EE475 Compression
Fall 2019

Abdülkadir Gökce
Sefa Kayraklık

SUB-SAMPLING	PCM	DPCM	VQ	DCT	JPEG	Total
						/200

1. **Try to give concise one or two sentence answers, explanations**
2. **Copy and past images from the VC Demo program to enhance your answers**

IMAGE COMPRESSION SOFTWARE

In the following, study the image processing function experimenting with various parameters, report your experience briefly in a few sentences. Copy and paste the images and graphs to accompany your commentaries.

SUBSAMPLING

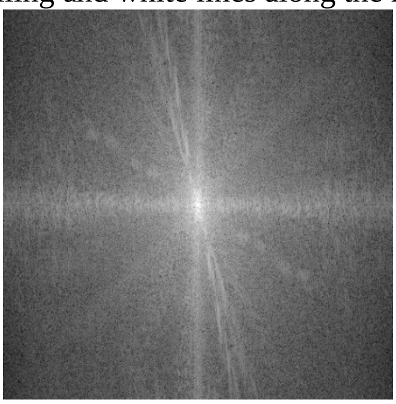
- Inspect the spectrum of the image Build512B.bmp. The DC component is in the center. Give explanations for the observed line structures. Note: *SS Button → Spectrum → Apply*

a) Why there are strong horizontal and vertical spectral lines passing through the origin?

As we know, FFT algorithms takes Discrete Fourier Transform of the image which regards the image of interest as replicated inifinitely along horizantal and vertical directions. Hence, this creates discontinuties along edges and cause high rate of change, therefore we observe strong response along axis.

b) Plot the image and its spectrum side by side, and identify two main image structures corresponding to two strong lines in the spectrum.

The image and its spectrum are shown in the following figures. From the spectrum we see that there are 2 strong lines tilted toward positive direction from *y-axis* are the results of high rate of changes induced from white side of the ceiling and white lines along the mini pool.



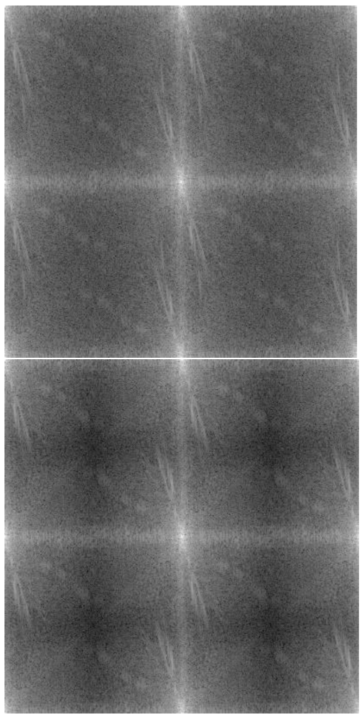
c) Which details in the image would cause the highest frequency components in the spectrum?

The highest repetition in the image was the iron railing along the ceiling and the staircase.

- Subsample by a factor of 2, once without anti-aliasing filter and then with 17-tap aliasing filter. Note: *SS → Factor → 2 , check Apply Subsampling box, check Blow-up Subsampled Image box*
SS → Filter → Check or uncheck Apply anti-alias filter, Set the number of taps to 17.

  The highest repetition in the image was the iron railing along the ceiling and the staircase.

d) Plot the images and their spectra side by side: Can you pinpoint the aliasing events in the subsampled image without anti-aliasing filtering? Where would you expect these distortions to occur?
The subsampled images with and without anti-aliasing and their spectrum are shown in the following figures.



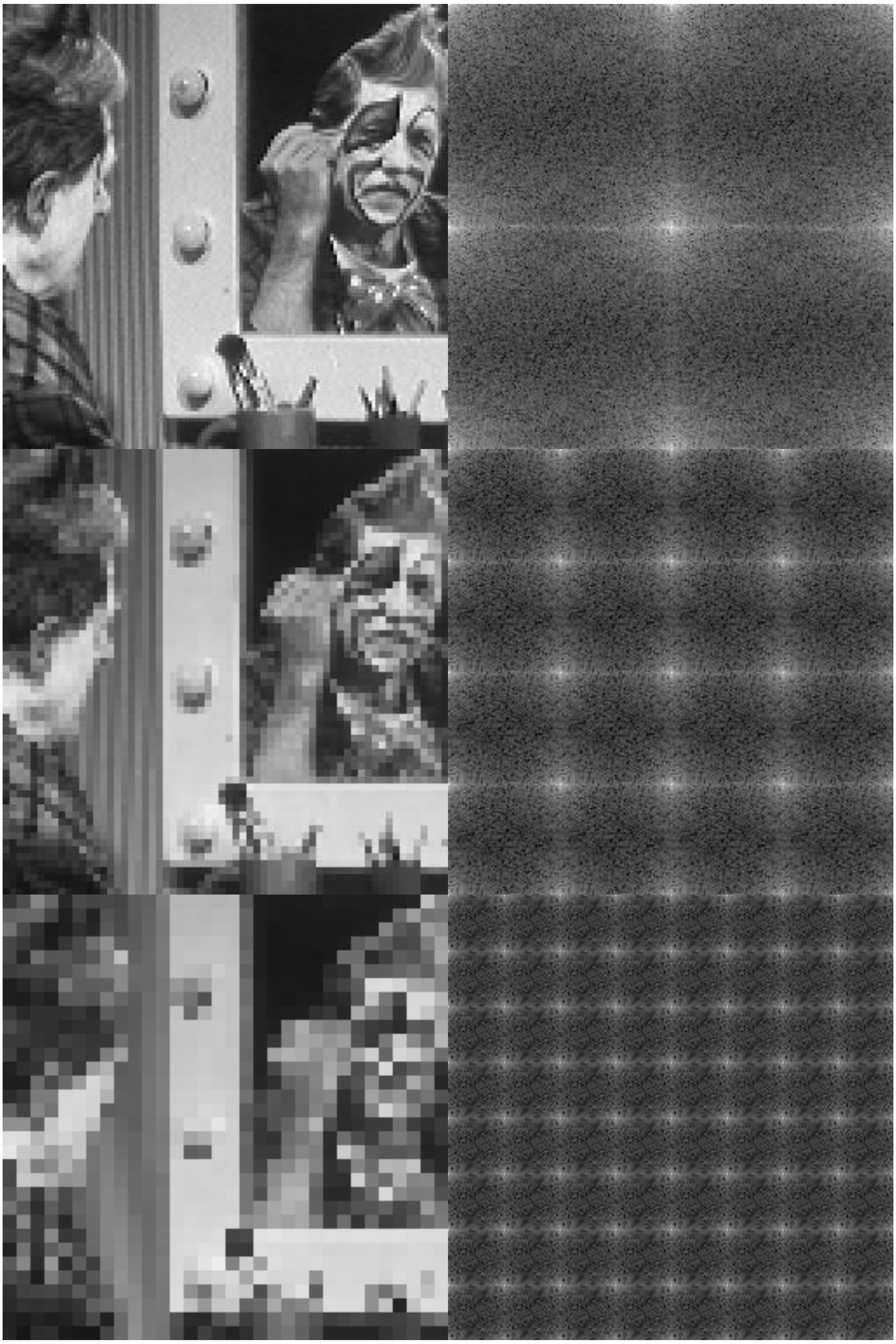
When we subsample the image, high frequency components overlap causing distortion to image at hand. Therefore, we expect impairment around the regions with high rate of change.

e) Does anti-aliasing filter improve it? Explain the improvements in the image and the changes in the spectrum.
It can be seen from the spectrum of the image without anti-aliasing that there exists aliasing in the high frequencies. For example, artifacts can be observed region around the staircase and along iron railing. After anti-aliasing filter, those part are improved. From spectrum, we saw that there are blackish lines there are the result of anti-aliasing filtering which prevented overlapping.

f) In turn, does the anti-aliasing filter cause any distortions? Explain the difference in the spectra of the subsampled images with and without anti-aliasing filters.

Anti-aliasing does also filter out information about the image. In particular, high frequency components are attenuated and blurred in the image. However, this is an acceptable level of data loss that prevents higher damage.

g) Subsample the Clown image by factors of 2, 4 and 8, and plot the spectra, side-by-side.
We can see that as we further subsample the resolution and the quality of the image decrease as expected.



PCM

- Study the images Lena256B, Clown256B, Odie256B. Note: *PCM button → Select bitrate → Apply. Do for bit rates 7 to 1.*

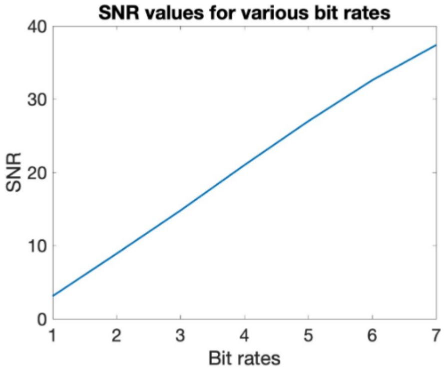
a) At which rate do the coding artifacts become objectionable? Where do you observe these artifacts most clearly? At this rate how many gray values are available to represent the image?

	Bit-rate at which artifacts become objectionable	Gray values at this rate
Lena256B	4	16
Clown256B	4	16
Odie256B	1	2

For Lena256B, Fill in the table.

Bit rate	SNR (dB)
1	3.1
2	8.9
3	14.8
4	21.0
5	27.0
6	32.6
7	37.8

b) Draw an SNR-versus-bit-rate plot for Lena256B. Explain the reason why you are getting such a slope for this curve.



As we decrease the number of available bits, we reduce the available gray levels and decrease the contrast, so the representation power is weakened. Hence false contouring and information loss occurs. Therefore, SNR increases with the number of pixel available

c) Explain the behavior of Odie under PCM, which is quite different than that of Lena. What distortions occur in Odie? What distortions occur in Lena?

Odie has 4 distinct colors, whereas Lena has more than 16 colors. Using 4 bit for Lena we saw false contouring on her shoulders and background. This image has make use of higher number of gray levels and more contrast between and within regions. On the contrary, Odie has only 4 distinct colors to represent and less color variations, hence we are good with less number of gray levels. Odie's shadow and tongue are shown as background in the pcm with 1 bit-rate.

DPCM

- Select the Lena256B image and the 1-D predictor. Note: *DPCM button → Model → Select the first prediction model. Select Bit Rate from 1 to 6.*

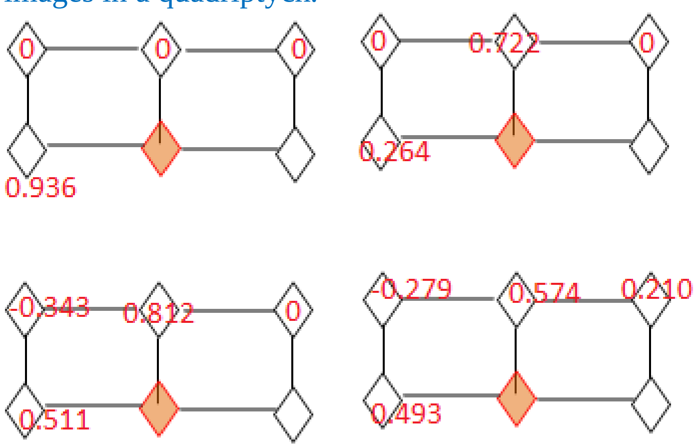
a) Carry out compression at bit rates 6 to 1 bpp and obtain the SNR – Observe the gain over PCM.

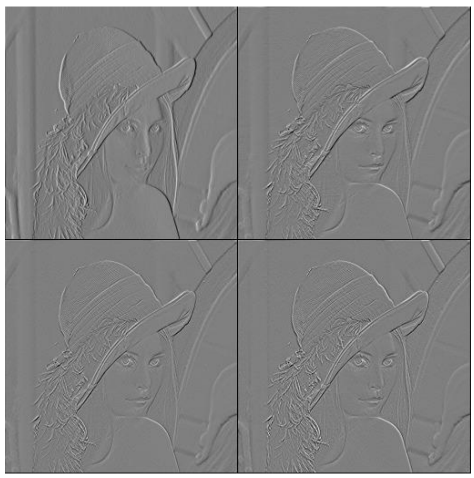
For Lena256B, Fill in the table.

Bit rate	DPCM SNR (dB)	PCM SNR (dB) Copy from previous table
1	7.4	3.1
2	13.2	8.9
3	18.5	14.8
4	24.3	21.0
5	31.2	27.0
6	36.5	32.6
7		37.8

b) Compare visually the images with their PCM version and state at what rate DPCM achieves a performance equal to that of PCM?
DPCM performs better that PCM with the same number of pixels. Not only it has higher SNR values with the same bit rates, its visual performance was even more with lower SNR value. For example, image with 3 bits is acceptable with some level of distortion whereas normal PCM was not.

c) Use four types of prediction region for DPCM at bit rate level 5 and comment on the performance differences. Sketch the prediction context (the orange diamond is the current pixel) and write the coefficient values. Plot the Lena image and its three prediction error images in a quadriptych.

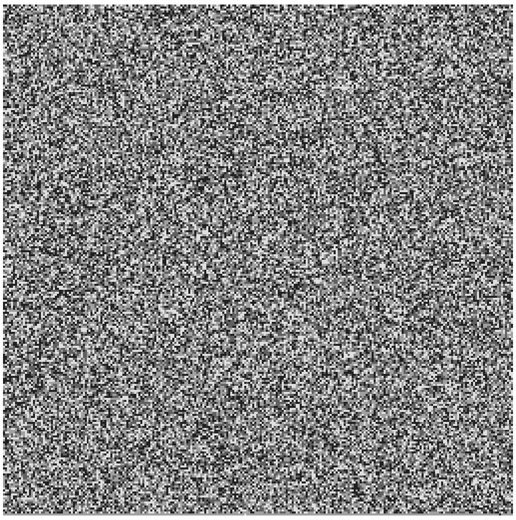




d) Observe the correlation matrix. What is the theoretical gain over PCM in terms of bit rate reduction at the same quality? Does the experimental result match the theory? Recall the prediction gain for one-tap predictor: $\frac{1}{1-r^2(1)}$ where r(1) is the normalized correlation coefficient.

The theoretical gain over PCM is 8.07 dB, where r(1)=0.936. So, it can be seen from the table of SNR in terms of bit rates that there exists approximately 4 dB gain over PCM in the DPCM. The experimental gain is half of the theoretical gain. Since image at hand is not fully correlated (i.e. hairs and eyes), it is reasonable to obtain less SNR improvement.

e) If DPCM is used on the noise image below, what do you expect the coding gain to be? What would be the prediction coefficients?



DPCM make use of local correlation in patches and only encodes difference levels. In this image, we saw uncorrelated components over the whole image and high rate of chance. Therefore, we would not obtain much of a gain using DPCM. Hence, the prediction coefficient would be very low, close to 0.

VQ

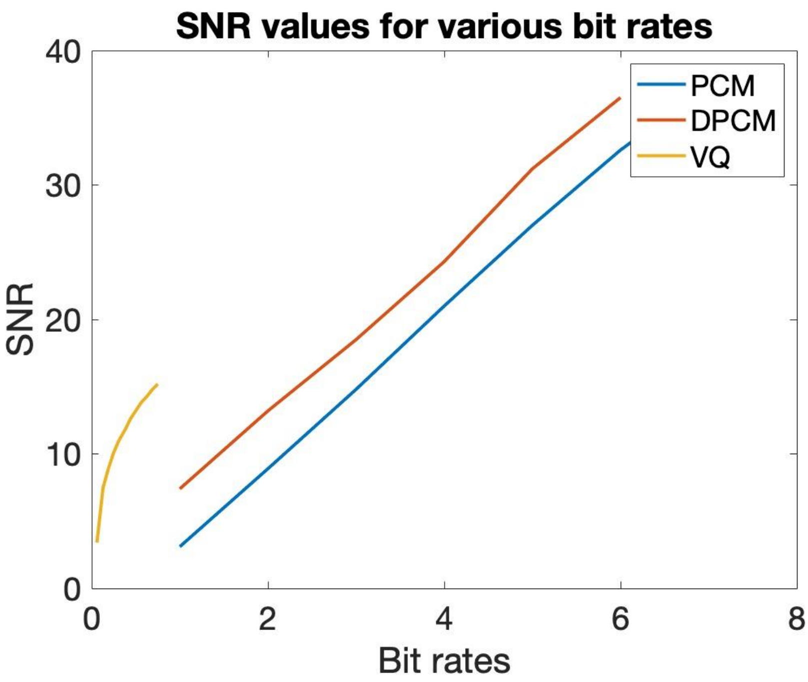
Load the predesigned codebook Standard_4x4_min1-max12.cbk. This codebook has been designed on 4 other images other than Lena.

Note: *VQ button* → *Codebook* → *Load codebook* → *Select Standard_4x4_min1-max12.cbk* → *Open*

Bits → *Select*

a) Make an SNR-bit rate plot of VQ on the same curve as PCM/DPCM, i.e., copy first the PCM-DPCM plot. Note: Be careful to first calculate the VQ bit rate (bits/pixel) before you plot.

The plot of the SNR versus bit rate is shown in the following figure.



b) What type distortion do you observe at low bit rates? Does it appear to be additive white Gaussian noise? Plot the VQ coded images.
The following images are the version of Lane with 1, 3, 5, 7, 9, and 11 bits per 4*4 vector respectively.

The following image is the version of Lena that is added white gaussian noise with 10 dB SNR.

By comparing it with vector quantized images, we observe that the noise is not quite similar. Quantized images have artifacts more similar to images with low resolution problems. Hence, we can say it is different than additive white Gaussian noise

c) Can you use codebooks obtained by Lena-type images on cartoons and maps? If yes, how? If not, why not?

We could try coding with it but that would be not efficient since cartoons and maps has homogenous, non-changing regions whereas Lena has varying features. Therefore, we would have low quality encodings.

DCT

Select the Lena256B image. Select PCM compression for all DCT coefficients. Pick c = 0.75 as the exponent power of the generalized Gaussian distribution.

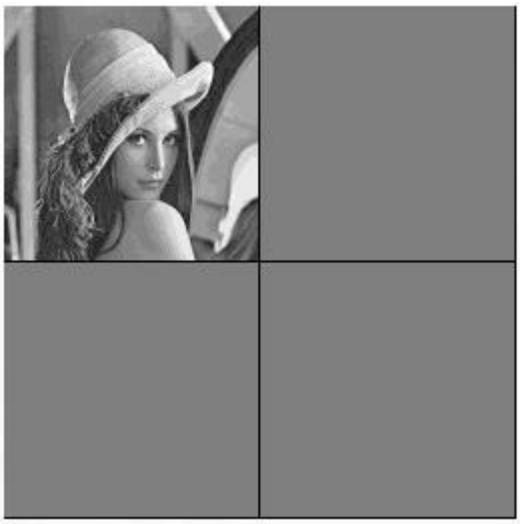
Note: *DCT button → Size tab → Check encoding of DCT coefficients → Set transform size as 2x2, 4x4 and 8x8 in turn and press apply*

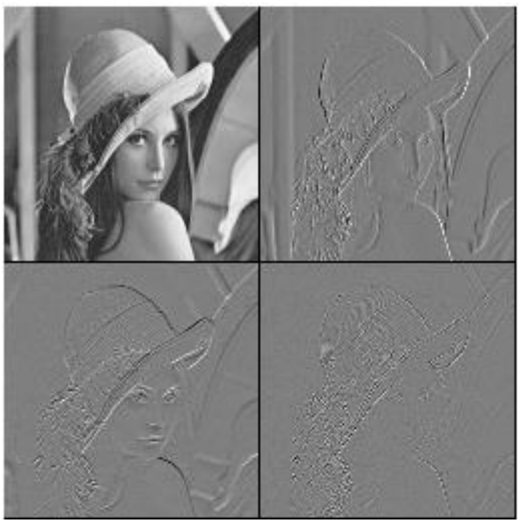
Coefs tab → Select PCM for First and Others, Select C value as 0.75

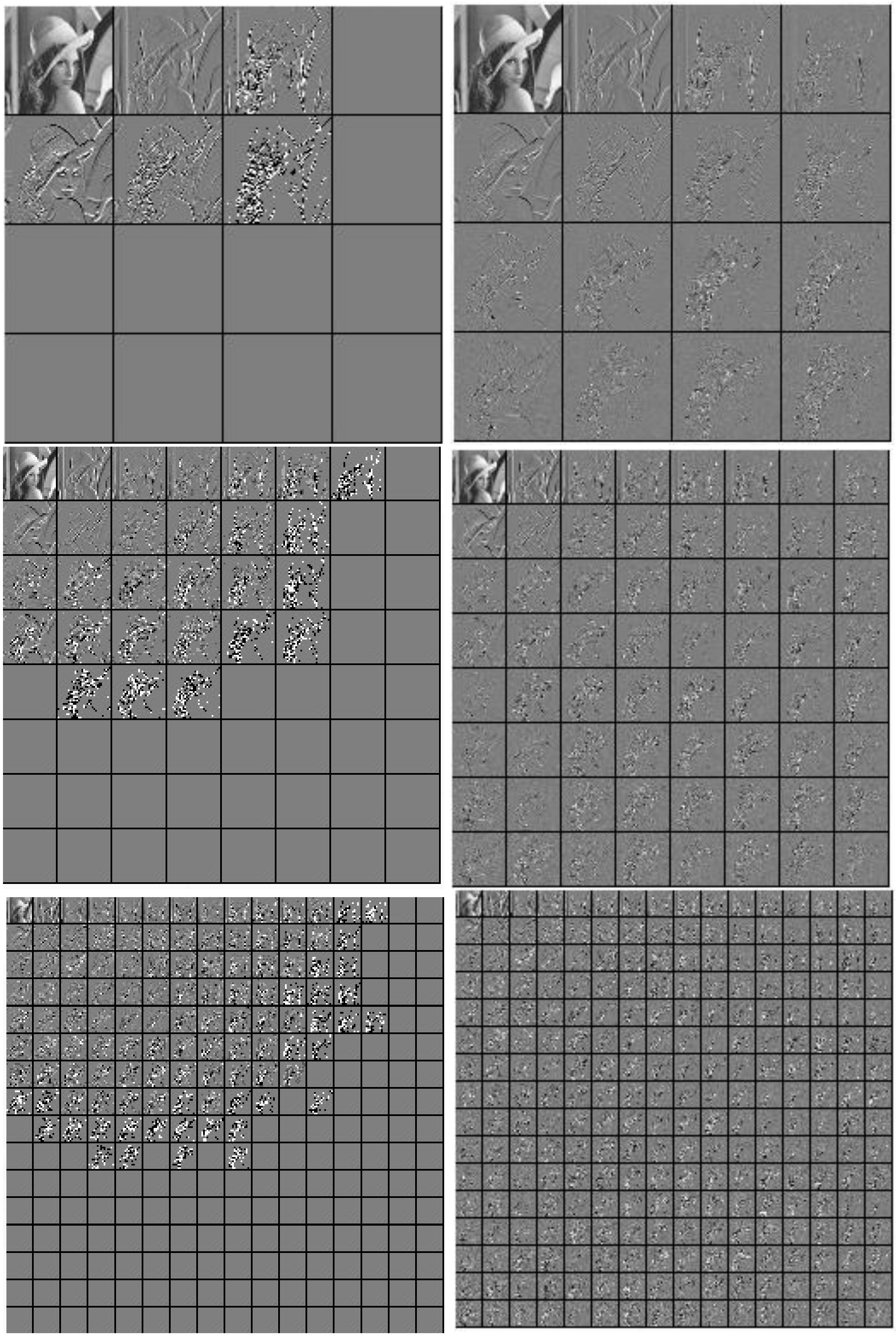
Bitrate → Check the entropy coding. Select each of the bit rates in turn.

a) Study the block sizes of 2x2, 4x4 and 8x8, 16x16. Explain what you observe in the "DCT: Original Coefficients" window. Plot both the un-coded and coded versions of the DCT coefficients, side by side for the four block size choices, as in the example below.

The DCT original coefficients window shows the corresponding coefficients of the discrete cosine transform basis functions. The following images show the un-coded and coded versions, respectively.







b) Use again entropy coding, but with the wrong signal model. Assume wrongly that the shape parameter of the generalized Gaussian distribution is a = 2 (hence pure Gaussian), and not a = 0.75. What do you observe? How is coding performance affected?

As we increased from 0.75 to 2, additional distortion is introduced to image. For example, we see more abrupt ridges present on Lena's shoulder and background noise on the wall. Moreover, coded DCT coefficient seems noisier and more number of non zero coefficient is required. Hence, it performs poorly and need more bits.

JPEG

- Use the Lena256B image and the standard luminance normalization matrix. Note: *JPEG button → Huffman tab → Select FLC, Standard VLC, Optimal VLC in turn*
  JPEG button → Bit rate tab → Bitrate → Select bit rate as 0.5, 0.8, 1.0, 1.5, 2.0, 2.5, 3.0 in turn → Apply
  Write down the Encoded Bit Rate in bpp, Optimized Quality Factor and the corresponding SNR. (FLC: Fixed Length Coding, VLC: Variable Length Coding)

a) Fill the SNR-bit rate tables, one for each entropy coding choice. At the same time, write down the corresponding QF: Quality Factors. How much additional SNR does entropy-coding give?

Bit rate	FLC			Standard VLC		
	Encoded Bit Rate	Quality factor	SNR (dB)	Encoded Bit Rate	Quality factor	SNR
0.5	0.7	2	7.4	0.5	17	15.7
0.8	0.8	4	10.4	0.8	38	18.2
1.0	1	6	12	1.0	55	19.6
1.5	1.5	13	14.9	1.5	78	22.4
2.0	2	23	16.6	2.0	87	24.8
2.5	2.5	34	17.8	2.4	91	26.7
3.0	3	48	19.0	3.0	94	28.8

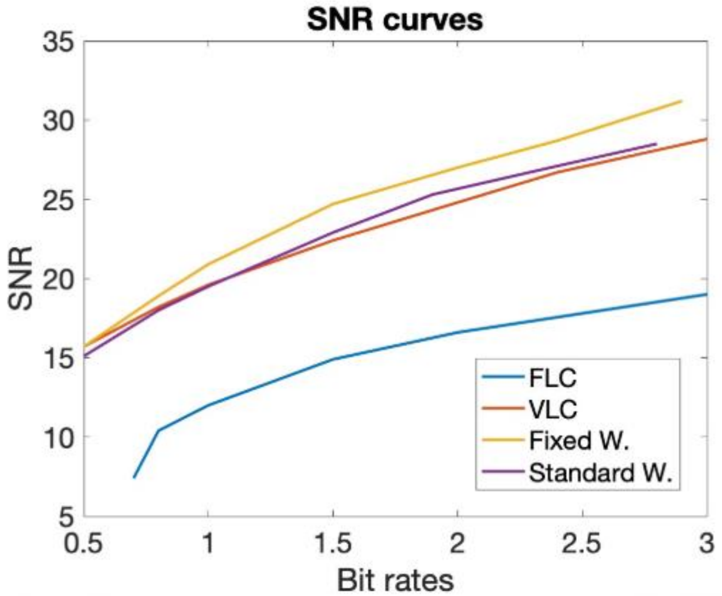
The SNR in the variable length coding increases by almost 10 dB, and the corresponding quality factor is nearly doubled for higher bit rates and approximately ten times higher for lower ones. So, with the VLC, we can encode JPEG images with much less bit rates and achieve superior performance.

b) Compare the resulting SNRs from a “flat normalization matrix” against that of the “standard luminance normalization” and comment.
Note: *Repeat above selecting: Quant → Choose Flat*

Bit rate	Fixed weighting			Standard weighting (psychovisual)		
	Encoded Bit Rate (b	Quality factor	SNR (dB)	Encoded Bit Rate	Quality factor	SNR
0.5	0.5	41	15.7	0.5	38	15.1
0.8	0.8	66	18.9	0.8	70	18
1.0	1	76	20.9	1	79	19.5
1.5	1.5	87	24.7	1.5	89	22.9
2.0	2.0	91	27	1.9	93	25.3
2.5	2.4	93	28.7	2.4	95	27.1
3.0	2.9	95	31.2	2.8	96	28.5

Although fixed weighting produces slightly higher SNR values then standard luminance normalization slightly lower quality factor which indicate they are comparable in terms of quantitative performances. However, the latter generates higher quality images in terms of human perception, hence preferable.

The four bit rate-SNR curves are drawn on the same plot, and shown in the following figure.



c) Observe the images with flat versus standard normalization matrix at 0.5 bpp. Their SNRs are equal, yet one looks better than the other one. Explain this dilemma.

The standard weighting matrix is tailored to human visual system so that we can discard information that our eyes cannot perceive. This matrix is adjusted with human psychovisual experiments that to imitate its bandpass response. With that property, we can select DCT coefficients that matter most and code them more elaborately while assigning less bits to others. Hence, we like more the one encoded with standard weighting scheme.