
第十四届中国研究生电子设计竞赛

技术论文

基于麦克风阵列的声源定位识别系统设计

Design of sound source localization recognition
system based on microphone array

参赛单位:西南交通大学

队伍名称:和尚洗头用飘柔

指导老师:蒋朝根

参赛队员:曾礼虎、付唐帅、赵维松

完成时间:2019.06.19

摘要

声源定位是指利用多个麦克风组成的具有一定空间结构的麦克风阵列获取声音信号，再由阵列的输出进行声源位置的估计。基于麦克风阵列的声源定位是声学信号处理中的一个重要问题，它在视频会议、智能和语音增强等领域有着广泛的应用。在真实的声场环境下的点声源定位研究是一项具有挑战性的课题，近年来受到越来越多的关注。人工智能语音交互技术的发展日新月异，逐渐引起了人们的广泛关注。在获取语音信号的过程中，不可避免的会受到室内环境噪声、混响以及其他说话人的干扰。

参与定位的麦克风数目越多，阵列接收到的声源信息就会越多，定位也就越准确。但随着麦克风数目的增多，不仅增加了麦克风阵列的复杂度，且接收到的信息量越大，计算复杂度越高，这样就可能无法保证系统的实时性。反之，若阵元数目太少，则麦克风接收到的信息量有限，又会降低系统的定位精度，使其无法达到预期的定位效果。

在设计麦克风阵列时，通常以结构简单、体积小为原则，且要能实现全方位定位。按照阵列的维数，可以将其划分为一维、二维和三维麦克风阵列。根据麦克风阵列的拓扑结构，可以将其划分为线性阵列、平面阵列和立体阵列等。

为了解决在有噪声和混响的复杂环境下，声源定位系统的定位精度低、实时性差和成本高等问题，本文对基于麦克风阵列的声源定位问题展开了研究。

(1) 介绍了几种常用的基于时延估计的定位方法，包括广义互相关时延估计算法和 LMS 自适应时延估计算法，在此基础上，引入了二次相关法对广义互相关时延估计算法。

(2) 建立了一个切实可行的实时声源定位系统，并详细地介绍了系统的平台组成；对麦克风阵列和动作执行模块的硬件结构进行了详细介绍，并给出了系统的软件设计流程；最后，通过实验验证了改进算法的性能和声源定位系统的可行性，并总结了影响定位精度的多种因素，包括算法本身的限制、硬件因素和环境因素等。

关键词：声源定位；噪声与混响；时延估计；广义互相关；二次相关

Abstract

The sound source localization refers to acquiring a sound signal by using a microphone array having a certain spatial structure composed of a plurality of microphones, and then estimating the sound source position by the output of the array. Sound source localization based on microphone array is an important issue in acoustic signal processing. It has a wide range of applications in video conferencing, intelligence and speech enhancement. Research on point source localization in a real sound field environment is a challenging topic and has received more and more attention in recent years. The development of artificial intelligence voice interaction technology has been changing with each passing day, which has gradually attracted people's attention. In the process of acquiring a voice signal, it is inevitable to be disturbed by indoor environmental noise, reverberation, and other speakers.

The more microphones that participate in the positioning, the more information the sound source receives from the array, and the more accurate the positioning. However, as the number of microphones increases, not only the complexity of the microphone array is increased, but also the greater the amount of information received and the higher the computational complexity, so that the real-time performance of the system may not be guaranteed. Conversely, if the number of array elements is too small, the amount of information that the microphone can receive is limited, which in turn reduces the positioning accuracy of the system, making it impossible to achieve the desired positioning effect.

When designing a microphone array, it is usually based on the principle of simple structure and small volume, and it is necessary to achieve all-round positioning. According to the dimensions of the array, it can be divided into 1D, 2D and 3D microphone arrays. According to the topology of the microphone array, it can be divided into a linear array, a planar array, a stereo array, and the like.

In order to solve the problems of low positioning accuracy, poor real-time performance and high cost in the complex environment with noise and reverberation, this paper studies the sound source localization based on microphone array.

(1) Several common positioning methods based on time delay estimation are introduced, including generalized cross-correlation delay estimation algorithm and LMS adaptive time delay estimation algorithm. On this basis, the quadratic correlation method is introduced to

generalized cross-correlation. Delay estimation algorithm.

(2) A practical real-time sound source localization system is established, and the platform composition of the system is introduced in detail. The hardware structure of the microphone array and the action execution module is introduced in detail, and the software design flow of the system is given. Finally, the performance of the improved algorithm and the feasibility of the sound source localization system are verified by experiments, and various factors affecting the positioning accuracy are summarized, including the limitations of the algorithm itself, hardware factors and environmental factors.

Key words: Sound source localization Noise and reverberation Time delay estimation

Generalized cross correlation Quadratic correlation

目录

摘要..... 2

Abstract..... 3

第 1 章 作品难点与创新..... 1

第 2 章 方案论证与设计..... 2

 2.1 原理..... 2

 2.2 模型建立..... 3

 2.2.1 理想模型建立..... 3

 2.2.2 单源多元混响模型..... 3

 2.3 麦克风阵列结构设计..... 4

第 3 章 原理分析与硬件电路图..... 6

 3.1 算法原理..... 6

 3.2 空间定位原理..... 7

 3.3 硬件电路图..... 9

 3.3.1 ESP32 原理图..... 9

 3.3.2 INMP441 电路图..... 10

第 4 章 软件设计流程..... 11

第 5 章 系统测试与误差分析..... 14

 5.1 系统测试结果..... 14

 5.2 误差分析..... 14

第 6 章 结语(结论)..... 16

 6.1 总结..... 16

 6.2 展望..... 16

参考文献..... 18

第 1 章 作品难点与创新

声源定位是指利用多个麦克风组成的具有一定空间结构的麦克风阵列获取声音信号，再由阵列的输出进行声源位置的估计。通常在没有噪声和混响、且声源距离麦克风较近的理想情况下，由单个麦克风就可以获得高质量的语音信号，但它的拾音范围有限，若声源位于麦克风的选择方向之外，就会引入大量噪声。相比于单一麦克风，麦克风阵列具有更多优势，可以弥补其在噪声处理、声源定位跟踪和语音分离等方面的不足，因此广泛应用于有嘈杂背景的语音通信环境，如仿人智能、视频会议系统、助听器和车载免提电话等。

基于麦克风阵列的声源定位是声学信号处理中的一个重要问题，它在视频会议、智能和语音增强等领域有着广泛的应用。在真实的声场环境下的点声源定位研究是一项具有挑战性的课题，近年来受到越来越多的关注。

基于麦克风阵列的声源定位问题是阵列信号处理技术的一个崭新方向，可以将其等效于宽带短时平稳信号的空间谱估计问题，其研究成果不仅可用于语音信号处理领域，还可以借鉴到其他超宽带信号处理领域以实现高精度定位，如无线通信、雷达等。随着智能语音交互技术的发展，人们对声源定位的期望越来越高，希望其应用更广泛、操作更人性化、使用更方便，这就极大地促进了有关声源定位课题的应用研究。在室内环境中，利用麦克风阵列进行语音信号的采集时，不可避免地会受到各种噪声、混响以及其他说话人的干扰等，这使得声源定位问题面临着极大的挑战。为了解决在有噪声和混响的复杂环境下，声源定位系统的定位精度低、实时性差和成本高等问题，本文对基于麦克风阵列的声源定位问题展开了深入研究。

声源定位算法主要可分为 3 类。第 1 类是基于波束形成的方法。该方法可以用于多声源的定位，但存在着对初始值敏感和需要声源和背景噪声先验知识等缺点。第 2 类是基于高分辨率谱估计的方法。该方法理论上能够对声源方向进行有效估计，但计算量大，且不适合处理人声等宽带信号。第 3 类是基于到达时间差的方法。由于该方法原理简单，计算量小，易于实现，在声源定位中得到了广泛应用。

本作品是基于麦克风阵列的声源定位识别系统设计，主要的运算工作需要在工业派上进行，因为工业派的外设接口文档较少。所以我们创新性的提出了一种使用串口连接的 ESP32 + 4 INMP441 麦克风的设计，巧妙的避开了兼容性的问题。

第 2 章 方案论证与设计

2.1 原理

声波是由于物体的振动而引起的一种波，当声源发生振动时，其周围的介质也会随之振动，声波随着介质的振动向周围扩散，因此，声波是球面波。按照声源与麦克风阵列中心的距离远近不同，可以把声场模型划分为近场模型和远场模型两类[35]。通常情况下，当声源与麦克风阵列中心的距离远大于信号波长时称为远场；反之，则为近场。设均匀线性阵列的相邻阵元间的距离（即阵列孔径）为 d ，声源的最小波长为 λ_{min} ，声源到阵列中心的距离为 r ，如图 2-1 所示，如果 $r > 2d^2/\lambda_{min}$ ，则是远场模型，否则为近场模型。

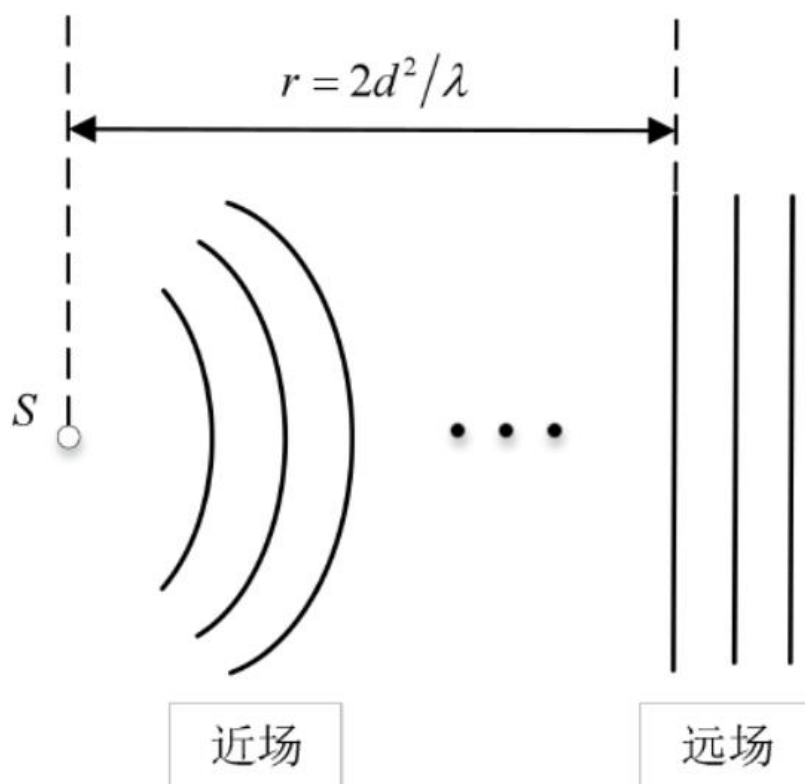


图 2-1 远场模型和近场模型

近场模型和远场模型的主要区别在于是否考虑各阵元接收信号的振幅衰减不同。近场模型把声波看作球面波，充分考虑到麦克风阵元接收的信号之间的振幅差异，而远场模型则把声波看作平面波，不考虑振幅差异，它近似地认为各接收信号之间仅存在简单的时延关系

已知通常人耳能够听到的语音信号的频率范围在 20Hz~20KHz 之间，又波长和频率的关系为 $\lambda = c/f$ ，其中，声速 $c=340\text{m/s}$ ，由此可知，语音信号的波长范围是 $[0.017\text{m}, 17.15\text{m}]$ 。但在语音信号处理中，规定的语音信号频率范围是 300Hz~3.4KHz，所以当 $d = 1.2\text{m}$ 时，区分近场和远场的分界距离 r 的范围为 $[2.54\text{m}, 28.8\text{m}]$ ，当 $d = 0.20\text{m}$ 时，近场和远场的分界距离 r 为 $[0.071\text{m}, 0.8\text{m}]$ 。

2.2 模型建立

2.2.1 理想模型建立

在理想环境下，假设室内只存在噪声，且噪声与信号之间和噪声与噪声之间互不相关，则麦克风阵列的信号接收模型如图 2-2 所示。

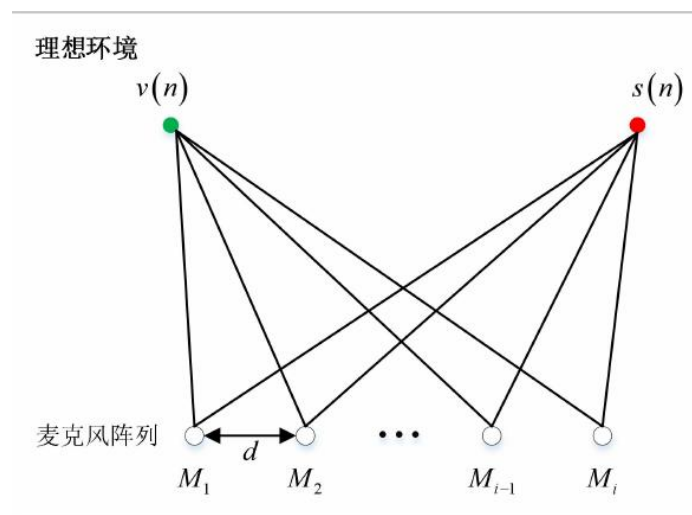


图 2-2 理想模型

2.2.2 单源多元混响模型

在室内环境中，利用麦克风阵列获取语音信号时，会受到环境噪声和多径效应的影响，此时，环境噪声、混响和语音信号将同时辐射在阵列平面上，麦克风接收到的信号

将不再是简单的语音信号，而是多个信号源的延迟与噪声信号的叠加。如图 2-3 为阵列的接收信号模型。

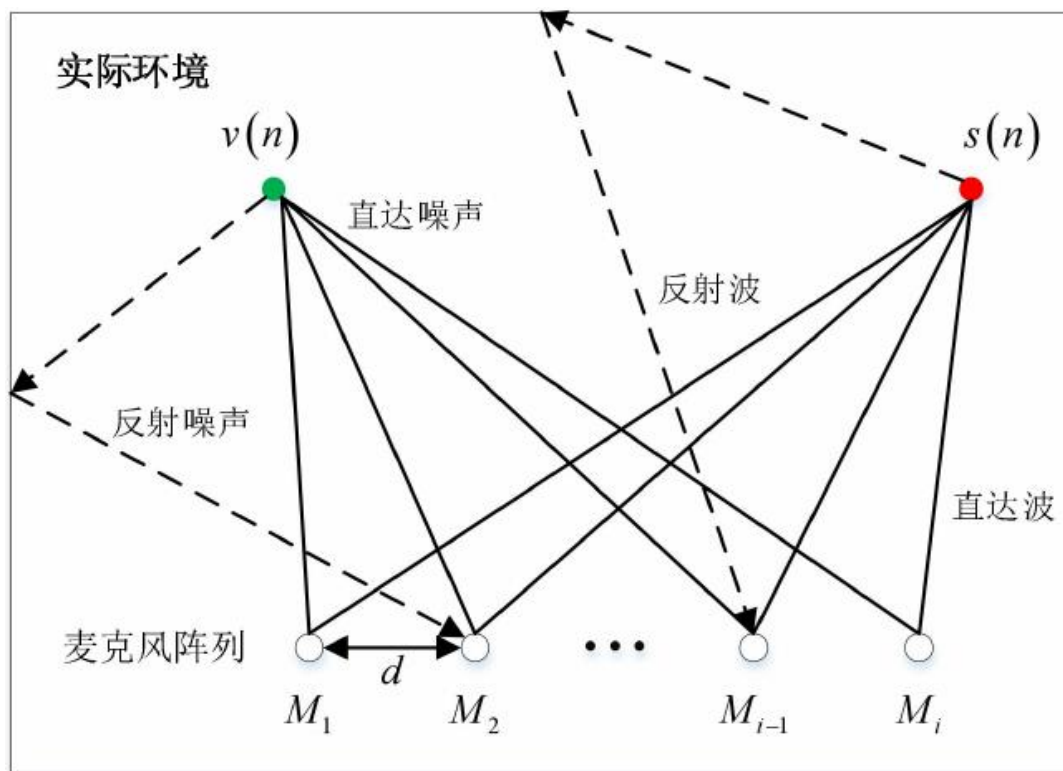


图 2-3 单源多元混响模型

2.3 麦克风阵列结构设计

麦克风阵列是由多个麦克风按照一定的形状规则放置而行成的，可以将其看作一种语音信号的采样装置。一般来说，麦克风阵列的参数包括阵元数目、阵列孔径大小、阵元间距和麦克风的分布形式等。阵元数目和阵列孔径决定了阵列实现的复杂性。理论上来说，应采用最小数量的麦克风来获得最佳的定位效果；而实际上，参与定位的麦克风数目越多，阵列接收到的声源信息就会越多，定位也就越准确。但随着麦克风数目的增多，不仅增加了麦克风阵列的复杂度，且接收到的信息量越大，计算复杂度越高，这样就可能无法保证系统的实时性。反之，若阵元数目太少，则麦克风接收到的信息量有限，又会降低系统的定位精度，使其无法达到预期的定位效果。阵列孔径用来描述空间阵列所占据的体积大小，一般阵列孔径越大，通常实现起来越难。此外，阵元数目还会影响阵列增益，当在噪声环境中获取语音信号时，通常使用阵列增益来描述阵列作为空间处理器所提供的信噪比改善程度。一般来说，麦克风的数量与阵列增益成正比。

若想要阵列有较好的分辨率，就要有较大的孔径 D ；阵列要有较高的截止频率，就需要有较小的阵列间距。然而，大孔径和小间距之间是互相矛盾的，如果两者都要满足，就只能通过增加麦克风的数量来实现。在实际应用中，通常针对特定的被测对象来权衡设计。

麦克风阵列的结构设计是决定声源定位系统定位精度的一个关键因素。在设计麦克风阵列时，通常以结构简单、体积小为原则，且要能实现全方位定位。按照阵列的维数，可以将其划分为一维、二维和三维麦克风阵列。根据麦克风阵列的拓扑结构，可以将其划分为线性阵列、平面阵列和立体阵列等。目前，常用的麦克风阵列结构一般有几类：线性阵列、十字型阵列、圆型阵列和螺旋型阵列等，如图 2-4 所示。其中，一维线性阵列中的所有麦克风呈线型排列，其阵元数目通常为 2~10 个，这种阵列结构几乎不受声源位置的影响，但它不能实现全方位定位，只能定位部分区域的声源；二维阵列是由一维阵列扩展而来的，弥补了一维阵列的不足，但其结构较一维阵列复杂，增加了系统的实现难度和计算复杂度；三维阵列定位更准确，但它的结构更复杂，计算复杂度更高，且可供选择的算法不多。

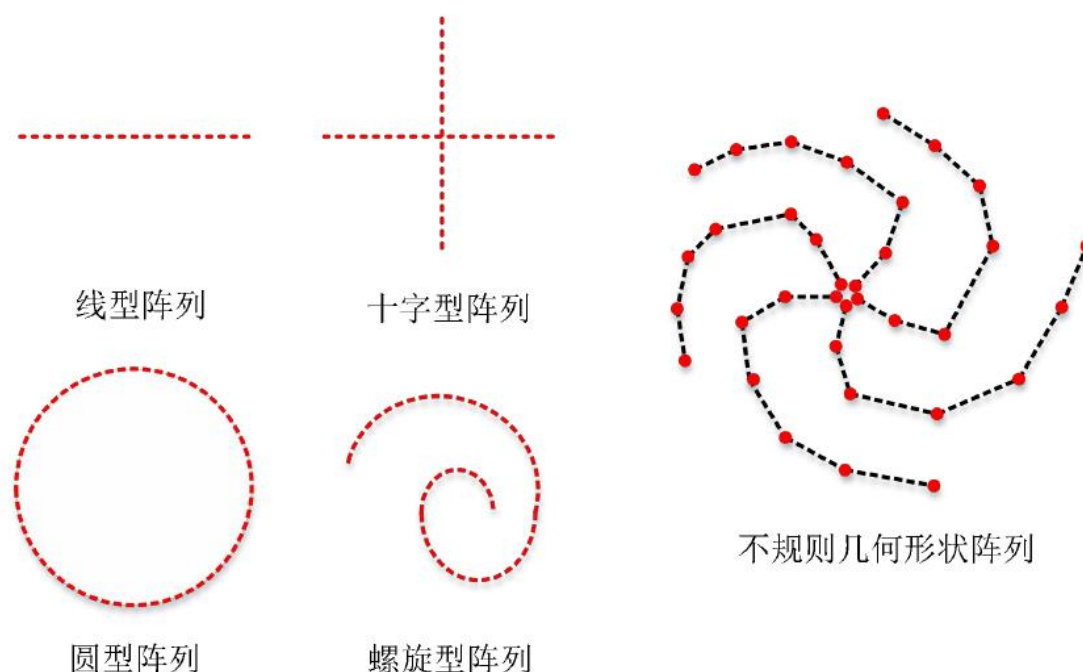


图 2-4 常用的阵列结构

综合考虑，为了满足系统对定位精度的要求，并降低系统的计算复杂度和实现难度，同时使其易于安装在上，本文采用二维平面阵列结构，设计了一种四元十字型平面麦克风阵列结构，用于下一步研究工作。

第 3 章 原理分析与硬件电路图

3.1 算法原理

我们使用的是快速傅里叶计算时间差的算法：根据麦克风到达时间差算出声程差，最后利用四路麦克风的声程差，计算物体在三维空间的具体位置。

首先我们用四路麦克风收集声音模拟数据，然后按采样定理对模拟信号进行等时间间隔采样，将得到的一系列时域上的样值，这些样值在时间上是离散的值，但在幅度上仍然是连续模拟量,然后将其进行 AD 转换，转化为数字信号。然后将数字信号进行傅里叶变换，将时域上的信号转化为频域上的信号，方便进行计算。最后将信号进行傅里叶逆变换，将频域信号转化为时域，进行峰值检测，算出每个麦克风与第一个麦克风的时间延迟。下面举例说明计算两个麦克风之间时间差的方法，如图 3-1 所示：

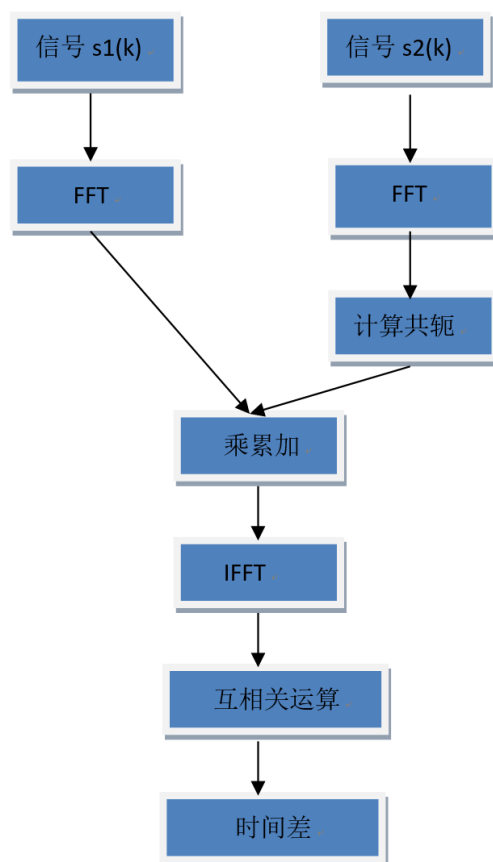


图 3-1 快速傅里叶计算时间差

说明:

1) 我们将信号进行快速傅里叶变换是因为:

傅里叶原理表明:任何连续测量的时序或信号,都可以表示为不同频率的正弦波信号的无限叠加。而根据该原理创立的傅里叶变换算法利用直接测量到的原始信号,以累加方式来计算该信号中不同正弦波信号的频率、振幅和相位。我们使用傅里叶变换将原始信号转化为便于处理的正余弦信号。一个正弦曲线信号输入后,输出的仍是正弦曲线,只有幅度和相位可能发生变化,但是频率和波的形状仍是一样的。变换后的信号我们称之为频域信号,或者叫做信号的频谱图,方便计算机后续计算处理。

2) 频域信号比时域信号更方便处理是因为:

傅里叶变换是线性算子,从而使得线性微分方程的求解可以转化为常系数的代数方程的求解,从而提供了计算卷积的一种简单手段;

3) 互相关运算估计最大值:

因为同一个声源信号,在传播过程中,它的波形图会产生变化,但是变化不大,互相关函数可以计算信号之间的相关性。利用极大似然估计原理,互相关函数算出最大值时,对应的时间差就是我们最后时间差的估计值。

3.2 空间定位原理

我们的麦克风是由 4 个麦克风组成的四元十字型麦克风阵列,具有结构小巧、便于安放的优点。阵列呈十字型结构对称分布在 x 轴和 y 轴上,采集同一声源信号。先估计各麦克风之间的延迟差,再利用几何关系计算声源的空间位置。由图 3.2 可见,麦克风 M_1, M_2, M_3, M_4 的坐标分别为 $M_1(a, 0, 0)$, $M_2(0, -a, 0)$, $M_3(-a, 0, 0)$, $M_4(0, a, 0)$, 声源信号的坐标为 $s(x, y, z)$, 声源 s 与坐标原点的距离用 R 表示; 声源 s 与 x 轴正方向的夹角, 即方位角用 ϕ 表示; 声源 s 与 z 轴正方向的夹角, 即俯仰角用 θ 表示。

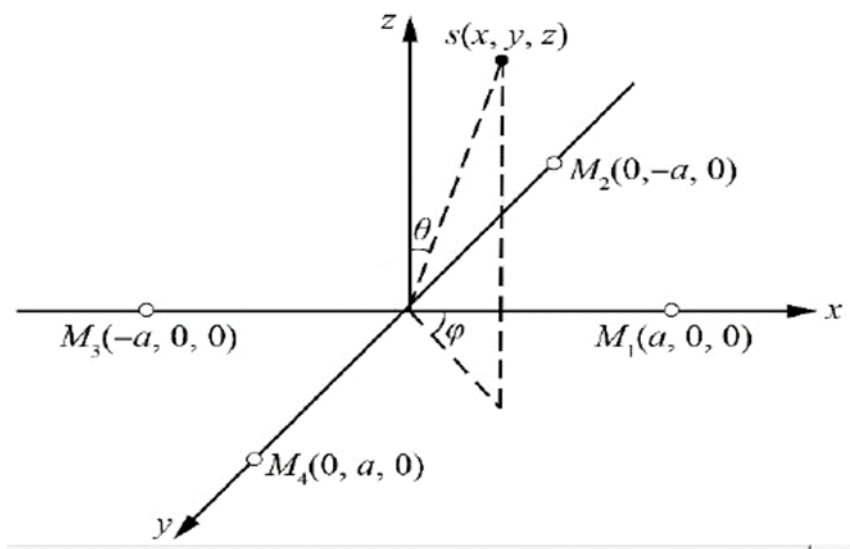


图 3-2 四元十字形麦克风阵列

假设分别用 r_i ($i = 1, 2, 3, 4$) 表示声源到麦克风的距离, 用 t_i ($i = 1, 2, 3, 4$) 表示声源到达麦克风的时间, 用 t_j^i ($i, j = 1, 2, 3, 4$) 表示声源到达麦克风 M_i 与 M_j 之间的时间延迟. 假设声源位于远场, 则近似求得目标声源的位置表达式为

$$\begin{cases} R \approx \frac{c(\tau_{12}^2 + \tau_{14}^2 - \tau_{13}^2)}{2(\tau_{13} - \tau_{12} - \tau_{14})}, \\ \varphi = \tan^{-1} \left(\frac{r_2^2 - r_4^2}{r_3^2 - r_1^2} \right), \\ \theta \approx \sin^{-1} \left(\frac{c \sqrt{\tau_{13}^2 - (\tau_{14} - \tau_{12})^2}}{2a} \right) \end{cases}$$

图 3-3

在该模型中, 我们使用 1 个麦克风与其他麦克风的时延差值进行定位计算, 运算非常简单, 计算量很小, 而且定位结果十分准确, 且能满足系统的实时性要求。

3.3 硬件电路图

3.3.1 ESP32 原理图

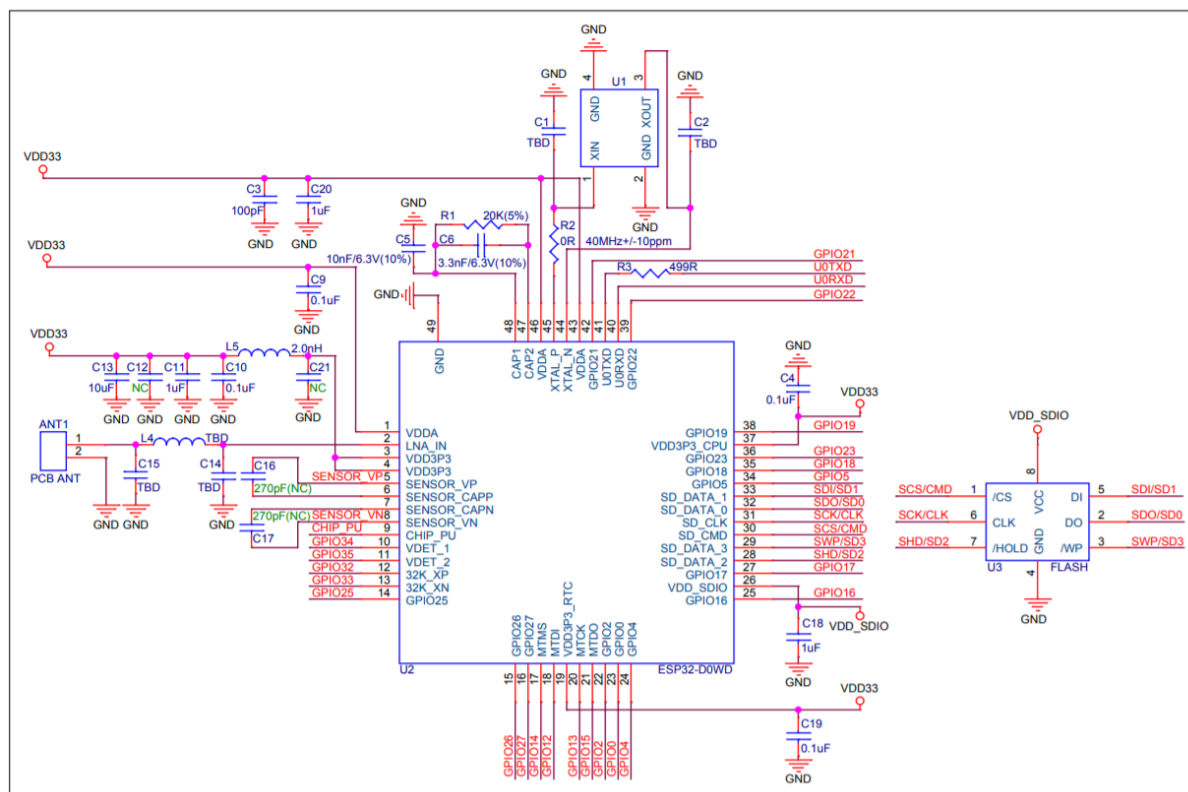


图 3-4 ESP32 的核心电路图

3.3.2 INMP441 电路图

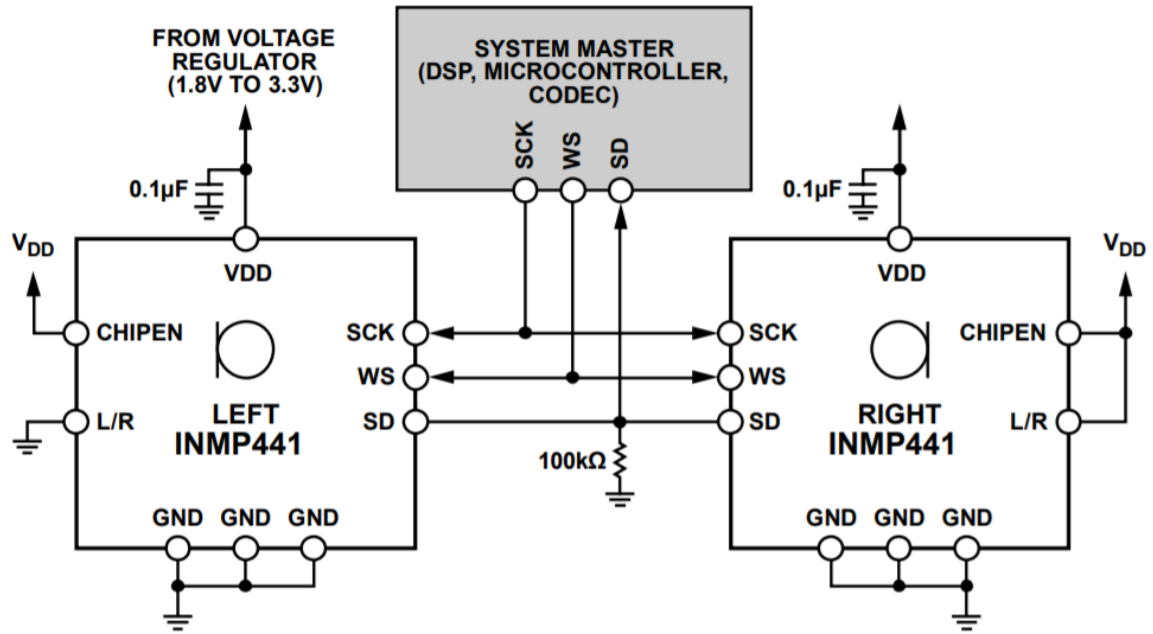


图 3-5 INMP441 电路图

第 4 章 软件设计流程

图 4-1 是整个麦克风阵列的整体工作流程图，首先使用 ESP32 通过 4 个全向数字麦克风收集环境声音数据，收集到之后将数据发送至工业派进行声源的判定，如果不是声源，则继续收集，如果确定为声源，则将收集到的数据在工业派上进行处理。最后将处理过程显示在控制台上。如图 4-2 所示

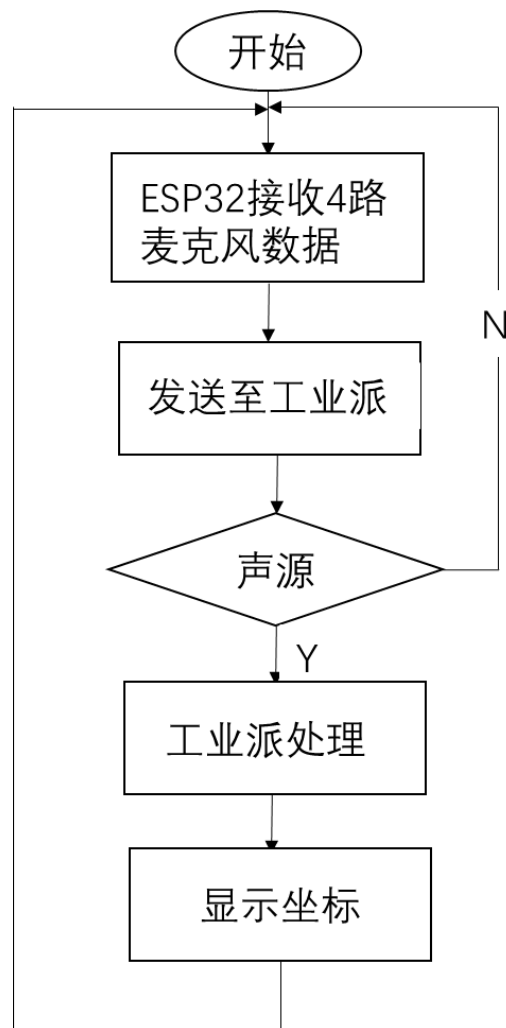


图 4-1


```
num of data2:39
data2:6591489.000000
data2:5353473.000000
data2:5371391.000000
data2:6293759.000000
data2:3115519.000000
data2:3215105.000000
data2:5310465.000000
data2:3657473.000000
data2:3497729.000000
data2:4665601.000000
data2:6771713.000000
data2:11514623.000000
data2:5316353.000000
data2:6192895.000000
data2:4563711.000000
data2:5382911.000000
data2:2959361.000000
data2:2836993.000000
data2:2397185.000000
data2:2250751.000000
data2:6581503.000000
data2:3815423.000000
data2:2033665.000000
data2:4482815.000000
data2:2872831.000000
data2:3153407.000000
data2:6006273.000000
data2:3352575.000000
data2:2392575.000000
data2:2590209.000000
data2:3509761.000000
data2:3118081.000000
data2:2720769.000000
data2:2873345.000000
data2:3790335.000000
data2:3613695.000000
data2:3019007.000000
```

图 4-2 二号麦克风的数据

图 4-3 是工业派处理声源数据的流程示意图，首先将 4 路麦克风数据分离出来，运用快速傅里叶变换确定每路麦克风的波形相似度，得到每路麦克风之间的时间差，然后将时间差传入空间定位算法中进行计算，最终得出声源相对基准麦克风在空间中的位置，并将最终结果显示在控制台上，如图 4-4 所示

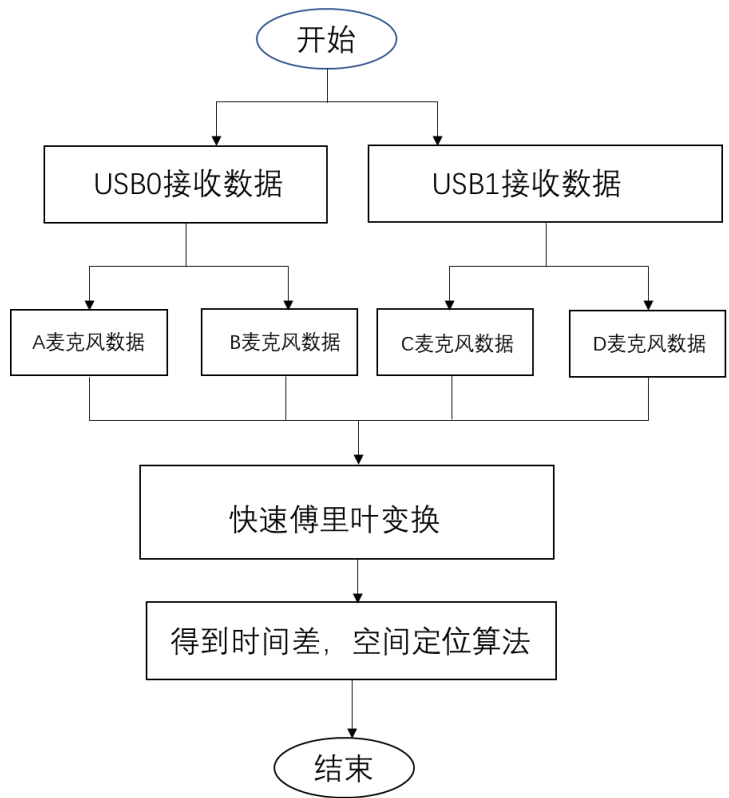


图 4-3

```
the target location:
Distance from the reference microphone:0.574000m
Direction angle with the X axis:39.800000
Direction angle with the z axis:73.600000
the target location:
Distance from the reference microphone:0.624000m
Direction angle with the X axis:46.700000
Direction angle with the z axis:67.400000
the target location:
Distance from the reference microphone:0.761000m
Direction angle with the X axis:285.400000
Direction angle with the z axis:70.800000
the target location:
Distance from the reference microphone:0.712000m
Direction angle with the X axis:278.900000
Direction angle with the z axis:61.300000
```

图 4-4 声源定位结果

第 5 章 系统测试与误差分析

5.1 系统测试结果

我们对系统进行了测试，结果如图 5-1，图中前两次结果为测试员在同一地点拍手得到的结果，后两次结果为测试员移动了位置后拍手得到的结果。

```
the target location:
Distance from the reference microphone:0.574000m
Direction angle with the X axis:39.800000
Direction angle with the z axis:73.600000
the target location:
Distance from the reference microphone:0.624000m
Direction angle with the X axis:46.700000
Direction angle with the z axis:67.400000
the target location:
Distance from the reference microphone:0.761000m
Direction angle with the X axis:285.400000
Direction angle with the z axis:70.800000
the target location:
Distance from the reference microphone:0.712000m
Direction angle with the X axis:278.900000
Direction angle with the z axis:61.300000
```

图 5-1 测试结果

在前两次拍手过程中测试员手部没有大范围的移动，但是得到的声源距离结果第一次为 0.574m，第二次为 0.624m，相差有 5cm；与 X 轴的夹角第一次为 39 度，第二次为 46 度，相差有 7 度；与 Z 轴夹角第一次结果为 73 度，第二次为 67 度，相差 6 度。后两次测试中，第一次结果为 0.761m，第二次为 0.712m，相差是 4.9cm；与 X 轴的夹角第一次为 285 度，第二次为 278 度，相差有 7 度；与 Z 轴夹角第一次结果为 70 度，第二次为 61 度，相差 9 度。

5.2 误差分析

(1) 软件算法：本系统中使用的时延估计算法不是对时延的准确估计，仅能得到近似的估计值。另外，在几何空间定位中，计算方位角和距离时，也存在一些近似计算。这些近似估计和近似计算也引入了误差，增加了系统最终的定位误差。

(2) 硬件因素：本系统中，使用的麦克风比较普通，其灵敏度和精度有限，加上麦

麦克风是有指向性的，对位于大部分位置的声源信号的接收效果比较好，但对于某些位置来说就会受到限制。其次，阵列结构的设计也会对定位结果产生一定影响。由于这些硬件因素的影响，导致最终的定位效果无法达到理想状态。

（3）环境因素：由于实际环境的不可预测性，每次实验过程中的随机噪声都是不同的。本系统虽然对噪声进行了一些处理，但并不能完全消除噪声，尤其是一些突发噪声，如突然的开门声等。

（4）测量误差：在对麦克风阵列位置和声源位置真实值的测量过程中都存在着不可避免的人为测量误差。此外，由于声音在不同条件下的传播速度是不一样的，而声速的测量又比较困难的，只能通过测量温度、湿度、风速、风向和气压等推算出来，其中温度对声速的影响最大，其他因素也有一定的影响，但是影响的程度要小得多，已知在温度为零度时，一般取值为 332m/s 。本实验中，为了方便计算取 $c=340\text{m/s}$ 。这就在一定程度上影响了时延估计的准确性。

综上所述，系统的硬件和软件因素是影响系统定位精度的主要因素。当这些因素导致时延值发生微小变化时，就会带来较大的定位误差，有时甚至会导致错误的结果或无法定位。选择性能较好的麦克风、设计合理的麦克风阵列结构以及采用适当的声源定位算法，能在很大程度上提高系统的定位精度，减小定位误差。

第 6 章 结语(结论)

6.1 总结

一个声源定位系统，不但要能够处理有噪声和混响的复杂环境，还应具备自动定位与跟踪目标声源的能力。而现有的声源定位系统在实际应用中往往存在着定位精度低、实时性差和成本高等问题，基于以上两个重点，本文对传统的基于时延估计的声源定位算法进行了研究，设计并实现了一个切实可行的基于麦克风阵列的声源定位系统，且利用该系统平台验证了本文算法的定位性能。主要工作总结如下：

(1) 建立了基于噪声和混响的阵列信号接收模型—单源多元混响模型；介绍了麦克风阵列的支架设计技术，并根据阵列结构的设计原则和声源定位系统的实际需求，设计了一种四元十字型平面麦克风阵列结构，且推导了其几何空间定位方法。

(2) 采用了滤波去噪和加窗分帧等预处理操作，并采用一种双门限语音端点检测算法以避免错误的定位，提高整个系统的运行效率。

(3) 采用二次相关法改进 GCC 算法，且提出了一种基于时延估计改进的声源定位算法，实验结果表明，本文提出的改进算法有效地提高了时延估计的精度，在噪声与混响的环境下，其定位性能更优越。

(4) 实现了一个切实可行的实时声源定位系统，通过三组实验，分别分析了采样频率、阵元间距和声源距离对于定位结果的影响，通过实验验证了改进算法的性能和声源定位系统的可行性。

6.2 展望

本文实现的基于麦克风阵列的机器人声源定位系统取得了很好的预期成果，如何进一步提高系统的定位精度，仍需深入研究，进一步工作总结如下：

(1) 本文实现的声源定位系统是将阵列采集的语音信号经过 AD 采样后，通过 USB 传至上位机进行数据处理与定位计算，而大量的数据传输过程必然会影响系统的实时性，因此，本文考虑直接通过硬件 来进行信号采集与算法处理，这样减少了传输时间，还能增加系统的实际应用性。

(2) 如何进一步提高声源定位的精度, 获得准确的声源位置, 可以考虑与图像识别和图像处理等方法相结合进行定位。

(3) 本文重点研究的是室内环境中的单声源定位, 然而实际应用中, 往往会出现多声源的情况, 如何在不成倍地增加计算量和提高系统复杂性的情况下, 进行多声源的定位与跟踪, 还需要进一步研究。

总而言之, 要提高系统的定位准确性, 必须将算法研究、语音信号的采集质量和硬件系统的性能结合起来。

参考文献

- [1] 王政, 刘君, 晏克俊,等. 基于四元麦克风阵列时延差值的全方位角声源定位方法研究[J]. 西安理工大学学报, 2017(4).
- [2] 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]. 传感器与微系统, 2012, 31(4):79-81.
- [3] 张明瀚. 基于时延估计的声源定位系统研究[D]. 重庆大学, 2009.
- [4] 赵圣. 基于麦克风阵列的仿人智能机器人声源定位技术的研究[D]. 哈尔滨工业大学, 2013.
- [5] 苏新霞. 分布式麦克风阵列跟踪算法研究[D]. 大连理工大学, 2011.
- [6] 刘超. 基于麦克风阵列的声源定位算法研究[D]. 南京大学, 2015.
- [7] 徐梵洁. 基于麦克风阵列的声源定位算法及系统实现[D]. 西安理工大学, 2017.
- [8] 周峰. 室内麦克风阵列声源定位算法研究和实现[D]. 复旦大学, 2009.
- [9] 彭奎. 一种基于麦克风阵列的声源定位系统及其应用研究[D]. 安徽大学, 2014.