

DOI: 10.19650/j.cnki.cjsi.J1905904

快速迭代收缩阈值声源识别算法及其改进^{*}

黄琳森, 徐中明, 张志飞, 贺岩松
(重庆大学汽车工程学院 重庆 400044)

摘 要:近场声全息是一种主流的声学成像技术,在技术实践中可以有效实现声源定位。等效源法可用于修正声源不同数值模型的 ℓ_2 范数正则化逆问题和 ℓ_1 范数正则化逆问题。为了提高传统等效源法的分辨力,减少旁瓣“鬼影”,提出了一种高阶波束形成改良快速迭代收缩阈值声源识别算法 mFISTESM- v ,该方法是在改进的快速迭代收缩/阈值等效源法与高阶波束形成方法相结合的框架下发展而来。通过单声源定位和相干声源定位的分辨力分析的数值模拟以及实验验证了该方法的实用性,结果表明,所提出的方法能较好地达到声源识别及定位的效果,该方法所得主瓣区小于 0.002。

关键词:等效源法;快速迭代收缩阈值;高阶波束形成;传声器阵列;近场声全息;声源识别

中图分类号: TH73 **文献标识码:** A **国家标准学科分类代码:** 460.4

A fast iterative shrinkage threshold sound source identification algorithm and its improvement

Huang Linsen, Xu Zhongming, Zhang Zhifei, He Yansong
(School of Automotive Engineering, Chongqing University, 400044 Chongqing, China)

Abstract: The near-field acoustic holography is an acoustic image method, which can be used to localize sound source effectively in technical practice. The equivalent source method can be applied to fix the ℓ_2 norm regularized inverse problem and the ℓ_1 norm regularized inverse problem for different numerical models of sound sources. To enhance the resolution quality of traditional ESM and narrow the sidelobe, we propose an improved fast iterative shrinkage/thresholding algorithm for sound source identification, which is based on the equivalent source method (mFISTESM- v). This method is a combined application of the modified improved fast iterative shrinkage thresholding equivalent source method and functional beamforming. Its practicability is evaluated by the numerical simulation of the resolution quality analysis, and the experiments of the single sound source location and the coherent sound source location. Results show that the proposed method can achieve excellent sound source identification and location, and the main lobe area obtained by this method is smaller than 0.002.

Keywords: equivalent source method; fast iterative shrinkage threshold; functional beamforming; microphone array; near-field acoustical holography; sound source identification

0 引 言

自从 Williams^[1]于 20 世纪 80 年代提出近场声全息 (near-field acoustical holography, NAH) 以来,其与波束形成 (beamforming, BF) 逐渐成为两种主流声学成像技术,在三维声场重建以及基于近场阵列的声源附近的声场测量方面具有显著优势^[2-4]。作为一种针对边界元法和近

场声全息求解方法的替代方法,等效源法 (equivalent source method, ESM) 有利于快速和稳健地进行求解运算,但是它存在着一个缺点,即没有标准规范来控制这些等效源的形成和方向。为此,宋玉来等^[5]讨论了非规则结构表面多声源声场的全息精度,提出一种基于目标深度识别与球面波叠加逼近的非规则表面多声源近场声全息方法。鉴于实际中采集通道的数量会固有限制,这将导致一个不适定的问题,该问题在逆问题求解过

收稿日期: 2019-12-16 Received Date: 2019-12-16

^{*} 基金项目: 国家自然科学基金 (11874096) 项目资助

程中经常是欠定的。因此,所有方法都需要解决不适定逆问题,以便获得声源或声场的模型来重建声源表面的三维声场。因此,如何进行快速求解成为了一个值得研究的问题。根据压缩感知(compressive sensing, CS)的原理,在信号稀疏的假设下,可以用比传统要求少得多的测量次数(即传声器数量)来重构稀疏信号;声场可以由声源振幅的稀疏向量很好地表示,并且应用不规则阵列几何结构,在声源和传声器之间的传递矩阵中提供足够的非冗余,然后可通过求解包含等效源强度向量的 ℓ_1 范数惩罚的目标函数进行求解^[6-7]。

值得注意的是,计算机视觉领域同样存在许多类似的线性逆问题。Beck 等^[8]在 2009 年提出了快速迭代收缩阈值算法(fast iterative shrinkage-thresholding algorithm, FISTA)。随后,Bhotto 等^[9]提出了改进的快速迭代收缩阈值算法(improved fast iterative shrinkage-thresholding algorithm, IFISTA),它是一种基于加权梯度的 FISTA 算法,通过在图像恢复算法中最小化问题的梯度函数中使用正定加权矩阵来预先设定 FISTA。利用该加权矩阵,IFISTA 算法和 FISTA 算法中的利普希茨常数分别保持关键值,并且对于 FISTA 算法保持稳定的所有步长值,同样保证了 IFISTA 的稳定运行。然而,可以通过使用一系列不满足 Nesterov 规则的过松弛参数来进一步加速 IFISTA。因此,Wang 等^[10]提出另一种改进的快速迭代收缩阈值算法(modified improved fast iterative shrinkage-thresholding algorithm, mIFISTA),该算法高效且稳健,而且保持了 IFISTA 和 FISTA 算法的简单性和清晰性,保留了一阶梯度方法的效率优势,更适用于大规模图像去模糊问题。Shen 等^[11]提出了一种基于傅立叶的改进快速迭代收缩阈值算法(FFT-IFISTA),该算法可以选择合适的加权系数来缩小主瓣并实现杂散源衰减。

本文将 FISTA、IFISTA 和 mIFISTA 方法应用于近场声全息领域中,并相应地将其命名为快速迭代收缩阈值等效源法(fast iterative shrinkage threshold equivalent source method, FISTESM)、改进快速迭代收缩阈值等效源法(improved fast iterative shrinkage threshold equivalent source method, IFISTESM)和改良快速迭代收缩阈值等效源法(modified improved fast iterative shrinkage threshold equivalent source method, mIFISTESM)。受高阶波束形成方法以及图像处理算法有效表现的启发,高阶波束形成改良快速迭代收缩阈值等效源法(modified improved fast iterative shrinkage threshold equivalent source method with functional beamforming, mFISTESM-v)是一种将 mFISTESM 和高阶波束形成(functional beamforming, FB)相结合的技术,适用于消除声源识别过程中存在旁瓣“鬼影”的问题。本文的目的是与传统的 Tikhonov 正则化方法相比,衡量算法在 NAH 中的声源重构性能。首

先,我们分别对算法在低频段和高频段的识别误差进行了评定;然后,我们将 mFISTESM 与高阶波束形成方法相结合,以分辨力为评价指标,进一步清晰地识别声源。

1 算 法

1.1 等效源法的基本概念

等效源法(又称波叠加法)的原理是假设平面上的不同的等效声源可以替代随机声源发出的声场,再解决逆问题。图 1 给出了整体试验的架构,包括全息测量面、声压重建平面和等效源面的布局。

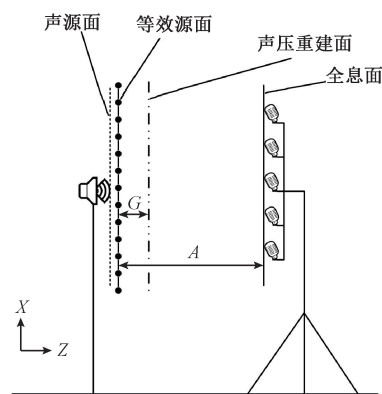


图 1 等效源法示意图

Fig. 1 The sketch of ESM

假设 M 个传声器分散在整个全息测量平面上,而在假设的等效源平面中存在的等效源的数量是 N 。相应地,在第 m 个点的声压矢量 $\mathbf{p}(m)$ 可以用式(1)中的矩阵表示:

$$\mathbf{p}(m) = \sum_{n=1}^N g(\mathbf{r}_m | \mathbf{r}_n) q_n \quad (1)$$

式中: $g(\mathbf{r}_m | \mathbf{r}_n)$ 代表自由场中的格林函数,也是亥姆霍兹方程的解,它在自由场中能起到传递函数的作用,可以建立第 n 个等效源点和第 m 个传声器测量点之间的关系,如式(2)所示。

$$g(\mathbf{r}_m | \mathbf{r}_n) = \frac{e^{-jk \|\mathbf{r}_{mn}\|}}{4\pi \|\mathbf{r}_{mn}\|} \quad (2)$$

式中: $k = 2\pi f/c$ 代表波数(单位为 rad/m); f 代表频率; c 代表声速; $j^2 = -1$; $\|\mathbf{r}_{mn}\|$ 表示声源到传声器之间的距离。以矩阵向量的形式表示为式(3)。

$$\mathbf{p} = \mathbf{A} \mathbf{q} \quad (3)$$

式中: $\mathbf{p}_{M \times 1}$ 是具有元素 $p(m)$ 的复数声压矢量; $\mathbf{A} = j\rho\omega g(\mathbf{r}_m | \mathbf{r}_n)$ 是阵列平面和假设平面之间的维数为 $M \times N$ 的传递矩阵; ω 是角频率(单位为 rad/s); ρ 是介质密度(单位为 kg/m³)。

点声源的数量通常要比传声器的数量多,这使得等

式(3)高度欠定。这种逆问题本质上是病态的,因为它尝试重建远离声源表面的倏逝波的成分。在声学工程实践中,测量过程中存在的噪声会影响测量结果的准确性,因此需要采用正则化方法来克服这一缺陷。通过计算等效声源 $\mathbf{q}$ 的最佳强度,重建声压 $\mathbf{P}_R$ 可由式(4)得出:

$$\mathbf{P}_R = \mathbf{G}\mathbf{q} \quad (4)$$

式中: $\mathbf{G}$ 是另一个连接重建面和等效源面的传递矩阵。

1.2 Tikhonov 正则化

基于最小 ℓ_2 范数准则的 Tikhonov 正则化方法由于其计算过程简单,不需要等效源强度的先验信息,因此被广泛用于处理不适定问题。Leclère^[12] 提出了一个选择正则化系数 λ 的综合标准。给定噪声输入数据 $\mathbf{p}$, 处理带有 ℓ_2 范数解惩罚的不适定逆问题的 Tikhonov 正则化方法可描述为式(5):

$$\underset{\mathbf{q}}{\text{minimize}} \|\mathbf{p} - \mathbf{A}\mathbf{q}\|_2^2 + \lambda^2 \|\mathbf{q}\|_2^2 \quad (5)$$

式中: $\|\cdot\|_2$ 表示 ℓ_2 范数, $\lambda \in [0, +\infty)$ 是正则化系数, λ 的选择对近似解会产生较大影响。当 λ 具有一定值时, Tikhonov 正则化有一个显式解析解,如式(6)所示。

$$\mathbf{q} = [\mathbf{A}^H \mathbf{A} + \lambda^2 \mathbf{I}]^{-1} \mathbf{A}^H \mathbf{p} \quad (6)$$

式中: $\mathbf{I}$ 等于对角矩阵,上标 H 与伴随算子相关。本文以 L 曲线法获得该正则化系数 λ 。

在传统的等效源法中,式(3)的类型尤其可以用 Tikhonov 正则化方法求解。然而,当声源在中频和高频范围内时,识别不准确的现象相对明显。因此,我们将注意力集中在处理另一种旨在更准确地检测声源的方法上。

1.3 快速迭代收缩阈值算法及其改进

迭代收缩阈值算法 (iterative shrinkage-thresholding algorithm, ISTA) 被广泛应用在信号及图像处理中的线性逆问题;但 ISTA 的非渐近全局收敛速度不太令人满意。相反, FISTA 作为一种加速梯度投影算法,因其优异的性能和保持 ISTA 一样的简单结构而被广泛研究。为提高 Tikhonov 正则化方法在中、高频段的分辨力,我们将快速迭代收缩/阈值算法应用到声源识别领域 (记为 FISTESM)。

初始化 $\mathbf{q}_0 = 0$, $\mathbf{y}_1 = \mathbf{q}_0$, $t_1 = 1$, 从第 n 步迭代结果 $\mathbf{q}_n$ 到第 $n+1$ 步迭代结果 $\mathbf{q}_{n+1}$ 的步骤如式(7)~(9):

$$\mathbf{q}_n = \text{prox}_{\frac{1}{L}\lambda} \left(\mathbf{y}_n - \frac{1}{L} \nabla f(\mathbf{y}_n) \right) \quad (7)$$

$$t_{n+1} = \frac{1 + \sqrt{1 + 4(t_n)^2}}{2} \quad (8)$$

$$\mathbf{y}_{n+1} = \mathbf{q}_n + \left(\frac{t_n - 1}{t_{n+1}} \right) (\mathbf{q}_n - \mathbf{q}_{n-1}) \quad (9)$$

式中: ∇f 是如式(10)所示的梯度函数, L 是等于 $\mathbf{A}^T \mathbf{A}$ 最大特征值的 ∇f 的利普希茨常数。

$$\nabla f(\mathbf{y}_n) = 2\mathbf{A}^T (\mathbf{A}\mathbf{y}_n - \mathbf{p}) \quad (10)$$

为了获得比 FISTESM 更快的收敛速度和更好的恢复效果, IFISTESM 的核心思想是在式(7)中增加一个正比重加权矩阵^[9], 即:

$$\mathbf{q}_n = \text{prox}_{\frac{1}{L}\lambda} \left(\mathbf{y}_n - \frac{1}{L} \mathbf{W}_n \nabla f(\mathbf{y}_n) \right) \quad (11)$$

式中: $\mathbf{W}_n$ 是固定且预先安排的正定加权矩阵, $\nabla f(\mathbf{y}_n)$ 在式(10)中给出。为了获得稳定的操作以及更快的收敛, 加权矩阵可以设置为式(12)表达的形式:

$$\mathbf{W}_n = \sum_{j=1}^n \binom{n}{j} (-1)^{j-1} \left(\frac{1}{L} \mathbf{A}^T \mathbf{A} \right)^{j-1} \quad (12)$$

如果单位矩阵等于加权矩阵, IFISTESM 算法可以转换为 FISTESM 算法。因此, IFISTESM 算法可视为 FISTESM 算法的改进。然而, IFISTESM 有一个缺陷, 即过度放松的参数不符合 Nesterov 规则。值得注意的是, IFISTESM 和 FISTESM 算法都采用 Nesterov 规则来提高收敛速度, 这意味着这两种算法中的序列 t 应该满足不等式(13):

$$t_{n+1}^2 - t_{n+1} \leq t_n^2, \quad \forall n \geq 1 \quad (13)$$

当 $t_1 = 1$ 时, mIFISTESM 算法^[10] 流程如下:

初始化 $\mathbf{q}_0 = 0$, $\mathbf{y}_1 = \mathbf{q}_0$, $t_1 = 1$, 从第 n 步迭代结果 $\mathbf{q}_n$ 到第 $n+1$ 步迭代结果 $\mathbf{q}_{n+1}$ 的步骤如式(14)~(16)所示:

$$\mathbf{q}_n = \mathbf{p} + \left(\mathbf{y}_n - \frac{1}{L} \mathbf{W}_n \nabla f(\mathbf{y}_n) \right) \quad (14)$$

$$\alpha_t = \frac{t}{t+a} \quad (15)$$

$$\mathbf{y}_{n+1} = \mathbf{q}_n + \alpha_t (\mathbf{q}_n - \mathbf{q}_{n-1}) \quad (16)$$

根据文献^[10], 研究人员使用 $a \in (0, 3)$ 来保证迭代结果收敛; 而且证明了式(15)中的参数 a 的变化对 mIFISTESM 的性能影响很小。在测试中, 本文设置 $a = 1.5$ 。该算法保留了基于一阶梯度方法的计算效率, 并且与 FISTESM 和 IFISTESM 算法一样简单有效。

1.4 高阶波束形成改良快速迭代收缩阈值等效源法

由于迭代算法能够收敛到最优解, 当迭代步长增加时, mIFISTESM 的分辨力会达到一定的水平。为了实现分辨力增强的目标, 重建面的重建声压 $\mathbf{P}_R$ 可用 FB 形成的系数矩阵来调整输出^[13-15]。本文结合 mIFISTESM 和高阶波束形成的优点, 提出一种高阶波束形成改良快速迭代收缩阈值等效源法 (mIFISTESM- v)。

式(17)定义了交叉互谱矩阵 $\mathbf{q}_n$:

$$\mathbf{C} = \mathbf{q}_n \mathbf{q}_n^H \quad (17)$$

式中: $\mathbf{C}$ 是非负的埃尔米特矩阵。引入了 $1/v$ 阶的高阶矩阵函数后并通过计算 $\mathbf{C}$ 的谱分解并将该函数应用于特征值来将其应用于 $\mathbf{C}$, 因此, 式(17)可以分解为式(18):

$$\mathbf{C} = \mathbf{U} \mathbf{\Sigma} \mathbf{U}^H \quad (18)$$

$$\mathbf{C}^{1/v} = \mathbf{U} \text{diag}(\sigma_1^{1/v}, \dots, \sigma_N^{1/v}) \mathbf{U}^H \quad (19)$$

式中: $\mathbf{U}$ 是维数为 $N \times N$ 的 $\mathbf{C}$ 的西矩阵。 $\mathbf{\Sigma} = \text{diag}(\sigma_1, \dots, \sigma_N)$ 是对角特征值矩阵, 对应着交叉互谱矩阵的相应特征值。

高阶波束形成方法在第 j 个网格的输出为:

$$\mathbf{Z}_j = \left(\frac{\mathbf{G}_j}{\|\mathbf{G}_j\|_2} \mathbf{C}^{1/v} \left(\frac{\mathbf{G}^j}{\|\mathbf{G}^j\|_2} \right)^H \right)^v \quad (20)$$

式中: $\|\cdot\|_2$ 是 ℓ_2 范数, $\mathbf{G}_j$ 是传递矩阵 $\mathbf{G}$ 的第 j 行。矩阵 $\mathbf{Z}$ 中的分量不是重建表面上的声强, 而是与重建声强的交叉谱矩阵相关的参数矩阵。

以单声源为例, 假设等效源点与旁瓣之间的传递向量为 $\mathbf{w}_1$, 等效源点与主瓣之间的传递向量为 $\mathbf{w}_2$, 实际声源与对应等效源点之间的传递向量为 $\mathbf{w}_0$ 。而且假设 $\mathbf{w}_0$ 、 $\mathbf{w}_1$ 、 $\mathbf{w}_2$ 均为单位向量, 则旁瓣与主瓣的相对值可以由式(21)表示:

$$r = \left| \frac{\mathbf{Z}_1}{\mathbf{Z}_2} \right| = \left| \frac{\mathbf{w}_1 \mathbf{w}_0^H \mathbf{w}_0^H \mathbf{w}_1^H}{\mathbf{w}_2 \mathbf{w}_0^H \mathbf{w}_0^H \mathbf{w}_2^H} \right|^v = \left| \frac{\cos(\theta_1)}{\cos(\theta_2)} \right|^{2v} \quad (21)$$

式中: θ_1 是 $\mathbf{w}_1$ 与 $\mathbf{w}_0$ 的夹角; θ_2 是 $\mathbf{w}_2$ 与 $\mathbf{w}_0$ 的夹角。由图 1 可知, 声压重建面平行于等效源面, 所以向量 $\mathbf{w}_2$ 与 $\mathbf{w}_0$ 的方向也平行, 即 $0 < \cos(\theta_1) < \cos(\theta_2) = 1$ 。因此, 随着 v 阶的增加, 旁瓣方向上的幅值会更快衰减; 高阶波束形成的输出将更多地集中在主瓣上。然而, 高阶波束形成的输出只是对应于 mFISTESM 产生的声压矩阵幅度的系数矩阵; mFISTESM 的最终输出是通过用 $\mathbf{Z}$ 修改 mFISTESM- v 输出来生成的。因此, 本文将 v 设置为 1, 以便更为直观看到高阶波束形成方法的效果。

$$\mathbf{p}_f(i) = \mathbf{p}_R(i) \cdot \sqrt{\frac{\text{abs}(\mathbf{Z}(i))}{\max(\text{abs}(\mathbf{Z}))}} \quad (22)$$

式(22)用 $\mathbf{Z}$ 修正了 $\mathbf{P}_R$ 中的每个分量以抑制旁瓣; 由于 mFISTESM- v 的良好重构精度, 振幅得以保持精确。

mFISTESM- v 结合了 mFISTESM 和高阶波束形成的优点, 以产生更好的重构结果的分辨力质量。所提出的方法能让重建声压聚焦于主瓣且抑制旁瓣, 更好缩小识别主瓣面积, 使得识别定位更加精准。本文将主瓣最大值的 -12 dB 区域被视作主瓣区, 用 -12 dB 主瓣面积 B/λ 表示分辨力 (λ 为声波波长)。

2 数值模拟

在数值模拟部分, 分别针对单声源和相干声源两种案例, 我们首先对包括 Tikhonov 正则化方法、FISTESM、IFISTESM、mFISTESM 和 mFISTESM- v 在内的 5 种方法针对声源识别在数值模拟中的声压进行了重建; 再对除 mFISTESM- v 外的 4 种方法进行了误差分析; 最后对前述的 5 种方法进行了分辨力分析。

论文中所有数值模拟结果都是通过 Windows 10

系统的 Intel Core 2 i5-8400 CPU @ 2.80 GHz 和 8 GB RAM 的台式电脑上运行的 MATLAB R2017a 软件来获取。在模拟中, 具有扬声器的声源平面、具有传声器的全息平面以及假想的等效源平面如图 1 所示排列。球体半径为 0.01 m, 振动速度为 2.5×10^{-2} m/s, 声速约为 340 m/s。在所有模拟中, 全息平面和声源平面之间的距离为 0.3 m, 声源平面和重建平面之间的距离为 0.02 m。等效源平面距声源平面 0.001 m。所提到的 4 个平面, 包括声源平面、全息平面、重建平面和等效源平面, 同时处于相同坐标系中。

重建声源平面和等效声源平面均采用 81×81 网格, 间距为 0.01 m。传声器阵列有 18 个分布在一个圆内的传声器, 图 2 标出了传声器以及声源的位置布置。重建误差 ε 可由式(23)评估算法性能。

$$\varepsilon = \sqrt{\sum_{i=1}^N (|p_{li}| - |p_{2i}|)^2} / \sqrt{\sum_{i=1}^N |p_{2i}|^2} \times 100\% \quad (23)$$

式中: p_{li} 和 p_{2i} 分别是重建平面上的重建声压值和理论声压值。

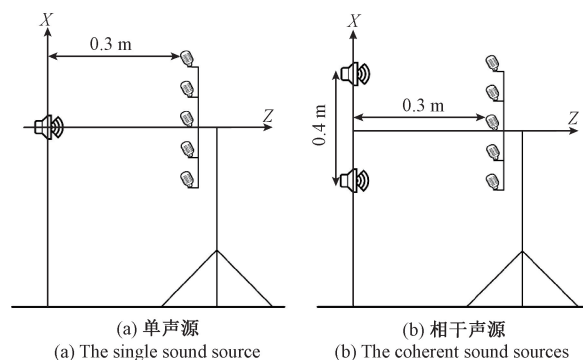


图 2 声源位置摆放示意图

Fig. 2 The position diagram of sound source

2.1 单声源模拟

单声源模拟的目标是比较针对单声源的 FISTESM、IFISTESM 和 mFISTESM 的重建性能。声源位于 (0, 0, 0) m。随机噪声以低于传声器平均声压 15 dB 的水平添加到传声器声压数据中, 同时动态显示范围设置为 3 dB。图 3 是在重建表面上由理论计算、FISTESM、IFISTESM、mFISTESM 和 mFISTESM- v 计算所得的声压级。当声源频率设置为 500 Hz 时, FISTESM、IFISTESM 和 mFISTESM 无法识别与理论声源图相对应的结果, 而 Tikhonov 正则化方法在给定位置形成了较大的热区。mFISTESM- v 则能有效缩小旁瓣, 达到接近理论声压级的效果。相比之下, 当声源频率设置为 2 500 Hz 时, FISTESM、IFISTESM 和 mFISTESM 在识别声源时可以保

持接近理论值的高分辨力。上述4种算法都随着频率的增加产生了更好的分辨力,也就是说随着频率的增加,识别效果逐渐增强。

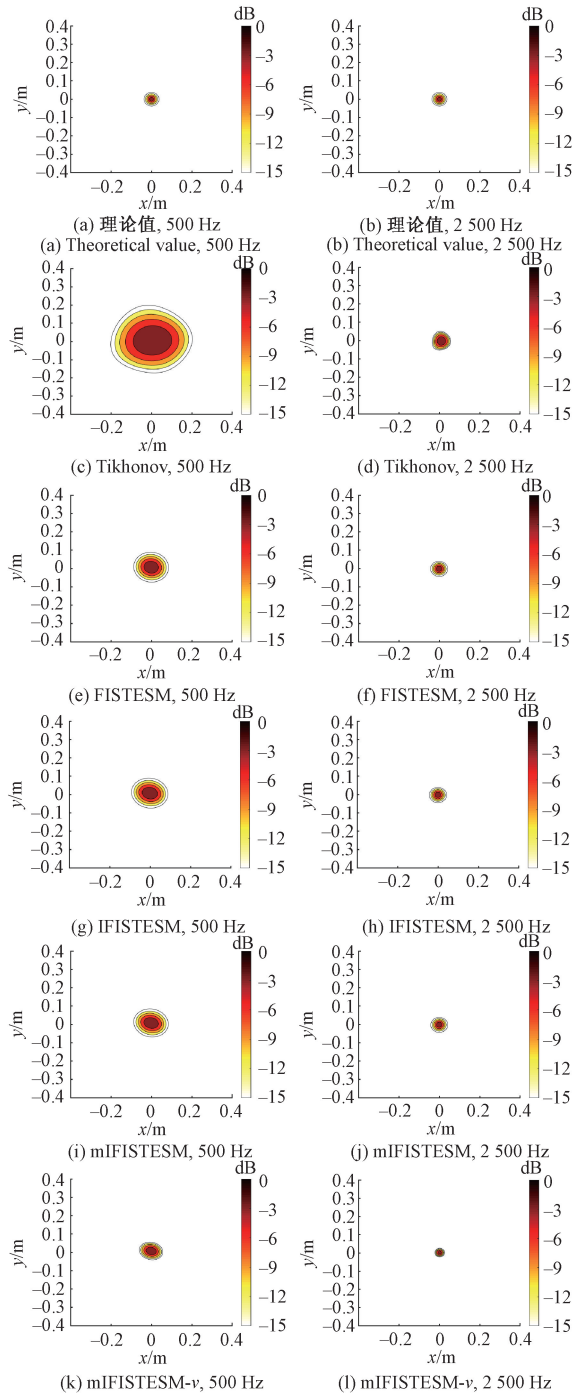


图3 500和2 500 Hz的单声源计算声压级图

Fig. 3 The calculated sound power level maps for the single source at 500 and 2 500 Hz

图4为沿重建面中间行的单声源重建声压幅值。可以看到,当声源频率设置为500 Hz时,Tikhonov正则化

在声源位置产生一个平坦的峰值,当声源频率设置为2 500 Hz时,其峰值也远离理论值,由FISTESM、IFISTESM和mFISTESM计算的重建声压的幅度接近理论值。此外,当声源的频率固定在500 Hz时,这些方法的效果并不明显。值得注意的是,在500和2 500 Hz,mFISTESM-v均能抑制旁瓣减少“鬼影”以此突出主瓣的识别效果。

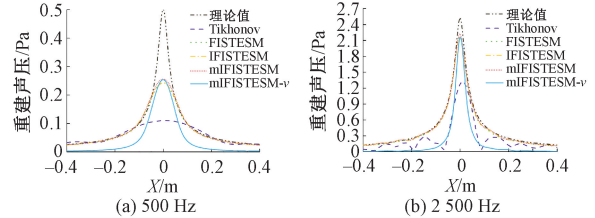


图4 沿重建面中间行的单声源重建声压幅值

Fig. 4 The single sound source reconstruction sound pressure along the middle row of reconstruction surface

根据式(22),使用30组数值模拟的平均值在500~2 500 Hz的频率范围内单声源重建误差如图5所示。从整体上看,FISTESM、IFISTESM、mFISTESM能够比Tikhonov正则化方法更稳定地重建声场。Tikhonov正则化方法的重构误差在超过1 500 Hz后急剧上升,对于高频段声源识别难以胜任。相反,其他3种算法FISTESM、IFISTESM和mFISTESM在500~2 500 Hz频段都具有相对稳定的性能,重构误差范围基本在30%以下。

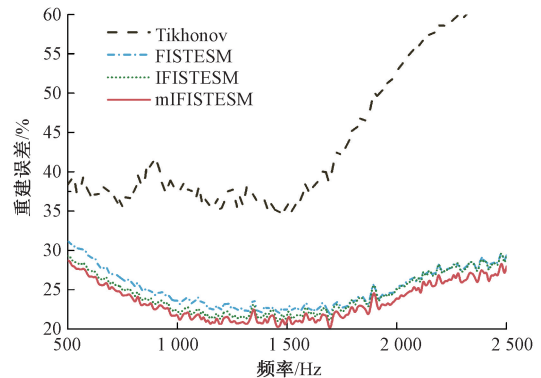


图5 Tikhonov、FISTESM、IFISTESM和mFISTESM之间的单声源重建误差

Fig. 5 The single sound source reconstruction error among Tikhonov, FISTESM, IFISTESM, and mFISTESM

如图6所示, B/λ 为-12 dB主瓣面积,以1 000 Hz为间隔,从500 Hz到2 500 Hz进行单声源分辨力分析,声源强度设为1。主瓣区越小表示分辨力越好。可以看出,除Tikhonov正则化方法所得到的主瓣区较大外,FISTESM、IFISTESM、mFISTESM和mFISTESM-v 4种方

法所得主瓣区均在 0.002 以下, mIFISTESM- v 所得主瓣区最小, 与单声源模拟结果较为接近。

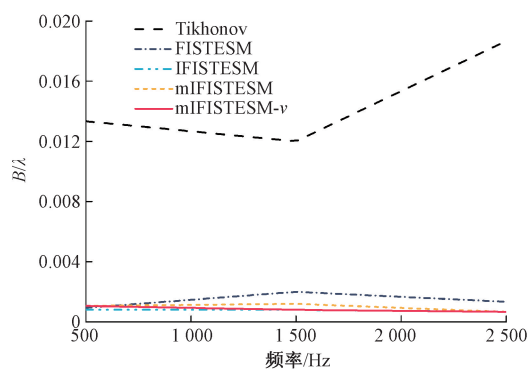


图6 单声源分辨力

Fig. 6 The resolution quality of the single sound source

2.2 相干声源模拟

本文也评估了上述 5 种声源识别方法对于相干声源的识别性能。相干声源模拟设置与单声源模拟设置相同, 区别在于, 如图 2(b) 所示坐标系中, 相干声源之间有 0.4 m 的分离距离。500 和 2 500 Hz 相干声源的声源识别结果呈现在图 7 中。当声源频率设置为 500 Hz 时, 相干声源没有被分辨, 两个真实声源被 Tikhonov 正则化方法中的大热区域覆盖。相比之下, FISTESM、IFISTESM、FISTESM 可以较为清楚地识别出两个声源, 此外, 它们形成的热区也逐渐变小。值得注意的是, mIFISTESM- v 能明显缩小旁瓣, 实现精确的定位效果。当相干声源的频率设置为 2 500 Hz 时, 这 5 种算法都有显著的改进, 可以清楚地发现给定声源的位置。FISTESM、IFISTESM、FISTESM 的识别效果非常接近。与 Tikhonov 正则化方法相比, mIFISTESM- v 获得的声压级图在热区大小上与理论声压级图吻合较好。

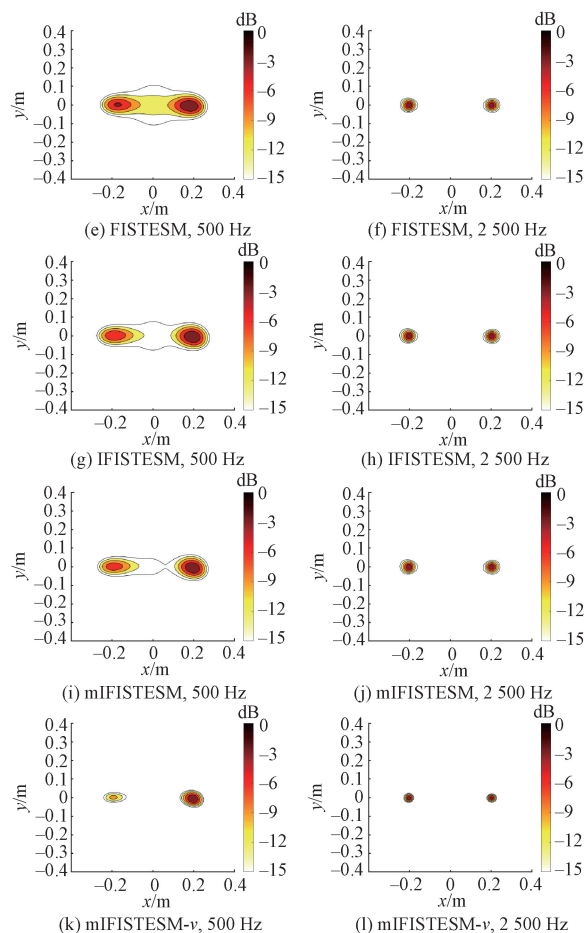
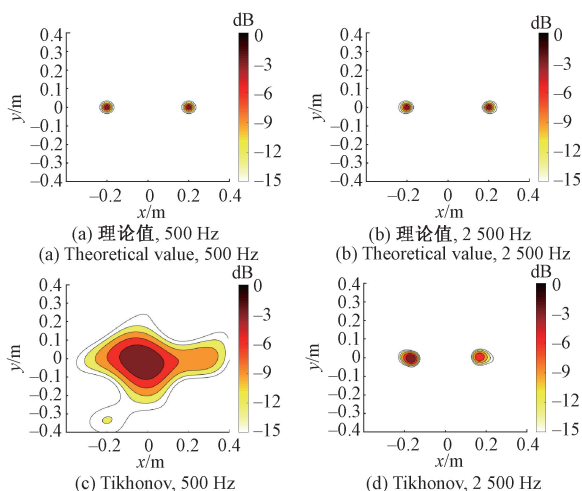


图7 500 和 2 500 Hz 相干声源数值模拟计算声压级图

Fig. 7 The calculated sound power level maps for the coherent sources at 500 and 2 500 Hz

图 8 为沿重建面中间行的相干声源重建声压幅值。当声源频率设置为 500 Hz 时, Tikhonov 正则化在两个真实声源位置之间只产生一个较小的平峰值, 而 FISTESM、IFISTESM 和 mIFISTESM 可以粗略识别低频双声源位置。当声源频率设置为 2 500 Hz 时, 用 FISTESM、IFISTESM 和 mIFISTESM 获得的声压幅值几乎与理论值一致, 除开两个峰值未能完全吻合。而 Tikhonov 正则化方法只能近似识别高频范围内的声源位置。相比之下, mIFISTESM- v 能明显压制旁瓣, 突出主瓣。

使用 30 组数值模拟的平均值在 500~2 500 Hz 的频率范围内相干声源重建误差如图 9 所示。可以看出, FISTESM、IFISTESM 和 mIFISTESM 都在 42% 的误差范围内, 并且比 Tikhonov 正则化方法都可以更稳定地重建声场。值得注意的是, Tikhonov 正则化方法在 1 700 Hz 以上的频段表现不稳定, 不能以明显的方式分辨误差, 误差超过 50%。

图 10 进行了相干声源分辨力分析, 声源强度均设为 1。可以看出, 除 Tikhonov 正则化方法所得到的主瓣区较大

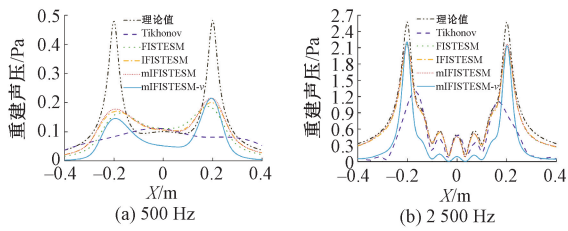


图8 沿重建表面中间行的相干声源重建声压幅值

Fig. 8 The coherent sound sources reconstruction sound pressure along the middle row of reconstruction surface

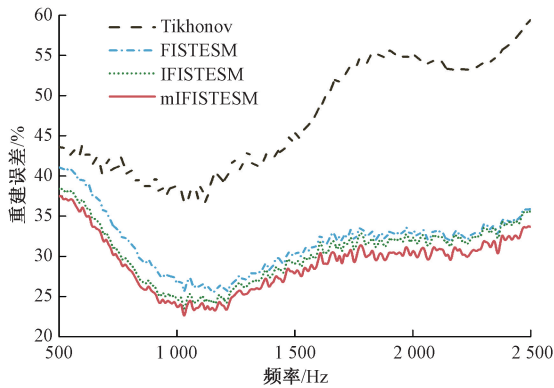


图9 相干声源重建误差

Fig. 9 The coherent sound sources reconstruction error

外, FISTESM、IFISTESM、mIFISTESM 和 mIFISTESM-v 4 种方法所得主瓣区均在 0.007 以下, mIFISTESM-v 所得主瓣区最小, 与相干声源模拟结果较为接近。

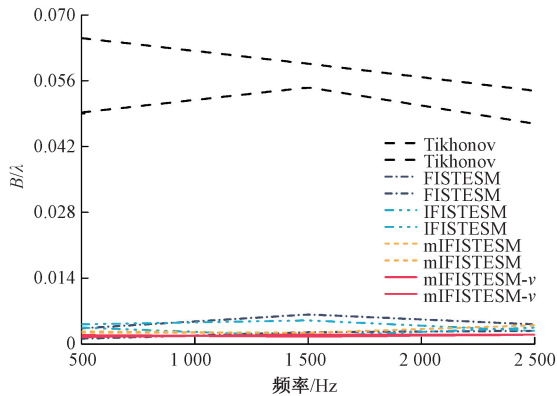


图10 相干声源分辨力

Fig. 10 The resolution quality of the coherent sound sources

3 验证实验

本文以实验验证了上述 5 种算法的效果。实验测量中采用了 Brüel & Kjær's 公司生产的 WA-1558-W-021 扇形阵列。扬声器被放置在目标声源位置, 由 18 通道组合传

声器阵列进行测量。传声器集成在一个直径为 0.38 m 的阵列中, Brüel & Kjær 4951 型传声器的坐标分布在图 11 中。

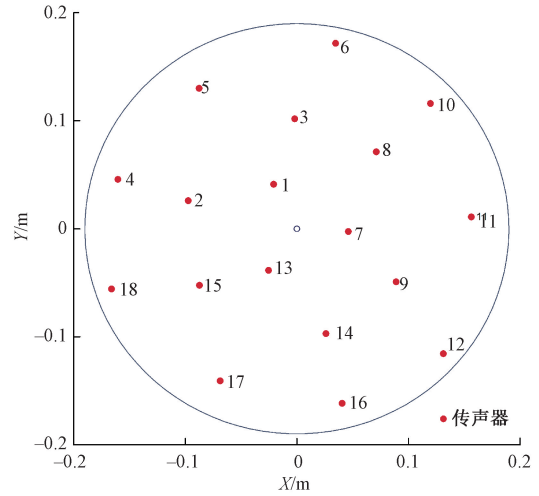


图11 传声器分布的坐标

Fig. 11 The coordinates of the microphones distribution

实验设置如图 12 所示。将单个声源设置在坐标原点, 声源表面距离传声器阵列 0.3 m, 全息平面的中心和声源的中心在同一轴线上。在实验过程中, 扬声器分别发出 500 和 2 500 Hz 的单频声。采样频率为 16 384 Hz, 采样时间为 5 s, 并加入 15 dB 的随机噪声, 以确保在测量过程中可以在相对敏感的条件下进行模拟。全息平面采集声压数据后, 利用 MATLAB 软件实现声学成像过程。以声压级作为输出参数, 并且显示范围设置为 5 dB, 梯度为 1 dB。

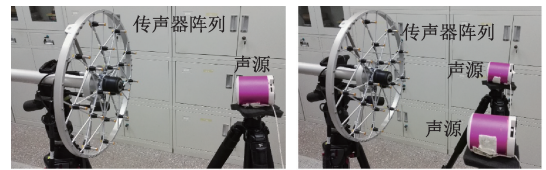
(a) 单声源实验布置 (b) 相干声源实验布置
(a) The single sound source case (b) The coherent sound sources case

图12 实验设置

Fig. 12 The experimental establishment

图 13 展示了单声源实验重建声压级图。当声源频率设置为 500 Hz 时, Tikhonov 正则化方法在中心周围单独形成一个较大的热区, 而 FISTESM、IFISTESM、mIFISTESM 在坐标原点处形成类似的热点。mIFISTESM-v 抑制旁瓣, 实现声源准确识别。当声源频率设置为 2 500 Hz 时, 情况会变得更好。这 5 种算法的分辨力都有明显的提升, 并都能较为准确地准确识别声源位置, mIFISTESM-v 效果较佳。

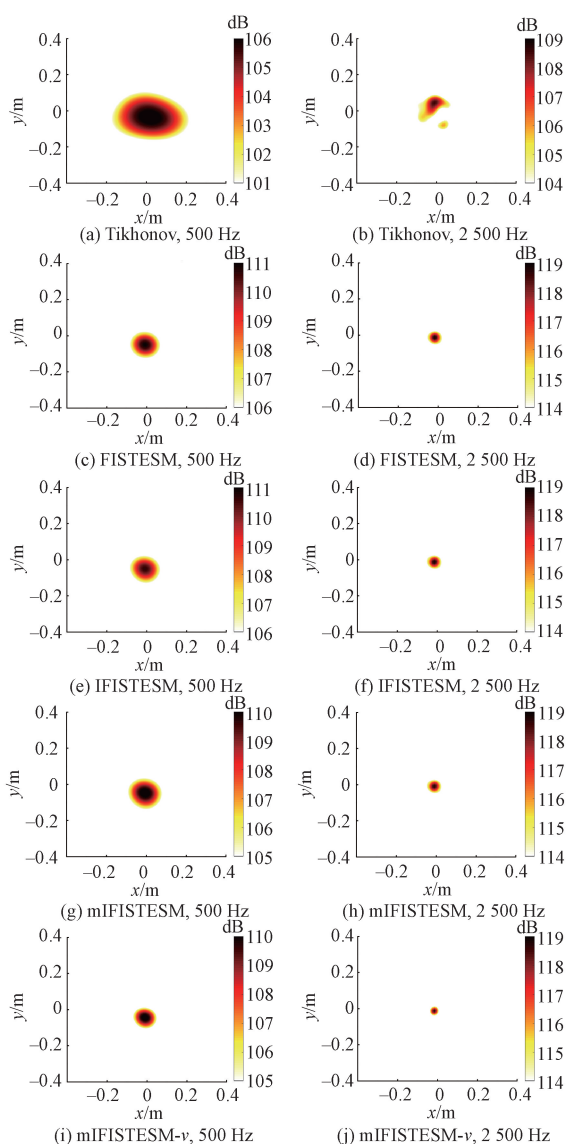


图 13 500 和 2 500 Hz 单声源重建声压级图

Fig. 13 The reconstructed sound power level maps for the single sound source at 500 and 2 500 Hz

图 14 展示了相干声源实验重建声压级图。当声源频率设置为 500 Hz 时,上述 5 种方法在识别相干声源时有不同的数值误差,总体而言,它们得到了错误的定位结果,与图 7 中的模拟结果有一定出入。然而, mFISTESM-v 仍能在一定程度上缩小旁瓣。相比之下,当声源的频率由于波长的减小而增加到 2 500 Hz 时,除了 Tikhonov 正则化方法之外的 4 种方法变得更精确,与图 7 中的模拟结果一致。所提方法仍能准确定位声源,聚焦主瓣抑制旁瓣。

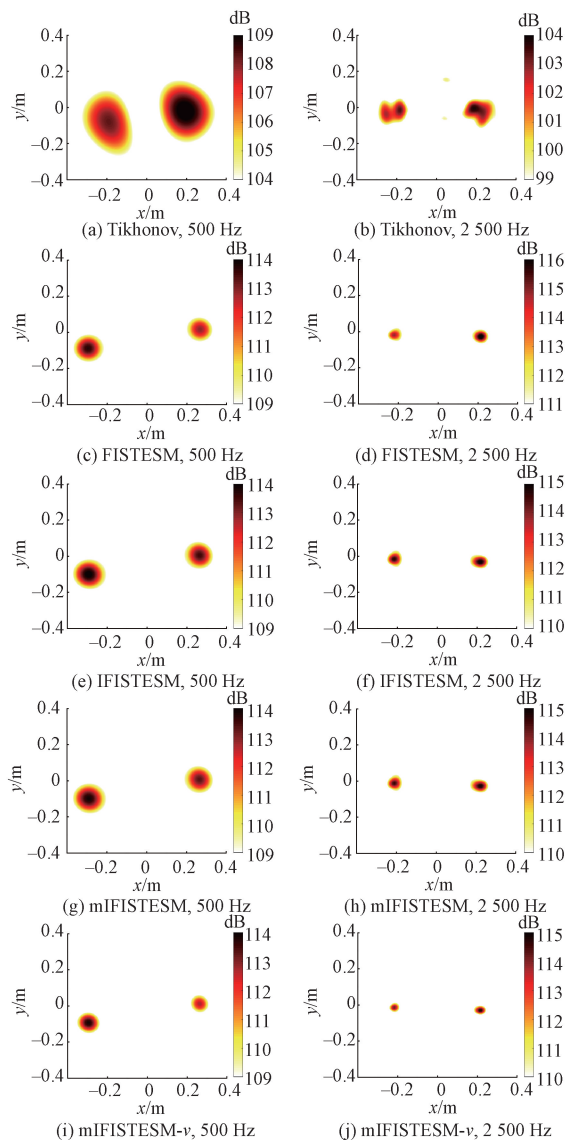


图 14 500 和 2 500 Hz 相干声源的重建声压级图

Fig. 14 The reconstructed sound power level maps for the coherent sound sources at 500 and 2 500 Hz

4 结 论

本文提出了一种高阶波束形成改良快速迭代收缩阈值声源识别算法 mFISTESM-v,能够获得更好的声源识别效果。首先,相比快速迭代收缩阈值等效源法 FISTESM,改进快速迭代收缩阈值等效源法 IFISTESM,改良快速迭代收缩阈值等效源法 mFISTESM 具有更高的计算效率和更好的收敛性;然后, mFISTESM-v 结合了 mFISTESM 和高阶波束形成的优点,通过选择合适的权重系数,可以实现聚焦主瓣、抑制旁瓣“鬼影”的效果。

参考文献

- [1] WILLIAMS E G, MAYNARD J D, SKUDRZYK E. Sound source reconstructions using a microphone array[J]. *Journal of the Acoustical Society of America*, 1980, 68(1): 340-344.
- [2] YU L, ANTONI J, ZHAO H, et al. The acoustic inverse problem in the framework of alternating direction method of multipliers [J]. *Mechanical Systems and Signal Processing*, 2021, 149: 107220.
- [3] YANG Y X, CHU ZH G, YANG Y, et al. Two-dimensional Newtonized orthogonal matching pursuit compressive beamforming [J]. *The Journal of the Acoustical Society of America*, 2020, 148 (3): 1337-1348.
- [4] MERINO-MARTÍNEZ 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.
- [5] 宋玉来, 胡红生. 非规则结构表面多声源近场声全息方法[J]. *仪器仪表学报*, 2017, 38(5): 1126-1134.
- SONG Y L, HU H SH. Near-field acoustic holography for multiple sound sources on irregular structure surface[J]. *Chinese Journal of Scientific Instrument*, 2017, 38(5): 1126-1134.
- [6] 黎术, 徐中明, 贺岩松, 等. 基于弹性网正则化的广义逆波束形成[J]. *仪器仪表学报*, 2015, 36(5): 1170-1176.
- LI SH, XU ZH M, HE Y S, et al. Generalized inverse beamforming via elastic net regularization[J]. *Chinese Journal of Scientific Instrument*, 2015, 36 (5): 1170-1176.
- [7] 徐中明, 黎术, 贺岩松, 等. 光滑 L0 范数广义逆波束形成[J]. *仪器仪表学报*, 2014, 35(6): 1276-1281.
- XU ZH M, LI SH, HE Y S, et al. Smoothed L0 norm generalized inverse beamforming[J]. *Chinese Journal of Scientific Instrument*, 2014. 35(6): 1276-1281.
- [8] BECK A, TEOULLE M. IEEE: A fast iterative shrinkage-thresholding algorithm with application to wavelet-based image deblurring [J]. 2009 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2009: 693.
- [9] BHOTTO M Z A, AHMAD M O, SWAMY M N S. An improved fast iterative shrinkage thresholding algorithm for image deblurring [J]. *SIAM Journal on Imaging Sciences*, 2015, 8(3): 1640-1657.
- [10] WANG ZH, WANG J J, WANG W D, et al. A novel thresholding algorithm for image deblurring beyond nesterov' s rule [J]. *IEEE Access*, 2018, 6: 58119-58131.
- [11] SHEN L B, CHU ZH G, ZHANG Y X, et al. A novel fourier-based deconvolution algorithm with improved efficiency and convergence[J]. *Journal of Low Frequency Noise Vibration and Active Control*, 2020, 39 (4): 866-878.
- [12] LECLERE Q. Acoustic imaging using under-determined inverse approaches: Frequency limitations and optimal regularization [J]. *Journal of Sound and Vibration*, 2009, 321(3-5): 605-619.
- [13] DOUGHERTY R. A new derivation of the adaptive beamforming formula [C]. 7th Berlin Beamforming Conference, 2018: 1-10.
- [14] XU ZH M, WANG Q H, HE Y S, et al. An improved algorithm for noise source localization based on equivalent source method[J]. *Shock and Vibration*, 2016(1): 1-13.
- [15] DOUGHERTY R. Functional beamforming [C]. Berlin Beamforming Conference, 2014: 1-25.

作者简介



黄琳森, 2015年毕业于内蒙古工业大学获得学士学位, 2017年毕业于台湾元智大学获得硕士学位, 2018年毕业于贵州大学获得博士学位, 现为重庆大学博士研究生, 主要研究方向为近场声全息和声学信号处理。
E-mail: 20183201008@cqu.edu.cn

Huang Linsen received his B. Sc. degree from Inner Mongolia University of Technology in 2015, received his M. Sc. degree from Taiwan YuanZe University in 2017, and received his M. Sc. degree from Guizhou University in 2018. He is currently pursuing his Ph. D. degree at Chongqing University. His main research interests include near-field acoustical holography and acoustical identification.



徐中明 (通信作者), 1983年毕业于重庆大学获得学士学位, 1992年毕业于清华大学获得硕士学位, 2002年毕业于日本东京大学获得博士学位, 现为重庆大学教授、博士生导师, 主要研究方向为振动噪声控制。
E-mail: xuzm@cqu.edu.cn

Xu Zhongming (Corresponding author) received his B. Sc. degree from Chongqing University in 1983, received his M. Sc. degree from Tsinghua University in 1992, and received his Ph. D. degree from Tokyo University in 2002. He is currently a professor and a Ph. D. advisor at Chongqing University. His main research interests include noise and vibration control.