

基于 DFT 滤波器组的低时延 FPGA 语音处理实现研究

薛一鸣¹, 陈 鹤¹, 何宁宁², 胡彩娥³, 王建平¹

(1. 中国农业大学信息与电气工程学院, 北京 100083; 2. 北京中电华大电子设计有限责任公司, 北京 102209;
3. 国网北京市电力公司, 北京 100031)

摘 要: 提出了 WOLA (Weighted Overlap-Add) 并行结构的低时延 DFT 滤波器组的设计和 FPGA 实现方法. 为降低系统总体时延, 在综合考虑传递失真、混迭失真的基础上, 将群时延引入系统目标函数, 并采用非对称综合原型滤波器设计方法, 提出迭代算法, 实现了 DFT 滤波器组低时延优化设计. 通过对 DFT 滤波器组中分析和综合功能的关键模块采用多路并行乘法、多级流水加法链设计, 实现了并行的 WOLA 结构 DFT 滤波器组, 降低 FPGA 实现的计算时延. 整个设计在 Xilinx 公司的 Zynq7020 型号 FPGA 芯片上进行实现. PESQ 测试表明, 设计的 DFT 滤波器组能取得较好的语音质量. 与串行 WOLA 结构的实现对比表明, 在 16kHz 语音采样率下, 并行的 WOLA 结构 FPGA 实现的总时延能降低 1.192ms, 其中群时延降低 12%, 计算时延降低 29.2%.

关键词: 语音处理; DFT 滤波器组; 低时延; FPGA; 并行设计

中图分类号: TN912.3 **文献标识码:** A **文章编号:** 0372-2112 (2018) 03-0695-07

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.3969/j.issn.0372-2112.2018.03.027

DFT Filter Bank-Based Realization of Low Delay Speech Processing on FPGA

XUE Yi-ming¹, CHEN Yao¹, HE Ning-ning², HU Cai-e³, WANG Jian-ping¹

(1. College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China;
2. CEC Huada Electronic Design Co., Ltd, Beijing 102209, China;
3. State Grid Beijing Electric Power Company, Beijing 100031, China)

Abstract: A low-latency parallel WOLA (Weighted Overlap-add) DFT filter bank design method and its implementation on FPGA are presented. System objective function combined with group delay, asymmetric synthesis window design and iterative algorithm are adopted to reduce the overall system delay during the optimization of DFT filter banks. Calculation delay of FPGA implementation is controlled through multichannel parallel multiplication, multistage pipeline addition chain in key modules of DFT filter banks. The whole design is implemented on a Xilinx FPGA chip of Zynq7020. PESQ test shows that the design can achieve good speech quality. Compared with the serial WOLA structure, the delay of parallel WOLA can be reduced by 1.192 ms at 16 kHz speech sampling rate, with the group delay reduced by 12% and the calculation delay reduced by 29.2%.

Key words: speech processing; DFT filter bank; low delay; FPGA; parallel design

1 引言

在语音实时处理应用中, 低时延一直是学者们所研究的热点问题之一^[1,2]. 以助听器为例^[3], 低延迟设计至关重要, 因为较大的延迟会导致助听器输出与声

音直接通过耳道到达鼓膜不同步现象的发生, 严重影响助听效果, 当语音处理时延超过 10ms 时人会有不适感, 当时延达到 3~5ms 时就能被人耳所察觉.

为减小延迟, Andersen K T^[4] 等人通过语音降噪算法的设计, 使降噪系统的整体时延缩短为滤波器组的

时延. Deng Ying 等人^[5]通过降低对原型滤波器的线性相位设计要求,换取语音增强应用中处理时延的降低. Markus Schnell 等人^[6]提出了 LD-AAC 结构,设计了专用的滤波器组,实现了低延迟音频编码. 但上述文献只是在算法级针对某种特定应用进行的时延优化,并未对语音分析和综合的时延进行系统分析,也未讨论硬件实现中的时延优化.

DFT 滤波器组 (DFT FB, Discret Fourier Transform Filter Bank) 是现代数字信号处理和通信领域中很重要的概念,作为一种均匀的复调制滤波器组,它在语音降噪^[4]、语音增强^[5]、回声消除^[7]等语音信号处理领域得到广泛应用. 因此,研究 DFT 滤波器组分析与综合实现时延的降低对语音处理系统整体时延的降低有重要意义.

滤波器组设计是传递失真和混叠失真的多目标优化问题,为满足滤波器组性能要求,在设计完全重建滤波器组时^[8,9],通常分析原型滤波器、综合原型滤波器的长度都比较长,系统时延较大. 为降低系统时延,文献^[10,11]采用窗函数法和等波纹法进行近似完全重构滤波器组设计,通过缩短综合原型滤波器的长度,降低了语音处理系统群时延. 然而受限于滤波器组的传递失真和混迭失真要求,所设计的滤波器长度不能过短. 为进一步降低群时延,文献^[4]引入非对称原型滤波器设计法来进行综合滤波器设计,但是在滤波器组通道数较少时原型滤波器性能退化严重. 文献^[12]基于多相设计方法,提出了单位非对角阵级联的算法理论,在优化分析滤波器、综合滤波器的同时,实现了滤波器组的低时延设计,但其算法复杂度偏高. 针对实时语音处理低时延和低计算复杂度要求,本文在滤波器组传递失真和混迭失真约束的基础上,在目标函数中引入时延约束,提出了简化的迭代设计算法,同时综合原型滤波器设计采用非对称设计方法. 进一步在 FPGA 上采用多路并行乘法、多级流水加法链等设计,减少 FPGA 实现所需的时钟周期数,降低计算时延. 测试表明,在 16kHz 语音采样的吞吐率需求下,并行实现的 WOLA 结构 DFT 滤波器组在语音分析综合处理的群时延和计算时延方面分别比传统结构实现小 12% 和 29.2%. 本方法及构建的语音分析和综合系统适用于实时语音处理应用领域.

2 DFT 滤波器组算法描述

2.1 原理分析

如图 1 所示. DFT 滤波器组包括语音分析滤波器组和语音综合滤波器组两部分,由于采用 FFT 快速算法来实现滤波器组,计算复杂度大大降低. 图 1 中 $x(n)$ 为输入语音信号序列, $h(n)$ 和 $g(n)$ 分别表示长度为 L_a 的分析窗和长度为 L_s 的综合窗序列, N 为滤波器组通道数, R 为降采样率, $y(n)$ 是经过分析综合后的输出语音

序列.

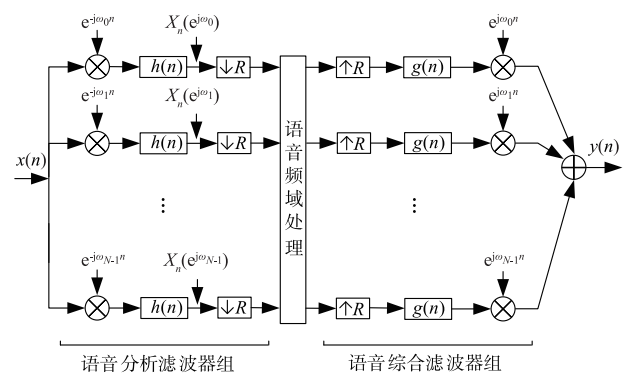


图1 DFT滤波器组框图

WOLA 结构是一种实现 DFT 滤波器组的高效且灵活的结构^[10]. 以语音分析部分为例,按处理的先后可划分为数据更新、乘窗序列、分段叠加及循环移位、FFT 运算共四个步骤,对输入语音序列 $x(n)$ 的离散时间离散频率的短时傅里叶变换可表示为

$$\begin{aligned} X_n(e^{j2\pi k/N}) &= X_n(e^{j\omega}) \Big|_{\omega=2\pi k/N} \\ &= \sum_{m=-\infty}^{\infty} x(m)h(n-m)e^{-j2\pi km/N} \\ & \quad 0 < k < N-1, \quad 0 < n < +\infty \end{aligned} \quad (1)$$

将 m 替换为 $n+m$, 并经过整理后进一步得到

$$X_n(e^{j2\pi k/N}) = e^{-j2\pi kn/N} \sum_{m=-\infty}^{\infty} x_n(m)e^{-j2\pi km/N} \quad (2)$$

式(2)中, $x_n(m)$ 是经过窗序列加权后的语音段; 对应于 WOLA 结构 DFT 滤波器组中语音分析部分的乘窗序列操作, 即

$$x_n(m) = x(n+m)h(-m) \quad (3)$$

令 $m = q + Nr$, 代入式(2)可得

$$\begin{aligned} X_n(e^{j2\pi k/N}) &= e^{-j2\pi kn/N} \sum_{q=0}^{N-1} u_n(q)e^{-j2\pi kq/N} \\ & \quad 0 < q < N-1, \quad -\infty < r < \infty \end{aligned} \quad (4)$$

其中, $u_n(q)$ 是 $x_n(m)$ 序列按每 N 点一段分割成 L/N 段后, 将每段对应位置上的序列值重新叠加求和得到的一个 N 点序列, 该分段叠加操作, 表示为

$$u_n(q) = \sum_{r=-\infty}^{\infty} x_n(q + Nr) \quad (5)$$

此时, 如式(4)所示, 语音短时频谱表示为序列 $u_n(q)$ 的 DFT (FFT) 计算结果乘以一个复指数常数 $e^{-j2\pi kn/N}$, 根据傅里叶变换的时移性质, 可视为 $u_n(q)$ 作循环移位 n 后再进行 DFT (FFT) 运算. 语音综合是语音分析的逆过程, 此处不再赘述.

2.2 低时延滤波器组设计

在 DFT 滤波器组系统中, 输出信号相对于输入信号的总时延由群时延和计算时延组成, 群时延的考量一般结合在滤波器组的原型滤波器设计过程中, 计算

时延则因硬件实现(或软件实现)的结构和系统时钟不同而有所差异. 本文所研究的实时语音处理应用要求在保证一定滤波器组性能的前提下,尽可能减小系统总时延,为此后文的内容将围绕着群时延和计算时延的降低展开.

2.2.1 群时延优化算法

DFT 滤波器组的设计其本质上是分析窗和综合窗(或称之分析和综合原型滤波器)的设计,通过分析窗和综合窗的优化设计可以有效降低群时延.

滤波器组的传递失真函数和混迭失真函数分别如式(6)和式(7)所示.

$$T(z) = \frac{1}{R} \sum_{k=0}^{N-1} H_k(z) G_k(z) \quad (6)$$

$$A(z) = \frac{1}{R} \sum_{l=1}^{R-1} \sum_{k=0}^{N-1} H_k(zW_R^l) G_k(z) \quad (7)$$

其中, $H_k(z)$ 表示 $h_k(n)$ 的 z 变换; $G_k(z)$ 表示 $g_k(n)$ 的 z 变换; 且 $h_k(n) = h(n) e^{-j\frac{2\pi}{N}k \cdot n}$; $g_k(n) = g(n) e^{-j\frac{2\pi}{N}k \cdot n}$; $W_R^l = e^{-j\frac{2\pi}{N}l}$.

在完全重建滤波器组设计中,分析窗与综合窗选取为相同的长度,系统总时延较大. 文献[13]将滤波器组的混叠失真分为分析滤波器组的混叠失真和综合滤波器组的镜像扰动,针对这两部分分量,分别提出了两个目标函数,从而将分析滤波器组和综合滤波器组分别用迭代的方法进行优化设计. 文献[11]考虑到分析滤波器组的设计直接关系到子带的划分,对语音的频域处理影响较大,采用了较长的分析原型滤波器以有效消除混叠影响,而综合原型滤波器则采用了较短的长度,以缩短系统的时延. 借鉴了文献[11,13,14]的滤波器组化简的设计思想,本文选择固定一组最优的分析窗系数 $h(n)$,在保证分析部分性能的前提下,在综合考虑传递失真 $T(z)$ 和混迭失真 $A(z)$ 约束的基础上,引入群时延 τ 约束到综合窗的设计过程中,将群时延的优化简化为对综合窗的优化. 具体地,综合窗优化的目标函数如式(8)所示.

$$E(g) = \alpha \varepsilon_1(g) + \beta \varepsilon_2(g) \quad (8)$$

其中, ε_1 为传递失真误差, ε_2 为平均相位误差, α 和 β 为加权系数. ε_1 和 ε_2 的定义如下, τ 为目标群时延.

$$\varepsilon_1(g) = \frac{1}{2\pi} \int_0^{2\pi} |1 - |T(e^{j\omega})||^2 d\omega \quad (9)$$

$$\varepsilon_2(g) = \frac{1}{2\pi} \int_0^{2\pi} |\text{angle}(T(e^{j\omega})) - \text{angle}(T(e^{j0})) + \omega\tau| d\omega \quad (10)$$

对于混迭失真的约束,简化为对综合窗频率响应的约束如式(11)所示.

$$\varphi(g) = \int_0^{\omega_p} |1 - |G_0(e^{j\omega})||^2 + \int_{\omega_s}^{2\pi} |G_0(e^{j\omega})|^2 d\omega \quad (11)$$

对 $E(g)$ 的求解是一个凸函数优化问题,可用迭代法等多种方法^[15]求解. 根据以上分析得到本文的非对称窗设计步骤如下.

步骤 1 基本参数初始化. 给定滤波器组通道数 N , 降采样率 R ,且满足过采样要求 $R < N$. 根据语音处理应用对群时延的要求 τ ,选择适当的分析窗和综合窗的长度 L_a 和 L_s ,采用等波纹法设计了线性相位分析窗序列,设置综合窗的通带截止频率初始值 $\omega_p = \omega_{p,0}$,迭代初始步长 Δ_0 ,步长更新参数 r ,阻带截止频率 ω_s . 根据对通带波纹 δ_p 和阻带衰减 δ_s 的要求确定误差加权比 W .

$$W = \delta_p / \delta_s \quad (12)$$

步骤 2 原型滤波器设计. 根据 $W, \omega_p, \omega_s, \tau$,采用非线性相位等波纹设计法设计综合原型滤波器,得到 $g(\omega_p, \omega_s, \tau), g(\omega_p, \omega_s, \tau)$ 应使式(11)所述目标函数值最小.

步骤 3 迭代阶段. 阻带截止频率 ω_s 固定,对通带截止频率 ω_p 进行梯度法迭代的过程为

$$\omega_{p,i} = \omega_{p,i-1} - \Delta_{p,i-1} \text{sign}(\hat{\omega}_{p,i-1}) \quad (13)$$

其中, $\Delta_{p,i-1}$ 为呈指数下降趋势的迭代步长, $\hat{\omega}_{p,i-1}$ 的符号特性决定迭代的方向,具体定义如下

$$\hat{\omega}_{p,i-1} = \frac{E_{i-1} - E_{i-2}}{\omega_{p,i-1} - \omega_{p,i-2}}, \Delta_{p,i-1} = \Delta_0 \cdot e^{-(i-1)/r} \quad (14)$$

步骤 2、步骤 3 迭代进行,直至满足迭代终止条件.

2.2.2 群时延数据比较

为了论证非对称窗设计法的性能,选择文献[11,16]方法以及最小相位窗设计法进行对比,并做标记如下:

- (1) 采用文献[11]所述窗函数法设计综合窗,记作对称窗 a;
- (2) 采用文献[16]所述等波纹法设计综合窗,记作对称窗 b;
- (3) 选择具有最小群时延的非对称窗作为综合窗,记作最小相位窗;
- (4) 采用本文 2.2.1 小节论述的非对称窗设计法设计综合窗,记作非对称窗.

选择滤波器组参数 $N = 16, R = 4, L_a = 128, L_s = 64, \tau = 20$,得到四种综合窗设计下的时域波形和幅频响应对比,传递失真、混迭失真以及群时延对比结果分别如图 2 和表 1 所示.

表 1 采用对称综合窗与非对称综合窗设计下的滤波器组性能对比

综合窗类型	长度/	传递失真/dB	混迭失真/dB	滤波器组群时延/samples
对称窗 a	64	< 4.3	< -75	100
对称窗 b	64	< 0.15	< -65	100
最小相位窗	64	< 0.75	< -68	83 ~ 91
非对称窗	64	< 0.79	< -63	88

从图 2 的幅频响应图可以看到,本文设计的非对称综合窗在通带和过渡带具有优于对称综合窗 a 且与对

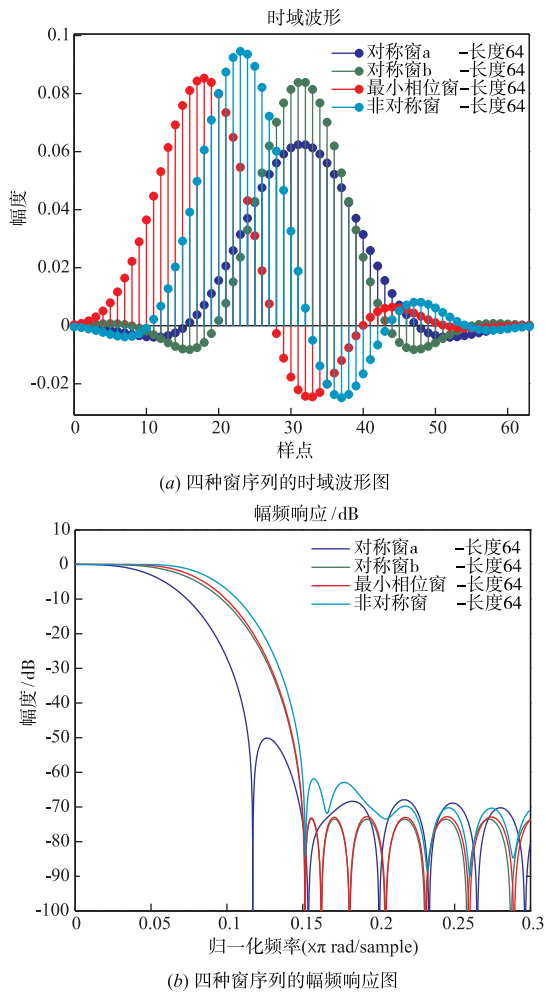


图2 对称综合窗与非对称综合窗的时域波形及幅频响应图

称综合窗 b 以及最小相位窗相当的幅频响应。

进一步分析表 1 中滤波器组的传递失真和混迭失真性能,显然非对称综合窗设计下的滤波器组混迭失真与对称综合窗 b 相近,传递失真稍大于对称综合窗 b. 从群时延比较可看出,最小相位窗在通带具有 8 个点左右群时延波动,相位失真明显,而在保证滤波器组性能的前提下,本文的非对称综合窗能够比对称综合窗 a、b 减少 12% 的群时延。

2.2.3 计算时延优化

计算时延用信号从输入到输出所经过的时钟周期数 Cycles 和系统时钟频率 F 来衡量,表示为

$$\text{Delay}_{\text{Calc}} = \text{Cycles} \cdot (1/F) \quad (15)$$

语音处理系统一般都是低功耗应用,要求系统时钟频率不能太高. 显然,在系统时钟频率固定的情况下,减少处理所需的时钟周期数 Cycles 能够带来计算时延的降低。

3 FPGA 实现

FPGA 平台所具有的并行特性使其在实时性要求

高或者计算复杂度大的研究中得到比 DSP 更为广泛的应用^[17]. 为减小系统的计算时延,本文利用 FPGA 设计并实现了具有并行特性的新型 WOLA 实现结构 DFT 滤波器组,系统设计如下。

3.1 系统的模块化设计与标准接口设计

系统架构如图 3 所示,整个体系结构遵循模块化设计和标准总线接口设计思想,内部总线为 AHB 总线,挂载 I²S 接收模块、预分析模块、FFT、IFFT 模块、后综合模块、I²S 发送模块和控制模块,根据应用的不同可增加语音增强、语音降噪、回声消除等在频域做语音处理的功能模块,外部总线为专用于音频数据传输的 I²S 总线, I²S 接收模块和 I²S 发送模块完成对所输入语音信号的接收和对经过系统处理后语音的发送;预分析模块和 FFT 模块实现语音分析操作;IFFT 模块和后综合模块实现语音综合操作;控制模块协调各模块的工作时序. 而语音分析后根据具体应用需求所进行的语音增强、语音降噪或回声消除等算法功能,可以单独作为一个功能模块加入到系统中。

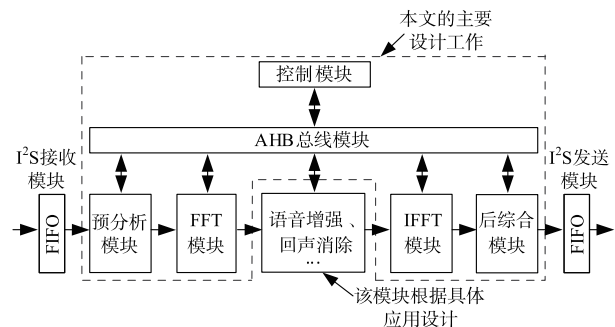


图3 DFT滤波器组在FPGA上实现的系统架构图

3.2 DFT 滤波器组关键部分的并行化

3.2.1 预分析模块并行化结构设计

预分析模块实现 R 点语音数据的更新、语音序列 $x(n)$ 与窗序列 $h(n)$ 的逐点相乘、分成每段 N 点长的 L_a/N 段子序列,并将对应点相加以及循环移位. 观察 WOLA 结构的分析部分可以发现,语音序列的乘法运算中,每个序列值之间是时序上相互独立的,即允许各序列的乘积运算并行进行. 根据这一特性,设计“RAM-乘法运算器-RAM”形式的基本单元,将 L_a 点数据乘窗操作分解为 L_a/N 路并行乘法以得到加权序列 $x_n(m)$,使乘法运算时间缩短为原来的 $(N-1)/L_a$ 倍,如图 4 所示. 为进一步提高处理速度,将求解 $u_n(q)$ 序列里每个序列值所需的 $L_a/N-1$ 次叠加运算设计成 $\log_2(L_a/N)$ 级流水加法链,使单点的叠加求值速度提升 $L_a/N-1-\log_2(L_a/N)$ 倍. 结合流水线处理机制的优点,可实现序列整体叠加求值速度提升 $(L_a/N-1-\log_2(L_a/N))\log_2(L_a/N)$ 倍. 此外,通过写入地址的循环更迭控制来实现对所求取序列 $u_n(q)$ 的循环移位操作,也避免了使用移位寄存器带来的额外处理时延。

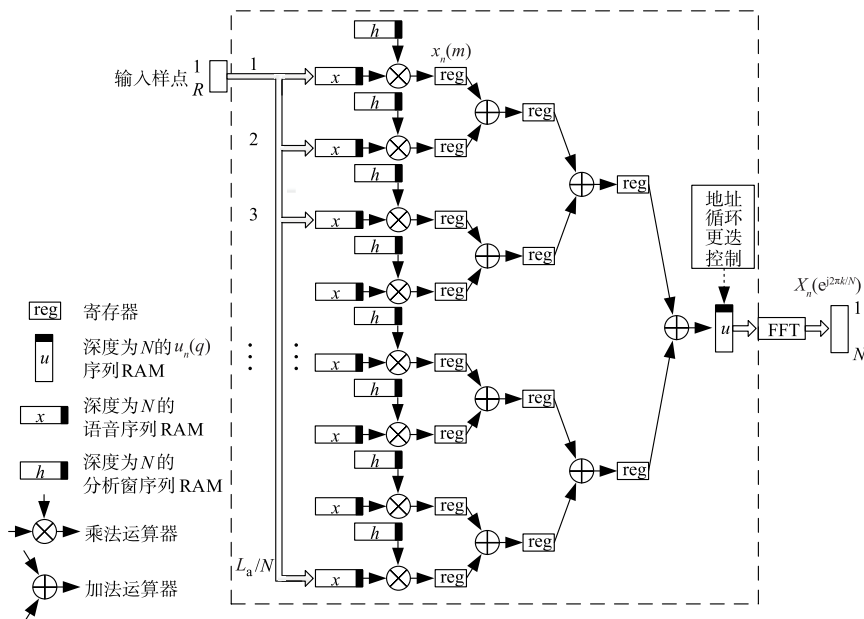


图4 预分析模块结构图

3.2.2 后综合模块设计与并行工作时序

后综合模块实现预分析模块的逆过程. 根据 WOLA 结构中的综合部分特点, 每次 IFFT 模块计算输出的 N 点数据, 在进入后综合模块后, 应进行循环移位得到序列 $u'_n(q)$, 将 $u'_n(q)$ 周期延拓至综合窗的长度, 并乘以综合窗序列 $g(-r)$ 实现加权, 加权结果会与预存的输出缓冲数据叠加并选择最新的 R 点数据作为后综合模块的输出. 如图 5 所示, 对于加权叠加操作实现采用与预分析模块相同的降低时延思路, 设计 L_s/R 路并行乘法器和加法器加快处理. 特别地, 在输出数据选择器相关设计中, y_fifo 地址标识动态分配机制会给 L_s/R 块 y_fifo 对应的 RAM 按循环规则分配标识号, 标识号用来区分当前乘法器计算结果是作为后综合模块处理结果输出或者是用于更新 y_fifo 中存储的语音缓冲数据. 动态分配标识可以实现无延时的 R 点输出语音缓冲区数据移位.

3.2.3 FFT/IFFT 模块

FFT 模块对预分析模块写入 FFT 数据双端口 RAM 的 N 点数据计算频谱, IFFT 模块实现 FFT 模块的逆过程并将处理结果输出给后综合模块. FFT/IFFT 模块采用 N 点的基-2 时间 FFT 运算, 在整个运算过程中, 蝶形运算单元的两个操作数通过双端口 RAM 可以同时取出, 通过并行运算减少运算实现所需的时钟周期数.

4 实验和结果

系统设计采用 Verilog 语言描述, 在 Xilinx 的 Zynq7020 型号 FPGA 上进行了实现. 实现选取本文算法描述部分非对称综合窗设计法所分析的参数 $L_a = 128, L_s = 64, N = 16, R = 4$, 并设定语音采样率为 $f_s =$

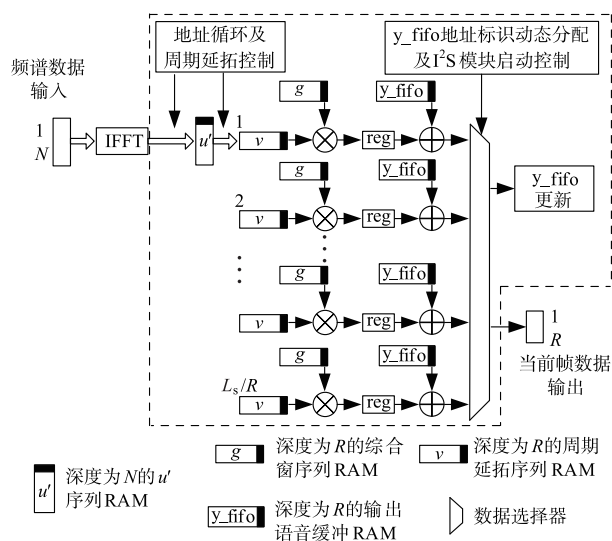


图5 语音后综合模块结构设计图

16kHz. FPGA 的综合报告如表 2 所示. 为验证系统语音分析和综合的正确性并测试系统时延的性能, 对系统的输出语音进行了 PESQ 语音客观质量测试^[18], 并将 WOLA 并行结构 DFT 滤波器组系统与串行 WOLA 结构 DFT 滤波器组在 FPGA 实现上的系统总时延进行对比.

表 2 FPGA 综合报告

Device	7z020clg484-3
Number of Slice Registers	11052 out of 106400 10%
Number of Slice LUTs	7358 out of 53200 14%
IO Utilization	16 out of 200 8%
Number of Block RAM/FIFO	6 out of 140 4%
Number of DSP48E1S	38 out of 220 17%

4.1 PESQ 测试

对本文中的非对称窗设计的 DFT 滤波器组 Matlab 仿真模型进行 PESQ 测试,得到在 $L_s = 128, L_a = 64, N = 16, R = 4$ 参数选取下,选择 ITU-T 组织在 P. 501 附录中所推荐的语料库作为测试语料库^[19],挑选库中 6 段长度为 5s,单声道,16kHz 采样的语音作为测试输入.得到输出语音的 PESQ 分数如表 3 所示.显然非对称窗设计下的 DFT 滤波器组能取得接近 4.425 分的语音质量,所设计的滤波器组性能良好.

表 3 PESQ 测试结果

测试语音样本	file1	file2	file3	file4	file5	平均值
PESQ	4.467	4.474	4.306	4.448	4.465	4.425

4.2 本文 WOLA 并行结构 DFT 滤波器组与串行结构的 FPGA 实现对比

选取短对称综合窗 a 和 b 的参数配置为 $L_a = 128, L_s = 64, N = 16, R = 4$,分别在 Zynq7020 型号 FPGA 上实现 WOLA 串行结构的 DFT 滤波器组.在 2MHz 系统时钟频率下,本文所提出的 WOLA 并行结构 DFT 滤波器组与串行 WOLA 结构在 FPGA 上实现时系统时延的对比结果如表 4 所示.

表 4 WOLA 并行结构 DFT 滤波器组与串行 WOLA 结构在 FPGA 实现时的时延对比

实现方式	群时延 /ms	计算时延		总时延 /ms
		一帧处理所需的时钟周期数	一帧处理时延/ms	
短对称窗 a/b + 串行 WOLA 结构	6.25	3028 个	1.514	7.764
非对称窗 + 并行 WOLA 结构	5.5	2144 个	1.072	6.572

从表 4 中可以看出,相比于短对称窗,本文的非对称窗除了在群时延上具有 0.75ms 的下降.通过所提的 WOLA 并行结构,计算时延比串行 WOLA 结构少 884 个时钟周期,总时延降低 15.4%.在 2MHz 的系统时钟频率下,对应于 0.442ms 的时延降低,而在时钟频率越低(一般是低功耗应用的场景)时,该时钟周期数减少所带来的计算时延降低效果会更加明显.并行结构相比串行结构所降低的 29.2% 的计算时延,对语音增强、语音降噪等语音实时处理系统的总时延降低具有重要意义.

5 结束语

本文提出了一种低时延 DFT 滤波器组设计方法,并采用 WOLA 并行结构在 FPGA 上实现.为降低系统总体时延,在综合考虑传递失真、混迭失真的基础上,将群时延引入系统目标函数,并采用非对称综合原型滤波器设计方法,提出了迭代算法,实现了 DFT 滤波器组低时延优化设计.在实现结构上采用模块化和标准

接口设计,通过多路并行乘法、多级流水加法链等完成对语音预分析和后综合模块的并行结构设计,减少硬件计算所需的时钟周期数.测试表明,群时延比短对称窗小 12%,在 16kHz 语音采样的吞吐率需求下,并行的 WOLA 结构设计在 FPGA 上的计算时延比串行结构小 29.2%.本文方法及构建的语音分析和综合系统适用于实时语音处理应用领域.在未来的研究工作中,将针对语音降噪等具体应用,实现频域的子带增益调控,并做低功耗应用场景的优化研究.

参考文献

- [1] SUGIURA R, KAMAMOTO Y, HARADA N, et al. Resolution warped spectral representation for low-delay and low-bit-rate audio coder [J]. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2015, 23(2): 288-299.
- [2] SEBASTIAN J, KUMAR M, MURTHY H A. An analysis of the high resolution property of group delay function with applications to audio signal processing [J]. Speech Communication, 2016, 81: 42-53.
- [3] AGNEW J, THORNTON J M. Just noticeable and objectionable group delays in digital hearing aids [J]. Journal of the American Academy of Audiology, 2000, 11(6): 330-336.
- [4] ANDERSEN K T, MOONEN M. Adaptive time-frequency analysis for noise reduction in an audio filter bank with low delay [J]. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2016, 24(4): 784-795.
- [5] DENG Y, MATHEWS V J, FARHANG-BOROUJENY B. Low-delay nonuniform pseudo-QMF banks with application to speech enhancement [J]. IEEE Transactions on Signal Processing, 2007, 55(5): 2110-2121.
- [6] SCHNELL M, GEIGER R, SCHMIDT M, et al. Low delay filterbanks for enhanced low delay audio coding [A]. Proceedings of IEEE Workshop on Applications of Signal Processing to Audio and Acoustics [C]. New Paltz: IEEE, 2007. 235-238.
- [7] SHEIKHZADEH H, BRENNAN R L, WHYTE K R L. Near-end distortion in over-sampled subband adaptive implementation of affine projection algorithm [A]. Proceedings of the 12th European Signal Processing Conference [C]. Vienna: IEEE, 2004. 413-416.
- [8] 石光明, 焦李成. 两通道完全重构滤波器组的设计方法: 因式分解法 [J]. 电子学报, 2001, 29(10): 1412-1414. SHI Guang-ming, JIAO Li-cheng. Design of PR two-channel filter bank based on factorization [J]. Acta Electronica Sinica, 2001, 29(10): 1412-1414. (in Chinese)
- [9] 石光明, 焦李成. 无约束法设计完全重构 M 带滤波器组

- [J]. 电子学报,2002,30(7):970-973.
SHI Guang-ming, JIAO Li-cheng. Design for PR M-channel filter banks without constraint[J]. Acta Electronica Sinica,2002,30(7):970-973. (in Chinese)
- [10] YIU K F C, GRBIC N, NORDHOLM S, et al. Multicriteria design of oversampled uniform DFT filter banks[J]. Signal Processing Letters IEEE,2004,11(6):541-544.
- [11] HERMANN D, CHAU E, DONY R D, et al. Window based prototype filter design for highly oversampled filter banks in audio applications[A]. Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing [C]. Honolulu: IEEE,2007. II-405-II-408.
- [12] MAKUR A, VIJAYAKUMAR A. Theory, design and application of arbitrary order arbitrary delay filterbanks[J]. IEEE Transactions on Signal Processing,2014,62(18):4811-4823.
- [13] STOCKER C, KURBIEL T, ALFSMANN D, et al. A novel approach to the design of oversampling complex-modulated digital filter banks[A]. Proceedings of the 17th European Signal Processing Conference [C]. Glasgow: IEEE,2009. 2648-2652.
- [14] GRBIC N, HAAN J M D, CLAESSEON I, et al. Design of oversampled uniform DFT filter banks with reduced in-band aliasing and delay constraints[A]. Proceedings of the Sixth International, Symposium on Signal Processing and ITS Applications [C]. Kuala Lumpur: IEEE,2001. 104-107.
- [15] DAM H H, NORDHOLM S, CANTONI A, et al. Iterative method for the design of DFT filter bank[J]. IEEE Transactions on Circuits and Systems II: Express Briefs,2004,5(11):581-586.
- [16] 蔡宇,侯朝焕,原建平,等. 适用于助听器语音信号处理的 DFT 调制滤波器组设计[J]. 声学技术,2012,31(5):517-521.
CAI Yu, HOU Chao-huan, YUAN Jian-ping, et al. Design of DFT modulated filter bank for speech processing in hearing aids[J]. Technical Acoustics,2012,31(5):517-521. (in Chinese)
- [17] CHEN J, LEI Y, PENG Y, et al. Configurable floating-point FFT accelerator on FPGA based multiple-rotation CORDIC[J]. Chinese Journal of Electronics,2016,25(6):1063-1070.
- [18] IUT-T P. 862. 3. Application Guide for Objective Quality Measurement Based on Recommendations P. 862, P. 862. 1 and P. 862. 2[S].
- [19] ITU-T P. 501. Test Signals for Use in Telephony Amendment 3: New Annex D-Speech Files with Male/Female Sentences Prepared for Use with Perceptual Based Objective Speech Quality Prediction[S].

作者简介



薛一鸣 男,1968 年出生,山西文水人. 中国农业大学信息与电气工程学院副教授、硕士生导师. 主要研究方向为多媒体信息处理、多媒体信息安全、大规模集成电路设计.
E-mail: xueym@cau.edu.cn



陈 鹄 男,1992 年 6 月出生,江西上饶人. 现为中国农业大学信息与电气工程学院硕士研究生,从事信息与通信工程相关研究.
E-mail: chenya@cau.edu.cn



何宁宁 男,1994 年 6 月出生,浙江丽水人. 现为北京中电华大电子设计有限责任公司数字电路设计工程师,从事算法实现与数字电路设计相关研究.
E-mail: henn@hed.com.cn