SF2: Image Processing Final Report

1. INTRODUCTION

Image compression is an important field today for a variety of reasons. It allows more efficient transmission and storage of the image. Given the high definition images prevalent today that consume a lot of memory, efficient and high quality image compression is vital. Various image compression methods have been explored so far:

- Laplacian pyramid (a multiresolution image compression scheme)
- Discrete Cosine Transform (DCT) (block encoding scheme)
- Lapped Biorthogonal Transform (LBT) (another block encoding scheme)
- Discrete Wavelet Transform (DWT) (another multiresolution image compression scheme)

It is desired to compress an image to 5kB with the best image quality possible. The basics of these methods have been covered in the earlier interim reports but for the sake of reference some key results are outlined in the next section.

Following this, the results from the tests that were part of the final design process will be presented. First the required compression will be ignored and the focus will be on image quality with RMS error used to simulate the eventual compression. Next, an attempt is made to compress the image to 5kB while maximising quality keeping the results from the earlier tests in mind.

2. OVERVIEW OF PAST RESULTS

The results stated in the earlier reports are re-stated here for reference. The maximum compression ratio ($\frac{bits\ in\ reference\ scheme}{bits\ in\ scheme\ under\ evaluation}$) achieved under each scheme (for the lighthouse image) is stated below along with the conditions required to achieve it.

- Laplacian pyramid $\approx$ 1.6 with non-uniform quantising under the MSE scheme where each layer is quantised such that it contributes an equal amount to the error in the reconstructed image.
- DCT $\approx$ 2.9 using a block size of 8
- LBT $\approx$ 3.4 using a block size of 8 and a scaling factor of 1.37.
- DWT $\approx$ 3 using 5 levels and non-uniform quantising under the MSE scheme where each layer is quantised such that it contributes an equal amount to the error in the reconstructed image.

The Laplacian pyramid gives very good image quality but a relatively poor compression ratio. The LBT and DWT give reasonably good image quality as well as compression whereas the DCT is riddled with some block artefacts but does give good compression.

More quantifiable measures of image quality will be discussed later in the report at which point the image quality from each of these schemes can be evaluated in more depth, more reliably and less arbitrarily. Significant additional improvements can be achieved from these schemes using more subtle quantisation schemes and suppressing some sub images that contain less information (lower energy). These improvements are investigated in the next section along with some additional methods to enhance the compression.

3. ADDITIONAL IMPROVEMENTS

Three additional methods of improvement are investigated in this section: (on the lighthouse image)

- **Centre clipped linear quantising:** this exploits the fact that the probability distribution of the intensities of band-pass sub images are highly peaked at zero. If more samples are quantised to zero, then high compression can be achieved.
- **Suppressing sub images:** Most of the information is concentrated in the DC co-efficient as is evident when these sub images are drawn or their energy is found, as has been done in earlier reports. Therefore, some information can be thrown away by suppressing sub images to achieve high compression for a relatively small loss in quality.
- **Singular Value Decomposition:** The Singular Value Decomposition (SVD) can be used as an efficient transform in image compression by representing the image as a low rank approximation of itself. This is not very useful in isolation when the image needs to be transmitted as it yields a poor quality reconstruction upon quantisation but it can be used as an effective pre-processing stage to enhance the performance of the compression scheme chosen.

3.1 CENTRE CLIPPED LINEAR QUANTISING

The change in compression ratio as the rise ratio of the ‘zero’ step changes is shown below.

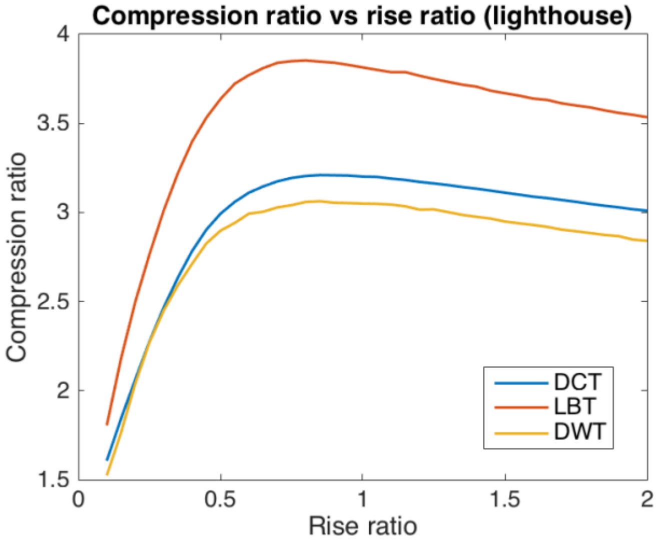

Figure 1. Variation in compression ratio as ratio of first rise changes

Optimal rise ratio

It can be seen that there is a peak in the compression ratio at a rise ratio of about 1 after which the compression falls again, albeit more slowly. It therefore seems that a rise ratio of about 1 is ideal as a compromise between effective compression and image quality.

3.2 SUPPRESSING SUB IMAGES

As stated earlier, most of the energy is carried by the DC components (low pass images) and very little in comparison is carried by the higher frequency components. Therefore, by suppressing some sub images after the forward DCT/LBT has been carried out, a much higher compression ratio can be achieved (If RMS error matching is not done after suppressing – it makes sense to investigate it this way as RMS error matching will not be done in the final scheme).

The relationship between a threshold of energy for sub images discarded (normalised by dividing by the highest energy) and the compression ratio is shown below for the DCT. The results are very similar for the LBT given the similarity in the procedures for the two.

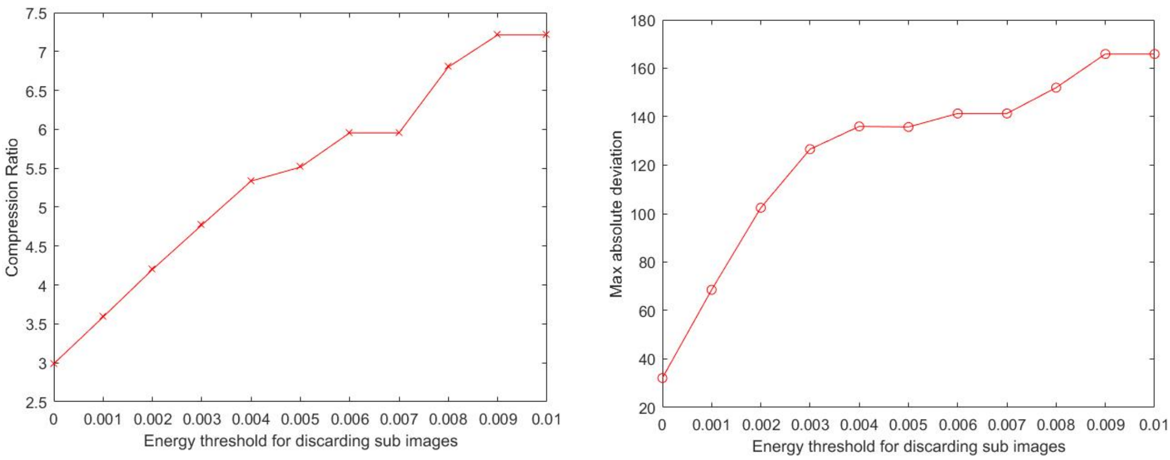

Figure 2. Variation in compression ratio and reconstruction error as sub images are discarded

3.3 SINGULAR VALUE DECOMPOSITION

Singular Value Decomposition (SVD) is not a useful compression scheme in isolation as it yields poor results upon reconstruction when quantised. However, it can make for a useful pre-processing scheme to enhance the compression of the chosen scheme. Again, the DCT has been chosen as the scheme to investigate this against as it representative of performance from LBT as well. The investigation is done by varying the ratio of the first rise during quantisation on the original image as well as on an image where a few singular values have been discarded.

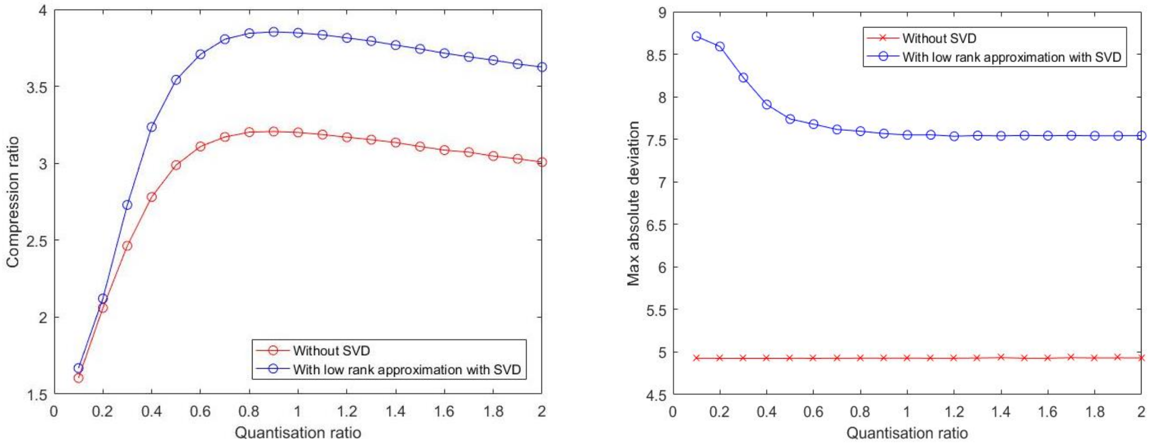

Figure 3. Compression ratio and RMS error: effect of SVD low rank approximation

As can be seen, the pre-processing step of discarding a few singular values to obtain a low rank approximation greatly enhances the compression ratio with a comparatively acceptable RMS error.

The higher RMS error can be minimised by varying the percentage of variance discarded through the singular values and also by applying the SVD on blocks rather than the entire image to exploit local variations and high local correlation in pixel intensities in the image [1]. For simplicity however, this will not be investigated here as the RMS error can be easily brought down by discarding fewer singular values before the compression scheme.

4. BRIEF LITERATURE REVIEW: INDUSTRY STANDARDS

Before choosing a final scheme to perform image compression, it would be useful to have a look at the existing literature on image compression schemes and evaluation to aid and guide the design process in choosing a final scheme.

4.1 Transform methods

Some of the most popular image compression schemes in industry are the JPEG and JPEG 2000 schemes. JPEG uses a DCT with a block size of 8 before doing a zigzag scan of transform co-efficients, quantising them according to quantisation tables based on psycho visual experiments and finally Huffman coding (discussed later) the result.

JPEG 2000 uses a 2D DWT before trellis coded quantisation and binary arithmetic bit-plane coding [2]. Most of the saving in these schemes comes from quantisation and the lossless coding after this at the final stage. It is worth investigating different means of coding as well.

4.2 Coding methods

Huffman coding: In JPEG, the DC co-efficient can be encoded using a pair of symbols that indicate the size of the co-efficient and the actual bits themselves. The AC co-efficients are then scanned in a zigzag way across the image and then Huffman coded.

5. METRICS TO JUDGE IMAGE QUALITY

Image compression is finally only really judged by the quality of the image compared to the original image. There are various measures to quantify this, some of which have been used earlier like RMS – Root Mean Square (L2 norm) and maximum absolute deviation (L1 norm). However, these do not really take into account image quality which is more subjective. There are some methods that attempt to quantify image quality using the original image as a reference:

- Peak Signal to Noise Ratio (PSNR)

$$PSNR = 20 \log_{10} \frac{\max pixel\ value}{\sqrt{MSE}}$$

  The higher the PSNR, the better the image has been reconstructed to match the original image and the better the reconstructive algorithm. [3] It may not be the best to use as it is just another way of representing the RMS error.
- Structural Similarity Index (SSIM)

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$$

  Where μ_x is the mean of X, μ_y is the mean of Y, σ_x is the standard deviation of X, σ_y is the standard deviation of Y, $c_1 = (k_1 L)^2$, $c_2 = (k_2 L)^2$, L is the dynamic range of pixel values, $k_1 = 0.01$ and $k_2 = 0.03$. [4]
  The SSIM can be seen as a quality measure of the image being compared to another image as long as the other image is seen as being of perfect quality. [5]
- Mean Opinion Score (MOS): an index calculated when people score an image on a scale of 1-5 depending on how good or bad it is. This is supposed to be one of the best methods of judging image quality. [6]

Given that MOS cannot be used in this context to judge images, it would be good to find a quantifiable index that is most correlated with MOS. Data shows that SSIM is a better measure of image quality than PSNR [5].

SSIM scores are much more highly correlated with MOS than PSNR values. Therefore, SSIM can be used to judge the quality of the compression algorithm. (The closer the value is to 1, the better the quality of the reconstruction)

6. SHORTLISTED METHODS AND RESULTS

6.1 Scheme 1:

- Low rank approximation with SVD: To enhance the compression as a pre-processing step
- DWT with an option to discard low energy images: Using the JPEG 2000 standard as an inspiration and also because DWT gives the best reconstruction quality
- LBT on each sub image with an option to discard low energy blocks because LBT gives the highest compression ratio
- Uniform Quantisation of LBT blocks

Results:

The results found by playing with some of the variable parameters above are quoted below.

Table 1. Initial Results from Scheme 1 (lighthouse)

SVD percentage variance	DWT levels	Discard DWT image	LBT levels	Discard LBT energy	Quantisation step	First rise ratio	SSIM
100	4	No	8	No	17	0.5	0.58
100	4	Yes	8	No	17	1.0	0.58
95	2	No	8	Yes	17	1.5	0.42
95	2	No	4	Yes	30	1.0	0.50
100	4	Yes	8	No	35	0.5	0.41
100	2	Yes	8	No	10	1.0	0.74
100	4	Yes	4	Yes	10	1.0	0.45

Images representative of the same are shown in the Appendix as Figure 4. Experimenting with different values, the optimal parameters are found to be: 2 levels of DWT and LBT without discarding any sub images and applying a rise ratio of 1 to the zeroth quantisation.

Quantisation is applied uniformly and in order to match an RMS error of 4.84, to some tolerance. (The same as quantising the image to 17 levels). This is done at this stage as we are only interested in the possible image quality from this scheme.

The following results are achieved for the three test images (lighthouse, bridge, flamingo)

Table 2. Results from Scheme 1 (quantised to match RMS error)

Image	SSIM	Max Deviation
Lighthouse	0.75	26.37
Bridge	0.93	22.91
Flamingo	0.86	24.81

Despite these seemingly acceptable values, a grid like artefact is seen in the reconstructed images that makes image quality poor. (Seen in the Appendix in Figure 5 most clearly in the lighthouse image) An attempt is made to see if this artefact can be avoided by using the DCT instead of the LBT. It is hypothesized that not much will change in terms of visual performance as the LBT is simply a pre-filtering operation (Photo Overlap Transform) on the image and then a regular DCT. The LBT has given better visual performance when used earlier in investigations with the test images.

This is exactly what is seen in the reconstructed images that can be found in Figure 6 in the Appendix. It is in fact worse than using LBT in the scheme which is also expected. Therefore a newer scheme is necessary.

6.2 Scheme 2:

Possibly the ordering of the DWT and LBT have caused the grid like artefact discussed above. It is worthwhile investigating if this artefact can be avoided by reversing the ordering of the schemes to apply the LBT first and then apply the DWT on the LBT sub images.

Results

The results from playing with some parameters are seen below.

Table 3. Initial results from Scheme 2

SVD percentage variance	LBT levels	Discard LBT energy	DWT levels	Discard DWT image	Quantisation step	First rise ratio	SSIM
100	8	No	2	No	30	0.5	0.76
100	4	No	4	No	17	1.0	0.70
95	8	No	2	Yes	10	1.0	0.80
95	4	No	4	Yes	10	1.5	0.71
100	4	Yes	2	No	30	0.5	0.72
100	2	Yes	2	No	35	0.5	0.62
100	8	Yes	4	No	17	0.5	0.84

Images representative of the same can be found in the Appendix as Figure 7. These images are of better quality than scheme 1 as can be seen from the higher SSIM index in general and for analogous cases to scheme 1 where the same parameters are used.

It was found the optimal parameters are 8 levels of LBT with no levels of DWT without discarding any sub images and throwing no singular values away. Quantisation is applied uniformly and in order to match an RMS error of 4.84, to some tolerance. (The same as quantising the image to 17 levels).

The following results are achieved for the three test images (lighthouse, bridge, flamingo)

Table 4. Results from Scheme 2 (quantised to match RMS error)

Image	SSIM	Max Deviation
Lighthouse	0.80	25.75
Bridge	0.94	21.58
Flamingo	0.88	23.57

These images can be seen in the Appendix in Figure 8. It is now hypothesized that further improvements can be made by using non uniform quantisation.

6.3 Scheme 3:

- LBT on each sub image with an option to discard low energy blocks
- Non uniform quantisation according to the JPEG quantisation tables to exploit the human visual system [7] along with a rise ratio of 0.8.
- Note: Scheme 2 has been changed to remove SVD and DWT as these are sub-optimal.

At this stage, it is reasonable to assume that this scheme will give the best image quality due to the tests carried out earlier. Therefore, RMS error matching is not done but rather the number of bits required to transmit the image are optimised.

Therefore, after quantisation:

- The co-efficients are coded using a combination of run length and Huffman coding as set out in the JPEG standard.
- This coding scheme is then used to optimise the number of bits.

Results

Table 5. Results from scheme 3 (optimised number of bits)

Image	SSIM	Max Deviation	PSNR
Lighthouse	0.67	78.24	39.29
Bridge	0.72	80.11	37.34
Flamingo	0.73	93.89	37.81

7. FINAL COMPRESSION METHOD SELECTED

It is also useful to compare the image quality in terms of SSIM for all other schemes as a final check to ensure that an 'optimal' method in some sense is being used for the same number of bits. This can be seen in the table below for the three test images.

Table 6. Comparison of SSIM for different schemes using the allocated 5kB

Scheme	Lighthouse		Bridge		Flamingo	
	PSNR	*SSIM*	*PSNR*	*SSIM*	*PSNR*	*SSIM*
1	30.04	0.60	25.83	0.73	25.87	0.67
2	39.29	0.66	37.40	0.72	37.77	0.72
3	39.29	0.67	37.30	0.72	37.80	0.72

Given that there isn't much of a difference between Scheme 2 and Scheme 3 in terms of this quantitative data, it is worthwhile comparing the actual pictures of the test images compressed under both schemes.

As can be seen in the Appendix, the image quality of the images in Figures 9 and 10 is fairly similar but there are some differences, highlighted on the image. In the lighthouse image, the sky is of better quality using the final scheme and there is a little more detail on the lighthouse. In the bridge image, the image from the final scheme appears sharper with a bit more detail. Finally, the flamingos image from the final scheme seems brighter and a bit more detail is visible to the eye. This would make sense as the quantisation is done based on tests on the human eye.

8. CONCLUSION

Various schemes and various combinations were tested in the course of the final design. Patrick (my colleague in the design process) and I shared the tasks involved. We created a GitHub repository for the code that we could edit and work on together. I mainly focussed on doing some research about possible methods to follow for the compression and coding and finding different means of quantifying visual quality. I also worked on the custom quantization of the LBT components in the final scheme and the SVD implementation for pre-processing. Patrick focussed his attention on the DWT and LBT experiments while also designing some functions to test.

The optimal (competition winning) solution was surprising in its eventual simplicity. Only the LBT was needed with custom quantisation tables to guarantee both good compression and visual quality. More efficient compression was achieved using a combination of run length and Huffman coding as outlined in the JPEG standard.

9. REFERENCES

[1] Y.-j. Jia, . P.-f. Xu and X.-m. Pei, "An Investigation of Image Compression Using Block SIngular Value Decomposition," *International Conference on Communication and Information Processing,* pp. 723-731, 2012.

[2] C. Lui, "A Study of the JPEG-2000 Image Compression Standard," Queen's University, Kingston, Ontario, Kingston, Ontario, Canada, 2001.

[3] National Instruments, "Peak Signal-to-Noise Ratio as an Image Quality Metric," National Instruments, 11 September 2013. [Online]. Available: http://www.ni.com/white-paper/13306/en/. [Accessed 29 May 2016].

[4] Z. Wang, A. . C. Bovik, . H. . R. Sheikh and E. P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity," *IEEE Transaction on Image Processing Volume 13,* pp. 600-613, April 2004.

[5] Z. Wang, "The SSIM Index for Image Quality Assessment," University of Texas, 2 June 2003. [Online]. Available: http://live.ece.utexas.edu/research/quality/SSIM/. [Accessed 30 May 2016].

[6] F. Ribeiro, D. Florencio and . V. Nascimento, "Crowdsourcing subjective image quality evaluation," in *2011 18th IEEE International Conference on Image Processing*, 2011.

[7] C.-Y. Wang, . S.-M. Lee and . L.-W. Chang, "Designing JPEG quantization tables based on human visual system," *Signal Processing: Image Communication Volume 16, Issue 5,* pp. 501-506, January 2001.

[8] L. Torres-Urgell and R. Lynn Kirlin, "Adaptive image compression using Karhunen-Loeve transform," *Signal Processing,* pp. 303-313, December 1990.

[9] M. K. Mathur and G. Mathur, "Image compression using DFT through Fast Fourier Transform Technique," *International Journal of Emerging Trends & Technology in Computer Science (IJETTCS),* pp. 129-133, July 2012.

[10] B. Girod, "web.stanford.edu," [Online]. Available: http://web.stanford.edu/class/ee398a/handouts/lectures/07-TransformCoding.pdf. [Accessed 29 May 2016].

[11] Data Compression, "Lossless Data Compression," 2016. [Online]. Available: http://www.data-compression.com/lossless.shtml. [Accessed 29 May 2016].

[12] MIT, "web.mit.edu," 13 February 2013. [Online]. Available: http://web.mit.edu/6.02/www/s2012/handouts/3.pdf. [Accessed 29 May 2016].

10.APPENDIX

Figure 4. Representative reconstructions from Scheme 1

Figure 5. Image reconstruction with grid like artefact from Scheme 1: quantised to match RMS error of 4.84

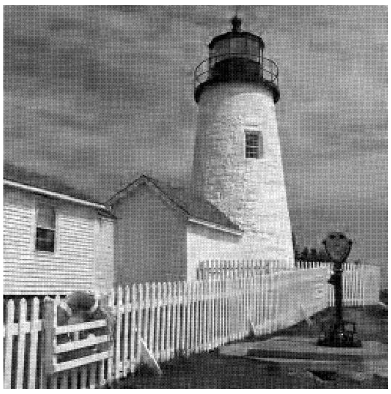

Figure 6. Image reconstruction with grid artefact from Scheme 1: LBT replaced by DCT

Figure 7. Representative reconstructions from Scheme 2

Figure 8. Image reconstruction from Scheme 2: quantised to match RMS error of 4.84

Figure 9. Image reconstruction from scheme 3: compressed to 5kB

Figure 10. Image reconstruction from scheme 2: compressed to 5kB