

基于麦克风阵列的室内三维声源定位优化算法^①



黄 静, 胡馨月

(浙江理工大学 信息学院, 杭州 310018)

通讯作者: 黄 静, E-mail: syhj_sy@163.com

摘 要: 针对基于相位变换加权的可控响应功率 (Steered Response Power-PHase Transform, SRP-PHAT) 定位算法精度高但实时性差的问题, 本文引入基于到达时间差 (Time Difference Of Arrival, TDOA) 的定位算法以提高实时性, 提出一种基于 TDOA 和搜索空间聚类 (Search Space Clustering, SSC) 优化的 SRP-PHAT 的组合算法——搜索空间收缩聚类算法. 该算法先利用 TDOA 定位算法经过离群值校正后得到声源在方向角和径向距离上的估计范围, 之后根据估计声源范围进行搜索区域收缩, 最后利用 SRP-PHAT-SSC 算法在收缩区域内进行细粒度 (5 cm) 的空间搜索计算, 得到估计声源的三维坐标. 本文采用五元麦克风阵列, 利用虚源法模拟室内声场, 通过 Matlab 对声源进行了三维定位仿真. 实验结果表明, 改进后的算法与基于 SRP-PHAT 的全网格搜索 (Full Grid Search, FGS) 算法和 SSC 算法相比, 在三维定位上的实时性和鲁棒性都得到了提高.

关键词: 声源定位; 三维空间; TDOA; SRP-PHAT-SSC; 搜索空间收缩聚类; 实时性; 鲁棒性

引用格式: 黄静, 胡馨月. 基于麦克风阵列的室内三维声源定位优化算法. 计算机系统应用, 2021, 30(9): 212-218. <http://www.c-s-a.org.cn/1003-3254/8072.html>

Indoor 3D Sound Source Localization Optimization Algorithm Based on Microphone Array

HUANG Jing, HU Xin-Yue

(School of Information Science and Technology, Zhejiang Sci-Tech University, Hangzhou 310018, China)

Abstract: The Steered Response Power-PHase Transform (SRP-PHAT) localization algorithm has high accuracy but poor real-time performance. In response to this problem, this study introduces a localization algorithm based on Time Difference Of Arrival (TDOA) to improve real-time performance and then proposes a combination algorithm based on TDOA and Search Space Clustering (SSC), search space shrinking clustering, to optimize SRP-PHAT. First, the TDOA localization algorithm is used to estimate the range of the sound source in the direction angle and radial distance after outlier correction. Then, the search area is shrunk according to the estimated sound source range. Finally, fine-grained (5 cm) space search calculations in the shrinking area are performed by the SRP-PHAT-SSC algorithm to obtain the three-dimensional (3D) coordinates of the estimated sound source. A five-element microphone array and the virtual source method are employed to simulate the indoor sound field, and the 3D localization of the sound source is simulated by Matlab. The experimental results show that compared with the Full Grid Search (FGS) algorithm and the SSC algorithm based on SRP-PHAT, the improved algorithm has great real-time performance and robustness in 3D localization.

Key words: sound source localization; three-dimensional space; Time Difference Of Arrival (TDOA); Steered Response Power-PHase Transform-Search Space Clustering (SRP-PHAT-SSC); search space shrinking clustering; real-time performance; robustness

① 收稿时间: 2020-12-01; 修改时间: 2021-01-04, 2021-01-18; 采用时间: 2021-01-20; csa 在线出版时间: 2021-09-02

信号源的定位是传感器阵列信号处理领域重要部分,曾应用于雷达、声呐等大型通信设备中,到20世纪80年代,随着语音通信发展,基于麦克风阵列的语音信号处理也逐步受到学者们关注,在会议发言人定位、机器人听觉系统、安防等场景都得到了广泛的研究与应用^[1]。目前基于小型麦克风阵列的声源定位大多是二维层面,要实现三维定位则需要增加麦克风组成大规模麦克风阵列采集声源信息,增加了设备成本和计算成本。因此,如何利用小型麦克风阵列进行声源的三维实时定位成了重要的研究课题。

在目前的研究中,近场声源定位算法主要分为3类:(1)基于高分辨谱估计的定位算法。该算法主要是针对窄带信号,若要运用到语音信号中,需要将宽带的语音信号先分解为窄带信号,因转换损失会造成偏差^[2,3]。(2)基于可控响应功率的定位算法。该算法对宽带信号和窄带信号都适用,但需要对整个空间的网格点进行搜索和功率计算,计算量大导致响应速度慢,不适用于实时的声源定位系统^[4]。(3)基于到达时间差的定位算法。该算法分为两个部分,第1步是时延估计,第2步是定位估计,计算量小,实时性高,在时延估计阶段有较高要求,否则会在定位估计中会放大时延估计的误差^[5]。

针对SRP-PHAT算法响应速度慢的缺陷,很多学者提出了改进方法。Cho等^[6]提出了针对小型麦克风阵列的SSC加速算法,利用相同到达时间差聚类取质心的方法减少搜索网格,但仅对俯仰角和方向角进行了计算,缺少测距,且聚类后的质心数目仍然;Yook等^[7]在SSC优化算法的基础上,提出了针对大型麦克风阵列的两级搜索空间聚类(Two-Level Search Space Clustering, TL-SSC)算法,将声源的候选位置划分成组,找到少数可能包含最大功率的组,进一步减少SSC算法的搜索网格,但对于小型麦克风阵列的计算量减小效果不明显;Salvati等^[8]利用搜索网格上TDOA信息的非均匀积累提出了基于灵敏度的区域选择SRP(Region selection SRP, R-SRP)算法以减少声源定位误差,但计算量较多,实时性差。

本文介绍了阵列信号模型和传统定位算法,改进了TDOA定位算法以提高定位准确性,改进了SRP-PHAT-SSC定位算法以简化聚类逻辑,同时针对小型麦克风阵列三维定位算法实时性不足的问题,提出了TDOA定位算法与SRP-PHAT-SSC算法相结合的搜索空间收缩聚类算法,并在Matlab平台对3种SRP-

PHAT搜索方式进行了仿真比较。算法拟针对五元小型麦克风阵列,在SRP-PHAT-SSC加速算法的基础上进一步减少搜索网格数,提高SRP-PHAT算法的实时性和三维定位精度。

1 阵列信号模型的建立

阵列模型的建立是声学研究的基础。本文针对室内的声源定位,声场是近场模型,声波以球面波的形式扩散,图1为近场中一个声源到两个麦克风的传播示意图^[9],其中,Source表示声源;Mic₁和Mic₂表示两个麦克风; τ 为声源信号到达两个麦克风的时延,单位为s; c 为声速,单位为m/s;声源到两个麦克风的距离 d 为 $\tau \times c$ 。

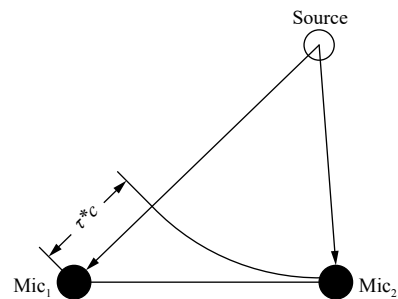


图1 近场声音传播示意图

根据图1,时延为 τ 的两个麦克风接收到的信号模型 $x_i(t)(i=1, 2)$ 可表示为^[10]:

$$x_i(t) = h_i * s(t - \tau_i) + n_i(t) \quad (1)$$

其中, $s(t)$ 为声源信号; h_i 为声源到麦克风 i 之间的脉冲响应函数,由式(2)表示; τ_i 为声源到达麦克风 i 的时间, $\tau = \tau_1 - \tau_2$; $n_i(t)$ 为第 i 个麦克风的环境噪声信号, $s(t)$ 、 $n_1(t)$ 、 $n_2(t)$ 之间互不相关。

$$h_i(r, r_s, t) = \sum_{p \in P} \sum_{l \in \Lambda} \beta_{x_1}^{|\lambda_x - q|} \beta_{x_2}^{|\lambda_x|} \beta_{y_1}^{|\lambda_y - j|} \beta_{y_2}^{|\lambda_y|} \beta_{z_1}^{|\lambda_z - k|} \beta_z^{|\lambda_z|} \frac{\delta(t - R/c)}{4\pi R} \quad (2)$$

其中, r 表示麦克风 i 的位置坐标; r_s 表示声源位置坐标; t 表示时间; c 为声速; β 表示一个房间6面墙的反射系数; R 表示虚源的位置坐标; $\delta\{\cdot\}$ 表示虚源 R 到麦克风 i 的直达声压,可替代为使用Hanning窗的理想低通滤波器,窗长为4ms,截止频率为奈奎斯特频率; $P = \{(q, j, k) \mid q, j, k \in \{0, 1\}\}$; $A = \{(\lambda_x, \lambda_y, \lambda_z) \mid \lambda_x, \lambda_y, \lambda_z \in Z: -N \leq \lambda_x, \lambda_y, \lambda_z \leq N\}$, N 表示虚源的级数; $R = ((1-2q)x_s + 2\lambda_x L_x - x, (1-2j)y_s + 2\lambda_y L_y - y, (1-2k)z_s + 2\lambda_z L_z - z)$ 。

在实际使用中,一般通过设定混响时间 RT_{60} 来改

变声源和麦克风 i 之间的脉冲响应, 根据 Sabin 经验公式可以由混响时间求出式 (2) 中的墙壁反射系数 β . 混响时间和墙壁反射系数的关系为:

$$RT_{60} = \frac{24 \ln(10) V}{c \sum_{i=1}^6 S_i (1 - \beta_i^2)} \quad (3)$$

其中, V 表示房间的体积, S_i 表示第 i 面墙的面积.

本文使用一个五元全向型麦克风阵列进行仿真, 以一个中型大小的会议室为场景, 房间大小为 $9 \text{ m} \times 8 \text{ m} \times 3 \text{ m}$; 5 个麦克风的位置为 $M_1(2, 2, 1.6)$ 、 $M_2(2, 1.9, 1.5)$ 、 $M_3(2, 2.1, 1.5)$ 、 $M_4(1.9, 2, 1.5)$ 、 $M_5(2.1, 2, 1.5)$, 阵列结构如图 2 所示. 声速 c 设定为 340 m/s ; 声源使用 Noizeus 中一段 3 s 的纯净男声, 采样频率为 8 kHz ; 声源坐标设为 $(1.6, 2.3, 1.7)$, 在 Matlab 平台根据式 (1) 模拟麦克风接收到的信号, 以加性高斯噪声作为环境噪声. 对麦克风模拟信号进行带通滤波和分帧加窗预处理, 窗函数选择典型的 Hamming 窗, 帧长取 256 点, 帧移为 50%.

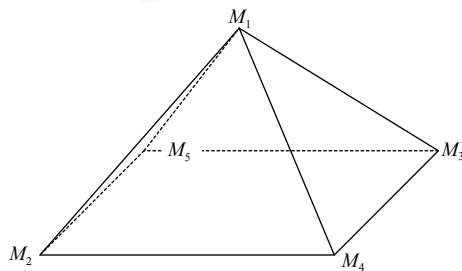


图 2 麦克风阵列结构模型

2 定位算法的分析与改进

2.1 TDOA 定位算法

2.1.1 广义互相关法时延估计

TDOA 定位算法的第一步是时延估计, 广义互相关法 (Generalized Cross-Correlation, GCC) 是一种最常用的时延估计算法^[11,12]. 通过最大化一对麦克风信号在频域上的互相关函数可估计出声音信号到达这对麦克风的时间差^[5].

信号 x_1 与 x_2 之间的时延 $\tau_{x_1 x_2}$ 定义为:

$$\tau_{x_1 x_2} = \underset{\omega}{\operatorname{argmax}} R_{x_1 x_2}(\tau) \quad (4)$$

信号 x_1 与 x_2 之间的广义互相关函数 $R_{x_1 x_2}(\tau)$ 为:

$$R_{x_1 x_2}(\tau) = \int_{-\infty}^{\infty} G_{x_1 x_2}(\omega) e^{j\omega\tau} d\omega \quad (5)$$

$$G_{x_1 x_2}(\omega) = X_1(\omega) X_2^*(\omega) \quad (6)$$

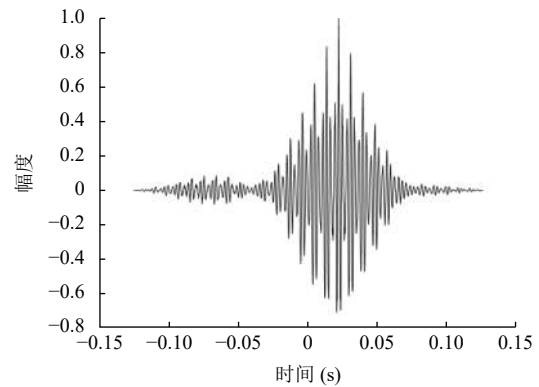
其中, ω 为频率; $j = -2\pi i$; $G_{x_1 x_2}(\omega)$ 为 x_1 与 x_2 的互功率谱; $X_1(\omega)$ 与 $X_2(\omega)$ 分别是 x_1 与 x_2 经过快速傅里叶变换 (FFT) 后的频域信号; $(*)$ 表示复共轭运算.

由于混响和噪声容易对广义互相关函数产生影响, 为了减少混响和噪声的影响, 本文引入 PHAT (PHase Transform, 相位变换) 加权函数对互功率谱函数作白化处理, PHAT 加权对混响有一定的抑制效果^[13,14]. 采用一对麦克风进行时延估计仿真, GCC-PHAT 函数与 GCC 函数的对比如图 3 所示, GCC-PHAT 函数只保留了相位信息, 从而锐化互相关函数的峰值. PHAT 加权函数的公式为:

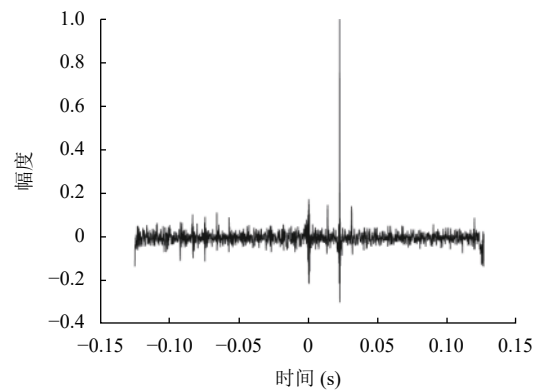
$$\phi_{\text{PHAT}} = \frac{1}{|G_{x_1 x_2}(\omega)|} \quad (7)$$

则式 (5) 引入 PHAT 加权可写为:

$$R_{x_1 x_2}(\tau) = \int_{-\infty}^{\infty} \frac{G_{x_1 x_2}(\omega)}{|G_{x_1 x_2}(\omega)|} e^{j\omega\tau} d\omega \quad (8)$$



(a) GCC



(b) GCC-PHAT

图 3 信噪比 0 dB、混响时间 0 s 时, GCC 与 GCC-PHAT 的比较

2.1.2 球形插值法定位估计

定位估计是 TDOA 定位算法的第 2 步, 为了保证 TDOA 定位算法的实时性, 本文选用球形插值法 (Spherical Interpolation, SI) 用于定位估计. 球形插值法属于最小二乘 (Least Square, LS) 算法, 复杂度低, 基本思想是以麦克风阵列中的一个麦克风为参考麦克风, 通过最小化声源到参考麦克风与其他麦克风之间距离差的估计值与实际值的误差平方和, 求出声源位置^[15,16].

假设有 N 个麦克风的坐标 (M_1, M_2, \dots, M_N) , 若以 M_1 为参考麦克风, 并作为空间坐标原点, 此时, M_i 的坐标为 q_i , 声源 s 到 M_i 与 M_1 的估计距离差为 \hat{d}_i , \hat{d}_i 是根据第一步的估计时延 τ 得到的值, M_i 到坐标原点的距离为 R_i , 声源位置坐标为 q_s , 则 M_i 到原点距离的平方可表示为:

$$d_i^2 = (R_s + \hat{d}_i)^2 = \|q_i - q_s\|^2 = R_i^2 + R_s^2 - 2q_i^T q_s.$$

由于估计值 \hat{d}_i 存在误差, 所以估计距离与实际距离的最小二乘误差函数可写为:

$$\varepsilon = 2R_s \hat{D} + 2Mq_s - \delta \quad (9)$$

$$\text{其中, } M = \begin{bmatrix} x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \\ \vdots & \vdots & \vdots \\ x_N & y_N & z_N \end{bmatrix}, \hat{D} = \begin{bmatrix} \hat{d}_2 \\ \hat{d}_3 \\ \vdots \\ \hat{d}_N \end{bmatrix}, \delta = \begin{bmatrix} R_2^2 - \hat{d}_2^2 \\ R_3^2 - \hat{d}_3^2 \\ \vdots \\ R_N^2 - \hat{d}_N^2 \end{bmatrix}.$$

式 (9) 可简化为:

$$\varepsilon = A\mu - b \quad (10)$$

$$\text{其中, } A = [M\hat{D}], \mu = \begin{bmatrix} q_s \\ R_s \end{bmatrix}, b = \frac{1}{2}\delta.$$

则式 (10) 的最小二乘解可表示为:

$$\hat{\mu} = (A^T A)^{-1} A^T b \quad (11)$$

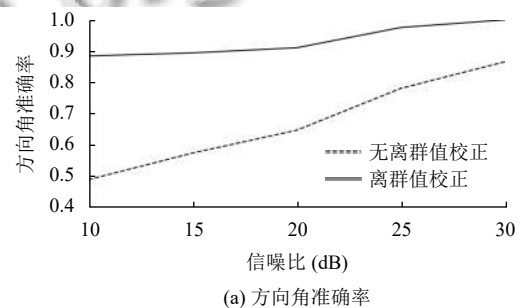
定义投影矩阵 $P_d = I - \frac{\hat{D}\hat{D}^T}{\hat{D}^T\hat{D}}$, 代入式 (11) 可得估计位置:

$$\hat{q}_s = (M^T P_d M)^{-1} M^T P_d b \quad (12)$$

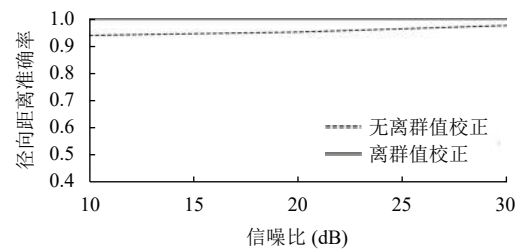
若估计声源的位置与实际位置中的方向角 φ 相差 $\Delta\varphi$ 度、俯仰角 θ 相差 $\Delta\theta$ 度、径向距离 r 相差 Δr 米, 则记为异常值, 否则记为正确值. 本文使用准确率作为评判准则: 准确率=正确值的帧数和/总帧数.

由于 GCC-PHAT 算法的时延估计误差会在球形插值法中放大, 定位错误率较高, 所以本文引入离群值校正, 对基于 TDOA 算法的定位结果进行校正处理, 将离群值校正为中位数, 其中, 离群值为中位数绝对偏差 (Median Absolute Deviation, MAD) 与 MAD 中位数相差超过 3 倍的值.

设置方向角、俯仰角和径向距离的容错范围分别是 $\Delta\varphi=15^\circ$ 、 $\Delta\theta=15^\circ$ 、 $\Delta r=1$. 在 5 元锥形麦克风阵列的情况下, TDOA 定位算法无法准确计算出俯仰角, 经过仿真实验得知, 俯仰角的准确率均在 20% 以下. 当无混响影响时, 如图 4 所示, 方向角和径向距离准确率随信噪比的减少而下降, 定位结果容易受到噪声影响, 做了离群值校正后, TDOA 定位算法的鲁棒性增强, 准确率都在 85% 以上; 当只有混响影响时, 由于 PHAT 加权函数对混响的抑制作用, 所以定位结果都在容错范围内, TDOA 在方向角和径向距离两个方面的准确率均为 100%, 不受混响时间的影响.



(a) 方向角准确率



(b) 径向距离准确率

图4 信噪比对 TDOA 定位算法的影响

2.2 SRP-PHAT 定位算法

SRP-PHAT 算法属于一步定位的算法, 其 PHAT 滤波被用于提高 SRP 算法鲁棒性^[17]. 通过对定位空间划分格子, 每一个格子进行所有麦克风对之间互相关函数的求和来精确功率的计算, 从而确定最大功率的空间点作为估计声源的位置^[18].

定义 q 为空间点位置坐标, $R_{lm}[\tau_{lm}(q)]$ 为麦克风 l 和麦克风 m 所接收信号的 GCC-PHAT 函数, 则可控响应功率 $P(q)$ 为:

$$P(q) = \sum_{l=1}^N \sum_{m=l+1}^N R_{lm}[\tau_{lm}(q)] \quad (13)$$

那么, 估计声源的位置就在可控响应功率值最大的空间网格:

$$\hat{q}_s = \underset{q}{\operatorname{argmax}} P(q) \quad (14)$$

可控响应功率取决于阵列中每个麦克风对的接收信号在频域上的相位差, 而频谱的相位差在时域上是离散的, 所以如果有两个空间点的位置很接近, 则这两个空间点到各个麦克风对的时间差相同, 那么就无法依靠式 (14) 求出声源位置.

SSC 算法是一种针对基于 SRP-PHAT 的小型麦克风阵列声源定位而提出的加速算法^[6]. SSC 算法将具有相同 TDOA 的空间点集合在一起形成集群, 然后求出每个集群的质心, 将每个集群的质心作为代表坐标存储在一个查找表中, 从而避免了上述问题的出现, 同时进行声源定位时大大减少了功率计算的时间, 因为对于小型麦克风阵列, 最终组成的集群的数目远小于初始的空间网格总数.

文献 [6] 中使用的是自顶向下的聚类方法构造查找表, 本文使用逻辑与实现更为简单的自底向上聚类方法, 其算法流程如算法 1.

算法 1. 自底向上聚类算法

- 1) 以球面坐标 (φ, θ, r) 划分空间网格, 形成网格点坐标集合 $\{b\}$;
- 2) 遍历 $\{b\}$, 计算每个网格点中每组麦克风对的 TDOA, 存储到 TDOA 查找表;
- 3) 将相同 TDOA 值的网格点划分成簇;
- 4) 计算每个簇的质心, 形成质心集合 $\{c\}$;
- 5) 遍历 $\{c\}$, 计算每个集群质心的 TDOA, 存储到 TDOA_SSC 查找表;
- 6) 计算每个质心的可控响应功率;
- 7) 功率最大的质心位置为估计声源位置.

2.3 搜索空间收缩聚类算法

本文改进后的 TDOA 算法提高了定位计算的鲁棒性, 在五元锥形麦克风阵列且信噪比大于 10 dB 的情况下, 可在一定容错范围内得到正确的方向角和径向距离, 但无法计算得到俯仰角.

SRP-PHAT-SSC 算法需要对室内空间的每个集群质心进行可控功率计算, 数量相比 SRP-PHAT 的全网格搜索虽然已经减少, 但算法的实时性仍然不高, 可进一步缩减搜索网格数目以减少可控响应功率的计算量, 提高算法实时性.

因此, 针对小型麦克风阵列的三维定位, 本文提出一种基于 TDOA 和 SRP-PHAT-SSC 的组合算法——搜索空间收缩聚类算法 (Search Space Shrinking Clustering, SSSC), 通过引入改进的 TDOA 算法以进一步减少 SRP-PHAT-SSC 算法中的候选声源位置, 增加算法

的鲁棒性并提高实时性. 该算法主要分为两个阶段, 粗定位阶段和细定位阶段, 整体流程如图 5 所示.

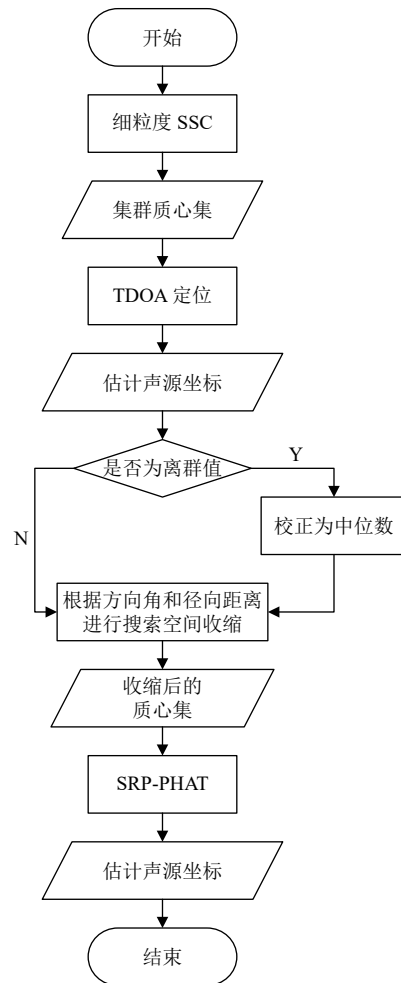


图 5 搜索空间收缩聚类算法流程

首先使用自底向上聚类的 SSC 算法得到整个空间的集群质心位置集合 B 并保存, 用于提供网格查询的功能, 采取 5 cm 的细粒度网格以提高 SRP-PHAT-SSC 算法的定位精度. 首先是粗定位阶段, 对音频信号帧进行基于 TDOA 的定位计算, 先利用式 (8) 求得估计时延, 再利用式 (12) 求得估计声源坐标, 得到估计声源位置集合 Q_s , 对 Q_s 进行离群值校正以提高 TDOA 定位的鲁棒性, 由于基于五元锥形麦克风阵列的 TDOA 定位算法无法准确判断俯仰角, 所以取其方向角 $\hat{\varphi}_{\text{TDOA}}$ 和径向距离 \hat{r}_{TDOA} 作为空间收缩的参考. TDOA 定位的容错范围设为 $\Delta\varphi=15$ 和 $\Delta r=1$, 根据设置的 TDOA 定位容错范围进行搜索空间收缩以减少可控功率的计算量, 收缩后的空间 B_s 为 $\{(\varphi, \theta, r) | \varphi \in [\hat{\varphi}_{\text{TDOA}} - \Delta\varphi_{\text{TDOA}}, \hat{\varphi}_{\text{TDOA}} + \Delta\varphi_{\text{TDOA}}], \theta \in B_\theta, r \in [\hat{r}_{\text{TDOA}} - \Delta r_{\text{TDOA}}, \hat{r}_{\text{TDOA}} + \Delta r_{\text{TDOA}}]\}$. 之

后进入细定位阶段, 根据 B_s 对集群质心坐标集合 B 进行筛选, 仅对收缩空间 B_s 内的集群质心使用式 (13) 进行可控响应功率计算, 之后根据式 (14) 取功率最大的质心位置作为声源。

3 仿真实验与结果分析

本文仿真实验中, SRP-PHAT 算法的全网格搜索方法和搜索空间聚类方法的网格尺寸分别取径向距离的精度为 20 cm、10 cm 和 5 cm, 搜索空间收缩聚类算法取 5 cm 的细粒度网格尺寸, 俯仰角和方向角的精度为 1° , 设置 $\Delta\varphi=10$ 、 $\Delta\theta=10$ 、 $\Delta r=1$, 若最终估计声源位置的球坐标在该设定范围内, 则标记为正确定位。

表 1 以信噪比为 30 dB, 混响时间为 0 s 的情况, 对 3 种算法进行了对比, 根据表 1 可以看出: 本文提出的 SSSC 算法搜索的格子数相比 SSC 算法减少了 96%, 从而使得计算时间比 SSC 算法减少了 97%, 具有较好的实时性, 且定位准确率达 95%。

图 6 和图 7 是对 3 种算法在球坐标系的 3 个维度

上分别进行准确率判断, 其中, FGS 和 SSC 算法选取与搜索空间收缩聚类算法相同网格粒度的数据作为展示。可以看出, 由于 PHAT 加权函数对混响的抑制作用, 3 个算法的准确率在仅有混响的情况下几乎不变。SSC 算法无法较好的估计出声源的径向距离, 改进后的算法通过引入 TDOA 算法计算出准确的径向距离, 从而弥补了 SSC 算法的缺点。在低信噪比的情况下, FGS 和 SSC 算法的准确率降低, 而改进后的算法在 3 个维度上的准确率都保持在 80% 以上, 这是因为其在 TDOA 定位阶段进行了离群值校正, 从而提高了鲁棒性。

表 1 不同 SRP-PHAT 算法的仿真结果对比

指标	全网格搜索(cm)			搜索空间聚类(cm)			搜索空间收缩聚类
	20	10	5	20	10	5	
定位准确率	0.6	0.65	0.66	0.51	0.53	0.54	0.95
搜索网格数	1501227	3002454	6004908	1885	2048	2161	70
运行时间(s)	77011	140621	269331	90	104	107	3.1

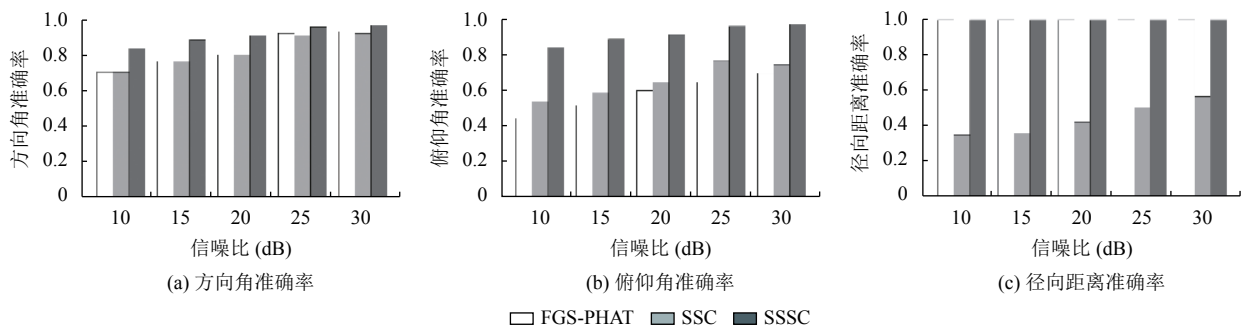


图 6 不同信噪比下 3 种算法的定位准确率

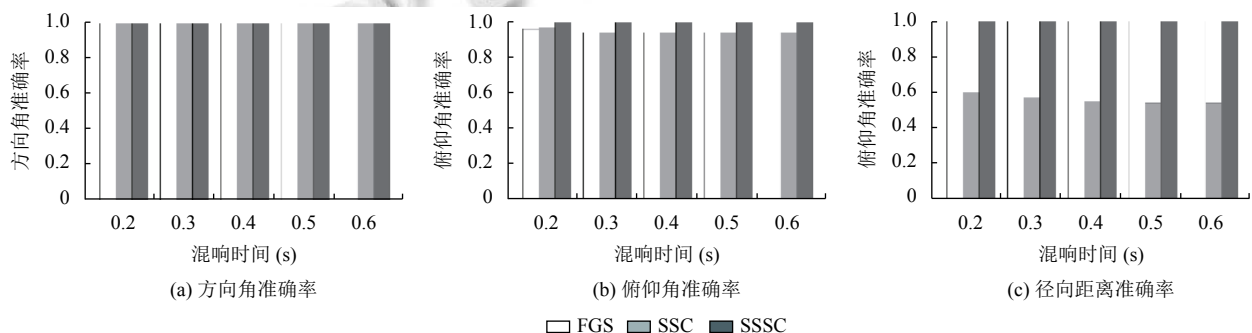


图 7 不同混响时间下 3 种算法的定位准确率

4 结论

本文分析了基于 TDOA 的声源定位算法和 SRP-

PHAT 声源定位算法, 在二者的基础上提出了基于 SRP-PHAT 的搜索空间收缩聚类优化算法, 本文选取了不

同的网格步长、信噪比和混响时间,对全空间搜索、搜索空间聚类 and 搜索空间聚类算法进行了对比分析,从实验结果可以看出搜索空间聚类算法减少了需要进行功率计算的网格数,从而减少了运行时间,具有较好的实时性,并且在球坐标的3个维度都有着高准确率,在实际应用中有一定的价值。

参考文献

- 1 王汀洲,朱梦尧,吴修坤.基于麦克风阵列的不一致性的校正.复旦学报(自然科学版),2018,57(3):344-349+357.
- 2 居太亮,彭启琮,邵怀宗.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.
- 3 郭业才,王超.基于支持向量机可控功率响应的 MUSIC-DOA 估计方法.南京理工大学学报,2019,43(2):237-243.
- 4 周峰,陈雄.基于卡尔曼滤波和预测的可控波束多声源跟踪.微电子学与计算机,2009,26(5):204-208.
- 5 茅惠达,张玲华.声源定位中广义互相关时延估计算法的研究.计算机工程与应用,2016,52(22):138-142. [doi: 10.3778/j.issn.1002-8331.1501-0090]
- 6 Cho Y, Yook D, Chang S, *et al.* Sound source localization for robot auditory systems. IEEE Transactions on Consumer Electronics, 2009, 55(3): 1663-1668. [doi: 10.1109/TCE.2009.5278040]
- 7 Yook D, Lee T, Cho Y. Fast sound source localization using two-level search space clustering. IEEE Transactions on Cybernetics, 2016, 46(1): 20-26. [doi: 10.1109/TCYB.2015.2391252]
- 8 Salvati D, Drioli C, Foresti GL. Sensitivity-based region selection in the steered response power algorithm. Signal Processing, 2018, 153: 1-10. [doi: 10.1016/j.sigpro.2018.07.002]
- 9 曹洁,吴尧帅,李伟,等.基于改进 TDOA 的近场声源鲁棒定位方法研究.计算机应用与软件,2018,35(4):102-108. [doi: 10.3969/j.issn.1000-386x.2018.04.019]
- 10 马晓红,陆晓燕,殷福亮.改进的互功率谱相位时延估计方法.电子与信息学报,2004,26(1):53-59.
- 11 余文晶,何琳,崔立林,等.声源定位中的时延估计方法研究进展.2016年度声学技术学术会议论文集.武汉,中国,2016.232-241.
- 12 郭培培,李建良.音频无人机定位的时延估计模拟分析.声学技术,2020,39(5):650-654.
- 13 金中薇,姜明顺,隋青美,等.基于广义互相关时延估计算法的声发射定位技术.传感技术学报,2013,26(11):1513-1518. [doi: 10.3969/j.issn.1004-1699.2013.11.009]
- 14 Knapp C, Carter G. The generalized correlation method for estimation of time delay. IEEE Transactions on Acoustics, Speech, and Signal Processing, 1976, 24(4): 320-327. [doi: 10.1109/TASSP.1976.1162830]
- 15 杨祥清,汪增福.基于麦克风阵列的三维声源定位算法及其实现.声学技术,2008,27(2):260-265.
- 16 于振华,付晓,王静,等.基于声学无线传感器网络的目标跟踪系统研究.电子科技大学学报,2011,40(4):568-572. [doi: 10.3969/j.issn.1001-0548.2011.04.019]
- 17 Velasco J, Martín-Arguedas CJ, Macias-Guarasa J, *et al.* Proposal and validation of an analytical generative model of SRP-PHAT power maps in reverberant scenarios. Signal Processing, 2016, 119: 209-228. [doi: 10.1016/j.sigpro.2015.08.003]
- 18 Zhao XY, Chen SW, Zhou L, *et al.* Sound source localization based on SRP-PHAT spatial spectrum and deep neural network. Computers, Materials & Continua, 2020, 64(1): 253-271.