

分类号: TP391

密 级: 公开

学校代码: 10363

学 号: 2210110111



安徽工程大学

Anhui Polytechnic University

硕士学位论文

题 目 工业机器人全局结构运动
视觉检测方法研究

论 文 作 者 邓湘琳

指 导 教 师 吴路路 副教授

学 科 (专 业) 机械工程

研 究 方 向 机器视觉

论文提交日期: 2024 年 5 月 30 日

类 号：TP391

单位代码：10363

密 级：公 开

学 号：2210110111

工业机器人全局结构运动视觉测量方法研究
RESEARCH ON GLOBAL STRUCTURAL MOTION
VISUAL MEASUREMENT METHODS FOR INDUSTRIAL
ROBOT

学生姓名： 邓湘琳

指导教师： 吴路路

专 业： 机械工程

研究方向： 机器视觉

答辩日期：2024 年 5 月 24 日

安徽工程大学学位论文原创性声明

本人郑重声明：我恪守学术道德，崇尚严谨学风。所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已明确注明和引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的作品及成果的内容。论文为本人亲自撰写，我对所写的内容负责，并完全意识到本声明的法律后果由本人承担。

学位论文作者签名：邓湘琳

日期：2024 年 5 月 30 日

安徽工程大学学位论文版权使用授权书

学位论文作者完全了解学校有关保留、使用学位论文的规定，同意学校保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅或借阅。本人授权安徽工程大学可以将本学位论文的全部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和汇编本学位论文。

保密 ☐，在_____年解密后适用本版权书。

本学位论文属于

不保密 ☒

学位论文作者签名：邓湘林

日期：2024年5月30日

指导教师签名：吴晓峰

日期：2024年5月30日

工业机器人全局结构运动视觉测量方法研究

摘 要

工业机器人作为一种刚柔耦合的串联机械体，其结构运动状态影响着机器人的工作性能。开展机器人结构运动性能测量与评估所提供的数据支撑，对于提升机器人运动学标定精度，优化机器人本体结构，具有十分重要的工程价值。在机器人结构运动测量领域，由于现有测量仪器的适用性和精度的限制，需寻求一种具备全局性、高精度、快速的测量技术。针对该问题，本文着眼于视觉测量技术，依托其测量的全局性、非接触及高灵活度的优势，将视觉系统应用于机器人全局结构运动的测量场景，完成机器人运动的可视化观测。本文的主要研究内容如下：

（1）构建了机器人结构运动视觉测量的整体方案，推导了相机成像模型与参数标定的相关理论，阐述了机器人运动表达形式。

（2）采用 YOLOv5 算法，通过对所设计的合作标识进行检测，实现了机器人结构运动追踪。为准确定位合作标识特征点的像素坐标，提出了一种亚像素坐标定位方法，融合了 Zernike 矩亚像素边缘点检测和最小二乘直线拟合，有效降低了边缘检测误差对特征点的影响，提升了特征点检测的鲁棒性，实现了合作标识特征点的亚像素级坐标提取，为后续机器人空间运动测量奠定基础。

（3）提出了基于单应特征矩阵的视觉测量模型，实现机器人平面结构运动的测量，该模型能够降低相机光轴与被测平面非正交偏角所产生的测量误差，从而提高工业机器人平面运动测量精度。

（4）构建了双目视觉测量与解析模型，通过三维重构和坐标变换，实现机器人空间结构运动测量。通过自制标识建立运动坐标系，实现了机器人姿态测量。为降低测量点误差对坐标变换求解精度的影响，提出基于误差加权的 RANSAC 三维坐标变换优化方法，该方法基于对各测量点数据赋予不同计算权重，优化了坐标系空间变换解，进而提高工业机器人空间运动测量精度。

（5）搭建了软硬件系统平台，结合本文提出的测量模型与方法，进行机器人结构运动视觉跟踪、定位与测量实验，验证了所提算法的有效性和精度。

关键词：工业机器人，结构运动，视觉测量，追踪定位，坐标变换优化

RESEARCH ON GLOBAL STRUCTURAL MOTION VISUAL MEASUREMENT METHODS FOR INDUSTRIAL ROBOT

ABSTRACT

The motion status of a robot intuitively reflects its working performance, thus conducting an evaluation of robot motion performance can provide valuable engineering support for optimizing the kinematic calibration and body structure of the robot. In the field of robot structural motion measurement, the limitations in applicability and precision of existing measuring instruments necessitate the search for a measurement technique that possesses global coverage, high precision, and rapid capabilities. This issue has been recognized as significant engineering value in providing data support for the optimization of robot body structure through calibrating the robot's kinematics. In response to this issue, this paper focuses on visual measurement technology, based on its advantages of non-contact measurement, high flexibility and accuracy. By applying the visual system to the measurement scenario of the global structural motion of the robot, the aim is to achieve visual observation of the robot's motion status and realize the two-dimensional and three-dimensional motion tracking and spatial positioning of the robot. The main research contents of this paper are as follows:

(1) A comprehensive framework for visual measurement of robot structural motion was established, encompassing the derivation of relevant theories on camera imaging models and parameter calibration, as well as the fundamental representation of robot motion.

(2) The YOLOv5 network framework was employed to track the cooperative marker designed for robot motion tracking. To accurately determine the pixel coordinates of the cooperative target's feature points, a robust sub-pixel coordinate localization method was proposed. This method integrates Zernike moment sub-pixel edge point detection and least squares line fitting, effectively reducing the impact of edge detection errors on feature points. Consequently, sub-pixel level coordinates of the cooperative marker's feature points were extracted, laying the groundwork for subsequent spatial motion measurement of the robot.

(3) A visual measurement model based on the homography feature matrix was proposed to achieve the measurement of robot planar structural motion. This model is capable of reducing measurement errors caused by non-orthogonal angular deviations between the camera optical axis and the measured plane, thereby enhancing the precision of industrial robot planar motion measurements.

(4) A binocular visual measurement and analysis model was constructed to measure the spatial structural motion of robots. Robot posture measurement was achieved by establishing a motion coordinate system using self-made markers. In order to mitigate the impact of measurement point errors on the precision of coordinate transformation solutions, an error-weighted RANSAC three-dimensional coordinate transformation optimization method was proposed. This method, based on assigning computational weights to the measurement point errors, optimized the solution for coordinate space transformation, thereby enhancing the precision of spatial motion measurements for industrial robots.

(5) A software and hardware system platform was established. By applying the measurement model and techniques proposed in this paper, visual tracking, positioning, and measurement experiments of robot structure motion were conducted, validating the measurement accuracy of the system and the effectiveness of the algorithms.

Deng Xianglin

Supervised by Associate Professor Wu Lulu

KEY WORDS: Industrial robots, Structural motion, Visual measurement, Tracking and positioning, Coordinate transformation optimization

目 录

摘 要	I
ABSTRACT	II
第 1 章 绪 论	1
1.1 课题研究背景与意义	1
1.2 机器人运动测量技术研究现状	2
1.2.1 机器人运动测量技术	2
1.2.2 视觉测量技术	4
1.3 本文研究内容与组织结构	6
第 2 章 机器人运动视觉测量系统构建	9
2.1 视觉测量系统方案	9
2.2 相机成像模型与参数标定	10
2.2.1 3D 到 2D 几何空间映射	10
2.2.2 单目相机参数标定	12
2.2.3 双目相机参数标定	15
2.3 机器人运动状态表达	16
2.3.1 机器人运动位置表达	16
2.3.2 机器人运动姿态表达	18
2.4 本章小结	20
第 3 章 机器人结构运动追踪与定位	21
3.1 基于深度学习的目标追踪	21
3.1.1 神经网络的基本结构	21
3.1.2 信息传递与反馈	22
3.2 基于 YOLOv5 的机器人运动追踪	23
3.2.1 YOLOv5 网络结构	24
3.2.2 损失函数与目标判别策略	25
3.2.3 算法部署	26
3.3 特征点精确定位	27
3.3.1 圆形十字标识特征点定位	27
3.3.2 三角形标识特征点定位	29
3.4 本章小结	30
第 4 章 机器人全局结构运动视觉测量	31
4.1 基于视觉的机器人平面运动测量	31
4.1.1 测量方法分析	31
4.1.2 基于单应特征矩阵模型的平面运动测量	31
4.2 基于视觉的机器人空间运动测量	33

4.2.1 3D 重建模型.....	33
4.2.2 坐标转换方法分析.....	35
4.2.3 基于 RANSAC 误差加权的坐标变换模型优化	36
4.3 机器人姿态测量.....	39
4.3.1 运动坐标系构建.....	39
4.3.2 旋转矩阵优化.....	40
4.4 本章小结.....	41
第 5 章 工业机器人运动视觉测量实验	42
5.1 实验平台搭建.....	42
5.1.1 系统硬件组成.....	42
5.1.2 软件环境.....	43
5.1.3 相机标定实验.....	44
5.2 机器人运动追踪与特征点定位实验.....	45
5.2.1 机器人运动追踪.....	45
5.2.2 特征点识别与定位.....	46
5.3 机器人全局结构运动视觉测量实验.....	48
5.3.1 机器人平面结构运动测量.....	48
5.3.2 机器人空间结构运动测量.....	50
5.3.3 机器人姿态测量.....	55
5.4 本章小结.....	56
第 6 章 总结与展望	57
6.1 总 结.....	57
6.2 展 望.....	58
参考文献	59
攻读硕士期间的主要成果	63
致 谢	64

第1章 绪论

1.1 课题研究背景与意义

至 21 世纪,我国制造业蓬勃发展,中国制造业规模不断扩大,俨然成为了亚洲乃至全球工厂。另一方面,随着工业智能化技术不断攀升,制造业正朝着高交互、高效率、高精度的发展与生产模式迈进。工业机器人在提高生产效率、保障产品一致性与质量、降低生产成本方面发挥着重要作用。我国相继出台了多项符合自身发展战略的纲领文件,《智能发展“十二五”专项计划》将高端制造业列入战略性新兴产业,《中国制造 2025》明确指出要以智能制造作为发展主攻方向,并重点突破航空航天、生物医药、高档数控机床和机器人等高新技术领域。同时,《新一代人工智能发展规划》与《机器人产业发展规划(2016-2020 年)》也明确指出,要发展智能机器人新型产业,在工业机器人领域攻克工业机器人关键技术。这些政策文件的制定表明我国对于工业机器人发展的重视,并为工业机器人的发展提供了政策支持和指导。



图 1-1 工业机器人应用场景

目前,在制造业、电子工业、物流运输中,工业机器人广泛应用于零件装配、部件焊接、物流仓储等任务。机器人的高速度、精确度和可编程性,对于处理小尺寸和复杂的电子组件至关重要,并且能够减少人工操作过程中的安全风险。

工业机器人的发展趋势将朝向协作化、灵活化、精细化,以更好地适应复杂多变的应用环境与大规模集成生产的需求。传统的工业机器人存在着缺乏感知和判断能力^[1]的局限性,导致无法适应变化的工作环境,也难以做出灵活的决策。另一方面,在高负载或进行高强度作业时,机器人的机械结构可能发生微小的变形、磨损和松动,这些因素会导致关节传动误差的积累,从而影响机器人的重复定位精度与绝对定位精度^[2]。在早期,解决该问题的主要方法是增强机器人本体结构性能和机器人装配精度,以实现加工制造阶段的误差预防。但该方法并不能解决工作磨损带来的精度损失问题,也无法提升机器人感知和判断的能力,适用性有限。

为此,研究人员提出了误差梳理的机器人精度补偿技术^[3]。该技术主要包含两

条路线,路线一为在线补偿方法,在机器人系统中融合多类型的高精度传感设备,获取机器人运动状态数据,形成机器人实时感知能力,以此建立闭环控制而实时优化机器人的运动状态。主要的伺服控制包括视觉与位置伺服^[4,5]、速度与加速度伺服^[6,7]以及力与扭矩伺服^[8]等;路线二是对机器人运动学模型进行标定,该方法通过确定机器人的几何参数和关节参数,以精确描述机器人的运动学模型。目前常见的运动学标定方法包括 D-H 法^[9]、MD-H 法^[10]、CPC 模型^[11]等。通过进行运动学标定,可以获得机器人的正逆运动学方程,从而实现机器人的精确控制和定位。

为了提高机器人运动控制精度,不论是采用在线补偿技术还是机器人运动学标定方法,都需要准确识别机器人的运动状态参数。因此,需要研究实现机器人运动精确测量的方法,以获取机器人运动过程中的状态参数,从而优化和调控机器人的运动精度。

本课题来源于国家自然科学基金青年基金项目“工业机器人 3D 全结构刚柔耦合交互特性视觉测量方法研究”(52005003),旨在开展针对机器人全局结构运动的视觉测量方法研究。视觉测量是一种非接触、全局测量方式,具有适应性好、抗电磁干扰能力强以及采集所得信息丰富等优点,可用于获取机器人结构空间多点同步运动数据,能够为进一步探究机器人刚柔耦合因素交互作用下的特性、优化机器人性能提供基础。

1.2 机器人运动测量技术研究现状

1.2.1 机器人运动测量技术

在早期机器人运动检测中,为检测机器人运动过程的加速度、负载受力和力矩信息、感知机器人振动特性,各类接触式传感器如加速度传感器、力传感器、编码器,惯性传感器等被安装在机器人的关节、连接杆或末端等部位,从而完成机器人结构参数标定、运动修正等工作。刘运毅^[12]等人针对机器人高速运动状态下的结构扰动,通过六维力传感器对机器人末端工具的力矩负载进行了测量,同时,以机器人结构力学关系在线计算末端重力负载参数,并基于极大似然估计方法,抑制机器人在高速运行的状态所产生的高噪振动对参数标定的影响,最终完成了对工业机器人末端负载的在线标定。

机器人底座安装倾角、末端执行器负载、重心坐标等参数,是机器人参数标定的重要内容。在一些较高精度的机器人工况下,传感器零点的偏移对测量偏差的影响是不可忽略的。为此,张春涛^[13]等学者重点考虑了六维力传感器的“零飘”误差,提出了零点漂移在线补偿策略,采用最小二乘法确定机器人力学基本参数,并根据机器人实时运动参数,建立了机器人运动的时序控制模型,来修正末端执行器重力

的影响和传感器零点的偏移,该方法能够抑制零点偏移误差,进一步提高测量准确度。

机器人在负载状态下的精确定位备受关注,在一些高负载状态下的作业中,会放大由高外力和扭矩而产生的机器人结构柔度特性,导致运动偏差。对此,Klimchik^[14]等学者在机器人每个驱动关节上使用双编码器,通过工具位置测量值和致动关节中的双编码器信息组合来识别刚度模型参数,并在此基础上,提出了智能顺应性误差补偿的策略,该策略使用测量过程的机器人运动体系结构,通过初级和次级编码器信号的线性组合来创建致动器反馈。所开发的技术被应用于工业机器人的刚度建模和精度补偿,降低了外部载荷和重力对定位精度的影响,机器人运动的柔度误差降低了 75%。

受机器人部分柔性构件所产生的弹性变形和非线性误差等因素的综合影响,机器人在运动到靠近奇异位置附近时,会产生自激振动现象。这种振动会降低机器人的定位和轨迹精度,同时加剧机器人结构的疲劳损伤。针对该问题,余龙焕等^[15]使用加速度传感器获取机器人加速度信号,并结合伺服电机信号实现闭环反馈,在此基础上,考虑加速度信号滤波以及信号传输和处理等因素引起的相位滞后,采用移相技术,提出加速度反馈的控制算法对机器人振动进行抑制和调控。

此外,拉线式传感测量系统、球杆仪也常用于监测机器人末端的姿态和位置信息,实现机器人运动标定与控制修正^[16-19]。当机器人执行运动时,拉线传感器可以感知和记录关键点的位置,从而帮助机器人进行运动轨迹修正;球杆仪通过测量机器人末端执行器的倾斜角度和方向,来获取机器人的姿态信息,进而可调控机器人运动姿态。

基于接触式传感器的机器人状态检测方法的普适性较好、检测成本低,能够适用于常见的测量场景。但接触式传感器存在信号传递延迟、零飘、噪声扰动大等不足,此外受传感器额外载荷影响,原始运动状态已发生改变,导致测量结果的精度降低、可靠性变差。因此,有必要采用一种既不改变机器人结构载荷,也能完成对机器人运动参数精确测量的方法,提高测量效率和准确度。

相较接触式测量方法,非接触式测量无需与被测对象接触,不会影响机器人结构的动态特性,可实现较高的测量精度。目前,非接触式测量多基于激光跟踪仪对运动目标测量,其原理是利用激光器发射出的激光束照射到目标物体表面,然后接收激光反射回来的信号,通过测量激光的时间延迟或相位变化等信息来计算目标物体与激光器之间的距离或位置。这种方式可以实现高精度的跟踪和测量,广泛应用于机器人位姿、运动追踪等领域。史晓佳^[20]等使用激光跟踪仪测量工业机器人位姿,并使用所设计的闭环控制系统实现了机器人运动误差的在线补偿,将机器人的绝对定位精度提升至 100um 以内;Zhao^[21]等使用激光跟踪仪识别工业机器人的

非线性残差，通过机器人控制器与激光跟踪器实时通讯的方式完成了对机器人的运动校准；Toni 等^[22]在运用激光跟踪仪测量的同时融合了 MIU 惯性测量单元，旨在得到更为全面的机器人运动状态信息。此外，激光干涉仪、三坐标测量仪也常用于机器人的运动学标定^[23-24]。

激光测量技术具有测量精度高、频带范围宽等优点，但是采用激光仪器测量过程中需要在被测结构上安装光学镜组，并且每套镜组只能测量一个位置点的运动信息。如果需要测量多个点或实现全局测量，则需要多套激光测量设备和靶球镜组，这受传感器安装空间限制，并且成本高昂。

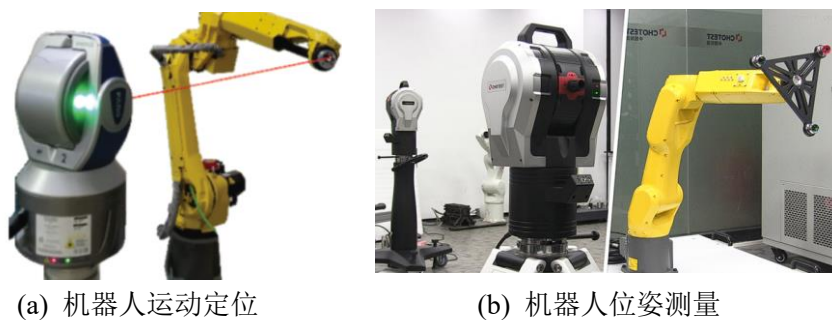


图 1-2 激光跟踪仪

1.2.2 视觉测量技术

随着相机光感元件制造精度的提升，以相机为主要传感元件的视觉测量技术在测量领域的应用范围不断扩大。视觉系统作为一种非接触式测量技术，具有以下优点：

(1) 测量过程可视化：视觉系统能够以图像的形式输出测量结果，通过直观的图像展示，方便用户进行观察和分析，使测量结果更容易理解。

(2) 高交互性：视觉系统能够融合其他多类传感器，以达到自动化测量和检测，实现高效、快速的自动化生产过程控制和质量检测，对提高生产效率和产品质量有很大的帮助。

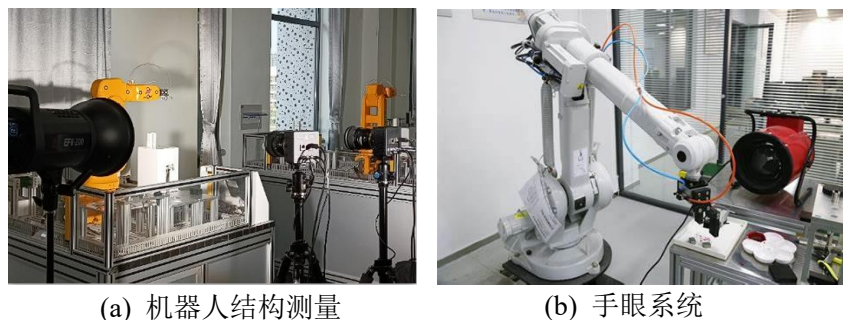


图 1-3 视觉检测技术

(3) 强泛化能力：视觉系统能够实现多维测量和非局限性测量，不受被测物体的形状、大小和位置等因素限制，具有较好的灵活性和通用性。

(4) 具备全局观测能力: 视觉系统无需在被测结构上安装复杂的传感单元, 每个像素点都能够作为被测点, 从而实现目标的全局测量。

视觉系统能够有效检测出目标的结构运动状态, 已经广泛应用于机器人领域的多个研究方向。在结构健康检测方面^[25], 视觉系统可以通过对土木结构表面的图像数据进行处理, 实现对结构变形、裂纹等损伤状态的检测和诊断。在点云重构^[26]方面, 视觉系统可以通过对周围环境的图像数据进行处理, 获取目标结构表面信息; 在机器人运动目标检测方面^[27], 视觉系统可以通过对机器人周围环境的动态图像数据进行处理, 实现对运动目标的跟踪和识别, 从而实现机器人智能感知和行为决策。

视觉系统能够获取目标的全局运动信息, 这为机器人运动观测提供了有力支持。严宏志等^[28]为修正几何参数误差对机器人绝对定位精度的影响, 使用双目测量系统对机器人基坐标进行标定, 通过坐标系变换结果构建误差辨识模型, 对误差进行补偿; 黄磊等^[29]为降低机器人系统的线性误差对机器人末端运动轨迹的影响, 使用视觉系统测量机器人运动轨迹, 通过高速相机实时图像定位机器人实际运动轨迹, 进行机器人运动轨迹的动态补偿; 徐秀秀等^[30]关注机器人柔性结构的振动特征, 通过视觉测量系统对自制 CDR 柔性机器人结构振动进行观测, 提取图像中光点质心来获取机器人表面结构的振动数据, 从而得到柔性臂的一阶和二阶振动模态; 郑继贵等^[31]关注机器人高速工况下热效应对运动精度的影响, 使用双目视觉系统对安装在机器人末端的基准球进行观测, 获取机器人实际位姿信息, 通过标定参数对时间微分曲线进行分析, 构建了机器人温度误差补偿模型, 将机器人重复定位精度控制在 0.1mm 以内; 尚何江^[32]使用双目视觉系统对机器人末端姿态和关节参数进行辨识, 以机器人逆运动理论解为基础, 采用位姿分离的方法进行逆解分析, 通过微分运动原理和微小变换, 推导了关节参数误差与机器人末端位姿误差之间的误差映射矩阵, 由此建立了高效准确的实际逆解分析方案, 经过标定后的机器人绝对位姿误差降低了 86%。

机器人运动视觉测量涉及目标检测、追踪、定位等过程。早期, 研究人员将运动视为目标区域在整体图像上的像素变化, 通过将目标像素值作为目标的外观特征, 实现追踪过程。基于该理论, 帧间差分^[33, 34]、背景差分^[35]、光流法^[36, 37]被运用在一些较为简单的追踪场景中, 然而, 该类方法抗噪能力差, 所得追踪结果难以适应遮挡、高速、多目标的场景。为解决该问题, 研究人员提出了基于运动模型的追踪方法, 该模型根据已观测到的运动数据, 预测目标在未来时间和空间上的位置和速度。常见的运动模型包括线性模型和非线性模型, 基于运动模型, 衍生出了许多鲁棒性较高的目标特征形式, 包括梯度直方图特征 (HOG feature)、尺度不变特征 (shift feature)、角度特征 (Harris feature) 等, 并基于此建立了诸多目标追踪

策略,如模板匹配^[38-40]、均值漂移^[41,42]等,均能实现目标有效追踪。为解决目标遮挡的问题,研究人员还将粒子滤波^[43]和卡尔曼滤波^[44]等算法用来提升视觉系统追踪运动目标的抗遮挡能力。

随着人工智能技术兴起,研究人员尝试运用机器学习与深度学习算法来提升视觉系统对运动目标的测量感知能力。随着 CNN 为主要结构的目标检测与追踪算法适用域不断扩大, BB8^[45]、YOLO-6D^[46]等算法被运用在机器人位姿估计中。该类算法通过提取目标锚框顶点坐标构建 2D-3D 点对信息,并基于 PNP 理论实现位姿测量,然而受限于锚框匹配的不确定度,所得结果的位姿精度难以得到保障。Chen^[47]等学者将 PNP 残差反馈给网络结构,引导参数实时更新,以提高测量精度,然而 PNP 测量模型中存在微分奇点,可能导致数据集训练困难或者不收敛。

在同一时期,Kehl^[48]等学者将 SSD 网络结构扩展到 3D 运动检测,提出了 SSD-6D 系列的位姿检测模型,该模型将锚框内的目标姿态信息分解为离散点的位置信息与 2D 视角下的旋转,最后通过投票机制得到得分最好的位姿估计值,该类回归模型无需考虑 PNP 模型的奇点问题,但检测效果依旧受锚框匹配的不确定度的制约。与该算法类似, PoseCNN^[49]、Fast R-CNN^[50]以及使用 FSM 机制融合 CNN 和 PointNet 的网络结构^[51]均在机器人运动检测中得到运用。

可以发现,基于深度学习的运动目标检测方法,能够基于预先训练数据的驱动,较为准确地提取目标的运动轨迹和运动状态。此外,该类方法能够同时对图像中不同目标实现同时序追踪,这与机器人全局结构运动所需的多点观测的需求相契合。然而,深度学习网络所得的观测数据并不能直接实现特定点的位置精确定位,需要额外的点特征像素定位算法进行补偿。

1.3 本文研究内容与组织结构

准确测量机器人运动是评估机器人动静态性能的重要内容。机器人结构运动的位置与姿态信息能够直观反映机器人运动精度、承载能力与结构稳定性。本文重点关注视觉测量技术在机器人结构运动测量方面的运用,针对机器人运动视觉测量,搭建了视觉测量系统,设计了机器人结构运动视觉检测方案,完成了机器人结构运动追踪与测量。

本文理论部分围绕机器人全局结构运动视觉测量这一主题,具体内容包括:构建了视觉系统测量方案,推导了相机成像与标定模型,建立了在几何空间中机器人运动的状态表达模型;采用 YOLO v5 算法实现机器人运动追踪,给出了合作标识特征点像素坐标精确定位的流程;提出了基于单应性特征矩阵的机器人结构平面运动视觉测量方法;推导了双目视觉三维重构模型,并提出了基于 RANSAC 误差加权约束的机器人结构空间运动测量方法,优化坐标变换模型求解精度;构建了运

动坐标系，推导了旋转矩阵优化方法，实现了机器人姿态测量。

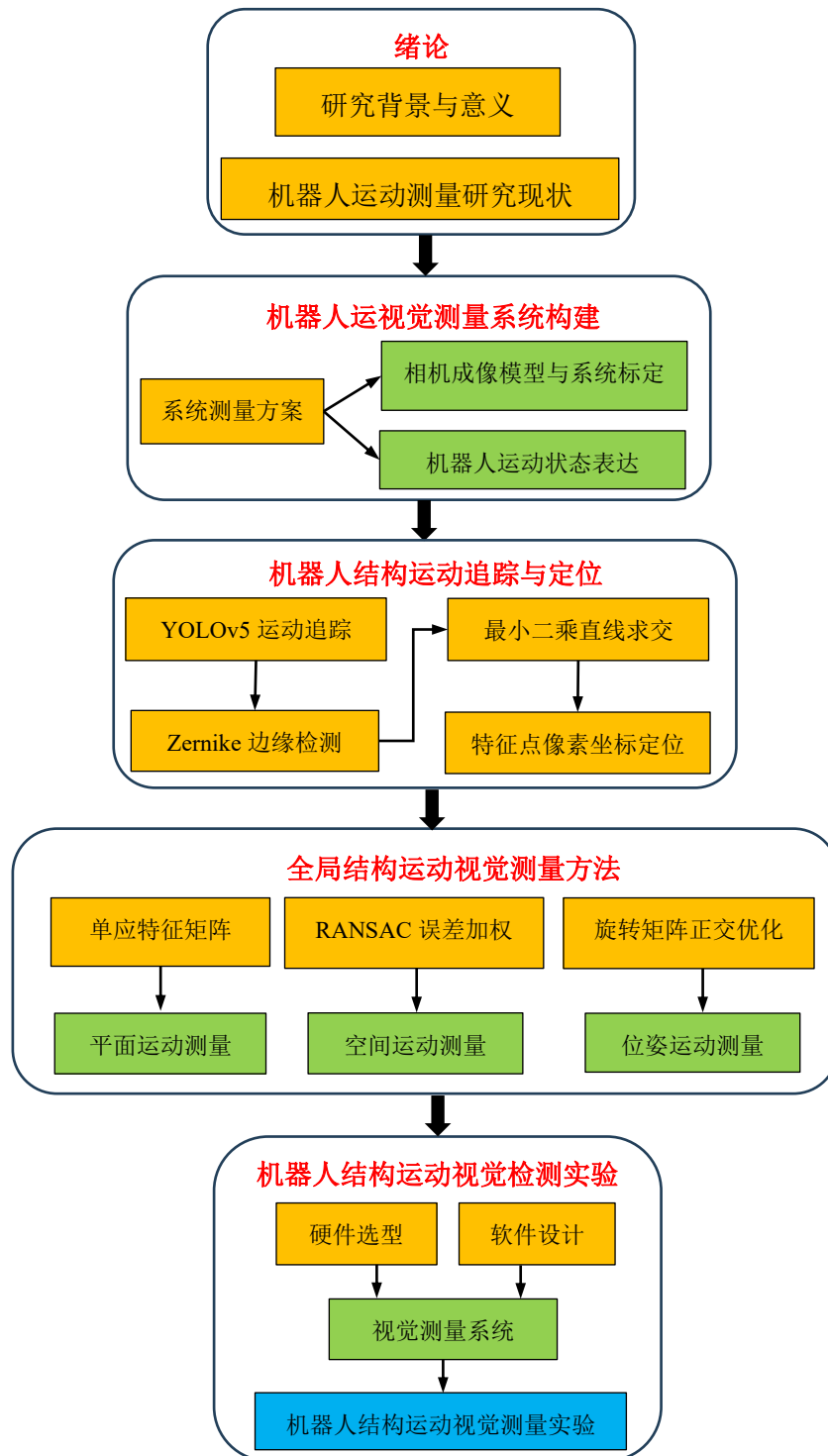


图 1-4 文章组织结构图

本文实验部分包括：搭建了视觉测量系统；设计了特定的合作标识实现对机器人运动追踪与量化；基于 YOLOv5 目标检测框架，实现机器人合作靶标的运动追踪；基于亚像素边缘特征点拟合求交，得到特征点像素坐标的精确位置；基于所提测量模型与方法完成机器人全局结构平面运动、空间运动视觉测量以及机器人姿态测量。文章的组织结构如图 1-4 所示，具体内容如下：

第一章：绪论。介绍了选题背景、课题研究的价值与意义；概述了机器人运动测量的主要方式，以及现阶段基于视觉测量的各类方法、研究概况及其优势和局限性；最后论述了本文研究的主要内容和论文的组织结构。

第二章：机器人运动视觉系统构建。搭建了机器人结构运动测量方案，推导了相机成像模型与参数标定的相关理论方法，阐述了机器人运动位置与姿态的表达方法。

第三章：机器人运动视觉追踪与特征点定位。设计了合作标识，使用 YOLOv5 算法完成了标识的运动追踪；研究了特征点亚像素定位方法，实现了特征点高精度像素定位；为后续机器人空间运动测量奠定基础。

第四章：机器人平面运动测量与空间位姿测量。为纠正平面测量中，相机光轴与测量平面不正交产生的测量误差，提出了基于单应性特征矩阵平面运动测量方法，推导了所提方法的测量模型，给出了平面运动测量流程；为实现机器人空间运动测量，推导了双目视觉三维运动测量模型，实现运动求解；为抑制 3D 坐标变换中测量噪声对模型精度的扰动，提出了基于 RANSAC 误差加权约束的坐标转换优化方法；给出了机器人三维位姿测量流程，运动坐标系构建方法和姿态解析表达。

第五章：机器人运动测量实验。搭建了视觉测量系统，以三自由度和六自由度机械臂为实验测量对象展开了综合实验。基于 MATLAB 软件平台完成了相机标定，基于 Python+pytorch 的虚拟环境下完成了 YOLOv5 算法的部署与应用，在此基础上，开展了特征点亚像素定位实验、机器人平面运动测量实验和三维位姿测量实验，验证了所提方法的可行性和测量效果。

第六章：总结与展望。总结了本论文的主要工作与成效，说明了研究存在的不足，指出了未来的研究方向。

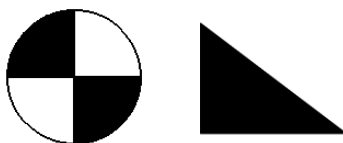
第2章 机器人运动视觉测量系统构建

为了实现机器人结构运动的观测，本章首先提出了基于视觉的测量方案；其次，由几何空间映射入手，详细推导相机的成像模型与参数标定方法；最后，对机器人的运动表达形式进行了阐述和分析。

2.1 视觉测量系统方案

视觉系统可分为图像采集模块、追踪与定位模块以及数据分析模块。其中，图像采集模块由高速相机和辅助光源组成，以适应较快速的机器人空间运动过程，收集机器人运动图像；追踪与定位模块负责提取机器人结构表面标识特征点的像素坐标；数据分析模块则利用预先标定的系统模型参数，将特征点的像素坐标恢复到物理空间中，从而获取和分析机器人在运动过程中的观测数据，实现对机器人全局结构运动的测量与评估。

视觉系统测量机器人运动的难点在于，机器人结构表面缺乏能够为视觉系统所识别的图像特征，比如圆、线边缘、角点、交点等几何特征。因此，需要在机器人结构表面布置人为设定的特征标识，明确视觉系统所检测的几何特征，提供视觉系统易于识别的、稳定的图像特征信息，从而降低视觉系统的测量难度。为此，本文设计了两种合作标识，如图 2-1 所示，其中，图 2-1 (a)所使用的十字标识用于机器人平面与空间运动的追踪与测量，其特征点为标识内黑白相间所形成十字线交点，该标识图 2-1 (b)所使用的三角形标识用于机器人六自由度位姿测量。两种标识被粘贴在机器人结构表面，视觉系统通过追踪定位标识特征点的坐标，达到测量机器人全局结构运动的目标。



(a) 圆形十字标识 (b) 三角形标识

图 2-1 所使用的合作标识

与几何特征对应的识别方法有圆心拟合、霍夫直线拟合求交、角点检测。其中，圆心拟合将标识圆心作为特征点，通过拟合外圆轮廓，进而计算圆心坐标。圆心拟合方法易受椭圆畸变影响，机器人发生位姿变化时，所提取的特征点不确定度急剧增大，鲁棒性差。霍夫直线拟合求交将两十字线的交点作为特征点，能够有效的适应位姿变化，但霍夫直线拟合算法的投票机制的约束能力差，需要繁杂的人工调参，并且抗光变能力弱。同样的，角点检测（Harris）及其改进方法（SIFT，SURF）类

别的点特征检测法对噪声敏感，单位标识图像上可能会得到多个错误特征点，并不适用于特定位置的特征点提取。

为准确观测机器人的运动，系统采用深度学习的网络框架，目标是实现对机器人运动的快速观测，并能够在多个观测点对其进行追踪，同时结合轮廓边缘线特征识别方法，拟合标识出十字线边缘与三角形各边，并求直线交点，定位对特征点的像素坐标，以追踪标识运动并实现精确定位的目的。这种方法结合了深度学习和图像处理技术，可以在面对观测不确定性较大的机器人运动过程时，提供较为鲁棒的运动追踪与定位性能，为机器人在复杂环境下的精确定位与运动分析给予有效支持。

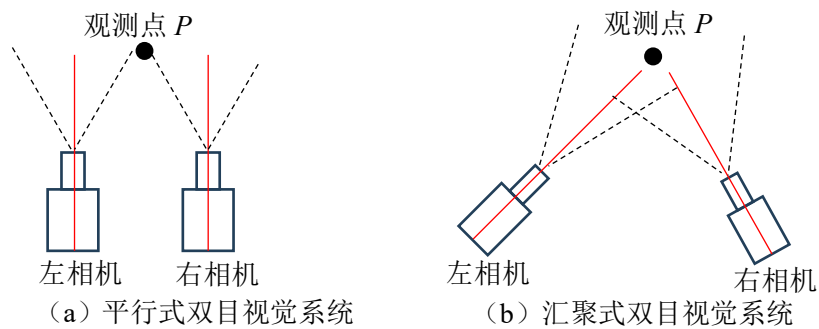


图 2-2 双目视觉系统结构类型

观测机器人在某一平面上的运动过程，可直接使用单个相机组成的单目视觉系统对机器人平面运动进行观测，完成机器人结构平面运动。而用于测量机器人于空间运动时，需使用两个参数一致的相机组成双目视觉系统，其常见的组合方式可分为平行式与汇聚式，如图 2-2 所示。较之平行式系统对安装精度的高要求，汇聚式系统具有测量范围广、深度感知精确、安装限制低等优势，更适合于测量机器人空间运动。使用合作标识完成机器人运动追踪与定位后，即可采用汇聚式双目系统，构建测量模型，将像素坐标恢复到物理坐标，完成机器人空间运动以及姿态的测量。

2.2 相机成像模型与参数标定

2.2.1 3D 到 2D 几何空间映射

相机捕捉 3D 物理空间的目标，将其转化为 2D 空间的图像信息。这种信息传递方式通过几何空间中的投影关系来实现。在不同的投影类型中，正交投影是最简单的模型，它通过垂直于观察平面（即投影平面）的线将物体上的点投影到平面上。然而，正交投影的约束条件非常严格，它不考虑透视效果，也就是说无论物体距离观察者远近如何，其大小和形状在投影过程中都不会发生变化，从而保持了物体的真实尺寸和比例。3D 的点 P 到 2D 的投影 p 的模型可表示为：

$$p = [\mathbf{I}_{2 \times 2} | 0] P \quad (2-1)$$

对应的齐次坐标的表达式为：

$$p = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} P \quad (2-2)$$

当物体与观测平面不正交，但满足平行于到物体中心的直线进行投影，式(2-2)可改写为：

$$p = \begin{bmatrix} a_{00} & a_{01} & a_{02} & a_{03} \\ a_{20} & a_{11} & a_{12} & a_{13} \\ 0 & 0 & 0 & 1 \end{bmatrix} P \quad (2-3)$$

该投影模型为类透视投影。可以发现，类透视投影约束条件宽泛，比正交投影更加具有泛化能力。继续解除约束，当物体与观测平面既不不正交，也不满足平行于到物体中心的直线进行投影，此时 3D 中的基元在 2D 投影面上的大小会随着距 2D 投影面的远近而发生改变，这一性质可称为尺度缩放。而在相机模型中，该性质可以基于相机焦距和被观测物体的距离来进行线性表达：

$$x = f \frac{X_c}{Z_c}, \quad y = f \frac{Y_c}{Z_c} \quad (2-4)$$

式中， (X_c, Y_c, Z_c) 表示物点在相机坐标系下的坐标， (x, y) 表示投影到图像坐标系下的坐标。同时，图像坐标系 O_{xy} 与像素坐标系 O_{uv} 的 2D 平移关系可表示为：

$$O_{uv} = O_{xy} + t(u_0, u_0) \quad (2-5)$$

则相机坐标系到像素坐标系的透视变换模型为：

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \frac{1}{Z_c} \begin{bmatrix} f_x/dx & 0 & u_0 \\ 0 & f_y/dy & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} \quad (2-6)$$

式(2-6)所给出的模型反映了质点的相机观测表达，此时，相机坐标系定义可为观测坐标系，那么，物体所在的物理坐标系到观测坐标系的转换关系可表示为如下形式：

$$P_c = T_c^w P_w \Leftrightarrow \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (2-7)$$

因此，完整的相机成像模型的数学表达为：

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \frac{1}{Z_c} \begin{bmatrix} f_x/dx & 0 & u_0 \\ 0 & f_y/dy & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{R}_{3 \times 3} & \mathbf{t}_{3 \times 1} \\ \mathbf{0}_{1 \times 3} & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \Leftrightarrow \mathbf{p} = \lambda \mathbf{K} \mathbf{M} \mathbf{P}_w \quad (2-8)$$

其中，外参矩阵 $\mathbf{M}$ 表示当前相机坐标系与物理坐标系的变换关系，外参矩阵随着与相机位置，物理坐标系定义而发生变化。 λ 表示尺度因子， $\mathbf{K}$ 表示相机的内参矩阵。

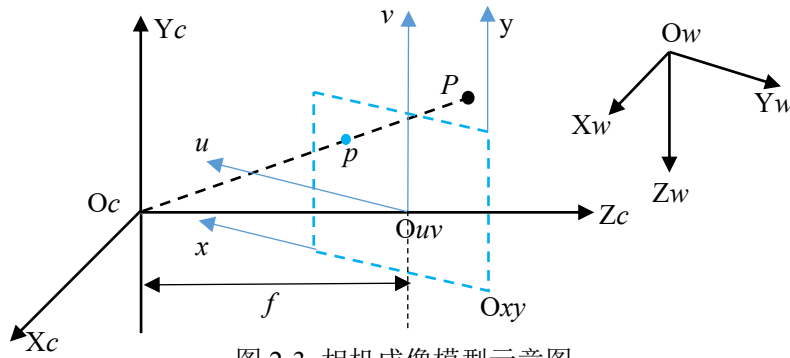


图 2-3 相机成像模型示意图

2.2.2 单目相机参数标定

当 3D 空间的物体能够被相机捕捉，那么在相机观测视角中，物体在像素坐标上的 2D 坐标往往是已知的。因此，对于式(2-8)的模型而言，确定相机内参 $\mathbf{K}$ 的精确解是完成测量的关键。为此，学者通常使用一个标定模型或者标定集合来估计相机的内参。

对于较小的运动标定设备，常见的标定物是在空间中摆放一个具有集合约束的 3D 标志物，通过运动相机至若干个被约束的位置点，求解成像模型，这种主动标定法标定过程简单，但通常无法精确计算标志物在 3D 空间的姿态信息，导致标定结果不收敛，适用范围有限。

当拥有精确的运动控制平台时，可控制平台来施加已知的运动约束，达到标定目的。视觉系统基于该约束求解成像模型，无需额外的标定物，然而，这种标定模式非常依赖运动约束的精确度，需要昂贵的标定设备。

为此，学者们开始寻求一种兼具标定精度和灵活性的标定模型。不难发现，各种标定模式的前提是在观测空间中建立起可控的约束，从而使得某些像素坐标的唯一 3D 射线能够匹配到空间对应的几何约束，在此基础上，通过相机基本成像模型来构建方程组，完成内参矩阵 $\mathbf{K}$ 的求解。基于该原理，N 平面标定法被验证并形成了一种新的标定模式，该标定模式依托于二维标定板，标定板上预先绘制了多

个确定几何约束的特征点，以此来建立起标定所需的解析方程组。同时，在计算过程中考虑了内参矩阵 $\mathbf{K}$ 的建模线型误差和径向畸变等非线性误差。具体的标定流程为：通过图像识别方法定位标定板上的特征点的像素坐标，并基于特征点的像素坐标和参考坐标系下的真实尺寸坐标建立求解方程组。通过固定相机，移动标定板，就能够不同标定图像对应的若干个平面约束。

通过该标定流程，每一个平面图像能够建立一个形如下式的解析方程：

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{K} \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{t} \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \quad (2-9)$$

令 $\mathbf{H} = \frac{1}{s} \mathbf{K} \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{t} \end{bmatrix}$ ，则式(2.9)可化为如下形式的单应特征相关式：

$$\mathbf{m} = \mathbf{H}\mathbf{M} \quad (\mathbf{H} \in \mathbb{R}^3) \quad (2-10)$$

可以发现，平面标定模式将相机内外参融合为了一个单应性矩阵，在该条件下，基于式(2-10)可以很容易得到 $\mathbf{H}$ 的正向解，但由 $\mathbf{H}$ 完成内参 $\mathbf{K}$ 与外参 $\begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{t} \end{bmatrix}$ 的解耦过程十分棘手。针对该问题，Zhang^[52]给出了一种线性约束方法，通过构建矩阵平方根和逆恢复模型来完成标定矩阵 $\mathbf{K}$ 的解耦过程。将单应矩阵 $\mathbf{H}$ 的每一行分解为 $\mathbf{H} = [\mathbf{h}_1 \quad \mathbf{h}_2 \quad \mathbf{h}_3]$ ，则 $\mathbf{H}$ 和 $\mathbf{K}$ 的关系可由下式表示：

$$\begin{cases} \mathbf{h}_1 = \lambda_1 \mathbf{K} \mathbf{r}_1 \\ \mathbf{h}_2 = \lambda_1 \mathbf{K} \mathbf{r}_2 \\ \mathbf{h}_3 = \lambda_1 \mathbf{K} \mathbf{t} \end{cases} \quad (2-11)$$

又因为 $\mathbf{r}_1$ 与 $\mathbf{r}_2$ 正交，则式(2-11)可变换为：

$$\begin{cases} \mathbf{h}_1^T \mathbf{K}^{-T} \mathbf{K}^{-1} \mathbf{h}_2 = 0 \\ \mathbf{h}_1^T \mathbf{K}^{-T} \mathbf{K}^{-1} \mathbf{h}_1 = \mathbf{h}_2^T \mathbf{K}^{-T} \mathbf{K}^{-1} \mathbf{h}_2 \end{cases} \quad (2-12)$$

式(2-12)即为内参矩阵的封闭的线型约束。在此基础上，构建一个对称矩阵

$$\mathbf{B} = \mathbf{K}^{-T} \mathbf{K}^{-1} = \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \\ B_{31} & B_{32} & B_{33} \end{bmatrix} \text{ 带入(2-13)可得:}$$

$$\begin{cases} \mathbf{h}_1^T \mathbf{B} \mathbf{h}_2 = 0 \\ \mathbf{h}_1^T \mathbf{B} - \mathbf{h}_2^T \mathbf{B} = 0 \end{cases} \quad (2-13)$$

因为 $\mathbf{B}$ 为对称矩阵，因此 $\mathbf{B}$ 元素满足 $B_{12} = B_{21}$, $B_{13} = B_{31}$, $B_{23} = B_{32}$ ，令线型方程的中间参量 $\mathbf{b} = [B_{11}, B_{12}, B_{22}, B_{13}, B_{23}, B_{33}]$ ，则式(2-13)可整理为齐次方程组如式(2-14)所示：

$$\begin{bmatrix} J_{12} \\ J_{11} - J_{22} \end{bmatrix} \mathbf{b} = 0 \Rightarrow \mathbf{J} \mathbf{b} = 0 \quad (2-14)$$

对 $\mathbf{J}$ 采用奇异值分解法得到如下形式 $\mathbf{J} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$, 则 $\mathbf{b}$ 的最小二乘解为 $\mathbf{V}^T$ 的最后一列向量。进一步, 使用 Cholesky 因式分解的逆恢复模型解析等式 $\mathbf{B} = \mathbf{K}^{-T}\mathbf{K}^{-1}$, 进一步得到对应相机内参的解析式如下:

$$\begin{aligned} v &= (B_{12}B_{13} - B_{11}B_{23}) / (B_{11}B_{22} - B_{12}^2) \\ \lambda_1 &= B_{33} - [B_{13}^2 + v(B_{12}B_{13} - B_{11}B_{23})] / B_{11} \\ f_x &= \sqrt{\lambda_1 / B_{11}} \\ f_y &= \sqrt{\lambda_1 B_{11} / (B_{11}B_{22} - B_{12}^2)} \\ \gamma &= -B_{12}\alpha^2\beta / \lambda_1 \\ u &= \gamma v / \alpha - B_{13}\alpha^2 / \lambda_1 \end{aligned} \quad (2-15)$$

将 $\mathbf{K}$ 带入式(2-13), 外参旋转矩阵 $\mathbf{R}$ 可基于下式确定:

$$\begin{cases} \mathbf{r}_1 = \lambda_2 \mathbf{K}^{-1} \mathbf{h}_1 \\ \mathbf{r}_2 = \lambda_2 \mathbf{K}^{-1} \mathbf{h}_2 \\ \mathbf{r}_3 = \mathbf{r}_1 \times \mathbf{r}_2 \end{cases} \quad (2-16)$$

式中, $\lambda_2 = 1 / \|\mathbf{K}^{-1} \mathbf{h}_1\| = 1 / \|\mathbf{K}^{-1} \mathbf{h}_2\|$ 。进一步可求得 $\mathbf{R} = [\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3]$, $\mathbf{t} = \lambda_2 \mathbf{K}^{-1} \mathbf{h}_3$ 。

理论上, 根据式(2-9)到式(2-16), 即可求得相机成像的内外参数, 然而, 该标定模式的参数求解精度受限于点线几何约束的规范程度, 具体的误差源于镜头畸变和内参矩阵 $\mathbf{K}$ 线型约束的封闭残差。对标定影响较大的镜头畸变类型有切向畸变和径向畸变, 如图 2-4 所示。

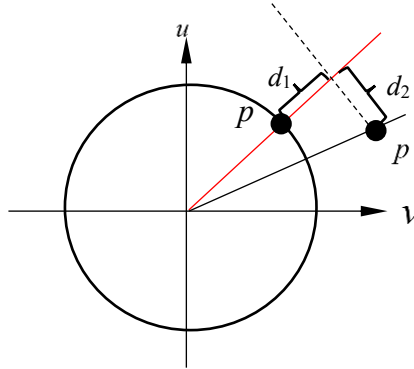


图 2-4 镜头畸变示意图

其中, 切向畸变 d_2 产生的原因在于, 镜片的装配误差使得它们的光轴不严格共线, 就目前镜片的安装工艺精度而言, 该误差畸变的影响远低于径向畸变; 径向畸变 d_1 产生的原因在于, 光学镜片沿径向的曲率变化导致像点沿着镜头的径向移动, 从而产生误差, 径向畸变误差与镜片的厚度有关, 无法完全消除。因此, 本研究只考虑径向畸变的非线性误差。根据 Polynomial Model 多项式模型得到相机的二阶径向畸变方程为:

$$\begin{cases} x = x(1 + k_1 r^2 + k_2 r^4) \\ y = y(1 + k_1 r^2 + k_2 r^4) \end{cases} \quad (2-17)$$

式中, $r^2 = x^2 + y^2$, 根据图像坐标系到像素坐标系的变换关系可得 $\begin{cases} u = x + u_0 \\ v = y + v_0 \end{cases}$, 同理

有 $\begin{cases} u = x + u_0 \\ \tilde{v} = y + v_0 \end{cases}$, 带入式(2-17)可得:

$$\begin{cases} u = u + (u - u_0)(1 + k_1 r^2 + k_2 r^4) \\ \tilde{v} = v + (v - v_0)(1 + k_1 r^2 + k_2 r^4) \end{cases} \quad (2-18)$$

式(2-18)化为非齐次线型方程矩阵形式如下:

$$\begin{bmatrix} (u - u_0)r^2 & (u - u_0)r^4 \\ (v - v_0)r^2 & (v - v_0)r^4 \end{bmatrix} \cdot \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} = \begin{bmatrix} u - u \\ \tilde{v} - v \end{bmatrix} \Leftrightarrow \mathbf{LD}_k = d \quad (2-19)$$

式(2-19)给出了单一点线约束的非线性误差模型解析, 而对于平面标定模式, 多个标记点构成的点线约束能够使得(2-19)组成超定方程组, 继而, 径向畸变系数的最小二乘解可表示为:

$$D = (\mathbf{L}^T \mathbf{L})^{-1} \mathbf{L}^T d \quad (2-20)$$

不难发现, 在模型解析的过程中, 大多数参数的解是不封闭的, 基于超定方程组估计所得的参数值为孤立优解, 并不能直接用于目标测量, 需要通过优化方法使得各参数在全局上到达最优, 该问题的有效办法是基于使用初始解的标定参数, 将标定板各特征点投影到像素空间下, 将投影所得的像素坐标与理论坐标之间的距离作为目标函数, 基于非线性优化方法使得重投影误差最小化, 从而逼近各参数的最优解。该优化的目标函数可表示为:

$$F = \sum_{i=1}^n \sum_{j=1}^m \|I_{ij} - I_{ij}(\mathbf{K}, k_1, k_2, \mathbf{R}, \mathbf{t})\| \quad (2-21)$$

式中, I_{ij} 表示重投影对应的坐标值, I_{ij} 表示特征点的理论坐标值。

2.2.3 双目相机参数标定

双目视觉标定是在单目标定的基础上, 对一组单目相机所构成的立体视觉系统进行标定, 确定两个相机之间的外部参数(如旋转和平移), 标定立体视觉系统外参是实现深度测量和三维重建的基础。

使用两个相机对 3D 空间中同一目标 P_w 进行观测，两个相机观测空间的几何转换关系可表示为：

$$\begin{cases} P_l = R_l P_w + t_l \\ P_r = R_r P_w + t_r \end{cases} \quad (2-22)$$

消融 P_w 于式(2-22)中有：

$$P_r = R_r R_l^{-1} (P_l - t_l) + t_r \quad (2-23)$$

以左相机为基准相机，则右相机到左相机的空间变换关系可表示为：

$$P_r = R_l^r P_l + t_l^r \quad (2-24)$$

联合式(2-23)，(2-24)即可得左右相机之间的位姿关系 R_l^r ， t_l^r 的函数表达：

$$\begin{cases} R_l^r = R_r R_l^{-1} \\ t_l^r = t_r - R_l^r t_l \end{cases} \quad (2-25)$$

2.3 机器人运动状态表达

2.3.1 机器人运动位置表达

机器人运动可以通过笛卡尔空间的正交矩阵 $SE(3)$ 进行直观描述。当将机器人末端执行器（工具）视为一个几何基元时，该基元的位置和姿态信息就可以由一个固定的坐标系进行度量。通常情况下，这个固定坐标系为机器人基坐标系。如图 2-5 所示，机器人末端执行器所在的运动坐标系为 $\{F\}$ ，机器人基坐标系为 $\{B\}$ ，该坐标系之间的位姿关系表示可为 ${}^B_F T$ 。

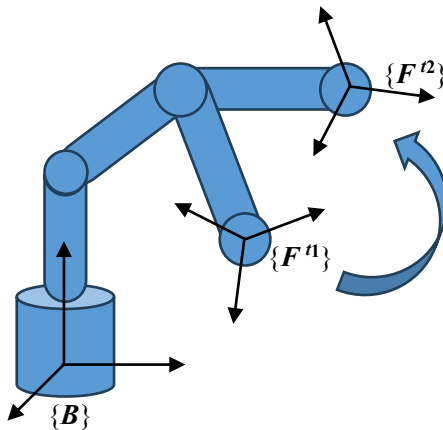


图 2-5 笛卡尔空间中的机器人运动表达

机器人的运动学模型可表示为：

$$x_{k+1} = f(x_k, u_k) + w_k \quad (2-26)$$

式中, x 表示物体位置, u 表示运动相关输入, w 表示运动噪声。将该运动扩展到如图 2-6 所示的三维空间中的刚性运动时, 其运动状态可解析某时刻在三维空间中的坐标变化, 对应的数学形如下式:

$$x_{k+1} = f_{SE(3)}(x_k, {}^B_F T) + w_k = {}^B_F R_k^{k+1} x_k + {}^B_F t_k^{k+1} + w_k \quad (2-27)$$

式中, ${}^B_F R_k^{k+1}$, ${}^B_F t_k^{k+1}$ 分别表示在 k 到 $k+1$ 时刻, 机器人末端执行器相对于机器人基坐标系的旋转运动量和位移运动量。

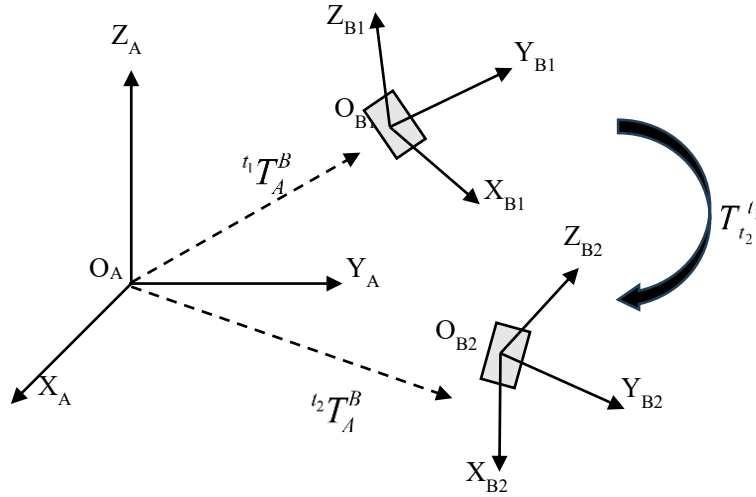


图 2-6 坐标系转换示意图

旋转矩阵的结构并不紧凑, 矩阵的 9 个分量仅表达了三个运动自由度的旋转, 这表明, 在计算的空间复杂度上, 旋转矩阵的表达形式并非最优。为此, 研究人员提出了旋转向量的空间旋转表达方式。

理论上, 任意物体的旋转均可视为物体绕一个旋转轴, 旋转任意角度。因此, 可以使用一个向量 n , 让向量的方向与选择轴保持一致, 向量的长度等于旋转角 θ , 这物体的旋转可由该向量进行表示。将该向量设定为旋转向量 θn 。旋转向量到 SO(3) 群元素的变换可由罗德里格斯公式(Rodrigues's Dormula)表示:

$$R = I \cos \theta + (1 - \cos \theta) n n^T + \sin \theta n^{\wedge} \quad (2-28)$$

式(2-28)中, I 为 3 维单位矩阵, $n^{\wedge}$ 表示 n 向量的反对称矩阵, 具体表示为:

$$n^{\wedge} = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix} \quad (2-29)$$

反向计算旋转矩阵到旋转向量的变换, 可先计算转角 θ , 对式(2-29)等式两边取矩阵迹可得:

$$tr(R) = \cos \theta tr(I) + (1 - \cos \theta) tr(n n^T) + \sin \theta tr(n^{\wedge}) \quad (2-30)$$

式(2-30)可化简为 $\text{tr}(\mathbf{R}) = 1 + 2\cos\theta$ ，因此，转角 θ 的解表示为：

$$\theta = \arccos \frac{\text{tr}(\mathbf{R}) - 1}{2} \quad (2-31)$$

使用旋转轴的几何性质，旋转轴方向上的向量不受旋转影响，由此可得到关系 $\mathbf{R}n = n$ ，因此，旋转轴 n 是矩阵 $\mathbf{R}$ 特征值为 1 对应的特征向量。

旋转向量的参数表达方式在数值计算中通常比旋转矩阵更加稳定，特别是在涉及到构建机器人运动学方程中数值优化和微分方程求解时，该参数化方式更为精确快捷。

2.3.2 机器人运动姿态表达

旋转矩阵与旋转向量提供了表示空间旋转的解析方法。但对于表达物理世界的个体的运动缺乏可视化，当给出 $\text{SO}(3)$ 群元素或旋转向量来对物体进行旋转时，并不能直观的体现旋转过程，这对描述机器人关节运动和建模过程带来了不便。为此，研究人员趋于将旋转过程拆分为绕参考坐标系不同轴的三次旋转。这种表示旋转的方式称为欧拉角。以 ZYX 顺序为例，把一次旋转过程拆解为以下过程：1) 物体绕参考坐标系的 Z 轴旋转，得到偏航角 yaw ；2) 绕旋转后的 Y 轴旋转，得到俯仰角 pitch ；3) 绕旋转后的 X 轴旋转，得到滚转角 roll 。这种旋转方式称为内旋，该过程可使用一个三维向量 $[r, p, y]$ 来表示三维空间中的任意旋转。那么，需要将绕每一个轴的旋转表示为旋转矩阵形式：

$$\begin{aligned} \mathbf{R}_x &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{bmatrix} \\ \mathbf{R}_y &= \begin{bmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{bmatrix} \\ \mathbf{R}_z &= \begin{bmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{aligned} \quad (2-32)$$

进而，旋转矩阵 $\mathbf{R}$ 可分解为 $\mathbf{R} = \mathbf{R}_z \mathbf{R}_y \mathbf{R}_x$ ，带入式(2-32)可得：

$$\mathbf{R} = \begin{bmatrix} \cos\gamma \cos\beta \cos\alpha - \sin\gamma \sin\alpha & \cos\gamma \cos\beta \sin\alpha + \sin\gamma \cos\alpha & -\cos\gamma \sin\beta \\ \sin\gamma \cos\beta \cos\alpha - \cos\gamma \sin\alpha & \sin\gamma \cos\beta \sin\alpha + \cos\gamma \cos\alpha & \sin\gamma \sin\beta \\ \sin\beta \cos\alpha & \sin\beta \sin\alpha & \cos\beta \end{bmatrix} \quad (2-33)$$

对式(2-33)进行逆变换，得到由旋转矩阵到欧拉角的逆变换解析如下：

$$\begin{cases} r_x = \arctan 2(\mathbf{R}(3,2), \mathbf{R}(3,3)) \\ r_y = \arctan 2(\mathbf{R}(3,1), \sqrt{\mathbf{R}(3,2)^2 + \mathbf{R}(3,3)^2}) \\ r_z = \arctan 2(\mathbf{R}(2,1), \mathbf{R}(1,1)) \end{cases} \quad (2-34)$$

工业机器人简化机械结构可视为由多个关节连接的多连杆结构，机器人连杆关节一般包括转动关节和棱柱联轴节，以常见的六连杆机械手为建模对象，其关节类型为转动关节，每一个连杆含有一个旋转自由度，拥有 6 个关节，自由度分配为 3 个位置自由度和 3 个方向自由度，这使得该类机器人能够在 3D 空间中在运动域达到任意的位置与姿态。

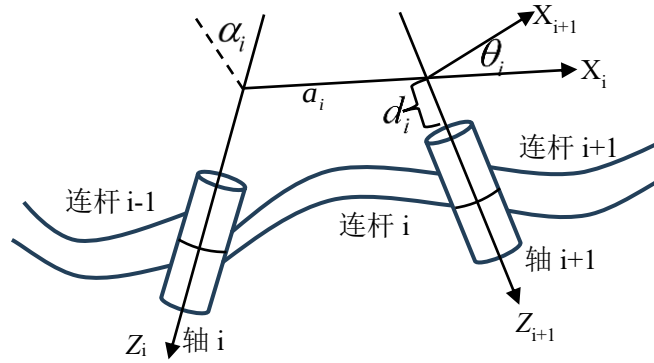


图 2-7 工业机器人 D-H 模型示意图

图 2-7 给出了机器人相邻关节的 D-H 模型。该模型通过分析两个相邻关节的运动状态，以完成对机器人整体运动的拆解分析。首先，在相邻的两个关节对应的连杆建立杆件坐标系，两个相邻的关节轴 i ，轴 $i+1$ 被连杆 i 连接，连杆长度为 a_i ，连杆偏距为 d_i 。构建关节坐标系，X 轴，Z 轴，Y 轴表示过关节坐标系原点且与 X 轴与 Z 轴正交。两关节之间的相对运动可表示为绕 Z 轴的转动，沿 Z 轴的平移；绕 X 轴的转动，沿 X 轴的平移。因此，可以基于 $SE(3)$ 矩阵来给出运动的齐次变换过程表达如下：

$${}^{n+1}_n \mathbf{T} = \mathbf{T}_{Rz} \mathbf{T}_{Lz} \mathbf{T}_{Rx} \mathbf{T}_{Lx} \quad (2-35)$$

式(2-35)中，各正交矩阵均表达了单自由度的运动，这对于分析机器人运动而言不够紧凑，因此，更倾向于使用欧拉角来解析该运动过程，式(2-35)可改写为：

$${}^{n+1}_n \mathbf{T} = \text{Rot}(x, \alpha) \text{Trans}(a, 0, 0) \text{Rot}(z, \theta) \text{Trans}(d, 0, 0) \quad (2-36)$$

整理即可得到 ${}^{n+1}_n \mathbf{T}$ 的变换通式为：

$${}^{n+1}_n \mathbf{T} = \begin{bmatrix} \cos \theta_n & -\sin \theta_n & 0 & a_n \\ \sin \theta_n \cos \alpha_n & \cos \theta_n \cos \alpha_n & -\sin \alpha_n & -d_n \sin \alpha_n \\ \sin \theta_n \sin \alpha_n & \cos \theta_n \sin \alpha_n & \cos \alpha_n & d_n \cos \alpha_n \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2-37)$$

进一步，任意两个关节之间的运动模型可根据齐次矩阵乘积，得到

$${}^sT_n = {}^sT_{s+1} {}^{s+1}T_{s+2} \cdots {}^{n-s+2}T_{n-1} {}^{n-1}T_n \quad (2-38)$$

式(2-38)表达了机器人关节坐标系之间的任意变换，该变换过程也称为机器人关节坐标系变换的链式法则。

2.4 本章小结

在使用视觉系统观测机器人运动时，涉及到机器人的运动状态与视觉系统观测的数学建模过程。本章首先提出了视觉测量方案，详细阐述了视觉系统测量的基本理论，介绍了 3D 空间到 2D 空间的映射类型，并通过透视变换关系引出相机成像模型。在此基础上，详细推导了相机参数标定的原理与过程，包括各参数解析所用的主流方法。最后，对机器人运动的基础理论和表达方式进行了推导和分析。

第3章 机器人结构运动追踪与定位

完成对机器人全局结构运动测量的关键工作在于，将视觉系统所捕捉到的机器人全局结构运动转换为几何空间的基本要素，即实现对标识特征点的跟踪与定位。为此，本章采用深度学习的多目标追踪算法，采集标识图片构建数据集进行训练，进而快速准确地追踪标记在机器人结构表面的多个标识；随后通过提出的特征点定位方法确定标识特征点的像素坐标，将机器人全局结构运动量化为几何空间的多点运动，从而实现机器人全局结构运动的初步定位。

3.1 基于深度学习的目标追踪

3.1.1 神经网络的基本结构

神经网络是一种计算模型，受生物神经元信息传递的启发，网络由大量相互连接的简单单元（称为神经元）组成，这些神经元通过学习和调整连接权重来处理输入数据并产生输出结果。神经网络通常存在多个层次，主要包括输入层、隐藏层和输出层，每一层都由多个神经元组成。每个神经元接收来自上一层神经元的输入，并通过线性变换和非线性激活函数将其转换为输出。这些输出随后传递到下一层的神经元，以此类推，直到达到输出层为止。

神经网络基于架构可区分不同类型，目前，主流的基本网络架构有前馈神经网络（Feedforward Neural Networks, FNN）、循环神经网络（Recurrent Neural Network, RNN）、生成对抗网络（Generative Adversarial Network, GAN）、卷积神经网络（Convolutional Neural Network, CNN）等。其中，CNN 对具有网格结构数据（如图像，视频）处理任务中所体现出高鲁棒性和优秀的泛化能力，广泛用于缺陷检测，多目标跟踪，视觉 SLAM 等领域。CNN 基本组成包括卷积层、池化层和全连接层，其结构如图 3-1 所示。

CNN 通过对输入数据进行卷积操作来提取特征。卷积操作利用一个滤波器（也称为卷积核）在输入数据上进行滑动，计算每个位置的内积，从而得到特征图。卷积操作的数学原理可以用二维离散卷积来描述，即将输入数据和卷积核进行逐元素相乘并求和。

假设有两个矩阵：图像的输入矩阵： I ，大小为 $M \times N$ ，通道数为 C ，其中元素表示输入图像的像素值。卷积核（滤波器）矩阵： F ，大小为 $K \times K$ ，输出特征图 O 的尺寸为 $OH \times OW$ ，当卷积操作使用零填充（padding）和步长（stride）为 S 时，具体的卷积过程可以用以下数学公式表示：

$$O_{m,n,c} = \sum_{i=0}^{k-1} \sum_{j=0}^{k-1} \sum_{k=0}^{C-1} I_{(m \times S+i), (n \times S+j), k} \cdot K_{i,j,k,c} \quad (3-1)$$

式中， $O_{m,n,c}$ 表示输出特征图中第 m 行、第 n 列、第 c 个通道的元素； $I_{(m \times S+i), (n \times S+j), k}$ 表示输入特征图中第 $(m \times S+i)$ 行、第 $(n \times S+j)$ 列、第 k 个通道的元素； $K_{i,j,k,c}$ 表示卷积核中第 i 行、第 j 列、第 k 个通道与输出特征图中第 c 个通道之间的权重。

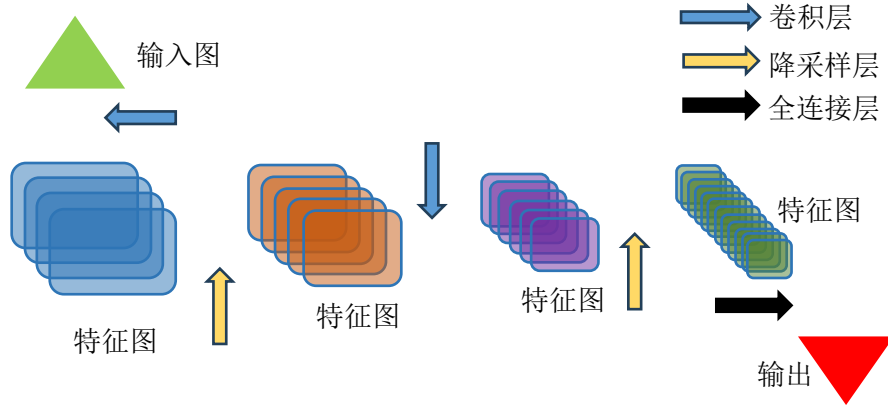


图 3-1 CNN 网络结构

池化层用于降采样，通过对输入数据进行池化操作来减小特征图的尺寸，同时保留重要的特征。常见的池化操作包括最大池化和平均池化。

最大池化是取池化窗口中的最大值作为输出，数学表达为：

$$f_{m,n,c} = \max x_{(m \times S+i), (n \times S+j), c} \quad (3-2)$$

平均池化则是取平均值作为输出，利用固定大小的窗口在特征图上进行移动，并根据窗口内的数值计算池化结果对应的数学解析为：

$$y_{m,n,c} = \frac{1}{k \times k} \sum_{i=0}^{k-1} \sum_{j=0}^{k-1} x_{(m \times S+i), (n \times S+j), c} \quad (3-3)$$

全连接层位于网络的尾部，起着整合特征、提取高级特征、模式识别和分类的重要作用，它将卷积层和池化层得到的特征图展开成一维向量，并通过权重连接每个神经元，以整合这些特征信息。最终，全连接层将学习到如何将这些特征映射到不同类别的输出。

3.1.2 信息传递与反馈

神经网络通过数据样本驱动，在给定的目标类型的基础上进行训练，网络不断度量并逼近预测信息和实际信息的距离，以提高网络对目标检测的准确度。CNN 的训练通常使用梯度下降法或其变种进行。在梯度下降法中，每次迭代需要计算损失函数关于模型参数的梯度，以更新参数，从而使损失函数逐渐减小。

梯度下降法的训练过程涉及到前向传播和反向传播两个关键步骤。在前向传

播过程中，将输入数据通过 CNN 中的各层网络进行计算，最终得到输出结果。前向传播可分为两个步骤进行：

(1) 将输入图像通过卷积层、池化层和激活函数等操作，得到一组特征映射。

(2) 将特征映射通过全连接层，将其映射到输出空间，得到最终的分类或回归结果。

假设有一个包含 L 层的卷积神经网络，其中第 l 层的输出表示为 a^l ，则前向传播的计算过程可以用如下公式表达：

$$\begin{cases} z^{l+1} = W^l a^l + b^l \\ a^{l+1} = g(z^{l+1}) \end{cases} \quad (3-4)$$

式中， z^{l+1} 表示第 $l+1$ 层的加权输入， W^l 和 b^l 分别表示第 l 层的权重和偏置项， $g(\cdot)$ 表示激活函数， a^{l+1} 表示经过激活函数后的第 $l+1$ 层的输出。

在反向传播 (Backward Propagation) 过程中，首先计算损失函数对网络参数的梯度，然后利用梯度下降法来更新参数。假设损失函数为 $Loss$ ，则反向传播的计算过程可以用如下公式表达：

$$\begin{cases} \frac{\partial Loss}{\partial W} = \frac{\partial Loss}{\partial z^{l+1}} \frac{\partial z^{l+1}}{\partial W} \\ \frac{\partial Loss}{\partial b^l} = \frac{\partial Loss}{\partial z^{l+1}} \frac{\partial z^{l+1}}{\partial b^l} \\ \frac{\partial Loss}{\partial a^l} = \frac{\partial Loss}{\partial z^{l+1}} \frac{\partial z^{l+1}}{\partial a^l} \end{cases} \quad (3-5)$$

式中， $\frac{\partial Loss}{\partial z^{l+1}}$ 表示损失函数对第 $(l+1)$ 层加权输入的偏导数， $\frac{\partial z^{l+1}}{\partial W}$ 表示表示第 $(l+1)$

层加权输入对权重的偏导数， $\frac{\partial z^{l+1}}{\partial b^l}$ 表示表示第 $(l+1)$ 层加权输入对偏置项的偏导数。

3.2 基于 YOLOv5 的机器人运动追踪

基于深度学习网络架构所成的目标检测系统具有快速，高鲁棒性的检测优势。YOLOv5 网络作为一种单阶段检测架构，其采用端到端的训练方式，即使单一的神经网络直接预测目标的类别和位置。相比双阶段网络结构，YOLOv5 具备轻量级的模型结构，可以更高效地学习目标的特征，训练和检测所占用的内存资源较少，这使得能够很好的适应小样本训练。同时，YOLOv5 对目标在图像占比变换具有较强的鲁棒性，这契合机器人在运动过程中的整体图像大小的远近变化，考虑上述因素，

本文选用 YOLOv5 算法进行机器人结构运动追踪。

3.2.1 YOLOv5 网络结构

YOLOv5 的网络结构组成可分输入层、主干网络（Backbone）特征融合模块（Neck）和检测头（Head）四个部分，如图 3-2 所示。

输入层接收原始图像，并通过 RandAugment 和 CutMix 两种增强技术，对原始图像进行裁剪、旋转、拼接、缩放等处理，旨在均衡化不同尺度的目标在总样本的占比，同时减少内存占用，同时，对图片的尺度大小规范化，并将处理后的图片传递给主干网络。

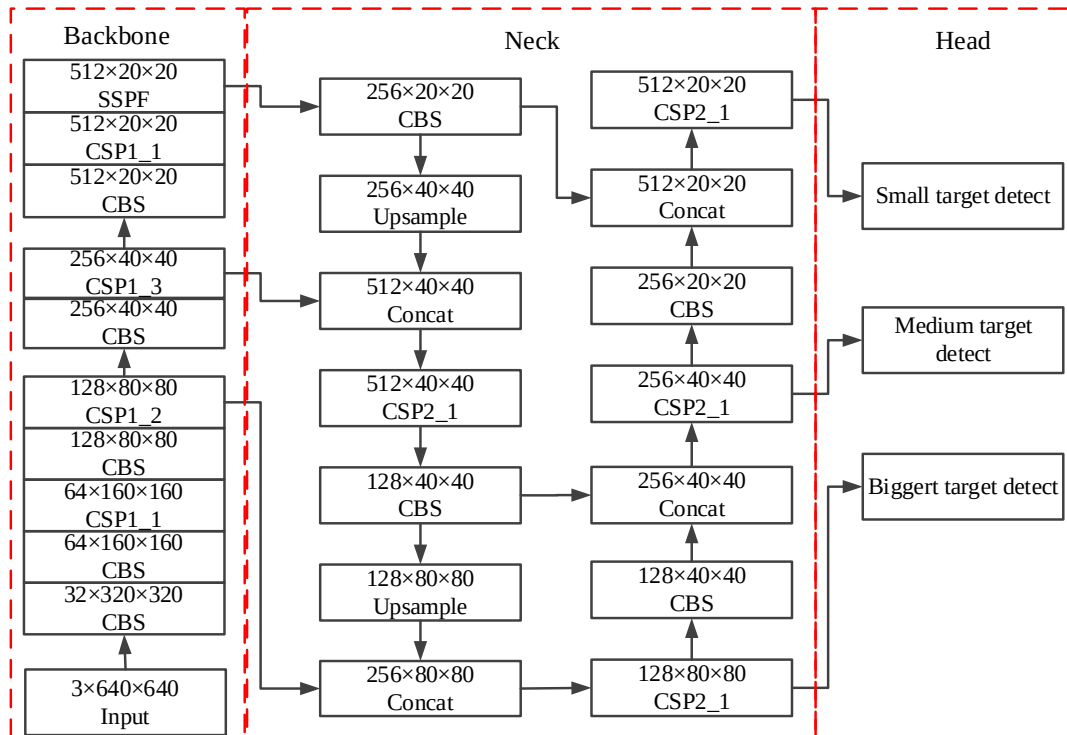


图 3-2 YOLOv5 网络结构简图

主干网络从输入图像中提取特征，通过多个卷积层和池化层逐渐降低特征图的尺寸，并在较高层次的特征图上进行目标检测和分类，初步形成 3 类大小不同的特征信息并传递给特征融合模块。

特征融合模块主要包括两个部分：跨尺度连接和特征上采样。首先，跨尺度连接通过将低分辨率的特征图与高分辨率的特征图进行连接，实现了不同尺度特征的融合。这样做的好处是在较高的分辨率上保留细节信息，并且利用低分辨率特征图的语义信息来增强检测性能。其次，特征上采样通过上采样操作将低分辨率的特征图恢复到与高分辨率特征图相同的尺寸。这样可以使得融合后的特征图具有相同的分辨率，从而更好地适应目标的多尺度检测需求。

训练完成后，检测头已赋有对特定目标的特征信息，基于所提取的特征信息，

对图像的目标进行检测，其具体作用可以分为以下几个方面：

目标位置预测：检测头通过卷积操作对特征图进行处理，产生目标的边界框位置预测。通常，每个边界框由 4 个坐标值（x、y、宽、高）表示。

目标类别预测：检测头利用卷积和全连接层对特征图进行分类操作，预测图像中目标的类别。一般情况下，目标类别由一个或多个预定义的类别标签表示。

目标置信度预测：检测头还会预测每个边界框的置信度，用于判断该边界框是否包含目标物体。置信度反映了模型对边界框是否准确的估计程度。

通过检测头的预测结果，可以得到检测到的目标的位置信息、类别信息以及相应的置信度。根据置信度的高低，可以进行目标筛选和后处理，最终得到准确的目标检测结果。

3.2.2 损失函数与目标判别策略

YOLOv5 网络结构将待检测目标分为 Biggert target、Medium target、Small target 三类目标，对应的像素网格大小分别为 80*80、40*40、20*20。检测头对这三类候选网格，通过（损失）函数对网格内的目标进行度量，以确定目标在图像中的位置。

损失函数是所有神经网络最主要的量化评价的数学模型，YOLOv5 的损失函数组成包括定位损失，置信度损失和分类损失。

YOLOv5 使用 CIOU 作为定位损失函数，该损失函数考虑了预测框 A 和目标位置 B 的重叠度，中心距离，宽高比和对角尺度，函数模型如下：

$$L_t = 1 - IoU + \frac{p^2}{c^2} + \alpha v \quad (3-6)$$

式中， IoU 表示 A 和 B 的交互比， p 表示中心点距离， c 表示 A 和 B 最大对角线距离。宽高一致比权重系数 α 的数学解析如下：

$$\alpha = \frac{v}{1 - IoU + v} \quad (3-7)$$

式中， v 为宽高一致比函数，表示如下：

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w^{true}}{h^{true}} - \arctan \frac{w}{h} \right)^2 \quad (3-8)$$

由式(3-8)可知，当预测框与目标区域宽高趋于一致时， v 的取值越小，进而，权重比值系数的取值越小，此时损失函数 L_t 着重优化目标距离，反之， v 的取值趋向于 1，此时损失函数 L_t 着重优化宽高比，该策略能够有效提升训练的稳定性和收敛速度。

YOLOv5 置信度损失和分类损失均采用交叉熵函数模型：

$$C = \frac{1}{n} \sum_x [y \ln a + (1-y) \ln(1-a)] \quad (3-9)$$

式中， a 表示预测输出， x 表示样本， y 表示样本标签， n 表示样本总量。其中，置信度损失函数使用二元交叉熵损失函数（Binary Cross Entropy Loss），用于衡量检测框中物体与背景的二元分类预测。假设 p_i 是预测的置信度值， y_i 是对应的真实标签（0 表示背景，1 表示物体）， N 是检测框数量，则二元交叉熵损失函数可以表示为：

$$L_j = -\frac{1}{N} \sum_{i=1}^N [y_i \ln p_i + (1-y_i) \ln(1-p_i)] \quad (3-10)$$

而分类损失函数则使用多类交叉熵损失函数（Categorical Cross Entropy Loss），用于衡量检测框中物体属于每个类别的概率预测。假设 t_{ij} 是预测的类别概率值， y_{ij} 是对应的真实标签（0 表示背景，1 表示物体），则多类交叉熵损失函数可以表示为：

$$L_c = -\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^C y_{ij} \ln t_{ij} \quad (3-11)$$

YOLOv5 的主态损失函数可表示为：

$$Loss = \lambda_t L_t + \lambda_j L_j + \lambda_c L_c \quad (3-12)$$

式中， λ 为模型的超参数表示损失权重，具体的取值依据现有的其中，置信度权重 λ_j 占比最大，矩形框权重 λ_t 居中，分类权重 λ_c 最小。

在目标识别过程中，检测头首先对每个预测框基于评价函数进行预测，判断预测置信度是否超过设定阈值，若超过则认为该预测框内存在目标。进而，根据非极大值抑制算法对存在目标的多个预测框进行筛选，提取出评分最高的预测框。最后，根据筛选后预测框的分类概率，取最大概率对应的类别索引作为目标的类别。此外，YOLO 网络在每一轮训练过程中对各预测框进行评价，通过计算和反向传播，不断更新损失函数，以迭代优化网络模型的参数，从而提高目标检测的精度。

3.2.3 算法部署

YOLO 整体的模型训练与检测流程如下所述：

（1）数据集准备：将标识粘贴在机器人表面，驱动机器人运动同时采集带有标签的机器人运动图像，并使用图像增强技术丰富原有数据组成数据集。

（2）模型选择：并确定模型的超参数使用（包括网络结构、滑动窗口大小、Anchor Box 数量）和训练方案（如学习率、优化器等）。

（3）损失函数定义：定义用于监督训练的多个损失函数，包括分类损失、位

置损失和置信度损失等，以最小化模型在训练集上的误差。

(4) 模型初始化：对 YOLO 模型的参数进行随机初始化，以便开始训练过程。

(5) 前向传播：将训练集中的图像输入到 YOLO 模型中，通过前向传播计算输出的检测框的位置、类别和置信度等信息。

(6) 损失函数计算：根据监督标签和 YOLO 模型输出的检测结果，计算各个损失函数的值，通过加权相加得到训练过程中使用的总损失函数。

(7) 反向传播：根据总损失函数对模型参数进行反向传播，得到每个参数的梯度，以用于后续的优化过程。

(8) 参数更新：根据优化算法（如梯度下降）和学习率，利用反向传播得到的梯度来更新 YOLO 模型的参数。

(9) 超参数调优：使用验证集对训练得到的模型进行评估，在训练过程中进行调整，以防止过拟合和欠拟合等问题。

(10) 测试模型：最后将训练得到的 YOLO 模型应用到测试集或实际场景中进行物体检测，反复调整模型参数，使之在检测准确率和速度上达到理想的平衡。

3.3 特征点精确定位

3.3.1 圆形十字标识特征点定位

完成 YOLOv5 模型的训练和部署后，检测器能够得到机器人结构表面粘贴的标识子图像块在整体图像上的区域，实现标识的初定位。然而，实现机器人运动视觉精确测量关键在于，能够准确提取每个图像中目标的像素坐标位置。因此，需要在初定位的基础上完成特征点定位以精确测量机器人的结构运动。

首先分析本文 2.1 节所设计图 2-1(a)所示的圆形内十字标识，特征点判定标识中心，几何约束表现为十字线的交点，如图 3-3 所示。

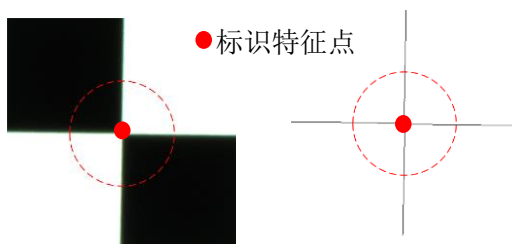


图 3-3 圆形十字标识特征点定位示意图

为确保特征点提取的准确性与精度，本文给出了一种高鲁棒性的特征点提取方法。该方法融合了 Zernike 矩亚像素点提取与最小二乘法直线拟合。使用高斯滤波对标识图像块进行平滑处理，以消除图像中低灰度部分的颜色突变。使用全阈值法灰度二值化将灰度图像转换为二值图像，通过 Canny 算子来检测像素级的十字线边缘。

在该边缘的基础上,使用 Zernike 矩亚像素检测进行进一步精确处理。处理过程如下:利用 Zernike 矩的模板 M11、M20 求解亚像素边缘,该模板的数学方程如式(3-13):

$$Z_{nm}(u,v) = \sum_{i=-t}^t \sum_{j=-t}^t I(u+j, v+j) M_{nm}(t+j, t+i) \quad (3-13)$$

式中,模板 M_{nm} 是当图像的 $f(u,v)=1$ 时的 n 阶 m 次幂矩。其大小为 $N \times N$ 。 t 的值为 $N/2$ 。 $I(u,v)$ 表示像素点 (u,v) 处的像素灰度值。模板 M_{nm} 使用 $N \times N$ 网格来划分单位圆。 $N \times N$ 系数可以使用以下公式(3-14)来计算。 S_{ji} 是指第 j 行和第 i 列中的小正方形。 N 取值可为 5、7、11 等。

$$M_{nm}(j,i) = \iint_{S_{ji} \cap C} V_{nm} dudv \quad (3-14)$$

考虑到边缘是渐变而不是阶跃变化,在这个过程中,设定边缘宽度为 $2w$, w 的值取为高斯函数标准偏差的两倍, $w=1.66\sigma$ 。下一步是计算子像素边缘点 canny 算子检测到的像素边缘点之间的距离。获取距离计算公式如式(3-15):

$$l = \left[1 - w^2 - \sqrt{(w^2 - 1)^2 - 2w^2 Z_{20}' / Z_{11}'} \right] / w^2 \quad (3-15)$$

其中, Z_{11}' 可通过如下方程求得:

$$\begin{cases} \phi = \arctan(\text{Im}(Z_{11}) / \text{Re}(Z_{11})) \\ Z_{11}' = \text{Im}(Z_{11}) \sin \phi + \text{Re}(Z_{11}) \cos \phi \end{cases} \quad (3-16)$$

$\text{Im}(Z_{11})$ 和 $\text{Re}(Z_{11})$ 分别表示 Z_{11} 的实部与虚部,由此,所得的亚像素边缘点可表示为:

$$\begin{bmatrix} u_{sp} \\ v_{sp} \end{bmatrix} = \begin{bmatrix} u \\ v \end{bmatrix} + \frac{Nl}{2} \begin{bmatrix} \cos \phi \\ \sin \phi \end{bmatrix} \quad (3-17)$$

在获得边缘点的基础上,根据边缘点对边缘点进行分类,具体而言,根据边缘点和中心点的斜率来区分各点所属的边缘线段(横线或者竖线),表示为:

$$k_l = \left| \frac{u_{spi} - u_{center}}{v_{spi} - v_{center}} \right| \Rightarrow \begin{cases} \delta_1 \leq k_l \leq \delta_2, (u_{spi}, v_{spi}) \subseteq \text{line1} \\ \text{else}, (u_{spi}, v_{spi}) \subseteq \text{line2} \end{cases} \quad (3-18)$$

式中, (u_{center}, v_{center}) 表示亚像素边缘点的质心, δ_1 , δ_2 为横线斜率的判断阈值。继而,通过最小二乘直线拟和得到横线和竖线的具体方程,最后,基于所拟合的直线求交定位特征点坐标。整体的特征点像素坐标定位流程如图 3-4 所示。

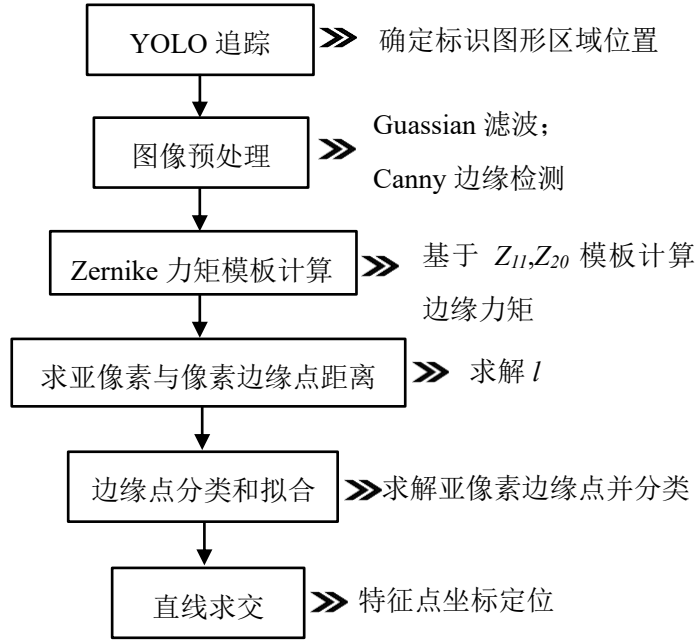


图 3-4 特征点像素坐标点位处理流程

3.3.2 三角形标识特征点定位

为测量机器人位姿，设计了直角三角形标识以建立机器人的运动坐标系，详细的建系过程将在文章的第四章阐述，本节仅讨论该标识的定位过程。如图 3-5 所示，所获取的特征点定义为三角形标识的三个顶点。

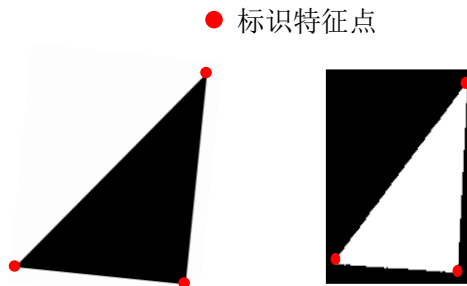


图 3-5 直角三角形标识特征点定位示意图

图像预处理的过程同 3.3.1 节保持一致：首先使用高斯滤波对三角形标识进行平滑处理，然后，使用全阈值法灰度二值化将灰度图像转换为二值图像，其次，通过 Canny 算子来得到三角形标识的三条边缘；再次，使用 Zernike 矩亚像素检测得到对应的亚像素边缘点。为准确对亚像素边缘点进行边分类，在 Canny 边缘检测的基础上，使用了霍夫直线变换以获取参考直线。

假设三角形三条参考直线分别为直角短边 L_1 ，直角长边 L_2 ，斜边 L_3 ，则直线的一般方程定义可 $A_i x + B_i y + C_i = 0$ ($i=1,2,3$)。因此，各亚像素边缘点到三条参考直线的距离可表示为：

$$d_i = |A_i x_0 + B_i y_0 + C_i| / \sqrt{A_i^2 + B_i^2} \quad (3-19)$$

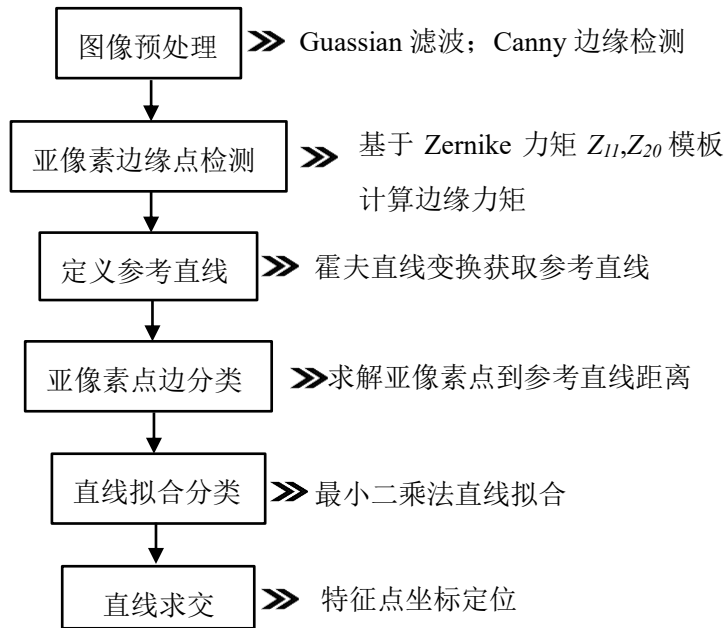


图 3-6 三角形标识特征点定位流程

进而，各亚像素点的边所属可根据点到三条直线的距离进行区分。在此基础上，使用最小二乘法分别对三角形三条边所属的亚像素点进行直线拟合，得到标准直线 L_{S1} , L_{S2} , L_{S3} 。最后，求三条标准直线的交点，其中，直角短边 L_{S1} 与斜边 L_{S3} 的交点定义为 p_1 ，直角短边 L_{S1} 与直角长边 L_{S2} 的交点定义为 p_2 ，直角长边 L_{S2} 与斜边 L_{S3} 的交点定义为 p_3 。图 3-6 展示了三角形标识特征点的定位流程。

3.4 本章小结

本章旨在追踪机器人的空间运动，通过追踪与定位标识最终得到标识特征点的像素坐标。为实现该目的，采用了 YOLOv5 算法完成了合作标识的运动追踪，初步定位合作标识的二维图像区域。简要阐述了 CNN 神经网络的基本原理，YOLOv5 的网络结构，训练损失函数模型以及 YOLOv5 的部署流程。在完成合作标识区域初定位后，为准确提取靶标特征点的像素坐标，给出了一种鲁棒性特征点定位方法。该方法在 Canny 边缘检测的基础上，首先通过 Zernike 矩亚像素处理拟合得到亚像素边缘点，并基于参考点或参考直线对亚像素边缘点分类；然后使用最小二乘法拟合亚像素边缘点对应的直线方程；最后联立直线方法求交得到特征点具体像素坐标。

第4章 机器人全局结构运动视觉测量

基于第3章所述内容,视觉系统可以将机器人结构运动表征为2D空间的点运动,得到合作标识在机器人运动过程中的像素坐标。本章将具体阐述机器人全局结构运动视觉测量原理与数学模型,给出由特征点像素坐标重构至机器人空间下的平面坐标,三维坐标的推导过程,从而完成对机器人结构面上任意点的平面运动、空间运动和姿态的运动解析。

4.1 基于视觉的机器人平面运动测量

4.1.1 测量方法分析

机器人平面运动视觉测量过程可以视为,由观测点的像素坐标重构出物理平面坐标系的对应位置。这种测量模型可以使用2D平面之间的正交投影进行表达:

$$\lambda_1 \begin{bmatrix} u \\ v \end{bmatrix} = \lambda_2 \begin{bmatrix} x \\ y \end{bmatrix} \quad (4-1)$$

式中, λ_1 、 λ_2 表示像素坐标系与物理坐标系正交投影的等比因子,该式可进一步简化为 $p = sP$ 形式,其中 s 表示像平面 p 到物理平面 P 的比例系数。

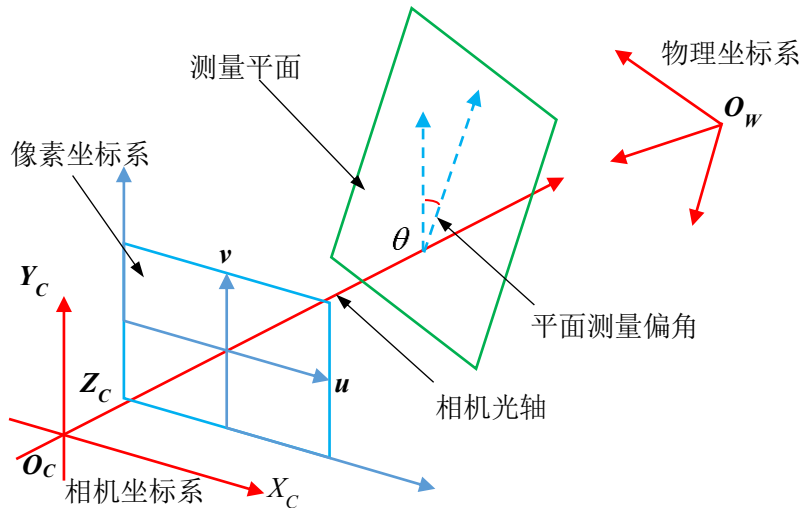


图 4-1 平面测量示意图

基于式(4-1)的平面重构模型的测量精度受限于相机光轴与被测物理平面的正交程度。如图4-1所示,当相机的光轴与物体运动平面存在倾斜偏差时,基于正交投影模型的测量精度难以得到保障。因此,该模型主要适用于对测量误差有较大容忍度的观测场所,比如大型土木结构位移等。然而,对于具备较高精度测量需求的机器人结构运动来说,该模型并不适用。

4.1.2 基于单应特征矩阵模型的平面运动测量

在使用视觉系统进行机器人平面结构运动测量时，机器人所处的平面难以确保与相机光轴垂直。为了解决这一问题，本章重新构建了一种基于单应特征矩阵变换的机器人平面运动视觉测量模型。在 2.2.1 节所提到的仿射变换模型具有更大的自由度，适用于物体与观测平面不正交的情况。因此，可以将基于式(2-3)的仿射变换投影模型推广为单应矩阵形式：

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = s \begin{bmatrix} h_{11} & h_{12} & h_{13} & h_{14} \\ h_{21} & h_{22} & h_{23} & h_{24} \\ h_{31} & h_{32} & h_{33} & h_{34} \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (4-2)$$

式中， s 表示透视变换的比例因子， h_{11} 至 h_{34} 表示单应矩阵系数。对式(4-2)施加平面约束，可以假设所有测量值均处于 O - XY 平面上，并且所有测量点 Z 坐标为 0，则可得到平面测量模型的单应态表达：

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = s \begin{bmatrix} h_{11} & h_{12} & h_{14} \\ h_{21} & h_{22} & h_{24} \\ h_{31} & h_{32} & h_{34} \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ 1 \end{bmatrix} \Rightarrow p = HP \quad (4-3)$$

式中，方程两侧均为向量形式并且相等，利用向量叉乘的基本原理，式(4-3)可以改写为 $p \times HP = 0$ ，将 H 矩阵解耦为 $H = [H_1; H_2; H_3]$ ，则式(4-3)可转换为向量叉乘形式：

$$p \times HP = \begin{bmatrix} vH_3X - H_2X \\ H_1X - uH_3X \\ uH_2X - vH_1X \end{bmatrix} = \begin{bmatrix} 0 & -X^T & vX^T \\ -X^T & 0 & -uX^T \\ -vX^T & -uX^T & 0 \end{bmatrix} \begin{bmatrix} H_1^T \\ H_2^T \\ H_3^T \end{bmatrix} = 0 \quad (4-4)$$

该等式运算涉及 3×9 的矩阵与 9 元素向量，由于像素坐标到空间坐标为 2D 空间的映射，因此，上式中的三个线性方程可以消除一个非独立约束，式(4-4)进一步简化为如下形式：

$$\begin{bmatrix} 0 & -X^T & vX^T \\ -X^T & 0 & -uX^T \end{bmatrix} \begin{bmatrix} H_1^T \\ H_2^T \\ H_3^T \end{bmatrix} = 0 \Rightarrow BH_h = 0 \quad (4-5)$$

式中， $B \in C^{2 \times 9}$ ，从方程(4-5)可以看出，一个点对可以给出两个方程，因此，至少需要 4 个点对联立方程可以得到 H_h 。由于不可避免的测量噪声，仅使用 4 点对所得的计算结果可能具有较大偏差。当机器人的运动被约束在单一平面时，就可以使用视觉系统收集获得组数大于 4 的 N 组已知点对，从而在参数求解阶段，构造超定方程组 $L_B H_h = 0$ ，其中， $L_B = [B_1; B_2; \dots; B_N]$ 。此时，求解 H_h 参数可视为齐次线

性方程组寻优问题，为此，可以使用线形最小二乘模型对方程组进行估计，则超定方程组可等价如下形式：

$$\mathbf{L}_B \mathbf{H}_h = 0 \Rightarrow \min_x \|\mathbf{L}_B \mathbf{H}_h\|^2 \quad \mathbf{L}_B \in C^{2N \times 9}, \mathbf{H}_h \in C^{9 \times 1} \quad (4-6)$$

对 $\mathbf{L}_B$ 进行奇异值分解为 $\mathbf{L}_B = \mathbf{U} \mathbf{D} \mathbf{V}^T$ 带入(4-6)得：

$$f(\mathbf{L}_B) = \min_x \|\mathbf{U} \mathbf{D} \mathbf{V}^T \mathbf{H}_h\|^2 \quad (4-7)$$

式中， $\mathbf{U} \in C^{2N}$ 为 $\mathbf{L}_B$ 的左正交矩阵，由 $\mathbf{L}_B$ 的 r 个正交向量 $(u_1, u_2, \dots, u_r)$ 和 $2N-r$ 个

$$\text{扩充正交向量 } (b_{2N-r}, b_{2N-r+1}, \dots, b_{2N}) \text{ 构成。} \mathbf{D} = \begin{bmatrix} \sqrt{\varphi_1} & & & 0 \\ & \ddots & & \\ & & \sqrt{\varphi_r} & \\ & 0 & & 0 \end{bmatrix}_{2N \times 9}, \mathbf{D} \text{ 中的 } \varphi \text{ 表}$$

示 $\mathbf{L}_B$ 特征值，各特征点大于零，且 $\varphi_1 > \varphi_2 > \dots > \varphi_r$ 。 $\mathbf{V} \in C^{2N}$ 为 $\mathbf{L}_B$ 的右正交矩阵，

由 $\mathbf{L}_B$ 的 9 个正交特征向量 $(v_1, v_2, \dots, v_9)$ 组成。令 $y = \mathbf{V}^T \mathbf{L}_B$ ，则(4-7)可整理为如下形式：

$$f(y) = \min_y \|\mathbf{U} \mathbf{D} y\|^2 = U^2 \min_y (\varphi_1 y_1^2 + \varphi_2 y_2^2 + \dots + \varphi_r y_r^2) \quad (4-8)$$

将 $\varphi_1 > \varphi_2 > \dots > \varphi_r$ 的约束带入(4-8)可以发现，当 $y = [0, 0, \dots, 1]^T$ 时， $\mathbf{H}_h$ 的最优非零解奇异值分解所得右正交矩阵 $\mathbf{V}$ 最后列向量。通过式(4-2)给出的单应特征矩阵平面测量模型，以及式(4-3)到式(4-8)的解析过程，能够实现对机器人结构上任一点的平面运动观测，该模型不受相机光轴与测量平面垂直度的限制，可以预见的是，平面运动视觉测量系统的精度与鲁棒性将得到提高。

4.2 基于视觉的机器人空间运动测量

4.2.1 3D 重建模型

机器人的运动精度、承载能力、振动、稳定性等性能指标都可以在机器人结构的运动所体现。而该类运动常以 3D 空间的线轨迹，圆弧轨迹，以及直线与圆弧的堆叠与积分运动形式出现，无法满足平面约束。因此，需要一个有效的数学模型来准确表达机器人任意时态下的三维位移信息。

在前文，透视变换中未知的尺度因子导致单目成像无法将 2D 信息恢复为 3D，因此，需要对视觉系统添加额外的约束。一个最为有效的方法是使用 2 个相机对同一目标进行观测，如图 4-2 所示，此时，基元 P_w 的空间位置为左右两相机逆透视射线的交点。基于该原理，双目视觉能够准确恢复目标的空间位置，从而实现机器

人 3D 运动的测量。

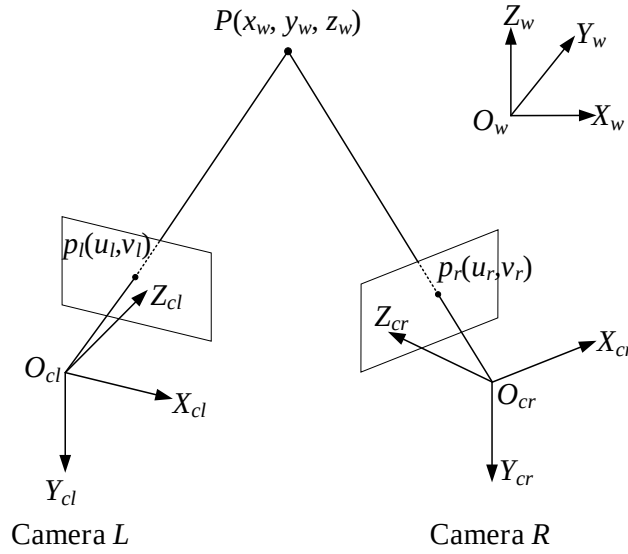


图 4-2 双目空间坐标测量模型示意图

结合相机成像与双目标定模型，可得到双目视觉系统的基本观测模型为：

$$\begin{cases} \lambda_l p_L = \mathbf{K}_L P_L \\ \lambda_r p_R = \mathbf{K}_R P_R \\ P_L = \mathbf{R}_l^r P_R + \mathbf{t}_l^r \end{cases} \quad (4-9)$$

式(4-9)中， $\mathbf{K}_L$ 和 $\mathbf{K}_R$ 是左相机和右相机的内参矩阵， λ_l 和 λ_r 表示左右相机测量模型的尺度因子。 p_L 和 p_R 表示像素坐标 p_l 和 p_r 的矢量形式， P_L 和 P_R 分别为重建在相机坐标系坐标 P_l 和 P_r 的齐次坐标的矢量形式。

以下将推导双目测量模型的解析过程。将式(4-9)的第三方程 $P_L = \mathbf{R}_l^r P_R + \mathbf{t}_l^r$ 带入第一方程 $\lambda_l p_L = \mathbf{K}_L P_L$ 中，可以得到如下形式：

$$\lambda_l p_L = \mathbf{K}_L (\mathbf{R}_l^r P_R + \mathbf{t}_l^r) \quad (4-10)$$

将式(4-10)等号两边同时叉乘 p_L 使得等号左边归零，以消除 λ_l ，此时有：

$$p_L \times \mathbf{K}_L \mathbf{R}_l^r P_R + p_L \times \mathbf{K}_L \mathbf{t}_l^r = 0 \quad (4-11)$$

对式(4-11)提取 P_R ，整理为齐次线性方程的矩阵乘积形式如下：

$$\begin{bmatrix} pKR_1 \cdot pKt_3 - pKR_3 \cdot pKt_1 \\ pKR_2 \cdot pKt_3 - pKR_3 \cdot pKt_2 \end{bmatrix} [P_R] = 0 \quad (4-12)$$

式中， pKR_i 表示 $p_L \times \mathbf{K}_L \mathbf{R}_l^r$ 叉乘的第 i 个行向量， pKt_i 表示 $p_L \times \mathbf{K}_L \mathbf{t}_l^r$ 结果的列向量第 i 个元素。同样的，将式(4-9)的第二方程 $\lambda_r p_R = \mathbf{K}_R P_R$ 等号两边同时叉乘 p_R ，可改写为：

$$\begin{bmatrix} K_{R11} & K_{R12} & K_{R13} - u_R \\ K_{R21} & K_{R22} & K_{R23} - v_R \end{bmatrix} \begin{bmatrix} P_R \end{bmatrix} = 0 \quad (4-13)$$

联立式(4-12)，式(4-13)可得：

$$\begin{bmatrix} pKR_1 \cdot pKt_3 - pKR_3 \cdot pKt_1 \\ pKR_2 \cdot pKt_3 - pKR_3 \cdot pKt_2 \\ K_{R11} & K_{R12} & K_{R13} - u_R \\ K_{R21} & K_{R22} & K_{R23} - v_R \end{bmatrix} \begin{bmatrix} P_R \end{bmatrix} = 0 \Rightarrow AP_R = 0 \quad (4-14)$$

在确定双目相机的成像模型参数和被观测 3D 基元在相机图像坐标系的映射后，即可使用式(4-14)重构出 3D 基元于相机坐标系下的 3D 坐标值 P_R 。该方程的求解过程可以使用 4.1 节所阐述的 SVD 分解法，得到 P_R 的重构值。

4.2.2 坐标转换方法分析

在平面测量模型中，通过单应性特征矩阵能够直接得到像素坐标与机器人坐标的转换关系。然而，在双目测量系统的基本观测模型下，所得的坐标表达是建立于相机坐标系下，无法直接得到机器人坐标系下的重构值，因此，需要得到相机坐标系到机器人坐标系的变换参数，统一测量数据与机器人运动数据。在先前的工作中，将相机坐标系到机器人坐标系的坐标变换过程视为 3D 坐标系的刚性变换，其核心是将坐标变换问题转化为点对匹配问题。具体而言，在机器人运动空间内设定若干个已知坐标的参考点，同时，参考点的位置通过视觉系统可重构于相机坐标系中，从而获得多组已知相对位置的点对，进而，点对匹配模型表达于 3D 坐标系中的转换关系可由式(4-15)所示：

$$p_i^r = R_c^r p_i^c + t_c^r \quad (4-15)$$

式中， p_i^r 表示机器人坐标系的坐标， p_i^c 对应相机坐标系的坐标， R_c^r ， t_c^r 为相机坐标系到机器人坐标系的变换旋转矩阵与平移向量。一般的，对于坐标系变换矩阵的求解采用两步法，首先得到 R_c^r 的最小二乘模型，将(4-18)改写为如下最小二乘形式：

$$f(R_c^r, t_c^r) = \arg \min \sum_{i=1}^n \|p_i^r - R_c^r p_i^c - t_c^r\|^2 \quad (4-16)$$

进一步，对最小二乘函数 $f(R_c^r, t_c^r)$ 中 t_c^r 变量求偏导，得到：

$$\frac{\partial f}{\partial t_c^r} = -2 \sum_{i=1}^n (p_i^r - R_c^r p_i^c - t_c^r) \quad (4-17)$$

观察式(4-17)可以发现，最小二乘函数关于 t_c^r 的偏导解析存在唯一零点，因此，函数存在唯一最小解，其解可表示为：

$$t_c^r = \frac{1}{n} \sum_{i=1}^n (p_i^r - R_c^r p_i^c) = \overline{p^r} - R_c^r \overline{p^c} \quad (4-18)$$

式中, $\overline{p}^r, \overline{p}^c$ 表示机器人坐标系与相机坐标系各点的质心, 此时(4-16)的极值求解式为:

$$f(\mathbf{R}_c^r, \mathbf{t}_c^r)_{\min} = \sum_{i=1}^n \left\| p_i^r - \overline{p}^r - \mathbf{R}_c^r (p_i^c - \overline{p}^c) \right\|^2 \quad (4-19)$$

令 $p_i^r - \overline{p}^r = x_{pi}^r$, $p_i^c - \overline{p}^c = x_{pi}^c$, 式(4-19)展开得:

$$f(\mathbf{R}_c^r, \mathbf{t}_c^r)_{\min} = \sum_{i=1}^n \left((x_{pi}^r)^2 + (x_{pi}^c)^2 - 2x_{pi}^r \mathbf{R}_c^r x_{pi}^c \right) \quad (4-20)$$

式中, 函数的唯一变量项为 $2 \sum_{i=1}^n x_{pi}^r \mathbf{R}_c^r x_{pi}^c$, 因此, 函数可以首先得到 $\mathbf{R}_c^r$ 的最小二乘解, 表示为 $\sum_{i=1}^n x_{pi}^r \mathbf{R}_c^r x_{pi}^c = 0$, 进而, 令 $\mathbf{X}_p^c = \{p_1^c - \overline{p}^c, p_2^c - \overline{p}^c, \dots, p_n^c - \overline{p}^c\}$, $\mathbf{X}_p^r = \{p_1^r - \overline{p}^r, p_2^r - \overline{p}^r, \dots, p_n^r - \overline{p}^r\}$, 使用奇异值分解法, 对方程组做最小秩分解, 所得结果如下:

$$\text{SVD}(\mathbf{X}_p^r \mathbf{R}_c^r \mathbf{X}_p^c) = \mathbf{U} \Sigma \mathbf{V}^T \Rightarrow \mathbf{R}_c^r = \mathbf{U} \mathbf{V}^T \quad (4-21)$$

将 $\mathbf{R}_c^r$ 带入式(4-18), 即可求得 $\mathbf{t}_c^r$ 。可以发现, 将坐标系变换模型求解问题转化为确定点对的匹配问题, 能够基于最小二乘法模型估计与奇异值分解法, 完成对模型参数的求解。该方法的优势在于, 最小二乘法模型估计通过优化误差平方和, 使得所得到的变换矩阵与真实的坐标系变换最为接近。这意味着在一定误差范围内, 最小二乘法能够找到使得测量数据与实际数据拟合最好的坐标系变换模型。

4.2.3 基于 RANSAC 误差加权的坐标变换模型优化

上节提到, 基于最小二乘法能在一定误差范围内求得一个最优的坐标系变换模型。然而, 在实际测量过程中, 本节推导了一个抗噪声, 高鲁棒性的双目视觉测量模型, 该模型的推导是基于对本领域以往工作的改进与延伸。使用点对的匹配来求解坐标系变换模型参数是一种常用的处理方式, 但该方法的求解精度受点对匹配的噪声约束。具体而言, 如果点集 $\mathbf{P}^r = \{p_1^r, p_2^r, \dots, p_i^r\}$ 和 $\mathbf{P}^c = \{p_1^c, p_2^c, \dots, p_i^c\}$ 的坐标值均为对应点在坐标系中的实际值, 不存在数据噪声, 通过 SVD 分解可求出理想的姿态旋转矩阵 $\mathbf{R}_c^r$ 和位移向量 $\mathbf{t}_c^r$, 实现测量点在两坐标系中的准确变换。但在实际的测量过程中, 由于受机器人运动精度、光照、特征点提取精度、相机畸变等因素的影响, 参与基准变换模型求解的数据会存在不同程度的噪声, 而最小二乘模型估计对于整体数据的权重不做区分, 这导致部分误差较大的离群点对模型求解精度影响较大, SVD 分解所得的姿态旋转矩阵 $\mathbf{R}_c^r$ 和位移向量 $\mathbf{t}_c^r$ 的解与真实值存在偏

差，进而造成基准变换模型失真，最终导致使运动测量结果不准确。

为此，本节在最小二乘模型估计的基础上嵌入了一个优化模块，该模块融合了随机抽样一致 (RANSAC)^[53]与迭代加权奇异值分解法 (IR-SVD)，首先，RANSAC 被运用于初始数据的过滤，目的是在初始测量点集中筛选出误差较小的优势子样本，过滤掉粗大误差；随后，提出了一个新的误差抑制算法：IR-SVD，该算法能够根据每个测量点的误差大小赋予最佳的计算权重，以提高最小二乘法模型估计的抗噪能力，并给出一个子样本对应的坐标变换模型。最终，通过 RANSAC 重复采样，能够得到每一个子样本对应的模型误差，并通过误差大小选出模型的最优解。

首先，RANSAC 算法从计算样本 $\mathbf{P}^r$ 与 $\mathbf{P}^c$ 中任意抽取 4 组测量点对组成最小计算子集。通过式(4-15)到(4-21)可以求解出一个当前样本下的初始坐标变换模型，进而，通过如下的误差函数计算总样本中各点于当前初始模型变换后的误差大小：

$$e_i = \|\mathbf{p}_i^r - \mathbf{PR}_i\| = \|\mathbf{p}_i^r - \mathbf{R}_c^r \mathbf{p}_i^c - \mathbf{t}_c^r\| \quad (4-22)$$

式中， $\mathbf{PR}_i = \{\mathbf{PR}_1, \mathbf{PR}_2, \dots, \mathbf{PR}_n\}$ 表示观测点集 $\mathbf{P}^c = \{\mathbf{p}_1^c, \mathbf{p}_2^c, \dots, \mathbf{p}_i^c\}$ 变换到机器人坐标系后的匹配点集。将最小样本模型对应的整体误差均值作为内外点判别值，测量点误差小于阈值的认为是内点，否则识别为外点。同时，给出了一个期望内点个数的统计公式来评估初始模型的优越性：

$$k \geq \frac{mn}{1+m} \quad (4-23)$$

式中， n 表示总样本的点对数量， m 表示采样数量，当内点个数大于 k 值时，视为当前的初始模型是趋优的，反之，认为当前采样值中存在粗大误差的值，将当前的模型舍去，重新抽样。接下来，IR-SVD 算法对数据做更深的数据误差评估和权重赋值。

假设初始模型通过式(4-23)能够评估出大于 k 的 N 个内点对，则这 N 个内点将被视为一个优势样本，通过 IR-SVD 算法来求解该优势样本对应的最优二乘模型。在使用奇异值分解法求解模型时，需要考虑样本误差影响，为此，本节给出了一个权重赋值函数，采用误差越大，权重越小的反比原则，突出误差较小的数据值在模型求解阶段的贡献，同时抑制误差较大的数据值参与求解过程的权重，赋值函数形如下式：

$$s_i = \left| \frac{1.5\bar{e} - e_i}{1.5\bar{e}} \right| / E \quad i = 1 \sim k \quad (4-24)$$

式中， s_i 为每个内点的计算权重， $E = \sum_{i=1}^k e_i$ 为 k 个内点于初始模型求解状态下的误差总和， $\bar{e}$ 表示 k 个内点误差的平均值。

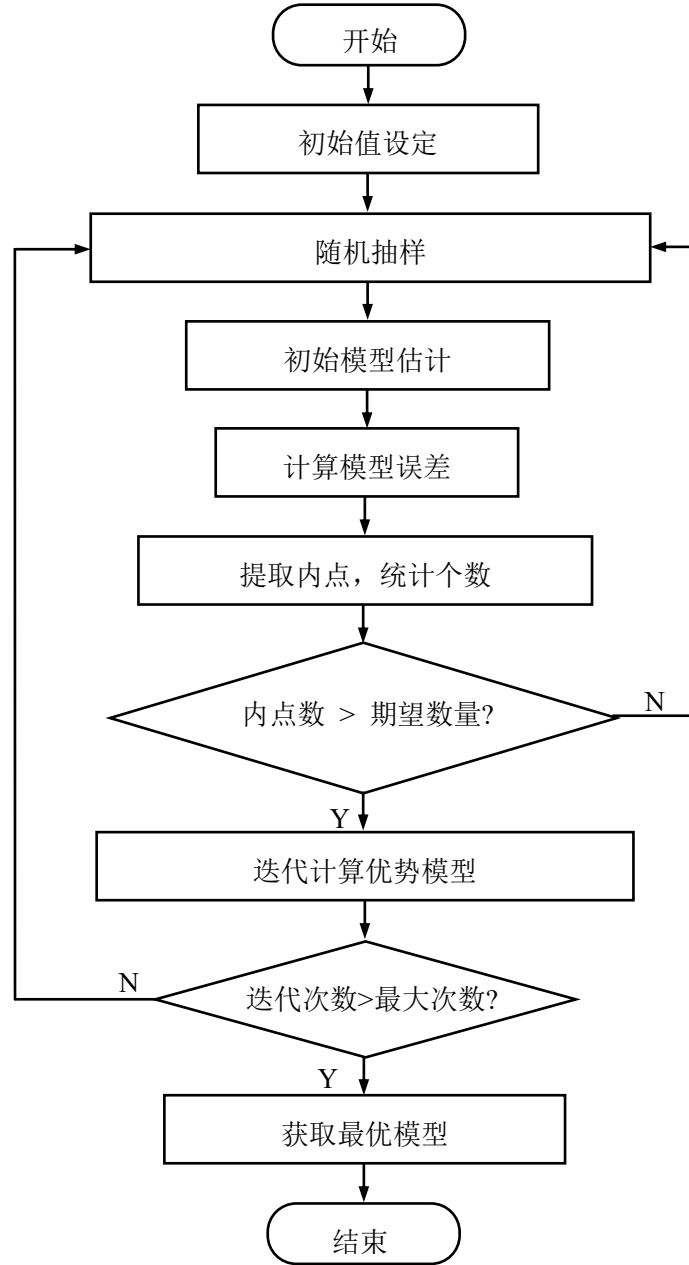


图 4-3 坐标变换优化流程图

在获得每个内点的计算权重 s_i 后, 将所得各点的权重值 s_i 与测量值的偏移量矩阵 $\mathbf{X}_p^r$ 进行乘积, 得到加权后的偏移量矩阵 ${}_s\mathbf{X}_p^r$, 对应的数学表达如下:

$${}_s\mathbf{X}_p^r = \left\{ s_1 \left(p_1^c - \mathbf{X}_p^r \right), s_2 \left(p_2^c - \mathbf{X}_p^r \right), s_3 \left(p_3^c - \mathbf{X}_p^r \right), \cdots, s_k \left(p_k^c - \mathbf{X}_p^r \right) \right\} \quad (4-25)$$

将更新后的 ${}_s\mathbf{X}_p^r$ 代入式(4-21)中, 就可以获得单次迭代中, 加权奇异值分解后的坐标变换模型。

为确保当前样本下的坐标变换模型最大程度逼近真实值, 在计算流程中设计了加权 SVD 的迭代过程, 通过更新测量权重值和坐标变换模型参数, 直至样本的

测量误差收敛，加权 SVD 具体的迭代步骤如下：

(1) 设置误差收敛阈值 ε ，该收敛阈值需满足 $\varepsilon \Rightarrow E_i - E_{i+1} \leq 0.05E_i$ ，通过式(4-22)和式(4-23)判断当前模型是否趋优，若是，开始迭代过程。

(2) 通过加权奇异值分解计算单次模型参数姿态旋转矩阵 $\mathbf{R}_c^r$ 和位移向量 $\mathbf{t}_c^r$ 。

(3) 将所得参数重新代入式(4-22)，判断相邻两次模型参数重构的误差值是否小于误差收敛阈值 ε ，若是，则终止迭代过程，否则返回步骤 (2)。

通过以上 3 个步骤，能够得到 RANSAC 单一采样次数下的结果：得到一个趋优样本对应的坐标变换模型和重构误差值，或者被舍弃。RANSAC 通过不断采样，重复这过程，并直至最大采样次数。算法通过审阅每一个趋优样本对应的重构误差值，将平均重构误差最小的趋优样本对应的坐标变换模型作为优化后的最终解，图 4-3 给出了坐标变换优化方法的具体流程。

4.3 机器人姿态测量

机器人姿态测量是机器人运动学参数标定的关键步骤，测量精度直接影响标定精度。本节在根据 4.2 节所述的空间定位方法的基础上，通过三角形合作标识来建立末端坐标系，给出了机器人末端姿态检测方法。

4.3.1 运动坐标系构建

典型的视觉测量方法为采用双目视觉系统获取安装在机器人末端上的多个靶标点位坐标，通过检测靶标点位的运动变化状态，从而实现末端姿态的计算与测量。这类测量方法通过测量机器人末端上任意不共线的 $n \geq 3$ 个靶标点位坐标，从而建立起以视觉系统为观测基准，机器人末端为 3D 基元的运动坐标系，实现机器人末端位姿测量。为此，设计了如图(4-4)直角三角形的合作标识，将直角三角形三个顶点作为特征点，得到其在相机坐标系下的空间坐标，表示为 $P_1 = (x_1, y_1, z_1)$ ， $P_2 = (x_2, y_2, z_2)$ ， $P_3 = (x_3, y_3, z_3)$ ，根据三角形的直角约束，则运动坐标系的基向量可表示为 $\vec{n}_x = \overrightarrow{P_1P_2} = [x_2 - x_1, y_2 - y_1, z_2 - z_1]$ ， $\vec{n}_y = \overrightarrow{P_1P_3} = [x_3 - x_1, y_3 - y_1, z_3 - z_1]$ ， $\vec{n}_z = \vec{n}_x \times \vec{n}_y$ 。进一步，运动坐标系的坐标轴可表示为各基向量的归一化：

$$\begin{cases} \vec{X}_{tcp} = \frac{\vec{n}_x}{\|\vec{n}_x\|} = \frac{[x_2 - x_1, y_2 - y_1, z_2 - z_1]}{\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}} \\ \vec{Y}_{tcp} = \frac{\vec{n}_y}{\|\vec{n}_y\|} = \frac{[x_3 - x_1, y_3 - y_1, z_3 - z_1]}{\sqrt{(x_3 - x_1)^2 + (y_3 - y_1)^2 + (z_3 - z_1)^2}} \\ \vec{Z}_{tcp} = \frac{\vec{n}_x}{\|\vec{n}_x\|} \times \frac{\vec{n}_y}{\|\vec{n}_y\|} \end{cases} \quad (4-26)$$

规定运动坐标系的原点为三点的质心位置，表示为 $CT = \frac{1}{3} \sum (P_1 + P_2 + P_3)$ ，则由相机坐标系到运动坐标系的旋转矩阵可表示为 $R_{TCP}^C = [\bar{X}, \bar{Y}, \bar{Z}]$ ，平移向量 $t_{TCP}^C = [CT]$ 。

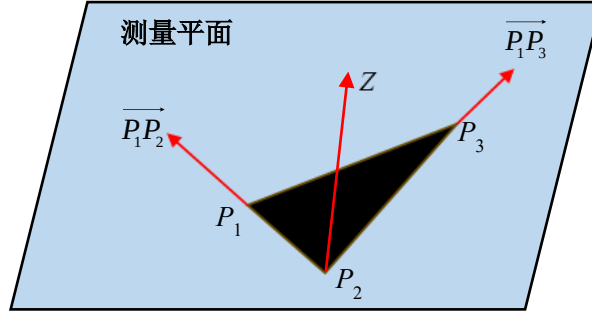


图 4-4 运动坐标系构建示意图

4.3.2 旋转矩阵优化

自建坐标系的模型精度受限于特征点定位误差，这导致旋转矩阵存在非正交误差。设无偏差的旋转矩阵为 R ，实际附有噪声干扰所得的旋转矩阵为 R_0 ，所得的优化正交矩阵为 R_1 则对应的误差模型可表示为：

$$R_0 = R + N \quad (4-27)$$

所优化的目标函数可表示为：

$$f(R_1) = \arg \min \|R_1 - R_0\|_F^2 \quad (R_1 R_1^T = I) \quad (4-28)$$

式(4-28)中， $\|R_1 - R_0\|_F^2$ 展开可得到：

$$\|R_1 - R_0\|_F^2 = \text{tr}((R_1 - R_0)^T (R_1 - R_0)) \quad (4-29)$$

因为 R_1, R_0 属于正交矩阵 $\mathbb{R}^3$ ，则式(4-32)可进一步整理为：

$$\|R_1 - R_0\|_F^2 = 3 + \text{trace}(R_0^T R_0) - 2\text{trace}(R_1^T R_0) \quad (4-30)$$

观测式(4-30)发现， $3 + \text{trace}(R_0^T R_0)$ 为定量，因此优化方程可转换为：

$$\arg \min \|R_1 - R_0\|_F^2 \Rightarrow \arg \min (\text{trace}(R_1^T R_0)) \quad (4-31)$$

使用 SVD 分解法将 R_0 展开为 $R_0 = USV^T$ ，带入(4-31)式可得：

$$\arg \min (\text{trace}(R_1^T R_0)) = \arg \min (\text{trace}(R^T R_1^T US)) \quad (4-32)$$

因为 R_1 需要满足正交性，则 $V^T R_1^T U = I$ ，因此式(4-32)可推导为：

$$\text{trace}(\mathbf{V}^T \mathbf{R}_1^T \mathbf{U} \mathbf{S}) = \text{trace}(\mathbf{I} \mathbf{S}) = \sum_{i=1}^3 \sigma_{ii} s_i \leq \sum_{i=1}^3 s_i \quad (4-33)$$

当且仅当 $\mathbf{R}_1 = \mathbf{R}^T$ 时等号成立，得到 $f(\mathbf{R}_1) = \arg \min \|\mathbf{R}_1 - \mathbf{R}_0\|_F^2$ 的最优解。进一步，令 $\mathbf{R}_{TCP}^C = \mathbf{R}_1 = \mathbf{U} \mathbf{V}^T$ ，则运动坐标系到 $O_{TCP}\text{-XYZ}$ 相机坐标系 $O_C\text{-XYZ}$ 的变换

$$\text{关系可表示为 } \mathbf{T}_{TCP}^C = \begin{bmatrix} \mathbf{R}_{TCP}^C & \mathbf{t}_{TCP}^C \\ \mathbf{0}^T & 1 \end{bmatrix}。$$

通过运动坐标系建系和旋转矩阵正交优化方法，可得观测坐标系与运动坐标系的转换关系。为统一观测值，一般的，工业机器人的末端姿态被表征于机器人基坐标系中。设机器人末端位姿在 i 时刻的位姿表示为 Rot_i ，在机器人基坐标的表达为 Rot_i^B ，在相机坐标下的观测值为 Rot_i^C ，理论上，位姿变换满足以下关系：

$$\text{Rot}_i^B = \mathbf{T}_{TCP}^C \mathbf{T}_C^B \text{Rot}_i \quad (4-34)$$

上述坐标系构建变换过程中， $\mathbf{T}_{TCP}^C$ ， $\mathbf{T}_C^B$ 均为已知量，因此可通过式(4-34)可得到任意时刻下机器人运动的观测姿态。

4.4 本章小结

在获得合作标识特征的亚像素坐标后，本章详细推导了机器人平面运动与空间位姿的测量原理和数学模型。提出了基于单应特征矩阵的机器人平面运动视觉测量模型，相较于基于正交投影的平面测量方法，所提方法能够消除相机光轴与测量平面非正交误差，提高视觉系统的平面测量精度与灵活性。提出了基于 RANSAC 加权约束的三维坐标变换模型优化方法，旨在抑制视觉系统的测量噪声对坐标变换模型精度的干扰，达到视觉系统的测量数据与机器人运动数据的配准。基于三角标识构建了机器人运动坐标系，并给出了坐标系正交优化方法，以实现对机器人末端姿态的准确观测。

第 5 章 工业机器人运动视觉测量实验

本章通过搭建视觉系统和实验平台，验证本文所提出的研究方法的正确性，测量方案的可行性，评估提出的方法在视觉测量任务中的性能。进行包括相机标定实验，机器人运动追踪与特征点像素定位实验，在此基础上，完成机器人平面运动测量以及空间运动测量，同时将本文所提的测量模型与优化方法与传统模型进行精度对比，分析测量误差，验证算法的可行性和有效性。

5.1 实验平台搭建

5.1.1 系统硬件组成

为验证视觉测量方案的可行性和所提测量算法的有效性，搭建如图 5-1 所示双目视觉机器人运动测量系统，系统包括：（1）组成双目测量单元的两台高速相机和镜头，参数如表 5-1、表 5-2 所示；（2）用于相机安装的三脚架；（3）用于图像采集和处理的上位机控制器；相机采用汇聚式安装方式。被测对象为三自由度直角坐标机器人，六自由度机器人。其中，三自由度直角坐标机器人作为平面运动与空间定位的测量对象。其空间运动最大范围为 400 mm*400 mm*300 mm（X、Y、Z 轴），机器人 XYZ 方向上运动最小分辨精度为可达 0.1um、0.1um、0.08um。六自由度机器人作为空间位姿测量的对象，其重复度定位精度为 0.05mm。

表 5-1 相机参数

传感器	分辨率 (H×V)	帧数 (s ⁻¹)	位深	像元尺寸 (um ²)	快门模式
COMS	2320×1720	<13000	8bit	7×7	Global

表 5-2 镜头参数

光圈范围	滤镜螺纹直径 (mm)	焦距 (mm)	放大倍率	视角
F/16-F/1.4	77	24	1:5.3	84.1°

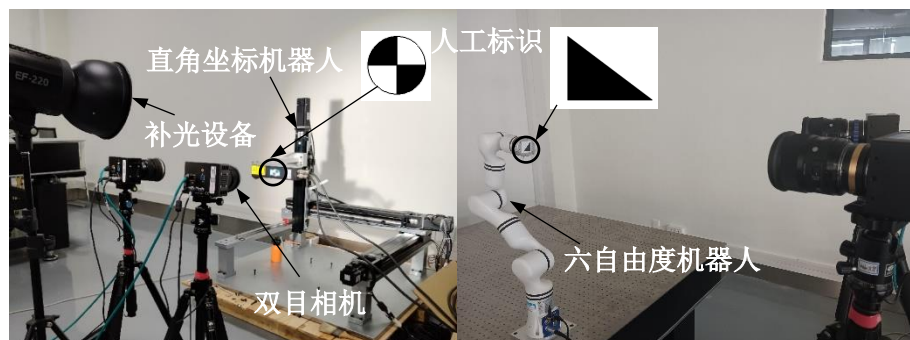


图 5-1 搭建的实验装置系统

实验开始前,调整相机位置,使其视野范围充分包含机器人运动范围。接着,调整镜头焦距,确保成像清晰。启用辅助光源,采用正向补光方式,调整光强,保证机器人运动过程中光强稳定。以上步骤确保了适宜的视觉环境,为后续实验提供了优良条件。

5.1.2 软件环境

本实验的主要编程语言包含 C++, Python, MATLAB。其中,使用 Visual Studio 编写了机器人运动控制程序,并基于 Windows 图像交互设计工具 MFC 设计了三自由度机器人运动控制界面;基于 Anaconda 生成虚拟环境,使用 Pycharm 联合 YOLOv5 的标注工具 Labelimg 完成了数据集的标注工作,采用 Pytorch 深度学习的框架完成了 YOLOv5s 模型部署。同时,图像处理所使用的库函数主要包括计算机视觉工具 OpenCV、矩阵与数据运算工具 Numpy 等。

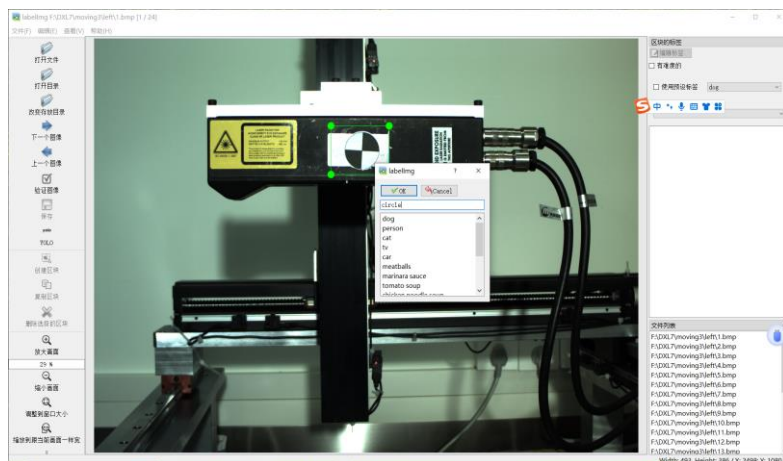


图 5-2 Labelimg 标注工具界面



图 5-3 机器人空间交互界面

在获得图像数据,并完成特征点定位提取等处理后,在 MATLAB 2020b 环境中编写了包括平面运动测量模型,空间运动测量模型以及位姿定位的具体实现算法,以对图像数据进行整合,得到机器人运动测量的分析结果。

5.1.3 相机标定实验

在测量伊始的首要任务是进行相机标定，得到视觉系统测量参数和镜头畸变系数。标定所使用的参考物为图 5-4 所示的黑白棋盘格标定板，其角点规格为 12×9 ，角点间距为 45mm。进行标定过程中，为确保标定结果的普适性，应初步估计测量所涵盖的空间范围，进而，标定范围应该涵盖所测量的空间，同时，在不同位置、角度和姿态下采集足够数量的图像。完成标定图像采集后，基于 MATLAB 相机标定工具箱计算标定结果，标定的空间示意图如图 5-5 所示。该工具箱提供了标定参数的计算功能，包括相机的内参数（焦距、主点坐标）和外参数（旋转矩阵、平移向量）。同时，工具箱提供了数据可视化交互界面，能够在标定过程中自由选择标定图像，以降低人为误差。

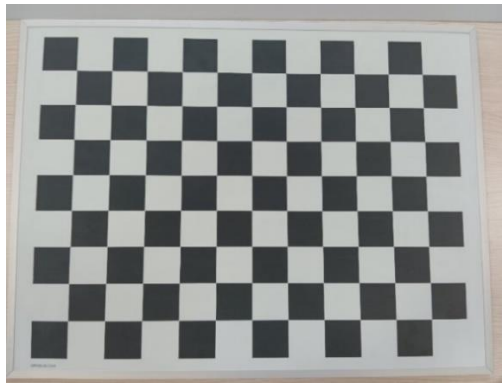


图 5-4 标定板实物图

得到标定结果后，基于标定的重投影误差来评估标定的精度。通过若干次标定实验，不断调整两相机的相对位姿，基线距离，光轴夹角，旨在尽可能降低系统结构对测量精度的影响。最终，将重投影误差最小的标定结果对应的相机参数如下 $\mathbf{K}_L = [3587.846, 0, 1201.472; 0, 3588.021, 844.081; 0, 0, 1]$; $\mathbf{K}_R = [3578.831, 0, 1172.704; 0, 3578.034, 843.969; 0, 0, 1]$; $\mathbf{R} = [0.936, -0.065, 0.346; 0.060, 0.998, 0.025; -0.346, -0.003, 0.938]$; $\mathbf{t} = [-233.144; -15.739; 50.715]$; 左相机的畸变系数为 $d_l = (-0.049, -0.111)$ ，右相机的畸变系数为 $d_r = (-0.074, 0.114)$ 。由图 5-6 可知，本次标定的最大重投影误差为 0.09 像素，平均重投影误差为 0.05 像素。

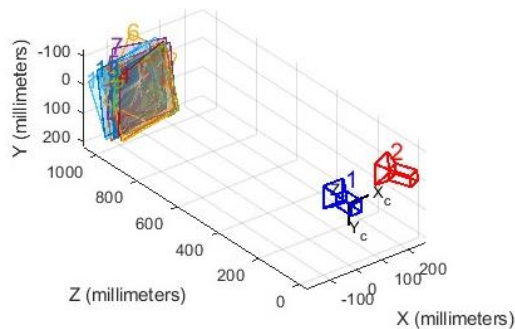


图 5-5 相机标定空间示意图

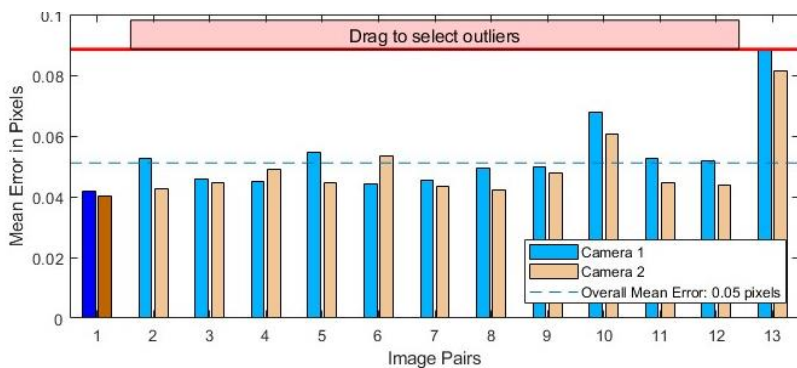


图 5-6 标定重投影误差结果

5.2 机器人运动追踪与特征点定位实验

5.2.1 机器人运动追踪

完成相机标定实验后，视觉系统开始采集机器人运动数据。使用机器人控制交互界面，将机器人运动控制设定为点动控制，将视觉系统的图像采集模型调整为手动抓拍，以获取机器人运动的空间离散点位置与平面离散点位置对应图像；在机器人控制交互界面编程设定机器人运动路径，同时视觉系统调整至高速采集模式，开始采集机器人连续空间运动，获取机器人连续直线轨迹与圆弧轨迹的图像序列。

在此基础上，将所采集部分图像作为 YOLOv5 训练的数据样本，采用离线训练的方式，将数据样本分为训练集，测试集与验证集。由于模型训练数据并非标准数据，存在数据量少，种类单一的局限。为此，在训练过程中增大了验证集的数据量，用于优化模型的超参数、评估模型性能和进行早期停止等操作，训练集，测试集与验证集的数据量比例为 6:3:1。

此外，为提升训练集的多样性，对训练集的图像进行了增噪，模糊，降低图像灰度值等图像增强的方法，以提升训练所得检测器的泛化能力。训练所采用的部分数据如图 5-7 所示。

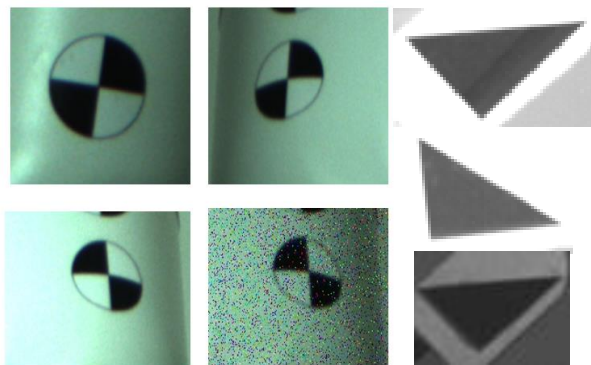


图 5-7 YOLO 训练集部分数据

改变原有 YOLO 模型的训练路径，将原有的 COCO 数据集改为自制数据集路

径，将 YOLO.yama 文件的目标类别调整为自制标识标签类别。模型训练轮次设定为 300 轮，调整 YOLO 参数如表 5-3 所示，训练过程使用 GPU 加速，最终，得到了针对人工标识的检测器模型。

表 5-3 YOLOv5s 超参数设置表

超参数设置					
定位损失系数	0.05	初始学习率	0.1	置信度阈值	0.45
分类损失系数	0.5	循环学习率	0.1	NMS 阈值	0.25
定位损失阈值	0.2	学习率动量	0.9	交互比阈值	0.45
训练轮数	300	底层网络模型	CNN	图像输入尺寸	640*640
权重衰减系数	0.005	目标类别	1	锚框宽高比	1:1

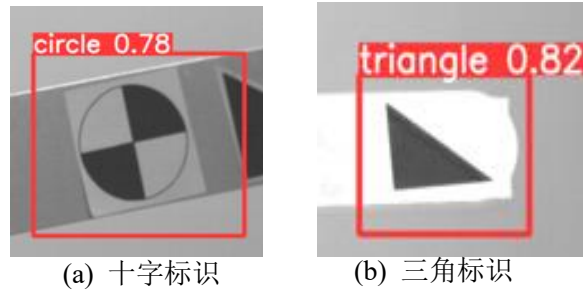


图 5-8 YOLO 检测器训练结果

由图 5-8 可知，训练所得的检测器模型已经能够准确识别人工标识。应当明确的是，在本次目标追踪实验中，本次所检测的目标对象仅为特定标识，因此目标类别的置信度对结果没有影响，故而不做进一步讨论。

使用 YOLO 检测器对机器人运动图像序列进行检测，得到标识在各图像上的具体位置。这些位置信息以归一化的形式表达，并使用目标框的左上角坐标以及整个图像的宽高比来描述。具体而言，检测器对每张图像进行目标检测，并提供以下位置信息：1) 目标框的左上角坐标：目标框是一个矩形边界框，表示目标物体在图像中的位置。通过提供目标框的左上角坐标，可确定目标框在整体图像中的位置。2) 目标框区域与整体图像的宽高比：目标框的宽高比反映了目标物体在图像中的形状。通过提供目标框的宽高比，可以得知物体相对于整个图像的尺度比例。

5.2.2 特征点识别与定位

使用 YOLOv5 对机器人运动进行追踪，能够初步定位标识于图像上的相对位置，在此基础上，同时，基于 3-3 节所提方法，完成特征点像素坐标定位。图 5-9 显示了圆形内十字特征点定位的示意图，针对标识十字交线部分，进行了以下处理：首先对标识区域进行二值化，并进行边缘检测，得到如图 5-9(a)所示边缘图像。其次，使用 Zernike 矩求解亚像素边缘如图 5-9(b)所示。

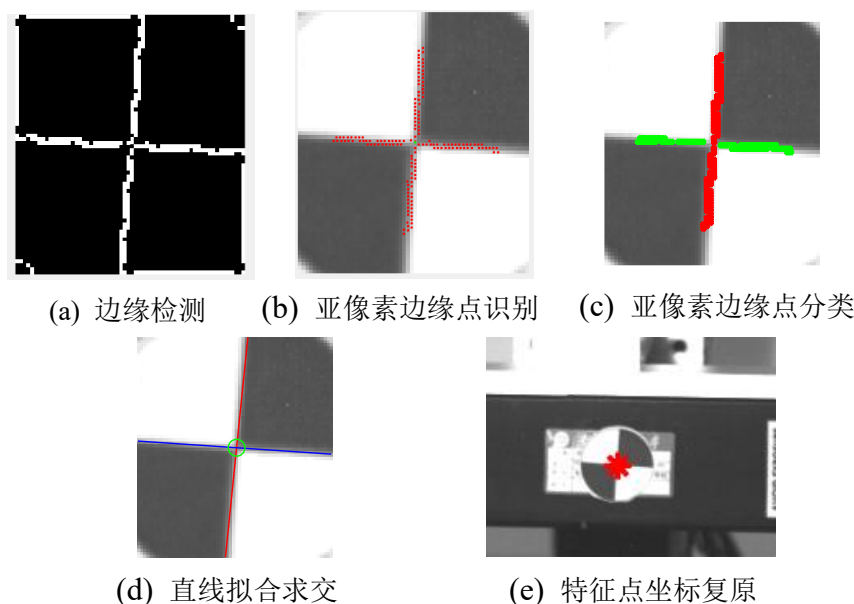


图 5-9 圆形十字特征点处理示意图

同时，对各亚像素边缘点的坐标数据求质心，得到参考点。再次，通过各边缘点到参考点连线的斜率区分出横线边缘点与竖线边缘点，如图 5-9(c)所示。进而，使用最小二乘直线拟合，得到横向与竖线的直线方程，并求两直线的交点坐标，如图 5-9(d)所示。最后，通过 YOLOv5 所得的标识图像子区域初定位的像素坐标，将特征点的坐标复原至图像上如图 5-9(e)所示。

同理，对于三角形标识特征点像素坐标定位的可视化处理过程如图 5-10 所示。

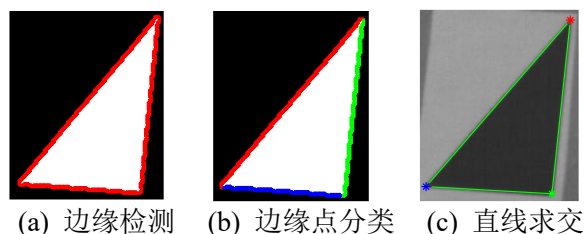


图 5-10 三角形标识图形处理过程

首先，通过直角三角形特征点定位方法得到标识的三个特征点，图 5-10 展示了部分图像处理的结果，其中，通过 Canny 得到三角形标识的初始边缘对应图 5-10(a)，通过 Zernike 矩计算得到亚像素边缘(对应的处理过程已在图 5-9 详细展示)，进而，对亚像素点进行分类，得到所属三角形的各边，如图 5-10(b)所示。其中，标记为蓝色的亚像素点群对应直角短边 Ls_1 ，绿色的亚像素点群对应直角长边 Ls_2 ，红色的亚像素点群对应斜边 Ls_3 。进而，图 5-10(c)展示了特征点定位的结果。其中，蓝色标记表示 Ls_1 与 Ls_3 的交点 p_1 ，绿色标记表示 Ls_1 与 Ls_2 的交点 p_2 ，红色标记表示 Ls_2 与 Ls_3 的交点 p_3 。最后，通过 YOLOv5 所得的标识图像子区域初定位的像素坐标，将直角三角形的三个顶点特征点的坐标复原至图像上。

5.3 机器人全局结构运动视觉测量实验

5.3.1 机器人平面结构运动测量

表 5-4 平面测量点测量坐标

机器人坐标			像素坐标		机器人坐标			像素坐标	
mm			pix		mm			pix	
	Y	Z	u	v		Y	X	u	v
1	0.024	-30.053	753.357	753.504	11	0.024	0.024	759.430	968.742
2	50.051	-30.053	930.154	745.273	12	0.024	50.044	939.136	959.720
3	100.074	-30.053	1109.111	737.831	13	0.024	100.067	1116.597	950.599
4	150.099	-30.053	1283.823	730.565	14	0.024	150.088	1291.183	941.804
5	200.123	-30.053	1455.592	723.132	15	0.024	200.111	1460.036	932.937
6	0.024	-60.08	754.913	861.053	16	0.024	0.024	764.237	1075.751
7	50.048	-60.08	934.496	852.496	17	0.024	50.047	945.107	1066.139
8	100.073	-60.08	1110.588	846.757	18	0.024	100.071	1120.965	1056.730
9	150.097	-60.08	1290.656	836.181	19	0.024	150.095	1296.928	1047.110
10	200.121	-60.08	1460.731	828.580	20	0.024	200.118	1467.266	1037.716

控制机器人进行二维平面的点运动，将运动平面约束在 $X=0$ 平面内，在机器人运动过程中使用视觉系统同步采集观测数据，得到机器人运动坐标与观测坐标如表 5-4 所示。基于所观测的 20 个测量点，通过单应性平面测量理论，计算得到单应性矩阵 H 的具体结果为 $[3.739, -0.169, 746.699; -0.042, -3.619, 645.056; 0, 0, 1]$

使用所得的单应性矩阵 H ，将特征点的像素坐标重构值机器人坐标系中，重构结果如图 5-11 所示，进一步，观察测量误差图 5-12 可知，重构坐标与机器人坐标保持一致，误差限制在 1mm 以下，反映了所提方法的正确性。

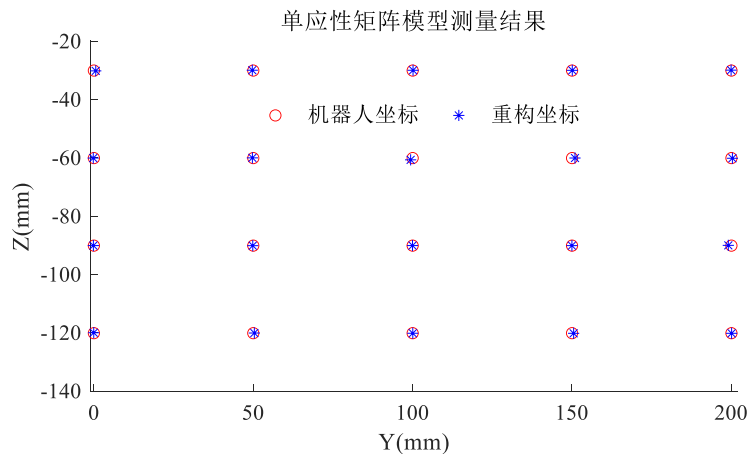


图 5-11 平面单应性模型测量结果

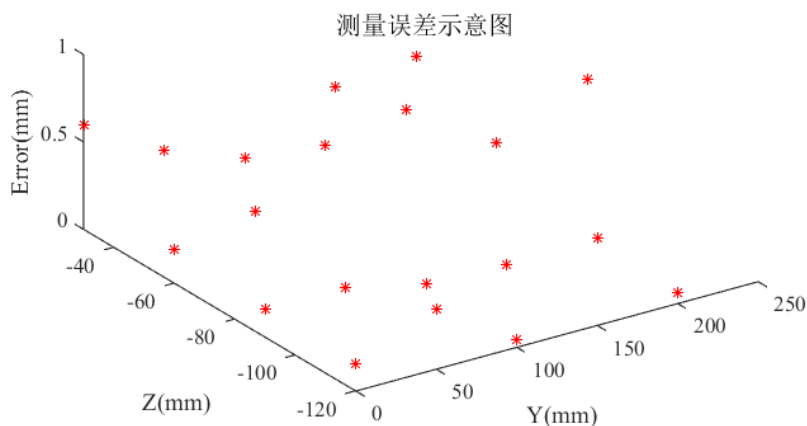


图 5-12 单应性模型测量误差图

为验证单应性矩阵方法的优越性，给出了一组对比实验。选择四个最大矩形点，序号为 1 和 5 的测量点计算像素坐标系 u 轴与 Y 轴的比值，序号为 1 和 16 的测量点计算像素坐标系 v 轴与 Z 轴的比值。通过比例模型计算其余各参考点对应的观测值，所得结果如图 5-13 所示。

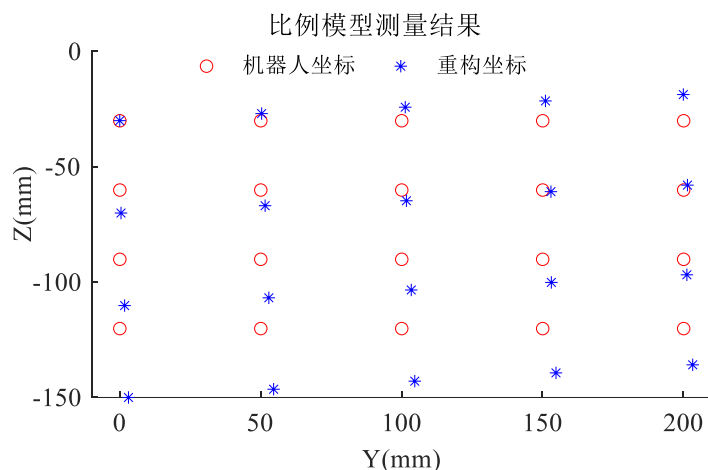


图 5-13 平面比例测量示意图

观察图 5-13，可以发现，基于比例模型计算所得的测量值与参考点的坐标发生了明显偏差，这证实了相机光轴与测量平面不垂直时，比例模型的测量精度会显著下降。

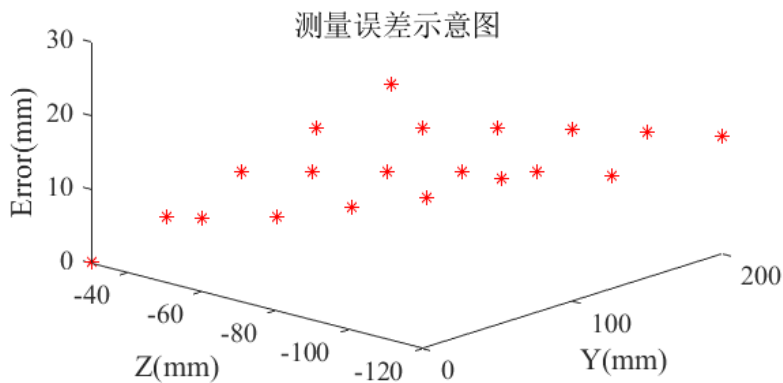


图 5-14 比例模型测量误差图

表 5-5 平面运动处理误差对比

所用方法	误差项	最大值	均值	标准差
比例模型		30.025mm	12.021mm	72.884mm
单应矩阵模型		0.976mm	0.322mm	0.089mm

图 5-14 给出了比例模型所得的测量误差数据，以测量的最大误差，平均误差和误差标准差来评价模型精度。可以发现，除规定矩形边角的参考点无误差外，其余测量的各点的误差均在 10mm 以上，表 5-5 报道了两种测量模型的误差对比，基于单应矩阵模型的最大测量误差为 0.976mm，平均误差为 0.322mm，误差标准差为 0.089mm，均优于比例测量模型。

5.3.2 机器人空间结构运动测量

根据 4.2 节内容叙述的机器人空间运动的测量原理，展开了机器人空间结构实验。控制机器人带动末端执行单元沿指定的空间位置运动，在机器人的运动空间范围内，共在 28 个位置采集标识图像作为参考点，对应的空间分布如图 5-15 所示。

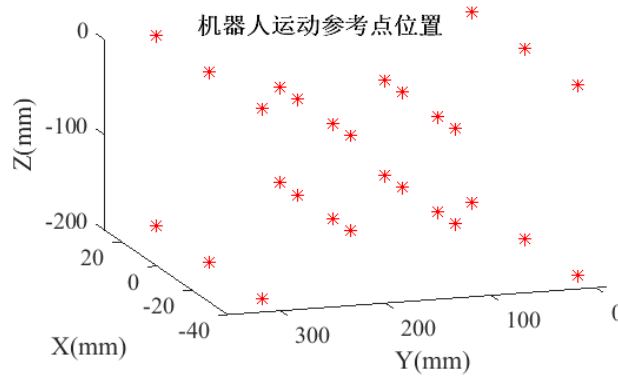


图 5-15 参考点空间位置分布

视觉系统同步采集标识图像，得到标识特征点对应的左右相机的像素坐标，通过 5.2.1 所给出的相机标定数据，重构于左相机坐标系下，参考点的机器人坐标和对应的相机坐标如表 5-6 所示。

表 5-6 参考点坐标数据

No.	左相机坐标系坐标(mm)			机器人坐标系坐标(mm)		
	X	Y	Z	X	Y	Z
1	-116.309	-37.338	1313.581	-29.955	0.024	-0.027
2	178.375	-43.095	1367.911	-29.955	300.048	-0.027
3	179.027	156.499	1385.171	-29.955	300.048	-200.054
...	...	...	...	...	...	...
26	89.409	13.0203	1304.798	20.046	200.065	-50.054
27	89.738	112.774	1313.730	20.046	200.065	-150.081
28	-8.5649	114.477	1295.628	20.046	100.087	-150.081

为验证坐标变换矩阵求解算法的正确性和精度，基于表 5-6 中的数据，分别采用传统算法和本文所提的坐标变换优化方法求取相机坐标系到机器人坐标系的变换模型。使用本文所提方法求解坐标变换矩阵参数过程中，为确保采样的充分性，RANSAC 算法的采样次数上限设置为 50000 次。以每次采样后所求变换矩阵将相机坐标系下的三维坐标变换到机器人坐标系后，与实际机器人基坐标值的欧式距离误差均值作为判定样本集中内外点的依据。所求 ${}^B_C\mathbf{T}$ 的结果如表 5-7 所示。为说明所提方法的模型求解精度，计算相机坐标系下坐标变换至机器人坐标系下的误差值，结果如图 5-17 所示，表 5-8 给出了误差最大值、均值和标准差的分析结果。

表 5-7 传统方法和本文改进方法结果

所用方法	参数	${}^B_C\mathbf{R}$	${}^B_C\mathbf{t}$
传统方法		[0.179, 0.087, -0.981;	[1281.732;
		0.984, -0.019, 0.178;	-120.579;
		-0.004 -0.996, -0.088]	78.306]
本文所提方法		[0.182, 0.086, -0.980;	[1281.176;
		0.983, -0.020, 0.181;	-124.583;
		-0.004, -0.996, -0.088]	78.116]

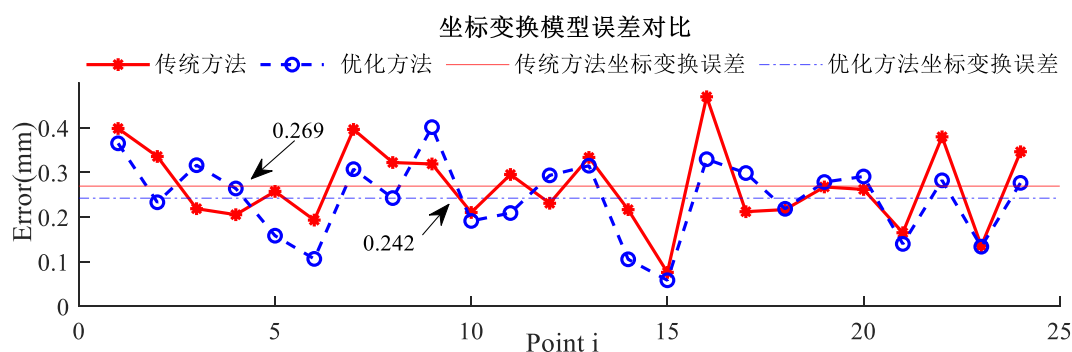


图 5-16 坐标变换模型误差对比结果

表 5-8 传统方法和本文改进方法变换误差比较

所用方法	误差项	最大值	均值	标准差
传统方法		0.469mm	0.269mm	0.093mm
本文所提方法		0.401mm	0.242mm	0.089mm

由图 5-16 和表 5-8 可知，传统方法所求模型精度受测量点误差的影响较为严重，相同数据下，使用本文所提算法的最大误差相比传统方法减少 14.5%，由 0.469mm 减小为 0.401mm；误差均值相比传统方法减少 10.0%，由 0.269mm 减小为 0.242mm；同时，优化后模型误差的标准差相比传统方法减少 4.3%，由 0.093mm 减小为 0.089mm，所提方法能够有效降低测量点噪声对基准变换模型的影响，相

较传统方法具有更高的求解精度和抗噪声干扰能力。

为进一步验证所提方法的模型求解精度和对视觉测量系统的提升效果，结合所提方法开展机器人连续轨迹运动测量实验。以直线和圆弧两种常见的运动轨迹形式，来验证基于所提方法的轨迹测量精度。所测量运动轨迹参数如表 5-9 所示，控制机器人末端沿指定轨迹运动，设定机器人运动的线速度为 450mm/min。

表 5-9 机器人运动轨迹参数

直 线	起点坐标 (-9.536, 52.929, -50.429)	终点坐标 (35.444, 142.887, -140.387)	方向向量 [0.373; 0.746; -0.746]
圆 弧	起点坐标 (10.047, 100.053, -0.027)	终点坐标 (10.047, 50.049, -50.054)	圆心, 半径 (10.047, 100.051, -50.041), $r=50.013$

图 5-17 显示了机器人末端标识的实际运行轨迹和视觉系统的测量轨迹。可以看出，测量轨迹和实际轨迹保持一致，改进方法能够实现机器人运动正确测量。

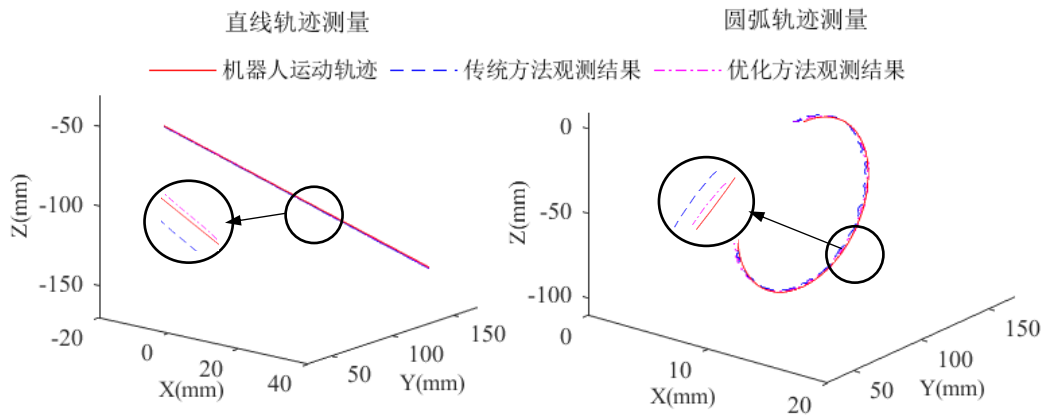


图 5-17 机器人末端运行轨迹和测量轨迹

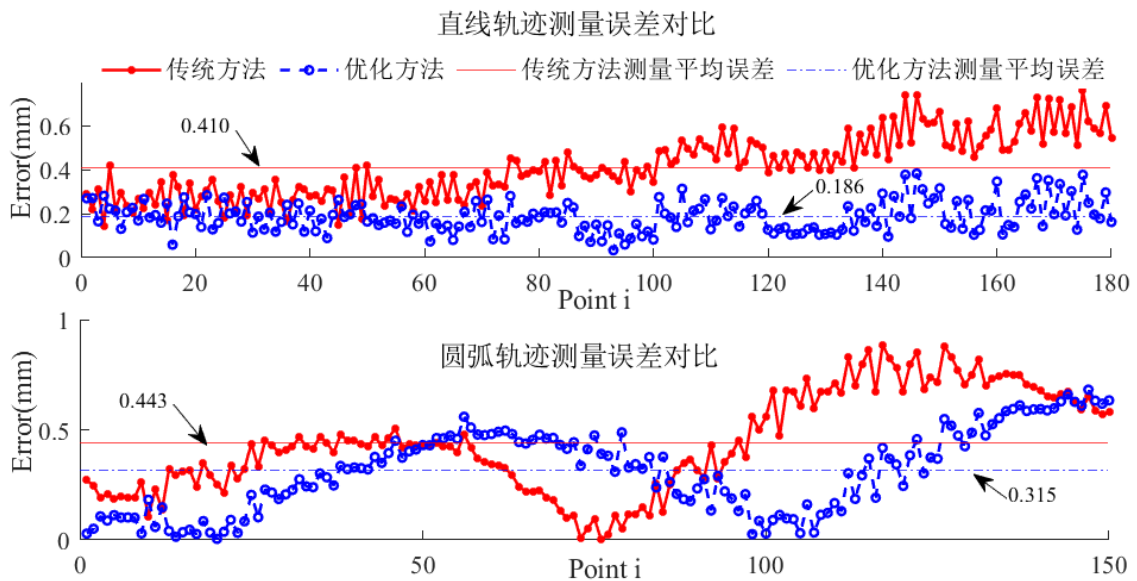


图 5-18 机器人运动轨迹测量偏差

表 5-10 机器人空间运动测量轨迹偏差比较

所用方法	误差项	最大值 (mm)	均值 (mm)	标准差 (mm)
直线轨迹	传统方法	0.759	0.410	0.142
	优化方法	0.382	0.186	0.070
圆弧轨迹	传统方法	0.885	0.443	0.224
	优化方法	0.682	0.315	0.185

为进一步验证测量精度，计算测量点至实际轮廓的轨迹偏差，结果如图 5-18 和表 5-10 所示。测量结果表明，采用传统方法直线轨迹测量轮廓偏差的最大值为 0.759mm，均值为 0.410mm，标准差为 0.142mm；基于所提方法的直线轨迹测量轮廓偏差的最大值为 0.382mm，均值为 0.186mm，标准差为 0.070mm，最大误差减小 49.7%，均值减小 51.3%，标准差减小 50.7%；采用传统方法圆轨迹测量轮廓偏差的最大值为 0.885mm，均值为 0.443mm，标准差为 0.224mm；基于所提方法的圆轨迹测量轮廓偏差的最大值为 0.682mm，均值为 0.315mm，标准差为 0.185mm，最大误差减小 22.9%，均值减小 28.9%，标准差减小 17.4%。

为进一步验证所提优化方法较之其他坐标变换方法的先进性，在原有的实验基础上给出了一组对比实验。重新驱动机器人在空间中运动，并改变机器人运动的轨迹参数、空间位置和速度。

表 5-11 显示了速度设置为 900 毫米/分钟，修改后运动轨迹参数。同时，在重复实验中添加了对比方法，使用四元数和 ICP 来求解坐标变换模型，并对机器人的运动进行测量。

表 5-11 重复实验轨迹参数

直线	起点坐标	终点坐标	方向向量
	(-20.350,150.520,0.530)	(20.070,50.540,-100.550)	[0.275; -0.679; -0.680]
圆弧	起点坐标	起点坐标	圆心，半径
	(0.047, 70.053, -0.027)	(20.043, 70.049, -0.054)	(-0.120,70.140,21.323), $r=20.002$

从图 5-19 可以看出，对于同一直线和圆轨迹的运动测量，基于四元数和 ICP 方法求解所得坐标变换模型，在用于机器人运动测量的误差均高于最小二乘解法。原因在于，四元数计算需要归一化，这会引入额外的计算误差，而 ICP 在没有预配准算法的辅助下，对初始匹配点误差干扰会影响迭代过程，导致误差累积。四元数和 ICP 的最大测量误差均超过 1mm，平均误差超过 0.9mm。进一步分析测量结果，得到了如图 5-20 的重复实验误差分析图，分析结果表明，本文所提出的优化方法的测量精度趋于最优，具体而言，优化方法所测量的直线轨迹偏差的最大值为 0.482mm，平均值为 0.248mm；所测量的圆弧轨迹偏差的最大值为 0.687mm，平均

值为 0.310mm。

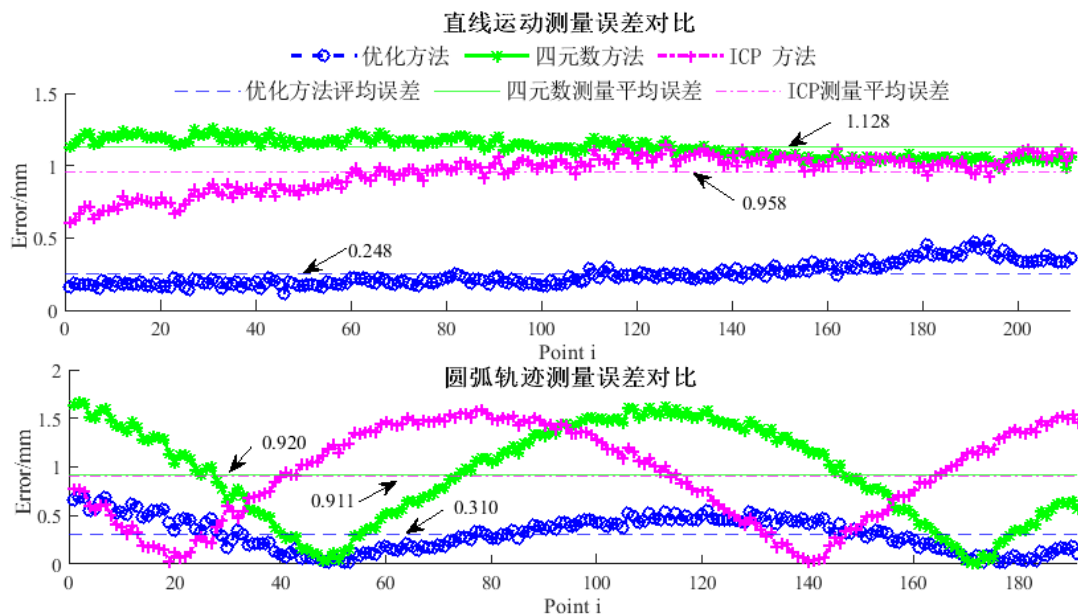


图 5-19 3D 测量重复试验误差对比

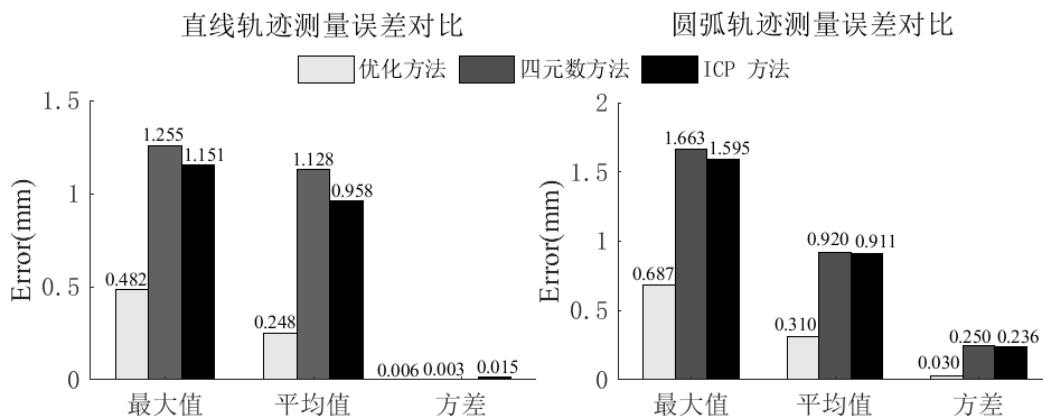


图 5-20 重复试验误差分析结果

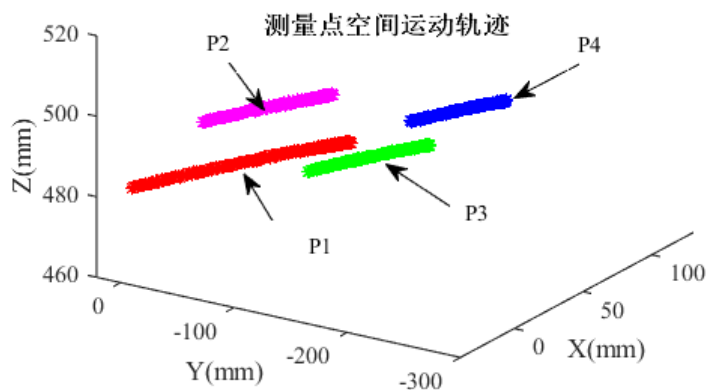


图 5-21 全局结构运动测量轨迹

为展示视觉系统在全局测量方面的能力，通过在机器人关节结构上安装 4 个标识，以模拟全局测量过程。视觉系统同时捕捉这 4 个标识的空间运动，通过本文第三章，第四章所给出的追踪，定位方法和测量模型将其重建到相机坐标系中。通

过这一过程，得到了机器人全局结构运动的测量轨迹，具体呈现在图 5-21 中。从全局结构运动的测量结果来看，本文提出的测量方案，能够有效且准确地完成对多个标识的追踪定位和测量需求。这进一步证明了本文所提方案的可靠性，能够对机器人全局结构的运动进行观测和精准测量。

5.3.3 机器人姿态测量

在 4.2 节的基础上，进行了姿态测量实验，使用三角形标识粘贴在机器人末端，驱动机器人进行位姿变换，机器人初始位姿 $[R_x, R_y, R_z; t_x, t_y, t_z]$ 数值为 $[1.845, 1.293, 1.672; 60.184, -39.861, 639.973]$ ，本次实验共观测了 12 组位姿运动。

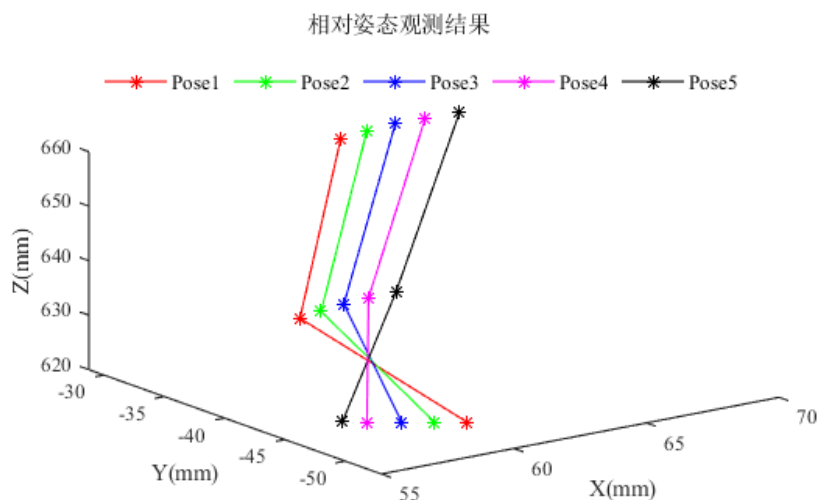


图 5-22 三角形标识重构示意图

进而，通过前文所叙述的姿态测量方法得到了三角形标识特征点的重构值，并将观测结果转换到机器人基坐标系中，图 5-22 给出了机器人基坐标系中的视觉系统的位姿观测结果。分析测量结果，通过对比观测值与机器人实际位姿运动参考值，得到了如表 5-11 所示的 12 组机器人姿态的相对运动数据和测量结果、图 5-23 所报道的机器人末端姿态测量误差。

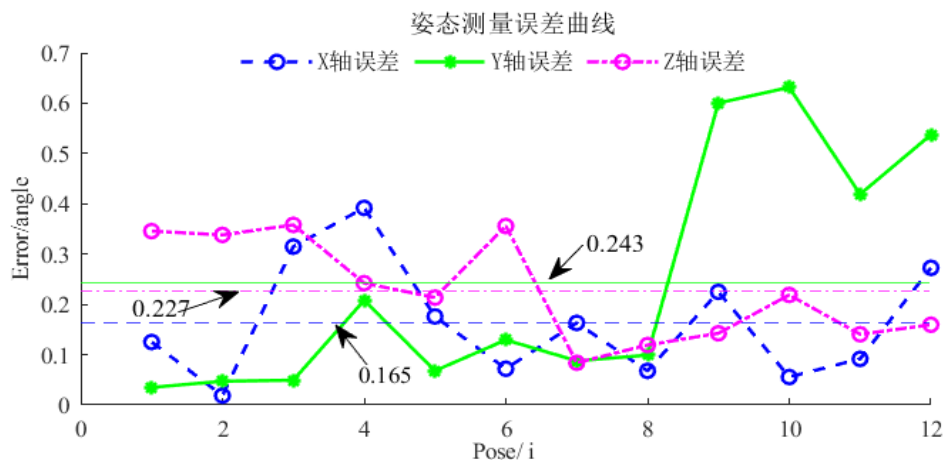


图 5-23 姿态测量误差示意图

表 5-11 姿态测量结果

序号	运动相对姿态			观测姿态		
	$\Delta R_x / ^\circ$	$\Delta R_y / ^\circ$	$\Delta R_z / ^\circ$	$\Delta R_x / ^\circ$	$\Delta R_y / ^\circ$	$\Delta R_z / ^\circ$
1	-5	0	0	-4.875	-0.035	-0.345
2	-5	0	0	-5.018	0.047	-0.338
3	-5	0	0	-5.315	0.049	-0.358
4	-5	0	0	-4.608	0.208	-0.242
5	2	4	1	2.176	3.932	1.213
6	-1.5	-2	-1	-1.570	-2.129	-0.644
7	0.5	1.5	0.5	0.663	1.412	0.416
8	1	1	0.5	0.932	1.099	0.381
9	0	-4	-1	0.224	-4.600	-1.142
10	0	-4	-1	-0.055	-4.631	-1.21
11	0	-4	-1	0.091	-4.419	-1.140
12	0	-4	-1	0.272	-4.537	-1.159

姿态测量实验中的 X 轴转角最大测量误差为 0.392 度，平均误差为 0.165 度；Y 轴转角的最大测量误差为 0.631 度，平均误差为 0.243 度；Z 轴转角的最大测量误差为 0.354 度，平均误差为 0.227 度。其中，Y 轴转角误差相对突出的原因在于，机器人末端以 Y 轴为旋转轴翻转时受光面积变小，标识清晰度下降，这使得标识特征点定位的不确定度增大，从而导致了部分数据误差增大。但整体而言，视觉系统的姿态测量误差均限制在 0.7 度以内，与同类位姿测量研究^[54,55]所得的机器人位姿测量精度趋于同误差量级，体现了姿态测量方法的合理性和可行性。

5.4 本章小结

为验证本文所提理论的可行性及相关算法的有效性，通过自建的视觉测量系统，以直角坐标机器人和六自由度机器人为测量对象展开了实验。进行了包括相机标定、基于 YOLOv5 的机器人运动追踪、机器人平面运动测量、机器人空间运动位姿测量的实验。实验结果表明，YOLOv5s 算法能够有效追踪机器人全局结构运动，所提出的特征点定位方法能够准确获取标识特征点的像素坐标；所提出的单应特征矩阵的平面测量模型能够有效消除传统正交投影模型测量的非正交误差；RANSAC 误差加权约束的坐标变换方法能够有效抑制测量误差，提高机器人 3D 运动测量的精度，在此基础上，实现了对机器人末端姿态的观测。各项实验结果验证了本文搭建的视觉系统与测量模型能够实现机器人全局结构运动观测。

第 6 章 总结与展望

6.1 总 结

工业机器人广泛应用于汽车制造业、电子电气、金属加工等领域。伴随着工业 4.0 发展,工业机器人作为一种高集成自动化的生产工具,能够有效保障产品质量,节约生产成本,提高生产效率。工业机器人技术已成为衡量国家工业水平的重要参考指标。

视觉测量技术得益于非接触、高交互性、高灵活度、全局观测优势已广泛应用于机器人领域的各个方面。采用视觉技术对机器人结构运动进行精确的跟踪测量,具有十分重要的工程应用价值。本文的主要工作如下:

(1) 给出了机器人结构运动的视觉测量方案,通过几何投影推导了相机成像模型和现有的相机参数标定方法;通过几何基元的运动形式与变换过程,阐述了机器人的运动表达形式。

(2) 为简化机器人运动过程,通过 YOLOv5s 算法对粘贴在机器人结构面上的标识进行追踪;提出了基于 Zernike 矩和最小二乘拟合的特征点定位方法,准确提取标识特征点的像素坐标。

(3) 深入研究了机器人运动视觉测量的数学模型,包括平面运动测量和 3D 运动测量理论与方法,相机成像模型和双目视觉模型,为提高工业机器人平面运动测量精度,提出了基于单应特征矩阵变换的视觉测量方法,为提高工业机器人空间运动测量精度,提出了基于误差加权的 RANSAC 三维坐标变换优化方法,基于所设计的三角形标识,给出了基于三角形特征标识的运动坐标系构建方法和姿态解析表达。

(4) 根据测量任务,完成了视觉系统的硬件选型和平台搭建,以六自由度机器人为测量对象,完成了 YOLOv5 算法的部署,实现了机器人平面运动视觉测量,机器人空间运动末端位姿视觉测量实验,实验结果表明:视觉系统能够实现机器人运动测量,位置误差限制在 0.4mm 以下,姿态误差限制在 0.7 度内,并达到了亚像素级别的检测精度,验证了所提测量方案的可行性,所提优化方法的有效性。

6.2 展望

获取精确的机器人结构运动测量数据，对实现机器人运动学标定、视觉伺服；优化机器人本体结构，探究机器人刚柔耦合特性具有重要意义。本文着眼于视觉测量技术，展开了机器人全局结构运动的视觉跟踪与测量研究。文章从几何建模出发，详细阐述了工业机器人运动的视觉测量方法与流程，提出了一系列有关机器人运动的视觉测量算法并进行了实验验证。实验结果验证了所提方法的可行性，所设计的视觉测量系统与算法能够完成机器人全局结构运动的测量需求。

然而，受限于时间与个人能力的不足，本文的研究内容仍有优化空间。改进工作可以参考以下两点推进：

(1) 视觉系统的标定精度极大影响后续运动测量的精度。虽然目前的标定算法趋于成熟，但就实际测量的适用性而言，仍有可优化角度。工业机器人的大空间运动要求视觉测量系统要具备大视场、较高精度的测量标定要求。就目前普遍采用的平面标定模型而言，使用大尺寸的标定板势必会受加工精度的干扰，导致标定精度下降，而小尺寸标定板的标定结果存在空间置信度低，与测量空间不匹配的问题。优化工作可以讨论组合标定的融合算法，实现多个平面标定物融合标定，以适应大视场测量需求。

(2) 提出的 RANSAC 误差加权坐标变换优化方法具有广泛的应用潜力，因为坐标变换过程适用于相机标定、3D 点云配准以及手眼系统标定等多个领域。然而，该算法的迭代循环过程具有较高的时间复杂度。尽管在机器人运动测量所需的坐标变换过程中，并未观察到时间复杂度对实验造成不利影响，但对于一些稠密点集的匹配场景，RANSAC 的循环过程可能会极大增加算法的时间成本。改进工作可以集中在降低该算法的时间复杂度上，以充分发挥该算法广泛的应用潜力。

参考文献

- [1] 王田苗, 陶永. 我国工业机器人技术现状与产业化发展战略[J]. 机械工程学报, 2014, 50(09): 1-13.
- [2] 吴锦辉, 陶友瑞. 工业机器人定位精度可靠性研究现状综述[J]. 中国机械工程, 2020, 31(18): 2180-2188.
- [3] 廖文和, 田威, 李波等. 机器人精度补偿技术与应用进展[J]. 航空学报, 2022, 43(05): 9-30+2.
- [4] 冯毅雄, 李康杰, 高一聪等. 面向视觉伺服的工业机器人轮廓曲线角点识别[J]. 浙江大学学报(工学版), 2020, 54(08): 1449-1456.
- [5] 赵丁选, 韩成浩. 遥操作机器人位置反馈型 H_{∞} 双向伺服控制系统研究[J]. 中国工程机械学报, 2010, 8(02): 147-151.
- [6] 吴震宇, 袁惠群, 罗宝佳等. 关节型工业机器人扭转振动的伺服补偿技术研究[J]. 振动与冲击, 2020, 39(09): 132-137.
- [7] 黄宣睿, 宋宇洋, 李秋生等. 一种基于内模控制的工业机器人关节伺服系统振动抑制算法[J]. 电工技术学报, 2019, 34(03): 497-505.
- [8] 吴炳龙, 曲道奎, 徐方. 基于力/位混合控制的工业机器人精密轴孔装配[J]. 浙江大学学报(工学版), 2018, 52(01): 165.
- [9] Denavit J, Hartenberg R S. A kinematic notation for lower-pair mechanisms based on matrices[J]. 1955, 22(2): 215-221.
- [10] Nubiola A, Bonev I A. Absolute calibration of an ABB IRB 1600 robot using a laser tracker[J]. Robotics and Computer-Integrated Manufacturing, 2013, 29(1): 236-245.
- [11] Zhuang H, Roth Z S, Hamano F. A complete and parametrically continuous kinematic model for robot manipulators[J]. IEEE Transaction on Robotics and Automation, 1992, 8(4): 451-463.
- [12] 刘运毅, 黎相成, 黄约等. 基于极大似然估计的工业机器人腕 6 维力传感器在线标定[J]. 机器人, 2019, 41(02): 216-221+231.
- [13] 张春涛, 王勇. 工业机器人六维力传感器在线标定方法研究[J]. 电子测量与仪器学报, 2021, 35(06): 161-168.
- [14] Klimchik A, Pashkevich A. Robotic manipulators with double encoders: accuracy improvement based on advanced stiffness modeling and intelligent control[J]. IFAC-PapersOnLine, 2018, 51(11): 740-745.
- [15] 余龙焕, 邱志成, 张宪民. 基于加速度反馈的平面 3-RRR 柔性并联机器人自激振动控制[J]. 机械工程学报, 2019, 55(21): 40-50.
- [16] Messay T, Ordóñez R, Marcil E. Computationally efficient and robust kinematic calibration

- methodologies and their application to industrial robots[J]. *Robotics and Computer-Integrated Manufacturing*, 2016, 37: 33-48.
- [17] Seo M, Yoo S, Choi M, et al. Vibration reduction of flexible rope-driven mobile robot for safe facade operation[J]. *IEEE/ASME Transactions on Mechatronics*, 2021, 26(4): 1812-1819.
- [18] Yang P, Guo Z, Kong Y. Plane kinematic calibration method for industrial robot based on dynamic measurement of double ball bar[J]. *Precision Engineering*, 2020, 62: 265-272.
- [19] Joubair A, Bonev I A. Kinematic calibration of a six-axis serial robot using distance and sphere constraints[J]. *The International Journal of Advanced Manufacturing Technology*, 2015, 77(1): 515-523.
- [20] 史晓佳, 张福民, 曲兴华等. KUKA 工业机器人位姿测量与在线误差补偿[J]. *机械工程学报*, 2017, 53(08): 1-7.
- [21] Zhao G, Zhang P F, Guocai Ma, et al. System identification of the nonlinear residual errors of an industrial robot using massive measurements[J]. *Robotics and Computer-Integrated Manufacturing*, 2019, 59: 104-114.
- [22] Toni Cvitanic, Shreyes Melkote, Stephen Balakirsky. Improved state estimation of a robot end-effector using laser tracker and inertial sensor fusion[J]. *CIRP Journal of Manufacturing Science and Technology*, 2022, 38: 51-61.
- [23] Gao Q, Zhang P C. The Errors Recognition and Compensation for the Numerical Control Machine Tools Based on Laser Testing Technology[J]. *Open Physics*, 2019, 17(1): 857-862.
- [24] Qiao Y, Chen Y, Chen B, et al. A novel calibration method for multirobots system utilizing calibration model without nominal kinematic parameters[J]. *Precision Engineering*, 2017, 50: 211-221.
- [25] 朱贺, 杨华, 尹周平. 纹理表面缺陷机器视觉检测方法综述[J]. *机械科学与技术*, 2023, 42(08): 1293-1315.
- [26] 孙红, 孙明, 王一鸣. 植物生长机器视觉无损测量研究综述[J]. *农业机械学报*, 2006(10): 181-185.
- [27] 尹仕斌, 任永杰, 刘涛等. 机器视觉技术在现代汽车制造中的应用综述[J]. *光学学报*, 2018, 38(08): 11-22.
- [28] 严宏志, 邵冬世, 韩奉林等. 基于双目视觉测量系统的机器人误差检测方法[J]. *现代制造工程*, 2020(09): 81-87+94.
- [29] 黄磊, 余峰. 基于视觉反馈的机器人末端轨迹跟踪控制研究[J]. *仪表技术与传感器*, 2021(08): 95-98.
- [30] 徐秀秀, 郭毓, 余臻, 张懿. 基于机器视觉的柔性臂振动测量研究[J]. *华中科技大学学报(自然科学版)*, 2013, 41(S1): 129-132.

- [31] 郝继贵, 张楠楠, 任永杰等. 基于双目立体视觉的工业机器人在线温度补偿[J]. 光学精密工程, 2018, 26(09): 2139-2149.
- [32] 尚何江. 基于双目视觉的 6R 串联工业机器人结构参数标定方法研究[D]. 哈尔滨工业大学, 2022.
- [33] Singla N. Motion detection based on frame difference method[J]. International Journal of Information & Computation Technology, 2014, 4(15): 1559-1565.
- [34] Szolgay D, Benois-Pineau J, Mégret R, et al. Detection of moving foreground objects in videos with strong camera motion[J]. Pattern Analysis and Applications, 2011, 14: 311-328.
- [35] Barnich O, Van Droogenbroeck M. ViBe: A universal background subtraction algorithm for video sequences[J]. IEEE Transactions on Image processing, 2010, 20(6): 1709-1724.
- [36] 周颖, 张立迅, 刘彤等. 基于计算机视觉的结构系统识别[J]. 土木工程学报, 2018, 51(11): 17-23.
- [37] Sun D, Roth S, Black M J. A quantitative analysis of current practices in optical flow estimation and the principles behind them[J]. International Journal of Computer Vision, 2014, 106: 115-137.
- [38] 王世勇, 乾国康, 李迪等. 面向边缘特征的实时模板匹配方法[J]. 华南理工大学学报(自然科学版), 2023, 51(09):1-10.
- [39] 张淑芳, 李亚阳, 张涛. 基于快速仿射模板匹配的 AMCL 算法[J]. 北京航空航天大学学报, 2023, 49(11): 2898-2905.
- [40] 王瑞洲, 梁炽聪, 蔡远馨等. 离面位移激励下微视觉运动追踪精度劣化特性[J]. 光学精密工程, 2023, 31(01): 119-128.
- [41] Comaniciu D, Meer P. Mean shift: a robust approach toward feature space analysis[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(5): 603-619.
- [42] Kim K I, Jung K, Jin H K. Texture-based approach for text detection in images using support vector machines and continuously adaptive mean shift algorithm[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(12): 1631-1639.
- [43] Ma D, Joo E M, Beng L H. A comprehensive study of Kalman filter and extended Kalman filter for target tracking in Wireless Sensor Networks[C]// IEEE International Conference on Systems, Man and Cybernetics, Singapore, 2008, pp. 2792-2797.
- [44] Shen X, Varshney P K. Sensor Selection Based on Generalized Information Gain for Target Tracking in Large Sensor Networks[J]. IEEE Transactions on Signal Processing, 2014, 62(2).
- [45] Rad M, Lepetit V. BB8: a scalable, accurate, robust to partial occlusion method for predicting the 3D poses of challenging objects without using depth [C] // Proceedings of the 2017 IEEE International Conference on Computer Vision. Piscataway: IEEE, 2017: 3848-3856.
- [46] Tekin B, Sinha S N, Fua P. Real-time seamless single shot 6D object pose prediction [C] //

- Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 292-301.
- [47] Chen B, Parra Á, Cao J W, et al. End-to-end learnable geometric vision by backpropagating PnP optimization [C] // Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2020: 80978106.
- [48] Kehl W, Manhardt F, Tombari F, et al. SSD-6D: making RGB-based 3D detection and 6D pose estimation great again[C]// Proceedings of the 2017 IEEE International Conference on Computer Vision. Piscataway: IEEE, 2017: 1530-1538.
- [49] 张键驰. 基于 RGB 图像和三维模型的物体姿态估计[D]. 华南理工大学, 2021.
- [50] 潘天峰. 基于改进 Faster R-CNN 的目标检测和姿态估计[D]. 华中科技大学, 2021.
- [51] 张耕慎. 基于深度学习的三维计算机视觉下的 6D 姿态估计[D]. 中国科学院大学(中国科学院深圳先进技术研究院), 2021.
- [52] Zhang Z Y. A flexible new technique for camera calibration[J]. IEEE Transactions on pattern analysis and machine intelligence, 2000, 22(11): 1330-1334.
- [53] Martin A. F, Robert C. B. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography[J]. Communications of the ACM, 1981, 24(6): 381-395.
- [54] 崔家山. 三维运动立体视觉测量方法研究[D]. 哈尔滨工业大学, 2016.
- [55] 张小迪. 面向机器人末端位姿的视觉测量跟踪技术研究[D]. 南京航空航天大学, 2020.

攻读硕士期间的主要成果

- [1] Lulu Wu, Xianglin Deng, Yuan Wang, Xin Du, Xinyan Xiong , Benchu Jiang. Research on vision-based robot planar motion measurement method[J]. Journal of the Brazilian Society of Mechanical Sciences and Engineering, 2023, 45(4): 216. (SCI, Q4)
- [2] Wu Lulu, Deng Xianglin, Wang Yuan, Du Xin, Xiong Xinyan, Jiang Benchu. Robot motion visual measurement based on RANSAC and weighted constraint method[J]. Measurement Science and Technology, 2023, 35(3):035012. (SCI, Q3)
- [3] 吴路路, 邓湘琳, 王园等. 一种基于视觉的机械臂全局性能测量系统及方法[P]. CN115294198A, 2022-11-04. (发明实审)

致 谢

踏入安徽工程大学校园那一刻，至今已近七年，时光荏苒如长江东逝，可谓弹指一挥间。在求学过程中，母校各位老师教学热情，平易近人，学院各位领导倾听关怀学生的生活和学习日常，这给予了我极大的帮助。

在硕士学习期间，导师吴路路治学严谨，辅导学生耐心细致，总能设身处地为学生着想，在我心目中，是一位十分称职的、优秀的青年导师。

无奈，所琢非玉，难成大器。学生胸无大志，朽木难雕，外不修行，内不修德，常怀小人之慧，不念忠恕之道；于家难尽天伦之孝，于国无立寸土之功，为此常深感愧疚，寝食难安。十九年的学生生涯临终之际，万般思绪涌上心头，难以言诉。

感谢父母的二十多年的养育之恩。

感谢党和国家对底层劳动人民的支持，对学生群体的呵护。

感谢十九年求学期间每一位老师。

感谢为人二十四年期间每一位亲人，朋友。

惟愿父母身体康健，社稷昌盛，江山太平。

尚德敏学 唯实惟新

