DCT Based Image Compression with Llyod's Quantization and Variable Block-Size

Prabal Devkota
BE graduate
Kathmandu University
Dhulikhel, Nepal
prabaldevkota19@gmail.com

Nabin Bhusal
BE graduate
Kathmandu University
Dhulikhel, Nepal
nabinbhusal101@gmail.com

Avimannu Bhandari
BE graduate
Kathmandu University
Dhulikhel, Nepal
avimannubhandari1997@gmail.com

Madhav Prasad Pandey*
Assistant Professor
Kathmandu University
Dhulikhel, Nepal
mpp@ku.edu.np

Abstract: The number of decision-levels in Llyod's quantization and DCT block-size affect the average number of bits to represent symbols as well as the visual quality of reconstructed image. In this paper, a model is presented that has an ability to adapt a wide range of block-size and different numbers of decision-level in quantization. The model overcomes the drawback of an existing JPEG-based image compression algorithm that deploys fixed block size DCT and quantization table. A test image was subjected to varying DCT block-size persisting the decision level constant, and varying decision level persisting the block-size constant. The trials show reconstructed image of diverse parameters having Peak Signal-Noise-Ratio (PSNR) 42.36 dB, Mean Square Error (MSE) 3.78 and Compression Ratio (CR) 3.33 to PSNR 8.4 dB, MSE 9415 and CR 5.71.

Keywords: Image compression, DCT, Llyod's quantization, DCT block-size

I. INTRODUCTION

A digital image is usually a two-dimensional array of pixels. In the raw form, the image may require a huge amount of memory. The objective of compression is to store or transfer image data efficiently by reducing the following redundancies [1], [2].

1. Coding redundancies: Each pixel in an image is mostly represented by 8-bit codes which contain more bits than actually needed to represent that intensity.
2. Interpixel redundancies: Redundancies existing in correlated values.
3. Psychovisual redundancies: Redundancies due to more than required spectral component for human vision.

The DCT transform widely used in image compression has two major properties that are decorrelation and energy compaction. The decorrelated coefficient reduces the interpixel redundancies. The energy compaction property of the DCT transform represents highly correlated input image data by a few numbers of uncorrelated DCT coefficients. Both these properties are significant to remove redundant information. The 2D DCT of N*N input sequence can be expressed mathematically as in Eq. (1) [3].

$$C(u,v) = \alpha(u)\alpha(v)\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y)\cos\left[\frac{\pi(2x+1)u}{2N}\right]\cos\left[\frac{\pi(2y+1)v}{2N}\right] \quad (1)$$

Similarly, the Inverse Discrete Cosine Transform (IDCT) can be obtained by using the Eq. (2).

$$f(x,y) = \sum_{u=0}^{N-1}\sum_{v=0}^{N-1} \alpha(u)\alpha(v)C(u,v)\cos\left[\frac{\pi(2x+1)u}{2N}\right]\cos\left[\frac{\pi(2y+1)v}{2N}\right] \quad (2)$$

Where, $\alpha(v) = \begin{cases}\sqrt{\frac{1}{N}} & for\ v = 0\\ \sqrt{\frac{2}{N}} & for\ v \neq 0\end{cases}$; $\alpha(u) = \begin{cases}\sqrt{\frac{1}{N}} & for\ u = 0\\ \sqrt{\frac{2}{N}} & for\ u \neq 0\end{cases}$;

$u, v = 0,1,2,\dots,N-1.$

The compressed image output performance can be evaluated in terms of MSE, PSNR and CR mathematically written as in Eq. (3), Eq. (4), Eq. (5) [4].

$$MSE = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}\left(A_{i,j}-B_{i,j}\right)^2}{M*N} \quad (3)$$

Where, A = Original matrix of size M×N, B = Reconstructed matrix of size M×N

$$PSNR = 20log_{10}\left(\frac{R}{MSE}\right) \quad (4)$$

Where, R = Maximum intensity level,

$$CR = \frac{n1}{n2} \quad (5)$$

Where, n1 = uncompressed file size, n2 = compressed file size

Llyod max quantization is an optimization technique that generates optimal codebooks for specified partitions. The algorithm continuously executes unless stopping criteria or threshold is achieved [5]. In MATLAB the function can be initialized with the length of the codebook vector which is the number of bits to represent L-decision level. The function gets terminated when the relative change in distortion between iterations is less than $1 * 10^{-07}$.

The existing JPEG-based image compression uses a fixed block-size of 8*8 pixels to perform forward DCT. The quantization table (Q-table) developed for specific block-size cannot adapt to variable block-size. In order to overcome such problems, a model is presented that can adopt varying block-size using Llyod's quantization. This research aims to study the effect of block-size and the number of decision-levels of Llyod's quantization in the output image.

The proposed model consists of parameters such as scale factor, block size and the number of decision level in Llyod's quantization. These parameters can be varied to get reconstructed image of different size and quality. The Llyod's quantization deployed in the model will not discard higher spectral component entirely that might contribute better PSNR and quality of the reconstructed image. The Llyod's quantization in a larger block size increases statistical redundancy that can decrease the size of output image. The reconstructed image of diverse quality can be incorporated in different application based on their usability.

II. IMAGE COMPRESSION FRAMEWORK

An image compression model shown in Fig. 1 consists of two major parts that are encoder and decoder. The encoder comprises various sub-blocks. An original image matrix of M*M pixel is segmented and divided into blocks of different sizes. The image matrix after segmentation contains N*N pixel elements in each block. The N*N element of every block undergoes 2D-Forward DCT transformation. The resulting DCT transformed coefficients of every block are concatenated to represent a single matrix of M*M elements. The concatenated matrix containing DCT coefficients is given a positive offset by the absolute minimum of the transformed coefficient. The resulting large value after offset is scaled down and quantized using Llyod's quantizer of a fixed decision-level. The output of Llyod's quantization is encoded using the statistical coding called Huffman encoder. The output of the Huffman encoder, Huffman dictionary and Huffman code gives the compressed output. The decoding stages follow Huffman decoding, scaling up and providing a negative offset to the transformed coefficients. The quantization used during encoding is permanent and lossy thus, cannot be recovered. The next step follows inverse DCT implemented in a block to get reconstructed images. The framework deployed doesn't require knowledge of Q-table for quantization during encoding and decoding as commonly used in JPEG standards.

III. Experimental Setup And Observation

The study is carried out in MATLAB of version R2016a. The custom and default functions were used to evaluate and analyze empirical evidences. The code is available at *https://github.com/Prabal1998/DCT-Image-compression*. An 8-bit image of TIFF format [6] is accessed publicly from the image database. The image matrix is reformed to subject study in a grayscale image of 200*200 pixels as shown in Fig. 2.

The 2D DCT coefficients are obtained for different blocks of size 4, 8, 20 and 50 respectively. These coefficients are scaled down by factor 5 and quantized using Llyod's quantization by fixing number of decision-level 512. The quantized coefficients are then encoded using Huffman encoding to get a compressed output data. The image is reconstructed following the decoding steps and image parameter shown in Table I is obtained. The observation shows that increase in block-size increases MSE and CR of the test image while PSNR decreases. The reconstructed images in Fig. 3 shows that blocking artifacts are being noticed in an image of large block size.

Similarly, fixing the DCT block-size by 4 and scale factor by 5, the observation is made for decision-level 512, 256, 128, and 8 respectively. The result in Table II demonstrates that increase in decision-level decreases MSE and CR in the output images while PSNR increases. The effect of coarse quantization in Fig. 3 is visually noticed in an image reconstructed using low number of decision level.

IV. Results And Conclusion

The DCT coefficients of a large block are highly uncorrelated spanning in a long-range interval while the small block-size coefficients are sparsely correlated constrained in a small-range as shown in Fig. 4. The use of 512 decision-level in Llyod's quantization represents DCT coefficients more appropriately in a small block-size. Therefore, a small block-size image yields better PSNR and lesser MSE during reconstruction. In addition to this, histogram plot in Fig. 5 shows that larger block-size images have greater counts of DCT coefficients concentrated around zero. The DCT coefficients of a larger block represented by a few numbers of decision-level can increase statistical redundancies and CR of the output image.

When the quantization level is increased for a given block-size, the MSE of the DCT coefficient after quantization is decreased as observed in Table III. This signifies that DCT coefficients after quantization are nearly correlated to DCT coefficients before quantization. So, during reconstruction most spectral components are preserved which produces a better PSNR and lower MSE. In contrast, CR may decrease to represent an additional decision level.

The trial shows that reconstructed image PSNR in the rage of 36 to 42 dB and MSE in the range of 3 to 17 doesn't exhibit noticeable artifacts. Also, blocking artifacts is not noticed when block size less than 20.

The quantitative parameters such as MSE and PSNR of the reconstructed image will not always ascertain visual quality of an image, so the subjective parameter that relies on human perception can also be introduced. The overall image subjective and quantitative parameter can be considered and evaluated in term of Mean Opinion Score (MOS). This is not included in this paper due to time constraint and encouraged in the future work.

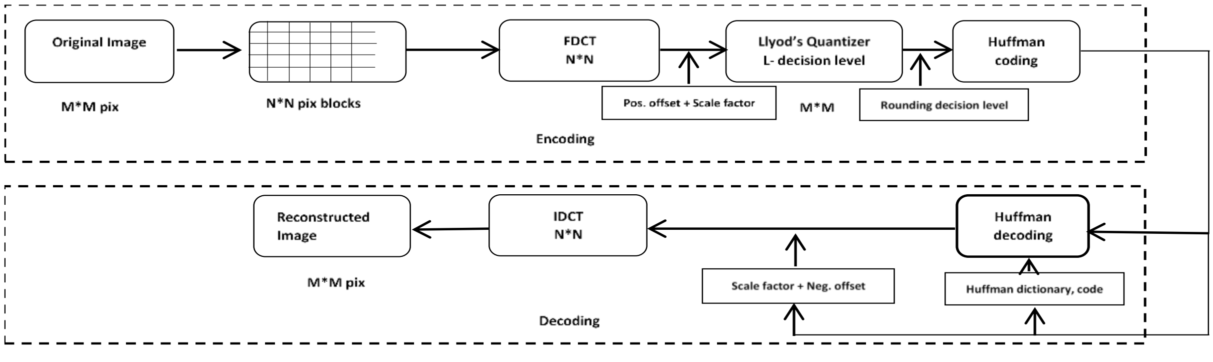

Fig. 1: Block diagram of deployed Image compression model

Fig. 2: An original grayscale image of 200*200 pix

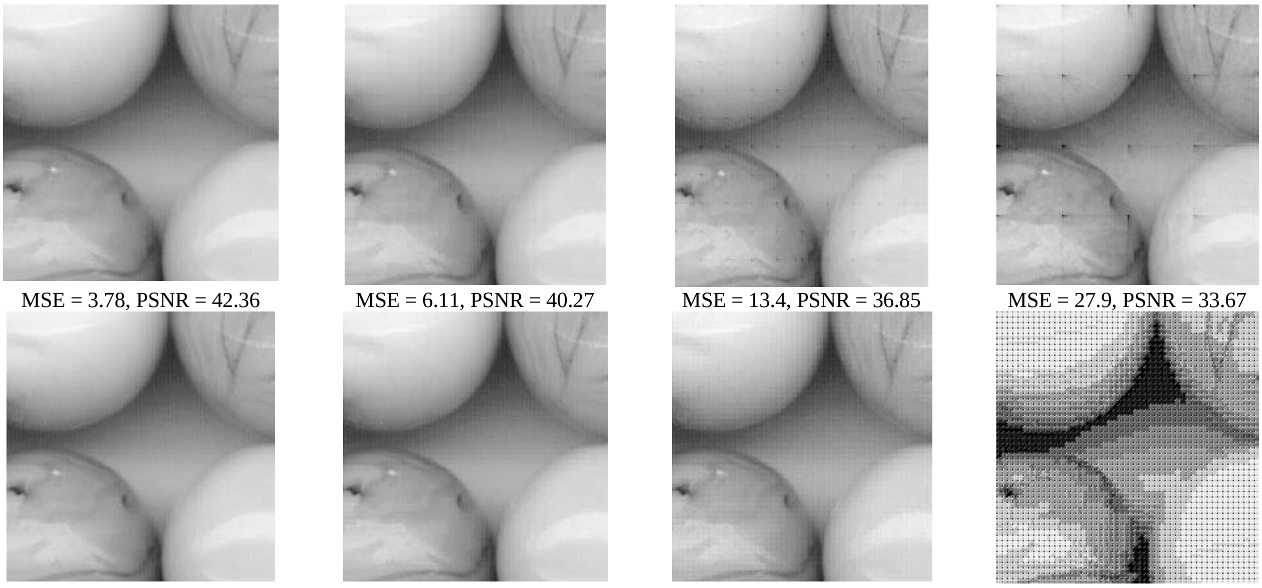

Fig. 3: Reconstructed images and parameters. (PSNR is in dB)
Top row: Reconstructed image for DCT block-size 4, 8, 20, 50 from left to right respectively when decision-level 512 was fixed.
Second row: Reconstructed image for Decision-levels 512, 256, 128, 8 from left to right respectively when block-size 4 was fixed.

TABLE I RECONSTRUCTED IMAGE PARAMETERS FOR DECISION-LEVEL -L 512 AND SCALE FACTOR 5

Test image	Block size	MSE	PSNR (dB)	CR
Grayscale Marbles4.tiff	4	3.78	42.36	3.33
	8	6.11	40.27	4.44
	20	13.4	36.85	5
	50	27.9	33.67	5.07

TABLE II RECONSTRUCTED IMAGE PARAMETER FOR BLOCK-SIZE 4 *4 AND SCALE FACTOR 5

Test image	L-level	MSE	PSNR (dB)	CR
Grayscale Marbles4.tiff	512	3.78	42.36	3.33
	256	6.15	40.24	3.63
	128	17.2	35.77	3.77
	8	9415	8.393	5.71

TABLE III QUANTIZATION PARAMETER FOR BLOCK-SIZE 4*4 AND SCALE FACTOR 5.

No. of L-level	MSE
512	0.1510
256	0.2460
128	0.6881
8	376.61

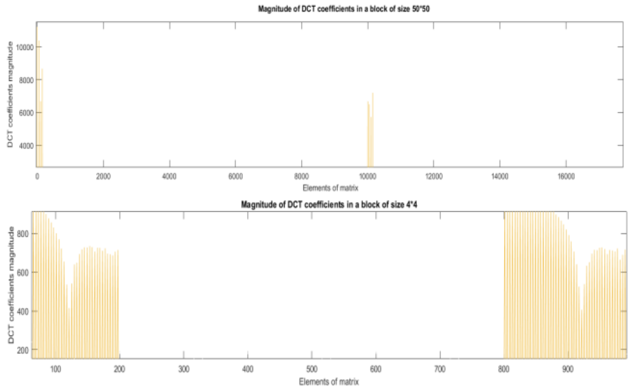

Fig. 4: Magnitude of DCT coefficients observed in a different block-size when decision-level 512 is fixed.

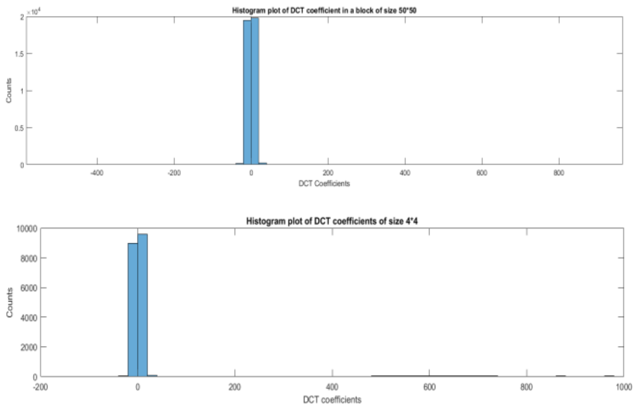

Fig. 5: Histogram plot of DCT coefficients observed in different block-size

References

[1] R. C. Gonzalez and R. E. Woods, *Digital image processing (2nd edition)*. Prentice Hall, 2002.

[2] R. R. Sindhu M, “Images and Its Compression Techniques - A Review,” *International Journal of Recent Trends in Engineering*, vol. 2, no. 4, pp. 71–75, 2009.

[3] S. Ali Khayam, “The Discrete Cosine Transform (DCT): Theory and Application1,” Jan. 2003. https://www.researchgate.net/publication/242275100_The_Discrete_Cosine_Transform_DCT_Theory_and_Application1 (accessed Mar. 05, 2023).

[4] K. Bindu, A. Ganpati, and A. K. Sharma, “A Comparative Study of Image Compression Algorithms,” *International Journal of Research in Computer Science*, vol. 2, no. 5, pp. 37–42, 2012, doi: 10.7815/ijorcs.25.2012.046.

[5] S. P. Lloyd, “Least Squares Quantization in PCM,” *IEEE Trans Inf Theory*, vol. 28, no. 2, pp. 129–137, 1982, doi: 10.1109/TIT.1982.1056489.

[6] “TIFF Sample Files: Metadata for MARBLES.TIF.” https://www.fileformat.info/format/tiff/sample/1f37bbd5603048178487ec88b1a6425b/download (accessed Mar. 17, 2023).