Waves Are Everywhere: The Fourier Kingdom and Signal Processing in Image Compression

Indy Lyness, Luca Araujo, Joshua Lee, Jiale Wan, Layla Oesper

Carleton College, Computer Science Department

1. Introduction

There are many different image compression algorithms that exist, but almost all of them share a common factor- a mathematical transform. By taking the image data to a different space, it is possible to drastically decrease the associated entropy of the pixels before applying compression techniques. For our project we focused on several so-called transform coding schemes, including several connected to a mathematical juggernaut- the Fourier Transform.

2. The Discrete Fourier Transform (DFT)

The DFT was originally developed as a way to deconstruct sound waves into their original frequency components. However, any ordered list of values

$$x = \{x_1, x_2, \ldots, x_N\},$$

called a **time signal**, can be represented as a sum of complex exponentials $e^{2\pi i k/N}$, where $0 \leq k < N$. The normalized coefficients for these exponentials are then given by

$$X_k = \frac{1}{\sqrt{N}} \sum_{m=0}^{N-1} x_m e^{-2\pi i k m/N},$$

forming a new signal $X = \{X_1, X_2, \ldots, X_N\}$. We say that X lives in the **frequency domain**. In this sense, the DFT can effectively be thought of as a **change of basis** in $\mathbb{R}^N$.

3. Transforming an Image with the 2D DFT

If we view an image as a 2D array of pixel values, the whole image can be transformed by applying the DFT to each row and column:

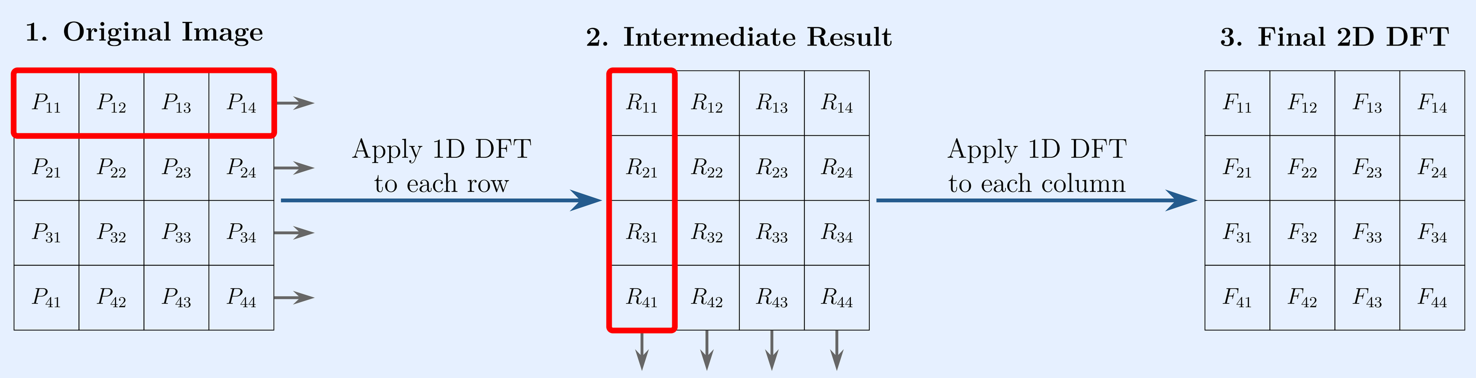

Fig. 1: Applying the 2D DFT on a sample 4x4 chunk

The transform is **lossless**, so the original image can be fully reconstructed. However, the energy of the image is now concentrated in the lower frequencies, resulting in decreased **entropy**. In practice, the image is first broken up into smaller chunks (the standard is 8x8).

Fig. 2: Original Image

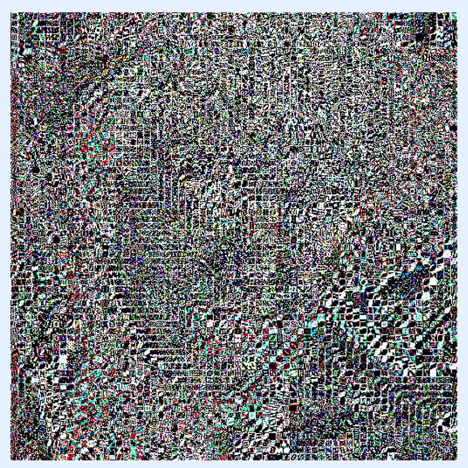

Fig. 3: Transformed Image Visualization

4. The Discrete Cosine Transform (DCT)

A drawback of the DFT is that it assume the original signal is periodic modulo N, which can cause more energy in higher frequencies due to a sharp change in values. To fix this, the original signal $x = \{x_1, \ldots, x_n\}$ is mirrored into a new signal

$$\hat{x} = \{x_1, \ldots, x_n, x_n, \ldots, x_1\}.$$

This results in a smoother signal with lower high frequency energy:

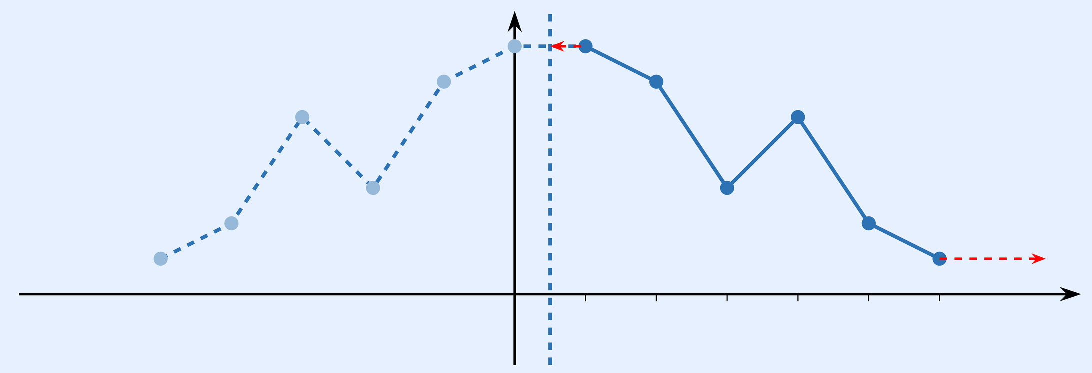

Fig. 4: Constructing the DCT

The DCT is then obtained by running the DFT on the resulting sequence. Thus, the kth component is then

$$X_k = \sum_{m=0}^{2N-1} \hat{x}_m e^{-2\pi i k m/2N} = 2e^{\pi i k/2N} \sum_{m=0}^{N-1} x_m \cos\left(\frac{\pi k(2m+1)}{2N}\right).$$

After applying a phase shift and normalizing, this gives the final version

$$X_k = \frac{c_k}{\sqrt{N}} \sum_{m=0}^{N-1} x_m \cos\left(\frac{\pi k(2m+1)}{2N}\right),$$

where $c_k = 1$ if $k = 0$ and $c_k = \sqrt{2}$ otherwise.

5. Results

To test the ability of our transform algorithms, we ran each on 4 different datasets of images and recorded several metrics. Figure 5 shows how the **Peak Signal to Noise Ratio (PSNR)** varied for each dataset as the **Compression Ratio (CR)** was increased.

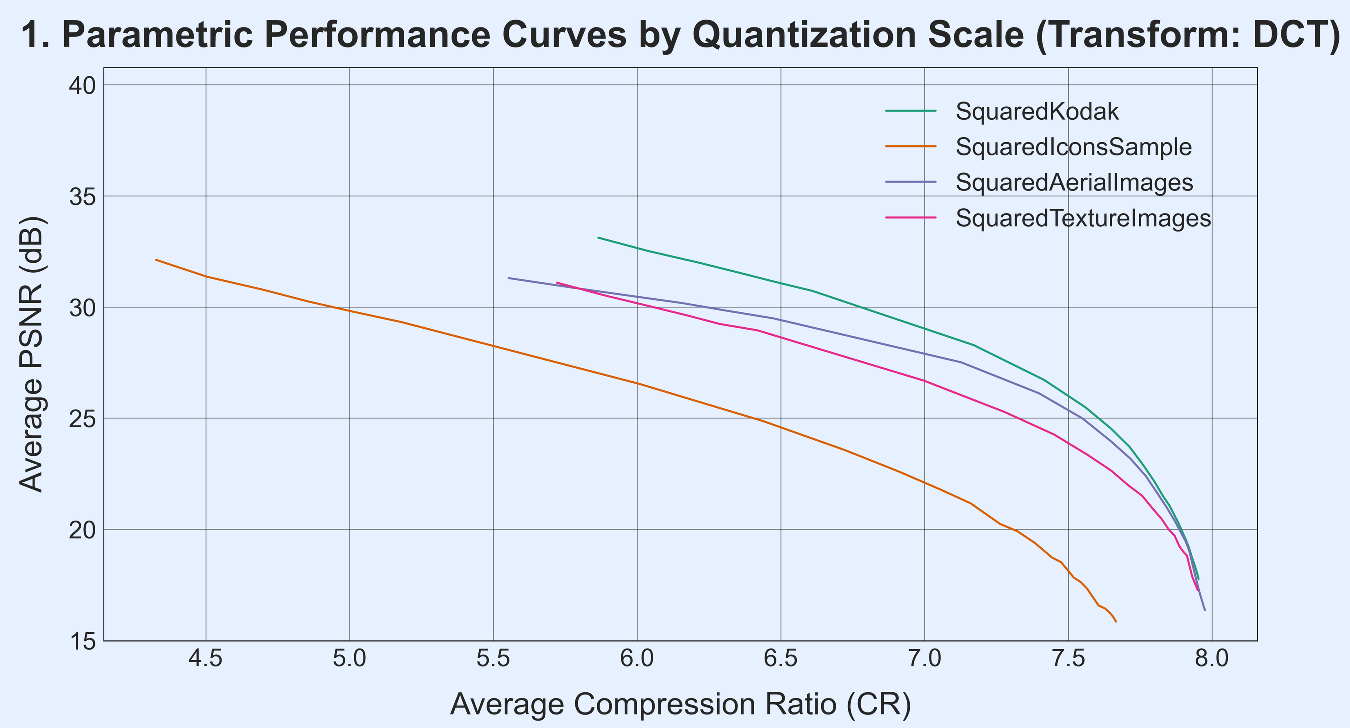

Fig. 5: Average CR vs PSNR by Dataset for DCT

Notably, there is some variance between the different datasets due to the typical properties of images in each. For example, the rapid changes of pixels in the **Textures** dataset is harder to compress than the much more gradual changes in **Aerial Images**.

6. More Results

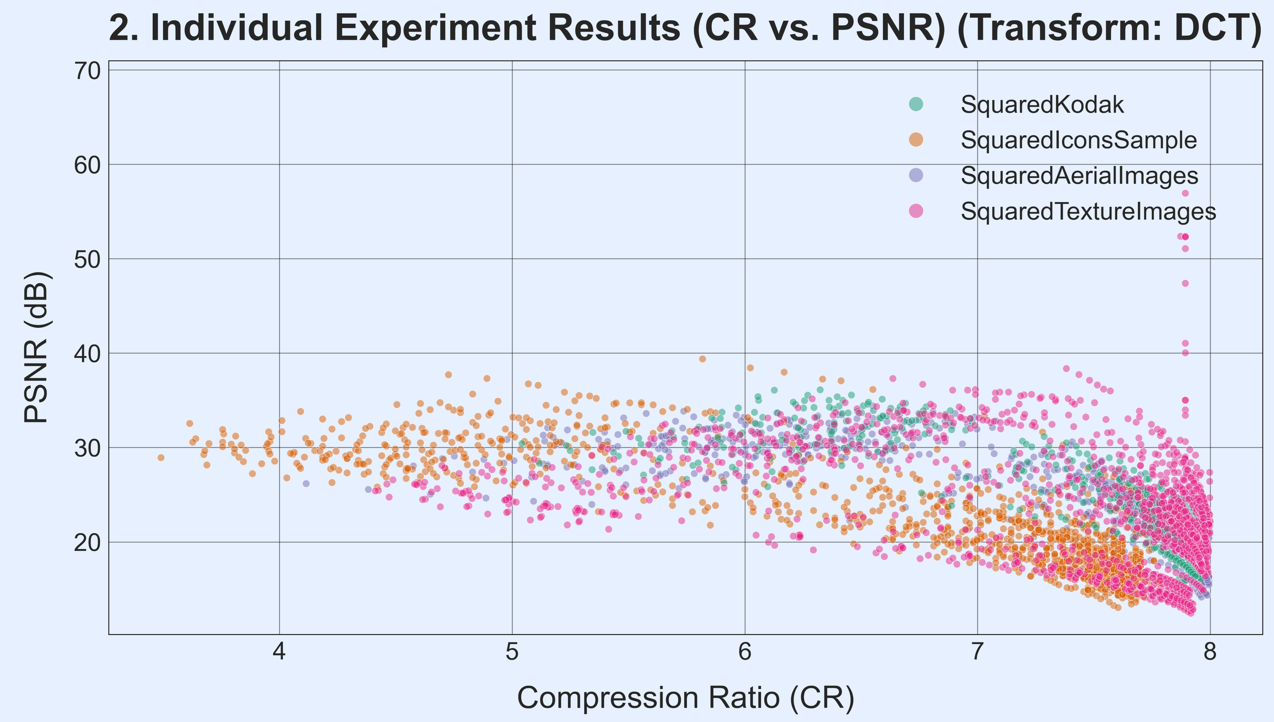

Fig. 6: CR vs PSNR by each image for DCT

Figure 6 shows the breakdown for each individual image. The line-like patterns present come from the same image as the compression ratio is increased. It is now more clear that there are several images with excellent PSNR statistics even at a very high CR, while other images already see a drop in quality even with a low CR.

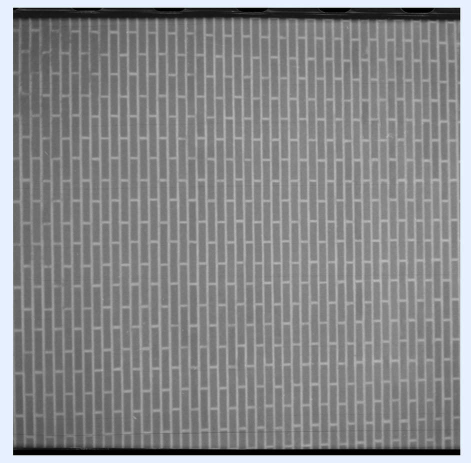

Fig. 7: An image suitable for compression

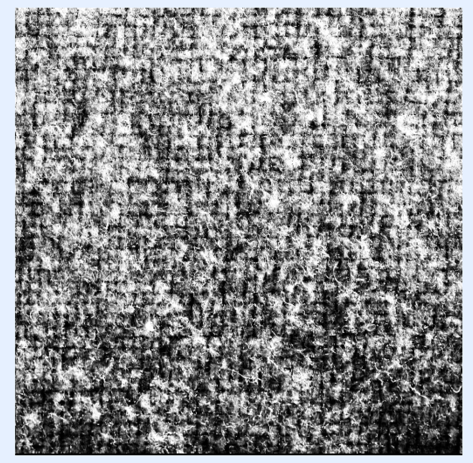

Fig. 8: A more difficult to compress image

Above are two images from the **Textures** dataset. Figure 7 shows an image that is mostly smooth at a small chunk size with repeating patterns. This makes it easy to encode with just a few low frequency components. Figure 8 shows a much more chaotic image, with much more high energy coefficients needed to represent it.

References

[Kha03] Syed Ali Khayam, *The discrete cosine transform (DCT): theory and application*, Michigan State University **114** (2003), no. 1, 31.

[Wal92] Gregory K Wallace, *THE JPEG STILL PICTURE COMPRESSION STANDARD*, IEEE Transactions on Consumer Electronics **38** (1992), no. 1 (en).

Acknowledgements

Thanks to my Comps groupmates Luca, Josh, and Wan. Also, a special thanks to my advisor Layla Oesper!