

HIGH SPEED PARALLEL MULTIPLIER – ACCUMULATOR (MAC)-A REVIEW

Deepika Setia¹, Mrs.Charu Madhu²

¹Post graduate (M.Tech) in UIET, Panjab University, Chandigarh

²Assistant Professor, UIET, Panjab University, Chandigarh

ABSTRACT

With the rapid advances in multimedia and communication system, high capacity signal processing are in demand . Since most DSP functions are accomplished by repetitive applications of multiplication and addition operations, so High Speed MAC are essential to improve performance of signal processing System. In parallel MAC, Accumulator that has the largest delay in MAC was merged with multiplication to enhance speed. Thus the performance of MAC is thus governed by i) Efficiency of MBA ii) Final Adder which comprises the result of accumulator and multiplication. Many hardware implementations were proposed .However the results are disperse through the literature. In this paper, we survey known methods and techniques for high speed parallel MAC and examine their strength and weaknesses and a new approach for the same is proposed.

Keywords: Modified Booth Algorithm (MBA), Carry Save Adder (CSA), Wallace Tree, Compressors.

1. INTRODUCTION

Increasing demands of high speed data signal processing motivated the researchers to seek fastest processors. The multiplier and multiplier-and-accumulator (MAC) [1] are the building blocks of the processor and has a great impact on the speed of the processor. MAC is the necessary element of the digital signal and image/audio processing system such as filtering, convolution and inner products hence high speed is crucial to develop for real processing applications. Many researchers have attempted in designing MAC for high computational performance and low power consumption. High throughput MAC is always a key factor to achieve high performance digital signal processing applications for real time signal processing applications. Since the multiplier requires the longest delay among the basic operation in digital system, the critical path is limited by the multiplier. Multiplier basically consists of three operational steps: Booth Encoder, Partial product reduction network (Wallace Tree) and final adder.

For high speed multiplication, Modified Booth Algorithm (MBA) [4] is most commonly used, in which partial product is generated from Multiplicand (X) and Multiplier (Y) .Booth multiplication allows for the smaller ,faster multiplication circuits through encoding the signed bits to 2’s complement which is also the standard technique in chip design and provide substantial improvement by reducing the partial products. Although the partial products are further reduced by using higher radix (4, 8, 16, 32) Booth Encoder which increases complexity and improves the performance [1].

In order to increase the speed, two major bottlenecks are considered .One is partial product reduction technique that is used in multiplication block and other is the accumulator. Both of the stages require the addition of large operands that require long paths for carry propagation. Wallace Tree [3] and compressors are used to add the partial products as parallel as possible; its process time is proportional to $O(\log_2 N)$, where N is the number of inputs. In the real implementations (3:2), (4:2), (7:3) counters are used to reduce the output in each pipeline step.

The last step is the final addition in which final result is produced by addition of sum and carry. If the process to accumulate is also considered then a MAC consists of four steps which are shown in Figure.1.

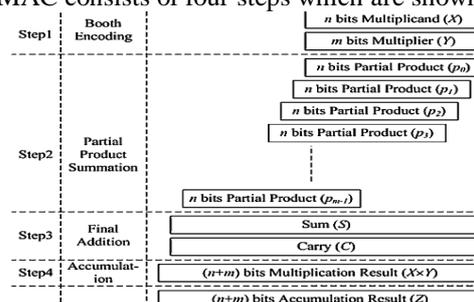


Figure.1 Hardware Architecture of Conventional MAC

The mathematical behavior of Multiplier Accumulate can be expressed as:

$$P = X \times Y + Z = \sum_{i=0}^{N/2-1} d_i 2^i Y + \sum_{j=0}^{2N-1} z_j 2^j. \tag{1}$$

Where $d_i = -2x_{2i+1} + x_{2i} + x_{2i-1}$

If the above equation is carried in Base 4 type redundant sign digit form in order to apply radix-2 Booth's encoding, it would be [4]. Each of the two terms on the right hand side of equation (1) can be calculated independently, and the final result is produced by adding the two results. The MAC architecture by equation (1) is called Standard MAC design [2].

1.1 MERGED/FUSED MAC:

The approach to speed up the operation implements both the multiplication and accumulation operation within one or same functional block by merging accumulator with the multiplication circuit. If an operation to multiply two N-bit numbers and accumulate into a 2N-bit number is considered, the critical path is decided by the 2N-bit accumulation operation. The overall performance of MAC is improved by eliminating the accumulator itself by merging it with the partial reduction tree used for Multiplication [6].

The Hardware architecture of the MAC as shown in Figure.3 executes the multiplication operation by input Multiplier (X) and Multiplicand (Y). This is added to previous result (Z) as accumulation step. The operation of MAC is performed in 3 steps instead of four steps by fusing multiplication and addition.

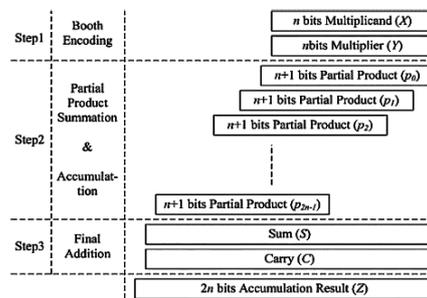


Figure.3 Hardware Architecture of Merged MAC

2. APPROACH TO DESIGN HIGH SPEED PARALLEL MAC

2.1. A FAST MAC DESIGN USING MODIFIED BOOTH ENCODING:

One of the types of MAC for general purpose signal processing was proposed by Elguibaly [8], 2000. It is an architecture where accumulation has been combined with the Carry Save Adder (CSA) Tree that compresses the partial products and provides fast possible implementation. In the architecture proposed by [8], the critical path delay is removed by eliminating the separate adder block and decreasing the number of input bit in the final adder. In order to avoid irregular wiring, CSA preferred to connect the neighbor interconnect structure. In conventional MBA, sign extension is used at all partial product rows. So each row becomes N+2, which is completely waste and it increases hardware as well. So in this implementation, it uniformly extend each summand by 1 bit only even through it can add three 2's complement number at each step.

In Fig.4, dotted line will interpret the addition of 1 in order to perform 2's complement operation. p0 - p16 be the result of overall MAC output. z0 - z15 be the accumulated input values to be added in the result of multiplication. Diamond icon shows the implementation of 2 or 4-bit fast CLA. Circular icon translates the Carry Save Adder used to add the partial product terms. MAC performs two pipeline stages i.e. firstly to generate, add the partial product and the part of z. Second stage performs the final addition and add last part of z using N+2 bits. After implementation it was conclude that pipelined MAC structure based on MBA was three times faster than the standard pipelined technique for construction of MAC.

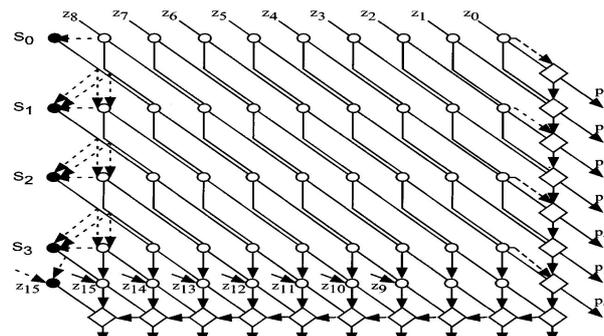


Figure 4 Dependence Graph of MBA [8]

This has better performance because of the reduced critical path delay as compared to conventional MAC as it divides the final adder into two stages i.e. one to add the least significant bits (N-1) bit words and one to add most significant (N+2) bit word which automatically improves the performance as well as delay at final addition step.

2.2 FAST MAC DESIGN USING 5:2 COMPRESSORS:

Effective increase in speed of MAC design was proposed by *Kwon, Nowka* [9] 2000. Compressor scheme was modified by fast 5:2 technique instead of 4:2 or 3:2 compressors formed a merged multiplier with accumulator. This design consists of 16 x 16 bits Input followed by Radix-4 Booth Encoding Algorithm, addition by Carry Save design which caused elimination of one row of partial product matrix before actual reduction method which rejected one 3:2 counter delay .

Conventional 3:2 counters have $2\Delta_{XOR}$ delay and 4:2 counters can provide $3\Delta_{XOR}$ delay on their critical path as actually it has 5 inputs due to neighbor carry-in signal and 3 outputs due to extra carry-out to one greater significant cell connected. Here Modified 5:2 compressor was implemented with only 5 inputs and 2 yield outputs which allow effective increase in speed and approximately one Δ_{XOR} less delay compared to other compressors.

Figure.5 depicts the basic block diagram of MAC which consists of two 16 -bits Inputs followed by Radix-4 Booth Recording Techniques, then partial product reduction is performed by 10:2 reduction bit slice that formed by two 5:2 compressors and then one 4:2 counters. This will result in the improvement of speed as delay is $7\Delta_{XOR}$ and $8\Delta_{XOR}$ delay in conventional compressors formed by 3:2 and 4:2 counters.

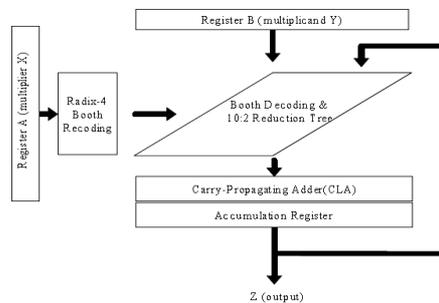


Figure.5 Block Diagram of MAC [9]

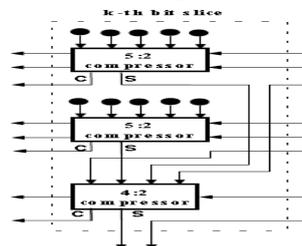


Figure.6 Bit Slice in 10:2 Reductions [9]

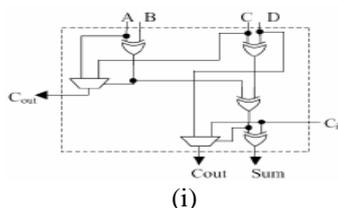
This provides the overall 14% increase in speed on MAC design due to use of fast 5:2 compressors. This compressor is widely utilized in applications regarding multiplications and can be effectively used in 10:2, 15:2, 20:2 compressors depending on the size of input bits.

2.3 IMPROVEMENTS IN SPEED AS WELL AS AREA:

A high speed and area efficient MAC was proposed by *Abdelgawad, Bayoumi* [12] ,2007 where the critical path delays and hardware complexities of Conventional MAC Architecture was deduced at a design with low delay and hardware requirements. The technique based on binary reduction tree constructed with modified 4:2 compressors and reducing the overall area by full utilization of compressors instead of putting zero in free inputs of the partial product rows terms.

Figure.7 described the circuit of original and modified 4:2 compressors. The modified compressor used will give an additional carry out bit that is required to be fed to the next column of partial products which in turn will require an additional modified 4:2 compressor in the next column at the first stage which will also have an additional carry out. Here $Z_0 - Z_{15}$ are the inputs of accumulator which are added to partial product of multiplying inputs.

Stage 1 performs the compression of two blocks A and B in a single clock cycle so speed and throughput is therefore improved and rest of the inputs of accumulator are utilized with the compressor output in the Stage 2.



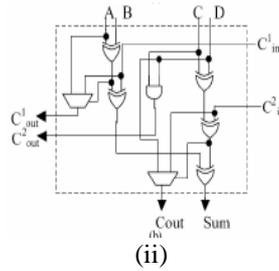


Figure.7 (i) Original Compressor (ii)Modified compressor [12]

The final decision to add Z7 bit of accumulator will decide the number of modified compressors to be used in the design. The accumulator stage has maximum critical delay path as it consists of large accumulated bits which cause large capacitive hardware and thus decreases the area efficiency. The final decision to add Z7 bit of accumulator will decide the number of modified compressors to be used in the design. The accumulator stage has maximum critical delay path as it consists of large accumulated bits which cause large capacitive hardware and thus decreases the area efficiency.

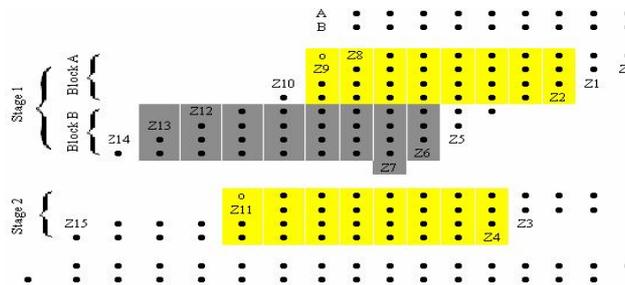


Figure 8 Data bits distribution of merged MAC [12]

Then a fast adder is designed at the last stage to add output from the compressor. Here Carry Look Ahead Adder [5] is used in the final adder stage which is theoretically one of the fastest adders and it has a hierarchical structure to make adders that have a multiple of 4 bits. This implementation designed for 8 bits, 16 bits, 32 bits which provide reduction in area by 6.25%, 3.2% and 2.5% and improvement in speed by 14%, 16% and 19% respectively.

2.4 PIPELINED MAC UNIT

Further implementation of pipelined MAC was proposed by *Shanthala, Cyril Prasanna, Kulkarni* [13], 2009. For designing the pipelined MAC, various architectures of multipliers (Booth algorithm and Array Multiplier) and one bit full adders are considered. The Static and Dynamic one bit Full Adder was implemented as the basic block for multiplier addition. They were compared with Pseudo Domino and N-P domino logic results in fastest and smallest adder. The MAC is designed by both Array and Booth encoding techniques (Figure. 9) where reduction is done by using Carry Save Adder.

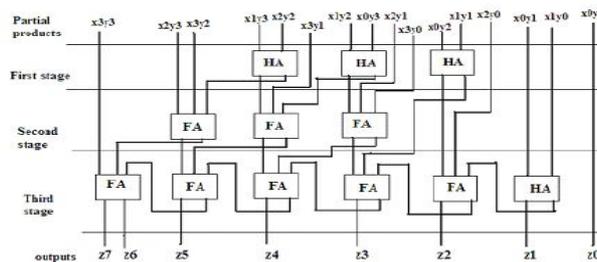


Figure 9 Array Multiplier [13]

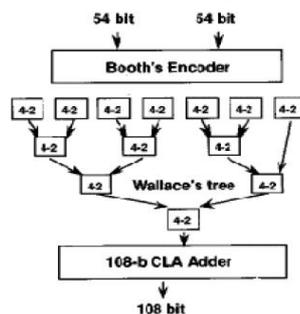


Figure 10 Inoue Multiplier [13]

The input was fetched from the memory location and fed to multiplier to perform multiplication and adder accumulated the result. The entire process was achieved in a single clock cycle. The performance was yield in terms of area and power; this design provided only 50.26 mW of dissipation and 3*1.05 mm² in area. The response time of design is just 6 clock cycles [13].

2.5 EFFECTIVE DESIGN OF FAST PARALLEL MAC USING RADIX-2 MBA

This advanced design of MAC was proposed by *Young Ho-Seo, Wook Kim*[16],2010 to provide a high speed architecture of MAC with accumulation fused with the CSA tree which performed the partial product reduction. Due to large path of multiplication, Radix-4 was less preferred over the radix-2 and for high speed multiplication, Booth recording technique was preferred as it compressed the partial products matrix to half and therefore less input left for summation [1].

Wallace Tree obtained by Carry Save Adder was utilized such that they add the encoding terms as parallel as possible and its operation time is proportional to O (log₂N), where N is the number of Inputs. The delay intensive path for MAC is accumulation stage. If it gets eliminated then critical path due to accumulator definitely will be reduced and delay left only will be due to the final adder at the last stage. So this delay due to final adder can be reduced by decreasing the input bits for addition. This method was proposed in [10] to improve the performance of final adder. In order to reduce the bits for final adder, the partial products rows are compressed into sum and carry outputs using CSA [6].

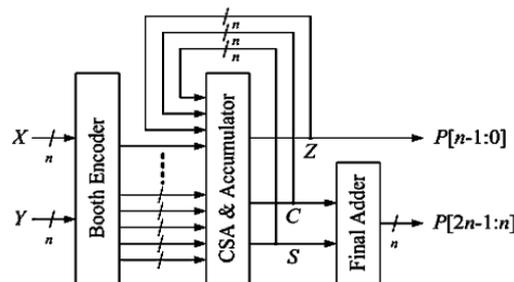


Figure.11 Block Diagram of MAC [16]

The number of bits for final adder was further foreshortened by advance calculating the lower bits within the range so that the overall performance should not get affected. Output rate can be further increased by inputting back the higher bits in the form of sum and carry into the CSA architecture such that final adder delay is not considered every time during the calculation of MAC output. Because of the feedback of sum and carry, number of inputs to CSA is increased. Mathematically, it can be easily understood as supposed X and Y be the two input to the Booth encoder and Z be the output accumulated value.

Multiplication can be performed such as:

$$X \times Y = d_0 2^0 Y + d_1 2^1 Y + d_2 2^2 Y + \dots + d_{N/2-1} 2^{N-2} Y. \tag{2}$$

This product is divided into three product terms such as sum, carry and pre-added lower bits of sum and carry.

$$X \times Y = d_0 2^0 Y + \sum_{i=1}^{N/2-2} d_i 2^{2i} Y + d_{N/2-1} 2^{N-2} Y. \tag{3}$$

Therefore the above equation is separated as first partial product, middle terms and upper bits partial products.

The accumulated output Z is also separated into lower and upper bits such that upper bits value are feedback as sum and carry and lower bits value are feedback as addition value of sum and carry.

$$Z = \sum_{i=0}^{N-1} z_i 2^i + \sum_{i=N}^{2N-1} z_i 2^i \tag{4}$$

The final output of MAC can be arranged such as:

$$P = \left(d_0 2^0 Y + \sum_{i=0}^{N-1} z_i 2^i \right) + \left(\sum_{i=1}^{N/2-1} d_i 2^{2i} Y + \sum_{i=0}^{N-2} c_i 2^i 2^N \right) + \left(d_{N/2-1} 2^{N-2} Y + \sum_{i=0}^{N-2} s_i 2^i 2^N \right). \tag{5}$$

The first right bracket term (in equation 5) is the procedure to add the first partial product with the added value of sum and carry. The second bracket term is the procedure which compiles the middle partial product terms with the sum of CSA that are feedback and the third bracket is the procedure that adds the lower partial product terms with the carry of the CSA. That's why CSA used in this method was called hybrid CSA as it provided advance output which improved the speed and also the execution steps. The hybrid CSA proposed was implemented as shown in Figure. 12 for 8 x8 bit MAC.

Such as S_i for sign extension, N_i is to convert 1's complement into 2's complement result, Z[i] be the addition of lower bits of sum and carry bits, Z'[i] be the previous result value of Z[i]. Due to MBA, partial product rows become four for

8-bit multiplier and grey row of partial product is the previous accumulated values. Rectangular block interpret the 2-bit CLA[5] block to add the lower bits of sum and carry as CLA is one of the fastest adder. The whole process of MAC by this technique was compared with *Elguibaly* [8] and difference of 1 clock cycle was noticed. Pipeline delay was also calculated to be half as compared to *Elguibaly* [8].

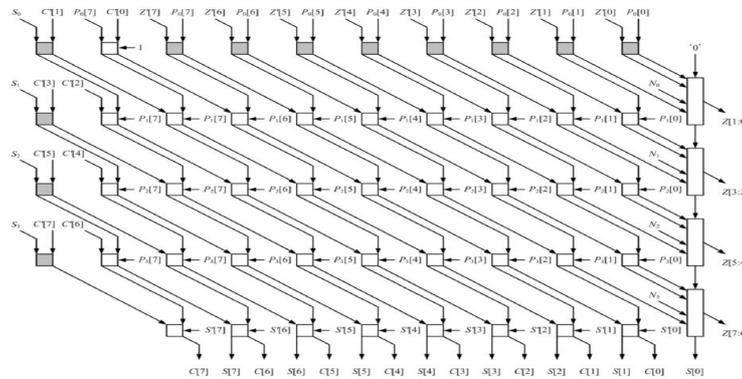


Fig.12 Architecture of proposed CSA Tree[16]

This is one of the best methods to perform parallel MAC. This design can be further improved using Higher Radix Multiplier and Fast Adder at the final addition stage.

3. PROPOSED WORK

By examining the various architectures, we have observed that delay can be improved by using higher radix MBA which reduces number of partial product rows that eventually reduces number of multiplication thereby improving speed. Thus we propose a new high speed and area efficient MAC architectures which will be an improvement over the existing Architecture [16] by replacing Radix-2 with Radix-4 and Radix-8 Modified Booth Encoder in the Multiplication Stage. The output of multiplication and accumulation stages will be combined using hybrid reduction through CSA, CLA and HA which enhancing speed and efficiency. In addition, the Final Stage of MAC will include a New Carry Select Adder (CSLA) [17] with Binary to Excess Convertor.

The new adder will be an improvement over the existing CSLA [18] by replacing Ripple Carry Adder in internal blocks by CLA.

The design will be implemented using VHDL language and simulated using Xilinx ISE10.1 Simulator. We expect the proposed MAC will be useful in high performance signal processing system.

4. CONCLUSION & FUTURE SCOPE

Considering all the design of MAC above, higher radix MBA and partial reduction technique by hybrid CSA tree can give good result in terms of speed as well as area. Its area and speed efficiency is the main advantage in VLSI design world. This work can be further extended for power and area.

REFERENCES:

- [1] A. R. Omondi, "Computer Arithmetic Systems". Englewood Cliffs, NJ: Prentice-Hall, 1994.
- [2] O. L. MacSorley, "High speed arithmetic in binary computers," *Proc.IRE*, vol. 49, pp. 67–91, Jan. 1961.
- [3] C. S. Wallace, "A suggestion for a fast multiplier," *IEEE Trans. Electron Computer.*, vol. EC-13, no. 1, pp. 14–17, Feb. 1964.
- [4] A. D. Booth, "A signed binary multiplication technique," *Quart. J. Math*, vol. IV, pp. 236–240, 1952.
- [5] Raahemifar, K. and Ahmadi, M., "Fast carry look ahead adder", *IEEE Canadian Conference on Electrical and Computer Engineering*, 1999.
- [6] Taewhan Kim and Junhyung Um "A Practical Approach to the Synthesis of Arithmetic Circuits Using Carry-Save-Adders" *IEEE transactions on computer-aided design of integrated circuits and systems*, vol. 19, no. 5, May 2000.
- [7] A. Weinberger; J.L. Smith, "A Logic for High-Speed Addition", *National Bureau of Standards, Circulation 591*, pp. 3-12, 1958.
- [8] F. Elguibaly, "A fast parallel multiplier-accumulator using the modified Booth algorithm," *IEEE Trans. Circuits Syst.*, vol. 27, no. 9, pp.902–908, Sep. 2000.
- [9] Kwon, Swartzlander. Nowka, "A 16 bit x 16 bit MAC design using 5:2 compressors "IEEE 2000.
- [10] A. Fayed and M. Bayoumi, "A merged multiplier-accumulator for high speed signal processing applications," *Proc. ICASSP*, vol. 3, pp.3212–3215, 2002.

- [11] Hwang-Cherng Chow and I-Chyn Wey, "A 3.3V 1GHz high speed pipelined Booth multiplier," Proc. of IEEE ISCAS, vol. 1, pp. 457-460, May 2002.
- [12] A. Abdelgawad, Magdy Bayoumi "High Speed and Area-Efficient Multiply Accumulate (MAC) Unit for Digital Signal Processing Applications" IEEE 2007.
- [13] Shanthala S, Cyril Prasanna Raj, Dr.S.Y.Kulkarni, "Design and VLSI Implementation of Pipelined Multiply Accumulate Unit", IEEE computer society, 2009
- [14] P. Zicari, S. Perri, P. Corsonello, and G. Cocorullo, "An optimized adder accumulator for high speed MACs," Proc. ASICON 2005, vol.2, pp. 757-760, 2005.
- [15] K. Babulu et al, G.Parasuram "FPGA Realization of Radix-4 Booth Multiplication Algorithm for High Speed Arithmetic Logics", (IJCSIT) Vol.2(5), 2011, 2102-2107.
- [16] Young-Ho Seo, Dong-Wook Kim, "A New VLSI Architecture of Parallel Multiplier-Accumulator Based on Radix-2 Modified Booth Algorithm" IEEE transactions on very large scale integration (VLSI) systems, vol. 18, no. 2, February 2010.
- [17] Samiappa, Sakthikumar, S. Salivahanan, V. S. Kanchana Bhaskaran, V. Kavinilavu, B. Brindha and C. Vinoth, "A Very Fast and Low Power Carry Select Adder Circuit", 978-1-4244-8679-3/11.2011 IEEE.
- [18] P. Devi and A. Girdher, "Improved Carry Select Adder with Reduced Area and Low Power Consumption", International Journal of Computer Applications (0975- 8887) vol. 3 - No.4, June, 2010.

AUTHORS:

Deepika Setia received the B.Tech degree in Electronics and Communication in 2010. She is pursuing M.Tech (Microelectronics) from UIET, Panjab University, Chandigarh. Her area of interest includes Image Processing and VLSI.

Mrs. Charu Madhu is M.E (Electronics and Communication) from Bant college of Engg. And Technology, Gurdaspur (PTU). Her Area of interest includes VLSI, Nanoscale Devices and Optoelectronics. Currently she is working as Assistant Professor (ECE). She has 3 publications in International journals /Conferences proceedings.