

# 基于数字麦克风阵列的声源定位系统研究

段雯<sup>1</sup>, 张永超<sup>1</sup>, 张安莉<sup>1</sup>, 赵录怀<sup>2</sup>, 李浩培<sup>1</sup>

(1. 西安交通大学 城市学院, 西安 710018; 2. 西安交通大学 电气学院, 西安 710048)

**摘要:** 针对当前多麦克风定位阵列需要较大体积、信号采集流程较繁琐、效率较低、实时定位效果差的问题, 设计了一个使用3个数字麦克风组成阵列进行声源定位系统; 该系统利用3个麦克风的时域信息, 采用广义互相关算法对达到时间延迟进行计算, 通过阵元几何关系并利用阵元信号的时延, 获得了在室内环境下较好的声源定位效果; 系统硬件以MAXI-Bit为主控芯片, 选用数字麦克风组成三元线阵采集声音数据, MCU采集麦克风信号并进行数据处理实现声源定位; 最后, 在LCD屏上实时显示声源的方位和距离; 此外, 为了更加直观地反映声源定位的准确度, 还设计了舵机控制模块, 舵机模块上固定激光灯指示声源, 舵机转动角度随声源位置改变而实时改变; 经测试, 该系统最大误差为4.68%, 且舵机上的激光灯可以随声源位置改变而实时指向声源; 该系统可以应用在机械设备故障诊断、鸣笛抓拍等领域, 从而提高生产效率。

**关键词:** 麦克风阵列; 声源定位; 广义互相关; MAXI-Bit

## Research on Sound Source Localization System Based on Digital Microphone Array

DUAN Wen<sup>1</sup>, ZHANG Yongchao<sup>1</sup>, ZHANG Anli<sup>1</sup>, ZHAO Luhuai<sup>2</sup>, LI Haoyu<sup>1</sup>

(1. Xi'an JiaoTong University City College, Xi'an 710018, China;

2. School of Electrical Engineering, Xi'an Jiaotong University, Xi'an 710048, China)

**Abstract:** Aimed at the problems of large volume, cumbersome signal acquisition process, low efficiency and poor real-time positioning effect in multi-microphone positioning array, this paper presents a sound source positioning system by which three digital microphones are used to construct an array. The system utilizes the time domain information of three microphones and the generalized cross-correlation algorithm to calculate the arrival time delay. The geometric relationship of the array element and the delay of the array element signal are used to obtain a better sound source localization effect in the indoor environment. Taken MAXI-Bit as a main control chip, the digital microphone in the system hardware is selected to form a three-way linear array and collect the sound data, the MCU collects the microphone signal and processes the data to realize the location of the sound source. Finally, the direction and distance of the sound source are displayed in real time on the LCD screen. In addition, in order to more directly reflect the accuracy of the sound source positioning, a servo control module is also designed, and the laser light is fixed on the servo module to indicate the sound source, and the steering gear rotation angle changes in real time with the change in the sound source position. Through testing, the system has a maximum error of 4.68%, and the laser light on the servo can point to the sound source in real time, with the position of the sound source changed. The system can be applied in mechanical equipment fault diagnosis, whistle capture and other fields, thereby improving production efficiency.

**Keywords:** microphone array; sound source location; generalized cross correlation; MAIX-Bit

## 0 引言

声源定位是声学的重要问题之一。例如: 利用水下滑翔机估计声源距离<sup>[1]</sup>, 利用接收信号时频谱干涉结构估计声源距离<sup>[2]</sup>, 测量弹药试验火炮的着陆点和空中炸点<sup>[3]</sup>,

在较远距离识别声音<sup>[4]</sup>, 在高噪声环境下对危险情况的探测和分类<sup>[5]</sup>, 高背景噪声情况下探测航空器的噪声源<sup>[6]</sup>, 智能车载电话、会议记录监控系统以及老年人的助听装置<sup>[7-8]</sup>等。

文献[9]中采用基于TODA定位算法的归一化三角定

收稿日期: 2023-09-19; 修回日期: 2023-10-31。

基金项目: 2023陕西省体育局常规课题(2023653); 陕西省科技厅重点研发计划项目(2022GY-089); 机器人与智能制造陕西省高校工程研究中心基金项目(2022GZ04)。

作者简介: 段雯(1984-), 女, 硕士, 工程师。

张永超(1977-), 男, 硕士, 副教授。

张安莉(1974-), 女, 硕士, 教授。

赵录怀(1963-), 男, 博士, 教授。

引用格式: 段雯, 张永超, 张安莉, 等. 基于数字麦克风阵列的声源定位系统研究[J]. 计算机测量与控制, 2024, 32(10): 194-200.

位方法, 由单片机获取 3 个声传感器与 3 个声传感器之间的时间差, 最后通过液晶显示屏显示出声源距离以及坐标。该方法采用传统模拟麦克风, 信号采集流程较繁琐、效率较低、实时定位的效果差。文献 [10] 中利用卡尔曼滤波技术对 RSSI 信号进行滤波, 对滤波后的数据进行三点定位, 获得待定位点的座标。该方法通过降低信号干扰, 实现声源定位, 定位精度差。文献 [11] 设计了一套由 5 个话筒构成的阵列, 主要负责声音信号的收集与放大, 实现声源定位。该方法需要设计复杂的外围信号调理电路。文献 [12] 利用基站之间的单向通讯来实现无线时间同步的方法, 并在此基础上, 设计了一套室内无线定位系统。该方法过于依赖 PC 平台, 便携性差、价格过高。

针对当前多麦克风定位阵列需要较大体积、信号采集流程较繁琐, 效率较低, 实时定位效果差的问题, 设计了一种便携式、成本低、精度较高, 且能够实时定位的数字麦克风阵列的声源定位系统。该方法采用数字麦克风组成三元线阵进行声音数据采集, MCU 对数据处理实现声源定位, 在 LCD 屏上实时显示声源的方位和距离。此外, 为了更加直观地反映声源定位的准确度, 还设计了舵机控制模块, 舵机模块上固定激光灯指向声源, 当声源位置改变, MCU 控制舵机实现激光灯总是指向声源。

### 1 声源定位系统方案和算法设计

声源定位系统结构如图 1 所示。1) 3 个数字麦克风传感器组成阵列, 将声音信号变为数字信号; 2) MCU 读取麦克风传感器的数字信号, 并进行数据处理; 3) LCD 显示屏显示麦克风和声源的角度和距离; 4) 舵机模块被 MCU 控制转向, 实现激光灯总是指向声源; 5) 锂电池为整个系统提供电源, 通过电源管理模块进行降压从而为麦克风阵列模块和人机交互模块提供相符电压。

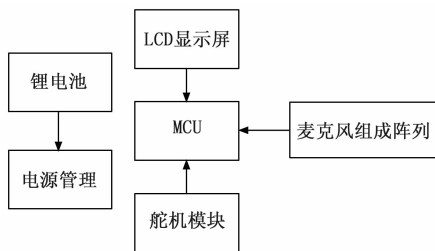


图 1 声源定位系统结构

#### 1.1 到达时间差法远场原理

根据声源的实际环境和位置, 声源可分为远场源和近场源。对麦克风阵列而言, 远场和近场的语音信号传输模型为: 远场模型为平面波, 近场模型为球面波。

一维均匀线性阵列远场和近场模型判断中, 若相邻麦克风之间的距离为  $d$ , 声源信号中最高频率分量为  $f_{\max}$ , 设声速为  $v_s$ , 则声源的最小波长为:

$$\lambda_{\min} = \frac{v_s}{f_{\max}} \quad (1)$$

则可以定义阈值距离为:

$$d_{\text{th}} = \frac{2d^2}{\lambda_{\min}} = \frac{2d^2 \cdot f_{\max}}{v_s} \quad (2)$$

当声源与麦克风阵列中心点的距离满足  $d > d_{\text{th}}$  时, 可判断为远场模型, 反之则为近场。假设某一实际应用场景中, 声速  $v_s = 342 \text{ m/s}$ , 声源最高频率分量为  $f_{\max} = 8 \text{ kHz}$  (对应  $16 \text{ kHz}$  的采样率), 对于单耳助听器与麦克风之间距离为  $d = 1 \text{ cm}$ , 经计算此时有  $d_{\text{th}} = 4.7 \text{ mm}$ , 因为  $d > d_{\text{th}}$ , 则为远场模型; 而对于双耳助听器取  $d = 20 \text{ cm}$ , 经计算  $d_{\text{th}} = 1.9 \text{ m}$ , 则为近场模型。远场信号源传播如图 2 所示。

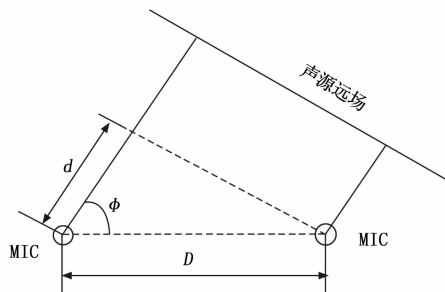


图 2 远场信号源传播示意图

#### 1.2 TDOA 算法

本文采用时延 TDOA 的定位算法, 其原理是根据声源信号到达不同麦克风的时间差, 计算出声源的相对位置。利用到达阵列上各麦克风的语音信号间的时间差来计算声源位置, 该算法计算量较小, 对硬件要求较低, 因此在实际应用中占有很大的比重<sup>[13]</sup>。声源发出声音信号, 传感器 MIC1 和 MIC2 接收到声音信号和背景噪音, 数学建模如下:

设任意两个麦克风接收信号为:

$$x_1(t) = A_1 s(t - \tau_1) + n_1(t) \quad (3)$$

$$x_2(t) = A_2 s(t - \tau_2) + n_2(t) \quad (4)$$

其中:  $s(t)$  为声源信号,  $A_1$  和  $A_2$  为信号衰减系数,  $n_1(t)$  和  $n_2(t)$  为实际环境中的噪声信号,  $\tau_1$  和  $\tau_2$  为声源信号从声源处传播到麦克风所花时间。计算出  $\tau_1$  和  $\tau_2$ , 就可计算出声源和两个传感器间的距离。

计算  $\tau_1$  和  $\tau_2$  算法为广义互相关 (GCC) 算法<sup>[14]</sup>。根据实际情况,  $A_1$ 、 $A_2$ 、 $\tau_1 - \tau_2$  是缓慢变化量, 给出数学模型后, 计算信号  $x_1(t)$ 、 $x_2(t)$  的广义互相关函数:

$$\begin{aligned} R_{x_1, x_2} &= \int_{-\infty}^{+\infty} x_1(t) x_2(t - \tau) dt = \\ &A_1 A_2 \int_{-\infty}^{+\infty} s(t - \tau) s(t - \tau_2 - \tau) dt + \\ &A_1 \int_{-\infty}^{+\infty} s(t - \tau_1) n_2(t - \tau) dt + \\ &A_2 \int_{-\infty}^{+\infty} s(t - \tau_2 - \tau) n_1(t) dt + \\ &\int_{-\infty}^{+\infty} n_1(t) n_2(t - \tau) dt \end{aligned} \quad (5)$$

能量信号  $s(t)$  的自相关函数定义表达式如下:

$$R(\tau) = \int_{-\infty}^{+\infty} s(t)s(t-\tau)dt \quad (6)$$

将式 (6) 代入式 (5), 由于声源信号  $s(t)$  与噪声信号  $n_1(t)$ 、 $n_2(t)$  都互不相关, 因此可认为式 (5) 中第二项和第三项值都为 0, 再假设噪声信号  $n_1(t)$ 、 $n_2(t)$  为互不相关的高斯白噪声<sup>[15]</sup>, 可以得到第四项为 0, 则可得式 (7):

$$R_{x_1x_2}(\tau) = A_1A_2 \int_{-\infty}^{+\infty} s(t)s(t-\tau_1-\tau_2)dt = A_1A_2R(\tau-\tau_{12}) \quad (7)$$

根据互相关函数性质, 当  $\tau_{12} = \tau_1 - \tau_2$  时,  $R_{x_1x_2}(\tau)$  为最大值, 可得到互相关函数波峰时的  $\tau$  值, 即  $\tau_{max}$ , 则可知两个麦克风接收到信号的时间  $\tau_{12}$ 。

## 2 系统硬件设计

### 2.1 麦克风模块选型

模拟信号输出的麦克风声源定位系统必须设计信号调理电路, 该电路使系统精度和稳定性大打折扣。

本系统选用数字麦克风作为声音传感器, 该传感器将声音信号直接处理成数字信号, 输出协议为音频传输协议 (I<sup>2</sup>S, inter-IC sound), 具有较强的抗干扰性<sup>[16-17]</sup>。本系统采用敏芯微推出的型号为 MSM261S4030H0 的数字 MEMS 麦克风, 该型传感器串行数据采用输出数据为 24bit 的 I<sup>2</sup>S 信号格式, 在立体声帧每个数据字有 64 个 SCK 周期, 或 32 个 SCK 循环。当  $L/R=0$  时, 输出数据在左通道中; 而  $L/R=1$  时, 数据在右通道中。输出数据引脚 (SD) 为在输出 LSB 之后三态, 使得另一个麦克风可以驱动公共数据线, 默认数据格式为 I<sup>2</sup>S, MSB 优先。在这种格式中, 每个字的 MSB 延迟从每半帧开始一个 SCK 周期。此外, MSM26S4030H0 是一种全向麦克风, 它具有高性能和高可靠性, 工作电压范围为 4~8.5 V, 工作电流为 750  $\mu$ A, 其原理如图 3 所示。

### 2.2 麦克风阵列结构设计

麦克风阵列的拓扑结构按照维度一般为一维线性、二维平面和三维立体阵列<sup>[18]</sup>。一维线阵的中心位于同一条直线上。根据相邻元素之间的距离是否相同, 可以分为均匀线阵和嵌套线阵。在麦克风阵列中, 均匀线阵是最简单的阵列拓扑, 每个相邻阵列元件之间的距离相同, 麦克风相位和灵敏度等参数相同<sup>[19]</sup>。二维麦克风阵列, 是由多个麦克风按照规则排列在二维平面上的声音接收系统, 其阵元中心分布在一个平面上。根据阵列的几何形状, 可以分为等边三角形阵列、十字形阵列、圆形或矩形区域阵列等<sup>[20]</sup>。

根据本设计的实际情况, 数字麦克风传感器阵列选用一维线性阵列。该结构数据处理简单, 能够测量定位 180° 平面, 3 个数字麦克风传感器构成线性阵列, 用于声源定位。阵列中各个数字麦克风传感器距离太小会产生耦合度增加的现象, 传感器距离和噪声对麦克风传感器的影响成反比, 从而影响声源定位的精确度, 因此, 设计为麦克风

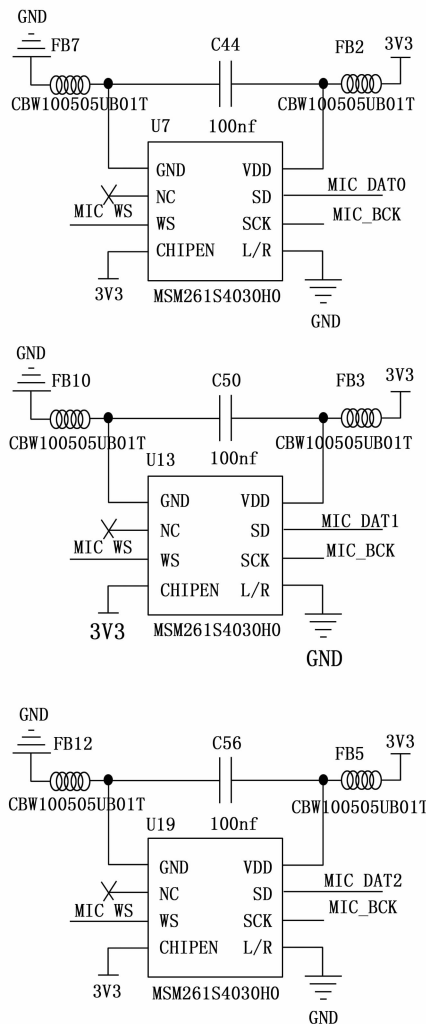


图 3 MSM261S4030H0 原理图

单元的距离小于接收声波的 1/2 波长。人可听到的频率范围为 20~2 000 Hz, 声音在空气中的传播速度为 340 m/s, 所以, 设计的麦克风单元距离间距  $d < 14.17$  cm。其次, 因为数字麦克风大小的因素, 将数字麦克风阵列中麦克风距离设计为 8 cm, 麦克风阵列分布如图 4 所示。

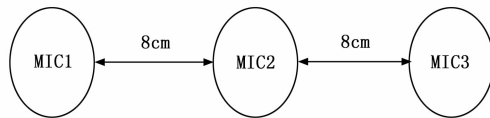


图 4 麦克风阵列分布示意图

### 2.3 MCU 选型

声源定位系统中 MCU 的选型, 主要考虑的影响因素首先是性能, 其次是实时性, 最后是功耗和拓展性, 基于以上综合考虑选择 MAXI-Bit 完成 A/D 转换、数据处理、声源定位、显示以及舵机的控制, MAXI-Bit 核心板结构如图 5 所示。

核心板的资源拓展由其母板承载, 板载了 8 路 12 bit (0

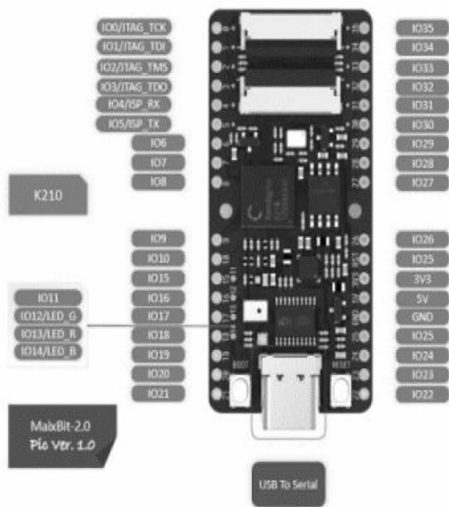


图 5 MAIX Bit 核心板结构

~4 095) ADC 通道, 可以直接连接麦克风阵列模块, 具有电源电压检测电路, 板载了一个屏幕接口, 适用于 LCD 进行人机交互等。同时, 也具有本设计所需的舵机接口, 有对应的电源接口, 3.3 V 和 5 V 可供选择。

### 2.4 人机交互模式

LCD 是一种应用非常广泛的显示屏, 用于实时显示, 可以提高用户的开发效率。本文使用 2.4 寸 LCD 屏幕实时显示相关参数和数据, 需 3.3 V 供电, 由 FPC 连接器直接插入母板对应接口中。

### 2.5 舵机控制

为了验证定位效果, 本设计使用舵机模块对声源信号的方向进行追踪, 舵机模块上固定激光灯指示声源。舵机引出 3 根排线, 其中两根为电源线, 小型舵机的工作电压为 4.8 V 或 6 V。另一条是用于控制操纵器方向的数据线, 给舵机输入一定频率、一定占空比的时间脉冲信号。而转向器的控制通过 50 Hz 的脉冲, 控制其高电平在 0.5 ~ 2.5 ms 之间的时间, 实现转向角的控制。

MCU 通过改变 PWM 占空比来控制舵机的转动角度, 当 PWM 占空比为 0 时, 舵机会保持在初始位置; 当 PWM 占空比为 100% 时, 舵机会转到最大角度; 而当 PWM 占空比为 50% 时, 舵机会转到中间角度位置。因此, 可以通过控制 PWM 占空比来实现对舵机的精细控制。该方法不仅减少了硬件电路的使用, 而且减少了软件的消耗, 提高了舵机的控制效率和精度。此外, 与 MCU 的硬件连接也相对简单, 舵机硬件接口电路如图 6 所示。

### 2.6 系统电源模块设计

本设计使用锂电池和 Type-C 供电, 为各模块提供稳定的直流供电, 确保各模块的正常运行。在设计时, 既要考虑基本的参数, 如电压范围、电流容量等, 还要优化功率转换效率、降噪、抗干扰等特性。除舵机和激光灯电路是由一块 7.4 V、2200 mA·h 的可充电锂电池提供, 其余硬件

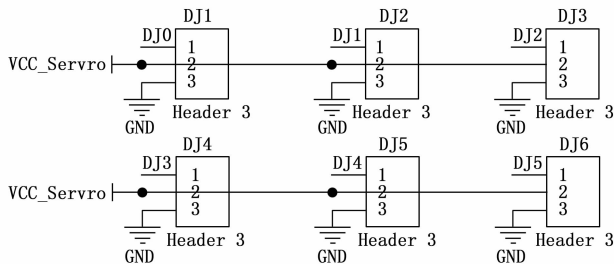


图 6 舵机硬件接口电路图

电路都是由 Type-C 串口提供的。

本设计中采用了多种稳压电路, 将稳定电源提供给各模块。通过 Type-C 串口得到稳定的 5 V 电压, 5 V 为 LCD 屏幕和舵机模块供电, 3.3 V 为 MCU、麦克风阵列模块供电。舵机接口电压通过跳线帽来进行调节, 实现 5 V 或 7.4 V 的电压值变换, 本设计使用的 TBS-K20 型号舵机需要 5 ~ 8.4 V 内供电, 满足设计要求, 舵机稳压电路如图 7 所示。Type-C 得到的 5 V 为 LCD 屏幕供电, 降压所得到的 3.3 V 为麦克风阵列模块。

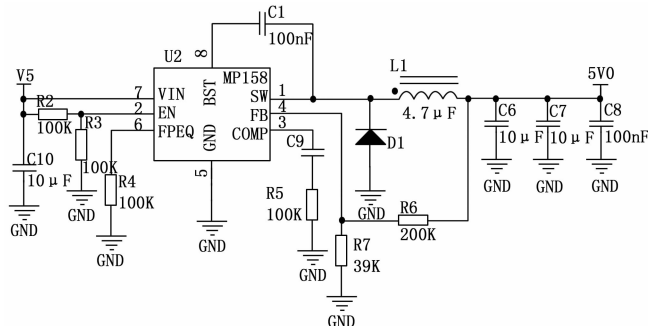


图 7 舵机稳压电路图

## 3 系统软件设计

### 3.1 主程序设计

由于所使用的 K210 为双核单片机, 所以根据实际情况, 将信号处理部分及信号采集置于 CPU0, 舵机控制置于 CPU1, 这样既充分利用了单片机资源, 又将两个 CPU 之间的通信量降至最低水平。

CPU0 中, 开启了一个 100 μs 的定时器中断, 在中断服务函数里对信号进行采集, 由于采用一维线性阵列, 故对三路 ADC 进行了初始化, 中断不断进行 ADC 采集, 同时也使用了双 buff 数据缓冲, 可以得到更好的效果。

CPU1 中, 只需对舵机进行控制, 因此开始是对舵机进行初始化处理, 根据 ATOM 的时钟频率以及最大占空比, 设置舵机的初始化频率为 300 Hz 和舵机直行中值。使用独立按键加减占空比测出舵机转动角度的最大值及最小值。

在主循环中, 根据声源角度给舵机对应的占空比, 让其可以准确追踪声源方向。在主循环中对采集到的数据进行处理, 即互相关运算。得到时间差后便可以根据余弦原理, 算出声源的角度位置信息。而最开始使用的是一般互

相关算法,再使用 FFT 算法对其进行了加速处理,但在其中还要考虑内存的分配,硬件 FFT 和软件 FFT 需配合使用,在进行软件 FFT 运算时可知,需要一个输入数组和一个输出数组,这两个数组的数据类型为 cfloat32 结构体类型的指针,cfloat32 结构体内部为两个 float32 类型的数据,一个用于保存实部,一个用于保存虚部。因此,一个 cfloat32 的结构体即占用 8 个字节,由于信号有 2 048 个点,考虑到互相关运算时的补零操作,实际上进行软件 FFT 运算时,我们需要计算 4 096 个点的 FFT,因此输入信号的数组大小即 4 096 个 cfloat32 类型的结构体。同理可得,输出信号的数组大小也为 4 096 个 cfloat32 类型的结构体,即需要  $4\ 096 \times 8\ \text{byte} = 32 \times 1\ 024\ \text{byte} = 32\ \text{kB}$  的 RAM。在此基础上,修改了互相关的运行流程(等待一次软件 FFT 运算结束之后,再进行下一次数据的变更),通过这一改动,使得数组的利用率极大地提升,最终只需要使用 3 个 32 kB 的数组,共  $3 \times 32\ \text{kB} = 96\ \text{kB}$  的 RAM 即可完成互相关的全部运算。

本文选用 MAXI-Bit 单片机为双核单片机,其中 CPU0 负责信号的采集和处理,为了满足实时更新声源的位置等信息,将声源信号互相关运算做了不同类型的优化,并进行了比较,最终选择了适合本设计的方案。CPU1 负责舵机控制以及 LCD 人机交互程序。控制程序放在了定时器中断中。当 CPU0 中对数据进行处理之后,会得到声源方向信息,此数据存储在—全局变量之中,在 CPU1 中可以根据该变量的值计算出舵机所需的 PWM 占空比,使转动方向刚好指向声源方向。具体流程如下:

- 1) CPU0 进行初始化,首先要关闭 CPU 总中断以及看门狗中断,读取总线频率。
- 2) GPIO 初始化,将独立按键需要用到的 I/O 口配置为输入模式,默认高电平。
- 3) 3 路 ADC 初始化,对 ADC0、ADC1、ADC2 进行初始化,且最高采样频率都设为 10 kHz。
- 4) 定时器中断 CCU6 初始化,配置为 100  $\mu\text{s}$  的定时器中断,即每隔 100  $\mu\text{s}$  进入一次中断服务函数,在中断服务函数中对数据进行采集。
- 5) 开启 CPU 总中断。告知 CPU1,CPU0 初始化完成。
- 6) 进入主循环,在主循环中进行数据处理,计算出声源方向信息。
- 7) 开启 CPU 总中断,关闭 CPU1 看门狗。
- 8) LCD 初始化,即配置模拟 SPI 通信所需的 I/O 口,5 个 GPIO 口配置为推挽输出模式,且默认输出低电平。
- 9) 舵机初始化,即初始化 ATOM 为 PWM 输出功能,舵机工作频率设置为 300 Hz,其占空比为  $(\text{Duty}/\text{ATOM PWM MAX}) \times 100$ 。
- 10) 等待 CPU0 初始化完成。在 CPU1 主循环中计算指向声源方向所需舵机占空比,并做出相应控制。数字麦克风声源定位系统软件流程如图 8 所示。声源定位系统上

电后,首先完成 ADC,定时器及舵机的初始化,其次,完成 ADC 采样指定次数,最后,完成对所有数据 FFT 变换,再进行互相关运算。得到时间差后便可以根据余弦定理原理,算出声源的角度位置信息并控制舵机指向声源方向。

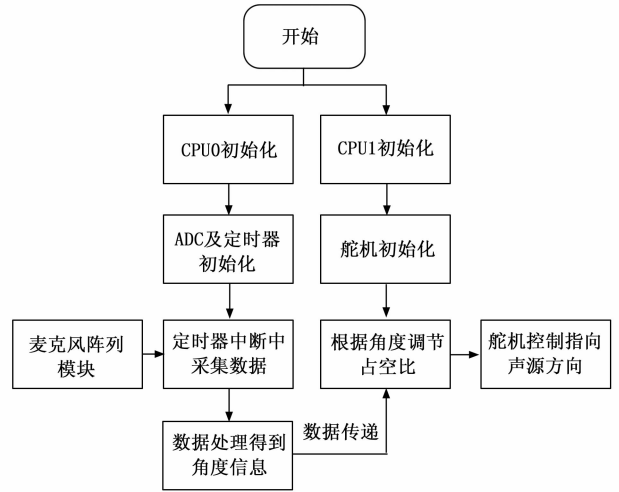


图 8 系统软件流程图

### 3.2 互相关程序设计

麦克风阵列组合模块上的三路麦克风数据被 MCU 采集,为了能达到更好的处理效果,使用了  $2 \times \text{ADC\_DATA\_LEN}$  的数组存放麦克风数据。这样做的优点是,可以边采集边处理(假设采集一帧完成,  $g\_adc0Data[0]$  存放采集完成的待处理数据,处理的时间一般较长,因此将处理时新采集到的数据存放在  $g\_adc0Data[1]$  中)。其程序中变量定义如下:

```

define ADC_DATA_LEN 2048
__attribute__((aligned(256))) volatile sint16 g_adc0Data[2][ADC_DATA_LEN];
__attribute__((aligned(256))) volatile sint16 g_adc1Data[2][ADC_DATA_LEN];
__attribute__((aligned(256))) volatile sint16 g_adc2Data[2][ADC_DATA_LEN];
  
```

数据采集完后,赋值相应的标志位后即可进行数据处理工作。由于采用了一维线性阵,需要对相邻的麦克风之间做互相关运算,通过互相关结果最大值可以得到信号从扬声器到麦克风之间的延时时间,从而获得两个麦克风到声源距离之差,如式(8)所示:

$$\Delta d = \tau_{12} \times v_2 \quad (8)$$

取扬声器信号一个周(前十个点)与麦克风 1 信号做互相关,结果如表 1 所示。

例如:互相关结果第 0 项 =  $(-10) \times 8 + (-10) \times 8 + (-10) \times (-8) + (-10) \times (-8) + (-10) \times (-8) + 10 \times (-8) + 10 \times (-8) + 10 \times 8 + 10 \times 8 + 10 \times 8 = 160$ ;

互相关结果第 8 项 =  $(-10) \times (-8) + (-10) \times (-8) + (-10) \times (-8) + (-10) \times (-8) + (-10) \times (-8) + 10 \times 8 +$

$$10 \times 8 + 10 \times 8 + 10 \times 8 + 10 \times 8 = 800;$$

.....

通过寻找互相关运算结果的最大值, 就可以计算出来信号延时的单位 (一个周期 10 减去互相关结果最大值下标 8 等于 2)。如表 1 中最大为第 8 项, 麦克风 1 的延时时间为:  $2 * \text{采样间隔时间 } 100 \mu\text{s} = 200 \mu\text{s}$ , 麦克风 1 与扬声器之间距离为  $200 \mu\text{s} * 346 \text{ m/s} = 7.2 \text{ cm}$ 。同理, 单片机 ADC 采集一帧为 2 048 点, 采样率为 10 kHz, 对两个麦克风数据进行互相关处理, 则可以得到两个麦克风距声源距离之差。

表 1 两路麦克风互相关结果

扬声器	-10	-10	-10	-10	-10	10	10	10	10	10
麦克风 0	8	8	-8	-8	-8	-8	-8	8	8	8
1	8	8	8	-8	-8	-8	-8	-8	8	8
2	8	8	8	8	-8	-8	-8	-8	-8	8
3	8	8	8	8	8	-8	-8	-8	-8	-8
4	-8	8	8	8	8	8	-8	-8	-8	-8
5	-8	-8	-8	8	8	8	8	-8	-8	-8
6	-8	-8	-8	8	8	8	8	8	-8	-8
7	-8	-8	-8	-8	8	8	8	8	8	-8
8	-8	-8	-8	-8	-8	8	8	8	8	8
9	8	-8	-8	-8	-8	-8	8	8	8	8
互相关结果	160	-160	-480	-800	-480	-160	160	480	800	480

### 3.3 优化后的互相关运算程序设计

由于采集到的数据点数过大, 对其直接做互相关运算计算量太大, 不能做到实时识别声源方向。需要对相关运算加速, 因此使用了 FFT 计算相关函数, 其具体操作步骤如下:

- 1) 将采集到的三路声源信号数据从 2 048 点补零成 4 096 点。
- 2) 对相离最远的信号 0 和信号 2 做 FFT (快速傅里叶变换)。
- 3) 然后再对信号 2 求共轭。
- 4) 将信号 0 的 FFT 结果同信号 2 FFT 结果的共轭做卷积运算。
- 5) 对上述结果进行 FFT 运算。
- 6) 求出 FFT 变换结果的幅值。
- 7) 用逐一比较法求出最大幅值数组所对应的下标。
- 8) 其信号之间的延时 = 最大下标 \* 1/采样频率。
- 9) 声源距两路麦克风距离之差 = 延时 \* 声速。

此外, MAX1-1024 还具有 8 MIB 的 EMEM 内存空间和内置硬件 FFT 的计算资源, 而硬件 FFT 的运算速度相对于软件 FFT 有了非常大的提升。根据在实际环境下的效果, 可以实现实时精确定位声源方向。

## 4 试验验证

### 4.1 信号互相关算法调试

通过 ADC0 和 ADC1 互相关、ADC1 和 ADC2 互相关、ADC0 和 ADC2 互相关。即可获取声源大致方向信息, 例

如: 当 ADC0 和 ADC2 的互相关结果最大值对应数组下标为 0 时, 声源距离麦克风 1 和麦克风 2 的距离相同, 再根据 ADC1 和 ADC0 的互相关结果可以知道声源在左还是右。而相邻位置 ADC 互相关又可将方向再一次细分。最终可以实现 180°范围内声源方向的确定, 但其精度偏低, 高达 45°且响应较慢。为了提高精度、运算速度, 使用了硬件 FFT 进行加速。

### 4.2 试验测试结果

声源的实际角度信息可以通过米尺和量角器来进行测量, 可以对比该算法定位的准确性。为了更好地做比较, 本设计对模拟麦克风和数字麦克风在实际中分别进行测试, 模拟麦克风测量结果如表 2 所示, 数字麦克风测量结果如表 3 所示, 可以得到数字麦克风在实现声源定位中有更好的优越性。本文使用的 TDOA 声源定位算法能够较准确地识别声源方向, 因此可以证明该实验方案的可行性。

表 2 模拟麦克风测量结果 (°)

序号	实际角度	测量角度	绝对误差	相对误差
1	22	29.60	7.60	25.6
2	14	22.45	8.45	37.6
3	11	9.47	-1.53	16.2
4	0	1.55	1.55	0.01
5	8	9.47	1.47	15.5
6	15	18.45	3.45	18.6
7	23	36.60	13.60	37.1

表 3 数字麦克风测量结果

位置	1	2	3	4	5
测试 1/(°)	66.20	86.97	91.68	93.58	105.50
测试 2/(°)	66.17	86.61	80.86	100.98	111.75
测试 3/(°)	66.34	76.88	90.52	98.91	100.76
测试 4/(°)	65.64	78.59	90.78	98.90	103.35
实际角度/(°)	70.00	80.00	90.00	100.00	110.00
平均角度/(°)	66.87	81.81	88.77	98.47	106.27
绝对误差/(°)	3.13	1.81	1.23	1.53	3.73
相对误差/%	4.68	2.21	1.38	1.55	3.51

## 5 结束语

本设计采用了以 MAIX Bit 开发板为基础的声学定位系统, 通过麦克风阵列, 可以对周围的声学信号进行准确的监控, 利用到达时间差理论, 分析了声源远场模型和广义互相关算法, 给出了利用时间差实现声源定位的方法。在出现异常的情况下, 可以迅速地发现问题所在, 并对问题进行迅速有效地解决, 其成果可以帮助工人快速发现并处理车间生产线的机械故障, 提高生产效率。经测试本系统可以有效地提取声源方位和距离。

### 参考文献:

[1] 吴禹沈, 秦继兴, 李整林, 等. 声学滑翔机联合的深海水下声

- 源定位 [J]. 声学学报, 2023, 48 (3): 437 - 446.
- [2] 刘与涵, 郭良浩, 章伟裕, 等. 深海声影区时频谱干涉结构与声源定位 [J]. 应用声学, 2024, 43 (1): 12 - 23.
- [3] 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]. 合肥: 安徽大学, 2014.
- [4] KUMATANI K, MCDONOUGH J, RAJ B. Microphone array processing for distant speech recognition: from close-talking microphones to far-field sensors [J]. IEEE Signal Processing Magazine, 2012, 29 (6): 127 - 140.
- [5] LOPATKA K, KOTUS J, CZYZEWSKI A. Evaluation of sound event detection, classification and localization in the presence of background noise for acoustic surveillance of hazardous situations [C] //International Conference on Multimedia Communications, Services and Security. Springer International Publishing, 2020: 96 - 110.
- [6] GENESCA M, ROMEU J, ARCOS R, et al. Measurement of aircraft noise in a high background noise environment using a microphone array [J]. Transportation Research Part D Transport Environment, 2013, 18 (1): 70 - 77.
- [7] AARABI P, ZAKY S. Robust sound localization using multi-source audiovisual information fusion [J]. Information Fusion, 2001, 2 (3): 209 - 223.
- [8] 郝张红, 段羽浩, 韩彬彬. 基于 TDOA 的声源定位算法研究及实现 [J]. 信息系统工程, 2020 (2): 140 - 141.
- [9] 徐乐, 韦玉科. 基于三点定位与加权坐标的三角定位算法 [J]. 计算机工程与应用, 2020 (9): 111 - 116.
- [10] 李彦璐, 沈一, 潘欣裕, 等. 基于麦克风阵列的声源定位系统研究 [J]. 物联网技术, 2021 (7): 26 - 28.
- [11] 石钦, 赵思浩, 崔晓伟, 等. 无线时间同步的 TDOA 室内定位系统 U 门 [J]. 导航定位与授时, 2018, 5 (3): 5358.
- [12] MORGENSTERN H R B. Far-field criterion for spherical microphone arrays and directional sources [J]. 4th Joint Work-
- shop on Hands-free Speech Communication and Microphone Arrays (HSCMA). IEEE, 2015, 34 (3): 276 - 280.
- [13] YAMADA T, NAKAMURA S, SHIKANO K. Distant-talking speech recognition based on a 3-D Viterbi search using a microphone array [J]. IEEE Transactions on Speech & Audio Processing, 2002, 10 (2): 48 - 56.
- [14] CHEN H, BAO Y, SER W. Effects of sensor position errors on farfield/nearfield wideband beam formers for microphone arrays N [J]. IEEE Sensors Journal, 2015, 15 (9): 4812 - 4825.
- [15] BENESTY J, SOUDEN M, HUANG Y. A perspective on differential microphone arrays in the context of noise reduction [J]. IEEE Transactions on Audio, Speech and Language Processing, 2012, 20 (2): 699 - 704.
- [16] HERBORDT W, BUCHNER H, NAKAMURA S, et al. Multichannel bin-wise robust frequency-domain adaptive filtering N [J]. IEEE Transactions on Audio Speech & Language, 2007, 13 (2): 1034 - 1051.
- [17] ZHANG Y, MENG S J. Sound source localization algorithm based on a helmet-mounted microphone array [C] // Proceedings of International Symposium on Parallel Architectures, Algorithms, and Programming, 2014: 183186.
- [18] YANGW X. Beamforming [J]. IEEE Transactions on Audio, Speech and Language Processing, 2007, 15 (4): 1340 - 1351.
- [19] SCHMIDT R O. Multiple emitter location and signal parameters estimation [J]. IEEE Transactions on Antennas & Propagation, 1986, 34 (3): 276 - 280.
- [20] ROY R, PAULRAJ A, KAILATH T. Estimation of signal parameters via rotationalinvariance techniques-ESPRIT [C] // Advanced Algorithms and Architectures for Signal Processing I. International Society for Optics and Photonics, 1986.
- tion of FPGA implemented CNN accelerators [J]. IEEE Transactions on Device and Materials Reliability, 2023, 23 (1): 116 - 126.
- [17] NGUTEN T T H, TATOWTJ A, COUSTATY M, et al. Survey of post-OCR processing approaches [J]. ACM Computing Surveys (CSUR), 2021, 54 (6): 1 - 37.
- [18] QIU J, WANG J, YAO S, et al. Going deeper with embedded FPGA platform for convolutional neural network [C] // Proceedings of the 2016 ACM/SIGDA International Symposium on Field-programmable Gate Arrays, 2016: 26 - 35.
- [19] SUDA N, CHANDRA V, DASIKA G, et al. Throughput-optimized OpenCL-based FPGA accelerator for large-scale convolutional neural networks [C] // Proceedings of the 2016 ACM/SIGDA International Symposium on Field-Programmable Gate Arrays. 2016: 16 - 25.
- [20] ZHANG C, WU D, SUN J, et al. Energy-efficient CNN implementation on a deeply pipelined FPGA cluster [C] // Proceedings of the 2016 International Symposium on Low Power Electronics and Design, 2016: 326 - 331.