Advanced Image Processing: Assignment 4 (Due Apr 10, 2023)

Note: Please provide detailed comments for code that may be written to solve the following problems. The assignment will be evaluated not just based on the final results but also how you obtained them. Late submissions will be penalized.

Problem 1: JPEG Implementation (20 points)
Implement a toy version of JPEG through the following steps:

1. Transform: Compute an 8x8 discrete cosine transform (DCT) for every non-overlapping block in the input grey scale image.

2. Quantization: Use the following quantization matrix to quantize each DCT coefficient in a given 8x8 block

$$Q = \begin{bmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \\ 72 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \end{bmatrix}.$$

   Note that the quantized index of the DCT coefficient $x(i, j)$ is given by

$$y(i, j) = \left\lfloor \frac{x(i, j)}{Q(i, j)} + 0.5 \right\rfloor$$

   and the reconstruction is given by $\hat{x}(i, j) = y(i, j)Q(i, j)$.

3. Lossless source coding: Use the following table to encode the quantized index corresponding to each DCT coefficient.

Quantized DCT index	Code
0	0
-1,1	10x
-3,-2,2,3	110xx
-7,-6,-5,-4,4,5,6,7	1110xxx
...	...

The output bitstream (or file) is given by the concatenation of the sequence of bits produced for each 8x8 block.

Using the JPEG implementation described above:

1. Compute the size of the output file generated for the cameraman.tif image provided to you. Also compute the mean squared error between the original image and reconstructed image. The reconstructed image is obtained by taking the inverse DCT for each block of quantized reconstructions of DCT coefficients. Calculate the compression ratio (defined as the ratio of the input image in bits and size of the output file in bits).

2. Present a mechanism to modify the file size by modifying the quantization matrix Q given to you. Note that while you modify the quanitzation matrix, the relative quantization parameters for different DCT coefficients should remain the same. (Hint: You can scale the entire matrix by constants to achieve this). Plot a curve between mean squared error and the compressed file size (in MB) computed based on the approach provided in this question.

Problem 2: Bit allocation for uniform scalar quantization of independent sources (10 points):
Consider a pair of uniformly distributed continuous independent sources X_1 and X_2, both with mean zero and variances 5 and 10 respectively. Let the rate of the uniform scalar quantizers used for X_1 and X_2 be R_1 and R_2 respectively where $R_i = \log_2 K_i$ for $i \in \{1, 2\}$ and K_i denotes the number of quantization points. Given a budget sum rate constraint of $R_1 + R_2 = 3$, compute the rate pair (R_1, R_2) that minimizes the sum of squared error in the respective reconstructions $\mathbb{E}\left[(X_1 - \hat{X}_1)^2\right] + \mathbb{E}\left[(X_2 - \hat{X}_2)^2\right]$. Comment on how you would allocate bits among different DCT coefficients based on this result.

Problem 3: YOLO Object Detection (15 Points):

- Train a YOLO Object Detection model on this dataset.

- The dataset has 10 classes. The classes are defined in data.yaml file. Each image has a corresponding label file with bounding box annotations. Each row in the annotation file is of the form class-id center-x center-y width height corresponding to an object in the image. The values are in zero to one normalized coordinate space. Refer YOLOv5 annotation format if any doubts.

- Evaluate the performance of the trained model on the given test set. Compute the following metrics: (i) mAP@0.5IoU; (ii) mAP@[0.5:0.95]IoU (take average of mAP at 10 IoU thresholds from 0.5 to 0.95 with a step size of 0.05). You should also give my qualitative results in your report.

Prepare a report containing the answers to all the problems and submit along with relevant code.