



Reconfigurable Discrete Wavelet Transform Processor for Heterogeneous Reconfigurable Multimedia Systems

PO-CHIH TSENG,* CHAO-TSUNG HUANG AND LIANG-GEE CHEN

*DSP/IC Design Lab, Graduate Institute of Electronics Engineering, Department of Electrical Engineering,
National Taiwan University, Taipei, Taiwan*

Received July 2003; Revised January 2004; Accepted March 2004

Abstract. In this paper, a novel reconfigurable discrete wavelet transform processor architecture is proposed to meet the diverse computing requirements of future generation multimedia SoC. The proposed architecture mainly consists of reconfigurable processing element array and reconfigurable address generator, featuring dynamically reconfigurable capability where the wavelet filters and wavelet decomposition structures can be reconfigured as desired at run-time. The lifting-based reconfigurable processing element array possesses better computation efficiency than convolution-based architectures, and a systematic design method is provided to generate the hardware configurations of different wavelet filters 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 0.35 μm 1P4M CMOS process. At 50 MHz, this chip can achieve at most 100 M pixels/sec transform throughput, together with energy efficiency and unique reconfigurability features, proving it to be a universal and extremely flexible computing engine for heterogeneous reconfigurable multimedia systems.

Keywords: discrete wavelet transform, lifting scheme, reconfigurable computing, heterogeneous reconfigurable multimedia systems, energy efficiency

1. Introduction

As the era of system-on-a-chip (SoC) coming, a wide range of complex functions can be combined on a single die. SoC designs that integrate embedded microprocessors, digital signal processors, embedded memory, and custom modules have been reported by a number of industrial companies and academic organizations in the past decade. Projections of future integration densities suggest that this trend will surely continue in the next decade. It is therefore reasonable to expect that a future generation multimedia SoC will combine all the functionality of a portable multimedia terminal, including not only the traditional computational functions and

operating system, but also the extensions for full multimedia support such as image, video, graphics, and audio. In order to meet the diverse computing requirements of future generation multimedia SoC, sufficient functional flexibility should be provided by the system implementation platform. Although the traditional programmable platforms such as embedded microprocessors and digital signal processors can provide ultimate flexibility, these solutions always experience the problems of computation and energy inefficiencies. In the literature, many researches addressed these problems and proposed to adopt the reconfigurable computing technology into the implementation platform for multimedia SoC. Among these proposals, the heterogeneous reconfigurable system proposed by Rabaey et al. [1] is likely to be the most promising one to meet the diverse computing requirements while achieving high computation and energy efficiencies.

*Present address: Department of Electrical Engineering, National Taiwan University, 1, Sec. 4, Roosevelt Rd., Room 332, Taipei 106, Taiwan.

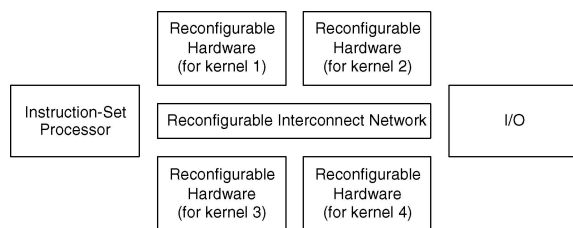


Figure 1. Conceptual view of the heterogeneous reconfigurable system.

1.1. Heterogeneous Reconfigurable Multimedia Systems

Figure 1 shows the conceptual view of the heterogeneous reconfigurable system, including the instruction-set processor for control-oriented tasks, the reconfigurable interconnect network for data communication tasks, several reconfigurable hardware modules for dominating computational kernels, and the I/O data interface. The principle of the heterogeneous reconfigurable system is to provide programmability or reconfigurability at just the right granularity so as to eliminate virtually all reconfiguration overhead. Since the digital signal processing applications typically have a few dominating computational kernels with high regularity, these regular computational kernels can be mapped to several function-specific reconfigurable hardware modules with minimum reconfiguration overhead. However, the mapping of computational kernel to function-specific reconfigurable hardware is not a trivial work, and the mapping results directly affect the overall system performance.

The mapping process mainly consists of three design steps. The first step is to identify the possible reconfigurable parameters of the computational kernel. The identification can minimize the hardware reconfigurability to a constrained set and reduce the reconfiguration overhead to the minimum. The second step is to investigate the algorithm with highly regular structure for the computational kernel. Regularity allows the algorithm to be decomposed into architecture patterns of computation, memory access, and interconnection. The third step is to develop the corresponding function-specific reconfigurable hardware architecture for the architecture patterns in second step to be mapped onto. The reconfigurable architecture is a combination of datapath units, specially partitioned and accessed memory blocks connected by dedicated links. The architecture should be modular and scalable in order to allow easily mapping of architecture patterns.

In most multimedia applications, three of the dominating computational kernels are discrete wavelet transform, motion estimation, and discrete cosine transform. A heterogeneous reconfigurable multimedia system consisting of these three function-specific reconfigurable processors is therefore capable to perform almost all the functionality of a portable multimedia terminal with high computation and energy efficiencies. In this paper, we focus on one of the computational kernels—the discrete wavelet transform.

1.2. Discrete Wavelet Transform

During the past decade, wavelets have been developed as an effective multiresolution signal analysis tool. Since the discrete wavelet transform (DWT) deduced by Mallat [2], many researches on wavelet-based image analysis and compression have derived fruitful results. Recently, emerging multimedia standards such as JPEG2000 still image coding [3] and MPEG-4 visual texture coding [4] have also adopted DWT as their transform coders. 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 enable DWT to decompose a signal into different subbands of well-defined time-frequency characteristics. Hence, it is clear to identify that two reconfigurable parameters of the DWT computational kernel are variable wavelet filters and variable wavelet decomposition structures. A reconfigurable DWT processor in heterogeneous reconfigurable multimedia systems is therefore supposed to sufficiently provide these two reconfigurable parameters in order to support the flexible functionality required by future generation multimedia SoC.

In the literature, there have been many proposals devoted to the hardware architecture of DWT [5–10]. Most of the proposals based on fixed wavelet filter and fixed wavelet decomposition structure. Some recent proposals [11–16] addressed the importance of flexibility and proposed programmable or reconfigurable DWT architectures for either variable wavelet filters [11–14] or variable wavelet decomposition structures [15, 16]. However, these proposals are still not flexible enough to meet the diverse computing requirements of future generation multimedia SoC. This situation attracts us to have the research motivation to

investigate a reconfigurable DWT processor which can be dynamically reconfigured as desired wavelet filter and wavelet decomposition structure, being a universal and extremely flexible DWT computing engine for heterogeneous reconfigurable multimedia systems.

1.3. Paper Organization

This paper is organized as follows. In Section 2, the preliminaries of DWT are first illustrated. The lifting scheme for DWT is then reviewed in the following. In Section 3, a systematic design method based on lifting scheme is proposed to derive the DWT algorithm with highly regular structure. The derived algorithm is also decomposed into architecture patterns of computation, memory access, and interconnection by proposed method. Two case studies are illustrated to show the effectiveness of proposed method. In Section 4, the proposed reconfigurable DWT processor architecture is described in detail, including the reconfigurable processing element array and the reconfigurable address generator. The chip implementation and architecture evaluation results are given in Section 5 to show the energy efficiency and architectural uniqueness of proposed reconfigurable DWT processor. Finally, a brief summary in Section 6 concludes this paper.

2. Discrete Wavelet Transform and Lifting Scheme

2.1. Preliminaries of 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 Figs. 2 and 3, where Fig. 2 represents the DWT analysis filter bank and Fig. 3 represents the 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

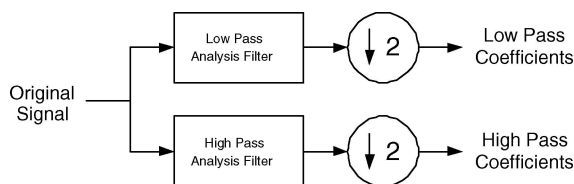


Figure 2. DWT analysis filter bank.

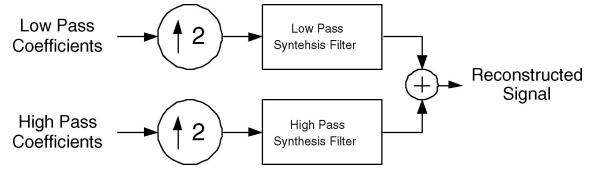


Figure 3. DWT synthesis filter bank.

and high pass coefficients. In the DWT synthesis, low pass and high pass coefficients are processed first by up-sampling 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 characteristics, 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. 4 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. 5 shows the 3-level dyadic type decomposition of test image Lena, and Fig. 6 shows the wavelet packet transform (WPT) of test image Barbara, where the DWT is performed according to image characteristics and special consideration with specific wavelet filter and wavelet decomposition structure to achieve best coding efficiency [17].

2.2. Lifting Scheme

The wavelet filter operation is a two channel filter bank, and this operation is conventionally implemented by convolution-based method. However, the convolution-based implementation method is computation-intensive when wavelet filter tap is long. Thanks to the appearance of lifting scheme [18] and a factorization method that factors wavelet transforms into lifting steps [19], the lifting scheme is widely used to speed up the DWT wavelet filter operation.

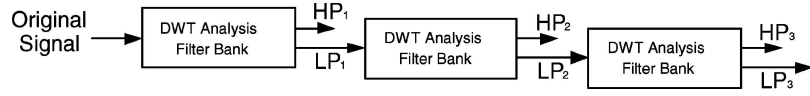


Figure 4. Dyadic type decomposition.

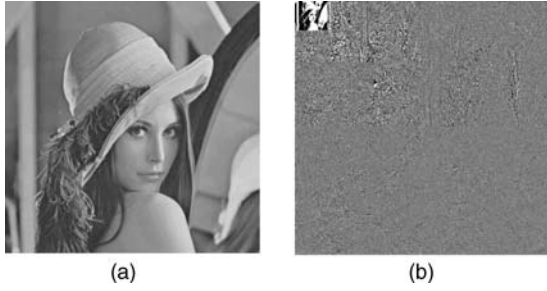


Figure 5. 3-level dyadic type decomposition of Lena.



Figure 6. Wavelet packet transform of Barbara.

The lifting scheme is a new method for constructing wavelets entirely by spatial approach [18]. 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 [19], 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. Figure 7 shows the generic block diagram of a wavelet filter. The forward transform uses two analysis filters \tilde{h} (low pass) and \tilde{g} (high pass) followed by subsampling, while the inverse transform first performs upsampling and then uses two synthesis filters h (low pass) and g (high pass).

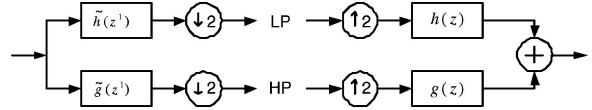


Figure 7. Generic block diagram of a wavelet filter.

Since the polyphase representation of a filter h is

$$h(z) = h_e(z^2) + z^{-1}h_o(z^2) \quad (1)$$

where h_e denotes the even coefficients and h_o denotes the odd coefficients. The polyphase matrix of a wavelet filter can be assembled as

$$P(z) = \begin{bmatrix} h_e(z) & g_e(z) \\ h_o(z) & g_o(z) \end{bmatrix} \quad (2)$$

In [19], it has been shown that if h and g is a complementary filter pair, then with the exploitation of Euclidean algorithm for Laurent polynomials, there always exist Laurent polynomials $s_i(z)$, $t_i(z)$ and a non-zero constant K so that

$$P(z) = \prod_{i=1}^m \begin{bmatrix} 1 & s_i(z) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ t_i(z) & 1 \end{bmatrix} \begin{bmatrix} K & 0 \\ 0 & 1/K \end{bmatrix} \quad (3)$$

In other words, any finite wavelet filter can be obtained by starting with the Lazy wavelet followed by several lifting steps with a scaling. Due to the exploitation of Euclidean algorithm for Laurent polynomials in this lifting factorization method, the factorization process is non-unique. That is, there exist many essentially different lifting factorizations, but which one more suitable for software and/or hardware implementations is still an open design issue.

3. Proposed Systematic Design Method

In this section, a systematic design method based on lifting scheme is proposed to derive the DWT algorithm with highly regular structure. The derived algorithm is also decomposed into architecture patterns of computation, memory access, and interconnection by proposed

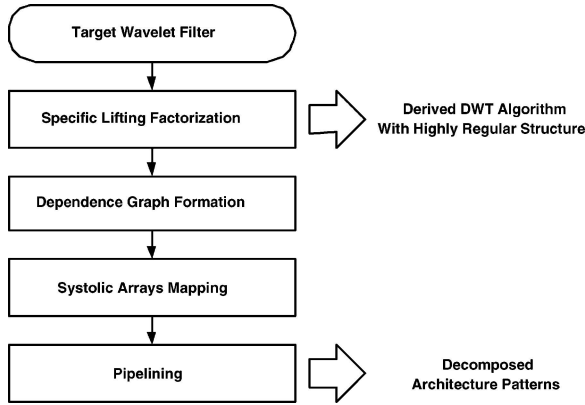


Figure 8. Proposed systematic design method.

method. As shown in Fig. 8, this design method consists of several design stages. Once a finite wavelet filter is targeted, four subsequent design stages are then performed to derive the corresponding DWT algorithm with highly regular structure and construct its decomposed architecture patterns. Detailed contents of each design stage are described in the following subsections.

3.1. Specific Lifting Factorization

As pointed out in Section 2, the lifting factorization process is non-unique. This freedom diversifies the design space of algorithm and corresponding architecture for lifting-based DWT. In the proposed systematic design method, a specific lifting factorization is chosen for all target wavelet filters. This factorization principle is to factor the Laurent polynomials $s_i(z)$ and $t_i(z)$ as symmetric or anti-symmetric as possible and allow at most two coefficients in each lifting step to achieve minimum lifting steps factorization. For instance, one lifting step can further be decomposed into two minimum lifting steps as

$$\begin{bmatrix} 1 & a(z) + b(z) \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & a(z) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & b(z) \\ 0 & 1 \end{bmatrix} \quad (4)$$

Following this principle, in each lifting step, an even location will only get information from two odd locations or vice versa. There exist only four possible categories of basic processing element in such factorization as shown in Fig. 9. Eqs. (5)–(8) show the four possible lifting step categories for $t_i(z)$, and each corresponds to the four basic processing element categories

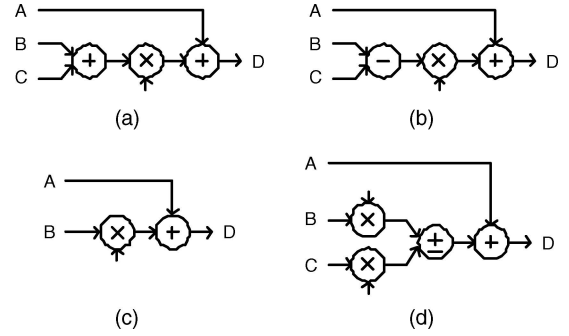


Figure 9. Four categories of basic processing element.

in Fig. 9. The case of $s_i(z)$ is similar.

$$\begin{bmatrix} 1 & 0 \\ \alpha(1 + z^{\pm N}) & 1 \end{bmatrix} \quad (5)$$

$$\begin{bmatrix} 1 & 0 \\ \alpha(1 - z^{\pm N}) & 1 \end{bmatrix} \quad (6)$$

$$\begin{bmatrix} 1 & 0 \\ \alpha & 1 \end{bmatrix} \quad (7)$$

$$\begin{bmatrix} 1 & 0 \\ \alpha \pm \beta z^{\pm N} & 1 \end{bmatrix} \quad (8)$$

The category (d) in Fig. 9 can be regarded as a general case of (a), (b), and (c). One can expect that, if the target wavelet filter is linear phase, namely, symmetric or anti-symmetric, then only the first three categories should appear by specific lifting factorization. In such cases, the number of multiplication in each lifting step can be reduced by at most a factor of two. In the following design stages, the scale factor K and $1/K$ will be excluded since it can be implemented exactly with two constant coefficient multipliers. After this design stage, the corresponding DWT algorithm with highly regular structure is derived.

3.2. Dependence Graph Formation

Once the specific lifting factorization is done, a dependence graph (DG) can be drawn for its corresponding lifting factored wavelet filter. However, in order to simplify the complexity of next design stage, the Systolic Arrays Mapping, a specific formation of the DG is performed to obtain a more regular and compact DG form.

As shown in Fig. 10(a), any lifting step constructed by specific lifting factorization can be depicted as a

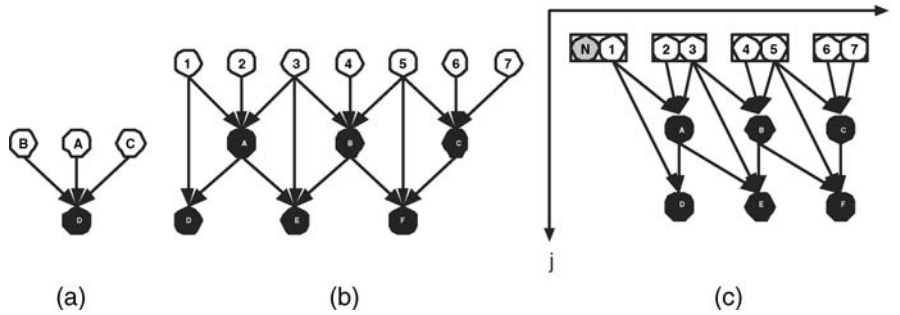


Figure 10. Dependence graph formation.

generic basic DG that is a combination of three input nodes (A, B, C) and one computation-output node (D). Due to the step-by-step serial connection property of specific lifting factorization, without loss of generality but for simplicity and regularity consideration, one slice of a DG can be depicted as shown in Fig. 10(b). In Fig. 10(b), the white node (tagged 1 to 7) denoted as the input node, and the black node (tagged A to F) denoted as the computation-output node. The formation principle is described as following two steps: First, merge one pair of even and odd input nodes into a new input node. As shown in Fig. 10(c), except for the first even node (tagged 1), one even and one odd node are merged into new single input node. The first even node 1 can be treated as merged with a virtual odd node N such that this merging step is regular. This step can make sure that the following systolic arrays mapped architectures have unified input-output ports and throughput. Second, move the computation-output nodes to the specific position such that there is no backward directional data flow existing in the DG. This step can make sure that the mapped architectures have unified data flow direction. Figure 10(c) is the DG form of Fig. 10(b) after these two formation steps.

3.3. Systolic Arrays Mapping and Pipelining

After the DG formation design stage, one set of unique systolic arrays mapping parameters are applied to the DG to obtain the corresponding signal flow graph (SFG). As the same systolic architecture definitions in [20], the DG in Fig. 10(c) is mapped by Processor Vector $p = (0, 1)^T$, Projection Vector $d = (1, 0)^T$, and Scheduling Vector $s = (1, 0)^T$.

The resulting SFG is depicted as shown in Fig. 11(a). The detailed architecture of each PE can be referred to

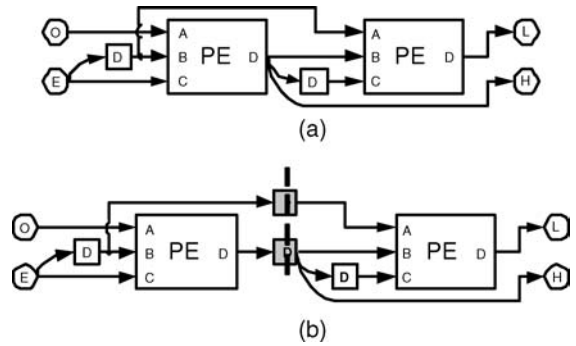


Figure 11. Systolic arrays mapped architecture.

Fig. 9, one of the four categories of basic processing element depending on its corresponding lifting step category. In this architecture, the critical path is two PE delay. In order to achieve modular architecture, pipelining is applied to the original SFG. As shown in Fig. 11(b), after the pipelining (dash line), two pipeline delay registers (D) are added and one PE critical path delay is achieved. By above four design stages, modular lifting-based architectures of any finite wavelet filter can be easily constructed. These constructed architectures are composed of several architecture patterns, including the PE for computations, delay registers for memory access, and dedicated data links for interconnection.

3.4. Case Studies

In this subsection, two practical examples are given to show the effectiveness of proposed systematic design method.

3.4.1. (9,7) Odd Symmetric Biorthogonal Filter. The first case to be studied is the popular (9,7) odd

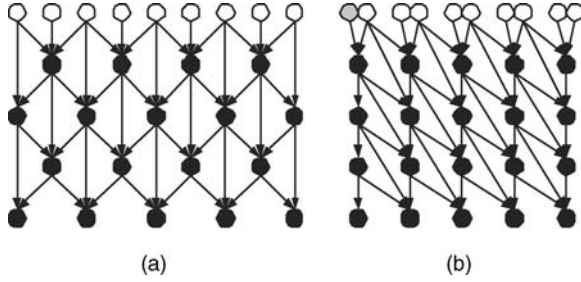


Figure 12. Dependence graph formation for (9,7) filter.

symmetric biorthogonal filter, which is adopted by JPEG2000 lossy coding. By specific lifting factorization, the polyphase matrix can be factored into four lifting steps and a scaling constant.

$$P(z) = \begin{bmatrix} 1 & \alpha(1+z^{-1}) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \beta(1+z) & 1 \end{bmatrix} \times \begin{bmatrix} 1 & \gamma(1+z^{-1}) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \delta(1+z) & 1 \end{bmatrix} \begin{bmatrix} \zeta & 0 \\ 0 & 1/\zeta \end{bmatrix} \quad (9)$$

$$\alpha = -1.586134342; \quad \beta = -0.05298011854;$$

$$\gamma = 0.8829110762; \quad \delta = 0.4435068522;$$

$$\zeta = 1.149604398$$

This factorization leads to the original and specific DG formation as shown in Fig. 12(a) and (b). After the systolic arrays mapping, the lifting-based architecture of (9,7) filter is shown in Fig. 13. The four PE architectures in this figure all correspond to Fig. 9(a). Finally, pipelining is made between each PE stages to construct the modular architecture.

3.4.2. (9,3) Odd Symmetric Biorthogonal Filter. The other case to be studied is the (9,3) odd symmetric biorthogonal filter, which is adopted by MPEG-4 visual texture coding. By specific lifting factorization, the polyphase matrix can be factored into three lifting

steps and a scaling constant.

$$P(z) = \begin{bmatrix} 1 & \alpha(1+z^{-1}) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \beta(1+z) & 1 \end{bmatrix} \times \begin{bmatrix} 1 & \gamma(1+z^{-1}) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \delta(1+z) & 1 \end{bmatrix} \begin{bmatrix} \zeta & 0 \\ 0 & 1/\zeta \end{bmatrix} \quad (10)$$

$$\alpha = -1.586134342; \quad \beta = -0.05298011854;$$

$$\gamma = 0.8829110762; \quad \delta = 0.4435068522;$$

$$\zeta = 1.149604398$$

This factorization leads to the original and specific DG formation as shown in Fig. 14(a) and (b). After the systolic arrays mapping, the lifting-based architecture of (9,3) filter is shown in Fig. 15. The three PE architectures in this figure also all correspond to Fig. 9(a). Again, pipelining is made between each PE stages to construct the modular architecture.

The lifting-based architectures constructed by this design method consist of several serially-connected basic processing elements, and the number of basic processing elements for a chosen wavelet filter depends on the number of lifting steps after specific lifting factorization. For instance, there are four basic processing elements in the (9,7) odd symmetric biorthogonal filter and three basic processing elements in the (9,3) odd symmetric biorthogonal filter.

4. Reconfigurable DWT Processor Architecture

In this section, a modular and scalable reconfigurable DWT processor architecture is proposed. The architecture patterns decomposed in Section 3 can be mapped onto proposed architecture, which is a combination of datapath units, specially partitioned and accessed memory blocks connected by dedicated links.

4.1. Reconfigurable DWT Processor Architecture Overview

In order to support variable wavelet filters and wavelet decomposition structures in a single architecture, a

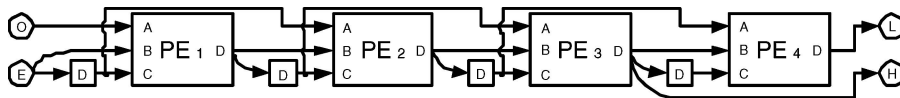


Figure 13. Lifting-based architecture of (9,7) filter.

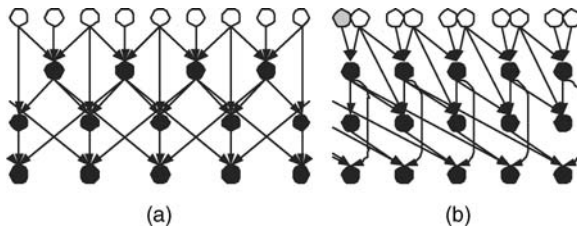


Figure 14. Dependence graph formation for (9,3) filter.

dynamically reconfigurable DWT processor architecture is proposed as shown in Fig. 16.

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. 16 act as the interface between the reconfigurable architecture and this frame memory. In a multimedia SoC, this virtual external frame memory can be implemented by a shared system memory or by a local frame memory tightly-coupled 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. 16, 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 computationally more efficient lifting scheme rather than conventional convolution approach. Besides, the proposed systematic design method in Section 3 is exploited to derive the reconfigurable DWT PE architecture and generate the corresponding hardware configurations of different wavelet filters 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. 16, 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.

4.2. Architecture of Reconfigurable DWT PE Array

According to the possible categories of basic processing element by proposed systematic design method in Section 3, if the target wavelet filters are linear phase, then only the first three categories should appear by specific lifting factorization. Therefore, the core cell, which is called the main computation unit (MCU), of reconfigurable DWT PE is derived as shown in Fig. 17. This core cell is a three-input (A, B, C) one-output (D) datapath, consisting of one adder/subtractor, one multiplier with coefficient α , and another adder. The datapath can be dynamically reconfigured as one of the three possible categories of basic processing element.

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 Section 3, since the number of basic processing elements is variable for different wavelet filter, a folding of systolic array technique can be exploited to fold variable number of basic processing elements into equal number of MCU with variable throughput. For instance, the (9,7) filter originally require four basic processing elements, 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

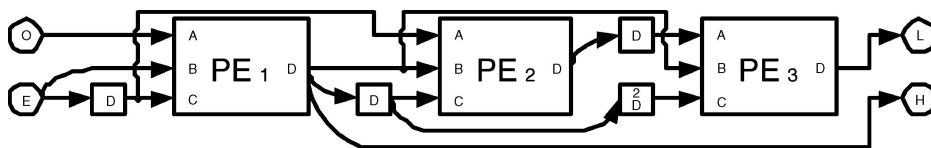


Figure 15. Lifting-based architecture of (9,3) filter.

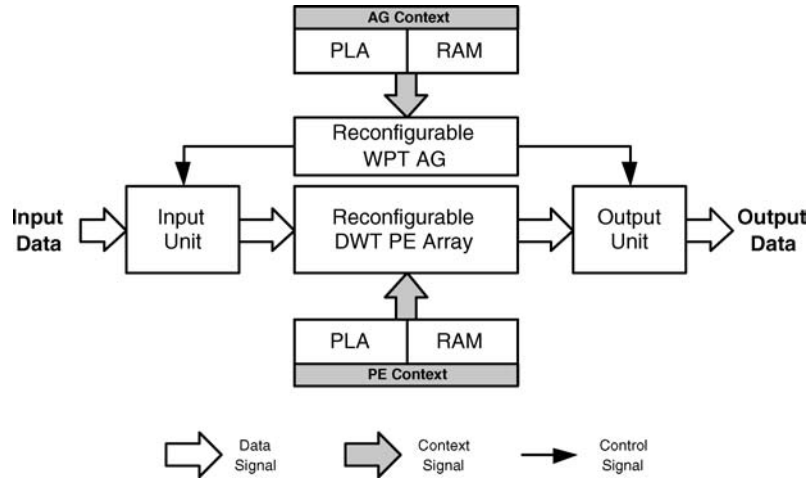


Figure 16. Proposed reconfigurable DWT processor architecture.

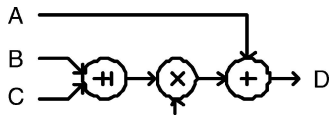


Figure 17. Core cell of reconfigurable DWT PE (MCU).

between each MCU, the reconfigurable DWT PE architecture is derived as shown in Fig. 18. In Fig. 18, 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. 17, 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.

By adding an additional design stage, 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 can be derived. Based on this modified design method, any finite wavelet filter can be mapped onto the Reconfigurable DWT PE Array with specific PE number through the generated hardware configurations.

4.3. Architecture of 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. 19, there are 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

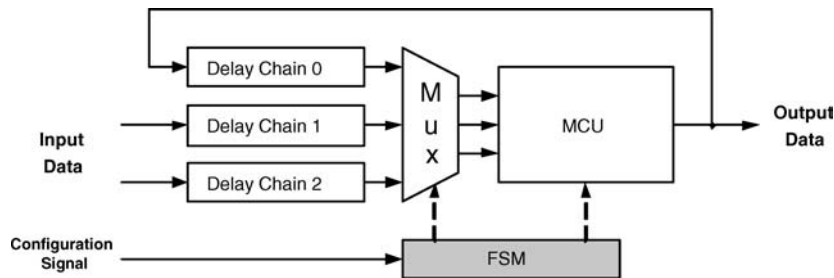


Figure 18. Reconfigurable DWT PE architecture.

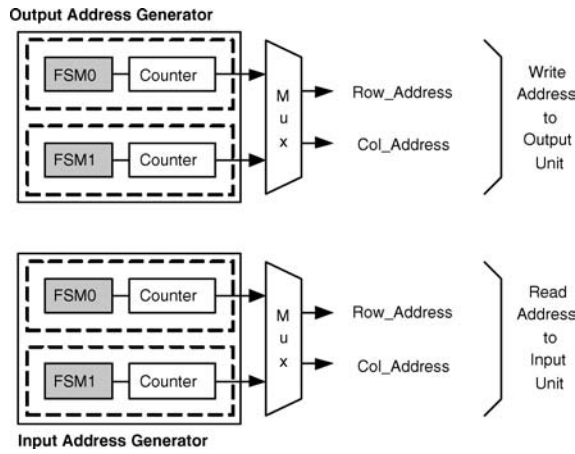


Figure 19. Architecture of Reconfigurable WPT AG.

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 decomposition structure.

5. Chip Implementation and Architecture Evaluation

5.1. Chip Implementation

In order to prove the feasibility of proposed reconfigurable DWT processor architecture, a prototyping chip has been implemented by cell-based design flow and fabricated by TSMC 0.35 μm 1P4M CMOS process. Two reconfigurable DWT PE are adopted to form the Reconfigurable DWT PE Array, and several useful wavelet filters and wavelet decomposition structures are stored in the PLA as default configurations. The key features of this prototyping chip is listed in Table 1 and the related performance is showed in Table 2, including

Table 1. Key features of the prototyping chip.

Technology	TSMC 0.35 μm 1P4M CMOS process
Package	100 CQFP
Die size	$2.86 \times 2.86 \text{ mm}^2$
Transistor count	168 K
Max clock rate	50 MHz
Power consumption	186 mW @ 3.3 V, 50 MHz

Table 2. Performance of the prototype chip.

Wavelet filter	Lifting steps	Through put (per cycle)	HW utilization (%)
(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

the wavelet filters, number of lifting steps, throughput per clock cycle, and corresponding hardware utilization. At 50 MHz, the prototyping chip can achieve at most 100M pixels/sec transform throughput (for (5,3) filter), which is capable to perform the CCIR 601 (720×576) format image signal at 30 frame/sec with two-level wavelet packet transform. The photograph of the prototyping chip is shown in Fig. 20.

5.2. Architecture Evaluation

Two categories of architecture evaluation have been made to show the energy efficiency and architectural uniqueness of proposed reconfigurable DWT processor.

5.2.1. Energy Efficiency Comparison with Programmable Solutions.

The proposed reconfigurable

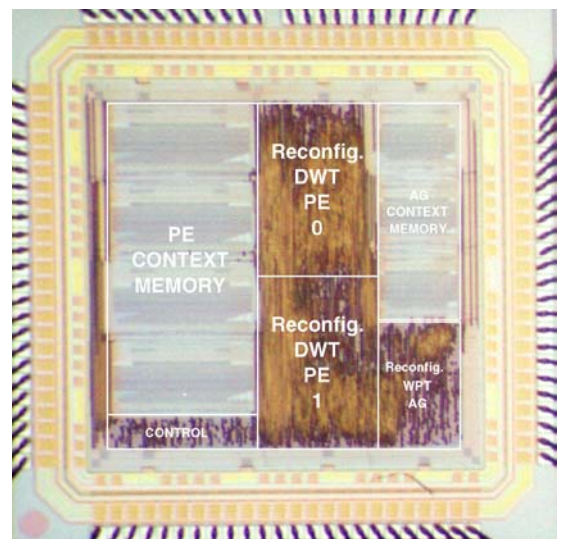


Figure 20. Photograph of the prototyping chip.

Table 3. Energy efficiency comparison.

Architecture candidates	Wavelet filter			
	(5,3)	(9,3)	(9,7)	(2,10)
Dedicated Hardware (lower bound)	0.32	0.65	0.99	0.87
Proposed reconfigurable DWT processor	0.57	0.86	1.14	1.14
Low-power DSP (TI C54)	9.99	10.53	11.07	10.26
High-performance DSP (TI C62)	60	72	96	72

DWT processor has been compared with two programmable digital signal processors (DSPs) from Texas Instruments [21], one is a low-power DSP—TMS320C54, and the other is a high-performance DSP—TMS320C62. For DSP implementation, the wavelet filters are realized by real symmetric FIR filter with polyphase decomposition. Besides, the dedicated hardware implementations are also included as a lower bound reference. In order to achieve a reasonable comparison, all the IC technologies of targeted architecture candidates are scaled to 0.18 μm . The comparison results are listed in Table 3, and the unit is mW/Msamples. According to the results, it is clear that our proposal achieves 10 to 100 times energy efficiency than programmable solutions and approaches the performance of dedicated hardware.

5.2.2. Comparison with Programmable/Reconfigurable DWT Architectures. In order to show the uniqueness of our proposal in terms of reconfigurability, the proposed architecture has been compared with several previous programmable or reconfigurable DWT architectures. The comparison

Table 4. Comparison with programmable or reconfigurable DWT architectures.

Architecture	Variable wavelet filter	Wavelet filter basis	Variable decomposition structure
Proposed	Yes	Lifting	Yes
Chen [11]	Yes	Convolution	No (dyadic only)
Ravasi [12]	Yes	Convolution	No (dyadic only)
Ferretti [13]	Yes	Lifting	No (dyadic only)
Andra [14]	Yes	Lifting	No (dyadic only)
Trenas [15]	No	Not specified	Yes
Wu [16]	No	Not specified	Yes

results are listed in Table 4. The results show that our proposal has the richest reconfigurability among all proposals and is the only solution that can provide sufficient functional flexibility desired by heterogeneous reconfigurable multimedia systems for future generation multimedia SoC.

6. Conclusion

We have proposed a reconfigurable DWT processor architecture to meet the diverse computing requirements of future generation multimedia SoC. The proposed architecture is dynamically reconfigurable in terms of the wavelet filters and wavelet decomposition structures. By proposed systematic design method, the DWT algorithm with highly regular structure is derived and decomposed into architecture patterns to be mapped onto proposed reconfigurable architecture. The lifting-based Reconfigurable DWT PE Array possesses better computation efficiency than convolution-based architectures, and 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, high energy efficiency, and unique reconfigurability, proving it to be a universal and extremely flexible computing engine for heterogeneous reconfigurable multimedia systems.

Acknowledgments

This work was supported in part by MOE Program for Promoting Academic Excellence of Universities under the grant number 89E-FA06-2-4-8, in part by National Science Council, Republic of China, under the grant number 91-2215-E-002-035, and in part by MediaTek Inc. The multiproject chip support from the National Science Council of Taiwan/Chip Implementation Center is also acknowledged.

References

1. J.M. Rabaey, A. Abnous, Y. Ichikawa, K. Seno, and M. Wan, "Heterogeneous Reconfigurable Systems," in *Proc. of IEEE Workshop on Signal Processing Systems*, 1997, pp. 24–34.
2. S.G. Mallat, "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, 1989, pp. 674–693.

3. *JPEG 2000 Part 1 Final Draft International Standard*, ISO/IEC FDIS15444-1, Dec. 2000.
4. *Information Technology—Coding of Audio-Visual Objects - Part 2: Visual*, ISO/IEC 14496-2, 1999.
5. K.K. Parhi and T. Nishitani, "VLSI Architectures for Discrete Wavelet Transforms," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 1, no. 2, 1993, pp. 191–202.
6. M. Vishwanath, R.M. Owens, and M.J. Irwin, "VLSI Architectures for the Discrete Wavelet Transform," *IEEE Transactions on Circuits and Systems—II: Analog and Digital Signal Processing*, vol. 42, no. 5, 1995, pp. 305–316.
7. A. Grzeszczak, M.K. Mandal, S. Panchanathan, and T. Yeap, "VLSI Implementation of Discrete Wavelet Transform," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 4, no. 4, 1996, pp. 421–433.
8. C. Chakrabarti, M. Vishwanath, and R.M. Owens, "Architectures for Wavelet Transforms: A Survey," *The Journal of VLSI Signal Processing*, vol. 14, 1996, pp. 171–192.
9. P.C. Wu and L.G. Chen, "An Efficient Architecture for Two-Dimensional Discrete Wavelet Transform," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, no. 4, 2001, pp. 536–545.
10. M. Weeks and M. Bayoumi, "Discrete Wavelet Transform: Architectures, design and Performance Issues," *The Journal of VLSI Signal Processing*, vol. 35, Sept. 2003, pp. 155–178.
11. C.Y. Chen, Z.L. Yang, T.C. Wang, and L.G. Chen, "A Programmable Parallel VLSI Architecture for 2-D Discrete Wavelet Transform," *The Journal of VLSI Signal Processing*, vol. 28, 2001, pp. 151–163.
12. M. Ravasi, L. Tenze, and M. Mattavelli, "A Scalable and Programmable Architecture for 2-D DWT Decoding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 12, no. 8, 2002, pp. 671–677.
13. M. Ferretti and D. Rizzo, "A Parallel Architecture for the 2-D Discrete Wavelet Transform with Integer Lifting Scheme," *The Journal of VLSI Signal Processing*, vol. 28, July 2001, pp. 165–185.
14. K. Andra, C. Chakrabarti, and T. Acharya, "A VLSI Architecture for Lifting-Based Forward and Inverse Wavelet Transform," *IEEE Transactions on Signal Processing*, vol. 50, no. 4, 2002, pp. 966–977.
15. M.A. Trenas, J. Lopez, and E.L. Zapata, "A Configurable Architecture for the Wavelet Packet Transform," *The Journal of VLSI Signal Processing*, vol. 32, Nov. 2002, pp. 255–273.
16. X. Wu, Y. Li, and H. Chen, "Programmable Wavelet Packet Transform Processor," *IEE Electronics Letters*, vol. 35, no. 6, 1999, pp. 449–450.
17. A. Bovik, *Handbook of Image and Video Processing*, Academic Press, 2000.
18. W. Sweldens, "The Lifting Scheme: A Custom-Design Construction of Biorthogonal Wavelets," *Applied and Computational Harmonic Analysis*, vol. 3, no. 15, 1996, pp. 186–200.
19. I. Daubechies and W. Sweldens, "Factoring Wavelet Transforms into Lifting Steps," *The Journal of Fourier Analysis and Applications*, vol. 4, 1998, pp. 247–269.
20. K.K. Parhi, *VLSI Digital Signal Processing Systems—Design and Implementation*, Wiley Interscience, 1999.
21. *Texas Instruments*, <http://www.ti.com>.



Po-Chih Tseng was born in Tao-Yuan, Taiwan in 1977. He received the B.S. degree in Electrical and Control Engineering from National Chiao Tung University in 1999 and the M.S. degree in Electrical Engineering from National Taiwan University in 2001. He currently is pursuing the Ph.D. degree at the Graduate Institute of Electronics Engineering, Department of Electrical Engineering, National Taiwan University. His research interests include VLSI design and implementation for signal processing systems, energy-efficient reconfigurable computing for multimedia systems, and power-aware image and video coding systems.
pctseng@video.ee.ntu.edu.tw



Chao-Tsung Huang was born in Kaohsiung, Taiwan, R.O.C., in 1979. He received the B.S. degree from the Department of Electrical Engineering, National Taiwan University, Taipei, Taiwan, R.O.C., in 2001. He currently is working toward the Ph.D. degree at the Graduate Institute of Electronics Engineering, National Taiwan University. His major research interests include VLSI design and implementation for signal processing systems.
cthuang@video.ee.ntu.edu.tw



Liang-Gee Chen (S'84–M'86–SM'94–F'01) received the B.S., M.S., and Ph.D. degrees in electrical engineering from National Cheng Kung University, Tainan, Taiwan, R.O.C., in 1979, 1981, and 1986, respectively. In 1988, he joined the Department of Electrical Engineering, National Taiwan University, Taipei, Taiwan, R.O.C.

During 1993–1994, he was a Visiting Consultant in the DSP Research Department, AT&T Bell Labs, Murray Hill, NJ. In 1997, he was a Visiting Scholar of the Department of Electrical Engineering, University of Washington, Seattle. Currently, he is Professor at National Taiwan University, Taipei, Taiwan, R.O.C. His current research interests are DSP architecture design, video processor design, and video coding systems.

Dr. Chen has served as an Associate Editor of IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY since 1996, as Associate Editor of the IEEE TRANSACTIONS ON VLSI SYSTEMS since 1999, and as Associate Editor of IEEE TRANSACTIONS CIRCUITS AND SYSTEMS II since 2000. He has been the Associate Editor of the *Journal of Circuits, Systems, and Signal Processing* since 1999, and a Guest Editor for the *Journal of VLSI Signal Processing Systems*. He is also the Associate Editor of the PROCEEDINGS OF THE IEEE. He was the General Chairman of the 7th VLSI Design/CAD Symposium in 1995 and of the 1999 IEEE Workshop on Signal Processing Systems: Design and Imple-

mentation. He is the Past-Chair of Taipei Chapter of IEEE Circuits and Systems (CAS) Society, and is a member of the IEEE CAS Technical Committee of VLSI Systems and Applications, the Technical Committee of Visual Signal Processing and Communications, and the IEEE Signal Processing Technical Committee of Design and Implementation of SP Systems. He is the Chair-Elect of the IEEE CAS Technical Committee on Multimedia Systems and Applications. During 2001–2002, he served as a Distinguished Lecturer of the IEEE CAS Society. He received the Best Paper Award from the R.O.C. Computer Society in 1990 and 1994. Annually from 1991 to 1999, he received Long-Term (Acer) Paper Awards. In 1992, he received the Best Paper Award of the 1992 Asia-Pacific Conference on circuits and systems in the VLSI design track. In 1993, he received the Annual Paper Award of the Chinese Engineer Society. In 1996 and 2000, he received the Outstanding Research Award from the National Science Council, and in 2000, the Dragon Excellence Award from Acer. He is a member of Phi Tan Phi.
lgchen@video.ee.ntu.edu.tw