

2020 年第十三届“认证杯”数学中国 数学建模网络挑战赛第一阶段论文

基于麦克风阵列的声源定位方法研究 摘要

本文针对麦克风阵列接收声源信号以获得声源实时位置及如何测量大厅几何参数的问题进行分析研究,通过分析比较,最终确定以十字形麦克风阵列为核心,以声达时间差和能量分析为方法的优化模型。

针对于问题一,我们将从两个方面拆解问题并给出分析过程,其中包括麦克风阵列的选择和优化定位模型的建立。

第一个方面是关于麦克风阵列的选择,我们首先建立了一维线性阵列模型,在对线性模型的计算改进的基础上建立了二维面阵模型其中包括方形阵列模型、等边三角形阵列模型、圆形阵列模型。接着我们对三个模型进行了合理的理论推导与证明,通过实验计算出原始位置与测定位置之间的误差,并分析其结果得出平均误差,证明出方形阵列为最适宜阵列模型。最后结合实际应用与实现的难易程度进行分析,且由于方形阵列可以近似看为十字形阵列,因此本文将十字形阵列确定为最优阵列的选择。

第二个方面是关于定位模型的建立,由于考虑到声源传播模式可以分为近场与远场,我们对远近场区分的常用公式进行了声源在近远场的比对分析,但因为在本题中可以忽略远近场的影响,于是在理想状态下,可以根据信号时延性进行计算以得到声源位置。阵元间距问题,我们通过给出一定的声源频率,结合声速与波长的关系,给出合适的阵元间距。最后,通过声源的位置信息,确定了以四个麦克风数量为基准的麦克风矩阵定位模型,再进行逐步分析增加,最终确定麦克风最优个数。

针对于问题二,我们需要解决如何在大厅几何参数未知的情况下,测定大厅的体积及实时确定人的位置。由于大厅几何参数未知,我们首先利用声音会逐渐减弱的特性确定了在仅有扩散衰减情况下的声音衰减模型。通过分析确定无法忽略吸收衰减和散射衰减对模型的影响,于是对模型进行了优化提出了衰减系数 α 。接下来,我们设计了一个实验通过利用 matlab 进行 power 幂函数方程拟合以测定在室内环境下衰减系数的值。其次,由于未确定大厅几何参数会对定位模型造成一定的影响,因此以声达时间差来确定大厅的定位模型无法精确得出大厅的大小,所以我们选择二选一声音强度(能量法)来帮助对未知大厅大小的定位模型的建立。通过得出声强级和传播距离的关系,设计出合理的麦克风安置方法。考虑到麦克风接收声源信息的误差会导致不同麦克风对声源位置的估计产生偏差,因此使用蒙特卡罗算法进行近似处理,从而通过计算可以得到的合理声源位置。

最后,本文对所有问题的模型进行了客观合理的评价,并且对建立的数学模型进行了改进分析,使得模型的应用面更加广泛。

关键词: 麦克风阵列 声源定位 时延性 声音衰减模型 蒙特卡罗算法

ABSTRACT

In this paper, the problem of microphone array receiving sound source signal to obtain the real-time position of sound source and measuring hall geometric parameters is analyzed and studied. Through analysis and comparison, the cross microphone array model and sound energy attenuation model are finally determined as the optimization model.

For problem 1: we first set up a one-dimensional linear array model, and on the basis of the improvement of the linear model, we set up a two-dimensional planar array model, including square array model, equilateral triangle array model and circular array model. Then we have carried on the reasonable theoretical proof and derivation to the three models, calculated the error of the original position measurement position through the experiment, and proved that the theoretical result is that the square array is the most suitable array model. Then, combining with the practical application and the degree of implementation, it is analyzed that the square array can be approximately regarded as a cruciform array, and the cruciform array is finally determined to be the optimal model.

Secondly, considering that the propagation mode of sound source can be divided into near field and far field, we carry out a comparative analysis of sound source in near field according to the commonly used formula of near field division. However, since the influence of near field can be ignored in this problem, the position of sound source can be obtained by calculating according to the time ductility of signal in an ideal state. Finally, according to the information required to determine the location of the sound source, the number of four microphones is determined as the benchmark, and the number of microphones is gradually analyzed and increased to determine the optimal number of microphones.

For question two: since the geometric parameters of the hall are unknown, we first determine the sound attenuation model in the case of only diffusion attenuation by using the feature that sound will gradually weaken. The influence of absorption attenuation and scattering attenuation on the model can not be ignored. Next, we designed an experiment using matlab to perform power function equation fitting to determine the value of attenuation coefficient in the indoor environment.

Secondly, because the undetermined room geometric parameters will have a certain impact on the positioning model, and it is impossible to accurately obtain the size of the room by determining the positioning model of the room with the sound arrival time difference, we choose to use the sound intensity (energy method) to help determine the size of the room. Finally, this paper makes an objective and reasonable evaluation of the models of all problems, and makes an improved analysis of the established mathematical models, so as to make the models more widely applied.

Key words: microphone array cruciform model, ductile sound attenuation model

一、问题重述

全向麦克风作为声传感器可以 360° 收集声音信号，通过对其收到的声音进行对比分析，可以得到声源的具体位置及其他声源信息。将多个麦克风安装到一个刚性支架上，对支架可以进行水平推动或是旋转，可以更好地接收声源信号，但单个麦克风无法接收声源信号。由于麦克风放置位置不同，故其接收的声源信号也有细微差别。

第一个问题可以分成两个方面来分析，其中包括麦克风阵列的选择和优化定位模型的建立。

将上述模型放在实际应用中，选择一个大厅放置建立的麦克风树阵列，希望通过声源信号的收集来进行测量以获得所需信息。根据已知和所求信息的不同，这个问题可以有不同复杂程度。但是，麦克风数量的不同所实现的难易程度及成本也有很大的区别。希望能在成本尽量低且能达到使用要求的情况下建立模型，并给出每支麦克风的相对位置以及相对于地面的高度。

第二个问题我们需要解决如何在大厅几何参数未知的情况下，测定大厅的体积及确定人的位置。

二、问题分析

对于问题一：已知平面是矩形的大厅且已知大厅轮廓的准确尺寸，在理想情况下，如何建立起一个在刚性的枝形架子上的麦克风树，忽略了近远场的影响及声音的传输损失，通过分析各种阵列的优缺点以选择建立最优的一个成本尽量低、而且可以达到使用要求麦克风阵列和确定阵元的数量给出每支麦克风的相对位置以及相对于地面的高度，能实时确定发出清晰声源（脚步声）的人的位置。

对于问题二：现实中往往并不能准确知道大厅的几何参数，并且无法忽略声音传输的损失。由于以上原因，我们可以首先在大厅体积未知的情况下，在问题一解决的基础上，用已经建立的模型下寻找借助声源传播求解大厅的几何参数，然后将建立一个新的模型，对结果分别进行预测，并将结果进行比较，于是可以得出在非理想情况下，如何实时确定人的位置。

三、模型假设

1. 在远场模型中，因为声源信息到麦克风阵列的幅度差较小，忽略不计。
2. 忽略鞋和体型等因素，一个正常人脚步声的频率在 $20\text{ Hz} \sim 40\text{ Hz}$ 。
3. 不考虑环境的噪声影响以及室内混响的影响。
4. 在第一问中，忽略远场和近场对麦克风阵形的影响。
5. 麦克风的品质（即对信息的收集能量）不随实验时间的进行而下降。
6. 持续行走发出的脚步声的声强级（分贝）不发生改变，始终为 50 分贝。

四、定义与符号说明

符号	说明
L	相邻麦克风阵元的长度
λ	声波的波长
r_0	声源到阵元参考点的距离

P	听阈声压
L_p	声压级
W	声功率
L_w	声功率级
I	声强

表 1 定义与符号说明

五、模型的建立与求解

5.1 麦克风阵列最优化模型

麦克风阵列的排列方式有许多种，可以依据阵列的维度将其分为线性阵列、平面型阵列、立体型阵列。在这里我们主要对比讨论线性阵列与平面型阵列中的正方阵与等边三角形阵列及圆形阵列以寻求最优化模型。

5.1.1.1 线性阵列

线性麦克风由 m 个麦克风以相等的间距均匀地摆放成一条直线，阵元间距一般是其波长的一半，由于是在远场模型中只需考虑声源信息的时延。如图 1 所示：

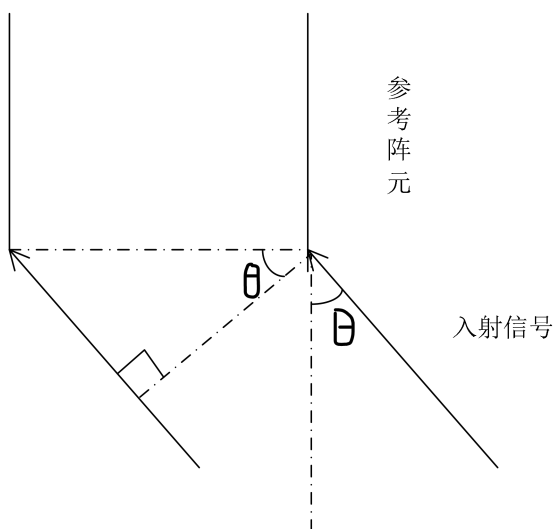


图 1 麦克风线性阵列

可以得出其的声源信息的时延式：

$$\tau_i = \frac{\langle \bar{p}_i, \bar{r} \rangle}{v}$$

\bar{p}_i 为第 i 个阵元的位置， \bar{r} 表示声源位置。

一维线性阵列易于实现且符合成本低的要求，但是声源传递过程中存在失真且只能给出声源的入射角，无法满足实时定位的需要。

5.1.1.2 平面阵列——方形拓扑阵列

如下图 2 所示为方形拓扑阵列，实心点处表示麦克风放置的位置，把一定区域划分成面积为同样大小 3 的正方形方格。这种方法得到的麦克风分布位置距离一定且在一定区域内分布地比较均匀。

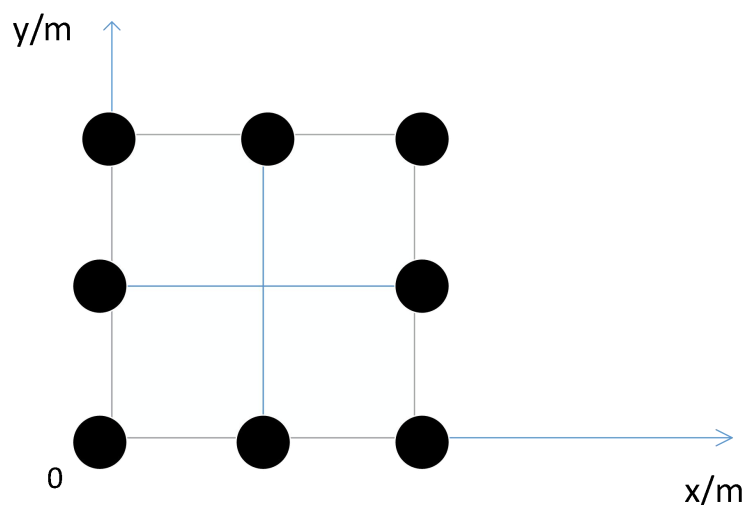


图 2 方形拓扑结构

绝对误差 E 即为原始位置 (x, y) 与测量位置 (x_i, y_i) 之间的距离，公式如下：

$$E = \sqrt{(x-x_i)^2 + (y-y_i)^2}$$

定位结果如下表 2：

原始坐标	方形拓扑阵列	
	测量位置	绝对误差
(1.25,0.00)	(1.35,0.15)	0.18
(1.50,0.00)	(1.50,0.00)	0.00
(1.75,0.00)	(1.81,0.11)	0.12
(1.25,0.15)	(1.34,0.31)	0.18
(1.50,0.15)	(1.36,0.91)	0.15
(1.75,0.15)	(1.83,0.15)	0.08
(1.25,0.50)	(1.31,0.62)	0.14
(1.50,0.50)	(1.50,0.50)	0.00
(1.75,0.50)	(1.82,0.64)	0.16
平均误差	0.112	
最大误差	0.18	

表 2 方形拓扑阵列定位结果

5.1.1.3 平面阵列——等边三角形拓扑阵列

等边三角形拓扑阵列结构是将等边三角形的三个顶点作为圆心，这样可以得到两个与彼此相交的三个圆，当等边三角形的边长是圆的半径的一定倍数时，三个圆相互覆盖的面积最大的重叠部分为最小，通过这样的方式可以用最少数量的圆来实现空间内的基

本区域的覆盖。通过上述方法可以得到如下图 3 所示的等边三角形拓扑阵列结构，其中把三角形的顶点作为麦克风放置点，将圆用形状十分贴合于圆并且之间相互排列无缝连接的正六边形替代，十分适合于一定区域的划分。这种拓扑阵列是可以用最少的点实现对一定区域的全方面的覆盖。

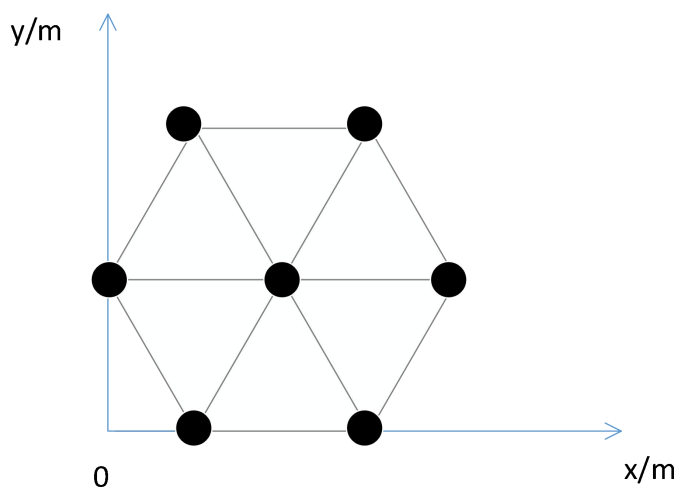


图 3 等边三角形拓扑结构

根据上文给出的绝对误差 E 的公式，可以计算出相应的定位结果如下表 3：

原始坐标	等边三角形拓扑结构	
	测量位置	绝对误差
(0.75,0.00)	(0.74,0.18)	0.18
(1.25,0.00)	(1.48,0.19)	0.30
(0.15,0.50)	(0.25,0.50)	0.10
(1.00,0.50)	(1.00,0.61)	0.11
(1.50,0.50)	(1.64,0.62)	0.18
(0.75,0.93)	(0.76,0.82)	0.11
(1.25,0.93)	(1.23,0.82)	0.11
平均误差	0.156	
最大误差	0.30	

表 3 等边三角形拓扑结构定位结果

5.1.1.4 平面阵列——圆形拓扑阵列

将一个半径为 R 的圆按照一定的角度等分成 m 份，在每个等分点的位置放置一个麦克风进行信号的收集。这是一种适合圆形区域或方形区域的拓扑结构，此结构的原理是取一定区域的中心点，将其作为第一个选定参考点，然后依据一定区域的面积和定位精度选定递增半径 R 。使认定的第一个参考点作为圆心，以 R 为半径画圆，取边长均为 $\sqrt{2}$ 的内接四边形的四个顶点为参考点，做第一个圆的半径为 $2R$ 的同心圆，其内接 8 边形边长为 $\sqrt{2}$ ，选取 8 边形的八个顶点作为参考点。如下图 4 所示：

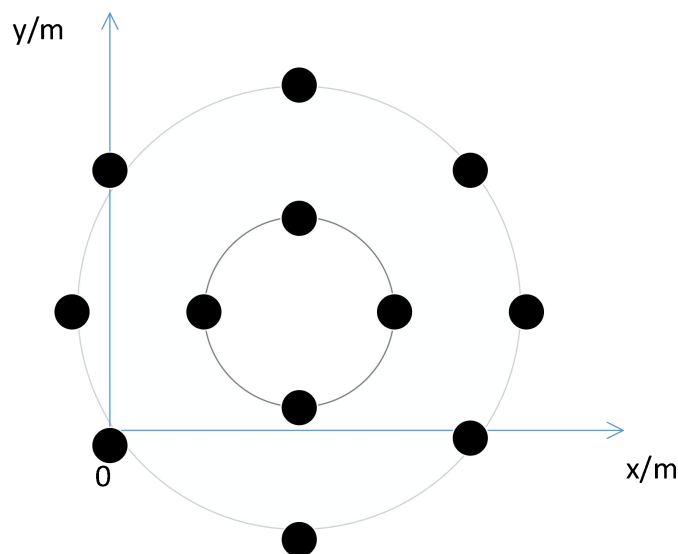


图 4 圆形拓扑阵列

根据上文给出的绝对误差 E 的公式，可以计算出相应的定位结果如下表 4：

原始位置	圆形拓扑结构阵列	
	测量位置	绝对误差
(0.50,0.00)	(0.35,0.20)	0.25
(1.00,0.00)	(1.12,0.17)	0.21
(1.50,0.00)	(1.46,0.21)	0.21
(1.00,0.15)	(1.11,0.20)	0.12
(0.15,0.50)	(0.15,0.50)	0.00
(0.75,0.50)	(0.65,0.58)	0.13
(1.25,0.50)	(1.20,0.50)	0.05
(1.75,0.50)	(1.77,0.62)	0.13
(1.00,0.70)	(1.00,0.66)	0.04
(0.50,0.93)	(0.35,0.80)	0.20
(1.00,0.93)	(1.17,0.83)	0.20
(1.50,0.93)	(1.65,0.80)	0.20
平均误差	0.145	
最大误差	0.25	

表 4 圆形拓扑结构定位结果

经过上述四种方法的对比利弊及平均误差的计算分析，可以得出方形拓扑阵列误差较小且易实现更为适宜，又由于方形阵列可以近似看作十字形阵列，如下图 5 所示。故本文采用十字型阵列。

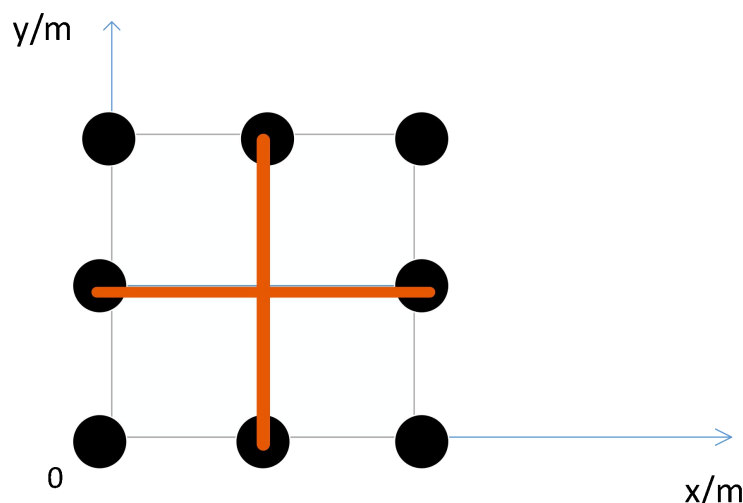


图5 十字型阵列结构

5.1.2 阵列间距

上文中已经确定了使用十字型阵列，接下来将通过分析确定阵列间距。确定阵列间距应该从以下两个方面来进行考虑：①所需定位的声源信号的频率；②麦克风系统内部所选择的声源信号的采样频率。在 1924 年，美国物理学家奈奎斯特已经提出这样的理论：当麦克风系统所选择的采样频率 F_s 大于声源信号最高频率的两倍时，经过采样转变成的数学信号就会十分完整的保留声源原始发出的信号。通过以上理论可以得出以下结论：声源信号的波长应该大于或等于两倍的相邻阵元间的最小距离。

通过资料，我们知道正常人的脚步声频率一般在 20Hz~40Hz 之间

波长 λ 与声源频率 f 及声速 v 的关系式如下：

$$\lambda = \frac{v}{f}$$

将已知数据代入，由此可以求得波长：

$$\lambda_{\max} = \frac{v}{f} = \frac{340\text{m/s}}{20\text{Hz}} = 17\text{m}$$

$$\lambda_{\min} = \frac{v}{f} = \frac{340\text{m/s}}{40\text{Hz}} = 8.5\text{m}$$

又需满足声源信号的波长大于或等于两倍的相邻阵元间的最小距离，即：

$$d_{\max} = \frac{\lambda_{\max}}{2} = 8.5\text{m}$$

$$d_{\min} = \frac{\lambda_{\min}}{2} = 4.25\text{m}$$

所以相邻麦克风阵元之间的最小距离 d 应在 4.25~8.5 m 之间

5.1.3 阵元的数目确定

阵元数目的确定应该选择从以下两个方面进行考虑：①麦克风阵列系统需要的定位声源的精度。②所使用的麦克风阵列模型体积。虽然从理论上说，当阵列的数目越多时，

声源定位系统的精度就越高，但是从实际应用上讲，当阵元数目过多时，一定程度上增加了麦克风阵列模型的实际体积，将模型实现的难度大大提高；从另一个角度上讲，阵元数目与定位算法的复杂程度有着一定的联系，阵元数目越多则声源定位的算法就越复杂，直接影响了系统的实时性。于是，在实时性系统中，当仅仅需要获取声源的方位角信息时，阵元的数目一般不超过四个。但在本模型中，还需考虑俯仰角及高度等信息，故阵元数目需要超过四个。

5.2 近场和远场模型区分

由上文的麦克风阵列排布模型可得出，麦克风阵列排布成十字形阵列最为理想。在第一阶段问题中，虽然已知了大厅的轮廓尺寸，但是麦克风的放置点坐标无法精确，就无法确定声源（该人走路脚步声）到各个麦克风的距离，因此需要将脚步声的传播模型分成两种，近场和远场模型。目前，对于近场和远场模型的划分，常用公式为：

$$r = \frac{2L^2}{\lambda} \quad (5.2.1)$$

在现实生活运用中，经常把声源到阵元参考点的距离 r_0 与上式中的所求的比较，当 $r_0 < r$ 时为近场模型，当 $r_0 > r$ 时为远场模型。在近场模型中，因为声源到麦克风阵列的距离较近，声源信号到麦克风阵列每个阵元的幅度差大，阵列波前应该使用球面波前模型。在远场模型中，因为声源到麦克风阵列的距离较远，声源信号到麦克风阵列每个阵元的幅度差较小，可忽略不计，阵列波前应该使用平面波模型。近场和远场模型如图 6：

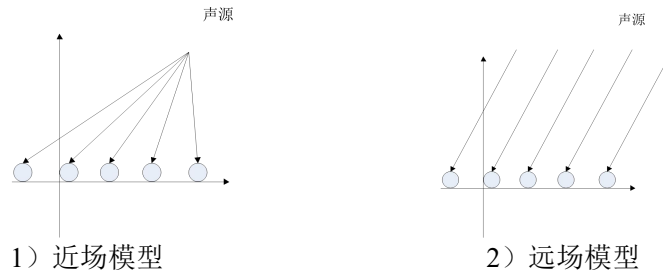


图 6 近场与远场模型

在本题中，声源是可移动的单声源（一个人的脚步声），忽略鞋和体型等因素，一个正常人的脚步声都在 20Hz~40Hz，声速取为 $v=340\text{m/s}$ ，由频率和波长的关系：

$$\lambda = \frac{v}{f}$$

可得到脚步声的波长 λ ，进而由式 (5.2.1) 求得 r 。由 r 与 r_0 比较大小，判断近场或者远场模型，故在本题中需要分两种模型讨论。

5.2.1 近场的信号模型

当信号模型是近场模型时，采用球面波模型更加精确一些，该问题中，确定声源（脚步声）是一个三维空间里确定声源的问题，设有 m 个麦克风阵元，分别是 $m_0, m_1, m_2 \dots m_{m-1}$ ，其中 m_0 麦克风阵元在原点 O 处，坐标是 $m_0 (0, 0, 0)$ ，声源坐标设为 $P(r_i, \theta_i, \Phi_i)$ ，其中 r_i 代表的是声源和麦克风阵元 m_i 的距离， θ_i 表示声源信息的方位角， Φ_i 代表的是声源信息的俯仰角。近场的信号模型如下图 7 所示：

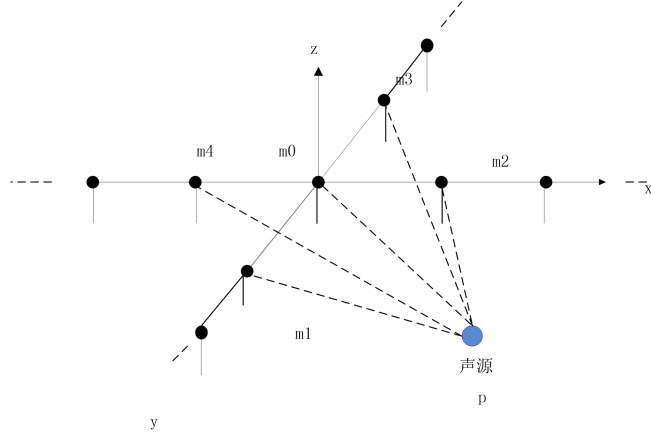


图7 近场信号模型

在第一个问题中因为墙壁、地面、天花板材料全是光滑的大理石，光滑的大理石的吸声系数近似于零。假设忽略环境的噪声影响以及室内混响的影响，则各个麦克风阵元 m_i 接收到的信号为 $z_i(t)$ ， $z_i(t)$ 有以下关系：

$$z_i = \gamma_i(r_i, \theta_i, \Phi_i) * p(t)$$

其中 $\gamma_i(r_i, \theta_i, \Phi_i)$ 代表的是声源到各个麦克风阵元 m_i 的衰减因子。

5.2.2 远场的信号模型

当信号模型是远场模型时，采用平面波前模型更加精确一些，设有 m 个麦克风阵元，分别是 $m_0, m_2, m_3 \dots m_{m-1}$ ，设 θ_x 表示声源对于十字麦克风阵列中 x 向线性阵列的入射角度， θ_y 表示声源对于十字麦克风阵列中 y 向线性阵列的入射角度，相邻两麦克风的距离为 d ，各个麦克风阵元 m_i, m_j 接收声源信息的延迟时间为 Δ_{ij} 。远场的信号模型如下图8所示：

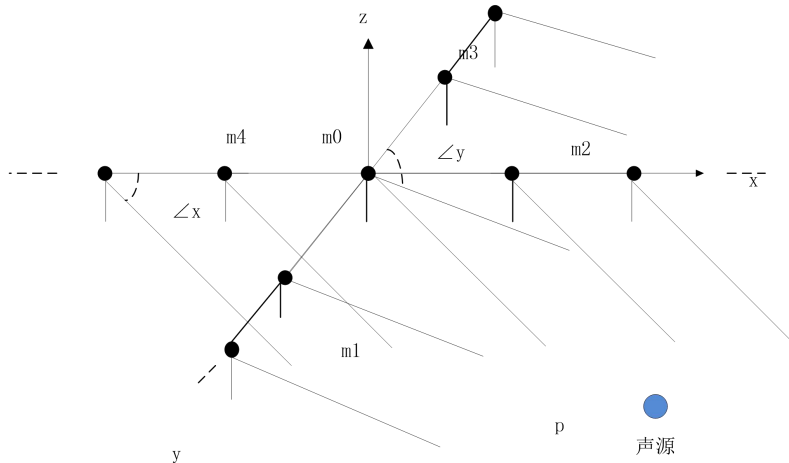


图8 远场信号模型

5.2.3 十字麦克风阵列模型

由上分析可知，要确定麦克风的相对位置，需要分近场跟远场模型讨论，在本题中，忽略远场和近场模型对阵形的影响，即无论近场、远场都选用十字形阵列，只考虑阵列波前平面波和球面波的不同。因为声源是来自于人的鞋与大厅地板的振动产生的，故声源始终来自于地面。脚步声到麦克风阵列各个阵元的传播简化线路图如下图9所示

(图中以 5 个阵元为例):

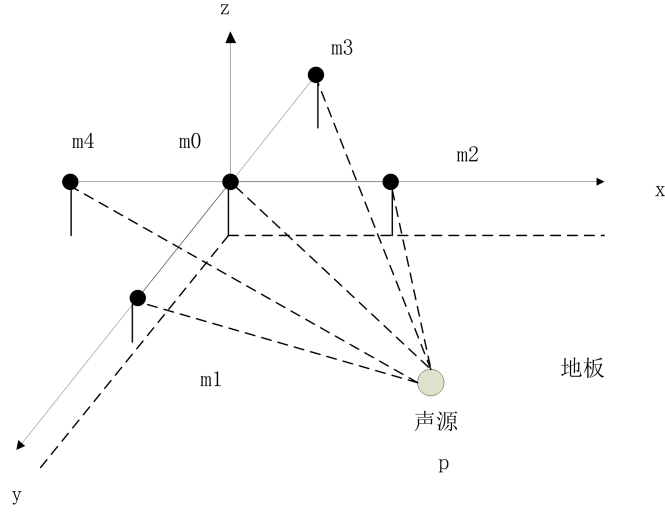


图 9 各个阵元的传播简化线路图

上图 9 是一个由五个麦克风阵元组成的麦克风阵列模型，五个麦克风阵元为同一相对高度，故取该平面为 XoY 面，在平面直角坐标系中，该图例中的五元十字形阵列中五个麦克风阵元的坐标分别为 $m_0 (0, 0, 0)$, $m_1 (0, a, 0)$, $m_2 (a, 0, 0)$, $m_3 (0, -a, 0)$ 和 $m_4 (-a, 0, 0)$ ，声源的坐标设为 (r, θ, Φ) ， r 表示声源信息到原点 O 的距离， θ 表示方位角， Φ 为俯仰角。 $S(r, \theta, \Phi)$ 与每个麦克风阵元的距离设为 r_i ，设各个的麦克风阵元 m_i 和 m_j 的传播延迟的时间为 \mathfrak{T}_{ij} ，声速 $v=340\text{m/s}$ ，根据声源与麦克风阵列的各个阵元的几何关系可得：

$$\begin{aligned} x^2 + y^2 + z^2 &= r^2 \\ x^2 + y^2 + z^2 &= r_0^2 \\ (x-a)^2 + y^2 + z^2 &= r_1^2 \\ (y-a)^2 + x^2 + z^2 &= r_2^2 \\ (x+a)^2 + y^2 + z^2 &= r_3^2 \\ (y+a)^2 + x^2 + z^2 &= r_4^2 \end{aligned}$$

因为原点 O 与麦克风阵元 m_0 在同一点，故 $r = r_0$ ，又根据各个麦克风阵元之间的时间延迟关系得以下关系式：

$$\begin{aligned} r_2 - r_1 &= \mathfrak{T}_{12} * v \\ r_3 - r_1 &= \mathfrak{T}_{13} * v \\ r_4 - r_1 &= \mathfrak{T}_{14} * v \end{aligned}$$

将下列各式代入上列各式中，其中 \mathfrak{T}_{ij} 是延迟的时间，由各个麦克风阵元接收到信号的时间差可测出，是已知量， r, θ, Φ 是未知量，联立式子可得：

$$\begin{aligned} r &\approx \frac{v}{2} \times \frac{\mathfrak{T}_{12}^2 + \mathfrak{T}_{13}^2 + \mathfrak{T}_{14}^2}{\mathfrak{T}_{13} - \mathfrak{T}_{12} - \mathfrak{T}_{14}} \\ \theta &= \arcsin\left(\frac{v}{2a} \sqrt{(\mathfrak{T}_{12} - \mathfrak{T}_{14})^2 + \mathfrak{T}_{13}^2}\right) \\ \Phi &= \arctan\left(\frac{\mathfrak{T}_{14} - \mathfrak{T}_{12}}{\mathfrak{T}_{13}}\right) \end{aligned}$$

由上式可知, r, θ, Φ 通过上式中的十字麦克风阵列模型可求, 因此由此可精确声源对麦克风阵列的相对位置。另外上面模型只给出了 5 个麦克风阵元作为一个麦克风阵列进行分析, 而在现实中, 如果仅仅需要求出声源的方位角信息时, 一般麦克风阵元的数目不超过四个。而本问题中, 是一个三维空间确定声源问题, 既要考虑方位角 θ , 又要考虑俯仰角 Φ , 因此在十字麦克风阵列模型中, 麦克风阵元需要超过四个。又麦克风阵元的数量还与大厅的轮廓有关, 故可根据需求在原来的十字麦克风阵列模型上增加麦克风数量, 具体增加麦克风阵元方式如下图 10 所示:

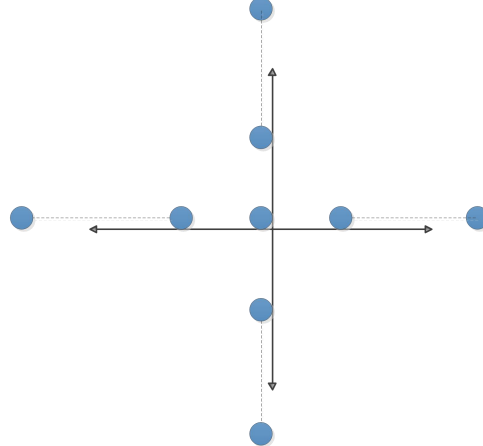


图 10 麦克风增加方式

5.3 声音能量衰减模型

5.3.1 声音传播机理的介绍

声音是由物体的振动产生的, 振动在弹性介质中的传播方式是以声波的形式。声波是一种机械波, 声波的传播空间就称为声场。人耳可以听到的声波的频率一般在 20 赫兹至 20000 赫兹之间。声音在我们的日常生活中必不可少, 因此我们需要描述声音的大小。国际上常用来描述声音强弱的物理量分别为: 声压, 声强, 声功率以及各自相对应的级。而响度和响度级则常用来描述声音的大小。

声压 P 就是大气压受到声波扰动后产生的变化, 即为大气压强的余压, 它相当于在大气压强上的叠加一个声压波扰动引起的压强变化。主要是用来度量声音的强弱, 通常我们所致的僧有复制记为声压, 我们将人耳所能听到的声压定义为听阈声压, 其值为 $P=2 \times 10^{-5} \text{ pa}$, 也称基准声压。与它相对应的声压级 L_p 为:

$$L_p = 20 \log \frac{P}{p_0} \quad (5.3.1)$$

声功率 W 声功率是声源在单位时间内向空间辐射声的总能量。是反映声源特性的物理量, 它的大小反应声源辐射声能的本领, 基准功率为 $W = 10^{-12} \text{ w}$, 其所对应的声功率级为:

$$L_w = 10 \log \frac{W}{w_0} \quad (5.3.2)$$

声强 I 是指声波平均能流密度的大小, 是以能量的方式度量声音的大小。声强即为理论上声音能量的定义, 其基准声强为 $I = 10^{-12} \text{ w/m}^2$ 。并且声强与声压的关系式可以表示为:

$$I = P^2 / \rho c \quad (5.3.3)$$

其中 ρ 为介质密度, c 为声音的传播速度。并且声强对应的声强级为:

$$L_1 = 10 \log \frac{I}{I_0} \quad (5.3.4)$$

5.3.2 声音衰减参数的分析

我们知道在现实生活中理想的介质是不存在的，因此声音在传播的过程中总会越来越小，我们将其称为声音的衰减。从生活经验我们可知：随着声音传播距离的增加，声音在不断的减小。而这一现象则是由于诸多因素引起的，我们溯本求源，将其总结为三种主要因素引起的衰减：扩散衰减，吸收衰减，散射衰减。

扩散衰减又称之为几何衰减，是指声波在传播过程中由于波阵面的面积不断扩大而引起单位面上能量的减少。扩散衰减主要与声源的波形和声束的形状有关，而与声音传播的介质无关。若不考虑其他衰减，则可以将扩散衰减理解为：扩散过程中声源的总能量不变，但是随着声波体积的增大而导致密度的减少，类似水涟漪。并且即便是在理想条件下，扩散衰减也同样存在。

本模型是建立通过麦克风探知声音能量的变化来达成对未知大厅大小的估算，因此若是假设没有其他衰减，只考虑扩散衰减。在一个大厅内建立一个点声源，声源的扩散波形为球面波，并且当声源确定时，其声压为定值，若已知距声源 d_1 处的声压级为 L_1 ，声源 d_2 处的声压级为 L_2 ，结合上式 (5.3.4) 则不难推导出声压级 L_1 和 L_2 存在以下关系：

$$L_2 = L_1 + 20 \log \frac{P_2}{P_1} \quad (5.3.5)$$

结合 (5.3.4)，(5.3.5) 式及波的球面扩散特性可知声压与距离 d 存在反比关系，同时也可以得出声音的能量 I 与距离的关系式：

$$\frac{P_1}{P_2} = \frac{d_2}{d_1}, \frac{I_1}{I_2} = \frac{d_2^2}{d_1^2} \quad (5.3.6)$$

则根据上式，我们假设距声源无穷小位置的能量为 S ，则可以得到声音的能量随着距离衰减的理想公式：

$$I = \frac{S}{d^2} \quad (5.3.7)$$

吸收衰减是指在声音的传播过程中，由于介质具有粘滞性，导致一部分能量会被介质吸收，转化为热能，另外介质的热传导和介质微观过程中的弛豫效应也会导致声波能量损耗。我们通常认为吸收衰减和声音频率的平方成正比关系，频率越高，声波就会越容易被吸收，并且随着传播距离的增加，声波也会吸收的越多。

散射衰减是指声音在传播过程中，由于遇到了另外的一种不同介质而导致部分声波的传播方向发生偏离，从而导致声波的衰减。散射衰减不仅仅与传播介质有关，还与障碍物的形状，分布等有关。在升学研究中，声波的散射衰减遵循指数衰减。

我们可以知道现实生活理想的介质是不存在，吸收衰减与散射衰减总是不同程度的存在着，由此可以明显得知在实际应用中我们无法忽略吸收衰减和散射衰减，因此实际中声音的能量并不是和距离的平方成反比。但是我们历遍现有文献，我们都无法找到具体的吸收衰减和散射衰减理论上的公式，因此本文通过实际测量来确定衰减系数 α ，所以我们可以将式 (5.3.7) 改进为：

$$I = \frac{S}{d^\alpha} \quad (5.3.8)$$

5.3.3 声音衰减系数 α 确定实验

在理想环境下声音的衰减系数为 2，但是这是仅仅考虑扩散衰减的情况，在实际应用中，吸收衰减和散射衰减也会的衰减系数产生一定影响。因此，为了减少衰减系数不

确定对实验造成的误差，我们设计一个实验去测定在室内环境下衰减系数的值。具体实验如下：

我们在离声源 15m 的距离内，每隔 1m 进行信号采集，将采集的数据通过电脑程序换算成能量值，再将处理后的数据通过 matlab 工具箱进行 power 幂函数方程拟合，拟合结果如下图 11：

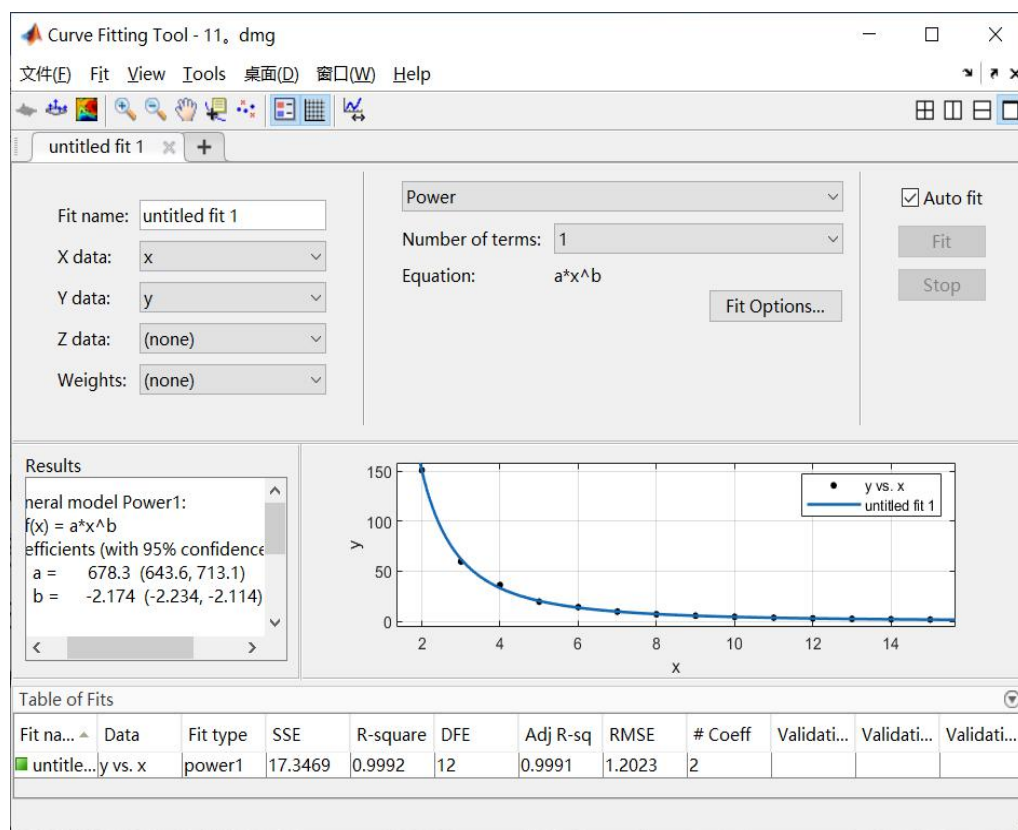


图 11 matlab 拟合结果

从图中我们可以从左侧的 Results 窗口得知幂函数的系数值 $a=678.3$, $b=-2.174$, 以及拟合曲线图，从而得出幂函数方程：

$$y = \frac{678.3}{x^{-2.174}} \quad (5.3.9)$$

其中 a 即为声源的能量强度， b 即为衰减系数。我们通过观察曲线可以明显的发现点与线之间的拟合程度很高，由此我们可以将室内的衰减系数确定为 $\alpha = 2.174$ 。由此得到声音在室内的衰减模型为：

$$I = \frac{S}{d^{2.174}} \quad (5.3.10)$$

5.3.4 确定大厅大小模型的建立

由于大厅大小的未确定可能会对定位模型造成一定的影响，因此我们需要测定大厅的大小，可是由于利用声达时间差的确定大厅下的定位模型虽然可以较为准确的测量出脚步的声源位置，但是无法精确得出大厅的大小，所以我们现在引用声音强度(能量法)来定位。

由上可知，我们已经得出声强和传播距离的关系。也就是说我们可以通过麦克风来确定声源（脚步声）与接收器（麦克风）之间的距离，因此设计出如下图 12 的麦克风安置方法：

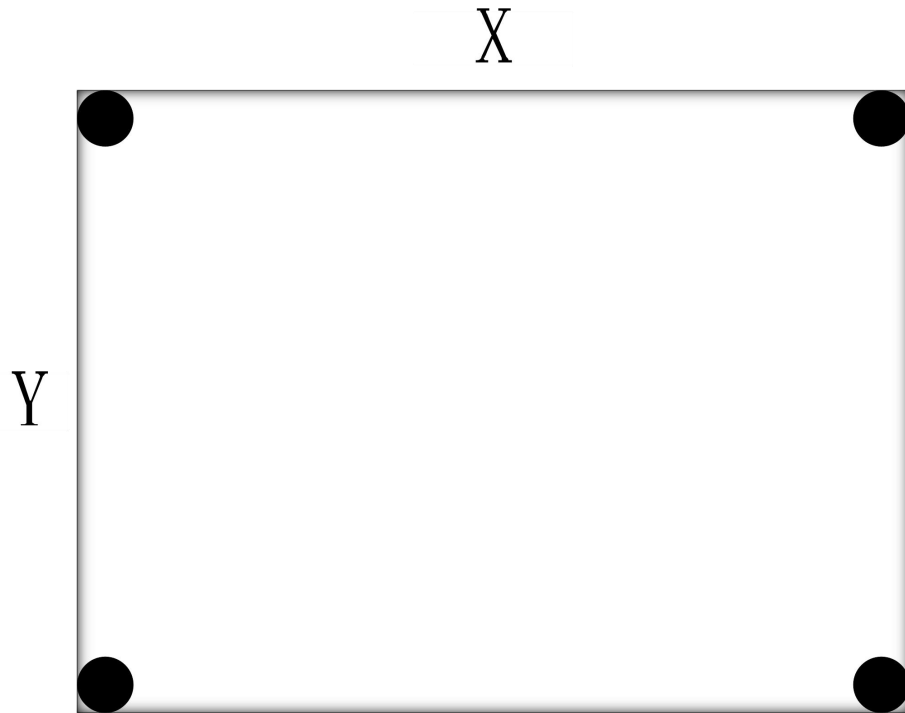


图 12 麦克风安置方法

如图 12 为该未知轮廓的大厅俯视图，我们分别将麦克风放置在大厅的四个角落，并设大厅的长为 X ，宽为 Y 。当声源（脚步声）响起时，麦克风接收到声音信息。但我们知道麦克风接收信息为声强级，因此我们需要将接收到的声强级信息转换为声强信息，由（5.3.4）式变换可知：

$$I = I_0 10^{\frac{1}{10} L_I} \quad (5.3.11)$$

同时考虑到麦克风接收信息的误差会导致不同麦克风对声源位置的估计产生偏差，因此这里使用蒙特卡罗算法进行近似处理，从而得到可以计算的合理声源位置。具体程序由附录给出。

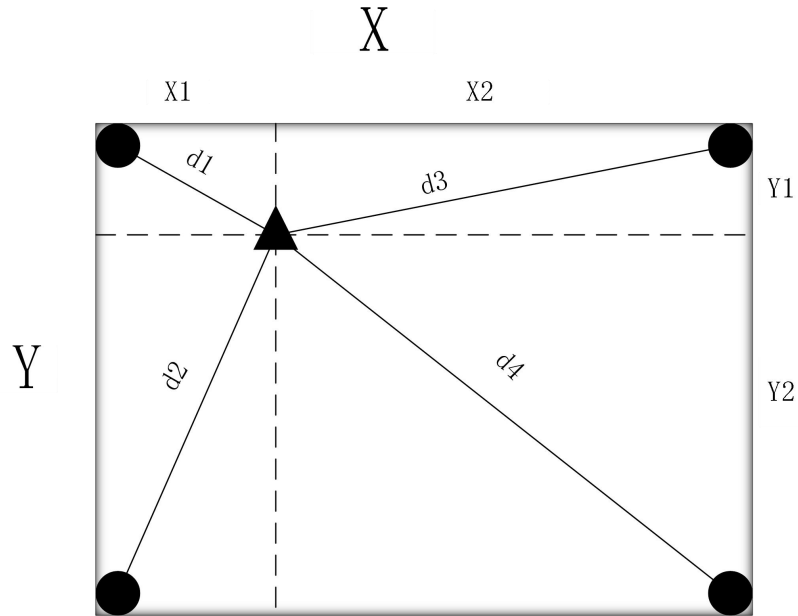


图 13 声源响起大厅的俯视分析图

如图 13 为声源（脚步声）响起后大厅的俯视分析图，图中三角形代表声源（脚步声）圆形代表检测装置（麦克风），我们将收集到的麦克风信息通过式（5.3.11）进行处理得到我们需要的声强信息。再通过式（5.3.10）可以得出声源到麦克风的距离 d_1, d_2, d_3, d_4 ，最后通过解方程组得出结果，方程组如下：

$$\left\{ \begin{array}{l} \sqrt{x_1^2 + y_1^2} = d_1 \\ \sqrt{x_1^2 + y_2^2} = d_2 \\ \sqrt{x_2^2 + y_1^2} = d_3 \\ \sqrt{x_2^2 + y_2^2} = d_4 \\ x_1 + x_2 = x \\ y_1 + y_2 = y \end{array} \right.$$

(5.3.12)

由式（5.3.10），（5.3.11），（5.3.12）我们可以通过 matlab 程序建立出一套关于测定未知大厅轮廓的模型，具体程序在附录中给出。

六、结果分析

6.1 十字形麦克风阵列检验

6.1.1 数值检验

已有实验数据表明，在大小为 $8*4*3.5$ 的室内实验室环境下，采用四元十字形麦克风阵列录制四路的声源信号，麦克风阵列的采样率是 10.2kHz ，噪声包括日光灯、风扇和窗户的噪声。麦克风各个阵元的摆放高度是 1.2m 。各个麦克风阵元之间的距离设为 0.6m 。麦克风的坐标分别为： $m_1(-0.3\text{m}, 0, 0)$ ， $m_2(0.3\text{m}, 0, 0)$ ， $m_3(0, 0.3\text{m}, 0)$ ， $m_4(0, -0.3\text{m}, 0)$ 。从测得的数据中选择 6 组实时声源定位的数据，如下表 5 所示：

序号	实际的声源坐标	测量的声源坐标
1	(1m, 25° , 30°)	(1.12m, 23.7° , 32.1°)
2	(1m, 40° , 35°)	(0.87m, 42.6° , 37.1°)
3	(1m, 60° , 50°)	(1.23m, 61.2° , 51.4°)
4	(1m, 25° , 35°)	(1.17m, 24.6° , 36.1°)
5	(1m, 60° , 60°)	(1.26m, 58.8° , 62.2°)
6	(1m, 45° , 60°)	(1.07m, 43.6° , 58.6°)

表 5 声源定位数据 (1)

由上表可知，距离定位的误差范围约为 $-0.13\text{m} \sim 0.17\text{m}$ ，角度定位误差范围约为 $\pm 2^\circ$ 左右，因此由该实验数据结果表明，距离定位和角度定位的误差不可忽略，这是因为实验环境中噪声等其它不可控因素的影响，会很大程度上影响我们十字形麦克风矩阵的测量各个麦克风阵元的时延，又因为时延的准确度会影响声源方位计算的精度，所以该实验的方案对声源位置的估计有着很大的误差。

6.1.2 改进

在原有的四元十字形麦克风阵列的基础上，四个麦克风阵元的中间位置再加一个麦克风阵元，其它实验环境、麦克风间距和高度跟上面一致，从测得的数据中选择 6 组实时声源定位的数据，如下表 6 所示：

序号	实际的声源坐标	测量的声源坐标
1	(1m, 25° , 30°)	(1.01m, 24.6° , 30.1°)
2	(1m, 40° , 35°)	(0.97m, 40.4° , 30.4°)
3	(1m, 60° , 50°)	(1.03m, 60.2° , 50.3°)
4	(1m, 25° , 35°)	(1.04m, 24.6° , 30.2°)
5	(1m, 60° , 60°)	(1.06m, 59.8° , 60.2°)
6	(1m, 45° , 60°)	(1.03m, 43.6° , 58.6°)

表 6 声源定位数据 (2)

由上表 6 可知，距离定位的误差范围约为 $-0.03\text{m} \sim 0.04\text{m}$ ，角度定位误差范围约为 $\pm 0.5^\circ$ 左右，因此由该实验数据结果表明，距离定位和角度定位的误差较小，可忽略不计。由于实验环境噪音和其它不可控因素的影响，该误差只能减小而不能消除。

由上面数据表明，十字麦克风阵列模型能够精确的定位声源目标的位置。

七、模型评价与推广

7.1 模型的评价

7.1.1 优点

1. 对于麦克风阵列排布方式,将已有的麦克风阵列的排布方式线性阵列与平面形阵列比较得到阵形的最优化阵形-十字麦克风阵形。对于麦克风的间距问题,在美国物理学家奈奎斯特理论基础上,得到声源信号的波长应该大于或等于两倍的相邻阵元间的最小距离,以此来得到麦克风的间距可信度强。

2. 对于建立麦克风阵列模型来精确人的位置信息,由于无法确定声源到各个麦克风的距离,因此将脚步声的传播模型分成两种,近场和远场模型分别讨论,近场模型中,使用球面波前模型;远场模型中,阵列波前应该使用平面波模型。在问题的求解中,我们选用十字麦克风阵列模型,采用角度距离定位方法,通过求各个麦克风阵元的延时确定声源的位置。

3. 对于确定大厅的轮廓问题,选用的是声音能量衰减模型,在现实生活中理想的介质是不存在的,因此声音在传播的过程中总会越来越小,该模型建立通过麦克风探知声音能量的变化来达成对未知大厅大小的估算,并且所用的麦克风数量也较少。

7.1.2 缺点

1. 麦克风阵列的排列方式有许多种,可以依据阵列的维度将其分为线性阵列、平面型阵列、立体型阵列。在我们的模型中我们只对比讨论线性阵列与平面型阵列中的正方形与等边三角形阵列及圆形阵列,并没有考虑到立体型的麦克风阵形。

2. 在精确人的位置信息的问题中,本模型没有考虑忽略远场和近场模型对阵形的影响,近场、远场都选用十字形阵列,只考虑阵列波前平面波和球面波的不同。显然这样的处理方法是否精确还有待商榷。

3. 在确定大厅的轮廓问题中,测量范围有所限制,当测量范围过大时,测量精度几乎成倍增长。麦克风的品质(即对信息的收集能量)不随实验时间的进行而下降,显然这只是在理想状态下考究的。最后声音能量衰减模型的计算过程复杂,需要大量的计算机资源。

7.2 模型的推广

7.2.1 改进

1. 不仅仅只考虑线性和平面的麦克风阵形,还考虑一下立体型的麦克风阵形,这样的改进方案麦克风阵列阵元会比普通的线性和平面的麦克风阵列覆盖范围更广,对声源信息的测量精度更高。

2. 将测方位与测大厅大小分开,采用十字麦克风阵列模型来进行大厅大小的测算,并确定声源的位置信息。

3. 我们都知道麦克风矩阵的选择会对模型的精度产生较大影响。声音能量衰减模型相当于将单个麦克风作为一个麦克风信息收集矩阵,那么我们可以通过增加四个角的麦克风矩阵数量或是改变麦克风矩阵排版来达到提高精度和范围的需要。

7.2.2 推广

由上可知我们的模型适合于测量一定范围内的位置并且拥有较高的精度,但是面对超出模型范围的测量则显得有些无能为力,因此我们提出了上述三种改进方法,都可以使得模型的精度和测量范围进一步扩大,但是相应的会增加经济负担。如何选择就需要通过实际的财力和测量要求来敲定。另外在现实生活中,基于麦克风阵列的声源定位技术早已在军事侦察,雷达和机器人听觉定位等方面得到了广泛的应用。随着人工智能水平的不断进步,听觉系统作为人类感官的重要组成部分,已经成为机器人研究领域的重要研究对象,麦克风阵列的声源定位的模型推广的价值很大。

八、参考文献

- 【1】 王月英;《基于声音位置指纹的室内声源定位研究方法》【D】,河北工业大学,2015。
- 【2】 姜锦云;《基于麦克风阵列的近场和远场混合声源定位》【D】,西南交通大学,2013。
- 【3】 潘菡;《随机声阵列地面预警定位系统研究》【D】,南京理工大学,2015。
- 【4】 卓金武;《MATLAB 在数学建模中的应用》【M】,北京航空航天大学出版社,2011。
- 【5】 朱哲民,龚秀芬,杜功焕;《声学基础》【M】,南京南京大学出版社,2001。
- 【6】 张海澜;《理论声学》【M】,高等教育出版社,2012。

九、附录

Matlab 程序代码 1:

/*本程序用于确定十字麦克风阵列模型*/

```
clear;clc;
v=340;z=11.33;
a=input(' please enter the time difference between unit 1 and unit
2:');
b=input(' please enter the time difference between unit 1 and unit
3:');
c=input(' please enter the time difference between unit 1 and unit
4:');
e=input(' please enter the hall width:');
p=sqrt(e*z/4);
r=abs((v/2)*((a^2+b^2+c^2)/(b-a-c)));
h=asin((v/(2*p))*(sqrt((a-c)^2+b^2)));
for i=0:pi/2:2*pi
k=atan((c-b)/a)+i;
if (k>0)&&(k<pi/2)
d=k;
end
end
disp(r);
disp(h);
disp(d);
```

Matlab 程序代码 2:

/*本程序用于确定大厅大小与定位*/

```
clear;clc;
d1=input(' please enter a value of d1:');
d2=input(' please enter a value of d2:');
d3=input(' please enter a value of d3:');
d4=input(' please enter a value of d4:');
a=1;b=1;c=1;d=1;e=0;f=0;g=0;h=0;k=0;j=0;
for x=1:0.5:20
for y=1:0.5:20
for i=1:10000
x1=round(x*rand);
```

```

        y1=round(y*rand);
        d5=sqrt(x1^2+y1^2);
        d6=sqrt(x1^2+(y-y1)^2);
        d7=sqrt((x-x1)^2+y1^2);
        d8=sqrt((x-x1)^2+(y-y1)^2);
        if
            (abs(d5/d1-1)<0.01)&&(abs(d5/d1-1)<a)&&(abs(d6/d2-1)<0.01)&&(abs(d6/d2-1)
            <b)&&(abs(d7/d3-1)<0.01)&&(abs(d7/d3-1)<c)&&(abs(d8/d4-1)<0.01)&&(abs(d8
            /d4-1)<d)
                a=abs(d5/d1-1);
                b=abs(d6/d2-1);
                c=abs(d7/d3-1);
                d=abs(d8/d4-1);
                e=x1;
                f=y1;
                g=d5;
                h=d6;
                k=d7;
                j=d8;
            end
        end
    end
end

if (e==0)&&(f==0)
    fprintf(' 所给数据误差过大无法分析\n');
else
    x1=e;
    y1=f;
    disp(x1);
    disp(y1);
    disp(x1+sqrt(k^2-y1^2));
    disp(y1+sqrt(h^2-x1^2));
end

Matlab 程序代码 3:
/*本程序用于通过声强级计算距离*/
clear;clc;
c=1*10^(-12);
a=input(' please enter an audio message:');
b=c*10^(a/10);
disp(b);
s=1.0000e-07;
d=s/b;
e=nthroot(d,2.174);

```

```
disp(e);
```