



## 逆模型下的高效声源定位算法研究

赵达达, 齐小刚, 冯海林

引用本文:

赵达达,齐小刚,冯海林. 逆模型下的高效声源定位算法研究[J]. 智能系统学报, 2023, 18(1): 117–123.

ZHAO Dada, QI Xiaogang, FENG Hailin. Efficient acoustic source localization algorithm based on an inverse model[J]. *CAAI Transactions on Intelligent Systems*, 2023, 18(1): 117–123.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202106010>

## 您可能感兴趣的其他文章

### 一种深度自监督聚类集成算法

A deep self-supervised clustering ensemble algorithm

智能系统学报. 2020, 15(6): 1113–1120 <https://dx.doi.org/10.11992/tis.202006050>

### 面对类别不平衡的增量在线序列极限学习机

Incremental online sequential extreme learning machine for imbalanced data

智能系统学报. 2020, 15(3): 520–527 <https://dx.doi.org/10.11992/tis.201904040>

### 结合谱聚类的标记分布学习

Label distribution learning based on spectral clustering

智能系统学报. 2019, 14(5): 966–973 <https://dx.doi.org/10.11992/tis.201809019>

### 结合稀疏表示与约束传递的半监督谱聚类算法

A semi-supervised spectral clustering algorithm combined with sparse representation and constraint propagation

智能系统学报. 2018, 13(5): 855–862 <https://dx.doi.org/10.11992/tis.201703013>

### 基于近似消息传递算法的压缩感知雷达成像方法

Compressed sensing radar imaging using approximate message passing

智能系统学报. 2015(4): 592–598 <https://dx.doi.org/10.3969/j.issn.1673-4785.201411025>

DOI: 10.11992/tis.202106010

网络出版地址: <https://kns.cnki.net/kcms/detail/23.1538.TP.20220506.1002.002.html>

# 逆模型下的高效声源定位算法研究

赵达达, 齐小刚, 冯海林

(西安电子科技大学 数学与统计学院, 陕西 西安 710071)

**摘要:** 与波束形成算法相比, 广义互相关逆模型宽带声源定位算法提供了空间的高分辨率, 但需要更高的计算量。为了提升广义互相关逆模型算法的计算效率, 并尽可能保留其分辨率优势, 本文提出了一种高效的声源定位算法。该算法首先去除麦克风阵列输出功率较小的网格点来压缩计算网格, 其次, 在几何因素条件下利用基于密度的聚类进一步压缩网格, 最后仅用保留下的点进行计算, 从而大大降低了计算规模。真实数据实验证明, 本文所提算法能有效提升广义互相关逆模型算法的计算效率。

**关键词:** 传统波束形成; 时域波束形成; 广义互相关; 压缩计算网格; 逆模型; 声源定位; 麦克风阵列; 基于密度的聚类

**中图分类号:** TP391.9; TN912.3; TN641 **文献标志码:** A **文章编号:** 1673-4785(2023)01-0117-07

**中文引用格式:** 赵达达, 齐小刚, 冯海林. 逆模型下的高效声源定位算法研究 [J]. 智能系统学报, 2023, 18(1): 117-123.

**英文引用格式:** ZHAO Dada, QI Xiaogang, FENG Hailin. Efficient acoustic source localization algorithm based on an inverse model[J]. CAAI transactions on intelligent systems, 2023, 18(1): 117-123.

## Efficient acoustic source localization algorithm based on an inverse model

ZHAO Dada, QI Xiaogang, FENG Hailin

(School of Mathematics and Statistics, Xidian University, Xi'an 710071, China)

**Abstract:** Compared with the beamforming algorithm, the wideband acoustic source localization based on the generalized crosscorrelation inverse model provides a higher spatial resolution and requires more computation. This paper proposes an acceleration algorithm to improve the computational efficiency of acoustic source localization and retain the existing resolution. By removing the grid points with the low output power of the microphone array, the computational grid is compressed. The geometric factor is also considered, and the density-based clustering algorithm is used to further compress the grid. Finally, only the reserved points are used for the calculation to reduce the calculation scale. Experiments with real data show that the proposed algorithm can improve computational efficiency significantly.

**Keywords:** conventional beamforming; time domain beamforming; generalized cross correlation; compression computational grid; inverse model; acoustic source localization; microphone array; density-based clustering

麦克风阵列是以特定方式排列, 能够准确获取空间信息的一组声音采集系统。麦克风阵列声源定位 (acoustic source localization, ASL) 就是在阵列几何形状已知条件下, 利用麦克风阵列输出信号中包含的空域-时域-频域信息, 对声源进行正

确的定位。近十几年来, 基于麦克风阵列的 ASL 技术得到了迅速发展, 并广泛应用于视频会议系统<sup>[1]</sup>、机器人听觉、语音增强<sup>[2]</sup>、语音识别<sup>[3]</sup>等领域。

与 ASL 算法相关联的麦克风阵列, 目的是补偿麦克风与虚拟扫描点之间的时间或相位延迟, 可以在时域或频域进行。频域技术使用麦克风信号的交叉频谱矩阵 (cross spectral matrix, CSM), 最常见的技术是波束形成。传统波束形成 (conventional beamforming, CBF) 具有鲁棒性且较为简单, 然而由于旁瓣效应<sup>[4]</sup>的存在, 使得其在低频率下

收稿日期: 2021-06-06. 网络出版日期: 2022-05-06.

**基金项目:** 国家自然科学基金项目 (61877067); 近地面探测技术重点实验室开放基金项目 (TCGZ2019A002, TCGZ2021C003); 重点实验室稳定支持项目 (6142414200511); 陕西省自然科学基金基础研究计划项目 (2021JZ-19).

**通信作者:** 齐小刚. E-mail: [xgqi@xidian.edu.cn](mailto:xgqi@xidian.edu.cn).

的空间分辨率很差。为了提高空间分辨率,其他波束形成算法也被开发出来。Li 等<sup>[5]</sup>使用矩阵方法保持恒定的波束宽度和低峰值旁瓣电平,同时保持自适应零干扰能力。Yan 等<sup>[6]</sup>提出了最小方差无失真响应 (minimum variance distortionless response, MVDR) 在球谐波域的波束形成,在低峰值旁瓣和窄主瓣宽度之间提供了一个合适的折衷。Kumar 等<sup>[7]</sup>利用球谐域波束形成解决了球形麦克风阵列的近场声源定位问题。Salvati 等<sup>[8]</sup>提出了一种基于对角卸载 (diagonal unloading, DU) 变换的波束形成方法,实现了低计算复杂度的鲁棒定位。Simard 等<sup>[9]</sup>研究了传声器阵列中的源识别问题,证明了压缩感知波束形成方法也适用于近场聚焦球波模型,提高了源定位分辨率。Long 等<sup>[10]</sup>从几何投影的角度阐述了声源定位问题,实现了在较短声信号帧内的声源定位,证明了其算法与传统算法之间的等价性。然而,这些算法对阵列误差敏感,且在低信噪比下受限。因此,研究一种鲁棒的高分辨率波束形成算法是非常必要的。

近年来,航空声学中许多反卷积方法被用于提高声源定位分辨率<sup>[11]</sup>。这是因为 CBF 输出可以表示为声源分布与波束形成器对点源的响应的卷积。通过对 CBF 波束形成器的波束图进行反卷积,可以得到清晰波束图下的实际声源分布,从而提高分辨率。Dougherty 等<sup>[12]</sup>首次将 CLEAN 算法应用于声波阵列测量中。Brooks 等<sup>[13]</sup>提出反卷积声源成像 (deconvolution approach for the mapping of acoustic sources, DAMAS) 算法用于声源定位。Sun 等<sup>[14]</sup>提出了一种 Ex-RL (extended richardson lucy) 算法,提升了在二维移变反卷积声学图像测量问题中的分辨率。Bai 等<sup>[15]</sup>提出了非凸  $L1/2$  正则化方法,提高了反卷积问题的空间分辨率和鲁棒性。DAMAS 使研究者充分认识到反卷积的优点和可行性,但不幸的是, DAMAS 需要较高的计算量,消耗较长运行时间和较大计算机内存。Ma 等<sup>[16]</sup>提出了一种基于波束形成算法的压缩计算网格 DAMAS 算法,通过去除非正波束形成点,减小了计算规模,提高了计算效率。

在时域中最常用的方法是基于传声器对之间时延的广义互相关方法<sup>[17]</sup>。这个时间延迟可以用来为扫描区域上可能的源位置生成一个双曲线。所有双曲线 (所有麦克风对) 的交集确定出声源位置。Noël 等<sup>[18]</sup>使用了与反问题相关的广义互相关 (generalized cross correlation, GCC) 来定位声源位置。反问题求解使理论互相关函数和测量互相关函数之间的差异最小化。他们最终得到了一

个声源图,其中包含了从各个方向接收到的与麦克风阵列相关的角能量流。具有良好的应用前景。Padois 等<sup>[19]</sup>提出了一种快速的声源定位技术,能够检测主要声源和反射信号,他们提出了一个基于 GCC 的最小化问题,但与 Noël 等<sup>[18]</sup>的理论公式和求解方法不同,利用了两种不同的稀疏表示来解决该最小化问题。Chu 等<sup>[20]</sup>提出了一种基于聚焦点与麦克风对之间时延估计差异的累积概率的方法,解决了传播模型矩阵最优阈值随聚焦点距离、阵列几何形状和阵列大小而变化的问题。

然而,传播模型矩阵计算以及求解反卷积所带来的高计算量问题始终没有解决。Ma 等<sup>[16]</sup>提出的压缩计算网格 DAMAS 算法为我们提供了优化思路,本文对 Ma Wei 等的压缩网格算法进行了改进,提出了一种基于密度聚类的压缩网格算法,利用几何信息对网格点进行聚类,去除噪声点,将该算法运用到 Padois 等<sup>[19]</sup>提出的声源定位技术中,最终取得了比原算法更高的计算效率。

## 1 声源定位技术

### 1.1 麦克风阵列信号

考虑如下声场,点声源位于坐标  $\mathbf{r}_s$  处,其产生的信号  $s(\mathbf{r}_s, t)$  被设置在  $\mathbf{r}_m (m = 1, 2, \dots, M)$  处的  $M$  个麦克风所记录。设在自由场条件下被第  $m$  个麦克风记录的声压信号  $p_m$  为

$$p_m(t) = \alpha_m(\mathbf{r}_s) s(\mathbf{r}_s, t - \Delta t_{ms}) + n_m(t)$$

式中:  $\alpha_m(\mathbf{r}_s)$  是传播过程中的衰减系数,  $n_m(t)$  是由背景或传感器噪声产生的不相关噪声。声源  $s$  和麦克风  $m$  之间的传播时间  $\Delta t_{ms}$  为

$$\Delta t_{ms} = \frac{1}{c} \|\mathbf{r}_m - \mathbf{r}_s\|_2$$

式中:  $c$  为声速,  $\|\cdot\|_p$  表示矩阵或向量的  $p$  范数。麦克风阵列信号  $y(\mathbf{r}_s, t)$  为麦克风信号的算数平均:

$$y(\mathbf{r}_s, t) = \frac{1}{N} \sum_{i=1}^N p_i(t)$$

### 1.2 时域波束形成

声源定位的麦克风阵列信号的输出功率  $y_e(\mathbf{r}_s)$  定义为

$$y_e(\mathbf{r}_s) = E\{y(\mathbf{r}_s, t)^2\} = \int_{-\infty}^{+\infty} \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N p_i(t) p_j(t) dt$$

其中  $E\{\cdot\}$  为数学期望。麦克风阵列信号的输出功率也可以写为

$$y_e(\mathbf{r}_s) = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N (p_i \cdot p_j)(\tau)$$

式中:  $(p_i \cdot p_j)$  表示对应的两个麦克风信号在时延  $\tau = (\Delta t_{js} - \Delta t_{is})$  时的互相关函数, 即

$$(p_i \cdot p_j)(\tau) = \int_{-\infty}^{+\infty} p_i(t) p_j(t + \tau) dt$$

由于在上式中的自相关项( $i = j$ )不包含时延估计的信息, 因此, 不予考虑。此外, 根据互相关函数的对称性, 可以将多余的麦克风对去除。因此, 如果在一组 $\mathbf{r}_l (l = 1, 2, \dots, L)$ 扫描点上搜索源位置, 则修改后的麦克风阵列信号的输出功率 $y'_e(\mathbf{r}_s)$ 定义为

$$y'_e(\mathbf{r}_s) = \frac{1}{N_p} \sum_{i=1}^N \sum_{j>i}^N (p_i \cdot p_j)(\Delta t_{ji} - \Delta t_{il})$$

式中 $N_p$ 为麦克风对的数量。

### 1.3 广义互相关 GCC

为计算 GCC 函数, 需要利用麦克风信号在角频率 $\omega$ 处的互谱的傅里叶逆变换, 计算公式为

$$(p_i \circ p_j)(\tau) = \int_{-\infty}^{+\infty} W_{ij}(\omega) C_{ij}(\omega) \exp(j\omega\tau) d\omega$$

式中:  $W_{ij}(\omega)$ 为权值函数,  $C_{ij}(\omega)$ 为

$$C_{ij}(\omega) = \left( \int_{-\infty}^{+\infty} p_i(t) \exp(-j\omega t) dt \right) \left( \int_{-\infty}^{+\infty} p_j(t) \exp(-j\omega t) dt \right)^*$$

因此, 麦克风阵列信号的修正功率可以表示为

$$y'_e(\mathbf{r}_s) = \frac{1}{N_p} \sum_{i=1}^N \sum_{j>i}^N \int_{-\infty}^{+\infty} W_{ij}(\omega) C_{ij}(\omega) \exp(j\omega\tau) d\omega$$

### 1.4 逆模型

对于包含源向量 $\mathbf{x}$  (对应源信号的功率) 的寻优问题 $J$ , 其最小化依赖于测量噪声源图 $\mathbf{y}'$  (CBF 所得结果) 和模型源映射 $\bar{\mathbf{y}}$ 的损失函数 $\rho$ , 是减小旁瓣和杂散瓣影响从而提高声源定位精度的一种方法。从而, 寻优问题定义为

$$J(\mathbf{x}) = \min_{\mathbf{x}} \rho(\mathbf{y}', \bar{\mathbf{y}})$$

定义源映射模型为线性系统, 即

$$\bar{\mathbf{y}} = \mathbf{A}\mathbf{x} \quad (1)$$

式中 $\mathbf{A}$ 为传播模型矩阵<sup>[21]</sup>。为计算传播模型矩阵, 考虑 $\mathbf{r}_k$ 位置处的声源, 时延估计 $\Delta t_{ijk}$ 为

$$\Delta t_{ijk} = \frac{1}{c_0} \|\mathbf{r}_i - \mathbf{r}_k\|_2 - \frac{1}{c_0} \|\mathbf{r}_j - \mathbf{r}_k\|_2 = \Delta t_{ik} - \Delta t_{jk}$$

矩阵 $\mathbf{A}$ 中元素定义为

$$a(\mathbf{r}_k, \mathbf{r}_l) = \frac{1}{N_p} \sum_{i=1}^N \sum_{j>i}^N \begin{cases} 1, & |\Delta t_{ijk} - \Delta t_{ijl}| \leq \varepsilon \\ 0, & \text{其他} \end{cases}$$

如果时延估计的差值很小, 则设置为 1, 否则设置为 0, 这意味着只需要考虑靠近源位置的扫描点的贡献<sup>[21]</sup>。令代价函数 $\rho$ 表示欧氏距离, 从而线性最小二乘问题可以定义为

$$J(\mathbf{x}) = \min_{\mathbf{x}} \|\mathbf{y}' - \mathbf{A}\mathbf{x}\|_2^2$$

## 2 基于密度聚类的压缩网格算法

广义互相关逆模型算法虽能有效提升分辨率, 但其计算复杂度较高。当网格点数目比较大

时, 无论是传播模型矩阵的计算, 还是该逆模型的求解都会耗费相当长的时间。因此, 本节提出了一种能有效提升计算效率的算法。图 1 为本文所提算法的算法流程图。

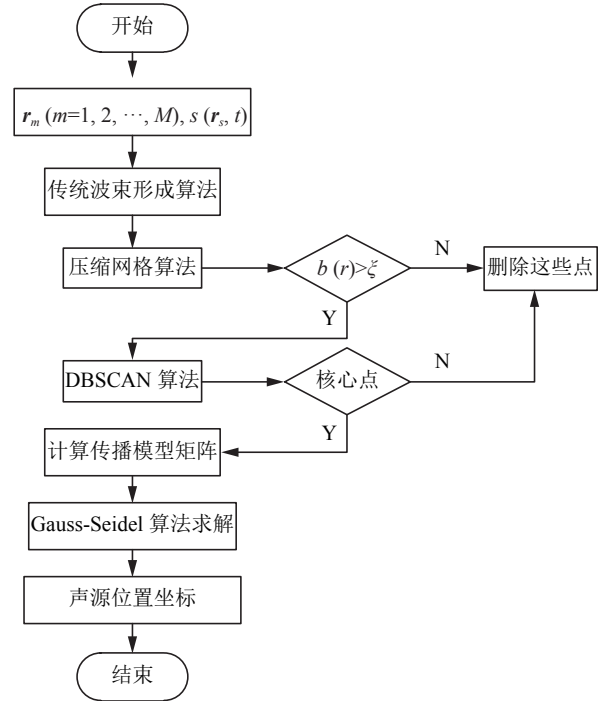


图 1 算法流程

Fig. 1 Algorithm flow

### 2.1 传统波束形成

CBF 公式为

$$b(\mathbf{r}) = \frac{1}{\|\mathbf{v}(\mathbf{r})\|^4} \mathbf{v}(\mathbf{r})^H \mathbf{E} \{ \mathbf{p} \mathbf{p}^H \} \mathbf{v}(\mathbf{r})$$

式中:  $\mathbf{E} \{ \mathbf{p} \mathbf{p}^H \}$  表示互谱矩阵,  $\mathbf{p}$  表示声压函数,  $\mathbf{v}(\mathbf{r}) = [v_1(\mathbf{r}) v_2(\mathbf{r}) \dots v_M(\mathbf{r})]$  为聚焦点 $\mathbf{r}$ 所对应的导向向量, 其中元素表达式为

$$v_m(\mathbf{r}) = \frac{\|\mathbf{r} - \mathbf{r}_m\|}{\|\mathbf{r}\|} \exp\{-jk\|\mathbf{r} - \mathbf{r}_m\|\}$$

式中:  $\|\mathbf{r}\|$  为聚焦点到麦克风阵列中心的距离,  $\|\mathbf{r} - \mathbf{r}_m\|$  为聚焦点到第 $m$ 个麦克风的距离,  $k$  为波数。

### 2.2 压缩网格算法

为了消除麦克风自噪声污染, 特别是由与麦克风相互作用的湍流引起的自噪声污染, 对角卸载 (DR) 通常应用于 CSM, 通过去除 (消零) CSM 的对角线项, 并解释分母中 CMS 项数的变化。

$$b(\mathbf{r}) = \frac{\mathbf{v}(\mathbf{r})^H \mathbf{E} \{ \mathbf{p} \mathbf{p}^H \}_{\text{diag}=0} \mathbf{v}(\mathbf{r})}{\|\mathbf{v}(\mathbf{r})\|^4 - \sum_{m=1}^M |v_m(\mathbf{r})|^4}$$

当 DR 应用于 CSM 时, CBF 会有负值出现。这些波束形成非正的网格点被认为是冗余点, 从而可以将其从扫描区域的原始计算网格中剔除<sup>[16]</sup>。在压缩网格算法中, 去除这些输出功率低于阈值的点, 仅保留满足条件的点, 即



$$b(\mathbf{r}) > \xi$$

则方程 (1) 转化为

$$\tilde{\mathbf{y}} = \tilde{\mathbf{A}}\tilde{\mathbf{x}}$$

式中:  $\tilde{\mathbf{A}} \in \mathbf{R}^{\tilde{R} \times \tilde{R}}$ ,  $\tilde{\mathbf{x}} \in \mathbf{R}^{\tilde{R} \times 1}$ ,  $\tilde{\mathbf{y}} \in \mathbf{R}^{\tilde{R} \times 1}$ ,  $\tilde{R}$  为保留的网格点的数目。

### 2.3 基于密度的聚类

通过上述步骤后保留下来的点都聚集在声源附近, 若声场中声源个数为  $n$ , 则应该有  $n$  块较为密集的点集以及一些零散点, 而这些零散点正好是我们所不需要的。所以, 运用基于密度的聚类, 将聚类结果的核心点保留, 去除远离声源的点。

一种比较有代表性的基于密度的聚类算法是 DBSCAN(density-based spatial clustering of applications with noise)<sup>[22]</sup>。它以元素之间的紧密程度作为判别条件, 以此来决定它们是否属于一类, 最终将所有紧密相连的元素划分为同一类。

DBSCAN 算法认为“最大相连”的点集可以构成一个簇, 此处的“最大相连”的意思是: 构成簇的点集必然是互相紧密相连的, 与簇中的点密度可达的点必然属于这个簇。DBSCAN 算法流程如下:

1) 筛选核心点: 对所有点进行邻域计算, 按照如上规则划分核心点、边界点和噪声点。

2) 删除噪声点: 将所有噪声点剔除, 只保留核心点和边界点;

3) 生成簇: 扫描整个数据集, 找到任意一个核心点, 对该核心点进行扩充。扩充的方法是寻找从该核心点出发的所有密度相连的数据点。遍历该核心点的邻域内的所有核心点 (因为边界点是无法扩充的), 寻找与这些数据点密度相连的点, 直到没有可以扩充的数据点为止, 这些点就形成了一个簇;

4) 循环执行步骤 3), 直到原始点集为空, 至此就划分出了所有簇。

### 2.4 含有正性约束的 Gauss-Seidel 迭代法

通过上述步骤, 则我们问题转化为求解

$$\hat{\mathbf{y}} = \hat{\mathbf{A}}\hat{\mathbf{x}}$$

式中:  $\hat{\mathbf{A}} \in \mathbf{R}^{\hat{R} \times \hat{R}}$ ,  $\hat{\mathbf{x}} \in \mathbf{R}^{\hat{R} \times 1}$ ,  $\hat{\mathbf{y}} \in \mathbf{R}^{\hat{R} \times 1}$ ,  $\hat{R}$  为经过两步去除后剩下的网格点数。我们用带有正性约束的 Gauss-Seidel 迭代法来求解  $\hat{\mathbf{x}}$ :

$$r_n^{(i)} = \sum_{n'=1}^{n-1} A_{nn'} x_{n'}^{(i+1)} + \sum_{n'=n+1}^L A_{nn'} x_{n'}^{(i)} - y_n$$

$$x_n^{(i+1)} = \max\left(x_n^{(i)} - \frac{r_n^{(i)}}{A_{nn}}, 0\right), n = 1, 2, \dots, L, \quad i = 1, 2, \dots, K$$

式中:  $L$  表示格点数,  $K$  表示迭代次数。  $r_n^{(i)}$  是每步迭代的残差, 第一次迭代时初始值  $r_n^{(1)}$  可以取做 0 或  $B_n$ 。

## 3 实验

本文所用数据来自多通道脉冲响应数据库<sup>[23]</sup>。该数据库数据测试场景设置如下, 传感器阵列为 8 通道等间隔线性阵列, 阵列间隔为 0.08 m; 声源到阵列中心 (0 m, 0 m) 的连线与阵列正方向成  $15^\circ \sim 345^\circ$ , 分辨率为  $15^\circ$ , 且声源到阵列中心 (0 m, 0 m) 的距离为 1 m 和 2 m, 混响为 160 ms, 360 ms, 610 ms。实验中, 将麦克风坐标按实测场景设置为 (-0.28 m, 0 m), (-0.20 m, 0 m), (-0.12 m, 0 m), (-0.04 m, 0 m), (0.04 m, 0 m), (0.12 m, 0 m), (0.20 m, 0 m), (0.28 m, 0 m), 从数据库中选取混响为 160 ms, 声源到阵列中心距离为 1 m 的数据进行实验。

图 2 为单声源在  $45^\circ$  时传统波束形成法的定位结果, 其主瓣较宽, 旁瓣较高, 空间分辨率低。图 3 为对使用压缩网格算法后的剩余点进行 DBSCAN 算法聚类的结果, 其中, 棕色区域为使用压缩网格算法所剔除的点, 深蓝色区域为使用 DBSCAN 算法计算出的噪声点, 浅蓝色区域为计算出的边缘点, 黄色区域为核心点区域, 即基于密度聚类的压缩网格算法所确定出的最终计算范围。

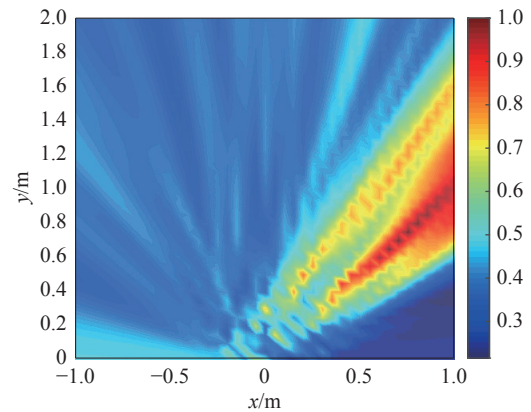


图 2 CBF 结果

Fig. 2 Result of CBF

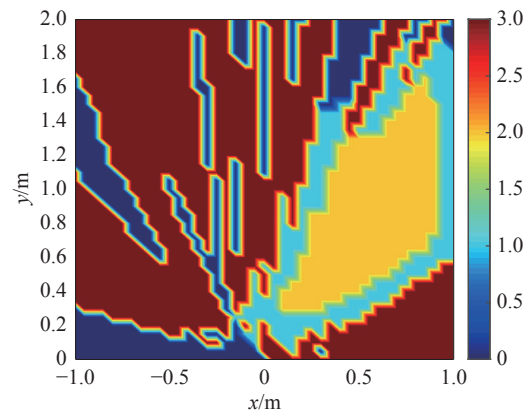


图 3 聚类结果

Fig. 3 Result of clustering

图 4 为单声源在  $30^\circ$ , 每个网格点大小为  $0.05\text{ m} \times 0.05\text{ m}$ , 总网格点数为 1 681 时的定位结果, 与图 2 的 CBF 定位结果相比, 图 4 中结果主瓣较窄, 有较高的空间分辨率。图 4(a) 为原算法<sup>[19]</sup>定位结果, 图 4(b) 为基于聚类的压缩网格算法定位结果, 图中 2 种算法的均方根误差 (RMSE) 相差不超过  $0.05\text{ m}$ , 定位效果基本一致。图 5 为双声源定位结果, 声源分别处于  $75^\circ$ 、 $90^\circ$  角, 每个网格点大小为  $0.04\text{ m} \times 0.04\text{ m}$ , 总网格点数为 2 601, 图中两算法 RMSE 总偏差不超过  $0.08\text{ m}$ 。基于聚类的压缩网格算法基本保留了原算法的高空间分辨率优势。

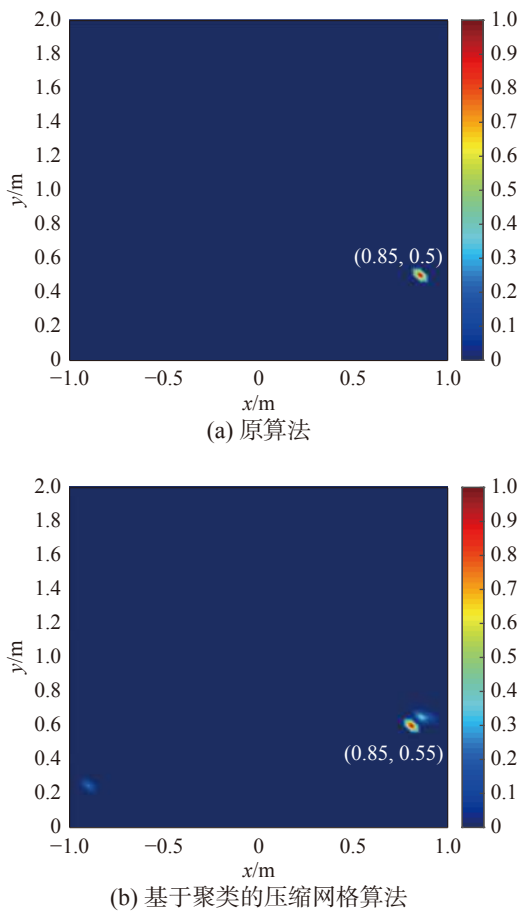


图 4 单声源定位结果

Fig. 4 Localization result of one source

表 1 为 3 种算法计算规模与计算时间的对比, 结果为 10 组数据结果的算术平均。基于密度聚类的压缩网格算法将计算规模降至最低, 矩阵  $A$  计算时间以及声源位置求解时间均为 3 种算法中最短。时间由 MATLAB R2018b 的 tic-toc 函数给出, 实验设备为  $2.40\text{ GHz}$  的四核处理器,  $16.0\text{ GB}$  内存的计算机。表 2 为 3 种算法在不同信噪比 (SNR) 下定位精度的对比, 评价参数选用均方根误差 (RMSE), 单位为  $\text{m}$ , 分辨率为  $0.05\text{ m}$ 。基于密度聚类的压缩网格算法与原算法以及压缩网格算法

相比, 在 3 种信噪比级别下 RMSE 相差均不超过  $0.05\text{ m}$ 。图 6 和图 7 分别给出了不同网格点个数下, 3 种算法求解传播模型矩阵与声源位置求解的时间对比图, 由于基于密度聚类的压缩网格算法很大程度的降低了计算规模, 所以计算时间最短。

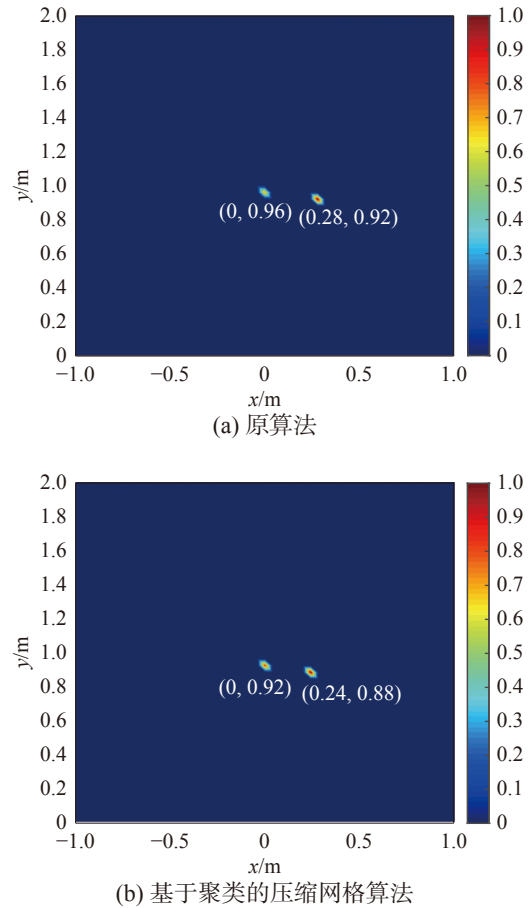


图 5 双声源定位结果

Fig. 5 Localization result of two sources

表 1 计算效率对比

Table 1 Comparison of computational efficiency

算法	原算法 <sup>[19]</sup>	压缩网格 算法 <sup>[16]</sup>	基于密度聚类 压缩网格算法
网格点的数量	1681	782	249
矩阵 $A$ 计算时间/s	45.52	15.01	1.60
声源 $x$ 求解时间/s	0.51	0.21	0.09

表 2 定位精度对比

Table 2 Comparison of localization accuracy

信噪比 SNR/dB	原算法 <sup>[19]</sup> RMSE/m	压缩网格算法 <sup>[16]</sup> RMSE/m	基于密度聚类压缩 网格算法RMSE/m
10	0.0977	0.1043	0.1156
20	0.0605	0.0732	0.0828
30	0.0160	0.0340	0.0525

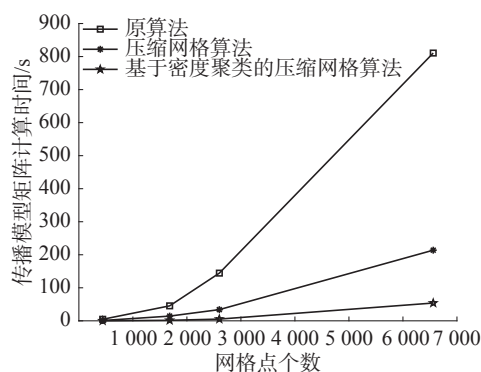


图 6 传播模型矩阵计算时间

Fig. 6 Computation time of the propagation model matrix

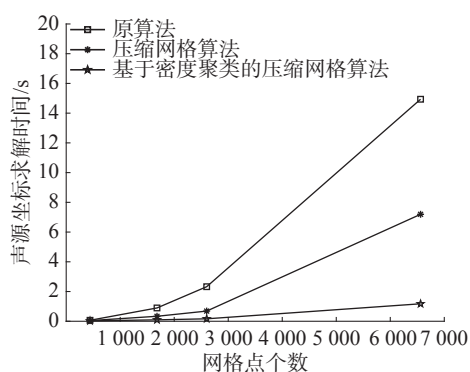


图 7 声源位置求解时间

Fig. 7 Computation time of the source location

基于密度聚类的压缩网格算法有效提升了定位效率,且在定位误差相差不大的前提下,保留了原始逆模型算法的高空间分辨率优势。图 3 中黄色区域是所提算法对网格点进行压缩后的最终计算区域,极大的缩小了计算规模,这也是所提算法能有效提升效率的关键所在。表 1 中数据以及图 6 图 7 证明,无论是传播模型矩阵的计算时间,还是声源坐标的求解时间,本文所提的新算法均优于所对比的两类算法。表 2 中数据表明,在所设条件下,本文所提算法在 3 种信噪比下与对比算法的 RMSE 相差不超过 0.05m,定位精度的小幅降低在所提升的计算效率前可以接受。图 2 与图 4 及图 5 的对比非常明显,传统波束形成算法主瓣宽,旁瓣高,分辨率较低,而逆模型算法能有效提升空间分辨率,且通过对比图 4(a) 与 (b) 以及图 5(a) 与 (b) 发现,本文所提算法很好地保留了逆模型算法的分辨率优势。

## 4 结束语

为了快速定位声源位置,提出了一种基于密度聚类的压缩网格算法。首先,移除传统波束形成结果中不满足阈值条件的点,然后将几何因素考虑在内,利用基于密度的聚类,删除噪声点,保

留核心点,以此来减小问题规模。通过实验与原算法进行对比,结果表明本文所提的算法可以有效的提升计算效率。

## 参考文献:

- [1] 李晓飞, 刘宏. 机器人听觉声源定位研究综述 [J]. 智能系统学报, 2012, 7(1): 9–20.  
LI Xiaofei, LIU Hong. A survey of sound source localization for robot audition[J]. CAAI transactions on intelligent systems, 2012, 7(1): 9–20.
- [2] 鲍长春, 项扬. 基于深度神经网络的单通道语音增强方法回顾 [J]. 信号处理, 2019, 35(12): 1931–1941.  
BAO Changchun, XIANG Yang. Review of monaural speech enhancement based on deep neural networks[J]. Journal of signal processing, 2019, 35(12): 1931–1941.
- [3] 高庆吉, 赵志华, 徐达, 等. 语音情感识别研究综述 [J]. 智能系统学报, 2020, 15(1): 1–13.  
GAO Qingji, ZHAO Zhihua, XU Da, et al. Review on speech emotion recognition research[J]. CAAI transactions on intelligent systems, 2020, 15(1): 1–13.
- [4] MICHEL U. History of acoustic beamforming[C]//1. st Berlin beamforming conference. Berlin: DLR, 2006, 1–17.
- [5] LI Jian, XIE Yao, STOICA P, et al. Beampattern synthesis via a matrix approach for signal power estimation[J]. IEEE transactions on signal processing, 2007, 55(12): 5643–5657.
- [6] YAN Shefeng, SUN Haohai, SVENSSON U P, et al. Optimal modal beamforming for spherical microphone arrays[J]. IEEE transactions on audio, speech, and language processing, 2011, 19(2): 361–371.
- [7] KUMAR L, HEGDE R M. Near-field acoustic source localization and beamforming in spherical harmonics domain[J]. IEEE transactions on signal processing, 2016, 64(13): 3351–3361.
- [8] SALVATI D, DRIOLI C, FORESTI G L. A low-complexity robust beamforming using diagonal unloading for acoustic source localization[J]. IEEE/ACM transactions on audio, speech, and language processing, 2018, 26(3): 609–622.
- [9] SIMARD P, ANTONI J. Acoustic source identification: experimenting the  $\ell_1$  minimization approach[J]. Applied acoustics, 2013, 74(7): 974–986.
- [10] LONG Tao, CHEN Jingdong, HUANG Gongping, et al. Acoustic source localization based on geometric projection in reverberant and noisy environments[J]. IEEE journal of selected topics in signal processing, 2019, 13(1): 143–155.
- [11] PADOIS T, SGARD F, DOUTRES O, et al. Acoustic

- source localization using a polyhedral microphone array and an improved generalized cross-correlation technique[J]. *Journal of sound and vibration*, 2017, 386: 82–99.
- [12] DOUGHERTY R, STOKER R. Sidelobe suppression for phased array aeroacoustic measurements[C]//4th AIAA/CEAS Aeroacoustics Conference. Toulouse: AIAA, 1998: 2242.
- [13] BROOKS T F, HUMPHREYS W M. A deconvolution approach for the mapping of acoustic sources (DAMAS) determined from phased microphone arrays[J]. *Journal of sound and vibration*, 2006, 294(4/5): 856–879.
- [14] SUN Dajun, MA Chao, MEI Jidan, et al. Improving the resolution of underwater acoustic image measurement by deconvolution[J]. *Applied acoustics*, 2020, 165: 107292.
- [15] BAI Baohong, LI Xiaodong. Acoustic sources mapping based on the non-negative L1/2 regularization[J]. *Applied acoustics*, 2020, 169: 107456.
- [16] MA Wei, LIU Xun. Compression computational grid based on functional beamforming for acoustic source localization[J]. *Applied acoustics*, 2018, 134: 75–87.
- [17] KNAPP C, CARTER G. The generalized correlation method for estimation of time delay[J]. *IEEE transactions on acoustics, speech, and signal processing*, 1976, 24(4): 320–327.
- [18] NOËL C, PLANEAU V, HABAUT D. A new temporal method for the identification of source directions in a reverberant hall[J]. *Journal of sound and vibration*, 2006, 296(3): 518–538.
- [19] PADOIS T, DOUTRES O, SGARD F, et al. Time domain localization technique with sparsity constraint for imaging acoustic sources[J]. *Mechanical systems and signal processing*, 2017, 94: 85–93.
- [20] CHU Zhigang, WENG Jing, YANG Yang. Determination of propagation model matrix in generalized cross-correlation based inverse model for broadband acoustic source localization[J]. *The journal of the acoustical society of America*, 2020, 147(4): 2098.
- [21] VELASCO J, PIZARRO D, MACIAS-GUARASA J. Source localization with acoustic sensor arrays using generative model based fitting with sparse constraints[J]. *Sensors*, 2012, 12(10): 13781–13812.
- [22] ESTER M, KRIEGEL H P, SANDER J, et al. A density-based algorithm for discovering clusters in large spatial databases with noise[C]// KDD'96: Proceedings of the Second International Conference on Knowledge Discovery and Data Mining. Portland: AAAI, 1996, 34: 226–231.
- [23] HADAD E, HEESE F, VARY P, et al. Multichannel audio database in various acoustic environments[C]//2014 14th International Workshop on Acoustic Signal Enhancement. Juan-les-Pins: IEEE, 2014: 313–317

### 作者简介:



赵达达, 硕士研究生, 主要研究方向为信号处理与声源定位。



齐小刚, 教授, 博士生导师, 主要研究方向为系统建模和仿真、资源管理和调度、性能评估和优化算法设计网络故障诊断。



冯海林, 教授, 主要研究方向为系统可靠性建模和生存数据分析。