ELE725 Lab 3 Report

Intra and Inter Frame Coding

Ashif Fahim
Department of Electrical, Computer and Biomedical Engineering
Ryerson University
Toronto, ON, Canada

Santos Bethany
Department of Electrical, Computer and Biomedical Engineering
Ryerson University
Toronto, ON, Canada

***Abstract*— In this lab, intra and inter frame coding is implemented using the DCT and DPCM method respectively. It was observed how the DC, low frequency and high frequency components of the DCT of an image can affect the reconstructed image. Also, after an image goes through DCT, quantization, and reconstruction, it still holds error from the original image, even though differences between the two images are not visible to the naked eye. This report will also explore the advantages of inter-frame DPCM coding in video compression using various prediction methods and the nature of smooth transition frames.**

I. INTRODUCTION

This lab will explore the core principles of image frames used in images and videos. It will focus on Intra-frame coding of 2D block based Discrete Cosine Transform (DCT) and Inter frame csoding using DPCM.

-

II. THEORY

A. Intra-Frame Coding (DCT)

Discrete Cosine Transform breaks down a finite sequence of data point to a sum of its cosine components. Below is the equation that describes 2D DCT function[1]:

$$F(u,v) = \frac{2}{\sqrt{MN}} C(u)C(v) \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} cos\frac{(2x+1)u\pi}{2M} cos\frac{(2x+1)v\pi}{2N} f(x,y)$$

The inverse DCT is described as follows:[1]

$$f(x,y) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \frac{2}{\sqrt{MN}} C(u)C(v) cos\frac{(2x+1)u\pi}{2M} cos\frac{(2x+1)v\pi}{2N} F(u,v)$$

Where M and N are the dimension of the image and C(·) is ...

In JPEG, after an image has been cosine transformed, it is then quantized for further compression. The quantization values used is a standard 8-by-8 table called the Q table. The Q table below is used for luminance and will be used for a the quantization of a grayscale image for this lab.

$$Q = \begin{bmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \\ 72 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \end{bmatrix}.$$

Figure 1. Q Table for luminance used in JPEG DCT [1]

B. Inter-Frame Coding (DPCM)

Inter-frame coding can be simply done by taking the difference of two consecutive frames where the previous frame is X'_{n+1} and current frame is X_n. Then the difference frame, would simply be $D = X_n - X'_{n+1}$. A more advanced method of inter-frame coding would be to first intra-frame code the previous frame, then take the difference between the encoded frame and the current frame. There are different ways to intra-frame code where Table 1 shows 4 different modes or methods, each with their own prediction operation.

TABLE 1. PREDICTION MODES AND OPERATIONS

Mode	Prediction Operation
1	A
2	B
3	C
4	A + B - C

The variables A, B, C and X correspond to the pixel locations indicated in the figure below. X is the result of the prediction operation.

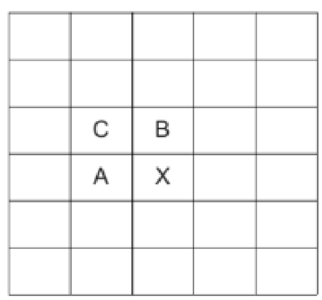

Figure 2. Pixel Location of Variables for Prediction Modes

III. Methodology

A. Intra Frame Coding (DCT)

The MATLAB function *dct2(X)* is used to calculate the DCT values of the input 2D matrix X. The MATLAB function *idct2(X)* is used to calculate the inverse DCT values of the input 2D matrix X. A grayscale image was cropped to a 4-by-4, 8-by-8, and 32-by-32 images to perform, test, analyze DCT behaviour. The following test were performed:

(1) Zero the DC component of the DCT image and apply IDCT
(2) Zero the low frequency components of the DCT image and apply IDCT
(3) Zero the high frequency components of the DCT image and apply IDCT

For test (2) and (3) the MATLAB function histogram was used to determine the low frequency and high frequency components.

Next, Intra-frame coding seen in JPEG images was implemented to a grayscale image in the following sequence:

1. Divide image into 8-by-8 pixel blocks
2. Apply DCT and obtain the DCT coefficients
3. Divide DCT coefficients with its respective quantization constant from the Q table
4. Round the quantized coefficients
5. De-quantize the image by multiplying the rounded coefficients with its respective quantization constant from the Q table
6. Apply inverse DCT to obtain reconstructed image

A. From here, the MATLAB functions *immse(), snr(), pscr()* are used to calculate the MSE, SNR and PSNR of the reconstructed image.

B. Inter-Frame Coding (DPCM)

The inter-frame coding was implemented in MATLAB by creating a function called *interframe_coding()* which accepts an input of two frames, first the previous frame followed by the concurrent frame, and a third input consisting of an integer value from 1-4 which determines the prediction mode to implement. The first frame, the reference frame, is intra-frame coded accordingly to the mode where the X pixel values are then placed into a temporary matrix corresponding to the pixel location of the previous frame. Next, the difference between this temporary matrix and the current frame is taken and placed into the output matrix of the function which is the DPCM coded difference frame.

IV. Results

Part 1. Intra-Frame Coding (DCT)

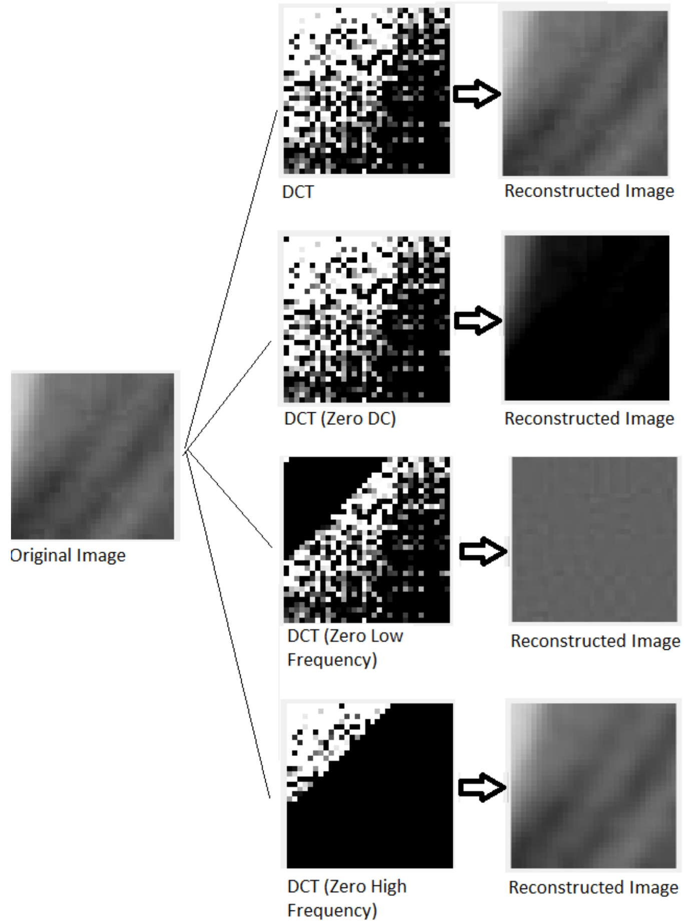

Figure 3. DCT and reconstructed image results of the 3 tests

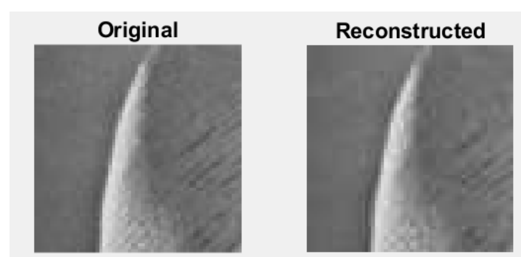

Figure 4. Original gray scale image and reconstructed image after DCT, quantization, and inverse DCT

TABLE 2. MSE, SNR, PSNR of Reconstructed Image

MSE	*SNR*	*PSNR*
18.39	-0.005	-12.66

Part 2. Inter-Frame Coding (DPCM)

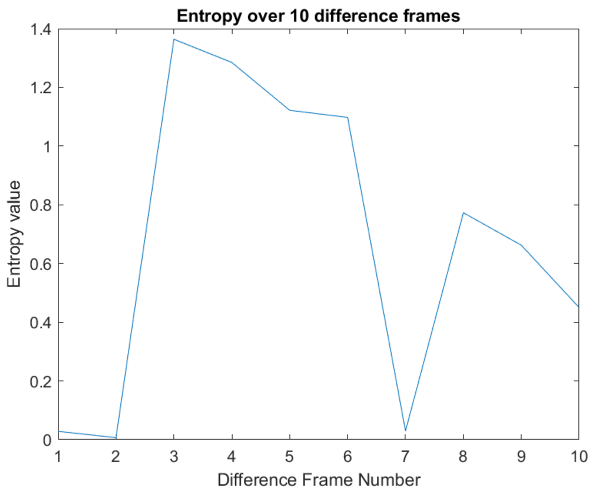

Figure 5. Entropy of First 10 Difference Frames

TABLE 3. CHARACTERISTICS OF FIRST 10 DIFFERENCE FRAMES

No Prediction Mode	
Average Entropy	0.6818
Compression Ratio	11.733 (8/0.6818)
Using Prediction Mode 4	
Average Entropy	0.6523
Compression Ratio	12.264 (8/0.6523)

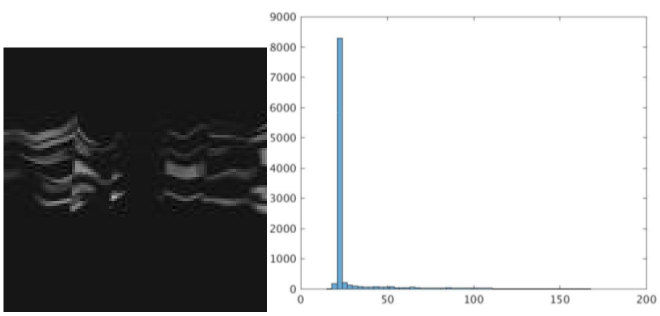

Figure 6. Sample Frame 1

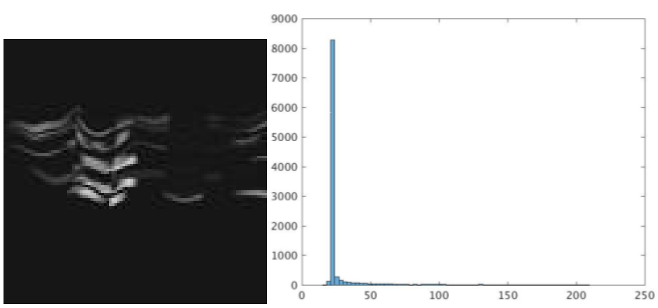

Figure 7. Sample Frame 2

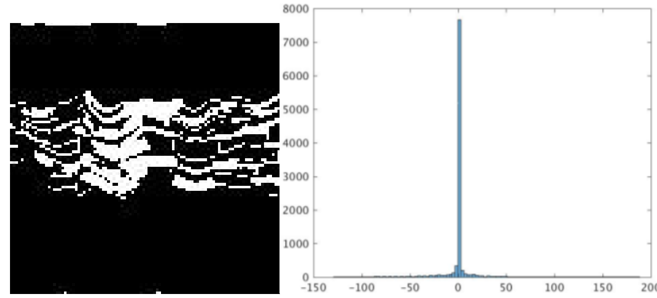

Figure 8. Difference Frame with No Prediction Mode

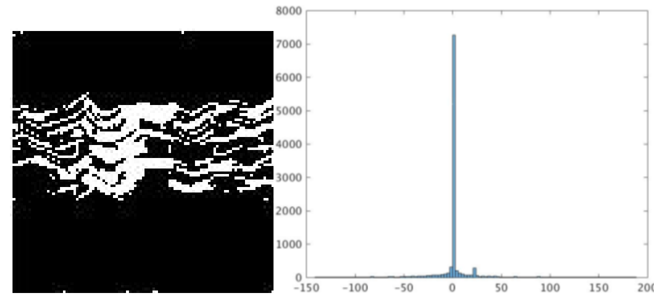

Figure 9. Mode 4 DPCM Coded Difference Frame

V. DISCUSSION

Part 1.. Intra Frame Coding (DCT)

Figure 3 shows the different image results when the three different tests described in the methodology are performed. During normal DCT and inverse DCT, we are able to get a reconstructed image that the same as the original. When the DC component is zeroed out, the reconstructed image is far from what the original image looks like. This is because the DC component represents the average colour/value of the image and when this information is removed, it causes great information loss when performing inverse DCT. For test (2) where the low frequency components of the image are zeroed out, the resulting reconstructed image were far from what the original image looks like. This is due to the fact that the low frequency components contains the most information regarding the image, thus when reconstructed many of the significant information is missing and the program is unable to reconstruct the image that is accurate to the original. On the other hand, when the high frequency components are zeroed, the reconstructed image look fairly similar to the original image. This is because the high frequency components of the image holds the least amount of information when reconstructing and deals little distortion to the reconstructed image. Figure 4, shows a gray scale image and a reconstructed image that has gone through the process of DCT, quantization, and inverse DCT. To the naked eye, the two image look the same however, since the MSE holds a non zero value, this means that the two images are not exactly equal. There are difference between them but, these differences are not visible to the human eyes.

Part 2. Inter-Frame Coding (DPCM)

Figure 5 shows the entropy values for the first ten difference frames for a sample input video file. The difference frames here are simple as the output is equal to subtracting the previous frame from the current frame; there is no intra-frame coding done. It is noticeable that consecutive frames that do not change much have low entropy values as seen in difference frames 1 and 2 which have values very close to 0, and consecutive frames that differ vastly will tend to have larger entropy values. The average entropy of these first 10 difference frames is 0.6818 shown in Table 3. Assuming this an ideal entropy coder, where the first frame does not require any compression and assuming 8 bits per pixel since the frames are in grayscale, the compression ratio was found to be 11.733 also shown in Table 3. Using prediction mode 4, another 10 difference frames were computed where the average entropy was slightly lower at 0.6523 and the compression was higher at 12.264 than using no prediction mode.

Figures 6 and 7 show two sample frames and their respective histograms. Figure 8 shows the difference frame when the first sample frame is subtracted from the second and its histogram. Figure 9 shows the difference frame when the first frame is intra-frame coded using mode 4 then subtracted from frame 2. Histograms of the raw sample frames show no negative integers as grayscale images have values from 0-255. In the histograms of the difference frames, there can be be negative values as the first frame may have higher pixel values than the second frame in the same pixel coordinates.

The histograms of the raw consecutive frames are very similar which is reflected in the difference frames' histograms where the majority of pixels are 0 or very close. This indicates the consecutive frames do not change much. The pixels that do vary largely or become negative after being intra-coded and subtracted are displayed as white colour pixels in the difference frames.

VI. CONCLUSION

In conclusion, intra-frame coding takes advantage of spectral redundancy to compress images/frames. The DC component and the low frequency component of the DCT holds the most significant information when reconstructing (inverse DCT). inter-frame coding does provide great compression ratios as in this experiment, as using a sample of 10 difference frames gave ratio values over 10. Intra-coding frames using various prediction modes before performing DPCM did provide slightly better compression metrics as to using no prediction modes supported by the results in Table 3. Inter-frame coding provides better compression ratio as pixels in the difference frames tend to have lower values as non-computer generated videos have frames that have smooth transitions. These frames will not have sharp transition where the difference of pixels in the same coordinates of the previous and current frame will be small. This is supported by the histograms in Figures 8 and 9 where the majority of pixel values are close to 0.

REFERENCES

[1] Z. Li and M. Andrew, *Fundamentals of Multimedia*, 1st ed. New Jersey: Pearson Education, 2004.