



南京大學

本科畢業論文

院 系 匡亞明學院

專 業 計算機科學與技術

題 目 基於超聲波和 IMU 的室內定位系統

年 級 2016 級 學 號 161240005

學生姓名 陳勇虎

指導老師 謝磊 職 稱 副教授

提交日期 2020 年 5 月 20 日

南 京 大 学

本科生毕业论文（设计）指导情况记录

开 题 简 况	论文题目：
	<div>1、选题质量（简述选题与专业培养目标、专业要求关系、题目难度、工作量、创新性、理论性、实用性）</div> <div>2、开题意见：</div> <div>指导教师签名：</div> <div>年 月 日</div>
中 期 检 查	<div>指导教师检查论文的进展情况：（指导和培养学生查阅文献资料、综合运用知识、研究方案设计、研究方法和手段运用和外文应用等能力简况）</div> <div>指导教师签名：</div> <div>年 月 日</div>

南 京 大 学

本科生毕业论文（设计）指导教师评阅意见

指导教师评语：

指导教师签名：

年 月 日

南 京 大 学

本科生毕业论文（设计）评阅教师评阅意见

评阅教师评语：

评阅教师签名：

年 月 日

南 京 大 学

本科生毕业论文（设计）答辩记录、成绩评定

答辩记录：

答辩记录人签名：

答辩小组评语：

答辩小组成员：_____

成绩_____

组长签名：

答辩时间： 年 月 日

南京大学本科生毕业论文(设计、作品)中文摘要

题目：基于超声波和 **IMU** 的室内定位系统

院系：匡亚明学院

专业：计算机科学与技术

本科生姓名：陈勇虎

指导老师（姓名、职称）：谢磊副教授

摘要：

基于 IMU 和智能手机的室内定位系统

关键词：IMU; 智能手机;

南京大学本科生毕业论文(设计、作品)英文摘要

THESIS: IMU

DEPARTMENT: School of Computer Science

SPECIALIZATION: Computer Science and Technology

UNDERGRADUATE: Yonghu Chen

MENTOR: Professor Lei Xie

ABSTRACT:

IMU

KEY WORDS: IMU, smart phone,

目 录

目 录	V
1 绪论	1
1.1 研究背景及意义	1
1.2 国内外研究现状	1
1.2.1 室内定位方案	1
1.3 研究内容与本文主要工作	1
1.4 论文的整体结构	1
2 声源定位相关研究工作	3
2.1 声源定位方案	3
2.1.1 声源定位技术的应用	4
2.1.2 远程视频会议及远程教学等	4
2.2 室内定位的信道模型	4
2.2.1 基于 RSS 的信道模型	4
2.2.2 基于 TOA 的信道模型	4
2.2.3 基于 TDOA 的信道模型	4
2.2.4 射线跟踪信道模型	4
2.3 室内定位技术原理	4
2.3.1 几何定位法	4
2.3.2 场景分析法	4
3 智能手机的感知和识别	5
3.1 双麦克风的节点定位	5
3.2 智能手机的传感器	6
3.2.1 加速计传感器	6
3.2.2 陀螺仪	6
3.2.3 重力传感器	6
3.2.4 磁力计	6

4	基于双麦克风手机的声源定位	7
4.1	offline 数据处理流程	7
4.2	数据采集	8
4.3	声音信号处理	12
4.3.1	信号除噪	12
4.3.2	异常点移除	12
4.3.3	有效声音提取	14
4.3.4	TDOA 计算	14
4.3.5	TDOA 处理	15
4.4	offline 定位方法	16
4.5	offline 到 online 的迁移	16
5	模型系统实现与实验分析	17
5.1	开发环境与平台	17
5.2	系统设计	17
5.2.1	多模块设计	17
5.2.2	界面与运行方式	17
5.3	实验与分析	17
5.4	本章小结	17
6	总结与展望	19
6.1	工作总结	19
6.2	前景展望	19
	参考文献	21
	致 谢	23

第一章 绪论

1.1 研究背景及意义

在室内环境 (办公室, 住所) 中想要寻找一个小物品, 诸如钥匙, 硬币, 可以说是一件令人费神的事情了^[1]。在游戏世界和虚幻的电影场景中, 我们都或多或少接触过类似于“寻宝罗盘”的神奇宝物, 它可以类似一个掌上的指南针, 时刻指示“宝物”的方向。随着各式各样移动智能设备的广泛涌现, 例如智能手机, 智能手表, 这些设备不仅自带了诸多传感器, 同时可以为用户提供很好的 UI 界面。因此, 如果可以利用这些设备实现类似于寻宝罗盘的功能, 用于寻找我们室内环境中的小物品, 无疑是有趣且富有创造性的。

随着对位置精度要求的提高, 对于定位技术的研究也越加深入。传统的全球定位系统 (Global Positioning System, GPS)^[2] 是现代定位系统的首选, 然而, 在室内环境中, 由于卫星信号受到建筑物的削弱, 遮蔽, 以及在复杂的室内环境下引起的信号反射, 折射, 透射等造成的多径和非视距传输现象, 导致定位误差较大, 所以卫星定位的精度已经无法满足室内定位的需求, 因此以无线传感网为媒介的定位系统逐渐发展起来。

1.2 国内外研究现状

1.2.1 室内定位方案

1.3 研究内容与本文主要工作

1.

1.4 论文的整体结构

本文的各章节组织结构如下:

第一章: 绪论

第二章: 相关工作

第三章：实验

第四章：总结与讨论

。 。 。

第二章 声源定位相关研究工作

2.1 声源定位方案

在现代城市中，利用机器视觉实现的机器人导航系统已是非常普遍。然而由于机器视觉本身存在的问题，使得利用机器视觉进行机器人导航在一些环境下表现的差强人意，比如，在光线不足的情况下，或者是在机器人的“视野”所不能到达的地方，机器人就找不到目标在空间中的位置，如果给机器人安装上“耳朵”，给予机器人声源采集装置，在进行人机交互的时候，人就可以通过语音告诉机器人自己现在所处的位置，机器人对声源信号进行处理，求解出声源的具体位置，然后制定关节运动决策，使交互对象重新进入其视野之内。

- 2.1.1 声源定位技术的应用
- 2.1.2 远程视频会议及远程教学等
- 2.2 室内定位的信道模型
 - 2.2.1 基于 RSS 的信道模型
 - 2.2.2 基于 TOA 的信道模型
 - 2.2.3 基于 TDOA 的信道模型
 - 2.2.4 射线跟踪信道模型
- 2.3 室内定位技术原理
 - 2.3.1 几何定位法
 - 2.3.1.1 多边测距法
 - 2.3.1.2 双曲线定位法
 - 2.3.1.3 三角定位法
 - 2.3.2 场景分析法

第三章 智能手机的感知和识别

3.1 双麦克风的节点定位

近些年来，基于声学传感器的定位系统越来越受到人们的关注，依赖于多个同步的麦克风，采集到的多路同步信号，用以后续的信号处理。随着无线传感网的不断发展进步，无线传感网定位问题也是一个比较热门的课题。而节点定位作为目标定位的前提条件，节点的定位精度也影响着目标定位的结果。然而，尽管目前有很多节点定位算法，但是因为功耗，成本等客观条件的限制，这些定位算法在很多场景中并不能很好的应用。

随着可移动智能设备的普及，尤以智能手机更为常见。自 2018 年起，大部分的智能手机都开始集成两个甚至多个麦克风，除了底部的通话话筒以外，智能手机的顶部都还有一个“小孔”(不同的手机可能有所区别，例如 iPhone 手机的次麦克风可能位于主摄像头的附近)(如图 3-1)，这就是手机的降噪麦克风，它的主要作用是实现有噪声环境下的高质量通话。



图 3-1: iPhone8 的次麦克风手机位于摄像头附近

手机上设有的两个电容式麦克风尽管作用不同，但是其性能相同，对声音信号的收集能力相同。利用双麦克风系统的时间同步(两个麦克风是集成于同一个设备的，因此完全时钟同步)特性，即使两个麦克风(如图 3-2)之间的距离较短(16cm 左右，不同手机略有差异)，也可以根据声音到达两个麦克风时刻的差异计算 TDOA，因此，利用双麦克风系统手机结点作为实验设备有效解决了研究过程中时钟同步的问题。

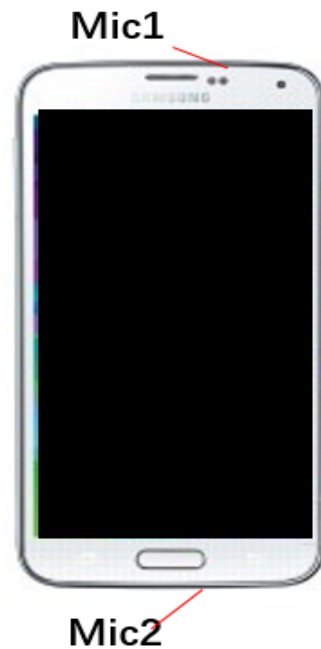


图 3-2: Dual-microphone phone

目前已经有很多利用双麦克风结点实现声源定位或者其他有趣的研究成果。例如利用双麦克风智能手机进行手势轨迹感知^[3];Aarab 使用双麦克风节点定位的思路测算说话人的位置^[4]。

3.2 智能手机的传感器

随着技术的进步，手机已经不再是一个简单的通信工具，而是具有综合功能的便携式电子设备。大多数的 Android 设备都有内置传感器，用来测量运动，环境变化等。例如重力传感器可以用来推断用户的手势和动作，温度传感器等可以用来探测环境，地磁传感器等可以用来指示方向。

3.2.1 加速计传感器

3.2.2 陀螺仪

3.2.3 重力传感器

3.2.4 磁力计

第四章 基于双麦克风手机的声源定位

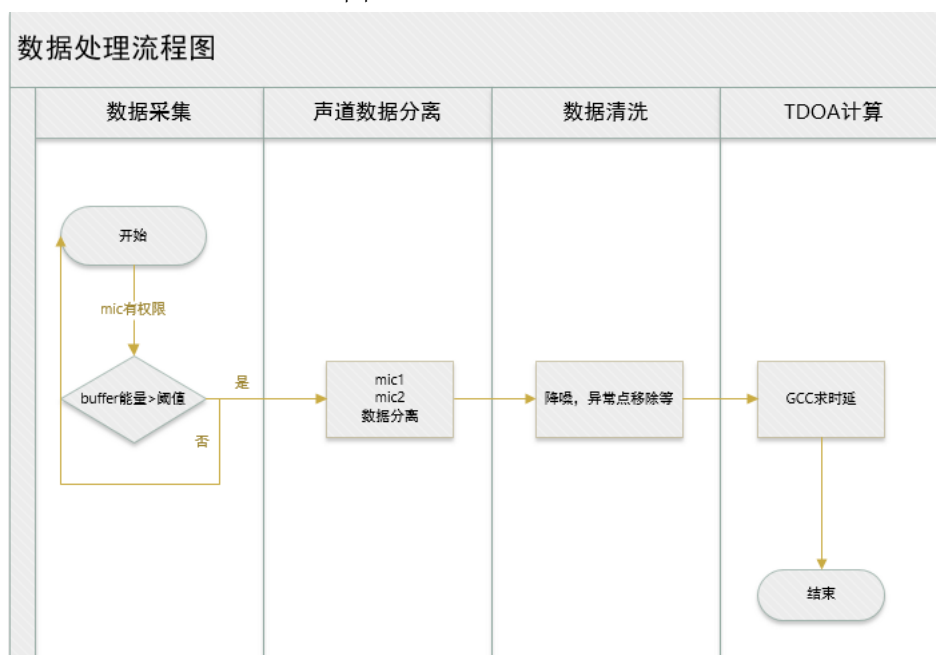
近些年来，智能手机已经成为人们生活中不可缺少的一部分，手机上集成的传感器的发展，使得智能手机的功蟹越来越强大。在前面我们已经简述到大部分智能手机已经配备两个麦克风，用于消除我们在日常通话中的环境噪声，并且可以很好的解决的时钟同步问题。因此，利用智能手机的优势，通过获取同一个手机两个麦克风的声数据，我们从中过滤出我们想要的声音信号，得到两路同步的声音信号，利用智能手机计算出两个麦克风的到达时间差 (TODA), 对 TDOA 信号进行后续的处理，就可以实现一些有意义的功能。

。 。 。

4.1 offline 数据处理流程

本小节中，首先介绍 offline 算法的处理流程，并逐渐过渡到最终系统使用的 online 算法的处理流程。offline 算法中，智能手机会收集声音文件并产生一个 wav 文件，通过 MATLAB 分析该 wav 文件，计算出 TDOA。该 offline 算法处理流程是为了通过处理一段时间窗口内，设备没有发生大幅度状态变化 (包括角度，位置等参数) 的情况下，计算出静态声源相对于该位置的方向 (Speaker Direction Finding (SDF))。在一段时间窗口内的处理流程如下图所示 (4-1)。

图 4-1: offline flowchart



4.2 数据采集

本文利用具有双麦克风系统的智能手机 (Samsung 5) 作为传感器结点的进行声音信号的手机, 在分析阶段, 使用 MATLAB 科学工具对采集到的信号进行解析显示。我们利用另一台手机发出了一个频率在 2000Hz-2700Hz 变换的线性调频信号, 如图 4-2 所示, 图中蓝色的标记为 LY 的数据以及红色的标记为 RY 的数据, 分别代指两个麦克风采集的数据。图中的上一部分图中, 横坐标代指数据采样点, 纵坐标是对数据进行标准化后的结果, 因此无量纲。但是在下部分图中, 横坐标是标准化的频率, 单位为 $\pi * \text{radians/sample}$ (频率标准化的相关概念不做赘述, 可以参考^[5]), 纵坐标为能量谱, 这里我们很容易看出两个麦克风的采集的数据的能量谱是非常接近的, 换句话说, 两个麦克风的采集能力非常接近。这里我们对原始数据取样其中发生部分连续的 400 个点。如图 4-3 所示, 同样我们依然可以看出两个麦克风的采集能力是非常接近的, 另一方面, 即便我们暂时无法确定 TDOA 的值, 但是我们可以清晰看出两路声音数据是差了几个样本点的, 因此只要计算出相差的样本点数, 根据采样率 (44100Hz) 我们就可以求解出 TDOA 的近似的结果。

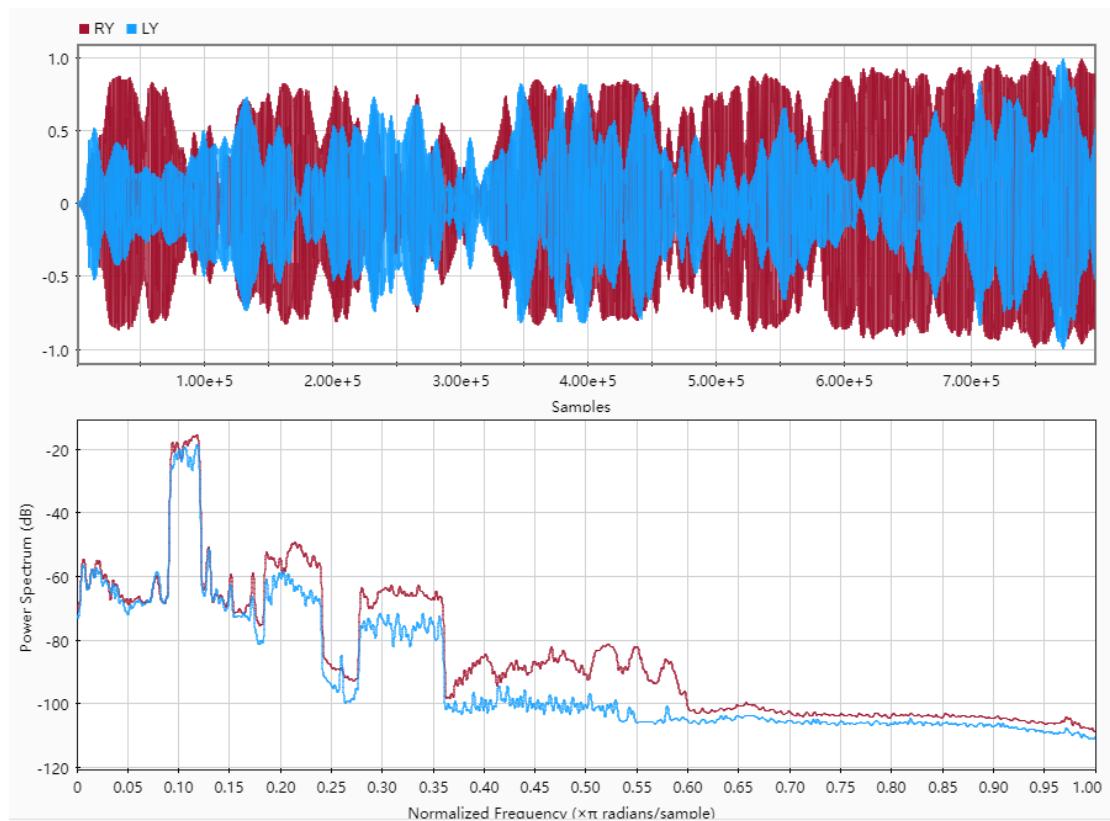


图 4-2: 双麦克风信号

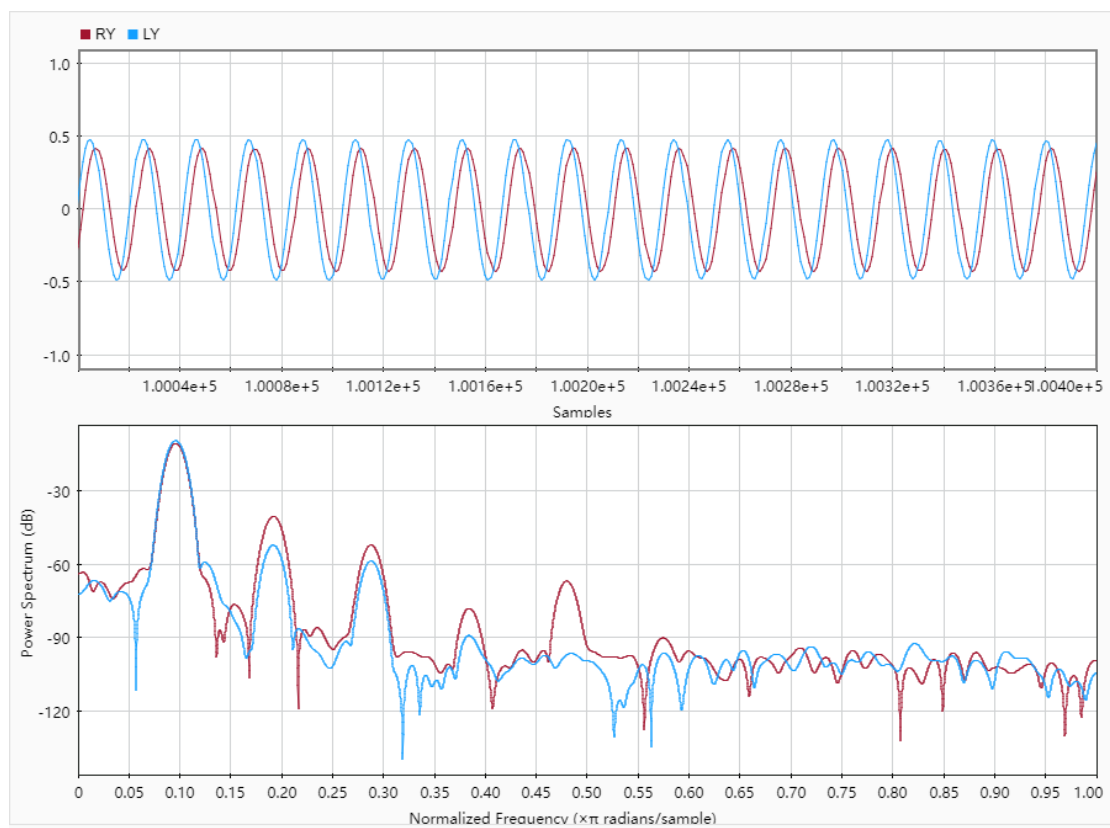


图 4-3: 双麦克风信号取样

下表中 4-1 展示了采集声音信号的手机参数，实验中使用了 SamSung S5 手机作为实验设备，S5 集成了两个同步的麦克风，分别位于手机的底部和顶部，而这个两个麦克风的语音采集能力完全相同。目前的智能手机最高采样率都可以达到 44100Hz，每秒钟就可以获得 44100 个声音信号的样本值。

表 4-1: SamSung S5 设备参数

参数	描述
灵敏度	-26dB
麦克风距离	12cm
采样频率	44100Hz

本文利用内置双麦克风的 S5 手机进行实时采集声音信号，首先介绍 **offline** 处理的，对信号进行汇总处理和分析。在底层获取的声音信号实际是个一维数组，当然利用 **MATLAB** 进行读取，会转换成两个一维数组，分别代表两个麦克风的数据。在原始的底层存储中，一维数组的值即是采集的到数据。如图 4-4 所示，实验中，我们采用 16bit 的 **stereo** 的格式进行采集，因此根据图中的存储方式很容易分别获得两个麦克风的数据，并且两组数据是完全满足时钟同步的。为了环境噪声或者其他声音的干扰，本文中设置了一个基于能量的硬阈值，当每个 **buffer** 信号的能量低于该阈值时，则认为没有声源发生，当平均能量高于该阈值时，认为有声源发生，从而进行后面的处理。



图 4-4: 数据存储方式

数据采集阶段核心代码如图 4.1 所示。手机调用麦克风不断检测周围环境的
声音信号，当信号端的平均能量高于阈值时，系统判断声源发声，并且将两路
的声音信号进行分离，用于后续的处理。

算法 4.1 dual-microphone signals acquisition(双麦克风信号采集)

```

1: mAudioRecord.startRecording()
2: //系统检测周围环境中的声音信号
3: while isRecording do
4:   //读取两路声音信号
5:   read = mAudioRecord.read(mBuffer, 0, BUFFER_SIZE)
6:   if read <= 0 then
7:     return false
8:   else
9:     //根据设定的阈值判断是否有声源发声
10:    sum = 0
11:    for i = 0 to read do
12:      sum += Math.abs(mBbuffer[i])
13:    end for
14:    //达到设定的阈值，认为有声源发声
15:    if sum / read > thresholdValue then
16:      //两路信号分离
17:      index = 0
18:      for i = 0 to mBuffer.length/2 do
19:        if i % 2 == 0 then
20:          System.arraycopy(mBuffer, i * 2, mBuffer1, 0, 2)
21:          LY[index] = ((mBuffer1[0]&0x000000FF) | (((int)mBuffer1[1])<<8))/
22:            32768.0
23:          LYList.add(LY[index])
24:          index = i / 2
25:        else
26:          System.arraycopy(mBuffer, i * 2, mBuffer2, 0, 2)
27:          RY[index] = ((mBuffer2[0]&0x000000FF) | (((int)mBuffer2[1])<<8))/
28:            32768.0
29:          RYList.add(RY[index])
30:        end if
31:      end for
32:    end if

```

```

31:      //other code for further data-processing
32:  end if
33: end while

```

4.3 声音信号处理

从底层获得的两路原始信号包括了真实声源信号，环境噪声以及其他各种异常点，本小节将会利用滤波器降低环境噪声的影响，同时使用加窗移动的方式去除采集中的异常值，对原始的信号进行清洗处理。

4.3.1 信号除噪

常见的去噪方式就是使用滤波器，这里我们使用带通滤波器，降低环境噪声的干扰。在 MATLAB 通过 `filter`, `designfilter` 函数我们很容易实现我们需要的功能，在后续的 online 算法处理流程中，因为程序运行在 Java 平台上，所以 Android 方面使用 `dsp-collection` 这一个 jar 包即可满足我们需要的功能。

例如当我们的声源频率范围在 2000Hz-2700Hz 时，那么在除噪的过程中，我们就需要通过一个带通滤波器，减益非该频率范围的成分。如图 (4-5) 所示，横纵坐标信息同图 (??) 通过滤波，不在制定频率范围的数据被大大减益，从而降低了对实验结果的影响。

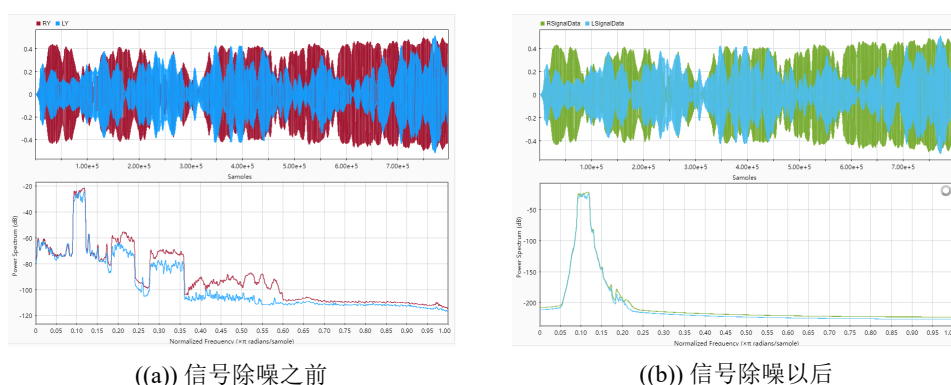


图 4-5: (a) 数据除噪之前。(b) 数据除噪以后。

4.3.2 异常点移除

对实验结果影响较大的还有异常数据，例如在稳定的信号中仍会有一些突变的信号。为防止信号之间的干扰，这里采用分别对两路信号进行处理。值得一提的是，由于采样率高达 44100Hz，所以处理过程中去掉一些数据

并不会造成较大的影响。这里，我们采用滑动窗口的方式，使用加权平均移动 (WMA^[6]) 的方式进行处理。加权平均移动认为：同一个窗口内的不同的数据对目标值都有着不相同的权重因子，因此目标值就可以看成窗口内数据和一个固定的权重函数的卷积。由于数据本身异常点并不是很多，并且使用传统的WMA就可以获得较好的效果，因此记 signal_t 为 t 时刻的数据，采用下面的公式更新 t 时刻的信号值：

$$\begin{aligned} \text{signal}_{t_{\text{new}}} = \text{WMA}_t &= \frac{np_t + (n-1)p_{t-1} + \cdots + 2p_{(t-n+2)} + p_{(t-n+1)}}{n + (n-1) + \cdots + 2 + 1} \\ &= \frac{n \times \text{signal}_{t-1} + (n-1) \times \text{signal}_{t-1} + \cdots + 1 \times \text{signal}_{t-n+1}}{n + (n-1) + \cdots + 2 + 1} \end{aligned} \quad (4-1)$$

经过上述处理后，数据变得更加光滑，并且从图4-6中已经可以隐约可见TDOA的大致范围。这里，LSD为第一个麦克风降噪后是数据，LSignaData代表降噪后数据经过异常点去除操作后的数据。

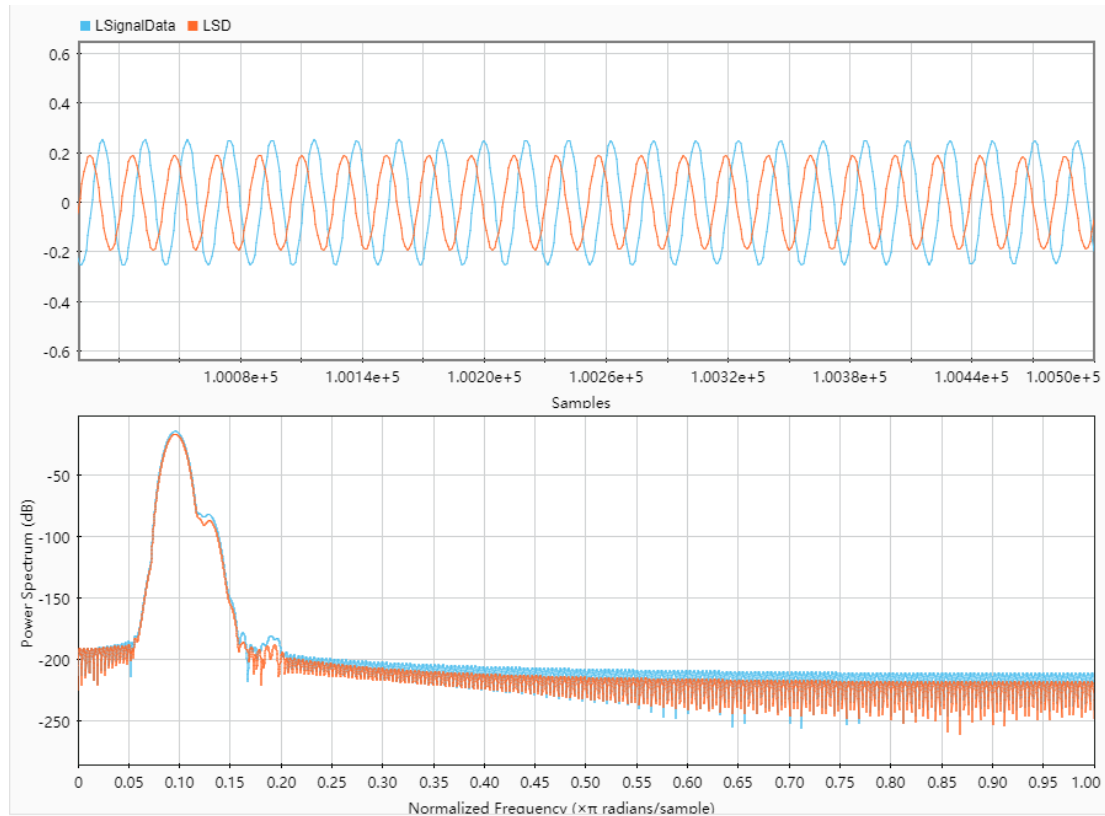


图 4-6: 异常点移除后的数据

4.3.3 有效声音提取

阈值中值判断，online 算法中去除

TODO

4.3.4 TDOA 计算

本小节基于前期的信号处理获取到两路同步的声音信号，调用传统的 GCC 算法对两路同步信号进行 TDOA 计算 (估计)。

由于接收到的声音信号来自于同一个声源，因此两路信号之间具有较强的相关性而与噪声无关。广义互相关方法恰好可以得到两路信号之间的时延差。其基本思想是对两路信号 $x_1(t)$ 和 $x_2(t)$ 进行预滤波，然后再求互相关函数，其原理框架如图 (4-7) 所示，公式表达为:

$$\tau_{\text{delay}} = \arg \max_{t \in \mathbb{R}} ((f \star g)(t)) \quad (4-2)$$

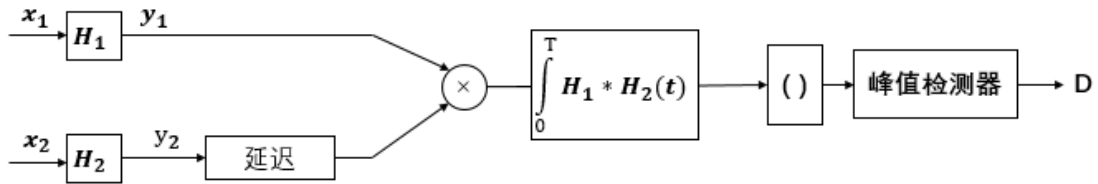


图 4-7: Principle diagram of the generalized cross-correlation based time delay estimation

需要注意的是，这里计算出的“时延”实际是两个信号相差的采样点数，但是，由于采样率为设定好的 44100Hz, 因此计算出相差的采样点数，经过转换就可以得到对应的时延，也就是所求的 TDOA，后面我们用相差的采样点数代指 TDOA，转换公式无需赘述。可以明确的是，当 $\text{TDOA} < 0$ 时，声音先到达第二个麦克风， $\text{TDOA} > 0$ 时，声音先到达第一个麦克风， $\text{TDOA} = 0$ 时，声音同时到达。降至二维空间下，详情如图 (4-8) 所示。

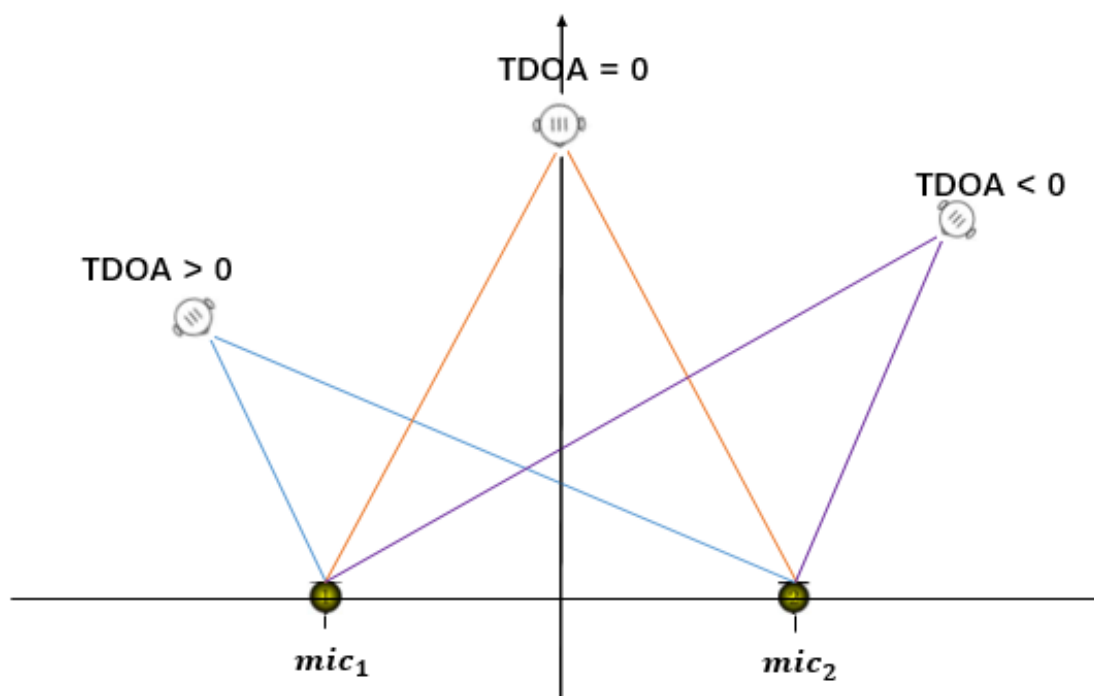


图 4-8: $TDOA > 0$, $TDOA < 0$, $TDOA = 0$

4.3.5 TDOA 处理

利用智能手机作为传感器，计算指定声源到达同一个手机麦克风的 TDOA，可以得到大量的数据并且计算出很多个数据段的 TDOA。因此鉴于数据量多的优势，那么我们就可以利用此优势进行数据的筛选。即便我们的某些 TDOA 测量并不是足够“精确”，但是依然可以辅助进行方向探测。

回到最初的系统设置，我们设置的采样率 44100Hz，编码为 16bit，我们暂时不考虑高编码编码的精确度，但显然的是，由于使用的 bit 数较大，因此数据量化的范围也更大，因此数据也是更精确的。对于 44100Hz 的采样率，麦克风每手机一秒钟，就可以采样到 44100Hz 的采样点。实验中发现，如果将太多的数据一次性进行计算，由于我们没有加窗等操作，加上声源功率较小，信号衰减等，大部分测量结果都是不准确的。然而，另一方面，即便数据过多造成了结果不够准确，但是对数据分段，对该一秒钟的部分数据进行 tdoa 计算，在原理上也是整段数据的 tdoa 的估计，因为延迟是“全局”的延迟，通过实验我们也发现这一个问题。因此，鉴于这种思想，我们可以将信号进行分段，当然，由于在后续实验中设备会一直处理数据，因此我们并不是对 1s 数据分段，而是每 k 个连续数据进行分段。从而，对两个麦克风相同次序的数据段进行 TDOA，依然可以得到全局的 TDOA，如图 (4-9)。

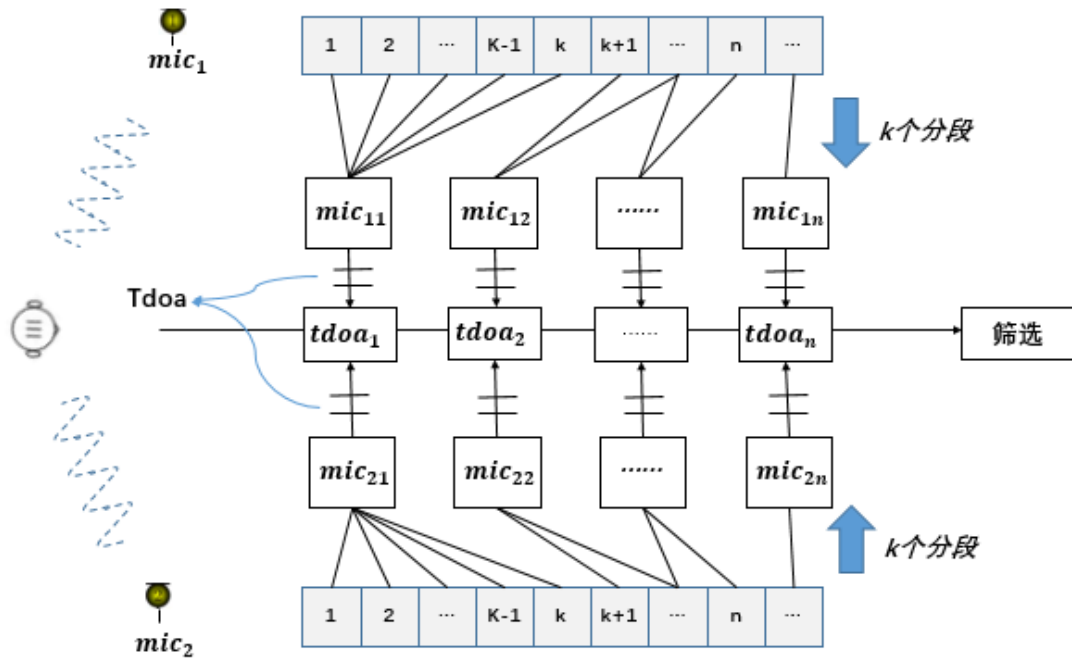


图 4-9: 数据分段示意

数据分段之后，并进行对应的数据段的 tdoa 计算，从而可以产生一系列的 tdoa 值。在实际计算中，我们设定 $k = 2048$ ，很明显的是，

4.4 offline 定位方法

4.5 offline 到 online 的迁移

第五章 模型系统实现与实验分析

5.1 开发环境与平台

5.2 系统设计

5.2.1 多模块设计

5.2.2 界面与运行方式

5.3 实验与分析

5.4 本章小结

第六章 总结与展望

6.1 工作总结

6.2 前景展望

参考文献

- [1] HONGZI Z, YUXIAO Z, ZIFAN L, et al. HyperEar: Indoor Remote Object Finding with a Single Phone[J]. IEEE 39th International Conference on Distributed Computing SysTems(ICDCS), 2019.
- [2] Wikipedia contributors. Global Positioning System[EB/OL]. Wikipedia, The Free Encyclopedia, 2020 (2020/4/2) [2020/5/20].
https://en.wikipedia.org/wiki/Global_Positioning_System.
- [3] KE S, TING Z, WEI W, et al. VSkin: Sensing Touch Gestures on Surfaces of Mobile Devices Using Acoustic Signals[J]. ACM MobiCom, 2018.
- [4] AARABI P. The Fusion of Distributed Microphone Arrays for Sound Localization[J]. EURASIP Journal on Applied Signal Processing, 2003, 2003(4): 338–447.
- [5] Wikipedia contributors. Normalized frequency[EB/OL]. Wikipedia, The Free Encyclopedia, 2020 (2017/7/16) [2020/5/20].
[https://en.wikipedia.org/wiki/Normalized_frequency_\(unit\)](https://en.wikipedia.org/wiki/Normalized_frequency_(unit)).
- [6] Wikipedia contributors. Moving averagery[EB/OL]. Wikipedia, The Free Encyclopedia, 2020 (2020/2/19) [2020/5/20].
https://en.wikipedia.org/wiki/Moving_averag.

致 谢

非常感谢在 Dislab 实验室度过的两年时光，让我在实验室里拓宽眼界的同时，也不断纠正了自己的学习方法和学习态度。我的毕业论文撰写和系统实现中，也或多或少的遇到许多困难，也走过很多弯路。我很感谢我的导师谢磊副教授，从我毕设选题阶段，到我研究工作的开展，再到最后的论文撰写中，谢老师都给予了我细心的指导和点拨，给我纠正了很多错误的方向，避免我走了很多弯路。在实验室的两年，谢磊老师也给我的研究工作不断指引方向，在此，我向谢磊老师致以最真挚的谢意。

同样，我也要感谢在 Dislab 实验室的同一课题组的各位师兄师姐，鲁欣然师兄在我的成长道路上指引了很多，从细节方面给予我帮助，王楚豫师兄同样，给了我研究工作上的很多帮助，让我的研究工作变的富有创新，也是他们给我提出的宝贵建议，让我可以不断改进自己。

感谢大学四年来四种相伴的同学和朋友，因为你们的存在，我的大学才会在忙碌中透露出精彩。感谢每一位悉心教导我的老师，感谢钱柱中教授的指引，让我能够去香港游学增长知识和阅历，让我在问题求解的课程中感受到计算机和算法的魅力。