DOI: 10.3901/JME.2020.22.046

二维动态网格压缩波束形成声源识别方法^{*}

樊小鹏¹ 余立超^{2,3} 褚志刚^{2,3} 杨洋^{2,3} 李丽¹

(1. 广东电网有限责任公司电力科学研究院 广州 510082;

2. 重庆大学机械传动国家重点实验室 重庆 400044;

3. 重庆大学汽车工程学院 重庆 400044)

摘要: 基于平面传声器阵列测量和固定网格的传统二维压缩波束形成通过建立阵列传声器测量声压信号和假想声源网格点未知源强之间的欠定线性方程组, 基于主声源通常为稀疏分布这一事实, 利用稀疏促进算法求解上述方程组从而获得声源波达方向和源强的准确估计, 进而准确识别声源。然而, 当声源波达方向与网格点不一致、即发生基不匹配时, 性能会劣化。为克服该问题, 提出二维动态网格压缩波束形成声源识别方法。定义网格坐标和源强分布矢量为变量并采用对数求和罚函数构建目标函数以促进解的稀疏性; 基于优化最小化框架在目标函数的基础上构造合适的替代函数以降低优化复杂度; 通过梯度下降法对替代函数进行迭代优化求解, 从而使网格坐标和源强分布矢量逐渐收敛至真实值附近。仿真和试验均表明, 相较于传统固定网格的二维压缩波束形成, 该方法能够克服基不匹配问题、获得更高的定位精度和量化精度; 该方法能够适用于传声器随机分布的平面阵列, 且无需先验的信噪比(噪声干扰)及声源稀疏度等信息, 即使在传声器数量较少的情况下也能得到低污染高分辨率的声源成像, 保证了高精度的二维波达方向估计和源强量化。

关键词: 压缩波束形成; 动态网格; 平面传声器阵列; 声源识别; 波达方向

中图分类号: TB52

Two-dimensional Dynamic Grid Compressive Beamforming for Acoustic Source Identification

FAN Xiaopeng¹ YU Lichao^{2,3} CHU Zhigang^{2,3} YANG Yang^{2,3} LI Li¹

(1. Electric Power Research Institute of Guangdong Power Grid Limited Liability Corporation, Guangzhou 510082;

2. State Key Laboratory of Mechanical Transmissions, Chongqing University, Chongqing 400044;

3. College of Automotive Engineering, Chongqing University, Chongqing 400044)

Abstract: Conventional two-dimensional compressive beamforming, which is based on planar microphone array measurements and fixed discrete grids, establishes an underdetermined linear system of equations between the sound pressure signals measured by microphones and the unknown source strengths of the assuming acoustic sources corresponding to grids. Based on the fact that the main sources are generally sparse, the above equations can be solved by the sparsity promotion algorithms, then accurate estimation of the source direction-of-arrivals(DOAs) and quantification of the source strengths are achieved, thus identifying sources accurately. However, its performance deteriorates when sources do not coincide with the grids, namely the basis mismatch occurs. To solve this issue, a two-dimensional dynamic grid compressive beamforming for acoustic source identification is developed. First, the grid coordinates and the source strength distribution vector are defined as variables, and the objective function is constructed using a log-sum penalty function to promote the sparsity of the solution. Subsequently, a suitable surrogate function is formulated based on the objective function to reduce the optimization complexity under the majorization-minimization framework. Finally, a gradient descent method is used to iteratively optimize the surrogate function, leading to a gradual process to refine the grid coordinates and

^{*} 国家自然科学基金(11874096,11704040)和中国南方电网科技(GDKJXM20180152)资助项目。20200222收到初稿, 20200908收到修改稿

the source strength distribution vector. The results of numerical simulations and experiments demonstrate that the proposed technology can circumvent the basis mismatch and thus achieve higher location and quantification accuracy, comparing to the conventional two-dimensional fix-grid approach. It can be applied to a planar array with microphones randomly distributing and does not require prior knowledge of signal-to-noise ratio (noise interference) or source sparsity. The dynamic grid compressive beamforming can provide high-resolution and low contamination imaging, allowing accurate estimation of two-dimensional DOAs and quantification of source strengths, even with a small number of microphones.

Key words: compressive beamforming; dynamic grid; planar microphone array; sound source identification; direction-of-arrival

0 前言

作为一种空间滤波技术,波束形成通过对阵列传声器测量声信号进行处理实现声源波达方向(Direction-of-arrival, DOA)估计和源强量化^[1-3]。压缩波束形成^[4-7]是近年来提出并得到迅速发展的一种新型波束形成方法。其基于压缩感知理论和主要声源通常为稀疏分布的事实,通过在声源源强分布和阵列传声器测量声压信号之间建立具有稀疏约束的欠定线性方程组并加以求解获得源强分布,进而实现主声源的准确识别。即便是在传声器数量很少、测量干扰较大的情况下^[8-9]也能获得高分辨率的声源识别结果。

依据是否对目标声源区域划分网格,压缩波束形成分为离散(有网格)和连续(无网格)两类。离散压缩波束形成将目标声源区域离散成有限数目的网格点,每个网格点代表一个观测方向,并假设声源 DOA 落在这些观测方向上,阵列传声器测量声信号和声源源强分布关联数学模型的建立和求解均在离散域实施,其模型简单、求解方便,但存在基不匹配缺陷,即声源识别结果仅在声源 DOA 与观测方向一致时才精准,在声源 DOA 与观测方向不一致时便失真,后者则是更一般的情况^[10-11]。细化网格点虽然能一定程度上缓解基不匹配缺陷,但以加重计算量为代价,且过密网格点会增大感知矩阵的列相关性反而可能劣化声源识别性能^[12-13]。为克服该问题,基于传声器均匀分布的线性阵列 XENAKI 等^[14]于 2015 年提出了无网格的一维连续压缩波束形成,适用于声源位于阵列前方且与传声器共面的情形。YANG 等^[15-16]在此基础上进一步发展出无网格的二维连续压缩波束形成,实现了阵列前方更广阔区域内声源 DOA 的精准估计。但已有的二维连续压缩波束形成仅适用于传声器均匀分布的矩形阵列(包括从中随机抽选传声器形成的阵列),对其他形式的平面阵列并不适用且需要准确估计噪声水平,应用受限。2016 年, FANG 等^[17]提出了一种用于一维线谱估计的同时实现参数学习和稀疏信号恢

复的迭代重加权算法。其在优化最小化框架^[18-20]下通过迭代缩减替代函数来优化目标函数,使字典参数和稀疏信号在交替的迭代过程中不断被更新,最终收敛于真实值。该方法不仅克服了基不匹配缺陷,并且相较于一维无网格线谱估计方法具有更高的重构精度和更加宽容的应用条件。本文受文献[17]启发,提出了基于平面传声器阵列的二维动态网格压缩波束形成声源识别方法,即通过将网格坐标设定为优化参数并使其在迭代求解过程中逐渐收敛于真实声源 DOA 附近。该方法首先建立阵列传声器测量声信号和被施加稀疏约束的源强分布之间的关联数学模型;采用对数求和罚函数^[21-22]构造可同时优化网格坐标和源强分布的目标函数,并基于优化最小化框架在目标函数的基础上构造合适的替代函数以降低优化复杂度;最后,通过梯度下降法对替代函数进行迭代优化,从而使声源 DOA 和源强分布矢量逐渐收敛至真实值附近。该方法能够获得高精度的二维 DOA 估计和源强量化、实现声源的准确识别,有效克服基于固定网格的离散压缩波束形成存在的基不匹配和现有二维连续压缩波束形成仅适用于矩形网格阵列(及其随机抽选阵列)的应用受限问题。

在本文中, $\mathbf{R}$ 和 $\mathbf{C}$ 分别代表实数集和复数集。加粗的小写字母表示矢量,加粗的大写字母表示矩阵,花体大写字母表示函数。上标“T”和“H”分别表示转置和共轭转置符。“ $|\cdot|$ ”表示标量的模。“ $\triangleq$ ”表示定义。定义矢量 $\mathbf{x} \in \mathbf{C}^n$ 的 l_p 范数为 $\|\mathbf{x}\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{1/p}$, 其中 l_0 范数定义为 $\|\mathbf{x}\|_0 = \sum_{i=1}^n \mathbf{1}_{x_i \neq 0}$ 。符号 $\text{tr}(\mathbf{X})$ 表示 $\mathbf{X}$ 的迹。符号 $\text{diag}\{\mathbf{x}\}$ 表示一个对角线元素为 $\mathbf{x}$ 、非对角线元素均为 0 的方阵。

1 观测模型

图 1 所示为非规则的扇形阵列^[23]在平面波传导声场中接收来自第 i 个声源产生的声压信号的采样模型,其中“ $\bullet$ ”代表传声器,假设传声器数目

为 M 。以阵列中心为原点建立如图 1 所示的笛卡儿坐标系并通过坐标 (θ_i, φ_i) 刻画声源的 DOA。 θ_i 为第 i 个声源的入射方向与 z 轴的夹角, φ_i 为第 i 个声源的入射方向在 $x-y$ 平面上的投影与 x 轴的夹角, 其中 $0^\circ \leq \theta_i \leq 90^\circ$, $0^\circ \leq \varphi_i < 360^\circ$ 。 q_i 为第 i 个声源的源强, 用其在坐标原点产生的声压复幅值表示。则坐标为 (x, y) 的传声器测得声压 $p(x, y)$ 为

$$p(x, y) = \sum_{i=1}^k q_i \exp(j2\pi(t_{1i}x + t_{2i}y)) \quad (1)$$

式中, k 为声源数目; $j = \sqrt{-1}$ 为复数单位; $t_{1i} = \sin\theta_i \cos\varphi_i / \lambda$; $t_{2i} = \sin\theta_i \sin\varphi_i / \lambda$; λ 为波长。

图 1 平面波传导声场中阵列采样声信号模型

令 $\mathbf{x} = (x_1, x_2, \dots, x_M)^T \in \mathbf{R}^M$ 和 $\mathbf{y} = (y_1, y_2, \dots, y_M)^T \in \mathbf{R}^M$ 分别表示由所有传感器的 x 轴和 y 轴坐标所构成的 M 维矢量。构造 $\mathbf{p}^* = (p(x_1, y_1), p(x_2, y_2), \dots, p(x_M, y_M))^T \in \mathbf{C}^M$ 表示由传感器接收信号所构成的 M 维矢量。令 $\mathbf{d}(t_{1i}, t_{2i}) = [\exp(j2\pi(x_1 t_{1i} + y_1 t_{2i})), \exp(j2\pi(x_2 t_{1i} + y_2 t_{2i})), \dots, \exp(j2\pi(x_M t_{1i} + y_M t_{2i}))]^T \in \mathbf{C}^M$, 则 $\mathbf{p}^*$ 可表示为

$$\mathbf{p}^* = \sum_{i=1}^k q_i \mathbf{d}(t_{1i}, t_{2i}) \quad (2)$$

考虑测量噪声干扰, 传声器实际接收的信号 $\mathbf{p}$ 应包含来自真实声源产生的信号 $\mathbf{p}^*$ 和测量噪声干扰信号 $\mathbf{n}$, 即

$$\mathbf{p} = \mathbf{p}^* + \mathbf{n} \quad (3)$$

式中, $\mathbf{n} \in \mathbf{C}^M$ 满足 $SNR = 20 \lg(\|\mathbf{p}^*\|_2 / \|\mathbf{n}\|_2)$ 。在后续的仿真中通过添加一定信噪比(Signal-to-noise ratio, SNR)的高斯白噪声来加以模拟。

综上, 压缩波束形成就是通过传声器阵列测量得到的声压信号 $\mathbf{p}$ 求解声源 DOA 及其源强, 从而准确识别声源。

2 稀疏恢复与基不匹配

传统离散压缩波束形成通常将目标声源区域沿

θ 和 φ 两个方向按照一定间隔离散为 N 个固定网格点, 并假设实际声源 DOA 与网格点一致, 从而将声源识别问题转变为求解如下的方程组

$$\mathbf{p} = \mathbf{A}\mathbf{q} \quad (4)$$

式中, $\mathbf{A} = [\mathbf{d}(t_{11}, t_{21}), \mathbf{d}(t_{12}, t_{22}), \dots, \mathbf{d}(t_{1N}, t_{2N})] \in \mathbf{C}^{M \times N}$ 表示感知矩阵, $\mathbf{q} = (q_1, q_2, \dots, q_N)^T \in \mathbf{C}^N$ 为未知的源强分布矢量。实际应用中往往有 $M \ll N$, 式(4)是严重欠定的线性方程组, 不存在解析解。然而, 当待求矢量 $\mathbf{q}$ 具备稀疏特性, 即仅有极少数元素不为零时, $\mathbf{q}$ 可通过求解如下的 l_0 范数最小化问题来实现准确恢复

$$\min_{\mathbf{q} \in \mathbf{C}^N} \|\mathbf{q}\|_0 \quad \text{s.t.} \quad \|\mathbf{p} - \mathbf{A}\mathbf{q}\|_2 < \xi \quad (5)$$

式中, ξ 表示对噪声信号 $\mathbf{n}$ 的容差。虽然式(5)是一个 NP-hard 问题, 通常难以求解, 但其等价于如下 l_1 范数最小化问题^[4, 7]

$$\min_{\mathbf{q} \in \mathbf{C}^N} \|\mathbf{q}\|_1 \quad \text{s.t.} \quad \|\mathbf{p} - \mathbf{A}\mathbf{q}\|_2 < \xi \quad (6)$$

以 l_1 范数最小化算法为代表的凸松弛法^[24-26]和以正交匹配追踪 (Orthogonal matching pursuit, OMP) 算法为代表的贪婪算法^[27-29]是求解上述稀疏恢复问题的两类常用方法, 然而无论哪一类方法都不能摆脱固定网格引起的基不匹配问题。

图 2 显示了传统延迟求和波束形成、基于 l_1 范数最小化算法求解和 OMP 算法求解的固定网格离散压缩波束形成仿真成像结果, 5 个声源的源强从左到右依次被设置为 2 Pa、 $\sqrt{2}$ Pa、 $\sqrt{2}$ Pa、1 Pa 和 2 Pa, 声源频率为 3 000 Hz, SNR=30 dB。仿真时采用阵列直径为 0.65 m 的 36 通道扇形阵列。 θ 和 φ 两个方向均以 5° 的间隔进行网格划分。其中图 2a~2c 被设置为基匹配的情况, 声源的 DOA 分别为 $(65^\circ, 110^\circ)$ 、 $(70^\circ, 130^\circ)$ 、 $(45^\circ, 160^\circ)$ 、 $(25^\circ, 210^\circ)$ 和 $(45^\circ, 245^\circ)$ 。图 2d~2f 被设置为基不匹配的情况, 声源的 DOA 分别为 $(63^\circ, 107.5^\circ)$ 、 $(68^\circ, 127^\circ)$ 、 $(44^\circ, 162^\circ)$ 、 $(26.5^\circ, 213^\circ)$ 和 $(47^\circ, 243^\circ)$ 。图 2a、2d 为传统延迟求和波束形成, 图 2b、2e 为 l_1 范数最小化算法求解的固定网格传统离散压缩波束形成, 图 2c、2f 为 OMP 算法求解的固定网格传统离散压缩波束形成。其中, 网格点用 “•” 进行表示, 真实声源用 “○” 进行表示, 压缩波束形成所识别出的声源用 “*” 进行表示。计算结果参照最大输出值进行 dB 缩放, 设置显示动态范围为 20 dB, 同时每幅图上方标出以基准声压 2×10^{-5} Pa 为参考的最大输出值, 其中最大的声源强度为 100 dB。后续仿真均采用相同显示设置, 不再赘述。

由图 2a~2c 可见,传统延时求和波束形成能够准确识别声源,但主瓣宽、空间分辨率低、旁瓣污染严重,而传统离散压缩波束形成具有极高的空间分辨率和旁瓣衰减能力,声源成像更清晰。由图 2e 可见,当声源 DOA 与网格点不一致、即发生基不匹配时, l_1 范数最小化算法求解的传统离散压缩波束形成结果中声源能量被分散到了若干相邻的网格点,难以准确定位真实声源位置,且源能量亦出现较大偏差。同样,由图 2f 可见,OMP 算法求解的传统离散压缩波束形成对声源的定位亦出现了偏差,只能识别为真实声源附近的网格点,不能准确识别声源位置。

图 2 传统延时求和波束形成和固定网格的传统离散压缩波束形成的仿真成像云图

3 动态网格压缩波束形成

3.1 构造替代函数

与传统离散压缩波束形成采用固定网格、且计算对应于固定网格的源强分布不同,所提出的二维动态网格压缩波束形成方法则同时优化动态网格坐标及其对应的源强分布,采用基于优化最小化框架的迭代重加权算法求解相应数学模型。优化最小化框架的核心是寻找比原始目标函数更容易优化的替代(目标)函数,使其最优解在满足一定条件时无限逼近于原始目标函数的最优解。二维动态网格压缩波束形成方法通过以目标函数为基础构建一个关于网格坐标和源强分布的合适

为解决基不匹配问题,连续压缩波束形成通过引入原子范数在连续的参数空间中计算 DOA,避免了网格划分。然而已有的二维无网格压缩波束形成由于基于半正定规划进行求解,仅适用于传声器均匀分布的矩形阵列(及其随机抽选阵列),且需要准确估计噪声水平,限制了其应用。为解决这些问题,本文提出了一种二维动态网格压缩波束形成,其放宽了连续压缩波束形成对阵列传声器布置形式的限制、适用于传声器随机分布的平面阵列,克服了离散压缩波束形成固有的基不匹配缺陷,且无需信噪比或声源稀疏度等先验信息,显示出优异的声源识别性能。

替代函数,可实现包括网格坐标和对应源强分布在内的多参数同步优化,从而规避了因网格固定引起的基不匹配问题。

由于采用动态网格、即网格坐标亦是优化参数,则感知矩阵不再是常矩阵,而是变成了关于网格坐标 θ 和 φ 的函数 $\mathbf{A}(\theta, \varphi) = [\mathbf{d}(t_{11}, t_{21})(\theta_1, \varphi_1), \mathbf{d}(t_{12}, t_{22})(\theta_2, \varphi_2), \dots, \mathbf{d}(t_{1N}, t_{2N})(\theta_N, \varphi_N)]$, 其中 $\theta = [\theta_1, \theta_2, \dots, \theta_N] \in \mathbf{R}^N$ 和 $\varphi = [\varphi_1, \varphi_2, \dots, \varphi_N] \in \mathbf{R}^N$ 。式(6)相应变为

$$\min_{q \in \mathbf{C}^N, \theta \in \mathbf{R}^N, \varphi \in \mathbf{R}^N} \|\mathbf{q}\|_1 \text{ s.t. } \|\mathbf{p} - \mathbf{A}(\theta, \varphi)\mathbf{q}\|_2 < \xi \quad (7)$$

为进一步促进解的稀疏性,在式(7)的基础上引入对数求和罚函数来代替 l_1 范数最小化优化目标函数,则式(7)可改写为

$$\min_{q \in \mathbb{C}^N, \theta \in \mathbb{R}^N, \varphi \in \mathbb{R}^N} \mathcal{L}(q) = \sum_{n=1}^N \ln(|q_n|^2 + \epsilon) \quad (8)$$

$$\text{s.t. } \|p - A(\theta, \varphi)q\|_2 < \xi$$

式中, ϵ 为一正参数, 它具有双重作用, 一方面可以保证对数函数被正确定义, 另一方面则是作为迭代进程的控制参数。通过将 ϵ 设置为某个较小的初值(比如 1), 并令其随着迭代过程逐渐缩减至 0, 可以保证式(8)的全局最优解能收敛至真实解附近^[17]。在式(8)的目标函数中增加正则化惩罚项 $\lambda \|p - A(\theta, \varphi)q\|_2^2$, 则式(8)可进一步表达成如下的无约束形式

$$\min_{q \in \mathbb{C}^N, \theta \in \mathbb{R}^N, \varphi \in \mathbb{R}^N} \mathcal{L}(q) + \lambda \|p - A(\theta, \varphi)q\|_2^2 \quad (9)$$

式中, λ 为正则化参数, 其大小会影响重构精度和解的稀疏度。为了获得更好的识别性能, λ 被设置为伴随迭代进程变化的参数(选取方式会在后面部分加以说明)。由于直接求解式(9)的最优解非常困难, 优化最小化框架通过优化合适的替代函数来间接实现对目标函数的优化。

用 $\hat{q}^{(t)} = [\hat{q}_1^{(t)}, \hat{q}_2^{(t)}, \dots, \hat{q}_N^{(t)}]$ 和 $\lambda^{(t)}$ 分别表示第 t 次迭代后获得的源强分布矢量和 λ 。在第 $t+1$ 次, 构造一个关于 $\mathcal{L}(q)$ 的替代函数 $\mathcal{P}(q|\hat{q}^{(t)})$, 满足 $\mathcal{P}(q|\hat{q}^{(t)}) - \mathcal{L}(q) \geq 0$ (当且仅当 $q = \hat{q}^{(t)}$ 时, 等号成立)。其中

$$\mathcal{P}(q|\hat{q}^{(t)}) \triangleq \sum_{n=1}^N \left(\frac{|q_n|^2 + \epsilon}{|\hat{q}_n^{(t)}|^2 + \epsilon} + \ln(|\hat{q}_n^{(t)}|^2 + \epsilon) - 1 \right) \quad (10)$$

相应地, 式(9)亦改写为如下的替代函数形式

$$\min_{q \in \mathbb{C}^N, \theta \in \mathbb{R}^N, \varphi \in \mathbb{R}^N} \mathcal{F}(q, \theta, \varphi|\hat{q}^{(t)}) \triangleq \mathcal{P}(q|\hat{q}^{(t)}) + \lambda^{(t)} \|p - A(\theta, \varphi)q\|_2^2 \quad (11)$$

根据优化最小化框架的优化流程, 通过迭代构造替代函数 $\mathcal{F}$ 并加以求解即可获得与式(7)中目标函数相近的全局最优解。

3.2 求解替代函数

首先移除式(11)中常数项将其简化为

$$\min_{q \in \mathbb{C}^N, \theta \in \mathbb{R}^N, \varphi \in \mathbb{R}^N} q^H D^{(t)} q + \lambda^{(t)} \|p - A(\theta, \varphi)q\|_2^2 \quad (12)$$

$$D^{(t)} = \text{diag} \left\{ \frac{1}{|\hat{q}_1^{(t)}|^2 + \epsilon}, \dots, \frac{1}{|\hat{q}_N^{(t)}|^2 + \epsilon} \right\}$$

固定 θ 和 φ , 通过拉格朗日乘子法可以得到式

(12)关于 q 的优化解 $\tilde{q}$ 为

$$\tilde{q}(\theta, \varphi) =$$

$$\left(A^H(\theta, \varphi) A(\theta, \varphi) + (\lambda^{(t)})^{-1} D^{(t)} \right)^{-1} A^H(\theta, \varphi) p \quad (13)$$

将 $\tilde{q}(\theta, \varphi)$ 代回式(12), 则问题被进一步简化为对网格坐标 θ 和 φ 的优化问题

$$\min_{\theta \in \mathbb{R}^N, \varphi \in \mathbb{R}^N} \mathcal{R}(\theta, \varphi) \triangleq -p^H A(\theta, \varphi) \cdot \left(A^H(\theta, \varphi) A(\theta, \varphi) + (\lambda^{(t)})^{-1} D^{(t)} \right)^{-1} A^H(\theta, \varphi) p \quad (14)$$

式(14)可采用梯度下降法求解, 从而获得新一轮网格坐标 $\hat{\theta}^{(t+1)}$ 和 $\hat{\varphi}^{(t+1)}$ 。每个网格点的更新如下

$$\begin{cases} \hat{\theta}_n^{(t+1)} = \hat{\theta}_n^{(t)} - \alpha \frac{\partial \mathcal{R}(\theta, \varphi)}{\partial \theta_n} \\ \hat{\varphi}_n^{(t+1)} = \hat{\varphi}_n^{(t)} - \alpha \frac{\partial \mathcal{R}(\theta, \varphi)}{\partial \varphi_n} \end{cases} \quad n=1, 2, \dots, N \quad (15)$$

式中, α 为初始步长, 本文取 $\alpha = 10^3$ 。 $\frac{\partial \mathcal{R}(\theta, \varphi)}{\partial \theta_n}$ 和

$\frac{\partial \mathcal{R}(\theta, \varphi)}{\partial \varphi_n}$ 为 $\mathcal{R}(\theta, \varphi)$ 关于 θ_n 和 φ_n 的一阶偏导数。在梯度下降法中, 步长按 0.1 倍的变化规律进行修正以确保更新后的 $\hat{\theta}^{(t+1)}$ 和 $\hat{\varphi}^{(t+1)}$ 满足

$$\mathcal{R}(\hat{\theta}^{(t+1)}, \hat{\varphi}^{(t+1)}) \leq \mathcal{R}(\hat{\theta}^{(t)}, \hat{\varphi}^{(t)}) \quad (16)$$

将 $\hat{\theta}^{(t+1)}$ 和 $\hat{\varphi}^{(t+1)}$ 代回式(13)即可得到对源强分布矢量的新一轮估计

$$\hat{q}^{(t+1)} = \tilde{q}(\hat{\theta}^{(t+1)}, \hat{\varphi}^{(t+1)}) \quad (17)$$

之后, 为了在迭代过程中保证重构精度和稀疏度之间的平衡, 根据式(18)对下一轮迭代中 λ 的取值进行更新^[17]

$$\lambda^{(t+1)} = \frac{\delta M}{\|p - A(\hat{\theta}^{(t+1)}, \hat{\varphi}^{(t+1)}) \hat{q}^{(t+1)}\|_2^2} \quad (18)$$

式中, δ 为一恒定缩放常数, 推荐取值为 0.05~1。

通过循环构造和迭代缩减替代函数, 网格坐标 θ 和 φ 逐渐收敛至真实声源所在的位置, 源强分布矢量 q 也在这个过程中得到愈加准确的估计。

3.3 仿真算例

迭代求解得到的 θ 和 φ 指示了声源所在方向, 源强分布矢量 q 量化了所识别声源的强度。为了保证结果准确性的同时兼顾计算成本, 迭代过程中伴随着一些控制措施的执行。定义 $\sigma^{(t+1)} = \|\hat{q}^{(t+1)} - \hat{q}^{(t)}\|_2 / \|\hat{q}^{(t)}\|_2$ 为第 $t+1$ 次迭代的收敛

误差, 则随着迭代结果的收敛, σ 应收敛到 0 附近。为避免冗余计算, 两种终止条件被同时应用于判断是否需要退出迭代, 包括 σ 是否小于设定阈值 σ_{tol} 和是否达到允许的最大迭代次数 t_{max} 。为进一步降低计算复杂度, 预设阈值 κ , 在每次迭代中通过剔除源强分布矢量 $\mathbf{q}$ 中小于 κ 的小值和其对应的网格点坐标可以实时动态减少感知矩阵 $\mathbf{A}(\theta, \varphi)$ 的维度, 从而极大地缩短更新感知矩阵所需的计算时间。实现该算法的主要步骤可总结如下。

(1) 输入声压信号 $\mathbf{p}$ 、初始网格坐标 $\hat{\theta}^{(0)}$ 和 $\hat{\varphi}^{(0)}$ 、正则化参数 $\lambda^{(0)}$ 。初始化源强分布 $\hat{\mathbf{q}}^{(0)} = \mathbf{0}$, $t=0$ 。

(2) 基于 $\hat{\mathbf{q}}^{(t)}$, $\hat{\theta}^{(t)}$, $\hat{\varphi}^{(t)}$, $\lambda^{(t)}$ 构建如式(11)所示的替代函数。

(3) 利用梯度下降法根据式(15)寻找满足不等式(16)的新网格坐标估计 $\hat{\theta}^{(t+1)}$ 和 $\hat{\varphi}^{(t+1)}$ 。

(4) 通过式(13)、(17)计算新的源强分布估计 $\hat{\mathbf{q}}^{(t+1)}$ 。

(5) 通过式(18)计算下一轮迭代中的正则化参数 $\lambda^{(t+1)}$ 。

(6) 如果 $\sigma^{(t+1)} > \sigma_{\text{tol}}$ 且 $t < t_{\text{max}}$, 令 $t=t+1$, 返回步骤(2); 否则, 结束迭代。

本文中设置 $\sigma_{\text{tol}} = 10^{-6}$, $t_{\text{max}} = 200$, $\lambda^{(0)} = 10^3$ 。设置声源条件、阵列形式与图 2d~2f 相同, 预设 $\kappa = 0.5 \text{ Pa}$, θ 和 φ 两个方向以 18° 的间隔进行初始网格划分, 模拟计算实际声源 DOA 与初始网格存在基不匹配时二维动态网格压缩波束形成的识别结果, 其迭代期间和迭代完成时的成像云图如图 3 所示。图 3a~3d 分别是初始化及迭代进行到第 5 次, 第 8 次和完成时的成像结果。

由图 3 可见, 动态网格压缩波束形成迭代过程中同时进行着对源强分布矢量和网格坐标的更新。原本偏离真实声源 DOA 的网格点会逐渐向真实声源 DOA 逼近, 并且其对应的源强估计亦会逐渐收敛至真实值。最终对应于真实声源 DOA 的网格点被保留, 其他网格点则会随着迭代的进行而被剔除。仿真表明动态网格压缩波束形成能够准确定位声源并量化源强。

使用搭载 3.5 GHz 6 核处理器和 32 GB 内存的计算机时, 所提出方法的计算耗时为 15.91 s。其计算复杂度主要来源于第二步求解替代函数过程中对网格坐标所进行的逐个优化。

图 3 动态网格压缩波束形成的仿真成像云图

4 试验验证

为验证所提出方法的正确性和有效性, 采用集成了 36 支 Brüel & Kjær 4958 型传声器、直径为 0.65 m 的扇形轮阵列进行给定扬声器声源的识别试验。图 4 为试验现场布置。2 个扬声器和 2 个由于地面反射形成的镜像声源的坐标分别为 (1.25, 0, 3) m, (1.25, -2.4, 3) m, (-1.25, 0, 3) m, (-1.25, -2.4, 3) m, 对应的 DOA 分别为 $(22.62^\circ, 0^\circ)$, $(42.05^\circ, 297.51^\circ)$, $(22.62^\circ, 180^\circ)$, $(42.05^\circ, 242.49^\circ)$ 。各传声器接收的声压信号经过 Brüel & Kjær PULSE 3660D 型数据采集系统同步采集并传输到 Brüel & Kjær PULSE LABSHOP 软件中进行 FFT 分析计算得到其频谱。采样频率为 16 384 Hz, 频率分辨率为 4 Hz。图 5 展示了 3 000 Hz 下不同算法的成像结果, 其中图 5a~5d 使用全部 36 支传声器的测量声压信号, 图 5e~5f 仅使用了随机挑选的 20 支传声器的

测量声压信号。对应于图 5a、5e 的传统延时求和波束形成、图 5b、5f 的 l_1 范数最小化算法求解的固定网格传统离散压缩波束形成以及图 5c、5g 的 OMP 算法求解的固定网格传统离散压缩波束形成, θ 和 φ 两个方向均以 5° 的间隔进行网格划分; 对应于图 5d、5h 的动态网格压缩波束形成, θ 和 φ 两个方向以 18° 的间隔进行初始网格划分。

图4 试验现场布置图

图5 试验成像结果

5 性能分析

该节通过仿真分析探讨声源频率、SNR、传声器数量以及声源间隔对动态网格压缩波束形成声源

识别性能的影响。引入定位误差 DOA_{error} 和量化误差 STR_{error} 来分别度量波达方向估计精度和源强量化精度。其中, DOA_{error} 定义为所有声源估计出的 DOA 与设定 DOA 之间夹角的平均值, STR_{error} 定义为所有声源量化出的源强与设定源强之间的相对误差。

由图 5a、5e 可见, 传统延迟求和波束形成方法空间分辨率低、旁瓣污染严重, 在传声器数量减少之后传统延迟求和波束形成的成像结果进一步恶化, 难以识别真实的声源位置。由图 5b、5c、5f、5g 可见, 基于 l_1 范数最小化算法和 OMP 算法求解的固定网格传统离散压缩波束形成具有优于传统延迟求和波束形成的空间分辨率和更低的旁瓣水平, 然而由于存在基不匹配, 二者均产生明显的定位偏差和量化偏差, 并且在传声器数量减少之后声源定位结果恶化。由图 5d、5h 可见, 动态网格压缩波束形成方法能够获得高精度的 DOA 估计和源强量化结果, 且在传声器数量较少的情况下仍然能够准确识别声源。试验结论与仿真结论一致。

差的平均值, 其表达式分别如下

$$\text{DOA}_{\text{error}} = \frac{1}{k} \sum_{i=1}^k \arccos(\cos \theta_i \cdot \cos \hat{\theta}_i + \cos(\varphi_i - \hat{\varphi}_i) \cdot \sin \theta_i \cdot \sin \hat{\theta}_i) \quad (19)$$

$$\text{STR}_{\text{error}} = \frac{1}{k} \sum_{i=1}^k \left| 20 \lg \left(\frac{|q_i|}{|\hat{q}_i|} \right) \right| \quad (20)$$

5.1 声源频率和 SNR 对声源识别性能影响

在可行空间域内, 随机生成声压级在 91 dB 到 100 dB 之间且夹角不小于 5° 的不等强同频率双声源, 并施加 SNR 为 30 dB 的高斯白噪声。仿真频率从 1 000 Hz 开始以 500 Hz 步长逐渐增加到 6 000 Hz。每个频率进行 100 次蒙特卡洛仿真, 得到的平均 $\text{DOA}_{\text{error}}$ 和平均 $\text{STR}_{\text{error}}$ 结果如图 6 所示。由图 6 可见, 所提出方法对声源频率为 1 500 Hz 以上的声源平均 $\text{DOA}_{\text{error}}$ 低于 1°, 平均 $\text{STR}_{\text{error}}$ 低于 0.2 dB。

图6 平均定位误差和量化误差随声源频率的变化曲线

进一步在 1 500 Hz 到 6 000 Hz 之间随机生成声源频率, SNR 从 15 dB 开始以 5 dB 步长逐渐增加到 40 dB, 其余仿真参数与图 6 相同。对每个 SNR 亦进行 100 次蒙特卡洛仿真, 得到的平均 $\text{DOA}_{\text{error}}$ 和平均 $\text{STR}_{\text{error}}$ 如图 7 所示。可见, 在 SNR 优于 20 dB 时, 所提出方法的平均 $\text{DOA}_{\text{error}}$ 低于 1°, 平均 $\text{STR}_{\text{error}}$ 低于 0.2 dB; 而当 SNR 进一步降低后, 识别误差随之迅速增大。究其原因, 如式(12)所示, 随着噪声增大, 残差范数 $\|p - A(\theta, \varphi)q\|_2^2$ 将会增大, 最小化 $q^H D^{(t)} q + \lambda^{(t)} \|p - A(\theta, \varphi)q\|_2^2$ 的过程中需要对 $q^H D^{(t)} q$ 和 $\|p - A(\theta, \varphi)q\|_2^2$ 进行平衡, 使得对解产生了过滤波, 进而导致偏差变大。

综上, 与其他波束形成方法类似, 动态网格压缩波束形成依然特别适用于中高频声源识别, 在信噪比达到 20 dB 以上的测试环境中能够获得极高的 DOA 估计精度和源强量化精度。

图7 平均定位误差和量化误差随信噪比的变化曲线

5.2 传声器数量对声源识别性能影响

传声器数量亦是影响波束形成声源识别性能的重要因素。使用传声器数量充足的阵列通常能够获得很好的识别性能, 如能在保证识别性能的前提下采用尽可能少的传声器则更具实际意义, 尤其是对传声器数量有限抑或是出现部分传声器损坏的情况。为探究动态网格压缩波束形成在采用少量传声器时的识别性能, 固定信噪比为 30 dB, 从上述扇形轮传声器阵列中随机挑选部分传声器进行仿真, 传声器数量从 15 支开始以 5 支步长逐渐增加到 35 支, 其余仿真参数与图 7 相同, 每种传声器数量进行 100 次蒙特卡洛仿真, 得到的平均 $\text{DOA}_{\text{error}}$ 和平均 $\text{STR}_{\text{error}}$ 如图 8 所示。

图8 平均定位误差和量化误差随传声器数量的变化曲线

由图 8 可见, 就本节的声源配置情况而言, 所提出方法在传声器数目低至 20 的时候仍然显示出极高的定位精度和量化精度, 而当传声器数目进一步降低时则平均 $\text{DOA}_{\text{error}}$ 和平均 $\text{STR}_{\text{error}}$ 均会骤增。事实上, 与基于固定网格的压缩波束形成相似, 动态网格压缩波束形成在实际应用时亦需兼顾声场的待识别声源数量和传声器数量^[7]。

5.3 声源间隔对声源识别性能影响

为探究所提出方法对相邻声源的分辨能力, 引入识别成功率这一指标, 定义为识别成功次数和总

独立运行次数的比值。如果在一次试验中每个识别声源的 DOA 与其真实 DOA 的夹角均小于 1° ；本节才认为该次识别结果为成功。令两个声源 DOA 的夹角 Δ 从 1° 开始以 1° 为步长逐渐增加到 15° 。固定信噪比为 30 dB，其余仿真参数与图 7 相同。对每一个夹角进行 100 次蒙特卡洛仿真，得到识别成功率随声源 DOA 夹角 Δ 的变化曲线如图 9 所示。

图9 识别成功率随声源 DOA 夹角 Δ 的变化规律

由图 9 可见，所提出方法对夹角较小的相邻声源分辨能力不足，但随着声源夹角的增大，能够高精度分辨声源的成功率会显著提高。其中，对夹角在 10° 以上的相邻声源的识别成功率超 80%。

6 结论

提出了二维动态网格压缩波束形成声源识别方法，与传统离散压缩波束形成采用固定网格、且只优化对应于固定网格的源强分布不同，二维动态网格压缩波束形成则同时优化动态网格坐标及其对应的源强分布，并采用基于优化最小化框架的迭代重加权算法求解相应数学模型。给定扬声器的仿真模拟和验证试验结果表明如下。

(1) 与传统固定网格的离散压缩波束形成相比，动态网格压缩波束形成能够克服基不匹配问题并获得更高的定位精度和量化精度。就本文的传声器非规则分布的 36 通道扇形阵列而言，基于蒙特卡洛方法的声源识别性能分析结果表明：在信噪比不小于 20 dB 且频率不小于 1 500 Hz 时，平均声源定位精度优于 1° ，平均声源源强量化精度优于 0.2 dB。

(2) 与现有二维连续压缩波束形成仅能适用于矩形网格平面阵列(及其随机抽选阵列)不同，动态网格压缩波束形成能够适用于传声器随机分布的平面阵列，且无需信噪比(噪声干扰)及声源稀疏度等

先验信息。

(3) 随着信噪比、声源频率以及传声器数量的降低，声源识别性能亦总体呈现降低的趋势，这与其他波束形成方法呈现的特性一致。

此外，需要指出的是，动态网格压缩波束形成虽然能对夹角较大的声源实现高概率准确识别，但对紧密相邻的小夹角声源的分离能力有限，探索提高该方法分离能力的措施是值得进一步努力的方向。

参 考 文 献

- [1] CHIARIOTTI P, MARTARELLI M, CASTELLINI P. Acoustic beamforming for noise source localization: Reviews, methodology and applications[J]. Mechanical Systems and Signal Processing, 2019, 120: 422-448.
- [2] 黎术, 徐中明, 贺岩松, 等. 基于函数广义逆波束形成的声源识别[J]. 机械工程学报, 2016, 52(4): 1-6.
LI Shu, XU Zhongming, HE Yansong, et al. Sound source identification based on functional generalized inverse beamforming[J]. Journal of Mechanical Engineering, 2016, 52(4): 1-6.
- [3] 褚志刚, 段云炆, 沈林邦, 等. 函数波束形成声源识别性能分析及应用[J]. 机械工程学报, 2017, 53(4): 67-76.
CHU Zhigang, DUAN Yunyang, SHEN Linbang, et al. Performance analysis and application of functional beamforming sound source identification[J]. Journal of Mechanical Engineering, 2017, 53(4): 67-76.
- [4] XENAKI A, GERSTOFT P, MOSEGAARD K. Compressive beamforming[J]. Journal of the Acoustical Society of America, 2014, 136(1): 260-271.
- [5] ZHONG S, WEI Q, HUANG X. Compressive sensing beamforming based on covariance for acoustic imaging with noisy measurements[J]. Journal of the Acoustical Society of America, 2013, 134(5): 445-451.
- [6] NING F, WEI J, QIU L, et al. Three-dimensional acoustic imaging with planar microphone arrays and compressive sensing[J]. Journal of Sound and Vibration, 2016, 380: 112-128.
- [7] 宁方立, 卫金刚, 刘勇, 等. 压缩感知声源定位方法研究[J]. 机械工程学报, 2016, 52(19): 42-52.
NING Fangli, WEI Jingang, LIU Yong, et al. Study on sound sources localization using compressive sensing[J]. Journal of Mechanical Engineering, 2016, 52(19): 42-52.
- [8] EDELMANN G F, GAUMOND C F. Beamforming using compressive sensing[J]. Journal of the Acoustical Society of America, 2011, 130(4): 232-237.

- [9] GERSTOFT P, XENAKI A, MECKLENBRÄUKER C F. Multiple and single snapshot compressive beamforming[J]. Journal of the Acoustical Society of America, 2015, 138(4): 2003-2014.
- [10] CHI Y, SCHARF L L, PEZESHKI A, et al. Sensitivity to basis mismatch in compressed sensing[J]. IEEE Transactions on Signal Processing, 2011, 59(5): 2182-2195.
- [11] CHAE D H, SADEGHI P, KENNEDY R. Effects of basis-mismatch in compressive sampling of continuous sinusoidal signals[C]// Proceedings of the 2010 2nd International Conference on Future Computer and Communication, May 21-24, 2010, Wuhan China. IEEE, Article number: 5497605.
- [12] CANDÈS E J. The restricted isometry property and its implications for compressed sensing[J]. Comptes Rendus-Mathématique, 2008, 346(9-10): 589-592.
- [13] FANNJIANG A, LIAO W. Coherence pattern-guided compressive sensing with unresolved grids[J]. SIAM Journal on Imaging Sciences, 2012, 5(1): 179-202.
- [14] XENAKI A, GERSTOFT P. Grid-free compressive beamforming[J]. Journal of the Acoustical Society of America, 2015, 137(4): 1923-1935.
- [15] YANG Y, CHU Z, XU Z, et al. Two-dimensional grid-free compressive beamforming[J]. Journal of the Acoustical Society of America, 2017, 142(2): 618-629.
- [16] YANG Y, CHU Z, PING G, et al. Resolution enhancement of two-dimensional grid-free compressive beamforming[J]. Journal of the Acoustical Society of America, 2018, 143(6): 3860-3872.
- [17] FANG J, WANG F, SHEN Y, et al. Super-resolution compressed sensing for line spectral estimation: An iterative reweighted approach[J]. IEEE Transactions on Signal Processing, 2016, 64(18): 4649-4662.
- [18] SUN Y, BABU P, PALOMAR D P. Majorization-minimization algorithms in signal processing, communications, and machine learning[J]. IEEE Transactions on Signal Processing, 2017, 65(3): 794-816.
- [19] CANDÈS E J, WAKIN M B, BOYD S P. Enhancing sparsity by reweighted l_1 minimization[J]. Journal of Fourier Analysis and Applications, 2008, 14(5-6): 877-905.
- [20] LANGE K, HUNTER D R, YANG I. Optimization transfer using surrogate objective functions[J]. Journal of Computational and Graphical Statistics, 2000, 9(1): 1-20.
- [21] COIFMAN R R, WICKERHAUSER M V. Entropy-based algorithms for best basis selection[J]. IEEE Transactions on Information Theory, 1992, 38(2): 713-718.
- [22] SHEN Y, FANG J, LI H. Exact reconstruction analysis of log-sum minimization for compressed sensing[J]. IEEE Signal Processing Letters, 2013, 20(12): 1223-1226.
- [23] SIMARD P, ANTONI J. Acoustic source identification: Experimenting the l_1 minimization approach[J]. Applied Acoustics, 2013, 74(7): 974-986.
- [24] CHEN S S B, DONOHO D L, SAUNDERS M A. Atomic decomposition by basis pursuit[J]. SIAM Review, 2001, 43(1): 129-159.
- [25] CANDÈS E J, TAO T. Decoding by linear programming[J]. IEEE Transactions on Information Theory, 2006, 51(12): 4203-4215.
- [26] DAUBECHIES I, DEFRISE M, DE MOL C. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint[J]. Communications on Pure and Applied Mathematics, 2004, 57(11): 1413-1457.
- [27] MALLAT S G, ZHANG Z F. Matching pursuits with time-frequency dictionaries[J]. IEEE Transactions on Signal Processing, 1993, 41(12): 3397-3415.
- [28] NEEDELL D, TROPP J A. CoSaMP: Iterative signal recovery from incomplete and inaccurate samples[J]. Applied and Computational Harmonic Analysis, 2009, 26(3): 301-321.
- [29] DAI W, MILENKOVIC O. Subspace pursuit for compressive sensing signal reconstruction[J]. IEEE Transactions on Information Theory, 2009, 55(5): 2230-2249.

作者简介: 樊小鹏, 男, 1986年出生, 博士。主要研究方向为电力系统噪声环境保护、电气设备故障诊断。

E-mail: 283128493@qq.com

余立超, 男, 1994年出生。主要研究方向为阵列声源识别技术及其应用。

E-mail: 810177064@qq.com

褚志刚(通信作者), 男, 1978年出生, 博士, 教授, 博士研究生导师。主要研究方向为振动噪声测量分析技术, 数字信号处理, 噪声源识别技术。

E-mail: zgchu@cqu.edu.cn

杨洋, 女, 1988年出生, 博士, 讲师。主要研究方向为噪声源识别技术理论及其应用, 工程信号处理。

E-mail: yangyang911127@cqu.edu.cn

李丽, 女, 1971年出生, 博士, 教授级高级工程师。主要研究方向为电力系统噪声环境保护、电气设备故障诊断。

E-mail: liligz@163.com