Low-Power and High-Performance 2-D DWT and IDWT Architectures Based on 4-tap Daubechies Filters

TZE-YUN SUNG
Department of Microelectronics Engineering
Chung Hua University
707, Sec. 2, Wufu Road
Hsinchu, 30012, TAIWAN

Abstract: - This paper proposes two architectures of 2-D discrete wavelet transform (DWT) and inverse DWT (IDWT). The first high-efficiency architecture comprises a transform module, an address sequencer, and a RAM module. The transform module has uniform and regular structure, simple control flow, and local communication. The significant advantages of the single transform module are full hardware-utilization and low-power. The second architecture features parallel and pipelined computation and high throughput. Both architectures are very suitable for VLSI implementation of new-generation image coding/decoding systems, such as JPEG-2000 and motion-JPEG. In the realization of 2-D DWT/IDWT, we focus on a VLSI implementation using 4-tap Daubechies filters, which saves power and reduces chip area.

Key-Words: - DWT/IDWT, Low-power, Image coding/decoding system, JPEG-2000, VLSI, 4-tap Daubechies filters, Multiplierless.

1 Introduction

In the field of digital image processing, the JPEG-2000 standard uses the scalar wavelet transform for image compression [1]; hence, the two-dimensional (2-D) discrete wavelet transform (DWT)/inverse DWT (IDWT) has recently been used as a powerful tool for image coding/decoding systems. Two-dimensional DWT/IDWT demands massive computations, hence, it requires a parallel and pipelined architecture to perform real-time or on-line video and image coding and decoding, and to implement high-efficiency application-specific integrated circuits (ASIC) or field programmable gate array (FPGA). At the heart of the analysis stage of the system is the DWT. In the synthesis stage, the inverse DWT recovers the original image from the coefficients of DWT.

Cohen, Daubechies and Feauveau proposed using 4-tap Daubechies coefficients for lossy analysis [1]. The symmetry of 4-tap Daubechies filters and the fact that they are almost orthogonal [2]–[4] make them good candidates for image analysis application. The coefficients of the filter are quantized before hardware implementation; hence, the multiplier can be replaced by limited quantity of shift registers and adders. Thus, the system hardware is saved, and the system throughput is improved significantly.

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

The remainder of the paper is organized as follows. Section 2 presents the 2-D discrete wavelet transform algorithm, and derives new mathematical formulas. In Section 3, the high-efficiency architecture for the 2-D DWT is proposed. In Section 4, the high-efficiency and low-power architecture for the 2-D IDWT is proposed. Section 5 applies the two proposed 2-D DWT/IDWT architectures to the coefficient quantization scheme for VLSI implementation, and analyzes their performance. Finally, comparison of performance between the proposed architectures and previous works is made with conclusions given in Section 6.

2 2-D Discrete Wavelet Transform Algorithm

The 2-D DWT is a multilevel decomposition technique. The mathematical formulas of 2-D DWT are defined as follows [5]:

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

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

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

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

where $0\le n,\ m<N_j$, $LL^0(m,n)$ is the input image, K denotes the length of filter, $h(i)$ denote the impulse responses of the low-pass filter, and $g(k)$ denote the impulse responses of the high-pass filter, which is developed from $(K\times K)$ -tap filters, and $LL^j(m,n)$, $LH^j(m,n)$, $HL^j(m,n)$, and $HH^j(m,n)$ denote respectively the coefficients of low-low, low-high, high-low and high-high subbands produced at the decomposition level j (also represented by LL^j , LH^j , HL^j , and HH^j). $N_j\times N_j$ denotes samples of LL^j .

According to the mathematical formulas (1), (2), (3) and (4), the decomposition is produced by four 2-D convolutions followed by the decimation both in the row and in the column dimension for each level. In the three-level analysis for 2-D DWT, the data set LL^{j-1} having $N_{j-1}\times N_{j-1}$ samples is decomposed into four subbands $LL^j, LH^j, HL^j,$ and HH^j each having $N_j\times N_j$ (equals to $(N_{j-1}/2)\times(N_{j-1}/2)$) samples.

Let, $h(i)h(2k)$, $h(i)g(2k)$, $g(i)h(2k)$ and $g(i)g(2k)$ be 1-D DWT consisting of the even-numbered samples, and $0\le n\le N_j$; $0\le k\le K/2$. Moreover,

let $LL_m^j(2n+1)$, $h(i)h(2k+1)$, $h(i)g(2k+1)$, $g(i)h(2k+1)$ and $g(i)g(2k+1)$ be 1-D DWT consisting of the odd-numbered samples, and $0\le n\le N_j$; $0\le k\le K/2$. $LL_{m,i}^j(n)$, $LH_{m,i}^j(n)$, $HL_{m,i}^j(n)$, and $HH_{m,i}^j(n)$ can be expressed as follows:

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

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

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

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

The above equations imply that $LL_{m,i}^j(n)$, $LH_{m,i}^j(n)$, $HL_{m,i}^j(n)$ and $HH_{m,i}^j(n)$ can be computed as the sum of two 1-D convolutions performed independently on the even part $LL_{2m-i}^{j-1}(2n-2k)$ and the odd part $LL_{2m-i}^{j-1}(2n-2k-1)$ .Please, leave two blank lines between successive sections as here.

3 The Proposed 2-D DWT Architecture

The proposed architecture performs parallel and pipelined processing. Each analysis level involves two stages: stage 1 performs row filtering, and stage 2 performs column filtering. In a one-level filter bank for 2-D DWT computation. At the first level, the size of the input image is $N\times N$, and the size of the output of each of the three subbands *LH*, *HL* and *HH* is $(N/2)\times(N/2)$. At the second level, the input is the *LL* subband whose size is $(N/2)\times(N/2)$, and the size of the output of each of the three subbands *LLLH*, *LLHL* and *LLHH* is $(N/4)\times(N/4)$. At the third level, the input is the *LLLL* subband whose size is $(N/4)\times(N/4)$, and the size of the output of each of the four subbands *LLLLLL*, *LLLLLH*, *LLLLHL* and *LLLLHH* is $(N/8)\times(N/8)$.

The coefficients of the low-pass filter and the high-pass filter have been derived in the biorthogonal 9/7 wavelet [6]-[7]. The coefficients are quantized before hardware implementation. We assume that the low-pass filter has four tapes: $h(0)$, $h(1)$, $h(2)$ and $h(3)$, and the high-pass filter also has four tapes: $g(0)$, $g(1)$, $g(2)$ and $g(3)$. The horizontal filter and the vertical filter for 1-D DWT are shown in Figures 1 and 2, respectively. According to Equations (5), (6), (7) and (8), each 1-D decimated convolution can be computed as the point-wise sum of two 1-D convolutions performed independently. The proposed architecture of the transform module for 2-D DWT composed of four

vertical filters and horizontal filters is illustrated in Figure 3. The proposed architecture of the single transform module for 2-D DWT comprises a ($N/2 \times N/2$) memory module, a transform module, a multiplex and an address generator is shown in Figure 4. It requires 42 clock cycles to perform the 2-D DWT transform in 8×8 image. Clock cycles 0 to 31 perform the level-1 analysis, clock cycles 32 to 40 perform the level-2 analysis, and clock cycles 40 to 41 perform the level-3 analysis.

4 The Proposed 2-D IDWT Architecture

In the proposed architecture of three-level synthesis for 2-D DWT, each synthesis level involves two stages: stage 1 performs column filtering, and stage 2 performs row filtering. The vertical filter and the horizontal filter for 1-D IDWT are shown in Figures 5 and 6, respectively. Figure 7 illustrates the transform module for 2-D IDWT. In Figure 7, the low-pass filter has four tapes: $\tilde{h}(0)$, $\tilde{h}(1)$, $\tilde{h}(2)$ and $\tilde{h}(3)$, and the high-pass filter also has four tapes: $\tilde{g}(0)$, $\tilde{g}(1)$, $\tilde{g}(2)$ and $\tilde{g}(3)$. The single transform module comprises an inverse transform module, a RAM module $(N/2 \times N/2)$ and a multiplexer. At the first-level composition, the inverse transform module comprises the $(LL)^{j-2}LL$ subband selected from the $(LL)^{j-1}LL$ subband by the multiplexer, as well as the $(LL)^{j-1}LH$, $(LL)^{j-1}HL$, and $(LL)^{j-1}HH$ subbands obtained by the entropy decoder. The $(LL)^{j-2}LL$ subband is stored in the RAM module. At the second-level composition, the inverse transform module comprises the $(LL)^{j-3}LL$ subband selected from $(LL)^{j-2}LL$ subband by the multiplexer of the RAM module, as well as the $(LL)^{j-2}LH$, $(LL)^{j-2}HL$, and $(LL)^{j-2}HH$ subbands obtained by the entropy decoder. At the last-level composition, the inverse transform module comprises the original image selected from the LL subband by the multiplexer, as well as the LH , HL , and HH subbands obtained by the entropy decoder. It requires 22 clock cycles to perform the 2-D IDWT transform. Clock cycles 0 to 1 perform the level-1 synthesis, clock cycles 2 to 9 perform the level-2 synthesis, and clock cycles 10 to 41 perform the level-3 synthesis.

5 Hardware Implementation and Performance Analyses of the Proposed 2-D DWT/IDWT Architectures

Filter coefficients of 4-tap Daubechies low-pass filter are quantized before implementation in the high-speed computation hardware [8]-[9]. In the proposed architectures, all multiplications are performed using shifts and additions after approximating the coefficients as a booth binary recoded format. The multiplier is replaced by a carry-save-adder (CSA), an adder and three hardwire shifters in processing element (PE). The replaced multiplier reduces power dissipation by *m* compared with conventional architectures in *m*-bit operand (low-power utilization).

Two proposed DWT and IDWT architectures have regular structure, local communication and simple control flow, so they are very suitable for VLSI implementation and scalable filter length. In both of the single transform modules, the hardware utilization are 100%, so the systems consumes ultra-low power. The total data processing times of 2-D DWT and IDWT are $(2/3)\cdot(1-2^{-2j})\cdot(N \times N)$, where $j = \log_2 N$.

Two architectures with fixed point operation applying on 3-level image compression and decompression by the proposed periodic extension method [10], which is shown in Figure 8. The peak-signal-to-noise ratio (PSNR) of the reconstructed image is 255.5dB. The original image and reconstructed image are shown in Figure 9. The log-log plot of the 2-D DWT/IDWT computations against different image width for each algorithm is shown in Figure 10 and we find that the overhead of handling borders is lower, when the image width is increased. Hence, the performances of the proposed 2-D DWT/IDWT architectures are suited for JPEG-2000 and motion-JPEG.

The proposed 2-D DWT processor has been synthesized by TSMC 0.18 μm 1P6M CMOS cell libraries. The core size is $588 \times 588 \mu m^2$ and the power dissipations is 28.53 mW at 1.8V with clock rate of 50MHz. Figures 11shows the layout view of the implemented 2-D DWT core processor.

6 Conclusion

In this paper, two high-speed and ultra low-power architectures for 2-D DWT and IDWT with single transform modules have been proposed. Both architectures perform analysis in $2\cdot(1-2^{-2j})\cdot N^2/3$

processing time and synthesis in $2\cdot(1-2^{-2j})\cdot N^2/3$. They are significantly faster than other commonly used architectures proposed by Chakrabati, Visshwanath, Wu and Chakrabati [11]-[14]. Tables 1 and 2 depict the comparisons of the proposed architecture to other commonly used architectures for 2-D DWT and IDWT [11]-[14]. As can be seen, the system performances of the two proposed architectures are significantly better than that of previous works.

Filter coefficients are quantized before implementation using 4-tap Daubechies filters. Both hardwares are cost-effective and the systems have high-speed. The architectures reduce power dissipation by *m* compared with conventional architectures in *m*-bit operand (low-power utilization).

The proposed architectures have been verified by Verilog-HDL [15] and implemented on VLSI [16]. 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 new-generation image coding/decoding systems, such as JPEG-2000 and motion-JPEG.

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, Jan. 1991, pp.171–173.

[3] I. Daubechies, "Orthonormal Based of Compactly Supported Wavelets," *Communications on Pure and Applied Mathematics*, Vol. 41, 1988, pp. 909-996.

[4] I. Daubechies, Ten Lectures on Wavelets, 1st edition, Philadelphia PA: *SIAM*, 1992.

[5] F. Marino, Two Fast Architectures for the Direct 2-D Discrete Wavelet Transform, *IEEE Transactions on Signal Processing*, Vol. 49, No. 6, June 2001, pp. 1248-1259.

[6] M. Antonini, M. Barlaud, P. Mathieu, I. Daubechies, Image Coding Using Wavelet Transform, *IEEE Transactions on Image Processing*, Vol.1, No.2, April 1992, pp. 205-220.

[7] K. A. Kotteri, A. E. Bell, J. E. Carletta, Design of Multiplierless, High- Peformance, Wavelet Filter Banks with Image Compression Applications, *IEEE Transactions on Circuits and Systems-I*, Vol. 51, No. 3, March 2004, pp.483-494.

[8] T.-Y. Sung, Y.-S. Shieh, C.-L Chiu, C.-W. Yu, VLSI Architectures for 2-D Forward and Inverse Discrete Wavelet Transform Using 4-tap Daubechies Filters," *2006 Conference on Electronic Communication and Application (2006-CECA)*, Kaoshiung, Taiwan, July 2006, BO-019.

[9] T.-Y. Sung, C.-L. Chiu, Y.-S. Shieh, C.-W. Yu, Low-Power Multiplierless 2-D DWT and IDWT Architectures Using 4-tap Daubechies Filters, *The 7th International Conference on Parallel and Distributed Computing, Applications and Technologies*, Taipei, Taiwan, R.O.C., December 2006, pp.185-190.

[10] M. Ferretti, D. Rizzo, Handling Borders in Systolic Architectures for the 1-D Discrete Wavelet Transform for Perfect Reconstruction, *IEEE Trans. Signal Processing*, Vol. 48, No. 5, May 2000, pp. 1365-1378.

[11] C. Chakrabarti, M. Vishwanath, "Efficient Realizations of the Discrete and Continuous Wavelet Transforms: From Single Chip Implementations to Mappings on SIMD Array Computers,' *IEEE Transactions on Signal Processing*, Vol. 43, No. 5, 1995, pp. 759-771.

[12] M. Vishwanath, R. M. Owens, M. J. Irwin, "VLSI Architectures for the Discrete Wavelet Transform," *IEEE Transactions on Circuits and Systems-II*, Vol. 42, No. 5, May 1995, pp.305-316.

[13] 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, May 2002, pp. II-312-II-315.

[14] C. Chakrabarti, C. Mumford, "Efficient Realizations of Analysis and Synthesis Filters Based on the 2-D Discrete Wavelet Transform," *International Conference on Acoustic, Speech and Signal Processing*, May 1996, pp. 3256-3259.

[15] D. E. Thomas, P. H. Moorby, *The Verilog Hardware Description Language*, Fifth Edition, Kluwer Academic Pub. 2002.

[16] *TSMC 0.18 μm CMOS Design Libraries and Technical Data*, v.3.2, Taiwan Semiconductor Manufacturing Company, Hsinchu, Taiwan.

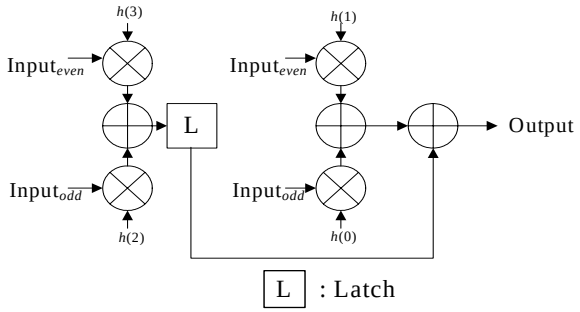


Fig.1. The horizontal low-pass filter for 1-D DWT

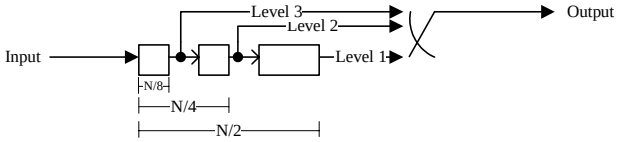


Fig.2(a). Line delay (LD) for Vertical Filter in 1-D DWT

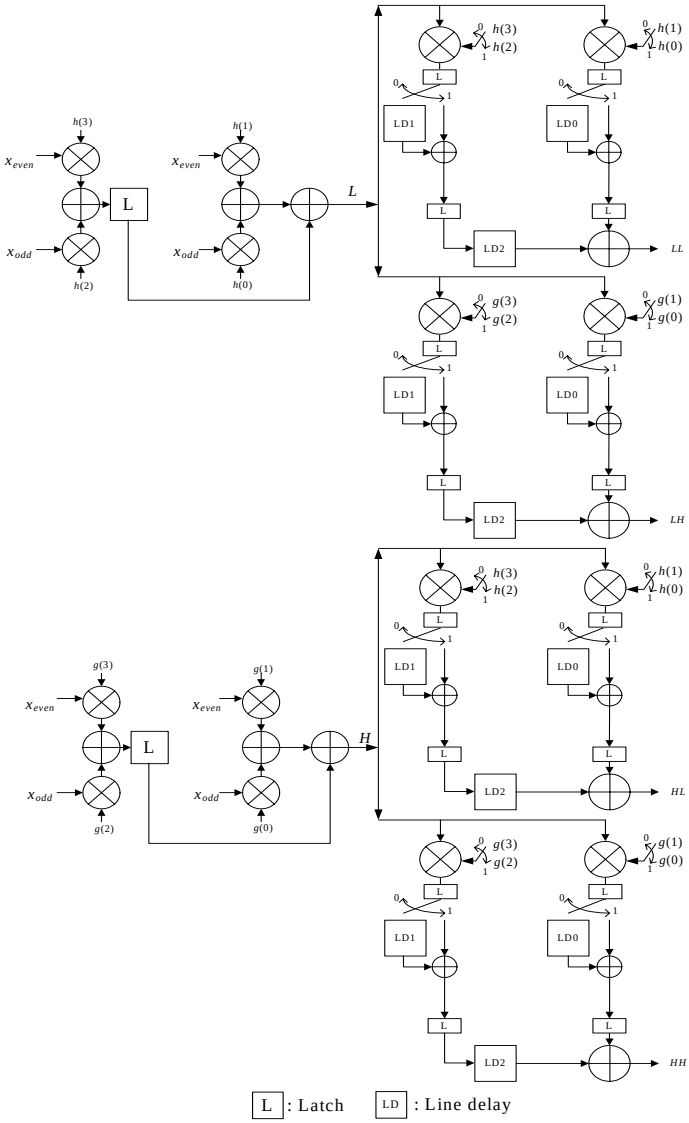


Fig.3. The proposed architecture of the transform module for 2-D DWT

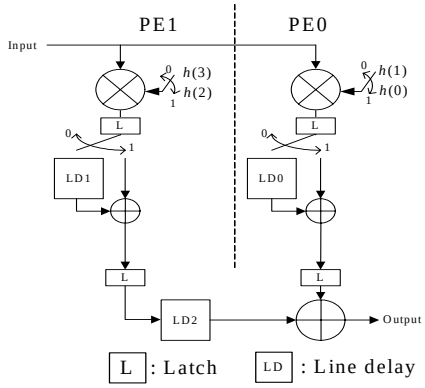


Fig.2(b). The vertical low-pass Filter for 1-D DWT

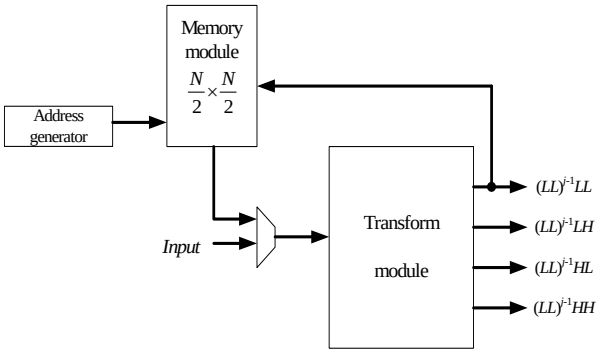


Fig.4. The proposed architecture of the single transform module for 2-D DWT

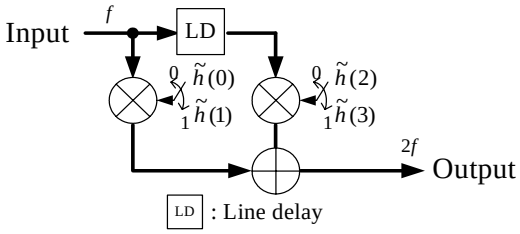


Fig.5. The vertical low-pass filter for 1-D IDWT

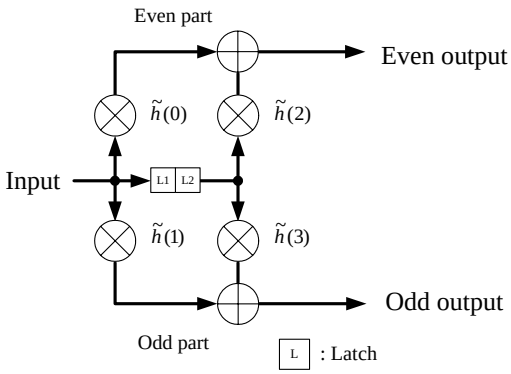


Fig.6. The horizontal low-pass filter for 1-D IDWT

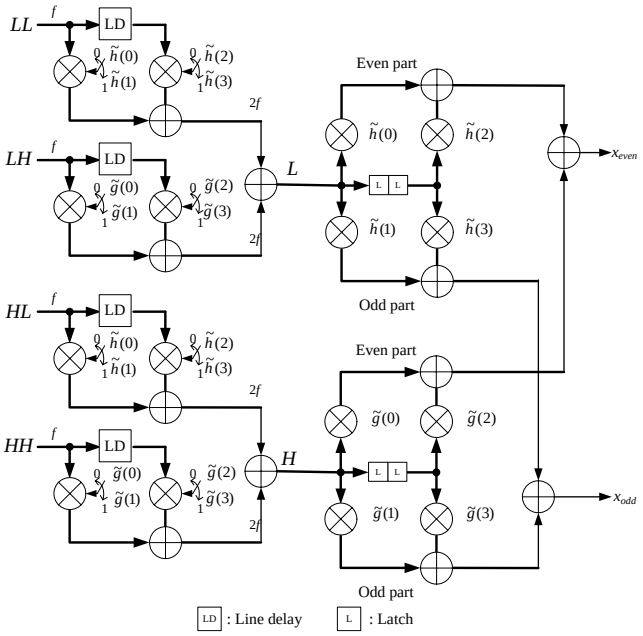


Fig.7. The proposed architecture of the transform module for 2-D IDWT

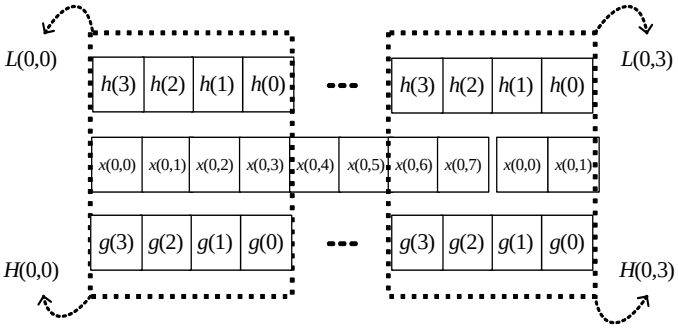
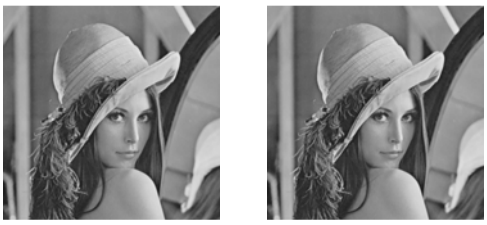


Fig.8. The proposed periodic extension method



(a) (b)

Figure 9. (a) Original image and (b) Reconstructed image

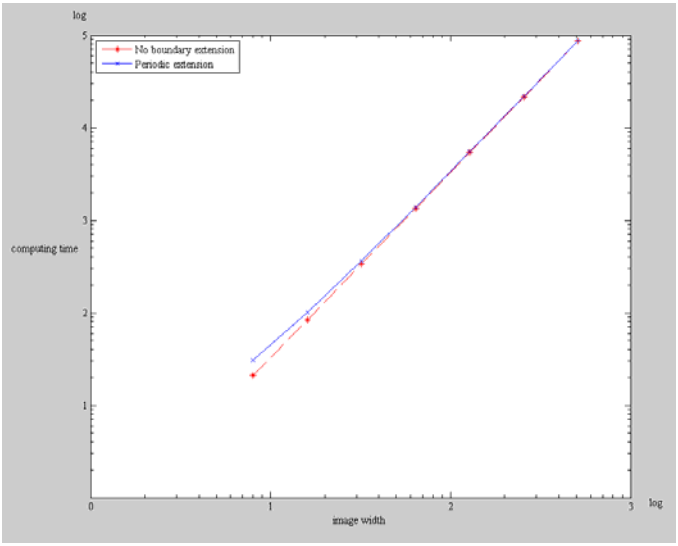


Fig.10. Log-log plot of the 2-D DWT/IDWT computations against different image width for each algorithm

Table 1 The comparison of the proposed architecture to other commonly used architectures for 2-D DWT

Architecture	Multiplier	Adder	Storage	Computing time	Control complexity
This work	--	$8K$	$\frac{N^2}{4}+N(K+2)$	$\approx\frac{2N^2}{3}$	Simple
Polyphase-Folded[13]	$4K$	$4K$	$\frac{N^2}{4}+KN+K$	$\approx\frac{2N^2}{3}$	Medium
Direct Approach[12]	K	K	N^2	$4N^2$	Complex
Systolic-Parallel[12]	$4K$	$4K$	$2KN+4N$	N^2+N	Complex
Non-Separable[11]	$2K^2$	$2K^2$-2	$2KN$	N^2+N	Complex
SIMD[11]	$2N^2$	$2N^2$	N^2	K^2+J	Complex
Parallel-Parallel[14]	$4K$	$4K$-4	$2KN+N$	N^2+N	Complex

K : filter length N^2 : image size J : level of decomposition

Table 2 The comparison of the proposed architecture to other commonly used architectures for 2-D IDWT

Architecture	Multiplier	Adder	Storage	Computing time	Control complexity
This work	--	$8K$	$\frac{N^2}{4}+N(K+2)$	$\approx\frac{2N^2}{3}$	Simple
Polyphase-Folded[13]	$4K$	$4K$	$\frac{N^2}{4}+(K-2)N+K$	$\approx\frac{2N^2}{3}$	Medium
Parallel-Parallel[14]	$4K$	$4K$-4	$2KN+N$	N^2+N	Complex

K : filter length N^2 : image size

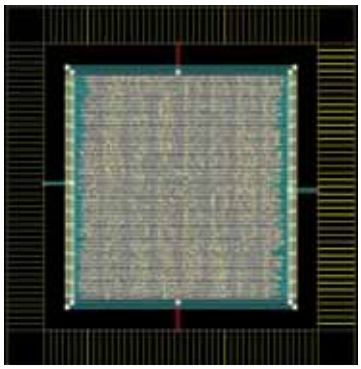


Fig.11. The layout view of the implemented 2-D DWT core processor