University of Cambridge

Engineering Tripos Part IIA

Girton College

SF2 Image Processing
Final Report

Clinton N. Igwegbu (cni22)

June 04, 2019

Project Summary

Each stage of this project has led up to the design of an image compressor and decompressor in the competition on Monday June 3rd 2019. The first stages of this project introduced a series of energy compaction techniques: the Laplacian Pyramid scheme, the Discrete Cosine Transform, the Lapped Bi-Orthogonal Transform and the Discrete Wavelet Transform. The resulting tranformed images were then quantised and an estimate of the compression ratios was found by calculating the entropy of each image. In the later stages of the project, practical run-length and Huffman coding techniques were explored to give real values of the compression ratios achievable. Additionally, centre-clipped quantisers were discussed as a means of improving compression. All the techniques learnt served as a basis for the design for the competition.

1 Introduction

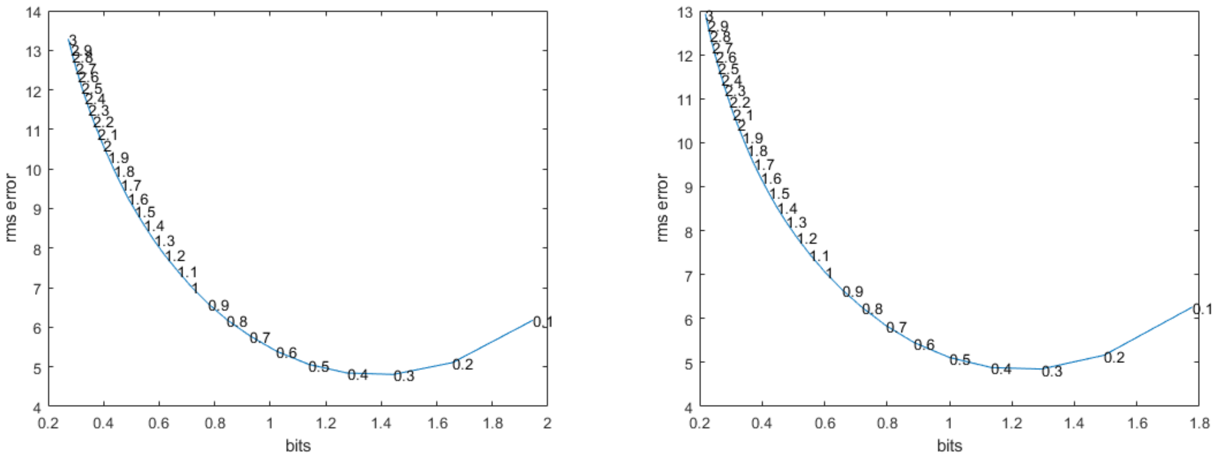
The first stages of this project introduced a series of energy compaction techniques: the Laplacian Pyramid scheme, the DCT, LBT and DWT. These techniques revealed image data in the frequency domain where the partitioning of the images energy could be more clearly the energy can be seen to be compacted at lower frequencies. From this point on the transform is quantised so that it can be encoded. This process leads to a smaller range of values that each pixel can take on and therefore compresses the data. Compression can be achieved more smartly by widening the quantisation step size around the 0-pixel values i.e. by varying the first rise. This process is known as centre-clipping and is examined in the first section of this report. Throughout the course of this project, entropy has been used as an estimate of the number of bits that an image can be compressed to. This report introduces practical encoding techniques: Huffman encoding of run-length-amplitude encoded values. Energy compaction, quantisation and encoding make up the three components necessary to build an image compression system. The first of these was explored in the two interim reports. Centre-clipped quantisation and encoding are explored in the Theory section of this report.

The report then details the work that went into the design of an image compressor for the June 3rd competition. It explains the designs considered, and the rationale behind choosing the final design.

2 Theory: Quantisation and Encoding

2.1 Centre-clipped Quantisers

It was noted that most of the energy of an image lies around the zero frequency i.e. dc. The entropy of an image is the minimum possible number of bits it can be compressed to while maintaining 100% reconstruction quality. A uniform distribution has the highest possible entropy as there is more uncertainty on what value each pixel takes. By making the first rise of the quantiser large, more pixels are quantised to 0 so that there is less uncertainty in the value of the quantised pixels. Therefore centre-clipped quantisers are used in the hope of increasing compression rates of images. However it must be noted that while making the first rise very large would produce a quantised image with very low entropy, it would also generate a reconstructed image that has greater error relative to the original. Plots of RMS reconstruction error versus the number of bits in a quantised image are shown in figure 1 below for the DCT and LBT schemes.



(a) DCT

(b) LBT

Figure 1: Reconstruction error vs Bits

It can be seen in all cases the plot takes on a parabola-like form with a minimum reconstruction error occurring at about a 0.4 rise (ratio between first rise and step size). As such there is a trade-off between minimising the RMS recontstruction error and maximimising compression. This trade-off must be considered while designing a suitable compressor. Many quantisers use a rise of 1. Across the energy compaction schemes seen in this project, tests with a rise of 1 led to a 40% increase in compression ratios compared to uniform quantisers. However a 40% in the RMS reconstruction error was also seen.

2.2 Combined Run-Length/Huffman Coding

While entropy served as an estimate of the number of bits needed to encode an image, a more practical scheme was needed. First the transformed images were run-length encoded. A string of image coefficients were therefore encoded as a non-zero coefficient with the number of zeros that preceeded it. This formed a single event. JPEG limits the number of zeros in a run-length to 15. Additionally it encodes amplitudes of non-zero coefficients as a base 2 logarithm rounded up to the nearest integer. Including the end-of-block (EOB) codeword and the run-of-16 codeword (ZRL) leads to 162 possible events.

Each event acts as a distinct symbol with its own probability of appearing in the image. A Huffman code works by assigning more probably symbols shorter code lengths. Since more probable symbols occur more often in an image, giving them shorter code lengths means that the whole image can be represented with a short code, thus leading to good compression ratios. JPEG performs a Huffman code on DCT compacted images. The standard JPEG Huffman code tables essentially assign each symbol a standardised pre-selected probability. It is also possible to tailor Huffman codes uniquely to each image such that the probability assigned to each symbol matches the distribution of symbols in that particular image. This would offer better compression rates but is more time-consuming.

At this point all fundamental image compression theory has been explored. The next section details the processes that went into designing a compressor for the competition.

3 Competition Design

The main design considered was a multi-stage LBT JPEG encoder. It was decided that I would implement the multi-stage LBT energy compaction scheme and that my teammate David de Oliveira would implement the scanning strategy and Huffman encoder compatible with this energy compaction scheme. The Multi-Stage LBT is detailed below.

3.1 Multi-Stage LBT

The DCT is the energy compaction scheme typically employed in JPEG. However, DCT is performed on adjacent image blocks and does not take into account correlations between image blocks. This has a tendency to leave 'blocky' artefacts in image reconstructions. The LBT overcomes this effect, and it was also seen to offer better compression rates in the second interim report [2]. The LBT used in this project consists of performing a photo-overlap transform (POT), to process correlations between blocks, followed by a DCT step which leads to blocks of frequency coefficients. These coefficients can then be regrouped so that coefficients for the same frequencies form sub-images. This regroup stage is an important part of the multi-stage LBT described below.

This design consisted of performing successive LBT transformation on the dc coefficients. In the first stage, the LBT is performed as normal, then the result is regrouped. The top left sub-image (DC coefficients) forms the basis of the next stage. The LBT is performed on this sub-image and then it is regrouped, just like the first stage. These stages are repeated iteratively until the LBT block size, N, matches the sub-image size. The result is a set of nested grids. This is illustrated in figure 2 using a block size of 4
.

(a) After 1st stage

(b) After 2nd stage

Figure 2: Nested Structure of Multi-Stage LBT

The decoder essentially reverses the process described above. So starting at the bottom and working upwards: the regroup process is reversed then the inverse LBT is performed. Going up one level this process is repeated until the original image is obtained.

A conservative estimate was made on the potential compression ratios achievable using this energy compaction scheme. This estimate was made by crudely quantising every LBT level with the same step size and using the function bpp.m to find the entropy of each image. Quantising with a step size of 17 gave compression ratios as high as 4.7 while keeping rms reconstruction errors as low as 0.2.

Though entropy cannot be achieved by practical encoders, the estimate given was still conservative as it did not vary the quantisation step size between levels and also it did not make use of the function dctbpp.m to calculate the entropy. Remember that dctbpp.m calculates the entropy of each sub-image in a regrouped image and then takes the average. Due to greater correlations between sub-images, this gives a lower entropy than that calculated by bpp.m. The multi-stage LBT JPEG encoder therefore had great potential but due to time constraints the accompanying scanner and Huffman encoder could not be implemented for this process.

3.1.1 Testing the LBT

Though a suitable scanner and encoder could not be implemented for this energy compaction scheme, there were attempts at using the scanning strategy already in place in jpecenc.m, while replacing the DCT step in this file with the multi-stage LBT step. While the test results indicated good compression ratios, of the order of 4.8, the RMS reconstruction errors were of the order of 12 and the subjective visual quality was deemed to be poor as shown in the figure below. It can be seen that this technique did not take kindly to high frequency features, hence the poor performance in the reconstruction of the tiny twigs and leaves in the image shown. These tests made use of the standard Huffman coding tables.

(a) Original (b) Reconstructed

Figure 3: Multi-LBT on Bridge (step=17)

3.2 Final Design and Testing

While preparing for the competition, simpler designs were made and evaluated. One of these simply involved changing the quantisation step of the jpegenc.m file provided. This involved using the standard JPEG luminance quantisation tables together with an 8x8 DCT compaction scheme. Upon compressing the test images to the required size of 5 KB, the RMS reconstruction error using this design was found to be 12.57. The reconstructed image was seen to be somewhat blocky and blurry, resulting in the loss of fine features such as the fishing line in the 2018 competition image and the partition of the monkey's lips in the 2019 competition image as shown in the figure below. Also notice that the bushed behind the monkey has formed patchy blotches.

(a) Original

(b) Reconstructed

Figure 4: 2019 competition image

3.3 Further Possible Improvements

As mentioned in section 3.1, the multi-stage LBT had great potential and a system centred around this scheme can be improved in three ways:

- Parallel scanning strategy:

  The scanning strategy in jpegenc.m goes goes left to right and top to bottom from one set of frequency coefficients to the next such that the frequencies are gradually increasing as the scanning strategy proceeds. This same scanning strategy can be used in the multi-stage LBT, except running in parallel for different stages.

- Non-uniform quantisation across stages

  The range of frequencies in each stage would be different. Therefore the quantisation step size required to filter out the higher frequencies of each stage are different. Therefore by varying the step size across stages, much of the high frequency data can be discarded. Since the human eye does not detect high frequencies too well then visual quality can be maintained to a high level.

- Non-uniform quantisation within a stage

  Within a single stage, frequency-dependent quantisation is possible. This would involve quantising lower frequency data more finely than higher frequency data, so that more low frequency information can retained while still maintaining good visual quality.

4 Project Conclusions

Energy Compaction:

- The Laplacian Pyramid scheme was disadvantageous as it increased the sample size of images.
- DCT maintained sample size and offered good compression ratios but generated 'blocky' artefacts in image reconstructions.
- LBT offered the best compression ratios out of all the schemes seen and ameliorated the blockiness seen in DCT methods.

Quantisation:

- Centre-clipped quantisation offered 40% increases in compression ratios but also 40% increases in RMS reconstruction error.
- Frequency-dependent quantisation is a more efficient way of removing unnecessary high-frequency content form an image, while achieving good compression ratios.

Encoding:

- Scanning strategy greatly determines how closely an encoding strategy can approach entropy.

Competition Design:

- Multi-stage LBT compaction scheme gave good entropy-based compression estimates up to 4.7 however there was not enough time to implement a good accompanying scanning strategy.
- Multi-stage LBT scheme with quantisation scheme and 'zig-zag' scanning strategy used in jpegenc.m produced images with poor visual quality.
- DCT compaction scheme with standard JPEG luminance quantisation tables produced acceptable visual quality in the competition.

5 References

- Project Handout

  https://www.vle.cam.ac.uk/mod/resource/view.php?id=805731

- Second Interim Report

  https://github.com/ClintonIgwegbu/Image-Processing/blob/master/SF2_cni22_report2.pdf