IZMİR INSTITUTE OF TECHNOLOGY

ELECTRICAL AND ELECTRONICS ENGINEERING DEPARTMENT

EE452 PROJECT REPORT

"Implementation of DCT based Image Compression Techniques at Different Compression Rates (P1)"

Group Members

Atakan Aygün	: 240206074
Hasan Hüseyin Harman	: 230206044
Mert Uztemur	: 230206009
Date	: 18/01/2021

1) INTRODUCTION

An image is a visual representation of something that has been created or copied and stored in electronic form and can be represented as vector graphics [1]. The digital image consists of pixels which are the smallest units of images that contain color intensities such as red, green, and blue.

During the transmission of an image or other media files, the files need to be compressed to gain bandwidth and power. Image compression is done by deleting unnecessary or irrelevant information in the image and encoding the remaining information in a way that is not noticed in this transformation [2]. Image compression can be examined under two main headings. One of them is a lossy compression, also called irreversible compression which is a data encoding method that approximates the original signal by removing some parts [3]. Another one is lossless compression that constructs the data without any loss so that there is no noise in the reconstructed signal. There are techniques of lossy compression, such as Quantization, Transform Coding, and techniques of lossless compression such as Linear Predictive Coding, Multi-resolution Coding. The compression ratio is the fundamental distinction point between lossy and lossless compression paths. The compression ratio is the ratio between the number of bits in the original file and the compressed file. The rise of CR provides a gain in bandwidth and power in transmission but decreases image quality. However, the human eye is not sensitive to higher frequencies. Thus, some frequency components are deleted without loss of quality for desired compression ratio in lossy compression techniques such as Discrete Fourier Transform, Discrete Cosine Transform, and Discrete Wavelet Transform.

Discrete Cosine Transform is one of the lossy compression techniques that represent the signal as cosine functions at different frequencies. Unlike Discrete Fourier transform, DCT of signal consists of real values, not complex values. In practice, DCT is preferred over DFT in most conversion systems because DFT coefficients require twice as much memory space as DCT coefficients [4].

This project uses DCT to implement image compression for varying block sizes and quality levels.

2) SYSTEM MODEL AND DESIGN

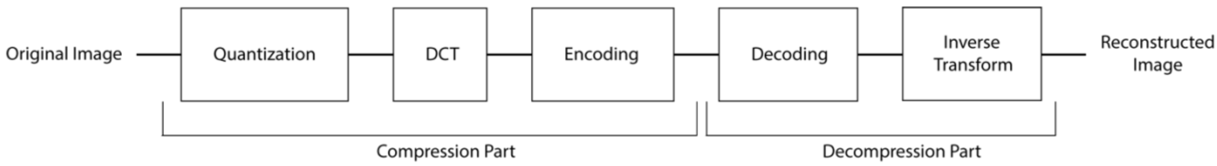


Figure 1: Block diagram of the project

Figure 1 shows the block diagram of this project. Each step explained in the below sections starting with Discrete Cosine Transform (DCT). DCT is a method of orthogonal transformation that decomposes an image to its spectrum of spatial frequency. To apply DCT, the image is divided into N-by-N pixel matrices. The N-by-N DCT basis matrix is constructed by using cosine basis functions for x and y dimensions. The formula below shows the implementation of DCT operation for each pixel on N-by-B submatrices.

$$\mathbf{DCT(i,j)} = \frac{\mathbf{1}}{\sqrt{\mathbf{2N}}}\mathbf{C(i)C(j)}\sum_{\mathbf{x=0}}^{\mathbf{N-1}}\sum_{\mathbf{y=0}}^{\mathbf{N-1}}\mathbf{I(x,y)}\cos\left[\frac{\mathbf{(2x+1)i\pi}}{\mathbf{2N}}\right]\cos\left[\frac{\mathbf{(2y+1)j\pi}}{\mathbf{2N}}\right] \quad [5]$$

where, I(x,y) is the intensity level for image, N is the block size and,

$$C(u) = \begin{cases} \frac{1}{\sqrt{2}} \; if\ u = 0 \\ 1 \quad otherwise \end{cases}$$

In matrix form,

$\boldsymbol{D} = \boldsymbol{T} \times \boldsymbol{A} \times \boldsymbol{T'}$, where **A** is the submatrix block of the image and **T** is the DCT basis matrix.

$$T_{i,j} = \left\{ \begin{array}{ll} \frac{1}{\sqrt{N}} & \text{if } i = 0 \\ \sqrt{\frac{2}{N}} \cos\left[\frac{(2j+1)i\pi}{2N} \right] & \text{if } i > 0 \end{array} \right\}$$

[6]

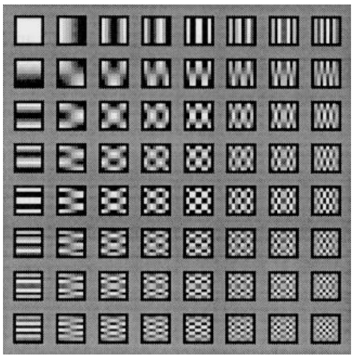


Figure 2: DCT basis matrix

Quantization is the process of mapping continuous input values to a set of discrete finite values. This process provides flexibility to compress the image at the desired compression ratio and adjust the quality of an image. The quality levels from 1 to 100 are chosen for "1" as lowest quality and highest compression and "100" as the highest quality and lowest compression.

Quantization starts as follows: DCT coefficient matrix is element-wise divided by a quantization matrix. The matrix obtained is rounded to the nearest integer value. In this way, redundant frequency components would be zero [7].

$$DCT_{quantized_{i,j}} = Round(\frac{DCT_{i,j}}{Q_{i,j}})$$

$$Q_{10} = \begin{bmatrix} 80 & 60 & 50 & 80 & 120 & 200 & 255 & 255 \\ 55 & 60 & 70 & 95 & 130 & 255 & 255 & 255 \\ 70 & 65 & 80 & 120 & 200 & 255 & 255 & 255 \\ 70 & 85 & 110 & 145 & 255 & 255 & 255 & 255 \\ 90 & 110 & 185 & 255 & 255 & 255 & 255 & 255 \\ 120 & 175 & 255 & 255 & 255 & 255 & 255 & 255 \\ 245 & 255 & 255 & 255 & 255 & 255 & 255 & 255 \\ 255 & 255 & 255 & 255 & 255 & 255 & 255 & 255 \end{bmatrix} \quad Q_{50} = \begin{bmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \\ 72 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \end{bmatrix} \quad Q_{90} = \begin{bmatrix} 3 & 2 & 2 & 3 & 5 & 8 & 10 & 12 \\ 2 & 2 & 3 & 4 & 5 & 12 & 12 & 11 \\ 3 & 3 & 3 & 5 & 8 & 11 & 14 & 11 \\ 3 & 3 & 4 & 6 & 10 & 17 & 16 & 12 \\ 4 & 4 & 7 & 11 & 14 & 22 & 21 & 15 \\ 5 & 7 & 11 & 13 & 16 & 12 & 23 & 18 \\ 10 & 13 & 16 & 17 & 21 & 24 & 24 & 21 \\ 14 & 18 & 19 & 20 & 22 & 20 & 20 & 20 \end{bmatrix}$$

Figure 3: Comparison of the quantization matrices

After the quantization process, the coefficients of the quantized matrix are encoded as a binary data sequence by starting from low-frequency components to high-frequency components. The method used is called the zigzag scanning method. In information theory, most probable streams are encoded as fewer bits. Since low-frequency components carry more energy than higher frequency components, it is logical to map low-frequency components to shorter codewords.

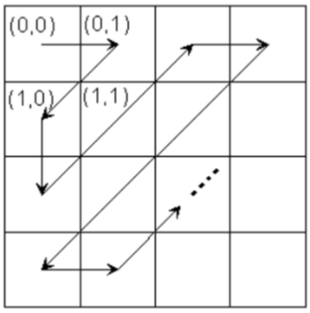


Figure 4: Zigzag encoding representation

The reconstruction process of the image occurs by applying the reverse of the procedures described in the previous topics. Firstly, the binary data sequence is decoded, and the inverse zig-zag method is applied to the transformed matrix. Each element of the matrix is multiplied by the corresponding element of the original quantization matrix. Finally, inverse discrete cosine transform is applied to the obtained matrix to get the original image. Since the DCT transformation matrix, denoted as **T**, is a real orthonormal matrix, the inverse of matrix **T** is equal to its transpose. Inverse DCT operation is implemented using formula below.

$\mathbf{IDCT = T' \times D \times T}$, D denoted as 8 by 8 block of the DCT of image.

The peak signal to noise ratio (PSNR) is used as the criterion for the quality of the image in lossy compression techniques. PSNR is a measure of the peak error between the compressed image and the original image. It is measured by:

$$\mathbf{PSNR\ dB = 20\log_{10}\left(\frac{MAX_I}{\sqrt{MSE}}\right)} \quad [8].$$

where $\mathbf{MAX_I}$ is the highest possible power of the original image signal, the cumulative square error between the compressed images and the original images is represented by MSE.

$$\mathbf{MSE = \frac{1}{m * n}\sum_{i=0}^{m}\sum_{j=0}^{n}[I(i,j) - K(i,j)]^2} \quad [9].$$

A lower MSE value means a lower error and as seen from the inverse MSE-PSNR relationship, this leads to a higher PSNR value.

3. NUMERICAL RESULTS

The test code was written in MATLAB environment using the steps described in the previous sections. Firstly, the picture taken as input was divided into RGB components. After being quantized, it passed through the encoding and decoding stages. In the first stage, the compression grade to three different quality levels was examined. Then, the block size effect on the compression level was examined with three different block sizes. Values of 10, 50, and 90 were used for the quality levels, and values 2, 4, and 8 were used for the block size. The examinations were compared using Compression Ratio (CR), Mean Square Error (MSE), and Peak Signal to Noise Ratio (PSNR) values.

First, the quality level was examined for the fixed block size. The results in Table 1 were obtained. Since there was no visible difference in the reconstructed pictures, DCT space images were also added to figure 6.

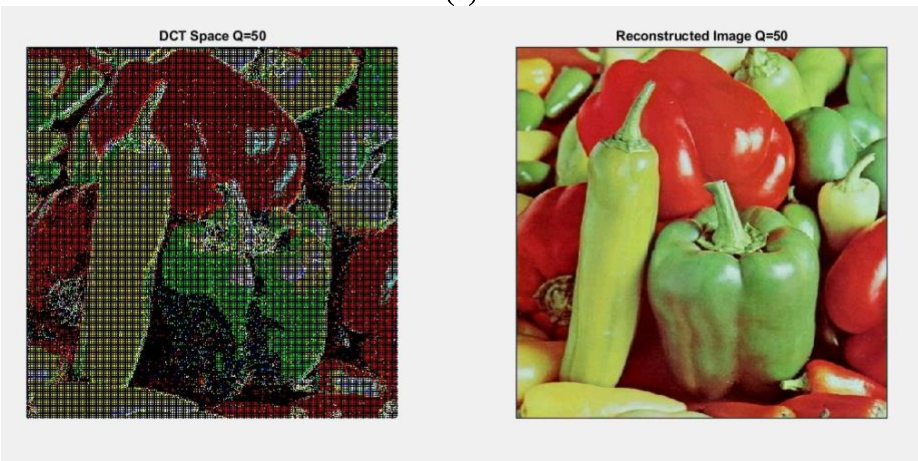
Quality Level	Compression Level
10	High compression
50	Standard compression
90	Low compression

Table 1: Quality level comparison

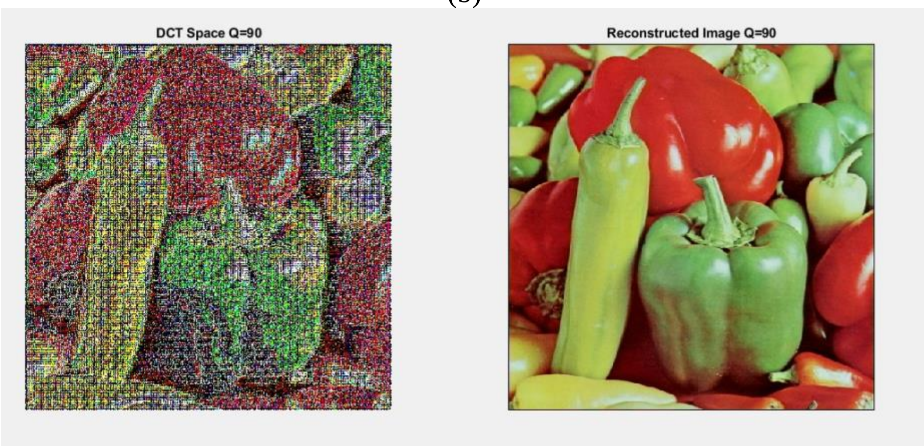
As a general evaluation, quality level 50 was observed to be the most optimal level. It produced a good size image without much loss of image quality.

Figure 5: Original peppers image.

(a)



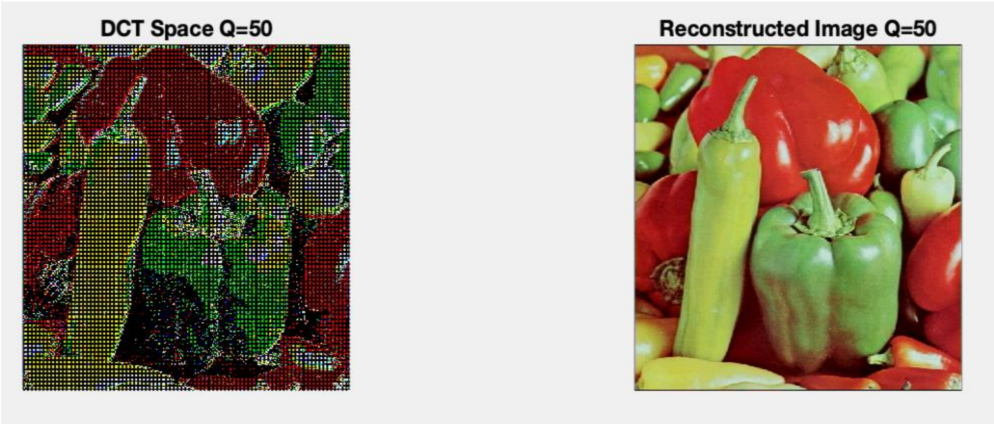
(b)



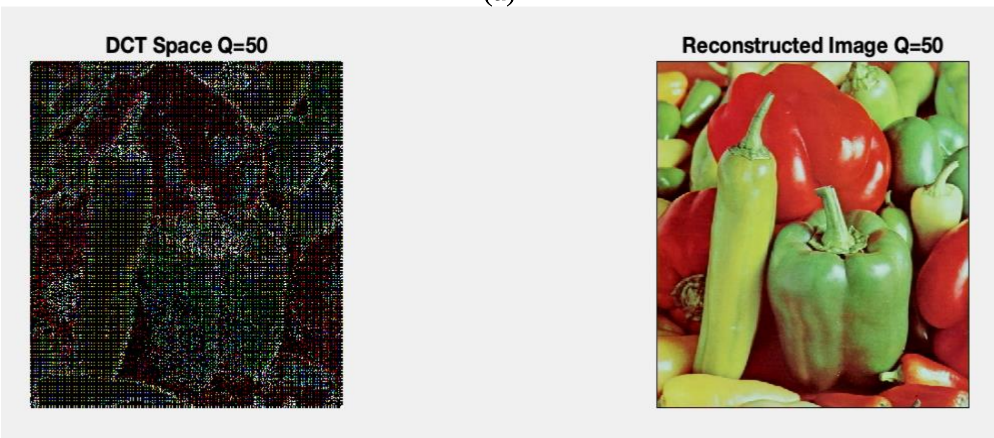
(c)

Figure 6: Quantized images in DCT space and Reconstructed images for Block size = 2 and **(a)** Quality level=10, **(b)** Quality level=50, **(c)** Quality level=90.

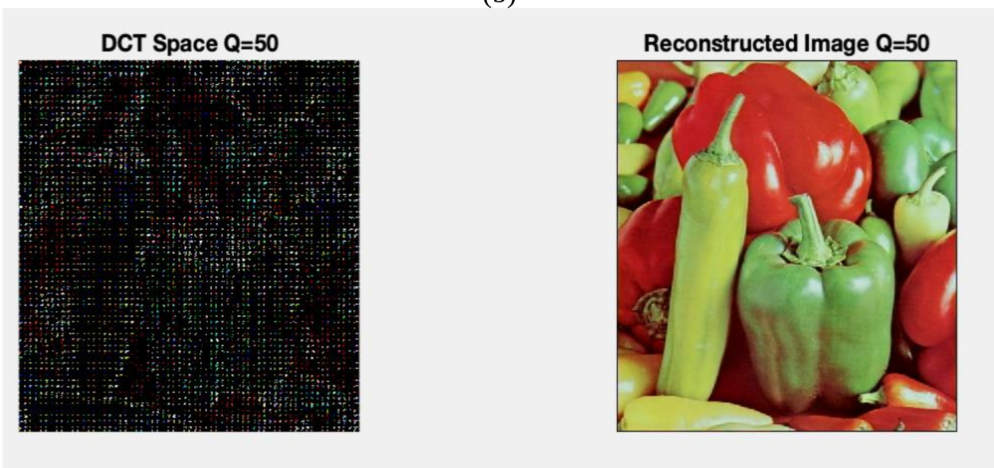
As seen in the figure, there is no visible difference between the original and reconstructed images. The amount of loss is visible on a quantized image on DCT space. The black dots in the picture show the 0s, that is, the data lost while quantizing. As can be seen from Figure 6, as the quality level decreases, the black spots in the pictures, i.e. the disappearing data, increase. This indicates more compression.



(a)



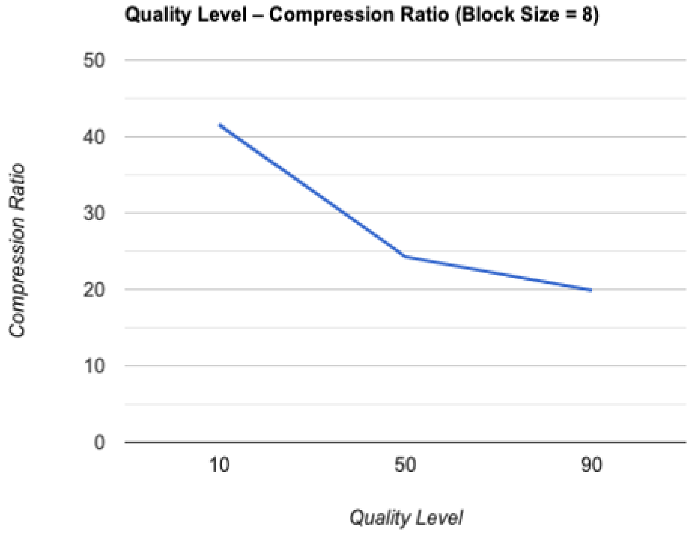
(b)



(c)

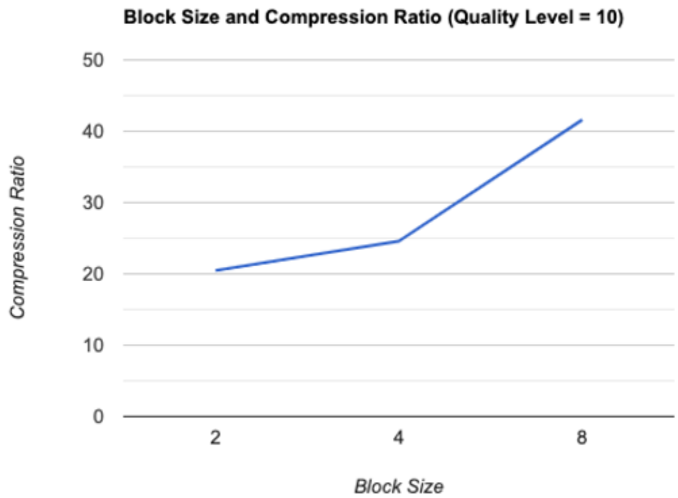
Figure 7: Quantized images in DCT space and Reconstructed images for Quality level = 50 and **(a)** Block size=2, **(b)** Block size=4, **(c)** Block size=8.

Then the quality level was fixed to see the effect of the block size. DCT space visuals have been added to figure 7, as it is difficult to observe the difference in the reconstructed images. The higher the block size, the quicker the image will be scanned and the earlier the process will be finished. As can be understood from here, as the block size increases, the compression ratio increases.



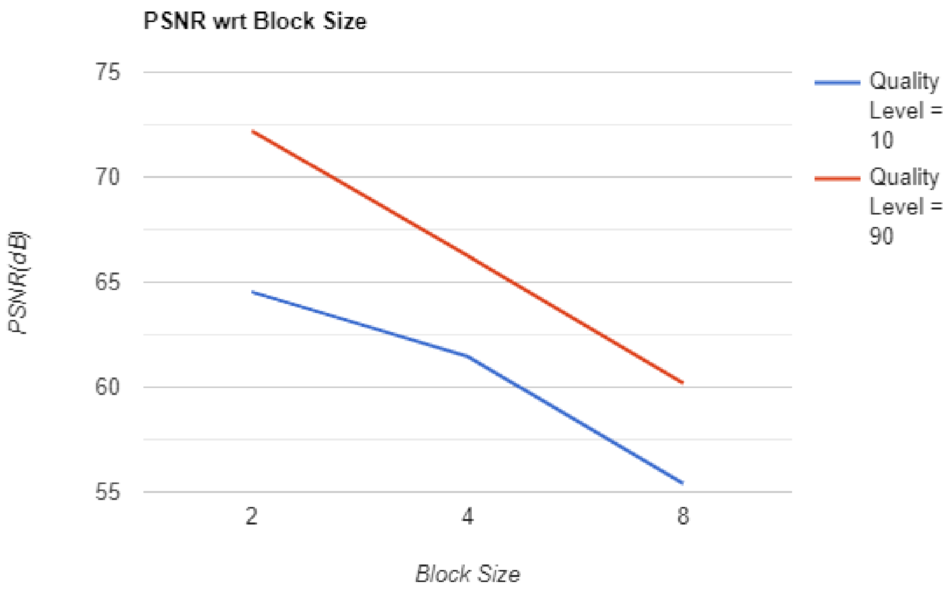
Graph 1: Quality level comparison with respect to the compression ratio with fixed block size.

Values are explained with graphics to be seen clearly. First, the effect of the quality level on the compression ratio was examined by keeping the block size value constant. As the quality level increases, the compression ratio should be low as the reconstructed picture will be clearer. Results in Graph 1 given above match the expected value.



Graph 2: Block size comparison with respect to the compression ratio with fixed quality level.

Next, the effect of block size on compression ratio was investigated. As block size increases, the compression ratio should increase in direct proportion. This situation is interpreted as the number of steps increases; the number of details decreases. This situation is clearly observed in Graphic 2.



Graph 3: PSNR comparison with respect to the compression ratio and block size.

PSNR values changes inversely comparing to MSE values. Block size positively affects the compression ratio, as seen in the previous graphics. Increasing block size also increases compression ratio. For this reason, increasing block size values should cause a decrease in PSNR values and Graph 3 confirms this statement. It is also showed with two different quality levels. As quality level increases, clearness of reconstructed image will increase, so the PSNR values for each block size must be highest at quality level 90.

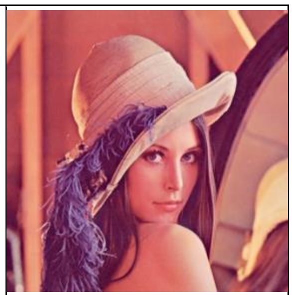
		 peppers			 lena
Quality Level	**Block Size**	**CR**	**PSNR**	**MSE**	**PSNR**
10	2	20.4758	64.5303	1489.81	63.8141
50	2	18.3937	71.7303	283.8768	70.1431
90	2	17.2259	72.2024	254.6374	70.5283
10	4	22.5889	61.4725	3012.39	60.7438
50	4	20.7115	65.8211	1106.74	64.2184
90	4	17.4176	66.2584	1000.73	64.5127
10	8	31.3211	55.4117	12161.7	55.1708
50	8	21.6877	58.4907	4754.37	58.0628
90	8	19.9749	60.1956	4042.04	58.4533

Table 2: General comparison

In the last stage, the effects of quality level and block size were examined simultaneously, and performance analyzes were made. The results are given in Table 2. As can be seen from the table, while the quality level increased, the PSNR values increased, MSE and CR values decreased. As expected, the highest compression ratio was observed at block size 8 and quality level at 10. Inversely, the lowest compression ratio was also observed at block size 2 and quality level 90. The test code was also repeated for a different picture. PSNR values were calculated for the Lena visual. A similar pattern with peppers appeared in PSNR values.

4) CONCLUSION

Image compression has become vital today. There are many different techniques, and these are divided into lossy and lossless compression techniques. The purpose of lossless compression is to compress the image while keeping the quality the same. However, in this technique, the compression ratio is very low. Since the aim of our project is to keep the compression rate high, we preferred the lossy compression technique. We used Discrete Cosine Transform to separate unnecessary information from the image. In this way, we tried to get pictures with a high compression ratio and lower PSNR value by removing unnecessary details in the picture. We used different parameters for this.

First of all, we focused on different quality levels. Quality levels range from 1 to 100. Value of “1” signifies the lowest quality, “100” represents the highest quality. We chose three quantization matrices with quality levels 10-50-90 to examine the details and concluded that the best quality image is with the “90” quantization matrix.

Then we tried to observe the effect of changing the block sizes on the picture. For this, we selected 3 different block sizes as 2, 4 and 8. We conclude that increasing the block size to certain point results in a higher compression ratio and a lower PSNR value. We found the highest compression ratio at the block size is 8 and quality level is 10.

RGB has some problems in itself. RGB is not the best match to represent visual perception. YUV stands out as better coding in this regard. In YUV capture, the bandwidth is lower compared to RGB. For this reason, YUV capture may be carried out in the future to further advance the project and increase the compression ratio [10].

At the end of these processes, we realized that the pictures contain too much unnecessary information. Thanks to the parameters used, we have served the purpose of storing and transmitting the pictures more efficiently.

REFERENCES

[1], [5], [6] Cabeen, Gent, K., n.d. *Image Compression and Discrete Cosine Transform*. 1st ed. [ebook] College of the Redwooks. Available at: <https://www.math.cuhk.edu.hk/~lmlui/dct.pdf> [Accessed 15 December 2020].

[2] Kaur, R. and Choudhary, P., 2016. *A Review of Image Compression Techniques*. 1st ed.

[3] En.wikipedia.org. 2020. *Lossy Compression*. [online] Available at: <https://en.wikipedia.org/wiki/Lossy_compression#:~:text=In%20information%20technology%2C%20lossy%20compression,%2C%20handling%2C%20and%20transmitting%20content.> [Accessed 15 December 2020].

[4] Malik, S. (2017). *Two-Dimestional DCT and Image Compression*. Http://Mason.Gmu.Edu/. http://mason.gmu.edu/~smalik17/Project%205/Project%205

[7] International Journal of Computer Science & Engineering Survey (IJCSES) Vol.5. (2015, April). *Jpeg Image Compression Using Discrete Cosine Transform - A Survey*. A.M.Raid1 , W.M.Khedr2 , M. A. El-dosuky1 and Wesam Ahmed1. https://arxiv.org/pdf/1405.6147.pdf

[8], [9] *How to Calculate PSNR (Peak Signal to Noise Ratio) in MATLAB?* (2014, February 5). Matlabclass. https://www.matlabclass.com/2014/02/how-to-calculate-psnr-peak-signal-to.html

[10] Sensoray.com. 2021. *Color Spaces In Frame Grabbers: RGB Vs. YUV*. [online] Available at:<http://www.sensoray.com/support/appnotes/frame_grabber_capture_modes.htm#:~:text=RGB%20formats%20are%20usually%20straightforward,0%2D255%20per%20color%20component.&text=YUV%20color%2Dspaces%20are%20a,more%20than%20RGB%20capture%20can> [Accessed 10 January 2021].