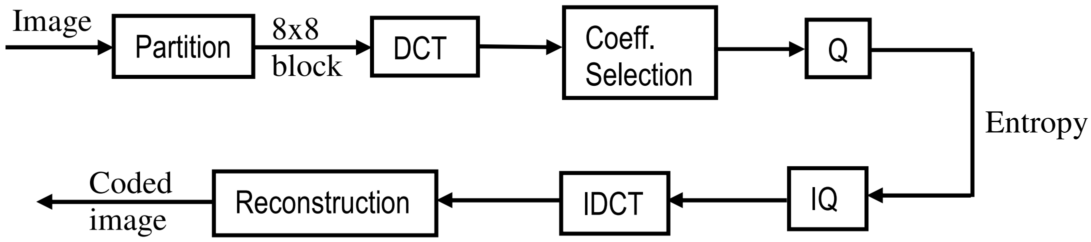
Video Signal Processing Homework #2 (Computer): DCT Coding

H.-M. Hang, October 2021

- You are asked to write your own programs to encode images using DCT (discrete cosine transform).
- You should be able to demonstrate your program on a workstation or a PC
- Please attach your program to your report.
- *Reminder:* The organization, clarity, etc. of your report contributes to 30% of the report score.

The overall procedure of encoding an image is shown in the following diagram.



At the encoder side, images are first partitioned into 8×8 blocks. The DCT operation is then applied to each image block. According to the international video standards (JPEG, H.263, ...), transform coefficients are rounded to the nearest integers and are clipped into 12-bit (-2048 to 2047). These coefficients are then arranged into 1-D data array according to the zigzag scan order. Then, some coefficients are selected and quantized. At the decoder side, the quantized coefficients are reconstructed by the inverse DCT (IDCT). The outputs of IDCT are rounded and clipped to 0 to 255. (We are dealing with Y signals only.) The specifications of DCT, quantizer, and zigzag scan order are given on the next page.

Solve each of the problems given below on 3 test images: lena (lena.raw), pepper (pepper.raw), and baboon (baboon.raw). The test images can be down-loaded from the Web. They are 512 by 512, 8-bit (per pixel), gray-level (black and white) images without any header. The image pixels are stored from left to right and from top to bottom into a file. (This is in raw image format, and an example of convert it to the BMP format is provided.) You need to compute the mean-square-error (MSE) between the original images and the reconstructed images.

(1) Energy compaction: What MSE will you get if you retain **one**, **3**, **5**, **10**, or **30** coefficients without quantization? Compute MSE of each picture in each case separately.

(2) Zonal coding: Use the given quantizer of step sizes **8**, **16**, and **32** to quantize the first 5 coefficients and the first 10 coefficients (6 cases in total). Compute the corresponding MSE and the total entropy of the 5 quantized coefficients. Repeat the same process for the first 10 coefficients. In the process of computing entropy, estimate the probability distribution of each coefficient independently, and then compute its entropy. The total picture block entropy equals to the sum of all the coefficient entropy. (Compute entropy of each picture separately for every step size.) Attach the reconstructed images in your report.

(3) Threshold coding: Quantize *all* the coefficients using the step sizes specified in **(2)**. Compute the average number of "nonzero" coefficients. Compute, also, the MSE using the procedure described in **(2)**. Attach the reconstructed images in your report. To count the number of nonzero coefficients, you start from the end of a block and count the number of zero coefficients (say, N_{zero}) until you hit the first nonzero coefficient. Then, the number of nonzero coefficient of that block is (N_{non} =64 − N_{zero}).

In doing research, it is important to interpret the data you obtained. At the end of your report, explain as much as you can the meaning of your data. For example, compare bits (entropy) needed for each picture, and explain why some pictures require more bits than the others. What are the advantages and the disadvantages of the methods used in **(2)** and **(3)**? You may want to look at the pictures for subjective comparison.

==================

DUE: November 18, 2021

==================

- **DCT:** 8×8

$$F(u,v)=\frac{1}{4}C(u)C(v)\sum_{m=0}^{7}\sum_{n=0}^{7}f(m,n)\cdot\cos\left(\frac{(2m+1)u\pi}{16}\right)\cos\left(\frac{(2n+1)v\pi}{16}\right)$$

$$f(m,n)=\frac{1}{4}\sum_{u=0}^{7}\sum_{v=0}^{7}C(u)C(v)F(u,v)\cdot\cos\left(\frac{(2m+1)u\pi}{16}\right)\cos\left(\frac{(2n+1)v\pi}{16}\right)$$

$$\text{where } C(u),C(v)=\begin{cases}1/\sqrt{2}, & \text{for } u,v=0;\\ 1, & \text{otherwise.}\end{cases}$$

- **Quantization:**

$$F_q(u,v)=\left\lfloor\frac{F(u,v)+\text{Sign}(F(u,v))\cdot\left\lfloor\frac{step\_size}{2}-2^{-12}\right\rfloor}{step\_size}\right\rfloor$$

Note: "2^{-12}" is included in the quantization formula to reduce machine precision problem.

Reconstruction:

$$F_{iq}=F_q(u,v)\cdot step\_size$$

- **Zigzag Scan Order:**

0	1	5	6	14	15	27	28
2	4	7	13	16	26	29	42
3	8	12	17	25	30	41	43
9	11	18	24	31	40	44	53
10	19	23	32	39	45	52	54
20	22	33	38	46	51	55	60
21	34	37	47	50	56	59	61
35	36	48	49	57	58	62	63