

969212



中国人民解放军军医进修学院
中国人民解放军总医院


硕士学位论文

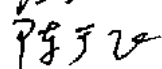
军医进修学院

研究生学位论文原创性声明

秉承我院“敬业、勤奋、求实、创新”的学风，本人声明：所呈交的是我本人在导师指导下进行的研究工作及取得的研究成果。据我所知，除了文中特别加以标注和致谢的地方外，论文中不包含任何其他个人或集体已经发表或撰写过的研究成果，也不含为获得我院或其他教育机构的学位及证书而使用过的材料，对本文的研究作出贡献的个人或集体，均已在文中做了明确的说明并表示谢意。

申请学位论文与资料若有不实之处，本人承担一切相关责任。


论文作者签名：  日期：2006.5.27

指导教师签名：  日期：2006.5.27

军医进修学院

研究生学位论文授权使用授权书

本人保证毕业离院后，发表论文或使用论文工作成果时署名为军医进修学院或解放军总医院。学院有权保留并向国家有关部门或机构送交论文原件、复印件和电子版本，可以采用影印、缩印、扫描或其它手段保存论文以供被查阅和借阅。学院可以公布学位论文的全部或部分内容（保密内容除外）。

论文作者签名：  日期：2006.5.27

指导教师签名：  日期：2006.5.27

基于 DSP 数字助听器关键技术的研究

摘 要

本文在学习国外当前先进技术的基础上,对数字助听器的软硬件开发进行了深入的探索,并结合听觉生理学和语音信号处理技术,讨论一种基于 TI 公司 DSP TMS320VC5416 数字助听器关键技术的设计思路和实现方法。

本文介绍的数字助听器系统之设计包括硬件电路和算法研究,硬件上主要由核心处理芯片 TMS320VC5416 和模拟接口芯片 TLV320AIC23B 两部分构成,系统的主要功能均在 DSP 中完成,硬件电路设计相对比较简单。本文重点在于助听器语音信号处理算法的研究,系统能够实现依频域宽动态压缩,移频压缩和自适应噪声消除等功能。首先分析目前数字助听器应用的数字信号处理算法,针对传统多频段压缩放大方法存在的不足,提出依频域的宽动态压缩算法。该算法灵活的实现患者在不同频段需要不同增益补偿的功能,在对不同频率成分进行非线性压缩放大的同时,尽可能保留较多的基本频谱特征信息。同时本文还专门针对重度听力障碍患者听力学特征所提出移频压缩方案,将语音频谱从高频部分按比例搬移到低频,保证语音频谱适应患者残留的动态范围而保持频谱的相对特征不变。此外本文还设计一种语音信号自适应噪声消除算法,利用 LMS 自适应算法的特点抑制语音信号中的宽带噪声信号,提高信噪比。系统所有的算法的目的都是避免语音失真,提高舒适度。而后本文还利用 MATLAB 软件仿真实现所有算法,对比实验结果,进一步确认算法的可行性。

最终,算法的设计方案被移植到 DSP 语音实验平台上,在 CCS 强大的编译环境中,采用 C 和汇编语言联合编译的方式实现所有的算法,并调试成功,通过对比实验结果,取得符合理论推算的结果,从而验证了系统的可操作性。本系统可根据患者的听力测试结果任意调节频响参数,满足临床听障患者对听力进行补偿的要求,也为进一步研究助听器高级算法搭建了较好的实验平台。

关键词: 数字助听器 DSP 宽动态压缩 移频压缩 自适应噪声消除

Study on Key Technology of Digital Hearing Aid

Based on DSP

Abstract

This paper presents research on hardware and software of DHA (digital hearing aid) based studying of overseas advanced technology, and it also describes the key technology for design of DHA based DSP TMS320VC5416 from TI with theory about physiological acoustics and speech signal processes.

The DHA system introduced consists of circuit and algorithms. Construction of circuit is include of DSP TMS320VC5416 as central MCU and TLV320AIC23B from TI as Analog Interface Circuit. Function of the system is performed in DSP mostly, so design of hardware is easier than algorithms in a certain extent. Emphasis of this paper is research in algorithms for speech signal processing. The DHA system applies WRDC (wide dynamic range compression), frequency transposition compression and adaptive noise reduction etc. Firstly we study the algorithms used for digital signal processing of speech for the hearing impaired currently, then we present devise of spectrum dependent WRDC which can remedy deficiency of previous algorithms. The devise perform functions for different loudness compensation in different frequency area facility. Most basic properties of speech signal remain invariant across procession, while it receives nonlinear amplification variously. we also give an algorithm of frequency transposition compression especially for persons with hearing loss seriously. Frequency spectrum in high frequency area is compressed linearly to low to fit impaired's residual dynamic range in accordance within their audiologic curves and properties of speech signal remain invariant relatively. A noise reduction algorithm used in speech signal processes which is referred to as LMS adaptive noise reduction. is also designed to reduce wide-band noise but improve Signal to Noise Ratio. All algorithms aims for less distortion and better comfort. Then we emulate them on MATLAB, and obtain evaluation result to prove availability of

the algorithms.

At last the algorithms are replanted on DSP experimental kit for speech signals processing. we build project with C and Compile language in CCS. All programmes runs correctly and exact result is gained to prove availability DHA system based DSP. The system can set parameter of frequency response depending impaired's audiogram and fit their need for loud compensation. This system is well available for study of advanced algorithms of DHA.

Key words: digital hearing aid, DSP, wide dynamic range compression, frequency transposition compression, adaptive noise reduction

第一章 绪 论

1.1 课题背景

随着社会的不断老龄化以及人们对耳聋问题的日益关注，助听器的发展逐渐受到人们的重视。据英国医学研究理事会听力研究学会估计，到 2005 年全球听力损失程度在 25 分贝以上的总人数将超过 5.6 亿，而中国听力障碍的患者有 1.2 亿多。调查表明，世界上每 10 人就有 1 人听力损失，均能够通过使用助听器获益。其中特别是感音神经性听力损失的患者，除突发性耳聋等少数情况外，目前尚无良好的药物、手术等治疗手段，而且许多情况下病因不明。但是人类发展到今天是以能够进行言语交流和具有思想意识为基础的，听力下降而导致不能良好的与他人交流，这对患者本人、家人和社会是及其痛苦的。在没有找到医疗的药物和手术方法以前，科学家们在几个世纪以来采用各种方法帮助听力障碍患者，使用助听器来获得听力补偿就是最常采用的方法之一，其已经成为听障者康复的主要手段。

因为听力障碍的患者病因各异，其听力损失情况存在着较大差异，每个患者的听力损失曲线几乎都是不一样的，加上患者在年龄、语言识别、以及对设备适应程度方面存在着差异，使得每位患者对于助听器的补偿有着不同的要求。传统的助听器存在线路功能的局限性，无法满足大部分患者的要求。目前国外助听器技术已经由传统的模拟放大电路逐步被功能强，运算快，功耗低，体积小的 DSP 处理芯片所取代，从而使现代助听器技术产生了一个质的飞跃^[1]。

数字信号处理（Digital Signal Processing, DSP）是一门涉及许多学科而又广泛应用于许多领域的新兴学科。二十世纪六十年代以来，随着计算机和信息技术的飞速发展，数字信号处理技术应运而生并得到迅速的发展。数字信号处理是利用计算机或专用处理设备，以数字形式对信号进行采集、变换、滤波、估值、增强、压缩、识别等处理，以得到符合人们需要的信号形式。数字信号处理技术是围绕着数字信号处理的理论、实现和应用等

几个方面发展起来的。数字信号处理在理论上的发展推动了数字信号处理应用的发展。反过来,数字信号处理的应用又促进了数字信号处理理论的提高^[5]。在生物医学工程领域中,DSP 技术在数字化助听器的开发过程中得到广泛的应用,数十年 DSP 技术在电子、信号处理、电信和供电等技术领域的革命性进展,都对 DSP 助听器的发展和广泛应用作出了贡献^[3]。

经过近 20 年的实验室研究,1986 年数字调节的模拟机问世,但真正实用的数字式助听器直到 1996 年才见于市场,助听器使用数字信号处理技术给听力障碍患者带来了新的希望,能够实现更好的保真度,更高的灵活性,改善噪声环境下的语言辨别能力,更好地“恢复”各种令人烦恼的听觉损伤。

1.2 国内外研究的进展

十八世纪以来,人们一直没有间断过助听设备的研究工作,助听器经历了集声器时代,碳精时代,电子管时代,集成电路时代,如今已经迈进数字时代。

数字技术的发展给助听器的开发注入了新的设计理念,将数字技术最早应用于语言处理是对复杂模拟系统(analog systems)的仿真处理(simulations)。早期的仿真处理需要很长的时间,对几秒钟的语言信号处理需要一个小时甚至更长。60 年代中期计算机处理速度的发展使得数字式硬件在可靠性和紧密度(compactness)方面的优势得以显现。其后逻辑装置(logic devices)的速度和集成电路(integrated circuits, ICs)得到了进一步的发展。这就使得在单一芯片上进行大量运算和对言语取样速率进行许多实时处理成为可能。

60 年代初期纽约市立大学(CUNY)和美国聋人中央研究院(CID)开展了助听器此方面的研究。CUNY 主要探索计算机化的心理物理测试,而 CID 侧重于诱发电位计算机化数字言语系统的应用。CUNY 最终在高速 DSP 芯片问世之前制造出了建立在阵列处理器(array processor)之上的数字主导听器。阵列处理器能够同时处理大量数字,这是最终产生 DSP 助听器的重要一步。而 CID 终于在 1985 年制备出了一台用电池为能源的盒式试验

性 DSP 助听器，并在取得专利后转让给制造商^[27]。

和其它较小的产业一样，助听康复设备的技术进步主要依赖于大型制造业（如半导体和编程硬件嵌入式软件制造业）的发展。因此 1989 年代小型高速 DSP 芯片技术的发展导致了耳级（wearable）DSP 助听器的快速发展。在此期间，Bell 实验室、MIT 等一些研究机构也作出了自己的努力。在商业制造方面，Auditone、Nicolet 和 Rion 也为制造出患者可以接受的 DSP 助听器作出了贡献。Auditone 的最初产品虽然不能商业使用，但在技术上确为早期的真正数字处理盒式助听器。Nicolet 为 Phoenix 研发的数字处理助听器初为盒式，后来试制了一款很大的耳背式。后者使用 3 个 675 型电池（3 个电池相互叠放在一起）。但是这些产品在体积和外观上均未能达到患者的实用要求。

近几年来，由于一些国家的政府机构、大学研究机构和生产企业的联合，数字式助听器有了实质性的发展，从常规的模拟机到编程机，进而发展到数字式。以下因素对 DSP 助听器的研制起到了主要作用：

- a) 硅片技术使制造低功耗计算机芯片成为可能，其耗能量低至可以考虑使用常规助听器电池作为电源；
- b) 用计算机分析研究外耳道和鼓膜对助听器放大的作用比传统方法更有优越性；
- c) 助听器放大声音的某些参数（如清晰度指数、量化的言语可懂度）的作用被广泛接受。

1996 年 Oitcon 和 Widex 几乎同时在美国市场推出了“真正的”DSP 助听器：DigifocusTM 和 SensoTM。目前几乎所有主要的生产厂商均推出了各自的“全数码”助听器，但是各自的产品在时间处理、压缩阈、压噪方式、抗反馈线路等方面均有很大不同，各成体系。数字助听器代替基于模拟电路的助听器成为趋势，据统计，2002 年，美国助听器总销量的 40% 为数字式产品，在我国随着低价位数字产品的问世，其选配数量也呈上升趋势。

我国的助听器生产和选配工作开始于 20 世纪 60 年代，近年来质量、种类、数量有较大快速的发展。90 年代，各大国际名牌企业开始纷纷到中国投资建厂。但国内的助听事业的发展与国外相比存在一定差距，特别是

在助听器的生产能力方面，国内停留在较低价位的盒式和中小功率的耳背式助听器，都属于模拟式的，产品与国际先进水平相比，在技术上存在着较大的差距，特别是在信号处理技术方面。国内厂家的产品仍采用模拟电路设计，对数字信号处理技术没有太多的研究，到目前为止没有国内自主知识产权的数字助听器产品面市，这些情况带来的直接结果就是国外的助听器产品几乎垄断了国内市场。因此，我国亟需继续投入人力、物力开发新型的国产助听器，特别是数字助听器。

1.3 数字式助听器的一般原理

助听器是一种增加声音强度的装置，它可以协助某些耳聋患者适当提高听力。从技术上讲，任何将有声环境与听力障碍个体联系起来的装置都可以认为是助听器。声音强度的扩大可以依靠增加声音集合、声能放大和减少声能消散等方面来获得。助听器产品从诞生到现在，经过了一百多年，但其基本结构没有发生变化，仍由三部分组成：传声器、系统电路、耳机。但各个部分的技术已经有了很大的发展。传统的助听器完全采用模拟器件来完成上述的对语音信号的滤波和压缩放大处理，助听器最初的设计只是考虑对声音敏感性下降的补偿，实际只是一个信号放大器。模拟信号处理的局限性很大，电路的动态范围受到电路噪声极限和输出最大值的限制；作为滤波器基本环节的放大器，其动态范围将会因器件耗能下降而降低，靠减少体积、降低功耗、增大数据处理能力的想法是不现实的；模拟电路的体积减小有很大困难，噪声干扰也较严重。因此，采用数字电路与数字信号处理技术来设计助听器成为必然。

随后出现的数字可编程助听器（digital programmable hearing aids）将数字控制器件和存储器件引入助听器，虽然对语音信号的处理仍由模拟的功能模块完成，但是这些模块的参数，如频带宽度、放大增益、压缩阈值和压缩比率等，是由系统中的数字控制器件决定的，这样就可以根据病人的具体情况方便地调节助听器的参数，甚至可以在存储器件中保存几套不同的参数以对应不同的使用环境等。采用模拟器件来完成对语音信号的处理存在一些固有的缺陷，主要是由于模拟器件的功能固定，造成系统的功能

比较单一，无法进行一些智能化的处理，如有效地区分噪音和语音信号等，而这些功能越来越成为人们衡量助听器性能的重要指标。

近年来，随着数字集成电路技术以及数字信号处理技术的发展，开始出现全数字化的助听器产品。数字助听器首先对麦克风输出的电信号进行模数转换（A/D），然后在数字域对语音进行处理，如进行频带划分、分频带压缩放大、减少背景噪音等，处理后的数字信号经数模转换（D/A）还原为语音信号，送往受话器。数字型助听器可以有效解决或减少某些模拟型助听器无法解决的问题，其根本区别就在于：数字型助听器在系统电路中可以通过 DSP 产品加入数字信号处理方法，来完成对声音信号的分离、降噪等功能使得使用者获得较为清晰的声音信号。DSP 助听器还具有明显优势有：

- 1) 由于更多的操作由软件完成，DSP 助听器需要较少的元件。这就意味着线路具有更大的稳定性，而且随着使用时间的延长产生混叠信号更少。
- 2) 由于滤波功能和增益值的调节是由软件完成的，所以 DSP 需要较少的机械调节，如各种调节钮和微调。减少空间和电路焊接点，降低故障的可能性。
- 3) 与使用高通滤波器的许多模拟机相比，DSP 在消除反馈啸叫方面更加行之有效，而不牺牲交谈语言的高频信息；
- 4) 由于数字技术的灵活性，为进一步开发助听器各类高级信号处理算法提供便捷，减少研发周期。

数字式助听器的基本结构图如图 1-1 所示：传声器→前置放大器→抗混叠（anti-aliasing）滤波器→模数转换 ADC→数字微处理器和数字滤波器→数模转换 DAC→抗映像（anti-imaging）滤波器→耳机。

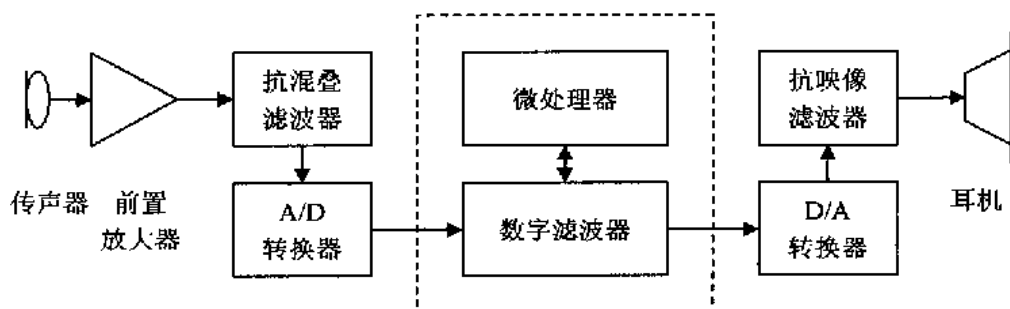


图 1-1 数字式助听器基本结构图

传声器作为助听器产品中的重要组成部分，在其中起到收集声音的功效。但长期以来，助听器中的传声器在采音过程中，只能全向采音，即接受声音不集中于某一方向，而是对整个区域的声音进行收集。这种情况在佩戴者身处一个较复杂的环境中并要与多方面的人进行交谈时比较适用，而在只有两者进行单独交谈的情况下，全向采音就显得不太适宜。正因为实际使用中的这种需求，当今先进的数字助听器又设计出了传声器的定向采音技术。这种传声器可以具有从全向采音到定向采音的功能。

目前，助听器中的耳机与传声器经过多年的发展，结构越变越小，功能越变越完善。很多国外厂家如斯达克、西门子等研制的耳机和传声器系统芯片体积仅如一粒芝麻般大，结构精巧。国内清华大学研制的硅微声学器件微麦克风灵敏度达到 15mV/Pa ，频响范围 $25\text{Hz}\sim 20\text{kHz}$ ，工作电压 3V ，芯片尺寸小于 $5\text{mm}\times 5\text{mm}\times 0.5\text{mm}$ ，完全满足数字助听器的设计要求。

数字助听器的核心是用于完成语音信号处理的 DSP 器件，图中微处理器和数字滤波器部分即为 DSP 芯片。DSP 提供的运算能力可以实时地进行频谱分析和处理，可以根据佩戴者的听力曲线进行复杂的非线性处理，可以根据不同的使用环境对环境噪声进行智能化的处理。

借助于高集成度芯片的生产技术已经实现了功能强大模拟接口芯片，其上集成了高精度 A/D、D/A、以及抗混叠和抗镜像滤波器。

助听器的小型化一直是助听器的主要发展目标，到 20 世纪 90 年代末完全耳道式助听器出现后，这一努力告一段落，研制的重点开始转向助听效果的优越和操作的便利性。表现在各产品将卖点转向数字式、通道数、

独立去噪功能和多助听方案选择，助听器的算法也得到人们更多的重视。

基于 DSP 的助听器能够实现软件控制功能，包括频率整形、反馈抑制、噪声抑制、双耳处理、耳廓和耳道滤波、混响抑制并备有用于接收来自数字电话、电视机或者其它音频设备直接数字输入的接口。DSP 还意味着无需改变硬件即可对助听器的算法及特征进行客户化设计或变更。甚至还有可能具备可由用户选择的程序，以便在恶劣的聆听场合切换到精细处理的声音，或在安静环境下切换回传统的失真较低的声音。除此各主要助听器生产厂家还纷纷开发各种复杂的语音增强算法，例如，为了更好的提高语言分辨能力，目前也出现了一种改变音素时间长度的算法，即将较不易辨识的辅音的时间加长，而相对地缩短较易辨识的元音的时间。这样的系统在电子耳蜗中使用得相当成功，目前也考虑使用到数字助听器中来。

1.4 本文研究内容及意义

本文的研究工作主要有以下几个部分内容：

- 1) 系统学习听觉生理学理解听觉产生过程和听觉障碍的机理；
- 2) 研究人声发音机理和语音信号处理的基本知识，并分析语音信号的性质，为后续的研究奠定基础；
- 3) 研究现有数字式助听器所应用的数字信号处理方法，提出基于频域的宽动态压缩算法、自适应消噪算法以及其他与助听器相关的功能算法；
- 4) 利用 MATLAB 软件进行算法仿真；
- 5) 在现有的 DSP 开发板上实现移植算法，对实验结果进行分析，得出结论，为下一步工作提出参考和依据，此步为本文的重点；
- 6) 自主设计基于 DSP 的助听器硬件系统。

目前国外助听器技术已经由传统的模拟放大电路逐步被功能强，运算快，功耗低，体积小的 DSP 所取代，从而使现代助听器技术产生一个质的飞跃。本文在学习先进技术的基础上，研制以 TI 公司 DSP TMS320C5416 为核心的助听器，对助听器的软硬件开发进行了深入的探索。我国助听器行业尚就处于十分落后的水平，至今还没有完全自主知识产权的数字助听

器生产，本文能够为国内基于 DSP 数字助听器的研发提供参考。

根据以上工作内容，本文共分为七部分。第一章为绪论，介绍数字助听器的发展现状；第二章从听觉生理学角度解释听障发生的机理和测定；第三章介绍语音信号的特征以及数字语音处理的发展概况；第四章介绍系统的硬件结构可行性设计方案和硬件开发环境，其中重点介绍 TMS320VC5416 的性能特征、体系结构；第五章详细介绍助听器语音信号处理的发展状况，以及本系统的信号处理方案，其中重点讨论宽动态压缩、自适应噪声消除和移频压缩算法；第六章对语音处理方案进行实验验证。首先通过 MATLAB 在计算机上模拟可行性，而后在 DSP 硬件平台上编程实现；第七章为结束语，分析系统设计存在的不足，并讨论下一步研究的改进方向。四、五章是本文的核心部分，也是课题研究的重点。

第二章 听觉生理学

今天我们周围有听力障碍的人越来越多，他们中有的的是因为年老听觉功能衰退所致，有的是先天性耳聋，有的是后天因药物所意外伤害所致，还有的则是由于长期工作在强噪声环境造成听力的下降。要解除患者的痛苦，最有效的方法就是给他们配戴一个助听器。助听器的主要功能是补偿耳朵灵敏度的下降，听力下降的患者只是在某些频段上敏感性下降，而且每个频段敏感性下降又具有其特异性，设计一个有效的助听器首先必须了解耳朵的生理特征。

2.1 听觉的一般生理学基础

听觉系统由听觉器官（耳）、听神经、各级听觉中枢及它们之间的连接网络组成，听觉就是接收声音，转换成为神经脉冲并经过大脑听觉中枢处理的过程。听觉是学习掌握言语—语言的最重要的感觉道，听觉减退就可能影响言语—语言的发挥运用。

外周的听觉系统一般指听觉器官（耳）和听神经。如图 2-1 所示：耳分为外耳、中耳和内耳。

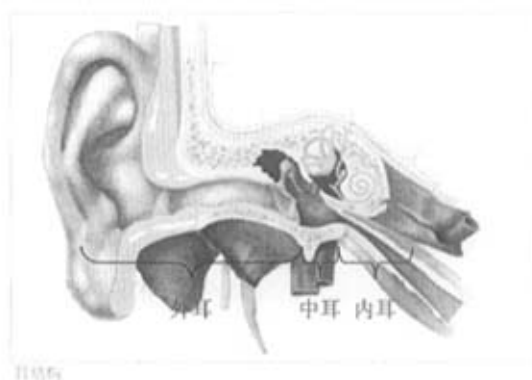


图 2-1 耳的结构示意图

外耳包括耳廓及外耳道，主要起集声和共振作用。两侧耳廓的协同集

声可以帮助判别声源的方向。耳廓结构对声音就有增益的作用,耳廓呈不规则的漏斗状,反射界面有利于集声。外耳道是一段相对密封的管腔,声音在此管腔内能产生共振的作用,特别是对 3000~4000Hz 之间频率的扩音作用达到高峰。

中耳包括鼓膜、听骨链、鼓室、中耳肌、咽鼓管等结构。鼓膜和听骨链实际上作为整体起着增高声压、减低振幅的作用。鼓膜振动的上限频率约为 20kHz,老年人稍低一些。听小骨的功能是阻抗变换和限幅。听骨链本身具有一种杠杆作用,它是鼓膜与卵圆窗之间的机械联系,其主要生理功能是实现有效的阻抗匹配,作为一个杠杆系统把声波的振动传递给内耳。限幅可保护耳在大音量时免受损害,它是由内耳肌完成的,因此内耳肌能利用听小骨的收缩功能使传输得到衰减,尤其是在音量很大的情况下,由于镫骨肌的收缩作用能改变镫骨的振动方向从而降低了卵圆窗的激励。中耳对所有的频率的传送是不同的,在 2kHz 附近的频率它传送最好,在高频、低频时传送效率将降低。

内耳由前庭、圆形窗、卵形窗及耳蜗组成。前庭包括半规管和有关器官,它们用于平衡感知。耳蜗是一条通过圆形窗和卵形窗与中耳相通的蜗形通道,内有将声振动转换成神经生物电信号的换能器。蜗形的管道一头较粗为蜗底,一头较细为蜗顶,耳蜗从蜗底伸到蜗顶,分别为前庭阶、鼓阶和蜗管。分隔蜗管和前庭之间的膜成为前庭膜,分隔蜗管和鼓阶之间的膜成为基底膜,由感觉细胞、听神经末梢及其他结构组成的听觉感受器(Corti's organ)就排列在基底膜上。

外界的声音传导至耳蜗引起基底膜振动,振动以波的形式沿着基底膜向前传播。按照物理学上的行波学说。振动在基底膜上从蜗底向蜗顶传播时,振幅逐渐增大,当达到共振频率与声波频率相一致是振幅最大,离开该部位后,振幅很快减小,稍远即停。耳蜗基底膜的部位不同所感受声音的频率也不同。高频引起的最大振幅靠近蜗底,低频则靠近蜗顶。基底膜振动时,毛细胞上与顶盖膜相连的纤毛发生弯曲,从而引起毛细胞兴奋,并将机械能转化为生物电能,而使附于毛细胞底部的蜗神经末梢产生神经冲动,经蜗神经及其中枢传导径路上传到听觉皮层。

声音传入耳内有两条途径，一是空气传导（气导），空气振动的声波由耳廓收集，经外耳道抵达鼓膜，引起听骨链和鼓室内的空气也发生振动。听骨链的振动最终传至基底膜，使位于基底膜上的螺旋器受到刺激，将冲动经听神经传至中枢产生听觉。另一为骨传导（骨导），声波直接经颅骨传导，使外淋巴发生振动，再刺激耳蜗的神经末梢产生听觉。骨导虽发生在气导的同时，但经骨导竟如耳蜗的声能很小，并无重要意义。因此，声音传入内耳的途径以气导为主。

2.2 听觉障碍的机制

听力系统是人体重要的感知器官之一，听觉系统从末梢到中枢任何一个部位发生的病变，都会引起程度不同的听觉障碍。听力残疾是我国最常见的五种残疾（智力、视力、肢体、精神及听力）中人数最多的一种。听力损失对健康的危害不容忽视，听觉反应迟钝，会使大脑思维缓慢，影响与人交流，影响中枢神经整体平衡发展。引起耳聋的原因很多，如：遗传、感染、耳毒性药物、噪声、老年性、外伤等等，仅 10% 的患者可通过医疗手段完全或部分恢复听力，其他则通过佩戴助听器达到听觉补偿的目的，全聋患者应考虑电子耳蜗植入，并接受相应的语言训练。

听力障碍主要表现为听力下降，听觉动态范围的减小以及对声音的频率分辨力和时间分辨力的下降。图 2-2 显示的分别是一个正常的听力状况和有听力损失的听力状况。图中上面的一条曲线称为听力曲线(audiogram)，听力曲线上的点表征的是在某个频率点上的听阈值，即人能够感知到声音时的最小输入声强。在听力损失的情况下，听阈普遍下移，从而造成听觉动态范围减小。同时还可以发现这种动态范围的减小程度与频率有关，一般高频损失较大。听力损失还会造成语言分辨能力下降等相应问题。

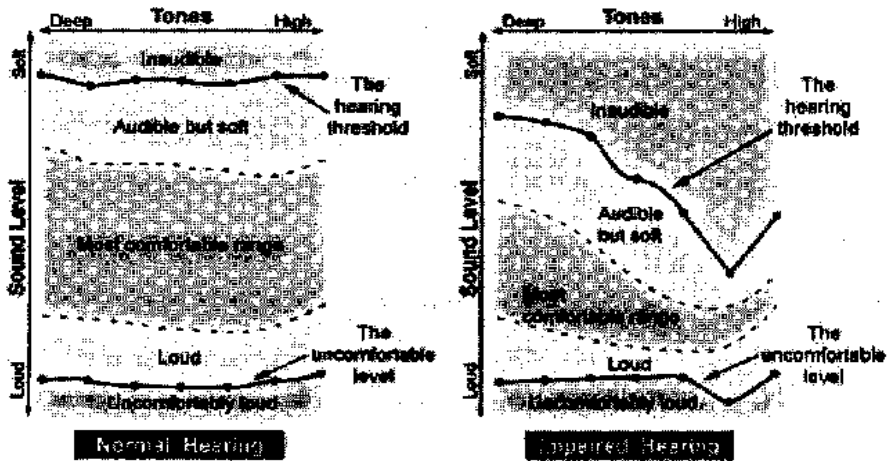


图 2-2 正常耳与听障耳的听力曲线

听力障碍轻者为失聪（hard of hearing），重者为耳聋（deaf）。当中耳腔内的质量增加会引起高频听力的下降，而中耳腔内的僵硬性增加则会引起低频听力下降。当听觉系统有器质性病变时称为器质性耳聋，它又分为传导性聋（conductive hearing loss）、感音性聋（sensorineural hearing loss）以及混合性聋（mixed hearing loss）。其中，传导性聋由于外耳和中耳病变引起，耳蜗以上的功能仍为正常，声音在抵达内耳之前的振动受阻，传导性聋一般采用手术治疗，之后若需要可以配戴适当的助听器恢复听力。

助听器主要的适用人群为不可治愈感音神经性损失患者，耳蜗、听神经或者听觉中枢发生病变引起对声音感觉和认知功能障碍的听力损失称为感音神经性聋，其中比较多见的是属于耳蜗病变为主的听力损失。耳蜗中连接基膜的毛细胞作用是将机械信号转换成神经信号，如果毛细胞遭到损坏，听觉系统就小能将声压波转换成神经脉冲，导致听觉丧失。毛细胞损坏经常是由某些疾病（诸如脑膜炎等），或者先天性不足，或其他的一些原因造成的。如果大量的毛细胞或听神经受到损伤，这种情况将导致深度耳聋。

感音神经性病变主要的临床表现主要有：

- 高频听力损失，频率分辨率差，言语辨别力低；
- 听觉动态范围明显缩小，有小声听不见，大声又嫌吵的重振现象；
- 高频耳鸣，多与听力损失同时发生。

目前应用药物或手术治疗感音神经性听力损失存在一定局限性，因此选配合适的助听器有助于提高“听力”，对于部分不明原因的听损伤疾病，简单的应用放大装置不能取得令人满意的效果，而数字式助听器灵活的听力补偿功能较好的满足患者要求，目前数字式助听器已成为中重度神经性聋听力恢复的重要手段。

2.3 临床听力测试

临床上检查听力的一个重要方法叫做纯音测听法。纯音测听的具体方法是在一个安静的环境里，让受试者通过测试及听不同频率的声音，并告诉测试者是否听见了声音。将受试者能听到的各种频率声音的最小强度用曲线连接起来，就构成“听力图”。典型的听力图如图 2-3 所示，通常以图表的形式表示，图的纵轴代表听力计，横轴代表频率。且横轴上一个倍频程的宽度相当于听力级轴上 20dB 的距离。因为正常人耳对各频率纯音的听阈声压级不同，图中听力零级是通过测定一定数量耳科正常人群的各个频率听阈求出均数而得到的基准等效声压级值。

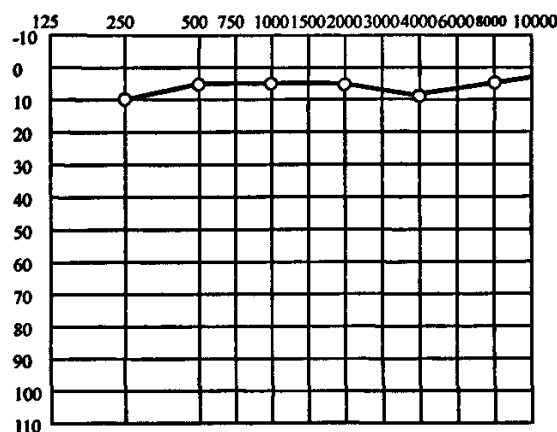


图 2-3 听力图

通常，采用国际标准组织（ISO）的标准来评价一个人听力损伤的成度。以一个人对于 500Hz，1000Hz，1500Hz 三个频率点的纯音听阈值的平均值作为评价依据。听力分级的标准如表 2-1：

表 2-1 听力分级表

平均听阈 HL (dB)	听力评价
≤ 25	正常
25~39	轻度听力损伤
40~54	中度听力损伤
55~69	中重度听力损伤
70~89	重度听力损伤
> 90	极度听力损伤

第三章 语音信号的特征

语言是人类在漫长的进化过程中逐渐形成的一种相互交流和沟通信息的方式,通过语音相互传递信息是人类最重要的基本功能之一。人类最重要、最精细的信息源只有语言、图像和文字三种。与用声音传递相比,显然用视觉和文字相互传递信息的效果要差得多,因为语音中除包含实际发音内容的语言信息外,还包括发音者是谁及喜怒哀乐等各种信息。语言是人类最重要、最有效、最常用和最方便的交换信息的形式。另一方面,语言和语音与人的智力活动密切相关,与文化和社会的进步紧密相连,具有最大的信息容量和最高的智能水平。所以,语音的采集、处理技术是数字助听器的研究重点。

3.1 语音基础

3.1.1 发音机理

语音是有声语言,是与人类同时发展进化而来的。在进化过程中,人类在语言的产生、理解、代表语言信息的声音接收上,形成了自己独特的器官:肺气管喉和嘴、唇、耳,从更深的意义上说大脑也应该包括在内。

人的发音器官由肺、气管、声门、咽腔、口腔和鼻腔等构成。口腔中又有非常灵活的舌头、上腭、软腭、上下齿和双唇等。上述一系列器官形成一条形状复杂的管道,称之为声道。声带位于喉的两侧,两声带之间的部分称为声门。

从物理学的角度而言,声音是由一定能量作用于物体,并使之振动所产生并通过媒质传播的波。当物体振动时,就会引起与之毗邻的媒质发生压力和质点速度等参量的变化。声音的源头是机械振动,发音时,气流吸入肺泡,经交换后,经声门由声带振动调制气流使成交变信号。再经过声道(口腔、鼻腔、咽腔)加以滤波形成不同的简正振动方式,由口和鼻辐射到空气环境中便形成语音。

3.1.2 语音的基本特性

声带的振动是发声的最重要的激励源，声带的张开和闭合使气流形成一系列的脉冲，这些脉冲的形状和周期与环境有关，这些脉冲的重复率称为基音。基音主要受声带张力控制，并受进入耳和脑的反馈语音调节。

声道在声学上是一截而小均匀的管子。通常声管一端为开口的，而另一端基本是闭合的。大约在管的中部有一分支，这就是鼻腔。这样的管子是一分布参数系统，它有许多自然谐振频率（即在这些频率上管子的传递函数具有最大值）。如果声道的截而是均匀的，则谐振频率的频率如下式：

$$f_n = \frac{(2n-1)c}{n} \quad n=1,2,3\cdots \quad (3.1.1)$$

在空气中声速 $c=340\text{m/s}$ ，而实际上的声道截而小是均匀的，这使谐振点之间的间隔也小相等，但声道谐振点的平均密度仍然大约为每 1kHz 有一个。这些谐振点称为共振峰，它是声道的重要声学特性。声门脉冲序列具有丰富的谐波成分，这些频率成分与声道的共振峰之间相互作用的结果对语音的音质有很大的影响。语音频谱中几个共振峰的绝对频率和相对分布对于人耳辨别各种音节起着重要的作用。

根据激励方式的不同可将语音分为以下三类：

- 浊音，当气流通过声门时，如果声带的张力刚好使声带发生振动，那么就能产生准周期的空气脉冲，这一空气脉冲激励声道得到浊音，如元音或浊辅音。浊音在时域上呈现出明显的周期性，在频域上有共振峰结构，而且能量大部分集中在较低频段内。
- 清音或摩擦音，使声道在某处发生收缩，同时迫使空气以高速冲过这一收缩部分而产生湍流，就得到这种音，如[f]；
- 塞音和爆破音，使声道完全闭合，在闭合后建立气压，然后突然释放，就得到爆破音，如[t]，[b]。

人的语音频率很广，一般认为人能听到的声音频率为 $20-20000\text{Hz}$ ，不同频率段有不同的特点。从声学角度看：

高频频段为 $6\text{kHz}\sim 20\text{kHz}$ ，这个频段的声音幅度影响音色的表现力。

中高频频段为 $600\text{Hz}\sim 6\text{kHz}$ ：这个频段是人耳听觉比较灵敏的频段，

它影响音色的明亮度、清晰度、透明度。

中低音频段为 200~600Hz: 这个频段是人声和主要乐器的主音区基音的频段。

低音频段为 20~200Hz: 这个频段如果比较丰满, 则音色会变得浑厚, 有空间感。

3.1.3 语音的人耳感知

语音的人耳感知特性主要表现在两个方面: 一个是声音的大小, 另一个是音调的高低。声音的大小由振动的幅度, 振动幅度越大, 声音就越大。声音的大小成为声音强度, 用“声压级 (SPL)”来表示, 单位是分贝 (dB)。声压级分贝值越高, 表示声音强度越大。声压与声压值的定义如下:

声压是指在声场中由于声波的存在, 媒质中的压力超过静压力的值。声压有正负值, 通常说的声压为其有效值。声压级表示的两个绝对声压的比值, 表达式为:

$$SPL = 20 \log \frac{p_1}{p_2} \quad (3.1.2)$$

其中, p_1 为当前声压, p_2 为基准声压, 这里的基准声压 (0dB SPL) 是 10^{-12}W/m^2 或 $20 \mu\text{Pa}$, 基准声压刚好是正常人对 1000Hz 纯音的听阈。

语音的另一个性质是音调的高低。音调是由振动的频率来决定, 音调的高低用“频率”来表示, 单位是赫兹 (Hz), 赫兹数越大, 频率越高, 声音越尖锐; 频率越低, 声音越低沉。人耳所能听见的频率范围, 是 20Hz-20000Hz。

正常人的各频率的绝对听阈和痛阈如图 3-1 所示, 正常人可听声音的强度范围为 0~120dB SPL (声压级)。纯音听阈又叫绝对听阈 (absolute threshold of hearing), 它是在无任何其他声音的情况下, 一个纯音刚被听到时的声压级。它是与频率有关的量。痛阈代表人耳可容忍的最高声压, 在声压级高到一定程度时, 耳朵会出现不适感觉, 或者具有痒、压迫及痛感。对正常人而言一般取 120dB 为不适阈, 140dB 为痛阈, 且认为它与频率无关。

绝对听阈可由非线性函数描述:

$$T_q = 3.64(f/1000)^{-0.8} - 6.5e^{-0.6(f/1000-3.3)^2} + 10^{-3}(f/1000)^4 \quad (dB SPL) \quad (3.1.3)$$

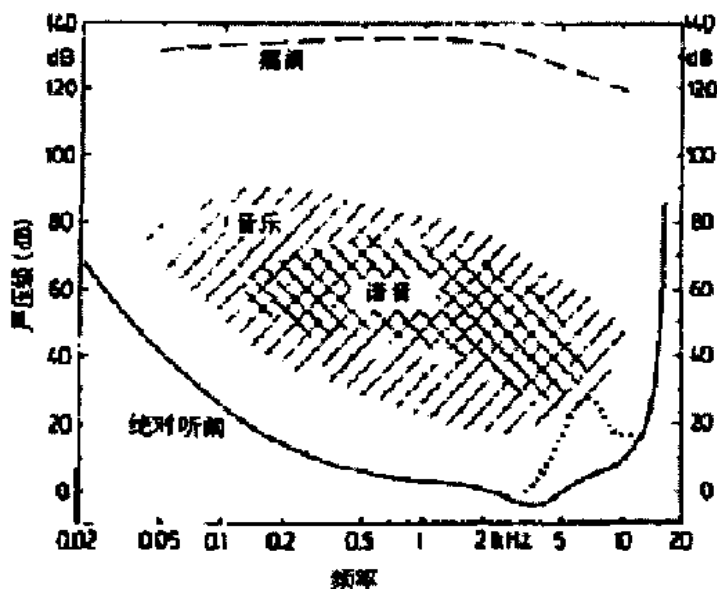


图 3—1 正常人的绝对听阈痛阈

表 3—1 语言能量%、语言可懂度%与语音信号频率之间的关系

频率 (Hz)	语言能量 (%)			可懂度 (%)	
60~125	5	60	95	1	5
125~250	13			1	
250~500	42			3	
500~1k	35	35		35	35
1k~2k	3	5	5	35	60
2k~4k	1			13	
4k~8k	1			12	

语音感知对语音信号研究有重要作用。人耳对语音的感知主要是通过语音信号频谱分量幅度获得的,对相位谱则不敏感。如表 3—1 所示语言能量、语言可懂度与语音信号频率之间的关系,低频 500Hz 以下部分约占 60

%的能量,但对于语言的可懂度帮不大,1000Hz 以下的低频虽占能量的 95%,但语言的可懂度仅有 40%。由于通常听障患者的高频损失较大,所以助听器的设计过程中必须提高高频部分的增益。人耳对频率高低的感受近似与该频率的对数值成正比。共振峰对语音的感知十分重要,特别是第二共振峰比第一共振峰更为重要。

3.2 数字语音处理

3.2.1 数字语音处理的发展现状

语音处理在早期基本上全部采用模拟电路,但是由于数字化形式的记录和处理方式具有无可伦比的灵活性和方便性的优点,在现代得到迅速的普及和发展。渐渐出现了数字化的专门芯片,从专用 DSP 芯片,到通用的 DSP 芯片,数字化的程度越来越高。对语音的处理同样遵循着电子设备通常的发展方向,即模拟到数字,专一到通用,硬件到软件。

通常的数字语音信号处理的实现方法一般有以下几种:

- 1) 在通用的计算机(如 PC 机)上用软件(如 Fortran.C 语言、Matlab)实现;
- 2) 在通用计算机系统中加上专门的加速处理机实现;
- 3) 用通用的单片机(如 MCS-51、96 系列、80x86 系列、ARM 等)实现,这种方法可用于一些不太复杂的数字信号处理,如数字控制领域;
- 4) 用通用的可编程 DSP 芯片实现。与单片机相比, DSP 芯片具有更加适合于数字信号处理的软件和硬件资源,可用于复杂的数字信号处理算法;
- 5) 用专用的 DSP 芯片实现。在一些特殊的场合,要求的信号处理速度极高,用通用 DSP 芯片较难实现,例如专用于 FFT、数字滤波、卷积、相关等算法的 DSP 芯片,这种芯片将相应的信号处理算法在芯片内部用硬件实现,无需进行编程。

在上述几种方法中,第 1 种方法的缺点是速度较慢,一般可用于基本

音频算法的模拟或者非实时的语音编解码；第 2 种和第 5 种方法专用性强，应用受到很大的限制，第 2 种方法也不便于系统的独立运行；第 3 种方法适用于 DSP 算法的开发，效率相对较低；第 4 种方法才使数字信号处理的应用打开了新的局面。

虽然数字信号处理的理论发展迅速，但在 20 世纪 80 年代以前，由于实现方法的限制，数字信号处理的理论还得不到广泛的应用。直到 20 世纪 70 年代末 80 年代初世界上第一片单片可编程 DSP 芯片的诞生，才将理论研究结果广泛应用到低成本的实际系统中，并且推动了新的理论和应用领域的发展，可以毫不夸张地说，DSP 芯片的诞生及发展对近 20 年来通信、计算机、控制等领域的技术发展起到十分重要的作用。

数字语音信号处理系统是以数字信号处理和通用 DSP 芯片为基础，因此具有数字处理的全部优点：

- a) 接口方便：DSP 系统与其他以现代数字技术为基础的系统和设备都是相互兼容的，与这样的系统接口以实现某种功能要比模拟系统与这些系统接口要容易很多；
- b) 编程方便：语音处理系统中的可编程 DSP 芯片，可使设计人员在开发过程中灵活方便地对软件进行修改和升级；
- c) 稳定性好：DSP 系统以数字处理为基础，受环境温度以及噪音的影响较小，可靠性高；
- d) 精度高：16 位数字系统可以达到 10^{-5} 的精度；
- e) 可重复性好：模拟系统的性能受到元器件参数影响性能变化比较大，而数字系统基本不受影响，因此数字系统便于测试、调试和大规模生产；
- f) 集成方便：DSP 系统的数字部件有高度的规范性，便于大规模集成。

目前数字化语音技术应用已经十分广泛，主要有语音编码、语音合成、语音识别、语音增强、说话人辨认、说话人确认、语音邮件、语音存储等。

3.2.2 数字语音的频率特性

模拟语音信号是连续信号，无法被 DSP 处理。因此语音处理的第一步

首先是将模拟信号转化为数字信号，也就是常说的模数转换，这其中主要包括两个步骤：采样和量化。

一个正常人语音的频率一般在 40Hz~4000Hz 的范围内，成年男子的语音频率较低，妇女和儿童的语音频率较高。在数字语音应用方面，用于电话通信的窄带语音频率范围为 60~3400Hz，带宽为 3.2K，采样率为 8kHz；用于远程会议的宽带语音频率范围为 50~7000Hz，带宽 7K，采样率为 16kHz；16 位声卡的采样频率为 22.05kHz；CD 的带宽为 20~20000Hz，采样率为 44.1kHz；数字音带（DAT）的带宽为 20K，采样率为 48kHz；现代的语音合成或识别系统中，需将语音频率的上限提高到 10kHz 左右。在信号处理时，通常处理的语音信号的频带在 100Hz~10000Hz，主要频段约为 100Hz~4000Hz。根据奈奎斯特采样定律（Nyquist Sampling Theorem），采样频率应为原始语音频率的两倍以上，考虑到滤波器性能的影响，这个阈值还应该提高。一般来说，普通语音的采样率在 15kHz~20kHz 左右。否则，如果采样率不满足采样定律，将会产生频谱混叠，使信号中的高频失真。

目前常见的音频信号的频率范围大致如下：8kHz、11.025kHz、16kHz、22.05kHz、37.8kHz、44.1kHz、48kHz 等。采样后的语音数据要能为 DSP 所存储和处理还必须进行量化处理。量化过程中，不可避免地会引入误差。量化时，如果采用较多的量化级数来记录样点的幅度，量化误差就较小，相应的比特（bit）数就会增多。但是这是以增加存储容量和处理时的计算量为代价的，因此必须根据应用场合合理地选择量化字长。常用的音频取样频率与量化精度如表 3-2 所示：

表 3-2 常用的语音采样频率与量化精度及数据率的关系

采样频率（kHz）	8	11.05	16	22.05	44.1	48
量化精度（bit）	8	8	16	16	16	16
数据率（kbps）	64	88.2	256	352.8	352.8	768

从上表可见，采样频率越高，数字化后的音频质量越高，但其数据量也越大，所以采样频率的选取要兼顾整个语音处理系统的信道容量和处理质量。

第四章 硬件平台

本文是基于 TI 的 TMS320VC5416 数字信号处理器完成的，下面对助听器算法开发的硬件环境作简要说明。

4.1 DSP 理论

4.1.1 DSPs 的发展

信号处理的发展随着时代的进步而进步，信号处理总是给芯片制造提出新的要求，芯片制造水平的提高反过来又推动信号处理向新的台阶发展。

自 80 年代以来，数字信号处理器（Digital Signal Processors, DSPs）发展突飞猛进。其运算的速度和计算能力，已经达到实时信号处理的要求。实时（real time）是指在实际运算中，输出的速度跟得上采样点输入的速度。对于某些运算（如滤波），这意味着每当接收到一次新的输入采样点时，就能产生一个新的输出。

DSPs 以其强大的功能获得开发者的青睐。DSPs 具有下面几个特点：

a) 哈佛结构

哈佛结构是不同于传统 PC 机所使用的冯·诺依曼（Von Neuman）结构的并行体系结构，其特点是 DSP 将程序和数据存储在不同的存储空间中，两者相互独立，独立编址，独立访问。与两个存储器相对应，系统中设置了两条总线：程序总线 and 数据总线，使数据的吞吐量提高一倍，系统的性能大大增强。

在哈佛结构中，由于程序和数据存储器在两个分开的空间中，因此取指和执行能完全重叠运行。为了进一步提高运行速度和灵活性，TMS320 系列 DSPs 在基本的哈佛结构上作了改进，一是允许数据放在程序存储器中，并被算术运算指令直接使用，增强了芯片的灵活性；二是指令存储在高速缓冲器（Cache）中，当执行此指令时，不需要再从存储器中读取指令，节约了一个指令周期的时间。

b) 流水线操作

DSP 芯片广泛采用流水线,以减少指令执行周期。TMS320 系列 DSP 处理器流水线深度从 2 到 6 级不等。流水线每个指令处于流水线上不同的阶段,使得像取指、译码、执行等操作可以独立的进行,使得指令的执行能完全重叠。在一个指令周期中,数条指令处于激活状态。只是当中断产生时,处于各级流水线上的指令才作废,重新开始新的流水操作。因此,一般情况下,流水线操作使 DSP 芯片的执行具有高效性。

c) 专用的乘法器

乘法是 DSP 算法的重要组成部分。对于滤波器算法中常用的每个抽头运算,都要做一次乘法和一次加法。这样,乘法的速度越快,DSP 处理器的性能就越高。在 TMS320 系列中,拥有专门的硬件乘法器是来实现单周期乘法,支持密集乘法运算,乘法可以在一个指令周期内完成,大大加快了运算速度。DSP 另外还有专门的多位累加器来处理多个乘积之和。极大提高了信号处理中经常使用的乘累加的运算效率。

d) 特殊的专用指令集

高度专用的指令系统以及优化的高级语言开发系统,使算法完成迅速。DSP 指令集是出于两个目的设计的:一是使处理器能够在每个指令周期内完成多个操作,从而提高每个指令周期的执行效率。二是将存贮 DSP 程序的存储器空间减到最小(由于存储器对整个系统的成本影响甚大,该问题在对成本敏感的 DSP 应用中尤为重要)。

e) 快速的指令周期

通过以上这些硬件及软件的优化设计,使得通常 DSP 芯片的指令周期在 200ns 以下。而 C5416 芯片的最快运行速度更是可以达到 160MIPS。

以上几个特点为开发复杂的助听器算法提供了硬件保障。

DSP 芯片制造商,目前有模拟器件(AD),德州仪器(TI),摩托罗拉(Motorola)等。DSP 芯片在中国市场上以往一直被国外所垄断。目前,国内自主研发了 DSP 汉芯二号、汉芯三号,依赖国外进口的局面被打破了。

TI 公司最早于 1982 年推出第一代 DSP 芯片 TMS32010 及其系列产品 TMS32011, TMS32C10/C14/C15/C16/C17 等,之后相继推出了 TMS32020,

TMS320C25/C26/C28, TMS32C30/C31/C32, TMS32C40/C44, TMS32C5x 以及集多个 DSP 于一体的高性能 DSP 芯 TMS32C80/C82 等。目前, TI 推出了 90nm, 1GHz 的 DSP, 标志着当今 DSPs 的发展水平。

DSPs 分为定点与浮点两种类型。一般在定点处理器中, 开发人员必须选择一种运算格式, 而且在整个过程中都必须注意格式问题。而浮点处理器在计算时几乎不需要考虑, 因此编写浮点 DSP 程序通常比定点程序简单。TI 的 C5000 系列支持定点运算, 定点处理器与浮点处理器相比, 优点在于完成同样的任务, 定点比浮点所需的运算时间少。另外, 定点处理芯片的成本比较低。

4.1.2 TMS320C54xxDSPs 概况

TMS320C54xx 是为实现低功耗、高性能而设计的定点 DSP 芯片, 属于 C5000 系列, 主要应用在通信系统方面。该芯片的内部结构及指令系统都是全新设计的, 它的主要特点是:

- 1) 运算速度快。指令周期为 25/20/15/12.5/10/6.25ns, 相应的运算能力为 40/50/66/80/100/160MIPS。其中 VC5416 指令周期为 6.25ns。
- 2) 优化的 CPU 结构。围绕 1 组程序总线、3 组数据总线和 4 组地址总线建立的改进型哈佛结构, 使得性能和多功能得以提高, 具有高度并行性和专业硬件逻辑的 CPU 设计, 使芯片性能大大提高。它内部有 1 个 40 位的算术逻辑单元, 2 个独立的 40 位的累加器, 1 个 17×17 的乘法器和 1 个 40 位的桶形移位器, 4 条内部总线和 2 个地址产生器。另外, 内部还集成了维特比加速器, 用于提高维特比编译码的速度。
- 3) 低功耗方式。TMS320C54xx 的主要特点是低功耗, 可以在 3.3V 或 2.7V 下工作, 有三个低功耗方式: IDLE1、IDLE2、IDLES 指令, 可以节省 DSP 的功耗。
- 4) 智能外设。除了标准的串行口和时分复用 (TDM 串行口外, 还提供了自动缓冲串行口 BSP (auto-Buffered Serial Port) 和与外部处理器通信的 HPI (Host Port Interface) 接口。BSP 可提供 2K 字数据缓

冲的读写能力,降低处理器的额外开销,当指令周期是 10ns 时, BSP 的最大数据吞吐量为 100Mbit/s,即使在 IDLE 方式下, BSP 也可以全速工作。HPI 可以与外部标准的微处理器直接接口。

- 5) 高度专业化的指令集,更适用于快速算法的实现和高级语言编程的优化。
- 6) 模块化结构设计理念,使派生器件得到更快的发展。
- 7) 高性能和低功耗的先进 IC 工艺技术,降低芯片功耗同时极大的缩小电路板的面积,使得 DSP 技术在耳级助听器设备中得到应用。
- 8) 新的静电设计结构从而获得的低功耗和增强的抗辐射能力。

笔者采用的 TMS320C5416 处理器在本系列中处于先进水平,是 TI 公司生产一款优秀的 16 位定点 DSPs,是 TMS320C54xx 系列中的第 3 代芯片,在单内核芯片中运算速度最快的一种。它本身具有丰富的片内存储资源和多种片上外设,在语音与通信领域得到广泛的应用。TMS320 系列中的同一代芯片具有相同的结构,但是片内存储器和片内外设的配置是不同的,通过把存储器和外围设备集成到一块片子上来,还可以降低系统的成本和节省电路板的空间, C5416 与该系列中其他部分芯片的具体比较如下表 4-1。

表 4-1C54xDSP 芯片配置比较

TMS320C54x	指令周期 (ns)	工作电压 (V)	片内 RAM (bits)	片内 ROM (bits)	串口	BSP	HPI
C541	20/25	5/3.3/3.0	5K×16	28K×16	2 个标准口		
C542	20/25	5/3.3/3.0	10K×16	2K×16	1 个 TDM	1	1
C543	20/25	3.3/3.0	10K×16	2K×16	1 个 TDM	1	
C545	20/25	3.3/3.0	6K×16	48K×16	1 个标准口	1	1
C546	20/25	3.3/3.0	6K×16	48K×16	1 个标准口	1	
C548	15/20/25	3.3/3.0	32K×16	2K×16	1 个 TDM	2	1
LC/VC549	10/12.5/15	3.3/2.5	32K×16	16K×16	1 个 TDM	2	1
C5402	10	3.3/1.8	16K×16	4K×16		2	1
C5416	6.25	3.3/1.6	128K×16	16K×16		3	1

由表可见, C5416 具有运算速度快,内部存储空间大,外部接口性能

好等优点。核心芯片 C5416 的低功耗特性又较符合在以电池为电源的助听器系统的能耗要求,所以笔者选择技术上比较先进,价格又较低廉的 C5416 作为硬件开发对象,并以北京瑞泰创新公司提供的基于 TMS320VC5416 的开发板 ICETEK-VC5416-C 进行改装后作为助听器的硬件开发平台。下面结合 C5416 的实际情况,介绍一下 C5416 的体系结构,并重点介绍内存的分配,这在编程及调试的过程中占有十分重要的位置。

4.1.3 TMS320VC5416DSPs 体系结构

4.1.3.1 总线结构

5416 体系结构围绕 8 组主要的 16 位总线(4 组程序/数据总线和 4 地址总线)构成,其中:

- 程序总线 (PB): 从程序存储器装载指令码和立即操作数。
- 3 组数据总线 (CB、DB 和 EB): 负责将片上的各个不同的部分相互连接,例如 CPU,数据地址产生逻辑,程序地址产生逻辑,片上外设和数据存储器。其中,CB 和 DB 从数据存储器读取操作数。EB 把操作数写到数据存储器。
- 4 组地址总线 (PAB, CAB, DAB, EAB): 负责装载指令执行所需要的地址。

PB 能加载保存于程序存储空间的操作数(如系数表)到乘法器和加法器进行乘/加操作或利用数据移动指令(MVPD 和 READA)把操作数移动到数据存储空间的目地址中。这种性能,与双操作数读取的特性一起,使 54xx 支持单周期三操作数指令。54xx 还有一条双向的片内总线用于访问片上外设,这条总线轮流使用 DB 和 EB 与 CPU 连接。对这组总线的访问,需要两个或更多的机器周期来进行读和写,具体所需周期数由片内外围设备决定。

4.1.3.2 存储器组织

5416 内部存储器被组织进三个独立的可选的空间:程序存储空间、

数据存储空间和 I/O 空间，大小都是 64K，总共是 192K 大小。C54xx 芯片都包括随机存储器 (RAM) 和只读存储器 (ROM)。其中，RAM 分为两种，双访问 RAM (DRAM) 和单访问 RAM (SRAM)。

1) 片内 ROM

C5416 包括 16K 的片内 ROM，内有一个引导程序，可以把用户代码调入程序存储器的任何一个位置。片内 ROM 的引导程序是芯片出厂的时候固化好的，用户一般不能直接使用。

2) 片内 RAM

C5416 有 128K×16bit 的片内 RAM，其中 64K×16bit 为片上双存取访问 RAM (DARAM)，DRAM 被组织在一些块上，因为每个 DARAM 块能够在每个机器周期中被访问两次，结合并行的体系结构，使得 C5416 得以在一个指定的周期内完成四个并发的存储器操作：一个取指操作、两个数据读操作和一个数据写操作。

C5416 的存储器映射寄存器被映射在数据存储空间。其中包括：

- a) 无等待状态访问的 CPU 寄存器 (共 26 个)
- b) 片上外设的控制和数据寄存器，存放在 0020h-005Fh 的地址中。
- c) 用于补充的 32 个字的 DARAM，这就不必将较大的 RAM 块分成小碎片。

下面重点介绍 C5416 几个重要的寄存器：

- a) 中断寄存器 (IMR、IFR)，地址：0h 和 1h 其中，中断屏蔽寄存器 (IMR) 可以个别的禁止或允许指定的可屏蔽中断。中断标志寄存器 (IFR) 可以指定当前的中断状态。
- b) 状态寄存器 (ST0、ST1)，地址：6h 和 7h。状态寄存器包含 C5416 的不同的状态和模式。其中，ST0 包括了算术运算和位操作使用的状态位 (OVA、OVB、C 和 TC) 及 DP 字段和 ARP 字段。ST1 反映了处理器和指令执行所依赖的模式和状态。
- c) 辅助寄存器 (AR0-AR7)，地址：10h 到 17h。共有 8 个 16 位的辅助寄存器可被 CPU 和辅助寄存器算术单元 (ARAUs) 修改。

C5416 提供了三个控制位用于在存储空间中配置片上存储器，利用这

三个控制位可以设置片上存储器怎样配置到不同存储空间，并指定是配置到程序存储空间还是数据存储空间。以上通过设置处理器模式状态寄存器（PMST）中的状态位，可以进行调整：

- MP/MC ：当此位是 1 时，禁止片上 ROM 配置到 5416 的程序存储空间中，即微处理器模式；当此位是 0 时，允许片上 ROM 配置到 5416 的程序存储空间中，即微计算机模式。
- OVLY：当 OVLY=1 时，片上 RAM 配置到程序和数据存储空间中；当 OVLY=0 时，片上 RAM 仅配置到数据存储空间。
- DROM：当 DROM=1 时，片上 ROM 配置到程序和数据存储空间。当 DROM=0 时，片上 ROM 配置到数据存储空间。DROM 与 MP/MC 状态无关

内存分配情况详见图 4-1：

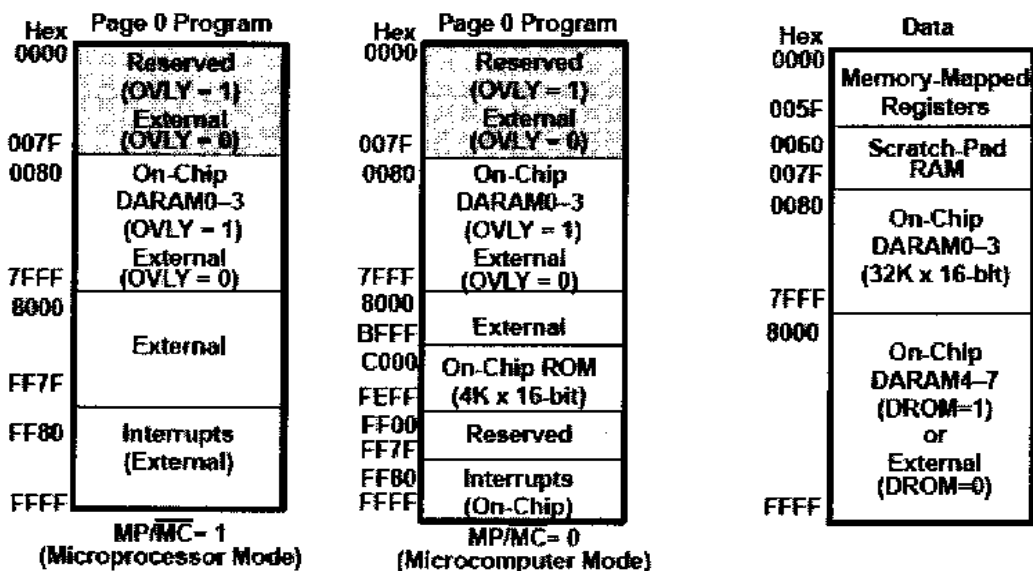


图 4-1 1C5416 存储器分配状况

3) I/O 空间

C5416 提供了 64K 的 I/O 空间，寻址范围是 0000H-0FFFFH，作用是与片外设备连结。PORTR 和 PORTW 两条指令可以访问这段存储空间。它适用于访问映射到 I/O 空间的设备而不是存储器。

4.1.3.3 中央处理器

C5416 芯片的 CPU 包括:

- 一个 40bit 的算术逻辑单元 (ALU)
- 两个 40bit 的累加器 (ACCA 和 ACCB)
- 一个 40bit 桶形移位器
- 17×17bit 的乘法器 (MAC)
- 40bit 加法器 (ADD)
- 比较, 选择和存储单元 (CSSU)
- 指数编码器
- 各种 CPU 寄存器 (CPU 寄存器由存储器映射的, 能够快速恢复和保存)

4.1.3.4 片上外设

C54xx 包含有相同的 CPU, 但是 CPU 连结不同的片上外设, C5416 包括的外设有:

- 通用目的输出引脚 (\overline{BIO} 和 XF)
- 软件等待状态发生器
- 可编程存储器切换逻辑
- 并行口
- 硬件定时器
- 串行口, 主要是多通道缓冲串行口

4.1.3.5 IEEE 1149.1 标准的逻辑扫描电路

片上的 JTAG 接口符合 IEEE 1149.1 标准。这个接口用于硬件仿真和测试。笔者所应用的开发套件中, 就包含了一个硬件仿真器。使用硬件仿真器, 可以缩短工程开发周期, 提高工作效率。

TI 提供的增强型 JTAG 连接, 是一种可与任意 DSP 系统相连的低侵扰式的连接, 使得 CCS 能够控制程序的执行, 实时监视程序运行。通过主机

和 DSP APIs 提供主机和 DSP 之间的双向实时数据交换, 能够使开发者实时连续的观察到 DSP 应用的实际工作方式。

4.2 DSP 开发环境

DSP 的硬件开发环境包括开发板、仿真器, 软件开发主要在 CCS 集成开发环境中完成。开发板选用的是北京瑞泰公司研制的以 TMS320VC5416 为核心的语音开发板 ICETEK-VC5416-C, 仿真器是瑞泰公司研制的 ICETEK-5100PP。

4.2.1 开发板与仿真器

开发系统的硬件资源分布如图 4-2:

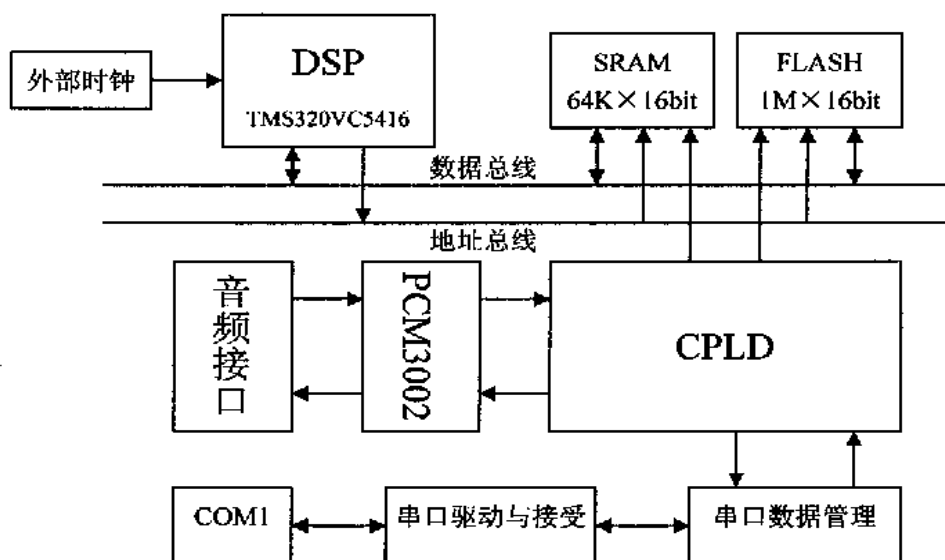


图 4-2 开发板硬件资源分布

ICETEK-VC5416-C 板所包含的硬件资源如下:

- 工作在 160MHz 的 TMS320VC5416 数字信号处理器芯片
- 片外扩展的只读存储器 (ROM) 共 128K×16bits
- 片外扩展的随机存储器 (RAM) 共 64K×16bits
- 系统自启动设计: 1M×16bits FLASH
- 立体声双路输入输出语音采集芯片 PCM3002

- 麦克风/扬声器接口
- RS232 串行数据接口
- 8 位可编程的发光二极管

笔者选用仿真器 ICETEK-5100PP 与开发板有着较好的硬件兼容性，其独特的优点如下：

- 兼容 TI 全系列 DSP 产品；
- 支持 Code Composer Studio2.2 集成调试环境；
- 仿真不占用任何 DSP 资源；
- 支持多 DSP 同时调试仿真；

4.2.2 模拟接口芯片的改进

4.2.2.1 AIC 芯片的特点

开发板采用的模拟接口芯片 AIC (Analog Interface Circuit) PCM3002 属于 TI 公司早期的产品，对于助听器的硬件设计要求而言，该片在功能与性能上显得有些不足，如无麦克接口，现有封装体积不够小，功耗偏大，采样率设置繁琐。为此我们改选用 TI 生产的新一款高性能立体声 AIC TLV320AIC23B，该芯片不但有着与 TI DSP 良好的硬件兼容，各方面的性能也明显提升了一级，目前广泛应用于语音设备。经过改进后的系统硬件框图 4-3 如下：

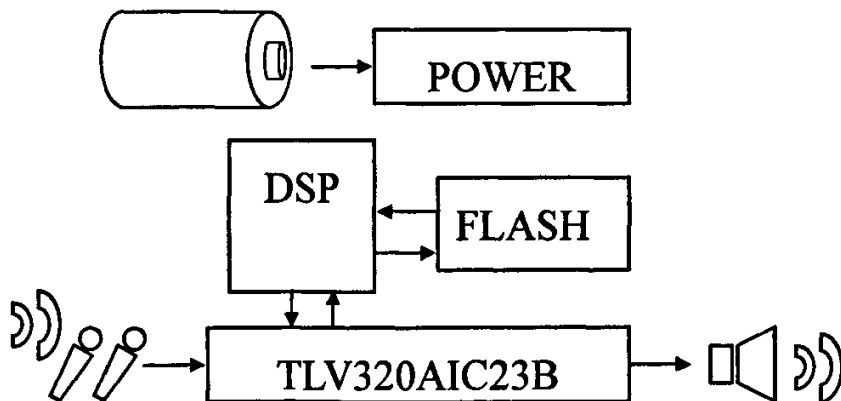


图 4-3 改进设计的助听器硬件开发系统结构

TLV320AIC23B 是 TI 公司近年新推出的低功耗 Σ - Δ 型 A/D、D/A 音频接口 (AIC) 芯片。模拟接口芯片 (AIC) 又称调制解调编解码器 (modem Codec), 以其高度可编程性, 高性能, 低功耗, 较少的外围器件等特点, 成为当前语音处理芯片的主流产品。该芯片工作方式、采样速率及通讯模式均可由编程设置。内部 ADC 之前有抗混叠滤波器, 之后有抽样滤波器, DAC 之前有插值滤波器。功能特点如下:

- 高性能, 采用先进的 Σ - Δ 过采样技术, 可以在 8K 到 96K 的频率范围内提供 16bit、20bit、24bit 和 32bit 的采样。ADC 和 DAC 的输出信噪比分别可以达到 97dB 和 100dB。
- 模拟接口丰富, 内置耳机输出放大器, 最大输出功率可达 30mW (耳机阻抗 32 欧), 支持 MIC 和 Line in 两种输入方式, 且对输入输出都具有可编程的增益调节。
- 数字接口灵活, 数字音频接口支持标准音频 I²S 以及专为 TI DSP 设计的 DSP Mode,。而配置接口采用 2 线制的 I²C 模式, 也支持 SPI 模式。无论是数字音频还是配置接口都可以和 DSP 的多通道缓存串口实现无缝连接。
- 功耗低, 回放模式下功耗仅 23mW, 省电模式下 <15uW。工作电压 2.7~3.6V, 与 TI C54xx 系列 DSP 工作电压保持一致。
- 芯片封装体积小, 外围元件少, 具有 28 脚的 TSSOP 封装和面积仅 25mm² 的 GQE 封装, 满足耳级助听器的设计要求。

4.2.2.2 AIC 与 DSP 的通信

AIC23B 与 C5416 的接口有两个, 一是数据接口用于输入输出数据, 另一个为配置接口, 用于寄存器初始化设置。AIC23B 的数据接口采用标准 I²S 或者专为 TI 的 DSP 设计的连接模式, 两者与 DSP 的 McBSP (Multi-channel buffered serial port, 多通道缓存串口) 连接相同, 连接如图 4-4:

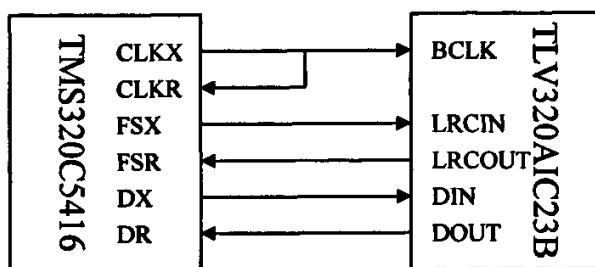


图 4-4.DSP 与 AIC23B 数字音频接口

其中 BCLK—数字音频接口时钟信号, 当 AIC 为从模式时(通常情况), 该时钟由 DSP 产生, AIC 为主模式时, 该时钟由 AIC 产生;

LRCIN—数字音频接口 DAC 方向的帧信号;

LRCOUT—数字音频接口 ADC 方向的帧信号;

DIN—数字音频接口 DAC 方向的数据输入;

DOUT—数字音频接口 ADC 方向的数据输出。

数字音频的时序如图 4-5, 4-6:

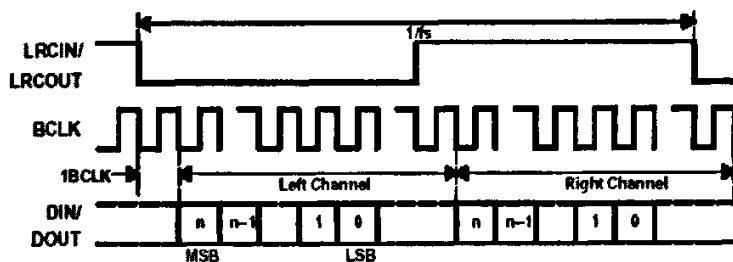
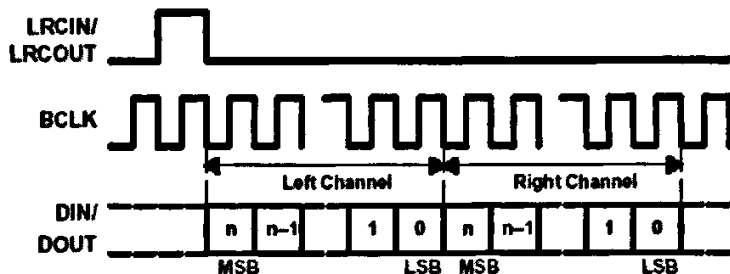

 图 4-5.DSP 与 AIC23B 数字音频接口 I²S 模式时序


图 4-6.DSP 与 AIC23B 数字音频接口 DSP Mode 模式时序

DSP 通过 McBSP 与 AIC 进行数据通信, DSP 接收音频数据过程采用中断的方式, AIC23B 每采集一次就发送数据, McBSP 每接受到一个 AIC23B

发来的数据包，产生中断，CPU 暂停主程序中，将读入的数据写入缓冲区内存后，继续主程序的进程。

AIC23B 的配置接口支持 I²C 模式，也支持 SPI 模式。通常比较简单的办法是利用 DSP 的一个 McBSP 用 SPI 模式跟 AIC23 连接。但是有些时候，如果 DSP 的 McBSP 串口资源比较紧张（比如需要跟近端 RS-232 和远端 RS-485 连接），也可以通过 DSP 模拟 I²C 总线与 AIC23 连接。采用 2 线制的 I²C 模式，因为 C5416 没有 I²C 接口，我们利用 DSP 的 GPIO（General Purpose Input/Output）来实现 I²C 时序。我们可以利用其中的两个管脚来作为 I²C 中的 SCLK 和 SDI。时序如图 4-7，4-8：

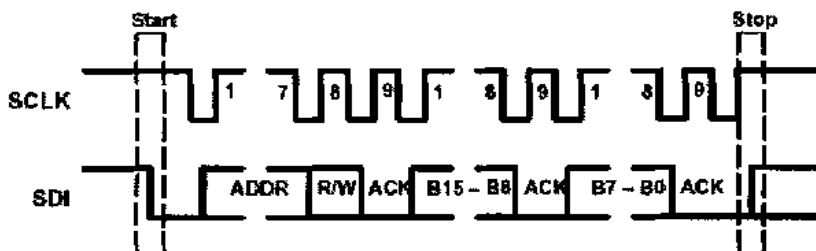


图 4-7. AIC23B 配置接口 I²C 时序

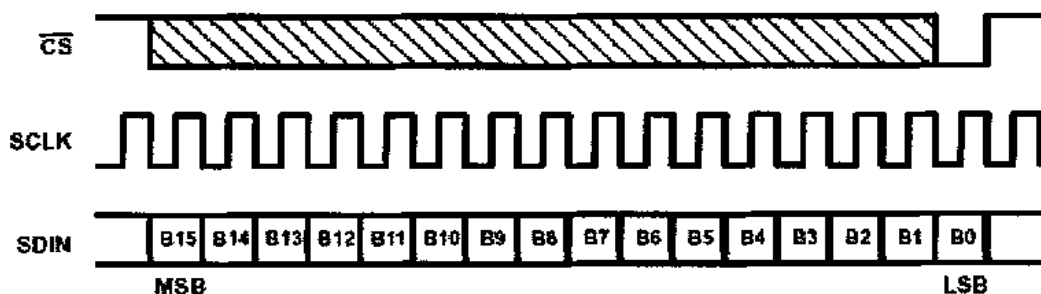


图 4-8. AIC23B 配置接口 SPI 时序

4.3 本章小结

在这章中，主要介绍了 DSP 芯片的基本结构和特征，C5416 的体系结构，以及 DSP 开发套件的硬件环境。片上资源决定了程序所能够用到的硬件环境，例如，开发板上外围辅助电路决定了信号输入输出的方式，扩展存储器接口可以适应更大程序代码的输入等。

我们还对硬件环境中的模拟接口芯片部分进行了改进，提出改进设计后的助听器开发系统结构。本章的重点是基于 DSP 可行性硬件方案的设计。

第五章 助听器语音处理方案

助听器可看作一个便携低功耗的声信号实时采集、处理、播放系统，其制造工艺、适配程序等都会影响助听效果，但起决定作用的是该系统采用的声处理算法。目前，人们对于听觉和听力损失的生理机制还不能提出肯定和精确的解释，这成为阻碍助听器发展的一个主要问题，如果解决了这个问题，助听器的性能将前进一大步。因此有专家预言，未来几年算法发展的重点可能集中在算法和生理基础的结合上，以及新算法的出现和应用上。

助听器信号处理算法可以说是一台助听器的“灵魂”，它为调整输入/输出特性和系统的频率响应特性提供了很强的灵活性。总的来说，这些算法要达到的最终目的有两个：

- 听觉最舒适
- 最高语言辨识度 (Speech Intelligibility, SI)

与一般语音算法相区别的是，助听器算法须以不破坏语音结构、不降低 SI 为前提。另外考虑到算法实时性和计算量的限制，算法设计成为了寻找性能和复杂度之间的平衡点的过程，这个平衡点的优化取决于硬件的不断进步和新算法的出现。此外考虑到助听器的实用需求，尽可能做到计算量小，存储量小，延迟小，是对算法的共同要求。

本章研究目前流行的各类处理算法，针对不足提出改进，确定适合本助听器的语音处理方案。设计助听器语音处理方案是本文的核心，也是 DSP 编程实现的理论基础。

5.1 短时平稳性预处理

语音是时变的、非平稳的随机过程，但是于语音的形成过程是与器官的运动密切相关的，人类发音系统的生理结构的变化速度是有一定限度的，这种物理运动比起声音振动速度来要缓慢得多，因此语音信号常常可假设为短时平稳的，即在 10~30ms 这样的时间段内，人的声带和声道形状是相对稳定的，其频谱特性和某些物理特征参量可近似地看作是不变的。因而

语音的短时谱具有相对稳定性，几乎所有的语音信号处理方法都是基于这个假设。这样，我们就可以采用平稳过程的分析处理方法来。

5.1.1 分帧处理

由于语音信号的准平稳特性，语音信号处理算法都建立在“短时”基础上，因此语音信号常被分段或分帧来处理。对于实时语音信号的处理，前一帧和后一帧的交叠部分称为帧移，帧移与帧长的比值一般取为 $0 \sim 1/2$ ，图 5-1 给出帧移与帧长的关系。这样做的好处是：在语音短时平稳允许的时间间隔内，增加可处理帧数，可减少语音分帧处理造成的不连续性。

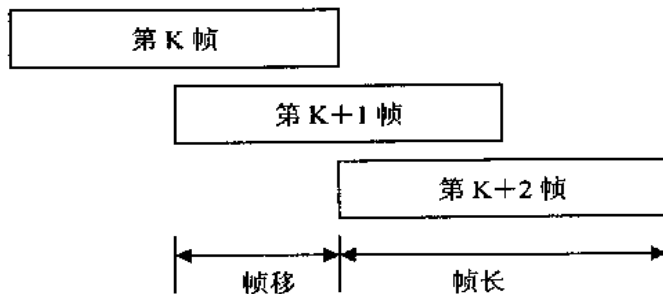


图 5-1 帧长与帧移的示意

5.1.2 窗函数的选择

一帧语音信号 $s(n)$ 还要经过加窗处理，形成加窗语音，

$$s_w(n) = s(n) \cdot w(n) \quad (5.1.1)$$

在语音信号的数字处理中，常用的窗函数有矩形窗、汉明（hamming）窗和汉宁（hanning）窗，其表达式如下（ L 为窗长）：

$$\bullet \text{ 矩形窗 } w(n) = \begin{cases} 1 & n = 0, 1, \dots, L-1 \\ 0 & n \text{ 为其它值} \end{cases} \quad (5.1.2)$$

$$\bullet \text{ 汉明窗 } w(n) = \begin{cases} 0.54 + 0.46 \cos \left[\left(\frac{2n}{L-1} - 1 \right) \pi \right] & n = 0, 1, \dots, L-1 \\ 0 & n \text{ 为其它值} \end{cases} \quad (5.1.3)$$

$$\bullet \text{ 汉宁窗 } w(n) = \begin{cases} 0.5 \left[1 - \cos(2\pi n / L) \right] & n = 0, 1, \dots, L-1 \\ 0 & n \text{ 为其它值} \end{cases} \quad (5.1.4)$$

一帧信号的选择实质上即是连续的语音信号乘以矩形窗的结构，为了减少频谱泄漏，我们选择的窗函数 $w(n)$ 满足两项要求：

- 窗谱主瓣尽可能的窄，以获得较陡的过渡带；
- 尽量减少窗谱的最大旁瓣的相对幅度，也就是能量尽量集中于主瓣，这样使肩峰纹波减少，增加阻带的衰减。

但是以上两项要求是不能同时满足的，往往我们是增加主瓣宽度以换取旁瓣的抑制。我们一般选择中间大两头小的光滑函数，这样的冲激响应所对应的滤波器具有低通特性。其带宽和频率就取决于窗函数的选择。以上三种常用的窗均有低通的特性，例如：汉明窗的频率幅度特性如图 5-2 所示，图中 b 代表主瓣宽度， a_1 代表旁瓣宽度。

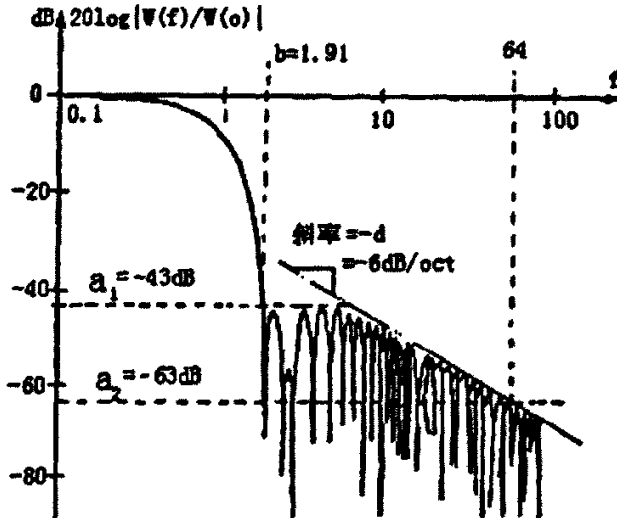


图 5-2 汉明窗的频率响应

表 5-1 为各种窗的主瓣宽度 (b) 和旁瓣宽度 (a_1) 的比较，我们可知矩形窗的主瓣宽度最小，具有较高的频率分辨率，但其旁瓣高度最高，使基音的相邻谐波之间的相互干扰比较严重，表现为相邻谐波间隔内有时叠加有时抵消，出现了一种随机变化的现象。相邻谐波之间的这种严重泄漏 (Gibbs)；汉明窗的主瓣最宽，而旁瓣高度最低，汉明窗得到的频谱要平滑得多。汉宁窗和汉明窗的差异在于前者随频率增加衰减很快，而后者基本保持一个常量。矩形窗的旁瓣太高，会产生严重的泄露现象，因此只在某些特殊场合中采用；汉宁窗衰减太快，低通特性不平滑；汉明窗旁瓣最

低，可以有效地克服泄露现象，具有更平滑的低通特性。汉明窗由于其平滑的低通特性和最低的旁瓣高度而得到最为广泛的应用。我们选用汉明窗来实现语音的预处理。

表 5—11s 长的各种窗的主瓣宽度 (b) 和旁瓣宽度 (a_1)

	矩形	汉明	汉宁
b	0.81Hz	1.91Hz	1.87Hz
a_1	-13dB	-43dB	-32dB

窗宽对短时频谱也会产生影响，比较如下：

- 1) 长窗具有较高的频率分辨率，但具有较低的时间分辨率。长窗起到了时间上的平均作用。
- 2) 短窗的频率分辨率低，但却具有较高的时间分辨率。
- 3) 在对语音信号进行短时傅立叶分析时，窗的长度的选择必须折衷考虑。选择提取语音信号的短时变化的短窗的同时，兼顾考虑语音信号频率分辨率。
- 4) 矩形窗和汉明窗的频谱特性都具有低通的性质，且在截止频率处都较尖锐。窗越宽其通带越窄。当其通带窄时，加窗后的频谱能够能够很好逼近短时语音信号的频谱。

长窗具有较高的频率分辨率，但具有较低的时间分辨率；短窗的频率分辨率低，但却具有较高的时间分辨率。对于同一种窗函数，主瓣宽度与窗长成反比，一般说来，窗长越长，它对信号的平滑作用越明显，如果想要反映变化快的信息，应该缩短窗长。在对语音信号进行短时傅立叶分析时，窗的长度的选择必须折中进行考虑。一方面，短窗具有较好的时间分辨率因而能够提取出语音信号中的短时变化，但同时却损失了频率分辨率。还应当注意到，语音信号的基音周期是有一个覆盖范围的，因此，窗宽的选择还应当考虑到这个因素。

5.2 宽动态压缩

在助听器的算法中，听力补偿算法是最基本，最重要的一种算法。

听力补偿算法的目的是对声音进行压缩放大，将正常人听力动态范围内的声音映射至聋人听域内，并尽可能地保持听觉舒适和提高声音的清晰度和辨识度。频谱中存在许多重要特征（如频谱包络、共振峰位置和相对强度等），对于言语识别具有重要意义。然而目前助听器产品中流行的多频段压缩放大方法几乎都忽略了对这些特征的保护，频段间处理的不衔接造成了语音结构变形、识别率降低。因此我们设计一种新的算法，希望在对不同频率成分进行非线性压缩放大的同时，尽可能保留较多的频谱特征信息。

5.2.1 压缩原理

动态范围是指人耳对某一种频率的声音或某类声音从刚刚听见到感到难以忍受的声强范围。动态压缩是数字式助听器最主要的信号处理方法，听力损伤者与正常人相比，在其损伤的频率范围内对声音的适应能力大大降低了如图 5-3 所示（THR 代表听阈；UCL 代表不适阈）。整个可听声音频域上的动态范围都缩小，其直接体现就是在这个频率范围内，正常人可以听到的较高或较低的声音。听力损伤者都无法听到。为了对这一部分频域进行补偿，使患者也能听到正常人能够听到的声音，同时为了满足患者对声音的动态接受要求，助听器设计者就提出了“压缩”的技术概念。

压缩放大是区别于线性放大而言，线性放大输入输出呈线性关系，线性放大技术通常应用在补偿传导性听力损失中。压缩放大为非线性放大，对于稳态信号，其输入输出函数曲线的斜率小于 1，助听器的增益随着输入信号强度的不同而自动的有所变化。感音神经性听力损失的动态范围变窄，要求助听器将强度范围为 30dB（50~80dB）的言语声压缩至听力损失者残余听力的动态范围内。也就是说给小声给予足够的放大，对于大声放大量要减小。这样输入信号较宽的动态范围就被压缩到输出信号较窄的范围之中，这就是“压缩”的内涵。正是由于数字信号处理技术的发展，不断提高了压缩的准确性，增加了多样性。

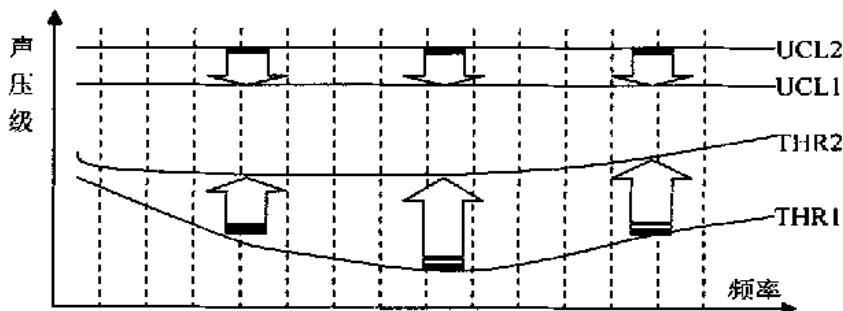


图 5-3 听力正常者与听力损伤者的听力指标对比

(下标 1 为正常人的听力指标 下标 2 为听力损伤患者的听力指标)

尽管人们已知人耳听觉是基于声音频率的敏感, 然而对听觉和听力损失的具体生理及听觉心理尚缺乏明确的解释, 因此目前通用的补偿算法在很大程度上仍然属于经验算法。

5.2.2 传统的多通道压缩放大算法

多通道压缩放大方法利用滤波器组将信号分频段处理后再综合, 利用多通道线路, 声音信号可以被分成数个独立的频率区域, 这些频率区域就被称之为通道。在各个通道内, 然后根据患者听力损伤的情况对于不同频段加以不同的放大处理, 对不同频率成分使用不同的压缩策略, 最后将合成后的声音再发放到患者的耳道中。多通道比单通道更有针对性, 使听力补偿更贴合患者听力损失情况, 且因为无须将信号在时域和频域间反复转换, 所以计算量相对较低。基本上目前所有的可编程式和数字式助听器都采用了多通道频响补偿的方案, 国外还出现过多达拥有 16 个通道的助听器。

多通道压缩方案的结构如图 5-4, 每个通道都有一个窄带滤波器, 提取本通道的信号, 然后信号经过压缩函数处理, 再用一个特征相同的窄带滤波器进行滤波处理。滤波器组的实现往往是通过数字有限冲击相应滤波器 (DFIR) 来实现。这是因为 FIR 滤波器具有线性相位特性, 而且稳定易实现。

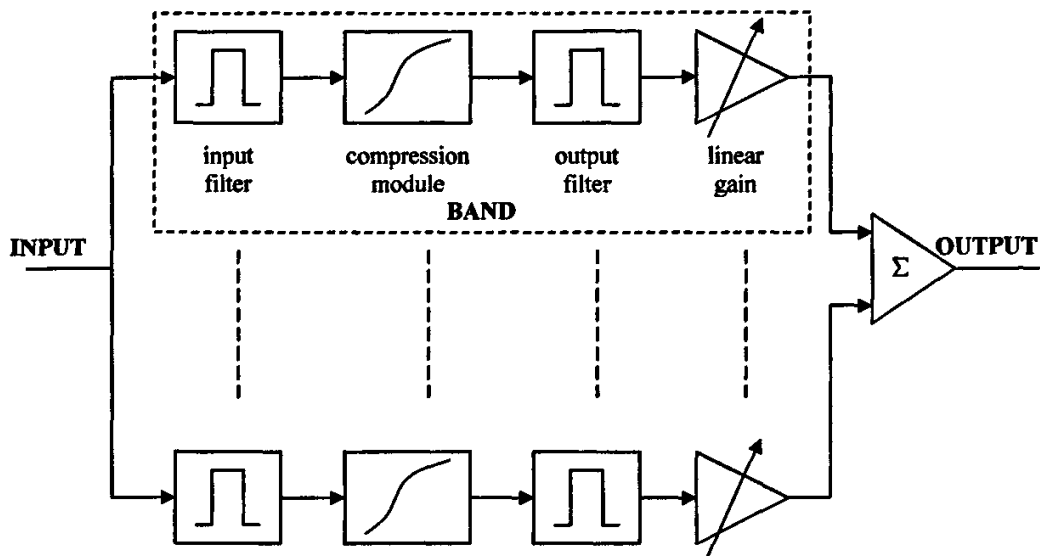


图 5-4 多通道压缩放大算法结构

然而该类方法频段交叠区的处理却带来了新的问题：

- 1) 当窄带信号经过多通道滤波器组进行放大时，由于助听器放大原理和叠加效果，会使得当信号频率远离交叠区时获得的放大倍数小于信号频率位于两频段交叠区时的放大倍数。如果按照半增益原理（认为增益值应约为听力损失的一半）计算，单位功率的正弦信号增益的差值将达到 3dB。且放大比越大，增益差值越大。这有可能使原有声谱波谷消失或人为制造出一个波峰，从而破坏频谱结构。频段越多，这样的变形就越多，语音失真就越严重。
- 2) 若共振峰正好位于频段交叠处（这是很有可能发生的），会被“劈”成两部分，分别被不同频段给予不同程度的放大。这有可能造成共振峰移动、变形，甚至消失。这些变化对信号获取和分析能力均弱于常人的耳聋患者来说是一个重大损失。

以上问题随着软硬件技术的不断发展和人们对助听器要求的不断提高而愈显重要。要解决这些问题，可以从 3 个方面着手。首先，有学者正在研究防拖尾滤波器（anti-smearing filters）并已取得了一些成果。其次，可以考虑将现有的固定频段滤波器组改为动态实时划分滤波器组，对每一帧信号实时划分频段以避开共振峰。这类方法对硬件系统存储空间和计算能

力都有较高要求,所以较少使用。第三,重新回到单频段方式,先找到频谱特征,并在压缩放大的过程中根据这些特征确定频谱增益。当然,此时的单频段方式在算法复杂程度和效果上已经远远超出早期的单频段方法。这种方式目前受到的关注最多,本论文也正是基于这类思路。

5.2.3 依频域宽动态压缩

压缩算法中最关键的一步即求出依从声压级的频率响应,我们所提出的宽动态压缩算法 WDRC (wide dynamic range compression) 是将整个言语动态范围按一定的比例均匀的完全压缩到患者的残余听力中。

5.2.3.1 助听器的频响模型

人的耳朵是一个生物滤波器,其作用也可以当成是一个线性移不变的系统,如图 5-5,设听力健康者的响应函数是 $C(j\omega)$, 输入声音信号 $a(j\omega)$, 通过人耳到达大脑后,输出信号为 $o(j\omega)$, 则有 $o(j\omega) = C(j\omega)a(j\omega)$ 。对于听障患者而言,耳朵的响应函数变为 $C'(j\omega)$, 且 $o(j\omega) \neq C'(j\omega)a(j\omega)$, 助听器的作用,在人耳外面级联一个线性移不变的系统,两个系统级联后的响应函数为 $N(j\omega) = C'(j\omega)H(j\omega)$ 。为了正确补偿患者的听力障碍,回复患者达到正常的听力,对于配戴助听器后,必须能满足 $o(j\omega) = N(j\omega)a(j\omega)$, 即 $C(j\omega) = N(j\omega)$ 。

我们可得到 $H(j\omega) = C(j\omega)/C'(j\omega)$, 如果用 dB 来表示响应函数则有 $|H(j\omega)|_{dB} = |C(j\omega)|_{dB} - |C'(j\omega)|_{dB}$, 其中 $|H(j\omega)|_{dB} = 20\log_{10}|H(j\omega)|$

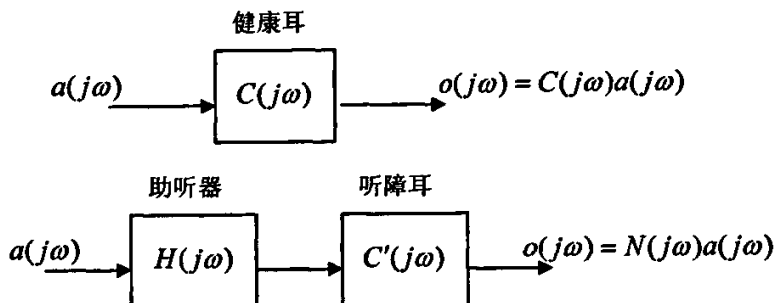


图 5-5 助听器与人耳响应函数模型

根据式 $|H(j\omega)|_{db} = |C(j\omega)|_{db} - |C'(j\omega)|_{db}$ ，我们如能够求得正常人耳和听障耳实时的频响函数，则就可以得到助听器的频响曲线。通过纯音测试，我们能够得到正常人和患者的听力图，也可以得到正常耳和听障耳的动态范围，从测得的数据即可得到正常耳与听障耳的响应函数，助听器的响应函数依从声压级的，整个频阈的增益补偿随着输入信号的声压级实时的变化。我们按照增益曲线修改语音信号的频谱，再变换到时域信号输出，那该助听器就能够有效的帮助患者达到听力补偿的目的。

5.2.3.2 频域增益的确定

对于宽动态压缩助听器而言，能够将输入信号的动态范围完全压缩到患者的残余听力范围内，而且保证输入声音的最适声压级，也能刚好映射到患者残余听力的最适声压级，既要保证助听器的高保真度，也提高使用的舒适性。我们的输入输出曲线如图 5-6 所示。

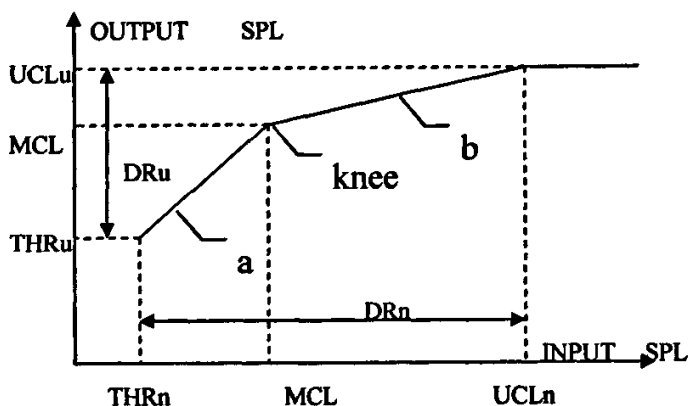


图 5—6 助听器 I/O 曲线

图中，横轴表示输入声压级，纵轴表示输出声压级，THR 表示听阈，MCL 表示最适级，UCL 表示不适阈，正常人耳的动态范围 DR_n 完全被映射在患者的动态范围 DR_u，为了保证正常的最适级 MCL_n 能够映射到患者的最适级 MCL_u，两者并不是满足线性对应的关系。I/O 曲线上对应的这一点被成为拐点 knee。拐点以下曲线 a 的压缩比 (CR_a) 拐点以上曲线 b 的压缩比 (CR_b)，我们可得下式：

$$CR_a = \frac{MCL_n - THR_n}{MCL_u - THR_u} \quad (5.2.1)$$

$$CR_b = \frac{UCL_n - MCL_n}{UCL_u - MCL_u} \quad (5.2.2)$$

由上图我们可得出，输入输出压缩的表达式为：

- 当 $SPL_{in} < THR_n$ 时， $SPL_{out} = 0$ ，助听器不启动压缩算法，配戴者感觉不到的声音；
- 当 $THR_n \leq SPL_{in} < MCL_n$ 时， $SPL_{out} = THR_u + \frac{SPL_{in} - THR_n}{CR_a}$ ，助听器启动 a 段压缩算法，配戴者开始感觉到较小的声音；
- 当 $MCL_n \leq SPL_{in} < UCL_n$ 时， $SPL_{out} = MCL_u + \frac{SPL_{in} - MCL_n}{CR_b}$ ，助听器启动 b 段压缩算法，配戴者开始感觉到较大的声音，当 $SPL_{in} = MCL_n$ 时， $SPL_{out} = MCL_u$ ，配戴者开始感觉到最舒适的声音；
- 当 $UCL_n \leq SPL_{in}$ 时， $SPL_{out} = UCL_n$ ，为了保护配戴者的听力，防止声音继续增强到痛阈，助听器启动削峰算法，输出保持稳定，配戴者感觉到的是持续处于不适阈的大声。

不同个体的在不同频率上的听力损失情况不同，则助听器在不同频率上的频响函数的参数也不同，但是我们无法得到患者在整个频谱上连续的响应函数，临床上进行听力测听时，通常测听几个特征频率点上的听阈，再将所得听阈用直线相连绘出被测者的听力曲线。通常纯音听阈测定选用声频倍频程频率 250、500、1000、2000、4000、8000Hz，如果在相邻两个倍频程频率之间的听阈相差超过 20dB 以上，则还加上测定另一组倍频程频率 750、1500、3000、6000Hz。在多通道动态压缩算法中，这些声频倍频

程频率也往往被选定为各通道的中心频率。因此我们对于频谱的补偿,也是通过得出特征点的增益值,我们选取 250、500、1000、2000、4000、8000Hz 为特征点,根据各频率点上的 I/O 曲线求出增益值,再以插值的方法求得特征点以外各点的增益值,原频谱上每个频点的值分别乘以增益序列相应的值即可得到补偿后的频谱序列。

插值具体算法如式 (5.2.3), 设 $G(\omega_n)$ 为各频率特征点 ω_n 增益值, 两个相邻特征点 ω_{n-1} 和 ω_n 之间的任意频率点 ω 上的增益值满足:

$$\frac{G(\omega) - G(\omega_{n-1})}{\omega - \omega_{n-1}} = \frac{G(\omega_n) - G(\omega_{n-1})}{\omega_n - \omega_{n-1}} \quad \omega_{n-1} < \omega < \omega_n \quad (5.2.3)$$

相邻特征点之间的所得各增益值满足等差数列。增益曲线表现为几段线段的首尾相连。

5.2.3.3 WDRC 算法结构

图 5-7 中, 虚线框部分即为在频域中完成压缩的过程, 我们采用频域补偿的方案与多通道动态压缩不同, 频域补偿避免了多通道滤波器之间存在频段交叠的现象, 而是对整个频域同时进行补偿, 处理后的频谱依然保持连续, 避免了语音的失真, 从而减少不适感, 提高语音识别率。利用 Fourier 变换, 将信号变换到频域, 对频谱进行补偿的方法更加灵活直观, 该思路也为下文所述的频域压缩算法提供了方便。

现代 DSP 的运算速度与精度大为提高以及 FFT 技术的应用, 成倍的降低了频域补偿算法所需运算时间。使得这种补偿算法能够在助听器这种实时语音处理系统中得到实现。

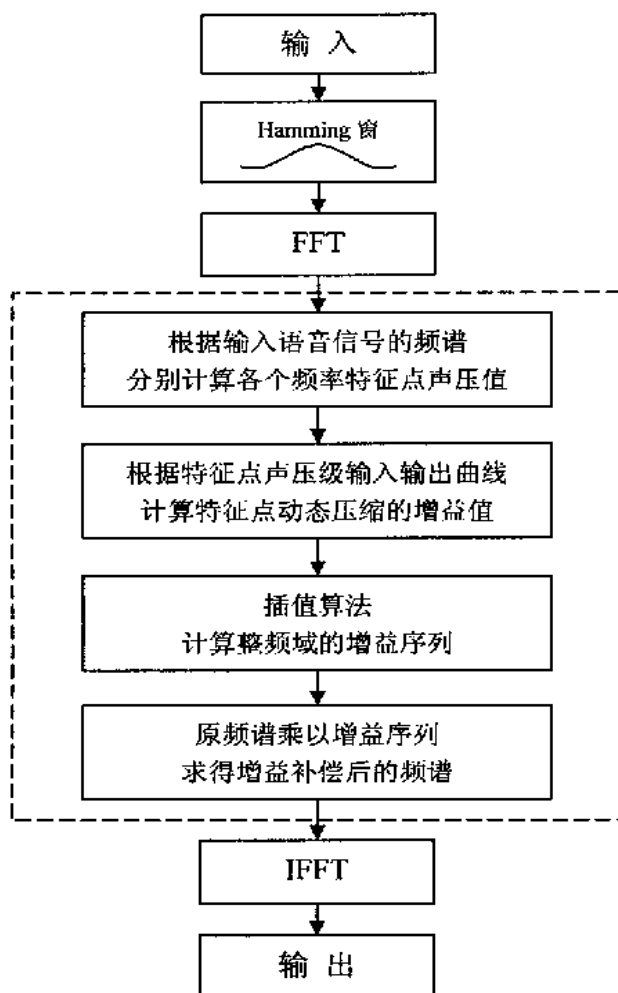


图 5—7WDRC 算法结构

5.3 自适应噪声消除

5.3.1 助听器噪声分析

人们在语音处理过程中不可避免地会受到来自周围环境、传输媒介引入的噪声、通信设备内部电噪声、乃至其它讲话者的干扰。这些干扰最终将使接收者接受到的语音已非纯净的原始语音信号，而是受噪声污染的带噪语音信号。研究表明，噪声影响，可使患者语言识别率大幅下降，因此去噪应该是助听器的一个重要功能。目前很多数字助听器产品中都带有相对独立的去噪部分。

我们所关心的噪声包括：周期性噪声、冲激噪声、宽带噪声、语音干扰等。周期性噪声的特点是有许多离散的窄谱峰，它往往来源于发动机等周期性运转的机械。电器干扰，特别是 50 或 60Hz 交流声也会引起周期性噪声。冲激噪声表现为时域波形中突然出现的窄脉冲，它通常是放电的结果。宽带噪声通常可以假定为高斯噪声和白噪声。它的来源很多，包括风、呼吸噪声和一般随机噪声源。量化噪声通常作为白噪声来处理，也可以视为宽带噪声。语音干扰可能是由于话筒拾得的其他语音引起的或传输时串话引起的。背景噪声能量主要集中在周期性噪声、宽带噪声上。只用一种方法就将它们很好的消除是不现实的，因此设计去噪算法时都有其针对性。考虑到对耳聋患者，不应降低语言识别率，所以人们只能希望通过去噪算法对于 SI 的促进，能够最大程度的克服算法使语音变形所带来的负面影响。

本文研究的主要是宽带噪声干扰。

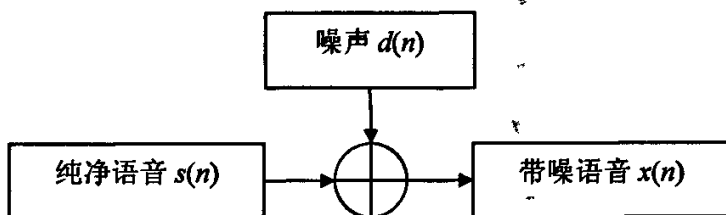


图 5-8 带噪语音模型

带噪语音的模型（见图 5-8）为：

$$x(n)=s(n)+d(n) \quad (5.3.1)$$

这里 $s(n)$ 和 $d(n)$ 分别代表纯净语音和干扰噪声。除此之外，还做如下假设：

- 1) 噪声是局部平稳的。局部平稳是指一段带噪语音中的噪声，具有和语音段开始前那段噪声相同的统计特性，且在整个语音段中保持不变。也就是说，可以根据语音开始前那段噪声来估计语音中所叠加的噪声统计特性。
- 2) 噪声与语音统计独立或不相关。
- 3) 系统内只有带噪语音可以利用，没有其他参考信号。

5.3.2 自适应滤波算法

5.3.2.1 自适应信号处理的特性

自适应信号处理 (Adaptive Signal Processing) 是近 40 年来发展起来的信号处理领域一个新的分支。随着适应信号处理的理论和技术日臻完善, 其应用的范围也愈来愈广泛。在包含语音信号处理的信息和信号处理学科中, 自适应信号处理是一个重要的分支学科。

自适应滤波理论和技术是统计信号处理和非平稳随机信号处理的主要内容, 它具有维纳滤波和卡尔曼滤波的最佳滤波性能, 但不需要先验知识的初始条件, 凭借测量信息不断在线估计即可自动调节系统参数完成最佳滤波, 它是通过自学习来适应外部自然随机环境, 因而自适应滤波器可以用来检测确定性信号, 也可以检测平稳的或非平稳的随机过程。自适应信号处理是研究一类结构可变或可以调整的系统, 它可以通过自身与外界环境的接触来改善自身对信号处理的性能。通常这类系统是时变的非线性系统, 可以自动适应 (即自适应) 信号传送变化的环境 and 要求, 无需详细知道信号的结构和信号的实际知识, 无需精确设计信号处理系统本身。

通过信号和信号的传递, 人们可以获取 (有用) 信息, 而经过对信号进行处理使蕴含于信号中的 (有用) 信息提取出来。由于语音信号形式及语音信号传送方式和传送环境的复杂性, 促使语音信号在传送过程中受到噪声和干扰的污染, 从而直接导致人们在用语音传递信息时无法获得信源方传递的所有有用信息, 所以只有对语音信号的处理, 才能提取出蕴含于信号中的 (有用) 信息。同时, 影响语音信号的噪声是多种多样的, 我们不可能了解每一种噪声的详细特征, 从而用普通的滤波方法消除噪声。所以, 对噪声的自适应处理成为人们有效地获取信息的关键。

5.3.2.2 LMS 自适应原理

图 5—9 显示了维纳滤波器的输入输出关系, 其输入是随机信号:

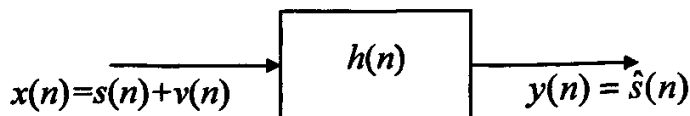


图 5—9 维纳滤波的输入输出关系

其中 $s(n)$ 表示信号的真值, $v(n)$ 表示噪声。其输出 $y(n)$ 等于 $s(n)$ 的估计值 $\hat{s}(n)$ 。维纳滤波器是具有这样的 $h(n)$ 或 $H(Z)$ 的滤波器, 它能使 $s(n)$ 与 $\hat{s}(n)$ 间的均方误差 $E[e^2(n)]$ 最小即

$$E[e^2(n)] = E[(s(n) - \hat{s}(n))^2] = \min \quad (5.3.2)$$

从而达到最好地从噪声中提取信号的目的。而自适应滤波器则能自动调节它的 $h(n)$ 值以满足上述最小均方误差的准则。

应用横向结构的 FIR 滤波器形式来实现自适应滤波是最常用的一种方法。如果 $h(n)$ 长为 N , 则可得到

$$y(n) = \sum_{m=0}^{N-1} h(m)x(n-m) = \sum_{i=1}^N h_i x_i \quad (5.3.3)$$

上式中 $i=m+1$; $h_i=h(i-1)$; $x_i=x(n-i+1)$

由此可见, 输出 $y(n)$ 是 N 个所有过去各输入的线性加权之和, 其加权系数就是 $\{h_i\}$ 。在自适应滤波器中这个加权系数常用符号 w_i 表示, 所希望的输出常用 d 表示。为了书写简化, 时间 n 用下标 j 表示, 于是上式成为

$$y_j = \sum_{i=1}^N w_{ij} x_{ij} \quad (5.3.4)$$

自适应滤波器可看成是自适应线性综合器, 如图 5—10 所示为最常用的横向 FIR 结构的自适应滤波器。

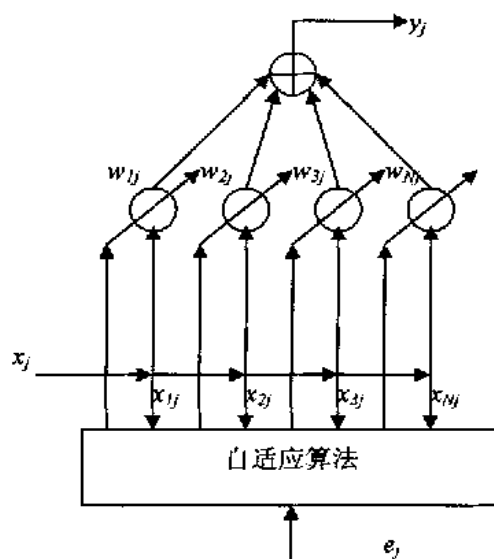


图 5-10 横向 FIR 结构的自适应滤波器

这种结构自适应滤波器的要害在于按照 e_j 及各 x_{ij} 值通过某种算法寻找 $E[e_j^2] = \min$ 时的各 w_j ，图中横向结构的自适应滤波器的权系数矢量和参考输入矢量为

$$\bar{W}_j = [w_{1j}, w_{2j}, \dots, w_{Nj}]^T \quad (5.3.5)$$

$$\bar{X}_j = [x_{1j}, x_{2j}, \dots, x_{Nj}]^T \quad (5.3.6)$$

式中 L 为自适应滤波器的长度。第 j 时刻线性综合器的输出以矩阵形式表达有

$$y_j = \bar{X}_j^T \bar{W}_j = \bar{W}_j^T \bar{X}_j \quad (5.3.7)$$

可以得到第 j 时刻自适应滤波器的误差信号

$$e_j = d_j - y_j = d_j - \bar{W}_j^T \bar{X}_j \quad (5.3.8)$$

显然， e_j 是一个随机变量，它有可能是正的，也有可能是负的。为求出最佳权系数，用它的最小均方值来表达误差，这就是最小均方误差 (LMS) 准则。采用这种准则作为最佳滤波准则的原因在于它的理论分析比较简单，不要求对概率的描述。同时，在这种准则下导出的最佳线性系统对其它很广泛一类准则也是最佳的。

我们可得到误差的均方误差为

$$E[e_j^2] = E[d_j^2] - 2E[d_j \bar{X}_j^T] \bar{W}_j + \bar{W}_j^T E[\bar{X}_j \bar{X}_j^T] \bar{W}_j, \quad (5.3.9)$$

$$\text{令 } \bar{P} = E[d_j \bar{X}_j] \quad (5.3.10)$$

$$\bar{R} = E(\bar{X}_j \bar{X}_j^T) \quad (5.3.11)$$

\bar{P} , \bar{R} 分别被称为 d_j 与 \bar{X}_j 的互相关矢量和参考输入 \bar{X}_j 的自相关矩阵。

将式 (5.3.10)、(5.3.11) 代入式 (5.3.9) :

$$E[e_j^2] = E[d_j^2] - 2\bar{P}^T \bar{W}_j + \bar{W}_j^T \bar{R} \bar{W}_j, \quad (5.3.12)$$

对于平稳输入, $E[e_j^2]$ 是权系数矢量 \bar{W}_j 的二次型函数。因此, $E[e_j^2] \sim \bar{W}_j$ 是一个上凹的超抛物面的曲面, 具有唯一的极小值点。可以用梯度方法沿着该曲面调节权系数矢量的各元素, 得到 $E[e_j^2]$ 的最小值。

均方误差的梯度 (用 $\bar{\nabla}_j$ 表示) 可以通过将式 (5.3.10) 对权矢量的各 w_j 进行微分得到

$$\bar{\nabla}_j = \left\{ \frac{\partial E[e_j^2]}{\partial w_1}, \frac{\partial E[e_j^2]}{\partial w_2}, \dots, \frac{\partial E[e_j^2]}{\partial w_N} \right\} = -2\bar{P} + 2\bar{R}\bar{W}_j, \quad (5.3.13)$$

置 $\bar{\nabla}_j = 0$ 就可得到最佳权矢量, 用 \bar{W}_j^* 表示, 即

$$-2\bar{P} + 2\bar{R}\bar{W}_j^* = 0 \quad \text{或者} \quad \bar{W}_j = \bar{W}_j^* = \bar{R}^{-1}\bar{P} \quad (5.3.14)$$

上式即是维纳-霍夫方程的矩阵形式。 \bar{W}_j^* 即为最佳权系数矢量。

$$(E[e_j^2])_{\min} = E[d_j^2] - \bar{W}_j^{*T} \bar{P} \quad (5.3.15)$$

自适应中的识别控制环节将输出 y_j 与所希望的值 d_j 比较, 看是否一样。如果存在误差 e_j , 则用 e_j 去控制 \bar{W}_j , 使 \bar{W}_j 为 $E[e_j^2] = \min$ 的 \bar{W}_j^* 。因此自适应的关键在与如何能简便的寻找 \bar{W}_j^* , 最常用的算法是所谓的最小均方 (Least Mean Square) 算法, 简称 LMS 算法。

Widrow-Hoff LMS 算法是求得最佳权矢量的一个简单有效的递推方法, 此法不要求相关矩阵, 也不涉及矩阵的求逆, 而是运用最优化的数学算法最陡下降法。

最陡下降法简单的推演过程如下, 由式 (5.3.10) 可知, $E[e_j^2]$ 是权系数矢量 \bar{w}_j 的二次型函数。 \bar{w}_j 是一个多维矢量, 因此 $E[e_j^2]$ 随着 \bar{w}_j 的变化可以画一个“碗形”的“曲面”, 我们称之为均方误差面或自适应滤波器的行为面。如图 5-11 所示。

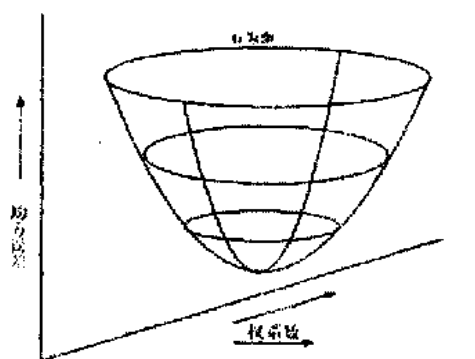


图 5-11 均方误差面

从非负二次均方误差函数的几何性质看来, 权向量 \bar{w}_j 从某一初值出发, 调整其自身之值使均方误差 $E[e_j^2]$ 达到最小, 相当于从均方误差面上的某点出发, 沿着此面下降直到“碗底”。这就可以使用梯度法来实现 \bar{w}_j 的调整。假定 $E[e_j^2]$ 对于 \bar{w}_j 的梯度 $\bar{\nabla}_j$ 是已知的, 令 \bar{w}_j 服从如下的递推方程:

$$\bar{w}_{j,t+1} = \bar{w}_j - \mu \bar{\nabla}_j \quad (5.3.16)$$

其中 μ 是正常数, 控制稳定性和收敛速度的参量, 称之为收敛因子。递推方程的含义是, 权向量在 $t+1$ 时之值等于它在 t 时之值加上一修正量, 后者正比于 $-\bar{\nabla}_j$ 。意味着, 在自适应过程中的任意时刻, 若总是沿着均方误差面最陡的方向下降的。由于均方误差面具有唯一的极小值, 采取这种下降策略在 μ 值选择适当时可使均方误差趋于其最小值或即使 \bar{w}_j 趋于维纳权, 不管初始值为何。控制 \bar{w}_j 变化步幅大小的量 μ 应适当选择。如果 μ 值过大, \bar{w}_j 未必收敛于 \bar{w}_j^* 。

然而, 在大多数需要应用自适应处理的场合, 输入信号的特性是未知

的或者随时间变化的。并不具备计算梯度 $\bar{\nabla}_j$ ，所需的知识，而且也不能测量，必须用 $\bar{\nabla}_j$ 的某种估计 $\hat{\bar{\nabla}}_j$ 代替。就是说

$$\bar{W}_{j+1} = \bar{W}_j - \mu \hat{\bar{\nabla}}_j \quad (5.3.17)$$

我们取单个样本误差的平方 e_j^2 的梯度作为均方误差梯度的估计。则有

$$\bar{\nabla}_j = \left[\frac{\partial e_j^2}{\partial w_{1j}}, \frac{\partial e_j^2}{\partial w_{2j}}, \dots, \frac{\partial e_j^2}{\partial w_{Nj}} \right]^T = 2e_j \left[\frac{\partial e_j}{\partial w_{1j}}, \frac{\partial e_j}{\partial w_{2j}}, \dots, \frac{\partial e_j}{\partial w_{Nj}} \right]^T \quad (5.3.18)$$

因为

$$e_j = d_j - \bar{W}_j^T \bar{X}_j \quad (5.3.19)$$

得到

$$\left[\frac{\partial e_j}{\partial w_{1j}}, \frac{\partial e_j}{\partial w_{2j}}, \dots, \frac{\partial e_j}{\partial w_{Nj}} \right]^T = -\bar{X}_j \quad (5.3.20)$$

所以有

$$\bar{\nabla}_j = -2e_j \bar{X}_j \quad (5.3.21)$$

我们可得到 Widrow-Hoff LMS 算法的权矢量递推方程组

$$\begin{cases} \bar{W}_{j+1} = \bar{W}_j - \mu \hat{\bar{\nabla}}_j = \bar{W}_j + 2\mu e_j \bar{X}_j \\ e_j = d_j - \bar{W}_j^T \bar{X}_j \end{cases} \quad (5.3.22)$$

这种算法对于每一个输入样本，只需对其进行两个乘法与两个加法运算，因此该算法适合用于实时系统实现。

5.3.2.3 自适应算法的收敛性质

为讨论 \bar{W}_j 的收敛过程，首先假设二次递推之间有充分大的时间间隔，以致可以认为二次输入信号 \bar{X}_j 与 \bar{X}_{j+1} 是不相关的，即

$$E[\bar{X}_j \bar{X}_{j+1}^T] = 0 \quad \text{当 } l \neq 0 \quad (5.3.23)$$

同时，由于 \bar{W}_j 仅是输入 \bar{X}_{j-1} ， \bar{X}_{j-2} ， \dots ， \bar{X}_0 的函数，故上述的假设导致 \bar{W}_j 与 \bar{X}_j 也不相关。

由方程组 (5.3.22) 得

$$\bar{W}_{j+1} = \bar{W}_j + 2\mu[e_j \bar{X}_j] = \bar{W}_j + 2\mu\{\bar{X}_j[d_j - \bar{X}_j^T \bar{W}_j]\} \quad (5.3.24)$$

由于 \bar{W}_j 是随机变量, 我们必须利用它的集合平均:

$$E[\bar{W}_{j+1}] = E[\bar{W}_j] + 2\mu E[d_j \bar{X}_j] - 2\mu E[\bar{X}_j \bar{X}_j^T \bar{W}_j] \quad (5.3.25)$$

考虑到 \bar{W}_j 与 \bar{X}_j 不相关, 故

$$E[\bar{X}_j \bar{X}_j^T \bar{W}_j] = E[\bar{X}_j \bar{X}_j^T] E[\bar{W}_j] \quad (5.3.26)$$

又考虑到 $\bar{P} = E[d_j \bar{X}_j]$ 及 $\bar{R} = E[\bar{X}_j \bar{X}_j^T]$, 代入式 (5.3.26), 得

$$E[\bar{W}_{j+1}] = [\bar{I} - 2\mu\bar{R}]E[\bar{W}_j] + 2\mu\bar{P} \quad (5.3.27)$$

设初试权矢量为 \bar{W}_0 , 则递推法可解得 \bar{W}_{j+1} 。有

$$E[\bar{W}_1] = [\bar{I} - 2\mu\bar{R}]E[\bar{W}_0] + 2\mu\bar{P}$$

$$E[\bar{W}_2] = [\bar{I} - 2\mu\bar{R}]E[\bar{W}_1] + 2\mu\bar{P} = [\bar{I} - 2\mu\bar{R}]E[\bar{W}_0] + 2\mu[\bar{I} - 2\mu\bar{R}]\bar{P} + 2\mu\bar{P}$$

...

$$E[\bar{W}_{j+1}] = [\bar{I} - 2\mu\bar{R}]^{j+1} E[\bar{W}_0] + 2\mu \sum_{i=0}^j [\bar{I} - 2\mu\bar{R}]^i \bar{P} \quad (5.3.28)$$

由于自相关矩阵是对称和正定的二次型矩阵, 总可以通过正交变换将其化成标准型

$$\bar{R} = \bar{Q}\bar{\Lambda}\bar{Q}^T \quad (5.3.29)$$

这里 \bar{Q} 是自相关矩阵 \bar{R} 的正交矩阵, 因此有

$$\bar{Q}\bar{Q}^T = \bar{I} \quad (5.3.30)$$

$\bar{\Lambda}$ 是由 \bar{R} 的特征值 λ_i 组成的对角矩阵

$$\bar{\Lambda} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_N \end{bmatrix} \quad (5.3.31)$$

将式 (5.3.29)、(5.3.30) 代入式 (5.3.27)

$$E[\bar{W}_{j+1}] = [\bar{Q}\bar{Q}^{-1} - 2\mu\bar{Q}\bar{\Lambda}\bar{Q}^{-1}]E[\bar{W}_j] + 2\mu\bar{P} = \bar{Q}[\bar{I} - 2\mu\bar{\Lambda}]\bar{Q}^{-1}E[\bar{W}_j] + 2\mu\bar{P}$$

$$E[\bar{W}_1] = \bar{Q}[\bar{I} - 2\mu\bar{\Lambda}]\bar{Q}^{-1}E[\bar{W}_0] + 2\mu\bar{P}$$

$$E[\bar{W}_2] = \bar{Q}[\bar{I} - 2\mu\bar{\Lambda}]\bar{Q}^{-1}E[\bar{W}_1] + 2\mu\bar{P}$$

$$= \bar{Q}[\bar{I} - 2\mu\bar{\Lambda}]^2\bar{Q}^{-1}E[\bar{W}_0] + 2\mu\bar{Q}\sum_{i=0}^1[\bar{I} - 2\mu\bar{\Lambda}]\bar{Q}^{-1}\bar{P}$$

...

$$E[\bar{W}_{j+1}] = \bar{Q}[\bar{I} - 2\mu\bar{\Lambda}]^{j+1}\bar{Q}^{-1}E[\bar{W}_0] + 2\mu\bar{Q}\sum_{i=0}^j[\bar{I} - 2\mu\bar{\Lambda}]^i\bar{Q}^{-1}\bar{P} \quad (5.3.32)$$

上式要收敛必须满足

$$|1 - 2\mu\lambda_{\max}| < 1 \quad (5.3.33)$$

$$\text{即 } \frac{1}{\lambda_{\max}} > \mu > 0 \quad (5.3.34)$$

即只有当 μ 大于零小于自相关矩阵 \bar{R} 的最大特征值的倒数时才有 $j \rightarrow \infty$, $[\bar{I} - 2\mu\bar{\Lambda}]^{j+1} \rightarrow 0$, 该算法收敛, 并保持稳定。

由于自相关矩阵 \bar{R} 的本征值 (包括 λ_{\max}) 通常是预先不知道的, 因此利用来确定 μ 的范围是不现实的, 然而我们知道矩阵 \bar{R} 的迹 $tr[\bar{R}]$ 等于 \bar{R} 对角线上的元素之和, 因此

$$tr[\bar{R}] = \sum_i E[\bar{X}_i^2] = \sum_i \lambda_i \quad (5.3.35)$$

由于正定矩阵的各 λ 值均大于零

$$\text{故 } \lambda_{\max} < tr[\bar{R}] \quad (5.3.36)$$

所以收敛的充分条件可以写成

$$\frac{1}{\sum_{i=1}^N E[\bar{X}_i^2]} \geq \mu > 0 \quad (5.3.37)$$

而 $\sum_i E[\bar{X}_i^2]$ 即为信号的总功率, 一般是已知的, 我们可以按照式

(5.3.37) 选取 μ 。

初始收敛速度、时变系统跟踪能力及稳态失调是衡量自适应滤波算法优劣的三个重要的技术指标。由于主输入端不可避免地存在干扰噪声，LMS 算法将产生参数失调噪声。干扰噪声越大，则引起的失调噪声就越大。减少步长因子 μ 可减少自适应滤波算法的稳态失调噪声，提高算法的收敛精度。然而步长因子 μ 的减少将降低算法的收敛速度和跟踪速度。即传统的 LMS 算法中步长因子 μ 较小时，算法收敛慢，但稳态失调误差小。 μ 取值在收敛范围内较大时，算法收敛速度快，但稳态失调误差大。因此，固定步长的自适应滤波算法在收敛速度、时变系统跟踪速度与收敛精度方面对算法调整步长因子 μ 的要求是相互矛盾的。

5.3.3 自适应消噪算法结构

在噪声干扰下进行语音通信经常会遇到对输入语音进行语音增强处理，提高语音通信质量的问题。目前应用的语音增强方法主要有基于噪声特性的自适应噪声抵消法、频谱减法、基于信号子空间处理的增强算法等。其中，带自适应滤波器的自适应噪声抵消法，是一种对含噪语音增强效果较好的算法。

对于宽带噪声，本文采用自适应噪声抵消法加以抑制。图 5-12 为自适应噪声消除算法原理图：

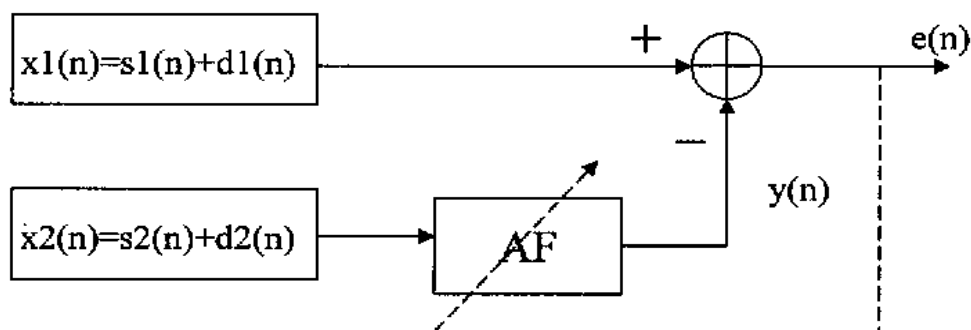


图 5-12 自适应算法原理图

我们采用两路语音输入 x_1 和 x_2 ，以 x_1 为输入信号，以为 x_2 参考输入信号，两路信号均包括两种成分噪声 d 和语音信号 s 。在这个系统中噪声为宽

带信号, 语音信号为窄带信号。由于 x_1 和 x_2 为同一个环境中采集的语音信号, 两路中的窄带信号 s_1 和 s_2 为强相关信号, 而宽带信号 d_1 和 d_2 彼此不相关。于是在 $e(n)$ 端输出中将是周期成分抵消只存在宽带成分, 而在 $y(n)$ 端输出只存在周期成分, 此时 AF 自动调节 \bar{w}_j 使其达到对周期成分起到选通的作用。

该噪声消除的递推公式如下

$$\begin{cases} y(n) = \sum_{i=0}^{N-1} w_i(n)x_2(n-i) \\ e(n) = x_1(n) - y(n) \\ w_i(n+1) = w_i(n) + 2\mu e(n)x_2(n-i) \end{cases} \quad (5.3.38)$$

其中 $i=0, 1, \dots, N-1$; $w_i(n)$ 为滤波器加权系数; $x(n)$ 由最近的 N 个采样点构成的序列; $y(n)$ 即为消噪后的语音信号。

μ 为收敛因子, 由于助听器系统对收敛的速度要求不太严格, 系统启动后, 自适应消噪算法就开始工作, 我们在满足 (5.3.37) 式后, 尽量减少 μ , 以减少自适应滤波算法的稳态失调噪声, 提高算法的收敛精度。

5.4 反馈消除

助听器扬声器发出的声音泄漏给麦克风 (泄漏可能和助听器与耳道的缝隙、助听器上的穿孔、助听器机械结构的振动传导有关), 可导致输出信号发生高强度的振荡 (即“啸叫”) 使得使用者不得不降低增益。通常反馈通路对低频声的衰减要大于助听器对其增益, 不易产生振荡, 因此振荡主要发生在高频段。当前的反馈消除主要有以下三种算法:

● 移相法

考虑到振荡产生的两个条件: 开环增益不小于 1, 相位延迟为 360 度的倍数, 提出最好能将相位延迟控制在 180 度附近, 这样可保持稳定的负反馈系统, 这一方法配合其他频谱幅度削减算法, 可大大提高其性能;

● 陷波法

增益和相位差都是造成振荡的条件, 但是在助听器中, 由于计算中的

延迟,造成反馈信号相位随频率变化很大,可认为相位作为频率的函数必在某些频率处为 360 度的倍数。这样,对稳定性的要求就转变为严格要求开环增益小于 1。因此人们考虑从增益入手,利用陷波滤波器(notch filter),对超出幅度阈值的波峰进行滤波:

- 自适应控制法

另一种思路是自适应反馈消除方法,是目前最流行的算法,以 FIR 滤波器,来估计声反馈通路 H 和反馈声,并将反馈声从输入声中减去。在实现这一思路时,不同的反馈声估计方法和不同的滤波器系数自适应更新方法,形成了不同的算法。

5.5 移频压缩

5.5.1 移频原理

传统助听器通过调控信号的振幅,使各频率的振幅变大,这种设计理念可以满足仅需要对声音进行简单放大的患者的需求。然而,最近的研究证明,高频听力损失超过 70dB 时,内毛细胞的功能已损失殆尽。此时病人即使可感知放大的高强度声刺激,但对语言理解不但没有作用反而有负面影响。造成该结果的原因是:

- 1) 传统助听器均采用振幅压缩放大原理,而其振幅放大量和频带宽度却受到麦克风和扬声器性能的限制,使其在高频尤其是 2kHz 以上的增益效果受限。仅采用振幅压缩放大原理的助听器,其助听效果不足以补偿高频听力损失超过 60dB 的病人对增益的需求。
- 2) 与声学上的上扩散掩蔽特性有关。上扩散掩蔽特性是一种低频声更易对高频声产生掩蔽作用的声学物理现象,因此现在已有很多助听器都采用了降低低频增益的方法来降低助听器的噪声。由于语言音素中元音的声谱多分布在低频,而辅音的声谱多分布在高频,如果助听器在中低频段有足够的增益而高频段增益不足,就会产生放大声音中元音对辅音的掩蔽。这也正是被国外许多研究证明的当听障人的高频听力损失超过 60dB 时,传统的助听器对听障者的言语辨

别力不但没有作用，反而有负面影响的原因所在。

- 3) 感音神经性聋病人需要更大（约比正常大4倍）的信噪比才能达到正常人在日常信噪比下的言语辨别力。正是由于以上的原因，就需要有一种助听装置能够在传统降噪的基础上，使语言中对言语辨别和理解至关重要的辅音音素的信噪比有进一步的提高。

大部分神经性耳聋的残余听力均集中在低频段100-1000Hz之间。如果一个正常全频声音进入耳蜗，根据行波理论，基底膜相应各段产生振动，但因为高频声—电转换神经元已不能发生脉冲，处于假象静止状态中，高频主要信息流中断，只有部分低频波能变成脉冲信号到达皮质，大脑无法完整地分析听觉信息，除了能觉出响度较大的部分低频声音外，自然就听不到高频声音了。临床有部分聋哑或严重耳聋者能听见打雷，火车轰鸣，鞭炮声也正是如此。

对于高频听力完全丧失的深重度听力损失患者，声强放大效果不明显，因此有学者提出移频算法和频率压缩算法，这种“移频”助听技术是一种避开增益限制和无用听力的信号处理策略，即按比例的频率压缩（proportional frequency compression），这样可以将关键的高频言语信息“移”到具有较好残余听力的低频区。用此技术有可能将输入言语信号的频带匹配到患耳最敏感的有限频带，而不是使已不存在的高频听力产生反应。这种改变并不影响语言的分辨，因为听辨语言最终靠大脑，大脑察觉频率是基于比率（ratios）而不是绝对的频率。实际上，可以说移频助听技术是一种介于传统助听器 and 电子耳蜗之间的助听技术。因此应用移频助听技术的助听器称为移频助听器（frequency transposition instrument）。

图5-13为是移频助听技术的按比例频率压缩工作原理图。a中两条曲线分别表示/f/和/s/的典型频谱图，阴影区代表助听器的频响区，与患者残余听力频率区一致。图5-13b中，将助听器的清辅音按比例压缩功能（FCVR）设置为2.25，则/f/和/s/的频率较初始音按比例压缩了2.25倍。由于两个声音的峰能量的关系不变，这样将放大的信号压缩到具有有效声增益且为病人有较好残余听力的频段（阴影区），可以听到和分辨高频的清辅音。只要峰能量关系不变，即使其绝对频率发生了改变，也可使听者识别其差异。

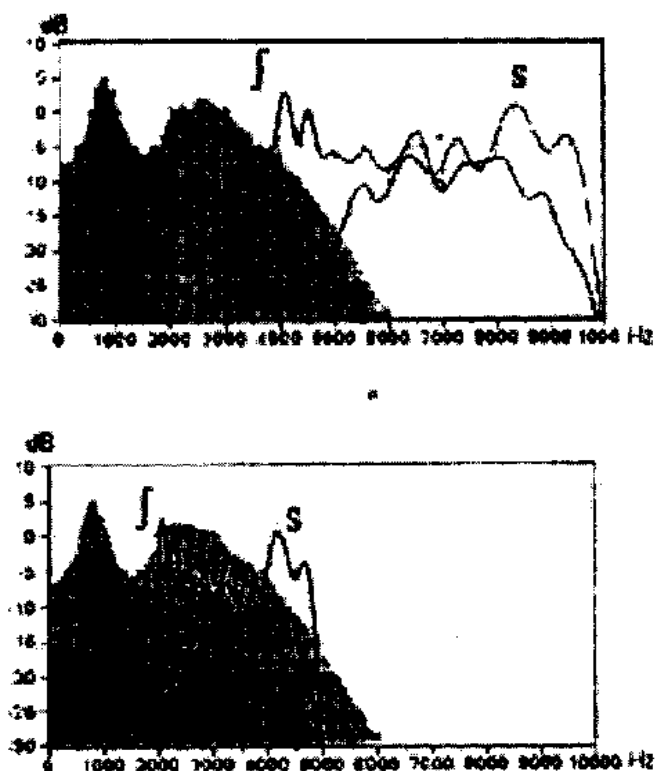


图 5-13 移频助听技术的按比例频率压缩工作原理图

(图 a 为辅音/sh/和/s/的典型频谱 图 b 为原始信号被向低频压缩 2.25 倍)

利用低频听觉细胞来听语频信息，必须满足两个条件：

- 信号信息流量必须经过压缩，因为残存的听细胞数量少，固有频率低。如果直将语频100-3000Hz的关键频段的信息量综合起来，是无法通过的。
- 必须使语频声波能振动基底膜低频段，也就是说必须使1000Hz以上的声波变成1000Hz以下的声波，使之携带的语言特征能被感觉到。

上述的按比例频率压缩功能只是一个具有选择功能的方式，故被称为动态言语重新编码（dynamic speech recoding）功能，它首先确定言语的特性，然后决定是否需要对言语声进行频率压缩。选择性按比例压缩器处理的设计可用于确保只改变语言的能量和高音（pitch）特性，而保持了语言的其他声学特性不变。

5.5.2 算法结构

移频和频率压缩算法与我们提出的宽动态压缩算法都通过频域处理进行补偿来实现的, 两者适用的患者群不同。算法结构如图 5-14:

我们以 $X(f)$ 表示信号的频谱函数, 需要压缩的原始频段为 $f_l \sim f_h$, 目标频段为 $f_{ol} \sim f_{oh}$, 则频域上的压缩比为 $p = \frac{f_h - f_l}{f_{oh} - f_{ol}}$ 。压缩后的频谱函数表

达式为

$$\begin{cases} X'(f) = 0 & f < f_{ol} \\ X'(f) = X[p * (f - f_{ol}) + f_l] & f_{ol} \leq f \leq f_{oh} \\ X'(f) = 0 & f_{oh} < f \end{cases} \quad (5.5.1)$$

由上式可知, 压缩过程中, 频谱特性完全按比例线性压缩到患者的残留听力的低频区域。而处于压缩范围以外的部分, 频谱都为零。

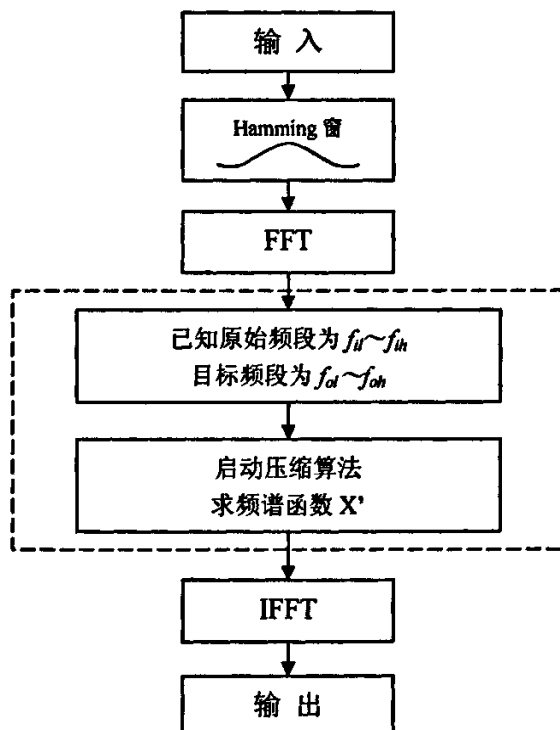


图 5-14 WDRC 算法结构

5.6 本章小结

在这章中，主要介绍了基于 DSP 助听器中应用的主要算法：宽动态压缩、自适应降噪、反馈压缩以及移频压缩。每种算法分别分析其在听觉生理学上的依据，算法原理，算法结构，以及具体算式。本章还在研究传统语音处理方案的基础上，针对原有算法存在的不足，而且提出基于频域宽动态压缩算法，并对算法的可行性进行论证。数字助听器信号处理算法的研究为下一章 DSP 编程实现和验证提供了理论依据。

本章的重点在基于频域宽动态压缩算法和自适应降噪算法的研究。动态压缩是助听器语音信号处理的核心，新的方案提出能够为国内基于 DSP 数字助听器的研究工作提供参考。

第六章 算法实现与验证

6.1 CCS 开发环境

在 DSP 上进行软件设计分为三个阶段：

- 1) 用汇编或 C 语言编写程序，再用 DSP 软件包中的汇编器等生成可执行的代码。
- 2) 用 PC 机上的 DSP 软件模拟器（Simulator）调试和验证程序及算法的功能。
- 3) 通过 PC 机以及 DSP 的仿真器和连接电缆对实际的 DSP 电路板进行在线仿真（ICE），直接对 DSP 电路的调试，

以上三个阶段都将在 DSP 调试环境代码编译器 CCS（Code Composer Studio）的帮助下完成，本节将介绍 CCS 的基本开发环境、软件开发过程、CCS 组件。CCS 提供了配置、建立、调试、跟踪和分析程序的工具，它便于实时、嵌入式信号处理程序的编制和测试，能够加速开发进程，提高工作效率。

6.1.1 CCS 概述

CCS 全称是：Code Composer Studio。它是德州仪器公司专为 TMS320 系列 DSP 开发的一个开发软件。CCS 在 WINDOS 环境下工作，类似于 VC++ 集成开发环境，它采用图形接口界面如图 6-1，整个窗口是包括主菜单、工具条、工程窗口、图形显示窗口、内存单元显示窗口以及寄存器显示窗口等。CCS 提供有编辑工具和工程管理工具，将基本的代码产生工具，如汇编器、链接器、C/C++编译器、建库工具整合为一个统一的开发平台。CCS 支持汇编语言、C/C++语言编程。能对 DSP 进行指令级的仿真和可视化实时数据分析，具有一定的调试、分析能力，极大地方便了 DSP 系统的软硬件开发。CCS 包括：CCS 代码生成工具；CCS 集成开发环境（IDE）；DSP/BIOS 插件程序和 API；RTDX 插件、主机接口和 API 等，下面将着重

介绍前两项。

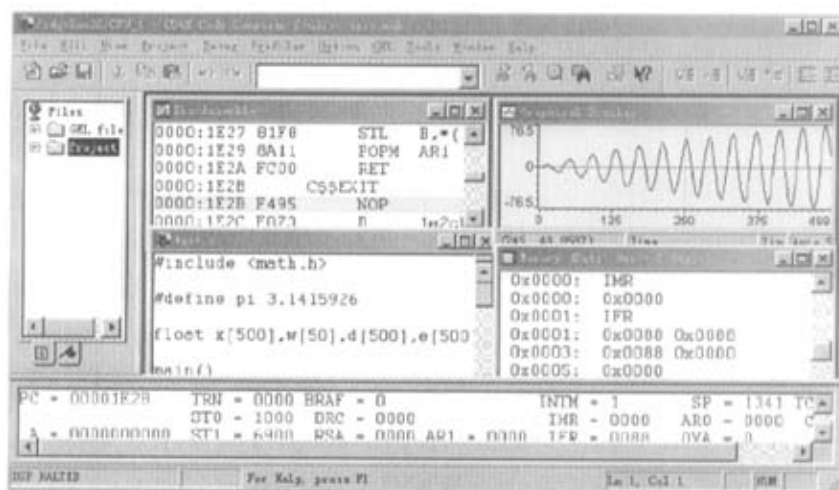


图 6-1 CCS 图形编程界面

6.1.2 代码生成工具

代码生成工具奠定了 CCS 所提供的开发环境的基础。典型的软件开发流程如图 6-2:

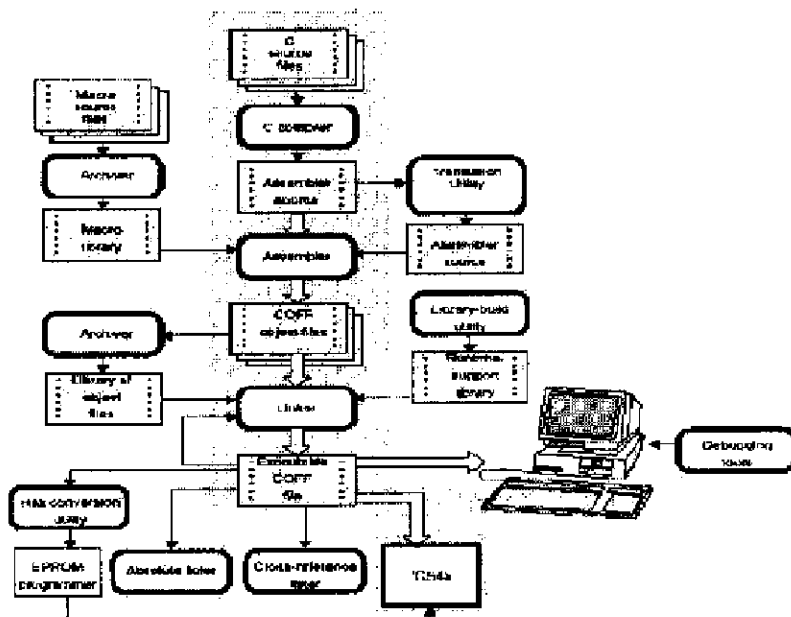


图 6-2 CCS 软件开发流程

图中阴影部分是通常的开发流程，其他部分是附加功能。基本流程的描述是：代码的开发主要分为代码生成和代码调试两个阶段，C 源代码和汇编源代码分别通过 C 编译器和汇编编译器生成公用目标文件格式（COFF）的代码，再通过连接器根据内存分配文件（如：recipe 文件或 cmd 文件），生成可执行的输出文件。之后就可以输出到目标板进行调试和功能实现了。

主要应用到的工具包括：

- 1) C 优化编译器（C compiler），将标准 C 语言源文件编译成为相应片种的汇编语言源代码文件，进而汇编和连接，产生可执行的目标文件，这些 C 编译器完全符合 ANSI C 标准。
- 2) 汇编器（assembler），将汇编语言源程序文件汇编成机器语言的目标程序文件，其格式为 COFF。汇编语言源程序可以包括汇编语言指令（Instruction）和宏指令（macro directives）。汇编器的输入文件为汇编语言源文件，其缺省的文件扩展名为“.asm”由汇编器所建立的目标文件，其缺省的文件扩展名为“.obj”。用汇编器可以建立列表文件，其缺省的文件扩展名为“.lst”。
- 3) 连接器（linker），连接器的基本任务是将目标文件连接在一起，产生可执行模块。连接器可以接受的输入文件包括汇编器产生的 COFF 目标文件、命令文件、库文件以及已部分连接好了的文件。它所产生的可执行 COFF 目标文件可以装入各种开发工具，或由 DSP 来执行。根据用户要求，该连接器还可以建立一个连接映射列表，用来描述存储器的分配、输入和输出程序断的位置、以及外部符号重新定位后的地址。
- 4) 归档器（archiver），可以把一组文件收集到一个归档文件中，这在长时间的调试时是十分有用的；运行支持库（runtime_support libraries），包括了 C 编译器所支持的 ANSI 标准运行函数、编译器公用程序函数、浮点运算函数和 C 编译器支持的 I/O 函数。
- 5) 十六进制转换公用程序（hax conversion utility），可以把 COFF 目标文件转换成 TI-Tagged、ASCII-hex、Intel、Motorola-S、Tektronix

等目标格式，之后把转换好的文件下载到 EPROM 编程器中。

6.2 总体结构

软件采用 C 语言和 DSP 汇编联合编译，体现了模块化的思路，提高程序调试的效率。如图 6—3 程序总体结构所示，整个程序包括初始化模块、语音预处理模块、傅里叶变换模块、宽动态压缩模块、移频压缩模块、自适应消噪模块和中断子程序。其中宽动态压缩和移频压缩模块在程序中同时只能运行其中一个。

我们采用的是将连贯的信号按帧处理的方案，主程序一直处于循环处理过程中，在处理完一帧信号后，等待下一帧接收完毕，重新循环。因此主程序的对一帧数据处理时间必须小于一帧数据接收的时间，才能够保证输出信号也是连贯。在系统采样率和帧长帧移的选取上必须兼顾以上原则，避免进程出错。系统中语音信号的采样率定为 16kHz，量化精度 16bit，数据率 256kbps，帧长为 256，帧移为帧长的 1/2。

系统中语音信号处理部分，我们主要实现了宽动态和移频两种压缩算法以及自适应消噪算法。DSP 接收模拟接口芯片发来的数据采用中断的方式，模拟接口芯片每一个采样点都会产生中断，主程序就转入中断子程序，因此中断的频率与采样率是一致的。中断子程序中 DSP 每接收一个数据到输入数据缓冲区，同时也从处理后的输出缓冲区中发送一个数据到模拟接口芯片。

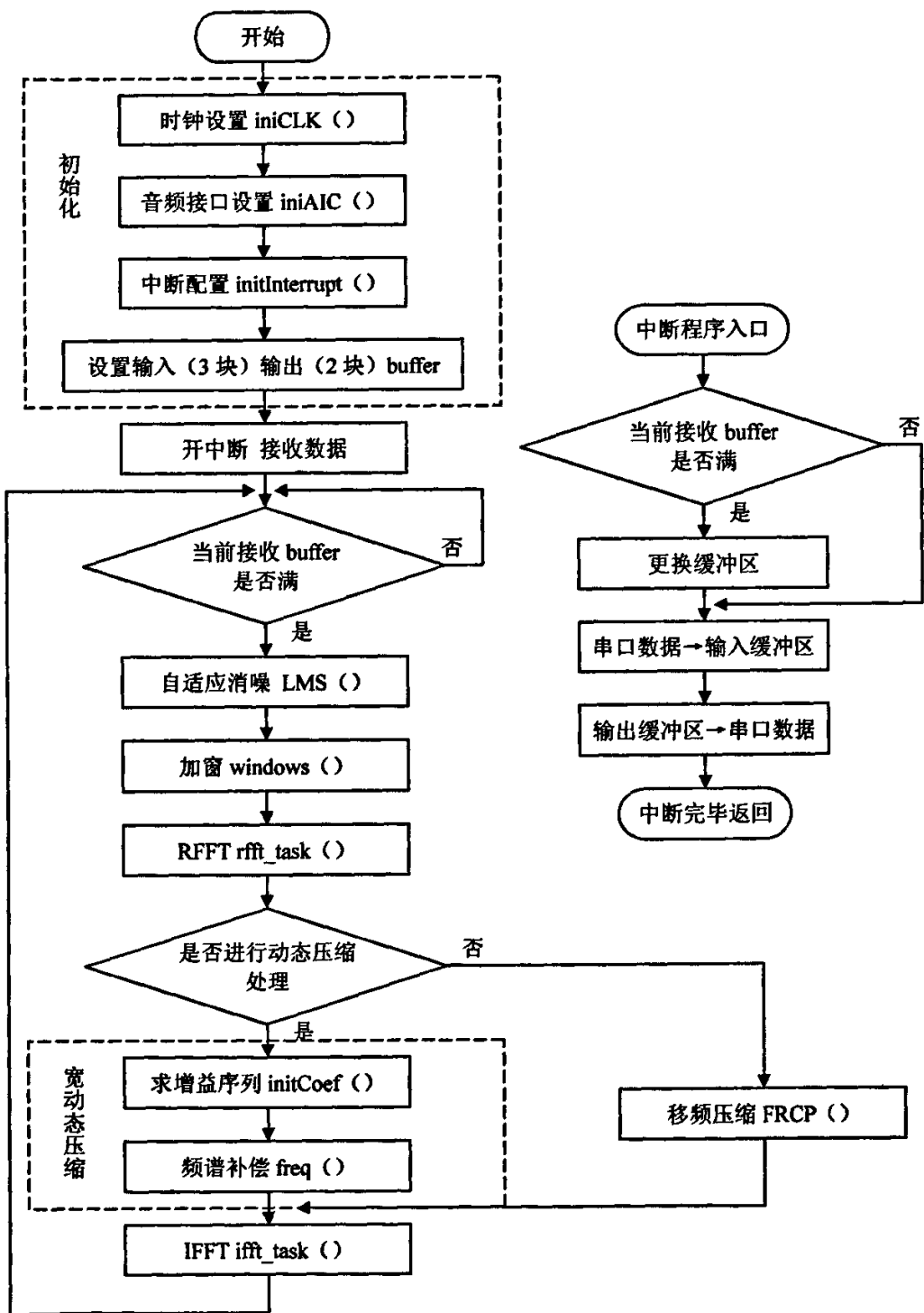


图 6-3 程序总体结构

6.3 RFFT 与 IFFT

动态压缩算法和移频压缩算法都是通过改变信号频谱特性来实现，在程序中 FFT 和 IFFT 成为最基本的模块。在传统的基 2FFT 算法的基础上，我们用汇编语言完成高效的 RFFT 和 IFFT。

6.3.1 时域抽取法 FFT

FFT 算法基本上分为两个大类：时域抽取法 FFT (Decimation-In-Time FFT)，简称 DIT-FFT 和频域抽取法 FFT (Decimation-In-Frequency FFT)，简称 DIF-FFT。考虑到 TI DSPs 支持位倒序的汇编指令，该指令极大能够提高 DIF-FFT 的运算效率，系统中我们采用 DIT-FFT 算法。

DIT-FFT 的基本原理如下：

设序列 $x(n)$ 的长度为 N ，且满足

$$N=2^M, M \text{ 为自然数}$$

按 n 的奇偶把 $x(n)$ 分解为两个 $N/2$ 点的子序列 $x_1(r)$ 和 $x_2(r)$

$$x_1(r)=x(2r) \quad (6.3.1)$$

$$x_2(r)=x(2r+1) \quad (6.3.2)$$

$$n=0,1,\dots,N/2-1$$

则 $x(n)$ 的 DFT 为

$$\begin{aligned} X(k) &= \sum_{n=\text{偶数}} x(n)W_N^{kn} + \sum_{n=\text{奇数}} x(n)W_N^{kn} = \sum_{r=0}^{N/2-1} x(2r)W_N^{2kr} + \sum_{r=0}^{N/2-1} x(2r+1)W_N^{k(2r+1)} \\ &= \sum_{r=0}^{N/2-1} x_1(r)W_N^{2kr} + W_N^k \sum_{r=0}^{N/2-1} x_2(r)W_N^{2kr} \end{aligned} \quad (6.3.3)$$

由于

$$W_N^{2kr} = e^{-j\frac{2\pi}{N}2kr} = e^{-j\frac{2\pi}{N}2kr} = W_{N/2}^{kr} \quad (6.3.4)$$

所以

$$X(k) = \sum_{r=0}^{N/2-1} x_1(r)W_{N/2}^{kr} + W_N^k \sum_{r=0}^{N/2-1} x_2(r)W_{N/2}^{kr} = X_1(k) + W_N^k X_2(k) \quad (6.3.5)$$

$$k = 0, 1, \dots, N-1$$

其中 $X_1(k)$ 和 $X_2(k)$ 分别为 $x_1(r)$ 和 $x_2(r)$ 的 $N/2$ 点 DFT, 即

$$X_1(k) = \sum_{r=0}^{N/2-1} x_1(r)W_{N/2}^{kr} = \text{DFT}[x_1(r)] \quad (6.3.6)$$

$$X_2(k) = \sum_{r=0}^{N/2-1} x_2(r)W_{N/2}^{kr} = \text{DFT}[x_2(r)] \quad (6.3.7)$$

由于 $X_1(k)$ 和 $X_2(k)$ 均以 $N/2$ 为周期, 且

$$W_N^{k+\frac{N}{2}} = -W_N^k \quad (6.3.8)$$

所以 $X(k)$ 又可以表示为

$$\begin{cases} X(k) = X_1(k) + W_N^k X_2(k) \\ X(k + \frac{N}{2}) = X_1(k) - W_N^k X_2(k) \end{cases} \quad (6.3.9)$$

$$k=0, 1, \dots, N/2-1$$

由上式可知 N 点 DFT 分解为两个 $N/2$ 点的 DFT, 上式即一个蝶形运算单元的表达式。因为 $N=2^M$, $N/2$ 仍然是偶数, 故可以对 $N/2$ 点 DFT 再作进一步的分解。以此类推, 经过 $M-1$ 次分解, 最后将 N 点 DFT 分解成 $N/2$ 个 2 点 DFT。下图为 N 点 DIT-FFT 运算流程图

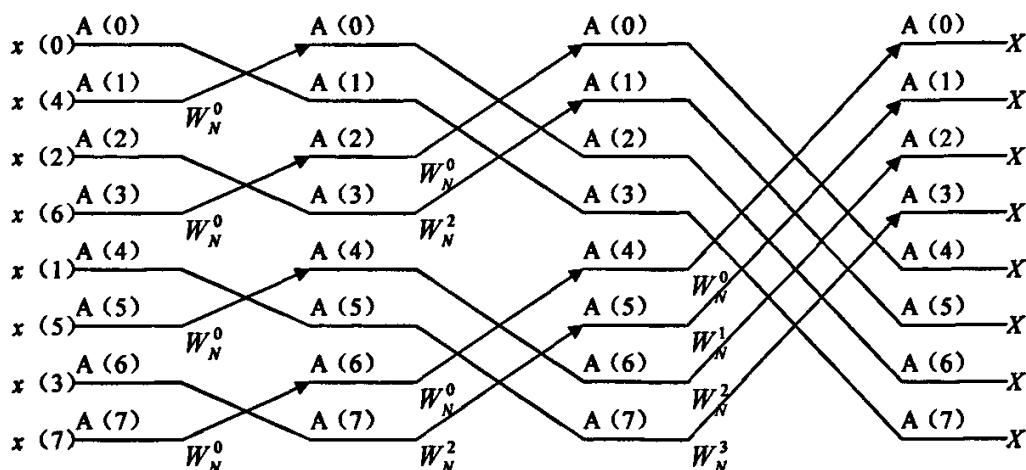


图 6-4 N 点 DIT-FFT 运算流程图 ($N=2^3$)

DIT-FFT 在汇编中实现还需要注意的运算规则如下:

- 输入序列倒序规律, 图 6-4 中输入序列不是顺序排列, 其排列是有规律的。0 (000), 4 (100), 2 (010), 6 (110), 1 (001), 5 (101), 3 (011), 7 (111) 满足倒序进位规律, 即二进制数从最高位加 1, 逢 2 要向右边的次高位进位。在 DSP 中, 用硬件电路和汇编指令产生倒序数很方便, 从而提高了 FFT 在 DSPs 中的运算效率。
- 原位计算, 从图 6-4 可以看出, DIT-FFT 运算很有规律。 $N=2^M$ 点的 FFT 共进行 M 级运算, 每级由 $N/2$ 个蝶形运算组成。同一级中每个蝶形的两个输入数据只对计算本蝶形有用, 而且每个蝶形的输入、输出数据结点又在同一条水平线上, 这就意味着计算完一个蝶形后, 所得输出数据可立即存入原输入数据所占用的存储单元。这样经过 M 级运算后, 原来存放 N 个存储单元中便存放 $X(k)$ 的 N 值。这种利用同一存储单元存储蝶形计算输入输出数据的方法称为原位计算。原位运算可以节省大量的存储空间, 从而使得设备成本降低。
- 旋转因子的生成, 在每一个蝶形运算中, 都乘以因子 W_N^p , 称其为旋转因子, p 为旋转因子的指数。旋转因子是随着级数 L 的变化而变化。第 L 级的旋转因子为

$$W_N^p = W_{2^L}^J = W_{N \cdot 2^{L-M}}^J = W_N^{J \cdot 2^{M-L}} \quad J=0,1,2,\dots,2^{L-1}-1 \quad (6.3.10)$$

在 FFT 运算中, 选择和因子 $W_N^m = \cos(2\pi m/N) - j \sin(2\pi m/N)$, 求正弦和余弦函数值的计算量是很大的。所以编程时, 产生旋转因子的方法直接影响运算速度。我们采用查表法来实现, 在 FFT 程序开始前预先计算出 $\cos(2\pi m/N)$ 和 $\sin(2\pi m/N)$, $m=0,1,\dots,N/2-1$, 存放在数组中, 作为正余弦表, 在程序执行过程中直接查表得到正余弦值直接参与计算。通过查找表的方法, 发现由于直接访问内存单元, 不用大量的运算, 速度有了极大的提高, 基本满足了计算速度的要求。

6.3.2 实序列 FFT

系统中音频芯片实际的输入信号为实数，对于实序列的傅里叶变化 RFFT 可以进一步采取节省运算量的措施。

对于序列 $x(n)$, $n=0,1,2,\dots, 2N-1$, 它的离散 FFT 变换的定义为:

$$X(k) = \sum_{n=0}^{2N-1} W_{2N}^{kn} x(n) \quad k=0,1,\dots,2N-1 \quad (6.3.11)$$

RFFT 的基本思想是将 $x(n)$ 看作长度为 N 的复序列进行 CFFT 变换。

当数据 $x(n)$ 为实序列时，如果直接进行 FFT 运算，则应把 $x(n)$ 看成虚部为零的复序列进行计算，增加了存储量和运算时间，我们采用 $N/2$ 点 FFT 计算一个 N 点实序列的 DFT 的处理方法。

设 $x(n)$ 为 N 点实序列，取 $x(n)$ 的偶数点和奇数点分别作为新构造 $y(n)$ 的实部和虚部， $y(n)$ 为 $N/2$ 点的虚数序列即

$$y(n) = x_1(n) + jx_2(n) \quad (6.3.12)$$

$$\text{在这里 } x_1(n) = x(2n) \quad (6.3.13)$$

$$x_2(n) = x(2n+1) \quad (6.3.14)$$

$$n=0,1,\dots,N/2-1$$

对 $y(n)$ 进行 $N/2$ 点 FFT，输出结果 $Y(k)$ 。

$$\text{则有 } X_1(k) = DFT[x_1(n)] = Y_{ep}(k) = \frac{1}{2}[Y(k) + Y^*(N-k)] \quad (6.3.15)$$

$$X_2(k) = DFT[x_2(n)] = Y_{op}(k) = \frac{1}{2j}[Y(k) - Y^*(N-k)] \quad (6.3.16)$$

$$X(k) = X_1(k) + W_N^k X_2(k) \quad (6.3.17)$$

$$k=0,1,\dots,N/2$$

由 $x(n)$ 为实序列，所以 $X(k)$ 具有共轭对称性，所以 $X(k)$ 另外 $N/2$ 的值为

$$X(N-k) = X^*(k) = X_1(k) - W_N^k X_2(k) \quad (6.3.18)$$

$$k=1,\dots,N/2-1$$

相对于一般的 FFT 算法，如果序列长 $N=2^M$ 为上述算法效率为

$$\eta = \frac{2M}{M+1} \quad (6.3.19)$$

6.3.3 逆 FFT 的快速实现

信号在频域进行变换后最终还要重新通过快速离散傅里叶逆变换 IDFT 恢复到时域, IDFT 算法也应用到 DFT 算法的基本结构, 我们使得 IDFT 与 DFT 公用一个 FFT 子程序, 进一步提高编码的效率。

比较 IDFT 和 DFT 的运算公式为

$$X(k) = DFT[x(n)] = \sum_{n=0}^{N-1} x(n)W_N^{kn} \quad (6.3.20)$$

$$x(n) = IDFT[X(k)] = \frac{1}{N} \sum_{k=0}^{N-1} X(k)W_N^{-kn} \quad (6.3.21)$$

从上式我们看出, 只要将 DFT 运算式中的系数 W_N^{kn} 改变为 W_N^{-kn} , 最后乘以 $1/N$, 就是 IDFT 的运算公式。由于

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X(k)W_N^{-kn} \quad (6.3.22)$$

所以

$$x^*(n) = \frac{1}{N} \sum_{k=0}^{N-1} X^*(k)W_N^{-kn} \quad (6.3.23)$$

对上式两边同时取共轭, 得

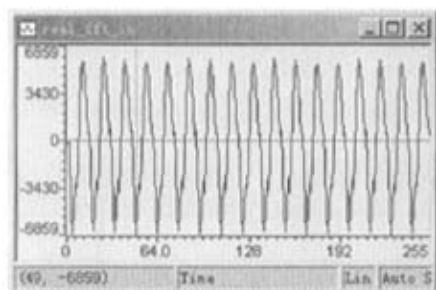
$$x^*(n) = \frac{1}{N} \left[\sum_{k=0}^{N-1} X^*(k)W_N^{-kn} \right]^* = \frac{1}{N} \{DFT[X^*(k)]\}^* \quad (6.3.24)$$

这样, 可以先将 $X(k)$ 取共轭, 然后直接调用 FFT 子程序, 最后的结果取共轭并乘以 $1/N$ 得到序列 $x(n)$ 。因为我们所处理的是实序列, 实际 IFFT 的结果序列其虚部都为零, 程序中, 无需对结果进行共轭处理, 只需提取其中的实部就是所需要的输出信号序列。这种方法能够方便实现与 FFT 共用同一子程序。

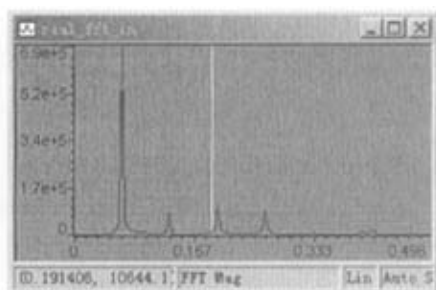
6.3.4 RFFT 和 IFFT 算法的验证

为较好的反映 RFFT 和 IFFT 算法, 本文采用输入纯音信号观测信号的时域频域波形。

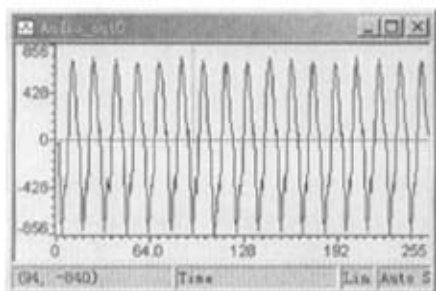
我们取采样率为 16k，输入长为 256 点频率为 1k 的标准纯音信号图 6-5a，信号首先通过 RFFT 变换称为频域信号，然后再通过 IFFT 返回到时域。输入的语音信号频谱应该是在 1k 点上表现为突出的峰值，而其它点上的值为均零。由于输入的纯音信号夹带一些干扰，所以在 CCS 给出的频谱图上有四个突出，但是在 1k 频点上有最大峰值，如图 6-5b 所示，输入信号经过 RFFT 变换到频域后的结果图 6-5c 与图 6-5b 几乎相同，在 1k 频点上显示较大的峰值。频域信号经过 IFFT 后得到的信号波形如图 6-5d，能够较好输入信号恢复。结果证明 DSP 上实现的 RFFT 和 IFFT 变换是有效的。



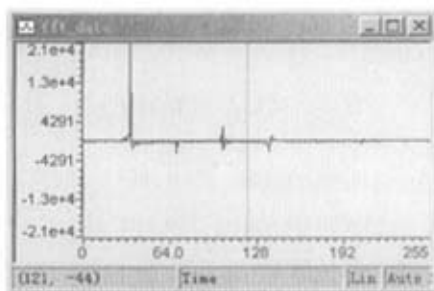
a 输入纯音信号



b 输入纯音信号频谱 (CCS 频谱图)



d IFFT 变换的结果



c 输入信号经过 FFT 变换的结果

图 6-5 RFFT 和 IFFT 变换结果

6.4 压缩算法的验证

长为 256 点的实序列经过 RFFT 变换后，得到同样长度的离散频谱序

列,系统离散信号的采样率定为 16kHz,离散频谱序列频率分辨率为 125Hz。压缩算法就是对该序列进行非线性变换实现合理的频谱补偿。我们在 MATLAB 和 DSP 平台上主要实现宽动态压缩和移频压缩。

6.4.1 宽动态压缩

如第五章所述宽动态压缩输入输出函数的变量和自变量都是声压级,输入信号的大小与声压满足线性关系,而与声压级满足对数关系。我们需要将频谱序列值转换到声压级,因此在运算过程中需要涉及 C 语言的对数与指数运算。

我们对于频谱的补偿,首先要求出特征点的增益值,然后通过插值算法求得频域所有点的增益,求得增益后,频域上各点乘以增益即可得到补偿后的频谱序列,我们选取的特征点为 250、500、1000、2000、4000、8000Hz。程序运行之前我们将这些点上的听阈、痛阈和最舒适声压级值组成数组存放在 ROM 中,程序运行时根据这些数值即可得到特征点上的输入输出特性。

设频率特征点 ω_n 输入声压值为 P ,则输入声压级 SPL_{in} 为

$$SPL_{in} = 20 \lg \left(\frac{P}{P_0} \right) \quad (6.4.1)$$

频率特征点增益值 $G(\omega_n)$ 为

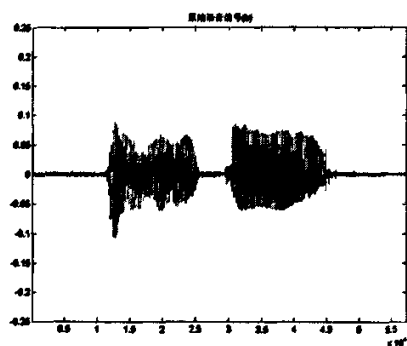
$$G(\omega_n) = 10^{\frac{SPL_{out} - SPL_{in}}{20}} \quad (6.4.2)$$

因为 CCS 编译器支持 C 语言中的浮点的对数和指数运算,我们能够较方便的求得声压增益。

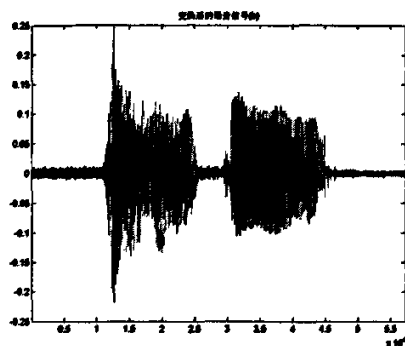
图 6—5 为 MATLAB 模拟 WDRC 算法下所得的结果, a1 为语音女声普通话“牙膏”的波形(16bit 44.1kHz), a1 为经过算法处理后的波形, b1 为语音女声普通话“牙膏”的频谱, b2 为经过算法处理后的频谱。

从图 6—6a1 与 a2 可以看出,经过压缩后的信号,声压幅度整体增加,输出声音的强度增大。图 b2 相对于 b1 的频谱在原有的基础上均有增加,且高频部分的增加的相对幅度高于低频部分。听障患者的高频损失大于低

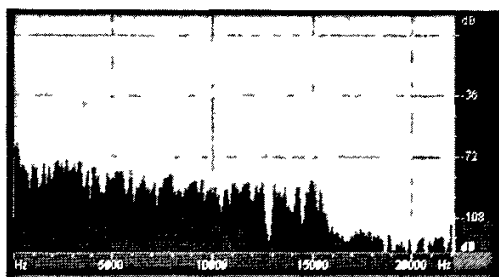
频，动态压缩中频谱在高频部分的补偿也应大于低频部分。上图中频谱上各点的增益组成的增益曲线如图 6-7，纵轴表示增益值，横轴为频率点间隔为一个倍频程频率。增益曲线是将个特征点的频率用线段相连，使之保持完整连续。



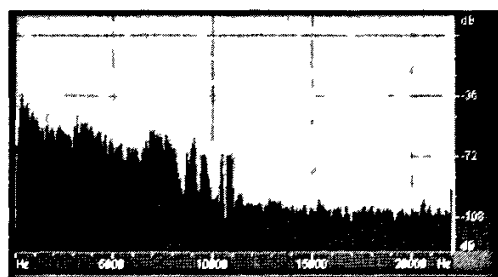
a1 女声普通话“牙膏”的时域



a2 经过算法处理后信号时域



b1 女声普通话“牙膏”的频域



b2 经过算法处理后信号频域

图 6-6 MATLAB 实现 WDRC 算法结果

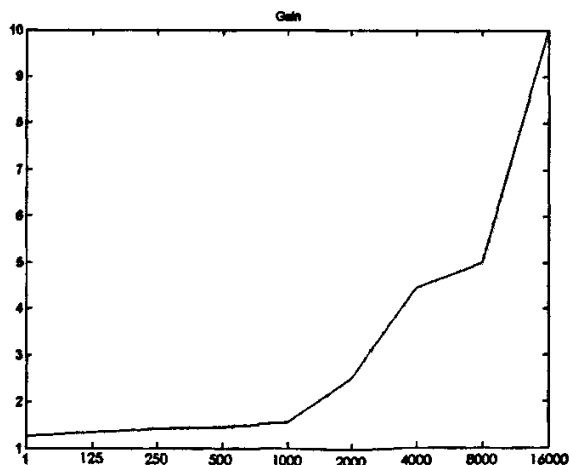
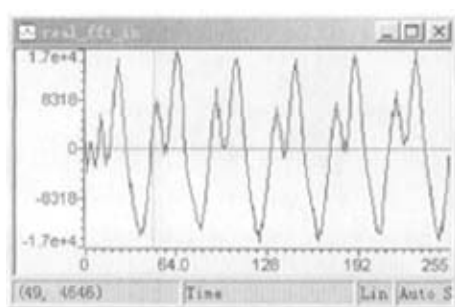


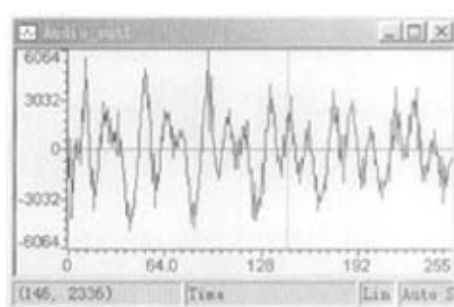
图 6-7 增益曲线图

本文最终在 DSP 开发平台上实现宽动态压缩算法,借助 CCS 调试工具提供的强大灵活的 DSP 存储器的观测窗口,我们得到信号波形图,能够直观看到缓冲区内信号的时域和频域波形的变化。

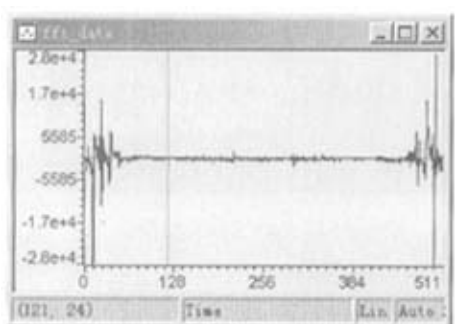
采集长度为 256 点的一帧数字语音经过 DSP TMS320VC5416 的信号处理,其处理前后结果对比所得如图 6—8。



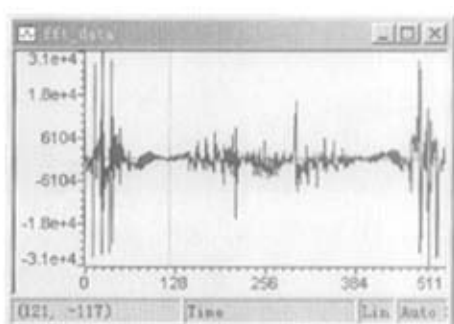
a1 输入信号时域



a2 输出信号时域



b1 输入信号频域



b2 输出信号频域

图 6—8 宽动态压缩在 DSP 中实现的结果

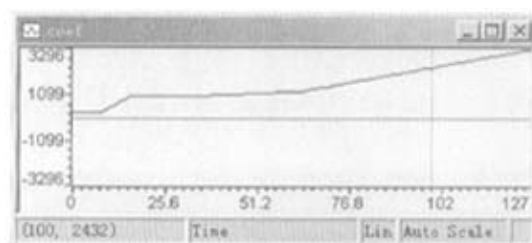


图 6—9DSP 平台宽动态压缩频域增益曲线图

上图可看出,输出信号较输入信号平滑,输出信号的高频能量增强,

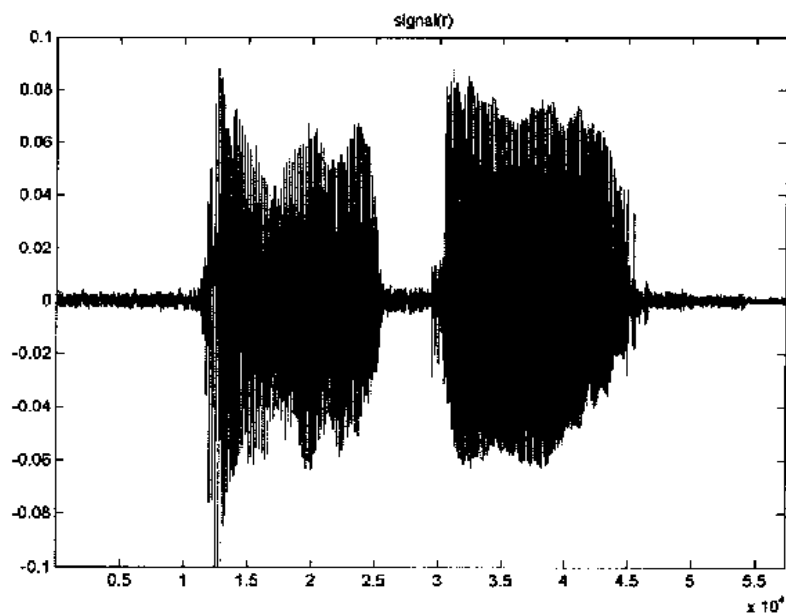
输出语音信号通过耳机听起来显得明显尖锐，音调增高，但正常人耳仍然可以分辨音节的变化（图 6-7a1, a2）。从 FFT 变换的结果对比来看（图 6-8b1, b2），低频高频的幅度均有增强，但是高频的增幅更大，整个频域上的增加的增益值是由增益曲线来获得的，如图 6-9 所示，增益曲线在整个频域上保持连续，且增益值在不同的频点上存在着差异。

6.4.2 移频压缩

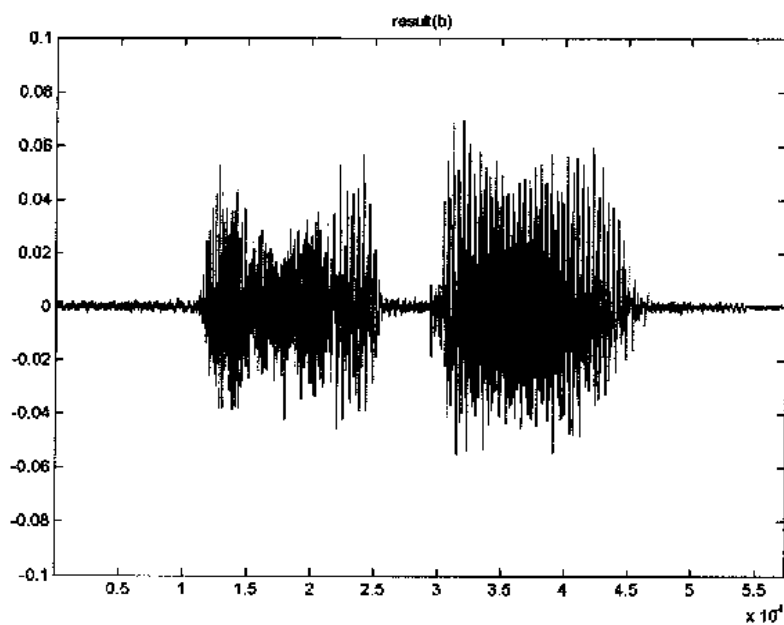
在 MATLAB 平台上，我们进行如下计算：

对于一段采样率为 44.1kHz 的语音信号，我们按帧加窗进行预处理，帧长为 256 点，再进行 256 点 FFT 运算。需要压缩的原始频段设为 0~22kHz，目标频段为 0~11kHz，压缩比为 $p=2$ 。

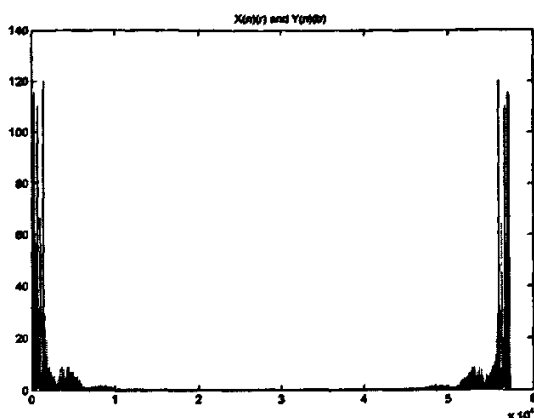
图 6-10 为变换的结果，图 6-10a1 为输入原始信号女声普通话“牙膏”的波形，图 6-7a2 为压缩后的波形，图 6-10b 为两个信号 FFT 变换结果图，其中红色为输入信号，蓝色为压缩后的信号。图 6-10a2 相对于图 6-10a1 波形变疏，高频成分减少，低频成分显现，但基本保持了时域波形的包络形状。两段信号的 FFT 变换图中显示，频谱的高频特性按比例压缩到低频，高频部分的能量降低直至为零。在助听器中，关键语频段的声音就能经变频及压缩后，基频发生改变，而保留其包络曲线基本不变，也就是保持时值，相位及幅度之间的对应关系。



a1 输入语音信号



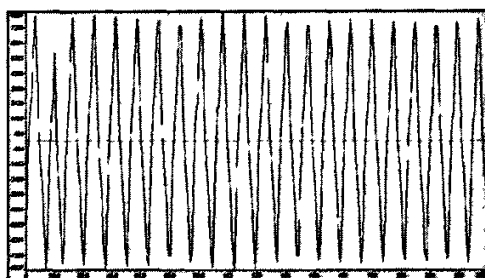
a2 压缩后的语音信号



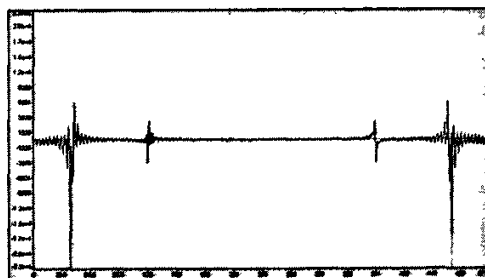
b FFT 变换结果

图 6-10 MATLAB 实现移频压缩算法结果

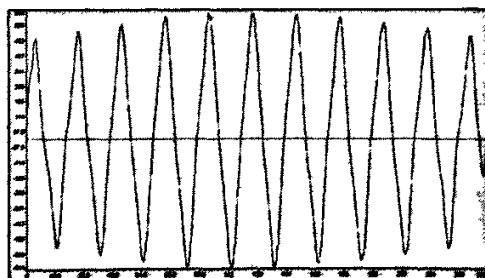
在 DSP 平台上, 我们进行压缩比为 2.0 的移频压缩计算, 将输入信号频段 0~8kHz 按比例压缩到频段 0~4kHz。如图 6-11a1 中我们输入 1.3kHz 纯音信号, 经过移频压缩后, 位于 1.3kHz 频点上的峰值, 按照压缩比 2.0 被移到了 0.65kHz 频点上。经过 IFFT 变换后, 信号频率改为 0.65kHz。输出端耳机听觉主观感觉是音调降低。



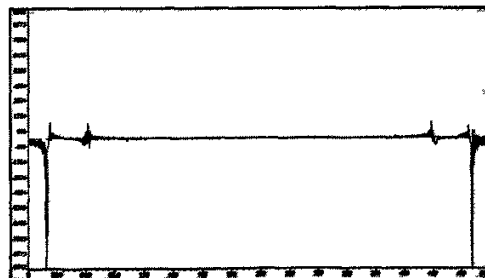
a1 输入 1.3kHz 纯音时域信号



a2 输入信号频谱



b1 经频域压缩后的时域信号



b2 压缩后的信号频谱

图 6-11 移频压缩结果

6.5 自适应算法的验证

MATLAB 模拟 DSP 自适应算法结果如图 6—12 所示,我们在两路语音“牙膏”的波形上叠加均值为 0.01 的白噪声信号得到图 6—12a1 和 a2,其中 a1 作为输入信号, a2 作为参考信号, 经过自适应算法后得到信号图 6—12b2 所示, 图 6—12b1 为未加噪声的信号波形。

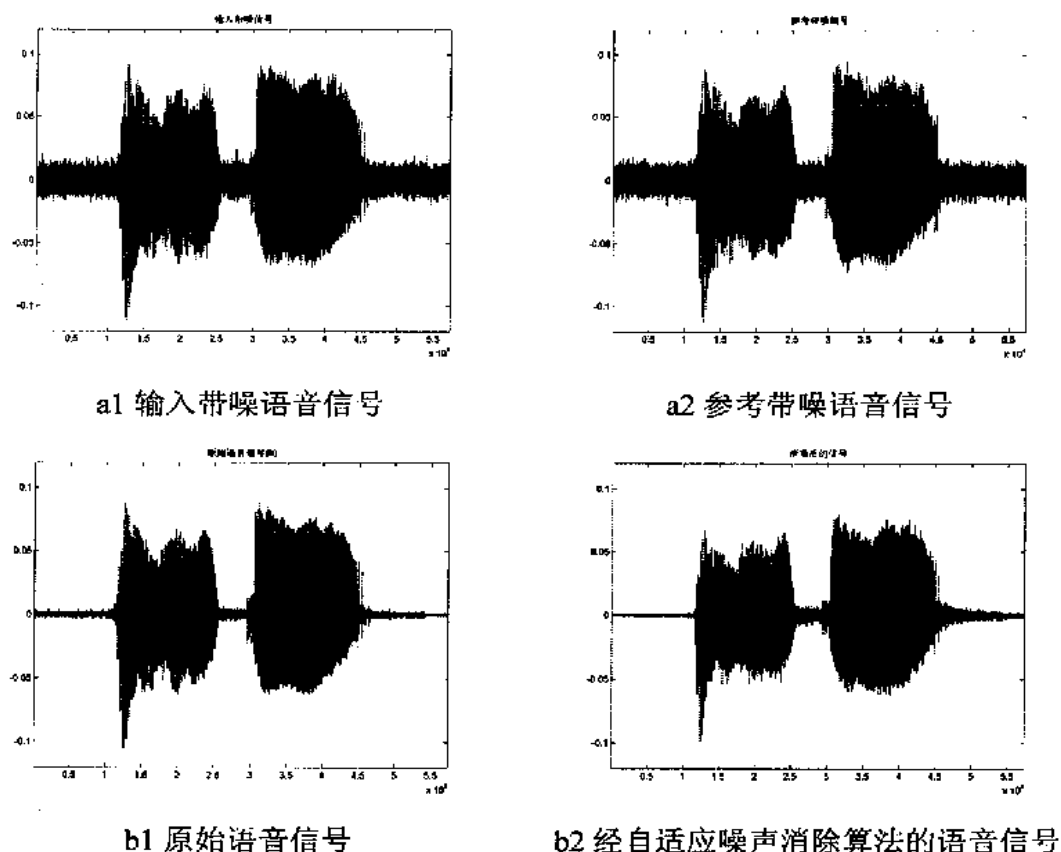


图 6—12 MATLAB 实现自适应消噪算法结果

经过自适应消噪, 信号中的噪声较好得到抑制, 特别是在没有语音输入信号的时段信噪比最小, 但是噪声的消除最为明显。图 6—12b 中, 当没有输入语音信号时, 能够看到噪声逐渐收敛最后始终保持在一个较小的幅度内。应用自适应消噪算法的助听器没有足够幅度的声音输入时候, 输出几乎是静音。

下图 6—13 为 DSP 平台上实现的自适应噪声消除结果, 语音信号的长度为 256 点, 采样率 16k。图 6—13a 为输入语音信号, 6—13b 为噪声消除

后的信号,从波形图上可看出后者波形变得平滑,锯齿状波收到抑制,输出语音听觉表现为噪音降低,语音清晰。

本文还对自适应算法对系统自身噪声消除的效果进行实验,图 6-14a 为 DSP 开发系统在没有音频输入情况下 A/D 端读入的白噪声信号,该信号随机变化,属于宽带信号。经过自适应消除后的波形如图 6-14b,噪声受到明显的抑制,输出的信号幅度限制在极小范围,输出端处于静音状态。

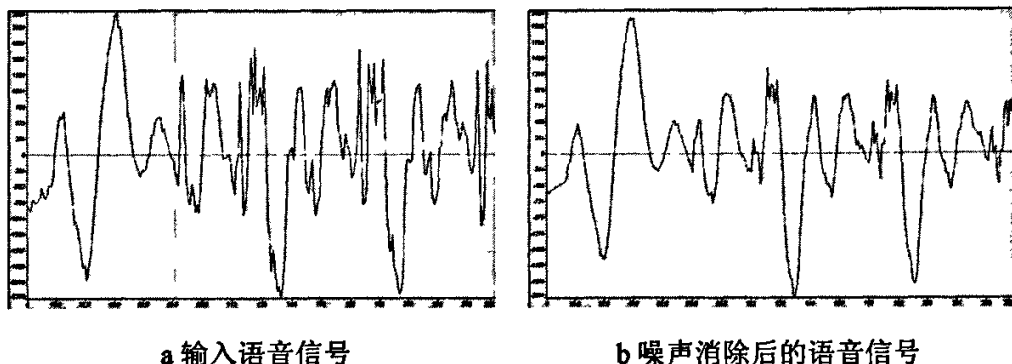


图 6-13 语音信号经噪声消除的结果对比

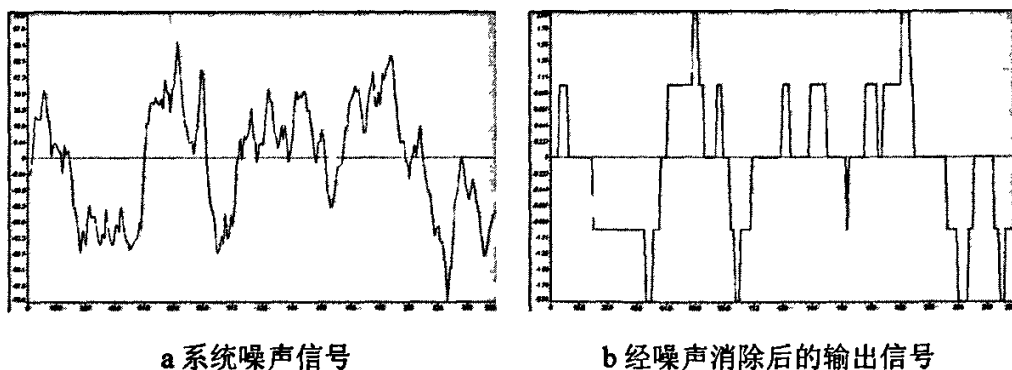


图 6-14 DSP 系统噪声经自适应噪声消除的结果对比

6.6 开发中的关键问题

下面就讨论助听器算法在此实验平台上实现上要注意的几个关键问题。所讨论的这些问题,在 DSP 系统设计中实际的参考和应用价值。

6.6.1 数的定标

DSP 分为定点和浮点两种, 在定点 DSP 芯片中, 采用定点数进行数值运算, 其操作数一般采用整型数来表示。一个整型数的最大表示范围取决于 DSP 芯片所给定的字长, 一般为 16 位或 24 位。显然, 字长越长, 所能表示的数的范围越大, 精度也越高。在算法的实现过程中, DSP 所要处理的数可能是整数, 也可能是小数或混合小数; 然而, DSP 在执行算术运算指令时, 并不知道当前所处理的数据是整数还是小数, 更不能指出小数点的位置在哪里。因此, 在编程时必须指定一个数的小数点处于哪一位, 这就是定标。通过定标, 可以在 16 位数的不同位置上确定小数点, 从而表示出一个范围大小不同且精度也不同的小数。例如: 在 Q15 中, $1080H=0.12890625$; 在 Q0 中, $1080H=4224$ 。

程序的编写中, 首先在处理前, 按要求确定 Q 值, 移位求出相应数值, 运算, 将结果反向移位得出结果。不同的 Q 值所表示的数不仅范围不同, 而且精度也不同。浮点数与定点数的转换关系可以表示为:

$$\text{浮点数}(x) \text{转换为定点数}(x_q): x_q = (\text{int})x * 2^Q \quad (6.6.1)$$

$$\text{定点数}(x_q) \text{转换为浮点数}(x): x = (\text{float})x_q * 2^{-Q} \quad (6.6.2)$$

下面结合标准的 C 语言的格式来讨论基本浮点算术运算的定点模拟方法:

1) 加法/减法运算的 C 语言定点模拟

设浮点加法运算的表达式为:

```
float x,y,z;
```

```
z=x+y;
```

将浮点加法/减法转化为定点加法/减法时, 最重要的一点就是必须保证两个操作数的定标值一样。若两者不一样, 则在做加法/减法运算前先进行小数点的调整。为保证运算精度, 需使 Q 值小的数调整为与另一个数的 Q 值一样大。此外, 在做加法/减法运算时, 必须注意结果可能会超过 16 位表示。如果加法/减法的结果超出 16 位的表示范围, 则必须保留 32 位结果, 以保证运算的精度。而后这一点最为重要, 一般程序中出错时, 也先想到

这个原因。所以，当笔者无法确定时，就用 long 型的数来装运算结果。

为了分清不同情况下的处理方法，下面分两种情况来分析：

● 结果不超过 16 位表示范围

设 x 的 Q 值为 Q_x ， y 的 Q 值为 Q_y ，且 $Q_x > Q_y$ ，加法结果 z 的定标值为 Q_z ，定点加法可以描述为：

```
int x,y,z;
long temp;
temp=y<<(Qx-Qy);
temp=x+temp;
z=(int)(temp>>(Qx-Qz)), /*若  $Q_x \geq Q_z$ , temp 结果右移  $Q_x - Q_z$  位*/
z=(int)(temp<<(Qz-Qx)), /*若  $Q_x \leq Q_z$ , temp 结果左移  $Q_x - Q_z$  位*/
```

● 结果超过 16 位表示范围

为了保留正确的运算结果，遇到这种情况时，输入值要保留全部的 32 位结果。所以运算时的大致结果变为如下：

设 x 的 Q 值为 Q_x ， y 的 Q 值为 Q_y ，且 $Q_x > Q_y$ ，加法结果 z 的定标值为 Q_z ，则定点加法为：

```
int x,y;
long temp,z;
temp=y<<(Qx-Qy);
temp=x+temp;
z=temp>>(Qx-Qz); /*若  $Q_x \geq Q_z$  */
z=temp<<(Qz-Qx); /*若  $Q_x \leq Q_z$  */
```

2) 乘法运算的 C 语言定点模拟

设浮点乘法运算的表达式为：

```
float x,y,z;
z=x*y;
```

定点表示的乘法为：

```
int x,y,z;
long temp;
```



```
temp=(long)x;
z=(temp*y)>>(Qx+Qy-Qz);
```

3) 除法运算的 C 语言定点模拟

设浮点除法运算的表达式为：

```
float x,y,z;
z=x/y;
```

假设经过统计后被除数 x 的定标值为 Q_x ，除数 y 的定标值为 Q_y ，商 z 的定标值为 Q_z ，则定点表示的除法为：

```
int x,y,z;
long temp;
temp=(long)x;
z=(temp<<(Qz-Qx+Qy))/y;
```

在使用定点 DSP 时，如何选择合适的 Q 值是一个关键性问题。就 DSP 运算的处理过程来说，实际参与运算的都是变量，有的是未知的，有的则在运算过程中不断改变数值，但它们在一个实际工程环境中作为一个物理参量而言都有一定的动态范围。只要这个动态范围确定了， Q 值也就确定了。可以看出， Q 值越大，数值的表示范围越小，但是精度越高，例如： Q_{15} 的数值范围是 -1 到 $+0.9999695$ ，但是精度却有 $1/32768$ 即 0.00003051 这么高；相反， Q 值越小，数值范围越大，但精度越低。如： Q_0 的数值范围是 -32768 到 $+32768$ ，精度为整数 1 。因此，对定点数而言，数值范围与精度是一对矛盾，一个变量要想能够表示比较大的数值范围，必须以牺牲精度为代价；而要想提高精度，则数的表示范围则就相应的减少。因此，在程序设计前，首先要通过细致和严谨的分析，找出参与运算的所有变量的变化范围，充分估计运算中可能出现的各种情况，然后确定采用何种定标标准才能保证运算结果正确可靠。般可以通过理论分析和统计分析方法确定变量绝对值得最大值，然后再根据它来确定 Q 值。但是，DSP 操作过程中的意外情况是无法避免的，即使采用统计分析法也不可能涉及到所有情况。因此，在定点运算过程中应该采取一些判断和保护措施（特别是在定点加法中）。另外，在数字信号处理中的大量运算是乘法和累加，应尽量

采用纯整数或纯小数运算，即全部变量都用 Q0 或 Q15 格式表示这样做的好处是操作简单、编程方便。只有当纯整数或纯小数运算不能满足变量的动态范围和精度要求时，才采用混合小数表示法进行定点运算。

6.6.2 溢出的问题

在编程的过程中，笔者还常常遇到溢出的情况。一般这种情况常发生于对数据进行运算处理的过程中的某一步或某几步。所以，当结果出现异常情况时，排除了算法出错的情况后，笔者第一个要想到的就是是否发生了溢出。之后就要用不同的方法来避免溢出的发生。

一般来说，现在出产的 DSP 芯片，都安排了溢出保护措施。例如，这款 C5416 芯片，就在状态寄存器 1 (ST1) 内，安排了 OXM 溢出模式位。其中，OXM=0 时：溢出结果不做变化，送回目的累加器；OXM=1 时：溢出结果为正时（上溢），目的累加器置最大正值，当结果为负时（下溢），目的累加器置最小的负值，也就是说取最大下值或最小负值。通常情况下，这款芯片的溢出方式是方式 0。

上而这种方法，是硬件设置的方法，有一定的局限性，因为一般发生溢出的时候，计算结果往往超出了变量所能承受的范围很多。这时即使取最大正值或最小负值，尽可能接近了真实值，但结果已经没有什么意义了。所以，在找到了溢出的来源后，应当采取软件的办法来修正它。一般的算法中，都是变量两两相互运算，即使是比较长的公式，也可以分成小的式子。这样，就可以理解为几个输入量的共同作用下，得到了一个输出值。所以，解决溢出的方法之一是减少输入值的大小。这里的输入值有真正的输入量的概念，如由麦克风采集进来的语音信号，也有在算法中加入的系数、参数等输入量的概念。为了尽量保持经过 A/D 转换后信号的保真性，通常小会把输入量减小，这样一来，就只好减小参数的大小了。幸运的是，一般参数是用笔者上而讨论过的定标，也就是 Q 值来表示的，这样，笔者可以把溢出也作为一个方面，考虑进 Q 值大小的决定因素中，通过 Q 值的调整，可能表示的范围下降了，但是避免了溢出。可是，问题又来了，毕竟 Q 值的调整是有限的，所以，笔者就又把目光集中到了输出值上。这里

的输出值是指运算的中间值。笔者可以让两个基本型 `int` 型的输入值相乘，把结果存放在一个长整形内，在全部运算完毕之后，再将其右移成 `int` 型（舍去低位）。这样就可以避免溢出了。这就是避免溢出的第三种方法。

实际应用中，还是第三种方法用的比较多，因为它比较灵活。有时，为了尽量避免出现意料之外的溢出，可以遵循 `int*int->long`（即 `int` 型的变量乘 `int` 型的变量，结果赋给 `long` 型的变量）的规则来进行运算，在下次运算开始之前，判断一下 `long` 型的数如果赋给 `int` 型是否会发生溢出，如果不会就赋给另一个 `int` 型的变量，开始下一次的运算。否则，也要保证与 `long` 型数相乘的另一个数一定要是 `int` 型。

6.6.3 量化与舍入误差

在数字信号处理中，需要把输入的离散时间信号和系统系数进行量化，量化的结果与原来的数值之间必然存在误差，误差大小要由计算机或专用硬件寄存器的字长而定。如果所选用字长太长，则对硬件要求太高，因而价格昂贵，且计算时间也长。另一种误差是由乘法引起的。如果两个 `b` 位的二进制数相乘，其结果是 `2b` 位的二进制数，在有限字长的限制下，我们需要对乘积进行截尾或舍入处理，这样也就产生了误差。

一般通用计算机的字长较长，量化步长很小，所以量化误差不大，常常可以忽略。因此，在用通用 DSP 来实现数字系统时，则因专门硬件设备的字长较短而必须考虑误差问题。在数字滤波器中，共有三种因量化而引起误差的因素：

- A/D 变换的量化效应；
- 系数的量化效应；
- 数字运算过程中的有限字长效应。

在定点制中，乘积尾数处理所产生的误差累计起来会使运算结果精度下降。在前而讨论算法的结构和具体实现时，并未考虑精度问题，在实际应用中，DSP 的有限字长效应在实现这些算法时就会使精度受到限制，因此编程者在程序设计时应考虑到有限字长所带来的各种不利因素以及对精度的影响。这些影响就包括输入信号的量化误差和算法中的舍入误差，它

们要么造成与设计原则相偏离，要么造成在滤波器输出中引入额外噪声，因此我们对运算中的量化噪声与舍入误差进行分析。

TMS320C5416 是一个 16/32 位的定点 DSP，处理的输入数据归一化成一个小数并用 15bit 来表示，1 个 bit 表示正负号。量化误差为

$$\delta = 2^{-b} \quad (b=15) \quad (6.6.3)$$

这就是系统的量化精度，每个输入以 δ 为单位量化成相应数值。

DSP 中每一次数据进行四则混合运算后都会面临截断，截断引入了舍入误差，在 C5416 中，乘法器和 ALU 是 40/32bit，计算结果值的尾数由 31bit 表示，外加 1bit 符号位，8bit 指数位，可以直接完成乘积精度累加。因此每次乘法后，数据只作一次四舍五入操作，极大的减少了舍入误差。而乘法器和 ALU 的量化精度 δ 等于 12^{-31} 。可以看出 δ 相对来讲是很小的，系统中的舍入误差为一个白噪声，分布在一个量化步长 δ 内，即舍入误差也很小。

6.6.4 C 语言和汇编联合编译

本文采用 C 与汇编指令联合编译的方式来完成软件，其中 RFFT 和 IFFT 部分采用汇编语言实现，目的是利用汇编特有的指令，提高软件效率，减少代码空间，系统的主程序和设计浮点数运算部分则采用 C 语言，目的是保证程序的可移植性，支持硬件系统的升级，同时提高可读性方便调试。

C 与汇编共同使用能简化开发的效率，但它们之间由于要交换一些数据，因此要遵守共同的接口规范，TI 公司对这些规范做了明确的定义，供 C 使用的寄存器中也要有严格的规定，下面结合调试过程中总结的经验来介绍一下。

TMS320C5402 定点处理器有两种类型的存储器：程序存储器和数据存储器。在程序存储器中主要包含可执行的程序代码；在数据存储器中，则主要包含外部变量、静态变量和系统堆栈。由 C 程序生成的每一块程序或数据存放于存储空间的一个连续块中。由 C54x 编译器编译生成六个可以进行重定位的代码和数据块，这些块可以用不同方式分配至存储器以符合不同系统配置的需要。这六个块可分为两种类型，一种是已初始化块，另一种是未初始化块。

已初始化块主要包括数据表和可执行代码。C 编译器共创建三个已初始化块：

.text 块包含可执行代码和字符串；

.cinit 块包含初始化变量和常数表；

.const 块字符串和 switch 表。在大存储器模式下，常数表也包含在.const 块中。

未初始化块用于保留存储器空间，程序利用这些空间在运行时创建和存储变量。C 编译器共创建三个未初始化块：

.bss 块保留全局和静态变量空间。在小模式中，.bss 块也为常数表保留空间。在程序开始运行时，C 初始化 boot 程序将数据从.cinit 块中拷贝至.bss 块。

.stack 块为系统堆栈分配存储器。这个存储器用于将变量传递至函数以及分配局部变量。

.system 块为动态存储器函数 malloc, calloc, realloc 分配存储空间。当然若程序没有用到这些函数，编译器就不创建.system 块。

一般的，.text, .cinit 和.const 连同汇编语言中的.data 块可链入到系统的 ROM 或 RAM，而.bss, .stack 和.system 块则应链入 RAM 中。

在 C 环境中，定点 C 编译器定义了严格的寄存器使用规则。这些规则对于编写汇编语言与 C 语言的接口非常重要。如果编写的汇编程序不符合寄存器使用规则，则 C 环境将会被破坏。要注意，C 编译器使用寄存器的方法在使用和不使用优化器时是不一样的。因为优化器需要使用额外的寄存器作为寄存器变量以提高程序的运行效率，但函数调用时保护寄存器的规则在使用和不使用优化器时是一样的。

汇编用于函数中，可以使用辅助寄存器、各种状态寄存器，还包括 T 寄存器和 P 寄存器。在使用时必须符合下列规则：

1) 辅助寄存器 (ARP 和 AR0-7)

对于 TMS320C54x，函数进入和返回时，ARP 必须为 0，也就是说当前的辅助寄存器为 AR0 在函数执行过程中，AR0-7 都可以自由使用，即可修改而不必恢复。当 AR6 和 AR7 被用作寄存器变量时，如果在函数中被修

改，则必须加以保护和恢复。对于 TMS320C54x，定点 C 环境中还保留了三个寄存器：AR0，AR2，AR2，它们的作用如下：

*AR0 帧指针

*AR1 堆栈指针

*AR2 局部变量指针，用于计算局部变量的地址

2) 状态寄存器

在 C 编译器中，始终假定 PM 状态位为 0。如果函数改变 PM 值，则在函数返回时必须将 PM 重新设为 0。在 TMS320C54x 中，TPM 为必须保持为 0，硬件复位时 TPM 为 0。对其它状态位既可以修改，也不必恢复。

3) 其它寄存器

累加器 ACC 可以自由使用，不必保护和恢复。ACC 可以用来返回整数、指针和浮点值。对于 TMS320C54x，P 和 T 寄存器也可以自由使用。

在一个函数中，定点 C 编译器可以自由使用多至两个寄存器变量。如果要在函数中使用寄存器变量，则应在函数的参数表或函数的第一块中定义，否则，作为一般变量处理。对于 TMS320C54x，编译器用 AR1 和 AR6 作为寄存器变量，其中 AR1 分配给第一个寄存器变量，AR6 分配给第一个寄存器变量。由于在运行时建立一个寄存器变量约需 4 个指令周期，因此，只有当一个变量访问 2 次以上，使用寄存器变量的效果才能明显的体现出来。

在程序中，是由汇编程序调用由 C 编写的函数，在编写独立的汇编程序时，要注意以下几点：

- 1) 不论是用 C 编写的函数还是用汇编编写的函数，都必须遵循寄存器使用规则。
- 2) 必须保护函数要用到的几个特定寄存器。这些特定的寄存器包括：AR0 (FP)、AR1 (SP)、AR6 和 AR7 其中，如果 SP 使用正常的话，则不必明确加以保护，换句话说，只要汇编函数在函数返回时弹出压入的现象，实际上就已经保护了 SP，因而 AR1 就可以自由的使用。其他寄存器则可以自由的使用。函数返回时 ARP 必须为 1。
- 3) 中断程序必须保护所有用到的寄存器。

- 4) 从汇编程序调用 C 函数时, 将从左向右第一个参数放在累加器 ACC 中, 其余的参数按照逆序依次压入堆栈。
- 5) 调用 C 函数时, 注意 C 函数只保护了几个特定的寄存器, 而其他寄存器 C 函数是自由使用的。
- 6) 长整型和浮点数在存储器中存放的顺序是低位字在低地址, 高位字在高地址。
- 7) 如果函数有返回值, 则返回值存放在累加器 ACC 中。
- 8) 汇编模块不能改变由 C 产生的 .init 块, 如果改变其内容则会引起不可预测的后果。
- 9) 编译器在所有的标识符 (函数名、变量名等) 前加一下划杠 “_”。因此, 编写汇编语言程序时, 必须在 C 程序可以访问的所有对象前加 “_”。
- 10) 任何在汇编中定义的对象或函数, 如果需要在 C 中访问或调用, 则必须压汇编指令 .global 定义。同样, 如果在 C 中定义的对象或函数, 需要在汇编中访问或调用, 在汇编中也必须用 .global 指令定义。

6.7 本章小结

本章介绍数字助听器算法程序的计算机模拟、DSP 开发、调试以及结果的验证过程, 应用到 MATLAB 和 TI 提供的 DSP 开发环境 CCS 软件。

MATLAB 是一款功能强大的数值计算软件, 拥有丰富的数据类型和结构、友善的面向对象、加快速精良的图形可视、广博的数学和数据分析资源、高效的应用开发工具。在国际学术界, MATLAB 已经被确认为准确、可靠的科学计算标准软件。我们的各项算法都能在 MATLAB 上通过编程实现模拟运算, 结果得到良好的验证与显示。CCS 软件同样是功能完善的开发环境, 对 C 语音和汇编进行联合编译, 它为开发者提供了灵活的内存观察窗口, 多样的调试模式, 实时的状态分析工具等, 通过使用这些调试工具, 笔者基本做到熟练掌握 DSP 系统的开发过程。

本章实现的算法主要有:

- 宽动态压缩

- 移频压缩
- 自适应噪声消除

这些算法全部在 DSP 开发板上实现了预期的语音处理功能，也得到符合理论的结果。我们在本章中给出了部分计算的结果。

需要说明的是：在说到相关程序时，限于篇幅，只是叙述关键部分和实现思想，具体的实现代码请参照有关源程序。

第七章 结束语

7.1 研究工作总结

根据导师与实验室的安排，本人在研究生下科作课题的两年期间，主要完成两项课题的任务，2004年5月下科到2005年主要从事“无线体温监测系统”的研究，而本课题的研究准备工作可以从2005年3月开始，最初主要是了解助听器原理和目前国内外算法研究的概况以及DSP芯片的原理应用，其中多次到临床科室学习。2005年5月开始论文开题的资料收集工作，收集了DSP应用资料和一些相关的设计范例。正式的研究工作从2005年6月开始，直至2006年5月份论文编写完成，整个课题展开约一年的时间。

本论文的题目是《基于DSP数字助听器关键技术的研究》，主要的研究目的是在DSP语音开发平台上实现几个数字助听器核心信号处理算法，包括宽动态压缩，移频压缩，自适应噪声消除等。论文还提出可行的助听器硬件设计方案。整个工作经过算法研究，计算机软件模拟，DSP算法实现等阶段。基本达到验证几个主要的助听器语音处理算法的目的，为DSP在听力补偿设备中的应用积累了经验，同时为下一步进行临床真人实验打下基础。

本论文的主要成果具体如下：

- 1) 我们查阅大量的文献，掌握数字助听器现在发展的状况，并进一步熟悉目前数字助听器应用的主要算法。在现有条件的基础上，完成基于DSP数字助听器的硬件和软件设计。我们研究对比多款DSP芯片和数据模拟接口芯片，选取TI DSP TMS320VC5416和TLV320AIC23B来构成助听器两个最核心的部分。即节省了成本又能够保证助听器硬件完整的功能，且硬件的体积与功耗均满足实际要求。
- 2) 对于助听器算法的设计是本论文的重点，本文研究目前数字助听器

主要的算法。由于目前可查关于数字助听器算法的文献中,较少有具体算法过程的内容。本文总结出一些算法思路,提出具体过程和算式,并通过计算机模拟来验证可行性,最后在 DSP 开发板上实现。对于算法中存在的不足,进一步修改完善。

- 3) 经过长时间摸索得出的一整套 DSP 芯片开发的方法与调试技巧。其中包括:熟悉了 TI 的 TMS320C54X 的硬件结构,CCS 环境下程序开发的方法、调试工具的使用、优化级别的选择等方面。在设计过程中,记录了开发日志,列举了遇到的关键问题及其解决方法,为今后的继续开发提供了重要参考资料。

回顾和总结毕业论文设计工作,虽然做了一些工作,但还很不够。许多可以深入研究的问题都没有得以深入下去,并且研究工作应更紧密地和实际应用结合,这样才能使得研究课题更有意义和价值。在研究工作中,也进一步体会到团队工作的重要性,一个系统的设计和完成,靠一个人的能力是不可以的,例如本研究工作,没有其它同学的帮助是无法完成的。

最后,作者深知自己所作的工作还很不够,离 DSP 数字助听器技术的实际应用还有一定的距离。

7.2 不足与展望

- 1) 目前 DSP 技术的快速发展,主要是 DSP 芯片运算速率的快速增强。TI 公司目前发布的最快速的 C64xx 系列 DSP 的工作频率可以达到 600MHz,指令密度可以达到 4800 万条/秒,再结合它的高级语言支持功能,其有效的降低了 DSP 开发的难度,扩大了应用领域。数字助听器的核心芯片影响到整个系统的发展,也保证了更复杂的助听器算法的实现。随着 DSP 技术的发展,综合考虑性能、成本、开发的难易程度前提下,核心芯片还可以进一步提高。
- 2) 助听器的算法虽已通过验证,本文的工作主要是提供一个可行性的算法框架和建立一个算法模型,实际应用中还需要考虑更多的因素,算法有待改进。数字助听器的功能还包括很多方面入方向性麦克技术,反馈消除技术等,由于受到时间和实验条件的限制,本

文中我们没有深入研究这些算法。程序代码的效率需要提高，若能够对编译后的代码进行手工优化，可以进一步提高处理速度。

- 3) 现有开发平台的资源尚未完全使用，还可以进行更加复杂算法程序的实验，数字助听器中可以移植更有针对性的汉语言声调强化处理等算法，专门针对汉语的发音特征，设计更适合国内听障患者使用的 DSP 数字助听器。
- 4) 本助听器系统的实现还需要通过真耳的验证。下一步工作将会考虑设计一套真耳测听方案，通过实际听障患者使用的对比，来判断系统的临床应用价值。

参考文献

- [1]. 张华.助听器.北京:人民卫生出版社,2004.134~140
- [2]. 姜泗长,顾瑞.临床听力学.北京医科大学中国协和医科大学联合出版社,1999.440~453
- [3]. 李小华,李雪琳,徐俊荣.基于 DSP 的数字助听器的研究.95 年生物电子学,医学传感器等联合学术会议文集,北京,1995:438~439
- [4]. 戴逸民,梁小雯,裴小平.基于 DSP 的现代电子系统设计.电子工业出版社.2002:22~34
- [5]. 刘松强.数字信号处理系统及其应用.清华大学出版社,1996:223~228
- [6]. 杨行俊,迟惠生,语音信号数字处理.北京:电子工业出版社.1995.
- [7]. 颜友钧,朱宇光.DSP 应用技术教程.中国电力出版社,2002:8~11
- [8]. 汪安民.TMS320C54XXDSP 实用技术.清华大学出版社,2002:92
- [9]. 彭启宗.TMS320C54X 实用教程.电子科技大学出版社 2000:3~7
- [10]. 张雄伟,曹铁勇.DSP 芯片的原理与开发应用.第二版,电子工业出版社,2001:37~38
- [11]. 肖宪波.数字助听器中若干主要算法的发展和现状.生物医学工程学杂志.2004;21 (4) :694~698
- [12]. 谭浩强.C 程序设计.第二版.清华大学出版社,2002:243~251
- [13]. 程佩青.数字信号处理教程.第二版.清华大学出版社,1997:215~228
- [14]. 胡广书.数字信号处理一理论、算法与实现.清华大学出版社.1997:129~23
- [15]. 宋仲康,胡智华.基于 TMS320VC5416 DSP 的音频接口设计.电子工程师.2005;31 (7) :24~26

- [16]. TMS320VC5416 Fixed-Point Digital Signal Processor Data Manual.Texas Instruments,2005
- [17]. TMS320C54X User's Guide. Texas Instruments Incorporated. 1999
- [18]. TMS320C54X DSP REFERENCE SET VOL.1 (CPU AND PERIPHERALS) .1996
- [19]. TMS320C54X DSP REFERENCE SET VOL.2 (MNEMONIC INSTRUCTION SET) .1996
- [20]. TMS320C54X Code Composer Studio Tutorial. Texas Instruments Incorporated.2001
- [21]. TMS320C54X Assembly Language Tools User' s Guide.Texas Instruments Incorporated. 2001
- [22]. TMS320C54X DSPBIOS User' s Guide. Texas Instruments Incorporated. 2001
- [23]. TMS320C54X Optimizing C Compiler User's Guide.Texas Instruments Incorporated.2001
- [24]. TLV320AIC23B Data Manual.Texas Instruments,2004
- [25]. Buechler M: How Good Are Automatic Program SelectionFeatures.The Hearing Review.2001;8 (9) :50,52~54,84.
- [26]. David Jiang.Application of digital technology in treating disorders of speech-language pathology.Chinese Scientific Journal of Hearing and Speech Rehabilitation.2005;13 (6) :52~55.
- [27]. N. Magotra,S. Sirivara.Real Time Digital Speech Processing Strategies for the hearing impaired. ICASSP'97. Munich Germany
- [28]. Keiichi Yasu, Kei Kobayashi, Kohshi Shinohara. Frequency compression of critical band for digital hearing aids.China-Japan Joint Conf. on Acoustics. Nanjing, China:2002.159~162
- [29]. TMS320C54x DSP Reference Set, Volume 1: CPU and Peripherals,Texas Instruments Incorporated,2001.24~36

- [30]. Vijaya Krishna G, Prasad SS, Patil KM. A New DSP-Based Multichannel EMG Acquisition and Analysis System. *Computers And Biomedical Research*, 1996, 29: 395~406
- [31]. Hsiang-Feng Chi, Shawn X. Gao, et al. Band-limited feedback cancellation with a modified filtered-X LMS algorithm for hearing aids. *Speech Communication*, 2003, 39 (1) : 147~161
- [32]. N. Magotra, S. Sirivara, "Real Time Digital Speech Processing Strategies for the hearing impaired," ICASSP'97, Munich, Germany
- [33]. C. T. Mullis, R. A. Roberts, *Digital Signal Processing*, "Addison Wesley, 1987
- [34]. L. R. Rabiner, R. W. Schafer, "Digital Processing of Speech Signals", Prentice-Hall, 1978
- [35]. S. Bangalore, "Quantization Analysis for Fixed-Point Implementation of Speech Processing for the Hearing Impaired," MS Thesis, EECE Department, University of New Mexico, 1999
- [36]. Marcio G. Siqueira and Abeer Alwan. Steady-state analysis of continuous adaptation in acoustic feedback reduction systems for hearing-aids, *IEEE Trans. on Speech and Audio Processing*, 2000, 8 (4) : 443~453
- [37]. Eri Moulines, Omar Ait Amrane, and Yves Grnier. The generalized multidelay adaptive filter: Structure and convergence analysis. *IEEE Trans on Sigal Processing*, 1995, 43 (1) : 14~28
- [38]. D. V. Anderson, R. W. Harris, D. Chabries, Evaluation of a hearing compensation algorithm, *IEEE ASSP-95*. 1995. 3531~3533
- [39]. D. Byrne, H. Dillon, The national acoustic laboratories' (nal) new procedure for selecting the gain and frequency response of a hearing aid, *Ear Hearing*. 1986. 257~265

文献综述

DSP 芯片的发展和应用

摘要：数字信号处理系统由于受运算速度的限制，其实时性在相当的时间内远不如模拟信号处理系统。从 80 年代至今的十多年中，DSP 芯片在运算速度、运算精度、制造工艺、芯片成本、体积、工作电压、重量和功耗方面取得了划时代的发展，开发工具和手段不断完善。DSP 芯片有着非常快的运算速度，使许多基于 DSP 芯片的实时数字信号处理系统得以实现。目前，DSP 芯片已应用在通信、自动控制、航天航空及医疗领域，取得了相当的成果。在生物医学工程领域，基于 DSP 芯片的技术具有广阔的应用前景。

关键词：数字信号处理 芯片 发展 应用

The Development and Applications of Digital Signal Processor

Abstract: Due to the limitation of operation speed, real time performance of digital signal processing system is far from that of analog signal processing system in decades ago. Since early 80's, DSP chips have been greatly improved in the following aspects: operation speed, computation precision, fabrication technics, cost, chip volume, operational power supply voltage, weight and power consumption. Furthermore, development tools and methods have been developed greatly. Modern DSP chips can be operated very fast, which make the implementation of many DSP based signal processing system possible. Now DSP chips have been widely applied successfully in communication, automatic control, aerospace and medicine. DSP based technology has very promising future in biological engineering area.

Key words: digital signal processing, chip, development, application

数字信号处理作为信号和信息处理的一个分支学科，已渗透到科学研究、技术开发、工业生产、国防和国民经济的各个领域，取得了丰硕的成果。对信号在时域及变换域的特性进行分析、处理，能使我们对信号的特性和本质

有更清楚的认识和理解,得到我们需要的信号形式,提高信息的利用程度,进而在更广和更深层次上获取信息。然而,数字信号处理系统由于受到运算速度的限制,其实时性在相当长的时间内远不如模拟信号处理系统,使得数字信号处理系统的应用受到了极大的限制和制约。自70年代末80年代初DSP(数字信号处理)芯片诞生以来,这种情况得到了极大的改善。DSP芯片,也称数字信号处理器,是一种特别适合进行数字信号处理运算的微处理器。DSP芯片的出现和发展,促进数字信号处理技术的提高,许多新系统、新算法应运而生,其应用领域不断拓展。

一、 DSP 芯片的发展

70年代末80年代初,AMI公司的S2811芯片,Intel公司的2902芯片的诞生标志着DSP芯片的开端。随着半导体集成电路的飞速发展,高速实时数字信号处理技术的要求和数字信号处理应用领域的不断延伸,在80年代初至今的十几年中,DSP芯片取得了划时代的发展。从运算速度看,MAC(乘法并累加)时间已从80年代的400ns降低到40ns以下,数据处理能力提高了几十倍。MIPS(每秒执行百万条指令)从80年代初的5MIPS增加到现在的40MIPS以上。DSP芯片内部关键部件乘法器从80年代初的占模片区的40%左右下降到小于5%,片内RAM增加了一个数量级以上。从制造工艺看,80年代初采用 $4\mu\text{m}$ 的NMOS工艺而现在则采用亚微米CMOS工艺,DSP芯片的引脚数目从80年代初最多64个增加到现在的200个以上,引脚数量的增多使得芯片应用的灵活性增加,使外部存储器的扩展和各个处理器间的通信更为方便。和早期的DSP芯片相比,现在的DSP芯片有浮点和定点两种数据格式,浮点DSP芯片能进行浮点运算,使运算精度极大提高。DSP芯片的成本、体积、工作电压、重量和功耗较早期的DSP芯片有了很大程度的下降。在DSP开发系统方面,软件和硬件开发工具不断完善。目前某些芯片具有相应的集成开发环境,它支持断点的设置和程序存储器、数据存储器 and DMA 的访问及程序的单部运行和跟踪等,并可以采用高级语言编程,有些厂家和一些软件开发商为DSP应用软件的开发准备了通用的函数库及各种算法子程序和各種接口程序,这使得应用软件开发更为方便,开发时间大大缩短,因而提高了产品开发的效率。

最成功的 DSP 芯片当数美国德州仪器公司 (Texas Instruments, 简称 TI) 的一系列产品。TI 公司在 1982 年成功推出其第一代 DSP 芯片 TMS32010 及其系列产品 TMS32011、TMS320C10/C14/C15/C16/C17 等。TI 将常用的 DSP 芯片归纳为三大系列, 即: TMS320C2000 系列 (包括 TMS320C2X/C2XX)、TMS320C5000 系列 (包括 TMS320C5X/C54X/C55X)、TMS320C6000 系列 (TMS320C62X/C67X)。如今, TI 公司的一系列 DSP 产品已经成为当今世界上最有影响的 DSP 芯片。TI 公司也成为世界上最大的 DSP 芯片供应商, 其 DSP 市场份额占全世界份额近 50%。

第一个采用 CMOS 工艺生产浮点 DSP 芯片的是日本的 Hitachi 公司, 它于 1982 年推出了浮点 DSP 芯片。1983 年日本 Fujitsu 公司推出的 MB8764, 其指令周期为 120ns, 且具有双内部总线, 从而使处理吞吐量发生了一个大的飞跃。而第一个高性能浮点 DSP 芯片应是 AT&T 公司于 1984 年推出的 DSP32。

与其他公司相比, Motorola 公司在推出 DSP 芯片方面相对较晚。1986 年, 该公司推出了定点处理器 MC56001。1990 年, 推出了与 IEEE 浮点格式兼容的浮点 DSP 芯片 MC96002。

美国模拟器件公司 (Analog Devices, 简称 AD) 在 DSP 芯片市场上也占有一定的份额, 相继推出了一系列具有自己特点的 DSP 芯片, 其定点 DSP 芯片有 ADSP2101/2103/2105、ADSP2111/2115、ADSP2161/2162/2164 以及 ADSP2171/2181, 浮点 DSP 芯片有 ADSP21000/21020、ADSP21060/21062 等。

目前各厂商生产的 DSP 芯片有: TI 公司的 TMS320 系列、AD 公司的 ADSP 系列、AT&T 公司的 DSPX 系列、Motorola 公司的 MC 系列、Zoran 公司的 ZR 系列、Inmos 公司的 IMSA 系列、NEC 公司的 PD 系列等。表 1-1 则是世界上主要 DSP 芯片供应商的代表芯片的一些数据。目前, 世界上的 DSP 有 300 多种, 其中定点的 200 种。TI 公司作为 DSP 生产商的代表, 生产的品种很多, 定点和浮点 DSP 大约占市场份额的 60%。可喜的是国内自主知识产权的 DSP 已经开发成功, 并推向市场。

二、通用 DSP 芯片的特点

DSP 芯片的内部采用程序和数据分开的哈佛结构, 具有专门的硬件乘法

器,采用流水线操作,提供特殊的 DSP 指令,可以快速的实现各种数字信号处理的算法。根据数字信号处理的要求,DSP 芯片一般具有如下特点:

- 1.在一个周期内可完成一次乘法和一次累加。
- 2.采用哈佛结构,程序和数据空间分开,可以同时访问指令和数据。
- 3.片内有快速 RAM,通常可以通过独立的数据总线在两块中同时访问。
- 4.具有低开销或无开销循环及跳转硬件支持。
- 5.快速中断处理和硬件 I/O 支持。
- 6.具有在单周期内操作的多个硬件地址产生器。
- 7.可以并行执行多个操作。
- 8.支持流水线操作,取指、译码和执行等操作可以重叠进行。

在目前市场上最常见的 DSP 芯片是美国德州仪器公司 TI (Texas Instruments) 的系列产品。主要有三大系列:

● C6000 TMS320C6000 系列为 TI 公司最高性能 16bit 定点 DSPs,低功耗主要应用于宽带网络和数字图像处理,包括定点处理器 C62x、64x 和浮点处理器 C67x。

● C5000 TMS320C5000 系列为 TI 公司性价比最高应用最广泛的 16bit 定点 DSPs,主要应用于个人、便携式网络和无线通信,包括定点处理器 C54x and C55x。

● C2000 TMS320C2000 系列为 TI 公司最优、最高性能的控制芯片 DSPs,最高的代码效率,对各种综合外设提供了最优的数字控制。包括定点处理器 C24x and C28x。

TI DSP 芯片具有更加适合数字信号处理的软件和硬件环境,可用于更复杂的数字信号的处理算法,是课题优先选择的处理器类型。

三、 DSP 芯片的分类

DSP 芯片可以按照下列三种方式进行分类。

1. 按基础特性分

这是根据 DSP 芯片的工作时钟和指令类型来分类的。如果在某时钟频率范围内的任何时钟频率上,DSP 芯片都能正常工作,除计算速度有变化外,没有性能的下降,这类 DSP 芯片一般称为静态 DSP 芯片。例如,日本 OKI

电气公司的 DSP 芯片、TI 公司的 TMS320C2XX 系列芯片属于这一类。

如果有两种或两种以上的 DSP 芯片,它们的指令集和相应的机器代码机管脚结构相互兼容,则这类 DSP 芯片称为一致性 DSP 芯片。例如,美国 TI 公司的 TMS320C54X 就属于这一类。

2. 按数据格式分

这是根据 DSP 芯片工作的数据格式来分类的。数据以定点格式工作的 DSP 芯片称为定点 DSP 芯片,如 TI 公司的 TMS320C1X/C2X、TMS320C2XX/C5X、TMS320C54X/C62XX 系列,AD 公司的 ADSP21XX 系列,AT&T 公司的 DSP16/16A, Motorola 公司的 MC56000 等。以浮点格式工作的称为浮点 DSP 芯片,如 TI 公司的 TMS320C3X/C4X/C8X,AD 公司的 ADSP21XXX 系列,AT&T 公司的 DSP32/32C, Motorola 公司的 MC96002 等。

不同浮点 DSP 芯片所采用的浮点格式不完全一样,有的 DSP 芯片采用自定义的浮点格式,如 TMS320C3X,而有的 DSP 芯片则采用 IEEE 的标准浮点格式,如 Motorola 公司的 MC96002、FUJITSU 公司的 MB86232 和 ZORAN 公司的 ZR35325 等。

3. 按用途分

按照 DSP 的用途来分,可分为通用型 DSP 芯片和专用型 DSP 芯片。通用型 DSP 芯片适合普通的 DSP 应用,如 TI 公司的一系列 DSP 芯片属于通用型 DSP 芯片。专用 DSP 芯片是为特定的 DSP 运算而设计的,更适合特殊的运算,如数字滤波、卷积和 FFT,如 Motorola 公司的 DSP56200, Zoran 公司的 ZR34881, Inmos 公司的 IMSA100 等就属于专用型 DSP 芯片。

四、 DSP 芯片的应用

随着 DSP 芯片性能的不断改善,用 DSP 芯片构造数字信号处理系统作信号的实时处理已成为当今和未来数字信号处理技术发展的一个热点。随着各个生产厂家研制的投入,DSP 芯片的生产技术不断更新,产量增大,成本和售价大幅度下降,这使得 DSP 芯片应用的范围不断扩大,现在 DSP 芯片的应用遍及电子学及与其相关的各个领域。典型应用主要有如下几点:

- (1) 通用信号处理:卷积,相关,FFT, Hilbert 变换,自适应滤波,谱分析,波形生成等。

- (2) 通信：高速调制/解调器，编/译码器，自适应均衡器，仿真，蜂房网移动电话，回声/噪声对消，传真，电话会议，扩频通信，数据加密和压缩等。
- (3) 语音信号处理：语音识别，语音合成，文字变声音，语音矢量编码等。
- (4) 图形图像信号处理：二、三维图形变换及处理，机器人视觉，电子地图，图像增强与识别，图像压缩和传输，动画，桌面出版系统等。
- (5) 自动控制：机器人控制，发动机控制，自动驾驶，声控等。
- (6) 仪器仪表：函数发生，数据采集，航空风洞测试等。
- (7) 消费电子：数字电视，数字声乐合成，玩具与游戏，数字应答机等。
- (8) 军事：保密通讯、雷达处理、声纳处理、导航、全球定位、调频电台、搜索和反搜索等。

DSP 在医学电子学方面的应用也越来越广，如同其它数字图像处理一样，DSP 芯片已在医学图像处理，医学图像重构等领域得到了广泛的应用，已取得了令人满意的效果。在助听器，电子耳蜗等方面也取得了相当的进展。国内、外也有关于脑电、心电、心音和肌电信号处理方面基于 DSP 芯片系统的报道。我们认为在生理信号处理领域基于 DSP 芯片的技术可以解决我们在实际工作中遇到的某些问题，如当生理信号数据量很大（如脑电，肌电等）且处理算法相对复杂时，现有的微机在实时采样、处理、存储和显示方面往往不能满足实际应用要求，而基于 DSP 芯片的高速处理单元和微机构成主从系统可以较好地解决这类问题。

结束语

以 DSP 芯片为核心构造的数字信号处理系统，可集数据采集、传输、存储和高速实时处理为一体，能充分体现数字信号处理系统的优越性，能很好地满足载人航天领域设备测量精度、可靠性、信道带宽、功耗、工作电压和重量等方面的要求。目前，DSP 芯片正在向高性能、高集成化及低成本的方向发展，各种通用及专用的新型 DSP 芯片在不断推出，应用技术和开发手段在不断完善。这样为实时数字信号处理的应用——尤其是在生物医学工程领域中的应用提供了更为广阔的空间。我们有理由相信，DSP

芯片进一步的发展和应用将会对生物医学信号处理领域产生深远的影响。

参考文献

- [1]. 刘松强.数字信号处理系统及其应用.北京:清华大学出版社,1996:223~228.
- [2]. TMS320C54x DSP Reference Set, Volume 1: CPU and Peripherals,Texas Instruments Incorporated,2001.
- [3]. 颜友钧,朱宇光.DSP 应用技术教程.中国电力出版社,2002: 8—11.
- [4]. 张雄伟,曹铁勇.DSP 芯片的原理与开发应用[M]. 第二版,北京:电子工业出版社,2001:37-38.

攻读学位期间发表文章情况

- [1]. 应俊,陈广飞.无线体温监测系统的研制.北京生物医学工程.2006,25(1):73~75
- [2]. 应俊,陈广飞. 基于 DSP 数字助听器设计和实现.医疗卫生装备.2006 (已录用)
- [3]. 陈广飞, 应俊. 基于 DSP 数字助听器算法研究. 北京生物医学工程.2006 (已录用)

致 谢

谨向所有指导过、帮助过我的老师和同学表示最真挚的感谢！

本文是在导师精心指导下完成的，我非常感谢导师陈广飞老师，从论文的选题、研究到撰写，陈老师都给予了我悉心的指导和无私的帮助，正是由于导师倾注了大量的精力和时间，我的论文才得以顺利的完成。三年来，给我印象最深的是导师渊博的学识，严谨的治学态度和高尚的精神风范。他让我看到一名科技工作者所应具备的那种踏实勤恳、一丝不苟、实事求是的优良品质和工作作风。在导师门下的这段日子里，导师首先教的是如何做人，其次教的才是如何作学问。师恩浩荡，只有在今后的人生路上，我将牢记师训，努力上作。

三年的学习过程中，我还得到我院耳鼻喉科郝昕老师和北京理工大学卜祥元老师的指导和帮助，老师的点拨使我茅塞顿开，无数关键的技术问题能够迎刃而解，在此表示由衷的感谢！

三年的学习过程中，我得到医学保障部以及医学工程保障中心各级领导的关心和照顾，他们在做人做学问上给予我的教诲使我终生铭记，在此表示衷心的感谢！

在论文的完成过程中，还得到了众多同窗的帮助，他们在平时的学习生活中给予我许多的帮助，让我体会到工作中和睦愉快的气氛，感受到真挚的友谊，为我的硕士生涯留下了美好的回忆，值此论文完成之际向他们表示深深的谢意。

感谢我的父母，感谢这么多年来父母对我的培养和支持！

感谢解放军总医院，感谢社会，为我提供了优越的成长环境，使我能够在不断汲取养分中取得进步。

应俊于解放军总医院医学计量研究室

2006.5