

PTcheck检测报告简明打印版

对比结果（相似度）：

总相似度：7%（相似字数占总字数的百分比）

红色相似度：1%（句子相似度70%-100%的字数占总字数的百分比）

橙色相似度：6%（句子相似度40%-70%的字数占总字数的百分比）

验真伪二维码



编号：6382feb0-c659-4e9d-b8c9-02759b8172ce

标题：2647220997972225

作者：26472209979722251

长度：44913字符（不计空格）

段落：498个

句子：1185句

词语：27453个

时间：2017-3-22 14:49:37

比对库：学术期刊、学位论文、互联网资源

相似资源列表：

(1) 7%即3143字来源于其他网络

全文简明报告：

绪论

课题来源及背景

{48%:随着现代社会的高速发展，听力损伤和听力障碍成为人们正常生活的一大困难。}出现听力问题的原因多样：工作或生活中的各式环境噪声；随着年龄增长而出现的听力衰减；新生儿的先天性听力障碍等[1]。{61%:世界卫生组织（World Health Organization, WHO）研究显示[2]，听力障碍患者的人数是当前残疾人中数目最多的一类。}{43%:早在2006年，国家发布的残疾人调查数据显示，当时我国有听力残疾人士2780万，其中很大一部分是老年人，然而仅三年之后（截止于2009年），}{46%:来自英国医学听力研究学会的一份统计报告表明，全世界大约有 6 亿人患有听力损伤疾病，而其中有 1.4 亿的听力障碍患者就来自我国[3]，由此可见，}听力障碍问题可见一斑。听力障碍人群数目增加一方面是由于轻微的听力损伤并不会对正常生活造成太大影响，普通人缺乏保护听力的意识，也不了解听力保护的相关知识，持续的听力损伤使得人们的听力水平下降到一定水平，最终影响人们的正常生活和交流[4]；{40%:另一方面，随着年龄的增长，身体生理机能的逐渐衰退也造成许多老年人的听力障碍，普通民众缺乏听力保护意识和知识，对老年人的听力障碍现象不闻不问、置之不理。}听力水平下降会导致人的交流障碍、听障人群参与社会活动的的能力降低、社交活动减少，严重的出现心理障碍、自闭，甚至会诱发老年痴呆[7]。因此，听力障碍人群的生活质量是社会亟待关注和解决的问题之一。

针对上述听力障碍的问题，人们使用助听器帮助听力障碍人群恢复听觉功能、提高生活质量[5]。然而助听器并非只是一个简单的音频放大器，它十分精密，需要根据听力损伤情况进行科学的验配，方可达到效果，如若验配不佳，甚至会反过来损伤听力[7]。此外，由于技术垄断等原因，只有少数的听力障碍患者能佩戴助听器[5]。数据显示，中国助听器选配比例仅在5%左右[7]；而在佩戴助听器的听力障碍患者中，由于听力检查和验配等不合适也导致许多患者不满意，效果不佳。{71%:据美国著名的助听器市场调查研究 MarkeTrak VII报告数据显示，只有71%的助听器使用者对于所使用的助听器比较满意[6]。}提高听力障碍患者的听觉功能面临巨大挑

战。助听器的验配对提高听障患者使用助听器的体验和效果至关重要，而听力检查和测试又是助听器验配的主要内容。目前而言，听力检查主要是在各省市大型医疗机构中完成，其所有到的设备大多需要进口，设备成本较高，覆盖面窄；检查过程中需要专业人士对待测者的反映进行判别，操作繁琐漫长[4]。限于这些困难，许多需要进行听力检查或是想通过听力检查来了解自身听觉功能的人群很难接受听力检查。

课题目的及意义

针对听力检查难，听觉保护意识薄弱等问题，本课题利用语音信号处理技术和听力检查相关标准，构建基于移动终端的听力测试系统，并研究听觉补偿算法以及如何在移动端实现。将基本的听力检查项集成至移动终端可使大多数人都能便捷地进行纯音听力测试、听力分辨力测试以及言语测听等基本听力测试项，有助于听障患者的助听器验配，也可方便普通人进行听力测试，提高大众的听觉保护意识，对听力障碍的预防和听力康复有指导意义[9]。通过广泛使用的移动终端设备（手机、平板等）传播听力保护知识，并可针对听力验配中较为简单的项目进行检测，帮助使用者了解自己的听力状态，并由此提高使用者的听力保护意识。此外，针对于助听器的语音补偿算法进行研究，有助于提高助听效果，通过和相应的听障人的听力测试结果进行结合分析，可达到更好、更精确的助听效果[8]。因此，对基于移动终端的听力测试系统和相应的提高助听器助听效果的语音处理算法的研究具有重要的实际意义。

国内外研究现状分析

1.3.1 听力测试

针对听力测试及其相关设备的研究开始于19世纪末。英国著名发明家David Edward Hughes最早与1879年在名为“感应电流平衡和实验研究”一文中提出听力计一词，并发明了世界上第一台听力计；{50%:实验家、作家Benjamin Ward Richardson对听力计给予高度评价并选在医学界宣传听力计；} {95%:1912年，Hawksley将听力计投入商业化生产；}此后听力计发展进入一个断档期，并在1931年由Brown等人重新提起；1932年，Watkyn Thomas和Yates提到电子管听力计的测试；二十世纪四十年代，B&K公司设计出首台听力计产品，并广泛应用与医疗和科研领域[10]。此后，随着电子技术的飞速发展，听力计的硬件和软件都得到了革新，集成电路、液晶屏、DDS信号合成技术，DSP技术等等，使得听力计越加精确和方便；{86%:美国、德国和丹麦等国家许多企业均在致力于听力计的研究和创新；} {58%:1999年已出现使用DDS技术的高精度听力计；2004年移动手机听力计出现；}美国智听公司研制出基于PC设备的听力检查系统，通过对产品软件进行设计，使其具有良好的图形界面，方便用户操作使用[11]；在进入信息时代的21世纪，随着无线通信技术和互联网技术的不断发展，Ykhlef Fayal等人利用计算机编程提供虚拟听力计，实现标准听力计的功能[12]；2013年，Daoyuan Yao和Gregg Givens等人将无线通信技术引入到远程听力评估系统中[13]；最近，Alba Fernandez和Marcos Ortega等人对存在认知下降状况的听障患者在自动听力评价中的手势反应及相关问题进行研究[14]。

国内对于听力测试和听力计的研究起步较晚，但是发展迅速。{74%:在20世纪50年代听力计便在我国一些医院中使用；} {62%:1964年我国第一部听力学著作《听力学概论》出版[15]；}之后我国听力康复研究不断发展，从进口听力检查设备到自主生产和设计听力计等医疗仪器；{56%:从模拟技术到数字技术、直接频率合成技术；}从听力检查的无人问津到测听国家标准的制定，我国听力学和听力康复治疗已得到较好发展。近年来，在国家测听标准的基础上，针对测听操作流程的优化，以及测听设备的数字化，测听软件的智能化研究日益广泛。周洋等人针对听力检查内容及流程进行优化，并开发相关的计算机软件，将听力检查操作集成至个人PC终端，使得听力检查操作更加便捷[4]。随着听力学研究者和听力康复临床工作者们的深入研究，综合国内外的听力测试研究不难看出，目前听力检查主要向着便捷化、多功能、便携化及智能化方向发展。

1.3.2 助听补偿

助听补偿算法的研究伴随着助听设备的发展而前进。人类使用手掌来汇聚传入人耳的声音，起到放大的效果。可以说，手掌是人类最原始的助听器[11]，而最原始的助听补偿算法即是简单的放大所接收到的声信号。20世纪初的磁晶助听器能提供20~30dB的补偿增益；到20世纪20年代，电子管类助听器获得了更强的功率，可以给出最高70dB的补偿增益，然而，电子管助听器直到世纪中叶才变得方便携带、实用，并借此迅速占领助听器市场。{42%:20世纪50年代末，集成电路、晶体管的迅猛发展使得助听器进入晶体管时代。}晶体管助听器的体积已经非常小巧，助听器的发展开始体现在其助听效果包括输出语音的清晰度、可懂度、舒适度等，此时也正是助听

补偿算法真正开始发展的时机。{62%:随着计算机技术的发展,数字技术引入助听器行业,相关的数字信号处理技术被渐渐地带入数字助听器。}助听补偿方案从最初的线性放大到压缩放大、分频段补偿、多通道补偿以及针对特殊听力损失患者的移频助听等,数字助听器的补偿效果随着补偿算法的不断提升、优化而越加进步。近年来,国外学者对多通道响度补偿、子带划分、子带滤波器设计[65]、窄带响度补偿[64]、非线性响度补偿[63]、助听器频移电路[62]、非线性频移压缩[45]以及降频助听器[57]均有相关研究成果。

国内关于助听器助听补偿算法的研究起步较晚,但是研究发展迅猛。{53%:至今国内关于语音信号及助听器相关数字处理算法的研究已取得较多成果。}{67%:关于响度补偿类算法特别是多通道响度补偿算法的研究,}国内东南大学[5][11]、南京邮电大学[66][67]等相关学者已有较为深入的研究。近年来,响度补偿类算法开始更多的关注听损患者本身的听力情况[66],并重视助听器的去噪能力,利用多种手段去除补偿后语音信号的噪声[68],提高输出语音的辨识度,同时关于响度补偿算法的自适应调整技术的研究也有学者取得一定进展[69]。此外,国内学者对于针对重度听损患者的移频助听及降频助听算法的研究也有较多进展[58][59][60],关于重度听损患者的移频助听效果对比也有学者进行详细分析对比[70][71]。随着助听补偿算法的不断发展,不难看出,患者本身的听力情况将得到更多的重视,助听算法将更智能化,针对具体患者的定制性更为明显,同时,{55%:重度听损患者的补偿方案的研究也将获得更多关注。}

本文研究内容和结构

本课题主要针对听力测试以及助听器补偿算法进行研究。其中听力测试系统是基于移动终端的软件,主要研究纯音听力测试,听力分辨率测试,言语识别率测试和言语识别阈测试等内容及其如何在移动端实现;研究助听器补偿算法,针对听力测试给出的结果合理的进行助听补偿使助听器更好的服务于听障患者,并在移动终端实现补偿算法。本文各章节内容安排如下:

{60%:第一章介绍课题背景意义以及国内外相关研究的进展现状。}本章主要阐述了听力障碍问题背景,包括听力障碍人群数目增长、听力障碍产生原因多样以及人们目前对待听力问题的状态等;随后阐述基于移动终端进行听力测试及助听补偿研究的意义;最后本章对听力测试和助听补偿相关的研究现状进行分析。

第二章介绍了助听器听力测试相关理论基础。本章首先介绍了人类听觉系统详细结构及声信号传导原理,掌握人类听觉系统构造以及声信号传导机制对分析听力损失原因、设计助听补偿算法均有原理性指导意义。随后本章对纯音听阈测试、频率分辨力测试以及言语测听基本定义及相关国家标准进行介绍。最后本章详细分析听力损伤定级标准以及常见听力损失类型。

第三章主要阐述基于移动终端的听力测试系统设计方案。本章首先分析听力测试系统基本需求和指标;随后描述听力测试系统的整体架构;然后,本章对纯音听阈测试、频率分辨力测试及言语测听等具体测试项的测试流程进行设计,详细描述各项测试的基本步骤和方法,并对某些测试流程结合移动终端特性进行改进,降低其测试复杂度。

第四章主要提出一种基于共振峰的谱对比增强补偿算法。本章首先对语音信号的基本特征和数字模型进行介绍,并对传统响度补偿算法基本原理进行研究。随后,将谱对比增强处理引入响度补偿算法,并对共振峰检测技术以及处方公式进行分析。最后,仿真实验表明基于共振峰的谱对比增强算法补偿效果较好,证明了该算法的可行性和有效性。

第五章主要提出一种频谱伸缩助听算法。本章首先对频率助听类算法的背景意义进行分析,表明频率补偿算法对重度听损患者的重要性。随后,对传统的移频降频补偿方法进行介绍,并分析其局限性和不足。然后本章对频率伸缩算法进行详细阐述,包括其具体原理及参数计算方案。最后,仿真实验表明该算法具有良好的适应性和兼容性,能够满足更多听损类型患者的需求。

第六章主要对基于移动终端的助听补偿系统进行分析,并给出听力测试及助听补偿系统整体的设计方案。同时,本章对系统中各项测试及功能进行测试。

第七章主要对本文工作进行回顾和总结,同时对系统的可完善方向以及助听算法的可优化方向进行展望。

助听器听力测试理论基础

生活因聆听变得更加便捷、美好，失去听觉功能的人生无疑是枯燥痛苦的。人类由耳接收外界声音，经由复杂而精妙的听觉系统使大脑感知声音。然而，由于长期暴露于噪声环境、先天缺陷、意外伤害等原因，人听觉系统受到破坏，使得听觉功能变弱，甚至丧失[16]。听力障碍和听力缺失困扰着听障患者的生活。随着科技不断发展，助听器被听障人群广泛使用，以此改善听力情况。随着助听器的不断推广和普及，其相关的验配方式和方法受到更多学者的关注，对听障患者合理有效的听力损失测定将帮助其完成助听器验配，从而更好的享受有声生活。本章首先介绍人耳听觉系统，然后给出助听器听力测试常见测试项进行简单介绍，最后介绍听力损失的定级及听损类型划分。

听觉系统

人类听觉系统主要由外周围听觉系统和听觉神经中枢组成[17]。外周围听觉系统主要由耳器官组成，负责将外界声信号转换成神经信号；神经中枢则负责将转换后的声信号传输至大脑，引起听觉。

外周围听觉系统

{53%:外周围听觉系统主要由耳器官构成。}人耳可感受声压的范围跨越近10个数量级，最微小的可感知声压约为，感知能力和分辨能力极强，为人接收外界声信号提供生理基础。人耳生理结构如图所示，按照相对位置关系可将耳器官分为：外耳、中耳和内耳三部分。

图2-1人耳结构

外耳包括耳廓和外耳道两部分。耳廓形状类似于扇形，具有一定的纹理，并有一定的凹陷形成一个半边的小腔，小腔汇集口便是外耳道入口。耳廓主要负责收集人前方和侧方向的声波信号，对来自后方的声信号具有一定的阻挡和抑制效果，通过对来自不同方向的声信号进行相应汇聚或是抑制使得耳廓给人听觉处理系统提供一定的声源定位信息。外耳道形状为一圆形小管，长约2.2cm，连接着耳廓中心的小腔口和中耳的鼓膜。外耳道负责将耳廓汇聚的外界声音传导至鼓膜处。由于外耳道的腔体特性，使得其具有一定的共鸣效果。研究表明，外耳道腔体的谐振频率中心约为3.4kHz，此时强度增益约为10dB，可见外耳道对语音信号高频部分具有一定的提升效果，而高频部分往往决定着语音信号的语义和细节信息，因此，外耳道的共鸣特性有助于提高言语可懂度，起到一定的放大作用。

经过外耳部分的聚集之后，声音传至中耳部分。中耳由鼓膜、听小骨和鼓室三部分构成。{65%:声波由外耳道传至鼓膜。鼓膜呈薄膜状，传入声波作用致使其进行机械振动。}听小骨连接于鼓膜，鼓膜的振动带动着听小骨的振动，听小骨所连接形成的听骨链为一杠杆系统，其另一端连接着镫骨，镫骨进一步将振动传送至卵圆窗膜。同时，鼓室内的两块小肌肉亦控制着听小骨的振动，当传入声音强度过大时，这两块肌肉提供消减听小骨振动幅度的机制，起到保护内耳的作用。听骨链的杠杆作用，以及镫骨底板所连接的卵圆窗膜与鼓膜面积的缩小，提供了较强的声压放大效果。

{51%:声振动信号传至鼓膜后经听骨链传递引起耳蜗中淋巴液和基底膜的振动，由此将声信号传至内耳。}{54%:内耳主要包含耳蜗、耳蜗神经、半规管和前庭等部分，其中耳蜗是人耳听觉传导中最重要也是最复杂的器官。}耳蜗由其独特的螺旋形状而得名。{85%:声波振动可以引起耳蜗中的淋巴液和基底膜的振动，使得耳蜗科蒂器官毛细胞产生兴奋，而听神经纤维就处在毛细胞下方的基底膜周围，}声波振动信号触发听神经纤维上的神经冲动。依附在基底膜周围的听神经纤维接收基底膜振动信号，而基底膜针对不同频率的声波会在不同段处产生共振，并将此振动信号经由听神经传至大脑皮层产生听觉。由此可见，基底膜及相关的听觉传导神经具有一定的频率选择特性，实际上这种选择特性十分精确，对听觉补偿研究具有指导意义，有学者根据基底膜的频率选择特性设计多通道响度补偿算法，取得较好的效果[11]。

听觉神经中枢

经过内耳处理后的声信号转变为神经信号，该信号从蜗神经节开始由听觉神经中枢传导至大脑。听神经中枢传导通路结构如图2-2所

示:

图2-20听神经中枢传导通路结构

听觉神经中枢传导线路主要包含耳蜗核、下丘脑和大脑皮层等部位[18]。在听觉中枢系统中,听觉神经信号的传递存在多条并行线路,各线路间互相交错,其传导系统十分复杂,对应着人听觉系统处理声神经信号的复杂多样性[19]。蜗神经节传导信号首先由蜗根进入耳蜗核,它是听觉神经中枢的起始位置。耳蜗核包括蜗神经前核、蜗神经后核以及上橄榄核等部分。耳蜗核所包含多种神经元,这些神经元的特性使得耳蜗核具有增强和保存声信号谱信息的作用。同时,上橄榄核对声音定位起着重要作用。听觉神经信号经耳蜗核处理后到达下丘脑。下丘脑中神经元通过轴突与内侧膝状体相连。内侧膝状体重新处理听觉神经信号后将其传入大脑皮层,由此引起听觉。

听力测试方法简述

对于听觉功能正常的人,其可听频率范围为20~20000Hz,所能承受声压范围从 数量级至 数量级,可听范围广阔。而对于听力有所损伤的患者而言,其听觉功能已经发生一些变化,如何合理定量表征听力损失患者的听力情况,是听力学研究的一个长久话题[20]。从19世纪末针对听力计和听力检查的研究开始,经过一个多世纪的发展,听力检查测试方法和操作步骤已基本标准化。通过对听障患者各项测听指标的测量,可以对其听力损失程度、类型以及采用何种治疗方式有更具具体更全面的认识。{42%:常见听力测试指标主要有听阈值、舒适阈值、痛阈值、言语识别率、言语识别阈和频率分辨率等,}{45%:相对应的测听内容为纯音听阈测试、言语测听和听力分辨率测试[21]。}

纯音听阈测试

纯音听阈测试是一项基本的听力测试项,其历史可追溯至听力计一词的诞生[22]。纯音测试在听觉障碍患者的临床诊断中是最普及、最基本的检查项,其测试结果对患者助听器验配具有指导意义。从第一台听力计开始便有此项测试内容,随着听力学的不断发展纯音听阈测试已经标准化,并在临床上广泛用来测试受试者的听觉灵敏度,获得受试者的听阈、舒适阈和痛阈值。

国家标准GBT 16403-1996为纯音听阈测试制定了详细的操作标准和流程,可分为加掩蔽和不加掩蔽的测量方式,其中不加掩蔽的听阈测试又可分为上升法和升降法[23]。测试过程中受试者处于安静环境之下,通过气导或骨导方式不断给声并通过受试者的主观反映来调节给声响度的大小,最终确定受试者在各个频率纯音的听阈。

言语测听

在实际生活中听障患者言语能力的缺失是其主要特征,而纯音测试等无法评价宽带的实际言语信号。实践发现,言语能力的下降有时并不一定伴有纯音听敏度的变化,表明纯音测试的结果不能准确的表达受试者的言语理解力[26]。人类听觉系统包括外周听觉的灵敏度和中枢听觉系统对听信息的处理,前者由纯音测试结果表征,后者可由言语测试结果来反映。言语测听一般以日常使用的双音节词作为测试材料,其测试结果不仅可以判断受试者的听觉敏感性,也反映出其对语言的理解能力。

{60%:言语测听的方法参照国家标准GBT17696-1999。}{47%:言语测听的测试内容项较多,其中最常被使用的是言语识别率和言语识别阈测试[24]。}前者是指,对一受试者,以一定的给声方式,在指定的声压级下,其能正确识别的检查项的数目所占总检查项数目的百分比,即为该语言级下的言语识别率;后者是指,对一受试者,以一定的给声方式,言语识别率为50%时所对应的声压级。

频率分辨力测试

听觉障碍患者听力能力下降一方面是由于听阈偏高,对正常响度的声信号无法感知,这类患者可以通过佩戴助听器,提高外界声音的响度,使其高于患者的听阈值从而使患者可以感知声音。然而,如果患者对于音频信号的频率分辨能力下降严重,或者说其频率分辨力过低,则会影响其对语言理解能力的下降,即患者可以听到声音,但是听不懂声音的语言信息[27]。针对频率分辨力缺失患者的助

听器验配难度较大，普通的助听器甚至无法达到良好的助听功效，一般对于频率分辨力及听力损失严重的患者需进行相应的频率补偿，以提高助听效果

受试者的频率分辨力一般由其心理物理调谐曲线给出。心理物理调谐曲线测量基本思想是固定参考信号的频率及声强，通过不断改变掩蔽声的声强和频率，并接收是否能听到参考声的反馈信息，以此判断受试者针对参考信号频率的辨别能力。心理物理调谐曲线的测量过程比较复杂，测试时间过长。针对这些缺陷，有学者提出快速心理物理调谐曲线测量方法，以改善其不足[35]。

听力损失定级和类型

听力损失 (Hearing Loss) 一般是指人听觉系统某处发生障碍，导致其听力能力与正常人相比出现不同程度的下降，即人们常说的耳聋、耳背[28]。听力损失常用纯音测听结果来描述，表现为听力图中某些频率点处听阈值高于正常人所测得听阈值。{40%:根据世界卫生组织 (WHO) 预防聋和听力损失项目会议报告 (1997, 日内瓦)，听力损失划分为三级，具体划分如下表[29]: }

表2- 1 WHO推荐听力损失级别划分

听损级别听损程度平均听阈值 (dB HL)

I极重度 81

II重度61~80

III中度40~61

其中，{66%:平均听阈值为500Hz、1000Hz、2000Hz和4000Hz四个频点处听阈值的均值。}由上表可见，世界卫生组织推荐的听力障碍等级划分为三个级。由于我国各类残疾分级均分为四个级别，故在WHO推荐的划分标准上，对极重度听障级别再进行细分，使其满足国内各类残疾标准划分的一致性，符合国情亦参考国际标准，其具体划分细节如下[29]:

{71%:听力残疾一级：较好耳平均听阈值大于等于91dB HL，听觉系统功能基本丧失，损伤极度严重，在没有外带设备协助情况下，听觉系统几乎无法感知外界声音，}不能通过言语信号进行正常交流；

{72%:听力残疾二级：较好耳平均听阈值在81dB HL~90dB HL之间，听觉系统功能损伤严重，正常言语交流存在严重障碍，无法聆听正常交流声强级别语音，}只能感知声强较大的声音；

{43%:听力残疾三级：较好耳平均听阈值在61dB HL~80dB HL之间，听觉系统中度损伤，正常声音交流亦存在障碍，能够聆听部分声强偏强的声音信号，}存在中等程度的言语交流障碍；

{40%:听力残疾四级：较好耳平均听阈值在40dB HL~ 60dB HL之间，听觉系统功能有损伤，并引起一定的言语声音交流障碍，能够聆听部分日常交流声，}然而无法完全听清，存在轻度言语信号交流障碍。

{62%:一般临床而言，听力损失被笼统地划分为传导性听力损失和感音神经听力损失两大类。}听力损失既可以是先天而得，也可以是后天获得性的，而先天性听力损失既可以是遗传而得也可能是非遗传的。而后天获得性的听力损失一般与患者生活环境、经历密切相关。长期暴露在噪声环境、药物中毒、年龄增长、自身免疫失调和前庭神经细胞瘤等因素是导致感音神经性耳聋的主要原因；而中耳炎、鼓室硬化、鼓穿孔和耳硬化等病症是造成后天传导性耳聋的主要因素。

随着对听力损失的研究，人们对其认识不断深入。根据不同类型听力损失在病理生理学与解剖学上所存在的差异，{48%:有学者将听力损失细分为：传导性听力损失、耳蜗性听力损失、神经性听力损失、反馈性听力损失和中枢性听力损失五类[30][31][32][33]，}

简要介绍如下：

传导性听力损失：主要是指外耳至内耳传导通路出现障碍而导致的听力损失。{49%:因此，很多引起此类阻碍的病因都会导致传导性耳聋发生，比如：外耳道被耵聍阻塞、先天性外耳道畸形、中耳积液以及外耳道塌陷等。}另外，鼓膜的损伤和缺陷也是引起传导性听力损失的主要原因。一般而言，传导性听力损失通过助听矫正等治疗手段可以恢复至正常听觉功能。

耳蜗性听力损失：主要是指由耳蜗损伤或缺陷而导致的内耳功能缺失引起的听力损失。常见病因为内外毛细胞损坏或缺失。由上人耳听觉系统分析可知，耳蜗对人听觉系统而言至关重要。耳蜗外毛细胞的损伤常引起听觉灵敏度丧失，高频听力损失严重。一般耳蜗损伤导致听力损失可通过人工耳蜗植入技术进行改善。

神经性听力损失：主要是指听觉传导中枢上听神经元损伤而造成的听觉功能缺失。其特点为外周围听觉系统工作正常，听觉中枢出现传导障碍，又被称为听神经病。此类听力损失常伴随听觉系统对语音时域处理能力的下降，造成言语感知困难。

反馈性听力损失：主要是指听觉系统反馈控制机制损坏导致的听力损失。此类损伤研究偏少，主要由听觉系统对听觉信息的反馈控制下行通路出现障碍造成，比如镫骨肌反馈控制听骨链故障、耳蜗核上橄榄核反射障碍等。

中枢性听力损失：主要是指非外周围听觉系统故障所造成的听力损伤。该类听力损失的研究属于新兴领域，与生理学、神经科学等多方面密切相关。

2.4 本章小结

本章主要对助听器听力测试相关基础理论进行介绍。首先介绍了人类听觉系统基本概况，并针对听觉系统的传导通路已经听觉神经中枢相关内容进行分析。人类听力障碍起因与人类听觉系统的生理特性及其构成密切联系，听觉系统中不同部位出现问题，所导致的听力障碍情况均有不同，了解人类听觉系统有助于理解听力障碍的本质。

随后本章对纯音听力测试、言语测听以及频率分辨力测试相关的理论及方法进行简要阐述，为后续章节设计测试流程提供理论基础。

最后本章介绍了听力损伤定级的详细内容，并对听力损失类型进行分析。听力损伤定级是一种共识性的人为定义标准，关于听力损伤方面的研究仅需遵照划分即可。不同程度听力损伤的定量划分对后续助听补偿算法的参数确定具有指导意义。

基于移动终端的听力测试系统

在听力测试相关理论的基础上，本章基于Android移动终端操作系统，设计听力测试系统指标以及各项测试操作流程，构建听力测试系统。

系统基本功能需求

综合考虑助听器验配的相关需求、补偿算法所需验配数据以及移动平台本身的特点，听力测试系统的功能需求主要有以下方面：

1. 系统具有纯音信号产生功能，可以根据频率信息产生对应的纯音信号；
2. 具有纯音听阈测试功能，可以给出受试者的听力图；
3. 具有音调分辨力测试功能，给出受试者的频率分辨力图；
4. {42%:具有一定的语料库数据，并能够测试受试者的言语测听指标，包括言语识别率和言语识别阈；}

5. 具有良好的人机交互界面，可以保存用户的验配数据。

系统设计所参照的技术指标约定如下：

1. 频率范围：125Hz-8KHz中11个频点（125Hz，250Hz，500Hz，750Hz，1KHz，1.5KHz，2KHz，3KHz，4KHz，5KHz，8KHz）；
2. 最大声压级100dB HL；声压级误差在 ± 3 dB内；
3. 频率误差小于1%；
4. 环境噪声小于40dB；

测听系统总体架构

测听系统的整体架构如图3-1所示。基于Android操作系统，利用移动终端硬件资源构建纯音信号生成器模块，并将其用于纯音听阈测试和频率分辨力测试；将专业的言语测听所用语音资源内置于系统中，提供言语测听平台；耳机、触摸屏等硬件完成受试者和软件系统的交互，得到检测结果。

图3-1测听系统架构框图

测听系统主要由硬件和软件两部分构成。硬件部分主要包括CPU、声卡、显示屏以及耳机等等移动设备所具有的硬件资源；{42%:软件部分主要基于Android操作系统编写。各测试模块流程主要通过软件实现；}交互部分由软硬件共同完成。

听力测试软件设计架构

听力测试系统主要功能通过软件实现，{43%:其功能主要包括：纯音听阈测试、频率分辨力测试（音调分辨力测试）和言语测听（包括言语识别率和言语识别阈）。}听力测试软件基于Android操作系统，在详细研究国家相关标准的情况下，设计软件流程，完成各听力测试项并保存测试结果，指导助听器验配，并为后续补偿算法提供参数。

纯音听阈测试

根据针对听力测试方法的国家标准[GBT 16403]，参考其气导测听方法可知，听阈测试的可分为上升法和升降法，两者测试流程稍有不同[23]。在此标准基础上，本测听系统对其测试流程稍作修改，在不改变任何测试结果的前提下，使该操作流程更加适合智能移动终端软件设计以及方便受试者操作。

上升法测听阈的操作流程如下：

从1KHz频点处开始测试，以上一频点的听阈或初始化的声压级作为初始声压级，以5dB一档逐渐增大给声声压级直至受试者反馈听到声音；

以10dB为一档降低纯音声压级，直至受试者反馈未听到声音，而后每5dB为一档上升给声声压级，直至受试者反馈听到声音；

判断是否已是在上升过程中在此声压级做出反应次数已达三次，若是则该声压级即为对应频点的听阈值；若不是则继续2)的操作；

保存测试信息，若还有未测试频点则切换频点从1)开始测试新频点；若已是最后的1KHz频点上的测试，则判断最初的1KHz的测试结果

与本次测试结果是否相差过大，若相差过大则本次所有频点的测试结果均无效，需重新测试，否则测试完毕。

上升法测试流程框图如图3-2所示：

图3-2上升法测听阈流程图

图中上升转换点的定义为，从受试者听不到测试音开始，在增强给声声压级的过程中受试者第一次做出听到测试音的反馈所对应的测试点为上升转换点；同理，下降转换点为从受试者听到测试音开始，在降低给声声压级的过程中，受试者第一次做出听不到测试音的反馈所对应的测试点为下降转换点。

升降法与上升法稍有不同，主要体现在第二步，操作流程如下：

从1KHz频点处开始测试，以上一频点的听阈或初始化的声压级作为初始声压级，以5dB一档逐渐增大给声声压级直至受试者反馈听到声音；

在受试者做出表示听到测试音的反馈后，将给声声压级增加5dB后开始测试，并以5dB一档逐渐减小给声声压级，当达到下降转换点（定义如上）时，再将给声声压级减小5dB并开始以5dB为一档增加给声声压级，直至上升转换点，再将给声声压级增加5dB并开始以5dB为一档减小给声声压级，如此反复三次（即降三次、升三次）；

三次完成，所处声压级即为听阈值。保存测试信息，若还有未测试频点则切换频点从1)开始测试新频点；若已是最后的1KHz频点上的测试，则判断最初的1KHz的测试结果与本次测试结果是否相差过大，若相差过大则本次所有频点的测试结果均无效，需重新测试，否则测试完毕。

升降法测试软件操作流程如图3-3所示：

图3-3升降法测听阈流程图

频率分辨力测试

纯音听阈测试可以反映出受试者在特定频点上的听力级别，但是它没能反映出受试者频率分辨特性，即受试者在此频点出的灵敏度。一般认为，对频率的分辨能力在人对语言的理解至关重要。

心理物理调谐曲线法

理论上而言，{83%:频率选择特性的评价方法主要采用心理物理调谐曲线（psychophysical tuning curves, PTCs）法[34]。}心理物理调谐曲线描述窄带掩蔽信号的中心频率和强度之间的关系。针对某个给定信号，利用窄带噪声信号对其进行掩蔽，为掩蔽原信号所需的窄带噪声信号中心频率和强度之间的对应关系即为一条PTC曲线。其具体测量方法描述如下：固定待测频点信号的声压级和频率，改变用于掩蔽的窄带噪声的中心频率和声压级，当受试者反馈恰好听不到该频点上的纯音信号时，则该窄带信号的中心频率点所需的掩蔽强度为此时窄带信号所对应的强度。通过在待测频点的左右各选若干中心频率点进行测试，可得到该频点所给纯音信号所对应的掩蔽信号的频率和强度对应关系，将这种对应关系表现在频率-强度图上即是该频点所对应强度下的PTCs。

在PTCs的测量过程中，测量软件以何种方式给声，以及受试者以何种方式反馈并没有一个统一的标准。有的学者根据传统的心理物理调谐曲线检测法，在纯音刺激声频率的上下倍频程上选若干频点作为掩蔽噪声中心频率，并调节掩蔽声的声强进行掩蔽测试；有的学者通过在刺激声频率上下通过正反向扫频的方式改变掩蔽噪声的中心频率并利用软件自动调节掩蔽噪声的强度，受试者只需反馈是否听到刺激纯音信号即可[35]。随着需要测试的中心频点数增加时，PTCs测量的繁琐情况则无法忽略，一次完整的PTCs测量长超过2h~3h，十分耗时，且长时间的测试会对受试者的生理造成一定伤害，也使得受试者的反应发生暂时性的改变，影响测试结果，这也

使得PTCs在临床上面临一些困难。

改进的频率分辨力测试法

针对PTCs所面临的测试繁琐问题，结合移动端测试平台的特性，以及测听系统的需求，本文对频率分辨力测试的方法做出改变，提出一种改进的频率测试方法。该方法大大降低测试时长，作为代价，它只测试固定频点上的频率分辨阈，不能得到强度辨别阈和相位辨别阈，同时该方法以将频率分辨能力离散化，用固定的频率分辨百分比表征频率分辨能力。该方法可以在一定程度上反应受试者的频率分辨程度，也比较适合移动平台的软硬件实现，因此本文采用该方法作为听力测试系统中频率分辨力评价的方法。

{46%:本方法的测量范围为125Hz至8000Hz上的11个频率点: 125Hz, 250Hz, 500Hz, 750Hz, 1000Hz, 1500Hz, } {60%: 2000Hz, 3000Hz, 4000Hz, 6000Hz, 8000Hz; }同时，频率辨别阈 f 分为 $[1/2, 1/4, 1/8, 1/16, 1/32, 1/64, 1/128, 1/256, 1/512]$ 九个档位。测试步骤如下：

选择本轮测试频点。从剩余待测频点中选择本轮测试频点 f_i ，并生成相应的纯音信号（声压级可由受试者调节固定）；

生成给声信号。当前频偏系数 f_c 初始化为 $1/2$ ，对比信号频率 f_c 的取值空间= $\{f_i \times (1-f_c), [f]_i, f_i \times (1+f_c)\}$ ，{81%:三者出现的概率分别为 $1/4, 1/2, 1/4$ 。}给声信号由该频点纯音和增加偏频之后的信号拼接而成，即“纯音信号——停顿——偏频信号”格式，三段信号的时常均为500ms；

受试者选择。受试者再听完所给声之后通过软件进行结果反馈，可选反馈项为，所听到的两段音段相同或不同。此阶段给声三次为一组；若受试者在第一组反馈均正确，则将 f_c 下调一档，继续从步骤二开始测试。若第一组的连续三次给声反馈中有错误，则再给一组测试，若全部正确则将 f_c 下调一档，继续从步骤二开始测试；若仍有误判，则认为当前 f_c 受试者无法分辨，并将 f_c 的上一档作为该频点的频率辨别阈，若还有为测完频点，步骤一开始继续测试；否则测试结束。

频率分辨力测试软件设计流程如图3-4所示：

图3-4 频率分辨力测试流程图

言语测听

{40%:言语测听在助听器验配和听力检查中具有明显的指导意义，为衡量受试者的言语识别能力提供重要依据。}言语识别能力可由言语识别率和言语识别阈两个角度表征，故在本软件系统中言语测听分为言语识别率测试和言语识别阈测试两部分。

言语测听所用语料为由专业播音员录制的双音节扬扬格词，测听方法参考言语测听国家标准GB/T 17696-1999[25]。

言语识别率测试

言语识别率是指在某一固定声压级下，受试者正确识别出的词数所占总测试词库的百分比[25]。测试时，由系统播放测试语料，并通过选择拼音和字的方式反馈，软件统计受试者的反馈结果，并在测试结束时给出测试结果。词库中的语料会以随机的顺序遍历完成，使每个语料均有测试的机会，并不会出现重复的现象。言语识别率测试软件流程如图3-5所示：

图3-5言语识别率测试流程图

言语识别阈测试

言语识别阈（言语接受阈）定义为言语识别率为50%时的给声声压级（dB SPL），故其值可通过言语识别率的测试方式先测出P-I（识

别率-强度)曲线,再从P-I曲线中50%识别率所对应的声压级读出[25]。

除了根据言语识别阈的定义方式进行测量之外,文献[21]参考美国言语听力学会(ASHA)提出的言语识别阈测试指南,给出言语识别阈的另一种测试方法。该方法首先测定受试者完全听懂5个扬扬格词的声压级,并将其作为初始给声言语级,在此声压级的基础上以5dB为一档构建阶梯下降词表,软件记录受试者的反馈信息,当在某一声压级上5个测试扬扬格词均未被正确识别则可终止测试,并由如下公式计算言语识别阈:

$$\text{言语识别阈} = \text{初始给声声压级} - \text{测试中正确应答的数量} + 2.5\text{dB (校正因子)} \quad (3-1)$$

参考上述测量方法,言语识别阈测试软件设计流程如图3-6所示:

图3-6言语识别阈测试流程图

声强标定

测听系统软件正常工作的前提的输出音频声压级与显示值相对应。{55%:测听仪器的校准中常使用声压级(Sound Pressure Level, dB SPL)作为声音强度描述单位,} {59%:然而在Android系统中使用整型数控制输出音频的音量。}因此,使用测听软件进行听力检查之前需要对系统输出音频声压级进行标定和校准。

测听软件需要以声压级为单位控制输出音频的大小,因此需要对移动设备输出的音频音量和声压级进行一一对应。但是整型数表示的音量数值范围过大,很难全部完成标定;而测听系统设计的输出音频声压级范围是0~100dB(考虑到移动设备的输出能力有限),故可以通过选取声压级范围中关键点的方式进行标定,非关键点的声压级对应的音量可通过线性差值的方法给出。

图3-7Android系统音量与耳机播放声信号声压级定性关系

受试者接受声信号的耳机端所播放音频的声压级与Android系统的音量对应关系如图3-7所示。软件系统中为标定0~100dB范围内的声压级-音量对应关系每隔5dB标定一个点,其他声压级点所对应的音量数据由线性差值得到。

3.4本章小结

本章主要介绍基于移动终端的听力测试系统整理架构设计。本章首先分析系统基本需求,对听力测试系统的功能指标及相应听力测试参数进行详细描述;随后,对系统的整体架构进行阐述,听力测试系统主要分为硬件系统和软件系统,其中硬件系统即是基于移动终端选取,软件系统主要包括各听力测试项以及相关信号提供给和产生源之间的联系。最后针对每种听力测试具体项进行详细分析和设计,听力测试系统主要涵盖纯音测听、频率分辨力测试以及言语测听相关内容。本文第六章将详细描述个软件测试项的软件界面及操作流程等细节。

响度补偿算法

现代助听器的核心功能是针对佩戴者的听力损失情况自动调节输出语音的响度、频率等参数,使患者能够在听觉上更好的感觉和理解外界发出的声音[36]。助听器完结合人耳特性、语音信号处理基本原理以及听障患者的听力损失情况对输入声信号进行恰当处理并输出,最终提升佩戴者对原声信号接收能力的过程称为响度补偿[37]。因此,响度补偿技术是数字助听器的核心算法,也是影响助听效果的关键因素。

一般而言,听力障碍患者的听阈较正常人有不同程度的提升,且其痛阈值也有相应的下降,即可听动态范围下降明显[38];另外针对伴有频率分辨损失的听障患者而言,其语音可懂度的提升需要对输出声音的特定频率进行补偿或伸缩;此外,外界声音中常夹杂着噪声,需要有效的抑制噪声以便提高输出语音的可懂度。传统助听器响度补偿算法最基本的处理方式是针对听障患者听觉动态范围的下

降进行宽动态范围压缩（WDRC），将声音的响度调节至听障患者剩余的听力范围内[5]。随着助听器相关技术和信号处理技术的不断发展，针对频率分辨力损失的相关补偿方法也被应用于助听器补偿算法中，以提升补偿效果。

本章首先分析人耳的听觉相关特性以及响度补偿的基本原理和算法，并针对提高语音的可懂度，对语音信号的共振峰进行相应补偿，结合谱对比增强处理提高响度补偿算法性能。

语音信号基本特性和数字模型

语音信号特性

短时平稳性

语音信号在一段较长的发音时间内是非平稳信号，许多针对平稳信号的数字信号处理方法无法使用，然而，语音信号的变化速度是比较缓慢的，研究表明，在10ms~30ms时间段内，语音信号可被视为平稳信号进行处理，即语音信号具有短时平稳的特性[39]。**{61%:短时平稳性是语音信号处理的基础，几乎所有传统语音信号处理方法均基于该特性，事实证明，语音信号也确实是具有短时平稳特性的。}**

清音和浊音

语音信号一般可分为清音、浊音和爆破音三个部分，其中爆破音所占能量小且无明显规律，一般语音信号处理中均将其忽略[40]。

{40%:清音是发音时声带不振动情况下产生的语音，相比于浊音，能量较小，不具备周期性且频率成分分布较广。}针对清音部分所提取的语音特征有助于说话人识别，因为不同说话人，特别是性别不同或年龄差别较大的情况下语音中清音成分往往相差很大。

浊音是发音时伴随声带振动的音，占据整个语音六成以上的能量，具有准周期特性，其周期被称为基音周期，该特征广泛的应用于语音相关模式识别中。**{65%:浊音能量主要集中在较低频部分，且其频谱有明显共振峰现象。}**由此可见，浊音部分中蕴含语音中许多关键特征，因此浊音在语音信号处理中至关重要。

语音发音数字模型

人发声过程由呼吸器官、声带和喉头以及口鼻腔等构成的声道和声腔共同作用完成。为了更加深入了解发声机理，学者根据发声器官作用原理对语音信号产生过程建模，得到语音信号发音数字模型。**{54%:从原理上可将发音模型分为三个部分：激励模型、声道模型和辐射模型[41]，分别对应于上述发声器官，其原理框图如下图所示：}**

{77%:图4-1语音信号产生数字模型}

{56%:由图4-1可知，语音信号产生的数字模型可分为激励源、声道模型和辐射模型三个部分。}激励模型部分主要模拟清音或浊音，声道部分模拟语音从声带至口鼻腔过程中的共振效应，辐射模型模拟口鼻腔的辐射效果。**{48%:可见，该数字模型以语音信号产生的生理学结构为依据构建。}**

激励模型

激励模型以语音信号中清音浊音产生机理为依据，通过生成周期性脉冲经由声门脉冲模型加权模拟浊音部分；利用白噪声序列模拟轻音部分。

周期性脉冲序列生成器时域可用冲击函数累加方式表述：

(4-1)

其中A表示脉冲加权的幅值。其Z变换频域表示如下：

(4-2)

声门脉冲模型通常使用周期性斜三角脉冲来表示，其时域表示为

(4-3)

其中N₁和N₂分别是单个三角脉冲的上升和下降时间。将上式进行Z变换得其系统函数G(z)如下：

(4-4)

其中g₁和g₂是常数，G(z)是一个二阶全极点模型。{50%:因此，模拟浊音激励部分的数字模型系统函数可以表示为：}

(4-5)

声道模型

声道模型模拟语音从声带处传出后经由声道腔并在其中共振的现象。声道对激励语音的加权和共振作用形成最终语音中的元音和辅音部分，并在最终语音的语谱图上表现出共振峰特性。所谓共振峰，从语谱图上看表现为在某些频率点上其能量相对较高，颜色较深；从发声原理上可解释为声带振动产生的声音在喉口鼻腔内传输时在不同频点出产生共振，使得最终语音某些频率点附近能量较高。

共振峰频率具体值因人而异，与共振腔体器官构造关系密切，因此共振峰也能在一定精度范围内作为声纹特征使用，同时这些频点附近的频率蕴含着语音信号的大部分能量，对语音语义的贡献较大。共振峰一般取前三至五个使用。

声道产生元音的原理可以看成是一组串联的谐振器，一个谐振器对应于一个共振峰频率，故元音声道模型可以用一组串联全极点IIR滤波器描述，其传输函数可如下表示：

(4-6)

{77%:辅音部分由并联型零极点IIR滤波器表示，其传输函数如下：}

(4-7)

辐射模型

辐射模型模拟口鼻对所发出的语音信号向外辐射的过程，唇齿结构特征对辐射结果有较大影响。辐射部分用如下所示的一阶差分模型描述：

(4-8)

响度补偿基本原理和常见算法

响度补偿基本原理

基于人听觉感知器官的特性，针对同一声压级也即单位面积上能量相同，但是频率不同的声波信号，人最终的主观感觉声强却大不相同。为了合理描述人对声强的感觉特性，人们利用响度表述声音在主观感受上的强弱。同一声压级的声音其响度值大小与声音频率有密切关系，不同频率，人耳所感受到的响度大不相同。国际标准化组织（ISO）通过大量实验研究，总结出在人耳可听频率范围内，声音的声压级单位和响度单位的对应关系——等响曲线[42]，如图4-2所示。响度级的单位为方（phon），从图中可以看出，响度级和声压级以频率为1KHz的声音所产生的感觉级为参考，测量不同频率不同声压级声音所能产生的响度级。可见，在低频处，若要产生与1KHz相对应的感觉级声音，则需较大声压级，即人耳对此低频声信号并不敏感；当频率大于500Hz之后，感觉级响度和声压级对应关系相差波动较小，人耳对3000KHz左右的声音感觉最为敏感。

图4-2 ISO-等响曲线

助听器的核心功能是：根据听障患者听力缺失情况，对外界声音进行合理的补偿，使佩戴者的对输出声音的感觉响度和实际声音的响度大致相同。**{56%:因此，完成此功能的响度补偿算法是助听器系统的核心部分。}**如前所述，听力损失患者在听阈值提高的同时会伴随着痛阈值的下降，其可听范围大大减小。响度补偿算法的首要目的便是将输入声音强度合理的映射为患者可听范围内，让患者首先在响度感觉上听到此声音，因此在响度补偿中动态范围压缩算法必不可少。

单通道的响度补偿在增益计算时，将全频带统一处理，根据各频点的增益插值得到全频域的增益，这样的处理在灵活性和输出语音的舒适度上均有限制。针对单通道的局限性，研究者们将频率划分成多个子带，利用多通道滤波器组完成频带划分，并在每个子带内单独进行补偿处理通过多通道划分之后补偿结果的舒适度、可听性和清晰度等都有显著提高[43]。因此，在目前基于多通道的响度补偿方法广泛应用于各类助听器。

随着响度补偿算法的不断发展，研究者们开始关注听损者频率分辨力上的损失情况。有的患者对高频信号的动态范围已所剩无几甚至基本丧失，那么针对输入信号处在高频段的信息需要进行变换处理，将其搬移至可听频率范围之内，这样虽然会使得语音有较大变化，但是对于听损患者而言能听到语音并理解是基本需求。因此，根据听损患者的听力图和频率分辨损失情况针对特定频率段的信号进行搬移，即频谱搬移算法，显得十分必要。

宽动态范围压缩（WDRC）

WDRC技术是响度补偿算法的核心部分，宽动态范围压缩算法基于频域分析处理，是一种典型的单通道补偿方法，其具体操作步骤如下[44]：

将信号变换至频域；

{62%:根据听力图和信号特征频率点的声压级计算相应的增益；}

利用线性插值方法得到全频域增益曲线；

增益补偿之后将信号变换至时域输出。

如图4-3所示，是某一频点对应的I/O曲线。其中nTHR、nMCL和nUCL分别对应于该频点上正常人的听阈、舒适阈和痛阈值；unTHR、unMCL和unUCL分别对应于改频点处听障患者的听阈、舒适阈和痛阈值。WDRC算法根据这些值构建对应关系，将输入信号的声压级对应成图中输出的SPL值，根据输入输出值的差异进行补偿。

图4-3动态范围压缩I/O曲线

由上I/O曲线可见，输出补偿增益与输入信号声压级呈分段线性的关系。其输出规律总结如下（输入信号声压级用 L_{in} 表示，输出信号声压级用 L_{out} 表示，增益用 G 表示）：

当 $L_{in} < L_{th}$ ，即当前外界输入信号声压级未达到正常人听阈值时，系统默认不使用任何补偿方式，即该I/O曲线输出增益 $G = 0$ ；

当 $L_{th} \leq L_{in} < L_{p}$ 时，输出信号声压级 L_{out} ，即如图4-3中（ L_{th} ， L_{p} ）所对应直线段，此时算法根据I/O曲线中的对应关系，对输入信号进行增益补偿输出，输出增益 G （单位为dB SPL），一般而言此段增益大于零；

当 $L_{in} \geq L_{p}$ 时，输出信号声压级 L_{out} ，如图中（ L_{p} ， L_{p} ）所示，可见I/O曲线将声压级处在正常人的舒适阈至痛阈范围内的声信号，映射至听障患者舒适阈至痛阈范围，此段变换效果中，既有对输入信号进行补偿的部分，也有对输入信号进行限幅衰减的部分，因为，一般而言听障患者会伴有痛阈值的下降，即在此段计算出的 G ，或大于零，或小于零；

当 $L_{in} \geq L_{p}$ 后，I/O曲线输出开始保持为听障患者的痛阈值 L_{p} ，即 $L_{out} = L_{p}$ ，避免过大声信号对助听器佩戴者造成伤害，起到一定的限幅效果，此时增益 $G < 0$ 。

上述增益曲线是针对某一特征频点所构建而得，通过计算所有特征频点的I/O曲线，得到对应频点、声压级所对应的增益值，通过插值的方式获得全频域的增益因子，由此对输入信号进行补偿输出，这边是WDRC的基本思路。

WDRC算法合理利用听障患者剩余的听力范围，使其可以得到在声压级上更为合理舒适的声信号，然而实际使用发现该算法处理后的语音舒适度非常低，达不到使患者满意的效果，究其原因，WDRC算法将全频域作为一个整体通道进行处理，对实际语音的频率成分分析不够，使得输出信号可听性过低。后述多通道响度补偿算法正是针对WDRC算法的痛点进行处理，获得不错的补偿效果。因此，WDRC算法一般而言与多通道响度补偿算法结合使用达到较好的响度补偿效果。

多通道响度补偿算法

针对上述基于单通道的WDRC算法的不足，研究者们通过细分频带范围思路对其进行改进，引入频带分解的思想，将全部可听频带划分为多个通道，分别进行响度补偿。每个子带内使用的补偿算法可以是简单的WDRC处理，也可以根据患者的先验信息进行其他更有针对性的响度补偿策略。多通道补偿思路使得算法的灵活性大大增加，在提升算法性能的同时也可以针对患者的实际听损情况，比如频率分辨缺失情况，实现更人性化的补偿方案，较好的改善输出语音信号的舒适度，提供佩戴者使用体验。

{47%:多通道响度补偿算法首先利用滤波器组将输入信号划分为多个子带，而后对针对子带进行响度补偿，最后将各子带信号合并输出[43]，算法主要流程如下图所示：}

图4-4多通道响度补偿流程

由图4-4可见，多通道处理算法首先需要一组用于分解信号和综合处理后的子带信号的滤波器组；另外一个核心便是每个子带进行子带补偿的算法。因此，针对多通道响度补偿的研究大都针对于频带划分方式以及如何在各子带内进行补偿处理。滤波器参数可提前设定，如果子带补偿中不涉及频域处理，那么多通道响度补偿算法主要运行在时域，算法计算量小，非常适合移动便携设备上运行。

诚然，多通道处理思路的补偿算法更加灵活有效，但也带来了新的问题。研究表明，多通道处理方式的效果与频带划分方式息息相关。各子带划分会使得相邻两子带边界频点最终获得增益不一致。如果分界频率恰好处在某一共振峰附近，那么该算法极有可能会使得语音共振峰被分割。共振峰的损坏将使得输出语音可懂度极度下降。因此，实际使用的多通道补偿算法需要改进。

基于共振峰的谱对比增强响度补偿算法

在语音信号处理当中，共振峰特征应用广泛，其蕴含着原始语音的语义信息。从语音信号数字模型角度出发，共振峰即是描述声道腔

共振特性的基本参数。说话人语义信息对提高耳聋患者的言语识别率有关键作用，因此语音信号中共振峰特性对提高语音补偿效果具有重要意义。传统多通道响度补偿、宽动态范围压缩等均未考虑语音共振峰特性，容易产生劈峰，虚假共振峰等情况，对共振峰造成破坏，导致语音可懂度、清晰度严重下降[46]。通过共振峰检测等方式对共振峰频率处信号进行保护，避免破坏语音信号的清晰度和可懂度[46]，使得补偿后语音可懂度等有所提升。但是在噪声情况下，仅仅保留共振峰信号的完整性不足以提高整体含噪语音的清晰度或可懂程度。

本文基于共振峰优化思想以及其重要意义，以及以上算法对含噪情况表现的不足，将谱对比增强引入响度补偿之中，提出一种基于共振峰谱增强方式的响度补偿处理算法。谱对比增强可以提高含噪语音信号的清晰度[]。一般的谱增强处理可能会产生虚假共振峰。通过在谱增强处理中以共振峰所在频率为中心增强其频带内信号，同时适当抑制共振峰带外信号，提高语谱对比度，使得语音信号清晰度得到提升。同时，处理也完整的保护并加强共振峰信息，使得处理后语音可懂度有所提升。

共振峰

{52%:语音信号产生数字模型的第二部分——声道模型描述语音发声腔体特性。}由喉、鼻、口腔等构成的声腔对从声带等激励部分发出的声音会在特定频率处产生共鸣效果，即发生共振[47]。共振发生的频率因人而异，同时同一人的声腔会在多个频率处产生共振效应，而人主动改变声腔形状也会使得共振的频率产生变化。将共振频率从小到大排列分别称为第一共振峰、第二共振峰...一般取前3至5个共振峰作为特征参数。从共振峰产生过程可以看出，一帧语音信号共振峰频率与该帧语音信号产生时共振腔（声道腔）特性密不可分，对语音产生和语义生成有重要意义。因此，共振峰合成技术常用于语音合成。

共振峰提取主要有谱包络检测法、倒谱法和线性预测（LPC）法等，其中谱包络法运算量最小，效果最差；{56%:而倒谱法运算量较大，故本文采用LPC法提取共振峰参数。}

{41%:线性预测基本思路是利用当前样点的前 个样点预测当前样点，即：}

$$(4-10)$$

式中 表示信号样点， 为线性预测系数。则预测误差可如下表示：

$$(4-11)$$

求解 ，即求解使得 最小的一组线性预测系数 即可得到相应AR模型。根据模型系数可得其传递函数：

$$(4-12)$$

其中 是预测模型阶数， 是增益常数。 又可以表示为 个极点级联的形式：

$$(4-13)$$

{57%:式中 ， 极点对应的共振峰频率即为 ，其中 为采样周期。}实际提取中常使用峰值检测的方法获取共振峰频率，LPC共振峰检测处理框图如下所示：

图4-6 LPC共振峰检测原理

图4-7为两例语音帧LPC预测分析结果对比：

图4-7 LPC预测结果对比

LPC谱包络中可以明显地看出共振峰信息，同时可以发现，由于增益G的存在，LPC谱和信号频谱间在元音和辅音间相差不同，但可以证明该差值均为常数。

通过获取对应帧语音信号共振峰信息，可指导该帧后续处理中针对共振峰频率处进行特别处理，降低处理带来的语音失真同时探寻提升语音清晰度可懂度的处理方式。

谱对比增强

在日常生活中，语音聆听过程常伴随许多背景噪声，使得听清目标的难度增大，对听障患者而言，对背景复杂的聆听环境显得更加力不从心。助听器普遍采用的处理方式是，选取合适的降噪算法降低背景噪声的能量，提高信噪比。通过降噪处理之后，语音的清晰度有较好的提升。但是，降噪一般只针对如何将信号中的噪声部分剔除以提高信噪比，并未直接针对如何提高语音信号清晰度进行处理。谱对比增强的思想是在基本降噪处理的基础上，对信号频谱进行处理，提高频谱对比度。研究表明[48]，谱对比增强可以提高耳蜗植入患者噪声环境下语音的可懂度。一般而言，听损患者需要更强的谱对比度以提高噪声环境下的言语识别率。

谱对比增强处理保持频谱波形的峰值并拉低频谱波形的谷值，使波峰与波谷对比更加明显，算法具体处理步骤如下：

对待处理帧进行FFT，并将其转换成对数谱；

寻找对数谱 的峰值序列 和谷值序列；

对于频谱上任意一点，确定其所在线段的峰值 和谷值，计算峰谷值差；

确立增强处理后新的峰谷值差，其中 是增强算法程度控制因子，取值一般为0到1之间，取0即代表无增强效应；

增强后谱序列各点值由如下公式算得：

$$(4-14)$$

谱对比增强算法流程图如下：

图4-8谱对比增强处理流程

处理算法中关键部分是利用峰谷值和调节参数对原频谱进行变换，处理细节如上步骤所述。下图为一帧语音增强前后频谱对比：

图4-9谱对比增强前后频谱对比

经过对比增强处理后频谱波谷更加深，频谱峰值保持不变，相邻峰谷值之间的值按比例对应于新的峰谷值间。处理后，频谱动态范围扩大，对微弱噪声有一定的抑制效果，而尽可能的原始语音的增大动态范围对听力动态范围所剩不多的听障患者而言是有好处的。因为，动态范围广的信号在经过WDRC处理后的剩余动态范围相对更宽泛，表现为语音在响度上更加容易分辨，因此，谱对比增强处理对提高语音清晰度和可懂度有一定的帮助。

处方公式及I/O曲线

处方公式帮助听力师选配助听器，常被用于计算补偿增益。它依据患者各特征频点处听阈参数，计算所需增益（被称为目标放大量）。这些增益参数常用于指导助听器对输出语音进行相应补偿得到输出语音，因此，处方公式对助听器计算患者对应I/O曲线有直接

影响，是助听器补偿功能的核心单元。

早在1935年，Knudsen和Jones所使用的镜像听力图法处理听损和增益间关系便是一种处方公式的应用，他们简单的将1dB听损对应需要1dB的增益以达到补偿效果[49]；1940年Watson和Knudsen开始基于患者舒适阈值(MCL)寻求补偿增益计算方法[50]；1944年Lybarger提出了基于MCL的“1/2增益原则”[51]；此后，大量针对不同特性需求的处方公式被提出，大部分处方公式也均已写入各商用助听器软件中。根据处方公式增益计算方法不同，可分为线性处方公式和非线性处方公式。

线性处方公式给出增益值基本为固定计算值，并不考虑输入声压级大小影响。线性处方公式有：1/2增益、Libby、Lybarger、Skinner、Pogo II、Berger、NAL-R和理想感觉级(DSL)等，常用的线性处方公式简单说明如下：

1/2增益公式：核心思想是助听器输出语音达到最适阈所需增益等于纯音听阈值的1/2，同时增加有10~15dB的保留增益，适于佩戴过助听器患者使用；

Libby公式：将1/2增益改为1/3，即取听阈值1/3作为增益，并将250Hz和500Hz频率处增益分别减去5dB和3dB，适于首次佩戴者；

Pogo II公式：原始Pogo公式在1/2增益公式的基础上对低频分量添加消减因子，以提高言语理解能力；Pogo II公式在其基础上针对严重听损情况(>65dB)，增益中增加高于65dB部分的一半作为附加增益；

NAL-R(National Acoustic Laboratories Revised)公式[52]：NAL-R于1986年由Byrne & Dillon提出，它是基于NAL公式的修正版本，针对言语频率，尤其是500Hz ~ 1000Hz频带内的听损，其所需提供的能量更加优化精细，并使用1/3增益原则，其具体计算步骤如下：

(4-15)

其中 分别为500Hz，1000Hz和2000Hz处的听阈值， 为修正因子值如下：

表4- 1修正因子值

{100%:频率 (Hz) 250500100015002000300040006000}

(dB)

-17-811-1-2-2-2

从其计算过程可以发现，针对重度听损患者，NAL-R并不能给出较大增益，因此，Byrne在其基础上又增加两个修正因子以达到提高低频声音能量并尽量抑制啸叫，修正后的公式称为NAL-RP。

{54%:理想感觉级 (the desired sensation level, DSL) 公式[53]：该公式针对感音神经性听障患者，}发现其感觉最佳舒适阈值的言语声必须放大到一定的感觉级 (sensation level, SL)。DSL公式目的是将言语声转换至患者动态范围内。DSL增益方式对幼儿尤其合适。

线性处方公式完成对输入语音信号放大的功能，使得患者可以听到未补偿时听不到的声音，但由于其并未考虑输入信号本身声压级大小，容易造成过补偿情况。非线性处方公式针对线性处方公式所存在的不足对补偿I/O曲线更加细化处理，综合考虑输入信号声压级和患者听力图信息，给出更加合理的增益补偿方案。使用非线性处方公式，选配人员可以更加细致的调节助听器补偿参数，验配效果也更加完美，常用的非线性处方公式有：FIG6处方公式、DSL[i/o]公式、NAL-NL 1、LGOB公式以及IHAFF公式等。

FIG6处方公式[54]: FIG6公式命名源于发表该公式的论文的图表6, FIG6并不基于个体响度测试, 它根据大量同等听力损失级人群的平均响度信息来计算增益, 这也表明该公式只需要患者听阈信息即可。其具体计算操作如下表:

表4- 2 FIG6增益计算

低强度声音增益: (1) $0 \sim 20\text{dB HL}$: $G = 0$

(2) $20 \sim 60\text{dB HL}$: $G = \text{TH} - 20$

(3) $\text{TH} \geq 60\text{dB HL}$: $G = \text{TH} - 20 - 0.5 (\text{TH} - 60)$

舒适阈声强增益: (1) $0 \sim 20\text{dB HL}$: $G = 0$

(2) $20 \sim 60\text{dB HL}$: $G = 0.6 (\text{TH} - 20)$

(3) $\text{TH} \geq 60\text{dB HL}$: $G = 0.8 (\text{TH} - 23)$

高强度声音增益: (1) $0 \sim 40\text{dB HL}$: $G = 0$

(2) $\text{TH} \geq 60\text{dB HL}$: $G =$

其中TH表示相应听阈值, G为算得的增益。

DSL[i/o]公式[53]: DSL公式分为DSL[i/o]线性公式和DSL[i/o]曲线公式两种, DSL[i/o]线性公式将输入声音信号的声强范围划分为三个区域:

当输入声强度小于压缩阈值, 即 , 此时输出声强 ;

当输入声强介于压缩上限和下限之间, 即 时, 输出信号为线性放大效果, 即 ;

当输入声强高于压缩上限, 即 时, 。

可见, DSL[i/o]线性公式理论基础和宽动态范围压缩一致。

DSL[i/o]曲线公式常用于非线性助听器选配过程, 其最终目的是使输出信号的响度更加合理化。相比于线性版DSL公式, 其针对压缩范围内的输出增益计算有所不同, 该区间I/O线形态取决于正常人响度增长的比率和听障患者响度增长的比率。

NAL-NL 1公式[52]: 该公式为NAL公式的非线性版, 其基本理论是在言语总体响度不超过正常人感觉到的水平的条件下, 使得输出语音可懂度尽可能高。NAL-NL 1公式以言语可懂度最大化为准则, 并提出响度均衡处理方案, 同时针对听损严重频率增益相对较小, 而听力残余情况较好的区域则提供较大增益。

{88%: IHAFF公式[55]: IHAFF是独立助听器选配论坛 (the independent hearing aid fitting forum) 的简称, }其技术专家组以开发适用于各类可调WDRC助听器程序为目的, 其指导思路总结为: 放大后的声音不改变其原有性质, 如柔和声仍旧是柔和声; 正常言语声补偿后处于患者舒适阈附近; 高强度语音在不超过痛阈的条件下, 补偿后仍为较响语音。

综合以上所述各类处方公式, NAL-NL 1和DSL[i/o]应用最为广泛, 分别针对于成人助听器和儿童助听器程序。各常使用的处方公式也

均已集成与助听器厂家软件中，因此，相比于处方公式的研发，更需要研究人员注意的是各处方公式的特性以及针对特定情况下处方公式的选择。

本文算法采用FIG6处方公式计算特征输入强度点所对应增益值和对输出声压级，并通过线性差值以及限幅输出原则计算全输入强度区间内输出声压级。

选取声压级为40dB SPL（柔和）、65dB SPL（舒适）和95dB SPL（过响）三种特征点声压级作为增益拐点，由FIG6处方公式给出此三处增益值， G_1 ， G_2 ，并算得对应输出声压级 O_1 ， O_2 。由三处特征声压级对应输出声压级为参考进行线性插值，算得全声级段内输出声压级。也即：

当时， $X \leq 40$ ，其中；

当时， $40 < X \leq 65$ ，其中；

当时， $65 < X \leq 95$ ，其中；

当时，

以上各式中 O 代表输出声压级， X 代表输入声压级，实际上 G 是各分段线性插值比例。图4-10为某听力图某频率点处所对应I/O曲线。

图4-10某频点I/O曲线

通过对听力图线性插值，可以得到频段内任一频点所对应I/O曲线。结合本章其他处理过程，算法提取语音信号FFT所对应频点在125Hz（听力图最小频点）以上各频点处I/O曲线，用于对输入信号响度补偿。

谱增强响度补偿算法

{47%:本节在多通道响度补偿算法的框架下，将语音共振峰信息和谱增强处理方法用于补偿算法之中，并对多通道响度补偿算法进行改进。}首先为了避免多通道滤波器设计所带来的大量计算开销以及滤波器不理想造成的频带混叠，本方法利用傅里叶变换直接将信号对应至频域；其次为了充分利用患者剩余听力范围并尽可能提高输出语音动态范围以提高语音清晰度和可懂度，本方法将谱对比增强用于频谱处理，优化补偿效果，并考虑到共振峰的特殊性质，本方法在谱对比增强处理中适当加强共振峰所在频率处信号，突出共振峰频率处信号，保护语音信号中的原始声道特性。最终，对经过谱增强处理的信号按照患者听力图进行响度补偿，并输出处理后的语音。

算法整体结构流程图如下：

图4-11共振峰谱增强算法流程

图4-11即是针对某语音片段中某帧信号在本文所述基于共振峰增强和频谱对比增强的响度补偿算法中所经历的流程。首先利用4.3.1节描述的LPC谱分析方法提取共振峰参数，供谱增强处理使用；并由输入听力图信息和所选处方公式生成输入输出增益曲线，供响度均衡处使用；而后，算法利用傅里叶变换将语音帧变换至频域，通过SCE算法以及共振峰加强处理，得到增强后的频谱信息，并在频域进行响度均衡处理；最后，利用傅里叶逆变换求得处理后输出语音帧。

算法具体步骤如下：

对第 帧语音信号，求其共振峰信息 和 ；

对该帧语音信号进行傅里叶变换，得到其频谱序列 ；

结合（1）中获取的共振峰信息，利用谱对比增强处理算法优化频谱序列，增加其动态范围，并适当增强共振峰所在处能量，得到增强后的频谱序列 ；

计算当前帧声压级，并根据听力图和处方公式，通过线性差值方式的到频谱序列个频点所在处补偿增益，并对频谱 进行补偿，得到补偿后频谱序列 ；

求 的逆傅里叶变换得到补偿后语音帧。

补偿算法结果仿真

I/O曲线

为了获得更加精细的频率与听阈间对应关系以便于后续针对频谱补偿，算法提取患者从125Hz至8KHz频率范围内的11个特征频点处的听阈，其测听结果如图（）所示。患者在125Hz、250Hz、500Hz、750Hz、1KHz、1.5KHz、2KHz、3KHz、4KHz、5KHz和8KHz处听损分别为30dB HL、40dB HL、50dB HL、60dB HL、60dB HL、70dB HL、70dB HL、75dB HL、80dB HL、80dB HL以及85dB HL。{48%:其平均听损在60~80dB HL范围，属于中重度耳聋。}

图4-12某患者听力图

图4-13为上述听力图所对应全频带内I/O曲线集合：

图4-13全频带I/O曲线集合

从4.3.3节处方公式和I/O曲线计算过程可以看出，本文三维I/O曲线被3个特征声压级处值分为四段。针对输入声压级为0~40dB SPL的声音，I/O曲线给出的补偿较大，因为此段为低声，听障患者想要听到此段声音需调整其为较大声压级；对于40dB SPL~65dB SPL声压级段，I/O曲线所给出增益相比低声段下降，因为此段声压级已基本接近听障患者可听范围，无需太强增益，既保证患者可听，同时使其感觉声音为正常舒适感，与正常人聆听此声压级声音的舒适阈感觉一致；针对65dB SPL~95dB SPL声压级段声音，I/O曲线所给出增益已非常小，输入声压级基本与输出声压级一致，因为此段声压级声音基本处在听障患者可听范围内，仅需较小增益；对于声压级大于95dB SPL的输入声，算法将取95dB SPL对应输出声压级作为其输出声压级，已达到限幅输出效果，就实际而言，听损患者听觉动态范围的缩小，一般伴随着痛阈的下降，合理限幅处理有助于保护助听器佩戴者剩余听力，同时提升患者对助听器的佩戴体验。

响度补偿实验

实验选取TIMIT语料库中dr1组mjsw0目录下“sa2.wav”语料进行测试，语音内容“Don't ask me to carry an oily rag like that.”。使用本章响度补偿处理方法对该语音进行补偿处理，结果如图4-14和图4-15所示：

图4-14补偿前后语音信号时域对比

图4-15补偿前后语音信号频域对比

算法处理详细过程及中间步骤结果图如图4-16和图4-17所示：

图4-16语音信号时域变化

图4-17语音信号频域变化

算法首先使用4.3.1节LPC法提取信号共振峰信息，选取语音帧前三个共振峰，并针对被选取的共振峰所在频率处信号进行增强处理。从图4-17可知，经过共振峰增强处理之后频谱共振峰明显突出，其频率所对应幅值变大，共振峰增强后语音波形幅值变大。提高共振峰分量一方面有助于保证语音在后续处理后仍有较好的清晰度，提高语音可懂度；另一方面，共振峰处频谱的增强，有助于加强谱对比增强效果。

之后算法使用SCE（谱对比增强）处理方法对频谱进行优化，SCE因子取1，其最终处理结果如图4-17中SCE语谱所示。从语谱图对比变化可见，经过谱对比增强处理之后语谱谱线更为清晰，强谱线中心处幅值仍较大，但两相邻强谱线之间频谱幅值变小，谱线之间区分度更为明显，究其原因SCE处理过程中对语谱中能量极大值处的频点值保持不变，并将介于相邻极大值和极小值之间的频谱被进行拉伸处理。从增加SCE处理之后的语音波形来看与共振峰增强后结果对比没有明显区别，可见SCE所提供的频谱微处理并不会大幅影响语音信号。然而，经过SCE处理之后的频谱却更适合后续的频域响度补偿算法。

在语音帧频谱经过共振峰增强和谱对比增强之后，算法利用经由用户输入的听损情况（听力图如图所示）以及处方公式所得I/O曲线对FFT各频点处频谱进行补偿，补偿结果如图4-17。从补偿后频谱可见，语音高频成分增强显著，这与I/O曲线补偿特点中高频多补偿相关；同时，原频谱中能量高的部分亦得到相应的补偿，其补偿量与所在频点差值听阈值和其本身能量大小相关；从语谱结果可见，原本高能的低中频部分补偿量小于高频补偿量，一方面是由于该部分能量原本较大无需大幅响度补偿便可达到佩戴者听损要求，另一方面由于I/O曲线限幅机制的存在，补偿处理中应不会出现过补偿现象，两方面原因均和补偿过程中I/O曲线增益计算机制密不可分；并且，得益于共振峰增强和SCE处理机制，补偿后频谱共振峰频率所对应幅值仍较大，其构成的共振峰谱线仍然清晰。同时，经过响度补偿后，语音波形变化也较大，其规律和频谱变化相对应。语音波形整体幅值得到提升，对应于频谱整体能量提升；波形中金属声加重，部分波形包络发生变化，对应于频谱中高频成分分量幅值提升明显，0.4s左右处波形变化最为明显，该段语音主要由高频成份构成，经过响度补偿后，从频谱看其频域幅值增加明显，从时域波形看其整体幅值均有较大提升，这符合感音性听损患者对语音高频成分多补偿的诉求。算法整体处理之后输出语音高频成分增强，金属声加重，对于听力正常人而言语音质量下降，然而，对于听损严重的患者而言，更强的高频成分有助于其理解语音中高频成分所含语义，意义显著。

4.5 本章小结

本章主要阐述助听方案最为核心亦是常被用到的响度补偿算法。首先，介绍语音信号的基本特征以及其数字模型，为后续基于语音特征的补偿方案以及共振峰提取奠定基础；随后，本章介绍了传统响度补偿的基本原理和常见算法；基于响度补偿算法的基本原理，本章针对提高补偿后语音可懂度提升以及助听算法的抗躁能力提升问题，提出基于共振峰增强和谱对比增强结合的响度补偿算法。算法对频谱进行整体处理，在共振峰增强的条件下将语音帧信号的频谱进行谱对比增强处理，使处理后语音频谱的清晰度和对比度有所提升，这种变化意味着语音信号的特征获得的补偿比其他噪声获得的补偿更多，语音信号的信噪比有良好的提升。最后，补偿实验结果亦体现算法针对语音信号信噪比提高的效果。

频谱伸缩助听算法

随着对助听器核心算法的研究不断深入，研究者们发现，针对部分听损特别严重的患者而言，佩戴常规的响度补偿助听器所能获取的言语识别率的提高甚微，甚至有一定的反作用[58]。出现此类问题的主要原因在于，听阈过大、听损严重的频率处其所对应的耳蜗内毛细胞功能几乎完全丧失。由于“偏位听力”效应的影响，听损患者可以通过基底膜上附近功能完好的频点部位感知到被高增益补偿过的声信号，然而这并没有对言语理解提供帮助，其原因有：

（1）传统的响度补偿助听策略，主要是针对语音信号幅值的补偿。当患者听阈值过大，对应频点所需的补偿增益也比较大，而这些大增益常会受到实际硬件条件的影响而无法输出相应的补偿信号；

(2) 许多严重听损患者其主要听力图呈陡降型, 高频区听损严重, 由于原因(1)的存在, 会导致这部分信号增益不足; 同时由于声学上扩散掩蔽特性的存在(低频声更容易对高频声产生掩蔽), 当低频增益足够, 而高频增益欠缺时, 高频部分信号很难被感知; 而语音中有近60%的关键信息存在于1KHz以上的频段, 约35%的重要言语信息分分布在2KHz以上的较高频段, 同时语音中大部分辅音所在频段亦是4KHz以上的高频段[72]。因此, 高频段语音的理解至关重要, 高频段信息可懂度的下降会使得患者言语分辨能力急剧下降。

{43%: (3) 一般感音神经性聋患者的高频听损严重、频率分辨力低、听觉动态范围小、重振现象明显, 使得其所能理解的语音的信噪比有更高的要求。}

针对传统的响度补偿策略所面临的这些问题, 现在助听器需要能够在幅值增益和降噪的基础上进一步提高存在耳蜗死区的严重听损患者的言语能力的办法, 移频助听技术应运而生[57]。移频处理概念早在20世纪中叶出现, 当时有对于用频率降低技术进行助听的研究, 但未能有相应的成熟助听工具产生[58]; 四十多年后, {51%: 澳大利亚国家听力实验室的研究人员提出“移频(frequency shifting, } frequency transposition) ”的助听技术以解决传统响度补偿助听方法在针对高频听损严重患者时所面临的问题[59]。

移频助听策略的主要思想是避开患者听力损失严重的区域, 把处在听力薄弱区域的信号通过频率搬移或压缩的方式转移至残余听力较好的其他频段内。一般而言, 严重听力损失发生在较高频段, 通过移频处理, 关键性的清辅音等高频段语音元素, 在原本听不到的情况下可以被听到甚至理解, 这有助于提高助听器佩戴者的言语识别率。从生理学上看, 人脑对声信号中频谱信息的感知并不依赖于其绝对频率而是各频率间的比率, 因此, 移频策略具有相应的生理学原理支持。

针对极重度听损的听力障碍患者而言, 移频处理是唯一可选、可行的助听方案, 全球有近一半的助听器生产商提供基于移频技术的助听器。由于移频助听技术所要解决的问题的特殊性, 其参数调节与具体的听障患者实际听损情况有密切联系, 移频参数设置严重影响最终佩戴者的获益程度, 因此, 关于移频助听器所合适的佩戴人群以及佩戴后能够获得的言语识别率的提高程度存在着争议。

本章首先介绍移频助听技术, 并在其基础上研究频谱伸缩算法及其算法参数选择策略, 最后对算法进行仿真实验。

移频助听算法

移频技术包含三个基本元素: 原始频段、目标频段和对应法则[60]。其中, 原始频段是指信号频谱中待搬移或待压缩的频段, 该频段下限频率被称为起始频率; 目标频段是指算法对原始频段处理后输出信号所在频段; 对应法则即是完成将原始频段映射至目标频段的函数。常见移频助听技术原理框图如下所示:

图5-1移频助听算法处理结构

输入语音帧信号经过FFT变换至频域后, 利用滤波器将原始频段从全频带中划分出来(若起原始频段即为全频带则无需滤波处理), 剩余部分保留, 不做处理合并至新频谱, 原始频段根据相应的对应法则转换至目标频段后与剩余频段合并成新频谱, 经由逆傅里叶变换后输出。算法核心部分是图5-1中移频处理部分, 根据该部分的处理细节的差异, 移频助听算法可以分为频谱压缩和频谱搬移两类。

频谱压缩

频谱压缩技术类似于动态范围压缩技术, 只是前者针对信号的频率范围, 而后者针对信号的幅值范围。频谱压缩时原始频段(f_{ol}, f_{oh})带宽一般比目标频段(f_{tl}, f_{th})的带宽要大, 即 $f_{oh} - f_{ol} > f_{th} - f_{tl}$, 且两个频段的起始频率一般为相同值, 即移频算法的起始频率。频谱压缩起始频段和目标频段的对应关系可用下图表示:

图5-2频谱压缩频率对应关系

根据原始频段的起始频率点的不同，频谱压缩算法又可以分为线性压缩和非线性压缩两类。线性压缩算法中压缩段起始频率为最低频（1Hz），而非线性压缩算法的起始频率会更高一些，这也意味着有部分低频区将保持不变。线性频率压缩可以保证压缩前后频谱间的比率关系，有助于听觉皮层对压缩后语音的感知，而指数型对应关系与耳蜗基底膜上频率分布的对数关系有关联，一般性函数的对应关系并没实际意义。为了尽可能提高压缩后的语音的可懂性，频谱压缩常采用线性压缩的方式。

设原始信号频谱为 X ，频谱压缩原始频段为 (f_{ol}, f_{oh}) ，目标频段为 (f_{tl}, f_{th}) 。原始频谱经由滤波器之后剩余频段部分为 X_r ，带压缩频段为 X_c 。则对压缩频段 X_c 处理后输出频谱

$$(5-1)$$

最终，输出频谱 $X' = X_c' + X_r$ ，一般而言，频谱压缩过程中有 $f_{ol} = f_{tl}$ ，故其叠加过程中不会存在频谱混叠的情况。

例如，对一段采样率 $f_s = 16000\text{Hz}$ 的语音进行线性频谱压缩，原始频段为，目标频段取为，线性压缩结果如图所示。

图5-3频谱压缩示例

原始语音信号频率分布在0-8000Hz的频段内，经由线性压缩处理之后信号频谱处在0-4000Hz以内。由于是线性压缩，频谱间的相对位置关系被保存下来。通过适当的选取频谱压缩的参数，可将不在可听范围内的语音变换至较低频的可听范围之内，对提升言语识别率具有一定的帮助，但是频谱的大幅压缩会导致变换后语音的一定失真，患者需要一段时间适应压缩后的语音；同时，频谱中许多时间段内其频谱分布范围本身就已经是在可听阈内，无需压缩处理，然而由于基本的移频压缩算法不存在参数动态自适应以及输入信息检测机制，这些本可以无需压缩的语音片段被强制压缩，造成一些没必要的语音失真。

频谱搬移

频谱搬移类算法的目的是将部分高频段频谱搬移至目标频段。与频谱压缩不同的是，原始频段所在频谱不会被压缩，只是转移或平移至目标区域。一般而言，由于频谱搬移算法中目标频段的最低频低于起始频率，故在搬移得到的频谱与剩余频谱叠加时会有一定的重叠，言语特征会发生改变。频谱搬移算法的输入频率与输出频率间对应关系如下图：

图5-4频谱搬移频率对应关系

频谱搬移过程中没有对原始频段进行压缩处理，所以被搬移部分频谱带宽没有改变，即 $f_{oh} - f_{ol} = f_{th} - f_{tl}$ ；且搬移的目的是将听损严重的高频段搬移至低频部分，所以对于频谱搬移而言一般存在 $f_{ol} > f_{tl}$ ，这也意味着搬移后频谱与未经任何处理的剩余部分存在重叠。常见频谱搬移类算法有线性频率转移（Linear Frequency Transposition, LFT）和频率平移（Frequency translation/spectral envelope warping）算法。

线性频率转移（LFT）技术将检测到的最重要的高频区频谱信息复制到1/2原始频率处，即降低一个倍频程。为了不过度影响低频信息的感知，原始频段不能太宽，只能选择部分信号进行搬移处理。同时，原始频段中重要信息的检测规则影响着目标频段的确定，因此，最强峰检测规则也至关重要。

频率平移技术原理与LFT技术类似，均是将部分高频信息移至低频。但是频率平移算法是一种自适应算法，它只在被处理的语音帧频谱中高频成分比重较大时才启动，依靠特征检测的方式监测输入频谱。检测到有比较明显的高频成分后，启动移频算法，将原始频段的高频特征复制到听力尚存的目标频段；若算法未检测到突出的高频信息，则搬移算法不会启动。

移频助听小结

移频助听算法的核心思想是利用特定的对应法则将原始频段中的频谱搬移至目标频段，其目的是使输出信号尽可能的避开听障患者的

严重听损区域，以此为患者争取更多的识别率和语义理解能力的提升。

传统移频助听算法主要针对高频听力损失严重的患者，通过对信号频谱进行压缩或者搬移实现。由于其目标人群听损发生在高频段，为了避免听损严重的区域算法需要将高频段信号移至中低频段，因此又被称为降频助听算法。频谱的变化给语音信号带来失真，佩戴者需要适应由信号频谱变化所带来的改变。作为重度听损患者所特需的，亦是为数不多的针对重度听损患者的助听算法，移频助听技术的研究具有重要意义。但是，目前而言，移频助听算法都只针对高频听损严重的情况，对于中低频的严重听损患者并没有给出类似的助听方案。

频率伸缩助听算法

在传统移频助听算法的基础上，本节针对其在特殊听损情况下的不足（类似V型严重听损）提出一种频率伸缩助听算法，使基于频谱的助听方案更全面。

5.2.1 频率伸缩策略

传统移频助听算法的核心思想是通过语音信号的频谱处理将信号频率成分尽可能的避开听损严重的高频段。然而，此类处理方式并不能使听力图呈岛型、V型、翻转型等的严重听算患者受益，因为他们的听损严重频段并不是仅仅存在于高频段，无法仅通过降频或是移频处理而使处理后的信号频谱避开听损严重的中频段甚至是低频段。

图5-5典型噪声性聋患者听力图

图是典型V型噪声性聋严重听损患者的听力图。该类患者的高频段听力情况相对较好，而在2000Hz至4000Hz的中高频段听损情况最为严重。如果采用传统的频谱压缩或是频谱搬移类算法，则信号压缩后频谱的整体动态范围将会大大缩小，过大的压缩比率将会导致信号可听度急剧下降，使用者很难再通过训练而熟悉语音信号的这类改变。

鉴于类似的非高频段听损严重情况，本文提出一种针对于某个频段的补偿策略：通过拉伸语音信号中听损最为严重的频段以提高患者对于该频段信号的感知，并在采取拉伸处理频段的两端进行相应的压缩处理保证处理后的信号频谱不存在混叠。该算法的目标是使患者的剩余听力范围利用最大化，根据患者听力图的实际情况决定频谱伸缩处理的参数，使得处理后的语音信号频谱与患者剩余听力情况较好的吻合，以此提高助听效果。

某V型听力图对应频谱伸缩策略的频率输入输出关系如下图所示（以信号采样频率为16KHz为例）：

图5-6频谱伸缩算法频率对应关系

图中 f_a 类似于传统频谱压缩策略中的截止频率，在此频率之下患者听力情况仍然良好信号无需处理；AB段类似于传统频谱压缩策略中需要进行压缩处理的频段，由压缩比参数控制压缩程度，患者在 $f_{(a'')}$ 至 $f_{(b'')}$ 频段内应处于轻中度听损，保证频段 $[f_a, f_b]$ 内的信号压缩至 $[f_{(a'')}, f_{(b'')}]$ 后能够理解大部分的言语信息；BC频段 $[f_b, f_c]$ 是整个语音信号频段内听损最为严重的区域，通过拉伸处理使该频段内的信号扩散至附近频段 $[f_{(b'')}, f_{(c'')}]$ ，拉伸处理的目的在于尽量减少处理后语音信号在听损严重频段内的能量，同时也考虑到语音信号本身频谱的连贯性以及频谱间相对位置的关系，在尽可能保证语音质量的情况下进行严重听损段的频谱扩散处理，试图增加处理后语音对于特定听力状况下的可听度和可懂度；CD频段则与AB频段相类似，听损较为严重的 $[f_c, f_d]$ 段被压缩至听损较轻的 $[f_{(c'')}, f_{(d'')}]$ 段，以增强 $[f_c, f_d]$ 频段内信号的可听可懂度；DE段则与起始频段 $[0, f_a]$ 类似，该频段所对应的听力状态良好，信号无需进行补偿处理。

根据上述原理性分析，总结得频谱伸缩原始信号频率与输出信号频率关系曲线确定过程如下：

确定听力损失最为严重的频段中心频率 f ，该参数可从听力图以及辅助频率分辨测试而得，该频点对应于听力图中听阈值最大的频段

;

确定频点 f 附近的严重听损频段 $[f_b, f_c]$ ，该频段信号会被进行拉伸处理，一般可以通过听损患者在频点 f 处的频率辨别阈 f 得到 f_b 和 f_c 参数，亦可从患者听力图折线情况间接推算出；

确定严重听损频段 $[f_b, f_c]$ 的拉伸系数 γ ， γ 取值大小影响严重听损频段补偿效果， γ 值越大则补偿后信号严重听损频段内所剩能量越小，但是拉伸后信号的失真更大， γ 取值小时信号失真小，但是针对严重听损频段的补偿小也随之下降；

确定频率拉伸区域两边的频率压缩区域 $[f_a, f_b]$ 和 $[f_c, f_d]$ ，频率压缩的起始频率 f_a 与终点频率 f_d 以及压缩后所对应的频点 $f_{(a'')}$ 和 $f_{(d'')}$ 可从听力图参数中确定，两频段的压缩比分别记为 β_1 和 β_2 。

各个压缩段以及拉伸段参数详细确定策略后续章节将继续讨论。在原始信号与输出信号的频率转换关系图的参数确定之后即可对输入的数字化声信号进行非线性频率伸缩处理，具体步骤如下：

对数字助听器输入的数字化声音信号进行分帧，帧长 $N=2^E$ ， E 为正整数，一般地可取 $E=10$ ，帧长 $N=1024$ ，也即每帧信号具有1024采样点；

对单帧信号进行快速傅里叶变换，可得长度为 N 的初始频谱序列 $X=[x_0, x_1, \dots, x_{(N-1)}]$ ；

计算频谱序列中频率 f_b 对应的点 x_b 的下标 $b=2N f_b/f_s$ ，类似的计算频率点 f_a 、 f_c 和 f_d 所对应的频谱点 x_a 、 x_c 和 x_d 的下标值 $a=2N f_a/f_s$ 、 $c=2N f_c/f_s$ 和 $d=2N f_d/f_s$ ，根据频谱的对称性，初始频谱 X 的 $[b, c]$ 段和 $[N-b, N-c]$ 段是拉伸段， $[a, b]$ 段、 $[c, d]$ 段、 $[N-d, N-c]$ 段和 $[N-b, N-a]$ 段是压缩段，其中 f_s 是数字声信号的采样频率；

计算频率伸缩后的谱序列 $M=[m_0, m_1, \dots, m_{(N-1)}]$ ，其中 $[m_{(a'')}, m_{(b'')}]$ 段、 $[m_{(c'')}, m_{(d'')}]$ 段和与之对应的 $[m_{(N-b'')}, m_{(N-a'')}]$ 段、 $[m_{(N-d'')}, m_{(N-c'')}]$ 段为频谱压缩后对应的谱序列， $[m_{(b'')}, m_{(c'')}]$ 段和与之对应的 $[m_{(N-c'')}, m_{(N-b'')}]$ 段为频谱拉伸后对应的谱序列，原始频谱序列与伸缩后序列对应关系如图所示；

图5-7伸缩处理前后频谱序列对比

计算频率伸缩后的谱序列 $M=[m_0, m_1, \dots, m_{(N-1)}]$ 的拉伸段，拉伸中心频率 f （见图5-6）保持不变，也即

$$(5-2)$$

根据拉伸系数的定义：

$$(5-3)$$

可计算得拉伸段起始频率点的下标：

$$(5-4)$$

同理可得拉伸段中止频点的下标：

$$(5-5)$$

在拉伸段频谱相应坐标确定后, $[m_{(b'')} \&\& m_{(c'')}]$ 段和 $[m_{(N-c'')} \&\& m_{(N-b'')}]$ 段的值可由 $[x_b \&\& x_c]$ 和 $[x_{(N-c)} \&\& x_{(N-b)}]$ 线性插值而得;

继续计算频谱伸缩后的谱序列 $M=[m_0 \&\& m_1 \&\& \dots \&\& m_{(N-1)}]$ 的压缩段, 压缩后的 $[m_{(a'')} \&\& m_{(b'')}]$ 段可由压缩前所对应的原始频段 $[x_a \&\& x_b]$ 均匀抽样而得, 压缩前后信号满足, 同理 $[m_{(c'')} \&\& m_{(d'')}]$ 段可由 $[x_c \&\& x_d]$ 均匀抽样得到, 压缩前后信号满足, 继而频谱对称部分的 $[m_{(N-d'')} \&\& m_{(N-c'')}]$ 段和 $[m_{(N-b'')} \&\& m_{(N-a'')}]$ 段亦可由 $[x_{(N-d)} \&\& x_{(N-c)}]$ 段和 $[x_{(N-b)} \&\& x_{(N-a)}]$ 抽样而得。

继续计算频谱伸缩后的谱序列 $M=[m_0 \&\& m_1 \&\& \dots \&\& m_{(N-1)}]$ 的剩余段, 其中处理前后频谱下标 $a''=a, d''=d$, 也即处理后频谱的 $[m_0 \&\& m_{(a'')}]$ 段与原始频谱的 $[x_0 \&\& x_a]$ 相同, $[m_{(d'')} \&\& m_{(N-d'')}]$ 段与 $[x_d \&\& x_{(N-d)}]$ 段相同, $[m_{(N-a'')} \&\& m_{(N-1)}]$ 段与 $[x_{(N-a)} \&\& x_{(N-1)}]$ 段相同, 至此频谱伸缩后谱序列 M 计算完成。

本节以一典型“V”型听力图患者频谱伸缩处理为例介绍频谱伸缩算法中频谱伸缩策略以及其具体计算方法, 为后续分析建立基础。

5.2.2 频谱伸缩算法参数选择

由上节频谱伸缩算法原理介绍可知, 该算法所需要确定的参数包括补偿前的频率点 f_a, f_b, f_c 和 f_d 以及补偿后频率点 $f_{(a'')}, f_{(b'')}, f_{(c'')}$ 和 $f_{(d'')}$ 。同时, 频谱伸缩算法是一种频率补偿算法, 只有听力损失较为严重的听损患者需要借助频率补偿类算法以获取更好的助听效果, 因此, 算法启用条件也是算法参数选择之一。这些参数均需从患者听力图中获取。

首先是算法启动条件参数确定, 也即何种听力情况下需使用此频谱伸缩算法。遵循频率补偿类算法的初衷, 其所针对的用户是中重度听损患者, 因此, 通过听力图确定该患者的听力损失情况是否属于较重度患者即可。根据第二章中WHO听力损伤定级方案, 本文选取70dB HL作为患者听损情况分界点, 即当患者的听力损失情况达到中重度损失及以上时, 频率补偿算法启用。

其次, 算法需要确定补偿前频谱上各压缩段和拉伸段频点参数 f_a, f_b, f_c 和 f_d 。根据算法原理可知, 频段 $[[0, f]_a]$ 及 $[[f]_d, f_s/2]$ 属于听力状况良好的频段, 参考WHO听力损伤定级标准, 选取中度听损起始声压级(40dB HL)作为参考线, 划定频点 f_a 和 f_d 。当患者的高频听损均较为严重时, 即高频段不存在听力状况良好的频段, 则算法中 $[[f]_d, f_s/2]$ 段退化, 图5-8所示即无高频听损良好段, 故 f_d 取为 $f_s/2$ 。拉伸段起始频率 f_b 以及拉伸段终止频率 f_c 则由中重度听损判定标准线(70dB HL)确定, 如图5-8所示。同理, 当听障患者的高频听力能力呈单调下降趋势(听力图无“V”型)时, 拉伸段终止频率 f_c 退化, 此时, 可以判断患者在听阈70dB HL所对应频率至最高频($f_s/2$)的频段内听力情况损失均较为严重, 不存在听力损失较轻微的频段 $[[f]_c, f_d]$, 因此, 此时对听损较为严重的频段进行拉伸处理的收效甚微。面对此种听力情况, 算法将第一压缩段的截止频率延伸至最高频, 算法退化为普通的频率压缩类算法, 与已有算法吻合。

图5-8 补偿前频谱各频点参数划定

最后确定补偿后频谱各频率节点参数。由算法原理描述可知补偿后第一压缩段起始频率 $f_{(a'')}$ 与补偿前第一压缩段起始频率 f_a 相同; 补偿后第二压缩段终止频率 $f_{(d'')}$ 与补偿前第二压缩段终止频率 f_d 相同。因此需要单独确定的参数为补偿后拉伸段起始频率 $f_{(b'')}$ 和终止频率 $f_{(c'')}$ 。频点 $f_{(b'')}$ 和 $f_{(c'')}$ 的值决定着听损严重段 $[[f]_b, f_c]$ 拉伸后所对应的频段。用 V_a 表示听力状态良好划分点(本文取40dB HL), V_b 表示听力损失中重度划分点(本文取70dB HL), 并用 H_{max} 表示听损最大值(如图中即为85dB HL), 用 V_m 表示补偿后拉伸段频点参数划分点。则 V_m 可用如下公式算得:

$$(5-6)$$

从参数 V_m 的计算过程可以看出, 其值取决于患者的听损最大值 H_{max} 和重度听损划分界限 V_b , 同时其值不会小于听力状态良好划分点 V_a 和听力损失中重度划分点 V_b 的均值。参数 V_m 值将作为拉伸段 $[[f]_b, f_c]$ 拉伸处理后的拉伸段参数 $f_{(b'')}$ 和 $f_{(c'')}$ 的划定标准, 因此该参数决定算法的拉伸度和压缩比。听损最大值 H_{max} 越大, 则拉伸比和压缩比越大; 同理, 当听损最大值 H_{max} 与重度

听损划分界限 V_b 所差无几时, V_m 值和 V_b 相近, 此时由 V_m 确定的频点 $f_{(b'')}$ 和 $f_{(c'')}$ 的值和补偿前相应的频点值 f_b 和 f_c 相近, 也即压缩比和拉伸比将接近1。此机制与听力图状态分布和理论上补偿方案吻合。

由 V_m 参数值确定频点 $f_{(b'')}$ 和 $f_{(c'')}$ 的值方法与其他节点频点值确定方法一致, 如图5-9所示:

图5-9伸缩后对应频点划定

至此, 算法中个节点频率参数均已确定, 在节点频率参数确定之后便可得到频率伸缩补偿方案的核心数据——补偿前后频率对应关系。在以上各算法参数确定方案介绍中有提到几类边缘型问题, 即当用户的听力图有所变化时, 补偿前后频率的对应关系也会有所变化。因此下文将对几类相关的听力图类型及相应的频率对应关系进行介绍。

高频听损良好（听力图呈“V”型）

此类听损患者最高频听力情况保存良好, 听损区域发生在中频段, 且听力损失严重, 典型听力图及相应的频率对应关系如下:

图5-10 V型听力图及相应频率对应关系

高频听损中度（听力图呈半“V”型）

此类患者高频听力损失较为严重, 高频段不存在听损良好段, 而且由于高频听损严重, 在听力图上无法得到补偿后压缩段的截止频率 $f_{(c'')}$ 。面对此种情况, 算法将 $f_{(c'')}$ 值取为补偿前压缩段的截止频率 f_c , 如此则第二压缩段退化, 而拉伸处理则会将拉伸段全部向低频段（由 f_b 至 $f_{(b'')}$ ）拉伸, 因为患者在第二压缩段所在频段的听力状况并不良好, 而低频段听损保存较好, 这样处理更有助于对听力完好区的利用。示例听力图及相应的频率对应关系如下:

图5-11半V型听力图及频率对应关系

高频听损重度（直降型听力图）

此类听损患者的听损情况随着频率上升而越发严重, 根据算法参数确定方案, 补偿前拉伸段截止频率 f_c 亦将无法得到, 此类听力图表明患者高频段听损严重, 补偿处理时, 无法利用高频段的听力能力。因此, 算法在面对此类听力图时, 仅采用第一压缩段, 并以其压缩比作为参考, 将 f_b 以上频段信号全部压缩, 此种处理方式与传统的频谱压缩补偿策略类似。示例听力图及相应的频率对应关系如下:

图5-12直降性听力图及频率对应关系

5.2.3 频谱伸缩算法小结

频谱伸缩算法作为频率补偿类助听算法之一, 其所面向的对象亦是听损较为严重、传统响度补偿类助听方案效果不佳的患者。传统的频率补偿类算法主要包括频谱压缩类和频谱搬移类, 专注于高频听损严重的听障患者, 而对于次高频段、中频段听损严重的患者, 此类助听方案并不合适。频谱伸缩算法正是针对传统的频率补偿策略的此类不足而提出的一种新的频率补偿策略。该算法以患者的听力图作为基础依据, 结合频谱压缩和拉伸处理, 生成相应的频谱补偿处理方案。

频谱伸缩算法首先以完整V型听力图患者作为目标对象, 从其听力图得出各个算法参数, 从而确定补偿前后的频谱对应关系。然而对于不同类的听力损失患者所需的频率补偿策略大相径庭, 因此算法对听力图类型进行较为细致的划分。对于听损并没有特别严重的患者, 无需采用频率补偿类算法, 该类算法会改变语音频谱结构使其听起来有失真感, 强行开启频率补偿类算法可能会适得其反, 因此, 算法的开启条件设定很有必要。同时, 参数确定方案并不能完美匹配所有类型的听力图, 因此, 针对不同类型的听力图, 参数确定

算法需要进行相应的适配以达到较好的补偿效果。参数确定和频谱关系的适配均遵循患者剩余听损最大化的原则。算法通过听力图的不同而进行参数和策略调整，使得其在合理应对V型听力图的基础上也能兼顾最高频听损严重的其他类患者，增加了算法的适应性。

本节主要介绍频谱伸缩算法基础理论和相关的参数选择策略以及面对各种不同类型听力图时算法的应对方式，后文将继续分析算法的补偿效果。

频率补偿实验

本节介绍以频率伸缩算法为处理方案的频率补偿实验。以一“V”型听力图的中重度听损患者为例，以TIMIT语料库的某段语音为补偿对象，展示算法各参数的计算以及最终的补偿效果。

5.3.1 语料库介绍

实验所选取语料来自TIMIT语料库。TIMIT的全称是The DARPA TIMIT Acoustic-Phonetic Continuous Speech Corpus, {74%:该语料库由德州仪器 (TI)、麻省理工学院 (MIT) 和斯坦福研究院 (SRI) 共同研制构建[61]。}语料由德州仪器 (TI) 录制，在MIT进行转录和后期处理，并由美国国家标准技术研究所 (NIST) 验证和刻录为CD-ROM。语料库建立的初衷是为了给语音音素学习学者以及语音识别系统的发展和评估提供语料。TIMIT语料库的录制包含了630名说话者以及全美的8大英语方言，每位参与录音者所朗读语句音节信息丰富。语料库中的语音均经过时间对齐正交处理，每段语音文件采用16KHz的采样率，16位编码处理。

TIMIT语料库应用广泛，相对而言公正客观。虽然本文不涉及语音识别等相关技术，但是为了试验的客观性，选取公认语料库的语音段作为实验数据依然有必要。因此，本章后续实验以TIMIT语料库中dr1组mjsw0目录下“sa2.wav”语料作为实验语料。

5.3.2 频率补偿实验

本节实验以一噪声性聋中重度耳聋患者为例，其听力损失严重区域主要在2KHz至5KHz的中高频段，其听力图状况如下图所示：

图5-13示例患者听力图及频率对应关系

从图5-13大致可以看出患者的严重听损段在2KHz之后的中高频段，在5KHz左右之后的最高频段听力情况稍有好转，属于5.2.2节中的第二种情形。从听力图可以看到，患者的最严重听损值 H_{max} 与 V_b 相差无几，因此 V_m 值亦与 V_b 相近，这便使得患者的频谱输入输出关系中压缩段和拉伸段的系数均接近于1，这也是频率补偿策略的初衷所导致的。只有当患者的听损情况非常糟糕时其压缩或拉伸处理力度才会较大，因为对频谱的拉伸和压缩必然会导致语音的可听度受到影响。

经过算法参数计算程序处理，得到各算法参数值如下：

将参数值导入频谱关系生成程序，并利用所得补偿前后频谱对应关系处理样例语料“sa2.wav”，最终补偿结果如下所示：

图5-14补偿前后语音信号时域及频域对比（半V型）

补偿结果的频谱变化从图中的补偿前后频谱对应关系便可预知，图中语谱图对比使得频谱处理更为直观。因为此例患者听损情况基本刚出于启用频率补偿的边缘，相对于听损特别严重的患者而言，其听损情况比较良好，所以其相应的应对方案对频谱的处理变化比较轻微。通过这种轻微的频谱处理可以使该患者对其听损严重段（2KHz-5KHz）内的语音信号的聆听和理解有些帮助，同时频谱拉伸后所造成的轻微压缩也不影响其他频段信号的正常理解。

在面对听损更为严重的患者时，算法所给出的频率补偿方案对最终输出语音的影响也将更为明显。以重度高频听损患者为例

（5.2.2节，第三类），该类患者整个高频段的听损均比较严重，且随着频率上升，听力状况愈发下降。针对此类患者，本算法给出

的频率补偿方案与传统频谱压缩类补偿方案类似，5.2.2节中已进行分析。选取5.2.2节中第三类患者的示例听力图及频谱对应关系为例，对实验语料段‘sa2.wav’进行补偿处理。选取图5-12中听力图为例计算算法各参数如下：

根据听力图类型确定算法补偿方案类型，并以上所得算法参数为基础得到频谱对应关系，最终对实验语音段‘sa2.wav’进行补偿，结果如下：

图5-15补偿前后语音信号时域及频域对比（直降型）

从频谱对应关系和最终的语谱图对比可以看出：频段0-900Hz段信号保持不变，频段900-8000Hz内的信号被压缩至900-5980Hz。算法处理后的语音极高频段信号全被压缩至近6KHz以内，听力正常人聆听处理后语音会显得比较闷，但是对于高频段听损严重的该患者而言，这样的压缩处理意味着有机会获得更多的高频段信息，对于帮助理解语音信号意义显著。

本章小结

本章主要针对助听器中频率补偿类算法进行研究。首先，本章对频率类补偿算法的基本背景的适用情形等相关内容进行介绍。随后，第一节主要介绍已有的传统移频助听类算法，包括频谱压缩和频谱搬移两类处理方案，这类算法的研究已颇具历史，在一些具有移频助听该功能的助听器中已投入使用；第二节主要介绍在传统移频助听算法基础上所提出的频谱伸缩助听算法，针对传统移频助听算法的不足和局限性，本节主要表述频谱伸缩算法的核心处理策略、参数选择策略，对频谱伸缩策略的基本原理和处理方案进行介绍；最后，第三节进行仿真实验，介绍了选取的实验语料库，以及补偿案例效果，从仿真实验的基础上可以看出，算法在传统移频助听的基础上进行拓展，能更好的适应不同类型的听力障碍患者。

系统综合架构及测试

听力测试及补偿系统主要包含测听及补偿两个部分。听力测试部分各功能原理相对简明，核心在于操作流程；补偿系统核心在于各补偿算法细节，系统操作流程相对简单。前文主要对听力测试部分和补偿算法进行介绍，本章将对系统的整体架构进行分析，同时对于系统中的各个功能进行测试。

6.1 测听及补偿系统综合架构设计

听力测试系统主要为使用者提供简单的听力情况自评测试功能，而补偿系统的目的则是提供助听补偿的功能，二者构成整个测听补偿系统。听力测试系统的结构在第三章已有详细分析，本节首先对补偿系统的结构进行介绍，然后介绍整个系统的综合架构。

补偿系统结构

助听补偿系统的整体结构如图6-1所示：

图6-1助听补偿系统结构

如前文所说，助听补偿部分的核心在于补偿算法的细节，因此其结构相对比较简明，主要包括算法参数计算和算法实施两部分。

其中算法参数计算部分的主要功能是给算法提供相应的参数值。该部分的数据源来自于测听系统给出的听力信息数据和算法提供给用户选择的接口。听力信息数据包括听力图信息、频率分辨能力等等；用户输入信息包括补偿算法中频率补偿是否开启、通道选择等功能。

补偿算法实施部分则是补偿算法的最终处理阶段。算法以实施输入语音作为程序输入、以算法参数为算法控制、以补偿处理后的语音作为输出。该部分是算法的核心处理部分，其效果影响着整个补偿系统的补偿效果。因此，该部分是补偿系统中理论研究最为关键

也是最为需要的部分。

然而，在补偿系统中使用的具体补偿算法不同时，其相应的参数计算、算法实施等过程可能会有写差别（例如，响度补偿中可能会增加共振峰提取过程），图6-1所示结构仅是助听补偿系统普适结构。

系统综合架构

测听及补偿系统整体架构如图6-2所示：

图6-2听力测试及助听补偿系统整体框架

整个系统以Android移动终端为基础，划分为听力测试和助听补偿两部分。如前文所述，听力测试部分包括纯音听阈测试、音调分辨力测试以及言语测听等相关内容；助听补偿部分则主要包含响度补偿和频率补偿等相关内容。两者之间通过以用户唯一标识命名的信息文件（userid.info）传递数据。听力测试部分将测试结果写入userid.info文件，在助听补偿部分需要读取某些数据时，则按照信息文件的格式读出相应的数据即可。信息文件格式如表6-1所示：

表6-1信息文件格式

编号 参数说明 变量类型 字节数

1~11 听阈值 int[] 4*11

12~22 舒适阈值 int[] 4*11

23~33 痛阈值 int[] 4*11

34~44 音调分辨力值 double[] 8*11

45 言语识别率值 double 8*1

46 言语识别阈值 int 4*1

在表6-1和图6-2所展示的功能及参数之外，测听系统以及助听补偿系统中还可以扩展相关的其他测试项及功能项，这时，信息文件跟随相关功能需求进行扩展修改即可。

6.2 系统功能测试

听力测试及助听补偿相关软件基于Android系统设计研发，运行于Android 4.0.3（API 15）及以上的安卓系统上。听力测试功能和助听补偿功能分别用单独的应用软件实现，两者通过文件系统存储和传递用户数据。下文将对听力测试软件及助听补偿软件相应功能进行介绍、分析及测试。

6.2.1 纯音听阈测试

纯音听阈测试以3.3.1节所述升降法为基础，稍作修改进行检测。测试流程修改点主要在于：1. 细化声压调节粒度至1dB HL；2. 增加舒适阈和痛阈选项，按照听阈测试的流程完成舒适阈和痛阈测试。

软件界面截图以及测试完成案例显示如下：

图6-3纯音听阈测试界面

图中左图为听阈测试操作界面，提供折线显示、频点切换、声强调节、用户反馈以及信息保存和历史记录查看等功能；右图为一已测量完成的听阈痛阈舒适阈折线图。本测试所选包括125Hz-8kHz频段上的共11个频点，声强控制粒度为1dB，为了良好的测量效果和更准确的测量结果，建议测试者使用耳机并处于安静的环境中进行测试。

6.2.2 频率分辨力测试

如3.3.2节所述，本文所述移动终端系统采用改进的频率分辨力测试法，旨在进行固定声强下的频率分辨力大致测量，以契合移动终端的实际条件并达到初步的频率分辨能力的检测效果，更专业的频率分辨力测试还需使用基于专业设备和心理物理调谐曲线的测试方法。

如图6-4所示是频率分辨力测试软件操作界面以及某测试者（听力正常）测试结果：

图6-4频率分辨力测试界面

受试者只需佩戴专业播放耳机，并选定声压级（即图中音量），测试过程中声压级固定，通过点击播放按钮，聆听短音信号，并做出选择“相同”或“不同”的选择，每个频点的每一级分辨力的每组测试包含三次短音测试，全部正确方可进入下一级。测试频点涵盖125Hz-8000Hz的11个频点，整个测试所需时长大约在10分钟。

6.2.3 言语识别率测试

言语识别率测试流程及原理如3.3.3节所述。测试按照先听后选的方式进行，受试者聆听语料，然后选择相应的文字及拼音所组成的选项中与所听到的语音相匹配的选项，软件将记录受试者的每次选项结果，并在测试结束后显示整体测试结果。

测试交互界面及某次测试结果如图6-5所示：

图6-5言语识别率测试流程界面

言语识别率测试亦是固定声压级下进行。测试开始首先选择本次测试题库数目；之后按照测试流程遍历完整个测试集，最终显示并记录测试结果。

6.2.4 言语识别阈测试

软件系统言语识别阈测试采用文献[21]所述方法，以双音节词组为测试语料，通过不断降低声压级遍历整个可听范围，最终使用公式（）计算言语识别阈。与言语识别率测试类似，受试者反馈方式为听词选字，选项由拼音和汉字组成。

图6-6为言语识别阈测试操作流程截图：

图6-6言语识别阈测试流程界面

测试开始时首先设定初始声强，初始声强影响着测试的准确性。由于词库容量有限，当初始声强选择过大时，遍历完所有词组后仍未达到测试中止标准（某族测试全部无法听清），则无法测出言语识别阈。之后测试流程与言语识别率测试类似，通过聆听语料选择与其对应的选项。最终，测试成功的情况下给出测试结果。图即为某位听力基本完好的受试者所测结果。当初始声强设定不合理（过

大)时,程序将会提示要求降低初始声压级后再次测试,如下图所示:

图6-7言语识别阈测试(过大初始声强)流程界面

6.2.5 助听补偿功能测试

系统所提供补偿类功能由助听补偿软件实现。软件集成第四章响度补偿及第五章频率补偿算法,根据听力测试所得参数定制补偿参数,并实现补偿功能。助听补偿所需大部分参数均通过中间信息文件(userid.info)获得,因此助听补偿软件界面主要实现补偿开关控制及部分参数设置功能。

助听补偿界面如下图所示:

图6-8补偿系统界面

软件第一栏提供助听器开启和关闭功能(需连接专业耳机后方可开启,否则单个扬声器同时充当播放和录音功能将产生明显的啸叫);后两栏分别为响度补偿和频率补偿控制及参数设定模块,响度补偿中SCE factor参数控制配对比增强的程度,值为0-1之间,0表示不进行谱对比增强,1则表示增强强度最强,详见4.3.2节;一般而言,频率补偿开启关闭控制由听力图参数便可确定,本系统考虑进行频率补偿的特殊性,提供其开启关闭功能旨在给使用者更多的选择余地,同时针对频率补偿中听力情况良好线 V_a 和听力损失中重度划分线 V_b 也提供调节能力,旨在为使用者提供更加匹配的频率补偿功能。

6.3 本章小结

本章以前文所述理论为基础,分析基于移动终端的听力测试及补偿系统的整体架构和实现。第一节主要分析助听补偿系统软件设计结构,并阐述整个系统的整体架构;第二节以具体功能细节为划分,展示软件系统各个部分功能,阐述实际操作流程和相应说明。

总结与展望

随着现代社会的不断发展,听力情况愈发的得到人们的重视,同时严重的听力损失也在影响着成千上万的听障患者的日常生活。一方面,长时间佩戴耳机、长期处于高分贝噪声环境、年龄增长机体功能衰退等等造成听力损失的原因不断的变得复杂;另一方面,轻微听力损失往往会被人们忽视,听力保护意识薄弱,使得因听力损失而影响正常生活的人群不断增加。鉴于此,针对听力损失的出现的助听器及其相关内容的研究日益受到国内外学者的重视。本文以移动终端为载体,研究听力测试和助听补偿相关算法和内容,一方面对助听算法进行研究和拓展,另一方面借助移动终端的普及型和便利性增加人们听力保护意识并增强对自身听力情况的了解。

7.1 本文主要工作总结

(1) 本文第三章阐述听力测试系统各部分结构及原理,在已有测听理论的基础上,基于终端完成各测试操作流程设计,并编程实现各个测试具体内容,同时考虑到系统设计的初衷以及使用者操作的便利性,部分测试以放弃测试结果精确程度的代价,极大降低测试复杂度和耗时,同时所提供的频率分辨精度能够使受试者对其自身听力情况有基本的了解。

(2) 本文第四章针对响度补偿算法处理后语音信噪比以及噪声环境下语音可懂度提高等问题,研究谱对比增强算法,并提出一种基于共振峰增强的谱对比增强算法。算法从语音基本特征入手,利用快速傅里叶变换模拟多通道处理,对每帧信号整体频谱进行补偿处理,同时对传统处方公式进行改进,增加限幅输出原则,以保护佩戴者听力能力。算法计算量较低,适合移动终端的通用计算能力,同时输出语音频谱共振峰频点处得到较多补偿,语音整体可懂度有所提高,并且经过共振峰增强和谱对比增强之后语音信号信噪比得到提升,频谱清晰度提升,处理后的语音最后经过响度增益增强,保证了听损严重区域的频谱强度以及语音的整体声强。算法旨在响度补偿算法的频谱微处理,通过语音特征的增强以及整体频谱的对比均衡处理,提高响度补偿算法所输出语音的可懂度,以及其抗噪能力,为助听器响度补偿算法研究特别是植入耳蜗噪声环境下补偿效果提升提供思路 and 方案。

(3) 本文第五章针对传统频率补偿方案的局限性提出一种基于频谱拉伸压缩处理结合的频率补偿方法。传统频率补偿方案均针对最高频段听损严重患者进行降频补偿,有的通过频谱压缩达到降频效果,有的则通过频谱搬移完成,其目的是充分利用患者剩余听力良好频段,将高频信息转移至较低频段。然而,这种转移的前提是听损患者是高频听损严重类型,且此类处理方法对频谱的影响较大,补偿后语音变化较大。频谱伸缩算法亦是以充分利用患者剩余听力良好频段为目的,设计一种新的频谱变化方案。根据患者听力图类型确定压缩频段和拉伸频段,尽量将听损严重频段的信号扩散至听力情况良好的频段,同时控制压缩和拉伸的程度,减小频谱压缩和拉伸处理带来的副作用。算法对听损严重患者听力图进行分类,并针对不同类型的听力图提供不同的补偿方案,使算法的适应性和适用性得到较好提高,同时补偿方案也更具有定制性,与待补偿患者的听力情况更为贴切。

(4) 本文第六章节首先阐述基于移动终端的助听补偿系统设计,并分析听力测试及助听补偿系统的整架构。随后给出听力测试系统和助听补偿系统间的联系文件的详细结构。最后,对系统中各个功能进行展示及测试,并对每个测试功能进行简要分析。至此,整个听力测试和助听补偿系统基本阐述完成。

7.2 展望

虽然本文在测听理论以及补偿理论的基础上完成听力测试及助听补偿系统的基本设计,但是其中相关测试内容及算法均处于初生阶段,本人研究时间和水平均有限,因此,这里对后续可研究内容和方向进行展望:

- (1) 听力测试系统可以继续完善。更加完善的听力测试系统可以加入PC端软件的支持。一方面,在有条件的情况下PC软件操作更为方便;另一方面,PC的计算能力和处理能力相对移动终端而言也有提升。
- (2) 听力测试系统可添加更多的听力测试。基于移动终端设计听力测试系统的一个重要因素和目的即是提高人们听力保护意识。增加更过基于听力测试和听力保护的相关内容会使得听力测试系统更加丰满,更加实用。
- (3) 目前助听补偿系统只提供了本文所述的两种补偿算法,缺少对比性,同时助听补偿系统所用算法过于单一。将响度补偿相关的已有经典算法继承入助听补偿系统,并为使用者开放算法切换接口。助听算法的增加和完善可以使助听补偿系统对助听器选配有提前体验和指导。
- (4) 目前本文并未研究基于频谱搬移的助听效果研究。对听损严重频段的信号是否通过搬移处理会获得更好的补偿效果还不得而知。基于频谱伸缩的助听补偿算法可以基于频谱搬移处理进行衍生,利用搬移的方法转移听损严重频段信号并对其他频段信号进行适当压缩或拉伸,并对比补偿效果。
- (5) 听力测试及助听补偿系统的软件界面设计及操作反馈流程可以优化,提升系统使用体验。

致谢

{59%:时间如白驹过隙,三年研究生时光即将结束。这三年的学习生活,收获丰富,受益良多。}一路走来,有许多指导、帮助过我的人相伴,没有你们就不会有现在的我。

{62%:首先我要感谢我的导师赵力教授。赵老师在我的学术研究上给予了许多帮助。}从论文选题到具体研究再到系统改进,导师都付出了大量心血,他的指导和帮助是我能够顺利完成课题研究的关键;导师治学严谨、平易近人,生活中亦给予了贴心的帮助和关心。转眼间就要毕业了,但是师生情谊不会随着时间的流逝而消失,由衷地感谢导师一直以来的关心和帮助。

感谢梁瑞宇、王青云博士,他们开拓的思维、对科研的执着和态度,感染者身边的每个人。在课题研究当中他们给予了宝贵的意见,并无私的帮助我完成课题研究。谢谢你们。

感谢实验室的同门和研究伙伴们，是大家共同的努力，使得实验室的氛围和环境如此温馨积极，也感谢你们给予的帮助和关怀。感谢你们每一个人：何旭、黄晨、陈晟、丁一坤、蒋浩、刘成宇、刘耘、皮慧、谢跃、张明阳、张奇、张祺威、邹炉庚、张希龙、黄玉洁、卢小芳，愿你们在新的人生阶段展翅高飞。

感谢我的父母，他们的关爱和支持是我一直以来的精神支柱，没有父母，就没有我。渐渐地长大了，很少陪在父母身边，希望他们健康平安。

感谢审阅论文的老师，感谢你们的宝贵意见，希望大家都有更多的学术成果出现。

最后，再次感谢所有给予关心和帮助的老师、同学、家人和朋友，希望你们平安健康，幸福快乐！

检测报告由PTcheck文献相似度检测系统生成
Copyright 2007-2016 PTcheck