

Overview of Non-line-of-sight Sound Source Localization Techniques

WANG Shuo, CHEN Zhe*, YIN Fuliang

(School of Information and Communication Engineering, Dalian University of Technology, Dalian 116024, China)

Abstract: Non-line-of-sight sound source localization refers to the technique of achieving accurate sound source positioning in environments with obstacles and occlusions by analyzing acoustic propagation characteristics and applying traditional digital signal processing or deep learning approaches. This technology has a wide range of application prospects in robot navigation, smart home, security monitoring and other fields. This paper presents a comprehensive review of non-line-of-sight sound source localization methods. Building on the fundamental concepts of non-line-of-sight localization, it classifies and discusses existing approaches and research progress according to three main modeling frameworks: Multi-sensor localization models, geometrical acoustics models, and statistical acoustics models. Moreover, this paper summarizes various non-line-of-sight source localization methods, identifies current challenges, and presents promising directions for future research.

Highlights:

1. The principles of sound propagation in non-line-of-sight (NLOS) environments are systematically reviewed, with a particular focus on sound reflection and diffraction, as well as representative propagation models and their applicable conditions.
2. Existing NLOS sound source localization methods are comprehensively reviewed, with indoor approaches mainly categorized into acoustic measurement-based, geometric acoustics-based, and statistical acoustics-based localization models, and their respective advantages and limitations are analyzed.
3. The key challenges in NLOS sound source localization are summarized, and future research directions are discussed, including multi-source localization, adaptation to dynamic environments, real-time processing, diffraction modeling, and multimodal sensor fusion.

Key words: non-line-of-sight (NLOS); sound source localization; acoustic reflection; acoustic diffraction; multi-sensor localization; microphone array

Foundation items: National Natural Science Foundation of China (Nos.62271103, 61871066); National High Technology Research and Development Program of China ("863" Program) (No.2015AA016306); Natural Science Foundation of Liaoning Province (No.20170540159); Fundamental Research Funds for the Central Universities (No.DUT17LAB04).

Received: 2025-08-15; **Revised:** 2025-12-06

***Corresponding author, E-mail:** zhechen@dlut.edu.cn.

非视域声源定位方法综述

王 硕, 陈 喆, 殷福亮

(大连理工大学信息与通信工程学院, 大连 116024)

摘 要: 非视域声源定位技术是指在障碍物遮挡环境中, 通过分析声波传播特性, 并结合信号处理理论, 采用传统的数字信号处理方法或深度学习方法实现声源精确定位。该技术在机器人导航、智能家居和安全监控等领域具有广泛应用前景。本文对非视域声源定位方法进行了系统综述, 以非视域声源定位的相关概念为基础, 按照不同的模型对非视域声源定位方法及其研究现状进行阐述, 其中主要包括多传感器定位模型、几何声学模型和统计声学模型, 比较、分析并总结了不同模型对应的非视域声源定位方法, 最后指出存在的问题并对未来研究方向进行展望。

关键词: 非视域; 声源定位; 声反射; 声衍射; 多传感器定位; 麦克风阵列

中图分类号: TN911

文献标志码: A

引用格式: 王硕, 陈喆, 殷福亮. 非视域声源定位方法综述[J]. 数据采集与处理, 2026, 41(4): 1026-1040. WANG Shuo, CHEN Zhe, YIN Fuliang. Overview of non-line-of-sight sound source localization techniques[J]. Journal of Data Acquisition and Processing, 2026, 41(4): 1026-1040.

引 言

声源定位(Sound source location, SSL)技术是指利用麦克风等声传感器采集声音信号, 通过信号处理技术估计声源在空间中的位置。声源定位^[1-3]作为感知目标的重要手段, 在机器人导航、智能家居、安全监控、无人驾驶、环境监测、灾害救援以及国防军事等领域得到了广泛应用。通常, 声源定位技术能让如服务机器人、智能音响等智能设备像人类一样感知声音的方位, 从而更加全面、准确地判断周围的环境情况, 作为后续决策和行动的依据。声源定位是声学信号处理领域的重要研究方向之一, 并日益受到高度关注。

使用单个麦克风难以获取空间信息, 因此, 当前声源定位大多基于麦克风阵列^[4-5]。在可视范围内, 尽管存在混响和背景噪声的干扰, 但声源到麦克风之间的直达声信号由于其强度较高, 能够保证获得较为理想的声源定位性能。然而, 在实际应用中, 声源与麦克风节点并非总位于可视范围内, 两者间常有障碍物遮挡。这种非视域(None-line-of-sight, NLOS)场景导致的“遮挡效应”, 使得声源发出的声波在传播过程中受到障碍物的影响, 产生反射、衍射及散射现象, 导致接收到的声波信号出现畸变和不完整, 从而使得传统声源定位方法的性能显著下降。尤其是在声源或麦克风移动的应用场景, 时变的反射和衍射现象更是加剧了信号的畸变程度, 严重时, 传统声源定位方法甚至无法正常定位。

非视域声源定位技术一般分为室内非视域声源定位和室外非视域声源定位。室内空间相对狭小且封闭, 环境内物体通常较为固定, 声波传播中反射与衍射效应显著, 此时可利用较稳定的空间结构信息进行建模与约束。但室内环境中混响和噪声通常较强, 多径传播会导致时延与方向信息混叠, 局部

基金项目: 国家自然科学基金(62271103, 61871066); 国家“863”计划项目(2015AA016306); 辽宁省自然科学基金(20170540159); 中央高校基本科研业务费专项资金(DUT17LAB04)。

收稿日期: 2025-08-15; **修订日期:** 2025-12-06

遮挡物对传播路径产生干扰,从而对定位精度造成明显影响。而室外环境更为开阔和复杂,场景常具有动态性与随机性,声波传播范围更大且遮挡物形态不规则,环境噪声、气象条件等因素影响更强,使得室外定位面临更大的挑战。尽管如此,通过设计合理的传感器阵列并引入鲁棒的算法,室外场景仍可实现较为可靠的声源定位。总体而言,室内环境相对易于建模与控制,但混响和噪声的影响是其关键难点,而室外环境则面临更强的不确定性和复杂性^[6-11]。近年来,室内非视域声源定位技术的研究已取得一定进展,而室外场景下的相关研究相对滞后,亟待加强系统性探索。随着声源定位技术在机器人、安全监控以及国防军事等领域的广泛应用,突破非视域场景中的声源定位技术瓶颈,不仅在学术上具有重要意义,在实际中也具有广泛应用前景。

本文对非视域场景中声源定位的相关研究进行梳理,依照原理将其分为几种类型,并对其国内外研究现状进行综述,同时指出存在的问题和未来发展方向。

1 非视域场景中声的传播原理

声音从声源出发,以波的形式在介质中传播。在均匀介质中,声波通常沿直线传播,但当遇到不同的介质或障碍物时,可能会发生反射、衍射和散射等现象。

1.1 声反射

声波的反射指声波在传播过程中遇到障碍物时,部分声波会从障碍物表面按照反射定律射向另一方向。此现象在许多场景中都可观察到,如山谷中的回声或封闭房间里的混响。

通常情况下,声波的反射分为镜面反射和漫反射。当反射面比较光滑时,声波遵循反射定律,以等同于入射角的角度反射出去;而当反射面比较粗糙时,声波会发生漫反射,即声线会射向四面八方,此时反射的声线是杂乱无章的,但每一条声线都遵循反射定律。由于漫反射的模拟过程较为复杂,不利于实时定位,且已有研究表明,在典型室内环境中,镜面反射通常占主导,而经由漫反射产生的声信号在整体声源能量中所占比例相对较小^[12-14],所以在利用反射声线进行声源定位时,通常只考虑镜面反射^[15]。

如图1所示,在室内环境下,声源 S 发射出声线 $r_i (i=1, 2, \dots, n)$,随机向房间内的某一点射去,入射角为 θ_i ,经墙面或室内物体 B_i 反射后,生成反射声线 r'_i ,由麦克风 M_i 接收。

1.2 声衍射

声波的衍射指声波在传播的过程中遇到障碍物或小孔时,能绕过障碍物的边缘或小孔继续传播。当障碍物的尺寸远大于波长时,声波近似为直线传播;但当障碍物或孔洞的尺寸与声波波长相近或比波长更小时,衍射现象较为显著。

如图1所示,声源 S 发射出声线 $d_i (i=1, 2, \dots, n)$, d_i 随机射到障碍物上的 P_i 点,假设室内障碍物 B_3 的尺寸相对于声波波长较小,此时 d_i 在点 P_i 处发生衍射, P_i 成为新的点波源,向不同的方向发射子声波 d'_{ij} , j 表示该衍射点产生的不同传播方向的子声波编号。

在非视域声源定位中,衍射建模直接决定了定位的精度与鲁棒性。常用方法主要分为两类:基于比奥-托尔斯托伊-梅德温(Biot-Tolstoy-Medwin, BTM)^[13,15-17]的时域

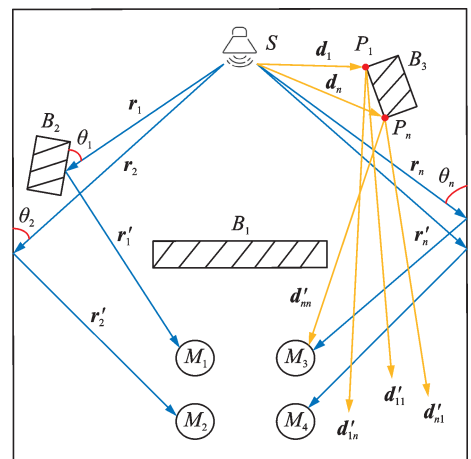


图1 简单的室内声反射和衍射场景

Fig.1 Simple indoor sound reflection and diffraction scene

全波模型和基于统一衍射理论(Uniform theory of diffraction, UTD)^[18-20]的高频渐近模型。BTM模型不依赖高频近似,能够在波长与障碍物尺寸可比拟的条件下再现瞬态响应,其幅度误差可小于0.2 dB^[16],因此常作为基准解,但其计算量随频率和几何复杂度急剧增加,限制了实时应用^[21]。UTD模型则通过解析衍射系数,可在高频情况下快速近似声场,常适用于大尺度场景^[22],但在频率低于140 Hz时精度下降,此时与精确解偏差较大^[16]。因此,在障碍物尺寸与波长相比拟的低频场景,BTM模型的精度较高,计算量较小;而在高频或对实时性要求较高的应用中,UTD模型及其与射线追踪结合的方法更为有效。在实际中,常采用边缘细分、预计算或混合建模策略,以便在精度与效率之间取得平衡^[21,23]。

由此可见,UTD衍射理论模型通常用于计算高频声波在遇到障碍物边缘时发生衍射的行为。该模型以几何光学为基础,通过引入绕射修正因子,使得其计算结果在过渡区更加准确和平滑^[22]。根据UTD模型,不同的衍射场景如楔形物、平面或曲面屏和弯曲楔形物,对应不同的衍射系数计算公式;不同的入射波如平面波、圆柱波、圆锥和球面波,决定了系数公式中参数的取值。在一般的室内场景中,声波遇到楔形物产生衍射的现象较为常见。对于楔形物,标量衍射系数的计算公式为^[22]

$$D_{s,h}(\phi, \phi'; \beta_0) = \frac{\exp\left[-j\frac{\pi}{4}\right] \sin \frac{\pi}{n}}{n\sqrt{2\pi k} \sin \beta_0} \quad (1)$$

$$D_{s,h}(\phi, \phi'; \beta_0) = \frac{\exp\left[-j\frac{\pi}{4}\right] \sin \frac{\pi}{n}}{n\sqrt{2\pi k} \sin \beta_0} \left[\cot\left(\frac{\pi + \phi_1}{2n}\right) F[kLa + \phi_1] + \cot\left(\frac{\pi - \phi_1}{2n}\right) F[kLa - \phi_1] \right. \\ \left. - \left\{ \cot\left(\frac{\pi + \phi_2}{2n}\right) F[kLa + \phi_2] + \cot\left(\frac{\pi - \phi_2}{2n}\right) F[kLa - \phi_2] \right\} \right] \quad (2)$$

式中: $\phi_1 = \phi - \phi'$, $\phi_2 = \phi + \phi'$, ϕ 和 ϕ' 分别为衍射射线和入射射线方向的角度; β_0 为入射射线与边缘切线的夹角; k 为波数; L 表示距离参数(当入射波为平面波时, $L = s \sin^2 \beta_0$;当入射波为圆柱波时, $L = r/(r + r')$, r 为从波源(发出入射波的点)到边缘上的衍射点的直线距离, r' 为从边缘上的衍射点到观测点(接收衍射波的点)的直线距离;当入射波为圆锥和球面波时, $L = ss' \sin^2 \beta_0 / (s + s')$, s 为沿入射射线方向(从波源指向边缘衍射点的射线)声波从波源传播到边缘衍射点的路径长度, s' 为沿衍射射线方向(从边缘衍射点指向观测点的射线)声波从边缘衍射点传播到观测点的路径长度); a 为边缘在衍射点处的曲率半径; $F(x)$ 为过渡函数,包含菲涅尔积分。

式(1)适用于非过渡区域。非过渡区是指远离阴影边界的区域,包括直射区和阴影区:直射区为观察点位于直射波覆盖范围内,衍射贡献相对较小且稳定;阴影区为观察点完全位于阴影中,没有直射波到达,衍射成为主要贡献波,但声场变化平缓。在非过渡区内,几何声学(即Keller高频衍射解)能够准确地描述声场,场强变化平滑连续,因此衍射系数无需过渡函数修正,可直接采用高频解。

式(2)适用于过渡区域及所有角度。过渡区是指阴影边界附近的区域(又称Fresnel区),此处声场随空间变化剧烈,从直射区快速衰减至阴影区。由于几何边界条件发生突变,Keller高频解可能出现奇异或不连续现象。为保证声场场强在边界处的连续性,UTD模型引入过渡函数 $F(x)$ 对衍射系数进行修正,以实现平滑过渡。

2 室内非视域声源定位方法

对于室内非视域声源定位方法,主要基于3种定位模型。(1)基于声信息测量的定位模型,即通过分析信号到达传感器的时间差或到达角等信息推算出声源位置。在复杂的环境中,多个传感器协同工作能减少遮挡物和噪声对定位结果的影响,但传感器的数量和布局对定位精度的影响较大。(2)几何声学定位模型,此模型主要考虑声波传播过程中发生的反射和衍射效应,在相对简单且规则的环境中表现

良好,但难以应对复杂和随机的遮挡效应,特别是当环境非常不规则或衍射效应非常复杂时,定位效果不尽如人意。(3)统计声学定位模型,这种模型基于统计学原理,对声反射、衍射等现象进行建模和分类,通过概率分布刻画声音传播中的不确定性,最终实现声源目标的定位。由于考虑了复杂环境中的随机性和遮挡效应,此类模型的适应性较强,但当遮挡物数量较多时,算法的计算复杂度较高。

除上述3大类模型之外,还有研究人员提出非视域信号的识别与分类^[24-29]方法,可以帮助定位系统在复杂的非视域场景中根据需要调整定位过程,这对遮挡声源定位的研究也有一定的启发意义。据此,本节系统梳理了室内非视域场景下各类声源定位模型的研究脉络,重点剖析了不同技术的优势与局限,可作为后续该领域研究的重要理论参考。

2.1 基于声信息测量的定位模型

基于声信息测量的定位模型是指通过测量声波在传播过程中所获得的信息来估算声源的位置。在非视距声源定位中,常通过时间、角度等信息来捕捉声波的传播特性。基于这些声学测量数据,可以构建定位模型,从而推算出声源的位置。常见的声学测量模型包括基于到达时间、到达角等方法,这些模型通过精确测量时间差、角度或其他传播特性,弥补非视距环境中信号的衰减和畸变,从而实现声源定位。此外,还有一些方法通过融合声传感器与光学传感器、惯性传感器等多种传感器的数据,进一步提升定位的精度和鲁棒性。

2.1.1 基于到达时间的非视域声源定位方法

到达时间(Time of arrival, ToA)指声音信号从声源出发,经传播后到达接收器所经历的时间。在声源定位中,常指声音信号从发出到被麦克风接收的绝对时间。在非视域场景中,由于缺乏直达路径,声源信号只能通过反射、衍射或组合路径到达接收端。如图2所示,这是一种基于ToA的NLOS SSL场景^[30],其中 $R_i(i=1,2,3,4)$ 为从声源传播到到麦克风的声线, t_i 为声线从声源到达麦克风的时间, H 为同步位置和麦克风之间的距离, t_0 为同步时间。非直达路径误差定义为接收信号相对于直达路径增加的额外距离,该误差是正偏差,其大小依赖于具体的传播环境。

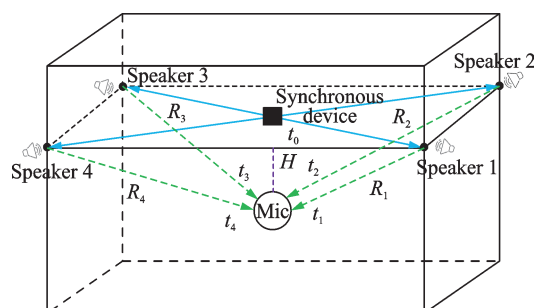


图2 基于ToA的NLOS SSL场景

Fig.2 ToA-based NLOS SSL scene

定位过程中,主要有3种基于ToA的方法来减小非直达声信号的影响:(1)通过已知的多径散射模型来预测并降低非直达路径误差,此方法的优势是不需要大量麦克风^[31];(2)结合所有非直达路径的测量值,采用不同权重因子进行目标定位,以提高定位精度,且不需要进行启发式的非直达路径麦克风识别^[32];(3)通过不同数学模型提供闭式的区位估计表达式,有效抑制非直达路径误差,且无需任何先验知识^[33]。

Lin等^[34]采用建模地图估计真实的ToA,将其与测得的ToA进行对比,以更新目标位置,实现低频遮挡声源的定位,其静态与动态声源的二维定位精度分别达到0.12 m和0.25 m。然而,由于系统采样率较低,该方法需等待0.2~0.5 s,以消除前期声音信号的混响。考虑到背景噪声会降低信号的信噪比,影响系统的精度、采样率和范围,Lin等^[35]对上述方法进行了改进,首先估计并校正ToA中的偏差,然后利用校正后的ToA递推估计声源位置,并在高达75.8 dB背景噪声条件下,成功定位了两个不同声源,精度范围为0.14~0.50 m。然而,该方法依赖于先验已知目标声源且具有无重复模式,因此在相对较小的空间环境中更加有效。Cao等^[30]提出一种音频信号ToA检测算法,在NLOS环境中,约90%定位结果的误差小于0.50 m。该方法采用由粗略到精细的搜索方式,粗略搜索利用余弦定理从移动设备的接

收信号中提取音频数据段(Audio data segment, ADS),而精细搜索通过源信号的自相关函数特性来估计ADS的ToA。

除此之外,Zuo等^[36]提出一种基于ToA的声学指纹室内定位架构,在低锚点部署密度的密集非视距环境下,能实现智能手机在室内的准确定位,具有良好的实际应用价值。Zhang等^[37]基于双曲调频(Hyperbolic frequency modulated, HFM)信号的多普勒不变性,引入到达时间和到达频率(Frequency of arrival, FoA)估计,提出一种基于声信号ToA和FoA的粒子滤波定位方法,可实现速度和距离估计,并适应不同的应用场景。

2.1.2 基于到达角的非视域声源定位方法

信号从发射端传播到接收端时,传播方向与接收端参考方向之间的夹角就是到达角(Angle of arrival, AoA),在三维空间中,通常需要用方位角和仰角来完整描述信号的到达方向。

Wang等^[38]提出了一种基于AoA的遮挡声源定位方法。该方法首先通过融合短时傅里叶变换与小波分析的多分辨率分析技术,来估计时变低频相干声音信号的AoA,解决了传统方法在处理相干信号和低频声音时的精度问题。在此基础上,利用发射声音信号和改进的3D多重信号分类(Multiple signal classification, MUSIC)算法完成室内结构参数估计,最终通过联合声源到达角与室内结构信息重构声波传播路径,实现遮挡环境下的声源定位,定位误差为0.47 m。

Xiong等^[39]针对室内环境中混合视距和非视距的三维AoA声源定位问题,提出了2种高效且易用的减小误差方案。方案1给出了闭式位置估计,即针对视距AoA分量,根据到达方向向量推导出加权线性最小二乘估计器,然后用残差平方和进行数据选择,以丢弃易出错的非视距线路。方案2基于模拟退火框架,应用稳健的 ℓ_1 最小化准则,减小了非视距定位误差,并给出了计算机模拟和超声现场实验定位结果。

2.1.3 多信息融合实现非视域声源定位

传感器阵列指利用多个传感器组成阵列,通过合理设置传感器的数量和布局来提高定位精度和空间分辨率。张弛^[40]基于传感器阵列实现遮挡声源定位,该方法首先得到声线到达阵元的时间,再根据阵列的空间位置建立非线性方程组,并构建适应度函数,最后用粒子群优化获得遮挡声源定位结果,但其计算量相对较大。

多信息融合是指将来自多个不同类型传感器(如声学传感器和光学传感器,或声学传感器和惯性传感器)的数据进行综合处理和分析,利用多种类型传感器的互补信息,发挥不同传感器独特的感知能力和优势,以获得更准确的声源位置信息。例如,Wang等^[41]结合声传感器与惯性传感器,成功实现了非视域声源定位,在距离锚点50 m的非视域场景中,以95%的置信度实现了约30 cm的定位精度。Sun等^[42]提出了一种基于麦克风阵列的声学-惯性融合定位方法,实现了室内无人机的高精度定位。该方法首先利用麦克风阵列采集声信号,结合惯性测量单元(Inertial measurement unit, IMU)获取运动数据,构建声学-惯性融合测量模型,基于卡尔曼滤波器对运动状态进行递推估计,并用四分位距(Interquartile range, IQR)规则剔除异常值。实验结果表明,该方法的平均定位误差较超宽带技术降低46%,且可用于任意大小和布局的室内环境。

2.2 几何声学定位模型

几何声学定位模型的核心思想是利用声波在环境中的传播特性,通过对直达、反射和衍射路径的建模来推算声源位置。该类方法通常依赖环境几何信息并结合射线追踪等技术,将声波传播近似为沿直线路径的能量传递过程,并通过交点计算、传播时延或入射角等参数进行定位。与仅依赖直达声的方法相比,几何声学模型能够同时刻画反射与衍射效应,从而在复杂遮挡场景中为非视域声源定位提

供物理可解释且直观有效的途径。如图3所示,这是一种基于反射和衍射的NLOS SSL场景。

2.2.1 基于反射信号的非视域声源定位方法

该类方法采用射线跟踪技术,可有效模拟声波的直达传播和反射路径,通过计算声射线的交点来估计声源位置。Kallakuri等^[43]将被动声学定位与环境的三维几何信息结合,利用射线投射法评估声波的传播和反射,以便机器人检测非视域内的移动物体。然而,多次反射后,该算法的声射线投射精度可能下降。Yoshinari等^[44]提出一种基于反射声的到达方向估计与光线追踪技术结合的方法,以定位人的脚步声,其平均距离误差为0.84 m,但前一次脚步声的混响会影响后续脚步声的检测。Xie等^[45]基于自然反射路径的发射角、到达角和相对飞行时间建立定位模型,并用最小二乘法估计目标位置的闭式解,但在角度和时间误差增大时,定位精度会降低。Inkyu等^[46]考虑了直接声与间接声信号,利用反向声线跟踪技术生成并跟踪直达声和反射声路径,同时结合蒙特卡洛定位(Monte Carlo localization, MCL)估计三维声源位置,平均距离误差为0.8 m。为提升算法的鲁棒性,Inkyu等^[47]假设声源位置附近不同声线路径的反向传播信号相似,通过测量反向传播信号之间的相似性来验证待定的声源位置,将平均定位误差减小到0.51 m。肖祥文^[48]分析了声反射,应用基于Refect加权函数的广义互相关时延估计算法,通过构造位置相关的非线性方程组并用列文伯格-马夸尔特算法求解来计算声源位置,其延时估计误差小于0.1 ms,定位距离误差为2~9 cm,但该方法的定位误差随声源与阵列的距离以及声源与阵列离墙面的距离的增大而增大。

2.2.2 基于衍射信号的非视域声源定位方法

在前述反射定位方法中,主要研究了单声源的定位技术,然而当多个声源同时存在时,通过声线交叉点识别声源位置的过程将变得复杂。声射线数量的增加,必将延长计算交叉点所需时间,这给实时计算声线交点带来了挑战^[15]。此外,此类反射假设声波沿直线传播,主要适用于高频声的镜面反射,而对低频声的模拟效果相对欠佳,特别是在衍射现象(即当障碍物尺寸与波长相当时的散射)起主导作用时。实际中,由于衍射频繁发生,利用衍射现象进行非视域声源定位也获得关注。

Inkyu等^[49]将所设计的算法与UTD模型结合,预先计算室内场景重建网格的楔形物以生成衍射声线,然后通过粒子滤波器判别生成声射线的收敛区域是否为估计的声源位置。该方法相较于不考虑衍射的算法,定位精度提高了37%~130%,但对于低频声源其定位依然存在精度不足的问题。Zhai等^[50]提出一种结合快速边缘化块稀疏贝叶斯学习与二阶边缘衍射的声源定位算法,利用BTM模型计算二阶边缘衍射传递函数,实现了高分辨率定位。肖祥文^[48]提出了一种基于声波衍射的非视域声源定位方法。该方法结合到达时间差和多基阵加权技术,通过声衍射点位置信息的求解,依据衍射点与场景空间几何关系来估计被遮挡机器人的初始位置,并利用K均值聚类的密度加权算法进行优化,最终得到最佳定位估计。该方法的定位误差在2~8 cm之间,体现了声衍射在非视域定位中的应用优势。

2.2.3 基于反射和衍射结合的非视域声源定位方法

Takami等^[51]提出了一种基于反射与衍射相结合的非视域声源跟踪方法。该方法利用麦克风阵列采集反射和衍射的观测值,以生成联合声学观测似然,其在未知环境中的目标定位误差小于1.5 m。

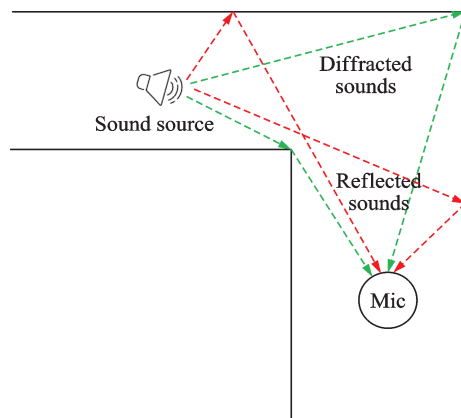


图3 基于反射和衍射的NLOS SSL场景

Fig.3 Reflection and diffraction-based NLOS SSL scene

Inkyu等^[52]提出一种声线追踪算法建模声音的高阶反射行为。该算法结合UTD模型,用粒子滤波器确定生成声射线的收敛区域,将其作为估计的声源位置,其单声源定位误差为0.7 m。在此基础上,他们进一步改进算法,在预计算阶段结合同步定位与地图构建(Simultaneous localization and mapping, SLAM)和基元拟合技术重建3D环境。当声线满足衍射准则时,基于UTD模型产生衍射声线,识别声线的会聚区域,实现多声源定位^[15]。改进后算法的单源和多源定位误差分别为0.65 m和0.74 m,但当障碍物表面光滑时定位精度下降,且混响对定位效果影响较大。

2.3 统计声学定位模型

与上述基于几何声学模型的方法不同,基于统计模型的方法并非简单地计算几何反射路径,而是强调环境中各类声波传播路径的统计特性。统计声学模型将声波传播视为随机过程,通过声源位置与接收信号之间的概率关系实现定位。如图4所示,这是一种基于概率估计的NLOS SSL场景。

Kumon等^[53]通过计算新观测的耳间声级差(Interaural level difference, ILD)与已存储ILD的相关性来估计声学观测似然,然后融合光学观测似然,构建联合观测似然,最后用递推贝叶斯估计更新目标的置信度。Takami等^[54]将该算法应用于跟踪医疗保健服务的老年人,在83%的位置实现了15 cm以内的定位精度。Takami等^[55]进一步提出在递推贝叶斯估计框架内,将光学与声学传感器的观测结果融合,在复杂室内场景中其平均定位误差小于2 cm。在此之前,Takami等^[56]也提出利用多传感器阵列测量反射时间差(Time difference of arrival, TDOA)以及与频率相关的衍射效应,并结合联合声学观测似然来优化定位过程。其主要目标是在递推贝叶斯估计框架下对非视域目标进行精确定位,但其仅支持单源定位。Zhai等^[50]用BTM模型计算二阶边缘衍射传递函数,对声源信号与麦克风阵列接收信号之间的关系建模,并基于声源信号的空间稀疏性,利用块稀疏模型实现声源定位。Sasaki等^[57]提出一种基于麦克风阵列的三维声源定位系统,该方法结合LiDAR(Light detection and ranging)生成环境的三维地图,在6个自由度内估计传感器的位置以及移动机器人在三维空间中的波达方向,并用MCL估计声源位置,两麦克风的测距误差分别为0.28 m和0.62 m。此外,Boger-Lombard等^[58]通过多个检波器对宽频噪声相关性进行检测,在混响环境中实现了拐角处人群的定位与跟踪。

2.4 基于手机的非视域声源定位模型

前文所述的定位模型大多依赖专用硬件、复杂环境建模或高算力算法,通常适用于机器人导航等场景。本节方法则以智能手机为核心载体,旨在实现非视域场景中的低成本、易部署的定位。该方法与前几类方法的本质区别体现在硬件依赖、算法复杂度和场景适配性等多个维度,而非仅仅是载体的不同。

(1) 针对手机硬件缺陷的信号鲁棒性优化

智能手机内置麦克风存在精度较低、易受机身噪声和环境混响影响的问题,且缺乏专用传感器的同步机制,导致传统的ToA/TDOA方法难以直接应用。一些研究聚焦于硬件缺陷补偿算法的设计。Haddad等^[59]提出了一种基于稳健飞行时间估计(Time of flight, TOF)和滑动窗口探测信号结构搜索的非视域声源定位方法。该方法通过滑动窗口机制锁定有效的探测信号结构,并结合几何约束修正

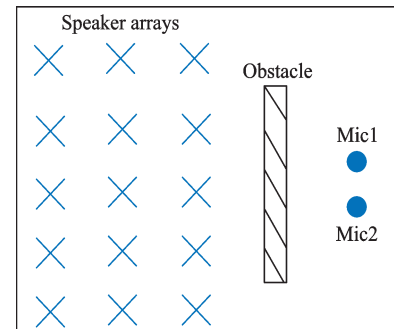


图4 基于概率估计的NLOS SSL场景

Fig.4 Probability-estimation-based NLOS SSL scene

TOF 误差,从而避免了对额外专用同步设备的依赖。Cheng 等^[60]提出了一种基于 Chirp 信号和两步式 ToA 检测的非视域声源定位方法。针对手机麦克风信号的低信噪比,预处理阶段采用带通滤波器滤除低频噪声,ToA 估计阶段通过“粗检测确定噪声范围-精检测捕捉非视域直达声”的分步策略,显著增强了多径干扰抑制能力。

(2) 手机端多模态融合降低算法复杂度

另一些算法的核心目标是适配手机实时算力,通过多模态融合降低算法复杂度。Yang 等^[61]提出了一种声学-惯性融合框架,结合手机内置的 IMU 与声传感器数据,无需外部高精度惯性设备。该算法在存在非视域错误数据时,通过 IMU 轨迹预测修正声学定位偏差,计算过程能够适配手机的 CPU 算力,且不依赖于额外硬件。Wang 等^[62]则提出了一种参数自更新卡尔曼滤波策略,针对手机端数据可能丢失的问题,设计了最小二乘反馈和贝叶斯参数估计的双融合机制,在数据部分缺失时通过 IMU 趋势预测调整滤波参数,避免了复杂的多传感器数据关联逻辑。

(3) 利用手机移动性的场景适配策略

智能手机的用户携带移动性是与前几类模型的一个关键区别,部分研究利用这一特性解决非视域场景中的动态定位问题,并且不需要预构建高精度环境地图。Bordoy 等^[63]提出了虚拟接收器动态建模的方法,针对手机移动过程中环境遮挡变化的问题,将非视域场景的声反射视为动态虚拟接收器,通过联合概率数据关联滤波器轻量化地建模虚拟接收器与 TDOA 的关系,无需预构建室内三维几何地图。Wang 等^[64]提出了动态定位优化方法,利用手机随用户移动的特性,在非视域场景中通过声学测距与行人航位推算的动态融合,无需固定锚点覆盖,即可在大范围移动过程中实现精确定位。

2.5 其他定位非视域声源的模型

Mita 等^[65]应用复杂反射声的时间特征-几何声音轮廓(Geometric sound profile, GSP),结合卷积神经网络,以较少的训练数据实现了声源定位。Zhang 等^[66]在短时间位置序列中联合估计多个位置,并利用利文贝格-马夸特(Levenberg-Marquardt, LM)算法进行优化求解,提出了一种多位置联合估计方法,实现了密集非视域环境下的精确定位。这种基于距离和相对速度的定位方法优于传统的基于距离的定位方法,并可通过调整序列长度来平衡定位精度和算法运行时间。

3 室外非视域声源定位方法

国内外对于室外非视域声源定位的研究内容虽然不多,但也取得了一定的成果,尤其是对自动驾驶汽车的非视域声源定位。

King 等^[10]针对自动驾驶应用,基于视听技术定位非视距声源。该方法先估计声源的波达方向(Direction of arrival, DoA)以及反射的图像源,然后将所有 DoA 估计与反向射线追踪相结合以估计声源位置,定位误差约为 0.5 m。Chen 等^[6]将非视距环境下匀速移动声源的定位问题,转化为由声衍射、多普勒频移和几何定位理论构建的方程,用多传感器融合算法减少估计误差,提出一种定位车辆发动机声音的方法。该方法对速度小于 36 km/h 的车辆,定位误差小于制动距离的 5%,弯道定位误差小于制动距离的 1%,但平均定位误差随车速的增加而增大。Hao 等^[8]基于深度学习实现了非视距车辆声检测。该方法将声音的到达方向和时频特征作为输入,用神经网络感知十字路口被遮挡的移动车辆的方向,以区分车的接近和离开状态。实验结果表明,该方法对 6 类交叉路通状况的分类准确率高达 96.71%,可在车辆进入视线前 1 s 对其进行检测,为驾驶员提供了额外的反应时间。Jeon 等^[9]首先结合声学 and 空间伪似然的粒子滤波,然后用鸟眼视图相机提供非视距车辆的实时位置作为地面实况,最后用所建立的隐形目标自我定位数据集对所提方法进行验证,实现了非视域环境下车辆声的定位。

4 总结与展望

根据具体的室内场景,选择合适的定位模型至关重要。不同的遮挡声源定位模型各具优势,但也都存在一定的局限,具体如表1所示。对于障碍物较少的环境,可通过简单的几何声学模型实现遮挡声源定位;对于复杂的、多样化的环境,统计声学模型则更合适。但若想在复杂环境中更好地实现精确的声源定位,多模态传感器定位模型是不二之选。

表1 不同声源定位方法优缺点总结
Table 1 Summary of advantages and disadvantages of different sound source localization methods

方法原理	优点	缺点
基于声信息测量的定位模型	基于 ToA 简单 ^[67] ,适用范围广 ^[34]	受背景噪声影响 ^[33] ,对同步要求严格 ^[35]
	基于 AoA 无需时钟同步 ^[38-39,68] ,可与波束形成结合 ^[38]	麦克风阵列布局影响精度 ^[38] ,计算复杂度高 ^[38-39]
	惯性传感器和声学传感器结合 融合多源信息,提高精度和鲁棒性 ^[62-63]	硬件成本高 ^[60,62] ,实时性差 ^[41] ,系统复杂度高 ^[63]
几何声学定位模型	基于反射 适用于动态场景 ^[43] ,可利用环境反射特性辅助定位 ^[45]	精度随反射次数增加而下降 ^[45] ,依赖精确的环境建模 ^[46] ,易受噪声和混响影响 ^[44] ,对环境几何参数敏感 ^[46]
	基于衍射 适用于低频声信号 ^[49] ,能处理复杂环境 ^[15]	定位精度较低 ^[69] ,计算复杂度高 ^[50] ,难以处理高频声 ^[49-50]
	基于反射和衍射结合 融合优势 ^[15,48,52] ,精度较高 ^[51] ,对多遮挡环境有较好的适用性 ^[15]	计算开销大 ^[51] ,对周围环境的结构和材质有较高的依赖性 ^[15] ,对混响和噪声敏感 ^[52]
统计声学定位模型	贝叶斯估计 鲁棒性强 ^[51] ,精度可调 ^[53] ,可处理部分噪声干扰 ^[51] ,可适用不确定性较强的动态场景 ^[51]	计算复杂度高 ^[51,56] ,需大量训练数据或先验知识 ^[53] ,精度受限于模型 ^[54]

总而言之,尽管非视域声源定位技术研究已经取得了重要进展,但未来仍面临诸多挑战。特别是在以下几方面,定位算法的实际性能与期望差距较大,亟待突破其局限,弥补其不足:

(1)在多个声源同时发声的情况下,由于声源间的相互干扰,已有方法难以准确定位所有声源。为此,可在定位算法前端集成适合的声源分离算法,从而提高多声源定位能力。此外,室内障碍物导致的声波反射、衍射和散射等物理现象引起的“虚声源”导致声源定位的区分度较低,对此应利用粒子滤波、联合概率数据关联等技术,以便在众多候选位置中确定真实声源的位置。

(2)实际应用中,声源与麦克风节点同时移动的情况时常发生,而现有定位方法大多都没有考虑此种场景,此时其性能明显退化。未来可考虑研究配备麦克风阵列的移动机器人,进行实际的移动目标估计。

(3)在实际中感兴趣的待定位声源有时是瞬态声源,如枪声、爆炸声和鸣笛声等。由于信号持续时间有限,现有定位方法大多对瞬态声源的定位性能较差,甚至无法提供有效的定位结果。

(4)部分定位算法计算复杂度过高,无法实现实时定位,不利于实时应用场景。此时,需对算法重构或优化,减少其计算复杂度。此外,也可考虑使用专用的硬件加速器,对算法进行定制化的硬件设计,提高算法的运行速度。

(5)由于UTD模型是近似模型,在不同的环境中可能不同,使得基于此模型的定位算法对低频声音的定位准确性较低。此外,UTD模型方法在障碍物表面光滑情况下定位精度降低,而衍射模型方法能提高精度但不易实时处理。因此,需要开发新定位算法,以平衡定位精度和计算复杂度。

(6)动态环境下IID数据库的自动更新机制有待完善,未来可结合双耳时间差(Interaural time difference, ITD)和双耳相位差(Interaural phase difference, IPD)来增强多模态声学感知能力,以适应复杂非视距场景的声源定位需求。

(7)针对声波传播过程中路径损耗导致有效作用距离受限问题,需通过优化信号处理算法(如多径补偿、分集接收)或改进硬件设计(如高灵敏度麦克风阵列)等,突破现有系统在大范围场景下的定位能力瓶颈。

(8)针对复杂应用场景,可采用多模态传感器和数据融合算法,以进一步提升定位系统的性能。例如,可考虑采用声学传感器、视觉传感器(如LiDAR/摄像头)、惯性导航等多类型传感器,构建多模态信息互补的协同定位系统,从而有效应对复杂环境中的遮挡、混响等干扰,有效提升非视距声源定位的精度与鲁棒性。

除算法外,外部环境也对定位结果产生一定程度的干扰:

(1)室内混响和设备噪声的干扰,会严重影响定位算法的性能。为此,建议在定位算法前端集成合适的去噪算法,如自适应滤波器、盲源分离等,构建“降噪-定位”一体化系统,有效抑制环境噪声对定位性能的影响。

(2)环境中不同材质(如吸声棉、玻璃和木质等)的墙壁、天花板、地板和障碍物的反射系数存在差异,且声源类型(人声、脚步声和机械声等)的辐射模式也存在差异,可能导致反射或衍射传播模型失配。因而,需对环境声反射特性和声源辐射特性的建模误差进行精细化处理。对此,未来可建立不同材质的声学数据库,结合深度学习对不同材质进行精确分类并判断声源辐射模式,从而自动匹配反射或衍射系数。

(3)现有定位算法在预计算阶段对应用场景进行三维重建的瓶颈需要突破。通常,三维场景重建依赖SLAM或人工建模的三维网格,难以反映实时的环境变化(如人员移动或物体移动等)。未来可采用多传感器融合(LIDAR+RGBD相机)技术实时更新环境信息,结合动态网格优化算法提升场景表示的准确性,为声线追踪提供更可靠的几何约束。

参考文献:

- [1] YASSIN A, NASSER Y, AWAD M, et al. Recent advances in indoor localization: A survey on theoretical approaches and applications[J]. IEEE Communications Surveys & Tutorials, 2017, 19(2): 1327-1346.
- [2] LIU Manni, CHENG Linsong, QIAN Kun, et al. Indoor acoustic localization: A survey[J]. Human-Centric Computing and Information Sciences, 2020, 10(1): 1-24.
- [3] APARICIO J, ÁLVAREZ F J, HERNÁNDEZ Á, et al. A survey on acoustic positioning systems for location-based services [J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71: 1-36.
- [4] 李剑汶,章宇栋,童峰,等.利用块稀疏特性的压缩感知麦克风阵列声源定位[J].数据采集与处理,2019,34(4): 682-688.
LI Jianwen, ZHANG Yudong, TONG Feng, et al. Microphone array direction of arrival estimation based on block sparse feature[J]. Journal of Data Acquisition and Processing, 2019, 34(4): 682-688.
- [5] 何赛娟,陈华伟,尹明婕,等.基于差分麦克风阵列和语音稀疏性的多源方位估计方法[J].数据采集与处理,2015,30(2): 372-381.
HE Saijuan, CHEN Huawei, YIN Mingjie, et al. Direction finding of multiple sound sources based on sparseness of speech

- signals and differential microphone array[J]. *Journal of Data Acquisition and Processing*, 2015, 30(2): 372-381.
- [6] CHEN Zexin, HE Yanxiao, WANG Qingsong, et al. A sound source localization method under NLOS environment for vehicles[C]//*Proceedings of IEEE 4th International Conference on Electronics Technology (ICET)*. [S.l.]: IEEE, 2021: 790-795.
 - [7] CHOI B. Acoustic source localization in 3D complex urban environments[D]. Blacksburg: Virginia Polytechnic Institute and State University, 2012.
 - [8] HAO Mingyang, NING Fangli, WANG Ke, et al. Acoustic non-line-of-sight vehicle approaching and leaving detection[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2024, 25(8): 9979-9991.
 - [9] JEON M, CHO J K, KIM H Y, et al. Non-line-of-sight vehicle localization based on sound[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2025, 26(2): 2321-2338.
 - [10] KING E A, TATOGLU A, IGLESIAS D, et al. Audio-visual based non-line-of-sight sound source localization: A feasibility study[J]. *Applied Acoustics*, 2021, 171: 1-13.
 - [11] MIŠKOVIĆ M, VUKMIROVIĆ N, GOLUBOVIĆ D, et al. Method for direct localization of multiple impulse acoustic sources in outdoor environment[J]. *Electronics*, 2022, 11(16): 1-16.
 - [12] ZENG Xiangyang, CHRISTENSEN C L, RINDEL J H. Practical methods to define scattering coefficients in a room acoustics computer model[J]. *Applied Acoustics*, 2006, 67(8): 771-786.
 - [13] VORLÄNDER M, EMBRECHTS J J, GEETERE L D, et al. Case studies in measurement of random incidence scattering coefficients[J]. *Acta Acustica united with Acustica*, 2004, 90(5): 858-867.
 - [14] HEIMES A, PAN L, VORLÄNDER M. Numerical simulation and analysis of surface scattering part 1[C]//*Proceedings of 10th Convention of the European Acoustics Association Forum Acusticum 2023*. Turin, Italy: [s.n.], 2024: 851-856.
 - [15] INKYN A, JUNG-WOO C, DINESH M, et al. Diffraction- and reflection-aware multiple sound source localization[J]. *IEEE Transactions on Robotics*, 2022, 38(3): 1925-1944.
 - [16] SVENSSON U P, FRED R I, VANDERKOOY J. An analytic secondary source model of edge diffraction impulse responses [J]. *The Journal of the Acoustical Society of America*, 1999, 106(5): 2331-2344.
 - [17] ANTANI L, CHANDAK A, TAYLOR M, et al. Efficient finite-edge diffraction using conservative from-region visibility[J]. *Applied Acoustics*, 2012, 73(3): 218-233.
 - [18] SCHISLER C, MEHRA R, MANOCHA D. High-order diffraction and diffuse reflections for interactive sound propagation in large environments[J]. *ACM Transactions on Graphics*, 2014, 33(4): 1-12.
 - [19] TAYLOR M, CHANDAK A, REN Z, et al. Fast edge-diffraction for sound propagation in complex virtual environments[C]//*Proceedings of the EAA Symposium on Auralization*. Espoo, Finland: [s.n.], 2009: 1-6.
 - [20] TAYLOR M, CHANDAK A, QI Mo, et al. Guided multiview ray tracing for fast auralization[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2012, 18(11): 1797-1810.
 - [21] CALAMIA P T, SVENSSON U P. Fast time-domain edge-diffraction calculations for interactive acoustic simulations[J]. *EURASIP Journal on Advances in Signal Processing*, 2006, 2007(1): 1-10.
 - [22] KOUYOUJIAN R G, PATHAK P H. A uniform geometrical theory of diffraction for an edge in a perfectly conducting surface[J]. *Proceedings of the IEEE*, 1974, 62(11): 1448-1461.
 - [23] LOKKI T, PULKKI V, CALAMIA P T. Measurement and modeling of diffraction from an edge of a thin panel[J]. *Applied Acoustics*, 2008, 69(9): 824-832.
 - [24] KAN Ruixiang, WANG Mei, LIU Xin, et al. An advanced artificial fish school algorithm to update decision tree for NLOS acoustic localization signal identification with the dual-receiving method[J]. *Applied Sciences*, 2023, 13(6): 1-26.
 - [25] WU Chi, HOU Hongwei, WANG Wenjin, et al. TDOA based indoor positioning with NLOS identification by machine learning[C]//*Proceedings of 10th International Conference on Wireless Communications and Signal Processing (WCSP)*. Hangzhou, China: [s.n.], 2018: 1-6.
 - [26] BAI Xujing, ZHANG Lei, YANG Tong, et al. Semi-supervised learning based acoustic NLOS identification for smartphone

- indoor positioning[C]//Proceedings of IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC). Dalian, China: IEEE, 2019: 1-6.
- [27] XUE Wentao, HU Zhixin, WANG Nan, et al. Unsupervised learning based acoustic NLOS identification for smart phone indoor positioning[C]//Proceedings of IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC). Macau, China: IEEE, 2020: 1-6.
- [28] KAN Ruixiang, WANG Mei. LOS and NLOS signal classification based on advanced particle swarm optimization for acoustic self-calibrating indoor localization[C]//Proceedings of 4th International Conference on Intelligent Information Processing (IIP). Beijing, China: [s.n.], 2019: 341-345.
- [29] KAN Ruixiang, WANG Mei, ZHOU Zou, et al. Acoustic signal NLOS identification method based on swarm intelligence optimization SVM for indoor acoustic localization[J]. *Wireless Communications and Mobile Computing*, 2022(1): 1-20.
- [30] CAO Shuai, CHEN Xiang, ZHANG Xu, et al. Effective audio signal arrival time detection algorithm for realization of robust acoustic indoor positioning[J]. *IEEE Transactions on Instrumentation and Measurement*, 2020, 69(10): 7341-7352.
- [31] LUECKEN H, WITTNEBEN A. UWB radar imaging based multipath delay prediction for NLOS position estimation[C]//Proceedings of IEEE International Conference on Ultra-Wideband (ICUWB). Bologna, Italy: IEEE, 2011: 101-105.
- [32] ZHANG Boliang, SHEN Lu, YAO Jiahua, et al. Automatic tracking based on weighted fusion back propagation in UWB for IoT devices[J]. *Sensors*, 2024, 24(4): 1-24.
- [33] XIONG W X, SCHINDELHAUER C, SO H C, et al. Maximum correntropy criterion for robust TOA-based localization in NLOS environments[J]. *Circuits, Systems, and Signal Processing*, 2021, 40(12): 6325-6339.
- [34] LIN C K, TOMONARI F. A ToA-based approach to NLOS localization using low-frequency sound[C]//Proceedings of Australian Conference on Robotics & Automation. Auckland, New Zealand: [s.n.], 2006: 1-9.
- [35] LIN C K, TOMONARI F. Non-line-of-sight localization of a controlled sound source[C]//Proceedings of IEEE/ASME International Conference on Advanced Intelligent Mechatronics. Singapore: IEEE, 2009: 475-480.
- [36] ZUO Wenbin, YANG Weiting, HU Zhixin, et al. Acoustic fingerprint based smart mobiles indoor localization under dense NLOS environment[C]//Proceedings of IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC). Xi'an, China: IEEE, 2021: 1-6.
- [37] ZHANG Lei, LIU Wenlei, ZHOU Jingyi, et al. TOA and FOA based acoustic indoor localization in NLOS environment[C]//Proceedings of IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC). Zhengzhou, China: IEEE, 2023: 1-6.
- [38] WANG Mei, SUN Wei, QIU Lili. MAVL: Multiresolution analysis of voice localization[C]//Proceedings of 18th USENIX Symposium on Networked Systems Design and Implementation (NSDI). [S.l.]: [s.n.], 2021: 845-858.
- [39] XIONG W X, BORDOY J, GABBRIELLI A, et al. Two efficient and easy-to-use NLOS mitigation solutions to indoor 3D AOA-based localization[C]//Proceedings of International Conference on Indoor Positioning and Indoor Navigation (IPIN). Lloret de Mar, Spain: [s.n.], 2021: 1-8.
- [40] 张驰. 基于声学传感器阵列的非视域声源目标探测与定位方法研究[D]. 长沙: 湖南大学, 2022.
- ZHANG Chi. Research on sound source detection and location method of non-line-of-sight object based on acoustic sensor array [D]. Changsha: Hunan University, 2022.
- [41] WANG Hucheng, LUO Xiaonan, ZHONG Yanru, et al. Acoustic signal positioning and calibration with IMU in NLOS environment[C]//Proceedings of Eleventh International Conference on Advanced Computational Intelligence (ICACI). Guilin, China: [s.n.], 2019: 223-228.
- [42] SUN Y M, WANG W G, MOTTOLA L, et al. Indoor drone localization and tracking based on acoustic inertial measurement [J]. *IEEE Transactions on Mobile Computing*, 2024, 23(6): 7537-7551.
- [43] KALLAKURI N, EVEN J, MORALES Y, et al. Using sound reflections to detect moving entities out of the field of view [C]//Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems. Tokyo, Japan: IEEE, 2013: 5201-5206.

- [44] YOSHINARI M, TAKAHASHI M, HIOKA Y. Non-line-of-sight footstep localization for autonomous mobile robots[C]//Proceedings of the 10th International Conference on Mechatronics and Robotics Engineering (ICMRE). Milan, Italy: [s.n.], 2024: 114-119.
- [45] XIE Liangbo, ZHANG Lin, TIAN Zengshan, et al. Indoor localization using reflection paths under non-line-of-sight environments[C]//Proceedings of International Conference on Microwave and Millimeter Wave Technology (ICMMT). Harbin, China: [s.n.], 2022: 1-3.
- [46] INKYU A, MYUNGBAE S, DINESH M, et al. Reflection-aware sound source localization[C]//Proceedings of IEEE International Conference on Robotics and Automation (ICRA). Brisbane, Australia: IEEE, 2018: 66-73.
- [47] INKYU A, DOHEON L, BYEONGHO J, et al. Robust sound source localization considering similarity of back-propagation signals[C]//Proceedings of IEEE International Conference on Robotics and Automation (ICRA). Paris, France: IEEE, 2020: 1574-1580.
- [48] 肖祥文. 基于声传感器的被遮挡机器人定位方法研究[D]. 长沙: 湖南大学, 2023.
XIAO Xiangwen. Research on the localization method of occluded robot based on acoustic sensors[D]. Changsha: Hunan University, 2023.
- [49] INKYU A, DOHEON L, JUNE-WOO C, et al. Diffraction-aware sound localization for a non-line-of-sight source[C]//Proceedings of International Conference on Robotics and Automation (ICRA). Montreal, Canada: [s.n.], 2019: 4061-4067.
- [50] ZHAI Qingbo, NING Fangli, WEI Juan, et al. Non-line-of-sight sound source localization based on block sparse Bayesian learning and second-order edge diffraction[J]. *Applied Acoustics*, 2025, 228: 1-10.
- [51] TAKAMI K, LIU H, MAKOTO K, et al. Recursive Bayesian estimation of NFOV target using diffraction and reflection signals[C]//Proceedings of 19th International Conference on Information Fusion. Heidelberg, Germany: [s.n.], 2016: 1-8.
- [52] INKYU A, JUNE-WOO C, DINESH M, et al. Reflection and diffraction-aware sound source localization[C]//Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops. [S.l.]: CVPR, 2019: 13-16.
- [53] KUMON M, KIMOTO D, TAKAMI K, et al. Bayesian non-field-of-view target estimation incorporating an acoustic sensor [C]//Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems. Tokyo, Japan: IEEE, 2013: 3425-3432.
- [54] TAKAMI K, FURUKAWA T, KUMON M, et al. Non-field-of-view acoustic target estimation in complex indoor environment[J]. *Field and Service Robotics*, 2016, 113: 577-592.
- [55] TAKAMI K, FURUKAWA T, KUMON M, et al. Estimation of a nonvisible field-of-view mobile target incorporating optical and acoustic sensors[J]. *Autonomous Robots*, 2016, 40(2): 343-359.
- [56] TAKAMI K, FURUKAWA T, KUMON M, et al. Non-field-of-view indoor sound source localization based on reflection and diffraction[C]//Proceedings of IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI). San Diego, CA, USA: IEEE, 2015: 59-64.
- [57] SASAKI Y, TANABE R, TAKEMURA H. Probabilistic 3D sound source mapping using moving microphone array[C]//Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Daejeon, Korea: IEEE, 2016: 1293-1298.
- [58] BOGER-LOMBARD J, SLOBODKIN Y, KATZ O. Non-line-of-sight passive acoustic localization around corners[J]. *Scientific Reports*, 2023, 13(1): 1-6.
- [59] HADDAD D B, MARTINS W A, DA COSTA M DO V M, et al. Robust acoustic self-localization of mobile devices[J]. *IEEE Transactions on Mobile Computing*, 2016, 15(4): 982-995.
- [60] CHENG Bingbing, HUANG Ying, ZOU Chuanyi. Robust indoor positioning with smartphone by utilizing encoded chirp acoustic signal[J]. *Sensors*, 2024, 24(19): 1-25.
- [61] YANG H, ZHANG R, BORDOY J, et al. Smartphone-based indoor localization system using inertial sensor and acoustic transmitter/receiver[J]. *IEEE Sensors Journal*, 2016, 16(22): 8051-8061.
- [62] WANG Hucheng, XUE Can, WANG Zhi, et al. Smartphone-based pedestrian NLOS positioning based on acoustics and IMU

- parameter estimation[J]. IEEE Sensors Journal, 2022, 22(23): 23095-23108.
- [63] BORDOY J, SCHINDELHAUER C, HÖFLINGER F, et al. Exploiting acoustic echoes for smartphone localization and microphone self-calibration[J]. IEEE Transactions on Instrumentation and Measurement, 2020, 69(4): 1484-1492.
- [64] WANG Hucheng, WANG Zhi, ZHANG Lei, et al. A highly stable fusion positioning system of smartphone under NLoS acoustic indoor environment[J]. ACM Transactions on Internet Technology, 2023, 23(2): 1-19.
- [65] MITA Y, MURAKAMI H, SASATANI T, et al. Geometric sound profile: Multipath-time-of-flight fingerprint for high-accuracy acoustic localization[C]//Proceedings of the 14th International Conference on Indoor Positioning and Indoor Navigation (IPIN). Hongkong, China: [s.n.], 2024: 1-6.
- [66] ZHANG Lei, WANG Hucheng, HE Wei, et al. Acoustic indoor localization based on range and relative velocity in non-line-of-sight environment[J]. IEEE Sensors Journal, 2024, 24(12): 19466-19481.
- [67] MAK L C, FURUKAWA T. Non-line-of-sight indoor localization of micro air vehicles using first-arrival sound[J]. International Journal of Micro Air Vehicles, 2010, 2(4): 221-237.
- [68] FISCHER G K J, SCHAECHTLE T, GABBRIELLI A, et al. A systematic survey and analysis of angular-based indoor localization and positioning[J]. IEEE Communications Surveys and Tutorials, 2026(28):3830-3869.
- [69] SÜDBECK S, KRAUSE T, OSTERMANN J. Non-line-of-sight time-difference-of-arrival localization with explicit inclusion of geometry information in a simple diffraction scenario[C]//Proceedings of IEEE 22nd International Workshop on Multimedia Signal Processing (MMSP). Tampere, Finland: IEEE, 2020: 1-6.

作者简介:



王硕(2002-),女,硕士研究生,研究方向:非视域声源定位、音频信号处理,E-mail:ws@mail.dlut.edu.cn。



陈喆(1975-),通信作者,男,教授,博士生导师,研究方向:音频信号处理、图像处理和宽带无线通信,E-mail:zhechen@dlut.edu.cn。



殷福亮(1962-),男,教授,博士生导师,研究方向:音频信号处理、图像处理和宽带无线通信。

(编辑:刘彦东)