FPGA Based High Speed, Low Power Matrix Multiplier using Urdva Tiryakbhyam Algorithm

Pranaji J. Rathod¹, S. S. Mungona²
ME student¹, Professor²
Sipna College of Engineering and Technology, Amravati, India

Abstract:
Matrix multiplication is a computational intensive operation and plays an important role in many scientific and engineering applications such as image processing, discrete signal processing. This paper presents architecture for the multiplication of two matrices using Field Programmable Gate Array (FPGA). This paper presents unsigned two 3x3 High-Speed matrix multiplier The hierarchical structuring has been used to optimize for multipliers using “Urdhava Trigyaagbhyam” sutra (vertically and crosswise) which is one of the sutas for Vedic mathematics. Each element of matrix is represented by 4-bit, output is of 8-bit. The coding has been done using VHDL and synthesized using Altera Quartus II. A concept of design is hierarchical structuring; this gives less computation time for calculating the multiplication result. We will synthesize the proposed designs and the existing design using Altera Quartus II tools. The proposed structure consumes less energy.

Key words: FPGA, Matrix multiplier, vedic, Urdhava Trigyaagbhyam, VHDL

I. INTRODUCTION

Matrix multiplication is frequently used operation in a wide variety of graphics, image processing, robotics, and signal processing applications. The increases in the density and speed of field-programmable gate arrays (FPGAs) make them attractive as flexible and high-speed alternatives to DSPs and ASICs. It is a highly procedure oriented computation, there is only one way to multiply two matrices and it involves lots of multiplications and additions. But the simple part of matrix multiplication is that the evaluation of elements of the resultant elements can be done independent of the other, this point to distributed memory approach. In this paper, we propose an architecture that is capable of handling matrices of variable sizes our designs minimize the gate count, area, improvements in latency, computational time, and throughput for performing matrix multiplication and reduces the number of multiplications and additions hardware required to get the matrices multiplied on commercially available FPGA devices. Matrix Multiplication is a frequently used kernel operation in a wide variety of graphics, image processing, robotics, and signal processing applications. Several signal and image processing operations can be reduced to matrix multiplication. Most of the previous work on matrix multiplication on FPGAs focuses on latency optimization. The proposed architecture realized on FPGA which is based on Vedic Multiplication sutra (algorithm) “Urdhava Trigyaagbhyam”. The implementation architecture is also use of IP CORE (Intellectual Property) for adders which allows us to be optimized for speed and space. The objective of this paper is to propose a low area, speed, energy efficient, maximum running frequency, low power, matrix multiplier.

II. LITERATURE REVIEW

A Number of experimental investigations have been reported for matrix multiplication and their blends to improve speed as well as reduce power consumption. The findings of few such studies are presented below. This paper presented the design methods and Field Programmable Gate Array (FPGA) implementation of matrix multiplier in image and signal processing applications. This paper gives the detailed review of matrix product implementation on FPGA where area, energy, speed, power dissipation, latency and time efficient methods are compared for different Xilinx families. Here also presented parallel architecture and parameterized system for matrix product used in FPGA. In this way, this literature will help in making performance characterization to implement high performance matrix multiplier on FPGA. [1] Paper presented new algorithms and architectures for matrix multiplication on configurable devices. These have reduced energy dissipation and latency compared with the state-of-the-art field-programmable gate array (FPGA)-based designs. By profiling well-known designs, they identify “energy hot spots,” which are responsible for most of the energy dissipation. Based on this, they develop algorithms and architectures that offer tradeoffs among the number of I/O ports, the number of registers, and the number of PEs. To avoid time-consuming low-level simulations for energy profiling and performance prediction of many alternate designs, they derived functions to represent the impact of algorithm design choices on the system-wide energy dissipation, area, and latency. These functions are used to either optimize the energy performance or provide tradeoffs for a family of candidate algorithms and architectures. [4] This paper presents computations involving matrices form the kernel of a large spectrum of computationally demanding applications for which FPGAs have been utilized as accelerators. Their performances related to their underlying architectural and system parameters such as computational resources, memory and I/O bandwidth. A simple analytic model that gives an estimate of the performance of FPGA-based sparse matrix vector and matrix-matrix multiplication is presented, dense matrix multiplication being a special case. The efficiency of existing implementations is compared to the model and performance trends for future technologies examined. [6] Paper presented 2x2 High-Speed matrix multiplier using Virtex 5 ML 507 Evaluation Platform (xc5vfx70t-iff1136) FPGA. The
hierarchical structuring has been used to optimize for multipliers using “Urdhava Triyagbhyam” sutra (vertically and crosswise) which is one of the sutras for Vedic mathematics. Each element of matrix is represented by 16-bit. The coding has been done using VHDL and synthesized using Xilinx 13.2. A concept of design is hierarchical structuring; This gives less computation time for calculating the multiplication result.[7] Paper presented Matrix multiplication is the kernel operation used in many transform, image and discrete signal processing application. They developed new algorithms and new techniques for matrix multiplication on configurable devices. In this paper, they have proposed three designs for matrix-matrix multiplication. These design reduced hardware complexity, throughput rate and different input/output data format to match different application needs. These techniques have been designed implementation on Virtex-4 FPGA. They have synthesized the proposed designs and the existing design using Synopsys tools.[8] Paper presented FPGA-based hardware realization of matrix multiplication based on distributed memory approach architecture. They propose an architecture that is capable of handling matrices of variable sizes our designs minimize the gate count, area, improvements in latency, computational time, and throughput for performing matrix multiplication and reduces the number of multiplications and additions hardware required to get the matrices multiplied on commercially available FPGA devices. [9]

III. SYSTEM ARCHITECTURE

The proposed work is aimed at finding high speed and low power consumption of matrix multiplication base on FPGA. Conventional method take more time for matrix multiplication calculation so we study different Vedic math’s sutras for multiplication and use “Urdhva Triyagbhyam Algorithm” which take less timing.

3.1 Conventional System

The common multiplication method is adding and shift algorithm. In parallel multipliers number of partial products to be added is the main parameter that determines the performance of the multiplier. To reduce the number of partial products to be added, Modified Booth algorithm [3] is one of the most popular algorithms. To achieve speed improvements Wallace Tree algorithm can be used, here the number of sequential adding stages can be reduced. However with increasing parallelism, the amount of shifts between the partial products and intermediate sums to be added will increase which may result in reduced speed, increase in silicon area due to irregularity of structure and also increased power consumption due to increase in interconnect resulting from complex routing. On the other hand, serial-parallel multipliers compromise speed to achieve better performance for area and power consumption. The selection of a parallel or serial architecture actually depends on the nature of application. In this lecture we introduce the multiplication algorithms and architecture and compare them in terms of speed, power and combination of these metrics.

Method for 3x3 matrix multiplication:

Two 3x3 matrix multiplies requires 27 multiplications and18 additions. Here, in 3x3 matrix multiplication, each bit is of 4-bit where 4-bit multiplication and 8-bit addition is takes place. For 32-bit adder uses 32 bitIP core (Intellectual Property).

Matrix A and Matrix B 3x3 matrices

\[
A = \begin{pmatrix}
a_{11} & a_{12} & a_{13} \\
a_{21} & a_{22} & a_{23} \\
a_{31} & a_{32} & a_{33}
\end{pmatrix}
\]

\[
B = \begin{pmatrix}
b_{11} & b_{12} & b_{13} \\
b_{21} & b_{22} & b_{23} \\
b_{31} & b_{32} & b_{33}
\end{pmatrix}
\]

MATRIX A X MATRIX B = MATRIX C

\[
C = \begin{pmatrix}
C_{11} & C_{12} & C_{13} \\
C_{21} & C_{22} & C_{23} \\
C_{31} & C_{32} & C_{33}
\end{pmatrix}
\]

C11 = a11b11 + a12b21 + a13b31
C12 = a11b12 + a12b22 + a13b32
C13 = a11b13 + a12b23 + a13b33
C21 = a21b11 + a22b21 + a23b31
C22 = a21b12 + a22b22 + a23b32
C23 = a21b13 + a22b23 + a23b33
C31 = a31b11 + a32b21 + a33b31
C32 = a31b12 + a32b22 + a33b32
C33 = a31b13 + a32b23 + a33

in this case 3x3 matrix multiplication each element in multiplication and addition takes place so it take more time for calculation time delay is occurred.

3.1.1 Array Multiplier

Array multiplier is well known due to its regular structure. Multiplier circuit is based on add and shift algorithm. Each partial product is generated by the multiplication of the multiplicand with one multiplier bit. The partial product are shifted according to their bit orders and then added. The addition can be performed with normal carry propagate adder. N-1 adders are required where N is the multiplier length. In this paper conventional array multiplier is synthesized using 16T full adder cell. Consider the multiplication of two unsigned n-bit numbers, where A=An-1, An-2,….,A0 is the multiplicand and B=Bn-1, Bn-2,….,B0 is the multiplier. The product of these two bits can be written as S7S6S5S4S3S2S1S0, where S0 is the LSB and S7 is the MSB. Figure shows multiplication of a 4x4 array multiplier using carry save adders.

**Figure 1. 4x4 Array Multiplication.**

The conventional array multiplier uses carry save addition to add the products. In the carry save addition method, the first
row will be either half adders or full adders. If the first row of the partial products is implemented with full adders, Cin will be considered “0”. Then the carries of each full adder can be diagonally forwarded to the next row of the adder. The resulting multiplier is said to be carry save array multiplier as the carry bits are not immediately added but rather saved for the next stage of addition. Hence the names carry save multiplier. In the design if the full adders have two input data the third input is considered as zero. The final adder which is used to add carries and sums of the multiplier is removed. Then the carries of the multiplier at the final stage is carefully added to the inputs of the multiplier. The carry of the fourth column of the multiplier is given to the input of the fifth column instead of zero. Then the carry of the fifth column is forwarded to the input of the sixth column and so on.

3.2 Urdhva Triyagbhyam Algorithm:

The Sanskrit word ‘Veda’ means ‘knowledge’. Vedic Mathematics is the name given to the ancient system of Indian Mathematics which was rediscovered from the Vedas between 1911 and 1918 by Sri Bharati Krishna Tirthaji (1884-1960). According to his research all of mathematics is based on sixteen Sutras, or word-formulae. Vedic Mathematics is much simpler and easy to understand than conventional mathematics. Urdhva Triyagbhyam is a multiplication Sutra (Algorithm) in Vedic Mathematics for the multiplication of decimal numbers. We used same idea to the binary numbers to make it compatible with the digital hardware. This is the general formula applicable to all cases of multiplication. The sixteen sutras and their corollaries are as follows:

- Ekadhikina Purvena - By one more than the previous one (Cor: Anurupena)
- Nikhilam Navatashcaramam Dashatah - All from 9 and the last from 10 (Cor: Sisyate Sesamajna)
- Urdhva-Tiryagbhyam-Vertically and crosswise (Cor: Adyamadyenantyamantyena)
- Paraavartya Yojayet-Transpose and adjust (Cor: Kevalaih Saptakam Gunyat)
- Shunyam Saamayasanuuccaye-When the sum is the same, that sum is zero. (Cor: Vestana)
- (Anurupye) Shunyamanyat-If one is in ratio, the other is zero (Cor: Yavadunam Tavadunam)
- Sankalana-yyavakalanabhyam-By addition and by subtraction (Cor: Yavadunam Tavadunikritya Varga Yojayet)
- Puranapuranabhyam-By the completion or non-completion (Cor: Antyayodashake)
- Chalana-Kalanabhyam-Differences and Similarities
- Yaavadunam-Whatever the extent of its deficiency (Cor: Samuccayagunitah)
- Vyashitsamanstih-Part and Whole (Cor: Lopanastha panabhyam)

- Shesanyankena Charamena-The remainders by the last digit (Cor: Vilokanam)
- Sopaantadvayamantyam-The ultimate and twice the penultimate (Cor: Gunitasamuccayay Samuccayagunitah)
- Ekanyunena Purvena-By one less than the previous one (Cor: Dhvajanka)
- Gunitasamuchyay-The product of the sum is equal to the sum of the product (Cor: Dwandwa Yoga)
- Gunakasamuchyay-The factors of the sum is equal to the sum of the factors

Urdhva means Vertical and Triyagbhyam means crosswise. Steps:

a) Vertical Multiplication (LSB):
Multiply LSBs (Least Significant Bits) ‘1’ and ‘0’; place product ‘0’ as LSB of the result.

b) Crosswise Multiplication:
Multiply Crosswise ‘1’ and ‘1’; ‘0’ and ‘1’; add the product terms and place obtained sum as middle term of the result.

c) Vertical Multiplication (MSB):
Multiply the MSBs (Most Significant Bits) ‘1’ and ‘1’, place product ‘1’ as MSB of the result:

Example: a1 a0 * b1 b0

Explanation for 2-bit Multiplication: a1 a0 * a1 a0

\[ a1 \times b1 ; a1 \times b0 + a0 \times b1 ; a0 \times b0 \]

Similarly multiplication can be done for any digital numbers using Urdhva Triyagbhyam algorithm of ancient Indian Vedic mathematics.

\[ 1 \times 1 = 1 \]
\[ 1 \times 0 = 0 \]
\[ 0 \times 1 = 0 \]
\[ 0 \times 0 = 0 \]

Figure 2. Illustration of Vertical and Crosswise algorithm.

3.2.1. 4-bit multiplier:
The 4-bit multiplier is made by using 4, 2-bit multiplier blocks and 3, 4-bit adders. Here the Multiplicands are of bit size (n=4) and the output is of bit size 8. As per the sutra, the 4-bit (n) input is divided into small block of size n/2=2 which reduces the complexity. The newly formed blocks of 2 bits are given as input to 2-bit Vedic multiplier blocks. It will leads to high efficiency multiplier architecture. The output obtained from output of 2-bit Vedic
multiplier block which is of 4 bits is given to the 4-bit adder. For 4-bit adder uses 4-bit IP core (Intellectual Property). The block diagram is shown in

\[
\begin{array}{c|c}
\text{a3} & \text{a1} \\
\text{a2} & \text{a0} \\
\end{array}
\begin{array}{c|c}
\text{b3} & \text{b1} \\
\text{b2} & \text{b0} \\
\end{array}
\]

\text{X} \rightarrow \text{P7 P6 P5 P4 P3 P2 P1 P0 (8-bit Result)}

Figure.3: 4 bit multiplier block

\begin{array}{c|c|c|c|c}
\text{a3 a2 a1 a0} & \text{b3 b2 b1 b0} & \text{P7 P6 P5 P4 P3 P2 P1 P0} \\
\end{array}

Figure.4: 4x4 bit multiplication.

In this case, A1A0 forms AL, A3A2 forms AM, and similarly B1B0 forms BLB2B forms BM of the two 4-bit inputs. Firstly, the vertical result of AL and BL is computed by passing it through a 2x2 vedic multiplier. Then, the crosswise result of (AMxBL + ALxBM) is computed using 2x2 multipliers and a 4-bit adder[8]. Finally the vertical result of AM and BM are computed using a 2x2 multiplier.

3.3. FPGA

A field-programmable gate array (FPGA) is an integrated circuit created to be configured by the customer after manufacturing—hence "field-programmable". The FPGA configuration is generally defined using a hardware description language (HDL), similar to that used for an application-specific integrated circuit (ASIC) (circuit diagrams were previously used to specify the configuration, as they were for ASICs, but this is increasingly rare). FGPGAs can be used to implement any logical function that an ASIC can perform. The ability to update the functionality after shipping, partial re-configuration of the portion of the design and the low non-recurring engineering costs relative to an ASIC design, offer advantages for many applications.

ADVANTAGES OF FPGA

FGPGAs have become very popular in the recent years owing to the following advantages that they offer:

1) Fast prototyping and turn-around time-
2) NRE cost is zero
3) High-Speed
4) Low cost

3.4. VHDL

The VHSIC Hardware Description Language (VHDL) is a formal notation intended for use in all phases of the creation of electronic systems. Because it is both machine readable and human readable, it supports the development, verification, synthesis, and testing of hardware designs; the communication of hardware design data; and the maintenance, modification, and procurement of hardware.

- BENEFITS OF USING VHDL
  - Executable specification
  - Validate spec in system context (Subcontract)
  - Functionality separated from implementation
  - Simulate early and fast (Manage complexity)
  - Explore design alternatives
  - Get feedback (Produce better designs)
  - Automatic synthesis and test generation (ATPG for ASICs)
  - Increase productivity (Shorten time-to-market)
  - Technology and tool independence (though FPGA features may be unexploited)
  - Portable design data (Protect investment)

IV. RESULT

In this chapter we present the test environment and the experimental results of our design modules. Our objectives of this project are to study and design a new method for high speed and low power model of Matrix Multiplication using Urdhava Trigayaagbhyam Algorithm to improve speed performance and time, reduction area. The design is implemented in VHDL, simulated using modelsim and synthesized by ALTERA QUARTUS II.

Simulation result as follows:

1) consider \(a = 15\) \(b = 12\)

hence its product is \(15x12=180\)

Figure. 5: Simulation result conventional method multiplication.
2) consider \( a = 2 \) \( b = 3 \) hence its product is \( 2 \times 3 = 6 \)

Figure 6: Simulation result Urdhava Trigya\-gbyam multiplication algorithm.

3) Consider matrix \( x \times \) matrix \( y = \) matrix \( z \)

\[
X = \begin{pmatrix} 3 & 4 & 2 \\ 2 & 5 & 3 \\ 3 & 2 & 5 \end{pmatrix} \quad Y = \begin{pmatrix} 3 & 4 & 2 \\ 2 & 5 & 3 \\ 3 & 2 & 5 \end{pmatrix} \quad Z = \begin{pmatrix} 23 & 36 & 28 \\ 25 & 39 & 34 \\ 28 & 32 & 37 \end{pmatrix}
\]

Figure 7: Simulation result Matrix multiplication using conventional method.

4) Consider matrix \( x \times \) matrix \( y = \) matrix \( z \)

\[
X = \begin{pmatrix} 3 & 4 & 2 \\ 2 & 5 & 3 \\ 3 & 2 & 5 \end{pmatrix} \quad Y = \begin{pmatrix} 3 & 4 & 2 \\ 2 & 5 & 3 \\ 3 & 2 & 5 \end{pmatrix} \quad Z = \begin{pmatrix} 23 & 36 & 28 \\ 25 & 39 & 34 \\ 28 & 32 & 37 \end{pmatrix}
\]

Figure 8: Simulation result Matrix multiplication using Urdhava Trigya\-gbyam algorithm.
Analysis and synthesis results

All the result has shown below based on ALTERA cyclone II EP2C20F484C7 FPGA platform. From above result also we get the High speed and low power matrix multiplier using Urdhava Trigyagbhyam algorithm.

Table 1: compare between multiplier using Conventional method and Urdhava Trigyagbhyam algorithm

<table>
<thead>
<tr>
<th>Sr. no.</th>
<th>Parameter</th>
<th>Conventional multiplier</th>
<th>Multiplier using Urdhava Trigyagbhyam algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Area</td>
<td>35</td>
<td>33</td>
</tr>
<tr>
<td>2</td>
<td>Time(ns)</td>
<td>70.678</td>
<td>20.166</td>
</tr>
<tr>
<td>3</td>
<td>Frequency (MHz)</td>
<td>14.148</td>
<td>49.588</td>
</tr>
<tr>
<td>4</td>
<td>Power(mw)</td>
<td>68.53</td>
<td>68.52</td>
</tr>
</tbody>
</table>

Table 2: compare between matrix multiplier using Conventional method and Urdhava Trigyagbhyam algorithm

<table>
<thead>
<tr>
<th>Sr. no.</th>
<th>Parameter</th>
<th>Matrix multiplication using Conventional multiplier</th>
<th>Matrix multiplication using Urdhava Trigyagbhyam sutra</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Area</td>
<td>563</td>
<td>542</td>
</tr>
<tr>
<td>2</td>
<td>Time(ns)</td>
<td>3.942</td>
<td>8.854</td>
</tr>
<tr>
<td>3</td>
<td>Frequency (MHz)</td>
<td>107.043</td>
<td>112.943</td>
</tr>
<tr>
<td>4</td>
<td>Power(mw)</td>
<td>84</td>
<td>83.97</td>
</tr>
</tbody>
</table>

Area: Area can be defined as number of logic resources utilized by the FPGA circuit. Area is directly proportional to number of Logical Elements (LE) utilized by the architecture. Mathematically, it can be written as

\[
\text{Area} = \text{Total Number of logic elements utilized by the architecture}
\]

Figure 9: Area for conventional method multiplication.

Figure 10: Area for Urdhava Trigyagbhyam algorithm.

Time: The propagation delay is the execution time required by the architecture after application of input to produce output. It ultimately called as process time required by the device.

\[
\text{Time} = \text{Required Time + Actual Time + From + To + From Clock + To Clock + Failed Paths}
\]

Figure 11: Time for matrix multiplication using conventional method.

Figure 12: Time for Matrix multiplication using Urdhava Trigyagbhyam algorithm.

Here, the propagation delay for both the system is as shown in table where Time for matrix multiplication using conventional method is 9.342 ns and Time for Matrix multiplication using Urdhava Trigyagbhyam algorithm is 8.854 ns. By comparing this above value we found that the propagation delay required for conventional method is more than the using Urdhava Trigyagbhyam algorithm.

Power:

The power required by the device is the summation of core static thermal power dissipation and I/O thermal power dissipation. Mathematically it can be written as,

\[
\text{Total Power} = \text{Core Static thermal power dissipation} + \text{I/O Thermal Power dissipation}
\]

Figure 13: Power for Matrix multiplication using Conventional method.
Figure. 14: Power for Matrix multiplication using Urdhava Trigyaagbhyam algorithm.

Here, power for both the system is as shown in table where power for matrix multiplication using conventional method is 84 mW and power for Matrix multiplication using Urdhava Trigyaagbhyam algorithm is 83.97 mW. By comparing this above value we found that the power required for conventional method is more than the using Urdhava Trigyaagbhyam algorithm.

RTLVIEW:

Figure. 15: RTL view for multiplication using Conventional method.

Figure. 16: RTL view for multiplication using Urdhava Trigyaagbhyam algorithm.

Figure. 17: ALTERA DE1 FPGA board and Quartus II

V. CONCLUSION

Proposed work presents architecture for the multiplication of two matrices using Field Programmable Gate Array (FPGA). This paper presents unsigned two 3x3 High-Speed matrix multiplier. The hierarchical structuring has been used to optimize for multipliers using “Urdhava Trigyaagbhyam” sutra (vertically and crosswise) which is one of the sutra for Vedic mathematics. Each element of matrix is represented by 4-bit, output is of 8-bit. The coding has been done using VHDL and synthesized using Altera Quartus II. A concept of design is hierarchical structuring; This gives less computation time for calculating the multiplication result. We will synthesized the proposed designs and the existing design using Altera Quartus II tools.

VI. REFERENCE


[6]. Colin Yu Lin, Philip H.W. Leong “A MODEL FOR MATRIX MULTIPLICATION PERFORMANCE ON FPGAS” 2011 21st International Conference on Field Programmable Logic and Applications

[7]. Ms.S.V.Mogre1,Mr. D. G.Bhalke2 “Implementation of High Speed Matrix Multiplier using Vedic Mathematics on FPGA” 2015 International Conference on Computing Communication Control and Automation

