**电视内置阵列麦克风设计与实现研究**

**摘要**

人工智能的快速发展为消费电子公司带来了新的活力，智能液晶电视的普及，其高使用频率和屏幕可视化的优势使其成为最常用的家庭终端设备之一。在智能电视上安装远程语音人工智能程序可让更多用户体验智能交互环境，因此内置麦克风阵列和语音增强功能已成为近期研究的重点之一。由于电子通信产品的飞速发展，人们对语音通信的质量提出了很高的要求，因此本文研究了基于电视内置扬声器的麦克风阵列的设计和实现，以改善各种扬声器中广泛使用的语音。达到。

在本文中，我们首先介绍声场模型和麦克风阵列信号模型，然后研究麦克风阵列拓扑以及如何确定阵列的几何形状。在实践中，语音信号的噪声较为复杂，因此引入了语音信号的预处理技术，并分析了延迟变压器的形成方法和语音信号的数字处理。

其次，在分析电视远程语音设计要求的基础上，提出了一种电视内置麦克风的四通道语音增强系统的设计方法。该系统使用芯片TMS320VC5509A作为处理器，并使用两个音频编解码器芯片来构建A/D（D/A）转换电路。除了需要DSP芯片和A/D转换电路外，系统还必须支持外围电路（例如信号采集电路，滤波器电路，音频转换电路，存储电路和电源电路）。该系统的功能通过使用多微配置线性阵列系统来消除背景噪声并改善语音选择和通信质量，从而在嘈杂的环境中很有用。该硬件系统结构简单，处理能力强。

此外，本文设计并实现了系统的各种功能模块，包括系统的软件开发过程，系统程序初始化设置，DSP芯片实现，DSP VC5509A程序启动程序以及其他相关任务。

最后，本文对设计和实现的系统进行了详细的测试，结果表明，系统的各种指标基本达到了预期的结果，验证了系统设计和实现的准确性。

**关键词**：阵列麦克风；语音增强；DSP；软件设计；测试

**Abstract**

The rapid development of artificial intelligence has injected new vitality into home appliance enterprises.Because the intelligent liquid crystal TV has the popularity, the use frequency is high, the screen visualization advantage and becomes the use frequency one of the highest dot.Installing remote voice artificial intelligence program on smart TV can realize more humanized intelligent interactive experience, so microphone array voice enhancement has gradually become one of the research hotspots in recent years.With the rapid development of electronic communication products, people have put forward higher requirements for the quality of voice communication. Therefore, this paper studies the design and implementation of microphone array based on embedded processor, so as to realize voice enhancement and provide technical means for its wide application in a variety of speakers, televisions, teleconferencing and other aspects.

This paper first introduces the sound field model and the signal model of microphone array, and studies how to determine the topology and geometry of microphone array.Considering that the noise in the speech signal is more complex in the actual environment, the pre-processing technology of the speech signal is introduced and the shaping method of the delay transformer and the digital processing of the speech signal are analyzed.

Secondly, on the basis of the analysis of the requirement of TV far-field speech design, the design scheme of 4-channel speech enhancement system for TV embedded array microphone is presented.The system USES the chip TMS320VC5509A as the processor and two audio codec chips to form A/D (D/A) conversion circuit.In addition to the DSP chip and A/D conversion circuit, the system also needs to support the peripheral circuit (such as signal collection circuit, filter circuit, audio conversion circuit, memory circuit and power circuit).The function of this system is to eliminate background noise by using a multi-micro-configured linear array system, and to improve the useful voice selection function and communication quality in noisy environment.Hardware system structure is simple, strong processing capacity.

Thirdly, this paper designs and implements various functional modules in the system, including the software development process of the system, initialization Settings of system program, realization of DSP chip, and bootstrap of DSP VC5509A program.

Finally, the paper carried out a detailed test of the system designed and implemented, and the test results indicated that the system's various indicators basically achieved the expected effect, which verified the correctness of the system design and implementation.

**Key words:** array microphone;Speech enhancement;DSP;Software design;test

**目录**

[摘要 1](#_Toc15838)

[Abstract 3](#_Toc16602)

[第一章 绪论 7](#_Toc28926)

[1.1研究背景 7](#_Toc4093)

[1.2国内外研究现状 7](#_Toc21678)

[1.2.1国外研究现状 8](#_Toc20928)

[1.2.2国内研究现状 9](#_Toc10378)

[1.3研究内容 10](#_Toc27611)

[第二章 预备知识 12](#_Toc30158)

[2.1声场模型 12](#_Toc7025)

[2.2阵列的拓扑结构 12](#_Toc25215)

[2.3语音信号的预处理 13](#_Toc7714)

[2.3.1预滤波 14](#_Toc20163)

[2.3.2归一化处理 15](#_Toc4210)

[2.3.3加窗分帧 15](#_Toc5661)

[2.4延时—求和波束形成 16](#_Toc5284)

[2.5语音信号的数字化处理 16](#_Toc23401)

[2.6电视远场语音设计需求 17](#_Toc11581)

[2.6.1远场语音定义 17](#_Toc15815)

[2.6.2电视内置远场语音设计实现方式 18](#_Toc20680)

[第三章 电视内置阵列麦克风设计方案 20](#_Toc1498)

[3.1 TMS320CSSx系列DSP结构简介 20](#_Toc10388)

[3.2 TMS320VC5509A及Techv-DSP5509A开发板介绍 20](#_Toc5638)

[3.2.1 TMS320VC5509A介绍 20](#_Toc14975)

[3.2.2 DSP核心电路的选取 21](#_Toc32145)

[3.2.3硬件系统设计说明 22](#_Toc32276)

[3.3系统的硬件设计 23](#_Toc19921)

[3.3.1语音采集模块 24](#_Toc8058)

[3.3.2滤波放大电路 26](#_Toc2512)

[3.3.3采样频率和采样精度 26](#_Toc1805)

[3.3.4 5V电源电路 28](#_Toc3084)

[3.3.5音频转换电路 29](#_Toc12955)

[3.3.6 MCBSP多通道缓冲串口 31](#_Toc23358)

[第四章 功能模块的设计与实现 33](#_Toc5521)

[4.1软件开发流程 33](#_Toc28238)

[4.2系统程序初始化设置 33](#_Toc8105)

[4.2.1音频编解码TLV320AIC23B的初始化设置 34](#_Toc8405)

[4.2.2DSP的I2C模块初始化设置 35](#_Toc17867)

[4.2.3多通道缓冲串口McBSP的配置 36](#_Toc25831)

[4.2.4 SDRAM初始化设置 37](#_Toc1863)

[4.2.5定时器原理和初始化 37](#_Toc25595)

[4.2.6中断处理流程 39](#_Toc31122)

[4.3DSP的实现 41](#_Toc4130)

[4.3.1自适应算法的实现 41](#_Toc30503)

[4.3.2存储空间分配 42](#_Toc28655)

[4.4DSP VC5509A程序自举 43](#_Toc12862)

[4.4.1自举引导模式的配置 44](#_Toc711)

[4.4.2芯片简介及常用命令 44](#_Toc12923)

[4.4.3DSP SPI程序引导烧写步骤 45](#_Toc18535)

[第五章 阵列麦克风性能测试 47](#_Toc15547)

[5.1阵列麦克风性能测试 47](#_Toc3556)

[5.1.1测试实验一 47](#_Toc24236)

[5.1.2测试实验二 48](#_Toc29560)

[5.1.3测试实验三 50](#_Toc25564)

[5.2阵列麦克风实际应用测试 52](#_Toc31181)

[5.2.1测试准备 52](#_Toc26703)

[5.2.2测试环境搭建 53](#_Toc6837)

[5.2.3实际应用测试 54](#_Toc7409)

[5.3测试结果分析 56](#_Toc20578)

[第六章 总结与展望 57](#_Toc13397)

[6.1总结 57](#_Toc18331)

[6.2展望 57](#_Toc25243)

[参考文献 59](#_Toc7312)

**第一章 绪论**

**1.1研究背景**

语音作为音频信号的非接触式传输，就像视频和无线电一样，是人类在没有工具帮助的情况下可以自由控制的唯一“无线”自然资源。然而，当在真实环境中进行语音交流时，环境中的噪声会干扰语音。在这种情况下，噪声是指除有用的语音信号[1]外的所有干扰信号。种类有很多噪音,在现实世界中,例如室内回响,外部噪音,周围人的复杂性、电磁噪声实验室etc.Ceci则是因为在使用的话音通信方法及引发的各类噪声影响的人们。因此，语音增强的功能是处理强语音并获得最纯净的语音信号。目前语音、加固技术被广泛用于视频会议、军事通信、窃听电话和语音识别,这实际上提高了语音清晰度视频内容,提高了语音通信设备的质量和处理语音的噪声污染问题。对于实际情况。它在其有用性方面发挥了重要作用，值得研究。

一组麦克风由两个或两个以上的麦克风组成，这取决于空间的形状。最近，使用一个麦克风来提高语音质量的效果非常好。但是在现实世界和应用中，例如在房间里说话时，由于声源位置的变化和多次反射，许多噪声信号与单个麦克风接收到的语音信号混合在一起。为了解决这些情况下的语音混淆问题，研究人员使用多个微配置麦克风网络进行语音通信，以弥补单个麦克风的缺点。轨道定位，有效抑制干扰和环境噪声。

与快速发展的云计算和大数据、人工智能已经逐渐浮出水面,虽然Siri语音控制功能的apple和Amazon ECHO智能扬声器、手机应用成为在全世界广泛使用。与此同时，京东与香港科技大学联合推出的犀牛语音助手、语音云等智能语音产品也由中国智能语音技术公司[3]推出和开发。Siri是一个短程语音应用程序，ECHO是一个长程语音应用程序。

智能设备与人的互动主要有三种方式:触摸、手势和声音。语音交互、更自然、更理性的功能以及家庭环境中更少的任务，使得语音成为一种越来越普遍的交互方式。短程语音根据使用的终端而变化，而远程语音提供了更多的开放和方便。智能液晶电视得到了广泛的应用。智能电视具有高频和可视显示器的巨大优势。它在形成使用习惯方面具有非常特殊的优势。

**1.2国内外研究现状**

**1.2.1国外研究现状**

自20世纪80年代以来，世界各地的研究人员逐渐提出了一种基于麦克风阵列和噪声条件[4]的语音放大算法。国外对麦克风语音增强算法的研究比国内快得多，取得了许多丰富而经典的实验成果。

基于语音的短期光谱估计的光谱减法，该方法是在语音信号和噪声信号是相互独立的，噪声信号是恒定或缓慢变化的附加白噪声的情况下提出的。该算法具有较好的语音增强效果，但由于噪声分量较大，在处理过程中会产生间歇性的音乐噪声。为纠正缺点谱减法,研究者们开发了许多办法来改善谱减法等基本谱减法得到改善,人耳听觉掩蔽效应的基础上,结合算法的波束成形和功率谱减法和一种改进的多谱减法。现有的光谱减法得到了改进，语音增强效果得到了显著改善，降噪不再局限于平滑时的白色噪声，还可以更好地抑制彩色噪声[5]。简易固定波束形成算法所耽搁,谁先参考麦克风,麦克风的位置,并根据这一基准,压缩每个元素所购买的相应延迟语音信号的装置,使信号的相位同步和处理装置。放大算法的语音延迟为简易的波束成形的放大与某些功能很简单、简单的言语,但大量的删除是必要的,以便获得更好的降噪麦克风和很少被用作治疗解决上游的语音信号处理。

自适应波束成形方法目前得到了广泛的应用。两种自适应波束成形算法中最典型的理论原型有两个方向，第一个是线性应力自适应波束成形算法。第二种是基于改进的LCM V算法的传统侧叶补偿器。gsc算法的优点是麦克风数量越少，降噪效果越理想，因此被研究人员[6]广泛使用。然而，有一些问题，如有用的语音信号组件泄漏，影响结果。

在过去的几年里，研究人员石武虎、Shiung D、Choi YH和Wi Wi Liu扩展了gsc的使用，并随机获得了强大的自适应波束形成。我们提出了一种传统的功能性侧叶清除剂。建立了具有比例通道传递函数的时变锁紧矩阵，利用余弦函数调制锁紧矩阵，对原广义侧叶消除算法[7]进行了改进。来探索和改进训练方法,提高自适应波束的语音结构的国家,研究人员使用了频段采用自适应技术和技术司的波段,以减少复杂性的降噪算法性能的同时保证。改进后的语音增强系统的收敛速度大大提高，输出信噪比也得到了提高。gsc算法具有良好的相干噪声抑制效果，但对于消除相干噪声和弱相干噪声并不理想。对于广义侧叶补偿器，算法在复杂的环境中不能很好地工作，不能补偿其缺陷。自适应波束形成器与维纳滤波器或卡尔曼滤波器相结合的语音增强算法提供了自适应语音增强技术。它已成为改进[8]的新研究中心。2003年，科恩提出了一种多通道过滤语音放大的方法。然而,即使经过治疗目前加强过滤后的语音处理、噪音问题仍然存在,并在回答这个问题时,中国一位研究员,zhang xujing语音增强方法,提出了一种基于卡尔曼滤波的组合subinverse subinverse技术进行分解。

这项实验始于1985年，当时美国电话电报公司贝尔实验室(ATT Bell labs)引进了外国设备，利用一组400多种微配置的麦克风构建了一个智能信号处理系统。其目的是跟踪声源在宽敞空间[9]中的位置。自21世纪以来，外国研究人员一直在使用大型麦克风网络来设计基于麦克风网络的语音增强系统。

**1.2.2国内研究现状**

一套麦克风语音放大系统的设计在中国起步较晚，麦克风的使用通常比国外少。2008年,李兴元、第十一和电子科技大学,一年解决了这一问题,采用固定宽带波束在频域,通过提高固定宽带波束在她的论文并进行固定宽带波束语音地带。降低波束对低频干扰[10]的抵抗力。12个电容麦克风用于在同一空间形成线性网络，德州仪器(TI)生产的数字浮点信号处理器TMS320C6713B用于设计语音放大系统。在延迟估计中，只分析了算法对无关噪声的性能。硬件系统采用波束成形算法。在同一年,大连理工大学提出了一个经验模态分解为组件分开报告中,噪声和方法经验模态分解(EMD)与收缩功能,提高了后处理的降噪能力。给出了4通道语音增强系统硬件电路配置的概念，但没有在硬件电路[11]中实现理论算法。此外，多通道子空间的分解方法计算成本高，难以在硬件上实现，这一缺点限制了其实际应用。

2009年，中国电子科技大学的陆志恒利用最小均方根误差算法改进了优先信噪比和噪声功率谱估计，以提高优先信噪比的估计精度，更新噪声。功率谱估计的新算法。该算法可应用于多个DSP系统族(it系统C\_5000和C6000)，有效消除多通道语音信号中的干扰噪声。2011年，北京大学、邮电陈晨改进了基于MMSE减法的频谱使用。该方法具有更大的信噪比范围，T1 TMS320VC5510为单片机DSP芯片的最终确定引入了TMS320VC5510，加强了系统向语言系统的过渡，缺点是系统[12]存在剩余噪声。2012年，重庆大学的Chan Chaofan提出了一种语言增强算法，用于人耳的掩体效应，为了降低计算和编程的复杂性，只使用DSP的内部内存，导致数据收集和观测的减少。

近年来，利用网络收集语音信号和利用语音增强方法获得高质量的语音已成为研究的新方向。此外，还对盲束分离方法、神经网络遗传算法和子空间分解方法进行了研究。随着麦克风网络技术算法的日益成熟，电子语音增强设备[13]的使用变得越来越迫切。DSP芯片的使用钛作为中央处理器的处理器芯片,引起行业越来越重视,但这些产品仍在发展和DSP可以提供硬件支撑低复杂性的语音增强算法,算法。DSP的处理速度和转换精度不能满足当前的要求。因此，正在积极进行研究和应用，以提高国内外数据收集系统的实时性能。

**1.3研究内容**

本文分为5章。

第1章简介。研究背景，发展现状及主要研究方向。

第2章相关概念。介绍该主题的基本理论。这主要包括语音信号的基本知识，麦克风子阵列的排列结构，多个语音信号的预处理方法，延迟1和波束成形方法以及语音信号的数字处理。

第3章电视远程语音设计要求。在学习和总结现有算法的基础上，提出了一种基于改进的动态阶跃因子最小均方误差的自适应麦克风阵列语音增强方法。该方法实现硬件简单方便。经MATLAB仿真验证的该方法显示出快速收敛速度和较小的稳态误差。使用主观和客观评估指标进行的分析和改进表明，输出声音的主观听觉良好，输出信噪比高。

第4章语言增强系统的设计。本章的内容是系统硬件设计和软件实现。

第5章测试内置阵列麦克风的语音增强系统。首先，初始化AD芯片和DSP外设寄存器。其次，CCS总结了DSP系统的软件开发过程，并提供了在处理过程中实现关键子功能（例如，自适应滤波器，中断处理功能和存储空间设计）的过程。最后，完成了DSP系统对语音的噪声去除处理，分析系统的处理结果和程序的增强。

**第二章 预备知识**

本章介绍了基于麦克风布局的电视语音增强研究的概念和技术。主要是模拟信号的声场和麦克风,确定网络的网络拓扑的麦克风和网络结构、语音信号的预处理方法调节变压器的落后和数字语音信号处理。

**2.1声场模型**

声场模型可根据声源的距离和距离以及麦克风网络的距离分为近场模型和远场模型两类。在近场中，声音之间的关系随着距离的增加而变得更加复杂，必须考虑接收到的信号之间振幅的差异，因此声波被认为是球面波和近场(14)。在远场模型中，声波被认为是平面波，它忽略了矩阵中不同元素接收到的信号之间振幅的差异，并认为它或多或少只是接收到的不同信号之间的延迟关系。显然，远场模型是对实际模型的简化，大大简化了处理困难。语音增强的一般方法是基于一个远程领域模型。

音源到麦克风的距离在这两者之间时，可以根据公式明确区分近场和远场。以均匀线性拓扑为例，阵列宽度为，采样频率是，对应语音信号最短波长，声源到麦克风阵列中学的距离为。若，则该模型为近场模型，否则是远场模型。声场模型如图2.1所示，本文的声场模型是远场模型

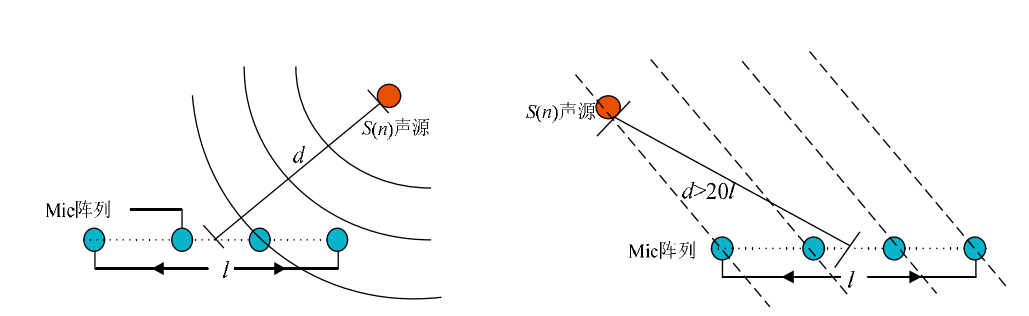


图2.1声场模型

**2.2阵列的拓扑结构**

一组麦克风由两个或两个以上的麦克风组成，按照给定的空间结构排列。根据空间位置的差异，将矩阵拓扑划分为三种主要类型:均质线性矩阵、非均质线性矩阵、一维线性矩阵和可叠加线性矩阵[15]。平面二维阵列，包括均匀环形阵列、非均匀环形阵列、矩形阵列和分布式阵列(具有更不规则阵列的麦克风阵列);图2.2、2.3和2.4显示了特定三维阵列(球面、立方体等)的结构图。

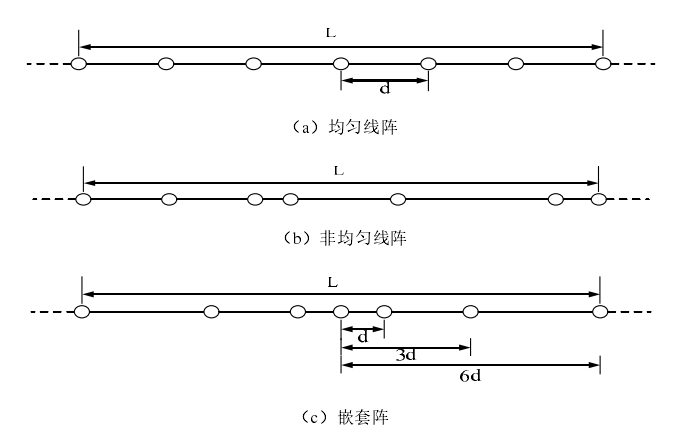


图2.2一维线阵

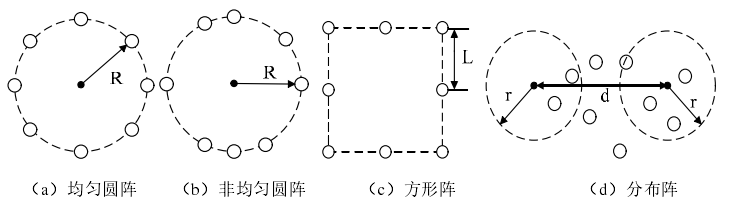


图2.3二维面阵

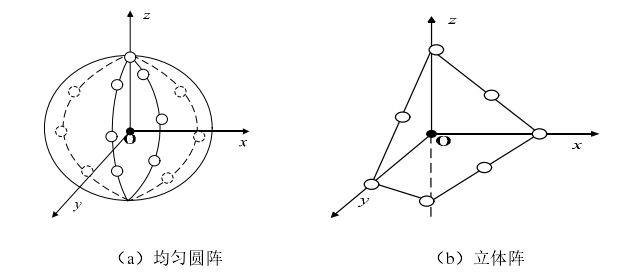


图2.4三维立体阵

三类麦克风阵列的拓扑结构应用于不同的场景，各有各自的优势。

**2.3语音信号的预处理**

**2.3.1预滤波**

语音信号中混有许多噪声，且频带范围宽。因此对语音增强处理之前，先预滤波。论文MATLAB实验中，设计FIR带通滤波器。根据语音信号的频带范围，设定、、和。阻带衰减为60dB，通带衰减为1dB。图2.5是该带通滤波器的幅度响应，图2.6是预滤波处理图。

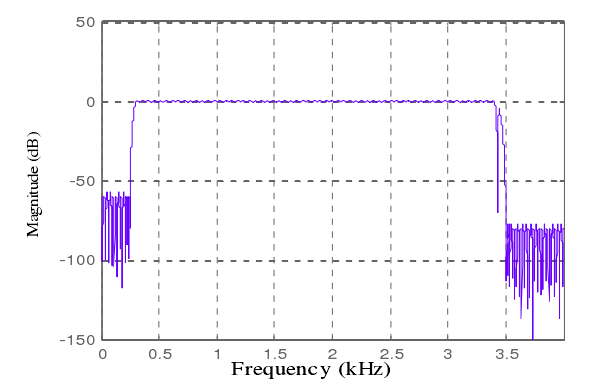


图2.5 FIR滤波器的幅度响应

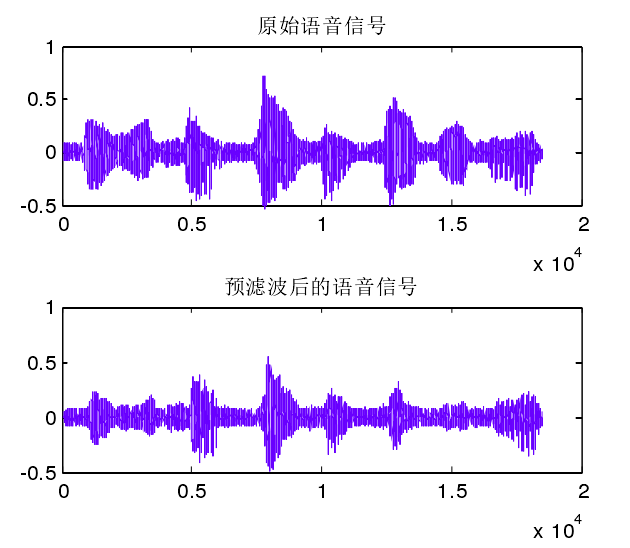


图2.6预滤波语音信号效果图

**2.3.2归一化处理**

在传输过程中阻塞的对象会干扰信号，导致信号处理和分析不准确。为了消除这种影响，改进的算法在处理前被标准化，即使语音信号的能量均匀地分布在x轴上。

**2.3.3加窗分帧**

语音信号是非静态的，随着时间的推移而变化，需要图像处理。由于消声器的惯性运动，语音信号在很短的时间内(通常在10到30毫秒之间)几乎保持不变。将一个完整的语音信号分成几个部分进行处理的过程称为电源[16]。此外，前后两幅图像之间的对应部分称为“纬纱运动”，纬纱运动值一般小于纬纱长度的一半。图2.7为纬纱长度和运动示意图。

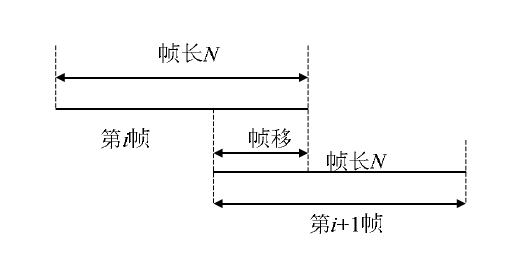


图2.7重叠分帧

**2.4延时—求和波束形成**

延迟交联法是在麦克风阵列中处理语音信号最简单、应用最广泛的改进方法。这种方法可以直接添加未经处理的麦克风网络接收到的信号，因此延迟成形和波束成形方法也称为单波束成形[17]。麦克风网首先被选中作为参考,然后借助网络的延迟时间估计(det)之间的时间差异,所收到的信号网络的另一个麦克风和参照所收到的信号网络的元素是延误的补偿计算是计算。最后，补偿信号是逐步产生的。偏移图如图2.8所示。

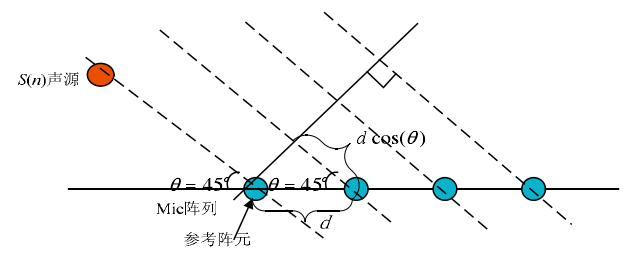


图2.8时延估计示意图

**2.5语音信号的数字化处理**

声信号是麦克风接收到的声波，并以相应的能量转换成电信号。然而，大多数语音增强系统用于数字语音处理，需要对模拟信号进行数字化，如图2.9所示。

首先，输入语音模拟信号通过A/D转换器转化为数字信号，为使转换后的信号能够无失真地还原出原始语音信号，采样时须遵循奈奎斯特定理。再由数字信号处理器DSP对进行各种算法处理，得到相应的输出信号，最后经过D/A转换器和平滑滤波器对数字信号进行解码，得到所需的模拟信号。在通信领域、消费电子产品和军事等领域，数字信号处理起着越来越重要的作用。

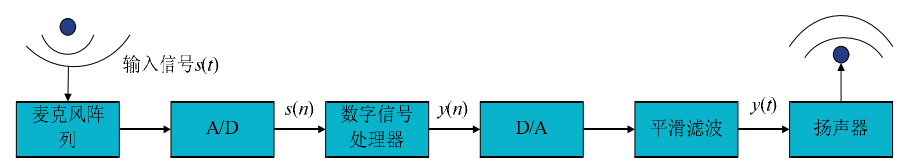


图2.9语音信号数字化处理流程

**2.6电视远场语音设计需求**

**2.6.1远场语音定义**

远程语音是一种集成语音搜索功能的语音交互技术，已广泛应用于智能电视和智能扬声器。它允许远程语音控制，而不需要远程控制从5米的距离。智能电视和智能扬声器可以直接由远程唤醒的单词激活，而不需要连续按下现有蓝牙语音遥控器上的语音按钮。

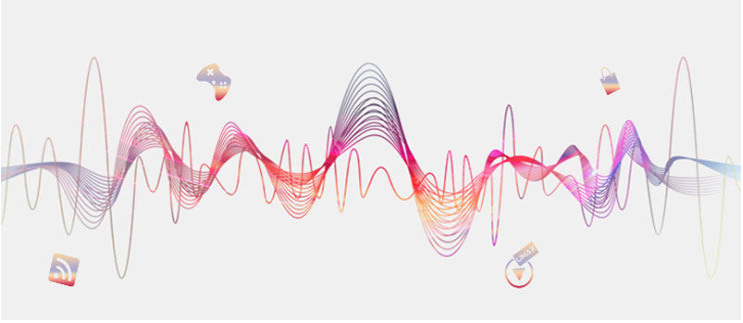


图2.10远场语音

**2.6.2电视内置远场语音设计实现方式**

麦克风传感器:语音信号由麦克风硬件模块接收。麦克风阵列可以使用多种形式，如线性、L形或球形。不同类型的网络使用不同数量的麦克风，以优化任何远程交互的准确性。麦克风的选择可以是ECM或MEMS麦克风，而MEMS麦克风可以是数字(集成ADC)或模拟麦克风。

该网络的设计基本上是两个或两个以上的脉冲，拜耳主要用于电视接收180度声源和其他产品。Domill用于电视机、扬声器、智能家用电器等产品，分布在可接收360度噪声源的线性多密度网络和多密度圆形网络中。

语音提取:利用话筒在环境中采集的多通道语音信号上的语音芯片进行一系列操作，消除回波，降噪处理，有效的功能提取控制[18]。通过模拟输出或IIS格式传输到主电视芯片。语音处理芯片有两种类型。一种是使用软件降噪算法的通用中央单元。一种是通过ASIC芯片上的材料处理获得语音控制的清晰信息。

语音识别，语义理解:语音识别是将语音转换为文本信号的过程。为了理解这意味着什么，命令从文本消息(上面的语音命令节点)中提取。语言识别和语义理解是涉及语言学、心理学、逻辑学、声学、数学和计算的新领域。充分利用现代需求、语音、语义和实践知识，使人与机器在自然语言中的交流成为可能，意味着计算机必须使用自然语言文本来理解自然语言文本的含义并表达给定的意图。随着思想等人机交互语言的进一步探索和大数据技术的发展，语言教学继续沿着语音识别和理解的道路发展，并取得了一定的成果。

智能电视:智能语音的语音服务提供定期服务,并根据实际需求定制产品,包括在线视频点播、操纵、控制家里的智能菜单购物、聊天、百科知识、气象和transport.Soutien问题。门户网站集中,当您的宽屏液晶电视的问题,在视频、搜索、音乐搜索项目,smart tv功能包括可执行的远程办公,提高用户体验和用户提供一个更好的舒适性。使用智能语音服务需要支持语音提供者的内容生态系统。语音识别和意义理解使公司成为一个大公司。换句话说，sci、百度du mi等完成了智能电视的“语音功能远程设计”。通过为电视产品选择远程语音解决方案，并将产品定位成本与开发周期相结合，可以采用三种设计，如图2.11所示。

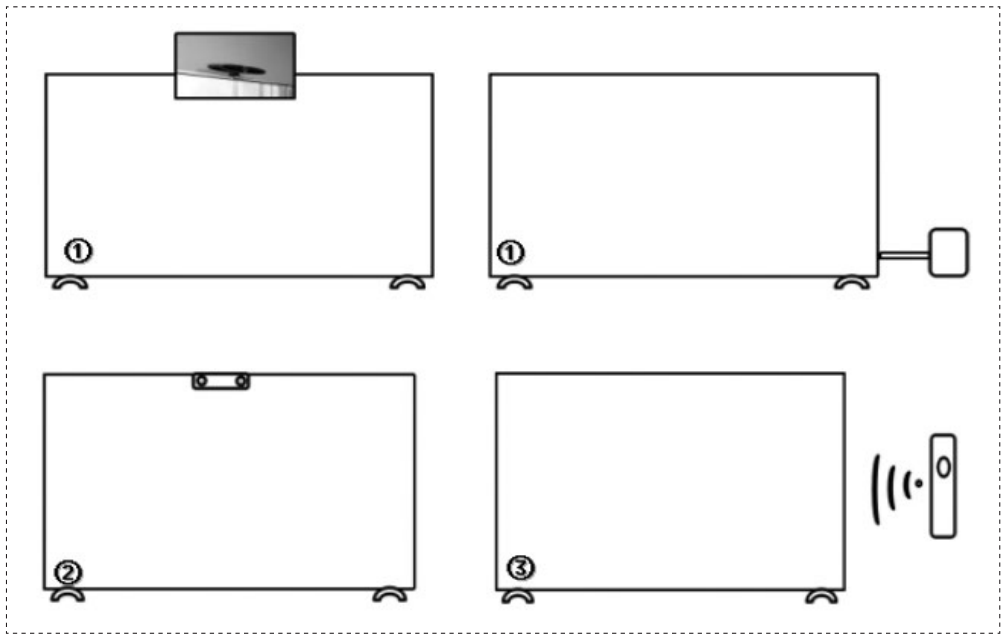
****

图2.11智能电视远距离语音功能设计

(1) USB端口可连接到远程语音功能模块，该模块可通过扩展数据电缆连接到计算机、扬声器或视频机箱中。你可以在购买时购买这幅画。

(2)远程语音模块集成在一个简单、美观、美观的电视机绝缘板中，如笔记本电脑摄像机。

(3)将远程语音模块放置在无线扬声器中或集成到无线遥控中。语音信号通过wi-fi、蓝牙、RF 2.4G等通信传输，提供无限的自由和感觉。

电视可以在客厅的任何地方提供快速响应，而不需要与遥控器对话。远程人工智能语音交互的操作系统是基于语音识别、自然语言处理、机器学习、大数据等技术,为用户提供一个容易和准确的语音互动体验,通过功能强大的云计算、多功能服务、信息、etc.Vous能全面了解[19]。用户不需要寻找遥控器，就可以在看电视的同时响应即时请求，从而获得完美的体验。随着技术的发展，语音内容商会将进一步扩大其活动，越来越多的智能液晶电视公司将接受支付资源的门槛。智能电视用户可以自由移动他们的嘴。你可以在一个方向上完成必要的任务。

**第三章 电视内置阵列麦克风设计方案**

本文介绍了一种集成在电视网络中的4通道麦克风语音增强系统的设计。系统采用TMS320VC5509A芯片作为处理器，两个音频编解码器芯片形成A/D (D/A)转换电路。除DSP芯片和A/D转换电路外，系统还应支持外围电路(包括信号采集电路、滤波电路、音频转换电路、存储单元和电源部分)。该系统的功能是利用多组线性网络消除环境噪声，提高噪声环境下的语音选择功能和通信质量。硬件系统简单易用。

**3.1 TMS320CSSx系列DSP结构简介**

C55x是TI C5000系列DSP的子系列。 C55x处理器的低功耗和高性能特性使其非常适合小型个人和移动设备服务，例如手机和可编程音频设备。

（1）C55x的结构特征

图4.1是C55x的框图。提供4个40位通用累加器，40位ALU和1个16位ALU，以优化并行工作。为了提高数据吞吐量，C55x CPU提供了12条内部总线。该总线允许CPU在一个时钟周期内执行3个数据读取操作和2个数据写入操作。

（2）C55x CPU

C55x的CPU主要包括4个主要单元的:指令缓冲设备IU C指令，程序流单元（PU），地址数据流单元（AU）（寻址-），该功能单元是程序总线，数据总线和地址总线交换数据和程序。该并行总线可以在一个时钟周期内读取错误位程序，可以读取3次16位数据，并且可以两次写入16位数据。

**3.2 TMS320VC5509A及Techv-DSP5509A开发板介绍**

**3.2.1 TMS320VC5509A介绍**

TMS320VC5509A是基于C55xCPU处理内核的定点数字信号处理器，具有较高的可行性和较低的功耗。

TMS320VC5509A具有许多外围设备。 5509A的片上外设的任务是完成与外部设备的通信。关键特性：控制外围设备，处理传感器数据收集，输出处理结果[21]。

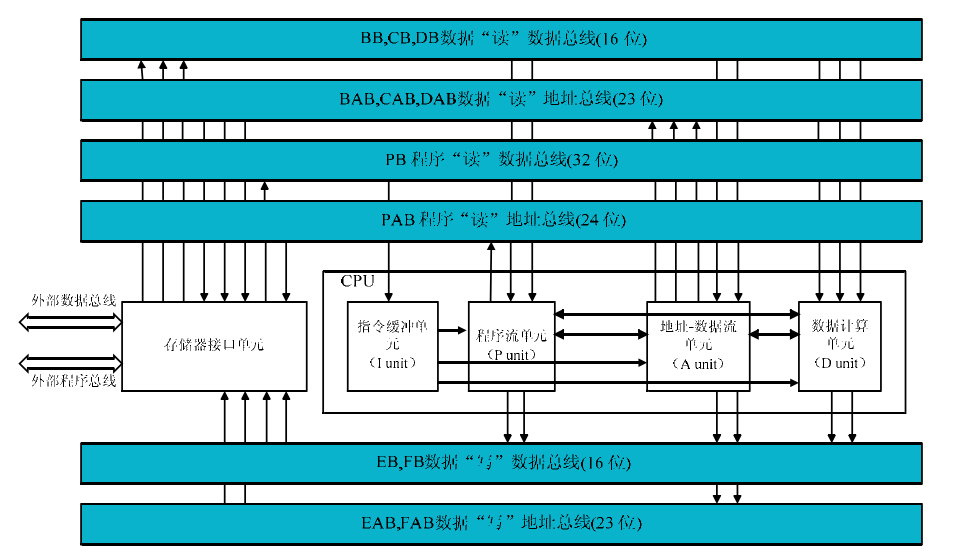


图3.1 C55x的结构图

系统设计主要包括六个:时钟发生器，命令缓冲区（高速缓存），外部存储器接口（EMIF），多通道缓冲区串行端口（MC bsp）实际I/O端口（RTC）。

**3.2.2 DSP核心电路的选取**

关于DSP核心电路的开发工作很多，周期长且工具不足。目前，市场上有很多基于TMS320VC\_5509A的开发板，除了最低的带DSP的系统外，还有一些基本的外设和扩展的I/O端口[22]。因此，决定暂时放弃独立的四通道采集系统的开发，并从北京Secorda Star购买Techv-DSP 5509A的发行版，并根据该主题的需要对模块电路进行了更改和添加。类型是：

（1）内存资源

Techv-DSP 5509ADSP芯片，RAM: 256K RAM除外，4M \* 16位SDRAM，设备类型MC 48 LC 4m 16；闪存: 4Mb串行闪存（sst25 VF 040）。

（2）可编程逻辑控制单元CPLD

（3）音频编解码器

编解码器: TLV320AIC23B，MIC模式输入，单通道采样。

（4）外部接口和总线:SCI，USB接口，与计算机的串行通信。图3.2显示了Techv-DSP 5509ADSP框图。

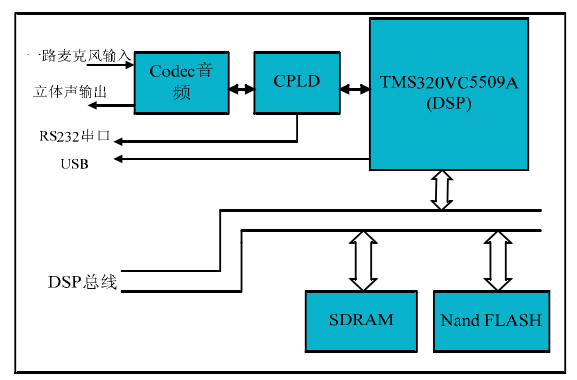


图3.2Techv-DSP5509A开发板原理框图

**3.2.3硬件系统设计说明**

techv-dsp5509a选中该条版本只允许单向获取语音和音频输入接口,可连接只特定的设备(例如,扬声器和耳机接口的),可以收集真实环境下的实时语音信号。本文对四向语音信号采集的要求不能满足，不存在语音扩增[23]。同样，原始的techv-dsp5509a系统，其硬件上有一个SDRAM内存芯片，还没有开发出来，也没有足够的空间来存储在处理语音增强算法时生成的数据和程序代码。

本文的目标是基于techv-dsp5509a系统的硬件资源，对四通道语音增强系统进行总体设计，包括硬件设计和软件编程。硬件设计工作包括修改音频芯片、语音采集麦克风、滤波电路、5V电源电路和音频转换电路的输入接口模式。软件设计部分的工作包括语音增强算法的设计、每个模块在正常运行时的初始化、外部存储空间的扩展和程序的增强。本章后面将更详细地讨论具体设备的研究工作。

**3.3系统的硬件设计**

四通道语音增强系统是专门为语音信号的选择而设计的，它采用了四种全点微配置的统一线路网络。在这个过程中，声音信号被转换成模拟电信号，并传输到转换芯片A/D TLV320AIC23B。

将四个转换通道扩展到语音信号传输的设计是将音频编解码器芯片AIC23作为A/D和D/A转换电路的核心，并引入麦克风和线路。文件是在网上输入的。布线方法包括两个布线输入引脚，每个引脚对应于右侧通道和右侧通道。一般渠道界面的左边和右边线同时使用单向语音习得,并通过试验,两针入口通道对于左和右线可以使用独立并且还提供视音频采集[24]。本文使用两个AIC23芯片，每个芯片的右通道和右通道被认为是音频处理过程。换句话说，麦克风网络收集的四种不同信号中的每一种都有四个A/D转换通道。

系统控制总线使用I2C命令。DSP \_5\_509A包含一个集成的I2C电路模块，因此I2C总线用于为两个AIC23分配不同的设备地址。

DSP 5509A上的两个串行McBSP1和McBSP2缓冲器分别与两个数字AIC23接口相关联，这两个接口构成了DSP和音频芯片的数据采集通道。在数据存储方面，需要为4通道语音数据处理提供更大的存储空间。5509A通过其自己的并行总线接口EMIF将数据发送到磁盘上的SDRAM。当执行语音增强算法时，可以从SDRAM中读取数据进行处理，也可以在算法执行[25]时，预定义在SDRAM空间中指定位置创建的大量中间变量。

本章设计的四通道语音增强系统如图3.3所示。主电路模块包括语音信号采集模块、DSP信号处理模块、A/D转换模块、存储模块和电源模块。

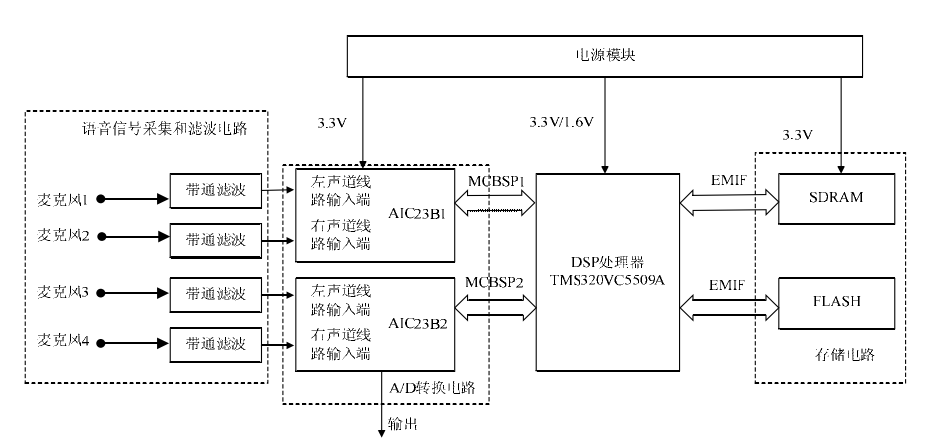


图3.3四通道语音增强系统框图

完整的系统设计完成后，DSP通过引导加载程序将整个系统程序存储在非易失性存储器FLASH中，并加载程序以使程序脱机工作。

**3.3.1语音采集模块**

论文设计的语音增强系统的最前面，即语音信号的阵列收集装置，由4个均匀排列的微构成[26]。那么，如何设置数组元素间距以使数组具有强大的声音选择功能呢？在确定物理传感器的间隔之前，进行了以下理论分析。隔开麦克风的间距是，语音信号的波长为。一般环境中，阵元间距取值较大时，阵列的几何分辨力较好。另外，的取值还要考虑语音信号的频率，通常取信号波长的一半，即：

（4.1）

（4.2）

其中表示声波在空气中的传播速度，一般，是语音信号的频率范围300—3400Hz。由式(4.1)和(4.2)计算可知，，，那么，。综合理论分析和实际语音采集传感器的灵敏度，论文设计阵元间距为10cm。此外，结合电视使用的场景，为使采集阵列具有更好的空间采集性，设置平台高度为70cm。实物结构图如图3.4所示。

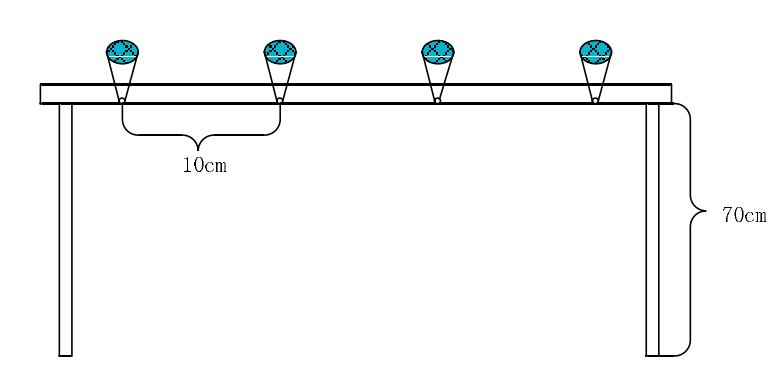


图3.4采集阵列实物结构图

论文使用的麦克风即拾音传感器，是KZ—501D迷你小型拾音器。KZ—501D是新型高灵敏度拾音器件，适合正常范围内的声音采集，采集范围是20-30m2。

图3.5为KZ—501D拾音器。图中，红线、黑线、绿线分别代表+5 V电源线、公共地线和信号线。



图3.5 KZ-501D拾音器

Kz-501d搓纸设备具有高灵敏度、低噪音、低失真、小尺寸、大工作电压范围、内置电源反向保护电路和低功耗等优点。性能参数包括:

(1)工作电压:3-5v；

(2)工作电流:5毫安-60毫安；

(3)频率响应:10Hz-20 khz；

(4)输出电压:500-600mv。

通常，麦克风收集的语音信号很弱，输出信号宽度在2mV左右浮动。但是，根据上述修正器的操作参数，输出电压约为500-600mV，kz-501d麦克风具有信号放大模块[27]。这样，从麦克风阵列收集的语音信号将直接发送到芯片AIC23，而不影响转换精度，从而实现音频编码。此外，设计5 V直流电源，为拾音器供电。

**3.3.2滤波放大电路**

该系统使用四个麦克风在真实环境中接收语音信号。一般来说，语音信号的频率范围在300到3400赫兹之间。考虑到语音中各种噪声和干扰的混合，设计了一个简单的人工带宽滤波器电路，用四个麦克风分别对混合语音进行预处理，消除了部分干扰和噪声。然后将其转发到AIC23音频芯片中。本章系统地研究了一种由一阶低压RC电路和一阶高压RC电路组成的手动RC带滤波器电路。具体电路示意图如图3.6所示。

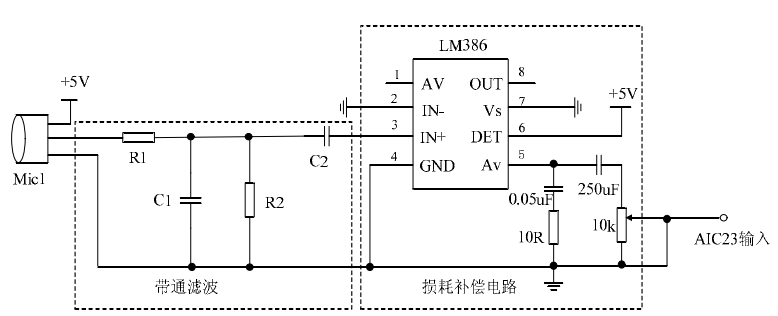


图3.6滤波放大电路

这种过滤电路的优点是简单可靠，不需要直流电源。然而，带内信号能量的缺点是能量损失，其振幅约为输入信号的一半。因此，在滤波电路之后，本文设计了一种信号放大电路，对处理后的滤波信号进行放大。

**3.3.3采样频率和采样精度**

需要改进300—3400Hz频率范围的语音信号的论文设计的语音增强系统。收集的混合信号用作信号处理的基准，需要收集的信号频率范围应为目标信号频率的上下2倍。也就是说，收集的混合信号的频率范围在150Hz—6.8KHz的频率范围内。因此，选择适当的r，c参数，获得通过频带频率计算过程。

（3.3）

（3.4）

上式中，、分别指带通滤波器的上·下限截止频率。

图4.7中，，，，。

计算得：





过滤后，收集信号有能量损失，因此设计了简单放大器电路来补偿收集信号。放大器电路是使用芯片的LMS386o LM386的内置电压增益为20，外部电子组件较少的音频放大器。LMS386的不准确程度较低，工作电压范围为4-12V。LMS386内部部件结构如图3.7所示。

该电路设计内部电压增益20、针脚3处的信号输入、通过内部放大器的针脚5处的信号输出、10K电位器调节输出信号振幅。电路工作电压为5 V。

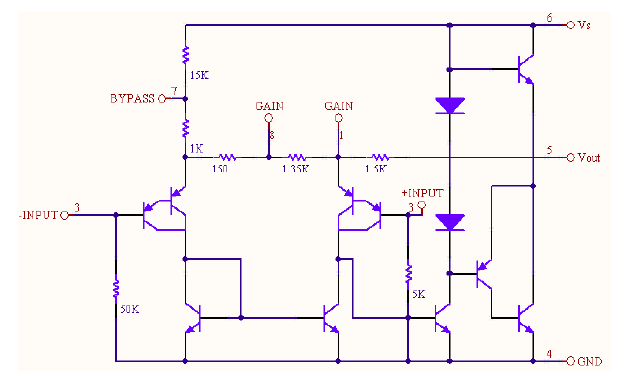


图3.7 LMS386内部件结构图

因此，根据奈奎斯特采样定理，最小采样频率应为13.6KHz。本论文将AIC23音频转换器的采样频率设置为16KHz，以获得良好的抗混音效果。

此系统的重点是麦克风将语音信号转换为电压信号的大小精度。搓纸传感器的输出电压范围为0-0-0.6v(0-600mv)。假设

预估精度为100uV。也就是说:



（1）期望精度为100uV，需要读数为n



上式中，是待测语音信号的最大动态范围，是系统可检测到的语音信号的最小电压信号值。

（2）由上式可知，在动态范围0—600mV内需要读6000个读数，即每个读数为：



由于论文设计扩展音频芯片AIC23通路，使得一片AIC23处理两路数据，这就造成最高精度减半，智能设置AIC23的采样精度为16位。



因此，论文系统及时考虑动态变化的遗失数据和电器干扰，设置AIC23为16位A/D也是足够使用的。

**3.3.4 5V电源电路**

本论文采用LB10—10B05电源模块，设计了为滤波器和搓纸传感器提供5 V工作电压的5 V电源电路。

LB10—10B05是Mornsun生产的高效环保电源模块。输入电压大，直流和交流集成在一起，效率高，可靠性高，损耗低，安全隔离等。通用输入信号为100-240 vac，50/60Hz。5 V电源电路图如图4.8所示。

在图3.8中，NTC是5 D-9热敏电阻，C1是输出滤波器高频铝电解电容器。C1=330uF。C2是高频陶瓷电容器或聚冷电容，C2=0.1uF/50V。Fuse是输入保险丝。

在电路中选择合适的保险丝必须由安全机制保证，这是非常必要的。保险丝额定值太高时，会发生错误[37]。级别太低时，错误的保险丝断了。适当的保险丝额定值是系统保护和电源保护的重要组成部分。以下公式计算了相应的熔断器额定值:

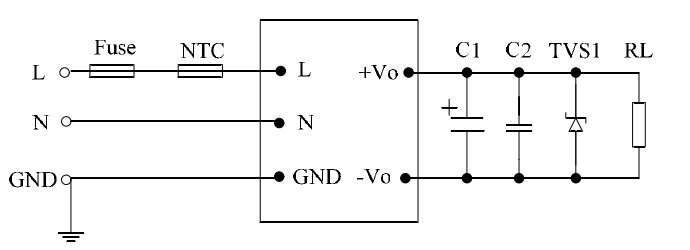


图3.8 SV电源电路图

（3.5）

 （3.6）

 （3.7）

  （3.8）

   （3.9）

根据LBxx—10xx SERIES数据手册中，LB 10-1OB05的相关参数如表3.1所示。

表3.1LBxx—10xx 特性表

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 型号 | Output（Vol/Iol） | Power | Efficiency（%） | Ropple and Noise |
| LBxx—10xx | 5V/2000mA | 10W | 73 | 50mV |

**3.3.5音频转换电路**

论文的模数转换电路采用TLU320AIC23B芯片(以下简称“AIC23”)，专为语音信号采集和广播的音频处理而设计。aic23的软件控制和音频输入输出通过DSP 5509A的McBSP接口。

iczm 23b的内部方案如图3.9所示。对于4路信号的传输，每个AIC23控制一个2路信号。基本原则如下。图3.9看看,发现TLU320AIC23B本身含有的双输入信道的每一行的左边和右边,而每个通道的左翼和右翼的包含额外的监管,监管的消声器,计算机和A / D转换器。在设计过程中，设计了两个模块化的音频转换电路(由aic 23组成)，分别将左右通道的输入分别处理为两个单通道，从而提供了四通道音频处理功能。

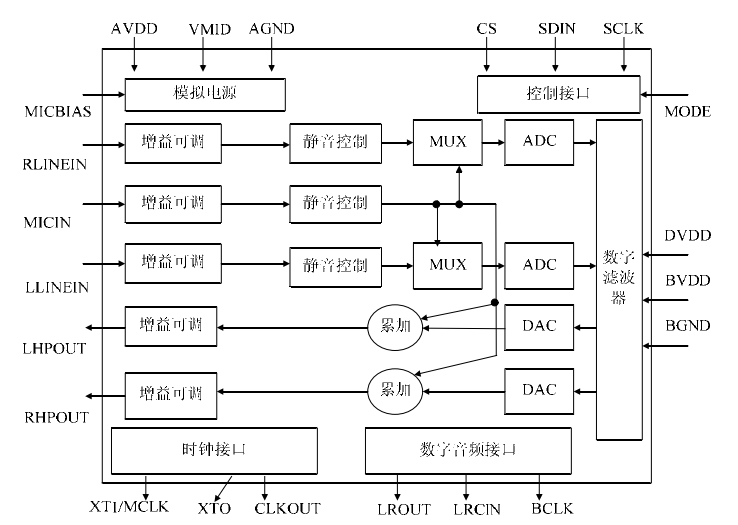


图3.9 AIC23内音区结构图

在本文使用的techv-5509a版本中，集成了AIC23音频，并添加了音频转换电路来接收四个语音信号。本白皮书使用一个简单的操作系统，通过5509A中包含的I2C模块的I2C总线，将控制引脚直接连接到模块电路的AIC23芯片上。添加到此设计中的音频转换电路的示意图设计位于编解码器中的Protel Prote199SEo软件下(见图3.10和3.11)。使用Sch方案和编解码器方案封装PCB组件。

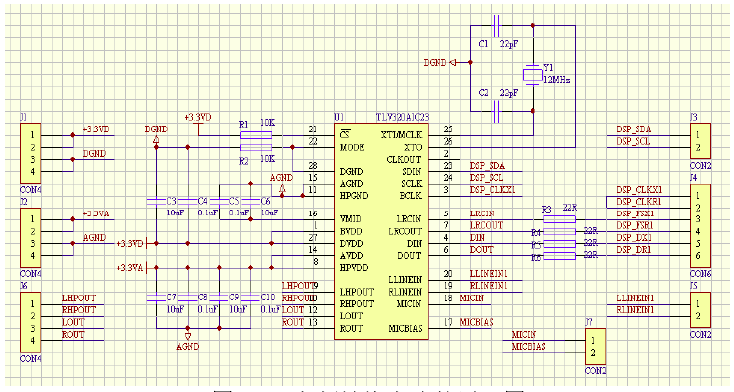


图3.10音频转换电路的原理图

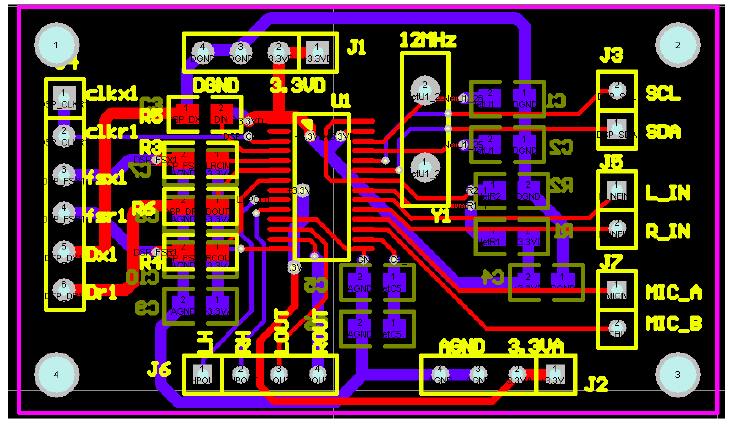


图3.11音频转换电路的元件封装图

**3.3.6 MCBSP多通道缓冲串口**

DSP 5509A为串行McBSPO、McBSP1和McBSP2提供三个多通道缓冲。每个McBSP都有7个引脚连接到外部设备。负责DX数据传输。DR负责接收数据。CLKX是调度时钟的指针。CLKR是接收时钟的指针。FSX是传输帧的同步引脚。自救者是接收帧的同步针。设计了一种音频转换电路，将McBSPl接口作为双向数据通道。如果McBSP1是aic23的欺骗设备，aic23接收同步信号并执行同步传输。AIC23与数字系统的公共接口模式为I2S模式和DSP模式。本白皮书使用DSP模式使McBSP1兼容，并与aic23完全连接。

McBSP的示意图如图4.12所示。当DSP向AIC23发送数据时，处理器首先将数据写入传输寄存器(控制寄存器DXR1、DXR2、(XRDY)位SPCR2)。如果数据长度不超过16位，则仅使用DXR 1。否则，使用DXR 1和DXR2。然后将XSR寄存器输出压缩并传输到DX引脚。DSP在读取aic23时接收数据的方式与传输时几乎相同，但是接收日志(RB R)被添加到移位日志(RSR)的接收和解压过程中，然后进行记录。不同的(rra)发送DSP。与数据传输过程相比，数据接收是一个三步缓冲区。

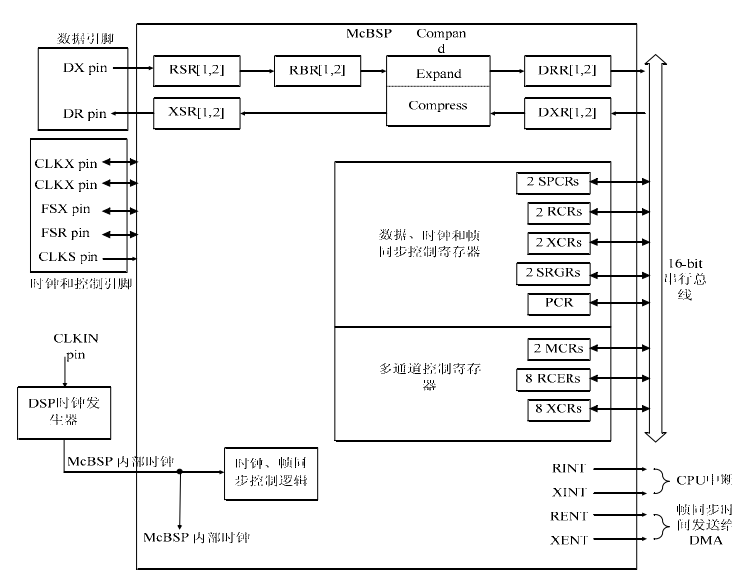


图3.12 McBSP结构框图

Speech is the most basic and fastest way of communication in human life.However, when voice communication is performed in the environment where you actually live, the presence of noise in the environment can interfere with the voice.In this context, noise refers to all interfering signals except for useful speech signals [1].There are many types of noise in the real world, such as indoor reverberation, outdoor noise, the complexity of the people around you, electromagnetic noise in the lab, and so on.This is because in the process of using communication and voice communication methods, people will be seriously affected by a variety of noise signals.Then, the speech enhancement function is to process the loud speech and get the purest speech signal.At present, speech enhancement technology is widely used in video conference, military communication, eavesdropping technology and speech recognition technology, which actually improves the clarity of speech video conference content, improves the communication quality of speech equipment, and solves the problem of speech noise pollution.Used in practical situations.It has played an important role in practicality and merit of study.A microphone array

Speech is the most basic and fastest way of communication in human life.However, when voice communication is performed in the environment where you actually live, the presence of noise in the environment can interfere with the voice.In this context, noise refers to all interfering signals except for useful speech signals [1].There are many types of noise in the real world, such as indoor reverberation, outdoor noise, the complexity of the people around you, electromagnetic noise in the lab, and so on.This is because in the process of using communication and voice communication methods, people will be seriously affected by a variety of noise signals.Then, the speech enhancement function is to process the loud speech and get the purest speech signal.At present, speech enhancement technology is widely used in video conference, military communication, eavesdropping technology and speech recognition technology, which actually improves the clarity of speech video conference content, improves the communication quality of speech equipment, and solves the problem of speech noise pollution.Used in practical situations.It has played an important role in practicality and merit of study.A microphone array Speech is the most basic and fastest way of communication in human life.However, when voice communication is performed in the environment where you actually live, the presence of noise in the environment can interfere with the voice.In this context, noise refers to all interfering signals except for useful speech signals [1].There are many types of noise in the real world, such as indoor reverberation, outdoor noise, the complexity of the people around you, electromagnetic noise in the lab, and so on.This is because in the process of using communication and voice communication methods, people will be seriously affected by a variety of noise signals.Then, the speech enhancement function is to process the loud speech and get the purest speech signal.At present, speech enhancement technology is widely used in video conference, military communication, eavesdropping technology and speech recognition technology, which actually improves the clarity of speech video conference content, improves the communication quality of speech equipment, and solves the problem of speech noise pollution.Used in practical situations.It has played an important role in practicality and merit of study.A microphone array Speech is the most basic and fastest way of communication in human life.However, when voice communication is performed in the environment where you actually live, the presence of noise in the environment can interfere with the voice.In this context, noise refers to all interfering signals except for useful speech signals [1].There are many types of noise in the real world, such as indoor reverberation, outdoor noise, the complexity of the people around you, electromagnetic noise in the lab, and so on.This is because in the process of using communication and voice communication methods, people will be seriously affected by a variety of noise signals.Then, the speech enhancement function is to process the loud speech and get the purest speech signal.At present, speech enhancement technology is widely used in video conference, military communication, eavesdropping technology and speech recognition technology, which actually improves the clarity of speech video conference content, improves the

Sp**第四章 功能模块的设计与实现**

**4.1软件开发流程**

利用数字信号处理器TMS320VC5509A设计的语音增强系统，可以选择最有效的语音增强算法，包括噪声抑制性能、有限的精度范围和算法的复杂性。因此，在确定论文算法之前，MATLAB必须对不同的自适应算法进行多次仿真。

本白皮书讨论了软件可移植性的问题，通常用C语言开发大多数部分。需要执行复杂任务或代码执行效率高的软件模块应尽可能用汇编语言编写。用户可以使用c语言和汇编语言通过c编译器和汇编程序转换源文件(.c和.asm)和目标文件(.obj)，然后通过链接在DSP5509A上创建可执行对象。用于模拟和调试的代码。

在调试后，使用十六进制代码转换工具(.hex)调试后，执行可以在FLASH (.out)中执行的代码，并对FLASH代码进行改进，使DSP目标系统能够在计算机上独立运行。论文系统的软件开发方案如图4.1所示。

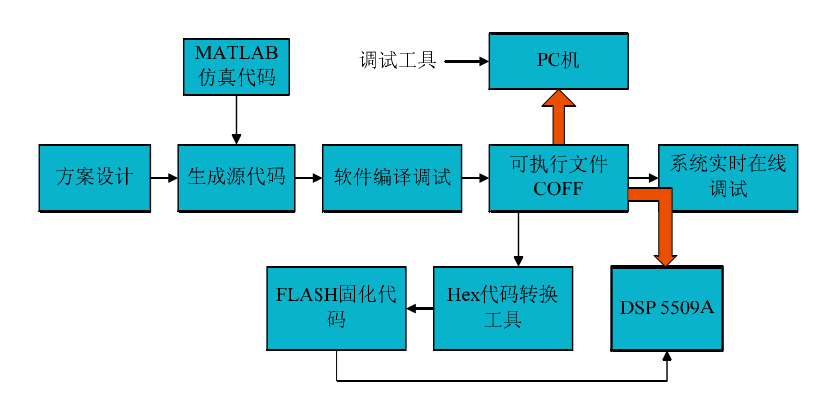


图4.1系统的软件开发流程图

**4.2系统程序初始化设置**

本节包括登记册的参数定义内部交际TLV320AIC23B系列芯片的音频、I2C模块的初始化l’aic23来读取数据,启动配置两个缓冲器为McBSP1和多渠道McBSP2 (McBSP1设置和McBSP2大体上是一样的)收入了。通过初始化外部存储的SDRAM、中断和定时器等参数，为数据接收和语音噪声抑制算法处理做好系统准备。

**4.2.1音频编解码TLV320AIC23B的初始化设置**

本文使用2个TLV320AIC23B作为楼梯设备，DSP 55009a作为主要的I2C总线设备，11个功能寄存器的配置被编程到2个AIC23B音频芯片中。为了将系统扩展到4通道语音处理系统，本文主要讨论立体声输入。aic23b的右通道和右通道的指针是Lin和llinein。aic23b的每个通道都有4个通道的输入，这两个通道被认为是一个用于音频处理的通道。换句话说，输入方法必须选择线路输入，在模拟音频通道控制寄存器中选择模拟输入必须将ISEL设置为零。表4.2显示了TLU320AIC23的地址和数据格式。

表4.2TLU320AIC23的地址和数据格式

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| TLU320AIC23内部从高到低15位 | | | | | | | | | | | | | | | |
| D15 | D14 | D13 | D12 | D11 | D10 | D9 | D8 | D7 | D6 | D5 | D4 | D3 | D2 | D1 | D0 |
| 寄存器地址 | | | | | | | | 寄存器内容 | | | | | | | |

每个aic23b注册中心的地址是唯一的，因此在填写地址时，每个注册中心必须输入每个地址。此外，要更改注册表的状态，必须首先替换重新初始化的注册表(注册表0的值)，然后重写所有相关的注册表。否则它就不能正常工作。根据TLU320AIC23接口的时间表，本白皮书将TLU320AIC23密钥寄存器委托给系统需求。系统由DSP5509A的I2C接口控制，两个AIC23在基本模式下运行，数据接口选择DSP模式，传输字符长度设置为16位。音频输入输出音量增益为0dB。USB模式下的采样率，可选的12MHz晶体振动控制，控制寄存器将采样率设置为8KHz。图4.2显示了11个aic23寄存器的详细系统配置值。

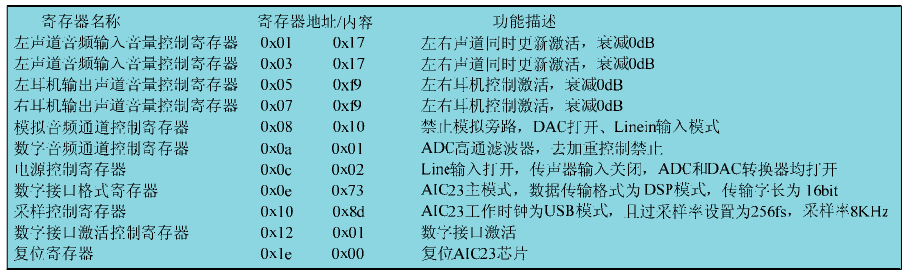


图4.2 TLU320AIC23各寄存器的详细配置

**4.2.2DSP的I2C模块初始化设置**

使用AIC配置AIC23寄存器的DSP系统55009应在AIC23配置之前在DSP上初始化I2C端口。I2C模块:7位转起,8位数据传输站(长度n是1比8的数字数据的范围由计数器I2CMDR BC)、数字闭环模式是开放和自由的方式是开放的。I2C模块的第二个频率划分产生了一个12 MHz的模块化时钟和一个100 KHz的主时钟。图4.3显示了I2C模块的主寄存器和系统初始化。

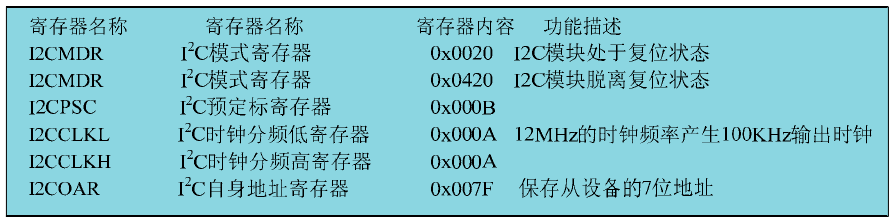


图4.3 I2C模块寄存器及初始化设置

系统时钟为144 MHz, I2C总线时钟频率为系统时钟的第二个频率。I2CCLKL和I2CCLKH寄存器都设置为10，用于12MHz的I2C模块化时钟，以便在I2C总线上获得100KHz的SCL时钟频率。I2C模块的内部时钟示意图如图4.4所示。

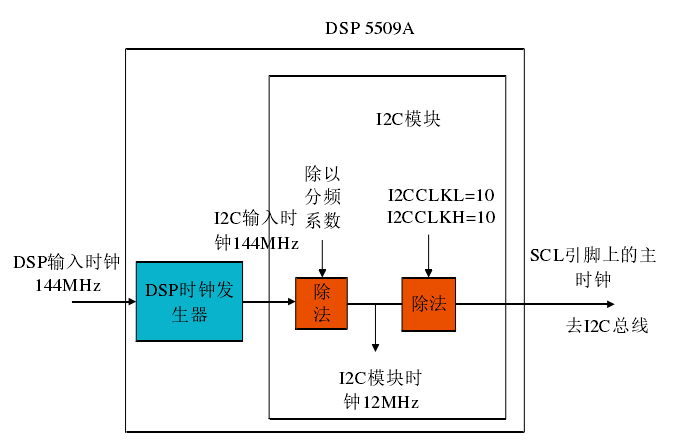


图4.4 I2C模块的内部时钟图

**4.2.3多通道缓冲串口McBSP的配置**

系统分别控制添加到McBSP1和McBSP2的代码模块，以及主开发板中包含的AIC23芯片。aic23的数字音频接口设计为与DSP模式兼容，与DSP的McBSP1和McBSP2接口兼容，确保无缝连接。McBSP1和McBSP2都是aic23b的斯拉夫设备，它们接收来自aic23的同步信号，以便进行同步传输。

本文所述的四通道语音增强系统将左右通道视为音频通道，而aic23则收集双向语音数据。当DSP5509A与aic23通信时，从音频芯片左侧通道收集的数据传输到McBSP的DRR2接收寄存器，从右侧通道收集的数据传输到DRR 1。纸上的DRR 1和DRR2是相互独立的，数据存储在不同的麦克风中。由于两个rra寄存器的最大存储容量为16位，因此aic23数据的长度应限制在16位或更高，以避免DRR1和DRR2寄存器接收到的数据溢出。

McBSP的初始化主要通过在内部密钥寄存器(spcr1、spcr2、xcr1、xcr2、rcr1、rcr2、PCR等)中输入相应的命令字，确保时钟和AIC23帧的同步信号。数据的接收和传输应在监督下进行。McBSP内部密钥寄存器中某些特殊功能位的配置及其功能描述如图4.5所示。

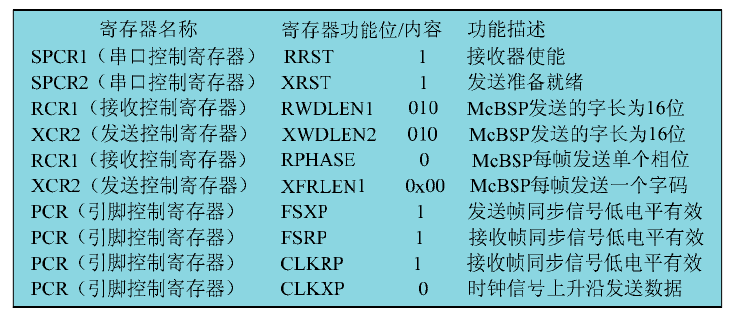


图4.5 McBSP关键寄存器和特殊功能位

McBSP1、McBSP2和AIC23B之间的数据传输在单个图像中为16位，没有延迟。数据传输应由信号或耦合信号的下降触发。图4.18为DSP模式下aic23b的串行时间图。在这种模式下，系统中两个AIC23的左通道数据首先发送数据字，然后发送右通道数据。

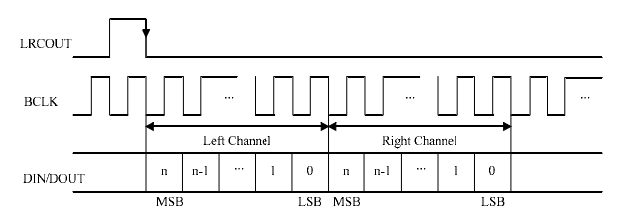


图4.6 AIC23B DSP模式串口时序图

**4.2.4 SDRAM初始化设置**

本白皮书使用CE2和CE3外部存储作为SDRAM。首先，将外部总线选择控制寄存器EBSR设置为1。换句话说，请确保SDRAM在完全EMIF模式下运行。要使CLKMEM频率等于SDRAM工作频率的一半和CPU时钟频率的一半，请将全局控制寄存器EGCR中的MEMFREQ字段设置为001b [38]。设置CE2的MTYPE ＝ 11b，CE3空间控制寄存器CE2\_1和CE3\_1为16位SDRAM。设置SDRAM控制寄存器和初始化寄存器，等待6个CPU时钟周期，然后EMIF接口自动初始化SDRAM。图4.7显示了与SDRAM相关的寄存器的详细配置，这些寄存器扩展了外部存储空间。

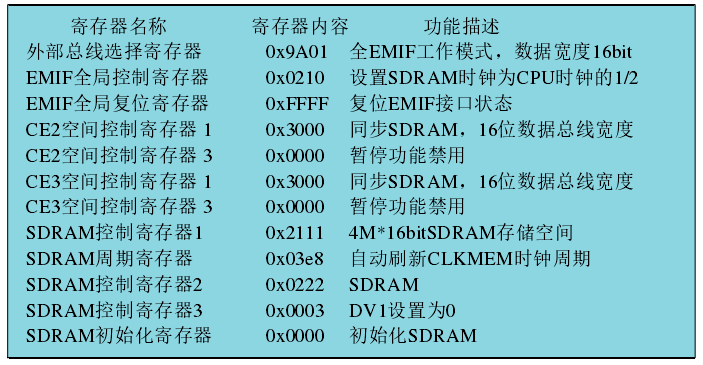


图4.7 SDRAM寄存器配置详情

**4.2.5定时器原理和初始化**

TMS320VC55909A具有相同的20位软件可编程定时器Timer0和Timer1。这会产生一个周期性的中断信号，该信号被发送到DSP外部的设备。常规计时器包括2个计数器和2个周期寄存器（PSC:前计数器，4位）和主计数器（TIM，16位），调度标准周期寄存器（TDDR: 4位）和主周期寄存器（PRD:perion）

本文仅激活计时器时间r0。这里包括的保留标准寄存器PRSCO是一个16位寄存器，周期寄存器TDDR和计数寄存器PSC分别是PRSCO的9-6位字段和3-0位字段。初始化计时器时，周期寄存器TDDR和PRD的内容被复制到计数寄存器（PSC和TIM0）。当PSC自动减小到0时，TIM0开始减小，而当减小到0时，定时器发送一个中断请求。中断信号的频率计算公式为例:4.10。

（4.10）

论文中，DSP5509A的CPU时钟为144MHz，准备在TIN/TOUT针脚上生成8KHz的时钟。计时器注释/输出端的级别每次计数器减小到0时转换一次。因此，计时器的中断信号TINTO的频率是TIN/TOUT针脚上时钟频率的两倍，因此中断信号TINTO频率=4MHz。将TINTO频率和CPU频率替换为4.10:



所以



因此，设置周期寄存器为18000 CPRD=17999，预定标计数器的值为1（TDDR=0）。图4.8为定时器结构图。

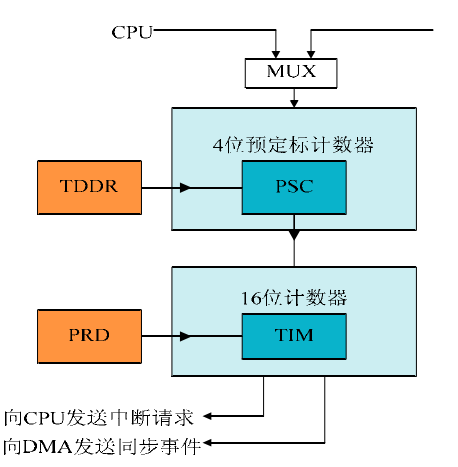


图4.8定时器结构图

初始化与计时器关联的寄存器。首先，设置定时器控制寄存器TCR段TSS=1停止计时，TLB=1定时器自动重装，FUNC=O 1时间源选择内部CPU时钟。通过将记录在分段TDDR中的数字记录在寄存器PRSC中以指示周期数，运行时TDDR的内容将直接传递到保留标记计数器PSC。最后，开始计时TSS清洁。

**4.2.6中断处理流程**

TMS320VC5509A支持32个支持的中断服务ISR。掩码中断是常用的。如果您的程序需要在CPU上触发屏蔽的中断，则需要在IFRO和IFR 1中断标志寄存器中找到对应的中断标志以指定1。该标志设置为1以创建一个中断。否则，将不处理中断。有关5509A可中断中断寄存器和位，请参见表4.3。

表4.3描述可屏蔽中断的寄存器和位

|  |  |
| --- | --- |
| 寄存器位 | 功能描述 |
| IVPD | 中断指针向量寄存器 |
| IVPH | 终端指针向量寄存器 |
| IFR0,IFR1 | 终端标志寄存器，包含所有可屏蔽中断的标志位 |
| IER0,IER1 | 中断使能寄存器 |
| DBIER0.DBIER1 | 调试终端使能寄存器，定义时间临界 |
| INTM | 中断模式位 |

本文在编程中使用定时器0中断处理可屏蔽定时器0中断的基本过程分为3个步骤。

（1初始化相应的中断寄存器

程序空间点的16位中断向量寄存器IVPD和IVP为32个中断向量。寄存器IVPD管理0-15优先级和24-31优先级向量，而寄存器IVP管理16-23优先级中断向量。要打开实时操作系统中断，请将中断向量指针寄存器IVPD设置为Ox00D0。检查中断功能寄存器IERO是否设置为0x0010。也就是说，Timer0的中断标志位设置为1，其余中断设置为0，而未打开。然后使用与Timer0对应的中断标志位1将DBIERO值设置为0x0010，并调试中断标志寄存器IFRO中的中断。最后，打开全局中断，CPU响应该中断。

（2中断服务编程

编程关键字是interrupt。使用此关键字来声明中断函数的名称。中断服务功能的名称可以是任意的，但与第三步中的中断方向指示符的名称相对应。

（3）中断向量表的创建

在本文中，TINTO的中断函数的名称为Timer0，并且中断向量表文件将中断向量声明为:

.ref c\_ into,\_ timer0;

.sect"vectors";

RS: .ivec\_ c\_ int0;

T0:.ivec\_timer0;

由于数据收集量很大，因此必须在纸张上设置定时器中断，因此可能会在设置的时间内生成一定数量的数据收集以进行处理。本文使用多通道缓冲区串行McBspl和McBsp2收集四个语音信号，包括对每个寄存器的详细介绍以及由寄存器中特定位字段表示的中断服务。

**4.3DSP的实现**

本节主要介绍自适应滤波算法在DSP上的加载实现和DSP存储空间的分配。

**4.3.1自适应算法的实现**

DSP算法处理任务主要分为:时域任务和频域任务。时域运算是数字滤波器，例如有限长度冷杉滤波器和无限长IIR滤波器，而频域运算主要是傅立叶变换。使用DSP处理计算任务的特征是:数据量大，相同的数据可重复使用多次，生成大量的中间数据以及生成大量的地址生成和乘法运算。本文中的自适应语音增强算法是一种时域算法，在DSP中需要引起关注。

（1）该软件程序是在CCS3.3环境中通过调试执行的，文件之间的函数引用不需要函数声明，可以进行编译和链接，但是在某些情况下，处理结果与预期结果之间会有差异。为了避免编程，已预先声明严格遵守c的标准语法和跨文件函数[39]。而且，软件设计与硬件系统一致。注意数组名称，数据类型和长度的一致性。在编程过程中，无法确定正确的数据类型，因此经常会发生数据泄漏并且计算结果不正确。乘以16位寄存器，需要在将数据加载到寄存器之前根据计算出的数据长度确定数据类型，以避免更多的溢出。表4.4中列出了用于纸张的特定数据类型和长度。

表4.4数据类型和长度

|  |  |  |
| --- | --- | --- |
| 数据类型 | 长度（字节） | 值域 |
| char | 1 | ASCII码字符 |
| unsigned char | 1 | -128-127 |
| unsigned short int | 2 | 0-65535 |
| shout int/signed int | 2 | -32768-32768 |
| float | 4 | 约精确到小数点后6位数 |
| double | 4 | 约精确到小数点后15位数 |

（2）DSP 5509A是16位定点DSP，其动态范围受到限制，因此在使用动态范围时要小心以避免溢出。在本文中，我们考虑了语音增强算法，可以将其理解为可改变加权因子的各种滤波器，应对滤波器算法溢出的最佳方法是设置适当的滤波器步长，并使滤波器权重向量小于1。

（3）如有必要，与汇编语言和硬件相关的特殊指令将直接应用于C程序，以简化代码和计算复杂性。

**4.3.2存储空间分配**

如果不使用DSP5509A磁盘存储空间，则可以分配给它自己。它设计了一个连接配置文件CMD，用于管理和分配存储系统中的所有物理寄存器和地址空间，并为用户界面生成一个CMD文件。文档主要用于在PAGEO和PAGE 1中分配程序存储和数据存储。CMD文件由三个部分组成:输入/输出定义、内存定义和节定义。内存命令用于创建一个目标存储模型，该模型定义存储空间的位置和长度。section命令(通常称为段)指定每个段的位置来分配地址空间，即内存设置的存储模型。

在将文档分配给DSP5509A部分时，c代码和设计数据应分为系统定义部分和用户定义部分。

(1)系统定义的段。初始化部分主要包含常量(.coastline)、初始常量和可执行代码。非初始化段为用于创建和存储变量(例如全局变量和静态变量)、堆栈和动态存储空间的程序保留存储空间。

(2)伪命令pragram用于将处理函数的自定义段通知编译器。高端盘TMS320VC5509A只包含128K的RAM和16位\* \* 16兆比特的四车道ROM.Les话音信号收集纸质导致搜集大量的数据处理系统的语音和扩建三年级用SDRAM等数据存储磁盘和ec3 8米\* 16位。DSP的斜坡内部更快的阅读和写作是坡道SDRAM,这样的外部数据缓冲打开系统来储存数据的输入和输出上面用个性化的部分的语法,# pragma DATA\_SECTION空间中定义,SDRAM的外部数据。在程序存储空间的第0页，数据占用尽可能多的连续物理存储空间，避免了空间浪费[40]。DSP的内部斜坡有系统定义的段、可执行的c代码和自适应滤波器权重数据变量的优先级，这些数据变量通常在编程过程中编写和编写，以确保系统的运行速度。图4.9显示了本文所述的四向语音增强系统中的一些内存分配。

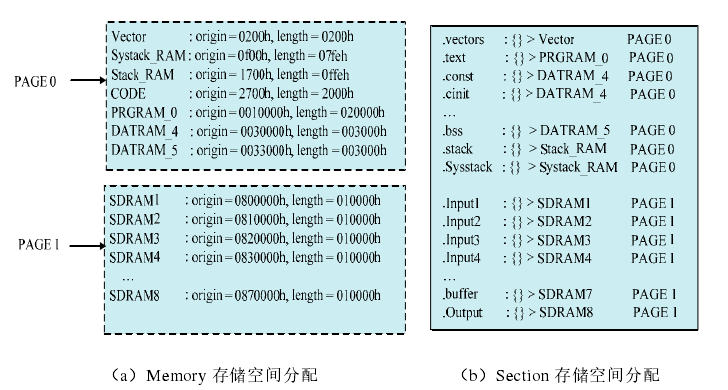


图4.9存储空间的分配

**4.4DSP VC5509A程序自举**

在CCS3.3开发环境中，可以使用基于纸张的语音增强系统的预测测试和调试，通过JTAG电缆与PC进行DSP通信。整个系统的软件已经完成。为了完全脱离PC的CCS开发环境，您需要将DSP系统连接到它自己的启动功能，即模拟器，打开电源，然后DSP启动并加载程序来完成软件设计功能。

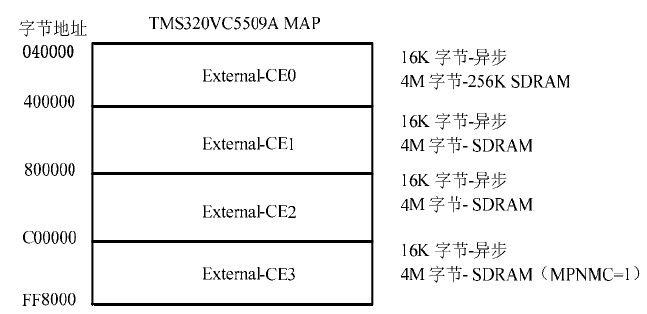


图4.22DSP片外存储空间映射图

数字信号处理器TMS320VC5509A} 42]由于用户没有可用的闪存用于内部存储，并且整个DSP系统开发的最终目标是在模拟器上运行，因此DSP需要额外的非易失性存储器来存储程序代码。需要。将SST的产品产生的SST2\_SVF040B4Mbit SPI FLASH用作系统的外部程序存储器，并通过5509A MCB SPO连接完成SPI协议通信。其中，DSP主设备和FLASH是从设备。

**4.4.1自举引导模式的配置**

TMS320VC5509A是基于RAM的设备。非易失性存储器的读写速度慢。接通系统电源后，内部ROM中的引导加载程序将存储在闪存中的代码移动到DSP内部RAM或DSP \_5\_509A自由烧写的扩展空间SDRAM中的相应地址。指针移动到主程序门户并自动启动程序以完成启动。

5509A支持多种引导模式，例如主机接口HPI引导方法，8/16位并行EMIF引导方法，8/16位标准串行引导方法和USB接口引导方法。 5509A的启动模式使用4引脚BOOT配置完成。 BOOT引脚分别连接到GPIO1,2,3,0。根据系统要求，本文的硬件系统将DSP设置为使用McBSP0串行24位地址SPI引导模式。如表4.5所示，将Boot [0333693]设置为0001。

表4.5 BOOT引导方式

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| BOOT资源 | BOOT与GPIO对应引脚 | | | |
| BOOT0 | BOOT1 | BOOT2 | BOOT3 |
| 串行24位地址 | GPIO0 | GPIO1 | GPIO2 | GPIO3 |
| EEPROM引导方式 | 0 | 0 | 0 | 1 |

**4.4.2芯片简介及常用命令**

SST2VF040B是由SST生成的4mbit（512k \* 8位）SPI串行闪存。时钟速度为20MHz，电源电压为2.7V-3.6V，可重复记录，访问时间快，数据保存时间长。通过命令读/写清除配置SST2\_5 VF040B可以达到内部操作的目的，以自动完成命令序列。表4.6列出了SST2\_SVF040B的完整命令格式。

WREN（雷恩）在写入Flash之前先激活Flash，然后发出芯片级擦除命令60H。多亏了芯片，您必须降低CE＃引脚。然后，在对所需的地址字节进行编程之后，执行wrdi（写保护）04h指令以退出AAI的AAI指令AFH [41]。最后，执行字节编程指令02H，写意图（WREN）指令06H，状态寄存器（EWSR）指令O1H等，以完成将数据写入FLASH的操作。然后，通过发出读取ID号命令90H，读取命令03H，读取状态寄存器（RDSR）命令05H等，完成了从闪存向DSP读取数据的RAM执行。

表4.6 SST25VF040B的全部指令格式

|  |  |  |
| --- | --- | --- |
| 指令操作 | 指令格式 | 说明 |
| Read | 03H | 从Flash读数据 |
| Sector-Erase | 20H | 整段擦除 |
| Block\_Erase | 52H | 扇区擦除 |
| Chip-Erase | 60H | 全片擦除 |
| Byte-Program | 02H | 字节编程 |
| Auto Address Increment Program | AFH | 自动增量地址的程序 |
| Read-Status-Register | 05H | 读状态寄存器 |
| Enable-Write-Status-Register | 50H | 使能写状态寄存器 |
| Write-Status-Register | 01H | 写状态寄存器 |
| Write-Enable | 06H | 写使能 |
| Write-Disable | 04H | 禁止向Flash写数据 |
| Read0-ID | 90H or ABH | 读厂商和器件编号 |

**4.4.3DSP SPI程序引导烧写步骤**

论文程序最终以SPI FLASH编写在硬件系统上，目的是实现离线操作。默认操作过程为:首先准备两个程序文件。一种是本文的语音增强算法程序。另一个是SPI FLASH读写程序flash\_burn.pjt，该程序将voice\_enhanced.c的.Out文件转换为CCS识别的.dat文件，然后转换为flash\_ burn程序。pjt将算法程序.dat写入外部FLASH。上电后，5509A会自动将FLASH的代码移至DSP芯片以执行。详细步骤如下所述。

（1）使用CCS编译最终程序voice\_enhancement.pit以生成.out格式文件。

（2）将.out文件转换为.hex文件。

使用文本编译器创建一个.cmd文件。该文件的内容如下，并将其另存为hex.cmd文件。 hex.cmd的文件内容如下。

-启动;如何编写启动文件

-v5510:2;创建C55xx系列DSP的启动文件格式

-serial8;使用串行SPI程序引导模式

-e\_c\_int00;计划参与地址

-reg\_config Ox1c00;0x2490;将0x2490写入Ox1c00累加器

-delay\_count 3000;延迟3000个CPU时钟周期

-b;用于生成.bin文件

-o voice\_ Enhance.bin;输出.bin文件;输入.out文件

-map voice\_ Enhance.map;创建.map文件

（3）首先，将输出文件hex.cmd文件和转换工具Hex55.exe放在同一目录中。然后在DOS环境中指定文件路径。最后，从该路径运行hex55 hex.cmd命令以生成所需的.bin文件。

（4）通过读写将.bin文件写入SPI FLASH。

（5）成功编写后，系统无需仿真器和软件仿真环境即可实现自己的启动，并在完成软件功能后完成整个DSP系统设计。

Speech is the most basic and fastest way of communication in human life.However, when voice communication is performed in the environment where you actually live, the presence of noise in the environment can interfere with the voice.In this context, noise refers to all interfering signals except for useful speech signals [1].There are many types of noise in the real world, such as indoor reverberation, outdoor noise, the complexity of the people around you, electromagnetic noise in the lab, and so on.This is because in the process of using communication and voice communication methods, people will be seriously affected by a variety of noise signals.Then, the speech enhancement function is to process the loud speech and get the purest speech signal.At present, speech enhancement technology is widely used in video conference, military communication, eavesdropping technology and speech recognition technology, which actually improves the clarity of speech video conference content, improves the communication quality of speech equipment, and solves the problem of speech noise pollution.Used in practical situations.It has played an important role in practicality and merit of study.A microphone array consists of two or more microphones, depending on the specific spatial shape.Recently, the use of a single microphone to improve speech quality has yielded very good results [2].However, in real-world environments and applications, such as when the speaker is walking in a small part of the room, many noise signals are mixed into the speech signals collected by a single microphone due to the location of the source and the multipath reflection.In order to deal with speech confusion in these situations, researchers are using multiple microconfigured microphone arrays for voice communication to compensate for the disadvantage of a single microphone.Track the location and effectively suppress interference and environmental noise.With the rapid development of cloud computing and big data, artificial intelligence is gradually on the rise. Apple's Siri voice control function and Amazon ECHO smart speaker also appear. Intelligent voice has become a widely used application worldwide.At the same time, intelligent voice products such as the setting speaker jointly released by jingdong and Hong Kong university of science and technology, rhinoceros voice assistant and voice cloud products have emerged successively, which were developed by China intelligent voice technology company [3].Siri is a short-range voice and ECHO is a long-distance voice application.There are three main ways smart devices interact with people: touch, gesture and voice.Voice interaction, more natural, more reasonable functionality, and fewer tasks in the home scene have made voice an increasingly mainstream method of interaction.Short-range speech varies according to the terminal used, while long-range speech provides more openness and convenience.Smart LCD TV has been widely used.Smart TV has the huge advantage of high frequency and visual screen.It has a unique advantage in developing user habits. Speech is the most basic and fastest way of communication in human life.However, when voice communication is performed in the environment where you actually live, the presence of noise in the environment can interfere with the voice.In this context, noise refers to all interfering signals except for useful speech signals [1].There are many types of noise in the real world, such as indoor reverberation, outdoor noise, the complexity of the people around you, electromagnetic noise in the lab, and so on.This is because in the process of using communication and voice communication methods, people will be seriously affected by a variety of noise signals.Then, the speech enhancement function is to process the loud speech and get the purest speech signal.At present, speech enhancement technology is widely used in video conference, military communication, eavesdropping technology and speech recognition technology, which actually improves the clarity of speech video conference content, improves the communication quality of speech equipment, and solves the problem of speech noise

**第五章 阵列麦克风性能测试**

**5.1阵列麦克风性能测试**

本文通过TDS510仿真器将硬件系统连接到PC主机。系统配置为在操作模拟器模式下启动McBSP1和McBSP2，并继续下一个周期。在测试中，可以通过示波器输入从四个不同麦克风收集的语音信号，并在四个输出通道上观察信号，或者打开CCS开发平台提供的图形显示工具来观察测试信号的图形。根据测试结果，验证系统算法是否有效。如果在线测试成功，请刻录Flash以完成系统的离线工作。

**5.1.1测试实验一**

背景噪音是风扇噪音。实验环境是一个室内工作室，对象的声音是室内扬声器的语音信号，而忽略了房间的反射。在这种环境下，纸张系统的前端麦克风阵列传感器收集的信号是一个有噪声的语音信号，它是DSP系统增强的处理，并输出最终处理的语音信号。该系统的采样频率为8Khz。图5.1是风扇噪声信号，图5.2是分别为4通道麦克风收集的噪声语音信号波形，处理后的输出声音如图5.3所示。

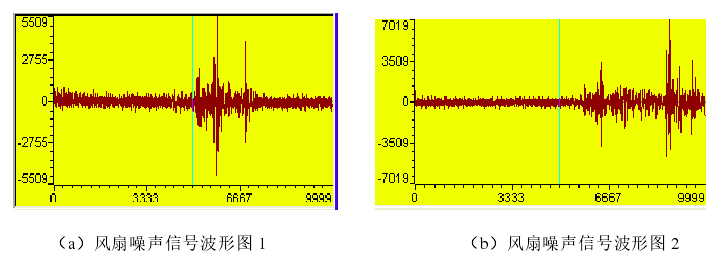
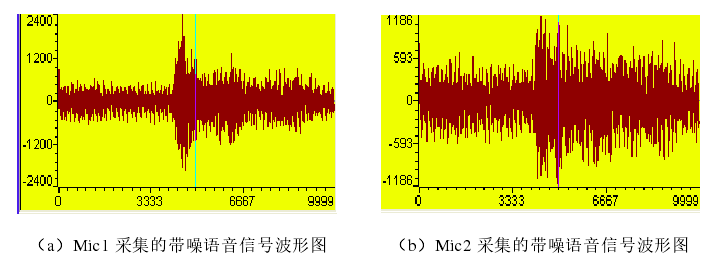


图5.1风扇噪声波形图



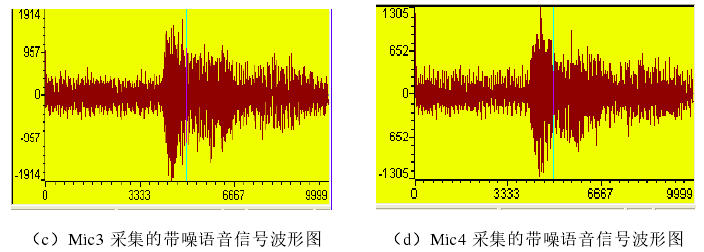


图5.2四路麦克风采集的带噪语音信号波形图

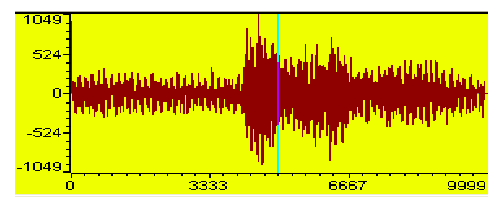


图5.3增强处理后输出信号波形图

在上图中，横坐标表示10000个数据采样点，纵坐标表示信号幅度。坐标幅度越大（或更小），信号量越大（或更小）。如图5.1Ca和（b噪声示例）所示，一般而言，简单的风扇噪声在理论上与白噪声非常相似，但是在实际环境中，由于与外界的摩擦，风扇在某些时候的运动会产生异常信号。 a），（c）和（d）是从四个不同的麦克风收集的语音语音信号波形，尽管观察结果显示四个麦克风收集的信号幅度和噪声有所不同，但信号的基本形状相似。图5.3是DSP处理后的输出语音波形，将该图与图5.2中的四个嘈杂语音图进行比较，可以很容易地看出，在保持一定噪声信号的同时，降低了声音的音量。

**5.1.2测试实验二**

室内噪音很大，所收集的声音是中文句子。其他实验条件与实验1相同，室内噪声波形图，麦克风采集信号和DSP处理输出信号分别如图5.4、5.5和5.6所示。

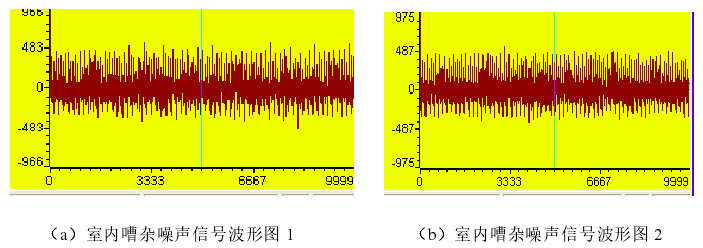
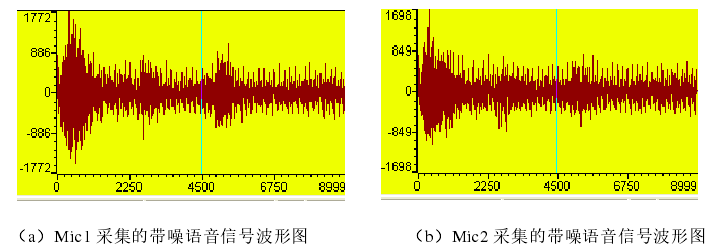


图5.4室内噪声波形图



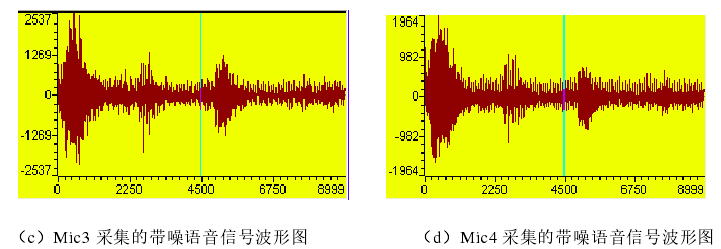


图5.5四路麦克风采集的带噪语音信号波形图

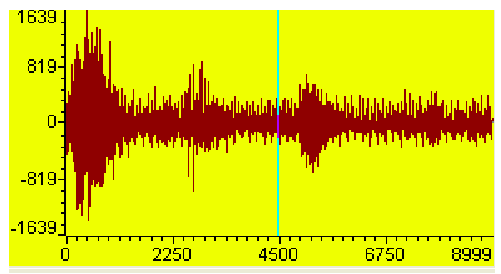


图5.6增强处理后输出语音信号波形图

实验环境中的噪音包括图5.4（a）和（b）所示的键盘打字声音，广播扬声器的嗡嗡声，人与人之间的耳语和行走。由于四个麦克风位于不同的位置，因此从每个阵列收集的目标语音能量和噪声是不同的，但总体波形几乎相同。 CCS视图窗口中显示的图形会根据找到的信号的幅度自动调整纵坐标，因此该图形的系统输出信号似乎并没有明显改善。实际上，该系统具有一定的噪声消除功能，但是噪声仍然存在，不会影响人耳。

**5.1.3测试实验三**

实验背景是一个相对安静的室内环境（没有机械噪音，没有扬声器噪音干扰），在这种环境下是男孩语音采集，使用PC向该音频添加噪音，信噪比为15dB，噪音类型为白色模拟。播放此合成频带噪声是理想的。白噪声信号，四个麦克风的信号波形以及来自系统的输出信号分别如图5.7、5.8和5.9所示。

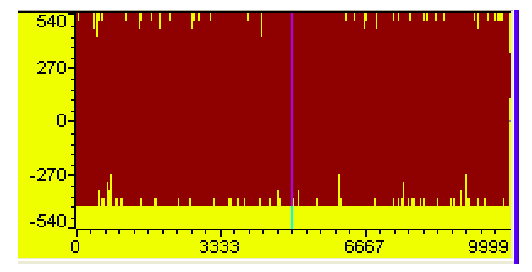
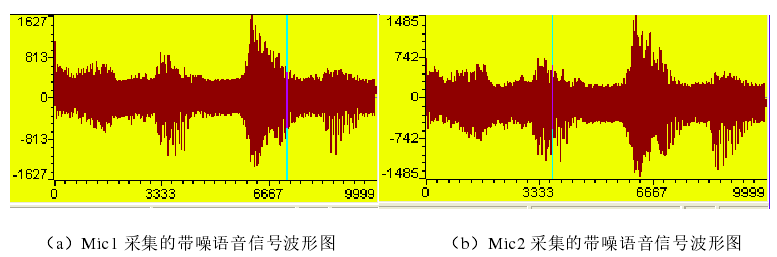


图5.7白噪声波形图



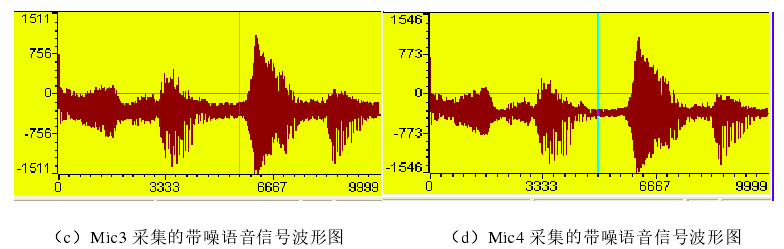


图5.8四路麦克风信号

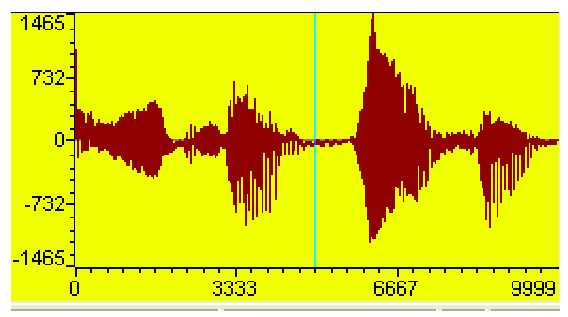


图5.9输出语音信号波形图

首先，在噪声波形图中，添加到本实验中的白噪声与前面两个实验不同，是非常平滑的噪声。在这种噪声环境下，显然系统的语音增强结果是理想的。然后根据MOS评价指标进一步验证算法性能。系统输出语音的主观听觉测试。以上三项实验结果见表5.1。

表5.1MOS评分表

|  |  |  |  |
| --- | --- | --- | --- |
| 测试实验 | 测试实验1 | 测试实验2 | 测试实验3 |
| 处理前一路信号MOS值 | 2.890 | 2.911 | 3.502 |
| 系统输出信号MOS值 | 2.983 | 3.051 | 3.780 |

表5.1中的比较结果表明，实验1和实验2为麦克风阵列收集实际环境中的噪声语音，语音增强效果不明显，但仍然具有一定的噪声抑制功能。从实验3数据比较结果看，第三章模拟实验输入信噪比为15dB时的主观评价大体一致。总的来说，论文将改进的可变步长语音增强算法移植到DSP系统中，可以达到一定的语音增强效果。实际情况下噪声抑制效果不如模拟条件下的处理能力，但根据人的耳主观感觉，噪声消除前具有较好的听觉效果，验证了系统实现的正确性。

**5.2阵列麦克风实际应用测试**

**5.2.1测试准备**

测试准备如下表5.2所示。

表5.2测试准备

|  |  |  |
| --- | --- | --- |
| 项目 | 用例 | 说明 |
| 环境 | 混响环境（模拟家庭客厅环境） |  |
| 器材 | 两个高保真音箱 | 1个用于播放语音 |
| 1个用于播放噪声 |
| 音响支架2个 | 1个用于放置语音播放设备 |
| 1个用于放置噪音播放设备 |
| 笔记本电脑2个 | 1个用于播放语音信号和噪声信号 |
| 1个用于抓取日志或录音 |
| 分贝仪1个 | 用于噪声、语音信号强度测试，计算信噪比等 |
| 卷尺1个 | 用于测试与设备的距离 |
| 语料 | 唤醒语料 | 用于测试唤醒率 |
| 命令词语料 | 用于语音识别，测试识别率 |
| 本机功放播放音频 | 回声消除测试使用 |
| 家庭环境噪声音频 | 可播放中央台新闻节目，约30分钟 |
| 硬件 | 讯飞demo板1个 |  |
| 裸板1个 |  |
| 整机1个 |  |
| 软件 | IPTV主板软件 | 可抓日志，准备至少两个串口线 |
| 可录音，可录15分钟以上。准备两个U盘 |
| 可手动打开/关闭唤醒模式。可手动设置波束 |
| 核心板固件 | 准备烧录工具 |
| 唤醒词 | 跟唤醒词音频一致 |

**5.2.2测试环境搭建**

麦克风阵列测试示意图如下：

唤醒源

噪声源

1m/3m/5m

待测区域

1.5m

麦克风阵列



图5.10麦克风阵列测试示意图

在安静环境下，放置阵列位于待测区域中间位置，唤醒源位于距阵列1m处，噪声源位于距阵列1.5m处，唤醒源和阵列在一条直线上。

通过高保真音箱播放语料，通过分贝仪在阵列处测试信噪比，要求噪声源、唤醒源在阵列处的响度均为55dB。安静环境下和噪声环境下分别测试唤醒率和识别率。

调整唤醒源的位置，距阵列的距离分别为3m和5m。要求唤醒源在阵列处的响度仍为55dB。安静环境下和噪声环境下分别进行唤醒率和识别率测试。

测试说明：测试环境因素影响非常大，唤醒源的位置角度调一调，响度校正时测试值的波动也很大。每次测试都要有对比物，只有同一时间同一环境对比测试的结果才有意义。

**5.2.3实际应用测试**

（1）声学效果测试

1）分别对音箱6麦克整机与音箱裸麦、音箱裸麦与评估板裸麦进行唤醒、声源定位测试，测试步骤为：

|  |  |
| --- | --- |
| 1 | 电脑上打开CoolEdit,准备播放唤醒词语音信号或噪声信号 |
| 2 | 1. 音箱上电后，语音唤醒音箱，确认唤醒功能是否打开，以及唤醒词与语料是否一致。可通过串口信息确认：   root@Hi3798MV200:/mnt/sdcard # logcat -v time |grep Hardware  01-01 08:23:22.805 D/Hardware( 2508): wakeAngle is 0 |
| 3 | 1. 串口登陆IPTV主板，日志保存为文件并命名清楚，执行以下命令开始抓取唤醒日志：   logcat -c  logcat -v time |grep Hardware  01-01 08:23:22.805 D/Hardware( 2508): wakeAngle is 344 |
| 4 | 开始播放唤醒词语音。语音播放结束后，保存好抓取的日志 |
| 5 | 通过查找日志中字符串wakeAngle的个数可以得出唤醒率 |

2）对音箱6麦克整机与音箱裸麦、音箱裸麦与评估板裸麦进行识别测试，测试步骤为：

|  |  |
| --- | --- |
| 1 | 电脑上打开CoolEdit,准备播放唤醒词语音信号或噪声信号 |
| 2 | 音箱上电后，加噪场景下需要手动把唤醒模式关闭，再设置波束定位0号MIC |
| 3 | 1. 串口登陆IPTV主板，执行以下命令开始录音，语料播放结束停止并保存录音。   alsa\_aplay -C -D hw:0,0 -r 16000 -c 1 -f S16\_LE /mnt/sdcard/zhengji\_production\_3m.wav |
| 4 | 采集到的录音数据需要发给讯飞，在服务器上进行语音识别 |

（2）回声消除效果测试

对音箱整机进行回声消除效果测试，在信噪比为（0dB，-5dB，-15dB，-25dB）环境下进行唤醒及识别测试。信噪比0dB是指唤醒源55dB与自身噪声信号55dB之比。也就是把音箱自身响度分为四挡：55、60、70、80dB。

调整音箱自身响度为55dB，进行唤醒及识别测试；

调整音箱自身响度为60dB，进行唤醒及识别测试；

调整音箱自身响度为70dB，进行唤醒及识别测试；

调整音箱自身响度为80dB，进行唤醒及识别测试；

测试说明：只在1m距离不加噪环境下测试。

（3）测试雷区

1）录音文件有数据，音箱播放没有声音。

检查音箱是否静音；检查音箱是否唤醒；

2）加噪声场景下测试识别率，录音文件后半部分，命令词听不清楚，噪声很大。

加噪场景下，音箱可能存在误唤醒，会将波束定位至噪音方向，导致命令词声音被抑制，噪声被增强。所以，加噪场景下可以把唤醒模式关闭，再设置波束。这样可以排除误唤醒的影响，只看降噪效果。需要手动关闭唤醒功能，并设置波束定位至0号MIC。注意，关闭唤醒功能掉电后失效。

**5.3测试结果分析**

从前面的测试可以看出，本章的实验结果与第三章MTALAB算法的仿真结果略有不同。主要原因是MATLAB是一个模拟环境，目标语音来自PC的word文件，噪声也是模拟的理想白噪声。。然而,在真实环境中进行了测试系统,讲话者靶随机走动并有许多类型的噪声环境中,从而造成一定的时间差,理论与实际应用的语音和噪声放大放大的输出信号中存在剩余的语音系统。但主观听力表明，与治疗前相比，这些噪声总体上不影响人耳的听觉感知，提高了清晰度。系统在很大程度上达到了预期的效果，表明系统设计是正确的。与此同时，该系统的缺陷为今后的研究指明了方向。

Speech is the most basic and fastest way of communication in human life.However, when voice communication is performed in the environment where you actually live, the presence of noise in the environment can interfere with the voice.In this context, noise refers to all interfering signals except for useful speech signals [1].There are many types of noise in the real world, such as indoor reverberation, outdoor noise, the complexity of the people around you, electromagnetic noise in the lab, and so on.This is because in the process of using communication and voice communication methods, people will be seriously affected by a variety of noise signals.Then, the speech enhancement function is to process the loud speech and get the purest speech signal.At present, speech enhancement technology is widely used in video conference, military communication, eavesdropping technology and speech recognition technology, which actually improves the clarity of speech video conference content, improves the communication quality of speech equipment, and solves the problem of speech noise pollution.Used in practical situations.It has played an important role in practicality and merit of study.A microphone array

Speech is the most basic and fastest way of communication in human life.However, when voice communication is performed in the environment where you actually live, the presence of noise in the environment can interfere with the voice.In this context, noise refers to all interfering signals except for useful speech signals [1].There are many types of noise in the real world, such as indoor reverberation, outdoor noise, the complexity of the people around you, electromagnetic noise in the lab, and so on.This is because in the process of using communication and voice communication methods, people will be seriously affected by a variety of noise signals.Then, the speech enhancement function is to process the loud speech and get the purest speech signal.At present, speech enhancement technology is widely used in video conference, military communication, eavesdropping technology and speech recognition technology, which actually improves the clarity of speech video conference content, improves the communication quality of speech equipment, and solves the problem of speech noise pollution.Used in practical situations.It has played an important role in practicality and merit of study.A microphone array Speech is the most basic and fastest way of communication in human life.However, when voice communication is performed in the environment where you actually live, the presence of noise in the environment can interfere with the voice.In this context, noise refers to all interfering signals except for useful speech signals [1].There are many types of noise in the real world, such as indoor reverberation, outdoor noise, the complexity of the people around you, electromagnetic noise in the lab, and so on.This is because in the process of using communication and voice communication methods, people will be seriously affected by a variety of noise signals.Then, the speech enhancement function is to process the loud speech and get the purest speech signal.At present, speech enhancement technology is widely used in video conference, military communication, eavesdropping technology and speech recognition technology, which actually improves the clarity of speech video conference content, improves the communication quality of speech equipment, and solves the problem of speech noise pollution.Used in practical situations.It has played an important role in practicality and merit of study.A microphone array Speech is the most basic and fastest way of communication in human life.However, when voice communication is performed in the environment where you actually live, the presence of noise in the environment can interfere with the voice.In this context, noise refers to all interfering signals except for useful speech signals [1].There are many types of noise in the real world, such as indoor reverberation, outdoor noise, the complexity of the people around you, electromagnetic noise in the lab, and so on.This is because in the process of using communication and voice communication methods, people will be seriously affected by a variety of noise signals.Then, the speech enhancement function is to process the loud speech and get the purest speech signal.At present, speech enhancement technology is widely used in video conference, military communication, eavesdropping technology and speech recognition technology, which actually improves the clarity of speech video conference content, improves the communication quality of speech equipment, and solves the problem of speech noise pollution.Used in practical situations.It has played an important role in practicality and merit of study.A microphone array Speech is the most basic and fastest way of communication in human life.However, when voice communication is performed in the environment where you actually live, the presence of noise in the environment can interfere with the voice.In this context, noise refers to all interfering signals except for useful speech signals [1].There are many types of noise in the real world, such as indoor reverberation, outdoor noise, the complexity of the people around you, electromagnetic noise in the lab, and so on.This is because in the process of using communication and voice communication methods, people will be seriously affected by a variety of noise signals.Then, the speech enhancement function is to process the loud speech and get the purest speech signal.At present, speech enhancement technology is widely used

**第六章 总结与展望**

**6.1总结**

电视的内置阵列麦克风需要实现语音增强系统，语音增强是语音信号处理的重要组成部分，具有重要的研究价值和实际意义。同时，与语音增强有关的设备也已经成为社会广泛需求的一种通信产品。在传统的通信设备中，单个麦克风通常用作语音拾取设备。随着社会的发展，对通信质量的需求日益增长，麦克风阵列技术具有出色的空间选择性，逐渐取代单个麦克风成为语音信号研究领域的重要方法。当前，有许多改善麦克风阵列语音的成熟方法，并且已经成功地应用于商业产品，但是在实际环境中存在许多类型的噪声和复杂性。改进本地语言的方法，甚至提出新的理论。本文的主要任务是：

（1）我们研究了麦克风阵列语音增强的基础知识，例如声场模型，阵列拓扑，语音信号的一些预处理方法，延迟和波束形成以及语音信号的数字处理。准备进一步的研究和研究。

（2）设计语音增强硬件系统。系统处理器选择C55x系列TMS320VC5509A数字信号处理器，使用两个A/D转换芯片TLV320AIC23B，释放传统的麦克风输入模式，使用线路输入端子，每个A/D转换芯片都有两个线路输入通道有。提供了四个模拟信号供系统转换为数字信号。另外，还设计了语音获取装置，滤波器电路，音频转换电路和5V电源电路的结构。

（3）采用C语言和汇编语言设计语音增强算法，并结合硬件系统寄存器设置，总线读写子程序项目文件。 CCS3.3用于联合调试代码和硬件系统，在各种室内环境中实时收集响亮的声音，提供实验结果以及分析系统的噪声消除能力。根据系统输出结果，语音增强算法可以满足系统的实时性能，语音增强效果更好。

（4）使用DSP 5509A内部ROM的启动代码编写SPI FLASH读写程序，将其存储在非易失性闪存中，保持仿真器完好，直接实现代码功能并完成整个系统设计。

**6.2展望**

由于能力和知识有限，自适应算法的改进是相对肤浅的，因此硬件设计需要进一步改进。未来的工作前景主要包括两个方面：

（1）本文将自适应语音增强算法应用于内置硬件系统，该系统结构清晰，硬件实现简单。考虑到DSP 5509A的处理能力的局限性，改进的自适应算法是更好的解决方案。但是，由于该方法的降噪功能有一定的局限性，模拟实验不能完全模拟真实环境中的噪声，因此在第五章中，真实声学环境测试对该算法进行了一定的增强，但效果不是很好。多于。研究更适合现实环境的语音增强算法并解决现实噪声问题是前进的道路。

（2）本白皮书未实现软件代码的优化设计，并且如果进行了优化，则可以减少DSP的执行周期，加快处理速度并节省特定的存储空间。因此，这是未来研究工作的方向。

（3）由于5509A是定点DSP，因此计算精度受到限制，并且算法实现过程中的反复迭代不可避免地会导致较大的错误，从而导致系统在现实条件下处理嘈杂的语音，从而进行了MATLAB仿真实验是不同的。因此，在以后的硬件工作中，您可以从基本的处理器系统更新开始。

Speech is the most basic and fastest way of communication in human life.However, when voice communication is performed in the environment where you actually live, the presence of noise in the environment can interfere with the voice.In this context, noise refers to all interfering signals except for useful speech signals [1].There are many types of noise in the real world, such as indoor reverberation, outdoor noise, the complexity of the people around you, electromagnetic noise in the lab, and so on.This is because in the process of using communication and voice communication methods, people will be seriously affected by a variety of noise signals.Then, the speech enhancement function is to process the loud speech and get the purest speech signal.At present, speech enhancement technology is widely used in video conference, military communication, eavesdropping technology and speech recognition technology, which actually improves the clarity of speech video conference content, improves the communication quality of speech equipment, and solves the problem of speech noise pollution.Used in practical situations.It has played an important role in practicality and merit of study.A microphone array consists of two or more microphones, depending on the specific spatial shape.Recently, the use of a single microphone to improve speech quality has yielded very good results [2].However, in real-world environments and applications, such as when the speaker is walking in a small part of the room, many noise signals are mixed into the speech signals collected by a single microphone due to the location of the sourcede voice an increasingly mainstream method of interaction.Short-range speech varies according to the terminal used, while long-range speech provides more openness and convenience.Smart Lart TV has the huge advantage of high frequency and visual screen.It has a unique advantage in developing user habits. Speech is the most basic and fastest way of communication in human life.However, when voice communication is performed in the environment where you actually live, the presence of noise in the environment can interfere with the voice.In this context, noise refers to all interfering signals except for

**参考文献**

[1].动态[J].家电科技,2018,02:6-10.

[2]李坚,徐遥令,袁新艳.一种采用远场语音控制的分体式OLED电视系统设计[J].电视技术,2017,41Z4:177-180.

[3]李坚,徐遥令,袁新艳.一种采用远场语音控制的分体式OLED电视系统设计[J].电子产品世界,2018,2506:47-49+80.

[4]苏亮.创维在京召开“探寻无限”秋季新品发布会，两款重磅智能电视登场[J].家用电器,2019,09:78-79.

[5].创维在京召开“探寻无限”秋季新品发布会,两款重磅智能电视登场[J].电信技术,2019,09:30.

[6]马振洲.AI智能音箱机顶盒的关键部件设计与实现[J].广播与电视技术,2019,4610:77-83.

[7]彭亚奇.基于互联网技术下智能音箱的设计研究[D].武汉理工大学,2018.

[8]舒治宇.基于麦克风阵列的声源定位算法研究[D].湖北工业大学,2019.

[9]吴晓静.家庭数字娱乐系统的智能语音服务设计研究[D].华东理工大学,2018.

[10]徐勋庭.新型语音交互式智能家居系统研究与设计[D].广东工业大学,2018.

[11].美国国家标准学会(ANSI)标准题录(下)(2005年版)[J].世界标准信息,2005,11:188-239.

[12]王晓东.基于麦克阵列的点声源定位及声源类型分析[D].东北石油大学,2016.

[13]聂欣欣.麦克风阵列的自适应降噪算法研究[D].哈尔滨理工大学,2016.

[14]张帅.多阵列宽频双极化天线的设计[D].武汉邮电科学研究院,2017.

[15].通信[J].中国无线电电子学文摘,2011,2705:93-132.

[16]陈露明.嘈杂背景下的声源定位及语音分离实现技术的研究[D].电子科技大学,2007.

[17]黄奔.气动噪声源的麦克风阵列识别定位技术研究[D].中国空气动力研究与发展中心,2014.

[18]肇毓.基于RFID的ETC前端系统的设计与实现[D].东北大学,2011.

[19]王宏敏.麦克风阵列语音增强及DSP实现[D].辽宁工业大学,2015.

[20]梅亚男.分布式麦克风阵列说话人定位与跟踪[D].辽宁工业大学,2015.

[21]刘超.基于麦克风阵列的声源定位算法研究[D].南京大学,2015.

[22]姚传涛.多频带手机天线的研究与实现[D].华东理工大学,2015.

[23]郭元曦.水声阵列信号处理的并行实现研究[D].哈尔滨工程大学,2008.

[24]吴博.无线移动终端的天线设计[D].北京邮电大学,2011.

[25]杜晓琳.3G上网卡天线OTA性能及低SAR终端研究[D].北京邮电大学,2012.

[26]李向阳.小区宽带接入工程基于EPON技术的研究与应用[D].北京邮电大学,2012.

[27]朱兴宇.基于麦克风阵列的语音增强系统设计[D].北京化工大学,2012.

[28]韩海强.基于RFID的公交多功能自动报站系统[D].兰州大学,2009.

[29].每月新品[J].世界电子元器件,2008,04:48-59.

[30]郭敬.基于听觉的室外小型移动机器人相对定位[D].天津大学,2012.

[31]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].安徽大学,2014.

[32]朱泰鹏.基于近场模型的麦克风阵列语音增强算法研究[D].华南理工大学,2013.

[33]Williams, M, & Le Du, Guillaume. (2000). Multichannel Microphone Array Design.

[34]Pirz, & F. (1979). Design of a wideband, constant beamwidth, array microphone for use in the near field. Bell System Technical Journal, 58(8), 1839-1850.

[35]Pirz, F. C. . (1979). Design of a Wideband, Constant Beamwidth, Array Microphone for Use in the Near Field. IEEE International Conference on Acoustics, Speech, & Signal Processing. IEEE.

[36]Do, A. F. R. , Carlos, S. R. J. , & De, M. M. A. F. . (2018). Design of microphone phased arrays for acoustic beamforming. Journal of the Brazilian Society of Mechanical Sciences & Engineering, 40(7), 354-.

[37]Ayllón, David, Gil-Pita, R. , & Rosa-Zurera, M. . (2013). Design of microphone arrays for hearing aids optimized to unknown subjects. Signal processing, 93(11), 3239-3250.

[38]Huang, G. , Benesty, J. , & Chen, J. . (2017). On the design of frequency-invariant beampatterns with uniform circular microphone arrays. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 25(5), 1140-1153.

[39]RABINKIN,Daniel V. (1996). A DSP Implementation of Source Location Using Microphone Arrays. Meeting of the Asa.

[40]Ben Hagai, Ilan, Fazi, Filippo Maria, & Rafaely, Boaz. (2011). Efficient dual-sphere microphone array design based on generalized sampling theory.

[41]SARIC, Zoran, JOVICIC, Slobodan, KOVACEVIC, Vladimir, TESLIC, Nikola, & KUKOLJ, Dragan. (2008). SYSTEM AND PROCEDURE OF HANDS FREE SPEECH COMMUNICATION USING A MICROPHONE ARRAY. WO.

[42]Seongkyu Mun, Suwon Shon, Wooil Kim, & Hanseok Ko. (2015). Robust speaker direction estimation with microphone array using nmf for smart tv interaction.

I Ben Hagai, F M Fazi, & B Rafaely. (2011). Efficient dual-sphere microphone array design based on generalized sampling theory. Forum Acusticum, Aalborg, Denmark.

[43]Bertrand, A. , Doclo, S. , Gannot, S. , Ono, N. , & Waterschoot, T. V. . (2015). Special issue on wireless acoustic sensor networks and ad hoc microphone arrays. Signal processing, 107(feb.), 1-3.

[44]STMicroelectronics. (0). St's mems microphone smoothes the way to natural voice control in smart consumer devices. stmicroelectronics.

[45]Nazareno Brier Cruzada. (2008). Electronic personal computing and videophone system consisting of a remote server system providing dynamic, subscription based virtual computing services & resources, a thin client hardware device connected to a television set and wireless keyboard & mouse, and a wireless mobile device (a Pocket PC Phone). US.

[46]Masakiyo Fujimoto, Yasuo Ariki, & Shuji Doshita. (2003). Hands-free speech recognition in real environments using microphone array and 2-levels mllr adaptation as a front-end system for conversational tv. Acoustical Science & Technology, 24(6), 379-381.