

# 基于可控波束响应的球形传声器 阵列定位方法研究

章康宁<sup>1</sup>, 王星稚<sup>1</sup>, 刘康<sup>1</sup>, 高志华<sup>1</sup>, 孙科学<sup>2,3</sup>

(1. 南京邮电大学通信与信息工程学院, 江苏南京 210023; 2. 南京邮电大学电子科学与工程学院, 江苏南京 210023;  
3. 江苏省射频集成与微组装工程实验室, 江苏南京 210023)

**摘要:** 球形传声器阵列因其完全对称的结构, 可在三维空间内实现有效的声源定位。文章重点关注基于球谐函数展开的可控波束响应算法, 推导了算法的实现流程, 并通过仿真验证了可控波束响应算法的性能。建立了包含 8 个传声器单元的实验系统, 其中球形结构在 8 个单元的条件下可以视为立方结构, 通过低阶球谐展开实验验证了可控波束算法的性能。在视听室的实验结果充分验证了定位算法的有效性, 定位误差小于 10°。

**关键词:** 声源定位; 球形传声器阵列; 可控波束响应; 定位算法

中图分类号: TN912.16

文献标识码: A

文章编号: 1000-3630(2015)-06-0489-04

DOI 编码: 10.16300/j.cnki.1000-3630.2015.06.003

## Study of controllable beam response algorithm for spherical microphone arrays

ZHANG Kang-ning<sup>1</sup>, WANG Xing-zhi<sup>1</sup>, LIU Kang<sup>1</sup>, GAO Zhi-hua<sup>1</sup>, SUN Ke-xue<sup>2,3</sup>

(1. College of Telecommunications & Information, Nanjing University of Posts and Telecommunications, Nanjing 210023, Jiangsu, China;  
2. College of Electronics Science and Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210023, Jiangsu, China;  
3. Jiangsu Province Engineering Laboratory of RF Integration & Micropackage, Nanjing 210023, Jiangsu, China)

**Abstract:** Spherical microphone arrays can realize effective beamforming and localization in 3-D space because of its completely symmetrical structure. This paper focuses on the controllable beam response algorithm based on spherical harmonic function expansion. The implementation process of the algorithm is derived, and the performance of the algorithm is verified by simulation. An experimental system containing eight microphone units is established. The spherical structure under the condition of 8 units can be seen as a cubic structure. Through the experiment of low order spherical harmonic expansion, the performance of the controllable beam algorithm is verified. The experimental results in the audiovisual room validate the effectiveness of the localization algorithm, and the error of localization is less than 10 degrees.

**Key words:** sound source localization; spherical microphone array; controllable beam response; location algorithm

## 0 引言

声源定位通常是指通过声场分析获得噪声的方位和频谱信息, 其在故障诊断和噪声治理等领域有广泛的应用。在较为复杂的声学环境中, 声信号往往来源于多个不同的声源, 如果各个声源不能独立工作, 简单的声学测量就无法对源信息做准确的分析。有时甚至同一个声源也会因体积较大或分割

振动等原因呈现出较为复杂的声辐射特性, 这同样给声源识别带来较大困难。在这些复杂条件下, 可以借助传声器阵列的声源定位特性对声源进行识别。传声器阵列是指按一定空间分布放置的一组传声器。通过利用声波抵达阵列中每个单元之间的幅度和相位差异, 传声器阵列可以实现对声源的有效定位<sup>[1,2]</sup>。球形传声器阵列因其完全对称的结构, 可在三维空间内实现有效的声源定位, 因此受到研究者的广泛重视。

基于可控波束响应的声源定位技术对传声器所接收到的声源信号滤波并求加权和来形成波束, 调节传声器阵列的波束方向, 在整个接收空间内扫描, 最终使波束输出功率最大的点就是声源的位置<sup>[3]</sup>。可控波束响应的定位技术主要有延迟累加波束算

收稿日期: 2015-05-25; 修回日期: 2015-08-20

基金项目: 国家级大学生创新训练计划(SZDG2014002)、南京邮电大学自然科学基金(NY214049)、江苏省高校自然科学面上项目(15KJD510001)资助。

作者简介: 章康宁(1994—), 男, 江苏南京人, 研究方向为通信工程、信号处理。

通讯作者: 孙科学, E-mail: sunkx@njupt.edu.cn

法和球谐函数波束算法<sup>[4,5]</sup>。延迟累加波束算法较为简单,通过利用阵列单元间的时延信息进行相应的加权累加,形成单一方向的波束,但波束的主瓣宽度较大,导致定位的分辨率较低。针对球形传声器阵列的特点,可以借助球谐函数对入射声波做正交分解<sup>[6]</sup>,并对阵列单元的信号做空间域上的变换加权,最终得到特定的波束。当球谐函数展开阶数较高时,可控波束响应的主瓣较窄,使定位结果较为精确。

球谐函数的展开阶数  $N$  与传声器个数  $M$  需要满足一定的关系<sup>[7]</sup>:

(1) 传声器等角分布:即球坐标中的水平角和仰角在相邻两个传声器是固定的,此时  $N$  与  $M$  的关系为:  $M=4(N+1)^2$ 。这种分布的优势在于其角度差是固定的,有利于球阵列的等角转动。其劣势在于需要较多的传声器并有随之而来的信息处理量。

(2) 传声器高斯分布:即球坐标中的水平角在相邻两个传声器是固定的,而仰角并不固定,但仰角方向上的相邻两个传声器的距离是相等的,此时  $N$  与  $M$  的关系为:  $M=2(N+1)^2$ 。相对于等角分布,其优势是减少了传声器的个数,但由于在仰角方向的角度差不相等,会带来测量的不便。

(3) 传声器均匀分布:即球坐标中的水平角和仰角在相邻两个传声器都不是固定的,但两传声器间的距离都相等,此时  $N$  与  $M$  的关系为:  $M \geq (N+1)^2$ 。这种分布的优点是更大地减少了传声器的个数,但由于两传声器的角度差不相等,会导致测量不方便,而且此种分布下的 FFT 计算也会受到影响。根据以上所述,均匀分布的阵列虽然测量时不太方便,但所需要的传声器单元数最少,因此本论文讨论的是均匀分布的阵列单元。尽管球谐函数展开波束算法需要进行全局搜索,运算量较大,且要依赖于声源信号的频谱特性,但其有较高的分辨率,能精确定位声源,而且还能适用于多声源场合,实际应用较为广泛。

本文针对可控波束响应定位算法在球形传声器阵列中的应用展开讨论,分析和推导了在球谐域实现可控波束响应的算法流程,并通过仿真和实验验证了定位结果的有效性。

## 1 基于可控波束响应的定位算法

理想情况下,由可控波束响应算法得到的波束是一个在空间上的冲激信号,此时的波束主瓣无限小,旁瓣为零,则波束的分辨率可以无限精确,用

公式表示为

$$b_{\text{ideal}}(\theta, \phi, \theta_0, \phi_0) = \delta(\theta - \theta_0) \delta(\phi - \phi_0) \quad (1)$$

其中:  $\theta$ 、 $\phi$  分别表示可控波束方向的仰角和水平角,  $\theta_0$ 、 $\phi_0$  分别表示入射波射向阵列方向的仰角和水平角。

由于球谐函数满足正交完备性<sup>[6]</sup>:

$$\int_0^{2\pi} \int_0^\pi Y_n^m(\theta, \phi) \bar{Y}_n^{m'}(\theta, \phi) \sin \theta d\theta d\phi = \delta_{nn'} \delta_{mm'} \quad (2)$$

其中,  $Y_n^m(\theta, \phi) = \xi_n^m P_n^m(\cos \theta) e^{im\phi}$ , 是球谐函数,  $\xi_n^m = \sqrt{\frac{2n+1}{4\pi} \frac{(n-|m|)!}{(n+|m|)!}}$ ,  $P_n^m(\cos \theta)$  为连带勒让德函数,任意函数都能等效为球谐函数的线性组合,即

$$f(\theta, \phi) = \sum_{n=0}^{\infty} \sum_{m=-n}^n C_n^m Y_n^m(\theta, \phi) \quad (3)$$

对式(1)做球谐函数展开,再根据冲激信号的积分特性,可得:

$$b_{\text{ideal}}(\theta, \phi, \theta_0, \phi_0) = \sum_{n=0}^{\infty} \sum_{m=-n}^n \bar{Y}_n^m(\theta_0, \phi_0) Y_n^m(\theta, \phi) \quad (4)$$

实际应用中,阶数不可能为无限大,取阶数为  $N$ ,则波束响应为:

$$b_N(\theta, \phi, \theta_0, \phi_0) = \sum_{n=0}^N \sum_{m=-n}^n \bar{Y}_n^m(\theta_0, \phi_0) Y_n^m(\theta, \phi) \quad (5)$$

又因为  $\bar{Y}_n^m(\theta_0, \phi_0)$  是声源的信息,在实际中无法获得,但可以利用球谐函数的正交性,得到:

$$\bar{Y}_n^m(\theta_0, \phi_0) = \frac{A_n^m}{R_n(a, r_0)} \quad (6)$$

其中

$$A_n^m = \int_0^{2\pi} \int_0^\pi \Phi_a(\theta, \phi, \theta_0, \phi_0) \bar{Y}_n^m(\theta, \phi) \sin \theta d\theta d\phi \quad (7)$$

其中,  $\Phi_a(\theta, \phi, \theta_0, \phi_0)$  是速度势,实际无法测量。但速度势含有幅值和相位的信息,等效于信号的频域信息,可以使用信号频域信息代替式中的速度势。

式(6)中  $R_n(a, r_0)$  是不同球阵列结构的参数,常用表达式如表 1 所示。

表 1 不同球阵列结构的参数  
Table 1 Parameters of different spherical arrays

	实心球体	透明球体
平面波	$i^n \cdot [j_n(ka) - a' \cdot h_n^{(2)}(ka)]$	$i^n \cdot j_n(ka)$
球面波	$i \cdot h_n^{(2)}(kr_0) [j_n(ka) - a' \cdot h_n^{(2)}(ka)]$	$i \cdot h_n^{(2)}(kr_0) j_n(ka)$

近场情况下,需要事先知道  $r_0$ , 即声源距离,但实际难以获得,需要进行粗略估计。

另外在数值计算中,公式(7)无法计算积分,但可以通过近似得到

$$A_n^m = \sum_q w_q \Phi_a(x_0, x_q) \bar{Y}_n^m(x_q) \quad (8)$$

其中:为便于书写,令  $(x) = (\theta, \phi)$  为  $w_q$  是把积分近似为累加后得到的权重,其计算可由以下得到:

$$\overline{G}w=e \quad (9)$$

其中： $w$  是包含数值积分的权重的向量， $e$  是一个全 1 向量， $\overline{G}$  被称为 reproducing kernel 的矩阵，其元素为：

$$G_N(x,y)=\sum_{n=0}^N \sum_{m=-n}^n Y_n^m(x)\overline{Y}_n^m(y) \quad (10)$$

最后，结合以上讨论和公式，计算某个角度下的响应公式为

$$b_N(x)=\sum_q \left[ \frac{1}{\sum_{n=0}^N R_n(a,r_0)} \sum_{m=-n}^n w_q \overline{Y}_n^m(x_q) \cdot Y_n^m(x) \right] \cdot \Phi_a(x_0, x_q) \quad (11)$$

根据式(11)，搜索空间的所有水平角和仰角(x)，可以得到全空间不同方向上的波束响应，响应最大值对应的角度即为声源的入射方向。

## 2 定位算法仿真

理论上传声器的个数越多，球谐函数展开的阶数越高，则最终的定位结果的分辨率就越高。为保证仿真精度，选择展开阶数为 6 阶。采用的阵列为以坐标原点为球心、半径为 0.5 m 的球阵列(选择较大半径的球阵列是因为当球谐函数展开的阶数较高时，阵列半径和声信号的主要频谱成分之间需要满足一定的约束条件<sup>[8]</sup>。这是由球谐函数近似而产生的正交误差所决定的。传声器位置采用均匀分布(两传声器间隔接近相等，位置计算采用文献[9]提到的非线性规划方法)，单元个数为 64(满足传声器个数和展开阶数的关系)，如图 1 所示。

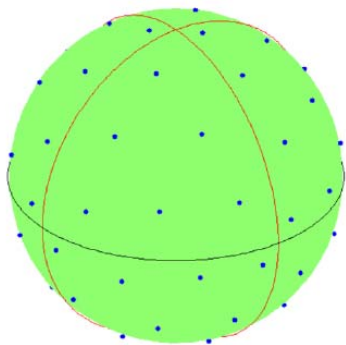


图 1 传声器单元分布图

Fig. 1 Element distribution on a spherical microphone array

假设声源位于水平角为  $80^\circ$ 、仰角为  $30^\circ$ (即球坐标  $\theta$  为  $60^\circ$ )，距离坐标原点 2 m 的地方。声源信号为播音员标准测试语音信号，采样率为 44100 Hz，各传声器接收信号可以利用镜像模型计算声源到达各传声器的传递函数来获得。以下所有仿真环境如未说明即为自由场，且不添加任何噪声。根据

Spherical Harmonic Beam(SHB)的算法流程，定位的结果如图 2 所示。

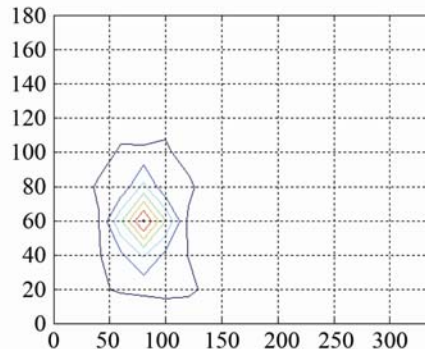


图 2 SHB 算法的仿真结果

Fig. 2 Numerical calculation results of SHB algorithm

由于实际环境中混响和噪声的影响，需要对算法做鲁棒性分析(假设房间中心在原点，长宽高分别为 8、6 和 4 m，声源及阵列位置保持不变)。评测六面墙相等的反射系数从小到大变化导致的定位方向的偏差，即混响的影响。由于全空间搜索间隔为  $20^\circ$ ，所以偏差以  $10^\circ$  近似(下文仿真搜索间隔相同)，结果如表 2 所示。

表 2 混响对 SHB 算法的影响  
Table 2 Effects of reverberation on SHB algorithm

反射系数	0.1	0.3	0.5	0.7	0.8	0.9
水平角误差/( $^\circ$ )	0	0	0	0	0	0
仰角误差/( $^\circ$ )	0	0	0	0	0	40
					出现伪峰	出现多个伪峰

较高混响下，不同方向上会出现伪声源，这些峰值是由各反射面的反射造成的。由于声源和阵列均处于房间中心附近，距离墙壁较远，混响对定位的影响不是很明显，所以当混响较高的时候才出现定位出错。在一般混响下，定位结果可有效反映实际声源的位置。

信噪比(Signal Noise Ratio, SNR)从大变小导致定位方向的平均偏差(多次偏差的绝对值求平均)，可反映噪声的影响。结果如表 3 所示。

表 3 噪声对 SHB 算法的影响  
Table 3 Effects of noise on SHB algorithm

SNR/dB	20	10	0	-10
水平角误差/( $^\circ$ )	0	0	0	0
仰角误差/( $^\circ$ )	0	0	0	0

可见在一般噪声影响下，算法的定位误差较小。在较强噪声背景下，定位的偏差也不是很大，这是因为仿真加入的噪声是直接添加到信号接收端，并没有附加位置的信息。假如噪声是来源于一

固定位置的噪声源, 定位的等高线图会出现声源方向和噪声源方向上的多个峰值。

### 3 实验与分析

由于受到实验条件的限制, 实验中无法用到 64 个传声器, 而是用 8 个传声器构建了一个基本的球阵列, 三维阵列实际的边长为 0.16 m, 半径约为 0.14 m。实验的信号采集设备为丹麦 B&K 公司的 B&K Pulse 3560D, 采样频率为 65536 Hz。声源为 TANNOY 同轴扬声器, 测试信号使用白噪声与播音员标准测试语音信号。实验环境为混响时间约 0.3 s 的视听室。实验时, 可以旋转阵列三脚架的旋钮上下转动或者左右转动, 等效于改变声源的方位而不需要移动声源。但从三脚架刻度得到的水平角和仰角并不能完全等同于坐标的水平角和仰角, 需要进行几何上的转换。另外, 转动时, 由于转动点不在阵列中心, 需要实时地对声源进行微调, 使声源与阵列中心保持在同一水平面上, 并且确保相互距离不变。实验中信噪比均高于 40 dB, 实现重复次数为 5 次。

#### 3.1 实验 1(测试信号: 白噪声)

声源距离阵列 1 m, 定位频率为 655 Hz 时的定位偏差如表 4 所示, 表 4 中定位误差小于  $10^\circ$  时记为 0。

表 4 噪声源与球阵列形成不同角度时的定位偏差  
(声源距离球阵列 1 m)

Table 4 Localization errors at different azimuth angles of the noise source relative to spherical array(distance 1 m)

水平角与仰角/(°)	定位误差/(°)
0, 0	0, 10
0, -30	-10, 0
0, -60	0, -10
-30, 0	0, 10
-60, 0	0, 10
-90, 0	-10, 10
-33.7, -25.7	0, 10
-49.1, -48.6	10, 0
-63.4, -14.5	-10, 10
-73.9, -25.7	-10, 0

#### 3.2 实验 2(测试信号: 语音信号)

声源距离阵列 1 m, 定位频率为 655 Hz 时的定位偏差如表 5 所示, 表 5 中定位误差小于  $10^\circ$  时记为 0。

从实验结果可以看出, 除了极个别情况, 球阵列可控波束响应算法的定位误差都在  $10^\circ$  以内, 考

虑到阵列单元的个数较少, 只能做低阶球谐展开, 这是一个很高的定位精度。此外, 算法对不同种类信号有很强的适应性。

表 5 语音声源与球阵列形成不同角度时的定位偏差  
(声源距离球阵列 1 m)

Table 5 Localization errors at different azimuth angles of the voice source relative to spherical array(distance 1 m)

水平角和仰角/(°)	定位误差/(°)
0, 0	0, 0
0, -30	-10, 10
0, -60	0, -10
-30, 0	0, 10
-60, 0	0, 10
-90, 0	-10, 10
-33.7, -25.7	0, 0
-49.1, -48.6	10, 0
-63.4, -14.5	-10, 20
-73.9, -25.7	0, 0

### 4 结论

本文讨论在球形传声器阵列中利用球谐函数正交展开, 使用可控波束响应方法进行声源定位的算法, 给出了算法的理论推导流程, 并通过仿真和实验验证了算法的有效性。实验结果表明: 球阵列可控波束响应算法在具有一定混响的声学场景中有较高的定位效果。

#### 参 考 文 献

- [1] Benesty J, Chen J D, Huang Y T. Microphone array signal processing[M]. Springer, 2008.
- [2] Koretz A, Rafaely B. Dolph-Chebyshev beam pattern design for spherical arrays[J]. IEEE Trans. Signal Processing, 2009, 57(6): 2417-2420.
- [3] Petersen S O. Localization of sound sources using 3D microphone array, University of Southern Denmark, 2004.
- [4] Thushara A, Prasanga S. Analysis and synthesis of room transfer function over a region of space using distributed spherical microphone arrays[J]. J. Acoust. Soc. Am., 2012, 131(4): 3209-3215.
- [5] Epain N, Jin C T. Independent component analysis using spherical microphone arrays[J]. Acta Acustica united with Acustica, 2012, 98(1): 91-102.
- [6] Park M, Rafaely B. Sound-field analysis by plane-wave decomposition using spherical microphone array[J]. J. Acoust. Soc. Am. 2005, 118(5): 3094-3102.
- [7] Rafaely B. Analysis and Design of Spherical Microphone Arrays[J]. IEEE Trans. Speech, Audio Proc. 2005, 13(1): 135-143.
- [8] Li Z, Duraiswami R. Flexible and optimal design of spherical microphone arrays for beamforming[J]. IEEE Trans. Audio, Speech, Lang. Proc. 2007, 15(2): 702-714.
- [9] Fliege J, Maier U. The distribution of points on the sphere and corresponding cubature formulae[M]. IMA J. Numer. Anal. 1999, 19(2): 317-334.