

◇ 研究报告 ◇

# 基于 Group Lasso 的多重信号分类声源 定位优化算法\*

吴江涛 胡定玉<sup>†</sup> 方宇 朱文发

(上海工程技术大学城市轨道交通学院 上海 201620)

**摘要** 多重信号分类算法因其抑制噪声能力强、计算速度快等优点,在声源定位领域得到广泛应用。但该算法在中低频段分辨率及聚焦性能较差。针对该问题,提出一种基于 Group Lasso 的多重信号分类优化算法。该算法将多重信号分类算法输出值作为初始值,并在 Group Lasso 算法组间计算时对目标信号进行稀疏、在组内计算时对该组信号进行平滑及阈值截断。仿真结果表明:该优化算法在中低频段可明显提高多重信号分类算法分辨率,同时改善因扫描位置与声源面位置不重合引起的聚焦性能下降问题。

**关键词** 多重信号分类算法, Group Lasso, 声源定位

中图分类号: TB52+5

文献标识码: A

文章编号: 1000-310X(2019)02-0261-06

DOI: 10.11684/j.issn.1000-310X.2019.02.016

## An optimized multiple signal classification algorithm based on Group Lasso for sound localization

WU Jiangtao HU Dingyu FANG Yu ZHU Wenfa

(School of Urban Rail Transportation, Shanghai University of Engineering Science, Shanghai 201620, China)

**Abstract** Multiple signal classification (MUSIC) algorithm is widely used in the field of sound source localization due to its robustness to noise and computation efficiency. However, this algorithm has poor resolution and focusing performance in the low and medium frequency bands. In view of this problem, a MUSIC algorithm optimized by Group Lasso algorithm is proposed. The output of MUSIC algorithm is used as the initial value. When the Group Lasso algorithm group is calculated, the target signal is sparse and calculated in the group. The set of signals is smoothed and the threshold is truncated. The simulation results show that the optimized algorithm can significantly improve the resolution of the MUSIC algorithm in the middle and low frequency bands, and at the same time, the problem of degraded focusing performance caused by the non-coincidence of scanning position and sound source surface position is improved.

**Key words** Multiple signal classification algorithm, Group Lasso, Source location

2018-05-21 收稿; 2018-10-15 定稿

\*国家自然科学基金青年基金项目(51605274), 上海工程技术大学展翅计划项目(RC152017), 上海工程技术大学研究生科研创新项目(17KY1012)

作者简介: 吴江涛(1993-), 男, 山东潍坊人, 硕士研究生, 研究方向: 噪声源识别与定位。

<sup>†</sup> 通讯作者 E-mail: dyhu1987@sues.edu.cn

## 0 引言

波束形成<sup>[1]</sup>是一种应用于信号波达方向(Direction of arrival, DOA)估计、声源定位的空间滤波处理方法,在汽车、航天、军事等领域都有着广泛应用,并在多声源定位领域得到发展<sup>[2]</sup>。目前已发展出多种波束形成算法,其中多重信号分类(Multiple signal classification, MUSIC)算法利用噪声子空间来最大化声源信号,较传统波束形成算法、Capon算法在抑制噪声能力、定位效果准确性上有明显优势。

MUSIC算法最早由Schmidt<sup>[3]</sup>提出,并迅速在语音识别、三维声源定位等方向得到研究并应用<sup>[4-5]</sup>。以MUSIC算法作为基础,加权MUSIC算法<sup>[6]</sup>及改进MUSIC算法<sup>[7-8]</sup>进一步提高了MUSIC算法的准确性及适用性。为提高MUSIC算法抗干扰能力,Gardner<sup>[9]</sup>将信号的循环平稳特性用于DOA估计,黄知涛等<sup>[10]</sup>利用信号的循环平稳特性提出了循环互相关MUSIC算法,王超等<sup>[11]</sup>改进了预滤波MUSIC算法,提高了目标声源分辨率。一系列的研究进一步提高了该算法的抗干扰能力及运算速度,但现有算法存在中低频段定位效果不准确及聚焦效果较差的问题。

针对MUSIC算法在中低频段分辨率低及聚焦性能差这一缺点,本文提出基于Group Lasso<sup>[12]</sup>改进的MUSIC算法,通过Group Lasso算法增强MUSIC算法声源定位聚焦效果。优化算法中,将MUSIC算法计算结果作为Group Lasso算法初始值,将导向矩阵与测量值作为Group Lasso算法输入变量,通过组间L1范数正则化,对扫描声源面信息进行稀疏筛选,且使用L2范数并加入阈值截断进行组内计算。通过仿真分析,证明基于Group Lasso的MUSIC算法具有增强中低频段声源定位分辨率及聚焦性能的优势。

## 1 基于Group Lasso的MUSIC算法声源定位

### 1.1 MUSIC算法声源定位原理

MUSIC算法在波束形成算法基础上,将由测量信号得到的协方差矩阵分为信号子空间和噪声子空间,利用噪声子空间来最大化空间扫描声源面声源点输出值,获取高分辨率结果和优良聚焦性能<sup>[3]</sup>。

MUSIC算法传声器阵列测量值 $\mathbf{Z}$ 表达式为

$$\mathbf{Z} = \mathbf{A} \times \mathbf{s} + \mathbf{n}, \quad (1)$$

式(1)中, $\mathbf{s}$ 为 $K \times 1$ 维声源信号, $K$ 为声源数; $\mathbf{n}$ 为 $M \times 1$ 维噪声信号, $\mathbf{Z}$ 是 $M \times 1$ 维向量。 $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_K]$ 为 $M \times K$ 维导向矩阵<sup>[13]</sup>,其中 $\mathbf{a}$ 为导向向量,对于近场声源定位问题,采用球面波假设,其表达式为

$$\mathbf{a} = \left[ \frac{1}{r_{k,1}} e^{-i2\pi f \tau_{k,1}}, \dots, \frac{1}{r_{k,M}} e^{-i2\pi f \tau_{k,M}} \right]^T, \quad (2)$$

式(2)中, $r_{k,m}$ 为第 $k$ 个声源到第 $m$ 个传声器的距离, $\tau_{k,m}$ 为第 $k$ 个声源到第 $m$ 个传声器的声辐射传播时间, $f$ 为声源频率。MUSIC算法原理是根据噪声子空间特征向量与导向矩阵列向量正交的关系来获取扫描点输出值,为得到噪声与信号子空间,首先得到协方差矩阵

$$\mathbf{R}_Z = \mathbf{Z}\mathbf{Z}^H/I, \quad (3)$$

式(3)中, $I$ 为总的采样数,协方差矩阵 $\mathbf{R}_Z$ 可分解为信号子空间 $\mathbf{R}_S$ 和噪声子空间 $\mathbf{R}_n$ ,即 $\mathbf{R}_Z = \mathbf{R}_S + \mathbf{R}_n$ 。其中, $\mathbf{R}_S = \mathbf{U}_S \mathbf{\Sigma}_S \mathbf{U}_S^H$ , $\mathbf{U}_S$ 表示信号子空间特征向量, $\mathbf{\Sigma}_S$ 表示信号子空间特征值; $\mathbf{R}_n = \mathbf{U}_n \mathbf{\Sigma}_n \mathbf{U}_n^H$ , $\mathbf{U}_n$ 表示噪声子空间特征向量, $\mathbf{\Sigma}_n$ 表示噪声子空间特征值。对 $\mathbf{R}_Z$ 进行特征值分解可得到

$$\mathbf{R}_Z = \mathbf{U}_Z \mathbf{\Sigma}_Z \mathbf{U}_Z^H, \quad (4)$$

式(4)中, $\mathbf{U}_Z$ 为协方差矩阵特征向量, $\mathbf{\Sigma}_Z = \text{diag}[\lambda_1, \dots, \lambda_M]$ 为对应于特征向量的特征值且特征值 $\lambda_1 > \lambda_2 > \dots > \lambda_M$ ,特征向量正定( $\mathbf{U}_Z \mathbf{U}_Z^H = \mathbf{U}_Z^H \mathbf{U}_Z = \mathbf{I}_M$ ), $\mathbf{I}_M$ 为 $M$ 维单位矩阵。

信号子空间的协方差矩阵为实对称阵,因特征值和特征向量的秩都为 $K$ (声源个数),则信号子空间特征向量 $\mathbf{U}_S$ 和噪声子空间特征向量 $\mathbf{U}_n$ 可表示为

$$\mathbf{U}_S = [\mathbf{u}_1, \dots, \mathbf{u}_k], \quad (5)$$

$$\mathbf{U}_n = [\mathbf{u}_{k+1}, \mathbf{u}_{k+2}, \dots, \mathbf{u}_{k+M}]. \quad (6)$$

因为 $\mathbf{R}_Z \mathbf{U}_n = \mathbf{A} \mathbf{R}_S \mathbf{A}^H \mathbf{U}_n + \sigma^2 \mathbf{U}_n$ , $\sigma^2$ 为噪声方差,且 $\mathbf{R}_Z \mathbf{U}_n = \sigma^2 \mathbf{U}_n$ ,可得到

$$\mathbf{A} \mathbf{R}_S \mathbf{A}^H \mathbf{U}_n = \mathbf{0}, \quad (7)$$

式(7)中, $\mathbf{R}_S$ 满秩,非奇异,存在逆矩阵,因此当且仅当 $\mathbf{A}^H \mathbf{U}_n = \mathbf{0}$ 时,式(7)满足导向矩阵 $\mathbf{A}$ 中各列

向量与噪声子空间正交的要求,  $\mathbf{U}_n^H \mathbf{a}_k = \mathbf{0}$ 。当  $\mathbf{a}$  为扫描声源面非声源点到传声器阵列的导向向量时,  $\mathbf{U}_n^H \mathbf{a} \neq \mathbf{0}$ , 通过噪声子空间法得到 MUSIC 算法在第  $k$  个扫描声源点处的声功率相对值

$$y_k = \frac{1}{\mathbf{a}_k^H \mathbf{U}_n \mathbf{U}_n^H \mathbf{a}_k}, \quad (8)$$

式(8)中,  $\mathbf{a}_k e^{i2\pi f r_k/c}$  为扫描声源到测量阵元面的导向向量,  $r_k$  为扫描声源点  $k$  到传声器阵列的距离向量。当  $\mathbf{U}_n^H \mathbf{a}_k = \mathbf{0}$  时, MUSIC 算法在该扫描点输出值较大, 反之输出值较小, MUSIC 算法最大化了声源点处的输出值, 因此有较高的空间分辨率<sup>[13]</sup>。

## 1.2 基于 Group Lasso 的 MUSIC 算法

MUSIC 算法在高频声源定位分辨率及聚焦性较好, 能够清晰地反映声源位置, 但在中低频时出现分辨率低, 聚焦性能差的缺点, 无法满足应用要求。本文对其使用 Group Lasso 算法进行优化, 需要输入信息有待求量初始值  $\mathbf{y}$ 、测量值  $\mathbf{Z}$ 、导向矩阵  $\mathbf{A}$ 。Group Lasso 算法第一步需对  $\mathbf{y}$  和  $\mathbf{A}$  进行分组, 即  $\mathbf{y} = [\mathbf{y}_1^H, \mathbf{y}_2^H, \dots, \mathbf{y}_N^H]^H$  和  $\mathbf{A} = [\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_N]$ 。Group Lasso 代价函数为

$$J_\beta = \frac{1}{2} \left\| \mathbf{Z} - \sum_{n=1}^N \mathbf{A}_n \mathbf{y}_n \right\|_2^2 + \beta \sum_{n=1}^N \|\mathbf{y}_n\|_2, \quad (9)$$

式(9)中,  $\beta$  为惩罚项的正则化系数, 通过  $\beta$  来控制  $\mathbf{y}$  的稀疏程度。

在 Group Lasso 算法组间应用 L1 范数正则化, 导致在组水平上拥有稀疏性, 此为提高 MUSIC 算法中低频分辨率及聚焦程度的关键因素。在组内应用 L2 范数正则化, 使得组内数据平滑<sup>[14]</sup>。MUSIC 算法声源定位在声源点位置扫描结果往往会有较突出的信号值, 因此为了提高结果的分辨率及聚焦度, 需要在组内循环中加入阈值, 以截断 MUSIC 算法中低频环境产生的虚影, 该阈值通过动态范围给出, 可保证信号位置信息不被破坏, 且提高算法的分辨率及聚焦性能。

优化算法过程如下:

(1) 定义组间迭代次数  $i$ , 以式(8)中  $\mathbf{y}_k$  为初始优化目标。

(2) 确定分组数, 本文以扫描点数量为分组依据, 设定组长度,  $n = 1, 2, \dots, N, n$  为分组数。

(3) 在第  $n$  组计算时, 首先确定本组残差值

$$\mathbf{r}_n^{(i)} = \mathbf{Z} - \sum_{k=1, k \neq n}^N \mathbf{A}_k \mathbf{y}_k, \quad (10)$$

式(10)中,  $\mathbf{r}_n^{(i)}$  为第  $n$  组残差值,  $\mathbf{A}_k$ 、 $\mathbf{y}_k$  分别为第  $n$  组的导向向量和 MUSIC 算法输出值。并通过交叉验证确定  $\beta$  值, 作为判断该组进行组内循环的依据。

(4) 若  $\left\| \mathbf{A}_n^H \mathbf{r}_n^{(i)} \right\|_2 \leq \beta$ , 则该组输出值为零; 否则进入该组组内循环, 通过坐标梯度下降方法<sup>[15]</sup>, 计算输出值, 计算公式为

$$\mathbf{y}_k^{(i,j+1)} = \mathbf{y}_k^{(i,j)} - \alpha \left( \mathbf{A}_n^H \mathbf{A}_k \mathbf{y}_k^{(i,j)} + \frac{\beta \mathbf{y}_k^{(i,j)}}{\left\| \mathbf{y}_k^{(i,j)} \right\|_2} - \mathbf{A}_n^H \mathbf{r}_n^{(i)} \right), \quad (11)$$

式(11)中,  $\alpha$  为梯度值, 当前后两次迭代误差小于  $10^{-3}$  时停止第  $j$  次迭代, 并使用动态范围计算阈值来截断输出值  $\mathbf{y}_k$ 。动态范围  $DR = 20 \cdot \lg \left( \frac{p_{\max}}{p_{\min}} \right)$ ,  $p_{\max}$  为初始输入值的最大值, 本文  $p_{\min}$  为初始输入值的最小值。选择 0.1 倍的动态范围作为截断阈值  $P_c$ , 即

$$P_c = 10^{\frac{DR}{20}}. \quad (12)$$

当输出值小于  $P_c$  时输出为零, 否则该输出值保持不变。

(5) 组内循环结束, 输出该组  $\mathbf{y}_n^{(i)}$  值。

(6) 重复步骤(3)~(5), 当前后两次迭代误差小于  $10^{-3}$  时停止第  $i$  次迭代, 得到输出值  $\mathbf{y}$ 。

## 2 基于 Group Lasso 的 MUSIC 仿真

### 2.1 仿真设计

Group Lasso MUSIC 优化算法的目的是提高中低频环境下声源定位分辨率及聚焦性能, 针对此目的设计仿真实验。第一, 在低信噪比且信噪比恒定、扫描面位置与仿真声源面位置重合情况下, 分别在单源和多源时分析 Group Lasso MUSIC 算法在中低频对声源定位聚焦性能和分辨率的优化效果; 第二, 在低信噪比且信噪比恒定、声源为中低频环境下, 改变扫描面到声源面的距离  $d$ , 验证 Group Lasso MUSIC 算法对声源定位聚焦性能的优化效果。

仿真采用如图 1 所示的 30 阵元随机传声器阵列, 阵列半径为 0.15 m。图 2 为仿真示意图, 包括声

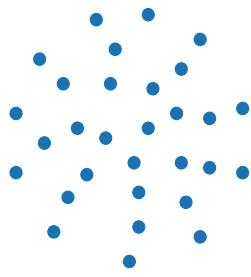


图1 随机传声器阵列阵元分布图

Fig. 1 Random microphone position map

源所在平面、扫描面、阵列面,其中,声源面到扫描面距离为 $d$ ,声源面到阵列面距离为 $D = 1\text{ m}$ 不变。距离 $d$ 是由于声源位置测量不准确导致的误差,在实际应用中,可能由于环境原因无法准确获取真实的距离 $D$ 。

## 2.2 仿真分析

### 2.2.1 仿真1

设置声源为点声源,位于原点位置,信噪比为5 dB,声源面到扫描面距离 $d = 0$ ,声源面到阵列面距离 $D = 1\text{ m}$ ,分析在中低频声源情况下,Group Lasso MUSIC算法的优化效果。单声源优化结果如图3所示,图3(a)、图3(c)、图3(e)分别为200 Hz、600 Hz、1000 Hz时MUSIC算法计算结果,从这三幅图中可看出,随着声源频率提高,声源位置的分辨程度逐渐提高,声源主瓣宽度逐渐缩小,聚焦性能呈

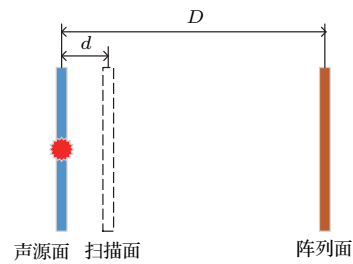


图2 仿真示意图

Fig. 2 Schematic diagram of simulation

增强趋势,但聚焦程度仍然低。

图3(b)、图3(d)、图3(f)为基于Group Lasso的MUSIC算法优化结果,从这三幅图中可以看出,优化后的MUSIC算法声源定位效果受频率影响极小。Group Lasso MUSIC算法经过对声源点稀疏筛选,聚焦性能较MUSIC算法有极大提高。

在单声源仿真基础上,将声源增加为双声源验证算法对多声源声场环境适应性及稳定性。图4(a)、图4(c)、图4(e)分别为双声源MUSIC算法200 Hz、600 Hz、1000 Hz的计算结果,相较于单声源,双声源MUSIC算法结果分辨率要差。经Group Lasso算法优化后,如图4(b)、图4(d)、图4(f),对于200 Hz仿真结果,稀疏程度并没有其他两者高,但主瓣宽度明显下降,分辨率显著提高。600 Hz与1000 Hz优化效果较好,分辨率提升明显,本文算法对中低频多声源环境有较好的适应性,效果稳定。

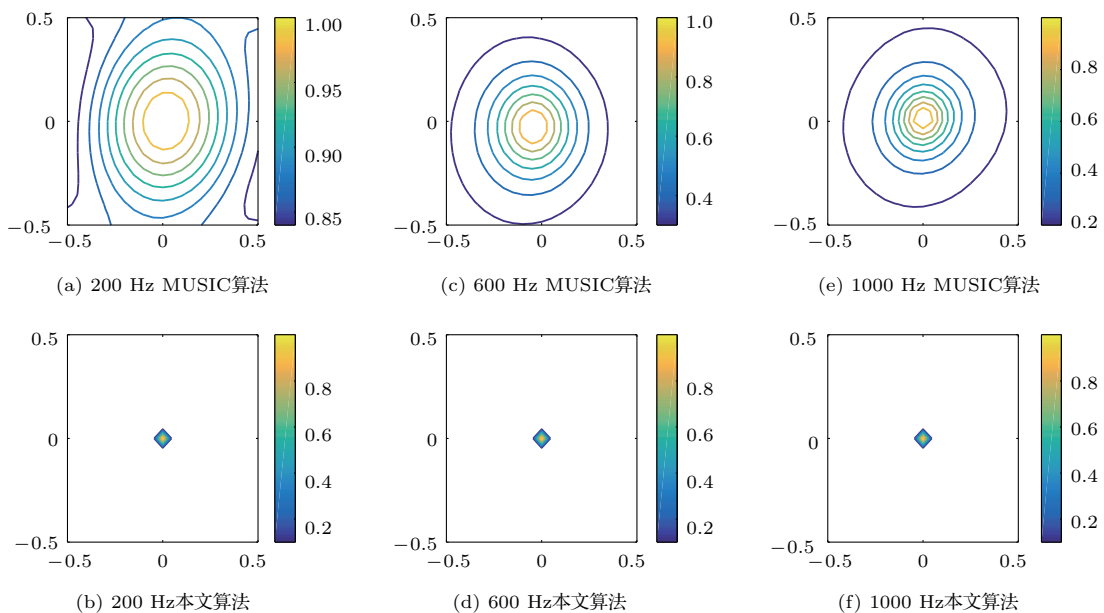


图3 仿真1 MUSIC算法与本文算法效果对比图

Fig. 3 Comparison of MUSIC algorithm in simulation 1 and algorithm in this paper

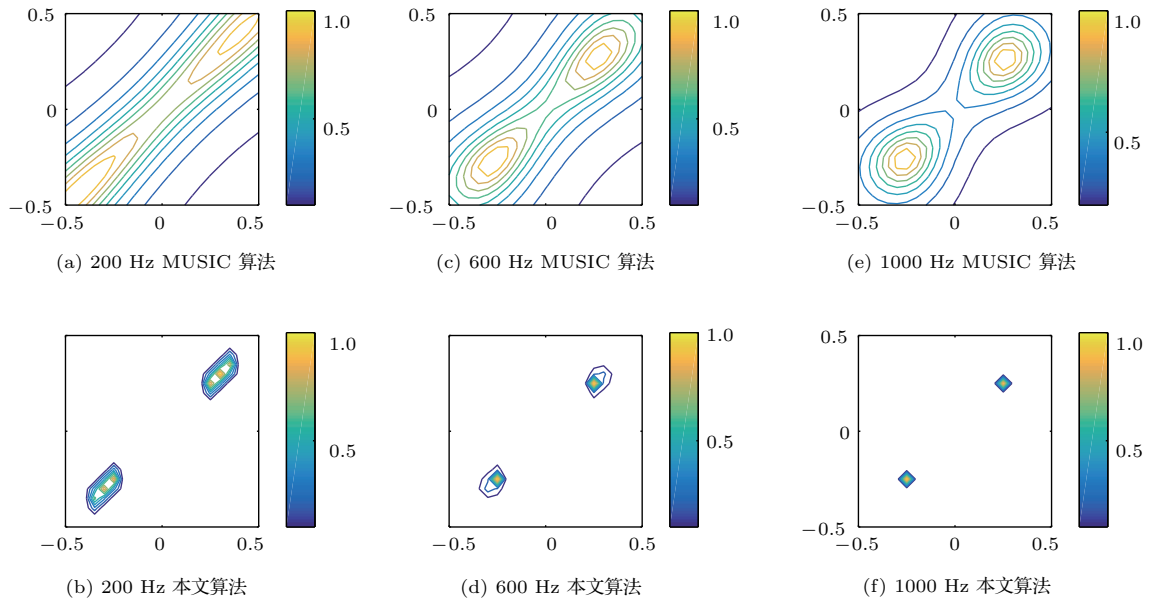


图4 仿真1双声源MUSIC算法与本文算法效果对比图

Fig. 4 Dual source comparison of MUSIC algorithm in simulation 1 and algorithm in this paper

2.2.2 仿真2

在仿真1单声源参数保持不变情况下，分析1000 Hz固定频率下、扫描面与声源面不重合(即  $d \neq 0$ )时 Group Lasso MUSIC算法优化效果。在多数实际测量中，无法准确获取阵列面到声源面距离，因此会存在误差，本文中误差量为  $d$ 。仿真2旨在说明随着误差  $d$  的变化，本文算法的稳定性。

如图5所示，图5(a)、图5(c)为MUSIC算法声源定位效果图，图5(a)中，聚焦面边缘出现轻微虚影，说明在  $d \neq 0$  情况下，MUSIC算法效果变差。

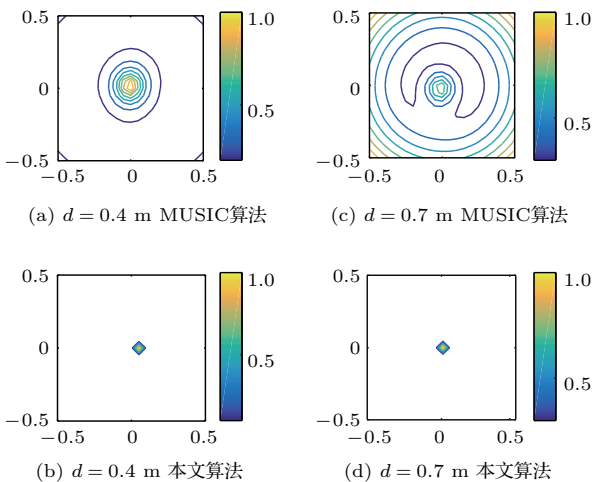


图5 仿真2 MUSIC算法与本文算法效果对比图

Fig. 5 Comparison of MUSIC algorithm in simulation 2 and algorithm in this paper

图5(c)中，当  $d = 0.7$  m时虚影值已超过声源值，对声源位置的判断造成了严重干扰，分辨率及聚焦性能极差。经 Group Lasso MUSIC算法优化后，如图5(b)、图5(d)所示，算法有效地抑制了虚影产生，并提高了聚焦性能及分辨率。

3 结论

Group Lasso MUSIC算法将传统MUSIC算法计算结果作为初始值，通过 Group Lasso算法，经组间稀疏、组内截断等处理，改进了MUSIC算法在中低频时，声源定位分辨率低、聚焦性能差的缺点。本文算法对于中低频环境单源和多源都有较好效果，对扫描面与声源面不重合时MUSIC算法出现的计算失真情况得到了有效控制。在低信噪比环境下，本文算法能够对声源准确定位，稳定性与适应性较高。

参 考 文 献

[1] van Veen B D, Buckley K M. Beamforming: a versatile approach to spatial filtering[J]. IEEE ASSP Magazine, 1988, 5(2): 4-24.  
 [2] 肖栋, 向阳, 卓瑞岩, 等. 基于波束形成的多类型多声源定位研究 [J]. 应用声学, 2017, 36(3): 220-227.  
 Xiao Dong, Xiang Yang, Zhuo Ruiyan, et al. Localization of multiple sound source with multi-type based

- on beamforming[J]. *Journal of Applied Acoustics*, 2017, 36(3): 220–227.
- [3] Schmidt R. Multiple emitter location and signal parameter estimation[J]. *IEEE Transactions on Antennas and Propagation*, 1986, 34(3): 276–280.
- [4] 居太亮. 基于麦克风阵列的声源定位算法研究[D]. 成都: 电子科技大学, 2006.
- [5] 居太亮, 邵怀宗, 彭启琮. 近场声源三维定位 MUSIC 算法研究[J]. *电子测量与仪器学报*, 2007, 21(1): 44–48.  
Ju Tailiang, Shao Huaizong, Peng qizong. Speech source 3D localization MUSIC algorithm in near field[J]. *Journal of Electronic Measurement and Instrument*, 2007, 21(1): 44–48.
- [6] Viberg M, Ottersten B, Kailath T. Detection and estimation in sensor arrays using weighted subspace fitting[J]. *IEEE Transactions on Signal Processing*, 1991, 39(11): 2436–2449.
- [7] 何子述, 黄振兴, 向敬成. 修正 MUSIC 算法对相关信号源的 DOA 估计性能[J]. *通信学报*, 2000, 21(10): 14–17.  
He Zishu, Huang Zhenxing, Xiang Jingcheng. The performance of DOA estimation for correlated signals by modified MUSIC algorithm[J]. *Journal on Communications*, 2000, 21(10): 14–17.
- [8] 石新智, 王高峰, 文必洋. 修正 MUSIC 算法对非线性阵列适用性的讨论[J]. *电子学报*, 2004, 32(1): 147–149.  
Shi Xinzhi, Wang Gaofeng, Wen Biyang. Discussion of the application for the monopole, Cross-Loops array based on MUSIC algorithm[J]. *Acta Electronica Sinica*, 2004, 32(1): 147–149.
- [9] Gardner W A. Signal interception: a unifying theoretical framework for feature detection[J]. *IEEE Transactions on Communications*, 1988, 36(8): 897–906
- [10] 黄知涛, 王炜华, 姜文利, 等. 一种基于循环互相关的非相干源信号方向估计方法[J]. *通信学报*, 2003, 24(2): 108–113.  
Huang Zhitao, Wang Weihua, Jiang Wenli, et al. A cyclic cross-correlation based direction-of-arrival estimation approach[J]. *Journal on Communications*, 2003, 24(2): 108–113.
- [11] 王超, 笄良龙, 韩梅, 等. 单矢量水听器的高分辨目标方位跟踪算法研究[J]. *应用声学*, 2017, 36(1): 59–66.  
Wang Chao, Da Lianglong, Han Mei, et al. High-resolution target azimuth tracking for a single vector hydrophone[J]. *Journal of Applied Acoustics*, 2017, 36(1): 59–66.
- [12] Simon N, Friedman J, Hastie T, et al. A sparse-Group Lasso[J]. *Journal of Computational and Graphical Statistics*, 2013, 22(2): 231–245.
- [13] Chu N. Bayesian approach in acoustic source localization and imaging[D]. University Paris Sud – Paris XI, 2013: 46–71.
- [14] Berry A, Grand jean P, Gauthier P A. Investigation of the group lasso algorithm for sound field reproduction: comparison with the lasso and elastic-net algorithms[C]. *International Congress on Sound & Vibration*, 2017: 23–27.
- [15] Lilis G N, Angelosante D, Giannakis G B. Sound field reproduction using the Lasso[J]. *IEEE Transactions on Audio, Speech, and Language Processing*, 2010, 18(8): 1902–1912.