

# Image Compression : Review and Comparative Analysis

Jaya S. Kulchandani<sup>#1</sup>, Suman H. Pal<sup>#2</sup>, Kruti J. Dangarwala<sup>#3</sup>

<sup>#</sup>Department of Information Technology  
Shri S'ad Vidya Mandal Institute of Technology  
Bharuch 392-001, Gujarat, India

**Abstract**— Image compression is the art of reducing the size of data that represents a digital image; without degrading the quality of the image. Large quantity of digital data is required to store an image. Though there is technological advancement in storage space required to store data but the demand for storage capacities is not satisfied by the availability. In such scenario image compression can be considered as the only solution. This paper addresses different image compression techniques and comparative analysis of Huffman encoding, Arithmetic coding, Run length coding, Transform coding and Wavelet coding is provided. Review will allow new researchers to study different techniques of image compression and provide them direction to choose best technique for their application.

**Keywords**— Image Compression ; Image ; Huffman encoding; Arithmetic coding; Run length coding; Transform coding; Wavelet coding

## I. INTRODUCTION

The main aim of Image compression is to reduce the storage space required to store an image. Reduction in the size of image allows to efficiently utilize the memory or disk space [1]. Over last few years analog television has been replaced by digital television which requires high transmission of digital data. Moreover, many images and videos are moved all over the world via internet. Demand of bandwidth for transmission of these digital data exceeds the availability. Hence image compression provides a way to satisfy the demand of transmission of data with available limited bandwidth. Image compression is widely used in field of image processing due to its wide range of applications like teleconferencing [2], multimedia, medical imaging [2,3], remote sensing [10], control of automatic system etc. Some recent research trends in image compression are:

- Lossy Compression and Iterative Reconstruction for Encrypted Image [4]
- Image Compression using K-Means Clustering And Nuclear Medicine Image Processing [5]
- Three-Dimensional Video Compression using Motion Estimation-Compensation and Discrete Wavelet Transform [6]
- New Network Bandwidth-limited Multi-view Video plus Depth Coding Method for 3D Video [7]

- Image Compression of Neural Network Based on Corner Block [8]
- Improved wavelet based compression with adaptive lifting scheme using Artificial Bee Colony algorithm [9]

General scenario of image compression is shown in Figure 1.

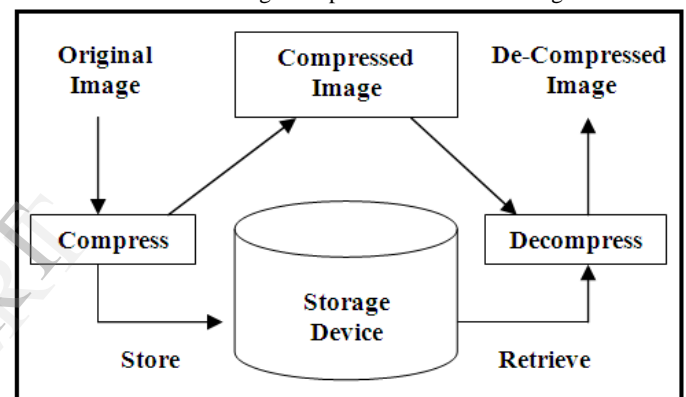


Fig .1 General scenario of image compression

Original Image represents the raw image or image obtained from sequence of video frames. Original image is compressed to required amount by image encoder using different image compression techniques. Compressed image then either stored on storage device or transmitted. If the compressed image is to be transmitted, extra bits are added to it which is then converted to radio frequency signals via digital modulator that is suitable for transmission. At receiver's end, image is decompressed by performing demodulation and decoding operations followed by entropy decoder and source decoder operations. In case, data is not transmitted then the compressed data stored in storage device is entropy decoded followed by source decoding [10].

Remainder of the paper is arranged as follows: Section II provides brief idea about fundamental concept of image compression. Section III presents classification of image compression techniques and focuses on different lossless and lossy image compression techniques. Section IV provides comparative analysis of Huffman encoding, Arithmetic coding, Run length coding, Transform coding and Wavelet coding. Section V concludes the paper.

## II. FUNDAMENTAL CONCEPTS OF IMAGE COMPRESSION

Image compression i.e., form of data compression is the aspect of representing qualitative information in reduced form. Central issue of image (data) compression is data redundancy. Data redundancy is the state where same piece of data is represented multiple times for given information  $n_1, n_2$  relative data redundancy  $R_D$  is represented as:

$$R_D = 1 - (1/C_R) \quad (1)$$

Where  $C_R$  is compression ratio i.e.  $C_R = n_1/n_2$   
In image compression redundancy can be mainly classified into three forms: a) Coding redundancy b) Interpixel redundancy and c) Psycho-visual redundancy [10]. Coding redundancy is present when less than optimal code words are used. Interpixel redundancy results from correlations between the pixels of an image. Psychovisual redundancy is due to data that is ignored by the human visual system (i.e. visually non essential information). The objective of image compression is to reduce the number of bits. It needs to represent an image by removing the Spatial and Spectral redundancies as much as possible, while keeping the resolution and visual quality of the reconstructed image as close to the original image. Efficient image compression can be obtained by reduction of one of the above redundancy.

## III. CLASSIFICATION OF COMPRESSION ALGORITHM

Compression algorithms come in two general flavors: Lossless and Lossy. As the name states, when Lossless data is decompressed, the resulting image is identical to the original where as Lossy compression algorithms result in loss of data and the decompressed image is not exactly the same as the original image.

### A. Lossless Compression

Lossless Compression is a form of compression that is used when it is necessary to preserve the original format of the image, since small loss of information in many applications can prove to be deleterious. In lossless image compression, image is considered as sequence of pixel in row major order and each pixel is processed in two steps: Firstly numerical value of the pixel is predicted. Secondly using an entropy coder and chosen probability distribution, difference between the predicted pixel value and actual intensity of next pixel is coded. Different lossless compression technique can be explained as:

#### 1) Huffman Coding

Huffman Coding is based on the principle of probability distribution of the pixel value where in probability distribution of the pixel value is used to develop code words. Firstly, frequency distribution is calculated for the pixel value in order to calculate probability distribution. Once the probabilities are known, code words are assigned on the basis of probabilities value i.e. pixel value with higher probability are assigned shorter codes, where as longer codes are assigned to pixels that have lower probability. Lastly binary tree is generated. Say suppose an image sequence is represented by four different symbols

{a, b, c, d} and the probability of occurrence of each symbol is {0.125, 0.25, 0.375, 0.25} respectively. Binary Tree is generated by considering two symbols with the minimum probability from left to right, adding them together in order to obtain equivalent symbol whose probability is equal to sum obtained by adding probability of corresponding symbols. The process is continued until there is just one symbol. At last bits are assigned to different branches from right to left, starting from root [13].

#### 2) Runlength Encoding

Run length encoding is considered to be simplest form of lossless compression technique. It allows compressing data that is made of any combination of symbols. The general idea behind this method is that it replaces identical symbols known as Runs. Observed identical symbols are replaced by single identical value along with its number of count. Run length encoding can be used with the file or data that contains many such identical values. [12] Run length encoding method is implemented by most bitmap files like TIFF, BMP, etc...

#### 3) Arithmetic Coding

In arithmetic coding instead of coding each image pixel (symbol) individually, entire image sequence is assigned single arithmetic code word [10]. Unlike Huffman coding, no one-to-one correspondence exist between code word and image sequence. An interval from 0 to 1 is defined by code word. With the increase in number of symbols, interval decreases. Say suppose image sequence is represented by symbols {a, b, c, d}. Initially interval [0, 1) is assumed for representation of image sequence. Based on the probability value for each symbol, interval is narrowed until last interval representing the entire image sequence is obtained.

#### 4) LZW Coding

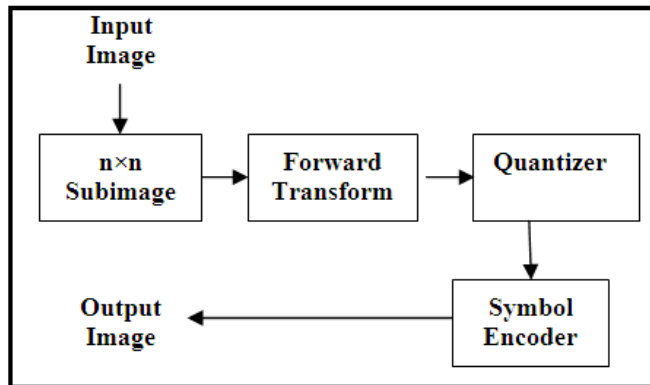
Lempel-Ziv-Welch is a dictionary-based compression technique. It allows mapping of a variable length of image sequence to fixed length of code [14]. The main working principle of LZW algorithm is the multiple occurrence of image sequence for a given image that needs to be encoded. LZW algorithm builds a dictionary by replacing the multiple occurrences of pattern by an index code. The dictionary is generally initialized with 256 values of the ASCII table. Different occurrence of pattern of image sequence is placed in the dictionary. In case pattern of image sequence already exist in the dictionary, pattern is replaced by the index value of the image sequence that already exist in the dictionary. In encoding process the image sequence is passed only if it has length smaller than the longest image sequence pattern that already exist in the dictionary. In decoding process, dictionary is rebuilt in reverse direction.

### B. Lossy Compression

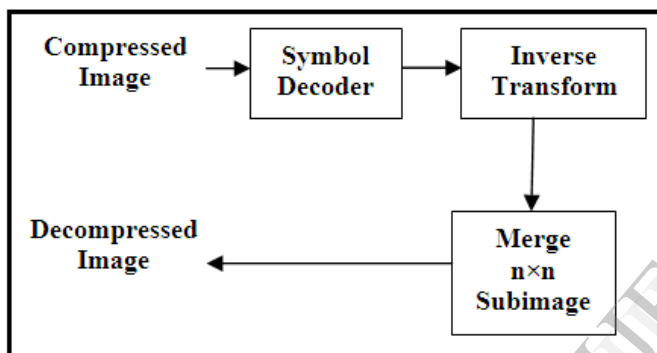
Lossy Compression, for a given image sequence tests the color values in order to determine the variations in pixel color value that are so minute that human eye cannot distinguish the difference. The color value that falls in the boundary range of human perception is replaced by lossy

compression algorithms. The finely ranked pixels are then discarded. Higher compression rate can be obtained by lossy compression but final quality of the product may degrade.

### 1) Transform Coding



(a)



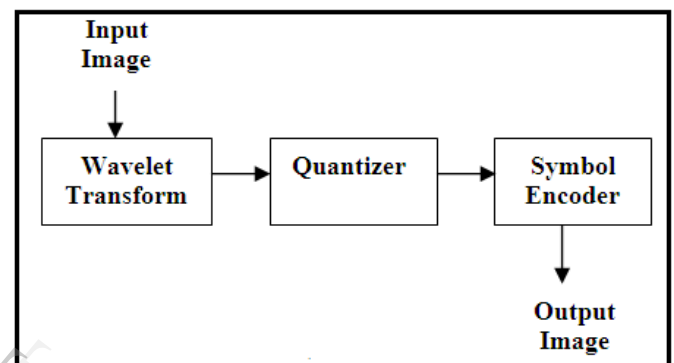
(b)

Fig. 2 (a) Encoding Process (b) Decoding process [10]

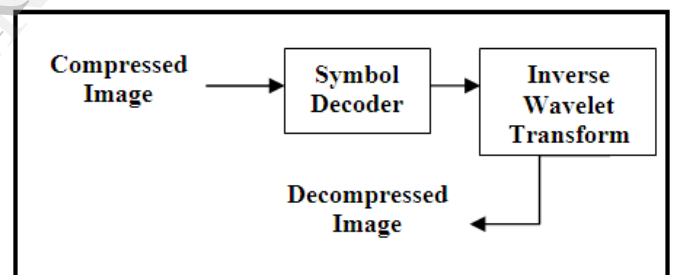
Transform coding is most well known form of lossy compression where in original form of the image is subdivided into smaller sub images. For each sub image obtained transform coefficients is calculated which converts the original array of sub image in array of coefficients. The arrays of coefficients are then quantized and obtained output is further treated by image encoding technique which provides the final output representing the encoded image. In decoding process reverse process is applied i.e. the decoded output obtained is dequantized. Decoding process does not provide exactly the same image as original; in other words, information lost during encoding process cannot be recovered. DCT, DFT, WHT etc... are frequently used transforms.

### 2) Wavelet Coding

Wavelet coding is general used to represent high frequency components in two-dimensional images, like images of star or satellite images by means of smaller amount of information[11]. Wavelet Coding is implemented in JPEG-2000. The main purpose of wavelet coding is to store an image of higher dimension in space as small as possible. In encoding process, initially wavelet transform is applied as a result of which, coefficients are obtained where in number of coefficients obtained are equal to number of pixels in the image. After that coefficients are quantized and quantized output pass through encoder that provides final output. In decoding process, Compressed is decoded by decoder followed by inverse wavelet transfer.



(a)



(b)

Fig. 3 (a) Encoding Process (b) Decoding process [10]

## IV. COMPARATIVE ANALYSIS

Comparative analysis for image compression techniques Huffman encoding, Arithmetic coding, Run length coding, LZW, Transform coding and Wavelet coding is done on the basis of the compression ratio. Compression ratio in image compression is defined as the ratio between uncompressed image and compressed image.

$$\text{Compression Ratio (CR)} = \frac{\text{Uncompressed Image}}{\text{Compressed Image}} \quad (2)$$

Table I shows comparative analysis for the methods discussed in the paper for the images in fig. 4 along with the application area.

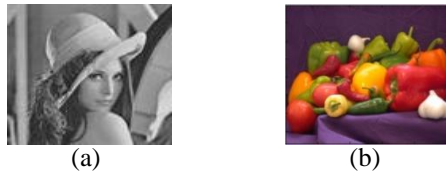


Fig.4 (a) Lena Image with original file size = 3051.52 Bytes (b) Peppers Image with original file size = 3747.84 Bytes

TABLE I

Algorithm	Lena Image	Peppers Image	Application Area
	CR	CR	
Huffman Coding	1.16	1.18	Used in JPEG.
Run Length Coding	1.21	1.60	Used in TIFF and GIF files.
Arithmetic Coding	1.76	1.52	Used mostly for frequently occurring sequences of pixels.
LZW	1.28	1.42	Used mostly for TIFF, BMP and PCX files.
Transform Coding	3.88	4.22	Used in JPEG and M-JPEG
Wavelet Coding	3.75	2.49	Used in JPEG-2000

## V. CONCLUSION

Image Compression is an important field of research due to its wide range of application in image processing area. This paper provides basic idea about Image Compression. Fundamental concepts of Image Compression are discussed briefly which includes redundancy concepts, general scenario of image compression process. Paper focuses mainly on review of different image compression techniques like Huffman encoding, Arithmetic coding, Run length coding, LZW, Transform coding and Wavelet coding. Comparative analysis is provided for the discussed techniques based on the compression ratio achieved by each technique.

## REFERENCES

- [1] Uli Grasemann and Risto Miikkulainen, "Effective Image Compression using Evolved Wavelets," ACM, pp. 1961-1968, 2005.
- [2] Gopinath R. Kuduvali and Rangaraj M. Rangayyan, "Performance Analysis of Reversible Image Compression Techniques for High-Resolution Digital Teleradiology," IEEE Transactions On Medical Imaging, Vol. 11, No. 3, pp. 430-445, September 1992.
- [3] H. Narfi Stefansson, Kevin W. Eliceiri, Charles F. Thomas, Amos Ron, Ron DeVore, Robert Sharpley and John G. White, "Wavelet Compression of Three-Dimensional Time-Lapse Biological Image Data," Microscopy and Microanalysis, pp. 9-17, 2005.
- [4] Xinpeng Zhang, "Lossy Compression and Iterative Reconstruction for Encrypted Image," IEEE Transactions On Information Forensics And Security, Vol. 6, No. 1, pp. 53-58, March 2011.
- [5] Himadri Nath Moulick and Moumita Ghosh, "Image Compression Using K-Means Clustering And Nuclear Medicine Image Processing," International Journal of Innovative Research in Computer and Communication Engineering Vol. 1, Issue 4, pp. 869-877 June 2013.
- [6] Harneet Kaur and Neeru Jindal, "Three-Dimensional Video Compression using Motion Estimation-Compensation and Discrete Wavelet Transform," ACEEE, Proc. of Int. Conf. on Recent Trends in Information, Telecommunication and Computing, ITC, pp. 208-219, 2014.
- [7] Yapei Zhu, Mei Yu, Gunjun Zhang, Zongju Peng, Feng Shao, Gangyi Jiang, "New Network Bandwidth-Limited Multi-View Video Plus Depth Coding Method For 3d Video," Journal Of Multimedia, Vol. 8, No. 1, pp. 8-15, February 2013.
- [8] Wenjing Zhang and Wei Yao Donglai Ma, "Image Compression of Neural Network Based on Corner Block," Journal of Multimedia, Vol. 9, No. 1, pp. 166-172, January 2014.
- [9] R.Ramanathan, K.Kalaiarasi and D.Prabha, "Improved wavelet based compression with adaptive lifting scheme using Artificial Bee Colony algorithm," International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 2, Issue 4, April 2013.
- [10] R. Gonzalez and R. Woods, Digital Image Processing, 2<sup>nd</sup> ed. Upper Saddle River, NJ, USA: Prentice Hall, Inc, 2002.
- [11] Wavelet Transform. [Online]. Available: [http://en.m.wikipedia.org/wiki/Wavelet\\_transform](http://en.m.wikipedia.org/wiki/Wavelet_transform).
- [12] Uchale Bhagwat Shankar, "Image Compression Technique," International Journal of Information Technology and Knowledge Management, Volume 2, No. 2, pp. 265-269, December 2010.
- [13] Mamta Sharma, "Compression Using Huffman Coding" International Journal of Computer Science and Network Security, Vol.10 No.5, pp. 133-141, May 2010.
- [14] Dalvir Kaur and Kamaljeet Kaur, "Data Compression on Columnar-Database using Hybrid Approach (Huffman And Lempel-Ziv Welch Algorithm)," International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 5, pp. 889-894, May 2013.