Image Compression

EQ2330 Image and Video Processing, Project 2

Federico Favia
favia@kth.se

Yue Song
yuesong@kth.se

December 4, 2019

Summary

In this project, we implemented two transform based image compression algorithms, namely the discrete cosine transform (DCT) and the fast wavelet transform (FWT). The achievable performances of both transforms are evaluated and compared against each other.

1 Introduction

For the prevalent DCT-II type, the $M \times M$ transform matrix A can be defined as

$$A_{i,j} = a_i \; cos(\frac{(2j+1)i\pi}{2M}), \quad i,j = 0, 1, \dots M-1 \tag{1}$$

where a_i depends on the row number of transform matrix:

$$a_i = \begin{cases} \sqrt{\frac{1}{M}} & i = 0 \\ \sqrt{\frac{2}{M}} & i > 0 \end{cases} \tag{2}$$

As DCT/IDCT is an orthogonal transform, the DCT transform can be expressed as $y = Tx$ and the inverse DCT transform is similarly calculated as $x = T^T y$, where $T^T T = I$, *i.e.*, T has orthogonal columns.

A uniform quantizer splits the mapped input signal into quantization steps of equal size, thus avoid using threshold. Let Q denotes the quantization step, the output of a quantizer is given by:

$$x_Q[k] = Q \cdot \lfloor \frac{x[k]}{Q} + \frac{1}{2} \rfloor \tag{3}$$

where $\lfloor \cdot \rfloor$ is the loss function which maps a real number to the largest integer not greater than its augment. The characteristic curve is displayed in the Fig. 5. Without restricting input in amplitude, the resulting quantization indexes $x[k]$ are countable infinite. For a finite number of quantization indexes, the input signal has to be restricted to a minimal/maximal amplitude $x_{min} < x[k] < x_{max}$ before quantization.

In the classical implementation of Fast Wavelet Transform (FWT), we have two-channel filterbanks which deal with approximation coefficients and detail coefficients respectively. Fig. 1 shows the cascaded analysis and synthesis step. The reconstructed signal can be calculated as

$$\hat{x}(z) = \frac{1}{2}[h_0(z)g_0(z) + h_1(z)g_1(z)]x(z) + \frac{1}{2}[h_0(-z)g_0(z) + h_1(-z)g_1(z)]x(z) \tag{4}$$

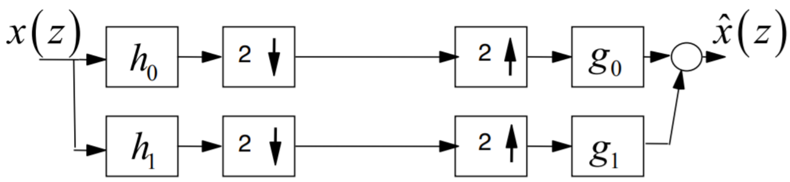

Figure 1: Decomposition step and reconstruction step of two-bank FWT implementation, taken from [2].

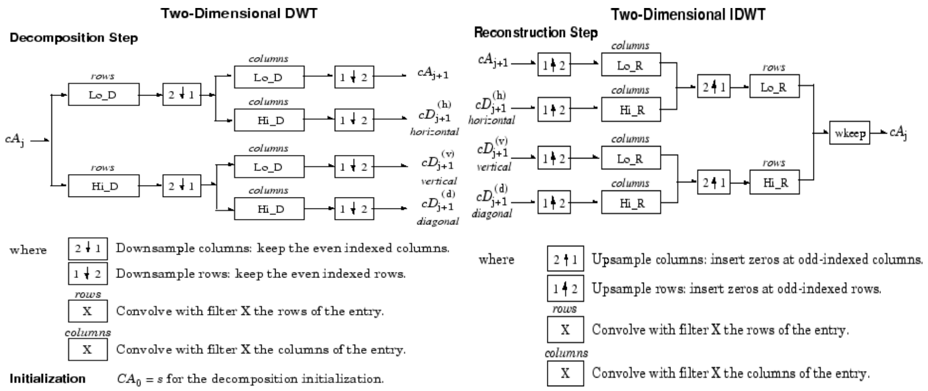

Figure 2: Two-dimensional FWT and inverse transform.

where $h_0(\cdot), h_1(\cdot)$ are low-pass and high-pass decomposition filters, and $h_0(\cdot), h_1(\cdot)$ denote low-pass and high-pass reconstruction filters. The second term in the rhs of Eq. 4 represents the aliasing effect. If the filters satisfy the conditions as

$$\begin{cases} g_0(z) = h_1(-z) \\ -g_1(z) = h_0(-z) \end{cases} \tag{5}$$

We can accomplish the conditions by quadrature mirroring filter (QMF) and flipping the filter in time/spatial domain:

$$\begin{cases} g_1(z) = QMF(g_0(z)) = g_0(-z) & \\ h_i(z) = -g_i(z) & i = 0, 1 \end{cases} \tag{6}$$

In the case of 2-D signal (Fig. 2), the FWT is just a recursive concatenation of 1-D FWT.

2 System Description

All the codes are implemented in Matlab 2019b. No additional library is further needed.

3 Result

3.1 Blockwise DCT and IDCT

We verified our DCT implementation with *Matlab* internal function and present the coefficient matrix in the Tab. 1. Some interesting conclusion can be drawn from the table. The row of odd index is symmetric, while the row of even index is anti-symmetric.

0.3536	0.3536	0.3536	0.3536	0.3536	0.3536	0.3536	0.3536
0.4904	0.4157	0.2778	0.0975	-0.0975	-0.2778	-0.4157	-0.4904
0.4619	0.1913	-0.1913	-0.4619	-0.4619	-0.1913	0.1913	0.4619
0.4157	-0.0975	-0.4904	-0.2778	0.2778	0.4904	0.0975	-0.4157
0.3536	-0.3536	-0.3536	0.3536	0.3536	-0.3536	-0.3536	0.3536
0.2778	-0.4904	0.0975	0.4157	-0.4157	-0.0975	0.4904	-0.2778
0.1913	-0.4619	0.4619	-0.1913	-0.1913	0.4619	-0.4619	0.1913
0.0975	-0.2778	0.4157	-0.4904	0.4904	-0.4157	0.2778	-0.0975

Table 1: The coefficient of 8×8 DCT transform matrix.

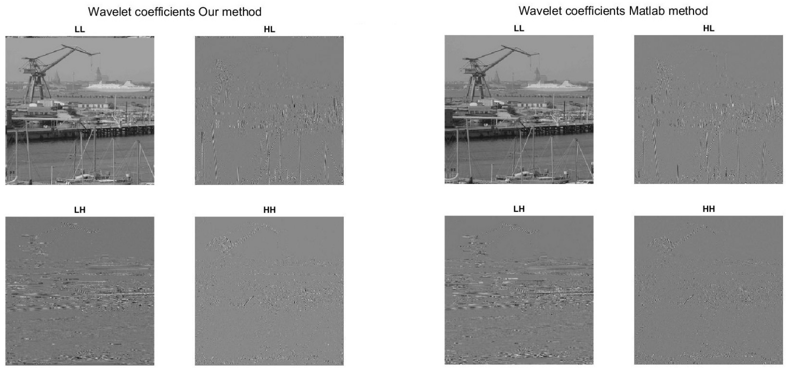

Figure 3: The wavelet coefficient matrix of our method (left) and MATLAB internal DWT implementation.

3.2 FWT-based Image Compression

To implement a Fast Wavelet Transform system, the one-dimensional two-band analysis and synthesis filter banks are fundamental. Our approach to build these blocks is based on direct implementation, with filters, down-sampling and up-sampling (Fig. 1), using a prototype scaling vector (e.g. Haar, Farras nearly symmetric, Daubechies 4, 8 and 10-tap filters...). The serial application of the analysis and synthesis filter banks to a random size 1-D input signal (of random values) produces in our case a reconstructed signal equal to the input down to machine precision (our code checks zero mse).

As for two-dimensional FWT and inverse, it is implemented by recursive application of the analysis and synthesis filter banks. Using the Daubechies 8-tap filter, we plot the wavelet coefficients for scale 4 of the image `harbour.tif` in Fig. 3. In the same figure we plot the obtained wavelet coefficients by the official MATLAB implementation `dwt2()`, showing that in our implementation a borders effect can be noticed. Note the official MATLAB implementation increases the dimension of FWT output signal (half of the scaling vector length). The boundary is well kept in the horizontal and vertical approximation part. We insisted using the original output dimension. Although a slight boundary effect exists on the top of horizontal and vertical approximation part, we still get a perfect reconstruction signal.

In Fig. 4, we demonstrated that our reconstructed image is both visually and metrically (MSE checked by our code) equal to the original input, in the same way of the official inverse FWT implemented by MATLAB.

Figure 4: Visual comparison of reconstruction image by our implementation and MATLAB function. Both methods have zero mean square error (MSE).

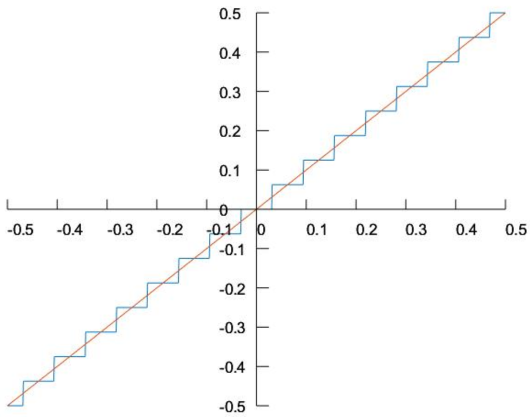

Figure 5: Mid-tread uniform distribution. The blue line represents the quantization curve, and the red line connects all the mid-treads.

3.3 Uniform Quantizer

A uniform quantizer splits the mapped input signal into quantization steps of equal size. Therefore, the mid-treads lie on the same line. We display the mid-tread quantization characteristics curve in the Fig. 5.

3.4 Distortion and Bit-Rate Estimation

If no quantization or any noise occur during DCT/IDCT transform, the reconstructed image can be calculated as

$$\hat{x} = T^T y = T^T T x = I x = x \tag{7}$$

where T is the DCT transformation matrix and I is the identity matrix. However, quantization process introduces distortion, we can only get the reconstructed image as $T^T y_q$. The MSE of reconsctructed image can be represented as

$$\begin{aligned}(\hat{x} - x)^T(\hat{x} - x) &= (T^T y_q - T^T y)^T (T^T y_q - T^T y) \\ &= (y_q - y)^T T T^T (y_q - y) \\ &= (y_q - y)^T (y_q - y)\end{aligned} \tag{8}$$

Hence, the mean square error of reconstructed image is equal to that of transform matrix. In the perspective of energy conservation, the only energy loss is from quantization. When evaluating PSNR or MSE of reconstructed image, it is

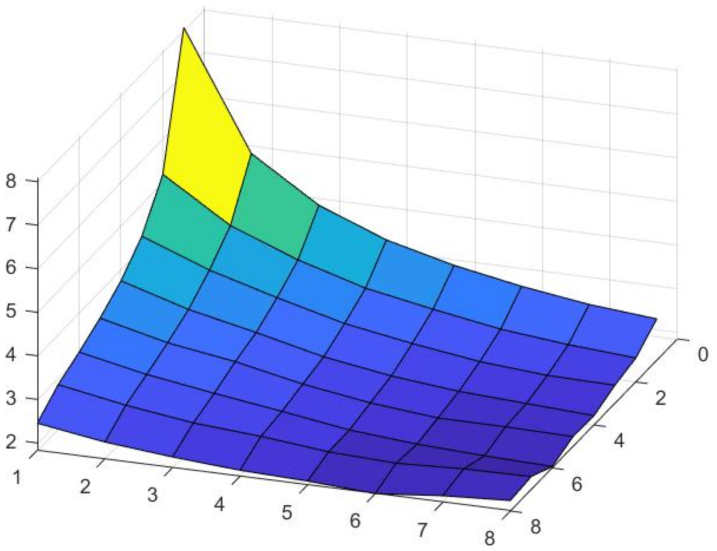

Figure 6: Average entropy of all 8×8 DCT coefficients in all the images. A quantization range of 2^9 is adopted.

sufficient to evaluate the quantization error. Since wavelet is also an orthogonal transform , the proof applies for DWT.

Compared with DWT, DCT owns the advantage of compacting most energies in the first few coefficients. It is in particular obvious in 2-D signals, where the DC components and first few coefficients have large values, and other AC components tend to zero. The first few coefficients convey most information, as shown by the entropy curve in Fig. 6. In the case of DWT, the entropy has a more dispersed distribution indicated in Fig. 7.

In Fig. 8, we varied the quantizer stepsize over the range $2^0, \ldots, 2^9$ and measure the PSNR versus bit-rates curve using the images *pepper, harbor, and boat*. In general more bits is utilized, the less distortion, *i.e.* higher PSNR value can be observed. As expected, DCT outperforms FWT by a large margin when having limited bits. This is because DCT is closer to idal KLT transform and has a more energy compact transform, thus becoming a mainstream for transmission. When number of bits increase, FWT outperforms DCT, mainly due to the precision and larger transform matrix. It is also worth to mention the gain in the performance is about 6 dB per pixel, which is confirmed by Shannon's Information theory [4]. If it is a Gaussian source, the MSE distortion gain will be -6 dB.

Quantization range	2^0	2^1	2^2	2^3	2^4	2^5	2^6	2^7	2^8	2^9
DCT MSE	45.89	4.55	0.74	0.19	0.06	0.01	5.92×10^{-4}	2.77×10^{-4}	1.29×10^{-4}	2.57×10^{-5}
DCT bits	0	0.01	0.03	0.11	0.46	0.50	0.62	1.16	2.11	2.80
FWT MSE	3.44	0.67	3.77×10^{-2}	5.07×10^{-3}	1.16×10^{-3}	9.07×10^{-4}	1.25×10^{-4}	4.14×10^{-5}	9.89×10^{-6}	2.49×10^{-6}
FWT bits	0	0.71	0.72	1.01	1.24	1.71	1.81	2.33	2.71	3.05

Table 2: The MSE error and bitrates for different quantization range of both DCT and FWT transform.

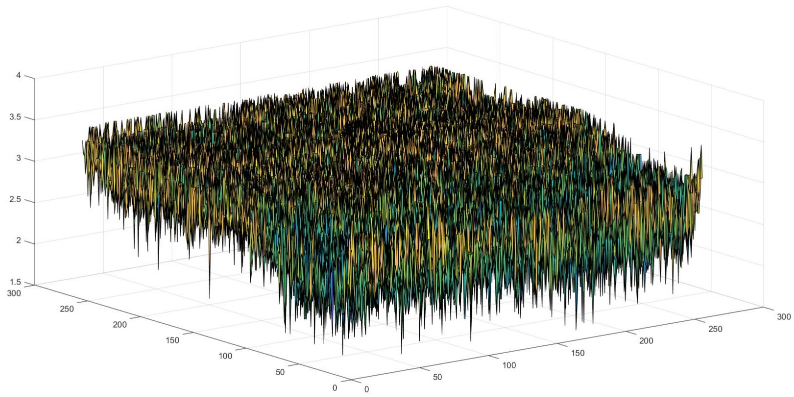

Figure 7: Average entropy of all 64×64 FWT coefficients in all the images. A quantization range of 2^9 is adopted.

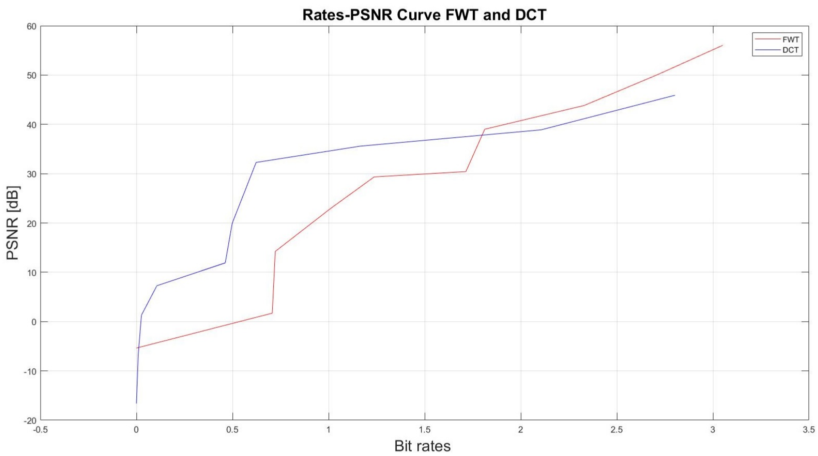

Figure 8: PSNR versus bit-rates curve of DCT and FWT transform.

4 Conclusion

In the project, we implemented both DCT and FWT transform and accordingly inverse transform, two of the most widely used techniques in the image compression. In particular for FWT, experimenting to avoid borders effect was not a trivial part, but in the end we obtained a perfect reconstructed signal. Experimental results demonstrated for orthogonal transform, the reconstruction error comes from quantization process. Moreover, DCT proved to be a more energy compact transform, packing the most information in fewest coefficients. It also minimizes the block-like appearance called blocking artifact that results when boundaries between sub-images become visible. Advantages of Wavelet Transform over the DCT can be observed too: firstly, it allows good localization both in time and spatial frequency domain. There is no need to divide the input coding into non- overlapping 2-D blocks, and it avoids blocking artifacts with higher compression ratios. As we can see in Table 2, for each quantization step the MSE is better in the case of FWT even though it is worth to mention that DCT has better PSNR with limited number of bits. Finally the wavelet transformation of the whole image introduces inherent scaling.

Appendix

Allocation of responsibilities

The project was done collaboratively by both authors.

References

[1] Rafael C. Gonzalez and Richard E. Woods, *Digital Image Processing*, Prentice Hall, 2nd ed., 2002

[2] M. Flierl. *Image and Video Processing [EQ2330], Course Slides, KTH, 2019*

[3] Wang, Zhou and Bovik, Alan C and Sheikh, Hamid R and Simoncelli, Eero P and others, *Image quality assessment: from error visibility to structural similarity*, IEEE transactions on image processing, Volume 13, 600-612, 2004

[4] Shannon Claude Elwood, *A mathematical theory of communication*, Bell system technical journal, Volume 27, 379-423, 1948