SF2: Image Processing Interim Report 2

1. INTRODUCTION

After looking at quantisation and the Laplacian pyramid, some advanced topics are covered in this report that build on to some extent from the Laplacian pyramid. These are the Discrete Cosine Transform, the Lapped Biorthogonal Transform and the Discrete Wavelet Transform. The performance of all these schemes is investigated in both qualitative and quantitative terms.

2. DISCRETE COSINE TRANSFORM (DCT)

The DCT operates on non-overlapping blocks of the image with a linear, reversible transform process defined by:

$$y(k) = \sum_{n=0}^{N-1} C_{kn} x(n) \text{ where } C_{0n} = \frac{1}{\sqrt{N}} \text{ and } C_{kn} = \sqrt{\frac{2}{N}} \cos\left(\frac{kn\pi + {}^{k\pi}/_{2}}{N}\right)$$

Both this and its inverse can be vectorised and applied as matrix operations. The DCT applied to an image in 8 by 8 blocks is seen in the Appendix as Figure 1 after the sub images have been regrouped. Most of the energy is found in the low pass images towards the top left. As the image is traversed, the energy reduces. This can be visualised using the `imagesc` function in Matlab. The redder the colour, the higher the energy content and the bluer, the lower the energy. This can be seen in Figure 2.

The DCT can also be thought of as consisting of basis functions, a superposition of whom gives each block in the image. In order to visualise these the bases are displayed in the Appendix as Figure 3. It can be seen that the bases in the top left correspond to low frequency components and towards the bottom right, they correspond to the higher frequency components.

This coding of images in blocks combined with quantisation achieves some degree of data compression. If the entropies of sub images are calculated after quantisation to a step size of 17 and multiplied by the number of pixels in the image, an estimate can be obtained of the number of bits required to transmit the image. This value is found to be 97468.5 compared to the otherwise 109746.35 required if the image is not coded in blocks. This is because the probability distribution of the pixels in the sub blocks is different from the original image and this can be exploited in the coding.

If the original image is reconstructed from the quantised image (ensuring that the quantisation step size gives the same RMS error as direct quantisation), the image obtained can be seen in the Appendix as Figure 4. Using different block sizes leads to slightly different results.

Table 1. Performance for DCT using different block sizes (4 s.f)

Block Size	**Compression Ratio**	**Max deviation**
4	2.706	26.00
8	2.997	32.03
16	2.678	33.87

NB. If the DCT image is quantised to a step size of 17 and then reconstructed, the compression ratio is 2.347 and the max deviation is 21.74.

The visual differences in the image (Figure 4) are mainly seen in the sky where a large block size seems to give block artefacts. A block size of 8 seems smoothest. This seems close to the original image and much better than the directly quantised image.

In the limit that images are coded such that each pixel is coded individually and entropy calculated individually, the number of bits required tends to 0. However, this is not a reasonable result because the entropies are calculated based on the probability distribution in a block. For very small images this would not yield reasonable results as the probability distribution would not be plausible for very small image sizes.

It seems that the best transform size for the lighthouse image is 4 to keep the deviation low but 8 to keep compression ratio high. Repeating a similar analysis for the bridge image which has considerably higher detail, it seems that a transform size of 8 gives better performance.

3. LAPPED BIORTHOGONAL TRANSFORM (LBT)

The DCT does not take into account any correlation between blocks in the image. The LBT can be used to overcome that. An LBT can be thought of as pre-filtering the image using a Photo Overlap Transform (POT) before applying the DCT. This can also be achieved using matrix operations.

The degree of bi-orthogonality or the scaling factor can be changed while applying the LBT. The compression ratio for the reconstructed image changes depending on the scaling factor used. For a block size of 8 for the DCT part of the process, a graph showing compression ratio against this scaling factor is shown in the Appendix as Figure 5. The scaling factor that maximises the compression ratio is 1.37. In order to ensure that the RMS error matches the quantisation step size is calculated for each scaling factor. The graph showing the variation in the step size for the same RMS error against scaling factor is shown in the Appendix as Figure 6.

The basis functions for the POT can also be visualised. It can be seen (in Figure 7) that as the scaling factor increases from 1 to $\sqrt{2}$ the image becomes brighter but also grainier with block artefacts. After this, from $\sqrt{2}$ to 2, the image becomes more and grainier with more and more block artefacts.

The effect of using different block sizes can also be investigated. The performance features are listed below.

Table 2. Performance for LBT for different block sizes (4 s.f) (Scaling factor = 1.37)

Block Size	Compression Ratio	Max deviation
4	3.067	25.70
8	3.430	25.40
16	2.573	22.70

The reconstructed images are seen in the Appendix as Figure 8. There is not that much difference in the visual quality of images as is evident in the small difference in max deviation of the reconstructed image. There are some small differences in the smoothness of the sky and the quantisation of the house. There are also some minor differences in the brightness of the sky.

4. DISCRETE WAVELET TRANSFORM (DWT)

The DWT combines some of the best features of the Laplacian pyramid and the DCT. They employ banks of band pass filters and construct a wavelet tree, which can be used to compress and reconstruct the image.

Applying the LeGall filters once to the image X to produce a low pass and high pass image produces a result like the one seen in Figure 9 in the Appendix. The standard deviation of the 2 images is as follows: 49.90 for the low pass image and 10.33 for the high pass one. Again, most of the energy is in the low pass image.

When the LeGall filters are applied to the images to produce 4 images as is seen in Figure 10 in the Appendix, it can be seen that they extract (from top left clockwise) the DC component, vertical edges, high frequency component and horizontal edges.

The DWT can be applied recursively to an image by applying it to the top left image at each stage and constructing it to an arbitrary number of levels. By quantising this image, data compression can be achieved. Quantisation can be done in two ways: uniformly or using the Mean Squared Error criterion where each sub image is quantised such that it contributes an equal amount to the total error.

The compression ratio achieved under both schemes for different levels of the DWT can be seen graphically in Figure 11. From this figure it can be seen that it would be best to use 5 levels under a non-uniform quantisation scheme. If the bridge image is considered, a different graph is obtained which suggests that the non-uniform scheme is not as good as the uniform scheme up to about 5 levels. This suggests that for images with higher detail it is better to use fewer levels and a uniform step. The flamingo image produces a shape similar to the lighthouse image but also suggests that the uniform quantising with fewer levels is better.

The uniform quantised images have a slightly higher gain in the sky. The non-uniformly quantised ones seem to be smoother and vary relatively little regardless of the number of levels used. If only 1 level is used, the image has block artefacts and seem a little patchy. The more the levels used, the smoother the image seems.

5. CONCLUSION

From the analysis above, it seems that all three methods provide some improvement on simply using a Laplacian pyramid. They offer similar max deviation errors but significantly higher compression ratio in some cases. Of the three transforms discussed, the LBT seems to give the best compression ratio but may be expensive to evaluate for high resolution images, although not as much as the wavelet transforms. The visual appearance does seem to be best for the wavelet transform but not by that much.

6. APPENDIX

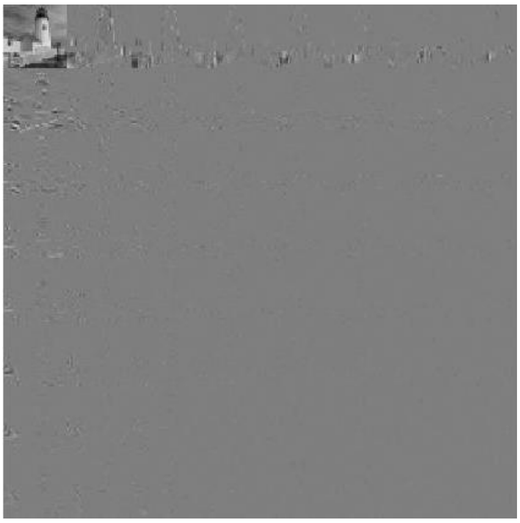

Figure 1. DCT applied to image. Regrouped images.

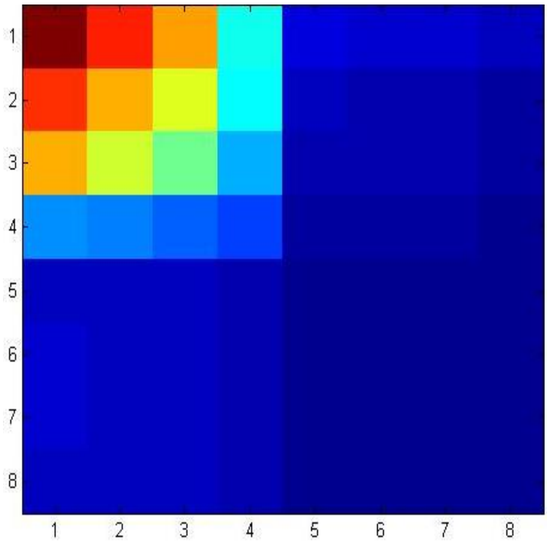

Figure 2. Energy of sub images

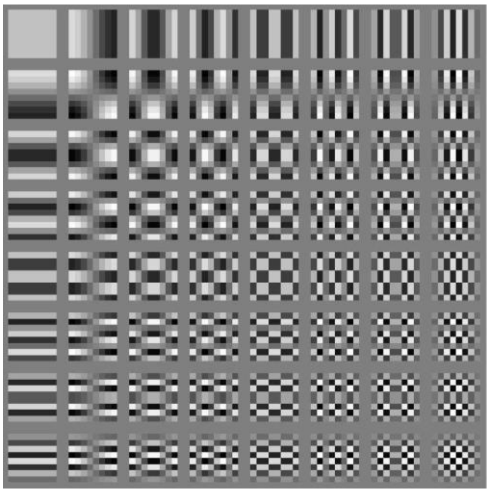

Figure 3. Bases for DCT

Figure 4a. DCT Reconstructed image using block size of 8

Figure 4b. DCT Reconstructed image using block size of 4

Figure 4c. DCT Reconstructed image using block size of 16

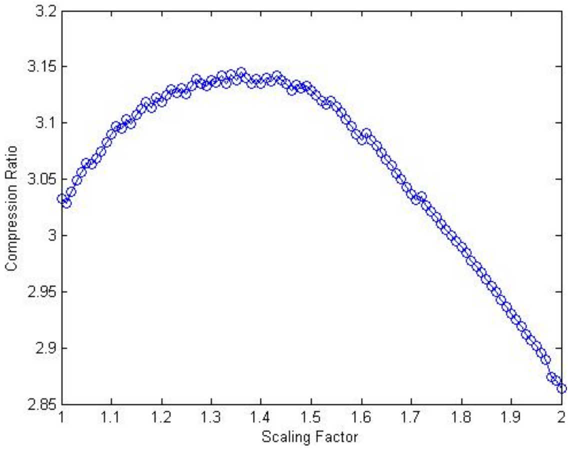

Figure 5. Compression Ratio vs Scaling factor for LBT

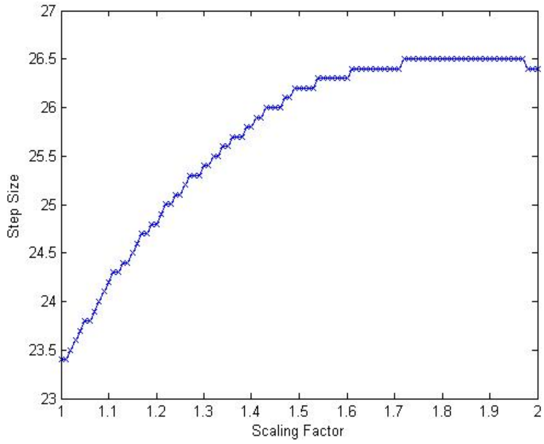

Figure 6. Quantisation step vs Scaling factor for LBT

Figure 7a. Pre-filtered image for scaling factor = 1

Figure 7b. Pre-filtered image for scaling factor = $\sqrt{2}$

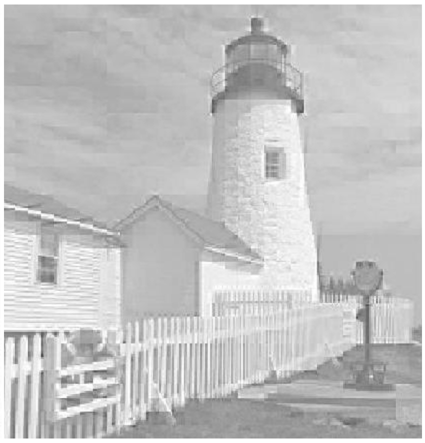

Figure 7c. Pre-filtered image for scaling factor = 2

Figure 8a. Reconstruction for LBT using block size 4

Figure 8b. Reconstruction for LBT using block size 8

Figure 8c. Reconstruction for LBT using block size 16

Figure 9. Single low pass and high pass image from LeGall filters

Figure 10. Low pass and high pass images from LeGall filters (scaled for display)

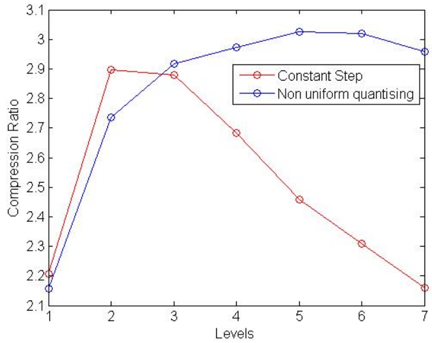

Figure 11. Compression Ratio for quantising under different schemes

Figure 12a. 1 level uniform quantise DWT

Figure 12b. 3 level uniform quantise DWT

Figure 12c. 5 level uniform quantise DWT

Figure 12d. 1 level non uniform quantise DWT

Figure 12e. 1 level non uniform quantise DWT

Figure 12f. 1 level non uniform quantise DWT