



ANALYSIS ON IMPLEMENTATION OF LIFTING SCHEME IN DWT ARCHITECTURE FOR EFFICIENT MEMORY

Ashok.P¹, Thirumaraiselvi.C²

¹PG Scholar, Department OF ECE, ²Assitant Professor, Department of ECE

^{1,2} Sri Krishna College of Engineering and Technology, Kuniamuthur P.O., Coimbatore-641008, Tamil Nadu, India.
¹glanceashok@gmail.com, ²thirumaraiselvi@skcet.ac.in

Abstract—A high-throughput scalable architecture for 2-D DWT is presented for efficient memory handling. Various existing DWT architectures was analyzed and observed that data scanning method has a significant impact on the memory efficiency of DWT architecture. Hence, a novel parallel stripe-based scanning method based on the analysis of the dependency graph of the lifting scheme is proposed. With the new scanning method for multi-level 2D DWT, a high memory efficient scalable parallel pipelined architecture is developed. The developed architecture requires no frame memory and 3-level DWT decomposition is adopted with an image of size $N \times N$ pixels with 32 pixels processed concurrently. The elimination of frame memory and the small temporal memory lead to significant reduction in overall size. Thus, this architecture has a regular structure and emphasizes the utilization of hardware. The synthesis results show that the proposed architecture achieves a better area-delay product by 60% and higher throughput by 97% when compared to the best existing design.

Keywords— Discrete Wavelet Transform, Parallel Stripe Based Scanning, Frame Memory, Temporal Memory

1. INTRODUCTION

Memory is a vital factor in designing any type of architecture. In regards with the read and write operation memory content has to be scanned, the scanning has the impact in the available hardware resources. So, in accordance with the scanning appropriate hardware resources has to be chosen.

Hence 2 factors has to be considered for the design of DWT architecture they are scanning technique and architecture implementation of scanning technique. The analysis technique include line based scanning, modified line based scanning, block based scanning, stripe based scanning, overlapped stripe based scanning,

lifting scheme for 1D and 2D DWT, reduce in the size of some registers say transposition memory, pipeline architecture, multilevel lifting scheme architecture.

The key point is the requirement of fast scanning such that it doesn't affect any hardware resources that are incorporated for the design of memory structure. This includes the consideration of limited hardware resources so that area could be reduced with the complete utilization of hardware resources.

2. SURVEY ANALYSIS FOR EFFICIENT MEMORY DESIGN

2.1 LINE BASED SCANNING METHOD

A high performance and memory-efficient pipelined architecture with parallel scanning method is introduced for 2-D lifting-based DWT in JPEG2000 applications. The Proposed 2-D DWT architecture is composed of two 1-D DWT cores and a 2×2 transposing register array. The proposed 1-D DWT core consumes two input data and produces two output coefficients per cycle, and its critical path takes one multiplier delay only. Moreover, we utilize the parallel scanning method to reduce the internal buffer size instead of the line-based scanning method. For the $N \times N$ tile image with one-level 2-D DWT decomposition, only $4N$ temporal memory and the 2×2 register array are required for 9/7 filter to store the intermediate coefficients in the column 1-D DWT core. And the column-processed data can be rearranged in the transposing array. According to the comparison results, the hardware cost of the 1-D DWT core and the internal memory requirements of proposed 2-D DWT architecture are smaller than other familiar architectures based on the same throughput rate. The implementation results show that the proposed 2-D DWT architecture can process 1080p HDTV pictures with five-level decomposition at 30 frames/sec. On analysis it is found that complete processing of row is done before Proceeding to next row and Data is processed as soon as it is scanned in. The drawback behind this is the temporal memory is needed in addition to store the intermediate results; also cDWT has to wait for some time to get output from DWT.

ARCHITECTURE	MULTI	ADDER	BUFFER – TRANS	TEMPORAL BUFFER	THROUGHPUT RATE
Generic RAM	10	16	1.5N	4N	2 – I/O
DSA	12	16	0	4N	2 – I/O
Flipping w/o pipeline	10	16	1.5N	4N	2 – I/O
Flipping with 5 stage pipeline	10	16	1.5N	11N	2 – I/O
Modified Lifting	6	8	1.5N	4N	1 – I/O
Cell based	12	16	0	7.5N	2 – I/O
Proposed	10	16	0	4N	2 – I/O

Table 1. Analysis of 2D DWT architectures for 9/7 filter

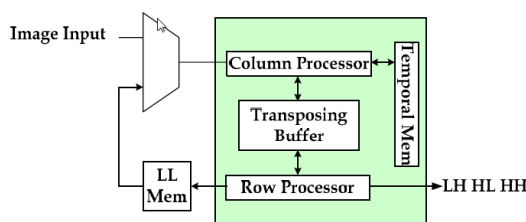


Fig 1. Block diagram of proposed 2D DWT

2.2 MODIFIED LINE BASED SCANNING METHOD

Efficient line-based architectures for two-dimensional discrete wavelet transform (2-D DWT) are presented in this paper. It is said that four-input/four-output architecture for direct 2-D DWT that 1-level decomposition of an $N \times N$ image could be performed in approximately $N^2/4$ intra-working clock cycles (ccs), where the parallelism among four sub bands transforms in lifting-based 2-D DWT is explored. By using this four-input/four output architecture, we propose a novel pipelined architecture for multilevel 2-D DWT that can perform a complete dyadic decomposition of image in approximately $N^2/4$ ccs.

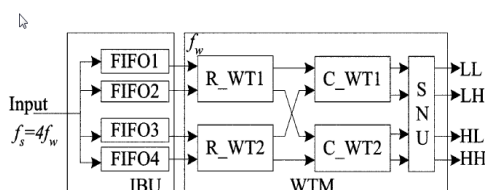


Fig 2 4in – 4out 1 level transform

Performance analysis and comparison results demonstrate that, the proposed architectures have faster throughput rate and good performance in terms of production of throughput rate and hardware cost, as well as hardware utilization. The proposed pipelined architecture could be an efficient alternative for high-speed and/or low-power applications.

It makes benefit of performing Simultaneous conduction of alternative rows and columns is done. Hence, cDWT need not wait for input from rDWT, enough input is provided for it by simultaneous conduction. The limiting factor is fixed size transposition memory and large temporal memory is needed and there is a resource increase for the storage of interleaving results.

2.3 BLOCK BASED SCANNING METHOD

A systematic high-speed VLSI implementation of the discrete wavelet transform (DWT) based on hardware-efficient parallel FIR filter structures is presented. High-speed 2-D DWT with computation time as low as $N^2/12$ can be easily achieved for an $N \times N$ image with controlled increase of hardware cost.

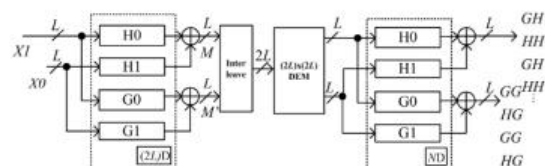


Fig 3. Proposed $N^2/2L$ 2D DWT Structure

Table 2 Lifting and Flipping Comparisons

2D DWT	MULTI	ADDER	SIZE – ON CHIP	SIZE – OFF CHIP	COMPUTING TIME
$N^2/4$	24	76	$30N+16+60J$	$2N+8(J-1)$	$N^2/3$
$N^2/2$	10	40	$14N+4+14J$	$N+2(J-1)$	$2N^2/3$
$N^2/2$	10	38	$24N+4+24J$	$N+2(J-1)$	$2N^2/3$
Lifting	8	8	$4N+14$	0	N^2+N+2J
Flipping	6	8	$12N$	0	$4N^2/3$

Compared with recently published 2-D DWT architectures with computation time of $N^2/3$ and $2N^2/3$, the proposed designs can also save a large amount of multipliers and/or storage elements. It can also be used to implement those 2-D DWT traditionally suitable for lifting or flipping-based designs, such as (9, 7) and (6, 10) DWT.

The throughput rate can be improved by a factor of 4 by the proposed approach, but the hardware cost increases by a factor of around 3. Furthermore, the proposed designs have very simple control signals, regular structures and 100% hardware utilization for continuous images. It is inferred that it has high throughput because image is divided into blocks and scanned row by row for each separate block and Convolution type of architecture is adopted here. It is said that temporal memory is large and it requires large amount of arithmetic resources.

2.4 STRIPE BASED SCANNING METHOD

This paper presents optimal data transfer and buffering schemes for JPEG 2000 encoder. The data transfer scheme Combines bit-level zero run-length coding and addressing mode to reduce data transfer time between Discrete Wavelet Transform (DWT) and Embedded Block Coding with Optimized Truncation (EBCOT) to 21%.

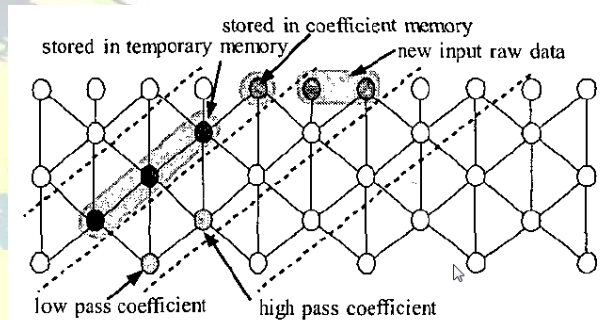


Fig 4. 1D DWT lifting scheme

Furthermore, this data transfer scheme also helps in word to-bit plane data format conversion from DWT to EBCOT. On the other hand, by jointly considering coding flow of both lifting-based DWT and EBCOT with different design parameters, the reduction in total buffer requirement for both DWT and EBCOT can up to a factor of 4.22.

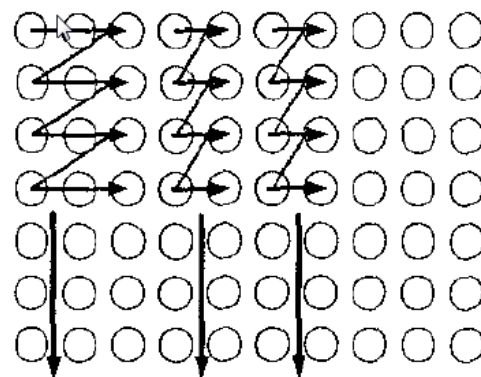


Fig 5 Scan order of Z scan 1D DWT

Hence Stripping is done such that image is completely scanned row by row, as a result it saves buffer size comparatively. But its Interleaving is done

between row and column, and it suffers from complex control scheme.

2.5 MODIFIED STRIPE BASED SCANNING

In this paper, a detailed analysis of very large scale integration (VLSI) architectures for the one-dimensional (1D) and two-dimensional (2-D) discrete wavelet transform (DWT) is presented in many aspects, and three related architectures are proposed as well. The 1-D DWT and inverse DWT (IDWT) architectures are classified into three categories: convolution-based, lifting based, and B-spline-based. They are discussed in terms of hardware complexity, critical path, and registers.

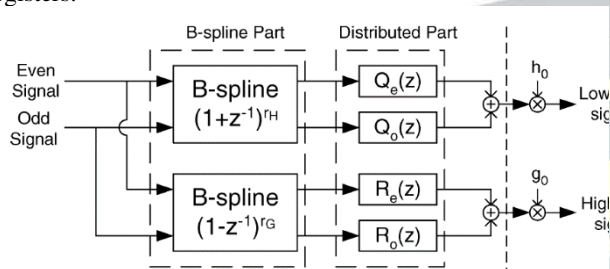


Fig 6 DWT architecture for B-spline

Fig 7 IDWT architecture for B-spline

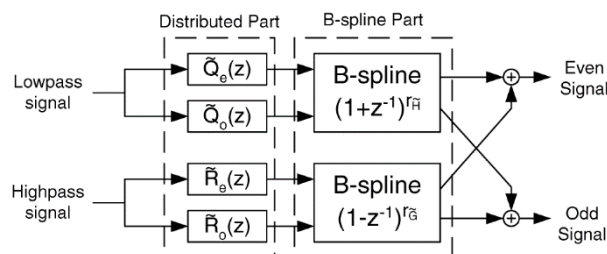
As for the 2-D DWT, the large amount of the frame memory access and the die area occupied by the embedded internal buffer become the most critical issues. The 2-D DWT architectures are categorized and analyzed by different external memory scan methods. The implementation issues of the internal buffer are also discussed, and some real-life experiments are given to show that the area and power for the internal buffer are highly related to memory technology and working frequency, instead of the required memory size only. Besides the analysis, the B-spline-based IDWT architecture and the overlapped stripe-based scan method are also proposed. Last, we propose a flexible and efficient architecture for a one-level 2-D DWT that exploits many advantages of the presented analysis.

The advantage of having this is the temporal memory is eliminated and partial results need not to be buffered. But it suffers from longer computation time and larger bandwidth and the Control scheme is larger.

2.6 2D DWT LIFTING

It is suggested a new data-access scheme for the computation of lifting two-dimensional (2-D) discrete wavelet transform (DWT) without using data transposition. We have derived a linear systolic array directly from the dependence graph (DG) and a 2-D systolic array from a suitably segmented DG for parallel

STRUCTURE	MULT	ADDER	REG	TP	ON CHIP	MUX	CT
Line	12	16	24	N	4N+24	2	MN/2
Parallel Scan	10	16	44	N	4N+24	2	MN/2



and pipeline implementation. These two systolic arrays are used as building blocks to derive the proposed transposition-free structure for lifting 2-D DWT. The proposed structure requires only a small on-chip memory of $(4N + 8P)$ words and processes a block of P samples in every cycle, where N is the image width.

Lifting	6P	8P	10P	$N(P+2)/2$	$8P+N(P+8)/2$	$P(6+(P-2)/2)$	MN/P
Multi-level	4.5P	8P	7.5P	2.5N	$7.5P+5.5N$		MN/P
Proposed	4.5P	8P	8P	N	$4N+8P$		MN/P

Table 3 Acquired results of proposed structure.

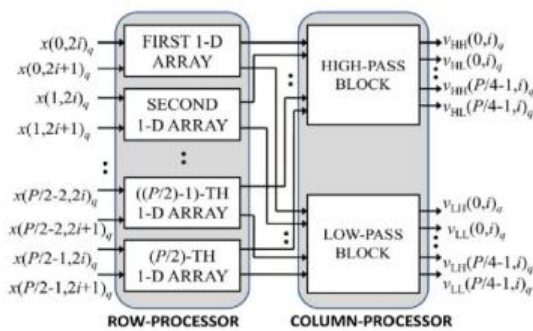


Fig 8 Block of Lifting 2D DWT

Moreover, it has small output latency of nine cycles and does not require control signals which are commonly used in most of the existing DWT structures. Compared with the best of the existing high-throughput structures, the proposed structure requires the same arithmetic resources but involves $1.5N$ less on-chip memory and offers the same throughput rate. ASIC synthesis result shows that the proposed structure for block size 8 and image size 512×512 involves 28% less area, 35% less area-delay product, and 27% less energy per image than the best of the corresponding existing structures. Apart from that, the proposed

structure is regular and modular; and it can be easily configured for different block sizes.

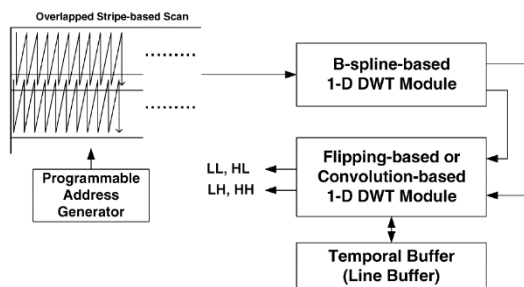


Fig 9 Proposed 1 level 2d DWT architecture

Here On chip memory consumption is less and it doesn't require control signals and has small output latency. This suffers from the fact that Critical delay is a

failure factor in this concept and computational time is more.

2.7 TRANSPOSITION MEMORY

High-speed and reduced-area 2-D discrete wavelet transform (2-D DWT) architecture is proposed. Previous DWT architectures are mostly based on the modified lifting scheme or the flipping structure. In order to achieve a critical path with only one multiplier, at least four pipelining stages are required for one lifting step, or a large temporal buffer is needed.

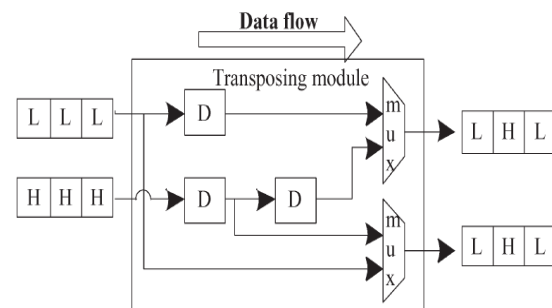


Fig 10 Architecture of transponding Module

In this brief, modifications are made to the lifting scheme, and the intermediate results are recombined and stored to reduce the number of pipelining stages. As a result, the number of registers can be reduced to 18 without extending the critical path. In addition, the two-input/two-output parallel scanning architecture is adopted in our design. For a 2-D DWT with the size of $N \times N$, the proposed architecture only requires three registers between the row and column filters as the transposing buffer, and a higher efficiency can be achieved. The advantage of using flipping structure that critical path is reduced and the Transposing buffer count is reduced for storing the data. But temporal buffer is increased in count for intermediate results. Although critical path is reduced, adders are increased.

2.8 PIPELINE ARCHITECTURE – TRANSPOSITION MEMORY

A high-performance and memory-efficient pipeline architecture which performs the one-level two-dimensional (2-D) discrete wavelet transform (DWT) in the 5/3 and 9/7 filters is proposed here. In general, the internal memory size of 2-D architecture highly depends on the pipeline registers of one-dimensional (1-D) DWT. Based on the lifting-based DWT algorithm, the primitive data path is modified and efficient pipeline architecture is derived to shorten the data path. Accordingly, under the same arithmetic resources, the 1-D DWT pipeline architecture can operate at a higher processing speed (up to 200 MHz in 0.25- m technology) than other pipelined architectures with direct implementation.

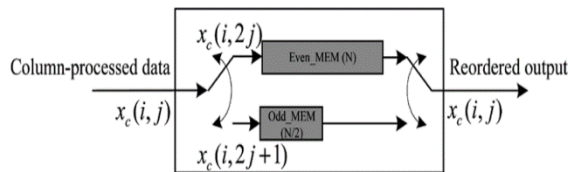


Fig 11: Block diagram of transposing buffer

The proposed 2-D DWT architecture is composed of two 1-D processors (column and row processors). Based on the modified algorithm, the row processor can partially execute each row-wise transform with only two column-processed data. Thus, the pipeline registers of 1-D architecture do not fully turn into the internal memory of 2-D DWT. For an image, $5N$ internal memory is required for the 5/3 filter, and $5N$ is required for the 9/7 filter to perform the one-level 2-D DWT decomposition with the critical path of one multiplier delay (i.e., and indicate the height and width of an image).

Table 4 Data flow of transposing buffer

Clk	Input	Output	Clk	Input	Output
0	$x_c(0,0)$		10	$x_c(2,2)$	$x_c(3,0)$
1	$x_c(1,0)$		11	$x_c(3,2)$	$x_c(3,1)$
2	$x_c(2,0)$		12	$x_c(0,3)$	$x_c(0,2)$
3	$x_c(3,0)$		13	$x_c(1,3)$	$x_c(0,3)$
4	$x_c(0,1)$	$x_c(0,0)$	14	$x_c(2,3)$	$x_c(1,2)$
5	$x_c(1,1)$	$x_c(0,1)$	15	$x_c(3,3)$	$x_c(1,3)$
6	$x_c(2,1)$	$x_c(1,0)$	16		$x_c(2,2)$
7	$x_c(3,1)$	$x_c(1,1)$	17		$x_c(2,3)$
8	$x_c(0,2)$	$x_c(2,0)$	18		$x_c(3,2)$
9	$x_c(1,2)$	$x_c(2,1)$	19		$x_c(3,3)$

The pipeline data path is regular and practicable. Finally, the proposed architecture implements the 5/3 and 9/7 filters by cascading the three key component. It is advantageous that data path of each pipeline stage is shortened and the internal memory size is minimized.



Fig 12 input and output orders of transposing buffer

The additional feature is that the Column processor, transposing buffer and row processor are cascaded together for efficient implementation. Although, internal memory size depends on the pipeline registers and the larger memory and more pipeline registers are used to increase the processing speed.

2.9 LIFTING BASED DWT

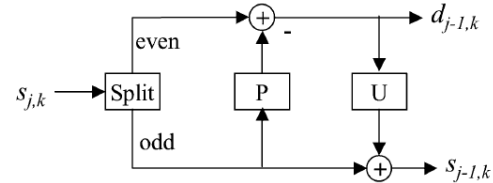


Fig 13 Lifting Scheme

The lifting scheme reduces the computational complexity of the discrete wavelet transform (DWT) by factoring the wavelet filters into cascades of simple lifting steps that process the input samples in pairs. Christo Ananth et al. [3] proposed a system which uses intermediate features of maximum overlap wavelet transform (IMOWT) as a pre-processing step. The coefficients derived from IMOWT are subjected to 2D histogram Grouping. This method is simple, fast and unsupervised. 2D histograms are used to obtain Grouping of color image. This Grouping output gives three segmentation maps which are fused together to get the final segmented output. This method produces good segmentation results when compared to the direct application of 2D Histogram Grouping. IMOWT is the efficient transform in which a set of wavelet features of the same size of various levels of resolutions and different local window sizes for different levels are used. IMOWT is efficient because of its time effectiveness, flexibility and translation invariance which are useful for good segmentation results.

The recursive and dual scan architectures can be readily extended to the 2-D case. The 2-D recursive architecture is roughly 25% faster than

conventional implementations, and it requires a buffer that stores only a few rows of the data array instead of a fixed fraction (typically 25% or more) of the entire array.

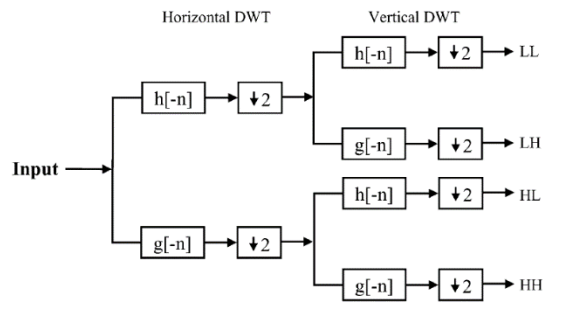


Fig 14 Block of 2D separable DWT

The 2-D dual scan architecture processes the column and row transforms simultaneously, and the memory buffer size is comparable to existing architectures. It is composed of simple processing units efficient for computing wavelet coefficients and it achieves efficiency gain by keeping data path hardware busy. The drawback behind this is, it will cause time complexity and reduction in speed and this could not be extended for multi wavelets.

2.10 MULTILEVEL 2D DWT

In this paper, a design strategy for the derivation of memory-efficient architecture for multilevel 2-D DWT is introduced. Using the proposed design scheme, we have derived a convolution-based generic architecture for the computation of three-level 2-D DWT based on Daubechies (Daub) as well as biorthogonal filters. The proposed structure does not involve frame-buffer. It involves line-buffers of size $3(K - 2)M/4$ which is independent of throughput-rate, where K is the order of Daubechies/biorthogonal wavelet filter and M is the image height.

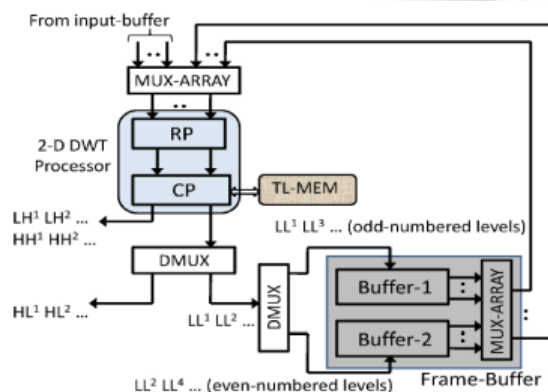


Fig 15 Generic structure of folded 2D DWT

This is a major advantage when the structure is implemented for higher throughput. The structure has regular data-flow, small cycle period TM and 100% hardware utilization efficiency. As per theoretical estimate, for image size 512×512 , the proposed structure for Daub-4 filter requires 152 more multipliers and 114 more adders, but involves 82 412 less memory words and takes 10.5 times less time to compute three-level 2-D DWT than the best of the existing convolution-based folded structures. Similarly, compared with the best of the existing lifting-based folded structures, proposed structure for 9/7-filter involves 93 more multipliers and 166 more adders, but uses 85 317 less memory words and requires 2.625 times less computation time for the same image size.

Table 6 Hardware complexities of various structures

Structures	MULT	ADD	MEM words	CT (in T_A)
Huang [6]	16	16	83 724	344 064
Meher [8]	16	12	83 462	344 064
Cheng [9] (8-parallel)	96	240	6752	65 535
Proposed (Daub-4)	168	126	1050	32 768
Xiong [11]	18	32	84 742	344 064
Lai [10]	10	16	84 780	516 096
Cheng [9] (2-parallel)	24	76	16 598	436 905
Mohanty [13] ($P = 16$)	99	176	5854	65 536
Tian [14] ($P = 16$)	96	128	88 448	86 016
Proposed (9/7-Filter)	189	294	3131	32 768

It involves 90 (nearly 47.6%) more multipliers and 118 (nearly 40.1%) more adders, but requires 2723 less memory words than the recently proposed parallel structure and performs the computation in nearly half the time of the other. In spite of having more arithmetic components than the lifting-based structures, the proposed structure offers significant saving of area and power over the other due to substantial reduction in memory size and smaller clock-period. ASIC synthesis result shows that, the proposed structure for Daub-4 involves 1.7 times less area delay-product (ADP) and consumes 1.21 times less energy per image (EPI) than the corresponding best available convolution based structure. It involves 2.6 times less ADP and consumes 1.48 times less EPI than the parallel lifting-based structure. The benefit of using this is it involves less area complexity and less computation time and involves significantly less on chip memory and provides high throughput. The drawback is flipping is done increasing



critical path delay and the data scanning impacts on memory size.

REFERENCES

- [1] K. Andra, C. Chakrabarti, and T. Acharya, "A VLSI architecture for lifting-based forward and inverse wavelet transform," *IEEE Trans. Signal Process.*, vol. 50, no. 4, pp. 966–977, 2002.
- [2] C. Cheng and K. K. Parhi, "High-speed VLSI implementation of 2-D discrete wavelet transform," *IEEE Trans. Signal Process.*, vol. 56, no. 1, pp. 393–403, 2008.
- [3] Christo Ananth, A.S.Senthilkani, S.Kamala Gomathy, J.Arockia Renilda, G.Blesslin Jebitha, Sankari @Saranya.S., "Color Image Segmentation using IMOWT with 2D Histogram Grouping", *International Journal of Computer Science and Mobile Computing (IJCSMC)*, Vol. 3, Issue. 5, May 2014, pp-1 – 7
- [4] C. Chrysafis and A. Ortega, "Line-based, reduced memory, wavelet image compression," *IEEE Trans. Image Process.*, vol. 9, no. 3, pp. 378–389, 2000.
- [5] C.-T. Huang, P.-C. Tseng, and L.-G. Chen, "Flipping structure: An efficient VLSI architecture for lifting-based discrete wavelet transform," *IEEE Trans. Signal Process.*, vol. 52, no. 4, pp. 1080–1089, 2004.
- [6] C.-T. Huang, P.-C. Tseng, and L.-G. Chen, "Analysis and VLSI architecture for 1-D and 2-D discrete wavelet transform," *IEEE Trans. Signal Process.*, vol. 53, no. 4, pp. 1575–1586, 2005.
- [7] Christo Ananth, A.S.Senthilkani, Praghash.K, Chakka Raja.M., Jerrin John, I.Annadurai, "Overlap Wavelet Transform for Image Segmentation", *International Journal of Electronics Communication and Computer Technology (IJECCCT)*, Volume 4, Issue 3 (May 2014), pp-656-658.
- [8] H.-Y. Liao, M. K. Mandal, and B. F. Cockburn, "Efficient architectures for 1-D and 2-D lifting-based wavelet transforms," *IEEE Trans. Signal Process.*, vol. 52, no. 5, pp. 1315–1326, 2004.
- [9] B. K. Mohanty and P. K. Meher, "Memory efficient modular VLSI architecture for highthroughput and low-latency implementation of multilevel lifting 2-D DWT," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 2072–2084, 2011.
- [10] B. K. Mohanty, A. Mahajan, and P. K. Meher, "Area- and power-efficient architecture for high-throughput implementation of lifting 2-D DWT," *IEEE Trans. Circuits Syst. II, Express Briefs*, vol. 59, no. 7, pp. 434–438, 2012.
- [11] B. K. Mohanty and P. K. Meher, "Memory-efficient high-speed convolution-based generic structure for multilevel 2-D DWT," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 23, no. 2, pp. 353–363, 2012.
- [12] B.-F.Wu and C.-F. Chung, "A high-performance and memory-efficient pipeline architecture for the 5/3 and 9/7 discrete wavelet transform of JPEG2000 codec," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 12, pp. 1615–1628, 2005.
- [13] C.-Y. Xiong, J. Tian, and J. Liu, "Efficient high-speed/low-power line-based architecture for two-dimensional discrete wavelet transform using lifting scheme," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 2, pp. 309–316, 2006.
- [14] C.-Y.Xiong, J. Tian, and J. Liu, "Efficient architectures for two-dimensional discrete wavelet transform using lifting scheme," *IEEE Trans. Image Process.*, vol. 16, no. 3, pp. 607–614, 2007.