International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-2 Issue-2, May 2013

# Efficient Optimization Of FPGA On-Chip Memory For Image Processing Algorithm

**S.M.Mehzabeen, I.Manju**

***Abstract-This paper is concerned with efficient optimization and low power implementation of FPGA on-chip memories in image processing algorithms. In recent years on chip memories are expected to increase continuously which depends upon the application for future generation portable devices and high performance processors. Memory plays a major role in image processing applications more than 90% of the consumed power in the system is by the memory part. This paper provides a novel approach by making SPSRAM to function like a DPSRAM. It supports most of the access schemes for Image processing algorithms and also when the readout changes the memories need not to be redesigned. It achieves high throughput, less hardware requirement and high bandwidth utilization. The full bandwidth utilization has been achieved by splitting the on-chip memory into four sub banks. The Optimization of power can be done by making any two banks active at a time.It is well suited for various image coding algorithms when compared to the typical SPSRAM and TDP SRAM.It finds applications in most of the parallel processing fields.GENERAL TERMS Design, measurement, performance, theory.***

***KEYWORDS- Bandwidth utilization,Field programmable gate array, Power optimization, SRAM, access schemes in image processing, TDP SRAM.***

## I. INTRODUCTION

Reconfigurable computing finds an application in various research fields. The main idea behind the reconfigurable systems is to increase system hardware performance .
Based on hardware design reconfigurable systems can be designed either in hardware mode or software mode and even though models comprising of both hardware and software. The hardware mainly classified into ASIC's or FPGA's. But they provide different values to the designers, they must be evaluated carefully one over another. ASIC a hardwired technology all the operations should be performed at board level. As the name implies Application specific it must be fast and efficient for the specific hardwired applications. Designing with ASIC's leads to the difficulty in replacing or remodification of the chips as it is concerned with the board level.
A reconfigurable system made up of microprocessors seems to have a set of instructions .These instructions should be fetched by the processors and decode the instructions into a series of actions and the processor has to execute. The systems flexibility has been increased but the performance level decreases.

**Revised Manuscript Received on 30 May 2013.**
* Correspondence Author
**Mrs. Mehzabeen S.M**, PG Student/Applied Electronics, Velammal Engineering College, Chennai, India.
**I. Manju**,Assistant Professor[Sr.Gr], Electronics and Communication Department, Velammal Engineering College, Chennai, India.



Reconfigurable computing has to effectively fill the gap between the hardware and software issues. Thus FPGA fills the gap between the hardware and software issues in reconfigurable systems.[1]. FPGA's comprising of configurable logic blocks along with programmable interconnect points to form the necessary circuits. An FPGA with reconfigurable computing finds enormous applications in various fields. In FPGA's the memories are implemented inside the logic blocks has no external access. When compared to the external access throughput and speed has been increased.
The typically selected memories are SRAM's because of the minimum energy consumption .SRAM's are operated in sub threshold regions with ultra low power voltages. The use of on-chip memories nowadays are expected to increase continuously based on the future generation high performance and portable devices. There are various on-chip memories are available depending upon the applications .The selection of memories mainly based on storage capacity and hardware capability Throughput is the most critical requirement in high performance operating systems. The memory types based on its volatility. Volatile memories hold the data until power is applied to the memory device. It is not suitable for applications that data must be retained when the memory is switched off. FPGA on-chip memory are a volatile memory. But non volatile memories retain its data when power is switched off. The disadvantage of using non volatile memory leads to complex writing and erasing procedures. Flash memories are used by the embedded systems for nonvolatile storage. Most of the embedded system applications need both the volatile and nonvolatile memories. The simplest type of memory to be used in the FPGA based embedded system is on-chip memory.The memory is implemented in the FPGA itself no external connections are necessary on the circuit board. Some modifications of on-chip memory can be accessed in dual port mode. The best application of on-chip memory can be used as a cache.
SRAM based FPGA's become the workhorse of numerous reconfigurable applications.[3].These devices contain an array of lookup tables for combinational functions , flip flops for the sequential finite state machine, memory blocks for data storage, DSP block for compact signal processing , clock management blocks for configurable system clocks and interconnection nets to link. The goal is to use behind the temporal and spatial partitioning of reconfigurable logic resources having maximum advantage of parallelism resource usage and flexibility. To achieve longer lifetime and high reliability battery operated devices the power consumed by the SRAM must be reduced [2].There may be some existing low power embedded processors.


Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Cache memories in embedded processors contribute to the major fraction. [7] In integrated circuits power specification plays a major role[4] Detection of open defects in static access memory SRAM cells including those caused data retention faults (DRF's) are known to be time consuming and difficult.Researchers are going on to detect the data retention faults occurring in the SRAM's. [5] As technology migrates to smaller geometry leakage contribution to total power consumption increases faster than dynamic power, indicating that leakage will be a major contributor to overall power consumption [6].

The previous survey clearly depicts the importance of low power processors for real time applications. It limits the performance of the system by frequent usage .Xilinx, and Altera nowadays use the strategy of high bandwidth on-chip memories operating at low frequencies [1]. There are different types of memories available for image processing algorithms. SPSRAM (less area) can be used to access one address at a time. It can either perform any one operation, whether read or write. DPSRAM (high speed) can be used to access two ports at the same time. DPSRAM can perform simultaneous read and write. To gain the advantage of both the SRAM's Pseudo dual ports is introduced which can read and write the data in the same clock using rising and falling edges of the clock

Based on the Table 1 gives the comparison of memories.Considering SPSRAM not suitable for image processing algorithms as it has less bandwidth, though power and hardware complexity is less. In DPSRAM bandwidth is high which is an essential for image processing algorithms. The hardware complexity and power will be more. Comparing the advantages and disadvantages the proposed system SPSRAM will act like a DPSRAM with moderate hardware complexity.

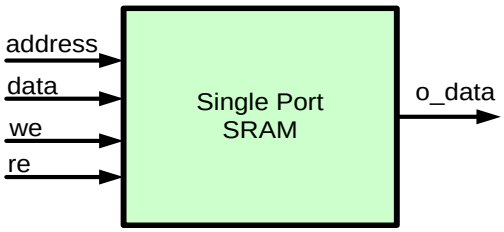


**Fig.1 Single port SRAM**

Image processing support various access schemes or readouts to retrieve the data from memory. The proposed system supports most of the scan orders used in the image processing algorithms. Major image scans orders include in the image processing are Row, Row prime, Column, Column prime, Morton and Peano Hilbert. The Peano Hilbert is a difficult access scheme to be used to retrieve the data in the blocks. The different complexities and challenges while generating the address for Peano Hilbert is extinct known by the paper [8]. The key features it can generate every Hilbert address in one clock cycle by using a simple incremental counter and multiplexers which require less hardware storage. The proposed systems generate the addresses without the requirement of any external hardware. This leads to optimization in power.

The image processing algorithms face the problem of redesigning when the access policies change. For particular applications in image processing there must be a need of using more than one access scheme. During that time when access policies change the memories need to be redesigned. The proposed system with the help of dual port SRAM can support the access change without redesigning and also without compromising the power. As according to the Table 1 the proposed system performs Dual read or write operations in a single clock cycle. It outperforms the typical SPSRAM and TDP SRAM's.

The organization of the paper is presented at Section 1deals with the introduction and the advantages of the proposed systems. Section 2 describes the memory architecture of the proposed sub bank systems. Section 3 describes the Proposed memory system and different access methods for image processing algorithms. Section 4 results and discussions of the proposed paper. Section 5 deals with a conclusion and future.

## II. ARCHITECTURE OF MEMORY

The proposed memory architecture has compared with the TDP SRAM in many aspects. It supports the most commonly used access schemes in image processing algorithms. In the proposed system the use of Dual port has the advantage of full bandwidth utilization which is a must for image processing algorithms. The throughput has been greatly increased with an effective reduction in area and power when compared to TDP SRAM.

**Table. 1 Comparison of Memories**

| Memory schemes/ Parametes | Bandwidth | Hardware complexity | Power |
|---|---|---|---|
| SP SRAM | Less | Less | Less |
| DP SRAM | High | More | More |
| TDP SRAM | High | Moderate | Moderate |
| Proposed sub-bank DP SRAM | High | Moderate | Less |

### *A. Typical Memory Models*

As shown in the Fig. 1, the SPSRAM has a single address line and data line. It performs two operations whether Read or Write .Depending on the selection of re and we the data has been written to the particular address line or the data is read from the particular address in desired pattern as per the mode selection signals.Through the single address and data line, there may be a provision getting only one output at a time. Comparing to the Table 1. It has less bandwidth,less hardware complexity ans less power.But as in the case of Image processing algorithms high bandwidth requirement is a major constraint.For the particular reason We are analysing the next type of RAM DPSRAM.As per the comparison table and as per our application for DPSRAM the band width is high and hardware complexity and power is more when compared to SPSRAM.In order to minimize the hardware complexity and power of the proposed system the on chip memory has to be subdivided into four banks .accessing any two banks at a time leads to power optimization. As shown in the Fig 2 the DPSRAM has two address and data lines and two output lines. At a time we can perform dual read or write operations .The read and write operation has to be following the required pattern as per the mode selection signals.



Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

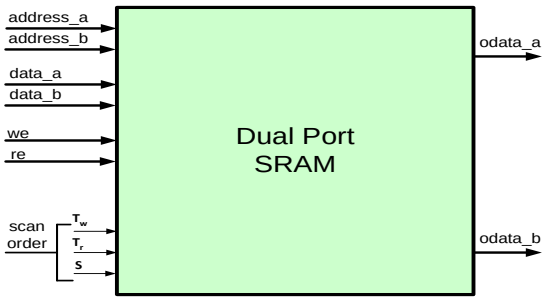


**Fig.2 Dual port SRAM**

This will provide the complete bandwidth utilization which is an essential criteria for image processing application. In the proposed system, the SPSRAM will act like a DPSRAM. The on-chip memory is divided into four sub banks. In order to achieve power optimization two banks enabled at one time keeping the other two in off modes.

### *B. Proposed Sub Bank Memory Model*

As shown in the Fig.3 The on-chip main memory is sub divided into four banks. (Bank 0,1,2,3). The sub banks will be indicated by BXX. The size of the main memory is indicated by MXN. The total memory size is divided by 4.

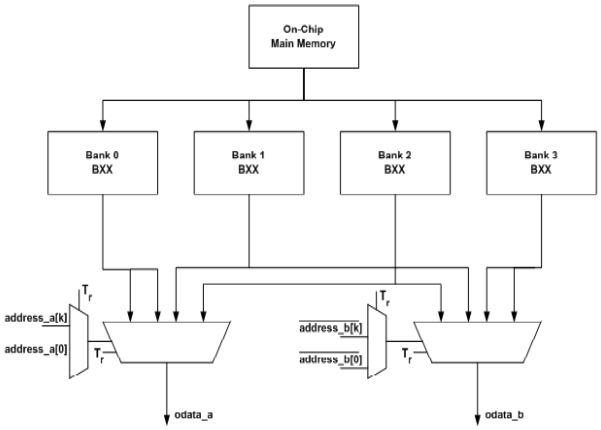


**Fig.3 Sub bank memory model**

Each sub bank will be about the size MXN/4.Each memory can hold MXN/4 pixels .A four input multiplexer is used to output any two banks at a time.The address generator generates the addresses where to perform READ and Write operations to and from the memory. The clock generator provides the necessary synchronization operations. The proposed sub bank memory produces a dual-port output. The address generation for particular
modes can be selected based on the Tr,Tw and S signals.

## III. PROPOSED MEMORY SYSTEMS FOR IMAGE PROCESSING

The Proposed memory systems for Image processing algorithms is as shown in the Fig.4.

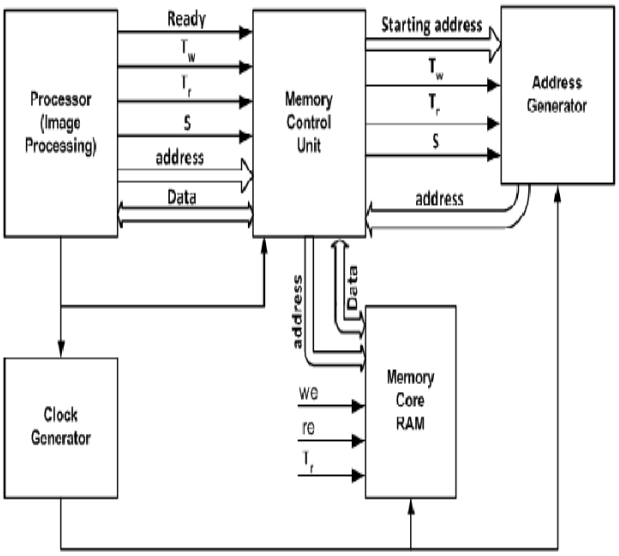


**Fig.4 Proposed system for Image processing**

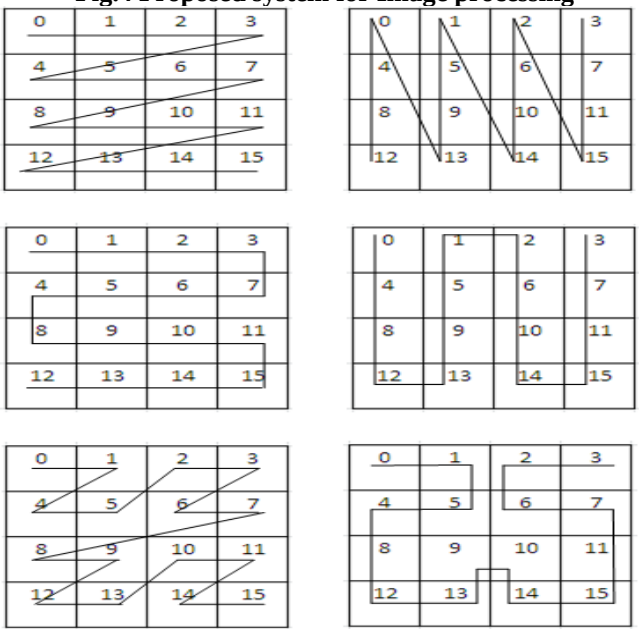

**Fig.5 Different access schemes**

It comprising of various sub blocks such as Processor for image processing depending upon the application, address generator, clock generator, memory control unit and memory core. Efficient utilization of memory is the major constraint while implementing in FPGA. The proposed memory system effectively utilizes the memory by splitting the n-chip memory into sub banks. Full bandwidth utilization is achieved which is an essential for image processing algorithms. The sequence of operations carried out by the proposed memory systems is as follows. The processor for image processing algorithm issues the READY signal, starting address and required scan pattern through which the read and write operations are performed. Based on the mode selection signals (Tw,Tr and S) the required modes has been selected to the subsystems. On receiving the READY signal the memory control unit produces the starting address to the address generator from which the address generator has to generate the address for the selected pattern.


Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

The MCU issues the address and data to the sub bank RAM. In the sub bank RAM depending upon the address and data it performs the READ and WRITE operations. The clock generator issues the clock signals to all the sub blocks in order to achieve proper synchronization.

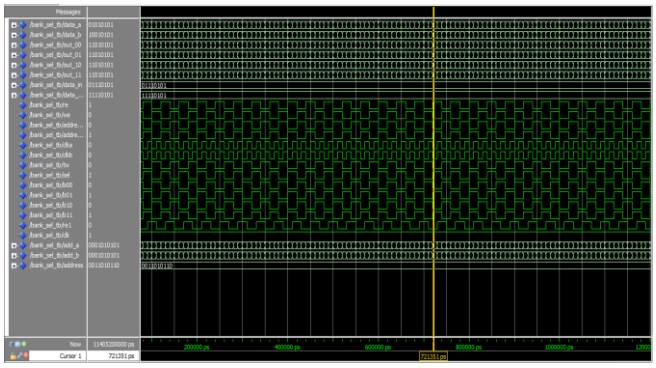

**Fig.6 Various control signal generation for memory banks**

### *A. Address Generator*

The address generator generates the particular address from which the data has to be READ/WRITE in the memory. The storing and retrieval of images in image processing is strongly patterned. Therefore the address generator controlled by MCU issued to arrange memory allocation in any one of the following ways A, The incoming data are written to consecutive locations and the consumed data to be read should follow the required pattern. B, The incoming data should written in a predetermined form and reading can proceed in sequential locations. Each memory location is denoted with four bits(0000). The bit in the zeroth and second denotes the memory bank selected.The bits in the first and third denotes the address .Starting from the particular address and pattern selected from the mode selection signals the memories has been readouts.The sub - bank will be chosen based on the bank selection factors for row and column index. In order to select a 16 number of pixels we need four address bits. The address 0000 will select bank B00 with starting address 00. Similarly depending upon the address the banks are selected.The starting address for the address generator is given from the MCU .Then depending on the starting address the control in the address generator starts generating the address based on the scan order selected. The scan order has been selected depending on the control signals Tw, Tr and S. The available scan orders Are Peano- Hilbert, Morton, Column prime, column, Row prime and Row. The address generator starting addresses from particular location based on the Tr,Tw and S with dual addresses at the output is shown in Fig.7

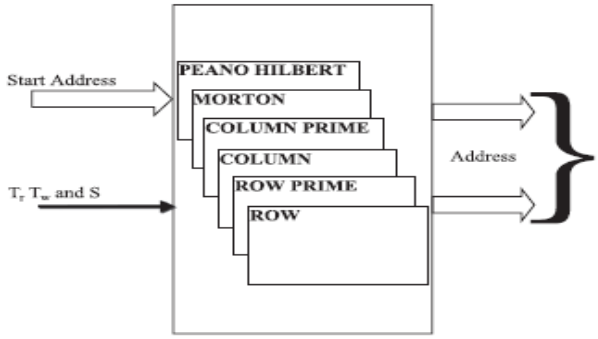


**Fig.7 Address generator**

### *B.Memory Control Unit*

The MCU comprising of Request interface unit which requests the data, address and scan access signals necessary to generate the addresses from which the address should be generated by the address generator. In between the MCU and the processors there are buffers used for proper synchronization and it act as a handshaking signal. The MCU issues the address and data to the memory sub bank module perform READ/WRITE.The output of the memory control unit is the control signals to Address generator and the memory core.The first block is the Request interface block comprising of control signal interface,address interface and data interface.

### *C.Finite State Machine*

The FSM in the memory control unit selects the different mode of operations .There are eight states in the FSM depending on the three mode select signals. The eight states are Idle, Write, Row, Column, Row Prime, Column Prime, Morton and Peano Hilbert access scheme selections. When the READY signal becomes 1 the processor issues the starting address of the MCU. Otherwise it will be in the idle state. The mode select signals are signals ( Tw, Tr and S).

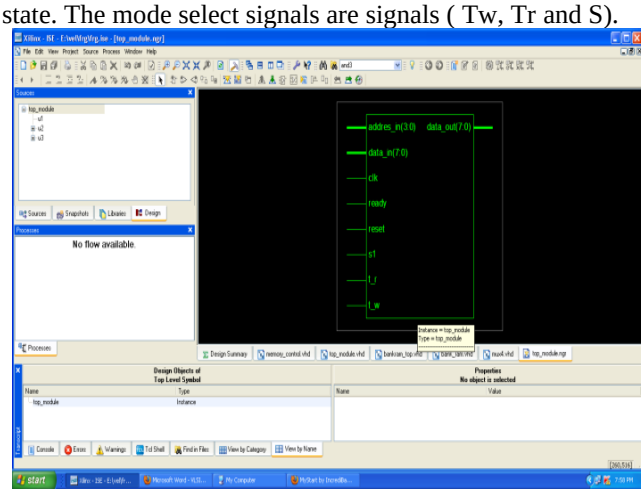

**Fig.8 Schematic of Top module**

An example Table.2 is shown below the clock selection for Row scan.

**Table.2 Generation of clock for Row scan.**

| Clock | Clock Generation | | Port A | | Port B | | Accessed Banks (On state) | Unaccessed Banks (Off state) |
|---|---|---|---|---|---|---|---|---|
| | Clk-a | clk-b | Decimal | Binary | Decimal | Binary | | |
| 1 | 1 | 0 | 0 | (0000) | 1 | (0001) | B00,B01 | B10,B11 |
| 1 | 1 | 0 | 2 | (0010) | 3 | (0011) | B00,B01 | B10,B11 |
| 1 | 0 | 1 | 4 | (0100) | 5 | (0101) | B10,B11 | B00,B01 |
| 1 | 0 | 1 | 6 | (0110) | 7 | (0111) | B10,B11 | B00,B01 |
| 1 | 1 | 0 | 8 | (1000) | 9 | (1001) | B00,B01 | B10,B11 |

## IV. SIMULATION RESULTS AND DISCUSSIONS

The Various sub blocks for the proposed memory system for image processing applications has been simulated using ModelSim individually and the Top RAM Memory module has been synthesized using Xilinx ISE. The necessary modules for the image processing algorithms has been designed and simulated. Fig.7 shows the Simulation result of various control signal generation for the different banks.



Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

## V. CONCLUSION AND FUTURE WORK

The sub-bank Dual port memory architecture uses four single port SRAMs which are proposed, supports several data access policies. This architecture throughput can be comparable with that of TDP SRAMs.

Using this proposed architecture the power consumption and speed can be improved. This DP SRAMs can support most of the scan orders and is most suitable for image processing applications. The future work, this memory system architecture can be extended to support the other area of applications for parallel data accesses.

## REFERENCES

1. K.Compton, S.Hauck, Reconfigurable computing: a survey of systems and software, ACM Comput. Surv. 34(2002)171–210.
2. Ranjith kumar volkankursun, Temperature adaptive voltage scaling for enhanced energy efficiency in subthreshold memory arrays ,Microelectron J.40(2009) 1013-1025.
3. Altera corporation,Embedded Design Hand book,Chapter 7,Memory System Design (2010).
4. F. Francisco, F. Mariano, C. Enrique, Run-time self-reconfigurable 2D convolver for adaptive image processing, Microelectron. J. 42 (2011) 204–217.
5. ARM Architecture Reference Model, ARM DDI 0100l (2005).
6. P.Deepa,C.vasanthanayaki,FPGA based efficient on-chip memory or image processing algorithm.Microelectron.J(2012).
7. Yan Wang, Shoushun Chen and Amine Berma k smart,Novel VLSI implementation of Peano-Hilbert curve Address Generator,Sensory Integrated Systems Lab Electronic and Computer Engineering Department,Hong kong University of science and technology.
8. G.S. Sohi, M. Franklin, High-bandwidth data memory systems for superscalar processors, SIGOPS Oper. Syst. Rev. 25 (Special Issue) (1991) 53–62.
9. S. Heithecker, A. do Carmo Lucas, R. Ernst, A mixed QoS SDRAM controller for FPGA-based high-end image processing, in: SIPS, IEEE Workshop Signal Process. Syst. 2003 (2003) 322–327.
10. Muhammad M.Khellah,Member,IEEE and Mohamed I.Elmasry,Fellow,IEEE,A Low power High performance current mode multiport SRAM,IEEE Transactions on VLSI,Vol.9,No .5,October 2004.
11. Houman Homayoun, member, IEEE, Avesta Sasan, Member, IEEE, Alexander V.Veidenbaum, member, IEEE,Hsin Cheng Yao, Shahin Golsan,and Payam Heydari, senior Member, IEEE MZZ-HVS Multiple sleep modes Zig-zag horizontal and vertical sleep transistor sharing to reduce leakage power in on-chip SRAM peripheral circuits,IEEE transactions on VLSI systems,Vol.19,No.12,December 2011.
12. Srinivas R.Sridhara,Member, IEEE,Michael Direnzo, Member, IEEE, Srinivas Lingam, Member, IEEE,Seok-Jun Lee, Member, IEEE, Raul Blazquez, Member, IEEE,Jay Maxey, Member, IEEE, Samer Ghanem,Yu-Hung Lee,Rami Abdallah, PrashantSingh, Member,IEEE, and Manish Goel, Member, IEEE, Microwatt Embedded Processor Platform for Medical System-on-Chip Applications,IEEE Journal of Solid-state circuits,Vol.46,No.4,April,2011.
13. P. Ranganathan, S. Adve, N.P. Jouppi, Reconfigurable caches and their application to media processing, Proceedings of the 27th International Symposium on Computer Architecture, USA. (2000) pp. 214–224.
14. Q. Liu, G. Constantinides, K. Masselos, P. Cheung, Automatic on-chip memory minimization for data reuse, in: 15th Annual IEEE Symposium on Field- Programmable Custom Computing Machines. (2007) pp. 251–260.
15. S.S. Ang, G.A. Constantinides, W. Luk, P.Y.K. Cheung, Custom parallel caching schemes for hardware accelerated image compression, J. Real-Time Image Proc. 3 (2008) 289–302.
16. ARM Architecture Reference Model, ARM DDI 0100l (2005).



Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)