Assignment 4

Sangeeta Yadav (15318)

April 4, 2022

All the source codes are available here

1 Ques 1

In this question we implement a toy version of JPEG. The steps to implement are:-

- Read the input image and cast it into double.
- Transform: Compute an 8x8 discrete cosine transform (DCT) for every non-overlapping block in the input grey scale image.
- Quantization: Used the same quantization matrix as mentioned in the question.
- Lossless source coding: Use the same table as mentioned in the question to encode the quantized index corresponding to each DCT coefficient.

The input is 256x256 size 8-bit gray scale image. Therefore the size of the image is 524288 bits.

Lossless source coding:- For the lossless source coding of an input x, based on the examples given in the table, the following logic was used.

- The coded bit will be zero for $x = 0$
- If $2^{n-1} < +|x| <= 2^n - 1 \forall x \neq 0$: code = n times '1' then 0 then x converted into n bits code with $max(x) = n$ times '1' and min(x) = n times '0'

1.1 With quantization

For [a , b , c] = [10 , 40 , 20] and Quantizing the DCT coefficients using above mentioned methods, the results are shown below. CR(compression ratio) is given by bit-length of input uncompressed image divided by bit length of compressed image.

Table 1: Results

Parameter	Values
Input filesize(in bits)	521920
Output file size (in bits)	100918
MSE	42.42529296875
CR	5.1717235775580175

The reconstructed image with and without quantization are shown below:

Figure 1: Original gray scale image

Figure 2: Reconstructed image with quantization

1.2 Without quantization

Figure 3: Reconstructed image with quantization

Figure 4: Reconstructed image without quantization

Table 2: Compare the quality of image compression for with and without quantization

	With Quantization	Without quantization
Input filesize (in bits)	521920	521920
Output file size(in bits)	100918	353276
MSE	42.42529296875	0.302584201097488
CR	5.1717235775580175	1.4773717999524452

Table 3: Minimum MSE

Parameter	Value
MSE	32.8125839233398
$[a, b, c]$	$[30, 30, 50]$
Output size(in bits)	100394
Output size(in bytes)	12549.25

1.3 Optimal a, b, c

Figure 5: Reconstructed image with optimal value of [a,b,c] = [30,30,50]

2 Ques 2

2.1 Find the optimal quantization points and decision boundaries for a 2-bit scalar uniform quan- tizer

Part 1) Since the given pdf is symmetric around x=0, so we will consider a symmetric quantizer. We need three thresholds and 2 representation levels y_i for this. The $2-$ bit uniform quantizer divides the range of input data in 2 regions, the inner region $(-x_{max}, x_{max})$ and the outer $(-\infty, -x_{max}) \bigcup (x_{max}, \infty)$. The total distortion will be the sum of distortions in these 2 regions.

$$\begin{aligned} &f_X(x) = \frac{1}{6}exp(\frac{-|x|}{3}) \\ &D_{in} = 2(\int_0^d (x-y_1)^2 * f_X(x)dx + \int_d^{2d} (x-y_2)^2 * f_X(x)dx) \\ &D_{out} = 2(\int_{x_{max}}^{\infty} (x-y_{max})^2 * f_X(x)dx \\ &D_{total} = D_{in} + D_{out} \end{aligned} \tag{1}$$

For calculating D_{in} and D_{out}, I have used scipy.integrate for integrating.

$$\begin{aligned} &\text{Inner Distortion: } \left(-6.0de^{\frac{2d}{3}} + \left(0.25d^2 - 3.0d + 18.0\right)e^d - \left(0.25d^2 + 3.0d + 18.0\right)e^{\frac{d}{3}}\right)e^{-d} \\ &\text{Outer Distortion: } \left(0.25d^2 + 3.0d + 18.0\right)e^{-\frac{2d}{3}} \\ &\text{Total Distortion:} 0.25d^2 - 3.0d - 6.0de^{-\frac{d}{3}} + 18.0 \end{aligned} \tag{2}$$

I have differentiated the D_{total} w.r.t d, I got

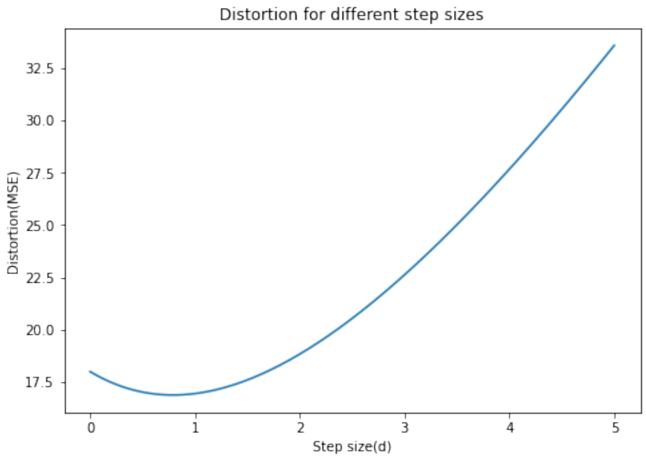
$$\left(2.0d + (0.5d - 3.0)e^{\frac{d}{3}} - 6.0\right)e^{-\frac{d}{3}} \tag{3}$$

Equating this to zero, we get :

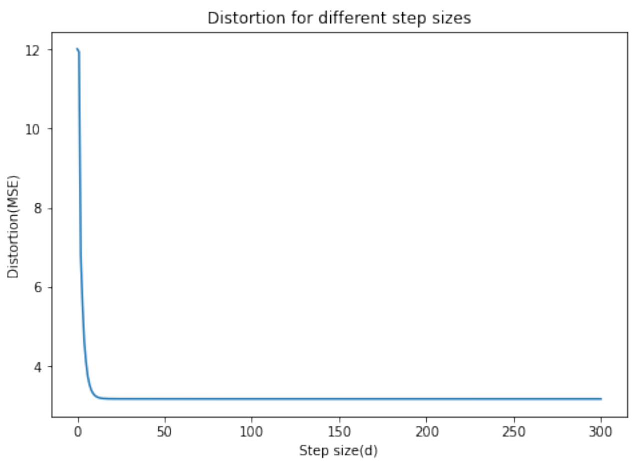
$$d(t+1) = 6 + \frac{(12 - 4d(t))}{exp(\frac{d(t)}{3})} \tag{4}$$

I have calculated the optimal value of d iteratively. $d(0) = \frac{ln(4)}{\sqrt{2}}$.
The optimal quantization pts are $y_i = (i - 0.5) * d$ for $i = 1, 2$.
The decision boundaries are $x_i = i * d$ for $i = 0, 1, 2$

- The optimal quantization pts are: [[-1.1858],[-0.3952],[0.3952], [1.1858]]
- The decision boundaries are: [[-1.5811],[-0.7905],[0.0], [0.7905], [1.5811]]
- Distortion achieved by the optimal quantizer 12.1399



2.2 Find the quantizaton points and decision boundaries for a 2-bit Lloyd-Max quantizer



The final quantization pts: [-7.7189, -1.7196, 0.0,1.8433, 7.8441]
The final decision boundaries: [-4.7189, -0.0622,0.0622, 4.8441]
Value of distortion at optimal step size d: 3.1715

2.3 Compare the distortion achieved by the uniform scalar quantizer and the Lloyd-Max quantizer and explain your observation

We can see that the distortion achieved by 2-bit Lloyd-Max quantizer is very less as compared to the scalar uniform quantizer.

3 Ques 3

Dataset is downloaded from here as given in the assignment.

Following metrics(as given in part A) on the dataset are calculated:

3.0.1 Mean squared error in pixel domain

MSE between 2 images is:

$$e_{mse} = \frac{1}{MN} \sum_{n=1}^{M} \sum_{m=1}^{N} [img(n,m) - ref_{img}(n,m)]^2 \tag{5}$$

3.0.2 Single scale structural similarity index(SSIM)

SSIM is used for measuring the similarity between two images. The SSIM index is a full reference metric; in other words, the measurement or prediction of image quality is based on an initial uncompressed or distortion-free image as reference.

3.0.3 Learned perceptual image patch similarity metric (LPIPS)

Higher means more different. Lower means more similar.

3.1 Spearman rank order correlation coefficient

Spearman's correlation measures the strength and direction of monotonic association between two variables. It takes values between 1 and -1.The sign of the Spearman correlation indicates the direction of association between X (the independent variable) and Y (the dependent variable). If Y tends to increase when X increases, the Spearman correlation coefficient is positive. If Y tends to decrease when X increases, the Spearman correlation coefficient is negative. A Spearman correlation of zero indicates that there is no tendency for Y to either increase or decrease when X increases. The Spearman correlation increases in magnitude as X and Y become closer to being perfect monotone functions of each other. When X and Y are perfectly monotonically related, the Spearman correlation coefficient becomes 1. A perfect monotone increasing relationship implies that for any two pairs of data values X_i , Y_i and X_j , Y_j , that $X_i X_j$ and $Y_i Y_j$ al- ways have the same sign. A perfect monotone decreasing relationship implies that these differences always have opposite signs. We calculate the Spearman rank order correlation coefficient between human opinion score ($blur_{dmos}$), MSE values and SSIM values for all images and tabulated below. Spearman rank order correlation coefficient between the dmos scores in "blur dmos" and each metric is as shown below:

Table 4: Spearman rank order correlation coefficient

Metrics	Correlation w.r.t $blur_{dmos}$	p-value
$blur_{dmos}$	0.9999999999999998	0.0
MSE	0.6607345299952763	1.530215622624923e-19
SSIM	-0.9213194772476775	1.5504169627935235e-60
LPIPS	0.8724137931034482	2.6518716703443e-46

3.1.1 Comment on the relative performances of all the indices

LPIPS is highly correlated to human opinion whereas SSIM is negatively correlated to human opinion. Thus LPIPS is a much better metric for checking image quality.