

基于声传播算子的声源定位实验模拟方法研究*

林志斌[†] 卢晶 徐柏龄

(南京大学声学研究所 近代声学重点实验室 南京 210093)

摘要 声传播算子是一种高效的时域声场计算方法,它能够很方便地计算出给定系统参数下任意时刻任意位置的声场变化情况,本文采用这种方法计算所得的二维房间声场信息进行传声器阵列的声源定位仿真实验。计算结果表明,用该方法获取的阵列数据能有效地应用于阵列信号处理算法中,准确地估计出初始高斯脉冲声源的方向。声传播算子声场计算方法能为传声器阵列声源定位的实验提供方便,使得传声器阵列声源定位算法在不同混响时间的鲁棒性实验研究变得更加简捷。

关键词 麦克风阵列, 延迟估计, 声源定位, 声传播算子

Numerical modeling of a sound localization experiment based on the acoustical wave propagator

LIN Zhi-Bin LU Jing XU Bo-Ling

(Key Laboratory of Modern Acoustics, the Institute of Acoustics, Nanjing University, Nanjing 210093)

Abstract The acoustical wave propagator is a highly efficient technique for the calculation of the time-domain evolution of the acoustical wave at any time and position with given acoustical parameters. In this paper, the acoustical wave propagator is applied to a sound localization problem in the environment of a two-dimensional room. Simulation results illustrate that the signals acquired by this technique can be used in the algorithm of array signal processing efficiently. As a result, the technique of acoustical wave propagator provides a feasible experiment simulation method for sound localization using microphone array, which makes the experiment research of the robustness of a sound localization algorithm at various reverberation time more conveniently.

Key words Microphone array, Time delay estimation, Sound localization, Acoustical wave propagator (AWP)

2007-01-19 收稿; 2007-05-09 定稿

* 国家自然科学基金(No. 60340420325 和 No. 10604030), 博士点基金项目(No. 20040284038)

作者简介:林志斌(1979 -),男,博士,讲师,研究方向:音频信号处理,计算声学。

卢晶(1977 -),男,博士,副教授。徐柏龄(1941 -),男,教授、博士生导师。

[†] 通信联系人 E-mail: zblin@nju.edu.cn

1 引言

传声器阵列声源定位是指用传声器阵列拾取声音信号,通过对多路声音信号进行分析与处理,在空间域中定出一个或是多个声源的空间坐标。它涉及到阵列信号处理和数据融合等理论知识,是语音信号处理中具有挑战性的研究课题。同时它还涉及到多路信号采集、实验系统搭建以及实验数据处理等方面,实验成本较高;在研究算法的鲁棒性问题上,比如在不同混响时间下,实验环境实现上困难较多,整个算法研究投入的时间和精力较多^[1-4]。

声源定位算法的实验模拟可以采用声场计算的方法产生相应的实验数据进行验证,这将使得算法实验研究的工作效率大大提高,比如对混响场中时延问题的研究,很多学者都采用 image 模型,通过设置 image 模型的边界吸收系数来模拟不同的混响时间,获取相应的数据进行时延估计,达到研究时延算法的抗混响能力的目的^[5-6]。其它常用的数值计算方法如有限差分法(FDTD)和边界元法(BETD)在阵列信号数据的模拟中也常常用到^[7]。但是如何在计算量和计算精度之间取得平衡是这些方法面临的共同难题。2000年,声传播算子方法引入到声场的时域计算中^[8],该算法将频域计算的谱方法和时域计算的切比雪夫展式相结合,是一种高效准确的时域计算方法。该方法有一个显著的优点:只需要知道空间声学参数(包括声速,密度等)的分布,不需要人为构造边界条件,这样便使得计算与实际情况匹配得很好,并能及时地追踪参数变化对声场的影响^[9]。

本文正是采用这种高效的时域计算方法,模拟了二维房间中传声器阵列各个传声器的信号,比较了采集信号和解析解的误差,然后通过相应的阵列信号处理的方法计算了各个传声器之间的时延,进一步计算初始声源的

间位置。

2 声源定位及声传播算子理论

2.1 声源定位及其时延估计方法

现有的声源定位技术基本上可以分为三类:基于最大输出功率的可控波束形成技术,高分辨率谱估计技术和基于声达时间差(TDOA)的定位技术。基于声达时间差声源定位方法一般分为两个步骤来进行^[1]:先进行声达时间差估计,并从中获取传声器阵列中阵元间的声延迟 TDOA,再利用获取的声达时间差,结合已知的传声器阵列的空间位置进一步定出声源的位置。这种方法的计算量一般比前两种要小,更有利于实时处理,所以它在语音信号的声源定位中占有很大的比重。在这种方法中,时延估计的精度是关系到声源定位精确与否的关键因素,时延估计的方法很多,运用较多的是广义互相关函数法(GCC)和最小均方差自适应滤波法(LMS)。本文采用了广义互相关函数法^[4, 10]。

假设系统中 $s(n)$ 为声源, $x_1(n)$ 和 $x_2(n)$ 为传声器 1 和传声器 2 拾取的信号, $b_1(n)$ 和 $b_2(n)$ 为传声器 1、2 拾取的除了声源信号外的附加噪声信号。传声器 1、2 得到的信号可以如下表示:

$$\begin{aligned} x_1(n) &= a_1 s(n) + b_1(n) \\ x_2(n) &= a_2 s(n - \tau) + b_2(n) \end{aligned} \quad (1)$$

其中 a_1 和 a_2 分别为声源信号到达传声器后的衰减系数, τ 为两个传声器拾取的信号的延迟时间,即声达时间差 TDOA。这时, TDOA 可以采用传统的相关法进行估计,相关方程为:

$$\begin{aligned} R_{x_1 x_2}(\tau) &= \int_{-\infty}^{+\infty} \varphi(f) S_{x_1 x_2}(f) e^{j2\pi f \tau} df \\ &= \int_{-\infty}^{+\infty} \Psi_{x_1 x_2}(f) e^{j2\pi f \tau} df \end{aligned} \quad (2)$$

其中 $S_{x_1 x_2}(f) = E\{X_1(f)X_2^*(f)\}$ 是传声器 1、2 拾取信号的相关谱, $\varphi(f)$ 是权函数, $\Psi_{x_1 x_2}(f) =$

$\varphi(f)S_{x_1x_2}(f)$ 是广义相关谱, 这样 TDOA 可用 $\hat{\tau}_\varphi = \operatorname{argmax}_\tau R_{x_1x_2}(\tau)$ 求解。在公式(2)中, 选取不同的权函数 $\varphi(f)$ 就可以得到不同的声达时间差 TDOA 的不同算法, 它的选取可以根据实际的声学环境选择相应的准则, 使得 $R_{x_1x_2}(\tau)$ 有个比较尖锐的峰值, 可以得到最好的估计效果。但是在实际应用中, 权函数的选取是一个难点, 目前用得较多的是基于互功率谱相位加权 (CSP) 法^[1-4,10], 其中加权函数选取为 $\varphi(f) = \frac{1}{|S_{x_1x_2}(f)|}$, 这种方法通过对信号功率谱的归一化, 去除了信号的幅度信息, 只保留了信号的相位特性, 对于噪声和混响都有较好的抑制效果。

2.2 二维声传播算子方法^[8,9]

根据声传播算子理论, 二维房间中的线性声波方程可用下面的偏微分方程表示:

$$\frac{\partial}{\partial t} \Phi(x, y, t) = -\hat{\mathbf{H}}\Phi(x, y, t), \quad (3)$$

其中 x, y 表示的是空间坐标, $\Phi(x, y, t)$ 是状态矩阵, $\hat{\mathbf{H}}$ 是空间微分算子, 对于二维的声波方程来说, $\Phi(x, y, t)$ 可用声压 $p(x, y, t)$ 和声速的 x 分量 $v_x(x, t)$ 和 y 分量 $v_y(y, t)$ 表示:

$$\Phi(x, y, t) = \begin{bmatrix} p(x, y, t) \\ v_x(x, t) \\ v_y(y, t) \end{bmatrix} \quad (4)$$

而 $\hat{\mathbf{H}}$ 可以表示为:

$$\hat{\mathbf{H}} = \begin{bmatrix} 0 & \rho c^2 \frac{\partial}{\partial x} & \rho c^2 \frac{\partial}{\partial y} \\ \frac{1}{\rho} \frac{\partial}{\partial x} & 0 & 0 \\ \frac{1}{\rho} \frac{\partial}{\partial y} & 0 & 0 \end{bmatrix} \quad (5)$$

二维房间的边界条件可以通过不同的密度函数 $\rho(x, y)$ 和声速函数 $\rho(x, y)$ 表示。将公式(3)在时间上积分可得:

$$\Phi(x, y, t) = e^{-(t-t_0)\hat{\mathbf{H}}}\Phi(x, y, t_0). \quad (6)$$

因此, 可以通过求解声传播算子 $e^{-(t-t_0)\hat{\mathbf{H}}}$ 并结合初始时刻的 $\Phi(x, y, t_0)$ 得到任意时刻的二

维房间任意位置的声压和两个分量的声速, 媒质的特性可以包含在算子矩阵 $\hat{\mathbf{H}}$ 中。

为了方便计算, 将声波算子矩阵归一化:

$$\hat{\mathbf{H}}' = \frac{\hat{\mathbf{H}}}{|\lambda|_{\max}} \quad (7)$$

其中 $|\lambda|_{\max}$ 是算子矩阵的最大特征值, 同时定义 $R = (t - t_0) |\lambda|_{\max}$, 该算子矩阵则可以用第一类切比雪夫多项式展开如下:

$$e^{-(t-t_0)\hat{\mathbf{H}}} = e^{-R\hat{\mathbf{H}}'} = \sum_{n=0}^{\infty} a_n(R) T_n(\hat{\mathbf{H}}'), \quad (8)$$

其中: $a_n(R) = 2J_n(R)$, 除 $a_0(R) = J_0(R)$ 外, $J_n(R)$ 是 n 阶贝塞尔函数; 零阶和一阶的第一类切比雪夫多项式分别定义为 $T_0(\hat{\mathbf{H}}') = \mathbf{I}$ 和 $T_1(\hat{\mathbf{H}}') = \hat{\mathbf{H}}'$, 其余的由下列递推公式递推:

$$T_{n+1}(\hat{\mathbf{H}}') = 2\hat{\mathbf{H}}'T_n(\hat{\mathbf{H}}') - T_{n-1}(\hat{\mathbf{H}}'), \quad (9)$$

则公式(3)变成:

$$\begin{aligned} \Phi(x, y, t) &= e^{-(t-t_0)\hat{\mathbf{H}}}\Phi(x, y, t_0) \\ &= \sum_{n=0}^{\infty} a_n(R) T_n(\hat{\mathbf{H}}')\Phi(x, y, t_0), \end{aligned} \quad (10)$$

同时将傅立叶变换引入到空间差分的计算当中:

$$\frac{\partial^n}{\partial x^n} \Phi(x, y, t) = F^{-1} \{ (jk)^n F[\Phi(x, y, t)] \}, \quad (11)$$

其中 $F[\]$ 和 $F^{-1}[\]$ 分别表示傅立叶变换和反变换, k 为波数。

由前可知, 声传播算子方法将频域计算的谱方法和时域计算的切比雪夫展式相结合, 是一种高效准确的时域计算方法。通过设置二维房间的声学参数的分布, 不需要人为构造边界条件, 便能很好地模拟任意时刻任意位置的自由场和不同混响时间的混响场。

3 算法测试与仿真

3.1 声传播算子数值计算精度

这里首先验证二维声传播算子计算二维

房间的计算精度问题,为后面的计算提供依据。为了检验数值计算精度,这里采用极坐标下的解析结果进行比较,在极坐标系中,二维的声波方程可以用下面式子表示:

$$\frac{\partial^2 p}{\partial r^2} + \frac{1}{r} \frac{\partial p}{\partial r} = \frac{1}{c^2} \frac{\partial^2 p}{\partial t^2}, \quad (12)$$

同时假设初始条件为:

$$p(r, 0) = p_0(r), \quad \frac{\partial p(r, 0)}{\partial t} = v_0(r). \quad (13)$$

这时 t 时刻的声压解可由下面展开得到^[11]:

$$p(r, t) = \int_0^\infty \xi \bar{p}(\xi) \cos c \xi t J_0(\xi r) d\xi + \frac{1}{c} \int_0^\infty \xi \bar{v}(\xi) \sin c \xi t J_0(\xi r) d\xi, \quad (14)$$

其中:

$$\bar{p}(\xi) = \int_0^\infty r p_0(r) J_0(\xi r) dr, \quad (15)$$

$$\bar{v}(\xi) = \int_0^\infty r v_0(r) J_0(\xi r) dr, \quad (16)$$

二维空间的初始波包采用如下形式:

$$p_0(r, 0) = e^{-r^2/a^2}, \quad (17)$$

$$v_0(r, 0) = 0. \quad (18)$$

将公式(17)和公式(18)代入公式(14)可得二维空间的解析解为:

$$p(r, t) = \frac{a^2}{2} \int_0^\infty \xi \exp\left(-\frac{\xi^2 a^2}{4}\right) \cos c \xi t J_0(\xi r) d\xi \quad (19)$$

此时定义声波传播 t 时间后 $[0, r]$ 空间区域内各点的相对误差为:

$$\delta(r, t) = \frac{|p_{\text{awp}}(r, t) - p_e(r, t)|}{|\max(p_e(r, t))|} \quad (20)$$

其中 $p_{\text{awp}}(r, t)$ 和 $p_e(r, t)$ 分别为数值计算结果和解析结果,当声波传播 0.01s 后计算空间内各点的相对误差如图 1 所示。图 1 表示的是数值计算和解析结果的相对误差随距离变化

情况,从图中可以看出,用声传播算子计算结果和解析结果的相对误差不大于 2×10^{-6} 量级,计算精度很高。

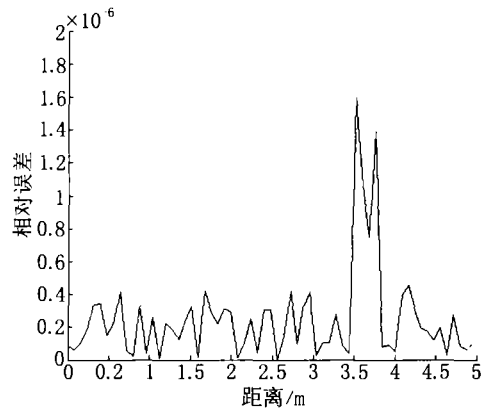


图 1 声传播算子数值计算结果和解析结果的相对误差图

3.2 自由场及不同混响场数值计算声场定位

这里测试了二维平面下由四个传声器阵列组成的脉冲声源定位系统,系统实验图如图 2 所示:

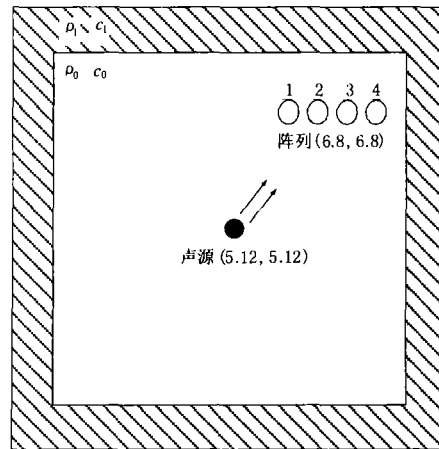


图 2 基于声传播算子的声源定位实验模拟仿真示意图

如图 2 所示,基于声传播算子的声源定位实验模拟系统的二维房间大小为 10.24m × 10.24m,初始脉冲声源位于二维房间的中心,即位置为(5.12m,5.12m),传声器阵列最左边的传声器 1 位置为(6.8m,6.8m),传声器间距为 8cm。同时,图中的阴影部分为声传播算子

模拟混响场时设置的不同媒质性质区域, ρ_1 、 c_1 和 ρ_0 、 c_0 为两种不同媒质的密度和声速, 可以根据实际实验设置需要进行设置, 本文分别模拟了自由场声源定位和两种混响场的声源定位情况, 其中自由场取声场参数 $\rho_1 = \rho_0$ 及 c_1 、 c_0 , 两种混响场声场参数分别取 $\rho_1 = 3\rho_0$ 和 $\rho_1 = 5\rho_0$, 两者的声速都是 $c_1 = c_0$ 。声源取如下的高斯脉冲波包:

$$p(x, y, 0) = e^{-\frac{(x-x_0)^2}{4a^2} - \frac{(y-y_0)^2}{4a^2}} \quad (21)$$

其中 a 为高斯波包的宽度系数, (x_0, y_0) 为高斯初始波包在二维房间中的位置坐标, 本文中 $a = 0.25$, $x_0 = 5.12\text{m}$, $y_0 = 5.12\text{m}$ 。自由场条件下 4 个传声器采集信号及三种声场参数下传声器 1 采集的信号分别如图 3 和图 4 所示:

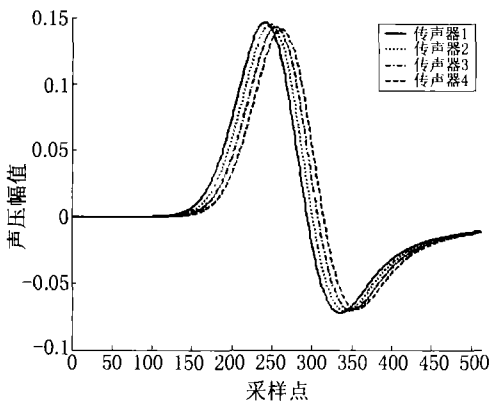


图3 自由场声场参数条件下各个传声器采集的信号

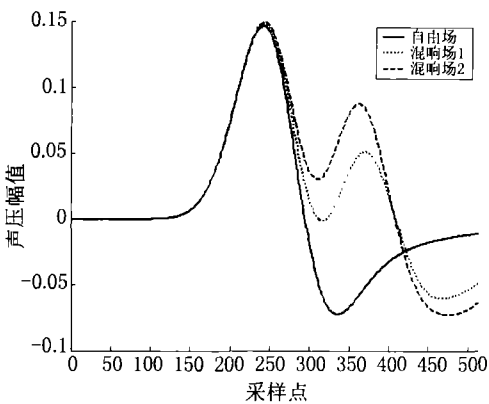


图4 三种声场参数下传声器1采集信号

图 3 是声传播算子计算所得的自由场下各个传声器的信号, 图 4 是三种声场参数条件下传声器 1 采集信号, 分别代表了自由场、混响场 1 和混响场 2 的信号情况, 其中采样频率为 40kHz, 同时取声速为 $c_0 = 344\text{m/s}$, 密度 $\rho_0 = 1.21\text{kg/m}^3$ 。从图 3 可以看出不同的传声器采集的信号延迟, 同时从图 4 可以看出, 在声传播算子方法中加入相应的边界媒质条件, 可以很方便地模拟出房间壁面的反射作用, 从图中也可以非常清楚地分辨出传声器 1 采集的直达声和反射声, 从而达到模拟不同混响时间的效果。

随后采用广义相关法 (GCC) 计算 TDOA 估计值以及由 TDOA 计算的声源方向, 算法帧长取 256 点, 采用了半重叠的汉宁窗 (Hanning)。表 1 和表 2 就是采用声传播算子获得的三种声场参数下传声器阵列信号计算相应的传声器 1 和各个传声器之间的 TDOA 和声源方向。

表 1 三种声场参数下传声器 1 和各传声器之间的 TDOA 计算值和理论值比较

| | 传声器 1 和 2 | 传声器 1 和 3 | 传声器 1 和 4 |
|----------------|--------------|--------------|--------------|
| 理论值(采样点) | 6.6 | 13.2 | 19.8 |
| 自由场计算值(采样点) | 6.4 | 12.9 | 19.7 |
| 混响场 1 计算值(采样点) | 6.3 | 13.3 | 19.9 |
| 混响场 2 计算值(采样点) | 6.3 | 13.2 | 19.9 |

表 2 三种声场参数下根据 TDOA 计算所得声源方向的计算值和理论值比较

| | 传声器 1 和 2 | 传声器 1 和 3 | 传声器 1 和 4 |
|-------------|--------------|--------------|--------------|
| 理论值(度) | 44.8° | 44.8° | 44.8° |
| 自由场计算(度) | 46.5° | 46.1° | 45.1° |
| 混响场 1 计算(度) | 47.3° | 44.4° | 44.5° |
| 混响场 2 计算(度) | 47.3° | 44.8° | 44.5° |

如表 1 和表 2 所示, 采用声传播算子获取的传声器阵列数据能很好地模拟实际情况的声场数据, 时延估计值和声源定位的准确性较高, 同

时还能模拟不同混响时间下的声源定位问题,计算速度快,精度高,这种实验模拟的方法能为声源定位提供较准确的数据,为算法的研究提供极大的方便。这里需要说明的是表 1 的采样点出现小数位是为了提高时延估计精度经过插值获取的。同时,为了提高声源定位的精度,还可以通过升采样等等方法对声传播算子方法获取的信号进行处理,得到更准确的声源方向。

4 结论

声传播算子方法是一种高效的声场计算方法,本文首先研究了该方法在二维房间中的数值计算和解析结果之间的误差,结果表明算法结果有很高的精度;接着采用这种方法模拟了三种不同声场参数下二维房间内的传声器阵列采集的信号,并进行声源定位的研究。模拟实验结果说明,无论是自由场还是混响场,将声传播算子的方法获取的信号用于传声器阵列信号处理中,通过相应的声源定位方法能准确的估计出声源的位置。声传播算子能准确方便地计算给定条件下传声器阵列的信号,从而为传声器阵列信号处理方法的研究、实验模拟提供方便。

参 考 文 献

- [1] Brandstein M. S., Adcock J. E. and Silverman H. F. A Close Form location estimation for use with room environment microphone arrays. *IEEE Transaction on Speech and Audio Processing*, 1997, 5(1):45-50.
- [2] Brandstein M. S. and Silverman H. F. A practical methodology for speech source localization with microphone arrays. *Computer Speech and Language*, 1997, 11:91-126.
- [3] Benesty J., Chen J. D. and Huang Y. T. Time delay estimation via linear interpolation and cross correlation. *IEEE Transaction on Speech and Audio Processing*, 2004, 12(5):509-519.
- [4] 林志斌,徐柏龄.基于传声器阵列的声源定位. *电声技术*, 2004, 5:19-23.
- [5] Allen J. B. and Berkley D. A. Image model for efficiently modeling small-room acoustics. *Journal of the Acoustical Society of America*, 1979, 65(4):943-950.
- [6] Brandstein M. S. Time delay estimation of reverberated speech exploiting harmonic structure. *Journal of the Acoustical Society of America*, 1999, 105(5):2914-2919.
- [7] Franek O., Pedersen C. F. and Andersen J. B. Numerical modeling of a spherical array of monopoles using FDTD method. *IEEE Transactions on Antennas and Propagation*, 2006, 54(7):1952-1963.
- [8] Pan J., Wang J. B. Acoustical wave propagator. *Journal of the Acoustical Society of America*, 2000, 108(2):481-487.
- [9] Sun H. M., Wang J. B. and Pan J. Acoustical wave propagator method for two-dimensional sound propagation. *Acoustics*, 2003, 9:381-390.
- [10] Knapp C. H. and Carter G. C. The generalized correlation method for estimation of time delay. *IEEE Transaction on Acoustics, Speech and Signal Processing*, 1976, 24(4):320-327.
- [11] Graff K. F. *Wave motion in elastic solids*. Oxford: Clarendon Press, 1975, 217-219.