EEE 507 - Spring 2018
Multidimensional Signal Processing
Semester Project

Image Compression

Under the guidance of Dr. Glen Abousleman

SATHYA SRI CHIKOTI

1213268041

1. Introduction:

1.1 Project Outline (Problem Statement):

The aim of this project is to perform the image compression of a 512 X 512 8-bit per pixel, raster-scanned image using Discrete Cosine Transform (DCT) and Sub-band Wavelet Transform (SWT) and compare these two methods.

1.2 Motivation:

Image compression is very important for the efficient transmission and storage of images. Demand for communication of multimedia data through the telecommunications network and accessing the multimedia data through the internet is growing explosively. Image data comprises of a significant portion of the multimedia data and occupies the major portion of the communication bandwidth. Therefore, development of efficient techniques for image compression has become quite necessary.

The need for image compression becomes apparent when the number of bits per image are computed which are resulted from typical sampling rates and quantization methods. For example, the amount of storage required for 14 x 17 inch radiograph scanned at 70 x 10^{-6} mm: 5000 x 6000 pixels, 12 bits/pixel nearly contains 360 x 10^{6} bits. Thus, storage of even a few images could cause a problem and hence are to be compressed [1].

1.3 Approach:

The number of bits required to represent the information in an image can be minimized by removing the redundancy present in it. Redundancy can be found across multiple dimensions including spatial, spectral and temporal. The process of removing this redundancy efficiently is called compression. Compression can either be lossless or lossy. Depending on the application, the appropriate scheme is used. Lossless compression is preferred for archival purposes and often medical imaging, technical drawings, clip art or comics. This is because lossy compression methods, especially when used at low bit rates, introduce compression artifacts.

Thus, by using the image compression techniques, lossy or lossless compression of images can be obtained which significantly reduces the amount of storage required and makes communication easier. These images can be reconstructed later to obtain the original size.

1.4 Theory:

In this project, Image Compression is being implemented using (1) Discrete Cosine Transform, (2) Sub band Wavelet Transform (16 band dyadic and 22 band modified pyramid).

1.4.1 JPEG compression using Discrete Cosine Transform (DCT):

JPEG:

JPEG is one of the most popular image compression techniques and is generally referred to as a lossy algorithm or JPEG baseline algorithm. The baseline algorithm is capable of compressing continuous tone images to less that 10% of their original size without visible degradation of the image quality [2].

Algorithm:

The JPEG lossy compression algorithm consists of three successive stages: DCT Transformation, Coefficient Quantization, and Lossless Compression [3].

The key to the JPEG baseline compression process is a mathematical transformation known as the Discrete Cosine Transform (DCT). The basic purpose of these operations is to take a signal and transform it from one type of representation to another. For example, an image is a two-dimensional signal that is perceived by the human visual system. The DCT can be used to convert the signal (spatial information) into numeric data ("frequency" or "spectral" information) so that the image's information exists in a quantitative form that can be manipulated for compression [3].

The purpose of the DCT transformation phase is to *identify* "pieces of information in the image's signal that can be effectively 'thrown away' without seriously compromising the quality of the image". No information is lost, nor is any compression achieved, in the DCT stage [3].

1.4.2 Discrete Cosine Transform:

A discrete cosine transform (DCT) expresses a finite sequence of data points in the terms of the sum of cosine functions oscillating at different frequencies. DCTs are important to numerous applications in the science and engineering, from the lossy compression of audio (e.g. MP3) and images (e.g. JPEG) (where small high-frequency components can be discarded), to spectral methods for the numerical solutions of partial differential equations (PDEs).

DCT logically follows that the horizontally oriented set of basis functions represent horizontal frequencies and the other set of basis functions represent vertical frequencies [4].

In image applications, DCT separates images into parts of different frequencies where less important frequencies are discarded through quantization and important frequencies are used to retrieve the image during decompression [4].

The forward 2D-DCT transformation is given by the following equation:

$$B_{pq} = \alpha_p \alpha_q \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} A_{mn} \cos \frac{\pi(2m+1)p}{2M} \cos \frac{\pi(2n+1)q}{2N}, \quad \begin{matrix} 0 \le p \le M-1 \\ 0 \le q \le N-1 \end{matrix}$$

The Inverse 2D-DCT transformation is given by the following equation:

$$A_{mn} = \sum_{p=0}^{M-1} \sum_{q=0}^{N-1} \alpha_p \alpha_q B_{pq} \cos \frac{\pi(2m+1)p}{2M} \cos \frac{\pi(2n+1)q}{2N}, \quad \begin{matrix} 0 \le m \le M-1 \\ 0 \le n \le N-1 \end{matrix},$$

Where,

$$\alpha_p = \begin{cases} \frac{1}{\sqrt{M}}, & p = 0 \\ \sqrt{\frac{2}{M}}, & 1 \le p \le M-1 \end{cases}$$

and

$$\alpha_q = \begin{cases} \frac{1}{\sqrt{N}}, & q = 0 \\ \sqrt{\frac{2}{N}}, & 1 \le q \le N-1 \end{cases}$$

M and *N* are the row and column size of A, respectively.

In JPEG image compression, the two-dimensional DCT of N x N blocks are computed, and the results are quantized, and entropy coded. In this case, N is typically 8 and the DCT-II formula is applied to each row and the column of the block. The result is an 8 x 8 transform coefficient array in which the element (top-left) is the DC (zero-frequency) component and the entries with the increasing vertical and horizontal index values represent higher vertical and horizontal spatial frequencies .

Quantization is achieved by compressing a range of values to a single quantum value. When the number of discrete symbols in a given stream is reduced, the stream becomes more compressible. A quantization matrix is used in combination with a DCT coefficient matrix to carry out transformation.

Quantization is the step where most of the compression takes place. DCT really does not compress the image because it is almost lossless. Quantization makes use of the fact that higher frequency components are less important than low frequency components. It allows varying levels of image compression and quality through selection of specific quantization matrices [3].

In our method, the DCT values of 512 x 512 pixels are leveled between 0 to 1023 levels to quantize them to a 10-bit uniform Scalar.

Our Method:

1. The Original image is segmented into 8 x 8 pixels sub-blocks.
2. 2D-DCT matrix is calculated for each sub-block.
3. One-quarter, one-half and three-quarters of the coefficients in each block are set to zero (High frequencies).
4. Remaining coefficients are quantized using a 10-bit uniform scalar quantizer.
5. One quarter, one-half and three-quarters of the coefficients of these quantized matrices are again set to zeros. (As the process of quantization may have replaced the zero coefficients).
6. Image is reconstructed by performing the Inverse DCT on the quantized matrices.

Finally, Signal to Noise Ratio (PSNR) between the original image and the images obtained from the One-quarter, one-half and three-quarter zeros of the DCT coefficients, is calculated to compare the quality of reconstruction.

1.4. 3 Wavelet Image Compression:

Wavelet Transform has become an important method for image compression. Wavelet based coding provides substantial improvement in picture quality at high compression ratios mainly due to better energy compaction property of wavelet transforms.

Sub-band coding:

A signal is passed through a series of filters to calculate DWT. Procedure starts by passing this signal sequence through a half-band digital low pass filter with impulse response h(n). Filtering of a signal is numerically equal to the convolution of the tile signal with impulse response of the filter.

$$x[n]*h[n]= x[k].h[n-k]$$

A half band low pass filter removes all frequencies that are above half of the highest frequency in the tile signal. Then the signal is passed through high pass filter. The two filters are related to each other as

$$h[L-1-n] = (-1)^n g(n)$$

Filters satisfying this condition are known as quadrature mirror filters. After filtering, half of the samples can be eliminated since the signal now has the highest frequency as half of the original frequency. The signal can therefore be subsampled by 2, simply by discarding every other sample. This constitutes the first level of decomposition and can mathematically be expressed as

$$y1[n]= x[k]h[2n-k]\ k$$

$$y2[n]= x[k]g[2n+1-k]\ k$$

where y1[n] and y2[n] are the outputs of low pass and high pass filters, respectively after subsampling by 2.

This decomposition halves the time resolution since only half the number of samples now characterizes the whole signal. Frequency resolution has doubled because each output has half the frequency band of the input [5]. This process is called as sub-band coding. It can be repeated further to increase the frequency resolution as shown by the figure 1.

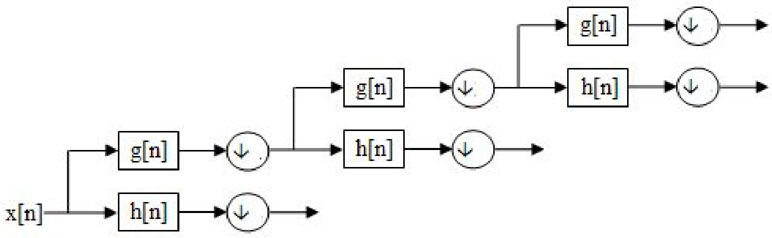

Fig 1. Example of a Filter Bank

Our Method:

1. The original image is decomposed using the 16-band dyadic decomposition and 22-band modified pyramid decomposition. 9-7 tap biorthogonal wavelet analysis filters are used to obtain the sub-bands and symmetric extension is used to eliminate the edge effects [6].
2. For the dyadic case, the highest frequency, three highest frequency and six highest frequency sub-bands are set to zeros. For the 22-band modified, three highest frequency, 10 highest frequency, 15-highest frequency sub-bands are set to zeros [6].
3. The resultant matrices with the remaining coefficients are quantized using a 10-bit uniform scalar quantizer (as mentioned in 1.4.2).
4. Image is reconstructed from the quantized matrices by using the 9-7 tap biorthogonal wavelet synthesis filter.

Finally, the Signal to Noise Ratio (PSNR) between the original image and the images obtained from the sub-band coding, is calculated to compare the quality of reconstruction.

2. Results- DCT:

Implementation:

For implementing the DCT, firstly Lina image has been extracted as a 2D array for processing it further. Calculation of the DCT has been done block-wise using the function 'dct_test' which is written from scratch to compute the 2D DCT. Then, the quarter, half and threequarters of DCT matrix have been set to zero using the function 'makezeros' which accesses the DCT matrix block by block and sets the assigned pixels to zero.

As the DC values of the image frequencies consists of huge information, they are extracted separately and are quantized to 10-bit. The same is repeated with the rest of the higher frequencies and a quantization matrix is obtained. Due to quantization, the zeroed elements will also be replaced with quantized values. Hence quarter, half and three-quarter portion of the quantized matrices have been set to zero again. Image is reconstructed by calculating the Inverse DCT of the quantized matrices.

In the later part, the Signal to Noise Ratio of the obtained images are calculated using "calc_psnr" function to compare the quality of reconstruction. And, the 2D DCT magnitude spectrum plots of the reconstructed images have been plotted.

Also, Number of bits required per pixel calculated below shows a significant reduction of data storage when few of the pixel values are made to zero.

Calculations:

Total number of bits used to code the image: 512*512*10 = 26,21,440 bits

Average number of bits per pixel:

1. One-quarter zero:
   Each 8 x 8 block has 64-pixel values. When One quarter of them are made to zero, 48 are the remaining non-zero-pixel values. And, there are 64 * 64 such 8 x 8 blocks. Therefore,
   Average number of bits/pixel = (48 * 10 * 64 * 64) / (512 * 512) = 7.5 ~ 8 bits/pixel

2. One-half zero:

   Each 8 x 8 block has 64-pixel values. When One half of them are made to zero, 32 are the remaining non-zero-pixel values. And, there are 64 * 64 such 8 x 8 blocks.
   Therefore,
   Average number of bits/pixel = (32 * 10 * 64 * 64) / (512 * 512) = 5 bits/pixel

3. Three-Quarters zero:

   Each 8 x 8 block has 64-pixel values. When One half of them are made to zero, 16 are the remaining non-zero-pixel values. And, there are 64 * 64 such 8 x 8 blocks.
   Therefore,
   Average number of bits/pixel = (16 * 10 * 64 * 64) / (512 * 512) = 2.5 ~ 3 bits/pixel

Observations:

PSNR Values:

1. PSNR of Original image to the reconstructed image from One-Quarter zero:
   Obtained PSNR = 42.9047
2. PSNR of Original image to the reconstructed image from One-half zero:
   Obtained PSNR = 39.2955
3. PSNR of Original image to the reconstructed image from Three-Quarters zero:
   Obtained PSNR = 34.5596

Quality of the reconstructed image:

From the PSNR values observed and the obtained DFT magnitude spectrum, it is clear that the best reconstruction has been obtained using the three-quarter zero matrix. This can also be told because, as three-quarter of the DFT values are made to zero, the low frequency values, which consists of most of the information are retained and help reconstruct the image appropriately. While in the other cases, there are still some high frequencies which act as noise elements and reduce the reconstructed image quality.

Observations on 2D- DFT Magnitude Spectrum of the reconstructed images:

The 2D-DFT Magnitude Spectrum of the original and the reconstructed images are as shown below. As observed in the plots, all the reconstructed images and their 2D DFT Magnitude spectrum look mostly the same. This is because, even if the quarter, half and three-quarter of the values of DCT are made to zero, as the most significant information lies in the DC and low frequency range. Hence, reconstruction is quite good and appropriate. Among the obtained reconstructions, their PSNR and the Magnitude Spectrum shows that the image reconstructed from three-quarter zero matrix is of the best quality.

2-D magnitude Spectrum of each reconstructed images:

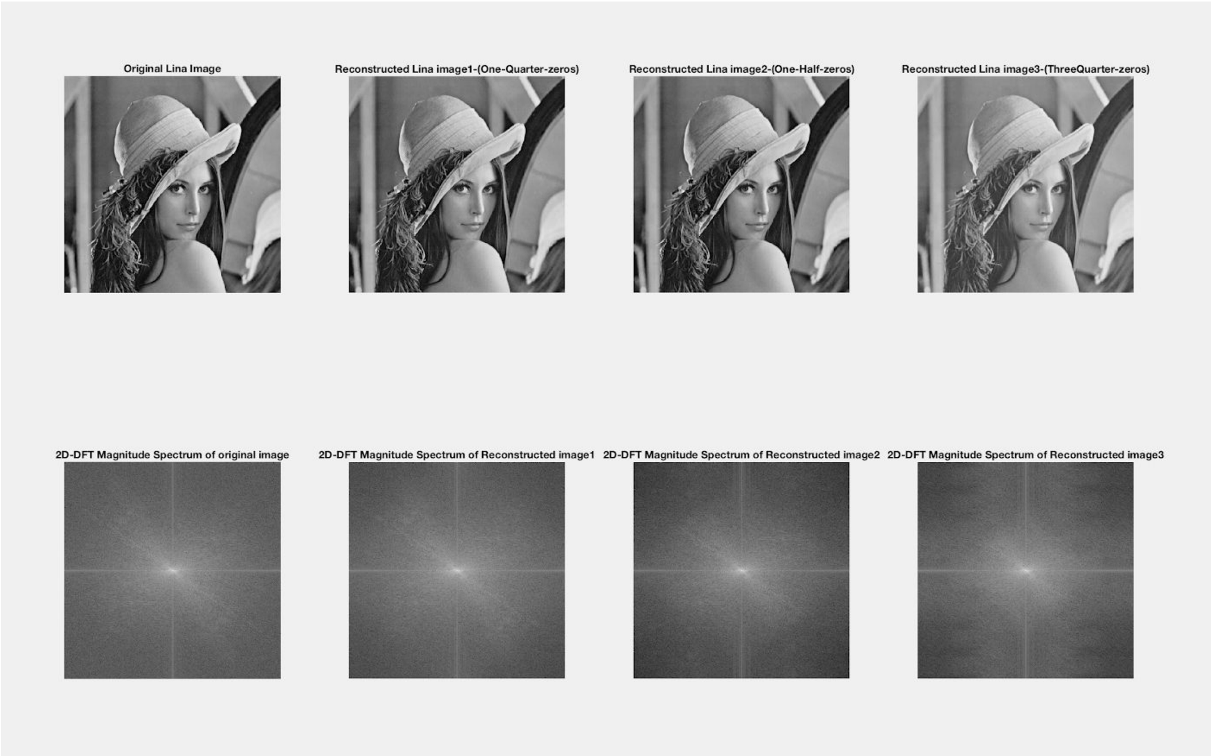

Fig 2: 2D DCT -Original and Reconstructed images and their Magnitude Spectrum

3. Results- SWT:

Implementation:

First, the analysis and synthesis filter coefficients for the 9-7 tap biorthogonal wavelet filter are obtained from [7]. Using these filter coefficients the analysis and synthesis filters are defined, wherein firstly, symmetrical extension is applied on the image to eliminate the edge effects.

To explain in detail, as in this method, filter is being applied or the convolution of filter with the image is being performed row-wise first and then column-wise. As in any convolution operation, filter is first flipped and then applied on each pixel, at the edge pixels, the filter length might exceed the normal image length. Hence to avoid such uncertainties, before each time the filtering is performed, symmetric extension of the image is done. Accordingly, sub-bands are extracted by applying the filter to the respective sub regions of the image.

After obtaining the sub-bands using the two methods, the highest-frequency, three-highest frequency and six-highest frequency sub-bands of the 16-band dyadic are made zeros. For, 22-band modified pyramid, three-highest frequency, 10-highest frequency and 15-highest frequency sub-bands to zeros.

After setting the above mentioned sub-bands to zeros, the remaining sub-bands are quantized using 10-bit uniform scalar and are made to zeros again. Thus obtained matrix is passed through the synthesis filter to reconstruct the image.

In the later part, the Signal to Noise Ratio of the obtained images are calculated to compare the quality of reconstruction. And, the 2D SWT, 2D DFT magnitude spectrum plots of the reconstructed images have been plotted.

Observations:

PSNR Values for 16-band dyadic Sub-band:

1. PSNR of Original image to the reconstructed image from highest-frequency zero:
   Obtained PSNR = 12.2624
2. PSNR of Original image to the reconstructed image from Three-highest frequency zero:
   Obtained PSNR = 12.2913
3. PSNR of Original image to the reconstructed image from Six-highest frequency zero:
   Obtained PSNR = 12.2541

PSNR Values for 22-band modified Sub-band:

1. PSNR of Original image to the reconstructed image from Three highest-frequency zero:
   Obtained PSNR = 11.7084
2. PSNR of Original image to the reconstructed image from Ten-highest frequency zero:
   Obtained PSNR = 11.7708
3. PSNR of Original image to the reconstructed image from 15-highest frequency zero:
   Obtained PSNR = 11.6891

Quality of the reconstructed image:

Although my code has some errors with the reconstruction, from the observed PSNR values, it can be concluded that reconstruction obtained from the Six-Highest frequency for 16 band and the 15-highest frequency for 22 band are the better reconstructions than the others.

Observations on 2D- DFT Magnitude Spectrum of the reconstructed images:

From Fig 3 and 4 mentioned below, it can be inferred that the low and high frequencies are being distinguished well but due to some disturbance in the recovered images, the DFT magnitude spectrum is also scattered. The reconstruction images are to be obtained accurately to conclude the performance.

2-D magnitude Spectrum of each reconstructed images:

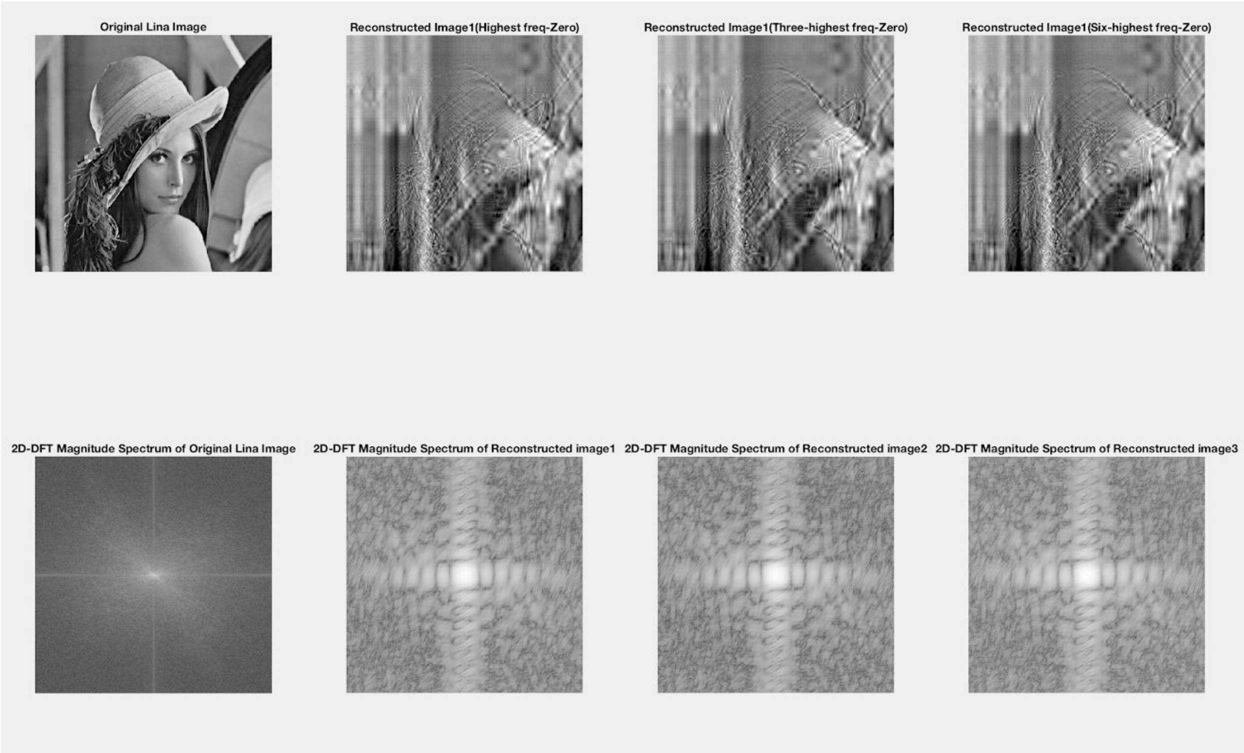

Fig 3: 22-Band- Original and Reconstructed images and their Magnitude Spectrum

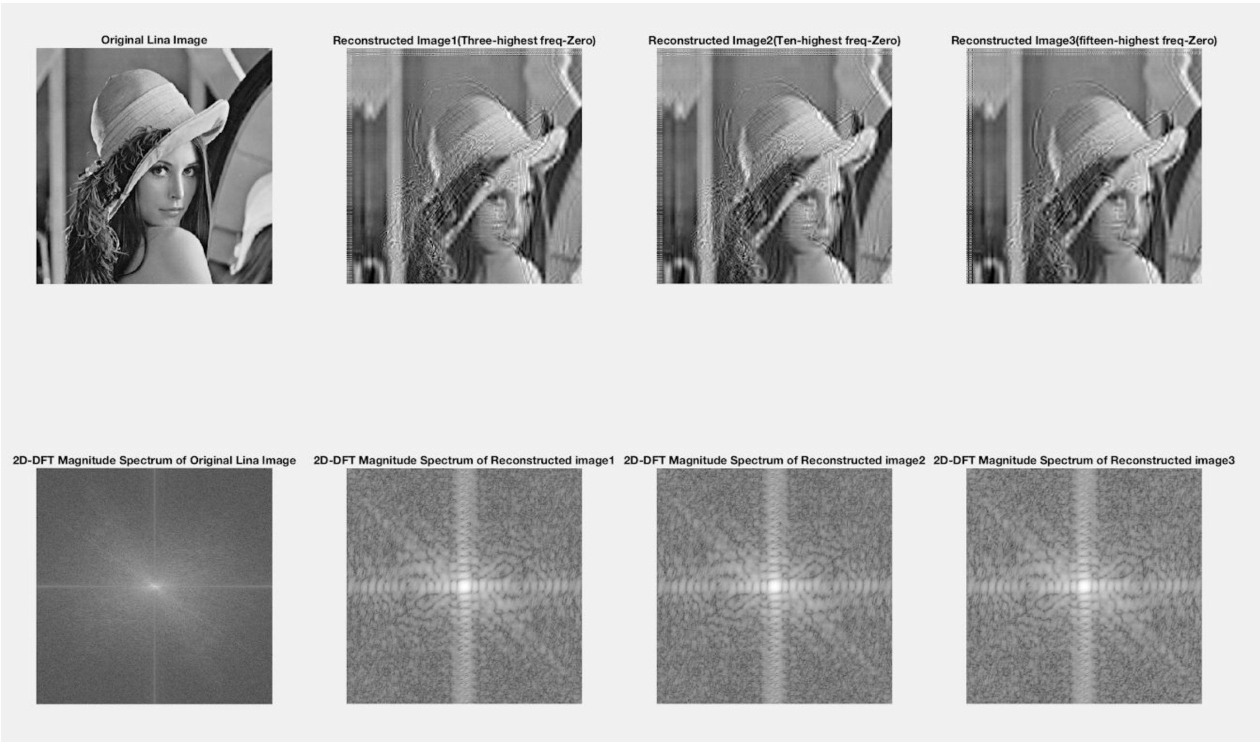

Fig 4: 22-Band- Original and Reconstructed images and their Magnitude Spectrum

4. Comparisons and Observations:

Image reconstruction has been performed on Lina image, using the 2D DCT, 16-band SWT and 22-band SWT. After comparing the 3 methods, I found that the DCT is computationally less costly and is quickly adaptable. In comparison, SWT requires much higher computations for the same purpose. Due to some coding errors, the reconstructed images obtained from DCT are much better in my case. But if executed properly, SWT gives an appreciable image compression than that of the DCT. Also, between 16 band and 22 band, the 22 band modified pyramid gives better results than the 16 band as more lower frequencies are being used for the reconstruction of the image in the 16 band case.

5. References:

1. Anitha S., Image Compression Using Discrete Cosine Transform & Discrete Wavelet Transform, *International Journal of Scientific & Engineering Research*, Volume 2, Issue 8, August-2011
2. Data Compression, JPEG, The Intellectual Excitement of Computer Science, https://cs.stanford.edu/people/eroberts/courses/soco/projects/data-compression/lossy/jpeg/index.htm
3. G. K. Wallace, The JPEG Still Picture Compression Standard, *IEEE Transactions on Consumer Electronics*, Vol. 38, No. 1, February 1992
4. G. Strang, The Discrete Cosine Transform, *Society for Industrial and Applied Mathematics Review.*, Volume 41, Issue 1, 135–147, 1999
5. S. A. Martucci, “Signal extension and noncausal filtering for subband coding of images,” in *Conf. Proc., 1991 SPIE Conf. Visual Commun. Image Processing*, Boston, MA, Nov. 1991.
6. A. Said, A., W. A. Pearlman, A new, fast, and efficient image codec based on set partitioning in hierarchical trees, *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 63, pp. 243-250, June 1996
7. M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, “Image coding using wavelet transforms,” *IEEE Trans. Image Processing*, vol. 1, pp. 205-220, Apr. 1992.