ECE 634 Digital Video Systems

Project 2: Vector Quantization and Discrete Cosine Transform

By

Amit Kumar Singh Yadav
Email ID: yadav48@purdue.edu

Table of Contents

Part 1: Vector Quantization

1.1 Theory

1.1.1 Encoder

- first pre-process all the training images and image to compress
    - convert them to grayscale
    - Resize them to 512x512
- Now, extract vectors from all the training image i.e., 4x4 (16-dimensional vector)
- Using these 16-dimensional training data create codebook
    - Start by creating a codebook of size = 1 (mean of all the training data)
    - Then keep extending the code book using the training data and LBG (k-means)
- Now, save the codebook, and extract vectors from image to be coded
- Find index of nearest codes in codebook for each vector block from image to be encoded
- Saved the index codes for each vector block and send it with codebook for decoding

1.1.2 Decoder

- Given an vector quantized encoded image, for each codebook index in image replace it with a 4x4 block using the codebook
- This is the Decoded Image

1.1.3 Performance Metric

- We use PSNR of reconstructed image w.r.t the original image as a performance metrics for evaluating the quality of the reconstructed image. We assumed that a PSNR of 25dB or above is satisfactory reconstruction.

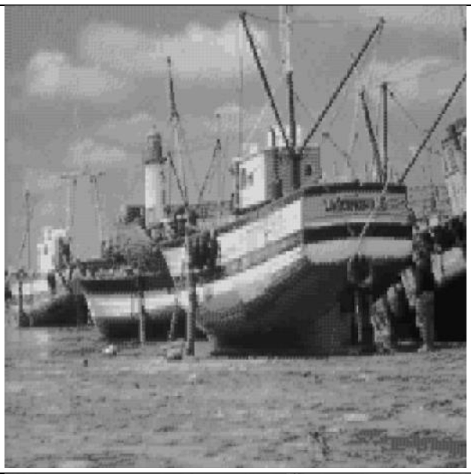
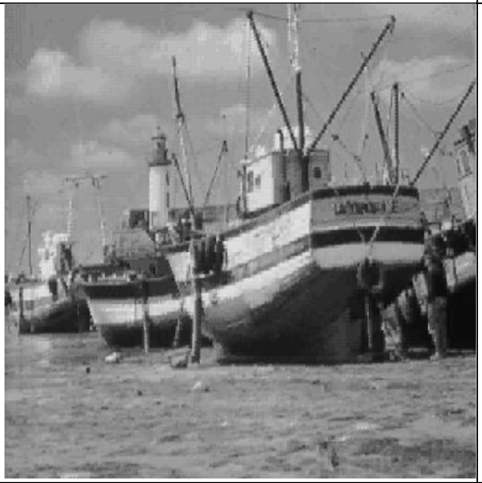
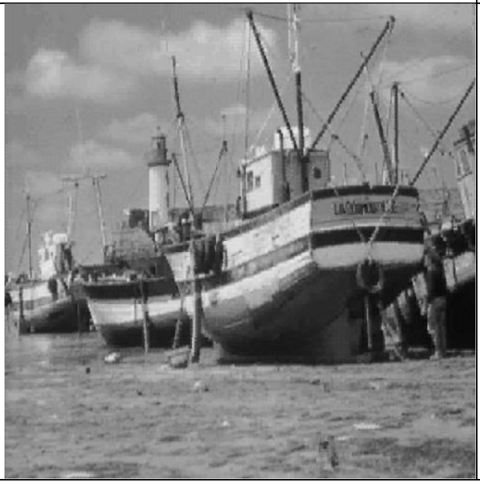
1.2 Results when using all the blocks in an image to make codebook

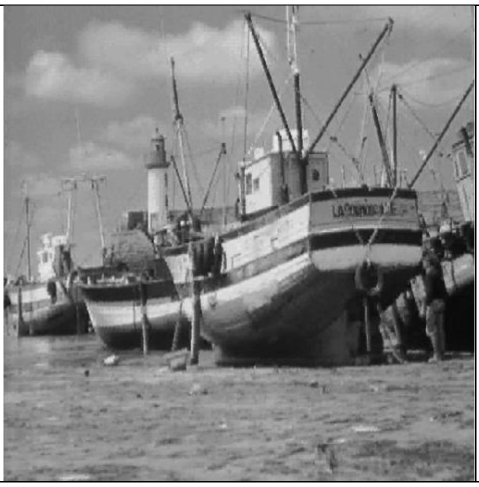
In this case we used the same image for creating the codebook

1.2.1 Experimenting impact of different codebook sizes

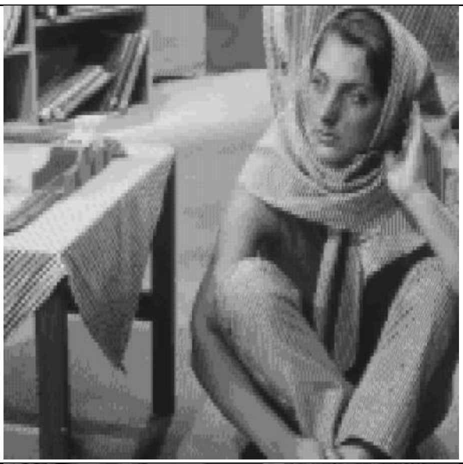
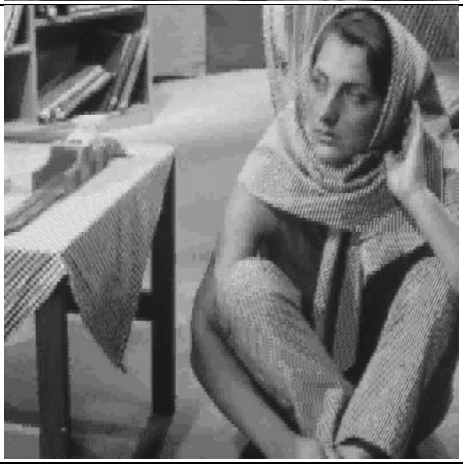
1.2.1.1 Boat Image

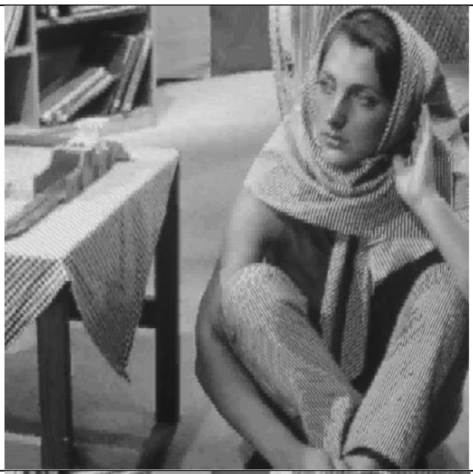
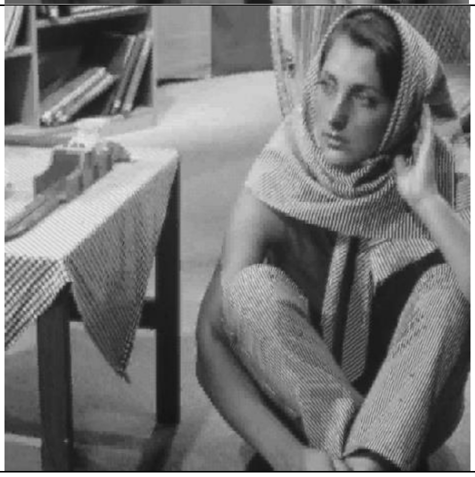
Codebook Size	**Original Image**	**VQ reconstructed Image**	**PSNR (dB)**

32			25.69
64			26.90
128			27.78

256			28.85

1.2.1.2 On Barbara Image

Codebook Size	Original Image	VQ reconstructed Image	PSNR (dB)
32			23.57
64			24.83

128			25.96
256			27.10

1.2.2 Discussion – impact of different codebook sizes using codebook trained with an image

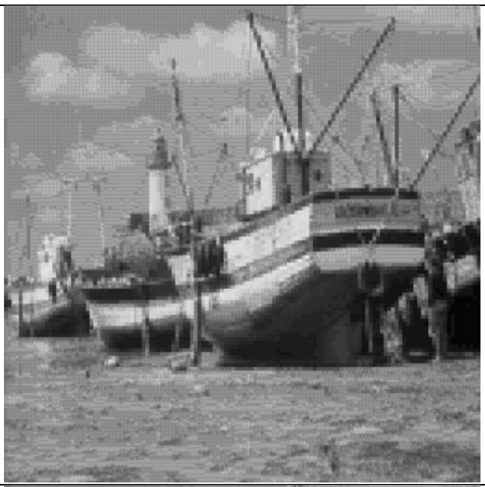
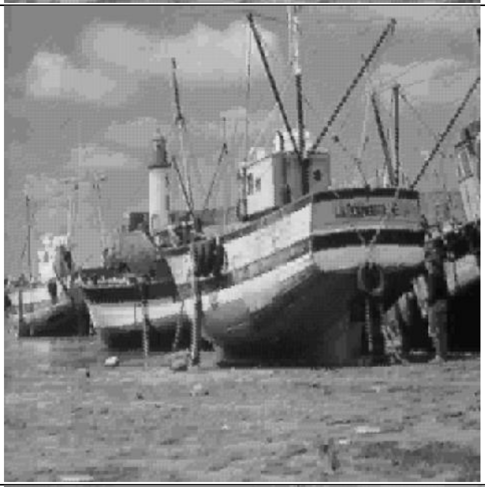
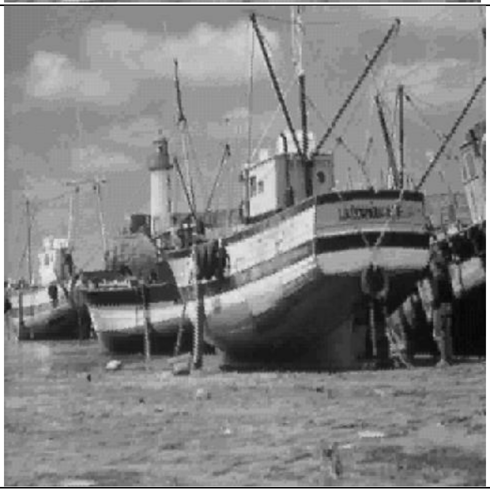
- We can see that as we increase the codebook size the PSNR increase and the reconstructed image quality improves
- Observe that the absolute value of PSNR is significantly dependent on the image content for fixed size of codebook

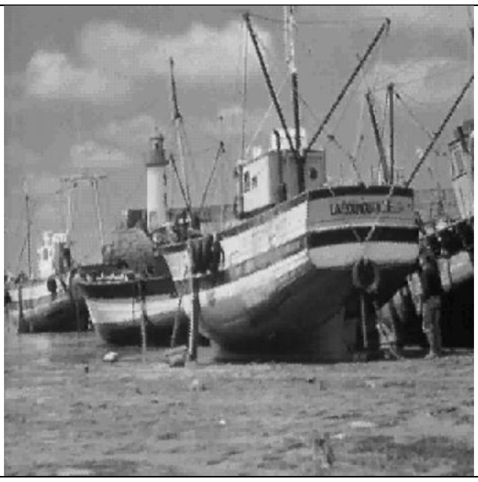
1.3 Results when trained Codebook using 10 images

1.3.1 Experimenting impact of different codebook sizes

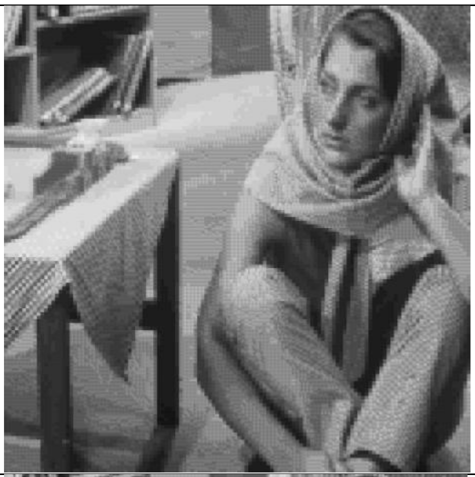
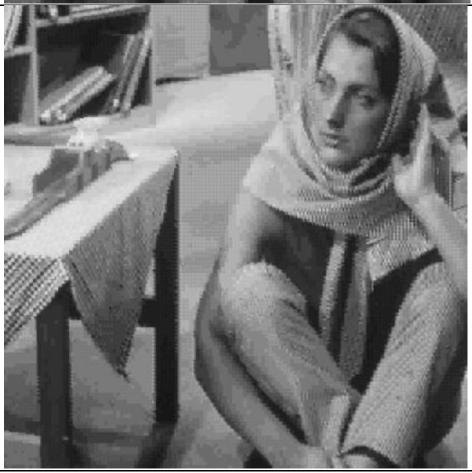
1.3.1.1 Boat Image

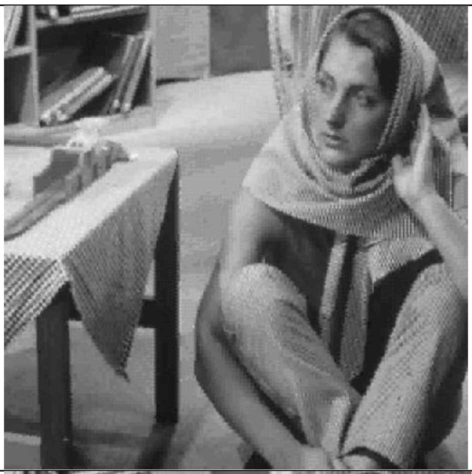
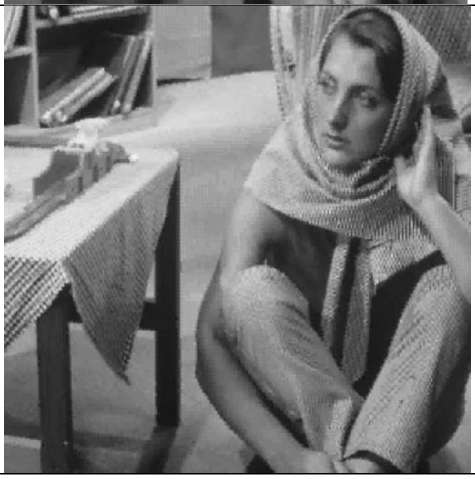
Code-book Size	**Original Image**	**VQ reconstructed Image**	**PSNR (dB)**

32			24.58
64			26.18
128			26.94

256			27.59

1.3.1.1 On Barbara Image

Code-book Size	Original Image	VQ reconstructed Image	PSNR (dB)
32			22.98
64			23.69

128			24.01
256			24.37

1.3.2 Discussion – impact of different codebook sizes for codebook trained with 10 image

- In this case too we can observe that the PSNR of reconstructed image w.r.t original image increase by increasing codebook size
- The PSNR is less w.r.t codebook of same size but created using same single image that is being encoded. This is expected. However, the reduction in performance is not very significant given the save in complexity. As now, the same codebook can be used for multiple images.
- The improvement in PSNR by increasing codebook size is not very significant as in case of codebook created from the same image.

1.4 Comparison of results using single image v/s 10 images to generate codebook

Code-book Size	**Single Image trained**	**10 Image trained**	**PSNR in dB 1 img PSNR/ 10 img PSNR**

32			**25.69** /24.58
64			**26.90** /26.18
128			**27.78** /26.94

256			**28.85** /27.59

1.4.1 Observation from comparison

- Codebook created from the same image unanimously for all the block size gave better results w.r.t codebook created from 10 images different than image being encoded.
- The change in PSNR is approximately 1dB by creating codebook for each image being encoded. Notice, it is not very significant given that the codebook prepared from 10 training image can work for any similar image.
- Creating codebook from 10 training image is saving us from computationally complexity to generate codebook for each image.
- In case of codebook created from 10 training images the blocking effect is comparatively more evident (see the cloud regions).

Part 2: Discrete Cosine Transform

2.1 Theory

2.1.1 Encoder

- We perform DCT using block of size 8x8, so first we generate 8x8 DCT coefficients
- Now, we take non-overlapping 8x8 blocks from image and perform DCT on it
- Once we have DCT transformed image, for each 8x8 block in image, we keep the first K DCT coefficients (we get it using zigzag order, refer code for more details) and make rest of them equals 0. Notice for 8x8 DCT, K at max can be equal to 64
- We experimented K = 2,4,8,16,32
- Now, we have an optional step for quantization, we take the quantization table from the lecture note, to quantize these coefficients

2.1.2 Decoder

- We dequantize the coefficients,
- perform inverse DCT on the dequantized image in block-wise manner
- reconstruct the original image

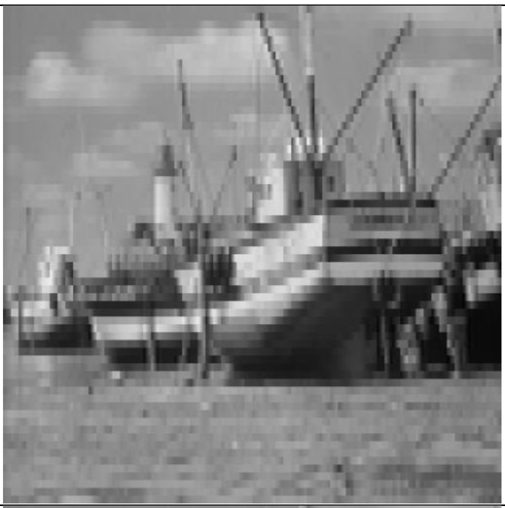
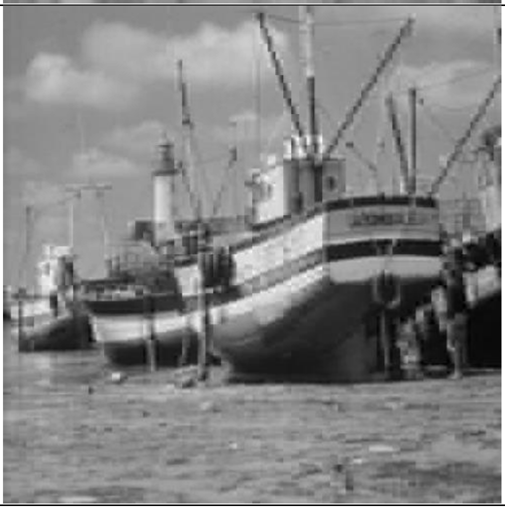
2.1.3 Performance Metric

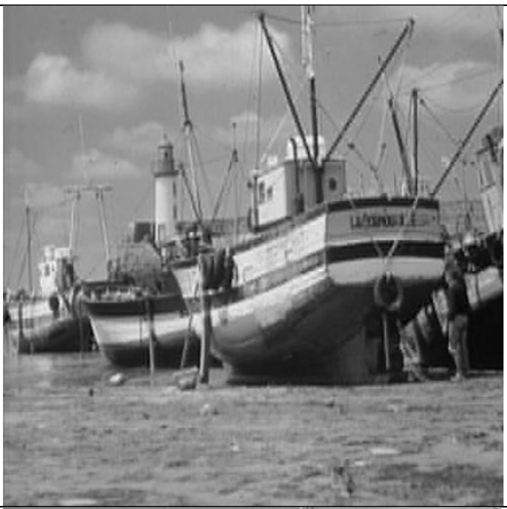
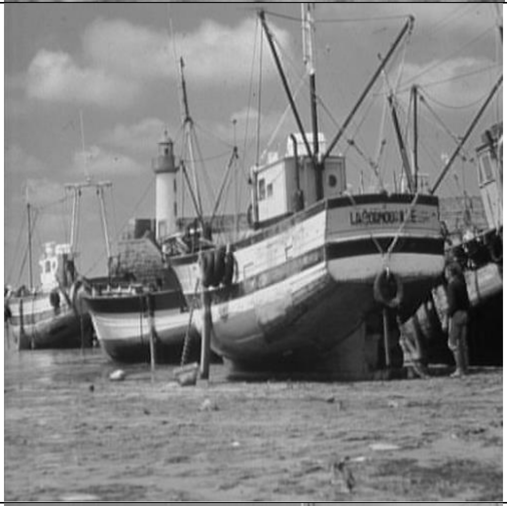
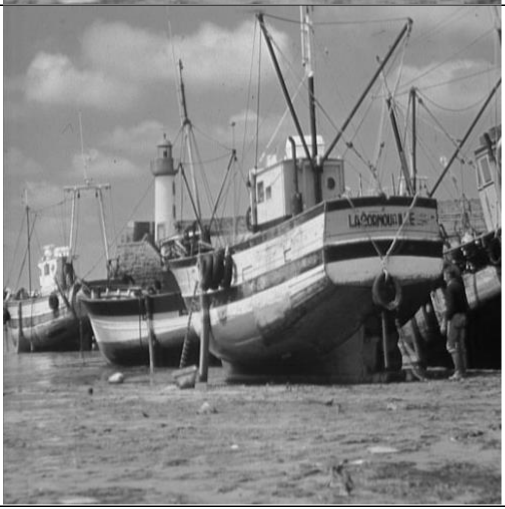
- we compare the PSNR value of reconstructed image w.r.t original image before compression as our metric for evaluating quality of reconstruction
- We assumed reconstruction with PSNR above 25 dB as acceptable/satisfactory reconstruction

2.2 Results when using K=2,4,8,16, and 32 without quantization

2.2.1 Reconstructed images for different K without quantization

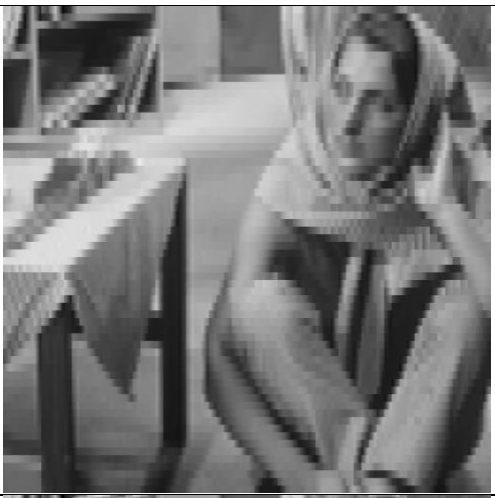
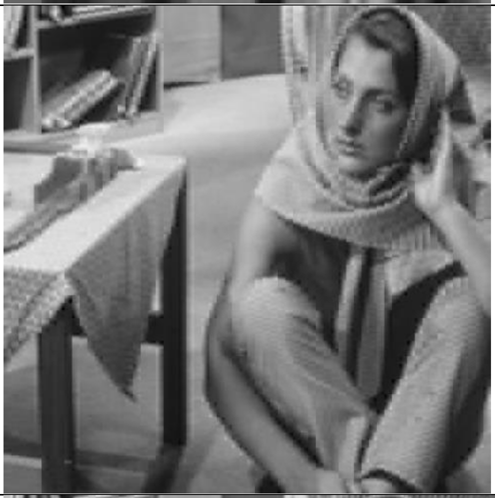
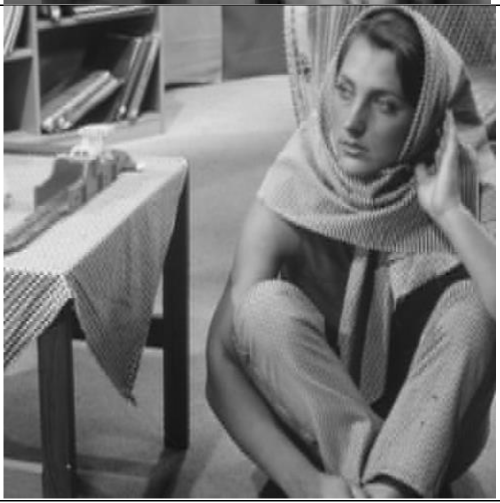
2.2.1.1 On Boat Image

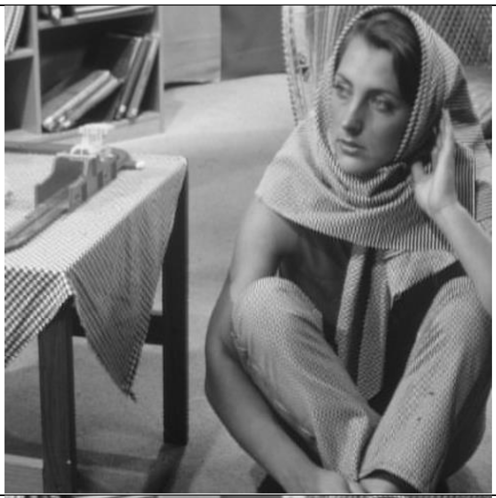
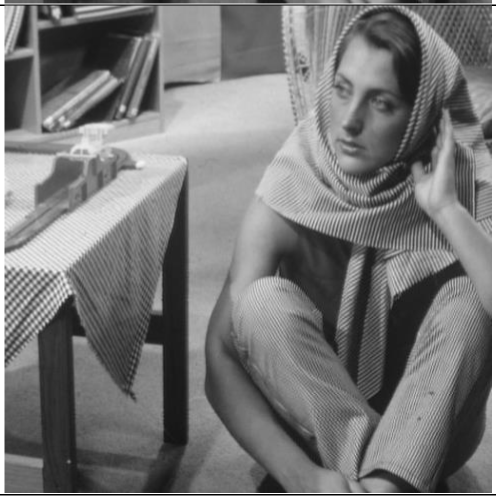
K	Original Image	Reconstructed Image	PSNR (dB)
2			23.11
4			25.27

8			28.17
16			31.46
32			36.57

2.2.1.1 On Barbara Image

K	**Original Image**	**Reconstructed Image**	**PSNR (dB)**

2			22.20
4			23.27
8			24.41

16			25.72
32			30.37

2.2.2 Discussion – reconstructed images for different K without quantization

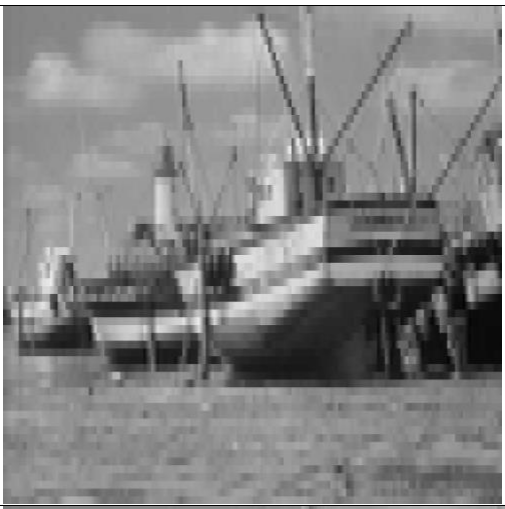
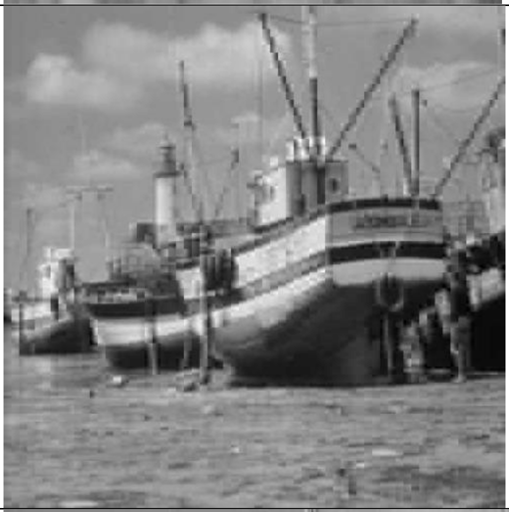
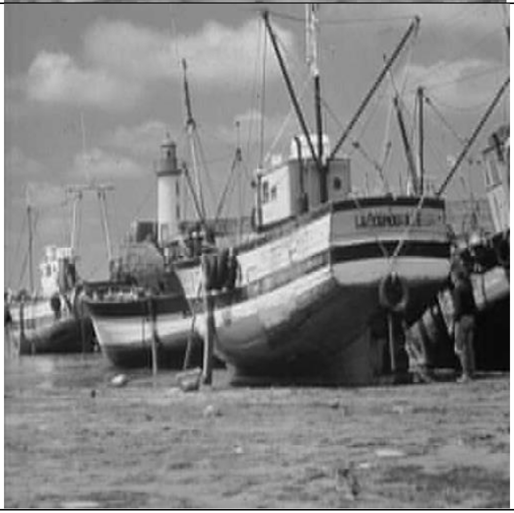
- The PSNR increases with increasing K
- For both images, for K>=16, we get PSNR >=25dB
- For boat image, even K=4,8 have PSNR >=25
- From visual results, we can see that with increasing K, details in image is increasing
- Even for lower values of K such as K=4,8 the object is perceptually visible

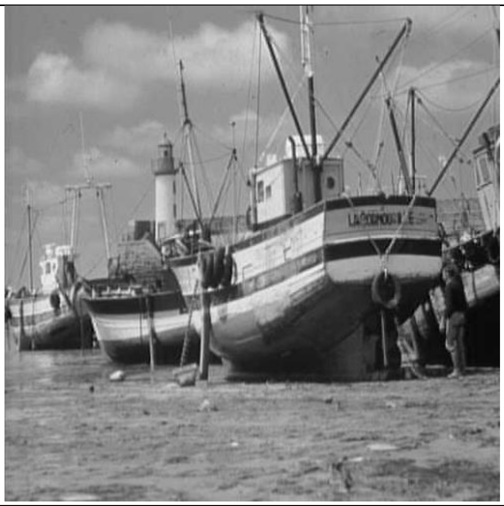
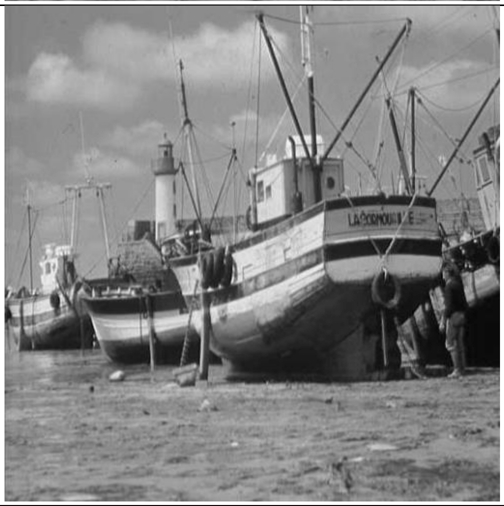
2.3 Results when using K=2,4,8,16, and 32 with quantization

2.3.1 Reconstructed images for different K with quantization

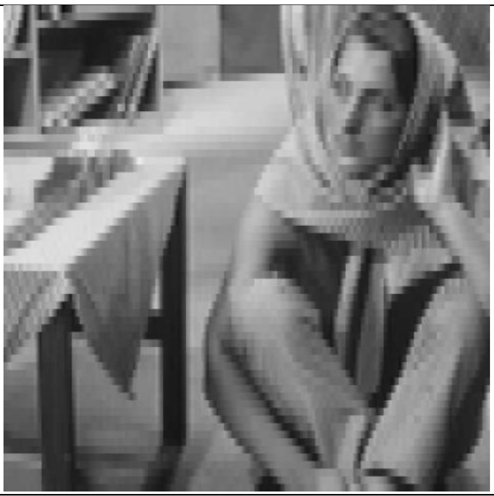
2.3.1.1 On Boat Image

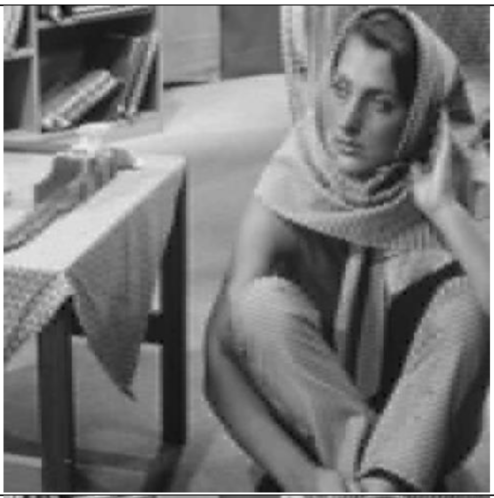
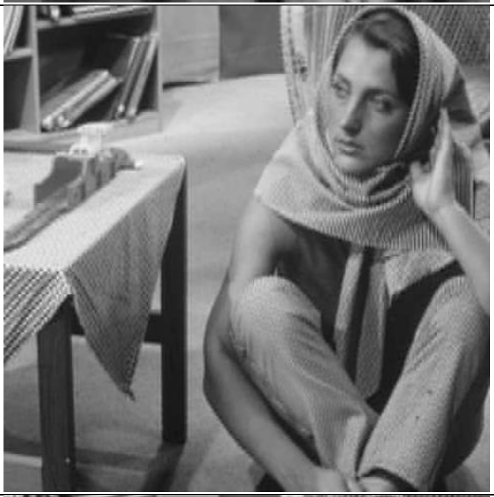
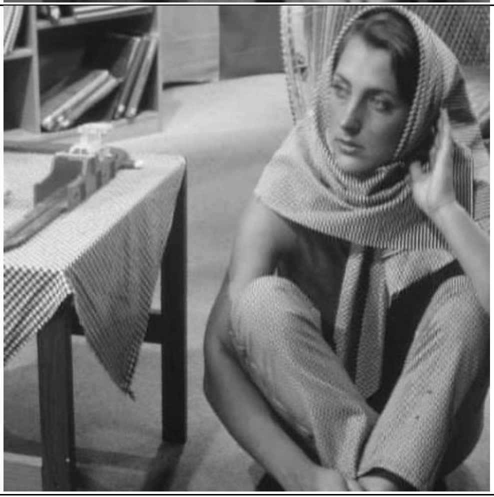
K	**Original Image**	**Reconstructed Image w/ Quantization**	**PSNR (dB)**

2			23.10
4			25.25
8			28.10

16			31.03
32			33.27

2.3.1.2 On Barbara Image

K	Original Image	Reconstructed Image	PSNR (dB)
2			22.20

4			23.26
8			24.38
16			25.60

32			29.33

2.3.2 Discussion – reconstructed images for different K with quantization

- With increasing K, PSNR and image quality still improves
- Notice quantization is making very less reduction in PSNR for K=2,4,8 (this may be due to the fact that quantization noise is suppressed from higher distortion at low value of K)
- For K =16,32 the reduction in PSNR due to quantization is more evident
- Overall, quantization is making DCT matrix more sparse but the reduction in PSNR w.r.t lossless DCT compression is not very significant at least at K<=16

Part 3: Observation and Discussion

- DCT gave better PSNR on same images w.r.t VQ
- Higher the codebook size for VQ better PSNR
- Higher the K for DCT, higher PSNR of reconstructed Image
- PSNR of VQ reconstructed image from single same image trained codebook is better in term of PSNR but not a significant gain in terms of PSNR given the increase in complexity
- Quantizing the DCT increase sparsity without very significant drop in

Part 4: References

1. Lecture Notes, ECE 634 Digital Video Systems, Professor Maggie Zhu, Purdue University
2. Lecture Notes, ECE 661 Computer Vision, Professor Avinash Kak, Purdue University
3. For Code Implementation, we initially tried LBG implementation mentioned here https://github.com/internaut/py-lbg/blob/master/lbg_test.ipynb , the time taken was very large
4. Discussed K-Means and LBG from colleagues