

摘要

基于运动成像平台的运动目标检测，文献中又称为独立运动目标（Independently Moving Objects）检测是计算机视觉的研究热点之一，在工业自动化、视频监控、人工智能、医学图像分析、军事等领域有着广阔的应用前景。基于单一视觉传感器的独立运动目标检测方法，由于成像平台运动、相机抖动和有限图像分辨率等多种因素的影响，检测实时性和稳定性仍需改进。

惯性传感器与视觉传感器具有很好的互补特性，两者的融合可以解决单一视觉或惯性传感器运动估计存在的模糊，从而改善独立运动目标检测性能。一方面，图像观测可以修正惯性传感器的误差积累，可以区分传感器的旋转、加速运动、重力场以及加速度测量中的偏差的影响；另一方面，惯性传感器数据可以解决图像场景退化（Degeneration）问题例如场景中包含的特征点少、特征点无限远，可以解决由于特征点的进入和离开视野所造成的运动估计不连续性问题，可以使运动估计对特征的误匹配具有鲁棒性。

本文首先回顾了基于视觉传感器的静止成像平台、运动成像平台以及基于多传感器融合的运动成像平台下的运动目标检测方法。在分析惯性传感器数据特点和融合检测策略的基础上，本文对三种典型应用场景下的独立运动目标检测进行了研究：(1) 针对成像平台沿成像光轴方向的平移运动可忽略或场景的深度变化远小于场景绝对深度的情形，本文提出了一种基于运动补偿的融合单目视觉传感器和惯性传感器的独立运动目标检测的新方法，在单目工业相机和微机电惯性传感器组成的传感器平台采集到的室内数据上进行了实验；(2) 针对成像平台旋转运动可忽略的情形，本文提出了一种基于扩展点光流一致性约束的融合单目视觉传感器和惯性传感器的独立运动目标检测新方法，在 VISAT 多传感器移动测量系统实测的单目序列图像和惯性传感器数据上进行了实验；(3) 针对无成像平台运动模式限制的复杂场景中独立运动目标检测情形，本文提出了一种基于时域空域双重采样的粒子滤波的融合双目立体视觉与惯性传感器的独立运动目标检测的新方法，在 VISAT 多传感器移动测量系统实测的双目序列图像和惯性传感器数据上进行了实验。不同应用场景下的实验结果与分析证明了本文提出的三种方法的有效性。

主题词：独立运动目标检测 多传感器融合 运动补偿 光流 粒子滤波

ABSTRACT

Moving objects detection with moving imaging platforms, also known as **independently moving objects (IMOs)** detection in the literature, has been the focus of computer vision research and is widely applied in areas such as industry automatization, video surveillance, artifical intelligence, medical image analysis and military applications. Independently moving objects detection which can be applied to practical situations can not be easily achieved by vision sensor alone considering the complicated effects of platform motion, imperfect camera calibration and limited image resolution.

Vision and inertial sensors are good candidates to be deployed together since each can be used to resolve the ambiguities in the estimated motion using the other modality alone. For instance, image measurements can help to counteract the error that accumulates when integrating inertial readings, and can be used to distinguish between the effects of sensor orientation, accerlation, gravity, and bias in accelerometer measurements. On the other hand , inertial data can help to resolve the ambiguities in motion estimation by a camera that sees a degenerate scene, such as one containing too few features, features infinitely far away; to remove the discontinuities in estimation motion that result from features entering or leaving the camera's field of view; to make motion estimation more robust to mistracked image features .

This paper firstly presents a general introduction to the methods involving detection of moving objects by stationary imaging systems, moving imaging systems with vision sensors alone and by moving imaging systems fusing multiple sensors. Based on the analysis of inertial data and moving objects detection strategy by vision and inertial sensors, we have devoted our research to three particular sets of applications. (1)For the case of moving platform has neglectible motion components along the optical axis or varations of scene depth are relatively small in comparion to the absolute depth of observed scene, we proposed a novel IMOs detection method using monocular vision and inertial sensors by motion compensation. Experiments have been carried on the data collecting indoors using a monocular industrial camera and MEMS (Micro Electronic Magetic System) IMU (Inertial Measurement Unit) on a moving platform. (2) For the case of moving platform has neglectible rotation motion components, we proposed a novel IMOs detection method using monocular vision and inertial sensors on the basis of FOE (Focus of Expansion)-related optical vectors constraint. Experiments have been carried on the real-road mobile mapping monocular vision and geo-referenced image sequence of VISAT systems. (3) For the general case of a moving platform with no assumption made on the motion characteristics or the observed scene, we proposed a

novel IMOs detection method using binocular vision and inertial sensors under a spatial-temporal parallel resample particle filtering framework. Experiments have been carried on the real-road mobile mapping binocular vision and geo-referenced image sequence of VISAT systems. Various experiments of particular sets of applications have proved the effectiveness of the proposed methods.

Key Words: Independently Moving Objects Detection Sensor Fusion
Optical Flow Motion Compensation Particle Filter

表 目 录

表 1.1 典型的多传感器融合运动目标检测方法.....	7
表 2.1 惯性导航数据	14

图 目 录

图 1.1 反透视投影变换.....	6
图 2.1 捷联惯性导航系统框图.....	12
图 2.2 微机电惯性传感器.....	13
图 2.3 捷联惯性导航系统数据解算框图.....	13
图 2.4 参考坐标系统.....	14
图 2.5 成像平台运动投影.....	15
图 3.1 透視投影模型.....	18
图 3.2 运动补偿法流程图.....	20
图 3.3 检测算法流程图.....	21
图 3.4 不同运动速度下的差分图.....	22
图 3.5 数据同步处理.....	23
图 3.6 系统实验平台.....	24
图 3.7 独立运动目标实验环境配置.....	24
图 3.8 室内简单背景下的独立运动目标检测.....	25
图 3.9 室内复杂背景下的独立运动目标检测.....	26
图 4.1 典型的平移运动模型.....	28
图 4.2 光流的形成.....	29
图 4.3 典型运动模式下的光流图.....	30
图 4.4 扩展点性质.....	32
图 4.5 扩展点位置.....	32
图 4.6 运动成像分析.....	33
图 4.7 光流法检测流程图.....	34
图 4.8 三帧共现性约束滤波.....	35
图 4.9 幅值比值统计滤波.....	36
图 4.10 运动场中的扩展点与光流.....	36
图 4.11 实验数据 VISAT TM 测绘平台	37
图 4.12 光流场稳健性滤波	37
图 4.13 独立运动目标检测	38
图 4.14 独立运动目标检测	39
图 5.1 立体视觉几何结构	41
图 5.2 会聚式立体视觉模型	42
图 5.3 三角解算原理	43

图 5.4 状态空间模型	46
图 5.6 极线几何约束	49
图 5.7 极线约束匹配区域	49
图 5.8 特征匹配搜索距离	49
图 5.9 粒子滤波独立运动检测系统	52
图 5.10 SIFT 特征提取及运动矢量提取	52
图 5.11 独立运动目标检测结果	53

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研究成果。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： 融合视觉和惯性传感器的独立运动目标检测

学位论文作者签名： 孙浩 日期： 2008 年 11 月 13 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： 融合视觉和惯性传感器的独立运动目标检测

学位论文作者签名： 孙浩 日期： 2008 年 11 月 13 日

作者指导教师签名： 王利 日期： 2008 年 11 月 13 日

第一章 绪论

运动成像平台下的运动目标检测，文献中又称为独立运动目标(**Independently Moving Objects**)检测^{[58][62]}是计算机视觉的研究热点之一，在工业自动化、视频监控、人工智能、商业、体育、医学和军事等领域有着广阔的应用前景。独立运动目标的正确检测是实现目标识别与跟踪、运动图像编码、安全监控、基于内容的检索等视频分析和应用的基础，具有十分重要的理论和现实意义。

1.1 课题背景和意义

近年来，基于运动成像平台的运动目标检测又称独立运动目标(**Independently Moving Objects**)检测^{[58][62]}逐渐成为计算机视觉研究领域的一个热点，它在工业自动化、视频监控、人工智能、军事等领域有着广阔的应用前景。基于单一视觉传感器的独立运动目标检测方法，由于成像平台运动、相机抖动和有限图像分辨率等因素的影响，检测实时性和稳定性仍需改进^[10]。基于多传感器信息融合的方法可以改善检测性能^[61]。惯性导航系统由于其不靠外界任何条件，能自主地提供姿态、位置、速度等信息的特点，在航天测控、智能导航和计算机视觉等领域有着广泛的应用。但由于传统的惯性传感器体积大、价格贵，很多实际应用场合中限制了惯性传感器的使用。近年来，随着微机电 MEMS (Micro Electro Mechanical System) 制造技术、微惯性测量器件逐渐成熟^[3]，微机电捷联惯导系统 MEMS SINS(Strapdown Inertial Navigation System) 发展迅速。但市场上广泛应用的中低端微机电惯性测量单元 MEMS IMU (Inertial Measurement Unit) 由于测量精度和数据漂移等问题，不能单独应用于自主定位与运动参数的获取。惯性传感器与视觉传感器具有很好的互补特性，两者的融合可以解决单一视觉或惯性传感器运动估计存在的模糊性^{[11][12][13]}，从而改善独立运动目标检测性能。一方面，图像观测可以修正惯性传感器的误差积累，可以区分传感器的旋转、加速运动、重力场以及加速度测量中的偏差的影响^[11]；另一方面，惯性传感器数据可以解决图像场景退化(Degeneration)问题^[11]例如场景中包含的特征点少、特征点无限远，可以解决由于特征点的进入和离开视野所造成的运动估计不连续性问题，可以使运动估计对特征的误匹配具有鲁棒性。

融合视觉和惯性传感器的独立运动目标检测具有十分重要的现实意义和应用价值，例如在智能交通系统中^[42]，运用基于视觉和惯性传感器融合的方法对独立运动目标进行报警，可以大大改善系统性能，减轻人工干预负担，实现自动化；在车辆智能驾驶过程中^[43]，对行人或其他运动车辆的实时检测，可以保证驾驶安

全；在工业智能机器人应用方面，快速有效地检测出独立运动目标，避免产生事故，是多种任务顺利完成的前提；在高精端武器导航领域，基于视觉和惯性传感器的运动姿态估计和独立运动目检测是实现精确打击的前提；在移动测绘领域^[47]，独立运动目标的检测和归类是评估测绘数据有效性的一个重要手段。

1.2 课题研究现状及挑战

目前文献中有大量的运动目标检测方法，根据所使用的传感器不同，可分为基于视觉传感器的运动目标检测和基于多传感器融合的运动目标检测。基于视觉传感器的运动目标检测根据成像平台是否运动，又可以分为两类：一类是成像平台静止情况下的运动目标检测，一类是成像平台运动情况下的运动目标检测。为区别于序列图像中的成像平台的运动投影，文献中常把成像平台运动情况下的运动目标运动目标的检测称为独立运动目标（Independently Moving Objects）检测^{[58][62]}。

1.2.1 基于视觉传感器的运动目标检测

● 静止成像平台下的运动目标检测

成像平台静止情况下的运动目标检测方法非常多，根据目标特性，背景的复杂度，具体应用场合的要求等不同情况，算法有很大的差别。但总体思路主要有三种，一是对序列图像的时域分析，二是对序列图像的空域分析，三是结合时空域的信息分析处理序列图像。基于时域分析主要是以时间变化检测为准则，利用帧间的相关性，例如帧间差、光流场、运动矢量等信息实现运动目标的检测和分割。基于空域分析主要是以空间同性为准则，利用图像的亮度、色度纹理或其他的统计量信息实现运动目标的检测和分割。时空域结合的方法继承了前两者的优势，综合利用序列图像可获取的信息，是目前的研究热点和未来的发展方向。文献中针对成像平台静止情况下的运动目标检测方法大致可以分为以下几类：

1. 基于变化的运动目标检测：固定空间位置时域的变化性包含了运动目标的信息，基于变化的方法利用差异性进行运动目标检测。目前比较常用和成熟的算法有：

◆ 背景减法：根据某种背景模型更新参考图像，计算当前图像与参考图像的差分图像，然后阈值化分割出运动物体，这种方法计算简单，如果参考图像选择得当，这种方法的优点是可以准确地分割出运动物体。背景模型的建立更新有很多方法，基于统计信息的自适应背景更新模型法如高斯模型，混合高斯模型，非参数化模型法；运用稀疏采样的方法估计背景模型的方法；基于预测的方法，如卡尔曼滤波和维纳滤波方法。文献[8]中，提出了一种基于概率的动态信息窗口的自适应背景更新方法解决背景的复杂性问题。文献[9]中将图像的每

象素点与该象素点对应的混合高斯模型的各个高斯模型进行比较，从而将图像分成前景和背景的部分。Haritaoglu 等^[23]利用最小、最大强度值和最大时间差分值为场景中每个像素进行统计建模，并且进行周期性地背景更新；McKenna 等^[24]利用像素色彩和梯度信息相结合的自适应背景模型来解决影子和不可靠色彩线索对于分割的影响；Karmann 与 Brandt^[25]、Kilger^[26]采用基于卡尔曼滤波（Kalman Filtering）的自适应背景模型以适应天气和光照的时间变化；Stauffer 与 Grimson^[27]利用自适应的混合高斯背景模型（即对每个像素利用混合高斯分布建模），并且利用在线估计来更新模型，从而可靠地处理了光照变化、背景混乱运动的干扰等影响；Toyama 等^[28]提出的 Wallflower 算法利用线性维纳滤波器学习和预测背景变化，并在像素级，区域级和相框级三个层次上进行背景的维护；Elgammal 等^[29]提出了一种非参数模型对动态背景进行建模；Hanzi wang 等^[30]在分析对比各种背景模型的基础上，提出了一种基于样本一致性的背景建模方法 SACON，对标准 Wallflower 序列图像实验表明背景的建模可以适应多种因素的变化。背景减法依赖于所采用的背景模型更新方法，以弥补动态场景中的光线变化等因素带来的不利影响，不适合在复杂环境下运动物体的检测，因为户外复杂环境条件下很难建立背景模型。

- ◆ 时域差分法：对相邻的两帧（或多帧）连续图像进行相减，以去除静止或移动较慢的物体和背景，它是消除图像中的静止物体以及提供运动物体运动轨迹的最直接的方法，优点是计算简单且不易受环境光线变化的影响，缺点是最终的效果与图像采样频率以及被检测物体的速度有关，且不能够完整的分割运动对象，不利于进一步的对象分析和识别，分割出的区域比实际的要大，在运动目标中会形成空洞。Lipton^[35]等利用两帧差分的方法从序列图像中检测出运动目标进行分类和跟踪。为提高差分图像法的性能，最近有学者又提出了累积图像差分法和对称图像差分法，这两种方法都可以很有效的改善简单帧差法的不足。Kameda^[34]采用概率统计的方法通过三帧连续图像间的运动信息分割运动物体，N Paragios^[35]等人提出了基于空间 Markov 随机场的连续帧图像间的差值统计量来描述帧间差异，此外还有一些学者提出结合单帧图像分割结果来修正差分图像法所得到的图像运动信息的方法。

2. 基于运动能量的运动目标检测：把序列图像看作由二维空域加上时间维构成的三维空间，使用时空梯度算子计算每一个像素点的各个时空梯度方向上的分量，经过高斯平滑后的时空梯度分量称为运动能量^{[29][36][37]}。运动对象经过的那些位置的像素，因所有像素点基本都沿某个一致方向运动，这一方向的运动能量值较大，而杂乱运动像素（树叶，烟火等）各个方向的运动能量基本均衡，计算能量梯度后，杂乱运动像素点处的能量梯度值基本为零，而一致运动像素的能量梯度很大，

这样真正的运动对象就被检测出来。运动能量法能消除背景中的抖动像素，使按某一方向运动的对象更加突出地显现出来，其缺点是只能估计出运动对象的大概位置，不能够精确提取对象。

3. 基于时空域联合信息的运动目标检测：基于时域的运动目标检测可以标记出静止和运动的区域，但一般无法提供准确的运动边界，基于空域的运动物体分割可以提供边界信息，但常常会造成过分割。融合时域和空域信息的运动目标检测方法，是在检测过程中通过区域捆绑逐步融合时域和空域信息，而不是在时域分割结束之后再融合空域信息，通过早期分割得到许多小区域，然后将这些小区域捆绑成一些捆绑核，再将剩下的区域通过强或弱的规则捆绑到相邻的捆绑核，从而实现目标区域的分割。Ma and Zhang^[33]提出了基于图象时域和空域熵的运动目标检测，基于象素点属性的时空不确定性进行运动目标的检测取得了较好效果，Guo 等^[34]对基于时空熵的检测算法进行改进，提出了基于差分图象熵运动目标检测算法。目前研究热点正是综合利用各种图像的时空域信息实现运动目标的自动检测和分割。

4. 基于高阶统计的运动目标检测^{[31][32]}：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。实际中遇到的许多信号具有非零的高阶统计，且噪声呈现高斯分布。视频序列图像中，由于目标的运动，相邻的两帧图像的帧差相当于一个零均值的、对称的概率密度函数，所以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波动、慢变的背景纹理都相当于高斯信号，所以可以用图像信号的四阶统计或者四阶矩来提取运动区域。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相适应变化地阈值，将计算出的四阶矩与其比较，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。

5. 基于特征聚类的运动目标检测^{[14][19]}：目标通常被认为是根据特定的条件具有一致性的一组区域，且根据多个特征能够区分这些不同的区域。图像中常用的一些特征如坐标、线条、拐点、轮廓、纹理、形状、颜色、运动、直方图、空间连同性、时间一致性等。此类算法可以简单地同时检测出具有类似性质的多个对象，但是由于对象或对象间的遮挡，会影响某些特征的出现或隐藏，这使得检测算法变得更复杂。

此外，近年来基于 MRF 场的方法，基于人工神经网络的方法，基于小波的方法，分形方法和遗传算法在运动目标检测中的应用也有大量的研究。

● 独立运动目标（Independently Moving Objects）检测

成像平台运动情况下的运动目标检测，又称独立运动目标检测^{[58][62]}，比成像平台静止情况下的难度更大，其很多检测方法以成像平台静止情况下的检测方法为

基础。目前文献中大量的独立运动目标检测方法大致可以分为以下几类：

1. 基于先验知识的方法；基于先验知识的方法利用先验信息获得运动目标在图像中的位置，常用的运动目标的先验信息有对称性、色彩信息、阴影信息和几何特征（例如角点、水平 / 垂直边缘）、纹理信息、运动物体特性（车灯信息等）。
 - ◆ 对称性：人造物体通常都有着较明显的对称性信息，这种对称性信息可以在计算机视觉领域被广泛的用来进行目标检测与识别。文献[40][41]利用对称性对运动车辆进行检测。
 - ◆ 色彩信息：在车辆道路检测等多种应用场合，运动目标与背景在色彩或亮度信息上存在差异性，这种差异性可用来运动目标的识别。文献[42][43]用色彩信息对道路进行检测从而将运动车辆从背景中分离。
 - ◆ 阴影信息：阴影信息的存在是由于目标的遮挡效应所引起的，遮挡区域的灰度特性与背景区域的特性存在不一致性，通过检测阴影信息的变化，可以检测出运动目标的存在。
 - ◆ 几何特征：不同的运动目标具有不同的几何特征，例如车辆通常可以表示成矩形具有四个显著角点的模型，人脸可以表示成具有椭圆特性的模型，通过模板匹配法可以利用几何特征进行运动目标的检测。Bertozzi 等^[40]提出了一种基于角点的方法来得到运动车辆的可能位置，Matthews 等^[41]利用边缘检测获取强的垂直边缘，在边缘图的基础上计算垂直积累定位运动车辆的左右边界。
 - ◆ 纹理信息：运动目标的存在会导致图像局部灰度特性的变化，不同的运动目标在图像中产生的灰度变化规律不同的纹理模型，这种纹理模型可用来检测运动目标。

2. 基于立体视觉的方法^{[43][44]}；基于立体视觉的方法利用空间立体信息进行运动目标检测，常采用的方法有差异图和反透视投影变换。

- ◆ 差异图：在双目视觉系统中，对应像素点在左视图和右视图中不一致性成为差异，所有的像素点的不一致性构成差异图。获得了差异图后，在同一感兴趣深度（根据差异间隔确定）的像素点可以构成相应的差异直方图，如果运动目标出现在某感兴趣深度范围，相应的差异直方图中块会出现突变，通过检测差异直方图的突变可以检测运动目标^[44]。
- ◆ 反透视投影变换：通过约束目标点在水平平面从而进行逆投影解算。假设 p 为三维空间中一点，透视投影模型下，其在图像平面中的投影点为 p' 与投影中心连线与投影平面的交点。反透视投影变换(见图 1.1)的步骤为： p' 为图像中一点，其与投影中心 N 连线与水平面的交点为反透视投影变换结果。如果进行透视变化和反透视变化，水平面将映射为其本身，高出水平面的场景会不同程度失真。通过检测失真度可以检测独立运动目标。在平坦道路的假设前提下，Zhao

和 Yuta^[42]在立体视觉框架下，利用反透视投影变换在给定左图像的情况下对右图像进行预测后进行运动目标的检测。

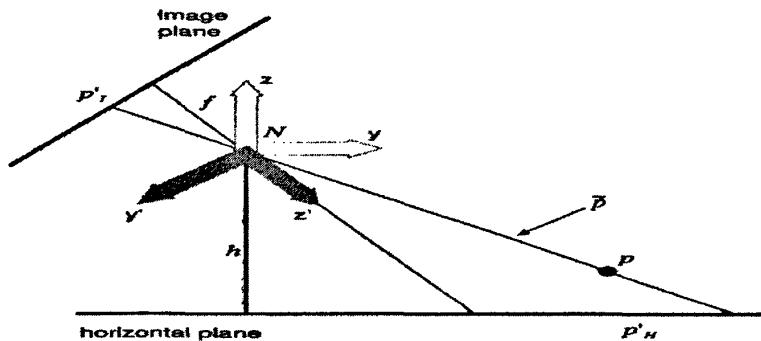


图 1.1 反透视投影变换

3. 基于光流的方法^{[43][44][62]}；光流是空间运动物体被观测面上的像素点运动产生的瞬时速度场，包含了物体 3D 表面结构和动态行为的重要信息。一般情况下，光流由相机运动、场景中目标运动，或两者的运动产生。当场景中有独立的运动目标时，通过光流分析可以确定运动目标的数目、运动速度、目标距离和目标的表面结构。光流研究已经在环境建模、目标检测与跟踪、自动导航及视频事件分析中得到了广泛的应用。

4. 基于运动补偿的方法^{[62][63]}；基于运动补偿的方法按照场景深度变化与场景绝对深度相对关系，可以分为基于二维的补偿方法和基于三维的补偿方法。

◆ 基于二维的补偿方法通常必须满足以下假设之一^[62]：i) 观测场景可以近似为平面，ii) 场景中的深度变化比场景的绝对深度要小得多，iii) 成像系统只存在旋转和缩放运动。基于二维的运动补偿方法是首先对全局背景运动进行运动估计和补偿校正，然后再利用成像平台静止情况下的运动目标检测方法实现检测。运动补偿校正技术的核心问题是平台运动估计，其包括三个方面：摄像机运动仿射参数模型、运动估计匹配准则和运动估计搜索策略。视频图像是三维成像空间在二维图像平面上的投影。图像坐标中各像素点的运动速度并不相同，而是遵循同一摄像机的运动参数，其为全局运动参数。典型的摄像机运动仿射参数模型有：基于平行投影的六参数模型、基于透视投影的八参数模型、基于旋转平移假设的四参数模型等。由于图像中各个背景点的运动参数都相同，因此无需求得所有背景点的运动场，只要选择稀疏特征点估计出图像背景的全局运动参数就可以得到各个背景点的运动位移。但目标点和噪声点与背景点的运动参数明显不同，当它们被选为特征点时，将得不到精确的结果。鲁棒性较差。在实际的系统中必须消除由于目标运动和噪声带来的全局运动参数估计误差。文献中已有大量关于二维运动估计方法的研究，可以归纳为光流场方

法、块匹配方法、像素递归法和相位相关法。

- ◆ 基于三维的补偿方法，主要分为两类，一类是直接利用成像平台运动估计消除相机运动所带来的影响；一类是间接的把场景分为多个平面，利用平面+视差的方法进行处理。Sawheny^[62]等提出一种在三维场景中利用平面+视差方法检测独立运动，利用极线几何约束和形状约束，从序列图像中渐进求解平面和视差参数。基于固定扩展点的假设，Hu 和 Uchimura^[50]通过从相邻图像特征对应性来估计相机运动参数信息的方法检测运动目标。

1.2.2 基于多传感器融合的运动目标检测

基于单一视觉传感器的独立运动目标检测，由于成像平台运动、相机抖动和有限图像分辨率等因素的影响，在实际很多应用场合中检测有效性和系统稳定性仍需改进^[10]。基于多传感器信息融合的方法可以改善检测性能^[63]。多种不同的传感器都可以用来进行独立运动目标检测^{[1][64]}，例如单目或多目视觉（彩色或灰度）传感器、红外（Infra Red）传感器、激光雷达（LIDAR）传感器、GPS/DGP 和惯性传感器（Inertial Sensors）等。融合检测的方法很多，按照数据抽象层次来分，可分为数据级融合检测，特征级融合检测和决策级融合检测^[11]。表 1.1 中列出了几种典型的融合检测方法^[65]。

表 1.1 典型的多传感器融合运动目标检测方法

研究者	传感器类型	融合特征或线索	融合方法
Handmann et al.	单目视觉，雷达	色彩、边缘和局部图像熵	最大似然估计
Stiller et al.	立体视觉，雷达，LIDARS，DGPS/INS	垂直边缘、立体差、光流	卡尔曼滤波
Kato et al.	单目相机，雷达	Kanade-Lucas-Tomasi 特征、距离信息	基于距离信息的序列图像特征耦合
Steaux et al.	彩色相机，雷达	阴影位置、对称性、颜色	置信网络
Thrun et al.	彩色相机（单目），GPS，LIDARs，加速度计，陀螺仪	色彩信息、平台位置、姿态和速度	攀升卡尔曼滤波

视觉传感器和惯性传感器的融合检测方法研究有很多，目前应用较广的方法为贝叶斯滤波方法和神经网络方法。

1.2.3 面临挑战

虽然对运动目标检测的研究已经经历了几十年的时间，但到目前为止仍面临

许多不能很好解决的问题^{[8][9][23][24][25][26][27][30][31][43][44]}：

1)在运动目标分割过程中,如果成像平台是静止的,则运动目标分割相对容易实现,然而在多种应用场合,存在成像平台的运动导致图像中存在多种模式的运动,使得运动目标检测十分困难。

2)视频场景中光线变化问题(渐变、突变、反光等)会影响运动目标的提取。

3)光照条件下,目标存在阴影。与背景不同,运动目标的阴影将随运动目标移动,检测分割的结果可能是阴影与被检测对象连为一个整体出现,使得运动目标形状失真,从而导致相关的聚类工作失败。另一种检测结果是阴影与目标分离成两个独立的运动目标,导致了虚假运动目标的出现。

4)遮挡也是运动目标检测过程中一个难以解决的问题,在运动目标前方的遮挡物可能会作为目标的一部分被提取出来,从而造成检测目标形变,严重时会造成运动目标检测的失败。

5)背景物体的干扰运动,背景中存在的如风中树叶的摇动、水面波光的闪动、车窗玻璃的反光以及天气的变化等许多细微活动,同样会影响到运动目标的检测。

6)另外目标运动过程是否连续、速度如何、是否存在往返运动的现象等问题也影响运动目标的检测结果。

1.3 主要研究内容和工作

本文在研究和分析融合视觉和惯性传感器的独立运动目标检测基本原理和融合策略的基础上,对三种典型应用场景下的独立运动目标检测进行了研究(1)针对成像平台沿成像光轴方向的平移运动可忽略或场景的深度变化远小于场景绝对深度的情形,本文提出了一种基于运动补偿的融合单目视觉传感器和惯性传感器的独立运动目标检测的新方法,在单目工业相机和微机电惯性传感器组成的传感器平台采集到的室内数据上进行了实验; (2)针对成像平台旋转运动可忽略的情形,本文提出了一种基于扩展点光流一致性约束的融合单目视觉传感器和惯性传感器的独立运动目标检测新方法,在 VISAT 多传感器移动测量系统实测的单目序列图像和惯性传感器数据上进行了实验; (3)针对无成像平台运动模式限制的复杂场景中独立运动目标检测情形,本文提出了一种基于时域空域双重采样的粒子滤波的融合双目立体视觉与惯性传感器的独立运动目标检测的新方法,在 VISAT 多传感器移动测量系统实测的双目序列图像和惯性传感器数据上进行了实验。

1.4 论文结构安排

论文共包含五章:

第一章，绪论。主要介绍了课题的来源、研究背景和意义，分析了运动目标检测的研究现状及面临挑战，最后说明论文的主要内容和组织结构。

第二章，技术背景与问题分析。介绍了相关的惯性技术背景和运动成像分析，分析了运动成像平台下融合惯性传感器与视觉传感器的基本原理和融合检测策略。

第三章，基于运动补偿的独立运动目标检测。首先介绍了运动补偿独立运动目标检测方法的基本原理，然后针对成像平台沿成像光轴方向的平移运动可忽略或场景的深度变化远小于场景绝对深度的情形，提出了一种基于运动补偿的融合单目视觉传感器和惯性传感器的独立运动目标检测的新方法，最后在单目工业相机和微机电惯性传感器组成的传感器平台采集到的室内数据上进行了实验。

第四章，基于光流的独立运动目标检测。首先介绍了光流基础和光流计算方法，并对光流场形成进行分析，提出了扩展点光流一致性约束，然后针对成像平台旋转运动可忽略的情形，提出了一种基于扩展点光流一致性约束的融合单目视觉传感器和惯性传感器的独立运动目标检测新方法，最后在 VISAT 多传感器移动测量系统实测的单目序列图像和惯性传感器数据上进行了实验。

第五章，基于粒子滤波的独立运动目标检测。首先介绍了立体视觉基础和宽基线条件下图像配准等有关问题，然后针对无成像平台运动模式限制的复杂场景中独立运动目标检测情形，提出了一种融合双目视觉传感器和惯性传感器的时域空域双重采样粒子滤波独立运动目标检测方法，最后在 VISAT 多传感器移动测量系统实测的双目序列图像和惯性传感器数据上进行了实验。

最后是结束语。主要是谈谈自己一年多的工作体会和下一步的工作展望。

第二章 技术背景与问题分析

2.1 引言

惯性导航系统由于具有自主性、隐蔽性和能获取运动载体完备运动信息的独特优点已应用于各种领域。然而不同级别的惯性传感器的数据精度有很大差别，近年来得到广泛应用的中低端微机电惯性传感器目前还很难独立地应用于自主运动参数的获取。视觉传感器和惯性传感器具有很强的互补特性。目前，惯性传感器与视觉传感器融合在道路障碍监测、自主导航、增强现实、姿态跟踪等多领域的应用是国内外研究的热点^{[10][13][61]}。

2.2 惯性技术

2.2.1 基本思想

惯性导航^[3]是一门综合了机电、光学、数学、力学、控制及计算机等学科的尖端技术。由于惯性是所有质量体的基本属性，所以建立在惯性原理基础上的惯性导航系统不需要任何外来信息，也不会向外辐射任何信息，仅靠惯性导航系统本身就能在全天候条件下，在全球范围内和任何介质环境里自主地、隐蔽地进行连续的三维定位和三维定向，这种同时具备自主性、隐蔽性和能获取运载体完备运动信息的独特优点是其他导航系统所无法比拟的。

惯性导航系统在很多领域里都获得了广泛的应用^[3]，如航天、火星探测器飞行试验、UAV (Uninhabited Aerial Vehicle 无人机)、AUV (Autonomous Underwater Vehicle 自主水下航行器)，其中在军事领域的应用尤为突出，如 GN&C (Guidance, Navigation, and Control 制导、导航与控制)、武器系统、SAR 运动补偿、IRST (Infra-Red Search and Track 红外搜索与跟踪)、探雷等，在民用领域也开始得到广泛的应用，如车辆控制、AVCSS (Advanced Vehicle Control and Safety Systems 先进车辆控制与安全系统)、ITS (Intelligent Transport Systems 智能运输系统)、井下定位和定向等。

惯性导航系统INS (Inertial Navigation System)是以陀螺和加速度计为敏感器件的导航参数解算系统，该系统根据陀螺的输出建立导航坐标系，根据加速度计输出解算出运载体的速度和位置。惯性导航系统通常由惯性测量装置、计算机、控制显示器等组成。惯性测量装置包括加速度计和陀螺仪，又称惯性导航组合，3个自由度陀螺仪用来测量三维转动运动分量，3个加速度计用来测量三维平移运动分量。计算机根据测得的加速度信号计算出飞行器的速度和位置数据。控制显示器

显示各种导航参数。

如果航行器的初始位置和初始速度已知，只要在惯性系中精确地连续测量出航行器的加速度，经过积分就可以得到航行器在任意时刻的位置和速度。陀螺和加速度计构成惯性测量组件 IMU，由速率陀螺得到测量系相对惯性系的转动角速度矢量，由导航计算机可以计算出测量系相对惯性系的姿态。在测量系中由加速度计得到航行器的加速度，按照测量系相对惯性系的姿态，可以将航行器在测量系中的加速度变换到惯性系。理论上，由航行器在惯性系的加速度就可以得到质点的速度和位置。

如果没有其它信息，完全由加速度确定航行器的运动轨迹是一个开环过程，由于 IMU 的精度有限，所计算出的航行器的运动轨迹会产生严重的漂移。在地球表面附近运动时，地球的重力场提供了垂直“测量参照”，而地球的自转牵连速度则提供了方位“测量参照”，惯性导航系统本质上是利用这两个“测量参照”，使惯性导航系统的误差保持为 Schuler 振荡。此处所称“测量参照”是指测量的零位，即水平面是加速度计的“测量参照”，东向是陀螺的“测量参照”。

在地球重力场中，加速度计敏感到的比力不仅有运动加速度，还有重力加速度；只有将运动加速度从比力中分离出来，才能计算出位置和速度。如果知道初始时刻测量系相对地球的姿态和位置，根据地球重力场模型就可以确定出重力加速度矢量，从而分离出运动加速度，并计算位置和速度。再按照地球的自转和航行器相对地球的运动不断地对重力加速度矢量进行修正，就可以继续计算位置和速度，从而实现惯性导航。

2.2.2 导航系统分类

根据构建导航坐标系方法的不同可将惯性导航系统分为两大类型：采用物理平台模拟导航坐标系的系统称为平台式惯性导航系统，采用数学算法确定出导航坐标系的系统称为捷联式惯性导航系统^[3]。

- 平台式惯性导航系统 根据建立的坐标系不同，又分为空间稳定和本地水平两种工作方式。空间稳定平台式惯性导航系统的台体相对惯性空间稳定，用以建立惯性坐标系。地球自转、重力加速度等影响由计算机加以补偿。这种系统多用于运载火箭的主动段和一些航天器上。本地水平平台式惯性导航系统的特点是台体上的两个加速度计输入轴所构成的基准平面能够始终跟踪飞行器所在点的水平面（利用加速度计与陀螺仪组成舒拉回路来保证），因此加速度计不受重力加速度的影响。这种系统多用于沿地球表面作等速运动的飞行器（如飞机、巡航导弹等）。在平台式惯性导航系统中，框架能隔离飞行器的角振动，仪表工作条件较好。平台能直接建立导航坐标系，计算量小，容易补偿和修正

仪表的输出，但结构复杂，尺寸大。

- 捷联式惯性导航系统 SINS 根据所用陀螺仪的不同，分为速率型捷联式惯性导航系统和位置型捷联式惯性导航系统。前者用速率陀螺仪，输出瞬时平均角速度矢量信号，后者用自由陀螺仪，输出角位移信号。捷联式惯性导航系统的出现是惯性技术和计算机技术等多项技术迅速发展的结果。目前，捷联式惯性导航系统在航空、航天、航海和陆地各领域中已得到广泛的应用，已逐步取代平台式惯性导航系统。在捷联式惯性系统中，陀螺仪和加速度计直接安装在运载体上，不需要稳定平台。捷联式惯性系统中的陀螺和加速度计在运载体运动时，要直接感受过载、冲击、振动、温度变化等恶劣环境，从而产生动态误差，所以与平台式相比它用的惯性仪表和器件有特殊的要求。运载体上直接安装的陀螺和加速度计，它们随运载体的运动必须能快速精确地测量并传输给计算机，计算机即时处理陀螺仪的姿态数据，从而在计算机内形成一个“数学解析平台”来取代物理稳定平台功能。捷联惯导系统由于去掉了机械框架系统，结构大大简化，体积小，重量轻，成本低，可靠性高，功耗小，使用方便灵活，维护简便，在运动过程中姿态不受限制，可以实现全姿态变化。捷联惯性系统由捷联惯性测量组合和导航计算机（或捷联计算机）两大部件组成，如图 2.1 所示：

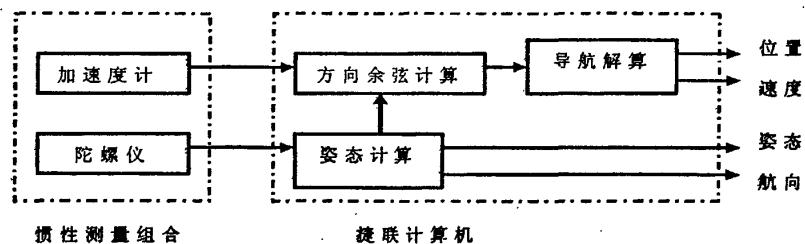


图 2.1 捷联惯性导航系统框图

- 微机电捷联式惯性导航系统 在现代及未来的高科技战场上，强技术对抗的复杂战场环境迫切需要微型飞行器、微型机器人等微小型侦查设备及大量战术武器，这种需求促使惯性导航系统向低成本、微型化、低功耗的方向发展。近年来，随着微机电（MEMS）制造技术，微机电惯性测量器件逐渐成熟，微机电捷联惯导系统（MEMS INS）发展迅速正在成为当前惯性技术领域的一个研究热点。微机电惯性传感器是采用 MEMS 技术研制的加速度计和陀螺，体积小、成本低，可靠性高。MEMS 加速度计是所有 MEMS 传感器中商业市场化应用最为成功的，它在精度方面已能满足战略导弹的应用要求，MEMS 陀螺性能也已接近或达到于战术级导航的水平。随着 MEMS 惯性传感器性能提高，MEMS INS 以及组合导航系统以其体积小、重力轻、功耗低、成本低的特点，大大拓展了惯性导航的应用范围。军事应用和民用领域对导航的需求共同构

成了促进 MEMS INS 不断发展的动力。微电子技术、微机械加工技术的发展将从技术上不断推动 MEMS INS 向高性能、微型化发展，目前整个系统已经集成到了一块电路板内。图 2.2 为几种典型的微机电惯性传感器。

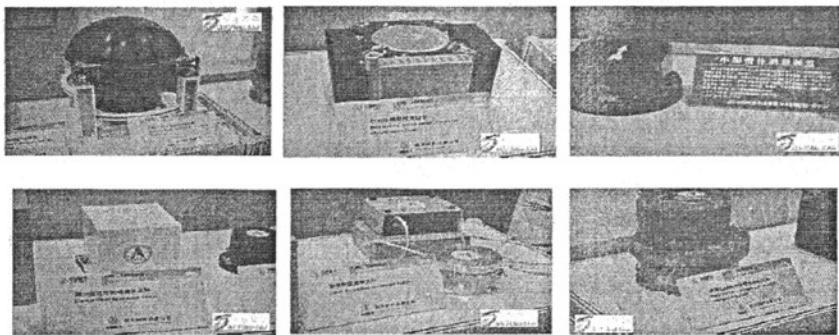


图 2.2 微机电惯性传感器

2.2.3 数据处理

惯性导航系统最基本的数据处理就是通过对加速度的时域二次积分对位置进行估计，存在旋转分量时，必须要进行坐标变换后处理。图 2.3 所示是捷联惯性导航系统的数据解算框图，系统具有三维的正交加速度计和陀螺仪的惯性测量单元(IMU)。

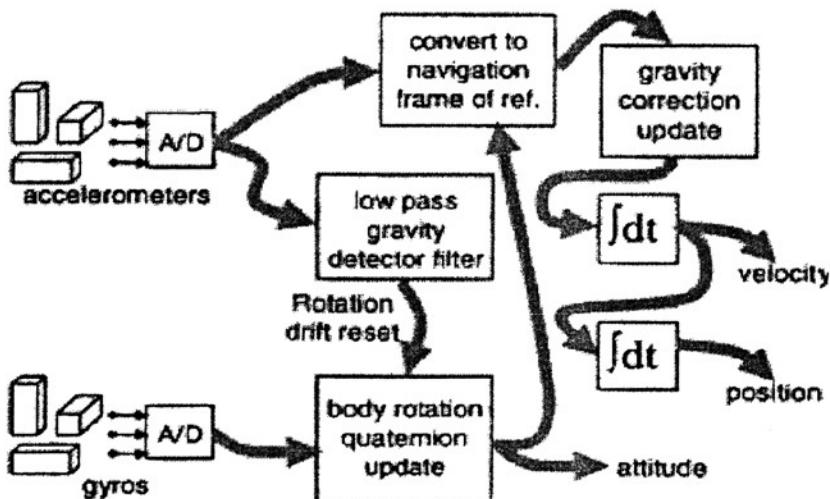


图 2.3 捷联惯性导航系统数据解算框图

不难看出,由惯性传感器可获得的数据如表 2.1,其中 j 表示线性加速度率, s 表示线性加速度, 其他变量意义见表中说明。

表 2.1 惯性导航数据

微分	$\varphi = \ddot{\theta}$ $j = \dot{a}$	角加速度 线性加速度率(jerk)
直接观测量	$w = \dot{\theta}$ $a = g + \ddot{x}$	角速度 线性加速度与重力加速度
积分	θ $v = \dot{x}$ $s = \ddot{x}$	角位置(attitude) 线速度 位置(position)

由于导航坐标系中重力加速度的存在,观测到的线性加速度分量不能直接应用到处理过程。当系统初始处于静止或匀速运动时,惯性传感器沿三个坐标轴方向的线性加速度分量 (a_x, a_y, a_z), 不难确定出重力加速度为 $g = (a_x, a_y, a_z)$ 。

2.3 视觉/惯性传感器融合分析

2.3.1 参考坐标系统

如图 2.4, 图 2.5 在融合视觉和惯性传感器数据分析中涉及到的坐标系统有五个: 1)世界坐标系(W), 地理参考坐标系, 2)导航坐标系(N), 惯性传感数据测量的参考坐标系; 3)机体坐标系(B), 多传感器负载平台坐标系统, 假设其初始位置与导航坐标系重合; 4)相机坐标系(C), 原点位于相机中心, Z 轴于相机光轴重合; 5)图像坐标系(I), 二维图像平面坐标系统, 原点与相机主点融合。

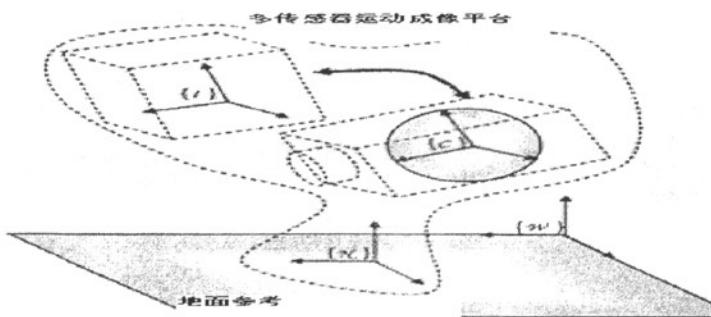


图 2.4 参考坐标系统

2.3.2 平台运动投影

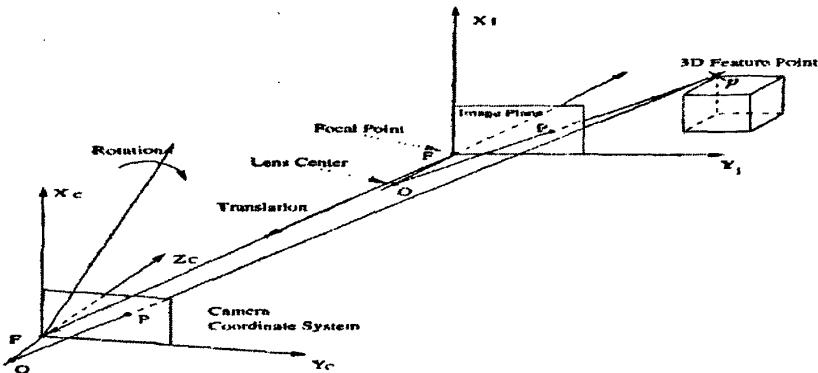


图 2.5 成像平台运动投影

在计算机视觉等领域中，相对于场景中独立运动目标的运动，成像平台的运动通常称为自运动(Ego-Motion)，成像平台自运动分析的参考坐标系通常为地理参考坐标系或导航参考坐标系。假设相机成像采用针孔成像模型，成像平台运动视为刚体运动可以分解为沿三个坐标轴方向的平移和旋转分量，分别表示为 (T_x, T_y, T_z) , $(\Omega_x, \Omega_y, \Omega_z)$ ，平台运动在图像平面中的二维投影为(如图 2.5):

$$\begin{aligned} u &= \left[\frac{-f_c T_x + x T_z}{Z} + \frac{xy}{f_c} \Omega_y - f_c \left(1 + \frac{x^2}{f_c^2}\right) \Omega_y + y \Omega_z \right], \\ v &= \left[\frac{-f_c T_y + y T_z}{Z} - \frac{xy}{f_c} \Omega_y + f_c \left(1 + \frac{y^2}{f_c^2}\right) \Omega_x + x \Omega_z \right] \end{aligned}, (u, v) = (x, y) \quad (2.1)$$

其中 (X, Y, Z) 为空间场景中点在机体坐标系中的坐标， (x, y) 为对应的图像平面上成像点坐标， f_c 为成像焦距。

2.4 融合策略

基于运动成像平台的独立运动目标检测的核心问题是如何消除成像平台自运动的影响，在序列图像分析中即为如何区分平台自运动和场景中独立运动目标运动在图像平面上的投影。由式 2.1 可以看出，平台自运动在图像平面上的投影与目标点的场景深度，平台的自运动参数和成像焦距有关。

惯性传感器可以提供平台的自运动参数的信息，避免了基于二维运动估计三维运动的难题。高精度惯性传感器数据信息可以直接应用于独立运动目标检测中，然而由于造价昂贵，应用受到了极大限制。中低端微机电惯性传感器，由于体积小，价格低廉、易于集成等特点得到了广泛应用，但其数据精度有限，且存在误差积累问题不能直接应用于独立运动目标检测。视觉传感器和惯性传感器有很好

的互补特性，融合视觉和惯性传感器的独立运动目标检测可以改善检测性能^[61]。

惯性传感器可以提供平台运动参数信息很好的近似，但由于精度和误差积累问题，不能直接应用检测问题中；单目视觉传感器不能提供场景的结构信息，立体视觉系统含有景深信息，如何有效地提取景深信息及结合惯性传感器数据应用到独立运动目标检测问题中，面临很多挑战。针对不同的应用场合，应选择不同的融合检测策略。在分析运动投影方程及典型应用场景基础上，本文针对三种不同的典型应用场景提出了三种不同的独立运动目标融合检测方法。

(1)针对成像平台沿成像光轴方向平移运动可忽略或场景的深度变化远小于场景绝对深度的情形，本文提出了一种基于运动补偿的融合单目视觉传感器和惯性传感器的独立运动目标检测的新方法，首先利用惯性传感器平台运动参数信息和景深信息对相邻帧图像进行粗配准，然后提取背景块进行图像精配准，最后采用帧间差法检测独立运动目标，同时利用图像观测修正惯性传感器数据，抑制了惯性传感器的数据漂移问题；

(2)针对成像平台旋转运动可忽略的情形，本文提出了一种基于扩展点光流一致性约束的融合单目视觉传感器和惯性传感器的独立运动目标检测新方法，利用惯性传感器的平台运动信息和成像相机内参数预测运动场中扩展点位置，然后在计算稳健光流场的基础上利用一致性约束对光流矢量进行分类，实现独立运动目标的检测；

(3)针对无成像平台运动模式限制的复杂场景中独立运动目标检测情形，本文提出了一种基于时域空域双重重采样的粒子滤波的融合双目立体视觉与惯性传感器的独立运动目标检测的新方法，利用极线约束下的SIFT特征点立体匹配和运动匹配、惯性传感器平台运动参数等先验信息，在粒子滤波框架下以理想运动矢量与实际运动矢量的差异为观测对独立运动特征点投影状态进行修正，最后进行空间聚类实现独立运动目标的检测。

2.5 本章小结

本章介绍了相关的惯性技术基础，介绍了惯性传感器的数据处理及融合检测分析中涉及到的各个坐标系统，分析了运动成像平台下的图像运动投影，分析了融合视觉传感器和惯性传感器的核心问题和融合策略。

第三章 基于运动补偿的独立运动目标检测

3.1 引言

序列图像中的运动是由相机自运动，相机参数变化（例如变焦等），目标独立运动和场景结构等多因素共同作用形成的。独立运动目标检测的核心问题就是如何消除或补偿成像平台运动的影响。由成像平台运动引起的图像运动不仅与平台运动参数信息有关，而且还与每个点的场景深度信息有关，然而在实际应用场合中很难同时精确的获得这些信息。运动补偿法假设场景深度变化远小于场景绝对深度，采用参数模型法对图像运动建模，补偿平台运动从而实现独立运动目标检测。

3.1.1 问题描述

针对远景航拍检测与跟踪、动态监控等多种实际应用场合中满足以下四种假设之一的情形：

- (1) 成像平台沿成像光轴的平移运动分量可以忽略或只存在旋转运动；
- (2) 场景深度变化远小于场景绝对深度；
- (3) 观测的场景点位于同一平面上；
- (4) 相机存在参数变化，例如变焦；

我们提出了一种基于运动补偿的融合单目视觉和惯性传感器的独立运动目标检测方法，首先利用惯性传感器平台运动信息和景深信息在平移运动模型下对相邻帧图像进行粗配准，然后提取背景块进行精配准，最后采用帧间差法检测独立运动目标，同时利用图像观测修正惯性传感器数据，抑制了惯性传感器的数据漂移问题。

3.1.2 基本原理

假设 (X, Y, Z) 表示原点为相机的笛卡尔直角坐标系统， (x, y) 表示成像平面上相应的坐标系统，如图 3.1 所示。透视投影下（相机模型采用针孔成像模型），空间场景中点 (X, Y, Z) 在图像平面上成像坐标 (x, y) 可以表示为：

$$x = f_c \frac{X}{Z}, y = f_c \frac{Y}{Z} \quad (3.1)$$

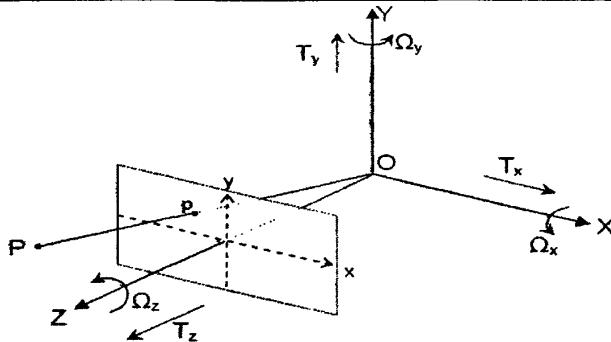


图 3.1 透視投影模型

相机运动可以视为刚体运动，可以分解为沿三个坐标轴方向的平移和旋转分量，分别表示为 (T_x, T_y, T_z) , $(\Omega_x, \Omega_y, \Omega_z)$ ，相机运动在图像的二维投影为：

$$\begin{aligned} u &= \left[\frac{-f_c T_x + x T_z}{Z} + \frac{xy}{f_c} \Omega_x - f_c \left(1 + \frac{x^2}{f_c^2}\right) \Omega_y + y \Omega_z \right] \\ v &= \left[\frac{-f_c T_y + y T_z}{Z} - \frac{xy}{f_c} \Omega_y + f_c \left(1 + \frac{y^2}{f_c^2}\right) \Omega_x + x \Omega_z \right] \end{aligned} \quad (3.2)$$

其中 $(u, v) = (x, y)$ ， Z 为空间场景点的深度， f_c 为相机焦距。

在场景几何结构或相机运动满足以下几种假设之一的前提下，图像中的场景点运动可以用二维的参数化模型近似表示。

1) 平面：观测的场景点都位于同一平面上，平面方程可表示为：
 $Z = A + BX + CY$ ，其中 (A, B, C) 为平面参数，则运动投影方程可以表示为：

$$\begin{aligned} u &= [a + b \cdot x + c \cdot y + g \cdot x^2 + h \cdot xy] \\ v &= [d + e \cdot x + f \cdot y + g \cdot xy + h \cdot y^2] \end{aligned} \quad (3.3)$$

其中

$$\left\{ \begin{array}{l} a = -f_c \alpha T_x - f_c \Omega_y \\ b = \alpha T_z - f_c \beta T_x \\ c = \Omega_z - f_c \lambda T_z \\ d = -f_c \alpha T_y + f_c \Omega_x \\ e = -\Omega_z - f_c \beta T_y \\ f = \alpha T_z - f_c \gamma T_y \\ g = -\frac{\Omega_y}{f_c} + \beta T_z \\ h = -\frac{\Omega_x}{f_c} + \gamma T_z \end{array} \right\}, \quad \left\{ \begin{array}{l} \alpha = \frac{1}{A} \\ \beta = \frac{-B}{Af_c} \\ \gamma = \frac{-C}{Af_c} \end{array} \right\}。图像运动可以表示为八参数$$

数的投影变换。

2) 远景：当场景离相机距离很远时，基于平面的变化相对于绝对景深比较小

时，平面模型是很好的近似。这种情况下，由二次方程变换描述的图像运动场可以达到亚像素的精度。而且，随着 $Z \rightarrow \infty$ ， $\frac{T_x}{Z}, \frac{T_y}{Z}, \frac{T_z}{Z} \rightarrow 0$ ，图像运动中的平移分量可以忽略。远程监控领域窄视角(Field-Of-View, FOV)相机（典型为5°或更小）远景（1km以上）对移动目标的监控通常都满足远景的条件。

3) 相机旋转：当相机只存在旋转运动($T = 0$)或者相机的平移运动可以忽略($|T| \ll Z$)，运动投影方程可以表示为：

$$\begin{aligned} u &= \left[\frac{xy}{f_c} \Omega_x - f_c \left(1 + \frac{x^2}{f_c^2} \right) \Omega_y + y \Omega_z \right] \\ v &= \left[\frac{xy}{f_c} \Omega_y + f_c \left(1 + \frac{y^2}{f_c^2} \right) \Omega_x + x \Omega_z \right] \end{aligned} \quad (3.4)$$

二维图像运动场这种情况下也可以用二次方程变换进行描述。

4) 相机变焦：当相机成像过程中存在变焦时，图像中会产生膨胀或缩放效应，这种情况下的图像运动场可以用八参数投影变换表示，变焦会对参数 b, f 产生影响。

3.1.3 运动补偿

成像平台运动补偿实际上就是一个坐标变换的过程，通过二维运动估计建立相邻图像间的几何变换关系。自上世纪八十年代初以来，人们已经提出了很多运动估计方法，如光流分析法、基于块匹配的运动估计方法、像素递归法、贝叶斯法等等。其中基于块的运动估计方法是目前为止最通用的一种，它也是目前一些主要的视频编码标准(MPEG-1/2/4、H.261/3等)所采用的运动估计方法。块运动估计方法的特点是：原理简单、稳定性好、硬件复杂度小、适用面广。常用的块匹配运动估计快速算法有三步搜索法 TSS、新三步搜索法 N3SS、四步搜索法 4SS 和菱形搜索法 DS。本文中我们采用块匹配法进行平台运动投影的精确估计，搜索方法采用惯性传感器信息和景深信息辅助下的局部搜索法。

3.2 运动补偿法独立运动目标检测

惯性传感器能提供运动平台六个自由度(Degree Of Freedom-DOF)的运动信息，通过坐标变换和误差纠正，可以得到相机与三维空间场景点之间的相对运动参数信息。理论上讲，已知成像焦距和固定场景深度，可以求解二维图像中的精确运动场，然而由于实际场景多样性、成像噪声和惯性传感器数据有限精度等因素的影响，精确运动场往往难以得到。实际应用中为保证系统工作的稳定性需要利用图像观测对惯性传感器数据进行修正以抑制误差漂移问题。

针对远景航空及目标跟踪拍摄等多种应用场合，运动成像平台的旋转分量可以忽略、不存在沿成像光轴方向的平移运动和场景平均深度信息可以获得（高度仪、雷达等）的情况，我们提出了一种基于运动补偿融合单目视觉和惯性传感器的独立运动目标检测新方法，设计了实验思路，并对室内采集到的简单和复杂背景下单目视觉和惯性传感器数据进行了实验，其流程图如图 3.2：

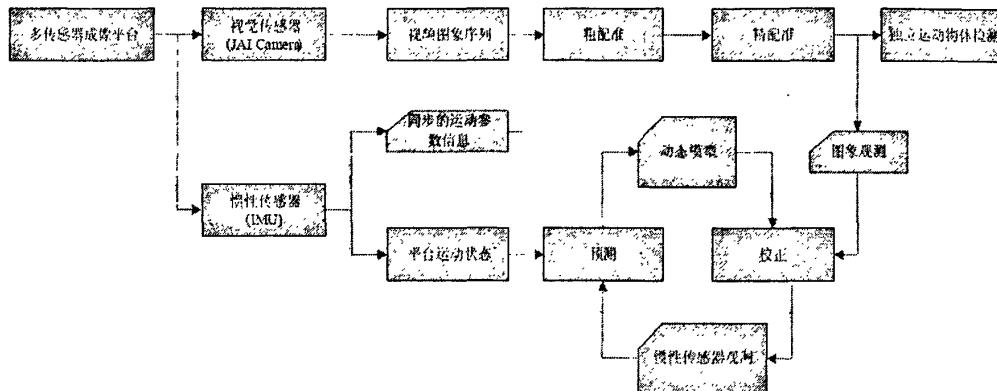


图 3.2 运动补偿法流程图

3.2.1 运动补偿法检测算法

已知相机的运动参数、场景深度信息及成像焦距，理论上图像中每一点由相机运动所形成的速度场可以准确地描述出来。然而由于成像条件变化、运动参数提取误差、场景深度测量误差、有限精度和插值误差等多因素的影响，速度场信息不能直接用于独立运动目标检测。

本章中我们假设运动成像平台的旋转分量可以忽略、不存在沿成像光轴方向的平移运动，场景深度变化远小于场景绝对深度且场景绝对深度已知，式 3.2 可以简化为：

$$u = \left[\frac{-f_c T_x}{Z} \right], v = \left[\frac{-f_c T_y}{Z} \right] \quad (3.5)$$

其中 $(u, v) = (x, y)$ ， Z 为空间场景点的深度， f_c 为相机焦距。

基于以上假设，我们提出一种从粗到精的运动目标检测的新方法，在多传感器信息框架下，首先利用惯性传感器提供的运动平台参数信息和景深信息进行相邻帧图像的粗配准，粗配准差分后提取独立运动目标的粗略位置，然后选择背景图像块局部搜索精配准，最后差分获取独立运动目标的精确位置，有效的克服了特征匹配-最小二乘或直接块匹配等方法中独立运动物体对求解背景运动参数的影响。实验系统中视觉传感器和惯性传感器的都具有高采样率，相邻帧的视野变化较小，我们采用相邻两帧图像时域差分的方法来提取独立运动目标。相邻帧图像

配准采用由粗到精的策略。其算法流程如图 3.3:

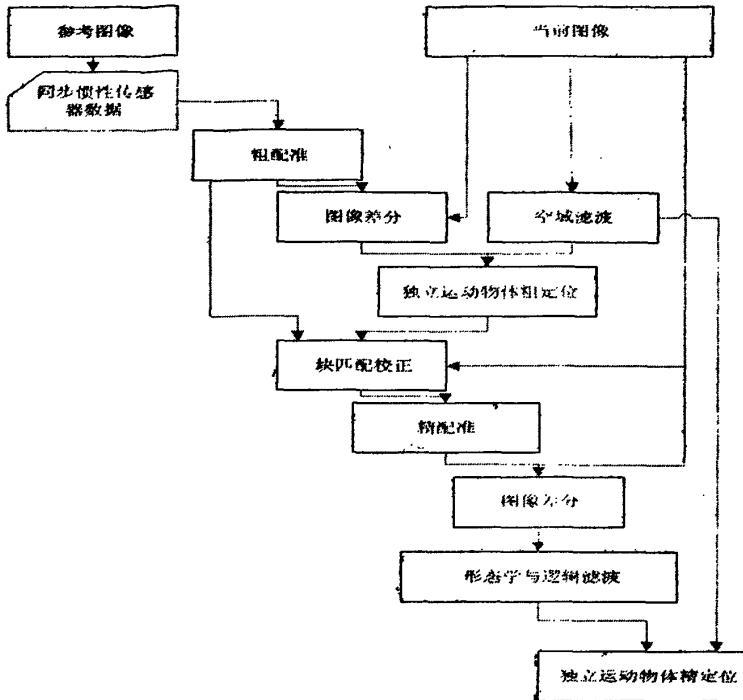


图 3.3 检测算法流程图

检测当前图像中独立运动目标的具体步骤为：

- 1) 获取当前帧图像及前一帧图像，获取两帧图像同步的惯性传感器运动信息及固定景深信息；
- 2) 利用惯性传感器数据信息和景深信息，采用平面运动投影模型进行粗配准后（实验中我们采用平移运动模型）进行差分；
- 3) 对粗配准差分结果图像进行形态学滤波，并与当前帧图像的空域分割图进行逻辑滤波；
- 4) 区域分割及空域滤波得到独立运动物体的粗略位置；
- 5) 在当前帧图像中选择背景块，在惯性传感器数据和景深信息辅助下，进行局部块匹配精配准；
- 6) 对精配准的相邻图像差分（如图 3.4），对差分后图像进行形态学滤波及与当前帧图像的空域分割图进行逻辑滤波，得到独立运动物体精确位置。

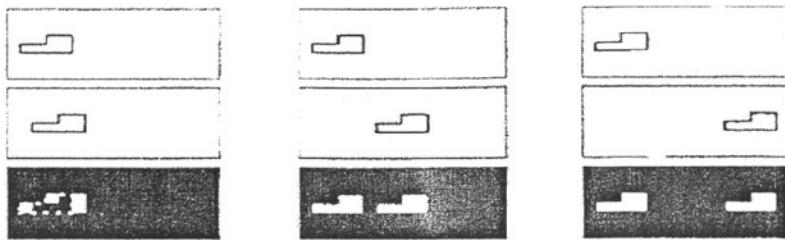


图 3.4 不同运动速度下的差分图

3.2.2 融合滤波器

惯性传感器能提供自主导航所需要的完备运动信息，但由于其精度和误差积累的问题，不能单独应用于自主定位与运动参数的获取，要使系统能够长期稳定的工作，必须要对惯性传感器数据进行修正。

本文中我们采用文献[10]中提出的状态表示与观测修正方程，在扩展卡尔曼滤波的体系下，将视觉传感信息与惯性传感信息融合。高采样率的惯性传感器数据在系统状态模型中以高频预测更新运动参数信息，精确配准后的图像观测信息在观测模型中以低频更新修正运动参数信息。

系统的状态向量表示为 $x_t = [p, v, q]^T$ ，其中 p 是运动平台位置参数， v 是平移速度， q 是方位姿态的四元数表示，非线性系统模型^{[10][11]}：

$$x_{k+1} = f(x_k, u_k, v_k), v_k \sim N(0, Q_k) \quad (3.5)$$

系统状态向量的动态特性为：

$$\left\{ \begin{array}{l} p_{k+1} = p_k + v \cdot \Delta t + (R(q) \cdot R_X \cdot a + g) \cdot \frac{\Delta t^2}{2} \\ v_{k+1} = v_k + (R(q) \cdot R_X \cdot a + g) \cdot \Delta t \\ \theta = w \cdot \Delta t \\ q_{k+1} = A(\theta) \cdot q_k \\ A(\theta) = \cos\left(\frac{\|\theta\|}{2}\right) I_4 + \sin\left(\frac{\|\theta\|}{2}\right) \Omega \end{array} \right. \quad (3.6)$$

其中 $u_t = [w, a]^T$ 为惯性传感器的观测数据， R_X 为运动平台坐标系与相机坐标系变换矩阵， Ω 为三个方向角速度构成的 4×4 旋转矩阵。

$$\Omega(\omega) = \begin{pmatrix} 0 & -\omega_x & -\omega_y & -\omega_z \\ \omega_x & 0 & \omega_z & -\omega_y \\ \omega_y & -\omega_z & 0 & \omega_x \\ \omega_z & \omega_y & -\omega_x & 0 \end{pmatrix} \quad (3.7)$$

背景块精确匹配后，很容易得到相邻两帧图像中的 2D 点对或特征对。假设一

点为 $[\xi_k, \psi_k]$ ，系统的观测方程可以表示为^[10]:

$$[-f\alpha, f, \dot{\xi}_k\alpha - \dot{\psi}_k]v_k = 0, \alpha = \frac{\dot{\psi}_{dis} - \dot{\psi}_{rot}}{\dot{\xi}_{dis} - \dot{\xi}_{rot}} \quad (3.8)$$

其中 $\dot{\xi}_{dis}, \dot{\psi}_{dis}$ 表示由于相机的平移所造成的特征速度， $\dot{\xi}_{rot}, \dot{\psi}_{rot}$ 表示由于相机的旋转所造成的特征速度， f 是成像的焦距。

3.2.3 数据同步

由视觉传感器获得数据必须与惯性传感器采集数据同步才能够提供正确的姿态运动参数估计，进而解算出相邻采样时刻间的平台运动分量，为独立运动目标检测提供平台运动信息的先验支撑。为获得可靠的同步性必须解决好两个问题。

多采样率下的传感器融合：视觉传感单元与惯性传感单元工作于不同的采用频率。为了能有效地对观测到的信息进行融合处理，在扩展卡尔曼滤波体系下在获得惯性观测数据时对状态方程进行修正，在获得图像特征点对时利用观测方程对系统状态进行修正，实现了不同采样率下的传感器数据的有效融合。

视觉传感数据滞后：由于对序列图像进行处理需要时间，这样获取的图像观测信息对于卡尔曼滤波传感器融合存在时间上的滞后。融合滤波器在每一个惯性传感数据采样时刻进行高速率的状态修正，没有图像观测信息的情况下惯性传感数据存入缓存中，当图像观测信息获得时，系统状态返回到视觉传感数据的时间标志点修正后，再利用缓存中的惯性传感数据对当前时刻的系统状态进行预测更新。数据同步处理如图 3.5:

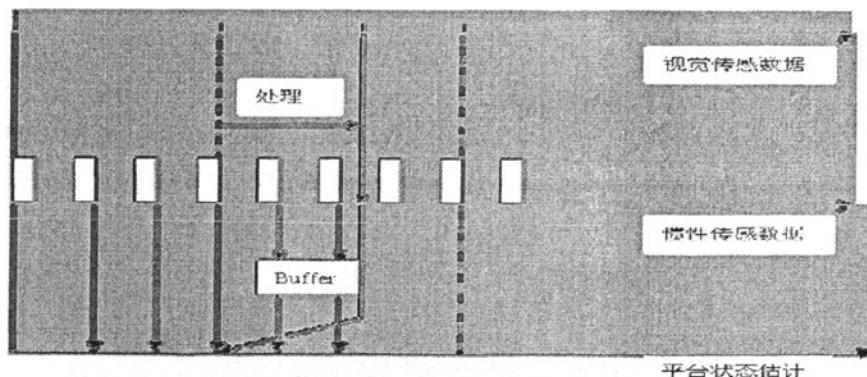


图 3.5 数据同步处理

3.3 实验与分析

本文的实验系统（图 3.6）中视觉传感器模块包括窄视角（5°）单目相机(CV-A10

CL From JAI Inc.)帧率为 50frame/sec，焦距为 55mm。惯性传感器包括一个微机电惯导测量单元(v3 From Spark Fun)，可以提供六个自由度（DOF）的运动参数信息，采样率为 200HZ，数据的传输采用蓝牙技术（Bluetooth Technology）。

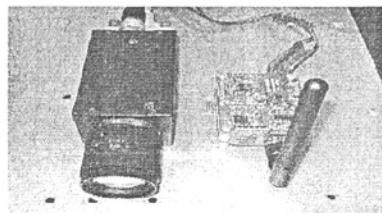


图 3.6 系统实验平台

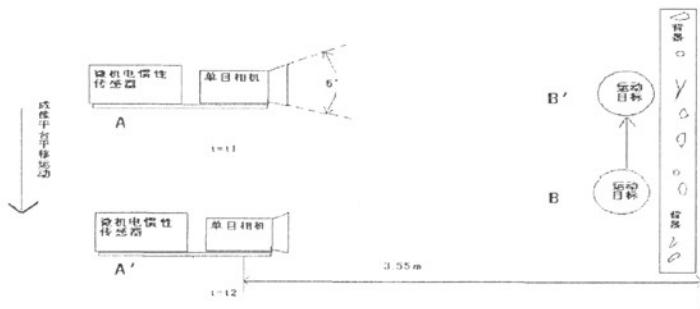
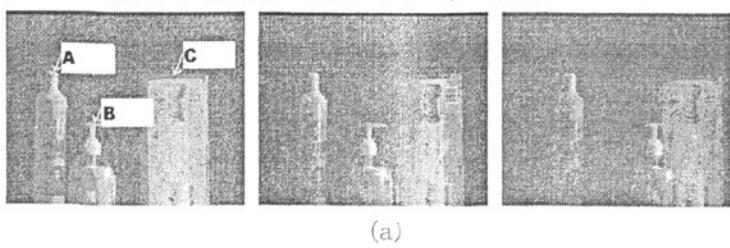


图 3.7 独立运动目标实验环境配置

图 3.7 为在室内利用单目视觉和微机电惯性传感器进行数据采集的实验环境配置：在惯性测量参考坐标系中 $t=t_1$ 时刻成像平台位于位置 A，独立运动目标位于位置 B； $t=t_2$ 时刻，成像平台运动到位置 A'，独立运动目标运动到位置 B'。在惯性测量参考坐标系中，背景静止不动，成像中心到背景的绝对景深为 3.55m，相机的视场角为 5°。

我们在室内进行实验数据的采集，考虑到相机为窄视角及室内场景深度与成像焦距的相对关系，我们采用小目标作为实验对象，分别在简单背景和复杂背景两种典型情况下进行了数据采集。实验中我们采用简化的平移运动参数模型对平台运动进行建模，忽略相邻采样间隔之间的旋转运动和场景深度变化，利用运动补偿法进行独立运动检测，其实验结果如图 3.8,3.9 所示：



(a)

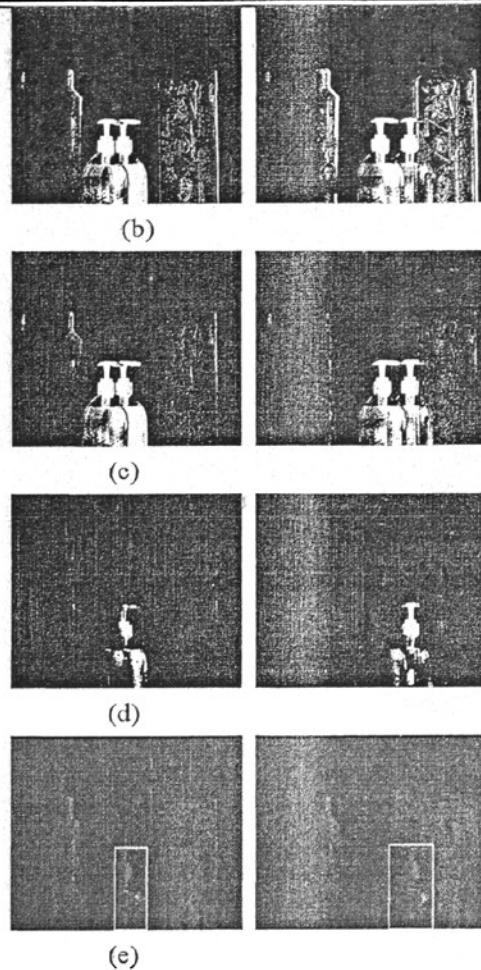


图 3.8 室内简单背景下的独立运动目标检测

图 3.8 为室内简单背景下独立运动目标检测结果。为其中(a)为原始序列图像中的等间隔的相邻三帧图像，为了便于表述在左图中对场景中的三个目标进行了标识，其中目标 A 和 C 为背景目标，目标 B 为独立运动目标。可以看出，由于成像平台的运动，图像中所有目标出现了位置上变化，然而由于独立运动目标位移量与背景目标的位移量不同，故三个目标间的相对位置发生了变化。(b)为利用惯性传感器的观测数据和室内测量的固定景深信息代入平移运动模型后，对相邻图像帧进行粗配准后帧间差法得到的差分图像，实验中我们考虑了垂直于成像光轴的二个方向的位移，其中一个方向的位移量绝对值较大，另一个方向的位移量绝对值较小。(c)为在粗配准差分结果图与原始图像帧的分割图进行逻辑滤波后，选择区域残差较小的图像块作为背景块进行相邻图像间的精确配准后，帧间差法得到的差分图像；(d)为(c)图像经过形态学和与原始图像帧的二值图分割图进行逻辑滤波后的结果图像；(e)为检测出的独立运动物体在原始图像中的标定。

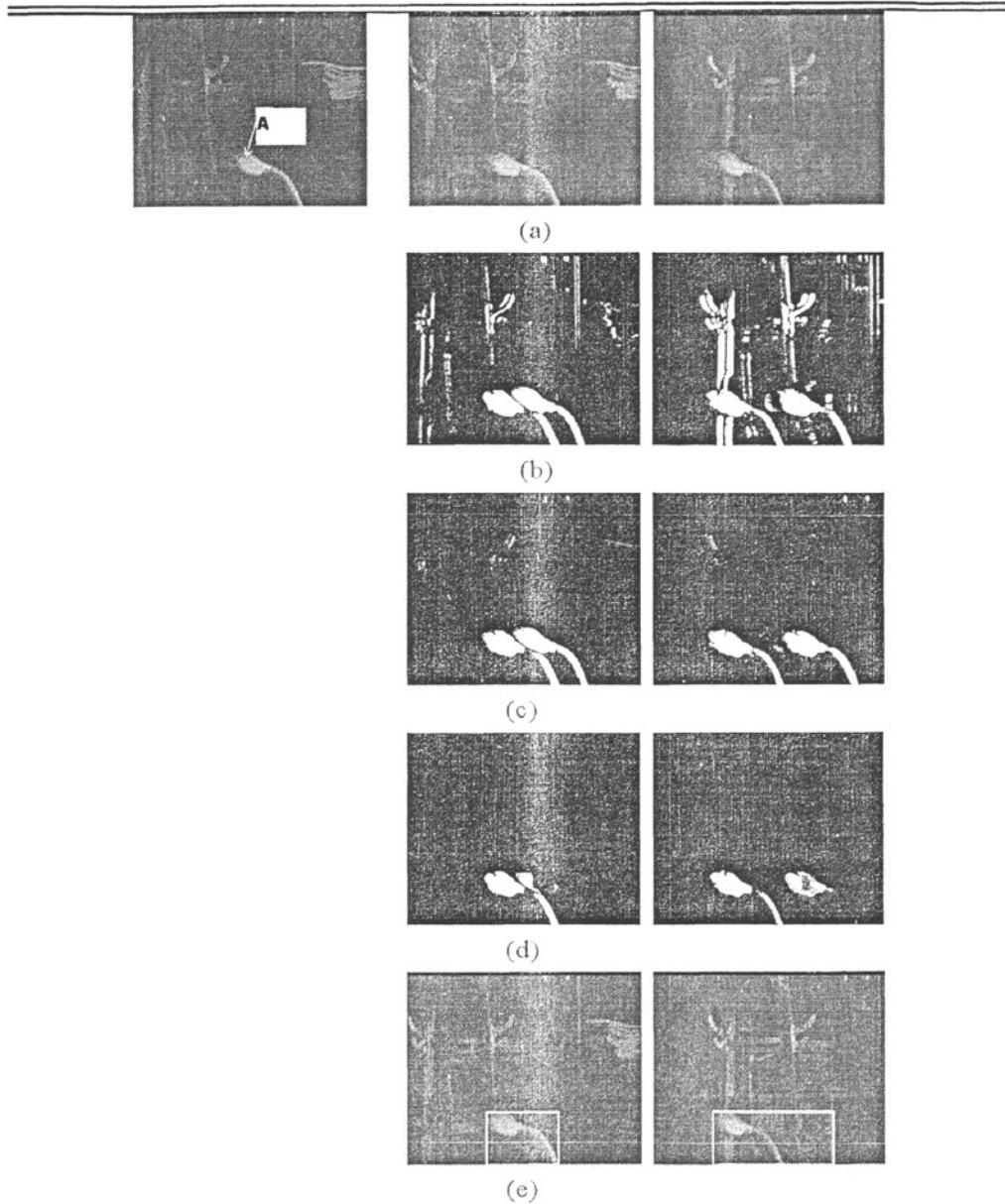


图 3.9 室内复杂背景下的独立运动目标检测

图 3.9 为室内复杂背景下独立运动目标检测结果, 其中(a)为原始序列图像中的等间隔的相邻三帧图像, 为了便于表述在左图中进行了标识, 其中目标 A 为独立运动目标, 其他部分为背景。可以看出, 由于成像平台的运动, 图像中所有目标出现了位置上变化, 然而由于独立运动目标位移量与背景目标的位移量不同, 故目标 A 与背景目标间的相对位置发生了变化。(b)为利用惯性传感器的观测数据和室内测量的固定景深信息代入平移运动模型后, 对相邻图像帧进行粗配准后帧间

差法得到的差分图像，实验中我们考虑了垂直于成像光轴的二个方向的位移，其中一个方向的位移量绝对值较大，另一个方向的位移量绝对值较小。(c)为在粗配准差分结果图与原始图像帧的分割图进行逻辑滤波后，选择区域残差较小的图像块作为背景块进行相邻图像间的精确配准后，帧间差法得到的差分图像；(d)为(c)图像经过形态学和与原始图像帧的二值图分割图进行逻辑滤波后的结果图像；(e)为检测出的独立运动物体在原始图像中的标定。

由两组实验结果看出，在室内简单背景和复杂背景下独立运动目标低速运动时检测结果较准确，在室内复杂背景下独立运动物体高速运动时由于差分断裂现象，保留了历史运动痕迹，检测出的运动区域比实际大。

3.4 本章小结

本章提出了基于运动补偿的融合惯性传感器和单目视觉传感器的独立运动目标检测的新方法，首先利用惯性传感器提供的平台运动参数信息和已知固定景深信息进行相邻帧图像粗配准，然后提取背景块进行局部搜索精配准，最后采用帧间差法检测独立运动目标，同时在扩展卡尔曼滤波体系下，利用图像观测信息修正惯性传感器数据，抑制了惯性传感器数据漂移。

本章的工作已经发表以下论文中。

孙浩,王程,王润生. 基于多传感器融合的运动平台运动目标检测[J]. 计算机应用 2008.4

第四章 基于光流的独立运动目标检测

4.1 引言

光流是空间运动物体被观测面上的像素点运动产生的瞬时速度场，包含了物体3D表面结构和动态行为的重要信息。一般情况下，光流由相机运动、场景中目标运动，或两者的运动产生。当场景中有独立的运动目标时，通过光流分析可以确定运动目标的数目、运动速度、目标距离和目标的表面结构^{[43][44]}。光流研究已经在环境建模、目标检测与跟踪、自动导航及视频事件分析中得到了广泛的应用。

4.1.1 问题描述

针对战场目标跟踪、识别和自主导航等多种应用场合下成像平台的旋转运动分量可以忽略的情形（如图4.1），本章提出了一种基于扩展点光流一致性约束的融合单目视觉和惯性传感器的独立运动目标检测方法，首先根据惯性传感器提供的平台运动参数信息和已知的成像参数信息预测扩展点在运动场中的位置，然后利用扩展点光流一致性约束对计算的稳健光流场进行一致性分类从而实现独立运动目标的检测。

图4.1中 T_1 为沿成像光轴(Z)的平移运动， T_2 为垂直于成像光轴(Y)方向的平移运动， T_3 为垂直于成像光轴(X)方向的平移运动。

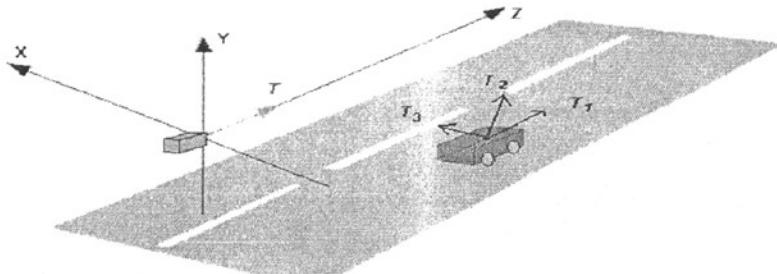


图4.1 典型的平移运动模型

4.1.2 光流基础

光流是指图像中模式运动的速度^{[65][66]}，光流场是二维的瞬时速度场，其中的2D速度场矢量是3D速度场在2D平面上的投影（图4.2）。光流不只包含了被观察物体的运动信息，而且携带着有关场景的3D结构的信息。当镜头与场景目标之间发生了相对运动的时候，所观察到的亮度模式运动即为光流。基于光流的序列

图像分析的直接目标是确定运动场，光流与运动场虽然有着密切的关系但是又不完全对应。场景中的目标运动导致图像中的亮度模式运动，而亮度模式的可见运动又产生了光流。在多数情况下，光流与运动场存在着相互对应的关系，可以根据光流与运动场的相互关系由图像的变化来估计相对运动。

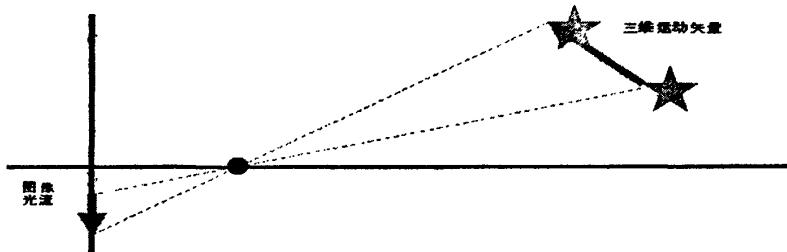


图 4.2 光流的形成

在亮度不变假设下，序列图像间的相对运动成为光流。亮度不变假设是指在相邻图像中运动像素的亮度值保持不变，即

$$I(x, y, t) = I(x + u\delta t, y + v\delta t, t + \delta t) \quad (4.1)$$

其中 $I(x, y, t)$ 为像素点 (x, y) 在 t 时刻的亮度值，所有像素点的速度矢量构成的矢量场 $\bar{V} = (u, v)$ ，称为光流场。令 $\delta x = u\delta t$ 、 $\delta y = v\delta t$ ，如果亮度随 x 、 y 和 t 平滑变化，可把上式按泰勒级数展开并丢弃高阶项得到：

$$\delta x \frac{\partial I}{\partial x} + \delta y \frac{\partial I}{\partial y} + \delta t \frac{\partial I}{\partial t} = 0 \quad (4.2)$$

用 δt 除等式两端和取 $\delta t \rightarrow 0$ 的极限后可求得：

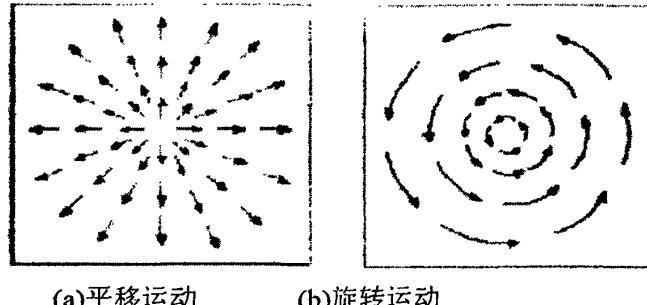
$$\frac{\partial I}{\partial x} \frac{dx}{dt} + \frac{\partial I}{\partial y} \frac{dy}{dt} + \frac{\partial I}{\partial t} = 0 \quad (4.3)$$

$$\text{令 } u = \frac{dx}{dt}, v = \frac{dy}{dt}, I_x = \frac{\partial I}{\partial x}, I_y = \frac{\partial I}{\partial y}, I_t = \frac{\partial I}{\partial t} \text{ 得}$$

$$I_x u + I_y v + I_t = 0 \quad (4.4)$$

上述方程即为光流亮度约束方程。

当 $I_x, I_y \rightarrow 0$ 时，可以看出无法从约束方程中得到关于 u 和 v 的信息，即在图像灰度是常数或变化缓慢的地方，像素的运动不能精确地确定。约束方程只对速度向量 (u, v) 给出一个约束，如果要根据它来估计运动，还必须施加其他的约束。一种方法是假设速度从一个像素到另一个像素的变化是平滑的，因此可以用相邻像素满足约束方程求解最小平方解。速度变化平滑的假设在物体内部是合理的，但在物体边界处是不正确的；另一种方法利用约束方程估计物体拐角的两个边缘的速度，拐角的两个边缘分别提供光流在两个方向上的分量信息，所以在拐角处的速度可以完全地确定。图 4.3 为典型运动模式下的图像光流图。



(a)平移运动 (b)旋转运动

图 4.3 典型运动模式下的光流图

文献中存在大量的光流计算方法^{[63][66][67][68][69][70][71]}，主要可以分为以下几类：

- 基于微分的方法

微分法^{[66][68]}，又称时空梯度法，假定图像在空间和时间上是连续的(可差分的)，利用图像强度的时空倒数来计算每一像素点的速度矢量。根据光流方程求解光流场的方法又可分为全局法和局部法。全局法利用附加的全局约束计算大面积图像区域中的稠密光流。局部法利用局部邻域中的正常速度信息进行最小平方优化求得最佳速度矢量。一般根据速度估计的邻域大小来确定所用技术是全局法还是局部法。表现为不连续光流的遮挡问题可以通过混合速度分布或参数模型进行分析。这些技术将光流划分成各种独立运动的对象区域或表面。分层估计框架配合图像变换方法可用于分析更大的 2D 运动。微分法还衍生出了其他算法，例如近年来出现的基于张量 (tensor) 的光流算法，其性能较已有的算法有较大的提高。一个像素在时空域中的运动速度可以用一个向量表示，而张量实际上是一种向量的表示方式。基于张量的光流算法大致可以分为三步：1) 逐点估计张量；2) 将图像分割为若干区域，在各个区域内计算运动模型参数；3) 利用运动模型参数，计算速度场。

- 基于匹配的方法

匹配法^{[69][70]}，在相邻图像中查找像素(块)之间的对应关系，两幅图像之间对应像素(块)之间的位移即是所求的光流。当图像信噪比很差时微分法得到的计算结果可能不正确，这时匹配法效果可能更好。基于匹配的光流计算方法包括基于特征和区域的两种。基于特征的方法不断地对目标主要特征进行定位和跟踪，对目标大的运动和亮度变化具有鲁棒性。存在的问题是所得到的光流通常较稀疏，而且特征提取和精确匹配也十分困难。基于区域的方法先对类似的区域进行定位，然后通过相似区域的位移计算光流。这种方法在视频编码中得到了广泛的应用。然而，它计算的光流仍不稠密。此外，基于匹配的方法很难得到亚像素精度的光流，计算量大。

- 基于频域的方法

能量法^{[63][64]},又称基于频域的方法,基于速度调谐滤波器的输出能量计算光流。能量方法是基于生物视觉系统而提出的运动感知模型,有一定的生物视觉合理性。能量法主要有三个优点:(1)运动可以在频域中进行分析,有些运算在频域中要比时域中容易处理;(2)可以分析运动的变化规律,如速度、加速度等,从而更好地利用运动的动态特性;(3)系统地使用多幅图像,可以提高速度和深度估计的稳健性和精度。在基于能量的模型中,首先要对输入图像序列进行时空滤波处理,这是一种时间和空间整合,对于均匀的流场,要获得正确的速度估计,这种时空整合是非常必要的。然而,这样做会降低光流估计的空间和时间分辨率。此外,基于能量的光流技术还存在高计算量问题。

● 基于相位的方法

基于相位的方法^{[65][66]}因速度是根据带通滤波器输出的相位特性确定的而称为相位方法。基于相位的光流技术的综合性能比较好,速度估计比较精确且具有较高的空间分辨率,对图像序列的适用范围也比较宽。与基于能量的光流技术一样,基于相位的模型既有一定的生物视觉合理性,又有较高的计算复杂性。

● 基于小波的方法

小波法^{[67][68]},利用小波变换灵活的多尺度框架,构造金字塔式多分辨率结构通过分层技术计算光流。分层技术通过对原图像多次下采样,得到自上而下按比例缩小的一系列图像层,这样原始图像中的大距离运动就相当于采样后低层图像中的小距离运动,从而可以计算光流值很大的光流场。

● 基于神经动力学的方法

神经动力学方法^{[69][70][71]},利用视觉运动感知神经动力学模型描述运动感知中视皮层简单细胞、复杂细胞、超复杂细胞以及视网膜双极细胞之间的相互作用,揭示了运动分割与组合、竞争与合作的神经整合机制。这个模型对于整体运动方向的判别非常有效,但不能给出运动速度的大小。Fay等人模仿视网膜中的时空处理和大脑的视觉运动通路,基于并联动力学提出了一个多层神经网络,涉及光适应、边缘增强和边缘速度提取等几个处理阶段。但这个神经网络仅能提供运动边缘的法向速度估计,为了恢复整个模式的光流场,还必须用速度泛函方法将估计的法向流整合成一个致密的光流场。

4.1.3 扩展点光流一致性约束

相机三维平移运动矢量与投影平面的交点称为扩展点(Focus Of Expansion-FOE)或收缩点(Focus Of Contraction-FOC)【取决于相机的运动方向】^[15]。如图4.5所示,扩展点为OL轴在成像平面上的交会点。扩展点在相机运动分析中有着很重要的作用。扩展点一条特有的属性就是在相机和场景只存在

平移运动时，图像平面上所有的运动矢量都相交于扩展点如图 4.4。

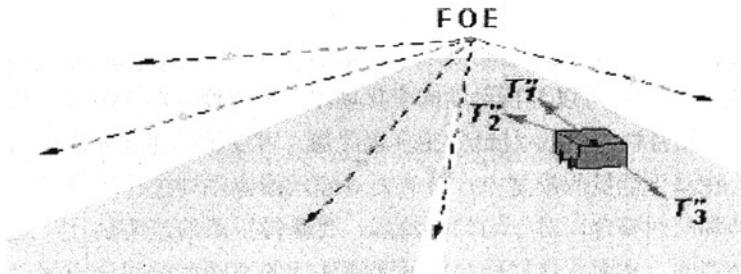


图 4.4 扩展点性质

假设扩展点在图像坐标系中的位置为 (x_0, y_0) ，则

$$\frac{x_0}{T_x} = \frac{y_0}{T_y} = \frac{f}{T_z} \quad (4.5)$$

其中 f 为相机的成像焦距，即

$$x_0 = f \frac{T_x}{T_z}, y_0 = f \frac{T_y}{T_z} \quad (4.6)$$

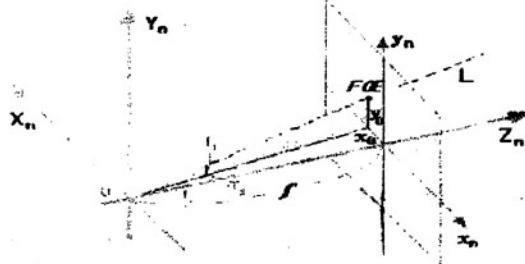


图 4.5 扩展点位置

假设空间点 P 在图像中的投影为 $p(x_n, y_n)$ ，如图 4.6 所示，根据透视投影模型其在图像坐标系中的坐标为：

$$x_n = f \frac{X_n}{Z_n}, y_n = f \frac{Y_n}{Z_n} \quad (4.7)$$

p 在图像中的瞬时运动速度 (u, v) 可由上式的时域微分得到：

$$\begin{aligned} u &= \dot{x}_n = \frac{d}{dt} \left(f \frac{X_n}{Z_n} \right) = f \left(\frac{dX_n}{Z_n} - X_n \frac{dZ_n}{Z_n^2} \right) \\ v &= \dot{y}_n = \frac{d}{dt} \left(f \frac{Y_n}{Z_n} \right) = f \left(\frac{dY_n}{Z_n} - Y_n \frac{dZ_n}{Z_n^2} \right) \end{aligned} \quad (4.8)$$

如果相邻图像间采样时间足够短，下列等式可用来近似点 p 的位移：

$$\left\{ \begin{array}{l} x_{n+1} - x_n = uT = f\left(\frac{dX_n}{Z_n}T - X_n \frac{dZ_n}{Z_n^2}T\right) = f\left(\frac{\Delta X_n}{Z_n} - X_n \frac{\Delta Z_n}{Z_n^2}\right) \\ y_{n+1} - y_n = vT = f\left(\frac{dY_n}{Z_n}T - Y_n \frac{dZ_n}{Z_n^2}T\right) = f\left(\frac{\Delta Y_n}{Z_n} - Y_n \frac{\Delta Z_n}{Z_n^2}\right) \end{array} \right\} \quad (4.9)$$

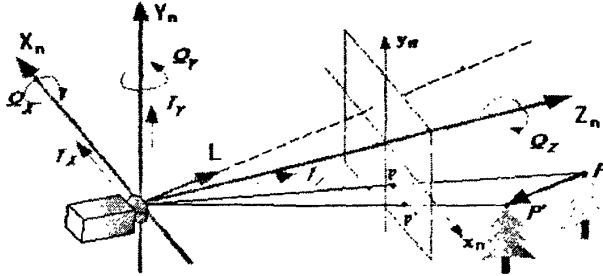


图 4.6 运动成像分析

假设相邻采样图像间，相机与场景点相对运动参数为 $(T_x, T_y, T_z), (\Omega_x, \Omega_y, \Omega_z)$

则有

$$\left\{ \begin{array}{l} \Delta X_n = X_{n+1} - X_n = \Omega_z Y_n - \Omega_y Z_n - T_x \\ \Delta Y_n = Y_{n+1} - Y_n = \Omega_x Z_n - \Omega_z X_n - T_y \\ \Delta Z_n = Z_{n+1} - Z_n = \Omega_y X_n - \Omega_x Y_n - T_z \end{array} \right\} \quad (4.10)$$

即

$$\begin{aligned} x_{n+1} - x_n &= \frac{x_n y_n}{f} \Omega_x - \frac{f^2 + x_n^2}{f} \Omega_y + y_n \Omega_z + \frac{T_z}{Z_n} (x_n - x_0) \\ y_{n+1} - y_n &= \frac{f^2 + y_n^2}{f} \Omega_x - \frac{x_n y_n}{f} \Omega_y - x_n \Omega_z + \frac{T_z}{Z_n} (y_n - y_0) \end{aligned} \quad (4.11)$$

假设 $T_z/Z \neq 0$ ，则有

$$\frac{(x_{n+1} - x_n) - \left(\frac{x_n y_n}{f} \Omega_x - \frac{f^2 + x_n^2}{f} \Omega_y + y_n \Omega_z \right)}{(y_{n+1} - y_n) - \left(\frac{f^2 + y_n^2}{f} \Omega_x - \frac{x_n y_n}{f} \Omega_y - x_n \Omega_z \right)} = \frac{x_n - x_0}{y_n - y_0} \quad (4.12)$$

由相机运动所引起的场景图像投影点运动矢量必须满足上式，文中成为扩展点光流约束方程。当相邻采样时刻间相机旋转运动分量可忽略的情况下时，约束方程可简化为：

$$\frac{x_{n+1} - x_n}{x_{n+1} - x_n} = \frac{x_n - x_0}{y_n - y_0} \quad (4.13)$$

4.2 光流法独立运动目标检测

针对运动成像平台不存在旋转运动分量或旋转运动分量可以忽略的情形，我们提出了基于扩展点光流一致性约束的独立运动目标检测方法，设计实验思路，并对 VISAT 多传感器测绘平台实测的单目图像序列和惯性传感器数据进行了实验，其算法流程图如图 4.7。

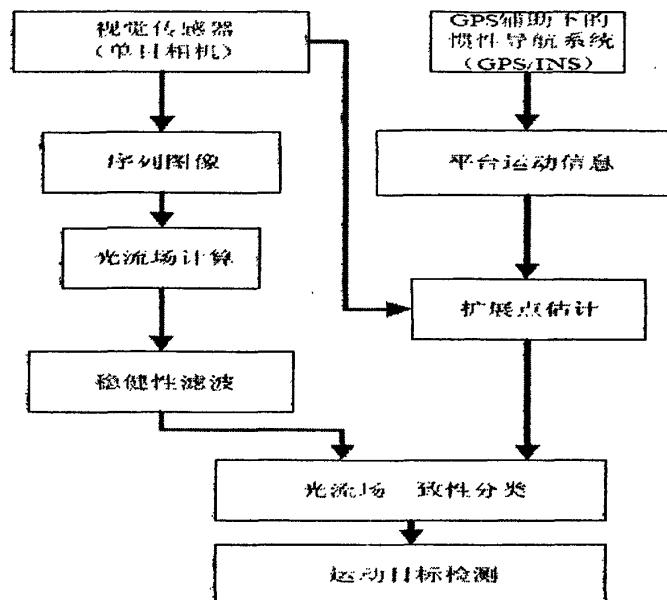


图 4.7 光流法检测流程图

4.2.1 光流场计算

基于特征匹配的光流计算方法检测图像中的显著特征采用匹配跟踪的策略来计算光流矢量，一般情况下，产生的光流场是稀疏光流场，计算量比基于微分方法小。基于特征匹配的方法可以消除孔径(Aperture)效应，对目标的运动限制较小，可以处理相邻图像间大位移，对噪声的敏感性降低，基于以上优点，本文采用基于特征匹配跟踪的方法来计算光流场。

角点特征是影像的重要特征，具有旋转不变性和不随光照条件变化的优点，在图像匹配、目标描述与识别以及运动估计、目标跟踪等领域，具有十分重要的意义。常用的角点提取算子有 Harris 算子，Moravec 算子，Forstner 算子，SUSAN 算子，Plessy 算子等，实验中我们对比发现 Harris 算子在处理时间和提取效果都优于其他算子^[53]。Harris 算子使用图像的一阶差分，计算每个像素处的平均平方梯度矩阵(Average square gradient matrix)，通过特征值分析给出角点响应，可以隐式地使用滑动窗口，适合矩阵运算。

获得图像中的角点特征后，采用金字塔迭代 Lucas-Kanade 算法计算角点处的光流矢量^[55]，得到稀疏光流场。

4.2.2 稳健性滤波

光流计算过程中由于噪声、区域灰度特性相似等多因素的影响，光流矢量计算结果可能存在误差，本文中采用基于多帧图像特征点共现约束、区域一致性约束和帧间最大运动幅值约束的稳健性滤波对估计的光流矢量进行处理。

- **多帧共现性约束：**只有在当前图像帧中被检测为特征点且在相邻两帧内都跟踪到匹配点的角点才被列为是需要处理的兴趣点。考虑到相机的高采样率，相邻帧的视野变化比较小，独立运动目标的运动投影点出现在多幅相邻图像帧中。此处，我们不关心由于小的视野变化而导致的在一帧图像中出现而在相邻的前向或后向相邻图像帧中没有出现的体积较小的独立运动目标。图 4.8(a) 为对 VISAT 实测序列图像中相邻三帧图像 ($t-1$ 时刻, t 时刻, $t+1$ 时刻) 计算得到的双向光流场在 t 时刻图像帧中的表示，图中 $t-1$ 时刻到 t 时刻的光流矢量标识为红色， t 时刻到 $t+1$ 时刻的光流矢量标识为绿色。图 4.8(b) 为对双向光流场进行三帧共现性约束滤波后的结果图。可以看到一些大位移的误匹配运动矢量、由于相机视野变化引起的进入和离开视野的背景运动矢量被滤除。

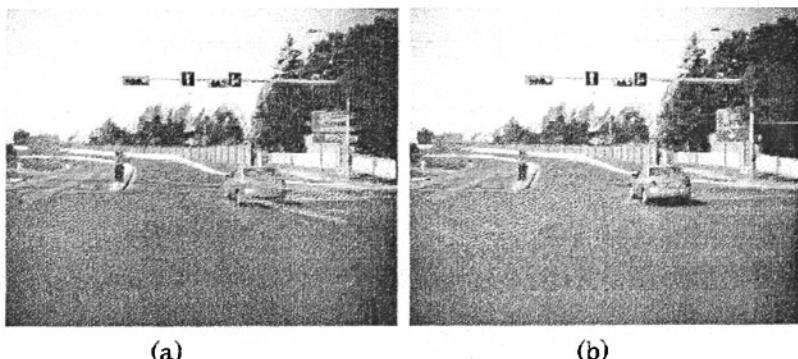


图 4.8 三帧共现性约束滤波

- **区域一致性约束：**由于场景中大部分的运动物体都可以视为刚体，独立运动目标检测的目的是确定出具有不同于背景区域运动属性的区域，属于同一区域的角点特征应具有相似的动态特性，在光流场中表现为属于同一区域的角点特征应具有相似的运动位移和运动方向，利用直方图统计可以滤除一些不具有区域特性的运动矢量。

区域一致性约束滤波的步骤为：

- (1) 计算出序列图像中相邻三帧 ($t-1$ 时刻, t 时刻, $t+1$ 时刻) 图像的光流场，

并对其进行三帧共现性约束滤波；

(2) 计算每个角点特征处的 t 时刻到 $t-1$ 时刻的后向光流矢量， t 时刻到 $t+1$ 时刻的前向光流矢量，对角点特征的双向光流矢量的幅度比值统计，滤除统计值较小的角点特征。

(3) 计算每个角点特征处的 t 时刻到 $t-1$ 时刻的后向光流矢量， t 时刻到 $t+1$ 时刻的前向光流矢量，对角点特征的双向光流矢量的夹角进行统计，滤除统计值较小的角点特征。

图 4.9 (a) 为 t 时刻图像角点特征的双向光流矢量幅值比值的统计直方图，(b) 为滤除双向光流矢量幅值比值统计量较小的角点特征后，统计的幅值比值直方图。

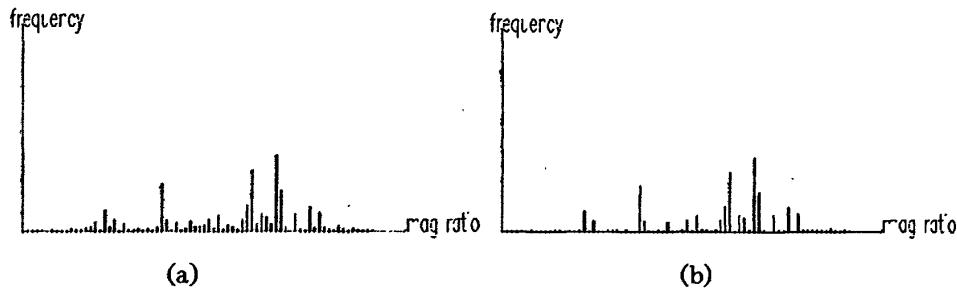


图 4.9 幅值比值统计滤波

- 运动幅值最大值抑制：考虑到相机系统的高采样率，认为相邻帧间运动位移较小，如果某角点光流矢量幅度大于给定门限值，滤除该光流矢量。

4.2.3 一致性分类

由 GPS 辅助下的惯性导航系统获得的平台运动信息和校正后的相机参数信息，根据扩展点的定义，通过计算和坐标变换，可以估计出其在图像中的位置。假设平台在相邻帧间的旋转运动可以忽略，利用简化后的约束方程可以对稀疏光流场进行分类。光流场中既有由于运动平台运动的投影运动矢量又有独立运动物体的投影运动矢量，如图 4.10 所示。

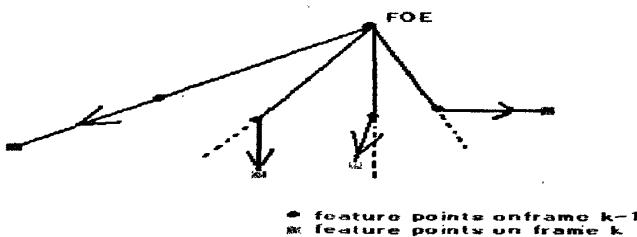


图 4.10 运动场中的扩展点与光流

为了减小检测结果对噪声、不精确扩展点估计和旋转分量的敏感度，分类规

则如下：假设第 k 帧特征点为 (x_k, y_k) ，其在 $k+1$ 帧对应的特征点为 $(x_{(k+1)i}, y_{(k+1)i})$ ，扩展点位置为 (x_0, y_0)

$$\begin{aligned} k_i &= (y_{(k+1)i} - y_k) / (x_{(k+1)i} - x_k) \\ k_B &= (y_k - y_0) / (x_k - x_0) \end{aligned} \quad (4.14)$$

如果 $\{k_i \notin [k_B(1-\lambda), k_B(1+\lambda)]\}$ ，则判决为独立运动矢量，否则归类为背景运动矢量。其中 λ 为调适因子，取值范围为 $(0,1)$ 。

4.3 实验与分析

实验中我们采用 VISAT™ 多传感器移动测绘实测数据进行基于光流的独立运动目标检测算法的验证。实验中用到的 VISAT™ 数据是王程副教授在加拿大卡尔加里大学做访问学者期间合作研究的数据。VISAT™^[46,47] 是加拿大卡尔加里大学研制的最早的陆地移动测绘系统 (Mobile Mapping System)。图 4.11 为 VISAT™ 运载平台。



图 4.11 实验数据 VISAT™ 测绘平台

VISAT™ 系统的数据获取部分主要包括捷联惯性导航系统 (SINS)，双频 GPS 接收机，6-12 个彩色数字相机和集成的速度控制单元。

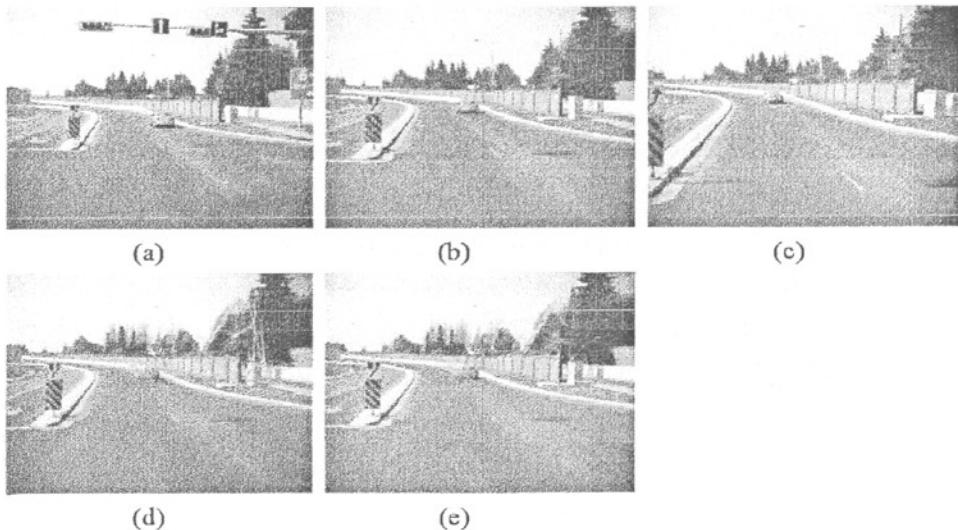


图 4.12 光流场稳健性滤波

图 4.12 为光流场稳健性滤波的实验结果，为了较清楚的显示结果，光流场结果图显示选用较大的尺度。(a)(b)(c)为相邻的三帧图像($t-1$ 时刻, t 时刻, $t+1$ 时刻)，(d)为采用角点特征跟踪方法计算的双向光流场图 (t 时刻到 $t-1$ 的后向光流场用红色标识, t 时刻到 $t+1$ 时刻的前向光流场用绿色标识, (e)为采用三个约束进行稳健性滤波的最终双向光流场结果图。可以看出，一些位移量较大的错误匹配和不满足区域一致性的光流矢量被滤除。

图 4.13, 4.14 为不同场景下的独立运动目标检测实验结果。为了较清楚的显示结果，光流场分类结果图显示选用较大的尺度。

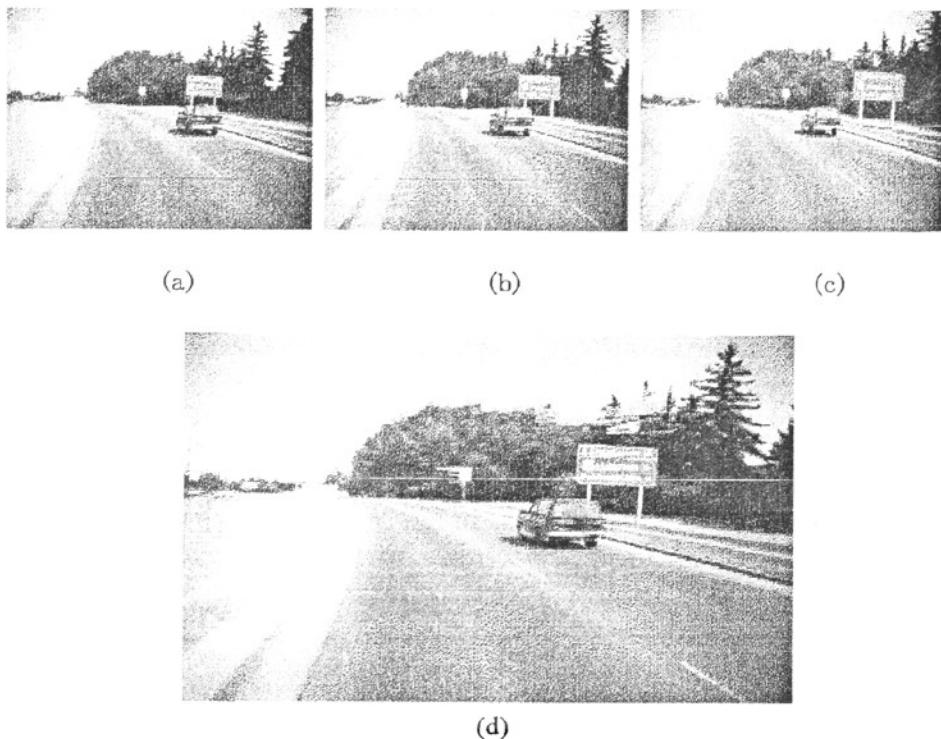


图 4.13 独立运动目标检测

(a)(b)(c)为原始序列图像, (d)为最终的光流场分类结果图，其中背景部分运动矢量用绿色标识，独立运动目标运动矢量用红色标识。可以看出，远场景中的护栏和树木的部分运动矢量由于噪声影响，误分类为独立运动目标运动矢量，近场景中的独立运动目标运动矢量正确分类。

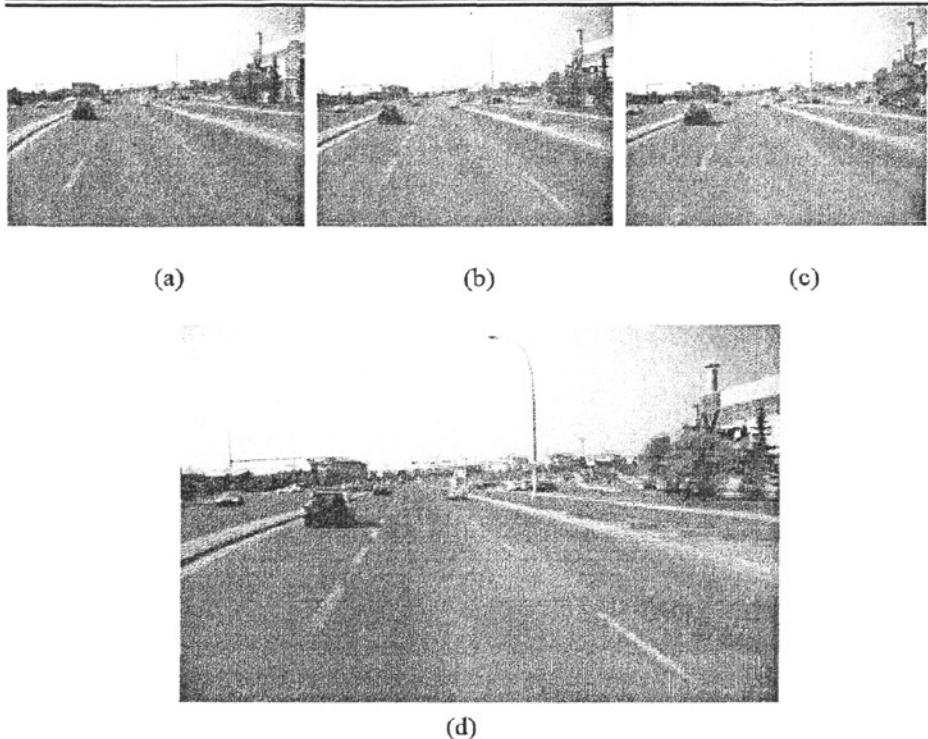


图 4.14 独立运动目标检测

(a)(b)(c)为原始序列图像, (d)为最终的光流场分类结果图, 其中背景部分运动矢量用绿色标识, 独立运动目标运动矢量用红色标识。远场景中位于视野边缘的沿相反方向行进的车辆的运动矢量由于三帧共现性约束被滤除, 背景中的路灯底座和路旁的草地中的运动矢量由于噪声的影响出现了误分类。

实验结果表明, 使用本文提出的基于光流扩展点一致性约束的方法可有效的检测实际运动场景中的独立运动目标, 但同时我们在实验中发现当运动目标较小或运动目标上可检测跟踪的特征点较少时, 可能会发生漏警。

4.4 本章小结

本章在介绍光流和动态场景中扩展点光流一致性约束的基础上, 针对成像平台旋转运动分量可以忽略的情形, 提出了一种新的独立运动目标检测方案, 首先采用基于角点跟踪的方法计算稀疏光流场, 结合运动矢量时空约束采用稳健性滤波得到稳定的光流矢量, 然后利用 GPS 辅助下的惯性导航系统获得成像平台相邻采样时刻间的运动参数信息对扩展点位置进行预测, 最后以扩展点一致性约束为准则得到分类后的光流场, 实现了多种场景下的独立运动目标检测。

本章的工作已经发表在以下论文中。

Hao Sun, Cheng Wang, Naser El-sheimy. Moving Objects Detection for Mobile mapping[C].IEEE Conference on Information Acquisition and Automation, 2008.

第五章 基于粒子滤波的独立运动目标检测

5.1 引言

随着工业摄像技术的发展和视觉传感器的不断成熟，廉价的多目立体视觉系统已越来越多应用到了计算机视觉的各个领域，基于立体视觉的独立运动目标检测也受到了广泛的关注。

5.1.1 问题描述

基于单目视觉传感器的独立运动目标检测由于缺乏场景结构信息，不能很好的解决复杂场景中多运动模式情形下的多个独立运动目标检测问题。

针对无成像平台运动模式限制复杂环境下的多个独立运动目标检测的情形，我们提出一种基于状态粒子数可变的时域空域双重重采样的粒子滤波的融合双目立体视觉和惯性传感器的独立运动目标检测的新方法。

基于粒子滤波的独立运动目标检测方法首先采用 SIFT 特征匹配进行立体配准和实际运动场的提取，然后在状态粒子数可变的时域空域双重重采样的粒子滤波体系下融合解算出的立体景深信息、惯性传感器平台运动参数信息及运动场信息，最后进行 SIFT 特征点的空域聚类，实现多个独立运动目标的正确检测。

5.1.2 立体视觉基础

立体成像和显示系统是模拟人类的立体视觉设计的，其基本原理是用不同位置的摄像机采集三维场景，进行配准后解算从而还原出三维场景信息。目前常见的双目视觉系统有平行结构和会聚结构两种，如图 5.1(a)、5.2。平行结构使用两个平行放置的摄像机，它们的光轴平行；会聚结构中两个摄像机的摆放有一定的夹角，光轴在远处相交。VISATTM系统的立体视觉系统结构为多目，如图 5.1(b)，本文中我们采用其前向的会聚结构双目视觉系统数据进行实验。

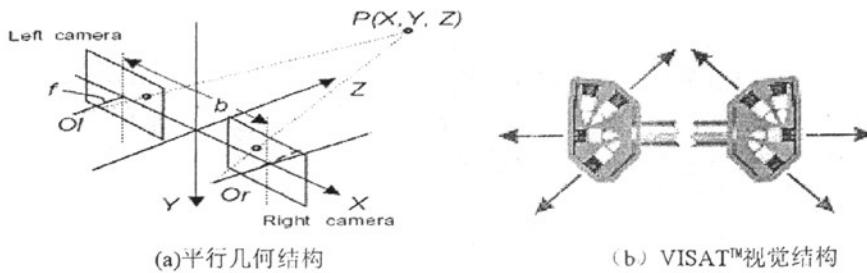


图 5.1 立体视觉几何结构

图 5.2 所示为在理想的针孔成像模型基础上考虑了透镜的径向畸变因素的会聚式双目立体视觉系统模型。图中 $A(X_w, Y_w, Z_w)$ 是三维世界坐标系 $O_w X_w Y_w Z_w$ 中成像点； $O_l x_l y_l z_l$ 和 $O_r x_r y_r z_r$ 分别为左右摄像机的坐标系；坐标中心点 O_l 和 O_r 分别为左右摄像机的光学中心； z_l 和 z_r 轴分别与左右摄像机光轴重合； $O_u X_l Y_l$ 和 $O_u X_r Y_r$ 分别是左右摄像机成像平面的坐标系，图像中心 O_u 为光轴 z_l 与左摄像机图像平面的交点，图像中心 O_{ur} 为光轴 z_r 与右摄像机成像平面的交点； X_l 和 Y_l 轴分别与左摄像机坐标系的 x_l 和 y_l 轴平行， X_r 和 Y_r 轴分别与右摄像机坐标系的 x_r 和 y_r 轴平行； $o_l u_l v_l$ 和 $o_r u_r v_r$ 分别是左右摄像机的成像在计算机中的图像坐标系。 (X_u, Y_u) 是 A 点在理想针孔成像模型下的图像坐标， (X_d, Y_d) 是 A 点的实际图像坐标，因透镜的径向变形而偏离了其理想图像坐标位置。有效焦距 f 是光学中心到图像平面的距离。

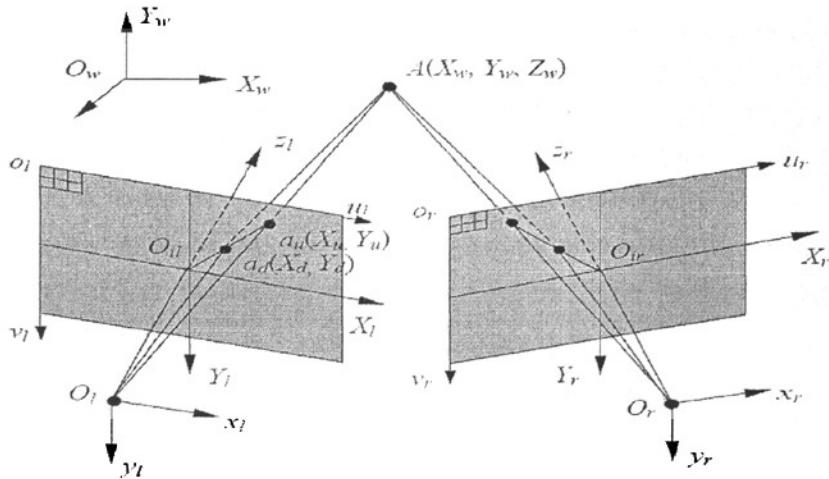


图 5.2 会聚式立体视觉模型

以左摄像机的成像过程为例，将物体点的三维世界坐标 (X_w, Y_w, Z_w) 变换到计算机图像平面坐标 (u, v) 的步骤如下：

(1) 物体三维世界坐标 (X_w, Y_w, Z_w) 到摄像机三维坐标 (x, y, z) 的变换：

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = R \cdot \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} + T \quad (5.1)$$

其中 R 为 3×3 的旋转变换矩阵， T 为 3×1 的平移矢量。 R 确定了摄像机相对于世界坐标系的方向，它可由三个欧拉角 γ, β, α 来表示。 $R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}$ 其中

$$\begin{aligned}
 r_{11} &= \cos \alpha \cos \beta \\
 r_{21} &= -\sin \alpha \cos \beta \\
 r_{31} &= \sin \beta \\
 r_{12} &= \sin \alpha \cos \gamma + \cos \alpha \sin \beta \sin \gamma \\
 r_{22} &= \cos \alpha \cos \gamma - \sin \alpha \sin \beta \sin \gamma \\
 r_{32} &= -\cos \beta \sin \gamma \\
 r_{13} &= \sin \alpha \sin \gamma - \cos \alpha \sin \beta \cos \gamma \\
 r_{23} &= \cos \alpha \sin \gamma + \sin \alpha \sin \beta \cos \gamma \\
 r_{33} &= \cos \beta \cos \gamma
 \end{aligned}$$

平移矢量 $\mathbf{T} = [T_x \ T_y \ T_z]^T$ 表示了摄像机相对于世界坐标系的位置。

(2) 摑像机三维坐标 (x, y, z) 在理想的针孔成像模型下经投影变换至图像坐标系

$$\begin{aligned}
 (X_u, Y_u) : \quad X_u &= f \frac{x}{z} \\
 Y_u &= f \frac{y}{z}
 \end{aligned} \tag{5.2}$$

(3) 虑摄像机透镜的径向畸变因素，将理想图像坐标 (X_u, Y_u) 转换成实际图像坐标

$$\begin{aligned}
 (X_d, Y_d) : \quad X_d &= X_u (1 + kr^2)^{-1} \\
 Y_d &= Y_u (1 + kr^2)^{-1}
 \end{aligned} \tag{5.3}$$

其中 $r = \sqrt{X_d^2 + Y_d^2}$ ，表示图像中心到实际图像坐标 (X_d, Y_d) 的距离； k 表示透镜径向畸变系数。

(4) 实际图像坐标 (X_d, Y_d) 到计算机图像坐标 (u, v) 的变换：

$$\begin{aligned}
 u &= X_d N_x + u_0 \\
 v &= Y_d N_y + v_0
 \end{aligned} \tag{5.4}$$

(u_0, v_0) 表示图像中心的计算机图像坐标 (Pixel)， N_x, N_y 表示图像平面上单位距离内所包含的像素数。

为了检测物体的三维运动，常常需要知道特征点的三维空间坐标，在双目立体视觉系统中，在已知相机空间几何相对关系时，利用三角测量可以获得目标特征点的三维空间坐标，其原理如图 5.3：

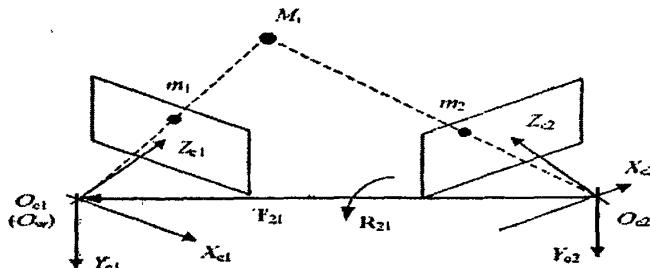


图 5.3 三角解算原理

假设已知相机间的旋转矩阵 R_{21} 和平移矢量 T_{21} , 定义左相机坐标系为世界坐标系, 空间点 M 的三维坐标可通过对应的图像点对 (x_1, y_1) , (x_2, y_2) 获得, 其表示为

$$\left\{ \begin{array}{l} x = zx_1 / f_1 \\ y = zy_1 / f_1 \\ z = \frac{f_1(f_2t_x - x_2t_z)}{x_2(r_7x_1 + r_8y_1 + f_1r_9) - f_2(r_1x_1 + r_2y_1 + f_1r_3)} \\ = \frac{f_1(f_2t_y - y_2t_z)}{y_2(r_7x_1 + r_8y_1 + f_1r_9) - f_2(r_4x_1 + r_5y_1 + f_1r_6)} \end{array} \right\} \quad (5.5)$$

5.1.3 宽基线特征匹配

立体视觉宽基线条条件下的特征点提取与稳定匹配^{[77][78]}是三维重建, 运动解算等多种高层计算机视觉任务的前提。宽基线条条件下的特征应该具有较高的稳定性和可分辨性。稳定性是指对旋转、尺度缩放、仿射变换、视角变化、光照变化等图像变换因素保持一定的不变性, 而对物体运动、遮挡、噪声等因素也保持较好的可匹配性, 从而实现差异较大的两幅图像之间特征的匹配。可分辨性是指不同的特征对应的描述算子不同。

目前, 基于局部特征分析的方法已成为国际上的研究热点, 在目标匹配与识别的实际应用中获得了很大成功。近年来, 局部特征分析方法获得了很大发展, 人们相继提出了多种有效的局部特征算子, 这些算子一般都具有较高的稳定性和较强的匹配能力。通常情况下, 局部特征分析方法分为特征提取 (detector) 和特征表达 (descriptor) 两步。局部特征分析方法大都致力于保持特征描述算子对图像变化的不变性上, 其基本思想是在目标图像上建立起一些不变区域作为支持区域, 继而在支持区域上建立起高稳定性的局部特征描述算子。目前文献中^{[77][78][79][80]}提出的局部特征检测算子有很多, 典型的有如下几种:

- Harris 算子: 该算子对图像旋转保持不变性, 所选择的支持区域是一个以感兴趣点为中心的大小固定的领域;
- Harris-Laplace 算子: 该算子对图像旋转、仿射变化和像素的灰度值变化保持不变性, 感兴趣点由一个尺度可变的 Harris 函数来检测, 在 Laplace 算子的尺度空间上做出最终选择, 所选的尺度决定了支持区域的大小;
- Harris-Affine 算子: 该算子的优点在于可以很好地保持对图像仿射变换的不变性, 感兴趣点的定位方法与 Harris-Laplace 算子相似;
- DoG (Difference of Gaussian) 算子: 该算子的最大优点是可以保持对图像尺度变化的不变性, 同时对图像的旋转和仿射变换也保持不变性, 其感兴趣点由基于差分高斯滤波器的尺度空间的局部极值上确定的, 所选的尺度决定了支持

区域的大小。

检测算子^{[76][80]}所选定的支持区域确定以后，接下来就要在支持区域上计算局部特征描述算子。近年来，人们提出了多种局部特征描述算子^{[75][76][79]}。典型的算子可分为以下几种：

- 基于颜色分布的算子：Johnson 与 Heber 提出了一种基于像素颜色分布的可用于三维目标识别的特征表达方法，在三维空间感兴趣点的邻域的相对位置建立起该点的直方图；
- 基于非参数变换的算子：Zabih 与 Woodfill 提出了一种建立在非参数统计基础上的局部变换，这种变换依据的数据的正序和倒序的信息而得，最终得到依据感兴趣点的邻域点的之间关系的特征描述；
- 基于差分技术的算子：Florack 等人推导出了依据局部导数的差分不变性，由此可以得到图像的旋转不变性，Freeman 和 Adelson 提出了可以对感兴趣点的梯度方向保持旋转不变的 Steerable 滤波器；
- 基于尺度不变的算子：Lowe^[74]提出了一种在感兴趣点的邻域利用多层次金字塔图像得到的对目标图像的尺度变化保持不变的 SIFT (Scale Invariant Feature Transform) 检测算子，在选定窗口内利用梯度直方图技术为每个感兴趣点来分配主方向，最终得到了一个可去除尺度变化、图像旋转等几何因素的特征向量。

5.1.4 粒子滤波基础

贝叶斯滤波原理的实质是试图用所有已知信息来构造系统状态变量的后验概率密度，即用系统模型预测状态的先验概率密度，再使用最近的量测值进行修正，得到后验概率密度。这样通过量测数据 $Z_{1:k}$ ，来递推计算状态 X_k 取不同值时的置信度 $p(X_k \mid Z_{1:k})$ ，由此获得状态的最优估计。若概率密度初值 $p(X_0 \mid Z_0) = p(X_0)$ ，对于一阶马尔可夫过程，由 Chapman-Kolmogorov 方程有先验概率密度

$$p(X_k \mid Z_{1:k-1}) = \int p(X_k \mid X_{k-1}) p(X_{k-1} \mid Z_{1:k-1}) dX_{k-1} \quad (5.6)$$

获得测量值 Z_k 后，通过贝叶斯公式更新先验值，得到后验概率密度

$$p(X_k \mid Z_{1:k}) = \frac{p(Z_k \mid X_k) p(X_k \mid Z_{1:k-1})}{p(Z_k \mid Z_{1:k-1})} \quad (5.7)$$

其中归一化常量 $p(Z_k \mid Z_{1:k-1}) = \int p(Z_k \mid X_k) p(X_k \mid Z_{1:k-1}) dX_k$ ，它取决于似然函数 $p(Z_k \mid X_k)$ 及测量噪声的统计特性。

对于系统状态的滤波估计通常采用状态空间模型法来描述系统特性和观测方程。经典的卡尔曼滤波算法及其衍生算法的推导都建立在状态空间模型法之上。

采用离散时刻状态空间模型表示动态系统，根据一直到当前时刻 t 的观测来估计隐含目标状态的概率分布。

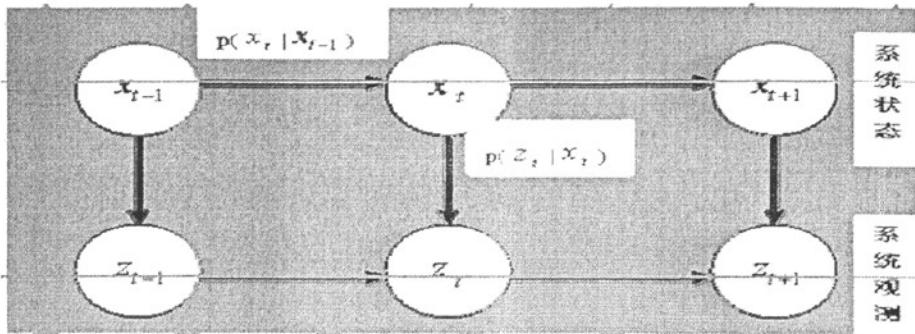


图 5.4 状态空间模型

描述系统动态特性的状态方程为

$$x_{t+1} = f_t(x_t) + w_t \quad (5.8)$$

描述观测特征的观测方程为

$$z_t = g_t(x_t) + v_t \quad (5.9)$$

式中 w_t 和 v_t 为具有已知概率分布的系统噪声和观测噪声向量，系统的初始状态概率密度函数为 $p(x_0)$ 。一般假设在状态-空间模型中，系统状态满足下式的马尔可夫属性。 $p(X_t) = p(x_0) \prod_{i=1}^N p(x_i | x_{i-1})$ 给定状态向量 X_t 时观测 Z_t 相互独立即满足 $p(Z_t | X_t) = \prod_{i=1}^N p(z_i | x_i)$ 。

粒子滤波方法^{[81][82]}是蒙特卡罗方法和贝叶斯方法的结合，也是求解贝叶斯估计问题的一种实用算法，适用于任何能用状态空间模型以及传统的卡尔曼滤波表示的非线性系统，精度可以逼近最优估计。它通过非参数化的蒙特卡罗模拟方法来实现递推贝叶斯滤波，其思想是利用一系列随机抽取的样本以及样本的权重来计算状态的后验概率分布。当样本数足够多时，可以近似于真实的后验概率分布。粒子滤波算法不用满足系统为线性、噪声高斯分布、后验概率也是高斯型的限制条件，因此应用面比卡尔曼滤波方法更广。同时由于粒子滤波同样具有贝叶斯滤波的时域递推结构，因此和卡尔曼滤波方法一样，不需要存储所有时刻的历史数据，计算时也仅需要上一时刻的估计值和当前时刻的观测值，对计算机的存储要求小，提高了运算速度，因此获得了广泛的关注，应用也越来越广。

粒子滤波的机制是：把状态空间分成许多部分，按照某种概率量度用粒子对这些部分进行填充。概率越大，粒子数就集中的越密。粒子系统根据状态方程随

时间演化。演化的概率密度函数满足 FPK(Fokker-Planck-Kolmogorov equation)等式。通过对状态空间的随机取样，获得一个可以代表演化概率密度函数的粒子集合。粒子滤波也就是通过寻找一组在状态空间中传播的随机样本对后验概率密度函数进行近似，以样本均值代替积分运算，从而获得状态最小方差估计的过程，这些样本即称为“粒子”。即

$$p(x_n | Z_n) \approx \frac{1}{N_p} \sum_{n=1}^{N_p} \delta(x_n - x_n^{(i)}) \equiv p(x_n | Z_n) \quad (5.10)$$

上式中 $x_n^{(i)}$ 是从 $p(x_n | Z_n)$ 中采样的独立同分布样本， N_p 充分大， $p(x_n | Z_n)$ 近似为真实的后验概率分布。

5.2 粒子滤波法独立运动目标检测

针对无成像平台运动模式限制的复杂环境下的多个独立运动目标检测的情形，我们提出了一种基于状态粒子数可变的时域空域双重采样的粒子滤波的融合双目立体视觉与惯性传感器的独立运动目标检测方法，其流程图如图 5.5。

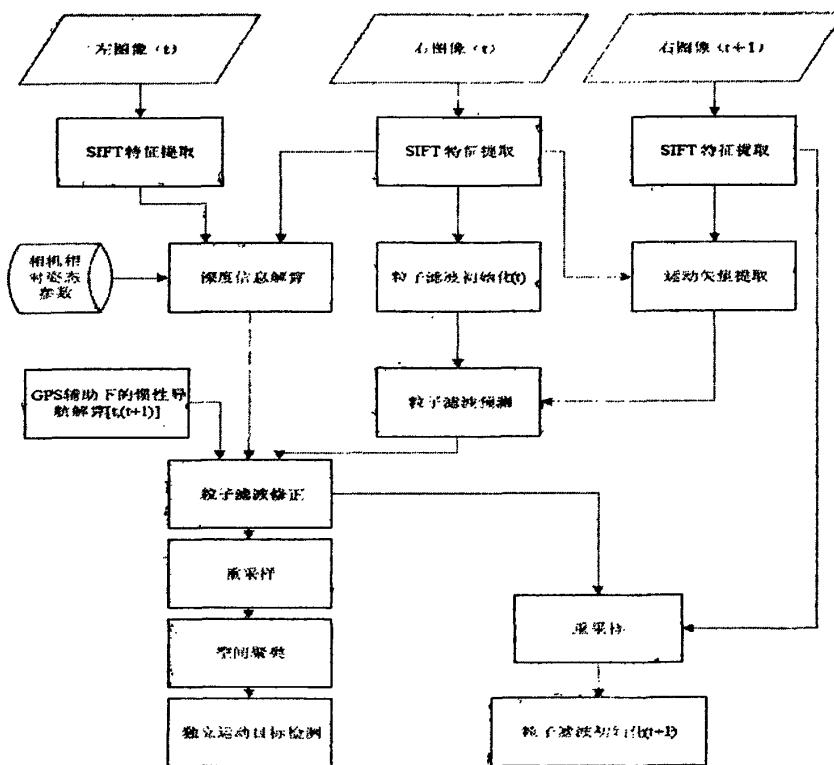


图 5.5 粒子滤波独立运动目标检测流程图

5.2.1 基于极线约束的 SIFT 特征匹配

SIFT 特征^{[74][75]}匹配方法是目前国内外局部特征分析领域的研究热点，该方法具有很强的局部特征匹配能力，可以处理图像之间发生平移、旋转、仿射变换情况下的匹配问题。SIFT 算法提取的 SIFT 特征向量具有如下特性：a) SIFT 特征是图像的局部特征，其对旋转、尺度缩放、亮度变化保持不变性，对视角变化、仿射变换、噪声也保持一定程度的稳定性；b) 区分性好，信息量丰富，适用于在海量特征数据库中快速、准确地匹配；c) 多量性，即使少数的几个物体也可以产生大量 SIFT 特征向量；d) 高速性，经优化的 SIFT 匹配算法甚至可以达到实时的要求；e) 可扩展性，可以很方便的与其它形式的特征向量进行联合。

本章中我们一方面采用 SIFT 特征进行双目立体像对的空域特征匹配，实现宽基线条件下的立体图像配准，利用已知的立体相机几何关系可以提取匹配特征点对的场景深度信息；另一方面采用 SIFT 特征进行单目图像序列的时域特征匹配，从而提取图像中的运动场信息。

SIFT 特征匹配算法包括两个阶段，第一阶段是 SIFT 特征向量的生成，即从多幅待匹配图像中提取出对尺度缩放、旋转、亮度变化无关的特征向量；第二阶段是 SIFT 特征向量的匹配。

本章中我们采用 Lowe^[74]提出的 SIFT 特征向量的生成算法进行特征向量的提取，由于已知立体相机的几何关系，我们采用基于空域极线约束的特征匹配算法对立体图像对 SIFT 特征组进行匹配；同理，惯性信息提供的平台运动参数信息可以确定相邻采样间隔之间单目相机的相对位置变化关系，我们采用基于时域极线约束的特征匹配算法对单目图像序列的时域 SIFT 特征组进行匹配。立体相机获取的立体图像对间的极线几何约束与单目运动相机获取的时域图像对间的极线几何约束本质是相同的^[4]，后文中我们不再做特别区分。

一幅图像 SIFT 特征向量的生成算法总共包括 4 步：

- (1) 尺度空间极值检测，以初步确定关键点位置和所在尺度；
- (2) 通过拟合三次多项式以精确确定关键点的位置和尺度，同时去除低对比度的关键点和不稳定的边缘响应点，以增强匹配稳定性、提高抗噪声能力；
- (3) 利用关键点邻域像素的梯度方向分布特性为每个关键点指定方向参数，使算子具备旋转不变性；
- (4) 利用关键点的邻域梯度信息生成特征向量。

假设三维空间中的 X 点在两个视角下成像，第一个视角成像为 x ，第二个视角成像为 x' 。由图 5.6 可以看出成像点 x 和 x' ，空间点 X 和相机的中心是共面的，把这个平面表示为 π ，显然由 x 和 x' 发出的反向投影光线相交于 X ，而且方向投影光线共面，位于平面 π 内。(a) 两个相机的位置由它们的成像中心 C 和 C' 及图像

平面表示。(b)由图像点 x 和相机成像中心 C 定义的反向投影光线在第二个视角中投影为直线 l' , 投影为 x 的三维空间点 X 必定位于反向投影光线上, 它在第二个视角中的成像投影必定位于 l' 上^[4]。

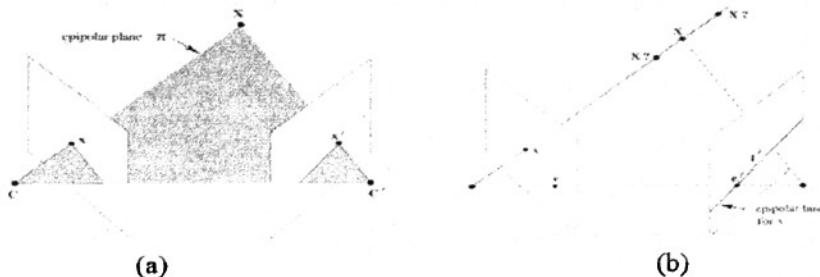


图 5.6 极线几何约束

考虑到惯性传感器数据的精度和极线求解过程中噪声影响, 我们在进行空域立体及时域单目图像对 SIFT 特征匹配时, 首先根据视角间的立体几何关系计算出 t 时刻右视图像中每个待匹配 SIFT 特征点在 t 时刻左视图像中的极线和 $t+1$ 时刻右视图像中的极线, 然后设定待匹配 SIFT 特征点基于极线约束的匹配区域(图 5.7), 最后在匹配区域内实现 SIFT 特征匹配。

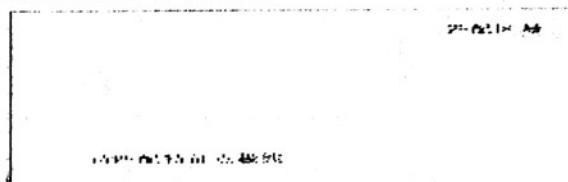


图 5.7 极线约束匹配区域

我们采用基于向量欧式距离定义两幅图像中关键点的相似性判定度量, 取图像 1 中的待匹配 SIFT 特征向量, 在待匹配区域中找出与图像 2 中欧式距离最近的前两个特征向量, 在这两个特征向量中, 如果最近的距离除以次近的距离小于比例阈值, 则接受匹配。图 5.8 为没有极线约束和极线约束匹配的搜索范围统计图。

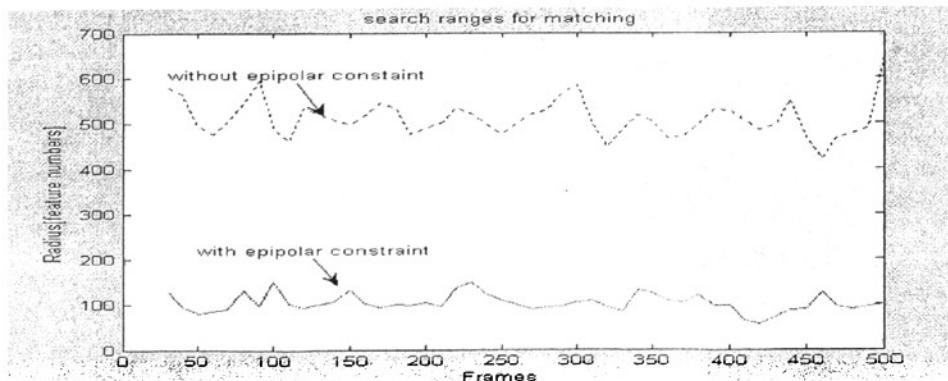


图 5.8 特征匹配搜索距离

5.2.2 时域空域双重重采样粒子滤波

假定以右相机坐标系为世界坐标系, t 时刻系统的状态表示为 θ_t , θ_t 表示世界坐标系中独立运动物体在右图像平面上的投影, 具体到图像平面坐标即为独立运动像素点。

假设系统状态是马尔可夫过程, 则条件传递函数 $p(\theta_{t+1} | \theta_0, \theta_1, \dots, \theta_t) = p(\theta_{t+1} | \theta_t)$, 本文中每个 SIFT 特征点的传递概率由该点的运动矢量与零均值归一化噪声两部分构成。

假设 t 时刻的系统观测表示为 y_t , 似然函数 $p(y_t | \theta_t)$ 对观测过程建模, 由于我们关注的是离散像素点属于独立运动物体投影点的概率, 因此似然函数应反映出像素点的运动属性。对 t 时刻左图像和右图像分别进行 SIFT 特征提取与匹配后, 利用已知的相机相对位置关系, 根据三角测量原理, 可以求出匹配特征点的深度信息。在 GPS 辅助下利用惯性测量单元获得的数据, 求解出时刻 t 与时刻 $t+1$ 之间相机的平移运动分量 $T = (T_x, T_y, T_z)^T$ 和 $\Omega = (\Omega_x, \Omega_y, \Omega_z)^T$, 三维物体 (X, Y, Z) 投影点 $(x = fX/Z, y = fY/Z)$ 的运动方程可表示为

$$\begin{aligned} u &= \left[\frac{-fT_x + xT_z}{Z} + \frac{xy}{f}\Omega_x - f\left(1 + \frac{x^2}{f^2}\right)\Omega_y + y\Omega_z \right], \quad (u, v) = (x, y) \\ v &= \left[\frac{-fT_y + yT_z}{Z} - \frac{xy}{f}\Omega_y + f\left(1 + \frac{y^2}{f^2}\right)\Omega_x + x\Omega_z \right] \end{aligned} \quad (5.11)$$

对每一个像素点, 有了深度和相机运动信息, 可以计算得到理想运动矢量, 由 t 时刻和 $t+1$ 时刻匹配 SIFT 特征可得到图像中实际的运动矢量, 两者之间的差异性反映了像素点的运动属性。

$$p(y | \theta) = 1.0 - \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\alpha^2}{2}\right) \quad (5.12)$$

其中 α 为理想运动矢量与实际运动矢量间的夹角。

粒子滤波独立运动目标检测的算法如下

- ◆ 初始化($t=0$), 由图像的 SIFT 特征点中产生 N 个状态粒子 $\theta_0^{(i)}$, $i \in \{1, 2, \dots, N\}$;
- ◆ 迭代进化:
 - (1) 通过传递概率密度函数 $p(\theta_{t+1} | \theta_t^{(i)})$ 预测 $\theta_{t+1}^{(i)}$;
 - (2) 计算粒子权值 $w_{t+1}^{(i)} = p(y_{t+1} | \theta_{t+1}^{(i)})$;
 - (3) 归一化权值;
 - (4) 空域重采样, 由大权值的粒子邻域产生粒子代替小权值粒子后进行空间聚类, 实现 t 时刻独立运动目标检测; 时域重采样, 由 $t+1$ 时刻检测的新的

SIFT 特征点粒子代替小权值粒子，结果产生 N^t 个状态粒子；

(5) 令 $t=t+1$ ，重复迭代：

空域重采样：由于在立体匹配和时域运动矢量提取过程中很多的 SIFT 特征向量由于没有可靠的立体匹配或时域匹配被滤除，而要获取精确的独立运动目标的位置信息，需要保留尽可能多的 SIFT 特征点。考虑到场景中的独立运动目标一般都具有区域特性，具有大权值运动属性的粒子的空间邻域内的特征点也应该具有相似的运动属性，故实验中我们空域重采样的方法为：在大权值粒子的 15×15 空间邻域内搜索 SIFT 特征点用来代替小权值的粒子，没有可替代的 SIFT 特征点时复制大权值粒子本身取代小权值粒子。

时域重采样：复杂场景运动成像平台下的独立运动目标检测，由于成像平台的运动和独立运动目标的运动，已检测出的独立运动目标可能离开视野或新的独立运动目标进入视野，如何及时有效的检测这种变化是实现多个独立运动目标连续正确检测的关键。在运动矢量提取过程中，我们滤除了那些在 t 时刻检测出的同时在 $t+1$ 时刻没有可靠匹配的特征点，解决了独立运动目标离开视野的问题。由似然函数我们可以得到当前 t 时刻图像中的特征点属于独立运动投影点的概率（粒子权重），小权值的粒子属于由成像平台运动引起的运动背景。为了能有效的对新进入视野的独立运动目标进行检测，实验中我们时域重采样的方法为：用 $t+1$ 时刻图像中检测出的同时在 t 时刻图像中没有可靠后向匹配的新的 SIFT 特征点粒子取代 t 时刻滤波后的小权值粒子，这里需要注意到典型的粒子滤波体系中状态粒子个数是固定的，而我们提出的时域空域重采样粒子滤波是状态粒子可变的，从而有效的解决了新的独立运动目标进入视野的问题。

5.3 实验与分析

采用 VISA™ 多传感器移动测绘实测的双目图像序列和惯性传感器数据进行实验，图 5.9 是本文开发的实验系统的软件界面。

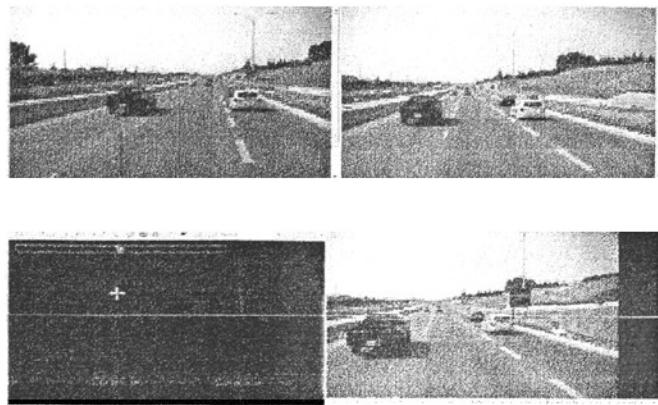


图 5.9 粒子滤波独立运动检测系统

实验系统如图 5.9 所示，上方两个窗口为原始双目序列图像浏览窗口，下左窗口为 GPS 辅助惯性导航解算得到的地理参考坐标，下右窗口为粒子滤波融合独立运动目标检测结果。

图 5.10 为 SIFT 特征点检测与运动矢量提取结果图。

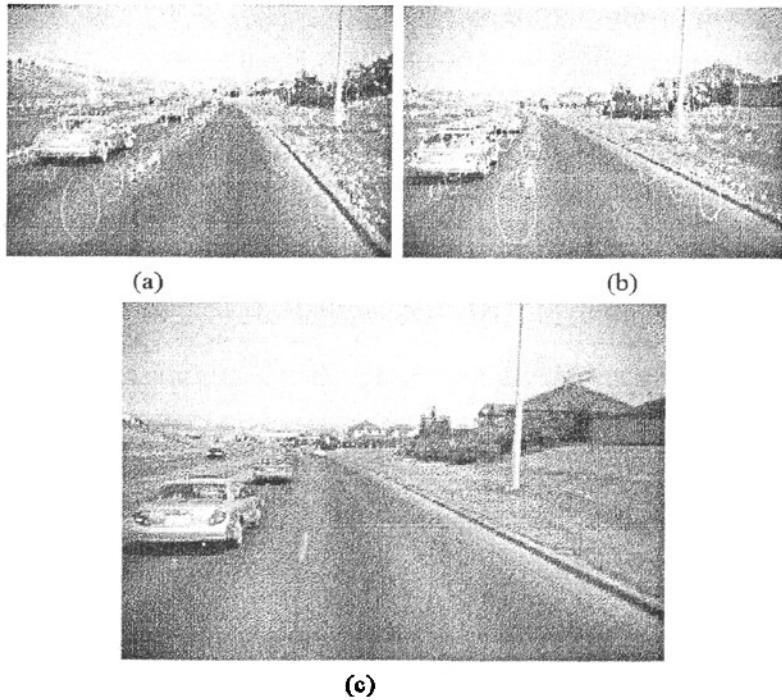


图 5.10 SIFT 特征提取及运动矢量提取

图 5.10 中(a)为 $t=1757$ 时左视相机图像，其中检测出 SIFT 特征点 2585 个，其中+表示特征点位置，椭圆大小代表特征点的尺度，(b)为 $t=1757$ 时右视相机图

像，其中检测出 SIFT 特征点 2347，其中+表示特征点位置，椭圆大小代表特征点的尺度。(c)为利用右视图像序列中对 $t=1757$ 和 $t=1758$ 时的帧图像进行 SIFT 特征检测和匹配后形成的实际图像运动场图，其中匹配的 SIFT 特征对为 683，为了显示细节，结果图像选用了大的尺度因子。从图中可以看出 SIFT 特征点的空间概率分布以及运动矢量的概率分布可以保证后序独立运动目标检测的有效性，但对远景小目标检测可能存在特征漏检。

图 5.11 为采用粒子滤波融合检测方法对 VISAT™多传感器移动测绘平台实测的多种复杂场景中多运动模式的数据进行独立运动目标检测的结果图。

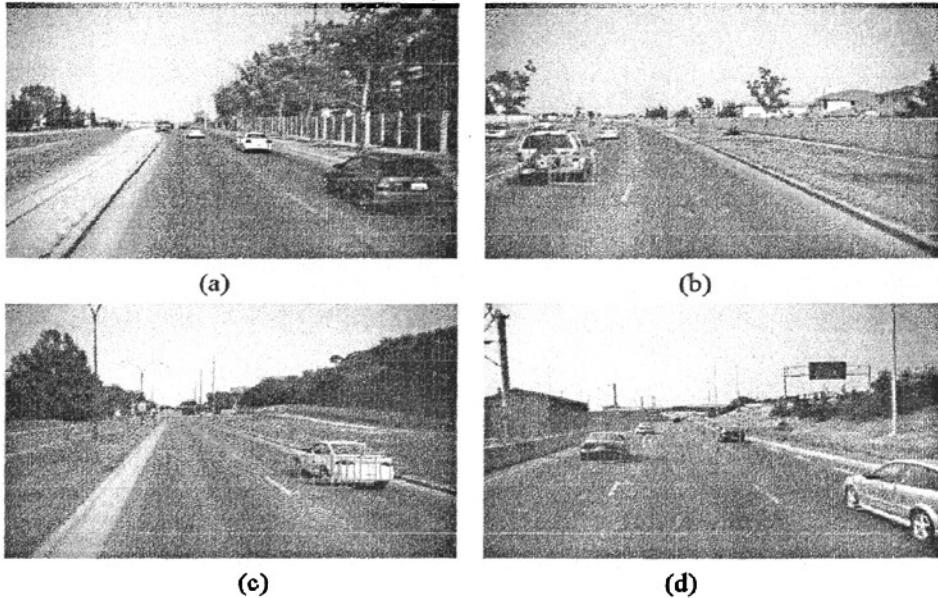


图 5.11 独立运动目标检测结果

在复杂环境多运动模式的场景中的实验结果证明了方法的有效性和稳健性。但从图 5.11(b)结果图中可以看出，远景深中的路边缘部分为虚警，这主要是由于远景深除草地和路面部分的特征点的跟踪匹配有一定的误差，而其本身的偏移量绝对值较小，在处理过程中检测判决易受到噪声的干扰。

5.4 本章小结

本章首先介绍了立体视觉的相关基础，介绍了宽基线条件下立体视觉特征检测及描述方法，介绍了粒子滤波的基本思想，然后针对复杂场景中多运动模式下的多个独立运动目标检测的情形，我们提出了一种融合双目立体视觉和惯性传感信息的状态粒子数可变的时域空域双重重采样的粒子滤波独立运动目标检测的新方法，在极线约束下利用 SIFT 特征匹配方法提取可靠的立体图像特征对和单目序

列图像的运动场，在分析理想光流场与实际光流场的差异性的基础上利用状态粒子数可变的时域空域双重重采样粒子滤波实现多个独立运动目标检测，并对 VISAT 多传感器移动测绘平台实测的多个复杂背景中多运动模式的数据进行了实验，结果表明，本方法能够有效的检测复杂场景中的多个独立运动目标。

结 束 语

本论文是我一年多来学习和工作经验的总结，希望有兴趣阅读它的人可以有所收获。下面是对论文工作的简要总结：

1、针对成像平台沿成像光轴方向平移运动可忽略或场景的深度变化远小于场景绝对深度的情形，本文提出了一种基于运动补偿的融合单目视觉传感器和惯性传感器的独立运动目标检测的新方法，首先利用惯性传感器平台运动参数信息和景深信息对相邻帧图像进行粗配准，然后提取背景块进行图像精配准，最后采用帧间差法检测独立运动目标，同时利用图像观测修正惯性传感器数据，抑制了惯性传感器的数据漂移问题。

2、针对成像平台旋转运动可忽略的情形，本文提出了一种基于扩展点光流一致性约束的融合单目视觉传感器和惯性传感器的独立运动目标检测新方法，利用惯性传感器的平台运动信息和成像相机内参数预测运动场中扩展点位置，然后在计算稳健光流场的基础上利用一致性约束对光流矢量进行分类，实现独立运动目标的检测。

3、针对无成像平台运动模式限制的复杂场景中独立运动目标检测情形，本文提出了一种基于时域空域双重重采样的粒子滤波的融合双目立体视觉与惯性传感器的独立运动目标检测的新方法，利用极线约束下的 SIFT 特征点立体匹配和运动匹配、惯性传感器平台运动参数等先验信息，在粒子滤波框架下以理想运动矢量与实际运动矢量的差异为观测对独立运动特征点投影状态进行修正，最后进行空间聚类实现独立运动目标的检测。

融合视觉和惯性传感器的独立运动目标检测是计算机视觉领域的一个比较前沿的课题，在移动测绘、智能交通系统、机器人导航等多学科中有广泛的应用前景。这一领域有很多问题还没有或者没有很好地解决，如融合策略、运动表示等，还有很多事情可以做，很多路需要走。希望我的工作能够为后续的工作者起到抛砖引玉的作用，更希望后续工作者能够取得更大的突破。

致 谢

本论文是在王程副教授的亲切关怀、悉心指导和耐心帮助下完成的。在我攻读硕士学位及论文撰写期间，得到了王老师的亲切关怀，特别是在论文的撰写过程中，王老师提出了许多宝贵的意见和建议，正是由于这样本文才得以顺利完成。两年多来，王老师严肃的科学态度、严谨的治学精神、扎实的科研作风、开阔的思路以及对待工作和学生高度的责任感给我留下了深刻的印象，使我受益匪浅。这一切对我今后的工作、学习和生活来讲都是宝贵的财富。在此谨向王老师致以最诚挚的谢意和最崇高的敬意。

感谢王润生教授、刘方副教授、朱长仁副教授、张鹏老师、周军红副教授，在科研项目和课题研究上给我的启迪、教育和帮助；老师们敏捷的思维、孜孜不倦的钻研精神、高尚的师德使我受到了很大程度的熏陶和影响。

非常感谢同教研室的各位师兄弟、师姐在学习和生活上给予的帮助，他们是：郭武、钟平、王威、裴立志、江悦、周晖、张志、程环环、陈进、孙玉康、陈宜稳、郭军、叶蓬、王建平、伍新科、胡粲彬。

感谢国防科技大学研究生院、电子科学与工程学院以及 ATR 国家重点实验室的领导和老师们，为我完成学业给予了不同程度的支持和帮助。

感谢我的父母家人，感谢他们在学习上无私的支持鼓励和生活上无微不至的关心，我成长的每一步都凝聚着他们的辛劳、汗水和希望，他们的支持与鼓励永远是我前进的动力。

最后，我要感谢黎超同学，感谢她对我学习和工作上的无私支持和关心，使我顺利的完成本文。

参考文献

- [1]. 王润生. 信息融合[M]. 北京: 科学出版社,2007.
- [2]. 王润生. 图像理解[M]. 国防科技大学出版社,1995.
- [3]. 秦永元. 惯性导航[M]. 北京: 科学出版社,2006.
- [4]. A.Z.R. Hartley. Multiple View Geometry in Computer Vision [M]. Cambridge university press, 2000.
- [5]. O.D. Faugeras. Three-Dimensional Computer Vision: A Geometric Viewpoint. MIT press, 1993.
- [6]. 邢慧, 颜景龙, 张树江. 基于相位相关的数字稳像算法研究[J]. 光电工程. 2007.34(3):37-41.
- [7]. 王嘉, 王海峰, 刘青山, 卢汉清. 基于三参数模型的快速全局运动估计[J]. 计算机学报. 2006.29(6):920-927.
- [8]. 吴思, 张勇东, 林守勋, 李豪杰. 动态场景视频序列中的前景区域自动提取[J]. 计算机辅助设计与图形学学报. 2005.17(2):359-363.
- [9]. Moscheni F. Spatio-temporal segmentation and object tracking: An application to second generation video coding [D]. Lausanne: Swiss Federal Institute of Technology, 1997
- [10]. Bleser. G , Becker. M, Stricker. D. Fast and Stable Tracking for AR fusing Video and Inertial Sensor Data[C]. Proceedings of WSCG 2006.
- [11]. Streloew .D, Singh.S. Motion Estimation from Image and Inertial Measurements[J]. The International Journal of Robotics Research, 2004, 23(12),1157-1194.
- [12]. Huster.A, Rock.S.M. Relative position sensing by fusing monocular vision and inertial rate sensors[C]. Proceedings of the 11th International Conference on Advanced Robotics(ICAR 2003), July, vol.3,1562-1567.
- [13]. Foxlin.E.M, Naimark.L. VIS-Tracker: a wearable vision-inertial self-tracker[C]. IEEE Virtual Reality Conference,2003.
- [14]. Kameda.Y, Minoh Michihiko. A Human Motion Estimation Method Using 3-successive Video Frames[C].Proceedings of International Conference on Virtual Systems and Multimedia,1996,135-140.
- [15]. Paragios. N, Geodesic. R.D. Active Contours and Level Sets for the Detection and Tracking of Moving Objects [J].IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(3):266-280.
- [16]. Guo Jing, Rajan. D, Chng Eng Siong. Motion Detection With Adaptive Background And Dynamic Thresholds[C]. In: Proc IEEE Conference on Information, Communications and Signal Processing, 2005, 41-45.

-
- [17]. Fang Shuai, Chi Jian-nan, Xu Xin-he. Moving Target Tracking Algorithm in Video Surveillance [J]. Control and Decision, 2005, 20(12):1388-1391.
 - [18]. Hu Wei-ming, Tan Tie-niu, Wang Liang, and S. Maybank. A Survey on Visual Surveillance of Object Motion and Behaviors[J]. IEEE Transactions on systems, man, and cybernetics—PART C: Applications and Reviews, 2004 , 34(3): 334-352.
 - [19]. Cheng, Y. MeanShift Model Seeking and Clustering[J]. IEEE Trans.on PAMI 1995,17(8):790-799.
 - [20]. 王圣男, 郁梅, 蒋刚毅.智能交通系统中基于视频图像处理的车辆检测与跟踪方法综述[J].计算机应用研究, 2005,9:9-14.
 - [21]. 王蕾, 张旭东.自动分割视频序列中运动物体的新算法: 多特征联合方法[J].中国图象图形学报, 2003,8A(11):1346-1351.
 - [22]. Michael Isard , Andrew Blake. CONDENSATION-conditional density propagation for visual Tracking [J]. Int..J.Computer Vision, 1998,1(29):5-28.
 - [23]. Haritaoglu I, Harwood D and Davis L. W4: real-time surveillance of people and their activities[J]. IEEE Trans Pattern Analysis and Machine Intelligence, 2000, 22 (8): 809-830.
 - [24]. McKenna S et al. Tracking groups of people[J]. Computer Vision and Image Understanding, 2000, 80 (1): 42-56.
 - [25]. Karmann K and Brandt A. Moving object recognition using an adaptive background memory.[C] In: V Cappellini, Time-varying Image Processing and Moving Object Recognition. Elsevier, Amsterdam, The Netherlands, 1990.
 - [26]. Kilger M. A shadow handler in a video-based real-time traffic monitoring system[C]. In: Proc IEEE Workshop on Applications of Computer Vision, Palm Springs, CA, 1992, 1060-1066.
 - [27]. Stauffer C and Grimson W. Adaptive background mixture models for real-time tracking[C]. In: Proc IEEE Conference on Computer Vision and Pattern Recognition, Fort Collins, Colorado, 1999, 2: 246-252.
 - [28]. K. Toyama, J. Krumm, B. Brumitt, and B. Meyers. Wallflower: principles and practice of background maintenance[C]. In Proc. Int. Conf. Computer Vision, 1999, 255–261.
 - [29]. Meyer D, Denzler J and Niemann H. Model based extraction of articulated objects in image sequences for gait analysis[C]. In: Proc IEEE International Conference on Image Processing, Santa Barbara, California 1997, 78-81.
 - [30]. Elgammal, A.D. Harwood, and L.S. Davis. Nonparametric Model for Background Subtraction[C]. ECCV. 2000.751-767.
 - [31]. Hanzi Wang and David Suter.. Background Subtraction Based on a Robust Consensus Method[C]. In:Proc IEEE 18th International Conference on Pattern
-

-
- Recognition 2006, 223-226.
- [32]. S.Araki T., Matsuoka H., and Yokoya N. Real-time tracking of multiple moving objects in moving camera image sequences using robust statistics[C]. Proceedings of the 14th International Conference on Pattern Recognition,1998,1433-1435.
- [33]. Murray D and Basu A. Motion tracking with an active camera[J]. IEEE Trans Pattern Analysis and Machine Intelligence, 1999, 16(5):449-459.
- [34]. Guo Jing, Chng Eng Siong and Rajan D. Foreground motion detection by difference-based spatial temporal entropy image[C].2004 IEEE Region 10 Conference.2004,A,1.379- 382.
- [35]. Lipton A, Fujiyoshi H and Patil R. Moving target classification and tracking from real-time video[C]. In: Proc IEEE Workshop on Applications of Computer Vision.Princeton.NJ.1998, 8-14.
- [36]. Stringa E. Morphological change detection algorithms for surveillance applications[C]. In: British Machine Vision Conference, Bristol, UK.2000, 402-411.
- [37]. Neri A, Colonnese S, Russo G, et al. Automatic moving object and background separation[J].Signal Processing.1998,66 (2):219-232.
- [38]. Pedersini F, Sarti A, Tubaro S. Combined motion and edge analysis for a layer-based representation of image sequences[C]. IEEE Proceedings of International Conference on Image Processing.1996, 1(1):921-924.
- [39]. Gouze A, DeRover C, Macq B, Herbulot A, Debrouve E, Barlaud M. Watershed-driven active contours for moving object segmentation[C]. IEEE Proceedings of International Conference on Image Processing.2005, 12(2):810-818.
- [40]. Bertozzi M, Broggi A, and Castelluccio S. A Real-Time Oriented System for Vehicle Detection[J]. Systems Architecture, 1997, 317-325.
- [41]. Matthews N, An P, Charnley D, and Harris C. Vehicle Detection and Recognition in Greyscale Imagery[J]. Control Eng.Practice, 1996,14, 473-479.
- [42]. Zhao G and Yuta S. Obstacle Detection by Vision System for Autonomous Vehicle[J]. Intelligent Vehicles, 1993,12, 31-36.
- [43]. Sun Z, Bebis G, and Miller R. On-road vehicle detection: A review[J]. IEEE Trans. Pattern Anal. Mach. Intell., 2006, 28(5): 694-711.
- [44]. Sun Z, Bebis G, and Miller R. Monocular precrash vehicle detection: features and classifiers[J]. IEEE Trans. Image Proeess., 2006,15(7):2019-2034.
- [45]. Y. Zhang, S. J. Kiselewich, W. A. Bauson, and R. Hammoud. Robust moving object detection at distance in the visible spectrum and beyond using a moving camera[C]. Proc. IEEE Int'l Conf. Computer Vision and Pattern Recognition Workshop, 2006.
- [46]. El-sheimy N, Schwarz P, and Gravel M. Mobile 3-D positioning using

-
- GPS/INS/Video Cameras[C]. In Proc. Mobile Mapping Symposium, OSU Center for Mapping, 1995. 236-240.
- [47]. El-Sheimy N and Schwarz K. Navigating urban areas by VISAT- A mobile mapping system integrating GPS/INS/Digital cameras for GIS application[J]. Navigation, 1999, 45(4): 275-286.
- [48]. Bossler J and Toth C. Feature positioning accuracy in mobile mapping: results obtained by the GPSVan™[C]. Int'l Archives of Photogrammetry and Remote Sensing, 1996, vol. XXXI, part B4, 139-142.
- [49]. Grejner-Brzezinska D. Airborne integrated mapping systems: positioning component[C]. In Proc. 53rd ION Annual Meeting, Albuquerque, NM, 1997, 225-235.
- [50]. Hu Z and Uchimura K. Moving object detection from time-varied background: an application of camera 3D motion analysis[C]. In Proc. IEEE Int'l Conf. 3-D Imaging and Modeling, 1999, 59-67.
- [51]. Takeda N, Watanabe M, and Onoguchi K. Moving object detection using residual error of FOE estimation[C]. In Proc. IEEE Int'l Conf. Intelligent Robots and Systems, 1996, 4:1642-1647.
- [52]. Schmid C, Mohr R, and Bauckhage C. Evaluation of interest point detectors[J]. Int J. Comput. Vis., 2000, 37(2):151-172.
- [53]. Shi J and Tomasi C. Good features to track[C]. In Proc. IEEE Comput. Soc. Conf. Comput. Vision and Pattern Recogn., 1994, 593-600.
- [54]. Franke U and Heinrich S. Fast object detection for urban traffic situations [J]. IEEE Trans. Intell. Syst., 2002, 3(3):173-181.
- [55]. Bruhn A, Weickert J, and Schnorr C. Lucas/Kanade meets Horn/Schunck: combining local and global optic flow methods[J]. Int J. Comput. Vis., 2005, 63(3):211-231.
- [56]. Alam S M, Bal A. Improved multiple target tracking via global motion compensation and optoelectronic correlation[J]. IEEE Transactions on Industrial Electronics, 2007, 54(1):522-529.
- [57]. Prati A, Seghedo F, Cucchia R. Fast dynamic mosaicing and person following[A]. Proc. of International Conference on Pattern Recognition[C], Hong Kong, 2006, 920-923.
- [58]. Yamaguchi K, Kato T, Ninomiya Y. Vehicle ego-motion estimation and moving object detection using a monocular camera [A]. Proc. of International Conference on Pattern Recognition[C], Hong Kong, 2006, 610-613.
- [59]. Schoen T, Gustafsson F. Integrated navigation of cameras for augmented reality[A]. International Federation of Automatic Control World Congress[C], Prague, Czech Republic, 2005.
- [60]. Hol J, Slycke P, Schoen T, Gustafsson F. 2d-3d model correspondence for camera

-
- pose estimation suing sensor fusion. Proc.of InerVis workshop at the IEEE International Conference on Robotics and Automation,2005.<http://www.xsens.com/>
- [61]. Chumerin N, Van H M M. Cue and sensor fusion for independent moving objects detection and description in driving scenes. <http://simone.neuro.kuleuven.be/~nick/papers/Chumerinbook-chapter-2007-draft-1.pdf>
- [62]. Sawhney H, Guo Y, Kumar R. Independent motion detection in 3D scenes. [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2000, 22(10):1191-1199.
- [63]. Irani M, Anandan P. A unified approach to moving object detection in 2D and 3D scenes[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(6):577-589.
- [64]. Grinstead B, Koschan A, Abidi M. A comparison of pose estimation techniques: Hardware vs. Video[A]. Proceedings SPIE Unmanned Ground Vehicle Technology[C], Orlando, FL, 2005, VII(5804) 166-173.
- [65]. Trucco. E, Tommasini. T, Roberto. V. Near-Recursive Optical Flow From Weightted Image Differences[J]. IEEE Transactions on Systems, Man, and Cybernetics-Part B: CYBERNETICS, 2005, 35(1):124-129.
- [66]. Chunke. Y, Oe.S. A new gradient-based optical flow method and its application to motion segmentation. IECON 2000, Vol.2, 1225-1230.
- [67]. Chivilo. G, Mezzaro. F, Sgorbissa. A, Zaccaria. R. Follow-the-Leader Behaviour through Optical Flow Minimization[C]. Proceedings of 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems,3182-3187.
- [68]. Giachetti. A, Campani. M and Torre. V. The Use of Optical Flow for Road Navigation[J]. IEEE Transactions on Robotics and Automation, 1998, 14(1):34-48.
- [69]. 王新余,张桂林. 基于光流的运动目标实时检测方法研究[J]. 计算机工程与应用, 2004.1, 43-46.
- [70]. 武理静, 明军. 一种改进的光流估计方法[J]. 国外电子测量技术, 2006, 25 (12) :6-9.
- [71]. 胡以静, 李政访, 胡跃明. 基于光流的运动分析理论及应用[J]. 计算机测量与控制, 2007, 15 (2) : 219-221.
- [72]. Hao Sun, Cheng Wang, Naser El-sheimy. Moving Objects Detection for Mobile mapping[C].IEEE Conference on Information Acquisition and Automation, 2008.
- [73]. 孙浩, 王程, 王润生. 基于多传感器融合的运动平台运动目标检测[J]. 计算机应用, 2008.4

-
- [74]. Lowe. D. Distinctive image features from scale-invariant keypoints, cascade filtering approach[J]. 2004, Internation Journal of Computer Vision, 60:91-110.
 - [75]. Ke. Y, Sukthankar. R. PCA-SIFT: A more distinctive represation for local image descriptors[J]. In CVPR(2). (2004):506-513.
 - [76]. Mikolajczyk.K, Schmid. C. An affine invariant interest point detector. In ECCV,2002, 128-142.
 - [77]. Tuytelaars. T , Van Gool. L. Wide baseline stereo based on local, affinely invariant regions. In :BMVC,2000, 412-422.
 - [78]. Matas. J, Chum. O.M.U, Pajdla. T. Robust wide baseline stereo from maximally stable extremal regions. In: BMVC, 2002, 384-393.
 - [79]. Mikolajczyk. K, Schmid. C. A performance evaluation of local descriptors [J]. PAMI, 2005(27):1615-1630.
 - [80]. Mikolajczyk. K, Tuytelaars. T, Schmid. C, Zisserman. A, Matas. J etc. A comparison of affine region detectors [J]. IJCV,2005(65),43-72.
 - [81]. Isard. M, Blake. A. Condensation-conditional density propagation for visual tracking. International Journal of Computer Vision, 1998:5-28.
 - [82]. Isard. M, Blake. A. Condensation-Unifying low-level and high-level tracking in a stochastic framework In ECCV'98:893-908.
 - [83]. Frykman. P. Applied particle filters in integrated aircraft navigation. Master's thesis, 2003.
 - [84]. Yong. G.J, Yong. J.L, Hoon. K. Segmentation tracking and Recognition based on Foreground-Background Absolute Features, Simplified SIFT, and Particle Filters[C]. IEEE Congress on Evolutionary Computation,2006, 1279-1284.
 - [85]. Cheng. C, Ansari. R. Kernel Particle Filter for Visual Tracking[J]. IEEE SIGNAL PROCESSING LETTERS, 2005,12(3):242-245.

作者在学习期间取得的学术成果

发表和已撰写的论文：

- [1] 孙浩, 王程, 王润生. 基于多传感器融合的运动平台运动目标检测[J]. 计算机应用 2008.4
- [2] Hao Sun, Cheng Wang, Naser El-sheimy. Moving Objects Detection for Mobile mapping .ICIA2008.

参与的科研项目：

- [1] 纵向项目. “可见光卫星图像舰船目标检测子系统”. 2007.8—2008.9
- [2] 国防科技重点实验室项目. “×××中基于异类传感器的信息融合关键技术研究”. 2006.12—2008.12