

基于声压幅度比的多声源分离定位决策研究^①

杨 洋^{②*} 王秀芹^{**}

(* 渤海大学工学院 锦州 121013)

(** 渤海大学信息科学与技术学院 锦州 121013)

摘 要 利用声压幅度比模型,提出了一种基于声压幅度比的多声源分离定位方法,该方法利用盲信号分离算法实现混合声源信号的分离,根据谱估计的相似度确定接收信号中各声源的分配情况,结合幅度差异因子获得传感器的声源信号分布,再通过单声源的声压幅度比模型确定声源位置,实现多声源定位。由于盲信号分离算法比较成熟,且实际中的声源信号大多为非高斯,因而满足盲信号分离条件。该方法具有实用强、应用性广等特点,对其它分离、定位问题也有借鉴作用。

关键词 声压幅度比,多声源定位,盲信号分离(BSS),谱估计

0 引 言

随着计算机技术的发展和人机交互需求的增加,声源定位在实际领域日显重要,快速、精确的声源位置识别成为目标分离和语音识别的基础。目前相关技术已在语音通信、语音控制、移动机器人、助听装置、消防安监、远程视频会议、公共场所异常声源定位及滑坡破裂面追踪等方面发挥重要作用^[1-3]。

声源定位包括空间坐标位置和方向的定位,经典方法如时间到达(time of arrive, TOA)法、到达时间差或时延估计法^[4,7]、双耳互相关函数(binaural cross-correlation functions, BCCF)法^[8]、极大似然法^[9,10]、波动方程法^[11]、声音衰减模型法^[12,13]、头部反应传送函数(head related transfer function, HRTF)^[14]模型法和声压幅度比模型等^[15,16]等。在众多算法中,到达时间差或时延估计被广泛采用,通过估计声源到各个麦克风之间的时间差来估计声源位置,计算量小、实时性好,但定位依赖于时延估计

的精度,在多声源定位时性能下降^[17,18];双耳互相关函数法包括离线和在线两个阶段,离线阶段需要计算所有头相关脉冲响应的双耳互相关函数,因而需已知声源信号^[8];文献[9]提出的极大似然法可以针对近场和远场进行单声源定位,在多声源系统则性能较低,运算量偏大;波动方程法实质是求解一类非线性优化反问题,算法在声源个数未知时求解非常困难^[11];基于声音衰减模型的声源定位计算量小,但精度依赖于实际问题中声音传播衰减规律的建模,环境复杂时该模型鲁棒性难以保证;HRTF考虑人头等身体及环境特征解决声源定位问题,当精度要求高时,计算量大,且与个体关系密切,应用具有局限性;声压幅度比模型在实际中应用方便,它依据传感器接收信号的幅度与声源和传感器之间距离二者存在的约束条件实现,可以单独或与延时估计等模型融合实现声源定位,但无法直接扩展到多声源定位^[15]。此外,文献[19]利用基本广义交叉-相关相位变换(generalized cross-correlation phase transform, GCC-PHAT)方法结合真实约束条件,如传感器空间分布,房间大小等,当声源个数增加时,

① 国家自然科学基金(61401043)资助项目。
② 女,1984年生,博士生,讲师;研究方向:信号处理;联系人,E-mail: yangy195@163.com
(收稿日期:2016-04-06)

性能可能下降;文献[20]给出一种基于听觉机制的声源定位算法,并在多声源定位过程中能够不断修正声源定位方向,由于语音的复杂性,该算法对多声源定位时精度可能下降。从上述分析可知,实际中单声源定位方法非常成熟,但将其应用于多声源定位时仍可能存在局限性。声源的声强或声压信息较易通过传感器测量,因此,本文基于声压信息,给出一种受听觉机制启发的基于声压幅度比的多声源分离和定位算法,以解决声压幅度比模型无法应用于多声源定位问题。由于同一声源信号在不同接收传感器上强度和信号特征会有差异,仿照人耳对声源的辨识,通过盲信号分离(blind signal separation, BSS)算法分离传感器接收的声源信号,再利用谱估计的相似度估计分离声源信号的幅度信息,将多声源定位转化为单声源定位,从而根据声压幅度比及相应约束条件唯一地确定声源位置。本文既解决声源分离中盲信号分离算法的幅度不确定性,又解决声压幅度比模型应用于多声源定位问题,算法具有应用方便、扩展性强等特点。

1 声压幅度比模型实现单声源定位

声压幅度比模型实现单声源定位的原理是利用接收传感器的声压幅度与对应声源到传感器距离的关系及约束条件唯一地估计声源位置。设声源在第 i 个传感器位置产生的声压可用下式表示^[15]:

$$p_i(t) = \frac{r_0 p_0(t)}{r_i} e^{j(\omega t - k(r - r_0))} \quad (1)$$

其中 ω 为声源简谐振角频率, p_0 为声源在基准距离 r_0 处产生的声压, $k = \frac{\omega}{c}$, c 为声源的传播速度, 声压在传感器处产生的电压输出为 $e_i(t) = p_i(t) H_i \cos \theta_i$, H_i 参量由传感器的传递特性确定, θ_i 为声波在传感器处的入射角。经过推导后可以近似得出 $\frac{e_1(t)}{e_2(t)} = \frac{r_2^2}{r_1^2}$ ^[15]。即传感器测量的声压幅度比与声源到传感器距离的平方成反比。假设某声源有效声压为 E , E_1, E_2, E_3, E_4 分别为声压传感器接收到的有效声压, 声源到传感器的距离分别为 r_1, r_2, r_3, r_4 。由于 $E = r_4^2 E_4 = r_3^2 E_3 = r_2^2 E_2 = r_1^2 E_1$, 可得到下式:

$$\frac{r_2^2}{r_1^2} = \frac{E_1}{E_2}, \frac{r_3^2}{r_1^2} = \frac{E_1}{E_3}, \frac{r_4^2}{r_1^2} = \frac{E_1}{E_4} \quad (2)$$

其中 $\frac{E_1}{E_2}, \frac{E_1}{E_3}$ 和 $\frac{E_1}{E_4}$ 为声压幅度比即有效声压比, 由于传感器位置已知, 根据式(2)可估计出单声源位置。当存在多个声源且满足盲信号分离条件时, 可通过盲信号分离算法解决混合信号中声源信号分离问题。

2 盲信号分离算法

根据输入源信号的统计特性, 从众多混合信号中提取出未知源信号的过程, 称为盲信号分离, 又称为盲源分离或盲分离^[21]。该算法的假设条件为: 源信号分量相互统计独立且最多只有一个服从高斯分布; 混合系数具有足够的差异且矩阵可逆; 接收传感器个数不少于声源信号个数^[22]。传感器接收信号的瞬时混合模型如下式所示^[23]:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{v}(t) \quad (3)$$

$$\text{其中 } \mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pm} \end{bmatrix}, \mathbf{s}(t) = [s_1(t),$$

$s_2(t), \dots, s_m(t)]^T$, $s_j(t)$ 表示第 j 个声源信号, $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_p(t)]^T$, $x_i(t)$ 为第 i 个传感器接收到的信号, p 为传感器个数, $p \geq m$ 。 $a_{ij} (i = 1, 2, \dots, p, j = 1, 2, \dots, m)$ 表示第 j 个声源传输到第 i 个声压传感器时的衰减幅度, $v_i(t)$ 表示第 i 个传感器接收到的干扰及噪声, 不考虑噪声情况其数值为 0。盲源分离的基本思路是寻求分离矩阵 $\mathbf{W}$, 观测信号 $\mathbf{x}(t)$ 经过 $\mathbf{W}$ 变换后得到输出 $\mathbf{y}(t)$, 即

$$\mathbf{y}(t) = \mathbf{W}\mathbf{x}(t) = \mathbf{W}\mathbf{H}\mathbf{s}(t) \quad (4)$$

若 $\mathbf{W}\mathbf{H} = \mathbf{I}$ ($\mathbf{I}$ 为单位阵), 则 $\mathbf{y}(t)$ 为估计信号或分离信号, 从而实现恢复源信号的目标。

快速固定点法作为经典的盲源分离算法具有快速收敛的特点, 设声源信号的方差为 1, 不关心源信号之间的相对顺序, 对混叠矩阵 $\mathbf{A}$ 的估计可简化为对观测向量进行预白化, 经过线性变换后有

$$\mathbf{f}(t) = \mathbf{M}\mathbf{x}(t) = \mathbf{M}\mathbf{A}\mathbf{s}(t) = \mathbf{B}\mathbf{s}(t) \quad (5)$$

其中 $\mathbf{B} = \mathbf{M}\mathbf{A}$ 是一个正交矩阵, 由于 $E\{f(t)f(t)^T\} = \mathbf{B}\mathbf{B}^T = \mathbf{I}$, 将估计任意满秩 $\mathbf{A}$ 转换为估计矩阵 $\mathbf{B}$ 的问题, 源信号可由 $\mathbf{s}(t) = \mathbf{B}^T \mathbf{f}(t)$ 得到^[24, 25]。盲源分离问题存在排列顺序和信号幅度两个内在的不确定性, 由于信息主要包含在信号的波形中, 利用信号谱估计和信号重构可以有效解决幅度不确定问题。

3 多声源声压幅度比模型定位

3.1 信号谱估计

谱估计是一类从一组观测值中估计信号能量和功率分布方法的统称, 其谱曲线一般是功率或能量随频率变化的情况。为从混合信号中恢复声源, 并获得声压幅度比例关系, 本论文采用文献[26-28]中方法研究声源谱曲线的配比情况。

以参数模型为基础的谱估计方法步骤为^[26]: (1) 对被估计的随机过程选择合理模型; (2) 根据已知观测数据估计模型参数; (3) 通过估计得到的模型参数计算功率谱。对于差分方程 $y(n) = \sum_{i=1}^p a_i y(n-i) + e(n)$, 模型激励源 $e(n)$ 是均值为 0, 方差为 σ^2 的白噪声序列, 称为 p 阶自回归模型或简称为 $AR(p)$ 模型, 其传输函数为

$$H_{AR}(z) = \frac{1}{A(z)} = \frac{1}{1 - \sum_{i=1}^p a_i z^{-i}} \quad (6)$$

输出功率谱如下式所示^[27]:

$$S_{xx}(e^{j\omega}) = \frac{\sigma^2}{|1 + \sum_{k=1}^p a_k e^{-j\omega k}|^2} \quad (7)$$

a_k 为 AR 系数, 对 AR 模型参数估计的常用方法有自相关法、Burg 法和最小二乘法。

3.2 多声源定位

传感器接收混合声源信号后, 采用快速固定点算法获得声源估计信号。由于利用信号到达不同传感器的声压比例差异, 使得分离后信号的幅度信息非常重要。本文给出一种利用合成信号与实际传感器测量信号之间谱曲线的相似程度的量化方法。由于分离信号的配比不同, 合成信号的特征曲线会随

着发生变化。采用谱估计的相似度法估计接收传感器中不同声源的分配情况^[28], 其步骤为: (1) 将合成信号与传感器接收混合信号谱曲线相比较, 通过相似度模拟混合信号并找出幅度差异因子; (2) 利用幅度差异因子, 修正合成信号; (3) 利用修正的合成信号得出各声压传感器中每个声源估计信号, 即对分离信号按照一定比例缩放; (4) 分别计算每个声源到达各个传感器的声压幅度比。

以两声源为例, 通过盲信号分离算法得到两个声源估计信号。利用谱估计的相似度确定混合信号中两个声源的比例分配, 结合幅度差异因子估计出单个声源到达每个声压传感器的信号, 即恢复分离后信号的幅度信息, 进而得出每个声源到达两个不同传感器的声压幅度比, 根据每个声源到达不同声压传感器的声压比例分别进行位置估计。式(2)中的非线性方程直接求解比较困难, 遗传算法在解决非线性问题中表现出很好的鲁棒性和全局最优性^[23], 论文采用遗传算法优化声源空间位置, 目标函数即适应度函数为

$$fitness = \min \sum_{i=2}^4 |r_i^2 - \frac{E_1}{E_i} r_1^2| \quad (8)$$

进而找到满足条件的空间位置。

4 仿真实验

空间内有两个声源信号, 4 个接收传感器, 设空间坐标原点在正方形的中心, 传感器位置在正方形的顶点, 正方形边长为 2m, 则传感器空间坐标位置分别为 $(-1, 1, 0)$ 、 $(1, 1, 0)$ 、 $(-1, -1, 0)$ 和 $(1, -1, 0)$ 。混合矩阵 $\mathbf{A} = \begin{bmatrix} 0.3 & 0.8 & 0.6 & 0.2 \\ 0.9 & 0.4 & 0.2 & 0.5 \end{bmatrix}^T$, 不考虑噪声情况。图 1 为传感器接收的混合信号, 图 2 为利用快速固定点算法得到的分离信号。盲信号分离性能可以从相关系数结果中看出, 实验中分离信号 1 与原始声源 1 的相关系数为 0.9998, 分离信号 2 与原始声源 2 的相关系数为 1.0000。可见分离信号与原始信号具有高度的相似性, 但同时也在幅度和信号的顺序方面存在不确定性。将分离后的两个信号根据谱估计的相似度分析, 求得分离信号

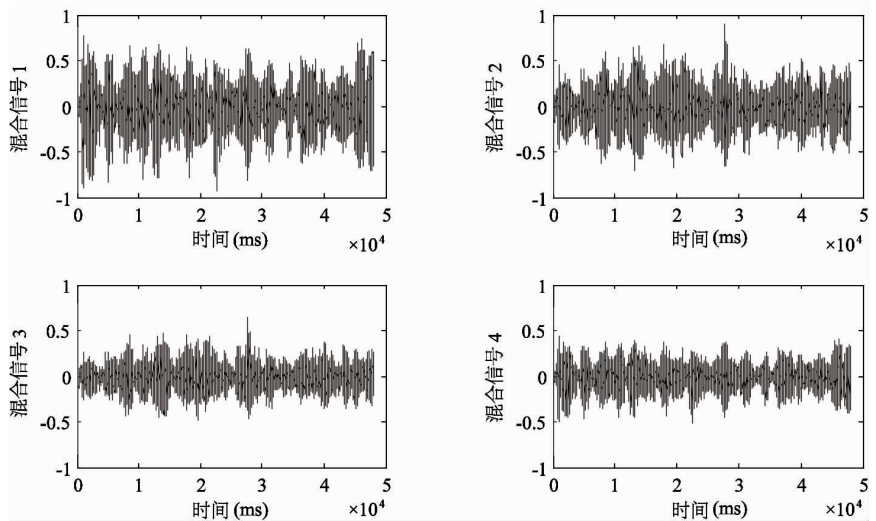


图 1 传感器接收到的混合信号

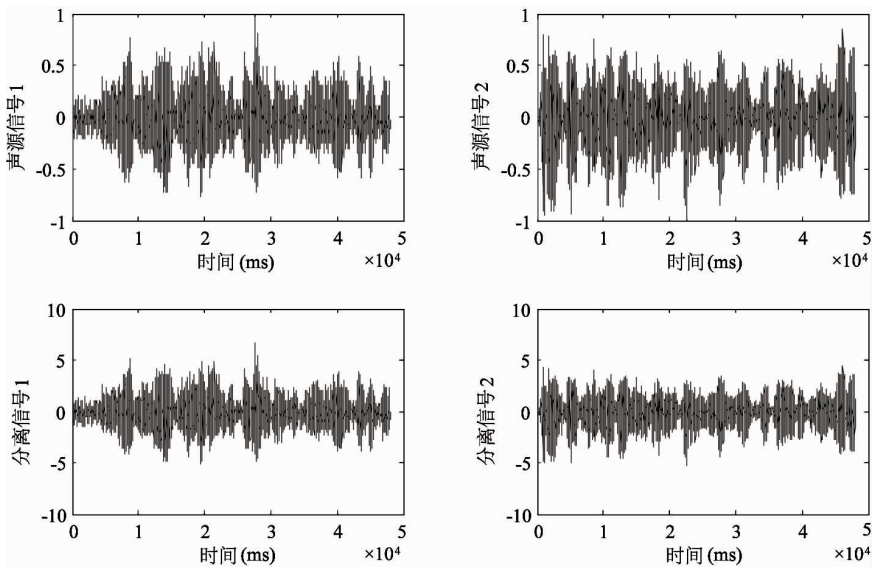


图 2 盲信号分离得到的分离信号

声压比例关系从 1 比 10 到 10 比 1 的实验结果,如图 3 所示。可以看出,关联度数值具有单调性,这也为求解实际信号中的混合比例提供了可靠的方法。将分离信号的声压比例关系图取值在一个小范围内,具体变化规律可见图 4 点线所示,虚线代表分离信号与传感器采集的声源信号谱曲线之间的相似程度。图 4 中,点线与虚线的交点代表实际传感器中的声源分离信号比例关系。

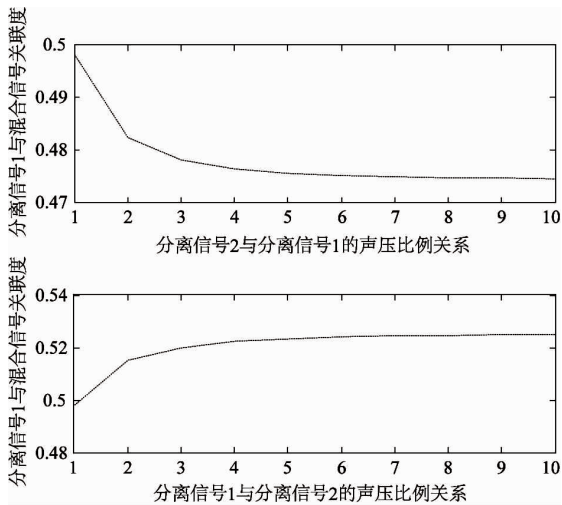


图 3 分离信号声压比例关系图

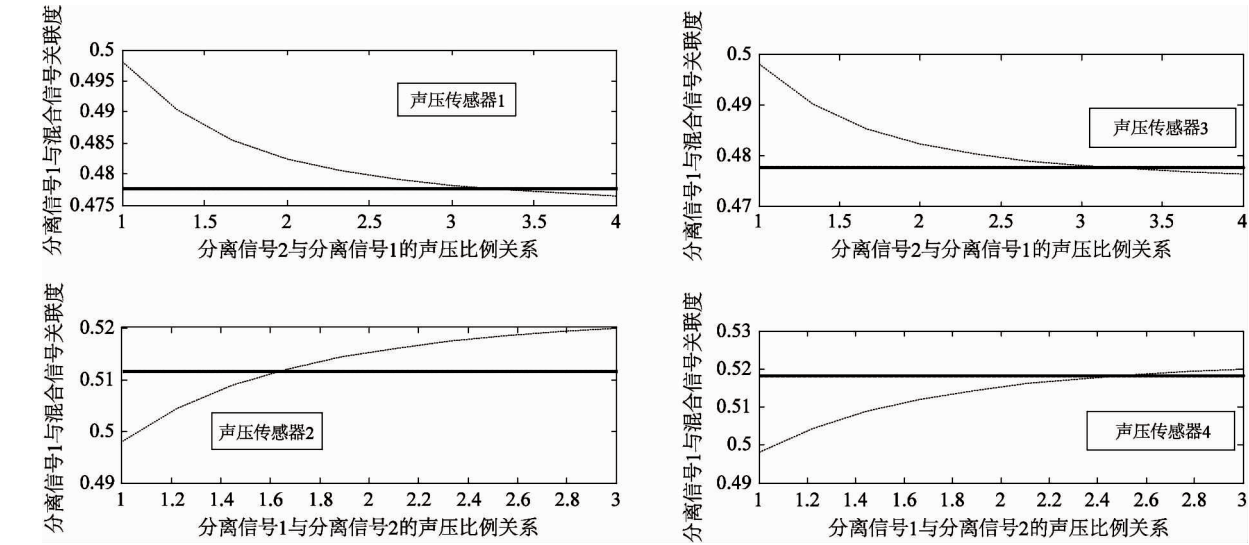


图 4 四个传感器下的分离信号的比例分配图

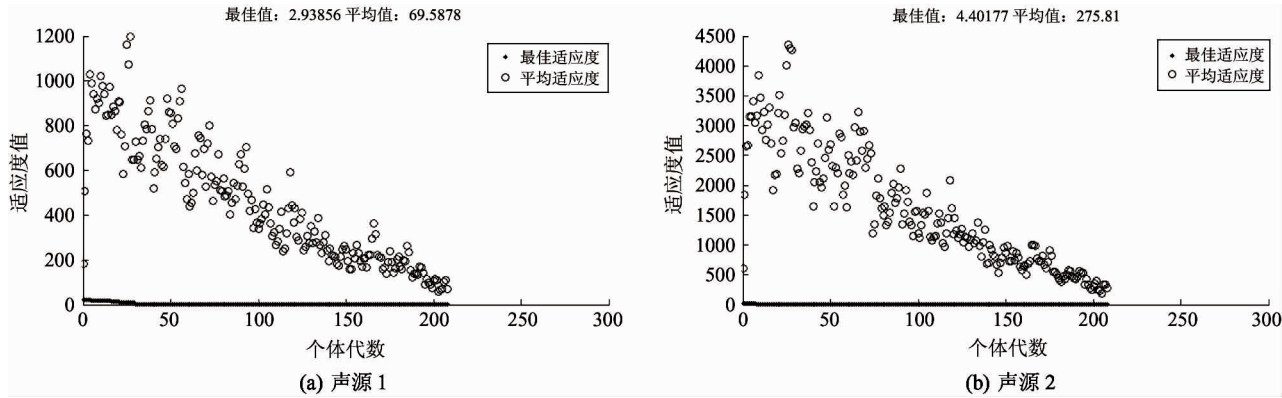


图 5 声源 1 和声源 2 的遗传算法适应度函数值与个体代数的关系图

表 1 为计算得到的两个声源在不同传感器中声压幅度比的计算结果,其中 E_{mi}/E_{mj} 表示估计得到的第 m 个声源信号在第 i 和 j 传感器处声压幅度比。根据式(2),利用遗传算法估计出每个声源位置,最

终得到 2 个声源位置坐标估计分别为 $(-0.1586, 0.0166, 0.1280)$, 和 $(0.0098, 0.9428, -0.0363)$ 。图 5 给出利用遗传算法估计声源位置的适应度函数值与个体代数的关系图。

表 1 声压幅度比的分析

声源	声压幅度比	仿真值	实际值	声源	声压幅度比	仿真值	实际值
分离信号 1	E_{11}/E_{12}	0.3809	0.3750	分离信号 2	E_{21}/E_{22}	2.3275	2.2500
分离信号 1	E_{11}/E_{13}	0.5078	0.5000	分离信号 2	E_{21}/E_{23}	4.7582	4.5000
分离信号 1	E_{11}/E_{14}	1.6261	1.5000	分离信号 2	E_{21}/E_{24}	1.7887	1.8000

5 结 论

本文给出了一种新的多声源分离与定位方法,

解决了原有声压幅度比模型无法应用于多声源定位以及盲信号分离算法中幅度不确定性两个问题。理论分析和仿真结果表明,该方法同其它声源定位算法相比,具有应用方便、计算量小和扩展性强等优

点,可广泛应用于多声源定位、盲信号辨识等问题中。由于分离信号具有顺序的不确定性,即使在没有原始声源信号先验信息的情况下,基于声压幅度比模型仍然可以求出的全部声源位置,但若只关心某一特定声源,受盲信号分离算法局限性,本文算法还没有突破这一问题。此外,随着盲信号分离技术的发展,已经出现很多关于信号源个数未知的分离算法,如基于特征值分解方法对声源数目进行估计,欠定盲信号分离问题等。因此,本文方法可以与其他定位算法融合解决多声源定位问题,更多改进算法还有待进一步研究。

参考文献

- [1] 张大威,鲍长春,夏丙寅. 复杂环境下基于时延估计的声源定位技术研究. 通信学报, 2014,1: 183-190
- [2] 李伟红,汤海兵,龚卫国. 公共场所异常声源定位中时延估计方法研究. 仪器仪表学报, 2012,33(4): 750-756
- [3] 潘忠晴,熊庆国,陈开端. 基于声源定位技术的滑坡破裂面追踪系统. 自动化与仪表, 2015,30(7): 81-84
- [4] 张毅,邢武超,罗元等. 基于耳蜗核模型改进双耳时间差的声源定位. 计算机应用, 2013, 33(11): 3280-3283
- [5] 王其华,郭戈. 基于到达时间差的半定松弛规划优化的定位算法. 大连海事大学学报, 2013,39(4): 59-62
- [6] 文小军,吉建华,钟林波等. 基于时延估计的声源定位算法实验研究. 计算机工程与应用, 2014, 50(23): 212-214, 233
- [7] 郭业才,朱赛男,张宁等. 自适应多通道双L型麦克风阵列声源定位算法. 系统仿真学报, 2015,27(7): 1556-1562
- [8] 万新旺,吴镇扬. 基于双耳互相关函数的声源定位算法. 东南大学学报(自然科学版), 2011,41(5): 903-906
- [9] Chen J C, Hudson R E, Yao K. Maximum-likelihood source localization and unknown sensor location estimation for wideband signals in the near-field. *IEEE Transactions on Signal Processing*, 2002, 50(8): 1843-1854
- [10] 居太亮,彭启琮,邵怀宗等. 基于麦克风阵列的近场声源定位子阵算法研究. 电子测量与仪器学报, 2006, 20(5): 50-55
- [11] 林娟. 基于布谷鸟算法的声源定位. 数学的实践与认识, 2015, 45(15): 230-236
- [12] 崔晓宗,全杰,刘弢等. 一种用于无线传感器网络的非线性最小二乘声源定位算法. 计算机研究与发展, 2011,48(s2): 65-70
- [13] 杨洋,王秀芹. 基于能量衰减的增量式声源跟踪算法研究. 计算机技术与发展, 2014, 12: 41-44
- [14] Argentieri S, Danes P, Soueres P. A survey on sound source localization in robotics: from binaural to array processing methods. *Computer Speech & Language*, 2015, 34(1): 87-112
- [15] 陈可,汪增福. 基于声压幅度比的声源定位. 计算机仿真, 2004,21(11): 85-88
- [16] 曹景敏,万群,魏合文等. 目标海拔已知的信号到达幅度比无源定位方法. 电子学报, 2016, 44(6): 1369-1375
- [17] 梁瑞宇,赵力,奚吉等. 基于听觉仿生和压缩感知的助听器声源定位方法. 仪器仪表学报, 2011, 32(6): 1390-1395
- [18] Athanasopoulos G, Verhelst W, Sahli H. Robust speaker localization for real-world robots. *Computer Speech & Language*, 2015, 34(1): 129-153
- [19] Brutti A, Omologo M. Multiple source localization based on acoustic map de-emphasis. *EURASIP Journal on Audio, Speech, and Music Processing*, 2009, 2010(1): 1-17
- [20] Huang J, Ohnishi N, Sugie N. A biomimetic system for localization and separation of multiple sound sources. *IEEE Transactions on Instrumentation and Measurement*, 1995, 44(3): 733-738
- [21] 刘琚,何振亚. 盲源分离和盲反卷积. 电子学报, 2002,30(4): 570-576
- [22] Zhou W L, Chelidze D. Blind source separation based vibration mode identification. *Mechanical Systems and Signal Processing*, 2007, 21(8): 3072-3087
- [23] Yang Y, Wang X, Zhang D. Blind source separation research based on the feature distance using evolutionary algorithms. *International Journal of Acoustics and Vibration*, 2014, 19(4): 276-286
- [24] Hyvärinen A, Oja E. A fast fixed-point algorithm for independent component analysis. *Neural computation*, 1997, 9(7): 1483-1492
- [25] 陈宏滨,冯久超,谢智刚. 传感器网络中的盲源分离与信号重构. 北京: 电子工业出版社, 2012. 20-80

[26] 杜丽萍,王双维,化晓茜等. 基于现代谱估计的噪声源在线识别方法研究. 噪声与振动控制, 2008, 28(1): 60-63

[27] 董浩, 迟学芬, 曲良东等. 现代互谱估计方法及仿真分析. 吉林大学学报(工), 2014, 44(5):1423-1428

[28] Yang Y, Wang S W, Hao N L, et al. On-line noise source identification based on the power spectrum estimation and grey relational analysis. *Applied Acoustics*, 2009, 70(3):493-497

Research on multi-sound source separation and localization
based on the pressure amplitude ratio

Yang Yang^{*}, Wang Xiuqin^{**}

(^{*} School of Engineering, Bohai University, Jinzhou 121013)

(^{**} School of Information Science and Technology, Bohai University, Jinzhou 121013)

Abstract

The sound pressure amplitude ratio model was applied to the study of multiple sound source localization, and a method for multi-sound source separation and localization based on the sound pressure amplitude ratio was proposed. The method uses the blind signal separation algorithm to separate mixed signals and uses the spectrum estimation similarity method to determine the distribution of each sound source in the mixed signals. Then, according to the amplitude difference factor, the method obtains the distribution of sound source signals of the sensor and uses the sound pressure amplitude ratio model for single sound sources to determine sound source positions to achieve the multi-sound source localization. The blind signal separation algorithm is relatively mature, and the actual sound source signals are mostly non-Gauss, so it satisfies the condition of blind signal separation. The proposed method is practical, and can be widely used to solve the similar problems in signal localization and separation fields.

Key words: sound-pressure amplitude ratio, multi-sound source localization, blind signal separation (BSS), spectrum estimation