

基于 TDOA 多声源定位的虚假声源消除方法

刘海涛^{1, 2}, 陈永华¹, 林艳明¹, 周新¹

(1. 华东交通大学机电与车辆工程学院 南昌, 330013) (2. 清华大学苏州汽车研究院 苏州, 215131)

摘要 基于到达时差(time difference of arrival,简称 TDOA)的多声源定位方法难以将麦克风获得的 TDOA 值与真实声源进行有效的关联。针对此问题,提出了一种基于 TDOA 的多声源空间定位方法。利用互相关算法估计声源的 TDOA 值,并基于 Chan 算法求解多目标声源的空间位置。为消除虚假声源,将阵列麦克风分为定位和校验两组子阵列,并构建阵列分组定位校验模型。定位麦克风用于所有可能声源的定位。校验麦克风用于虚假声源的消除,并获得多声源的初始位置。根据初始真实声源位置,构建全阵列 TDOA 序列校验模型,并获得最终真实声源位置。搭建了仿真及实验平台,对提出的方法进行验证。仿真及实验结果表明,提出的方法有效地消除了基于 TDOA 多声源定位的虚假声源,并能充分利用阵列麦克风数量来提升多声源定位精度。

关键词 多声源定位;到达时差;虚假声源;互相关算法;Chan 算法

中图分类号 TH73;TB52⁺⁹

引言

近几十年来,声源定位一直是研究热点,引起了众多学者的关注。声源定位在噪声源识别、目标声源跟踪、远程会议系统及智能机器人等诸多领域中得到广泛应用^[1-4]。声源定位一般采用麦克风阵列来估计声源位置。声源定位方法可归纳为 3 类:①波束形成^[5-6],传统波束形成是将各个麦克风接收到的声压信号进行对应的延时求和,在真正的声源方向的各个延时信号将会同相位叠加形成一个峰值,从而可以识别声源的方向;②声全息方法^[7-8],它通过求解声源的逆传播问题,并重建声场;③参数测量^[9-12],如到达时差、到达时间(time of arrival,简称 TOA)、达波方向(direction of arrival,简称 DOA)和接收信号强度(received signal strength,简称 RSS)等,利用获得的定位参数构建非线性多元方程组求解,来获取声源位置。传统的波束形成方法和声全息方法都是在重建的二维声场平面上识别声源,为获得精准的声源三维位置坐标,需要进行大量的复杂运算^[13]。另外,为了获得高分辨率的声场重构平面图,阵列中需要大量的麦克风数量,而 TDOA 方法用少量的麦克风即可实现声源定位。

对于单声源,Chan 等^[14]提出两步加权最小二乘算法,可求解出较精确声源位置。但对于多声源,阵

列中任意一麦克风对,利用互相关算法估计的 TDOA 与每个声源的关联是未知的,从而产生虚假声源,并降低声源定位的精度^[9]。针对这一问题,相关研究人员已经提出一些解决方法。Mukae 等^[15]利用检测到的声源信号和麦克风位置信息来构造一致性函数,通过快速搜索算法估计函数的最大值,最大值对应的位置为真实声源位置,但该方法计算效率低,且没有直接消除虚假声源问题,导致的位置估计不准确。Stotts 等^[16]基于结构分析实现多声源定位,该方法的计算时间随着监测区域的增大而增加,定位精度受网格大小的影响。Shen 等^[17]基于凸优化方法来消除虚假声源,消除虚假声源的本质是向量中每列的测量值相同或者近似相同对应同一声源。Venkateswaran 等^[18]采用并行和分层的方法消除虚假声源。该方法的定位精度会受到传播距离的影响,而且假设声源发射的次数多,会影响计算效率。以上消除虚假声源定位方法都存在定位精度不高和计算效率低等问题。

但佳壁^[19]采用了在麦克风阵列中构造校验麦克风,并利用校验麦克风与声源之间的相对位置关系,获得真实声源位置。然而该研究中校验阵列麦克风未能参与声源定位计算,降低了阵列麦克风的利用效率,多声源的定位精度还有提升空间。在但

* 国家自然科学基金资助项目(51765017);江西省自然科学基金资助项目(20202BABL204043);江西省主要学科学术和技术带头人培养计划资助项目(20204BCJL23034);华东交通大学科研基础经费资助项目(26541022)

收稿日期:2020-05-07;修回日期:2020-06-13

佳壁的研究基础上,笔者提出一种改进多声源定位的虚假声源消除方法。通过构建麦克风阵列分组定位校验模型有效消除虚假声源,获取初始真实多声源位置。再构建全阵列TDOA序列校验模型,计算初始真实多声源位置到全阵列麦克风的TDOA序列来近似匹配出全阵列正确TDOA序列。将全阵列的正确TDOA序列带入空间声源定位模型来获取最终真实多声源位置,充分利用阵列麦克风的数量来提升多声源定位精度。仿真和实验结果表明,本方法有效地消除了虚假声源,同时提升了多声源定位精度。

1 基于TDOA的空间声源定位模型

基于TDOA方法建立了空间声源定位模型。TDOA方法实质上是三角测量方法。假设声源的坐标为 $S=(x, y, z)$,声源与麦克风的距离差表达式为

$$\begin{cases} r_0 = \|M_0 - S\| \\ r_i = \|M_i - S\| \end{cases} \quad (i=1, 2, \dots, N_m) \quad (1)$$

其中: r_0 为声源与参考麦克风 M_0 之间的距离; r_i 为声源与其他麦克风 M_i 之间的距离; N_m 为除参考麦克风外的其他麦克风的数量。

通过构建声源到参考麦克风 M_0 和其他麦克风 M_i 的TDOA,可以得到声源的空间定位模型。表达式为

$$r_{i,0} = \|M_i - S\| - \|M_0 - S\| = c\tau_{i,0} \quad (2)$$

其中: c 为声速; $\tau_{i,0}$ 为声源到 M_i 和 M_0 的TDOA。

$\tau_{i,0}$ 可以用互相关算法^[20]来估计,设 $u_i(t)$ 和 $u_0(t)$ 分别为麦克风 M_i 和 M_0 采集到的声源信号。两个信号之间的互相关函数为

$$R_{i,0}(\tau) = \int_{-\infty}^{\infty} u_i(t)u_0(t+\tau)dt \quad (3)$$

$$\tau_{i,0} = \arg \max [R_{i,0}(\tau)] \quad (4)$$

空间声源定位模型由一组非线性多元方程组成。通常是将其转化为线性多元方程组来求解,对式(1)中 r_i 进行平方处理,表达式为

$$r_i^2 = (r_{i,0} + r_0)^2 \quad (5)$$

由式(4)进行展开得

$$2r_{i,0}r_0 + 2x_{i,0}x + 2y_{i,0}y + 2z_{i,0}z = K_i - K_0 - r_{i,0}^2 \quad (6)$$

其中: $K_i = x_i^2 + y_i^2 + z_i^2$; $K_0 = x_0^2 + y_0^2 + z_0^2$; $x_{i,0} = x_i - x_0$; $y_{i,0} = y_i - y_0$; $z_{i,0} = z_i - z_0$ 。

式(6)可写成矩阵形式,如式(7)所示

$$AX = B \quad (7)$$

$$\text{其中: } X = [x, y, z]^T; A = \begin{bmatrix} x_{1,0} & y_{1,0} & z_{1,0} \\ \vdots & \vdots & \vdots \\ x_{N,0} & y_{N,0} & z_{N,0} \end{bmatrix}; B = \begin{bmatrix} r_{1,0} \\ \vdots \\ r_{N,0} \end{bmatrix} r_0 + \frac{1}{2} \begin{bmatrix} r_{1,0}^2 - K_1 + K_0 \\ \vdots \\ r_{N,0}^2 - K_N + K_0 \end{bmatrix}。$$

当其他麦克风数量 $N_m = 3$ 时,理论上满足直接求解空间声源位置条件。当 $N_m \geq 4$ 时,Chan等^[14]提出了一种精度合理、有效的求解算法,该算法使用两次加权最小二乘法给出较准确的声源位置。因此,文中采用Chan的方法求解来获取声源位置。

2 多声源定位的虚假声源消除方法

2.1 虚假声源的产生原理

当只有一个声源时,式(3)只有一个峰值,将获取的麦克风阵列所有TDOA值可直接来定位目标声源。当有多个声源时,式(3)会产生多个峰值,每个峰值代表一个真实声源信息。然而,在缺乏其他声源特性信息的情况下,这些峰值的顺序是不确定的,不能将这些峰值与声源一一对应。两个麦克风检测3个声源的混合信号的互相关估计结果如图1所示。

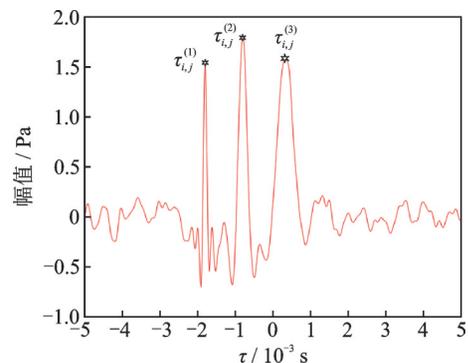


图1 3个声源的互相关结果

Fig.1 The cross-correlation results for three sound sources

在图1中,峰值的横坐标值 $[\tau_{i,j}^{(1)}, \tau_{i,j}^{(2)}, \tau_{i,j}^{(3)}]$ 表示从3个声源到麦克风 M_i 和 M_j 的TDOA。然而, $\tau_{i,j}^{(i)}$ 的顺序与声源对应不确定, $\tau_{i,j}^{(i)}$ 错误序列的组合会产生虚假声源,从而降低声源定位精度。基于排列组合的知识,可能产生的定位结果的数量为 $N_s^{N_m}$,其中 N_s 为声源的数量。虚假声源产生如图2所示。

从图1可以看出,声源数量与出现峰值的个数

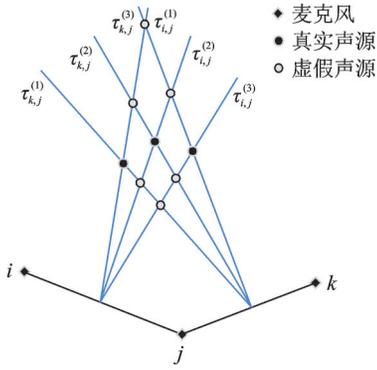


图2 虚假声源产生的示意图

Fig.2 The diagrammatic sketch of false source generation

相等,解决本研究提出的虚假声源消除方法过程中所需要的声源数量先验知识。从图2可以看出, $\tau_{k,j}^{(i)}$ 和 $\tau_{i,j}^{(k)}$ 的组合一共生成9个声源,其中只有3个是真实声源,另外6个虚假声源需要识别和消除,这也是本研究的重点。

2.2 麦克风阵列分组定位校验模型

对于多声源定位,可以利用声源与麦克风的相对位置关系来分离目标声源。本研究将阵列麦克风分为两组。第1组麦克风用于定位多个声源,为了减少计算量,第1组只需满足基于Chan算法求解的阵列麦克风最小数量要求。第2组麦克风用于消除虚假声源,通过阵列麦克风与声源之间的相对位置关系,构建麦克风阵列分组定位校验模型,并确定真实声源的位置。

第1组麦克风可能的定位结果数为 $N_p = N_s^{N_m}$, 由于计算时间随着麦克风数量的增加呈指数增长,除了参考麦克风外,还有4个麦克风构成第1组麦克风阵列,这是基于Chan的方法能够在三维空间中给出具有合理精度定位结果的最小麦克风数量要求。从真实声源到第1组麦克风阵列的TDOA如式(8)所示

$$T_{\text{first}} = \begin{bmatrix} \tau_{1,0}^{(1)} & \tau_{1,0}^{(2)} & \cdots & \tau_{1,0}^{(N_p)} \\ \tau_{2,0}^{(1)} & \tau_{2,0}^{(2)} & \cdots & \tau_{2,0}^{(N_p)} \\ \tau_{3,0}^{(1)} & \tau_{3,0}^{(2)} & \cdots & \tau_{3,0}^{(N_p)} \\ \tau_{4,0}^{(1)} & \tau_{4,0}^{(2)} & \cdots & \tau_{4,0}^{(N_p)} \end{bmatrix} \quad (8)$$

从真实声源到第2组麦克风阵列的TDOA如式(9)所示

$$T_{\text{second}} = \begin{bmatrix} \tau_{5,0}^{(1)} & \tau_{5,0}^{(2)} & \cdots & \tau_{5,0}^{(N_p)} \\ \tau_{6,0}^{(1)} & \tau_{6,0}^{(2)} & \cdots & \tau_{6,0}^{(N_p)} \\ \vdots & \vdots & & \vdots \\ \tau_{N_m,0}^{(1)} & \tau_{N_m,0}^{(2)} & \cdots & \tau_{N_m,0}^{(N_p)} \end{bmatrix} \quad (9)$$

一个可能的声源可以从 T_{first} 矩阵的每一列中选择一组TDOA序列,再基于Chan算法求解得到。基于排列组合的知识,可以从矩阵 T_{first} 中得到 N_s^4 可能的声源。所有可能的声源都由 S_p 表示,即

$$S_p = \left\{ \bar{s}_p^u | \bar{s}_p^u (\tau_{1,0}^{(k)}, \tau_{2,0}^{(l)}, \tau_{3,0}^{(n)}, \tau_{4,0}^{(m)}) \right\}$$

$$(k, l, n, m = 1, 2, \dots, N_s; u = k \ln m) \quad (10)$$

其中: $u = k \ln m$ 为所有可能的声源数,即目标声源数的4次方。

计算从 \bar{s}_p^u 到第2组麦克风的TDOA,如式(11)所示

$$\tilde{T}_{\text{second}} = \begin{bmatrix} \tilde{\tau}_{5,0}^{(1)} & \tilde{\tau}_{5,0}^{(2)} & \cdots & \tilde{\tau}_{5,0}^{(N_p)} \\ \tilde{\tau}_{6,0}^{(1)} & \tilde{\tau}_{6,0}^{(2)} & \cdots & \tilde{\tau}_{6,0}^{(N_p)} \\ \vdots & \vdots & & \vdots \\ \tilde{\tau}_{N_m,0}^{(1)} & \tilde{\tau}_{N_m,0}^{(2)} & \cdots & \tilde{\tau}_{N_m,0}^{(N_p)} \end{bmatrix} \quad (11)$$

需要一个校验矩阵来识别 N_p 个声源中的真实声源。校验矩阵 \hat{T}_{second} 的元素从矩阵 T_{second} 中选择。选择 \hat{T}_{second} 矩阵的准则定义为

$$\begin{cases} \text{For } \forall k_a \in \{1, 2, \dots, N_s\} \\ \hat{\tau}_{i,0}^{(u)} = \arg \min \left| \tilde{\tau}_{i,0}^{(u)} - \tau_{i,0}^{(k_a)} \right| \end{cases} \quad (u = 1, 2, \dots, N_p; i = 5, 6, \dots, N_m) \quad (12)$$

校验矩阵 \hat{T}_{second} 表达式为

$$\hat{T}_{\text{second}} = \begin{bmatrix} \hat{\tau}_{5,0}^{(1)} & \hat{\tau}_{5,0}^{(2)} & \cdots & \hat{\tau}_{5,0}^{(N_p)} \\ \hat{\tau}_{6,0}^{(1)} & \hat{\tau}_{6,0}^{(2)} & \cdots & \hat{\tau}_{6,0}^{(N_p)} \\ \vdots & \vdots & & \vdots \\ \hat{\tau}_{N_m,0}^{(1)} & \hat{\tau}_{N_m,0}^{(2)} & \cdots & \hat{\tau}_{N_m,0}^{(N_p)} \end{bmatrix} \quad (13)$$

矩阵 \hat{T}_{second} 的元素是从真实声源得到的。然而,矩阵 $\tilde{T}_{\text{second}}$ 的元素是所有声源获得的,其中包括真实声源和虚假声源。利用矩阵 \hat{T}_{second} 与 $\tilde{T}_{\text{second}}$ 对应列向量的相似性可以用来判断声源的真假,评价准则函数如式(14)所示

$$U_{N_s}(u) = \text{ARank}_{N_s} \left[\left\| \hat{T}_{\text{second}}(:, u) - \tilde{T}_{\text{second}}(:, u) \right\| \right] \quad (u = 1, 2, \dots, N_p) \quad (14)$$

其中: $U_{N_s}(u)$ 为前 N_s 个 u 的值,代表真实声源; $\text{ARank}_{N_s}[\cdot]$ 表示按升序排列,并获取前 N_s 个结果的函数; $\hat{T}_{\text{second}}(:, u)$ 为矩阵 \hat{T}_{second} 的列向量, $\tilde{T}_{\text{second}}(:, u)$ 为矩阵 $\tilde{T}_{\text{second}}$ 的列向量; $\|\cdot\|$ 为2-范数。

获得的初始真实声源表示为

$$S'_{\text{ir}} = \left\{ \bar{s}_p^u | u \in U_{N_s} \right\} \quad (15)$$

2.3 全阵列TDOA序列校验模型

以上仅有5个麦克风用于定位来获取初始的真

实声源,增加定位的麦克风数量会提升定位精度。文中利用初始真实声源位置信息,构建了全阵列TDOA序列校验模型,可以利用所有阵列麦克风来提升多声源的定位精度。

由式(8)和式(9)可以得到,真实声源到所有阵列麦克风的TDOA,如式(16)所示

$$T_{\text{all}} = \begin{bmatrix} \tau_{1,0}^{(1)} & \tau_{1,0}^{(2)} & \cdots & \tau_{1,0}^{(N_s)} \\ \tau_{2,0}^{(1)} & \tau_{2,0}^{(2)} & \cdots & \tau_{2,0}^{(N_s)} \\ \vdots & \vdots & & \vdots \\ \tau_{N_m,0}^{(1)} & \tau_{N_m,0}^{(2)} & \cdots & \tau_{N_m,0}^{(N_s)} \end{bmatrix} \quad (16)$$

矩阵 T_{all} 中每一行元素的顺序是不确定的,这就需要识别真实声源对应的正确TDOA序列,而利用初始真实声源位置可以识别出全阵列正确TDOA序列。初始真实声源到所有阵列麦克风的TDOA,如式(17)所示

$$\check{T}_{\text{all}} = \begin{bmatrix} \check{\tau}_{1,0}^{(1)} & \check{\tau}_{1,0}^{(2)} & \cdots & \check{\tau}_{1,0}^{(N_s)} \\ \check{\tau}_{2,0}^{(1)} & \check{\tau}_{2,0}^{(2)} & \cdots & \check{\tau}_{2,0}^{(N_s)} \\ \vdots & \vdots & & \vdots \\ \check{\tau}_{N_m,0}^{(1)} & \check{\tau}_{N_m,0}^{(2)} & \cdots & \check{\tau}_{N_m,0}^{(N_s)} \end{bmatrix} \quad (17)$$

由于初始真实声源的坐标与最终真实声源的坐标很接近,因此可以利用矩阵 \check{T}_{all} 对矩阵 T_{all} 的每一行进行重新排序。重新排序准则如式(18)所示

$$\begin{cases} \text{For } \forall k_a \in \{1, 2, \dots, N_s\} \\ \check{\tau}_{i,0}^{(n)} = \arg \min |\check{\tau}_{i,0}^{(n)} - \tau_{i,0}^{(k_a)}| \\ (n = 1, 2, \dots, N_s; i = 1, 2, \dots, N_m) \end{cases} \quad (18)$$

最后,获得的最终真实声源定位的重排TDOA矩阵为

$$\dot{T}_{\text{all}} = \begin{bmatrix} \dot{\tau}_{1,0}^{(1)} & \dot{\tau}_{1,0}^{(2)} & \cdots & \dot{\tau}_{1,0}^{(N_s)} \\ \dot{\tau}_{2,0}^{(1)} & \dot{\tau}_{2,0}^{(2)} & \cdots & \dot{\tau}_{2,0}^{(N_s)} \\ \vdots & \vdots & & \vdots \\ \dot{\tau}_{N_m,0}^{(1)} & \dot{\tau}_{N_m,0}^{(2)} & \cdots & \dot{\tau}_{N_m,0}^{(N_s)} \end{bmatrix} \quad (19)$$

选择矩阵 \dot{T}_{all} 的每一列,再基于Chan算法求解可以获得最终的真实多声源定位结果。

3 仿真分析

3.1 定位场景和声源信号构造

实际场景中的绝大多数音频噪声都是有色噪声。为了更好地模拟实际声源定位场景,文中采用有色噪声来构建各种目标声源信号。阵列几何关系与声源频率特性有关^[4]。对于单频声源定位,为了避免相位缠绕,阵列设置需要满足以下条件: $d_{\text{max}} <$

$c/2f_{\text{single}}$,其中 d_{max} 为参考麦克风与其他麦克风之间的最大距离, c 为声速, f_{single} 为单频声源的频率。对于宽带和脉冲声源,可以忽略相位缠绕现象,阵列几何关系将不受限制。文中选用麦克风阵列中的最大间距约为0.5 m,仅对单频声源具有频率上限,可实现绝大部分声源频率特性的声源定位。为了模拟实际声源的复杂多样性,不同声源设置为不同频带且部分频带重叠。图3为由高斯白噪声构成的4个声源的声谱图,分别由2,3,4个声源构成3种多声源定位场景。利用这3种多声源定位场景,来验证笔者提出的方法有效性。表1为3种场景中的声源坐标位置和频带信息。文中一共采用8个麦克风,以建立的参考坐标系的坐标原点为参考麦克风,另选4个麦克风构建定位麦克风阵列实现多声源定位,其余3个麦克风为校验麦克风。麦克风阵列的坐标位置如表2所示。

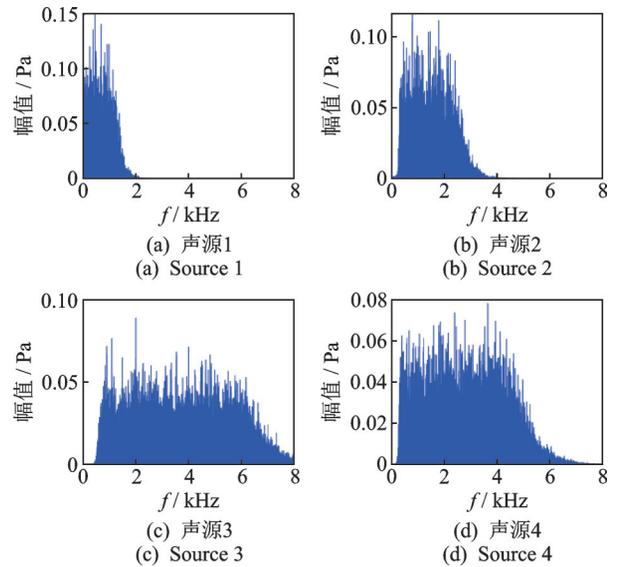


图3 4个声源的声谱图

Fig.3 The spectrogram of four sound sources

表1 3种定位场景的声源坐标和频带信息

Tab. 1 Coordinate position and frequency band information of the three localization scenarios

场景类型	声源数	$(x, y, z)/\text{m}$	频带/Hz
场景1#	2	$(-1.200, 4.000, 1.700)$	1~1 500
		$(1.500, 4.600, 1.500)$	300~2 800
场景2#	3	$(-1.200, 2.200, 0.800)$	1~1 500
		$(-0.100, 2.800, 0.500)$	300~2 800
		$(1.500, 3.000, 0.800)$	600~6 500
场景3#	4	$(-0.800, 2.200, 0.800)$	1~1 500
		$(-0.100, 2.800, 0.500)$	300~2 800
		$(1.500, 3.000, 0.800)$	600~6 500
		$(0.400, 2.500, 0.300)$	300~5 000

表 2 麦克风阵列的坐标

Tab. 2 The coordinates of the array microphones

麦克风	$(x, y, z)/\text{m}$
参考麦克风	$M_0(0.000, 0.000, 0.000)$
第 1 组麦克风	$M_1(0.800, -0.200, 0.500)$
	$M_2(0.693, 0.400, 0.000)$
	$M_3(0.693, 0.600, 0.800)$
	$M_4(-0.693, 0.400, 0.000)$
第 2 组麦克风	$M_5(-0.693, 0.400, 0.500)$
	$M_6(0.800, -0.200, -0.400)$
	$M_7(-0.800, -0.400, -0.300)$

3.2 多声源定位结果分析

3 个声源和阵列麦克风构成的定位场景如图 4 所示。在某一合理噪声水平下,分别对 3 种定位场景,将参考麦克风与另一定位麦克风构成的一组麦克风对,利用互相关算法估计的 TDOA 如图 5 所示,其中图(a)~(c)表示不同的定位场景。基于互相关算法估计所有 TDOA,再利用文中提出的消除虚假声源方法实现多声源定位,第 1 组麦克风获取的初始真实多声源定位的结果如图 6 所示,其中图(a)~(c)表示不同的定位场景。将第 1 组麦克风经蒙特卡罗循环处理,目的是消除随机噪声对单次定

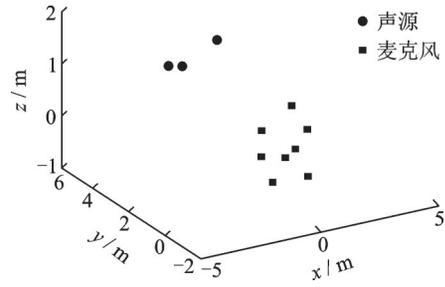


图 4 阵列麦克风和定位场景 2

Fig.4 Microphone array and localization scenario 2

位结果的影响,再利用所有阵列麦克风获得的最终真实多声源定位结果,与第 1 组麦克风的获得初始真实多声源定位结果进行比较,对比结果如表 3 所示。通过计算定位结果的平均误差来衡量定位的精度。平均定位误差定义为

$$ME = \frac{1}{N_s} \sum_{i=1}^{N_s} \|s_i - s'_i\| \quad (20)$$

其中:ME 为定位结果的平均定位误差; s_i 为原始声源的坐标; s'_i 为定位结果的坐标。

经蒙特卡罗过程可以表示为

$$MEL = \frac{1}{q} \sum_{j=1}^q ME_j \quad (21)$$

其中:MEL 为经蒙特卡罗处理后的多声源定位误差,MEL 值越小多声源定位精度越高; q 为仿真实验过程中蒙特卡罗次数,仿真过程令 $q=1\ 000$ 。

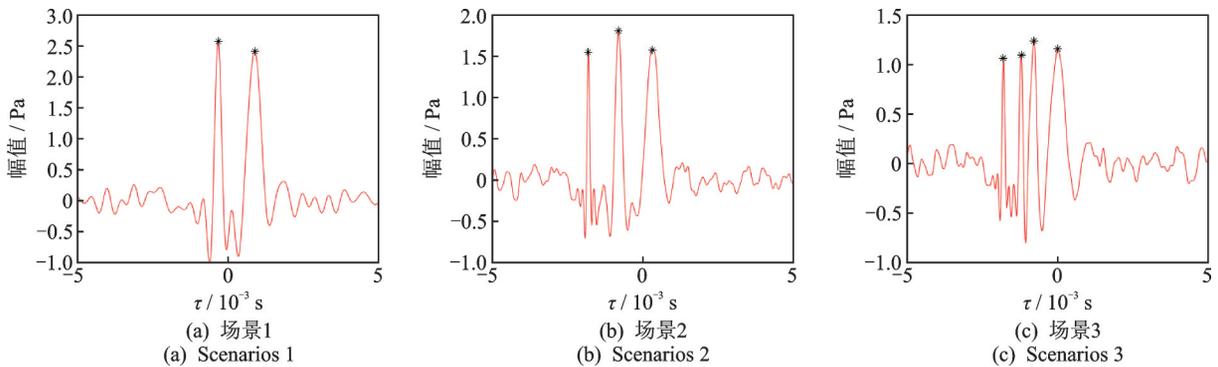


图 5 3 种定位场景的互相关结果

Fig.5 The cross-correlation results for the three localization scenarios

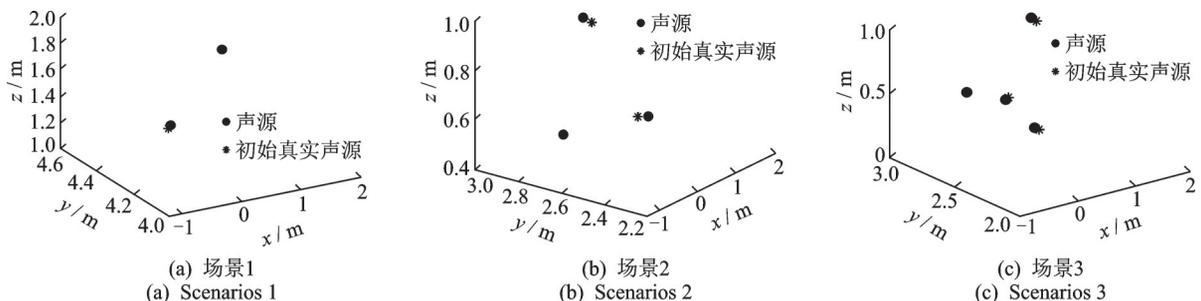


图 6 初始真实多源定位结果

Fig. 6 The results of initial real multi-sources localization

表 3 定位结果对比

Tab.3 The comparison of the localization results

场景	初始定位结果	最终定位结果	定位精度
	(x,y,z)/m	(x,y,z)/m	提升率/%
场景 1#	(-1.030, 3.456, 1.512)	(-1.284, 4.273, 1.799)	79.04
	(1.203, 3.709, 1.260)	(1.624, 4.991, 1.615)	
场景 2#	(-1.213, 2.146, 0.791)	(-1.231, 2.168, 0.791)	86.04
	(-0.094, 2.608, 0.494)	(-0.095, 2.708, 0.494)	
场景 3#	(-0.7017, 1.844, 0.713)	(-0.935, 2.079, 0.892)	71.62
	(-0.050, 2.200, 0.657)	(-0.160, 2.495, 0.615)	
	(1.040, 2.591, 0.657)	(1.741, 2.723, 0.743)	
	(0.313, 2.005, 0.329)	(0.363, 2.280, 0.303)	

从图 5 可以看出,在不同定位场景下,声源数量等于峰值的数量。从图 6 中可以看出,3 个定位场景中的多个声源有效地分离和定位,且定位误差合理,但是仅用 5 个麦克风来获得初始真实声源,没有充分利用麦克风数来定位,声源定位精度还有待提升。从表 3 可以看出,3 种定位场景下,所有阵列麦克风参与定位计算的最终定位结果精度均大幅提升,验证了笔者方法的有效性。

4 实验分析

为了验证笔者提出的基于 TDOA 多声源定位的虚假声源消除方法,在空旷的运动场上搭建实验测试系统对笔者提出方法进行验证。构造具有重叠频带的不同声源信号,所有声源的频带范围为 300~1 650 Hz,将构造声源信号的音频文件导入至扬声器播放。整个实验过程中忽略风速,设置声速 $c = 343 \text{ m/s}$ 。实验麦克风及声源位置如表 4 所示,在声源 $s_1 \sim s_4$ 中分别选择前 2 个、前 3 个和前 4 个,

构成定位场景 1、场景 2 和场景 3。实验过程的采样频率 $f_s = 10 \text{ kHz}$ 。实验场景由声源、信号采集装置和麦克风阵列三部分组成,如图 7(a)~(c)所示。图 8 是对前 5 个麦克风通道经傅里叶变换获得的频谱图,其中图 8(a)~(e)表示每个通道的图谱图。将每个麦克风检测的信号利用互相关算法估计出 TDOA,并基于 Chan 算法求解出所有可能声源位

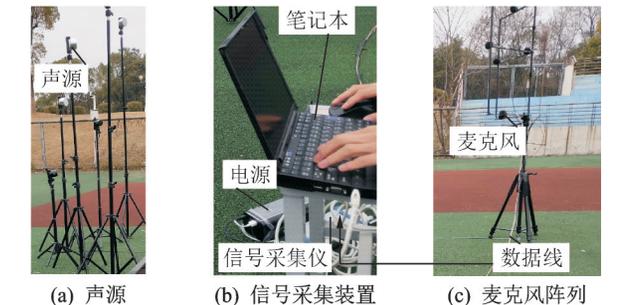


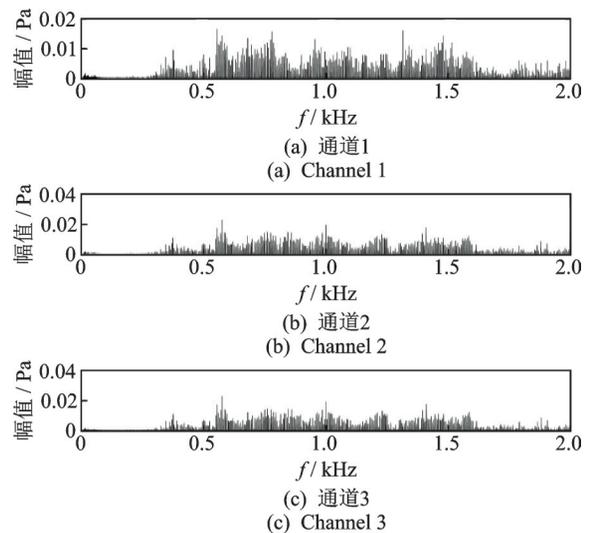
图 7 实验场景

Fig.7 The experiment scene

表 4 阵列麦克风和声源空间位置坐标

Tab.4 The coordinates of the array microphones and sound sources

麦克风	麦克风坐标 (x,y,z)/m	声源	声源坐标 (x,y,z)/m
M_0	(0.004, 0.028, 1.469)	s_1	(-0.830, 1.370, 1.527)
M_1	(0.465, 0.294, 1.485)	s_2	(0.000, 1.730, 1.230)
M_2	(-0.467, 0.303, 1.818)	s_3	(0.860, 1.400, 1.940)
M_3	(0.243, 0.167, 1.816)	s_4	(-1.121, 1.121, 1.690)
M_4	(0.313, 0.207, 1.022)		
M_5	(-0.447, 0.290, 1.474)		
M_6	(0.011, 0.032, 1.812)		
M_7	(-0.247, 0.176, 1.017)		



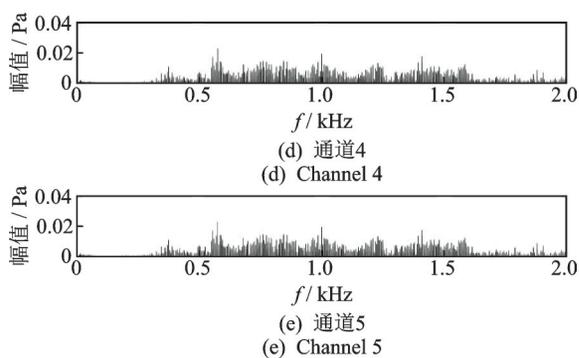


图 8 前 5 个麦克风通道经过傅里叶变换后的频谱图

Fig.8 Spectrum of the first five microphone channels after Fourier transform

置,再利用校验麦克风与声源的空间位置关系来消除虚假声源,得到初始的真实声源位置,定位结果如

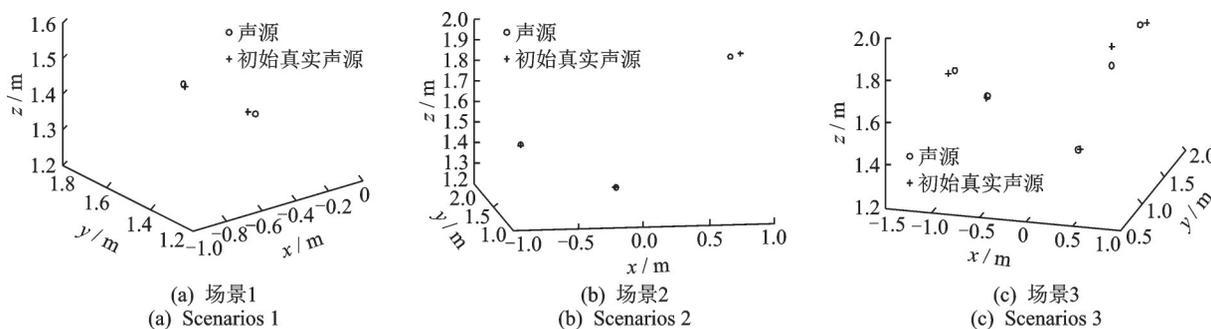


图 9 本方法获取的初始真实声源的实验定位结果

Fig.9 The experimental results of the initial real sound source obtained by the proposed method

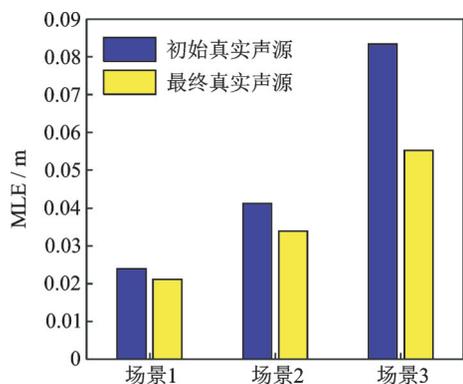


图 10 不同场景下最终真实声源位置与初始真实声源位置的平均定位误差对比

Fig.10 Comparison of the average positioning errors between the final real source position and the initial real source position under different scenarios

5 结论

1) 利用麦克风阵列分组定位校验模型有效地消除了虚假声源,实现多声源分离和定位。

2) 利用全阵列 TDOA 序列校验模型获得的最

图 9 所示,其中图 9(a)~(c)表示不同的定位场景。图 10 是在 3 种定位场景下,本方法获得的最终真实声源与初始真实声源的平均定位误差对比。

由图 8 可以看出,检测的声信号主频带在 300~1 650 Hz 之间,测试环境存在背景噪声。从图 9 可以看出,本方法成功消除了虚假声源,实现了多声源定位,且定位误差合理。从图 10 可以看出,在不同的定位场景下,本方法获得最终真实声源位置相比初始真实声源位置的平均定位误差均有所减小,本方法提升了多声源定位精度,实验与仿真结论一致。然而,实验过程中存在麦克风位置测量误差的影响,而且声源和阵列的相对位置与仿真有所不同,因而本方法在实验验证中的定位精度提升率比仿真稍有下降。

终真实多声源平均定位误差比分组定位校验模型获得的初始真实多声源平均定位误差更小。本方法充分利用了阵列麦克风,并有效提升了多声源定位精度。

3) 文中多声源定位方法的实验验证仅在空旷的测试环境下进行,该场景混响和其他干扰噪声影响较少;同时仅实现近场多声源定位。今后可研究存在麦克风位置误差以及远场和复杂噪声环境下的多声源定位方法。

参 考 文 献

[1] RASCON C, MEZA I. Localization of sound sources in robotics: a review [J]. Robotics and Autonomous Systems, 2017, 96: 184-210.

[2] PADOIS T, BERRY A. Application of acoustic imaging techniques on snowmobile pass-by noise [J]. Journal of the Acoustical Society of America, 2017, 141 (2):EL134-EL139.

[3] 张俊红,汤周杰,林杰威,等.IVMD 融合 Robust ICA 的内燃机噪声源分离[J].振动、测试与诊断,2020,40

- (1):28-34.
ZHANG Junhong , TANG Zhoujie, LIN Jiewei, et al. Noise source separation of internal combustion engine based on IVMD-RobustICA method[J]. Journal of Vibration , Measurement & Diagnosis, 2020, 40(1) : 28-34. (in Chinese)
- [4] MIAO F, YANG D, WEN J, et al. Moving sound source localization based on triangulation method [J]. Journal of Sound and Vibration, 2016, 385: 93-103.
- [5] CHIARIOTTI P, MARTARELLI M, CASTELLINI P, et al. Acoustic beamforming for noise source localization-reviews, methodology and applications[J]. Mechanical Systems and Signal Processing, 2019, 120: 422-448.
- [6] 季浩然, 马晓川, 张舒皓, 等. 功率受限的常规波束形成后处理拟合方法[J]. 声学学报, 2020, 45(1): 1-14.
JI Haoran , MA Xiaochuang, ZHANG Shuhao, et al. Power constraint conventional beamforming post-processing fitting method[J]. Acta Acustica, 2020, 45(1) : 1-14. (in Chinese)
- [7] 毛荣富, 朱海潮, 杜向华. 近场声全息中的声场分解技术[J]. 振动、测试与诊断, 2011, 31(6): 120-125.
MAO Rongfu, ZHU Haichao, DU Xianghua. Partial sound field decomposition techniques in near-field acoustic holography [J]. Journal of Vibration, Measurement & Diagnosis, 2011, 31(6) : 120-125. (in Chinese)
- [8] HOFFMANN F, FAZI F M, WILLIAMS E G, et al. A general radiation model for sound fields and nearfield acoustical holography in wedge propagation spaces [J]. Journal of the Acoustical Society of America, 2017, 142(3): 1249-1260.
- [9] SUNDAR H, SREENIVAS T V, SEELAMANTULA C S, et al. TDOA-based multiple acoustic source localization without association ambiguity [J]. IEEE Transactions on Audio, Speech, and Language Processing, 2018, 26(11): 1976-1990.
- [10] ENNASR O N, TAN X. Time-difference-of-arrival (TDOA) -based distributed target localization by a robotic network [J]. IEEE Transactions on Control of Network Systems, 2020, 7(3): 1416-1427.
- [11] 李红光, 郭英, 睦萍, 等. 基于稀疏贝叶斯的多跳频信号二维波达方向估计[J]. 上海交通大学学报, 2020, 54(4): 359-368.
LI Hongguang, GUO Ying, SUI Ping, et al. Two-dimensional DOA estimation for multi-hopping signals based on sparse bayes [J]. Journal of Shanghai Jiao Tong University, 2020, 54(4) : 359-368. (in Chinese)
- [12] 熊杰, 陈俊, 宁静, 等. 移动辐射源 AOA-TDOA-FDOA 联合定位闭合解算法[J]. 电子科技大学学报, 2020, 49(2): 219-227.
XIONG Jie, CHEN Jun, NING Jing, et al. Closed-form AOA-TDOA-FDOA solution for moving source location [J]. Journal of University of Electronic Science and Technology of China, 2020, 49(2) : 219-227. (in Chinese)
- [13] PADOIS T, BERRY A. Two and three-dimensional sound source localization with beamforming and several deconvolution techniques [J]. Acta Acustica United With Acustica, 2017, 103(3): 392-400.
- [14] CHAN Y T, HO K C. A simple and efficient estimator for hyperbolic location [J]. IEEE Transactions on Signal Processing, 1994, 42(8): 1905-1915.
- [15] MUKAE J, ISHIDA Y, MURAKAMI T. Improvement of the histogram in the degenerate unmixing estimation technique algorithm [J]. Journal of the Acoustical Society of America, 2014, 136(4) : 2214-2215.
- [16] STOTTS S A, MARTIN J L, BEDFORD N R. Multiple-source localization using GPS technology and received arrival time structure analysis in an air-deployed system [J]. IEEE Journal of Oceanic Engineering, 1997, 22(3): 576-582.
- [17] SHEN H, DING Z, DASGUPTA S, et al. Multiple source localization in wireless sensor networks based on time of arrival measurement [J]. IEEE Transactions on Signal Processing, 2014, 62(8): 1938-1949.
- [18] VENKATESWARAN S, MADHOW U. Localizing multiple events using times of arrival: a parallelized, hierarchical approach to the association problem [J]. IEEE Transactions on Signal Processing, 2012, 60(10): 5464-5477.
- [19] 但佳壁. 汽车排气外声场可视化方法研究[D]. 北京: 清华大学, 2017.
- [20] KNAPP, C, CARTER. The generalized correlation method for estimation of time delay [J]. IEEE Transactions on Acoustics, Speech and Signal Processing, 1976, 24(4): 320-327.



第一作者简介: 刘海涛, 男, 1986年1月生, 博士、副教授、硕士生导师。主要研究方向为阵列信号处理、信息融合、声源识别与定位、声源追踪及振动噪声控制。曾发表《Arbitrary microphone array optimization method based on TDOA for specific localization scenarios》(《Sensors》2019, Vol.19, No.19)等论文。
E-mail: 2860@ecjtu.edu.cn

通信作者简介: 陈永华, 男, 1994年12月生, 硕士生。主要研究方向为阵列信号处理、声源识别与定位。
E-mail: 2018038080200012@ecjtu.edu.cn