

基于光学 ToF 与被动麦克风阵列 AoA 的多模态测距与定位传感系统设计

November 6, 2025

Contents

1	设计背景	2
2	总体设计思路	2
2.1	目标	2
2.2	系统构成	2
3	设计选择理由	2
4	物理原理概述	2
4.1	光学 ToF	2
4.2	单目视觉	3
4.3	被动声学 TDOA / AoA	3
5	核心算法概念	3
5.1	GCC-PHAT TDOA	3
5.2	视觉检测与 bearing	3
5.3	ToF 读取与质量指标	3
5.4	融合策略	3
6	关键物料及作用	4
7	AoA 原理	4
8	TDOA / GCC-PHAT 参数考量	4
9	置信度设计	4
10	卡尔曼滤波 (KF/EKF)	4
11	误差分析	5
11.1	误差来源	5
11.2	误差传播公式	5
12	校准流程	5

1 设计背景

在常用测距测位传感系统中多为光学测量或声学测量，单一传感模态在复杂环境（如玻璃、低反射表面、噪声环境）下易失效。多模态融合可提升鲁棒性与精度。光学 ToF 模块提供高精度距离测量，但对透明或镜面物体失效；被动麦克风阵列通过声源方向估计（AoA）补充信息。结合视觉检测进一步增强系统能力。

2 总体设计思路

2.1 目标

设计并实现一个“以光学 ToF 为主、被动麦克风阵列 AoA 为辅”的多模态传感系统，用于室内近距（约 0.2-3 m）物体/声源的定位。系统通过模态冗余与失效补偿提高在复杂表面（玻璃、低反射、暗色）及噪声环境下的鲁棒性。重点在传感器链路（物理 → 电信号 → 数字采样）、时序/校准与误差分析，算法采用可解释、计算量可控的方案。

2.2 系统构成

- 主传感：光学 ToF 模块（主动，直接测距）
- 被动声学：MEMS 麦克风阵列 → AoA（方向）估计（GCC-PHAT）
- 主控平台：STM32F4/F7 MCU（实时采样 + DSP）
- 数据融合：置信度加权为主，卡尔曼滤波（KF/EKF）可选

3 设计选择理由

1. 互补性：ToF 在多数材质/照明下精度高但对透明/镜面失效；被动声学可提供方向信息。
2. 资源可控：现成 ToF 模块 + MEMS 麦克风阵列 + STM32 可完成硬件链路、信号处理与基础融合。
3. 契合度高：覆盖传感器物理、模拟前端、采样时序、校准与误差分析。
4. 可扩展性：可后续增加 EKF、更多麦克风阵列或视觉模块，形成多模态融合

4 物理原理概述

4.1 光学 ToF

发送短脉冲光，测量往返时间 Δt ：

$$r_{\text{opt}} = \frac{c \cdot \Delta t}{2}$$

其中 c 为光速。模块输出距离和信号质量指标。特点：精度高，响应快；对透明或低反射表面可能失效。

4.2 单目视觉

用于目标检测和角度估计。若已知目标尺寸，可粗略估计深度：

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

f 为焦距（像素）， W 为目标实际宽度， p 为像素宽度。

4.3 被动声学 TDOA / AoA

- 原理：声源发声，声波到达各麦克风的时间不同，测量通道间到达时间差 Δt (TDOA)，由阵列几何推算入射角 (AoA)。
- 限制：只能测方向，不能直接测距；需要目标或环境发声。

5 核心算法概念

5.1 GCC-PHAT TDOA

1. FFT 得到每通道频域表示 $X_k(f)$
2. 计算互谱 $R_{ij}(f) = X_i(f)X_j^*(f)$
3. PHAT 加权：
$$R_{ij}^{PHAT}(f) = \frac{R_{ij}(f)}{|R_{ij}(f)| + \varepsilon}$$
4. IFFT 得到互相关函数 $r_{ij}(\tau)$ ，峰值对应延迟 $\hat{\tau}_{ij}$
5. 根据阵列几何求 AoA，结合多通道加权平均生成最终方向

5.2 视觉检测与 bearing

$$\theta_v = \arctan\left(\frac{u - c_x}{f_x}\right)$$

5.3 ToF 读取与质量指标

ToF 输出距离 r 与回波强度 S 或信号质量 Q ，映射为置信度 C_{opt} 。

5.4 融合策略

角度融合：

$$\theta_f = \frac{C_v \theta_v + C_a \theta_a}{C_v + C_a}$$

距离选择：若 C_{opt} 高，用 ToF 距离；否则使用视觉或其他策略。
线性 KF 或 EKF 对时间序列观测进行平滑/跟踪。

6 关键物料及作用

- STM32F4/F7 MCU: 采样、DSP 运算、融合逻辑
- 光学 ToF 模块: 主动测距, 提供置信度
- 单目摄像头: 目标检测、方向信息
- MEMS 麦克风 $\times 4$: 阵列采集声学信号, 用于 AoA
- 前端/ADC: 信号放大滤波
- 电源/时钟/连接线: 稳定供电和时序

7 AoA 原理

对于两麦克风线性阵列:

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

- 无法直接测距, 只提供方向线
- 必须有声源 (目标或环境音)

8 TDOA / GCC-PHAT 参数考量

- 采样率 f_s : 分辨率 $1/f_s$, 可提升亚样本精度
- 窗长 N : 平衡时域/频域分辨率
- 多径处理: 选择直接到达峰
- 阵列基线 B : 越长角度分辨率越高, 但受物理限制

9 置信度设计

- 光学 ToF 置信度 C_{opt} : 回波强度/信号质量
- 视觉置信度 C_v : 检测器置信度 + bbox 稳定性
- 声学置信度 C_a : 互相关峰值/背景能量比 + 多通道一致性

10 卡尔曼滤波 (KF/EKF)

- 线性 KF: 适用于线性状态与观测
- EKF: 处理非线性观测 (如 bearing, r)
- 实现难度中等, STM32F4 完全可用

11 误差分析

11.1 误差来源

- ToF: 时间测量噪声、表面反射、背景光
- 视觉: 相机内参、像素量化、目标检测错误
- AoA: 采样/量化误差、阵列基线测量误差、多径、声速偏差

11.2 误差传播公式

若 $x = d \cos \theta, y = d \sin \theta$:

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

12 校准流程

1. 相机内参标定
2. ToF 偏置校准
3. 麦克风阵列时延和增益校准
4. 传感器外参标定
5. 温度校正（声速）
6. 测试与验证