

第一章 绪论

1.1 研究背景及意义

随着 3D 影像技术热潮的来临，人们对音频系统的要求也日益提升，房间冲击响应（扬声器-房间系统响应）均衡成为国内外研究的热点^{[1]-[7]}。衡量音频系统品质的一个客观标准是其幅频特性的平坦性，由于房间反射、组件品质等因素的影响，普通音频系统的幅频响应平坦特性难以得到保证，因此需要对音频系统进行均衡处理。最常见的音频重放系统是扬声器-房间系统，它的声学特性会受到扬声器的系统响应、房间反射等影响，是音频系统均衡研究的重点和难点。运用数字信号处理技术对房间冲击响应进行补偿和优化，不仅可以弥补扬声器组成部件的不足，提升低成本扬声器系统的性能，还可以消除房间属性对音频信号的影响，对声音进行修正，改善聆听感受，获得更加真实的声重放效果，使听者有不一样的听音感受。

所谓均衡是指在信号传输过程中，为了减少传输信号的线性失真，针对信道传递函数偏离其理想现状而进行的补偿过程^[8]。在许多声学的应用场景中都会使用到均衡技术，而在房间冲击响应均衡的特定语境中，我们所说的均衡指的是为了使信号无失真地通过扬声器-房间系统重放出来而对扬声器-房间系统传递函数所做的修改工作。封闭的声学空间可视为一个线性系统，其传输特性可以用冲击响应 $h(n)$ 或其频域形式 $H(e^{j\omega})$ 来表征。一般来说， $H(e^{j\omega})$ 也称为房间传输函数（room transfer function, RTF）。冲击响应 $h(n)$ 可以完整描述一个声音信号通过传输发生的变化，包括直达声和各个方向反射声的影响。显而易见，房间冲击响应会随着声源和听者的位置发生变化。

假设房间系统的冲击响应是 $h(n)$ ，房间系统的输入信号为 $x(n)$ ，均衡滤波器为 $f(n)$ ，如果：

$$x(n) * h(n) * f(n) = x(n) \quad (1-1)$$

则实现了信号的无失真重放，即对音频重放系统 $h(n)$ 做了均衡。从式(1-1)可以得出： $f(n) = h(n)^{-1}$ 。因此均衡的实质是要求解音频重放系统逆系统的过程。

房间冲击响应均衡包括单点均衡和多点均衡。其中，房间冲击响应单点均衡是指利用在房间某固定一点测量得到的房间冲击响应数据进行均衡滤波器的设计。因此仅在测量位置或其附近有均衡效果，在其他位置均衡效果不明显甚至会比未均衡的系统效果还要差^[9]。为了简单了解单点均衡对其他位置点房间冲击响应的影响，只考虑房间中的一阶反射， $h_1(n)$ 和 $h_2(n)$ 是某房间中两个不同位置 P_1 ， P_2 的冲击响应，其中

$$h_1(n) = \delta(n) + \alpha_2 \delta(n-1); \quad |\alpha_2| < 1 \quad (1-2)$$

$$h_2(n) = \delta(n) + \beta_2 \delta(n-1); \quad \alpha_2 \neq \beta_2 \quad (1-3)$$

根据 P_1 点的房间冲击响应 $h_1(n)$ 设计的均衡滤波器为 $h_{eq}(n)$ ，

$$h_{eq}(n) = (-\alpha_2)^n u(n) \quad (1-4)$$

因为

$$h_{eq}(n) * h_1(n) = \delta(n) \quad (1-5)$$

对于位置点 P_2 ，均衡后的房间冲击响应为：

$$h_{eq}(n) * h_2(n) = \delta(n) - (\alpha_2 - \beta_2)(-\alpha_2)^{n-1} u(n-1) \quad (1-6)$$

对于位置点 P_2 ，均衡后的房间冲击响应误差为：

$$\begin{aligned} \varepsilon &= \frac{1}{I} \sum_{n=0}^{I-1} e^2(n) \\ &= \frac{1}{I} \sum_{n=0}^{I-1} (\delta(n) - h_{eq}(n) * h_2(n))^2 \\ &= \frac{(\alpha_2 - \beta_2)^2}{I} \sum_{n=0}^{I-1} (-\alpha_2)^{2n-2} \end{aligned} \quad (1-7)$$

很明显位置点 P_2 的房间冲击响应并未得到均衡，因为误差 ε 是大于零的，如图 1-1 所示。随着 $|\alpha_2 - \beta_2|$ 增大，即随着 P_1 点与 P_2 点房间冲击响应之间的差异越来越大，应用于 P_1 点的均衡滤波器在 P_2 点产生的均衡误差会逐渐增大，甚至会出现比未进行均衡处理效果更差的情况。因此不能就一点的房间冲击响应数据进行整个房间的均衡处理，必须进行多点均衡。

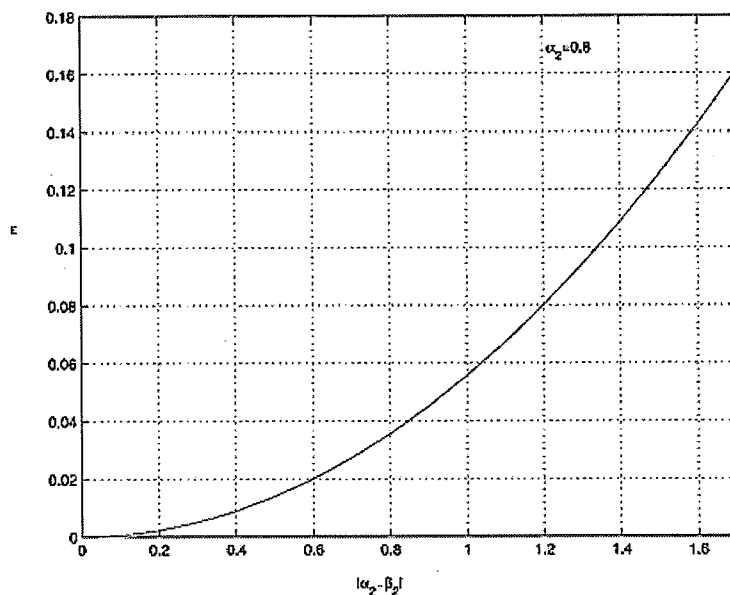


图 1-1 随着 P_1 、 P_2 房间冲击响应差异变化 P_2 点的均衡误差

1.2 研究现状

1.2.1 房间冲击响应单点均衡

1982 年, Jensen 提出了全数字前置放大器的设计原则^[10]。其中就已经包含了均衡器和前置放大两重意思, 只是由于当时所有的数字均衡滤波器技术都还没有被发明, 因此 Jensen 在文中只阐述了分频段进行均衡的思想, 并没有具体研究和说明均衡的技术。

1985 年, Mourjopoulos 等人提出用 FIR 滤波器均衡音频系统^{[11]-[14]}, 1988 年, Mourjopoulos 详细地说明了设计 FIR 滤波器进行音频系统均衡的过程。Mourjopoulos 最早提出设计扬声器系统 FIR 均衡滤波器的时候, 使用的是最小均方误差逼近法, 为了提高对混合相位系统的均衡性能, 对期望信号引入了纯延时 k 。

1991 年, Hawksford 等人首次提出了设计 IIR 均衡滤波器的想法^{[15]-[16]}。IIR 滤波器能在阶数相对较小的情况下获得与 FIR 滤波器一样好的幅度均衡效果。IIR 数字滤波不具有理想的线性相位响应。通过级联一个全通系统进行相位网络校正, IIR 均衡滤波器可以获得近似的线性相位。1996 年, Hawksford 又进一步提出了用 Matlab 软件设计 IIR 均衡滤波器的具体步骤。Hawksford 等人指出在求均衡滤波器之前, 对扬声器响应的预处理也是扬声器均衡中重要的一环。这些预

处理操作包括了扬声器响应的时域修形、频域修形以及目标函数的设计等。IIR 滤波器由于其高效的性能而在扬声器均衡技术中被广泛使用,也吸引了更多的人去研究,不同的 IIR 均衡器设计算法其设计思想和 Hawksford 等人大体上是一致的,只是求解 IIR 滤波器系数的方法不同而已。

1996 年, Matti 等人首次提出了运用一种基于频域的双线性保角映射的方法来设计均衡滤波器,即是所谓的频率弯折滤波器 (warped filters), 又称拉格里滤波器 (Laguerre filters) [17]-[18]。频率弯折滤波器的设计思想是将滤波器中的单位延时单元 z^{-1} 用一个一阶全通系统来代替, 其中, 全通系统中的零点 λ 被称为频率弯折参数。1995 年, 斯坦福大学的 Smith 教授和美国航天局艾姆斯研究中心的 Abel 研究员研究了频率弯折参数 λ 和采样频率的关系, 使得对于 Bark 和 ERB 刻度都有特定的弯折参数 λ 与之相符合。得到频率弯折参数 λ 后, 频率弯折滤波器的设计是一个很直观的过程: 先将扬声器的频率响应通过上述频率变换映射到所需要的非线性空间, 然后采用 FIR 或 IIR 滤波器设计方法设计已经经过弯折的扬声器频率响应, 最后再做反映射从而得到所需要的均衡滤波器。

2006 年, Matti 等人提出了利用 Kautz 滤波器均衡扬声器系统的算法 [19]-[21]。Kautz 滤波器同样是在非线性刻度上设计的均衡滤波器, 它是对频率弯折滤波器的发展, 频率弯折滤波器只能通过一个频率弯折系统来调整保角映射的非线性度, 而 Kautz 滤波器控制频率弯折非线性度的则是一组参数 $\{z_j\}$ 。 $\{z_j\}$ 实际上是 Kautz 滤波器的一组极点, 通过调整极点的放置位置, 可以调整不同频段的均衡权重, 从而实现非线性刻度下的均衡。当 $\{z_j\}$ 参数组都取同一值时, Kautz 滤波器便退化成了频域弯折滤波器。显然, Kautz 滤波器比频率弯折滤波器的自由度更高, 结构更加灵活, 但计算复杂度也更高。

同样是 2006 年前后, Ramos 和 Lopez 提出了利用双二阶 IIR 滤波器级联结构对扬声器系统进行均衡的算法 [22]-[26]。算法利用先进的滤波器级联结构, 通过每一级双二阶滤波器的最小相位来保证均衡器的稳定性。此外, 通过控制每一级滤波器极点的位置来控制滤波器的设计刻度。如果极点是均匀分布的, 那么均衡滤波器就是在线性刻度下设计的, 低频和高频段的权重一致; 如果极点是对数分布的, 那么均衡滤波器就是在对数刻度下设计的, 低频的权重更大一些。Ramos 和 Lopez 放置极点位置的方法是: 在偏离频谱平衡位置最远的峰或谷处布置极点。每一级双二阶 IIR 滤波器都是一个峰谷滤波器, 可以对应实现一个频谱峰谷的均

衡功能，因此这样布置极点的方法能够使均衡滤波器在低阶时效率最高。然而，由于算法中包含了大量随机搜索和适配的过程，因此计算复杂度比上述所有算法都要高。

2013 年，Bank 提出了定极点组并联均衡滤波器（fixed-pole parallel filter）的设计算法^{[27]-[30]}。Bank 从理论上证明了定极点组并联均衡滤波器和 Kautz 滤波器是等价的，但是求解滤波器系数的方法不同。Bank 为了降低 Romas-Lopez 算法的计算复杂度提出了定极点组形式的并联均衡滤波器，同样是级联的结构，但由于极点组固定，因此计算复杂度下降，但均衡效率也随之下降。如何设计极点组是这种级联型均衡滤波器结构避不开的研究课题。

1.2.2 房间冲击响应多点均衡

1982 年，Mourjopoulos、Clarkson 和 Hammond 提出了用最小二乘方法来均衡混合相位的房间冲击响应^[31]。最小相位信号的求逆比较容易理解，但是对于非最小相位信号而言，对其求逆或是解卷积就相对复杂。Mourjopoulos 等人对应用于有限长序列求逆的两种技术：同态技术和最小二乘技术进行了比较。应用同态技术求取混合相位信号的逆，需要先将信号分解为最小相位部分和全通部分。而最小二乘方法可以直接用来求取有限长序列的逆。

1988 年，Miyoshi 和 Kaneda 提出了基于 MINT (multiple-input/multiple-output inverse theorem) 求取房间冲击响应多点均衡精确解的方法^[32]。一般来说，房间冲击响应都是非最小相位的，利用最小二乘等方法只能够去逼近而无法得到精确的逆系统系数。该方法通过增加扬声器或麦克风的个数实现增加音频信号传输通道数，再利用矩阵代数知识得到由 FIR 滤波器构成的均衡器系数的精确解。

1989 年，Elliot 和 Nelson 利用自适应方法实现房间冲击响应多点均衡^[33]。在该方法中，作者采用梯度下降算法，不断更新均衡滤波器系数，使各点输入信号与参考信号之间均方误差的和达到最小。其中，输入信号为在各点测得的受扬声器-房间特性影响后的信号，参考信号为加入适当时延的原始电信号。这种方法可以达到仅采用一个均衡滤波器实现房间冲击响应多点均衡的目的。

1997 年，Haneda、Makino 和 Kaneda 提出通过建立房间传输函数 (RTF) 的共极点模型实现房间冲击响应多点均衡滤波器的设计^[34]。他们认为各个位置点的房间冲击响应的共同极点对应整个房间的谐振频率而与位置无关。通过测量得

到多个点的房间冲击响应并建立共极点模型,针对此模型进行整个房间均衡滤波器的设计,虽然不能均衡所有点的频响特性,但是可以覆盖这些频率特性中的共同峰值谷值,得到与位置无关的均衡效果。

2004 年, Sunil Bharitkar 和 Chris Kyriakakis 利用赛曼映射 (Sammon Map) 形象展示房间冲击响应并将其应用于房间冲击响应多点均衡^[35]。赛曼映射主要用于房间冲击响应建模,因为通过聚类分析进行房间冲击响应建模时,往往不能准确确定聚类个数,赛曼映射针对此问题能够提供一个有效的方法。另外,赛曼映射能体现出房间内多点均衡结果的一致性。再加上该映射方法计算复杂度较低,因此该方法可以取得理想的房间冲击响应多点均衡结果。

2007 年, Sascha Spors 和 Herbert Buchner 提出可以利用自适应方法进行房间内多通道系统的均衡^[36]。一般情况下房间内多通道间存在耦合问题,算法利用奇异值分解解决去耦合问题,但是由于针对变化的房间冲击响应进行广义奇异值分解矩阵的连续自适应运算计算复杂度很高,所以很难进行单纯的自适应处理。该方法通过将数学方法与声波传播的物理描述进行结合,实现了具有很好去耦合特性的房间多通道系统的自适应。该方法称为:波域自适应滤波 (wave domain adaptive filtering)。

2012 年, Alberto Carini、Stefania Cecchi 与 Francesco Piazza 等人提出在频域实现房间冲击响应多点均衡的方法^{[37]-[40]}。该方法主要思路是利用多点的房间传输函数进行房间冲击响应建模,然后根据此模型进行均衡滤波器的设计。作者采用多种频域算法替代原有时域的基于模糊 C 均值聚类算法进行房间冲击响应建模。同时在频域对房间冲击响应模型进行均衡中采用频域弯折技术,改善了低频段的均衡效果。与其他在时域实现的房间冲击响应多点均衡方法相比,该方法大大降低了计算复杂度并且取得了理想的均衡效果。

1.3 本文主要研究内容

本文针对基于人耳听觉特性的房间冲击响应均衡算法展开,如图 1-2 所示主要分为五部分:(1)预处理,主要包括房间冲击响应的测量、平滑与建模;(2)主要的房间冲击响应均衡算法总结;(3)人耳听觉特性,主要包括响度及等响曲线、掩蔽效应及临界频带宽度;(4)基于响度的房间冲击响应均衡;(5)基于临界频带宽度的房间冲击响应均衡。

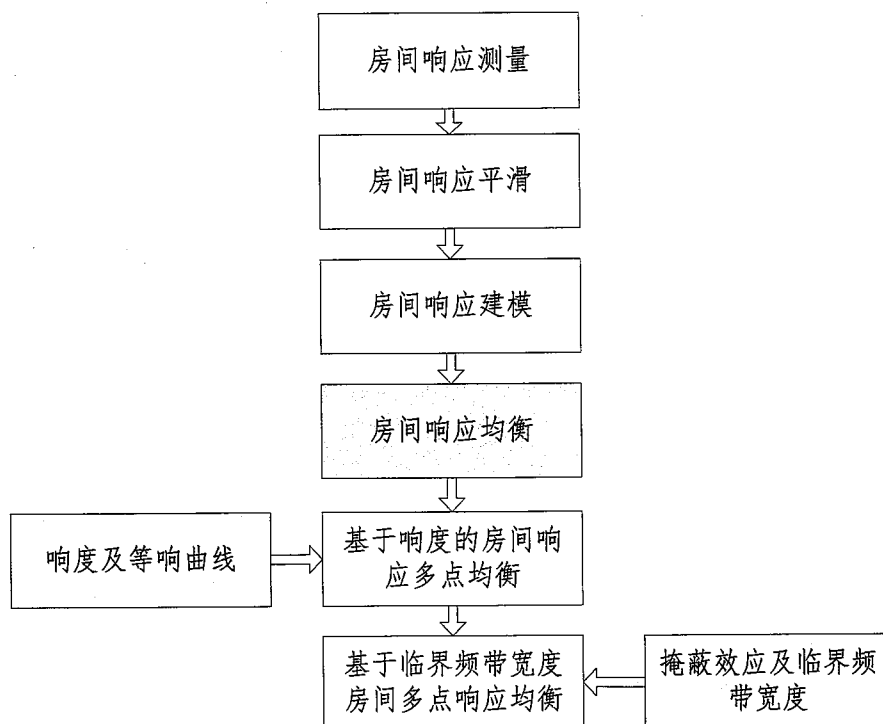


图 1-2 论文组织结构

论文组织结构为：

第一章：绪论。课题相关内容的介绍。具体包括研究的背景，课题所涉及技术的发展现状以及课题的关键内容，最后介绍了论文的组织结构。

第二章：单点均衡技术及多点均衡技术总结。介绍了多种单点均衡和多点均衡算法的原理和实现过程。

第三章：人耳听觉特性。介绍了包括人耳听觉结构、掩蔽效应、临界频带宽度和等响曲线等在内的听觉特性相关概念与原理。

第四章：预处理。本章主要介绍正式进行房间冲击响应均衡之前的预处理过程，包括：房间冲击响应测量、房间冲击响应平滑和房间冲击响应建模。

第五章：基于响度度量的房间冲击响应均衡。设计响度转换滤波器，将误差由幅度转化为响度，实现以响度为度量的房间冲击响应均衡。

第六章：基于临界频带宽度的房间冲击响应均衡。利用临界频带宽度对第五章提出的响度转化滤波器进行优化与改进，实现基于临界频带宽度的房间冲击响应均衡。

第七章：总结与展望。总结论文工作，对课题中的不足以及未来可以补充的工作分析。

第二章 房间冲击响应单点均衡与多点均衡算法

正如前一章节所述均衡的概念与意义,假设房间系统响应的系统函数是 $h(n)$, 房间系统的输入信号为 $x(n)$, 均衡滤波器为 $f(n)$, 如果:

$$x(n) * h(n) * f(n) = x(n) \quad (2-1)$$

则实现了信号的无失真重放, 即对房间系统响应 $h(n)$ 做了均衡。从上式可以得出: $f(n) = h(n)^{-1}$, 因此均衡的实质是要求房间传输系统的逆系统。房间冲击响应均衡又可以分为单点均衡和多点均衡。单点均衡仅利用房间内一个点的房间冲击响应数据进行均衡, 多点均衡是在单点均衡基础上为了扩大均衡有效区域进行的改进与优化。本章主要总结一下绪论中提到的常用的几种单点均衡与多点均衡算法, 为后边的研究奠定基础。

2.1 房间冲击响应单点均衡算法

2.1.1 FIR 均衡滤波器的设计

FIR 滤波器的设计方法有很多种, 而应用于房间冲击响应均衡系统的 FIR 均衡滤波器设计的时候, 用的是最小均方误差逼近法, 代价函数如下式所示:

$$E = \sum_{n=0}^{N-1} e^2(n) = \sum_{n=0}^{N-1} [d(n) - y(n)]^2 \quad (2-2)$$

其中,

$$y(n) = x(n) * f(n) * h(n) \quad (2-3)$$

$d(n)$ 是系统参考信号, 为测量房间冲击响应时采用的 MLS 信号并带有一段合适纯延时, $y(n)$ 是系统输出函数, $x(n)$ 是输入信号即测量房间冲击响应时采用的 MLS 信号, $f(n)$ 是设计的均衡滤波器系统响应, $h(n)$ 是测量得到的房间系统的传递函数。利用 LMS 算法或者其他算法求取最佳的均衡滤波器系数, 实现均衡器设计。

2.1.2 IIR 均衡滤波器的设计

IIR 滤波器能在滤波器阶数相对 FIR 滤波器较低的情况下获得与 FIR 滤波器一样好的幅度均衡效果, 但 IIR 数字滤波不具有理想的线性相位响应, 通过增加一些计算复杂度, IIR 均衡滤波器可以获得近似的线性相位。一般情况下, 房间

冲击响应多为非最小相位的，可以将房间系统分解为最小相位系统和全通系统，再进一步进行房间冲击响应均衡。基于 IIR 滤波器的房间冲击响应均衡步骤如下：

(1) 房间冲击响应系统的自回归-滑动平均模型（ARMA 模型）为：

$$H(z) = \frac{\sum_{i=0}^{N-1} b_i z^{-i}}{1 - \sum_{j=0}^{M-1} a_j z^{-j}} \quad (2-4)$$

(2) 将房间冲击响应的频率响应化为最小相位和全通系统：

$$\begin{aligned} H(z) &= \prod \frac{(z - n_m)(z - n_e)}{z - d} \\ &= \prod \frac{(z - n_m)(zn_e^* - 1)}{z - d} \cdot \frac{z - n_e}{zn_e^* - 1} \\ &= H_{\min}(z) \cdot H_{\text{ap}}(z) \end{aligned} \quad (2-5)$$

式 (2-5) 中 n_m 、 n_e 分别表示单位圆内、外的零点， d 表示极点。

(3) 幅度均衡：由于最小相位系统的逆是稳定的，因此求解幅度均衡滤波器只需将分离得到的最小相位系统直接求倒数即可。

(4) 相位均衡：全通系统的逆就是房间冲击响应的相位均衡滤波器。全通系统的幅频响应恒等于 1，只有相位部分 $e^{j\omega\sigma}$ ，可以看出，相位均衡的方法是 $e^{j\omega\sigma} \cdot e^{-j\omega\sigma} = 1$ ，为此可以引入纯延时 τ ，相位均衡变为 $e^{j\omega\sigma} \cdot e^{j\omega(\tau-\sigma)} = e^{j\omega\tau}$ ，实现房间冲击响应的相位均衡。

2.1.3 频率弯折均衡滤波器的设计

FIR 和 IIR 均衡滤波器的设计过程都是在线性刻度上进行的，低频和高频的分辨率一样，可是这与人耳的听觉特性不符，相对于高频段来说，人耳对低频的分辨率及敏感程度要更高，也就是说人耳对频率的敏感程度并不是符合线性规律。因此许多学者就想在一个更符合人耳听觉特性的刻度上进行均衡滤波器的设计。频率弯折滤波器的设计思想是将滤波器中的单位延时单元 z^{-1} 用一个一阶全通系统来代替：

$$D(z) = \frac{z^{-1} - \lambda}{1 - \lambda z^{-1}} \quad (2-6)$$

其中， λ 是弯折参数，控制着滤波器的非线性程度。如果 λ 是正数， λ 越大，低频段的分辨率在弯折域呈现地越高，如果 λ 是负数， λ 绝对值越大，高频段的分

辨率在弯折域呈现地越高。因此频率弯折滤波器的设计过程是：先将扬声器的频率响应通过上述频率变换映射到所需要的非线性空间，然后此前 FIR 和 IIR 均衡滤波器的设计方法都可以应用到已经弯折的扬声器频率响应上面，最后再做反射从而得到所需要的均衡滤波器。

2.1.4 Kautz 均衡滤波器的设计

基于 Kautz 滤波器的房间冲击响应算法是对频域弯折均衡滤波器的优化与改进，在频域弯折滤波器中，只有 λ 一个参数控制着频率的弯折程度，无法实现更加复杂的频率映射。Kautz 滤波器的表达式如下：

$$\begin{aligned} H(z) &= \sum_{i=0}^N \omega_i G_i(z) \\ &= \sum_{i=0}^N \omega_i \left(\frac{\sqrt{1 - z_i z_i^*}}{1 - z_i z^{-1}} \prod_{j=0}^{i-1} \frac{z^{-1} - z_j^*}{1 - z_j z^{-1}} \right) \end{aligned} \quad (2-7)$$

由式(2-7)中可以看出 Kautz 滤波器是一组定极点 (fixed-pole) 滤波器的级联。当 $z_j = 0$ 时，Kautz 滤波器转换为普通的 FIR 滤波器；当 $z_j = \lambda$ 时，Kautz 滤波器则转换成频域弯折滤波器。利用 Kautz 滤波器进行房间冲击响应均衡的步骤为：

(1) 确定极点组 $\{z_j\}_{j=0}^N$

极点组的固定设置法是指在频率轴上按照对数刻度来放置极点，这种极点组的设置方法与响应曲线的特征没有任何关系，对于任何形状的频率响应曲线都可以使用。而自动搜索方法则是结合了待均衡的频率响应曲线的特征，通过迭代算法求出最佳极点组。

(2) 确定均衡滤波器系数 ω_i

极点组确定之后，即 $G_i(z)$ 已经确定，权系数 ω_i 就可以利用一般的 LMS 等算法求得。

2.1.5 双二阶 IIR 级联结构的均衡滤波器的设计

双二阶滤波器 (second-order sections, SOS) 链的均衡滤波器结构，其中链的每一级 SOS 都是最小相位的 IIR 峰谷滤波器 (peak filter)，其表达式如下：

$$H(s) = \frac{s^2 + \frac{A}{Q} \cdot s + 1}{s^2 + \frac{1}{A \cdot Q} \cdot s + 1} \quad (2-8)$$

式(2-8)中, A 是均衡滤波器的增益参数, Q 是均衡滤波器峰谷的品质因数。从上述式可以看出, 峰谷滤波器只要求频率 s , 增益 A 和品质因数 Q 三个参数, 而以前 IIR 滤波器的设计方法则需要延迟 k 、ARMA 模型的阶数 M 、 N 和滤波器系数 a 、 b 五个系数, 从这点上来看, SOS 链的方法与经典算法相比是有简化的。

算法设计的代价函数定义如下:

$$e_{\omega_{\log}} = \frac{1}{n_f - n_i + 1} \sum_{k=n_i}^{n_f} |W(\omega_k) \cdot e(\omega_k)_{(dB)}| \quad (2-9)$$

其中,

$$e(\omega_k)_{(dB)} = H_{\text{target}}(\omega_k)_{(dB)} - H_{\text{ls}}(\omega_k)_{(dB)} - H_{\text{eq}}(\omega_k)_{(dB)} \quad (2-10)$$

式(2-9)中, 代价函数衡量的是待均衡频率响应和目标响应在某个关注频段之间的差异, n_i 和 n_f 分别代表关注频段起始频率和终止频率。式(2-10)中, $H_{\text{target}}(\omega_k)_{(dB)}$ 表示均衡目标响应, $H_{\text{ls}}(\omega_k)_{(dB)}$ 表示房间系统的频率响应, $H_{\text{eq}}(\omega_k)_{(dB)}$ 表示均衡滤波器的频率响应。下标 dB 表示数学式是在 dB 的单位下进行运算的。 ω_k 表示 1/48 倍频程对数分布的频率空间, 因此 $e(\omega_k)_{(dB)}$ 又称对数误差函数, 它的评价体系是与人耳的听觉特性相符合的。 $W(\omega_k)$ 是加权系数, 引入加权系数可以让设计者选择不去均衡频谱上一些人耳并不会察觉到的尖锐峰谷。SOS 链结构的均衡算法的主要思想是寻找代价函数 $e(\omega_k)_{(dB)}$ 的峰谷然后对其进行补偿。考虑效率, 第一级 SOS 用于补偿最大的峰或谷, 第二级 SOS 用于补偿次大的峰或谷, 以此类推。每一级 SOS 的设计过程都是相同的。

2.1.6 定极点组均衡滤波器的设计

定极点组并联均衡滤波器的形式如下:

$$H(z^{-1}) = \sum_{k=1}^K c_k \frac{1}{1 - p_k z^{-1}} + \sum_{m=0}^M b_m z^{-m} \quad (2-11)$$

其中, c_k 和 b_m 为滤波器系数, p_k 为设计者给定的复数极点组。因此设计定极点并联滤波器只需要执行以下两个步骤:

- (1) 合理布置极点位置，例如在频率轴上均匀分布或对数分布等。
- (2) 利用最小均方误差逼近方法求解上式的滤波器系数。

式(2-11)可以改写成以下实数极点的形式：

$$H(z^{-1}) = \sum_{k=1}^K \frac{d_{k,0} + d_{k,1}z^{-1}}{1 + a_{k,1}z^{-1} + a_{k,2}z^{-2}} + \sum_{m=0}^M b_m z^{-m} \quad (2-12)$$

式中，

$$a_{k,1} = p_k + p_k^*, \quad a_{k,2} = |p_k|^2, \quad d_{k,0} = c_k + c_k^*, \quad d_{k,1} = -c_k p_k^* - c_k^* p_k \quad (2-13)$$

上标*表示复数共轭运算。设计如上式形式的定极点组并联滤波器同样执行上述两个步骤，只是求解滤波器系数的方法稍有不同。 $h_{\text{target}}(n)$ 表示均衡目标响应， $h_{\text{ls}}(n)$ 表示房间系统的频率响应， $h_{\text{eq}}(n)$ 表示均衡滤波器的频率响应。采用上式的形式，则有：

$$\begin{aligned} h_{\text{target}}(n) &= h_{\text{eq}}(n) * h_{\text{ls}}(n) \\ &= \sum_{k=1}^K (d_{k,0} u_k(n) * h_{\text{ls}}(n) + d_{k,1} u_k(n-1) * h_{\text{ls}}(n)) + \sum_{m=0}^M (b_m \delta(n-m) * h_{\text{ls}}(n)) \\ &= \sum_{k=1}^K (d_{k,0} s_k(n) + d_{k,1} s_k(n-1)) + \sum_{m=0}^M (b_m h_{\text{ls}}(n-m)) \end{aligned} \quad (2-14)$$

式中， $u_k(n)$ 是 $1/(1 + a_{k,1}z^{-1} + a_{k,2}z^{-2})$ 的冲击响应 $s_k(n) = u_k(n) * h_{\text{ls}}(n)$ 。滤波器系数的求解同样采用最小均方误差逼近法。

2.2 房间冲击响应多点均衡算法

2.2.1 基于 MINT 均衡滤波器的设计

由于房间冲击响应多为非最小相位的，利用最小二乘等算法求取房间冲击响应多点均衡滤波器的系数只能实现逼近而无法得到精确解。为了得到房间系统均衡滤波器的精确解，需要通过增加声源或者接收点个数方式来增加声音信号传输通道个数，以此得到均衡滤波器的精确解。以图 2-1 所示系统为例，

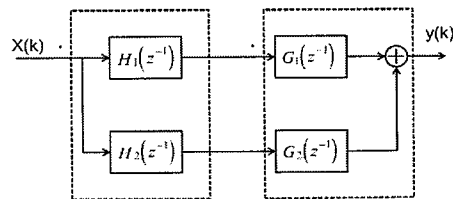


图 2-1 2 输入 1 输出系统示意图

为了实现对房间冲击响应精确求逆即进行均衡，必须满足：

$$D(z^{-1})=1=G_1(z^{-1})H_1(z^{-1})+G_2(z^{-1})H_2(z^{-1}) \quad (2-15)$$

式中， $G_1(z^{-1})$ 、 $G_2(z^{-1})$ 、 $H_1(z^{-1})$ 、 $H_2(z^{-1})$ 是 z^{-1} 的多项式， $H_1(z^{-1})$ 、 $H_2(z^{-1})$ 的解集满足以下 2 个特性：

(1) 只有在 $G_1(z^{-1})$ 、 $G_2(z^{-1})$ 没有共同零点的情况下，式(2-15)有解；

(2) 当 $H_1(z^{-1})$ 比 $G_2(z^{-1})$ 阶数低而且 $H_2(z^{-1})$ 比 $G_1(z^{-1})$ 阶数低时，式(2-15)有唯一解。

因此，存在一对 FIR 滤波器， $H_1(z^{-1})$ 、 $H_2(z^{-1})$ 能够实现 2 个输入 1 个输出的系统得到精确的均衡滤波器系数。

2.2.2 多点自适应均衡滤波器的设计

房间冲击响应单点均衡时，利用理想信号与输出信号之间误差构成代价函数，然后选择合适的自适应算法求得最佳均衡滤波器系数。将该思想应用于多点均衡时，需要考虑各点的误差，综合各点的误差构造代价函数，然后利用使代价函数最小的算法进行均衡滤波器最佳系数的求取，如图 2-2 所示。

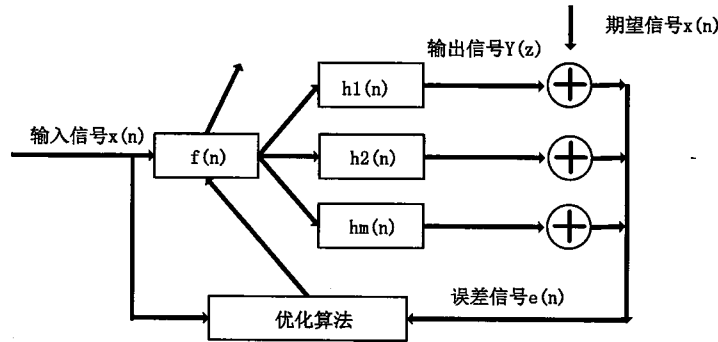


图 2-2 多点均衡中自适应算法框图

由上图我们可以得到以下关系：

$$r_1(n) = h_1(n) * x(n) \quad (2-16)$$

$$y_1(n) = r_1(n) * f(n) \quad (2-17)$$

每一个位置点的误差信号为：

$$e_1(n) = x(n) - y_1(n) \quad (2-18)$$

令：

$$R^T(n) = [r_1(n), r_2(n), r_3(n), \dots, r_m(n)] \quad (2-19)$$

$$e^T(n) = [e_1(n), e_2(n), e_3(n), \dots, e_m(n)] \quad (2-20)$$

新的代价函数为:

$$J = E\{e^T(n)e(n)\} \quad (2-21)$$

根据最大梯度算法, 我们可以得到均衡滤波器系数的迭代关系:

$$f(n+1) = f(n) + \mu \frac{\partial J}{\partial f(n)} = f(n) + \mu R^T(n)e(n) \quad (2-22)$$

2.2.3 基于共声学极点的均衡滤波器设计

尽管房间内各点的房间传输函数不尽相同, 但是, 所有传输函数的共同部分是其共振频率和品质因数 Q 。一般情况下, 房间传输特性的幅频特性中的峰值只与整个房间的共振频率相关而与位置点无关。房间传输函数中的零点部分引起房间内各点声学特性的差异, 因此, 在进行整个房间声学特性均衡时, 应该更加侧重于多个房间的传输函数的共同极点部分而忽略各个声学传输函数中的零点部分。基于上述分析, 利用共声学极点 (Common-Acoustical-Pole, CAP) 进行均衡滤波器设计首先进行房间声学特性共极点模型的建立, 如下式所示:

$$H_i(z) = \frac{B_i(z)}{A_{CAP}(z)} = \frac{\sum_{n=0}^N b_i(n)z^{-n}}{1 - \sum_{n=0}^P a_{CAP}(n)z^{-n}} \quad (2-23)$$

其中, $A_{CAP}(z)$ 与位置点无关, 是整个房间的共同极点, $B_i(z)$ 是第 i 点房间传输函数的零点部分。进而利用此共极点模型进行全极点 (Multiple-Point All-Pole, MPAP) 均衡, 如式(2-24):

$$F_{MPAP}(z) = A_{CAP}(z) = 1 - \sum_{n=0}^P a_{CAP}(n)z^{-n} \quad (2-24)$$

均衡滤波器的系数根据最小均方误差方法求得。

2.2.4 基于归一化误差的均衡滤波器设计

在进行房间冲击响应多点均衡时, 可以选择使多点的误差之和构成的代价函数最小的方法进行均衡滤波器最佳系数的求取。由于各点之间的误差信号差别较大, 个别点的误差可能会以很大的贡献影响整体误差。为了减小均衡结果对位置

的依赖性，应该首先对各点测量得到的频率响应进行归一化。 $H^m(f)$ 是在第 m 个位置点测量得到的房间冲击响应，为了得到房间冲击响应的归一化形式，可以计算幅频特性的均方根，如式 (2-25) 所示：

$$\hat{H}^m(f) = \frac{H^m(f)}{\sqrt{\frac{1}{f_{nq}} \int_0^{f_{nq}} |H^m(f)|^2 df}} \quad (2-25)$$

式中 f_{nq} 是奈奎斯特频率，多个点的误差构成的代价函数为：

$$E_{\text{mag}}(f) = \sum_{m=1}^M |e^m(f)|^2 = \sum_{m=1}^M \left(\left| a_{\text{mag}} - |\hat{H}^m(f)| Q_{\text{mag}}(f) \right|^2 \right) \quad (2-26)$$

式中 a_{mag} 是目标响应的幅频特性，它是一条直线， $Q_{\text{mag}}(f)$ 是待求的均衡滤波器频率响应。在滤波器系数求取过程中，为了防止出现病态，需要将代价函数进行正则化，如式(2-27)：

$$\begin{aligned} E_{\text{mag}}(f) &= \sum_{m=1}^M \left(|e^m(f)|^2 + \beta(f) |Q_{\text{mag}}(f)|^2 \right) \\ &= \sum_{m=1}^M \left(\left| a_{\text{mag}} - |\hat{H}^m(f)| Q_{\text{mag}}(f) \right|^2 + \beta(f) |Q_{\text{mag}}(f)|^2 \right) \end{aligned} \quad (2-27)$$

对式(2-27)进行求导，使偏导数为零得到最佳的均衡滤波器系数为：

$$|Q_{\text{mag}}(f)| = \frac{a_{\text{mag}} \sum_{m=1}^M |\hat{H}^m(f)|}{\sum_{m=1}^M |\hat{H}^m(f)|^2 + \beta(f)} \quad (2-28)$$

2.2.5 频域多点均衡滤波器的设计

频域多点均衡滤波器的设计步骤为：

- (1) 利用多个频域房间冲击响应数据建立频域房间冲击响应模型。
- (2) 将建立的频域房间冲击响应模型经弯折处理映射到非线性域。
- (3) 在该非线性域进行均衡滤波器的设计。
- (4) 把设计的均衡滤波器反映射回线性域，得到正常滤波器系数。

其中在非线性域进行均衡滤波器设计时，相关处理是：求取房间冲击响应模型 $H_p(e^{j\omega})$ 的自相关；利用 Levison-Durbin 算法进行 LPC 分析；对建立的 LPC 模型进行求逆。如图 2-3 所示：

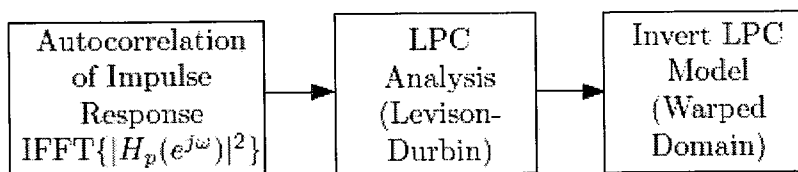


图 2-3 频域均衡滤波器设计算法框图

利用 LPC 分析建立全极点模型，然后直接求逆求得 FIR 均衡滤波器的系数。当然，均衡滤波器的系数是在弯折域的，想要得到线性域的均衡滤波器系数，还需要进行反映射处理。如图 2-4 所示：

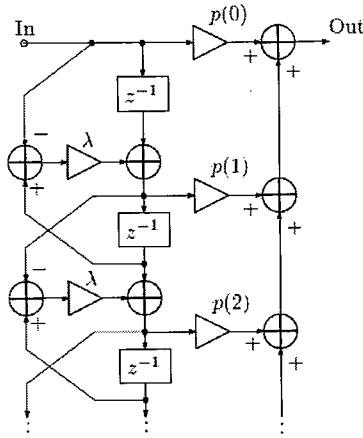


图 2-4 IIR unwarped 滤波器结构

在反映射滤波器结构中， λ 是将线性域映射到非线性域时利用的弯折参数数， $p(\cdot)$ 是弯折域的滤波器系数，经过此变换，弯折域的均衡滤波器就能转换为线性域的均衡滤波器。

2.3 本章总结

第二章介绍了多种房间冲击响应单点均衡及多点均衡算法，通过对这些算法的学习研究，可以对此领域有一个深入的理解，有利于研究的进行和深入。不同的均衡算法思路、原理和处理过程都是不同的，得到的结果也是各有优势与劣势。对他人算法的学习与分析有助于后边均衡工作的顺利进行，同时可以对比这些研究得到的结果与论文得到的实验结果，这对评价论文算法的好坏与算法的进一步优化与提高有很大的帮助。

第三章 人耳听觉特性

3.1 听觉系统与功能

从结构上来说，人类的听觉系统^[41]从外到内大体可以分成三部分：外耳、中耳与内耳，具体结构如图 3-1 所示。其中外耳主要由耳廓和外耳道构成，听小骨、鼓室和鼓膜等结构组成中耳，内耳除了包括前庭和半规管，还有由一卷成 2.75 圈的类似蜗牛壳的骨质管状结构组成，称为耳蜗。

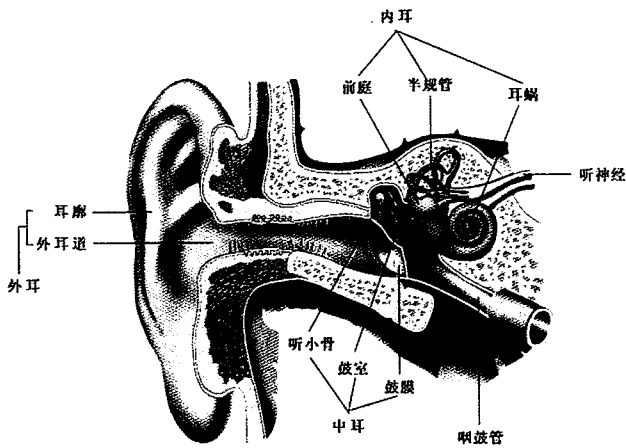


图 3-1 人类听觉器官结构示意图

在功能方面，隶属外耳的耳道部分的主要功能是声波的传输。耳廓的功能是声音信号的收集和低频信号的散射与反射，除此之外它还能与耳道相互耦合生成高频共振。中耳的主要功能是匹配声阻抗。内耳中最重要的结构是耳蜗。耳蜗的结构如图 3-2 所示

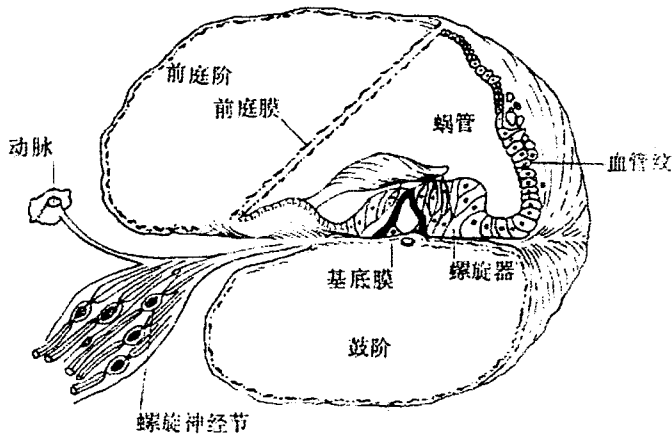


图 3-2 耳蜗结构示意图

内耳之所以是人耳听觉器官不可或缺的一部分，最重要的原因是它能够进行频率分析。内耳中的重要结构是基底膜，它是进行频率分析的重要结构。这是因为该结构的力学相关性质会随着它长度的延伸方向不断变化，因此基底膜上不同位置对于不同频率声音信号产生的响应也是不一样的。以平稳的正弦声音信号为例，该声音信号会引起基底膜的振动，而信号频率是振动幅度包络最大点在基底膜上出现位置的決定因素。这是因为底端附近的基底膜结构比较薄且比较窄，相反顶端附近的基底膜结构比较宽且比较厚。因此，高频段的聲音信号会在靠近基底膜底端的部分引起位移包络最大而其他位置位移包络较小的振动。相反的，低频段的聲音信号会在靠近基底膜顶端的部分引起位移包络最大而其他位置位移包络较小的振动。所以能得到结论：基底膜有频率-位置转换的功能，不同频率的聲音信号对应基底膜上不同的振动位移包络最大位置。图 3-3 是几个不同频率的正弦波在基底膜产生的位移包络以及 200 Hz 的正弦行波的示意图。

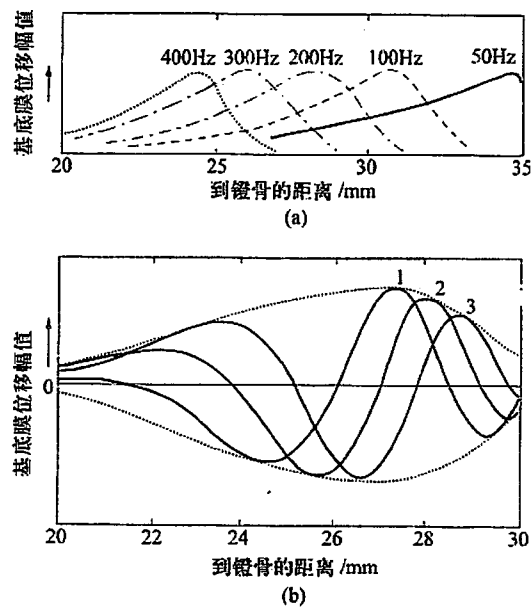


图 3-3 不同频率正弦波在基底膜产生的位移包络示意图

(a) 行波包络；(b) 基底膜上 200 Hz 的行波

若声音信号的频率高于 500 Hz，从基底膜顶端开始，最大位移包络的线性距离与频率的对数接近正比关系，但是相对带宽近似为恒量。正是由于这种相对带宽近似为恒量的性质，使得人耳听觉滤波器在分析处理声音信号时，在低频段有较高的分辨率，相反在高频段有较低的分辨率。当两个不同频率的纯音信号同时

作用于人耳听觉系统时,听觉系统对这两个纯音信号的处理要取决于这两个信号的频率。如果这两个声音信号的频率相差比较多,他们产生的最大的位移包络将出现在基底膜不同位置上,由于相互之间影响不大,基底膜能够分开处理两个声音信号;但是如果两个信号的频率越来越接近,他们在基底膜产生最大位移包络的位置也会越来越接近,基底膜对这两个声音信号产生的响应也会因此逐渐融合;如果两个信号的频率非常非常相近时,基底膜无法分辨他们在基底膜产生最大位移包络的位置,基底膜对这两个声音信号产生的响应融合在一起,不能分辨出单独的频率成分。内耳尤其是基底膜的频率分析性能与音调、听觉掩蔽效应等很多听觉现象是相关的。

3.2 掩蔽效应

听觉掩蔽效应^[42]是指在同一个环境中,某个声音信号会降低听者对另外一个声音信号的听力的现象,或者说一个声音信号对于人耳听觉系统的听阈因为另外一个更强的声音信号的存在而上升的现象。一个声音信号恰好被另一个声音信号掩蔽时,该声音信号的能量级被称为它的掩蔽阈值。如图 3-4 所示。

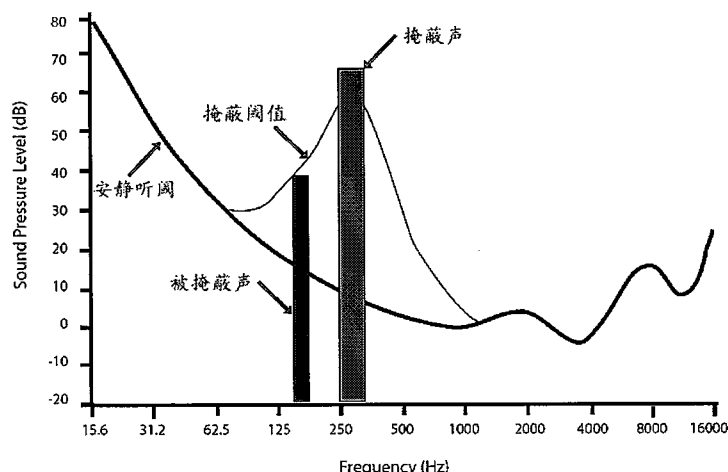


图 3-4 掩蔽阈值示意图

无论是语音增强、语音识别还是其他音频相关的信号处理中,听觉掩蔽效应都是在研究中不可忽略的人耳听觉主观感知特性。当人耳听觉系统对声音信号进行解析和处理时,无论是在时域还是在频域,处理结果都会受到分辨率或者动态范围存在上限的影响,因而导致听觉掩蔽现象。如果两个声音信号的时间和频率都特性接近只是响度有所不同,当人耳听觉系统处理这两个声音时,由于响度比

较高的声音信号的存在响度较低的声音信号就会受到影响,从而不容易被人耳听觉系统察觉。因为响度较高而影响了另一个声音信号的声音成为掩蔽音,因为响度低而受到影响的声音称为被掩蔽音。根据掩蔽音出现时间的先后掩蔽又分为前向、后向和同时掩蔽。如图 3-5 所示。

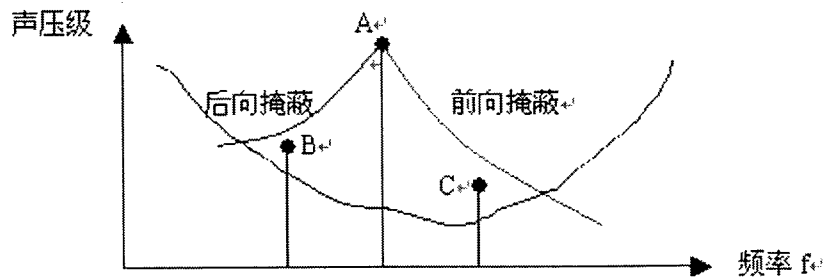


图 3-5 掩蔽效应示意图

3.3 听觉滤波器及临界频带宽度

早在 1940 年, Fletcher 等学者研究了带通噪声对纯音的掩蔽影响^[43], 实验结果表明, 噪声对纯音的掩蔽作用是有条件的: 以纯音频率 f 为中心的一定频带宽度为 Δf , 在 Δf 内的噪声对该纯音有掩蔽作用; 而频带宽度 Δf 以外的噪声对纯音是没有掩蔽作用的。 Δf 称为 f 的临界频带宽度 (critical band width), 利用基底膜的频率分析功能可以解释这个结果。基底膜上特定位置点是对某一特征频率的响应最大, 当声波的频率偏离特征频率的时候, 该点的响应减小。因而基底膜上的每一点可以等效为具有特定中心频率的带通滤波器。而整个基底膜 (严格来说是听觉系统) 可以等效为一系列具有连续特征频率的、互相重叠的带通滤波器, 能够对不同频率的声音信号产生不同的作用, 我们称之为听觉滤波器。

听觉滤波器有关频率听觉的分辨率是与听觉滤波器的频带宽度和形状有关的。如果听觉滤波器的有效带宽大于或等于噪声的频带宽度, 那么该噪声对在自己频带宽度内的纯音存在掩蔽作用; 如果听觉滤波器的有效带宽小于噪声的频带宽度, 那么该噪声仅仅对在听觉滤波器的有效带宽内的纯音存在掩蔽作用, 对听觉滤波器的有效带宽以外的纯音没有掩蔽作用。因而临界频带宽度^[44]给出了听觉滤波器带宽的一种估计。心理声学的实验结果表明, 临界频带宽度与其中心频率的关系^[45]为:

$$\Delta f_{CB} = 25 + 75 \times \left[1 + 1.4 \left(\frac{f}{\text{kHz}} \right)^2 \right]^{0.69} \quad (3-1)$$

由临界频带宽度可以引出一个与听觉有关的新频率标度：临界频带率，临界频带率的单位是 Bark^[46]。一个临界频带的宽度就是一个 Bark，另外 1 Bark 大约相当于基底膜上 1.3mm 的距离。Bark 标度 ν 与 f （单位 kHz）的关系为：

$$\nu = 13 \times \arctan \left(0.76 \frac{f}{\text{kHz}} \right) + 3.5 \times \arctan \left(\frac{f}{7.5 \text{ kHz}} \right)^2 \quad (3-2)$$

一些较新的试验结果表明，实际的听觉滤波器的形状是非对称的，但是为了方便分析，可以用具有连续特征频率的、互相重叠的带通滤波器来表示非对称的听觉滤波器。矩形滤波器的通带传输系数等于相应听觉滤波器的最大传输系数，而选择矩形滤波器的带宽使其传输白噪声时具有和听觉滤波器相同的输出功率。这种等效的听觉滤波器的带宽称为等效矩形带宽（Equivalent Rectangular Bandwidth, ERB）^{[48]-[51]}。对心理声学的实验结果分析表明，听觉滤波器的等效矩形带宽 ERB 与其中心频率 f 之间的关系^[47]可以用式(3-3)表示：

$$ERB = 24.7 \left(4.37 \frac{f}{\text{kHz}} + 1 \right) \quad (3-3)$$

其中 ERB 的单位为 Hz，中心频率 f 的单位为 kHz。

利用等效矩形滤波器的概念，可引入一个新的频率标度-ERB 数，它与频率 f （单位 kHz）的关系为：

$$ERBN = 21.4 \lg \left(1 + 4.37 \frac{f}{\text{kHz}} \right) \quad (3-4)$$

由上面的公式(3-4)可以验证，当频率在 500 Hz 以下时，临界频带宽度 Δf_{CB} 为 100 Hz 左右，当频率在 500 Hz 以上时，临界频带宽度 Δf_{CB} 是其中心频率的百分之二十左右。当频率大于 500 Hz 时，等效矩形带宽 ERB 与临界频带宽度 Δf_{CB} 结果是基本一致的，当频率小于 500 Hz 时，等效矩形带宽 ERB 是和频率相关的，这是等效矩形带宽 ERB 与临界频带宽度 Δf_{CB} 主要的不同。

无论是临界频带宽度 Δf_{CB} 还是等效矩形带宽 ERB 都是随频率而增加，也就是说听觉系统的频率分辨率是随着频率增加而下降的。因而临界频带宽度 Δf_{CB} 和等效矩形带宽 ERB 模型反映了听觉系统对频率的非均匀分辨率。声音的响度感觉、掩蔽效应等心理声学现象也都是与听觉滤波器密切相关的。

3.4 响度及等响曲线

响度^{[52]-[54]}是一种主观心理量，是人耳听觉器官在分析和处理声音信号的时候，为了表示人类对声音信号强度的主观感受而提取的具有声音信号幅度性质的物理量。通常情况下，若音频信号的频率不发生改变时，音频信号的响度会随着音频信号的强度变大而变大。但是，声音信号的响度是与声音信号的频率有关的物理量，具有同样声压的声音信号可能由于他们频率的不同而出现响度不相等的情况。这是因为人类的听觉器官十分复杂关键是还具有频率、时间等独有性质。所以，声音信号的响度不单单受信号强度的影响，还要考虑声音频率和瞬态时间特性，这些都是与响度变化相关的物理量。响度和响度级是两个不同的概念，响度是一个相对量，单位是“宋”(Sone)。而响度级是一个相对量，单位是“方”(Phon)，参考国际协议，定义 40 dB、1 kHz 纯音的响度为 1 Sone，同时 Phon 0 dB 声压级的 1 kHz 纯音的响度定义为 0 Phon，n dB 声级的 1 kHz 纯音的响度级就是 n Phon。将整个频段上全部响度级相同的声压级标注的点连接起来，就能得到对应的等响曲线，如图 3-6 所示：

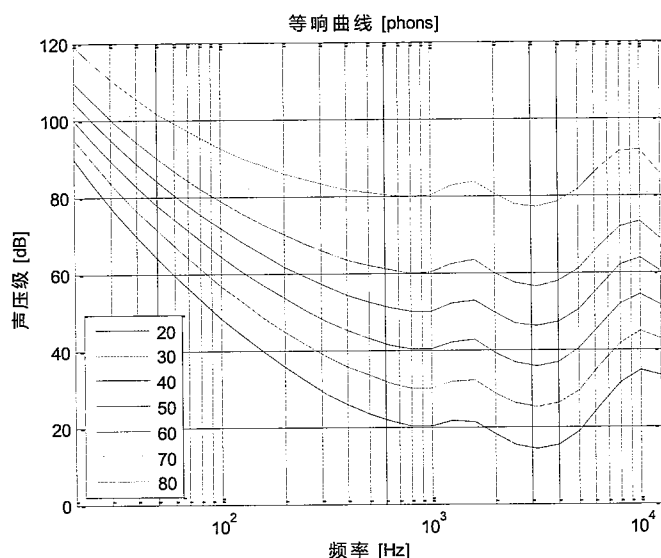


图 3-6 等响曲线（来自 ISO 226:2003 修正版）

由图 3-6 可以看到：

- (1) 声压级和响度级当且仅当信号频率为 1 kHz 的情况下才相同。
- (2) 听觉器官在低频段不够敏感，因此可能无法听见该频段内高声压信号。
- (3) 听觉器官在高频段足够灵敏，尤其频率为 3000 Hz 至 5000 Hz 的信号。

(4) 若声音信号的声压高于 100 dB, 各响度级的等响曲线变得越来越平直。

3.5 本章总结

第三章介绍了人耳听觉特性相关的几个关键概念: 掩蔽效应、等响曲线、临界频带宽度等。只有对这些概念、特性了解地比较清楚, 才能站在这个角度去优化均衡滤波器。人耳听觉特性相关知识在第四章、第五章、第六章都有应用, 是基于人耳听觉特性房间冲击响应均衡研究的基础。

第四章 预处理

在绪论中提到了在实施房间冲击响应多点均衡处理前，首先需要进行房间冲击响应测量、房间冲击响应平滑、房间冲击响应建模预处理。房间冲击响应测量是为了得到各点理想的房间传输函数，房间冲击响应平滑是为了减小均衡结果受时间、位置变化的影响，房间冲击响应建模是通过多点房间冲击响应数据建立能代表整个房间声学特性的模型，使单点均衡转化为多点均衡。本章将详细介绍三个预处理过程的原理、方法和实验结果，为房间冲击响应多点均衡的实现奠定基础。

4.1 房间冲击响应测量

扬声器或者房间声学响应的测量和其频率特性的确定，是声学或音频信号处理相关工作的重要过程。在本文中，理想的房间冲击响应的测量结果是房间冲击响应均衡的基础和前提。房间冲击响应测量工作的整体思路是：选择合适的源信号作为激励输送给整个房间系统，通过传声器记录输出信号，对输入信号和输出信号进行解相关运算求得房间传输函数。在整个测量工作中最重要的就是选择理想的源信号。Muller 和 Massaran 在文献^[55]中提到多种应用比较广泛的房间声学响应测量方法。其中，应用最大长度序列（maximum length sequence, MLS）信号^[56]和扫频信号^[57]作为源信号进行房间冲击响应测量的方法最受大家青睐。在本章中，只对基于 MLS 信号的房间冲击响应测量方法进行分析。

首先对 MLS 信号进行简单的介绍，MLS 信号是伪随机二进制序列，它拥有与白噪声类似的性质，但应用于房间冲击响应测量时，性质却优于白噪声信号。MLS 信号的优越性主要体现在，它易于生成且重复性好。具体来说，周期为 $L = 2^n - 1$ 的 MLS 信号，其中 n 为正整数，拥有以下特性：

$$\sum_{k=0}^{L-1} s(k) = -1 \quad (4-1)$$

$$R_{ss}(k) = \frac{1}{L} \sum_{n=0}^{L-1} s(n)s(n+k) = \begin{cases} 1 & k = 0, L, 2L, \dots \\ -\frac{1}{L} & k \neq 0, L, 2L, \dots \end{cases} \quad (4-2)$$

图 4-1 是 MLS 信号的自相关函数和频谱密度。从图中可以看出，MLS 信号拥有与白噪声类似的优良性质，尤其是自相关函数，可以用来进行解相关处理求出房间冲击响应。

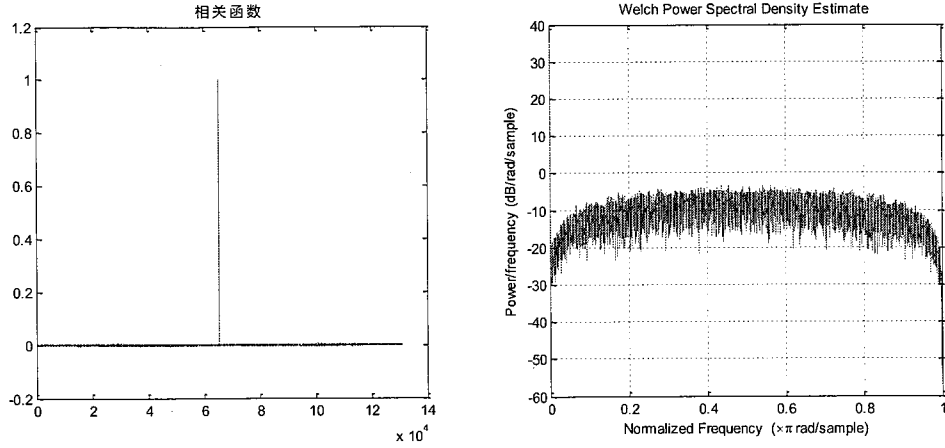


图 4-1 MLS 信号的自相关函数及频谱密度

在房间冲击响应测量中， $x(t)$ 是输入信号， $h(t)$ 是待测的房间冲击响应， $y(t)$ 是输出信号，三者满足以下关系：

$$y(t) = \int_{-\infty}^{\infty} x(t-k)h(k)dk \quad (4-3)$$

对式(4-3)等式两端与输入信号 $x(t)$ 作互相关运算，得到

$$\begin{aligned} R_{yx}(\tau) &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} \int_{-\infty}^{\infty} x(t+\tau-k)h(k)x(t)dkdt \\ &= \int_{-\infty}^{\infty} R_{xx}(\tau-k)h(k)dk \end{aligned} \quad (4-4)$$

由式(4-4)容易知道，若 $R_{xx}(\tau-k) = \delta(\tau-k)$ 那我们就可以得到：

$$R_{yx}(\tau) = h(\tau) \quad (4-5)$$

即：若输入信号的自相关函数是冲击函数，那么输出信号与输入信号的互相关等于系统的冲击响应。因此，将 MLS 信号作为输入信号 $x(n)$ 测量房间冲击响应 $h(n)$ ，用得到的输出信号与输入信号 $x(n)$ 进行互相关运算，我们可以得到：

$$y(n) = x(n) * h(n) = \sum_{k=0}^{L-1} s(n-k)h(k) \quad (4-6)$$

$$\begin{aligned}
R_{yx}(\tau) &= \sum_{k=0}^{L-1} x(n)y(n+\tau) \\
&= \sum_{k=0}^{L-1} \sum_{n=0}^{L-1} x(n)x(n+\tau-k)h(k) \\
&= \sum_{k=0}^{L-1} R_{xx}(\tau-k)h(k) \\
&= h(\tau) - \frac{1}{L} \sum_{k=1}^L h(\tau-k)
\end{aligned} \tag{4-7}$$

在式(4-7)中,当 L 足够大时,输入信号与输出信号的互相关结果就趋近于我们所需的房间冲击响应 $h(n)$ 。在实际的测量过程中,我们选择合适周期的 MLS 信号即可得到理想的房间冲击响应。值得注意的是 MLS 信号是周期信号,如果 MLS 信号的周期小于房间冲击响应的长度,那么将发生混叠现象,影响最终的测量结果。为了避免混叠现象的发生,应该选择足够大的周期 L 。

在正式测量之前首先要进行仿真工作,确定解相关运算的正确性。仿真思路是:让 MLS 信号经过一个已知的扬声器系统 $h(n)$ 得到输出信号 $y(n)$,在利用上述的解相关运算求得 $h'(n)$ 。对比 $h(n)$ 与 $h'(n)$,结果如图 4-2 所示,可以确定此解相关运算的有效性。

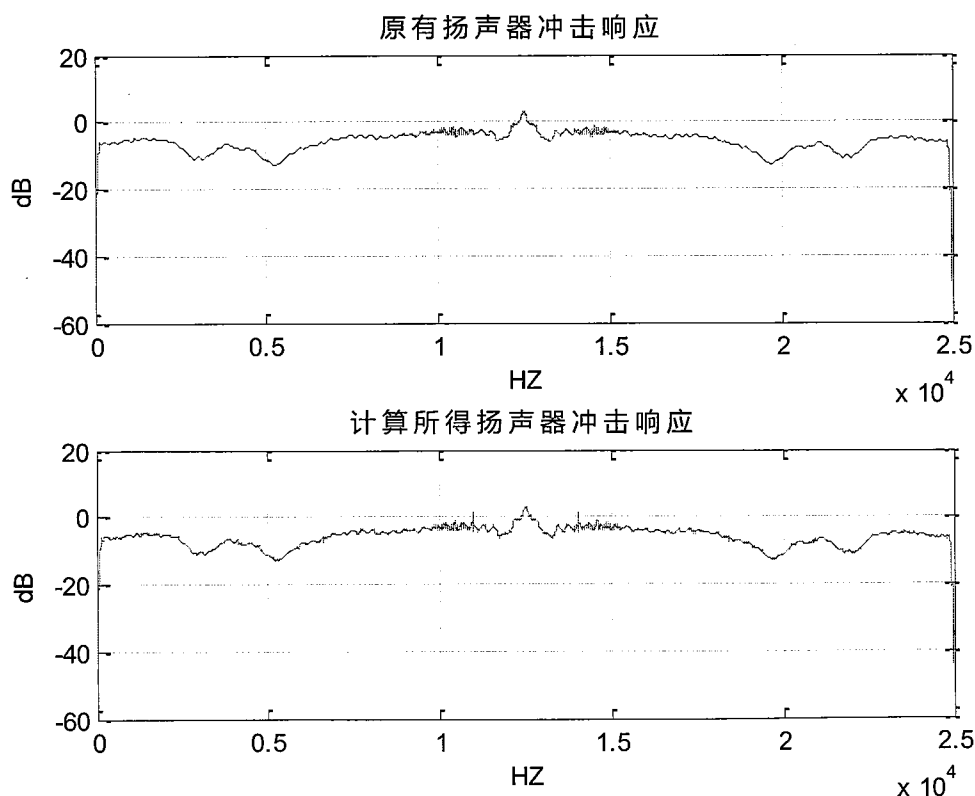


图 4-2 利用 MLS 信号进行测量仿真结果

4.2 房间冲击响应平滑

利用在房间内多点测量得到的房间传输函数进行均衡器设计的过程中,考虑到人耳听觉特性和实践经验,有一个步骤至关重要,那就是对测量得到的数据进行频率响应平滑。房间冲击响应会随着时间和位置发生很大变化,为了得到理想的均衡效果,选择适当的均衡滤波器长度和预先对房间冲击响应进行平滑,是解决均衡效果受位置影响问题的重要途径。研究表明,人耳对声音变化的感知不仅与房间冲击响应幅频特性的波动有关,还受相位响应的影响。因此,为了使房间冲击响应均衡结果受位置变化的影响尽可能小,房间冲击响应的相位响应也需要进行平滑处理。前边的章节提到,人耳听觉系统对声音频率的分辨率不是线性的,人耳在低频段的频率分辨率要高于在高频段的分辨率。因此,无论是平滑处理还是均衡处理,都要更加重视低频段的修正。就这一点而言,平滑处理的频带在低频段应该设计较窄的频带宽度相反在高频段就可以设计较宽的频带。

倍频程是在人的可听频率范围内,将声音的频谱进行一定规则的集中,变成有限的几个频点对应的强度,这样描述比较起来容易,是一种公约的描述形式。分析频谱时,对于连续谱而言,分析某频率点上的声功率是没有意义的,因此有必要统计某一频带内的声功率。对于频带划分,倍频程和 1/3 倍频程是常用的划分方法之一,它们都是相对恒定带宽,例如 1/3 倍频程的带宽是中心频率的 23%。

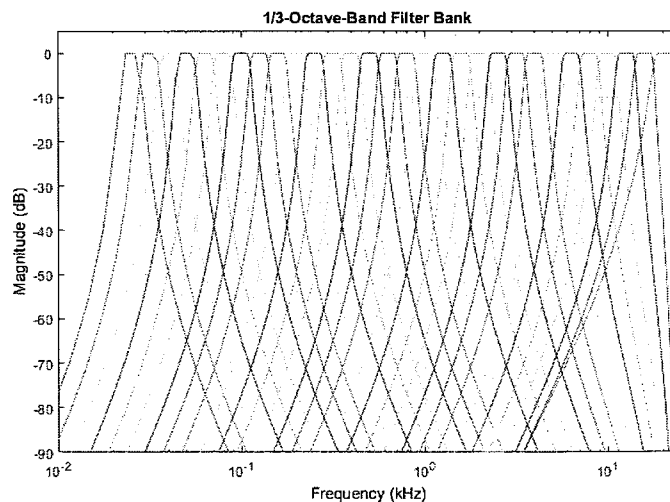


图 4-3 1/3 倍频程滤波器带宽

本文采用 1/3 倍频程的复数平滑(也称为频率响应平滑)方法进行房间冲击响应的平滑。使用 1/3 倍频程主要是因为人耳对声音的感觉,其频率分辨能力不

是单一频率，而是频带，而 1/3 倍频程曾经被认为是比较符合人耳特性的频带划分方法，不过现在心理声学里提出了 Critical Band 这么个频带划分方法，虽然更符合人耳特性，但 1/3 倍频程仍在广泛使用，图 4-3 展示了 1/3 倍频程滤波器的带宽。1/3 倍频程频带宽度为：

$$\Delta f = f_U - f_L \quad (4-8)$$

其中，

$$f_U = 2^{\frac{1}{2 \times 3}} \times f_{\text{cnt}} \quad (4-9)$$

$$f_L = 2^{-\frac{1}{2 \times 3}} \times f_{\text{cnt}} \quad (4-10)$$

在一个频带内的平滑结果为：

$$\begin{aligned} H_{\text{mean12}} &= \frac{1}{f_2 - f_1} \int_{f_1}^{f_2} H(f) df \\ &= \frac{1}{f_2 - f_1} \left(\int_{f_1}^{f_2} \text{Re}[H(f)] df + j \int_{f_1}^{f_2} \text{Im}[H(f)] df \right) \end{aligned} \quad (4-11)$$

4.3 房间冲击响应建模

在第二章中提到过多种房间冲击响应多点均衡方法。整体来说可以分为两大类：（1）将各个点测量得到的房间冲击响应直接应用于均衡算法，通过使各个点的误差之和达到最小等思路求得均衡滤波器；（2）将各点测量得到的房间冲击响应进行分析、处理，建立一个房间冲击响应模型代表整个房间的声学特性，将房间冲击响应多点均衡问题转换为单点均衡，然后利用经典的单点均衡算法进行处理。本文采用第二种思路，房间冲击响应建模也成为重要的研究工作之一。

在一个普通房间即一个封闭的声学空间中，可以用冲击响应 $h(n)$ 或其频域形式 $H(e^{j\omega})$ 来描述声音信号从声源到接收点的变化。在房间的多个位置点上进行测量得到多个房间冲击响应数据，利用这些数据进行房间冲击响应建模，进而对此模型进行均衡处理，是实现房间冲击响应多点均衡的有效方法。在房间冲击响应建模中，可以对频域数据进行数学处理，也可以利用时域数据进行模糊 C 均值聚类分析。

4.3.1 房间冲击响应建模的数学方法

1、RMS (Root-Mean-Square) 方法

由于 RMS 方法简单，在许多建模相关算法中应用较多。整个房间的声学特性模型由各点房间冲击响应的均方根求得，如式(4-12)。应用大量的房间冲击响应数据平均操作可以得到房间冲击响应的整体趋势，去除点一些偏离较大数据的影响。

$$\left| H_{\text{rms}}(e^{j\omega}) \right| = \sqrt{\frac{1}{M} \sum_{k=1}^M \left| H_k(e^{j\omega}) \right|^2} \quad (4-12)$$

2、均值方法

房间冲击响应模型由测量得到的各点房间冲击响应的均值得到，如式(4-13)。与 RMS 相比，直接求平均能更有效地去除各个房间冲击响应中峰值与谷值的影响，得到一个更理想的代表整个房间属性的模型。

$$\left| H_{\text{mean}}(e^{j\omega}) \right| = \frac{1}{M} \sum_{k=1}^M \left| H_k(e^{j\omega}) \right| \quad (4-13)$$

3、最大-最小方法

建立的房间冲击响应模型由某频点各房间冲击响应中最大值和最小值决定，如式(4-14)。应用这种建模方法得到的房间冲击响应模型经均衡滤波器 $\left| H_{\text{mm}}(e^{j\omega}) \right|^{-1}$ 补偿后，其包络均集中在 0 dB 附近，尽管各点房间冲击响应之间的差异会缩小，但是建模结果仍对测量结果中偏离数据十分敏感。

$$\left| H_{\text{mm}}(e^{j\omega}) \right| = \frac{1}{2} \left(\max_k \left| H_k(e^{j\omega}) \right| - \min_k \left| H_k(e^{j\omega}) \right| \right) \quad (4-14)$$

4、中值方法

中值方法是一种源于非线性信号处理概念的简单有效的建模方法。房间冲击响应的整体趋势可以由测量得到各点房间冲击响应的中值提取，如式(4-15)所示，在上述的所有建模方法中，中值方法得到的建模结果受偏离整体趋势的数据影响最小。

$$\left| H_{\text{median}}(e^{j\omega}) \right| = \text{median}_k \left| H_k(e^{j\omega}) \right| \quad (4-15)$$

4.3.2 房间冲击响应建模的模式识别方法

除了利用数学方法还可以将模式识别技术应用于房间冲击响应建模。根据房间冲击响应的相似之处对在不同位置点上测量得到的房间冲击响应进行分类，进而建立一个能表征由于位置不同而产生差异的房间冲击响应一般化的模型。

首先简单介绍一下应用于房间冲击响应建模的聚类分析。从宏观角度来讲，聚类算法根据数据之间的相似程度将其分为有图心的几类或几簇。聚类分析算法一般会应用一个代价函数对数据进行分类，这个代价函数可以是各个数据到图心距离的平方和或者其他参数。通过聚类，拥有较高相似性的房间冲击响应分到一簇中，这些数据间的相似性用来确定图心，而图心则是代表本簇的模型。聚类分析算法有硬 C 均值聚类和模糊 C 均值聚类两种。在硬 C 均值聚类分析中，一个房间冲击响应只能属于一个簇，可以用二进制函数 $\mu_i(h_j)$ 表征此关系：房间冲击响应 h_j 属于或不属于簇 i ，其中 $\mu_i(h_j) \in \{0,1\}$ 。而在模糊 C 均值聚类分析中，一组数据可以同时属于几个簇，并且数据以不同程度属于一个簇， $\mu_i(h_j) \in [0,1]$ 。

在本文中采用的是模糊 C 均值聚类分析来建立房间冲击响应模型。如图 4-4 所示，

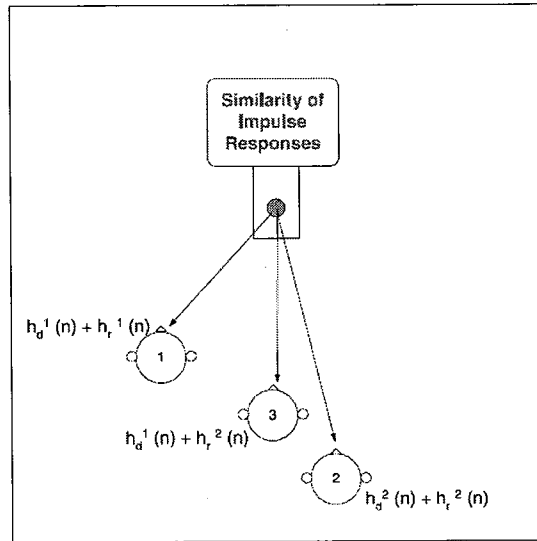


图 4-4 房间冲击响应聚类分析

图 4-4 中， P_1 ， P_2 ， P_3 三个位置点的房间冲击响应分别为：

$$h_1 = h_d^1(n) + h_r^1(n) \quad (4-16)$$

$$h_2 = h_d^2(n) + h_r^2(n) \quad (4-17)$$

$$h_3 = h_d^3(n) + h_r^3(n) \quad (4-18)$$

由于与声源距离相同， h_3 与 h_1 具有相同的直达声部分的房间冲击响应 $h_d^1(n)$ ，同样的，由于与声源的夹角相近， h_3 与 h_2 具有相同的反射声部分的房间冲击响应

$h_r^2(n)$ 。因此, 可以将 h_3 与 h_1 分为一类, 也可以将 h_3 与 h_2 分为一类。综上所述, 在房间冲击响应建模时, 模糊 C 均值聚类方法更为恰当。

利用模糊 C 均值聚类方法迭代求取每一个聚类的图心, 公式如式(4-19)、(4-20)、(4-21):

$$\hat{h}_i^* = \frac{\sum_{k=1}^M (\mu_i(h_k))^2 h(k)}{\sum_{k=1}^M (\mu_i(h_k))^2} \quad (4-19)$$

$$\mu_i(h_k) = \left[\sum_{j=1}^c \left(\frac{d_{ik}^2}{d_{jk}^2} \right) \right]^{-1} = \frac{\frac{1}{d_{ik}^2}}{\sum_{j=1}^c \frac{1}{d_{jk}^2}} \quad (4-20)$$

$$d_{ik}^2 = \|h_k - \hat{h}_i^*\|^2 \quad i = 1, 2, 3, \dots, c; \quad k = 1, 2, 3, \dots, M \quad (4-21)$$

最后的房间冲击响应模型为:

$$h_{\text{final}} = \frac{\sum_{j=1}^c \left(\sum_{k=1}^M (\mu_j(h_k))^2 \right) \hat{h}_j^*}{\sum_{j=1}^c \left(\sum_{k=1}^M (\mu_j(h_k))^2 \right)} \quad (4-22)$$

4.4 实验结果及分析

论文选择在一个普通的 $6 \text{ m} \times 5 \text{ m} \times 3 \text{ m}$ 房间进行实验, 在房间内选择五个点 P1、P2、P3、P4、P5 (如图 4-5) 来实现房间冲击响应的测量、平滑、建模及后面章节的均衡处理。实验装置包括一台笔记本电脑、一个扬声器(型号: HiVi X4) 和一个传声器(型号: ISK TRM9)。利用扬声器播放声音, 利用传声器录取声音, 通过本章介绍的解相关运算求取房间冲击响应, 再进行平滑和建模工作。图 4-6 到图 4-10 为实验结果。

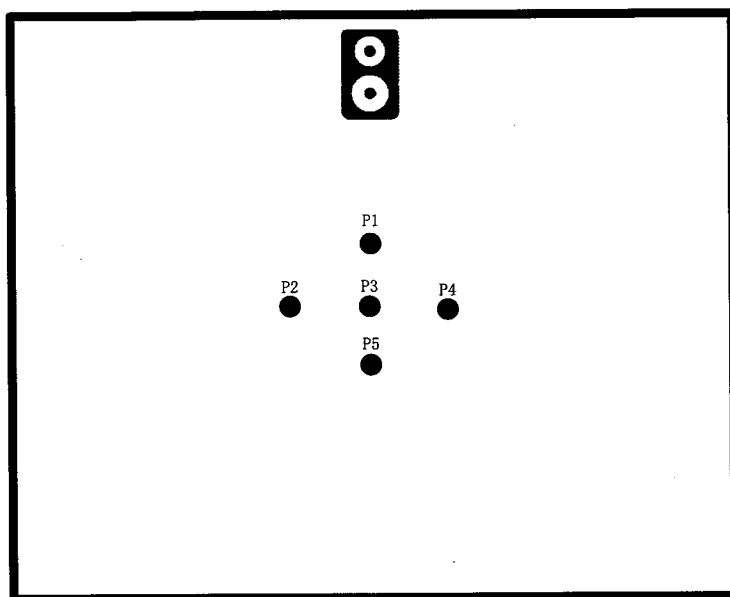


图 4-5 房间内 5 点选择示意图

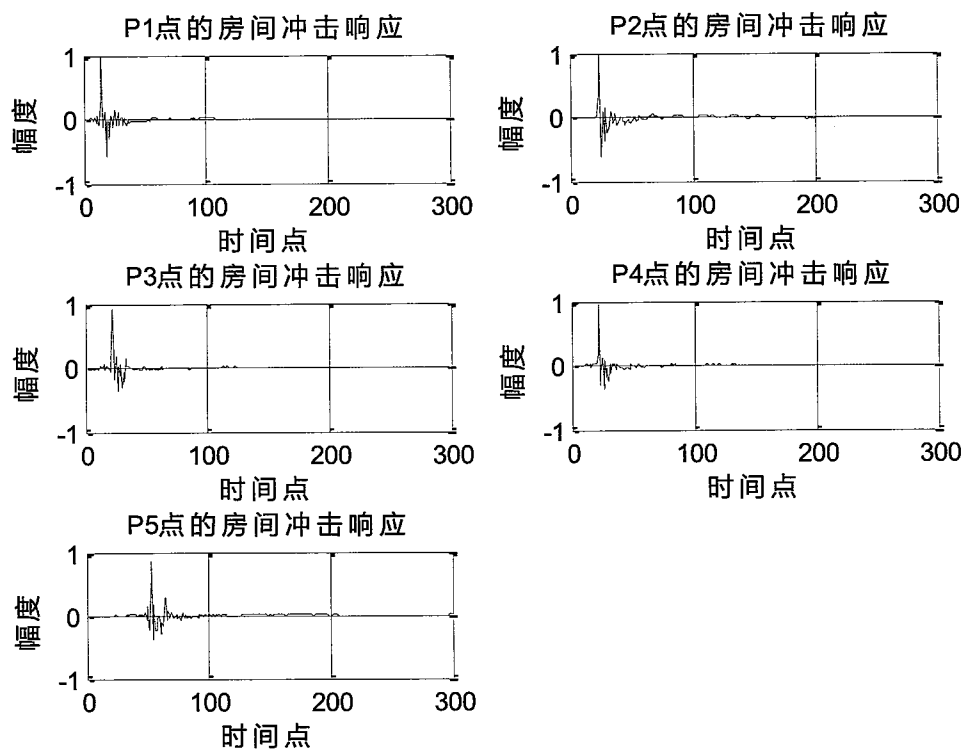


图 4-6 房间冲击响应时域测量结果

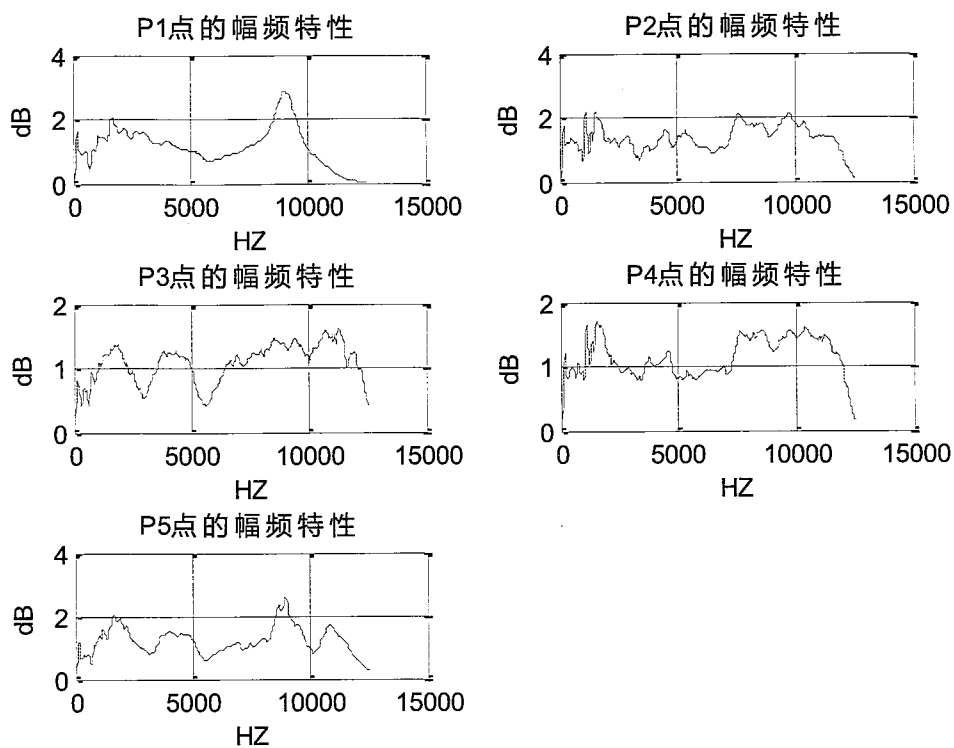


图 4-7 房间冲击响应频域测量结果

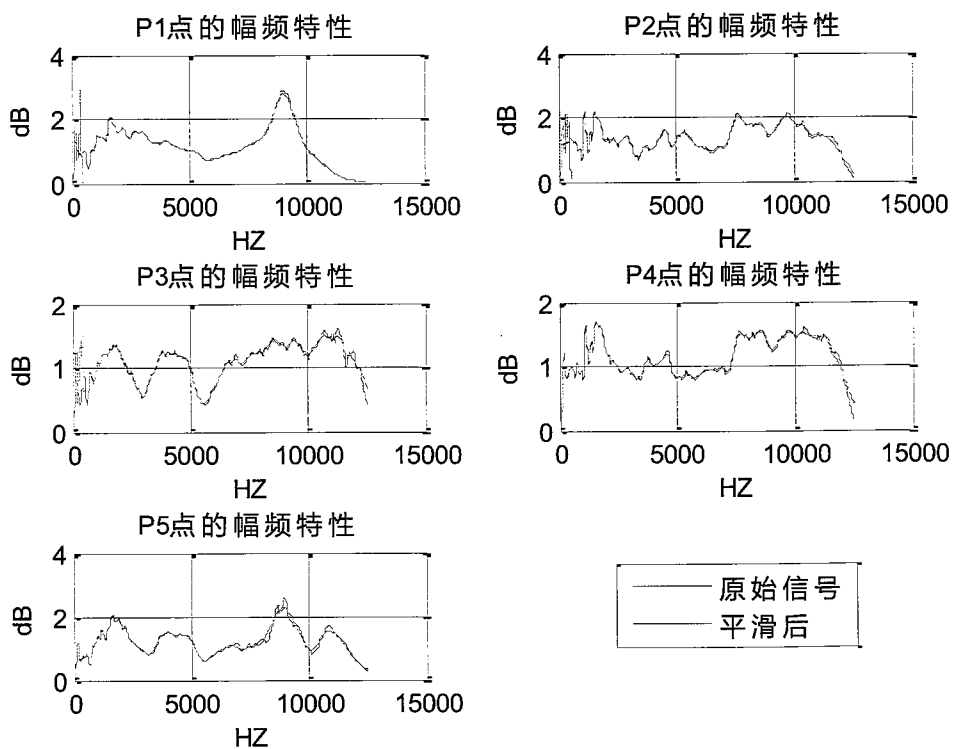


图 4-8 房间冲击响应平滑结果

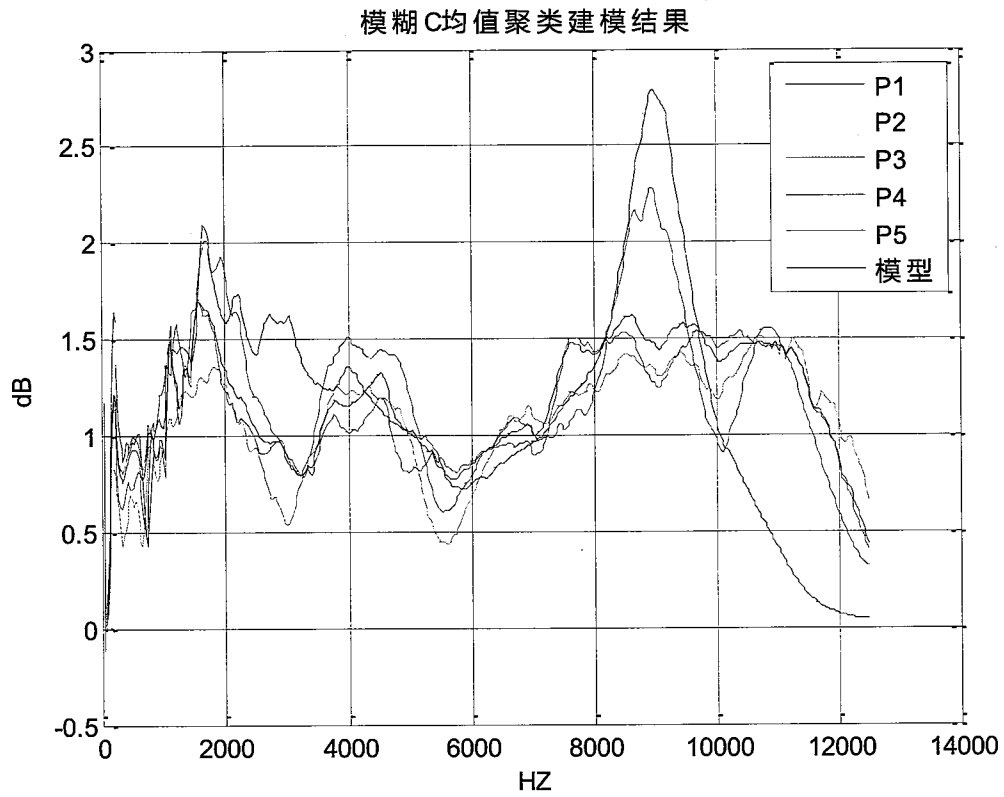


图 4-9 房间冲击响应建模结果

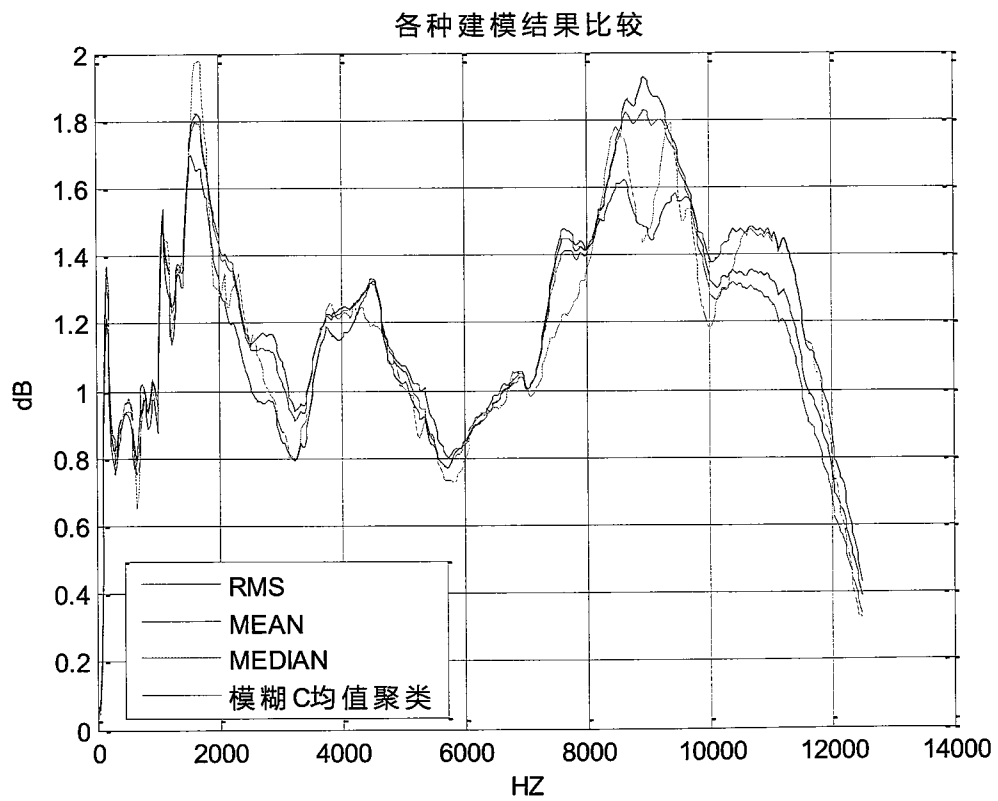


图 4-10 房间冲击响应建模结果比较

图 4-6 是在房间内五个点测量到的时域结果，由于受到扬声器声学特性、房间反射折射等影响，不是理想的冲激函数；这一点在图 4-7 中也有体现，各点房间冲击响应的幅频特性不是平直的，也就是说在不同频率信号的增益是不同的即产生了失真；图 4-8 是房间冲击响应平滑结果，由于采用 1/3 倍频程复数平滑，带宽随着频率增大，因此在高频段的平滑结果更为明显；图 4-9 给出了建模结果，由于位置的不同，各点房间冲击响应有很大的不同，房间冲击响应建模力求提取多点房间冲击响应数据中的共同趋势来表征整个房间的声学特性；图 4-10 则是对比了几种建模方法的结果，所有建模结果都能抓住整体趋势，但本文采用的基于模糊 C 均值聚类的建模方法能更好的去除大波动对整体趋势的影响。

4.5 本章总结

第四章介绍了房间冲击响应测量、房间冲击响应平滑、房间冲击响应建模三个预处理过程及结果，为第五章、第六章研究工作的开展提供了数据也奠定了基础。为了保证整个房间冲击响应均衡研究顺利进行，预处理工作不可或缺，因为只有得到可靠的房间冲击响应数据，均衡结果才有意义。这三步预处理工作仍有很多不足，在今后均衡研究的不断优化过程中也要对这部分工作进行改进。

第五章 基于响度度量的房间冲击响应多点均衡

传统房间冲击响应均衡方法都是以声音信号的幅值为度量标准的,但是人耳听觉系统是非常复杂的,涉及到声学、语音学和听力学,并且与频率、声强等诸多因素相关。另外,人耳听觉系统对声音的感受具有主观性,声压大小与人耳对声音强弱的感觉并不完全一致,甚至存在声压降低但人耳感觉更响的情况,这是因为人对声音的感受不仅与声压大小有关,还与声音的频率成分有关。前边的章节提到,响度是描述声音大小的主观量,它不仅取决于声音的声压级,还与声音时间、频率结构以及人耳听觉机理有关。响度是一种主观心理量,是人耳在分析和处理声音时提取的反映声音信号振幅特性的特征量,表征人类主观感觉到的声音强弱的程度。一般来说,当声音频率一定时,声音越强则响度也就越大。但是,响度是与频率有关的,相同的声强,频率不同时响度也不可能相同。响度是人耳听觉判别声音强度的等级概念。由于听觉具有复杂的强度特性、频率特性及时间特性,所以响度不仅取决于声音的强度,还与它的频率及瞬态时间特性有关。响度是听觉的基础。

基于人耳听觉系统的复杂性与响度的物理特性,本章主要介绍以响度为度量的房间冲击响应多点均衡算法。首先介绍算法原理,然后阐述响度转换滤波器的设计,接着介绍利用 LMS 算法求取均衡滤波器的最佳系数过程,进而给出实验结果及分析,最后对本章内容进行总结。

5.1 基于响度度量的房间冲击响应多点均衡算法思路

在传统的房间冲击响应均衡过程中,常采用自适应算法求取最佳的均衡滤波器系数,如图 5-1 所示: $X(z)$ 为测量房间传输函数时采用的 MLS 信号(见 4.1 节)作为自适应算法的输入信号; $H(z)$ 为由在房间各点测量得到的房间传输函数 $H_1(z)$, $H_2(z)$, ..., $H_8(z)$ 建立的房间冲击响应模型(见 4.3 节),使其尽量能反映整个房间的声学特性; $F(z)$ 为设计均衡滤波器特性; $Y(z)$ 为经过房间传输特性和均衡滤波器特性影响后的输出信号;期望信号 $X(z)$ 为加入适量延时的输入信号,即理想的均衡效果是输出信号无限逼近输入信号; $E(z)$ 为期望信号与输出信号之间的误差,

利用 LMS 算法不断调整均衡滤波器系数使误差减小，由此求得均衡滤波器的最佳系数。

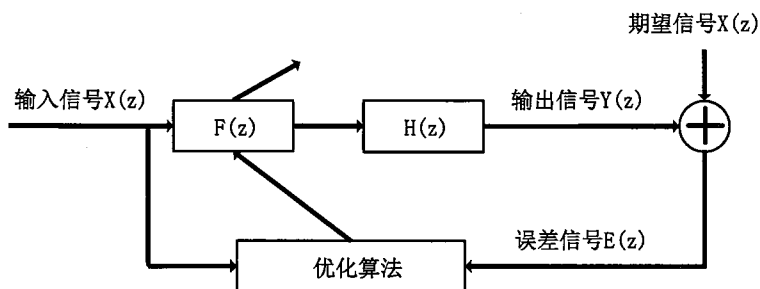


图 5-1 传统房间冲击响应均衡算法

由上述算法中，对不同的频率是等同处理的，所得结果在整个频率范围的总体算法最小，这样在数学上是最佳的，从听觉上未必如此。为了得到听觉上最佳的均衡效果，可以在计算误差时对不同的频率进行加权处理。在听觉系统分辨率较高的频率范围，应加大其对误差的贡献。相反在听觉系统分辨率较低的频段上容许其误差较大，可减少其对总体误差的贡献。由于响度是表征人耳听觉系统对声音的主观物理量，通过研究声压与响度的关系，设计响度转换滤波器 $T(z)$ ，如图 5-2 所示，对误差信号 $E(z)$ 进行加权处理，使其转换为响度误差 $E1(z)$ 。这样在房间冲击响应均衡过程中，控制目标由声压变为响度，得到的均衡结果更贴近听觉最佳。

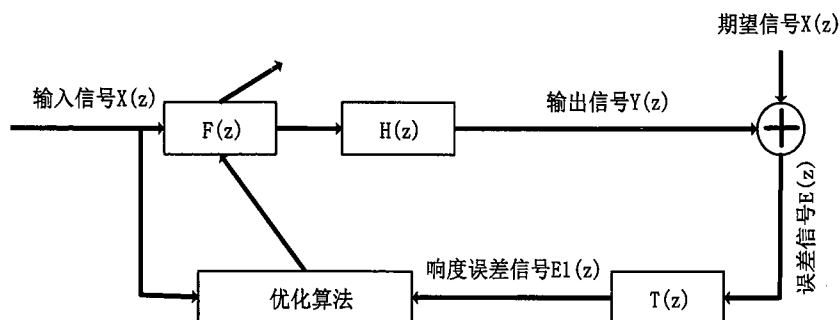


图 5-2 基于响度度量的房间冲击响应均衡算法

因此，整个均衡过程分为三部分：

- (1) 通过测量、平滑、建模（第四章）得到房间冲击响应模型 $H(z)$ ；
- (2) 设计响度转换滤波器，将误差信号由声压转化为响度；
- (3) 利用 LMS 算法求取均衡滤波器系数，对 $H(z)$ 进行均衡。

通过第四章的介绍，我们可以得到理想的房间冲击响应模型 $H(z)$ ，在本章接下来的内容将重点介绍响度滤波器的设计和最佳均衡滤波器系数的求取。

5.2 响度转换滤波器的设计

如图 5-3 所示, 等响曲线有很多条, 但对任意一条等响曲线而言共同点是: 不同频率声音的声压级是不同的, 而且整体趋势一致。从图 5-3 我们可以看出, 随着频率的增大, 声压级变化引起的响度变化关系并不是线性的。举例来说, 在频率为 1 kHz 时, 声压级 10 dB 的差距只能引起响度级为 10 phon 的变化, 而在频率为 100 Hz 时, 声压级 10 dB 的差距却能引起响度级为 20 phon 的变化。响度是反应人耳对声音大小的感受的主观物理量, 对于不同的频率上的声压级误差, 人耳的主观感受也是不同的。为了将声压刻度的误差转化为响度刻度的误差, 必须要研究不同频率上声压级变化引起的响度变化的关系, 然后利用此关系设计响度转换滤波器, 以响度为度量设计均衡滤波器。

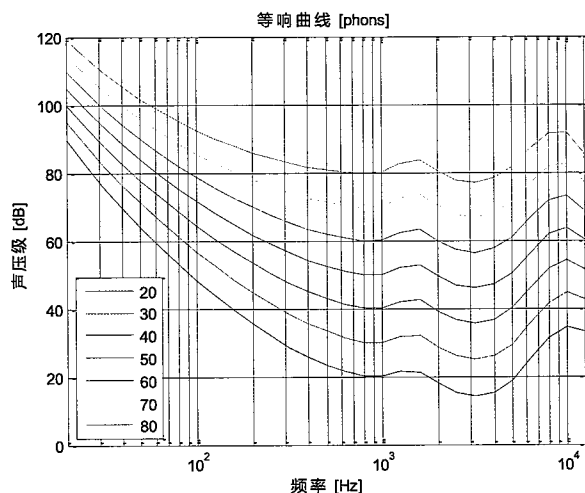


图 5-3 等响曲线 (来自 ISO 226:2003 修正版)

选择 20 phon 和 80 phon 的两条等响曲线为研究对象, 观察在不同频率下, 响度级从 20 phon 升至 80 phon 需要声压级做出多少改变。如图 5-4 所示, 同样是发生 60 phon 的响度级变化, 不同频率所需的声压级是不同的, 这个关系可以视为物理的声压误差对人耳主观的响度误差对应关系, 类似于声压级误差经过一个滤波器放大或缩小了的该误差对人耳主观听觉的影响。引起同样的响度级变化, 需要比较小声压级变化的频段是人耳比较敏感的频段, 即人耳可以轻易捕捉到这个频段的声压级变化; 需要比较大声压级变化的频段是人耳不太敏感的频段, 即人耳在这个频段可以容忍较大的声压级变化而听起来差不多。可以根据这个关系进行响度转换滤波器的设计。

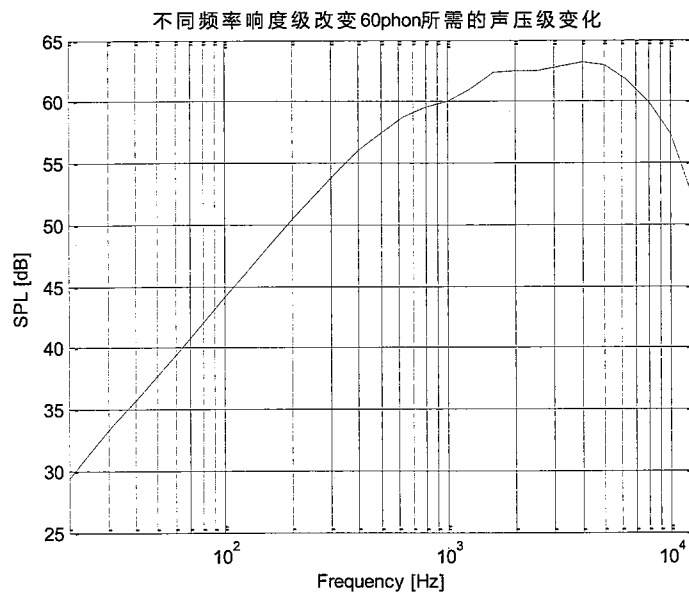


图 5-4 不同频率响度级改变 60 phon 所需的声压级变化

设计响度转换滤波器,是为了将声压误差转化为更符合人耳听觉特性的响度误差,根据图 5-4 级其分析,我们可以得到声压级变化与响度级变化之间的关系。设想将响度转换滤波器的幅频特性设计成图 5-4 曲线的倒置型,即需要较小的声压级变化就能引起较大响度级变化的频段(人耳较为敏感的频段)有一个较大的增益,放大这个频段的响度误差以增大和这个频段对总体误差的贡献;需要较大的声压级变化就能引起较大响度级变化的频段(人耳不太敏感的频段)有一个较小的增益,也就是允许该频段有较大的误差。得到的响度转换滤波器幅频特性如图 5-5。

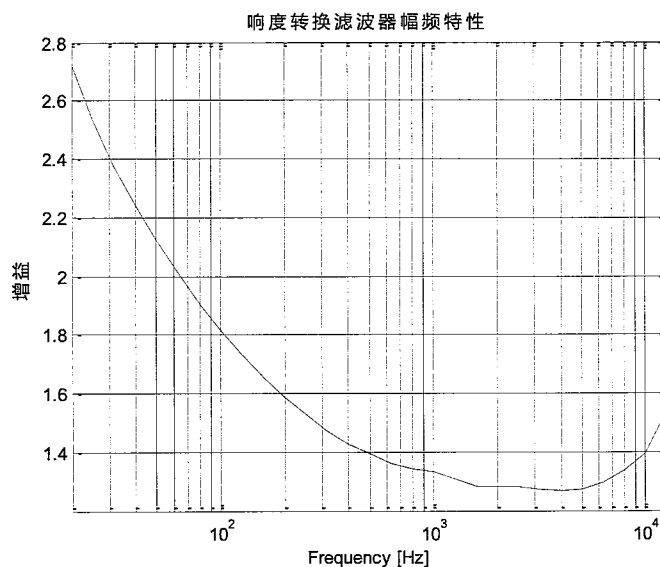


图 5-5 响度滤波器设计结果

5.3 最佳均衡滤波器系数求取

本文利用 LMS 算法进行均衡滤波器系数的求取, 本节首先简单介绍 LMS 算法, 然后阐述均衡器系数的求解过程。

5.3.1 LMS 算法

LMS (Least Mean Square) 算法是自适应滤波器中常用的简单有效的递归算法, 它不要求相关矩阵, 也不涉及矩阵求逆, 而是运用了最优化的数学算法——最陡下降法求取自适应滤波器的最佳权系数, LMS 算法是用均方误差作为性能指标, 也就是使滤波器的期望输出值和实际输出值之间的均方误差 (Mean Squared Error, MSE) 达到最小。设自适应滤波器的输入矢量为:

$$x(n) = [x(n), x(n-1), \dots, x(n-M+1)]^T \quad (5-1)$$

加权矢量 (即滤波器参数矢量) 为:

$$W(n) = [w_1(n), w_2(n), \dots, w_M(n)]^T \quad (5-2)$$

滤波器的输出为:

$$y(n) = W^T(n)X(n) = X^T(n)W(n) \quad (5-3)$$

误差信号 $e(n)$ 定义为期望输出 $d(n)$ 与滤波器实际输出之间的误差, 即

$$e(n) = d(n) - y(n) = d(n) - W^T(n)X(n) \quad (5-4)$$

$W(n)$ 是常数矢量的情况下, 其均方误差可表示为

$$J(n) = E[e^2(n)] = E[(d(n) - W^T(n)X(n))^2] \quad (5-5)$$

权矢量递推公式, 即 LMS 算法递推公式为:

$$\begin{aligned} J(n) &= E[e^2(n)] \\ &= E[(d(n) - W^T(n)X(n))^2] \end{aligned} \quad (5-6)$$

$$= E[d^2(n)] - 2E[d(n)X(n)]W^T(n) + W^T(n)E[X^2(n)]W(n)$$

$$\begin{cases} P = E[d(n)X(n)] \\ R = E[X(n)X^T(n)] \\ J(n) = E[d^2(n)] - 2PW^T(n) + W^T(n)RW(n) \end{cases} \quad (5-7)$$

利用最陡下降法的 LMS 算法, 滤波器系数迭代关系为:

$$W(n+1) = W(n) + 2\mu X(n)e(n) \quad (5-8)$$

5.3.2 最佳均衡滤波器系数求取

将 LMS 算法应用于基于人耳听觉特性的房间冲击响应均衡滤波器的设计，经过响度转换滤波器的作用，控制目标由声压误差转换为响度误差。由于系统加入响度转换滤波器，LMS 算法需要进行改进才能实现最佳均衡滤波器系数的求取。如上图所示： $X(z)$ 为测量房间传输函数时采用的 MLS 信号（见 4.1 节）作为自适应算法的输入信号； $h(n)$ 为房间冲击响应模型特性； $f(n)$ 为设计均衡滤波器特性； $y(n)$ 为经过房间传输特性和均衡滤波器特性影响后的输出信号；期望信号 $x(n)$ 为加入适量延时的输入信号，即理想的均衡效果是输出信号无限逼近输入信号； $e(n)$ 为期望信号与输出信号之间的误差； $t(n)$ 为响度转换滤波器； $e_1(n)$ 为误差信号 $e(n)$ 经过响度滤波器的处理转换为响度误差。

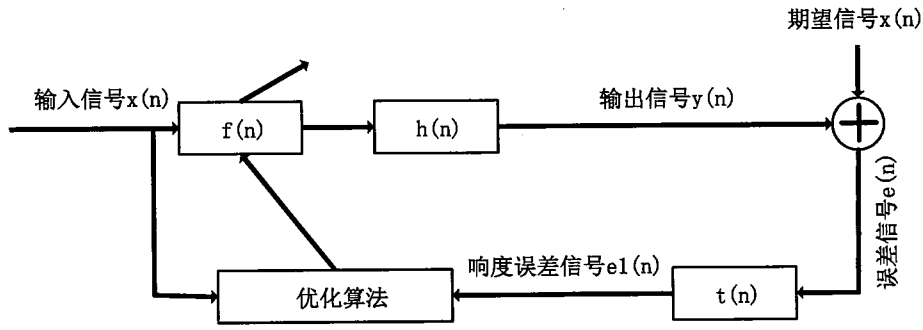


图 5-6 基于响度度量的均衡算法

经过房间系统响应和均衡滤波器响应的影响，输出信号为：

$$y(n) = x(n) * f(n) * h(n) \quad (5-9)$$

理想的均衡结果是输出信号无限逼近于输入信号，即近似于无失真传输，因此利用加入适量延时的输入信号作为理想信号，进行误差的求取：

$$e(n) = x(n) - y(n) \quad (5-10)$$

通过响度滤波器的处理，原来的误差信号转化为响度误差，即：

$$e_1(n) = e(n) * t(n) \quad (5-11)$$

将式(5-10)代入式(5-11)，可以得到式(5-12)。我们也可以理解为输出信号和理想信号都经过响度转换滤波器的预滤波处理：

$$e_1(n) = x(n) * t(n) - y(n) * t(n) \quad (5-12)$$

利用该误差求相应的代价函数，并利用最陡下降法得到均衡滤波器系数的迭代公式：

$$f(n+1) = f(n) + \mu * x(n) * t(n) * e_1(n) \quad (5-13)$$

5.4 实验结果及分析

利用 5.3 节求得的均衡器系数对房间冲击响应模型及每个点的房间冲击响应进行均衡处理，图 5-7 中黑色曲线为房间冲击响应模型的幅频特性，蓝色曲线为均衡滤波器幅频特性，红色曲线为均衡后的房间冲击响应模型幅频特性。从图 5-7 可以看出，该方法取得了理想的均衡结果，即便是在低频段均衡结果较差，也将幅频特性的波动均衡至人耳不能察觉的范围之内。

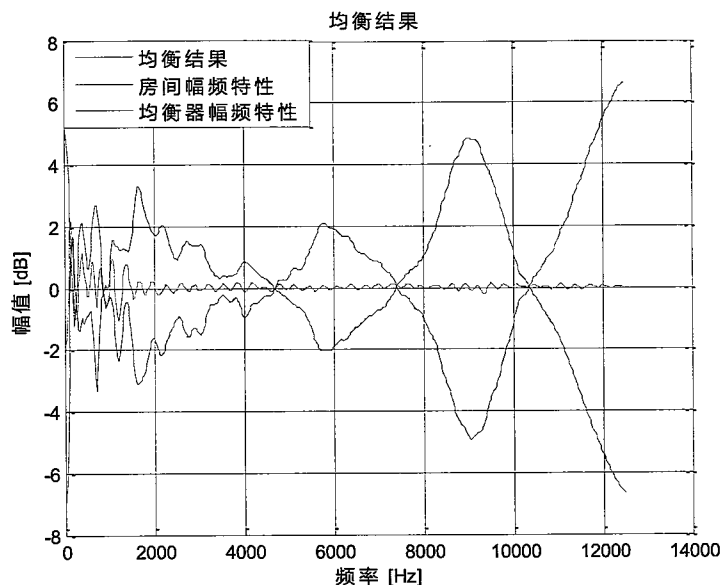


图 5-7 均衡结果

图 5-8 展示了多种均衡方法的到的均衡结果。对比 IIR 均衡滤波器、FIR 均衡滤波器、Kautz 均衡滤波器，论文提出的均衡算法有明显优势，尤其是在低频段，均衡结果明显优于其他三种方法。针对 WIIR 均衡滤波器，本文提出的方法在低频段的均衡结果虽然有所逊色，但是 WIIR 均衡滤波器在高频段的均衡结果却不及本文方法，这也验证了弯折均衡滤波器是牺牲高频段均衡结果来改善低频段均衡结果的结论。

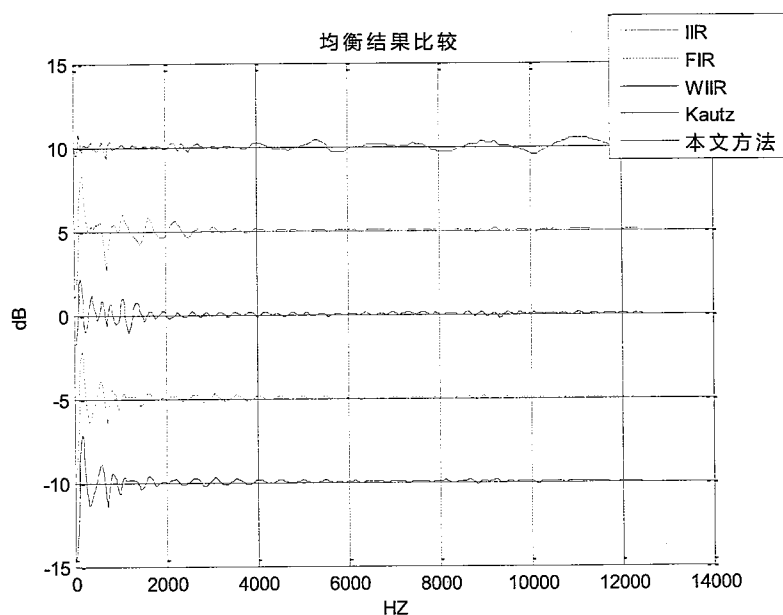


图 5-8 均衡结果比较

为了进一步对各类方法的均衡结果进行比较, 本文引入两种幅度均衡误差评价标准, 一种是传统线性均衡误差, 另一种是加权误差。分别表示为:

$$E1 = \sqrt{\frac{1}{f_H - f_L + 1} \sum_{i=f_L}^{f_H} (20 \log |E(f)| - d)^2} \quad (5-14)$$

$$E2 = \frac{1}{f_H - f_L + 1} \sum_{i=f_L}^{f_H} W(f) (20 \log |E(f)| - d) \quad (5-15)$$

其中, d 为幅频特性的平均值, 两种误差计算方法都是用来衡量幅频特性偏离平均水平的程度。 $E2$ 中的加权函数 $W(f)$ 是响度转换滤波器的幅频特性, 目的也是将误差从幅度转化为响度, 用更符合人耳听觉特性的准则来评价均衡结果。另外均衡比例为均衡前的误差减去均衡后的误差相对均衡前误差的百分比, 即:

$$p = \frac{E_{\text{before}} - E_{\text{after}}}{E_{\text{before}}} \quad (5-16)$$

五种方法的均衡误差汇总, 如表 5-2 所示。

计算得到的误差结果与图 5-8 分析结果一致, 本文方法得到的均衡后的频谱基本平坦, 计算得到的线性误差和加权误差分别为 0.2917 dB 和 0.8861 dB, 幅频特性的峰谷起伏在 ± 1.5 dB 左右, 这种程度的起伏人耳是无法察觉的。无论是线性误差还是加权误差, 本文方法都优于 IIR 均衡滤波器、FIR 均衡滤波器、Kautz 均衡滤波器。对于线性误差, 本文提出的方法虽然不及弯折 IIR 滤波器, 但是加

权误差却有明显优势,正是因为本文方法在提高低频段均衡效果的同时没有影响其他频段的均衡结果。

表 5-2 均衡误差比较

	$E1$	均衡比例 p	$E2$	均衡比例 p
房间冲击响应模型误差	2.4355		11.4652	
IIR 均衡误差	0.4920	79.8%	1.0873	90.5%
FIR 均衡误差	0.4862	80%	1.0010	91.2%
WIIR 均衡误差	0.2269	90.7%	1.0978	90.4%
Kautz 均衡误差	0.4868	80%	1.0087	91.2%
本文方法均衡误差	0.2917	88.1%	0.8861	92.3%

以上实验结果证明该方法在人耳听觉特性方面的可取性,为了证明多点均衡的正确性,即对整个房间冲击响应模型进行均衡可以实现在其他点也有均衡效果,实验对各点的均衡前后误差进行比较。图 5-9、表 5-3 的结果说明房间各点都有均衡效果,无论是线性误差还是加权误差,都能反映出各点得到一定程度的均衡,但是远没有对房间冲击响应模型均衡结果那么理想。

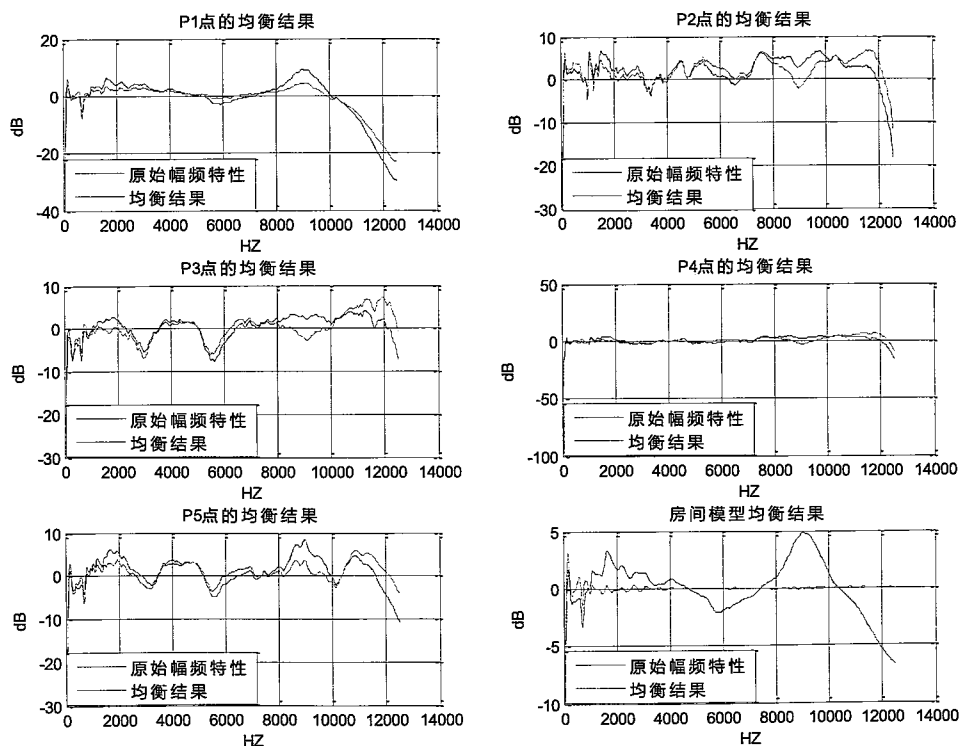


图 5-9 房间各点均衡结果

表 5-3 房间各点均衡结果

		<i>E1</i>	均衡比例	<i>E2</i>	均衡比例
房间冲击响应模型	均衡前	2.4355	88%	11.4652	92.3%
	均衡后	0.2917		0.8861	
<i>P1</i> 点	均衡前	8.1993	24.5%	34.3315	25.1%
	均衡后	6.1938		25.7419	
<i>P2</i> 点	均衡前	3.9136	14.4%	15.2291	6.1%
	均衡后	3.3505		14.3103	
<i>P3</i> 点	均衡前	3.3948	10.2%	16.1777	12.5%
	均衡后	3.0477		14.1486	
<i>P4</i> 点	均衡前	4.2298	14.7%	14.4879	16.7%
	均衡后	3.6099		12.0745	
<i>P5</i> 点	均衡前	3.7821	29.8%	18.3201	32.2%
	均衡后	2.6555		12.4200	

为了验证基于人耳听觉特性的房间响应均衡效果，主观的测听实验是不可或缺的。本实验则是选用 10 名测听人员来进行均衡结果主观测听实验，他们都是本实验室的学生，从事语音或音频信号处理方面的研究工作，对声音信号方面都比较熟悉。每名测听人员随机听取四段不同类型的声音频文件，每段音频文件又有五种不同的处理方法，最后根据主观听觉感受的好坏从 0 到 10 分打分，0 分表示最差，10 分表示最好，然后 10 个人的打分取平均，实验结果如表 5-4，。

表 5-4 主观测听结果

	音频 1	音频 2	音频 3	音频 4	平均
IIR 均衡方法	6.6	7.5	7.0	8.2	7.325
FIR 均衡方法	8.5	7.4	6.9	8.1	7.725
WIIR 均衡方法	9.3	6.8	7.9	9.0	8.25
Kautz 均衡方法	7.6	7.1	8.4	8.8	7.975
本文均衡方法	9.1	8.0	7.2	8.7	8.2

从表 5-4 中汇总的结果来看, 主观测听实验结果不是很理想, 在五种均衡方法打分中排在第二位, 没有明显优势。分析原因为: 1、无论应用什么设备进行重放, 重放效果都会受到设备自身性能的影响, 所以没有理想的重放效果作为测听人员的参考, 仅凭主观感受打分, 影响因素较多且不稳定; 2、几种均衡方法的处理结果都已达到人耳无法察觉的误差范围内, 所以测听人员无法分辨出能够在数据上体现出的差别; 3、主观测听实验设计的不合理, 测听人员没有经过严格专业的训练。

5.5 本章总结

本章给出了基于响度度量的均衡方法, 关键是设计响度转换滤波器将误差从幅度转换为响度, 力求从更符合人耳听觉特性的角度设计均衡滤波器和评价均衡结果。实验结果表明, 引入响度滤波器确实可以在一定程度上改善低频段的均衡效果而且不影响高频的均衡结果。但是考虑到人耳听觉系统的另一特性临界频带宽度, 在对误差进行加权时应该注意在一个临界频带内的加权函数不宜发生跳变且在不同频段人耳对频率的分辨率不同, 因此响度转换滤波器系数的精确度也应该有相应的调整, 因此需要对响度转换滤波器的设计进行改进, 这正是第六章将要进行的内容。

第六章 基于临界频带宽度的房间冲击响应多点均衡

第五章基于响度的房间冲击响应多点均衡是为了将均衡器设计过程中的度量由声压转化为响度,更贴近人耳的主观听觉感受。人耳听觉特性的另外一个体现是人耳听觉系统对频率的分辨率不是线性的,均衡过程中,如果对低频和高频的分辨率等同处理,是不符合人耳的听觉特性的。人耳对低频的分辨率要高于对高频的分辨率,因此许多研究人员就以进一步改善在低频段的均衡效果为目标进行房间冲击响应均衡的研究。第五章提到,为了取得听觉最佳均衡结果,引入响度转换滤波器,响度转换滤波器的设计是核心工作。由于人类听觉系统有关频率的分辨率是与听觉滤波器的频带宽度和形状有关的,在响度转换滤波器的设计过程中不仅需要考虑响度特性还有考虑人耳频率分辨率,因此需要将临界频带宽度列入研究范围之内,进一步优化响度转换滤波器。优化目标有两个:1、降低求取转换滤波器系数时的计算复杂度;2、在低频段提高设计所得响度滤波器幅频特性的精确度。

本章主要介绍应用临界频带宽度进行响度滤波器的优化的房间冲击响应多点均衡算法。首先介绍利用 Bark 模拟的临界频带宽度,然后阐述基于临界频带宽度的响度转换滤波器优化,进而给出实验结果及分析,最后对本章内容进行总结。

6.1 临界频带宽度

临界频带是听觉学和心理声学的一个专业名词,它于 19 世纪 40 年代年被 Harvey Fletcher 提出。耳蜗是内耳中听觉的传感器官,临界频带指的是由于耳蜗构造产生的听觉滤波器的频率带宽。概况地说,临界频带是声音频率带,在临界频带中第一个单音感知性会被第二单音的听觉掩蔽所干扰。声学研究中,人们使用听觉滤波器来模拟不同的临界频带。后来研究者发现人耳结构大致会对 24 个频率点产生共振,根据这个结论 Eberhard Zwicker 在 1961 年针对人耳特殊结构提出:信号在频带上也呈现出 24 个临界频带,分别从 1 到 24。这就是 Bark 域。根据相关的心理声学的研究,由于人耳的特殊结构,在同一个临界频带里信号容易发生掩蔽效应,即主要信号容易被能量大并且频率接近的掩蔽信号所掩蔽。因此我

们可以认为 Bark 域约近的信号越容易产生掩蔽效应。然而，Bark 域表示必须通过查表获得，在上个世纪 80 年代左右到 90 年代左右学者们提出了各种近似函数来近似表示 Bark 域。如表 6-1:

表 6-1 表征 Bark 域的近似函数

Zwicker & Terhardt (1980)	$\text{Bark} = 13 \times \tan^{-1}(0.76f/1000) + 3.5 \times \tan^{-1}(f/7500)^2$
	$\text{Bark} = 8.7 + 14.2 \times \log_{10}(f/1000)$
Terhardt (1979)	$\text{Bark} = 13.3 \times \tan^{-1}(0.75f/1000)$
	$\text{Bark} = 12.82 \times \tan^{-1}(0.78f/1000) + 0.17(f/1000)^{1.4}$
Wang, Sekey & Gersho (1992)	$\text{Bark} = 6 \times \sin^{-1}(f/600)$
Schroeder (1977)	$\text{Bark} = 7 \times \sin^{-1}(f/650)$
Traunmüller (1990)	$\text{Bark} = 26.81/(1+(1960/f)) - 0.53$

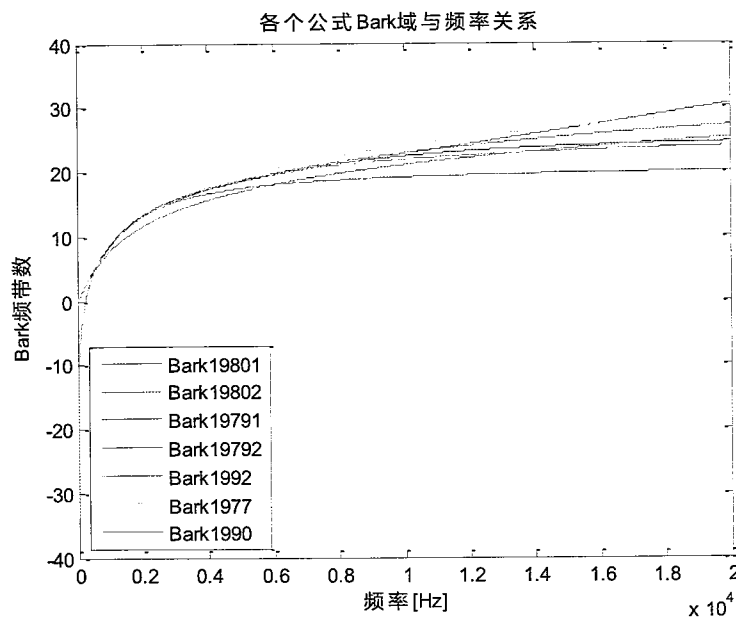


图 6-1 Bark 与频率关系

图 6-1 展示的偏差无法代表模型对 Bark 域描述的精度，也并不能精确表示各个公式换算临界频带的准确度。这是因为每个人的主观听感是不一样的，因此 Bark 近似描述曲线的效果会因人而异，在实际研究中可以根据需要选择相关的

Bark 域近似曲线。站着另外一个角度，一个声音信号在线性频率域和 Bark 域的表现分别如图 6-2、图 6-3：

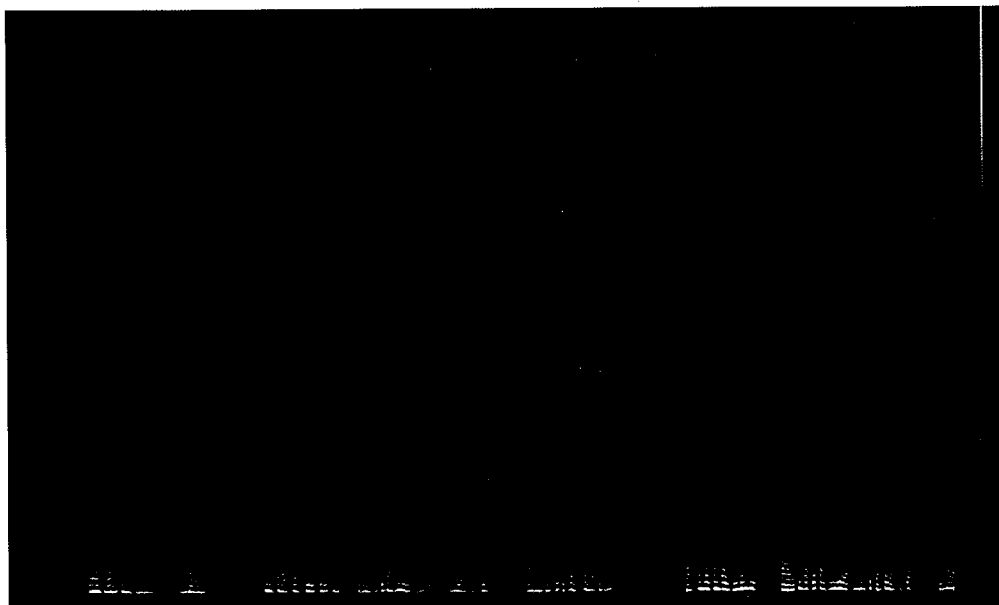


图 6-2 声音信号在线性频率域表现

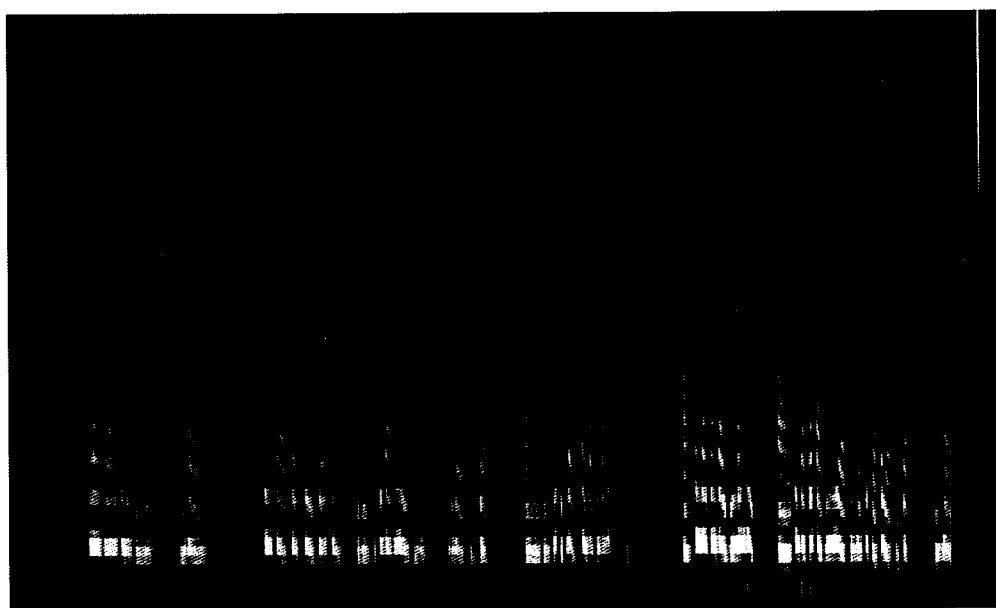


图 6-3 声音信号在 Bark 域表现

从图 6-2、图 6-3 两个图容易看出，线性频率域对低频信号不够直观，Bark 域对低频具有放大作用，高频具有压缩作用。这点能够从上面的转换公式明显看出来。回顾声音的掩蔽效应，从 Bark 域来看，能够更清晰地分析出哪些信号容易产生掩蔽，哪些噪声比较明显。同时对于某些声音特征例如基音、音调等在

Bark 域分析也有独特的效果。总体来说在 Bark 域中能够更加真实地反应人耳对信号产生的感觉。所以本章选择在 Bark 域进行响度转换滤波器的优化。

6.2 响度转换滤波器优化

为了使响度滤波器对误差的处理结果更贴近人耳听觉特性，并且保证在一个临界频带宽度内加权系数没有突变的现象发生，首先对现在的响度转换滤波器幅频特性在 Bark 域进行平滑。如式(6-1)所示， $H(f)$ 是由等响曲线设计的响度转换滤波器的幅频特性， f_1 、 f_2 是某个 Bark 域的截止频率， H_{12} 是这个 Bark 带的平滑结果，即保证在一个 Bark 域内所有频率的加权系数相同。

$$H_{12} = \frac{1}{f_2 - f_1} \int_{f_1}^{f_2} H(f) df \quad (6-1)$$

在前一章提到的响度滤波器设计过程中，分析等响曲线及其声压误差与响度误差的关系可以得到大体的响度滤波器幅频特性，为了简化运算减低计算复杂度，可以在曲线上均匀取到 20 个点进行滤波器系数的求取。如图 6-4 所示：

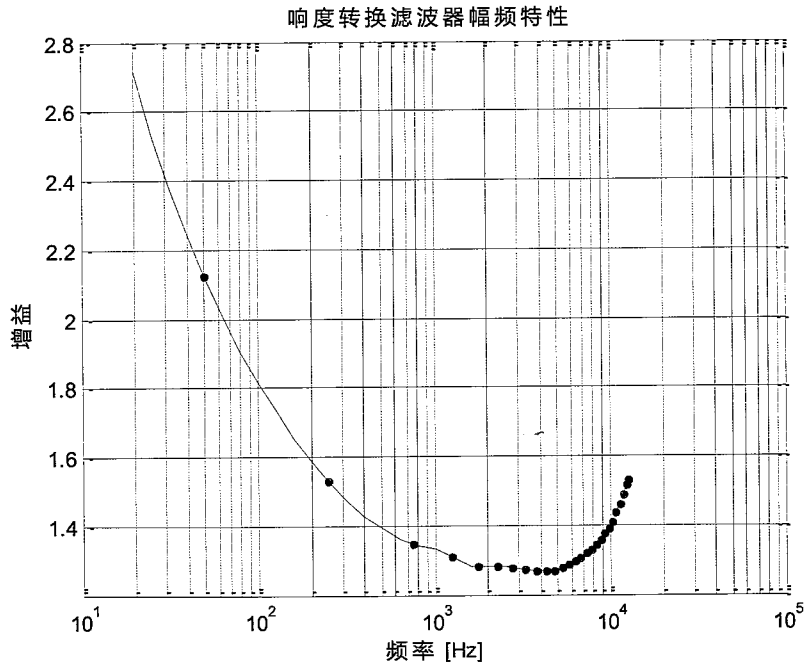


图 6-4 响度滤波器设计线性域采样点

但是，为了进一步考虑人耳听觉系统对频率分辨率的特性：随着频率的增大，人耳对频率的分辨率逐渐降低。临界频带宽度反映了听觉系统对频率的非线性分辨率。1 Bark 等于一个临界频带宽度，因此本文在 Bark 域进行响度转换滤波

器的设计。具体实现是改变对响度转换滤波器幅频特性的采样方式，在 Bark 域上进行采样，如图 6-5 所示，

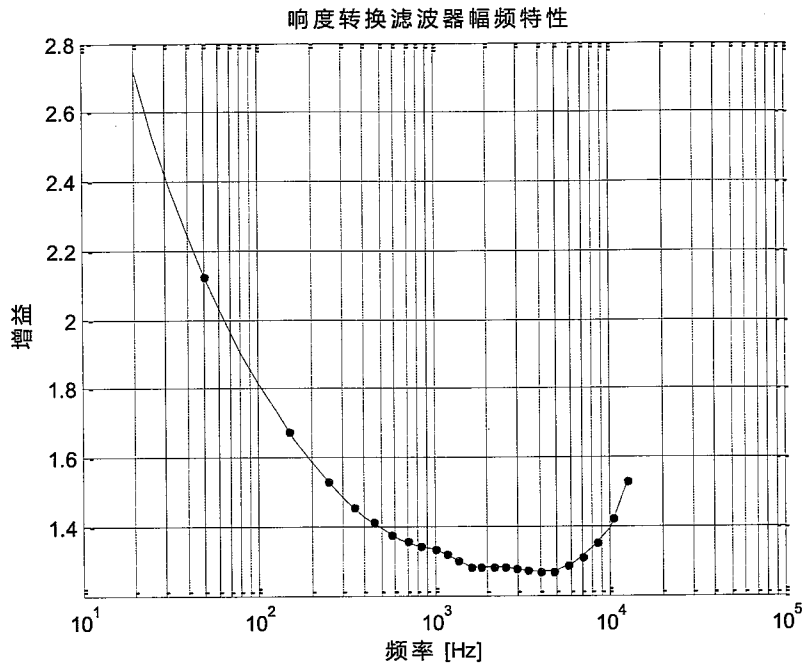


图 6-5 响度滤波器设计 Bark 域采样点

采样后求得每个频段的增益，汇总到表 6-2

表 6-2 Bark 域上响度转换滤波器采样

序号	频段	增益	序号	频段	增益
1	0-0.1 kHz	2.1218	13	1.72-2.0 kHz	1.2821
2	0.1-0.2 kHz	1.6748	14	2.0-2.32 kHz	1.2809
3	0.2-0.3 kHz	1.5292	15	2.32-2.7 kHz	1.2800
4	0.3-0.4kHz	1.4563	16	2.7-3.15kHz	1.2760
5	0.4-0.51kHz	1.4106	17	3.15-3.7kHz	1.2714
6	0.51-0.63 kHz	1.3768	18	3.7-4.4 kHz	1.2664
7	0.63-0.77 kHz	1.3548	19	4.4-5.3 kHz	1.2694
8	0.77-0.92kHz	1.3410	20	5.3-6.4kHz	1.2853
9	0.92-1.08kHz	1.3332	21	6.4-7.7kHz	1.3117
10	1.08-1.27kHz	1.3181	22	7.7-9.5kHz	1.3509
11	1.27-1.48kHz	1.3015	23	9.5-12.0kHz	1.4223
12	1.48-1.72kHz	1.2833			

取前 22 个点的数据确定一条响度转换滤波器的拟合曲线。代入公式：

$$\left\{ \begin{array}{l} A_1 = \frac{b_1 z_1^{-1} + b_2 z_1^{-2} + b_3 z_1^{-3} + b_4 z_1^{-4} + b_5 z_1^{-5} + b_6 z_1^{-6} + b_7 z_1^{-7} + b_8 z_1^{-8} + b_9 z_1^{-9} + b_{10} z_1^{-10} + b_{11} z_1^{-11}}{1 + a_1 z_1^{-1} + a_2 z_1^{-2} + a_3 z_1^{-3} + a_4 z_1^{-4} + a_5 z_1^{-5} + a_6 z_1^{-6} + a_7 z_1^{-7} + a_8 z_1^{-8} + a_9 z_1^{-9} + a_{10} z_1^{-10} + a_{11} z_1^{-11}} \\ A_2 = \frac{b_1 z_2^{-1} + b_2 z_2^{-2} + b_3 z_2^{-3} + b_4 z_2^{-4} + b_5 z_2^{-5} + b_6 z_2^{-6} + b_7 z_2^{-7} + b_8 z_2^{-8} + b_9 z_2^{-9} + b_{10} z_2^{-10} + b_{11} z_2^{-11}}{1 + a_1 z_2^{-1} + a_2 z_2^{-2} + a_3 z_2^{-3} + a_4 z_2^{-4} + a_5 z_2^{-5} + a_6 z_2^{-6} + a_7 z_2^{-7} + a_8 z_2^{-8} + a_9 z_2^{-9} + a_{10} z_2^{-10} + a_{11} z_2^{-11}} \\ \vdots \\ A_{22} = \frac{b_1 z_{22}^{-1} + b_2 z_{22}^{-2} + b_3 z_{22}^{-3} + b_4 z_{22}^{-4} + b_5 z_{22}^{-5} + b_6 z_{22}^{-6} + b_7 z_{22}^{-7} + b_8 z_{22}^{-8} + b_9 z_{22}^{-9} + b_{10} z_{22}^{-10} + b_{11} z_{22}^{-11}}{1 + a_1 z_{22}^{-1} + a_2 z_{22}^{-2} + a_3 z_{22}^{-3} + a_4 z_{22}^{-4} + a_5 z_{22}^{-5} + a_6 z_{22}^{-6} + a_7 z_{22}^{-7} + a_8 z_{22}^{-8} + a_9 z_{22}^{-9} + a_{10} z_{22}^{-10} + a_{11} z_{22}^{-11}} \end{array} \right. \quad (6-2)$$

$$\omega_k = 2\pi \frac{F_k}{f_s} = 2\pi \frac{F_k}{12500}$$

$$Z_k = \cos(\omega_k) + i \sin(\omega_k)$$

$$F_k = \{50, 150, 250, 350, 450, 570, 700, 840, 1000, 1170, 1370, 1600, 1850, 2150, 2500, 2900, 3400, 4000, 4780, 5800, 7000, 8500\}$$

$$A_k = \left\{ \begin{array}{l} 1.2856, 1.7313, 1.9842, 2.1519, 2.2731, 2.3746, 2.4508, 2.4968, 2.4994, 2.4258, 2.3780, 2.3525, \\ 2.4754, 2.6061, 2.7390, 2.7817, 2.7848, 2.7286, 1.9602, 2.3042, 2.0785, 1.9085, 1.8622 \end{array} \right\}$$

$$1 \leq k \leq 22$$

根据这 22 个点的数据确定一条响度转换滤波器的拟合曲线，令 $M=N=11$ ，方程组如式(6-2)，由方程组可以求得响度转换滤波器系数，得到设计的响度转换滤波器。将求得的系数带入滤波器方程，可以得到根据等响曲线及声压差与响度差关系设计的响度转换滤波器，其频响特性如图 6-6 所示。本文正是用此响度滤波器将幅度误差转化为响度误差，力求得到更符合人耳听觉特性的均衡结果。

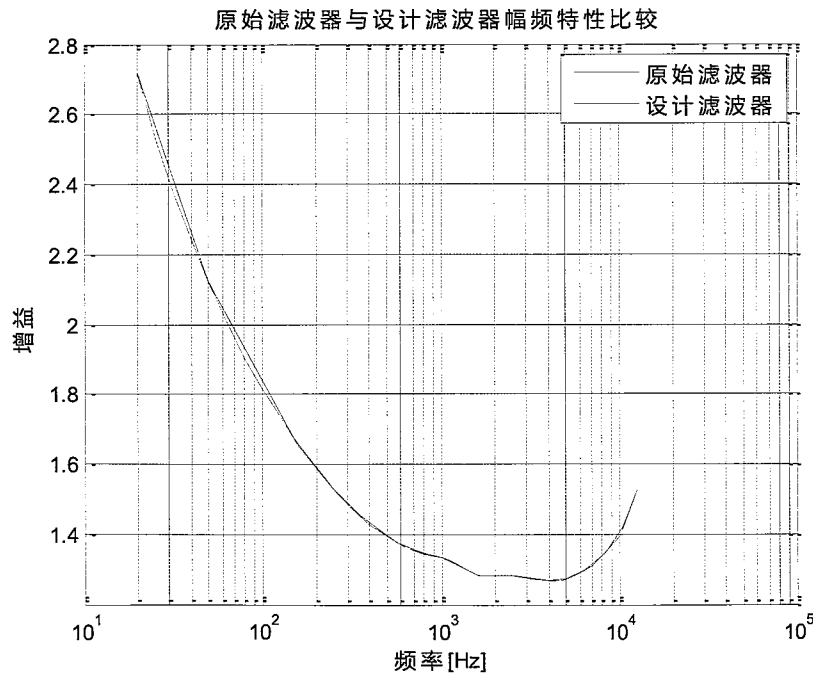


图 6-6 响度滤波器设计结果

6.3 实验结果及分析

利用 6.2 节求得的均衡器系数对房间冲击响应模型及每个点的房间冲击响应进行均衡处理，图 6-7 中黑色曲线为房间冲击响应模型的幅频特性，蓝色曲线为均衡滤波器幅频特性，红色曲线为均衡后的房间冲击响应模型幅频特性。从图 6-7 可以看出，利用临界频带宽度 Bark 对响度转换滤波器优化之后，再应用到均衡滤波器设计当中，同样取得了理想的均衡结果，均衡后的幅频特性已经十分平直，低频段也能到达人耳分辨不出的效果。

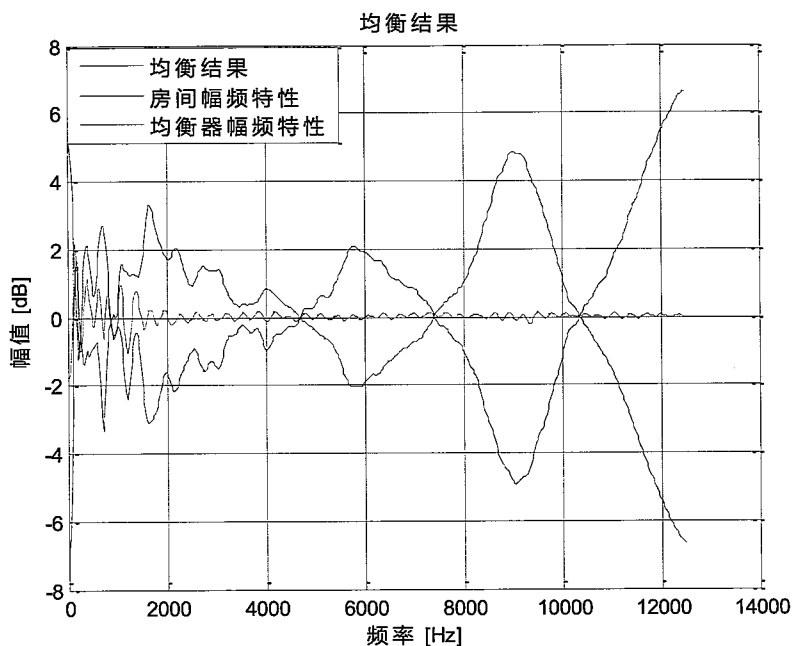


图 6-7 房间冲击响应模型均衡结果

图 6-8 展示的是多种均衡方法的到的均衡结果。与第五章的结果类似，对比 IIR 均衡滤波器、FIR 均衡滤波器、Kautz 均衡滤波器，论文提出的均衡算法有明显优势，尤其是在低频段，均衡结果明显优于其他三种方法。同样的 WIIR 均衡滤波器在低频段的均衡效果要优于文章提出的均衡方法，但是 WIIR 均衡滤波器在高频段的均衡结果却仍然不及本文方法，这也验证了弯折均衡滤波器是牺牲高频段均衡结果来改善低频段均衡结果的结论。

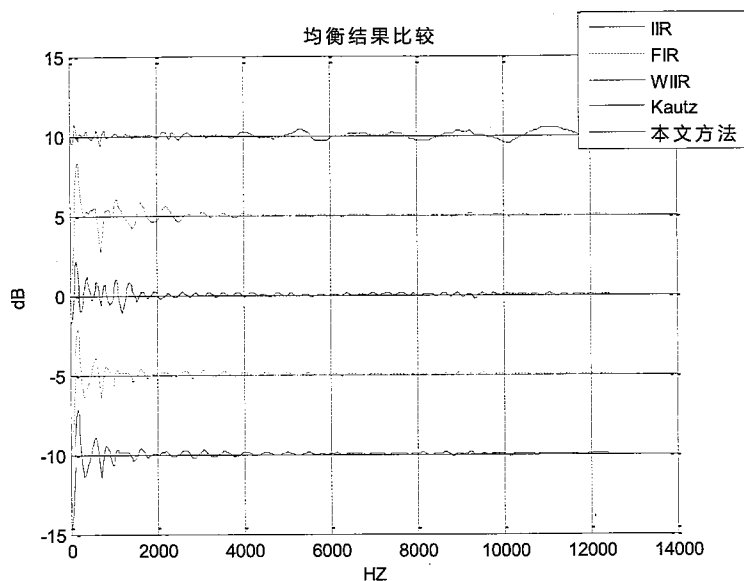


图 6-8 房间冲击响应模型均衡结果比较

为了进一步对各类方法的均衡结果进行比较,分析每种方法的性能,本章同样引入两种幅度均衡误差评价标准,一种是传统线性均衡误差,另一种是加权误差。分别表示为:

$$E1 = \sqrt{\frac{1}{f_H - f_L + 1} \sum_{i=f_L}^{f_H} (20 \log |E(f)| - d)^2} \quad (6-3)$$

$$E2 = \frac{1}{f_H - f_L + 1} \sum_{i=f_L}^{f_H} W(f) (20 \log |E(f)| - d) \quad (6-4)$$

另外均衡比例为均衡前的误差减去均衡后的误差相对均衡前误差的百分比,即:

$$p = \frac{E_{\text{before}} - E_{\text{after}}}{E_{\text{before}}} \quad (6-5)$$

五种方法的均衡误差汇总,如表 6-3 所示:

表 6-3 房间冲击响应模型均衡结果比较

	$E1$	均衡比例 p	$E2$	均衡比例 p
房间冲击响应模型	2.4355		11.5368	
IIR	0.4920	79.8%	1.1072	90.4%
FIR	0.4862	80%	1.0229	91.1%
WIIR	0.2269	90.7%	1.1057	90.4%
Kautz	0.4868	80%	1.0308	91%
本文方法	0.2718	88.8%	0.8007	93.1%

表 6-3 结果与图 6-8 结果一致，在图 6-8 中我们可以看出本文方法得到的均衡后的频谱基本平坦，计算得到的线性误差和加权误差分别为 0.2718 dB 和 0.8007 dB，幅频特性的峰谷起伏在 ± 1.5 dB 左右，这种程度的起伏人耳是无法察觉的。本章的基于临界频带宽度的均衡方法无论是线性误差还是加权误差，都比 IIR 均衡滤波器、FIR 均衡滤波器、Kautz 均衡滤波器更理想。线性误差本文提出的方法不及弯折 IIR 滤波器，但是加权误差却有明显优势，正是因为本文方法在提高低频段均衡效果的同时没有影响其他频段的均衡结果。

以上实验结果证明该方法在人耳听觉特性方面的可取性，为了证明多点均衡的正确性，即对整个房间冲击响应模型进行均衡可以实现在其他点也有均衡效果，实验对各点的均衡前后误差进行比较。图 6-9、表 6-4 的结果说明房间各点都有均衡效果，无论是线性误差还是加权误差，都能反映出各点得到一定程度的均衡，但是远没有对房间冲击响应模型均衡结果那么理想。

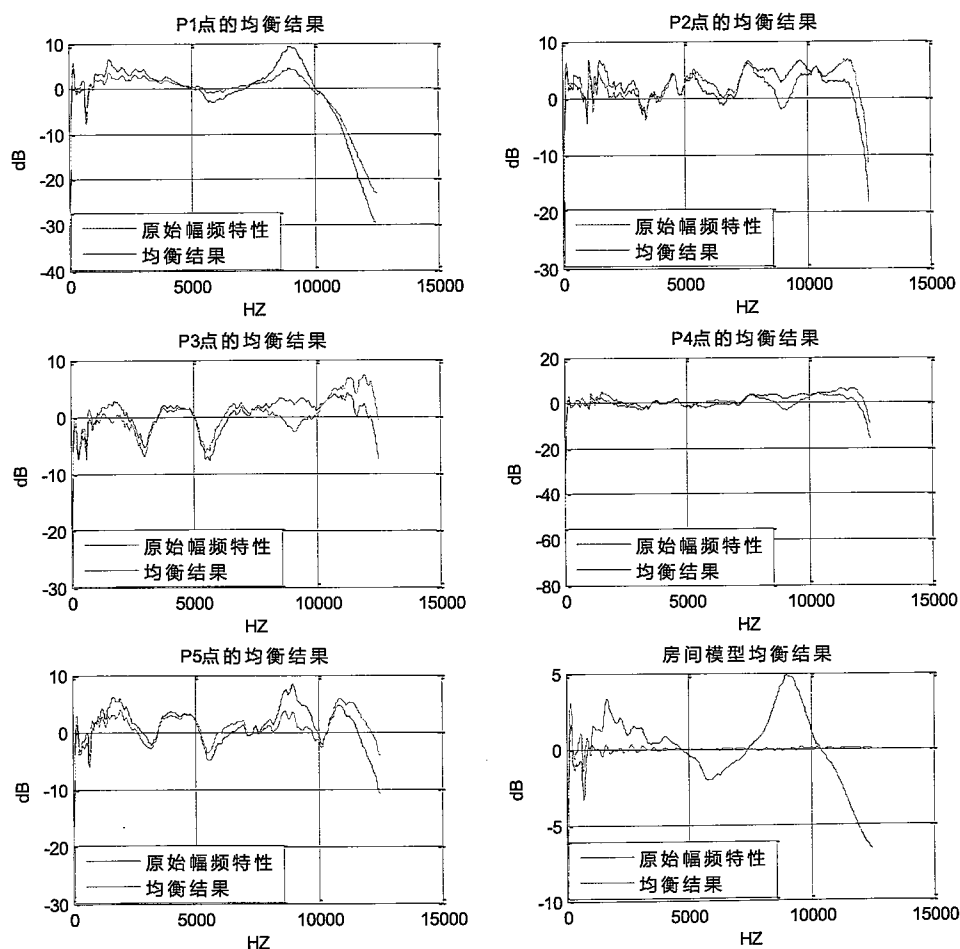


图 6-8 各点房间冲击响应均衡结果

表 6-3 各点房间冲击响应均衡结果

		$E1$	均衡比例	$E2$	均衡比例
房间冲击响应模型	均衡前	2.4355	88.8%	11.5368	93.1%
	均衡后	0.2718		0.8007	
$P1$ 点	均衡前	8.1993	24.5%	34.4938	25%
	均衡后	6.1937		25.8776	
$P2$ 点	均衡前	3.9136	14.4%	15.3564	6%
	均衡后	3.3505		14.4273	
$P3$ 点	均衡前	3.3948	7%	16.3162	10.8%
	均衡后	3.1475		14.5542	
$P4$ 点	均衡前	4.2298	13.6%	14.6242	16.6%
	均衡后	3.6543		12.2015	
$P5$ 点	均衡前	3.7821	29.8%	18.4431	32.2%
	均衡后	2.6553		12.5046	

另外，对比第五章基于响度度量的房间冲击响应均衡，第六章基于临界频带宽度的房间冲击响应均衡的均衡结果，汇总如表 6-4，通过利用临界频带宽度对响度转换滤波器的优化，均衡后的线性误差由 0.2917 降低为 0.2718，均衡比例由 88.1% 升为 88.8%，同样的均衡后的加权误差由 0.8861 降低为 0.8007，均衡比例由 92.3% 升为 93.1%。考虑到计算响度时已经利用了临界频带宽度的概念，所以均衡结果没有提升特别多，但是在利用其它加权函数进行均衡滤波器的优化时，临界频带宽度是必须要考虑的，优化效果也会更加明显。

表 6-4 各点房间冲击响应均衡结果

	$E1$	均衡比例 p	$E2$	均衡比例 p
房间冲击响应模型	2.4355		11.5368	
第五章均衡方法	0.2917	88.1%	0.8861	92.3%
第六章均衡方法	0.2718	88.8%	0.8007	93.1%

同样的，为了验证本章中基于人耳听觉特性的房间响应均衡效果，也需要进行主观的测听实验。与第五章实验方法相同，选用 10 名测听人员来进行均衡结

果主观测听实验，每名测听人员随机听取四段不同类型的声音频文件，每段音频文件又有五种不同的处理方法，最后根据主观听觉感受的好坏从0到10分打分，0分表示最差，10分表示最好，实验结果如表6-5。

表 6-5 主观测听结果

	音频 1	音频 2	音频 3	音频 4	平均
IIR 均衡方法	8.5	7.3	7.6	7.8	7.8
FIR 均衡方法	7.8	9.1	8.2	7.1	8.05
WIIR 均衡方法	7.2	8.3	7.5	8.8	7.95
Kautz 均衡方法	6.9	8.2	8.3	7.6	7.75
本文均衡方法	8.3	8.0	7.5	8.1	7.975

从表 6-5 中汇总的结果来看，主观测听实验结果依然不是非常理想，在五种均衡方法打分中仍然排在第二位，但是现在排在第一位的不是在第五章实验中排在第一位的 WIIR 均衡方法，而是 FIR 均衡方法，而 FIR 均衡方法在第五章的实验结果中排在第四位。正如在第五章的实验结果中分析的三个原因，主观实验的设计需要进一步的改进。

6.4 本章总结

本章利用临界频带宽度对第五章提出的响度转换滤波器进行优化，并给出了优化后的实验结果。本文提出的基于人耳听觉特性的均衡方法，关键是设计响度转换滤波器将误差从幅度转换为响度，力求从更符合人耳听觉特性的角度设计均衡滤波器和评价均衡结果。考虑到临界频带宽度 Bark 域是更适合分析人耳主观听觉的刻度，为了保证误差加权系数在一个 Bark 域内不发生跳变，另外为了提高在人耳更敏感的频带保证加权函数的精确性，本章利用临界频带宽度 Bark 进行响度转换滤波器的优化。实验结果表明，对响度滤波器进行优化确实可以在一定程度上改善均衡结果，提高了线性均衡比例和加权均衡比例。为了进一步提高均衡效果，仍然需要深入研究人耳听觉特性，以便在今后的研究中更好地应用人耳听觉特相关的知识来改善音频相关处理结果。

第七章 结论与展望

论文在三维音频迅速发展的背景下,针对房间冲击响应会影响音质和三维声音产生的沉浸感与包围感,提出了一种基于人耳听觉特性的房间冲击响应均衡方案。这样既解决了均衡有效区域存在局限性的问题,又可以将均衡结果由数学最佳转为听觉最佳,充分考虑了人耳听觉的主观性,以最大限度地满足在房间内某个区域的听觉需求。

论文首先介绍了有关房间冲击响应均衡的基本算法:房间单点均衡以及多点均衡;然后引入人耳听觉特性:听觉器官结构、掩蔽效应、临界频带宽度以及等响曲线等;进而进行了预处理相关工作的介绍,为后边的均衡工作提供数据并奠定基础;随后论文阐述了基于响度度量的房间冲击响应均衡,提出将幅度转化为响度进行均衡滤波器的设计;最后论文介绍了利用临界频带宽度对基于响度度量的房间冲击响应均衡进行优化改进的方法,进一步改善了均衡结果。

其中,预处理是整个均衡研究中最基本也是不可或缺的一部分,预处理部分完成了以下功能:

1. 房间冲击响应测量,得到房间冲击响应数据;
2. 房间冲击响应平滑,减小均衡结果随时间、位置变化的影响;
3. 房间冲击响应建模,得到能反映整个房间声学特性的房间冲击响应模型;

论文中最重要的部分是基于响度度量的房间冲击响应均衡的实现,这是得到听觉上最佳均衡结果的关键,这部分完成的工作有:

1. 响度转换滤波器的设计;
2. 最佳均衡滤波器系数求解;
3. 均衡实验及其分析;

另外,在均衡过程中考虑临界频带宽度,利用它去改进响度转换滤波器实现基于临界频带宽度的房间冲击响应均衡,可以防止出现加权系数在一个 Bark 域内发生跳变的现象得到更有效的响度转换滤波器,这部分完成的工作有:

1. 临界频带宽度 Bark 的学习;
2. 响度转换滤波器的优化;
3. 均衡实验及其分析;

实验结果表明,文章提出的基于人耳听觉特性的均衡方法在能取得理想均衡结果,尤其在人耳敏感的频段均衡效果较其他均衡方法有明显的优势。但针对具体应用仍有一些方面需要改进。

1. 房间冲击响应测量。在进行房间冲击响应测量时,考虑最多的是选择一个合适的源信号实现有噪声的情况下得到较为准确的房间冲击响应。但是不可否认的是测量过程中传声器或者麦克风的传输特性也应该考虑。如何将传声器或麦克风的影响消除也应该作为研究重点;

2. 房间冲击响应建模。无论是在频域进行数学方法的房间冲击响应建模还是在时域进行基于模式识别的房间冲击响应建模都不能很好地实现提取整个房间的声学特性。想要得到更准确的房间冲击响应模型,必须从每个点的房间冲击响应本身出发,认真分析并从声学特性角度进行建模;

3. 人耳听觉特性的应用。本文旨在从人耳听觉特性出发设计均衡滤波器和衡量均衡滤波器均衡效果,但是有关人耳听觉特性的应用不够完善。在以后的相关研究中,应当更加全面的调研人耳听觉特性进而能够更好的应用它来改善优化房间冲击响应均衡工作;

4. 非线性失真的考虑。利用数字信号处理知识只能解决房间冲击响应的线性失真问题,但是非线性失真也会影响重放系统性能、声音信号音质及听者各方面的感受,也需要列入房间冲击响应均衡的研究范畴。但是与线性失真相比,非线性失真问题更为复杂,值得深入得学习与分析;

5. 主观测听实验设计不够合理。主观测听实验并没有取得理想的结果,除了实验室中分析的原因外,主观测听实验设计的不合理也是影响最后结果的重要因素。实验中从测听人员到进行测听时利用的声源都应该经过严格的筛选,力求实验结果准确有参考价值。

参 考 文 献

- [1] Mourjopoulos, John. "Digital equalization methods for audio systems." Audio Engineering Society Convention 84. Audio Engineering Society, 1988.
- [2] Ifeachor, Emmanuel C., and Barrie W. Jervis. Digital signal processing: a practical approach. Pearson Education, 2002.
- [3] Greenfield, Richard, and Malcolm J. Hawksford. "Efficient filter design for loudspeaker equalization." Journal of the Audio Engineering Society 39.10 (1991): 739-751.
- [4] Karjalainen, Matti, et al. "Loudspeaker Response Equalization Using Warped Digital Filters." Proc. of NorSig-96. 1996.
- [5] Karjalainen, Matti, and Tuomas Paatero. "Equalization of loudspeaker and room responses using Kautz filters: Direct least squares design." EURASIP Journal on Applied Signal Processing 2007.1 (2007): 185-185.
- [6] S Cecchi, L Palestini, E Moretti, and F Piazza. A new approach to digital audio equalization. In Applications of Signal Processing to Audio and Acoustics, 2007 IEEE Workshop on, pages 62 – 65. IEEE, 2007.
- [7] Alfred Mertins, Tiemin Mei, and Markus Kallinger. Room impulse response shortening/reshaping with infinity-and-norm optimization. Audio, Speech, and Language Processing, IEEE Transactions on, 18(2):249 – 259, 2010.
- [8] 李贤胜.扬声器系统均衡及在线保护方法研究[J]. 北京:中国科学院大学声学研究所, 2015
- [9] S. Bharitkar and C. Kyriakakis, Immersive Audio Signal Processing. New York: Springer, 2006.
- [10] Jorgen Arendt Jensen. A new principle for an all-digital preamplifier and equalizer. In Audio Engineering Society Convention 80. Audio Engineering Society, 1986.
- [11] Peter M Clarkson, John Mourjopoulos, and JK Hammond. Spectral, phase, and transient equalization for audio systems. Journal of the Audio Engineering Society, 33(3):127–132, 1985.
- [12] John Mourjopoulos. Digital equalization methods for audio systems. In Audio Engineering Society Convention 84. Audio Engineering Society, 1988.

- [13] Soterios Salamouris, Kostas Politopoulos, Vasilis Tsakiris, and John Mourjopoulos. Digital system for loudspeaker and room equalization. In Audio Engineering Society Convention 98. Audio Engineering Society, 1995.
- [14] John Mourjopoulos, P Clarkson, and J Hammond. A comparative study of least-squares and homomorphic techniques for the inversion of mixed phase signals. In Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP'82. , volume 7, pages 1858–1861. IEEE, 1982.
- [15] Richard Greenfield and Malcolm J Hawksford. Efficient filter design for loudspeaker equalization. Journal of the Audio Engineering Society, 39(10):739–751, 1991.
- [16] Malcolm J Hawksford. Matlab program for loudspeaker equalization and crossover design. Journal of the Audio Engineering Society, 47(9):706–719, 1999.
- [17] Julius O Smith and Jonathan S Abel. The bark bilinear transform. In Applications of Signal Processing to Audio and Acoustics, 1995. IEEE ASSP Workshop on, pages 202–205. IEEE, 1995.
- [18] Julius O Smith and Jonathan S Abel. Bark and ERB bilinear transforms. Speech and Audio Processing, IEEE Transactions on, 7(6):697–708, 1999.
- [19] Tuomas Paatero, Matti Karjalainen, and Aki Harna. Modeling and equalization of audio systems using kautz filters. In Acoustics, Speech, and Signal Processing, 2001. Proceedings. (ICASSP'01). 2001 IEEE International Conference on, volume 5, pages 3313–3316. IEEE, 2001.
- [20] Matti Karjalainen and Tuomas Paatero. Equalization of audio systems using kautz filters with log-like frequency resolution. In Audio Engineering Society Convention 120. Audio Engineering Society, 2006.
- [21] Matti Karjalainen and Tuomas Paatero. Equalization of loudspeaker and room responses using kautz filters: Direct least squares design. EURASIP Journal on Applied Signal Processing, 2007(1):185–185, 2007.
- [22] German Ramos and Jose Javier Lopez. Direct method with random optimization for parametric IIR audio equalization-applications to one way and multi-way systems. In Audio Engineering Society Convention 116. Audio Engineering Society, 2004.
- [23] Germán Ramos, José J López, and Jaime Lloret. Direct method with random optimization for loudspeaker equalization using IIR parametric filters. In

- Acoustics, Speech, and Signal Processing, 2004. Proceedings. (ICASSP'04). IEEE International Conference on, volume 4, pagesiv–97. IEEE, 2004.
- [24] German Ramos and Jose Javier Lopez. Subjective evaluation of an equalization method for loudspeakers based on random parametric optimization of IIR filters. In Audio Engineering Society Convention 116. Audio Engineering Society, 2004.
- [25] Jose Lopez and German Ramos. Low order IIR parametric loudspeaker equalization, a psychoacoustic approach. In Audio Engineering Society Convention 118. Audio Engineering Society, 2005.
- [26] German Ramos and Jose J Lopez. Filter design method for loudspeaker equalization based on IIR parametric filters. *Journal of the Audio Engineering Society*, 54(12):1162–1178, 2006.
- [27] Balázs Bank. Direct design of parallel second-order filters for instrument body modeling. In Proc. International Computer Music Conference (ICMC2007), Copenhagen, Denmark, pages 458–465, 2007.
- [28] Balázs Bank. Perceptually motivated audio equalization using fixed-pole parallel second-order filters. *Signal Processing Letters, IEEE*, 15:477–480, 2008.
- [29] Balázs Bank and Germán Ramos. Improved pole positioning for parallel filters based on spectral smoothing and multiband warping. *Signal Processing Letters, IEEE*, 18(5):299–302, 2011.
- [30] Balázs Bank. Audio equalization with fixed-pole parallel filters: An efficient alternative to complex smoothing. *Journal of the Audio Engineering Society*, 61(1/2):39–49, 2013.
- [31] J. Mourjopoulos, P. Clarkson, and J. Hammond, “A comparative study of least-squares and homomorphic techniques for the inversion of mixed phase signals,” in Proc. ICASSP 1982, Int. Conf. Acoust., Speech, Signal Process., Paris, France, May 1982, pp. 1858–1861.
- [32] M. Miyoshi and Y. Kaneda, “Inverse filtering of room acoustics,” *IEEE Trans. Signal Process.*, vol. 36, no. 2, pp. 145–152, Feb. 1988.
- [33] S. J. Elliot and P. A. Nelson, “Multiple-point equalization in a room using adaptive digital filters,” *J. Audio Eng. Soc.*, vol. 37, no. 11, pp. 899–907, Nov. 1989.
- [34] Y. Haneda, S. Makino, and Y. Kaneda, “Multiple-point equalization of room

- p transfer functions by using common acoustical poles,” IEEE Trans. Speech Audio Process., vol. 5, no. 4, pp. 325–333, Jul. 1997.
- [35] J. Mourjopoulos, “Digital equalization of room acoustics,” J. Audio Eng. Soc., vol. 42, no. 11, pp. 884–900, Nov. 1994.
 - [36] S. Spors, H. Buchner, R. Rabenstein, and W. Herbordt, “Active listening room compensation for massive multichannel sound reproduction systems using wave-domain adaptive filtering,” J. Acoust. Soc. Amer., vol. 122, no. 1, pp. 354–369, Jul. 2007.
 - [37] S Cecchi, L Palestini, E Moretti, and F Piazza. A new approach to digital audio equalization. In Applications of Signal Processing to Audio and Acoustics, 2007 IEEE Workshop on, pages 62–65. IEEE, 2007.
 - [38] Alberto Carini, Ivan Omicciuolo, and Giovanni L Sicuranza. Multiple position room response equalization: Frequency domain prototype design strategies. In Image and Signal Processing and Analysis, 2009. ISPA 2009.Proceedings of 6th International Symposium on, pages 638–643. IEEE, 2009.
 - [39] S Cecchi, L Palestini, P Peretti, F Piazza, and A Carini. Multipoint equalization of digital car audio systems. In Image and Signal Processing and Analysis, 2009. ISPA 2009. Proceedings of 6th International Symposium on, pages 650–655. IEEE, 2009.
 - [40] Carini A, Cecchi S, Piazza F, et al. Multiple position room response equalization in frequency domain [J]. Audio, Speech, and Language Processing, IEEE Transactions on, 2012, 20(1): 122-135.
 - [41] 谢菠荪.头相关传输函数与虚拟听觉.国防工业出版社,2008.
 - [42] Moore, Brian CJ. "Frequency analysis and masking." Hearing (1995): 161-205.
 - [43] Fletcher H.1940. Auditory patterns. Rev. Mod. Psys. ,12(1);47-65.
 - [44] Scharf, Bertram. "Critical bands." Foundations of modern auditory theory 1 (1970): 157-202.
 - [45] Sek, Aleksander, and Brian CJ Moore. "The critical modulation frequency and its relationship to auditory filtering at low frequencies." The Journal of the Acoustical Society of America 95.5 (1994): 2606-2615.
 - [46] Fastl, Hugo, and Eberhard Zwicker. Psychoacoustics: Facts and models. Vol. 22. Springer Science & Business Media, 2007.

- [47] Fastl, Hugo, and Eberhard Zwicker. "Psychoacoustics: facts and models." (2001).
- [48] Moore, Brian CJ, and Brian R. Glasberg. "Formulae describing frequency selectivity as a function of frequency and level, and their use in calculating excitation patterns." *Hearing research* 28.2 (1987): 209-225.
- [49] Scharf, Bertram. "Critical bands." *Foundations of modern auditory theory* 1 (1970): 157-202.
- [50] Moore, Brian CJ, Robert W. Peters, and Brian R. Glasberg. "Auditory filter shapes at low center frequencies." *The Journal of the Acoustical Society of America* 88.1 (1990): 132-140.
- [51] Glasberg, Brian R., and Brian CJ Moore. "Derivation of auditory filter shapes from notched-noise data." *Hearing research* 47.1 (1990): 103-138.
- [52] American National Standards Institute, "American national psychoacoustical terminology" S3.20, 1973, American Standards Association.
- [53] Olson, Harry F. (February 1972). "The Measurement of Loudness" (PDF). *Audio*: 18–22.
- [54] Lenk, John D. (1998). *Circuit Troubleshooting Handbook*. McGraw-Hill. p. 163. ISBN 0-07-038185-2.
- [55] Muller S, Massarani P (2001), *J. Audio Eng. Soc.*, 49(6):443–471.
- [56] Schroeder MR (1975), *J. Acoust. Soc. Amer.*, 57:149–150.
- [57] Farina A (Apr. 2000), 108th Conv. of Audio Eng. Soc. (preprint 5093).

硕士期间发表的论文与研究成果

- [1].**Qianqian FANG**, Huaxing XU, Risheng XIA, Junfeng LI and Yonghong YAN.
“Equalization of sound reproduction system based on the human perception characteristics” in IIHMSP-2015, Adelaide, Austria, Sep. 2015. (EI)
- [2].房倩倩, 徐华兴, 李军锋, 夏日升, “基于人耳听觉特性的音频系统均衡” 声学技术, 2015,12. (核心期刊)

致 谢

时间总在不经意间悄悄从指缝中溜走，第一次走进声学所的场景还是历历在目，一转眼，三年的研究生生活却已经接近尾声。总结自己几年间的成长与收获，从一个懵懂的初学者接触工程研究到现在经历锻炼，积累了一定的科研学习方法。在此过程中，有很多人给过我无私的帮助和支持，即将完成硕士论文之际，我向他们致以最真挚的感谢！

首先特别感谢实验室主任颜永红老师，颜老师知识渊博，工作认真负责，一言一行都对我产生了深刻的影响。无论什么时间来实验室似乎都能看到您，而且每次都是活力满满、斗志昂扬的状态，您是我一生学习的榜样。同时感谢颜老师为学生提供了一流的工作和学习环境，让我们无后顾之忧地科研学习。还要感谢潘接林老师、徐为群老师、应冬文老师等等曾经指导过我、帮助过我的实验室每一位老师。您们对工作兢兢业业、对学生热情谦和，用心地为我们解答每一个疑惑、解决每一个困难，为我们的科研学习保驾护航。

感谢我的导师李军锋老师。从进入实验室开始，无论是学术上还是工程上您时刻都关心着我，从论文的选题、撰写、修改到定稿，都凝聚了您的心血。三年间，李老师的悉心指导和帮助让我如沐春风、受益终身。您渊博的学识以及乐观向上的工作态度，对我影响深远。借此机会，向李老师表达我最衷心的感谢！

感谢夏日升老师，您不仅在学习研究中给了我很多教诲和启发，也在生活中给了我无微不至的关怀。每当我遇到困难时，夏老师悉心指导，让我一步步取得进步，提高了理论实践等多方面能力。

感谢徐华兴、王志超、姚鼎鼎等师兄师弟们，实验室偶尔枯燥乏味的生活因为有你们也变得十分有意义，一起讨论一起想办法出主意的时光点亮了我的科研生活，用李老师的话就是“we are 伐木累”。

感谢梁好、王惊亦、陈艳霞、姚海涛等同窗好友们，很幸运能够与你们分到一个班、一个实验室或是一个宿舍，一起经历怀柔集中教学，一起经历找工作等各种或美好或苦难的事情，培养了别人读不懂看不透的默契，革命友谊铁铁的。

感谢张敏、安冬、陈钦钦、吴金晓等小伙伴们，从本科到现在，跟你们的友谊是我的人生财富，当时伴着《凤凰花开的路口》大学毕业，时间长河往前流，

咱们却没有分头走，有你们真好。

感谢我的男朋友郝文昌，谢谢你几年如一日地照顾与包容，正能量满满的你总是能在我遇到困难时给予我鼓励和支持，漫漫人生路还有很多风雨和风景，我们一起面对。

最为感谢的是我的父母、我的家人，谢谢你们多年来对我无私的爱，正是因为你们的养育与爱护，无论遇到什么困难坎坷我都觉得有所依靠，在以后的日子里开始换我来做你们的臂膀，为你们遮风挡雨。

毕业论文的完成也标志了我学习生涯即将结束，借此机会，我也向所有教育过我的老师们致以最衷心的感谢，是你们给予我理想、知识以及学习能力，让我有了更好的发展机会。

最后，再次向所有给予本人关怀与帮助的老师、朋友和为审阅本论文而付出辛勤劳动的专家学者表示诚挚的感谢！