SF2 Image Processing - First Interim Report

Kingson Chan - ksc37

1 Introduction

Image data compression systems are widely used in various modern technologies and devices, with the aim to compress the size of the data file required to store an image as much as possible while preserving the quality of the image to an acceptable level. In this project, we will investigate the design of three main processes in an image compression system: the input filtering process, the quantisation process, and the coding process.

In this first interim report, we will focus on basic image filtering, the use of Laplacian Pyramid for energy compaction, and basic schemes for quantisation. An image of a lighthouse will be used throughout our investigation to test the effects of various image processing techniques.

2 Simple image filtering

We use a sampled half-cosine pulse as an image low-pass filter. The impulse response $h(n)$ is given by

$$h(n) = G\cos\left(\frac{n\pi}{N+1}\right) \qquad \text{for } \frac{-(N-1)}{2} \leq n \leq \frac{N-1}{2} \tag{1}$$

where G is the gain factor. For unity DC gain, G should be calculated such that the components of the FIR filter sums to one.

Using a half-cosine filter of length 15, the image is filtered in the row direction, column direction, and then in both directions. The image becomes blurred in each direction and there is a gradual fade to black at the edges caused by the convolution treating the signal as zero outside the range of the input vector, as shown in fig 1.

The process of filtering row and column separately is known as separable 2D filtering. This is much more computationally efficient than general non-separable 2D filtering. The maximum absolute pixel difference between a row first then column filtered image, and a column first then row filtered image is found to be 1.14×10^{-13}. This is insignificant and most likely error due to the handling of very small floating point numbers by MATLAB. This demonstrates that there is no difference whether the rows or columns are filtered first.

In order to minimise the edge effects, the technique of symmetric extension is then used, where it assumes that the image is surrounded by a flat mirror along each edge. The filtered image extends symmetrically in all directions with the same period as the original image. A high-pass filter can be obtained by subtracting the low-pass result from the original image. As shown in fig 2, low-pass and high-pass filtered images are generated using 3 different odd-length half-cosine filters with the symmetric extension technique applied. It can be seen that the black edge artifacts are no longer present. A short length filter creates a weakly blurred low-pass image and faint edges in the high-pass image, compared to a longer length filter which creates a very blurry low-pass image and very pronounced edges in the high-pass image.

A metric to assess filtered images is by comparing it's energy content, defined as the sum of the squares of the individual pixel values. It is found that low-pass images have much higher energy content compared to high-pass images. In the case of a 15 sample half-cosine filter, the low-pass image has energy 1.26×10^9 and the high-pass image has energy 4.80×10^7. The low-pass image energy content is two orders of magnitude larger than that of the high-pass image.

3 Laplacian pyramid

By exploiting the fact that low-pass images have much higher energy and that the low-pass image is much lower bandwidth than the original image, we can split the original image into smaller and smaller pairs of high-pass and low-pass filtered images through the process of sub-sampling. The image can be perfectly reconstructed using the smallest low-pass image and all of the high-pass images. A 4-layer pyramid for the lighthouse is shown in fig 3. By creating a Laplacian pyramid, we have achieved image compression if the decimated and filtered images can be transmitted with fewer bits than the original image. This will usually be the case since the filtered images contain much less energy than the original image.

We will be using a 3-tap filter $\mathbf{h}$ with coefficients $\frac{1}{4}[1\ 2\ 1]$ and a longer 5-tap filter $\mathbf{h}$ with coefficients $\frac{1}{16}[1\ 4\ 6\ 4\ 1]$, which has a lower cut-off frequency. We will investigate and quantify the effects of using different number of layers in the pyramid in the next section.

4 Quantisation and coding efficiency

The entropy represents the minimum average number of bits per sample needed to code samples with a given probability distribution $p(i)$. Using a 1-layer pyramid scheme with a quantiser of step size 17, the number of bits required to encode $\mathbf{X}$ is 2.29×10^5, compared to only 1.83×10^5 bits required to encode both $\mathbf{X1}$ and $\mathbf{Y0}$. There is diminishing returns with increasing layers, reaching a plateau of about 1.37×10^5 bits after 4 layers.

As shown in fig 4, using more layers to reconstruct the output image results in greater RMS error. This is due to the quantisation error being magnified and propagating through the layers when the image is reconstructed. The quantised original image has an RMS error of 4.934, which acts as a reference lower bound for the RMS error using a quantiser of constant step size 17. By inspecting the images in fig 5, we can see that the images become very pixelated due to the limited grayscale colours available, and becomes progressively blurry as the number of pyramid layers increase. In the reconstructed image for 7 layers, the different shades of colour in the sky can barely be differentiated, showing the extent of loss of detail in the image.

Since image quality is reduced by quantisation, coding schemes can only be compared when they produce similar quantisation error. We will use the compression ratio to compare the performance of various schemes, defined as the ratio between the total bits required for the reference scheme and the total bits required for the compressed scheme. The reference scheme is chosen to be the direct pixel quantisation method with its quantiser adjusted to give the same RMS error as the compressed scheme.

4.1 Constant step size

Under the constant step size scheme, we find the quantiser step size needed to achieve similar error as using direct quantisation with a step size of 17. As shown in fig 6a, the optimal step size decreases with increasing layers but seems to reach a plateau. From fig 6b, we find that the compression ratio peaks at two layers for both the 3-tap and 5-tap filter. The optimal 3-tap filtered image has a better compression ratio of 1.40 compared to 1.35. However, we can see from fig 7 that the 5-tap filter produces less artifacts and appears smoother, whereas the 3-tap filter is more pixelated, especially across the sky. This is expected as the longer filter has a lower cut-off frequency so the image becomes blurrier, but the higher energy in the high-pass is able to compensate and maintain the details in the edges for a clear overall image.

4.2 Equal MSE criterion

The quantiser step sizes can vary for different levels of the pyramid. In the equal MSE scheme, step sizes are chosen such that quantisers in each layer contribute equally to the Mean Squared Error of the reconstructed image. We choose the ratio of step sizes for each layer such that an impulse of that size will contribute the same energy to the decoded image at each layer. Similar to the constant step size scheme, we find the set of quantiser step sizes which produces similar RMS error as the reference scheme.

From fig 6b, we find that there is a jump in compression ratio from using one to two pyramid layers, but quickly levels off past 2 layers for both the 3-tap and 5-tap filter. By visual inspection, we can see that image quality decreases as the number of layers increase. Hence we will choose the two layer depth as optimal. Similar to the constant step size scheme, the 3-tap filter has a better compression ratio of 1.57 compared to 1.44 by the 5-tap filter. From fig 7, we find that the 3-tap and 5-tap images look very similar, with the 5-tap image looking marginally smoother. However, I think that the slightly improved image quality is not sufficient to compensate for the much higher compression ratio by the 3-tap filter.

Overall, the equal MSE scheme drastically outperforms the constant step size scheme. The equal MSE scheme has a much higher compression ratio when using the same pyramid depth and filter. The image quality is also much better in the equal MSE scheme. The difference is especially pronounced with the 3-tap filter, where the image compressed with the constant step size scheme looks very pixelated with a patchy sky and many artifacts.

Appendix

(a) Original image (b) Row filtered (c) Column filtered (d) 2D filtered

Figure 1: Low-pass filtering with half-cosine filter of length 15

(a) Length 5 (b) Length 15 (c) Length 25

Figure 2: Low-pass (top) and high-pass (bottom) filtering with half-cosine filter of various length

Figure 3: 4-layer Laplacian Pyramid for Lighthouse image

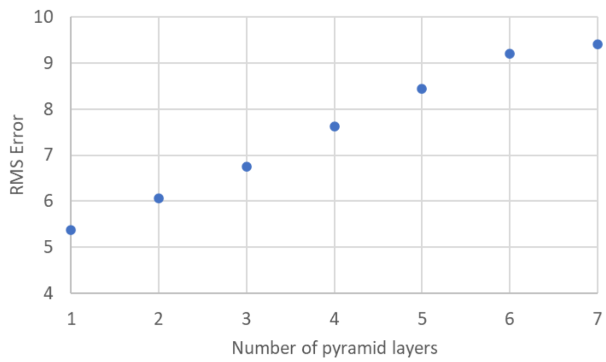

Figure 4: RMS error for constant quantisation step size

(a) Original quantised

(b) 1 layer

(c) 4 layers

(d) 7 layers

Figure 5: Quantised Laplacian pyramid reconstruction images

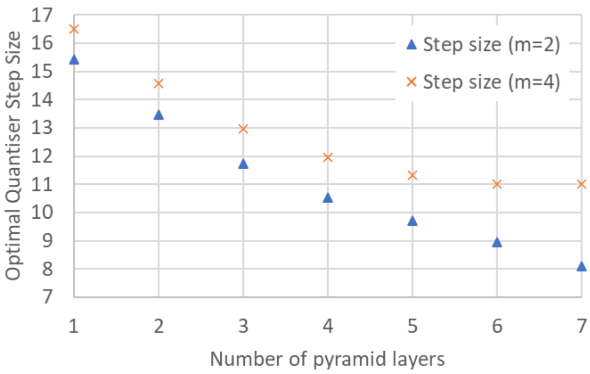

Figure 6a: Optimal quantisation step size

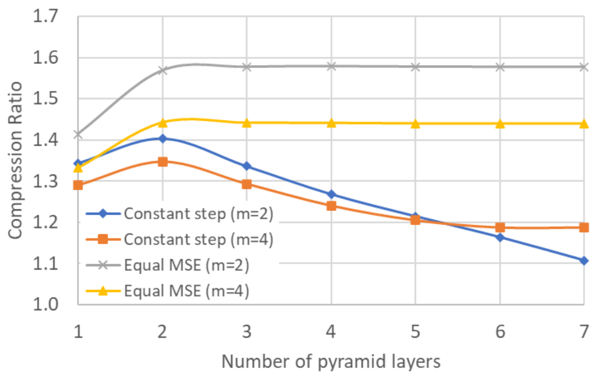

Figure 6b: Compression ratio over different layers using various schemes

Constant step size
m=2, 2 pyramid layers
Compression ratio = 1.404

Constant step size
m=4, 2 pyramid layers
Compression ratio = 1.347

Equal MSE
m=2, 2 pyramid layers
Compression ratio = 1.568

Equal MSE
m=4, 2 pyramid layers
Compression ratio = 1.442

Figure 7: Optimal decoded images for various compression schemes