Second Interim Report

Devang Agrawal

May 21, 2015

1 Introduction

This report follows on from the First Interim Report. In this report mainly the compression ratio and the visual features produced by various transformations are explored. The transforms explored are Discrete Cosine Transform, Lapped Bi-Ortogonal Transform and the Discrete Wavelet Transform.

2 Discrete Cosine Transform (DCT)

DCT performs energy compaction by operating on non overlapping blocks of pixels.

After transforming the image using DCT different frequency components of various blocks can be grouped together to form sub-images representing the frequency content of the entire image. These are shown in Fig. 1a. Most of the energy of the image can be seen to be concentrated in the DC component, the energy keeps decreacing for higher spatial frequencies. This is very clearly shown in Fig.1b .

Images of the different frequency DCT coefficients correspoding to different spatial frequencies were generated using the transformation matrix, they are given in Figure 2 in the Appendix. The top left image represents the DC component, moving rightwards increaces the horizontal spatial frequency and moving downward increaces the vertical spatial frequency.

Quantization and Coding efficiency

The transformed image can be quantized to decreace the bits required to store it. The transformed image was quantized using a step-size of 17. The normalized regrouped image is given in Figure3. It can be seen that there are slightly different probabililty distributions of pixels in each of the sub-images. The distribution of the DC component based subimage can be seen to be very different from the others.

The reconstruction was done after quantizing the transformed image with a step size of 17. This then means that the sum of the total entropy of the sub-images is less than the total entropy of the entire image. Some compression can hence be achieved simply by using different encodings for each of the subimages. Table 1 shows the total entropy of the image for different transform sizes calculated using either sub-images or the entire image.

The reconstructed image from the Quantized transformed image is given in Figure4 (a) shows a part of an image reconstructed from the transformed image, block aretefacts have been introduced into the image and are clearly visible hence affecting the visual quality. However the overall visual quality is much better than that of the directly quantized image which has unrealistic intensity jumps. The standard deviation between the reconstructed image and the original image is less than that of a directly quantized image . The compression ratio of this scheme was then calculated to be around ***3.*** Compression ratios for various transform sizes are given in Table 2.

Alternative transform sizes

Some alternative block sizes are used for making the transformed images. The images obtained by using 4 x 4 blocks look very different from the ones wih 16 x 16 blocks. Parts of the images are given in Figure 5. The block artefacts in these images are really bad. The image with 4 x 4 DCT has really small blocks which are all visible and make the image look very "pixelated", in the image with 16 x 16 DCT , the blocks are really big and adversely affect the appearance of the image. The compression ratios for various transform sizes are given in Table 2. It can be seen that the compression ratio is the largest for the 8 x 8 blocks. 8 x 8 blocks provide the highest compression ratio and the best resultant image and are hence the best choice for the sample "lighthouse" image used. However it can be possible that different transform sizes might work better for different types of images. For a very smooth image, with no abrupt changes, a bigger transform size might be better and for image which have very rapidly varying spatial features and sharp edges, a smaller transform size might be better. In Figure 5, it can be seen that the 4 x 4 transform captures the sharp edges near the boundary of the lighthouse and the sky much better than the others, the 16 x 16 transform captures the smooth sky much better than the others. Also artificial images consisting of simple blocks of various sizes would give the best performance with blocks corresponding to their size.

<<review>>When comparing different transform sizes there is a concern that using `dctbpp` gives smaller entropy values for larger transform sizes since the regrouped images are smaller and more deterministic. Using `dctbpp(Yr,256)` effectively calculates the entropy individually for each pixel. Each pixel considered as a sub-image is deterministic and hence its entropy is zero. This is a completely unreasonable result as it suggests that zero bits are required to store the image. An alternative would be to use the same sub-image size while calculating bits per pixel. The compression ratio for this scheme is given in Table 2, as expected the compression ratio for 4 x 4 blocks increaces and that for 16 x 16 blocks decreaces

The actual coding scheme would determine which method is closer to the actual number of required bits for the transformed image.

3 Lapped Bi-orthogonal Transform (LBT)

LBT transforms overlapping blocks in X to generate smaller non-overlapping blocks in Y. It involves a prefiltering stage which transforms the image to make blocks overlap with nearby blocks. The bases for this prefiltering stage are shown in Figure 6, parts of some prefiltered images for different POT factors (s) are shown in Figure 7 . The block artefacts become much more visible with increasing 's', this is since with increasing s the amount of overlap between adjacent blocks increaces in the prefiltering and hence the block pattern becomes more obvious.

Figure <<7>> shows how the Compression ratio varies with POT scaling factor (denoted s), it can be seen that the Compression Ratio increaces and then decreaces, the Compression ratio is maximized at s=1.3 with Compression ratio = 3.197. As 's' increaces the overlap increaces which initially helps in reconstruction and increaces the Compression ratio but after a point the overlap becomes a hindrance and reduces the Compression ratio.

Figure 7 shows the optimized step size for each POT scaling factor, this increaces almost uniformly with slightly wobbles around s=2. Figure <<8>> shows reconstructed images with optimum step size for s=1 , 1.3 and 2 . The 3 images look fairly similar but comparing these to images obtained in Figure 5 for DCTs, LBT can be seen to have produced a much better looking image despite having the same standard deviation relative to the original image. The blocking artefacts are mostly gone due to the overlapping blocks used. The lighthouse still distorts the sky around it, but the effect is much weaker and affects visual experience less.

Alternative Transform Sizes

The LBT was then investigated using different transform sizes. When calculating the number of bits `dctbpp(Yr,16)` was used regardless of the block size to reflect simultaneous coding of several blocks used in actual algorithms. For all experiments in this part 's' was fixed to 1.3 . The Compression ratios are given in Table 3. It can be seen that the 4 x 4 transform achieves the best Compression ratio. The images obtained using the different transform sizes are given in Figure 11. The one using the 4 x 4 block can be seen to provide the best image, the image doesnt have very prominent block artefacts like the 2 x 2 and also the area of the sky which is blurred in the vicinity of the tower is not as much as the 8 x 8 or 16 x 16.

4 Discrete Wavelength Transform (DWT)

The discrete wavelength transform combines best features of the Laplacian pyramid and the DCT. It works by splitting the image into 4 images, a 2D lowpass image UU, a 2D highpass image VV , a horizontal lowpass and a vertical low pass horizontal highpass image VU and a horizontal low pass vertical highpass image UV.

The energies of the highpass images are much lesser than the ones for the 2D low pass image. This justifies recursively splitting the image UU to achieve energy compaction. The energies of the high-pass images are so low that their intensities need to be scaled up on occasions to display them properly. Figure 12 shows a one layer dwt pyramid, the images UU VU VV and UV can be seen by going clockwise going from the the top-left. UU is a 2D lowpass image, VU picks the vertical edges and UV the horizontal edges. VV is like a 2D highpass image which picks up diagonal edges and very small features (the brick pattern in the lighthouse) in the image.

To do the final reconstruction an equal MSE based scheme can be used to find ratios of step sizes, this is similar to the one used for Laplacian pyramids in the First Interim Report. An alternative is to use simple equal step sizes. Table 4 shows some values of `dwtstep` as defined in the lab handout.

Figure 13 shows the compression ratios for using either constant step size, or the equal MSE scheme. For the equal MSE scheme the Compression ratios increase with increasing n, however increacing n over 4 does not change the Compression ratio that much and hence does not provide much benefit. Hence a 4 level DWT is reasonably optimum for the bridge and lighthouse images. The compression ratio for the equal step size scheme increaces on increasing the levels in the DWT since the impulse response of the deeper layers keeps increasing and hence quantization in them affects the final image greatly.

The compression ratios obtained for the bridge image are much smaller than that obtained for the lighthouse image. The bridge image has a lot of energy in the high pass images due to the plants present around the bridge. This reduces the effectiveness of recursively compressing the low-pass image.

For the bridge image it was seen that there is not that much visual difference between the original bridge image and the image reconstructed using DWT with either equal MSE or equal step sizes. Changing the number of levels does not impact the visual features of the image by much either.

Figure 14 shows the lighthouse for 3 layer DWT using equal MSE scheme or equal step sizes. The clouds have decome very distorted in these images as compared to the original. The clouds form the low frequency part of the image and are hence affected the most when DWT is used.

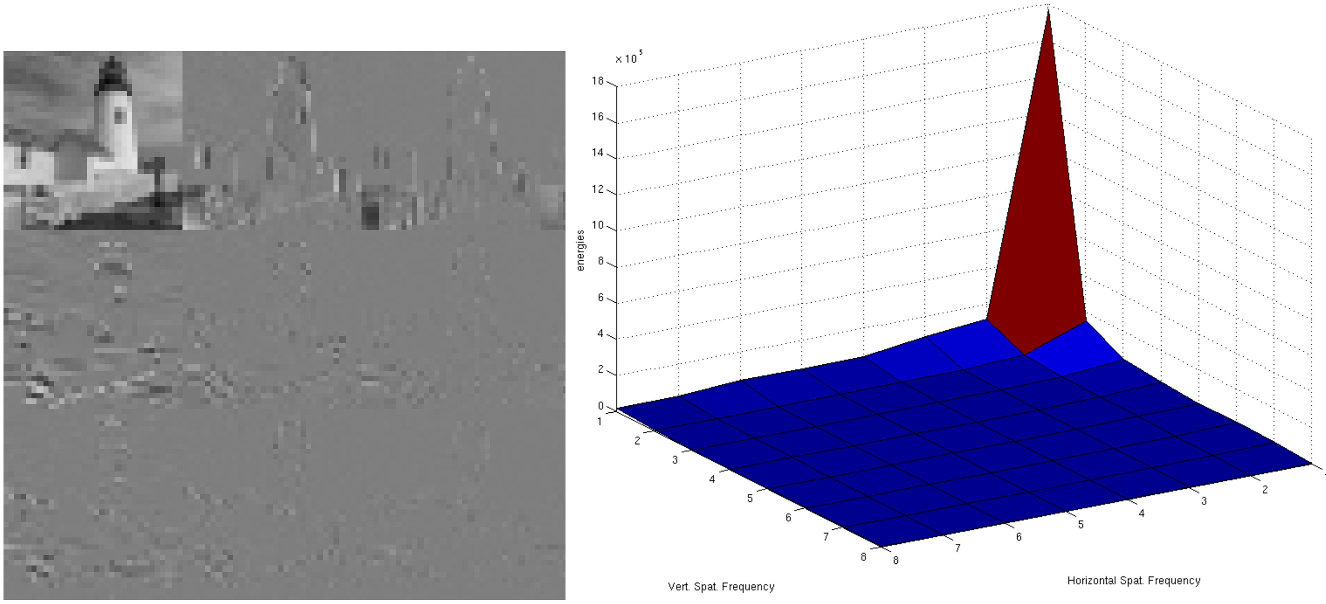
No significant visual difference is observed by varying the number of layers.

Overall the image reconstruction for both the cases is reasonably good pointing towards the strength of DWT as a compression method.

5 Conclusion

All the three transforms eplored in the lab demonstrated excellent compression ratios. However the visual features displayed by the three transforms were very different. The Discrete Cosine transform resulted in unpleasant blocking artefacts, the Lapped Bi-Orthogonal Transform improved on these artefacts and reduced their effect, but there was still a considerable distortion around object boundaries. The discrete wavelet transform proved to be the best transform explored in the lab. It achieved a Compression Ratio comparable to the other 2 transforms but also managed to reduce most of the artefacts present in the other 2.

Appendix



(a) Some subimages formed by grouping different frequency components from various transformed image blocks together

(b) Energies of subimages for different spatial frequencies

Figure 1

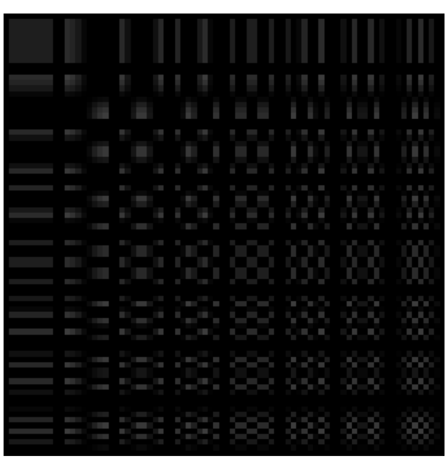


Figure 2: The 2D DCT coefficients

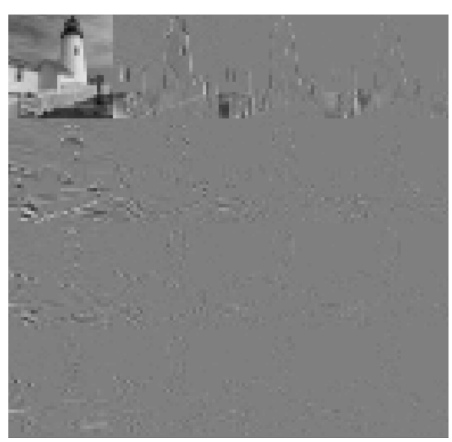


Figure 3: Part of the Normalized Regrouped transformed image formed after quantizing the transformed image

(a) Reconstructed image (b) Directly Quantized (c) Original

Figure 4: Part of original image with various transforms

(a) 4 x 4 (b) 8 x 8 (c) 16 x 16

Figure 5: Reconstructed images using different Block sizes with same deviation from the original image

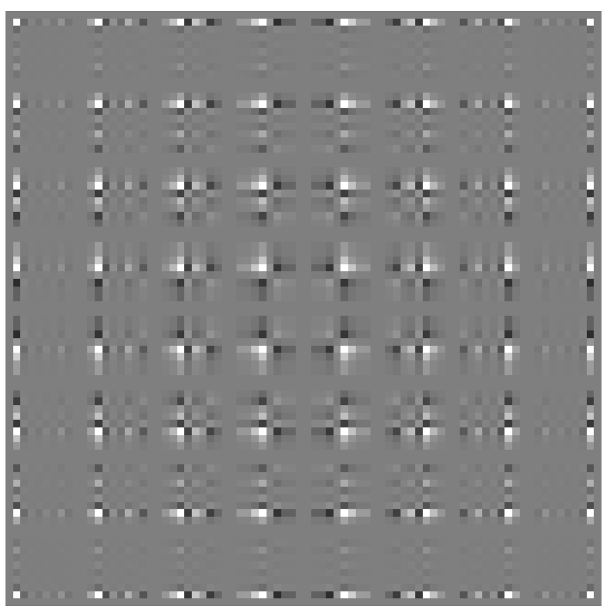
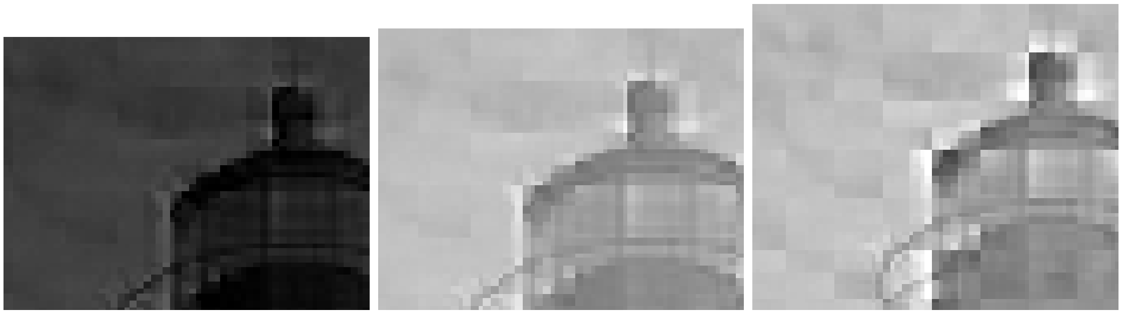


Figure 6: LBT prefiltering bases



(a) s=1 (b) s=1.3 (c) 2

Figure 7: Prefiltered Images for different POT scaling factor (s)

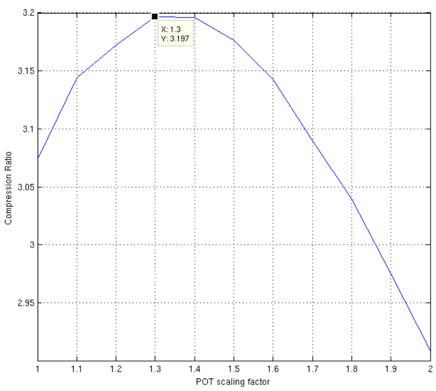


Figure 8: Compression ratios vs POT scaling factors

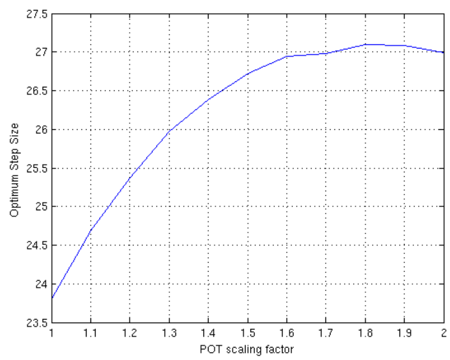
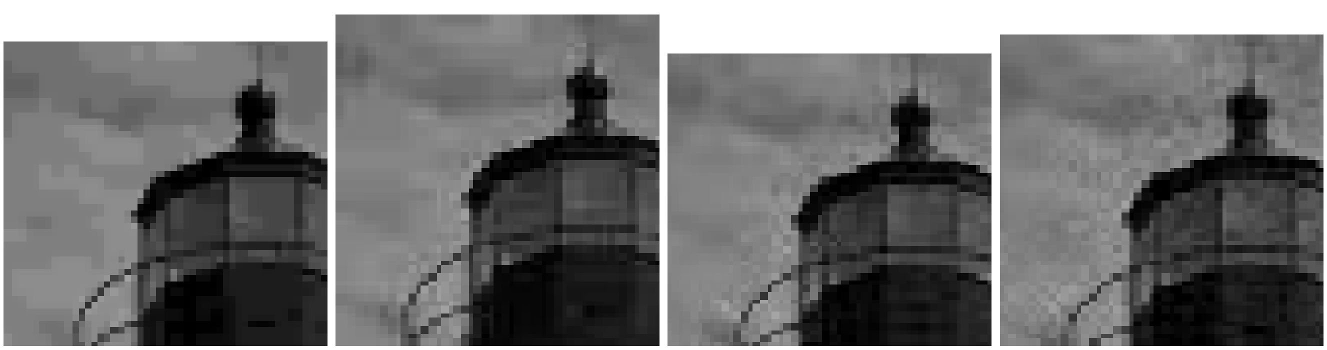


Figure 9: Optimized Quantization step vs POT scaling factors

(a) s=1 (b) s=1.3 (c) s=2

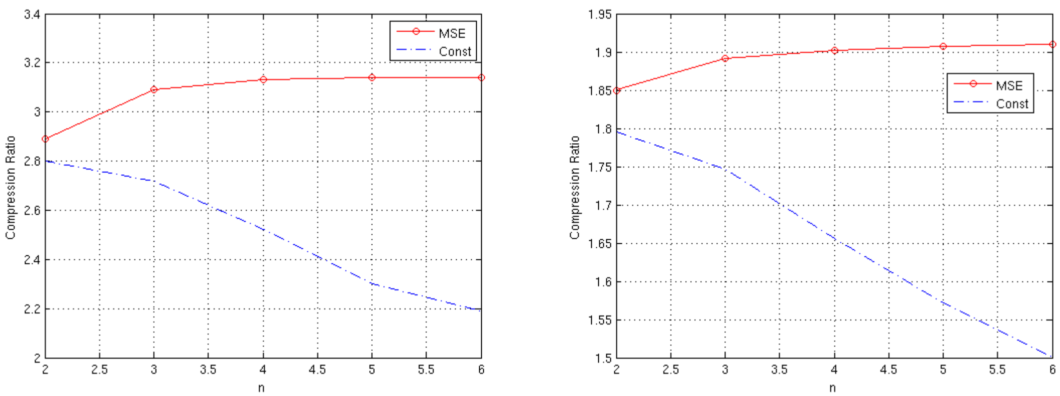
Figure 10: Reconstructed LBT images with different POT scaling factors



(a) 2 x 2 (b) 4 x 4 (c) 8 x 8 (d) 16 x 16

Figure 11: Reconstructed images for Different Transfrom Sizes for LBT

Figure 12: Images produced from a 1 layer DWT pyramid



(a) For the lighthouse image

(b) For the bridge image

Figure 13: Values of Compression ratio for either constant step size or equal MSE strategies with DWT

(a) Equal MSE

(b) Equal Step Size

(c) Original

Figure 14: Reconstructed lighthouse image using 3 layer DWT

	Total Entropy(in kbits)
Entire Image	97.4
Using Sub-images	109.63

Table 1: Entropy of image

Transform-Sizes	4x4	8x8	16x1
Compression Ratio variable N	2.69	2.99	2.93
Compression ratio with N=8	2.92	2.99	2.851
Optimum-Step Size	24.35	24.15	22.7

Table 2: Compression Ratio with either Variable N or N=8 in `dctbpp` and Optimum step-sizes

Size	Compression Ratio
2 x 2	2.99
4 x 4	3.62
8 x 8	3.50
16 x 16	3.01

Table 3: Compression ratios for various transform sizes with LBTs

1	0.6521	0.3566	0.3864
0.7233	0.5632	0.3274	0
1	0.6521	0.3556	0

(a) n=3

1	0.6521	0.3566	0.1821	0.1943
0.7233	0.5632	0.3274	0.1706	0
1	0.6521	0.3566	0.1821	0

(b) n=4

Table 4: `dwtstep` ratios based on equal MSE error.