**基于移动终端的听力测试和补偿系统设计与实现**

目录

[**第 1 章** **绪论** 2](#_Toc441132667)

[1.1 课题来源及背景 2](#_Toc441132668)

[1.2 课题目的及意义 3](#_Toc441132669)

[1.3 国内外研究现状分析 3](#_Toc441132670)

[1.4 主要研究内容 4](#_Toc441132671)

[**第 2 章** **助听器测听基本知识** 5](#_Toc441132672)

[2.1 纯音听阈测试 5](#_Toc441132673)

[2.2 频率分辨力测试 5](#_Toc441132674)

[2.3 言语测听 6](#_Toc441132675)

[**第 3 章** **基于移动终端的测听系统实现** 7](#_Toc441132676)

[**第 4 章** **响度补偿算法** 7](#_Toc441132677)

[**第 5 章** **频率补偿算法** 7](#_Toc441132678)

[**第 6 章** **移动终端算法实现与测试** 7](#_Toc441132679)

1. **绪论**
2. 课题来源及背景

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]。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。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），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[8]；另一方面，随着年龄的增长，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，对老龄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[4]。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[6]。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[4]。此外，由于技术垄断等原因，使得只有少数的听力障碍患者能佩戴助听器[6]。数据显示，中国助听器选配比例仅在5%左右[4]，提高听力障碍患者的听觉功能面临巨大挑战。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。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[7]。可见，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[8]。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。

1. 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（比如手机）的听力检查系统，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、更精确的验配效果[10]。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。

1. 国内外研究现状分析

针对听力检查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；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；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。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；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；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；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 Fayçal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；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；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。

国内对于听力检查和听力计的研究起步较晚，但是发展迅速。在20世纪50年代听力计便在我国一些医院中使用；1964年我国第一部听力学著作《听力学概论》出版；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；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。随着听力学研究者和听力康复临床工作者们的深入研究，综合国内外的听力检查研究不难看出，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. 主要研究内容

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。主要工作内容可概括为以下几个方面：

1. 对助听器测听理论的基本知识进行研究；
2. 实现基于移动终端的测听软件系统开发；
3. 对助听器补偿算法进行研究；
4. 基于移动终端的听觉补偿算法实现和测试。
5. **助听器测听基本知识**

助听器验配是助听器佩戴中十分关键的步骤，通过听力检查和言语测听可对患者的听觉功能进行分析描述，为助听器验配提供数据基础，有助于康复治疗。常见的测听项有：纯音听阈测试、频率分辨力测试和言语测听，其中，言语测听包括言语识别率和言语识别阈测试。

1. 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。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。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。

1. 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，或者说其频率分辨阈过高，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。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，需针对患者的频率分辨情况对助听器的频率补偿算法进行调节以提高助听效果。

受试者的频率分辨阈一般可由其心理物理调谐曲线给出，但是心理物理调谐曲线的测量过程过于漫长，长时间的测量也会给受试者造成一定的伤害。一般而言，针对某频点的频率分辨测试可使用类似于听阈测试的方式进行，通过不断给予源频率信号和差频信号，根据受试者的主观判决来大致测量其在此频点的频率分辨能力。

1. 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。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。

言语测听的方法参照国家标准GBT17696-1999。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。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%是所对应的声压级。

1. **基于移动终端的测听系统**

在深入研究测听相关国家标准后，基于广泛使用的Android移动端操作系统，根据本测听系统的相关技术指标和要求构建测听系统。

1. 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，测听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率图；
4. 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；声压级误差在内；
3. 频率误差小于1%；
4. 环境噪声小于40dB；
5. 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；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。

1. 测听软件设计架构

测听系统主要功能通过测听软件实现，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）。测听软件基于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，完成各测听项并保存测试结果，指导助听器验配。

1. 纯音听阈测试

根据针对测听方法的国家标准[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。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：

1. 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；
2. 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；
3. 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；
4. 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：

1. 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；
2. 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）；
3. 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图如图3.3所示：



图3.3 升降法测听阈流程图

1. 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。

1. 心理物理调谐曲线法

理论上而言，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。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。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未能很好的在临床上得到应用。

1. 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，我们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，因此我们采用该方法作为测听系统中频率分辨力评价的方法。

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, 2000Hz, 3000Hz, 4000Hz, 6000Hz, 8000Hz；同时，频率辨别阈分为九个档位。测试步骤如下：

1. 选择本轮测试频点。从剩余待测频点中选择本轮测试频点*，*并生成相应的纯音信号（声压级可由受试者调节固定）；
2. 生成给声信号。当前频偏系数初始化为，对比信号频率的取值空间，三者出现的概率分别为。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；
3. 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，若全部正确则将下调一档，继续从步骤二开始测试；若仍有误判，则认为当前受试者无法分辨，并将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

1. 言语测听

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GBT 17696-1999。

1. 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。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

1. 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。

除了根据言语识别阈的定义方式进行测量之外，文献[成人言语测听的基本内容及其临床价值]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：

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

1. 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。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，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，无法一一标定完成；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3.7 Android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。

1. **响度补偿算法**

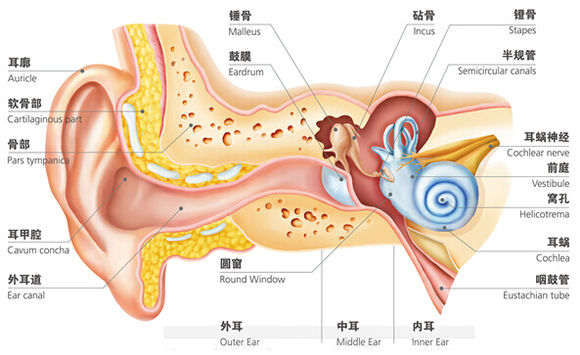
现代助听器的核心功能是针对佩戴者的听力损失情况自动调节输出语音的响度、频率等参数，使患者能够在听觉上更好的感觉和理解外界发出的声音。助听器完结合人耳特性、语音信号处理基本原理以及听障患者的听力损失情况对输入声信号进行恰当处理并输出，最终提升佩戴者对原声信号接收能力的过程称为响度补偿。因此，响度补偿技术是数字助听器的核心算法，也是助听设计的首要目标。

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；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。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，提高响度补偿算法性能。

1. 语音信号听觉基础
2. 人耳听觉系统

人耳听觉系统是人接收外界声信号的生理器官，其声压感知能力和分辨能力极其灵敏，为人接收和感知声信号提供生理基础。人耳听觉器官可分为三个部分：外耳、中耳和内耳，其生理结构如下图所示。



处在耳器官声信号接收最外层部分是耳廓（也叫耳郭），它主要负责接收来自人前方和侧方的声波信号，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郭给人听觉处理系统提供一定的声源定位信息。声信号经耳郭汇集后由外耳道传至中耳鼓膜处。外耳道长约2.2厘米，始于耳郭，终于耳膜，具有一定的共鸣特性，对语音信号高频部分（3KHz以上）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。声振动信号传至鼓膜后经听骨链传递引起而耳蜗中淋巴液和基底膜的振动，由此将声信号传至内耳。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。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。

通过对人耳生理结构的分析以及大量听觉实验研究可知，人耳感知语音信号具有一些特性，称为听觉感知特性，主要描述人对耳所接收语音信号的主观感受。这些主观感受与人耳生理结构、语音信号产生机理等多方面因素密切相关，主要特性如下：

1. 语音信号的共振峰信息对语音理解和人耳感知至关重要；
2. 某一频率的单频信号在人耳中产生的感知效果与该信号的声压级成比例；
3. 人耳的频率分辨能力与信号的声强有关，声强太小或太大均会影响频率分辨力。
4. 语音信号基本特性和数字模型
5. 语音信号特性
6. 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。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。

1. 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。

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。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。

1. 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。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，分别对应于上述发声器官，其原理框图如下图所示：



语音信号产生数字模型

由上图可知，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。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。

1. 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：



其中A表示脉冲加权的幅值。其Z变换频域表示如下：



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



其中和分别是单个三角脉冲的上升和下降时间。将上式进行Z变换得其系统函数如下：



其中和是常数，是一个二阶全极点模型。因此，模拟浊音激励部分的数字模型系统函数可以表示为：



1. 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：



辅音部分由并联型零极点IIR滤波器表示，其传输函数如下：



1. 辐射模型

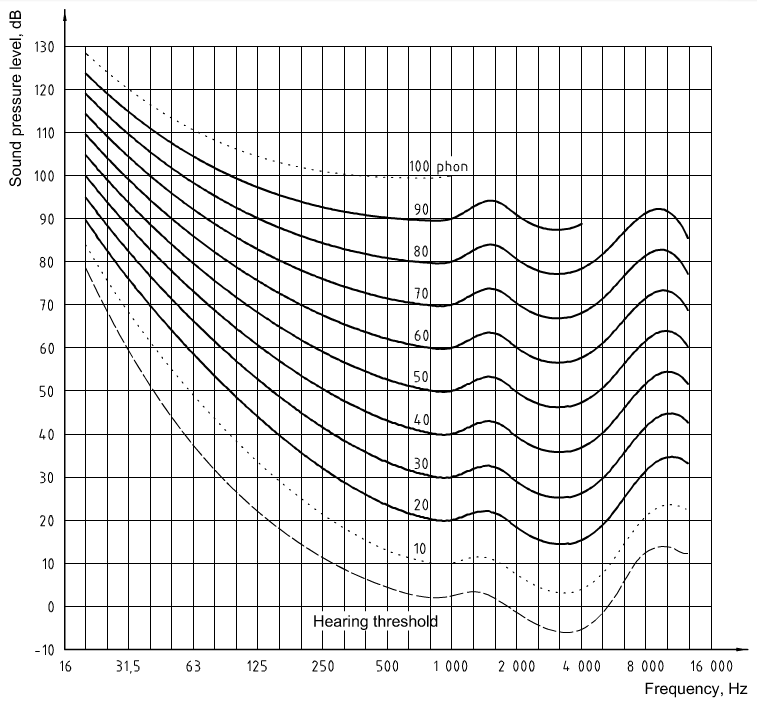
辐射模型模拟口鼻对所发出的语音信号向外辐射的过程，唇齿结构特征对辐射结果有较大影响。辐射部分用如下所示的一阶差分模型描述：



1. 响度补偿基本原理和常见算法
2. 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，如图所示。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。



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。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。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中，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。

1. 宽动态范围压缩（WDRC）

WDRC技术是响度补偿算法的核心和灵魂。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：

1. 将信号变换至频域；
2. 根据听力图和信号特征频率点的声压级计算相应的增益；
3. 利用线性插值方法得到全频域增益曲线；
4. 增益补偿之后将信号变换至时域输出。

如图所示，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 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（输入信号声压级用表示，输出信号声压级用表示，增益用表示）：

1. 当，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；
2. 当时，输出信号声压级，即如图中（，）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（单位为dB SPL），一般而言此段增益大于零；
3. 当时，输出信号声压级，如图中（，）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，或大于零，或小于零；
4. 当后，I/O曲线输出开始保持为听障患者的痛阈值，即，避免过幅声信号对助听器佩戴者造成伤害，起到一定的限幅效果，此时增益小于零。

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，配合其主要思想完成多通道响度补偿。

1. 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的思路对其进行改进，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。

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，算法主要流程如下图所示：



由图可见，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，多通道处理思路是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。

1. 频谱搬移算法

针对高频区听力损失严重的患者，如果输入信号频谱在高频区能量较大，则其能够理解该语音的几率将会大大降低。为了提高高频区听力缺失患者对高频成分较多信号的理解能力，研究者们将高频区信号搬移压缩至可听频率范围内。移频压缩之后，语音本身会出现一些信息混叠、丢失等情况，使得听起来音调稍有变化，但是移频过程中保持着各频点的相对位置关系，使得语音失真在可接受范围内。但是，对于患者而言，从基本听不高频部分声音到听到该声音移频后的语音是质的变化。因此，对于高频信号听觉能力缺失患者而言移频效果利大于弊，具有实际意义和研究价值。

移频压缩算法首先确定需要搬移的频率范围以及目标频率范围。用和分别表示被搬移部分移频前后的频谱，则移频关系可如下表达：



令，为搬移前后频率范围比值，表示频谱搬移前后压缩的程度，称为压缩比。移频压缩处理算法一般流程如下：



图中频谱搬移处理部分即是使用上述搬移算法对信号频谱精细调整，完成搬移压缩工作，经过逆变换至时域与未搬移部分合并输出。

1. 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。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[][]，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躁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，以及以上算法对含躁情况表现的不足，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躁语音信号的清晰度[]。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。

1. 共振峰

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。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；而倒谱法运算量较大，故本文采用LPC法提取共振峰参数。

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：



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：



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：



又可以表示为个极点级联的形式：



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AR模型峰值检测的方法获取共振峰频率。

1. 谱对比增强
2. 共振峰谱增强响度补偿算法

1. 人耳听觉系统及其特性（分析人耳听觉系统和耳蜗的频率选择特性）

2. 响度补偿基本原理和常见算法（听力图对比、不同频率补偿不同（等响曲线）、WDRC、多通道补偿等）

3. 基于FFT的共振峰响度补偿

4. 补偿效果实验和仿真

5. 小结

1.语音信号时频特性

时域信号含义明确，清晰直观；频域反映响度感知特性；人耳对相位不敏感；

短时平稳性、清音浊音、频率和幅值、人耳感知频率范围、声压级

2.语音信号的听觉感知特性

响度补偿 即借助人耳听觉特性，根据输入语音信号的频率和声强，调整助听器输出语音增益，提高输出语音的被感知能力。

响度补偿原理：

正常人和听障患者听力图对比；

声强和增益对比，（不同声强所需补偿增益不同），声强增益曲线。

传统多通道响度补偿：分子频带分别补偿

宽动态压缩算法（WDRC）：动态范围指的是听阈到不适阈。分段线性插值已得到全频域的增益系数；单通道的性能一般，语音舒适度不够，被多通道取代。

移频压缩，深度耳聋患者，高频段听力丧失，语音高频 相对应于辅音音节，提高语音可懂度，合理利用患者剩余听力范围，改善助听质量。

1. **频率伸缩算法**
2. **移动终端算法实现与测试**