Middle-East Journal of Scientific Research 23 (5): 957-961, 2015
ISSN 1990-9233

DOI: 10.5829/idosi.mejsr.2015.23.05.22224

Review of Memory Efficient 2-D Finite Impulse Response (FIR) Filter Architecture

[1] *P.G Punitha and* [2]*R. Ramesh*

[1]M.E VLSI Design, Saveetha Engineering College, Chennai, India
[2]Professor, ECE Depatment, Saveetha Engineering College, Chennai, India

Abstract: Images which are blurred can be enhanced by using Image filtering Technique. This technique can be performed by using 2-D FIR filter. 2-D FIR filter can be extensively utilized in Image Processing, Video Processing, Biomedical Signal Processing (especially Medical Image Processing) fields. Filtering the distorted medical images which are not clear may help the technician or a doctor to diagnose the patients' diseases with the processed noise free image. Many literatures have been surveyed which describe the various methods of implementation of the filtering technique. Full Insight of these techniques can be obtained by doing an exhaustive study of all the literatures related to it. The concepts learnt have been presented below. All these concepts rely on the fact that key terms of VLSI (area, speed and power) are influenced by the combinational complexity of the filter. Based on these facts, the proposal of the survey makes the concentration to turn towards memory. Memory complexity comprises of the major portion of the filter design. The solution arrives after the analysis that tends to reduce the memory complexity which involves in the reduction of the area and power consumption.

Key words: VLSI · Image Processing · Biomedical Signal Processing · Memory Complexity

INTRODUCTION

A filter is a device that eliminates unwanted components from asignal. Filtering is a class of signal processing, the defining feature of filters being the complete or partial filtering of some aspects of the signal. Digital filter is a structure that performs mathematical operations on a sampled, discrete-time signalto change certain aspects of that signal. Digital filters can be classified into two types based on its impulse response. They are: 1.Infinite Impulse Response (IIR), 2.Finite Impulse Response (FIR). FIR (Finite Impulse Response) filter is a filter whose impulse response is of finite duration. 2-D FIR filter performs filtering in two dimensions (both in x and y directions). 2-D FIR filter has been extensively used in the applications of image processing, video processing fields such as image enhancement, image restoration, template matching, face recognition, feature extraction for bio-metric systems. 2-D FIR filter is more popular than IIR filter, because it has its own advantages. It is stable, more accurate and has linear phase [1].

Literature Overview: Many literatures have been reviewed and the main scope of the work is illustrated below. Different techniques studied are stated below. The paper is organized as follows: Systolic architecture and its various forms are presented in section III. Section IV is presented with symmetry architecture, Section V explains the arithmetic techniques used in filters to reduce its complexity [2].

Systolic Architecture: Systolic architectures of 2-Dimensional recursive digital filters are constructed on the basis of Local State Space (LSS) model. This approach requires more no. of multipliers. A minimization procedure is adapted to reduce the no. of multipliers equal to that of Transfer function model. Systolic architecture derived directly based on Transfer function model is suggested. This technique overcomes the disadvantage of extra procedure required to reduce the no. of multipliers. Processing Elements (PEs) utilized in Transfer function model are much simpler than other. Systolic architectures have a speed advantage which leads to a flexible structure co-operative to VLSI design. The implementation of 2-D

Corresponding Author: P.G Punitha, M.E VLSI Design, Saveetha Engineering College, Chennai, India.

FIR filter requires finding the solution for 2-D convolutional problem. Although solutions are available, it requires data in non-raster scan format and cascading of these structure is complex. This structure is designed which allows the input data to be entered in row-wise fashion which can be transmitted to different Processing Elements (PEs). To build the entire structure, many no. of PE's are required in accordance with the order of the filter.

Systolic realization can be improved by performing a transformation, which is used to reduce the errors that occur during quantization. The systolic realization can be represented in Signal Flow Graph (SFG), which can be transformed by a technique called Systolic Transformation (ST) [3]. This can be done by rearranging the delay elements presented in the structure in such a way that should perform parallel realization. In this ST technique it introduces two delay elements for every delay element in canonical SFG. This systemization technique requires more no. of delay elements. The present architecture developed using ST is compared with the previous systolic architectures based on the parameters such as number of adders, multipliers, registers, the clock period, the latency and the speed-up factor (SUF) for both 2-D IIR and 2-D FIR filter. Latency is reduced to zero than previous existing architectures. A complete error analysis is performed which concludes that our architecture has least storage error.

The recursive equations are derived in order to accelerate the level of parallelism in the realization of 2-D FIR filter and 2-D Linear Phase FIR filters. As a consequence of the newly derived recursive equations, the fully systolic array architecture for high throughput has been developed [4]. It would be ideal to obtain L times more throughputs by using L times more hardware. The proposed architecture is designed in a flexible manner in such a way that, if high throughput is needed (L value may be high) more hardware will be required, otherwise less hardware is used. Thus the cost effective and flexible architecture is realized for processing the real time images as per their needs.

The combination of revised reordering scheme and new systolic transformation leads to a new systolic architecture of both 2-D IIR and FIR filters with local broadcast, zero latency and lower quantization error [5]. The fusion of these two techniques utterly excludes global broadcasts of both input and output signals i.e eliminates global broadcast in both i and j dimensional paths. The critical period of the filter is adequate which is not violating the hardware requirements (no. of multipliers and delay elements).

Systolization of filters is done to attain high throughput and low latency implementation. Systolization requires more hardware structures to realize the filter. It requires more delay elements which leads to the global broadcast of signals. This leads to increased consumption or requirement of VLSI design parameters (area, speed, power).

Symmetry Architecture: Symmetry is the property of the magnitude and frequency response of the filter design which is used to lessen the hardware and time complexities of the filter design. Symmetry property of the impulse (frequency) responses in time (frequency) domain is used to reduce the complexity of the design. The restriction imposed on the filters' Transfer function coefficients by the symmetry property of the filter produces the filter with less design complexity.

Symmetry architectures in 2-D Filter possess various magnitude symmetries such as: 1. Diagonal Symmetry, 2. Quadrantal Symmetry, 3. Four fold rotational Symmetry, 4. Orthogonal Symmetry [6]. This property is applied to both IIR and FIR filters.

For a 2-D polynomial $Q(z_1,z_2)$, the magnitude squared function of the frequency response of the form $F(\theta_1,\theta_2)=Q\left(e^{j\theta_1},e^{j\theta_2}\right).Q^*\left(e^{-j\theta_1},e^{-j\theta_2}\right)=Q(z_1,z_2).Q^*(z_1,z_2)$, where $z_i=e^{j\theta_i}$ for $i=1,2$. Q^* is the complex conjugate of the co-efficients of Q.

Diagonal Symmetry: The magnitude squared function of the diagonal symmetry function can be defined as $F(\theta_1,\theta_2)=F(\theta_2,\theta_1)=F(-\theta_1,-\theta_2)=F(-\theta_2,-\theta_1)$. The property of real polynomials satisfies that, then $F(\theta_1,\theta_2)=F(\theta_2,\theta_1)$ is sufficient to prove the diagonal symmetry.

Fourfold Rotational Symmetry: The magnitude squared function of the four fold rotational symmetry function can be defined as $F(\theta_1,\theta_2)=F(-\theta_2,\theta_1)=F(-\theta_1,-\theta_2)=F(\theta_2,-\theta_1)$. The property of real polynomials satisfies that $F(\theta_1,\theta_2)=F(-\theta_1,-\theta_2)$ and then $F(\theta_1,\theta_2)=F(-\theta_2,\theta_1)$ is sufficient to prove the four fold rotational symmetry.

By using diagonal symmetry property, design technique is derived for both 2-D IIR and FIR filter design.

In this proposal, no of variables and latency are the two factors which are of great importance. In this, the total no of variables required to implement the filter gets

reduced from 17 $(N=9, D=8)$ to 11 $(N=6, D=5)$ due to the diagonal symmetry property as shown in the matrices below.

$$N(z_1,z_2)=\begin{array}{c} \\ z_1^0 \\ z_1^{-1} \\ z_1^{-2} \end{array}\begin{array}{c} \begin{array}{ccc} z_2^0 & z_2^{-1} & z_2^{-2} \end{array} \\ \begin{bmatrix} a_{0,0} & a_{0,1} & a_{0,2} \\ a_{1,0} & a_{1,1} & a_{1,2} \\ a_{2,0} & a_{2,1} & a_{2,2} \end{bmatrix} \end{array}$$

$$D(z_1,z_2)=\begin{array}{c} \\ z_1^0 \\ z_1^{-1} \\ z_1^{-2} \end{array}\begin{array}{c} \begin{array}{ccc} z_2^0 & z_2^{-1} & z_2^{-2} \end{array} \\ \begin{bmatrix} 1 & b_{01} & b_{02} \\ b_{10} & b_{11} & b_{12} \\ b_{20} & b_{21} & b_{22} \end{bmatrix} \end{array}$$

According to the diagonal symmetry property, register reordering takes place which recasts the older form and implements the filter with (M-P) shift registers [7]. The two different architectures such as Type-I and Type-II are proposed. Type-I (with P=0) architecture experiences global broadcast and has fewer no. of multipliers, higher critical period. Since the critical period is large and the signals are broadcasted globally, it is destructive to the restricted fan-out VLSI design. To get rid of this problem, the architecture should be improved to higher speed and local broadcast can be obtained by substituting P=1. The architecture derived by this means is called as Type-II filter[8]. Thus new 2-D diagonal symmetry architectures are implemented with less no of multipliers to achieve high throughput and zero latency without the necessity of extra delay elements.

Similarly the four fold rotational symmetry property [9] is embedded in the 2-D digital filter architecture as mentioned above. The anticipated architecture is derived with the constraint applied to its filter co-efficients based on its symmetry property which reduces the complexity in design. The structures obtained are based on register reordering and it creates Type-I structure. To improve the performance in terms of its latency, critical period and with the motto of avoiding the global broadcast incurred in signals, Type-II structures are designed with the alteration of P value with 1 instead of 0 in Type-I.

In this paper a new transposed structure is derived by transposing the original systolic structure [1] which necessitates few delay elements than the original structure. Two new systolic architectures are derived by combining the basic sub-blocks of original structure with new transposed structure. These systolic structures obtained can be realized with separable denominators which require fewer multipliers [8]. Symmetry properties mentioned above can also be applied to the suggested architecture which gives an added advantage of even lesser no. of multipliers used when compared to the original symmetry architectures presented in [7]. Transfer functions with Separable Denominator have the following properties such as:

The assurance of the Stability can be done easily by resolving the poles of two 1-D polynomials separately. Poles which are unstable can be made to attain their stability by substituting with its inverse pole without affecting the magnitude. ii. Requires less no. of multipliers. iii. Separable denominator aids the system to possess all the symmetry properties in magnitude response. Separable denominator structures implied in the original systolic and new transposed structure gains the advantage of requirement of less no. of multipliers for the latter than the earlier structure obtained. The presence of symmetry in magnitude response in the separable denominator leads to the new architecture with fewer multipliers [9]. These structures are derived based on the symmetry property which imposes constraints to the filter coefficients. These structures derived can have lower critical period of $(T_m + 2T_a)$.

2-D Filters without global broadcast are designed by using filter frameworks of 1-D FIR filter. The generalized method is developed that allows the derivation of new 2-D VLSI filter structures by the addition of different filter sub-blocks and interconnections (frameworks) [10]. The sub-blocks of 1-D FIR are made up of direct form and lattice form by applying the sub-blocks in various frameworks, transfer functions of 2-D filter structures is obtained. The filter formulation using sub-blocks and interconnections uses special arrangement of delays to exclude the global broadcast of the signals and to maintain the critical period. Sub-blocks of lattice form does not have any constant term, so realization of bottom sub-block should be in direct form. The lattice form needs more delay elements which produce z^{-n} latency in the sub block. This latency can be exempted in the final reconfiguration of the structure during the entire filter frame work construction. Filter interconnection frame work means constructing the 2-D filter template using the 1-D filter sub blocks and shift registers in such a way that it realizes 2-D Filter transfer function. The filter framework using separable denominator is constructed. This method is more advantageous, which reduces the multiplier

requirement.The symmetry applications are also included in the filter interconnection framework construction. Using quadrantal symmetry property the no of multipliers required is very much less. The filter framework using direct form has the lowest round off noise and their benefit increases with the filter order. The lattice-type possesses better multiplier sensitivity.

Two dimensional VLSI filter architectures are designed to possess multiple symmetry in a single filter rather than using four different symmetry filter structure [2]. Even though each symmetry filter consumes less chip area, for using the anticipated filter in a multiple symmetry system it requires very low chip area than the overall chip area. Actually in this paper, Type–I and Type-II (obtained by taking the transpose of Type-I structures) for all the four symmetries totally 8 structures have been suggested. The transfer function is expressed by the separable form which is to split the filter structure into the two blocks as block1 and block 2 for the ease of implementation. The recommended filter structure also requires less power for four symmetry modes. Though each distinct symmetry filter structure has compact area and lesser power consumption than the proposed filter structure, these individual structures are single-mode systems. So by using multimode filter structure, we can save larger chip area when the four symmetric individual structures are required in a multi-symmetry system. Compared to the sum of the chip areas of the four individual filters, the chip area could be saved up to 63.25%. Therefore, the proposed architecture is power and cost effective. The proposed multimode filter can yield up to 63.25% of area when compared to sum of areas of the four individual (Type-1) symmetry filters [11].

Arithmetic Techniques: Multiple Constant Multiplication (MCM) is a method in which the input samples are multiplied with the set of constant coefficients. MCM technique can be realized in Multiplier less filter implementation by using additions, subtractions and shift operations. Algorithms suggested for MCM technique are very fast but its performance is worst. In order to overcome the disadvantages, a new Distributed Arithmetic (NEDA) Architecture is proposed [12]. It diminishes both power consumption and area, but it preserves the high speed and accuracy which inherits the novelty of DSP applications. In NEDA the filter coefficients are applied in parallel fashion which is contrast to the distributed arithmetic technique where the coefficients are applied serially. The benefit of using the technique (NEDA) to perform inner product computation is the use of simpler structure. In this technique the adder structure is replaced with butterfly matrix, which is a sparse matrix. The matrix represents the fixed set of filter coefficients having 0's and 1's. The adder butterfly matrix has the opportunity of redundancy which could be reduced by the matrix compression scheme. The compression techniques look for the similar rows and columns and compress to a single row and column respectively interpreting that the outcome of the other rows (columns) can be obtained from the same adder array. Then, it looks for the pairs of rows (columns) which share maximum no. of 1's. Likewise the search of the algorithm continues until there is no common pattern. The NEDA Architecture is suitable for hybrid filters. The grouped filter co-efficients are made into separate blocks and the compression scheme is applied. As a result, the sub matrices formed lead to a significant reduction in the no. of adders used. The no. of adders and the sub-blocks formed are inversely proportional to each other. NEDA approach uses only adders and shifters contrary to the MCM accomplish the reduction of hardware resource and its modularity[13].

MAC Units present in the FIR filter are replaced by the use of Configurable LUT's to reduce the memory size. As the input word size increases, LUT size also increases. This major problem can be optimized by using Combined APC-OMS (Anti-symmetric Product Coding-Odd Multiple Storage) [14]. In Anti-symmetric Product Coding technique LUT size is reduced by half than the conventional LUT. This technique stores only the upper half of conventional LUT blocks. In Odd Multiple Storage (OMS) technique, only odd multiple values are stored in the LUT and the even multiples are obtained by shifting operation. Shifting operation is performed by using Barrel Shifter. LUT size is reduced to half from its original size. The combination of APC and OMS technique will reduce the LUT size by 75% or ¼th from its original size. The main advantage of using this technique is it reduces area through the reduction of LUT size. The power consumption also gets reduced due to the fact that memory access requires less memory.

CONCLUSION

The insight obtained from reviewing all the literatures reveals that many researchers believe in arithmetic or combinational complexity that it has major impact on the key factors of VLSI, mainly area and power. But FIR filter

comprises of two major modules such as 1. Arithmetic module and 2.Memory or Storage module. All the literatures focus mainly on Combinational complexity. The idea that came into existence for the proposed work involves the optimization of memory module by using Memory Reuse and Memory Sharing Technique. Block based filters are used for memory Reuse technique with effective reuse and appropriate scheduling of computation. By using these techniques, Memory usage is reduced which ultimately achieves area and power savings.

Further improvements in area and power savings can be attained by optimizing conventional methods using advanced techniques. It has been planned to use Distributed Arithmetic Technique in 2-D FIR filter for the sake of providing improvements in multiplication. This technique replaces conventional MAC unit by series of LUT accesses and summation. Memory access requires less time than performing inner product computation, which consumes lower power. Also it is suggested to extend the work to memory efficient 3-D FIR filter which processes real time videos and 3-D images with less power and area requirements.

REFERENCES

1. Sid-Ahemed, M.A., 1989. A systolic realization for 2-Ddigital filters, in IEEE Trans. Acoustic Speech Signal Process., 37: 560-565.
2. Chen, P.Y., L.D. Van, I.H. Khoo, H.C. Reddy and C.T. Lin, 2011. in Powerefficient cost-effective 2-D symmetry filter architecture, IEEE Trans. Circuit Syst. I, Reg. Papers, 58(11): 12-125.
3. Sanbhag, N.R., 1991. In An improved systolic architecture for 2-D digital filters, IEEE Trans. Signal Process. 39: 1195-1202.
4. Mohanty, B.K. and P.K. Meher, 1996. in Cost-effective novel flexible cell-vlevel systolic architecture for high-throughput implementation of 2-D FIR filters, Proc. IET Comput. Digit. Tech., 143(6): 436-439.
5. Mohanty, B.K. and P.K. Meher, 1999. In High-throughput and low-latency implementation of bit-level systolic architecture for 1-Dand 2-Ddigital filters, IET Proc. Computer Digital Technique, 146(2): 91-99.
6. Van, L.D., 2002. In A new 2-D systolic digital filters architecture without global broadcast, IEEE Trans. Very Large Scale Integr. Syst., 10(4): 477-486.
7. Chen, P.Y., L.D. Van, H.C. Reddy and C.T. Lin, 2008. In A new VLSI 2-D diagonal-symmetry filter architecture design, in Proc. IEEE APCCAS, Macao, China, pp: 320-323.
8. Khoo, I.H., H.C. Reddy, L.D. Van and C.T. Lin, 2009. In 2-D digital filter architectures without global broadcast and some symmetry applications, in Proc. IEEE Int. Symp. Circuis Syst. ISCAS, pp: 952-955.
9. Chen, P.Y., L.D. Van, H.C. Reddy and C.T. Lin, 2009. In A new VLSI 2-D fourfold-rotational-symmetry filter architecture design, in Proc. IEEE Int. Symp. Circuis Syst. (ISCAS), pp: 93-96.
10. Khoo, I.H., H.C. Reddy, L.D. Van and C.T. Lin, 2010. In Generalized formulation of 2-D filter structures without global broadcast for VLSI implementation, in Proc., IEEE MWSCAS, Seattle, WA, USA, pp: 426-529.
11. Yao, C.Y., H.H. Chen, C.J. Chien and C.T. Hsu, 2004. In novel common subexpression elimination method for synthesizing fixed-point FIR filters, IEEE Trans. Circuit Syst. I, Reg. Papers, 51(11): 2215-2221.
12. Luis, J., T.X., R.M. A.P. and M. Bayoumi, 2007. in Hybrid multiplierless FIR filter architecture based on NEDA, in Proc. IFIP Int. Conf. Very Large Scale Integration (VLSI-SoC 2007), pp: 316-319.
13. Meher, P.K., S. Chandrasekaran and A. Amira, 2008. FPGA realization of FIR filters by efficient and flexible systolization using distributed arithmetic, in IEEE Trans. Signal Process., 56(7): 3009-3017.
14. Meher, P.K., 2010. New approach to look-up table design and memorybased realization of FIR digital filters, in IEEE Trans. Comput.-Aided Design Integr. Circuits Syst., 57(3): 592-603.