



基于广义互相关时延估计 (TDOA) 算法的声源定位跟踪系统设计与实现

徐雪慧¹, 梅振龙²

(1. 武汉职业技术学院 电子信息工程学院, 湖北 武汉 430074;
2. 国家脉冲强磁场科学中心, 湖北 武汉 430073)

摘 要: 为实现在二维平面内对声源进行实时定位和动态跟踪, 运用时延估计 (TDOA) 互相关算法设计一种原声监听头阵列的 CC-TDOA 声源定位跟踪系统, 实现对多个原声监听头进行同步采样, 再进行信号放大及运算处理, 运用 LCD 屏实时显示目标声源的距离和方位角度, 同时运用二维云台控制激光笔对准声源, 并持续动态跟踪声源。模拟仿真及真实环境实验测试表明, 采用基于到达时间差的互相关定位算法, 计算量小, 精度较高, 测试角度误差小于 2° , 距离误差小于 1.2%, 可以满足多种智能应用场合中声源实时定位与跟踪的要求。

关键词: 声源定位; 互相关; TDOA 定位算法; 跟踪系统设计

中图分类号: TN641

文献标识码: A

文章编号: 1671-931X (2024) 02-0025-10

DOI: 10.19899/j.cnki.42-1669/Z.2024.02.004

一、引言

声源定位系统是一种用于确定声音来源位置的技术, 随着信息技术和人工智能的快速发展, 声源跟踪定位技术广泛应用于智能机器人^[1]、智能视频监控、智能汽车、智能家居、智能国防、智能故障检测等方面^[2]。

常见的声源定位技术有基于最大输出功率的可控波束形成技术、基于高分辨率谱估计技术的声源定位技术、基于到达时间差 (Time Difference of Arrival, TDOA, 也称时延估计) 的定位技术等^[3-5]。

其中, 基于时延估计的声源定位技术的基本思想是先估算出各个麦克风的时延, 再将时延转化为距离差, 最后结合麦克风阵列的空间几何位置确定声源的方位。基于时延估计 TDOA 的算法是应用最为广泛的声源定位算法, 计算量小、硬件易于实现, 具有定位精度较高、硬件成本低、鲁棒性好等优点, 比较适合实时处理, 实际应用非常广泛。

本文基于广义互相关平滑相干加权时延估计 (TDOA) 算法设计和研制了一款嵌入式声源定位跟踪系统 (简称为 gccscot-TDOA 声源定位系统), 系

收稿日期: 2023-10-11

作者简介: 徐雪慧 (1978—), 女, 湖北浠水人, 武汉职业技术学院电子信息工程学院副教授, 研究方向: 职业教育、物联网; 梅振龙 (1999—), 男, 湖北黄梅人, 国家脉冲强磁场科学中心技术支持人员, 研究方向: 电子技术、重频电源设计。

统实现了对真实环境声源的定位与跟踪,在声源定位、语音增强、异音检测、鸣笛抓拍、自动语音识别等领域有重要的应用价值^[6~7]。

二、声源定位理论分析与算法

TDOA 声源定位是一种利用时间差进行定位的方法,主要分为两个步骤:时延估计与定位估计,即先通过测量信号到达接收阵的时间差,再运用接收阵的几何关系确定声源的距离。由于声源距离接收阵有一定的距离且方向非垂直,两个阵元间由于

声程差必然存在到达时间差,可以运用 TDOA 声源定位算法是通过计算两个或多个麦克风接收到声源信号之间的时间差。而在二维平面中,声源位置与两个麦克风之间的相对位置呈现一定的几何关系,通过解算几何关系式,可以确定声源的位置。

(一)TDOA 声源定位算法

直线麦克风阵列中,取 3 个原声监听头的麦克风阵列模型如图 1 所示。假设没有混响且噪声较小;而且声源离麦克风较近,为近场声波,波阵面为球面波^[8]。

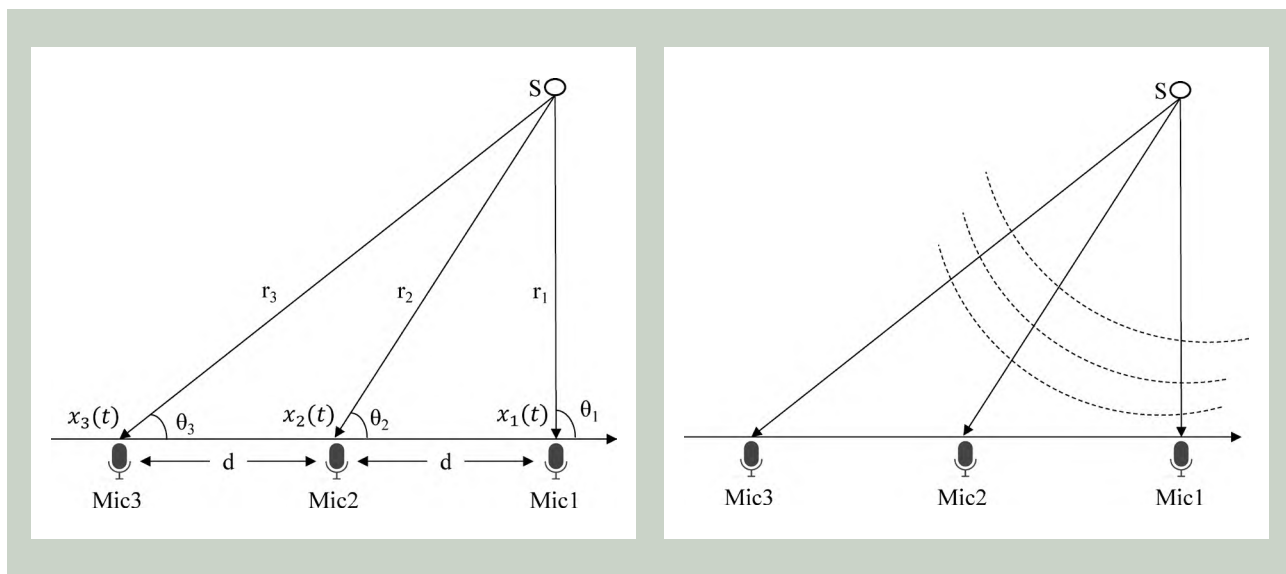


图 1 3 个原声监听头的麦克风阵列近场球面波 / 模型图

其中, s 为声源, $Mic1$ 、 $Mic2$ 和 $Mic3$ 为原声监听头麦克风阵列, r 为声源与监听头的距离, d 为原声监听头麦克风之间的间距(在实验时可以根据测量条件设置为固定值)。阵列中两个原声监听头 $Mic1$ 和 $Mic2$ 阵元的接收信号为:

$$x_1(t)=a_1s(t-\tau_1)+n_1(t) \quad (1)$$

$$x_2(t)=a_2s(t-\tau_2)+n_2(t) \quad (2)$$

其中 $s(t)$ 为声源信号, a 为幅度衰减因子, $n_1(t)$ 和 $n_2(t)$ 为环境噪声(假设为互不相关的高斯白噪声), τ_1 和 τ_2 是信号从声源处传播到两个原声监听头 $Mic1$ 和 $Mic2$ 阵元的传播时间。

互相关函数时延估计 TDOA 算法根据两个原声监听头信号的互相关函数峰值来估计时延值。在声源定位系统中,麦克风阵列的每个阵元接收到的目标信号都来自于同一个声源,因此,各通道信号之间具有较强的相关性。理想情况下,通过计算

每两路信号之间的相关函数,对两路信号作互相关之后再进行峰值检测即可得到 τ_1 和 τ_2 的时间差,就可以确定两个原声监听头观测信号之间的时延。

在基本互相关算法中,互相关函数峰值不够尖锐,即在周围旁瓣或者噪声幅值很大、信噪比下降时时延估计成功率明显下降;而且当采样率不够高时,相关函数峰值可能位于两次采样点之间,从而漏掉峰值,使得时延估计精度较低。

因此,我们采用对噪声的抑制能力较强的能够减少信号波动的广义互相关平滑相干加权 TDOA 算法(gccscot-TDOA),通过对互相关函数的互功率谱进行平滑相干加权。平滑相干(SCOT)加权函数本质上相当于白化滤波,当噪声水平较高时作用明显,增大主峰和次峰的相对值,让峰值更加突出。这种算法,在较高的信噪比的时候,表现出了波动小、峰值尖锐的特性,在降低信噪比时,也表现出了

较强的抗干扰性,抑制了低信噪比情况下噪声和混响对信号的干扰,解决了基本互相关算法的时延估计成功率低和精度上的局限性^[9-10]。

(二)gccscot-TDOA 声源定位算法

广义互相关平滑相干加权时延估计 gccscot-TDOA 算法如图 2 所示。

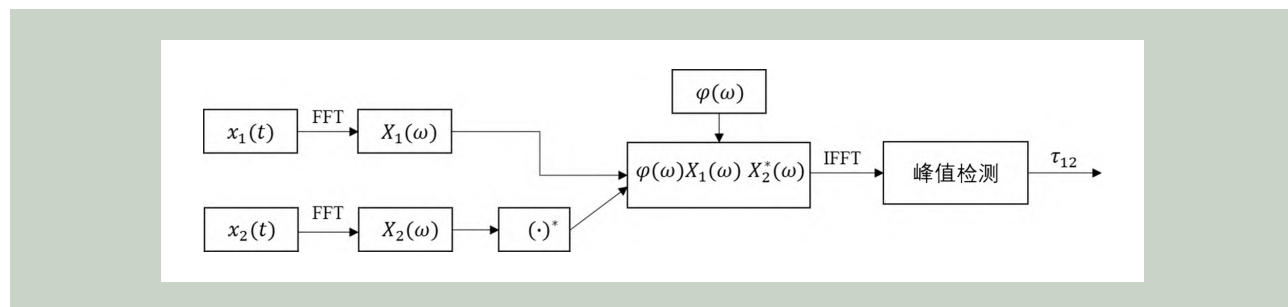


图 2 广义互相关平滑相干加权时延估计算法框图

运用互相关算法做时延估计,再进行傅立叶变换得到互相关函数和互功率谱的关系式,为了锐化互相关函数的峰值,在频域内对互功率谱进行平滑相干加权,可抑制噪声和混响干扰,最后进行傅里叶逆变换,得到互相关函数:

$$R_{x_1 x_2} = \int_0^\pi \varphi_{12}(\omega) X_1(\omega) X_2^*(\omega) e^{-j\omega t} d\omega \quad (3)$$

其中 $\varphi_{12}(\omega)$ 表示频域平滑相干加权函数(SCOT),表达式见式(4)。

$$\varphi_{12}(\omega) = \frac{1}{\sqrt{X_1(\omega) X_2^*(\omega)}} \quad (4)$$

$X_1(\omega)$ 和 $X_2(\omega)$ 分别为 $x(t)$ 、 $y(t)$ 的自功率谱密度函数。当信噪比较大时,平滑相干法对噪声的抑制能力较强,能够减少信号波动对 TDOA 估计的影响。

由相关函数的性质可知,当 $t=\tau_{12}$ 时,两阵元信号相似度最高,相关性最强, $R_{x_1 x_2}$ 取最大值即峰值。因此,找到互相关函数的峰值所对应的时刻,就能计算出两个原声监听头阵元的时延 τ_{12} 。

(三)基于 gccscot-TDOA 的声源定位数学模型

在图 1 麦克风阵列中,不同麦克风阵元之间声音到达的时间差可由下述方程(5)(6)表示。其中 r 是麦克风与声源之间的距离(各参数如图 1 中所示), c 是声音在空气中的传播速度。

$$\tau_{12} = \frac{r_2 - r_1}{c} \quad (5)$$

$$\tau_{13} = \frac{r_3 - r_1}{c} \quad (6)$$

同时,根据麦克风阵列的几何关系,可以得到式(7)。

$$\begin{cases} r_2^2 = r_1^2 + d^2 + 2r_1 d \cos\theta_1 \\ r_3^2 = r_1^2 + 4d^2 + 4r_1 d \cos\theta_1 \end{cases} \quad (7)$$

运用广义互相关时延估计 TDOA 算法可以得到的 τ_{12} 和 τ_{13} ,从而可以求解(5)(6)(7)方程组,进而使用余弦定理可求得声源与麦克风阵列中的原声监听头之间距离及相应夹角,从而实现对声源的位置定位。

$$\begin{cases} r_2 = \frac{c^2(\tau_{12}^2 + \tau_{13}^2) - 2d^2}{2c(\tau_{12} + \tau_{13})} \\ \theta_1 = \arccos \left[\frac{(c^3 \tau_{12} \tau_{13} + d^2 a)(\tau_{13} - \tau_{12})}{d^2(\tau_{12}^2 + \tau_{13}^2) - 2d^3} \right] \end{cases} \quad (8)$$

三、gccscot-TDOA 声源定位系统设计

(一)gccscot-TDOA 系统硬件设计

gccscot-TDOA 声源定位跟踪系统的功能电路由声源和声源定位两大部分组成,系统结构图如图 3 所示。其中,声源部分主要是声源信号模块,主要由 ICL8038 单片函数信号发生器组成,声源参数可控,方便测试;声源定位部分主要由麦克风阵列拾取声源模块、音频信号放大模块、采样 A/D 转换模块、声源定位算法处理模块、结果显示 LCD 模块和声源跟踪模块构成。

在保证系统定位精度的前提下, gccscot-TDOA 声源定位跟踪系统采用原声头麦克风阵列实现对声音的拾取;嵌入式 MCU 对采集到的数据进行 A/D 转换和处理,依据声源定位算法,进而计算出声源的位置信息;位置信息输出至 LCD 屏进行位置参数实时显示,同时通过二维云台控制激光笔实时动态定位和跟踪声源。

系统实现电路是由原声监听头阵列、四个 LM324 组成的音频功放电路、STM32F103VET6 主控器 MCU、LCD-JLX256128G 液晶显示屏、四相五

徐雪慧,梅振龙:基于广义互相关时延估计(TDOA)算法的声源定位跟踪系统设计与实现

线步进电机、二维云台、电源模块组成,主要模块电 路如图 4、5、6 所示。

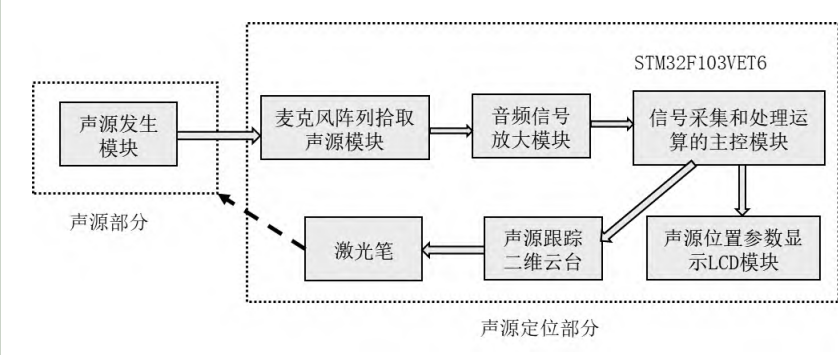


图 3 gccscot-TDOA 声源定位跟踪系统结构框图

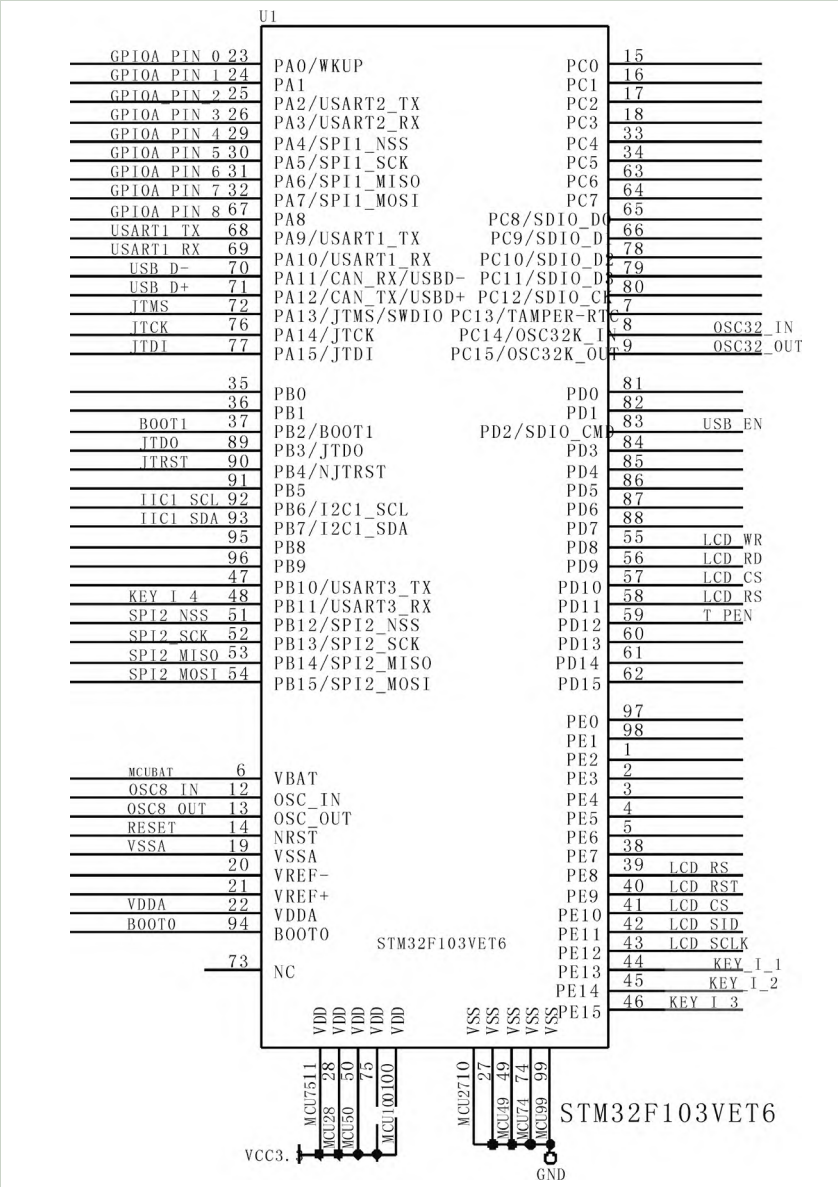


图 4 MCU 芯片 STM32F103VET6 引脚连接图

其中,系统采用原声头麦克风阵列实现对声音的拾取。麦克风阵列的拓扑结构是决定声源定位结果的一个重要因素,系统使用10个麦克风组成的均匀直线阵麦克风阵列。系统采用电容式驻极体话筒(咪头),该咪头能将一般的声音信号转化为电信号,接收范围广,有较高的输入灵敏度和输出信噪比,非常适用于空间内声源定位。为满足空间采样定理,麦克风阵元的间距设置需要小于信号最小

波长的一半,因此麦克风阵元的间距设置为10cm。声音信号经麦克风转换为电压信号以后是毫伏级的,考虑到MCU的ADC参考电压范围和麦克风原始信号幅值大小,系统单独设计了音频放大电路对麦克风接收信号进行调理,该电路正向端的设计使得输出电压具较大的直流偏置,解决了MCU采样电压局限性的缺陷,同时提高了声源信号的信噪比,进一步提高系统的灵敏度。

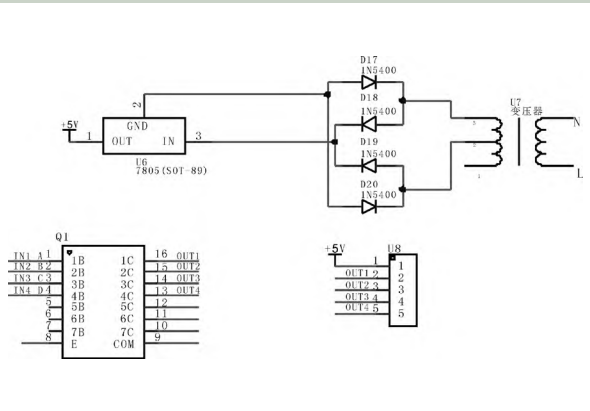
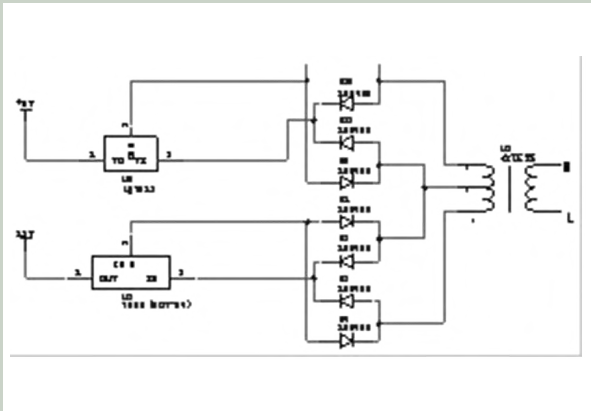


图5 电源模块电路+四相五线步进电机电路

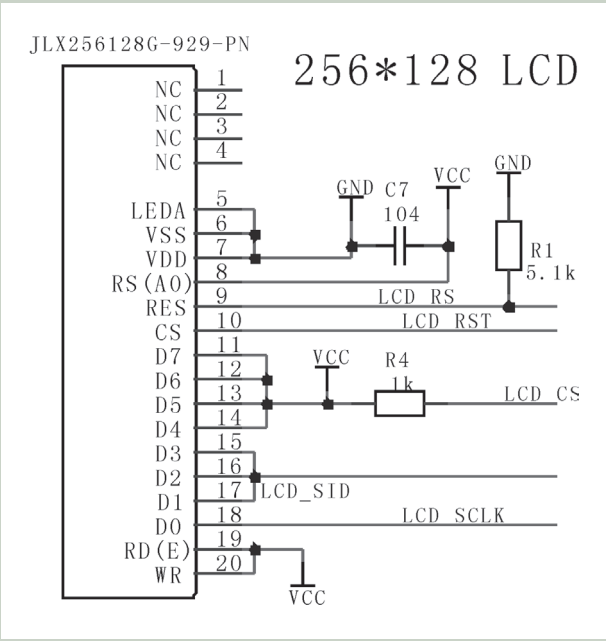
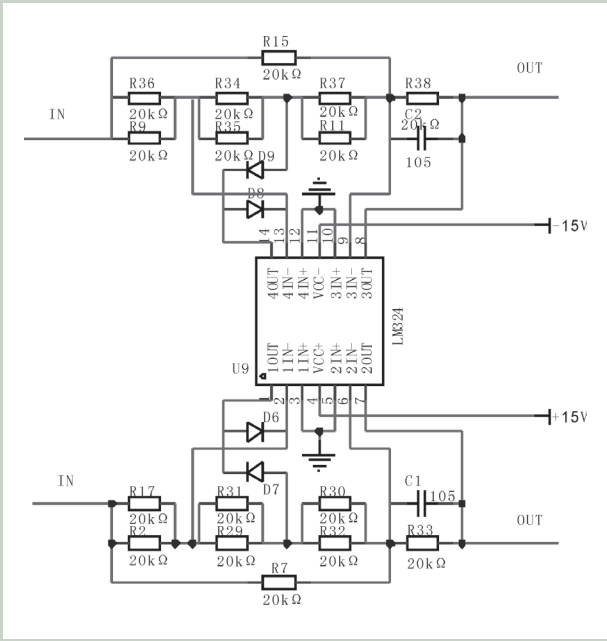


图6 LM324 音频放大器电路+256*128LCD 显示模块电路

(二) gccscot-TDOA 系统程序流程设计

gccscot-TDOA 声源定位算法包括信号接收、信号预处理、特征提取、声源位置解算、位置显示和动态跟踪等步骤。系统运用原声头麦克风阵列实现

对声音的拾取;信号预处理可以包括滤波、去噪、放大等操作,以增强信号的质量和可读性;特征提取可以从信号中提取出有用的特征,这里的特征是到达时间差,即时延;声源位置解算可以通过三角几

徐雪慧,梅振龙:基于广义互相关时延估计(TDOA)算法的声源定位跟踪系统设计与实现

何计算公式来确定声源的位置;再运用显示和跟踪设备进行位置显示和动态跟踪。gccscot-TDOA 系统的工作流程如图 7 所示,简洁有序,同时系统的整体程序部分算法设计合理,为整个系统的稳定性和精确性提供了坚实的保障。

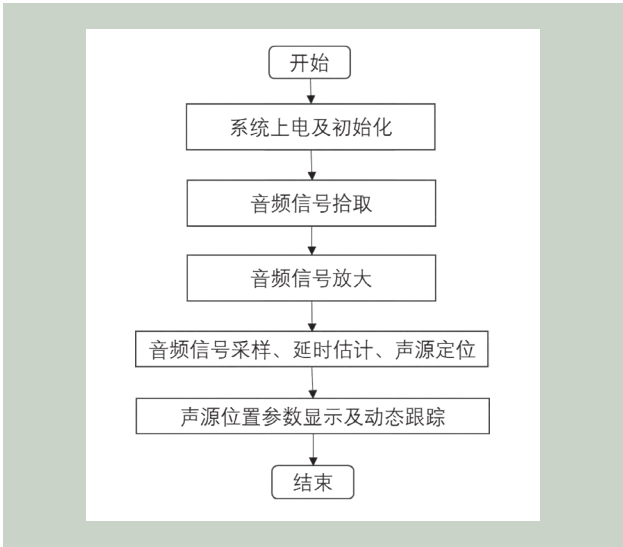


图 7 gccscot-TDOA 系统工作流程图

gccscot-TDOA 系统流程执行步骤如下:(1)声源信号被原声监听头麦克风阵列的麦克风接收;(2)通过前置放大电路放大;(3)进入 MCU 中进行采样

和 A/D 转换,进行一些滤波和去噪预处理操作,同时根据短时平均幅度,判断采集到的数据是否含有有效的声音信号,再利用广义互相关时延估计 TDOA 算法得到时延,运用时延参数结合几何定位法计算出声源的位置参数;(4)MCU 控制 LCD 显示测试位置参数,同时控制二维云台使激光笔对准声源并持续跟踪。

三、实验测试

(一)运用 MATLAB 仿真测试

为验证 gccscot-TDOA 声源定位算法系统设计的有效性和算法的正确性,在电路系统具体测试之前,针对系统电路中声源产生电路部分、信号放大、采样、A/D 转换及广义互相关平滑相干加权延时估计 TDOA 算法,均运用 MATLAB 软件进行了仿真测试。仿真效果如图 8、9、10 所示。

图 8 中的声波是依据声源信号产生器产生的声源信号完全相同特性进行仿真,得到一段持续性的声波,声波幅度大小变化,能仿真出接近于自然环境的声源,充分验证声源产生电路的合理性和有效性。在声源信号中,在检测到声波的峰峰值时开始计时以测量时延,降低了环境噪声干扰,可以有效减小测量误差,提升时延精度。

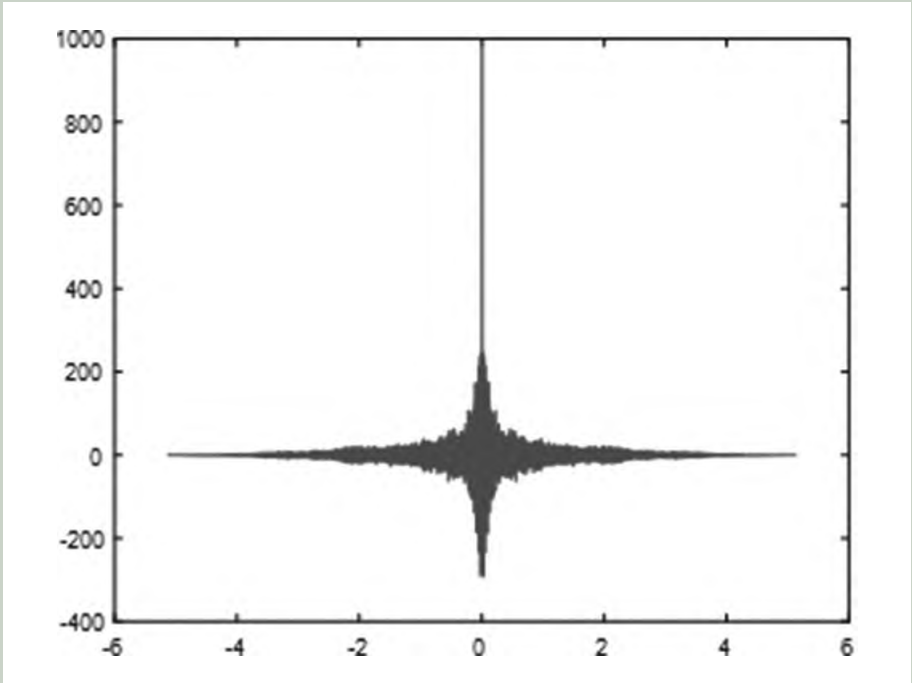


图 8 接近自然声源的声波仿真图

图 9 是运用广义互相关平滑相干加权时延估计 TDOA 声源定位算法仿真得出的抽检声源的仿

真图。仿真时运用与实际实验系统相同的麦克风阵列拓扑结构,设置了麦克风直线阵列,各麦克风阵元之间的间距也一样为 10cm。仿真时,随机设定多组声源点和麦克风阵列坐标,运用广义互相关时

延估计 TDOA 声源定位算法仿真确定的声源点坐标与模拟声源的坐标一致(仿真实验数据如表 1 所示),充分验证了算法的可靠性。

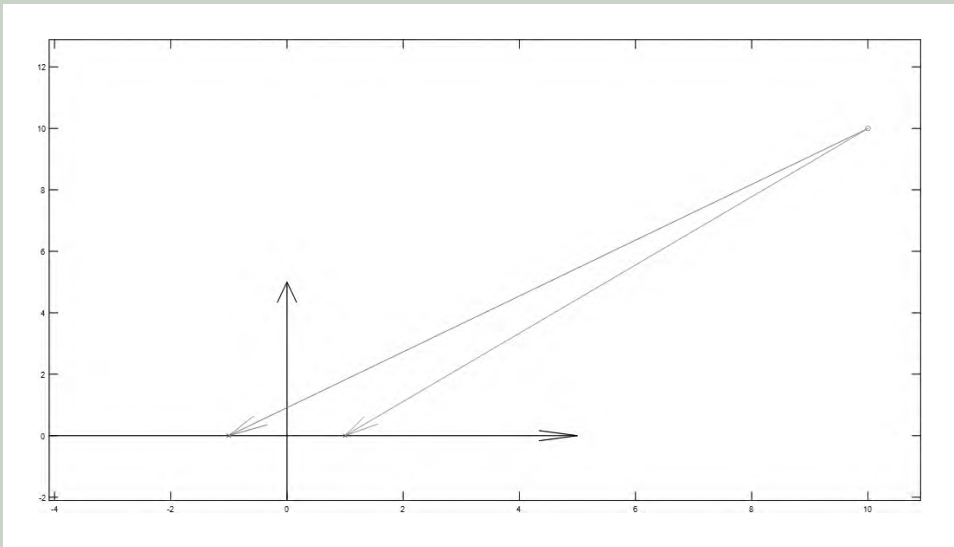


图 9 检测声源位置仿真图

表 1 MATLAB 仿真测试数据表

实验次数	预设声源位置坐标(m)	仿真定位声源位置坐标(m)	误差分析
1	(-2,2)	(-2.01,2.02)	(0.01,0.02)
2	(-4,7)	(-4,7.01)	(0,0.01)
3	(3,5)	(3,5.02)	(0,0.02)
4	(6,12)	(6.03,12)	(0.03,0)
5	(-4,13)	(-4,12.07)	(0,0.03)
6	(-2,9)	(-2.01,9.02)	(0.01,0.02)
7	(12,4)	(12.01,4)	(0.01,0)
8	(-1,11)	(-1.01,11)	(0.01,0)
9	(9,5)	(9.02,5.01)	(0.02,0.01)
10	(11,3)	(11.03,3.01)	(0.03,0.01)

同时,我们还运用 LevenBerg-Marquardt 的 BP 网络算法训练函数^{[10][12]}进行 MATLAB 仿真训练和测量,进一步验证了互相关时延估计 TDOA 算法的可靠性。用麦克风阵列对空间随机生成的声源进行采集,结合 BP 神经网络模型进行训练,利

用训练好的模型,输入四个麦克风的坐标数据,进行声源位置检测,测试结果如图 10 所示。仿真测试中互相关系数 R 接近 1,表明相关性好,进一步验证了 gccscot-TDOA 声源定位系统运用的算法可行性高。

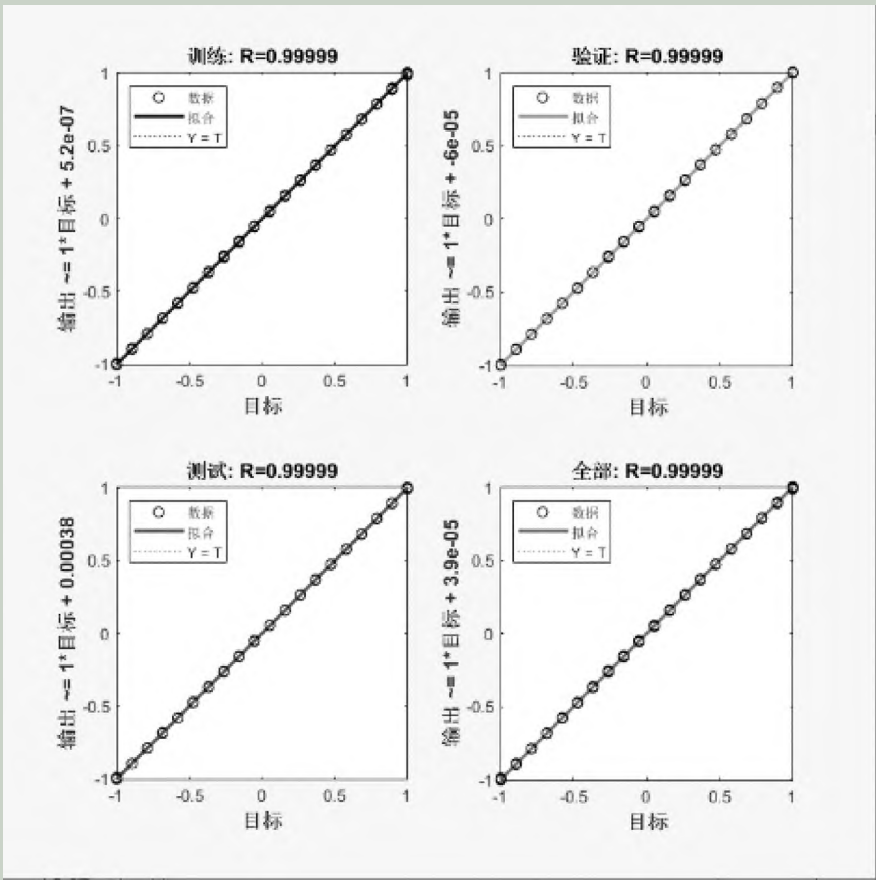


图 10 神经网络预测图

(二) 真实环境实验系统成品测试结果

经过仿真验证后,在真实环境进行了 gccscot-TDOA 声源定位系统的实验和测试。实验环境为没有强干扰噪声的普通实验室,将声源 S 和麦克风元

Mic1、Mic2、Mic3 直线阵置于同一水平面上,依次将声源放到经过准确测量的点进行实验,麦克风元之间的间距 d 为 10cm,实验系统测量区域俯视图如图 11 所示。

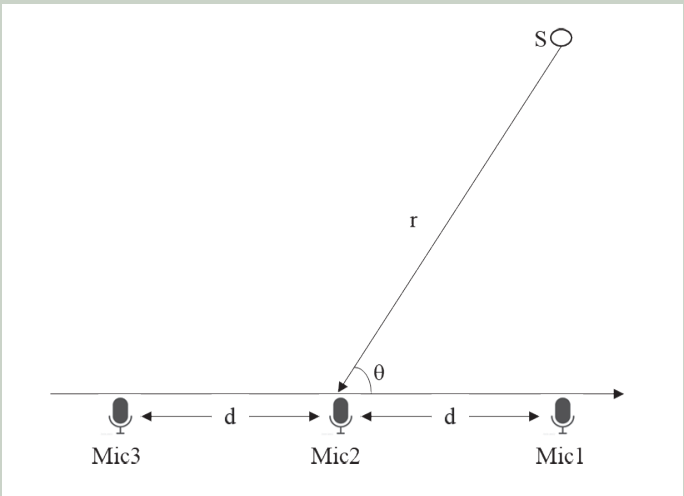


图 11 实验系统测量区域俯视图

用定时器 25 微秒计时,原声监听头阵元接收到声源模块的信号时,电压比较器产生跳变,MCU

进行处理并运用广义互相关时延估计 TDOA 算法计算到达不同监听头阵元之间的时间差。通过时间差计算出声源与麦克风元之间的距离 r 和夹角

θ , 并运用 LCD 屏进行测试结果显示。共进行了 5 次测试,实验数据和测试结果见表 2。

表 2 声源定位实验数据和测试结果记录表

项目 / 记录有效值 / 次数	第 1 次	第 2 次	第 3 次	第 4 次	第 5 次
设定角度 θ	20°	10°	-30°	0°	-25°
LCD 屏显示的测量结果 θ	19°	11°	-29°	2°	-23°
测量绝对误差	1°	1°	1°	2°	2°
设定声源距离 r	268cm	258cm	288cm	250cm	278cm
LCD 屏显示的测量结果 r	266cm	255cm	230cm	248cm	281cm
绝对误差	2cm	3cm	2cm	2cm	3cm
相对误差 /%	0.75	1.16	0.69	0.8	1.07

在运用二维云台进行移动声源的动态跟踪测试时,声源以 0.2 m/s 左右的速度移动,测量二维云

台的激光笔光电指向声源的偏差,测试结果数据见表 3。

表 3 声源动态跟踪测试结果记录表

项目 / 记录有效值 / 次数	第 1 次	第 2 次	第 3 次	第 4 次	第 5 次
位置偏差(cm)	1.5	1.3	1.9	1.6	1.1
反应时间 (s)	1.3	1	1.6	1.4	0.8

(三)测试误差分析

根据实验测量结果显示,在声源定位测试时,设定角度与测试角度之间误差为 1-2°,设定距离与测试距离之间的误差为 2-3cm,距离相对误差小于 1.2%,误差较小,表明定位精度高;在进行声源动态跟踪时,位置最大偏差不超过 2cm,最大反应时间是 1.6 秒,表明定位跟踪反应快,跟踪效果好。测试中发现,当声源离原声监听头麦克风阵元较远时,产生的误差比较大。误差产生的主要原因是环境杂声干扰和不同介质中声音传播速度不同,另外还有实验条件的限制和影响。

四、结论

随着人工智能的普及,利用麦克风阵列处理音频信号,并从中获得声源的相关信息已被广泛应用到生活的方方面面,声源定位技术有非常现实的研究意义和应用前景。团队设计开发的 gccscot-TDOA 声源定位系统,能够对声源进行实时定位、声源位置显示和声源动态跟踪。从仿真测试及现场实验发现,系统算法和电路设计经过模拟仿真及真

实环境实验测试表明,算法可靠性高,计算量小,精度较高,能有效地实现在室内环境下声源目标的精确定位。后续的研究可以从算法、电路的设计及器件的选择方面进一步优化和改进,实现灵敏度和定位精度更高的声源定位跟踪系统,以适应更多智能场合的应用。

参考文献:

[1] Yang Xu, Xing Hongyan, Su Xin. AI-based sound source localization system with higher accuracy [J]. Future Generation Computer Systems,2023(141):1-15.
[2] 张超洋,周桂宇,蒋立华,等.一种声源定位跟踪系统实现方法[J].物联网技术,2023(9):22-25.
[3] 郝张红,段羽浩.基于TDOA的声源定位算法研究及实现[J].天津理工大学学报,2020(2):1-2.
[4] 王松.基于TDOA的声源定位算法研究与实现[D].济南:山东大学,2020:6-13.
[5] 谭亮.基于麦克风阵列TDOA声源定位技术研究[D].阜新:辽宁工程技术大学,2019:27-47.
[6] 韩文革,苏淑靖,薛彦杰.基于麦克风阵列的声源被动定位系统设计[J].电子技术应用,2017(12):61-64.

徐雪慧,梅振龙:基于广义互相关时延估计(TDOA)算法的声源定位跟踪系统设计与实现

[7] 吉爱国,刘伟平,刘志强,等.麦克风阵列近场声源定位与跟踪系统[J].信息记录材料,2019(7):149-152.

[8] 岑林泽,徐一祺,杨宇帆,等.基于麦克风阵列的真实声场环境声源定位跟踪系统设计[J].电子制作,2023(8):17-21.

[9] 余萍,杨乘,王紫薇.基于联合加权的广义二次相关时延估计算法[J].计算机仿真,2023(3):400-404.

[10] 曾庆宁,苏盼,王红丽,等.基于互相关序列和BP网络的声源定位算法[J].科学技术与工程,2023(16):1-3.

[11] 焦琛,张涛,孙建红.基于卷积神经网络的室内麦克风阵列声源定位算法[J].激光与光电子学进展,2020(8):195-200.

[12] 王松.基于TDOA的声源定位算法研究与实现[D].济南:山东大学,2020:6-13.

[责任编辑: 卢艳红]

Design and Implementation of Sound Source Localization and Tracking System Based on Time Difference of Arrival (TDOA) Cross-correlation Algorithm

Xu Xuehui¹, Mei Zhenlong²

(1. College of Electronic Information Engineering, Wuhan Polytechnic, Wuhan, Hubei 430074, China;
2. National High Magnetic Field Center, Wuhan, Hubei 430073, China)

Abstract: To achieve real-time localization and dynamic tracking of sound sources in a two-dimensional plane, a CC-TDOA sound source localization and tracking system is designed and implemented based on the Time Difference of Arrival (TDOA) cross-correlation algorithm. The system synchronizes the sampling of multiple sound sensors, amplifies and processes the signals, and uses an LCD screen to display the distance and azimuth angle of the target sound source in real time. Additionally, a two-dimensional gimbal is used to control a laser pointer to align with the sound source and continuously track it. Simulation and real-world experimental tests demonstrate that the TDOA-based cross-correlation localization algorithm has low computational complexity and high accuracy, with an angle error of less than 2 degrees and a distance error of less than 1.2%. This system can meet the requirements of real-time localization and tracking of sound sources in various intelligent applications.

Key words: sound source localization; cross-correlation; Time Difference of Arrival (TDOA) algorithm; tracking system design