University of Cambridge

Engineering Tripos Part IIA

SF2 Image Processing

Second Interim Report

Clinton N. Igwegbu May 22, 2019

Introduction

The first interim report explored the Laplacian Pyramid image compression scheme. This report explores the Discrete Wavelength Transform (DCT), Lapped Biorthogonal Transform (LBT) and the Discrete Wavelength Transform (DWT). Then it compares these three new methods to the methods introduced in the first report.

1. Discrete Cosine Transform

The DCT is advantageous compared to the laplacian pyramid method in that it preserves the number of samples in the original image while a 4-stage laplacian pyramid causes a 1.33 sample expansion. Since the aim is to compress the image, the DCT therefore has a headstart on the pyramid method. As mentioned in the project handout, the DCT is similar to the Fourier transform, effectively giving the frequency content of the image. The transform is found by multiplying the rows and columns of the image by a matrix C as described in the project handout. This method evaluates the images in blocks of size NxN where N is the size of the square transform matrix. Quantisation allows the removal of fine bit value transitions from one pixel to the next due to the finite step used to capture the data. Quantisation in the frequency domain is therefore a form of low-pass filtering. It was observed that the higher frequency images had lower energy, as seen in figure 1, where energy is defined here as the sum of the squared pixel values. Given that the energy content of an image wanes with increasing frequency, low-pass filtering (quantisation) rejects the frequencies of the image that have very little data. The DCT results in a matrix with adjacent samples having different frequency components. By regrouping the samples in the transform such that similar frequency components are placed adjacently, a grid of subimages is formed with similarly correlated values. The function dctbpp takes advantage of this by coding each subimage independently then averaging the entropy of each subimage to find the overall quantised transformed image entropy. Compression ratios are therefore calculated by dividing the entropy of the directly quantised image by that of the transformed image. It was ensured that the directly quantised image and the transformed image had the same rms error. Keeping this in mind, a compression ratio of 2.9416 was obtained using this method, for N=8 and step_X=17. It was noted that the method of individual subimage coding was flawed as the sample size that the probability distribution of a subimage is obtained from decreases as the number of subimages (NxN) increases.

Giving N=256 creates a grid of 256 by 256 subimages i.e. each sample in the transform is its own image meaning that there is no uncertainty in the value of a sample in each subimage. This falsely enables dctbpp to calculate an entropy of 0 for each subimage which gives an infinite compression ratio. It was suggested in the project handout to code on fixed 16x16 blocks. This suggestion was considered in the LBT section. For the lighthouse image, it was seen that the N=8 offered the best compression ratio (2.9416) compared to N=16 and N=4.

The original image is reconstructed by multiplying the transformed image, Y, by the transformation matrices C appropriately. Note that the DCT reconstruction in figure 2 is considerably less grainy than the directly quantised image and matches the original closely. Since the DCT treats images in blocks the reconstructed image can be seen to reveal discontinuities with a period of N, as can be seen by close examination of figure 2c. The LBT aims to reduce these artefacts.

2. Lapped Biorthogonal Transform

The blocking artefacts seen in DCT methods come as a result of neglecting correlations between image blocks (adjacent NxN regions). The LBT overcomes this problem by using overlapping blocks in X, the image. These transform to smaller non-overlapping blocks in Y, the transform. The LBT performed in this project had a simple solution as it was represented by a prefiltering operation (Photo Overlap Transform [POT]) followed by a DCT. Similarly, the inverse LBT was taken as a DCT followed by a postfiltering operation. A scaling factor, s, gave the degree of bi-orthogonality of the transforom such that a non-unity s produced Pf and Pr coefficients that are unequal, where Pf is the forward POT and Pr is the reverse POT. For an 8x8 DOT and equivalent direct quantisation step size of 17 (preserving rme), s was varied to find the value which maximised the compression ratio and the quantisation step size. A maximum compression of 3.1358 was found at s = 1.4051 and maximum step size of 27.0792 was found at s = 1.80202. This compression maximising case was visualised, giving the reconstruction in figure 3. As can be seen by examining figure 3 closely, the blocks seen in figure 2c have been eliminated but a 'fuzzy aura' can still be seen around some objects, most notably around the tower. As mentioned in the DCT section it was decided to code on fixed 16x16 blocks. It was then found that the N=8 case gave higher compression ratios than the N=4 and N=16 cases.

The improvements due to this method are described below:

3. Discrete Wavelet Transform

This transform does not expand the number of image samples (like DCT) but it represents an image as a sum of high pass subimages with one lowpass image (like laplacian pyramid). The basic principle behind this energy method is that the first DWT stage splits an image into low and high frequency components (referred to here as U and V). Remember that the energy of an image is centred around low frequencies. So it therefore makes sense to iteratively split the low frequency band into lower and upper frequency ranges so and then quantise the resulting set of subimages using different step sizes such that more bits may be used to represent the energy dense frequencies of an image (lower frequencies) and fewer bits for the sparse frequencies (higher frequencies). The 5 and 3 tap pair of LeGall filters is used to perform this 'frequency halving' instead of square pulse filters so as to prevent blocking artefacts in the image reconstruction. The result of one stage of the 'frequency halving' (output of first filter pair) can be seen in figure 4a. One stage DWT is formed by 'havling' each of U and V again to form an image [UU VU; UV VV] as given in figure 4b. Since the energy of the whole image was split into a series of subimages, the data was scaled up by a factor of 1.8 to amplify it for viewing, as otherwise the pixel amplitudes would be too small for the image to be seen clearly. One DWT stage was performed by the function dwt.m. A function nlevdwt.m was written to perform n dwt stages. Quantisation was performed by quantdwt.m which quantised each subimage using the step sizes given in the matrix dwtbpp. It's code may be seen in appendix B. Just as in the Laplacian Pyramid case, in the first interim report, the equal MSE and constant step size approaches were attempted.

A. Images

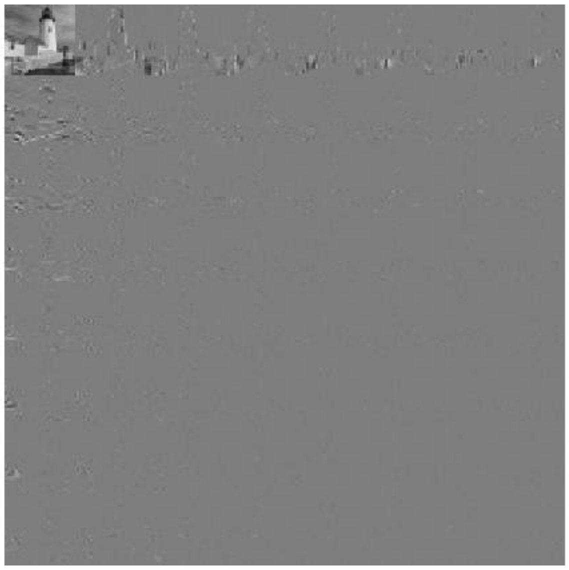
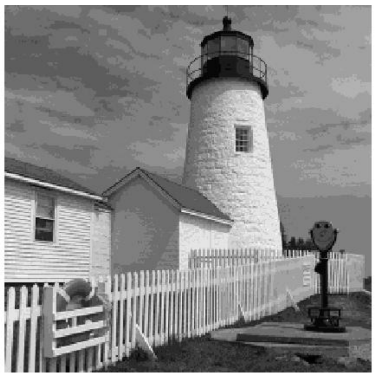


Figure 1: Regrouped DCT Transform

(a) Original


(b) Directly quantised image

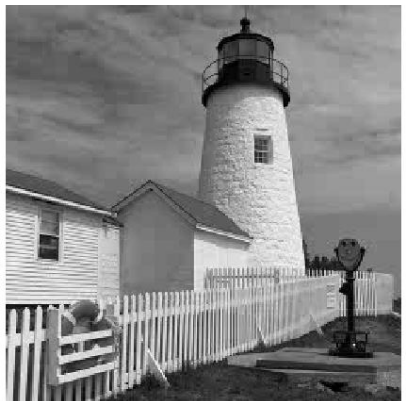

(c) DCT Reconstruction

Figure 2: Visual Effects of DCT Reconstruction

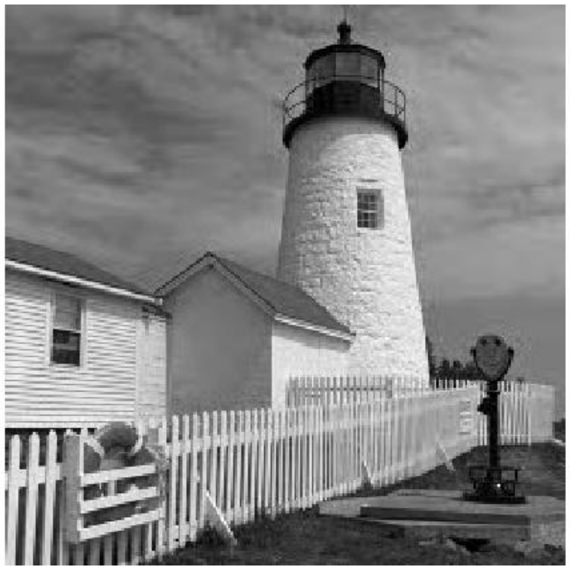
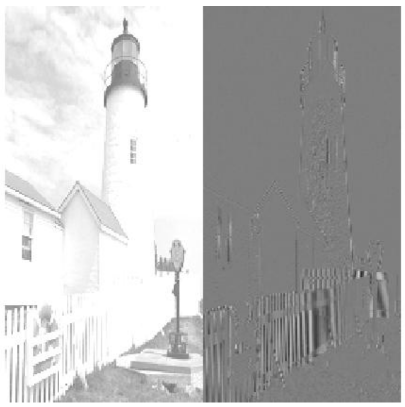


Figure 3: LBT Reconstruction


(a) Original

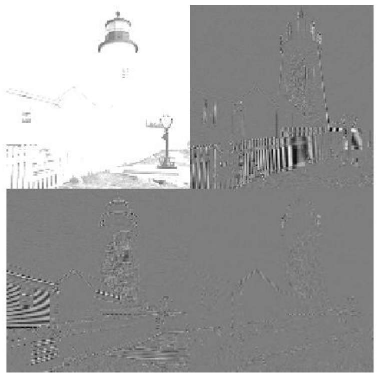

(b) Directly quantised image

Figure 4: First Stage DWT

B. Code Listings

```
function [Yq, dwtent] = quantdwt(Y, dwtstep)

m = length(Y)/2;
Yq=Y; % initialise Yq
dwtent = dwtstep; % initialise dwtent
n = size(dwtstep) - 1; % num stages of DWT

% quantise each highpass image according to step defined in
dwtstep
% then store entropies
for i = 1:n
    t=1:m;
    Yq(t,t+m) = quantise(Y, dwtstep(1,i)); dwtent(1,i) = bpp(
Yq(t,t+m));
    Yq(t+m,t) = quantise(Y, dwtstep(2,i)); dwtent(2,i) = bpp(
Yq(t+m,t));
    Yq(t+m,t+m) = quantise(Y, dwtstep(3,i)); dwtent(3,i) = bpp
(Yq(t+m,t+m));
    m=m/2;
end

% quantise final lowpass image then store its entropy
t=1:m;
Yq(t,t) = quantise(Y, dwtstep(1,n+1)); dwtent(1,n+1) = bpp(Yq(
t,t));


return
```