RECONFIGURABLE DISCRETE WAVELET TRANSFORM ARCHITECTURE FOR ADVANCED MULTIMEDIA SYSTEMS

Po-Chih Tseng, Chao-Tsung Huang, and Liang-Gee Chen

DSP/IC Design Lab, Graduate Institute of Electronics Engineering and
Department of Electrical Engineering, National Taiwan University,
Taipei, Taiwan, R.O.C.
E-Mail:{pctseng, cthuang, lgchen}@video.ee.ntu.edu.tw

ABSTRACT

In this paper, a novel reconfigurable discrete wavelet transform architecture is proposed to meet the diverse computing requirements of advanced multimedia systems. The proposed architecture mainly consists of reconfigurable processing element array and reconfigurable address generator, featuring dynamically reconfigurable capability where the wavelet filter kernels and wavelet decomposition structures can be reconfigured at run-time with little overhead. The lifting-based reconfigurable processing element array possesses better computational efficiency than convolution-based architecture, and a systematic design method is provided to generate the hardware configurations of different wavelet filter kernels for it. The reconfigurable address generator handles flexible address generation for data I/O access in different wavelet decomposition structures. A prototyping chip has been fabricated by TSMC 0.35μm 1P4M CMOS process, and at 50MHz, it can achieve at most 100M pixel/sec transform throughput, proving it to be a universal and extremely flexible computing engine for advanced multimedia systems.

1. INTRODUCTION

Discrete Wavelet Transform (DWT) [1] has been widely used in many multimedia applications, including multimedia coding and signal processing. Recently, emerging multimedia standards such as JPEG2000 still image coding and MPEG-4 still texture coding have also adopted DWT as its transform coder. The computations of DWT can be divided into two parts, one is the wavelet filter operation which performs the signal analysis and subsampling, and the other is the wavelet decomposition operation which recursively decomposes the signal according to specific decomposition structure. These two computational parts flexibly combine to make DWT able to decompose a signal into different subbands of well-defined time-frequency characteristics.

In the advanced multimedia systems, such as portable multimedia everything device or home entertainment center, the computing requirements must be quite diverse. Many kernel tools such as DWT, Motion Estimation (ME), and Discrete Cosine Transform (DCT) should be integrated into the systems with flexible functionality to support rich multimedia applications. For instance, a universal and extremely flexible DWT computing engine which can support various wavelet filter kernels and wavelet decomposition structures would become a necessary for advanced multimedia systems.

In the literature, there have been many proposals devoted to the hardware architecture of DWT. However, these proposals usually based on fixed wavelet filter kernel and/or fixed wavelet decomposition structure. There are no flexible enough architectures existing to meet the diverse computing requirements of advanced multimedia systems. This situation attracts us to have the research motivation to investigate a reconfigurable discrete wavelet transform architecture which can be dynamically reconfigured as various wavelet filter kernels and wavelet decomposition structures. In the following of this paper, the proposed architecture is overviewed in section 2. In section 3, the reconfigurable DWT processing element array is presented, and then the reconfigurable address generator is detailed in section 4. The chip implementation results are given in section 5, and finally, a brief summary in section 6 concludes this paper.

2. RECONFIGURABLE DWT ARCHITECTURE

2.1. Discrete Wavelet Transform

As mentioned in section 1, one of the computational parts of DWT is the wavelet filter operation, which is a two channel filter bank as shown in Fig. 1 and Fig. 2,

0-7803-7795-8/03/ $17.00 © 2003 IEEE

where Fig. 1 represents the DWT analysis filter bank and Fig. 2 represents the DWT synthesis filter bank.

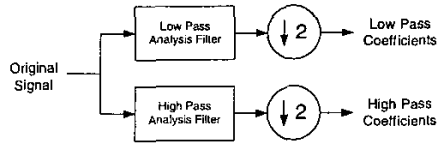


Fig. 1. DWT analysis filter bank

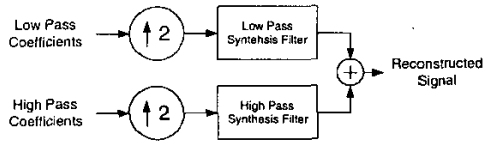


Fig. 2. DWT synthesis filter bank

In the DWT analysis, original signal is processed first by two analysis filters, low pass and high pass, and then followed by subsampling to decompose the low pass and high pass coefficients. In the DWT synthesis, low pass and high pass coefficients are processed first by upsampling and then followed by two synthesis filters to reconstruct the signal. This basic operation is called the one-level DWT decomposition (reconstruction). For multi-resolution analysis (synthesis), multi-level DWT decomposition (reconstruction) is performed.

The multi-level DWT decomposition, which is namely the other one of the computational parts of DWT, is very flexible, and according to the original signal characteristic, a specific wavelet decomposition structure can be performed to achieve best-suited multi-resolution analysis result. Among all possible decomposition structures, the dyadic type decomposition as shown in Fig. 3 is the most common case due to its regular and recursive structure. In the dyadic type decomposition, the output low pass coefficients of previous level are treated as current input signal to form a recursive chain. However, beyond the dyadic type decomposition, many other decomposition structures are possible but may be more irregular. Take the 2-D image signal as examples, Fig. 4 shows the 3-level dyadic type decomposition of test image *Lena*, and Fig. 5 shows the wavelet packet transform of test image *Barbara*, where the DWT is performed according to image characteristics and special consideration with specific wavelet filter kernel and wavelet decomposition structure to achieve best coding efficiency.

2.2. Proposed Reconfigurable DWT Architecture

In order to support various wavelet filter kernels and wavelet decomposition structures in single architecture, a dynamically reconfigurable DWT architecture is proposed as shown in Fig. 6.

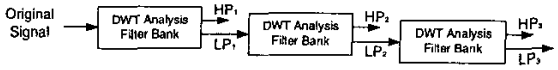
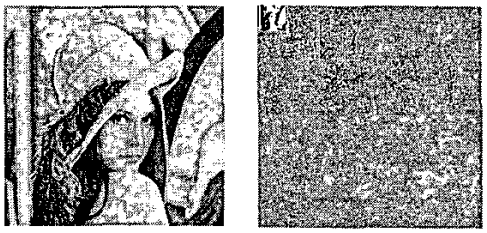
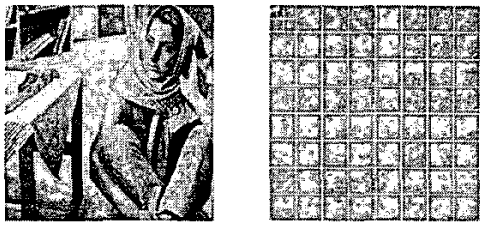


Fig. 3. Dyadic type decomposition



(a) (b)

Fig. 4. 3-level dyadic type decomposition of Lena



(a) (b)

Fig. 5. Wavelet packet transform of Barbara

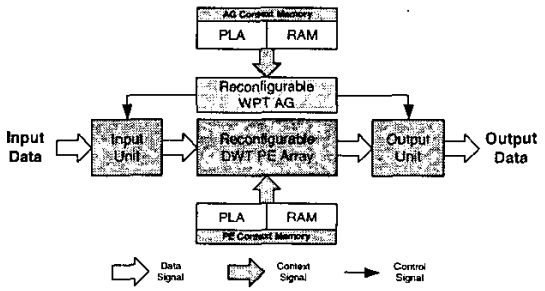


Fig. 6. Proposed reconfigurable DWT architecture

The proposed architecture is a general and scalable computational model, and the computational resources inside it can be flexibly scalable according to target application specification. A virtual external frame memory is required to buffer the data signal under processing, and the Input Unit and Output Unit depicted in Fig. 6 act as the interface between reconfigurable architecture and this frame memory. In a multimedia system-on-chip (SOC), this virtual external frame memory can be implemented by a shared system memory or by a local frame memory tightly-attached to the reconfigurable architecture. In addition to the I/O Units, the proposed architecture mainly consists of two functional blocks. One is the reconfigurable processing element array, and the other is the reconfigurable address generator. The reconfigurable processing

element array, depicted as Reconfigurable DWT PE Array in Fig. 6, is responsible for the wavelet filter operation and composed of a 1-D linear array of reconfigurable DWT processing elements (PE). The reconfigurable DWT PE is based on the more computationally efficient lifting scheme rather than conventional convolution approach. Besides, a systematic design method is exploited to derive the reconfigurable DWT PE architecture and generate the corresponding the hardware configurations of different wavelet filter kernels for it. The hardware configurations of Reconfigurable DWT PE Array are stored in the PE Context Memory, where the PLA part stores several default configurations and the RAM part stores user-programmable configurations.

The reconfigurable address generator, depicted as Reconfigurable WPT AG in Fig. 6, is responsible for the wavelet decomposition operation. By generating specific memory read/write address to I/O Units, flexible data access between external frame memory and I/O Units is performed for different wavelet decomposition structures. The hardware configurations of Reconfigurable WPT AG are stored in the AG Context Memory, and the PLA and RAM have the same features as those in the PE Context Memory.

3. RECONFIGURABLE DWT PE ARRAY

Before detailing the reconfigurable architecture, the lifting scheme and a systematic design method to derive efficient hardware architecture of 1-D lifting-based DWT are discussed first in subsection 3.1 and 3.2, respectively.

3.1. Lifting Scheme

The lifting scheme is a new method for constructing wavelets entirely by spatial approach [2]. Using lifting scheme to construct wavelets has many advantages, such as allowing a faster and fully in-place implementation of the wavelet transforms, immediately to find the inverse transform, easily to manage the boundary extension, and possibly of defining a wavelet-like transform that maps integer-to-integer. According to [3], any DWT with finite filter can be decomposed into a finite sequence of simple filtering steps, which is called the lifting steps. This decomposition corresponds to a factorization of the polyphase matrix of target wavelet filter into a sequence of alternating upper and lower triangular matrices and a constant diagonal matrix as below.

$$P(z)=\prod_{i=1}^{m}\begin{bmatrix}1 & s_i(z)\\ 0 & 1\end{bmatrix}\begin{bmatrix}1 & 0\\ t_i(z) & 1\end{bmatrix}\begin{bmatrix}K & 0\\ 0 & 1/K\end{bmatrix}$$

3.2. Systematic Design Method

In [4], a systematic design method to derive hardware architecture of 1-D lifting-based DWT is presented. By this systematic design method, an efficient 1-D lifting-based DWT architecture based on systolic array can be easily constructed. As shown in Fig. 7, this design method consists of several design stages: specific lifting factorization, dependence graph formation, systolic array mapping, and optionally pipelining. Once a finite DWT filter is chosen, four subsequent design stages can then be performed to construct the corresponding hardware architecture. The hardware architecture constructed by this design method consists of several serially-connected basic computing units. The possible three structures of the basic computing unit are shown in Fig. 8, and the number of basic computing units for a chosen DWT filter depends on the number of lifting steps after specific lifting factorization. For instance, there are four basic computing units in the (9,7) odd symmetric biorthogonal filter and three basic computing units in the (9,3) odd symmetric biorthogonal filter.

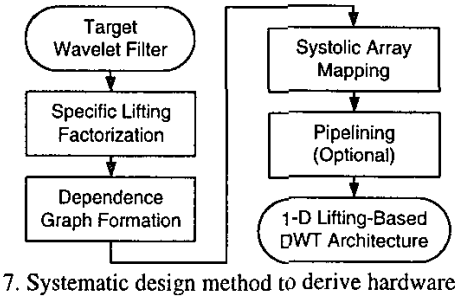


Fig. 7. Systematic design method to derive hardware architectures of 1-D lifting-based DWT

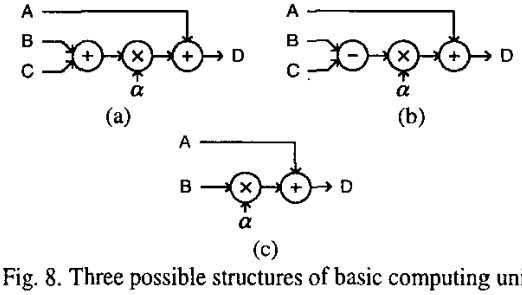


Fig. 8. Three possible structures of basic computing unit

3.3. Architecture of Reconfigurable DWT PE Array

According to the possible structures of basic computing unit in previous subsection, the core cell, which is called the MCU, of reconfigurable DWT PE is derived as shown in Fig. 9. This core cell is a three-input (A, B, C) one-output (D) datapath, consisting of one adder/subtractor, one multiplier with coefficient a, and another adder. The datapath can be dynamically reconfigured as one of the three possible structures of basic computing unit.

The Reconfigurable DWT PE Array is composed of a 1-D linear array of several reconfigurable DWT PE, and the number of the PE is scalable according to target application specification. As mentioned in subsection 3.2, since the number of basic computing units is variable for different DWT filter, a folding of systolic array technique can be exploited to fold variable number of basic computing units into equal number of MCU with variable throughout. For instance, the (9,7) filter originally require four basic computing units, after a fold by 2 operation, the required MCU number becomes two while the throughput becomes one half. The folding technique will induce feedback loop from the output to the input, therefore some feedback registers are necessary to buffer the feedback signal. Together with the lifting registers and pipeline registers between each MCU, the reconfigurable DWT PE architecture is derived as shown in Fig. 10. In Fig.10, the delay chain 0 contains feedback registers, the delay chain 1 and 2 contain lifting registers and pipeline registers, the MCU represents the core cell in Fig. 9, the Mux selects suitable input data from three delay chains, and the FSM receives configuration signal from PE Context Memory to decode necessary hardware configurations for MCU and Mux. Due to the regularity and modularity of reconfigurable DWT PE architecture, several PE can be cascaded serially to form a 1-D linear array as the Reconfigurable DWT PE Array.

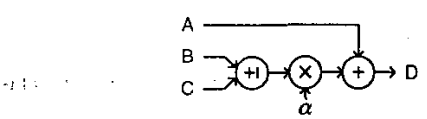


Fig. 9. Core cell of reconfigurable DWT PE (MCU)

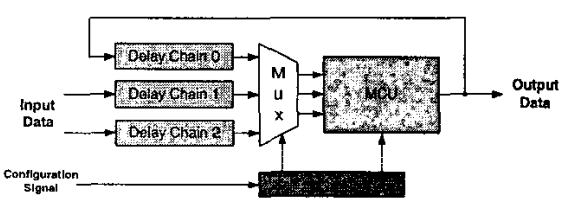


Fig. 10. Reconfigurable DWT PE architecture

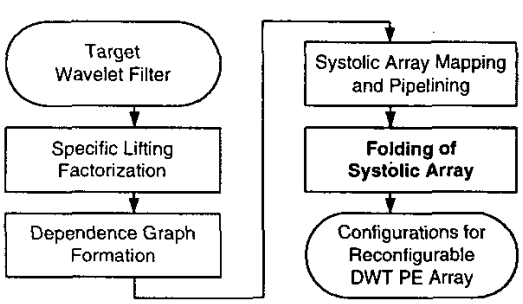


Fig. 11. Modified systematic design method to generate the hardware configurations for the Reconfigurable DWT PE Array

By adding an additional design stages, folding of systolic array, into original systematic design method, a modified systematic design method to generate the hardware configurations for the Reconfigurable DWT PE Array is shown in Fig. 11. By this design method, any finite DWT filter can be mapped onto the Reconfigurable DWT PE Array with specific PE number through the generated hardware configurations.

4. RECONFIGURABLE WPT AG

Compared to the architecture of Reconfigurable DWT PE Array, the architecture of Reconfigurable WPT AG is much simple and straightforward. As shown in Fig. 12, there two address generators in the architecture, one is the output address generator which generates the corresponding row or column address to Output Unit as write address to external frame memory, and the other is the input address generator which generates the corresponding row or column address to Input Unit as read address to external frame memory. The start time slot of four FSMs, the initial value of four counters, and the select signal of two Muxs are controlled by the configuration signal from AG Context Memory for specific wavelet packet transform.

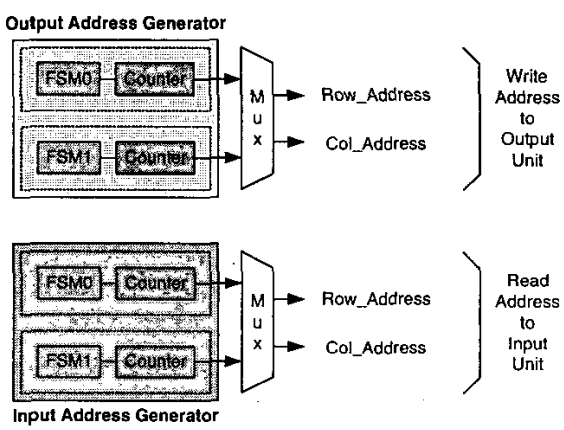


Fig. 12. Architecture of Reconfigurable WPT AG

5. CHIP IMPLEMENTATION RESULTS

In order to prove the feasibility of proposed reconfigurable architecture, a prototyping chip has been fabricated by TSMC 0.35μm 1P4M CMOS process. Two reconfigurable DWT PE are adopted to form the Reconfigurable DWT PE Array, and several useful wavelet filter kernels and wavelet decomposition structures are stored in the PLA for default configurations. The performance of this prototype

architecture is listed in Table I, including the wavelet filter kernels, number of lifting steps, throughput per clock cycle, and corresponding hardware utilization. At 50MHz, the prototyping chip can achieve at most 100M pixel/sec transform throughput (for (5,3) filter), which is equal to perform the CCIR 601 (720x576) format image signal at 30 frame/sec with four-level wavelet packet transform. The photograph of prototyping chip is shown in Fig. 13.

Table I. Performance of the prototype architecture

Wavelet Filter	Lifting Steps	Throughput (per cycle)	HW Utilitization
(5,3)	2	2	100%
(9,3)	3	1	75%
(9,7)	4	1	100%
(2,10)	4	1	100%
(13,7)	4	1	100%

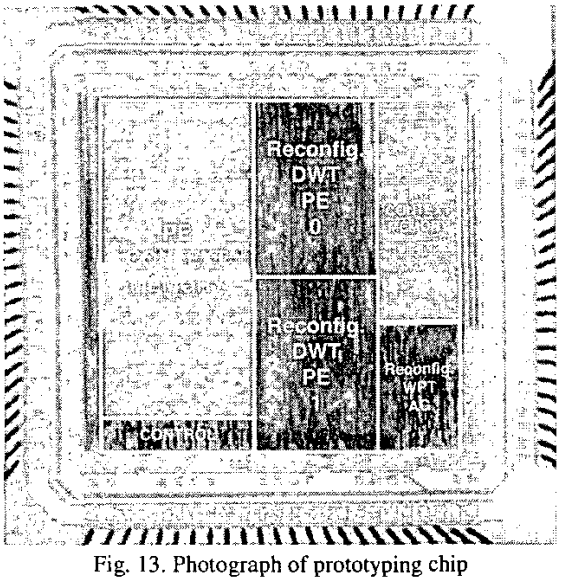


Fig. 13. Photograph of prototyping chip

6. CONCLUSION

We have proposed a novel reconfigurable DWT architecture to meet the diverse computing requirements of advanced multimedia systems. The proposed architecture is dynamically reconfigurable in terms of the wavelet filter kernels and wavelet decomposition structures. The lifting-based Reconfigurable DWT PE Array possesses better computational efficiency than convolution-based architecture, and a systematic design method is provided to generate the hardware configurations of different wavelet filter kernels for it. The Reconfigurable WPT AG handles flexible address generation for data I/O access in different wavelet decomposition structures. A prototyping chip has been fabricated with high performance and proved the proposed architecture to be a universal and extremely flexible computing engine for advanced multimedia systems.

7. REFERENCES

[1] S. Mallat, "A theory for multiresolution signal decompostion: The wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, July 1989.

[2] W. Sweldens, "The lifting scheme: A custom-design construction of biorthogonal wavelets," *Applied and Computational Harmonic Analysis 3*, pp. 186-200, 1996.

[3] I. Daubechies and W. Sweldens, "Factoring wavelet transforms into lifting schemes," *The Journal of Fourier Analysis and Applications*, vol. 4, pp 247-269, no. 3, 1998.

[4] C. T. Huang, P. C. Tseng, and L. G. Chen, "Efficient VLSI Architectures of Lifting-Based Discrete Wavelet Transform by Systematic Design Method," *IEEE International Symposium on Circuits and Systems*, vol. 5, pp 565-568, 2002.