Energy and Area Efficient Approximate Image Compression Technique using Discrete Wavelet Transform 2D DWT and VLSI

Dr. Sandip R. Udawant[1], Dr. Satyawati S. Magar[2]

[12]*Associate Professor, Dept of E & TC Dr.Vithalrao Vikhe Patil College of Engineering, Ahmednagar*

[2]*Associate Professor, Dept of E & TC Dr.Vithalrao Vikhe Patil College of Engineering, Ahmednagar*

[1]udawant_etc@enggnagar.com , [2]magar_etc@enggnagar.com

Abstract

In this study, the wavelet-based VLSI architecture for image compression is studied. The most popular method of picture compression is discrete wavelet transform, which is also the most effective algorithm for JPEG image compression. A discrete wavelet architecture is suggested that can do both forward and inverse lifting. As an improved approach, lifting scheme is utilized instead of the conventional method, which needs more memory, space, and power. The DWT architecture, a potent image compression technique, is developed using a lifting-based strategy. Reduced memory referencing, low power consumption, low latency, and high throughput are all benefits of this architecture. In a manner similar to this, the Inverse Discrete Wavelet Transform (IDWT) is also obtained to obtain the visual matrix. Both lossy and lossless compression can be achieved using the concept. Linear algebra views of DWT and IDWT have been used to simplify the design. The hardware optimization, quick computation times, consistent data flow, and decreased complexity of the suggested architecture are its benefits. The proposed architecture can easily be scaled with the filter length and 2D DWT level due to its regular nature. The 2D DWT VLSI architecture is built with FPGA in VHDL. Applications that employ a lot of picture and video coding are becoming more and more popular, making this a busy and dynamic sector. A foundation for study and development is provided by Image and Video Compression for Multimedia Engineering.

Keywords:- *Discrete Wavelet Transform(DWT), Inverse Discrete Wavelet Transform(IDWT), Field programmable Gate Array(FPGA),Very Large Scale Integration(VLSI).*

1. INTRODUCTION

Many processors based on text data and picture data signal processing have been developed in the industry as a result of the quick growth of VLSI design technology. A key component of image or video compression standards like JPEG 2000 is the two-dimensional discrete wavelet transform (2D DWT). An alternate method for processing signals is the Wavelet Transform (WT), which is particularly effective for studying spectral and spatial localization. A useful method for signal decomposition into multiple resolution sub-bands is the discrete wavelet transform. One of the quickest wavelet transform computations is the DWT. It is simple to implement and uses fewer resources and time during computation. In the instance of DWT, digital filtering techniques are used to produce a time-scale representation of the digital signal. The signal to be analyzed is run through filters at various scales and with various cut off frequencies. Depending on the application, the scale factor and frequency parameter can be changed. Due to its good localization in both time and frequency, the discrete wavelet transform (DWT) is a very helpful tool in time-frequency analysis. It has shown great success in fields of study like picture compression. In order to reduce the amount of redundant data that must be stored and, ultimately, the cost of communication, data compression is a strategy. Reducing the amount of storage needed equates to improving the storage medium's capacity and, consequently, transmission bandwidth. Therefore, creating effective compression algorithms will remain a design issue for advanced multimedia applications and communication systems in the future. The two categories of data compression algorithms are lossless and lossy, respectively. Usually, written data or scientific data are the subjects of lossless data compression strategies. For picture coding, the discrete wavelet transform (DWT) is being employed

more and more. Because DWT offers more advanced capabilities like progressive picture transmission by quality or by resolution, this is the case. The lifting strategy, which is now frequently used for DWT, is faster and requires less processing than the traditional convolution-based approach. To simplify difficult procedures, Daubechies and Sweldens first develop the lifting-based discrete wavelet transform. To address the needs of real-time processing, numerous VLSI architectures for the 2-D DWT have now been proposed. There are certain problems with the practical application of DWT. First, compared to DCT, the complexity of the wavelet transform is much higher. Second, additional memory is required by DWT to store the computed results at intermediate stages. Additionally, DWT must analyze enormous amounts of data at rapid speeds in order to do real-time picture compression. Software-based DWT image compression offers flexibility for manipulation, however it might not adhere to specific temporal requirements in some applications. However, there are issues with DWT's hardware implementation as well. The first issue is the high cost of hardware multiplier implementation. To construct a delay element, 415 transistors are needed for an adder, and 6800 transistors are needed for a multiplier[4]. For DWT, a number of VLSI architectures have been suggested. The 2-D DWT is currently utilized in a wide range of image processing tasks, including texture discrimination, fractal analysis, and image and video compression. Numerous architecture solutions have been proposed in the past for the effective implementation of the 2-D DWT since it necessitates a significant amount of computing. Demands for visual data and higher-quality video frames have increased as a result of advanced technologies like 3D movies, games, and HDTV. Because of this, there is a gap between the capabilities of the still-limited technology and the enormous amount of visual data needed for multimedia applications. With clear, current video and picture coding knowledge, images and video compression for multimedia engineering fills the gap. The utilization of enormous volumes of visual data in applications is increasing, making picture and video coding a vibrant and ever-evolving field. The use of image and video compression in multimedia engineering creates a foundation for further research, development, and study. The frequency domain representation is the foundation for the high performance architecture for image compression proposed in this research. The proposed architecture has undergone testing for still images and was created using the verilog Hardware Descriptive Language.

2. DISCRETE WAVELET TRANSFORM

We discover that a fast computation of the discrete wavelet transform, which is based on sub-band coding, is possible. It requires less time and resources to compute and is simple to implement. The multi-resolution time-frequency plane of the discrete wavelet transform is built using filter banks. By breaking down the signal into approximation and detail information, the Discrete Wavelet Transform analyses the signal at various frequency ranges and resolutions. The signal is divided into several frequency bands using a series of high pass and low pass filters on the time domain signal. A pair of analyzing filters are made up of the high pass filter g[n] and the low pass filter h[n]. Each filter's output has the same number of samples as the input signal, but only half as much frequency content. Although the amount of data is doubled, the two outputs together have the same frequency content as the input signal. Therefore, the outputs of the filters in the analysis bank are down sampled by a factor of two, represented by ↓2, before being applied to the data. With the help of the synthesis filter bank, the original signal can be recreated. Signals are up-sampled (by a factor of 2) represented by ↑2 and routed via the filters g[n] and h[n] in the synthesis bank. The analysis bank's filters serve as the foundation for the filters in the synthesis bank. Perfect reconstruction is possible with the right mix of analyzing and synthesizing filters combined. The output, which is typically an estimate of the input and is exactly equal to the input applied, defines a perfect reconstruction. One signal can be divided into numerous lower resolution components by repeating the decomposition process with successive approximations being decomposed in turn. Decomposition can be carried out as needed. A multiple level decomposition method is the Two-Dimensional DWT (2D-DWT). It changes the frequency domain of images from the spatial domain. Four filtered and sub-sampled images are created by one level of wavelet decomposition, or sub bands. There are many benefits to applying the wavelet transform to decompose sub-band images. In general, it has a high compression rate and is profitable for non-stationary picture signal analysis. And to allow for progressive data transfer over a low transmission rate line, its transform field is portrayed with many resolutions similar to the human visual system. In order to do a multi-resolution analysis, DWT processes data on a changing time-frequency plane that gradually matches the lower frequency components to coarser time resolutions and the high frequency components to finer time resolutions. The Discrete Wavelet Transform has developed into a potent tool in a variety of applications, including the processing of images and videos, mathematical analysis, and telecommunication. Since DWT provides a multi-resolution analysis of a signal with localization in both the time and frequency domain, it has an advantage over existing transforms like the discrete Fourier transform (DFT) and DCT.

3. PROPOSED DESIGN:

Figure 1 displays the suggested design's block diagram. It is made up of two external dual-port memory and a DWT processor. The pixel values from a grayscale image serve as the initialization for the two memory. In the suggested design, the DWT processor receives input by importing an image from the workspace in Matlab. The memory controller, crossbars, and DWT filter are all components of the DWT processor. The crossbars are used to alternately transfer the output of the high pass and low pass filters to the two memory banks, interleaving the image pixels.

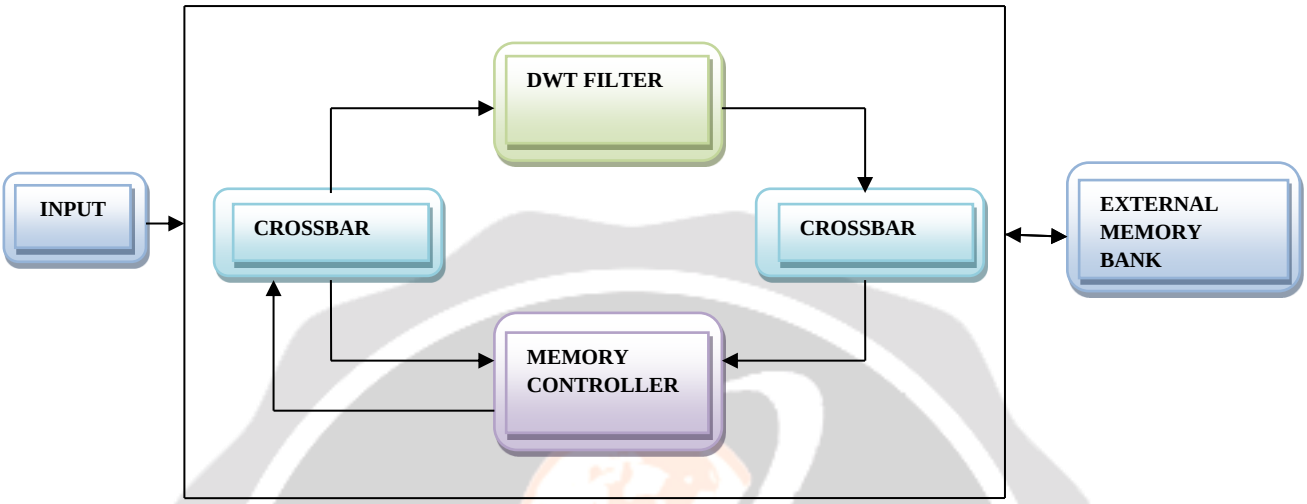

Figure 1. Block diagram of proposed design

Both high pass and low pass filters can be used to accomplish the discrete wavelet transform. The following transformations are used in the design of the high pass and low pass filters:

$$H(2n+1) = X(2n+1) - floor([X(2n) + X(2n+2)]/2) \quad \ldots(1)$$

$$L(2n) = X(2n) + floor([H(2n-1) + H(2n+1) + 2]/4) \quad \ldots(2)$$

These filters transform each pixel on a per-line basis, with lines being determined by the start-of-line (sol) and end-of-line (eol) markers. The image is divided into detail and approximation information by the high pass and low pass filters, respectively. The detail information mainly consists of high frequency, low scale picture components that add nuance. While the crucial portion of the image is conveyed by the high scale, low frequency components that make up the approximation information. New inputs are approved at one end of the high pass and low pass filters before previously accepted inputs start to show as outputs at the other end. Pipelining, a method that boosts the processor's speed, is this procedure. The output of the H and L filters will be allocated to the two memory banks in alternating fashion. In comparison to the 'L' outputs, the data on the 'H' outputs are 32 cycles later. The data coming from the "H" and "L" filters would always be attempting to write to the same memory bank if there were no delay. They always write to the opposing banks after the delay. A memory controller simultaneously executes the read and write operations. It does not take into account the filter's or the data's memory retrieval delay. Two free-running counters are the source of all memory control signals. The reset maintains the counts at 0 until the start pulse occurs. The master counter is used to identify the majority of control on a phase-by-phase basis. The number of phases is specified in the state registration. By recombining bits from the master counter during each phase, the address logic is created. The memory read for this phase is actually just a stride 1 loop around the entire memory bank; the read addresses are simply the count value. Each address is repeated twice in the write addresses for this phase.

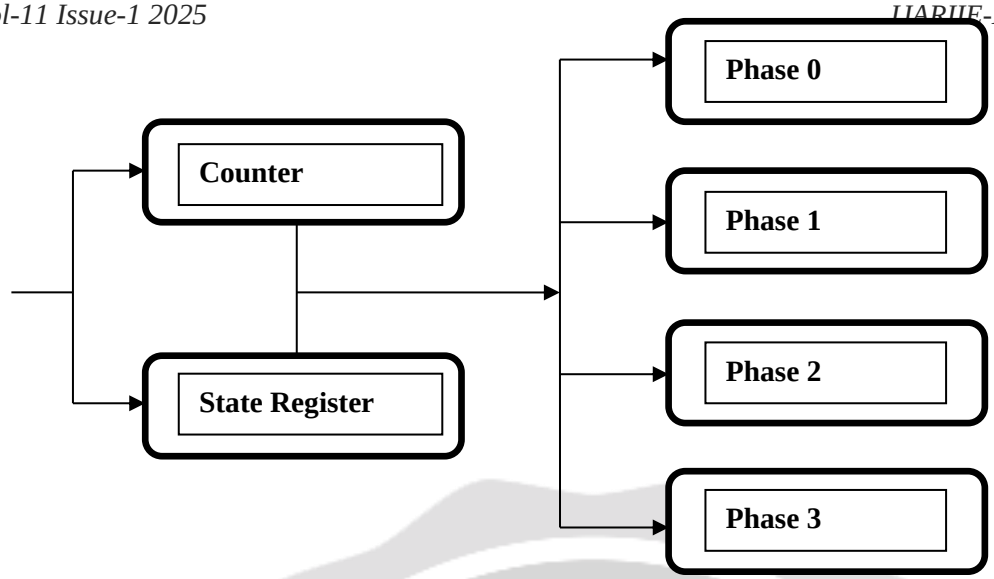

Figure 2. Memory Controller

The outside memory bank where the variable selector block's write enable flag is asserted. A subset of rows from the input are extracted by the variable selection, and P1 and P2 are given the result. These two products, P1 and P2, divide and multiply their inputs before passing them through a write inserter. Depending on the value of the second input, the write inserter passes the first or third input, and output is given to the read section. This indicates that one word has been added to a particular address point in an external memory bank. The suitable word is then selected by the read section from the memory vector. When addresses overlap, the read occurs before the write modifies the stored word.

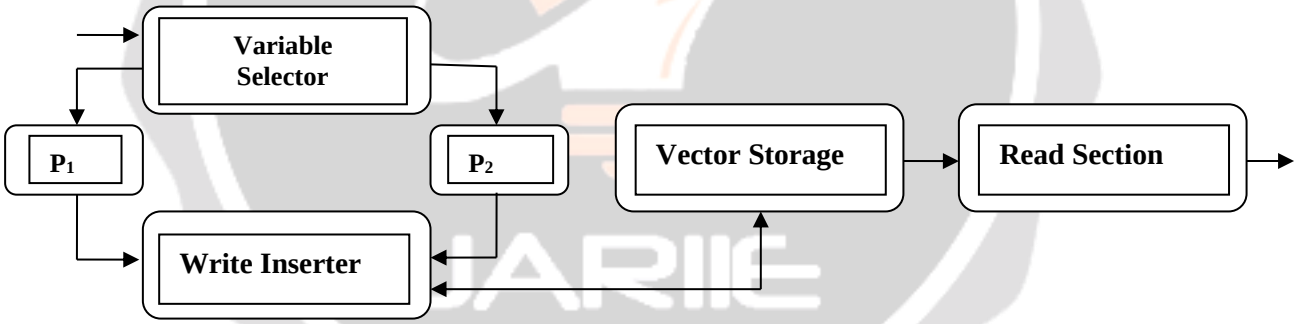

Figure 3. External Memory Bank

4. HARDWARE IMPLEMENTATION RESULTS

Xilinx System Generator block sets and Simulink were used to design and simulate the suggested model. The proposed model's DWT filter was used to simulate the process. The DWT filter divides the image into its detail and approximation information using high pass and low pass filters, respectively. Figure 4 depicts how the image decomposed.

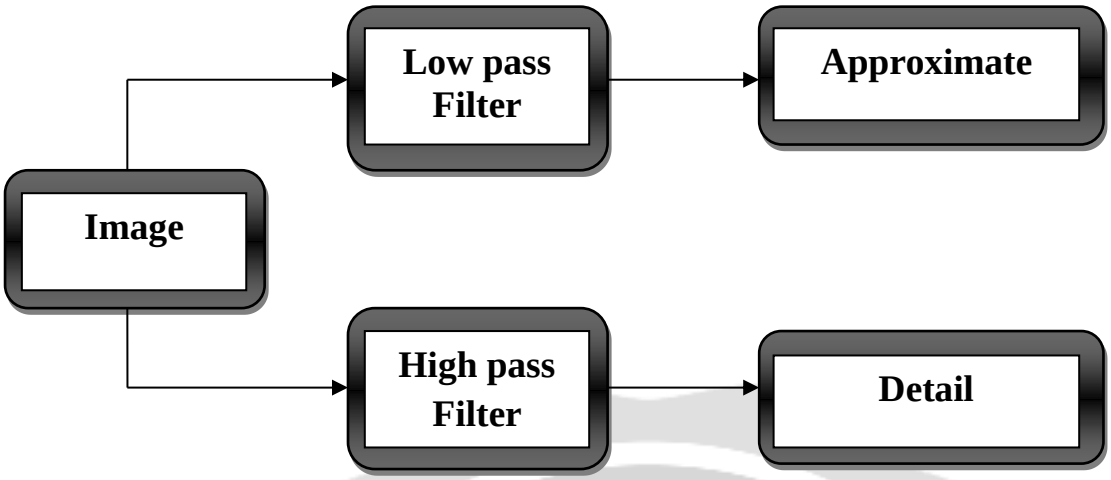

Figure 4. Decomposition of the Image

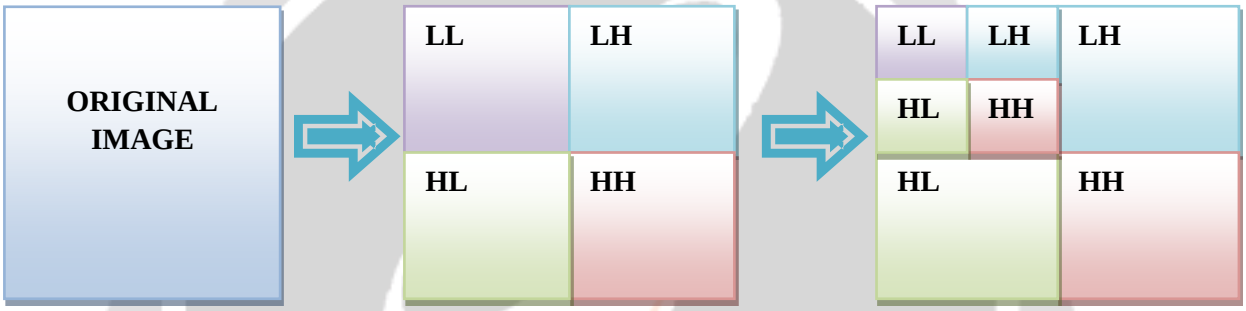

Figure 5 Original Image

Iteratively breaking down successive approximations into a large number of lower resolution components is possible. Alternatively, this is known as the wavelet decomposition tree. On the grayscale image displayed in figure 5, 2D-DWT is used. It sub-bands an image in a way that often results in lower level sub-band wavelet coefficients having greater energy than higher level sub-band wavelet coefficients.

Figure 6. Each row replaced by 1DWT

One-dimensional DWT filter can be used to do it in a separable way. In the first stage of the DWT, low-pass and high-pass filters are used to separate a picture into four sub-bands. Two phases make up the first stage of breakdown. A 1D vertical analysis filter bank is used to transform each row of a picture in the first stage. Figure 6 illustrates the initial phase.
Each column of the modified image is again horizontally changed in the second stage of the first level of decomposition using the same filter bank. Figure 7
displays the next phase. Four filtered and sub-sampled pictures are thus produced by the first level of decomposition.

The lowest sub-band is further divided for the second stage of decomposition by DWT using the same filtering strategy as before. Four additional sub-bands have been created from the lowest sub-band.

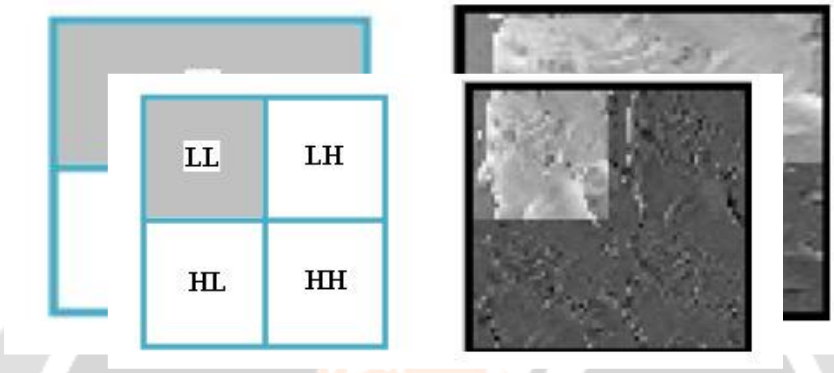

Figure 7. Each row replaced by 1DWT

In the lowest sub-band, 1D-DWT has been used to replace each row and column. Figure 8 depicts the outcome of the second degree of breakdown.

Modelsim was used to simulate the produced VHDL code; Xilinx Synthesiser Tool (XST) was used to synthesise it; and Spartan 2 and Spartan 3E-based FPGA devices—XC2S100-5tq144 and XC3S500E-4fg320, respectively—were

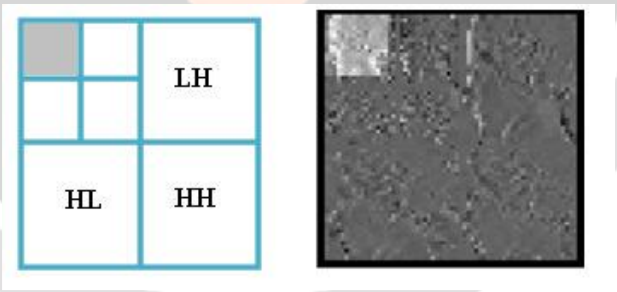

Figure 8. Second level of Decomposition

used to implement it. Table 1 displays the resource usage of an FPGA device based on Spartan 2 and 3E. Table 1 shows that the Spartan 3E based FPGA target device utilises fewer resources than the Spartan 2 based FPGA device, which improves speed and resource utilisation.

Table 1. Device Utilization Based on Spartan 2 and 3 FPGA

Device Utilization Summary						
Logic Utilization	Used		Available		Utilization	
	Spartan 2	Spartan 3	Spartan 2	Spartan 3	Spartan 2	Spartan 3
Number of Slice	278	272	1200	4656	23%	5%

Number of Slice Flip Flops	491	498	2400	9312	20%	5%
Number of 4input LUTs	363	334	2400	9312	15%	3%
Number of bonded IOBs	91	91	92	232	98%	39%
Number of GCLKs	1	1	4	24	25%	4%

Table 2 displays the timing summary for Spartan 2 and 3 based on XC2S100-5tq144 and XC3S500E-4fg320 FPGA devices, respectively. The proposed design may operate at a maximum operating frequency of 143 MHz when it is built on a Spartan 2 based FPGA. Additionally, the proposed design implemented on a Spartan 3E-based FPGA may operate at a maximum frequency of 231.27 MHz while using significantly fewer LUTS, flip-flops, and slice resources on the target device.

Table 2. Timing Summary Based on Spartan 2 and 3 FPGA

Description	Speed			
	Spartan 2		Spartan 3	
Minimum Period	6.990(ns)	143.04(MHz)	4.323(ns)	231.27(MHz)
Minimum input arrival time before clock	2.826(ns)	353.72(MHz)	1.945(ns)	513.86(MHz)
Maximum output required time after clock	8.188(ns)	122.10(MHz)	4.311(ns)	232.02(MHz)

XPower has been used to calculate the proposed design's overall power consumption based on the XC3S500E-4fg320 FPGA device. The proposed design consumed 117.86 mW at 28º Celsius.

After comparing with [8], proposed design results based on Spartan 2 & 3E FPGA have been released. Table 3 compares the proposed design's performance to those of [8] and Spartan 2 based designs, with the suggested design showing a 15% speed boost while using significantly less slices on the target FPGA device.

Table 3. Performance Comparison with existing models

Device Utilization Summary	DWT/IDWT Prototype Design	Proposed Design	
	Virtex 2	Spartan 2	Spartan 3E
Number of Slices	1907 out of 9280	278 out of 1200	272 out of 4656
Minimum period(ns)	4.973	6.991	4.324
Maximum Frequency(MHz)	201.092	143.041	231.267

Total estimated power consumption(mW)	861	429.93	117.86

5. CONCLUSIONS

In this paper, a DWT processor-based image compression model with great speed and area efficiency has been introduced. The introduction of the pipelined partially serial architecture boosts both speed and space effectiveness. The proposed design has a maximum operating frequency of 231.27 MHz and 117.86 mW power consumption at junction temperature of 28º C. By using significantly fewer resources on the Spartan 3E-based XC3S500E-4fg320 FPGA chip, a 15% speed boost has been made in order to offer affordable solutions for real-time image processing applications.

REFERENCES

1] Mungona and Ladhake, "VLSI Design approach for image compression using wavelet" International Journal of Application or Innovation in Engineering & Management (IJAIEM) Volume 4, Issue 8, August 2015.
[2] Usha Bhanu. N1 and Dr.A.Chilambuchelvan2 " Efficient VLSI Architecture for Discrete Wavelet Transform" IJCSI International Journal of Computer Science Issues, Special Issue, ICVCI-2011, Vol. 1, Issue 1, November 2011
[3] Joshni, C.George, T. Jayachandran, Dr. C.N Marimuthu "2D-DWT lifting based implementation using VLSI architecture"
[4] International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE) Volume 2, Issue 3, March 2013
[5] S. Jayachandranath, P. Suresh Babu " A High-performance VLSI Architecture for Image Compression Technique Using 2-D DWT" International Journal of Engineering Research and Applications (IJERA Vol. 2, Issue 6, November-December 2012, pp.1002-1006 1002
[6] Tarun Kumar and karun verma presented a document "A Theory Based on Conversion of RGB image to Gray image" International Journal ofComputer Applications (0975 – 8887) Volume 7– No. 2, September 2010
[7] Karthikeyan kamal kannan and Duraiswami proposed a paper "Conversion of Gray-scale image to Color Image with and without Texture Synthesis" IJCSNS International Journal of Computer Science and Network Security, VOL.7 No.4, April 2007
[8] Dr. Sudhir S. Kanade, Mrs. Nayna Vijaykumar Bhosale"a Lifting based DWT scheme for image compression using VHDL" International Journal of Computer Engineering and Applications, Volume IX, Issue VI, June 2015
[9] S.Vijayaraghavan 1, Dr.C.Parthasarathy gave a novel approach on "Fast DCT Based image compression Using FPGA" international journal of innovative research in electrical, electronics, instrumentation and control engineering vol. 2, issue 1, january 2014
[10] A. Mansouri, A. Ahaitouf, and F. Abdi, "An Efficient VLSI Architecture and FPGA Implementation of High-Speed and Low Power 2-D DWT for (9, 7) Wavelet Filter" IJCSNS International Journal of Computer Science and Network Security, VOL.9 No.3, March 2009
[11] Chien-Yu Chen, Zhong-Lan Yang, Tu-Chih Wang, Liang-Gee Chen "A Programmable VLSI Architecture for 2-D Discrete Wavelet Transform", ISCAS 2000 - IEEE International Symposium on Circuits and Systems, May 28-31, 2000, Geneva, Switzerland
[12] Anilkumar Katharotiya Swati Patel Mahesh "Comparative Analysis between DCT & DWT Techniques of Image Compression" Journal of Information Engineering and Applications ISSN 2224-5758 (print) ISSN 2224-896X (online) Vol 1, No.2, 2011
[13] Heba Raouf, Heba Yousef, Atef Ghonem "An Analytical Comparison between Applying FFT and DWT in WiMAX Systems" IJCSI International Journal of Computer Science Issues, Volume 11, Issue 6, No 2, November 2014 ISSN (Print): 1694-0814 | ISSN (Online): 1694-0784