

DOI: 10.7652/xjtuxb201804016

一种高效存储多级二维 9/7 离散小波变换结构

王鑫¹, 高家明², 梁煜², 张为²

(1. 公安部天津消防研究所, 300381, 天津; 2. 天津大学微电子学院, 300072, 天津)

摘要: 针对 JPEG2000 图像压缩编码硬件实现中多级离散小波变换模块(DWT)与优化截断嵌入式模块(EBCOT)之间大量小波系数存储带来的缓存问题,提出了一种新型高效存储的多级二维 9/7 离散小波变换结构。首先,基于 9/7 小波变换提升算法,设计了可满足并行结构要求的基本处理单元,并在此基础上搭建了二维变换模块;然后采用一种新型的数据扫描方式,通过对图像分块并进行块间组合扫描,大幅降低了 DWT 模块与 EBCOT 模块间的小波系数缓存;最后根据数据扫描特点完成了非折叠结构多级二维小波变换模块的构建。仿真实验结果表明,三级 9/7 离散小波变换结构处理一幅大小为 512×512 像素的图像,对比已有硬件结构可节约存储资源 40%以上。

关键词: 高效存储;离散小波变换;块扫描;图像压缩

中图分类号: TN919.81 **文献标志码:** A **文章编号:** 0253-987X(2018)04-0111-06

A Multi-Level 2-D 9/7 DWT Architecture with Efficient Memory

WANG Xin¹, GAO Jiaming², LIANG Yu², ZHANG Wei²

(1. Tianjin Fire Research Institute of Ministry of Public Security, Tianjin 300381, China;

2. School of Microelectronics, Tianjin University, Tianjin 300072, China)

Abstract: A novel multi-level two-dimensional 9/7 discrete wavelet transform (DWT) architecture with efficient memory is proposed to solve the problem that a large number of wavelet coefficients between the multi-level 2-D 9/7 DWT (2-D 9/7 DWT) module and the embedded block coding with optimized truncation (EBCOT) module need to be cached in the hardware implementation of JPEG2000 image compression and to decrease the cost of cache. A parallelizable basic processing unit is firstly designed based on the 9/7 DWT lifting scheme, and then a 2-D transform module is built. A novel block-based data scanning method is then adopted to scan combined blocks of an image so that the storage of wavelet coefficients between the modules is dramatically reduced. Finally, a unfolded multi-level 2-D DWT module is constructed according to the characteristics of data scanning. Experimental results and comparisons with existing hardware structures show that the proposed 3-level 2-D 9/7 DWT architecture saves at least 40% of storage resources for an image with the size of 512×512 pixels.

Keywords: efficient memory; discrete wavelet transform; block-based scanning; image compression

JPEG2000 是一种在图像压缩领域广泛应用的 新型图像压缩标准。二维离散小波变换(discrete

收稿日期: 2017-06-30。 作者简介: 王鑫(1983—),男,助理研究员;张为(通信作者),男,教授,博士生导师。 基金项目: 国家自然科学基金资助项目(61474080);公安部技术研究计划竞争性遴选项目(2016JSYJD04-03)。

网络出版时间: 2018-02-26 网络出版地址: <http://kns.cnki.net/kcms/detail/61.1069.T.20180226.1946.006.html>

wavelet transform, DWT) 由于其良好的时频局部化性能^[1] 成为 JPEG2000 压缩的核心技术之一。二维 DWT 将图像信息分解为多级高频分量与低频分量, 为优化截断嵌入式块编码(embedded block coding with optimized truncation, EBCOT) 提供原始数据。EBCOT 中位平面编码模块以码块形式读取小波系数并将其处理为上下文信息, 之后经熵编码和率失真截断编码形成压缩码流。由于算法计算量较大, 为满足高速处理要求, 高效的硬件 DWT 设计优化逐渐成为研究热点。

对于单级 9/7 DWT 硬件实现, 文献[2]提出了一种横向 Z 型扫描的双输入/双输出结构, 其关键路径为一个乘法器延时, 但是该结构由于并行度较低导致整体处理时间较长, 难以满足高速需求。文献[3]对输入数据进行重排, 通过翻转结构进行模块设计降低了图像整体处理时间, 然而其内部数据缓存较高同时控制逻辑较为复杂。Hu 和 Jong 在文献[4]中引入基于纵向扫描的并行结构, 提高了系统的处理速度, 但是该结构需要整张图像存储完成后才能开始处理数据, 故存储需求较大。文献[5]中采用流水线并引入正则乘法结构降低关键路径为一个加法器延时, 不过其结构中的系数精度较低运算误差较大, 难以满足应用要求。文献[6]采用了高系数精度的无乘法器结构以降低数据误差, 并消除了片外存储资源消耗, 但是该结构并行度同样较低, 因此整体处理时间较长不利于高速数据处理应用环境。

对于多级结构优化, 文献[7]提出一种通过并行度调整消除级间存储的多级结构, 但关键路径延时仍旧较大。文献[8]通过高并行度减低存储面积, 同时减低关键路径为一个乘法器延时, 不过该结构基于卷积算法进行设计运算复杂度较高。文献[9]通过引入重复扫描的方式设计了新型高并行结构, 消除了第一级变换模块内的级内变换缓存, 但是该结构运算资源较高且级间处理较为复杂。Lai 和 Chung 提出了一种码块优化的扫描方式^[10], 生成的多级小波系数可由并行 EBCOT 直接处理, 降低了模块间的存储。然而, 该方法由于采用了折叠多级结构, 导致模块内暂存极大, 同时结构中码块边界的近似处理也降低了压缩质量。文献[11]也基于块扫描对扫描方式进行优化以消除模块间缓存, 然而级间缓存过大仍然严重制约了系统性能。

本文通过对已有结构的分析, 设计了一种应用于 JPEG2000 硬件实现的新型高效存储的非折叠多级二维 9/7 DWT 结构, 通过对扫描方式的改进、数

据并行度的调整以及模块内数据传递的优化, 极大的减少了系统面积, 提升了硬件效率。

1 基本模块设计

1.1 优化提升算法

文献[12]根据已有 9/7 DWT 提升算法, 通过一系列变换提出的优化提升算法公式如下

$$D_1(n) = \frac{1}{\alpha}x(2n+1) + x(2n) \quad (1)$$

$$D_2(n) = \left(\frac{1}{\alpha\beta} + 1\right)x(2n) + \frac{1}{\alpha}x(2n-1) + x(2n-2) \quad (2)$$

$$D_3(n) = \frac{1}{\gamma}y(2n+1) + y(2n) \quad (3)$$

$$D_4(n) = \left(\frac{1}{\delta\gamma} + 1\right)y(2n) + \frac{1}{\gamma}y(2n-1) + y(2n-2) \quad (4)$$

$$\frac{1}{\alpha}y(2n+1) = D_1(n) + x(2n+2) \quad (5)$$

$$\frac{1}{\alpha\beta}y(2n) = D_2(n) + D_1(n) + x(2n+2) \quad (6)$$

$$\frac{1}{\gamma}H(2n+1) = D_3(n) + y(2n+2) \quad (7)$$

$$\frac{1}{\delta}L(2n) = D_4(n) + D_3(n) + y(2n+2) \quad (8)$$

$$S_H = KH(2n+1) \quad (9)$$

$$S_L = \frac{1}{K}L(2n) \quad (10)$$

式中: $D_1(n) \sim D_4(n)$ 为 4 个中间变量; $H(2n+1)$ 为归一化前的高频小波系数; $L(2n)$ 为归一化前的低频小波系数; $x(n)$ 代表原始输入像素值; n 为像素点坐标; S_H 和 S_L 代表变换最终产生的高频小波系数和低频小波系数; $\alpha, \beta, \gamma, \delta$ 和 K 为变换系数。对于二维变换, 数据经过行列变换处理后最终得到 4 个小波系数即低频-低频系数 S_{LL} 、高频-低频系数 S_{HL} 、低频-高频系数 S_{LH} 和高频-高频系数 S_{HH} 。

1.2 基本处理单元

列变换模块与行变换模块均采用双输入/双输出流水线结构, 两个模块的关键路径延时均为一个乘法器延时, 其中列变换模块结构如图 1 所示。

本文设计的数据基本输入方式为横向并行双输入, 列变换模块需要对中间变量 $D_1 \sim D_4$ 进行存储, 对于大小为 $N \times N$ 的输入图像, 所需的存储空间为 $4N$ 。图 1 中 RAM1、RAM2、RAM3、RAM4 分别存储中间值 D_1, D_2, D_3, D_4 。当设计并行结构时, 4 个中间值可由该级的上一个并行模块提供, 只需要在

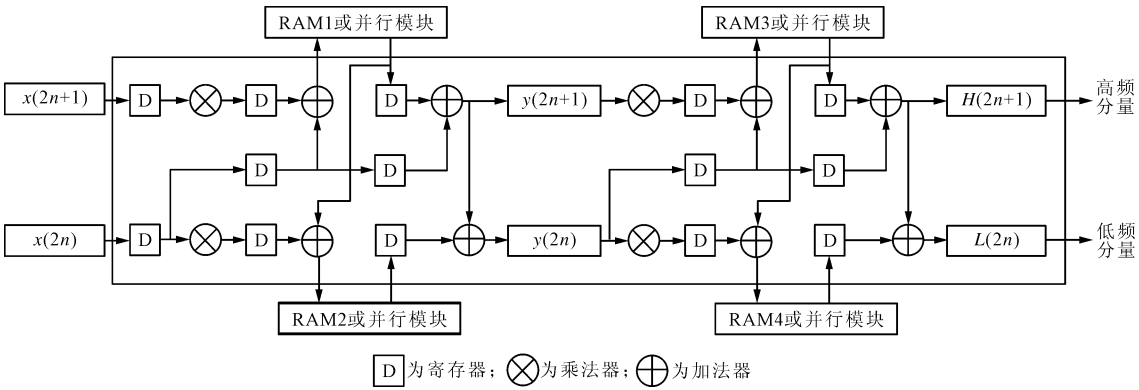


图 1 列变换模块结构

并行结构的首尾模块与 RAM 中进行中间值数据读写。列变换与行变换模块的内部结构一致,对于行变换模块而言,横向并行输入的数据为连续数据,因此每个中间值只需要 2 个数据暂存即可满足数据缓存需求。

由列变换模块输出的低频与高频系数需要经过转置后进入行变换模块进行处理,该转置模块结构如图 2 所示。由于列变换与行变换均为双输入双输出结构,且扫描方式为横向并行,因而需 3 个寄存器即可对列变换输出数据进行转置,满足行变换输入需求。

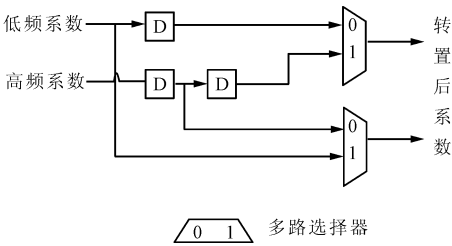


图 2 转置模块结构

由行变换输出数据进行缩放后得到最终的二维离散小波系数。缩放模块将待缩放的小波系数进行单次缩放,缩放模块设计结构如图 3 所示。

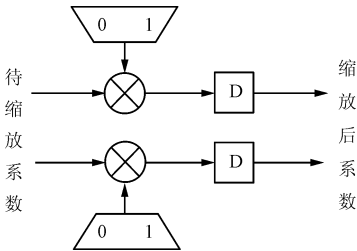


图 3 缩放模块结构

1.3 二维变换模块

本文所设计的二维 9/7 DWT 变换模块结构如图 4 所示,整体结构为双输入双输出,数据进入后,

依次由列变换模块、转置模块、行变换模块和缩放模块进行处理,最终输出离散小波系数。

上述各个模块组成的逻辑单元可单独对输入数据进行离散小波变换处理,也可并行搭建满足任意并行度下的高吞吐率的数据处理。当并行搭建逻辑单元以处理高并行度数据时,任意两单元间的列变换模块所需的大小为 $4N$ 的暂存可被单元间数据传递所取代,因此无论系统并行度大小,系统所需变换缓存恒为 $4N$ 。

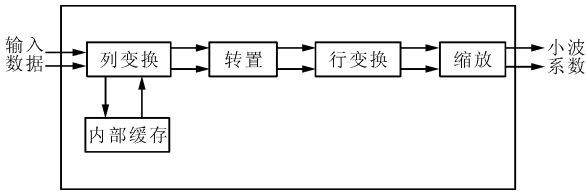


图 4 本文设计的二维 9/7 DWT 变换模块结构

2 整体结构设计

2.1 非折叠结构

根据 JPEG2000 算法,输入图像数据经过多级二维 9/7 DWT 模块处理后得到的小波系数会进入 EBCOT 模块进行进一步处理。多级 9/7 DWT 模块分为折叠结构与非折叠结构:折叠结构为每级变换依次处理,级间的 S_{LL} 分量需要进行全部存储;非折叠结构为多级变换共同进行,通过运算处理单元的增加来加快处理速度、减小缓存面积。由于非折叠结构处理速度更快、数据缓存更小,因此本设计采用非折叠结构。

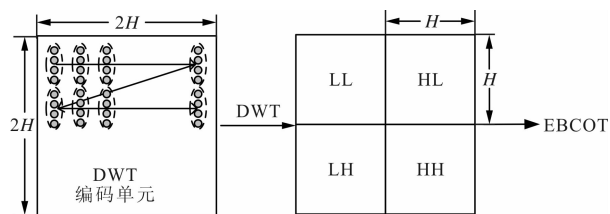
非折叠结构中多级离散小波系数同时产生,因此可以由并行 EBCOT 模块即时处理。然而,由于 EBCOT 采用的编码为基于码块的条带读取编码方式,码块大小为 $H \times H$,传统的基于行扫描、列扫描或并行多输入扫描方式所产生的离散小波系数无法

满足 EBCOT 即时编码的需求,因此,仍要经过存储后进行 EBCOT 编码,通常需 N^2 存储空间进行缓存。

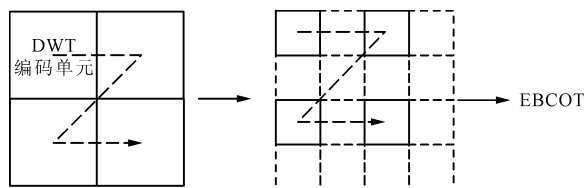
基于块扫描的 9/7DWT 结构则可以在单级中直接产生块状分布的离散小波系数用于 EBCOT 即时编码。但是,在多级 DWT 结构中,每一级的 S_{LL} 分量在下一级继续处理时数据流破碎,无法在每一级均产生恒定大小的 EBCOT 编码块。因此传统的基于块的扫描方式需要进行进一步的优化。

2.2 扫描方式优化

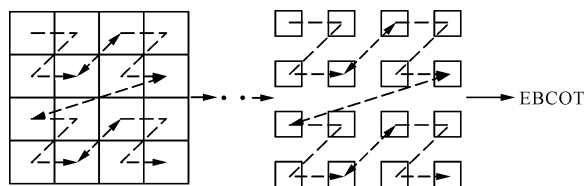
对于一级二维结构,基于块扫描的结构可以直接得到离散小波系数码块。因此设定扫描的 9/7 DWT 码块大小为 $2H \times 2H$,则可获得 4 个大小为 $H \times H$ 的可用于即时处理的 EBCOT 编码块,如图 5a 所示。图像划分为若干 $2H \times 2H$ 的 DWT 编码单元扫描,得到的系数只需占用 $4H^2$ 的存储空间即可被 EBCOT 模块处理。



(a) 一级变换扫描方式



(b) 两级变换扫描方式



(c) 三级变换扫描方式

图5 优化块组合扫描方式

对于两级二维结构,第一级产生的码块可直接供 EBCOT 模块处理。而第二级变换的输入数据为 $H \times H$ 的 S_{LL1} 码块,经变换处理后会得到 4 个大小为 $H/2 \times H/2$ 的小波系数块,这显然不能满足 EBCOT 的编码需求。需对块扫描方式进行优化,调整 DWT 码块的扫描顺序使其满足在第二级最快速度产生大小为 $H \times H$ 的 EBCOT 编码块的需求。因

此在数据扫描时,将相邻 4 个大小为 $2H \times 2H$ 的码块组成 DWT 基本编码单元,以 Z 型方式进行扫描如图 5b 所示。第一级处理后每个基本单元均产生大小为 $H \times H$ 的码块,并以 Z 型顺序进行第二级处理,这样相邻 4 个大小为 $H \times H$ 的 S_{LL1} 码块可以在最短时间内产生 EBCOT 编码块。

对于三级二维结构,为保证在三级变换中同时生成 EBCOT 编码块,对 16 个大小为 $2H \times 2H$ 的码块进行 Z 型嵌套扫描,即在 4 个基本编码单元扫描顺序基础上,以其为单位整体采用 Z 型方式进行扫描,如图 5c 所示。则在第三级变换中,毗邻的 16 个大小为 $H/2 \times H/2$ 的 S_{LL2} 码块可以以最快速度产生 EBCOT 编码块。

综上,对于级数为 L 的多级离散小波变换模块,为最大化非折叠结构的优势,多级同时产生可直接用于 EBCOT 编码的码块,可将图像划分为若干大小为 $4^{(L-1)} \times 2H \times 2H$ 的 DWT 基本编码单元,在单元内部按照 Z 型嵌套的扫描方式完成模块扫描。DWT 基本编码单元之间横向顺次进入 DWT 模块即可完成一副图像的高效编码。

2.3 并行度优化

本文所设计的码块内数据采用硬件效率较高的横向并行扫描方式。由于数据每经过一级数据处理并行度即降低为上一级的 $1/2$,因此本结构中每级的二维 DWT 模块并行度也降为上一级的 $1/2$ 。以三级系统为例,第一级并行度为 S ,则第二级并行度为 $S/2$,第三级并行度为 $S/4$ 。

由于数据并行度与结构并行度一致,因此非折叠结构级间的结构缓存得以消除。对于每一级而言,码块的扫描次序不变而只有码块的大小发生了改变,因此每一级的控制只需修改码块大小参数,极大的降低了系统控制的复杂性。

2.4 边界处理

由于采用优化码块扫描方式,需要对码块数据进行妥善处理以避免降低压缩质量。在本文设计中,每一个码块都需在其上边界和左边界接收数据,在其下边界和右边界传出数据,如图 6a 所示。由于存储单元为复用,因此总体边界缓存相当于在缓存每个基本编码块两边界,如图 6b 所示。

边界的数据传输即为变换模块中所传输的中间值 D_1 、 D_2 、 D_3 与 D_4 。由于缓存空间为复用,对于一幅大小为 $N \times N$ 像素的图像,列变换模块每一级所需暂存大小为 $4 \times \frac{N}{2^{j-1}}$, j 为当前变换级数,因此 L

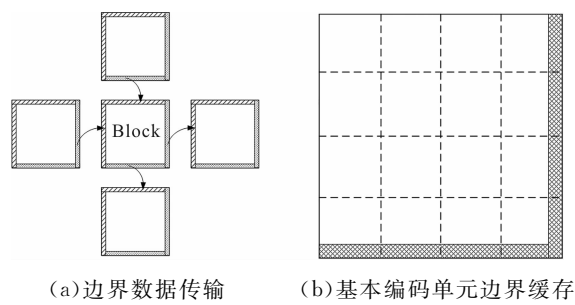


图 6 边界处理方式

级变换所需列变换暂存为

$$n_{\text{col}} = \sum_{j=1}^L 4 \frac{N}{2^{j-1}} = \sum_{j=1}^L \frac{N}{2^{j-3}} \tag{11}$$

由于 DWT 基本编码单元横向顺次进入 DWT 模块列变换,左右边界的数据传输即行变换模块数据寄存只需深度为 8 个基本处理单元大小的存储空间,用于分别存储 H 分量与 L 分量的中间值 D_1 、 D_2 、 D_3 与 D_4 。则 J 级变换所需行变换暂存为

$$n_{\text{row}} = \sum_{j=1}^L 8 \frac{2^L H}{2^{j-1}} = \sum_{j=2}^L 2^{L+4-j} H \tag{12}$$

在所设计的多级二维非折叠 DWT 结构中对离散小波系数的存储进行了改进,采用了乒乓存储的策略。这样在并行 EBCOT 处理速度不足的情况下,当前基本编码单元的系数存入不会对 EBCOT 读取上一单元的小波系数产生干扰,保证了 DWT 模块与 EBCOT 模块数据传输的稳定性与正确性,因此所需离散小波系数存储空间为

$$n_{\text{result}} = 2H^2(3(L-1)+4) = (6L+2)H^2 \tag{13}$$

以三级 9/7 DWT 模块为例,处理单元将会同时产生 10 个大小为 $H \times H$ 的 EBCOT 编码块,占据大小为 $20H^2$ 模块间存储空间。

小波系数组织输出后,位平面编码模块以编码块形式读取不同子带的小波系数值并处理为上下文信息对。由于上下文信息对的生成只由当前小波系数编码块决定,因此新型扫描结构不会影响后续算数编码的上下文形成和编码率-失真性能。

3 硬件分析与对比

3.1 硬件消耗

为合理估算硬件资源消耗,设定处理的图像大小为 $N \times N$,EBCOT 编码块大小为 $H \times H$,二维 9/7 DWT 模块共 L 级,数据并行度为 $S(S>1)$ 。

系统中列变换模块间缓存、行变换模块间缓存、级间结构缓存以及 DWT-EBCOT 模块间数据缓存为大规模数据存储,通过 RAM 读写实现。由式

(11)~(13)可知,所需 RAM 大小为

$$n_{\text{RAM}} = n_{\text{col}} + n_{\text{row}} + n_{\text{result}} = \sum_{j=1}^L \frac{N}{2^{j-3}} + \sum_{j=1}^L 2^{L+4-j} H + (6L+2)H^2 \tag{14}$$

基本处理模块内部缓存通过寄存器实现,根据表 1 及整体结构设计, L 级并行度为 S 的 9/7 DWT 系统寄存器数量为

$$n_{\text{reg}} = 45 \sum_{j=1}^L \frac{S}{2^{j-1}} = \sum_{j=1}^L \frac{45S}{2^{j-1}} \tag{15}$$

加法器与乘法器在列变换、行变换与缩放模块中被使用,与计算寄存器数量相同,系统所需加法器与乘法器数量为

$$n_{\text{add}} = 16 \sum_{j=1}^L \frac{S}{2^{j-1}} = \sum_{j=1}^L \frac{S}{2^{j-5}} \tag{16}$$

3.2 硬件效果对比

文献[11]也通过块扫描优化的方式来减少小波系数缓存。由于其未提供具体模块结构与运算时间等信息,因此通过对比存储资源消耗得出硬件效果的对比。在图像大小为 1024×1024 像素、EBCOT 码块为典型的 32×32 像素和 64×64 像素^[13]的情况下,两级小波变换存储资源消耗对比如表 1 所示。由表 1 可以看出,本文结构可节约存储资源消耗 40% 以上。

表 1 两种结构存储资源对比

码块大小	结构	模块内 存储量/kB	小波系数 存储量/kB	总体 存储量/kB
32×32	文献[11]	23.7	73.5	97.2
	本文	9.7	28.6	38.3
64×64	文献[11]	45.7	159	204.7
	本文	13.3	114.7	124.9

将本文结构与文献[7-9]的多级离散小波变换结构的硬件资源消耗作对比,结果如表 2 所示,其中所列结构均为输入数据并行度 $S=8$ 的三级变换结构,输入图像大小为 512×512 像素,EBCOT 码块大小为 64×64 ,存储资源消耗单位为 kB。表中处理时间为关键路径延时×处理周期数,单位为 ns。本结构的寄存器、加法器、乘法器数量以及资源消耗情况根据式(14)~(16)得出。通过表 2 的综合对照可以看出,与已有最优结构相比,虽然本文设计的结构运算资源略有增加,但总体存储资源消耗减少了 60% 以上,同时处理时间优于大部分结构,因此可以认为本文结构的硬件综合效果最优。并且随着图片尺寸的增大或 EBCOT 码块大小的减小,本文结构

表 2 三级 9/7 DWT 硬件资源消耗的综合对比

结构	加法器 数量	乘法器 数量	寄存器 数量	模块内 存储资源/kB	小波系数 存储资源/kB	总体 存储资源/kB	处理 时间/ns
文献[7]	168	94.5	157.5	11.4	524.3	535.7	209 879.04
文献[8]	294	189	443	5.7	524.3	530.0	111 247.36
文献[9]	180	111	341	3.1	524.3	527.4	160 563.20
本文	224	140	630	21.5	163.8	185.3	111 247.36

对 DWT 模块与 EBCOT 模块间的存储优化会进一步体现。

4 结束语

本文提出了一种新型多级二维离散小波变换的硬件结构,优化了数据扫描方式,调整了硬件架构与数据处理方式。各项对比结果表明,本文结构对比已有最优结构,可以节约 40% 以上的存储资源消耗,有着重要的实用价值。

参考文献:

[1] 邵婷婷,白宗文,周美丽. 基于离散小波变换的信号分解与重构 [J]. 计算机技术与发展, 2014, 24(11): 159-161.
SHAO Tingting, BAI Zongwen, ZHOU Meili. Decomposition and reconstruction of signal based on DWT [J]. Computer Technology and Development, 2014, 24(11): 159-161.

[2] DARJI A, AGRAWAL S, OZA A, et al. Dual-scan parallel flipping architecture for a lifting-based 2-D discrete wavelet transform [J]. IEEE Transactions on Circuits & Systems: II Express Briefs, 2014, 61(6): 433-437.

[3] WANG Jianxin, ZHU En. A high-throughput VLSI design for JPEG2000 9/7 discrete wavelet transform [J]. Journal of Southeast University (English Edition), 2015, 31(1): 19-24.

[4] HU Y, JONG C C. A Memory-efficient scalable architecture for lifting-based discrete wavelet transform [J]. IEEE Transactions on Circuits & Systems: II Express Briefs, 2013, 60(8): 502-506.

[5] 王雨,马军山,王华. 基于 FPGA 的 9/7 小波变换算法实现 [J]. 光学仪器, 2014, 36(5): 403-408.
WANG Yu, MA Junshan, WANG Hua. FPGA implementation of 9/7 wavelet transform algorithm [J]. Optical Instruments, 2014, 36(5): 403-408.

[6] 贾琦,梁煜,张为. 一种无片外存储的高性能二维 DWT 架构 [J]. 西安电子科技大学学报, 2017, 44

(4): 150-155.
JIA Qi, LIANG Yu, ZHANG Wei. Hardware efficient 2-D DWT architecture without Off-Chip RAM [J]. Journal of Xidian University, 2017, 44(4): 150-155.

[7] MOHANTY B K, MEHER P K. Memory efficient modular VLSI architecture for high throughput and low-latency implementation of multilevel lifting 2-D DWT [J]. IEEE Transactions on Signal Processing, 2011, 59(5): 2072-2084.

[8] MOHANTY B K, MEHER P K. Memory-efficient high-speed convolution-based generic structure for multilevel 2-D DWT [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2013, 23(2): 353-363.

[9] HU Y, JONG C C. A memory-efficient high-throughput architecture for lifting-based multi-level 2-D DWT [J]. IEEE Transactions on Signal Processing, 2013, 61(20): 4975-4987.

[10] WU B F, LIN C F. Memory-efficient architecture for JPEG 2000 coprocessor with large tile image [J]. IEEE Transactions on Circuits & Systems: II Express Briefs, 2006, 53(4): 304-308.

[11] YE L, HOU Z. Memory efficient multilevel discrete wavelet transform schemes for JPEG2000 [J]. IEEE Transactions on Circuits & Systems for Video Technology, 2015, 25(11): 1773-1785.

[12] ZHANG W, JIANG Z, GAO Z, et al. An efficient VLSI architecture for lifting-based discrete wavelet transform [J]. IEEE Transactions on Circuits & Systems: II Express Briefs, 2012, 59(3): 158-162.

[13] GAVVALA R, GOPAL M M, CHANDRA S S. Pass-parallel VLSI architecture of BPC for embedded block coder in JPEG2000 [C]// Asia Pacific Conference on Postgraduate Research in Microelectronics and Electronics. Washington, DC, USA: IEEE Computer Society, 2012: 111-117.

(编辑 刘杨)