Lossless Image Compression through Huffman Coding Technique and Its Application in Image Processing using MATLAB

Vikash Kumar, Sanjay Sharma

Abstract: ***Images include information about human body which is used for different purpose such as medical examination security and other plans Compression of images is used in some applications such as profiling information and transmission systems. Regard to importance of images information, lossless or loss compression is preferred. Lossless compressions are JPEG, JPEG-LS and JPEG2000 are few well-known methods for lossless compression. We will use differential pulse code modulation for image compression with Huffman encoder, which is one of the latest and provides good compression ratio, peak signal to noise ratio and minimum mean square error. . In this paper we try to answer the following question. Which entropy coding, Huffman, is more suitable compared to other from the compression ratio, performance, and implementation points of view? We have implemented and tested Huffman algorithms. Also we compare it with other existing methods with respect to parameter compression ratio, peak signal noise ratio.***

Keywords: Lossless Compression, PSNR, Compression-Ratio, Encoding Technique, Huffman Coding, JPEG2000, JPEG-LS, JPEG

I. INTRODUCTION

Images are important documents nowadays; images include various information e.g., human bodies in medical images which are used for different purpose such as medical security and other plans. Compression of images is used in some applications such as profiling data and transmission system. To work with them in a few applications they need to be compressed more or less depending on the purpose of the application. Regard to importance of images information lossless or loss compression is preferred. Lossless compressions are JPEG, JPEG-LS and JPEG2000 are some well-known methods for lossless compression. In this paper we proposed the Lossless method of image compression and decompression using a simple coding technique called Huffman coding. This technique is simple in implementation and utilizes less memory. A software algorithm has been developed and implemented to compress and decompress the given image using Huffman coding techniques in a MATLAB platform. Image compression plays a key role in many important applications, including image database, image communications, remote sensing The image to be compressed are gray scale with pixel values between 0 to 255. Multimedia data, especially images have been increasing every day.

Revised Version Manuscript Received on January 27, 2017.

Vikash Kumar, M.Tech Scholar, Department of Economic and Community, Millennium Institute of Technology & Science, Bhopal (M.P) India.

Dr. Sanjay Sharma, Associate Professor, Department of Economic and Community, Millennium Institute of Technology & Science, Bhopal (M.P) India.

Because of their large capacity, storing and transmitting are not easy and they need large storage devices and high bandwidth network systems. In order to alleviate these requirements, compression techniques and standards such as JPEG, JPEG2000, MPEG-2 and MPEG-4 have been used and proposed. The JPEG standard constructed from several functions such as DCT, quantization, and entropy coding. Huffman and arithmetic coding are the two most important entropy coding in image compression standards.

Image compression, more or less depends on the purpose of the user interests. There are some algorithms that perform this compression in different ways; some are lossless and keep the same information as the original image, some others loss information when compressing the image. Some of these compression methods are designed for specific kinds of images, so they will not be so good for other kinds of images. Some algorithms even let you change parameters they use to adjust the compression better to the image. There are different formats works for each of the images. There are some formats that match some images better than others depending in what you are looking for to obtain, and the type of image you are working with. The image compression technique is mostly classified into two techniques depending whether or not an exact copy of the original image could be reconstructed using the compressed image. Multimedia data, especially images have been increasing every day. Because of their large capacity, storing and transmitting are not easy and they need large storage devices and high bandwidth network systems.

A. Need for Compression

Compression is a technique to reduce the quantity of data without excessively reducing the quality of the multimedia data. The transition and storing of compressed multimedia data is much faster and more efficient than original uncompressed multimedia data.

B. Principle behind Compression

a. Redundancies reduction aims at removing duplication from the signal source (image/video).

b. Irrelevancy reduction omits parts of the signal that will not be noticed by the signal receiver, namely the Human Visual System.

C. Types of Compression

(i) Lossless Compression
(ii) Lossy Compression

Lossless Compression: Data is compressed and can be reconstituted (uncompressed) without loss of detail or information.

These are referred to as bit-preserving or reversible compression systems also Lossless compression frequently involves some form of entropy encoding and are based in information theoretic techniques.

Lossy Compression: The aim is to obtain the best possible fidelity for a given bit-rate or minimizing the bit-rate to achieve a given fidelity measure. Video and audio compression techniques are most suited to this form of compression. If an image is compressed it clearly needs to uncompressed (decoded) before it can viewed/listened to. Some processing of data may be possible in encoded form however. Lossy compression use source encoding techniques that may involve transform encoding, differential encoding or vector quantization.

II. METHODOLOGY

There are following methods for lossless and lossy data compression techniques:

A. Lossless Coding Techniques

a. Run length encoding
b. Huffman encoding
c. Arithmetic encoding
d. Entropy coding
e. Area coding

B. Huffman Encoding

Huffman Encoding Algorithms use the probability distribution of the alphabet of the source to develop the code words for symbols. The frequency distribution of all the characters of the source is calculated in order to calculate the probability distribution. According to the probabilities, the code words are assigned. Shorter code words for higher probabilities and longer code words for smaller probabilities are assigned. For this task a binary tree is created using the symbols as leaves according to their probabilities and paths of those are taken as the code words.

C. Development of Huffman Coding and Decoding Algorithm

Step1- Read the image on to the workspace of the MATAB.
Step2- Convert the given colour image into grey level image.
Step3- Call a function which will find the symbols (i.e.pixel value which is non-repeated).
Step4- Call a function which will calculate the probability of each symbol.
Step5- Probability of symbols are arranged in decreasing order and lower probabilities are merged and this step is continued until only two probabilities are left and codes are assigned according to rule that: the highest probable symbol will have a shorter length code.
Step6- Further Huffman encoding is performed i.e. mapping of the code words to the corresponding symbols will result in a compressed data.
Step7- The original image is reconstructed i.e. decompression is done by using Huffman decoding.
Step8- Generate a tree equivalent to the encoding tree.
Step9- Read input character wise and left to the table II until last element is reached in the table II.
Step10-Output the character encodes in the leaf and return to the root, and continue the step9 until all the codes of corresponding symbols are known.

D. Lossy coding techniques

a. Predictive coding
b. Transform coding (FT/DCT/Wavelets)

Proposed System Diagram

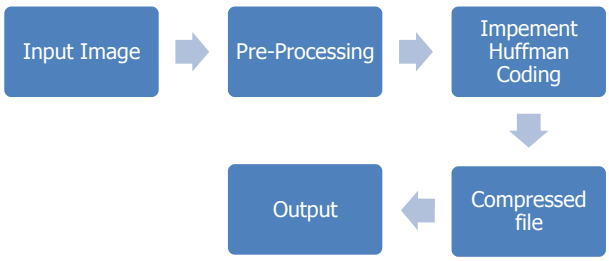

Figure 1: System Diagram for Compression Technique

III. RESULT AND DISCUSSIONS

In this paper MATLAB Simulator has been used to evaluate the performance of Huffman coding technique in the field of image compression behalf of compression parameters. The proposed algorithm has been applied on the different images having TIFF and JPEG format.

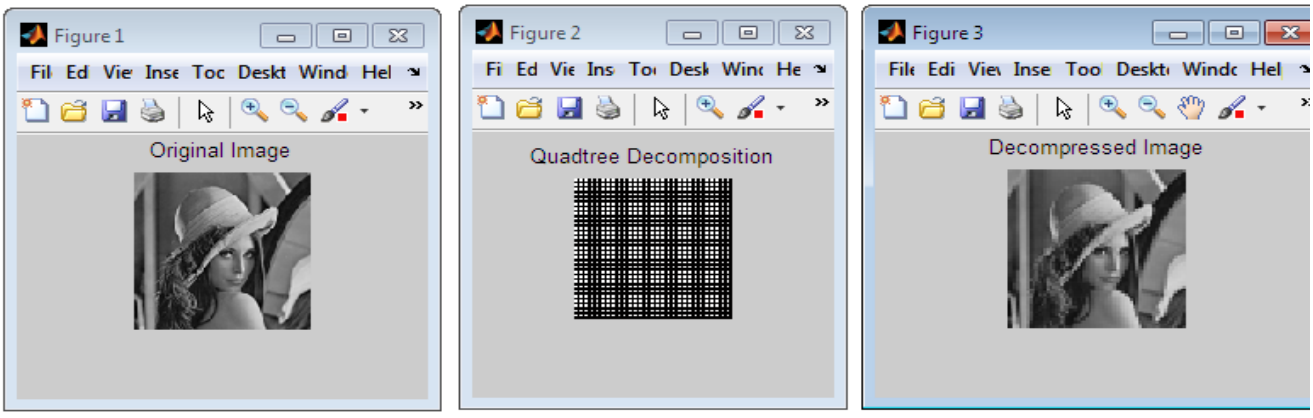

Figure 2: Lenna Original Image, Quadtree Decomposition Image & Decompressed Image

Time taken for compression = 7.554787 seconds
Compression ratio= 0.637963
Time taken for Decompression = 92.334621 seconds
PSNR= 25.506879

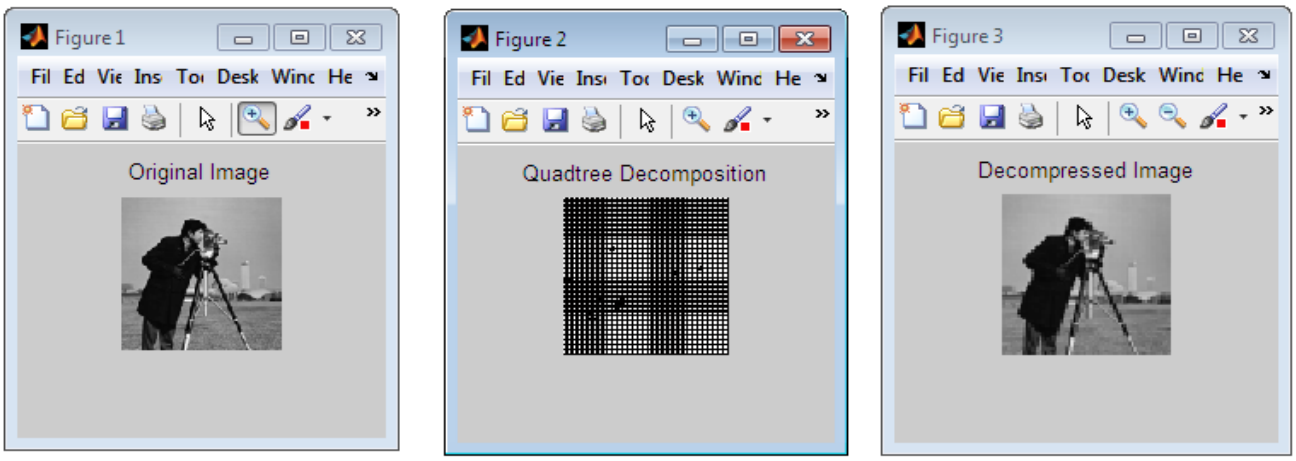

Figure 3: Cameraman Original Image, Quadtree Decomposition Image & Decompressed Image

Time taken for compression = 9.076491 seconds
Compression ratio= 0.645484
Time taken for Decompression = 91.595833 seconds
PSNR= 24.763435

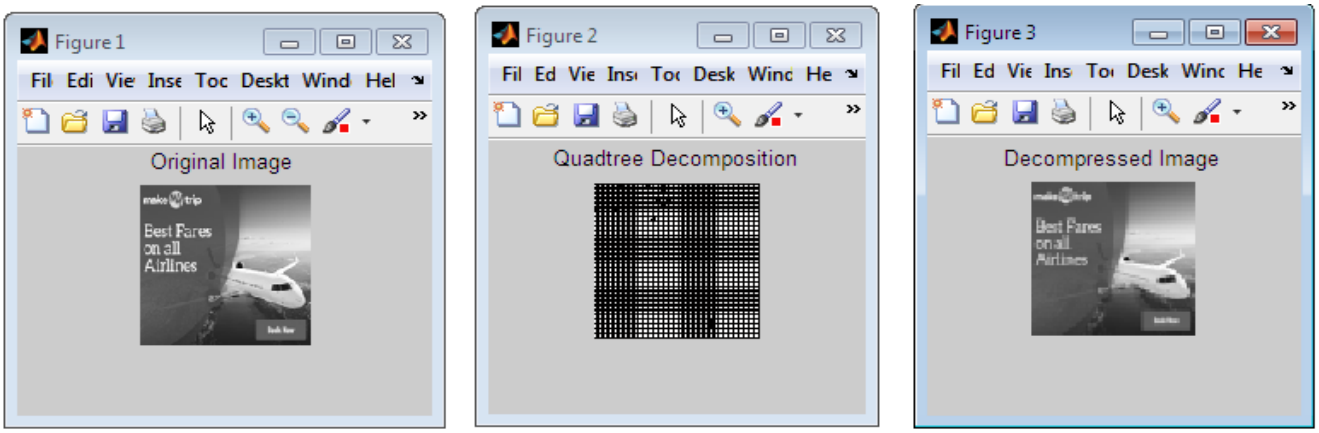

Figure 4: Aero plane Original Image, Quad Tree Decomposition Image & Decompressed Image

Size= 16 kb, 4kb, 10 kb
Time taken for compression = 6.152462 seconds
Compression ratio (CR) = 0.642711
Time taken for Decompression = 80.964061 seconds
PSNR= 24.986462

Result Comparison

	Images	Time Taken for Compression (Sec.)	Compression Ratio (CR)	Time Taken for Decompression (Sec.)	PSNR
Previous Result	**Lena**	156.1369	3.6893	23.1200	36.5806
Present Results	**Lena**	7.554787	0.637963	92.334621	25.506879
	Cameraman	9.076491	0.645484	91.595833	24.763435
	Aero-plane	6.152462	0.642711	80.964061	24.986462

IV. CONCLUSION & FUTURE SCOPE

As a future work more focus can be on improvement of compression ratio using the new techniques. The proposed technique can be experimented on different kinds of data sets like audio, video, text as till now it is restricted to images. New methods can be combined and proposed that decreases the time complexity incurred in creating dictionary. The experimental dataset in this research is somehow limited; so applying the developed methods on a larger dataset could be a subject for future research which may lead to new observations and conclusions. Further the work can be extended to video compression. Video data is basically a three dimensional array of color pixels, that contains spatial and temporal redundancy.

Similarities can thus be encoded by registering differences within a frame where data frame is a set of all pixels that correspond to a single time moment. With the advancements in compression technology, it is now very easy and efficient to compress video files.

REFERENCES

1. N. Parvatham and Seetharaman Gopalakrishnan, 2012 Third International Conference on Intelligent Systems Modelling and Simulation "A Novel Architecture for an Efficient Implementation of Image compression using 2D-DWT"
2. Giridhar Mandyam, Nasir Ahmed, Neeraj Magotra, "Lossless Image compression using Discrete Cosine Transform", Journal of Visual Communication and Image Representation ,Vol.8, No.1, March, pp.21-26, 1997, Article no. VC970323.
3. Donapati, S. Yagain "A Comparative Study of Effects of CSC on Image Compression Ratios While Using JPEG-XR", Year of Publication **(2013)**, pp. 158-161.
4. J. Wang, "Shot Cut Detection Based On The Statistical Parameter Histogram With The Discrete Walsh Transform", Second International Conference on MultiMedia and Information Technology, **(2010)**.
5. J. Ziv and A. Lempel, "A Universal Algorithm for Sequential Data Compression", IEEE Transactions on Information Theory, May 1977
6. Dr. T. Bhaskara Reddy, Miss.Hema suresh yaragunti, Dr.S.kiran, Mrs.T.Anuradha " A novel approach of lossless image compression using hashing and Huffman coding ",International Journal of Engineering research and technology ,vol.2 issue 3,march-2013.
7. G.C Chang Y.D Lin (2010) "An Efficient Lossless ECG Compression Method Using Delta Coding and Optimal Selective Huffman Coding" IFMBE proceedings 2010, Volume 31, Part 6, 1327-1330, DOI: 10.1007/978-3-642-14515-5_338.