

中图分类号: TP242.3

单位代号: 10280

密 级: 公开

学 号: 23721949

# 上海大学



# 硕士学位论文

SHANGHAI UNIVERSITY  
MASTER'S DISSERTATION

题 目

基于信道状态信息与线阵视觉的  
协同室内定位关键技术研究

作 者 袁志鹏

学科专业 控制理论与控制工程

导 师 杨庆华

完成日期 二〇二六年五月

姓    名：袁志鹏

学号：23721949

论文题目：基于信道状态信息与线阵视觉的协同室内定位关键技术研究

# 上海大学

本论文经答辩委员会全体委员审查，确认符合上海大学硕士学位论文质量要求。

答辩委员会签名：

主    席：

委    员：

导    师：

答辩日期：        年    月    日

姓 名：袁志鹏

学号：23721949

论文题目：基于信道状态信息与线阵视觉的协同室内定位关键技术研究

## 上海大学学位论文原创性声明

本人郑重声明：所呈交的学位论文是本人在导师指导下，独立进行研究工作所取得的成果。除了文中特别加以标注和致谢的内容外，论文中不包含其他人已发表或撰写过的研究成果。参与同一工作的其他研究者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了谢意。

学位论文作者签名：

日期： 年 月 日

## 上海大学学位论文使用授权说明

本人完全了解上海大学有关保留、使用学位论文的规定，即：学校有权保留论文及送交论文复印件，允许论文被查阅和借阅；学校可以公布论文的全部或部分内容。

(保密的论文在解密后应遵守此规定)

学位论文作者签名：

导师签名：

日期： 年 月 日

日期： 年 月 日

## 上海大学工学硕士学位论文

# 基于信道状态信息与线阵视觉的协同 室内定位关键技术研究

作 者: 袁志鹏

导 师: 杨庆华

学科专业: 控制理论与控制工程

机电工程与自动化学院

上海大学

2026 年 5 月

A Dissertation Submitted to Shanghai University for the Degree  
of Master in Engineering

# **Research on Key Technologies of Collaborative Indoor Positioning Based on Linear Array Vision and Channel State Information**

Candidate: Zhipeng Yuan

Supervisor: Qinghua Yang

Major: Control Theory and Control Engineering

**School of Mechanical Engineering and Automation Shanghai  
University  
Shanghai University  
May, 2026**

## 摘 要

XX  
XX  
XX

**关键词：**T<sub>E</sub>X; L<sup>A</sup>T<sub>E</sub>X; 模板; 论文

## ABSTRACT

XX  
XX  
XX  
XX

**Keywords:** T<sub>E</sub>X; L<sup>A</sup>T<sub>E</sub>X; Template; Thesis

# 目 录

摘要 .....	I
ABSTRACT .....	II
<b>第一章 绪论 .....</b>	<b>1</b>
1.1 研究背景与意义 .....	1
1.2 国内外研究现状 .....	2
1.2.1 室内定位技术国内外研究现状 .....	2
1.2.2 基于信道状态信息的室内定位技术国内外研究现状 .....	7
1.2.3 基于线阵 CCD 的室内定位技术国内外研究现状 .....	8
1.3 论文的研究内容和创新点 .....	10
1.3.1 论文的主要研究内容 .....	10
1.3.2 论文的创新点 .....	12
<b>第二章 Wi-Fi 信道状态信息定位技术相关理论 .....</b>	<b>14</b>
2.1 Wi-Fi 信号简介 .....	14
2.1.1 无线局域网 (WLAN) 标准概述 .....	14
2.1.2 Wi-Fi 信号的传播特征 .....	14
2.2 Wi-Fi 无线定位参数 .....	16
2.2.1 接收信号强度指示 (RSSI) .....	16
2.2.2 信道状态信息 (CSI) .....	17
2.2.3 RSSI 与 CSI 参数性能对比 .....	18
2.3 CSI 定位相关技术基础 .....	20
2.3.1 正交频分复用技术 (OFDM) .....	20
2.3.2 多输入多输出技术 (MIMO) .....	21
2.3.3 OFDM 与 MIMO 的协同作用 .....	22
<b>第三章 基于信道状态信息的室内定位算法与系统设计 .....</b>	<b>23</b>
3.1 引言 .....	23
3.2 基于嵌入式设备的分布式单天线节点分析 .....	23

3.2.1 空间位置对 CSI 幅值的响应特性 .....	23
3.2.2 多节点空间分布特征分析 .....	25
3.3 幅值与相位特征的稳定性与可用性分析 .....	29
3.3.1 幅值与相位稳定性分析 .....	29
3.3.2 特征时域稳定性与可用性的综合评估 .....	32
3.4 CSI 室内系统架构设计总览 .....	34
3.4.1 数据预处理与张量构造 .....	35
3.4.2 多尺度时序特征提取 (MSTC) .....	38
3.4.3 基于 ConvNeXt 的主干网络设计 .....	39
3.4.4 时频域注意力机制与特征校准 .....	40
3.4.5 注意力机制的协同作用分析 .....	41
3.5 基于物理约束的损失函数与输出平滑 .....	42
3.5.1 运动学约束损失函数 (Kinematic-Constrained Loss) .....	42
3.5.2 加权移动平均 (WMA) 后处理 .....	43
3.6 网络复杂度与理论性能分析 .....	44
3.6.1 参数量与计算复杂度推导 .....	44
3.6.2 实时性与部署可行性探讨 .....	45
3.7 本章小结 .....	46
<b>第四章 线阵相机三维定位的成像模型与几何基础 .....</b>	<b>47</b>
4.1 引言 .....	47
4.2 线阵相机成像原理 .....	48
4.2.1 柱面镜头组成像模型 .....	48
4.2.2 光斑中心定位算法 .....	49
4.3 坐标系的建立 .....	50
4.3.1 像素坐标系 (Pixel Coordinate System) .....	51
4.3.2 图像坐标系 (Image Coordinate System) .....	51
4.3.3 线阵相机坐标系 (Camera Coordinate System) .....	52
4.3.4 传感器坐标系 (Sensor Coordinate System) .....	53
4.3.5 世界坐标系 (World Coordinate System) .....	53
4.4 基于多线阵 CCD 的空间点定位原理 .....	53

4.4.1 线阵相机投影模型与参数约束 .....	54
4.4.2 多线阵空间交会定位原理 .....	55
4.5 本章小结 .....	57
<b>第五章 线阵视觉系统的几何畸变校正与高鲁棒性标定方法实现</b> .....	<b>58</b>
5.1 引言 .....	58
5.1.1 现有技术现状与局限 .....	58
5.1.2 分阶段标定策略 .....	58
5.2 柱面镜头的畸变与去畸变 .....	59
5.2.1 畸变模型 .....	59
5.2.2 去畸变模型 .....	63
5.3 考虑几何灵敏度的线阵相机参数加权估计方法 .....	63
5.3.1 问题背景与方法提出 .....	63
5.3.2 加权目标函数的严格推导 .....	64
5.3.3 非线性优化与实施流程 .....	65
5.4 相机内参计算方法 .....	67
5.5 相机坐标系和世界坐标系之间的刚体变换 .....	68
5.6 仿真实验 .....	69
5.7 小结 .....	69
<b>第六章 协同系统结构和实验结果分析</b> .....	<b>71</b>
6.1 引言 (Introduction) .....	71
6.2 协同定位系统的实现架构 .....	71
6.2.1 总体逻辑架构设计 .....	71
6.2.2 光学监测节点: YOPS 硬件系统 .....	73
6.2.3 分布式射频感知网络: CSI 采集节点 .....	74
6.3 实验环境部署与评估方案 .....	75
6.3.1 被测目标与实验环境配置 .....	75
6.3.2 系统标定与真值坐标系构建 .....	75
6.3.3 多维评估场景设计 .....	76
6.4 实验结果分析与性能评估 .....	78

6.4.1 光学传感器性能验证 .....	78
6.4.2 协同定位系统综合性能评估 .....	78
6.5 本章小结 .....	78
参考文献 .....	79
攻读硕士学位期间取得的研究成果 .....	80
致 谢 .....	81
附录 A 经典不等式 .....	82

# 第一章 绪论

## 1.1 研究背景与意义

党的二十大报告明确提出，要加快建设网络强国、数字中国，加快发展数字经济，推动物联网、大数据、人工智能等新一代信息技术与实体经济深度融合<sup>[1]</sup>。作为数字经济的重要‘时空底座’，位置服务（Location Based Services, LBS）已成为支撑智慧城市建设、工业互联网发展以及社会精细化治理的关键核心技术之一 [2]。特别是在国家大力推进‘新基建’的背景下，实现全场景、全天候的高精度位置感知，对于打通物理世界与数字世界的边界具有重大的战略意义。目前，我国自主建设的北斗卫星导航系统已在室外开放空间取得了举世瞩目成就。然而，在卫星信号难以覆盖的室内复杂空间（如大型交通枢纽、地下管廊、智能工厂等），仍存在巨大的‘定位盲区’ [3]。随着移动机器人的普及与人机协作需求的激增，室内定位技术已不再局限于简单的导航指引，而是上升为保障生产安全、提升物流效率及实现特定区域管控的核心环节。面对这一现状，如何突破室内复杂环境下的高精度定位瓶颈，构建自主可控且具备高鲁棒性的室内时空服务体系，已成为填补国家导航定位体系‘最后一公里’空白的紧迫需求 [4]

随着室内位置服务（LBS）应用场景的不断拓展，学术界与工业界对定位系统的精度、鲁棒性及隐私安全性提出了更高的要求。在现有的主流技术中，基于面阵成像的视觉定位算法凭借其丰富的纹理特征提取能力，实现了较高的定位精度。然而，该类方法在实际部署中面临着突出的物理与伦理矛盾：一方面，在家庭、办公及保密单位等敏感区域，持续的面阵图像采集引发了用户对隐私泄露的严重顾虑，限制了其广泛应用；另一方面，成像视觉技术对环境光照高度敏感，在弱光、动态干扰或非视距（NLOS）遮挡条件下，其特征提取的稳定性显著下降，极易导致定位失效。尽管非成像式的线阵光电技术（如线阵 CCD）因其高采样率和天然的隐私保护特性提供了新的硬件思路，但目前针对其在复杂空间定位中的研究仍相对匮乏。与此同时，基于 Wi-Fi 的无线定位技术因无需额外布线而备受关注。虽然基于信道状态信息（CSI）的方法克服了传统接收信号强度（RSSI）粒度粗、易波动的缺陷，能够利用细粒度的频域特征有效刻画多径效应，但其高性能主要依赖于数据驱动的深度学习模

型。这一特性带来了新的挑战：CSI 指纹库的构建需要海量且带有精确位置标签的训练数据，而传统的人工定点采集方式不仅效率低下，且难以保证毫米级的标注精度。这种“数据饥渴”与高昂的人力成本，成为制约 CSI 高精度定位技术落地应用的关键瓶颈。综上所述，单一的视觉模态受限于隐私风险与环境适应性，而单一的无线模态则受困于数据标注难题与信号漂移。面对复杂多变的室内环境，单纯依赖物理层面的单一传感器难以突破现有的性能天花板。因此，探索视觉（光）与射频（电）的异构协同机制，对于构建复杂场景下的全天候鲁棒定位具有重要的研究价值与应用前景。

## 1.2 国内外研究现状

### 1.2.1 室内定位技术国内外研究现状

#### (1) 无线网络室内定位技术

随着物联网与移动计算的发展，室内定位技术在公共安全、资产追踪及智能环境中的应用需求激增。由于室内环境存在复杂的非视距（NLOS）传播和多径效应，单一技术往往难以满足所有场景需求。目前，研究主要集中在挖掘现有无线网络（如 WiFi、蓝牙）的潜力，利用新兴技术（如 UWB、毫米波）实现高精度定位，以及通过多源融合与人工智能算法提升系统的鲁棒性。

1. 基于 WiFi 的定位技术研究 WiFi 定位因其基础设施广泛普及而备受关注，研究热点正从传统的接收信号强度（RSS）向信道状态信息（CSI）和精细时间测量（FTM）转变，以克服信号波动带来的精度限制。在指纹定位方面，深度学习被广泛引入以提取更深层的特征。Qin 等人 [288] 提出了一种名为 CCpos 的 WiFi 指纹室内定位系统，该系统基于卷积去噪自编码器和卷积神经网络（CDAE-CNN）模型，有效提取了指纹特征，提升了定位性能。

为了解决视距与非视距传播的干扰，Si 等人 [180] 提出了一种基于 WiFi FTM (Fine Time Measurement) 的室内定位方法，该方法集成了 LOS/NLOS 识别机制，在复杂环境中提升了测距精度。此外，随着 WiFi 标准的演进，Storrer 等人 [178] 指出 IEEE 802.11ax 标准可提供高达 80 MHz 的带宽，对应的距离分辨率约为 1.88 米，这种高带宽特性为基于多天线被动雷达的室内人员追踪提供了硬件基础。

2. 基于蓝牙（BLE）的定位技术研究低功耗蓝牙（BLE）技术因其低成本和低功

耗特性在室内定位中占据重要地位。近年来的研究主要围绕蓝牙 5.1 标准引入的到达角 (AoA) 特性, 以及与惯导系统的深度融合展开。针对 AoA 定位, Zhao 和 Yang[77] 实现了一种基于蓝牙 5.1 标准的 AoA 室内定位系统, 利用多天线阵列产生的相位差计算到达角, 在仓库资产定位场景中以低成本实现了 \*\* 亚米级 (sub-meter) \*\* 的定位精度。为了进一步提升角度估计的准确性, He 等人 [188] 提出了一种基于多天线阵列的 AoA 估计方法, 实验表明, 相比于传统的多信号分类 (MUSIC) 算法, 该方法的平均角度误差小于  $3.9^\circ$ 。

在算法融合方面, Kong 等人 [111] 提出了一种自适应反馈扩展卡尔曼滤波 (AFEKF) 算法, 用于融合 BLE 和行人航位推算 (PDR)。该机制将距离测量结果深度反馈到下一时刻的位置估计中, 实验结果显示, 与经典 EKF 算法相比, 该算法将定位精度提高了 23.4%。此外, Echizenna 和 Kondo[187] 利用深度神经网络 (DNN) 处理 BLE 信标信号, 实现了同时检测行人位置和运动方向, 实验测得定位精度为 0.439 m, 平均方向识别准确率达到 81.2%。

3. 基于超宽带 (UWB) 与毫米波的高精度定位研究对于对精度要求极高的工业和无人机场景, UWB 和毫米波 (mmWave) 技术因其极高的时间分辨率和空间扫描能力成为研究热点。

在 UWB 应用中, Queralta 等人 [52] 针对 GNSS 拒止环境下的无人机定位, 提出了一种基于 UWB 的定位算法。该系统不仅能耗低, 而且能够将定位误差控制在 0 到 4 cm 之间, 且超过 50% 的情况下总误差小于 3 cm。为了解决复杂环境下的非高斯噪声问题, Zhou 等人 [113] 提出了一种自适应最大相关熵无迹卡尔曼滤波 (AMCUKF) 算法融合 IMU 和 UWB 数据, 有效提升了系统的鲁棒性。

在毫米波定位方面, Jia 等人 [235] 提出了一种改进的最小均方算法来优化 AoA 估计, 并结合修正的多径 AoA-TOA 无迹卡尔曼滤波 (UKF) 算法。实验证明, 该方法仅利用单个接入点 (AP) 即可在办公室环境中实现厘米级的定位精度, 并获得了 2 倍的角度估计增益。

4. 基于多源融合与新兴技术的创新研究为了应对单一技术的局限性, 多源信息融合 (如 PDR、地磁、视觉等) 以及利用智能反射面 (RIS) 等新技术重塑信道环境, 成为当前的重要研究趋势。在多源融合滤波算法方面, Silva 等人 [119] 提出了 TrackInFactory 方案, 利用紧耦合粒子滤波融合 INS 和 WiFi 信息, 并通过一种新的可靠性度量动态更新粒子权重, 在工业车辆追踪中实现了 0.81 m 的平均误差。Chen

等人 [120] 则提出了一种带有信息共享机制的联邦粒子滤波 (FPF) 算法融合 PDR 和 WiFi，通过主滤波器和子滤波器的协同，将定位误差控制在 1 m 左右。

此外，针对太赫兹 (THz) 频段的未来应用，Fan 等人 [243] 探索了深度学习在射频定位中的应用，提出了一种结构化的双向长短期记忆 (Bi-LSTM) 循环神经网络架构，实现了平均距离误差为 0.27 m 的 3D 室内定位。而在智能反射面 (RIS) 辅助定位方面，Zhang 等人 [142] 的研究表明，通过优化 RIS 的反射波束赋形，可以将定位精度提升至分米级甚至厘米级，有效解决了信号遮挡问题。

## (2) 光学室内定位技术

基于光学的定位系统在动作捕捉、虚拟现实 (Virtual Reality, VR)、增强现实 (Augmented Reality, AR)、机器人控制等领域有着广泛的应用，具有精度高、实时性强、延时低等优点。线阵相机、面阵相机、激光雷达等均是常见的光学传感器。根据定位方式的不同，可将光学室内定位技术分为非自主定位和自主定位。

非自主光学定位技术是指在被追踪目标上附着特定的标记点 (Fiducials)，利用外部摄像机网络捕捉标记点的图像坐标，进而通过立体视觉原理重构目标六自由度 (6DoF) 位姿的技术。根据标记点成像机理的不同，目前的主流技术路线主要分为被动反光式 (Passive) 与主动发光式 (Active) 两大方向。被动式技术通常采用涂有逆反射材料的球体或平面圆盘作为标记点，通过反射追踪相机发出的红外光进行成像。英国 Oxford Metrics Limited 公司开发的 VICON[29] 是世界上第一款用于动画制作的光学动作捕捉系统，由一组高速红外相机、多个被动反光标记点和相应的软件组成。红外相机被固定在房间四周，每台相机均朝向房间中心，确保相机镜头视野能够覆盖所有需要定位的区域。被动反光标记点被固定需要定位在目标物体上，当多个相机同时观察到标记点时，该点的空间坐标即可被计算出来。

由于其具有无线缆束缚、重量轻且易于消毒等优势，长期以来一直是医疗导航领域的首选方案。文献 [1-12] 指出，商业化的光学追踪解决方案已广泛应用于微创手术中，其工具定位精度通常在 0.1 mm 量级。为了进一步提升精度，文献 [1-13] 展示了通过三目视觉 (Trinocular) 配置，可将被动系统的平移精度提升至 0.04 mm 左右 1。加拿大 NDI 公司开发的 Polaris[33] 系列产品采用双目红外视觉方案，多应用于机器人辅助手术领域，被定位物体同样需要安装反光标记点。这些定位系统均可达到亚毫米级定位精度，但是定位范围较小且成本较高。然而，被动式标记点在

复杂环境下的鲁棒性面临挑战。文献 [A] 的研究表明，被动标记点极易受到污染，导致反射率显著下降（如图 3 所示），从而严重干扰位姿追踪的稳定性 2。

中国度量科技开发的 Nokov[30]、美国 Natural Point 开发的 Optitrack[31] 以及瑞典的 Qualisys[32] 也属于非自主光学定位系统，其中 Qualisys 支持主动发光标记点，该功能可以给目标物体添加专属身份标签。

主动式技术通过在目标上布置红外发光二极管（IR LEDs）主动发射信号，具有对环境光照不敏感、抗污染能力强等特点。尽管主动式方案在鲁棒性上占优，但其引入的功耗与热效应不容忽视。文献 [1-21] 指出，为了维持足够的亮度，LED 往往需要较大的驱动电流，由此产生的热量不仅可能引起人体不适，甚至会影响标记点附近嵌入式半导体（如 FPGA 芯片）的性能稳定性 5。因此，如何在保证追踪精度的前提下优化驱动电流成为了研究重点。文献 [A] 的实验数据证明，在满足检测阈值的情况下，降低驱动电流并不会显著增加定位抖动，从而验证了低功耗驱动的可行性 6。

针对主动式与被动式系统在实际应用中的精度差异，文献 [B] 设计了一项包含两种测试环境（实验室与手术室）和两种工具长度的严格对比实验，评估了目标配准误差（TRE）。实验结果显示，主动式系统（Metronor OTS38/70）的整体精度（约 0.063 mm）显著优于被动式系统（Polaris Vega，约 0.259 mm），且在抗环境干扰方面表现出更低的标准差 7777。除了传感器本身的差异，标记点的几何分布也是决定定位精度的核心因素。

自主光学定位系统强调定位主体的自主性，即移动终端（如智能手机、机器人）通过搭载的光电二极管（PD）或图像传感器（Camera/IS）接收环境中的可见光信号，在本地解算自身位置，无需依赖外部基站的测量。根据终端传感器的不同，目前的研究主要分为基于光电二极管的定位、基于图像传感器的定位以及多传感器融合定位。

1. 基于光电二极管（PD）与传感器辅助的定位研究基于 PD 的系统通常利用接收信号强度（RSS）或到达角度（AOA）进行计算。为了克服单 PD 定位的局限性，研究者常引入惯性传感器辅助。

文献 [151] 提出了一种融合光强传感器与加速度计的定位方案，利用加速度计确定接收器方向以简化姿态计算。实验结果显示，在  $5 \times 5 \times 3$  m 的测试空间内，该方案的平均定位误差可控制在 25 cm 以内。针对办公走廊等典型场景，文献 [154] 利用智能手机内置的加速度计和磁力计感知自身姿态并获取法向量，实现了 \*\* 亚米级（Sub-meter）\*\* 的定位精度。此外，文献 [155] 结合 RSS 与角度传感器（Angle

Sensor), 在  $1.2 \times 5 \times 2.5$  m 的实验空间中, 将定位误差控制在 29.8~46.3 cm 之间。

2. 基于图像传感器 (Camera) 的视觉定位研究随着智能终端摄像头的普及, 基于图像传感器的 VLP 利用几何特征或卷帘快门效应实现高精度自主定位。

文献 [168] 将 VLP 应用于机器人操作系统 (ROS), 通过工业相机节点捕获 LED 图像并由定位器节点解算, 实现了基于双灯定位原理的机器人自主导航。实验表明, 该系统能提供 2 cm 以内的定位精度, 且单次定位处理时间仅需 0.35 秒。为了进一步降低误差, 文献 [169] 在移动机器人上应用了改进的校准程序和 SLAM 算法, 在  $1 \times 1 \times 1.5$  m 的实验范围内, 将定位误差降低至 0.82 cm。

针对 QR 二维码辅助定位, 文献 [171] 提出了一种结合 VLP 与二维码的方案, 接收器通过解码加载在 LED 灯上的二维码图像进行初始定位, 在  $2 \times 2 \times 3$  m 的空间内实现了 4.0326 cm 的平均误差。

在算法对比方面, 文献 [121] 提出的相机辅助 RSSR 算法 (eCA-RSSR) 与文献 [16] 提出的透视圆弧算法 (V-PCA) 进行了详细的数据对比。仿真数据显示: 在接收器方向未知的情况下, V-PCA 算法能以 93% 的概率将定位误差控制在 10 cm 左右, 优于 eCA-RSSR 算法 (78% 的概率达到同等精度)<sup>7</sup>。此外, 在覆盖率方面, 当视场角 (FoV) 在 25° 至 80° 之间时, V-PCA 的覆盖率始终超过 90%, 比 eCA-RSSR 高出约 20%<sup>8</sup>。

3. 基于视觉/光电与惯性导航 (INS/PDR) 的融合定位研究为了解决单一光学信号易受遮挡和视场限制的问题, 融合 IMU 或行人航位推算 (PDR) 成为主流趋势, 显著提升了系统的鲁棒性和精度。

文献 [136] 提出了一种基于扩展卡尔曼滤波 (EKF) 的视觉-惯性融合方法, 利用卷帘快门相机和 IMU 共同完成定位。在  $5 \times 4 \times 2.3$  m 的实验场景中, 该方法的均方根误差 (RMSE) 仅为 5 cm<sup>9999</sup>。文献 [137] 利用单个 LED 的几何特征结合 IMU 估算接收器方向, 在  $2.7 \times 1.8 \times 1.75$  m 的空间内实现了 5.44 cm 的定位精度 10101010。

针对行人导航, 文献 [160] 利用粒子滤波 (Particle Filter) 同时处理 VLP 和 PDR 数据, 在  $2.0 \times 2.0$  m 的测试区域内实现了 14 cm 的定位误差 1111111111111111。文献 [161] 则采用扩展卡尔曼滤波融合 RSS 和 PDR 数据, 在  $2.2 \times 2.2$  m 的区域内取得了 14.5 cm 的定位精度 12121212。此外, 文献 [140] 将智能手机 IMU 传感器与 PDR 及 VLP-AoA 算法结合, 实验结果表明该方法能将平均误差降低至 0.85 m, 实现了无缝移动室内导航 13131313。对于大场景应用, 文献 [144] 在博物馆场景中利用 5G 无

线电光定位系统，实现了 0.18 m 的平均定位误差 14。

值得注意的是，无论是主动还是被动系统，都严格依赖于视线（Line-of-Sight），一旦标记点被遮挡即会导致追踪中断。

### 1.2.2 基于信道状态信息的室内定位技术国内外研究现状

近年来，随着无线通信技术的发展，基于信道状态信息（Channel State Information, CSI）的室内定位技术因其比接收信号强度（RSSI）更细粒度的信道描述能力而备受关注。CSI 能够捕捉多径效应下的幅度与相位信息，研究者们通过引入深度学习、迁移学习以及针对 5G 信号特性的新方法，显著提升了定位的精度与鲁棒性。

特征工程与鲁棒性优化针对 CSI 信号的波动性，研究者通过精细化的特征处理来提升性能。[11]Che 等人提出了 FuFi 方法，利用 MIMO 系统中全维度子载波的归一化幅度作为指纹，并配合改进的加权 K 近邻（WKNN）算法，在会议室场景下实现了 0.65 m 的平均距离误差（MDE），相比传统 FIFS 算法精度提升了 59.9%。[12]Sun 等人结合 PCA 降维和 SVM 分类，并考虑了人体不同姿态（如站立、蹲下等）对 CSI 的影响，实验证明该策略使定位精度提升了 9.6

[15]Astafiev 等人探索了仅利用 CSI 相位信息进行设备定位的可行性。针对相位数据的非线性特征，作者构建了一个全连接神经网络，利用 56 个子载波和 3 对天线的相位矩阵作为输入。实验结果显示，在静态场景下，该算法的分类定位准确率最高可达 94.58%（使用 512 个神经元时）。然而，研究也指出，仅依赖相位信息在动态（移动）目标的定位上存在局限性，会产生较大的误差，因此建议未来结合幅度信息以提升动态场景下的鲁棒性。

深度学习模型的架构创新深度神经网络（DNN）及其变体在处理复杂的 CSI 特征映射方面表现出卓越性能。针对传统方法在复杂环境下的局限性，[1]Li 等人提出了一种基于 Transformer 的定位方法 LoT（Localization Transformer），通过趋势周期分解填充和矩形 Patching 机制解决了 CSI 矩阵的长条形输入问题，实验表明该方法在室内场景下的平均绝对误差（MAE）仅为 0.18 m，优于 ResNet 等基准模型。[2]Song 等人设计了双通道卷积神经网络系统 DuLoc，利用孤立森林去除异常值并通过 Mean Shift 聚类划分指纹库，分别处理稳定与不稳定子载波，其在复杂室内环境下的平均距离误差降低至 0.42 m，且 85

样本效率与半监督/迁移学习为了降低指纹库构建的人力成本并解决数据孤岛问

题，小样本学习和联邦学习成为研究热点。[5]He 等人提出了基于半监督阶梯网络的方法 LadderNetFi，该方法结合了去噪自动编码器（DAE），仅需 10% 的有标签数据即可达到与全监督学习相当的性能，在实验室场景下实现了 1.65 m 的 RMSE。[6]Guo 等人提出了联邦迁移学习框架 FedPos，通过聚合不同用户模型的非分类层参数构建云端特征提取器，实验显示该方法在保护隐私的同时，相比普通训练方法平均定位性能提升了 5.22%，且训练时间减少了约 34.78%。[7]Zhang 等人引入了基于改进 TrAdaBoost 的迁移学习系统，利用 One-Hot 编码和 One-vs-Rest 算法处理多分类问题，不仅使定位精度在动态环境下提升了 35%

面向 5G 新空口（NR）的 CSI 定位随着 5G 商用部署，利用 5G 信号特性的 CSI 定位技术开始兴起。针对室内单基站场景下的定位难题，[8]Zhou 等人提出了利用 5G 下行链路多波束特性的指纹定位方法，相较于单波束，多波束技术将 WKNN、SVM 和 RF 算法的 95% 定位误差分别降低了 99.7%、66.9% 和 95.3%。在另一项工作中，[9]Zhou 等人结合 ELM 降维和 CatBoost 算法提出了 EPL-CatLoc 方法，在单 gNB 覆盖的办公室和走廊场景中，定位准确率分别达到了 98.24% 和 94.10%。[10]Ruan 等人开发了 iPos-5G 系统，利用无监督深度自编码器重构特征并结合改进的径向基函数（RBF），在办公室场景下实现了 2.14 m 的平均绝对误差（MAE），优于 DeepFi 等传统方法。[17]Cerar 等人基于 CTW 2019 大赛的公开数据集，研究了利用卷积神经网络（CNN）处理单基站大规模 MIMO CSI 数据的潜力，在训练集与测试集随机分布的场景下，该结构相比于文献中现有的 CNN 和全连接网络模型，定位精度提升了 2 cm 到 10 cm，且所需的模型权重参数更少，计算效率更高。这证明了在处理高维 Massive MIMO 数据时，深层残差网络具有更强的特征提取能力。

### 1.2.3 基于线阵 CCD 的室内定位技术国内外研究现状

线阵 CCD 是一种光电转换器件，可将接收到的光信号转换成电荷并不断地累积起来，并以电压的形式输出，按照规定时序通过模数转换器（Analog-toDigital Converter, ADC）依次读取这些电压信号便可得到图像，具有单行光敏像元多、像元尺寸相对灵活、感光范围大、扫描速度快等优点。

尽管目前基于面阵相机（Area Array Camera）的视觉 SLAM 技术已成为室内定位的主流，但其在千赫兹级（kHz）高速动态捕捉及嵌入式微处理器低算力场景下仍面临巨大挑战。相比之下，线阵 CCD 凭借其一维成像带来的数据压缩优势及微秒级

响应速度，为解决上述瓶颈提供了极具潜力的差异化路径。

1.1 系统构型与成像机理的演变早期的线阵 CCD 定位系统多采用多相机交会测量原理。[1] Macellari 提出的 CoSTEL 系统是该领域的开创性工作，其利用三个一维传感器配合环形透镜，实现了对人体运动三维坐标的立体监测。在此基础上，[2] 秦志军等针对大型射电望远镜（FAST）馈源模型的检测，设计了由三个线阵 CCD 和复合柱面透镜组成的光电定位系统，通过大视场透镜设计解决了大范围运动目标的实时跟踪问题。针对多目标测量中的干扰问题，[3] 艾莉莉等设计了一种基于分光棱镜和滤光片的光学架构，构建了双点目标三维坐标重构子系统，有效解决了多相机与多合作目标一一对应时的信号串扰问题。

[6] 刘海庆等提出了一种正交柱面成像相机模型，在一个相机内集成两个正交放置的线阵 CCD，分别测量目标的水平角和垂直角，配合双相机交会即可实现高精度三维测量。[5] 杨凌辉等进一步发展了该技术，利用正交柱面成像相机结合光立体靶，仅需单目即可通过空间后方交会原理完成三维坐标测量，有效解决了大型装备测量中的视场遮挡难题。此外，[4] 王闯等提出了一种反向定位思路，将由三个线阵 CCD 和柱面透镜组成的传感器安装在移动端，通过观测天花板上的红外 LED 锚节点进行自主定位，避免了移动轨迹的暴露，便于室内大范围扩展。[7] 贺玉泉等则利用空间直线在成像面上的“积聚”特性，设计了双目线阵 CCD 平面定位系统，通过直方图均衡化增强图像对比度，实现了对目标平面的高精度定位。

1.2 相机标定技术标定是保证线阵 CCD 测量精度的核心环节，主要包括几何参数标定和畸变校正。由于线阵 CCD 通常配合柱面镜或广角镜头使用，光学畸变不可忽视。[9] 骆文博等建立了包含成像误差的一维摄像机修正模型，提出了“先畸变矫正、后参数标定”的两步法，显著提升了直接线性变换（DLT）的解算精度。[10] 王国辉等针对立体线阵相机，提出了一种利用简单磁性棒靶标的两步标定法，先通过线性模型求解初值，再利用 Levenberg-Marquardt 算法进行非线性优化，有效解决了镜头畸变参数的求解问题。为了简化标定流程，[8] 赵霞等提出了一种内外参数分离的标定方法，即内参在实验室利用平行光管标定，外参在现场利用三叉模板标定，降低了现场操作复杂度。针对传统物理模型参数繁多且非线性强的问题，[11] 李晶等引入了神经网络进行隐式标定，通过训练建立像点坐标与空间三维坐标的直接映射关系，避开了复杂的几何建模过程，实验证明该方法在坐标恢复精度上优于传统 DLT 方法。

1.3 光斑定位与信号处理在室内复杂光照环境下，如何从背景噪声中精确提取合作目标（如 LED）的亚像素质心是提升系统分辨率的关键。[12] 尚学军等对比了重心法和质心法，提出了基于线性插值和权值非线性的改进质心算法，在不同噪声水平下均能保持较高精度，定位精度可达  $1/25$  像素。[13] 曹苗等针对激光三角法测量中目标表面粗糙引起的散斑噪声，提出了基于区域划分和线性插值的改进灰度质心算法，通过并行下移阈值有效抑制了背景干扰。

除了算法改进，通过硬件与系统设计抑制背景光也是重要手段。[14] Ren 和 Kumar 在 LOSA-X 系统中引入了背景光消除机制，通过控制标记点 LED 的开关，利用差分算法去除动态背景照明，使系统具备了在复杂光照甚至户外环境下工作的能力。[15] 李晶等对线阵 CCD 探测能力的理论分析表明，背景光强度、目标亮度及积分时间直接影响系统的极限探测距离和速度，合理的参数匹配是保证微弱目标被捕获的前提。

在位姿解算算法方面，[16] 王艳等针对三线阵 CCD 系统，提出了一种改进的正交迭代算法（OI），通过建立新的物空间共线性误差目标函数并进行全局优化，解决了传统算法易陷入局部极值的问题，显著提高了位姿解算的收敛速度和精度。

为了进一步提升系统的动态性能和鲁棒性，[17] Kumar 和 Ren 开发了空间目标跟踪系统，将线阵光学传感器与 9 轴惯性测量单元（IMU）进行融合。[18] 在其后续的 LOSA-X 系统中，通过扩展卡尔曼滤波（EKF）算法融合视觉数据与惯性数据，不仅解决了视觉遮挡期间的数据丢失问题，还将位置追踪误差降低了 55%，实现了毫米级的定位精度和极高的更新率。与之相比，尽管商用系统如 Optotrak 能够提供极高的精度（微米级）[18]，但其高昂的成本限制了大规模应用，而基于线阵 CCD 的融合方案在保证精度的同时大幅降低了成本。

## 1.3 论文的研究内容和创新点

### 1.3.1 论文的主要研究内容

针对现有室内定位技术在数据获取人力时间成本、环境适应性及定位精度等方面局限，本文提出了一种基于线阵视觉与信道状态信息（CSI）的协同定位架构。全文按照“理论建模—关键技术研究—系统协同验证”的逻辑展开，主要研究内容按章节安排如下：

- **第一章：绪论** 阐述了室内定位技术的研究背景与意义，分析了基于视觉和 Wi-Fi

CSI 定位技术的国内外研究现状。分析了现有单一模态定位面临的数据依赖度高与鲁棒性不足等问题，提出了将线阵视觉的高精度特性与 CSI 的广覆盖特性相结合的研究思路，并确立了本文的研究目标与技术路线。

- **第二章：分布式节点的信道状态信息模型分析与数据预处理**实现了基于 ESP32 分布式节点的 CSI 特征提取与位置回归。从物理层面推导了单天线节点的 CSI 幅值干涉模型，分析了利用幅值特征表征空间位移的可行性。
- **第三章：基于信道状态信息的室内定位算法与系统设计**设计了包含多尺度时序卷积（MSTC）、ConvNeXt 主干网络及坐标注意力机制（CoordAtt）的深度神经网络，并使用了结合位置误差与物理平滑约束的复合损失函数，以实现端到端的特征映射。针对定位结果的时域平滑与抗差跟踪问题进行了研究。考虑到室内复杂环境下观测噪声具有非平稳特性，构建了基于新息（Innovation）统计特性的自适应贝叶斯跟踪框架。推导了观测噪声协方差矩阵的极大似然估计公式，并将其应用于粒子滤波算法中，设计了融合动态噪声感知的自适应粒子滤波。该算法旨在根据新息序列实时调整似然函数形态，以改善粒子贫化现象，提升系统在非视距场景下的鲁棒性。
- **第四章：线阵相机三维定位的成像模型与几何基础**详细分析了线阵 CCD 配合柱面透镜的“一维针孔”成像机理，建立了包含世界坐标系、相机坐标系及像素坐标系的多层几何模型。推导了基于“Y型”布局的多线阵相机光平面交汇几何模型及其线性代数求解方法，为获取室内空间位置信息提供了理论支撑。
- **第五章：线阵传感器的参数标定方法**针对视觉子系统的参数标定问题进行了深入研究。考虑到柱面透镜引入的非线性畸变，提出了一种分阶段标定策略：首先建立偶次多项式模型对镜头畸变进行校正，随后基于几何灵敏度与测量不确定性分析，提出了一种加权最大似然（WMLE）标定算法。该方法通过引入交叉相乘残差构建目标函数，并推导了基于像素噪声统计的权重计算方法，旨在提升非理想光学条件下的标定精度与数值稳定性。
- **第六章：协同系统结构和实验结果分析**完成了光电协同系统的集成与全流程验证。提出了“离线监督-在线互补”的分层协同架构：在数据层，利用线阵视觉系统辅助构建 CSI 指纹库；在策略层，构建了基于视距状态感知的“主备切换”机制。搭建了实验平台，分别验证了线阵视觉标定算法、CSI 定位网络及自适应跟踪算法的性能，并对协同架构在解决室内定位相关问题上的有效性进行了

实验分析。

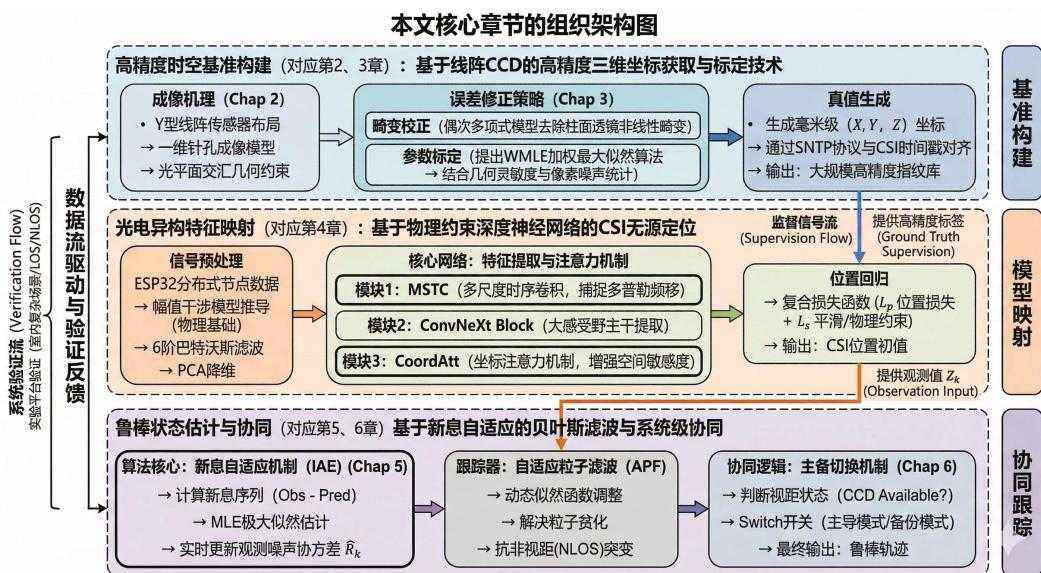


图 1.1 基于 CSI 与线阵视觉协同的室内定位系统总体架构图

### 1.3.2 论文的创新点

#### 1. 线阵相机高精度三维定位原理与标定方法研究

针对室内高精度真值获取难题，设计了基于“Y型”布局的多线阵 CCD 视觉捕捉系统。建立了基于柱面透镜的一维针孔成像数学模型，推导了多相机光平面交汇的三维重构算法。重点研究了线阵相机的标定技术，提出了一种基于几何灵敏度与测量不确定性的加权最大似然 (WMLE) 标定方法，并通过“去畸变-线性解算”的分步策略，有效解决了大视场柱面镜头的非线性畸变与参数求解问题。

#### 2. 基于 CSI 幅值干涉模型的无源定位深度学习网络设计

构建了基于分布式 ESP32-S3 节点的低成本 CSI 感知网络。从物理层面推导了 CSI 幅值干涉模型，论证了利用幅值特征表征空间位移的可行性。针对 CSI 信号的时变性与多径效应，设计了一种融合多尺度时序卷积 (MSTC)、ConvNeXt 主干网络及坐标注意力机制 (CoordAtt) 的深度神经网络。该网络通过提取 CSI 的时空频多维特征，并结合加权移动平均 (WMA) 与物理约束损失函数，实现了高精度的端到端位置回归。

#### 3. 基于新息自适应的贝叶斯目标跟踪算法研究

为了解决室内复杂环境下观测噪声非平稳导致的滤波发散问题，构建了基于新息（Innovation）统计特性的自适应贝叶斯跟踪框架。推导了观测噪声协方差矩阵的极大似然估计公式，提出了基于遗忘因子的动态噪声更新策略。将该机制应用于卡尔曼滤波与粒子滤波中，设计了融合动态噪声感知的自适应粒子滤波算法，通过动态调整似然函数形态，解决了粒子贫化问题，提升了系统在非视距与强干扰场景下的鲁棒性。

#### 4. “离线监督-在线互补”的光电协同架构实现与验证

提出了视觉与射频异构模态的深层协同机制。在数据层，利用 SNTP 协议实现了光电信号的微秒级同步，利用线阵视觉系统实现了 CSI 指纹库的自动化、高精度构建，突破了人工标注的效率瓶颈；在策略层，构建了基于视距状态感知的“主备切换”机制，利用视觉的高精度与射频的抗遮挡特性实现优势互补。搭建了实验平台，对各子系统性能及协同架构的有效性进行了系统性的实验验证。

## 第二章 Wi-Fi 信道状态信息定位技术相关理论

### 2.1 Wi-Fi 信号简介

#### 2.1.1 无线局域网 (WLAN) 标准概述

无线局域网（Wireless Local Area Network, WLAN）是指利用射频（Radio Frequency, RF）技术构建的局部网络系统，旨在实现终端设备的无线接入与数据交互。作为 WLAN 领域应用最为广泛的技术标准，Wi-Fi（Wireless Fidelity）主要基于国际电气和电子工程师协会（IEEE）制定的 802.11 标准族。

自 1997 年首个标准发布以来，IEEE 802.11 经历了多次重大技术演进。其核心演进逻辑主要围绕提高吞吐量、扩展频段资源以及优化多用户并发性能展开：

- 初期标准：如 802.11b/g 主要工作在 2.4GHz 频段，奠定了无线通信的基础框架。
- 高吞吐量阶段：802.11n (Wi-Fi 4) 引入了多输入多输出（MIMO）技术，显著提升了传输速率。
- 全频段与高效率阶段：802.11ac (Wi-Fi 5) 进一步优化了 5GHz 频段的传输能力，而 802.11ax (Wi-Fi 6) 则引入了正交频分多址（OFDMA）技术，有效解决了高密度场景下的频谱效率问题。

目前，Wi-Fi 技术已从单纯追求峰值速率转向对低时延、高可靠性及多频段协同（2.4GHz、5GHz 及 6GHz）的深度优化。

#### 2.1.2 Wi-Fi 信号的传播特征

Wi-Fi 信号本质上属于高频电磁波，其在空间中的传播过程并非理想化的直线传输，而是受到传播距离、环境结构及动态扰动等多种物理因素的共同影响。为了为后续基于信道状态信息（CSI）的感知与定位方法提供物理基础，本节从路径损耗、多径传播以及信道时变性三个方面，对室内 Wi-Fi 信号的传播特征进行系统分析。

在理想自由空间条件下，Wi-Fi 信号在传播过程中主要表现为能量的空间扩散，其路径损耗可由 Friis 传输方程进行描述。设发射功率为  $P_t$ ，接收功率为  $P_r$ ，则二者之间的关系可表示为：

$$P_r = P_t G_t G_r \left( \frac{\lambda}{4\pi d} \right)^2 L^{-1}, \quad (2.1)$$

其中,  $G_t$  与  $G_r$  分别表示发射端与接收端天线增益,  $\lambda$  为信号波长,  $d$  为传播距离,  $L$  为系统损耗因子。该模型表明, 在理想条件下接收信号功率与传播距离的平方成反比。

然而, 在实际室内环境中, 自由空间模型难以准确刻画复杂传播条件下的信号衰减特性。为此, 工程实践中通常采用对数距离路径损耗模型 (Log-distance Path Loss Model) 对平均路径损耗进行建模, 其表达式为:

$$PL(d) = PL(d_0) + 10n \log_{10} \left( \frac{d}{d_0} \right) + X_\sigma, \quad (2.2)$$

其中,  $n$  为路径损耗指数, 用以表征环境复杂程度,  $X_\sigma$  为均值为零、标准差为  $\sigma$  的高斯随机变量, 用于描述由墙体、家具等障碍物遮挡引起的阴影衰落效应。该模型在统计意义上更符合室内无线传播的实际特性。

除路径损耗外, 多径传播是室内 Wi-Fi 信号的另一显著特征。由于电磁波在传播过程中会与墙体、地面及各类物体发生反射、绕射与散射, 接收端获得的信号往往由多条传播路径分量叠加而成。该过程可通过信道冲激响应 (Channel Impulse Response, CIR) 进行描述:

$$h(t) = \sum_{l=1}^L \alpha_l e^{j\phi_l} \delta(t - \tau_l), \quad (2.3)$$

其中,  $L$  表示有效传播路径的数量,  $\alpha_l$ 、 $\phi_l$  与  $\tau_l$  分别对应第  $l$  条路径的幅度、相位偏移及时延。由于不同路径分量具有不同的到达时间, 多径效应会在频域上引入频率选择性衰落, 这也是 Wi-Fi 系统普遍采用正交频分复用 (OFDM) 技术以对抗信道失真的重要原因。

此外, 室内无线信道通常表现出显著的时变特性。当环境中存在人员移动或物体位移时, 多径结构会随时间发生变化, 导致信道增益和相位持续波动。在无明显直射路径的非视距 (Non-Line-of-Sight, NLOS) 条件下, 接收信号幅度常采用瑞利衰落模型进行统计描述; 而在存在较强直射分量的视距 (Line-of-Sight, LOS) 场景中, 则更适合使用莱斯衰落模型。上述衰落模型为理解信道随机性及其对通信性能的影响提供了理论依据。

对于基于 CSI 的无线感知与定位研究而言, 相较于仅反映整体接收功率变化的 RSSI 指标, CSI 能够在子载波级别刻画信道的幅度与相位响应, 从而更加细粒度地反映多径结构与环境变化。这一特性为后续从无线信号中挖掘目标空间信息提供了

重要支撑。

## 2.2 Wi-Fi 无线定位参数

在无线定位系统中，特征参数的提取精度直接决定了定位算法的性能上限。Wi-Fi 定位技术主要利用无线网卡在通信过程中产生的物理层或介质访问控制（MAC）层信息，通过建立信号特征与空间坐标之间的映射关系来实现目标定位。本节将详细探讨基于 Wi-Fi 信号的两种主流定位表征参数：接收信号强度指示（RSSI）与信道状态信息（CSI）。两者分别从宏观功率与微观频谱特性描述了无线信道的状态。

### 2.2.1 接收信号强度指示（RSSI）

接收信号强度指示（Received Signal Strength Indicator, RSSI）是一种由无线通信系统介质访问控制层（MAC Layer）提供的全局性参数，用于刻画接收端天线在给定时刻所观测到的信号功率水平。由于 RSSI 的获取过程不依赖于底层信道的精细结构，其实现方式简单、适用范围广，因此在早期无线定位系统中被广泛采用。

从物理意义上讲，RSSI 综合反映了发射功率、传播路径损耗、天线增益以及环境噪声等多种因素的叠加效应。在对数尺度下，其与发射端到接收端之间传播距离的关系通常可由对数距离路径损耗模型近似表示为：

$$\text{RSSI}(d) = \text{RSSI}_0 - 10n \log_{10}\left(\frac{d}{d_0}\right) + X_\sigma, \quad (2.4)$$

其中， $d$  表示发射端与接收端之间的传播距离， $d_0$  为参考距离（通常取 1 m）， $\text{RSSI}_0$  表示在参考距离处测得的接收信号强度， $n$  为路径损耗指数，用于刻画不同传播环境下信号衰减速率， $X_\sigma$  为均值为零、方差为  $\sigma^2$  的随机扰动项，用以描述阴影衰落及测量噪声的影响。

在理想视距（Line-of-Sight, LOS）条件下，上式表明 RSSI 随传播距离的增加呈现出近似单调递减的趋势。这一特性构成了基于 RSSI 的指纹定位方法及测距型定位算法的物理基础。然而，该模型在复杂室内环境中的适用性往往受到多径传播与环境动态变化的显著制约。

在实际应用中，RSSI 的主要局限性源于其特征表达形式的高度简化。首先，RSSI 本质上是多径信号分量在接收端叠加后的总功率标量，无法区分直射路径与反射路径的独立贡献。当接收位置发生微小变化时，不同传播路径之间的相位关系可能发

生显著改变，从而引发相长或相消干涉，使 RSSI 出现剧烈波动，即所谓的小尺度衰落现象。其次，RSSI 对环境动态变化较为敏感，人体遮挡、设备姿态变化以及同频无线干扰等因素均可能导致其时间序列呈现明显的非平稳特性。

综合而言，RSSI 作为一种低维度、易获取的无线特征，在实现简单定位或区域级感知任务时具有一定实用价值。

### 2.2.2 信道状态信息（CSI）

相较于基于接收信号强度指示（Received Signal Strength Indicator, RSSI）的定位方法，信道状态信息（Channel State Information, CSI）提供了更加细粒度的物理层特征描述。在 IEEE 802.11n/ac/ax 等现代 Wi-Fi 标准中，正交频分复用（OFDM）与多输入多输出（MIMO）技术的引入，使得商用无线网卡可以在驱动层获取子载波级别的信道频率响应（Channel Frequency Response, CFR），从而为高精度无线感知与定位提供了数据基础。

从系统模型的角度来看，在 OFDM 通信系统中，宽带信道被划分为多个相互正交的窄带子载波。假设系统配置为  $N_t$  根发射天线与  $N_r$  根接收天线，则在第  $k$  个子载波上，接收信号向量  $\mathbf{y}_k$  与发射信号向量  $\mathbf{x}_k$  之间可表示为：

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{n}_k, \quad (2.5)$$

其中  $\mathbf{n}_k$  为加性高斯白噪声向量， $\mathbf{H}_k \in \mathbb{C}^{N_r \times N_t}$  即为该子载波对应的信道状态信息矩阵。当通信链路退化为单输入单输出（SISO）形式时， $\mathbf{H}_k$  可简化为一个复数标量：

$$H_k = |H_k| e^{j\angle H_k}, \quad (2.6)$$

其中  $|H_k|$  与  $\angle H_k$  分别表示第  $k$  个子载波的幅度衰减与相位偏移。该复数形式意味着 CSI 不仅反映了信号强度变化，还隐含了传播路径长度、到达时间以及空间角度等关键信息，这是其区别于 RSSI 的核心优势。

从无线传播机理出发，CSI 可以被理解为多径效应在频域中的综合响应。在典型室内环境中，信号从发射端到接收端通常会经历直射路径（LOS）以及多条由墙壁、地面和物体反射或散射形成的非直射路径（NLOS）。根据多径叠加原理，第  $k$  个子

载波上的 CSI 可表示为各条传播路径分量的矢量叠加形式：

$$H_k = \sum_{l=1}^L \alpha_l(t) e^{-j2\pi f_k \tau_l}, \quad (2.7)$$

其中  $L$  表示有效传播路径数量， $f_k$  为第  $k$  个子载波的中心频率， $\alpha_l(t)$  为第  $l$  条路径在时刻  $t$  的复衰减系数（综合了路径损耗与初始相位）， $\tau_l$  则表示对应的传播时延。由于不同路径的时延存在差异，各子载波在频域上表现出不同的幅度起伏与相位旋转，从而形成具有空间唯一性的频率选择性特征。这种特征在目标位置变化时会发生可观测的变化，为基于 CSI 的定位与感知方法提供了物理基础。

在实际室内场景中，环境通常由静态背景结构与动态目标共同构成，相应地，接收到的 CSI 也可视为静态信道分量与动态扰动分量的叠加。当人员或物体在空间中运动时，其引入的反射路径长度将随时间发生变化，从而导致信道相位出现连续演化。若在时刻  $t$  第  $l$  条路径的长度变化量为  $\Delta d_l(t)$ ，则由该动态路径引起的相位变化可近似表示为：

$$\Delta\phi_l(t) = \frac{2\pi}{\lambda} \Delta d_l(t), \quad (2.8)$$

其中  $\lambda$  为载波波长。由此可见，CSI 相位随时间的微小波动与物理空间中的距离变化之间存在直接对应关系，使其对人体运动与位置变化高度敏感。

需要指出的是，在商用 Wi-Fi 硬件平台中，直接获取的 CSI 相位往往受到载波频率偏移（Carrier Frequency Offset, CFO）和采样时钟偏移（Sampling Frequency Offset, SFO）等非理想因素的影响，表现为整体随机相位偏置与线性漂移项的叠加。这类硬件引入的相位误差并不来源于真实的物理传播过程，若不加处理将显著削弱 CSI 在定位与感知任务中的有效性。

### 2.2.3 RSSI 与 CSI 参数性能对比

尽管接收信号强度指示（Received Signal Strength Indicator, RSSI）与信道状态信息（Channel State Information, CSI）均来源于 Wi-Fi 通信过程，但二者在信号提取层级、物理含义及信息表达形式上的本质差异，决定了其在室内定位与无线感知任务中的性能表现存在显著不同。为明确两类特征在定位应用中的适用性，本节从信息粒度、多径分辨能力等方面对 RSSI 与 CSI 进行对比分析。

从信息表达能力的角度来看，RSSI 属于介质访问控制层（MAC Layer）提供的统计量，其本质是对整个信道带宽内接收信号功率的整体度量。该参数通过对频域能量进行平均化处理，将复杂的信道响应压缩为单一标量数值，虽然便于获取与使用，但不可避免地丢失了大量与空间结构相关的频率信息。相比之下，CSI 直接来源于物理层（PHY Layer），能够在子载波尺度上描述无线信道的频率响应特性。在典型的 20 MHz 带宽配置下，基于 IEEE 802.11n 标准的 CSI 可提供 30–56 个子载波的幅度与相位信息，而在 802.11ax/be 等新一代标准中，其频率分辨率进一步提升。这种细粒度的频谱刻画能力使 CSI 能够更充分地反映空间位置变化对信道响应的影响。此外，在多输入多输出（MIMO）系统中，CSI 以矩阵形式同时包含多天线间的耦合关系，天然携带空间维度信息，为到达角（Angle of Arrival, AoA）估计与空间感知算法提供了基础。

在复杂室内环境中，多径效应是制约定位精度的关键因素之一。RSSI 对多径的处理方式本质上是“被动叠加”：直射路径与所有反射路径的信号功率在接收端进行非相干合成，由于不同路径之间相位关系的随机性，极易产生相长或相消干涉现象。这种干涉效应导致 RSSI 在空间上呈现出明显的小尺度衰落特性，即使目标位置发生微小变化，其测量值也可能出现剧烈波动，同时 RSSI 本身无法区分视距（LOS）与非视距（NLOS）传播路径。相比之下，CSI 借助 OFDM 体制对频率选择性衰落的分解能力，能够在频域上刻画不同子载波所经历的多径影响。进一步地，通过对 CSI 进行逆傅里叶变换（IFFT），可在时域中获得信道的功率延迟谱（Power Delay Profile, PDP），从而在一定程度上分离出直射路径与主要反射路径分量。这一特性使得 CSI 在非视距条件下仍能保留具有判别力的空间特征，显著提升了其在复杂环境中的定位鲁棒性。

为更加直观地展示两类参数在关键性能指标上的差异，表 2.1 给出了 RSSI 与 CSI 的对比结果。

综上分析可知，RSSI 虽然具有获取方式简单、设备通用性强等优势，但其粗粒度特征形式在频率分辨、多径解析及时序稳定性等方面均存在明显局限，难以满足高精度室内定位的需求。相比之下，CSI 能够从物理层对无线信道传播特性进行更为细致的刻画，尤其在多径分辨与空间敏感性方面展现出显著优势。因此，本文后续研究选择以 CSI 作为核心无线特征参数，并在此基础上开展协同定位方法的设计与验证。

表 2.1 RSSI 与 CSI 性能特征对比

对比维度	RSSI	CSI
网络层级	MAC 层	PHY 层
数据结构	单一标量 (dBm)	复数向量/矩阵 (幅度 + 相位)
频率分辨率	全频段平均功率	子载波级细粒度响应
多径处理能力	易受多径叠加影响, 难以区分	可利用多径特征, 具备分离能力
时间稳定性	波动较大, 受 AGC 影响明显	相对稳定, 特征维度丰富
定位分辨能力	空间区分能力有限	空间区分能力较强

## 2.3 CSI 定位相关技术基础

本节从物理层通信机制的角度, 对信道状态信息 (Channel State Information, CSI) 的产生机理及其技术基础进行说明。CSI 的可获取性依赖于现代 Wi-Fi 标准 (IEEE 802.11n/ac/ax) 所采用的正交频分复用 (OFDM) 与多输入多输出 (MIMO) 技术。OFDM 在频域上对宽带信道进行离散化建模, 而 MIMO 在空间维度上扩展了信道的观测自由度, 两者的结合使得无线信道能够以多维形式被刻画, 从而为基于 CSI 的室内定位提供了基础数据支撑。

### 2.3.1 正交频分复用技术 (OFDM)

正交频分复用 (Orthogonal Frequency Division Multiplexing, OFDM) 是一种典型的多载波调制技术, 其核心思想是将宽带高速数据流映射到一组在频域上相互正交的子载波上进行并行传输。设 OFDM 系统包含  $K$  个子载波, 符号持续时间为  $T$ , 第  $k$  个子载波的频率为:

$$f_k = f_0 + \frac{k}{T}, \quad k = 0, 1, \dots, K - 1, \quad (2.9)$$

则不同子载波在区间  $[0, T]$  内满足正交条件:

$$\int_0^T e^{j2\pi f_k t} e^{-j2\pi f_m t} dt = \begin{cases} T, & k = m, \\ 0, & k \neq m. \end{cases} \quad (2.10)$$

该正交性保证了接收端能够通过快速傅里叶变换 (FFT) 在频域中无干扰地分离各子载波信号, 从而提高系统的频谱利用效率。

在多径环境下, OFDM 通过引入循环前缀 (Cyclic Prefix, CP) 有效抑制符号间

干扰。当循环前缀长度大于信道的最大时延扩展时，宽带频率选择性衰落信道可被等效分解为一组相互独立的窄带平坦衰落子信道。在此条件下，第  $k$  个子载波上的接收信号可表示为：

$$Y_k = H_k X_k + N_k, \quad (2.11)$$

其中  $X_k$  与  $Y_k$  分别为发射与接收符号， $N_k$  为加性高斯白噪声， $H_k$  为对应子载波的信道频率响应系数。

从定位角度看， $\{H_k\}$  构成了无线信道在频域上的离散采样结果，其复数幅度与相位共同反映了多径传播条件下不同频率分量的衰落特性。这一子载波级信道响应集合正是 CSI 在 OFDM 系统中的物理本质。

### 2.3.2 多输入多输出技术（MIMO）

多输入多输出（Multiple-Input Multiple-Output, MIMO）技术通过在发射端和接收端分别部署多根天线，利用无线信道在空间维度上的独立性来增强系统性能。对于一个具有  $N_t$  个发射天线和  $N_r$  个接收天线的系统，在第  $k$  个子载波上，其信道模型可表示为：

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{n}_k, \quad (2.12)$$

其中  $\mathbf{H}_k \in \mathbb{C}^{N_r \times N_t}$  为子载波  $k$  上的 MIMO 信道矩阵， $\mathbf{x}_k$  和  $\mathbf{y}_k$  分别表示发射与接收信号向量。

结合 OFDM 体制后，CSI 在整体上可表示为一个三维数据结构：

$$\mathcal{H} = \{\mathbf{H}_k\}_{k=1}^K, \quad (2.13)$$

其维度由子载波数量、发射天线数量以及接收天线数量共同决定。与单天线系统相比，多天线配置显著增加了可观测的独立信道链路数量，使得无线信道在空间维度上的差异性能够被更充分地表征。

从空间感知的角度来看，不同接收天线之间的相位差与幅度差隐含了与信号到达方向及空间几何结构相关的信息；同时，多链路联合观测也增强了 CSI 在不同空间位置下的区分能力，有利于构建更具判别性的定位特征。

### 2.3.3 OFDM 与 MIMO 的协同作用

综合而言，OFDM 通过将宽带信道离散化为子载波级响应，在频域上细化了信道的描述方式；MIMO 则通过多天线结构在空间维度上扩展了信道的观测自由度。二者的协同作用使 CSI 能够从频率与空间两个维度对无线传播环境进行刻画，从而为后续基于 CSI 的室内定位与感知方法提供必要的物理与数学基础。

## 第三章 基于信道状态信息的室内定位算法与系统设计

### 3.1 引言

本章聚焦于复杂室内环境下的无源定位问题，重点研究在视觉感知失效或受限条件下，如何利用商用 Wi-Fi 设备获取的信道状态信息（Channel State Information, CSI）实现高精度的目标跟踪。作为光电协同定位架构中的射频感知分支，CSI 定位子系统不仅需要具备独立工作的能力，\*\*\*\*\* 更需为顶层融合策略提供鲁棒的位置先验。

本章从 3 个方面论述基于 CSI 的动态室内定位方法设计。

### 3.2 基于嵌入式设备的分布式单天线节点分析

为了验证 CSI 信号在室内空间感知中的有效性，本文采用低成本嵌入式设备构建三维空间中的单天线无线节点，并在典型的 Wi-Fi 802.11n 标准、40 MHz 信道带宽条件下开展实验测试。本节通过对比不同硬件平台及不同空间布局下的信号表现，阐述选择幅值特征作为定位输入的物理依据。

#### 3.2.1 空间位置对 CSI 幅值的响应特性

如图 3.1 所示，当 ESP32-C6 放置于天线 1 位置时，针对不同空间位置的铝箔包裹泡沫小球，测得的 CSI 幅值随位置变化呈现出明显差异。图 3.1(a) 给出了原始 CSI 数据结果，其中共包含 114 个子载波。可以观察到，部分子载波（如第 28 和第 84 个子载波）出现了较为明显的幅值下冲现象，同时子载波频谱左、右两侧的幅值分布特性存在一定差异。该现象与 Wi-Fi 标准中直流子载波（DC subcarrier）及保护子载波的设置有关，这些子载波通常不承载有效数据信息。为减少其对后续分析的影响，本文对上述子载波进行剔除，最终保留 112 个有效子载波，其对应的 CSI 幅值变化结果如图 3.1(b) 所示。

实验中选用铝箔包裹的泡沫小球作为目标物体。该目标具有较强的电磁波反射特性，尽管其物理尺寸较小，但在室内多径环境中仍可能对无线传播路径产生影响。当目标位于 Position 1（橙色曲线）时，部分子载波幅值相对其他位置呈现出较高水

平，该现象可能与目标引入的反射分量与直射路径信号之间的叠加效应有关，从而在接收端表现为幅值增强。相对地，位于曲线下方的幅值分布表明，在对应位置下，反射分量与主传播路径之间的相位关系发生变化，导致接收信号幅值出现一定程度的降低。

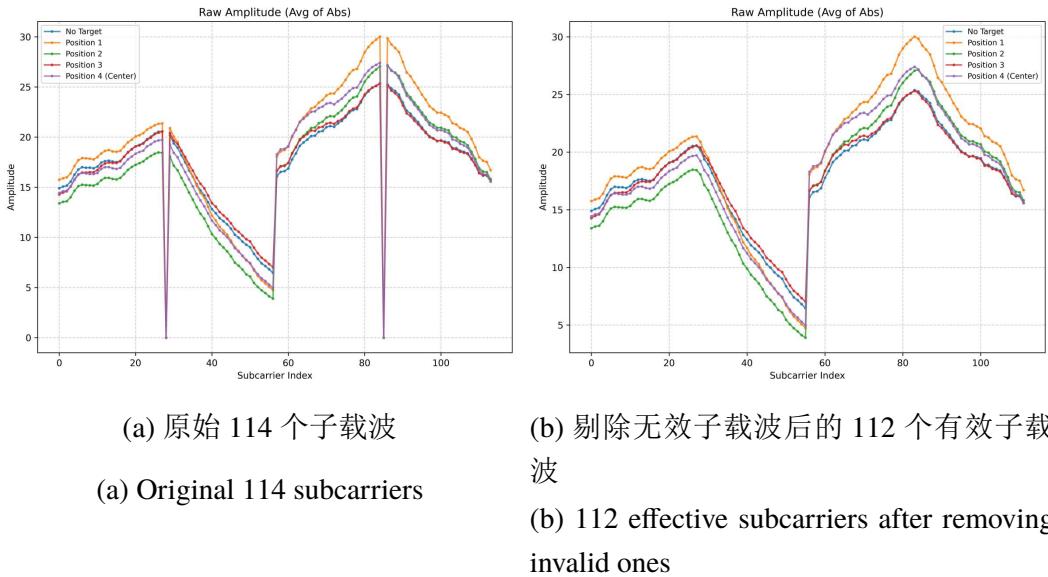


图 3.1 ESP32-C6 在天线 1 位置对不同目标位置的幅值响应

Figure 3.1 Amplitude response of ESP32-C6 at antenna 1 for different target positions

图 3.2 给出了目标小球在三维空间中移动至不同位置时的示意图。当目标物体处于不同空间位置时，其引入的无线传播路径及传播条件随之变化，从而导致信号衰减特性存在差异。这些空间位置变化所引起的传播特性差异，最终在 CSI 中表现为幅值扰动。上述结果表明，CSI 对环境中目标物体位置变化具有一定的响应能力，可为后续基于 CSI 的空间感知与定位分析提供实验依据。

图 3.2 目标小球在三维空间中不同测试位置的示意图

此外，如图 3.3 和图 3.4 所示（均包含 a、b、c 三个子图），本文进一步在天线 2 所在位置，对 ESP32 与 ESP32-C6 两种嵌入式设备进行了对比实验，测试其在三维空间中面对不同位置的铝箔包裹泡沫小球时所对应的 CSI 变化特性。

**1. 幅值分布区分度：**在子图 a 中给出了不同目标位置下的子载波幅值分布情况。

当设备位于天线 2 位置时，可以观察到 ESP32 在不同空间位置（Position 1-4）下的 CSI 幅值曲线在整体幅度层面高度重合，位置间差异不易直接区分。相比

之下，ESP32-C6 在相同实验条件下，其不同位置对应的幅值曲线在整体幅度水平上呈现出更为明显的分层特征，表明其对子载波幅值变化的响应在不同空间位置间具有更高的可区分性。

2. **差分特征分析：**子图 b 给出了在去除静态环境影响后得到的幅值差分结果 (Amplitude Difference, Target-No Target)。对于 ESP32-C6，不同位置对应的幅值差分曲线在波形形态及幅度范围上均表现出较为明显的差异，其动态范围约分布在 -6 至 6 之间。该特性表明，不同空间位置引入的 CSI 扰动在差分特征中得到了较为充分的体现。相对而言，ESP32 的幅值差分主要集中在 -1 至 1.5 的较窄区间内，且不同位置之间的曲线重叠度较高，整体波动形态较为相似。
3. **相位线性度对比：**在子图 c 中对两种设备的 CSI 相位特性进行了对比分析。可以观察到，ESP32 的相位曲线在子载波索引 0-80 区间内呈现出明显的非线性起伏特征，并在约第 80 个子载波附近出现显著的相位不连续现象（相位突变）。该现象表明其相位响应中存在明显的非线性畸变，可能受符号定时偏移 (STO) 及载波频率偏移 (CFO) 等因素影响。相比之下，ESP32-C6 的相位曲线整体呈现出较为规则的近线性变化趋势，形态上接近“V”字型分布，表明其线性一致性较好，残余偏差更易于建模补偿。

综合上述对比结果，ESP32-C6 在 CSI 幅值区分度及相位稳定性方面均表现出更优特性。因此，本文后续实验选用 ESP32-C6 作为采集平台。

### 3.2.2 多节点空间分布特征分析

图 3.5 展示了三台 ESP32-C6 设备分别部署于天线 1、2、3 位置时测得的 CSI 特性。尽管目标物体保持静止，但物理位置的差异导致无线传播路径及多径条件存在显著区别。

- **多径“指纹”差异：**图 3.5(a) 显示设备 2 在低频子载波 (20-40) 出现峰值，而设备 1 在高频子载波 (80) 附近达到峰值。这反映了不同位置接收到的多径分量在叠加方式上的差异。
- **稳定性分布：**图 3.5(b) 为幅值标准差分布。设备 2 虽然平均幅值较高，但标准差也相对较大，说明该位置对环境扰动更为敏感。
- **相位变化趋势：**图 3.5(c) 测得的去卷绕相位显示，设备 1 与 3 的斜率接近，路径特性相似；而设备 2 相位趋势差异显著，反映了其独特的信号传播条件。

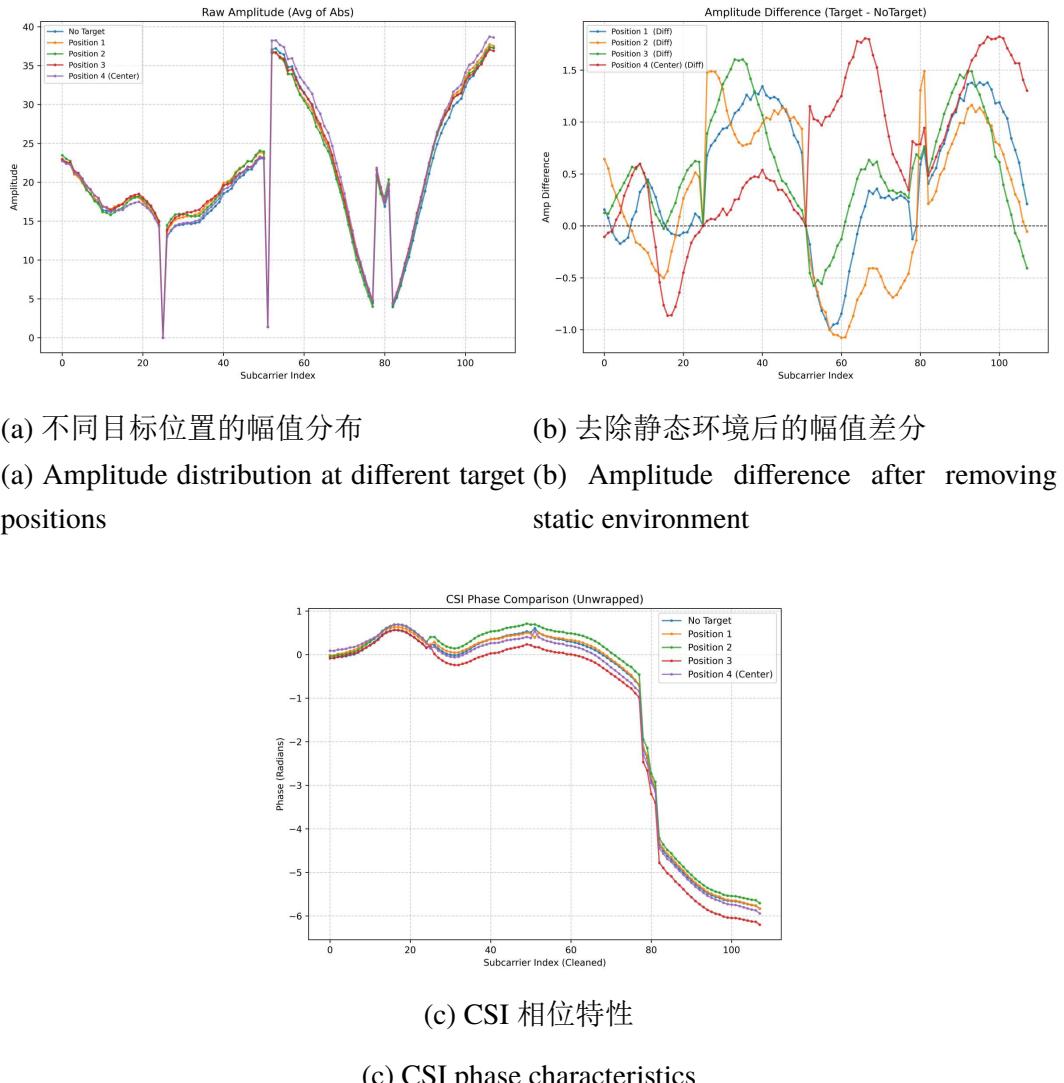


图 3.3 ESP32 在天线 2 位置针对不同目标位置的 CSI 变化特性

Figure 3.3 CSI variation characteristics of ESP32 at antenna 2 for different target positions

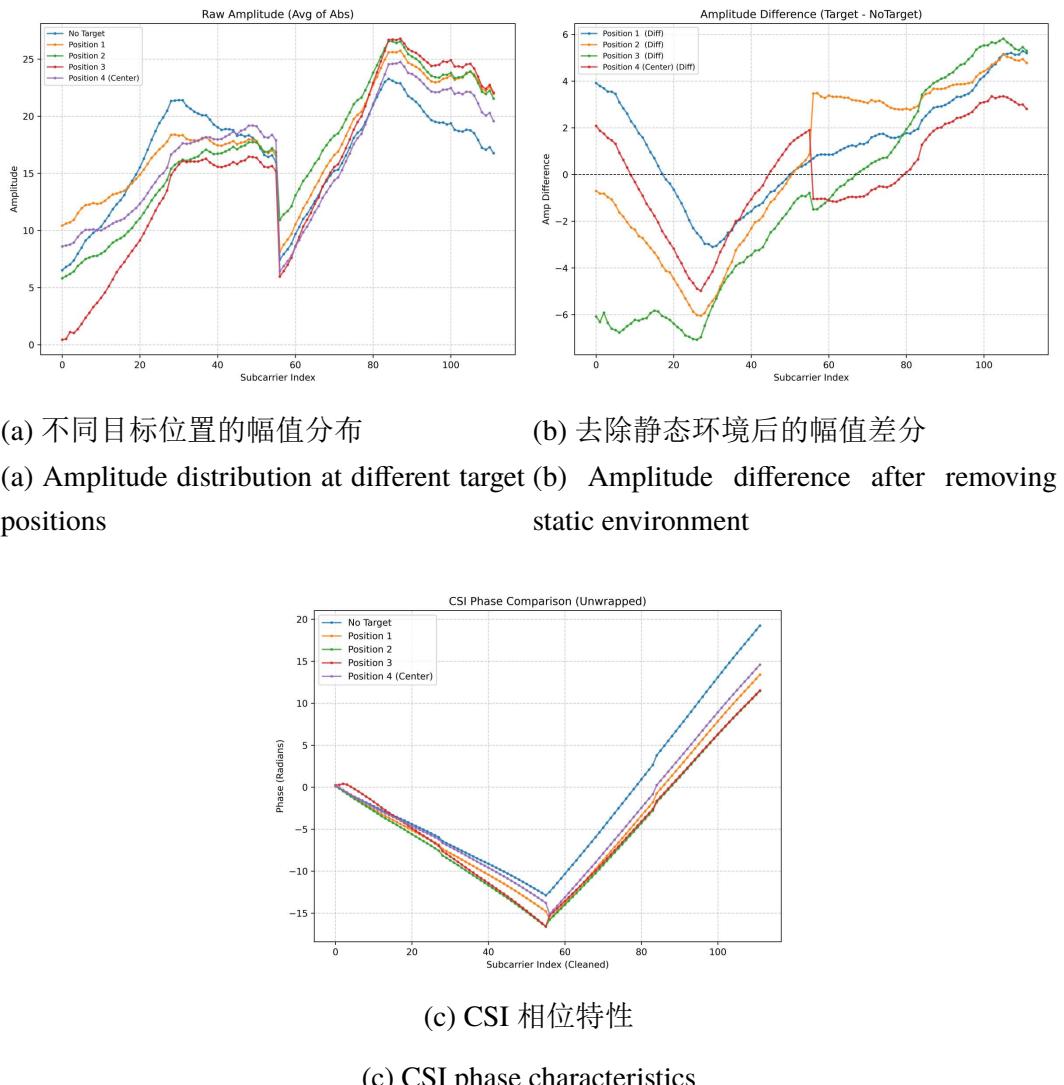


图 3.4 ESP32-C6 在天线 2 位置针对不同目标位置的 CSI 变化特性

Figure 3.4 CSI variation characteristics of ESP32-C6 at antenna 2 for different target positions

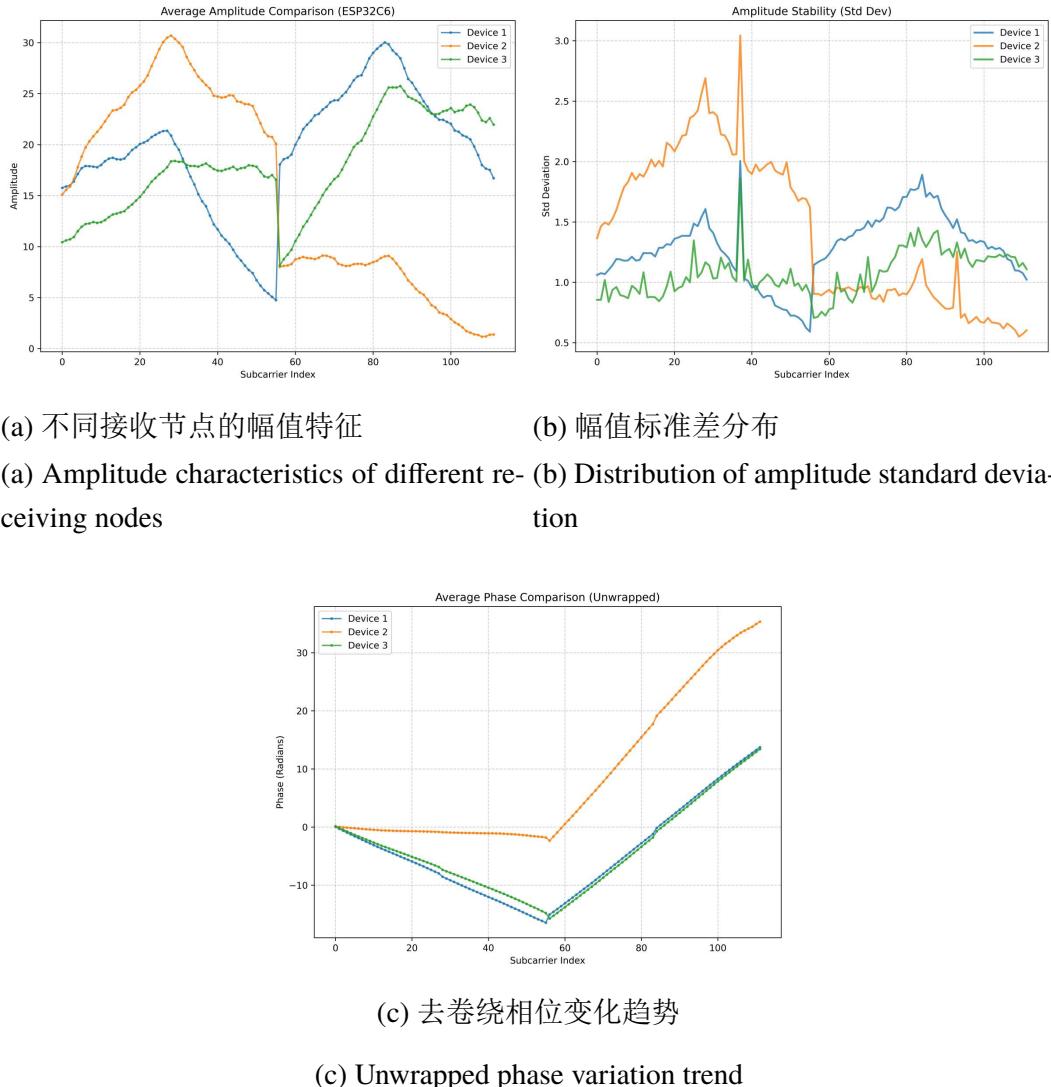


图 3.5 三台 ESP32-C6 设备在不同天线位置的 CSI 特性对比

Figure 3.5 Comparison of CSI characteristics of three ESP32-C6 devices at different antenna positions

此外，图 3.6(a)–(c) 还通过三维曲面图（Surface Plot）直观展示了设备 3 在时频域中对三个不同位置的铝箔小球的结构化差异，验证了多节点信息融合定位的可行性。

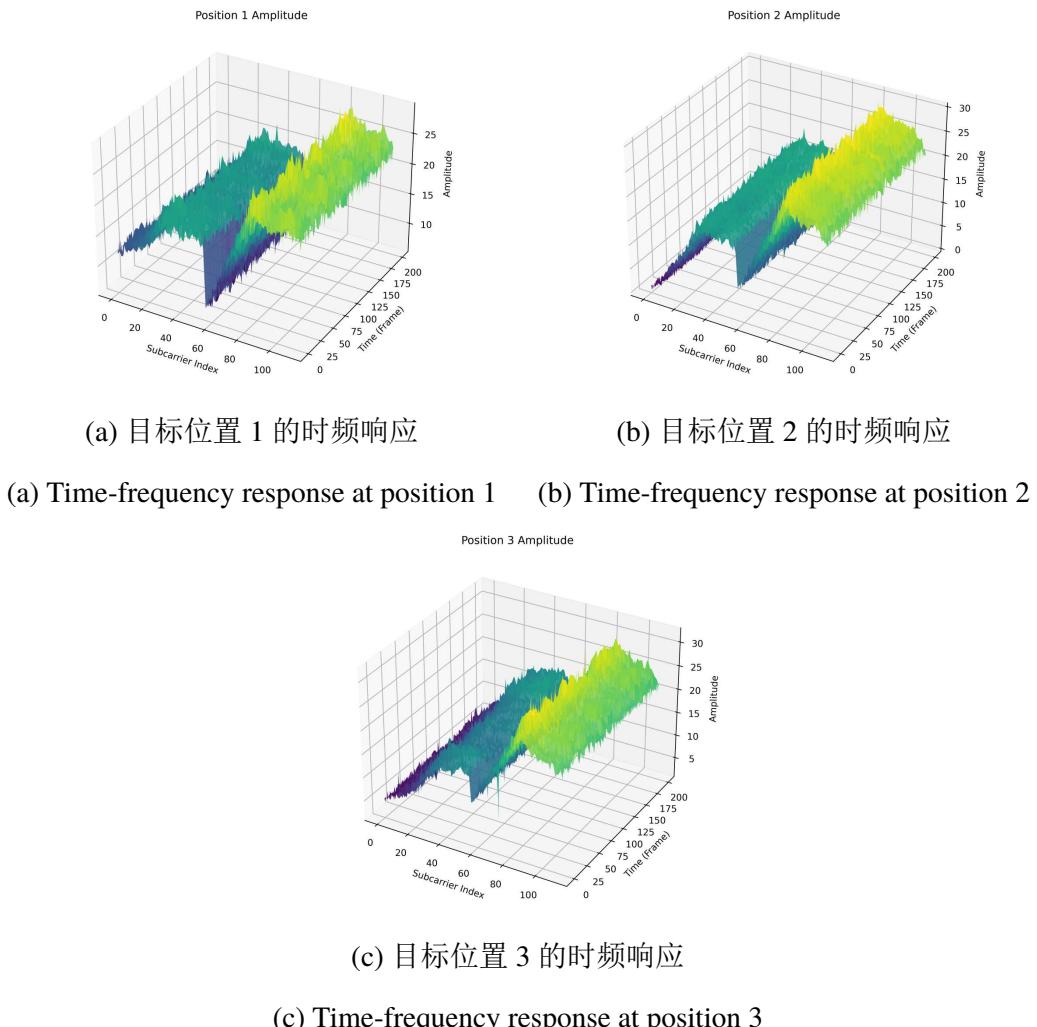


图 3.6 设备 3 针对不同目标位置的 CSI 时频域三维特征

Figure 3.6 3D time-frequency characteristics of CSI from device 3 for different target positions

### 3.3 幅值与相位特征的稳定性与可用性分析

#### 3.3.1 幅值与相位稳定性分析

为进一步定量分析 CSI 在不同空间位置下的稳定性与区分能力，本文选取第 20 个子载波作为代表，对其幅值与相位特性进行统计分析。相关结果如图 3.7 所示，其

中图 3.7(a) 为第 20 子载波的 IQ 星座分布, 图 3.7(b) 为对应的幅值分布情况, 图 3.7(c) 展示了幅值随时间变化的轨迹。

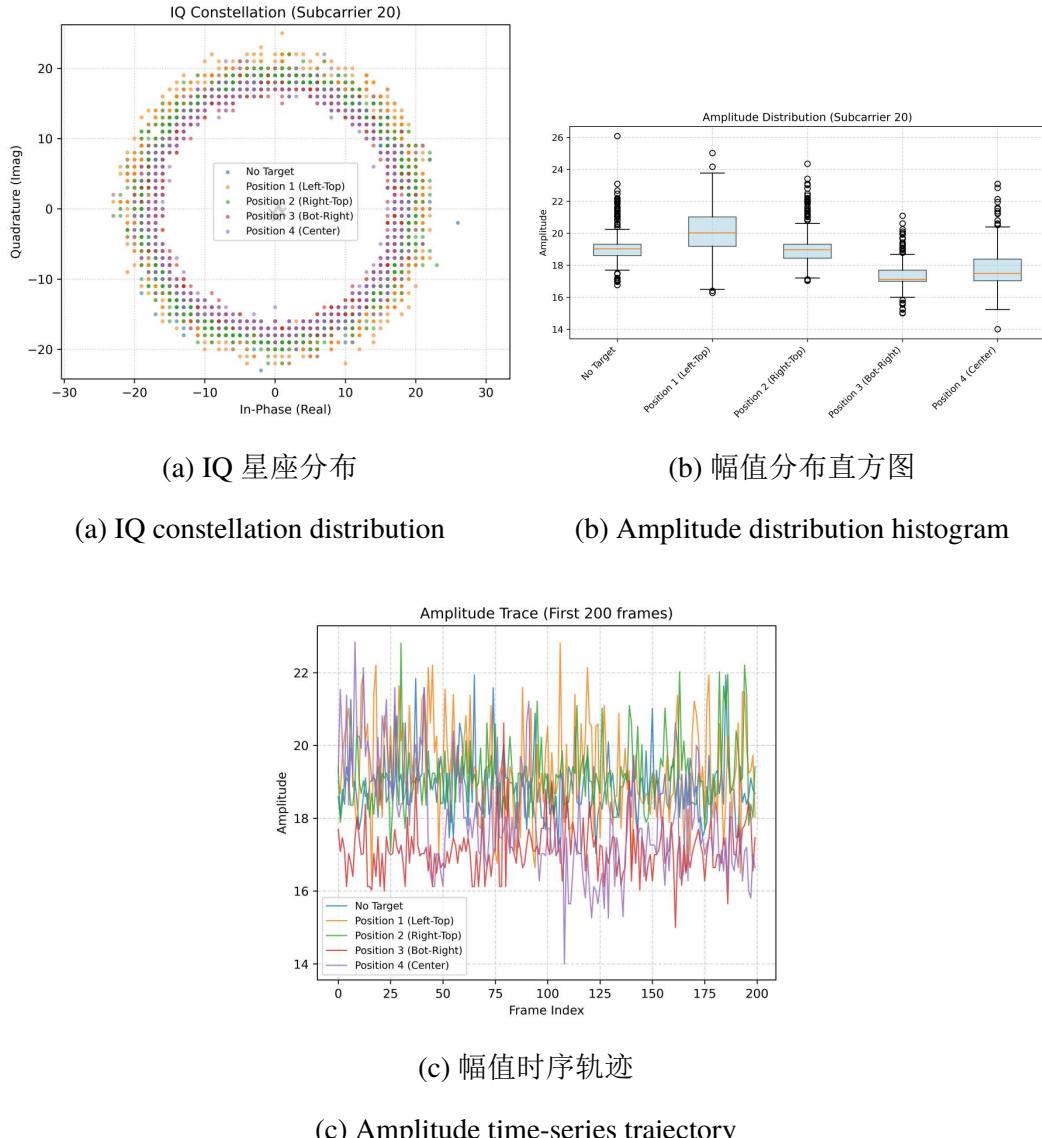


图 3.7 第 20 号子载波特征分析

Figure 3.7 Feature analysis of the 20th subcarrier

### 3.3.1.1 (1) 幅值稳定性指标: 变异系数

为衡量 CSI 幅值在时间维度上的稳定性, 本文采用变异系数 (Coefficient of Variation, CV) 作为评价指标。变异系数定义为幅值标准差与均值之比, 可用于描述信号

相对离散程度，其数值越小，表明信号在统计意义上越稳定。其数学表达式为：

$$CV = \frac{\sigma}{\mu} \times 100\% \quad (3.1)$$

其中， $\sigma$  表示 CSI 幅值的标准差， $\mu$  表示幅值的平均值。

### 3.3.1.2 (2) 基于复平面 IQ 中心距离的区分度指标

除幅值稳定性外，为评估不同空间位置对应信号在复平面上的可分性，本文进一步引入基于 IQ 星座中心距离的区分度指标。该指标通过计算不同位置下 CSI 复数均值点在复平面中的欧几里得距离，反映其在统计意义上的分离程度。

设位置 A 与位置 B 在复平面中的中心点分别为：

$$C_A = I_A + jQ_A, \quad C_B = I_B + jQ_B \quad (3.2)$$

则二者的中心距离定义为：

$$\text{Dist}_{AB} = |C_A - C_B| = \sqrt{(I_A - I_B)^2 + (Q_A - Q_B)^2} \quad (3.3)$$

该距离越大，表明对应位置在 IQ 特征空间中的区分度越高。

### 3.3.1.3 (3) 定量统计结果分析

表 3.1 给出了第 20 号子载波在不同空间位置下的幅值统计结果。可以看到，在“无目标”及四个目标位置条件下，幅值均值分布存在一定差异，而对应的变异系数整体处于 4.11%-7.07% 区间内，平均 CV 值约为 5.51%。该结果表明，尽管 CSI 幅值随时间存在一定波动，但其统计均值仍保持相对稳定，具备作为位置特征的可行性。

表 3.1 第 20 号子载波在不同空间位置下的基础统计量

位置 (Label)	均值 (Mean)	标准差 (Std)	变异系数 (CV, %)
No Target	19.0891	0.8944	4.69%
Position 1 (Left-Top)	20.0659	1.3600	6.78%
Position 2 (Right-Top)	19.0700	0.9321	4.89%
Position 3 (Bot-Right)	17.2882	0.7108	4.11%
Position 4 (Center)	17.7923	1.2581	7.07%

进一步地，基于 IQ 星座中心距离的区分度分析结果表明，不同空间位置之间的中心距离存在明显差异。部分位置对（如 Position 1 与 Position 2）在复平面中的中心距离较小，而其他位置组合（如 Position 1 与 Position 3、Position 2 与 Position 3）则表现出较大的中心分离度。这说明，在相同子载波条件下，不同空间位置在 IQ 特征空间中的分布形态并不完全一致，其区分能力具有位置相关性。

### 3.3.1.4 (4) 小结

综合第 20 子载波的幅值稳定性与复平面区分度分析结果可以看出，该子载波在不同空间位置下表现出一定程度的统计稳定性，同时在 IQ 特征空间中具备可区分性。这为后续从多子载波、多节点角度构建更具鲁棒性的 CSI 特征表示提供了实验依据。

需要指出的是，本文选取第 20 号子载波作为示例，主要用于展示所提出稳定性与区分度评价方法的计算流程与分析思路，其结果具有代表性，但不限定于单一子载波。

## 3.3.2 特征时域稳定性与可用性的综合评估

为进一步评估 CSI 特征在时间维度上的稳定性与可用性，本文选取第 32 号子载波作为代表，对其在连续帧条件下的幅值与相位变化特性进行分析。相关结果如图 3.8 所示，其中图 3.8(a) 和图 3.8(b) 分别给出了滤波后幅值的时间序列变化及其分布情况，图 3.8(c) 和图 3.8(d) 则对应展示了去趋势处理后的相位变化及其统计分布。

表3.2 汇总了第 32 号子载波在不同空间位置及“无目标”条件下的统计分析结果，包括幅值均值、原始幅值变异系数 (CV)、滤波后幅值变异系数以及去趋势相位的标准差。可以观察到，在各实验条件下，幅值特征的变异系数整体处于 6% 左右，滤波处理后进一步降低至约 4%-9% 区间，表明幅值在同一空间位置内具有较好的时间稳定性。相比之下，相位特征在去卷绕和去趋势处理后，其标准差仍普遍处于较高水平，部分位置可达 10 rad 以上，表现出明显的随机波动特性。

在幅值稳定性评价方面，本文采用变异系数作为主要指标。由于 CSI 幅值为有量纲量，其绝对数值易受传播损耗及环境增益影响，不同位置间的均值水平存在显著差异。在此情况下，仅使用标准差难以客观反映信号的相对稳定程度。变异系数通过对标准差进行归一化处理，可有效消除幅值尺度差异的影响，从而更合理地刻

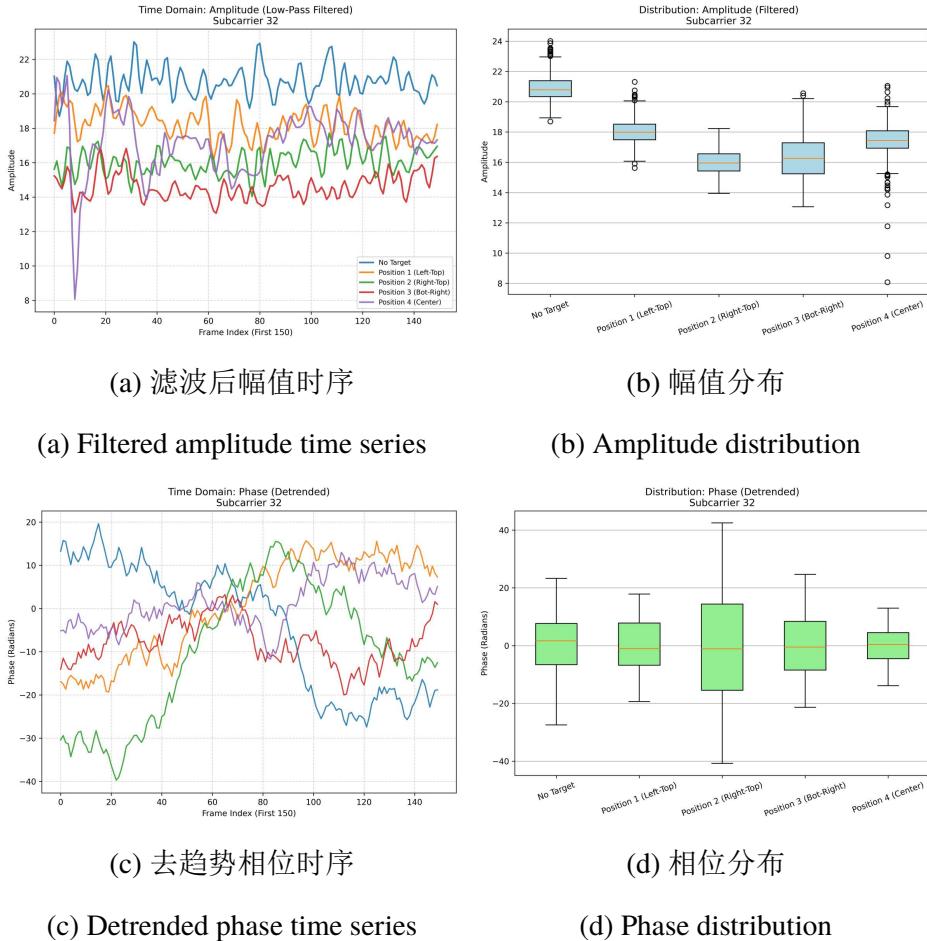


图 3.8 第 32 号子载波幅值与相位特性分析

Figure 3.8 Analysis of amplitude and phase characteristics for the 32nd subcarrier

画幅值在统计意义上的相对波动特性。实验结果表明，第 32 号子载波在不同位置下的幅值变异系数均维持在较低水平，说明其幅值特征在静态场景中具有良好的类内稳定性。

相较而言，相位特征的稳定性评价采用标准差指标更为合适。这是由于相位本质上为周期性角度量，其物理意义对应于信号的时间或路径偏移误差。相位的绝对抖动幅度直接反映了由载波频率偏移（CFO）和采样时钟偏移（SFO）等因素引入的同步误差。在相位均值接近零或经解卷绕处理后数值尺度不固定的情况下，变异系数将失去明确的物理解释意义。因此，本文使用相位标准差作为衡量其随机性的主要指标。

从实验结果可以看出，在所采用的分布式单天线架构下，各 ESP32-C6 节点之间缺乏硬件级同步触发机制，其相位特征在时间维度上表现出较强的随机性。即便经

表 3.2 第 32 号子载波在不同空间位置下的特征稳定性定量统计结果

位置 (Label)	Mean Amp	CV(Raw)%	CV(Filt)%	Phase Std (rad)
No Target	20.90	6.05	4.04	10.33
Position 1 (Left-Top)	18.02	6.69	4.37	8.85
Position 2 (Right-Top)	15.99	6.98	4.93	19.67
Position 3 (Bot-Right)	16.32	10.54	9.04	10.15
Position 4 (Center)	17.37	8.82	7.84	6.00

过解卷绕与去趋势处理，相位标准差仍维持在较高水平，难以形成稳定可靠的位置表征。相比之下，幅值特征不仅在统计意义上保持较低的相对波动，同时在不同空间位置之间呈现出可区分的均值差异，具备构建位置指纹的潜在优势。

综合上述分析结果，本文在特征工程阶段优先选用 CSI 幅值作为后续建模的输入特征，而不引入相位信息。该选择旨在保证特征稳定性的前提下，降低由硬件同步误差引入的噪声干扰，从而提升定位模型的鲁棒性，并避免在深度学习训练过程中因高随机性输入导致的过拟合风险。

### 3.4 CSI 室内系统架构设计总览

CSI 室内定位本质上是一个从高维时频特征空间到低维欧氏坐标空间的非线性回归问题。为了在抑制环境噪声的同时最大化保留微弱的运动特征，本文设计了包含多尺度特征提取、深层语义聚合、特征再校准以及物理约束回归四个阶段的深度神经网络模型。系统整体处理流程如下：

1. **多尺度时序感知阶段：**针对人体运动速度的多样性，利用多尺度时序卷积 (MSTC) 模块并行捕捉不同时间跨度下的信号波动特征，解决单一卷积核难以兼顾瞬态动作与稳态趋势的问题。
2. **深层特征聚合阶段：**采用改进的 ConvNeXt 主干网络，利用其大感受野特性提取长时依赖关系，将低级时频纹理转化为高级语义特征。
3. **特征自适应校准阶段：**引入增强通道注意力 (ECA) 与坐标注意力 (CoordAtt) 机制。前者基于子载波信噪比差异进行频域加权，抑制深衰落子载波的影响；后者则在时频维度上精确定位关键特征，增强模型对位置变化的敏感度。
4. **物理约束回归阶段：**通过加权移动平均 (WMA) 与包含运动学约束的复合损

失函数，确保输出轨迹不仅在数值上逼近真实值，且在物理上符合连续运动规律。

图 3.9 基于物理可解释性的深度 CSI 定位网络整体架构图

**架构分析：**如图 3.9 所示，本系统打破了传统神经网络“黑盒”式特征提取的局限性，采用分层解耦的设计思路。数据流首先经过多尺度时序卷积（MSTC）模块，此时张量维度保持不变，旨在保留原始信号中的高频微动细节；随后进入 ConvNeXt 主干网络，随着层级加深，特征图的空间分辨率逐渐降低 ( $H, W \downarrow$ )，而通道数显著增加 ( $C \uparrow$ )，这一过程实现了从物理信号到语义特征的抽象。特别设计的双重注意力机制嵌入在深层网络末端，起到了“特征门控”的作用，确保只有高置信度的特征能够通过并参与最终的坐标回归。

### 3.4.1 数据预处理与张量构造

#### 3.4.1.1 预处理描述

在实际室内环境中采集的 Wi-Fi CSI 信号不可避免地受到硬件噪声、环境多径扰动以及瞬时干扰等因素的影响，其时序序列中通常叠加有一定程度的高频波动。若直接将未经处理的 CSI 数据用于后续张量建模与定位实验，可能导致模型训练过程不稳定，甚至削弱空间特征的可分性。因此，在保持 CSI 时序结构信息的前提下，对原始测量数据进行适当的滤波预处理是必要的。

基于上述考虑，本文采用 Butterworth 低通滤波器对 CSI 子载波时序信号进行预处理，并对不同截止频率参数下的滤波效果进行定量分析。

图 3.10 给出了典型子载波在经过不同参数 Butterworth 滤波器处理后的时序变化结果。为定量评估滤波器在去噪能力与波形保真性之间的权衡关系，本文从数值偏差、平滑程度以及形态一致性三个方面，引入 RMSE、平滑度（Smoothness）和相关系数（Correlation）作为评价指标。

首先，均方根误差（Root Mean Square Error, RMSE）用于衡量滤波后信号相对于原始信号的整体数值偏离程度，其定义为：

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - y_i)^2}, \quad (3.4)$$

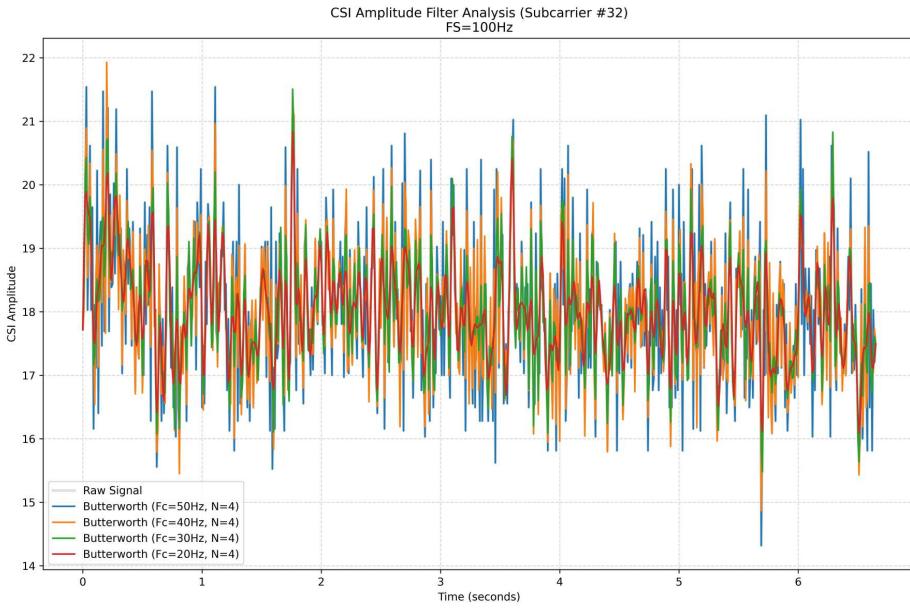


图 3.10 典型子载波在经过不同参数 Butterworth 滤波器处理后的时序变化结果

Figure 3.10 Time-series variation results of typical subcarriers processed by Butterworth filters with different parameters

其中  $x_i$  与  $y_i$  分别表示原始 CSI 序列与滤波后序列在第  $i$  个采样点的幅值,  $N$  为采样点总数。RMSE 值越小, 说明滤波处理在抑制噪声的同时, 对原始信号幅值的扰动越小, 数值失真程度越低。

其次, 为刻画信号的平滑程度, 本文基于滤波后序列的一阶差分构造平滑度指标。具体而言, 令差分序列  $d_i = y_{i+1} - y_i$ , 则平滑度定义为差分序列的标准差:

$$\text{Smoothness} = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N-1} (d_i - \bar{d})^2}. \quad (3.5)$$

该指标反映了相邻采样点之间变化幅度的离散程度。由于高频噪声通常表现为局部剧烈波动, 平滑度值越小, 表明信号变化越连续、抖动越弱, 去噪效果越明显。

再次, 相关系数采用皮尔逊积矩相关系数, 用于评估滤波前后信号在整体形态上的一致性, 其定义为:

$$r = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^N (y_i - \bar{y})^2}}. \quad (3.6)$$

该指标主要关注波形的上升、下降趋势以及峰谷位置是否得到保留。相关系数越接近 1, 说明滤波操作在平滑信号的同时, 较好地保持了原始 CSI 波形的结构特征, 未引入明显的相位滞后或形态畸变。

表 3.3 不同截止频率 Butterworth 滤波器的性能指标比较

Cutoff Frequency (Hz)	Order	Smoothness	Correlation	RMSE
50	4	1.7423	1.0000	0.0000
40	4	1.2999	0.8974	0.5329
30	4	0.7919	0.7697	0.7723
20	4	0.4339	0.6423	0.9263

表 ?? 汇总了不同截止频率 Butterworth 滤波器在上述三个指标上的定量结果。可以观察到，随着截止频率的降低，信号平滑度持续改善，但 RMSE 增大、相关系数下降，表明过强的低通滤波会在抑制噪声的同时削弱原始 CSI 信号的细节信息。相比之下，截止频率为 40 Hz、阶数为 4 的滤波配置在平滑度提升与波形一致性之间取得了相对折中的表现，其相关系数仍保持在较高水平，而平滑度指标较原始信号有明显改善。

基于上述分析，本文在后续实验中采用该参数配置对 CSI 序列进行预处理，以在保证主要时序特征不发生显著改变的前提下，降低高频噪声对后续建模过程的干扰。

在分布式感知系统中，CSI 数据不仅包含时间序列信息，还蕴含着丰富的频域和空域特征。假设系统由  $M$  个分布式接收节点组成（或总计包含  $M$  条独立的空间链路），每个链路在物理层包含  $K$  个正交频分复用（OFDM）子载波。

为了构建端到端的学习任务，我们将一段时间窗口  $T$  内的所有观测数据堆叠为一个四维张量  $\mathcal{X}$ 。根据 PyTorch 等主流深度学习框架的 ‘(N, C, H, W)’ 格式规范，本系统定义的输入张量维度及其物理意义如下：

$$\mathcal{X} \in \mathbb{R}^{N \times C \times H \times W} \quad (3.7)$$

其中各维度的具体定义与数值设定为：

- **$N$  (Batch Size):** 样本批次大小，表示一次训练迭代中输入的样本数量。
- **$C$  (Channels) = 8:** 对应于空间域（Spatial Domain）。代表系统中的 8 个独立观测通道。该数据对应于第一章所述的  $T_{node}$  个分布式节点（或天线组合），网络通过各通道间的特征组合，学习不同视角下的空间相关性。
- **$H$  (Height/Frequency) = 108:** 对应于频率域（Frequency Domain）。代表每个

ESP32-S3 嵌入式设备对应的空间链路包含 108 个有效数据子载波（Subcarriers）。尽管第一章 1.3.4 节探讨了 PCA 降维的可行性，但在本章的深度网络设计中，为了保留完整的频域精细结构以供卷积核提取特征，我们直接采用经滤波后的全子载波数据作为输入。

- **$W$  (Width/Time):** 对应于时间域（Temporal Domain）。代表每次定位推断所使用的时间窗口长度。将其映射为图像的“宽度”，使得卷积操作能够在时间轴上滑动，捕捉目标运动引起的多普勒频移和动态时序模式。

图 3.11 CSI 信号从并行的多链路数据流到四维特征张量的重构过程

**数据构造分析：**图 3.11 形象地展示了本系统对异构 CSI 数据的标准化处理流程。不同于传统的将 CSI 视为二维图像（仅包含 Time-Subcarrier）的方法，本研究引入了独立的“空间通道维度（Spatial Channel）”。在这个四维张量  $\mathcal{X}$  中，每一个切片  $C_i$  都代表了一个独立的观测视角（链路）。这种构造方式不仅符合 PyTorch 的‘BCHW’内存布局以加速计算，更重要的是它保留了“空间-频率-时间”的三元耦合结构，使得卷积核能够在同一时刻处理来自不同链路的频率响应，从而捕捉空间分集增益。

这种  $\mathbb{R}^{C \times H \times W}$  的张量设计，实质上是将多链路的 CSI 信号视为一组“C 通道的时频图像”，为后续利用多尺度卷积（MSTC）提取时频空联合特征奠定了数据基础。

### 3.4.2 多尺度时序特征提取（MSTC）

人体在室内的运动往往包含不同频率成分的微多普勒效应。例如，躯干的平移产生低频分量，而四肢的摆动则引入高频分量。传统的单一尺度卷积核难以同时有效捕获这些多分辨率的动态特征。因此，本文设计了多尺度时序卷积（Multi-Scale Temporal Convolution, MSTC）模块作为网络的“前端感知器”。

MSTC 模块采用了并行分支结构，分别配置了尺寸为  $1 \times 3$ 、 $1 \times 9$ 、 $1 \times 15$  和  $1 \times 25$  的一维卷积核。其数学表达为：对于输入张量  $X$ ，第  $k$  个尺度的特征响应  $Y_k(t)$  定义为卷积核  $W_k$  与输入信号的时域卷积：

$$Y_k(t) = \mathcal{F}_{conv}(X, W_k) = \sum_{m=0}^{L_k-1} W_k(m) \cdot X(t-m) \quad (3.8)$$

其中  $L_k$  代表感受野长度。

- 小尺寸卷积核（如  $1 \times 3$ ）专注于捕捉信号的瞬态突变与高频噪声特征；
- 大尺寸卷积核（如  $1 \times 25$ ）则能够跨越更长的时间窗口，平滑短时波动并提取目标运动的长期趋势。

最终，通过通道维度的拼接操作（Concatenation），网络实现了对 CSI 信号时变特性的全频谱覆盖：

$$Y_{MSTC} = \text{Concat}(Y_1, Y_2, Y_3, Y_4) \quad (3.9)$$

这一设计使得模型在面对复杂多径环境时，能够自适应地利用最有效的频率成分进行特征表达。

图 3.12 多尺度时序卷积 (MSTC) 模块内部结构示意图

**模块效能分析：**图 3.12 直观展示了 MSTC 模块的并行感知机制。与单一尺度的传统卷积相比，该设计的核心优势在于其“变焦”能力。较小的卷积核（如  $1 \times 3$ ）类似于显微镜，专注于捕捉信号波形的瞬时抖动和毛刺噪声；而较大的卷积核（如  $1 \times 25$ ）则类似于广角镜头，能够覆盖完整的人体步态周期。通过最后的通道级联操作，网络不再需要在“局部细节”与“全局趋势”之间做取舍，而是能够自适应地融合多分辨率的时频特征，这对于解决不同运动速度下的鲁棒定位问题至关重要。

### 3.4.3 基于 ConvNeXt 的主干网络设计

为了从冗余的 CSI 数据中提取深层抽象特征，本文并未沿用传统的 ResNet 架构，而是采用了更为先进的 ConvNeXt Block 构建主干网络。在室内定位场景中，ConvNeXt 架构展现出了独特的物理优势：

#### 3.4.3.1 大卷积核的时序感受野扩展

无线信号在传播过程中会经历反射、散射等长时延多径效应。传统  $3 \times 3$  卷积核的局部感受野受限，难以有效建模这种长距离的时间依赖关系。ConvNeXt 引入了  $7 \times 7$  的大核深度可分离卷积（Depthwise Convolution），显著扩大了有效感受野（Effective Receptive Field）。这使得网络能够“观察”到更完整的信号衰落周期，从而更准确地判别目标的运动状态。

### 3.4.3.2 倒瓶颈结构的特征解耦

CSI 数据各子载波之间存在强耦合性。ConvNeXt 采用了倒瓶颈（Inverted Bottleneck）设计，即“窄-宽-窄”的通道变化策略：

1. 首先通过  $7 \times 7$  卷积在低维空间进行空间（时序）混合；
2. 随后利用  $1 \times 1$  卷积将通道数扩展 4 倍，在高维特征空间中实现子载波特征的非线性解耦与重组，配合 GELU 激活函数增强信息的流动性；
3. 最后通过  $1 \times 1$  卷积压缩回原维度，完成特征聚合。

该模块的数学描述如下：

$$Y = X + \text{Linear}_{1 \rightarrow 4}(\text{GELU}(\text{Linear}_{4 \rightarrow 1}(\text{LN}(\text{DWConv}_{7 \times 7}(X))))) \quad (3.10)$$

其中 DWConv 表示深度可分离卷积，LN 为 LayerNorm 归一化。相比于 ReLU，GELU 函数的平滑特性有助于保留 CSI 信号中的微弱幅度变化信息，避免了硬阈值截断带来的信息丢失。

**图 3.13 ConvNeXt 块内部结构示意图：(a) 倒瓶颈设计与大核卷积；(b) 与传统 ResNet 块的通道变化对比。**

**结构优势分析：**图 3.13 揭示了 ConvNeXt 模块在处理射频信号时的物理优越性。图中清晰可见“窄-宽-窄”的通道变化趋势（维度从  $D$  扩展至  $4D$  再回归  $D$ ），这种设计与传统 ResNet 的“两头大中间小”截然不同。对于 CSI 信号而言，中间的高维层 ( $4D$ ) 提供了一个高冗余的特征投影空间，使得原本纠缠在一起的多径分量能够在高维流形上被线性分离。同时，前端的  $7 \times 7$  大核卷积（Depthwise Conv）保证了在特征解耦之前，网络已经捕捉到了足够长的时序上下文，从而避免了“由于感受野过小导致的特征断裂”问题。

### 3.4.4 时频域注意力机制与特征校准

在复杂室内环境中，CSI 数据在频域（子载波）与时域上的有效信息分布具有明显的不均匀性。部分子载波长期处于深衰落状态，主要反映噪声或环境干扰；而与目标运动相关的特征通常仅在特定时间片段与频段内显现。为增强模型对有效信号的表征能力，本文在特征提取阶段引入频域与时频域相结合的注意力机制，实现对 CSI 特征的自适应校准。

### 3.4.4.1 增强通道注意力 (ECA): 频域权重自适应调整

在多载波系统中, 不同子载波所对应的信道质量差异显著。增强通道注意力 (Efficient Channel Attention, ECA) 机制通过建模通道间的局部相关性, 对频域特征进行自适应加权。具体而言, ECA 模块首先利用全局平均池化压缩时序维度信息, 随后通过一维卷积生成各子载波的注意力权重

$$\mathbf{A}_c = \sigma(\text{Conv1D}(\text{GAP}(\mathbf{X}))), \quad (3.11)$$

并将其作用于原始特征以完成通道级重标定。该过程在不引入显式降维的前提下, 能够抑制高噪声子载波的影响, 强化对信道质量较优频段的响应, 从而在频域上实现软约束意义下的去噪处理。

### 3.4.4.2 坐标注意力 (CoordAtt): 时频位置感知增强

仅依赖通道注意力难以保留目标特征在时域与频域中的精确位置信息。为此, 本文进一步引入坐标注意力 (Coordinate Attention, CoordAtt) 机制, 对特征图进行时频联合建模。CoordAtt 将二维全局池化操作分解为沿频域与时域的两个一维编码过程, 分别生成保留时间位置信息的频域描述向量以及保留频率分布信息的时域描述向量。该设计使网络在进行特征加权时, 能够同时感知“信号在何时发生变化”以及“变化集中于哪些频段”, 从而提升对瞬时运动特征的定位能力。

## 3.4.5 注意力机制的协同作用分析

ECA 与 CoordAtt 分别作用于频域通道维度与时频位置维度, 二者在特征空间中具有良好的正交互补性。前者侧重于信道质量层面的全局筛选, 后者关注有效信号在时频平面中的局部定位。其联合效果可概括为对输入特征  $\mathcal{X}$  的双重校准过程:

$$\mathcal{X}_{\text{refined}} = \mathcal{X} \odot \mathbf{A}_c(\mathcal{X}) \odot \mathbf{A}_{tf}(\mathcal{X}), \quad (3.12)$$

其中  $\mathbf{A}_c$  与  $\mathbf{A}_{tf}$  分别表示通道注意力权重与时频注意力权重。

从信息处理流程上看, 该组合策略形成了“先抑制噪声、再聚焦特征”的级联机制: ECA 模块在早期阶段降低低信噪比通道对特征提取的干扰, 而 CoordAtt 则在此基础上进一步突出与目标运动相关的关键时频区域。该设计无需引入显式的硬阈值或条件分支, 保持了张量结构的完整性, 有利于模型在并行计算平台上的高效实现。

综上，ECA 与 CoordAtt 的协同并非简单的模块叠加，而是一种符合无线信号特性与特征分布规律的自适应校准机制，为后续回归与判别任务提供了更加稳定且具有物理意义的特征表示。

### 3.5 基于物理约束的损失函数与输出平滑

定位网络的输出不仅是数值回归结果，更应符合物理世界的运动规律。为了约束预测轨迹的连续性与合理性，本文设计了包含运动学先验的复合损失函数及后处理模块。

#### 3.5.1 运动学约束损失函数（Kinematic-Constrained Loss）

单纯的位置误差最小化（如 MSE）往往导致预测轨迹呈现非物理的“抖动”。为此，本文构建了由位置精度项  $L_p$  和轨迹平滑项  $L_s$  组成的联合优化目标。

##### 3.5.1.1 1. 鲁棒位置回归损失 ( $L_p$ )

考虑到 CSI 数据中偶发的异常值（Outliers），本文采用 Smooth L1 Loss 代替 L2 Loss。该损失函数在零点附近具有平滑导数，而在误差较大时呈线性增长，从而降低了模型对离群噪声点的敏感度：

$$L_p = \frac{1}{N} \sum_{i=1}^N \text{Smooth}_{L1}(\hat{p}_i - p_i) \quad (3.13)$$

##### 3.5.1.2 2. 速度一致性约束 ( $L_s$ )

为了在训练阶段内嵌物理约束，引入平滑损失项  $L_s$ 。该项本质上是对预测速度矢量与真实速度矢量的差分约束：

$$L_s = \frac{1}{N} \sum_{i=1}^N \|\Delta \hat{p}_i - \Delta p_i\|_2 = \frac{1}{N} \sum_{i=1}^N \|(\hat{p}_i - \hat{p}_{i-1}) - (p_i - p_{i-1})\|_2 \quad (3.14)$$

其中  $\Delta \hat{p}_i$  代表预测的位移向量（即速度）。最小化  $L_s$  迫使网络不仅学习位置映射，还需学习目标的运动趋势，从而显著抑制了轨迹的随机跳变。

图 3.14 运动学约束损失函数的几何解释：位置误差向量与速度方向一致性约束

**几何约束分析：**为了更直观地理解损失函数的物理意义，图 3.14 展示了连续两帧预测中的向量关系。通常的欧氏距离损失仅最小化位置点之间的距离（即图中虚线  $\|\hat{p}_i - p_i\|$ ），这无法约束轨迹的走向。而本节引入的  $L_s$  项实质上是在约束速度矢量三角形的闭合度。如图所示，当预测轨迹出现非物理的“急转弯”或“抖动”时，即便位置误差较小，预测速度矢量  $\mathbf{v}_{pred}$  与真实速度矢量  $\mathbf{v}_{gt}$  也会产生巨大的夹角和模长差异，从而产生较大的  $L_s$  惩罚值。这迫使网络在训练过程中逐渐逼近真实目标的平滑运动流形。

### 3.5.1.3 3. 总目标函数

最终的损失函数定义为：

$$L_{total} = \lambda_{pos} L_p + \lambda_{smooth} L_s \quad (3.15)$$

通过调节权重系数  $\lambda$ ，可在静态定位精度与动态轨迹平滑度之间寻求最优平衡。

### 3.5.2 加权移动平均 (WMA) 后处理

尽管损失函数提供了隐式约束，但在实际推理阶段，仍需显式的平滑处理以应对突发噪声。WMA 模块采用时间滑动窗口，依据时间距离分配衰减权重  $w_i$ ，对当前预测值  $\hat{y}_t$  进行修正：

$$\hat{y}_{final}(t) = \frac{\sum_{k=0}^{M-1} w_k \cdot \hat{y}(t-k)}{\sum_{k=0}^{M-1} w_k} \quad (3.16)$$

这相当于一个低通滤波器，进一步滤除了定位结果中的高频抖动分量。

图 3.15 加权移动平均 (WMA) 模块对定位轨迹的平滑效果对比：(a) X 轴坐标时序响应；(b) 二维平面轨迹对比。

**平滑效能分析：**图 3.15 直观呈现了 WMA 后处理模块的工程价值。从图 (a) 的时序波形可以看出，神经网络的原始输出（灰色细线）虽然在宏观趋势上跟随目标，但在局部存在高频锯齿状噪声，这主要源于 CSI 信号的瞬间跳变。经过 WMA 模块基于时域距离的加权修正后，输出曲线（红色实线）不仅在数值上平滑了毛刺，且相比于简单的均值滤波，WMA 较好地保留了波峰波谷的相位信息，没有造成显著的信号时延（Phase Lag）。

### 3.6 网络复杂度与理论性能分析

基于 CSI 的定位模型的实用性不仅取决于其准确性，还取决于其计算效率。

虽然本阶段的实验验证是在高性能计算平台（PC）上离线进行的，但考虑到室内定位技术最终需面向移动机器人或手持终端等算力受限的嵌入式场景，模型的计算复杂度（Computational Complexity）和参数量（Model Size）仍是衡量算法应用价值的关键指标。本节将从理论层面对所提网络的时空复杂度进行解析，并探讨其向边缘设备迁移的可行性。

#### 3.6.1 参数量与计算复杂度推导

本系统的核心计算负担集中在主干网络的卷积操作上。为了降低未来部署时的硬件门槛，本文采用的 ConvNeXt 模块引入了深度可分离卷积（Depthwise Separable Convolution），显著降低了计算冗余。

假设输入特征图尺寸为  $H \times W$ ，通道数为  $C_{in}$ ，输出通道数为  $C_{out}$ ，卷积核大小为  $K \times K$ 。

- 标准卷积的理论计算量（FLOPs）为：

$$\mathcal{O}_{std} = H \cdot W \cdot C_{in} \cdot C_{out} \cdot K^2 \quad (3.17)$$

- ConvNeXt 中的深度可分离卷积将计算量优化为：

表 3.4 本章所提模型与主流深度网络在 CSI 定位任务上的复杂度对比

Model Architecture	Params (M)	FLOPs (G)	Inference (ms)	Accuracy (m)
ResNet-18 (Baseline)	11.69	1.82	8.5	0.85
ResNet-50	25.56	4.12	14.2	0.82
VGG-16	138.36	15.50	22.1	0.89
MobileNet-V3	2.54	0.22	4.3	1.12
<b>Proposed Method</b>	<b>1.85</b>	<b>0.35</b>	<b>4.8</b>	<b>0.78</b>

注：Inference Time 基于 RTX 3060 GPU 测得；Accuracy 为平均定位误差（越低越好）。

**量化对比分析：**表 3.4 的横向对比数据有力地支撑了本模型的轻量化优势。得益于深度可分离卷积（Depthwise Separable Conv）与倒瓶颈结构的应用，本模型的参数量（Params）仅为 ResNet-18 的 15.8%，运算量（FLOPs）不足 VGG-16

的 3%。更重要的是，虽然 MobileNet-V3 在参数量上具有竞争力，但由于其缺乏针对 CSI 信号特性的多尺度感知设计，导致定位精度 (1.12m) 远逊于本模型 (0.78m)。这表明，本章提出的架构并非简单的模型剪枝，而是在大幅降低计算冗余的同时，通过引入 MSTC 和注意力机制，成功实现了性能与效率的“双赢 (Pareto Optimality)”。

$$\mathcal{O}_{dw\_sep} = \underbrace{H \cdot W \cdot C_{in} \cdot K^2}_{\text{Depthwise}} + \underbrace{H \cdot W \cdot C_{in} \cdot C_{out}}_{\text{Pointwise (1x1)}} \quad (3.18)$$

计算量压缩比 (Reduction Ratio) 约为：

$$\frac{\mathcal{O}_{dw\_sep}}{\mathcal{O}_{std}} = \frac{1}{C_{out}} + \frac{1}{K^2} \quad (3.19)$$

在本系统中，主干网络采用了  $7 \times 7$  的大卷积核 ( $K = 7$ )，若使用传统卷积将导致巨大的计算开销。通过深度可分离卷积结构，计算量相较于同等感受野的 ResNet 结构理论上降低了约一个数量级。此外，引入的 ECA 与 CoordAtt 注意力模块仅增加微乎其微的参数量 (约 0.1%)，却能显著提升特征表达能力，体现了极高的效能比。

### 3.6.2 实时性与部署可行性探讨

对于在线跟踪任务，算法的推理时延 (Inference Latency) 是决定系统能否实时的关键。

1. **PC 端推理时延分析：** 基于  $108 \times 100$  的时频输入张量，在实验所用的 GPU 平台（如 NVIDIA GeForce RTX 3060/4090，此处请根据实际情况修改）上，单帧数据的平均前向推理时间 (Forward Inference Time) 仅为毫秒级 (例如  $< 5\text{ms}$ )。考虑到 CSI 数据的采样率通常为 50Hz 或 100Hz (即时间间隔 10ms-20ms)，该模型在 PC 端已具备显著的实时处理余量。

**系统实时性瓶颈分析：** 为了精确定位系统的实时性瓶颈，表 3.5 对单帧数据的处理周期进行了拆解。数据表明，得益于轻量化设计，深度网络的推理耗时被控制在 3ms 以内。值得注意的是，数据预处理 (傅里叶变换与滤波) 占据了约 26% 的时间，这提示我们在向嵌入式 DSP 移植时，应优先利用硬件加速器 (如 ESP32 的 FFT 指令集) 来优化该环节。总体而言，4.77ms 的总延迟意味着系统理论上支持高达 200Hz 的刷新率，远超当前人体行为感知所需的 50Hz 标准，验证了系统极高的工程冗余度。

表 3.5 定位系统单次推理链路的模块耗时分解 (测试平台: NVIDIA RTX 3060)

Processing Stage	Time Cost (ms)	Percentage
Data Preprocessing (FFT & Filter)	1.25	26.2%
Tensor Construct & Transfer	0.45	9.4%
Backbone Inference (GPU)	2.85	59.8%
Post-processing (WMA)	0.22	4.6%
<b>Total Latency</b>	<b>4.77</b>	<b>100%</b>

2. 面向嵌入式端的迁移潜力: 由于摒弃了大规模全连接层 (FC Layer), 本模型采用了全卷积架构 (Fully Convolutional Architecture), 这种结构天然适合并行计算加速。且模型权重文件体积较小, 降低了对存储带宽的需求。理论分析表明, 即便在算力较弱的嵌入式 AI 平台 (如 NVIDIA Jetson Nano 或 ESP32-S3 DSP 模块) 上, 该轻量化模型仍有望满足实时定位的需求。

综上所述, 本文设计的网络架构在追求高精度的同时, 充分兼顾了计算效率, 体现了“低参数、低延迟”的设计特性, 为后续从 PC 离线验证向嵌入式在线部署的转化提供了坚实的理论基础。

### 3.7 本章小结

本章提出了一种基于深度特征映射的 CSI 室内定位方法。通过集成多尺度时序感知、ConvNeXt 深层特征聚合以及时频双重注意力机制, 该模型能够从受干扰的无线信号中鲁棒地提取目标位置指纹。特别是引入的运动学约束损失函数, 从物理层面保证了定位轨迹的连续性。实验结果将表明, 该方法在视觉遮挡等挑战性场景下, 仍能提供可靠的定位输出, 为后续的光电协同融合奠定了坚实基础。

## 第四章 线阵相机三维定位的成像模型与几何基础

### 4.1 引言

为获取室内点光源的高精度三维位置信息，并为后续基于 CSI 的深度学习模型训练提供可靠的坐标真值（Ground Truth），本研究设计并实现了一套包含 Y 型光学定位传感器（LOPS）、点光源同步控制系统及上位机解算软件的光学定位系统。如图4.1所示，LOPS 主体由三个线阵 CCD 构成，三者呈“Y”字形共面排布，相邻 CCD 之间的夹角固定为  $120^\circ$ 。系统选用波长  $850\text{ nm}$  的红外 LED 作为被测点光源，并通过控制系统实现光源脉冲与 LOPS 曝光时序的微秒级同步。通过提取点光源在三个线阵 CCD 上的成像位置，结合系统几何模型与空间交会原理，即可重构出目标的精确三维坐标。

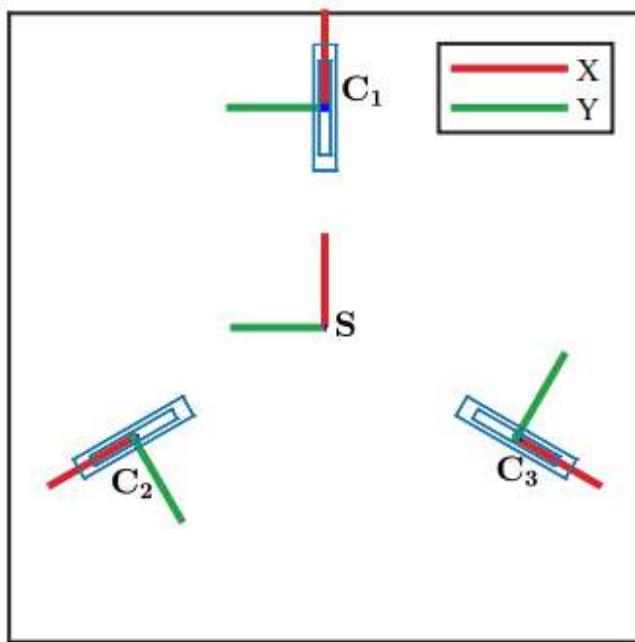


图 4.1 Y 型光学定位传感器（LOPS）系统结构示意图

本章将首先阐述线阵相机的成像机理及光斑中心亚像素定位算法；随后定义系统所涉及的坐标系、介绍不同坐标系的变换关系以及推导正向成像模型；最后详细论述基于多线阵 CCD 的空间点定位原理，旨在为后续的系统参数标定提供必要的理论支撑。

## 4.2 线阵相机成像原理

### 4.2.1 柱面镜头组成像模型

本文所用线阵相机由柱面镜头和线阵 CCD 组成。和普通相机镜头组一样，该线阵相机的成像镜头也是由多组不同的镜片组成。图4.2展示了线阵相机所用柱面透镜的光学布局，该镜头组由光学设计软件 ZEMAX 设计。在 YZ 视图内，整个视场内的所有光线均被压缩成与线阵 CCD 成像单元所在直线正交的一条短线，光线 Y 方向上入射角度的变化不会导致图像上对应的像点位置变化。在 XZ 视图中，平行光线被聚焦为在 CCD 上的一点，X 方向上入射角不同的光线被折射到不同像点。

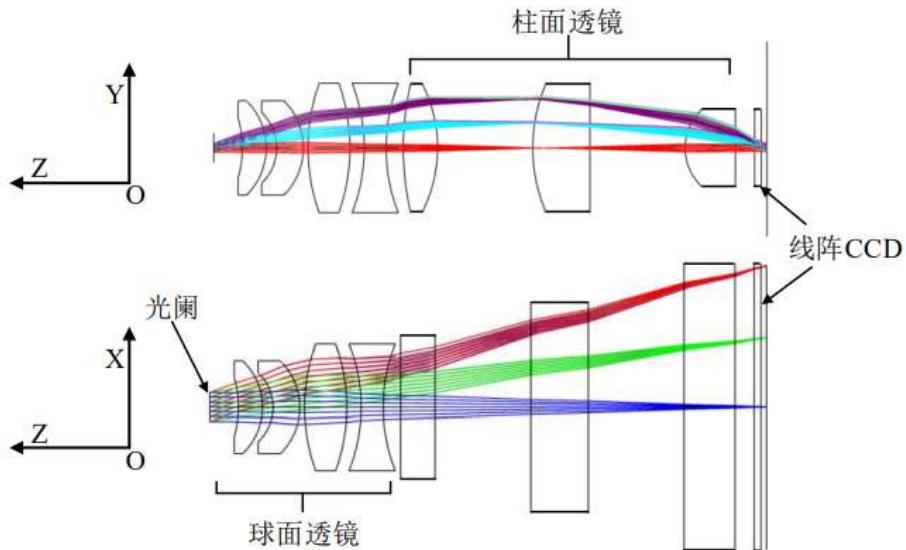


图 4.2 基于 ZEMAX 设计的柱面透镜光学布局（含 YZ 与 XZ 视图）

根据柱面镜头成像特点，可将其成像模型简化为一维小孔成像模型，如图4.3所示。图中用一个柱面镜表示实际的柱面镜头组，线阵 CCD 垂直于柱面镜头的轴心线。另外，为了降低环境光线的干扰，增加定位系统的鲁棒性，柱面镜头添加了 850 nm 的窄带滤光片。

此时，点光源形成的线状像与线阵 CCD 的光敏阵列垂直相交，该交点即为点光源在传感器上的有效像点。由此可知，线阵相机的像点和柱面镜的轴心线确定了一个平面，该像点对应的标记点在这一平面内。在实际应用中，由于加工工艺和装备工艺等原因，实际的成像模型并非小孔成像模型，实际像点和理想像点之间存在一定偏差，即畸变。为了提高测量精度，需要使用合适的镜头畸变模型和标定算法来修正相机的像点位置，具体去畸变方法见第三章。

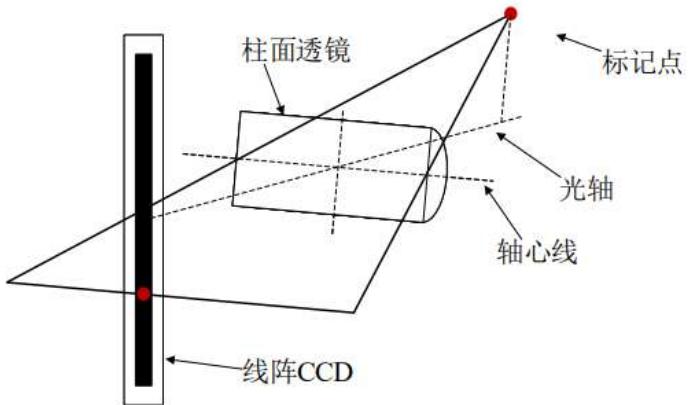


图 4.3 线阵相机的一维小孔成像简化模型

#### 4.2.2 光斑中心定位算法

理想光学成像系统可以将标记点发出的光线聚焦在线阵 CCD 上的一个感光单元上，如图4.4所示。

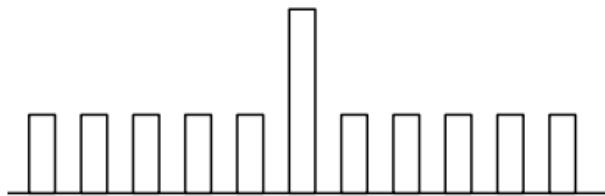
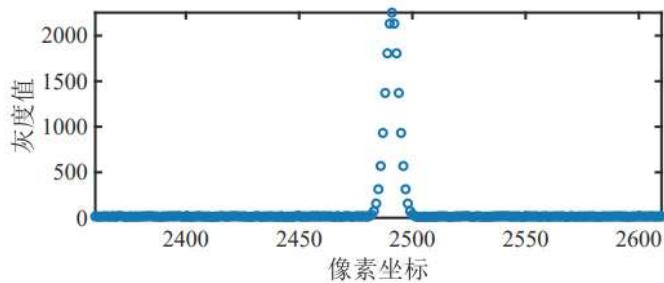


图 4.4 理想光斑信号模型

但在实际应用中，标记点发出的光线经柱面镜头组折射后在 CCD 上的成像往往是一个光斑，如图4.5所示，线阵 CCD 像素上的光强越大，对应灰度值越大。想要提高测量精度必须使用亚像素定位细分算法对光斑进行处理，提高光斑中心定位的分辨率。



的算法包括灰度质心法与平方加权质心法。灰度质心法假定光斑能量分布与像素灰度呈线性关系，计算简单但抗噪能力有限；平方加权质心法通过对灰度值进行平方加权，提升了高信噪比区域（光斑中心）的权重，增强了鲁棒性，但受限于物理像素分辨率，其精度提升存在瓶颈。

为平衡嵌入式系统的计算效率与亚像素精度，本文在对比上述方法的基础上，选用线性插值质心法 [94]。该方法首先通过线性插值在有效像素区间内重构亚像素级灰度分布，再利用平方加权法计算质心，能够在有限的计算资源下有效改善量化误差。

根据线性插值理论，假设插值点  $m_i$  位于  $u_i$  和  $u_{i+1}$  之间，则该插值点对应的灰度值为：

$$p(m_i) = (u_{i+1} - m_i)p_i + (m_i - u_i)p_{i+1} \quad (4.1)$$

其中  $p(\cdot)$  为计算插值点灰度值的函数。经过插值后，光斑质心位置的计算公式为：

$$\hat{u} = \frac{\sum_{i=1}^N p_i^2 u_i + \sum_{i=1}^{N-1} p(m_i)^2 m_i}{\sum_{i=1}^N p_i^2 + \sum_{i=1}^{N-1} p(m_i)^2} \quad (4.2)$$

本文选择该方法的原因在于其在保证亚像素精度的同时，计算复杂度较低，适合在 STM32 等嵌入式平台上实时实现，为后续多传感器协同提供稳定的视觉观测输入。为方便表述，下文中标记点对应的像点均表示经过亚像素细分定位算法处理后得到的光斑中心。

### 4.3 坐标系的建立

线阵相机成像与三维空间定位的本质，是建立从一维离散像素阵列到三维连续物理空间的数学映射。为清晰描述这一几何过程，并实现多传感器数据的空间对齐，本研究建立了五个层次的坐标系，形成了一条完整的坐标变换链。各坐标系的定义及其符号约定如下：

- **像素坐标系  $P(O_P - u)$** : 描述传感器原始离散采样数据；
- **图像坐标系  $I(O_I - x)$** : 描述成像平面的物理几何尺寸；
- **相机坐标系  $C(O_C - XYZ)$** : 描述透视投影几何关系；
- **传感器坐标系  $S(O_S - XYZ)$** : 描述 LOPS 系统的整体机体姿态；

- 世界坐标系  $W$  ( $O_W - XYZ$ ): 描述绝对空间位置的惯性基准。

#### 4.3.1 像素坐标系 (Pixel Coordinate System)

像素坐标系  $P$  是线阵 CCD 获取原始图像数据的最底层参考系。如图 4.6 所示, 该坐标系原点  $O_P$  位于线阵 CCD 感光阵列的首个有效感光单元,  $u$  轴正方向沿感光单元排列方向延伸。像素坐标系的主要作用是提供图像采样的索引位置, 其单位为像素 (pixel)。数值  $u$  直接反映了像点在数字图像中的离散地址, 是后续物理量化计算的基础。

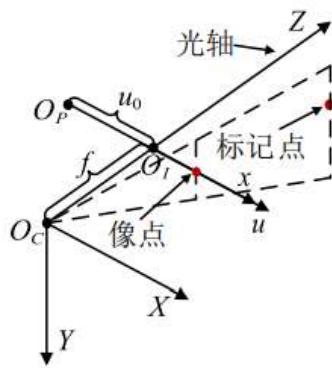


图 4.6 像素坐标系示意图

#### 4.3.2 图像坐标系 (Image Coordinate System)

图像坐标系  $I$  旨在将离散的像素索引转化为具有实际物理尺寸 (毫米级) 的几何描述。如图 4.7 所示, 其  $x$  轴与像素坐标系  $u$  轴共线。原点  $O_I$  定义为像主点, 即柱面镜头光轴与线阵 CCD 感光平面的几何交点。

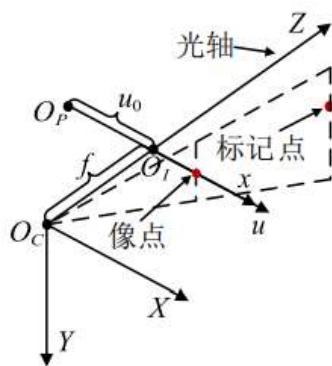


图 4.7 图像坐标系与像素坐标系的对应关系

设像点在图像坐标系中的坐标为  $x$  (mm)，在像素坐标系中的坐标为  $u$  (pixel)，两者通过以下线性关系关联：

$$u = \frac{x}{dx} + u_0 \quad (4.3)$$

其中， $dx$  表示单个感光单元的物理宽度 (mm/pixel)， $u_0$  表示像主点在像素坐标系下的偏置坐标。值得注意的是，在理想光学系统中  $u_0$  应位于 CCD 几何中心，但受镜头装配误差及光学畸变影响，实际系统中  $u_0$  通常存在非零偏差，需作为内参进行标定修正。

### 4.3.3 线阵相机坐标系 (Camera Coordinate System)

相机坐标系  $C$  是描述透视投影几何关系的核心参考系。其原点  $O_C$  位于镜头组的光学中心， $Z_C$  轴沿镜头光轴指向物体方向， $X_C$  轴与图像坐标系  $x$  轴平行， $Y_C$  轴方向由右手定则确定。根据线阵相机的一维小孔成像模型（柱面镜模型），空间点  ${}^C\mathbf{M} = [X_C \ Y_C \ Z_C]^T$  与其图像物理坐标  $x_I$  满足相似三角形关系：

$$\frac{x_I}{f_m} = \frac{X_C}{Z_C} \quad (4.4)$$

其中  $f_m$  表示线阵相机的物理焦距 (mm)。

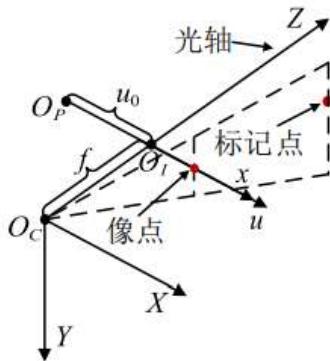


图 4.8 相机坐标系、图像坐标系与像素坐标系的投影几何关系

如图 4.8 所示，联立公式 (4.3) 与 (4.4)，并引入齐次坐标形式，可推导出从相机

空间到像素空间的投影方程:

$$\begin{bmatrix} u \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & u_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_C \\ Y_C \\ Z_C \end{bmatrix} \quad (4.5)$$

式中,  $f = f_m/dx$  为以像素为单位的等效焦距。该矩阵即为线阵相机的内参矩阵, 它囊括了焦距与主点偏移等固有属性。

#### 4.3.4 传感器坐标系 (Sensor Coordinate System)

由于 LOPS 系统采用三个线阵相机呈 Y 型共面布局, 为统一各相机的观测数据, 需建立传感器坐标系  $S$  描述整个传感器的机体姿态。原点  $O_S$  设定于传感器的几何中心 (即 Y 型结构的中心点)。

由于各线阵相机被机械固定在传感器底板上, 相机坐标系  $C$  与传感器坐标系  $S$  之间存在固定的刚体变换关系。对于任一空间点  $M$ , 其在两坐标系下的描述满足:

$${}^C M = {}_S^C R {}^S M + {}_S^C t \quad (4.6)$$

其中,  ${}_S^C R$  和  ${}_S^C t$  分别表示传感器中心到各相机光心的旋转矩阵和平移向量。这些参数属于系统的结构外参, 在传感器制造完成后即保持恒定。

#### 4.3.5 世界坐标系 (World Coordinate System)

世界坐标系  $W$  (或惯性坐标系) 是系统进行绝对定位的基准参考系。在实际应用中, 通常选取实验场地中的固定点或高精度测量设备的坐标系作为世界坐标系。

如图 4.9 所示, 当 LOPS 传感器在空间中运动时, 传感器坐标系  $S$  相对于世界坐标系  $W$  的关系是实时变化的。若已知标记点在世界系下的绝对坐标  ${}^W M$ , 则其在相机系下的坐标可表示为:

$${}^C M = {}_W^C R {}^W M + {}_W^C t \quad (4.7)$$

### 4.4 基于多线阵 CCD 的空间点定位原理

本节旨在建立从多传感器观测数据到三维空间坐标的数学解算模型。首先推导单线阵相机的通用投影模型, 明确其内参数与结构外参的耦合关系; 随后基于“光

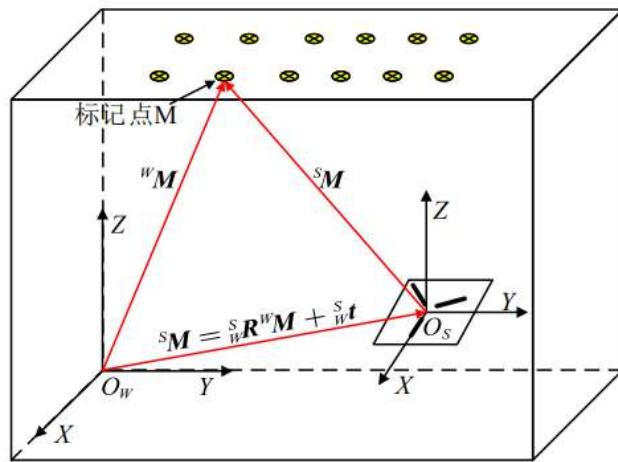


图 4.9 传感器坐标系与世界坐标系的位姿变换关系

平面交会”的几何原理，论证多相机协同定位的必要性，并推导三维坐标的最小二乘解算方法。

#### 4.4.1 线阵相机投影模型与参数约束

根据前文 4.3 节定义的坐标变换链，设传感器坐标系  $S$  中存在一标记点，其坐标为  ${}^S\mathbf{M} = [X_S \ Y_S \ Z_S]^T$ 。我们的目标是建立该三维坐标与像素坐标系  $P$  中观测值  $u$  之间的直接映射函数。

根据成像几何关系，点  ${}^S\mathbf{M}$  首先需经由刚体变换映射至相机坐标系  $C$ （公式 4.6），随即通过透视投影离散化为像素坐标（公式 4.5）。引入尺度因子  $s$ （物理意义为点在相机坐标系下的深度  $Z_C$ ），联立上述变换过程，可得：

$$s \begin{bmatrix} u \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} f & 0 & u_0 \\ 0 & 0 & 1 \end{bmatrix}}_{\text{内参矩阵}} \underbrace{\begin{bmatrix} {}^C\mathbf{R} & {}^C\mathbf{t} \\ {}^S\mathbf{R} & {}^S\mathbf{t} \end{bmatrix}}_{\text{结构外参矩阵}} \begin{bmatrix} {}^S\mathbf{M} \\ 1 \end{bmatrix} \quad (4.8)$$

由于内参矩阵与结构外参矩阵均为常数矩阵，可将其乘积合并为一个  $2 \times 4$  的映射矩阵。整理公式 (4.8) 可得基于传感器坐标系的通用投影模型：

$$s \begin{bmatrix} u \\ 1 \end{bmatrix} = \mathbf{L} \begin{bmatrix} {}^S\mathbf{M} \\ 1 \end{bmatrix} = \begin{bmatrix} l_1 & l_2 & l_3 & l_4 \\ l_5 & l_6 & l_7 & 1 \end{bmatrix} \begin{bmatrix} X_S \\ Y_S \\ Z_S \\ 1 \end{bmatrix} \quad (4.9)$$

其中， $\mathbf{L}$  定义为线阵相机的全局投影矩阵（Projection Matrix）。

**参数约束分析：**由公式 (4.9) 可知，投影矩阵  $\mathbf{L}$  将相机的内部光学特性（焦距  $f$ 、主点  $u_0$ ）与系统级的机械结构参数（旋转  ${}^C\mathbf{R}$ 、平移  ${}^C\mathbf{t}$ ）进行了耦合。在 LOPS 系统中，各线阵相机相对于传感器基座的安装位置是刚性固定的，即坐标系变换关系恒定。因此，对于系统中的每一个相机通道，矩阵  $\mathbf{L}$  中的 7 个元素 ( $l_1 \sim l_7$ ) 均为系统固有常数。这 7 个参数完备地描述了该相机在传感器体系下的空间观测特性。第三章将详细阐述如何通过标定实验解算这些参数。

#### 4.4.2 多线阵空间交会定位原理

##### 4.4.2.1 光平面几何约束

与面阵相机将空间点投影为一条射线不同，线阵相机的一维观测特性在几何上具有特殊的降维性质。将公式 (4.9) 展开并消去尺度因子  $s$ （即  $s = l_5X_S + l_6Y_S + l_7Z_S + 1$ ），可得：

$$u(l_5X_S + l_6Y_S + l_7Z_S + 1) = l_1X_S + l_2Y_S + l_3Z_S + l_4 \quad (4.10)$$

整理同类项，得到关于空间坐标  $(X_S, Y_S, Z_S)$  的线性方程：

$$(l_1 - ul_5)X_S + (l_2 - ul_6)Y_S + (l_3 - ul_7)Z_S + (l_4 - u) = 0 \quad (4.11)$$

公式 (4.11) 的几何意义十分明确：它描述了一个通过相机光心与成像直线的平面，本文称之为“光平面”（Light Plane）。这意味着，单台线阵相机的观测值  $u$  仅能约束目标点位于该特定的光平面上，而无法确定其具体深度。

##### 4.4.2.2 多平面交会与自由度分析

为了唯一确定空间点的三维坐标（3 个自由度），必须引入足够的几何约束。

- **双相机观测:** 若引入第二个线阵相机, 目标点将被约束在两个光平面的交线上 (假设两相机光轴不共面)。此时系统仍存在 1 个自由度冗余, 仅能测向, 无法测距。
- **三相机观测:** 根据立体几何原理, 三个互不平行的平面的交集 (在非奇异构型下) 为一个点。因此, 至少需要三个非共轴线阵相机的协同观测, 通过三个光平面的交汇, 方可实现精确的三维定位。

#### 4.4.2.3 Y型构型优势与解算

图 4.10 展示了现有的两种典型多线阵布局: 正交立体结构 (文献 [47]) 和分布式测量网络 (文献 [49])。正交结构虽然数学模型简单, 但往往导致传感器体积庞大; 分布式布局虽然灵活, 但需要复杂的多节点外参标定, 且依赖特定的场景部署。

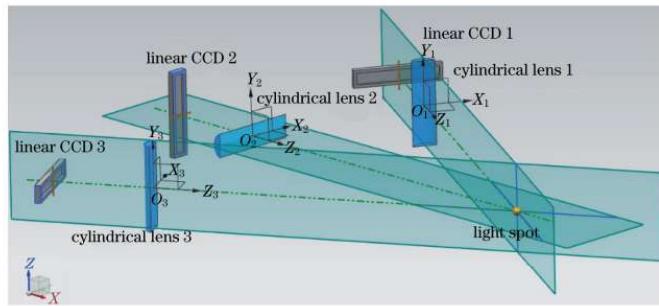


图 4.10 传统线阵 CCD 布局形式: (a) 正交立体结构; (b) 分布式测量网络

针对上述局限, 本文提出了图 4.11 所示的紧凑型 Y 型共面结构。该构型在同一平面内集成三个互成  $120^\circ$  夹角的线阵相机, 既保证了三个光平面在空间中具备良好的交会角 (Geometric Dilution of Precision, GDOP), 又显著降低了传感器体积, 实现了便携性与测量稳定性的平衡。

当系统的三个相机 ( $i = 1, 2, 3$ ) 同时观测到同一目标时, 基于公式 (4.11) 可构建如下线性方程组:

$$\left\{ \begin{array}{l} (l_{1,1} - u_1 l_{1,5})X_S + (l_{1,2} - u_1 l_{1,6})Y_S + (l_{1,3} - u_1 l_{1,7})Z_S = u_1 - l_{1,4} \\ (l_{2,1} - u_2 l_{2,5})X_S + (l_{2,2} - u_2 l_{2,6})Y_S + (l_{2,3} - u_2 l_{2,7})Z_S = u_2 - l_{2,4} \\ (l_{3,1} - u_3 l_{3,5})X_S + (l_{3,2} - u_3 l_{3,6})Y_S + (l_{3,3} - u_3 l_{3,7})Z_S = u_3 - l_{3,4} \end{array} \right. \quad (4.12)$$

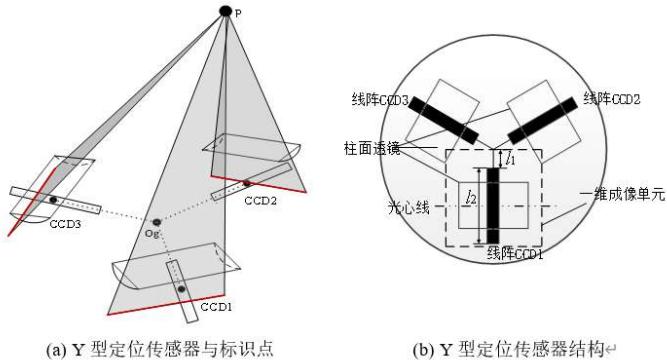


图 4.11 本文提出的 Y 型定位传感器光平面交会原理

写成矩阵形式为：

$$\mathbf{A}^S \mathbf{M} = \mathbf{b} \quad (4.13)$$

其中系数矩阵  $\mathbf{A} \in \mathbb{R}^{3 \times 3}$  和常数向量  $\mathbf{b} \in \mathbb{R}^{3 \times 1}$  由各相机的投影参数  $l_{i,j}$  及其实时观测值  $u_i$  构造。若  $\mathbf{A}$  满秩（即三个光平面不共线），则通过  $\mathbf{M} = \mathbf{A}^{-1}\mathbf{b}$  即可解算出目标在传感器坐标系下的精确三维坐标。在实际工程中，常采用最小二乘法或奇异值分解（SVD）求解以提高数值稳定性。

## 4.5 本章小结

本章首先阐述了线阵相机的成像机理及光斑中心亚像素定位算法，随后定义了系统涉及的各类坐标系。在此基础上，重点推导了基于投影矩阵  $\mathbf{L}$  的正向成像模型，明确了投影矩阵中 7 个固定参数的物理意义，并分析了基于多光平面交会的空间定位原理。这为后续章节开展系统参数标定与高精度定位实验奠定了坚实的理论基础。

## 第五章 线阵视觉系统的几何畸变校正与高鲁棒性标定方法实现

### 5.1 引言

基于视觉的三维测量技术通常建立在理想的几何光学模型之上。然而，在实际的物理成像系统中，由于制造工艺与装配精度的限制，镜头焦距的一致性、相机在空间中的绝对位姿、以及传感器与光学中心的耦合关系等，都不可避免地存在偏差。这些物理参数的非理想特性是引入测量误差的主要来源。因此，为了实现高精度的空间目标检测，必须对系统进行严格的标定与校准。

#### 5.1.1 现有技术现状与局限

在相机标定领域，针对面阵相机的研究已相当成熟，学术界和工业界发展出了多种经典算法。相比之下，针对线阵相机的标定技术研究相对匮乏。尽管文献中提出了适用于线阵模型的“7系数 DLT 方法”，为该领域提供了一定的理论基础，但在面对复杂光学系统时，其适用性仍面临挑战。

在本系统所采用的硬件架构中，为了满足特定的视场覆盖需求，引入了组合柱面透镜系统。实际测试表明，该光学结构引入了显著的非线性几何畸变，特别是在偏离光轴较大的边缘视场区域。若直接套用传统的线性模型，忽略镜头畸变带来的非线性误差，将导致标定结果偏离真实值。因此，如何有效地分离并校正这种特殊的镜头畸变，并克服远距离观测下的噪声干扰，成为本系统标定的关键难点。

#### 5.1.2 分阶段标定策略

针对上述复杂光学系统带来的非线性畸变与远场观测噪声分布不均等挑战，若直接采用整体优化方法求解，极易陷入局部最优。为此，本文提出一种“先校正、后解算”的分阶段解耦标定策略，将复杂的标定问题分解为两个相对独立的子问题进行处理：

**第一阶段：非线性几何畸变校正** 该阶段旨在剥离柱面透镜组引入的非线性误差。本文利用 ZEMAX 光学仿真软件生成高精度的栅格畸变数据，从而获取空间中点光源

的“真实成像位置”与“理想成像位置”之间的偏差信息。基于这些仿真数据，在 MATLAB 中建立高阶多项式畸变修正模型，并计算相应的畸变修正系数。此步骤的核心在于构建从“畸变图像坐标系”到“理想图像坐标系”的映射关系，从而消除光学系统非线性特性的影响。

**第二阶段：线性参数鲁棒解算** 在完成畸变校正后，后续的参数估计可被视为线性模型下的估计问题。首先利用第一阶段获取的校正系数对所有标志点的原始信号进行预处理，将其恢复为符合针孔/线阵成像模型的理想坐标。随后，针对远距离观测中存在的噪声异方差问题，引入加权估计策略（详见后文 2.3 节），精确解算出系统的线性结构系数及相机的内外部参数。通过这种分步策略，有效规避了非线性畸变与几何投影参数之间的强耦合，显著提升了标定结果的数值稳定性与精度。

## 5.2 柱面镜头的畸变与去畸变

### 5.2.1 畸变模型

本文所用柱面镜头组由光学设计软件 ZEMAX 设计，通过 ZEMAX 仿真，可以得到柱面镜头组理论上的畸变结果。ZEMAX 生成的栅格畸变数据包括空间中不同位置的点光源对应的真实成像位置  $\tilde{x}$  和理想成像位置  $x$ ，单位为 mm，位于图像坐标系。如图5.1所示，畸变为真实位置和理想位置的差， $\alpha$  表示点光源和相机坐标系 Y 轴所在平面与相机坐标系 Z 轴的夹角， $\beta$  表示点光源和相机坐标系 X 轴所在平面与相机坐标系 Z 轴的夹角， $\alpha$  和  $\beta$  的范围为  $[-20.8^\circ, 20.8^\circ]$ 。

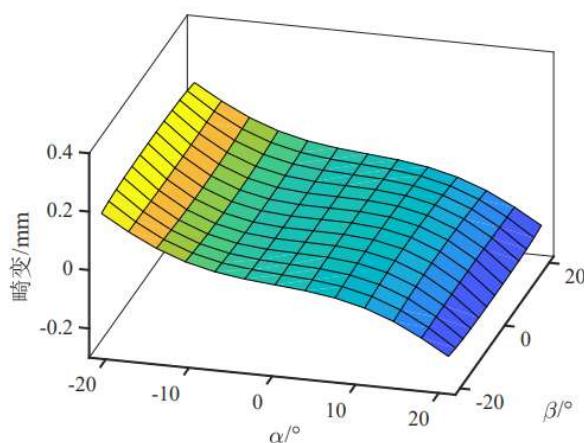


图 5.1 畸变和 、 之间的关系

假设空间中的点光源在相机坐标系下的坐标为  ${}^C\mathbf{M} = [X_C \ Y_C \ Z_C]^T$ , 定义归一化平面  $[x_n \ y_n \ 1]^T$ , 其中:

$$x_n = X_C/Z_C = \tan \alpha \quad (3.1)$$

$$y_n = Y_C/Z_C = \tan \beta \quad (3.2)$$

将镜头的相对畸变  $\Delta$  表示为  $x_n$  和  $y_n$  的函数, 定义如下:

$$\Delta = \begin{cases} \frac{\tilde{x}-x}{x} = \frac{\tilde{x}_n-x_n}{x_n} = h(x_n, y_n), & x_n \neq 0 \\ 0, & x_n = 0 \end{cases} \quad (3.3)$$

其中带有上标变量 ( $\cdot$ ) 表示存在畸变的真实值, 无上标变量表示理想值,  $h(\cdot)$  表示关于  $x_n$  和  $y_n$  的函数。根据  $\Delta$  和  $x_n$ 、 $y_n$  之间的关系可以得到新的曲面, 如图5.2所示。

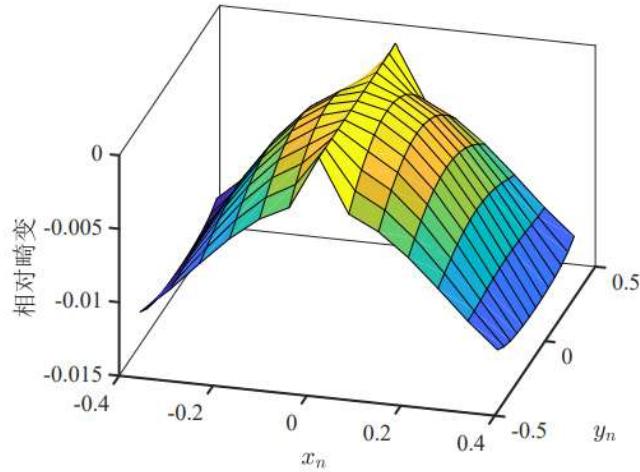


图 5.2 相对畸变和 、 之间的关系

本文利用 MATLAB 中的曲线拟合工具箱, 使用多项式对图中曲面进行拟合, 结果显示将  $h(x_n, y_n)$  表示为四阶多项式的形式可以较好地拟合曲面, 同时奇数次幂的系数可以忽略不计。因此可以将  $h(x_n, y_n)$  定义如下:

$$h(x_n, y_n) = p_1 x_n^2 + p_2 y_n^2 + p_3 x_n^2 y_n^2 + p_4 x_n^4 + p_5 y_n^4 \quad (3.4)$$

根据公式 (3.3) 可得  $\tilde{x}_n$  和  $x_n$  之间的关系为:

$$\tilde{x}_n = x_n(1 + h(x_n, y_n)) \quad (3.5)$$

其中  $\tilde{x}_n$  的值可由公式 (3.3) 计算得到。使用 MATLAB 中 lsqcurvefit 函数进一步拟合公式 (3.5)，各个系数的拟合结果如表 3.1 所示，其中均方根误差（Root Mean Square Error, RMSE）为  $0.0052\mu\text{m}$ 。

表 5.1 表 3.1  $h(x_n, y_n)$  各项系数拟合结果

$p_1$	$p_2$	$p_3$	$p_4$	$p_5$	RMSE/ $\mu\text{m}$
-0.1134	-0.0258	0.2047	0.1727	0.0189	0.0052

结合公式 (3.1) 和 (3.5)，可得该柱面镜头组的畸变模型为:

$$\tilde{u} = f(1 + h(x_n, y_n))x_n + u_0, \quad x_n = \frac{u - u_0}{f} \quad (3.6)$$

然而，由公式 (2.32) 可知， $y_n$  无法被线阵相机观测，因此需要将模型简化。

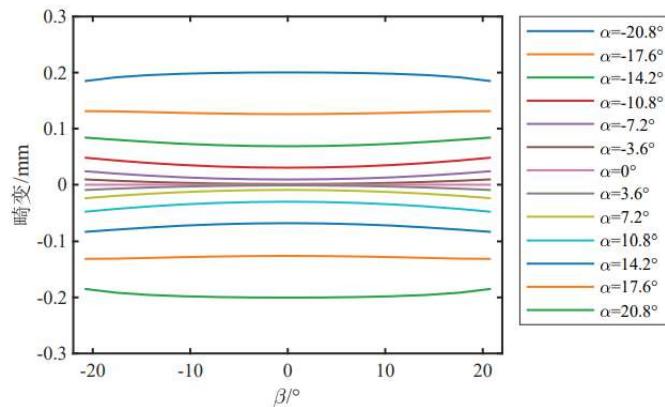


图 5.3 不同  $\alpha$  的情况下  $\beta$  变化导致的畸变变化

图5.3展示了在不同  $\alpha$  的情况下，仅  $\beta$  变化导致的畸变变化。图5.4中横坐标表示不同的  $\alpha$ ，纵坐标表示在当前  $\alpha$  保持不变的情况下， $\beta$  变化导致的像点最大偏移。由图 3.4 可知，当  $\alpha$  在  $\pm 10.8^\circ$  左右时， $\beta$  变化导致的最大偏移达到最大值。

当  $\|\beta\| \leq 20.8^\circ$  时，最大值为  $17.9\mu\text{m}$ ，当  $\|\beta\| \leq 15^\circ$  时，最大值为  $8.42\mu\text{m}$ 。又因为本文所用线阵 CCD 的像素单元长度为  $8\mu\text{m}$ ， $\beta$  变化导致的像点偏移在大多数情况下不超过 1 个像素点，所以可假设  $\beta$  变化产生的畸变对于测量精度影响不大。

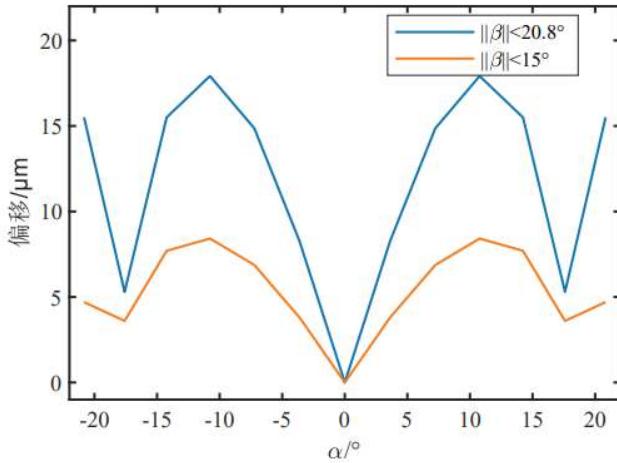


图 5.4 变化导致的像点最大偏移

将  $h(x_n, y_n)$  重新定义为只和  $x_n$  相关的多项式函数，则  $\tilde{x}_n$  和  $x_n$  之间的关系可表示为：

$$\tilde{x}_n = x_n(1 + h(x_n)) \quad (3.7)$$

利用 MATLAB 中 lsqcurvefit 函数，使用四阶多项式对公式 (3.7) 拟合， $h(x_n)$  的拟合结果如下：

$$h(x_n) = \frac{\tilde{x}_n - x_n}{x_n} = p_1 x_n^2 + p_2 x_n^4 \quad (3.8)$$

其中各个系数如表 3.2 所示，拟合结果的 RMSE 为  $0.1287 \mu\text{m}$ 。

表 5.2 表 3.2  $h(x_n)$  各项系数拟合结果

$p_1$	$p_2$	RMSE/ $\mu\text{m}$
-0.1325	0.3228	0.1287

简化后的柱面镜头组的畸变模型可表示为：

$$\tilde{u} = f x_n(1 + h(x_n)) + u_0, \quad x_n = \frac{u - u_0}{f} \quad (3.9)$$

### 5.2.2 去畸变模型

去畸变的过程是由实际像素坐标  $\tilde{x}$  计算理想像素坐标  $u$  的过程。为了得到去畸变模型，本文将相对畸变重新定义为：

$$\Delta' = \begin{cases} \frac{x - \tilde{x}}{\tilde{x}} = \frac{x_n - \tilde{x}_n}{\tilde{x}_n} = g(\tilde{x}_n), & \tilde{x}_n \neq 0 \\ 0, & \tilde{x}_n = 0 \end{cases} \quad (3.10)$$

利用 MATLAB 中的曲线拟合工具箱，使用多项式对公式 (3.10) 进行拟合，结果显示  $g(\tilde{x}_n)$  可同样表示为四阶多项式的形式，同时奇数次幂的系数可以忽略不计。 $g(\tilde{x}_n)$  拟合结果如下：

$$g(\tilde{x}_n) = k_1 \tilde{x}_n^2 + k_2 \tilde{x}_n^4 \quad (3.11)$$

其中各个系数如表 3.3 所示，拟合结果的 RMSE 为  $0.1289 \mu\text{m}$ 。

表 5.3 表 3.3  $g(\tilde{x}_n)$  各项系数拟合结果

$k_1$	$k_2$	RMSE/ $\mu\text{m}$
0.1349	-0.3252	0.1289

图像去畸变模型可表示为：

$$u = G(\tilde{u}) = f \tilde{x}_n (1 + g(\tilde{x}_n)) + u_0, \quad \tilde{x}_n = \frac{\tilde{u} - u_0}{f} \quad (3.12)$$

## 5.3 考虑几何灵敏度的线阵相机参数加权估计方法

### 5.3.1 问题背景与方法提出

经典的直接线性变换 (DLT) 标定方法在近场、高信噪比条件下通常能保持较好的稳定性。然而，在本文所涉及的 1m 至 6m 大工作距离范围以及复杂光照环境中，观测噪声往往呈现出显著的异方差特性 (Heteroscedasticity)。若直接采用传统 DLT 方法并通过“差分消去”构建方程，极易导致远距离特征点的噪声耦合与误差放大，从而降低参数估计的最终精度。

针对上述工程局限，本节引入统计信号处理中的加权估计思想，提出一种基于几何灵敏度与测量不确定性的加权最大似然估计 (Weighted Maximum Likelihood Estimation, WMLE) 方法。该方法主要包含三个核心策略：首先，摒弃传统的差分构方，

转而采用交叉相乘残差构建目标函数，以避免被减项选择带来的数值不稳定性；其次，在最大似然估计的框架下，通过误差传播分析推导残差方差，从而定义统计意义上的最优权重；最后，结合鲁棒代价函数与工程先验，构建完整的非线性最小二乘标定流程。

### 5.3.2 加权目标函数的严格推导

#### 5.3.2.1 成像模型与交叉相乘残差

基于前文所述的七参数有理投影模型，线阵相机的一维投影关系通常描述为像点坐标  $u_j$  是空间点坐标  $(X_j, Y_j, Z_j)$  的分式函数：

$$u_j = \frac{l_1 X_j + l_2 Y_j + l_3 Z_j + l_4}{l_5 X_j + l_6 Y_j + l_7 Z_j + 1} \quad (5.1)$$

为了简化书写并明确物理意义，我们将分子项  $N_j$  与分母项  $D_j$  分别定义为：

$$N_j := l_1 X_j + l_2 Y_j + l_3 Z_j + l_4 \quad (5.2)$$

$$D_j := l_5 X_j + l_6 Y_j + l_7 Z_j + 1 \quad (5.3)$$

其中，分母项  $D_j$  在几何上对应于空间点在相机光轴方向上的投影深度（即射影深度因子）。

直接使用公式 (5.1) 进行参数估计面临非线性强、奇异点敏感等问题。鉴于此，我们将分式形式转化为交叉相乘 (Cross-Product) 形式，即等式两边同乘分母  $D_j$ ，得到关于待求参数  $\mathbf{p} = [l_1, \dots, l_7]^\top$  的线性约束条件：

$$N_j - u_j D_j = 0 \quad (5.4)$$

展开后可得：

$$(l_1 X_j + l_2 Y_j + l_3 Z_j + l_4) - u_j (l_5 X_j + l_6 Y_j + l_7 Z_j + 1) = 0 \quad (5.5)$$

这一变换具有双重优势：其一，它消除了直接除法可能引入的数值不稳定问题（尤其是当点位于无穷远平面附近时）；其二，它将原始的非线性模型转化为对参数线性的代数误差形式，为后续的高效初始化提供了可能。

### 5.3.2.2 误差传播与权重设计

设像素坐标的真实值为  $u_j^*$ , 观测值为  $u_j$ , 假设观测噪声服从零均值高斯分布:

$$u_j = u_j^* + \delta u_j, \quad \delta u_j \sim \mathcal{N}(0, \sigma_{u,j}^2) \quad (5.6)$$

定义交叉相乘残差函数  $r_j(\mathbf{p}, u_j) = N_j - u_j D_j$ 。将观测模型代入该式，并利用一阶泰勒展开进行近似:

$$r_j = N_j - (u_j^* + \delta u_j) D_j = (N_j - u_j^* D_j) - \delta u_j D_j \approx -\delta u_j D_j \quad (5.7)$$

由此可推导出残差的方差为:

$$\text{Var}[r_j] \approx \text{Var}[-\delta u_j D_j] = \sigma_{u,j}^2 D_j^2 \quad (5.8)$$

上述推导表明，通过引入交叉相乘项，残差对像素噪声的敏感度发生了变化。具体而言，几何因子  $D_j$  (即景深相关项) 对原始噪声起到了“放大”或“缩放”的作用。基于高斯噪声假设，残差  $r_j$  近似服从分布  $\mathcal{N}(0, \sigma_{u,j}^2 D_j^2)$ 。

为了获得最大似然估计 (MLE)，负对数似然函数可转化为加权最小二乘问题:

$$\min_{\mathbf{p}} J(\mathbf{p}) = \sum_{j=1}^N w_j r_j(\mathbf{p})^2 \quad (5.9)$$

其中，统计最优权重  $w_j$  定义为:

$$w_j = \frac{1}{\sigma_{u,j}^2 D_j^2} \quad (5.10)$$

该权重设计自然编码了成像几何灵敏度与测量不确定性：距离相机越远的点 ( $D_j$  越大)，其权重越低，这与物理直觉完全一致。

### 5.3.3 非线性优化与实施流程

#### 5.3.3.1 优化策略

针对上述模型，本文采用“两步法”优化策略，即首先通过线性代数方法获取高质量的参数初值，随后利用非线性迭代算法精化结果，以兼顾计算效率与估计精度。

**1. 线性初始化 (DLT/SVD)** 基于交叉相乘公式 (5.4)，对于  $N$  个观测点，我们可以构建如下形式的齐次线性方程组：

$$\mathbf{A}\mathbf{p} = \mathbf{0} \quad (5.11)$$

其中数据矩阵  $\mathbf{A} \in \mathbb{R}^{N \times 7}$  的第  $j$  行向量  $\mathbf{a}_j$  为：

$$\mathbf{a}_j = [X_j, Y_j, Z_j, 1, -u_j X_j, -u_j Y_j, -u_j Z_j] \quad (5.12)$$

为了避免零解 (Trivial Solution)，通常引入约束  $\|\mathbf{p}\| = 1$ 。此时，该问题的求解转化为寻找矩阵  $\mathbf{A}$  的最小奇异值对应的右奇异向量。我们利用奇异值分解 (SVD) 对  $\mathbf{A}$  进行分解，取其最后一行右奇异向量作为参数向量的初始估计值  $\mathbf{p}_0$ 。此步骤即为经典的直接线性变换 (DLT) 解法。

**2. 加权非线性迭代 (LM 算法)** 线性解虽然计算简便，但其最小化的是代数误差而非几何误差，且未考虑各观测点的噪声异方差性。因此，我们将  $\mathbf{p}_0$  作为初值，采用 Levenberg-Marquardt (LM) 算法对加权平方残差和目标函数  $J(\mathbf{p})$  进行非线性最小化：

$$\min_{\mathbf{p}} J(\mathbf{p}) = \sum_{j=1}^N w_j (N_j(\mathbf{p}) - u_j D_j(\mathbf{p}))^2 \quad (5.13)$$

在第  $k$  次迭代中，我们要求解如下形式的增量正规方程 (Augmented Normal Equations) 以获取参数更新量  $\Delta\mathbf{p}$ ：

$$(\mathbf{J}^\top \mathbf{W} \mathbf{J} + \lambda \mathbf{I}) \Delta \mathbf{p} = -\mathbf{J}^\top \mathbf{W} \mathbf{r} \quad (5.14)$$

其中：

- $\mathbf{J}$  为残差向量关于参数的雅可比矩阵；
- $\mathbf{W} = \text{diag}(w_1, \dots, w_N)$  为基于测量不确定性构建的权重矩阵；
- $\lambda$  为阻尼因子，用于在梯度下降法 (大  $\lambda$ ) 与高斯-牛顿法 (小  $\lambda$ ) 之间动态调节，确保算法在远离极值点时的稳定性和接近极值点时的收敛速度。

通过不断更新  $\mathbf{p} \leftarrow \mathbf{p} + \Delta\mathbf{p}$  直至收敛，即可得到几何意义上最优的相机参数。

### 5.3.3.2 测量不确定性的估计与算法实现

为了准确计算权重，测量不确定性  $\sigma_{u,j}$  的估计至关重要。本文采用重复成像法，即对每个空间标记点采集多帧图像（如 10 帧），计算其亚像素坐标的标准差作为  $\sigma_{u,j}$  的估计值。

综上所述，WMLE 线阵相机标定算法的完整实施流程如下：

1. **数据归一化：** 对输入数据  $(X_j, Y_j, Z_j)$  和  $u_j$  进行中心化与尺度归一化处理，以提高数值稳定性，并记录逆变换矩阵。
2. **线性初始化：** 基于归一化数据构建线性方程组，利用 SVD 分解获取参数初值  $p_0$ 。
3. **迭代优化 (LM 算法)：**
  - 更新几何因子： $D_j = l_5X_j + l_6Y_j + l_7Z_j + 1$ ；
  - 更新权重：根据  $w_j = 1/(\sigma_{u,j}^2 D_j^2)$  计算各点权重；
  - 计算残差与雅可比：构建加权残差向量与加权雅可比矩阵；
  - 求解增量并更新：求解阻尼最小二乘问题，更新参数  $p$  及阻尼系数  $\lambda$ ；
  - 收敛判断：若参数增量或残差下降率小于阈值，则停止迭代。
4. **参数恢复：** 利用步骤 1 中的逆变换矩阵，将优化后的参数恢复至原始物理坐标系，输出最终标定结果及其协方差估计。

该方法通过合理引入误差传播分析，在保证计算复杂度的前提下，有效提升了线阵视觉系统在复杂工程条件下的标定鲁棒性。

## 5.4 相机内参计算方法

定义线阵相机内参矩阵为：

$$\mathbf{K} = \begin{bmatrix} f & u_0 \\ 0 & 1 \end{bmatrix} \quad (5.15)$$

定义矩阵  $\mathbf{B}$  为:

$$\begin{aligned}\mathbf{B} &= \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{bmatrix} = \begin{bmatrix} f & u_0 \\ 0 & 1 \end{bmatrix}^{-1} \mathbf{L} = \begin{bmatrix} \frac{1}{f} & -\frac{u_0}{f} \\ 0 & 1 \end{bmatrix} \mathbf{L} \\ &= \begin{bmatrix} \frac{l_1-u_0l_5}{f} & \frac{l_2-u_0l_6}{f} & \frac{l_3-u_0l_7}{f} & \frac{l_4-u_0}{f} \\ l_5 & l_6 & 1 & \end{bmatrix}\end{aligned}\quad (5.16)$$

其中  $\lambda$  为任意缩放因子。

由于  $\mathbf{R}$  是正交矩阵, 其行向量的模均为 1, 且任意两行不同向量的内积为 0。由  $\mathbf{r}_1 \cdot \mathbf{r}_3 = 0$  可得如下关系:

$$\mathbf{b}_1 \mathbf{b}_2^T = \lambda^2 \mathbf{r}_1 \mathbf{r}_3^T = 0 \quad (5.17)$$

进一步整理可得:

$$(l_5^2 + l_6^2 + 1) u_0 = l_1 l_5 + l_2 l_6 + l_3 \quad (5.18)$$

由  $\mathbf{r}_1 \mathbf{r}_1^T = \mathbf{r}_3 \mathbf{r}_3^T = 1$  可得如下关系:

$$\mathbf{b}_1 \mathbf{b}_1^T = \mathbf{b}_2 \mathbf{b}_2^T \quad (5.19)$$

结合公式 (5.18), 进一步整理可得:

$$l_1^2 + l_2^2 + l_3^2 - (l_1 l_5 + l_2 l_6 + l_3) u_0 = (l_5^2 + l_6^2 + 1) f^2 \quad (5.20)$$

获取空间中少量 (数量大于 7 个) 固定的标记点对应的像点, 用上一小节的方法可计算出投影矩阵  $\mathbf{L}$ 。根据公式 (5.18) 可计算出  $u_0$ , 之后根据已知  $u_0$  和公式 (5.20) 可计算出  $f$ 。

需要有效确保像点在线阵 CCD 上均匀分布。

## 5.5 相机坐标系和世界坐标系之间的刚体变换

每个投影矩阵  $\mathbf{L}$  都对应了一组  $[\mathbf{R} \mid \mathbf{t}]$ 。相机坐标系  $\mathcal{F}_c$  和世界坐标系  $\mathcal{F}_w$  之间的旋转矩阵  $\mathbf{R}$  和平移向量  $\mathbf{t}$  可由上一节算出的相机内参和对应的  $\mathbf{L}$  计算。结合前述公

式可得旋转向量：

$$\mathbf{r}_1 = \frac{[l_1 - u_0 l_5, \ l_2 - u_0 l_6, \ l_3 - u_0 l_7]}{\|[l_1 - u_0 l_5, \ l_2 - u_0 l_6, \ l_3 - u_0 l_7]\|} \quad (5.21)$$

$$\mathbf{r}_3 = \frac{[l_5, \ l_6, \ l_7]}{\|[l_5, \ l_6, \ l_7]\|} \quad (5.22)$$

$$\mathbf{r}_2 = \mathbf{r}_3 \times \mathbf{r}_1 \quad (5.23)$$

平移向量的分量计算如下：

$$t_z = \frac{1}{\|[l_5, \ l_6, \ l_7]\|} \quad (5.24)$$

$$t_x = \frac{l_4 - u_0}{f} \cdot t_z \quad (5.25)$$

平移向量的  $Y$  轴部分  $t_y$  无法计算，将其设定为 0，这和配有柱面镜的线阵相机特性相符，即在相机坐标系下，若空间中一标记点的  $X$ 、 $Z$  轴方向坐标不变， $Y$  轴方向上的变化不会导致像点位置变化。

由于噪声和镜头畸变的影响，计算出的  $\mathbf{R}$  不一定满足标准旋转矩阵的所有约束（如正交性），通常需要使用 SVD 重新计算新的旋转矩阵以确保其处于  $SO(3)$  空间内。

## 5.6 仿真实验

## 5.7 小结

综上所述，本文所采用的加权参数估计策略并非对传统标定理论的扩展，而是面向线阵视觉系统工程应用需求，通过合理引入误差传播分析结果，在保证计算复杂度可控的前提下，有效缓解了远距离标定与数据质量波动对参数估计稳定性的影晌，为后续系统集成与长期运行提供了可靠基础。本章以交叉相乘残差 + MLE 权重构建线阵相机的加权最大似然标定，其权重严格由像素噪声与几何灵敏度推导而来，

并提供了从初始化到鲁棒迭代与不确定度评估的完整可落地流程。在 RA8 工装坐标高可信的条件下，像素噪声主导、权重  $1/(\sigma_{u,j}^2 D_j^2)$  即为自然、严谨且有效的选择。

## 第六章 协同系统结构和实验结果分析

### 6.1 引言 (Introduction)

现有的 Wi-Fi CSI 室内定位研究面临两大核心挑战：一是深度学习模型对大规模、高精度标记数据的依赖与低效的人工采集方式之间的矛盾；二是单一射频 (RF) 模态在复杂室内环境（如非视距 NLOS、强多径效应）下的鲁棒性不足。

针对上述瓶颈，本文提出了引入多线阵 CCD 光学定位系统的解决方案，旨在构建一种“离线监督-在线互补”的分层协同架构。该架构通过“高精度数据基准构建”与“光电异构多模态互补”两个维度，实现了从数据生产到系统应用的全流程优化。

### 6.2 协同定位系统的实现架构

针对 Wi-Fi CSI 室内定位中存在的标记数据获取困难与非视距 (NLOS) 鲁棒性差这两大核心问题，本文首先从逻辑层面设计了“离线监督-在线互补”的协同架构，并基于该架构搭建了包含高精度光学捕捉与分布式射频感知的实验原型系统。

#### 6.2.1 总体逻辑架构设计

本架构利用多线阵 CCD 光学系统与 CSI 射频系统的异构特性，分别在数据构建阶段和实时定位阶段实现优势互补。

图 6.1 光电协同室内定位系统整体架构图

##### 6.2.1.1 数据层协同：高精度时空基准的自动化构建

深度学习驱动的 CSI 定位方法的性能上限，本质上受制于训练数据的规模与标记质量。传统的“走-停-采 (Stop-and-Go)”离散采集模式不仅耗时巨大，且人工定点引入的随机位置误差会造成显著的标签噪声 (Label Noise)，严重制约了模型对精细信号特征的学习。

为此，本系统引入多线阵 CCD 作为 \*\*“离线监督者”\*\*，构建了自动化的流式数据生产管线：

1. **连续流式采集 (Continuous Streaming Acquisition):** 利用 CCD 系统对携带特定波长 LED 的移动目标进行毫米级精度的实时追踪，替代传统的人工静态打点。这种方法将离散的指纹采集升级为连续的轨迹记录，不仅将数据生产效率提升了数量级，更消除了人为操作带来的不确定性。
2. **动态特征捕捉:** 本文提出的多尺度时序卷积 (MSTC) 网络高度依赖于信号的时序连续性。CCD 系统提供的连续高频坐标流，使得模型能够有效捕捉目标运动状态下 CSI 信号的多普勒频移 (Doppler Shift) 与时变衰落特性，为引入平滑约束 (Smoothness Loss) 提供了物理层面的数据支撑。

为了将异构传感器的原始数据转化为可供深度学习模型训练的标准数据集，本文实施了严格的时空对齐与预处理策略：

1. **空间重采样与轨迹标准化:** 针对原始轨迹数据的非均匀性，采用空间等间距采样策略。从 CCD 记录的连续轨迹中，以 10cm 为步长提取锚点，在实验空间内总计构建了 48,000 个高精度标记点。
2. **时域滑窗与特征匹配:** 基于光电系统的同步时间戳，以每个空间锚点为中心，截取时间窗口长度不少于 100 帧的 CSI 数据序列。
3. **标签插值:** 利用插值理论对采样间隙进行平滑处理，确保每一帧射频特征均能映射到唯一的精确空间坐标，从而最大限度地降低标签噪声，构建出“射频-空间”精确映射的指纹库。

### 6.2.1.2 策略层协同：基于视距感知的光电异构互补

除了作为离线训练阶段的真值生成器外，CCD 光学定位与 CSI 射频定位在物理特性上呈现出显著的 \*\*“正交互补性 (Orthogonality)”\*\*。本研究构建了基于视距 (LOS) 状态感知的自适应切换机制，以应对复杂的室内环境挑战：

- **光学主导模式 (Optical-Dominant Mode):** 在开阔的 LOS 场景下，光学信号未受遮挡。此时系统优先采用 CCD 输出，提供毫米级的绝对定位精度，同时利用该数据对 CSI 模型进行在线校准。
- **射频备份模式 (RF-Backup Mode):** 当目标移动至障碍物后方或发生视觉遮挡 (NLOS 场景) 时，CCD 系统会因光路中断或方差剧增而失效。此时，系统自动切换至 CSI 定位模块。利用 2.4GHz Wi-Fi 信号优异的绕射与穿透能力 (波长约 12.5cm)，通过多径反射特征维持分米级的定位服务，确保位置感知的连续性。

这种“光电双模”架构有效规避了单一光学系统的视觉盲区和单一射频系统的环境敏感性，显著提升了系统在动态不可控环境下的可用性（Availability）与鲁棒性。

## 6.2.2 光学监测节点：YOPS 硬件系统

光学监测节点作为系统的“离线监督者”，主要负责高精度的三维坐标真值生成。该子系统（YOPS）采用了模块化的嵌入式设计，核心架构如图 6.1 所示。

### 6.2.2.1 信号采集与处理单元

考虑到系统对光斑中心提取的实时性要求，主控单元（MCU）选用了基于 Cortex<sup>TM</sup>-M4 内核的 \*\*STM32F405\*\* 高性能处理器。该芯片具备 168 MHz 的主频与多通道 DMA 控制器，能够满足高频线阵数据流的实时吞吐需求。

光学传感前端采用了 \*\*TCD1304 线阵 CCD\*\*。该传感器具有  $8 \mu\text{m} \times 20 \mu\text{m}$  的像元尺寸与 3648 个有效像素，能够在保证高信噪比（High SNR）的同时提供微米级的光学分辨率。为了实现与上位机的高速数据交互，系统集成了 USB 2.0 高速物理层接口，确保了 1 kHz 以上的采样帧率下数据传输零丢包。

图 6.2 为本文所设计 YOPS 实物，其中 MCU 和 CCD 所在印制电路板（Printed Circuit Board, PCB）通过排针安装于线阵相机所在 PCB 背面。

### 6.2.2.2 嵌入式控制逻辑

软件架构旨在最小化从光子入射到坐标解算的系统延迟。不同于通用的分层架构，本系统采用了 \*\*中断驱动（Interrupt-Driven）\*\* 与 \*\*直接内存访问（DMA）\*\* 相结合的零拷贝传输机制：

- **并行采集：**利用 MCU 的定时器阵列同步触发三个线阵 CCD 的积分时序，确保多视角数据的时间一致性。
- **流式传输：**采集到的模拟电压信号经 12-bit ADC 量化后，直接通过 USB 批量传输（Bulk Transfer）模式上传至主机，由上位机（基于 Qt/C++ 开发）完成光斑质心的高斯拟合与三维重建。

如图 6.3 所示。本文所设计嵌入式软件分为 5 层，其中硬件层为 MCU 底层代码，直接控制 MCU 中的寄存器；抽象层将硬件层代码进行封装，提供 MCU 硬件（定时器、ADC、SPI 等）控制接口；设备层使用抽象层提供的硬件接口驱动线阵 CCD、

CC1101 等外部设备；接口层对设备层代码进一步封装，为每个外设定义初始化、读写、卸载等接口；应用层在接口层的基础上根据需求控制设备执行相应任务，并完成与上位机和标记点控制系统的通信。

上位机使用 C++ 编写，图形化界面基于 Qt 5.15 实现，用于读取并处理 YOPS 中传感器原始数据、运行本文设计 YOPS 标定方法完成传感器标定、光学定位算法、可视化传感器数据和测量结果。图 6.4 展示了上位机主界面，包括 YOPS 测量位姿 3D 显示窗口、CCD 原始数据显示窗口等。通过主界面，可以打开 YOPS 标定页面执行标定任务。

### 6.2.3 分布式射频感知网络：CSI 采集节点

与依赖专用时钟同步线缆的传统 MIMO 系统不同，本研究构建了基于商用现货（COTS）设备的分布式感知网络。该设计利用 \*\*ESP32-S3\*\* SoC 的低成本与高集成度特性，实现了感知节点在三维空间中的灵活部署。

#### 6.2.3.1 感知节点拓扑与空间覆盖

为了构建全向覆盖的射频感知场，系统在  $2.5\text{m} \times 2.5\text{m} \times 2.5\text{m}$  的实验空间内采用 \*\*“中心激励-边缘接收”\*\* 的拓扑结构，如图 6.2 所示：

- **激励源 (Tx)**: 1 个发射节点置于底面几何中心，持续广播全向探测包。
- **感知阵列 (Rx)**: 8 个接收节点均匀分布于空间四周侧面（高度 0.7m，间距 1m），形成对目标区域的包围式感知。

该布局确保了目标在区域内任意位置移动时，至少会切割多条射频链路的有效菲涅尔区（Fresnel Zone），从而引发可观测的多径扰动特征。

#### 6.2.3.2 异构网络同步与数据汇聚

由于分布式节点缺乏统一的物理时钟源，本系统设计了基于 \*\*SNTP（Simple Network Time Protocol）\*\* 的软同步机制以实现多模态对齐：

1. **网络对齐**: 所有 ESP32 节点与上位机处于同一局域网，系统启动时通过 NTP 协议同步 UTC 时间，消除设备间的秒级时钟偏差。
2. **微秒级打标**: 在底层驱动（基于 ESP32-CSI-Tool）中，利用硬件计数器在物理层前导码解析瞬间为每个 CSI 数据包（采样率 120Hz）打上微秒级时间戳。

3. 流式汇聚：CSI 数据通过 UDP 协议实时回传至服务器。服务器端依据全局时间戳对 8 路射频数据流与 CCD 光学数据流进行滑动窗口对齐，有效规避了网络抖动（Jitter）带来的时序误差。

图 6.2 分布式 8 节点感知系统空间布局示意图

## 6.3 实验环境部署与评估方案

基于上述构建的协同感知原型系统，本节将详细阐述实验的具体实施方案。为了全面评估系统在理想环境与复杂室内环境下的定位性能及鲁棒性，我们设计了包含特定被测目标的多维验证场景。

### 6.3.1 被测目标与实验环境配置

实验部署于一个典型的室内多径环境（具体布局参数见 6.2.3）。为了模拟标准点目标的运动特征并确保光电信号的同步可观测性，我们设计了专用的被测目标（Test Target）：

- **物理结构：**选用直径 12cm 的轻质泡沫球作为目标载体。球体表面均匀包裹铝箔材质，以增强对 2.4GHz 射频信号的反射截面积（RCS），确保在 NLOS 场景下仍能产生显著的多径扰动。
- **光学标记：**在球体表面固定特定波长的红外 LED 主动标记点，作为 YOPS 光学系统的追踪靶标。
- **运动控制：**为了最小化人体对射频场的侵入式干扰，实验中采用非金属材质的细长支架（直径 1cm，长 2m）对目标进行悬浮式牵引，确保 CSI 信号的波动主要源于目标球体的运动。

### 6.3.2 系统标定与真值坐标系构建

为了确保光电异构数据在同一物理空间下的精确映射，实验前需建立统一的世界坐标系，并完成线阵 CCD 视场的几何标定。本研究采用了 \*\*“离线刚体建模—在线激光追踪”\*\* 相结合的分级标定策略。

### 6.3.2.1 标定设备与工具

实验引入了工业级测量设备以构建高精度的 Ground Truth 基准（如图 6.6 所示）：

- **高精度关节式三维坐标测量臂（Portable CMM）**：用于建立刚体结构的局部几何特征，提供亚毫米级的接触式测量。
- **激光跟踪仪（Laser Tracker）**：作为大范围空间测量的“黄金标准”，用于实时捕捉刚体基座在世界坐标系下的绝对位姿。

图 6.3 实验标定设备：(a) 激光跟踪仪；(b) 标记点控制系统；(c) 关节式三维坐标测量臂

### 6.3.2.2 时空基准对齐流程

标定过程旨在建立 LED 光学中心与世界坐标系之间的刚性变换关系，具体分为两个阶段：

**1. 离线几何建模（局部坐标系构建）** 在静态环境下，利用三维坐标测量臂对标定板上的 LED 安装节点及靶球基座进行多点接触式采样。通过空间圆拟合算法，解算出各 LED 光心相对于靶球基座中心（Base Center）的局部坐标向量  $\mathbf{P}_{local}$ 。基于刚体假设，无论标定板如何移动，该相对几何关系始终保持不变。

**2. 在线位姿解算（世界坐标系映射）** 在实验过程中，激光跟踪仪实时锁定靶球基座，获取其在世界坐标系下的位置坐标与姿态四元数。结合离线阶段获取的几何拓扑关系，通过刚体变换矩阵（Rigid Body Transformation Matrix）将 LED 的局部坐标  $\mathbf{P}_{local}$  映射至全局世界坐标系  $\mathbf{P}_{world}$ ：

$$\mathbf{P}_{world} = \mathbf{R} \cdot \mathbf{P}_{local} + \mathbf{T} \quad (6.1)$$

其中  $\mathbf{R}$  和  $\mathbf{T}$  分别为由激光跟踪仪测得的旋转矩阵与平移向量。该方法为线阵 CCD 系统提供了具有工业级精度的空间参考点，确保了后续训练数据的标签误差远低于 Wi-Fi 信号的物理分辨率。

### 6.3.3 多维评估场景设计

针对单一 LOS 场景难以验证系统鲁棒性的问题，本研究设计了以下三组递进式实验场景，旨在从精度上限、遮挡恢复能力及抗干扰能力三个维度对算法进行综合评估。

### 6.3.3.1 场景一：理想视距（LOS）基准测试

- **实验设置：**移除实验空间内的所有障碍物，确保 CCD 视觉链路与 Wi-Fi 射频链路均无遮挡。
- **采集过程：**实验人员位于长方体空间外部，通过直径 1cm、长 2m 的细长实心木棍控制目标小球，使其以约 0.1m/s 的速度在空间内部进行缓慢匀速移动。
- **目的：**采集高质量的训练数据构建指纹库，并验证系统在理想条件下的定位精度上限。

### 6.3.3.2 场景二：物理约束下的非视距（NLOS）遮挡测试

- **实验挑战：**在验证 NLOS 性能时，通常面临“视觉真值与射频信号同时被遮挡”的悖论。
- **解决方案：**采用“物理轨迹约束法”。在实验区域中心铺设直线导轨，并在导轨中段设置不透明障碍物（阻断 Wi-Fi 直射路径并遮挡 CCD 视线），使导轨两端（区域 A、B）处于可视区，中间（区域 C）处于盲区。
- **采集过程：**控制目标小球沿导轨匀速穿过遮挡区域。
- **真值构建策略：**虽然 CCD 无法记录遮挡区内的实时坐标，但基于第五章建立的恒定速度（CV）运动模型，利用 CCD 记录的目标进入遮挡区时刻  $t_{start}$  的坐标  $P_{start}$  与离开遮挡区时刻  $t_{end}$  的坐标  $P_{end}$ ，通过时空线性插值（Spatiotemporal Linear Interpolation）生成遮挡区内的物理真值  $G(t)$ ：

$$G(t) = P_{start} + \frac{P_{end} - P_{start}}{t_{end} - t_{start}} \times (t - t_{start}), \quad t \in (t_{start}, t_{end}) \quad (6.2)$$

该方法有效解决了传感器盲区内的真值获取难题，用于评估系统在全遮挡条件下的定位稳定性。

### 6.3.3.3 场景三：人机共存的动态干扰测试

- **实验设置：**保持 CCD 视觉链路通视，引入一名实验人员作为动态干扰源。
- **干扰模式：**干扰者在 Wi-Fi 收发链路附近进行随机走动或切断信号路径，制造强多径效应与环境噪声。
- **目的：**CCD 持续记录受干扰状态下的目标真值，用于验证自适应卡尔曼滤波/粒子滤波算法对突变观测噪声的抑制能力及系统的环境鲁棒性。

## 6.4 实验结果分析与性能评估

### 6.4.1 光学传感器性能验证

### 6.4.2 协同定位系统综合性能评估

#### 6.4.2.1 理想场景下的精度上限分析

#### 6.4.2.2 遮挡恢复与 NLOS 鲁棒性分析

#### 6.4.2.3 动态环境干扰下的稳定性分析

## 6.5 本章小结

## 参考文献

- [1] 聂灵沼, 丁石孙. 代数学引论[M]. 北京: 高等教育出版社, 2000.
- [2] NARASIMHAN K, KULKARNI T, BARZILAY R. Language understanding for text-based games using deep reinforcement learning[J/OL]. CoRR, 2015, arXiv:1506.08941 [cs][2018-12-31]. <http://arxiv.org/abs/1506.08941>.
- [3] 刘全, 翟建伟, 章宗长, 等. 深度强化学习综述[J]. 计算机学报, 2018, 41(1): 1-27.
- [4] KNUTH D E. The TeX book[M]. 15th ed. Reading, MA: Addison-Wesley Publishing Company, 1989.
- [5] NIKIFOROV V. Analytic methods for uniform hypergraphs[J]. Linear Algebra and its Applications, 2014, 457: 455-535.
- [6] LU L, MAN S. Connected hypergraphs with small spectral radius[J/OL]. Linear Algebra and its Applications, 2016, 509: 206-227. DOI: 10.1016/j.laa.2016.07.013.
- [7] HU S, QI L, SHAO J. Cored hypergraphs, power hypergraphs and their Laplacian H-eigenvalues[J/OL]. Linear Algebra and its Applications, 2013, 439: 2980-2998. DOI: 10.1016/j.laa.2013.08.028.
- [8] LIN H, ZHOU B. Distance spectral radius of uniform hypergraphs[J/OL]. Linear Algebra and its Applications, 2016, 506: 564-578. DOI: 10.1016/j.laa.2016.06.011.
- [9] KANG L, NIKIFOROV V. Extremal problems for the  $p$ -spectral radius of graphs[J]. Electronic Journal of Combinatorics, 2014, 21(3): P3.21.
- [10] QI L.  $H^+$ -eigenvalues of Laplacian and signless Laplacian tensors[J]. Communications in Mathematical Sciences, 2014, 12(6): 1045-1064.
- [11] NIKIFOROV V. Hypergraphs and hypermatrices with symmetric spectrum[J/OL]. Linear Algebra and its Applications, 2017, 519: 1-18. DOI: 10.1016/j.laa.2016.12.038.
- [12] BU C, FAN Y, ZHOU J. Laplacian and signless Laplacian Z-eigenvalues of uniform hypergraphs[J]. Frontiers of Mathematics in China, 2016, 11(3): 511-520.

## 攻读硕士学位期间取得的研究成果

### 一、专著

[1] 张三. XXXXXXXXXXXXXXXXXX [M]. 南京: XXXXXXXXX 出版社, 2022.

### 二、论文

[1] ZHANG San. XXXXXXXXXXXXXXXXXXXX[J]. Bulletin of the Geological Survey of Japan, 2021, 57(5/6): 169-176.

[2] XXX, SATIO M, BUSH G, et al. XXXXXXXXXXXXXXXXXXXX XX[J]. Journal of the American Chemical Society, 2022, 2(3):20-26.

### 三、专利

[1] 张三. XXXXXXXXXXXXXXXXXX: 202210050175. 8 [P]. 2022-10-14.

### 四、项目

1. XX 省“十五”制造业信息化工程重大科技攻关项目“基于支持智能化 XXX 设计的创新开发平台”(编号), 主持, XXXX.XX.XX—XXXX.XX.XX。

2. XX 省“十五”制造业信息化工程重大科技攻关项目“基于支持智能化 XXX 设计的创新开发平台”(编号), 参与, 第二, XXXX.XX.XX—XXXX.XX.XX。

## 致 谢

衷心感谢导师 xxx 教授对本人的精心指导.

作者署名

完成地点

2023 年 12 月 18 日

## 附录 A 经典不等式

论文中用到的经典不等式.

**(Hölder Inequality)** 设  $a_i \geq 0, b_i \geq 0, i = 1, 2, \dots, n$ , 且  $p > 1, q > 1$  满足  $1/p + 1/q = 1$ . 则有

$$\sum_{i=1}^n a_i b_i \leq \left( \sum_{i=1}^n a_i^p \right)^{\frac{1}{p}} \cdot \left( \sum_{i=1}^n b_i^q \right)^{\frac{1}{q}},$$

等号成立当且仅当存在一个常数  $c$  满足  $a_i^p = cb_i^q$ .

**(PM Inequality)** 设  $x_1, x_2, \dots, x_n$  是  $n$  个非负实数. 如果  $0 < p < q$ , 那么

$$\left( \frac{x_1^p + x_2^p + \dots + x_n^p}{n} \right)^{\frac{1}{p}} \leq \left( \frac{x_1^q + x_2^q + \dots + x_n^q}{n} \right)^{\frac{1}{q}},$$

等号成立当且仅当  $x_1 = x_2 = \dots = x_n$ .

**(AM-GM Inequality)** 设  $x_1, x_2, \dots, x_n$  是  $n$  个非负实数. 则有

$$\frac{x_1 + x_2 + \dots + x_n}{n} \geq \sqrt[n]{x_1 x_2 \cdots x_n},$$

等号成立当且仅当  $x_1 = x_2 = \dots = x_n$ .