

DOI: 10.19650/j.cnki.cjsi.J2514574

基于自适应卡尔曼滤波的动态声源追踪方法

刘昊¹, 李文国², 徐明², 王晨², 席旭刚¹

(1. 杭州电子科技大学智能控制与机器人研究所 杭州 310018; 2. 杭州奇点感知技术有限公司 杭州 310009)

摘要:针对复杂环境下低空无人机的动态轨迹难以高精度跟踪问题,提出了一种基于分布式麦克风阵列的残差驱动融合自适应卡尔曼滤波(RD-AKF)方法。该方法采用分布式麦克风阵列框架,通过两种不同维度互补的算法获取声源观测信息:基于几何关系的基线交汇定位法提供方向约束,基于信号能量的频域相位补偿波束形成(FD-PCB)算法提供能量空间分布特征。这两类在物理本质上互补的观测信息,被同步作为联合观测输入进一个统一的卡尔曼滤波框架。为克服固定参数模型在应对动态目标突发噪声时的脆弱性,提升追踪精度,该方法进一步在卡尔曼滤波中引入残差驱动的自适应机制,通过实时计算各观测通道的预测残差,并依据其统计特性动态重构观测噪声协方差矩阵,优化数据权重,实现声源位置与速度的联合估计。将该方法与单一的基线交汇、FD-PCB定位方法和固定协方差矩阵的标准卡尔曼滤波方法在相同环境下进行对比实验,结果表明,经过卡尔曼框架融合的追踪方法较单一算法均方根误差降低至少24.3%,引入残差驱动后,RD-AKF较标准卡尔曼滤波均方根误差降低了18.5%,最大定位误差降低了15.6%,并在保持合理计算成本的条件下,显著提高了动态声源追踪的精度与稳定性,为在复杂场景下动态声源轨迹追踪提供了高精度的、稳定的解决途径。

关键词:声源定位;自适应卡尔曼滤波;算法融合;分布式麦克风阵列

中图分类号: TH89 TB52 **文献标识码:** A **国家标准学科分类代码:** 510.40

Dynamic sound source tracking method based on adaptive Kalman filtering

Liu Hao¹, Li Wenguo², Xu Ming², Wang Chen², Xi Xugang¹

(1. Institute of Intelligent Control and Robotics, Hangzhou Dianzi University, Hangzhou 310018, China;
2. Hangzhou Qidian Sensing Technology Co., Ltd., Hangzhou 310009, China)

Abstract: Addressing the challenge of achieving high-precision tracking for low-altitude unmanned aerial vehicle (UAV) trajectories in complex environments, this paper proposes a residual-driven adaptive Kalman filter (RD-AKF) method based on a distributed microphone array. Within this framework, acoustic source observations are acquired through two complementary algorithms of different dimensions: a geometry-based triangulation method that provides directional constraints, and a frequency-domain phase compensation beamforming (FD-PCB) algorithm that captures energy spatial distribution characteristics. These physically complementary observations are integrated as joint measurements into a unified Kalman filtering framework. To mitigate the limitations of fixed-parameter models in handling abrupt noise from dynamic targets and to improve tracking accuracy, the method incorporates a residual-driven adaptive mechanism into the Kalman filter. This mechanism computes the prediction residuals of each observation channel in real-time and dynamically adjusts the measurement noise covariance matrix based on their statistical properties, thereby optimizing data weighting for the joint estimation of the acoustic source's position and velocity. Comparative experiments were conducted in an indoor environment by tracking the same low-altitude UAV flight trajectory, comparing the proposed method against standalone baseline intersection, FD-PCB localization, and standard Kalman filtering with fixed covariance matrices. The results show that the Kalman-fused approach reduces the root mean square error (RMSE) by at least 24.3% compared to individual algorithms. With the residual-driven adaptation, the RD-AKF further reduces the RMSE by 18.5% and the maximum positioning error by 15.6% compared to the standard Kalman filter, achieving significantly improved tracking accuracy and stability while maintaining reasonable computational cost. The proposed method provides a high-precision and robust solution for dynamic acoustic source tracking in complex scenarios.

Keywords: sound source localization; adaptive Kalman filter; algorithm fusion; distributed microphone array

0 引言

近年来,“低空经济”发展迅速,无人机的广泛应用使其在空域安全管理、环境监测及军事侦察等领域的声源定位需求日益增长^[1-2]。而声学研究中,动态声源轨迹追踪面临着诸多技术难题,声源运动过程中,声音信号易受多种因素影响,如声源与传感器之间的距离变化、环境噪声以及声源自身的特性变化等^[3-4],这些因素易导致声源信号获取困难,且结果极不稳定。

针对空间中的动态声源,单一平面阵列难以依靠传统的定位算法获得其空间信息,三维立体阵列需要遍历整个区域,搜索点数随分辨率立方级增长^[5],计算成本是巨大的。多年来学者们提出多种方法来追踪动态声源,Zhang等^[6]创造性地将声源定位到达方向(direction of arrival, DoA)估计与视觉同时定位与地图构建(visual simultaneous localization and mapping, SLAM)系统融合,以此减少动态物体对于追踪系统的干扰,解决“声源有方向但无深度”的经典难题;Verreycken等^[7]提出灵活的分布式拓扑阵列,并结合到达时间差(time difference of arrival, TDoA)和到达角(angle of arrival, AoA)信息用于追踪鸟类移动声源;刘望生等^[8]提出通过构建一种基于延时估计和波束能量的多特征优化机制,并辅以改进的交互式多模型粒子滤波算法,解决了二维平面内的鲁棒跟踪问题,但难以判断声源目标所处的高度;Chen等^[9]提出了基于概率数据关联的分布式容积卡尔曼滤波(probabilistic data association-based distributed cubature Kalman filter, PDA-DCKF)方法,通过概率数据关联筛选分布式阵列观测,并结合立方体卡尔曼滤波进行非线性状态估计,提升了在嘈杂环境下的动态声源跟踪精度;Thakur等^[10]提出了一种基于时间延迟和能量比的实时移动声源定位技术,突破了动态声源定位需要声源移速、所在平面等先验信息等限制。

前人研究证实结合多方信息是实现动态声源定位的关键,但仅依靠麦克风设备获取多模态信息是困难的,本文提出了一种基于分布式麦克风阵列的残差驱动融合自适应卡尔曼滤波(residual-driven fusion adaptive Kalman filter, RD-AKF)方法。

在算法实现上,分布式麦克风阵列系统使得基线交汇法得以应用于声源定位领域,通过测量声信号到达不同阵元的时延差,为定位系统提供可靠的几何约束^[11-12],结合由频域相位补偿的波束形成(frequency-domain phase-compensation beamforming, FD-PCB)算法获得与几何信息独立且互补的声压能量空间分布特征,输入基于历史残差统计的卡尔曼滤波(Kalman filter, KF)框架,动态调整观测权重,得到高精度定位结果。

在计算成本上,由多个独立子平面阵列构成的分布式麦克风阵列架构避免了构建实体大型三维阵列的高昂成本;支持基线交汇将复杂的三维定位问题转换为多个子阵列上的二维波达方向估计问题,实现了计算量的级次降低;支持FD-PCB中等效构建一个虚拟的大孔径三维阵列,且对搜索空间实现灵活覆盖,从而在系统层面优化了计算资源的使用。

1 算法原理

1.1 频域相位补偿的波束形成

波束形成(beamforming, BF)算法能有效地在混合信号中分离并找到目标信号^[13]。基于BF的FD-PCB方法融合子阵列通道数据,对声源位置进行独立估计。视户外声场为自由场,依赖分布式架构可根据先验信息确定声源可能存在的三维空间进行分辨率划分,候选位置网格点 $\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_K$ 。计算从假设声源点 \mathbf{q} 到每个麦克风 m 与到参考麦克风 $m=0$ 之间的TDoA,即:

$$\tau_m(\mathbf{q}) = \frac{|\mathbf{q} \cdot \mathbf{p}_m| - |\mathbf{q} \cdot \mathbf{p}_0|}{c} \quad (1)$$

式中: \mathbf{p}_m 是麦克风 m 的坐标; c 是声速。

同时将每个麦克风接收到的时域信号转换为频域信号 $X_m(f)$ 。将理论时延转换为频域中的相位补偿因子,构建导向向量 $\mathbf{v}(f, \mathbf{q})$ 为:

$$\mathbf{v}(f, \mathbf{q}) = [e^{j2\pi f \tau_0(\mathbf{q})}, \dots, e^{j2\pi f \tau_{M-1}(\mathbf{q})}]^T \quad (2)$$

式中: f 表示频率; M 表示阵列中麦克风阵元总数。将实际接收到的第 i 帧信号 $X_m^{(i)}(f)$ 与导向向量中对应元素的共轭相乘,即:

$$X'_m(f, \mathbf{q}) = X_m^{(i)}(f) \cdot e^{-j2\pi f \tau_m(\mathbf{q})} \quad (3)$$

即在数学上等效于将各通道信号在时间上“回溯”理论时延 τ_m ,从而在假设声源位于 \mathbf{q} 的前提下,使各通道信号趋于同相。将式(3)所有通道补偿后的信号相干叠加,得到该假设点 \mathbf{q} 处的波束形成器输出,即:

$$Y(f, \mathbf{q}) = \sum_{m=0}^{M-1} X'_m(f, \mathbf{q}) \quad (4)$$

计算总能量,作为该假设点 \mathbf{q} 的置信度得分,即:

$$P^{(i)}(\mathbf{q}) = \sum_f |Y(f, \mathbf{q}) \cdot W(f)|^2 \quad (5)$$

式中: $W(f)$ 是频率依赖性权重函数。遍历完所有网格点后,得到整个搜索空间的能量分布 $P(\mathbf{q})$ 。其全局最大值点,即为本帧信号估计的声源位置。

1.2 基于多重信号分类的基线交汇

基线交汇定位方法依赖可靠的DoA信息得到声源的几何定位约束。在高分辨率场景下,多重信号分类(multiple signal classification, MUSIC)算法的DoA更准确^[14]。采用MUSIC算法进行空间谱估计,对于单个五元

十字阵,使用维度为 5×1 的 \mathbf{x} 表示阵列接收到的多通道信号向量,对该向量构建协方差矩阵 \mathbf{R} ,并通过特征值分解划分信号子空间与噪声子空间:

$$\mathbf{R} = E[\mathbf{x}\mathbf{x}^T] = \mathbf{U}_s \mathbf{\Lambda}_s \mathbf{U}_s^T + \mathbf{U}_n \mathbf{\Lambda}_n \mathbf{U}_n^T \quad (6)$$

式中: $E[\mathbf{x}\mathbf{x}^T]$ 表示对随机变量 $\mathbf{x}\mathbf{x}^T$ 取统计平均; \mathbf{U}_s 和 \mathbf{U}_n 分别表示信号子空间和噪声子空间的特征向量矩阵; $\mathbf{\Lambda}_s$ 和 $\mathbf{\Lambda}_n$ 分别表示信号子空间和噪声子空间的特征值矩阵。计算空间普函数:

$$P_{\text{MUSIC}}(\theta, \phi) = \frac{1}{\mathbf{a}^T(\theta, \phi) \mathbf{U}_n \mathbf{U}_n^T \mathbf{a}(\theta, \phi)} \quad (7)$$

式中: (θ, ϕ) 表示入射方向, θ 为方位角, ϕ 为俯仰角; $\mathbf{a}(\theta, \phi)$ 为入射方向对应的阵列导向矢量。搜索空间谱最大值,即可获得声源的 DoA 估计。

通过将两组阵列所估计的 DoA 信息转化为空间方向矢量,进行三维几何定位^[15],完成基线交汇估计。

设子阵列 1 和子阵列 2 的参考坐标原点分别为 \mathbf{p}_1 和 \mathbf{p}_2 , 对应的 DoA 单位方向向量为:

$$\begin{cases} \mathbf{d}_1 = \begin{bmatrix} \cos \varphi_1 \cos \theta_1 \\ \cos \varphi_1 \sin \theta_1 \\ \sin \varphi_1 \end{bmatrix} \\ \mathbf{d}_2 = \begin{bmatrix} \cos \varphi_2 \cos \theta_2 \\ \cos \varphi_2 \sin \theta_2 \\ \sin \varphi_2 \end{bmatrix} \end{cases} \quad (8)$$

实际声源位置 \mathbf{s} 应同时位于两条方向射线上,理想情况下满足:

$$\mathbf{s} = \mathbf{p}_1 + \lambda_1 \mathbf{d}_1 = \mathbf{p}_2 + \lambda_2 \mathbf{d}_2 \quad (9)$$

式中: λ_1 和 λ_2 分别是各阵列中心沿方向向量到达声源点的距离参数; \mathbf{p}_1 和 \mathbf{p}_2 即各自阵列中心。由于噪声与估计误差,射线通常并不相交,计算两条射线间的最短距离,取中点作为声源估计位置,如图 1 所示。

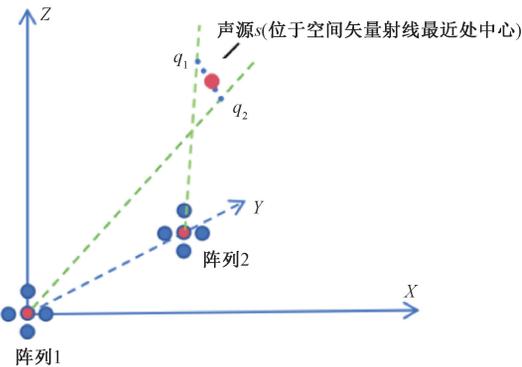


图 1 基线交汇几何示意图

Fig. 1 Diagram of baseline intersection geometry

射线最短距离对应的两个点 \mathbf{q}_1 和 \mathbf{q}_2 可通过求解最小化问题得到,即找到 λ_1^* 和 λ_2^* 作为 λ_1 和 λ_2 的最优解

满足式(10)。

$$\| \mathbf{q}_1(\lambda_1^*) - \mathbf{q}_2(\lambda_2^*) \| = \min_{\lambda_1, \lambda_2} \| \mathbf{p}_1 + \lambda_1 \mathbf{d}_1 - \mathbf{p}_2 - \lambda_2 \mathbf{d}_2 \| \quad (10)$$

从而找到求解两射线间的最短连线,并以其中点作为观测声源 \mathbf{s} ,即:

$$\mathbf{s} = \frac{1}{2}(\mathbf{q}_1 + \mathbf{q}_2) \quad (11)$$

该方法将复杂的三维空间搜索定位问题,分解为在子阵列上并行进行的、低维度的 DoA 估计问题,计算负担显著低于直接进行三维定位搜索。

1.3 融合基线交汇与波束形成的 RD-AKF 算法

建立如下状态空间模型,对于飞行无人机,定义状态向量为其在时刻 k 的位置和速度组成的六维向量:

$$\mathbf{x}_k = [x_k \ y_k \ z_k \ v_{x,k} \ v_{y,k} \ v_{z,k}]^T \quad (12)$$

式中: x_k, y_k 和 z_k 表示空间位置坐标; $v_{x,k}, v_{y,k}$ 和 $v_{z,k}$ 表示线速度。假设小时间步内无人机近似做匀速直线运动,则状态转移模型可表示为:

$$\mathbf{x}_k = \mathbf{F}\mathbf{x}_{k-1} + \mathbf{w}_{k-1} \quad (13)$$

式中: $\mathbf{F} = \begin{bmatrix} \mathbf{I}_3 & \Delta t \cdot \mathbf{I}_3 \\ \mathbf{0}_{3 \times 3} & \mathbf{I}_3 \end{bmatrix}$; $\mathbf{w}_{k-1} \sim N(\mathbf{0}, \mathbf{R}_k)$; \mathbf{Q} 表示过程噪声协方差。

观测模型的设计核心在于同时接纳来自如下两种互补的观测数据。

1) FD-PCB 定位融合两个子阵列通道,合并声音数据,利用时延差连接空间位置假设与信号处理验证,寻找最能解释实际信号间相位差的那个空间点作为估计位置,设为 \mathbf{z}_k^{bf} 。

2) 基线交汇定位利用一对阵列各自 DoA 估计结果构建方向向量,通过空间交汇最小化估计源点位置,设为 $\mathbf{z}_k^{\text{intersect}}$ 。两者均构成观测模型,即:

$$\mathbf{z}_k = \mathbf{H} \cdot \mathbf{x}_k + \mathbf{v}_k \quad (14)$$

$$\begin{cases} \mathbf{H} = [\mathbf{I}_3 \ \mathbf{0}] \\ \mathbf{v}_k \sim N(\mathbf{0}, \mathbf{R}_k) \end{cases} \quad (15)$$

式中: \mathbf{z}_k 表示观测位置; \mathbf{H} 为观测矩阵,仅取前 3 项作为测量项。引入观测融合思想,将两个观测视作独立测量,构造联合观测向量 $\tilde{\mathbf{z}}_k$ 与联合协方差矩阵 $\tilde{\mathbf{R}}_k$,即:

$$\begin{cases} \tilde{\mathbf{z}}_k = \begin{bmatrix} \mathbf{z}_k^{\text{bf}} \\ \mathbf{z}_k^{\text{intersect}} \end{bmatrix} \\ \tilde{\mathbf{R}}_k = \begin{bmatrix} \mathbf{R}_k^{\text{bf}} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_k^{\text{intersect}} \end{bmatrix} \end{cases} \quad (16)$$

为提高融合适应性,本文引入残差驱动的动态协方差调整机制,在每一时刻 k , 计算各观测项的残差,即:

$$\mathbf{e}_k^{(j)} = \mathbf{z}_k^{(j)} - \mathbf{H} \cdot \hat{\mathbf{x}}_k^-, \quad j \in \{\text{bf}, \text{intersect}\} \quad (17)$$

式中: $\mathbf{z}_k^{(j)}$ 为 k 时刻 j 项观测值; $\hat{\mathbf{x}}_k^-$ 为 k 时刻 j 项预测值。

基于残差模值对 $\mathbf{R}_k^{(j)}$ 进行自适应调整,更新为:

$$\mathbf{R}_k^{(j)} = \alpha \cdot \mathbf{I}_3 + \beta \cdot \mathbf{e}_k^{(j)} (\mathbf{e}_k^{(j)})^T \quad (18)$$

式中: α 和 β 为调节参数; $\mathbf{R}_k^{(j)}$ 是基于当前观测残差动态计算出来的协方差。 α 代表观测噪声的最小基准协方差,保证 $\mathbf{R}_k^{(j)}$ 始终正定; β 控制残差对协方差的调节权重。该机制通过增大异常观测的协方差,从而降低其在数据融合中的权重,抑制其对最终估计结果的影响。

由此卡尔曼滤波器的关键输入参数设计完成,卡尔曼滤波器的具体执行步骤如下。

预测步骤为:

$$\hat{\mathbf{x}}_k^- = \mathbf{F} \cdot \hat{\mathbf{x}}_{k-1} \quad (19)$$

$$\mathbf{P}_k^- = \mathbf{F} \cdot \mathbf{P}_{k-1} \cdot \mathbf{F}^H + \mathbf{Q}_k \quad (20)$$

更新步骤为:

$$\mathbf{K}_k = \mathbf{P}_k^- \cdot \tilde{\mathbf{H}}^T \cdot (\tilde{\mathbf{H}} \cdot \mathbf{P}_k^- \cdot \tilde{\mathbf{H}}^T + \tilde{\mathbf{R}}_k)^{-1} \quad (21)$$

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k^- + \mathbf{K}_k \cdot (\mathbf{z}_k - \tilde{\mathbf{H}} \cdot \hat{\mathbf{x}}_k^-) \quad (22)$$

$$\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k \cdot \tilde{\mathbf{H}}) \cdot \mathbf{P}_k^- \quad (23)$$

式中: $\tilde{\mathbf{H}} = \begin{bmatrix} \mathbf{H} \\ \mathbf{H} \end{bmatrix}$; \mathbf{P}_k 为时刻 k 的后验状态协方差矩阵;

\mathbf{P}_k^- 为时刻 k 的先验状态协方差矩阵; \mathbf{K}_k 为基于最小化估计误差的均方误差推导的卡尔曼增益矩阵,如式(24)所示。

$$\mathbf{K}_k = \mathbf{P}_k^- \cdot \tilde{\mathbf{H}}^T \cdot (\tilde{\mathbf{H}} \cdot \mathbf{P}_k^- \cdot \tilde{\mathbf{H}}^T + \tilde{\mathbf{R}}_k)^{-1} \quad (24)$$

可见观测噪声协方差矩阵 $\tilde{\mathbf{R}}_k$ 是增益公式(式(24))中的关键组成部分,直接影响 \mathbf{K}_k 的大小从而决定状态估计主要依赖于模型预测还是观测数据,间接影响误差协方差矩阵的更新:

$$\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k \tilde{\mathbf{H}}) \mathbf{P}_k^- \quad (25)$$

因此 $\tilde{\mathbf{R}}_k$ 直接影响卡尔曼滤波器的性能和定位结果,动态调整 $\tilde{\mathbf{R}}_k$ 更好地平衡观测数据和模型预测的权重,也可避免“欠拟合”^[16],在缺乏声音信号模型的情况下也能保持良好的性能。

整个融合框架以 KF 为基础,主要计算负担来自于式(20)和(23)矩阵乘法和式(21)的逆运算,复杂度为 $O(n^3)$,引入的残差驱动自适应机制仅增加 $O(m^2)$ 的额外计算,其中状态向量维度 $n = 6$,联合观测向量维度 $m = 6$ 。整体复杂度与 KF 相同,但通过动态调整观测权重减少了迭代次数。

2 实验和分析

2.1 实验环境

实验硬件平台为 Intel Core i9-14900K CPU,软件环境为 Python 3.9。采用基于型号 PGL50H-6IFBG484 FPGA 的 24 bit 音频采样阵列,支持 WiFi6 模块,可直接用于网络传输,设备以 5 个 MEMS 麦克风传感器排列成五元十字阵,阵元间距为 15.3 mm,频率响应范围

为 5~11 kHz,采样率为 96 kHz,帧长为 1 024。麦克风阵列设备如图 2 所示。



图 2 麦克风阵列设备

Fig. 2 Microphone array device illustration

实验以一台多旋翼无人机作为动态声源目标进行数据采集时的飞行状态如图 3 所示。



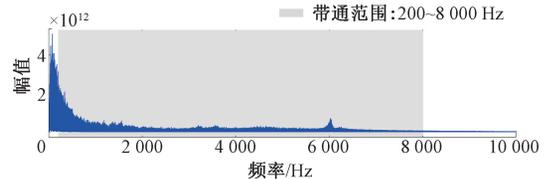
图 3 声源无人机及声音采样场景

Fig. 3 Demonstration of the drone and scene

2.2 声源信号处理

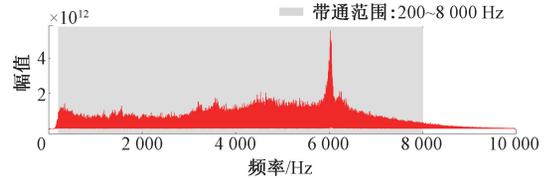
为深入探究多旋翼无人机的声源特性,实验对无人机设备静止悬停状态下进行实地音频数据采集。

原始信号频谱图如图 4(a)所示,可见其在低频段 (< 1 kHz) 具有极强能量集中性,同时在 6 和 15.5 kHz 附近存在明显的窄带调制峰值,高频部分虽能量较低但仍可感知显著特征峰。



(a) 原始信号频谱

(a) Original signal spectrum



(b) 处理后信号频谱

(b) Processed signal spectrum

图 4 声信号处理

Fig. 4 Acoustic signal processing

对每个通道信号的原始波形信号去直流操作,取 200~8 000 Hz 的带通范围。取预加重系数为 0.95,进行一阶预加重处理后信号频谱如图 4(b)所示,可见预处理有效抑制了低频干扰,使无人机噪声 5 kHz 附近关键频段能量集中。

2.3 定位算法精度与执行时间分析

验证算法性能,在信噪比(SNR)为 25 dB,声速 343 m/s 的实验室内,放置将阵列在全局坐标系下,阵列参考麦克风坐标分别为 (2.82 m, 0.00 m, 0.79 m), (1.60 m, 6.50 m, 0.81 m)。阵列 1 面向 Y 轴正方向,阵列 2 面向 Y 轴负方向,无偏转。悬停声源无人机设备在全局坐标系 (3.60 m, 2.30 m, 1.20 m) 处。

阵列面竖直放置如图 5 所示,常规方位角与俯仰角定义容易混淆,实验定义“角度 1”,即 $\angle a$,为局部坐标系下声源方向在 XOZ 面投影 S_2 与 X 轴正方向夹角 ($0^\circ, 180^\circ$);“角度 2”,即 $\angle b$,为局部坐标系下声源方向 S 在 YOZ 面投影 S_3 与 Y 轴正方向夹角 ($0^\circ, 180^\circ$)。

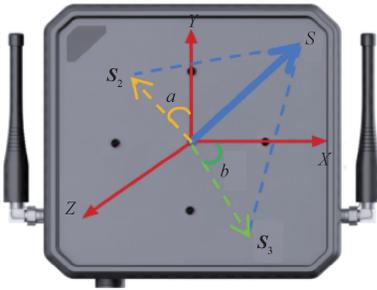


图 5 阵列面局部坐标系

Fig. 5 Local coordinate system on the array plane

MUSIC 算法输出定义角度。每个阵列计算出的 DoA 信息基于其局部坐标系,因此在融合定位前需进行坐标统一^[17]。阵列 1 的局部坐标系到全局坐标系的完整齐次变换矩阵为:

$$T_1 = \begin{bmatrix} 1 & 0 & 0 & 2.82 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0.79 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (26)$$

阵列 2 为:

$$T_2 = \begin{bmatrix} -1 & 0 & 0 & 1.6 \\ 0 & 0 & -1 & 6.5 \\ 0 & 1 & 0 & 0.81 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (27)$$

同一坐标系到全局基线交汇得到定位结果如图 6 所示。

图 6 中,实线箭头表示阵列 1 第 1 帧数据构建的 DoA 信息角 ($70^\circ, 85^\circ$);虚线箭头表示阵列 1 第 1 帧数据构建的 DoA 信息角 ($122^\circ, 117^\circ$)。叉点为真实位置 (3.60 m, 2.30 m, 1.20 m),星标为基线交汇估计结果 (3.72 m, 2.55 m, 1.12 m)。

在第 1 帧进行 FD-PCB 声源估计,搜索空间设置为 $[1.00 \text{ m}, 1.50 \text{ m}, 0.50 \text{ m}] \sim [5.00 \text{ m}, 4.00 \text{ m}, 2.00 \text{ m}]$,空间分辨率 0.10 m,带宽取 2 000 Hz,总网格点数约 120 000 个。

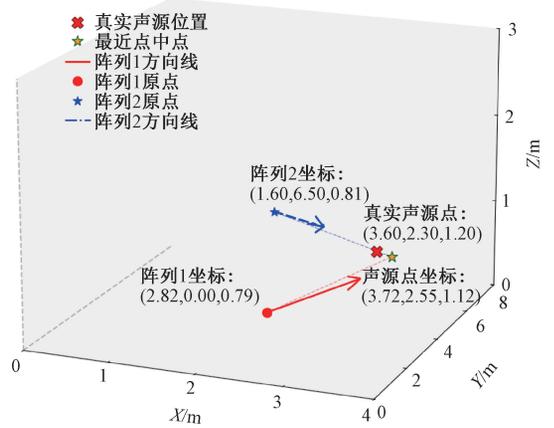


图 6 基线交汇定位第 1 帧

Fig. 6 Baseline intersection localization-first frame

图 7 所示为 FD-PCB 声源估计第 1 帧在 XY 面估计结果,能在各坐标面能量分布最高处找到星标作为声源估计点定位。可见两种不同的定位方法,在声源空间定位中均表现出一定的有效性。

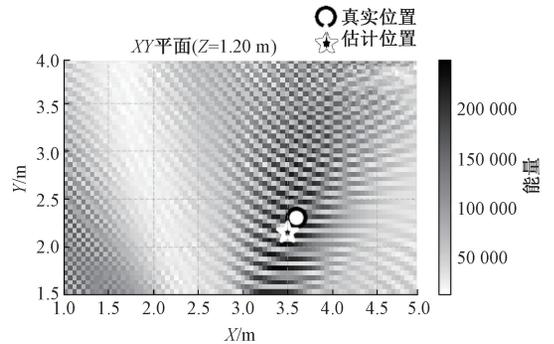


图 7 XY 平面视角能量分布

Fig. 7 Angular energy distribution in the XY-plane

全面评估 RD-AKF 方法性能,设计轨迹追踪实验。无人机飞行轨迹包括直线、曲线、晃动以及变速运动,全面检验算法在不同运动状态下的适应能力。

无人机飞行轨迹设定为从起点 (1.50 m, 2.00 m, 1.00 m) 至终点 (4.50 m, 3.50 m, 1.50 m),飞行高度维持在 1.0~2.0 m,飞行平均速度约为 0.5 m/s,整个轨迹持续时间为 10 s。在无人机上安装高精度 UWB 定位模块实现厘米级定位,以 100 Hz 频率记录真实轨迹。

提取 UWB 模块的三维轨迹数据,结果如图 8 所示。

设计如下实验作为对照组:

- 1) 基线交汇定位;
- 2) FD-PCB 定位;
- 3) 使用固定观测噪声协方差矩阵 KF 融合定位;
- 4) 本文提出的 RD-AKF 方法定位。

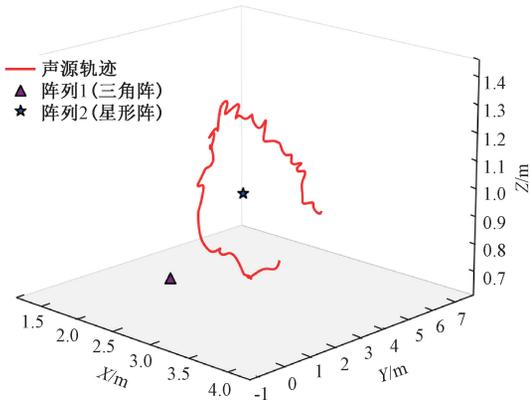


图8 UWB 模块提取的真实轨迹

Fig. 8 Real trajectory obtained from the UWB module

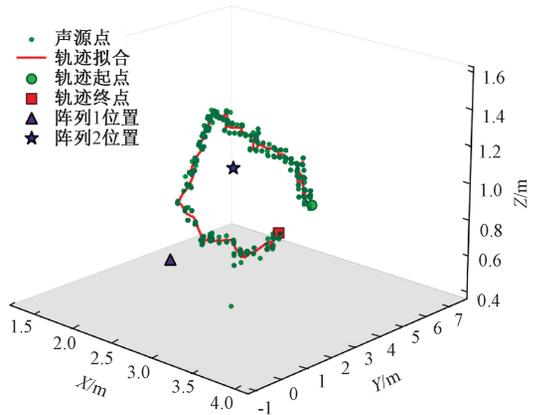


图10 基线交汇拟合轨迹结果

Fig. 10 Fitted trajectory from baseline intersection

采用追踪与真实位置之间的欧氏距离均方根误差 (root mean square error, RMSE) 和平均绝对误差 (mean absolute error, MAE) 作为评估指标。

处理双阵列采集的音频数据,通过互相关计算节点间的时间偏移实现严格的时间同步,截取对应的 912 帧数据进行实验测试。

图 9 所示为仅依靠基线交汇实现的全部定位点空间分布,使用基于密度的噪声应用空间聚类算法,将低密度区域中的点标记为噪声点,保留定位点分布密集的区域拟合轨迹,得到结果如图 10 所示。

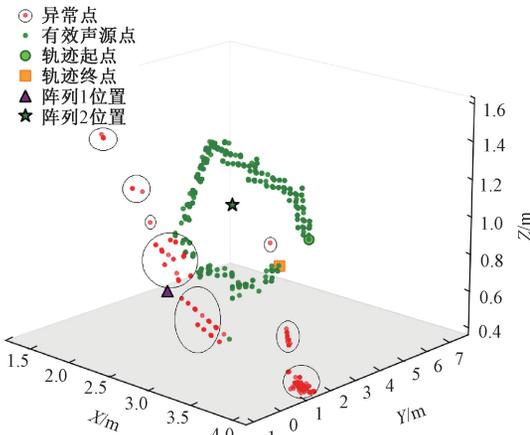


图9 基线交汇结果

Fig. 9 Results of baseline intersection

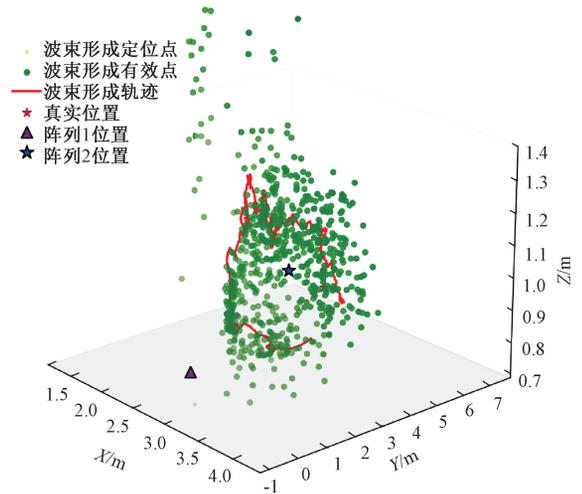


图11 双阵列融合波束形成定位结果

Fig. 11 Positioning results of FD-PCB

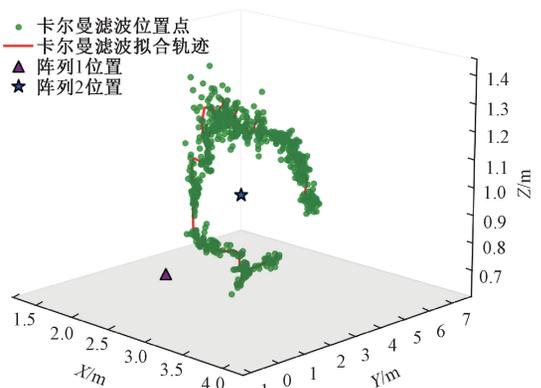


图12 KF 定位结果

Fig. 12 Fusion results of KF

对每一帧计算双阵列数据融合的 FD-PCB,寻找最大能量的空间网格,得到 912 个定位点分布结果如图 11 所示,使用同样的方法聚类并拟合得到轨迹。

融合两种定位结果,使用标准 KF 融合得到定位结果如图 12 所示, RD-AKF 方法得到定位结果如图 13 所示。

4 种定位对比实验结果被置于同一坐标系下如图 14 所示,轨迹线 1 为真实轨迹;轨迹线 2 对应基线交汇结

果;轨迹线 3 对应 FD-PCB 结果;轨迹线 4 对应 RD-AKF 结果;轨迹线 5 对应固定噪声协方差矩阵进行卡尔曼滤波融合结果。

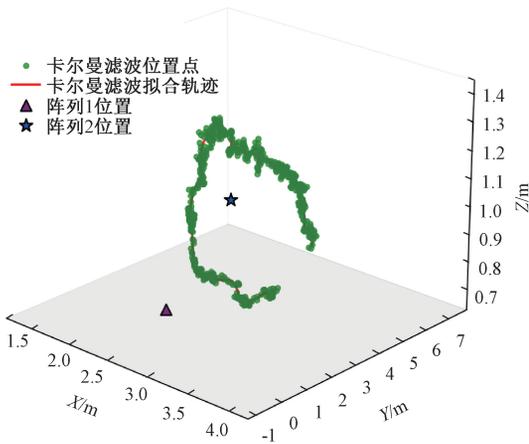


图 13 RD-AKF 定位结果

Fig. 13 Localization results of RD-AKF

可见 RD-AKF 轨迹几乎覆盖了真实轨迹线, 此时 $\alpha = 0.1, \beta = 1.0$, 其次与真实轨迹最接近的是固定噪声协方差矩阵进行卡尔曼滤波融合方法结果。

从 4 个维度分析说明算法融合的卡尔曼滤波定位方法可知, RD-AKF 在提升定位精度方面具有明显优势。

其中误差分布的统计如图 15 所示, 说明本文提出的 RD-AKF 方法在各项指标上均展现出显著优势。由

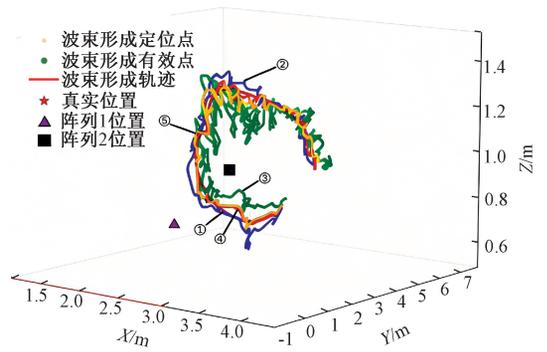
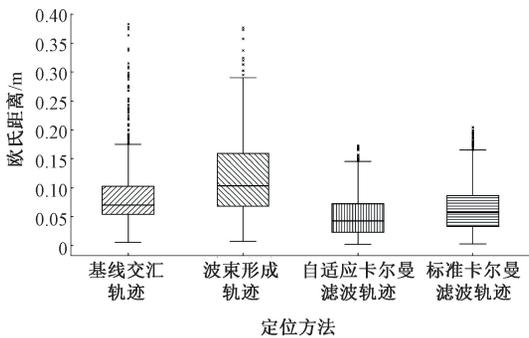


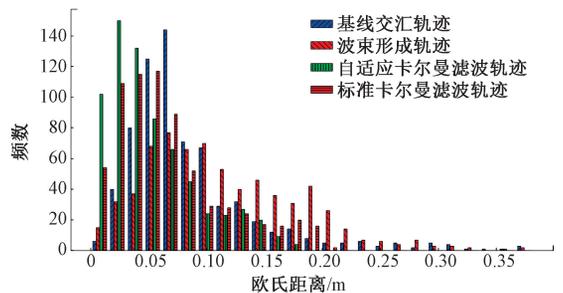
图 14 4 种定位结果与真实轨迹 (SNR=25 dB)

Figure. 14 Comparison of four localization results and the ground truth trajectory (SNR=25 dB)

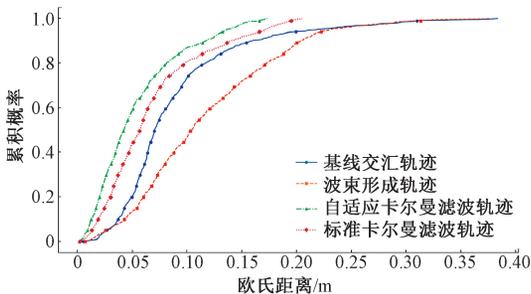
图 15(a) 可知, RD-AKF 的误差箱体最短、位置最低, 且无明显异常值, 说明其估计结果不仅精度高, 而且稳定性与鲁棒性最佳, 相比之下, 单一算法的误差分布离散, 存在大量高位异常值。由图 15(b) 和 (c) 可知, RD-AKF 的误差集中分布在区间 $[0.02 \text{ m}, 0.08 \text{ m}]$, 约 80% 的轨迹点误差低于 0.07 m , 呈现出高度集中的分布特性, 收敛速度远优于其他方法。由图 15(d) 可知, RD-AKF 的输出轨迹平滑性最好, 无明显抖动, 反映出滤波框架对观测噪声的有效抑制。



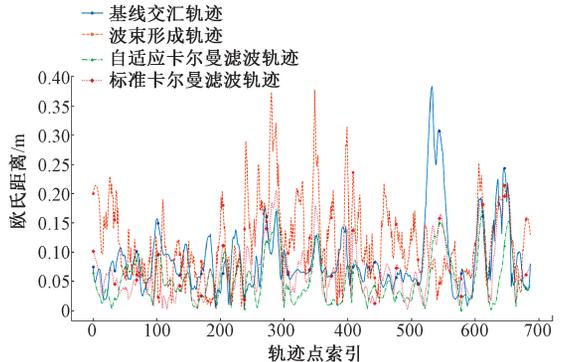
(a) 箱线图
(a) Boxplot



(b) 直方图
(b) Histogram



(c) 累积分布函数
(c) Cumulative distribution function



(d) 轨迹定位误差
(d) Trajectory error index

图 15 各方法误差分布对比

Fig. 15 Comparison of error distributions across methods

上述分布特性具体误差数值如表1所示,可见在综合误差指标上,RD-AKF的平均欧氏距离降至0.053 m,较其他单一算法其RMSE至少降低24.3%。较KF,其RMSE降低18.5%,体现了其更优的整体估计精度。尤为关键的是,RD-AKF的最大欧氏距离被控制在

0.173 m,较KF方法的0.205 m降低了约15.6%,证明了其有效抑制严重异常值的能力。其0.039 m的标准差在所有方法中最小,低于KF的0.046 m,这定量地印证了箱线图与分布函数的结论,即其估计结果具有极高的集中度和可重复性。

表1 数据误差统计总汇

Table 1 Statistical summary of data errors

(m)

| 方法 | 平均欧氏距离 | 欧氏距离 RMSE | X 轴 MAE | Y 轴 MAE | Z 轴 MAE | 最大欧氏距离 | 欧氏距离标准差 |
|--------|--------|-----------|---------|---------|---------|--------|---------|
| RD-AKF | 0.053 | 0.066 | 0.022 | 0.038 | 0.017 | 0.173 | 0.039 |
| KF | 0.067 | 0.081 | 0.030 | 0.044 | 0.023 | 0.205 | 0.046 |
| 基线交汇 | 0.089 | 0.107 | 0.041 | 0.057 | 0.033 | 0.383 | 0.060 |
| FD-PCB | 0.116 | 0.134 | 0.058 | 0.062 | 0.054 | 0.377 | 0.066 |

同时,为统计算法执行时间,对包含912帧的同一组音频数据连续处理10次以获取统计稳定的时间数据,如表2所示。

表2 算法执行时长统计

Table 2 Running-time statistics of the algorithm

| 算法 | 单帧平均耗时/ms |
|--------|-----------|
| FD-PCB | 28.5 |
| 基线交汇 | 3.8 |
| KF | 29.6 |
| RD-AKF | 29.8 |

数据说明基线交汇作为纯几何方法,计算效率最高;FD-PCB波束形成法因需遍历三维网格,计算负载最大。由于FD-PCB与基线交汇可并行执行,融合框架的总耗时由较慢的FD-PCB主导。相较于单一的FD-PCB算法,融合框架在仅增加约4%的时间开销下,将精度提升了超过50%;RD-AKF的计算成本与KF相当,约0.7%的相对时间增加换取了误差18.5%的误差降低。

2.4 鲁棒性分析

为验证算法在复杂噪声环境下的鲁棒性,本文设计对比实验旨在分析算法性能随SNR下降的变化情况。在实验室中通过人员活动和移动设备发出高斯白噪声,将环境SNR维持在15 dB,阵列布置与2.3节一致。避免实验结果偶然性,对目标无人机取10次10 s的随机复杂飞行音频,每次飞行同2.3节使用RD-AKF与标准卡尔曼滤波进行轨迹追踪分析,第1次轨迹如图16所示。

其中,实线轨迹线表述真实轨迹,虚线轨迹线表述RD-AKF轨迹,点划线轨迹线表述标准卡尔曼滤波轨迹。

统计10次误差数据,其中3组数据如表3所示。数据表明,随着SNR从25 dB降到15 dB,除X轴两种方法

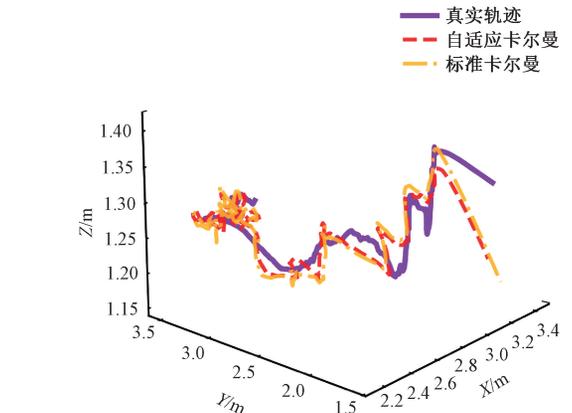


图16 两种定位结果与真实轨迹(SNR=15 dB)

Fig. 16 Comparison of two localization results and the ground truth trajectory(SNR=15 dB)

的误差均有明显增大,取10次数据平均值计算,RD-AKF的欧氏距离标准差从0.039 m增至0.046 m,增幅为9.1%,KF欧氏距离标准差从0.046 m增至0.053 m,增幅为15.2%,充分说明RD-AKF结果更加平滑、可靠,对环境噪声的变化不敏感;RD-AKF的最大误差稳定控制在0.3 m以内,而KF的最大误差多次超过0.4 m,表明RD-AKF更有效地遏制由于瞬时强噪声引起的滤波发散现象。

多次实验也证明X轴的误差变动不明显并非偶然,根源主要在于轨迹的几何特性与算法的工作机理共同作用的结果,实验的运动轨迹在X轴方向上具有更丰富的动态变化,观测信息量也更为充足,有利于依赖观测信息的卡尔曼滤波方法。

另外,所有方法在Y轴方向上的误差均显著高于X和Z轴,究其原因,传感器阵列的空间构型限制了方位辨识能力。两个竖直放置的平面阵列均与XOZ平面平行,导致其对Y轴方向的波达角度估计灵敏度较低,从而

表 3 数据误差统计 ($SNR=15$ dB)
Table 3 Statistic of data errors ($SNR=15$ dB)

(m)

| 组别 | 方法 | 平均欧氏距离 | 欧氏距离 $RMSE$ | X 轴 MAE | Y 轴 MAE | Z 轴 MAE | 最大欧氏距离 | 欧氏距离标准差 |
|----|--------|--------|-------------|-----------|-----------|-----------|--------|---------|
| 1 | RD-AKF | 0.075 | 0.088 | 0.026 | 0.055 | 0.028 | 0.292 | 0.046 |
| | KF | 0.087 | 0.101 | 0.028 | 0.064 | 0.033 | 0.413 | 0.051 |
| 2 | RD-AKF | 0.076 | 0.089 | 0.026 | 0.055 | 0.029 | 0.269 | 0.046 |
| | KF | 0.089 | 0.103 | 0.028 | 0.066 | 0.034 | 0.410 | 0.052 |
| 3 | RD-AKF | 0.061 | 0.072 | 0.021 | 0.045 | 0.024 | 0.247 | 0.038 |
| | KF | 0.086 | 0.103 | 0.027 | 0.064 | 0.032 | 0.486 | 0.055 |

影响该方向上的定位精度,使其成为系统整体定位精度的主要瓶颈。

3 结 论

本文提出的基于分布式麦克风阵列的残差驱动融合自适应卡尔曼滤波方法,在定位精度、稳定性和计算效率之间实现了良好平衡。实验表明,相较于单一算法及标准卡尔曼滤波,RD-AKF 在均方根误差、最大定位误差及轨迹平滑性方面均具有明显优势,且在较低信噪比环境下仍保持稳定性能。然而,该方法仍存在一定局限性,系统对 Y 轴方向的定位精度相对较低,受限于阵列布设构型对特定方向的感知灵敏度。未来可进一步优化阵列布局与自适应策略,提升系统在复杂场景下的适用性与稳健性。

参考文献

[1] 武晓雯, 郭孟莹, 胡阿建, 等. 基于改进遗传粒子群算法的无人机路径规划[J]. 仪器仪表学报, 2025, 46(4): 315-325.
WU X W, GUO M Y, HU A J, et al. Path planning for UAV based on improved hybrid genetic particle swarm algorithm[J]. Chinese Journal of Scientific Instrument, 2025, 46(4): 315-325.

[2] 何敬玉, 刘吉铮, 杨志晨. 多旋翼无人机噪声源跟踪定位系统研发及测试[J]. 现代防御技术, 2024, 52(3): 1-8.
HE J Y, LIU J ZH, YANG ZH CH. Development and testing of multi rotor unmanned aerial vehicle noise source tracking and positioning system[J]. Modern Defense Technology, 2024, 52(3): 1-8.

[3] RAHAMAN A, KIM B. An mm-sized biomimetic directional microphone array for sound source localization in three dimensions[J]. Microsystems & Nanoeengineering, 2022, 8: 66.

[4] 李永福, 文跃洲, 黄龙旺. 基于改进扩展状态观测器的四旋翼无人机轨迹鲁棒跟踪控制[J]. 仪器仪表学报, 2023, 44(4): 129-139.
LI Y F, WEN Y ZH, HUANG L W. Robust trajectory tracking control of quadrotor UAV based on the improved extended state observer[J]. Chinese Journal of Scientific Instrument, 2023, 44(4): 129-139.

[5] GRONDIN F, MICHAUD F. Lightweight and optimized sound source localization and tracking methods for open and closed microphone array configurations[J]. Robotics and Autonomous Systems, 2019, 113: 63-80.

[6] ZHANG T W, ZHANG H Y, LI X F, et al. AcousticFusion: Fusing sound source localization to visual SLAM in dynamic environments[C]. 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2021: 6868-6875.

[7] VERREYCKEN E, SIMON R, QUIRK-ROYAL B, et al. Bio-acoustic tracking and localization using heterogeneous, scalable microphone arrays[J]. Communications Biology, 2021, 4: 1275.

[8] 刘望生, 刘艳梅. 多特征优化下室内声源鲁棒跟踪算法[J]. 仪器仪表学报, 2024, 45(8): 316-325.
LIU W SH, LIU Y M. Robust tracking algorithm for indoor sound source based on multi-feature optimization[J]. Chinese Journal of Scientific Instrument, 2024, 45(8): 316-325.

[9] CHEN Y, CAO Y D, WANG R. Acoustic source tracking based on probabilistic data association and distributed cubature Kalman filtering in acoustic sensor networks[J]. Sensors, 2022, 22(19): 7160.

[10] THAKUR S, SINGH S. Moving sound source localization in 3D space based on time delay and energy ratio technique[C]. INTER-NOISE and NOISE-CON

- Congress and Conference Proceedings, 2024: 7156-7165.
- [11] CAKMAK B, DIETZEN T, ALI R, et al. A distributed steered response power approach to source localization in wireless acoustic sensor networks[C]. 2022 International Workshop on Acoustic Signal Enhancement, 2022: 1-5.
- [12] 林嘉睿, 荆伟杰, 任永杰, 等. 基于合作靶标的分布式测量系统定向方法[J]. 仪器仪表学报, 2021, 42(11): 10-16.
- LIN J R, JING W J, REN Y J, et al. Orientation method for distributed measurement system based on cooperative target [J]. Chinese Journal of Scientific Instrument, 2021, 42(11): 10-16.
- [13] HAGHIGHATSHOAR S, MUIR D R. Low-power spiking neural network audio source localisation using a Hilbert transform audio event encoding scheme[J]. Communications Engineering, 2025, 4: 18.
- [14] WANG H, GU ZH Y, CHEN K, et al. Robust direction-of-arrival estimation for a target speaker based on multi-task U-net based direct-path dominance test [J]. JASA Express Letters, 2021, 1(2): 024801.
- [15] 骆卉子, 曲长文, 徐征. 多机只测角无源定位中的空间对准方法[J]. 控制与决策, 2017, 32(5): 892-898.
- LUO H Z, QU CH W, XU ZH. Spatial alignment method for airborne multi-plane bearings-only passive localization [J]. Control and Decision, 2017, 32(5): 892-898.
- [16] 万俊豪, 左建勇, 丁景贤, 等. 基于新息自适应卡尔曼滤波地铁测速定位方法[J]. 仪器仪表学报, 2025, 46(1): 236-246.
- WAN J H, ZUO J Y, DING J X, et al. A speed measurement and positioning method of metro based on innovation-based adaptive Kalman filter[J]. Chinese Journal of Scientific Instrument, 2025, 46(1): 236-246.
- [17] 杜坡, 段振云, 张静, 等. 基于双目视觉的大直径零

件转角高精度测量方法[J]. 仪器仪表学报, 2025, 46(4): 35-43.

DU P, DUAN ZH Y, ZHANG J, et al. A high-precision measurement method for rotation angles of large-diameter components based on binocular vision[J]. Chinese Journal of Scientific Instrument, 2025, 46(4): 35-43.

作者简介



刘昊, 2023 年于杭州电子科技大学获得学士学位, 现为杭州电子科技大学硕士研究生, 主要研究方向为声源定位。

E-mail: liuhao86960@163.com

Liu Hao received his B. Sc. degree from Hangzhou Dianzi University in 2023. He is currently a master's candidate at Hangzhou Dianzi University. His main research interest includes sound source localization.



李文国, 分别在 2005 年、2009 年和 2022 年于杭州电子科技大学获得学士学位、硕士学位和博士学位, 现任职于咸亨国际(杭州)电气制造有限公司, 为正高级工程师, 主要研究方向为生物电信号处理与多模态信息融合。

E-mail: wgli_hz@163.com

Li Wenguo received his B. Sc., M. Sc. and Ph. D. degrees all from Hangzhou Dianzi University in 2005, 2009 and 2022, respectively. He is currently a senior technical engineer in Xianheng International (Hangzhou) Electrical Manufacturing Co., Ltd. His main research interests include bioelectrical signal processing and multimodal information fusion.



席旭刚(通信作者), 分别在 1999 年、2002 年、2019 年于杭州电子科技大学获学士学位、硕士学位、博士学位, 现为杭州电子科技大学教授, 主要研究方向为信号处理, 人机交互和康复机器人。

E-mail: xixugang@hdu.edu.cn

Xi Xugang (Corresponding author) received his B. Sc., M. Sc., and Ph. D. degrees all from Hangzhou Dianzi University in 1999, 2002 and 2019, respectively. He is currently a professor at Hangzhou Dianzi University. His main research interests include signal processing, human-robot interaction, and rehabilitation robot.