

# 块浮点FFT处理器的有限字长效应分析

乔树山, 黑 勇, 吴 斌, 王晓琴

(中国科学院微电子研究所 北京 朝阳区 100029)

**【摘要】**研究了基于基8算法的块浮点FFT处理器的有限字长效应问题, 提出了一种基于理论统计分析的静态模型。在不考虑输入信号的量化误差和系数量化误差情况下, 对基8单元和加权过程的误差进行了分析; 给出了有限字长效应所造成的误差随着频率点数和级数的变化趋势。通过SPEED开发平台得到的硬件仿真结果验证了该方法估计字长效应的正确性, 可以将其应用于工程分析。

**关键词** FFT处理器; 有限字长效应; 基8单元; SPEED开发平台  
**中图分类号** TN402 **文献标识码** A

## Finite Word-Length Effects in Implementation of a Block Floating-Point FFT Processor

QIAO Shu-shan, HEI Yong, WU Bin, WANG Xiao-qin

(Institute of Microelectronics, Chinese Academy of Science Chaoyang Beijing 100029)

**Abstract** This paper discusses issues of finite word-length effects in the block floating-point radix-8 FFT processor and presents a novel static model based on statistical analysis. When the input signal is white noise, the error caused by radix-8 unit and weight operation shows that the noise signal ratio (NSR) is proportional to  $r = \log_8 N$  without the quantization error of input signal and coefficients, where  $N$  is the length of FFT. The paper also gives the error trend with the frequency point and transform stages. The simulation result from SPEED platform has a good consistent with the finite word-length effects analysis results. Consequently this scheme has better value in practice.

**Key words** FFT processor; finite word-length effects; radix-8 algorithm; SPEED platform

快速傅里叶变换(FFT)在数字信号处理中应用极为广泛。然而随着FFT点数的增加, 有限字长效应引起的误差也越来越大, 为了保证合理的精度, 数字信号处理器的字长也必须相应增大。但是字长越大, 数字信号处理器的逻辑越复杂, 器件性能就越低, 面积、功耗也随之变大, 芯片成本增高, 可靠性降低<sup>[1-5]</sup>。因此, 设计时必须对有限字长效应进行充分的考虑, 进而在处理器的精度、性能、面积、功耗以及成本之间做出适当的折中方案。

### 1 分裂基算法原理

FFT变换为<sup>[6]</sup>:

$$y(k) = \sum_{i=0}^{N-1} x(i)W_N^{ki} \quad (1)$$

式中  $k = 0, 1, \dots, N-1$ ;  $i = 0, 1, \dots, N-1$ ;  $W = e^{-j2\pi ki/N}$  为旋转因子;  $N$  为FFT变换的数据长度。

设  $i = i_1 N_1 + i_2$ ,  $i_1 = 0, 1, \dots, N_2 - 1$ ,  $i_2 = 0, 1, \dots, N_1 - 1$ 。如果  $N$  是一个合数, 且  $N = N_1 N_2$

$(N_1 \neq 1, N_2 \neq 1)$ 。

设  $k = k_1 N_2 + k_2$ ,  $k_1 = 0, 1, \dots, N_1 - 1$ ,  $k_2 = 0, 1, \dots, N_2 - 1$ , 可得:

$$y(k) = y(k_1, k_2) = \sum_{i_1=0}^{N_2-1} \sum_{i_2=0}^{N_1-1} x(i_1, i_2) W_N^{(i_1 N_1 + i_2)(k_1 N_2 + k_2)} = \sum_{i_1=0}^{N_2-1} T(i_1, k_2) W_N^{i_1 k_2} W_{N_2}^{i_1 k_1} \quad (2)$$

由式(2)可以看出, 一个大点数的FFT运算可以转化为两个小点数的FFT运算<sup>[7]</sup>, 即 $N$ 点FFT运算可以转换为 $N_1$ 个 $N_2$ 点的FFT和 $N_2$ 个 $N_1$ 点的FFT, 总运算量从 $N^2$ 次乘法累加减少为 $N(N_1+N_2)$ 次, 大大降低了处理器的运算量。当分裂时 $N_i=8$ , 即为基8FFT算法。 $N_i$ 越大, 计算量越小, 但控制越复杂。

### 2 误差分析

基2与基4算法及其误差分析是较完善的方法

[8-9], 但是在实现大点数的FFT变换时, 基2与基4算法计算量仍然很大, 基8算法成为很好的选择。在采用基8算法实现FFT处理器时, 其误差来源主要有输入量化误差、系数量化误差、加权误差和混合基单元运算过程中的误差等。输入量化误差和系数量化误差与采样、A/D以及芯片的精度都有很大关系。下面分析硬件实现时在加权和混合基单元运算过程中舍入误差对计算精度的影响。

在基8算法的处理器中, 核心运算单元是基8单元。利用FFT运算结果的共轭对称性, 8点的FFT只需计算其中4点数就可以得出全部结果, 大大缩短计算时间, 其中的旋转因子可以利用硬查表的方法或者递归计算实现。在递归实现时, 只需要在初始时储存 $\log_2 N$ 个复数值, 而查表法需要储存 $N/2$ 个复数值。但是查表法比递归法运算时造成的系数量化误差小, 其最大输出幅度误差 $\max|e(k)|$ 分别为<sup>[10]</sup>:

$$\max_k |e(k)| \leq \sqrt{2} \Delta \frac{N}{2} (\log_2 N - 1) \|x\|_\infty \quad (3)$$

$$\max_k |e(k)| \leq \sqrt{2} \Delta \frac{N}{2} (N - 1 - \log_2 N) \|x\|_\infty \quad (4)$$

式中  $N$  为FFT运算的数据长度。由式(4)可见, 当使用查表法时, FFT的输出误差与 $M \log_2 N$ 成正比, 而递归算法与 $N^2$ 成比例。

### 2.1 基8单元误差分析

在浮点结构中, 最基本的运算是复数乘法, 该单元是由四个乘法与两个加法(包括减法)组成。而浮点的加乘法都可能使字长增加, 必须进行尾数处理。为保证性能采用舍入处理, 会造成舍入误差。建立的模型如图1所示。 $X_{m-1}(j)$ 与 $W_N$ 分别为被乘数和乘数,  $e_1(m)$ 、 $e_2(m)$ 为乘法误差,  $e_3(m)$ 为加法误差。

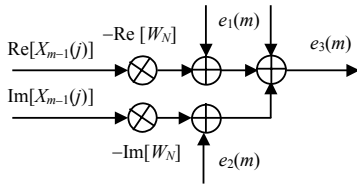


图1 浮点复数乘法实部误差模型

如果设输入信号为随机信号, 各误差源为白色且方差为 $\sigma^2$ , 可得图1中各误差信号的方差为<sup>[8]</sup>:

$$\sigma_{e_1}^2 + \sigma_{e_2}^2 = \sigma_{e_3}^2 = \frac{1}{2} \sigma^2 E[|X_{m-1}(j)|^2] \quad (5)$$

即浮点乘法复舍入误差的实部绝对均方值为:

$$\sigma_r = \sigma_{e_1}^2 + \sigma_{e_2}^2 + \sigma_{e_3}^2 = \sigma^2 E[|X_{m-1}(j)|^2] \quad (6)$$

进而浮点复数乘法复舍入误差绝对均方值为:

$$\sigma_f = 2\sigma_r = 2\sigma^2 E[|X_{m-1}(j)|^2] \quad (7)$$

对于8点的FFT单元, 任意输出点为:

$$X_m(j) = \sum_{i=0}^7 X_{m-1}(i) W_8^{ij} \quad (8)$$

所以基8单元的实部的绝对均方值为式(8)中8个复数乘法的累加和为:

$$\begin{aligned} \sigma_{8r}^2 &= 8\sigma^2 E[|X_{m-1}(j)|^2] + \\ &(1+1.5+\dots+4)\sigma^2 E[|X_{m-1}(j)|^2] = \\ &25.5\sigma^2 E[|X_{m-1}(j)|^2] \end{aligned} \quad (9)$$

进而基8单元的复舍入误差绝对均方值为:

$$\sigma_8^2 = 2\sigma_{8r}^2 = 51\sigma^2 E[|X_{m-1}(j)|^2] \quad (10)$$

### 2.2 加权过程误差分析

8点FFT变换的实部为:

$$\begin{aligned} \text{Re}[e_7(m)] &= \text{Re}[X_{m-1}(j^0)W_N^0] + \\ &\text{Re}[X_{m-1}(j^1)W_N^1] + \dots + \text{Re}[X_{m-1}(j^7)W_N^7] \end{aligned} \quad (11)$$

基8单元之后加权的实部为:

$$\begin{aligned} \text{radix8\_e} &= \text{Re}[e_7(m)] \text{Re}[W_N^K] - \\ &\text{Im}[e_7(m)] \text{Im}[W_N^K] = \\ &\text{Re}[e_7(m)W_N^K] \end{aligned} \quad (12)$$

基8单元之后加权的实部 radix8\_e 所引起的误差绝对均方值为:

$$\begin{aligned} E\{|\text{Re}[e_7(m)W_N^K]|^2\} &= E\{|\text{Re}[e_7(m)] \text{Re}[W_N^K]|^2\} + \\ E\{|\text{Im}[e_7(m)] \text{Im}[W_N^K]|^2\} &= 4\sigma^2 E[|X_{m-1}(j)|^2] \end{aligned} \quad (13)$$

所以加权的复舍入误差绝对均方值为:

$$\begin{aligned} \sigma_w^2 &= 2E\{|\text{Re}[e_7(m)W_N^K]|^2\} = \\ &8\sigma^2 E[|X_{m-1}(j)|^2] \end{aligned} \quad (14)$$

### 2.3 输出噪声信号比

对于8点的FFT单元, 由式(8)得知, 第 $m-1$ 级迭代输出信号与输入信号 $x(n)$ 之间关系为:

$$E[|X_{m-1}(i)|^2] = 8^{m-1} E[|x(n)|^2] \quad (15)$$

因为FFT运算的每个输出与第 $m$ 列的 $8^{r-m}$ 个8点单元相关, 其中 $r = \log_8 N$ ,  $N$ 为FFT变换的点数。由此可得输出噪声的总误差方差由 $r$ 级组成, 从1到 $r-1$ 级为基8单元之后加权的误差, 第 $r$ 级为基8单元之后的误差为:

$$\begin{aligned} E[|T(k)|^2] &= \sigma_8^2 + \sum_{m=1}^{r-1} 8^{r-m} (\sigma_w^2 + \sigma_8^2) = \\ &\frac{59r-8}{8} N \sigma^2 E[|x(n)|^2] \end{aligned} \quad (16)$$

由上述推导可得出浮点FFT系统白噪声输入信号时, 输出噪声信号比为:

$$\frac{E[|T(k)|^2]}{E[|X(k)|^2]} = \frac{59r-8}{8} \sigma^2 = \frac{59 \log_8 N - 8}{8} \sigma^2 \quad (17)$$

因此, 噪声信号比与 $r = \log_8 N$ 成线性关系。

### 3 FFT计算中的有限字长效应的分析验证

当输入信号为白噪声时,只考虑加权误差和混合基单元运算过程中的误差,可以得出输出噪声信号比理论值与硬件运算值曲线,硬件运算值是在SPEED平台上获得的,如图2所示。从图2可以看出,理论值与硬件运算值基本一致。但在硬件实现时,硬件运算值小于理论值,这是因为当 $W_N^j=1$ 或 $j$ 时,其实部与虚部可能为0,即在浮点乘法中 $e_1=e_2=e_3=e_4=e_5=e_6=0$ ,由于乘数或者加数为0,这些操作都不会带来误差,即浮点乘法不会引入噪声;旋转因子为0时也不会产生误差,所以实际噪声要比理论值小; $\sigma^2$ 的影响导致每次运算的比值不一样。

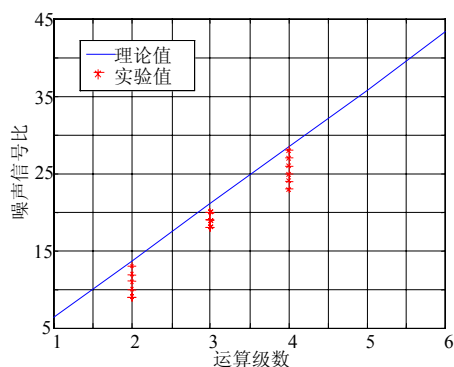


图2 FFT计算出噪声信号比曲线

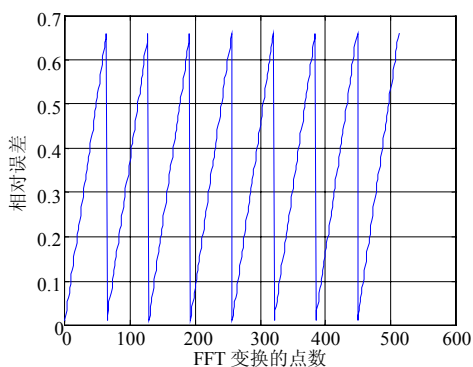


图3 512点FFT相对误差

当输入数据是白噪声时,用硬件实现512点FFT运算结果由有限字长效应引起的各频率点数的相对误差为 $E(n)=|f_{\text{mat}}-f_{\text{chip}}|/|f_{\text{mat}}|$ ,如图3所示。其中, $f_{\text{mat}}$ 为MATLAB模型运算结果; $f_{\text{chip}}$ 为硬件运算结果。在

点数 $N$ 固定时,误差随着各频率点数呈周期性变化,这是由旋转因子与加权因子的周期性变化造成的。

### 4 结论

本文分析了基8 FFT变换的各种误差来源,提出了分析有限字长效应的静态模型,得出了输出噪声信号比与FFT分裂的级数成线性关系以及误差随着各频率点数呈周期性变化的结论。硬件设计工程师可以在满足精度要求的情况下减小字长,提高硬件性能。本文的硬件运算结果都是在中科院微电子所开发的SPEED平台上获得的,进一步验证了分析的正确性。

#### 参考文献

- [1] 王林泉,皮亦鸣,陈晓宁,等.基于FPGA的超高速FFT硬件实现[J].电子科技大学学报,2005,34(2):152-155.
- [2] 龙宁,张凤荔.基于FFT的数字多波束测向算法研究[J].电子科技大学学报,2005,34(1):16-18.
- [3] 程郁凡,韩玮,李少谦.基于FFT的快速跳频同步信号存在性检测[J].电子科技大学学报,2003,32(5):535-540.
- [4] 孙世新,陈平安.与FFT并行算法相适应的体系结构探讨[J].电子科技大学学报,2000,29(5):535-539.
- [5] MOROZOV J A. Effects of quantization in FFT algorithms[C]//International Conference on Actual Problems of Electron Devices Engineering.[S.l.]:IEEE Press,2006:526-533.
- [6] 王世一.数字信号处理[M].北京:北京理工大学出版社,2003.
- [7] PENG Yong-jun. A parallel architecture for VLSI implementation of FFT processor[C]//5th ASIC International Conference. [S. l.]:IEEE,2003:748-751.
- [8] TRAN T, BED E L. Accumulation of roundoff errors in floating point FFT[J]. IEEE Transactions on Circuits and Systems,1977,24(3):132-143.
- [9] HUI C C, TIONG J D. Error analysis of FFT architecture for digital video applications[C]//3th IEEE International Conference on Electronics, Circuits and Systems. Greece:IEEE CAS Society Press,1996:820-823.
- [10] TAE H J, ALAN V O. Effects of FFT coefficient quantization on sinusoidal signal detection[C]//Acoustics, Speech and Signal Processing. New York, USA:IEEE Press,1988:1818-1821.

编辑 黄莘