E9 222 Signal Processing in Practice
Assignment 2

January 21, 2026

Dwaipayan Haldar

---

1. Plot 1D DCT Basis Functions

(a) Construct the basis

Ans: A function `dct_matrix(N)` was implemented that constructs the $N \times N$ orthonormal DCT-II matrix D with entries:

$$D_{k,n} = \phi_k[n] = \alpha_k \cos\left(\frac{\pi}{N}\left(n + \frac{1}{2}\right)k\right)$$

where $\alpha_k = \sqrt{1/N}$ for $k = 0$ and $\alpha_k = \sqrt{2/N}$ for $k = 1, 2, \ldots, N-1$.

For $N = 32$, The orthonormality of the DCT matrix is verified by computing the Frobenius norm:

$$\|DD^\top - I\|_F = 1.8845 \times 10^{-14}$$

Ideally, it should be 0. But the small error is there due to the rounding off effect in machine and machines cannot be infinite precision.

(b) Visualize basis functions

Ans: Figure 1 shows the first eight DCT basis functions ($k = 0, 1, \ldots, 7$) for $N = 32$.

Observations:

- The $k = 0$ basis function is a constant (DC component), representing the average value of the signal.
- As k increases, the basis functions oscillate with increasing frequency.
- The basis functions form a complete orthonormal set, enabling perfect reconstruction of any signal.

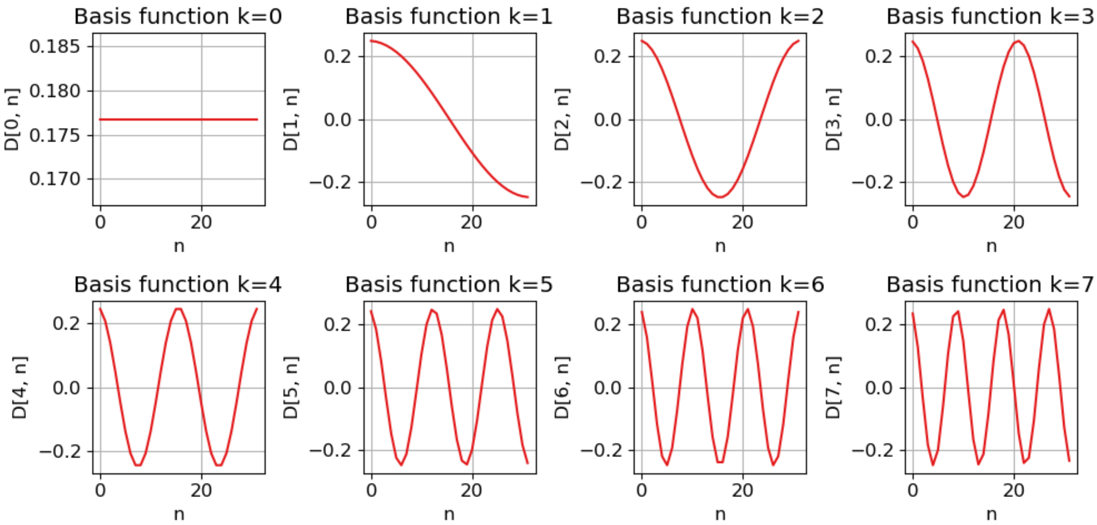

Figure 1: First eight 1D DCT-II basis functions ($k = 0, 1, \ldots, 7$) for $N = 32$.

2. Plot 2D DCT Basis Functions

(a) Separable 2D basis

Ans: The 2D DCT basis functions are constructed using the separable property of the DCT. For indices (u, v), the 2D basis pattern is defined as:

$$\Phi_{u,v}[m, n] = \phi_u[m] \cdot \phi_v[n], \quad m, n \in \{0, \ldots, 7\}$$

A function `dct2d_basis(u, v, M, N)` was implementated that computes the outer product of two 1D basis vectors to generate the 2D basis image for any (u, v).

(b) Visualize the full 8×8 set

Ans: Figure 2 displays all 64 basis patterns $\{\Phi_{u,v}\}$ arranged in an 8×8 grid, where rows are indexed by u (vertical frequency) and columns by v (horizontal frequency).

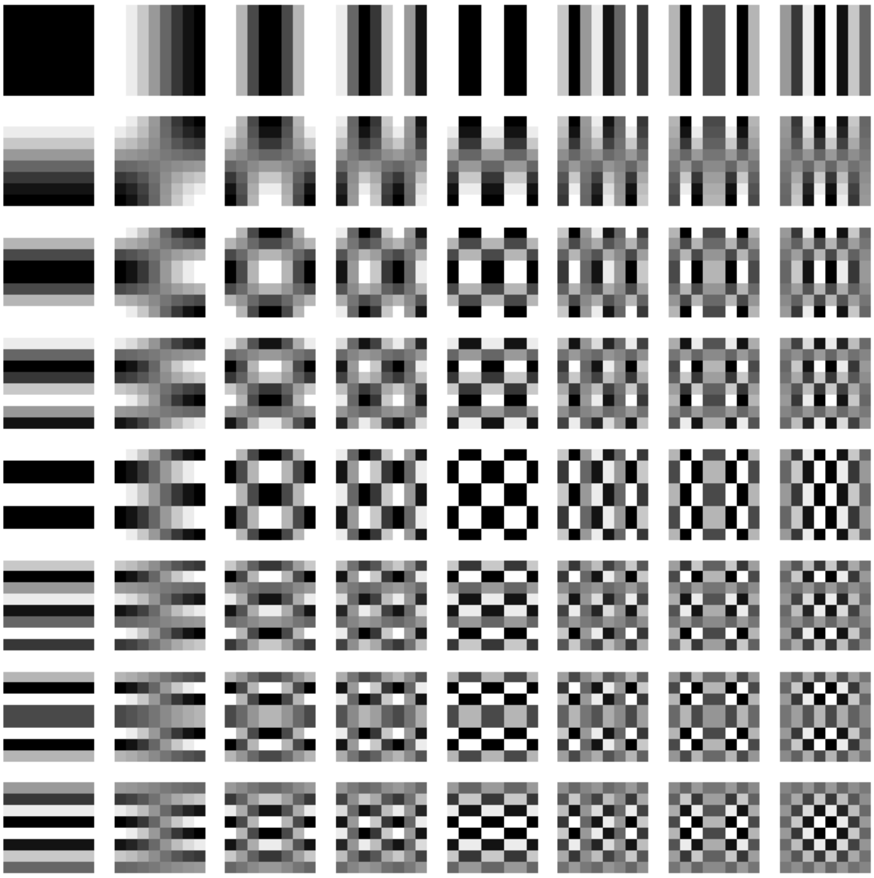

Figure 2: Complete set of 64 2D DCT basis patterns (8×8 grid) for $M = N = 8$.

(c) 2D DCT on image blocks

Ans: (i) 2D DCT on an 8×8 image block:

The 2D DCT was applied to an 8×8 grayscale image block by projecting onto the 2D basis functions. The DCT coefficients $\hat{x}[u, v]$ are computed as:

$$\hat{x}[u, v] = \sum_{m=0}^{7} \sum_{n=0}^{7} x[m, n] \cdot \Phi_{u,v}[m, n]$$

(ii) Perfect reconstruction:

The image was reconstructed from its DCT coefficients using:

$$x[m, n] = \sum_{u=0}^{7} \sum_{v=0}^{7} \hat{x}[u, v] \cdot \Phi_{u,v}[m, n]$$

Figure 3 shows the original and reconstructed images, demonstrating perfect reconstruction.

(iii) Energy distribution in DCT coefficients:

Figure 4 shows the distribution of DCT coefficients for a typical image block. The energy is highly concentrated in the low-frequency coefficients (top-left corner), with the DC coefficient $(0, 0)$ containing the largest magnitude. Higher frequency coefficients (bottom-right) have significantly smaller values.

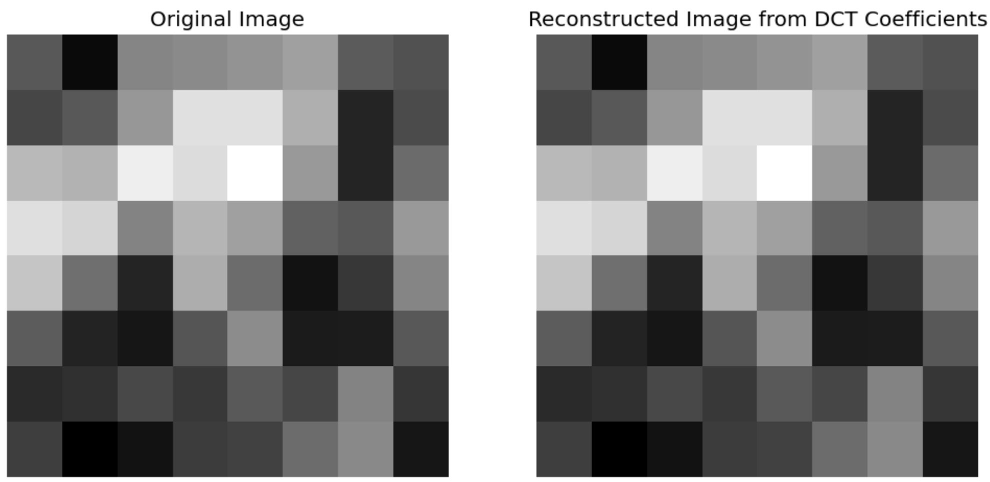

Figure 3: Original 8×8 image block and its reconstruction from DCT coefficients.

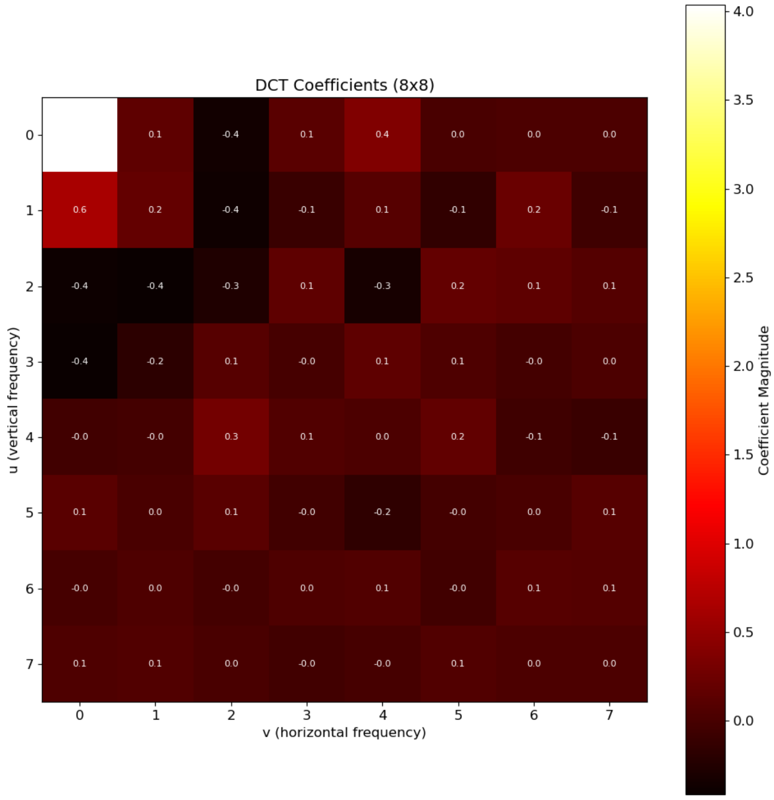

Figure 4: DCT coefficient magnitudes for an 8×8 image block. Energy is concentrated in low-frequency components (top-left).

3. Implement JPEG-style Compression and Investigate Energy Compaction

(a) Block DCT and reconstruction (no quantization)

Ans: Block based DCT was implemented and error between reconstructed and original is calculated. It is approximately 0.

The maximum absolute reconstruction error without quantization was:

$$\max |x_{\text{original}} - x_{\text{reconstructed}}| = 5.329 \times 10^{-15}$$

This error is at machine precision, confirming perfect reconstruction when no quantization is applied.

(b) Quantization (JPEG idea)

Ans: JPEG-style was implemented quantization using a parametric quantization matrix:

$$Q[u, v] = 1 + s(u + v)$$

where $s > 0$ is the compression strength parameter. The quantization process is:

$$\hat{x}_Q[u, v] = \text{round}\left(\frac{\hat{x}[u, v]}{Q[u, v]}\right), \quad \tilde{x}[u, v] = \hat{x}_Q[u, v] \cdot Q[u, v]$$

This quantization matrix applies stronger quantization to higher frequency components (larger $u+v$), which aligns with the JPEG philosophy of preserving low-frequency information while aggressively quantizing high-frequency details that are less perceptually important.

(d) Quality vs compression study

Ans: Five compression strengths was tested: $s \in \{0.01, 0.1, 1, 10, 100\}$. Figures 5–9 show the reconstructed images at each compression level.

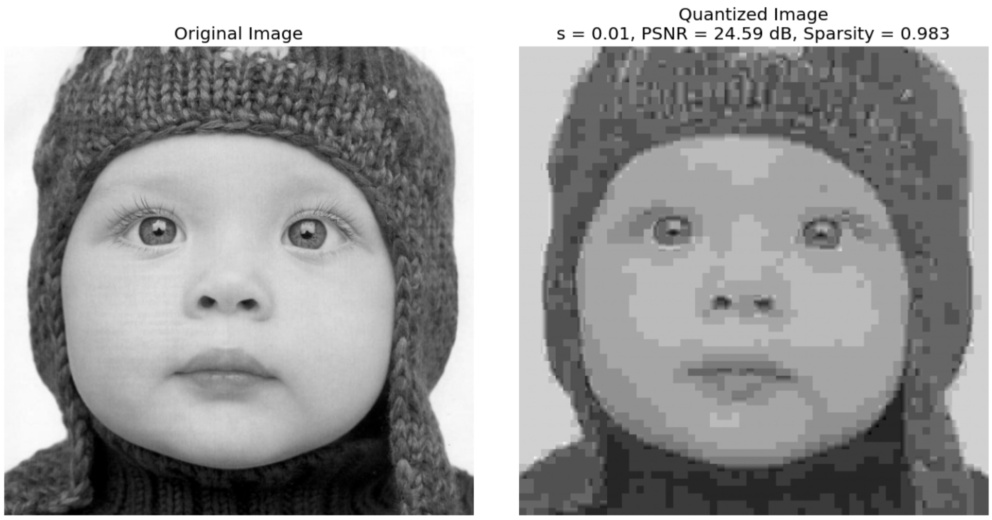

Figure 5: Compression with $s = 0.01$: Near-lossless reconstruction.

Figure 6: Compression with $s = 0.1$: High quality reconstruction.

Figure 7: Compression with $s = 1$: Moderate compression with some visible artifacts.

Figure 8: Compression with $s = 10$: High compression with noticeable blocking artifacts.

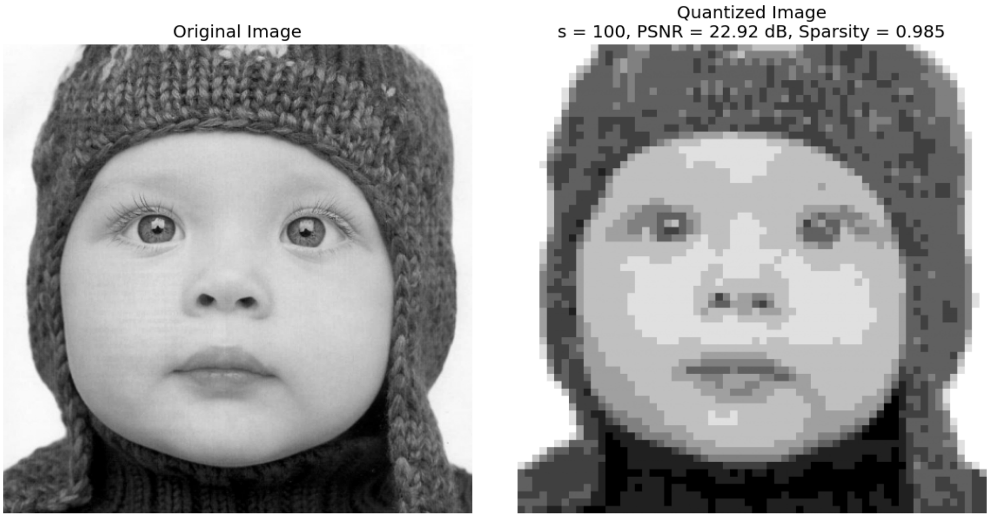

Figure 9: Compression with $s = 100$: Severe compression with significant quality degradation.

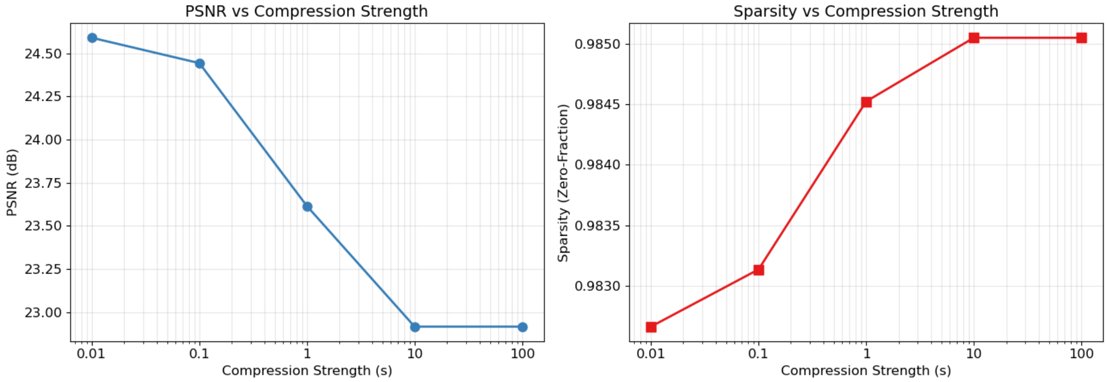

Figure 10: PSNR and Sparsity vs Compression Strength. Left: PSNR decreases as s increases. Right: Sparsity (zero-fraction) increases with s.

(e) Energy compaction (core concept)

Ans: I computed the energy compaction for all 8×8 blocks in the image:

$$E_{\text{total}} = \sum_{u,v} |\hat{x}[u,v]|^2, \quad E_K = \sum_{(u,v)\in S_K} |\hat{x}[u,v]|^2$$

where S_K contains the indices for a $K \times K$ region.

(a) Low-frequency energy fractions (top-left $K \times K$):

K	$E_K/E_{\textbf{total}}$
1	0.9584
2	0.9824
3	0.9917
4	0.9952
5	0.9974
6	0.9987
7	0.9995
8	1.0000

Table 1: Energy fraction in top-left $K \times K$ coefficients.

(b) Interpretation:

The energy accumulates extremely rapidly in the low-frequency coefficients:

- Just the DC coefficient ($K = 1$) captures **95.84%** of the total energy.
- The top-left 2×2 coefficients capture **98.24%** of the energy.

- By $K = 4$, we have **99.52%** of the total energy using only 25% of the coefficients.

This demonstrates the excellent **energy compaction** property of the DCT for natural images.

(c) High-frequency comparison (bottom-right $K \times K$):

K	$E_K/E_{\textbf{total}}$
1	0.0000
2	0.0000
3	0.0001
4	0.0004
5	0.0012
6	0.0032
7	0.0110
8	1.0000

Table 2: Energy fraction in bottom-right $K \times K$ coefficients.

The high-frequency coefficients contain negligible energy. Even the bottom-right 7×7 region (49 out of 64 coefficients, excluding only the DC and first row/column) contains only **1.1%** of the total energy. This stark contrast between low and high frequency energy distribution is why JPEG compression works so well—we can heavily quantize (or even discard) high-frequency coefficients with minimal perceptual impact.

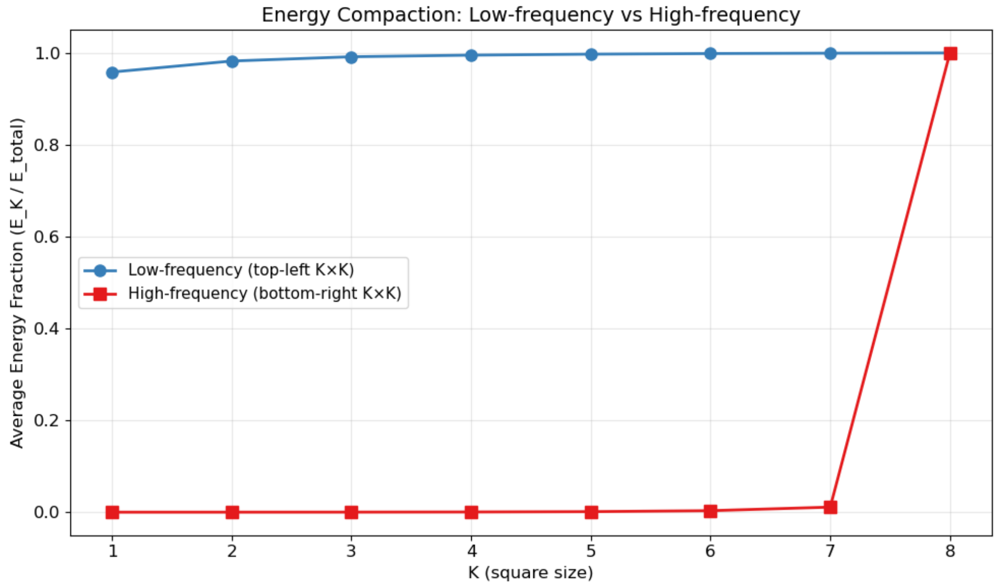

Figure 11: Energy compaction comparison: Low-frequency (top-left $K \times K$) vs High-frequency (bottom-right $K \times K$). The DCT exhibits excellent energy compaction, with most energy concentrated in low-frequency coefficients.