A Reconfigurable Hardware Implementation
of the One-Dimensional Discrete Wavelet Transform

ALI M. Al-HAJ
Department of Electronics & Computer Engineering,
Princess Sumaya University for Technology,
Al-Jubeiha P.O.Box 1438, Amman 11941,
JORDAN

*Abstract —*ASIC hardware implementations of the discrete wavelet transform are required to cope with the intensive real-time computations of the transform. In this paper, we describe a parallel implementation of the wavelet transform using one type of reconfigurable ASICs; Filed Programmable Gate Arrays (FPGAs). The implementation is based on reformulating the transform using the distributed arithmetic and ployphase decomposition techniques, so that the ample inherent parallelism of the transform can be well exploited by the fine-grained parallel architecture of Virtex FPGAs. Performance results demonstrate the applicability of   FPGAs with distributed   arithmetic and polyphase decomposition   to achieve the required high computational speeds of the discrete wavelet transform.

Key words:—Discrete wavelet transform, FPGA implementation, Polyphase filters, Distributed arithmetic.

1   Introduction

The discrete wavelet transform is a powerful new mathematical method with a broad spectrum of potential applications [1]. It has already been used successfully in   signal   processing [2] and numerical analysis [3], among many other audio-visual applications. In particular, the area of data compression has benefited incredibly from   the wavelet transform [4]. However, the transform 's high   computational requirements may limit its wide-spread, especially in applications requiring real time performance. Consequently, there has been a great demand on high-speed computing devices to meet the real-time computational requirements of the transform.

Fortunately, the discrete wavelet transform is inherently parallel, and it lends itself to hardware implementations on VLSI devices [5]. Indeed, many VLSI implementations of the transform have appeared in literature [6-9], however, most of the proposed architectures require   complex control units, and are not easily scaled up for different wavelets filters   and   different octave levels.

Recently, Filed programmable gate arrays (FPGAs) have become an attractive implementation platform for many digital signal processing algorithms [10-13]. FPGAs are programmable   ASICs,   offer   intermediate capabilities between those offered by   custom VLSI   ASICs and digital signal processors. Indeed, programmability of FPGAs makes them a perfect choice for implementing the discrete wavelet transform since this would allow easy modification of different wavelet types.

In this paper, we describe a parallel FPGA implementation of the discrete wavelet transform using Virtex FPGAs [14]. The implementation is based on achieving high execution speeds by exploiting the abundant inherent parallelism of the transform using the distributed arithmetic [15] and plolyphse decomposition [16]. Section 2 reviews the   Mallat's   pyramid   algorithm.   The implementation is described in section 3, and simulated in section 4. Results are presented in section 5 and conclusions in section 6.

2 Mallat's Pyramid Algorithm

Wavelets are special functions which, in a form analogous to sines and cosines in Fourier analysis, are used as basal functions for representing signals. The coefficients of the discrete wavelet transform can be calculated recursively and in a straight forward manner using the well-known Mallat's pyramid algorithm [17]. Based on Mallat's algorithm, the discrete wavelet coefficients of any stage can be computed from the coefficients of the previous stage using the following iterative equations:

$$W_L(n, j) = \sum_m W_L(m, j-1)h_0(m-2n)........(1)$$

$$W_H(n, j) = \sum_m W_L(m, j-1)h_1(m-2n)........(2)$$

Where $W_L(n,j)$ is the nth scaling coefficient at the jth stage, $W_H(n,j)$ is the nth wavelet coefficient at the jth stage, and $h_0(n)$ and $h_1(n)$ are the dilation coefficients corresponding to the scaling and wavelet functions, respectively. In order to reconstruct the original data, the DWT coefficients are upsampled and passed through another set of low pass and high pass filters, which is expressed as

$$W_L(n, j) = \sum_k W_L(k, j+1)g_0(n-2k) +$$

$$\sum_l W_H(l, j+1)g_1(n-2l)............................(3)$$

where $g_0(n)$ and $g_1(n)$ are respectively the low-pass and high-pass synthesis filters corresponding to the mother wavelet. It is observed from Equation (3) that the jth level coefficients can be obtained from the $(j+1)^{th}$ level coefficients.

Daubechies 8-tap wavelet has been chosen for this implementation. This wavelet type is known for its excellent special and spectral localities which are useful properties in image compression [18]. The filters coefficients corresponding to this wavelet type are shown in Table 1. H_0 and H_1 are the input decomposition filters and G_0 and G_1 are the output reconstruction filters.

Table 1.
Daubechies 8-tap wavelet filter coefficients.

H_0	H_1	G_0	G_1
-0.0106	0.2304	-0.2304	-0.0106
-0.0329	0.7148	0.7148	0.0329
0.0308	0.6309	-0.6309	0.0308
0.1870	-0.0280	-0.0280	-0.187
-0.0280	-0.1870	0.1870	-0.0280
-0.6309	0.0308	0.0329	0.6309
0.7148	0.0329	-0.0329	0.7148
-0.2304	-0.0106	-0.0106	0.2304

3 The Implementation

In this section, we describe an efficient implementation of the discrete wavelet transform. The implementation is based on re-arranging the QMF discrete wavelet transform multirate structure, described in [19] and shown in Figure 1, as a polyphase tree, and then implementing each sub-filter of the polyphase tree as a distributed arithmetic filter. This combination of the two efficient filter structures; polyphase and distributed arithmetic based structures, leads to an efficient implementation of the discrete wavelet transform. In what follows, we first describe the polyphase reformulation of the analysis and synthesis filter banks, and then we describe the distributed arithmetic implementation of the sub-filter structure of the polyphase filter banks.
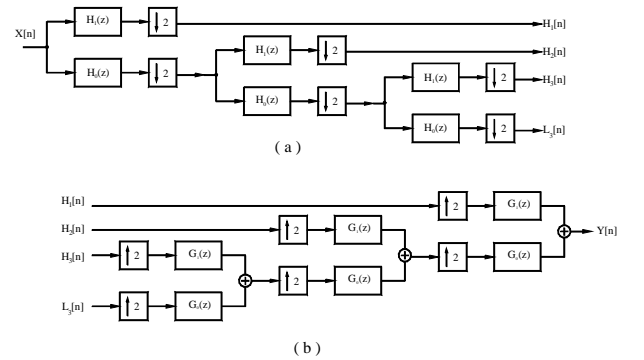

Fig. 1. Mallat's quadratic mirror filter tree (a). forward DWT tree; (b). inverse DWT tree.

3.1 Polyphase Filter Banks

Filters used in the parallel pyramid tree architecture of Figure 1, are constructed using FIR filters because of their inherent stability [20]. A direct form realization of FIR filter is shown in Figure 2a. A computationally efficient realization of the filter consists of decomposing the filter into two sub-filters executing in parallel. This realization is based on the polyphase decomposition algorithm [21], and results in a parallel architecture useful for real time applications. Consider the transfer function given in Equation (4). By separating the even numbered coefficients from the odd numbered ones, we can rewrite the transfer function as in Equation (5):

$$H(z) = \sum_{k=0}^{N-1} h[k] z^{-k} \qquad (4)$$

$$H(z) = E_0(z^2) + z^{-1} E_1(z^2) \qquad (5)$$

where $E_0(Z^2)$ and $E_1(Z^2)$ are the even sub-filter and odd sub-filter, respectively. Figure 2b shows the two-branch parallel realization of the direct-form FIR shown in Figure2a.
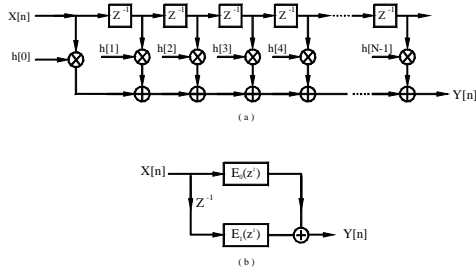

Fig. 2. (a). direct FIR structure; (b). polyphase realization.

3.1.1 Analysis Filter Bank

The analysis filter bank shown in Figure 3a represents the basic building block of the forward discrete wavelet transform. It consists of two decimators connected in parallel; the upper decimator is a low pass filter followed by a down sampler, and the lower decimator is a high pass filter followed by a down sampler. The down-sampler operates by taking a filtered sequence $x[n]$ and generating an output sequence $y[n]$ according to the relation $y[n] = x[2n]$.

All filtered elements in the subsequence $x[2n+1]$ are discarded. Consequently, the direct structure shown in Figure 3a is computationally inefficient since it unnecessarily calculates the values $x[n]$ for $n \neq 2n$, which are discarded by the down sampler after being calculated. To avoid such unnecessary calculation, a more efficient, but equivalent, implementation of the analysis filter bank exists, and its based on the concept of polyphase decomposition. If we represent the low pass transfer function in polyphase form, as explained earlier, we obtain the final polyphase structure of the analysis filter bank shown in Figure 3b. Analysis of the polyphase structure reveals that we have achieved a significant computation reduction (roughly a quarter of the computational complexity) in exchange for a modest increase in algorithm complexity and control.
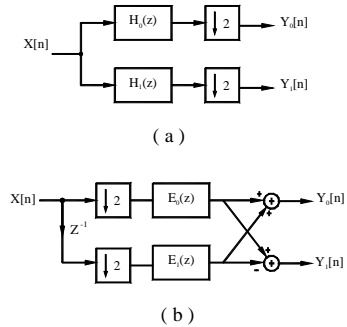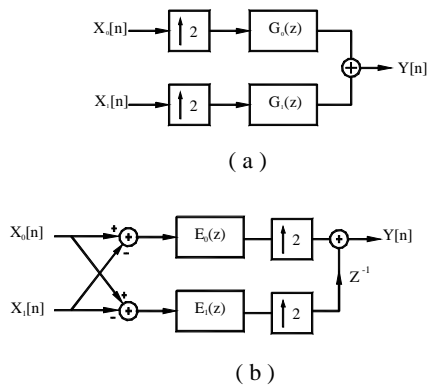

Fig. 3. Polyphase realization of the analysis filter bank : (a). direct structure; (b). polyphase structure.

3.1.2 Synthesis Filter Bank

The synthesis filter bank shown in Figure 4a represents the basic building block of the inverse discrete wavelet transform. It consists of two interpolators connected in parallel; the upper is a low pass filter proceeded by an up-sampler, and the lower is a high pass filter proceeded by an up-sampler. The up-sampler inserts an equidistant zero-valued sample between every two consecutive samples on the input sequence $x[n]$. An output sequence $y[n]$ is developed such that $y[n] = x[n/2]$ for even indices of n, and 0 otherwise. This makes the sampling rate of the output sequence $y[n]$ twice as large as the sampling rate of the original sequence $x[n]$.

3

We immediately observe a source of inefficiency in this simple interpolation scheme. One out of every two samples presented to the filter represents the actual data sample, and the other sample is zero. Therefore, its clear computation power is being wasted in performing arithmetic operations on zero values. To avoid such multiply-by-zero arithmetic operations, a more efficient ,but equivalent, implementation of the synthesis filter bank exists, and its based on the concept of polyphase decomposition. If we expand G_0 (z) to its polyphase form as explained earlier, we obtain the polyphase structure of the synthesis filter bank shown in Figure 4b. Analysis of the polyphase structure reveals that we have achieved a significant computation reduction (roughly a quarter of the computational complexity) in exchange for a modest increase in algorithm complexity and control.

(a)

(b)

Fig. 4. Polyphase realization of the synthesis filter bank: (a).direct structure (b). polyphase structure.

3.2 Distributed Arithmetic Sub-Filters

Distributed arithmetic is an efficient method for computing the inner product operation which constitutes the core of the discrete wavelet transform. Mathematical derivation of distributed arithmetic is extremely simple; a mix of Boolean and ordinary algebra [22]. Let the variable Y hold the result of an inner product operation between a data vector x and a coefficient vector a. The distributed arithmetic representation the inner product operation is given as follows:

$$Y = \sum_{j=1}^{B-1} \left[\sum_{i=1}^{N} x_{ij} a_i \right] 2^{-j} + \sum_{i=1}^{N} a_i (-x_{i0})$$

$$= \sum_{j=1}^{B-1} F_j 2^{-j} - F$$

(6)

where the input data words x_i have been represented by the 2's complement number presentation in order to bound number growth under multiplication. The variable x_{ij} is the jth bit of the x_i word which is Boolean, B is the number of bits of each input data word and x_{0i} is the sign bit. Distributed arithmetic is based on the observation that the function F_j can only take 2^N different values that can be pre-computed offline and stored in a look-up table. Bit j of each data x_{ij} is then used to address this look-up table. Equation (11) clearly shows that the only three different operations required for calculating the inner product. First, a look-up to obtain the value of F_j, then addition or subtraction, and finally a division by two that can be realized by a shift. FIR filters and the inner product operation described so far differ only in how they handle the input data.

Each sub-filter in the polyphase DWT analysis and synthesis structures, described in the previous subsection, is implemented as a distributed arithmetic FIR filter consisting of a look-up table (LUT) to store all possible partial products over the FIR filter coefficient of Table 1, a cascade of shift registers and a scaling accumulator, as shown in Figure 5a. Input samples are presented to the input parallel-to serial shift register at the input signal sample rate. As the input sample is serialized, the bit-wide output is presented to the bit-serial shift register cascade,1-bit at a time. The cascade stores the input sample history in a bit-serial format and is used in forming the required inner-product computation. The bit outputs of the shift register cascade are used as address inputs to the look-up table. Partial results from the look-up table are summed by the scaling accumulator to form a final result at the filter output port.

Since the LUT size in a distributed arithmetic implementation increases exponentially with the number of coefficients, the LUT access time can be a bottleneck for the speed of the whole system when the LUT size becomes large. Hence we decomposed the 8-bit LUT shown in Figure 5a

into two 4-bit LUTs, and added their outputs using a two-input accumulator. The modified partitioned-LUT architecture is shown in Figure 5b. The total size of storage is now reduced since the accumulator is less costly than the larger 8-bit LUT. Furthermore, partitioning the larger LUT into two smaller LUTs accessed in parallel reduces access time. In addition, throughput of the filter is maintained regardless of the length of the FIR filter. This feature is particularly attractive for flexible implementations of different wavelet types since each type has a different set of filer coefficients.
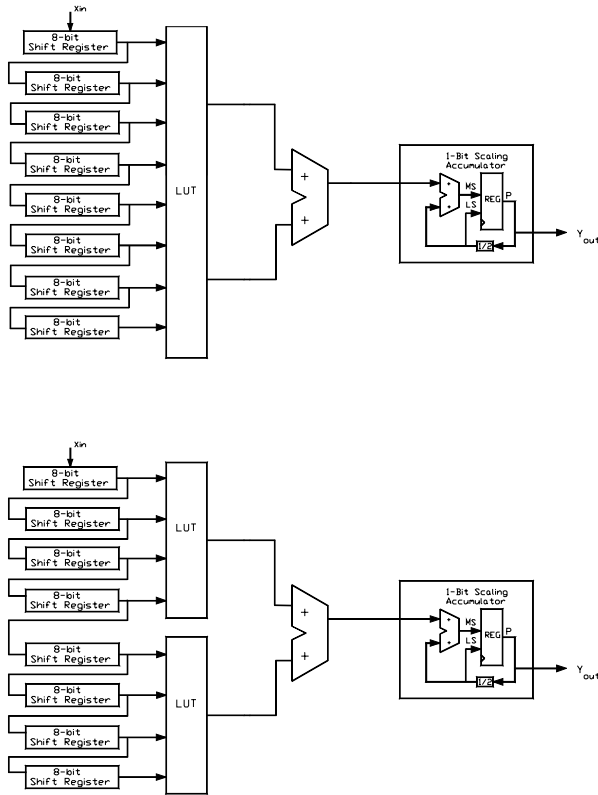
Fig. 5. Distributed arithmetic realizations of the FIR filter: (a). single-LUT realization; (b) efficient partitioned-LUT realization.

4 Functional Simulation

Functional simulation is a major prerequisite step towards a correct and efficient FPGA implementation of the discrete wavelet transform. Therefore, the implementation described in the previous section, was modeled by the Verilog hardware description language and verified by its functional simulator [23]. Simulation waveforms of the forward and inverse wavelet transforms are displayed in Figure 6. The waveforms prove that the implementation execute the operation of the wavelet transform correctly.

We used uniformly distributed 8-bit random input samples to generate the simulation waveforms. We also maintained sufficient precision of the intermediate and output coefficients since allocating sufficient bits to the intermediate and output coefficients is a necessary step to keep the perfect reconstruction capabilities of the discrete wavelet transform. If we allocate fewer bits than necessary, the output of the inverse discrete wavelet transform will not be the same as a delayed version of the input of the forward discrete wavelet transform. Also, if we're dealing with an image compression application, the decompressed image will suffer form some defects, such as ringing effects and blurring artifacts.

Simulation waveform of the forward wavelet transform architecture of Figure 1a is illustrated in Figure 6a. As an input sample X enters the first filter bank stage at a rate of 1sample/ clock, one sample (H_1) leaves to the output, and another sample (L_1) leaves to the second stage, both at a rate of 1sample/ 2 clocks. Similarly, the second stage sends a sample to the output (H_2), and another sample (L_2) to the third stage, both at a rate of 1sample/ 4 clocks. Finally, the third stage generates two samples (L_3 and H_3) at a rate of 1 sample/ 8 clocks.

Simulation waveform of the inverse wavelet transform architecture of Figure 1b is illustrated in Figure 6b. The first filter bank stage receives two inputs (H_3 and L_3), both produced from the third stage of the forward DWT at a rate of 1sample/ 8 clocks. The stage up-samples each of them by a factor of 2, and sends out their filtered summation at the rate of 1sample/ 4 clocks to the second stage, to be processed with an input sample coming from the second stage of the forward DWT stage at a rate of 1sample/ 4 clocks (H_2). Similarly, the second stage up-samples both by a factor of 2, and then sends out their filtered summation at a rate of 1sample/ 2 clocks to the third stage to be processed with an input sample coming from the first stage of the forward DWT at a rate of 1sample/ 2 clocks (H_1). Finally, the third

stage up-samples both by a factor of 2, and then sends out their filtered summation at a rate of 1sample/ clock to the output. This last output represents the reconstructed signal.
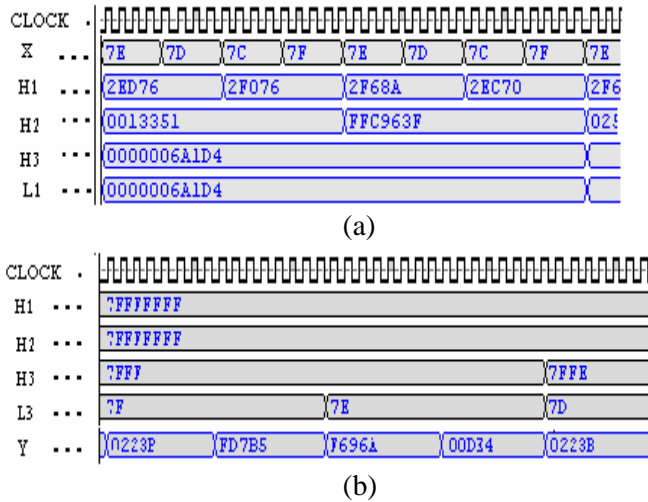

(a)

(b)

Fig. 6. Simplified functional Verilog simulation of the discrete wavelet transform : (a). forward DWT; (b). inverse DWT.

5 Discussion

In this section, we present performance results of the parallel polyphase & DA implementation described in section three. We also show how the results exceed considerably those obtained for other implementations of the transform.

5.1 Experimental Results

We carried out the implementation using a prototyping board called XSV-300 FPGA Board. The board is developed by XESS Inc. [24], and is based on a single XCV300 FPGA chip [25]. This chip contains 3072 slices (322,970 gates), where each slice contains 4-input, 1-output LUTs and two registers. The LUTs allow any function of five inputs or two functions of four inputs to be created within a CLB slice. Furthermore, The chip can operate at a maximum clock speed of 200 MHz. Performance is evaluated with respect to two metrics; throughput (sample rate) and is given in terms of the clock speed, and device utilization, and is given in terms number of logic slices used by the implementation. Using these two metrics, the results of the polyphase & DA

Implementation are as follows. The forward discrete wavelet transform implementation operated at a throughput of 131.7 MHz, and required 374 Virtex slices which represents 12 % of the total slices. Throughout of the inverse discrete wavelet transform implementation was 119.6 MHz, and the hardware requirement was 461 slices which represents 15 % of the total Virtex slices.

5.3 Performance Analysis

In what follows, we study the effects of using the polyphase decomposition and the distributed arithmetic techniques, separately. Therefore, we carried out three different implementations, and recorded their results in Tables 2 & 3. The first is a direct implementation is in which all filters in the DWT tree were implemented using the direct FIR structure shown in Figure 2a. The second is a polyphase implementation in which all filters in the DWT tree were implemented using the polyphase structure shown in Figures 3b and 4b. The third is a distributed arithmetic implementation in which all filters in the DWT tree were implemented using the distributed arithmetic FIR structure shown in Figure 5b.

Referring to Table 2, its noted that the throughput of the distributed arithmetic implementation is higher than the throughput of the direct implementation. This is expected since the distributed arithmetic implementation replaced the time-consuming conventional multiply accumulate operations with fast look-up tables and shift operations. Furthermore, partial products of all multiply accumulate operations were pre-computed offline and stored in the LUTs, thus saving a great a mount of real-time computation. As for Virtex slice utilization, Table 3 indicates that the distributed arithmetic implementation, uses less hardware resources than the direct implementation which uses conventional arithmetic. This is also expected since the conventional arithmetic multiplier requires more logic resources than the distributed arithmetic multiplier which requires small LUTs, simples adders and shift registers.

It is also noted from the results illustrated in Table 2, that the throughput of the polyphase implementation is much higher than the throughput of the direct implementation. This is

6

expected since the polyphase implementation avoids unnecessary decimator and interpolator computations as explained in section 3. Furthermore, realizing the different filter banks of the transform in parallel contributed significantly to the reduction of the total computation time, and in turn to the considerable increase in the sample throughput. As for the hardware resources requirements of the two implementations, Table 3 indicates that the requirements are comparable, with the polyphase implementation requiring slightly more than the direct implementation. This is due to the fact that parallelizing sequential structures necessitates using more hardware resources.

**Table 2.**
Throughput performance comparison.

Implementation	Forward DWT (MHz)	Inverse DWT (MHz)
Direct	14.11	11.6
Distributed Arithmetic	26.0	23.7
Polyphase Decomposition	104.6	98.5
Distributed Arithmetic & Polyphase Decomposition	131.7	119.6

Table 3.
Utilization performance comparison.

Implementation	Forward DWT (Slice)	Inverse DWT (Slice)
Direct	560 (18%)	619 (20%)
Distributed Arithmetic	374 (12%)	461 (15%)
Polyphase Decomposition	651 (21%)	708 (23%)
Distributed Arithmetic & Polyphase Decomposition	830 (27%)	922 (30%)

Finally, the wavelet transform was implemented on the TMS320C6711; a Texas Instrument digital signal processing board with a complex architecture suitable for image processing applications [26]. The board can operate at 150 MHz, with a peak performance of 900 MFLOPS [27]. It is noted from the results illustrated in Figure 7, that all the FPGA implementations perform much better than the TMS20C6711 software implementations The superior performance of the FPGA-based implementations is attributed to the highly parallel, pipelined and distributed architecture

of Xilix Virtex FPGA. Moreover, it should be noted that the Virtex FPGAs offer more than high speed for many embedded applications. They offer compact implementation, low cost and low power consumption; things which can't be offered by any software implementation.
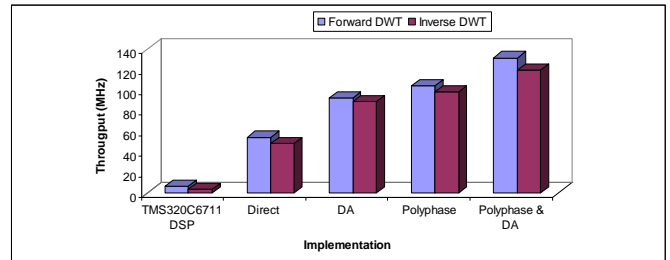
Fig. 7. Throughput comparison between different DWT implementations.

6 Concluding Remarks

In this paper, FPGA implementations of the discrete wavelet transform and its inverse were simulated and realized in a reconfigurable computing hardware board based on the advanced Xilinx Virtex FPGAs. According to the results obtained for the various implementations, we observed that the implementation which was based on the distributed arithmetic and ployphase decomposition techniques achieved the best performance results. We also observed that the performance of the TMS320C6711 digital signal processor was much lower the performance of the least efficient, direct FPGA implementation. One final remark is that the implementation is applicable to image-based applications where the image data is two dimensional. The 2-D transformation is straightforward, and can be easily achieved by inserting a matrix transpose module between two 1-D discrete wavelet transform modules. The 1-D discrete wavelet transform is first performed on each row of the 2-D image data matrix. This is followed by a matrix transposition operation. Next, the discrete wavelet transform is executed on each column of the matrix.

References

[1] C. Burrus, R. Gopinath and H. Guo, *Introduction to Wavelets and Wavelet Transforms.* New Jersey: Prentice Hall, 1998.

[2] O. Riol and M. Vetterli, " Wavelets and signal processing," *IEEE Signal Processing Magazine*, vol. 8, no. 4, pp. 14-38, October 1991.

[3] G. Beylkin, R. Coifman and V. Rokhlin, Wavelets in Numerical Analysis in *Wavelets and Their Applications*. New York: Jones and Bartlett, 1992, pp. 181-210.

[4] T. Ebrahimi and F. Pereira, *The MPEG-4 Book*. Prentice *Hall, July 2002*

[5] M. Smith, *Application-Specific Integrated Circuits*. USA: Addison Wesley Longman, 1997.

[6] C.Chakabarti, M. Vishwanath, and R. Owens, "Architectures for Wavelet Transforms: A Survey," *Journal of VLSI Signal Processing*, vol. 14, no. 2, Nov. 1996, pp.171-192.

[7] G. Knowles, "VLSI architecture for the discrete wavelet transform," *Electron Letters*, vol. 26, no. 15, July 1990, pp. 1184-1185.

[8] K. Parhi and T. Nishitani, VLSI architectures for discrete wavelet transforms, *IEEE Trans. VLSI Systems*,June 1993, pp. 191-202.

[9] C.Chakabarti and M. Vishwanath, "Efficient realizations of the discrete and continuous wavelet transforms: from single chip implementations to mappings on SIMD array computers ," *IEEE Trans. Signal Processing*, vol. 43, no. 3, Mar. 1995, pp. 759-771.

[10] R. Seals and G. Whapshott, *Programmable Logic: PLDs and FPGAs*. UK: Macmillan, 1997.

[11] P. Kollig, B. Al-Hashimi and K. Abbot, " FPGA implementation of high performance FIR filters," In *Proc. International Symposium on Circuits and Systems*, 1997.

[12] M. Shand, " Flexible image acquisition using reconfigurable hardware," In *Proc. of the IEEE Workshop on Filed Programmable Custom Computing Machines*, Napa, Ca, Apr. 1995.

[13] J. Villasenor, B. Schoner, and C. Jones, "Video communication using rapidly reconfigurable hardware," *IEEE Transactions on Circuits and Systems for Video Technology,* vol. 5, no. 12, Dec. 1995, pp. 565-567.

[14] Xilinx Corporation. "Xilinx breaks one million-gate barrier with delivery of new Virtex series," October 1998.

[15] S. White, " Applications of Distributed Arithmetic to Digital Signal Processing: A Tutorial", In *IEEE ASSP Magazine*, July 1989, pp. 4-19.

[16] M. Bellanger, G. Bonnerot and M. Coudreuse, "Digital Filtering By Polyphase Network: Application to Sample Rate Alteration and Filter Banks," *IEEE Acoustic Speech Signal Proc.*, vol.24, April 1976, pp.109-114.

[17] S. Mallat, " A theory for multresolution signal decomposition: The wavelet representation, *IEEE Trans. Pattern Anal. And Machine Intell.*, vol. 11, no. 7, July 1989, pp. 674-693.

[18] I. Daubechies, "Orthonomal bases of compactly supported wavelets," *Comm. Pure Appl. Math*, vol. 41, 1988, pp. 906-966.

[19] G. Strang and T. Nguyen, *Wavelets and Filter Banks*. MA: Wellesley-Cambridge Press, 1996.

[20] A. Oppenheim and R. Schafer, *Discrete Signal Processing*. New Jersy: Prentice Hall, 1999.

[21] P. Vaidyanathan, *Multirate Systems and Filter Banks*. New Jersey: Prentice Hall, 1993.

[22] L. Mintzer, "The Role of Distributed Arithmetic in FPGAs," Xilinx Corporation.

[23] J. Bhasker, *A Verilog HDL Primer*. PA: Star Galaxy Publishing, 1997.

[24] Xess Corporation. www.xess.com.2002.

[25] Xilinx Corporation. *Virtex Data Sheet*, 2000.

[26] N. Kehtarnavaz and M. Keramat, *DSP System Design Using the TMS320C6000*. New Jersey: Prentice Hall, 2001

[27] Texas Instruments Corporation. *TMS320C6711 Data Sheet*, 2000.