

DOI: 10.19650/j.cnki.cjsi.J2412995

一种面向高分辨率声学成像的频带加权方法*

白宗龙^{1,2}, 张君燕¹, 刘成刚¹

(1. 华北电力大学电子与通信工程系 保定 071003; 2. 河北省电力物联网重点实验室 保定 071003)

摘要: 声学成像是噪声源定位、异响诊断等应用的关键技术。由于声信号为非调制宽带信号,所以现有声学成像方法将麦克风阵列数据划分为若干子频带,然后分别对每个子频带进行声学成像。但是声信号在各频带的能量分布不均匀,导致部分子频带可能因信噪比过低产生错误估计,严重影响声学成像准确度。针对该问题,开展了基于复高斯混合模型的频带加权方法研究。通过联合利用多频带数据赋予各频带权重,降低出现错误估计频带对声学成像准确度的影响。为验证提出的方法的有效性,进行实验验证,利用误判率、漏检率、均方根误差等指标衡量声学成像准确度。实验结果显示本方法有效提高了声学成像准确度,特别是在信噪比低于 10 dB 条件下降低误判率 2.1% 以上。

关键词: 声学成像; 宽带信号; 复高斯混合模型; 频带加权

中图分类号: TH701 文献标识码: A 国家标准学科分类代码: 510.40

A band-weighting method for high-resolution acoustic imaging

Bai Zonglong^{1,2}, Zhang Junyan¹, Liu Chenggang¹

(1. Department of Electronic and Communication Engineering, North China Electric Power University, Baoding 071003, China;
2. Hebei Key Laboratory of Power Internet of Things Technology, North China Electric Power University, Baoding 071003, China)

Abstract: Acoustic imaging is a key technology for applications such as noise source localization and abnormal sound diagnosis. Since the acoustic signals are non-modulated broadband signals, existing acoustic imaging methods divide microphone array data into several sub-bands and then perform acoustic imaging on each sub-band separately. However, the energy distribution of the acoustic signals across different frequency bands is uneven, leading to potential estimation errors in some sub-bands due to low signal-to-noise ratios, significantly impacting the accuracy of acoustic imaging. To address this issue, research was conducted on band-weighting methods based on complex Gaussian mixture models. By jointly utilizing data from multiple frequency bands to assign weights to each sub-band, the impact of sub-bands with erroneous estimates on the accuracy of acoustic imaging is reduced. To validate the effectiveness of the proposed method, experimental verification was conducted, measuring the accuracy of acoustic imaging using indicators such as the false alert rate, miss detection rate, and root mean square error. Experimental results demonstrate that the method effectively improves the accuracy of acoustic imaging, particularly reducing the false alert rate by more than 2.1% under conditions where the signal-to-noise ratio below 10 dB.

Keywords: acoustic imaging; broadband signals; complex gaussian mixture model; band-weighting

0 引言

声学成像是阵列信号处理领域的关键技术,通过分析麦克风接收到的信号之间的时延估计声源的空间位置,是水下声源成像^[1-2]、机械故障异响定位^[3-4]、语音源定位^[5-6]、气体泄漏点定位^[7-10]、气动噪声源分析^[11-14]等

应用的核心关键技术。除此之外,声学成像技术广泛应用于电力设备健康状态监测和故障诊断,是电力听觉的关键技术。声信号是非调制宽带信号,所以声学成像需要考虑多个频带的信息。为联合利用多频带数据进行方位角估计,研究人员提出了多种方法,例如相干信号子空间法(coherent signal subspace method, CSSM)^[15-16]和投影影子空间正交性测试法(test of orthogonality of projected

收稿日期:2024-06-26 Received Date: 2024-06-26

* 基金项目:国家自然科学基金(12404545)、中央高校基本科研业务费专项资金(2024MS118)项目资助

subspace, TOPS)^[17]。然而,这些方法需要足够的测量数据和传声器数量,否则其性能会严重下降,制约了实际应用。

为提高声学成像准确度,稀疏信号恢复方法被应用到声学成像领域中,通过利用声源在空间中的稀疏性提高了分辨率,并且对测量数据和传声器数量要求低于传统方法^[18]。近年来,稀疏信号恢复方法在声学成像实际应用中受到广泛关注。具体而言:为解决语音增强过程中方位角估计分辨率不足的问题,Xenaki等^[19]提出一种稀疏贝叶斯学习方法,提高了语音源的定位分辨率;另外,白宗龙等^[20-22]提出两种稀疏贝叶斯学习框架,用于提高不同信噪比条件下声学成像准确度,并在此基础上,提出一种近似零范数约束的稀疏贝叶斯学习方法;为解决气体泄漏点定位准确度不足的问题,Ning等^[23-24]提出一种基于同步正交匹配追踪(simultaneous orthogonal matching pursuit, SOMP)和特征值分解(singular value decomposition, SVD)的宽带声学成像方法,用于提高多源声学成像的准确度;在此基础上,利用截断核函数正则化策略扩展麦克风阵列孔径,提高了声学成像准确度;在宽带声学成像方面,Shi等^[25-27]引入宽带声全息(wideband acoustical holography, WBH)和l1范数凸优化进行宽带声学成像。在此基础上,分析和比较了统计优化近场声全息、宽带声全息、l1范数最小化、混合压缩采样等方法的声学成像性能。为解决风力机叶片损伤检测中声学成像分辨率不足的问题,Sun等^[28]提出一种基于广义极小极大凹惩罚函数的声学成像方法,用于提高损伤的细微定位度。为解决水下声学成像分辨率不足的问题,Guo等^[29]提出了一种稀疏贝叶斯波束形成方法,提高了水下声学成像分辨率和准确度。此外,在频带加权方面,现有基于稀疏信号恢复的声学成像主要采用频带平均加权方法^[30-32]。文献[33]提出一种基于频率信噪比加权的可控响应功率定位算法,通过估计每个频率信噪比对频带进行加权,然后利用可控响应功率方法定位,提高了鲁棒性和正确率。

通过现有研究可知,稀疏信号恢复技术可有效提高声学成像的分辨率和准确度。然而,现有基于稀疏信号恢复的声学成像方法主要采用频带平均加权,缺少针对出现错误估计频带的处理机制。为解决该问题,本文开展基于复高斯混合模型的频带加权方法研究,在高分辨率声学成像的基础上,构建复高斯混合模型进行频带加权,降低错误估计频带的影响,提高声学成像准确度。具体包括:在分别估计各频带声源位置的基础上,利用复高斯混合模型描述多频带测量数据与各频带声源位置之间的关系,并赋予各频带独立权重。然后引入隐藏变量并采用期望最大化(expectation maximization, EM)方法进行变量更新,通过学习多频带

数据赋予各频带以不同的权重,降低错误估计频带对声学成像准确度的影响。

1 问题分析及数学模型构建

1.1 问题分析

声信号为非调制宽带信号,所以在声学成像过程中需要将麦克风阵列数据分割为若干子频带并分别进行处理。如图1所示,目标平面存在多个声源的情况下,声源在不同频带内的空间分布存在差异,又由于声源的能量在各频带分布不均匀,极易导致部分频带因信噪比偏低造成错误估计,直接影响声源数量和位置的估计结果,造成成像准确度严重降低。因此,研究面向高分辨率声学成像的频率加权方法对提高声学成像的准确度具有重要意义。

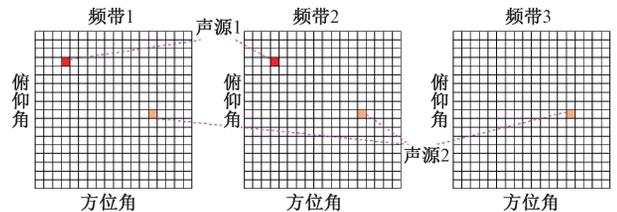


图1 宽带声源空间分布

Fig. 1 Diagram of spatial distribution of wideband sound sources

1.2 数学模型构建

1) 声信号传播模型

声学成像利用声信号传播到不同麦克风之间的时延进行声源的空间位置估计。假设目标平面存在 K_0 个宽带声源,使用阵元数为 M 的多臂螺旋麦克风阵列接收数据,那么声信号的传播模型为

$$y_m = \sum_{k_0=1}^{K_0} s_{k_0}(t - \tau_m(\mathbf{P}_{k_0})) + w_m \quad (1)$$

其中, y_m 表示第 m 个麦克风接收到的数据, m 为麦克风的索引, $s_{k_0}(t)$ 为第 k_0 个声源的时域信号, k_0 为声源的索引, $\tau_m(\mathbf{P}_{k_0})$ 表示第 k_0 个声源传播到第 m 个麦克风的时延, \mathbf{P}_{k_0} 表示第 k_0 个声源的空间位置坐标,可采用极坐标系或笛卡尔坐标系进行描述, w_m 表示第 m 个麦克风的测量噪声。

2) 过完备表示观测模型构建

实际过程中,需要利用麦克风阵列采集到的数据估计声源在目标平面的位置。通常先采用短时傅里叶变换(short time fourier transform, STFT)将时域数据转换为频域数据,并构建如下过完备表示观测模型:

$$\mathbf{Y}_f = \mathbf{A}_f \mathbf{S}_f + \mathbf{W}_f \quad (2)$$

其中,

$$\begin{cases} \mathbf{Y}_f = [\mathbf{y}_{f,\cdot,1}, \dots, \mathbf{y}_{f,\cdot,t}, \dots, \mathbf{y}_{f,\cdot,T}] \in \mathbb{C}^{M \times T} \\ \mathbf{A}_f = [\mathbf{a}_{f,\cdot,1}, \dots, \mathbf{a}_{f,\cdot,k}, \dots, \mathbf{a}_{f,\cdot,K}] \in \mathbb{C}^{M \times K} \\ \mathbf{S}_f = [\mathbf{s}_{f,\cdot,1}, \dots, \mathbf{s}_{f,\cdot,t}, \dots, \mathbf{s}_{f,\cdot,T}] \in \mathbb{C}^{K \times T} \\ \mathbf{W}_f = [\mathbf{w}_{f,\cdot,1}, \dots, \mathbf{w}_{f,\cdot,t}, \dots, \mathbf{w}_{f,\cdot,T}] \in \mathbb{C}^{M \times T} \end{cases} \quad (3)$$

其中, f 表示子频带的索引, $\mathbf{y}_{f,\cdot,t} \in \mathbb{C}^M$ 表示第 t 个观测向量, \mathbb{C} 表示复数空间, \mathbf{Y}_f 表示由 T 个观测向量组成的观测矩阵, $\mathbf{a}_{f,\cdot,k}$ 表示第 k 个扫描点对应的导向矢量, K 表示扫描点数量, \mathbf{A}_f 表示由所有导向矢量组成的传播矩阵, $\mathbf{s}_{f,\cdot,t}$ 表示第 t 个信号向量, 其中的元素代表每个扫描点上的复值信号, \mathbf{S}_f 表示由所有信号向量组成的信号矩阵, 由于预定义扫描点数量远大于实际声源数量, 所以 \mathbf{S}_f 是一个块稀疏矩阵, $\mathbf{w}_{f,\cdot,t}$ 表示第 t 个噪声向量, \mathbf{W}_f 表示由噪声向量构成的噪声矩阵。观测模型式 (2) 中测量数 M 远小于 K , 因此, 称式 (2) 为过完备表示模型, 其构建方法如图 2 所示。

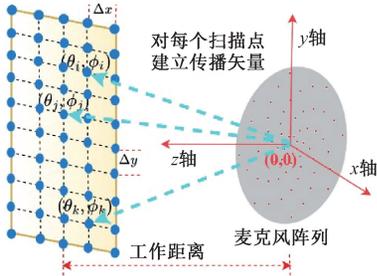


图 2 过完备表示观测模型构建

Fig. 2 Diagram of over-completed observation model

通过构建过完备观测模型, 声学成像转化为稀疏信号恢复问题, 即根据观测数据 \mathbf{Y}_f 和传播矩阵 \mathbf{A}_f 求解信号矩阵 \mathbf{S}_f 的问题。

稀疏贝叶斯学习方法具有自回归特性和不确定估计特性, 稀疏信号恢复的准确度高, 抗噪声干扰能力强, 因此本文采用稀疏贝叶斯学习方法进行单频带信号恢复, 因此构建如下稀疏贝叶斯模型。

3) 稀疏贝叶斯模型

式 (2) 所示的过完备表示观测模型中, 噪声矩阵 \mathbf{W}_f 中的变量服从独立的零均值复高斯分布, 即:

$$p(\mathbf{W}_f) = \prod_{t=1}^T \mathcal{CN}(\mathbf{w}_{f,\cdot,t} | \mathbf{0}, \rho_f^{-1} \mathbf{I}_M) \quad (4)$$

其中, $\mathcal{CN}(\cdot)$ 表示复高斯分布, $\rho_f^{-1} \mathbf{I}_M$ 表示协方差矩阵, \mathbf{I}_M 表示维度为 M 的单位矩阵。为自适应更新参数 ρ_f , 假设 ρ_f 服从参数为 c 和 d 的 Gamma 分布:

$$p(\rho_f) = \mathcal{G}(\rho_f | c, d) \quad (5)$$

其中, $\mathcal{G}(\cdot)$ 表示 Gamma 分布。根据式 (2) 和 (4), 似然函数为:

$$p(\mathbf{Y}_f | \mathbf{S}_f, \rho_f) = \prod_{t=1}^T \mathcal{CN}(\mathbf{y}_{f,\cdot,t} | \mathbf{A}_f \mathbf{s}_{f,\cdot,t}, \rho_f^{-1} \mathbf{I}_M) \quad (6)$$

由于信号矩阵 \mathbf{S}_f 为块稀疏矩阵, 为有效利用该先验知识, 构建如下多层贝叶斯框架:

第 1 层, 使用多维复高斯分布描述信号, 即:

$$p(\mathbf{S}_f | \boldsymbol{\lambda}) = \prod_{t=1}^T p(\mathbf{s}_{f,\cdot,t} | \boldsymbol{\lambda}) = \prod_{t=1}^T \mathcal{CN}(\mathbf{s}_{f,\cdot,t} | \mathbf{0}, \boldsymbol{\Lambda}^{-1}) \quad (7)$$

其中, $\boldsymbol{\lambda}$ 表示方差向量, $\boldsymbol{\Lambda}$ 表示以方差向量为对角元素的 diagonal 矩阵。

第 2 层, 将方向向量中的元素视为变量, 并使用 Gamma 分布描述, 则有:

$$p(\boldsymbol{\lambda} | a, b) = \prod_{k=1}^K p(\lambda_k | a, b) = \prod_{k=1}^K \mathcal{G}(\lambda_k | a, b) \quad (8)$$

其中, a 和 b 为 Gamma 分布的参数。根据式 (6) 和 (7), 通过对 $\boldsymbol{\lambda}$ 积分可以得到信号 \mathbf{S}_f 关于 a 和 b 的边缘分布:

$$p(\mathbf{S}_f) = \prod_{t=1}^T \int p(\mathbf{s}_{f,\cdot,t} | \boldsymbol{\lambda}) p(\boldsymbol{\lambda} | a, b) d\boldsymbol{\lambda} = \prod_{t=1}^T \prod_{k=1}^K \frac{b^a \Gamma(a+1)}{\pi \Gamma(a)} (b + s_{f,k,t}^2)^{-(a+1)} \quad (9)$$

其中, $\Gamma(\cdot)$ 表示 Gamma 函数。根据式 (9) 所示边缘函数, 多层贝叶斯框架实际赋予信号变量 Student-t 分布, 当 a 和 b 取值接近于 0 时, 可有效鼓励稀疏解。

式 (5) ~ (8) 构成稀疏贝叶斯学习模型, 通过模型求解可获得信号矩阵的估计, 具体过程见 1.3 节第 1 部分。由于信噪比低等因素, 部分频带可能出现错误估计, 严重影响声学成像准确度。为降低出现错误估计频带对声学成像准确度的影响, 进一步构建如下复高斯混合模型。

4) 复高斯混合模型构建

通过对稀疏贝叶斯学习模型求解, 可求得信号矩阵的估计值, 即 $\langle \mathbf{S}_f \rangle = \boldsymbol{\Omega}_f$ 。其中, $\langle \cdot \rangle$ 表示变量的期望。为得到估计 $\boldsymbol{\Omega}_f$ 的可信度, 构建如下复高斯分布模型, 从而赋予出现错误估计频带以小权重。

根据可控波束响应的概率 (probabilistic steered response power, PSRP) 模型^[34-35]可知, 测量数据变量的分布可表示为:

$$p(\{\mathbf{Y}_f\}_{f=1}^F | \{\mathbf{S}_f\}_{f=1}^F) = \prod_{f=1}^F \omega_f \prod_{t=1}^T \mathcal{CN}(\mathbf{y}_{f,\cdot,t} | \boldsymbol{\mu}_{f,\cdot,t}, \boldsymbol{\eta} \mathbf{I}_M) \quad (10)$$

其中, $\{\cdot\}$ 表示变量的集合, ω_f 表示第 f 个频带的权重, 可视作 $\boldsymbol{\Omega}_f$ 的可信度, $\boldsymbol{\mu}_{f,\cdot,t}$ 为 $\boldsymbol{\Omega}_f$ 的第 t 列, $\boldsymbol{\eta}$ 为预定义的协方差矩阵参数。通过模型 (10) 可知当 $\boldsymbol{\Omega}_f$ 与测量数据出现偏差时, 权重 ω_f 变小。通过求解权重 ω_f 并进行频带加权, 可降低错误估计频带对声学成像的影响。具体过程见 1.3 节第 2 部分。

1.3 数学模型求解

为降低宽带声学成像中错误估计频带的影响, 提高

声学成像的准确度,本文采用稀疏贝叶斯学习模型进行信号估计,并提出一种复高斯混合模型用于赋予错误估计频带小权重。下面进行模型参数求解。

1) 稀疏贝叶斯学习模型参数求解

由于设计涉及多重积分,隐藏变量的实际后验分布难以求解,因此本文采用变元贝叶斯推断 (variational bayesian inference, VBI) 技术进行参数更新。变元贝叶斯推断技术使用变元分布 $q(\Theta)$ 近似实际后验,其中 Θ 表示所有隐藏变量的集合,并且变元分布 $q(\Theta)$ 可以被因子化表示,即:

$$q(\Theta) = \prod_{f=1}^F q(S_f) q(\rho_f) q(\lambda) \quad (11)$$

由于本文构建的稀疏贝叶斯学习模型中,所有变量均采用共轭先验分布建模,所以各隐藏变量的近似分布均可以通过最小化变元分布和实际分布之间的 Kullback-Leibler 距离求得,即有:

$$\ln q(\theta_i) = E_{q(\theta_i)}(\ln p(\{Y_f\}_{f=1}^F, \Theta)) + c_j \quad (12)$$

其中, θ_i 表示集合 Θ 中的第 i 个隐藏变量, $E(\cdot)$ 表示期望, θ_i 表示集合 Θ 去除元素 θ_i 后的子集, c_j 表示归一化因子。由稀疏贝叶斯学习模型可计算联合分布,如式 (14) 所示。其中, $\|\cdot\|$ 表示 Euclidean 范数。

根据式 (12) 和 (14) 可以推导出所有隐藏变量的近似后验分布以及变量更新规则,具体步骤如下:

变量 S_f 的更新: 变量 $s_{f,\cdot,t}$ 的近似后验分布如式 (13) 所示。

$$\ln q(s_{f,\cdot,t}) = -s_{f,\cdot,t}^H (\langle \rho_f \rangle \mathbf{A}_f^H \mathbf{A}_f + \langle \Lambda \rangle) s_{f,\cdot,t} + \langle \rho_f \rangle s_{f,\cdot,t}^H \mathbf{A}_f^H \mathbf{y}_{f,\cdot,t} + \langle \rho_f \rangle \mathbf{y}_{f,\cdot,t}^H \mathbf{A}_f s_{f,\cdot,t} - \langle \rho_f \rangle \mathbf{y}_{f,\cdot,t}^H \mathbf{y}_{f,\cdot,t} + c_s \quad (13)$$

其中, $\langle \cdot \rangle$ 表示变量的期望, $(\cdot)^H$ 表示共轭转置, c_s 表示近似分布的归一化因子。根据式 (13) 可知变量 $s_{f,\cdot,t}$ 的近似后验分布为多维复高斯分布,因此, $s_{f,\cdot,t}$ 的更新规则如式 (15) 和 (16) 所示。

$$\begin{aligned} \ln p(\{Y_f\}_{f=1}^F, \Theta) &= \ln p(\{Y_f\}_{f=1}^F | \{S_f\}_{f=1}^F, \{\rho_f\}_{f=1}^F) \times \\ & p(\{\rho_f\}_{f=1}^F) p(\{S_f\}_{f=1}^F | \lambda) p(\lambda) = \\ & TM \sum_{f=1}^F \ln \rho_f - \sum_{f=1}^F \sum_{t=1}^T \rho_f \| \mathbf{y}_{f,\cdot,t} - \mathbf{A}_f s_{f,\cdot,t} \|^2 + \\ & FT \ln \Lambda - \sum_{f=1}^F \sum_{t=1}^T s_{f,\cdot,t}^H \Lambda s_{f,\cdot,t} + \sum_{k=1}^K ((a-1) \ln \lambda_k - \\ & b \ln \lambda_k) + \sum_{f=1}^F ((c-1) \ln \rho_f - d \rho_f) \end{aligned} \quad (14)$$

$$\Omega_f = \langle S_f \rangle = \langle \rho_f \rangle \Sigma_f \mathbf{A}_f^H \mathbf{Y}_f \quad (15)$$

$$\begin{aligned} \Sigma_f &= \langle \Lambda \rangle^{-1} - \langle \Lambda \rangle^{-1} \mathbf{A}_f^H \times (\langle \rho_f \rangle^{-1} \mathbf{I}_M + \\ & \mathbf{A}_f \langle \Lambda \rangle^{-1} \mathbf{A}_f^H)^{-1} \mathbf{A}_f \langle \Lambda \rangle^{-1} \end{aligned} \quad (16)$$

其中, Ω_f 和 Σ_f 分别为变量 S_f 的期望和方差,式 (16) 的推导利用了 Woodbury 矩阵恒等式降低计算复杂度。

变量 λ 的更新: 变量 λ 的近似后验分布如式 (17) 所示。

$$\begin{aligned} \ln q(\lambda_k) &= (a-1 + FT) \ln \lambda_k - \\ & \lambda_k \left(\sum_{f=1}^F \text{tr}(\langle S_f^H S_f \rangle) + b \right) + c_\lambda \end{aligned} \quad (17)$$

其中, $\text{tr}(\cdot)$ 表示矩阵的迹。变量 λ 的近似后验分布为 Gamma 分布,其更新规则为:

$$\langle \lambda_k \rangle = \frac{FT + a}{\sum_{f=1}^F \text{tr}(\langle S_f^H S_f \rangle) + b} \quad (18)$$

变量 ρ_f 的更新: 变量 ρ_f 的近似后验分布如式 (19) 所示。

$$\begin{aligned} \ln q(\rho_f) &= (FMT + c - 1) \ln \rho_f - \\ & \left(\sum_{f=1}^F \| \mathbf{Y}_f - \mathbf{A}_f \langle S_f \rangle \|^2_F + d \right) \rho_f + c_\rho \end{aligned} \quad (19)$$

其中, $\|\cdot\|_F$ 表示 Frobenius 范数。变量 ρ_f 的近似后验分布为 Gamma 分布,其更新规则为:

$$\langle \rho_f \rangle = \frac{FMT + c}{\sum_{f=1}^F \| \mathbf{Y}_f - \mathbf{A}_f \langle S_f \rangle \|^2_F + d} \quad (20)$$

2) 复高斯混合模型求解

一旦最大化式 (10) 所示的复高斯混合模型,权重 ω_f 即可表示 Ω_f 的可信度,但是由于该复高斯混合模型涉及高维度运算,难以直接最大化,因此本文采用期望最大化 (Expectation Maximization, EM) 算法进行变量更新。首先,引入隐藏变量 $\mathbf{z} \in \{0, 1\}^F$, \mathbf{z} 中有且仅有一个元素 z_f 为 1,其他元素均为 0。 z_f 可视为 Ω_f 的指示变量。令 $p(z_f = 1) = \omega_f$, 那么有:

$$p(\mathbf{z}, \boldsymbol{\omega}) = \prod_{f=1}^F \omega_f^{z_f} \quad (21)$$

根据式 (21), 观测变量关于隐藏变量 \mathbf{z} 的条件概率为:

$$p(\{Y_f\}_{f=1}^F | \mathbf{z}) = \prod_{f=1}^F \left(\prod_{t=1}^T \mathcal{CN}(y_{f,\cdot,t} | \mathbf{A}_f \boldsymbol{\mu}_{f,\cdot,t}, \boldsymbol{\eta} \mathbf{I}_M) \right)^{z_f} \quad (22)$$

根据式 (10) 和 (22), 有:

$$\begin{aligned} p(\{Y_f\}_{f=1}^F, \boldsymbol{\omega}, \mathbf{z}) &= \\ & \prod_{f=1}^F \left(\omega_f \prod_{t=1}^T \mathcal{CN}(y_{f,\cdot,t} | \mathbf{A}_f s_{f,\cdot,t}, \boldsymbol{\eta} \mathbf{I}_M) \right)^{z_f} \end{aligned} \quad (23)$$

下面使用 EM 算法求解复高斯模型参数。

E 步: 根据式 (23), E 步如式 (24) 所示。

$$\begin{aligned} \mathcal{Q}(\boldsymbol{\omega} | \boldsymbol{\omega}^{\text{old}}) &= E(\log p(\{Y_f\}_{f=1}^F, \mathbf{z} | \boldsymbol{\omega}^{\text{old}})) = \\ & E(z_f | \{Y_f\}_{f=1}^F, \boldsymbol{\omega}^{\text{old}}) (\ln \hat{\omega}_f + \psi_f) \end{aligned} \quad (24)$$

其中, $\hat{\omega}^{\text{old}}$ 为上一次迭代的估计值, ψ_f 为:

$$\begin{aligned} \psi_f &= \sum_{t=1}^T \ln \mathcal{CN}(y_{f,\cdot,t} | \mathbf{A}_f \boldsymbol{\mu}_{f,\cdot,t}, \boldsymbol{\eta} \mathbf{I}_M) = \\ & \sum_{t=1}^T \left(-M(\ln \boldsymbol{\eta} + \ln \boldsymbol{\pi}) - \frac{1}{\boldsymbol{\eta}} \| \mathbf{y}_{f,\cdot,t} - \mathbf{A}_f \boldsymbol{\mu}_{f,\cdot,t} \|^2 \right) \end{aligned} \quad (25)$$

并且有:

$$E(z_f | \{Y_f\}_{f=1}^F, \hat{\omega}^{\text{old}}) = \frac{\hat{\omega}_f^{\text{old}} \exp(\psi_f)}{\sum_{f=1}^F \hat{\omega}_f^{\text{old}} \exp(\psi_f)} = \langle z_f \rangle \quad (26)$$

M 步:在 M 步,需在约束下最大化式(24),如式(27)所示。

$$\omega^{\text{new}} = \underset{\omega}{\operatorname{argmax}} \mathcal{Q}(\omega | \omega^{\text{old}}) \text{ s. t. } \sum_{f=1}^F \omega_f = 1; 0 \leq \omega_f \leq 1 \quad (27)$$

求解式(27)可得:

$$\hat{\omega}_f^{\text{new}} = \frac{\langle z_f \rangle}{\sum_{f=1}^F \langle z_f \rangle} \quad (28)$$

给定变量 ω 和 z 初始值,通过 E 步和 M 步的迭代更新直到收敛,即可实现复高斯混合模型的参数求解。求解后的 $\hat{\omega}_f$ 作为权重与相应频带的恢复信号 Ω_f 相乘即可降低错误估计频带对声学成像的影响。

3) 计算复杂度分析

为直观展示算法,给出如下总结。请注意为提高计算速度,本文采用基剔除的方法进行降维。

输入:观测数据 $\{Y_f\}_{f=1}^F$ 、传播矩阵 $\{A_f\}_{f=1}^F$ 、模型参数 $\{a, b, c, d, \eta\}$ 、最大迭代次数 I_{\max} 、收敛误差 e_0 、基剔除阈值 T_p ;

输出:声成像矩阵 Ω ;

(1) 初始化参数,令 $\lambda = I_K$, ω 中的元素随机生成并归一化,基索引集合为 $\{1, \dots, K\}$,使用特征值分解对观测数据降维;

(2) 如果 λ 中最小元素小于 T_p ,则查找 λ 中大于该阈值的元素的索引,并更新基索引集合。根据基索引集合选取基向量并更新传播矩阵;如果基索引集合为空集,则算法停止;

(3) 利用式(15)和(16)更新所有频带的变量 Ω_f 和 Σ_f ;利用式(18)更新变量 λ ;利用公式(20)更新所有频带的 ρ_f ;

(4) 计算所有频带 Ω_f 和上一次迭代 Ω_f^{old} 的差值,如果差值小于 e_0 则停止迭代,否则重复步骤(2)和(3);如果迭代次数大于 I_{\max} ,则停止迭代;

(5) 根据式(25)和(26)更新所有频带的 $\langle z_f \rangle$;根据式(28)更新所有频带的 $\hat{\omega}_f$;

(6) 计算所有频带的 $\hat{\omega}_f$ 和上一次迭代的 $\hat{\omega}_f^{\text{old}}$,如果差值小于 e_0 则停止迭代,否则重复步骤(6);如果迭代次数大于 I_{\max} ,则停止迭代;

(7) 利用 $\hat{\omega}_f$ 作为权重对所有频带的恢复信号进行加权求和,得到声成像矩阵 Ω 。

根据以上流程可知,算法的计算复杂度主要分为两部分,第 1 部分为稀疏贝叶斯学习模型参数更新,其计算复杂度为 $\mathcal{O}(FMK^2I_1)$,其中 K_i 为第 i 次迭代的基数量, I_1 为稀疏贝叶斯学习算法迭代次数;第 2 部分为复高斯混

合模型参数更新,其计算复杂度为 $\mathcal{O}(FTMI_2)$, I_2 为 EM 算法迭代次数。由于算法中使用特征值分解对观测数据降维,所以 T 一般比较小,第 2 部分的计算复杂度小于第 1 部分计算复杂度,即相比于单纯的稀疏贝叶斯学习算法,本文提出的算法可以在不显著增加计算量的条件下降低错误估计频带对声学成像准确度的影响。

本文提出的算法中剔除阈值建议设置为一个较小的值,如 10^{-5} ,用于快速剔除绝大部分无信号源网格点。超参数 η 设置过大时第 2 项的作用减弱,设置过小时会导致对第 2 项敏感,本文建议设置在 $[0.01, 0.1]$ 以内。

2 实验验证

为验证验证提出算法的有效性,分别进行仿真实验和麦克风阵列数据实验。

2.1 仿真实验设计及结果分析

1) 仿真实验设计及参数设定

仿真实验中,采用 64 通道多臂螺旋阵列布局,阵列直径 0.16 m,最内圈直径 0.03 m,阵列布局如图 3 所示。实验中采用 3 个声源,每个声源由不同频率的信号叠加而得,其中,第 1 个声源的频率为 800 和 1 000 Hz,第 2 个声源的频率为 1 000 和 1 500 Hz,第 3 个声源的频率为 800 和 1 500 Hz。麦克风阵列平面和目标平面平行,距离为 2 m,将目标平面按方位角范围 $(-60^\circ, 60^\circ)$ 和俯仰角范围 $(-60^\circ, 60^\circ)$ 均匀划分,间隔均为 2° 。根据声源和麦克风之间的空间关系生成时域阵列数据并叠加高斯白噪声,采样频率设置为 16 kHz,然后利用 STFT 将数据转换为频域数据,每 64 ms 分割为一快拍,FFT 点数设置为 1 024 点,重叠率设为 25%,每 20 快拍设置为 1 帧。利用每帧数据进行声学成像。实验中采用函数反卷积(F-DAMAS)算法^[9]、基于稀疏贝叶斯学习波束形成(FSBL-BF)算法^[19]进行对比,其中,本文算法的参数设置如下: $\{a, b, c, d\}$ 设为 10^{-3} , η 设为 0.1,最大迭代次数设置为 2 000,收敛误差设为 10^{-3} ,基剔除阈值设为 10^{-5} 。

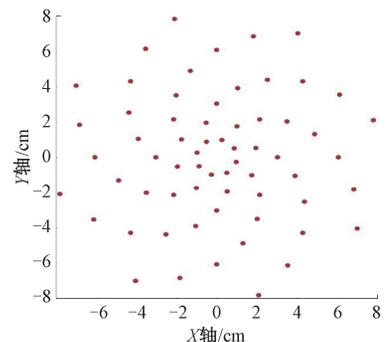


图 3 麦克风阵列布局

Fig. 3 Diagram of multi-arm spiral microphone array

为衡量各算法的声学成像准确性,本文采用误判率 (false alert ratio, FAR)、漏检率 (miss detection ratio, MDR)、均方根误差 (root mean square error, RMSE) 3 个评价指标^[35-36]。如果估计值与实际值的误差在 5° 以内则认为式是正确估计,否则认为是错误估计。FAR 的定义为错误估计的声源数量与实际声源总数的比值;MDR 为未检测出的声源数量与实际声源总数的比值;RMSE 为所有正确检测出的声源位置与实际位置的均方根误差。

2) 实验结果分析

在生成麦克风阵列数据后,分别对 800、1 000、和 1 500 Hz 频带进行信号恢复,并将所有频带的恢复信号进行加权叠加,其中 F-DAMAS 和 FSBL-BF 算法进行平均加权。首先分别给出信噪比 (signal to noise ratio, SNR) 为 20 和 0 dB 条件下各算法的单个成像的伪彩图,如图 4 所示。其中,圆圈为 3 个声源的实际位置。

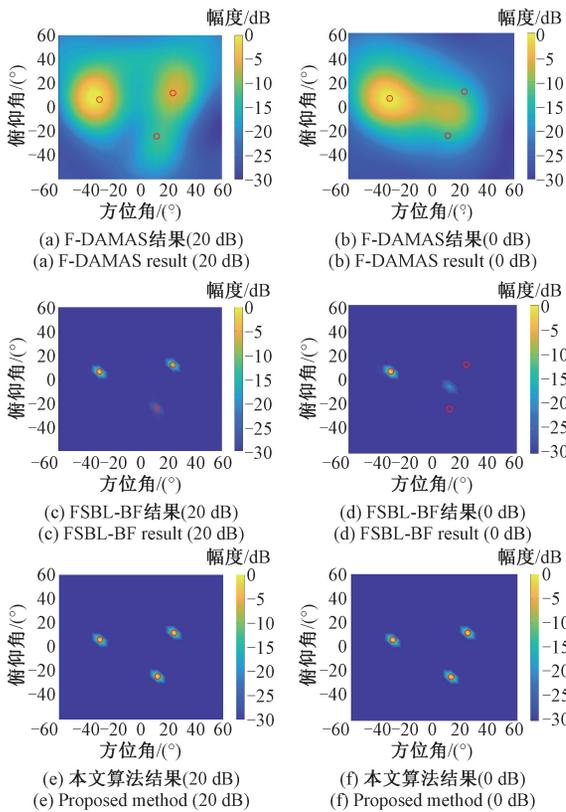


图4 信噪比为 20 dB 和 0 dB 条件下单次成像结果

Fig. 4 Single results under SNR ratios of 20 dB and 0 dB

图 4(a) 和 (b) 分别是 F-DAMAS 在 SNR 为 20 和 0 dB 情况下单次实验结果;图 4(c) 和 (d) 分别是 FSBL-BF 在 SNR 为 20 和 0 dB 情况下的单次实验结果;图 4(e) 和 (f) 分别是本文算法在 SNR 为 20 dB 和 0 dB 情况下的单次实验结果。由图 4 可知,在相同麦克风阵

列尺寸的条件,稀疏贝叶斯学习方法比函数反卷积方法具有更高的分辨率。进一步根据图 4(d) 可知,信噪比降低的情况下,会出现伪峰,严重影响声源数量和位置估计结果。相比于 FSBL-BF 算法,本文中提出的算法通过赋予出现错误估计频带以小权重,有效提高了声学成像准确度。

为定量验证本文提出算法的有效性,进一步开展不同 SNR 条件下的蒙特卡洛实验研究,其中,每次实验中声源在目标平面内的位置随机确定,其它参数设置与第 2.1 节第 1 部分设置相同。每设定一次 SNR 运行 500 次独立实验并取结果的平均值。利用 FAR、MDR、RMSE 衡量各算法的声学成像性能。结果如图 5 所示。

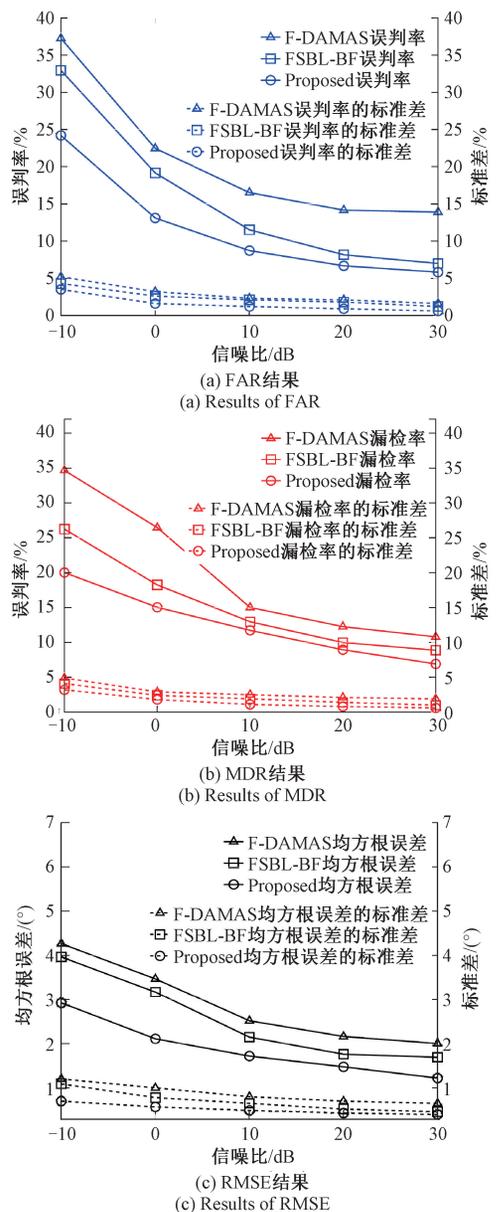


图5 不同 SNR 条件下各算法性能

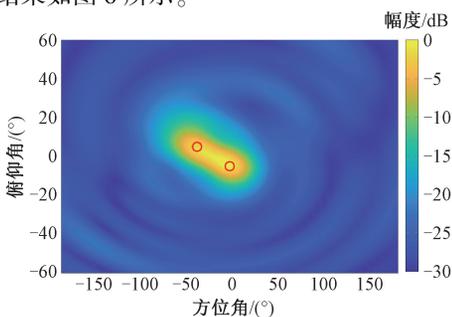
Fig. 5 Results of algorithms versus SNR

根据图 5 可知,随着 SNR 降低,所有算法的声学成像准确性均有所下降,但在相同 SNR 条件下本文提出的算法的 FAR、MDR 和 RMSE 均低于其它算法,特别是 FAR 在 SNR 低于 10 dB 的条件下比其它算法降低了 2.1% 以上,显示出更高的声学成像准确度。

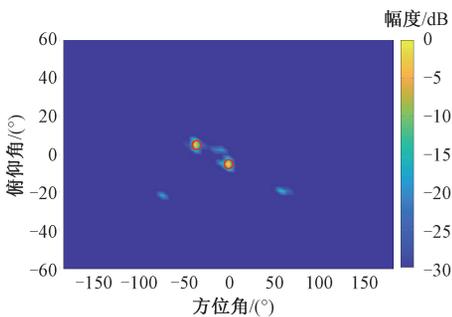
为直观比较不同算法的计算复杂度,给出 20 dB 条件下,不同算法运行 500 次的 CPU 耗时。本文中所有算法通过 MATLAB 2023a 实现,操作系统为 Windows 11, CPU 型号为 Intel(R) i9-10900, CPU 主频为 2.8 GHz, 内存大小为 32 G。F-DAMAS 算法的平均耗时为 0.32 s, 标准差为 0.05 s。FSBL-BF 算法的平均耗时为 0.40 s, 标准差为 0.05 s。本文提出的算法平均耗时为 0.43 s, 标准差为 0.06 s。相比于其它两种算法,本文中提出的算法耗时轻微增加,但声学成像准确度显著增加。

2.2 麦克风阵列数据实验结果分析

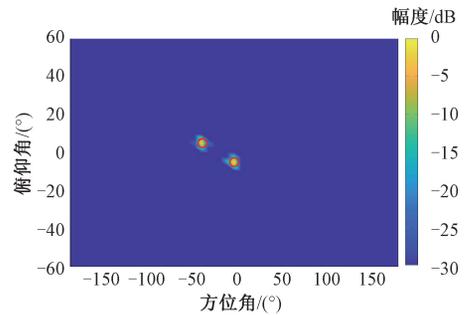
为验证本文提出算法在实际应用的有效性,本文利用声源定位和追踪数据库(acoustic source localization and tracking, LOCATA)中的麦克风阵列数据进行算法比较。LOCATA 数据库是国际上通用的声源定位方面的数据库,给出了不同条件下多种麦克风阵列数据,并且提供了所有声源的实际位置。数据库中 Task6 包含两个声源在不同位置的数据,因此本文采用 Task6 中的数据进行声学成像实验。首先将麦克风阵列数据降采样到 16 kHz, 然后利用 STFT 将时域数据转换为频域数据,STFT 参数与 2.1 节第 1 部分相同。然后利用算法进行声学成像。声学成像结果如图 6 所示。



(a) F-DAMAS 结果
(a) Results of F-DAMAS



(b) FSBL-BF 结果
(b) Results of FSBL-BF



(c) 本文算法结果
(c) Results of the proposed method

图 6 LOCATA 数据的一组成像结果

Fig. 6 Results from LOCATA data

根据图 6 可知 F-DAMAS 算法在两声源距离较近时无法进行准确分辨;FSBL-BF 算法可准确分辨两声源,但存在伪峰;本文提出的方法在准确分辨两声源的基础上减少了伪峰数量。

进一步给出声学成像准确度结果,如表 1 所示。

表 1 麦克风阵列数据实验结果

Table 1 Results of algorithms using real data

数据样本号	算法	误判率/ %	漏检率/ %	均方根 误差/(°)
Task 6, Recording 1	F-DAMAS	21.2	29.4	4.3
	FSBL-BF	23.5	24.7	3.7
	本文算法	16.3	22.6	3.1
Task 6, Recording 2	F-DAMAS	26.7	30.9	4.2
	FSBL-BF	24.6	26.1	3.9
	本文算法	13.5	23.2	3.2
Task 6, Recording 3	F-DAMAS	24.4	29.3	4.6
	FSBL-BF	25.1	27.6	3.8
	本文算法	12.8	21.7	3.0

根据表 1 可知,在使用实际麦克风阵列数据的情况下,本文所提出的算法可有效降低声学成像的 FAR、MDR 和 RMSE,有效提高声学成像的准确度。

3 结 论

针对现有声学成像技术缺少应对出现错误估计频带方法的问题,本研究在稀疏贝叶斯学习框架的基础上,提出一种复高斯混合模型,通过联合利用多频带数据赋予各频带相应的权重,降低错误估计频带对声学成像准确度的影响。仿真和实验结果表明,稀疏贝叶斯学习方法可有效提高声学成像的分辨率,结合本研究提出的复高

斯混合模型可实现低信噪比下高准确度的声学成像。与单独使用稀疏贝叶斯学习算法相比,本研究提出的算法轻微增加了耗时,但低信噪比下声学成像准确度显著增加,提高了声学成像算法的实用价值。本文假设测量噪声为高斯白噪声,未考虑相干噪声影响,后续将开展相干噪声条件下基于稀疏信号恢复的声学成像方面的研究,进一步提高声学成像准确度。

参考文献

- [1] 徐嘉璘, 郭良浩, 任云. 深海海底声反射区多途时延差分析与近海面声源定位 [J]. 声学学报, 2023, 48(4): 618-631.
XU J L, GUO L H, REN Y. Analysis of multipath time delay difference in deep ocean bottom bounce area and its application in near-surface target location [J]. Acta Acustica, 2023, 48(4): 618-631.
- [2] 吴禹沈, 秦继兴, 李整林, 等. 声学滑翔机联合的深海水下声源定位 [J]. 声学学报, 2023, 48(3): 437-446.
WU Y SH, QIN J X, LI ZH L, et al. Source localization based on two underwater acoustic gliders in deep water [J]. Acta Acustica, 2023, 48(3): 437-446.
- [3] 杨红波, 郭磊, 史文库, 等. 道路试验下某重型商用车驾驶室啸叫声源定位 [J]. 振动与冲击, 2022, 41(20): 307-314.
YANG H B, GUO L, SHI W K, et al. Noise source localization of a heavy commercial vehicle cab under a road test [J]. Journal of Vibration and Shock, 2022, 41(20): 307-314.
- [4] ZHANG ZH X, WANG Y Y, LI ZH H, et al. Localization of mechanical and electrical defects in dry-type transformers using an optimized acoustic imaging approach [J]. Plos One, 2023, 18(11): e0294674.
- [5] 刘望生, 潘海鹏, 王明环. 噪声混响下说话人跟踪的多特征自适应 UPF 算法 [J]. 仪器仪表学报, 2022, 43(4): 224-233.
LIU W SH, PAN H P, WANG M H. Adaptive unscented particle filter algorithm based on multi-feature for speaking tracking in noisy and reverberant environments [J]. Chinese Journal of Scientific Instrument, 2022, 43(4): 224-233.
- [6] LI X F, BAN Y T, GIRIN L, et al. Online localization and tracking of multiple moving speakers in reverberant environment [J]. IEEE Journal of Selected Topics in Signal Processing, 2019, 13(1): 88-103.
- [7] 刘轩宇, 李鹏, 张立豪, 等. 联合 TCT 和 W-SpSF 的气体泄漏定位 [J]. 电子测量与仪器学报, 2023, 37(5): 215-222.
LIU X Y, LI P, ZHANG L H, et al. Gas leak location using TCT and W-SpSF [J]. Journal of Electronic Measurement and Instrument, 2023, 37(5): 215-222.
- [8] NING F L, CHENG ZH H, MENG D, et al. A framework combining acoustic features extraction method and random forest algorithm for gas pipeline leak detection and classification [J]. Applied Acoustics, 2021, 182: 108255.
- [9] 赵慎, 李伟, 覃业梅, 等. 传声器阵列函数反卷积声源成像算法研究 [J]. 仪器仪表学报, 2023, 44(10): 112-119.
ZHAO SH, LI W, TAN Y M, et al. Functional deconvolutional approach for the mapping of acoustic sources algorithm of microphone array [J]. Chinese Journal of Scientific Instrument, 2023, 44(10): 112-119.
- [10] 王强, 薛生, 郑晓亮, 等. 埋地管道泄漏三维坐标-波速联合估计方法 [J]. 仪器仪表学报, 2023, 44(3): 69-79.
WANG Q, XUE SH, ZHENG X L, et al. A joint 3D coordinate and wave-speed estimation method for the buried pipe leak localization [J]. Chinese Journal of Scientific Instrument, 2023, 44(3): 69-79.
- [11] MERINO-MARTINEZ R, SIJTSMA P, SNELLEN M, et al. A review of acoustic imaging methods using phased microphone arrays [J]. CEAS Aeronautical Journal, 2019, 10(1): 197-230.
- [12] 初宁, 黄乾, 余亮, 等. 一种基于相位平均的旋转声源高分辨率定位方法 [J]. 振动与冲击, 2021, 40(19): 125-136.
CHU N, HUANG Q, YU L, et al. A high-resolution positioning method of rotating sound source based on phase average [J]. Journal of Vibration and Shock, 2021, 40(19): 125-136.
- [13] 高建正, 吴海军, 蒋伟康. 旋转叶片高频涡脱落噪声

- 源的双阵列声成像试验研究[J]. 振动与冲击, 2021, 40(7): 148-153.
- GAO J ZH, WU H J, JIANG W K. Dual-array acoustic imaging tests of high frequency vortex shedding noise source for rotating blades[J]. Journal of Vibration and Shock, 2021, 40(7): 148-153.
- [14] 高建正, 陈正武, 莫品西, 等. 旋转运动声源的频率波动修正波束形成方法[J]. 声学学报, 2021, 46(4): 576-583.
- GAO J ZH, CHEN ZH W, MO P X, et al. A frequency fluctuation correction beamforming method of rotary motion sound source[J]. Acta Acustica, 2021, 46(4): 576-583.
- [15] WANG H, KAVEH M. Coherent signal-subspace processing for the detection and estimation of angles of arrival of multiple wide-band sources [J]. IEEE Transactions on Acoustics, Speech, and Signal Processing, 1985, 33(4): 823-831.
- [16] 周英钢, 邵佳伟. 对数螺旋阵列的相干信号 DOA 估计研究[J]. 电子测量与仪器学报, 2023, 37(2): 220-227.
- ZHOU Y G, SHAO J W. Research on DOA estimation of coherent signals from logarithmic spiral arrays [J]. Journal of Electronic Measurement and Instrument, 2023, 37(2): 220-227.
- [17] YOON Y S, KAPLAN L M, MCCLELLAN J H. TOPS: New DOA estimator for wideband signals [J]. IEEE Transactions on Signal Processing, 2006, 54(6): 1977-1989.
- [18] MALIOUTOV D, CETIN M, WILLSKY A S. A sparse signal reconstruction perspective for source localization with sensor arrays [J]. IEEE Transactions on Signal Processing, 2005, 53(8): 3010-3022.
- [19] XENAKI A, BOLDT J B, CHRISTENSEN M G. Sound source localization and speech enhancement with sparse Bayesian learning beamforming[J]. The Journal of the Acoustical Society of America, 2018, 143(6): 3912-3921.
- [20] 白宗龙, 师黎明, 孙金玮. 基于自适应 LASSO 先验的稀疏贝叶斯学习算法[J]. 自动化学报, 2022, 48(5): 1193-1208.
- BAI Z L, SHI L M, SUN J W. Sparse Bayesian learning using adaptive LASSO priors [J]. Acta Automatica Sinica, 2022, 48(5): 1193-1208.
- [21] BAI Z L. Sparse Bayesian learning for sparse signal recovery using L1/2 norm[J]. Applied Acoustics, 2023, 207: 109340.
- [22] BAI Z L, ZHANG J Y, SHI L M, et al. Approximating the zero-norm penalized sparse signal recovery using a hierarchical Bayesian framework [J]. Signal Processing, 2024, 218: 109361.
- [23] NING F L, LIU ZH, SONG J H, et al. A robust and efficient compressed sensing algorithm for wideband acoustic imaging [J]. Chinese Journal of Mechanical Engineering, 2020, 33(6): 87-102.
- [24] NING F L, HU J L, HOU H J, et al. Sound source localization of non-synchronous measurements beamforming based on the truncated nuclear norm regularization [J]. Applied Acoustics, 2022, 191: 108688.
- [25] SHI T Y, LIU Y F, BOLTON J S. Spatially sparse sound source localization in an under-determined system by using a hybrid compressive sensing method [J]. The Journal of the Acoustical Society of America, 2019, 146(2): 1219-1229.
- [26] SHI T Y, THOR W, BOLTON J S. Near-field acoustical holography incorporating compressive sampling: Effect of measurement distance and array density [J]. Noise Control Engineering Journal, 2020, 68(6): 470-489.
- [27] SHI T Y, BOLTON J S, THOR W. Acoustic far-field prediction based on near-field measurements by using several different holography algorithms[J]. The Journal of the Acoustical Society of America, 2022, 151(3): 2171-2180.
- [28] SUN SH L, WANG T Y, YANG H X, et al. Damage identification of wind turbine blades using the microphone array under different parametric and measuring conditions: A prototype study with laboratory-scale models [J]. Structural Health Monitoring, 2023, 22(1): 201-215.
- [29] GUO Q J, YANG S Y, ZHOU T, et al. Underwater acoustic imaging via online Bayesian compressive beamforming[J]. IEEE Geoscience and Remote Sensing Letters, 2022, 19: 1-5.
- [30] 陈果, 卢永刚. 宽带声源方位估计的多频稀疏贝叶斯学习改进算法[J]. 仪器仪表学报, 2023, 44(5):

- 302-312.
CHEN G, LU Y G. Improved multi-frequency sparse Bayesian learning method for DOA estimation of the wideband sound source[J]. Chinese Journal of Scientific Instrument, 2023, 44(5): 302-312.
- [31] GEMBA K L, NANNURU S, GERSTOFT P. Robust ocean acoustic localization with sparse Bayesian learning[J]. IEEE Journal of Selected Topics in Signal Processing, 2019, 13(1): 49-60.
- [32] OLIVIERI M, BASTINE A, PEZZOLI M, et al. Acoustic imaging with circular microphone array: A new approach for sound field analysis [J]. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2024, 32: 1750-1761.
- [33] 赵小燕,陈书文,周琳. 基于频率信噪比加权的麦克风阵列声源定位算法[J]. 信号处理, 2020, 36(3): 449-456.
ZHAO X Y, CHEN SH W, ZHOU L. Sound source localization using SNR-based frequency weighting with microphone array [J]. Journal of Signal Processing, 2020, 36(3): 449-456.
- [34] TRAA J, WINGATE D, STEIN N D, et al. Robust source localization and enhancement with a probabilistic steered response power mode [J]. IEEE/ACM Trans. Audio Speech. Lang. Process, 2016, 24(3): 493-503.
- [35] BAI Z L, SHI L M, JENSEN J R, et al. Acoustic DOA estimation using space alternating sparse Bayesian learning[J]. EURASIP Journal on Audio, Speech, and Music Processing, 2021, 14: 1-16.
- [36] DORFAN Y, GANNOT S. Tree-based recursive expectation maximization algorithm for localization of acoustic sources[J]. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2015, 23(10): 1692-1703.

作者简介



白宗龙(通信作者), 2013年于哈尔滨工业大学获得学士学位, 2016年于哈尔滨工业大学获得硕士学位, 2021年于哈尔滨工业大学获得博士学位, 2018年至2020年赴奥尔堡大学联合培养, 现为华北电力大学讲师, 主要研究方向为声源定位和麦克风阵列声学成像。

E-mail: baizongyao@163.com

Bai Zonglong (Corresponding author) received his B. Sc. degree in 2013 from Harbin Institute of Technology, received his M. Sc. degree in 2016 from Harbin Institute of Technology, received his Ph. D. degree in 2021 from Harbin Institute of Technology, engaged in a joint training program at Aalborg University from 2018 to 2020. Now he is a lecturer in North China Electric Power University. His main research interests include sound source localization and acoustic imaging.