

# Analysis and Implementation of Image Compression Approach Inexact Computing

Anshu Soni<sup>1</sup> and Mr. Naveen Khare<sup>2</sup>

M. Tech Scholar, Babulal Tarabai Institute of Research and Technology, Sagar, MP, India<sup>1</sup>

Assistant Professor, Babulal Tarabai Institute of Research and Technology, Sagar, MP, India<sup>2</sup>

anshusoni524@gmail.com and naveenkhare90@gmail.com

**Abstract:** This study proposes a novel approach for digital image processing, specifically targeting the issue of inaccurate calculations in discrete cosine transformation (DCT) compression. The technique aims to address and resolve certain difficulties related with DCT compression. The proposed system consists of three processing stages. In the first step, an approximated Discrete Cosine Transform (DCT) is used for picture compression. This eliminates the need for computationally demanding floating-point multiplication and instead utilizes integer additions and, in some cases, logical right/left shifts for DCT processing. The second level further reduces the amount of data (from the first level) that has to be processed by isolating the frequencies that are imperceptible to human senses. Ultimately, the third stage includes flawed circuit level adders to calculate the DCT, with the intention of reducing power use and delay. The suggested three-level method is used to compress a collection of structured photographs for measurement. The study compares many performance indicators, including energy consumption, delay, power-signal-to-noise ratio, average difference, and absolute maximum difference, with current compression techniques. Additionally, an error analysis is conducted to validate the simulation findings. The results demonstrate significant improvements in energy and time reduction, while simultaneously retaining acceptable levels of accuracy for image processing applications.

**Keywords:** Approximate computing, DCT, inexact computing, image compression

## I. INTRODUCTION

The quantity of information that modern computational and power computing systems handle nowadays is often substantial and resource-intensive. Digital Signal Processing (DSP) systems are extensively used for the manipulation of image and video data, particularly in mobile and wireless settings. These digital signal processing (DSP) devices use techniques and algorithms for compressing images and videos. Nevertheless, the requirements for power and performance continue to be quite strict. Compression techniques are often used to reduce such demands. Image and video compression technologies may be broadly classified into two categories: lossless and lossy. The latter group is more hardware efficient, but this comes at the cost of the quality of the final decompressed images/videos. The Joint Photographic Experts Group (JPEG) approach is commonly used for lossy image processing, whereas the Moving Picture Experts Group (MPEG) method is widely utilized for lossy video processing. Both standards use the Discrete Cosine Transform (DCT) algorithm as the fundamental processing step. Several diverse efficient methods for DCT computation [1], [2] have been devised for picture and video applications. However, all of these algorithms still need floating point multiplications, making them computationally demanding and requiring substantial hardware resources. In order to tackle these problems, the coefficients in various algorithms, like the one mentioned in reference [3], may be adjusted and estimated as integers. This allows for the substitution of floating-point multiplications with integer multiplications, as discussed in references [4] and [5].

The resultant algorithms exhibit greatly enhanced speed compared to the original versions, making them widely used in real applications. Therefore, there has been significant focus in recent years on developing efficient approximations of the Discrete Cosine Transform (DCT) that may be implemented using lower bus width and simpler arithmetic operations, such as shifting and addition [6]. One positive characteristic of image/video processing is its high tolerance for errors. This means that humans typically cannot detect any decrease in performance, such as a decrease in the

quality of visual and auditory information. Hence, imprecise computing may be used in many applications that can accommodate a certain level of imprecision and uncertainty, [7], [8], such as image/video processing. Introducing inaccuracies at the circuit level in the DCT calculation is aimed at certain performance measures, such as power dissipation, latency, and circuit complexity [9], [10], [11], [12], [13], [14]. This task is very demanding. This technique aims to address issues related to sluggish power and process tolerance by a comprehensive redesign of a specific circuit at the logic, gate, and transistor levels. An technique to circuit design, known as logic synthesis, has been suggested in order to create circuits that can implement a slightly different version of a given function. This approach takes into account the error rate (ER) as a measure of how much mistake may be tolerated. Truncating the circuits at the lowest bit positions in an adder circuit reduces its complexity at the transistor level. This reduction in complexity leads to a higher reduction in power dissipation compared to conventional low power design techniques. Additionally, new figures of merit have been introduced in [11] to estimate the error in an inexact adder. This study introduces a novel framework for approximation discrete cosine transform (DCT) picture compression. The system is built upon the principles of inexact computation and is structured into three distinct layers. Level 1 comprises a Discrete Cosine Transform (DCT) without any multiplication, therefore just requiring addition operations. Level 2 involves filtering the high frequency components (coefficients). Level 3 involves performing computations with imprecise adders. Level 1 has been extensively examined in the technical literature, as referenced in [16], [17], [18]. Level 2 is an intuitive method that aims to decrease computational complexity in picture compression while only experiencing a little decrease in performance. Level 3 utilizes a circuit level method that involves pursuing inexact computation. This is achieved by using innovative and efficient inexact adder cells in this article. Hence, the significance of this book lies in the cumulative impact of these three tiers. The suggested framework has undergone thorough analysis and evaluation. The findings of the simulation and error analysis demonstrate a significant level of concurrence in the outcomes of picture compression while using inexact computation. To prevent any ambiguity, the term "approximate" will be only used to refer to DCT methods, while the term "inexact" will be used for circuits and designs that use non-exact hardware to compute the DCT.

## II. REVIEW OF DCT

For manuscript completeness, preliminaries to approximate DCT and a review of relevant topics are presented next.

### 2.1 Discrete Cosine Transform (DCT)

To obtain the  $i$ th and  $j$ th DCT transformed elements of an image block (represented by a matrix  $p$  of size  $N$ ), the following equation is used:

$$D(i, j) = \frac{1}{\sqrt{2N}} C(i) C(j) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} p(x, y) \cos \left[ \frac{\pi(2x+1)i}{2N} \right] \cos \left[ \frac{\pi(2y+1)j}{2N} \right] \quad (1)$$

$$C(u) = \begin{cases} \frac{1}{\sqrt{2}}, & u = 0 \\ 1, & u > 0 \end{cases}$$

Where  $p(x, y)$  is the  $x, y$ th element of the image. This equation calculates one entry ( $i, j$ th) of the Transformed image from the pixel values of the original image matrix. For the commonly used  $8 \times 8$  Block for JPEG compression,  $N$  is equal to 8 and  $x$  and  $y$  from 0 to 7. Therefore  $D(I, j)$  is also given

By the following equation:

$$D(i, j) = \frac{1}{4} C(i) C(j) \sum_{x=0}^7 \sum_{y=0}^7 p(x, y) \cos \left[ \frac{\pi(2x+1)i}{16} \right] \cos \left[ \frac{\pi(2y+1)j}{16} \right] \quad (2)$$

For matrix calculations, the SCT matrix is obtained from the following:

$$T_{DCT}(i, j) = \begin{cases} \frac{1}{\sqrt{N}}, & i = 0 \\ \sqrt{\frac{2}{N}} \cos \left[ \frac{\pi(2j+1)i}{2N} \right], & i > 0 \end{cases} \quad (3)$$

So, DCT is computation intensive and may require floating-point operations for processing, Unless an approximate algorithm is utilized.

## 2.2 Joint Photographic Experts Group (JPEG)

The JPEG processing begins by converting a picture to the frequency domain using the Discrete Cosine Transform (DCT), which splits the image into different frequency components. Next, the process of quantization is carried out to eliminate frequencies that are considered less significant. This demonstrates the capacity of humans to have a pretty adept perception of subtle variations in brightness throughout a rather expansive region. However, they often struggle to discern the precise intensity of fast fluctuating changes in brightness. Compression occurs during the quantization process, when each frequency domain component is split by a constant and then rounded to the closest integer. As a consequence, several high frequency components exhibit negligible or perhaps zero magnitudes, at most extremely tiny magnitudes. During the decompression process, the picture is recovered by using just the preserved essential frequencies. The execution of JPEG processing requires the following steps:

1. An image (in color or grey scales) is first subdivided into blocks of  $k \times k$  pixels (usually  $k = 8$ ).
2. Then from left to right and top to bottom, the DCT is applied to each block.
3. This generates  $k \times k$  coefficients (so 64 for  $k = 8$ ) that are then quantized to reduce the magnitudes.
4. The resulting array of compressed blocks represents the compressed image, i.e., the stored or transmitted image.
5. To retrieve the image, the compressed image (array of blocks) is decompressed using Inverse DCT (IDCT).

## III. INEXACT ADDITION AND APPROXIMATE DCT

Table 1: Approximate DCT Methods Applied to Image Compression; Number of Operations Required to Calculate the DCT for an  $8 \times 8$  Block Size

	Method	Additions	Multiplications	Shifts	Total Operations
<b>Multipliers</b>	DFT by definition [1]	$56^a$ (432)	$64^b$ (192)	0	624
	DFT, Cooley-Tukey [1]	$24^a$ (58)	$2^b$ (6)	0	64
	DCT by definition [2]	56	64	0	120
	Arai algorithm [3]	29	5	0	34
<b>Multiplier-less</b>	SDCT [23]	24	0	0	24
	BAS08 [25]	18	0	2	20
	BAS09 [26]	18	0	0	18
	BAS11 [27] with $a=0$	16	0	0	16
	BAS11 [27] with $a=1$	18	0	0	18
	BAS11 [27] with $a=2$	18	0	2	20
	CB11 [28]	22	0	0	22
	BC12 [29]	14	0	0	14
	PEA12 [16]	24	0	6	30
	PEA14 [17]	14	0	0	14

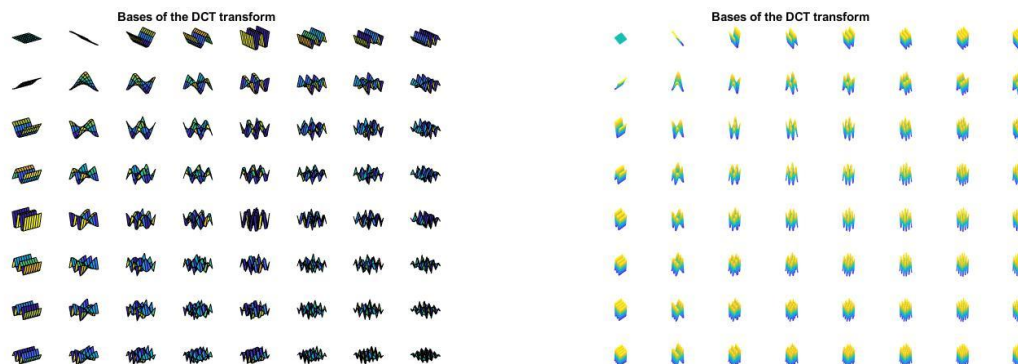
Arithmetic circuits are highly compatible with inexact computing. Addition has been thoroughly examined in technical literature and is considered a basic arithmetic operation in several applications of inexact computing [13]. A decrease in the complexity of the circuit at the transistor level in an adder circuit often results in a significant decrease in power dissipation, frequently surpassing the reduction achieved by standard low power design approaches [10]. The evaluation of inexact adder designs has been conducted in [12]. In these designs, the precise cell of a modular adder is either replaced with an approximation cell of reduced circuit complexity or the production and propagation of the carry in the addition process is modified. In [14], three novel inexact adder cell designs have been introduced (referred to as InXA1, InXA2, and InXA3). These cells possess advantageous electrical and error characteristics that are very suitable for approximation computation. The adder cells, as shown in Table 1, possess the following favorable characteristics in comparison to prior designs [10], [13]: (i) A minimal number of transistors; (ii) A minimal number of incorrect outputs at the two outputs (Sum and Carry); (iii) Reduced switching capacitance (measured as  $C_{gn}$  gate capacitance of the smallest NMOS), resulting in significant decreases in both delay and energy dissipation (as shown in Table 3) (and their combined metric). The table labeled as Table 1 presents metrics such as latency, energy wasted, and EDP (energy delay product) of the inexact cells for both average and worst instances. Among the imprecise cells, InXA1 exhibits the lowest average and most unfavorable delays, whereas InXA2 has the lowest average power dissipation and the lowest average energy-delay product (EDP). The average and worst case latency and energy dissipation of the adder cell have been calculated by rigorous simulation. The delay of each input signal is recorded at the point when the output reaches 90 percent of its maximum value, while the energy wasted is measured in all transistors at the period when the output also hits 90 percent. Considering these benefits, the InXA1 and InXA2 based adders are selected for the DCT application.

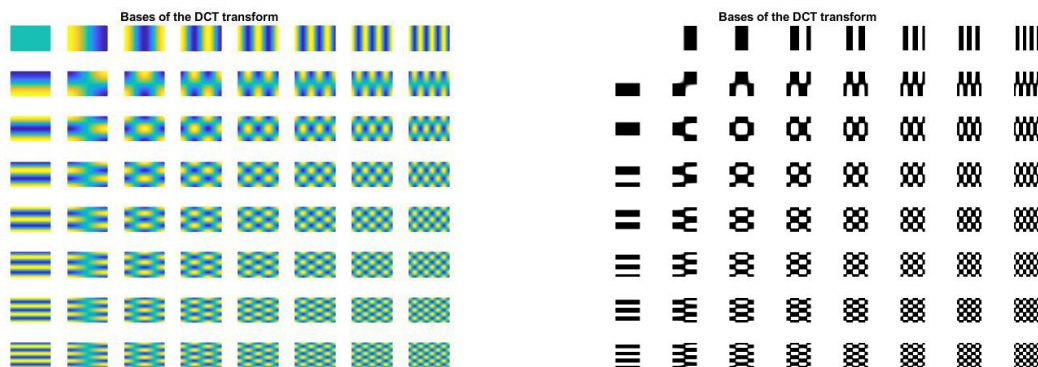
#### IV. PROPOSED APPROXIMATE FRAMEWORKS

This paper presents a new image compression framework that consists of three levels of approximation as follows. Level 1 refers to the DCT transformation without any multiplier, Level 2 involves high frequency filters, and Level 3 involves inexact calculation. The explanations for Levels 1 and 3 were provided in earlier sections. While high frequency filtering (Level 2) is not a novel idea, it is necessary to explain it in order to provide a comprehensive understanding, since it plays a role in the suggested framework for decreasing execution time and energy consumption. Hence, rather than applying the quantization process to all the DCT transformation coefficients obtained, it is only applied to the coefficients corresponding to the low frequency components of the modified block.

##### 4.1 High Frequency Filtration

When the high frequencies are filtered, the resulting picture is difficult for the human eye to detect since it is only sensitive to low frequency details. This functionality may be used to condense a picture.





As previously explained, a Discrete Cosine Transform (DCT) converts the picture into the frequency domain, allowing us to disregard the coefficients that represent high frequency components (which are not noticeable to the human eye) while preserving the remaining coefficients. Various numbers of preserved coefficients are taken into account for image compression purposes. It has been shown that a mere 0.34-24.26 percent of the 92112DCT coefficients are enough for high-speed face recognition applications [24], [25]. Here are some examples of 8x8 picture blocks:

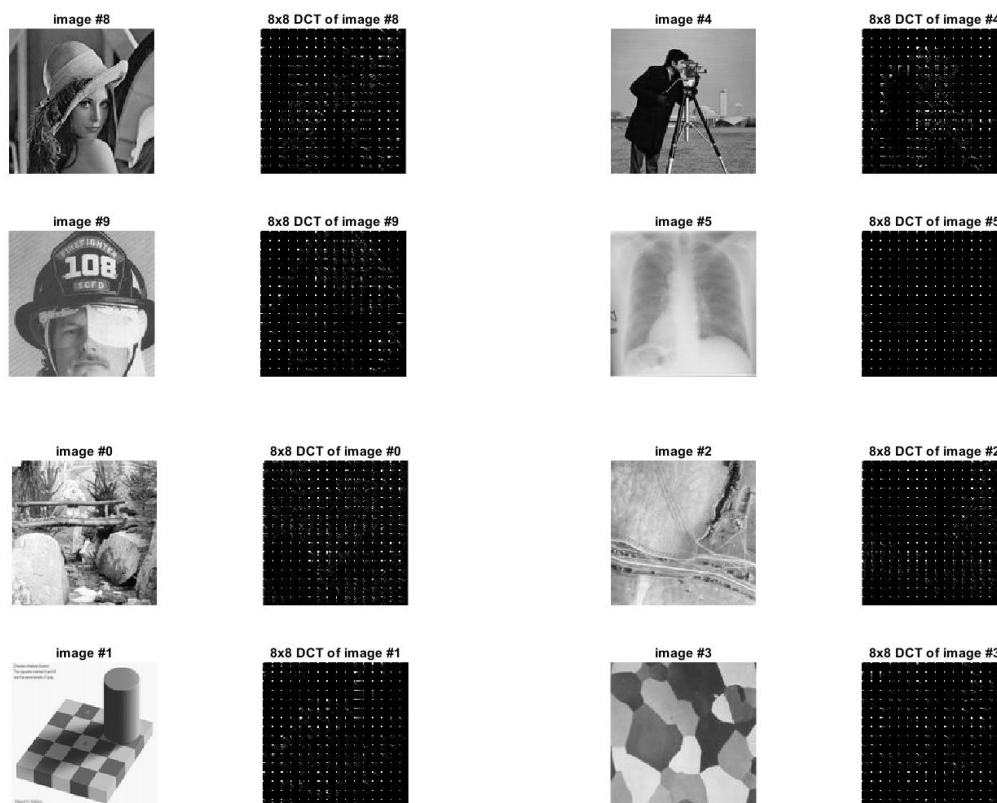


Image compression is achieved by using a support vector machine, which takes into account just the first 8-16 coefficients [16]. The suggested approach for picture reconstruction relies just on three coefficients, as said in [17]. The evaluation and assessment of different image compression methods, on the other hand, only use 10 coefficients, as stated in [25] and [23].



#### 4.2 Approximate DCT Implementation

In contrast to the approximation DCT methods shown in Table 1, all necessary computations (addition and subtraction) in this technique are performed at the bit level using the appropriate logic functions. All operators have a length of 32-bits and MATLAB simulates their implementations using Boolean logical functions. The Lena picture is used to simulate several DCT algorithms. The results are shown in Figure 1, which shows the Power Signal to Noise Ratio (PSNR) plotted against the number of Retained Coefficients (RC) utilized in the compression's quantization step.

The PSNR is calculated from the Mean Square Error(MSE) as follows:

Mean Square Error (MSE):

$$MSE = \frac{1}{m \times n} \sum_{x=1}^m \sum_{y=1}^n (p_{x,y} - \hat{p}_{x,y})^2 \quad (5)$$

Where  $p_{j,k}$  is the accurate pixel value at row  $x$  and column  $y$  of the image,  $\hat{p}_{x,y}$  is the approximate value of the same pixel,  $m$  and  $n$  are the size of the image (rows and columns respectively).

Peak Signal to Noise Ratio (PSNR):

$$PSNR = 10 \log \frac{(2^n - 1)^2}{MSE} \quad (6)$$

The results show that, except for the non-orthogonal SDCT method, compression using CB11 generally produces the best outcome in terms of PSNR. Three types of behavior are observed. Increasing output quality with an increase of the number of retained coefficients (RC). This occurs for CB11, BAS08, BAS09, BAS11 ( $a = 0$  and  $a = 1$ ), an almost constant PSNR by increasing the RC. This occurs for BC12 and PEA14, Degradation in output quality with an increase of RC. This occurs for both BAS11 ( $a = 2$ ) and PEA12. Two additional measures are used for a better insight on the resulting quality, i.e., the Average Difference (AD) and the Maximum Absolute Difference (MD).

These metrics are defined as:

Average Difference (AD):

$$AD = \frac{1}{m \times n} \sum_{j=1}^m \sum_{k=1}^n (p_{x,y} - \hat{p}_{x,y}) \quad (7)$$

Maximum Absolute Difference (MD):

$$MD = \max_{m,n} \{|p_{x,y} - \hat{p}_{x,y}|\} \quad (8)$$

Figs. 2 shows the resulting AD and MD for all methods; the average difference between the uncompressed and inexact-compressed images become smaller as RC increases except for BAS11( $a = 2$ ) and PEA12 (further confirming the PSNR results)

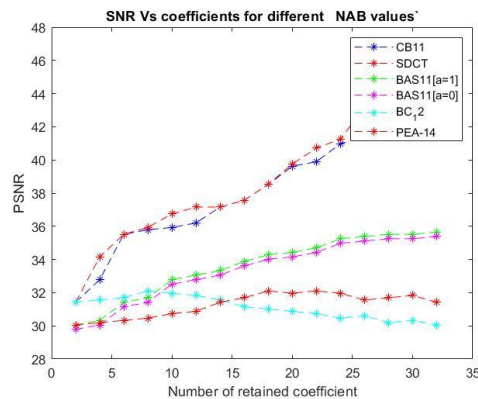


Fig1: Compression of an image using approximate DCT and bit-level exact computing.

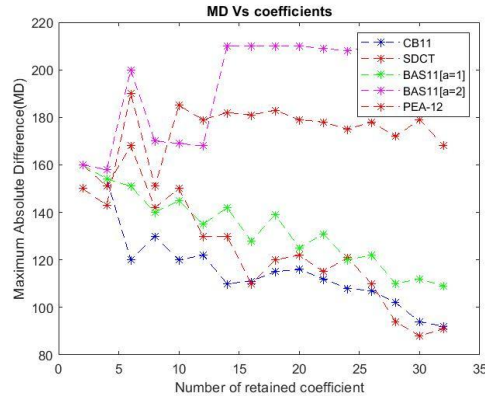


Fig2:Maximum absolute difference (MD) for compression of an image using approximate DCT

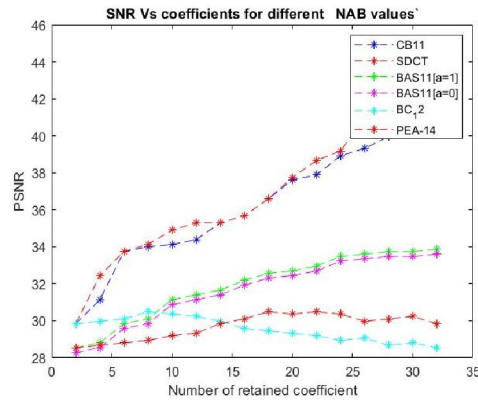


Fig3:Approximate DCT compression of an image using inexact adders with different NAB values; (Number of Approximate bits) NAB =3

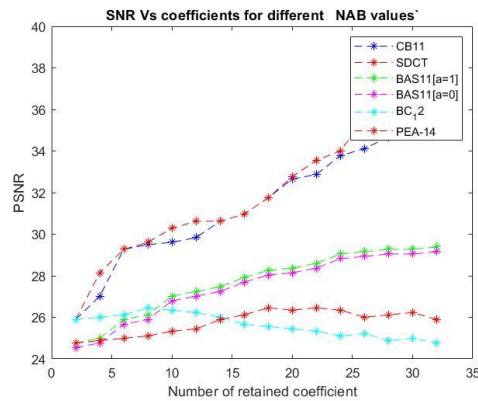


Fig4: Approximate DCT compression of an image using inexact adders with different NAB values; NAB =4

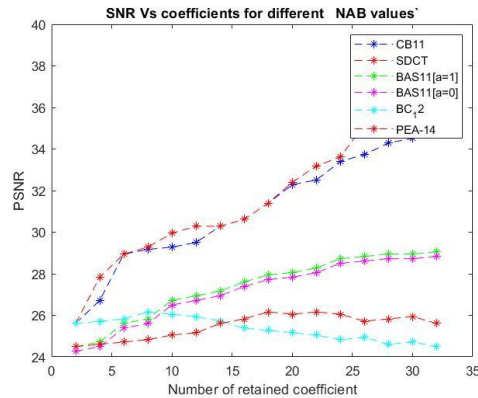


Fig5: Approximate DCT compression of an image using inexact adders with different NAB values; NAB =5

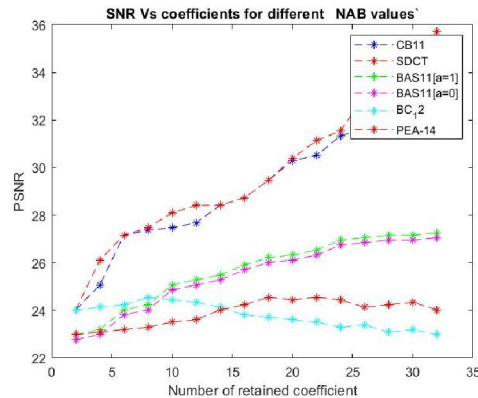


Fig6: Approximate DCT compression of an image using inexact adders with different NAB values; NAB =6

In Fig. 1). Fig. 2 shows that the MD between the uncompressed and inexact-compressed image pixels is reduced as more retained coefficients are used, the exceptions are PEA12 and BAS11 (a =2). This further confirms the previous results.

Fig. 4 depicts the compressed Lena image using the most accurate CB11 method for three RC values, i.e., 4, 10 and 16 retained coefficients. This figure also shows for comparison purpose the exact DCT compression results with RC = 16.

### 4.3 Approximate DCT Using Inexact Computing

Consider next the approximate DCT compression of Lena using inexact adders; as previously, the value of the NAB is increased from 3 to 6. The PSNR results are shown in Fig. 3,4,5,6 versus RC; the PSNR of the compressed images (a measure of quality) is plotted by executing all approximate DCT methods using only one inexact adder (for example the top row uses AMA1 as the inexact adder). Each column plots the quality of the compressed images by executing all approximate DCT methods with only inexact adders (for a NAB value). For example the left most column are for NAB = 3. As expected, the PSNR deteriorates as the NAB increases (an acceptable level of PSNR is reached at a NAB value of 4). 4.4 Truncation is one of the possible inexact computing techniques that may be utilized; the results of using truncation are also shown. The use of inexact adders results in more accurate results (truncation is performed at values of 3 and 4 bits).

## V. CONCLUSION

This study introduces a novel method for compressing pictures via the use of the Discrete Cosine Transform (DCT) algorithm, which employs approximation compression. The suggested solution has a 3-level structure. Firstly, a multiplier-less DCT transformation is performed, which only requires adds and shift operations. This is followed by



filtering and calculation of high frequency components using inexact adders. Research has shown that using  $8 \times 8$  picture blocks at each level significantly aids in approximating the compression process, ultimately resulting in a remarkably superior image quality. This work has verified that the collective impacts of these three levels are well comprehended; simulation and error analysis have shown a noteworthy consensus in outcomes for picture compression as an implementation of inexact computing.

The effectiveness of the suggested framework for a DCT approach that combines approximation at all three recommended levels has been shown. This publication confirms and validates the following particular conclusions via simulation and error analysis. Out of all the approximate Discrete Cosine Transform (DCT) techniques, CB11 demonstrates the greatest quality compression, as shown by the highest Peak Signal-to-Noise Ratio (PSNR) values. This is achieved when correct 16-bit adders are used, as shown in Figure 1. The PSNR findings were corroborated by other image alteration quality measurements, namely AD and MD. The figures labeled as 1 and 2. The techniques BAS08, BAS11 with  $a = 0$ , and BAS11 with  $a = 1$  are the most optimal approaches to use. Out of all the imprecise adders that have been mentioned [14], it has been determined that InXA2 achieves the highest performance. When approximation DCT JPEG compression is implemented using inexact adders, non-truncation-based approaches provide superior performance compared to truncation schemes, particularly when greater NABs are considered. The figures numbered 3, 4, 5, and 6. The results of the Discrete Cosine Transform (DCT) calculated with imprecise adders remain constant across diverse pictures. Typically, satisfactory compression may be achieved with NAB values of up to 4. Subsequent studies have shown a significant decline in the quality of the outcomes when higher NAB values are used. Based on the analysis of 4 image benchmarks, the BC12 and PEA14 approaches provide the lowest execution time and energy consumption while compressing a picture, as opposed to employing an exact adder. Regarding the measure of picture quality, the CB11 approach using the approximation DCT technique yields the greatest PSNR value. However, when considering both reductions in execution time and energy consumption, the BAS09 method is deemed the best approximate DCT method.

## REFERENCES

- [1]. R. E. Blahut, Fast Algorithms for Digital Signal Processing. Reading, MA, USA: Addison-Wesley, 1985
- [2]. V. Britanak, P. Yip and K. R. Rao, Discrete Cosine and Sine Transforms. New York, NY, USA: Academic, 2007
- [3]. Y. Arai, T. Agui, and M. Nakajima, "A fast DCT-SQ scheme for images," IEICE Trans., vol. E-71, no. 11, pp. 1095–1097, Nov. 1988.
- [4]. D. Tun and S. Lee, "On the fixed-point-error analysis of several fast DCT algorithms," IEEE Trans. Circuits Syst. Vid. Technol., vol. 3, no. 1, pp. 27–41, Feb. 1993.
- [5]. C. Hsu and J. Yao, "Comparative performance of fast cosine transform with fixed-point roundoff error analysis," IEEE Trans. Signal Process., vol. 42, no. 5, pp. 1256–1259, May 1994.
- [6]. J. Liang and T. D. Tran, "Fast multiplierless approximations of the fast DCT algorithms," IEEE Trans. Circuits Syst. Vid. Technol., vol. 3, no. 1, pp. 27–41, Feb. 1993.
- [7]. DCT with the lifting scheme," IEEE Trans. Signal Process., vol. 49, no. 12, pp. 3032–3044, Dec. 2001.
- [8]. Y. Dote and S. J. Ovaska, "Industrial applications of soft computing: A review," Proc. IEEE, vol. 89, no. 9, pp. 1243–1265, Sep. 2001.
- [9]. K. V. Palem, "Energy aware computing through probabilistic switching: A study of limits," IEEE Trans. Comput., vol. 54, no. 9, pp. 1123–1137, Sep. 2005.
- [10]. D. Shin and S. K. Gupta, "Approximate logic synthesis for error tolerant applications," in Proc. Des. Automat. Test Europe, 2010, pp. 957–960.
- [11]. V. Gupta, D. Mohapatra, A. Raghunathan, and K. Roy, "Low power digital signal processing using approximate adders," IEEE Trans. Comput.-Aided Des. Integrated Circuits Syst., vol. 32, no. 1, pp. 124–137, Jan. 2013.
- [12]. J. Liang, J. Han, and F. Lombardi, "New metrics for the reliability of approximate and probabilistic adders," IEEE Trans. Comput., vol. 62, no. 9, pp. 1760–1771, Sep. 2013.

- [13]. H. Jiang, J. Han, and F. Lombardi, "A comparative review and evaluation of approximate adders," in Proc. ACM/IEEE Great Lakes Symp. VLSI, May 2015, pp. 343–348.
- [14]. Z. Yang, A. Jain, J. Liang, J. Han, and F. Lombardi, "Approximate xor/xnor-based adders for inexact computing," in Proc. IEEE Int. Conf. Nanotechnology, Aug. 2013, pp. 690–693.
- [15]. H. A. F. Almurib, T. Nandha Kumar and F. Lombardi, "Inexact designs for approximate low power addition by cell replacement," in Proc. Conf. Des. Autom. Test Europe, Mar. 2016, pp. 660–665.
- [16]. N. Banerjee, G. Karakonstantis, and K. Roy, "Process variation tolerant low power DCT architecture," in Proc. Des. Automat. Test Europe, 2007, pp. 1–6.
- [17]. U. S. Potluri, A. Madanayake, R. J. Cintra, F. M. Bayer, and N. Rajapaksha, "Multiplier-free DCT approximations for RF multi-beam digital aperture-array space imaging and directional sensing," Meas. Sci. Technol., vol. 23, no. 11, pp. 1–15, Nov. 2012.
- [18]. U. S. Potluri, A. Madanayake, R. J. Cintra, F. M. Bayer, S. Kulasekera, and A. Edirisuriya, "Improved 8-point approximate DCT for image and video compression requiring only 14 additions," IEEE Trans. Circuits Syst. I Regular Papers, vol. 61, no. 6, pp. 1727–1740, Jun. 2014.
- [19]. N. Merhav and B. Vasudev, "A multiplication-free approximate algorithm for the inverse discrete cosine transform," in Proc. Int. Conf. Image Process., Oct. 24–28, 1999, pp. 759–763.
- [20]. C. Loeffler, A. Lightenberg, and G. Moschytz, "Practical fast 1-DDCT algorithms with 11 multiplications," in Proc. IEEE Int. Conf. Acoustics Speech Signal Process., Feb. 1989, pp. 988–991.
- [21]. P. Duhamel and H. H. Mida, "New S DCT algorithms suitable for VLSI implementation," in Proc. IEEE Int. Conf. Acoustics Speech Signal Process., 1987, pp. 1805–1808.
- [22]. E. Feig and S. Winograd, "On the multiplicative complexity of discrete cosine transform," IEEE Trans. Inform. Theory, vol. 38, no. 4, pp. 1387–1391, Jul. 1992.
- [23]. V. Lecuire, L. Makkaoui, and J.-M. Moureaux, "Fast zonal DCT for energy conservation in wireless image sensor networks," Electron. Lett., vol. 48, no. 2, pp. 125–127, Jan. 19, 2012.
- [24]. T. I. Haweel, "A new square wave transform based on the DCT," Signal Process., vol. 82, pp. 2309–2319, 2001.
- [25]. K. Lengwehasatit and A. Ortega, "Scalable variable complexity approximate forward DCT," IEEE Trans. Circuits Syst. Video Technol., vol. 14, no. 11, pp. 1236–1248, Nov. 2004.