

适用于助听器语音信号处理的 DFT 调制滤波器组设计

蔡 宇^{1,2}, 侯朝焕¹, 原建平¹, 鄢社锋¹

(中国科学院声学研究所, 北京 100190; 2. 中国科学院研究生院, 北京 100049)

摘要: 提出了一种近似重构、过采样 DFT 调制滤波器组的设计方法。根据滤波器组的系统失真和输出混迭方面的要求, 构造一个简单的目标函数, 并采用等纹波法优化设计原型滤波器。与现有方法相比, 该设计方法将复杂的非线性优化问题转化为对单一参数的选取, 简化了设计过程, 并且可在满足 DFT 调制滤波器组设计准则的条件下使得通带范围内误差逼近最小。利用加权叠接相加(Weighted Overlap and Add, WOLA)的快速实现结构, 使整个系统更加高效、易行, 可用于助听器语音信号分频带处理与综合。

关键词: DFT 调制滤波器组; 等纹波法; WOLA 结构; 语音信号处理

中图分类号: TN713

文献标识码: A

文章编号: 1000-3630(2012)-05-0517-05

DOI 编码: 10.3969/j.issn1000-3630.2012.05.014

Design of DFT modulated filter bank for speech processing in hearing aids

CAI Yu^{1,2}, HOU Chao-huan¹, YUAN Jian-ping¹, YAN She-feng¹

(1. Institute of Acoustics, Chinese Academy of Sciences, Beijing 100190, China;

2. Graduate University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract: This paper proposes a simple method of designing an oversampled, near-perfect reconstruction DFT modulated filter bank. According to the requirement of system distortion and aliasing power, a simple objective function is introduced, and the prototype filter is designed with the equiripple method. Compared with traditional methods, the proposed method optimizes a single parameter instead of the complex nonlinear optimization, and the design process is simplified. This approach can minimize the passband error under the filter-bank design criteria. By using weighted overlap and add (WOLA) structure, the system can be implemented efficiently, which is suitable for speech signal analysis and synthesis in hearing aids.

Key words: DFT modulated filter bank; equiripple method; WOLA structure; speech signal processing

0 引言

滤波器组和频谱分析、综合的概念在科学和工程的许多领域中都要用到, 在语音分析、频带压缩、通信等系统中更是尤其重要。DFT 调制滤波器组是一类重要的、易于实现的滤波器组, 它的概念主要是从对一些准平稳信号如语音信号进行频谱分析和综合中发展起来的^[1]。DFT 调制滤波器组具有设计简单的特点: 它的分析和综合滤波器各只需要一个原型滤波器, 其他各通道滤波器可由该原型滤波器调制得到。DFT 调制滤波器组还具有一个重要的

优点就是它具有高效的实现方式。加权叠接相加(Weighted Overlap and Add, WOLA)即是这样的一种结构, 它利用 FFT 技术, 大大减小了计算量, 适合语音信号的块处理方式和处理要求, 可以方便地在 DSP 上实现。

原型滤波器的设计很大程度上影响了滤波器组的性能, 为了避免混叠、失真等, 滤波器的设计必须满足一定的准则。目前大多采用的一类方法是构造一个目标函数, 最优原型滤波器的设计最终转化成为了一个有约束或无约束的非线性优化问题, 其求解过程非常复杂^[2,3]。另一类采用迭代算法, 由两个原型滤波器确定 DFT 调制滤波器组, 但很难保证找到全局最优解^[4,5]。在语音信号处理应用中, 一些学者也考虑从选择目标函数和减少约束条件来简化设计方法^[6-9]。

在本文中, 设计了一种近似重构、过采样的

收稿日期: 2011-08-07; 修回日期: 2011-12-01

基金项目: 国家自然科学基金资助项目(60802072)

作者简介: 蔡宇(1985-), 男, 四川自贡人, 博士研究生, 研究方向为语音信号处理。

通讯作者: 蔡宇, E-mail: caiyu@mail.ioa.ac.cn

DFT 调制滤波器组, 分析滤波器组和综合滤波器组各由一个原型滤波器产生, 该原型低通滤波器采用等纹波法设计, 将复杂的非线性优化问题简化为对通带截止频率单一参数的选取。在满足滤波器组设计准则的条件下, 通过加权可使得通带范围内误差逼近最小, 并且可以减小滤波器长度。在设计中运用了 WOLA 结构, 实现更高效、易行。整个系统适用于助听器语音信号的分频带处理与综合。

1 DFT 调制滤波器组原理

DFT 调制滤波器组的结构如图 1 所示, $h_0(n)$ 为分析原型低通滤波器, 占有带宽 $2\pi/K$, K 为通道总数, M 为降采样因子。当 $M=K$ 时, 称为临界采样; $M<K$ 时, 称为过采样。临界采样的 DFT 调制滤波器组很难获得满意的阻带衰减, 所以经常使用设计灵活、噪声抑制性能更好的过采样滤波器组。

记 $W_K = \exp(-2\pi j/K)$ 为第 k 条支路上的分析滤波器 $h_k(n)$ 为 $h_0(n)$ 被 W_K^{kn} 调制产生。滤波器组分析过程相当于输入信号 $x(n)$ 被 W_K^{kn} 调制后经过低通滤波器 $h_0(n)$ 滤波, 然后再将抽样率降低 M 倍后产生通路信号:

$$X_k(m) = \sum_{n=-\infty}^{\infty} h_0(Mm-n)x(n)W_K^{kn} \quad (1)$$

在没有任何通道处理的情况下, DFT 综合滤波器组先将通路信号 $X_k(m)$ 进行 M 倍的内插和低通滤波变回原来的抽样率, 然后再调制回原来的谱位置, 最后将他们加和输出单一信号:

$$y(n) = \sum_{m=-\infty}^{\infty} f_0(n-Mm) \sum_{k=0}^{K-1} X_k(m)W_K^{-kn} \quad (2)$$

式(2)中, $f_0(n)$ 是综合原型滤波器, 通道滤波器 $f_k(n) = f_0(n)W_K^{-kn}$, $y(n)$ 是输出信号。

从频域考虑, 背对背滤波器组的输入输出关系为

$$\begin{aligned} Y(z) &= \sum_{k=0}^{K-1} F_k(z) \frac{1}{M} \sum_{l=0}^{M-1} X(zW_M^l) H_k(zW_M^l) = \\ & \frac{1}{M} \sum_{k=0}^{K-1} F_k(z) H_k(z) X(z) + \\ & \frac{1}{M} \sum_{l=1}^{M-1} \left[\sum_{k=0}^{K-1} F_k(z) H_k(zW_M^l) \right] X(zW_M^l) \end{aligned} \quad (3)$$

式(3)中, 第一项表示希望得到的信号, 而第二项表示混迭分量, 应想办法去除。令

$$T(z) = \frac{1}{M} \sum_{k=0}^{K-1} F_k(z) H_k(z) \quad (4)$$

$$A(z) = \frac{1}{M} \sum_{l=1}^{M-1} \sum_{k=0}^{K-1} F_k(z) H_k(zW_M^l) \quad (5)$$

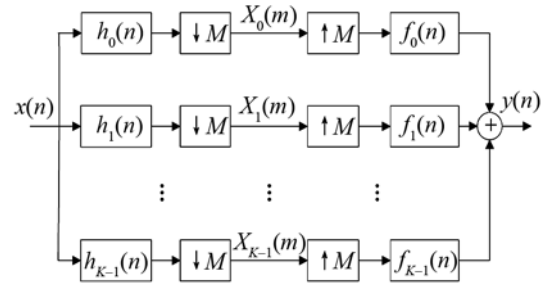


图 1 DFT 调制滤波器组结构

Fig.1 The structure of DFT modulated filter bank

$T(z)$ 称作系统的失真传递函数, $A(z)$ 称作系统的输出混迭函数^[10]。在近似重构滤波器组设计中, 希望失真传递函数尽可能接近于全通, 而输出混迭函数尽可能接近于零。特别地, 若失真传递函数具有纯延迟的形式, 即 $T(z) = cz^{-k}$, 滤波器组将实现准确重构。准确重构是一个很强的约束, 往往以牺牲滤波器组的整体性能为代价, 而在语音信号处理等一些特定应用中, 近似重构的滤波器组能够获得更好的效果。

2 原型滤波器的设计

在原型滤波器的设计过程中, 一般考虑的指标是系统传递失真和输出混迭失真, 使得设计的滤波器组能满足重构要求, 并且通带纹波小、阻带衰减大。正如前 1 节所述, 需要满足的条件为

$$\frac{1}{M} \sum_{k=0}^{K-1} F_k(e^{j\omega}) H_k(e^{j\omega}) = 1 \quad (6)$$

$$\text{且 } \sum_{k=0}^{K-1} F_k(e^{j\omega}) H_k(e^{j\omega} W_M^l) = 0, \quad l=1, 2, \dots, M-1 \quad (7)$$

在助听器语音信号处理中, 一般采用的是各通道少量交叠的结构, 即只有两个相邻滤波器叠接, 那么在任意频率上, 式(6)中只有两项相互作用。故在实际设计中只需要考虑前两个通道的传递失真条件就可以了。本文提出方法采用 Parks-McClellan 方法设计等纹波(最佳)原型滤波器, 在滤波器长度、阻带截止频率和误差加权比固定的情况下, 调整滤波器通带截止频率, 使其满足式(6)和式(7)所示条件。优化目标函数选取为

$$\begin{aligned} \phi &= \max_{\omega} \|F_0(e^{j\omega}) H_0(e^{j\omega}) + |F_1(e^{j\omega}) H_1(e^{j\omega}) - 1|, \\ & 0 < \omega < \frac{2\pi}{K} \end{aligned} \quad (8)$$

其中, ϕ 的物理意义表示在前两个滤波器组的叠接范围内, 系统输出幅度失真的最大值(在一个实际语音系统中, 如果滤波器有些时延或相位失真, 考虑幅度失真的条件就足够了)。优化过程就是通过

选取合适的通带截止频率 ω_p ，使得目标函数 ϕ 最小。理论经验表明，在一定频率范围内，如此构造的目标函数 ϕ 应是通带截止频率 ω_p 的凸函数。图2显示了目标函数与通带截止频率关系的一个实例，其中分析和综合滤波器使用同一个原型滤波器，通道数为 $K=16$ ，滤波器长度为 $N=128$ ，阻带截止频率为 $\omega_s=2\pi/K$ ，等比误差加权。事实上，在第一个分析滤波器占有带宽内，即 $0\sim\pi/K$ 的频率范围内，总能找到目标函数的一个全局最优解。解此类优化问题可用多种方法^[11]，这里就不赘述了。

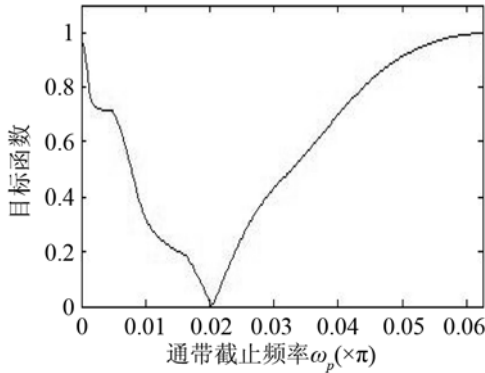


图2 目标函数与通带截止频率的关系
Fig.2 Objective function versus the passband edge

为了保证式(7)所述的避免混迭失真的条件，要求分析滤波器和综合滤波器的阻带截止频率分别满足^[1]：

$$\omega_{sh} \leq \frac{\pi}{M}, \quad \omega_{sf} \leq \frac{\pi}{M} \quad (9)$$

考虑到滤波器组通道间叠接且只有两个相邻滤波器叠接的要求，需将式(9)中对分析滤波器的阻带截止频率限制为：

$$\frac{\pi}{K} < \omega_{sh} \leq \frac{2\pi}{K} \quad (10)$$

而当 ω_{sh} 明显小于 π/M 时，混迭的分量间出现了间隙，这就放宽了对式(9)中对综合滤波器的阻带截止频率的要求，只要满足式(11)就可以了。

$$\omega_{sh} + \omega_{sf} \leq \frac{2\pi}{M} \quad (11)$$

总之，一般情况下本文提出的原型滤波器设计方法可分为以下几步：

步骤一：确定滤波器组的通道数 K 和降采样率 M ，满足过采样的条件 $M < K$ 。

步骤二：根据式(10)~(11)选择分析滤波器和综合滤波器的阻带截止频率 ω_{sh} 和 ω_{sf} ，选取适当的滤波器长度 N_h 和 N_f ，根据需要确定误差加权比 W 。

步骤三：分析滤波器和综合滤波器使用同样的通带截止频率 ω_p ，设计原型滤波器，并根据式(8)解出最优的 ω_p 。

需要说明的是，该设计方法具有一定的灵活性：在性能要求不是特别高的情况下，分析滤波器和综合滤波器可使用同一个原型滤波器，简化设计过程。而在更为精确的情况下，可通过通带和阻带的纹波误差要求，以及通带、阻带截止频率，确定出最小的滤波器长度。此外，还可以根据通带和阻带性能的相对重要性，调整误差加权因子，以获得满意的效果。

3 WOLA 结构实现方法

加权叠接相加(WOLA)结构的分析滤波器组工作流程如图3所示^[12]。输入数据每次将 M 个样点输入 N_h 长的缓冲区中， N_h 也是分析窗的长度。缓冲区中的数据与时间反转的分析窗加权，然后将加权序列进行时间混叠分为 K 点一组的块，再将这些块相加后用循环移位进行 W_K^{kmM} 的相移，最后进行 K 点的 FFT。

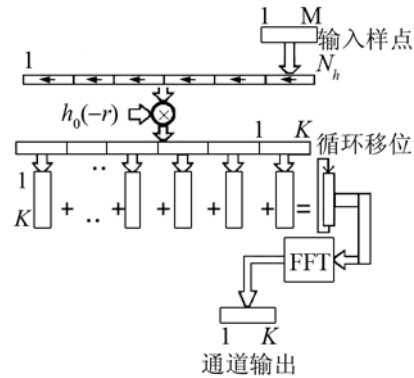


图3 WOLA 分析结构工作流程图
Fig.3 The processing flow of WOLA analysis

WOLA 综合结构则是分析结构的逆过程，如图4所示，首先将通道信号按 K 点一组输入，进行 K 点的 IFFT，用循环移位实现 W_K^{-kmM} 的相移，然后将序列进行周期延拓至长度为 N_f (综合窗的长度)，并用综合窗处理。最后与输出缓冲区里的数据叠

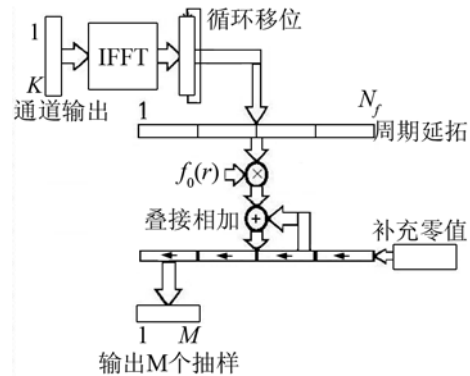


图4 WOLA 综合结构工作流程图
Fig.4 The processing flow of WOLA synthesis

加, 从队首输出 M 个抽样, 同时队尾补充 M 个零值。

WOLA 快速实现结构利用 FFT 技术, 大大减小了计算量, 适合语音信号的块处理方式和处理要求, 可以方便地在 DSP 上实现, 也可以研制成专用芯片。同时, 它的分析和综合各只需一个原型低通滤波器, 要求 $M \leq K$ 就可以了, 而并不要求 K 为 M 的整数倍, 这也方便了计算。

4 实验和结果

为了验证本文提出的滤波器组设计方法的性能, 仿真实验设计滤波器组的通道数为 $K=16$, 降采样率为 $M=4$ 。根据式(10), 分析滤波器的阻带截止频率设计为 $\omega_{sh}=2\pi/K=0.125\pi$ 。根据式(11), 可稍放宽综合滤波器的阻带截止频率, 设 $\omega_{sf}=0.15\pi$, 这样适当放宽过渡带可使在达到同样通带或阻带要求的前提下, 减小滤波器的长度, 从而减小滤波器组的群延迟。设分析和综合滤波器的长度分别为 $N_h=128$, $N_f=64$, 通带和阻带的误差加权比为 $W=10:1$ 。通过式(8)的优化计算, 得出通带截止频率的最优解为 $\omega_{p_opt}=0.01057\pi$ 。分析和综合原型滤波器的幅频响应如图 5 所示, 其中分析原型滤波器的阻带衰减能达到 120dB 以上, 由于综合原型滤波器长度减小了一半, 故阻带衰减要小一些, 但这并不影响滤波器组处理语音信号的整体性能。

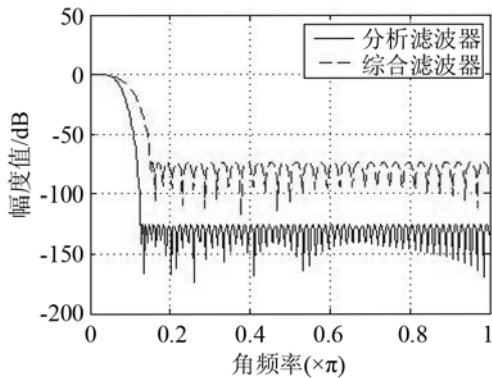


图 5 原型滤波器的幅频响应
Fig.5 Magnitude frequency response of prototype filter

图 6 是滤波器组的通道分布响应情况, 可以看到, 相邻的两个通道之间的过渡带交叠而不相邻的通道没有交叠, 这和理论分析一致。

式(4)和式(5)所示的系统失真传递函数和输出混迭函数是衡量滤波器组性能的两个重要指标。如图 7 所示, 失真传递函数的最大纹波约为 0.018dB, 且在零频率和 Nyquist 频率处平坦, 没有大的凸起或凹陷, 输出混迭量约为 -70 dB。该设计滤波器组

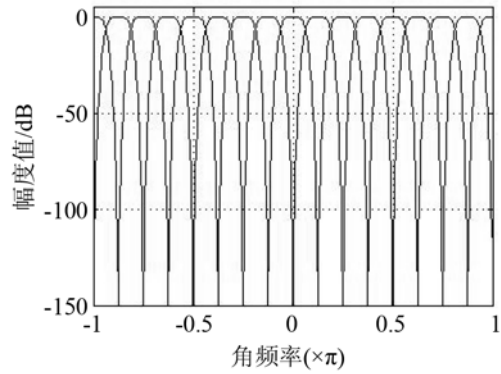


图 6 滤波器组各通道幅频响应
Fig.6 Frequency response of the filterbank

具有良好的重建和抗混迭性能。

选用一段语音信号, 男声英语朗读, 采样率为 8kHz, 时长 2.81s, 16bit 量化, 用设计的滤波器组进行分析和综合, 其中不做通道处理。图 8 显示了原信号和重建后的信号对比, 其中图 8(a)是原信号波形, 图 8(b)是重建信号波形, 可以看出, 两者波形几乎完全一致, 没有出现镜像和混叠效应, 实际听起来两段语音效果相同, 没有出现失真。

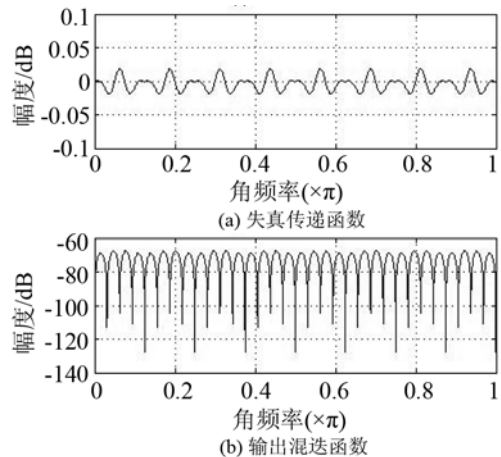


图 7 系统失真传递函数与输出混迭函数
Fig.7 Distortion and aliasing function

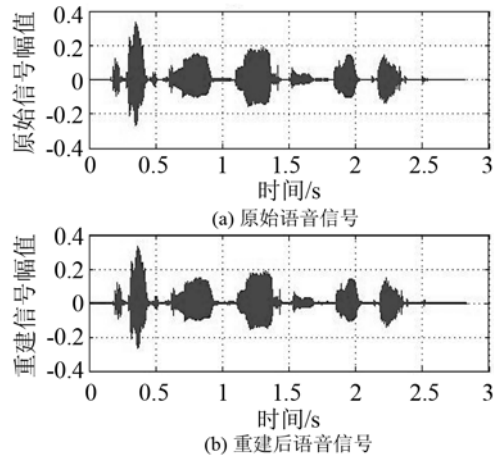


图 8 原始信号和重建后信号的对比
Fig.8 Comparison between the original and the reconstructed signals

5 结论

本文设计了一种可用于助听器语音信号分析处理的DFT调制滤波器组,原型滤波器的设计采用了等纹波法,相比于现有复杂的优化方法,将最优滤波器设计问题转化为对单一参数的选取,简化了设计过程。该方法在最大优化滤波器性能的同时考虑尽量降低其群延迟,以满足语音信号处理的要求。分析和综合滤波器组各由一个原型滤波器产生,并采用高效的WOLA结构实现,可用于实时处理。整个设计简单、易行,从实验结果中看,该方法取得了良好的效果。

参 考 文 献

- [1] Crochiere R E, Rabiner L R. Multirate digital signal processing[M]. Englewood Cliffs, NJ: Prentice-Hall, 1983.
- [2] Yiu K F C, Grbic N, Nordholm S. Multicriteria design of oversampled uniform DFT filter banks[J]. IEEE Signal Processing Letters, 2004, 11(6): 541-544.
- [3] Wilbur M R, Davidson T N, Reilly J P. Efficient design of oversampled npr gdf filterbanks[J]. IEEE Trans on Signal Processing, 2004, 52(7): 1947-1963.
- [4] De Haan J M, Gribbe N, Claesson I. Filter bank design for subband adaptive microphone arrays[J]. IEEE Trans on Speech Audio Processing, 2003, 11(1): 14-23.
- [5] Dam H H, Nordholm S, Cantoni A. Iterative method for the design of dft filter bank[J]. IEEE Trans on Circuits and System, 2004, 51(11): 581-586.
- [6] Creusere C D, Mitra S K. A simple method for designing high-quality prototype filters for M -band pseudo QMF banks[J]. IEEE Trans on Signal Processing, 1995, 43(4): 1005-1007.
- [7] Hermann D, Chau E, Dony R D. Window based prototype filter design for highly oversampled filter banks in audio application[C]// ICASSP, 2007, 2: 405-408.
- [8] Li Q, Chen W G, He C. Design of oversampled DFT modulated filter banks optimized for acoustic echo cancellation[C]// ICASSP, 2009, 197-200.
- [9] Alfsmann D, Gockler H G. Frequency-domain magnitude constraints for oversampling complex-modulated NPR filter bank system design ensuring prescribed signal-to-distortion ratio[J]. Electronics Letters, 2010, 46(3): 205-206.
- [10] 胡广书. 现代信号处理教程[M]. 北京: 清华大学出版社, 2004. HU Guangshu. Modern signal processing[M]. Beijing: Tsinghua University Press, 2004.
- [11] 何坚勇. 最优化方法[M]. 北京: 清华大学出版社, 2007. HE Jianyong. Optimization theory[M]. Beijing: Tsinghua University Press, 2007.
- [12] Brennan R, Schneider T. A flexible filterbank structure for extensive signal manipulations in digital hearing aids[C]// IEEE International Symposium on Circuits and Systems, NW: 1998: 569-572.

美国加州豪斯研究所 Sigfrid D. Soli 博士在东海站做学术报告

作为东海站的特聘技术顾问, Soli 博士于2012年9月10日在我站做了题为“Early hearing aid intervention: An international perspective”的学术报告。早期听力检测与干预对有听力障碍的成人与儿童都有帮助。小儿早期听力检测与干预的项目旨在为听障儿童在言语语言成长的关键时期提供适合的干预,并帮助听障儿童有效地参与到言语语言的发展过程中来。本项目也是对言语语言发展干预有效性的验证。早期儿童听力干预首先由助听器开始,但是针对佩戴助听器的低龄幼儿的行为学验证十分困难。

Soli 博士详细介绍了对120名1~5岁听力受损儿童的测试过程、数据分析,这些数据和比较能够为儿童助听器干预提供客观验证,同时,为采用循证医学原则建立一种最有效的实践方法提供帮助。

报告会进行了互动交流和提问。这次学术交流活动拓宽了科研人员和研究生的知识面,对人工耳蜗的临床研究起到了积极的推动作用。