

耳蜗滤波器的设计和 DSP 实现

赵 鹤 鸣

(苏州大学电子工程系, 苏州 215006)

1995 年 6 月 2 日收到

1995 年 9 月 6 日定稿

摘要 耳蜗的主要功能之一相当于一组并联的且频率特性重叠的带通滤波器, 该滤波器组的设计用一般数字滤波器设计方法难以奏效, 本文提出该滤波器组的设计方法, 并在数字信号处理器 (DSP) 上加以实现。

Design and DSP realization of cochlear filter

ZHAO Heming

(Department of electronic engineering, Suzhou University, Suzhou 215006)

Received Jun. 2, 1995

Revised Sept. 6, 1995

Abstract The main function of cochlear filter is a set of bandpass filters, which are paralleled and overlapping in frequency response. It is very difficult to design this filter set by means of normal method. In this paper, the author suggest a new way to implement it on DSP.

一、引 言

耳蜗是接受声音刺激的听觉器官, 耳蜗的基底膜 (BM) 对外来声信号有频率选择和调谐作用, 对不同的中心频率, 在相应临界频带 (critical band) 内的信号会引起基底膜上不同位置的振动。听觉系统依赖于它将声波转变为代表声音信息的神经冲动, 在整个复杂的听觉功能中, 耳蜗对声音的感受和换能作用是非常重要的第一环节。随着语音信号处理研究的不断深入, 利用人耳的听觉特性建立听觉模型并用于语音识别是当前人们深入研究的热点之一, 而耳蜗功能模型则是听觉模型中最主要的一部分。

研究表明^[1,2], 耳蜗的主要功能之一是一种串 / 并转换器的功能, 它相当于一组并联的且频率特性重叠的带通滤波器, 串行输入的声音信号在耳蜗中被分解并以多路并行的方式输出。

耳蜗滤波器的频率特性可从耳蜗模型^[3]或用实验方法^[1,4]得到。本文所设计的耳蜗滤波器由 24 个带通滤波器组成, 它相应于 20 Hz 至 16 kHz 听阈内的 24 个临界频带, 并以临界频带率 z (critical band rate, 单位为 Bark) 标度划分, 其频率特性则以德国著名声学专家 Zwicker 给出的频率响应曲线为依据^[1,5]。与一般数字滤波器的设计 (满足给定的几个指标) 不同, 该组滤波器以拟合耳蜗滤波频率特性曲线为设计准则, 所以, 一般数字滤波器设计方法及 CAD 设计软件难

以满足要求或设计结果阶数太高而不能实用,为此,本文提出了该组滤波器的设计方法,并在数字信号处理器 (DSP) 上加以实现,该耳蜗滤波器作为听觉模型的一部分构成了一个实用的语音信号处理系统。

二、设计预处理

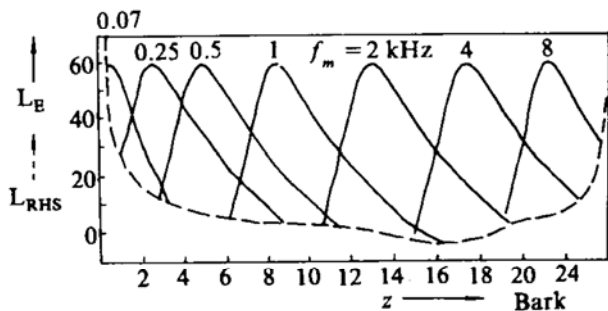


图 1 耳蜗滤波器的频率特性曲线

图 1 给出了耳蜗滤波器对不同中心频率激励 (强度为 60 dB) 下以 Bark 为标度的频率响应曲线,虚线对应听阈临界曲线。如前所述,耳蜗滤波器的设计应以拟合该组重叠的带通滤波器的频率特性曲线为基本考虑。为此,需将该组曲线转换成设计滤波器所需的对应衰减特性曲线,图 2 例示了中心频带为 10 Bark 的滤波器对应的衰减特性曲线的构造方法。由于临界频带在 Bark 标度上的带宽几乎是恒定的 (这也是耳蜗滤波器频率特性常用 $z(\text{Bark})$ 而不用 $f(\text{Hz})$ 标度的原因之一),

所以,不同临界频带率 z 上的衰减量可由中心频带对应的曲线水平移动测得的衰减量来决定,例如,图中 8 Bark 点对应的衰减量为 a_8 。由此方法反复进行即可得到构成衰减特性曲线所需的各临界频带上的衰减量。

显然,用此方法只能在有限个点上给出衰减量,而数字滤波器的设计需要衰减特性曲线的函数表达,为此,利用三次样条内插方法加以实现,即对于 $n+1$ 个样点 $\{a_j, z_j\}$ 构造 n 个最高阶数为 3 的多项式,并且根据临界频带率 z 与频率 f 的关系^[1] 构成实际设计滤波器所需的衰减特性函数。

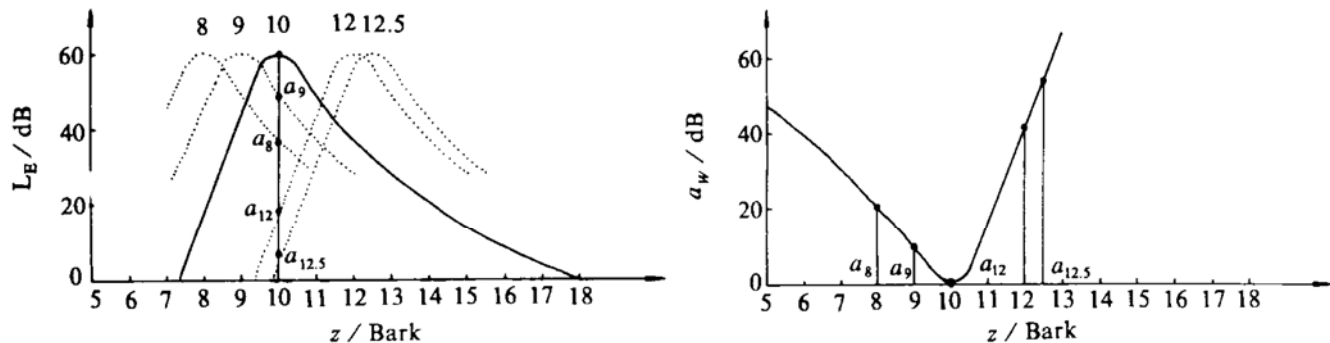


图 2 滤波器衰减特性曲线的构造示例

在滤波器的设计中,因滤波器阶数的限制,衰减特性函数的精确实现是不可能的,为此要确定其容差范围,方法是:将衰减特性函数分段设置容差,中心频带邻域误差小于 2 dB,止带衰减量大于 72 dB,并保证整个函数值容差的上下限不超过由激励强度为 40 dB 和 80 dB (与一般语音强度范围对应) 所得的衰减函数值。

三、设计方法

为使滤波器所需的计算量尽可能小,必须对各种可能的滤波器结构加以选择。首先考虑采样频率 f_s ,由于数字滤波器的计算量随 f_s 平方数量级增长,而最高临界频带 24 Bark 对应 16 kHz,所以 f_s 应大于 32 kHz,但对耳蜗滤波器组中的绝大多数而言,这是不必要的,为此,对各滤波器

采用不同的采样频率 $f_{si} = f_s/2^m$ (m 为正整数) 进行设计, 也即各滤波器的上限频率小于 $f_{si}/2$, 为此需另加低通滤波器来保证。基于上面的考虑, 耳蜗滤波器可有以下两种不同的结构: (a) 低通 (LP)–高通 (HP) 级联; (b) 半带 (HB: Halfband)–带通 (BP) 级联, 如图 3 所示。

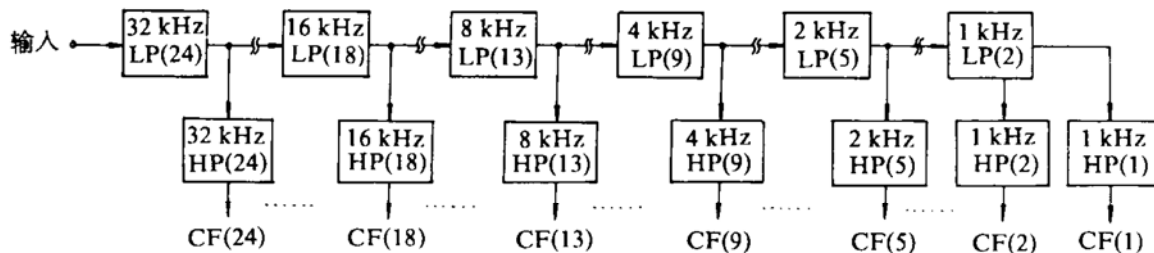


图 3(a) 耳蜗滤波器的 LP-HP 结构

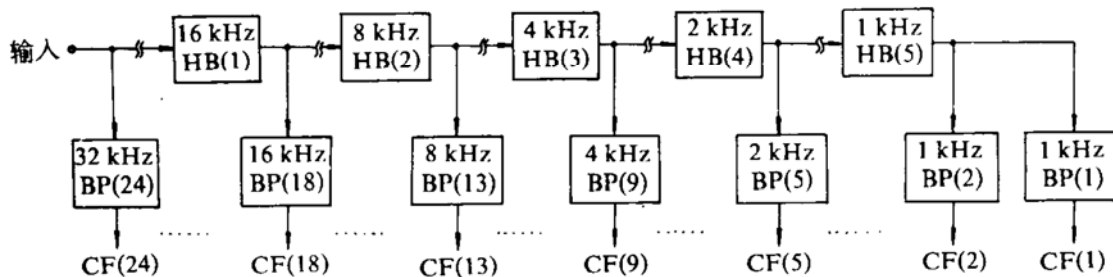


图 3(b) 耳蜗滤波器的 HB-BP 结构

LP-HP 结构存在以下几个缺点: (a) 前级 LP 的误差将传到下级; (b) 一个 LP 频率特性的改变将影响后续各级的特性; (c) 各 LP(i) 的阶数较高。为此, 选用图 3(b) 的结构。

对于图 3(b) 所示的结构, 需根据本文第二节中给出的频率特性 (为叙述方便称为理想曲线) 设计一组数字带通滤波器, 并使其频响曲线应尽可能在整个频率范围内与理想曲线一致。为此, 首先尝试用线性预测法设计递归型数字滤波器 (全极点模型), 该方法需先将理想频响特性转变为时域冲激响应 $h(n)$, 这里采用快速傅里叶反变换 IFFT 来实现, N 取得足够大, 以考虑整个频率特性对应的频谱。图 4 例示了 6 阶线性预测器设计的中心频带为 10 Bark 对应的 BP。

由图 4 可知, 该滤波器通带内存在过冲而与理想曲线偏差很大, 其原因为线性预测系数 (滤波器系数) 是基于整个频段内使预测误差最小来优化确定的。为改善特性, 采用极零模型线性预测方法, 而这种方法同时对滤波器零点系数和极点系数优化是难以做到的, 为此采用次优化设计方法。在我们的设计中, 根据选择性线性预测原理^[6], 提出选择性频谱线性预测 (SSLP) 设计算法, 以保证通带内频响特性的最佳逼近, 算法如下:

(1) 对冲激响应 $h(n)$ 进行 N 点 FFT, 并计算功率谱。

(2) 选择功率谱的一部分 (L 点, L 小于 N) 作为新的谱, 并通过 L 点 IDFT 计算自相关系数。

(3) 用线性预测方法计算滤波器系数。

显然, 上述算法保证了在选择谱范围内预测误差最小, 因而改善了原来设计中通带内的滤波特性。对于选择谱所选的频带宽度, 一方面考虑整个滤波器频响特性的大部分通过 SSLP 算法能很好逼近, 选择谱要选得宽; 另一方面为避免一般线性预测设计滤波器所遇到的问题, 选择谱要选得窄。这两者是矛盾的, 我们通过反复循环, 把设计结果与理想曲线对比求得最优结果。

通过上面 SSLP 算法设计的递归滤波器逼近了整个带通 (BP) 滤波器频率特性的主要部分, 我们把 BP 理想频率特性曲线与该递归滤波器频率特性曲线的差称为剩余谱, 其频率特性相对平坦, 为此同样采用 SSLP 设计方法用一个 m 阶的非递归滤波器加以逼近。这样, 整个带通滤波器

用一个 n 阶递归滤波器和一个 m 阶非递归滤波器加以实现。图 5 为中心频带为 10 Bark BP 对应的理想频率特性 (虚线) 和设计结果 (实线)。

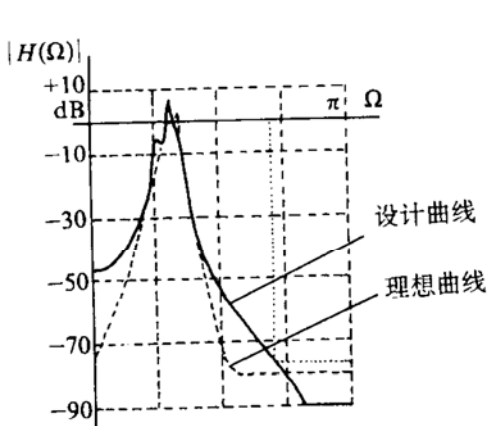


图 4 中心频带为 10 Bark 对应的理想曲线与 6 阶线性预测器设计的滤波器频响曲线

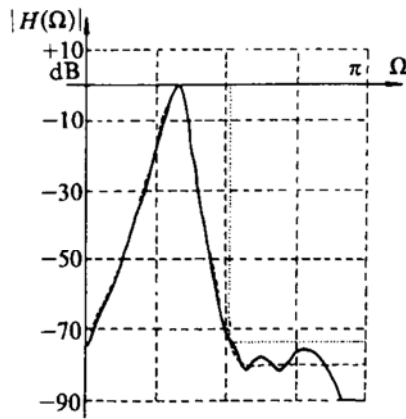


图 5(a) 中心频带为 10 Bark BP 对应的理想特性 (虚线) 和所设计滤波器 ($n=6, m=9$) 的逼近结果

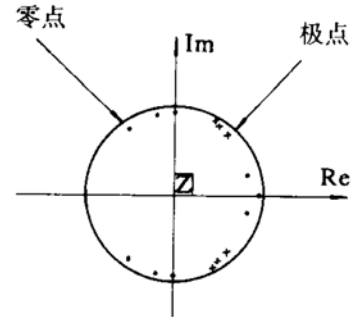


图 5(b) 所设计滤波器的极零点分布

图 3(b) 中半带 (HB) 滤波器的设计比较简单, 考虑到该组滤波器应具线性相位, 所以采用 FIR 型滤波器并用 McClellan-Parks 算法设计实现。

四、耳蜗滤波器的 DSP 实现

以上所设计的滤波器在数字信号处理器上实现, 主要应考虑两个问题: 一是实时处理; 二是当 DSP 为定点处理时的有限字长效应。要使系统能实时处理, 即两个 A/D 采样间隔内能完成所有滤波计算, 必须估计其运算量: 设平均采样率为 16 kHz, DSP 运算速度为 10 MIPS (与 TMS320C25 对应), 则每通道 (与临界频带相对应, 共 24 个) 在两个采样点间隔内允许的运算次数为 26 次, 若每阶滤波需 3 ~ 4 次运算, 那么, 每通道滤波器总阶数为 6 ~ 8, 在我们以上设计的耳蜗滤波器中, 每通道滤波器的总阶数不超过 17, 所以, 采用高速 DSP (例 TMS320C50 运算速度达 28.6 MIPS) 或用几个 DSP 构成多信号处理器系统即能实时实现。

尽管新一代 DSP 字长越来越长, 但当 DSP 为定点 (一般为 16 位或 32 位) 处理时, 仍必须考虑其有限字长效应。为此, 我们通过在计算机上模拟各滤波器系数舍入或截尾处理后对滤波性能的影响, 根据通带误差最小, 止带衰减量最大的优化原则, 选择一组最佳量化系数。

五、应用及讨论

用上述方法设计和实现的耳蜗滤波器作为听觉模型^[7]的主要部分用于限定人、较大词汇表的连续汉语语音识别。语音特征参数除选用听觉模型输出作为主要参数外, 还辅以过零率参数, 识别单元为声、韵母, 识别算法 (另文介绍) 采用基于知识的多级识别策略, 对二十三个连续语句 (共一百四十二个字, 含汉语全部声、韵母) 的正确识别率为 86.6%。同样的系统, 如采用典型的 LPC 系数 (14 阶) 则识别率有所下降, 主要是声母的误识增多。

对语音识别而言, 用听觉感知相关参数表示语音特征比用语音生成模型参数 (如 LPC 系数) 更接近人类对声音感受和语音辨识的实际情况, 因而这方面的研究也越来越为人们所重视, 但目前的研究结果尚未表明前者的识别率已有显著的提高, 这一方面有待于对听觉感知本质研究的进一步深入; 另一方面也受实现技术的限制。在本文提出的耳蜗滤波器设计方法中, 为避免非线性

运算, 对耳蜗声频识别与强度的非线性关系 (如带通滤波器的频响特性随不同强度的激励有一定的变化等) 近似为线性关系, 这对语音识别会有所影响。

六、总 结

本文所设计的耳蜗滤波器虽把实际耳蜗对声音的处理作了简化, 但仍能较好地体现其对语音的辨识功能。另外, 本文提出的数字滤波器设计方法也具有实用意义。

致 谢

感谢德国慕尼黑科技大学教授 E. Zwicker 博士和 G. Ruske 博士给予的帮助和指导。

参 考 文 献

- [1] Zwicker E. Psychoakustik, Springer-Verlag, 1982.
- [2] Allen J B. Cochlear modeling, *IEEE ASSP Magazine*, 1985, 1: 3 ~ 29.
- [3] 万旺根、樊昌信. 一种新的二维耳蜗模型, 通信学报, 1991, 12: 24 ~ 29.
- [4] Rhode W S and Robles L. Evidence from mossbauer experiments for nonlinear vibration in the cochlear, *J. Acoust. Soc. Am.*, 1974, 55: 588 ~ 596.
- [5] Entenmann W. Handbuch für das Funktionsmodell des Gehörs, Technische Universität München, 1972.
- [6] Markel J and Gray A. Linear prediction of speech, Springer-Verlag, 1976.
- [7] ZHAO Heming and ZHOU Xudong. A new acoustic perceptual model and its applications, *Journal of electronics*, 1995, 12: 73 ~ 78.