Article

An Internal Folded Hardware-Efficient Architecture for Lifting-Based Multi-Level 2-D 9/7 DWT

Wei Zhang [1], Changkun Wu [1], Pan Zhang [1] and Yanyan Liu [2],*

[1]   School of Microelectronics, Tianjin University, Tianjin 300072, China; tjuzhangwei@tju.edu.cn (W.Z.); tjuwuck@tju.edu.cn (C.W.); zhangpantju@tju.edu.cn (P.Z.)

[2]   Tianjin Key Laboratory of Photo-electronic Thin Film Devices and Technology, Nankai University, Tianjin 300071, China

*   Correspondence: lyytianjin@nankai.edu.cn

check for updates

Featured Application: The architecture proposed in this paper is intended for use in designs that require hardware acceleration and optimization.

Abstract: In this paper, a novel internal folded hardware-efficient architecture of multi-level 2-D 9/7 discrete wavelet transform (DWT) is proposed. For multi-level DWT, the unfolded structure is more extensively used compared with the folded structure, because of its low memory consumption and low time delay. However, a set of input data valid every few clock cycles caused the mismatch between clock and data in the unfolded structure. The mismatch usually needs to be solved by multi-clock or complex data adjustment, which increases the consumption of hardware resources and the complexity of the overall system. To solve the above problem of the unfolded structure, we adjust the data input timing by using a single clock domain and folding the DWT architecture of different levels in varying degrees, according to their own clock-to-data ratios. For an image of size of N × N pixels and 3-level DWT, the proposed architecture requires only 6N words temporal memory. For 3-level DWT with an image of size 512 × 512 pixels, the hardware estimation and comparison of the existing architectures show that, the hardware estimation result shows at least 30.6% area-delay-product (ADP) decrease, and at least 22.4% transistor-delay-product (TDP) decrease for S = 8, and 25.77% transistor-delay-product (TDP) decrease for S = 16.

---

1. Introduction

The discrete wavelet transform (DWT), as a multi-resolution analysis tool, is commonly used for image analysis, image compression and digital signal processing. To reach a high compression ratio, the input image is usually required to be decomposed by multi-level DWT. For hardware implementation of multi-level 2-D DWT, how to reduce the hardware consumption and improve the system performance becomes the hot research topic of structural optimization.

The existing DWT architectures can be classified into the lifting scheme and convolution-based scheme. Compared with the convolution-based scheme, the lifting scheme has been popularly researched because of lower computation complexity and less memory requirement. In order to optimize the lifting scheme, Zhang [1] proposed a two-input/two-output pipelined architecture by rearranging with the expressions to reduce the critical path delay (CPD) to Tm, where Tm is the delay of a multiplier, and limit the size of the temporal memory to $4N$. A dual-scan parallel flipping architecture of high hardware utilization efficiency was presented in [2], which can be folded to minimize the CPD

to *Tm*. [3] introduced the overlapping scanning method and decreased the temporal buffer to 3*N* by overlapping one pixel. Also, the overlapping scanning method is seen in [4–7]. By overlapping one pixel, three pixels and five pixels, it can economize the temporal memory of *N*, 2*N* and 3*N*, respectively. Nevertheless, the consumption of external memory will largely increase.

For multi-level DWT, which is usually used in the main application environment. Large demand of the frame memory that stores the low-low (LL) sub-band result for the folded architecture or complex data adjustment for the unfolded architecture is generally needed. Tian [8] proposed a block based high-throughput structure of 1-level 2-D DWT. Folded architecture based on this method needs $5N^2/16$ frame buffer words. Thus, to reduce the consumption of hardware resources, the unfolded multi-level 2-D DWT structure is more widely researched. Mohanty and Meher [9] presented a line-based parallel lifting structure without line and frame buffers. However, its CPD is *Tm* + 2*Ta*, where *Ta* is the delay of an adder. Later, [10] proposed a parallel convolution architecture, using more computing resources to reduce the memory demand for multi-level 2-D DWT. Consequently, high-throughput and less memory requirement is achieved at the expense of more area. Then, a block-based architecture was proposed to [11], which achieves lower demand for external memory access and higher energy efficiency. Furthermore, [12] presented a scalable parallel architecture of multi-level 2-D DWT based on lifting scheme. Temporal memory of this architecture is reduced to zero in the first level, by overlapping seven pixels. A creative processing method of different levels is used to decrease the temporal RAM to 3*N*. However, the control logic of the architecture is complicated. Besides, [13] discussed different data scanning methods and optimized the scanning sequence to decrease the area of the frame memory for the unfolded structure. Recently, Wu [14] introduced the CSD multiplier to decline the critical path to Ta. Nonetheless, multi-clock control system is necessary. [15] used an innovative block based Z types memory scanning method of their own way for reducing the total processing time, but it's not a multi-level architecture. The authors of [16] proposed bit-serial Distributed Arithmetic (DA) based VLSI architectures of 1-D/2-D DWT, which makes the designs Multiplierless and consumes less area, but the article does not mention the architecture's throughput rate. In [17], the authors proposed a look-up-table (LUT) based structure of high-throughput implementation of multilevel lifting DWT. The proposed structure can process one block of samples to achieve high-throughput rate. However, it requires 5210 more words and 21,504 words for block size 16 and 64 respectively, and the critical path of the proposed structure involves 3*Ta* delay for block size 16 and 64.

From the researches of the existing 2-D DWT architectures, it can be observed that, compared with the folded structure, the unfolded structure has smaller critical path delay and lower requirement for external memory accesses. However, only LL sub-band produced by the previous level needs to be further decomposed, which results in the mismatch problem between clock and data for the next level. For multi-level 2-D DWT architecture, the mismatch is usually solved by adopting multi-clock processing or inter-level data adjustment, which will add additional hardware consumption. Hence, based on the lifting scheme, we attempt to develop a high-throughput and hardware-efficient internal folded multilevel 2-D DWT architecture without complex multi-clock processing and complex inter-level data adjustment. Further analysis and optimizations are proceeding to overcome the mismatch problem, minimize the size of the logical units and the memory, and markedly improve the hardware efficiency.

The rest of this paper is organized as follows. Section 2 reviews the mathematical foundation of the lifting scheme and the flipping structure of the DWT. Section 3 presents the proposed architecture of the entire multi-level 2-D DWT, and Section 4 provides hardware estimation and comparison with previous architectures. Finally, Section 5 concludes.

2. Lifting Scheme

The lifting scheme was first proposed by Daubechies and Sweldens, and then modified into flipping structure by Huang et al. [18], shown as follows:

$$1/\alpha \times y(2n+1) = 1/\alpha \times x(2n+1) + x(2n) + x(2n+2) \tag{1}$$

$$1/\beta \times y(2n) = 1/\beta \times x(2n) + y(2n-1) + y(2n+1) \tag{2}$$

$$1/\gamma \times H(2n+1) = 1/\gamma \times y(2n+1) + y(2n) + y(2n+2) \tag{3}$$

$$1/\delta \times L(2n) = 1/\delta \times y(2n) + H(2n-1) + H(2n+1) \tag{4}$$

$$H^\circ(2n+1) = K \times H(2n+1) \tag{5}$$

$$L^\circ(2n) = 1/K \times L(2n) \tag{6}$$

where x represents the input pixel, y, H and L mean temporal variables. H° and L° stand for the final results of high frequency and low frequency, respectively. α, β, γ, δ and K are the constant coefficients. $\alpha = -1.586134342$, $\beta = -0.052980118$, $\gamma = 0.882911075$, $\delta = 0.443506852$, $K = 1.230174105$.

The flipping structure can provide a variety of hardware implementations to improve and possibly minimize the critical path as well as the memory requirement of the lifting-based discrete wavelet transform [18]. Moreover, the flipping structure has less computational complexity and the forms of flipping formulas are highly consistent. From the observation of Equations (1)–(4), we can find that each equation involves a multiplication operation and two addition operations. And the four formulas are extremely similar with the same basic operational structure, which of (1) and (2) are shown in Figure 1. Thus, they can be considered as basic operations and can be reused to reduce the computing resources. For further analysis of the pipeline structure shown in Figure 1, considering that $Tm \approx 2Ta$ [14], the multiplication and addition operations in each basic operation can be performed simultaneously without any bad effect on the critical path, the multiplication items of the flipping method (2)–(4) arrive at least one clock cycle ahead of their respective addition items. For instance, if the multiplication item $x(2n)$ in (2) arrives at the Xth clock cycle, the addition items $y(2n + 1)$ and $y(2n - 1)$ will be obtained through (1) at the $Xth + 1$ clock cycle at least, where X is defined as the number of clock cycles. So in each lifting scheme for (2)–(4), multiplication items will arrive one clock cycle advanced than the addition items as well.
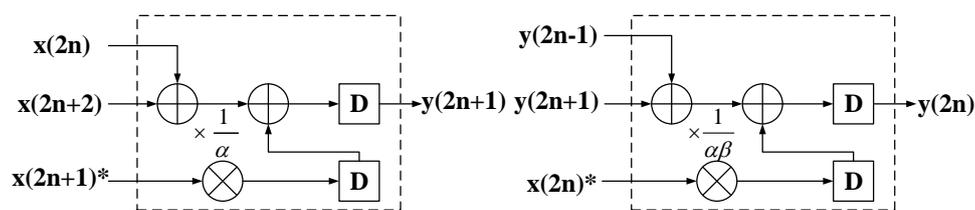
Figure 1. Basic operation structures of (1) and (2). Note: The items with * are the multiplication terms that should arrive ahead of time.

From Figure 1, it can be perceived that each basic operation module is completed within one clock cycle, so that the addition terms arrive just one clock cycle behind the multiplication term in the next formula. Thus, when the addition terms arrive in the current clock cycle, the two additions will be implemented with the multiplication item which has been multiplied. This method can not only minimize the number of registers, but also ensure the critical path of $2Ta$. Meanwhile, the multiplication factors can be selected to achieve the reuse of basic operation modules, since the four basic operations only differ in multiplication factors.

3. Proposed Architecture for Muti-Level 2-D DWT

3.1. Data Scanning Method

As each basic operation has three inputs, and [14] proved that it can reach the best comprehensive effect by overlapping one pixel, the proposed architecture uses multiple 3-input parallel line-based scanning method. The condition of parallelism S = 1 is presented in Figure 2, in which the gray pixels are overlapping pixels and CLKX represents the *Xth* clock cycle of current line. In order to meet the parallel execution requirement of multiplications and additions in each basic operation, this design adjusts the data input timing. Based on the particular scanning method, we adopt clock-to-data misalignment method to make $x(2n + 1)$ feed to the first-level 2-D DWT one clock cycle ahead of $x(2n)$ and $x(2n + 2)$. Then, each basic operation can be executed in parallel.
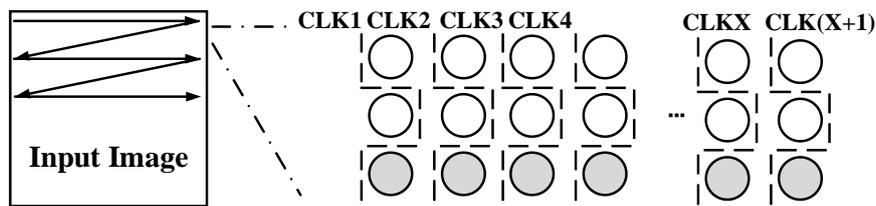

Figure 2. Scanning method for S = 1.

3.2. Unfolded Architecture

Fractal sets are characterized by their self-similarity property, that is each part of the set has the same or approximate shape of the whole set [19]. In the decomposed results of the lower level in the unfolded architecture, namely four sub-bands low-high (LH), high-low (HL), high-high (HH) and low-low (LL), only the LL sub-band is fed to the above DWT level, while the others output directly. Hence, the ratio of throughput between the lower level and the above level is 4:1. This means if we adopt the same clock for both the lower level and the above level, there will be a waste of many clock cycles. Besides, using multi-clock method will increase the area of the clock tree and the complexity of the system. On account of this, the overall module designed with a 2:1 parallelism ratio between the lower level and the above level is proposed, as shown in Figure 3. As a result, a working set of data is fed to the first-level DWT every clock cycle and the clock-to-data ratio is 1:1. Later, a working set of data is fed to the second-level DWT every two clock cycles and the clock-to-data ratio is 2:1. Then, a working set of data is fed to the third-level DWT every four clock cycles and the clock-to-data ratio is 4:1.
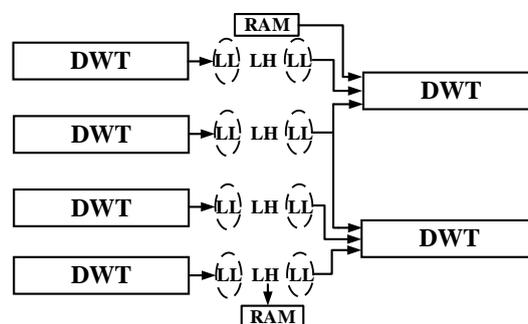

Figure 3. Structure of the lower level and the above level with a 2:1 parallelism ratio.

3.3. Proposed Multi-Level DWT Architecture

For the first-level 2-D DWT, in which the clock-to-data ratio is 1:1, since the clock match the input data, we adopt an internal unfolded structure. Namely, four basic operation structures mentioned in the previous presentation, are connected in turn to constitute the 1-D 9/7 DWT structure, as demonstrated in Figure 4. In it, the term $x(2n + 1)$ with * is the item which is fetched ahead and D_x means the data obtained at the *Xth* clock cycle. And this architecture can be implemented in the column and row filter by correctly selecting the RAM or Buffer. Moreover, the 2-D DWT consists of column filter, transposing buffer, row filter and scaling module. Meanwhile, if multiple 2-D DWT modules are carried out in parallel, the intermediate variables, $y(2n + 1)$, $y(2n)$ and $H(2n + 1)$, will be transferred to the next column filter in parallel without being stored in RAM. So the intermediate variables fetched from RAM previously will be obtained directly from the preceding column filter. Given the above, the structure of a single-level 2-D DWT with parallelism S is presented in Figure 5. And the 2-D DWT structure is shown in the dotted box of Figure 5.
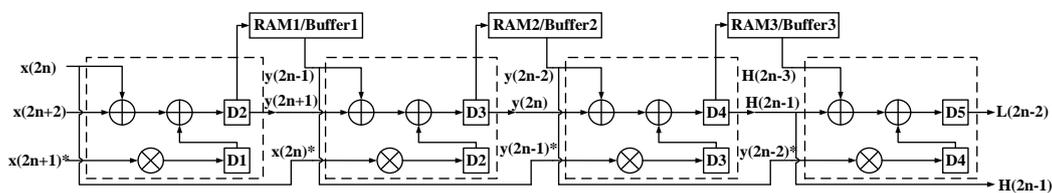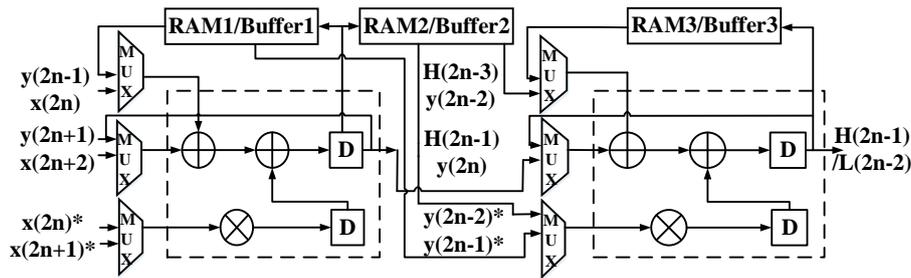

Figure 4. The 1-D discrete wavelet transform (DWT) structure for the first level. Note: The items with * are the multiplication terms that should arrive ahead of time.
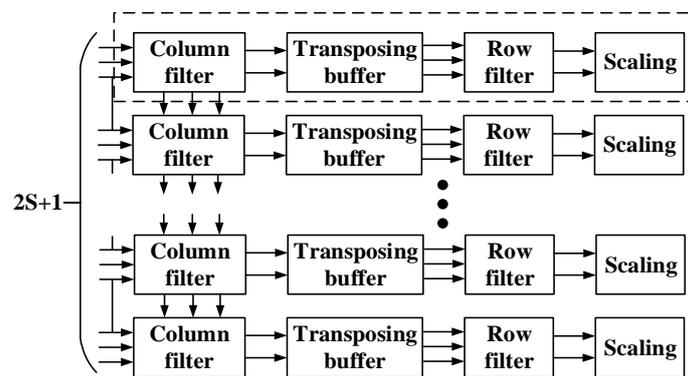

Figure 5. Architecture of a single-level 2-D DWT with parallelism S.

For the second-level 2-D DWT, since a set of data is valid every two clock cycles, the 1-D 9/7 DWT structure can be partially folded. That is, for the four basic calculations of a set of data, two basic operation modules are needed. As in the 1-D DWT module shown in Figure 6, after $x(2n)$, $x(2n + 1)$* and $x(2n + 2)$ enter the first basic module, $y(2n + 1)$ will be obtained at the second clock cycle through (1). Meanwhile, the intermediate variable $y(2n + 1)$ will be re-entered into the first basic module as input. Then, $y(2n)$ can be figured out through (2) at the third clock cycle. Similarly, the Formulas (3) and (4) also use this method to reuse the basic operation module. This ensures that the processing of data in each clock cycle is effective against adopting multi-clock. Furthermore, this architecture can be applied for the column and row filter by properly selecting the RAM or Buffer. It should be noticed that the Buffer in the second level has four buffers. Moreover, transposing module is needed to adjust the order for the output data onto the column filter.

Figure 6. The 1-D DWT structure for the second level. Note: The items with * are the multiplication terms that should arrive ahead of time.

For the third-level 2-D DWT, since a set of data is valid every four clock cycles, the 1-D DWT module can be fully folded. Namely, for the four basic operations of a set of data, only one basic operation module is used, as shown in Figure 7. After a set of valid data entering the module, the L and H coefficients can be obtained from the reused basic operation module within four clock cycles. And once the four basic operations completed, the next set of valid data exactly arrives and the same processing will be done. Moreover, by accurately selecting the RAM or Buffer, the 1-D module can be applied in the column and row filter. Similarly, the Buffer here has eight buffers. Meanwhile, the transposing module is also demanded.
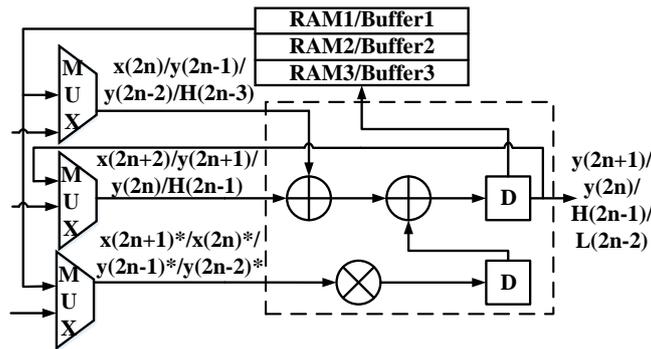
Figure 7. The 1-D DWT structure for the third level. Note: The items with * are the multiplication terms that should arrive ahead of time.

As mentioned above, in order to achieve the right clock-to-data ratio and meet the order for the data flow required by the row filter, the transposing buffer is needed in each 2-D DWT. Hence, it is necessary to design suitable transposing modules for the DWT architectures of different clock-to-data ratios, as shown in Figure 8, where the blanks represent the invalid data. Once the output data onto the column filter successively enter the transposing buffer, they will be temporally stored by different numbers of registers, and selected by the multiplexers according to the sequence demanded by the row filter. Besides, only one scaling module is used to finish the scaling computation in each level of the 3-level DWT architecture. That is, the data onto LL and HH sub-bands should be multiplied with the factor $(\alpha \times \beta \times \gamma \times \delta/K)^2$ and $(\alpha \times \beta \times \gamma \times K)^2$, respectively. Moreover, the data onto LH and HL sub-bands should be multiplied with the factor $(\alpha \times \beta \times \gamma)^2 \times \delta$.
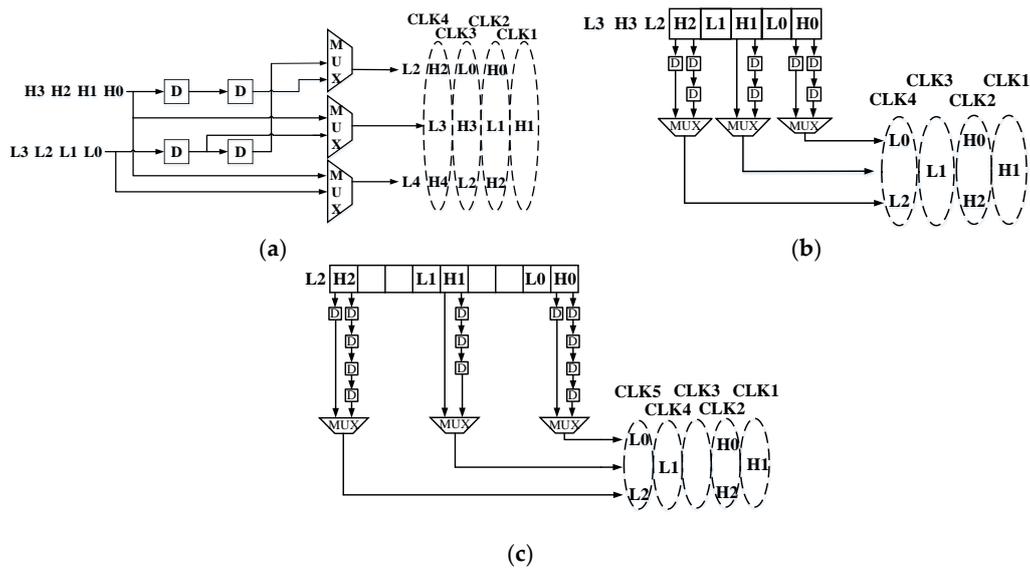
Figure 8. Structure of the transposing module and the orders of the input and output. (**a**) For the condition of 1:1 clock-to-data ratio. (**b**) For the condition of 2:1 clock-to-data ratio. (**c**) For the condition of 4:1 clock-to-data ratio.

It should be noticed that, the proposed architecture can integrate each 3-level DWT system into a clock domain, so it can extend to higher levels, by dividing the entire multilevel DWT system into multiple 3-level DWT systems. For example, in practical applications, such as JPEG2000 image compression, 5-level 2-D DWT can reach the nearly ideal compression performance for full-resolution image [20]. Thus, the entire 5-level DWT system can be divided into two 3-level DWT systems. Namely, the first-level, second-level and third-level DWT constitute the first 3-level DWT system in the first clock domain, and their structures are as shown in the previous presentation. Then, the fourth-level and fifth-level DWT constitute the second 3-level DWT system in the second clock domain. And the fourth-level DWT has the same structure as the first-level DWT and the fifth-level DWT has the same structure as the second-level DWT. Moreover, the overall module design with a 2:1 parallelism ratio between the lower level and the above level always works. It is a fact that is preferable to analyze the scattering problems of the TD framework rather than in the frequency domain (FD) [21]. Hence, the proposed architecture has strong application value.

4. Hardware Estimation and Comparison

4.1. Hardware Estimation

On the assumption that the input image is $N \times N$ with 8-bit depth, the hardware consumption of the entire 3-level 2-D DWT architecture is listed in Table 1, where 1:1, 2:1 and 4:1 represent the structures with 1:1, 2:1 and 4:1 clock-to-data ratio respectively, the clock-to-data ratio represents the number of clock cycles it takes to get a valid LL component of data, and S represents the parallelism. Totally, the architecture has $6N$ words temporal buffer, $53S/4$ multipliers, $21S$ adders and $229S/4$ registers.

Table 1. Hardware consumption of 3-level 2-D DWT architecture for 9/7 filter with $N \times N$ image size.

Architecture	Multiplier	Adder	Register	Temporal RAM (in Word)	Parallelism
1:1	10	16	28	$3N$	S
2:1	5	8	33	$2N$	$S/2$
4:1	3	4	51	N	$S/4$
3-level	$53S/4$	$21S$	$229S/4$	$6N$	–

4.2. Performance Comparison

In order to more intuitively to compare the hardware efficiency of each architecture, the comprehensive evaluation criteria area-delay-product (ADP) is proposed to [12]. However, ADP suffers high relevance to constraint rules and technologies. Thus, we adopt the transistor-delay-product (TDP) [14] to assess the hardware efficiency of architectures. The equation of TDP is shown in (7), where TC (Transistor Count) stands for the count of transistors and ACT (Active Cycle Time) is the computation time of an image in clock cycles, which can be calculated by ACT = N^2/throughput.

$$TDP = TC \times CPD \times ACT \text{ (transistor} \cdot \text{s)} \tag{7}$$

Hence, it can be considered that TDP takes into account the hardware consumption and total computing time for an image of the proposed architecture. Furthermore, the smaller the TDP is, the better hardware efficiency the architecture achieves.

For the assessment of the transistor count, a method for calculating the number of transistors based on the ripple carry adders (RCA), RCA-based multipliers, D flip-flops register and single-port SRAM for all the memory words, which are assumed to be implemented for all our structures is proposed to [9]. Also, it is assumed that $Tm = 2Ta$ and $Ta = 3.01$ ns.

After paired with these factors, Table 2 shows the comprehensive comparison of the 3-level 2-D DWT using different architectures of 512 × 512 image size. As discussed previously, the proposed structure adopts an overall module design with a 2:1 parallelism ratio between the lower level and the above level, which reduces the temporal memory between two levels and eliminates the frame buffer compared with the folded structure. And the internal folded architecture reduces the consumption of computing resources. Meanwhile, the CPD of the proposed structure is also reduced to 2Ta. Besides, all of these optimizations are reflected on TDP. For S = 8, the proposed architecture has the highest throughput rate. This architecture increases the hardware efficiency by more than 22.4% for S = 8 and 25.77% for S = 16 in TDP, compared with the existing parallel architectures.

Table 2. Hardware estimation and performance comparison of 3-level 2-D architecture for 9/7 filter with 512 × 512 image size.

Architecture	S	Throughout Rate	Multiplier	Adder	Register	MEM Words	CPD	TC (×10⁶)	ACT	TDP
[14] *	1	2/Ta	0	123	167	3840	Ta	0.509	131,072	200.93
[8]	8	4/Ta	96	128	6304	82,144	4Ta	10.13	21,504	2621.69
[9]	8	4/Ta	99	176	158	5696	4Ta	1.26	16,384	247.63
[10]	8	8/Ta	189	294	443	2688	2Ta	1.62	16,384	160.13
[12]	8	16/3Ta	111	180	341	1536	3Ta	0.975	16,384	144.18
Proposed	8	8/Ta	106	168	458	3072	2Ta	1.12	16,384	110.17
[12]	16	32/3Ta	216	348	682	1536	3Ta	1.76	8192	130.30
[17]	x	64/3Ta	0	1280	x	30,016	3Ta	x	4096	x
Proposed	16	16/Ta	212	336	916	3072	2Ta	1.94	8192	95.64

MEM: memory, where MEM represents the sum of temporal RAM and frame buffer. * Extra 66 subtractors used in [14] are not listed. x: letter x, where x represents the unknow data.

Then the synthesis results and comparison with the existing architecture in the same TSMC 90 nm CMOS library with Synopsys Design Compiler are tabulated in Table 3. Meanwhile, the power is estimated at 20 MHz frequency. EPI (energy per image), the energy consumption of decomposing an image, is also calculated and listed in Table 3. It can be seen that, the proposed architecture achieves the least ADP, about 30.6% less than the others, and the least EPI.

Table 3. Synthesis results of 3-level 2-D architecture for 9/7 filter with 512×512 image size.

Architecture	S	DAT (ns)	Area (μm^2)	Power (mw)	ADP (μm^2)	EPI (μJ)
[8]	8	42.66	3,377,870.70	24.45	3098.72	26.28
[9]	8	45.58	3,104,371.05	22.59	2318.29	18.50
[10]	8	25.42	2,139,397.29	15.26	891.01	12.50
Proposed	8	27.70	1,362,035.87	12.94	618.14	10.60

5. Conclusions

In this brief, we have proposed an internal folded multilevel 2-D DWT architecture of better hardware efficiency. Three-input parallel clock-to-data misalignment line-based scanning method is used and architectures with different clock-to-data ratios for different levels are implemented. We have applied this architecture to the compression chips. Compared with the non-parallel structure of [14], the proposed architecture sacrifices 2.03 times TC for 4 times increase in throughput rate, which results in 49.16% less TDP. Compared with the folded structure of [8], the proposed architecture involves 2 times higher throughput rate and 9.71 times less TC, which result in 94.85% less TDP. Compared with the structure of [9], the proposed architecture involves 2 times higher throughput rate and 1.139 times less TC, which result in 56.09% less TDP. Even compared with the structure of [10], which has the same throughput rate, the proposed architecture involves 1.43 times less TC, which results in 30.04% less TDP. Compared with the structure of [12], the proposed architecture sacrifices 1.17 times TC for 1.5 times increase in throughput rate, which results in 22.4% less TDP. For S = 16, compared with the structure of [12], the proposed architecture sacrifices 1.11 times TC for 1.5 times increase in throughput rate, which results in 25.77% less TDP. On the whole, the proposed architecture achieves less TDP than others, about at least 22.4% for S = 8 and 25.77% for S = 16 better than all the existing architectures. Moreover, the ASIC synthesis result shows that the proposed architecture is 30.6% smaller in ADP than the existing parallel architectures of S = 8.

References

1. Zhang, W.; Jiang, Z.; Gao, Z.; Liu, Y. An efficient VLSI architecture for lifting-based discrete wavelet transform. *IEEE Trans. Circuits Syst. II Exp. Briefs* **2012**, *59*, 158–162. [CrossRef]
2. Darji, A.; Agrawal, S.; Oza, A. Dual-Scan parallel flipping architecture for a lifting-Based 2-D discrete wavelet transform. *IEEE Trans. Circuits Syst. II Exp. Briefs* **2014**, *61*, 433–437. [CrossRef]
3. Hu, Y.; Jong, C. A memory-efficient scalable architecture for lifting-Based discrete wavelet transform. *IEEE Trans. Circuits Syst. II Exp. Briefs* **2013**, *60*, 502–506. [CrossRef]
4. Mohanty, B.K.; Meher, P.K. Area-delay-power-efficient architecture for folded two-dimensional discrete wavelet transform by multiple lifting computation. *IET Image Process.* **2014**, *8*, 345–353. [CrossRef]
5. Mohanty, B.K.; Meher, P.K.; Srikanthan, T. Critical-path optimization for efficient hardware realization of lifting and flipping DWTs. In Proceedings of the IEEE ISCAS, Lisbon, Portugal, 24–27 May 2015; pp. 1186–1189.
6. Todkar, S.; Shastry, P.V.S. Flipping based high performance pipelined VLSI architecture for 2-D discrete wavelet transform. In Proceedings of the IEEE iCATccT, Davangere, India, 29–31 October 2015; pp. 832–836.
7. Darji, A.; Limaye, A. Memory efficient VLSI architecture for lifting-based DWT. In Proceedings of the IEEE MWSCAS, College Station, TX, USA, 3 August 2014; pp. 189–192.

8.   Tian, X.; Wu, L.; Tan, Y.H.; Tian, J.W. Efficient multi-input/multioutput VLSI architecture for 2-D lifting-based discrete wavelet transform. *IEEE Trans. Comput.* **2011**, *60*, 1207–1211. [CrossRef]

9.   Mohanty, B.K.; Meher, P.K. Memory efficient modular VLSI architecture for high throughput and Low-Latency Implementation of Multilevel Lifting 2-D DWT. *IEEE Trans. Signal Process.* **2011**, *59*, 2072–2084. [CrossRef]

10. Mohanty, B.K.; Meher, P.K. Memory-Efficient High-Speed Convolution-Based Generic Structure for Multilevel 2-D DWT. *IEEE Trans. Circuits Syst. Video Technol.* **2013**, *23*, 353–363. [CrossRef]

11. Hu, Y.; Prasanna, V.K. Energy- and area-efficient parameterized lifting-based 2-D DWT architecture on FPGA. In Proceedings of the IEEE HPEC, Waltham, MA, USA, 9–11 September 2014; pp. 1–6.

12. Hu, Y.; Jong, C. A Memory-Efficient High-Throughput Architecture for Lifting-Based Multi-Level 2-D DWT. *IEEE Trans. Signal Process.* **2013**, *61*, 4975–4987. [CrossRef]

13. Ye, L.; Hou, Z. Memory Efficient Multilevel Discrete Wavelet Transform Schemes for JPEG2000. *IEEE Trans. Circuits Syst. Video Technol.* **2015**, *25*, 1773–1785. [CrossRef]

14. Wu, C.; Zhang, W.; Jia, Q.; Liu, Y. Hardware Efficient Multiplier-less Multi-level 2D DWT Architecture without off-chip RAM. *IET Image Process.* **2017**, *11*, 362–369. [CrossRef]

15. Chakraborty, A.; Chakraborty, D.; Banerjee, A. A multiplier less VLSI architecture of modified lifting based 1D/2D DWT using speculative adder. In Proceedings of the International Conference on Communication and Signal Processing, Chennai, India, 6–8 April 2017.

16. Anirban, C.; Ayan, B. Low Area & Memory Efficient VLSI Architecture of 1D/2D DWT for Real Time Image Decomposition. In Proceedings of the 2018 8th International Symposium on Embedded Computing and System Design (ISED), Cochin, India, 13–15 December 2018.

17. Abhishek, C.; Basant, K.M. A Block based Area-Delay Efficient Architecture for Multi-Level Lifting 2-D DWT. *Springer Nat.* **2018**, *37*, 4482–4503.

18. Huang, C.-T.; Tseng, P.-C.; Chen, L.-G. Flipping structure: An efficient VLSI architecture for lifting-based discrete wavelet transform. *IEEE Trans. Signal Process.* **2004**, *52*, 1080–1089. [CrossRef]

19. Emanuel, G. Harmonic Sierpinski Gasket and Applications. *Entropy* **2018**, *20*, 714.

20. Taubman, D.S.; Marcellin, M.W. *Jpeg2000: Image Compression Fundamentals, Standards and Practice*; Kluwer: Norwell, MA, USA, 2001.

21. Frongillo, M.; Gennarelli, G.; Riccio, G. TD-UAPO diffracted field evaluation for penetrable wedges with acute apex angle. *J. Opt. Soc. Am. A* **2015**, *32*, 1271–1274. [CrossRef] [PubMed]