

High-Speed and Low-Power Architectures for Forward and Inverse Discrete Wavelet Transform Using 4-Tap Daubechies Filters

Tze-Yun Sung¹ Chun-Wang Yu¹ Yaw-Shih Shieh¹ Hsi-Chin Hsin²

¹Department of Microelectronics Engineering, Chung Hua University, Hsinchu, Taiwan 300-12

²Department of Computer Science and Information Engineering, Formosa University, Hu-Wei, Taiwan 632-08

Abstract

This paper presents two architectures for 2-D discrete wavelet transform (DWT) and inverse DWT (IDWT). The first consists of transform module, RAM module and address sequencer. The transform module, which takes uniform and regular structure with local communication, offers many advantages, e.g. simple control flow, full hardware-utilization and low-power. The second architecture features parallel and pipelined computation for high throughput. Both architectures are suitable for VLSI implementations of the wavelet based image coders, e.g. JPEG-2000. The FPGA realization and VLSI implementation of the proposed 2-D DWT/IDWT architectures with 4-tap Daubechies filters shows significant improvements in terms of power saving and chip area.

Keywords: DWT/IDWT, low-power, image coder, JPEG-2000, 4-tap Daubechies filters, multiplierless.

1. Introduction

Discrete wavelet transform (DWT) is adopted in the JPEG-2000 standard [1]. However, 2-D DWT requires massive computations. For the real-time and/or on-line applications, parallel and pipelined architectures are to be taken into account to implement high-efficiency application specific integrated circuits (ASIC) or field programmable gate array (FPGA).

In [2], the wavelet with 4-tap Daubechies coefficients that is symmetric and almost orthogonal has been used for lossy image compression. As these coefficients are quantized for simplicity, multipliers can be replaced by simple shift registers and adders to improve the throughput.

In this paper, two high-efficiency architectures for the even and odd parts of 1-D decimated convolution involved in DWT are presented. The advantages of the proposed architectures are 100% hardware-utilization, multiplierless, regular structure, simple control flow and high scalability.

The remainder of the paper proceeds as follows. Section 2 presents the 2-D DWT algorithm with new derived mathematical formulas. In Section 3, the high-efficiency architecture of 2-D DWT is proposed. The high-efficiency and low-power architecture of 2-D IDWT is proposed in Section 4. The FPGA and VLSI implementations with performance evaluation are presented in Section 5. Conclusions and performance comparisons are given in Section 6.

2. 2-D DWT Algorithm

The 2-D DWT algorithm is as follows.

$$LL^j(m, n) = \sum_{i=0}^{K-1} \sum_{k=0}^{K-1} l(i) \cdot l(k) \cdot LL^{j-1}(2m-i, 2n-k) \quad (1)$$

$$LH^j(m, n) = \sum_{i=0}^{K-1} \sum_{k=0}^{K-1} l(i) \cdot h(k) \cdot LL^{j-1}(2m-i, 2n-k) \quad (2)$$

$$HL^j(m, n) = \sum_{i=0}^{K-1} \sum_{k=0}^{K-1} h(i) \cdot l(k) \cdot LL^{j-1}(2m-i, 2n-k) \quad (3)$$

$$HH^j(m, n) = \sum_{i=0}^{K-1} \sum_{k=0}^{K-1} h(i) \cdot h(k) \cdot LL^{j-1}(2m-i, 2n-k) \quad (4)$$

where $0 \leq n, m < N_j$, $LL^0(m, n)$ is the input image at the finest resolution 0, K denotes the length of filter, $l(i)$ and $h(k)$ are the low-pass and high-pass filters, respectively. $LL^j(m, n)$, $LH^j(m, n)$, $HL^j(m, n)$ and $HH^j(m, n)$ are the filter outputs representing the low-low, low-high, high-low and high-high subbands with size of $N_j \times N_j$ at resolution j , respectively. The decomposition equations (1) - (4) involve convolution followed by decimation by 2 in the horizontal and vertical directions.

Denote $LL_{m,i}^j(2n)$ with $l(i)l(2k)$, $l(i)h(2k)$, $h(i)l(2k)$ and $h(i)h(2k)$ the 1-D DWT for the even samples; $LL_{m,i}^j(2n+1)$ with $l(i)l(2k+1)$, $l(i)h(2k+1)$, $h(i)l(2k+1)$ and $h(i)h(2k+1)$ for the odd samples; $0 \leq n \leq N_j$ and $0 \leq k \leq K/2$. $LL_{m,i}^j(n)$, $LH_{m,i}^j(n)$, $HL_{m,i}^j(n)$, and $HH_{m,i}^j(n)$ can be written by

$$LL_{m,i}^j(n) = \sum_{k=0}^{\lceil K/2 \rceil - 1} l(i)l(2k) \cdot LL_{2m-i}^{j-1}(2n-2k) + \sum_{k=0}^{K-\lceil K/2 \rceil - 1} l(i)l(2k+1) \cdot LL_{2m-i}^{j-1}(2n-2k-1) \quad (5)$$

$$LH_{m,i}^j(n) = \sum_{k=0}^{\lceil K/2 \rceil - 1} l(i)h(2k) \cdot LL_{2m-i}^{j-1}(2n-2k) + \sum_{k=0}^{K-\lceil K/2 \rceil - 1} l(i)h(2k+1) \cdot LL_{2m-i}^{j-1}(2n-2k-1) \quad (6)$$

$$HL_{m,i}^j(n) = \sum_{k=0}^{\lceil K/2 \rceil - 1} h(i)l(2k) \cdot LL_{2m-i}^{j-1}(2n-2k) + \sum_{k=0}^{K-\lceil K/2 \rceil - 1} h(i)l(2k+1) \cdot LL_{2m-i}^{j-1}(2n-2k-1) \quad (7)$$

$$HH_{m,i}^j(n) = \sum_{k=0}^{\lceil K/2 \rceil - 1} h(i)h(2k) \cdot LL_{2m-i}^{j-1}(2n-2k) + \sum_{k=0}^{K-\lceil K/2 \rceil - 1} h(i)h(2k+1) \cdot LL_{2m-i}^{j-1}(2n-2k-1) \quad (8)$$

Thus, $LL_{m,i}^j(n)$, $LH_{m,i}^j(n)$, $HL_{m,i}^j(n)$ and $HH_{m,i}^j(n)$ can be computed by summing the two 1-D convolutions, which are performed independently on the even part $LL_{2m-i}^{j-1}(2n-2k)$ and odd part $LL_{2m-i}^{j-1}(2n-2k-1)$.

3. The Proposed 2-D DWT Architecture

The proposed architecture takes parallel and pipelined processing. At each decomposition level of 2-D DWT, there are two stages involved: stage 1 performs row filtering and stage 2 performs column filtering. At the first level, the size of the input image is $N \times N$; the size of the three output subbands: LH , HL and HH is $(N/2) \times (N/2)$. At the second level, the input is subband LL with size of $(N/2) \times (N/2)$; the size of the three output subbands: $LLLH$, $LLHL$ and $LLHH$ is $(N/4) \times (N/4)$. At the third level, the input is subband $LLLL$ with size of $(N/4) \times (N/4)$; the size of the four output subbands: $LLLLL$, $LLLLLH$, $LLLLHL$ and $LLLLHH$ is $(N/8) \times (N/8)$.

Figure 1 shows the proposed transform module for 2-D DWT. In which, 4-tap Daubechies wavelet with the low-pass and high-pass filters represented by coefficients $a(0) \sim a(3)$ and $b(0) \sim b(3)$, respectively, is used for demonstration. Based on equations (5) - (8), the splitter rearranges the even and odd parts of data by using the processing element (PE). f denotes the input frequency, $f/2$ the output frequency of L and H, and $f/4$ the output frequency of LL , LH , HL and HH . This single transform module consists of RAM of size $N/2 \times N/2$, multiplex, splitter and address sequencer. It takes 21 clock cycles to perform 2-D DWT. In which, clock cycles 0 to 15 are for the level-1 decomposition, clock cycles 16 to 19 for the level-2 decomposition, and clock cycle 20 for the level-3 decomposition. Three transform modules can be cascaded for parallel and pipelined processing.

4. The Proposed 2-D IDWT Architecture

In the proposed architecture of 3-level 2-D IDWT, two stages are performed at each reconstruction level, i.e. stage 1 performs column filtering and stage 2 performs row filtering. Figure 2 shows the transform module for 2-D IDWT. f denotes the input frequency and $4f$ the

output frequency. This single transform module consists of the inverse transform module, RAM module of size $N/2 \times N/2$, and multiplexer. At the first level, subband $(LL)^{j-2}LL$ is reconstructed from subbands $(LL)^{j-1}LL$ (that is obtained from the multiplexer), $(LL)^{j-1}LH$, $(LL)^{j-1}HL$, and $(LL)^{j-1}HH$ by using the inverse transform module. The intermediate output $(LL)^{j-2}LL$ is stored in the RAM module. In a similarly way, the original image can be reconstructed in 21 clock cycles. Clock cycle 0 is for the first level reconstruction; Clock cycles 1 to 4 for the second level reconstruction; Clock cycles 5 to 20 for the final reconstruction.

5. Hardware Implementation and Performance Analysis of The Proposed 2-D DWT/IDWT Architectures

In the proposed architectures, the multiplier can be replaced by simple shifters and adders, in which the filter coefficients are approximated and represented by using booth binary format. The multiplier has been replaced by a carry-save-adder (CSA), an adder and three hardware shifters in the processing element (PE).

The proposed architecture of 64×64 DWT has been synthesized by Xilinx FPGA Express tools, written in Verilog[®] [3], and emulated on the Xilinx XC2V4000 FPGA platform [4]-[5]. The chip has been synthesized by TSMC 0.18 μm 1P6M CMOS cell libraries. The core size and power consumption can be obtained from the reports of Synopsys[®] design analyzer and PrimPower[®] [6], respectively. The reported core size is $588 \times 588 \mu m^2$, and the power dissipation is 28.5 mW at 1.8V with clock rate of 50MHz. Figure 3 shows the layout view of the implemented 64×64 DWT processor. Due to the limitation of paper size, the reports and layout view of the proposed 2-D IDWT processor are not presented here. The proposed 2-D DWT/IDWT processors have been applied to many images with great satisfactions.

The decimation filter for 1-D DWT takes eight adders, thirteen shifters and three registers per PE, whereas it takes seven adders, thirty shifters and three registers per PE for 1-D IDWT. Both hardwares are very cost-effective requiring only three and two additions to perform each PE computation in DWT and IDWT, respectively. The proposed architectures can reduce the power dissipation by m in comparison with the conventional architectures in m -bit operand (low-power utilization). They have regular structure, local communication, simple control flow, and therefore are very suitable for VLSI implementation with scalable filter length. The hardware utilization is 100% in the transform modules involved. Thus, the s

systems consume ultra-low power. The total processing time of 2-D DWT can be calculated as follows.

$$\sum_{i=0}^{j-1} 2^{-(2i+2)} \cdot (N \times N) = (2^{-2} + 2^{-4} + \dots + 2^{-2j}) \cdot (N \times N) \quad (9)$$

$$= \frac{1}{3} (1 - 2^{-2j}) \cdot (N \times N)$$

where $j = \log_2 N$.

For the single transform module of 2-D IDWT, the total processing time is the same.

By the proposed architectures with fixed point operation, the peak-signal-to-noise ratio (PSNR) of the reconstructed image Lena is 255.5dB. The 3-level decomposition subband image, the original image and the reconstructed image are shown in Figures 4, 5(a) and 5(b), respectively.

6. Discussions and Conclusions

In this paper, two high-speed and ultra low-power architectures with processing time $(1 - 2^{-2j}) \cdot N^2 / 3$ for 2-D DWT and IDWT are proposed. They are much faster than the conventional architectures proposed by Wu and Chen [7]-[8], and Marino [9]. Table 1 shows the comparisons in terms of the system performance defined by AT^2 [8]-[12], where A is the area and T is the time or latency in clock cycles. It demonstrates that the proposed architectures are superior to the previous works. They are cost-effective, high-speed, and can reduce the power dissipation by m in comparison with the conventional architectures in m -bit operand (low-power utilization).

The proposed architectures have been verified by Verilog-HDL and implemented on FPGA. The advantages of the proposed architectures are 100% hardware utilization and ultra low-power. The architectures have regular structure, simple control flow, high throughput and high scalability. Thus, they are very suitable for the JPEG-2000 standard.

References

- [1] ITU-T Recommendation T.800. JPEG2000 image coding system – Part 1, ITU Std., July 2002. <http://www.itu.int/ITU-T/>.
- [2] A. S. Lewis and G. Knowles, "VLSI architecture for 2-D Daubechics wavelet transform without multipliers," *Electron. Lett.*, vol. 27, pp.171–173, Jan. 1991.
- [3] D. E. Thomas, P. H. Moorby, *The Verilog Hardware Description Language, Fifth Edition*, Kluwer Academic Pub. 2002.
- [4] Model ModelSim Products: <http://www.model.com/products>.
- [5] Xilinx FPGA products, <http://www.xilinx.com/products>.
- [6] Synopsys FPGA Express, <http://www.synopsys.com/products>.
- [7] P. –C. Wu, L. –G. Chen, "An Efficient Architecture for Two-Dimensional Discrete Wavelet Transform," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 11, No. 4, pp. 536-545, April 2001.
- [8] P. –C. Wu, C. –T. Liu, L. –G. Chen, "An Efficient Architecture for Two-Dimensional Inverse Discrete Wavelet Transform," *IEEE International Symposium on Circuits and Systems*, Vol. 2, pp. II-312-II-315, May 2002.
- [9] F. Marino, "Two Fast Architectures for the Direct 2-D Discrete Wavelet Transform," *IEEE Transactions on Signal Processing*, Vol. 49, No. 6, pp. 1248-1259, June 2001.
- [10] S. Y. Kung, *VLSI Array Processors*, Prentice-Hall, New Jersey, USA, 1989.
- [11] T. Y. Sung, C. S. Chen, "A Parallel-Pipelined Processor for Fast Fourier Transform," *The Fourth IEEE Asia-Pacific Conference on Advanced System Integrated Circuits (AP-ASIC-2004)*, Fukuoka, Japan, pp.194-197(10-1), August 3-5, 2004.
- [12] T. Y. Sung, "A Memory-Efficient and High-Speed Split-Radix FFT/IFFT Processor Based on Pipelined CORDIC Rotations," to appear in *IEE Proceedings – Vision, Image and Signal Processing*, 2006.

Table 1. Comparison with the previous works

Single Transform Module	This work	Wu & Chen [7] [8]	Marino[9]
Algorithm	Convolution	Convolution	Convolution
Latency (clock-cycle)(DWT)	$(1 - 2^{-2j}) \cdot N^2 / 3$	$2 \cdot (1 - 2^{-2j}) \cdot N^2 / 3$	$2 \cdot N^2 / 3$
Latency (clock-cycle)(IDWT)	$(1 - 2^{-2j}) \cdot N^2 / 3$	$2 \cdot (1 - 2^{-2j}) \cdot N^2 / 3$	---
Multiplierless	Yes	No.	No.
DWT/IDWT	DWT/IDWT	DWT/IDWT	DWT
Hardware Utilization	100%	<100%	100%
Low-power	Better	Good	Good
High-speed	Better	Poor	Good
AT^2 (System Performance)	AT^2 (DWT) AT^2 (IDWT)	$2.667 AT^2$ (DWT) $2.667 AT^2$ (IDWT)	$> AT^2$ (DWT) ---

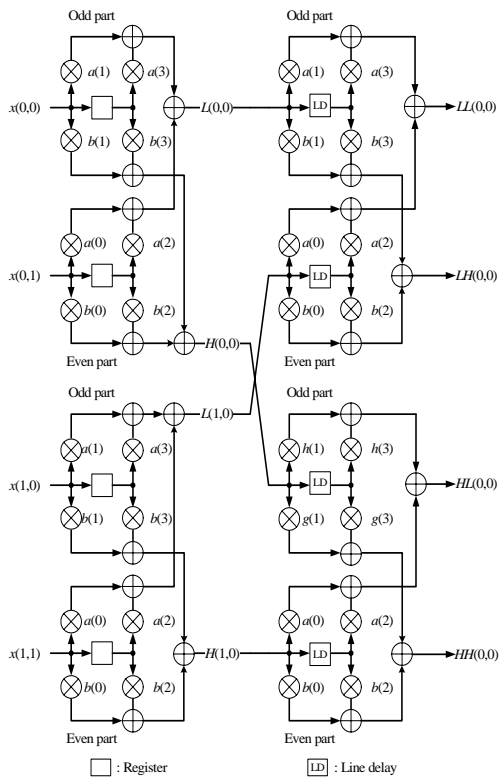


Fig. 1. The transform module for 2-D DWT

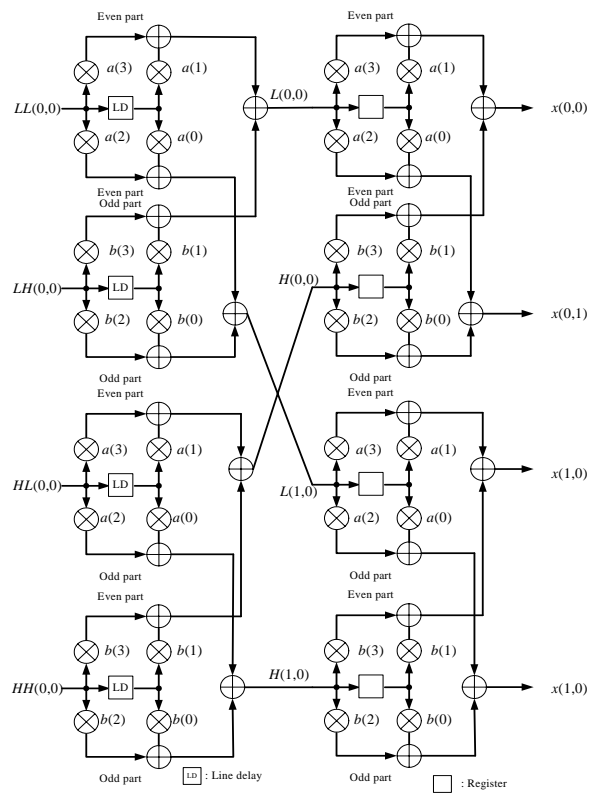


Fig. 2. The transform module for 2-D IDWT

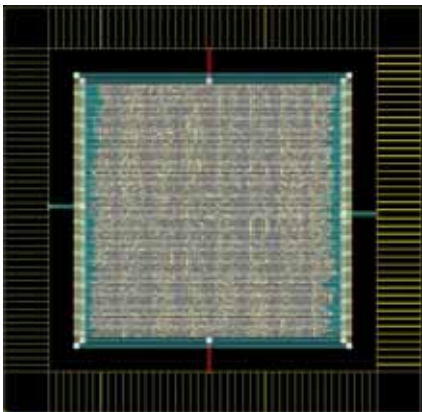


Fig. 3. The layout view of the 64x64 2-D DWT processor

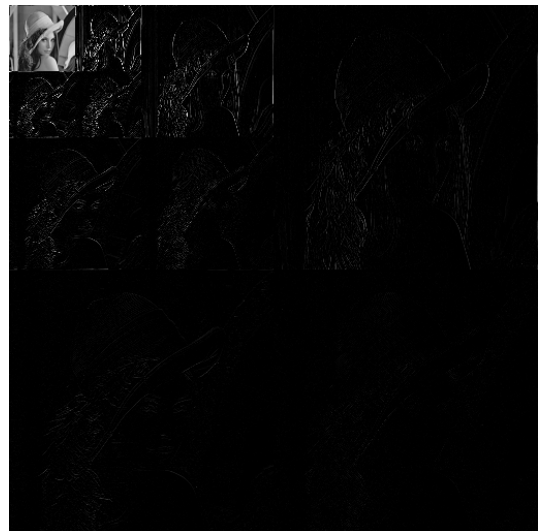


Fig. 4. The 3-level compressed image



(a)

(b)

Fig. 5. (a) Original image (b) Reconstructed image