A Comparative Study of DCT and DWT-SPIHT

Nivedita[1], Pardeep Singh[2], Sonika Jindal[3]
[1]PG Student, Shaheed Bhagat Singh College of Engineering and Technology (SBSCET), Ferozepur, Punjab(India)
nivedita.makrandi@gmail.com

[2]PG Student, Indo Global College, Mohali, Punjab(India)
pardeepsinghdhot@gmail.com

[3]Assistant Professor, Shaheed Bhagat Singh College of Engineering and Technology (SBSCET), Ferozepur, Punjab(India)
sonikamanoj@gmail.com

Abstract

Discrete Wavelet Transform (DWT) and Discrete Cosine Transform (DCT) are the common methods used in signal and image compression. Wavelet transform are very powerful compared to Fourier transform because its ability to describe any type of signals both in time and frequency domain simultaneously. In this paper, we will discuss the use of Discrete Cosine Transform (DCT) and Wavelet Based Image compression Algorithm-Set Partitioning in Hierarchical Tree (SPIHT). We do the numerical experiment by considering various types of images and by applying DCT and DWT-SPIHT to compress an image. We found that DWT yields better result as compared to DCT. For DWT, various wavelet filters such as Haar (2 filters) and Daubechies (up to 10 filters) are used. All the numerical results were done by using Matlab programming.

Keyword- Discrete Cosine Transform, Discrete Wavelet Transform, filters, Image Compression.

1. INTRODUCTION

Multimedia data requires considerable storage capacity and transmission bandwidth. The data are in the form of graphics, audio, video and image. These types of data have to be compressed during the transmission process. Large amount of data can't be stored if there is low storage capacity present. The compression offers a means to reduce the cost of storage and increase the speed of transmission. Image compression is used to minimize the size in bytes of a graphics file without degrading the quality of the image. There are two types of image compression is present. They are lossy and lossless. [1] In lossless compression, the reconstructed image after compression is numerically identical to the original image. In lossy compression scheme, the reconstructed image contains degradation relative to the original. Lossy technique causes image quality degradation in each compression or decompression step. In general, lossy techniques provide for greater compression ratios than lossless techniques i.e. Lossless compression gives good quality of compressed images, but yields only less compression whereas the lossy compression techniques [2] lead to loss of data with higher compression ratio. The approaches for lossless image compression include variable-length encoding, Adaptive dictionary algorithms such as LZW, bit-plane coding, lossless predictive coding, etc. The approaches for lossy compression include lossy predictive coding and transform coding. Transform coding, which applies a Fourier-related transform such as DCT and Wavelet Transform such as DWT are the most commonly used approach. [3] Over the past few years, a variety of powerful and sophisticated wavelet based schemes for image compression have been developed and implemented. The coders provide a better quality in the images. There are several algorithms for wavelet based compression such as Embedded Zerotree Wavelet (EZW) , Set Partitioning in Hierarchical Trees (SPHIT), Wavelet Difference Reduction (WDR), Adaptively Scanned Wavelet Difference Reduction (ASWDR) etc. [4]

In this paper, we will do comparison with discrete cosine transform (DCT) which is heart of JPEG (Joint Photographic Experts Group) standard and widely used wavelet based image compression algorithm set partitioning in hierarchical tree based on different performance measure such as Peak to Noise Ratio (PSNR), Mean Square Error (MSE) and CR.

The paper is organized as follows: Section II explains image compression; Section III explains Discrete Cosine Transform (DCT); Section IV Discrete Wavelet Transform (DWT); Section V explains Set Partitioning in Hierarchical Trees (SPHIT) Algorithm; Section V include Experiment Results and Discussion and Section VI gives the conclusion.

2. IMAGE COMPRESSION

Image compression means minimizing the size in bytes of a graphics file without degrading the quality of the image to an unacceptable level. The reduction in file size allows more images to be stored in a given amount of disk more memory space. It also reduces the time required for image to be sent over the internet or downloaded from web pages. The recent growth of data intensive multimedia based web application have not only sustained the need for more efficient ways to encode signals and images but have made compression of such signal central to storage and communication technology.

The principle behind image compression is that the neighbouring pixels are correlated and therefore contain redundant information. The foremost task then is to find less correlated representation of the image.

Two fundamental components of compression are redundancy and irrelevancy reduction.

- Redundancies reduction aims at removing duplication from the signal source (image/video).
- Irrelevancy reduction omits parts of the signal that will not be noticed by the signal receiver, namely the Human Visual System.

In digital image compression, three basic data redundancies can be identified and exploited:

- Coding redundancy
- Inter pixel redundancy
- Psycho visual redundancy

Data compression is achieved when one or more of these redundancies are reduced or eliminated.

2.1 Coding Redundancy

Use shorter code words for the more common gray levels and longer code words for the less common gray levels. This is called Variable Length Coding. To reduce this redundancy from an image we go for the Huffman technique were we are, assigning fewer bits to the more probable gray levels than to the less probable ones achieves data compression.

2.2 Inter pixel Redundancy

Another important form of data redundancy is inter pixel redundancy, which is directly related to the inter pixel correlations within an image. Because the value of any given pixel can be reasonable predicted from the value of its neighbours, the information carried by individual pixels is relatively small. Much of the visual contribution of a single pixel to an image is redundant; it could have been guessed on the basis of its neighbour's values. A variety of names, including spatial redundancy, geometric redundancy, and interframe Redundancies have been coined to refer to these interpixel dependencies.

2.3 Psycho visual Redundancy

Human perception of the information in an image normally does not involve quantitative analysis of every pixel or luminance value in the image. In general, an observer searches for distinguishing features such as edges or textural regions and mentally combines them into recognizable groupings. The brain then correlates these groupings with prior knowledge in order to complete the image interpretation process. Thus eye does not respond with equal sensitivity to all visual information. Certain information simply has less relative importance than other information in normal visual processing. This information is said to be psycho visually redundant. To reduce psycho visual redundancy we use quantizer. Since the elimination of psycho visually redundant data results in a loss of quantitative information. [5]

3. DISCRETE COSINE TRANSFORM (DCT)

The discrete cosine transform (DCT) represents an image as a sum of sinusoids of varying magnitudes and frequencies. The DCT has the property that, for a typical image, most of the visually significant information about the image is concentrated in just a few coefficients of the DCT. The DCT work by separating images into the parts of different frequencies. During a step called Quantization, where parts of compression actually occur, the less important frequencies are discarded, hence the use of the lossy. Then the most important frequencies that remain are used retrieve the image in decomposition process. As a result, reconstructed image is distorted [6]

The Process

In the DCT compression algorithm

- The input image is divided into 8-by-8 or 16-by-16 blocks
- The two-dimensional DCT is computed for each block.
- The DCT coefficients are then quantized, coded, and transmitted.
- The receiver (or file reader) decodes the quantized DCT coefficients, computes the inverse two-dimensional DCT (IDCT)of each block.
- Puts the blocks back together into a single image. [7]

The first step is to obtain matrix M from equation:

$$D(i,j) = \frac{1}{\sqrt{2N}} C(i)C(j) \sum_{x=0}^{N-1}\sum_{y=0}^{N-1} p(x,y) \cos\left[\frac{(2x+1)i\pi}{2N}\right] \cos\left[\frac{(2y+1)i\pi}{2N}\right]$$

One must use the following equation:

$$T(i,j) = \left\{ \frac{1}{\sqrt{N}} \text{ if } i = 0 \text{ and } \sqrt{\frac{2}{N}} \cos\left[\frac{(2j+1)i\pi}{2N}\right] \right\} \text{ if } i > 0$$

The next step is to apply DCT to 8X8 block. The particular block was chosen from the very upper left hand corner of an image. As DCT is designed to work on pixel values ranging from -128 to 127, the original block is levelled off by subtracting 128 from every entry. This results in the generation of matrix 'M'.

$$\mathbf{D = TMT'} \qquad (1)$$

The matrix 'M' is first multiplied on the left by the DCT matrix 'T' from the previous segment; this transforms the rows. The columns are then transformed by multiplying on the right by the transpose of the DCT matrix. This yields the matrix 'D'. The top-left co-efficient, C00, correlates to the low frequencies of the original image block. As we move away in all direction the dct coefficients correlates to high frequency of an image block where C_{77} corresponds to highest frequency. It is significant to note that the human eye is most sensitive to low frequencies.

The next step is to use the quantization technique of level Q50

$$\mathbf{C(i,j) = round} \left(\frac{D(i,j)}{Q(i,j)}\right) \qquad (2)$$

The integers situated near the upper-left corner correspond to the lower frequencies to which the human eye is most sensitive of the image block. In addition, the zeros represent the less important, higher frequencies that have been discarded, giving rise to the lossy part of compression. As mentioned earlier, only the remaining non-zero co-efficient will be used to reconstruct the image. [8]

Due to efficient energy compaction property of the DCT, many coefficients, especially the higher frequency coefficients become zero after quantization. Therefore, zigzag scanning the $N \times N$ DCT array is used to maximize the zero run-lengths. The zigzag scanning pattern used in JPEG is shown in Figure. After zigzag scanning the DCT coefficients, the DC and AC coefficients are entropy coded.

The quantized matrix is now ready for final step of compression. Before storage, all the coefficient are converted by encoder to the stream of binary data (1100011). The baseline system uses Huffman coding.

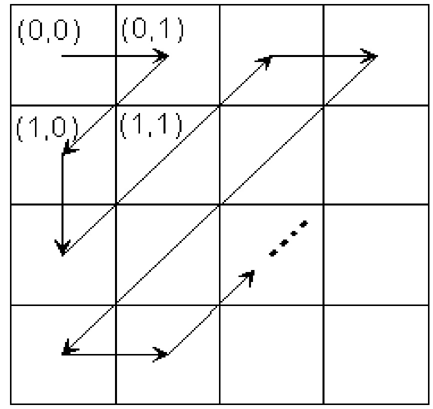

Fig1: Zig Zag Scanning

The final step of image compression using DCT is to decompose the matrix 'C' [14]

$$\mathbf{R(i,j) = Q(i,j) \; X \; C(i,j)} \qquad (3)$$

The IDCT is next applied to matrix R, which is rounded to the nearest integer. Finally, 128 is added to each element of that result, giving the decompressed version N of original 8 X 8 image block M. [9]

$$\mathbf{N = round(T'RT) + 128} \qquad (4)$$

4. DISCRETE WAVELET TRANSFORM

Wavelet transform is the latest method of compression where its ability to describe any type of signals both in time and frequency domain. JPEG2000 which is the standards of international image coding is adopted the method of wavelet transform coding. An M*N image is decomposed using wavelet transform. The image is decomposed into four sub-bands after passing a high- pass filter and low- pass filter. The four sub-bands are LL, HL, LH and HH respectively. The one obtained by low pass filtering rows and columns is referred as LL sub band contains horizontal details of the image. The one obtained by low pass filtering the rows and high pass filtering the columns is referred to as the LH sub band contains vertical details of the image and HH sub band contains the diagonal details of the image. The process is called the first level of wavelet decomposition. The low frequency sub-band can be continually decomposed into four sub-bands.

The image of low frequency sub-band contains major information. The values of high frequency sub-band approximate zero, the more high frequency the more obvious this situation. For image, the part of the low frequency is primary part which can represent the image information. So researchers take full advantage of the characteristic after wavelet transform and employ proper method to process the image coefficients for achieving effective compression.

5. SET PARTITIONING IN HIERARCHICAL TREE (SPIHT) ALGORITHM

The SPIHT algorithm was developed by Said and Pearlman in 1996. The SPIHT uses the fundamental idea of zero-tree coding from the EZW but is able to obtain a more efficient and better compression performance in most cases without having to use an arithmetic encoder. It uses wavelet sub band decomposition and imposes a quad tree structure across the sub bands in order to exploit the inter-band correlation.

SPIHT algorithm uses a special data structure – spatial orientation trees (SOT). This particular structure is not only made full use of different scales the correlation between the wavelet coefficients, but also give full consideration to the correlation of the same scale wavelet coefficients.

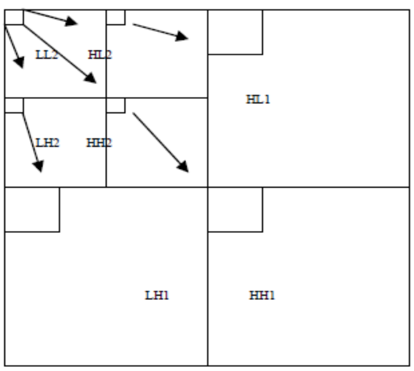

Fig2: Tree structure used in the SPIHT algorithm.

The algorithm searches each tree, and partitions the tree into one of three lists: 1) the list of significant pixels (LSP) containing the coordinates of pixels found to be significant at the current threshold; 2) the list of insignificant pixels (LIP), with pixels that are not significant at the current threshold; and 3) the list of insignificant sets (LIS), which contain information about trees that have all the constituent entries to be insignificant at the current threshold. [10-11]

The SPIHT algorithm consists of three stages: initialization, sorting pass and refinement pass. At the initialization stage the SPIHT first defines a start threshold according to the maximum value in the wavelet coefficients pyramid, then sets the LSP as an empty list and puts the coordinates of all coefficients in the coarsest level of the wavelet pyramid (*LL* band) in the LIP and those which have descendants to the LIS.

In the sorting pass, the elements in the LIP then in the LIS are sorted. For each pixel in the LIP it performs a significance test against the current threshold and outputs the test result (0 or 1) to the output bitstream. If a coefficient is significant, its sign is coded and then its coordinate is moved to the LSP. During the sorting pass of LIS, the SPIHT does the significance test for each set in the LIS and outputs the significance information (0 or 1). If a set is significant, it is partitioned into its offspring and leaves. The current threshold is divided by 2 and the sorting and refinement stages are continued until we achieve the target bit-rate. [12]

Steps of SPIHT Algorithm as follows:

1. Initialize the LIP, LIS, LSP table and determine the maximum threshold

   Threshold $T=2^n$,
   Where, n= [lg2max| (i, j) |]
2. For a given threshold, searches LIP, LIS table to determine the importance of each wavelet coefficient in LIP table.

3. If coefficient (i,j) is significant, then the output "1" and the sign bit is sent out , the node removed from the LIP form, added to the LSP end of the table.

4. If coefficient (i,j) is not significant, then the output should be "0", do not remove this node, the corresponding coordinates are moved to the LIP or LIS respectively, for subsequent testing at a lower bit level.

5. For the same threshold value, scan each node fine in turn in LSP table fine: output not newly added node LSP table wavelet coefficients corresponding to the first binary representation of n +1 bits, the scan end.

6. For the next scan: threshold $T \leftarrow T/2$, $n \leftarrow n-1$, repeat step (3) and step(4) step(5), until the threshold values or bit rate compliance encoder requirements.

6. EXPERIMENTAL RESULTS

The paper selects 8bit image of fishing boat and crowd image to stimulate for decomposition and reconstruction, and compare DCT and DWT-SPIHT algorithm. The simulation result showed in TABLE I, TABLE 2, Figure3 and Figure4. TABLE I show compressed size,

ISSN (Online): 2230-7893
www.IJCEM.org

Result on Images:

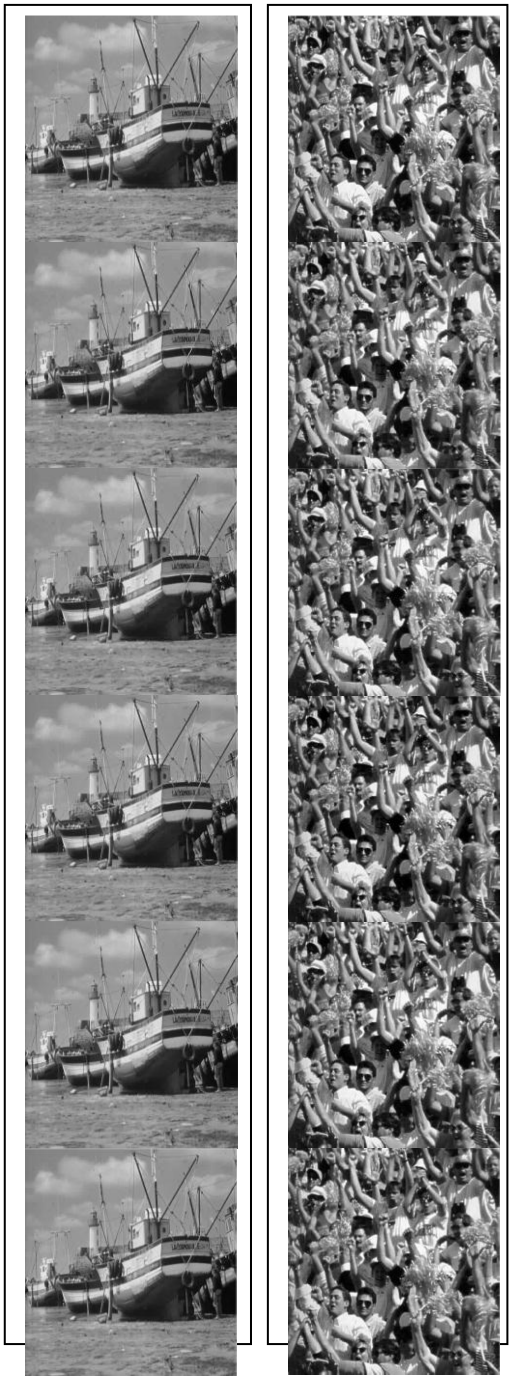

Figure 3&4: Original image b) level = 3 c) level =4 d) level =5 e) level =6 f) level =7

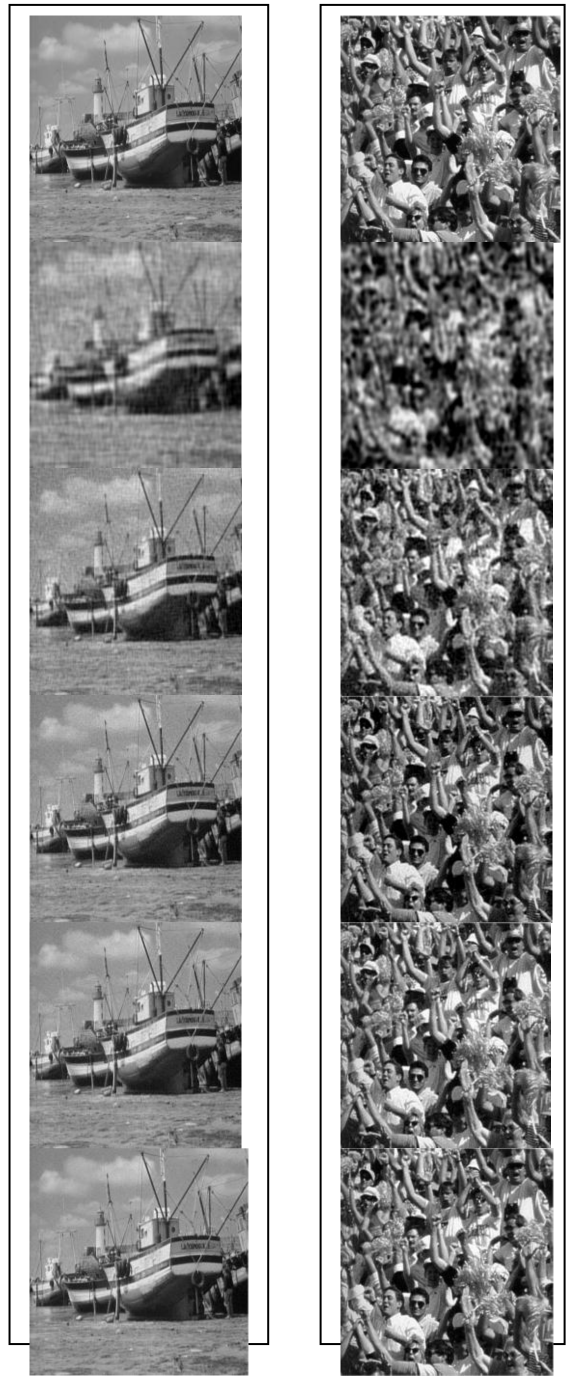

Figure 5 & 6: a) Original image b) c= 1000 c) c=5000 d) c=15000 e) c=20000 f) c=30000

TABLE I

PERFORMANCE EVALUATION OF DWT-SPIHT ALGORITHM

	Fishingboat.tif Original Size-257508 bytes				Crowd.tif Original Size-262571 bytes			
Decomposition level	Compressed size	Compression ratio	PSNR	MSE	Compressed Size	Compression Ratio	PSNR	MSE
3	48441	5.32	30.72	55.07	54056	4.86	26.78	136.34
4	75970	3.39	33.54	28.78	85463	3.07	29.07	80.48
5	83003	3.10	34.06	25.54	90668	2.90	29.44	74.00
6	84513	3.05	34.19	24.80	91464	2.87	29.50	72.92
7	84889	3.03	34.22	24.62	91619	2.87	29.52	72.65

TABLE II

PERFORMANCE EVALUATION OF DCT ALGORITHM

	Fishingboat.tif Original Size-257508 bytes				Crowd.tif Original Size-262571 bytes			
Coefficients	Compressed size	Compression ratio	PSNR	MSE	Compressed Size	Compression Ratio	PSNR	MSE
1000	151942	1.69478	22.3759	376.264	183709	1.42928	18.0576	1017.01
5000	195165	1.31944	25.5244	182.237	238443	1.10119	22.9739	327.861
15000	210591	1.22279	28.8355	85.0214	250040	1.05012	28.3232	95.6668
20000	213006	1.20892	29.9526	65.7386	250604	1.04775	30.1671	62.5977
30000	215930	1.19255	31.7795	43.1747	250815	1.04687	33.0722	32.0525

Compression Ratio, Peak to Noise Ratio (PSNR), Mean Square Error (MSE) for different level of decomposition. TABLE II show compressed size, Compression Ratio, Peak to Noise Ratio (PSNR), Mean Square Error (MSE) for different coefficients represents the coefficients with maximum energy. Figure 3 shows images of fishingboat.tiff at different level of decomposition. Figure 4 shows images of crowd.tif at same level of decomposition as that of fishing boat. It can be seen from Figure 3 and figure 4 that quality of images in figure 3 is much closer to the quality of the original image. Similarly Figure 5 and Figure 6 shows the images of fishingboat.tif and crowd.tif at different coefficients which contain most of energy. It has been analysed that result obtained from the set partition in hierarchical tree algorithm provide good visual quality and PSNR value and CR as compared to discrete cosine transform.

7. CONCLUSION

In this paper, the results of different transform coding techniques are compared i.e. Discrete Cosine Transform (DCT) and Wavelet based compression algorithm set partition in hierarchical tree (SPIHT). The effects of different number of decompositions, image contents and compression ratios are examined. The results of the above techniques DCT and DWT-SPIHT are compared by using two parameters such as Compressed Size, Compression Ratio, PSNR and MSE values from the reconstructed image. These compression algorithms provide a better performance in picture quality at higher compression ratio. These techniques are successfully tested on fishingboat.tif and crowd.tif images. It is observed that SPIHT provides a better result when compare to DCT. The SPIHT algorithm is coupled with the power of multiresolution analysis, yields significant compression with little quality loss. Because of the inherent multiresolution nature, wavelet-based coders facilitate progressive transmission of images. The above algorithms can be used to compress the image that is used in the web applications.

REFERENCES:

[1] Rafael C. Gonzalez, Richard Eugene; “Digital image processing”, Edition 3, 2008, page 466
[2] M. Ghanbari "Standard Codecs: Image Compression to Advanced Video Coding" Institution Electrical Engineers, ISBN: 0852967101, 2003
[3] Chuanwei Sun 1Quanbin Li,Jingao Liu “The Study of Digital Image Compression Based on Wavelets”, ICALIP2010
[4] Kamrul Hasan Talukder and Koichi Harada, Enhancement of Discrete “Wavelet Transform (DWT) for Image Transmission over Internet”, 2011 Eighth International Conference on Information Technology: New Generations
[5] Still Image and video compression with MATLAB, K. S. Thyagarajan, A JOHN WILEY & SONS, INC., PUBLICATION
[6] Ken Cabeen and Peter Gant“Image Compression and Discrete Cosine Transform “
[7] “DCT-BASED IMAGE COMPRESSION”, Vision Research and Image Sciences Laboratory http://visl.technion.ac.il
[8] Mr. Mahendra M. Dixit1, Prof. Priyatamkumar, Comparative Analysis of Variable Quantization DCT and Variable Rank Matrix SVD Algorithms for Image Compression Applications
[9] http://www.scribd.com/doc/26156035/Image-Compression-Using-DCT-Implementing-Matlab
[10] ANALYSIS OF SPIHT ALGORITHM USING TILING OPERATIONSG.Sadashivappa1, Mahesh Jayakar1.a, K.V.S Ananda Babu2, 2010 International Conference on Signal Acquisition and Processing
[11] , Chunlei Jiang Shuxin Yin “ A Hybrid Image Compression Algorithm Based on Human Visual System”, 2010 International Conference on Computer Application and System Modeling (ICCASM 2010)
[12] Li Zhu, Yi_min Yang “ Embeded Image Compression Using Differential Coding and Optimization Method”