Analysis of Image Compression Methods using DCT and Wavelet Transform in Matlab

Ashwani Kr. Yadav¹, Dr. R. D. Roy², Vaishali³
¹Lecturer, Amity School of Engineering
²Assistant Professor, Amity School of Engineering
³Amity University Rajasthan, Jaipur, India

Abstract: Image compression reduces the amount of data required to represent an image. To compress an image efficiently we use various techniques to decrease the space and to increase the efficiency of transfer of the images over network for better access. This paper explains about compression methods such as Discrete Cosine Transform (DCT), Discrete Wavelet Transform (DWT), JPEG 2000, EZW and SPIHT (Set Partition in Hierarchical Trees) on the basis error comparison, mean square error, peak signal to noise ratio and compression ratio. Provides substantial improvements in the quality of picture because of multi resolution nature. DWT is multi resolution in nature. So, it provides substantial improvements in the quality of picture. Image compression reduces the memory storage space of image and also maintains the quality information of the image. In this paper we have analysed some wavelet based compression techniques like Embedded Zero tree Wavelet (EZW), Set Partitioning in Hierarchical Trees (SPIHT) and JPEG2000. These techniques are analysed on lena image. The performance evaluation is made on the basis of parameters like PSNR, MSE and CR etc.

Keywords: DCT, DWT, EZW, JPEG 2000 and SPIHT.

1. Introduction

Images contain large amounts of information that requires much storage space, transmission bandwidths (larger) and transmission times (higher). So, compression is advantageous to compress the image by storing only the essential information needed to reconstruct the original image. An image can be thought of as a intensity values. In order to compress the image, redundancies must be utilise, for example, areas where there is little or no change between pixels. Therefore images having large areas of uniform colour will have large redundancies, and conversely images that have frequent and large changes in colour will be less redundant and difficult to compress. Wavelet analysis is used to divide the information of an image into approximation and detail sub-signals. The approximation sub-signal shows the general trend of pixel values, and three detail sub-signals show the vertical, horizontal and diagonal details of the image. If the details are very small then they can be set to zero without any change in the image. The value below which details are considered small enough to be set to zero is known as the threshold. The greater the number of zeros greater compression can be achieved. The amount of information retained by an image after compression and decompression is known as the energy retained. Energy retained is proportional to the sum of the squares of the pixel values. If the energy retained is 100% then the compression is known as lossless. So, the image can be reconstructed exactly. This occurs only when the threshold value is set to zero, i.e. the detail has not been changed. If any values are changed then energy will be lost and this is known as lossy compression. Ideally, during compression the number of zeros and the energy retention will be as high as possible. However, as more zeros are obtained higher energy loss occurs. So a balance between the two needs should be there. Discrete Wavelet Transform (DWT) can be efficiently used in Image Coding Applications because of their data reduction capabilities. Unlike the case of Discrete Cosine Transform (DCT), DWT can be composed of any function wavelet that satisfies the concept of multiresolution analysis [3].

1.1 The Need for Wavelets

Often signals to be process are in the time-domain, but in order to process them more easily other information is required, such as frequency [4]. The analogy cites the problem of multiplying two roman numerals. In order to do this calculation we would find it easier to first translate the numerals in to number system, and then translate the answer back into a roman numeral. Result is same, but taking the detour into an alternative number system made the process easier and quicker. Similarly we can take a detour into frequency space to analysis or process a signal.

A. Wavelets Definition

Wavelets are mathematical functions, wavelets cut up data into different frequency components. A wavelet is a wave-like oscillation with an amplitude that begins at zero, increases, and then decreases back to zero. The fundamental idea behind wavelets is to analyze the signal at different scales or resolutions, which is called multiresolution.
B. Discrete Cosine Transform

A DCT represents the input data points in the form of sum of cosine functions that are oscillating at different frequencies and magnitudes. There are mainly two types of DCT: one dimensional DCT and two dimensional DCT. The 2D DCT for an NxN input sequence can be defined as follows [3]:

$$d_{dct}(i,j) = \frac{1}{\sqrt{2N}} B(i)B(j) \sum \sum M(x,y) \cdot \cos \left( \frac{2x+1}{2N} i \pi \right) \cos \left( \frac{2y+1}{2N} j \pi \right)$$

(1)

Where B (u) = 1 \sqrt{2} if u = 0 and 1 if u > 0. M (x,y) is the input data of size xxy. The input image is first divided into 8x8 blocks; then the 8-point 2-D DCT is performed. The DCT coefficients are then quantized using an 8x8 quantization table. The quantization is achieved by dividing each element of the transformed original data matrix by corresponding element in the quantization matrix Q and rounding to the nearest integer value as shown in equation (2):

$$D_{quant}(i,j) = round \left( \frac{d_{dct}(i,j)}{Q(i,j)} \right)$$

(2)

After this, compression is achieved by applying appropriate scaling factor. Then in order to reconstruct the data, rescaling and de-quantization is performed. The de-quantized matrix is then transformed back using the inverse - DCT.

C. Discrete Cosine Transform from Subband Computation

Let us present here briefly the computation technique for DCT’s of an image from its subbands from [5]. It has been demonstrated that image compression carried out on the subimages obtained by a subband decomposition can be more effective than compressing the full band image. DCT computation scheme consisting of two parts: A subband decomposition of the input data followed by the computation of the DCT coefficients of the input data from the subband samples. As it will be shown, this approach provides a simultaneous solution to both block artifact reduction and fast computation. The definition of DCT for a 2D images x(m,n) of size N×N is as follows:

$$C(k,l) = \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \left[ 4x(m,n)\cos \left( \frac{2m+1}{2N} k \pi \right) \cos \left( \frac{2n+1}{2N} l \pi \right) \right]$$

(3)

The low-low subband x_{LL} (m,n) of the image be obtained as:

$$x_{ll}(m,n) = \frac{1}{4} \{x(2m,2n) + x(2m+1,2n) + (2m,2n+1) + x(2m+1,2n+1)\}$$

(4)

$$0 \leq m, n \leq \frac{N}{2} - 1.$$

Let C_{LL} (k,l), 0 < k,l < N/2 be the 2D DCT of x_{LL}(m,n). Then the subband approximation of DCT of x(m,n) is given by:

$$C(k,l) = \left\{ \begin{array}{ll} 4\cos \left( \frac{nk}{2N} \right) \cos \left( \frac{nl}{2N} \right) C_{LL}(k,l), & k,l = 0,1,\ldots, \frac{N}{2} - 1 \\ 0, & \text{otherwise} \end{array} \right.$$  

(5)

According to the definition of DCT, subband DCT’s are multiplied by a factor (in this case $4\cos \left( \frac{nk}{2N} \right) \cos \left( \frac{nl}{2N} \right)$). The definition of inverse DCT (IDCT) should also be modified accordingly. Refer this as subband approximation of DCT. On the other hand, the approximation as carried out in the approach proposed by Dugad and Ahuja [4], is as follows:

$$C(k,l) = \left\{ \begin{array}{ll} 4C_{LL}(k,l), & k,l = 0,1,\ldots, \frac{N}{2} - 1 \\ 0, & \text{otherwise} \end{array} \right.$$  

(6)

We refer this approximation as the low-pass truncated approximation of DCT. Consider, the multiplication factor 4 in Eq. (6) appears due to the definition of DCT used (Eq. (3)). Yet, this factor does not have any effect in the final results obtained by them (PSNR values (halved) and then (doubled) images in [5]). While halving an image, DCT coefficients for N/2-point DCT are obtained by dividing the N-point DCT coefficients with the factor. Afterwards, during image doubling, the N-point DCT coefficients are obtained by multiplying the N/2-point DCT coefficients with the same factor. For the better knowledge about subband computation of DCT, the readers are requested to study the work presented in paper [3].

1.2 Discrete Wavelet Transform

In DWT, an image is represented by sum of wavelet functions, which are known as wavelets, having different location and scale. Discrete Wavelet Transform represents the data into a set of high pass (detail) and low pass (approximate)
coefficients. Image is first divided into blocks of 32×32. Then each block is passed through two filters: in this the first level, decomposition is explained to breakdown the input data into an approximation and detail coefficients. With the help of transformed matrix, detail coefficients and approximate coefficients are separated as LL, HL, LH and HH coefficients. After that all the coefficients are discarded, except the LL coefficients. LL coefficient transformed into the second level. These coefficients are then passed through a constant scaling factor to achieve the desired compression ratio. Following fig. 2 is an illustration of DWT. Here, X[n] is the input signal, D[n] is the high frequency component, and A[n] is the low frequency component. To reconstruct the data, the coefficients are rescaled and with zeros, and passed across the wavelet filters. Daubechies filters coefficients are considered [6]:

![Figure 1: Block diagram of the 2-level DWT](image)

2. Baseline JPEG

The JPEG compression standards in detail. Figure 2 describes the JPEG compression process. In this process the image divides into 8 by 8 pixel blocks, after that calculates the discrete cosine transform (DCT) of each block. A quantizer rounds off the DCT coefficients according to the quantization matrix (can be image dependent). This step produces the “lossy” nature of JPEG, but good for large compression ratios. JPEG’s compression technique uses a variable length code such as run length codes or LZW on these coefficients. For decompression, it retain the quantized DCT coefficients from the compressed data stream, takes the inverse transform and shows the image. The basic idea behind this scheme is that the discrete cosine transform provides good energy compaction properties and hence we would require less number of bits to quantize and store these coefficients. Further, the quantized AC coefficients are mostly zeros and therefore can be run-length coded to give better compression ratios[4]. The JPEG method is a sophisticated implementation of block Discrete Cosine Transform encoding [12].

![Figure 2: The Baseline JPEG algorithm](image)
3. **JPEG-2000 Image Coding**

JPEG 2000 is the international standard for still images. This is the enhancement to the current JPEG system. Compression of images will be implemented in a new way with the help of JPEG 2000 based on the wavelet transform. This supports lossy and lossless compression of grayscale as well as color images. Here we have 2 processes encoding and decoding process.

**A. Encoding Process**
The input image is given to preprocessor, in this step tiling is done. Tiling is nothing but dividing the image into rectangular and non-overlapping tiles of equal size and then each tile is compressed independently using its own set of specified compression parameters. The preprocessed image is given to DWT here wavelet decomposition is done. The output of DWT is given to uniform quantizer which reduces precision of sub band coefficients so that fewer bits will be needed to encode transformed coefficients. Then output of quantizer is given to embedded block coding which uses wavelet transform to generate the sub band. This sub band is partitioned into small blocks of samples called code blocks. EBCOT generates a separate highly scalable bit stream for each code block. Next is the rate control block which adjusts the quantizer step size or discards some coding pass information.

**B. Decoding Process**
Decoder performs opposite to encoder. The output of encoder i.e. the code stream is received by the decoder. The coefficients in packets are decoded, then dequantized and inverse ICT is performed to get reconstructed image. Important point of jpeg 2000 is here compression can be done in one way but decompression is done in many ways.

4. **Embedded zero tree wavelet coding**

EZW stands for embedded zero tree wavelet. An EZW encoder is specially designed to use with wavelet transforms. The is based on progressive coding to compress an image into bit stream with increasing accuracy so when more bits are added to a bit stream image will be more detailed. The Embedded Block coding algorithm as described in JPEG2000 [18] was studied.

**Approach**
Before the processing of image data the image is preprocessed. In preprocessing step tiling on the original image is done. All operations, including component mixing, wavelet transform and quantization are performed independently on the image tiles. Tiling reduces memory requirements, and as they are also reconstructed independently, instead of the whole image they can be used for decoding specific parts of the image. All tiles have almost the same dimensions, except some, only at the boundary of the image. This unit transforms the input image from time domain to frequency domain and decomposes the original image into its fundamental components. The wavelet transform uses filter banks for the decomposition of preprocessed image. The Embedded ZeroTreeWavelet (EZW) encoder encodes the decomposed image by recognizing the priority of decomposed image pixel. Initial threshold for coding calculated by the encoder module given by T0= 2(log2cmax). The encoding process is performed using 2 passes namely
1. dominant pass and
2. Subordinate pass.

The dominant pass generates any one of four possible combinations like they are significant positive (SP), significant negative (SN), isolated zero (IZ) and zerotree root (ZR). Subordinate pass where the coefficients are encoded as 0 or 1 depending on the current threshold. The decoding unit reconstructs the values by identifying the symbols as positive, negative, zero tree and isolated zero tree. Inverse transformation is the process of retrieving back the image data from the obtained image values. The image data transformed and decomposed under encoding side is rearranged from higher level decomposition to lower level with the highest decomposed level been arranged at the top. Fig 5 shows the reconstruction of the obtained decomposed component.

4.1 **The Algorithm**

1. Set the initial threshold T0= 2(log2cmax). Here xmax is the maximum coefficient value.
2. set k=0 3. Conduct dominant pass by scanning the data. Outputs will be any of 4 conditions below i.e.
   a. If value of coefficient is greater than the threshold and the value is positive, it means the output significant positive.
   b. If value of coefficient is greater than the threshold and the value is negative, it means the output significant negative.
   c. If magnitude of coefficient is less than the threshold and all its descendants have magnitudes less than the threshold then coefficient is labeled as zero tree roots.
   d. If magnitude of coefficient is less than the threshold and all its descendants have values greater than the threshold then coefficient is labeled as isolated zero.
3. Conduct a subordinate pass or refinement pass by scanning through the data to refine the pixels already known to be significant in current bitplane.
4. Set k=k+1 and the threshold Tk=Tk-1/2.
5. Stop if the stopping criterion is met or go to step 3.

5. SPIHT Algorithm

Set Partitioning in Hierarchical Trees (SPIHT) is a wavelet based image compression method. SPIHT introduces three lists: a. List of Significant Pixels (LSP), b. List of Insignificant Pixels (LIP) and c. List of Insignificant Sets (LIS). First initialization is done, and then algorithm takes two stages for each level of threshold 1. The sorting pass (in which lists are organized) and 2. The refinement pass. LIS is further divided into two types of sets of insignificant pixels. They are Type A (all descendant are zero) Type B (all grandchildren and further descendants are zero). SPIHT algorithm defines four types of sets, which are sets of coordinates of coefficients: O(i,j); set of coordinates of all offspring of node (i,j); children only D (i,j): set of coordinates of all descendants of node (i,j); children, grandchildren, great-grand, etc. H (i,j): set of all tree roots (nodes in the highest pyramid level); parents L (i,j): D (i,j) – O(i,j) (all descendants except the offspring). To find the number of passes we use the formula n = \[ \log2cmax \]. We find initial threshold as T0=2n. The SPIHT algorithm forms a hierarchical quad tree data structure for the wavelet transformed coefficients. The set of root node and corresponding descendants are together called as spatial orientation tree (SOT). The method that SPIHT uses for arithmetic compression is quite involved and space does not permit a discussion of the details here. Some details are provided in [9].

6. Performance evaluation parameters

Two popular measures of performance evaluation are, Peak Signal to noise Ratio (PSNR) and Compression Ratio (CR). Which are described below: A. PSNR It is the most popular tool for the measurement of the compressed image and video. It is simple to compute. The PSNR in decibel is evaluated as follows [8]:

\[
PSNR = 10 \log_{10} \frac{F}{MSE} \quad (7)
\]

Where, F is allowable image pixel intensity level. MSE is mean squared error. It is another performance evaluation parameter of Image Compression Algorithms. It is an important evaluation parameter for measuring the quality of compressed image. It compares the original data with reconstructed data and then results the level of distortion. The MSE between the original data and reconstructed data is:

\[
MSE = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} (A_{ij} - B_{ij})^2 \tag{8}
\]

Where, A = Original image of size MxN
B = Reconstructed image of size MxN B. CR It is a measure of the reduction of detail coefficient of data.
CR = Discarded Data / Original Data

In the image compression process, it is important to know how much important coefficient one can discard from input data in order to preserve critical information of the original data.

7. The Wavelet Transform

Wavelet based techniques for image compression have been increasingly used for image compression. The wavelet uses subband coding to selectively extract different subbands from the given image. These subbands can then be quantized with different quantizers to achieve good compression. Specifically designed wavelet filters are used to satisfy certain constraints called the smoothness constraints. The wavelet filters are designed in such a way, so that the coefficients in each subband are almost uncorrelated from the coefficients in other subbands [8]. The wavelet transform achieves better energy compaction than the DCT and hence can help in providing better compression for the same Peak Signal to Noise Ratio (PSNR). A lot of research has been done on the performance comparison of the DWT and DCT for image compression. A comparative study of DCT and wavelet based image coding can be found in [9].

The Embedded Zerotree Wavelet or popularly known as EZW is an efficient coding scheme developed by Shapiro [12]. The author in this landmark paper introduces an efficient coding scheme based on the multi-resolution nature of wavelet transforms. The resulting algorithm gave a better results at low bit rates over the then other schemes. The EZW is considered as a beginning of a new era of wavelet coding and ignited a lot of research work in this field. Significance map coding and successive approximation quantization are the two important features of the EZW coding. This algorithm exploits the energy compaction properties and the self-similar and hierarchical nature of the wavelet transform. The hierarchical nature facilitates coding as it forms a tree structure. Inter band prediction is used to code the positions of the significant coefficients. The EZW algorithm does not code the location of significant coefficients but instead codes the location of zeros. The EZW algorithm was further extended by Amir et. al to give a new scheme.
called the Set Partitioning in Hierarchical Trees (SPIHT) [11]. SPIHT gave better results than the EZW without using the arithmetic encoder and so the algorithm was computationally more efficient. Extensive research has shown that the images obtained with wavelet-based methods yield very good visual quality. At first it was shown that even simple coding methods produced good results when combined with wavelets. SPIHT belongs to the next generation of wavelet encoders, employing more sophisticated coding. In fact, SPIHT exploits the properties of the wavelet-transformed images to increase its efficiency. Many researchers now believe that encoders that use wavelets are superior to those that use DCT or fractals.

8. Results and performance comparison

It is observed that with the limitations of our implementation that the Baseline JPEG performs better than the DCT-Based Embedded Image coding. The Wavelet based EZW scheme performs (for the same compression ratio) better than the corresponding DCT based version. A performance improvement due to wavelet based techniques over DCT based schemes has been shown in various works [5, 8]. This improvement can be attributed to the better energy compaction properties of the wavelet transform compared to the discrete cosine transform.

8.1 The effect of transforms used

The effect of wavelet filters was also studied. The results indicate that the PSNR obtained in each of the cases is almost the same for a given compression ratio. Figure 4 shows the difference between them. Where figure 3 shows the wavelet transform of the lena image and inverse of the thresholded wavelet transform.

Figure 3: (a) Lena image, 8 bpp. (b) Wavelet transform of image, threshold = 8. (c) Inverse of thresholded wavelet transform.
8.2. The Effect of the encoding scheme

For comparison purposes, the results of the SPIHT algorithm were obtained. It was observed that the SPIHT algorithm gave around 2 dB better performance than our corresponding implementation of the Wavelet based Embedded image coder. This can be assign better compression methods used in SPIHT. The algorithm gives comparable results with SPIHT and full-fledged JPEG2000 for compression. This confirms the correctness of the implementation. In Figure 5 SPIHT compression of lena image shown at different PSNR values.

![Figure 5: SPIHT compressions of Lena image. PSNR values: (a) 27.96 dB. (b) 30.85 dB.](image)

But according to the implementation as the compression ratio increases, the quality of reconstructed image degrades as compared to that of full-fledged JPEG2000. Remove bit planes (starting from LSB) to reduce the bit rate. Alternatively the optimal contribution from each block for each plane should be considered to find which bits should be used for decoding.

Finally, it is interesting to compare SPIHT compressions with compressions obtained with the JPEG method. It is used extensively for compression of images, especially for transmission over the Internet.
In Figure 6, we compare compressions of the Lena image obtained with JPEG and with SPIHT at two different compression ratios. (JPEG does not allow for specifying the bpp value in advance; the 59:1 compression was the closest we could get to 64:1.) It is clear from these images that SPIHT is far superior to JPEG. It is better both in perceptual quality and in terms of PSNR. Notice, actually the 59:1 JPEG compression is very distorted. The SPIHT compression, even at the slightly higher ratio of 64:1, exhibits none of unwanted features.

<table>
<thead>
<tr>
<th>Method</th>
<th>Compression Ratio</th>
<th>PSNR</th>
</tr>
</thead>
<tbody>
<tr>
<td>JPEG 2000</td>
<td>59:1</td>
<td>24:16 dB</td>
</tr>
<tr>
<td></td>
<td>32:1</td>
<td>30:11 dB</td>
</tr>
<tr>
<td>SPIHT</td>
<td>64:1</td>
<td>30:85 dB</td>
</tr>
<tr>
<td></td>
<td>32:1</td>
<td>33:93 dB</td>
</tr>
</tbody>
</table>

Table 1 shows the summary of the result shown in figure 6. DCT based implementation too show the Importance of the concept. The JPEG2000 scheme that has been implemented is better than the EZW scheme but not quite as good as the SPIHT scheme. This could mainly be because, That implementation has not been done at full functionality of the scheme. Finally table 2 shows the various methods used and their performance analysis in terms of MSE, PSNR with a particular value of CR. According to this DWT is better than DCT and SPIHT is better than EZW.

<table>
<thead>
<tr>
<th>Method</th>
<th>MSE</th>
<th>PSNR</th>
<th>Compression Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>DCT with JPEG 2000</td>
<td>0.27</td>
<td>31</td>
<td>4</td>
</tr>
<tr>
<td>DWT with JPEG 2000</td>
<td>0.26</td>
<td>34</td>
<td>4</td>
</tr>
<tr>
<td>DCT with EZW</td>
<td>0.24</td>
<td>27</td>
<td>4</td>
</tr>
<tr>
<td>DWT with EZW</td>
<td>0.23</td>
<td>32</td>
<td>4</td>
</tr>
<tr>
<td>DWT with SPIHT</td>
<td>0.23</td>
<td>35</td>
<td>4</td>
</tr>
</tbody>
</table>

9. Advantages of Image Compression

i) It reduces size of memory required for data storage.

ii) The good-quality signals for audio-visual data representation.

iii) Data security can be achieved by encrypting the decoding parameters and transmitting them separately from the compressed database.

iv) High speed input-output operations can be achieved in computing device due compressed data.

v) Image Compression reduces the cost of backup and recovery of data in computer systems.
10. Conclusion and Future Scope

The performance of algorithm with various transforms used was considered. It was observed that the wavelet transform gave a 1-2 dB performance improvement over the DCT due to its better energy compaction properties. Various methods of encoding such as Embedded Zero tree, SPIHT and JPEG2000 were considered.

The results of the above techniques EZW and SPIHT are compared by using four parameters such as CR, PSNR and MSE values from the reconstructed image. We found from our experimental results that SPIHT is better algorithm than EZW. Its results are better than EZW as we can see from Table 1-2. The above algorithms can be used to compress the image that is used in the web applications. Furthermore in future we can analyze different Image coding algorithms for improvement of different parameters.

11. References