DTA096A - Lab 3 Report

Visual Representation and Information Theory

Name: Andrea Faccioli
Date: 19/10/2025

1 Experimental Setup

The experiments have been carried out using Python within a Jupyter Notebook environment. Core libraries included NumPy for numerical operations, OpenCV and scikit-image for image processing. All scripts and functions used are provided in Appendix A.2.

2 Scalar Quantization: The Heart of Lossy Compression

The uniform scalar quantization experiment clearly demonstrates the trade-off between compression and image quality. As the number of quantization bits decreases, the number of available intensity levels is reduced, leading to visible banding and loss of fine detail. For high bit depths (e.g., 6 bits), the quantized image is visually close to the original, with minimal perceptual distortion. At lower bit depths (e.g., 1–2 bits), severe posterization occurs, and smooth gradients are replaced by large uniform regions, producing strong visual artifacts. The quantization error image confirms this behavior: as the bit depth decreases, the error magnitude and spatial extent increase significantly. Quantitatively, the Mean Squared Error (MSE) increases and the Peak Signal-to-Noise Ratio (PSNR) decreases with fewer quantization levels, indicating a direct correlation between quantization precision and reconstruction fidelity. The results are summarized in Table 1.

Table 1: Quantization performance summary for the *Camera* image.

Bits	Levels	MSE	PSNR (dB)	Max Error
1	2	1229.22	17.23	64.00
2	4	282.04	23.63	32.00
4	16	20.77	34.96	8.00
6	64	1.51	46.33	2.00

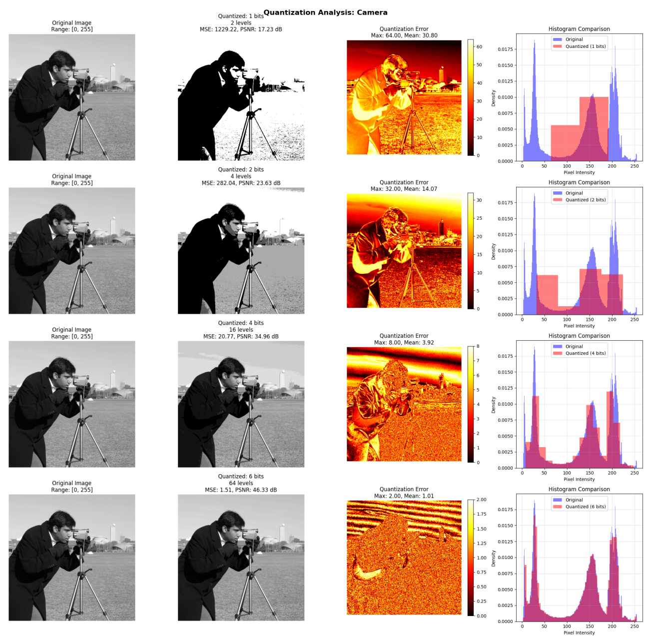

Figure 1: Analysis of the Camera image

From the visual results in Figure 1, it is evident that quantization affects smooth and textured areas differently. In the smooth regions of the image (such as the sky), coarse quantization (1–2 bits) produces strong banding artifacts, where continuous gradients are replaced by abrupt intensity transitions between a few discrete gray levels. These regions are highly sensitive to quantization because small intensity variations are visually noticeable when the number of available levels is limited. In contrast, textured areas (such as the background grass) tend to mask quantization errors more effectively, as the natural high-frequency content already contains sharp variations that disguise the introduced errors. As the bit depth increases to 4 or 6 bits, the image gradually regains smooth tonal transitions, and the quantization error images show a clear reduction in both magnitude and spatial extent. The histograms further illustrate this process: at low bit depths, the distributions collapse into a few discrete peaks, while at higher bit depths, they more closely approximate the continuous distribution of the original image.

3 Transform Coding and Quantization

Applying the Discrete Cosine Transform (DCT) to image blocks allows most of the signal energy to be concentrated in a few low-frequency coefficients, enabling more efficient quantization. The experiments compare two approaches: uniform quantization using a constant step size Q and a JPEG-style frequency-dependent quantization matrix. As shown in Table 2, increasing the quantization step Q progressively reduces precision in the DCT coefficients, leading to greater distortion in the reconstructed image and a significant drop in PSNR. At small Q values (e.g., $Q = 1$), the reconstruction is nearly lossless, while for large Q (e.g., $Q = 50$), high-frequency details are heavily suppressed, causing noticeable blurring and blocking artifacts. The number of unique quantized values per block also decreases as Q increases, indicating more coefficient clustering around zero.

Table 2: PSNR results for uniform quantization with different step sizes Q.

Quantization Step Q	**PSNR (dB)**
1	58.94
5	46.55
10	41.50
20	36.38
50	30.54

Table 3: PSNR results for JPEG-style quantization at different quality levels.

JPEG Quality	**PSNR (dB)**	**Est. percentage saved**
10	28.36	95.41%
50	32.58	87.48%
90	40.39	71.35 %

When using a JPEG-style quantization matrix (Table 3), lower frequencies are preserved more accurately while higher frequencies are coarsely quantized, resulting in a more perceptually efficient compression. As the quality parameter increases from 10 to 90, PSNR improves markedly—from 28.36 dB to 40.39 dB—demonstrating the effectiveness of adaptive frequency weighting in maintaining visual quality at reduced bitrates. Overall, transform coding combined with frequency-dependent quantization achieves a much better trade-off between compression and visual fidelity than uniform quantization.

Figure 2: Comparison of JPEG compression at different quality levels

The cropped section from the "JPEG Quality 10" image in Figure 4 clearly exhibits severe blocking artifacts, which manifest as a visible 8×8 grid. This distortion is a direct result of the coarse quantization applied during heavy compression and is particularly prominent along the high-contrast edge of the man's coat.

Figure 3: JPEG quality 10 crop visualization

4 Rate-Distortion Optimization

This experiment evaluates the rate-distortion (R-D) performance of a block-based Discrete Cosine Transform (DCT) compression scheme. The distortion was measured using the Peak Signal-to-Noise Ratio (PSNR), while the bitrate (in bits per pixel, bpp) was estimated using the first-order entropy of the quantized coefficients. Two quantization strategies were compared: a simple uniform quantizer (UQ) with a constant step size Q, and a frequency-dependent JPEG-style quantization matrix (JPEG-Q) controlled by a quality parameter.

Table 4: Rate-Distortion (R-D) results comparing Uniform Quantization (UQ) and JPEG-style Quantization (JPEG-Q) on the 'Cameraman' image. Rate is estimated via entropy of quantized coefficients.

Uniform Quantization (UQ)

Quant. Step (Q)	Est. Rate (bpp)	PSNR (dB)
1	4.703	58.94
5	2.598	46.55
10	1.875	41.50
20	1.258	36.38
50	0.568	30.54

JPEG-Style Quantization (JPEG-Q)

Quality Setting	Est. Rate (bpp)	PSNR (dB)
10	0.182	35.23
30	0.322	39.52
50	0.462	41.15
70	0.669	42.83
90	1.320	46.89

The results, summarized in Table 4, demonstrate the fundamental trade-off: increasing the compression (i.e., lowering the estimated bitrate) invariably results in higher distortion (a lower PSNR). However, the two methods show a significant difference in efficiency. The JPEG-style quantization, which exploits the human visual system's lower sensitivity to high-frequency noise, achieves a much better R-D trade-off. For instance, the JPEG-Q method at Quality 90 achieves a PSNR of 46.89 dB with an estimated rate of only 1.320 bpp. To achieve a comparable quality (46.55 dB), the uniform quantizer requires nearly double the bitrate (2.598 bpp). This superiority is consistent across the data range, indicating that an R-D curve plotting the JPEG-Q results would lie significantly above and to the left of the curve for uniform quantization, confirming its higher compression efficiency.

5 Image Quality Assessment

While PSNR is a widely used metric for image quality assessment, it does not always correlate well with human visual perception, as it measures only pixel-wise differences without considering structural or perceptual aspects. To address this limitation, the Structural Similarity Index (SSIM) was computed between the original and compressed images at various JPEG-style quantization levels. SSIM evaluates luminance, contrast, and

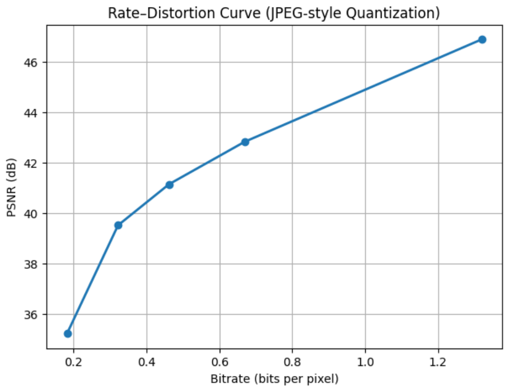

Figure 4: RD curve of a JPEG quantizer

structural similarity, providing a more perceptually relevant measure. As shown in the results, both PSNR and SSIM increase with bitrate, reflecting improved reconstruction quality at higher compression quality levels. However, their scales and sensitivities differ: as reported in Figure 5, the plots have a different growth behaviour. Overall, SSIM provides a more consistent reflection of visual improvements, and cases may arise where PSNR suggests similar quality between images, while SSIM better captures subtle structural differences perceived by the human eye.

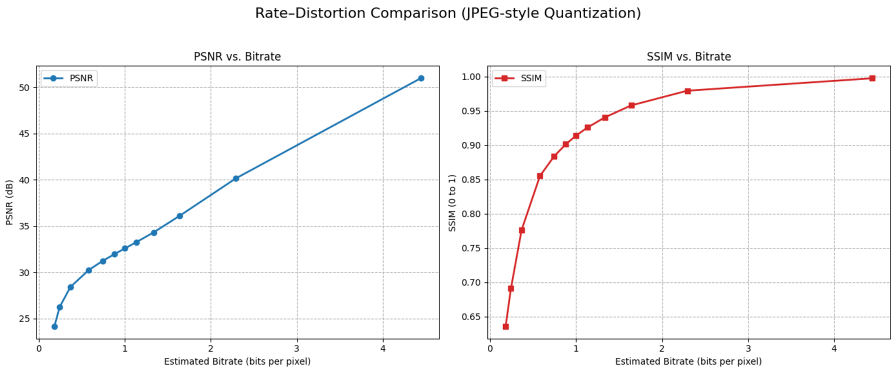

Figure 5: PSNR and SSIM curve for JPEG compression (Camera image)

6 A Simplified JPEG-like Codec

The objective of this task was to design, implement, and evaluate a complete lossy image compression pipeline, emulating the core principles of the JPEG standard. The encoder was architected to first process the input image in 8×8 blocks. Each block was transformed from the spatial domain to the frequency domain using the 2D Discrete Cosine Transform (DCT). Following the transform, a quantization step was applied by dividing the DCT coefficients by a predefined quantization matrix. This is the primary lossy stage, reducing the precision of high-frequency components that are less perceptible to the human eye. The quantized coefficients were then linearized using a zigzag scan and efficiently grouped using Run-Length Encoding (RLE) to represent the many resulting zero-value coefficients. Finally, these RLE symbols were entropy coded using a custom Huffman algorithm to generate the final compressed bitstream. The decoder reverses this entire process: it performs Huffman decoding, decodes the run-length data, dequantizes the coefficients using the same matrix, and applies the inverse 2D DCT to reconstruct the image.

The implemented codec was evaluated on a test image, and the performance metrics are summarized in Table 5. The codec demonstrated substantial compression capabilities. The original 512×512 8-bit grayscale image, representing 2,097,152 bits, was successfully compressed into a bitstream of only 47,064 bits. This equates to a remarkable storage saving of 97.76% and a high compression ratio of approximately 44.57:1.

To assess the fidelity of the reconstructed image, the Peak Signal-to-Noise Ratio (PSNR) was calculated. The resulting PSNR of 39.54 dB is an excellent score, indicating a very high degree of similarity between the original and reconstructed images. A value in this range suggests that compression artifacts are minimal and likely imperceptible to the average observer. This demonstrates that the codec achieves a very strong

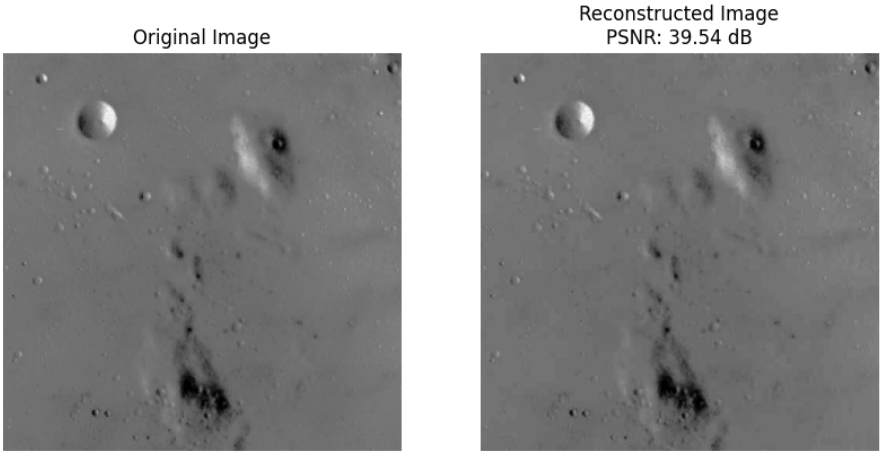

Figure 6: Original and reconstructed image (Moon)

Table 5: Codec Performance Metrics on Test Image

Metric	Value
Original Image Size	2,097,152 bits
Compressed Bitstream Size	47,064 bits
Compression Ratio	44.57:1
Storage Saving	97.76%
PSNR (Reconstructed)	39.54 dB
Unique Huffman Codes	124

balance, managing to attain a high compression ratio while preserving excellent visual quality. The Huffman coding stage's generation of 124 unique codes further confirms its role in effectively compressing the statistical redundancies present in the quantized RLE symbols. Further analysis could be conducted by experimenting with different quantization matrix or encoding strategies (e.g. arithmetic coder).

A Appendix 1

A.1 Images collection

The images processed in Lab 3 session can be found here.

A.2 Source code

The complete code of Lab 3 session can be found here.