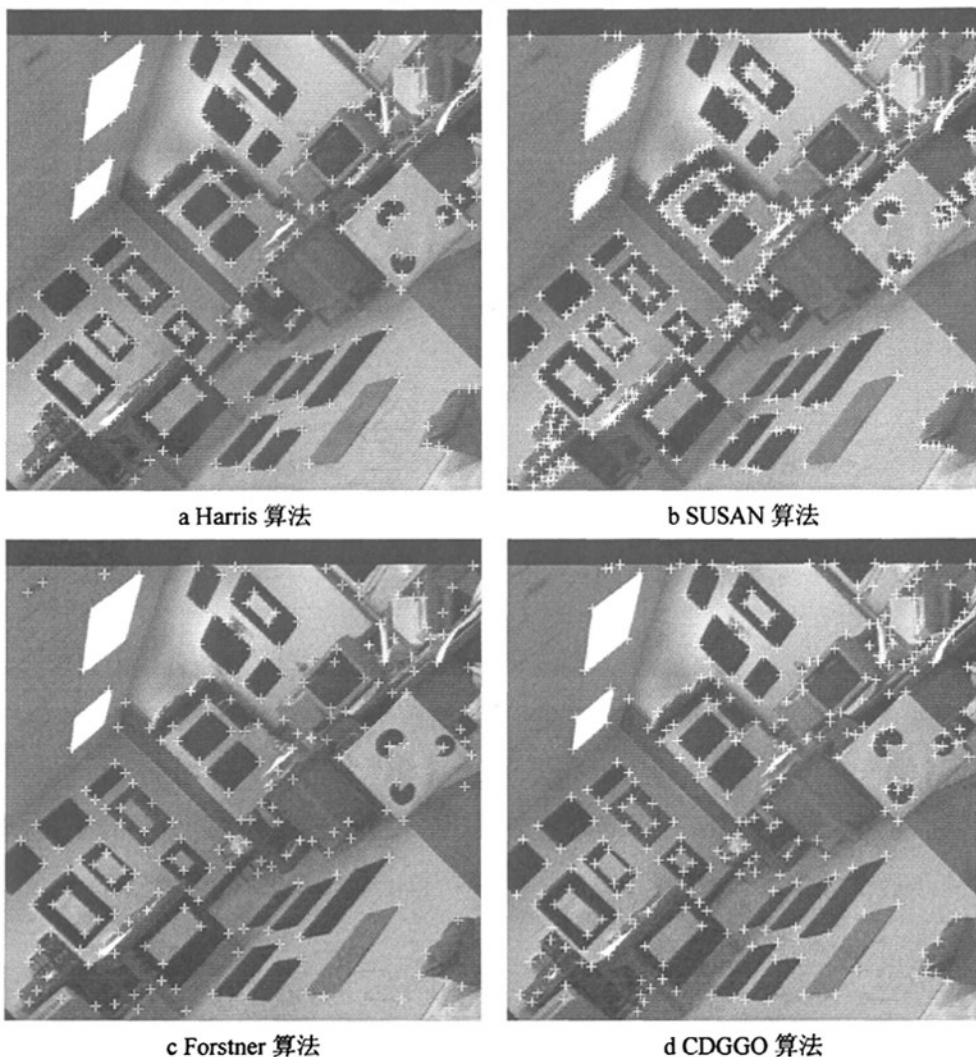


图 3-20 实验室图像检测结果

3.6 本章小结

本章针对图像角点特征提取算法进行了研究。本章简要介绍了图像角点特征分类，给出了角点特征检测算法的基本准则，详细介绍了角点检测算法分类，着重介绍四种经典的角点检测算法及其优缺点。

针对传统梯度算子的缺陷，在分析 Gabor 小波的基础上实现了一种 Gabor 梯度算子，有效地提高了梯度计算精度，同时具有较好的抗噪性。在 Gabor 梯度算子基础上，提出了一种新型角点检测算法。仿真图实验和真实图像实验均表明，新角点检测算法比 Harris 算法、SUSAN 算法和 Forstner 算法在角点检测性、噪声抑制等方面有了一定的提高。

第四章 基于 Gabor 特征描述的像机自运动消除方法研究

运动目标检测广泛应用于武器制导、交通流量监测、人体运动分析、视频压缩等诸多方面，涉及到计算机视觉、模式识别、统计学、图像理解等诸多研究领域，是图像处理研究领域的热点和难点之一。特别当像机自身也存在运动时，该问题变得更加难以解决。本章针对摄像机运动时的运动目标检测问题，提出了一种基于 Gabor 特征描述的像机自运动消除方法，为摄像机运动时检测运动目标奠定了基础。

本章内容的结构安排如下：4.1 节介绍了图像仿射变换的数学模型；4.2 节介绍基于特征匹配的运动目标检测算法的总体框架；4.3 节给出了基于 Gabor 特征描述的像机自运动消除算法原理；4.4 节给出了基于 Gabor 特征描述的像机自运动算法详细步骤；4.5 节给出了实验结果及分析；4.6 节对本章进行了小结。

4.1 图像仿射变换的数学模型

摄像机是获取图像的重要设备之一，建立摄像机的几何成像模型是消除像机自运动的基本前提。通常来说，摄像机几何成像模型可以用中心透视投影模型描述，在摄影测量学和计算机视觉等相关学科中，透视投影模型得到了充分研究，大多数算法都是基于透视投影模型开发的^[181,182]。但是，由于透视投影的非线性，这些算法相对比较复杂，算法数值稳定性较差，这一点在与图像测量相关的领域已经得到共识^[183]。

当拍摄过程中摄像机的内参数不发生变化、无人机距离拍摄区域较远时，视场远远大于目标的尺度，不共面的特征物点在远场图像中可以被近似认为是共面的。此时仿射投影模型是透视投影模型的很好近似。事实上，基于仿射投影模型的像机自运动消除算法不仅能取得足够的精度，而且能在很大程度上克服基于透视投影算法的数值不稳定的缺点^[184,185]。下面简要介绍不同视点图像间的仿射近似模型，为了方便读者理解，先简要介绍常用坐标系及其相互之间的变换关系。

4.1.1 常用坐标系及其关系

4.1.1.1 三个层次的坐标系统

摄影测量常用坐标系一般采用右手准则定义。图 4-1 表示了三个不同层次的坐标系：世界坐标系、摄像机坐标系和图像坐标系，其定义分别如下。

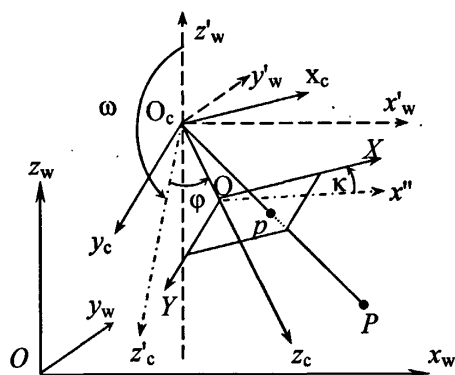


图 4-1 三个不同层次的坐标系统

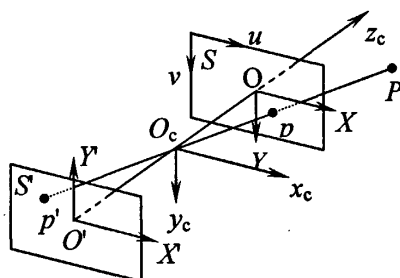


图 4-2 图像与摄像机坐标系

(1) 世界坐标系 (x_w, y_w, z_w)

也称作全局坐标系。它是由用户任意定义的三维空间坐标系，通常是将被测物体和摄像机作为一个整体考虑。图 4-1 中 (x'_w, y'_w, z'_w) 坐标系是以摄像机光心为原点，通过平移世界坐标系而得到的辅助坐标系。图 4-1 中的 ω , φ , κ 为决定摄像机光轴方位的三个角度量。在航空摄影测量中，其定义分别为： ω 是旁向倾角，定义为摄像机光轴 z_c 在平面 $y'_w z'_w$ 上的投影 z'_c 同 z'_w 之间的夹角； φ 是航向倾角，定义为摄像机光轴 z_c 同它在平面 $y'_w z'_w$ 上的投影 z'_c 之间的夹角； κ 是图像旋角，定义为 x_w 轴在图像平面 XY 内的投影 x''_w 和 X 轴的夹角。三个角度的正向按右手法则规定为： ω , κ 以逆时针为正， φ 以顺时针为正。在不同的应用领域中，这三个角度的定义可能不同。

(2) 摄像机坐标系 (x_c, y_c, z_c)

摄像机坐标系的原点为摄像机光心， z_c 轴与摄像机的光轴重合，且取摄影方向为正向。 x_c 、 y_c 轴通常与图像物理坐标系 X 、 Y 轴平行，如图 4-2 所示。其中平面 S' 和 S 定义为图像的负片和正片位置，分别位于摄像机坐标系的 $z_c = -f$ 平面

和 $z_c = f$ 平面内。

(3) 图像坐标系

在摄影测量学中，为了便于像点和对应点空间位置的相互换算，图像坐标系一般都建立在正片平面 S 中，如图 4-2 所示。图像坐标系分为图像像素坐标系 (u, v) 和图像物理坐标系 (X, Y) 两种，其定义分别为：

① 图像像素坐标系 (u, v)

图像像素坐标系是以图像左上角为原点，以像素为坐标单位的直角坐标系， u 、 v 分别表示像素在数字图像中的列数与行数。

② 图像物理坐标系 (X, Y)

图像物理坐标系是以光轴与象平面的交点为原点，以毫米为单位的直角坐标系，其 X 、 Y 轴分别与图像像素坐标系的 u 、 v 轴平行。

4.1.1.2 坐标系变换关系

在定义了上述各种空间直角坐标系后，为了达到摄影测量的目的，需要建立不同坐标系之间坐标变换的关系。

(1) 世界坐标系与摄像机坐标系的变换关系

世界坐标系中的点到摄像机坐标系上点的变换可由一个正交旋转变换矩阵 \mathbf{R} 和一个平移变换矩阵 \mathbf{T} 表示为

$$\begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R} & \mathbf{T} \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} = \mathbf{M}_2 \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \quad (4.1)$$

其中 $\mathbf{T} = [t_1, t_2, t_3]^T$ 是世界坐标系原点在摄像机坐标系中的坐标，矩阵

$$\mathbf{R} = \begin{bmatrix} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \\ r_7 & r_8 & r_9 \end{bmatrix} \text{ 中各项由(4.2)式计算得到。}$$

$$\begin{aligned} r_1 &= \cos \kappa \cos \varphi \\ r_2 &= \sin \kappa \cos \varphi - \cos \kappa \sin \varphi \sin \omega \\ r_3 &= \sin \kappa \sin \varphi + \cos \kappa \sin \varphi \cos \omega \\ r_4 &= -\sin \kappa \cos \varphi \\ r_5 &= \cos \kappa \cos \omega + \sin \kappa \sin \varphi \sin \omega \\ r_6 &= \cos \kappa \sin \omega - \sin \kappa \sin \varphi \cos \omega \\ r_7 &= -\sin \varphi \end{aligned} \quad (4.2)$$

$$r_8 = -\cos\varphi \sin\omega$$

$$r_9 = \cos\varphi \cos\omega$$

(2) 图像坐标系与摄像机坐标系的变换关系

空间中任意一点与其对应的图像点之间可用透视投影成像模型描述,如图 4-2 所示,所以摄像机坐标系中的物点 $P(x_c, y_c, z_c)$ 在图像物理坐标系中的像点 $p(X, Y)$ 坐标为:

$$\begin{cases} X = fx_c/z_c \\ Y = fy_c/z_c \end{cases} \quad (4.3)$$

其中 f 表示像机的焦距。

将上式的图像物理坐标系进一步转化成图像像素坐标系:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} 1/d_x & 0 & u_0 \\ 0 & 1/d_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \quad (4.4)$$

其中 u_0 、 v_0 是图像中心 (光轴与图像平面的交点) 的坐标, d_x 和 d_y 分别为一个像素在 X 与 Y 方向上的物理尺寸。

(3) 图像坐标系与世界坐标系的变换关系

(4.3)式表示空间点与对应的像之间的几何关系,是一种非线性的表示。为了利用坐标系间的线性变换关系,可将(4.3)式用齐次坐标表示为线性形式,即

$$\begin{bmatrix} fx_c \\ fy_c \\ z_c \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} = \mathbf{P} \begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} \quad (4.5)$$

注意到(4.5)式中等式右侧的向量 $[x_c \ y_c \ z_c \ 1]^T$ 实际上为空间点 p 的齐次坐标,因此(4.5)式表示了空间点与对应的像之间的齐次坐标关系,并且是一种线性的关系。由(4.1)式和(4.5)式可以建立图像坐标系中的点与世界坐标系中的点的齐次坐标表示形式:

$$\begin{bmatrix} fx_c \\ fy_c \\ z_c \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} = \mathbf{PM}_2 \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \quad (4.6)$$

其中

$$\mathbf{M}_2 = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}$$

又由于

$$\begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} = \begin{bmatrix} fx_c/z_c \\ fy_c/z_c \\ 1 \end{bmatrix} = \frac{1}{z_c} \begin{bmatrix} fx_c \\ fy_c \\ z_c \end{bmatrix} \quad (4.7)$$

将(4.6)式、(4.7)式代入(4.4)式中得到图像坐标 (u, v) 和世界坐标 $[x_w \ y_w \ z_w \ 1]^T$ 的齐次表示形式:

$$z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} 1/d_x & 0 & u_0 \\ 0 & 1/d_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \mathbf{M}_2 \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \quad (4.8)$$

由于齐次表示形式不唯一, 因此令 $s = z_c$, s 表示深度因子。令

$$\mathbf{M}_1 = \begin{bmatrix} 1/d_x & 0 & u_0 \\ 0 & 1/d_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} f_x & 0 & u_0 & 0 \\ 0 & f_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (4.9)$$

其中 $f_x = f/d_x$ 表示 u 轴上的归一化焦距, $f_y = f/d_y$ 表示 v 轴上的归一化焦距。 \mathbf{M}_1 矩阵中共有四个变量, 表示了像机焦距、分辨率以及像平面偏移量, 这些量只与像机的内部参数有关, 称之为像机的内参数。(4.8)式中的 \mathbf{M}_2 矩阵表示像机坐标系与世界坐标系之间的旋转平移关系, 称之为像机的外参数。因此, 图像上任意一点的像素坐标与世界坐标系间的关系可以表示为:

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{M}_1 \mathbf{M}_2 \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \quad (4.10)$$

由(4.10)式知, 对于世界坐标系中的一点, 由像机的内参数和外参数即可确定其在图像上的像素坐标。在世界坐标系中, 用两个不同的像机, 从不同的角度和距离获取空间中同一点的图像时, 可以用像机的成像模型确定该点在两幅图像上像素坐标之间的关系, 下面将在成像模型的基础上, 详细讨论不同视点图像之间的变换关系, 分析用仿射变换近似这种关系的条件。

4.1.2 不同视点图像间的仿射近似模型

上一节建立了在世界坐标系下空间中的一点与其对应的图像上点的坐标变换关系, 这种关系表示了单个像机单次拍摄获取图像时的情况。而对同一场景获取的不同图像, 可以认为是空间中的同名点, 在位于不同位置或角度的不同像机所成的像。可以通过空间同名点在不同像机的成像关系及像机之间几何约束, 得到

空间同名点在这些图像上对应的像素之间的变换关系。

4.1.2.1 多视角成像模型

4.1.1 节中对单个像机的成像模型进行了分析，当用不同像机从不同角度或距离对空间同一点成像时，可以用几何约束建立图像间的关系，下面以两个像机成像为例，分析同名点在不同图像上对应像素间的几何关系。

设两个像机为 C_1 和 C_2 ，空间中一点 P 在两个像机中所成的像分别为 p_1 、 p_2 ，如图 4-3 所示。

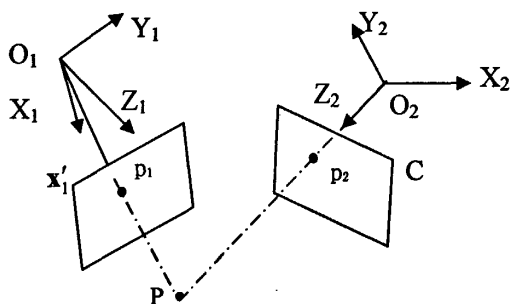


图 4-3 两个像机从不同视角获取图像模型

为了方便表述，设世界坐标系与 C_1 的像机坐标系重合，则 P 在世界坐标系和 C_1 像机坐标系中的坐标为 $p(x, y, z)$ ，像机外参数矩阵为单位阵，因此由(4.10)式知，对于像机 C_1 ，在该像机的图像坐标系中表示 P 点的像 $p_1(u_1, v_1)$ 与世界坐标中 $p(x, y, z)$ 之间的关系为：

$$s_1 \begin{bmatrix} u_1 \\ v_1 \\ 1 \end{bmatrix} = \begin{bmatrix} f_x^{(1)} & 0 & u_0^{(1)} & 0 \\ 0 & f_y^{(1)} & v_0^{(1)} & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (4.11)$$

其中 $s_1 = z$ 为深度因子，表示沿像机光轴方向上目标点的 z 轴坐标。 $f_x^{(1)}$ 、 $f_y^{(1)}$ 、 $u_0^{(1)}$ 、 $v_0^{(1)}$ 为像机 C_1 的内参数。点 P 在像机 C_2 的像素坐标 (u_2, v_2) 与世界坐标中 $p(x, y, z)$ 之间的关系为：

$$s_2 \begin{bmatrix} u_2 \\ v_2 \\ 1 \end{bmatrix} = \begin{bmatrix} f_x^{(2)} & 0 & u_0^{(2)} & 0 \\ 0 & f_y^{(2)} & v_0^{(2)} & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (4.12)$$

其中 $f_x^{(2)}$ 、 $f_y^{(2)}$ 、 $u_0^{(2)}$ 、 $v_0^{(2)}$ 为像机 C_2 的内参数， s_2 为深度因子，表示在 C_2 像机坐标系下 z 轴的坐标， \mathbf{R} 、 \mathbf{t} 分别表示 C_2 的像机坐标系与世界坐标系间的旋转矩

阵和平移分量，由(4.11)式和(4.12)式消除 x 和 y 分量可以得到：

$$s_2 \begin{bmatrix} u_2 \\ v_2 \\ 1 \end{bmatrix} = \begin{bmatrix} f_x^{(2)} & 0 & u_0^{(2)} & 0 \\ 0 & f_y^{(2)} & v_0^{(2)} & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} \frac{s_1 u_1 - z u_0^{(1)}}{f_x^{(1)}} \\ \frac{s_1 v_1 - z v_0^{(1)}}{f_y^{(1)}} \\ z \\ 1 \end{bmatrix} \quad (4.13)$$

由(4.13)式解得空间中同名点在两个图像像素坐标系中 (u_1, v_1) 与 (u_2, v_2) 之间的关系如下：

$$\begin{cases} u_2 = \frac{k_1 z}{f_x^{(1)} s_2} u_1 + \frac{k_2 z}{f_y^{(1)} s_2} v_1 - \frac{k_1 z}{f_x^{(1)} s_2} u_0^{(1)} - \frac{k_2 z}{f_y^{(1)} s_2} v_0^{(1)} + \frac{k_3 z}{s_2} + \frac{k_4}{s_2} \\ v_2 = \frac{k_5 z}{f_x^{(1)} s_2} u_1 + \frac{k_6 z}{f_y^{(1)} s_2} v_1 - \frac{k_5 z}{f_x^{(1)} s_2} u_0^{(1)} - \frac{k_6 z}{f_y^{(1)} s_2} v_0^{(1)} + \frac{k_7 z}{s_2} + \frac{k_8}{s_2} \end{cases} \quad (4.14)$$

其中

$$\begin{aligned} k_1 &= r_1 f_x^{(2)} + r_7 u_0^{(2)} \\ k_2 &= r_2 f_x^{(2)} + r_8 u_0^{(2)} \\ k_3 &= r_3 f_x^{(2)} + r_9 u_0^{(2)} \\ k_4 &= t_1 f_x^{(2)} + t_3 u_0^{(2)} \\ k_5 &= r_4 f_y^{(2)} + r_7 v_0^{(2)} \\ k_6 &= r_5 f_y^{(2)} + r_8 v_0^{(2)} \\ k_7 &= r_6 f_y^{(2)} + r_9 v_0^{(2)} \\ k_8 &= t_2 f_y^{(2)} + t_3 v_0^{(2)} \end{aligned} \quad (4.15)$$

通过以上推导，得出两幅图像像素坐标之间的关系式(4.14)，由表示形式可以看出，两幅图像坐标之间是一种线性表示形式，下面将详细分析这种关系所对应的模型。

4.1.2.2 仿射近似模型

(4.14)式表示像机从不同视角获取空间中同一点的像素坐标之间的关系，从表示形式可以看出，像素之间的关系与像机的内参数、外参数、 z 以及 s_2 有关。空间中三维目标的不同位置在两幅图像所成像的坐标将随着三维目标表面的起伏变化而变化，如图 4-4 所示。

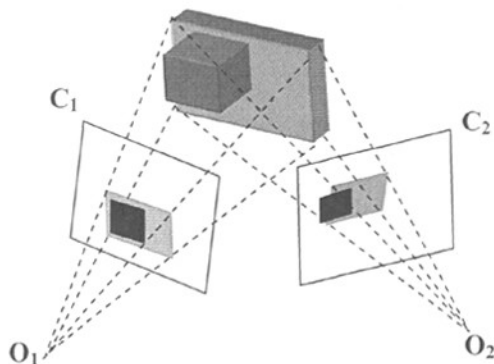


图 4-4 对应点在不同视点图像中的变化关系

当同一像机在不同角度或距离对空间同一目标成像后，对于目标上每一点，像机的内、外参数可视为已知的常数。由(4.14)式知，空间非共面点在不同图像上像素坐标变换关系并不相同，这种差别主要由参数 z 和 s_2 造成，将(4.14)式简化为如下形式：

$$\begin{cases} u_2 = r_0 u_1 + r_1 v_1 + r_4 \\ v_2 = r_2 u_1 + r_3 v_1 + r_5 \end{cases} \quad (4.16)$$

其中 r_i 为与 z 和 s_2 有关的系数，(4.16)式说明在 z 和 s_2 已知的条件下，两幅图像上像素点之间满足一种线性的几何关系，即仿射变换模型。但由于仿射变换系数与 z 和 s_2 有关，而图像像素所对应的 z 和 s_2 因目标的起伏而存在差别，因此两幅图像所有对应像素点对并不服从同一个仿射变换模型。

无人机图像定义为无人机上的像机对其他景物拍摄的图像。当像机在不同角度或距离拍摄目标时，如果目标深度的变化相对于平均深度可以忽略，则 z 和 s_2 近似为常数。在这种情况下，两幅图像中所有对应像素点对之间的坐标变换关系可以近似用同一个仿射变换模型进行描述。

对于无人机图像，当满足一定条件时，可以用仿射变换模型代替透视投影模型，分析不同视角或不同距离拍摄的两幅图像上像素之间的关系。本节主要介绍了不同视点图像间的仿射变换模型及其适用条件。从仿射变换公式可以看出，仿射变换参数未知数有六个，而一对对应像素点可以建立两个关于仿射变换参数的方程，所以当获得三对以上对应的像素点对时，可以估计出仿射变换参数，从而对其中一幅图像进行变换，达到消除像机自运动的目的。本章的后续章节将围绕基于特征匹配的像机自运动消除算法进行阐述。

4.2 基于特征匹配的像机自运动消除算法总体框架

由于无人机自身姿态或位置的改变，其上摄像机的姿态或位置也会有所变化。对于运动平台上运动目标检测问题，常用的一种方法需要先消除像机自身运动，即将同一场景的不同角度或位置拍摄的图像进行配准，然后检测运动目标。

按照配准基元的不同，图像配准算法可以分为以下三类^[186,187]：基于区域(Area-based)的算法、基于特征(Feature-based)的算法和基于结构关系(Structure Relation-based)的算法。其中，基于特征的算法因具有受几何变形和灰度变化影响小、可匹配特征类型多等优点而备受关注。

无人机图像中蕴含着大量的角点、边界线和纹理等信息。这些信息较好地满足基于特征的图像配准算法的要求，因此该类图像配准算法非常适用于无人机图像配准。

下面简要介绍基于特征匹配的像机自运动消除算法的总体框架，具体包括：1、提取特征，即提取图像中的点特征、线特征或面特征等；2、对每一个特征（点、线或面等）建立特征描述；3、根据特征描述及相似性函数对特征进行匹配；4、利用匹配结果估计图像变换模型参数；5、配准图像，消除像机自运动。其流程图如图 4-5 所示。

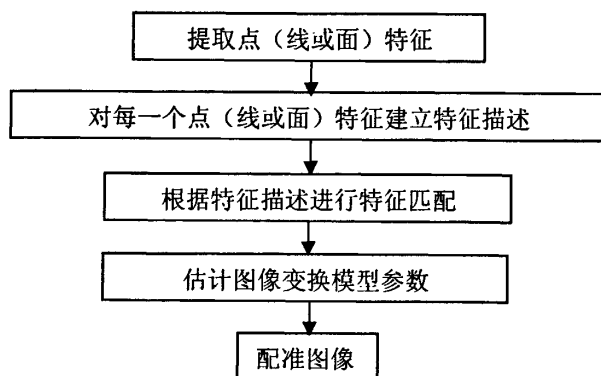


图 4-5 基于特征匹配的像机自运动消除算法总体框架

4.3 基于 Gabor 特征描述的像机自运动消除算法原理

为了消除像机自运动，本文提出了一种基于 Gabor 特征描述的像机自运动消除算法 (Camera Motion Compensation based on Gabor Feature Description, 简称 CMCGFD 算法)。新算法首先对两帧待配准的图像提取兴趣点，并对每个兴趣点建立特征描述，然后通过特征匹配获得对应的兴趣点对，接着利用已对应的兴趣点对估计出仿射变换模型参数，最后配准两帧图像，消除像机自运动。下面就算

法原理进行简要介绍。

4.3.1 兴趣点提取

如果对图像上所有点建立特征描述,再对前后视图进行特征匹配,计算量较大,而且通常匹配结果并不理想。其主要原因是图像中存在很多非稳定点,如平坦的平面或直线边缘上一些点,对这些点建立的特征描述非常相似,容易造成错误匹配,所以应尽量剔除这些非稳定点。在工程应用中,常采用 Harris 算法或 Forstner 算法提取图像上角点,减少非稳定点。然而,对于宽基线两帧图像进行配准时,Harris 算法或 Forstner 算法提取的角点会因成像角度不同或遮挡等问题而不容易得到正确匹配。Mikolajczyk^[188]提出了一种适用于宽基线匹配的多尺度 Harris 兴趣点提取算法。Lowe^[75]提出了 SIFT 匹配算法。这两种方法较好地解决了 Harris 算法和 Forstner 算法关于宽基线匹配的问题。这里,我们选取类似于 SIFT 算法的思路获得候选兴趣点。

获取图像候选兴趣点后,为了加速算法且降低误匹配率,应排除一部分不适合进行匹配的兴趣点。由 2.1.2 节的特征尺度选择理论可知,图像中任一像素特征尺度的三倍定义了该点的最佳描述区域大小。如果兴趣点的特征尺度太小,该点对应的最佳描述区域相应较小,对该点建立的特征描述可能因包含的信息比较少而不具备较好的可区分性;相反,如果其特征尺度太大,该点对应的最佳描述区域会较大,虽然包含的信息较多,但是计算量也相应增加了。所以,应排除特征尺度太小或太大的兴趣点。同时为了提高图像配准的准确性,还应尽量剔除不稳定的兴趣点,如平坦平面区域中点、直线边缘上的点等。

4.3.2 特征描述区域确定

提取图像兴趣点后,通常需要以其为中心选择一个区域,根据区域内像素点的颜色、梯度或梯度方向等信息建立特征描述,为特征匹配做准备。在区域选择上,常用的一类做法是对每个兴趣点选用相同大小的矩形或圆形区域,这样做存在两个问题:

- 1、对一个兴趣点,如果任意指定一定大小的区域建立特征描述,则该特征描述可能因包含的信息较少而不具备较好的可区分性,或者可能包含冗余信息而增加额外的计算量;

- 2、当序列图像上目标区域有缩放变化时,采用固定大小区域描述不能适应目标缩放变化。

根据 2.1.2 节的 Lindeberg 的特征尺度选择理论,目标的最佳描述区域大小与其特征尺度直接相关,而且特征尺度会根据图像中目标区域的缩放变化自适应作

出调整。该理论为建立尺度自适应特征描述提供了理论基础。

对于图像上兴趣点,可以根据 Lindeberg 特征尺度选择理论计算出该点的特征尺度参数 σ_c 。以三倍 σ_c 为半径、该点为中心确定一个圆形区域,该圆形区域定义了描述该点的最佳区域。由 Lindeberg 特征尺度选择理论可知,圆形区域增大则增加了冗余信息,而减小则描述信息不够。当不同图像上的同一目标存在缩放变形时,依据 Lindeberg 尺度选择理论,该目标的特征尺度 σ_c 会根据图像上目标的缩放程度进行自适应调整,即最佳描述区域大小具有尺度不变性。

对于图像上任意一个兴趣点,选用由其三倍特征尺度确定的最佳描述区域作为该点的特征描述区域,该区域定义了描述兴趣点的最佳区域大小,同时具有尺度不变性,因此可以较好地解决传统做法在目标描述区域确定上的两个问题。

4.3.3 特征方向确定

对于一个兴趣点,其特征方向 θ_c 定义为:当图像产生旋转变化时, θ_c 应根据图像的旋转自适应地作出调整,即特征方向应具有旋转不变性。

一个兴趣点的特征描述区域确定后,可以统计出该区域内所有像素点的梯度方向直方图,从而获得梯度方向直方图峰值。图像局部区域的梯度方向直方图定义为该区域内每个像素的梯度方向按其对应的角度进行投票,统计 360 度中每个角度的投票数。当图像只存在旋转变化时,特征区域内每个像素点的梯度方向均产生变化,重新统计旋转后该区域内所有像素的梯度方向直方图,可以看出该梯度方向直方图的峰值也相应地产生了变化,而这种变化与图像的旋转变化一致。因此,对于一个兴趣点,选择其特征区域的梯度方向直方图峰值所对应的方向作为该点的特征方向 θ_c , 该方向具有旋转不变性。

4.3.4 特征描述向量建立

传统的特征描述是以兴趣点为中心选取一定区域,通过统计该区域内的颜色(或灰度)直方图、梯度直方图、梯度方向直方图等信息建立特征描述。这些特征描述虽然建立简单,但是特征区分性不是很好。

Gabor 函数被认为是高等动物视觉皮层简单细胞感受野的最佳近似。对图像上任意一点,可以通过选择参数不同的二维 Gabor 滤波器与该点局部区域图像进行卷积,获得一组 Gabor 特征向量,该特征向量具有较好的区分性。(关于 Gabor 函数的介绍,请参考 2.2 节)

从二维 Gabor 滤波器的数学表达式(4.17)上可以看出,该函数不但具有方向信息描述参数 θ , 而且还具有尺度信息描述参数 σ_x 和 σ_y 。

$$g(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)\right] \exp[\omega_0(x\cos\theta + y\sin\theta)] \quad (4.17)$$

对于图像上兴趣点,通过 4.3.2 节和 4.3.3 节的分析可知,该点的特征尺度 σ_c 和特征方向 θ_c 可以根据图像的缩放、旋转变化自适应地进行调整,而三倍 σ_c 定义了兴趣点的最佳描述区域。如果将三倍 σ_c 赋给 Gabor 滤波器的尺度信息参数 σ_x 和 σ_y ,而将 θ_c 赋给 Gabor 滤波器的方向信息描述参数 θ ,则可以定义一个尺度、方向自适应的二维 Gabor 滤波器。将该滤波器与兴趣点局部图像进行内积运算,可以获得该点的一个特征描述。为了获得该点的特征描述向量,可以按一定规律修改参数 θ_c (这里不能修改 σ_c ,因为 σ_c 和最佳描述区域大小有关),构造一系列二维 Gabor 滤波器与兴趣点局部图像进行内积运算,将所得的结果构成一个特征描述向量。由于该特征描述向量是以特征尺度 σ_c 和特征方向 θ_c 为基础建立的,而这两个参数可以随着图像旋转、缩放变化自适应地进行调整,因此该特征描述向量具备尺度不变性和旋转不变性。

4.3.5 像机自运动消除

获得前后两帧图像的兴趣点后,通过选用相似性函数对兴趣点建立匹配关系。然而由于噪声等因素的影响,通常会存在少量错匹配点对。如果直接利用包含有误匹配点对的匹配结果估计模型参数,得到模型参数存在一定误差,所以需要采用一种鲁棒的参数计算方法估计图像间的仿射模型参数。本文采用随机采样一致性算法(Random Sample Consensus, 简称 RANSAC)估计仿射变换参数,该方法对误匹配点对有较好的抑制作用。最后根据获得的仿射变换参数对其中一幅图像进行变换,达到消除像机自运动的目的。

4.4 基于 Gabor 特征描述的像机自运动消除算法步骤

CMCGFD 算法首先计算前后两帧图像中每个像素的特征尺度,然后对两帧图像分别提取兴趣点,并对每个兴趣点建立特征描述,接着利用交叉匹配算法对两幅图像兴趣点进行匹配,最后利用对应的兴趣点对估计出两幅图像之间的仿射变换模型参数,根据得到的仿射变换消除像机自运动。算法步骤如图 4-6 所示。

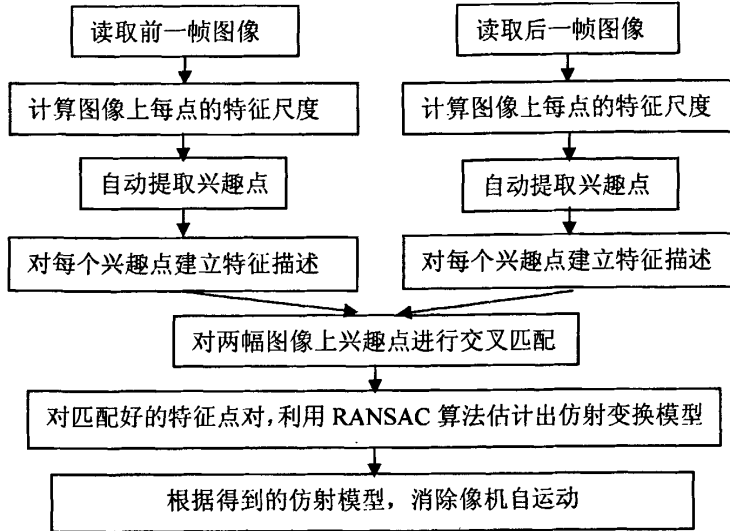


图 4-6 CMCGFD 算法流程图

下面简要介绍算法各个步骤。

4.4.1 计算图像上每一点的特征尺度

用不同尺度参数的高斯函数与原始图像进行卷积，获得该图像的不同尺度下的高斯图像（高斯图像定义为用参数 σ 的二维高斯函数对图像进行滤波的结果），其中高斯函数的尺度参数选取方法为 $\sigma_n = k^n \sigma_{n-1}$ ， $k = \sqrt{2}$ 。将尺度参数相邻的两个高斯图像相减，得到该图像不同尺度的DOG图像。对于图像上任意一点 \mathbf{x} ，其尺度参数为 σ 的DOG算子响应值定义为该点DOG图像的“灰度值”。由2.1节的尺度空间理论可知，该点的不同尺度的DOG算子响应曲线的最大值所对应的参数 σ 即为该点的特征尺度 σ_c 。因此，图像上任意一点 \mathbf{x} 的特征尺度 σ_c 应满足以下条件：

$$DOG(\mathbf{x}, \sigma_c) > DOG(\mathbf{x}, \sigma_n) \quad (4.18)$$

其中 $DOG(\mathbf{x}, \sigma_c)$ 为点 \mathbf{x} 处特征尺度 σ_c 对应的DOG算子响应值， $DOG(\mathbf{x}, \sigma_n)$ 表示点 \mathbf{x} 处特征尺度以外的参数所对应的DOG算子响应值。

4.4.2 自动提取兴趣点

如果对图像上所有点建立特征描述，再对前后视图进行特征匹配，计算量大，而且通常匹配结果不理想。因此，获取图像所有像素点的特征尺度后，应从中排除一些不适于进行特征匹配的兴趣点，具体做法如下：

- (1) 排除特征尺度过大或过小的像素点。如果图像上某个像素点的特征尺度

太小, 该点的最佳描述区域相应也就很小, 对该点建立特征描述可能会因区域内信息太少而不具备较好的可区分性。而当像素点的特征尺度太大, 对其建立特征描述的计算量会较大。所以, 应对图像上像素点进行排除, 保留特征尺度满足(4.19)式的像素点, 其中 th_1 和 th_2 为预先选取的阈值, σ_c 为像素点的特征尺度参数。

$$th_2 > \sigma_c > th_1 \quad (4.19)$$

(2) 对经过第一步排除后剩下的点继续进行排除。设图像上一点 \mathbf{x} 的特征尺度为 σ_c 。在点 \mathbf{x} 的特征尺度 σ_c 对应的 DOG 图像上, 如果该点的 $DOG(\mathbf{x}, \sigma_c)$ 大于其 5×5 邻域像素 \mathbf{x}_j 的 $DOG(\mathbf{x}_j, \sigma_c)$, 即满足公式(4.20), 则保留; 否则排除。

$$DOG(\mathbf{x}, \sigma_c) > DOG(\mathbf{x}_j, \sigma_c) \quad (4.20)$$

(3) 剔除非稳定点。Harris 等人指出, 图像点的 Hessian 矩阵 (简称 \mathbf{H} 矩阵) 的特征值正比于图像点的主曲率^[189]。假设 λ_1 和 λ_2 为 \mathbf{H} 矩阵的两个特征值, 且 $\lambda_1 > \lambda_2$ 。如果 λ_1 和 λ_2 都非常小, 说明该点为平坦区域中的一点; 如果 λ_1 非常大而 λ_2 非常小, 说明该点为边缘上的一点; 如果 λ_1 和 λ_2 都很大, 说明该点为角点。根据以上原理, 可以排除平坦区域和直线边缘上点。具体做法为: 假设 \mathbf{x} 点为经过前两步排除后剩下的候选点, σ_c 为该点的特征尺度, $L(\mathbf{x}, \sigma_c)$ 为特征尺度 σ_c 对应的高斯图像。计算该点的 2×2 的 \mathbf{H} 矩阵, 即

$$\mathbf{H} = \begin{bmatrix} L_{xx}(\mathbf{x}, \sigma_c) & L_{xy}(\mathbf{x}, \sigma_c) \\ L_{xy}(\mathbf{x}, \sigma_c) & L_{yy}(\mathbf{x}, \sigma_c) \end{bmatrix} \quad (4.21)$$

其中 $L_{xx}(\mathbf{x}, \sigma_c)$ 和 $L_{yy}(\mathbf{x}, \sigma_c)$ 分别表示 $L(\mathbf{x}, \sigma_c)$ 关于 x 方向和 y 方向的二阶导数, $L_{xy}(\mathbf{x}, \sigma_c)$ 表示 $L(\mathbf{x}, \sigma_c)$ 先关于 x 方向求导, 再关于 y 方向求导的二阶导数。

如果该点 \mathbf{H} 矩阵满足公式(4.22)则将该点作为兴趣点; 否则剔除。其中 th_3 和 th_4 为预先选取的阈值。

$$\begin{aligned} Tr(\mathbf{H}) &> th_3 \\ \frac{Tr(\mathbf{H})}{Det(\mathbf{H})} &< th_4 \end{aligned} \quad (4.22)$$

4.4.3 确定兴趣点的特征方向

假设 \mathbf{x} 为提取的兴趣点, σ_c 为该点的特征尺度, $L(\mathbf{x}, \sigma_c)$ 为特征尺度 σ_c 对应的高斯图像。将 360° 间隔 10° 划分为 36 个方向栅格。在高斯图像 $L(\mathbf{x}, \sigma_c)$ 上以三倍 σ_c 为半径、 \mathbf{x} 为中心选取一个圆形区域, 按公式(4.23)计算出圆形区域中每点的角度。因为反正切函数对应的角度范围为 $(-90^\circ, 90^\circ)$, 所以需要根据 $(L_{i,j+1} - L_{i,j-1})$ 和 $(L_{i+1,j} - L_{i-1,j})$ 的正负以及是否为零进行判断, 将角度计算结果调整到 0° 到 360° 。

$$\theta = \arctan\left(\frac{L_{i,j+1} - L_{i,j-1}}{L_{i+1,j} - L_{i-1,j}}\right) \quad (4.23)$$

然后按公式(4.24)统计角度方向栅格直方图。

$$q_u = \sum_{i=1}^N \delta[b(\theta_i) - u] \quad (4.24)$$

$$u = 0, \dots, 35 \quad b(\theta_i) = [\theta_i / 360] \quad \delta(x) = \begin{cases} 1 & x = 0 \\ 0 & x \neq 0 \end{cases}$$

其中 N 为圆形区域中包含的像素点总数, θ_i 为区域内第 i 像素的角度。函数 $b(\theta_i)$ 将角度 θ_i 映射到对应的方向栅格。 δ 函数为迪拉克函数。找寻角度栅格直方图的最大峰值对应的角度作为该兴趣点的特征方向 θ_c 。

4.4.4 建立特征描述

分别获得待配准两帧图像的兴趣点后, 需要对两帧图像上每个兴趣点建立特征描述, 才能进行匹配。具体做法为: 假设图像上某一兴趣点 \mathbf{x} 坐标为 (x_0, y_0) , σ_c 为该点的特征尺度, θ_c 为该点的特征方向。令 $s_x = s_y = 3\sigma_c$ 、 $\theta = \theta_c + n \times \pi/8$ (其中 $n = 0 \sim 7$)、 $c_x = x_0$ 、 $c_y = y_0$ 、 $\omega_0 = 0.283 \times 3\sigma_c$, 按公式(4.25)定义 8 个 Gabor 奇函数, 按公式(4.26)定义 8 个 Gabor 偶函数。

$$\begin{aligned} \psi_n(x, y | c_x, c_y, \theta, s_x, s_y) = & \frac{1}{2\pi s_x s_y} \exp\left\{-\frac{1}{2}\left[\frac{1}{s_x^2}((x-c_x)\cos\theta - (y-c_y)\sin\theta)^2\right.\right. \\ & \left.\left. + \frac{1}{s_y^2}((x-c_x)\sin\theta + (y-c_y)\cos\theta)^2\right]\right\} \times \sin\{\omega_0[(x-c_x)\cos\theta + (y-c_y)\sin\theta]\} \end{aligned} \quad (4.25)$$

$$\begin{aligned} \bar{\psi}_n(x, y | c_x, c_y, \theta, s_x, s_y) = & \frac{1}{2\pi s_x s_y} \exp\left\{-\frac{1}{2}\left[\frac{1}{s_x^2}((x-c_x)\cos\theta - (y-c_y)\sin\theta)^2\right.\right. \\ & \left.\left. + \frac{1}{s_y^2}((x-c_x)\sin\theta + (y-c_y)\cos\theta)^2\right]\right\} \times \cos\{\omega_0[(x-c_x)\cos\theta + (y-c_y)\sin\theta]\} \end{aligned} \quad (4.26)$$

设 $L(\mathbf{x}, \sigma_c)$ 为特征尺度 σ_c 所对应的高斯图像。在高斯图像 $L(\mathbf{x}, \sigma_c)$ 上以 \mathbf{x} 点为中心, 以该点特征尺度 σ_c 的三倍为半径选取一个圆形特征区域, 分别将定义的八个 Gabor 奇函数和八个 Gabor 偶函数分别与 \mathbf{x} 点的圆形特征区域进行内积运算, 利用得到的内积值建立该点的 16 维 Gabor 特征描述向量。

4.4.5 前后视图交叉匹配

获得待配准的两帧图像兴趣点并建立特征描述后,需要对兴趣点进行特征匹配,用以确定兴趣点之间对应关系。判断两个特征描述是否相似,可以通过比较它们之间的欧氏距离实现。设 \mathbf{f}_1 与 \mathbf{f}_2 分别为两幅图像上待匹配的兴趣点 \mathbf{x}_1 和 \mathbf{x}_2 的16维特征描述向量,其欧氏距离 $d(\mathbf{f}_1, \mathbf{f}_2)$ 表示为

$$d(\mathbf{f}_1, \mathbf{f}_2) = \sqrt{\sum_{i=1}^{16} (f_{1i} - f_{2i})^2} \quad (4.27)$$

其中 f_{1i} 和 f_{2i} 分别为特征描述向量 \mathbf{f}_1 与 \mathbf{f}_2 的第 i 个分量。欧氏距离越小,说明两个兴趣点的特征描述越相似。为了降低误匹配率,本文采用交叉匹配算法,即先以第一幅图像中的兴趣点为参考,在第二幅图像中找寻最匹配的点,接着再以第二幅图像中的兴趣点为参考,在第一幅图像中找寻最匹配的点。如果两幅图像上两个兴趣点 \mathbf{x}_1 和 \mathbf{x}_2 互为最匹配点,则认为该点对为匹配点对。对两幅图像所有兴趣点做以上判断,找出所有对应点对。

4.4.6 建立图像之间变换模型

设 \mathbf{x}_1 和 \mathbf{x}_2 为同一平面或近似平面场景由不同视点拍摄的两幅图像中的对应两点, $(x_1, y_1, 1)$ 、 $(x_2, y_2, 1)$ 分别为图像像素坐标系中 \mathbf{x}_1 和 \mathbf{x}_2 两点的齐次坐标。通过4.1节可知,空间同名点在不同视角或不同距离拍摄的图像上,其对应像素坐标之间的变换关系可以用仿射变换模型进行描述,即

$$\begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} = \begin{bmatrix} r_0 & r_1 & r_4 \\ r_2 & r_3 & r_5 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} = \mathbf{A} \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} \quad (4.28)$$

$$\mathbf{A} = \begin{bmatrix} r_0 & r_1 & r_4 \\ r_2 & r_3 & r_5 \\ 0 & 0 & 1 \end{bmatrix}$$

其中 \mathbf{A} 称为仿射变换矩阵, $\{r_0, r_1, \dots, r_5\}$ 为仿射变换参数。由式(4.28)可以看出,两幅图像中的每个匹配点对可以建立两个关于仿射变换参数 $\{r_0, r_1, \dots, r_5\}$ 方程,如果可以得到三对以上的正确匹配点对,就可求解出仿射变换参数。

4.4.7 鲁棒估计仿射变换模型参数

将经过前面几步得到的兴趣点对称为假设对应点对,其总数为 N 。下面介绍RANSAC算法估计仿射变换参数的具体做法:

- (1) 从假设对应点对中随机选取三组对应点对估计出一个仿射变换模型 A 。
- (2) 设 x_1 和 x_2 为任意一组假设对应点对，点 x'_1 和 x'_2 分别为通过估计的仿射变换模型 A 计算出在对应图像中的对应点，其关系如图 4-7 所示。

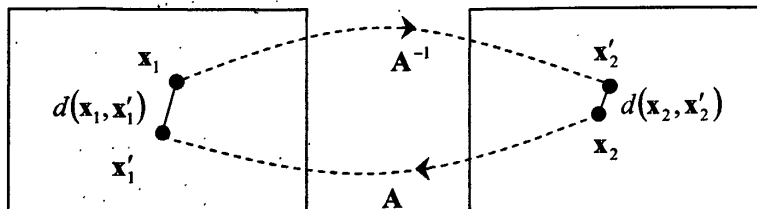


图 4-7 两幅图像对应点及估计点示意图

按公式(4.29)计算距离 $d(x_1, x'_1)$ 和 $d(x_2, x'_2)$ 。如果 $d(x_1, x'_1)$ 和 $d(x_2, x'_2)$ 均小于给定的阈值，则认为 x_1 和 x_2 为内点，否则为外点。统计所有假设对应点对中内点数。

$$\begin{aligned} d(x_1, x'_1) &= \|x_1 - Ax'_1\| \\ d(x_2, x'_2) &= \|x_2 - A^{-1}x'_2\| \end{aligned} \quad (4.29)$$

- (3) 重新随机选三组假设对应点对，重复前两步骤。
- (4) 选择具有最多内点数的 A 作为最终估计的仿射变换模型。

4.4.8 消除像机自运动

获得两幅图像之间的仿射变换参数后，根据仿射变换模型对应关系对其中一幅图像进行仿射变换，达到消除像机运动的目的。

4.5 像机自运动消除实验结果及分析

为了验证算法的有效性，选取四组无人机平台上像机对其他景物拍摄的序列图像进行计算。实验中分别利用 CMCGFD 算法和 SIFT 算法对两帧图像进行配准，以消除像机自运动。然后将已配准的图像相减得到与像机存在相对运动的目标，并与直接相减算法结果进行比较。

4.5.1 机场背景图像

图 4-8 中 a1 和 b1 是无人机平台上像机对地面拍摄的序列图像中的第一帧和第十五帧，图中有五辆汽车和像机存在相对运动。图 a2 和 b2 分别为图 a1 和图 b1 利用 CMCGFD 算法提取兴趣点的结果，其中箭头方向表示该点的特征方向、箭头

长度表示三倍特征尺度大小。图 a3 和 b3 为两帧图像兴趣点匹配结果。

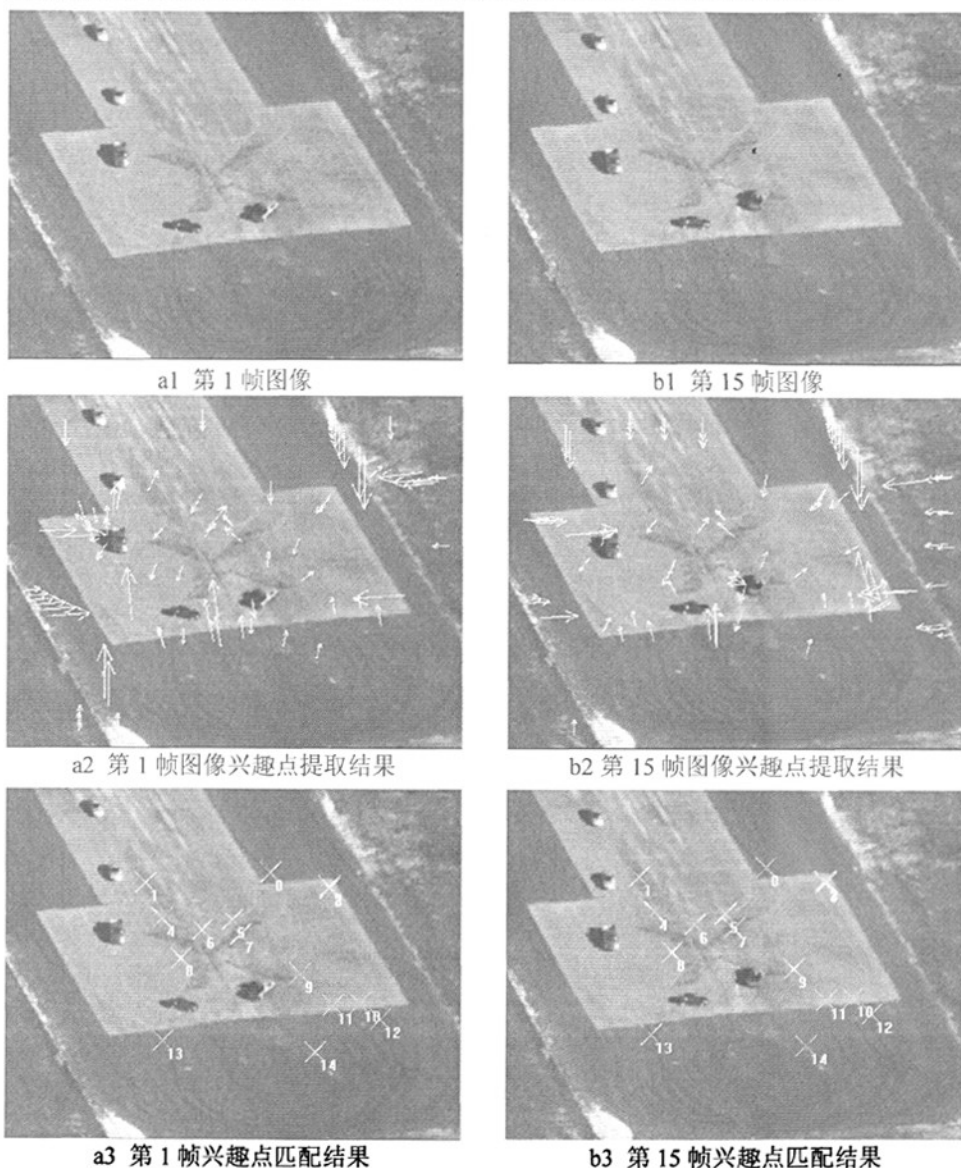


图 4-8 无人机平台图像兴趣点提取匹配结果

图 4-9 为无人机平台像机自运动消除结果对比。图 4-9.a 为图 4-8.a1 和图 4-8.b1 直接相减的结果，可以看出，由于像机自身存在运动，除了运动的汽车以外，机场 T 型头等也被检测出来。图 4-9.b 为图 4-8.a1 和图 4-8.b1 经过 SIFT 算法配准后，将图 4-8.a1 进行仿射变换再与图 4-8.b1 做差运算的结果，可以看出像机自运动消除不够准确，虽然所有运动目标均被检测出来，但是仍然存在较多背景杂波。图 4-9.c 为图 4-8.a1 和图 4-8.b1 经过 CMCDFD 算法配准后，将图 4-8.a1 进行仿射变换再与图 4-8.b1 再做差运算的结果，所有运动目标均被检测出来，背景杂波相对

于 SIFT 算法较少，像机自运动消除效果优于 SIFT 算法。



a 直接相减



b SIFT 算法配准后相减结果



c CMCGFD 算法配准后相减结果

图 4-9 无人机平台像机自运动消除结果对比

4.5.2 丛林背景图像

图 4-10 为丛林图像像机自运动消除结果对比，其中图 4-10.a 和图 4-10.b 分别为丛林序列图像中的第一帧与第十帧。可以看出，有两部汽车和像机存在相对运动。图 4-10.c 为两幅图像直接相减的结果，结果中除了目标以外，还包含较多背景杂波。图 4-10.d 为 SIFT 算法配准后再做差运算的结果，可以看出像机自运动消除不够准确，图像右侧仍然存在较多背景杂波。图 4-10.e 为采用 CMCGFD 算法配

准后再做差运算的结果，像机自运动消除结果要优于 SIFT 算法结果。

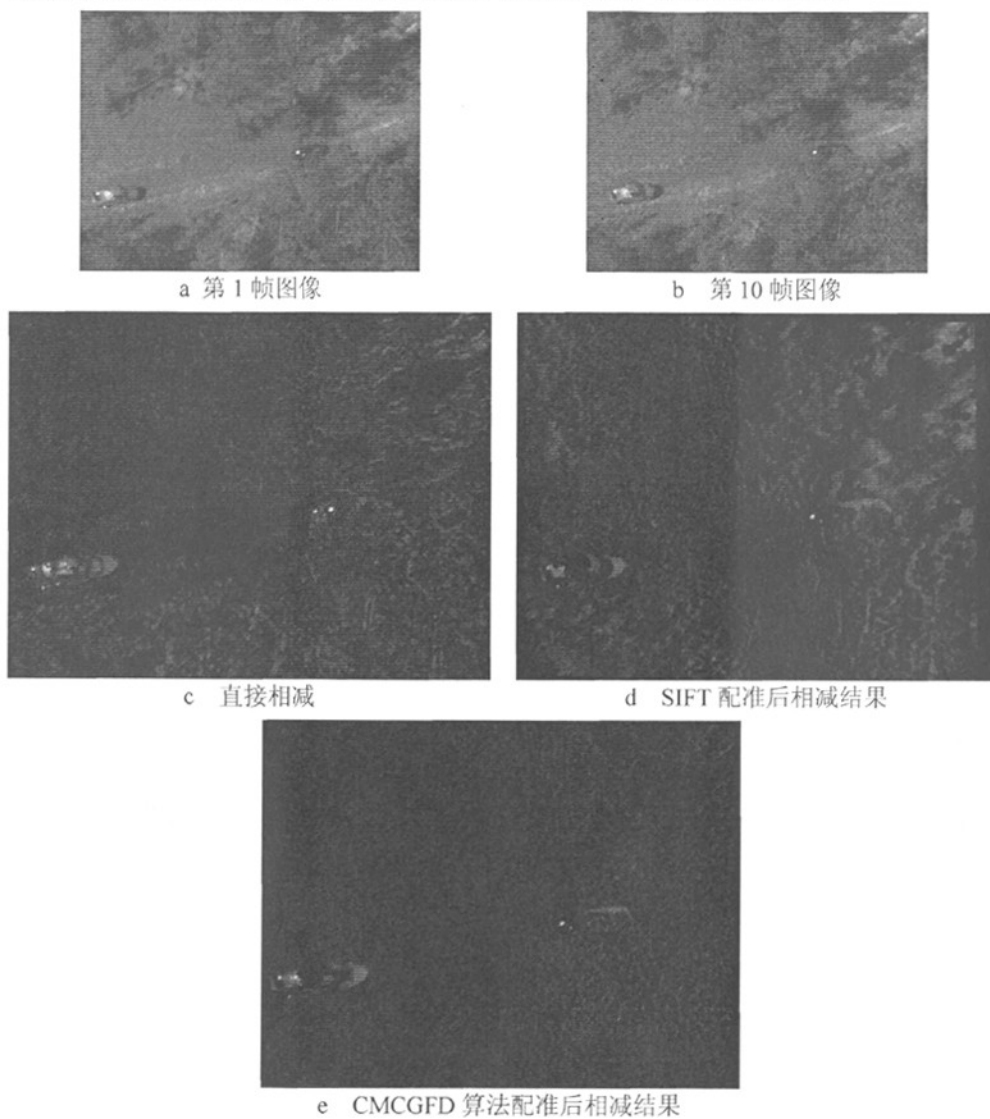


图 4-10 丛林背景图像像机自运动消除结果对比

4.5.3 山地背景图像

图 4-11 为山地图像像机自运动消除结果对比，其中图 4-11.a 和图 4-11.b 分别为高山背景下飞机目标序列图像中第一帧与第十五帧。图中有一架飞机与像机存在相对运动。图 4-11.c 为两帧图像直接相减的结果，检测结果中基本看不出飞机目标所在位置，包含较多背景杂波。图 4-11.d 为 SIFT 算法配准后再做差运算的结果，像机自运动消除存在一定误差，虽然可以清晰地分辨飞机目标，但检测结果

中仍存在一些背景杂波。图 4-11.e 为采用 CMCGFD 算法配准后再做差运算的结果，背景杂波少于 SIFT 算法检测结果，像机自运动消除效果优于 SIFT 算法。

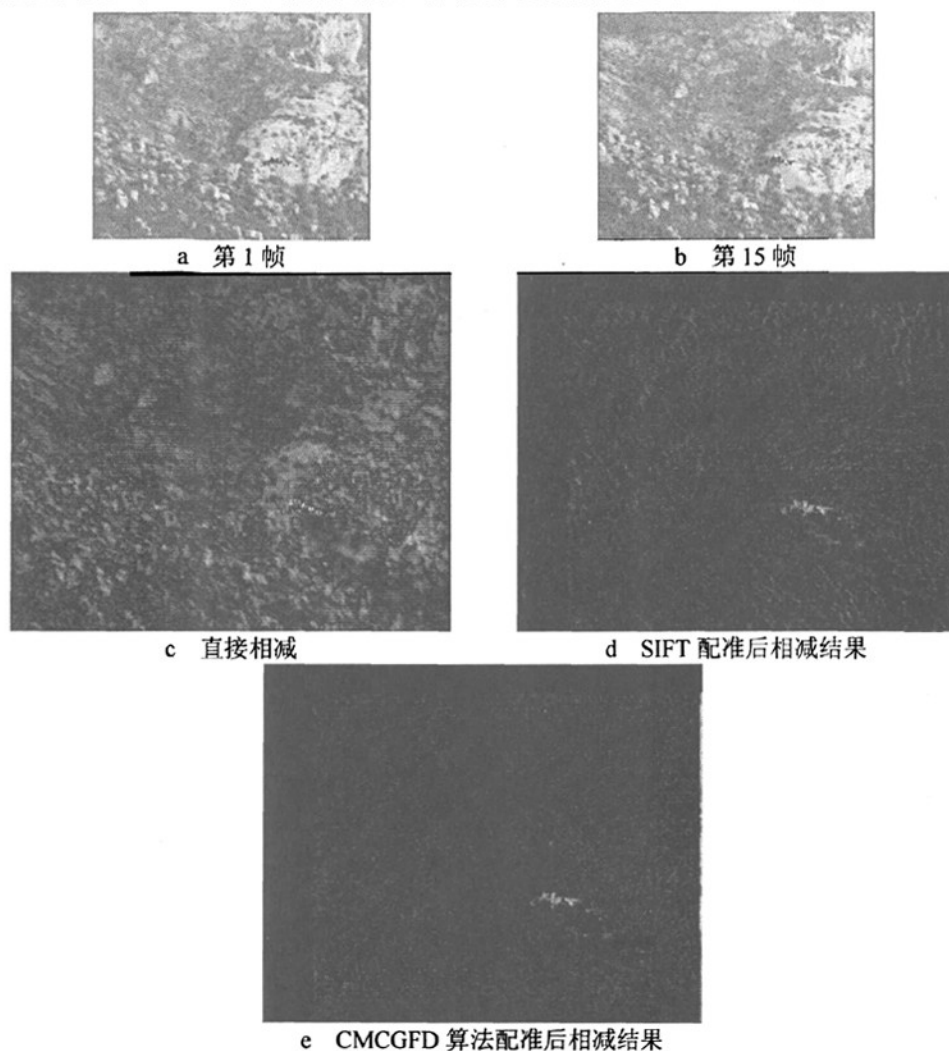


图 4-11 高山背景图像像机自运动消除结果对比

4.5.4 岛屿图像

图 4-12 为岛屿图像像机自运动消除结果对比，其中图 4-12.a 和图 4-12.b 分别为岛屿序列图像中第一帧与第十帧。图 4-12.c 为两帧图像直接相减的结果，其中包含较多背景杂波。图 4-12.d 为图 4-12.a 和图 4-12.b 经 SIFT 算法配准后，再将图 4-12.a 进行仿射变换与图 4-12.b 做差运算的结果，可以看出像机自运动消除存在一定误差，岛屿右边缘的匹配误差较大。图 4-12.e 为为图 4-12.a 和图 4-12.b 经 CMCGFD 算法配准后，再将图 4-12.a 进行仿射变换与图 4-12.b 做差运算的结果，

可以看出像机自运动消除效果优于 SIFT 算法。

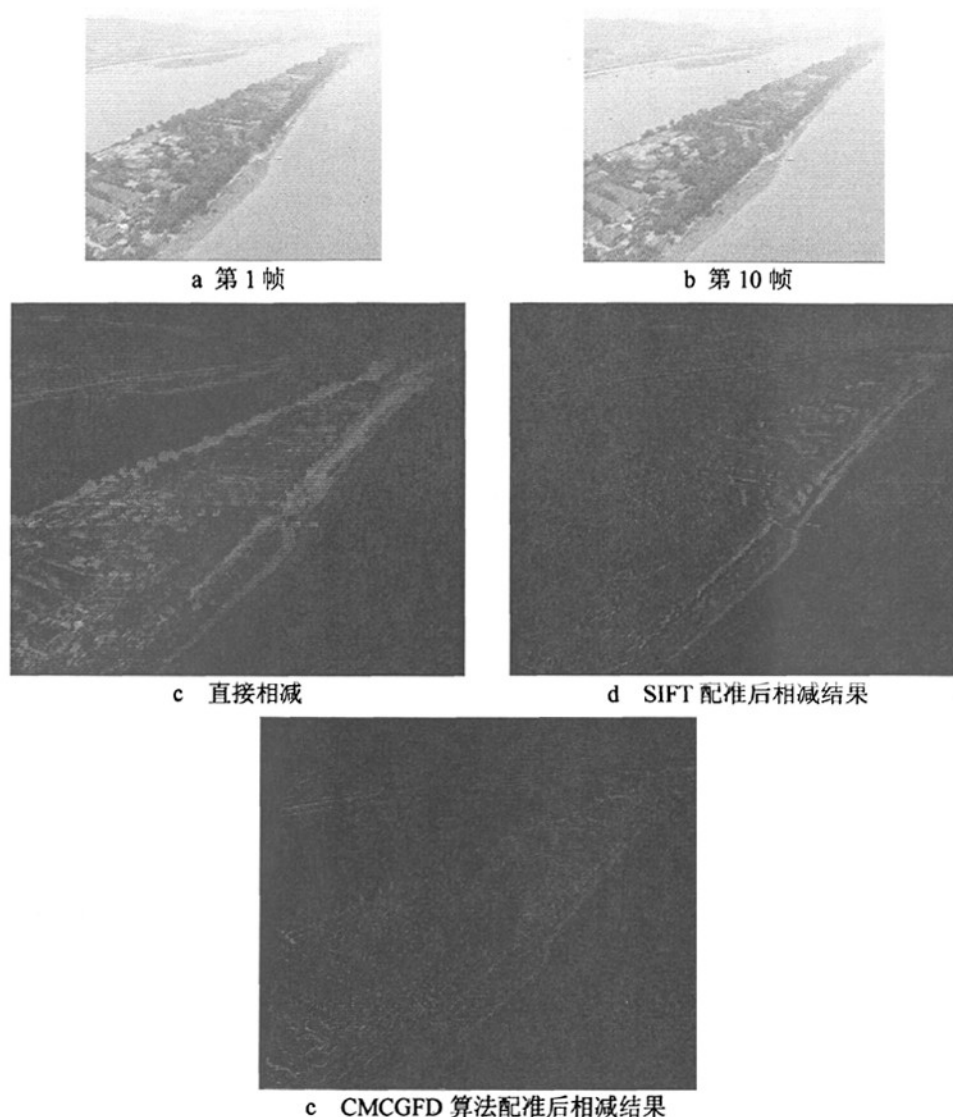


图 4-12 岛屿背景图像像机自运动消除结果对比

4.6 本章小结

像机自运动消除是机器视觉、图像理解等方面非常重要的研究内容之一，在军事及工业等领域都有着广泛的应用前景。本章针对该问题进行了系统研究。

本章首先简要介绍了常用坐标系及其变换关系，分析了图像仿射变换的数学模型及适用条件，然后在分析尺度空间理论基础上，提出了一种基于 Gabor 特征描述的像机自运动消除算法。由论文分析可知，兴趣点的特征尺度具有尺度不变

性，其特征方向具有旋转不变性。而 Gabor 函数被认为是高等动物视觉皮层简单细胞感受野的最佳近似，利用 Gabor 函数构造的特征描述具有很好的区分性。新算法将兴趣点的特征尺度和特征方向作为基准参数，对兴趣点构造 Gabor 特征描述向量，该特征描述向量不但具有尺度、旋转不变性，而且还具有较好的区分性，所以新算法可以有效地提高图像配准可靠性和准确性。为了检验新算法的性能，将 CMCGFD 算法用于实际图像配准以消除像机自运动，并将结果与 SIFT 配准算法结果进行了比较分析。实验结果表明，CMCGFD 算法比 SIFT 算法的检测结果更为理想，可以更有效地消除像机自身运动。

第五章 基于差异性加权的运动目标跟踪方法研究

运动目标跟踪技术是计算机视觉研究中的热点,在精确制导^[190]、战场侦察^[191]、安控^[192]、智能交通^[193,194]以及医疗图像处理^[195,196]等领域都有重要的实用价值。复杂背景下对指定运动目标进行稳健跟踪,又是此关键技术中的一个极具挑战性的难题。论文就此问题进行了研究,提出了一种复杂背景下目标描述区域差异性权值计算方法,并在此基础上提出了两种跟踪算法。实验结果表明,权值计算结果基本符合人类视觉反应,两种新的跟踪算法均比传统跟踪算法更有效稳定。

本章内容安排如下:5.1节提出了单帧目标差异性权值计算方法;5.2节介绍差异性加权最小二乘匹配算法原理及大致步骤;5.3节给出了差异性加权的最小二乘匹配算法的实验结果,并与经典算法进行了比较分析;5.4节介绍了加权 Mean Shift 算法,并对其收敛性进行了证明;5.5节介绍了差异性权值 Mean Shift 跟踪算法原理;5.5.2节给出了差异性权值 Mean Shift 跟踪算法的实验结果,并与经典算法进行了比较分析;5.7节对本章进行了小结。

5.1 复杂背景下单帧目标区域差异性权值确定方法

对于目标跟踪问题,通常在初始时刻以目标为中心选择一个矩形或圆形区域,建立目标特征描述。对于非均匀目标,由于目标区域自身存在不一致性,目标中某些区域和背景差异较大,而某些区域和背景差异较小。其中目标区域中和背景差异大的区域对于跟踪算法的有效性、稳定性贡献比较大。而矩形或圆形区域中的背景部分对目标跟踪会带来一定的误差,应该尽量排除。因此,在初始选取矩形或圆形范围内,应对目标中和背景差异性较大的区域赋予较大的权值,差异性较小的区域赋予较小的权值,而对背景区域赋予大小为零的权值。

基于以上原则,本节提出了一种单帧目标区域差异性权值计算方法。该方法以目标为中心选择内区域和外围区域,其中内区域要求将整个目标包含在内,如图5-1左图所示,包含飞机目标的红色小矩形为内区域,而红色大矩形与红色小矩形之间的区域为外围区域。



图 5-1 内外区域选择示意图

因为外围区域和内区域直接相连，而且外围区域不包含目标信息，所以外围区域内的信息和内区域中背景信息最为接近，如图5-1右图所示，绿色线包围的外围区域和内区域中两个区域的颜色特性非常相近，而且空间位置相邻。同理黑色、黄色、深蓝色和天蓝色包围均具有上述性质。如果在外围区域中统计颜色特性相近的区域作为参考背景（如绿色包络线中的外围区域），则可以有效地剔除内区域中具有相似特性的背景部分（如绿色包络线中的内区域）。

为了更为有效地统计内区域和外围区域中目标信息和背景信息，我们选择C-Mean聚类算法^[197]将空间相近颜色相似的区域划分为同一区域。根据内、外区域中聚类区域的差异性，剔除内区域中背景部分，同时给内区域中不同部分赋予不同权值。

下面以彩色图像为例，说明权值计算方法的具体步骤。对于灰度图像，思路基本一样，只需要将特征空间降为三维即可，其中灰度图像三维特征空间包括像素点的 x 坐标、 y 坐标和该点的灰度值。

5.1.1 目标区域差异性权值的具体计算步骤

5.1.1.1 选择目标内外区域

以指定目标为中心选取内区域和外围区域，如图5-1所示。

5.1.1.2 确定聚类数目及初始聚类中心

对内区域中每个像素点建立五维特征向量 $\mathbf{v}=(x,y,r,g,b)$ ，其中 x 和 y 分别为该像素点的横、纵坐标， r 、 g 、 b 分别为该点的红、绿、蓝颜色分量。点集 $S=(\mathbf{v}_1,\mathbf{v}_2,\dots,\mathbf{v}_n)$ 中任意一点 \mathbf{v}_i 的邻域定义为以 \mathbf{v}_i 为中心、 r 为半径的超球（ r 为预先选取的阈值）的区域。 \mathbf{v}_i 邻域内的数据点密度定义为：

$$D(\mathbf{v}_i) = \sum_{j=1}^N u(r - \|\mathbf{v}_i - \mathbf{v}_j\|) \quad (5.1)$$

其中 N 为选取区域中像素点的总数， $u(a) = \begin{cases} 1 & a \geq 0 \\ 0 & a < 0 \end{cases}$ 。如果数据点 \mathbf{v}_p 具有

最大的分布密度，即 $D(\mathbf{v}_p) = \max_{i=1}^N D(\mathbf{v}_i)$ ，那么第一个聚类中心为：

$$Z_1 = \frac{\sum_{\mathbf{v}_j \in S_1} [\mathbf{v}_j \cdot D(\mathbf{v}_j)]}{\sum_{\mathbf{v}_j \in S_1} [D(\mathbf{v}_j)]} \quad (5.2)$$

其中 S_1 为点 \mathbf{v}_p 邻域内的数据点集。找出第一个聚类中心后，接着在集合 $\{S - S_1\}$ 里查找具最大密度函数值的数据点，并将其邻域内的数据点的质心作为第

二个聚类中心 Z_2 。重复这样的过程，直至 $\{S - \sum_k S_k\}$ 为空集， Z_i ($i=1,2,3,\dots,k$) 为聚类中心， k 为目标区域的聚类数目，。

5.1.1.3 C-Mean聚类及区域分割

根据初始聚类数目及聚类中心，首先利用C-Mean聚类算法分别对内区域和外围区域进行聚类，然后利用形态学闭运算^[198]填充内区域聚类结果 S_i ($i=1,\dots,k$) 中的细小“空洞”。接着利用序贯算法^[199]将空间不连通的同一聚类区域重新分割为不同区域，最终形成的区域聚类结果记为 C_i ($i=1,\dots,n$)， n 为分割后的区域数目，其中 $n \geq k$ 。

5.1.1.4 确定内区域中背景部分

假设分割好的区域如图5-2所示，小正方形内部表示内区域，而大小正方形之间的区域为外围区域。

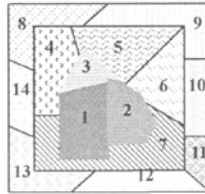


图 5-2 分割后内外区域示意图

由于外围区域和内区域直接相连，而且外围区域不包含任何目标区域，所以内区域中和外围区域相邻的分割区域为背景的可能性较大。根据以上原则可以在一定程度上分割出背景部分。具体做法为：对经过5.1.1.3节分割后的内区域（如区域1~区域7），判断其内每个区域是否包含内区域矩形边缘。如果包含内区域矩形边缘，则进行如下处理（以区域7为例进行说明）：

1、判断区域7和外围区域中哪些区域相邻接，这里为10、11、12、13和14号区域；

2、分别计算区域7、10、11、12、13和14的红、绿、蓝三种颜色的均值 Mr 、 Mg 、 Mb ，按(5.3)式判断区域7和这五个区域中哪个区域最为接近。

$$T = \min(|Mr_7 - Mr_i| + |Mg_7 - Mg_i| + |Mb_7 - Mb_i|) \quad i=10,11,12,13,14 \quad (5.3)$$

3、如果 T 大于预先指定的阈值，则认为区域7应属于背景部分。

对内区域中所有分割区域进行判断，将所有属于背景部分的区域划分为 C_{bg} 区域，其权值 $W(x_i)$ 设为0， x_i 表示内区域像素点坐标。

5.1.1.5 计算初始帧中目标的差异性权值

为了降低光照变化对权值计算的影响，对像素点的三个颜色分量进行归一化

处理。按(5.4)式计算内矩形区域中背景部分 C_{bg} 三个归一化颜色分量的均值, 分别记为 $\bar{r}_{C_{bg}}$ 、 $\bar{g}_{C_{bg}}$ 、 $\bar{b}_{C_{bg}}$ 。同样按(5.4)式分别计算内矩形区域中每个目标分割区域 C_i ($i=1, \dots, n_1$) 三个归一化颜色分量的均值, 分别记为 \bar{r}_i 、 \bar{g}_i 、 \bar{b}_i , 其中 n_1 为内区域中属于目标的区域总数。

$$\begin{aligned}\bar{r}_i &= \sum_{i \in C} \frac{r_i}{(r_i + g_i + b_i)} \\ \bar{g}_i &= \sum_{i \in C} \frac{g_i}{(r_i + g_i + b_i)} \\ \bar{b}_i &= \sum_{i \in C} \frac{b_i}{(r_i + g_i + b_i)}\end{aligned}\quad (5.4)$$

按(5.5)式计算每个目标分割区域 C_i 的权值 $W(\mathbf{x}_i)$, \mathbf{x}_i 表示内区域像素点坐标。

$$W(\mathbf{x}_i) = 1 - 1 / (1 + |\bar{r}_i - \bar{r}_{C_{bg}}| + |\bar{g}_i - \bar{g}_{C_{bg}}| + |\bar{b}_i - \bar{b}_{C_{bg}}|) \quad (5.5)$$

经过 5.1.1.4 节处理后可以剔除内区域中大多数背景部分, 但剩下的区域中可能还存在少量的残余背景部分。由于目标区域特性和内区域中背景部分特性差别较大, 而经过 5.1.1.4 节方法剔除背景后剩下的区域中背景部分的特性和内区域中背景部分的特性差别较小, 所以(5.5)式会给残余的背景赋予较低的权值, 而给目标区域赋予较高的权值。

5.1.2 目标区域差异性权值的计算结果

图 5-3 为复杂背景下飞机目标权值计算示意图。图 5-3.a 为原始图像, 其中小矩形表示包含目标的内区域, 大矩形和小矩形之间的区域表示外围区域。图 5-3.b 为内区域放大图像。图 5-3.c 为内区域图像经过 5.1.1.2 节和 5.1.1.3 节处理后的结果, 用不同的灰度代表不同的聚类区域。图 5-3.d 为经过 5.1.1.5 节计算出来的内区域权值结果, 其中越接近白色的区域, 说明该区域的权值越高。可以看出, 权值越大的区域相对于背景的差别越大。由图 5-3.b 和图 5-3.d 可以看出, 权值计算结果与人类视觉反应基本一致。

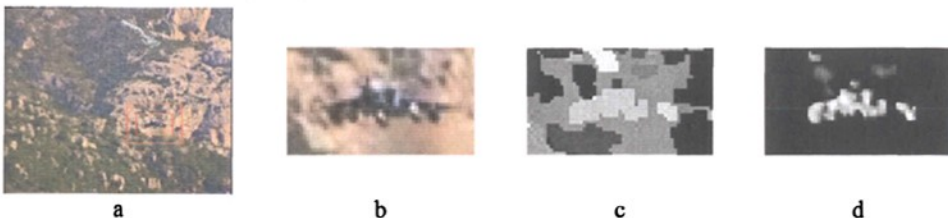


图 5-3 复杂背景图像中飞机目标差异性权值计算示意图

图 5-4 为不同复杂背景图像中飞机目标差异性权值计算结果。权值计算结果图

灰度值越接近 255 的区域，说明该区域权值越高。可以看出，权值计算结果与人类视觉反应基本一致。

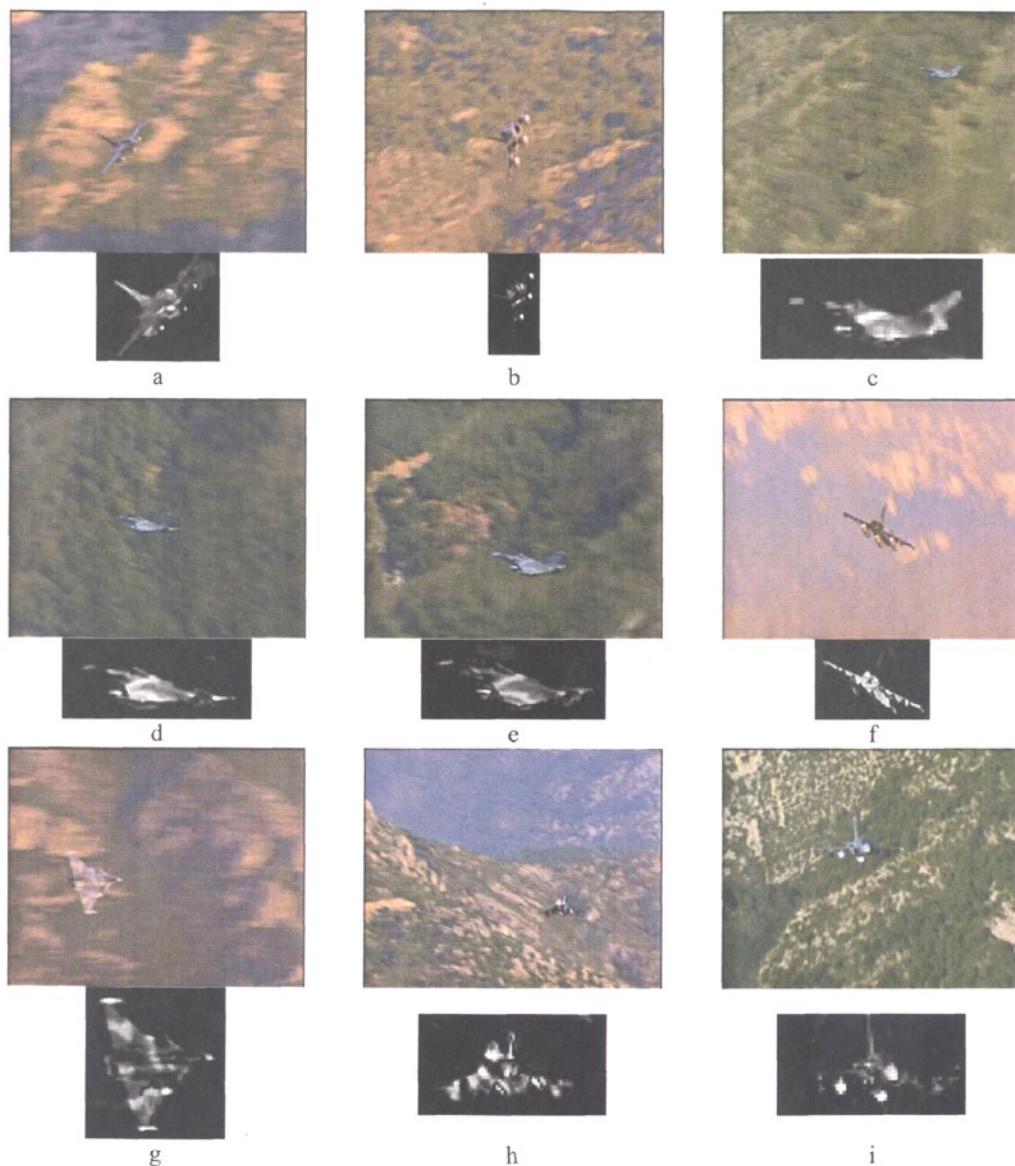


图 5-4 不同复杂背景图像目标差异性权值计算结果

5.2 差异性加权最小二乘图像匹配跟踪算法

最小二乘图像匹配方法具有灵活、可靠和精度高的优点，因而受到了广泛的重视。一些学者在 LSIM 算法的匹配精度和匹配速度方面做了卓有成效的工作，如 Zhang 等采用分级匹配策略来提高匹配的可靠性和精度^[200]；Lai 等给出了一种基于

自适应加权的匹配算法^[201]，以解决光照变化对匹配的影响；李峰等给出了一种基于变形系数相关的 LSIM 算法来提高匹配速度^[202]。

针对复杂背景下目标稳定跟踪问题，本文将 5.1 节的差异性权值计算方法和最小二乘匹配算法结合起来，提出了一种差异性加权最小二乘影像匹配运动目标跟踪算法（Diversity Weighted Least Squares Image Matching，简称 DWLSIM）。该算法根据目标与背景的差异大小对跟踪窗口中的像素点进行加权，有效地降低了窗口区域中背景部分的变化对目标跟踪的影响。

5.2.1 DWLSIM 的算法原理

运动目标跟踪过程中，一般取初始帧目标区域或理想目标模板作为当前图像跟踪匹配的基准图。由于目标相对像机运动，在当前图像上的目标与基准图上目标区域之间存在随机误差和系统误差。随机误差一般来源于随机噪声；系统误差主要为几何畸变。

当拍摄过程中摄像机的内参数不发生变化、无人机距离拍摄区域远时，视场远远大于目标的尺度，不共面的特征物点在远场图像中可以被近似认为是共面的，此时可以用仿射变换描述当前图中目标区域与基准图中目标区域之间的变换关系，具体表示如下：

$$\begin{cases} x_{t+1} = r_0 x_t + r_1 y_t + r_4 \\ y_{t+1} = r_2 x_t + r_3 y_t + r_5 \end{cases} \quad (5.6)$$

其中 (x_t, y_t) 为基准图上目标区域中像素点的坐标， (x_{t+1}, y_{t+1}) 为当前图上目标区域中对应于基准图上目标同名像素点的坐标， r_0 、 r_1 、 r_2 、 r_3 、 r_4 和 r_5 为仿射变换参数。忽略噪声影响后，当前图中目标区域与基准图中目标区域之间的灰度分布关系可用如下模型近似表示：

$$I_t(x, y) \approx I_{t+1}(r_0 x + r_1 y + r_4, r_2 x + r_3 y + r_5) \quad (5.7)$$

I_t 、 I_{t+1} 分别为基准图与当前图的灰度分布函数。

对于目标跟踪的情况，在窗口区域内同时有目标点和背景点，而只有目标点是满足(5.7)式。在此情况下，目标上任一点 (x, y) 的灰度误差为：

$$v = W \times (I_{t+1} - I_t) = W \times [I_{t+1}(r_0 x + r_1 y + r_4, r_2 x + r_3 y + r_5) - I_t] \quad (5.8)$$

其中 W 是通过 5.1 节方法获得目标区域差异性权值。将(5.8)式线性化后得到：

$$\begin{aligned} v = W \times [(I_{t+1})_x dr_4 + x(I_{t+1})_x dr_0 + y(I_{t+1})_x dr_1 + \\ (I_{t+1})_y dr_5 + x(I_{t+1})_y dr_2 + y(I_{t+1})_y dr_3 - \Delta I] \end{aligned} \quad (5.9)$$

其中 $(I_{t+1})_x$ 与 $(I_{t+1})_y$ 分别为当前图 I_{t+1} 的 x 方向和 y 方向梯度， dr_0, \dots, dr_5 是待定的参数改正值， $\Delta I = I_{t+1}(x, y) - I_t(x, y)$ 为图像灰度差。

当相邻两帧中目标区域变化不大时, 后续帧图像 I_{t+1} 中目标区域仿射变换参数初值 $[r_0^{t+1}, r_1^{t+1}, r_2^{t+1}, r_3^{t+1}, r_4^{t+1}, r_5^{t+1}]$ 可以取为 $[1, 0, 0, 1, 0, 0]$ 。当相邻两帧中目标的位置和形状变化较大时, 特别当目标平移运动很大时 (即仿射变换参数中的 r_4 和 r_5 较大), 并不满足泰勒展开的基本条件, 仍然通过对 (5.8) 式线性化求解目标的位置, 会导致跟踪失败。这时需要对目标运动进行预测, 获得当前帧仿射变换参数的初始值。

对目标区域内所有点统计灰度误差, 根据最小二乘图像匹配原理: $\sum v^2 = \min$, 可以得到关于参数改正值 dr_0, \dots, dr_5 的线性方程组, 由此方程组解出 dr_0, \dots, dr_5 , 最终可以得到 $r_0, r_1, r_2, r_3, r_4, r_5$ 的迭代解。最后根据式 (5.6) 式得到当前图中目标的位置。

在整个跟踪过程中, 对当前帧均以前一帧中的目标区域图像作为目标模板。获得当前帧目标位置后, 利用当前帧最终得到的目标区域信息, 结合前一帧目标差异性权值, 更新当前帧目标的差异性权值。

5.2.2 DWLSIM 算法步骤

算法大致步骤如图 5-5 所示, 下面对各个步骤进行说明如下:

1. 利用 5.1 节提出的方法计算目标区域差异性权值;
2. 根据当前的仿射变换参数, 利用 (5.6) 式对当前图像 I_2 进行仿射变换;
3. 由于换算得到的目标阵列的坐标点不是整像素点, 所以需要重采样得到 I_2' ;
4. 计算基准图 I_1 与仿射变换后的目标图像 I_2' 的相关系数, 判断是否需要继续迭代;
5. 采用最小二乘图像匹配, 求解本次迭代中仿射变换参数的改正值 $dr_0^i, dr_1^i, dr_2^i, dr_3^i, dr_4^i, dr_5^i$;
6. 利用 (5.10) 式计算本次迭代的仿射变换参数;

$$\begin{cases} r_0^i = r_0^{i-1} + r_0^{i-1} dr_0^i + r_2^{i-1} dr_1^i \\ r_1^i = r_1^{i-1} + r_1^{i-1} dr_0^i + r_2^{i-1} dr_1^i \\ r_2^i = r_2^{i-1} + r_0^{i-1} dr_2^i + r_2^{i-1} dr_3^i \\ r_3^i = r_3^{i-1} + r_1^{i-1} dr_2^i + r_3^{i-1} dr_3^i \\ r_4^i = r_4^{i-1} + dr_4^i + r_4^{i-1} dr_0^i + r_5^{i-1} dr_0^i \\ r_5^i = r_5^{i-1} + dr_5^i + r_4^{i-1} dr_2^i + r_5^{i-1} dr_3^i \end{cases} \quad (5.10)$$

7. 根据本次仿射变换参数, 利用式 (5.6) 对当前图像进行仿射变换并重采样, 返回第四步。如果停止迭代, 输出当前帧中目标的位置及仿射变换参数, 更新权值。

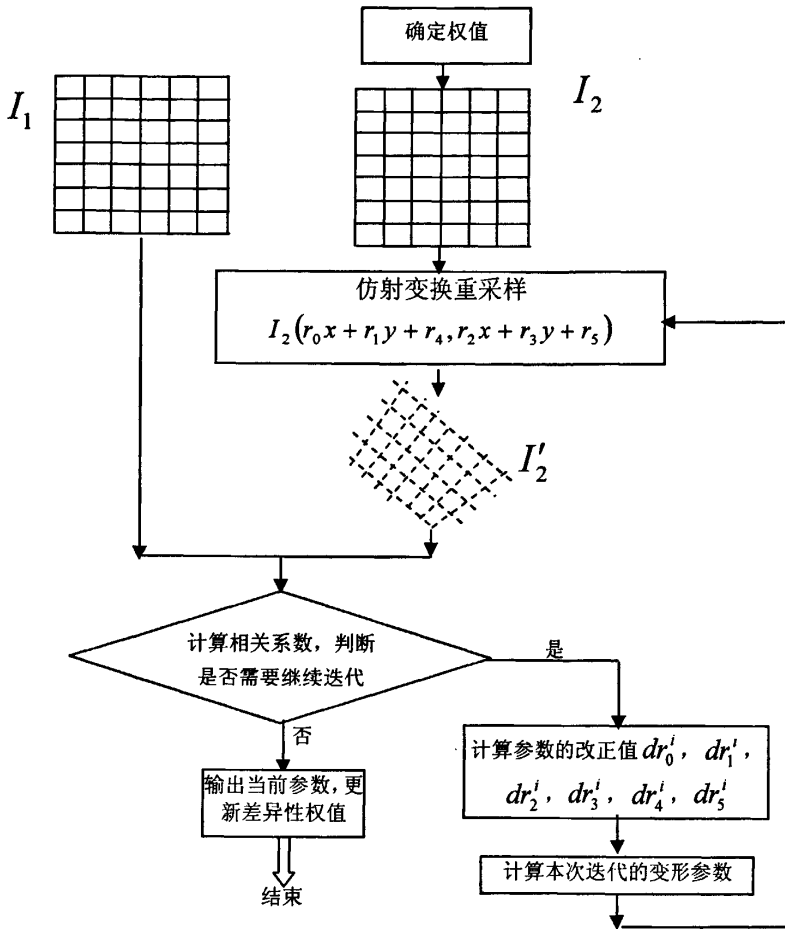


图 5-5 DWLSIM 算法步骤

5.3 DWLSIM 算法实验结果分析

仿射变换是一个非奇异线性变换与一个平移变换的复合, 公式(5.6)可用矩阵表示为:

$$\begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} = \begin{bmatrix} r_0 & r_1 & r_4 \\ r_2 & r_3 & r_5 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (5.11)$$

其中 \mathbf{A} 是一个 2×2 的非奇异矩阵, 其几何效应相当于两个基本变换—旋转和非均匀缩放的复合。根据矩阵奇异值分解原理^[203], 仿射矩阵 \mathbf{A} 总能分解为:

$$\mathbf{A} = R(\theta)R(-\phi)DR(\phi) \quad (5.12)$$

其中 $R(\theta)$ 和 $R(\phi)$ 分别表示转角为 θ 和 ϕ 的旋转, 而 \mathbf{D} 为对角矩阵:

$$\mathbf{D} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \quad (5.13)$$

因此,仿射矩阵 \mathbf{A} 可看成是一个非均匀缩放变换 $R(-\phi)DR(\phi)$ 和一个旋转 $R(\theta)$ 的复合变换。而非均匀缩放变换 $R(-\phi)DR(\phi)$ 是一个旋转 $R(\phi)$, 加上对已旋转的 x 和 y 方向分别进行比例因子 λ_1 和 λ_2 的缩放, 再加上一个回转 $R(-\phi)$ 的复合。仿射变换的本质是在一个特定角的两个垂直方向上进行缩放, 如图 5-6 所示。

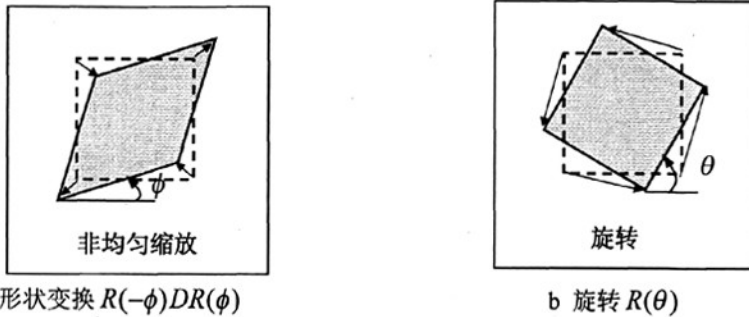


图 5-6 平面仿射变换失真

5.3.1 跟踪精度仿真实验

根据上述分析, 仿真图像实验中分别考察了平移、平移加旋转、完整仿射变换以及加噪声的完整仿射变换下 LSIM 算法和 DWLSIM 算法的跟踪精度。

仿真图像中的目标用十字丝模拟, 同时为了模拟复杂背景, 对每一帧图像人为地加入不同的无规则的灰色区域。十字丝区域在不同参数条件下按照仿射公式 (5.11) 变换。

5.3.1.1 纯平移变换下跟踪结果

平移变换是最常见的一种变换。初始帧中十字丝中心坐标为 (50,50), 取 $\theta = \phi = 0^\circ$, $\lambda_1 = \lambda_2 = 1$, r_4, r_5 取不同的值, 按 (5.11) 式对十字丝目标区域进行变换, 同时对每帧图像加入不同干扰模拟复杂背景, 如图 5-7。实验中, 对第 0 帧计算权值并建立目标模板, 分别用 LSIM 算法和 DWLSIM 算法对后续帧中十字丝中心进行跟踪, 实验结果见表 5-1, 跟踪误差随平移量变化的曲线见图 5-8。由实验结果可以看出, 复杂变化背景对跟踪精度有一定影响, DWLSIM 方法的跟踪误差低于 LSIM 方法。

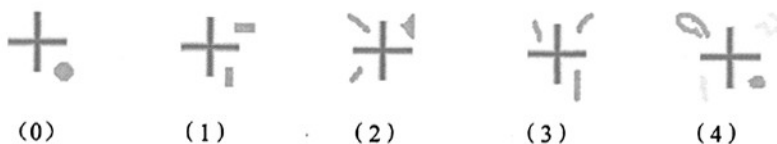


图 5-7 纯平移变换仿真序列图

表 5-1 纯平移变换下十字丝定位结果比较

| 序号 | 真值 | | LSIM 算法跟踪结果 | | | DWLSIM 算法跟踪结果 | | |
|----|-------|-------|-------------|-------|-------|---------------|-------|-------|
| | X(像素) | Y(像素) | X 方向 | Y 方向 | 误差 | X 方向 | Y 方向 | 误差 |
| 1 | 53.0 | 53.0 | 52.90 | 52.86 | 0.172 | 52.96 | 53.08 | 0.089 |
| 2 | 55.0 | 55.0 | 54.87 | 54.87 | 0.184 | 54.94 | 55.08 | 0.100 |
| 3 | 57.0 | 57.0 | 56.82 | 56.82 | 0.255 | 56.82 | 56.97 | 0.182 |
| 4 | 59.0 | 59.0 | 58.80 | 58.76 | 0.312 | 59.13 | 58.84 | 0.206 |

(表中“误差”栏为 x 和 y 方向跟踪误差的平方和的开方值, 以下各表同义, 不再赘述)

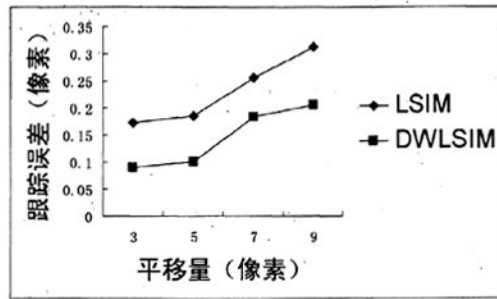


图 5-8 纯平移变换十字丝定位误差曲线

5.3.1.2 平移加旋转变换下跟踪结果

初始帧中十字丝中心坐标为(50,50), 同时加载平移和旋转变换。取 $\lambda_1 = \lambda_2 = 1$, $\phi = 0^\circ$, 而 r_4 , r_5 和 θ 取不同的值, 按公式(5.11)对十字丝目标区域进行变换, 同时对每帧图像加入不同干扰模拟复杂背景, 如图 5-9。实验中, 对第 0 帧计算权值并建立目标模板, 分别用 LSIM 算法和 DWLSIM 算法对后续帧中的十字丝中心进行跟踪, 实验结果见表 5-2, 跟踪误差随平移量变化的曲线见图 5-8。由实验可知, DWLSIM 方法的跟踪误差明显低于 LSIM 方法。

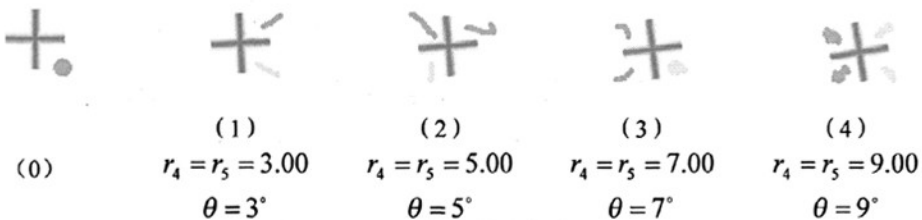


图 5-9 平移+旋转变换仿真序列图

表 5-2 平移+旋转变换下十字丝定位结果比较

| 序号 | 真值 | | 旋转角度 θ (度) | LSIM 算法跟踪结果 | | | DWLSIM 算法跟踪结果 | | |
|----|-------|-------|----------------------|-------------|-------|-------|---------------|-------|-------|
| | X(像素) | Y(像素) | | X 方向 | Y 方向 | 误差 | X 方向 | Y 方向 | 误差 |
| 1 | 53.0 | 53.0 | 3.0 | 53.23 | 53.13 | 0.264 | 53.11 | 52.92 | 0.136 |
| 2 | 55.0 | 55.0 | 5.0 | 54.74 | 54.81 | 0.322 | 54.75 | 55.09 | 0.266 |
| 3 | 57.0 | 57.0 | 7.0 | 56.66 | 57.21 | 0.400 | 56.90 | 56.75 | 0.287 |
| 4 | 59.0 | 59.0 | 9.0 | 58.55 | 59.32 | 0.552 | 58.68 | 59.21 | 0.383 |

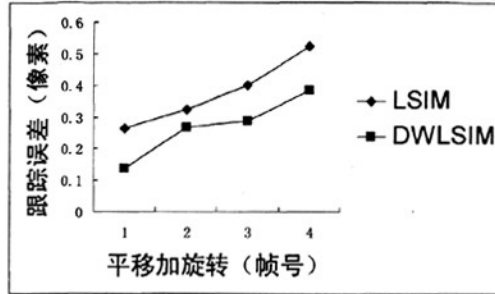


图 5-10 平移旋转变换十字丝定位误差曲线

5.3.1.3 相同噪声影响下不同模拟复杂背景仿射变换跟踪结果

考虑仿射变换情况， λ_1 ， λ_2 ， ϕ ， r_4 ， r_5 和 θ 取不同的值，同时对仿射变换序列图像均加入方差为 20%的高斯噪声，同时加入不同的干扰模拟复杂背景，如图 5-11 所示。实验中，对第 0 帧计算权值并建立目标模板，对后续帧分别采用 LSIM 方法和 DWLSIM 方法进行跟踪，跟踪结果如表 5-3 所示。由实验结果可知，两种方法均能跟踪目标，但是在相同的噪声影响下，DWLSIM 方法的跟踪误差明显低于 LSIM 方法。

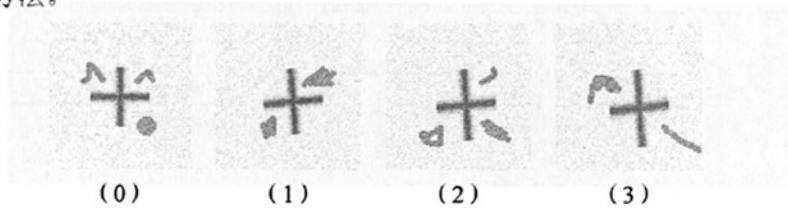


图 5-11 20%高斯噪声仿射变换下仿真序列图

表 5-3 20%高斯噪声仿射变换下十字丝定位结果比较

| 序号 | 十字丝中心真实坐标 | | LSIM 算法跟踪结果 | | | DWLSIM 算法跟踪结果 | | |
|----|-----------|-------|-------------|-------|-------|---------------|-------|-------|
| | X(像素) | Y(像素) | X 方向 | Y 方向 | 误差 | X 方向 | Y 方向 | 误差 |
| 1 | 53.0 | 53.0 | 52.75 | 53.27 | 0.367 | 52.85 | 58.87 | 0.198 |
| 2 | 55.0 | 55.0 | 54.70 | 54.67 | 0.446 | 54.78 | 54.75 | 0.333 |
| 3 | 57.0 | 57.0 | 57.42 | 56.63 | 0.560 | 57.2 | 56.64 | 0.412 |

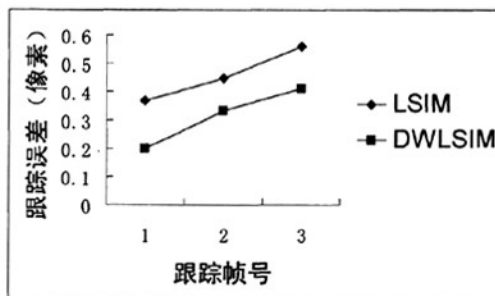


图 5-12 无噪声仿射变换下十字丝定位误差曲线

5.3.1.4 不同噪声影响下仿射变换跟踪结果

为了检验算法的抗噪性能,对初始帧图像进行仿射变换后,加入相同的干扰模拟复杂背景,并加入不同级别的高斯噪声,如图 5-13 所示。实验中,对第 0 帧计算权值并建立目标模板,对后续帧分别采用 LSIM 方法和 DWLSIM 方法进行跟踪,跟踪结果如表 5-4 所示。由实验结果可知,两种方法均能有效跟踪目标,但在相同的噪声影响下,DWLSIM 方法的跟踪误差明显低于 LSIM 方法。

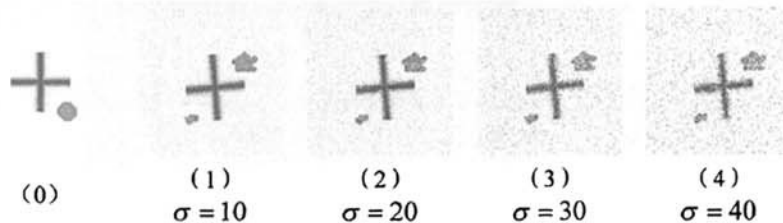


图 5-13 不同噪声仿真序列图

表 5-4 不同噪声下十字丝定位结果比较

| 序号 | LSIM 算法跟踪结果 | | | DWLSIM 算法跟踪结果 | | |
|----|-------------|-------|-------|---------------|-------|-------|
| | X 方向 | Y 方向 | 误差 | X 方向 | Y 方向 | 误差 |
| 1 | 52.70 | 52.66 | 0.340 | 53.01 | 53.00 | 0.010 |
| 2 | 52.68 | 52.63 | 0.489 | 53.02 | 52.97 | 0.036 |
| 3 | 52.71 | 52.51 | 0.569 | 52.98 | 52.95 | 0.054 |
| 4 | 52.59 | 52.59 | 0.580 | 53.01 | 53.08 | 0.081 |

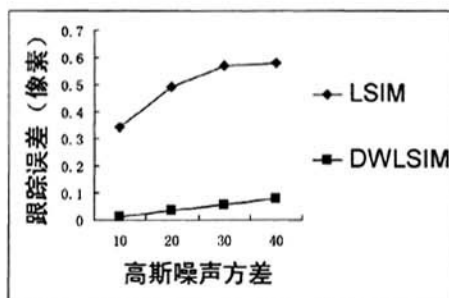


图 5-14 不同噪声下十字丝定位误差曲线

5.3.2 实际图像跟踪结果比较

利用最小二乘匹配跟踪法 (LSIM)、Mean Shift跟踪算法 (MS)、粒子滤波跟踪算法 (Partical Filter, 简称PF) 以及本文提出的DWLSIM跟踪算法,对两组复杂背景序列图像中的运动目标进行跟踪,其中用红色矩形框表示跟踪结果。

5.3.2.1 复杂背景下对飞机目标跟踪结果 (一)

图 5-15 表示为序列图像中第 0、20、40、60 帧的跟踪结果,其中 a 组为 LSIM 跟踪结果, b 组为 MS 跟踪结果, c 组为粒子滤波算法跟踪结果, d 组为 DWLSIM

跟踪结果，e 组为 DWLSIM 算法跟踪结果的局部放大图。e 组最左边图像中的红色十字为初始指定的跟踪点，其余图像中红色十字丝为该点的跟踪结果。实验中，LSIM 算法、MS 算法和粒子滤波算法均跟踪失败。而 DWLSIM 算法可以有效地跟踪运动目标。从 e 组结果可以看出，对于初始指定的一点，新算法可以一直准确地进行跟踪。

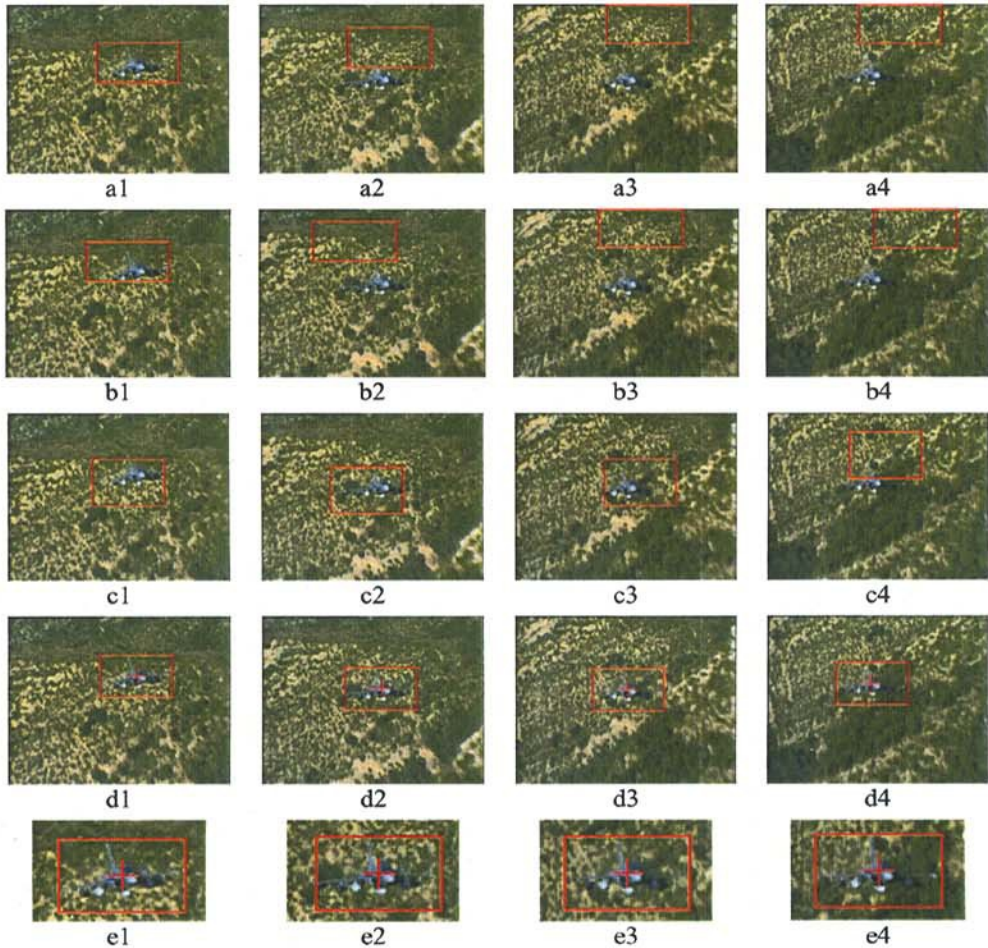


图 5-15 复杂背景下运动飞机跟踪结果比较（一）

5.3.2.2 复杂背景下对飞机目标跟踪结果（二）

图 5-16 表示为序列图像中第 0、10、20、30 帧的跟踪结果，其中 a 组为 LSIM 跟踪结果，b 组为 MS 跟踪结果，c 组为粒子滤波算法跟踪结果，d 组为 DWLSIM 跟踪结果。结果表明，LSIM 算法、MS 算法和粒子滤波算法均跟踪失败。而 DWLSIM 算法可以有效地跟踪运动目标。

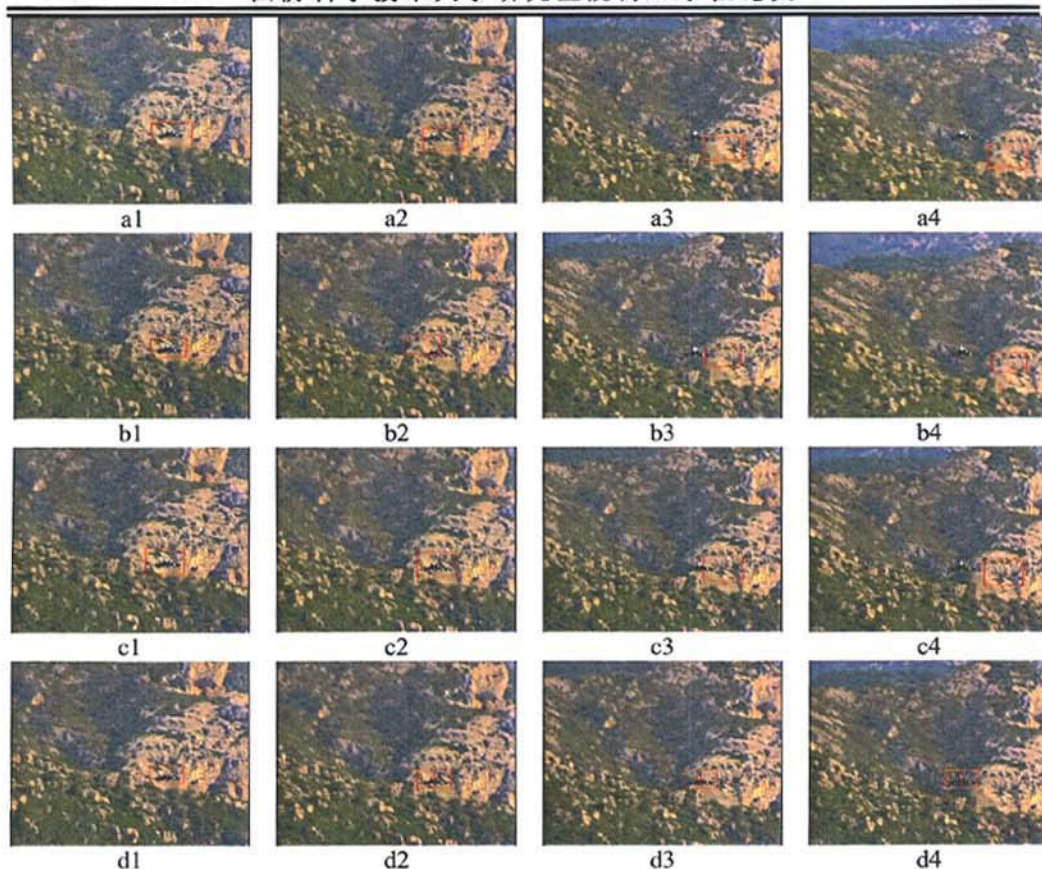


图 5-16 复杂背景下运动飞机跟踪结果比较 (二)

图 5-17 为 DWLSIM 算法对整个序列图像飞机目标跟踪结果的局部放大图像, 其中从图 5-17.1 到图 a11 分别对应序列图像的第 0、8、16、24、32、40、48、56、64、72、80 帧。图 5-17.1 中的白色十字为初始选取的一点, 而图 5-17.2 至图 5-17.11 为后序帧中该点的跟踪结果。可以看出, 虽然背景变化较大且目标存在一定的变形, DWLSIM 算法仍然可以准确地跟踪目标。

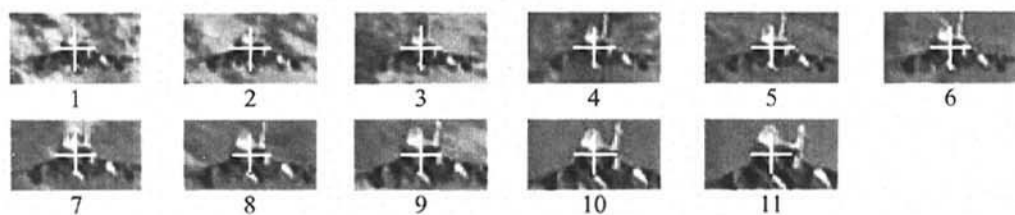


图 5-17 DWLSIM 跟踪算法结果局部放大

5.3.3 四种算法平均处理时间比较

对于两组图像，分别用 MS、LSIM、PS 和 DWLSIM 算法进行运动目标跟踪，四种算法均在 VC++6.0 平台上编程实现，计算机主频为赛扬 2.4GHz、内存为 512MB。在初始条件相同的条件下，算法平均处理时间如表 5-5 所示。可以看出，MS 虽然耗时较短但跟踪很容易失败，而 PS 算法耗时最长。本文提出的 DWLSIM 算法和 LSIM 算法耗时基本相当。

表 5-5 四种算法的平均处理时间比较 单位：毫秒

| | Ms 算法 | LSIM 算法 | PS 算法(30 个粒子) | 本文算法 |
|-------|-------|---------|---------------|-------|
| 第一组图像 | 16.5 | 47.8 | 150.7 | 30.8 |
| 第二组图像 | 17.5 | 49.2 | 200.4 | 56.01 |

5.4 加权 Mean Shift 算法

Mean Shift 算法以其无需参数、快速模式匹配等特性而被广泛应用到目标跟踪领域。关于 Mean Shift 跟踪算法的介绍请参考 2.3.4 节。对于复杂背景下运动目标跟踪问题，如果目标初始特征建立不准确，将严重影响 Mean Shift 跟踪算法的稳定性，甚至可能导致跟踪失败。针对这个问题，通常的一类做法是尽量获得目标的准确描述范围，或者对目标覆盖区域内像素点进行加权处理。

经典 Mean Shift 跟踪算法以目标为中心选择一个矩形或圆形区域，统计该区域的颜色直方图作为目标特征描述向量（这里以颜色直方图进行说明，对于灰度直方图下面介绍的理论完全适用）。当选择区域内存在较多背景区域时，如果直接将选择整个区域的颜色直方图作为目标特征描述向量，会降低目标初始特征描述的准确性，从而影响跟踪算法的可靠性。由 5.1 节知道，可以根据目标区域和背景差别的大小，得到选择区域差异性权值，其中对于选择区域中背景部分权值为零，而对于选择区域中和背景差别较大的区域权值较大。如果在统计选择区域的颜色直方图时，引入目标区域差异性权值信息，则可以得到一种突出目标且对背景有一定抑制的加权颜色直方图。将加权颜色直方图作为目标特征描述向量，可以有效提高 Mean Shift 跟踪算法的可靠性和稳定性。基于以上思路，本文提出了一种差异性权值 Mean Shift 跟踪算法，新算法可以对复杂背景中运动目标进行有效跟踪。

本节首先阐述加权 Mean Shift 过程，然后分析了加权 Mean Shift 过程收敛的权值取值条件，最后对加权 Mean Shift 过程的收敛性进行了证明。本节主要为下一节的复杂背景下差异性加权 Mean Shift 跟踪算法奠定理论基础。

5.4.1 加权 Mean Shift 过程

给定 n 个点 $\{\mathbf{x}_i\}_{i=1,\dots,n}$, 对(2.30)式进行加权得到使用核 $K(x)$ 的加权多变核密度估计为:

$$f_K(\mathbf{x}) = \frac{1}{nh^d} \sum_{i=1}^n \left[K\left(\frac{\mathbf{x}-\mathbf{x}_i}{h}\right) W(\mathbf{x}_i) \right] \quad (5.14)$$

其中 $K(x) = k(\|x\|^2)$, 而 $k(x)$ 称为 $K(x)$ 的轮廓函数, $W(\mathbf{x}_i)$ 为 \mathbf{x}_i 的权值. 核密度估计可改写为:

$$f_K(\mathbf{x}) = \frac{1}{nh^d} \sum_{i=1}^n \left[k\left(\left\|\frac{\mathbf{x}-\mathbf{x}_i}{h}\right\|^2\right) W(\mathbf{x}_i) \right] \quad (5.15)$$

如果 $k(x)$ 在 $[0, \infty)$ 上除有限个点外都可导, 那么可以定义: $g(x) = -k'(x)$, $g(x)$ 对应于一个新核 $G(x) = Cg(\|x\|^2)$, C 为正则化常数. 使用该核的加权多变量核密度估计为:

$$f_G(\mathbf{x}) = \frac{C}{nh^d} \sum_{i=1}^n \left[g\left(\left\|\frac{\mathbf{x}-\mathbf{x}_i}{h}\right\|^2\right) W(\mathbf{x}_i) \right] \quad (5.16)$$

那么使用核 $K(x)$ 的加权核密度梯度估计为:

$$\begin{aligned} \nabla f_K(\mathbf{x}) &= \frac{2}{nh^{d+2}} \sum_{i=1}^n \left[(\mathbf{x}-\mathbf{x}_i) k'\left(\left\|\frac{\mathbf{x}-\mathbf{x}_i}{h}\right\|^2\right) W(\mathbf{x}_i) \right] \\ &= \frac{2}{nh^{d+2}} \sum_{i=1}^n \left[(\mathbf{x}_i-\mathbf{x}) g\left(\left\|\frac{\mathbf{x}-\mathbf{x}_i}{h}\right\|^2\right) W(\mathbf{x}_i) \right] \\ &= \frac{2}{nh^{d+2}} \sum_{i=1}^n \left[g\left(\left\|\frac{\mathbf{x}-\mathbf{x}_i}{h}\right\|^2\right) W(\mathbf{x}_i) \right] \times \left[\frac{\sum_{i=1}^n \left[\mathbf{x}_i g\left(\left\|\frac{\mathbf{x}-\mathbf{x}_i}{h}\right\|^2\right) W(\mathbf{x}_i) \right]}{\sum_{i=1}^n \left[g\left(\left\|\frac{\mathbf{x}-\mathbf{x}_i}{h}\right\|^2\right) W(\mathbf{x}_i) \right]} - \mathbf{x} \right] \end{aligned} \quad (5.17)$$

定义(5.18)式为一般意义下的 Mean Shift 向量。

$$M_{h,G}(\mathbf{x}) = \frac{\sum_{i=1}^n \left[\mathbf{x}_i g\left(\left\|\frac{\mathbf{x}-\mathbf{x}_i}{h}\right\|^2\right) W(\mathbf{x}_i) \right]}{\sum_{i=1}^n \left[g\left(\left\|\frac{\mathbf{x}-\mathbf{x}_i}{h}\right\|^2\right) W(\mathbf{x}_i) \right]} - \mathbf{x} \quad (5.18)$$

结合(5.16)式, 可以得到:

$$\nabla f_K(\mathbf{x}) = \frac{2}{Ch^2} f_G(\mathbf{x}) M_{h,G}(\mathbf{x}) \quad (5.19)$$

也即:

$$M_{h,G}(\mathbf{x}) = \frac{Ch^2}{2} \frac{\nabla f_K(\mathbf{x})}{f_G(\mathbf{x})} \quad (5.20)$$

5.4.2 加权 Mean Shift 过程收敛性证明

定理: 设 Mean Shift 过程中, 位置序列 $\{\mathbf{y}_j\}_{j=1,2,\dots}$, 如果初始位置为 \mathbf{y}_j , 则按用核 $G(x)$ 公式(5.18)计算 \mathbf{y}_j 处的加权平均为:

$$\mathbf{y}_{j+1} = \frac{\sum_{i=1}^n \left[\mathbf{x}_i g\left(\left\|\frac{\mathbf{x}-\mathbf{x}_i}{h}\right\|^2\right) W(\mathbf{x}_i) \right]}{\sum_{i=1}^n \left[g\left(\left\|\frac{\mathbf{x}-\mathbf{x}_i}{h}\right\|^2\right) W(\mathbf{x}_i) \right]} \quad (5.21)$$

对应的用核 $K(x)$ 计算的核密度估计为: $\{f_K(j)\}_{j=1,2,\dots}$, Mean Shift 向量为 $M_{h,G} = \mathbf{y}_{j+1} - \mathbf{y}_j$. 如果 $K(x)$ 的轮廓函数 $k(x)$ 为凸的, 单调递减, 对于每个 \mathbf{x}_i 权值 $W(\mathbf{x}_i)$ 为大于等于 0 的常数, 那么 $\{f_K(j)\}_{j=1,2,\dots}$ 和 $M_{h,G} = \mathbf{y}_{j+1} - \mathbf{y}_j$ 收敛。

证明:

因为 n 有限, $\{f_K(j)\}_{j=1,2,\dots}$ 有界, 所以只要证明 $\{f_K(j)\}_{j=1,2,\dots}$ 严格单调递增, 则它一定收敛。即要证明, 如果 $\mathbf{y}_{j+1} \neq \mathbf{y}_j$, 则 $f_K(j) < f_K(j+1)$ 。不失一般性, 假定 $\mathbf{y}_j = \mathbf{0}$, 则

$$f_K(j+1) - f_K(j) = \frac{1}{nh^d} \sum_{i=1}^n \left\{ \left[k\left(\left\|\frac{\mathbf{y}_{j+1} - \mathbf{x}_i}{h}\right\|^2\right) - k\left(\left\|\frac{\mathbf{x}_i}{h}\right\|^2\right) \right] W(\mathbf{x}_i) \right\} \quad (5.22)$$

因为轮廓函数是凸的, 则对任意的 $x_1, x_2 \in [0, \infty)$, $x_1 \neq x_2$ 有: $k(x_2) \geq k(x_1) + k'(x_2)(x_2 - x_1)$, 注意到 $g(x) = -k'(x)$, 则有 $k(x_2) - k(x_1) \geq g(x_2)(x_1 - x_2)$, 代入公式(5.22)有:

$$\begin{aligned} f_K(j+1) - f_K(j) &\geq \frac{1}{nh^{d+2}} \sum_{i=1}^n \left[g\left(\left\|\frac{\mathbf{x}_i}{h}\right\|^2\right) W(\mathbf{x}_i) (\|\mathbf{x}_i\|^2 - \|\mathbf{y}_{j+1} - \mathbf{x}_i\|^2) \right] \\ &= \frac{1}{nh^{d+2}} \sum_{i=1}^n \left[g\left(\left\|\frac{\mathbf{x}_i}{h}\right\|^2\right) W(\mathbf{x}_i) (2\mathbf{y}_{j+1}^T \mathbf{x}_i - \|\mathbf{y}_{j+1}\|^2) \right] \end{aligned} \quad (5.23)$$

$$= \frac{1}{nh^{d+2}} \left\{ 2\mathbf{y}_{j+1}^T \sum_{i=1}^n \left[\mathbf{x}_i g \left(\left\| \frac{\mathbf{x}_i}{h} \right\|^2 \right) W(\mathbf{x}_i) \right] - \|\mathbf{y}_{j+1}\|^2 \sum_{i=1}^n \left[g \left(\left\| \frac{\mathbf{x}_i}{h} \right\|^2 \right) W(\mathbf{x}_i) \right] \right\}$$

结合式(5.21)可以得到:

$$f_K(j+1) - f_K(j) \geq \frac{\|\mathbf{y}_{j+1}\|^2}{nh^{d+2}} \sum_{i=1}^n \left[g \left(\left\| \frac{\mathbf{x}_i}{h} \right\|^2 \right) W(\mathbf{x}_i) \right] \quad (5.24)$$

因为 $k(x)$ 单调递减且 $W(\mathbf{x}_i) \geq 0$, 所以 $\sum_{i=1}^n \left[g \left(\left\| \frac{\mathbf{x}_i}{h} \right\|^2 \right) W(\mathbf{x}_i) \right] > 0$, 又因为 $\mathbf{y}_{j+1} \neq \mathbf{y}_j = \mathbf{0}$, 可以得到 $f_K(j+1) - f_K(j) > 0$, 所以 $\{f_K(j)\}_{j=1,2,\dots}$ 收敛。

现在证明 $M_{h,G} = \mathbf{y}_{j+1} - \mathbf{y}_j$ 也收敛。假设 $\mathbf{y}_j \neq \mathbf{0}$ 重写式(5.22), 可以得到:

$$f_K(j+1) - f_K(j) \geq \frac{\|\mathbf{y}_{j+1} - \mathbf{y}_j\|^2}{nh^{d+2}} \sum_{i=1}^n \left[g \left(\left\| \frac{\mathbf{y}_j - \mathbf{x}_i}{h} \right\|^2 \right) W(\mathbf{x}_i) \right] \quad (5.25)$$

因为 $\{f_K(j)\}_{j=1,2,\dots}$ 收敛, 所以 $\lim_{j \rightarrow \infty} [f_K(j+1) - f_K(j)] = 0$, 由于 $K(x) = K_E(x)$, 则 $g_E(x)$ 为正常数, 而 $W(\mathbf{x}_i)$ 为预先指定大于等于 0 的常数, 所以可以得到 $\|\mathbf{y}_{j+1} - \mathbf{y}_j\|^2$ 收敛于 0, 所以 $\{\mathbf{y}_j\}_{j=1,2,\dots}$ 为 Cauchy 序列, 因此 $M_{h,G} = \mathbf{y}_{j+1} - \mathbf{y}_j$ 也收敛, 证毕。

5.5 差异性加权 Mean Shift 跟踪算法原理及步骤

经典 Mean Shift 跟踪算法在初始帧中以目标为中心选择一个矩形或圆形区域, 统计该区域内的颜色 (或灰度) 直方图作为目标特征描述向量。在后续帧中跟踪目标时, 以上一帧目标位置为中心选择一个更大的搜索区域, 统计搜索区域的颜色 (或灰度) 直方图作为候选目标特征。通过不断移动搜索区域, 找到一个位置, 使得候选颜色 (或灰度) 直方图和目标特征描述向量最为匹配, 即达到目标跟踪目的。在统计颜色 (或灰度) 直方图时, 不可避免地包含部分背景信息, 而背景信息对 Mean Shift 跟踪算法的性能有一定影响。特别当运动目标在复杂变化背景中运动时, 变化的背景部分会严重影响经典 Mean Shift 跟踪算法的稳定性和可靠性。针对该问题, 本文在 5.1 节目标差异性权值计算方法和 5.4 节加权 Mean Shift 算法的基础上, 提出了一种差异性加权 Mean Shift 跟踪算法 (Diversity Weighted Mean Shift, 以下简称 DWMS)。

由于在相邻两帧图像中目标区域特性变化较小, 所以在计算当前帧目标位置时, 可以用上一帧的目标差异性权值作为权值统计加权直方图信息。因此在当前帧中, 利用加权 Mean Shift 算法迭代计算目标位置时, 差异性权值满足的加权 Mean Shift 过程收敛的权值要求, 即 DWMS 跟踪算法是收敛的。下面以彩色图像为例,

简要介绍一下 DWMS 算法的原理和步骤。对于灰度图像，新算法仍然适用，但要将颜色直方图换为灰度直方图。

5.5.1 差异性加权 Mean Shift 跟踪算法原理

5.5.1.1 确定颜色直方图栅格权值

相邻两帧中目标区域的颜色信息变化不大，因此在统计下一帧目标候选颜色直方图时，可以根据初始帧目标的差异性权值计算得到颜色直方图权值，再利用颜色直方图权值对每个像素点颜色信息进行加权处理。这样做可以有效降低后续帧中变化背景部分的影响。但是由于整个跟踪过程中背景区域颜色信息变化较大，所以在每一帧目标跟踪结束后，需要更新颜色直方图权值，具体更新方法将在后续章节进行介绍。

为了加快算法单帧处理速度，将内区域的颜色直方图每 d 个相邻的颜色划分为同一栅格，从而得到 m 个直方图栅格，其中 $d \times m = 256^3$ 。对于初始帧的内区域，按公式(5.26)计算每个颜色直方图栅格的权值 $W_p(u)$ ($u=1, \dots, m$)，该权值仅与颜色直方图栅格参数 u 有关，而与区域内像素坐标无关。

$$W_p(u) = \frac{\sum_{i=1}^N \{W_q(\mathbf{x}_i) \cdot \delta[b(\mathbf{x}_i) - u]\}}{\sum_{i=1}^N \delta[b(\mathbf{x}_i) - u]} \quad u \in \{1, \dots, m\} \quad (5.26)$$

其中函数 $b: \mathbf{x}_i \rightarrow \{1, \dots, m\}$ 将坐标为 \mathbf{x}_i 的像素根据其颜色信息映射到相应的颜色直方图栅格中， $\delta(x)$ 为迪拉克函数。 $W_q(\mathbf{x}_i)$ 为利用 5.1 节提出的方法计算初始帧中目标区域的差异性权值。

5.5.1.2 计算初始帧的目标特征描述模板

为了统计初始帧目标特征描述模板，经典 Mean Shift 算法是以目标为中心选择一定大小的区域，统计该区域的颜色直方图栅格。然而选择的区域中不可避免的存在背景像素。为了抑制背景像素的影响，我们先对目标区域像素利用 5.1 节的方法进行加权处理，然后再统计加权颜色直方图栅格作为目标的特征描述模板。具体做法为：假设 $\{\mathbf{x}_i\}_{i=1, \dots, N}$ 为目标窗口中第 i 个像素的坐标， N 为目标窗口的像素点总数，其中 \mathbf{x}_c 为目标窗口中心， m 为颜色直方图栅格总数。因此目标加权颜色直方图栅格特征向量中第 u 个分量为：

$$q_u = C \sum_{i=1}^N \left\{ k(\|\mathbf{x}_i - \mathbf{x}_c\|^2) \cdot W_q(\mathbf{x}_i) \cdot \delta[b(\mathbf{x}_i) - u] \right\} \quad u \in \{1, \dots, m\} \quad (5.27)$$

式中 $k(\|x\|^2)$ 是核函数（通常为高斯核函数）， $W_q(\mathbf{x}_i)$ 为坐标 \mathbf{x}_i 处像素的目标区域颜色差异性权值。而 C 为归一化常数，定义为：

$$C = 1 / \sum_{i=1}^N [k(\|\mathbf{x}_i - \mathbf{x}\|^2) \cdot W_q(\mathbf{x}_i)] \quad (5.28)$$

5.5.1.3 计算当前帧候选特征向量

由于相邻两帧同一目标的颜色特性变化不大，所以可以用前一帧目标直方图栅格权值作为当前帧目标区域权值统计候选特征向量，这样做可以一定程度上抑制背景影响。具体做法为：假设 $\{\mathbf{x}_i\}_{i=1, \dots, N_h}$ 为当前帧中搜索窗口中第 i 个像素的坐标， N_h 为搜索窗口中像素总数，其中 \mathbf{y}_0 为搜索窗口中心。因此当前帧中搜索窗口的加权颜色直方图栅格特征向量的第 u 个分量为：

$$p_u(\mathbf{y}_0) = C_h \sum_{i=1}^{N_h} \left\{ k \left(\left\| \frac{\mathbf{x}_i - \mathbf{y}_0}{h} \right\|^2 \right) \cdot W_p(u) \cdot \delta[b(\mathbf{x}_i) - u] \right\} \quad u \in \{1, \dots, m\} \quad (5.29)$$

式中 $W_p(u)$ 为当前帧中第 u 个直方图栅格对应的权值，归一化常数 C_h 的计算方法类似于 C ，而 h 为核函数的带宽。

$$C_h = 1 / \sum_{i=1}^{N_h} \left[k \left(\left\| \frac{\mathbf{x}_i - \mathbf{y}_0}{h} \right\|^2 \right) \cdot W_p(u) \right] \quad (5.30)$$

5.5.1.4 选择相似性函数

获得初始帧目标特征描述模板和当前帧目标候选特征向量后，需要利用相似性函数判断两个模式是否相似，即判断目标是否匹配。这里类似经典 Mean Shift 算法，仍选用 Bhattacharyya 系数描述初始帧目标模板和当前帧候选向量的相似程度，其定义为：

$$\rho(\mathbf{y}) \equiv \rho[\mathbf{p}(\mathbf{y}), \mathbf{q}] = \sum_{u=1}^m \sqrt{q_u \cdot p_u(\mathbf{y})} \quad (5.31)$$

式中 \mathbf{y} 表示当前帧候选模板的中心。

5.5.1.5 当前帧跟踪结果求解

为了使(5.31)式最大，在当前帧中以前一帧跟踪结果作为当前帧搜索窗口的位置，设 \mathbf{y}_0 为窗口中心，在 \mathbf{y}_0 邻域内寻找最优目标位置 \mathbf{y}_1 ，反复迭代直到找到某个位置使得当前帧目标候选特征向量和初始帧目标特征描述模板最为相似，即找到当前帧中目标的最优位置。对(5.31)式在 $\mathbf{p}(\mathbf{y}_0)$ 处进行泰勒展开，相似性函数可以近似表示为：

$$\rho[\mathbf{p}(\mathbf{y}), \mathbf{q}] \approx \frac{1}{2} \sum_{u=1}^m \sqrt{q_u \cdot p_u(\mathbf{y}_0)} + \frac{C_h}{2} \sum_{u=1}^{N_h} \left[\omega_i k \left(\left\| \frac{\mathbf{y} - \mathbf{x}_i}{h} \right\|^2 \right) \right] \quad (5.32)$$

$$\text{其中} \quad \omega_i = \sum_{u=1}^m \left\{ \delta[b(\mathbf{x}) - u] \sqrt{q_u / p_u(\mathbf{y}_0)} \right\} \quad (5.33)$$

这里的采样权值 ω_i 和经典 Mean Shift 算法虽然在形式上一样, 但实质上存在较大差别, 这种差别主要在于在计算 q_u 和 $p_u(\mathbf{y}_0)$ 时, 引入了权值信息。

通过对相似性函数求最大值, 根据 Mean Shift 向量迭代得到当前帧目标的新位置:

$$\mathbf{y}_{j+1} = \frac{\sum_{u=1}^{N_h} \left[\mathbf{x}_i \omega_i g \left(\left\| \frac{\mathbf{y}_j - \mathbf{x}_i}{h} \right\|^2 \right) \right]}{\sum_{u=1}^{N_h} \left[\omega_i g \left(\left\| \frac{\mathbf{y}_j - \mathbf{x}_i}{h} \right\|^2 \right) \right]} \quad (5.34)$$

5.5.2 差异性加权 Mean Shift 跟踪算法步骤

(1) 对初始帧图像, 利用 5.1 节的方法计算目标区域的差异性权值, 根据差异性权值统计目标特征描述向量 (即加权颜色直方图栅格), 将初始帧目标位置作为下一帧中目标初始位置 \mathbf{y}_0 ;

(2) 为了提高算法的可靠性, 利用(5.26)式计算初始帧目标区域颜色直方图栅格权值;

(3) 对于当前帧, 以 \mathbf{y}_j 为中心、带宽参数 h 为半径选择一个圆形区域, 利用(5.29)式统计当前帧的候选特征向量;

(4) 根据(5.33)式得到 Mean Shift 采用权值;

(5) 利用(5.34)式得到当前帧目标的位置 \mathbf{y}_{j+1} ;

(6) 如果 $\|\mathbf{y}_{j+1} - \mathbf{y}_j\| < \varepsilon$ (ε 为预先选择的阈值), 则停止迭代, 输出当前帧目标位置 \mathbf{y}_{j+1} , 更新目标差异性权值; 否则令 $\mathbf{y}_j = \mathbf{y}_{j+1}$, 返回 (3) 步。

5.6 DWMS 算法实验结果分析

利用最小二乘匹配跟踪法 (LSIM)、Mean Shift跟踪算法 (MS)、粒子滤波跟踪算法 (PF) 以及本文提出的DWMS跟踪算法, 对两组序列图像中的运动目标进行跟踪, 其中用红色矩形框表示跟踪结果。

5.6.1 复杂背景下运动飞机跟踪结果 (一)

该组实验序列图像中, 飞机相对于背景差别较大。图 5-18 表示为序列图像中第 0、20、40、60 帧的跟踪结果, 其中 a 组为 LSIM 跟踪结果, b 组为 MS 跟踪结果, c 组为粒子滤波算法跟踪结果, d 组为 DWMS 跟踪结果。结果表明, LSIM 算

法和 MS 算法均跟踪失败。粒子滤波算法开始可以有效地跟踪目标，但是后来也跟踪失败。而 DWMS 算法依然可以一直有效、准确地跟踪运动目标。

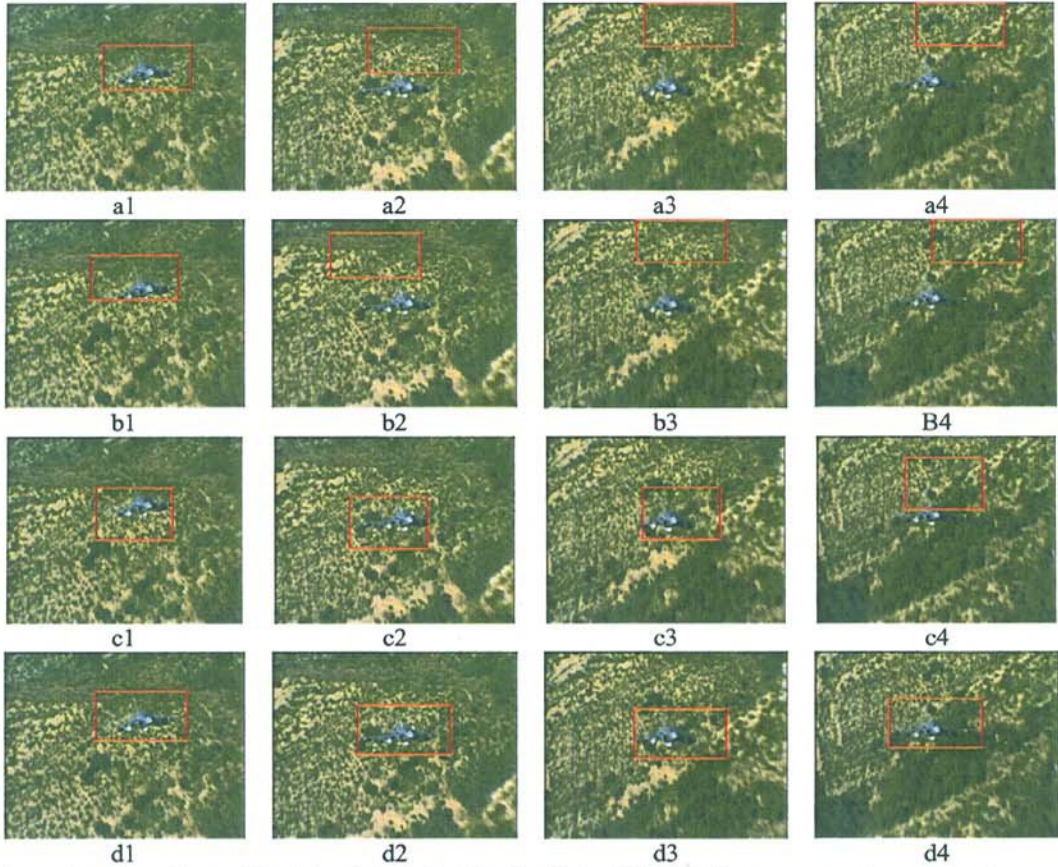


图 5-18 复杂背景下运动飞机跟踪结果比较（一）

5.6.2 复杂背景下运动飞机跟踪结果（二）

该组实验序列图像中，飞机相对于背景差别较小。图 5-19 表示为序列图像中第 0、8、16、24 帧的跟踪结果，其中 a 组为 LSIM 跟踪结果，b 组为 MS 跟踪结果，c 组为粒子滤波算法跟踪结果，d 组为 DWMS 跟踪结果。结果表明，LSIM 算法和 MS 算法均跟踪失败。粒子滤波算法开始可以有效地跟踪目标，但是后来也跟踪失败。而 DWMS 算法依然可以有效地跟踪运动目标。



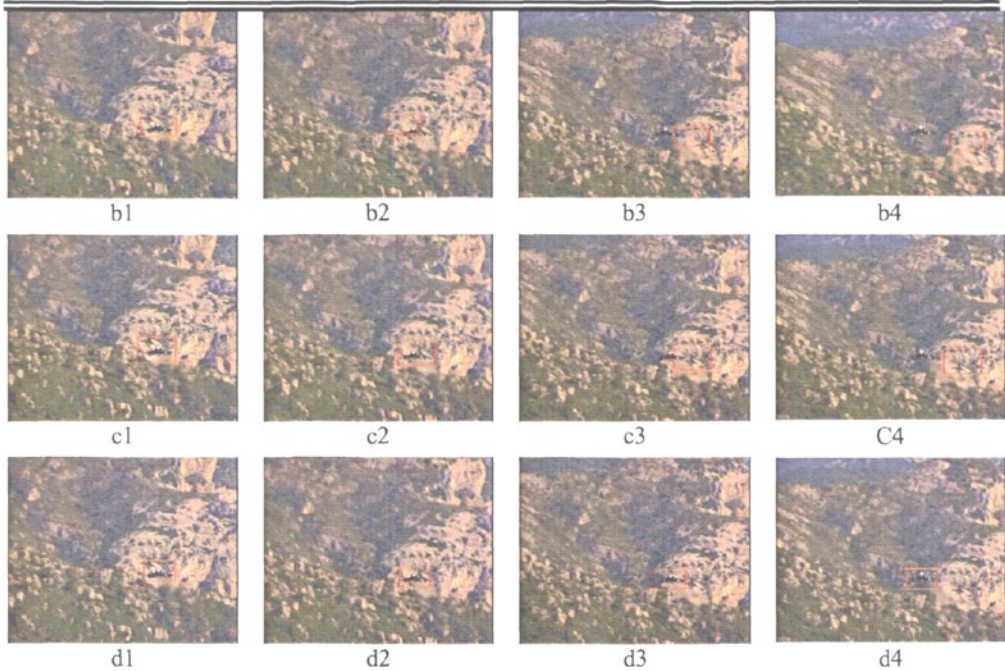


图 5-19 复杂背景下运动飞机跟踪结果比较 (二)

5.6.3 四种算法平均处理时间比较

对于两组图像, 分别用 MS、LSIM、PS 和 DWMS 算法进行运动目标跟踪。四种算法在 VC++6.0 平台上编程实现, 计算机主频为赛扬 2.4GHz、内存为 512MB。在初始条件相同的条件下, 算法平均处理时间如表 5-6 所示。可以看出, MS 虽然耗时较短但跟踪很容易失败, PS 算法耗时最长, 本文提出的 DWLSIM 算法较短。

表 5-6 四种算法平均处理时间比较 单位: 毫秒

| | Ms 算法 | LSIM 算法 | PS 算法(30 个粒子) | DWMS 算法 |
|-------|-------|---------|---------------|---------|
| 第一组图像 | 16.5 | 46.3 | 150.7 | 27.3 |
| 第二组图像 | 17.5 | 47.2 | 200.4 | 28.6 |

5.7 本章小结

本章针对准确建立目标区域描述和复杂背景下运动目标跟踪两个问题进行了研究。其中为了准确建立目标区域特征描述, 本章提出了一种单帧目标差异性权值计算方法。实验结果表明, 该权值可以较好地描述目标区域特性, 同时对选择区域中的背景部分有较好的抑制作用。

本章针对传统跟踪算法不能较好地解决复杂背景下运动目标跟踪问题, 提出了一种基于差异性加权的最小二乘影像匹配的运动目标跟踪算法。该算法根据目

标的形状及和背景的差异性对跟踪窗口中的点进行加权，降低了窗口区域中背景点对匹配的影响。仿真和实际跟踪结果证明了该算法的跟踪有效性及精度。

近些年，Mean Shift 跟踪算法因其独有的特性而成为研究热点。但是对于复杂背景下运动目标跟踪问题，该方法跟踪结果并不理想。本章针对这个问题，提出了一种基于差异性加权的 Mean Shift 运动目标跟踪算法。实验表明，新跟踪算法和 Mean Shift 算法、粒子滤波算法及 LSIM 算法相比，可以更为稳健、有效地、准确地跟踪复杂背景下运动目标。

第六章 尺度方向自适应 Mean Shift 跟踪算法研究

精确制导与精确打击是现代高科技武器发展的主要趋势之一。在众多精确制导技术中，基于图像的精确制导技术，因其分辨率高、弹载体积小、重量轻以及采用被动工作方式等优点受到越来越多的重视。

精确制导和精确打击技术按打击目标固定与否可分为打击固定目标精确制导技术和打击运动目标精确制导技术。对运动目标（如坦克、装甲运兵车等）进行精确打击时，为了获得最佳的打击效果，需要实时获得目标重要部件的精确位置信息，比如装甲运兵车的中心、坦克的履带等。为了实时获得目标精确位置信息和方向信息，本章针对同时存在位置、缩放和旋转变化的目标自适应跟踪定位问题进行研究，提出了一种尺度方向自适应的 Mean Shift 跟踪算法。

本章内容的结构安排如下：6.1 节就传统 Mean Shift 及其改进算法对目标尺度方向自适应跟踪问题进行阐述；6.2 节介绍一种目标最佳描述区域计算方法；6.3 节和 6.4 节详细介绍一种尺度方向自适应 Mean Shift 目标跟踪算法；6.5 节给出了实验结果及分析；6.6 节对本章进行了小结。

6.1 引言

如何对感兴趣目标进行有效跟踪是计算机视觉中一个极具挑战性的课题。特别是在跟踪过程中目标同时存在旋转、缩放、遮挡等情况时，对目标进行准确跟踪变得更加困难。

近些年，Mean Shift 算法作为一种高效的跟踪算法，因其不需要参数、不需要进行穷尽搜索等优点，已经成功地应用于目标跟踪领域。Mean Shift 跟踪算法中核函数带宽参数起着非常重要的作用，因为它不但决定了参与 Mean Shift 迭代的样本数量，而且也反映了跟踪窗口的大小。通常核函数带宽参数由初始跟踪窗口的尺寸决定，而且在整个跟踪过程中不再发生改变。然而，当目标存在明显缩放变化时，固定不变的核窗口带宽参数可能会导致目标跟踪失败。Comaniciu 和 Meer^[168]提出了用正负 10% 的增量分别对核窗口带宽参数进行修正的方法。该方法需要在当前帧中用三个不同大小带宽参数的核函数进行三次独立的 Mean Shift 跟踪计算，选择最大的 Bhattacharyya 系数所对应的带宽参数作为最佳核带宽参数。当目标逐渐缩小时，该方法可以得到较好的效果，但是当目标逐渐增大时，核窗口宽度很难被扩大，反而逐渐变小。这是因为基于 Bhattacharyya 系数的相似性度量经常会在较小的跟踪窗口中达到局部最大。

尺度空间理论自从提出后就备受关注，许多学者以该理论为基础提出了一些

非常有实际意义的算法。Collins 提出了一种尺度空间下目标 Mean Shift 跟踪算法^[204]，该方法使用 DOG-Epanichnikov 函数作为 Mean Shift 核函数，通过在尺度空间和坐标空间中搜索极大值，实现目标跟踪。此方法虽然有效地解决了目标缩放问题，但是对目标角度参数描述不够准确。

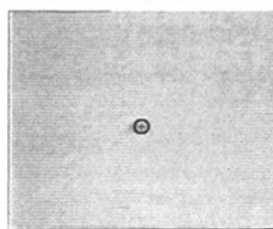
为了对同时存在旋转、缩放变化的运动目标进行尺度方向自适应跟踪，通常需要准确获得初始帧中目标大小和方向信息。而准确获得目标大小和方向信息，又为准确建立目标初始特征描述提供了条件。为此，本章提出了一种目标最佳椭圆描述区域计算方法。对于均匀或非均匀目标，该算法均可以准确地获得目标大小和方向信息。获得目标初始大小方向信息后，可以通过 Mean Shift 迭代过程逐帧更新目标位置、大小和方向信息，达到准确跟踪目标的目的。基于以上思路，本章提出了一种尺度方向自适应 Mean Shift 跟踪算法。下面先介绍目标最佳椭圆描述区域及算法，然后详细介绍提出的尺度方向自适应 Mean Shift 跟踪算法。

6.2 目标最佳椭圆描述区域计算方法

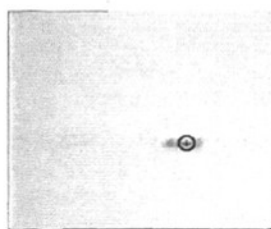
跟踪算法的准确性和稳定性很大程度上取决于目标初始特征建立是否准确。而目标初始特征建立又取决于目标描述区域确定和目标特征选择两个方面。本文就如何确定目标最佳描述区域问题进行了研究，在分析 Lindeberg 尺度空间理论基础上，提出了一种目标最佳椭圆描述区域计算方法。目标最佳椭圆描述区域尽可能将目标包含在内，同时给出了目标区域的大小和方向信息。该描述区域为准确提取目标特征和实现目标尺度方向自适应跟踪奠定了基础。下面先简要介绍新算法用到的椭圆正则化 LOG 算子，然后对新算法计算步骤进行介绍。

6.2.1 椭圆型正则化 LOG 算子

圆形正则化 LOG 算子是 Lindeberg 尺度理论中经常用到的算子之一。该算子虽然可以较好地描述对于圆形或正方形目标区域局部特征，但是对长条型目标描述并不准确。而且该算子没有方向信息描述能力。图 6-1. a 和图 6-1. b 分别为距坦克不同距离拍摄的图像，深灰色区域为坦克目标。



a 距坦克较远拍摄的图像



b 距坦克较近拍摄的图像

图 6-1 正则化 LOG 算子计算得到目标内部描述区域

图 6-1. a 中黑色十字为中心、圆形正则化 LOG 算子计算得到特征尺度为半径的黑色圆圈较好地覆盖了坦克目标区域。而图 6-1. b 中黑色十字为中心、圆形正则化 LOG 算子计算得到特征尺度为半径的黑色圆圈并不能较好地描述长条型坦克目标区域。但是该算子计算得到的特征尺度参数可以较好地描述目标宽度信息，如图 6-1. b 中黑色圆圈较好地描述了坦克目标宽度信息。

为了较好地描述长条形目标区域，我们在分析椭圆高斯算子基础上提出了椭圆型正则化 LOG 算子。在提出椭圆形正则化 LOG 算子之前，先简单介绍一下带方向的椭圆高斯函数，通常，该函数定义为：

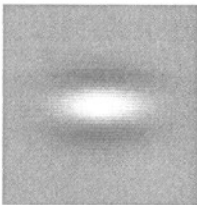
$$E(x_0, y_0, \sigma_a, \sigma_b, \theta) = \frac{1}{2\pi\sigma_a\sigma_b} e^{-\frac{1}{2} \begin{bmatrix} x-x_0 & y-y_0 \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \frac{1}{\sigma_a^2} & 0 \\ 0 & \frac{1}{\sigma_b^2} \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x-x_0 \\ y-y_0 \end{bmatrix}} \quad (6.1)$$

(x_0, y_0) 为椭圆高斯的中心， σ_a 椭圆高斯长半轴参数， σ_b 为椭圆短半轴参数， θ 为椭圆长轴与 x 坐标轴的夹角。由(6.1)式可以看出椭圆高斯函数具有方向描述能力。

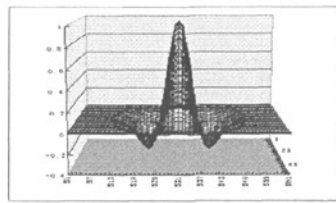
类似于圆形正则化 LOG 算子构建方法，定义椭圆正则化 LOG 算子 (Ellipse Normalized LOG, 简称 ENLOG) 为：

$$ENLOG = \delta_a \delta_b [E_{xx}(x_0, y_0, \sigma_a, \sigma_b, \theta) + E_{yy}(x_0, y_0, \sigma_a, \sigma_b, \theta)] \quad (6.2)$$

其中 E_{xx} 和 E_{yy} 分别为椭圆高斯算子关于 x 方向和 y 方向的二阶导数。图 6-2 为 $\sigma_a = 10, \sigma_b = 5, \theta = 0^\circ$ 的 ENLOG 算子示意图。图 6-2.a 中偏黑的部分表示 ENLOG 算子负的部分，白色部分表示 ENLOG 算子正的部分。图 6-2.b 为 ENLOG 算子三维示意图，为了方便读者理解，将图像进行了一定的旋转。



a ENLOG 算子平面图像



b ENLOG 算子三维示意图

图 6-2 ENLOG 算子示意图

6.2.2 目标最佳椭圆描述区域计算步骤

上一节介绍了 ENLOG 算子，可以看出 ENLOG 算子同时考虑了长轴信息、短轴信息和角度信息，因此该函数不但具有类似圆形 LOG 算子的尺度选择能力，还具有类似椭圆高斯算子的方向描述能力。本节主要介绍目标最佳椭圆描述区域计算方法。

对于非均匀目标, 直接对目标区域图像计算最佳描述区域, 其结果可能仅包含目标的局部区域。为了准确获得包含整个目标区域的描述区域, 首先应尽可能将整个目标从背景分割出来。由于目标的非均匀性, 分割后目标区域图像中常存在一些“空洞”, 为了进一步提高计算区域的准确性, 用形态学膨胀算法填充目标分割图像中“空洞”, 取膨胀后图像的形心作为最佳椭圆描述区域的中心坐标参数。由于圆形正则化 LOG 算子可以较好地描述目标的宽度信息, 所以可以利用该算子求取目标最佳椭圆描述的短半轴参数。而 ENLOG 具有类似椭圆高斯算子的方向描述能力, 因此可以人为指定一个长半轴参数, 而令角度 θ 从 0° 间隔 1° 增加到 179° , 构造 180 个 ENLOG 算子, 分别将 180 个 ENLOG 算子与目标分割膨胀后图像进行内积运算得到每个 ENLOG 算子的响应值, 最大响应值对应的角度 θ 与目标方向一致。最后利用已获得最佳椭圆描述区域中心坐标、短半轴参数和角度参数, 选择一系列长半轴参数, 构造一组 ENLOG 算子, 分别将该组 ENLOG 算子与目标区域分割图像进行内积运算, 最大响应值对应的长半轴参数作为最佳椭圆描述区域的长半轴参数。经过前面几步获得最佳椭圆描述参数可能存在一定误差, 还需要对每个参数进行求精运算。

本文提出的目标最佳椭圆描述区域计算方法不仅适用于彩色图像, 而且也适用于灰度图像。下面以灰度图像为例, 简要介绍目标最佳椭圆描述区域算法步骤。

6.2.2.1 确定目标区域分割图像

由于光照变化、背景遮挡等原因, 目标在图像上并不一定是连通的区域。如果将目标的部分区域误认为是整个目标, 而对其建立目标特征描述, 势必存在一定误差。因此尽可能将整个目标从背景中分割出来, 是准确建立目标特征的首要环节。这里利用 5.1 节提出的方法计算目标差异性权值, 然后对目标区域权值进行归一化处理, 将权值再拓展到 $0 \sim 255$, 即可得到目标区域分割图像。最后对分割图像进行形态学膨胀运算得到目标区域。图 6-3.A 为原始图像, 选择图像中心的瓶子作为目标。图 6-3.B 为原始图像中瓶子目标局部区域经 5.1 节方法计算得到的加权图像结果。图 6-3.C 为加权图像经过形态学膨胀运算后的结果。



A 原始图像



B 目标区域加权图像



C 目标区域加权图像膨胀运算后结果

图 6-3 目标区域加权图像计算结果

6.2.2.2 确定目标区域最佳椭圆描述

获得目标区域分割图像后, 根据 Lindeberg 尺度空间理论, 确定目标区域的最佳椭圆描述的五个参数, 即椭圆中心 (x_0, y_0) 、椭圆长半轴参数 σ_a 、椭圆短半轴参数 σ_b 及椭圆长轴与 x 坐标轴的夹角 θ 。由于同时确定五个参数计算量大, 所以我们采用逐一求取, 逐步求精的方法, 即先固定一个或几个参数, 仅对剩下的一个或几个参数进行迭代计算。参数计算顺序为: 1) 椭圆中心; 2) 椭圆短半轴参数; 3) 椭圆角度参数; 4) 椭圆长半轴参数。

1 椭圆中心位置求取

对目标区域分割图像进行形态学膨胀运算, 将膨胀后目标区域的形心作为最佳椭圆描述的中心坐标参数初始值。

2 椭圆短轴大小确定

利用 Lindeberg 尺度空间理论, 利用圆形正则化 LOG 算子计算出指定点的特征尺度, 将该特征尺度作为最佳椭圆描述的短半轴参数 σ_b 的初始值。具体做法为: 让尺度参数 σ 从 1.0 开始间隔 1.0 递增到 N (N 的大小根据实际情况进行选择), 构造一系列圆形正则化 LOG 算子。对于目标分割膨胀后的图像, 以椭圆中心初始值为当前点, 计算该系列算子的响应, 选取最大响应值对应的尺度参数 σ_c 作为 σ_b 的初始值, 其中特征尺度同时应满足以下条件:

$$F(\mathbf{x}, \sigma_c) > F(\mathbf{x}, \sigma_n) \quad (6.3)$$

$F(\mathbf{x}, \sigma_c)$ 表示 \mathbf{x} 处尺度参数为 σ_c 的响应值。 $F(\mathbf{x}, \sigma_n)$ 表示 \mathbf{x} 处除去 σ_c 参数以外的尺度参数为 σ_n 的响应值。

3 椭圆角度确定

获得椭圆短半轴参数 σ_b 后, 令椭圆长半轴 $\sigma_a = k\sigma_b$ (k 取 1.5), 而让 θ 从 0° 开始间隔 1° 增加到 179° , 构造一系列 ENLOG 算子, 将该系列算子与目标分割膨胀后图像进行内积运算, 得到各个 ENLOG 算子的响应值, 选取最大响应值对应的角度 θ 作为该目标最佳描述方向的初始值。

4 椭圆长轴求取

在获得椭圆中心、短半轴大小及椭圆角度四个参数后, 令长半轴 σ_a 为 $\sigma_b + 0.5n$ ($n=1, \dots, M$) 构建一系列 ENLOG 算子, 对目标分割膨胀后图像, 以椭圆中心初始值为当前点计算每个算子对于指定点的响应值, 其响应曲线类似于图 2-2.b, 选取最大响应值所对应的尺度 σ^c 作为椭圆长半轴 σ_a 的初始值, 其中特征尺度应满足类似(6.3)式的条件。

5 逐步求精

在获得椭圆描述五个参数的初始值后, 再逐一对每个参数进行求精运算, 获

得更为准确的椭圆描述。具体做法如下：首先固定椭圆中心、角度和长半轴参数，将椭圆短半轴参数 σ_b 设定为 $\sigma_b \pm 0.1 \times n$ ($n=5$)，对目标分割图像求取 ENLOG 算子响应，将最大响应值所对应的 σ'_b 作为新的短半轴参数。同理，依次更新椭圆的中心坐标、角度参数和长半轴参数。

6.2.3 实验结果及分析

图 6-4.a 为地面上放置了一个瓶子，图 6-4.b 中的白色椭圆为本文算法计算最佳椭圆描述。瓶子中上部与背景较为接近，而瓶子两端与背景有较大差别。可以看出计算得到的最佳椭圆描述区域的大小、方向和瓶子基本一致。

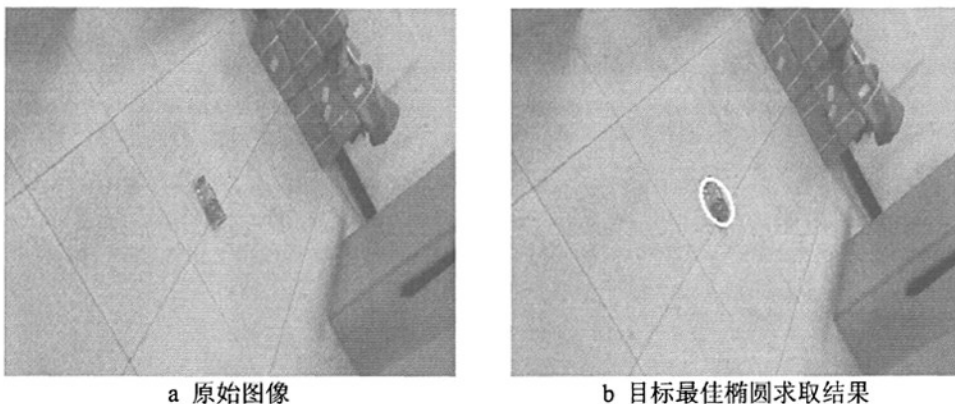


图 6-4 目标最佳描述椭圆计算结果

图 6-5 为无人机对某城市郊区拍摄图像，在该图像中选取了八个目标，用本节提出的算法计算最佳椭圆描述区域。可以看出，计算得到的最佳椭圆描述较好地描述了目标区域特性，即椭圆的大小、方向和目标区域基本一致。



图 6-5 目标最佳描述椭圆计算结果

6.3 尺度方向自适应 Mean Shift 算法原理

由 6.2 节可知, 目标最佳椭圆描述区域准确地描述了目标区域的位置、大小和方向。当相邻两帧图像中目标区域变化较小时, 同一目标的位置、大小和方向变化不大, 所以可以将上一帧的目标椭圆描述区域的五维参数 (即空间位置参数 \mathbf{x} 、尺度维参数 σ_a 和 σ_b 、角度维参数 θ) 作为下一帧目标描述的初始值, 通过 Mean Shift 迭代过程逐帧更新这五个参数, 达到尺度方向自适应跟踪目标的目的。

本文在获得目标最佳椭圆描述区域的基础上, 提出了一种尺度方向自适应的 Mean Shift 跟踪算法, 下面简要介绍一下新跟踪算法的原理。为了方面说明, 本节以灰度图像为例进行阐述, 对于彩色图像只需将灰度直方图换为颜色直方图即可, 新算法仍然适用。

6.3.1 核函数选择

在应用 Mean Shift 跟踪算法时, 通常选用高斯核、Epanechnikov 核或平台核作为核函数。当目标区域形状接近圆形或正方形时, 利用以上三种核函数可以获得较为准确目标特征描述。然而, 当目标区域形状为长方形时, 如果上述三种核函数带宽参数选择太大时, 会将部分背景区域信息错误地当作目标统计目标特征, 从而导致目标跟踪定位不够准确, 甚至跟踪失败; 如果三种核函数带宽参数选择太小时, 仅能将目标的部分区域信息作为目标整体特征进行计算, 也会导致目标跟踪定位不够准确。

为了准确地跟踪目标, 我们在分析 Collins 理论和尺度空间理论基础上, 分别选用 ENLOG 函数和椭圆高斯函数作为核函数, 其中在计算目标位置参数时用椭圆高斯算子作为核函数; 在计算目标尺度参数和角度参数时, 选用 ENLOG 函数作为核函数。椭圆高斯核函数和高斯核函数相似, 其轮廓函数满足 Mean Shift 算法要求, 使用该类核函数的 Mean Shift 迭代过程的收敛性已经得到了证明^[168]。而在新算法中使用了 ENLOG 作为核函数, 该函数存在负的部分, 并不满足 Mean Shift 算法对核函数的基本要求, 下面介绍一下 Mean Shift 算法与负核函数的关系。

Mean Shift 算法的一个限制是核函数 $K(\mathbf{x})$ 的轮廓函数必须非负、凸且单调。当核函数的轮廓函数存在负的部分或者非凸时, 有可能使得 Mean Shift 矢量为负矢量, 导致算法迅速发散, 在这种情况下 Mean Shift 算法有可能“收敛”到目标之外的任何地方。Collins 在其论文中阐述了 Mean Shift 算法与负核函数的关系^[204], 为了保证算法收敛, Collins 对原始的 Mean Shift 公式进行了修改, 即:

$$\Delta_{\mathbf{x}} = \frac{\sum [K(\mathbf{a} - \mathbf{x}) \times (\mathbf{a} - \mathbf{x})]}{\sum |K(\mathbf{a} - \mathbf{x})|} \quad (6.4)$$

其中 \mathbf{x} 为当前点, \mathbf{a} 为采样点。

以上讨论了核函数选择及 Mean Shift 收敛性问题, 下面开始介绍尺度方向自适应 Mean Shift 跟踪算法原理。

6.3.2 建立初始帧的目标特征灰度直方图

由 6.2 节可知, 目标最佳椭圆描述区域准确描述了目标区域的大小和方向信息, 所以在该区域内统计目标的特征描述更为准确。假设 $\{\mathbf{x}_i\}_{i=1,\dots,N}$ 为初始帧图像上目标最佳椭圆描述区域中像素点坐标, N 为目标最佳椭圆描述区域的像素总数, \mathbf{x}_c 为目标最佳椭圆描述区域中心。类似于(2.38)式计算初始帧目标特征直方图的第 u 个分量为:

$$q_u = \frac{1}{C} \sum_{i=1}^N \left\{ k(\|\mathbf{x}_i - \mathbf{x}_c\|^2 | \sigma_{a0}, \sigma_{b0}, \theta_0) \cdot \delta[b(\mathbf{x}_i) - u] \right\} \quad u \in [0, \dots, 255] \quad (6.5)$$

$$\text{其中:} \quad C = \sum_{i=1}^N \left[k(\|\mathbf{x}_i - \mathbf{x}_c\|^2 | \sigma_{a0}, \sigma_{b0}, \theta_0) \right]$$

k 是核函数的轮廓函数, 这里选用带方向的椭圆高斯算子其参数选为 σ_{a0} 、 σ_{b0} 、 θ_0 。 C 为归一化常数。

6.3.3 计算当前帧候选特征直方图向量

由于目标在图像上同时存在位置、旋转、缩放变化, 所以需要以前一帧目标位置为中心选择一个更大的范围作为搜索区域建立候选特征直方图。假设 $\{\mathbf{x}_i\}_{i=1,\dots,N_h}$ 为当前帧中搜索窗口中像素点坐标, N_h 为搜索窗口中像素总数, \mathbf{y}_0 为搜索窗口中心, 类似于(2.40)式计算当前帧中搜索窗口的候选特征直方图的第 u 个分量为:

$$p_u(\mathbf{y}_0) = \frac{1}{C_i} \sum_{i=1}^{N_h} \left\{ k(\|\mathbf{x}_i - \mathbf{y}_0\|^2 | h\sigma_{a0}, h\sigma_{b0}, \theta_0) \cdot \delta[b(\mathbf{x}_i) - u] \right\}$$

$$\text{其中:} \quad u \in [0, \dots, 255] \quad (6.6)$$

$$C_i = \sum_{i=1}^{N_h} \left[k(\|\mathbf{x}_i - \mathbf{y}_0\|^2 | h\sigma_{a0}, h\sigma_{b0}, \theta_0) \right]$$

k 是核函数的轮廓函数, 这里选用带方向的椭圆高斯算子, 其参数选为 $h\sigma_{a0}$ 、 $h\sigma_{b0}$ 、 θ_0 。 C_i 为归一化常数。而 h 定义了搜索窗口的大小。

6.3.4 相似性函数

当获得初始帧目标特征模板和当前帧候选特征直方图后,可以用 Bhattacharyya 系数描述两者之间的相似程度,其定义为:

$$\rho(\mathbf{y}) \equiv \rho[\mathbf{p}(\mathbf{y}), \mathbf{q}] = \sum_{u=0}^{255} \sqrt{q_u \cdot p_u(\mathbf{y})} \quad (6.7)$$

式中 \mathbf{y} 表示当前帧搜索区域中心。

6.3.5 Mean Shift 采样权值

由 Mean Shift 算法知,在计算 Mean Shift 矢量时,需要得到采样权值。下面以计算空间位置参数为例说明采样权值是如何获得的。当计算目标的空间位置参数 \mathbf{x} 时,在每一步迭代过程中固定 σ_a 、 σ_b 和 θ 不变,为了使(6.7)式最大,在当前帧中,以前一帧跟踪结果作为当前帧搜索窗口的位置,设 \mathbf{y}_0 为窗口中心,在 \mathbf{y}_0 邻域内寻找最优目标位置 \mathbf{y}_1 ,反复迭代,最后得到在当前帧目标的最优位置。对(6.7)式在 $\mathbf{p}_u(\mathbf{y}_0)$ 处进行泰勒展开,相似性函数可以近似表示为:

$$\rho[\mathbf{p}(\mathbf{y}), \mathbf{q}] \approx \frac{1}{2} \sum_{u=0}^{255} \sqrt{q_u \cdot p_u(\mathbf{y})} + \frac{C_h}{2} \sum_{u=0}^{N_i} \left[\omega'(\mathbf{x}_i) \cdot k(\|\mathbf{x}_i - \mathbf{x}_*\|^2) \right] h\sigma_{a0}, h\sigma_{b0}, \theta_0 \quad (6.8)$$

其中:

$$\omega'(\mathbf{x}_i) = \sum_{u=0}^{255} \left\{ \delta[b(\mathbf{x}_i) - u] \sqrt{q_u / p_u(\mathbf{y}_0)} \right\} \quad (6.9)$$

类似于经典 Mean Shift 算法,可以将 $\omega'(\mathbf{x}_i)$ 作为权值。同理,可以求出计算尺度维参数 σ_a 、 σ_b 和角度维参数 θ 时的采样权值。

6.3.6 更新目标特征直方图

由于光照等因素的影响,目标的灰度特性会发生变化,如果一直用固定目标特征灰度直方图,可能会导致跟踪失败。当获得当前帧目标的跟踪结果后,统计该帧中目标椭圆描述区域的灰度直方图 $h'(n)$ ($n=0 \sim 255$)。当满足(6.10)式时,说明当前帧中的目标灰度特性变化较大,应更新目标特征灰度直方图,即将 $h'(n)$ 作为新的目标特征灰度直方图。

$$\sum_{n=0}^{255} [h'(n) - h_c(n)]^2 > th \quad (6.10)$$

其中 $h_c(n)$ 为上一帧用到的目标特征直方图模板, th 为预先取的阈值参数。

6.4 尺度方向自适应 Mean Shift 的算法步骤

上一节介绍了尺度方向自适应 Mean Shift 跟踪算法的原理, 本节简要给出算法的大致步骤。

6.4.1 确定初始帧中目标最佳椭圆描述并统计目标特征直方图

利用 6.2 节中的算法确定初始帧中目标区域的最佳椭圆描述及统计目标特征直方图模板。

6.4.2 对每帧图像, 逐一求取目标的位置、尺度和角度参数

对目标进行尺度方向自适应跟踪时, 同时计算五维参数计算量较大, 这里采用交叉迭代方法逐一求取当前帧中目标的描述参数, 即每次固定一些参数不变, 针对另一个或几个参数进行 Mean Shift 迭代运算直至该参数收敛。

1、固定 σ_a 、 σ_b 和 θ , 以上一帧空间位置参数 \mathbf{x}_0 为中心, 选用椭圆高斯算子作为核函数, 利用(6.11)式进行空间位置参数的 Mean Shift 迭代运算, 直至收敛到新的目标位置, 其中 σ_{a0} 、 σ_{b0} 和 θ_0 为上一帧初始参数。

$$\mathbf{x}_1 = \frac{\sum_y [E(\mathbf{y} - \mathbf{x}_0, \sigma_{a0}, \sigma_{b0}, \theta_0) \times \omega'(\mathbf{y}) \times (\mathbf{y} - \mathbf{x}_0)]}{\sum_y [E(\mathbf{y} - \mathbf{x}_0, \sigma_{a0}, \sigma_{b0}, \theta_0) \times \omega'(\mathbf{y})]} \quad (6.11)$$

2、获得新的目标空间位置参数 \mathbf{x}_1 后, 固定 \mathbf{x}_1 、 σ_a 和 σ_b 参数, 选取 ENLOG 算子作为核函数, 利用(6.12)式对角度参数进行 Mean Shift 迭代直至收敛, 从而获得 S_{θ_1} , 再通过(6.13)式求出新的角度参数 θ_1 ;

$$S_{\theta_1} = \frac{\sum_s \sum_y [ENLOG(\mathbf{x}_1, \sigma_{a0}, \sigma_{b0}, \theta_s) \times \omega''(\mathbf{y}) \times s]}{\sum_s \sum_y |ENLOG(\mathbf{x}_1, \sigma_{a0}, \sigma_{b0}, \theta_s) \times \omega''(\mathbf{y})|} \quad (6.12)$$

$$\text{其中} \quad \theta_s = \begin{cases} \theta_0 + 10 \times e^{-s^2} & 0 \leq s \leq n \\ \theta_0 - 10 \times e^{-s^2} & -n \leq s < 0 \end{cases} \quad (6.13)$$

3、获得新的目标空间位置参数 \mathbf{x}_1 和角度参数 θ_1 后, 固定 \mathbf{x}_1 、 θ_1 和 σ_a 参数, 选取 ENLOG 算子作为核函数, 利用(6.14)式对尺度参数 σ_b 进行 Mean Shift 迭代直至收敛, 从而获得 S_{b_1} , 再通过(6.15)式求出新的尺度参数 σ_{b_1} ;

$$S_{b_1} = \frac{\sum_s \sum_y [ENLOG(\mathbf{x}_1, \sigma_{a0}, \sigma_{bs}, \theta_1) \times \omega''(\mathbf{y}) \times s]}{\sum_s \sum_y |ENLOG(\mathbf{x}_1, \sigma_{a0}, \sigma_{bs}, \theta_1) \times \omega''(\mathbf{y})|} \quad (6.14)$$

$$\text{其中} \quad \sigma_{bs} = \begin{cases} \sigma_{b0}(2 - e^{-s^2}) & 0 \leq s \leq n \\ \sigma_{b0} \times e^{-s^2} & -n \leq s < 0 \end{cases} \quad (6.15)$$

4、固定 x_1 、 θ_1 和 σ_{b1} 参数, 选取 ENLOG 算子作为核函数, 利用(6.16)式对尺度参数 σ_a 进行 Mean Shift 迭代直至收敛, 从而获得 S_{a1} , 再通过(6.17)式求出新的尺度参数 σ_{a1} :

$$S_{a1} = \frac{\sum_s \sum_y [ENLOG(x_1, \sigma_{as}, \sigma_{b1}, \theta_1) \times \omega''(y) \times s]}{\sum_s \sum_y |ENLOG(x_1, \sigma_{as}, \sigma_{b1}, \theta_1) \times \omega''(y)|} \quad (6.16)$$

$$\text{其中} \quad \sigma_{as} = \begin{cases} \sigma_{a0}(2 - e^{-s^2}) & 0 \leq s \leq n \\ \sigma_{a0} \times e^{-s^2} & -n \leq s < 0 \end{cases} \quad (6.17)$$

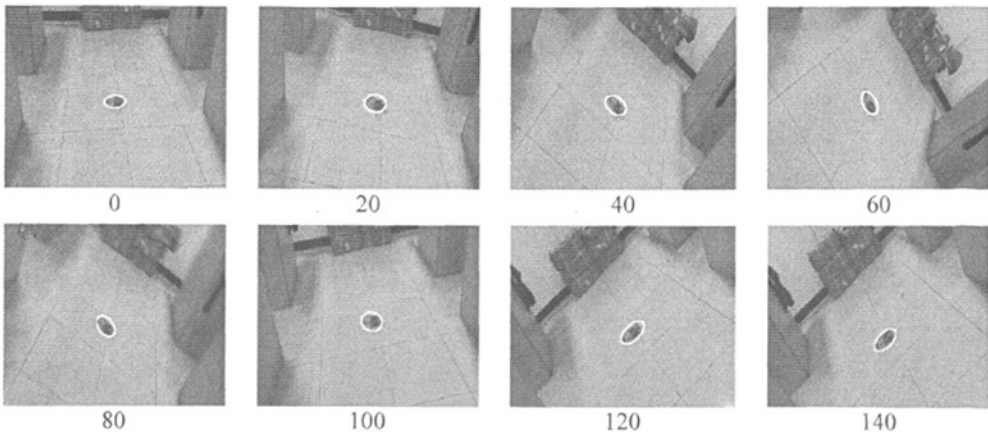
6.4.3 更新目标特征直方图

按 6.3.6 节方法判断是否需要更新目标特征直方图。如果需要更新, 则按 6.3.6 节的方法更新目标特征直方图。

6.5 实验结果及分析

6.5.1 跟踪结果及分析 (一)

在地面上放置了一个瓶子, 拍摄过程中像机相对于瓶子有一定的距离和角度变化。图6-6.0为序列的第0帧图像, 其中的白色椭圆为利用6.2节中的算法计算得到的初始最佳椭圆描述。获得目标初始描述后, 利用本文跟踪算法对后续图像进行逐帧跟踪, 其中图6-6.20至图6-6.280分别为序列图像中从第0帧开始间隔20帧的跟踪结果(用白色椭圆表示)。可以看出, 本文提出的新算法可以较好地对目标进行跟踪, 跟踪结果的位置、大小和方向和目标基本一致。



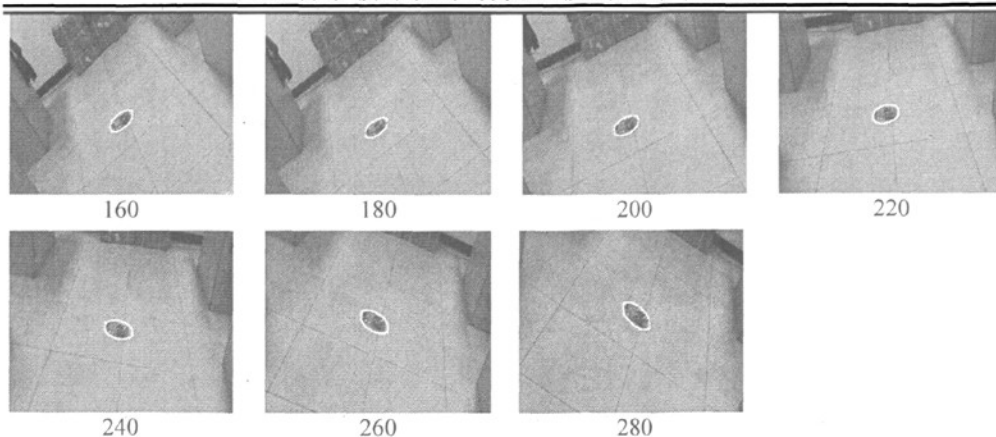


图6-6 瓶子尺度方向自适应跟踪结果

图6-7.0为序列的第0帧图像，其中的白色矩形为初始手动选取目标。用加减10%Mean Shift方法对目标进行跟踪。其中图6-7.20至图6-7.280分别为序列图像中从第0帧开始间隔20帧的跟踪结果（用白色矩形表示）。可以看出加减10%Mean Shift方法跟踪结果并不理想，当目标变大时，白色矩形反而变小。

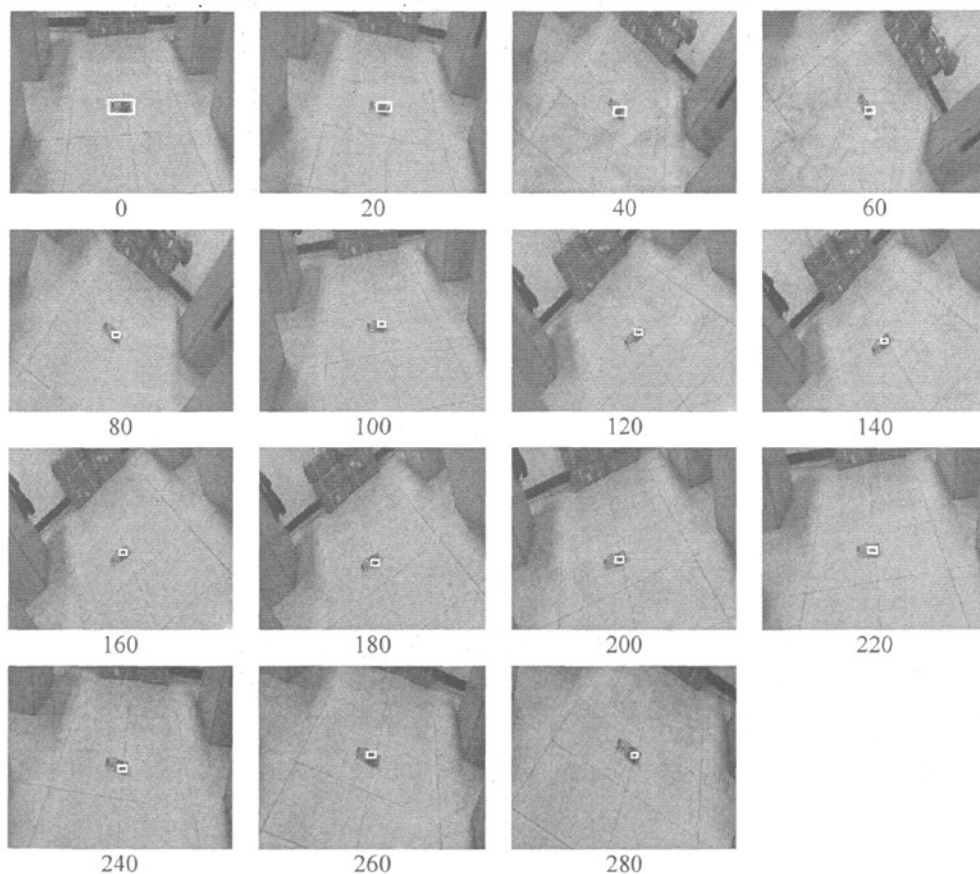


图6-7 Mean shift 加减10%跟踪结果

图6-8.0为序列的第0帧图像，其中白色矩形为初始手动选取目标。用Collins提出的Mean Shift方法对目标进行跟踪，其中图6-8.20至图6-8.280分别为序列图像中从第0帧开始间隔20帧的跟踪结果（用白色矩形表示）。可以看出，Collins的方法可以根据目标区域大小变化做出相应调整，但方向信息不准确。

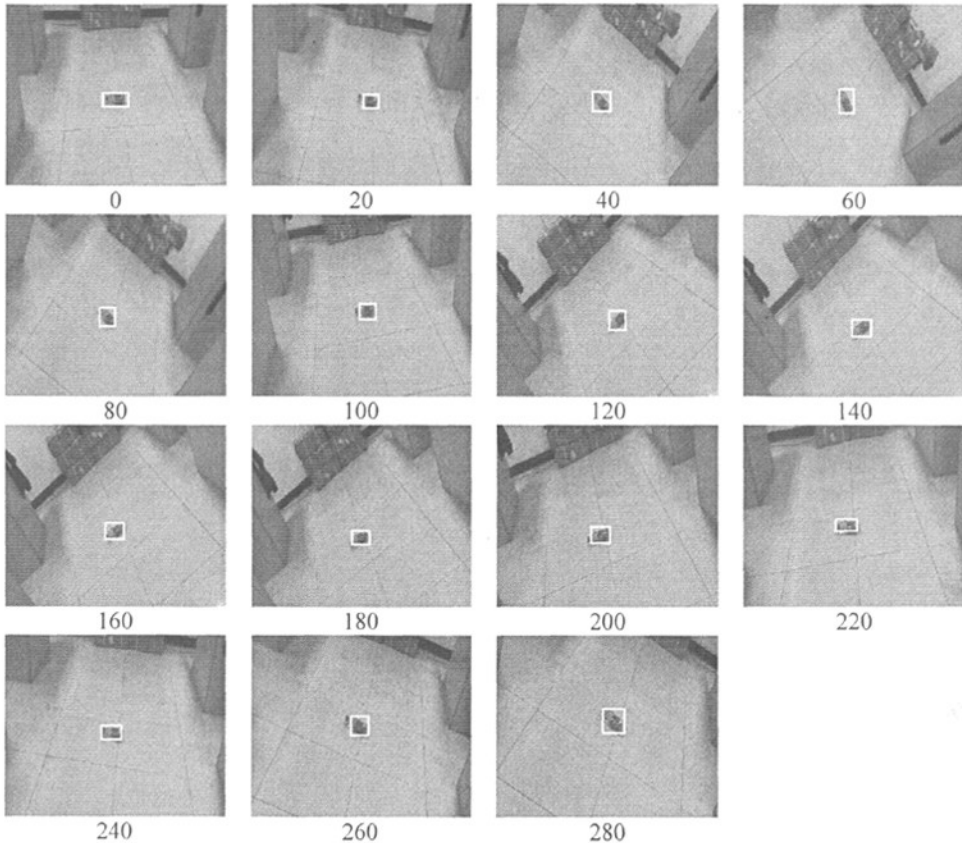


图6-8 Collins算法跟踪结果

6.5.2 跟踪结果及分析（二）

在桌子上放置了不同大小颜色的若干目标，从中选取U盘和读卡器两个目标进行跟踪，拍摄过程中摄像机相对于目标有一定的角度和距离变化，其中U盘的跟踪结果用白色椭圆表示，读卡器的跟踪结果用灰色椭圆表示。图6-9为序列的第0帧图像，白色椭圆和灰色椭圆分别为利用6.2节算法计算得到的两个目标初始最佳椭圆描述区域，可以看出椭圆的大小和方向与目标较为一致。获得目标初始最佳描述后，利用本文提出的跟踪算法对后续图像进行逐帧跟踪，其中图6-9.20至图6-9.140分别为序列图像中从第0帧开始间隔20帧至140帧的跟踪结果。可以看出，新算法可以较好地跟踪目标，且跟踪结果和目标的位置、大小及方向基本一致。图6-10

为正负10%Mean Shift跟踪算法的跟踪结果，图6-11为Collins的算法跟踪跟踪结果。其中正负10%Mean Shift跟踪算法结果出现类似于瓶子目标的情况，而Collins算法结果虽然可以适应目标缩放变化，但方向信息不够准确。

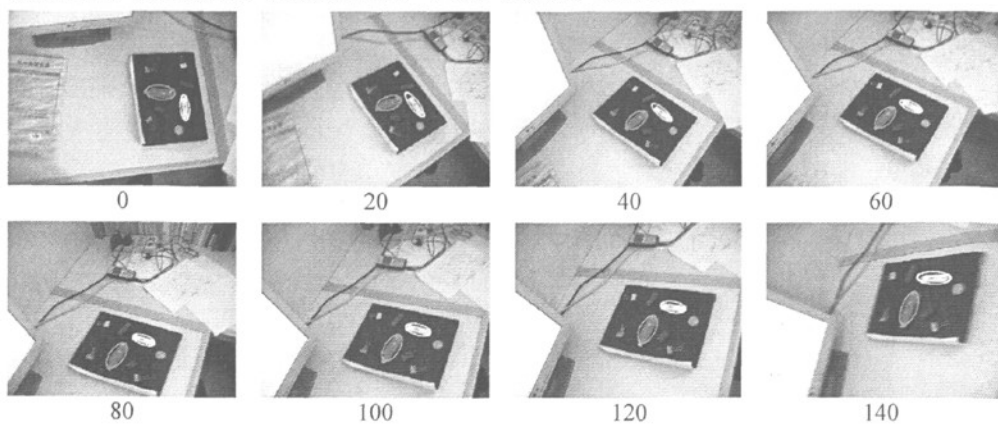


图6-9 本文算法跟踪结果

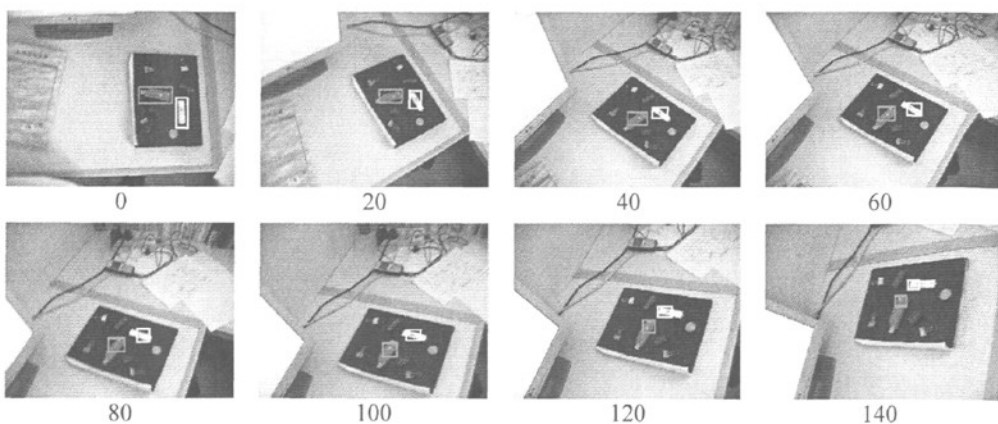


图6-10 加减10%Mean Shift算法跟踪结果

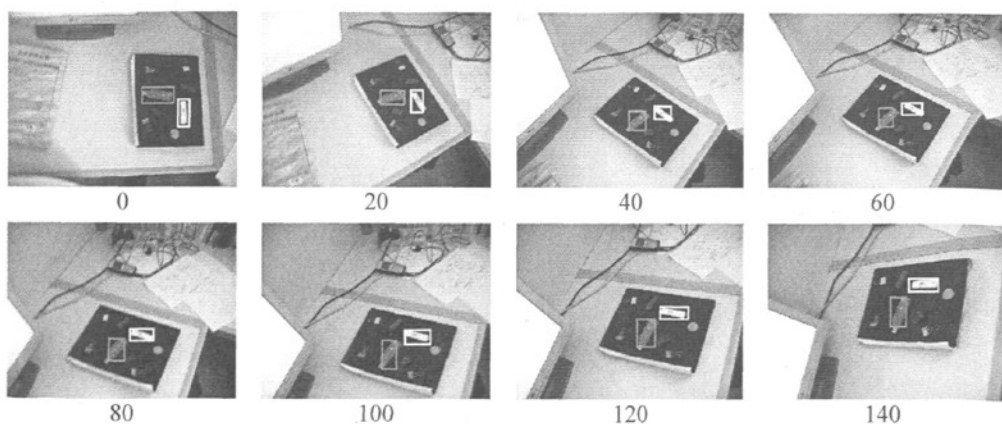


图6-11 Collins算法跟踪结果

6.6 本章小结

本章简要分析了现有 Mean Shift 算法及其改进跟踪算法对于存在缩放、旋转变化目标存在进行跟踪时的处理方法。为了准确得到目标描述区域，提出了一种目标最佳椭圆描述区域计算方法。新算法可以较好地计算出目标的最佳描述区域，同时给出目标大小和方向信息。针对 Mean Shift 经典算法及其改进算法不能同时给出目标大小和方向信息的缺点，提出了一种尺度方向自适应 Mean Shift 跟踪算法。实验表明，新算法可以稳健、有效地适应目标的旋转、缩放等复杂运动。但是新算法还有一些不足，比如由于计算量较大、还不能满足实时性要求、复杂背景下弱目标跟踪效果不理想，这些将是我们下一步的研究重点。

第七章 基于摄影测量的无人机辅助着陆系统

目前, 如何实现无人机安全平稳着陆是无人机技术发展迫切需要解决的问题之一。着陆时, 由于操纵复杂、地面干扰因素多, 而事故频繁。要解决这个问题, 必须高精度地获取在着陆过程中无人机的位置和姿态信息。本章针对这个问题进行研究, 设计并实现了一种基于摄影测量的无人机辅助着陆导航系统, 同时提出了一种基于正负正则化LOG算子小目标实时高精度检测算法。

本章内容的结构安排如下: 7.1 节主要介绍了基于摄影测量的无人机辅助着陆导航系统的总体方案; 7.2 节介绍了系统涉及的关键技术; 7.3 节给出了无人机辅助着陆导航系统及小目标实时检测跟踪算法的实验结果及分析; 7.4 节对本章进行了小结。

7.1 基于摄影测量的无人机辅助着陆航系统总体方案

7.1.1 需求分析

本项目的研究目的是建立稳定、可靠的辅助着陆导航系统, 即在无人机着陆过程中实时确定无人机的三维位置为其顺利着陆提供支持, 同时实时保存无人机着陆过程图像为事后分析提供数据支持。

实时高精度地测量无人机在降落过程中的空间位置是有效解决无人机辅助着陆问题的关键之一。在无人机降落过程中, 可以实时采集并存储无人机着陆过程的图像, 并通过光测方法实时高精度测量得到无人机相对于跑道的三维位置信息。该系统的价值主要体现在以下两个方面:

(1) 将测量结果以图像和数据的方式提供给着陆指挥官, 为着陆指挥官引导无人机着陆提供帮助;

(2) 实时记录无人机着陆过程, 为事后分析提供数据支持。

综合以上分析, 将无人机辅助着陆系统的需求总结如下:

1、功能性要求

要求实时高精度地测量无人机着陆过程中飞机的空间三维位置, 同时为了方便事后分析, 需要实时保存无人机降落过程的图像。

2、主要技术指标

(1) 图像存储: 实时存储无人机着陆过程图像, 丢帧率小于 0.1%。

(2) 图像处理帧频: 25 帧/秒。

(3) 空间定位精度: X、Y、Z 三个方向定位精度均优于 1 米。

7.1.2 总体方案

无人机在着陆过程中，需要高精度地保持在着陆下滑面内，才能安全、顺利地着陆。无人机下滑面如图 7-1 所示。

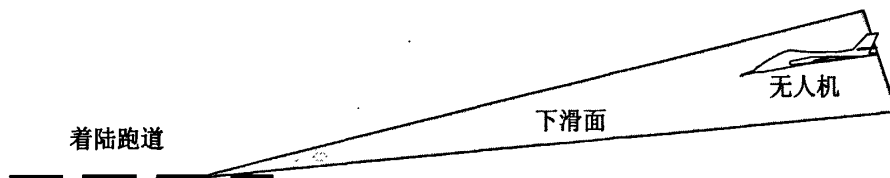


图 7-1 无飞机沿下滑面着陆示意图

实验中，在着陆地点附近设置两台摄像机（跑道两侧各设一台），利用两台摄像机同时拍摄同一空间目标，通过目标检测跟踪算法得到目标在图像上的成像位置，再根据事先标定像机得到的像机参数，通过双目交会测量无人机的三维位置，如图 7-2 所示。

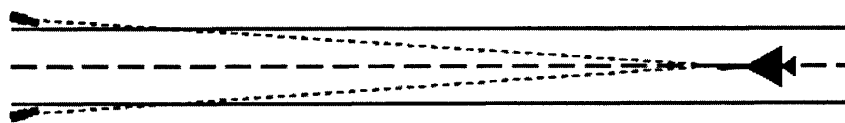


图 7-2 双目交会测量无人机三维位置

在获得无人机的位置参数后，即可根据无人机在图像上的位置和下滑面在图像上的成像区域，判断飞机所处位置与下滑面的相对关系，为无人机平稳着陆提供参考信息。

7.1.2.1 硬件系统组成

本实验需要沿着跑道方向架设两台摄像机拍摄着陆区域，同时需要三台计算机，两台用于目标跟踪，一台用于交会测量，还需要一套网络传输设备。

为了提高无人机位置参数求解的可靠性、实时性和精度，特别是为了高效可靠的自动识别、提取目标特征，需要在无人机上设置特性已知的合作标志。合作标志要有显著的易于识别的特征。合作标志应该满足以下几点要求：

- 1、固定在无人机上的稳定点；
- 2、与无人机上 GPS 测量点关系明确；
- 3、满足拍摄要求，即在无人机进入视场到最后成功着陆，整个过程都应能清晰的成像。

考虑以上因素，合作标志选择为在无人机上加装着陆辅助灯。为了验证无人机辅助着陆系统的有效性，在实验中通过测量着陆辅助灯的空间位置，并将结果与无人机上 GPS 数据进行比对，其中着陆辅助灯与 GPS 安装点的相对位置关系可

以通过预先测量获得。

7.1.2.2 系统硬件布置

无人机辅助着陆导航系统硬件布置如图 7-3 所示,由三台连网的计算机及两台摄像机组成。两套摄像机系统分别由一台分计算机控制,通过网络传输系统与主控计算机相连。

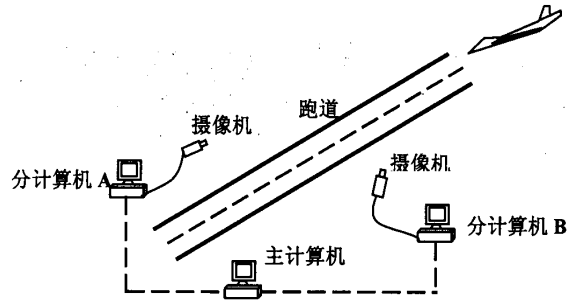


图 7-3 无人机辅助着陆系统硬件布置示意图

7.1.2.3 系统工作流程

先初始化设备,等待无人机进入视场。当无人机进入两台摄像机视场时,分计算机 A、B 开始实时采集保存无人机着陆过程图像,并对无人机上辅助着陆灯进行自动检测和实时跟踪,将跟踪结果传给主计算机。主计算机对跟踪结果进行交会处理,实时获取无人机的位置信息。工作流程如图 7-4 所示。

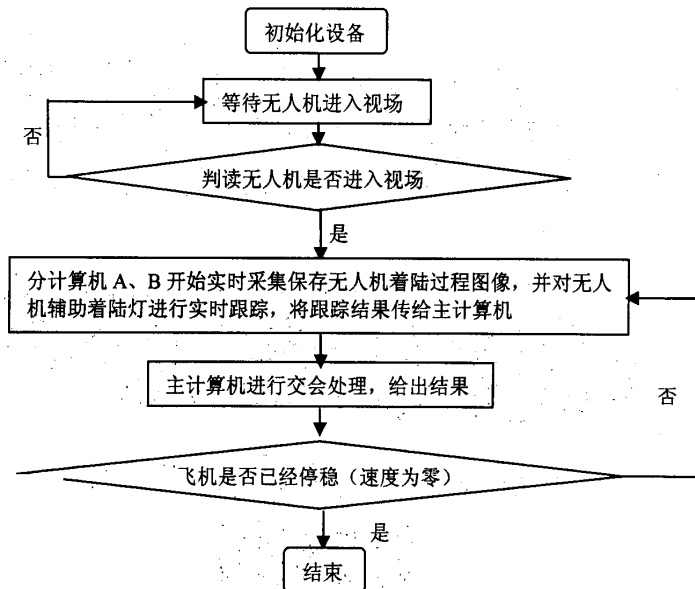


图 7-4 系统工作流程图

7.1.2.4 影响测量性能的因素

影响系统测量性能的因素有很多，其中最重要的有以下几个方面：

1、摄像机标定

摄像机标定是摄影测量的一项最基本、也是最重要的任务。因此，摄像机标定优劣直接关系到最后测量结果精度的高低。摄像机标定是一项系统而复杂的工作，特别是在本实验中，由于视场大、测量精度要求高，摄像机标定工作更加困难。因此，摄像机标定是影响系统测量精度的一个重要因素。

2、小目标检测与跟踪精度

测量过程中，无人机由远及近飞向跑道。辅助着陆灯在图像上开始表现为弱小亮点目标。随着无人机逐渐接近摄像机，该弱小亮点目标慢慢从小到大演变为光团目标。目标检测算法不仅需要较高的定位精度，而且需要适应目标尺寸变化。因此，目标高精度定位算法是影响系统测量精度的又一个重要因素。

3、系统实时性

在整个测量过程中，需要实时测量出无人机位置参数，同时还需要实时保存图像。实验中，我们采用的是帧频为每秒 25 帧的像机。除去图像传输和图像存储等时间消耗，留给单帧图像处理的时间大约为 12 毫秒，因此要求目标检测定位算法具有极高的实时性。

7.2 系统涉及的关键技术研究

7.2.1 双目交会测量技术

4.1.1 节简要介绍了常用坐标系及其相互之间的变换关系。下面在此基础上介绍双目交会测量基本理论。

双目交会测量简单的说就是在对摄像机参数进行正确标定的情况下，利用两台摄像机从不同角度对同一物体成像，通过线线交会求解物体三维位置，其原理图如图 7-5 所示。

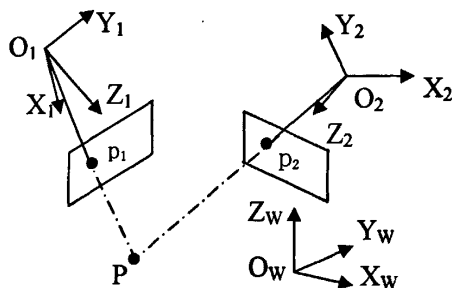


图 7-5 双目交会测量示意图

其中物点 P 为射线 $\overline{O_1P_1}$ 和射线 $\overline{O_2P_2}$ 的交点。 O_1 、 O_2 分别为位于两个不同角度的摄像机的光心， p_1 、 p_2 是物点 P 分别在两台摄像机上成像的像点。 (x_w, y_w, z_w) 为物点 P 的世界坐标系。 (X_i, Y_i, Z_i) 为物点 P 在第 i 个摄像机的成像坐标系。

设空间目标 $P(x_w, y_w, z_w)$ 由第 i 个摄像机拍摄的图像点为 $p_i(X_i, Y_i)$ ， 则共线方程可转化为如下关于 $P(x_w, y_w, z_w)$ 的方程组：

$$\begin{cases} (r_{31}X_i - f_i r_{11})x_w + (r_{32}X_i - f_i r_{12})y_w + (r_{33}X_i - f_i r_{13})z_w = f_i t_x - t_z X_i \\ (r_{31}Y_i - f_i r_{21})x_w + (r_{32}Y_i - f_i r_{22})y_w + (r_{33}Y_i - f_i r_{23})z_w = f_i t_y - t_z Y_i \end{cases} \quad (7.1)$$

其中 r_{ij} (i 取 1、 2、 3， j 取 1、 2、 3)、 f_i 、 t_x 、 t_y 和 t_z 均为通过标定得到的摄像机参数， (X_i, Y_i) 是通过目标检测跟踪算法得到的目标在图像上的成像坐标。 因此(7.1)式中的未知数只有目标的世界坐标 (x_w, y_w, z_w) 。 对于单像机， (7.1) 式表示的方程组为不定方程组， 方程有两个， 而未知数有三个， 没有唯一解。 当两台像机进行交会时， 方程的个数变为四个， 多于未知数的个数， 方程组变为超定方程组， 可以用最小二乘法求解得到目标空间坐标。

7.2.2 像机远场高精度标定技术

在摄影测量中， 为了从图像中提取空间物体的几何信息， 必须建立图像中像点位置和空间物体表面点位置的对应关系。 这种对应关系由像机成像模型决定， 其中像机参数需要通过实验确定。 确定像机内外参数的过程称为摄像机系统标定。 像机标定是摄影测量的一项最基本、 也是最重要的工作。 下面简要介绍一下本系统的标定方案。

7.2.2.1 摄像机外参数标定

标定所需设备为： 电子经纬仪、 标杆、 皮尺。 首先在像机塔架上安装电子经纬仪， 并按照事先计算的预定姿态角调整电子经纬仪的角度。 然后通过电子经纬仪在视场范围内确定标杆位置和标志点的位置。 再换上摄像机， 将摄像机图像中心对准标杆上的标志点。 此时摄像机的方位角即为电子经纬仪确定的姿态角。 此方法不能标定出像机的滚转角， 所以事先要调整像机， 使像机的滚转角为零。 而摄像机光心的坐标可用 DGPS 测出。 通过以上步骤可以获得摄像机外参数， 即光心坐标和光轴角度。

7.2.2.2 摄像机内参数标定

将视场范围内三个或三个以上的可测标志点的大地坐标和像机自身的坐标参数作为已知参数， 可以按附录 A 算法标定出摄像机内参数。 其中标志点布局为： 在跑道上摄像机视场范围内竖立两根标杆， 并在每根标杆上贴上两个标志。 每个

标杆的位置坐标由 DGPS 测量。标定布局如图 7-6 所示。

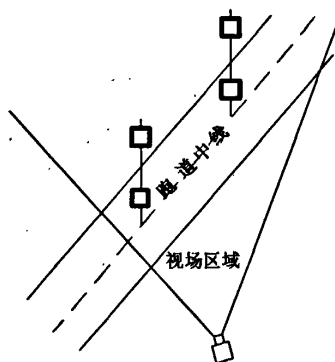


图 7-6 标定布局示意图

1、标杆位置确定

根据视场设计及镜头选择，像机聚焦点定在距离跑道坐标系原点 300 米远处。由视场计算结果知像机测量范围为最远距离 871.4 米、最近距离为 139.2 米。为了保证标定精度，相应地标定点应设在聚焦点的两侧，通过 DGPS 测得两个标定点距坐标原点的距离分别为：245.46 米和 345.61 米，如图 7-7 所示，其中 $OXYZ$ 为跑道坐标系。

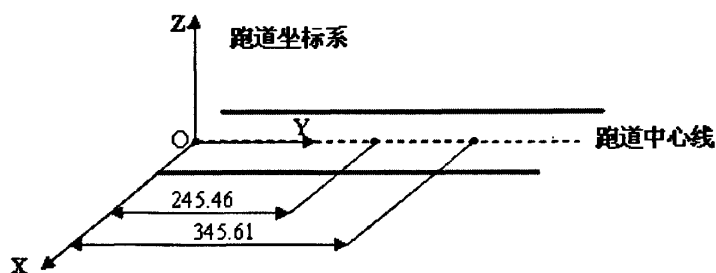


图 7-7 标定点布局示意图 (单位: 米)

2、标志点制作

两根标杆上分别固定两个标志块，作为标定参考标志点。根据视场分辨率，应保证每个标志块在图像上至少有 3×3 像素大小。为了增强对比度从而提高标志点提取精度，将每个标志块的中心涂成白色标记，底板涂成黑色。实验中，根据图像质量（受天气影响较大）选择标志块的上边缘、下边缘或者标志块的中心作为标定标志点。

精确测量标志点坐标是保证标定精度的前提。实验中，需要先在跑道中线上确定标杆放置的位置，标定点坐标由 DGPS 测得，精度为 1 厘米。标志点固定在标杆上的位置可通过标尺测得，精度为 1 毫米，其中标志点相对于标杆的关系如

图 7-8 所示。

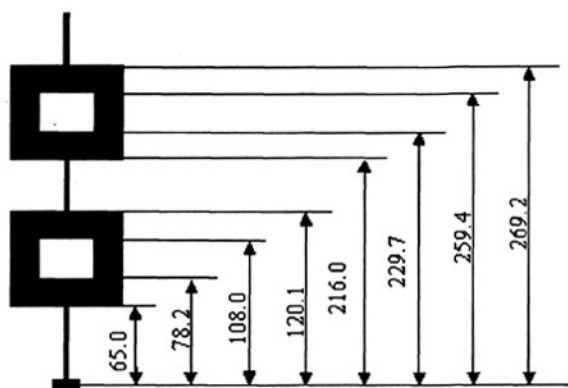


图 7-8 标志点与标杆的相对位置示意图 (单位: 厘米)

3、标定精度验证

图 7-9 为一幅实际标定图像, 白框中的白点即为标定标志点。从图像上提取标志点位置, 根据已知标志点空间坐标和像机光心坐标, 可以解算摄像机的内参数。

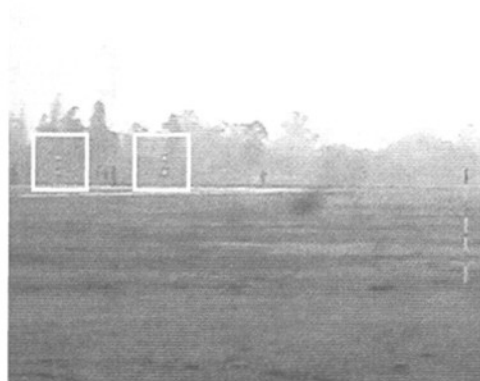


图 7-9 实拍标定图像

为了验证对摄像机系统标定的精度是否达到实验的要求, 我们在跑道设立一些验证点 (验证点坐标已知), 在验证点上竖立标杆, 进行双像机交汇测量标志点坐标, 与真实坐标进行比较。如果 Y 方向测量误差小于 15 厘米, X, Z 方向测量误差小于 5 厘米, 则认为标定成功, 否则需重新进行标定。表 7-1 给出像机标定验证结果。

表 7-1 像机标定结果验证 (单位: 米)

| 验证点 | 坐标 | 光测值 | 真实值 | 误差 | 各点坐标 | 误差 | |
|-----|----|-------|-------|--------|------|--------|-------|
| | | | | | | 平均值 | 标准差 |
| 1 | X | -0.02 | 0.000 | -0.020 | X | -0.012 | 0.044 |

| | | | | | | | |
|---|---|--------|---------|--------|---|--------|-------|
| | Y | 230.46 | 230.55 | -0.090 | | | |
| | Z | 0.99 | 0.931 | 0.059 | | | |
| | X | -0.03 | 0.000 | -0.030 | | | |
| 2 | Y | 230.48 | 230.55 | -0.070 | | | |
| | Z | 2.55 | 2.445 | 0.105 | | | |
| | X | 0.06 | 0.000 | 0.060 | | | |
| 3 | Y | 320.65 | 320.75 | -0.100 | Y | -0.105 | 0.125 |
| | Z | 0.88 | 0.931 | -0.051 | | | |
| | X | 0.04 | 0.000 | 0.040 | | | |
| 4 | Y | 320.68 | 320.750 | -0.070 | | | |
| | Z | 2.56 | 2.445 | 0.115 | | | |
| | X | -0.03 | 0.000 | -0.030 | | | |
| 5 | Y | 400.18 | 400.100 | 0.080 | Z | 0.059 | 0.058 |
| | Z | 1.08 | 1.02 | 0.060 | | | |
| | X | -0.05 | 0.000 | -0.050 | | | |
| 6 | Y | 400.2 | 400.100 | 0.100 | | | |
| | Z | 2.6 | 2.535 | 0.065 | | | |
| | X | -0.05 | 0.000 | -0.050 | | | |

4、实际标定时遇到的的问题及解决措施

标定时出现的问题主要有两方面：一是标志点的提取精度，二是标杆的放置精度。由于距离较远，视场较大以及天气影响下（大雾等情况），在采集的图像上，标志点边缘模糊，有时标志块白色区域很清楚，有时黑色区域比较清楚。解决措施：一是调整光圈及焦距，尽量提高拍摄质量；二是选择清晰的边界中心或区域中心作为标志点。标杆的放置精度对标定结果的影响也不容忽视。为获取标志点的高精度坐标，标杆需垂直竖于跑道中心线上。如果偏离，对标定结果有较大影响。解决措施：用固定底座将标杆竖立在事先选好的标定点上，再在跑道中线处架设一台电子经纬仪，校正标杆垂直度，调整标杆底座直至标杆垂直于跑道平面。

7.2.3 小目标自动检测与跟踪技术

小目标检测与跟踪技术是计算机视觉领域研究的热点和难点之一。近些年，许多国内外学者致力于该问题的研究，提出了一些具有实际意义的算法。但是本实验除了对实时性有较高要求之外，对目标定位精度以及目标大小自适应也有很高的要求。

实验中选用的摄像机帧频为 25 帧/秒，除去图像传输存储等方面消耗的时间，留给单帧图像处理时间仅有 12 毫秒左右。为了达到实时性要求，即在 12 毫秒内对无人机进行检测和跟踪定位，我们在无人机上加装辅助着陆灯作为合作标志点，将无人机检测与跟踪问题简化为亮点目标检测与跟踪问题。

无人机着陆过程，飞机会逐步接近像机。在图像上无人机辅助着陆灯从远处

点目标慢慢变大, 逐步演变一个较大的光团目标。因此, 还要求小目标检测与跟踪技术应具有适应点目标到光团目标变化的能力。

为此, 我们对现有经典算法, 就稳定性、目标尺寸变化适应性、跟踪精度及实时性等方面进行了认真分析和实验检验, 发现现有经典算法均难以满足系统要求。针对该问题, 本文提出了一种基于正则化正负 LOG 算子(positive and negative normalized LOG, 简称 PNNLOG 算子)的点目标实时检测与跟踪算法。该算法对亮点检测具有较好的稳定性、实时性, 同时检测精度高, 而且还具有适应目标尺寸变化能力。下面先简要介绍一下 PNNLOG 算子, 然后详细介绍对基于 PNNLOG 算子小目标实时高精度检测跟踪算法。

7.2.3.1 正则化正负 LOG 算子及其性质

(1) LOG 算子 (Laplacian of Gaussian)

LOG 算子的基本实现是首先用归一化后的高斯函数对图像进行平滑, 其中归一化高斯函数表达式如下:

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (7.2)$$

对图像 $f(x, y)$ 卷积得:

$$L(x, y) = f(x, y) * G(x, y, \sigma) \quad (7.3)$$

其中 * 表示卷积运算符。第二步是用拉普拉斯算子, 获取平滑图像的二阶方向导数图像 $M(x, y)$:

$$M(x, y) = \nabla^2 L(x, y) = \nabla^2 [f(x, y) * G(x, y, \sigma)] = \nabla^2 [G(x, y, \sigma)] * f(x, y) \quad (7.4)$$

其中 $\nabla^2 [G(x, y, \sigma)]$ 即为 LOG 算子, 其表达式可写为:

$$LOG(x, y, \sigma) = \frac{x^2 + y^2 - 2\sigma^2}{2\pi\sigma^6} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (7.5)$$

(2) 正则化正负 LOG 算子

由 Lindeberg 尺度空间理论可知, LOG 算子进行正则化以后得到的新算子即 NLOG (Normalized Laplacian of Gaussian) 算子具有尺度不变性。由于 NLOG 算子为中心负周围正的倒置墨西哥草帽型, NLOG 算子并不能用于亮点目标检测。为了检测跟踪亮点目标, 需要对 NLOG 进行负运算, 即正的变负、负的变正。由此可以定义正则化正负 LOG 检测算子(Positive and Negative Normalized LOG, 简称 PNNLOG 算子)为:

$$PNRLOG = \begin{cases} \frac{2\sigma^2 - x^2 - y^2}{2\pi\sigma^4} e^{-(x^2+y^2)/2\sigma^2} \text{亮目标} \\ \frac{x^2 + y^2 - 2\sigma^2}{2\pi\sigma^4} e^{-(x^2+y^2)/2\sigma^2} \text{暗目标} \end{cases} \quad (7.6)$$

其中 σ 为 PNNLOG 算子的尺度参数。两个正则化 PNNLOG 算子的检测原理一样，只是对于不同的光团目标选用不同的算子，其中对于相对于背景为亮光团用负 PNNLOG 算子进行检测，而对于相对于背景为暗光团用正 PNNLOG 算子进行检测。

(3) PNNLOG 算子性质研究

性质一：对同一尺度的高斯亮点用不同 PNNLOG 算子做卷积，当 PNNLOG 算子的尺度参数 σ 和高斯亮点参数 σ_g 相等时，响应值最大。其数学证明请参考附录 B PNNLOG 算子性质证明。

图 7-10 中每条曲线代表着同一尺度高斯点的不同尺度的 PNNLOG 算子响应曲线，其中四条响应曲线从左到右对应的高斯点的尺度逐渐增大。PNNLOG 算子的尺度参数 σ 从 0.25 开始，间隔 0.25 增加至 25，计算这些算子对该高斯点的响应。实验结果表明，对于同一尺度高斯点的响应最大值处，PNNLOG 算子的 σ 和高斯点的 σ_g 相等。同时还发现，响应最大值的大小与高斯点的 σ_g 大小无关。

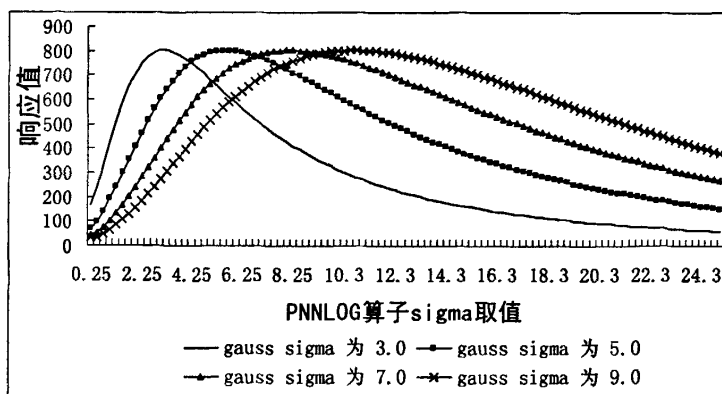


图 7-10 不同尺度的 PNNLOG 算子对同一高斯点的响应

性质二：对不同尺度高斯亮点用同一尺度参数的 PNNLOG 算子做卷积，PNNLOG 算子 σ 与高斯亮点参数 σ_g 相等时，响应值最大。特性二的证明过程与特性一的证明过程相似。

图 7-11 为同一尺度的 PNNLOG 算子对不同尺度的高斯点进行运算的响应曲线，其中五条响应曲线从左到右所对应的 PNNLOG 算子的尺度逐渐增大。高斯点参数 σ_g 从 0.2 开始，间隔 0.2 增加至 12。实验中发现，当 PNNLOG 算子的 σ 和高斯点的 σ_g 相等时，响应值最大。同时不同尺度的 PNNLOG 算子的响应最大值与

PNNLOG 算子参数 σ 大小无关。

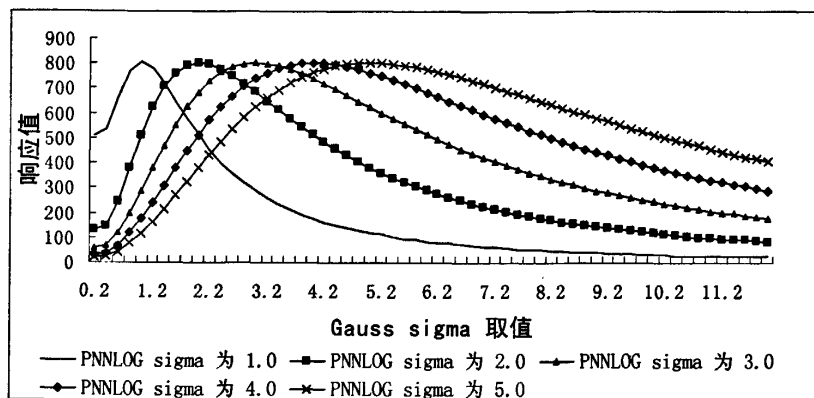


图 7-11 同一尺度的 PNNLOG 算子对不同尺度高斯点的响应

性质三：对于圆形平台目标，PNNLOG 算子也具有类似性质一的特性，即对同一半径的圆形平台目标用不同 PNNLOG 算子做卷积，当 PNNLOG 算子的尺度参数 σ 和圆形平台目标半径相等时，响应值最大。

图 7-12 中每条曲线表示同一半径的圆台不同尺度的 PNNLOG 算子响应曲线，其中五条响应曲线从左到右所对应的圆台的半径逐渐增大。实验中发现，当圆形平台目标半径小于 11 的时候，在同一圆形平台目标的响应最大值处，PNNLOG 算子的 σ 和圆台半径相等。而当圆形平台目标半径大于 14 的时候，在同一圆台的响应最大值处 PNNLOG 算子的 σ 略小于圆形平台目标半径。同时对于不同大小的圆形平台目标，PNNLOG 算子的响应最大值的大小相差不大。

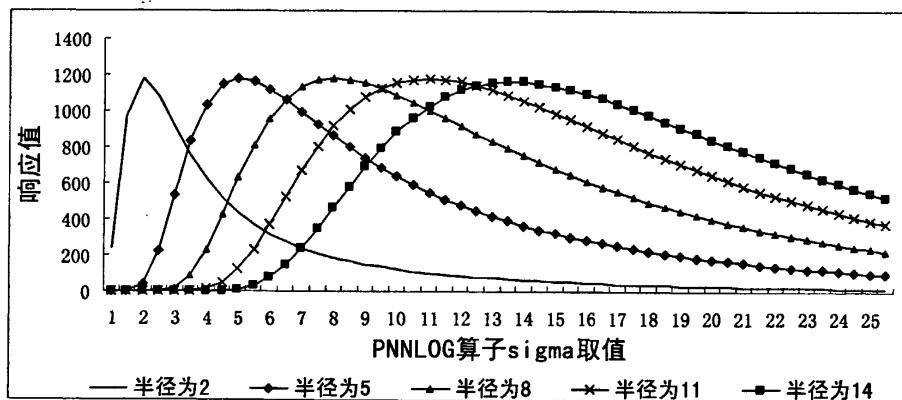


图 7-12 不同尺度 PNNLOG 算子对同一半径圆台的响应

性质四：对于灰度不同的尺度参数相同的高斯亮点，当 PNNLOG 算子的尺度参数 σ 和高斯亮点参数 σ_g 相等时，响应值最大。

构造仿真高斯点图像，参数 σ_g 为 3.0，灰度从 1 开始间隔 1 增加到 255，背景灰度为 0。对每一帧高斯亮点图像，构造 σ 从 0.5 开始间隔 0.2 逐步增加到 20.3 的

PNNLOG 算子, 计算这些算子对该高斯点的响应。为了方便显示, 这里只抽取了几个响应曲线, 如图 7-13 和图 7-14 所示。当高斯点灰度从 2 增加到 255 时, 每个高斯点的不同尺度 PNNLOG 算子响应曲线趋势类似, 但是响应值大小存在差别。当 PNNLOG 算子的尺度参数和高斯点尺度参数相同时, 响应值最大。然而当高斯点灰度等于 1 时, PNNLOG 算子尺度参数在 0.5 时响应最大。实验结果说明当高斯点灰度大于 1 时, PNNLOG 算子可以对高斯点作出正确响应。

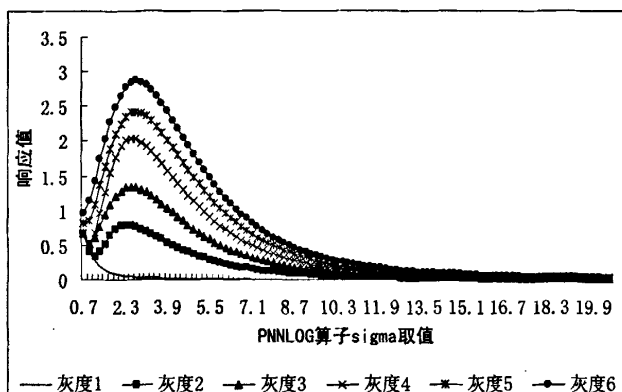


图 7-13 不同灰度高斯点的不同尺度 LOG 算子响应曲线 (一)
(灰度从 1 到 6)

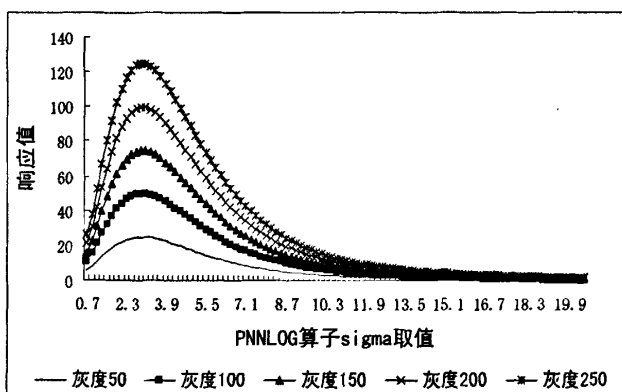


图 7-14 不同高斯点不同尺度 LOG 算子响应曲线 (二)
(灰度分别为 50、100、150、200 和 250)

7.2.3.2 基于正则化正负LOG算子的小目标检测跟踪算法

从光学角度来看, 当目标距离像机非常远时, 目标可模型化为一个二维空间的点源, 在图像上通常表现为相对于背景的明 (或暗) 高斯点或高斯平台。当图像中混杂着不同大小的虚假目标, 而只对某些特定尺寸的目标感兴趣时, 或者当目标迅速接近像机, 图像上目标尺寸从小迅速变大时, 现有方法难以适应。针对此问题, 本文提出了一种基于 PNNLOG 算子的小目标实时高精度检测与跟踪算法。

该算法步骤如下:

(1) 检测真实目标

初始时刻, 由于目标距离摄像机较远, 图像上目标较小, 而且短时间内可以认为连续几帧中目标做近似直线运动。因此可以先利用 Top-Hat 算法检测出初始连续几帧中包含真实目标的候选区域。然后利用连续几帧内目标做近似直线运动的性质, 可以判断出真实目标。检测出目标后, 以目标区域的形心作为目标点初始位置, 记为 A。

(2) 真实目标响应值及 PNNLOG 算子参数的确定

为了获得 PNNLOG 算子的初始参数 σ , 利用 PNNLOG 算子参数 σ 与圆台目标半径相等时响应最大的性质求取 PNNLOG 初始参数。具体做法为: 利用 QTSU 算法^[205]对包含真实目标的小区域进行二值处理, 统计目标区域面积大小, 利用圆面积公式计算出目标区域的等效圆半径 r , 将 PNNLOG 算子的 σ 设定为 $1.2 \times r$ 。选取一组 $\sigma + 0.5n$ (其中 n 取 $-3, -2, -1, 0, 1, 2, 3$), 构造一组 PNNLOG 算子。对于 A 点的 5×5 邻域内每一点, 计算新构建的一组 PNNLOG 算子的响应 (共 25×7 个响应值), 其中最大响应值所对应的 PNNLOG 算子相对于真实目标最为适合, 该算子所对应的参数 σ 为最佳响应算子尺度, 同时图像上所对应的具有最大响应值的像素点为真实目标点。

(3) 目标高精度定位

利用 PNNLOG 算子检测目标, 仅能得到整像素精度。为了得到更高的定位精度, 我们以检测出的目标点为中心选取 3×3 的区域, 利用更新后的 PNNLOG 模板计算该 3×3 区域内每点的响应值, 对计算出的响应值进行曲面拟合, 将拟合曲面的极值位置作为当前帧的目标位置, 其定位精度为 0.1 像素。其中曲面拟合算法请参考文献[206]。

(4) 后续帧目标检测定位

由于相邻两帧时间间隔很短, 可以认为相邻两帧图像上目标大小变化并不大。因此当检测出真实目标后, 可以利用上一帧最佳响应 PNNLOG 算子对后续帧进行检测得到真实目标的位置及响应值。具体做法如下:

(a) 目标检测。在当前帧中, 以上一帧真实目标位置为中心, 选取搜索区域。在搜索区域内利用上一帧的最佳 PNNLOG 算子对区域内每点进行卷积运算, 获得每点的响应值, 将具有最大响应值的像素点作为当前帧中真实目标初始位置。

(b) PNNLOG 算子参数更新。由于目标逐渐接近像机, 图像上目标越来越大, 如果用固定尺度的 PNNLOG 算子对整个序列图像进行检测, 后期将会失效, 所以必须要逐帧更新 PNNLOG 算子的尺度参数。假设前一帧最佳 PNNLOG 算子参数为 σ_{i-1} , 选取一组参数 $\sigma_{i-1} + 0.5n$ (其中 n 取 $-3, -2, -1, 0, 1, 2, 3$), 构造一组 PNNLOG

算子,对初始点计算响应值,选取具有最大响应值的 PNNLOG 算子参数作为当前帧目标的 PNNLOG 算子参数;

(c) 目标高精度定位。利用更新后的 PNNLOG 算子按第三步方法对目标进行高精度定位。

(5) 目标丢失判断

由于光照、遮挡等原因,可能出现目标丢失情况。在考虑目标空间位置和响应大小两个因素的基础上,本文提出了一种判断目标丢失的简单准则。具体思路如下:相邻两帧同一真实目标变化不大,由 PNNLOG 性质可知其 PNNLOG 响应值差别也应不大;而且由于目标是连续平稳运动,在相邻两帧中同一目标不会移动较大距离。具体做法为:假设上一帧目标没有丢失,按(7.7)式计算当前帧检测出的“目标”的响应值及位置和上一帧目标的响应值及位置之间的相似值。若相似值大于预先指定的阈值,则认为当前帧中目标没有丢失,否则目标丢失。

$$Z = 1 / \left[1 + \left(\frac{D_{i,j-1}}{C} \right)^2 \times \left(\frac{R_{i-1} - R_i}{R_{i-1}} \right)^2 \right] \quad (7.7)$$

其中 $D_{i,j-1}$ 表示当前帧检测出的“目标”和上一帧真实目标之间的图像上的距离, C 为距离常数(一般取为 30,具体视目标运动情况选取)。 R_{i-1} 为上一帧真实目标的响应值,而 R_i 为当前帧中检测出的“目标”的响应值。

(6) 目标丢失后处理

当目标丢失后,可以以最后一帧未丢失目标时小目标的位置为中心,选取一个更大的搜索区域,并根据最后一帧未丢失目标时刻的最佳 PNNLOG 算子选取一组 $\sigma + 0.5n$ (其中 n 取 $-3, -2, -1, 0, 1, 2, 3$) 对选取区域进行检测,找出具有最大响应值的像素点作为真实目标点候选位置。若该点满足第五步要求,则认为该点为真实目标;若不满足,则对下一帧进行上述处理。如果连续若干帧没有检测出真实目标,则认为目标脱离摄像机,检测结束。

7.2.4 图像实时存储技术

为了方便事后进行分析,在测量过程中需要实时保存无人机着陆过程的图像。因此,研究图像实时存储技术有着重要的意义。

在本系统中,图像存储主要分为三种模式,分别为硬盘存图模式、内存存图模式以及实时处理与存图模式。其中前两种模式都为只存图模式,而后一种模式则是边实时处理边存储图像。

7.2.4.1 硬盘存图模式

1、硬盘存图原理介绍及存在的问题

硬盘存图就是利用图形采集板从摄像机获得图像信号以后，将图像数据存储到硬盘的一个过程，如图 7-15 所示：

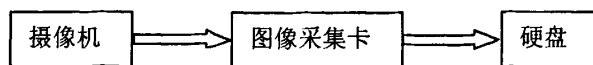


图 7-15 硬盘存图工作流程图

该存图模式工作流程非常简单，易于实现。但是在实际操作中，该存图模式会出现以下两个问题：

(1) 某些时刻图像数据不能在规定时间内存储，出现丢帧；

(2) 图像写入硬盘时，占用 CPU 和内存，从而造成系统资源紧张，影响其他工作。

造成这些问题的原因较多，经分析主要有以下两个方面：

1、目前市面上的主流硬盘，规格为 7200r/min，其最大内部数据传输率在 50Mb/s 左右，工作在 ATA100 接口标准下，可以满足 30Mb/s 以下的图像数据实时存储的任务。但是由于硬盘采用了机械部件，存在存储速度不稳定的缺陷，在某些时刻的存储速度会明显的小于正常情况下的存储速度。

2、在存图的过程中，需要频繁的进行文件操作，每一幅图像都需要保存为一个文件。而在硬盘存储的工作原理中，当直接向一个文件中导入数据时，硬盘可以处于连续的数据传输工作状态，但是当要重新生成一个新的文件时，硬盘则需要重新寻找空白地址，磁头需要切换磁道搜索新的地址或是等待盘片旋转到空白扇区。在等待和搜索的过程中，有可能存图时间会超出满足实时系统正常工作的时间。以 0.5Mb/帧图像为例进行计算，假定硬盘存储速度 30Mb/s，则硬盘完全工作状态下存储一幅图像的时间需要 16ms，一帧图像的间隙为 40ms，其中还包括图像数据从摄像机到采集卡、从采集卡到微机硬盘这部分时间，因此，当硬盘的寻址时间超过约 10ms，便会产生时间延迟。一旦此延迟时间超过了下一幅图像进入时的时间，则会出现丢帧。

2、改进的硬盘存图模式

针对硬盘存图模式存在的问题，我们展开了深入的研究。近些年，有学者提出了双缓冲池的方法来解决硬盘的数据实时存储问题。这种模式的主要原理是在内存空间中开辟双缓冲池，即使用两块内存区域，假设分别为 A、B。当一帧数据进入时，存放在 A 区域，然后将 A 区域中数据写入硬盘，而下一帧图像进入内存则存放在 B 区域，待 A 区域中数据传输完毕，再将 B 区域中数据写入硬盘。采用此种模式的好处，是将原先每帧 40ms 存图和处理时间提高到了 80ms，即留给硬盘存写的时间大大提高。但是在实际实验中，发现该模式仍然存在问题，会出现丢帧的情况。原因是此种模式没有解决硬盘频繁进行文件操作，需要寻址时间这

一根本问题。

在理论分析及实验测试的基础上,我们认为造成硬盘实时存图难以实现的根本原因在于硬盘在频繁进行文件操作时的寻址时间可能会严重影响系统的实时性。针对这一问题,我们提出了大文件存图模式。这种模式的工作原理就是在硬盘中开辟一段空间,新建一个 dat 格式的文件,当图像数据进入时,则将数据存放在 dat 文件中,所有图像数据都存放在一个 dat 文件中。存图结束后,再将 dat 文件分割成若干图像文件。这种模式的好处就在于避免了硬盘频繁的文件操作。对此种模式,我们进行了实验测试,在原硬件设备不变的情况下,较好的解决了丢帧等问题。在实验中,我们连续拍摄 25 分钟图像,存图文件达到 15.8Gb,未出现丢帧现象。

7.2.4.2 内存存图模式

内存存图模式的工作原理非常简单,就是在内存中开辟一段内存空间用于图像存储,在存图完毕以后,再将内存空间中的图像导入到硬盘,如图 7-16 所示。

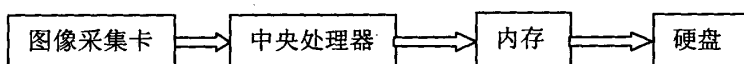


图 7-16 内存存图模式工作示意图

此模式大大节省了图像存储的时间。由于微机工作原理,中央处理器即 CPU 不能直接与硬盘交换数据,因此在前一节中硬盘存图模式,数据从采集板到硬盘,实际上中间经过了内存数据交换。相对于硬盘,内存的存储速度比较快,其读写速度可达到硬盘的十倍甚至几十倍以上。以 DDR400 内存为例,其读写最大理论速度可达到 6400M/s,实际读写速度达到约 1500M/s,而随着技术革新,这一数字仍在不断被刷新,目前市面上主流的 DDR667 内存,其实际存储速度,可达到 2000M/s 以上。在内存存图模式下,中央处理器只需将图形采集卡上得到的图像数据写入已经开辟的内存空间即可,免去了从内存到硬盘的时间,存储效率大大提高。同时,由于不需要内存与硬盘间的数据交换,这种存图模式也大大节省了系统资源,使得中央处理器和其他内存资源可以同时完成其他任务。

内存存图模式具有存图时间短、占用资源低等优点。但其存在的缺点也较为明显。主要是因为内存作为微机的动态、临时存储设备,其本身作用主要是实现数据在中央处理器与硬盘之间进行交换,不是用于数据的长时间存储。因此其存储量非常有限,并且易丢失。目前市面上主流微机,其内存支持度在 4~8Gb 之间,对于需要长时间存图的情况,这种方法不适用。

7.2.4.3 实时处理与存图模式

在前两节中已经分别介绍了硬盘存图和内存存图两种存图模式及其优缺点。

硬盘存图模式其优点是直接写入了终端存储设备，可以存储较长时间的图像。但其缺点也非常明显，在现有硬件设备情况下，受硬盘本身存储速度的影响，在存储效率上较低，同时由于硬盘读写需要占用大量的 CPU 和内存等系统资源，使得微机难以完成其他任务。

内存存图模式的优点在于存储速度非常快，大大节省了用于存储的时间和资源，使得计算机同时可以完成其他任务。以 4Gb 内存空间的计算机为例，开辟 2Gb 的内存空间作为存图空间，以每帧图像 0.5Mb 大小计算，可以存储 4000 帧图像，如果帧频 25 帧/s，则可以存储 160 秒图像，而实验中无人机飞机从 1 公里远处到着陆这段时间仅为 30 秒左右，因此在存储容量上完全满足需要。而另 2Gb 内存作为系统资源则完全满足实时处理的要求。

从以上的对比分析中可以看出，就目前的微机硬件设备，要实现实时处理与实时存图同时进行，硬盘存图这种模式暂时不能满足要求。因此我们采用了内存存图，具体实现如图 7-17 所示。在图像数据进入内存后，进行实时处理之前，将图像数据拷贝一份到开辟的内存存储空间中。仅考虑数据拷贝这一步骤，所需时间不到 1 毫秒。用此种模式，能够在实时处理实时存图的任务。

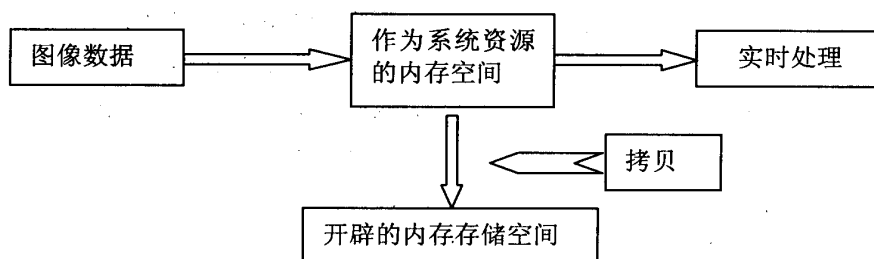


图 7-17 实时处理与存图模式工作示意图

7.2.5 天气影响及解决措施研究

7.2.5.1 天气影响问题

在实验过程中，我们遇到了不同的天气：大雾、阴天和晴天等，而且一天中不同时刻光线的变化也很大。从实验采集的图像中可以看出，光线强弱变化对拍摄效果影响较大，图 7-18 为阳光强烈时拍摄的图像，图 7-19 为阳光较弱时拍摄的图像，黑框内亮点为目标。由于受到太阳光的影响，在图像上无人机辅助着陆灯在阳光强时比较微弱，而在阳光弱时比较强。图 7-20 为飞机场大雾天气条件下所摄图片。

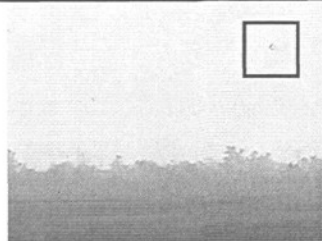


图 7-18 阳光强烈时拍摄的图像

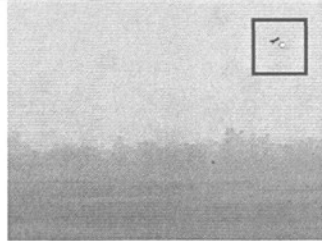


图 7-19 阳光较弱时拍摄的图像



晴天



阴天

图 7-20 大雾天气拍摄的图像

由于天气变化，对实验造成的影响总结如下：

- 1、强光直射时，捕捉不到信号灯；
- 2、图像变亮，影响目标（飞机信号灯）检测与跟踪锁定；
- 3、灯光亮度与周围环境亮度反差不大，形状不规则，影响提取精度，从而影响测量精度；

对上述影响，实验中我们主要的应对措施为：在飞机试验场，实验设备架设好后，雾还未散去，镜头光圈必须调到较大档位才能完成采图标定。而雾散去后，太阳光逐渐增强，图像变的很亮，我们适当调整光圈，这种做法虽然可适应测量要求，但是调整时应格外小心，因为调整光圈时，容易带动焦距档，引起焦距变化，先前的标定结果即不适用，影响实时测量结果。

而且标定时一般会遇到以下问题：

- 1) 当镜头对准亮处时，暗处影像不可见；当镜头对准暗处时，亮处影像变得太白；
- 2) 当镜头对准近处时，远处影像不可见；当镜头对准远处时，近处影像又很模糊；
- 3) 当镜头迎着强光时，强光周边部分图像会比较模糊。

产生这些问题的根本原因是：自然光线照度范围非常宽，从 $100,000\text{Lux}$ 到 0.00035Lux 。而传统 CCD 摄像机不能适应这种情况。

7.2.5.2 解决措施

要解决太阳光影响的问题，我们必须找到一种方便、稳定、适合工程应用的方法。通过分析调研，我们提出三种主要解决措施：一是在无人机上寻找适当位置安装辅助着陆灯，使得灯光与周围光线差别明显；二是在镜头前加载过滤片，尽量消除因光线强弱变化等带来的影响；三是选择新型的摄像机。

1、过滤片选型与实验

为了减少由于环境中的可见光等光源对人工合作标志成像的影响，摄像测量系统使用红外 LED 人工合作标志，拍摄时需要在摄像机的 CCD 前放置一块窄带滤光片，只让人工合作标志中 LED 发出的光谱范围的红外光通过。经过市场的调查研究，上海摩锐光电有限公司生产的 850 带通滤光片能够满足需求，该带通滤光片的光谱响应范围如图 7-21 所示。

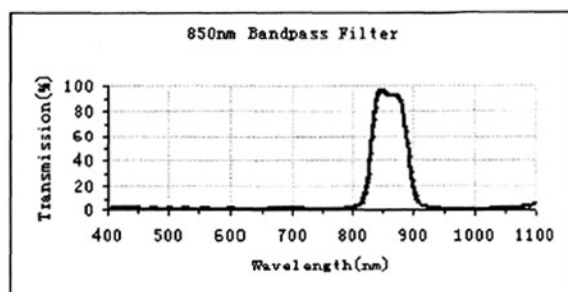


图 7-21 850 纳米带通滤光片的光谱透过率曲线

为了测试过滤片效果，我们做了一些测试实验。实验中人工合作标志的制作成十字型红外 LED。实验过程：选择一个阳光充足的天气，将四个红外 LED 十字丝装在一个矩形方框四周，摄像机迎着太阳光拍摄一组图像，在镜头前装上 850 带通的红外窄带滤光片再拍摄一组图像。图 7-22 是加滤光片前拍摄的图像，图 7-23 是加滤光片后拍摄的图像。由实验结果可以看出，加滤光片前，受太阳光的影响，四个红外 LED 十字丝在图像上显示不清；加滤光片后，极大地消减了太阳光的影响，十字比较清晰，易于精确提取十字丝。



图 7-22 加滤光片前拍摄的图像

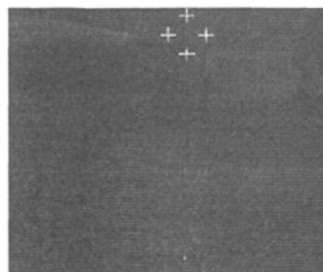


图 7-23 加滤光片后拍摄的图像

上述实验中，摄像机距红外 LED 十字丝约 30 米。对于无人机着陆实验情况，

拍摄时像机距离无人机最远可达一千多米，因此需要重新设计合作标志形状。实验中将十字丝标志改为圆形标志，大小可以依据辅助着陆灯尺寸进行调整。

2、新型摄像机的选择

基于数字像素系统 (Digital Pixel System, 简称 DPS) 技术的宽动态摄像机很好地克服了传统 CCD 摄像机不足, 它采用逐点高清晰成像, 水平分辨率可达到 480 线, 垂直分辨率可达到 400 线。即使在比较苛刻的光照条件下, 也可捕捉到清晰图像, 不会因为阴影、眩光、反射和太阳光而使图像发暗或被破坏。

DPS 摄像机就是有数字像素系统的摄像机。传统的 CCD 和 CMOS 摄像机传感器都是为每一列或每一行像素点配备一个模数转换器(A/D), 每个像素点的输出都是模拟光信号, 存在着噪声大和输出时间长等缺点。而 DPS 是基于美国 PIXIM 公司的 DPS 专利技术的图像传感器, 其核心技术是在图像传感器的每一个像素点上包含了一个 10 位 A/D 转换器, 即在 CMOS 摄像机图像传感器上的有源像素捕捉到光信号时, 直接将其放大并转换为数字信号, 从而可将阵列上的信号退化和串扰降到最小, 并允许采用更好的降噪方法。一旦数据以数字格式捕获, 就可以采用各种数字信号处理技术来重现真实图像。显然, DPS 技术中的图像传感器和图像处理器是全数字式的, 并采用 32 位 ARM CPU 精确控制每个像素, 使每个像素独立完成采样和曝光, 并直接转换为数字信号。该技术革命性的图像捕获和处理方式将摄像机的动态范围提升到史无前例的 120dB, 而传统的 CCD 摄像机的动态范围最大才为 60dB。

DPS 实现了每个像素的独立采样和曝光, 而且每个像素完成采样后直接转换为数字信号, 由于数字处理可减少信号传输噪声和损耗, 因此它可以实现快速低噪声的视频帧输出。

图 7-24 是普通 CCD 摄像机和 DPS 摄像机拍摄图像对比结果。可以看出 DPS 摄像机在成像清晰度、对光线适应能力等方面要优于 CCD 摄像机。



a1 CCD 摄像机拍摄图片



a2 DPS 摄像机拍摄图片

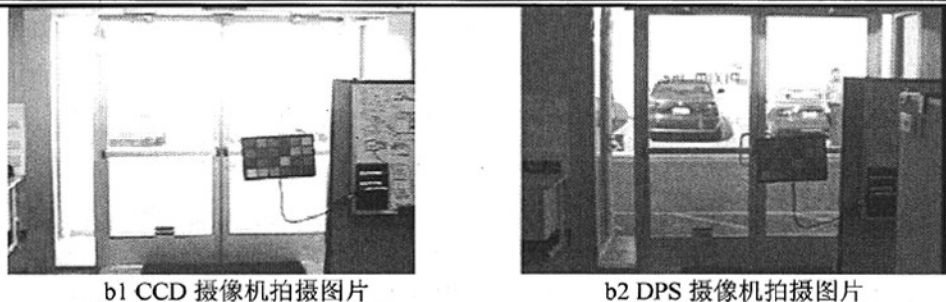


图 7-24 CCD 摄像机和 DPS 摄像机拍摄结果对比

由以上分析及实拍图像可知,对于太阳光变化较大的环境中,选用 DPS 摄像机拍摄有助于精确提取无人机辅助着陆灯标志。

3、云雾影响分析及解决措施

云雾对可见光成像的影响主要有三个方面:

一是云雾对非成像光线的散射。这一部分光在无云雾时不参与成像,有云雾时成为云雾本身的图像,使整个图像亮度增加,对比度降低。

二是云雾对成像光线的透射衰减。这部分光在有云雾和无云雾时均是有效的成像光线,其衰减强度与穿过云雾的距离呈指数规律增长,使目标图像的亮度和对比度降低。云雾较厚时目标甚至不能在传感器上成像;

三是成像光线被云雾散射后继续成像。在无云雾时,这一部分光是有效的成像光线;有云雾时,这一部分光被云雾散射,偏离原来的方向,其中的一部分不能成像,另一部分的成像偏离了本来位置,使图像模糊,对比度降低。

以上三个方面的影响最终导致目标影像亮度增加、对比度和分辨率下降。我们通过在镜头前加红外滤片方法,可大大消减云雾影响。图 7-25 为大雾情况下加滤片前后所拍摄图片。可以看出大雾天气情况下,加滤片后可大大提高图像清晰度,消减云雾的影响。

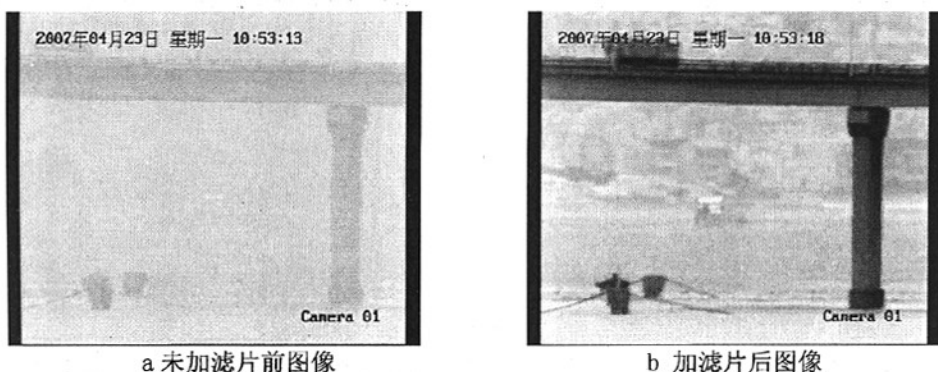


图 7-25 云雾影响消除结果对比

7.3 无人机辅助着陆系统实际测量结果分析

7.3.1 相同尺度参数的 PNNLOG 算子和高斯算子的检测结果对比

图 7-26.a 为无噪声原始仿真图像，其中包含不同尺度参数的高斯点。白框所包围的两个高斯点的参数 σ 均为 3.0，假定这两个点为真实目标，其中心点坐标分别为 (72, 89) 和 (172, 89)。分别选取 σ 均为 3.0 的 PNNLOG 算子和高斯算子进行检测，最大响应点位置用白色十字表示。图 7-26.b1 为高斯算子对加入方差为 40% 的椒盐噪声图像的检测结果；图 7-26.b2 为 PNNLOG 算子对加入方差为 40% 的椒盐噪声图像的检测结果；图 7-26.c1 为高斯算子对加入方差为 30% 的高斯噪声图像的检测结果。图 7-26.c2 为 PNNLOG 算子对加入方差为 30% 的高斯噪声图像的检测结果。实验结果表明，PNNLOG 算子可以准确地检测出两个真实目标，最大响应点坐标分别为 (72, 89) 和 (172, 89)，且对噪声有较好的抑制作用，而相同尺度的高斯算子错误地检测到最大的高斯亮点上。

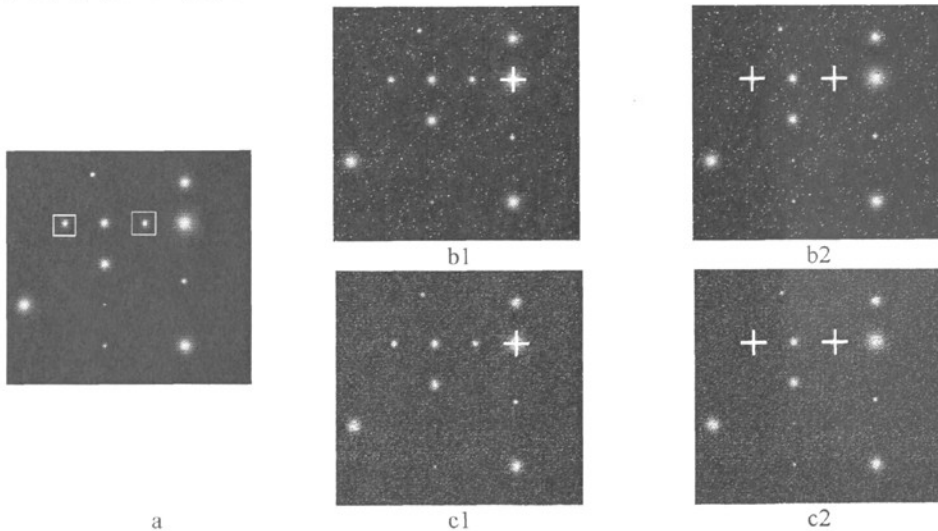


图 7-26 PNNLOG 算子对仿真图检测结果

7.3.2 正则化正负 LOG 算子对小目标检测跟踪结果

图 7-27 为无人机着陆过程中 PNNLOG 算子对无人机辅助着陆灯检测跟踪的结果，左图为无人机距离摄像机较远时，右图则为无人机着陆前。算法在无人机整个着陆过程都能很好的跟踪锁定目标。实验时，系统硬件配置为：CPU 为奔腾 2.8G，内存为 4Gb，算法在 VC++6.0 平台上开发的，处理的图像分辨率为 768×576。单帧处理平均时间为 8.5 毫秒左右，完全可以满足实时要求。同时实时保存了整个无人机着陆过程的图像，没有丢帧。

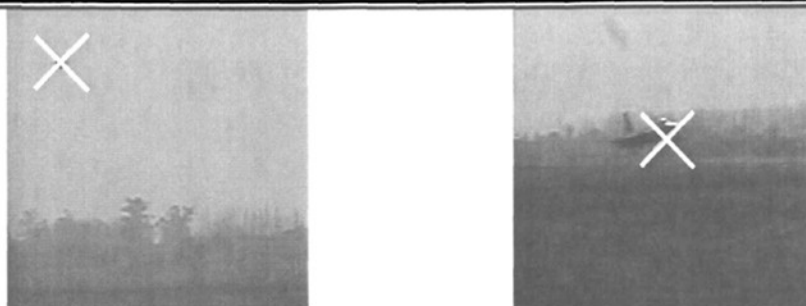
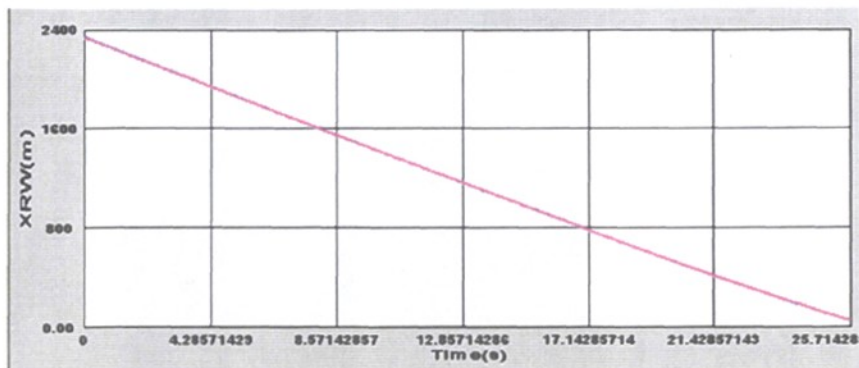


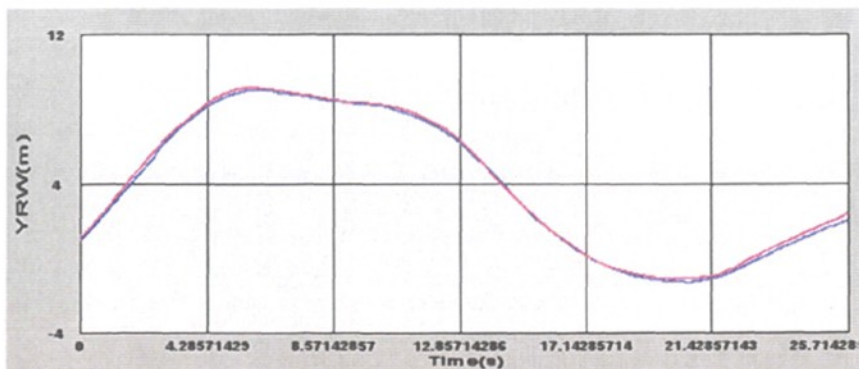
图 7-27 PNNLOG 算子对真实序列图像检测结果

7.3.3 无人机着陆过程定位测量结果

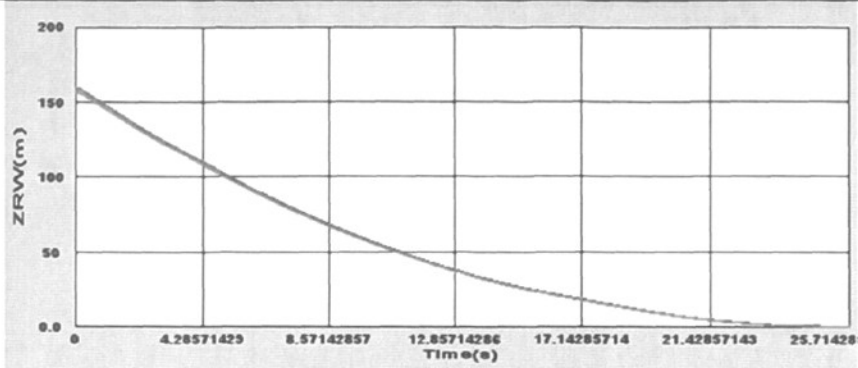
为了验证本文提出的无人机辅助着陆系统的有效性，对实际无人机进行测试，将实验测量结果与无人机上的 GPS 数据进行比对。图 7-28 为其中一组无人机辅助降落系统测量数据与 GPS 测量数据的对比结果，其中 a、b、c 分别为 X、Y、Z 方向位置参数对比结果，蓝色为本文系统的测量数据，红色为 GPS 测量数据。从图中可以看到，本文设计的系统的结果和 GPS 数据基本吻合。X 方向定位误差为 15 厘米，Y 方向定位误差为 30 厘米，Z 方向定位误差为 80 厘米，完全满足项目要求。



a X 方向



b Y 方向



c Z 方向

图 7-28 测量数据与 GPS 数据比对图

7.4 本章小结

本章针对无人机着陆系统进行研究，设计并实现了一种基于摄影测量的无人机辅助着陆导航系统。本章着重介绍了系统涉及的几个技术方面的内容，包括摄像机标定技术、小目标实时检测与跟踪技术、小目标高精度定位技术以及硬盘实时存储技术等，其中提出了一种小目标尺度自适应实时高精度检测与跟踪算法。实验证明，小目标检测新算法在实时性、检测准确性以及抗噪声性上均有良好的表现。无人机辅助着陆系统可以实时高精度地提供无人机位置信息，同时可以实时存储无人机着陆过程的图像。

第八章 结束语

8.1 本文主要研究成果及创新点

本文以无人机平台运动目标检测与跟踪以及无人机辅助着陆系统为研究背景。其中，无人机平台运动目标检测与跟踪技术主要涉及像机自运动消除、复杂背景下运动目标稳健跟踪以及目标尺度方向自适应跟踪等关键问题。而无人机辅助着陆系统主要涉及双目交会、小目标实时高精度检测跟踪和图像实时存储等核心问题。论文主要针对这些关键问题进行研究，主要研究成果归纳如下：

1. 提出并实现了一种图像角点检测方法。

针对图像角点特征提取问题，在分析 Gabor 小波的基础上，提出并实现了一种基于 Gabor 梯度的角点检测算法。新的角点检测算法比 Harris 算法、SUSAN 算法和 Forstner 算法在定位精度、噪声抑制等方面均有一定的提高。

2. 提出并实现了一种基于 Gabor 特征描述的像机自运动消除算法。

针对无人机图像背景复杂、像机视场变化大的特点，提出并实现了一种基于 Gabor 特征描述的像机自运动消除算法。该方法首先对图像自动提取有利于匹配的兴趣点，然后对每个兴趣点建立基于 Gabor 函数的特征描述，接着利用交叉匹配建立兴趣点之间的初始匹配关系，最后通过 RANSAC 算法鲁棒地估计出变换模型参数。新算法可以有效地消除像机自身运动，检测效果优于基于 SIFT 匹配的算法。

3. 提出并实现了一种目标区域差异性权值计算方法。

对于目标跟踪问题，目标特征建立的准确与否直接关系到跟踪算法性能的优劣。论文提出并实现了一种目标区域差异性权值计算方法。实验结果表明，目标区域差异性权值基本符合人类视觉反应。

4. 提出并实现了一种差异性加权最小二乘影像匹配运动目标跟踪算法。

为了适应复杂背景下目标在时间序列上发生的平移、旋转、缩放变化，提出并实现了一种差异性加权最小二乘影像匹配 (DWLSIM) 运动目标跟踪算法。通过仿真实验，比较了 LSIM 算法和 DWLSIM 算法在目标发生各种变化及噪声影响下的跟踪结果。从自然图像跟踪结果可以看出，DWLSIM 算法比 Mean Shift 算法、粒子滤波算法和 LSIM 算法更为有效稳定。

5. 提出并实现了一种差异性加权 Mean Shift 运动目标跟踪算法。

针对复杂背景下运动目标问题，将差异性权值和 Mean Shift 跟踪算法相结合，提出并实现了一种基于差异性加权 Mean Shift 运动目标跟踪算法，并对算法的收敛性进行了证明。实验结果表明，改进的 Mean Shift 算法比 Mean Shift 算法、粒子滤波算法和 LSIM 算法在目标跟踪稳定性、可靠性等方面有一定的提高。

6. 提出并实现了一种目标最佳椭圆描述区域计算方法。

对于目标跟踪问题，目标初始特征建立准确与否是整个跟踪过程中至关重要的一环。其中，准确的目标描述区域对于目标特征建立、目标跟踪、目标轮廓提取具有非常重要的意义。针对这个问题，在分析 Lindeberg 尺度空间理论的基础上，提出并实现了一种目标最佳椭圆描述区域计算方法。该目标区域最佳椭圆描述定义了描述目标区域的最佳区域，给出了目标区域的大小和方向信息，为确定目标区域特性提供基础。

7. 提出并实现了一种目标尺度方向自适应跟踪算法。

由于在跟踪过程中目标存在旋转、缩放、遮挡等复杂情况，如何对感兴趣目标进行有效准确跟踪变得更加困难。在分析 Mean Shift 跟踪算法和尺度空间理论的基础上，提出并实现了一种尺度方向自适应 Mean Shift 跟踪算法。新算法首先计算出目标的最佳椭圆描述区域，然后分别将椭圆正则化 LOG 算子和椭圆高斯算子作为核函数，通过 Mean Shift 过程实现目标尺度方向自适应跟踪。实验表明，新算法可以较好地适应目标旋转缩放变化。

8. 设计并实现了一种基于摄影测量的无人机辅助着陆系统。

无人机安全进场着陆过程中，其引导控制尤为关键。论文设计并实现了一种基于摄影测量的无人机辅助着陆引导系统。该系统可以实时给出无人机的空间三维位置信息，同时实时存储无人机着陆过程的图像，为无人机安全着陆提供有效的参考数据。针对系统涉及的小目标跟踪问题，提出并实现了一种基于 PNNLOG 算子的尺度自适应小目标实时检测跟踪算法。实验证明，新的小目标检测跟踪算法在实时性、检测准确性以及抗噪声性上均有较好的表现。

8.2 有待进一步研究的几个问题

论文对无人机平台运动目标跟踪以及无人机辅助着陆系统进行了研究，虽然已经取得了上述研究成果，但还存在一些问题，需要进一步研究，其中主要问题有：

(1) 图像角点特征提取是计算机视觉研究中非常基础的问题之一。本文提出的新算法虽然在提取准确性、抗噪声等方面比经典算法有一定提高，但在实时性上还存在不足。另外，论文仅研究了图像角点特征，而对图像中普遍存在线特征、面特征没有进行研究。

(2) 像机自运动消除是运动平台上进行运动目标检测的重要步骤之一。论文针对该问题提出了基于 Gabor 特征描述的像机自运动消除算法，新算法对大部分图像计算结果比较理想，但还需要继续完善以达到有效、准确地消除像机自运动。论文仅对像机自运动消除进行系统的研究，还需要进一步对运动目标提取进行研

究。

(3) 针对复杂背景下运动目标跟踪问题, 仅仅提出了一种单帧目标权值计算方法, 该方法虽然对彩色或灰度图像上目标计算出基本符合人类视觉的权值, 但在弱小目标权值计算上还有一定问题。论文针对复杂背景上运动目标跟踪问题提出了两种基于差异性加权的跟踪算法, 虽然两个新算法在跟踪稳定性上均比传统方法有一定的提高, 但仍需要针对不同情况提出更为有效的目标权值计算方法和新的目标跟踪算法。

(4) 本文在目标尺度方向自适应跟踪问题进行初步研究, 提出了一种尺度方向自适应 Mean Shift 跟踪算法, 该算法虽然对目标缩放、旋转变换具有一定的适应能力, 但是还不够完善, 尤其是当目标与背景差别较小时, 还存在一定的问题。

(5) 无人机辅助着陆系统虽然可以实时准确地获得飞机相对跑道的位置信息, 但是还不能获得无人机的姿态、速度和加速度等信息。这一问题是需要下一步着重努力解决的问题。

致 谢

即将为博士生生活画上句号，也即将翻开我人生新的一页，在这令人兴奋而激动的时刻，回首起往昔岁月，不禁无限感慨。最让我觉得幸福的是，在我欢乐时、忧愁时，胜利时、失败时，总有我的老师、同学、领导和亲朋好友陪伴我，关心我，在这里我衷心感谢你们！

首先，衷心感谢我的恩师于起峰教授！于老师治学严谨，学识渊博，对国防事业和科研工作始终充满激情，是令我非常钦佩的一位师长。于老师十分善于引导学生，从整体上和高层次上把握问题，每每聆听其教诲都有醍醐灌顶、茅塞顿开之感。于老师夜以继日、废寝忘食的工作作风，积极向上、永不言败的坚韧性格也都深深影响了我，将让我终生受益。

由衷感谢雷志辉副教授！雷老师是我的硕士导师，他工作细致认真、对学生严格要求，他平易近人而乐于助人。无论是在我的硕士阶段还是博士阶段，雷老师都在多方面给予了我极大的帮助。

感谢课题组的孙祥一副教授、张小虎副研究员，尚洋老师和杨夏老师等，在课题组的科研工作一直得到他们的耐心指导与无私帮助。感谢教研室的陈晓芒女士和刘明焕工程师的诸多帮助。

非常感谢陆宏伟、邱志强、伏思华、丁晓华、魏继华、冯卫东、邸慧等师兄和师姐，他们就像我的兄长一样，在我的科研上和生活上给予了我许多无私的帮助。感谢课题组的傅丹、张小苗、李立春、李由、朱宪伟、李健兵、孙俊峰、徐一丹、王鲲鹏、张政、周剑、周翔、吉芳芳、马敢干、李强和刘晓春等同学的帮助。

感谢王培云队长、蔡学武政委等学员队领导的关怀与教诲！

感谢我周围所有的亲朋好友，我的学习生活因为有你们而精彩！

谨以此文献给我的父母和珍！

2008年9月5日于长沙

参考文献

- [1] Marr D. Vision:a computational investigation into the human representation and processing of visual information[J]. Freeman W.H. and Company, San Francisco, 1982
- [2] Aloimonos J, Weiss I, Bandopadhyay A. Active vision. International Journal of Computer Vision, 1988: 333~356
- [3] Armen F, Aggarwai J K. Model-based object recognition in dense range images a review, ACM Computer Surveys, 1993, 25(1): 5~43
- [4] Pau L F. Behavioral knowledge in sensor/data fusion system. Robotic System, 1990, 7(3): 295~372
- [5] Pal N R, Pal S K. A review on image segmentation techniques. Pattern Recognition, 1993, 26(9): 1227~1294
- [6] Jain A K, Zhong Y, Lakshmanan S. Object Matching using deformable templates. IEEE Trans. On Pattern Analysis and Machine Intelligence, 1996, 18(3): 267~278
- [7] Starovoitov V V, Kose C, Sankur B. Generalized distance based matching of nonbinary images, IEEE International conference on Image Processing, 1998: 803~907
- [8] Barron J L, Fleet D J, Beauchemin S S. Systems and experiment: performance of optical flow techniques. International Journal of Computer Vision, 1994, 12(1): 43~77
- [9] Lee K W, Ryu S W, Lee S J, Park K T. Motion based object tracking with mobile camera. Electronics Letters, 1998, 34 (3): 256~258
- [10] Song S W, Liao M Y, Qin J M. Multiresolution image motion detection and displacement estimation. Machine Vision and Applications, 1990, 3(1): 17~20
- [11] Erdem C E, Karabulut G Z, Yanmaz E, Anarim E. Motion estimation in the frequency domain using fuzzy c-planes clustering. IEEE Trans. On Image Processing, 2001, 10(12): 1873~1879
- [12] Rosenfeld A. From image analysis to computer vision: An annotated bibliography [J]. Computer Vision and Image Understanding, 2001, 84(2): 298~324
- [13] Horn B K, Shunck B G. Determine Optical Flow[J]. Artificial Intelligence, 1981, 17(1-3): 185~204
- [14] Hou Z Q, Han C Z. A background reconstruction algorithm based on pixel intensity classification[J]. Chinese Journal of Software, 2005, 16(9): 1568~1576
- [15] Wang S, Ai H.Z. Difference image based multiple motion targets detection and

- tracking[J]. Chinese Journal of Software, 1999, 4(6): 470~475
- [16] Pavlidis I, Morellas V, Tsiamyrtzis P, Harp S. Urban surveillance system: From the laboratory to the commercial world[J]. Proceedings of the IEEE, 2001, 89(10): 1478~1496
- [17] Collins R, Lipton A, Fujiyoshi H, Kanade T. Algorithms for cooperative multisensor surveillance[J]. Proceedings of the IEEE, 2001, 89(10): 1456~1477
- [18] Masoud O, Papanikolopoulos N P. A novel method for tracking and counting pedestrians in real-time using a single camera[J]. IEEE Trans on Vehicular Technology, 2001, 50(5): 1267~1278
- [19] Bar-shalom Y, Fortmann T. Tracking and data association[M]. Academic Press, 1988
- [20] Rabiner L. A tutorial on hidden Markov models and selected applications in speech recognition[J]. Proceedings of IEEE, 1989, 77(2): 257~285
- [21] Salmond D, Gordon N, Smith A. A novel approach to nonlinear/non-Gaussian Bayesian state estimation[J]. IEEE Proceedings on Radar, Sonar and Navigation, 1993, 140(2): 107~113
- [22] Arulampalam M, Maskell S, Gordon N, Clapp T. A tutorial on particle filters for online/non-Gaussian Bayesian tracking[J]. IEEE Trans on Signal Processing, 2002, 50(2): 174~188
- [23] Isard M, Blake A. CONDENSATION-conditional density propagation for visual tracking[J]. International Journal of Computer Vision, 1998, 29(1) : 5~28
- [24] Meyer F, Boutheymy P. Region-based tracking using affine motion models in long time sequences[C]. CVGIP Image understanding, 1994, 60(2): 119~140
- [25] Bascle B, Deriche R. Region tracking through image sequences[J]. Proceedings of IEEE International Conference of Computer Vision, 1995: 302~307
- [26] Jepson A, Fleet D, El-Maraghi T. Robust online appearance models for visual tracking[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2003, 25(10): 1296~1311
- [27] Nguyen H, Smeulders A. Fast occluded object tracking by a robust appearance filter[J]. IEEE Trans. On Pattern Analysis and Machine Intelligence. 2004, 26(8): 1099~1104
- [28] Kaneko T, Hori O. Feature selection for reliable tracking using template matching[C]. Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003, 1: 796~802
- [29] Mitra P, Murthy C, Pal S. Unsupervised feature selection using feature

- similarity[C]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 2002, 24(3): 301~312
- [30] Koller D, Daniilidis K, Nagel H. Model-based object tracking in monocular image sequences of road traffic scene[J]. International Journal of Computer Vision, 1993, 10(3): 257~281
- [31] Smith S, Brady J. SUSAN- A new approach to low level image processing[J]. International Journal of Computer Vision, 1997, 23(1): 45~78
- [32] Comaniciu D, Ramesh V, Meer P. Real-time tracking of non-rigid objects using mean Shift[J]. IEEE Computer Vision and Pattern Recognition, 2000: 142~149
- [33] YANG C J, et al. Similarity measure for nonparametric kernel density based object tracking[C]. Eighteenth Annual Conference on Neural Information Processing Systems. Victoria, British Columbia, Canada: NIPS, 2004.
- [34] Collins R T. Mean-shift blob tracking through scale space[C]. 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003, 2: 234~240
- [35] Nummiaro K, Koller - Meier E, Van Gool L. An adaptive color-based particle filter[J]. Image and Vision Computing, 2003, 21(1): 99~110
- [36] Jurie F. Tracking objects with a recognition algorithm[J]. Pattern Recognition Letters, 1998, 19(3-4): 331~340
- [37] Drummond T, Cipolla R. Real-time visual tracking of complex structures[J]. IEEE Trans. On Pattern Analysis and Machine Intelligence, 2002, 24(7): 932~946
- [38] Gerard P, Gagalowicz A. Three dimensional model-based tracking using texture learning and matching[J]. Pattern Recognition Letters, 2000, 21(13-14): 1095~1102
- [39] Jung S K, Wahn K Y. A model-based 3-D tracking of rigid objects from a sequence of multiple perspective views[J]. Pattern Recognition Letters, 1998, 19(5-6): 499~512
- [40] Kass M, Witkin A, Terzopoulos. Snakes: Active contour models[J]. International Journal of Computer Vision, 1988, 1(4): 321~331
- [41] Cohen L D. On active contour models and balloons[C]. CVGIP:Image Understanding, 1991, 53(2): 211~218
- [42] Xu C, Prince J L. Snakes, shapes and gradient vector flow[C]. IEEE Transactions on Image Processing, 1998, 7(3): 359~369
- [43] Szeliski R, Terzopoulos D. Physically-Based and probabilistic modeling for computer vision[C]. In: Vemuri B C, Proceedings SPIE 1570, Geometric Methods in

- Computer Vision, 1991: 140~152
- [44] Vieren C, Cabestaing F, Postaire J. Catching moving objects with snakes for motion tracking[J]. Pattern Recognition Letters, 1995, 16(7): 679~685
- [45] Osher S, Sethian J A. Fronts propagating with curvature dependent speed: algorithms based on Hamilton-Jacobi formulation[J]. Journal of Computational Physics, 1998, 79(1): 12~49
- [46] Malladi R, Sethian J A, Vemuri B C. Shape modeling with front propagation: A level set approach[J]. IEEE Trans. On Pattern Analysis and Machine Intelligence, 1995, 17(2): 158~179
- [47] Kitchen L, Rosenfeld A. Gray level corner detection[J]. Pattern Recognition Letters, 1982, 3(1): 95~102
- [48] 田原, 梁德群, 吴更石. 直接基于灰度图像的多尺度角点检测方法[J]. 信号处理, 1998, 14(7): 6~9
- [49] Witkin A. Scale-space filtering[J]. In IJCAI, 1983, 1019~1023
- [50] Kitchen L, Rosenfeld A. Analysis of gray level corner detection[J]. Pattern Recognition Letters, 1999, 20(2): 149~162
- [51] Harris C, Stephens M. A Combined corner and edge detector[J]. Proc 4th Alvey Vision Conference, 1988: 189~192
- [52] Tomasi C, Kande T. Detection and tracking of point features[R]. Carnegie Mellon University, 1991
- [53] Smith S M, Brady J M. SUSAN-A new approach to low level image processing[R]. Internal Technical Report TR95SM S1, Defence Research Agency, Chobham Lane, Chertsey, Surry, UK, 1995
- [54] 杨莉, 初秀琴, 李玉山. 最小亮度变化角点自适应检测算法研究[J]. 西安电子科技大学学报, 2003, 30(4): 530~533
- [55] 费旭东, 荆仁杰. 基于知识的快速角点提取[J]. 计算机学报, 1994, 17(1): 30~36
- [56] 陈燕新, 戚飞虎. 一种新的提取轮廓特征点的方法[J]. 红外与毫米波学报, 1998, 17(3): 171~176
- [57] 王栓 等. 基于差分图像的多运动目标的检测与跟踪[J]. 中国图像图形学报, 1999, 4(6): 470~475
- [58] Qu Y S, Tian W J, Li Y C. Detecting small moving target in images sequences using optical flow based on the discontinuous frame difference[C]. Proceedings of the Third SPIE Multi-spectra Image Processing and Pattern Recognition Conference, 2003, 52(6): 915~918

- [59] Kim C , Hwang J N. Fast and automatic video object segmentation and tracking for content- based applications[J] . IEEE Trans on Circuits and Systems, 2002, 12 (2) : 122~129
- [60] Duda R O, Hart P E. Pattern classification and Scene analysis[M]. Wiley, New-York, 1973
- [61] Rosin P L, Ellis T. Image difference threshold strategies and shadow detection[C]. 6th British Machine Vision Conf., Birmingham, 1995: 347~356
- [62] Jain R, Nagel H. On the analysis of accumulative difference of picture from image sequences of real world scenes[J]. IEEE Trans. PAMI, 1979: 206~214
- [63] 陈朝阳,张桂林. 基于图像对称差分运算的运动小目标检测方法[J]. 华中理工大学学报, 1998, 26(9): 34~38
- [64] Lipton A, Fujiiyoshi H and Patil R. Moving target classification and tracking from real time video[C]. Proc of the 1998 DARPA Image Understanding Workshop, 1998
- [65] Changick K. Fast and automatic video object segmentation and tracking for content based application[J]. IEEE Transaction on Circuits and System for Video Technology, 2002, 12(2): 122~129
- [66] 杨志华, 曾禹村. 背景移动补偿技术的研究[J]. 北京理工大学学报, 2000, 20(3) : 333~337
- [67] Araki S, Matsuoaka T, Yokoya N, Takemura H. Real-time tracking of multiple moving object contours in a moving camera image sequence[J]. IEICE Trans. Inf.&Syst., 2000, (7): 131~137
- [68] Collins R T, Lipton A J, Kanade T. Special section on vidieo surveillance[J]. IEEE Trans On Pattern Analysis and Machine Intelligence, 2000, 22(8): 745~746
- [69] Viola P, Jones M J, Snow D. Detecting pedestrians using patterns of motion and appearancep[C]. International Conference on Computer Vision, 2003: 1234~1245
- [70] Eun J H, Robyn O. Segmenting occluded objects using a motion snake[C]. Asian Conference on Computer Vision, 2004
- [71] Alexei A E, Alexander C B, Greg M, Jitendra M. Recognizing Action at a distance[C]. International Conference on Computer Vision, 2003: 560~566
- [72] Mae Y, Shirai Y, Miura J, et al. Object tracking in cluttered background based on optical flow and edges[C]. Proceedings of the 13th International Conference on Pattern Recognition, 1996, 1: 196~200
- [73] Mae Y, Yamamoto S, Shirai Y, Miura Jun. Optical flow based realtime object

- tracking by active vision system[C]. Proc. 2nd Japan-France Congress on Mechatronics, 1994, 2: 545~548
- [74] Okada R, Shirai Y, Miura J. Object tracking based on optical flow and depth[C]. Proc. of RWC, 1997: 432~439
- [75] Lowe D G. Object recognition from scale-invariant features[C]. International Conference of Computer Vision, 1999: 1150~1157
- [76] 魏波. 点时空约束图像目标跟踪理论与实时实现技术研究[D]. 西安: 电子科技大学, 2000
- [77] Rosenberg Y, Werman M. Representing local motion as a probability distribution matrix and object tracking[C]. Image Understanding Workshop, 1997: 153~158
- [78] Grimson, Stauffer C, Romano C and Lee L. Using adaptive tracking to classify and monitor activities in a site[C]. Proc. IEEE conference on Computer Vision and Pattern Recognition, 1998: 758~764
- [79] Ahmed E, David H, Larry D. Non-parametric Model for background subtraction[C]. European Conference on Computer Vision, 2000: 654~659
- [80] Henry S and Takeo K. Object detection using the statistics of parts[J]. International Journal of Computer Vision, 2004, 56(3): 151~177
- [81] 谢敏. 视频序列图像中运动目标的分割和识别研究[D]. 苏州大学硕士论文, 2002
- [82] Mallat S. A theory for multiresolution signal decomposition: the wavelet representation[J]. IEEE Trans on PAMI, 1989, 11(7): 674~693
- [83] Yu X, Reed I S, Kraske W. A robust adaptive multispectral object detection by using wavelet transform[C]. Proc. ICASSP-92, 1992: 141~144
- [84] Casasent D P, Smokelin J S, Ye A. Wavelet and Gabor transforms for detection[J]. Opt. Eng., 1992, 9, 31(4): 1893~1898
- [85] 李红艳. 图像低信噪比小目标检测与跟踪算法研究[D]. 西安电子科技大学博士论文, 2000
- [86] Mohanty N C. Computer tracking of moving point targets in space[J]. IEEE Trans. on PAMI, 1981, 3(5): 606~611.
- [87] Pohlig S C. An algorithm for detection of moving optical targets[J]. IEEE Trans on AES, 1989, 1, 25(1): 56~63
- [88] Blostein S D, Huang T S. Detecting small moving objects in image sequences using sequential hypothesis testing[J]. IEEE Trans. on SP, 1991, 39(7): 1611~1629.
- [89] Blostein S D, Richardson H S. A sequential detection approach to target tracking[J].

- IEEE Trans.on AES, 1994, 30(1): 197~211
- [90] Barniv Y. Dynamic programming solution for detecting dim moving targets[J]. IEEE Trans. on AES, 1985, 21(1): 144~156
- [91] Wang G, Inigo R M and Mcvey E S. A pipeline algorithm for detection and tracking of pixel-sized target trajectories[J]. Single and Data Processing of Small Targets 1990, Proceedings of the SPIE, 1990, 1305: 53~62
- [92] Ren L J, Mahmood R, Dim Target Detection and Clutter Rejection using modified high order correlation[C]. IEEE International joint conference on Neural Network, 1993, 29(3): 841~856
- [93] Nanbara J H, Mahmood R. Target detection utilizing neural networks and modified high order correlation method[C]. SPIE, 2496(1995): 687~697
- [94] Ren J L, Mahmood R. Multiple target detection using modified high order correlation[C]. IEEE Trans on Aerospace and Electronic Systems, 1993, 34(2): 553~568
- [95] Liou R, Azimi_sadjadi. Dim target detection using high order correlation method[J]. IEEE Trans on AES. 1993, 29(3): 841~856
- [96] Bolstein S D, Huang T S. Detecting small, moving objects in image sequences using sequential hypothesis testing[J]. IEEE Trans. On SP-39, 1991, 7: 1611~1629
- [97] Blostein S D, Richardson H S. A sequential detection approach to target tracking[J]. IEEE Trans on AES. 1994, 30(1):197~212
- [98] 李红艳, 吴成柯. 一种基于小波与遗传算法的小目标检测算法[J]. 电子学报. 2001, 21(4): 81~83
- [99] 李红艳, 吴成柯. 一种基于小波变换的序列图像中小目标检测与跟踪算法[J]. 电子与信息学报. 2001, 23(10): 943~948
- [100] Barniv Y. Dynamic programming solution for detecting dim moving targets[J]. Part II, Analysis, IEEE Trans On AES-23, 1987, 6: 776~788
- [101] Arnold J, Shaw S, Pasternack H. Efficient target tracking using dynamic programming[J]. IEEE Trans on AES. 1993, 29: 44~56
- [102] Tonissen S M, Evans R J. Target tracking using dynamic programming algorithm and performance[C]. Proceedings of IEEE Conferece on Decision and Control. 1995
- [103] Tonissen S M, Evans R J. Performance of dynamic programming track before detect algorithm[J]. IEEE Trans on AES. 1996, 32(4): 1440~1451
- [104] Johnston L A, Krishnamuthy V. Performance of a dynamic programming track

- before detect algorithm[J]. IEEE Trans on AES. 2002, 38(1): 228~242
- [105] 李斌, 彭嘉雄. 基于动态规划的红外小目标检测与识别[J]. 华中理工大学学报. 2000, 28(6): 68~70
- [106] Qiang Yong, Jiao Licheng, Bao zheng. Study on mechanism of dynamic programming algorithm for dim targets[C]. ICSP 2002 Proceedings. 2002: 1403~1406
- [107] 陈华明, 孙广富, 卢焕章等. 基于动态规划和置信度检验的小目标检测[J]. 系统工程与电子技术. 2003, (4): 472~476
- [108] 陈尚峰, 陈华明, 卢焕章. 基于加权动态规划和航迹关联的小目标检测技术[J]. 国防科技大学学报. 2003, 25(2): 46~50
- [109] 钟圣芳, 张兵, 卢焕章. 一种基于动态规划的点目标轨迹关联算法[J]. 计算机测量与控制. 2004, 12(8): 772~775
- [110] 黄林梅, 张桂林, 王新余. 基于动态规划的红外运动小目标检测算法[J]. 红外与激光工程. 2004, 33(3)
- [111] Reed I S, Gagliardi R M, Shao H M. Application of three-dimensional filtering to moving target detection[J]. IEEE Trans On AES-19, 1983, 6: 898~905
- [112] Stocker A. D, Jensen P. Algorithm and architectures for implementing large velocity filter banks[C]. SPIE, 1991, 1481: 140~155
- [113] Chen Y. On suboptimal detection of 3-dimensional moving targets[J]. IEEE Trans on AES. 1989, 25(3): 343~350
- [114] Porat B, Friedlander. A frequency domain algorithm for multiframe detection and estimation of dim targets[J]. IEEE Trans on Pattern Analysis and Machine Intelligence. 1990: 398~401
- [115] Reed I S, Gagliardi R M., Stotts L. A recursive moving target indication algorithm for optical image sequence[J]. IEEE Trans on AES. 1990, 26(3): 434~440
- [116] Xiong Y, Ding M Y, Peng X. An extended track before detect algorithm for infrared target detection[J]. IEEE Trans on AES. 1997, 33(3): 1087~1092
- [117] Jauffret C, Bar-Shalom Y. Track formation with bearing and frequency measurements in clutter[J]. IEEE Trans on AES. 1990, 26: 999~1010
- [118] 宋柳平, 孙仲康, 沈振康. 红外制导系统中的点目标检测技术[J]. 电子学报. 1994, 22(1): 78~81
- [119] Qian Z B, Liu X H, Li J G. A real-time system for detecting moving point target in low SNR image sequence[C]. IEEE International Conference on Intelligent Processing Systems, 1997, 1005~1008

- [120] 廖斌, 杨卫平, 沈振康. 低信噪比线性运动红外小目标检测方法[J]. 红外技术. 2001, 23(5): 11~13
- [121] 艾斯卡尔, 李在铭. 最优分布变换与微弱点状目标检测技术[J]. 系统工程与电子技术. 2003, 25(1): 103~106
- [122] Gordon N J, Salmond J. Smith A F M. Novel approach to nonlinear/non Gaussian Bayesian state estimation[J]. IEEE Proceedings on Radar, Sonar and Navigation. 1993, 140(2): 107~113
- [123] Doucet A, Freitas, Gordon N J. Sequential Monte Carlo methods in practice[M]. New York: Springer-Verlag. 2001
- [124] Samond D J, Brich H. A particle filter for track-before-detect[A]. IEEE Proceedings of the American Control Conference. Washington. 2001: 3755~3760
- [125] Ristic B. Detection and tracking of stealthy targets, beyond the kalman filter: particle filter for tracking application[M]. Artech House. 2004
- [126] Rutten M G, Gordon N J. Maskell S. Efficient particle based track before detect in Rayleigh Noise[A]. In Fusion 2004: Proceedings of the 7th International Conference on Information Fusion. Sweden. June. 2004
- [127] 杨小军, 潘泉, 张洪才. 基于粒子滤波和似然比的联合检测与跟踪[J]. 控制与决策. 2005, 20(7).
- [128] Seera J. Image analysis and mathematical morphology[M]. NewYork : Academic Press, 1982.
- [129] Wang Y, Zhen Q B, Zhang J P. Realtime detection of small target in IR grey image based on mathematical morphology[J]. Infrared and Laser Engineering, 2003, 32(1): 28~31
- [130] Yan Z, Guilin Z. A morphology based on detection of small moving target in infrared image sequence[J]. Journal of Data Acquisition & Processing, 1997, 12(4): 276~279
- [131] Cassasent D. Sub-pixel target detection and tracking[C]. Proc of the SPIE, 1986, 726: 1206~1210
- [132] Rauch H E, Fetterman W I, Kemmer D B. Background suppression and tracking with a staring Mosaic sensor[J]. Optical Engineering, 1981, 20(1): 128~133
- [133] Horn B K, Schunck B G. The image flow constraint equation[J]. CVGIP, 1986, 35(3): 26~31
- [134] Horn B K, Schunck B G. Determining optical flow[R]. MIT : AIMemoAI Lab MIT, 1986: 88~933

- [135] Uras S , Girosi F , Verri A. A computational approach to motion perception[J]. Biol Cybern, 1988, 60: 79~87
- [136] Reichardt W E , SchlËgl R W. A two dimensional field theory for motion computation[J]. Biol Cybern, 1988, 60: 23~25
- [137] Markandey V. Motion estimation for moving target detection[J]. IEEE Trans on AES, 1996, 32(3): 866~874
- [138] 李晓春,孙陪懋,金国藩,等. 复杂场景中动目标图像分割的一种快速算法[J]. 清华大学学报(自然科学版), 1997, 37(8): 59~64
- [139] Verri A , Poggio T. Motion field and optical flow:qualitative properties[J]. IEEE Trans PAMI , 1989, 11 (5): 477~489
- [140] Li G K, Peng J X, Li H. Small target detection based on multiwavelet transform[J]. J Huazhong Univ of Sci &Tech, 2000, 28(1): 72~75.
- [141] Zuo Z, Zhang T X. Algorithm based on wavelet transform for small target detection on sea surface in infrared image [J]. Infrared and Laser Engineering, 1998, 27(3) : 42~45
- [142] Sheng W , Deng b ,Liu J. Multi2resolution distance map based small target detection in infrared image [J] . Acta Electronica Sinica, 2002, 30 (1): 42~45
- [143] Li H , Zheng C Y, Gao J L. Detecting small moving object based on wavelet multi-scale image registration[J]. Infrared and Laser Engineering, 2003, 32(3): 268~270
- [144] Giuseppe B. Small target detection using wavelets[A]. IEEE 1998 Fourteenth International Conference on Pattern Recognition, 1998: 1776~1778
- [145] Hongyan L, Chengke W. Detecting dim small targets in image sequences based on wavelet transforms and genetic algorithms[J]. Acta Electronica Sinica, 2001, 29(4): 439~442
- [146] Mahmoud S A. Motion analysis of multiple moving object using Hartley transform[J]. IEEE Trans on Systems , Man and Cybernetics, 1991, 21: 280~287
- [147] Mahmoud S A. Motion detection and estimation of multiple moving object in an image sequence using cosine area transform(CAT) [J] . IEEE Proc, 1991, 138 (5): 351~356
- [148] Falconer D G. Target tracking with the Hough transform[A]. Proceedings of 11th Asilomar Conference on Circuits , System, and Computer, 1977: 249~252
- [149] Snyder W E , Rajola S A. Tracking acquisition of sub-pixel target[J]. BATI ASUS Series, 1983: 227~233
- [150] Yongqiang D, Ran T, Siyong Z, et al. SAR moving target detection and imaging

- based on WHH Transform[J]. Journal of Beijing Institute of Technology, 1999, 8(1): 94~100
- [151] 张良云. 惯性导航系统[M]. 北京: 国防工业出版社, 1992
- [152] 周忠谟, 易杰军. GPS 卫星测量原理与应用[M]. 北京: 测绘出版社, 1992
- [153] 董绪荣, 张守信, 华仲春. GPS/INS 组合导航定位及其应用[M]. 长沙: 国防科技大学出版社, 1998
- [154] Courtney S. Sharp, Omid Shakenia, S.Shankar Sastry. A Vision System for Landing an Unmanned Aerial Vehicle[C]. Proceedings of the 2001 IEEE International Conference on Robotics and Automation, 2001: 1720~1727
- [155] Srikanth S, James F M and Gaurav S S. Vision-based Autonomous landing of an Unmanned Aerial Vehicle[C]. Proceedings of the 2002 IEEE International Conference on Robotics & Automation, 2002: 2799~2804
- [156] Srikanth S, James F M, Gaurav S S. Visually Guided Landing of an Unmanned Aerial Vehicle[J]. IEEE Transactions on Robotics and Automation, 2003, 19(3): 371~380
- [157] 邱力为, 宋子善, 沈为群. 用于无人直升机着舰控制的计算机视觉技术研究[J], 航空学报, 2003, 24(4): 351~354
- [158] Marr D C and Hildreth E C. Theory of edge detection[J]. Proc Roy Soc, London, 1980, 207: 187~217
- [159] Lindeberg T. Junction detection with automatic selection of detection scales and localization scales[C]. Proceedings of IEEE International Conference on Image Processing, 1994: 924~928
- [160] Gabor D. Theory of communication[J]. Journal of the Institute of Electrical Engineers, 1946, 93(26): 429~457
- [161] Marcelja S. Mathematical description of simple cortical cells [J]. J. Opt. Soc. Am A, 1980, 70: 1297~1300
- [162] Daugman J. Uncertainty relation for resolution in space, spatial frequency and orientation Optimized by two-dimensional visual cortical filters[J]. J. Opt. Soc. Am A, 1985, 2: 1160~1168
- [163] Pollen D A, Ronner S F. Visual cortical neurons as localized spatial frequency filters[J]. IEEE Transaction in Systems, Man and Cybernetics, 1983, 13(5): 907~916
- [164] Daugman J G. Uncertainty Relation for Resolution in Space, Spatial Frequency, and Orientation Optimized by Two-Dimensional Visual Cortical Filters[J].

- Journal of the Optical Society of America A, 1985, 2(7): 1160~1169
- [165] Jones J P, Palmer L A. An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex[J]. Journal of Neurophysiology, 1987, 58 (6): 1233~1258
- [166] Fukunaga K and Hostetler L D. The estimation of the gradient of a density function with application in pattern recognition[J], IEEE Trans. Inform. Theory, 1975, 21: 32~40
- [167] Cheng Y. Mean Shift, mode seeking, and clustering[J]. IEEE Trans. Pattern Anal. 1995, 17(8): 790~799
- [168] Comaniciu D and Meer P. Mean Shift: A robust approach toward feature space analysis[J]. IEEE Trans. Pattern Anal. 2002, 24(5): 603~619
- [169] Maneesh K. Mean shift segmentation with wavelet-based bandwidth selection[C]. Sixth IEEE Workshop on Applications of Computer Vision, 2002 : 43~51
- [170] Comaniciu D and Meer P. The variable bandwidth Mean Shift and data-driven scale selection[C]. 8th International Conference on Computer Vision, Vancouver, BC, Canada, July 2001, 1: 438~445
- [171] Collins R T. Mean-Shift blob tracking through scale space[C]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003 : 1063~1069
- [172] Georgescu B, Shimshoni I and Meer P. Mean shift based clustering in high dimensions: A texture classification example[C]. In Proc. Int. Conf. Computer Vision, 2003: 456~463
- [173] Yang C J, Duraiswami R, DeMenthon D and Davis L. Mean Shift Analysis Using Quasi-Newton Methods[C]. In IEEE International Conference on Image Processing, 2003, 3: 447~450
- [174] Yang C J, Duraiswami R, Gumerov N and Davis L. Improved Fast Gauss Transform and Efficient Kernel Density Estimation[C]. In IEEE International Conference on Computer Vision, 2003: 464~471
- [175] Parzen E. On estimation of a probability density function and mode[J]. Ann. Math. Stat., 1962, 33(3): 1065~1076
- [176] Benjamin J R. The Detection of 2D Image Feature Using Local Energy[D]. Australia : The University of Western Australia, 1996
- [177] Moravec H P. Towards automatic visual obstacle avoidance[C]. In 5th International Joint Conference on Artificial Intelligence, 1977: 584~590
- [178] Forstner W, Gulch E. A fast operator for detection and precise location of distinct

- points[C]. Corners and Centres of Circular Feature. Intercommision Conference on Fast Processing of Photogrammetric Data, Interlaken, Switzerland, 1987
- [179] Mehrotra R, Namuduri K R, Ranganathan N. Odd Gabor filter-based edge detection[J]. Pattern Recognition, 1992, 25(12): 1479~1494
- [180] LEE Tai Sing. Image representation using 2D Gabor wavelet[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 1996, 18(10): 959~971
- [181] 陈泽志, 吴成柯. 一种高精度估计的基础矩阵的线性算法[J]. 软件学报, 2002, 13(14)
- [182] Peter Sturm and Bill Triggs. A Factorization Based Algorithm for Mult-Image Projective Structure and Motion [C]. 4th European Conference on Computer Vision, Cambridge, 1996: 709~720
- [183] L. S. Shapiro, A. Zisserman, and M. Brady. 3D Motion Recovery via Affine Epipolar Geometry [J]. International Journal of Computer Vision, 1995, 16(2): 147~182
- [184] Paulo R. S. Mendonc and Roberto Cipolla. Estimation of epipolar geometry from apparent contours: Affine and circular motion cases [C]. In Proceeding Conf. Computer Vision and Pattern Recognition, 1999: 9~14
- [185] Ilan Shimshoni, Ronen Basri, Ehud Rivlin. A Geometric Interpretation of Weak-Perspective Motion [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1999, 21(3): 252~257
- [186] G L Brown. A Survey of Image Registration Techniques[J]. ACM Computing Survey, 1992, 24(4): 325~376
- [187] Barbara Zitova, Jan Flusser. Image registration methods : A survey[J]. Image and Vision Computing, 2003, 24(11): 977~1000
- [188] Mikolajczyk K, Schmid C. Scale and affine invariant interest point detectors[C]. International Journal of Computer Vision, 2004, 60(1): 63~86
- [189] Harris G, Stephens M. A Comblined corner and edge detector[J], 1988
- [190] 黄世奇, 禹春来. 成像精确制导技术分析与研究[J]. 导弹与航天运载技术, 2005, 278: 20~24
- [191] Yilmaz A, Shafique K, Shah M. Target tracking in airborne forward looking infrared imagery[J]. Image and Vision Computing, 2003, (21): 623~635
- [192] Haritaoglu I, Harwood D, Davis L. W4: Real-time surveillance of people and their activities[J]. IEEE Trans Patern Analysis and Machine Intelligence, 2000, 22 (8) : 809~830

- [193] Tsuji T, Hattori H, Watanabe M, et al. Development of night vision system[J]. IEEE Transactions on Intelligent Transportation Systems, 2002, 3 (3): 203~209
- [194] Yang Dakai, Cai Baigen, et al. An improved map-matching algorithm used in vehicle navigation system[C]. Proc IEEE Intelligent Transportation Systems, 2003, 2: 1246~1250
- [195] Aggarwal J K, Cai Q. Human motion analysis: A Review[J]. Computer Vision and Image Understanding, 1999, 73 (3); 428~432
- [196] 赵骅,吕科. 中国虚拟人数据的配准算法[J].计算机工程与应用, 2004, (34) : 64~67
- [197] 孙即祥等. 现代模式识别[M]. 长沙: 国防科技大学出版社, 2002: 31~34
- [198] Kenneth R Castleman. Digital image processing[M]. 北京: 电子工业出版社, 2001 : 408~410
- [199] 贾云得.机器视觉[M]. 北京: 科学出版社, 2000: 36~37
- [200] Zhang Z X, Zhang J Q, Wu X L. Global image matching with relaxation method[C]. Proceedings of the Intercommission Colloquium on Photogrammetry, Remote Sensing and Geographic Information System. Wuhan: WTUSM Press, 1992: 11~14
- [201] Lai S H , Fang M. Robust and efficient image alignment with spatially varying illumination models[C]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition.Fort Collins, Colorado:Colorado State University Press, 1999, 2: 167~172
- [202] 李峰, 周源华. 变形系数相关的最小二乘匹配算法[J]. 上海交通大学学报, 1999, 23(11) : 1391~1394
- [203] Hartley R, Zisserman A. Multiple view geometry in computer vision[M]. Cambridge University Press 2000
- [204] Collins R T. Mean-shift blob tracking through scale space[C]. In proceedings of the conference on computer vision and pattern recognition, 2003: 1069~1076
- [205] Rosenfeld A K. Digital picture processing[M]. New York: American Press, 1982
- [206] Kenneth R C. Digital Image Processing[M]. Prentice-Hall Press, 1996

附录 A 摄像机标定

标定原理：在标定过程中必须已知空间若干（三个以上）参考点在某参考坐标系下的坐标，已知像机光心在该参考坐标系下的坐标。根据摄影量学中的共线方程：

$$\begin{cases} \frac{X}{f} = \frac{r_{11}x_w + r_{12}y_w + r_{13}z_w + t_x}{r_{31}x_w + r_{32}y_w + r_{33}z_w + t_z} \\ \frac{Y}{f} = \frac{r_{21}x_w + r_{22}y_w + r_{23}z_w + t_y}{r_{31}x_w + r_{32}y_w + r_{33}z_w + t_z} \end{cases} \quad (11.1)$$

其中 f 为像机焦距， (X, Y) 为像点坐标， (t_x, t_y, t_z) 为光心偏移量， (r_{ij}) 为旋转矩阵， (x_w, y_w, z_w) 为物点坐标。已知空间若干参考标志点坐标及相应的图像点坐标，即可标定像机的焦距和姿态角。

中心透视投影成像关系

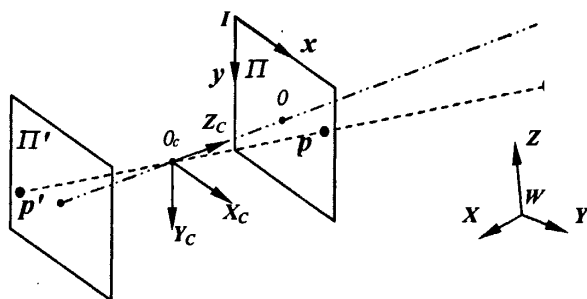


图 11-1 中心透视投影成像模型

图 11-1 中 O_c 为像机光心，双点划线为光轴， Π' 平面为反片位置， Π 平面为正片位置。 O 为图像主点， P 为物点， p 为 P 的像点，虚线为 P 点成像的光线。

像点 p 的图像坐标 (\tilde{x}, \tilde{y}) 与物点 P 的参考坐标系坐标 (X, Y, Z) 的关系可用齐次坐标表达为

$$Z_c \begin{bmatrix} \tilde{x} \\ \tilde{y} \\ 1 \end{bmatrix} = \begin{bmatrix} F_x & 0 & C_x & 0 \\ 0 & F_y & C_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{R} & \mathbf{T} \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (11.2)$$

其中 (C_x, C_y) 为图像主点； (F_x, F_y) 为横、纵向等效焦距， (X, Y, Z) 为物点 P 在参考坐标系 $W-XYZ$ 中的坐标， $\mathbf{T} = (T_x, T_y, T_z)^T$ 为平移向量， \mathbf{R} 为旋转矩阵，它的

元素 $r_0 \sim r_8$ 是旋转角 (A_x, A_y, A_z) 的三角函数组合, (\tilde{x}, \tilde{y}) 为像点 p 的图像坐标。

共面点的中心透视投影成像关系

将参考坐标系的 $O-XY$ 平面取为这些点所共的平面, 即对于这些共面点, $Z \equiv 0$, 因而可将中心透视投影关系式(1)简化为

$$\lambda \begin{bmatrix} \tilde{x} \\ \tilde{y} \\ 1 \end{bmatrix} = \begin{bmatrix} F_x & 0 & C_x \\ 0 & F_y & C_y \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} r_0 & r_1 & T_x \\ r_3 & r_4 & T_y \\ r_6 & r_7 & T_z \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \quad (11.3)$$

去掉了值恒为零的 Z 和与之相乘的旋转矩阵 R 的第三列, 此时投影矩阵简化为

$$M = \begin{bmatrix} m_0 & m_1 & m_3 \\ m_4 & m_5 & m_7 \\ m_8 & m_9 & m_{11} \end{bmatrix} = \begin{bmatrix} F_x r_0 + C_x r_6 & F_x r_1 + C_x r_7 & F_x T_x + C_x T_z \\ F_y r_3 + C_y r_6 & F_y r_4 + C_y r_7 & F_y T_y + C_y T_z \\ r_6 & r_7 & T_z \end{bmatrix} \quad (11.4)$$

而成像共线方程式可以简化为

$$\begin{cases} \tilde{x} = \frac{m_0 X + m_1 Y + m_3}{m_8 X + m_9 Y + m_{11}} \\ \tilde{y} = \frac{m_4 X + m_5 Y + m_7}{m_8 X + m_9 Y + m_{11}} \end{cases} \quad (11.5)$$

移项可得到关于 m_0 、 m_1 、 m_3 、 m_4 、 m_5 、 m_7 、 m_8 、 m_9 、 m_{11} 的线性齐次方程组

$$\begin{cases} X_i m_0 + Y_i m_1 + m_3 - \tilde{x}_i X_i m_8 - \tilde{y}_i Y_i m_9 - \tilde{x}_i m_{11} = 0 \\ X_i m_4 + Y_i m_5 + m_7 - \tilde{y}_i X_i m_8 - \tilde{y}_i Y_i m_9 - \tilde{y}_i m_{11} = 0 \end{cases} \quad (11.6)$$

两步法标定像机:

第一步, 线性求解像机参数的初值。

基于简化像机模型, 根据共面控制点的成像列出投影矩阵部分元素满足的线性方程组(11.6)进行求解, 再从中分解出像机的等效焦距、平移向量和旋转角。由于进行平差优化时需要对像机的各个参数进行改正, 因而在本步骤中必须对投影矩阵和旋转矩阵进行分解。

第二步, 平差优化求解像机参数的精确值。

控制点数为 n , 采用光束法平差优化求解待标定像机参数的观测方程为

$$\begin{cases} x_i = F_x \frac{\bar{X}_i}{\bar{Z}_i} + C_x + \delta_{x,i} \\ y_i = F_y \frac{\bar{Y}_i}{\bar{Z}_i} + C_y + \delta_{y,i} \end{cases}; (i = 0, 1, \dots, n-1) \quad (11.7)$$

由于控制点共面, $Z_i \equiv 0$, 因而式(11.7)中

$$\begin{cases} \bar{X}_i = r_0 X_i + r_1 Y_i + T_x \\ \bar{Y}_i = r_3 X_i + r_4 Y_i + T_y \\ \bar{Z}_i = r_6 X_i + r_7 Y_i + T_z \end{cases} \quad (11.8)$$

观测方程中的像差为

$$\begin{cases} \delta_{x,i} = \left[(k_0 d_{x,i} + k_1)(d_{x,i}^2 + d_{y,i}^2) + k_3 d_{x,i}^2 + k_4 d_{x,i} d_{y,i} \right] F_x \\ \delta_{y,i} = \left[(k_0 d_{y,i} + k_2)(d_{x,i}^2 + d_{y,i}^2) + k_3 d_{x,i} d_{y,i} + k_4 d_{y,i}^2 \right] F_y \end{cases} \quad (11.9)$$

其中

$$\begin{cases} d_{x,i} = (x_i - C_x) / F_x \\ d_{y,i} = (y_i - C_y) / F_y \end{cases} \quad (11.10)$$

将像差系数初值取为 0，将图像中心指定为像机主点，像机其它参数的初值取为第一步线性求解的结果。

附录 B PNNLOG 算子性质证明

证明：尺度参数 σ 不同的 PNNLOG 算子对同一高斯点的响应最大值处，PNNLOG 算子的 σ 和高斯点的 σ 相等，同时响应最大值与高斯点的 σ 大小无关。

二维高斯函数可以表示为：

$$e^{-(x^2+y^2)/2\sigma_1^2} \quad (12.1)$$

而二维 PNNLOG 算子为：

$$\frac{2\sigma_2^2 - x^2 - y^2}{2\pi\sigma_2^4} e^{-(x^2+y^2)/2\sigma_2^2} \quad (12.2)$$

那么响应可以以下积分进行表示：

$$R = \iint_D \frac{2\sigma_2^2 - x^2 - y^2}{2\pi\sigma_2^4} e^{-\frac{(x^2+y^2)}{2\sigma_2^2}} e^{-\frac{(x^2+y^2)}{2\sigma_1^2}} dx dy \quad (12.3)$$

令 $\sigma_1 = \sigma_2/k$ 则

$$R = \iint_D \frac{2\sigma_2^2 - x^2 - y^2}{2\pi\sigma_2^4} e^{-\frac{(1+k^2)x^2}{2\sigma_2^2}} e^{-\frac{(1+k^2)y^2}{2\sigma_2^2}} dx dy \quad (12.4)$$

由于 PNNLOG 算子的空间对称性，且 PNNLOG 算子只在一定区域以内不为 0，而该区域外均为 0，所以对(12.4)式积分得：

$$R = \frac{2}{1+k^2} - \frac{2}{(1+k^2)^2} \quad (12.5)$$

(12.5)式分别对 k 求一阶导和二阶导，可以得到，当 $k=1$ （排除等于 0 和-1 的情况）的时候， R 取极大值。从而说明不同尺度的 PNNLOG 算子对同一个高斯点的响应，当 $\sigma_{Pnlog} = \sigma_{gauss}$ 时取极大值，同时响应最大值与高斯点的 σ 大小无关。

作者在学期间取得的学术成果

学术论文

- [1] 张恒, Lichun LI, You LI, Qifeng Yu. A scale rotation adaptive new Mean Shift tracking method[C], Photonics Asia 2007 Conference and Exhibition, 2007, 6833: 68330S-1~68330S-8. (本文被 EI 全文检索: 082111273644) 论文集由美国 SPIE 出版, 并在 SPIE 数字图书馆中出版。
- [2] 张恒, 于起峰, 丁晓华, 李由, 李立春. 基于加权 Gabor 梯度的新型多尺度角点检测方法[J], 中国图象图形学报, 2007, 12(8): 1377~1382
- [3] 张恒, 李立春, 李由, 于起峰. 基于差异性加权的最小二乘图像匹配的跟踪方法[J], 光电工程, 2008, 35(4): 23~27. (本文被 EI 检索: 082111273644)
- [4] 张恒, 李由, 李立春, 于起峰. 基于差异性加权的 Mean Shift 跟踪方法[J], 光学技术, 2008, 34(3): 404~407. (本文被 EI 检索: 082611337812)
- [5] 张恒, 李立春, 于起峰. 一种目标区域最佳椭圆描述确定方法[J], 应用光学, 2008, 29(4): 498~502.
- [6] 张恒, 李由, 李立春, 雷志辉. 一种尺度自适应小目标实时检测方法[J], 应用光学, 2008, 29(1): 9~13.
- [7] 张恒, 李立春, 于起峰. 一种运动平台下运动目标检测方法[J], 光电工程, 2008, 35(9): 11~16
- [8] 李由, 张恒, 李立春. 基于多测量融合的粒子滤波跟踪算法. 国防科技大学学报[J]. 2007, 29(5): 26~30. (本文被 EI 全文检索: 074910963243)
- [9] LI You, 张恒. Region and boundary based level set method for contour tracking[C]. International symposium on photoelectronic detection and imaging 2007 (ISPDI' 07), Beijing, 2007, 66230Q. (本文被 EI 全文检索: 082011259697)
- [10] 李由, 张恒, 雷志辉. 基于生物视觉 Center-surround 机制的光团目标检测与跟踪[J]. 应用光学. 2008, 29(2): 283~288.
- [11] 李由, 张恒, 李立春. 基于粒子滤波的多测度融合视觉跟踪方法[J]. 国防科技大学学报. 2007, 29(5): 26~30. (本文被 EI 检索: 074910963243)
- [12] Lichun LI, 张恒, Dan Fu, You LI, Qifeng Yu. Image matching based on epipolar and local homography constraints[C], Photonics Asia 2007 Conference and Exhibition, 2007, 6833: 68330Z-1~68330Z-8 (本文被 EI 全文检索: 081411180001)

参加的科研项目

- [1] 某型号无人机着陆实时分析系统, 主要完成人, 2006 年;
- [2] 宇航员运动分析系统, 主要完成人, 2006 年;
- [3] 无人机对地定位系统, 主要完成人, 2004 年;

作者: [张恒](#)
学位授予单位: [国防科学技术大学](#)
被引用次数: 1次

本文读者也读过(1条)

1. [祝国浩](#) [基于无人机机载电视成像的地面目标跟踪](#)[学位论文]2009

引证文献(1条)

1. [孙祥一](#), [胡建](#), [王鲲鹏](#), [朱肇昆](#), [张小虎](#) [高速摄像三维图像分析技术与应用](#)[期刊论文]-[宇航计测技术](#) 2010(6)

本文链接: http://d.g.wanfangdata.com.cn/Thesis_Y1523979.aspx