

文章编号: 1673-3193(XXXX)XX-0001-10

# 低信噪比下改进聚焦频率选择的宽带声阵列 DOA 估计方法

申东<sup>1</sup>, 兰志强<sup>2\*</sup>, 朱坤<sup>2</sup>, 王艳杰<sup>2</sup>, 王洁<sup>2</sup>, 何剑<sup>1</sup>, 丑修建<sup>1</sup>

(1. 中北大学 仪器与电子学院, 山西 太原 030051; 2. 苏州大学 未来科学与工程学院, 江苏 苏州 215222)

**摘要:** 针对声阵列信源波达方向估计中相干子空间方法现有聚焦频率选择在低信噪比、多目标环境下面临的性能显著衰减问题, 设计了一种全频域特征值分布的聚焦频率选择方法, 优化了聚焦频率选择策略, 从而提升了估计效率和准确性。通过对各频点阵列接收信号协方差矩阵进行特征值分解, 定义信号特征值累加值和噪声特征值累加值的比值为信噪分离度指标, 选取较大的频点作为聚焦频点。针对多目标场景, 构建与信源数目匹配的多频率聚焦集合, 充分利用宽带信号内外的所有频率成分信息。通过仿真实验, 在低信噪比下本文方法的空间谱谱峰更尖锐、旁瓣更低, 近邻角度分辨率更优; 针对多信源目标, 其估计精度更具稳定性, 计算耗时也显著低于相最佳聚焦频率法。通过实验验证, 在办公室和商场复杂环境下, 本文方法与最佳聚焦频率、中心频率相比平均绝对误差分别降低 10.2%、36.4%, 均方根误差分别降低 41.0%、50.0%, 本文方法不仅平均误差小, 且结果更稳定、大误差情况更少。本文所提聚焦频率选择策略可充分利用全频域信息, 能有效平衡计算效率和计算精度, 适配低信噪比、多目标宽带声学场景。但计算复杂度会随频点数量线性上升, 动态实时性表现有待优化。

**关键词:** 宽带声学估计; 低信噪比; 聚焦频率; 特征值分布

中图分类号: TN911

文献标识码: A

doi: 10.62756/jnuc.issn.1673-3193.2025.09.0021

**引用格式:** 申东, 兰志强, 朱坤等. 低信噪比下改进聚焦频率选择的宽带声阵列 DOA 估计方法[J]. 中北大学学报(自然科学版), DOI:10.62756/jnuc.issn.1673-3193.2025.09.0021.

Shen Dong, Lan Zhiqiang, Zhu Kun, et al. DOA Estimation Method for Wideband Acoustic Array with Improved Focus Frequency Selection under Low Signal-to-Noise Ratio[J]. Journal of North University of China (Natural Science Edition), DOI:10.62756/jnuc.issn.1673-3193.2025.09.0021.

## DOA Estimation Method for Wideband Acoustic Array with Improved Focus Frequency Selection under Low Signal-to-Noise Ratio

Shen Dong<sup>1</sup>, Lan Zhiqiang<sup>2\*</sup>, Zhu Kun<sup>2</sup>, Wang Yanjie<sup>2</sup>, Wang Jie<sup>2</sup>, He Jian<sup>1</sup>, Chou Xiujuan<sup>1</sup>

(1. School of Instrument and Electronics, North University of China, Taiyuan 030051, China;

2. School of Future Science and Engineering, Soochow University, Suzhou 215222, China)

**Abstract:** To address the significant performance degradation of the existing focusing frequency selection in the coherent subspace method for direction-of-arrival (DOA) estimation of acoustic array sources under

收稿日期: 2025-09-29

基金项目: 极限环境光电动态测试技术及仪器全国重点实验室基金资助; 国家自然科学基金资助项目(U2341210, 62171414, 52205608); 山西省基础研究计划项目(202303021223008, 20210302123059); 青年科学基金项目(62301349)

作者简介: 申东(1999—), 男, 硕士生, 主要从事声阵列定位的研究。

\* 通信作者: 兰志强(1993—), 男, 副教授, 博士, 主要从事微弱声信号处理及其系统集成研究。E-mail: zqlan@suda.edu.cn。

low signal-to-noise ratio (SNR) and multi-target environments, a focusing frequency selection method based on full-frequency-domain eigenvalue distribution is designed to optimize the focusing frequency selection strategy and improve estimation efficiency and accuracy. Eigenvalue decomposition was performed on the covariance matrix of array received signals at each frequency point, and the ratio of the sum of signal eigenvalues to the sum of noise eigenvalues was defined as the signal-noise separation index; frequency points with higher indices were selected as focusing frequencies. For multi-target scenarios, a multi-frequency focusing set matching the number of sources was constructed to fully utilize all frequency component information of broadband signals. In simulation experiments, the proposed method exhibits sharper spatial spectrum peaks, lower sidelobes, and better adjacent angle resolution under low SNR; for multi-source targets, its estimation accuracy is more stable, and the computation time is significantly shorter than that of the optimal focusing frequency method. In experimental validations in complex office and shopping mall environments, compared with the optimal focusing frequency method and the central frequency method, the proposed method reduces the mean absolute error by 10.2% and 36.4%, and the root mean square error by 41.0% and 50.0%, respectively. It not only achieves smaller average errors but also yields more stable results with fewer large-error cases. The proposed focusing frequency selection strategy can fully utilize full-frequency-domain information, effectively balance computation efficiency and accuracy, and is suitable for low-SNR and multi-target broadband acoustic scenarios. However, its computational complexity increases linearly with the number of frequency points, and its dynamic real-time performance needs further improvement.

**Key words:** broadband acoustic estimation; low signal-to-noise ratios; focus frequency; eigenvalue distribution

## 0 引言

波达方向估计(Direction of Arrival, DOA)是阵列信号处理领域的核心研究内容之一,被广泛应用于声呐探测、语音定位、环境声监测及智能声学系统等多个声学领域<sup>[1-2]</sup>。然而,传统窄带DOA方法的前提是远场窄带信号模型,即信号带宽远小于中心频率,入射波近似为平面波,阵列接收信号差异仅体现为相位差。而声学信号多为宽带信号,其频散效应会使不同频率分量的导向矢量产生显著差异,导致窄带算法的相位补偿失效;且导致各频率分量的信号能量分布分散,不同频率分量的导向矢量不正交,直接导致子空间分离模糊,最终出现估计角度偏移、虚假峰或漏峰问题。因而宽带阵列信号处理技术在声学领域中显得尤为重要,依据所处理信号的相关性,可分为非相干子空间法(Incoherent Subspace Method, ISM)、相干子空间法(Coherent Subspace Method, CSM)两类<sup>[3-4]</sup>。ISM仅能处理非相干信号,具备原理简单且在高信噪比条件下有效的特点,但在低信噪比环境中估计精度显著下降。相比之下,CSM能处理相干信号和非相干信号,

通过聚焦变换操作对信号进行DOA估计,但聚焦矩阵形式和聚焦频率选择会影响到估计精度<sup>[5-7]</sup>。

为了提升CSM方法估计精度,现有多种针对聚焦矩阵构造的改进方案。例如文献<sup>[8]</sup>提出了基于聚焦信号子空间的估计方法,文献<sup>[9]</sup>利用阵列流形矩阵的分解表示,都避免了聚焦过程对初始DOA的依赖。然而,宽带信号处理需通过聚焦矩阵变换到聚焦频点上,现聚焦频点主要选择方法可分为两类:一是以中心频率为代表的固定频点法,二是以聚焦拟合误差最小化为目标的最佳聚焦频率法。例如文献<sup>[10]</sup>采用聚焦拟合误差最小化确定最佳聚焦频率。然而,最佳聚焦频率并非唯一,在实际应用中,由于信号和噪声的能量在带宽内分布的不均匀性,宽带内最佳的聚焦频率也会随环境变化,且最佳聚焦频率计算量较大<sup>[11]</sup>。中心频率<sup>[12]</sup>、中心频率附近最大能量频点<sup>[13]</sup>、子频带中心频率<sup>[14]</sup>等被研究并被选为聚焦频率,频率选择过程变得简单且不易受环境影响。然而,中心频率的固定频点法未能充分考虑信号的全部频带信息,如中心频点可能不是信号能量峰值位置,信号有效信息也会分布在高频或低频区域。这些问题最终导致将信号信息聚焦变换到聚焦频点时,信源信息利用不足,聚焦损失累积,影响

DOA精度。

综合上述分析,为了解决相干子空间方法中现有聚焦频率选择存在的问题,本文提出一种低信噪比下改进聚焦频率选择策略的宽带声阵列DOA估计方法。在聚焦频点选择逻辑方面,现有方法通过被动选取固定频点或遍历搜索最佳参考频点,而本文方法通过分析整个频域上阵列接收信号协方差矩阵的特征值分布,主动选取信号与噪声分离度较高的频点作为参考频点,且有效克服低信噪比影响。在信息利用方式方面,固定频点法仅利用单一频点信息,而本文方法通过分离度指标量化频点贡献,构建高质量多频点聚焦集合,减少聚焦损失,充分利用宽带内外所有频率成分信息,既避免单一频点的信息局限,又排除低质量频点的噪声干扰。在性能平衡方面,最佳聚焦频率法牺牲计算效率换取精度,固定频点法牺牲精度换取效率,本文通过聚焦频点集合的设计,实现精度与效率的主动平衡。

## 1 数学模型

### 1.1 阵列模型

本文采用的信源接收装置如图1所示,麦克风个数为 $M$ ,阵元间距为 $d$ 。

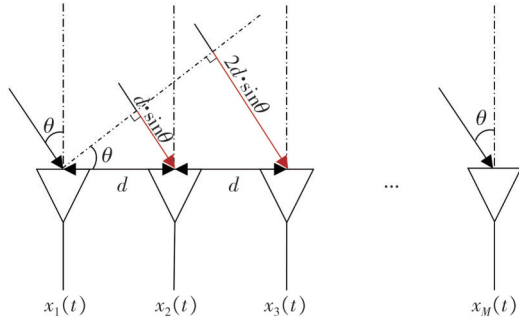


图1 麦克风阵列接收信号模型

Fig. 1 Microphone array signal receiving model

存在远场宽带信号 $\{s_k(t)\}_{k=1}^K = [s_1(t), \dots, s_K(t)]^T$ ,  $K$ 为信源个数,分别从角度 $\{\theta_k\}_{k=1}^K = [\theta_1, \dots, \theta_K]^T$ 入射到接收阵置,且假设信源与噪声彼此独立。选择某一阵元为参考阵元,令 $t$ 时刻下第 $m$ 个阵元接收到的信号表示为

$$x_m(t) = \sum_{k=1}^K s_k(t - \tau_{mk}) + n_m(t) \quad (1)$$

$m = 1, 2, \dots, M,$

式中: $x_m(t)$ 表示第 $m$ 个阵元在 $t$ 时刻接收到的信

号数据; $s_k(t)$ 表示第 $k$ 个信号源在时刻 $t$ 的信号波形; $n_m(t)$ 表示第 $m$ 个阵元所受的加性噪声; $\tau_{mk}$ 表示第 $m$ 个阵元接收第 $k$ 个信号时相较于参考阵元的时间延迟。

$$\tau_{mk} = \frac{(m-1)d \sin \theta_k}{c}, \quad m = 1, 2, \dots, M, \quad k = 1, 2, \dots, K, \quad (2)$$

式中: $c$ 为声音传播速度。

### 1.2 宽带阵列信号处理原理及问题分析

将宽带信号 $\{s_{k,j}(t)\}_{k=1}^K$ 、噪声分量 $n_m(t)$ 在时域进行分段,分别划分为 $\sum_{j=1}^J \{s_{k,j}(t)\}_{k=1}^K$ 、 $\sum_{j=1}^J n_{m,j}(t)$ ,  $J$ 个子段不重合且均满足窄带信号定义。式(1)可表示为

$$x_{m,j}(t) = \sum_{k=1}^K s_{k,j}(t - \tau_{mk}) + n_{m,j}(t), \quad j = 1, 2, \dots, J. \quad (3)$$

对式(3)中的每个子带进行傅里叶变换,可得

$$X_{m,j}(f) = \sum_{k=1}^K S_{k,j}(f) e^{-j2\pi f \tau_{mk}} + N_{m,j}(f), \quad j = 1, 2, \dots, J. \quad (4)$$

将式(4)展开成矩阵形式为

$$X_j(f) = A(f, \theta) \cdot S_j(f) + N_j(f), \quad j = 1, 2, \dots, J, \quad (5)$$

式中: $X_j(f)$ 、 $S_j(f)$ 、 $N_j(f)$ 分别为第 $j$ 个子带在频率 $f$ 处的阵列接收矢量、信源数据矢量与噪声数据矢量; $A(f, \theta)$ 为频率 $f$ 处的阵列流行矩阵。

$$A(f, \theta) = [a_1(f, \theta), \dots, a_K(f, \theta)], \quad (6)$$

式中: $a_k(f, \theta)$ 为第 $k$ 个信源对应频率 $f$ 下的导向矢量。

$$a_k(f, \theta) = [e^{-j2\pi f \tau_{1k}}, \dots, e^{-j2\pi f \tau_{MK}}]^T. \quad (7)$$

将式(5)关于频率 $f$ 的所有 $J$ 个子段数据按列排放,得到宽带信号阵列接收模型中关于频率 $f$ 的多频域快拍下的模型:

$$X(f) = A(f, \theta) \cdot S(f) + N(f). \quad (8)$$

在频率 $f$ 处,对应的阵列接收数据协方差矩阵定义为

$$R_x(f) = \frac{1}{J} X(f) X^H(f), \quad (9)$$

式中: $(\cdot)^H$ 为共轭转置。相干子空间法的核心思想是通过构造聚焦矩阵,将不同频率点上的数据聚焦映射到聚焦频率点 $f_0$ 上<sup>[15]</sup>。如采用旋转信号

子空间(Rotational Signal-Subspace, RSS)<sup>[16]</sup>方法进行聚焦, 聚焦矩阵  $T$  满足式(10)。

$$T = \arg \left\{ \min_T \|T(f)A(f, \theta) - A(f_0, \theta)\|_F^2 \text{ subject to } T^H T^H T = I_M \right\}, \quad (10)$$

式中:  $A(f, \theta)$  和  $A(f_0, \theta)$  分别为频点  $f$  和聚焦频率点  $f_0$  处的阵列流行矩阵。式(10)的聚焦矩阵显示解为

$$T(f) = V(f)U^H(f), \quad (11)$$

式中: 矩阵  $U^H(f)$ 、矩阵  $V(f)$  分别为矩阵  $A(f, \theta)A^H(f_0, \theta)$  经奇异值分解后的左奇异矩阵、右奇异矩阵。

在宽带阵列信号处理方法中都需要选择聚焦频率, 如 RSS 方法需要将阵列流行矩阵  $A(f, \theta)$  变换到聚焦频点  $f_0$  处的  $A(f_0, \theta)$ , 如式(12)所示。

$$T(f) \cdot A(f, \theta) = A(f_0, \theta). \quad (12)$$

聚焦频点  $f_0$  的选取至关重要, 若依赖信号中心频率或其邻近频率范围作为聚焦频率, 因宽带信号能量也会分布在非中心频段的高频或低频区域, 中心频率处的信号能量只能代表部分信息, 可能导致在有效频带利用不足。若依赖最佳聚焦频率作为聚焦频率, 最佳聚焦频率通过遍历搜索使得拟合误差最小, 计算量大且不唯一。这些问题都将导致聚焦误差变大, 降低 DOA 估计的精度。

## 2 特征值分布的多频率聚焦

针对上述问题, 本文提出基于特征值分布的多频率选择策略。通过分析各频点对应协方差矩阵的特征值分布, 量化信号子空间与噪声子空间的分离度, 有效识别出信号能量集中且噪声干扰小的频段, 选其频点作为参考频点集合, 具体分6个步骤。

步骤 1: 在每个频率点  $f$ , 对式(9)进行特征值分解, 具体为

$$R_X(f) = U(f)\Lambda(f)U^H(f), \quad (13)$$

式中: 矩阵  $\Lambda(f)$  是一个对角矩阵, 其特征值满足降序排列。

$$\Lambda(f) = \text{diag}(\lambda_1(f), \dots, \lambda_M(f)),$$

$$\lambda_1(f) > \lambda_2(f) > \dots > \lambda_M(f) = \begin{bmatrix} \Lambda_S(f) & 0 \\ 0 & \Lambda_N(f) \end{bmatrix}. \quad (14)$$

式(13)中特征向量矩阵  $U(f)$  可以分解为信号子空间  $U_S(f)$  和噪声子空间  $U_N(f)$ , 具体表示为

$$U(f) = [U_S(f) \ U_N(f)], \quad (15)$$

$$U_S(f) = [u_1(f), \dots, u_K(f)] \in \mathbb{C}^{M \times K}, \quad (16)$$

$$U_N(f) = [u_{K+1}(f), \dots, u_M(f)] \in \mathbb{C}^{M \times (M-K)}, \quad (17)$$

式中:  $K$  为信源个数,  $K$  通过信息论中最小描述长度准则(Minimum Description Length, MDL)验证确定。

$$MDL(P) = L(M-P) \ln \Lambda(P) +$$

$$\frac{1}{2} P(2M-P) \ln(L), \quad (18)$$

$$\Lambda(P) = \frac{1}{M-P} \frac{\sum_{i=P+1}^M \lambda_i}{\left( \prod_{i=P+1}^M \lambda_i \right)^{\frac{1}{M-P}}}, \quad (19)$$

$$K = \text{argmin}[MDL(P)], \quad (20)$$

式中:  $P$  为待估计信源数目;  $\Lambda(P)$  为似然函数;  $L$  为快拍数;  $M$  为阵元数目。

步骤 2: 定义分离度指标为

$$\text{Separation}_{\text{sum}}(f) = \frac{\sum_{i=1}^K \lambda_i(f)}{\sum_{i=K+1}^M \lambda_i(f)}, \quad (21)$$

式中:  $\lambda_i(f)$  表示在频率点  $f$  处协方差矩阵的第  $i$  个特征值。式(21)用于量化信号特征值与噪声特征值之间的分离度。分离度指标数值越大, 说明信号子空间  $U_S(f)$  与噪声子空间  $U_N(f)$  之间的区分度越明显, 信号能量越突出, 该频点对 DOA 估计的贡献越大。反之, 若分离度指标数值越小, 则说明该频点信号与噪声相互混叠严重, 该频点不利于 DOA 估计。

步骤 3: 根据式(21)定义的信号特征值与噪声特征值之间的分离度, 在候选频点  $f$  中搜索计算分离度, 保证频率点  $f_i$  的分离度不小于候选频点中任何其他频点的分离度, 具体表达式为

$$\text{Separation}(f_r) \geq \text{Separation}(f). \quad (22)$$

步骤 4: 通过式(22)选取分离度指标值最高的  $N$  个频点构建频率聚焦集合

$$f_r = \{f_{r_1}, \dots, f_{r_N}\} \quad i = 1, 2, \dots, N, \quad (23)$$

式中:  $N=K$ (真实信源数目)。选取分离度最高的  $N$  个频点, 可确保每个信源的关键能量成分均被



纳入聚焦集合,避免因 $N < K$ 导致信源信息丢失,或因 $N > K$ 引入低分离度、高噪声占比的频点,减少聚焦误差累积,这与相干子空间方法中信号子空间维度等于信源数的核心理论保持自洽。当 $N=K$ 时,复杂度仅与信源数正相关,可避免 $N > K$ 时的冗余运算。从而在保证DOA估计性能的前提下,有效控制计算量,实现计算效率与估计性能之间的良好平衡。

步骤 5: 通过式(11)求得聚焦矩阵 $T(f_r)=[T(f_{r_1}), \dots, T(f_{r_N})]$ ,对原频点 $f$ 处的阵列输出矩阵 $X(f)$ 进行聚焦变换,得聚焦后频点 $f$ 处的阵列输出矩阵 $Y(f)$ ,具体为:

$$Y(f) = T(f_r) \cdot X(f) \quad i=1, 2, \dots, N_o. \quad (24)$$

步骤 6: 经聚焦变换后,对各频点处的阵列接收数据协方差矩阵进行累加平均,具体为

$$\begin{aligned} R_Y &= \frac{1}{J} \sum_{j=1}^J Y(f) Y^H(f) = \\ &= \frac{1}{J} \sum_{j=1}^J T(f_r) \cdot X(f) X^H(f) \cdot T^H(f_r) = \\ &= \frac{1}{N} \sum_{i=1}^N \left( \frac{1}{J} \sum_{j=1}^J T(f_{r_i}) R(f) T^H(f_{r_i}) \right), \\ & \quad i=1, 2, \dots, N_o. \end{aligned} \quad (25)$$

综上所述,本文算法流程如下:

- 1) 通过式(13)计算特征值对角矩阵;
- 2) 通过式(21)、式(22),计算特征值分离度;
- 3) 通过式(23)选取聚焦频点 $f_r$ ;
- 4) 通过式(6)计算 $f_r$ 处的阵列流行矩阵 $A(f_r, \theta)$ ;
- 5) 通过式(6)计算全频段 $f$ 的阵列流行矩阵 $A(f, \theta)$ ;
- 6) 通过式(11)计算聚焦矩阵 $T(f_r)$ ;
- 7) 通过式(24)将信号数据聚焦到参考频点;
- 8) 通过式(25)计算聚焦后的协方差矩阵;
- 9) 最后采用传统窄带DOA方法处理协方差矩阵。

### 3 仿真测试及分析

#### 3.1 实验设置

本实验采用十六元均匀线性麦克风阵列( $M=16$ ),阵元间距 $d$ 为信号最高频率对应波长的二分

之一。信号源为 $K(K < M)$ 个远场宽带线性调频信号,参数设定如下:中心频率 $f_m=500$  Hz,带宽 $B=900$  Hz,满足宽带信号定义条件<sup>[17]</sup>。噪声环境设置为两种典型场景:零均值平稳高斯白噪声与 $1/f$ 特性粉红噪声,噪声带宽与信号带宽一致。

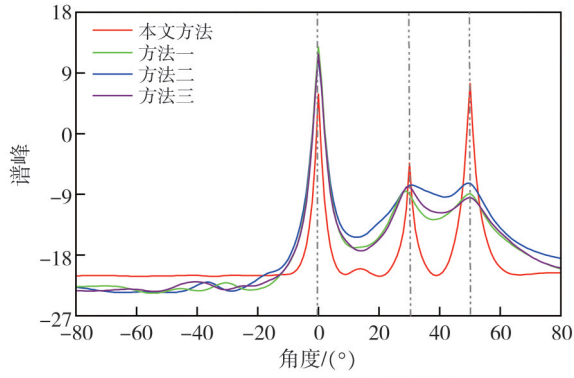
文献[14]选取子频带中心频率作为参考频率,文献[10]采用最佳聚焦频率作为参考频率。后文将最佳聚焦频率、中心频率作为参考频率的方法分别称作为方法一、方法二。仿真将选择本文方法、方法一和方法二分别作为聚焦频率选择方法,接着按照表1算法流程进行计算,最后选择多重信号分类(Multiple Signal Classification, MUSIC)进行DOA估计。另外将与CSM方法进行对比,后文称作方法三。

本文使用CPU为Intel(R) Core(TM) i5-10210U CPU @ 1.60GHz 2.11 GHz、RAM为8.00GB和OS为Windows 10 LTSC的硬件平台,软件环境为Python 3.7.6。在后续的验证阶段,将验证本文方法与方法一、方法二和方法三在DOA估计谱峰图的优劣,并对比四种方法估计角度与实际角度的均方根误差(Root Mean Square Error, RMSE),最后分析不同聚焦频率选择的计算耗时。

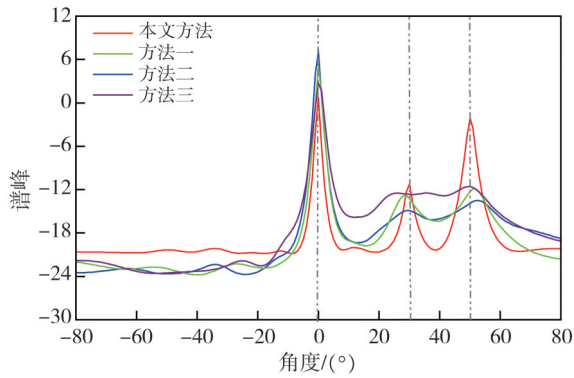
#### 3.2 不同信噪比下的空间谱图对比

假设有 $K=3$ 个人射信号,到达角度分别为 $\theta=[0^\circ, 30^\circ, 50^\circ]$ ,分析每个方法的空间谱峰图。如图2(a)、图2(b)分别为SNR等于-5 dB、-10 dB时下白噪声谱峰图;如图2(c)、图2(d)分别为SNR等于-5 dB、-10 dB时下粉红噪声谱峰图。

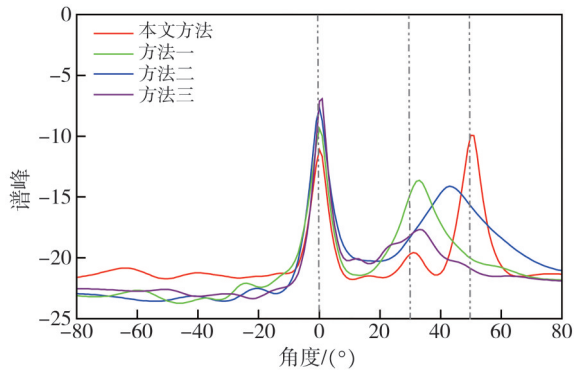
如图2(a)、图2(b),在白噪声中,本文方法显示出更尖锐、更窄峻的谱峰,保持较高的信号强度,减少了能量在主峰外的扩散;而方法一、方法二和方法三主峰更宽且下降速度较慢,且在 $\theta=50^\circ$ 时旁瓣扩散严重、峰值放缓。如图2(c)、图2(d),在粉红噪声环境中,SNR从-5 dB到-10 dB时,本文方法能保持稳定的峰值位置和较高的谱峰清晰度,在其他角度的谱值较低,能有效地抑制噪声和不相关的信号成分;而方法一、方法二和方法三的谱峰不完整且会出现虚假峰,谱峰角度偏移也相对较大。



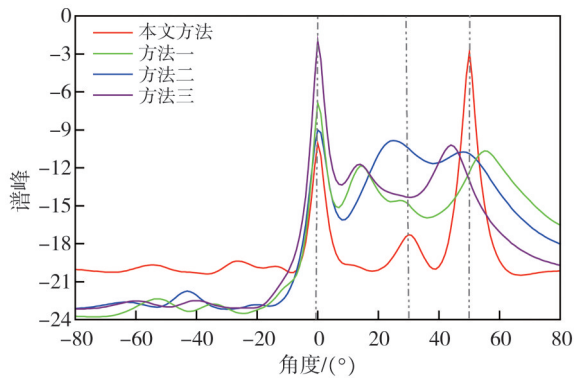
(a) SNR=-5 dB时白噪声谱峰图



(b) SNR=-10 dB时白噪声谱峰图



(c) SNR=-5 dB时粉红噪声谱峰图



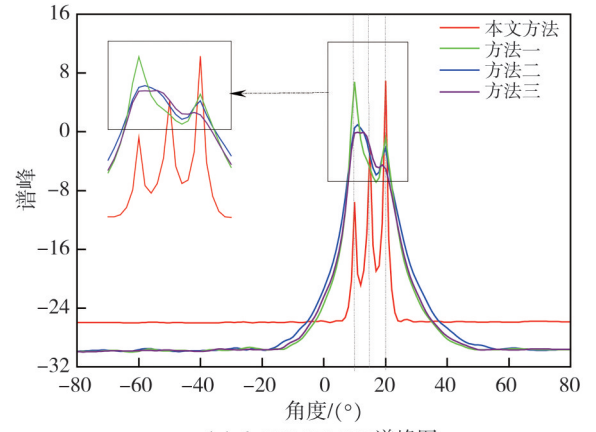
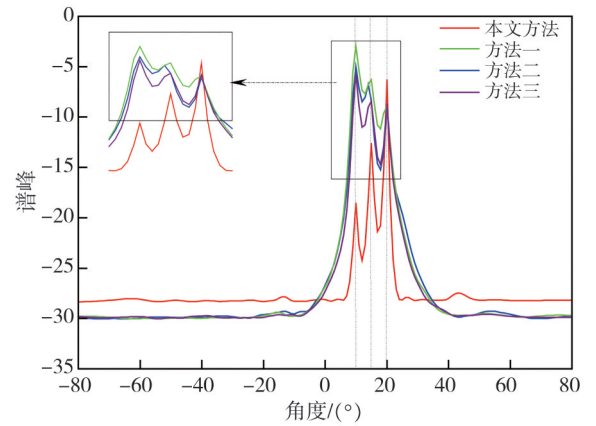
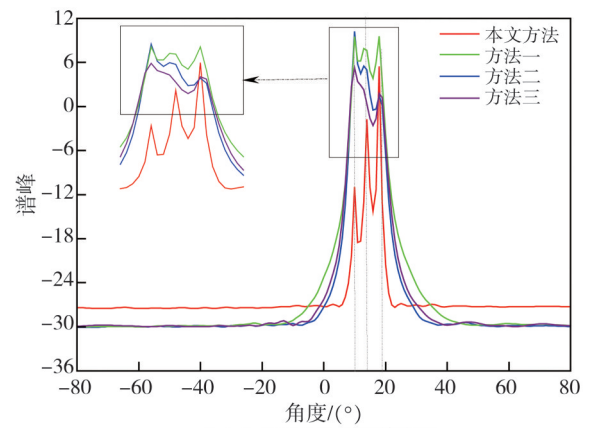
(d) SNR=-10 dB时粉红噪声谱峰图

图2 不同信噪比下的空间谱图对比

Fig. 2 Spatial spectrum comparison at different SNR

### 3.3 近邻角度下空间谱分辨率对比

对比在信号角度接近时不同方法的空间分辨率,如图3(a)、图3(b)、图3(c)、图3(d)分别为白噪声环境下SNR等于-5, 0, 5和10 dB时的谱峰图,图3(a)、图3(b)为 $\theta=[10^\circ, 15^\circ, 20^\circ]$ ,图3(c)为 $\theta=[10^\circ, 14^\circ, 18^\circ]$ 、图3(d)为 $\theta=[10^\circ, 13^\circ, 15^\circ]$ 。

(a)  $\theta=[10^\circ, 15^\circ, 20^\circ]$ 谱峰图(b)  $\theta=[10^\circ, 15^\circ, 20^\circ]$ 谱峰图(c)  $\theta=[10^\circ, 14^\circ, 18^\circ]$ 谱峰图

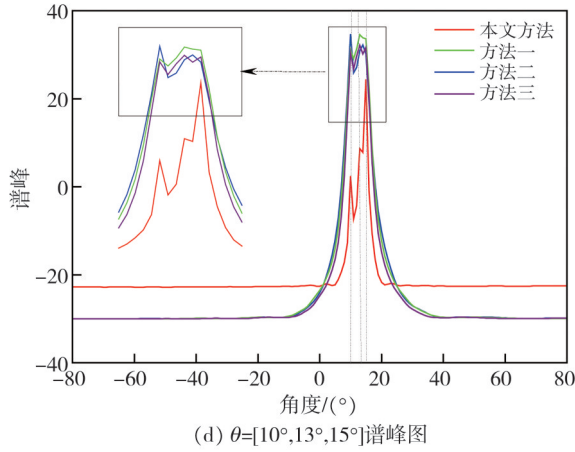


图3 近邻角度下空间谱图对比

Fig. 3 Spatial spectrum comparison under the nearest neighbour angle

如图3(a),在角度差距为 $5^\circ$ 时,本文方法能清晰分辨出来波方向,且谱峰尖锐清晰,但方法一、方法二和方法三出现谱峰融合,出现漏峰问题。但随着SNR从 $-5$  dB提高到 $0$  dB,方法一、方法二和方法三的角度分辨率提高(如图3(b)),也能分辨出三个来波方向,但谱峰尖锐程度不如本文方法。从图3(c)、图3(d)的放大区能看出,本文方法谱线主瓣更尖锐,相比之下方法一、方法二和方法三的谱峰存在一定展宽;在图3(d)更小角距场景中,方法一、方法二和方法三出现谱峰融合现象,且谱峰高度差异不明显,但本文方法仍具有清晰尖锐的谱峰。

### 3.4 不同信噪比下的均方根误差

根据空间谱图对比、空间谱分辨率对比,可以看出本文方法在谱峰清晰度、分辨率优于方法一、方法二和方法三。现再比较进行600次蒙特卡洛实验时,本文方法、方法一、方法二、方法三在不同信噪比和不同快拍数下的RMSE。

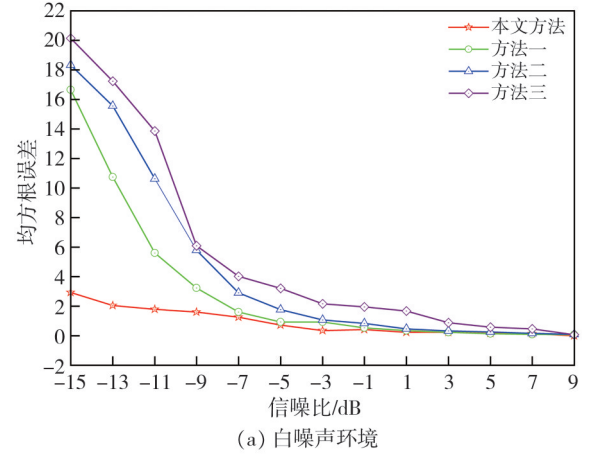
假设有 $K=3$ 个人射信号,到达角度分别为 $\theta=[0^\circ, 20^\circ, 35^\circ]$ , SNR动态范围 $[-15$  dB,  $10$  dB],步进 $2$  dB,在每个SNR点进行600次蒙特卡洛模拟实验,计算每种方法的RMSE,定义为

$$RMSE = \sqrt{\frac{1}{N} \sum_{n=1}^N |\hat{\theta}_k(n) - \theta_k|^2}, \quad (26)$$

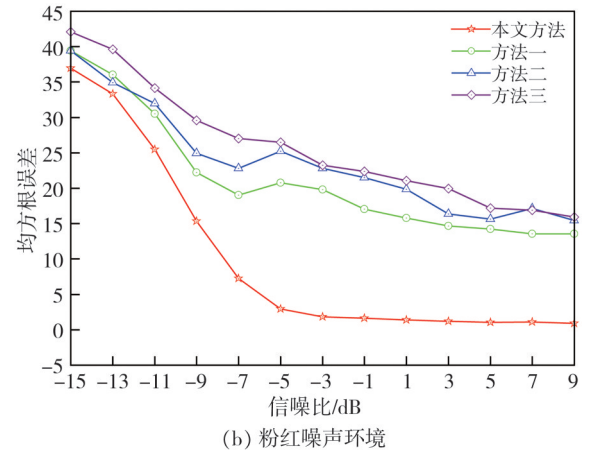
式中: $N$ 为蒙特卡洛实验次数; $\hat{\theta}_k(n)$ 为在第 $n$ 次实验对第 $k$ 个信号的估计值; $\theta_k$ 为真实角度。

如图4(a)、图4(b),分别为白噪声、粉红噪声环境下在各SNR下不同方法对应RMSE。在白噪声下整个信噪比范围内,尤其是在低SNR

( $-15$  dB到 $-5$  dB)时,本文方法始终保持最小的RMSE;且在整个信噪比范围内,本文方法的RMSE始终处于较低水平。在色噪声下, $-15$  dB至 $-10$  dB的SNR时,本文方法的RMSE低于方法一、方法二和方法三,且RMSE曲线呈快速下降趋势,且在 $-3$  dB后趋于平稳。



(a) 白噪声环境



(b) 粉红噪声环境

图4 不同信噪比下的均方根误差

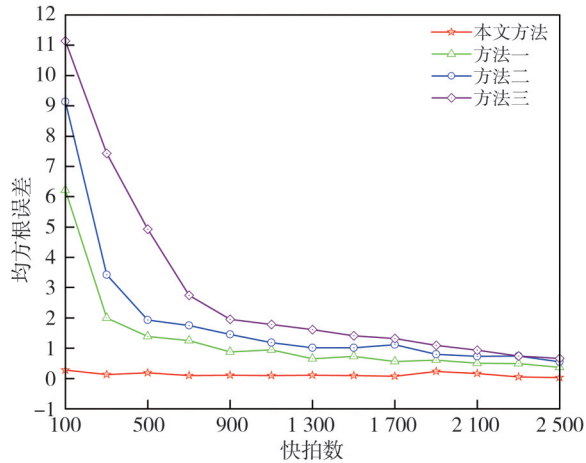
Fig. 4 RMSE at different signal-to-Noise Ratios

### 3.5 不同快拍数下的均方根误差

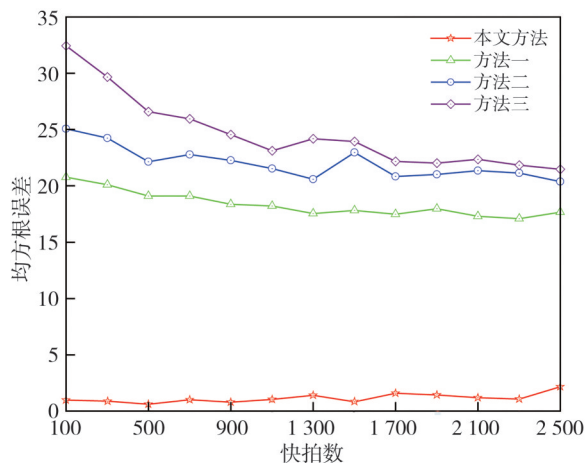
假设有 $K=3$ 个人射信号,到达角度分别为 $\theta=[-10^\circ, 25^\circ, 40^\circ]$ ,在每个快拍数目下进行600次蒙特卡洛模拟实验,计算在不同快拍数目下的RMSE。

图5(a)、图5(b)分别为SNR等于 $0$  dB时在白噪声、粉红噪声环境下,不同快拍数情况的RMSE。如图5(a)白噪声下,随着快拍数增加,方法一、方法二和方法三的RMSE呈下降趋势,最终RMSE接近于本文方法,但本文方法曲线相较于方法一、方法二和方法三的误差曲线波动更小,表现更为平滑稳定。如图5(b)粉红噪声下,

随着快拍数增加三种方法的RMSE的值较为稳定,但本文方法的RMSE值在快拍数范围内显著低于方法一、方法二和方法三。且在低快拍(如100)时本文方法的性能也有良好性能。



(a) 白噪声环境下不同快拍数



(b) 粉红噪声环境下不同快拍数

图5 不同快拍数下的均方根误差

Fig. 5 RMSE at different Snapshots

### 3.6 不同信源个数下的均方误差

分别假设有 $K=[2, 3, 4, 5]$ 个信源,其入射角度分别为 $[10^\circ, 80^\circ]$ 、 $[10^\circ, 45^\circ, 80^\circ]$ 、 $[10^\circ, 33^\circ, 56^\circ, 80^\circ]$ 、 $[10^\circ, 27^\circ, 44^\circ, 61^\circ, 80^\circ]$ 时,在不同信源数目下进行600次蒙特卡洛模拟实验,计算不同信源数目环境下的RMSE。

表1是白噪声环境下在 $\text{SNR}=0\text{ dB}$ 时对应的RMSE,由表格数据可知,在信源数目为2,3,4,5的各类场景中,本文方法计算得到的均方误差均小于方法一、方法二和方法三,在信源数较少时,四种方法的均方误差相差不大;且随着信源数目增多,本文方法的误差明显小于方法一、方法二

和方法三,精度优势更为突出。

表1 不同信源个数下的RMSE

Tab. 1 RMSE for different numbers of signal sources ( $^\circ$ )

| 方法   | 信源数目    |         |          |          |
|------|---------|---------|----------|----------|
|      | 2       | 3       | 4        | 5        |
| 本文方法 | 0.225 4 | 0.327 1 | 1.361 1  | 5.316 2  |
| 方法一  | 0.238 9 | 0.416 9 | 8.851 8  | 14.607 3 |
| 方法二  | 0.338 5 | 0.744 9 | 13.711 3 | 23.405 6 |
| 方法三  | 0.478 2 | 1.065 3 | 17.857 9 | 32.478 5 |

### 3.7 计算耗时

在每个快拍数进行100次蒙特卡洛实验,计算各种聚焦频率选择方法的耗时,如表2所示。在不同快拍数条件下,本文方法的耗时明显低于方法一,且整体耗时增长较为平缓,因方法一求最佳聚焦频率要计算阵列流行矩阵的乘积,当快拍数增加时计算耗时呈几何增加。虽然本文方法耗时略高于方法二,由3.2节到3.6节的仿真测验,方法二精度不如本文方法。

表2 计算耗时

Tab. 2 Calculation time s

| 方法   | 快拍数     |         |          |           |
|------|---------|---------|----------|-----------|
|      | 100     | 1 000   | 2 000    | 5 000     |
| 本文方法 | 0.119 3 | 1.110 3 | 2.185 7  | 5.269 2   |
| 方法一  | 0.143 3 | 9.225 3 | 36.493 8 | 212.967 0 |
| 方法二  | 0.065 9 | 0.579 0 | 1.172 6  | 2.798 8   |

### 3.8 实验对比总结

由3.2节到3.7节的对比实验可知,在计算精度准确度方面,因本文方法有效识别出信号能量集中且噪声干扰小的频段,整体效果是本文方法优于方法一、方法二和方法三。因方法二直接选取中心频率,而方法一遍历搜索使得聚焦拟合误差最小,能有效降低聚焦误差,方法一效果优于方法二。然而在实际工程运用中,也需考虑计算速度与精度的平衡。同时在3.7节的计算耗时对比上,耗时由少到多分别是:方法二、本文方法、方法一,因方法一需要计算阵列流行矩阵乘积耗时最长;而方法二直接选取信号中频耗时最短;而本文方法需在全频段计算分离度,耗时长于方法二。所以本文方法兼顾了较低的计算时间和较高的信息特征表达能力,提升了算法的准确性和稳定性。

## 4 实验验证

为验证所提算法在实际复杂声学场景下的普



适性与鲁棒性,本文采用真实麦克风阵列及采集的数据开展实验验证。实验所用麦克风阵列模型与声源数据均来自公开声学数据集 RealMAN<sup>[18]</sup>,该数据集以真实复杂环境为采集场景,包含多类型阵列拓扑与标注完备的声源数据,具体实验设计如下:

阵列模型选取 RealMAN 中的线性阵列拓扑结构,线性阵列的麦克风个数  $M=10$ ,其麦克风标号分别为 25, 21, 13, 5, 0, 1, 9, 17, 26 和 27,麦克风间距  $d=3\text{cm}$ 。其麦克风频响范围 20 Hz~20 kHz,采样率 48 kHz,满足语音信号与环境噪声的宽频采集需求。

声音数据选取来自商场和办公室两类典型复杂环境的混合噪声语音数据,声源位置真值由配套标注文件提供。一共选取 236 条数据,总时长为 1 121.95 s。

如表 3 所示,计算每种方法的平均绝对误差(Mean Absolute Error, MAE)和 RMSE, MAE 定义为

$$MAE = \frac{1}{N} \sum_{n=1}^N |\hat{\theta}_k(n) - \theta_k|, \quad (27)$$

式中:  $N$  为蒙特卡洛实验次数,这里的  $N=236$ ;  $\hat{\theta}_k(n)$  为在第  $n$  次实验对第  $k$  个信号的估计值;  $\theta_k$  为真实角度。

由表 3 可见,本文方法的 MAE 是 3 种算法中最低的,相比方法一降低了约 10.2%,相比方法二降低了约 36.4%,在复杂环境下的角度估计准确性更优。利用 RMSE 评估算法的整体误差分布,本文方法的 RMSE 远低于其他两种算法,相比方法一降低了约 41.0%,相比方法二降低了约 50.0%,说明本文方法不仅平均误差小,且在复杂环境下的定位结果更稳定、大误差情况更少。

表 3 不同方法的 MAE 和 RMSE

Tab. 3 MAE and RMSE for different methods (°)

| 算法名称 | MAE   | RMSE  |
|------|-------|-------|
| 本文方法 | 17.53 | 26.00 |
| 方法一  | 19.53 | 44.05 |
| 方法二  | 27.58 | 52.03 |

## 5 结 论

本文针对声阵列宽带 DOA 估计中相干子空间方法现有聚焦频率选择在多目标、低信噪比条件下性能显著衰减问题,改进提出了一种基于全频域特征值分布的聚焦频率点选取方法,

以提高宽带信号 DOA 估计性能。本文方法兼顾计算效率与准确性,计算耗时相对于最佳聚焦频率选择方法大幅降低,定位精度优于中心频率和最佳聚焦频率选择方法。在仿真和实验验证中均表现出优越估计性能,验证了本文方法的有效性。

### 利益冲突声明/Conflict of Interests

所有作者声明不存在利益冲突。

All authors declare no relevant conflict of interests.

### 参考文献:

- [1] 禹秀梅,郑文康,王立府,等. 基于声压振速联合处理的稀疏协方差 DOA 估计[J]. 中北大学学报(自然科学版), 2024, 45(3): 301-308.  
YU Xiumei, ZHENG Wenkang, WANG Lifu, et al. Sparse covariance DOA estimation based on combined information processing of sound pressure and particle velocity [J]. Journal of North University of China (Natural Science Edition), 2024, 45(3): 301-308. (in Chinese)
- [2] 黄嘉俊,张靖奇,何伟嘉,等. 基于迂回式匹配追踪算法的 DOA 估计方法[J]. 中北大学学报(自然科学版), 2022, 43(5): 424-430.  
HUANG Jiajun, ZHANG Jingqi, HE Weijia, et al. DOA estimation method based on detouring matching pursuit algorithm [J]. Journal of North University of China (Natural Science Edition), 2022, 43(5): 424-430. (in Chinese)
- [3] DING X R, XU W B, WANG Y, et al. Wideband 2d DoA estimation with uniform circular array [J]. IEEE Sensors Journal, 2024, 24(7): 11585-11598.
- [4] LIN Y X, XIE C C, XIE X M, et al. An improved DOA algorithm for wideband coherent signals [C]// 2023 IEEE International Conference on Mechatronics and Automation (ICMA), 2023: 2436-2441.
- [5] ZHANG J, BAO M, CHEN Z F, et al. Markov chain-based frequency correlation processing algorithm for wideband DOA estimation [J]. Signal Processing, 2023, 208: 108968.
- [6] ZHANG W T, MIAO C, JIANG M J, et al. Single-source DOA estimation for wideband LFM signal: Time-delay mixing and enhanced self-mixing music methods [J]. Circuits, Systems, and Signal Processing, 2025, 44: 218-238.
- [7] DONG J, WANG L F, QIAO S H, et al. House swallow optimizer: A new metaheuristic algorithm and its application for DOA estimation of broadband

- signal invector hydr-o-phone array[J]. IEEE Access, 2025, 13: 83717-83737.
- [8] MA F Q, ZHANG X T. Wideband DOA estimation based on focusing signal subspace[J]. Signal, Image and Video Processing, 2019, 13(4): 675-682.
- [9] BO L K, XIONG J Y, LUO L Y. A novel wideband DOA estimation method using direction-free focusing matrix[C]//3rd International Conference on Computer Science and Network Technology, 2013: 1065-1069.
- [10] 毕杨, 王英民, 王奇. 双重优化的宽带聚焦波束形成算法研究[J]. 兵工学报, 2017, 38(8): 1563-1571.  
BI Yang, WANG Yingmin, WANG Qi. Research on dual optimized broadband beamforming algorithm[J]. Acta Armamentarii, 2017, 38(8): 1563-1571. (in Chinese)
- [11] 张玉峰, 张应慧, 王华奎. 聚焦频率对相干子空间算法性能的影响[J]. 科技情报开发与经济, 2007, 17(26): 170-171.  
ZHANG Yufeng, ZHANG Yinghui, WANG Huakui. Influences of focusing frequency on algorithm performance of CSM(Coherent Signal-subspace Method)[J]. Sci-Tech Information Development & Economy, 2007, 17(26): 170-171. (in Chinese)
- [12] 陈洪光, 余圣发, 沈振康. 基于一致聚焦的宽带信号波达方向估计[J]. 通信学报, 2006, 27(1): 64-68.  
CHEN Hongguang, YU Shengfa, SHEN Zhenkang. Direction finding method for wideband signals base on consistent focusing [J]. Journal on Communications, 2006, 27(1): 64-68. (in Chinese)
- [13] 李建峰, 朱珂慧, 江航, 等. 基于多频率聚焦与融合
- 的相干辐射源时差直接定位方法[J]. 电子学报, 2023, 51(8): 2110-2117.
- LI Jianfeng, ZHU Kehui, JIANG Hang, et al. A direct position determination method of TDOA based on multi-frequency focusing and fusion for coherent radiation sources[J]. Acta Electronica Sinica, 2023, 51(8): 2110-2117. (in Chinese)
- [14] 贾思宇, 路茗, 丁华泽, 等. 一种改进的信号子空间聚焦宽带DOA估计算法[J]. 计算机工程, 2022, 48(1): 175-181.  
JIA Siyu, LU Ming, DING Huaze, et al. A modified wideband DOA estimation algorithm for focusing signal subspace [J]. Computer Engineering, 2022, 48(1): 175-181. (in Chinese)
- [15] CAO J, XIONG J J, XU W Z, et al. Broadband signal DOA estimation based on two-sided correlation transformation using array manifold interpolation[J]. IET Radar, Sonar & Navigation, 2025, 19: e12669.
- [16] HUNG H, KAVEH M. Focussing matrices for coherent signal-subspace processing[J]. IEEE Transactions on Acoustics, Speech, and Signal Processing, 1988, 36(8): 1272-1281.
- [17] 张爱莲. 基于阵列稀疏与目标稀疏的宽带DOA估计算法[D]. 杭州: 浙江大学, 2018.
- [18] YANG B, QUAN C S, WANG Y B, et al. Real-MAN: A real-recorded and annotated microphone array dataset for dynamic speech enhancement and localization[J]. Advances in Neural Information Processing Systems, 2024, 37: 105997-106019.