Transformers Zoo: Wavelet Transforms, what role to they play in Image Compression?

Luca Araujo Indy Lyness Joshua Lee Jiale Wan

Data Compression 101

Image compression algorithms are a subclass of a more general class of algorithms which are data compression algorithms. In general, those aim to efficiently encode data in the least amount of bits possible. This can be useful in a plethora of scenarios, from storage purposes to network transmission. An important distinguishing property of compression algorithms is whether they are lossless or lossy:

- **Lossless encodings**: all the original data can be perfectly reconstructed.
- **Lossy encodings**: only a selected "most relevant" subset of the data can be reconstructed. Those encodings can usually achieve much higher compression ratios but are specific to each data type.

Image compression algorithms are tailored to the special properties of images to achieve better compression ratios and good lossy encodings by selecting which areas are least relevant and thus prone to be removed.

What is special about images?

Image data is highly correlated! Take an arbitrary pixel in an image that happens to be red. Most likely the pixel right next to it will also be of some shade of red, with very similar colors.

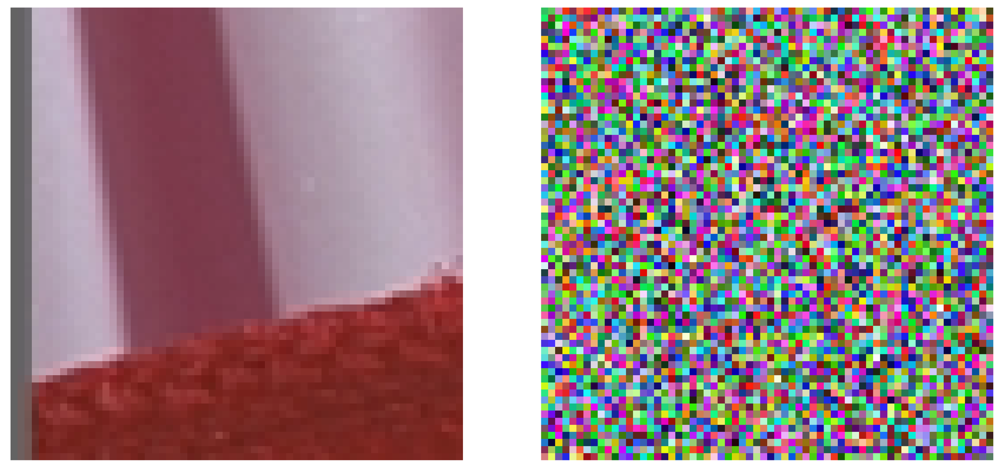


Figure 1. Zoomed in 64x64 image showing pixel correlation (left) vs. randomly generated 64x64 image with no correlation (right).

Transforms: What are they?

Transform coding is the key to exploiting image correlation! As we've seen, neighboring pixels in an image tend to be very similar, creating redundancy. Transforms are mathematical operations that change the representation of image data — essentially a change of basis that redistributes information.

The goal is to *decorrelate* the pixel data and concentrate most of the image information into fewer independent coefficients. Instead of storing similar values for many adjacent pixels, transforms allow us to store a few large coefficients (representing important features) and many small coefficients (representing fine details) that can be coarsely quantized for lossy compression.

We investigate several transform coding techniques in our work:

- Discrete Wavelet Transforms (DWT)
- Discrete Cosine Transform (DCT)
- Discrete Fourier Transform (DFT)
- S+P Transform

Discrete Wavelet Transforms

DWT with Haar wavelet in practice

Let there be a 1D signal with a resolution of 8 pixels. An idea to compress it might be to store the average of adjacent values.

However, we note that this is clearly losing information. In order to be able to perfectly reconstruct the original signal, we also must save the difference between adjacent values.

We can keep recursively applying this reduction to the averages until we reach only one pixel.

Resolution	Averages	Details
8	$[5,\ 8,\ 9,\ 7,\ 4,\ 2,\ 1,\ 3]$	–
4	$[6.5,\ 8,\ 3,\ 2]$	$[-1.5,\ 1,\ 1,\ -1]$
2	$[7.25,\ 2.5]$	$[-0.75,\ 0.5]$
1	$[4.875]$	$[2.375]$

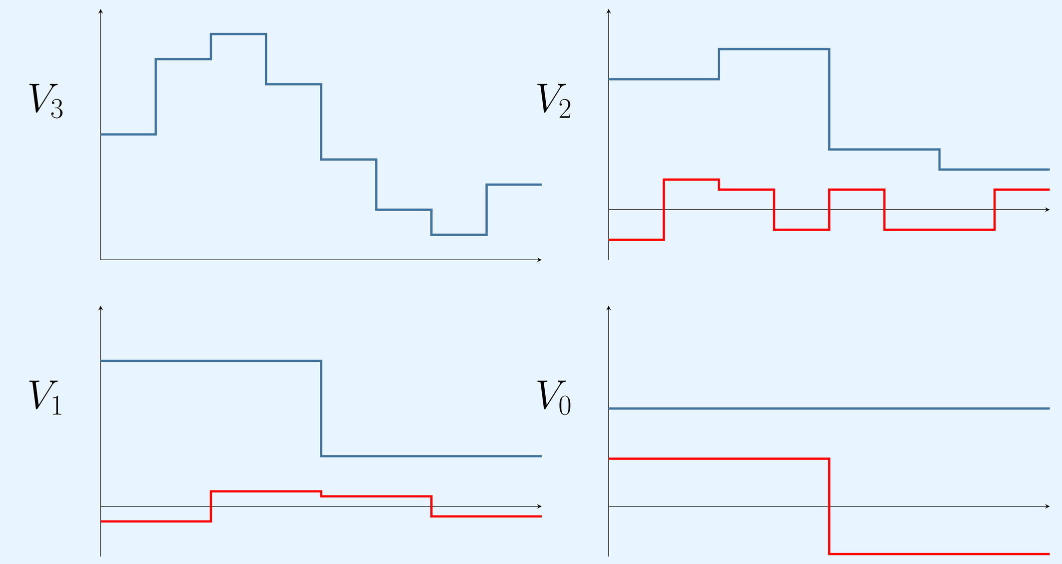


Figure 2. Haar wavelet decomposition showing signal (blue) and wavelets (red) at each resolution level.

We can then apply this to 2D signals (images!) by applying 1D reductions to each row and column one at a time.

Figure 3. Original image (left) and encoded visualization after Wavelet Transform using the Haar Wavelet (right).

Theory: Haar basis function

We can represent signals of size 2^k with vector spaces of piecewise constant functions from the half-open interval $[0, 1)$ to $\mathbb{R}$.

A 1-pixel image can be represented as a function mapping the entire $[0, 1)$ interval to a number, we call that space V_0.

A 2-pixel image is a function that maps $[0, 1/2)$ to a number and $[1/2, 1)$ to a (maybe distinct) number, we call that space V_1.

In general, we can define V_k to be the space of all piecewise constant functions from $[0, 1)$ with exactly 2^k constant pieces:

$$[0, 1/2^k), [1/2^k, 2/2^k), \cdots, [2^k - 1/2^k, 1)$$

A natural basis for such space (using V_3 as example) is the set of shifted and scaled Haar scaling functions:

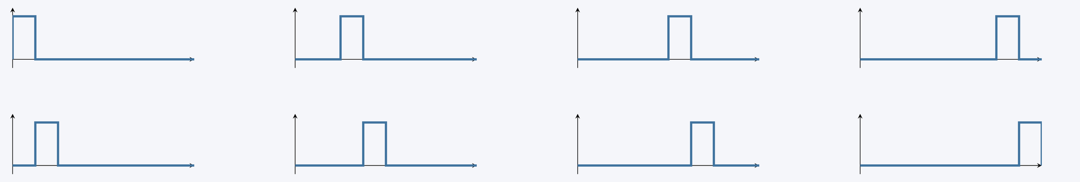


Figure 4. Scaling functions (basis) for V_3: eight piecewise constant functions.

Note that

$$V_0 \subset V_1 \subset V_2 \subset \cdots \subset V_k$$

This is a necessary property for the transform and the reason why we represent the signals as this family of vector spaces.

For the compression we essentially want to change the representation from the natural basis for V_k to the basis for V_{k-1}, which has a condensed version of the information, plus some other functions representing the detail, which are the wavelets.

We can formally define the wavelet space W_{k-1} as the orthogonal complement of V_{k-1} in V_k. The wavelet functions are a basis for this wavelet space. It can be shown that the union of a basis of V_{k-1} and a basis of W_{k-1} forms a basis for V_k.

If you are wondering what orthogonal complements even are, instead of explaining it is more effective to see the image of a wavelet basis along its corresponding scaling basis.

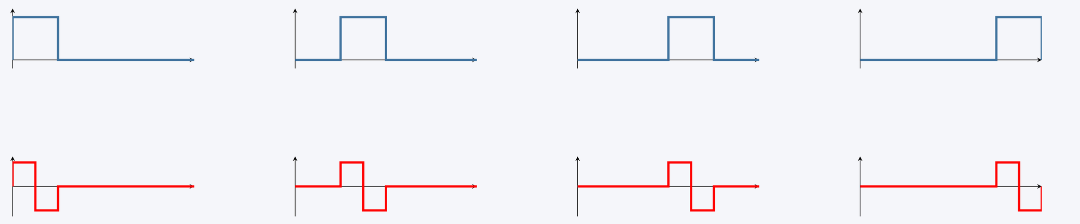
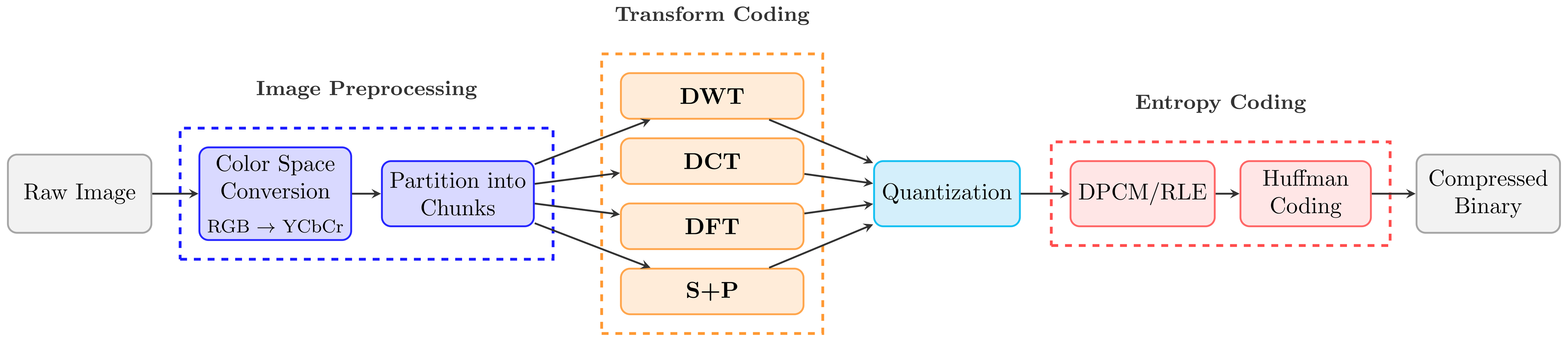


Figure 5. Basis for V_2 (blue) and Haar wavelets basis for W_2 (red).

Image Compression Encoding Pipeline



Experimental Setup

Datasets

- **Kodak** Standard set of high quality photographic images
- **Aerial Images** Aerial and satellite imagery
- **Textures** Images with repetitive patterns and fine details
- **Icons** Icons and graphic images with sharp edges

Metrics

- Compression ratio
- Mean Squared Error (MSE)
- **Peak Signal-to-Noise Ratio (PSNR)** log scale quality metric based on MSE
- Encoding/decoding time
- **Entropy**: Theoretical information content measured at each pipeline stage (original, transformed, quantized)

Results

Compression x Quality (PSNR) Performance Analysis

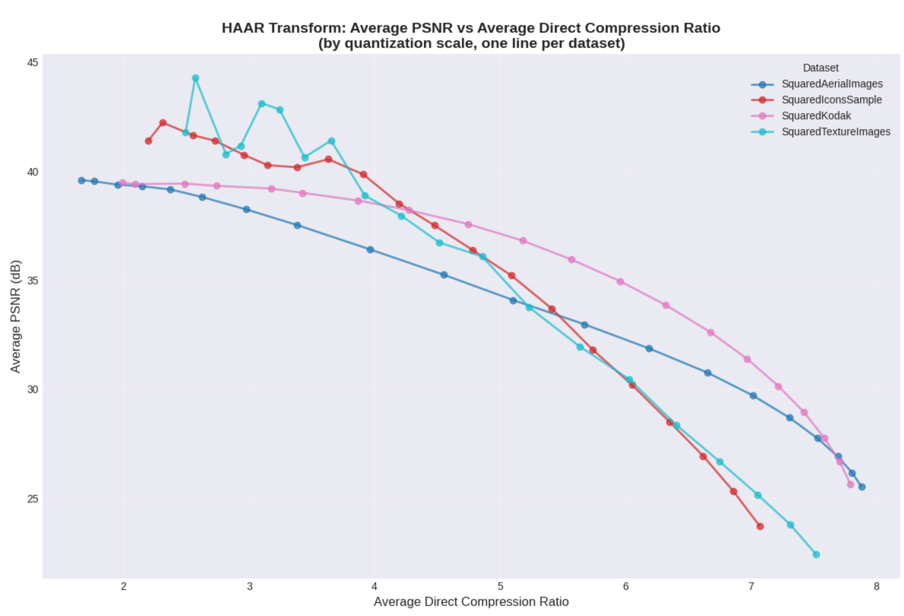


Figure 6. Curve showing Compression Ratio x PSNR performance for Haar Wavelet Transform for each dataset.

Visual Quality Assessment

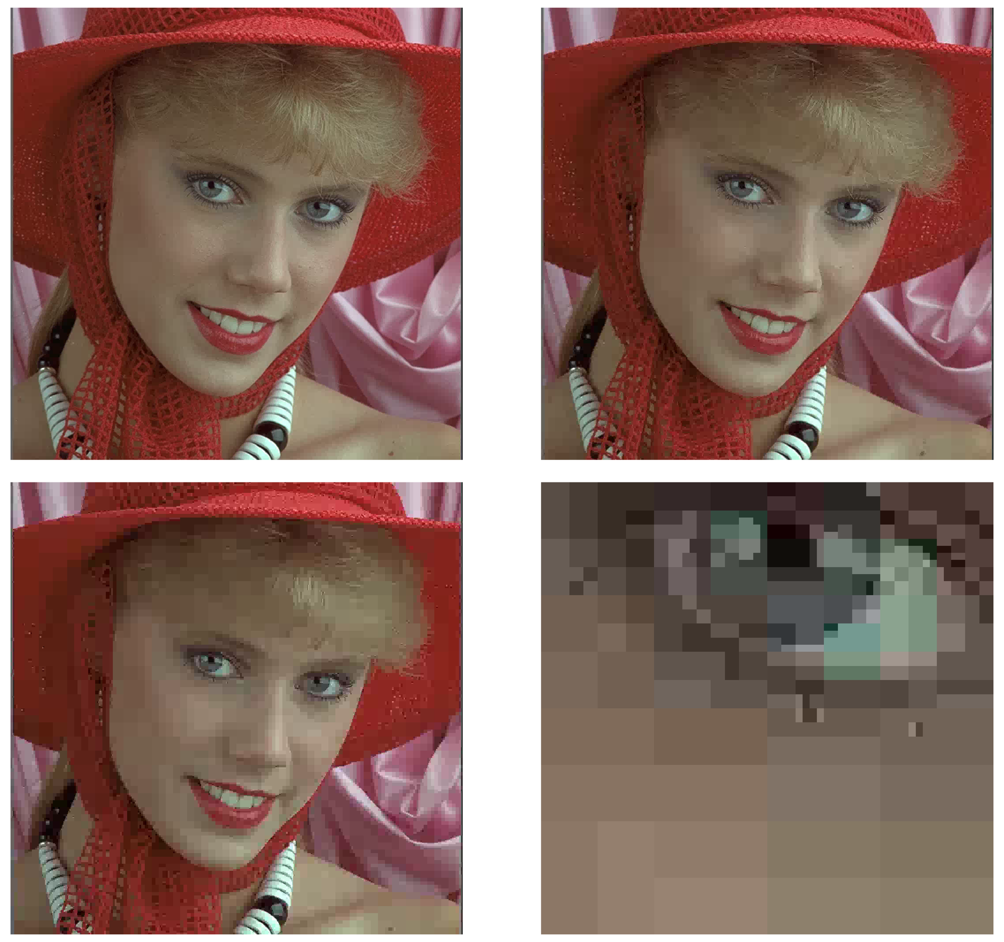


Figure 7. Compressed Kodak images: Compression Ratio=6 (top-left), Compression Ratio=7 (top-right), Compression Ratio=8 (bottom-left), and zoomed into eye of image with Compression Ratio=8.

November 2025

araujol@carleton.edu