

基于麦克风阵列和神经网络的声源定位实验设计

汤子涵 罗云荣 胡锦涛 杨雨睿

(湖南师范大学物理与电子科学学院, 湖南长沙 410081)

摘要 本文基于神经网络和广义互相关时延方法, 结合麦克风阵列技术搭建了一套声源定位实验装置。该装置首先利用卡尔曼滤波对原始声音信号进行预处理, 然后通过广义互相关时延算法提取麦克风间的时间差, 最后采用单隐藏层神经网络构建时间差与二维坐标之间的映射关系来确定声源的位置。通过利用神经网络替代传统数学模型, 既有效减弱了环境噪声、声速变化及设备延迟等系统性误差, 又大幅降低了计算复杂度和内存需求。实验结果表明, 该装置具有测量精度高、成本低, 并能实时显示与记录测量数据等优点。

关键词 麦克风阵列; 声源定位; 神经网络; 广义互相关时延

EXPERIMENTAL DESIGN OF SOUND SOURCE LOCALIZATION BASED ON MICROPHONE ARRAYS AND NEURAL NETWORK

TANG Zihan LUO Yunrong HU Jincheng YANG Yurui

(School of Physics and Electronics, Hunan Normal University, Changsha, Hunan 410081)

Abstract This paper presents a sound source localization experimental setup that integrates microphone-array technology with neural networks and the generalized cross-correlation time-delay (GCC-TD) method. The system first preprocesses raw acoustic signals with a Kalman filter, then estimates the time-difference-of-arrival between microphones via GCC-TD, and finally employs a single-hidden-layer neural network to map these time differences onto two-dimensional coordinates, thereby pinpointing the sound source. Replacing conventional mathematical models with a neural network markedly mitigates systematic errors arising from ambient noise, sound-speed variations, and hardware latency, while simultaneously cutting computational complexity and memory requirements. Experimental results demonstrate that the apparatus offers high accuracy, low cost, and real-time display and logging of measurement data.

Key words microphone arrays; sound source localization; neural network; generalized cross-correlation time delay

当前麦克风阵列在许多实用技术领域都有着广泛的应用, 在自动驾驶领域, 麦克风阵列能使智

能汽车检测到环境中的喇叭声、警笛声等, 并对其做出正确的反应^[1]; 在智能城市和智能家居中, 麦

收稿日期: 2025-04-09

基金项目: 国家自然科学基金项目(11747034); 国家级大学生创新创业训练计划项目(S202410542012, 202510542020); 湖南师范大学教学改革研究项目(JG2025020)。

通信作者: 罗云荣, lyr_1982@hunnu.edu.cn。

引文格式: 汤子涵, 罗云荣, 胡锦涛, 等. 基于麦克风阵列和神经网络的声源定位实验设计[J]. 物理与工程, 2025, 35(6): 190-199.

Cite this article: TANG Z H, LUO Y R, HU J C, et al. Experimental design of sound source localization based on microphone arrays and neural network[J]. Physics and Engineering, 2025, 35(6): 190-199. (in Chinese)

克风阵列可用于机器监听和音频监控;在应用软件中,麦克风阵列可用于实现会议、通话等功能^[2]。传统的通过麦克风阵列进行声源定位的方法大致分为三类:基于到达时延估计的声源定位算法^[3]、基于最大输出功率的可控波束形成声源定位算法^[4]和基于高分辨率谱估计技术算法^[5]。

在基于到达时延估计的声源定位算法中,最常用的方法是固定麦克风坐标,构建麦克风接收时间差与声源坐标的函数。如果把测得的时间差当作自变量 t ,把坐标当作因变量 y ,这就是一个求解时间差与坐标之间转换函数的问题。对于处理这种函数问题,我们可以用神经网络训练出一个矩阵来代替这个函数,与一般的方法相比较,运用神经网络的方法是唯数据论的,这使得系统性误差如混响、声速、装置本身延迟等可以被减小。

在本文的实验设计中,我们先搭建一个神经网络,然后用大量已测得的数据作为训练集来训练神经网络,再把训练好的神经网络的矩阵提取出来放入单片机内^[6],便可以达到用时间差信息来得到声源坐标的目的。

1 声源定位实验原理

1.1 卡尔曼滤波

在声源定位系统的构建中,首要且关键的一步是捕获并预处理出精确无误的声音信号。鉴于环境噪声的普遍存在,对原始声音信号进行有效滤波以提取纯净信号成为必要步骤。本实验中,我们采用卡尔曼滤波算法对麦克风阵列采集到的声音信号进行预处理^[7]。卡尔曼滤波以其高效的递归滤波器设计,能够在动态环境中准确估计信号状态,从而显著减少噪声干扰,为后续的时延估计奠定坚实基础,其采用反馈控制的方法来估计过程状态:它首先估计某一时刻的过程状态,随后通过包含噪声的测量变量来获取反馈,并进行状态更新^[8]。因此卡尔曼滤波器可分为两个部分:时间更新方程和测量更新方程。时间更新方程负责及时向前推算当前状态变量和误差协方差的估计值,以便为下一个时间状态构造先验估计。测量更新方程负责反馈,也就是说,它将先验估计和新的测量变量结合以构造改进的后验估计^[9]。时间更新方程可视为预估方程,测量更新方程可视为校对方程,最后的估计算法为一种具有数值解

的预估-校正算法,如图 1 所示。

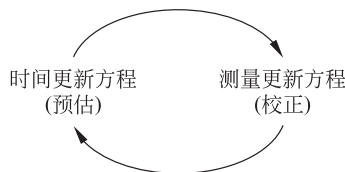


图 1 卡尔曼滤波循环示意图^[9]

此外,卡尔曼滤波器还具有如下优点:首先,该滤波器无需保留除前一个状态量之外的其他历史数据,这使得它具有内存占用小、推理速度快的显著特点。其次,卡尔曼滤波器巧妙地融合了观测数据与估计数据,并对误差进行闭环管理,这种机制使其能够将误差有效地控制在一定范围内,从而提高了定位的准确性和稳定性。这些优点对于我们的装置来说十分重要,速度快才能保证实时的声源定位,占用内存小才能够把足够大的空间留给矩阵运算,误差的限定则可以提升我们搭建神经网络的鲁棒性。

1.2 广义互相关时延

在获取了经过卡尔曼滤波优化的声音信号后,接下来的核心任务是将这些声音信号转换为可用于声源定位的时间差数据。为此,我们引入了广义互相关时延算法^[10]。该算法计算每两个信号间的互相关函数,通过寻找使互相关函数达到最大值的点,即可有效地确定这两个信号之间的时间差。离散信号的互相关函数^[11]

$$R(\tau) = E[x_1(x_1(m)x_2(m+\tau))] \quad (1)$$

求时间差就是找到互相关函数取最大值时的点

$$D = \operatorname{argmax} R(\tau) \quad (2)$$

理论上,采用上述方法即可有效获得两个信号之间的时间差,但是实际信号往往会受到噪声的干扰。低信噪比条件下,互相关函数的峰值可能变得不够明显,这在进行极值寻找时会引入误差,从而影响时间差的准确计算。为了得到具有更陡峭极值的互相关函数,我们选择使用一个加权函数来白化输入信号,这就是广义互相关时延的方法。信号的功率谱密度函数和互相关函数服从一对傅里叶变换关系

$$P(\omega) = \int_{-\infty}^{+\infty} R(\tau) e^{-j\omega\tau} d\tau \quad (3)$$

$$R(\tau) = \int_{-\infty}^{+\infty} P(\omega) e^{j\omega\tau} d\omega \quad (4)$$

由此可见,式(3)和式(4)把互相关函数变换到了频域,由于求时间差只需要相位信息,因此,我们可以利用相位变换法舍弃不相关的幅度信息以提高声音信号的健壮性^[11],即将式(4)变为

$$\tilde{R}(\tau) = \int_{-\infty}^{+\infty} A(\omega) P(\omega) e^{j\omega\tau} d\omega \quad (5)$$

其中

$$A(\omega) = \frac{1}{|P(\omega)|} \quad (6)$$

1.3 神经网络原理

基于前述步骤获得的时间差数据,我们进一步利用神经网络算法来建立时间差与声源二维坐标之间的映射关系。神经网络是一种模拟人脑的神经网络以期能够实现类人工智能的机器学习技术^[6],主要结构分为输入层、隐藏层和输出层,如图2所示。

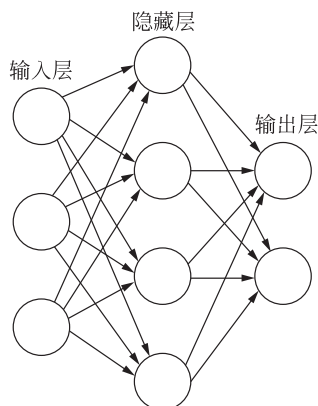


图2 神经网络结构图

在本次设计中,神经网络架构的输入层负责接收由三个麦克风测得的两两时间差数据,该层配置有三个神经元以适配这些输入特征。隐藏层作为连接输入层与输出层的关键部分,通过对输入数据进行非线性变换和特征提取,学习并揭示数据的内在规律和模式,其具体的结构设计细节将在3.3节详细讨论。输出层是神经网络的终端层,包含两个神经元,分别负责预测并输出目标物体的距离与角度信息^[12]。

在图2中,每一条线都代表一个权重,神经网络的训练算法,就是让权重的值调整到最佳,使整个神经网络的预测达到最好的效果。这一优化过程的基础在于神经元模型^[13],它是一个包含输入、输出与计算功能的模型。神经元的结构如图3所示,输入信号过程、输出信号过程和计算可以分

别类比为图3中神经元的树突、轴突和细胞核。

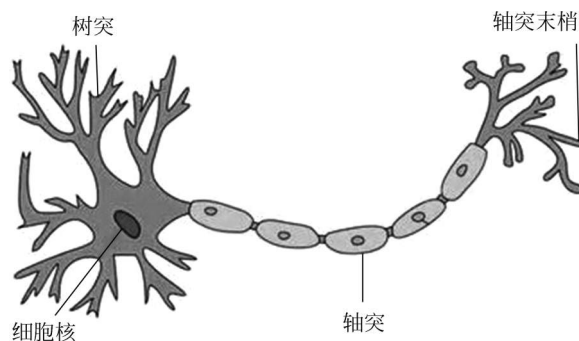


图3 神经元

基于这一神经元模型,我们所采用的前向神经网络模型为^[6]

$$f = a + \sum_{j=1}^h b_j \tanh(c_j + \sum_{i=1}^l d_{ij} t_i) \quad (7)$$

其中, h 为隐藏层神经元个数; l 为神经网络中输入参数个数; $\theta = \{a, b_j, c_j, d_{ij}\}$ 为模型参数集。根据数据集的规模来确定神经元的个数,进而决定神经网络的复杂性。

\tanh 是我们选用的神经网络的激活函数。几乎所有的神经网络都需要有激活函数以增加非线性因素,解决线性模型表达能力不足的缺陷。没有激活函数的神经网络实质上是一个线性回归模型,只能解决线性可分的问题。常见神经网络中的激活函数有sigmoid、ReLU、 \tanh 等^[6]。我们针对这些激活函数进行了对比实验,在相同训练集(1000组数据)和网络结构(单隐藏层4个神经元)条件下,分别记录了各激活函数在测试集(200组数据)中的表现于表1。实验结果显示:sigmoid函数因梯度消失问题导致平均距离误差达12cm;ReLU函数在负区间零激活特性导致约21%的样本无法收敛;而 \tanh 函数凭借其零中心对称特性^[14],在距离和方位角预测中均表现出更优性能。故本文选择 \tanh 作为激活函数。

表1 不同激活函数性能对比(测试集200组数据)

激活函数	平均距离 误差/cm	平均方位角 误差/(°)	收敛 失败率
sigmoid	12	15	0%
ReLU	9.8	11	21%
tanh	4.2	5.1	0%

本实验构建的神经网络模型旨在通过训练学

习,自动发现并提取时间差数据中隐含的声源位置信息,从而实现从时间差到声源坐标的高效、准确转换。这一技术的应用,不仅提升了声源定位的智能化水平,也为复杂环境下的声源定位提供了一种新的解决方案。

2 实验装置

2.1 实验装置总图

声源定位实验装置由蜂鸣器(点声源)、声音传感器、单片机、外接电脑等装置组成,分别对应发声装置、接收装置、计算装置和显示装置四个部分,其实际装置图片如图 4 所示。

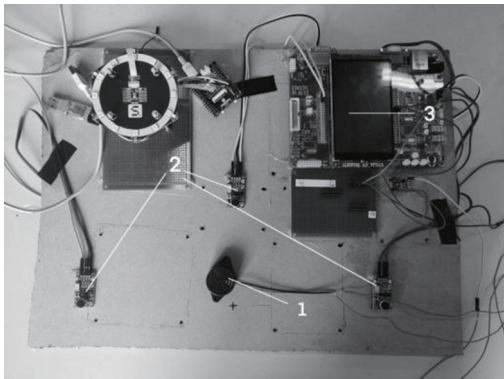


图 4 装置图片

1—蜂鸣器; 2—声音传感器; 3—STM32f407ZGT6 单片机

2.2 单片机

我们选用的是 STM32f407ZGT6 单片机,这是一款经典的单片机,功能集成度和性价比都比较高,如图 5 所示,它的最高工作频率为 168MHz,包含 17 个定时器^[15]。

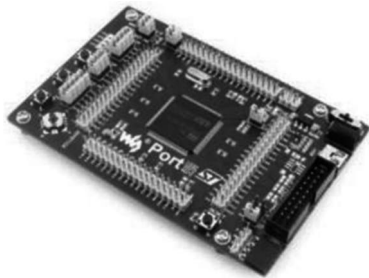


图 5 STM32f407ZGT6 单片机

2.3 声音传感器

图 6 是我们实验中所用的声音传感器,它的麦克风灵敏度为 52dB,工作频段为 50Hz~20kHz。

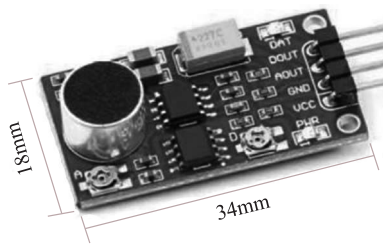


图 6 声音传感器

3 实验实施和测量分析

3.1 实验前准备

为了表示声源的空间坐标,我们在平面中选取 A、B、C 三点构成一个底角为 30° 的等腰三角形,腰长为 10.00cm,顶点为 A 点,底边中点为 O 点,点 A、B、C 位置实际为三个声音传感器的放置位置。以 O 点为极点、OC 为极轴正方向建立极坐标,如图 7 所示,通过测量声源到极点实际的距离和方位角来确定声源位置。

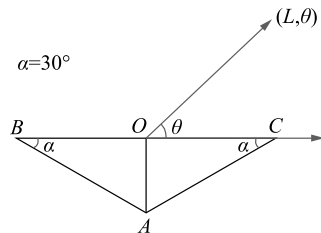


图 7 极坐标示意图

3.2 声源定位实验内容

(1) 在实验平台上,根据预设方案连接所有实验装置,并将各传感器(声音接收设备)安装至指定的 A、B、C 三点位置。确保所有设备电源连接无误,启动实验系统。

(2) 通过按下启动按钮触发蜂鸣器发声,利用麦克风阵列捕捉声音信号并计算各麦克风接收到声音的时间差,如图 8 所示。同时,采用卷尺测量声源至极点 O 的实际距离,并利用细棉线与量角器组合测量方位角,为方便起见,这里方位角我们用度($^\circ$)表示。本实验中,卷尺的最小分度值为 0.1cm,量角器的最小分度值为 1° 。将这一系列测量的距离、方位角以及时间差数据整理成数据集,作为训练集输入至计算机中的神经网络模型,通过训练过程生成一个用于声源定位的矩阵模型。

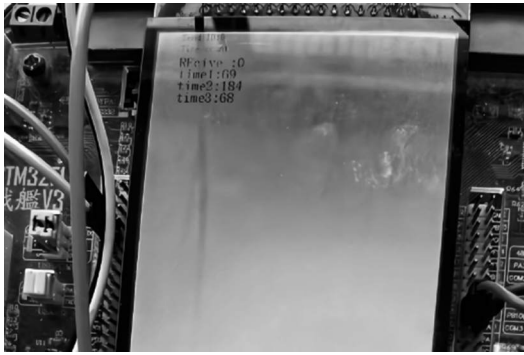


图8 单片机测得的声源到三个点的时间差

(3) 将训练好的矩阵模型输入至单片机中。随后,再次激活蜂鸣器,单片机即时收集麦克风阵列的信号数据,并依据内置矩阵和时间差数据进行快速运算。运算结果通过内置算法处理,直接转换为声源相对于极点 O 的距离和方位角信息。这些信息随后被传输至计算机显示屏上,以直观展示,如图 9 所示,预测结果中首个数据代表距离,紧随其后的数据代表方位角。

```

请输入一组数据 (以逗号分隔): 161.9,152.1,155.7
1/1 [=====] - 0s 87ms/step
预测结果: ['154.22 cm', '24.0 °']
是否继续预测? (Y/N):
请输入一组数据 (以逗号分隔): 232.3,224.9,221.9
1/1 [=====] - 0s 21ms/step
预测结果: ['166.93 cm', '157.3 °']
是否继续预测? (Y/N):
请输入一组数据 (以逗号分隔): 74.1,64.9,66.2
1/1 [=====] - 0s 20ms/step
预测结果: ['63.75 cm', '29.6 °']
是否继续预测? (Y/N):
请输入一组数据 (以逗号分隔): 113.0,103.1,108.3
1/1 [=====] - 0s 21ms/step
预测结果: ['108.59 cm', '3.8 °']
是否继续预测? (Y/N):
请输入一组数据 (以逗号分隔): 51.9,46.1,41.0
1/1 [=====] - 0s 21ms/step
预测结果: ['27.82 cm', '39.6 °']
是否继续预测? (Y/N):

```

图9 输入声音信号后程序预测图

3.3 声源定位实验模型构建

3.3.1 训练集的准备

在设定的极坐标系中,我们随机均匀地在以极点为圆心,半径为 300.00cm 的圆内测量声源到三个点 A 、 B 、 C 的两两时间差,总共测试了大约 1000 组数据当作训练集,在实际数据的测量中每个点位测量 3 次,记录 3 组数据计算平均值作为

该次测量的最终结果。

3.3.2 隐藏层结构的搭建

在神经网络搭建的过程中,所有需要人为来调整的参数统称为超参数,其中很重要的两个超参数就是隐藏层的层数和隐藏层的神经元个数。

实验设计中我们选择一层隐藏层,因为每多加一层隐藏层所需要的内存将大大增加。而隐藏层中的神经元数量也是装置的重要参数,我们选取 2 到 6 个神经元对声源位置进行了测试,其测试结果如表 2 所示。

表 2 不同神经元数目声源距离与方位角的预测值与实际值
距离单位:厘米/cm,方位角单位:度/(°)

神经元数目			
2		3	
预测值 距离, 方位角	实际值 距离, 方位角	预测值 距离, 方位角	实际值 距离, 方位角
18.12,4.1	10.00,0.0	6.74,0.3	10.00,0.0
42.75,11.8	30.00,15.0	19.45,16.7	30.00,15.0
62.06,43.3	50.00,30.0	41.83,20.8	50.00,30.0
58.39,49.4	70.00,45.0	64.02,47.4	70.00,45.0
73.40,27.0	90.00,60.0	85.75,63.2	90.00,60.0
88.11,94.3	110.00,70.0	108.11,66.1	110.00,70.0
157.40,99.0	130.00,80.0	132.58,89.4	130.00,80.0
182.21,145.1	150.00,110.0	139.41,102.1	150.00,110.0
145.85,98.2	170.00,135.0	158.20,126.5	170.00,135.0
163.37,182.1	190.00,170.0	164.15,158.1	190.00,170.0
4		5	
预测值 距离, 方位角	实际值 距离, 方位角	预测值 距离, 方位角	实际值 距离, 方位角
13.38,4.4	10.00,0.0	9.54,3.2	10.00,0.0
23.62,11.7	30.00,15.0	25.77,12.6	30.00,15.0
45.99,23.2	50.00,30.0	43.92,21.9	50.00,30.0
73.25,41.2	70.00,45.0	68.16,40.0	70.00,45.0
87.95,57.9	90.00,60.0	85.94,56.3	90.00,60.0
110.17,71.7	110.00,70.0	108.34,65.8	110.00,70.0
125.07,75.1	130.00,80.0	125.07,75.2	130.00,80.0
143.20,116.2	150.00,110.0	156.26,121.1	150.00,110.0
161.46,124.4	170.00,135.0	156.73,117.7	170.00,135.0
168.84,155.0	190.00,170.0	165.30,144.4	190.00,170.0

续表

神经元数目	
6	
预测值	实际值
距离, 方位角	距离, 方位角
9.32,2.7	10.00,0.0
26.65,15.3	30.00,15.0
46.39,27.0	50.00,30.0
65.27,39.1	70.00,45.0
83.94,62.5	90.00,60.0
106.04,67.3	110.00,70.0
123.37,75.8	130.00,80.0
136.63,119.2	150.00,110.0
179.01,152.0	170.00,135.0
171.89,151.4	190.00,170.0

根据表 2 中的测试数据,我们作了不同神经元数目下测试声源的距离误差图和方位角误差图,分别如图 10 和图 11 所示。从图 10 和图 11 可以看出,2个神经元时,神经网络的偏差极大,

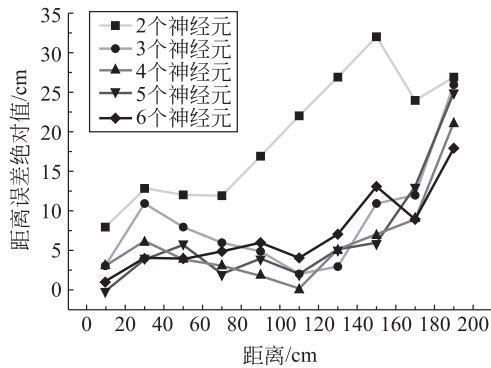


图 10 不同神经元数目下测试声源距离误差

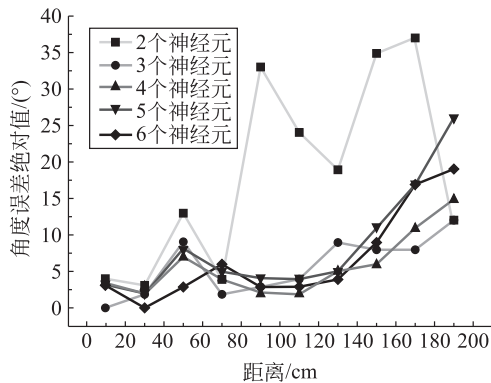


图 11 不同神经元数目下测试声源方位角误差

是因为神经元太少无法学习到规律。而神经元数目过多,误差并没有显著减小,是因为过多的神经元导致了过拟合。因此,我们最终选择使用 4 个神经元内置于单层隐藏层中。

3.4 实验数据测量和处理

为了测试装置对声源定位的准确性,我们分别测量了声源到极坐标 O 点的距离在 30.00~100.00cm,100.00~200.00cm,200.00~300.00cm 范围内,并在每个测量范围内随机选择 10 个位置,记录真实的位置信息与测量数据于表 3、表 4 和表 5,之后通过 Origin 绘制图像进行对比分析^[15],直观地反映测量数据与真实情况的吻合度,如图 12~图 14 所示。

表 3 30.00~100.00cm 范围声源距离与方位角误差

测量 距离/ cm	实际 距离/ cm	距离 误差/ cm	测量 方位角/ (°)	实际 方位角/ (°)	方位角 误差/ (°)
29.52	29.40	0.1	39.6	39.2	0.4
40.32	40.30	0.0	120.1	118.8	1.3
42.68	44.70	-2.0	107.1	106.4	0.7
58.56	56.60	2.0	277.8	277.2	0.6
62.52	62.60	-0.1	202.6	206.0	-3.4
73.83	70.80	3.0	282.8	284.4	-1.6
80.30	80.40	-0.1	20.9	22.8	-1.9
86.85	86.90	-0.1	313.2	312.0	1.2
93.48	92.50	1.0	271.7	273.6	-1.9
98.42	98.50	-0.1	60.9	61.2	-0.3

表 4 100.00~200.00cm 范围声源距离与方位角误差

测量 距离/ cm	实际 距离/ cm	距离 误差/ cm	测量 方位角/ (°)	实际 方位角/ (°)	方位角 误差/ (°)
103.03	103.10	-0.1	265.8	268.8	-3.0
119.27	119.30	0.0	53.1	54.4	-1.3
123.72	127.80	-4.1	66.9	69.2	-2.3
135.28	136.20	-1.1	239.0	240.0	-1.0
146.84	144.80	2.0	286.5	289.2	-2.7
152.04	153.00	-1.0	255.0	257.6	-2.6
164.28	168.20	-4.0	242.7	246.4	-3.7
174.76	174.90	-0.1	70.6	73.2	-2.6
185.47	185.60	-0.1	104.3	106.4	-2.1
198.39	195.50	2.9	77.9	78.8	-0.9

表5 200.00~300.00cm 范围声源距离与方位角误差

测量 距离/ cm	实际 距离/ cm	距离 误差/ cm	测量 方位角/ (°)	实际 方位角/ (°)	方位角 误差/ (°)
204.85	204.80	0.1	267.5	269.2	-1.7
210.43	213.60	-3.2	85.3	88.0	-2.7
227.17	222.30	4.9	109.4	110.4	-1.0
238.32	235.50	2.8	176.8	179.2	-2.4
248.15	248.20	-0.1	205.6	210.0	-4.4
258.29	258.40	-0.1	188.8	192.0	-3.2
255.81	260.90	-5.1	174.2	176.8	-2.6
269.19	265.40	3.8	168.0	170.8	-2.8
280.91	276.80	4.1	68.2	62.4	5.8
298.36	289.60	8.8	174.9	168.2	6.7

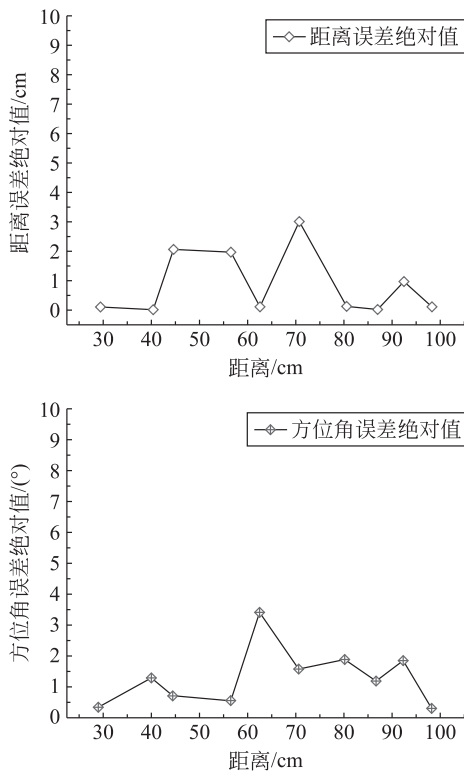


图12 30.00~100.00cm 声源距离与方位角误差分析图

由图12我们发现在30.00~100.00cm范围内,声源定位装置距离误差绝对值最大为3.0cm,方位角误差绝对值最大为3.4°,总体来说误差较小。

从图13可以看出,尽管距离和方位角误差相比图12均有稍稍增大的趋势,但总体而言误差依然较小,其距离误差绝对值最大为4.1cm,方位角

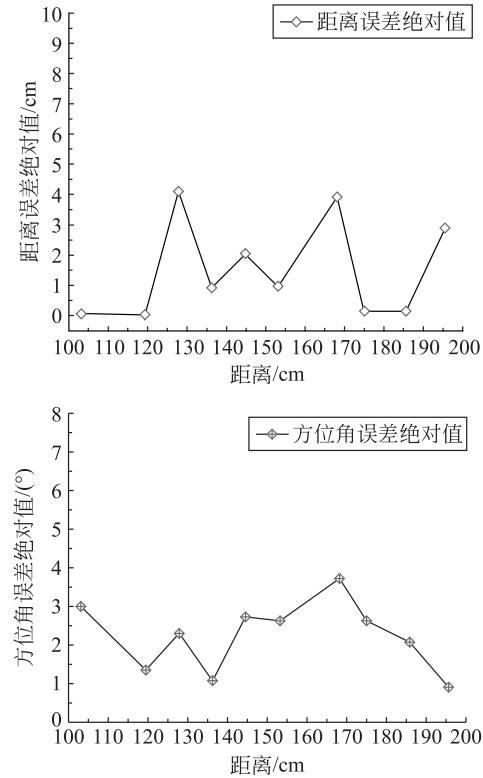


图13 100.00~200.00cm 声源距离与方位角误差分析图

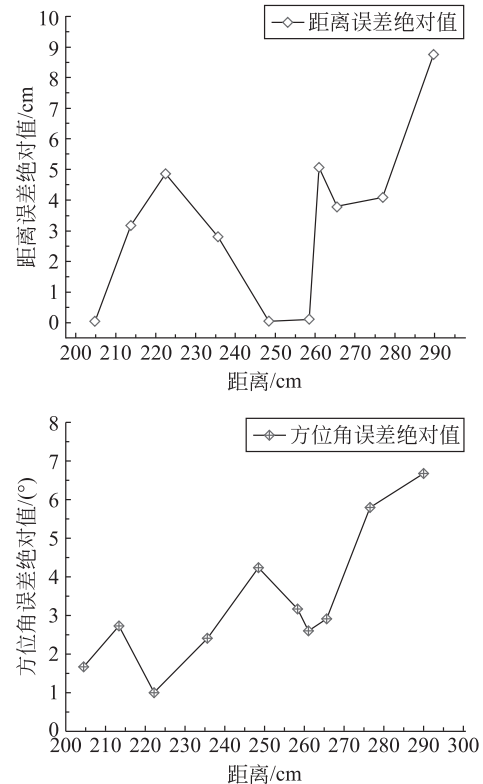


图14 200.00~300.00cm 声源距离与方位角误差分析图

误差绝对值最大为3.7°。

由图14我们可以看到,当声源距离大于270.00cm

时,麦克风阵列已逐渐不能稳定的接收声音信号,根据声波传播理论,声压级随传播距离呈对数衰减

(衰减公式为 $L_p(r) = L_{p0} - 20 \log_{10} \left(\frac{r}{r_0} \right)$ [16])。

当声源距离达到 270.00cm 时,原始信号幅度相比 30.00cm 处衰减约 19.1dB,导致麦克风接收信号的信噪比低于设备灵敏度阈值,直接影响时延估计精度。由此我们可以得到结论:装置的最佳测量范围是 30.00~270.00cm,且距离误差和方位角误差都有随距离增大而增大的趋势。而在 200.00~270.00cm 的范围内,距离误差绝对值最大为 5.1cm,方位角误差绝对值最大为 4.4°。

3.5 装置可行性分析

为了进一步验证装置的可行性,我们扩大了采集点的数量,一共在极坐标系中随机选取了 200 个位置作为测试集。为了减少测量误差,我们借助固定架等设备进行多次测量,并记录真实的位置信息与测量数据,之后通过 Origin 绘制距离和方位角实际值与预测值的拟合图像,如图 15 和图 16 所示。

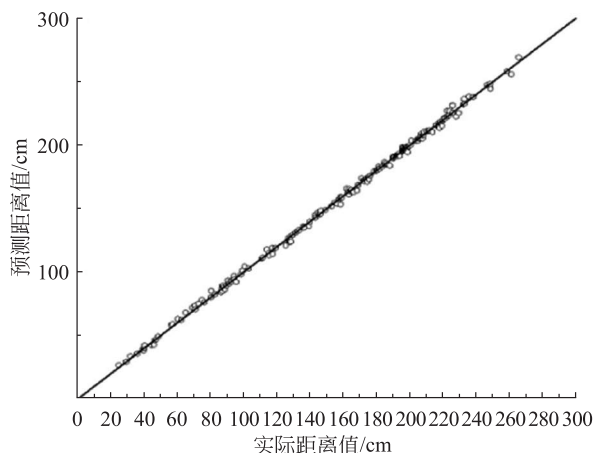


图 15 预测距离与实际距离拟合图

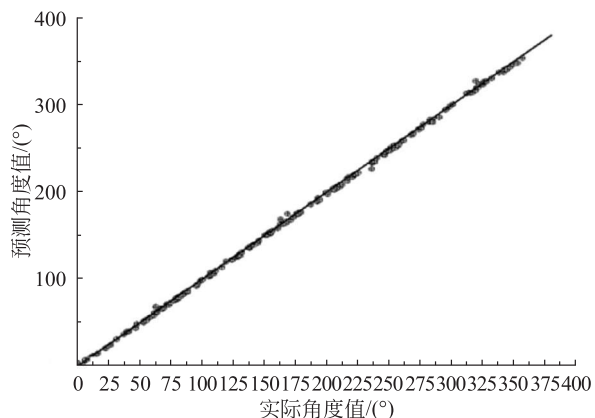


图 16 预测方位角与实际方位角拟合图

从图 15 和图 16 我们可以得到:距离和方位角的实际值与预测值的拟合曲线与直线 $y=x$ 高度吻合,数据密集且均匀的分布在 $y=x$ 直线附近,测试集的大量数据证明了该套装置用于声源定位是可行的,展现了神经网络唯数据论的优越处理性能。

3.6 装置测量范围

该套装置在二维空间内方位角 $0.0^\circ \sim 360.0^\circ$ 的范围都可以进行测量,在实际测量中,通过采集 200 个测试集数据,我们发现,声源距离装置在 30.00~270.00cm 的范围内测量误差较小且稳定,因此,这一距离范围为该装置的最佳测量距离范围。

3.7 装置精确度分析

我们将 200 个测试集数据的距离误差和方位角误差分别绘制成散点图,如图 17 所示。从图 17 我们不难看出,随着声源距离的增大,其测量所得的距离误差和方位角误差都呈增大的趋势。在装置的最佳测量范围内,我们统计了 200 个测试集数据,在声源距离测试装置的不同距离范围内,其测量得到的最大距离误差与方位角误差,整理于

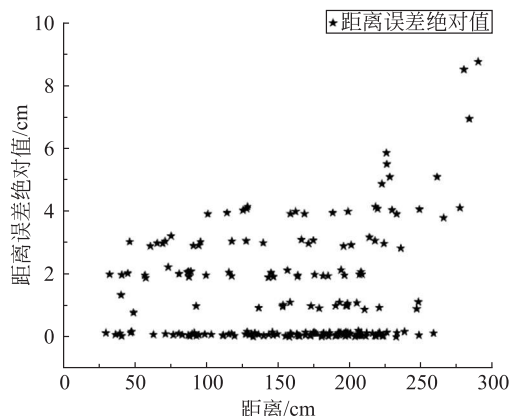
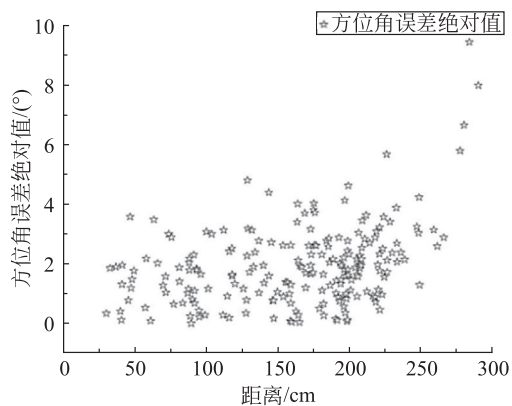


图 17 测试集误差散点图

表6之中。从图表中我们可以看到,该装置对声源的定位误差小,精度高。

表6 200个测试数据集误差统计表

距离范围/cm	30.00~	100.00~	200.00~
	100.00	200.00	270.00
距离误差/cm	±3.0	±4.2	±5.9
方位角误差/(°)	±3.6	±4.8	±5.7

4 装置创新点

(1) 神经网络模型相较于传统方法,在准确性、高效性及实时性方面展现出了显著优势,具体体现在:

① 准确性:神经网络模型摒弃了复杂的人为数学模型设定,凭借大量数据的训练,自动学习并精准捕捉声音信息与声源位置之间的复杂映射关系,这种数据驱动的模式不仅优化了算法流程,还有效削弱了声速变化、装置延时等系统性误差的干扰,从而显著提升了声源定位的准确性;

② 高效性:神经网络模型在完成训练后,展现出了极高的预测速度,能够迅速根据输入的时间差信息等数据进行声源位置的高效预测;

③ 实时性:神经网络模型结合上位机的实时显示和记录功能,使得实验装置成功实现了实时的声源定位,确保装置在复杂环境下依然能够展现出卓越的声源定位性能。

(2) 相较于普遍采用树莓派等设计的声源定位装置,本装置在计算量与内存需求上实现了显著优化。传统设计常需存储大量中间数据及结果,且易受环境噪声干扰影响声源定位精度。而本装置引入卡尔曼滤波算法,通过高效反馈控制机制估计信号状态,即首先估计某一时刻的过程状态,随后通过包含噪声的测量变量来获取反馈,并进行状态更新,这不仅能够使装置内存需求降低,计算量大大减少,还能够在动态环境中准确追踪信号变化,显著减少了噪声的干扰,有效提高了声源定位的鲁棒性。

5 讨论

本文基于神经网络和广义互相关时延估计方

法,将相关算法内置于单片机中,结合麦克风阵列技术搭建了一套声源定位实验装置。该装置实现了30.00~270.00cm范围内的声源精准定位,具有测量误差小、装置成本低、能实时显示与记录测量数据等优点。

在表7和表8中我们分别列举了利用传统方法^[17-22]与本实验方法进行声源定位在测量误差和实时性方面的测量数据。通过数据对比,容易发现本文基于神经网络进行声源定位在误差抑制和计算效率上较传统方法均具有明显优势。

表7 本实验方法与传统方法声源定位测量误差数据对比

测量距离/cm	测量方法	测量距离误差/cm	测量角度误差/(°)
30.00~	传统方法[17]	\	3.0
40.00	本实验方法	\	2.4
80.00~	传统方法[18]	5.0	\
120.00	本实验方法	3.2	\
150.00	传统方法[19]	\	7.0
	本实验方法	\	2.7
150.00~	传统方法[20]	10	\
250.00	本实验方法	5.9	\

表8 本实验方法与传统方法声源定位实时性对比

文献	开发平台	定位方法	单次定位平均耗时/ms
文献[21]	PC机	时延估计算法	502
文献[22]	xCORE-200+ PC机	可控响应功率-相位变换 算法	525.2
本文	STM32	时延估计+ 神经网络 算法	34.1

6 结语

综上所述,本套装置在声源定位方面展现出了优异的性能,唯数据论的方法有效减少了由装置和环境本身导致的误差。然而,随着距离的增加,误差逐渐增大,这提示我们在未来的研究中需要进一步优化训练集数据的采集和处理方法,以

扩大装置的有效测量范围并提高远距离声源定位的测量精度。

参 考 文 献

- [1] NANDWANA M K, HASAN T. Towards smart-cars that can listen: Abnormal acoustic event detection on the road [C]. *Interspeech*, 2016: 2968-2971.
- [2] CROCCO M, CRISTANI M, TRUCCO A, et al. Audio surveillance: A systematic review [J]. *ACM Computing Surveys*, 2016, 48(4): 52.
- [3] 刘涛. 基于多径信号时延估计算法的 TDOA 室内定位系统 [D]. 武汉: 华中科技大学, 2016.
LIU T. TDOA indoor localization system based on multipath time delay estimation [D]. Wuhan: Huazhong University of Science and Technology, 2016. (in Chinese)
- [4] SHI W L, LI Y S, ZHAO L Y, et al. Controllable sparse antenna array for adaptive beamforming [J]. *IEEE Access*, 2019, 7: 6412-6423.
- [5] ZHA X, SHAO Z, DAI Z, et al. Accurate frequency estimation for removal of orbital fringes in SAR interferograms [J]. *International Journal of Remote Sensing*, 2020, 41(14): 5305-5320.
- [6] ANDERSON J A. An introduction to neural networks [M]. MIT press, 1995.
- [7] 薛为, 栾小丽, 赵顺毅, 等. 基于影响力函数的自适应鲁棒卡尔曼滤波 [J]. *中国科学: 信息科学*, 2025, 55(3): 601-618.
XUE W, LUAN X L, ZHAO S Y, et al. An adaptive robust Kalman filter based on influence function [J]. *SCIENTIA SINICA Informationis*, 2025, 55(3): 601-618. (in Chinese)
- [8] MAGNI L, RAIMONDO D M, ALLGÖWER F. Nonlinear model predictive control [M]. Berlin: Springer, 2009, 384: 305-313.
- [9] WELCH G, BISHOP G. An introduction to the Kalman filter [R]. Technical Report TR 95-041, Department of Computer Science, University of North Carolina at Chapel Hill, 1995.
- [10] 董晓红, 康平. 广义互相关时延估计法声源定位研究 [J]. *无线互联科技*, 2020, 17(22): 5-6.
DONG X H, KANG P. Research on the location of sound sources by generalized cross-correlation delay estimation method [J]. *Wireless Internet Technology*, 2020, 17(22): 5-6. (in Chinese)
- [11] 屈顺彪, 俞华, 芦竹茂, 等. 面向声源定位的改进广义互相关时延估计方法 [J]. *导航定位与授时*, 2021, 8(6): 118-124.
QU S B, YU H, LU Z M, et al. Improved generalized cross-correlation time delay estimation method for sound source localization [J]. *Navigation Positioning & Timing*, 2021, 8(6): 118-124. (in Chinese)
- [12] RAWAT W, WANG Z H. Deep convolutional neural networks for image classification: A comprehensive review [J]. *Neural Computation*, 2017, 29(9): 2352-2449.
- [13] 张帅, 许家悦, 李梦迪, 等. 基于皮层神经元模型的经颅磁声电刺激神经网络放电活动仿真分析 [J]. *电工技术学报*, 2021, 36(18): 3851-3859.
ZHANG S, XU J Y, LI M D, et al. Simulation of the discharge activity of neural network under transcranial magneto-acousto-electrical stimulation based on cortical neuron model [J]. *Transactions of China Electrotechnical Society*, 2021, 36(18): 3851-3859. (in Chinese)
- [14] LECUN Y, BOTTOU L, ORR G B, et al. Efficient Back-Prop [M]. In: ORR G, MILLER K R, (eds). *Neural Networks: Tricks of the Trade*. Berlin: Springer, 2000.
- [15] 刘念, 左绪忠, 何恩节, 等. Origin 软件在大学物理实验教学中的应用 [J]. *科技风*, 2024(26): 124-126.
LIU N, ZUO X Z, HE E J, et al. The application of Origin in college physical experiments [J]. *Science and Technology Wind*, 2024(26): 124-126. (in Chinese)
- [16] 杜功焕, 朱哲民, 龚秀芬. 声学基础 [M]. 南京: 南京大学出版社, 2001.
DU G H, ZHU Z M, GONG X F. *Acoustics fundamentals* [M]. Nanjing: Nanjing University Press, 2001. (in Chinese)
- [17] 陶巍, 刘建平, 张一闻. 基于麦克风阵列的声源定位系统 [J]. *计算机应用*, 2012, 32(5): 1457-1459.
TAO W, LIU J P, ZHANG Y W. Sound source localization system based on microphone array [J]. *Journal of Computer Applications*, 2012, 32(5): 1457-1459. (in Chinese)
- [18] 陈星辉, 杨媛, 邓杰, 等. 基于物理原理的声源定位实验设计 [J]. *大学物理*, 2024, 43(3): 20-24.
CHEN X H, YANG Y, DENG J, et al. Experimental design of sound source localization based on physical principle [J]. *College Physics*, 2024, 43(3): 20-24. (in Chinese)
- [19] 张芳. 基于 STM32 的麦克风阵列声源定位系统的研究 [D]. 秦皇岛: 燕山大学, 2014.
ZHANG F. Research of microphone arrays sound source location system based on STM32 [D]. Qinhuangdao: Yanshan University, 2014. (in Chinese)
- [20] 陆灏铭, 陈玮, 刘寿宝. 基于麦克风阵列的声源定位系统设计 [J]. *自动化技术与应用*, 2012, 31(4): 79-83.
LU H M, CHEN W, LIU S B. Design of acoustic source localization system based on microphone array [J]. *Automation Technology and Application*, 2012, 31(4): 79-83. (in Chinese)
- [21] 李扬. 基于小型麦克风阵列的声源定位系统设计与实现 [D]. 哈尔滨: 哈尔滨工业大学, 2014.
LI Y. The design and implementation of sound source localization system based on small size microphone array [D]. Harbin: Harbin Institute of Technology, 2014. (in Chinese)
- [22] 马亚南. 基于小型阵列 MIC 的声源定位系统设计与实现 [D]. 南京: 东南大学, 2018.
MA Y N. The design and implementation of sound source localization system based on small microphone array [D]. Nanjing: Southeast University, 2018. (in Chinese)