高吞吐率双模浮点可重构FFT处理器设计实现

魏 星[①②] 黄志洪[①] 杨海钢*[①②]
[①](中国科学院电子学研究所 北京 100190)
[②](中国科学院大学 北京 100190)

摘 要：高吞吐浮点可灵活重构的快速傅里叶变换(FFT)处理器可满足尖端雷达实时成像和高精度科学计算等多种应用需求。与定点FFT相比，浮点运算复杂度更高，使得浮点型FFT的运算吞吐率与其实现面积、功耗之间的矛盾问题尤为突出。鉴于此，为降低运算复杂度，首先将大点数FFT分解成若干个小点数基2^k级联子级实现，提出分别针对128/256/512/1024/2048点FFT的优化混合基算法。同时，结合所提出同时支持单通道单精度和双通道半精度两种浮点模式的新型融合加减与点乘运算单元，首次提出一款高吞吐率双模浮点可变点FFT处理器结构，并在28 nm标准CMOS工艺下进行设计并实现。实验结果表明，单通道单精度和双通道半精度浮点两种模式下的运算吞吐率和输出平均信号量化噪声比分别为3.478 GSample/s, 135 dB和6.957 GSample/s, 60 dB。归一化吞吐率面积比相比于现有其他浮点FFT实现可提高约12倍。

关键词：快速傅里叶变换；双模浮点；混合基；融合运算单元
中图分类号：TN47 **文献标识码**：A **文章编号**：1009-5896(2018)12-3042-09
DOI: 10.11999/JEIT180170

High Throughput Dual-mode Reconfigurable Floating-point FFT Processor

WEI Xing[①②] HUANG Zhihong[①] YANG Haigang[①②]
[①](*Institute of Electronics, Chinese Academy of Sciences, Beijing* 100190, *China*)
[②](*University of Chinese Academy of Sciences, Beijing* 100190, *China*)

Abstract: In the advanced applications of real-time radar imaging and high-precision scientific computing systems, the design of high throughput and reconfigurable Floating-Point (FP) FFT accelerator is significant. Achieving high throughput FP FFT with low area and power cost poses a greater challenge due to high complexity of FP operations in comparison to fixed-point implementations. To address these issues, a serial of mixed-radix algorithms for 128/256/512/1024/2048-point FFT are proposed by decomposing long FFT into short implementations with cascaded radix-2^k stages so that the complexity of multiplications can be significantly reduced. Besides, two novel fused FP add-subtract and dot-product units for dual-mode functionality are proposed, which can either compute on a pair of double precision operands or on two pairs of single precision operands in parallel. Thus, a high throughput dual-mode floating-point variable length FFT is designed. The proposed processor is implemented based on SMIC 28 nm CMOS technology. Simulation results show that the throughput and Signal-to-Quantization Noise Ratio (SQNR) in single-channel single precision and dual-channel half precision floating-point mode are 3.478 GSample/s, 135 dB and 6.957 GSample/s, 60 dB respectively. Compare to the other FP FFT, this processor can achieve 12 times improvement of normalized throughput-area ratio.

Key words: Fast Fourier Transform (FFT); Dual-mode floating point; Mixed-radix; Fused arithmetic unit

收稿日期：2018-02-08；改回日期：2018-07-05；网络出版：2018-07-24
*通信作者： 杨海钢 yanghg@mail.ie.ac.cn
基金项目：国家自然科学基金(61704173, 61474120)，北京市科技重大专项课题(Z171100000117019)
Foundation Items: The National Natural Science Foundation of China (61704173, 61474120), The Major Program of Beijing Science and Technology (Z171100000117019)

1　引言

在高精度科学计算和高分辨率雷达成像等系统中，浮点(Floating Point, FP)型快速傅里叶变换(Fast Fourier Transform, FFT)由于具备较高的信号量化噪声比(Signal to Quantization Noise Ratio, SQNR)而得到广泛的应用[1,2]。不同场景对FFT的要求各不相同，如在雷达实时成像处理等应用中，原始回波数据的前端处理需要进行多种不同点数的高精度FFT以实现脉冲压缩[3]；而雷达图像的后处理部分需要将2维图像分解成大量1维向量进行处理，这对FFT的处理速度有较高要求。因此，为保证整个处理系统的实时性能，探索并设计出高吞吐率点数与精度可灵活重构的FFT处理器成为关键。

当前高吞吐率的FFT设计通常是通过增加并行处理数据路径获得，但其所需的运算资源数目与设计并行度成正比[4]。同时，与定点数运算相比，浮点数运算需要执行多个额外操作步骤，使得单个浮点运算消耗更多的硬件资源和功耗[5]。采用多操作数融合的浮点运算单元，通过共享中间算术逻辑可有效减少浮点运算的面积、功耗和延迟[6]。因此，通过优化设计FFT的实现算法来降低其运算复杂度，同时采用高计算密度的融合浮点运算单元以提高其利用效率，这在高吞吐浮点型FFT的设计实现中尤其关键。

优化FFT的实现算法一直是国内外的研究热点，文献[7]和文献[8]针对512点FFT分别提出两种改进型基2^5和基2^4-2^2-2^3的算法，以降低乘法运算的复杂度，其实际上都属于混合基算法。文献[9]采用基4算法提出一种混合结构实现的低面积开销浮点FFT，但其所采用的分时复用策略无法满足高吞吐应用中运算资源连续占用的需求。相比于采用传统分立浮点单元的实现，文献[10]初步展示了在FFT中引入融合浮点运算单元后，在速度和面积上的全面优势。本文针对浮点FFT设计中普遍存在的高吞吐率和所需面积、功耗等互相矛盾的问题，基于优化的混合基算法提出一款针对128/256/512/1024/2048 5种点数可重构的FFT处理器，同时为进一步提高运算单元的面积利用率，本文提出并设计了同时支持单通道单精度浮点(Single Precision, SP)和双通道半精度浮点(Half Precision, HP)两种工作模式的新型融合加减与点乘单元，并采用标准ASIC流程完成了该高吞吐率双模浮点可变点FFT处理器的设计实现和性能验证。

本文第2节重点介绍混合基算法的原理与核心思想，同时给出本文所采用的算法细节；第3节根据优化的混合基算法，重点阐述了双模浮点可变点FFT的设计实现细节，主要包括顶层实现结构、双模浮点蝶形单元、以及双模浮点融合加减与点乘运算单元等设计；第4节通过综合与仿真的实验结果，对所设计两个浮点运算单元以及双模FFT的整体性能进行比较与分析；最后总结全文。

2　FFT混合基分解算法

2.1　基本混合基算法原理

FFT本质上是一种快速计算离散傅里叶变换(Discrete Fourier Transform, DFT)的优化方法，经典Cooley-Tukey算法利用旋转因子的对称性和周期性，将N点DFT分解成更小的N_1点和N_2点的先后两次DFT，其中N=N_1N_2，表达式为

$$\begin{aligned} X(k) &= \sum_{n_2=0}^{N_2-1}\sum_{n_1=0}^{N_1-1} x(n_2+N_2n_1)\, W_N^{(n_2+N_2n_1)(N_1k_2+k_1)} \\ &= \sum_{n_2=0}^{N_2-1}\left\{\left[\sum_{n_1=0}^{N_1-1} x(n_2+N_2n_1)\, W_{N_1}^{n_1k_1}\right] W_N^{n_2k_1}\right\} W_{N_2}^{n_2k_2} \end{aligned} \tag{1}$$

其中，n_1, k_1∈[0, N_1−1]，n_2, k_2∈[0, N_2−1]。在式(1)中，如果N_2不是质数，则N_2点DFT还可进一步被分解。以此类推，当N=$N_1N_2N_3\cdots N_m$时，N点DFT可被连续分解成多个级联的N_1点，N_2点，…，和N_m点DFT组合。如果N_1=N_2=N_m=R，称为固定基算法(fixed-radix algorithm)，如基2算法等。而如果N点DFT被分解成不同大小点数的组合，则称为混合基算法(mixed-radix algorithm)。特别地，以固定基算法为基础，当R=2^k时，每个R点DFT又由k个级联的基2子级构成，这类算法称为基2^k算法(radix-2^k algorithm)，其乘法运算复杂度与基4算法相同，同时蝶形运算数据流图仍然保持与基2算法一致，具有硬件实现开销小的特点。因此，本文针对不同点数的FFT算法优化围绕基2^k混合基算法展开。

2.2　所提出的优化混合基算法

不同的混合基算法减小运算复杂度的关键是降低旋转因子部分的乘法复杂度。当旋转因子W_M^{Φ}满足Φ=0, M/4, M/2或3M/4时，此时的旋转因子乘法可简化为实部虚部交换与取反等操作。本文将M统称为基底。本文所采用优化混合基算法的核心思想是将大点数FFT分解成若干个小点数，再对每个小点数FFT采用基2^k算法进行实现，从而尽量增加W_4旋转的数目，以降低乘法的运算复杂度。据此提出的算法如表1所示，以128点FFT为例，其可被分解成8点，4点和4点3个级联的主级，每个主级又分别由3, 2和2个子级构成。即相应的混合基算法

表 1 本文所提出的混合基算法

点数	优化算法	每个子级相应的基底									
		1	2	3	4	5	6	7	8	9	10
128	2^3-2^2-2^2	4	8	128	4	16	4				
256	2^4-2^2-2^2	4	16	4	256	4	16	4			
512	2^4-2^2-2^3	4	16	4	512	4	32	4	8		
1024	2^5-2^2-2^3	4	8	32	4	1024	4	32	4	8	
2048	2^5-2^3-2^3	4	8	32	4	2048	4	8	64	4	8

可表示为2^3-2^2-2^2，前6个子级所对应的基底分别是4, 8, 128, 4, 16和4。其他点数的算法推导也均采用上述相同的方法。

3 高吞吐率双模浮点可重构FFT详细设计

3.1 可重构FFT顶层结构

基于上一节中表1给出的混合基优化算法，本文提出一种针对128/256/512/1024/2048点FFT运算的优化实现结构。如图1所示，其整体由3个主级构成，每个主级又分别由5, 3和3个级联的子级构成，通过中间的多路选择器进行数据流处理路径切换以实现可变基。为实现高吞吐率，同时考虑高并行度引入的额外面积开销[11,12]，本文采用并行度为8的多路径反馈(Multi-path Delay Feedback, MDF)流水线并行结构。

第1, 2主级中每个子级包含8路并行双模浮点蝶形运算单元、同步控制的并行FIFO(First-Input First-Output)以及末端用于旋转操作的双模浮点复数乘法器。第3主级采用全并行结构实现8或4点FFT，每个子级包含4个双模浮点蝶形运算单元。不同点数的FFT可通过配置数据路径长度、FIFO访问深度以及旋转因子生成来实现。例如，对于128点FFT, 3个主级中的多路选择器均选通下支路数据，即输入序列依次通过第1主级中的第1, 3, 5子级，以及第2主级中的第2, 3子级，最后经过第3主级中的第1, 3子级运算后输出。模式选择信号sp_hp用于控制蝶形运算单元在单通道单精度浮点(sp_hp=0)和双通道半精度浮点(sp_hp=1)两种工作模式之间的切换。

图1中同时支持两种浮点精度模式的蝶形运算单元，其主要由双模浮点融合加减单元、末端的双模浮点融合点乘单元以及路径交叉选择器构成。为进一步提高运算单元的面积利用率，本文采用高计算密度的多操作数融合浮点运算单元，并对其进行结构优化使其支持SP和HP两种浮点工作模式。两种运算单元的设计均插入3级流水线寄存器以提高运算吞吐率。

其中，双模浮点融合加减运算单元在当sp_hp等于0时，可完成单路SP浮点加法和减法运算，对于输入的SP浮点操作数A和B，同时输出$A-B$和$A+B$；当sp_hp等于1时，同时进行两路HP浮点加和减运算，即输出$A[31:16]-B[31:16]$，$A[31:16]+B[31:16]$和$A[15:0]-B[15:0]$，$A[15:0]+B[15:0]$。对于双模浮点融合点乘单元，当sp_hp等于0时可实现一路单精度点乘运算，即输入SP浮点数A，B，C和D，输出$AB+CD$；而当sp_hp等于1时，同时执行两路半精度点乘运算，即输出$A[31:16]B[31:16]+C[31:16]D[31:16]$和$A[15:0]B[15:0]+C[15:0]D[15:0]$。

符合IEEE754-2008标准[13]的单精度浮点数据位宽为32(包含1 bit符号位、8 bit指数位和23 bit尾数位)，半精度浮点数为16(包含1 bit符号位、5 bit指数位和10 bit尾数位)。基于此，本文设计了如图2所示的数据格式结构，对于一个完整的SP浮点复数，其实部与虚部分别填充一个64 bit数据的高32位和低32位；当处于双通道HP浮点工作模式时，HP浮点复数的两路16 bit实部和两路16 bit虚

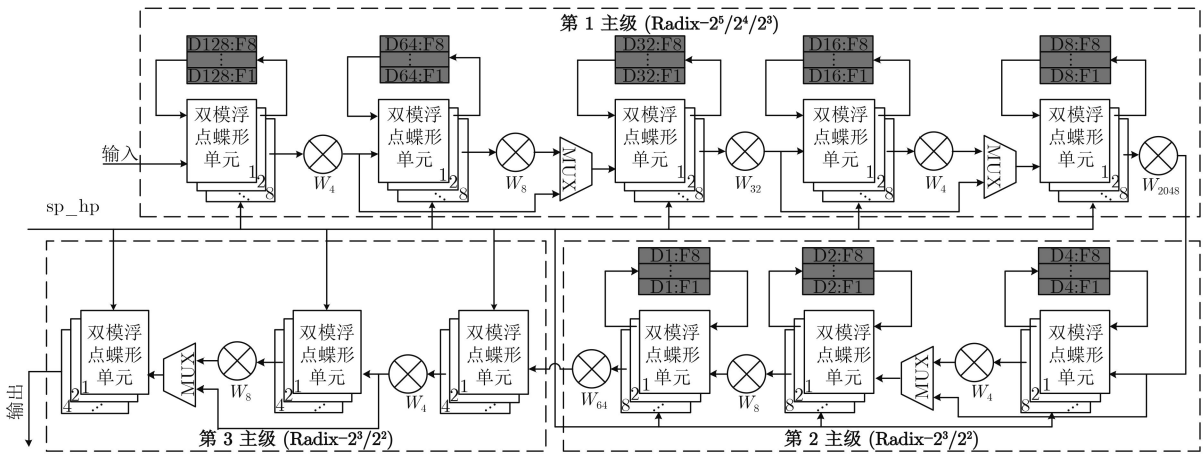

图 1 双模浮点128/256/512/1024/2048点FFT顶层结构框图

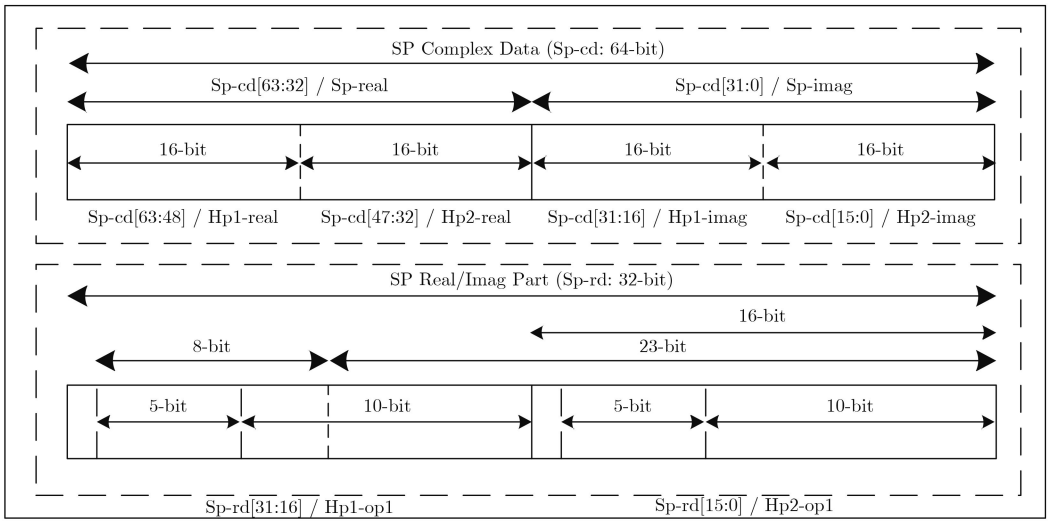

图 2 双模浮点数据格式示意图

部分别占据其高、低32位，即两个紧密排列的HP浮点数正好组成一个SP浮点数的32 bit。相应地，每个蝶形运算单元的输入输出、双模浮点运算单元的输入输出以及FIFO数据总线上的数据也均采用此排列方式。

3.2 双模浮点融合加减运算单元

图3所示为本文所提出的双模浮点融合加减运算单元结构图，为简化描述，图中只给出了尾数部分的处理过程。输入的操作数A和B在经过支持双位宽模式的尾数比较、交换和对齐操作后，产生两个尾数分别同时进行定点加法、减法以及舍入运算。为减少舍入操作模块面积，本文采用复合加、减法器直接计算出正常结果与比正常值大1的结果，最后通过舍入逻辑的输出进行选择。

同时，对于浮点减法，其尾数部分在对齐后作差可能会出现多个高位同时变零的情况，因此需要统计其差值中前导零的个数以进行后续尾数的规格化左移。鉴于此，本文提出了同时支持单通道24 bit和双通道12 bit两种工作模式的前导零预测器(Leading Zero Anticipator, LZA)。如图4所示，其可划分成前导零检测和同步纠正检测两个并列部分。输入的2个(单通道)或4个(双通道)操作数的每个bit成对经过24路并行编码器之后，生成24 bit的W编码和72 bit的ABC编码。W编码的高12位和低12位经过两个相同的前导零检测器(Leading Zero Detector, LZD)后分别产生HP浮点模式下对应两个通道的前导零个数Hp1_Lshift, Hp2_Lshift。LZA本质上是通过特殊的编码方式“等效地”对输入操作数进行作差，同时并不产生进位传播。但其生成的W编码在特定情况下会出现前导0个数比实际小1的问题[14]，因此还需同步产生纠正信号标志此类特例的发生。本文分别将HP浮点模式两个通道36 bit的ABC编码进行树型压缩运算(具体运算规则可参考文献[14])；而对于SP浮点模式，直接将上述双通道树型压缩运算的两个3 bit输出进行一次压缩运算。最后通过末端的简单逻辑运算得到两种模式下的3路纠正信号Hp1_cret, Hp2_cret和Sp_cret。

3.3 双模浮点融合点乘法运算单元

文献[5]中给出了一种单精度浮点融合点乘运算单元的实现结构，文中为提高运算速度，提出了将

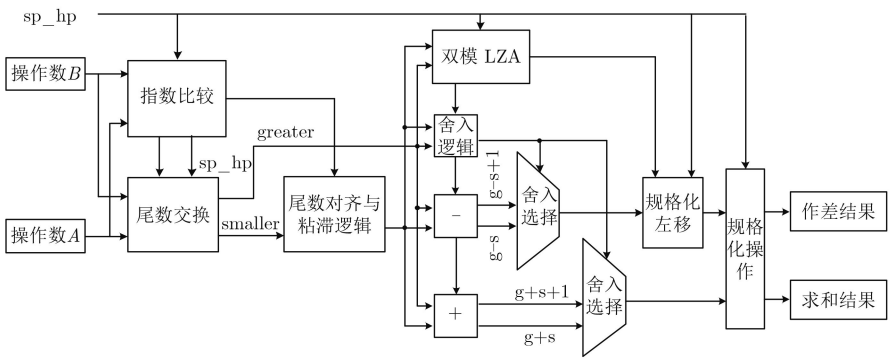

图 3 双模浮点融合加减运算单元结构

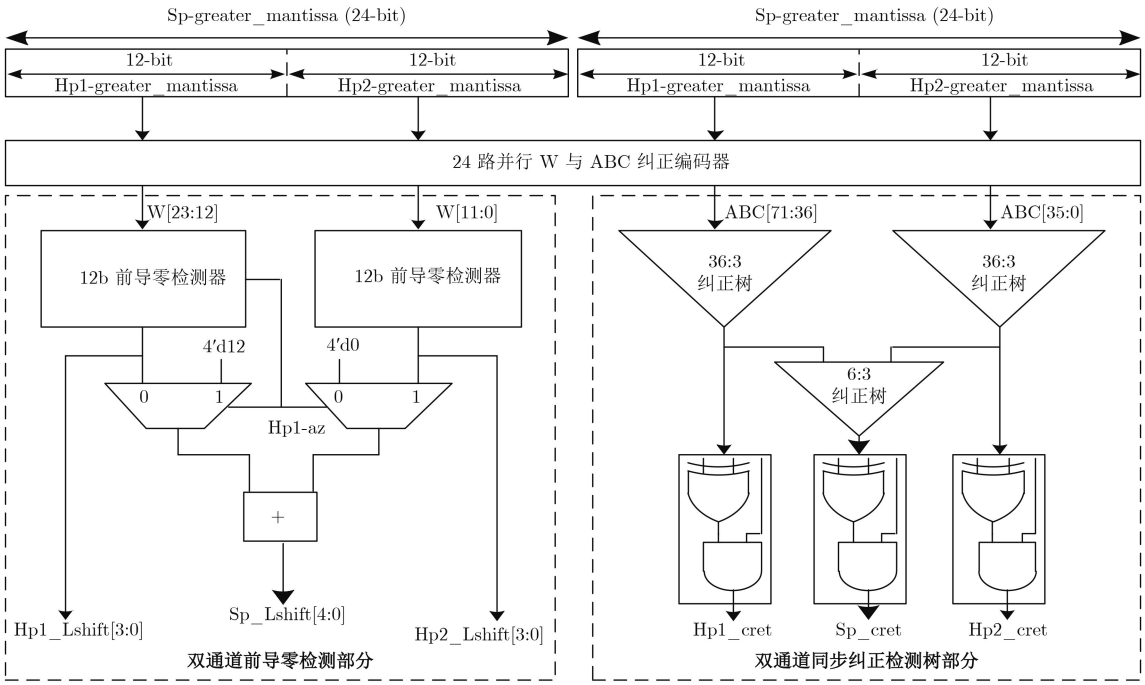

图 4 双模式前导零预测电路结构

规格化左移操作放置在尾数加法器之前，同时引入并行处理的4输入LZA，这直接导致大量额外的面积和功耗开销。针对上述问题，本文首次提出一种双加法舍入处理路径的新型结构，可同时完成绝对值求和、大小比较和舍入等操作，在保持一定工作速度的同时实现较高的面积和能量效率。

所提单元整体结构如图5所示，对于输入的4个操作数A, B, C和D，输入级的数据提取和规格化检查模块完成双浮点模式下符号位、指数位和尾数位的对应提取。然后经过4路2选1数据选择器，分别产生24 bit的尾数Ma, Mb, Mc和Md，作为两个部分积乘法器的输入。与普通乘法器相比，部分积乘法器的输出由于采用CS对(Carry-and-Sum pair)的形式来表示最终乘积结果，可以减少末级一个额外48 bit进位传播加法器的路径延时与面积。

在完成两个CS对的比较和对齐操作后，为提高面积利用率，本文采用两条并列的加法舍入处理路径同时进行绝对值求和、大小比较和舍入等操作，并将规格化左移放置在尾数求和之后。具体原理如下：较大和较小的CS对分别根据AB和CD是否异号进行按位取反，当同号时，均不取反，左右支路加法树的输入相同，其实质上是都在进行两个CS对的加法；反之，当异号时，对于左支路，较大的CS对取反，对于右支路，较小的CS对取反，其实质上是进行两个CS对相互的减法。经过取反选择器后的CS对分别作为左右两个双模式4:2进位保留加法器(Dual-Mode Carry Save Adder, DM-CSA)的输入。

在半精度模式下，每个DM-CSA产生的CS对输出按照从高位到低位被等分成4组(L1-L4和R1-R4)，其中L1, L2, R1, R2和L3, L4, R3, R4分别代表两个半精度通道加法树的中间结果，此时两个通道相互独立。然后L2与L1, L4与L3, R2与R1, R4与R3分别执行进位传播加法，同时L2和L4, R2和R4均产生中间进位和向上舍入标志作为其对应高位加法器的进位输入。而在单精度模式下，L1与L2, L3与L4之间仅有进位标志传递，且L3和L2之间有同时有进位和向上舍入标志的传递，等效地将DM-CSA产生的CS对进行二等分，右支路的处理也类似。左右两条路径分别将多个求和结果合并成两个大位宽的输出(sum_L和sum_R)，然后通过对两个“和”的高位进行简单逻辑运算以生成最终大小比较信号，从而选择其中一路作为绝对值输出。最后经过双模LZD、规格化左移等操作，即可生成最终的尾数。

图5中的双模CSA将两路CS对进行4:2加法树压缩，并输出一个CS对作为单通道单精度或双通道半精度尾数乘积压缩求和的结果。图6所示为同时支持51 bit和25 bit两种位宽模式的4:2 CSA设计框图，包含左边2个26 bit和右边2个25 bit的3:2进位保留加法器。设计思想是当处于双通道25 bit位宽模式时，每列的两个CSA级联组合成一个独立的4:2 CSA，左右两列的进位输入也分别选择对应通道的进位信号，进行减法情况下的求补操作；而当

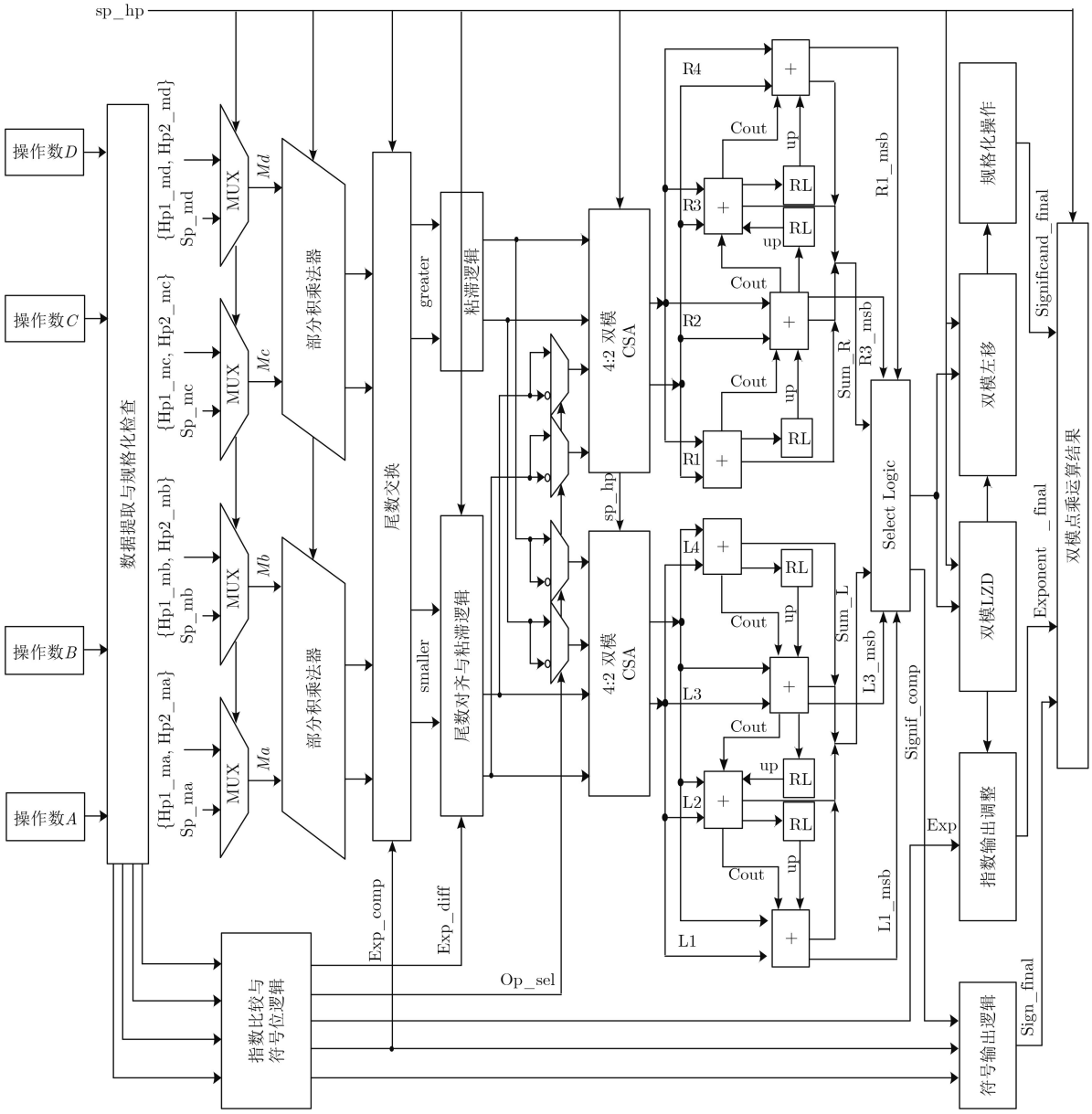

图 5 双模浮点融合点乘运算单元结构

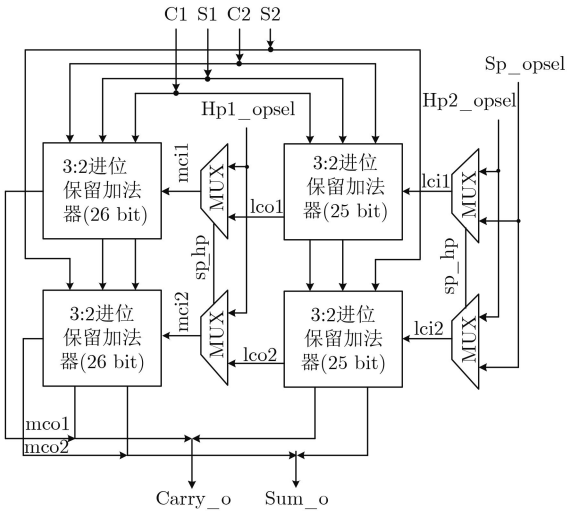

图 6 双模式4:2 CSA结构

处于单通道51 bit位宽模式时，中间的两个多路选择器分别选择右边低位3:2 CSA产生的进位信号，相应地作为左边高位3:2 CSA的进位信号。同时，右列低位3:2 CSA的进位输入选择单通道模式下的异号标志信号，从而实现大位宽4:2加法树压缩的功能。

4　实现与结果分析

4.1　实现环境与结果评估

本文所提出的双模浮点可重构FFT处理器通过Verilog HDL语言描述，采用标准ASIC设计流程，在SMIC 28 nm标准工作电压为1.05 V的CMOS工艺下，通过Design Complier工具的Topographical模式在PVT参数为(SS, 0.945 V, 125°C)下进行物

理逻辑综合，并对整个设计的面积、功耗以及工作频率等进行评估分析。对于输出SQNR，采用连续输入5组叠加高斯白噪声的信号，然后通过与Matlab中双精度浮点FFT的计算结果进行比较计算，最后取平均值得到。为了分别评估本文提出的双模浮点运算单元和优化混合基算法实现的双模可重构FFT结构的性能优劣，本节分别展开其与前人工作的各项性能指标对比与分析。

4.2 双模浮点运算单元性能比较分析

对于本文设计的双模浮点融合运算单元，均兼容IEEE754-2008标准。为了更加公平地与其他相关工作进行对比，本文分别选择了文献[5]和文献[6]中两种高速型浮点融合运算结构，并均采用其综合结果作为对比参照。同时在相同结构与实现环境下，对应设计并实现了只支持单精度浮点单一模式的融合加减与点乘运算单元，分别对应表中的参考结构T1, T2。考虑到不同工艺实现的因素，本文统一将运算单元的实现面积按照式(2)进行归一化处理[15]。同时，引入了功耗周期乘积来衡量单元的综合能量效率水平，这个乘积结果越小，代表能效水平越高。

$$\text{运算单元归一化面积} = \text{有效面积}/(L_{\min}/28\ \text{nm})^2 \tag{2}$$

其中，$L_{\min}$为实现工艺的特征尺寸(单位：nm)。

通过对双模浮点融合加减运算单元综合后的网表进行时序分析，表2中给出了3个流水级中根据模块划分的关键路径分析结果。由于存在多个功能模块并行处理的情况，故仅展示处于关键路径上的模块延时。表3给出了其与两个参照设计的性能对比结果，文献[6]中提出的双处理路径结构需要根据两个操作数指数阶差的大小，对尾数部分分别进行处理。其在关键路径上可分别省去规格化左移和舍入操作的延时，使工作频率得到提升，然而额外增加的处理路径导致较大的面积和功耗开销。因此，与文献[6]进行对比，本文的设计实现面积减少56%，虽然工作频率有所下降，但整体能效可优化69%；与仅单精度的设计实现T1相比，所需面积和能效分别优化13%，30%。

对于本文提出的双模浮点融合点乘运算单元设计，其关键路径与具体实现结果分别如表4和5所示。作为优化其能效和面积的关键因素，双加法舍入处理路径的延时与一个53 bit全加器相当，且仅需约两个53 bit全加器的面积开销即可完成对两组尾数部分积的求和操作，以及大小比较和舍入运算。与文献[5]相比，实现面积和能效分别优化17%和50%；与参考结构T2相比，本文的设计扩展支持了双通道半精度浮点的工作模式，所需面积基本相当，仅增加4%，同时能效仅增加7%。

4.3 双模浮点FFT整体性能比较分析

为了对本文与现有其他FFT设计方案的性能进行对比，本文选择了不同定点位宽[2,16]与单精度浮点[9]的多种类型FFT设计实现作为比较对象，分别与其版图和逻辑综合结果进行对比。考虑到不同点数FFT处理器实现的工艺、并行度和架构实现存在差别，为公平比较，本文利用文献[17]提出的面积归一化公式进行评估，如式(3)所示。同时定义运

表 2 双模浮点融合加减运算单元关键路径

模块名	流水级	延时(ns)
数据提取与尾数生成	1	0.43
指数比较与尾数交换	1	0.77
指数阶差与尾数对齐	2	1.98
尾数求和差	3	1.47
LZA	3	0.35
规格化左移	3	0.16

表 3 双模浮点融合加减运算单元性能比较结果

参数	文献[6]	参考结构T1	本文
工艺 (nm)	45	28	28
归一化面积 (μm^2)	5226 (100%)	2665 (51%)	2317 (44%)
工作频率 (MHz)	1920	500	435
计算延迟 (ns)	1.0	6.0	6.9
功耗 (mW)	5.2	0.6	0.4
功耗×周期	2.70 (100%)	1.20 (44%)	0.84 (31%)

表 4 双模浮点融合点乘运算单元关键路径

模块名	流水级	延时(ns)
数据提取与指数比较	1	0.33
尾数部分积相乘	1	0.87
指数阶差与乘积对齐	2	1.25
双模4:2 CSA	2	0.20
双加法舍入路径	2	0.53
LZD与规格化左移	3	1.98

表 5 双模浮点融合点乘运算单元性能比较结果

参数	文献[5]	参考结构T2	本文
工艺 (nm)	45	28	28
归一化面积 (μm^2)	12865 (100%)	10336 (80%)	10701 (83%)
工作频率 (MHz)	1493	500	435
计算延迟 (ns)	2.1	6.0	6.9
功耗 (mW)	16.9	2.7	2.5
功耗×周期	11.3 (100%)	5.3 (47%)	5.6 (50%)

算吞吐率与归一化面积的比例作为衡量FFT综合性能的参数。

具体实现与对比结果如表6所示，本文设计的双模浮点可重构FFT最高工作频率为435 MHz，平均功耗仅104 mW。单精度和半精度模式下的运算吞吐率分别可高达3478 MSample/s和6957 MSample/s，平均输出SQNR为135 dB和60 dB，2048点浮点FFT的处理时间仅为2.2 μs。

表 6　双模浮点FFT整体性能对比

性能参数	文献[2]	文献[9]	文献[16]	本文
工艺 (nm)	90	65	55	28
FFT结构	基于存储器	Hybrid	MDF	MDF
FFT点数	512	1024	1024	128/256/512/1024/2048
并行度	8	1	1	8
数据类型：字长 (bit)	块浮点：12	浮点：32	定点：16	浮点：32/浮点：16
平均输出SQNR (dB)	57	139	55	SP：135/HP：60
时钟频率 (MHz)	324	400	200	435
计算时间 (μs)	0.3	2.6	5.1	0.2/0.3/0.6/1.1/2.2
运算吞吐率 (MSample/s)	2592	400	200	SP：3478/HP：6957
平均功耗 (mW)	42	417	8	104 @435 MHz
有效面积 (mm^2)	0.93	1.19	0.15	1.41
归一化面积	10.0	22.1	1.9	16.0
归一化吞吐率面积比	259	18	103	SP：220/HP：440

归一化面积相比文献[9]中的浮点FFT减少了27.6%，与部分定点FFT实现相比，归一化面积也维持相当的水平。综合归一化运算吞吐率面积比在单精度模式下是文献[9]的12倍，与文献[2]中采用块浮点技术提高运算精度的高吞吐率FFT设计相比，本文半精度模式下的输出SQNR在维持一定优势的同时，归一化运算吞吐率面积比是其1.7倍。此外，相比超高速全并行结构实现的定点FFT[4](处理字长16 bit，输出SQNR为23 dB)，本文设计的浮点FFT处理器在两种浮点模式下输出SQNR均表现出明显的优势。

$$\text{FFT 整体归一化面积} = 1000A \Big/ \left[M\left(\log_2 N\right)\left(L_{\min}/28\right)^2\right] \tag{3}$$

其中，A为有效面积(单位：mm^2)，M为设计并行度，N为所支持的最大FFT点数。

5　结束语

本文首次提出了一款高吞吐率双模浮点128/256/512/1024/2048点可重构FFT处理器，并在28 nm标准CMOS工艺下进行设计并实现。实验结果表明，本文所设计的浮点FFT在最高工作频率435 MHz下平均功耗仅为104 mW。单通道单精度和双通道半精度浮点两种模式下的输出SQNR分别为135 dB和60 dB。同时，基于所提出的优化混合基算法，以及所提出的同时支持单通道单精度和双通道半精度两种浮点模式的融合加减与点乘单元，归一化面积和吞吐率面积比等性能相比前人其他设计实现均有明显的优势。

参 考 文 献

[1] 吕倩, 苏涛. 基于改进型快速双线性参数估计的复杂运动目标ISAR成像[J]. 电子与信息学报, 2016, 38(9): 2301–2308. doi: 10.11999/JEIT151359.
LÜ Qian and SU Tao. ISAR imaging of targets with complex motion based on the modified fast bilinear parameter estimation[J]. *Journal of Electronics & Information Technology*, 2016, 38(9): 2301–2308. doi: 10.11999/JEIT151359.

[2] HUANG Shenjui and CHEN S. A high-throughput radix-16 FFT processor with parallel and normal input/output ordering for IEEE 802.15.3c systems[J]. *IEEE Transactions on Circuits and Systems* Ⅰ: *Regular Papers*, 2012, 59(8): 1752–1765. doi: 10.1109/TCSI.2011.2180430.

[3] LAN G and FRANK H. Digital Processing of Synthetic Aperture Radar Data: Algorithms and Implementation[M]. Boston: Artech House Publishers, 2005: 154–210.

[4] 陈杰男, 费超, 袁建生, 等. 超高速全并行快速傅里叶变换器[J]. 电子与信息学报, 2016, 38(9): 2410–2414. doi: 10.11999/JEIT160036.
CHEN Jienan, FEI Chao, YUAN Jiansheng, *et al*. An ultra-high-speed fully-parallel fast Fourier transform design[J].

Journal of Electronics & Information Technology, 2016, 38(9): 2410–2414. doi: 10.11999/JEIT160036.

[5] JONGWOOK S and EARL E. Improved architectures for a floating-point fused dot product unit[C]. IEEE Symposium on Computer Arithmetic, Austin, USA, 2013: 41–48. doi: 10.1109/ARITH.2013.26.

[6] JONGWOOK S and EARL E. Improved architectures for a fused floating-point add-subtract unit[J]. *IEEE Transactions on Circuits and Systems* Ⅰ: *Regular Papers*, 2012, 59(10): 2285–2291. doi: 10.1109/TCSI.2012.2188955.

[7] CHO T and LEE H. A high-speed low-complexity modified radix-2^5 FFT processor for high rate WPAN applications[J]. *IEEE Transactions on Very Large Scale Integration* (*VLSI*) *Systems*, 2013, 21(1): 187–191. doi: 10.1109/TVLSI.2011.2182068.

[8] WANG Chao, YAN Yuwei, and FU Xiaoyu. A high-throughput low-complexity radix-2^4-2^2-2^3 FFT/IFFT processor with parallel and normal input/output order for IEEE 802.11ad systems[J]. *IEEE Transactions on Very Large Scale Integration* (*VLSI*) *Systems*, 2015, 23(11): 2728–2732. doi: 10.1109/TVLSI.2014.2365586.

[9] WANG Mingyu and LI Zhaolin. A hybrid SDC/SDF architecture for area and power minimization of floating-point FFT computations[C]. IEEE International Symposium on Circuits and Systems, Montreal, Canada, 2016: 2170–2173. doi: 10.1109/ISCAS.2016.7539011.

[10] EARL E and HANI H. FFT implementation with fused floating-point operations[J]. *IEEE Transactions on Computers*, 2012, 61(2): 284–288. doi: 10.1109/TC.2010.271.

[11] TANG S N, TSAI J W, and CHANG T Y. A 2.4-GS/s FFT processor for OFDM-based WPAN applications[J]. *IEEE Transactions on Circuits and Systems* Ⅱ: *Express Briefs*, 2010, 57(6): 451–455. doi: 10.1109/TCSII.2010.2048373.

[12] NIE Zedong, ZHANG Fengjuan, LI Jie, *et al*. Low-power digital ASIC for on-chip spectral analysis of low-frequency physiological signals[J]. *Journal of Semiconductors*, 2012, 33(6): 67–70. doi: 10.1088/1674-4926.

[13] IEEE 754-2008. IEEE Standard for Floating-Point Arithmetic[S]. 2008. doi: 10.1109/IEEESTD.2008.5976968.

[14] PETER K. Correcting the normalization shift of redundant binary representations[J]. *IEEE Transactions on Computers*, 2009, 58(10): 1453–1439. doi: 10.1109/TC.2009.38.

[15] YANG C H, YU T H, and DEJAN M. Power and area minimization of reconfigurable FFT processors: A 3GPP-LTE example[J]. *IEEE Journal of Solid-State Circuits*, 2011, 47(3): 757–768. doi: 10.1109/JSSC.2011.2176163.

[16] MARIO G, HUANG S J, CHEN S G, *et al*. The serial commutator FFT[J]. *IEEE Transactions on Circuits and Systems* Ⅱ: *Express Briefs*, 2016, 63(10): 974–978. doi: 10.1109/TCSII.2016.2538119.

[17] YANG K J, TSAI S H, and CHUANG G. MDC FFT/IFFT processor with variable length for MIMO-OFDM systems[J]. *IEEE Transactions on Very Large Scale Integration*(*VLSI*) *Systems*, 2013, 21(4): 720–731. doi: 10.1109/TVLSI.2012.2194315.

魏 星：男，1991年生，博士，研究方向为算法硬件加速设计、可重构计算芯片架构设计.

黄志洪：男，1984年生，助理研究员，研究方向为可编程逻辑结构设计、新型卷积神经网络芯片体系架构开发.

杨海钢：男，1960年生，研究员，研究方向为数模混合信号集成电路设计、超大规模集成电路设计等.