

基于麦克风阵列的高信噪比定向收音系统

何强, 杨朱杰, 郭晨, 严世涛, 孔鹏

(华中科技大学 电子系, 湖北 武汉 430074)

摘要: 研究讨论了用于麦克风阵列的高信噪比定向收音算法, 设计实现了麦克风阵列语音采集系统。通过对采集到的空间中不同方向音频进行数字信号处理, 使阵列形成的波束主瓣指向目标语音, 零陷指向干扰源, 提高收音信噪比, 实现对声源的定向收音等。测试结果表明, 本系统收音效果良好, 采集到的声音信号主瓣很窄, 能够实现高信噪比定向收音。

关键词: 麦克风阵列; 语音处理; SMI 算法; 定向收音

中图分类号: TP368.1

文献标识码: B

文章编号: 1674-7720(2013)01-0014-03

Microphone-array based high NSR oriented receiving system

He Qiang, Yang Zhuojie, Guo Chen, Yan Shitao, Kong Peng

(Department of Electronic Science & Technology, Huazhong University of Science & Technology, Wuhan 430074, China)

Abstract: Different oriented sound collecting algorithms which have high signal-to-noise ratio (SNR) are compared and a sound collecting system based on microphone array is introduced in the article. Using digital voice processing technologies to deal with the signal collected from different directions, the array utilizes beamforming to enhance the signal from the direction of interest, and attenuates unwanted signal by nulling other directions to increase SNR. Testing results show that this system can collect the signal with good effect and high SNR, and has a narrow main lobe beamwidth.

Key words: microphone array; voice processing; SMI arithmetic; directional sound collecting

目前, 在语音采集场合中(如舞台、大型会议室、电视会议等)通常使用孤立麦克风作为语音拾取工具。但是孤立麦克风会采集环境噪音, 在多音源场合中相互干扰难以避免, 这些缺陷严重影响了语音采集质量。随着阵列信号处理技术^[1]的发展, 利用麦克风阵列拾取语音信号为提高收音质量提供了可能^[2]。通过阵列信号处理的方法能够实现智能的语音信号优化效果, 实现语音定向采集, 提高信噪比。

目前, 麦克风阵列语音信号处理技术是语音采集技术领域的一个研究热点。CHANG A C 和 HUNG J C 等^[3]研究了 MUSIC 语音信号阵列处理算法, 在理论上证明了能够提高收音精度, 具备信号选择和提取等性能。GANNOT S 和 COHEN I 等^[4]研究了基于广义旁瓣抵消器结构的语音增强的算法, 认为 GSC 算法能够很好地提高收音信噪比。邵怀宗等^[5]设计了一种 12 阵元麦克风阵列, 提高了收音精度。杨祥清等^[6]提出的三维声源定位系统减少了阵元数量, 同时保持了一定的收音精

度。但国内尚无具有自主知识产权的产品, 所以研究麦克风阵列语音采集系统具有较高的市场价值。

本文从上述应用背景出发, 分析了基于麦克风阵列的高信噪比定向收音系统所涉及的相关算法, 重点讨论了软硬件系统的工程实现。麦克风阵列定向收音算法主要有自适应波束形成技术中的最小均方(LMS)算法和采样自相关矩阵求逆(SMI)算法^[7]等。自适应阵列的性能与算法的收敛速度密切相关。为了加快收敛速度并解决收敛速度依赖于特征值分布的问题, 常采用基于信号环境的采样自相关矩阵求逆(SMI)算法。本文采用 SMI 算法, 应用于易于生产、精度高于二维定位、实用性更强的 4 阵元麦克风阵列, 并使用 DSP 进行阵列信号处理, 以满足对声源信息定向采集的需求。

1 算法模型

自适应波束形成算法应用于麦克风阵列语音采集系统时, 能够随信号源的变化自动调节有关参数, 从而达到调节方向图主瓣方向的目的。该算法主要是对采集

到的麦克风阵列信号运行内部反馈控制,并根据一定的准则形成权向量,通过对接收到的信号进行加权叠加,使阵列方向图的波束主瓣指向有用信号,零陷或较低的旁瓣指向干扰信号方向^[8],从而将不同的信号从空间上实现分隔,实现定向采音。

1.1 阵列模型

本系统采用等距线性麦克风阵列^[9]。对于实际使用的阵列结构要求方向向量 $a(\theta)$ 与入射角 θ 一一对应,不能出现模糊现象。因此,阵元间距 d 不能任意选择,有时甚至需要非常精确地校准。假设 d 很大,则相邻阵元的相位延迟会超过 2π ,此时,阵列方向向量无法在数值上分辨出具体的相位延迟,就会出现相位模糊^[10]。对于等距线性阵列来说,其阵元间距不能大于半波长 $\lambda/2$ 。

语音的主要频率范围为 340 Hz~4 000 Hz,空气中声速约为 $C=340$ m/s,可得波长的范围为 0.085 m~1 m,因此 d 的范围为 4.25 cm~50 cm。而对于低旁瓣或零深陷的复杂波束,要求 $r=10L^2/\lambda$ (r 为声源到基阵的距离, L 为等距线性阵列长度)或更大距离^[11],考虑应用环境,取 r 范围为 2 m~10 m。经测试发现,一般人说话的声音频率在 1 000 Hz 左右,即 λ 在 0.34 m 左右。由此可以推算出 L 为 0.26 m~0.58 m。因此取 $L=45$ cm。

建立等距线性阵列模型,该信号在发射端表示为 $s(t)$,信道复增益(包括幅度和相位影响)为 $h(t)$,入射角为 θ ,以图 1 表示 M 阵元直线型麦克风阵列。

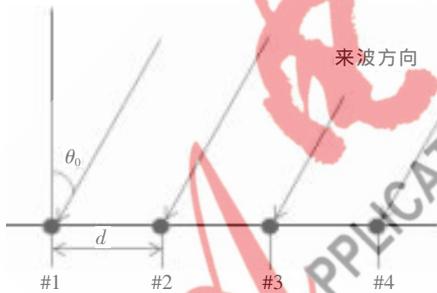


图1 M阵元直线型等间距阵元

该模型中,来波信号在阵元 #1(参考点)之间的相对相位为:

$$\phi_i(\theta) = -(i-1)kdsin\theta$$

阵元噪声为:

$$\mathbf{n}(t) = (n_1(t), n_2(t), \dots, n_M(t))^T$$

阵列连续信号模型为:

$$\mathbf{x}(t) = \mathbf{a}(\theta)s(t) + \mathbf{n}(t)$$

离散信号模型(此时假设有 k 路信号入射,入射角分别为 $\theta_0, \theta_1, \theta_2, \dots, \theta_k, \theta_0$ 为期望信号)为:

$$\mathbf{x}(n) = \sum_0^k \mathbf{a}(\theta_i)s_i(n) + \mathbf{n}(n) = \mathbf{A}(n) + \mathbf{n}(n)$$

其中,

$$\mathbf{A}(n) = (a(\theta_1), a(\theta_2), \dots, a(\theta_k))$$

$$\mathbf{s}(n) = (s_1(n), s_2(n), \dots, s_k(n))$$

由此得出的 $\mathbf{x}(n)$ 就是阵列接收的信号。

1.2 采样自相关矩阵求逆(SMI)算法

在实际应用中,接收数据是有限长的,很难得到自相关矩阵 \mathbf{R}_{xx} ,因此常采用自相关矩阵的估计值,即

$$\hat{\mathbf{R}}_{xx} = \frac{1}{N} \sum_{i=1}^N \mathbf{X}(i)\mathbf{X}^H(i)$$

其中, N 为采样数, $\mathbf{X}(i)$ 为阵列接收到的信号矩阵。

则该波束形成器要使噪声以及来自非 θ_0 方向的任何干扰的功率最小,同时又能保持在观测方向上的信号功率不变,其代价函数为:

$$\min_w \mathbf{w}^H \hat{\mathbf{R}}_{xx} \mathbf{w} \quad \text{subject to} \quad \mathbf{w}^H \mathbf{a}(\theta_0) = 1$$

利用拉格朗日因子法求解上式得到采样自相关矩阵求逆算法的权重向量:

$$\mathbf{w}_{SMI} = \frac{\hat{\mathbf{R}}_{xx}^{-1} \mathbf{a}(\theta_0)}{\mathbf{a}^H(\theta_0) \hat{\mathbf{R}}_{xx}^{-1} \mathbf{a}(\theta_0)}$$

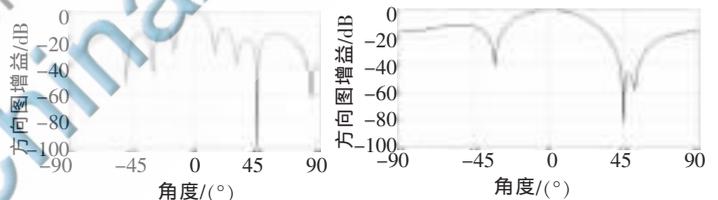
进而得到阵列的输出信号:

$$\mathbf{y}(n) = \mathbf{w}_{SMI}(n) \mathbf{x}(n)$$

$\mathbf{y}(n)$ 即为期望方向的声音信号,可以通过硬件电路处理后播放。

1.3 算法仿真

使用 MATLAB 软件对 SMI 算法进行仿真。图 2 所示为信号源方向为 0° , 主要干扰信号方向为 45° 的 8 阵元(图 2(a))与 4 阵元(图 2(b))麦克风阵列仿真结果。



(a) 8阵元 MATLAB 仿真图

(b) 4阵元 MATLAB 仿真图

图2 MATLAB 仿真图

由图 2 可以看出,该算法在 45° 方向形成了零陷,有效地抑制了主干扰信号。而在 0° 方向形成了具有一定宽度的主瓣,由于主瓣的增益大于所有旁瓣的增益,因此该算法能有效地采集到期望信号,抑制其他信号。

虽然,8阵元的仿真结果比4阵元的要好,主瓣较窄,零陷明显,但是算法复杂度较高,硬件实现较为困难。综合仿真结果和硬件电路复杂度,认为采集信号的麦克风的个数为4个,每个麦克风的间距为15 cm时,该算法的性能较好,且硬件电路较容易实现。

2 硬件实现

2.1 系统整体方案

图 3 为系统硬件结构图,包括4阵元直线型麦克风阵列、4路音频放大滤波电路、DSP 处理器以及音频编解码器。

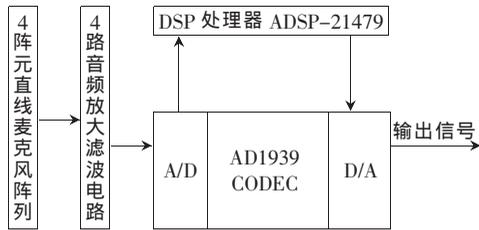


图3 系统硬件结构图

麦克风阵列采用4个驻极体式麦克风构成阵列;放大滤波电路对麦克风阵列采集到的信号进行预处理,通过RCA端子将预处理后的信号送往音频编解码器;处理器采用ADSP-21479,用SMI算法对量化编码后的4路音频信号进行处理,得到期望信号;音频编解码器AD1939对经过放大滤波后的4路音频信号进行量化编码,随后将DSP的处理结果经D/A转换后输出。

2.2 关键模块设计

2.2.1 DSP处理器与系统程序

数字波束形成^[12]的过程是一系列矩阵相乘的过程,其运算的数据量大,而信道环境是不断变化的,导致最优权值也处于不断的变化中,因此实际权值必须进行不断的调整,因而要求瞬时处理速度要快。

DSP处理技术^[13]可以运用在对瞬时处理能力要求更加苛刻的环境,DSP处理器和通用处理器最大的不同在于数据处理能力的增强,其核心是对连续存储的数据依次作重复的乘加运算^[14]。

另外,由于浮点型DSP处理器具有运算精度高等特点,因此本系统选择ADI公司的高性能浮点DSP处理器ADSP-21479芯片作为整个系统的核心。ADSP-21479是高性能32/40bit浮点处理器,具有高性能音频处理的功能;工作频率高达300MHz,满足实时性的要求;另外,还具有精简的指令集,编程较容易。

由于阵列信号处理是在信号的复基带进行的,需要进行大量的复数运算,因此如果没有简洁、优化的执行程序,算法的运算时间就会比较长。在本设计中,考虑到矩阵运算的复杂性,采用C语言进行编程。采用这种方式,可缩短软件开发的时间,提高程序的可读性和可移植性,但是在满足系统实时运算的要求上会有所缺陷。

基于上述讨论,DSP采用图4所示的流程图实现自适应波束形成。图4 DSP软件流程图



图4 DSP软件流程图

2.2.2 前置放大滤波电路

由于驻极体麦克风采集到的信号存在严重噪音,为了获得高质量的音频信号,在DSP板载RCA输入端子前加了前置放大滤波电路对语音信号进行预处理。经测试,放大滤波电路通频带为0~1300Hz,放大倍数为0~54dB

可调,典型值约为46dB,该电路可以有效降低噪声的干扰,从而提高音质。

电路原理图如图5所示。

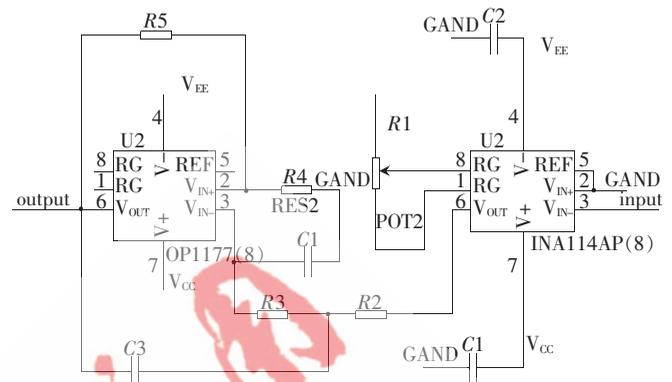


图5 放大滤波电路原理图

4 测试结果

为了减小混响以及其他因素对测试结果的影响,将测试环境选在一个大的房间中,且周围环境很安静。测试者手持一个特定音源,当测试者分别位于与麦克风阵列相距2m的0°角和45°角方向时,记录输出波形如图6所示。

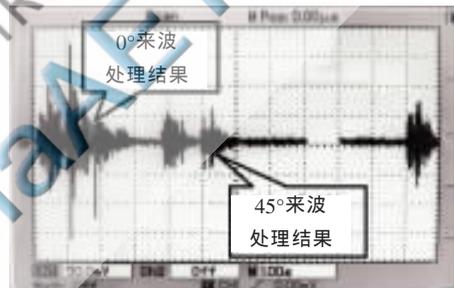


图6 测试结果输出波形

从图6可以看出,对于同样距离的同一声源,在主瓣(即0°方向)可以实现最大输出,而在形成零陷的45°方向将实现抑制。这个测试结果与图2所示的仿真结果相吻合。这样就实现了波束的形成,进而实现定向采音。

基于麦克风阵列的定向采音技术是一个新兴的领域,具有深刻的技术背景和广阔的应用前景。本文从算法模型到硬件实现详细介绍了基于麦克风阵列的高信噪比定向采音系统。本系统可以给出较好的采音效果,硬件实现也不复杂。在基于麦克风阵列的定向采音技术上,本系统还可以进行一些改进。在理论上,可以进一步提高定向采音精度,更快速地跟踪及更有效地去噪;在实现上,因为涉及多通道语音处理和更为复杂的核心算法,需要实现更加苛刻的实时信号处理要求。

参考文献

- [1] KRIM H, VIBERG M. Two decades of array signal processing research: the parametric approach [J]. Signal Processing

- Magazine, IEEE, 1996, 13(4):67-94.
- [2] KANEDA Y, OHGA J. Adaptive microphone-array system for noise reduction [J]. IEEE Transactions on Acoustics, Speech and Signal Processing, 1986, 34(6):1391-1400.
- [3] CHANG A C, HUNG J C. DOA estimation using iterative MUSIC algorithm for CDMA signals [J]. IEICE Transactions on communications, 2009, 92(10): 3267-3269.
- [4] GANNOT S, COHEN I. Speech enhancement based on the general transfer function GSC and postfiltering [J]. IEEE Transactions on Speech and Audio Processing, 2004, 12(6):561-571.
- [5] 邵怀宗, 林静然, 彭启琮, 等. 基于麦克风阵列的声源定位研究[J]. 云南民族大学学报(自然科学版), 2004, 13(4): 256-258.
- [6] 杨祥清, 汪增福. 基于麦克风阵列的三维声源定位算法及其实现[J]. 声学技术, 2008, 27(2): 260-265.
- [7] 桑怀胜, 李峥嵘. 智能天线的原理、自适应波束形成算法的研究进展与应用 [J]. 国防科技大学学报, 2001, 23(6): 83-89.
- [8] GRIFFITHS L, JIM C. An alternative approach to linearly constrained adaptive beamforming [J]. IEEE Transactions on Antennas and Propagation, 1982, 30(1): 27-34.
- [9] 王冬霞, 赵光, 郑家超. 麦克风阵列拓扑结构对语音增强系统性能影响的理论分析[J]. 辽宁工业大学学报(自然科学版), 2010, 30(1): 1-4.
- [10] BALLAL T, BLEAKLEY C. Phase-difference ambiguity resolution for a single-frequency signal in the near-field using a receiver triplet [J]. IEEE Transactions on Signal Processing, 2010, 58(11):5920-5926.
- [11] KENNEDY R A, ABHAYAPALA T D, WARD D B. Broadband nearfield beamforming using a radial beampattern transformation[J]. IEEE Transactions on Signal Processing, 1998, 46(8): 2147-2156.
- [12] TURCOTTE R L, MA S C H, AGUIRRE S. Method and intelligent digital beam forming system with improved signal quality communications[P]. Google Patents, 5856804: 1999-01-05.
- [13] RABINER L R, GOLD B. Theory and application of digital signal processing [M]. Englewood Cliffs, NJ, Prentice-Hall, Inc., 1975.
- [14] 刘书明, 苏涛, 罗军辉. TigerSHARC DSP 应用系统设计[M]. 北京: 电子工业出版社, 2004.

(收稿日期: 2012-07-01)

作者简介:

何强, 男, 1992年生, 本科, 主要研究方向: 数字电路设计, 信号处理。

杨朱杰, 男, 1990年生, 本科, 主要研究方向: 数字信号处理, 算法仿真。

郭晨, 男, 1990年生, 本科, 主要研究方向: 氧化石墨烯, 无铅压电陶瓷。