

基于USB2.0的MEMS数字麦克风阵列采集系统设计

陈洁, 丁磊

广东工业大学, 广东 广州

Email: 2580497655@qq.com, 287931951@qq.com

收稿日期: 2020年12月26日; 录用日期: 2021年1月20日; 发布日期: 2021年1月28日

摘要

本文设计并实现了一种使用USB2.0接口同步传输多通道音频数据的MEMS数字麦克风阵列, 通过将MEMS数字麦克风的时钟、控制、数据线, 连接到USB2.0转并口的高速传输芯片, 结合上位机端的接收与解码程序, 实现低成本、轻量化的麦克风阵列采集系统, 解决了传统麦克风阵列方案价格昂贵、结构复杂、体积庞大、灵活性低的缺点。并保证了该阵列在采样率、动态范围、信噪比、总谐波失真、阵元一致性、和各阵元通道之间的采样同步性上均表现良好。且本系统最多可以支持14路麦克风数据的同步采集。模块化的设计让用户可以在1到14个阵元的范围, 任意设计麦克风阵元数及阵列形状, 提高了阵列结构灵活性。并且, 同时支持PDM和I2S数字麦克风, 使得用户麦克风型号的选择范围更大。

关键词

麦克风阵列, 多声道同步传输, 多声道同步采样, 声音采集系统

MEMS Digital Microphone Array Acquisition System Based on USB2.0 Design

Jie Chen, Lei Ding

Guangdong University of Technology, Guangzhou Guangdong

Email: 2580497655@qq.com, 287931951@qq.com

Received: Dec. 26th, 2020; accepted: Jan. 20th, 2021; published: Jan. 28th, 2021

Abstract

This paper designs and implements a MEMS digital microphone array that uses a USB2.0 interface to synchronously transmit multi-channel audio data. By connecting the clock, control, and data lines of the MEMS digital microphone to the USB2.0 to parallel port high-speed transmission chip, and combining the receiving and decoding program of the host computer, a low-cost, lightweight

microphone array acquisition system is realized, which solves the disadvantages of traditional microphone array solutions such as expensive, complex structure, bulky, and low flexibility. It also ensures that the array performs well in terms of sampling rate, dynamic range, signal-to-noise ratio, total harmonic distortion, array element consistency, and sampling synchronization between each array element channel. And this system can support up to 14 simultaneous microphone data collection. The modular design allows users to design the number of microphone array elements and the shape of the array arbitrarily in the range of 1 to 14 array elements, which improves the flexibility of the array structure. In addition, PDM and I2S digital microphones are supported at the same time, allowing users to choose a wider range of microphone models.

Keywords

Microphone Array, Multi-Channel Synchronous Transmission, Multi-Channel Synchronous Sampling, Sound Acquisition System

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着近代声学中, 语音识别、语音增强、声源定位[1]、声场成像、噪声抑制等音频方面技术的不断完善与应用, 声音处理技术广泛地出现在了人们的生活当中。比如天猫精灵一类的智能音箱及实现定向增益的智能会议麦克风, 又比如工业生产中的声音与振动分析。而这些技术的实现, 多数依赖于麦克风阵列系统。同时, 因为应用方向的不同, 这些技术对于麦克风阵列系统的采样率、动态范围、信噪比、总谐波失真、阵元一致性、采样传输实时性, 和通道间数据采样的同步性, 麦克风的拓展兼容性以及设备体积, 都有要求。

自上世纪八十年代以来, MEMS (微机电系统)技术发展迅速, 使得微机械结构与微电路能够在硅片上集成, 从而出现了使用沉积和选择性蚀刻工艺, 将振动膜片和静态背板制作在晶圆基板上, 再与配套的集成电路 IC 共同封装的 MEMS 数字麦克风。与传统麦克风相比, MEMS 数字麦克风体积小, 外围电路简单, 且性能稳定, 敏感性不受温度、湿度, 时间和焊接加工工艺的影响。

而 USB2.0 接口, 也广泛的存在于各类硬件设备之上, 且 USB2.0 电缆结构简单, 并能够提供 480 Mbps 的传输带宽。并且, 随着微处理器技术的发展, 出现了集成各种外设模块的微处理器芯片, 其中不乏集成 USB2.0 高速接口的微处理器芯片。这些微处理器芯片中, 多数都已有成熟的应用案例。而 USB3.0 普及程度较 USB2.0 差, 且 USB3.0 电缆相对复杂昂贵, USB1.0 则是无法提供设计所需要的带宽, 且 USB1.0 接口几乎已经被淘汰。因此, 本文的麦克风阵列采集系统选择 USB2.0 作为数据传输接口。

本文将从硬件和软件的层面, 讲述如何利用微处理器与 MEMS 数字麦克风, 在低成本和低复杂度且保证阵列性能参数的设计思路下, 实现一种符合上述要求的麦克风阵列采集系统。

2. 麦克风阵列采集系统的硬件设计

麦克风阵列采集系统的硬件总体框架是通过一颗集成了 FIFO 并口转 USB 2.0 控制器的微处理器 (CY7C68013A), 以实现 MEMS 数字麦克风接口的数据与上位机之间的传输, 从而保证了各麦克风通道之间采样和传输的同步性。并使用一颗集成 USB 2.0 全速接口和 I2S 接口的 Cortex-M3 单片机 (STM32F103)接收上位机的指令参数, 产生麦克风时钟和 WS 声道选择信号, 同时兼容 PDM (脉冲密度

调制)和 I2S 两种接口的 MEMS 数字麦克风, 可以支持 1 到 14 个 PDM 数字麦克风, 或 1 到 12 个 I2S 数字麦克风。整体硬件设计框图如图 1 所示:

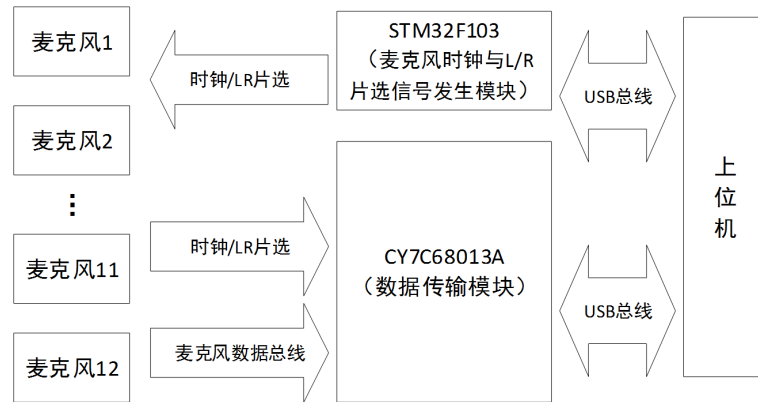


Figure 1. Block diagram of overall hardware design
图 1. 整体硬件设计框图

2.1. 数据传输模块

由于本麦克风阵列采集系统, 同时采样 14 个 PDM 数字麦克风时, 各通道为 3 MHz 同步采样率, 所以本麦克风采集系统总带宽最高为 42 Mbps。因此, 数据传输模块选择 Cypress 的 CY7C68013A 微控制器芯片。CY7C68013A 提供了一个能够工作在 8 位或者 16 位, 且传输速度最高 96 Mbyte/s, 带 4K FIFO 的并行接口。同时提供了一个智能 USB 控制器和 USB2.0 高速收发器, 将并行接口接收到的数据传输到 USB2.0 接口, 并集成了一个增强型 8051 内核用来运行程序[1], CY7C68013A 内部架构如图 2 所示[2]:

Logic Block Diagram

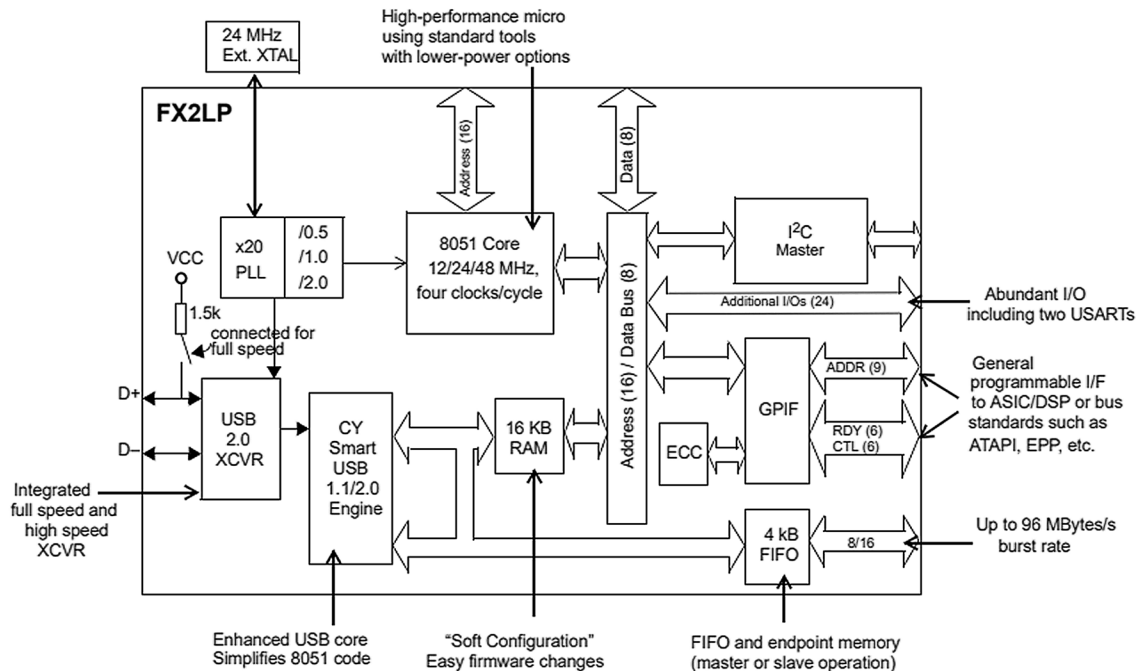


Figure 2. CY7C68013A internal architecture diagram
图 2. CY7C68013A 架构图

PDM 数字麦克风的数据线左右 2 个通道为一组, 连接到 CY7C68013A 的 FIFO 接口 D1:D7, 然后将全部的麦克风的时钟线与 CY7C68013A 的 FIFO 接口的 D0 连接。然后, 通道 1~7 的麦克风左右片选线全部接 VCC, 通道 8~14 的麦克风左右片选线全部接 GND, 硬件设计如图 3 所示:

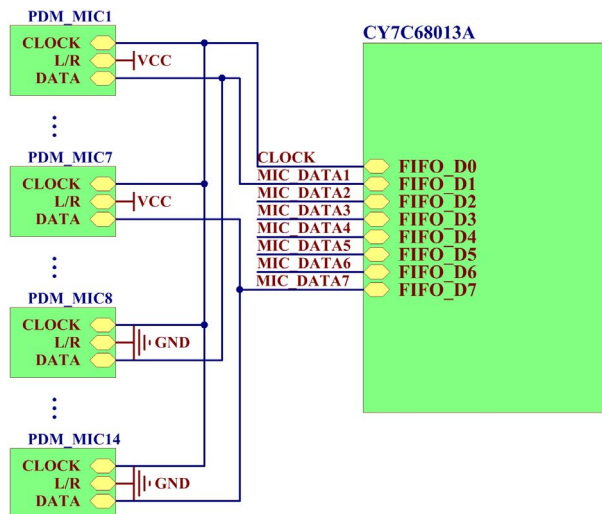


Figure 3. PDM microphone and CY7C68013A parallel port connection diagram

图 3. PDM 麦克风与 CY7C68013A 并口连线图

I2S 数字麦克风的数据线也是左右 2 个通道为一组, 连接到 CY7C68013A 的 FIFO 接口 D2:D7, 然后将全部麦克风的时钟线与 CY7C68013A 的 FIFO 接口的 D0 连接, 再将全部麦克风的 WS 声道选择信号线与 CY7C68013A 的 FIFO 接口的 D1 连接。然后, 通道 1~7 的麦克风左右片选全部接 VCC, 通道 8~14 的麦克风左右片选全部接 GND。硬件设计如图 4 所示:

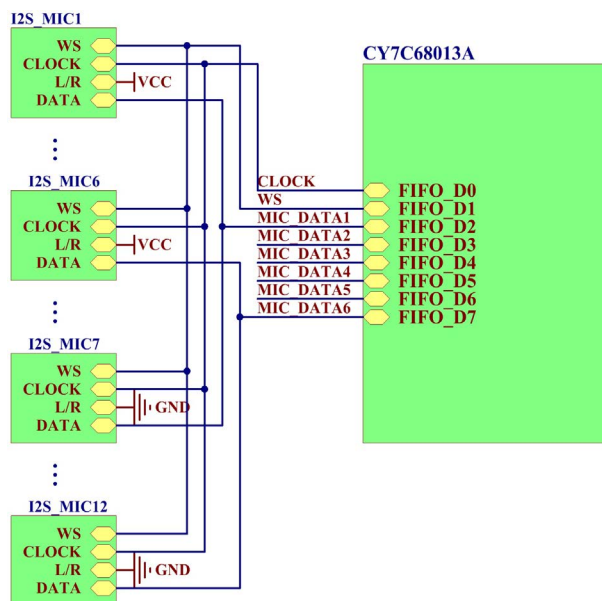


Figure 4. I2S microphone and CY7C68013A parallel port connection diagram

图 4. I2S 麦克风与 CY7C68013A 并口连线图

2.2. 麦克风时钟与片选信号发生模块

单片机型号选择 STM32F103, 其最高主频为 72 MHz, 支持外部晶振, 并提供 I2S 接口, 其 I2S 接口可工作在 8 k~192 kHz 的采样率下, 最高时钟可达到 12 MHz。同时片内集成 USB 全速控制器和 USB PHY, 可以通过 USB 接收来自上位机的关于麦克风种类以及麦克风时序逻辑参数的指令, 给麦克风提供时钟和 WS 声道选择信号。

2.3. 麦克风模块

为了提高对麦克风的兼容性, 麦克风模块设计成同时支持 PDM 和 I2S 两种数字麦克风。

PDM 数字麦克风的接口由时钟线与数据线以及麦克风声道选择线组成, 其接口图如图 5 所示。其中 PDM 数字麦克风的时钟信号, 由外部的 STM32F103 单片机提供, 麦克风 Select 线接高电平时, 麦克风工作在右声道模式, 麦克风 Select 线接低电平时, 麦克风工作在左声道模式。左声道与右声道麦克风, 可以共用一条数据线。在时钟信号上升沿时, 右麦克风数据线输出数据, 左麦克风进入高阻状态, 当时钟信号下降沿时, 左麦克风数据线输出数据, 右麦克风数据线进入高阻状态。PDM 数字麦克风接口时序逻辑图如图 6 所示。

PDM 麦克风采样原理, 是将声波转换为电压, 然后经过 Σ - Δ 调制得到数字信号, 如果把这一连串数字信号放在同一标尺上和声波信号相比, 会发现数字“0”和“1”随着声波的增减程度而产生密度的相应变化, 所以称为脉冲密度调制[3], 如图 7 所示[4]:

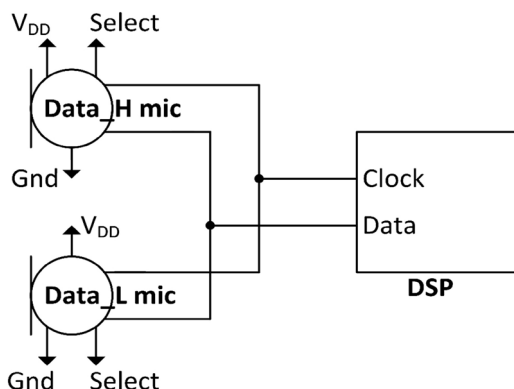
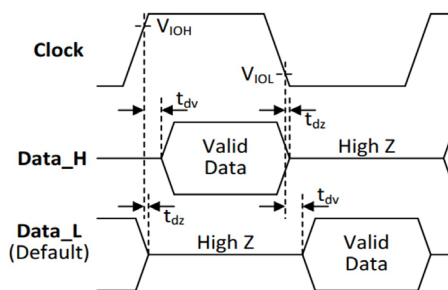


Figure 5. PDM digital microphone interface diagram

图 5. PDM 数字麦克风接口图

Timing Diagram



Typical Performance Characteristics

Figure 6. PDM digital microphone interface timing logic diagram

图 6. PDM 数字麦克风接口时序逻辑图

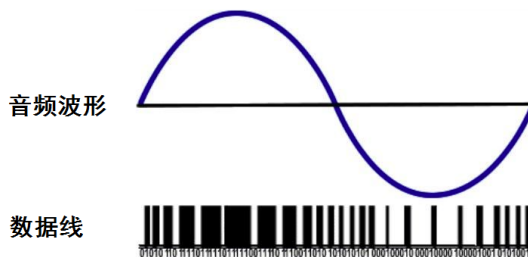


Figure 7. Pulse width coded modulation diagram
图 7. 脉宽编码调制图

I2S 数字麦克风的接口由时钟线、WS 声道选择信号线、数据线、麦克风声道选择线组成, 其接口图如图 8 所示。其中 I2S 数字麦克风的时钟信号, 由外部的 STM32F103 单片机提供, 麦克风 Select 线接高平时, 麦克风工作在右声道模式, 麦克风 Select 线接低平时, 麦克风工作在左声道模式。左声道与右声道麦克风, 可以共用一条数据线。WS 声道选择信号线为高电平时, 右麦克风数据线输出数据, 左麦克风进入高阻状态, WS 声道选择信号线为低电平时[3], 左麦克风数据线输出数据, 右麦克风数据线进入高阻状态。I2S 数字麦克风接口时序逻辑图如图 9 所示。其中每次采样对应的是一个 24 Bit 精度的数据对应为波形的幅值, 即脉冲编码调制, 如图 10 所示[5]。

PDM 数字麦克风选择 ZTS6032, ZTS6032 的灵敏度差异仅为±0.75 dB, 信噪比可达 61 dB, 电源抑制比可以达到-80 dB, 最高时钟频率为 3.25 MHz [4]。

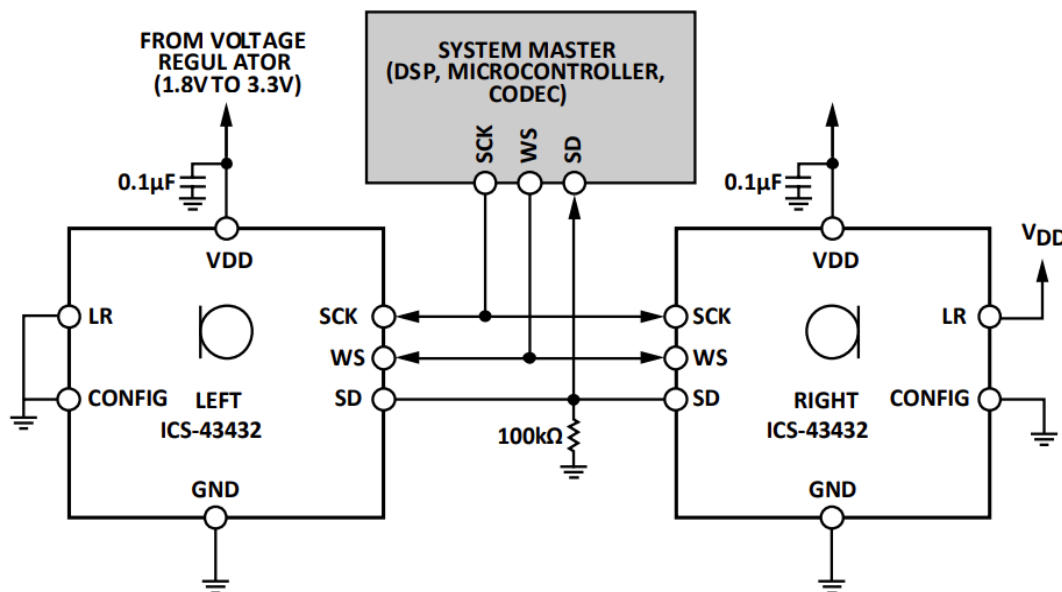


Figure 8. I2S digital microphone interface diagram
图 8. I2S 数字麦克风接口图

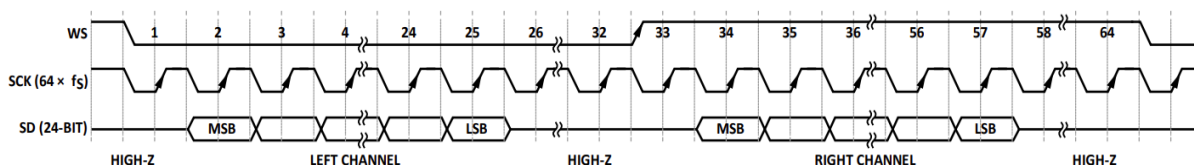


Figure 9. I2S digital microphone interface timing logic diagram
图 9. I2S 数字麦克风接口时序逻辑图

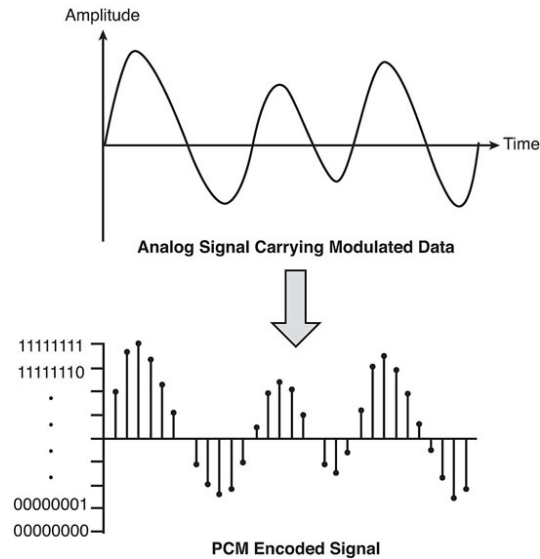


Figure 10. Pulse code modulation
图 10. 脉冲编码调制

I2S 数字麦克风模块选择 MSM261S40H0R, MSM261S40H0R 的灵敏度差异仅为 ± 1 dB, 其信噪比可达到 61 dB, 电源抑制比可以达到 -72 dB, 最高时钟频率为 4 MHz, 采样精度为 24 Bit [5]。

由于所选两款型号的数字麦克风, 在同型号麦克风个体之间灵敏度和频响参数差异极小, 同时两款型号的麦克风均在电源抑制比以及信噪比参数上都比较优秀, 所以, 能够保证麦克风阵列阵元间的一致性, 并降低耦合到声音信号中的电源噪声。

2.4. 电源模块

本设计使用上位机的 USB 接口供电, 所以其主电源为 5 V, 而 STM32 以及 CY7C68013A 以及 MSM261S40H0R 和 ZTS6032 的工作电压均为 3.3 V, 所以, 必须增加一级稳压电路。

而本设计不同的硬件模块, 对电源噪声的要求是有区别的, 其中 CY7C68013A 和 STM32 对电源噪声要求不高, 而 MSM261S40H0R 和 ZTS6032 的电源噪声过大会导致其在声音数据采样过程中耦合进电源的噪声。所以, 麦克风模块的稳压芯片选择低噪声的 LP3993-33B3F。

电源部分的原理图设计如图 11 所示:

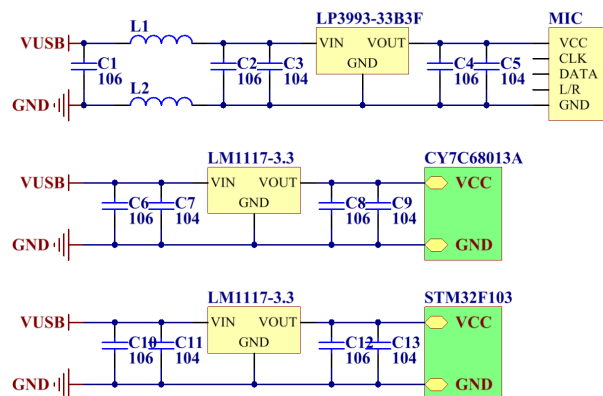


Figure 11. Schematic of the power supply section
图 11. 电源部分原理图

3. 麦克风阵列采集系统的软件设计

本设计的软件分为两大部分, 即下位机部分和上位机部分[4]。

下位机部分由 CY7C68013A 微控制器固件, 以及产生麦克风时钟信号和 WS 声道选择信号的 STM32F103 单片机固件组成。

上位机由负责接收 CY7C68013A 的 USB 接口数据的开源逻辑分析仪软件 PluseView, 以及对麦克风接口时序逻辑数据进行解码的解码软件组成, 以及串口助手组成。

其中 PluseView [6]是一款开源的逻辑分析仪软件, 由开源组织 sigrok 开发和维护, 旨在创建一个可移植的, 跨平台, 免费/自由/开源的信号分析软件套件, 并支持各种硬件设备类型。其基于 GNU GPL, version 3 授权许可。

程序部分结构及流程图如图 12 所示。首先, 上位机中的 STM32-USB 驱动程序检测到 STM32 的 USB-VCOM 设备。然后, 打开上位机的串口助手, 按照程序设计时的指令格式, 输入数字麦克风型号和采样率的参数, STM32 的固件程序会根据 USB 接口收到的数字麦克风型号以及采样率, 自动生成 I2S 参数, 并初始化 I2S 接口的时钟以及是否使能片选信号。然后, PluseView 判断是否存在 CY7C68013A 硬件设备, 如果存在, 将用于采集传输麦克风接口的固件下发到 CY7C68013A 中, 并下发采集指令给 CY7C68013A, CY7C68013A 按照所设置的采样率开始采集麦克风接口的时序信号, 并通过 USB 传输给 PluseView, PluseView 将所接收到的数据打包成二进制文件后, 交给解码程序, 按照用户输入的数字麦克风型号和采样率的参数, 解析成 WindowsPCM [7]编码的 wav 音频文件。

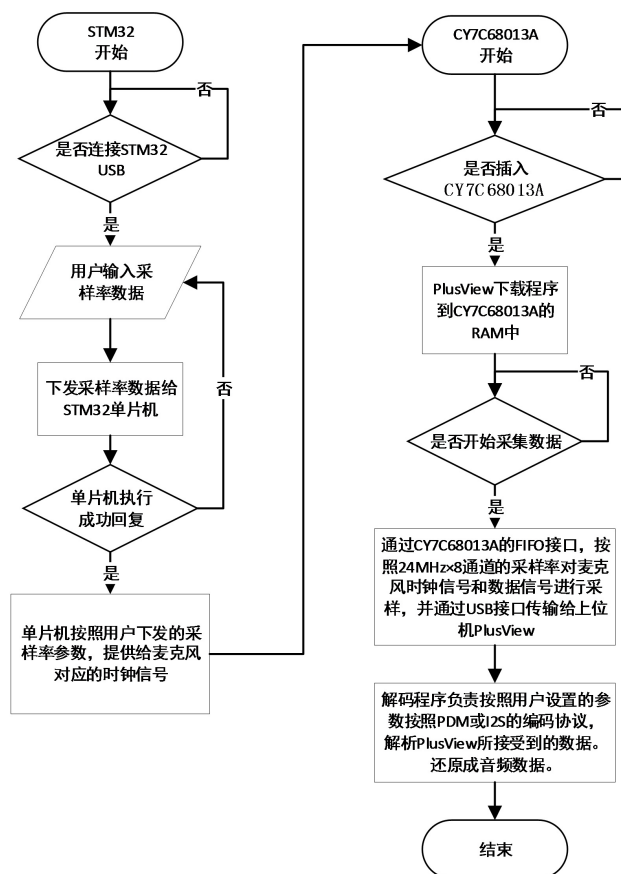


Figure 12. Part of the program structure and flowchart

图 12. 程序部分结构及流程图

3.1. 下位机软件

传输模块(CY7C68013A)的下位机固件, 由 PulseView 通过 USB 下载到 CY7C68013A 的 RAM 中并执行, 使 CY7C68013A 工作在 $24\text{ MHz} \times 8$ 位总线的工作模式, 采集并口总线上的麦克风的时钟线、WS 声道选择线和数据线的时序数据, 并传输到上位机中。因为 24 MHz 的采样速度远大于麦克风模块最高 4 MHz 的时钟速度, 所以此方案对于麦克风接口数据的采集传输是可靠的。

麦克风时钟信号和 WS 声道选择信号模块(STM32F103)的下位机固件, 负责通过 USB 虚拟串口的方式, 接收来自上位机的指令, 然后按照上位机下发的数字麦克风型号、采样率的参数初始化 I2S 外设控制器的输出时钟, 及是否使能 WS 声道选择信号, 以适配不同采样率的 PDM 或者 I2S 麦克风。

3.2. 上位机软件

首先, 通过上位机串口助手程序, 向麦克风时钟与片选信号发生模块(STM32F103)发送麦克风型号和采样率参数。

其次, PulseView 负责向 CY7C68013A 发送采样参数指令, 并接收来自 CY7C68013A 的 USB 数据包, 并将 USB 数据包中采样到的麦克风接口的时序数据, 按照 Raw binary 格式保存为二进制文件。

最后, 解码程序加载 PulseView 所保存的 Raw binary 二进制文件, 先将其解码为各路麦克风对应的时序逻辑信号, 然后再根据 I2S 或者 PDM 的时序逻辑, 解析出各路麦克风所采集到的数据, 如果是 PDM 编码, 会再进行抽值滤波处理, 然后将所得到的音频波形数据, 按照 WindowsPCM 编码, 保存为各路麦克风所对应的 wav 音频文件。

4. 数据采集与性能及可靠性分析

数据采集分析

首先, 在声学暗箱中使用 3×4 平面方阵采集 1 kHz 的标准音频信号, 其中麦克风之间间距小于 1.5 CM 。为了保证在近距离情况下, 阵列采集到的声音数据为平面波, 声源采用面盆大小为 10 寸的全频扬声器, 扬声器的振动面盆大于麦克风阵列的面积, 阵列中心与扬声器的面盆中心在同一直线, 阵列面与扬声器的面盆保持平行。以保证在测试过程中, 声波物理传播过程的同步性和一致性。

然后使用本麦克风阵列采集系统, 按照 48 K 的采样率, 24 Bit 的采样精度, 采集 12 路麦克风的音频数据, 并保存成 12 个麦克风通道对应的 wav 文件, 将 wav 文件导入到 Audition 中分析。

麦克风阵列所采集到的音频信号与扬声器所播放的音频信号的波形 - 时域图[5], 如图 13 所示。

以麦克风通道 1 作为参考通道, 使用广义互相关算法, 计算麦克风通道 1 与其他 11 路麦克风通道之间的时延差, 得到表 1 所示的时延差数据。

其中, 图 13 中, “Out_1kHz” 通道是扬声器播放的音频信号的波形 - 时域图, 该音频信号为 1 kHz 的正弦波, 在时域上, 存在振幅变化。“Channel 1” 到 “Channel 12” 分别为麦克风阵列的阵元 1 至阵元 12 所采集的声音的波形 - 时域图。观察所有麦克风阵元的波形变化, 可知在同一时刻中, 各阵元采集到的信号的振幅基本一致, 且各阵元间, 信号振幅变化规律与扬声器播放的音频信号的规律一致, 由此可证, 各阵元通道间的灵敏度及信噪比参数的一致性良好。又通过观察各阵元间, 所采集信号的振幅变化的起止时间基本一致, 以及通过广义互相关, 得到各通道间最大时延差为 20.833 微秒, 约为 1 个采样点的误差, 由此可证, 各阵元通道间的数据同步性良好。

而对于麦克风阵列对音频数据的还原性, 可通过对比阵元 1 与扬声器所播放的音频信号在同一段时域范围的波形 - 时域图, 与频谱 - 时域图考证, 其中图 14 为波形 - 时域图, 图 15 为频谱 - 时域图。

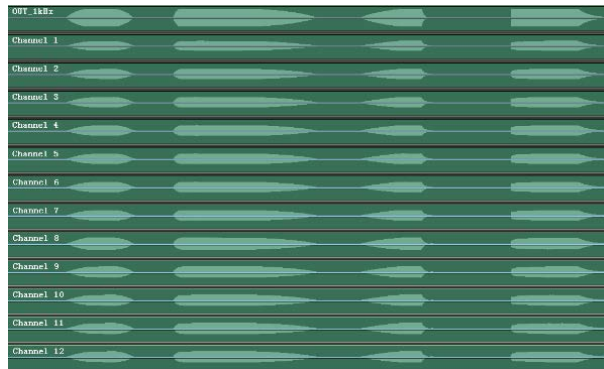


Figure 13. The waveform-time domain diagram of the audio signal collected by the microphone array and the audio signal played by the speaker

图 13. 麦克风阵列所采集到的音频信号与扬声器所播放的音频信号的波形 - 时域图

Table 1. Time delay difference data between microphone channel 1 and other 11 microphone channels

表 1. 麦克风通道 1 与其他 11 路麦克风通道之间的时延差数据

| | 时延差 |
|-------------|-------------|
| 通道 1 与通道 2 | 0 |
| 通道 1 与通道 3 | 0 |
| 通道 1 与通道 4 | 0.000020833 |
| 通道 1 与通道 5 | 0 |
| 通道 1 与通道 6 | 0 |
| 通道 1 与通道 7 | 0 |
| 通道 1 与通道 8 | 0 |
| 通道 1 与通道 9 | 0.000020833 |
| 通道 1 与通道 10 | 0 |
| 通道 1 与通道 11 | 0.000020833 |
| 通道 1 与通道 12 | 0.000020833 |

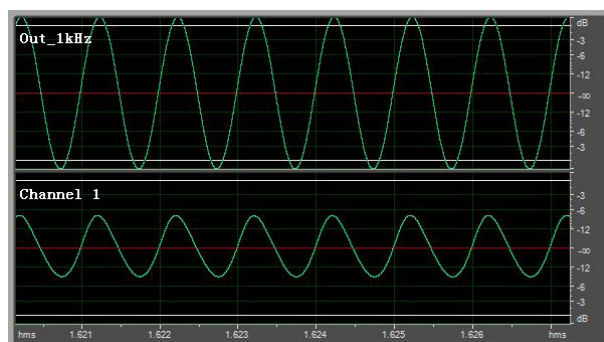


Figure 14. The waveform-time domain diagram of the audio signal played by the speaker and the audio signal received by the microphone 1

图 14. 扬声器所播放的音频信号与阵元 1 接收到的波形 - 时域图

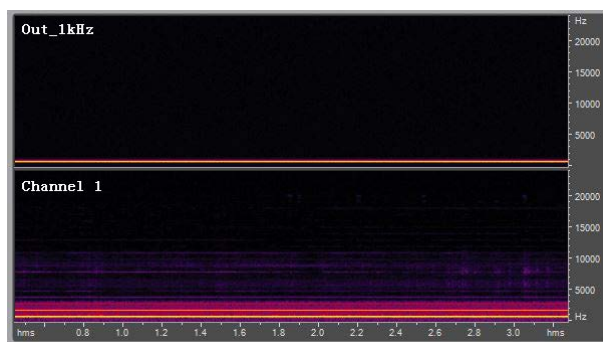


Figure 15. The spectrum-time domain diagram of the audio signal played by the speaker and the audio signal received by the microphone 1

图 15. 扬声器所播放的音频信号与阵元 1 接收到的频谱 - 时域图

由图 14 可以观察到, 阵元 1 所采集的声音数据在与扬声器播放的音频数据, 在同一段域时中, 其波形的频率与变化趋势均一致, 仅有波形振幅衰减。由图 15 可以观察到, 在相同的时域范围内, 阵元 1 所采集的声音数据的频谱 - 时域图, 与扬声器播放的音频数据的频谱 - 时域图, 基本一致。在阵元 1 的频谱 - 时域图中, 存在少量 2 kHz 和 3 kHz 的谐波成分, 而这一现象与扬声器性能及声学暗箱的性能相关。由此可证, 此麦克风阵列采集系统能够较大程度还原音频信号, 保留声音原有的特征。

5. 结论

本文根据现有的麦克风阵列采集设备所存在的价格昂贵, 结构复杂, 体积庞大的不足, 结合 USB 并口转换微控制器, 灵活地将 MEMS 数字麦克风的接口连接到 USB 接口芯片的并口, 从而在采样准确性、传输速度、多通道同步性等方面, 均达到了所期望的设计效果。并且, 该系统综合实现成本极低, 全部硬件成本在 120 元以下, 模块化的设计可以让用户在 1 到 14 个阵元的范围内, 任意选择麦克风阵元数及阵列形状。且该系统可同时支持 PDM 和 I2S 两种数字麦克风, 具有极大的灵活性。本麦克风阵列采集系统所传输的均为麦克风原始数据, 数据的后期处理均在上位机中实现, 保证了音频数据的完整性, 便于后期算法开发及音频数据的处理。

参考文献

- [1] 陶庭兴. 基于 FPGA 和 CY7C68013A 的数据采集系统[J]. 卷宗, 2015, 3(10): 367.
- [2] EZ-USB® FX2LP™ USB Microcontroller High-Speed USB Peripheral Controller. <https://www.cypress.com/file/138911/download>
- [3] 百度百科. DSD 直接比特流数字[Z/OL]. <https://baike.baidu.com/item/DSD/3873923?fr=aladdin>
- [4] Omni-Directional Microphone with High SNR Digital Output (Top Port). http://www.zilltek.com/Product/Product_Show.aspx?ProductSeriesNumber=ZT50
- [5] MSM261S4030H0R I2S Digital Output MEMS Microphone with Multi-Modes. <https://www.memensing.com/webupfile/15736999081895038243.pdf>
- [6] sigrok API documentation. <https://sigrok.org/api>
- [7] 音频设备设计指南[Z/OL]. <https://docs.microsoft.com/zh-cn/windows-hardware/drivers/audio>, 2017-04-20.