



A SURVEY ON VLSI ARCHITECTURE FOR 2-D DWT USING LIFTING SCHEME

Swapnil Mahto¹, Virendra Singh²

¹M. Tech. Scholar, SIRT Bhopal, ²Professor, SIRT Bhopal

¹swapnil.mahto@gmail.com, ²virendrasingh1180@gmail.com

Abstract

Evaluating the previous work is an important part of developing new hardware efficient methods for the implementation of DWT through Lifting schemes. The aim of this paper is to give a review of VLSI architectures for efficient hardware implementation of wavelet lifting schemes. The inherent in place computation of lifting scheme has many advantages over conventional convolution based DWT. The architectures are represented in terms of parallel filter, row column, folded, flipping and recursive structures. The methods for scanning of images are the line-based and the block-based and their characteristics for the given application are given. The various architectures are analyzed in terms of hardware and timing complexity involved with the given size of input image and required levels of decomposition. This study is useful for deriving an efficient method for improving the speed and hardware complexities of existing architectures and to design a new hardware implementation of multilevel DWT using lifting schemes.

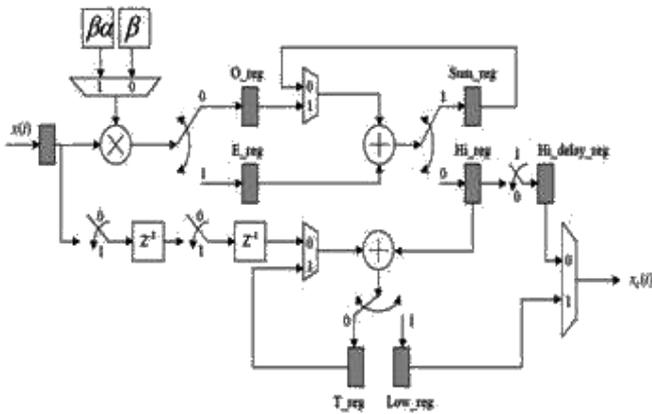
Index Terms: Lifting-based DWT, two-dimensional discrete wavelet transform, JPEG 2000.

I. INTRODUCTION

The compensation of the wavelet transform to conventional transforms, like fourier transform, are recognized fine. Since it is having good locality in time-frequency domain, wavelet transform is broadly used for analysis and compression of the signal. Mallat introduced prospect of its implementation. The discrete wavelet transform (DWT) perform a multi-resolution signal analysis, which has adjustable locality in both the space (time) and frequency domains. The decomposition of signals in to various sub bands with frequency and time information can be possibly by using DWT. Comparing to DCT, image restoration quality and coding efficiency is high for DWT. More over DWT has high compression ratio. So DWT is widely using for image compression and signal processing such as JPEG2000. . By using FIR filters and then sub sampling is the usual implementation method of DWT. A DWT using lifting scheme can be simply implemented due to significantly fewer Computations. This process is fully based on a spatial justification of the wavelet transform.

Moreover, it is having the ability of producing new mother wavelets. DWT implementation on field programmable gate array (FPGA) and DSP chips has been widely developed. The structural processing elements are set successively in the lifting scheme [1].

The DWT in image compression approaches has a property which enables it to overcome the blocking artifact that occurs in DCT-based or block-based image compression techniques. This advantage is due to the DWT acting on the whole image rather than on part of it, as in other block-based algorithms. The JPEG2000 image compression standard is one of the most important applications of the 2-D DWT. The wavelet filters used in JPEG2000 lossy and lossless compression systems are Cohen-Daubechies-Feauveau (9/7) (CDF 9/7) and integer CDF 5/3, respectively. The advantages of the DWT are obvious in many applications; however, the computation complexity and memory requirement are its main drawbacks. These drawbacks have an impact on speed, power consumption and hardware resources. Accordingly, introducing efficient and high speed DWT architectures is still a big and important challenge. Thus, various



architectures for different wavelet filters to elevate all or part of these drawbacks are introduced.[2]

The existing VLSI 2- D DWT architectures can be broadly classified into two main categories, namely convolution-based and lifting-based. While the convolution-based architectures are implemented with FIR filter banks, the lifting-based architectures are implemented by factorizing the filter banks into several lifting steps followed by a scaling step.

Figure1. 1-D DWT architecture for column processor

Both types of architectures perform the 2-D DWT of a 2-D image in two stages, the row-wise DWT (rDWT) followed by the column-wise DWT (cDWT), or vice versa. Both types of architectures are composed of arithmetic resources such as multipliers, adders and multiplexers, and storage resources.

The storage resources include transposition memory, temporal memory and frame memory. Transposition memory is used in the 2-D DWT to transpose the intermediate results produced by the rDWT for the input to the subsequent cDWT. Temporal memory is required for storing the partial results produced in both the rDWT and the cDWT. Frame memory is needed in multi-level DWT, which transforms successively the low-low subband outputs of more than one level, to store the subband coefficients produced at each level for the succeeding level.

Many techniques have been proposed for reducing the memory size. They can be categorized into the line-based, modified line-based, block-based and stripe-based, according to their data scanning methods. The

line-based scanning method was introduced for memory reduction. Since then, many architectures based on the line-based scanning method have been developed. The line-based scanning method scans the image data line by line.

One row of the image is completely processed before its succeeding row is scanned and the data is processed as soon as it is scanned in. But the cDWT is performed in an interleaved manner because it has to wait until sufficient intermediate results are generated by the rDWT. As such, a transposition memory is needed to store the intermediate results of sufficient number of rows for the inputs of the cDWT. In addition, a temporal memory is needed to store the partial results generated by the interleaved cDWT for several rows. Among the line-based designs, achieves the smallest memory size of $55N$ (words), with $25N$ and $3N$ for the transposition and temporal memory [3]

Despite the memory-efficiency advantage of the lifting-based DWT over its convolution-based counterpart, memory requirement is still a major concern in 2-D lifting-based DWT architecture design as it is a size-dominant factor. The memory in 2-D DWT architectures is mainly composed of temporal memory and transposition memory. New parallel stripe-based data scanning method, which enables the tradeoff between the external memory bandwidth and the internal buffer size. We then develop a regular operation unit, termed the Cell, for building a parallel lifting-based 2-D DWT architecture based on the flipped data flow graph (DFG). With the newly developed data scanning method, a novel memory-efficient parallel 2-D DWT architecture with a short CPD of $T_m + T_a$ is proposed.[4].

II. LIFTING SCHEME

Different kinds of lifting-based DWT architectures can be constructed by combining the three basic lifting elements. Most of the applicable DWTs like (9, 7) and (5, 3) wavelets consist of processing units, as shown in Fig.4, which is simplified as Fig.3. This unit is called the processing element (PE). The processing nodes A, B and C are input samples which arrive successively. To implement the predict unit, A and C receive

even samples while B receives odd samples. On the other hand, for the update unit, A and C are odd samples and B receives even samples. Now, the structure can be used to implement (5, 3) and (9, 7) wavelets is shown in Fig.3 & Fig.4. In this architecture each white circle represents a PE.

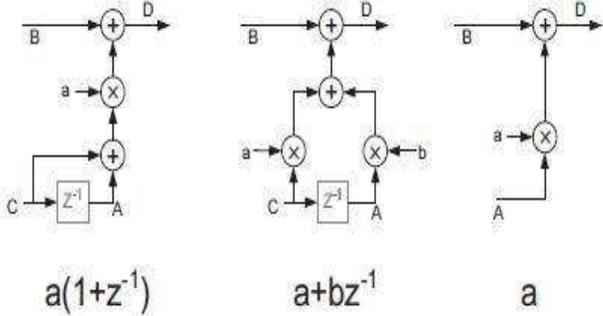


Figure 2. Basic functional units of lifting schemes

The input and output layers are essential (basic) layers and are fixed for each wavelet type, while by changing the number of extended layers, the type of wavelet can be changed accordingly. For example, omission of a single extended (added) layer in the Fig.4 structure will change the related architecture from (9, 7) type to (5, 3) type as in Fig.3. The black circles represent needed stored data for computing outputs (s, d). R0, R1 and R2, are registers that get their values from new input samples and are called data memory. The other three black circles which store the results of previous computations are known as temporary memory.

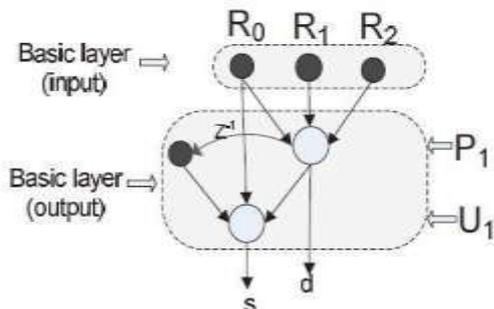


Fig 3.Lifting structure for (5,3) wavelet

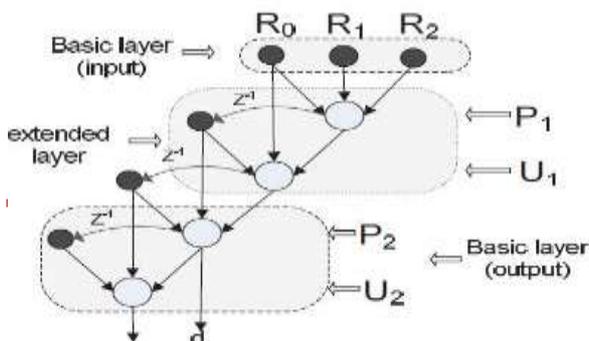


Fig 4.Lifting structure for (9,7) wavelet

III. LITRATURE REVIEW

From adapted algorithm, 2-D DWT pre-processing stage carry out serial-parallel translation of the original sample sequence and then data are given to column processor for doing the operation of column transform.

Then data output of column filter are given to transposing buffer, where transposition of data occurs to meet the dataflow order for the operation of row filter. Finally, scaling computation is done by scaling module. which helps to understand the series of operation concerned in this process.

Every even and odd row of sample is reading alternatively because of parallel scanning method .This way, column transform can be processed by column filter for the sample of neighbouring column alternatively. By adopting the two input/two output structural design, it is possible to decrease the transpose buffer size among column processor and row processor and also improvement in speed of operation. When column filter starts its duty, the input sample getting from pre-processing module, the odd sample $x_i(2n + 1)$ and the even sample $X_i(2n)$ are sending to column filter at the same time in every cycle.. A thorough study is carried out to evaluate the presented structural design with existing architecture. Thus the complexity in hardware, delay in critical path, and throughput of various architectures are compared. From the results, this work achieves better speed with lesser complexity in hardware and lesser storage space.[1]

The 2-D CDF 5/3 DWT architecture consists of two stages.. Each stage consists of

a 1-D DWT processor with different length of delay elements. The input image ($N \times N$ -pixel) is fed to the proposed architecture pixel by pixel using row by row scanning. Each clock cycle single pixel is fed. Thus, in the first stage (Stage1-row processor) a 1-D DWT for each row is computed. This process computes the low (L) and high (H) frequency components of each image row. The required delay element for this stage is 1 register or $R(n)=1$.

The second stage (Stage2) computes the full set of 2-D DWT components of the input image - Low-Low (LL), High-Low (HL), Low-High (LH) and High-High (HH) frequency components. It starts its computation process after N -clock cycle; single image row.

The proposed models are designed to be parameterized to tackle different word length and image sizes. Full analysis of power consumption, speed, hardware utilization and accuracy of the proposed architecture is carried out. The low-complexity of the proposed architecture, due to its construction from identical units, offers an easy way to compose higher DWT dimensions. In addition, the results of the 2-D DWT synthesis reveal that an operating frequency of up to 198 MHz can be achieved with a power consumption of 23 to 131 mWatts for operating frequencies of 25 to 100 MHz, respectively.[2]

The 2-D DWT processor consists of two 1-D DWT processors, namely row processor and column processor, and a transpose unit. According to the scanning scheme the row-wise DWT is performed first. The row processor requires $2N$ temporal memory to store intermediate data d^1 and d^2 . The line buffers used for storage are initialized with all zeros in the reset state, and later they are filled by the temporal data in first input first output (FIFO) manner. The outputs of row processor are fed to transpose unit to change the order of data required by column processor. In the column processor, a line buffer is not needed, the intermediate coefficients can be stored in registers because of the output order of transpose unit. Thus, only $2N$ temporal

memory is required for the complete 2-D DWT operation.

The temporal memory is reduced because of the modified overlapped stripe-based scanning method proposed. The implementation resulted 512 registers as line buffers to process input image of size 256 256. This suggests that the proposed 2-D DWT architecture uses only $2N$ temporal memory, which is lowest among all the other existing architectures and matches with theoretical estimation. The FPGA implementation is done to judge the hardware efficacy of the proposed algorithm for ASIC development.[3]

A less computationally intensive lifting-based DWT has been presented to carry out the biorthogonal wavelet filtering. By factorizing the conventional filter banks into several lifting steps, the computational complexity can be reduced effectively. Moreover, based on the line-based architecture, the memory requirement of lifting-based DWT can also be decreased compared to the convolutional DWT. Although the lifting scheme involves less computation and lower memory, the longer and irregular data path are the major limitations for the efficiency of hardware implementation. In addition, more pipeline registers would increase the internal memory size of 2-D DWT architecture. Several 1-D pipeline architectures have been presented to implement the different lifting step computations. A spatial combinative lifting algorithm (SCLA) to advance the arithmetic efficiency of multiplication for 2-D DWT. Based on the method, the SCLA-based architecture uses fewer multipliers to process the 2-D image data and only uses the on-chip memory up to $12N$ size to perform the multi-level DWT. A systematic design method to construct several efficient architectures of 1-D and 2-D DWT with the systolic array mapping. A general 2-D architecture to implement the various DWT filters proposed in JPEG2000. To perform the computations for different lifting steps, a general hardware scheduler and memory organization are proposed to implement the different factorization matrices. Tseng *et al.* derived a generic RAM-based architecture to optimize the internal memory size for the 2-D DWT

with the line-based method. The recursive and dual scan architectures to implement the 2-D DWT performing the multi-level and single-level decompositions. Based on the asymmetric and symmetric MAC, the two architectures are constructed in an efficient way to carry out the various lifting structures. The flipping structure to shorten the critical path without hardware overhead. With less pipeline registers of the 1-D DWT architecture, the internal memory size of 2-D architecture can also be decreased. Based on the direct implementation of lifting structure and line-based architectures, the critical issue is that using more pipeline registers can

improve the processing speed but requires larger memory size for 2-D DWT. To ease the tradeoff between the pipeline stages of 1-D architecture and memory requirement of 2-D implementation, a modified algorithm is implemented for the designs of 1-D and 2-D pipeline architectures. Based on the modified data path of lifting-based DWT, the architecture achieves the one-multiplier delay constraint but uses less internal memory compared to the related architectures. Moreover, the proposed architecture implements the $5/3$ and $9/7$ filters by cascading the three main component [4]

REVIEW TABLE

SR NO	NAME OF AUTHOR	PUBLISHING YEAR	WORK DONE	RESULT
1	MithunR, Ganapathi Hegde	2015	Reduced area and high speed 2-D DWT structural design	Better speed with lesser complexity in hardware and lesser storage space.
2	Saad Al-Azawi, Yasir Amer Abbas and Razali	2014	Low Complexity Multidimensional CDF $5/3$ DWT Architecture	The proposed models are designed to be parameterised to tackle different wordlength and image sizes.
3	Yusong Huand Ching Chuen Jong,,	Oct 2013	A Memory-Efficient High-Throughput Architecture for Lifting-Based Multi-Level 2-D DWT	Proposed a novel overlapped stripe-based scanning method for the multi-level decomposition and developed a scalable pipelined lifting-based DWT architecture for high throughput.
4	Yusong Hu and Ching Chuen Jong	August 2013	A Memory-Efficient Scalable Architecture for Lifting-Based Discrete Wavelet Transform	A new stripe-based scanning method has been proposed, enabling the tradeoff between the external input bandwidth and the internal buffer size
5	A D Darji, Ankur Limaye	2014	Memory efficient VLSI Architecture for Lifting-based DWT	The temporal memory is reduced because of the modified overlapped stripe-based scanning method

				proposed.
6	Yusong Huand Viktor K. Prasanna	2013	Energy- and Area-Efficient Parameterized Lifting- Based 2-D DWT Architecture on FPGA	Proposed architecture achieves high energy and area efficiency by introducing an overlapped block-based image scanning method which optimizes the number of external memory reads and the on-chip memory size.

V. CONCLUSION

The Discrete Wavelet Transform provides a multi resolution representation of signals. The transform may be implemented using filter banks. This work may presents the simulation work for column processor, transposing buffer and row processor of 2D DWT architecture for JPEG 2000 and the study of high-performance and low-memory pipeline architecture for 2-D lifting-based DWT of the 5/3 and 9/7 filters. By merging the predictor and updater into one single step, we can derive efficient pipeline architecture. The study may provide the same number of arithmetic units, the architecture may have shorter pipeline data path. In this paper, and architectures for the Lifting based Discrete Wavelet Transform have been studied. For each of them, parameters such as memory requirement and speed were discussed. Based on the application and the constraints imposed, the appropriate architecture can be chosen.

REFERENCES

- [1] "High Performance VLSI Architecture for 2-D DWT Using Lifting Scheme" Mithun R, Ganapathi Hegde, IEEE, 2015
- [2] "Low Complexity Multidimensional CDF 5/3 DWT Architecture" Saad Al-Azawi, Yasir Amer Abbas and Razali Jidin, IEEE, 2014
- [3] "A Memory-Efficient High-Throughput Architecture for Lifting-Based Multi-Level 2-D DWT" Yusong Hu, Student Member, IEEE, and ChingChuen Jong, Member, IEEE, 15 OCT 2013
- [4] "A Memory-Efficient Scalable Architecture for Lifting-Based Discrete Wavelet Transform" Yusong Hu and ChingChuen Jong, IEEE, AUGUST 2013
- [5] "Memory efficient VLSI Architecture for Lifting-based DWT" A D Darji, Ankur Limaye, IEEE, 2014
- [6] "Energy- and Area-Efficient Parameterized Lifting-Based 2-D DWT Architecture on FPGA" Yusong Huand Viktor K. Prasanna, IEEE,
- [7] X.Lan, N. Zheng and Y.Liu, "Low-power and high-speed VLSI architecture for lifting-based forward and inverse wavelet transform", *IEEE Transactions on Consumer Electronics*, Vol. 51, No. 2, pp. 379- 385, July 2005.
- [8] C. Christopoulos, A. Skodras and Ebrahimi, "The JPEG2000 still image coding system: an overview", *IEEE Trans. On Consumer Electronics*, Vol. 4, No. 4, pp.1103-1127, July 2000
- [9] T. Acharya and P. Tsai, "JPEG2000 Standard for Image Compression Concepts, Algorithms and VLSI Architectures", *Wiley Inter science—A John Wiley & Sons*, June 2005.
- [10] K.K. Parhi and T. Nishitani, "VLSI architectures for discrete wavelet transforms," *IEEE Trans. on VLSI Syst.*, vol.1, pp.191- 202, June 1993.
- [11] Andra, K., Chakrabarti, C., and Acharya, T.: "A high performance JPEG 2000 architecture", *IEEE Trans. on Circuits Syst. for Video Technol.*, vol. 8, No. 9, pp. 209-218, June 2003.
- [12] I. Daubechies and W. Sweldens, "Factoring wavelet transform into lifting steps," *The J. of Fourier Analysis and Applications*, vol.4,

- pp.247-269, April 1998.
- [13] G. K. Wallace, "The JPEG Still Picture Compression Standard", *IEEE Trans. On Consumer Electronics*, Vol. 38, No 1, Feb. 1992.
- [14] W. Sweldens, "The new philosophy in bi orthogonal wavelet constructions," in Proc. *SPIE*, vol.2569, pp.68-79, 1995.
- [15] S.-C. B. Lo, H. Li and M.T. Freedman, "Optimization of wavelet decomposition for image compression and feature preservation," *IEEE Trans. on MedicalImaging*, vol.22, pp.1141-1151, September 2003.
- [16] H. Liao, M. Kr. Mandal and B.F. Cockburn, "Efficient architectures for 1-D and 2-D lifting-based wavelet transforms," *IEEE Trans.on Signal Processing*, vol. 52, no. 5, pp. 1315-1326, May 2004.