ISSN 2395-695X (Print)
ISSN 2395-695X (Online)
International Journal of Advanced Research in Biology Engineering Science and Technology (IJARBEST)
Vol. 2, Special Issue 10, March 2016

ANALYSIS OF DWT ARCHITECTURE FOR EFFICIENT MEMORY USING LIFTING SCHEME

Ashok.P[1], Thirumaraiselvi.C[2]
[1]PG Scholar, Department OF ECE, [2]Assitant Professor, Department of ECE
[1, 2] Sri Krishna College of Engineering and Technology, Kuniamuthur P.O., Coimbatore-641008, Tamil Nadu, India.
[1]glanceashok@gmail.com,[2]thirumaraiselvi@skcet.ac.in

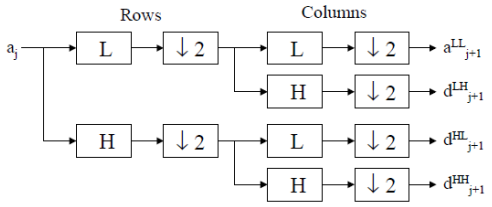
Abstract— A high-throughput scalable architecture for 2-D DWT is presented for efficient memory handling. Various existing DWT architectures was analyzed and observed that data scanning method has a significant impact on the memory efficiency of DWT architecture. Hence, a novel parallel stripe-based 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*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

1.1DISCRETE WAVELET TRANSFORM

The discrete wavelet transform (DWT) is a linear transformation that operates on a data vector whose length is an integer power of two, transforming it into a numerically different vector of the same length. It is a tool that separates data into different frequency components, and then studies each component with resolution matched to its scale. DWT is computed with a cascade of filtering's followed by a factor 2 subsampling (Fig1.1). H and L denote high and low-pass filters respectively, ↓ 2 denotes subsampling.



ISSN 2395-695X (Print)
ISSN 2395-695X (Online)
International Journal of Advanced Research in Biology Engineering Science and Technology (IJARBEST)
Vol. 2, Special Issue 10, March 2016

Fig.1.1 DWT TREE

Elements aj are used for next step (scale) of the transform and elements dj, called wavelet coefficients, determine output of the transform. l[n] and h[n] are coefficients of low and high-pas filters respectively One can assume that on scale j+1 there is only half from number of a and d elements on scale j. This causes that DWT can be done until only two aj elements remain in the analyzed signal these elements are called scaling function coefficients.DWT algorithm for two-dimensional pictures is similar. The DWT is performed firstly for all image rows and then for all columns (Fig 1.2).

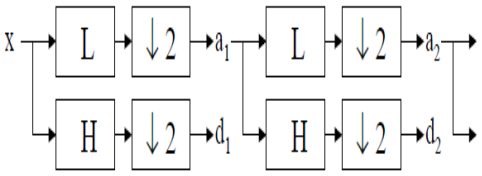


Fig.1.2 wavelet decomposition of 2D pictures

1.2 DWT IN LIFTING

DWT has traditionally been implemented by convolution or FIR filter bank structures. Such implementations require both a large number of arithmetic computations and a large storage—features that are not desirable for either high speed or low power image/video processing applications. This new approach is called the lifting-based wavelet transform or simply lifting. The main feature of the lifting-based DWT scheme is to break up the high-pass and low-pass wavelet filters into a sequence of upper and lower triangular matrices, and convert the filter implementation into banded matrix multiplications. This scheme in Fig.1.3 often requires far fewer computations compared to the convolution based DWT and offers many other advantages. The popularity of lifting-based DWT has triggered the development of several architectures in recent years. These architectures range from highly parallel architectures to programmable DSP-based architectures to folded architectures. In this paper we present a survey of these architectures. We provide a systematic derivation of these architectures and comment on their hardware and timing requirements.

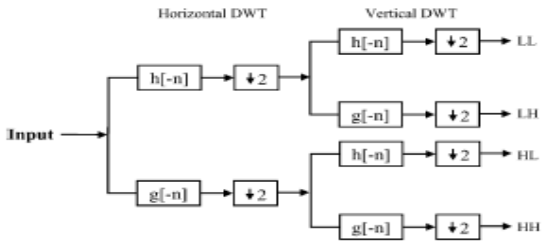


Fig.1.3 2-D separable DWT

2. RELATED WORK

ISSN 2395-695X (Print)
ISSN 2395-695X (Online)
International Journal of Advanced Research in Biology Engineering Science and Technology (IJARBEST)
Vol. 2, Special Issue 10, March 2016

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×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. 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 rDWT.

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*N image could be performed in approximately N2/4intra-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 N2/4 ccs. 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 N2/ 12 can be easily achieved for an N*N image with controlled increase of hardware cost. Compared with recently published 2-D DWT architectures with computation time of N2/3and 2N2/ 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

ISSN 2395-695X (Print)
ISSN 2395-695X (Online)
International Journal of Advanced Research in Biology Engineering Science and Technology (IJARBEST)
Vol. 2, Special Issue 10, March 2016

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.

3. PROPOSED SYSTEM

3.1 LIFTING SCHEME AND FLIPPING METHOD

The lifting scheme [1]is an alternative way of constructing the wavelet filters by lifting steps, namely, split, predict, update and scaling . The polyphase matrix of the low-pass and high-pass FIR filter bank can be factorized into the lifting steps.

The 2-D DWT can be decomposed into two steps, namely rDWT and cDWT. The rDWT generates the high-pass (H) and low-pass (L) intermediate results from the input samples and sends them to the cDWT. The cDWT then decomposes the high-pass and low-pass intermediate results into four subbands, namely the high-low(HL) , high-high(HH) , low-low(LL) and low-high (LH)subbands.

The outputsH and L of the rDWT are fed into the cDWT alternately. The formulation of the cDWT can be obtained by substituting Either H or L into x(m,n).

When H is the input to the cDWT, its outputs are HH(m,n) and HL(m,n).When L is the input, the outputs LH(m,n) and LL(m,n) can be similarly obtained.

For the higher level DWTs, the equations are the same as but with the low-low subband generated by the precedinglevel as the input. As the low-low subband is down-sampled in both the column and row direction, its size is only one quarter of the input of its preceding level and consequently, the ranges of and are halved after each level. Here, we use the superscript j to denote the signals in Level DWT.

Despite having several favorable characteristics, the lifting scheme suffers from a long critical path. The flipping method [5] was proposed to shorten the critical path length by flipping [8] the computation nodes with the reciprocalsof the lifting coefficients.

3.2 PROPOSED INPUT DATA SCANNING METHOD

3.2.1 OVERLAPPED STRIPE BASED SCANNING

The temporal memory can be eliminated if the partial results are not stored but produced as and when they are needed. We propose a new overlapped stripe-based scanning method [6], [11] hereto eliminate the temporal memory at the expense of additional arithmetic resources by regenerating the partial results when they are needed. As a result, memory efficient multilevel 2-DDWT architecture is

ISSN 2395-695X (Print)
ISSN 2395-695X (Online)
International Journal of Advanced Research in Biology Engineering Science and Technology (IJARBEST)
Vol. 2, Special Issue 10, March 2016

achieved. The overlapped stripe-based scanning method is first presented, followed by the descriptions of the data input sequencing for Level 2 and Level j DWT.

In the proposed overlapped stripe-based scanning method [6],an image of size N*N is divided into R=N/2S stripes each of width 2S columns and height N rows, with S being the number of parallel processing units in the rDWT that processes pixels concurrently. Fig.3.1 shows three stripes surrounded by thick borders r-1, r, r+1of a partial image, where the gray and white squares representing respectively the overlapped and non-overlapped pixels. The image is scanned into the rDWT stripe by stripe, and row by row within each stripe in a top-down direction as indicated by the arrows. In each cycle, 2S pixels from the current stripe with 7 pixels from the preceding stripe r-1are scanned into the rDWT. For the first stripe, seven padding columns of zeros are used as the overlapped pixels.

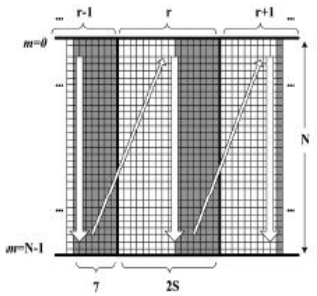


Fig.3.1 overlapped stripes based scanning

This method is developed based on the analysis of the dependency graph of the lifting scheme derived shown in Fig.3.2 where the two rectangular boxes at the top containing the input pixels of the current stripe r=0
and 7 columns of input pixels of the preceding striper=-1. The circle nodes marked with the computation nodes, whereas the dotted oblique boxes contain the parallel processing units, each composing of four computation nodes. Each processing unit has three pixel inputs and three partial result inputs. The shaded triangular box encloses the computation nodes that are needed for generating the three partial results. For clarity, the current stripe shown in Fig. 10 is the stripe r=0 whereas r=-1is its proceeding stripe. The negative stripe number and subscripts denote the paddings of zeros. In general, if the current stripe isr, the column indices m of the input pixels will be offset +2rsby and the column indices of the intermediate results and the partial results will be offset. Unlike the 3 existing stripe-based scanning methods[3], [6], [11], the proposed scanning method takes as input7 additional (overlapped) pixels per row for the computation of the partial results so that no temporal memory is needed to store them in Level 1 DWT. The original stripe-based method [3]does not have overlapped pixels but needs a large memory. The modified stripe-based method has 8 overlapped columns per stripe for the 9/7 filter, resulting in longer computation time. Their method can be used for both single-level convolution-or lifting-based DWT [10] but not for high throughput architecture. Coincidently, the scanning method also has 7 overlapped columns per stripe but the stripe width is fixed at 16 pixels. The corresponding multi-level convolution-based architecture [11] is the most efficient existing design in terms of ADP. However, the scanning method is proposed for the convolution-based design [11] and the resulting architecture consumes more arithmetic and memory resources compared to our proposed design. The

ISSN 2395-695X (Print)
ISSN 2395-695X (Online)
International Journal of Advanced Research in Biology Engineering Science and Technology (IJARBEST)
Vol. 2, Special Issue 10, March 2016

proposed method is extended for multi-level decomposition and applied to a newly designed lifting-based multi-level 2D DWT architecture as described below.

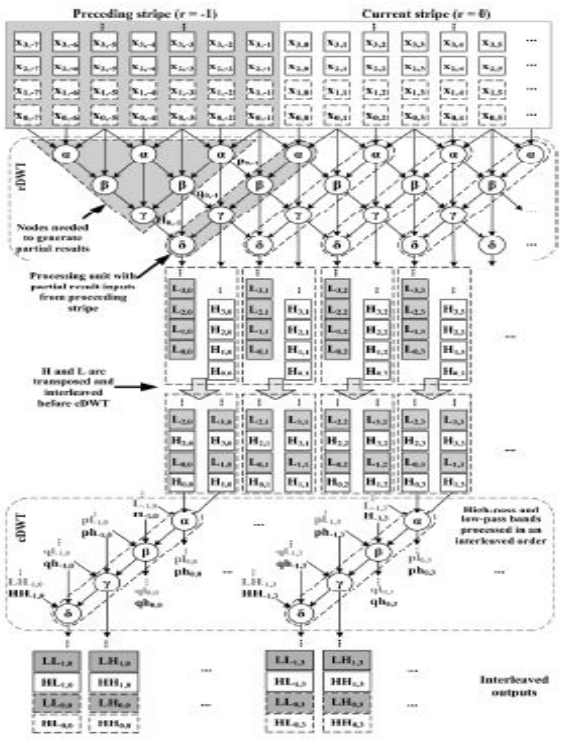


Fig.3.2 dependency graph of 2D lifting DWT

3.2.2 DATA PROCESSING PIPE

The flipped data flow graph (DFG) of lifting scheme [1] can be derived as shown in Fig. 11, where the computation nodes are flipped to eliminate the multiplier on the critical path. Instead of multiplications, only the overflow prevention factors are on the critical path. They are implemented as -

ISSN 2395-695X (Print)
ISSN 2395-695X (Online)
International Journal of Advanced Research in Biology Engineering Science and Technology (IJARBEST)
Vol. 2, Special Issue 10, March 2016

bit right shifts with hard-wired interconnection and do not incur delay on the critical path. In this design, we choose K=1.

The basic operation nodes of the flipped DFG are implemented Cells as shown in the dotted box of Fig.3.3. All the Cells are functionally and structurally identical and each consists of one multiplier, two adders and three -bit right shifters. Its critical path length of consists of the delay of one multiplier and one adder.

The data path that consists of 4 operation nodes in Fig.3.3 is implemented as a Data Processing Pipe (DPP) that comprises four Cells, each with a constant coefficient as shown in Fig.3.4. The DPP is used to realize the parallel processing units. All the DPPs in both the rDWT and cDWT at all the levels are the identical but with different inputs and outputs.

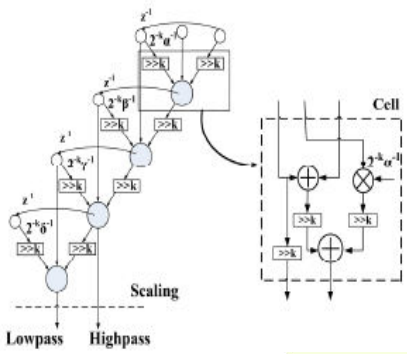


Fig.3.3 lifting scheme

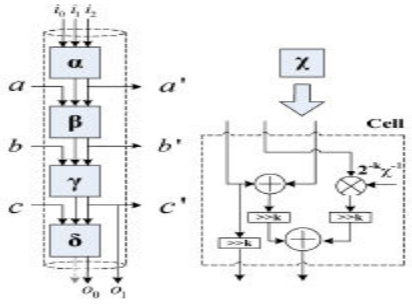


Fig.3.4 DP

3.2.3 DWT ARCHITECTURE FOR LEVEL 1DECOMPOSITION

With the proposed data scanning method for Level 1 decomposition based on the DPP structure proposed above, the architecture for Level 1 decomposition, Arch I, is presented below.

Without loss of generality, let the stripe width of the input image be. The 7 overlapped columns are generated by an external input buffer as shown in Fig.3.5. Hence, pixels are given as input into the rDWT concurrently. The rDWT is realized by a Row-PU (Processing Unit) and an Auxiliary-PU as shown in Fig.3.5 for processing of 7 pixels respectively. The Row-PU in Fig. 3.5(b) consists of parallel DPPs with each DPP consumes 2 pixels every clock cycle. The 3 partial results generated by the last DPP are ignored. The partial result inputs of the first DPP are generated from the 7 overlapped pixels. The Auxiliary-PU in Fig.3.5 (a), derived from the nodes in the triangle of is used to process these 7pixels. It does not contain any DPP but consists of only six Cells. It consumes 7 pixels and produces 3 partial results every clock cycle. These 3 partial results are not needed in the subsequent cycles and not

ISSN 2395-695X (Print)
ISSN 2395-695X (Online)
International Journal of Advanced Research in Biology Engineering Science and Technology (IJARBEST)
Vol. 2, Special Issue 10, March 2016

stored. As a result, a temporal memory of size is saved. Replacing the temporal memory with the Auxiliary-PU leads to significant area reduction as observed from the performance analysis presented.

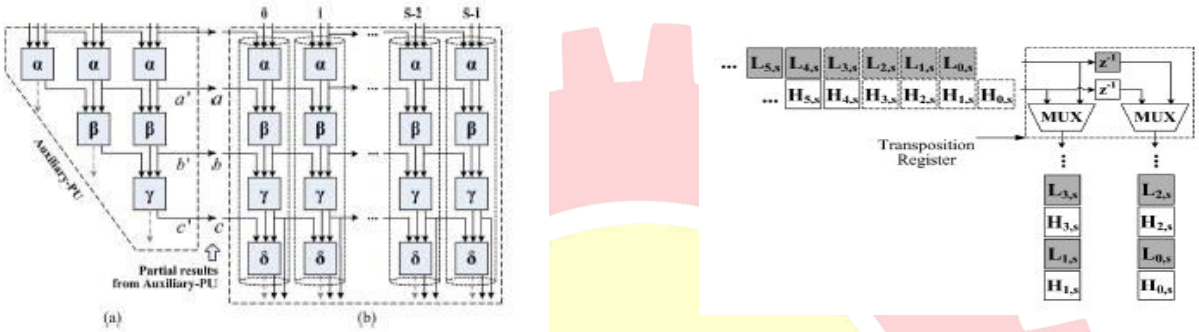


Fig.3.5rDWTa)Auxiliary PUb)Row PU

Fig.3.6 Transposition Register

transposition memory is needed to store and interleave the intermediate results according to the scanning method illustrated in 3.2. Each DPP of the rDWT of Arch I needs one transposition register at its output. The transposition register corresponding to one DPP is depicted in Fig.3.6, where four pairs of and before (in the dotted boxes) and after passing through the transposition registers are shown.

The cDWT is realized with independent DPPs. In every clock cycle, it alternately consumes intermediate results produces a subband pair processed in an inter leaved order, one intermediate result and three partial results generated by each DPP will be consumed by the four Cells only two cycles later, i.e., one extra cycle for interleaving the partial results in addition to the cycle needed in the original lifting scheme [1]. The generated coefficients of the four subband are scaled by the Scaling Units(SUs), of which the structure is shown in Fig.4.8.The structure of Arch I is composed of one rDWT, one cDWT, transposition registers and SUs as shown in Fig.3.8.

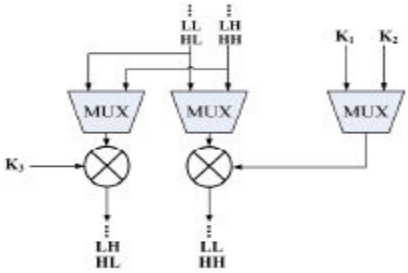


Fig.3.7 scaling unit

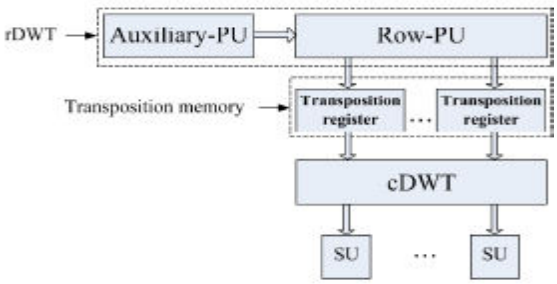


Fig.3.8 Structure of Arch

ISSN 2395-695X (Print)
ISSN 2395-695X (Online)
International Journal of Advanced Research in Biology Engineering Science and Technology (IJARBEST)
Vol. 2, Special Issue 10, March 2016

3.2.4 PROPOSED PIPELINED MULTI LEVEL DWT ARCHITECTURE

The proposed pipelined multi-level DWT [11] architecture is shown in Fig.3.9, where the decomposition at Level is performed by Arch j. The subbands is fed as the input to the succeeding level, while the other subbands and are output directly. Between every pair of processors, there is a Splitter, which realizes the sequencing scheme by splitting the output rows into halves. The structure of Splitter is shown in Fig.3.10. Splitter receives coefficients of every two clock cycles. Assuming the coefficients arriving at Splitter in the first clock cycle.The first half of the inputs is immediately fed as input to Arch through the MUX, while the second half is stored in the Segment register for one cycle. In the second cycle, they are output through the MUX.

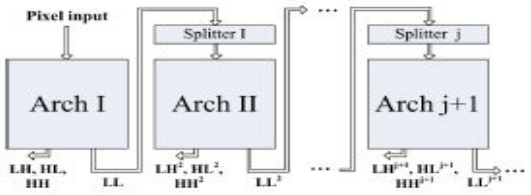


Fig.3.9 Pipelined DWT

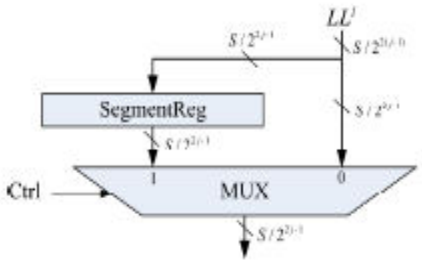


Fig.3.10 Sstructure of splitter j

Table.4.1 Characteristics of Existing and Proposed architectures

Architecture	Lai	Xiong	Tian	Mohanty	Mohanty	Mohanty	This work
DWT	Liftin	Lifting	Liftin	Lifting	Lifting	Convoluti	Lifting

ISSN 2395-695X (Print)
ISSN 2395-695X (Online)
International Journal of Advanced Research in Biology Engineering Science and Technology (IJARBEST)
Vol. 2, Special Issue 10, March 2016

Category	g		g			on	
Data scanning	Line based	Line based	Line based	Line based	Line based	Strip based	Strip based
Multi level	Folded	Folded	Folded	Folded	Pipelined	Pipelined	Pipelined
Frame buffer	Yes	Yes	Yes	Yes	No	No	No
Level1 Temporal Memory	Yes	Yes	Yes	Yes	Yes	No	No
Parallel architecture	No	Yes	Yes	Yes	Yes	Yes	Yes
Scalable Throughput	No	No	Yes	Yes	Yes	Yes	Yes
Flipping method	No	No	No	No	No	No	Yes

5. RESULTS AND DISCUSSION

TABLE 5.1 Result analysis of existing and proposed DWT architectures

SCANNING	SIZE	AREA	MAX FREQ	POWER
CONVOLUTION (EXISTING)	8	2109754	289	72.6
PARALLEL STRUCTURED	8	2702024	240	86.4
LINE BASED (PROPOSED)	8	1195925	285	35.2
CONVOLUTION (EXISTING)	16	3075430	240	93.2
LINE BASED (PROPOSED)	16	1655988	285	42.8

ISSN 2395-695X (Print)
ISSN 2395-695X (Online)
International Journal of Advanced Research in Biology Engineering Science and Technology (IJARBEST)
Vol. 2, Special Issue 10, March 2016

4 CONCLUSIONS

The existing DWT architectures were studied and observed that the area is dominated by the memory
size and the data scanning method has a significant influence on the memory size of the DWT architecture as it decides how the data flows and how the computation is scheduled. It is also observed that the lifting-based architectures generally consume less arithmetic resources than the convolution-based architectures do and the pipelined architectures are more memory efficient than the folded architectures are. Based on the observation a novel overlapped stripe-based scanning method for 3 level decomposition was developed with pipelined lifting-based DWT architecture for high throughput. With the newly proposed scanning method for 3 level decomposition, the basic Cell, the processing units (Data Processing Pipe), the row and column DWTs of each level, the transposition memory and the interface between every two levels (Splitter) is designed. Due to the new scanning method, the size of temporal memory is significantly reduced and the frame buffer is eliminated, resulting in significant saving of memory and consequently the overall size.

REFERENCES

[1] K. Andra, C. Chakrabarti, and T. Acharya, "A VLSI architecture forlifting-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] M.-Y. Chiu, K.-B.Lee, and C.-W. Jen, "Optimal data transfer and buffering schemes for JPEG2000 encoder," in Proc. Signal Process.Syst., 2003, pp.177-182.

[4] C. Chrysafis and A. Ortega, "Line-based, reduced memory, waveletimage 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] C.-T. Huang, P.-C.Tseng, and L.-G. Chen, "Generic ram-based architecturesfor two-dimensional discrete wavelet transform with linebasedmethod," IEEE Trans. Circuits Syst. Video Technol., vol. 15, no.7, pp. 910–920, 2005.

[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. SignalProcess., 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 multilevellifting 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 andmemory-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.