Image Compression – Decompression using Polynomial Based Wavelet Transformation in PDF Document

[1]D.Selvanayagi, [2]Dr. S. Pannirselvam
[1]Ph.D. Reseach Scholar, [2]M.Sc., M.Phil., Ph.D., Associate Professor and Head,
Erode Arts and Science College, Erode, Tamilnadu, India

***Abstract*— Image compression technique is a request of data compression on digital images which encodes the original image with few bits. The intended of image compression is used to minimizing the redundancy of the picture and to keep or broadcast information in an efficient manner. It is carried out the problem of minimizing the huge data needed to characterize a digital image. In PDF application, extraction of images becomes more highly required to analyze the image content. But, the Lossless Compression is insufficient to measure the correlation amid the images and the DCT is not efficiently considered. The sparse representation of double discrete wavelet transform (DDWT) is a generative method of image blur that determines wavelet analysis of a blurred image. The DDWT offers less reliable for identifying object motion blur. Then, presented an image decomposition model is used to reduce the noise in the images but is less optimization provided. Then, High quality, noise removed image contents are stored in image file and subjected for image compression and decompression in PDCT-LIC technique. Finally, Polynomial Discrete Cosine Transformation Lossless Image Compression (PDCT-LIC) technique is developed to improve compression ratio, quicker compression/ decompression time, reduce compressed file size and minimize information loss.**

***Keywords*— DCT, PCDCT-LIC, DDWT**

I. INTRODUCTION

a. Image compression model

Image compression is an application of data compression which encodes the original image with a small number of bits. Image compression is minimizing the size in bytes of a graphics folder without loss some quality of the image to an unacceptable level. The goal of image compression is to decrease the redundancy of the image and to store or broadcast data in an efficient manner. Image compression technique is to identifying the problem of minimizing the amount of information that needs to represent a digital image.

All images are redundant data. Redundancy means the duplication of data in the image. Reduction of redundancy presents helps to achieve a saving of storage space of an image which is achieved through one or more of redundancies are decrease or removed. In image compression, three basic data redundancies are specified and exploited. This image compression technique is widely classified into basic two methods.

Lossless Compression Techniques is compresses the image by encoding all the information from the original file, hence, when the image is decompressed, it will be exactly equal to the original image. A novel lossless image compression system is derived from DCT method produced a significant reduction in entropy. Consequently, it create possible to get compression using a traditional entropy coder. The method performed well better than popular lossless JPEG method.

A lossy Compression technique presents a higher compression ratio than lossless compression. Lossy compression scheme include major consideration such as:

- Compression ratio
- Signal to noise ratio
- Speed of encoding & decoding

Lossy compression is mostly applied to compress multimedia data such as audio, video, and images particularly in many applications for streaming media and internet telephony. Vector quantization (VQ) is appropriate for lossy data compression method achieve lower bit rate and higher performance. It is also employed for lossy data correction and density calculation.

b. Image compression on DCT

The discrete cosine transform (DCT) is a technique that used to translate a signal into basic frequency mechanism. It is extensively employed in image compression. The JPEG image compression standard use DCT. DCT is utilized by many Non-analytical applications for example image processing and signal-processing (DSP) applications for video conferencing and to map an image space into a frequency.

Many advantages for DCT in image compression are ability to achieve low frequencies for image data and reduce the blocking artifact outcome. The discrete cosine transform is a very fast transform and achieve robust method for image compression. It has excellent compaction for highly correlated data. DCT has fixed foundation images provides better compromise connecting information packing ability and computational complexity.

DWT can be exploited to decrease the image size and values less than a pre-specified threshold are discarded. Therefore, it reduces the total of memory need to represent given image. DCT-based image compression performed on two methods for reducing the data required to represent the image. First, the quantization of the image and then entropy coding of the quantized coefficients are used. Quantization is the process of reducing the amount of possible values of a quantity. Therefore, reducing the amount of bits required to represent it.

c. Images decomposition model

Image decomposition model provides novel framework for image denoising. Decompression is achieved by applying the inverse of each of the preceding steps in opposite order. Hence, the decoding process starts with entropy decoding and proceeds to convert the run lengths to a sequence of zeros and coefficients. Coefficients are de-quantized and Inverse Discrete Cosine Transform (IDCT) is executed to retrieve the decompressed image. It is inverse processes of compression that is getting back the original image from compressed image.

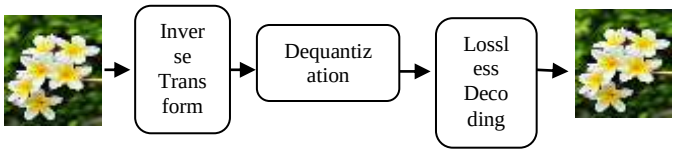

Fig: 1. Decompression model

II. REVIEW OF LITERATURE

A. Related works on compressed and decompressed model

Qionghai Dai et al., [QIO12] introduced a Robust joint reconstruction is compressed the Multiview imaging. A compressed multi-view imaging with robust joint reconstruction is developed to discover the pixel-level interview correlation. The pixel-to-pixel correspondence between inter-views image is known as disparity map efficiently formulate correlations among multi-view images. In addition to develop the quality of reconstructed image, the outliers results in disparity discontinuity, mismatching and illumination variance.

Xianming Liu et al., [XIA15] developed a new data-driven sparsity-based approach for the restoring JPEG-compressed images in the dual DCT-pixel domain. The main advantage is in the approach using residual redundancies of JPEG code streams and sparsity assets of latent images. The restoration is a sparse coding procedure performed together in the DCT coefficients of the latent image to remove the dispersal of quantization faults into the pixel domain. In addition to enhancing compressed images and videos by sophisticated post processing after decompression.

Tanaya Guha et al., [TAN13] presented the objective of sparse representation strategy is provided to encode the information content of an image. Sparsity of the representation is measured the compressibility regarding the other image. Therefore, the several sparse representation of an image is compressed through the maximum precise attained in image clustering, retrieval and classification and it more related to the other image.

Athira B. Kaimal et al., [ATH13] developed different types of image compression techniques. Lossy compression is applied to compact multimedia data of audio, video and images particularly in functions of streaming media. In lossless technique, image decoded with minimal information loss is improved for variety of applications. By differentiate lossless compression is needed for text articles and data files of bank details.

Gaurav Vijayvargiya et al., [GAU13] provide various Compression method is extensively different and then compression of binary raw data. The advantages of image compression are providing low cost reliable with broadcasting less data. In addition, it preserves low storage requirements and execution time and avoids the transmission errors while fewer bits are transferred.

Hanaa ZainEldin et al., [HAN15] discussed about the Image compression algorithms in WMSN. It is also provides the Image compression techniques for enhancing the performance to reduce computational load. With the objective of preserves the low power whereas acceptable visual quality is maintained.

B. Related works on Discrete Cosine Transform

Swati Das et al., [SWA15] explained about the Fast Image compression Technique using Hybrid method. Then, DWT achieve minimum size of the image at each level. It is useful to compressing a signals and multi resolution image analysis. In DCT+DWT Transform, the image compression is to decrease the storage size with high compression and minimum information loss.

Deepak Kumar Jain et al., [DEE14] presented a Discrete Cosine Transform (DCT) for compressing the image with good quality. By obtaining the results, the wavelet scheme performed to identify the wavelet filter caused by the smallest number of non- zero coefficients. Moreover, using a human visual system (HVS) based Processing Module to quantize wavelet coefficients and efficiently execute the Wavelet Transform.

Zahra Khanmirza et al., [ZAH15] presented a procedure for an appropriate compression through the fractal compression to attain high compression rate with suitable image quality. In addition, this procedure employs genetic algorithm to search for the best block with a high speed.

III. METHODOLOGY

A. Existing methodology identifying in the image compression and decompression

Lossy Image Compression is present limit of tolerance. This feature maximum the usefulness of compressed data to applications in which image content is continually decompressed other than never recompressed.

- ***Display only*** - Compressed images are viewed but not controlled. They are read-only.
- ***No selection*** - Compressed images are not changed and it pixels cannot be choose.
- ***Limited on-the-fly re-projection*** - A compressed image viewed only within a map the coordinate system which is more or less equal to that image.
- ***Reduced information content*** - The compression process reduces the original pixel information.
- ***No transparent pixels*** - ECW and JPEG2000 compression failed to preserve four channel (RGBa) information. Result is no channel accessible to collect transparent pixel information,images include transparent pixels are converted to a compressed image the transparent pixels will be modernized as black regions.

B. Proposed methodology

In order to overcome the drawbacks, the proposed method is developed to improve the image quality through the Vector Quantized Wavelet Transformation De-Noising (VQ-WTD) technique. First, Multi-Cluster Layout Segmentation technique is used to extract the image content from ebook in PDF files. Then, the extracted image contents are subjected to VQ-WTD technique that improves the image quality very quickly. Finally, Polynomial Discrete Cosine Transformation Lossless Image Compression (PDCT-LIC) technique is developed to minimize the information loss and improve compression ratio.

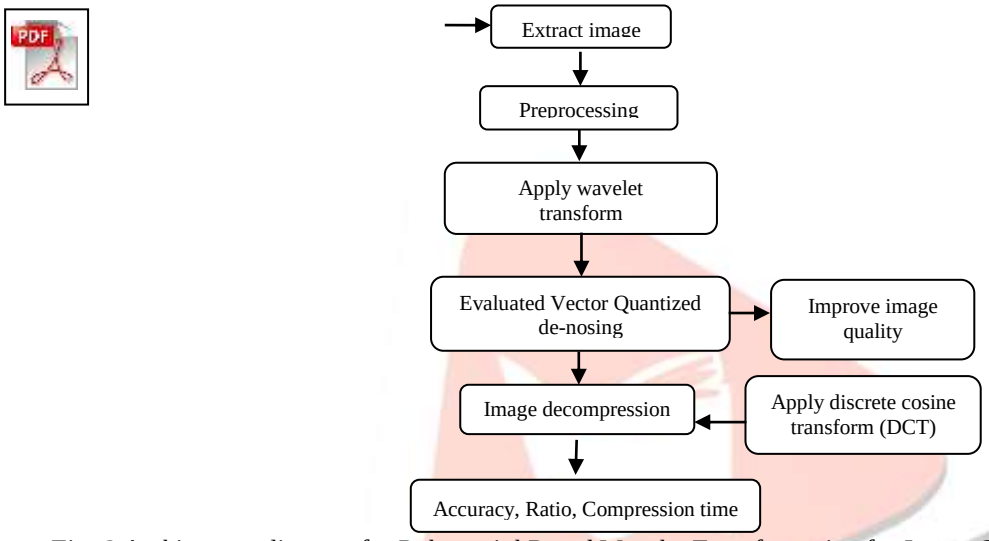

Fig: 2 Architecture diagram for Polynomial Based Wavelet Transformation for Image Compression-Decompression

C. Polynomial Discrete Cosine Transformation Lossless Image Compression (PDCT-LIC) technique

PDCT-LIC technique is used for image compression and decompression with higher compression ratio and minimal compression time and space. Discrete Cosine Transform (DCT) is used to every block and it translates pixels from left to right. The image is reconstructed during decompression, a process that employs the Inverse Discrete Cosine Transform (IDCT). Images are divided into parts of dissimilar frequencies by the DCT.

The DCT is employed in transformation for data compression. DCT is an orthogonal transform connected set of basis function utilized to map an image space into a frequency. This method produces an efficient decompression process with minimum information loss and less compression time. The output decompression images are approximately the same as input images. The DCT below taken as pixel (i,j) values and execute *DCT values (x,y) is expressed as,*

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

Input *: Images (img 1,img 2,….img n).*
Output : *Compressed images (c1,c2,…cn).*
Step 1 *: Image is divided into blocks of 8 x 8.Calculate DCT matrix.*
Step2 *:DCT is applied to each block by multiplying the modified block with DCT matrix on the left and transpose of DCT matrix on its right.*
Step4: *Each block is then compressed through quantization.*
Step 5:*Quantized matrix is then entropy encoded.*
Step6: *Compressed image (ci) is reconstructed through reverse process.*

Algorithm 1: Polynomial Discrete Cosine Transformation Lossless Image Compression (PDCT-LIC)

Input : *Compressed images (c1,c2,….cn).*
Output: *Images (img 1,img 2,….img n).*
Step 1: *Decode strings of bits into a 64-element array by using run length decoding technique.*
Step 2 : *Reform the 64-element array into an 8 × 8 block.*
Step 3: *Multiply each element in the block by a corresponding quantization value.*
Step 4: *Performs the IDCT on the de-quantized block.*
Step 5: *Add 128 to each of the de-transformed pixels to recover an approximation of the original image (img[i]*

Algorithm 2: Polynomial Discrete Cosine Transformation Lossless Image Decompression(PDCT-LIC)

IV. EXPERIMENTAL RESULTS

A. Purpose of MATLAB

MATLAB is extensively employed for computational device in science and engineering about the areas of physics, chemistry, math and all engineering streams. It is used in many fields such as Signal Processing and infrastructure, Image and video accessing, Control schemes, analysis and Measurement, Computational investment and Computational Biology. While debugging, we can change the value of a variable to see if the *new* value produces expected results. While the program is paused, assign a new value to the variable in the Command Window, Workspace browser, or Array Editor. Then continue running and stepping through the program. After identifying a problem, end the debugging session. It is best to quit debug mode editing an M-file. Clear breakpoints.

B. Results analysis

High quality, noise removed image contents are stored in image file and subjected for image compression and decompression using Polynomial Discrete Cosine Transformation Lossless Image Compression (PDCT-LIC) technique. The proposed method using following parameters are:

- Improve compression ratio
- Quicker compression/decompression time
- Reduce compressed file size
- Accuracy of image

The proposed method efficiently compared with existing four methods of Lossless Compression method (LCM) by Hao Wu et al., (HAO16], integer Discrete Tchebichef Transform (IDTT) by Bin Xiao et al., [BIN16], double discrete wavelet transform (DDWT) by Yi Zhang and Keigo Hirakawa [YIZ12] and Decomposition Framework for Image Denoising Algorithms (DF-IDA) by Gabriela Ghimpeteanu [GAB16].

a. Image Compression Ratio

Compression ratio is used to measure the ratio of compression being made with respect to the uncompressed size (of image) to the compressed size (image). The actual image size is taken that represents the uncompressed size of the image and the compressed size using different methods are obtained. Finally, the ratio of uncompressed size to the compressed size gives the actual compression ratio. The Compression ratio is mathematically formulated as,

$$Compression\ ratio = \left(\frac{Compresson\ being\ made\ with\ respect\ to\ uncompreesed\ size}{compressed\ size\ (image)}\right) * 100 \quad \text{...equ no.2}$$

TABLE 1. TABULATION FOR IMAGE COMPRESSION RATIO

Image size (KB)	Compression Ratio (%)				
	Existing LCM	Existing IDTT	Existing DDWT	Existing DF-IDA	Proposed VOD-PWT
15	95	96	89	97	101
30	98	98	91	99	103
45	101	100	93	101	106
60	103	103	95	103	108
75	105	105	99	105	111
90	107	106	101	107	115
105	100	109	103	109	117

Figure 3 shows the compression ratio based Wavelet Transformation image comparison based on the size of image in the dataset. As shown in the figure, by applying the VQD-PWT technique, compression ratio is improved even with the increase in the size of images in the dataset.

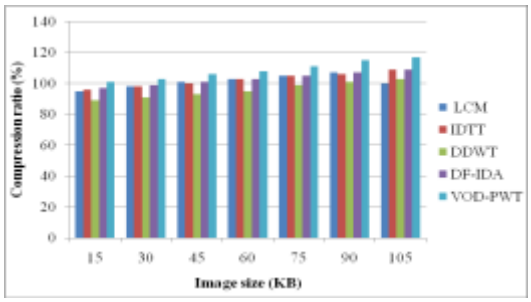

Fig: 3. Measures of Image Compression Ratio

On average, the proposed DF-IDA outperforms the LCM by 7%, 6% compared to IDTT, 12% compared to DDWT and 5% compared to DF-IDA respectively. In the proposed DF-IDA technique, Discrete Cosine Transformation is performed that improves the compression ratio for different set of images.

b. Quicker Compression/Decompression Time

Compression/decompression time is measured by the total amount time taken for the image compression and decompression based on different image size. It is measured in terms of (ms). It also reduces the time required for images compression to be sent over the Internet or downloaded from Web pages.

TABLE 2. QUICKER COMPRESSION/DECOMPRESSION TIME (MS)

Image size (KB)	Compression/Decompression Time (ms)				
	Existing LCM	Existing IDTT	Existing DDWT	Existing DF-IDA	Proposed VOD-PWT
15	10	8	11	9	4
30	12	12	13	11	6
45	14	14	14	13	8
60	16	16	16	15	11
75	18	18	19	17	13
90	21	20	21	18	15
105	23	21	23	22	17

Figure.4. illustrates the quicker compression / decompression time based Wavelet Transformation image comparison based on the size of image in the dataset. As shown in the figure, by applying the VQD-PWT technique, quicker compression/decompression time is reduced even with the increase in the size of images.

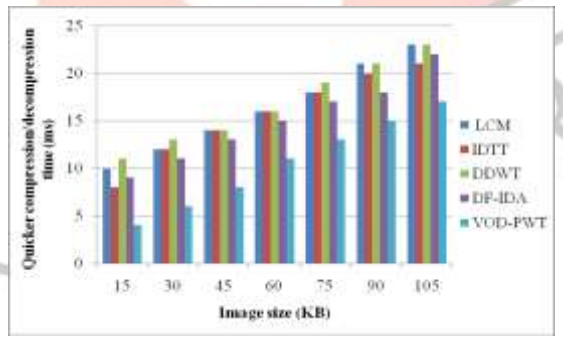

Fig:4. Measures of Quicker Compression/Decompression Time

Subsequently, the proposed VQD-PWT technique outperforms the LCM by 69%, 59% compared to IDTT, 76% compared to DDWT and 43% compared to DF-IDA respectively. In the proposed VQD-PWT technique, Discrete Cosine Transformation is performed that reduces the compression/decompression time.

c. Impact of Compressed File Size

The image compression file size is measured that allows more images to be stored in a given amount of disk or memory space. In addition to reduce the file size by making selections about the resolution of an image and the quality or compression of an image is provided. Reducing image size doesn't reduce the image quality.

$$Compressed\ file\ size = \frac{more\ images\ to\ be\ stored\ in\ a\ memory\ space\ (KB)}{total\ memory\ space\ (KB)}$$

TABLE: 3. COMPRESSED FILE SIZE

Image size (KB)	Compressed File Size (KB)				
	Existing LCM	Existing IDTT	Existing DDWT	Existing DF-IDA	Proposed VOD-PWT
15	22	25	27	29	18
30	26	29	30	35	21
45	30	31	33	37	23

60	34	34	37	40	28
75	38	37	40	42	31
90	41	41	43	44	33
105	43	43	46	47	36

Figure.5 demonstrates the compressed file size using Wavelet Transformation image comparison based on the size of image in the dataset. As shown in the figure, by applying the VQD-PWT technique, compressed file size is reduced even with the increase in the size of images.

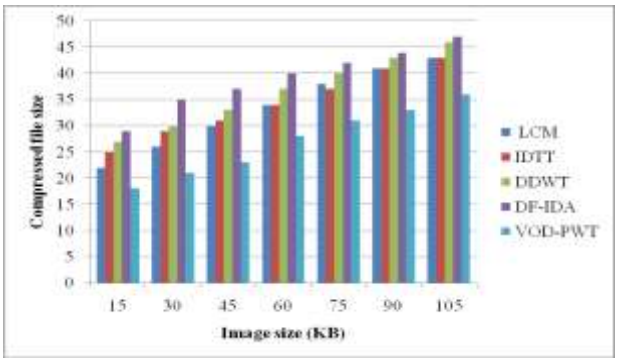

Fig: 5. Measures of Compressed File Size

Subsequently, the proposed VQD-PWT technique outperforms the LCM by 23%, 28% compared to IDTT, 36% compared to DDWT and 47% compared to DF-IDA respectively. In the proposed VQD-PWT technique, Discrete Cosine Transformation is performed that reduces the compressed file size.

d. Accuracy as on image decompression

Information loss is an error condition in information systems where information is destroyed by failures or avoid in storage, transmission, or processing. Information systems employ backup and disaster recovery equipment and procedures to avoid Information loss or restore lost Information. The Information loss is difference between the total number of image content and loss of image content information.

TABLE. 4. ACCURACY AS ON IMAGE DECOMPRESSION

No. of information in images	Accuracy (%)				
	Existing LCM	Existing IDTT	Existing DDWT	Existing DF-IDA	Proposed VOD-PWT
5	13	15	12	14	10
10	16	16	15	16	12
15	19	18	18	18	14
20	22	21	21	21	16
25	24	22	23	24	19
30	26	24	24	27	21
35	28	26	29	28	22

Figure.6 expresses the information loss using Wavelet Transformation image comparison derived from the No. of information in images in the dataset. As revealed in the figure, by applying the VQD-PWT technique, information loss is reduced even with the increase in the No. of information in images.

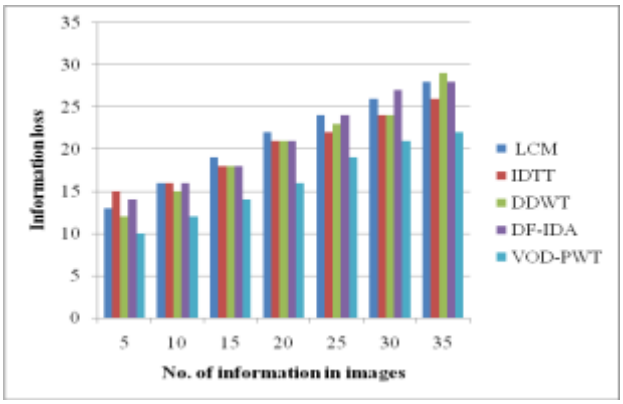

Fig: 6. Measures of Accuracy

Consequently, the proposed VQD-PWT technique out performs the LCM by 30%, 27% compared to IDTT, 24% compared to DDWT and 31 % compared to DF-IDA correspondingly. In the proposed VQD-PWT technique, Discrete Cosine Transformation is performed that reduces the information loss then the existing four methods.

V. CONCLUSION

Vector Quantized De-Noising and Polynomial Based Wavelet Transformation is used for Image Compression-Decompression of Image Contents. The proposed work presents a Multi-Cluster Layout Segmentation scheme for extracting image content from in PDF file. Then, Vector Quantized Wavelet Transformation De-Noising (VQ-WTD) technique is used to enhance the image quality with maximum compression ratio. Finally, Polynomial Discrete Cosine Transformation Lossless Image Compression (PDCT-LIC) technique is developed to achieve high quality and noise removed image contents are stored in image file and subjected for image compression and decompression with minimum information obtained.

References

[1] Qionghai Dai, Changjun Fu, Xiangyang Ji, YongbingZhang, "Robust Joint Reconstruction in Compressed Multi-view Imaging", Picture Coding Symposium, 2012, Pages 13-16.

[2] Xianming Liu, Xiaolin Wu, Jiantao Zhou and Debin Zhao, "Data-driven Soft Decoding of Compressed Images in Dual Transform-Pixel Domain", IEEE Transactions on Image Processing, pages 1-13.

[3] Tanaya Guha and Rabab K Ward, "Image Similarity Using Sparse Representation and Compression Distance", JOURNAL DRAFT, VOL. X, NO. X, APR 2013, pages 1-20.

[4] Athira B. Kaimal, S. Manimurugan, C.S.C .Devadass, "Image Compression Techniques: A Survey", International Journal of Engineering Inventions, Volume 2, Issue 4 (February 2013) PP: 26-28.

[5] Gaurav Vijayvargiya, Sanjay Silakari and Rajeev Pandey, "A Survey: Various Techniques of Image Compression", (IJCSIS) International Journal of Computer Science and Information Security, Vol. 11, No. 10, October 2013.

[6] Hanaa ZainEldin, Mostafa A. Elhosseini, Hesham A. Ali, "Image compression algorithms in wireless multimedia sensor networks: A survey", Ain Shams Engineering Journal (2015) 6, pages 481–490.

[7] Swati Das, Chandrashekhar Kamargaonkar, Dr.Monisha Sharma, "An efficient and Fast Image compression Technique using Hybrid methods", International Journal of Engineering and Technical Research (IJETR), Volume-3, Issue-4, April 2015 pages 497-500.

[8] Deepak Kumar Jain, Devansh Gaur, Kavya Gaur, Neha Jain, "Image Compression using Discrete Cosine Transform and daptive Huffman Coding", International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), Volume 3, Issue 1, January – February 2014, pages 90-94.

[9] Zahra Khanmirza, Farhad Ramezani, Homayon Motameni, "OPTIMIZATION OF FRACTAL COMPRESSION OF 3D IMAGES USING A GENETIC ALGORITHM", Advances in Science and Technology Research Journal, Volume 9, No. 26, June 2015, pages 124–128.

[10] Gabriela Ghimpe¸teanu, Thomas Batard, Marcelo Bertalmío, and Stacey Levine, "A Decomposition Framework for Image Denoising Algorithms", IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL: 25, NO: 1, JANUARY 2016, pages 388-399.

[11] Hao Wu, Xiaoyan Sun, Jingyu Yang, Wenjun Zeng, and Feng Wu, "Lossless Compression of JPEG Coded Photo Collections", IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 25, NO. 6, JUNE 2016, Pages: 2684-2696.

[12] Bin Xiao, Gang Lu, Yanhong Zhang, Weisheng Li and Guoyin Wang, "Lossless Image Compression Based on Integer Discrete Tchebichef Transform," Year:2016,Pages1-16.

[13] Yi Zhang and Keigo Hirakawa, "Blind Deblurring and Denoising of Images Corrupted by Unidirectional Object Motion Blur and Sensor Noise," JOURNAL OF LATEX CLASS FILES, VOL. 11, NO. 4, DECEMBER 2012.