

光学与声学融合测距测位系统技术总览报告

设计技术方向与方案比较

November 7, 2025

Contents

1	总体说明	2
2	光学测距与测位技术	2
2.1	主动光学 (Active Optical Sensing)	2
2.1.1	1. 飞行时间法 (ToF, Time-of-Flight)	2
2.1.2	2. 相位调制 ToF	2
2.1.3	3. 主动结构光	2
2.2	被动光学 (Passive Optical Sensing)	2
2.2.1	1. 单目视觉几何法	3
2.2.2	2. 双目立体视觉	3
2.2.3	3. 视觉算法	3
3	声学测距与测位技术	3
3.1	主动声学 (Active Acoustic Sensing)	3
3.1.1	1. 超声波 ToF 测距	3
3.1.2	2. 多点超声阵列 TDoA	4
3.2	被动声学 (Passive Acoustic Sensing)	4
3.2.1	1. AoA (Angle of Arrival) 原理	4
3.2.2	2. TDoA 估计算法	4
3.2.3	3. 环境声要求	4
4	多模态融合策略	4
4.1	融合目标	4
4.2	置信度加权融合	4
4.3	卡尔曼滤波 (KF / EKF)	5
5	方案比较与分析	5
5.1	误差来源与分析	5
6	关键硬件与作用	5
7	综合设计思路	6
8	结论与建议	6

1 总体说明

本报告系统整理了主动光学、被动光学、主动声学、被动声学四类测距测位技术，以及其常用的核心算法（如 ToF、立体视觉、超声 TDoA、声源定位 AoA、GCC-PHAT 等），并比较其优缺点、难度、成本、典型误差来源与融合策略。目标是为课程设计中的多模态测距定位系统提供系统性的技术参考与方案选型依据。

2 光学测距与测位技术

2.1 主动光学（Active Optical Sensing）

主动光学系统通过向目标发射光信号（通常为红外或激光），测量反射光的时间、相位或强度变化以获得距离信息。

2.1.1 1. 飞行时间法（ToF, Time-of-Flight）

- 原理：测量光脉冲往返时间 Δt ，距离 $r = \frac{c\Delta t}{2}$
- 典型器件：VL53L0X, VL53L5CX, AD-355 系列
- 优点：精度高（mm 级），响应快（ms 级）
- 缺点：对透明/吸光材料精度下降；多径效应明显
- 输出：距离 + 回波强度（信号置信度）

2.1.2 2. 相位调制 ToF

- 采用连续调制波光源，测量调制相位差：

$$r = \frac{c\Delta\phi}{4\pi f_m}$$

- 精度更高，适合中短距离

2.1.3 3. 主动结构光

- 发射特定图案，通过相机观测其畸变重建深度
- 常用于 RGB-D 相机（如 Kinect、RealSense）
- 缺点：算法与光学结构复杂

2.2 被动光学（Passive Optical Sensing）

通过自然光或反射光成像获取信息，不主动发射信号。

2.2.1 1. 单目视觉几何法

- 若目标尺寸 W 已知，则距离可由相似三角形：

$$d = \frac{f \cdot W}{p}$$

其中 f 为焦距， p 为像素尺寸

- 优点：低成本
- 缺点：精度低，需目标特征已知

2.2.2 2. 双目立体视觉

- 通过两相机的视差 Δx 计算深度：

$$Z = \frac{f \cdot B}{\Delta x}$$

- 优点：精度高，完全被动
- 缺点：校准复杂，对光照与纹理敏感

2.2.3 3. 视觉算法

常用算法：

- 特征匹配：SIFT, ORB
- 目标检测：YOLO, OpenCV contour-based
- 跟踪与位姿估计：PnP, optical flow

3 声学测距与测位技术

3.1 主动声学（Active Acoustic Sensing）

主动发射声波并接收回波，通过时间差计算距离。

3.1.1 1. 超声波 ToF 测距

$$r = \frac{v_s \Delta t}{2}$$

其中 $v_s \approx 343$ m/s 为声速。特点：

- 成本低（常用 HC-SR04）
- 精度 cm 级
- 易受温湿度影响

3.1.2 2. 多点超声阵列 TDoA

多个发射/接收器计算时间差（TDoA）以三角定位。优点：可定位静止物体；缺点：同步与串扰难处理。

3.2 被动声学（Passive Acoustic Sensing）

系统仅接收声音，不主动发射信号。通过声源信号到达各麦克风时间差（TDoA）估计声源方向（AoA）。

3.2.1 1. AoA（Angle of Arrival）原理

$$\Delta t = \frac{B \sin \theta}{v_s} \Rightarrow \theta = \arcsin \left(\frac{v_s \Delta t}{B} \right)$$

- B : 麦克风间距
- Δt : 到达时间差
- θ : 入射角

3.2.2 2. TDoA 估计算法

1. 交叉相关法（CCF）
2. 广义互相关（GCC）
3. PHAT 加权（GCC-PHAT）:

$$R_{ij}^{PHAT}(f) = \frac{X_i(f)X_j^*(f)}{|X_i(f)X_j^*(f)| + \varepsilon}$$

4. IFFT 得时域互相关函数 $r_{ij}(\tau)$ ，峰值对应延迟 $\hat{\tau}$

3.2.3 3. 环境声要求

- 被动声学需声源存在（目标发声或背景音）
- 安静环境下需主动声源辅助

4 多模态融合策略

4.1 融合目标

- 光学与声学互补：ToF 对透明/暗面失效时，声学仍可用
- 增强鲁棒性：多源信息互证、失效检测与补偿

4.2 置信度加权融合

$$x_f = \frac{\sum_i C_i x_i}{\sum_i C_i}$$

其中 C_i 为各传感器输出置信度（基于信号质量、互相关峰值、检测置信度等）。

4.3 卡尔曼滤波 (KF / EKF)

- **KF**: 线性状态估计, 低复杂度
- **EKF**: 扩展至非线性系统
- 主要用于连续跟踪与噪声平滑
- 可选模块, 不影响基本功能

状态与观测模型:

$$x_{k+1} = Ax_k + w_k, \quad z_k = Hx_k + v_k$$

通过预测与更新循环最小化误差协方差。

5 方案比较与分析

维度	主动光学	被动光学	主动声学	被动声学
主要原理	光 ToF	图像几何	声 ToF	AoA/TDoA
能否主动测距	是	否	是	否
典型应用	LiDAR, ToF 相机	视觉定位	超声测距	声源定位
精度	高 (mm)	中 (cm)	中 (cm)	低 (角度)
抗环境干扰	一般	差	一般	中等
实现复杂度	中	高	低	中
成本	中	中	低	中
课程适配性	高	中	高	高

Table 1: 四类测距测位技术比较

5.1 误差来源与分析

- 光学: 表面反射率、噪声光、温度漂移
- 声学: 声速变化、麦克风间距误差、多径效应
- 视觉: 相机标定误差、像素量化、特征检测误差

误差传播 (二维):

$$\sigma_x^2 = (\cos \theta)^2 \sigma_d^2 + (d \sin \theta)^2 \sigma_\theta^2, \quad \sigma_y^2 = (\sin \theta)^2 \sigma_d^2 + (d \cos \theta)^2 \sigma_\theta^2$$

6 关键硬件与作用

- **STM32F4/F7 MCU**: 主控采样、信号处理、融合逻辑
- **ToF 模块 (VL53L5CX)**: 主动光学测距
- **摄像头模块 (OV2640)**: 视觉识别与角度估计
- **MEMS 麦克风阵列**: 声源方向估计 (AoA)
- **音频前端/ADC**: 信号放大与模数转换
- **电源/时钟模块**: 系统供电与同步

7 综合设计思路

1. 以主动光学 ToF 为主，提供高精度距离测量；
2. 以被动声学 AoA 为辅，提供方向与失效补偿；
3. 采用置信度加权融合策略；
4. 若时间允许，加入卡尔曼滤波进行动态平滑；
5. 重点聚焦传感链路、误差分析与校准。

8 结论与建议

- 主动光学 + 被动声学是在成本、难度、课程适配性上最佳平衡点；
- 系统可扩展性强，可后续加入视觉模态与 EKF；
- 重点应放在信号链、时序、同步与误差标定；
- 算法实现以可解释与可演示为主，追求系统完整性。