High performance matrix multiplication using exact systolic array with Wallace multiplier-A Survey

[a]Praphulla T R, [b]Vinutha C B, [c]M Z Kurian
[a] M.Tech Student, VLSI and embedded system, Dept. of ECE, SSIT, Tumkur, India, 572105.
[b] Assistant Professor, Dept. of ECE, SSIT, Tumkur, India, 572105.
[c] HOD, Dept. of ECE, SSIT, Tumkur, India, 572105.

Abstract

Multiplication is most frequently and commonly used operation in arithmetic and logical computations, multiplication based operations like MAC is very much used in computation of DSP application. In DSP algorithm such as convolution, FFT, filtering and microprocessor, maximum power consumption takes place due to multiplication process, thus this is very much essential to have high speed multiplier. In addition to this, multiplication process dominates the execution time specifically. In some applications like image processing, matrix multiplication is widely used, matrix multiplication is one in which two matrix are multiplied by considering row of one matrix with the column of another.

To result with fast performance of matrix multiplication one of the method called systolic array is very popular. Systolic array is one in which data flows synchronously across the array between two processes and is a form of pipelining. We focus to improve speed of execution by employing Wallace multiplier, which use the concept of column compression and pipelining specifically for larger operands. In contrast to other multipliers such as booth multiplier and array multiplier, Wallace multiplier improves the processing steps and hence the speed of processing as it is a type of tree based multiplier. In this research paper a detailed investigation of improving the system performance in terms of processing steps, speed of operations, delay, scalability, and cost is presented.

KEYWORDS- Matrix multiplication, Parralel processing, pipelining, Systolic array, Wallace multiplier.

I. Introduction

In the past decades, VLSI system design is evaluated with its performance, specifying the Response time, power consumption, scalability, and the cost. Multiplication is the basic fundamental aspect in both arithmetic and logical operations. There are different types of multipliers Booth multiplier, array multipliers, combinational multiplier, sequential multiplier, Wallace tree multiplier. An efficient multiplier should give correct results that is accuracy, should perform operation at high speed, should consume less power, and it should accupies less number of steps. There are three main steps in multiplier process,

1. Partial product generation.
2. Partial product reduction.
3. Final addition.

In DSP application most of the power consumed by the multiplier, Therefore High speed multiplication is necessary. Multiplication dominates the execution time of most DSP applications because of computational performance in DSP algorithm [1];

Systolic array is a MAC protocol consists of adder, multiplier and accumulator. In parallel computer architectures, a systolic array is a homogeneous network of tightly coupled data processing units called cellsor nodes. Systolic algorithm is the form of pipelining, which is in more than one dimension. In this array, data flows from a memory in a rhythmic way and it passes through many processing elements before it returns to memory. Systolic arrays have balanced, uniform, grid like architectures in which each line indicates a communication path and each intersection represents a cell or systolic elements[2]; Execution time for Parallel multiplication has been decreased by using the technique of column compression. There are two classes of parallel multipliers, which are array multipliers and tree multipliers. Tree multipliers are also known as column compression multipliers. Column compression multipliers are faster than array multipliers [8];
The Wallace tree multiplier belongs to a family of multipliers called column compression multipliers. Wallace multiplier is an efficient parallel multiplier. A Wallace tree multiplier offers faster performance for large operands. In Wallace‘s scheme, the partial products are reduced as soon as possible [8];

In computer architecture, a systolic architecture is a pipelined network arrangement of Processing Elements (PEs) called cells. It is a specialized form of parallel computing, where cells compute the data which is coming as input and store them independently. In systolic multiplication, to carry out the multiplication and get the final product, the multiplicand and multiplier are arranged in the form of array.

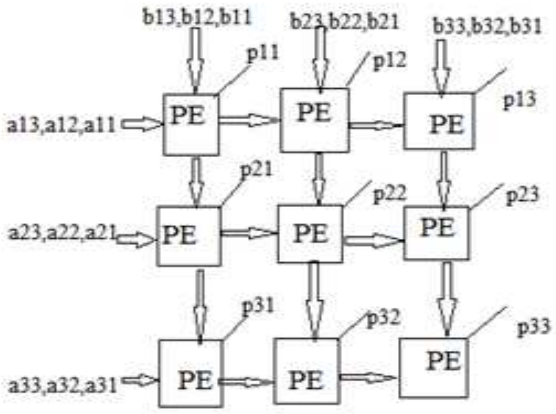
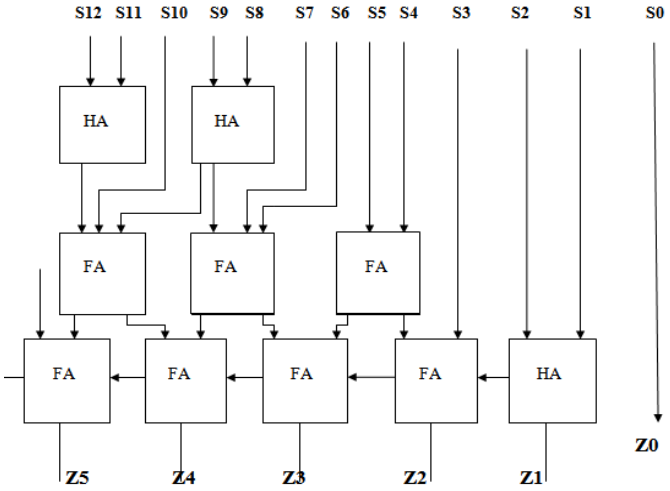


Figure 1: architecture of systolic array.

In the SA of the PEs are interconnected in a two dimensional array with communication between PEs occurring in cycles. The ith column incurs in a delay of i-1 cycles. Each PE stores the accumulated sum P (i,j). After 2N+1 operational cycles, the stored P (i,j) value forms the resulting product matrix P [2]; Systolic design for matrix multiplication is proposed to cater for low latency and high performance for high level applications. There is steps wise improvements in conventional multipliers where as in Wallace multipliers is used for non square multiplications and it’s done the performance in single step. In systolic design the design of multipliers is always above GF (2^m) and it is very suitable.



Figre 2 architeture of Wallace multiplier

Design of high level applications like digital filters design, image processing, non-numeric applications, complex arthemetic operations necessisites fast response, scalable and low power algoritms.The combinational design techniques are not suitable for such high level applications. The systolic array with Wallace multiplier design approach is one of the solutions to such issues [8];

II. Literature Review

1. “Synthesis of orthogonal systolic array for fault tolerant matrix multiplication”.

In [3] M K Stojcev, E Milovanovi’c, S R Markovi’c, proposes the systolic array for tolerate the fault in matrix multiplication.

For matrix multiplication, procedure for designing fault tolerant systolic array with orthogonal interconnect and bidirectional data flow. To achieve fault tolerance, method employs space time redundancy. Systolic array can tolerate single transient error and the majority of multiple error patterns with high probability.

The number of inputs and output pins is reduced to thirty percentages approximately compare to hexagonal array.

Systolic array is a type of special purpose system, it can be used to fill the desired throughput rate for time critical and computationally intensive problems, special purpose, high-speed computing system optimized for processing a specific task.

The main source of error in integrated circuit is permanent faults, transient faults and intermittent faults.

Permanent faults are due to irreversible physical changes, typical example faults are shorts and opens.

Transient faults are generated by environment conditions like cosmic rays. These transient faults are most common and their number is increasing due to high complexity, smaller transistor size, high operational frequency etc.

For masking error, Varity of techniques have been designed, in the error masking N-type modular redundancy majority voters are used [3];

In this system at least three modules are necessary called triple modular redundancy (TMR).

In triple modular redundancy at least two hundred percentage hardware overhead for fault tolerance is need.

In our approach we generate redundant information by means of hardware, involving two extra column of processing element in 2D-array and time redundancy along with

2D array, bidirectional orthogonal array is used drive three equivalent algorithm with disjoint space.
In this single transient error can be detected and corrected.
For matrix multiplication, synthesize fault tolerant systolic array with orthogonal interconnect and bi-directional data flow is performed to tolerate the single transient error and majority of multiple faults.

2. "Configurable systolic matrix multiplication".
In [4] Parastoo Kamranfar, S Ali Shahabi, Ghazaleh Vazhbakht, and Zainalabedin Navabi proposes Design and implementation of new matrix multiplication module using configurable systolic array.
New matrix multiplication hardware can be reconfigured easily to accept a pair of input matrix that are allowed to multiply, the proposed hardware multiplies both square and non-square matrix with scalable systolic architecture to enhances the speed in terms of clock cycle compared to previous work.
The new matrix multiplication is to multiply non-square multiplication and re-configurability and this configurable systolic matrix multiplication is used in high level system applications such as digital filter and image processing.
To reduce the area, to improve the speed and minimize power consumption in matrix multiplication, most of them considers square matrix and it is not suitable everywhere.
To properly extract the filter output from the given input sequence, a set of non square matrix multiplication operation will be used and number of required processing elements is constant in non square matrix multiplication.
In Square Matrix multiplication the required number of processing element is increased.
In this paper a configurable matrix multiplication is presented and it is able to multiply two valid input matrix and able to enhance the computation speed [4];

3. "Matrix-vector multiplication on a fixed size unidirectional systolic array".
In [5] E I Milovanovic, N M Stojanovic, I Z Milovanovic, M K Stojcev, proposes how to overcome the problem of multiplication, using a vector multiplication on unidirectional linear systolic array comprises of P<[n/2].
To match dimensions of matrix multiplication to unidirectional linear systolic array size, the partitioning of matrix into quasi diagonal is performed and memory interface subsystem architecture is proposed and it is located between host and unidirectional linier systolic array (ULSA).
It provides corresponding data transfer to ULSA or from ULSA and performance of synthesized ULSA is discussed, and compared with fixed size of bidirectional systolic array.
This type of systolic array is always less number of input and output pins required for input and output communications.
In order to design the ULSA, the number of PE's of index space of an algorithm projection's directions is performed.
In this paper, to realize the matrix multiplication by vector on ULSA comprising of (P<n/2) where P<n, and goal is to minimize the execution time with the given number of processing element in ULSA [5];

4. "Matrix multiplication by inexact systolic array".

In [6] Ke Chen, Fabrizio Lombardi, Jie Han, proposes matrix multiplication by inexact systolic array, to compute the approximate matrix multiplication systolic array with different scheme are presented.
The processing element utilizes the inexact full adder cell for the multiplier and the final adder implementation using two computational steps in matrix multiplication.
In this paper, with respect to circuit level performance three inexact schemes for the processing element are pursued such as delay, power consumption and number of transistors.
The extension of inexact computation to a systolic array, it is also performed with limited error, the execution of matrix multiplication in each PE, it results in inexact computation. The main application of the proposed inexact systolic array is discreet cosine transform. With small error, simulated results show that proposed inexact systolic array is very effective and this approximate computing reduces circuit complexity and delay [6];

5. “Design and implementation of VLSI systolic array multiplier for DSP application”.
In [2] Bairuk Saptalakar, Deepak Kale, Mahesh Rachannavar, Pavankumar M K proposes asimplementation of systolic arrayfor DSP applications. We know multiplication is most commonly used operations in mathematics like arithmetic and logical operations. Integer multiplication is commonly used in real word binary multiplication is the basic multiplication used for the integer multiplication.
To perform the binary multiplication the systolic array is an efficient one.
Execution time of the multiplication is dominates in DSP algorithm, so high speed of multiplier is needed to reduce the delay.
The application of systolic array DSP is convolution, FFT filtering and in microprocessor, in this multiplication is performed in array manner, to get the partial products.
To get the partial products the multiplication and multiplier are arranged in the form of array, each bit of multiplication is multiplier.
Multiplication is one of the important function in DSP such as convolution, FFT filtering and in microprocessor. For this reason systolic array plays important role in multiplication in DSP application [2];
There will be number of advantages in DSP while using systolic array.
6. “Delay-power performance comparison of multiplier in VLSI circuit design”.
In [7] Sumit Vaidya and Deepak Dandekar proposes the technical aspects of fundamental multipliers, Power, area and delay are the three main important consideration of the VLSI design.
Booth multiplier, array multiplier, Wallace multiplier, Vedic multipliers are some important multiplication types used in VLSI.
In terms of speed and power, there are many proposed logic or low power dissipation, high speed and each logic style has its own advantages.
ARRAY MULTIPLIER: Array multiplier is a combinational multiplier with efficient layout. Using a combinational circuit multiplication of two binary number can be obtained with one micro operation that forms the product bit. Array multiplier mainly focuses for shorter operands multiplication not for larger operands.
WALLACE MULTIPLIER: Wallace is fast processor for multiplication of two numbers. Two multiply these two numbers less steps are used compare to other multipliers. In this multiplier, in the matrix the bit of product is reduced to two row

matrix where sum of the bit products is equal to the sum of rows, the fast adder is summed with resulted row to get a final product.

BOOTH MULTIPLIERS: The main improvement in the booth multiplier is reducing the number of partial products, in this multiplier to reduce the number of partial products; it scans the three bits at a time.

The scanned three bits are one from high order bit of an adjacent lower order pairs, anther two bits from the present pair.

VEDIC MULTIPLIER: To reduce the typical calculations in conventional mathematics to simple one, Vedic multiplier is used[7];

	ADVANTAGES	DISADVANTAGES
1.Orthogonal systolic array	Reuction in the number of processing element	Time overhead due to involment of fault tolerance
2.Configurable systolic array	Multipliy both squre and non squre matrix also enanches the computation speed	Complexity
3.Unidirectional systolic array	Combines both matrix and vector for Fixed Size	Reqires more processing time
4.Inexact systolic array	Used in DSP applications	Not suitable for larger operands

Table 1: Advantages and Disavantages of litreture survey papers.

III. Conclusion

In this paper, To meet the requirement of minimum delay, high speed and powerful data has great influence in current technology.The concept of pipelining is used to overcome a problem of complexity For fast architecture and high scalability systolic array used and this systolic array is a form of pipelining. This systolic array reduces a delay, increases the speed, increase the efficiency in memory. By this parameter, performance of overall project is good. Along with systolic array Wallace multiplier is used to overcome a disadvantage of more delay, less speed and this Wallace multiplier uses the systolic array for larger operands.

References

[1] “A Review of Diffrent types of multipliers and multiplier accumulator unit” Saoniya, Suresh Kumar Intenational Journal of Emergimg Trends and Technology in computer science -2013

[2] “Design and implementation of VLSI systolic multiplier for DSP applications” Deepak kale, Mahesh Rachannavar, pavankumar M. K International Journal of scientific Engineering and Technology (ISSN: 2277-1581) Volume 2 Issue 3 -2013

[3] “Synthesis of orthogonal systolic arrays for fault-tolerant matrix multiplication” S.R.Markovi’c, I.Z. Milovanovic’c PROC. 27th International conference on microelectronics, IEEE 2010

[4] “Configurable systolic matrix multiplication” Parastoo karnarnfar, s. Ali shahabi, Ghazaleh vazhbakht, and zainalabedin Navabi CAD research group, School of electrical and computer engineering, collage of engineering, University of Tehran Iran IEEE 2014.
[5] “Matrix-vector multiplication on a fixed size unidirectional systolic array” N.M. stojanovic, I.Z. milovanovi, M.K. stojecev, E.I. Milovanovic IEEE 2007.
[6] “Matrix multiplication by an Inexact systolic array”, Ke Chen, Fabrizio Lombadi, Northeastren University Boston. Jie Han, University of Alverta Edmonton, Canada IEEE 2015.
[7] “Delay-power performance comparision of multiplier in International Journal of Computer Networks & Communications (IJCNC), Vol.2, No.4, July 2010.
[8] “Desgin and Analysis of a conventional wallace multiplier in 180nm^2 CMOS technology Inamul Hassain, Manishkumar- IOSR Journal of VLSI and Signal Processing (IOSR-JVSP) Volume 5, Issue 1, (Jan - Feb. 2015), PP 60-65 e-ISSN: 2319 – 4200, p-ISSN No: 2319 – 4197.