



## IMAGE COMPRESSION USING DECISION TREE TECHNIQUE

Amanpreet Kaur

Student Yadavindra College of Engineering,  
Talwandi Sabo Bathinda,  
Punjab (India)

Balkrishan Jindal

A. P., CE Section, Yadavindra College of Engineering,  
GKC, Punjabi University, Talwandi Sabo,  
Punjab (India)

**Abstract**— Image compression is the process of minimizing the size of file without degrading the file quality. The reduction in file size permits number of files to be stored in a memory space. The proposed compression method used the decision tree algorithm for compression of a grey scale images. The Decision tree reads an image which needs to compress. Patches are created and assign weight to each patch on the basis of similarity between patches. Decision tree is formed based on the similarity between the pixels. Then data is merged to form compressed image otherwise define root node which has maximum weight, left node which has less weight than root node and right node which has minimum weight and decision tree is created. The proposed method is evaluated using the Image Quality Measures (IQM) like Peak-Signal to Noise Ratio (PSNR), Compression Ratio (CR), Mean Square Error (MSE) and Elapsed Time (ET). The experimental results of the proposed method are compared with Huffman method and Wavelet Difference Reduction method. From the experimental results it has been concluded that the proposed method is better than Huffman method and Wavelet Difference Reduction.

**Keywords**— Lossless compression, Lossy compression, Peak Signal to Noise Ratio, Compression Ratio, Mean Square Error.

### 1. INTRODUCTION

Grey scale digital image is combination of black and white colors. Black shade represents weakest intensity and White represents strongest intensity value. Grey scale pictures have several shades of gray between weakest and strongest intensity values.

Image compression is a technique to remove the irrelevant and redundant data from image without lowering the quality of the picture to an undesirable form. Image compression is used for storing and transmitting the data in more efficient form. When the size of file is reduce, the number of images to be stored in a specific amount of memory. It also decrease the time needed for pictures to be sent over the web or downloaded from sites [1]. For compression, one shortcut trick to decrease redundant data, which means storing data one time rather than six times. Another one is to seek out those parts of the data that is not required and simply remove those away.

#### Image Compression Techniques

Image compression methods are divided into two strategies one is lossless compression and second one is lossy compression. Lossless, as the name suggests "no loss of data" after decompression. Error free compression is most popular for ZIP file format and image file format like PNG or GIP [2]. In Lossy compression, some information from the image removed at the time of compression. The various compression techniques are show in Figure 1.

The following are lossless and lossy compression techniques that are explained below:

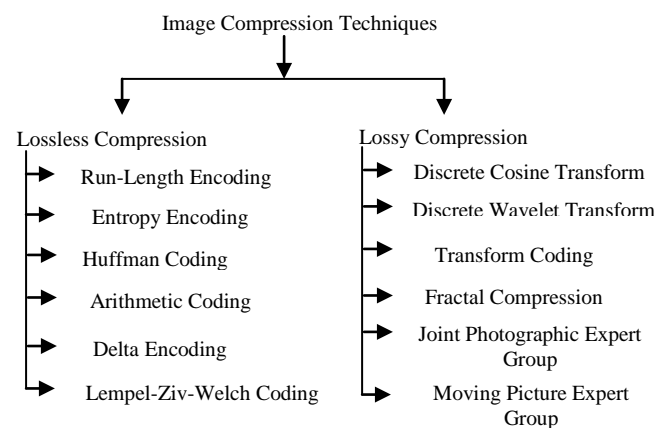


Figure 1. Various Compression Techniques

#### RLE

Run Length encoding is an error free compression technique. It is an easy method of compression [3]. In this method information expressed within the kind of value and run length. Run length is defined as number of successive equal values. Content Records are used numerous spaces to isolate sentence, arrange tables and graphs, indent passages and so on.

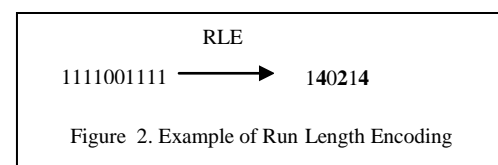


Figure 2. Example of Run Length Encoding

1 0 1 are called value and 4 2 4 are called Run Length.

Run length encoding compression is only useful for files that contain lot of repetitive data. Computer generated color images can also give fair compression ratios.

#### Huffman Coding

David A. introduce the Huffman encoding in 1952. Minimum redundancy codes are developed using this method. This coding technique can reduce the size 10% to 50% by removing the unnecessary information. It is used in JPEG compression. The Compression and decompression methods represent every character which follows the binary codes. The various steps are performed in [4]

- Step 1:** Arrange the pixel values in descending order of their probability.
- Step 2:** Two values N1 and N2 are join which has least probabilities, mark one with zero and another with one.
- Step 3:** Two least probabilities symbols are connect with each other.
- Step 4:** Repeat step 2 until you got a single symbol with probability 1.
- Step 5:** The coding tree trace from root.

### Entropy Encoding

It used to check the resemblance in information streams. Encoding strategy works as follows. It generates a distinct prefix code and allots one symbol to this code within the range of source data. This can be of contrasting size ensuing to influencing the prefix code. After quantization step, it is performed on an image. The main feature of entropy encoding is that it stores any picture or image in small memory space. The most widely recognized techniques are AC (Arithmetic Coding) and Huffman coding [5].

### Arithmetic Coding [6]

In AC, any message is assigned a arithmetic code word. The range of code word defines between real numbers zero and one. Zero represents low value and one represents high value. AC does not give coordinated correspondence between code words and input image. The point of AC is to characterize a strategy that gives symbols a optimal length. The normal code length is near the conceivable least given by data hypothesis. Output of this method is string of bits. The best method provided by the arithmetic coding is to code symbols according to the probability of their occurrence. It is a paid algorithm. Arithmetic coding is superior for black and white images and provides higher efficiency and lot of flexibility when compared with Huffman coding. It is used for compression of any video and pictures. This technique provide a particular range to each symbol that range show the probability in which that symbol presence neat.

### Delta Encoding

Delta encoding is a lossless compression technique. Any modification in a variable including letters, characters and digits etc. are expressed using the Greek letter delta ( $\Delta$ ). (In output, just the difference between back to back digits is store in memory instead of the considerable number of digits are put away themselves. In delta encoding strategy first value is exactly same that are used in input data streams. All the remaining values in the output equal to the difference between the next values in the input file and antecedent values. It is used for solving differential equations in mathematics.

### LZW Coding

It is developed by A. Lempel and J. Ziv. Some changes are also done by the Terry A. Welch. LZW coding replaces string of characters with single codes. Lempel Ziv Welch

technique is working supported the prevalence repetitive of character sequences within the string to be encoded. Its working comprises in substituting designs with a record code, by continuously assembling a dictionary. The word reference is computerize with the 256 estimations of the ASCII table. It is used in TIFF, Postscript and in GIF file formats [6, 7].

### Discrete Cosine Transform (DCT)

DCT is a lossy compression strategy. It is applied for analyzing the texture features. In Discrete Cosine Transform first off all breaks an image into dimension of  $8 \times 8$  or  $16 \times 16$ . DCT change for each sub square, last encodes, transmits and quantized Discrete Cosine Transform coefficients after change for shaping compacted picture organize. Discrete Cosine Transform is utilized to change over information into the summation of grouping of cosine waves moving at various frequencies. Image is divided into parts of different frequencies [12, 15].

In quantization step those subparts that are not required are discarded. So it is known as lossy compression. It is used in mathematics for solving fractional differential equations, JPEG standard [8].

### Discrete Wavelet Transform (DWT)

It is a Fourier based transform technique. The DWT expressed an image as a summation of wavelet coefficients, called wavelets with totally extraordinary scale and area. The Discrete Wavelet Transform speaks to the image information into a gathering of explained way (detailed) and shorter way passes. JPEG2000 is based on DWT (Discrete Wavelet Transform). The Joint Photographic Experts Group2000 compression strategy was outlined by an ISO institutionalization board of trustees in Jan. 2001[9]. JPEG2000 utilizes 'tiling' which alludes to the dividing the input image into rectangular no- covering blocks that are compressed severally, similar to they were completely extraordinary pictures. Tiling decreases memory necessities and furthermore enhances the compression execution which makes JPEG2000 a cutting edge picture compression method.

### Transform Coding

It separates the first image into sub blocks of small size. In this basically data is transformed into any form. For every image or sub block coefficients values are measured. Changing over the first  $8 \times 8$  exhibit of pixel stream into a variety of coefficients nearer to the upper left corner for the most part contain the greater part of the data expected to quantize and encode the image with minimal perceptual damaged. The subsequent coefficients are then quantized and the yield of the quantizer is utilized by an image encoding techniques to create the yield bit stream expressing to the encoded image [17].

### Fractal Compression

Fractal compression is also used for image compression as well as data based on fractals of image. But in this method some data is lost. It takes a shot at the way that components of a image regularly take after different parts of a similar picture. This technique changes over these parts into numerical information. This information is called "fractal codes". Encoded image is produced using this method. During this initially parcels an image into non covering  $8 \times 8$  pieces known as limit blocks and structures a space pool containing all of

perhaps covered  $16 \times 16$  pieces, related with 8 isometrics from reflections and pivots, called area squares.

### Joint Photographic Expert Group (JPEG)

Its main purpose is to compress the graphics and photos. In Joint Photographic Expert Group, a grey-scale image is breaks into  $8 \times 8$  pixel squares reduce the quantity of estimations. The complete Joint Photographic Expert Group technique will acknowledge an extensive variety of bits per pixel, including the utilization of shading data.

### Moving Pictures Experts Group (MPEG)

It is a compression method for computerized video groupings. It was designed in 1988 and its first meeting was held on May 1988 in Canada, Ottawa. MPEG additionally accommodates the compression of sound track related with the video. For example, utilized as a part of PC video and advanced TV stations. It has a few highlights the motion picture can be played forward or in invert, and at either typical or quick speed. It can be partitioned into two sorts of pressure: inside the-outline and between-outline. Inside the-outline pressure implies that individual edges making up the video grouping are encoded as though they were normal still pictures. It provides good quality video.

### Wavelet Difference Reduction

It is wavelet based strategy that is presented by Tian and Well [10]. WDR technique can detect both the texture and colour features. It is based on Run length encoding. In this method wavelet transform passed to an image. A wavelet transform is connected to a picture, and after that the bit-plane based WDR encoding calculation for the wavelet weights is done. The term distinction reduction alludes to the route in which Wavelet Difference Reduction encodes the areas of wavelet transform coefficients. It is a weight based method. During this, patch is input image that is made and weights assigned to every patch of the image. The patch that has dissimilar weight is removed from the image and different part of the image is left behind that is compressed image. This method retains the features like PSNR, Region of interest and Compression Ratio. Maximum compression ratio is required for high quality image.

## II.RELATED WORK

Benchikh and Corinthios [8] discussed the combination of two strategies Discrete Wavelet Transform and Discrete Cosine Transform for image compression. In this method they compared their methods with DCT compression technique and DWT compression technique. They showed the experimental results of the DCT-DWT better than DCT and DWT in terms of compression and quality of image.

Raja et al. [10] compared the WDR and ASWDR techniques. WDR technique contained all of the essential parameters like ROI (Region of Interest), Embeddedness, low complexity and dynamic SNR. The performances of the two methods were compared with three parameters PSNR, MSE and SNR values. The results shown that ASWDR performance is better than WDR when comparing SNR, Mean Square Error (MSE) and Peak Signal-to Noise Ratio (PSNR). That provides high PSNR, lower MSE.

Kau and Lin [13] described the switching coding strategy for compression based on Run Length and Adaptive Linear Predictive coding. In this method they proposed that this prediction results were good that can be obtained in both

slowly varying area and pixels around boundary. They used only casual pixels to estimate coding pixels and no other information needed to be transmitted. They also reduced the complexity of the algorithm with only a negligible degradation in entropy.

Jasmi et al. [14] presented the compression methods based on Huffman method, Discrete Wavelet Transform and Fractal coding. They measured the performance of their method using different parameters like Compression Ratio, Mean Square Error, Peak Signal to Noise Ratio and Bit per pixel etc. They concluded that fractal coding was better method for image compression as compared to Huffman coding and Discrete Wavelet Transform in terms of PSNR and Compression Ratio parameter.

Parmer [15] described the wavelets based image compression strategy. In this paper, the author also compared their method with Discrete Cosine Transform & Discrete Fourier Transform in terms of compaction of energy parameter. Wavelet based compression obtained from blocking artifacts that was problem in DCT.

## III.PROPOSED METHOD

The proposed lossless compression method used the decision tree for compression of a grey scale images.

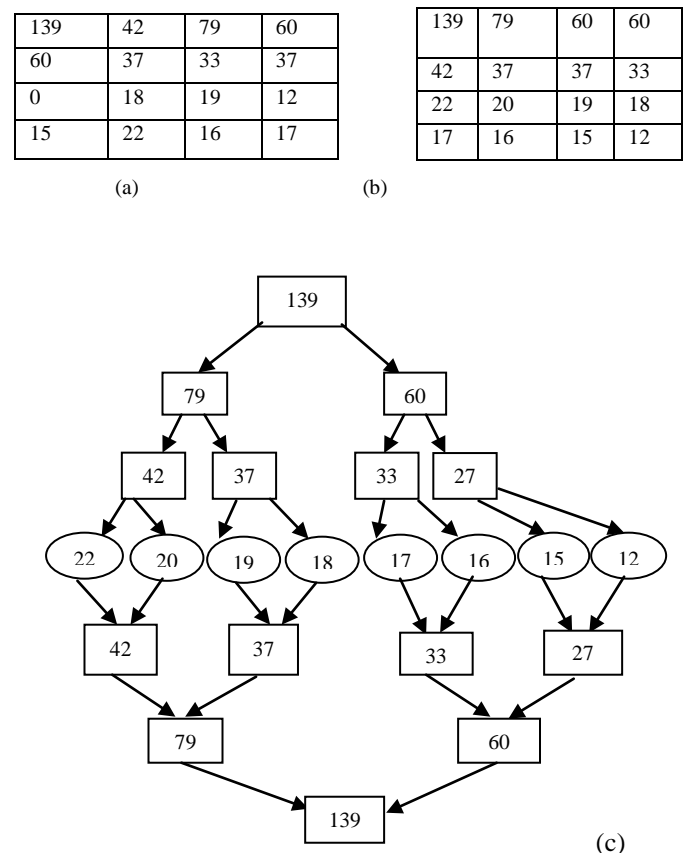


Figure 3. (a) Representation original data in patch. (b) Patches are sorted, (c) Decision tree of the patch

# Proposed algorithm for image compression and decompression

Input: Image need to compress

Output: Compressed Image

**Step 1:** Select the Input images which need to perform the compression

**Step 2:** Traverse the input image in the diagonal directional to cover each pixel in the image

**Step 3:** Check similarity of each pixel

for (k=0;K =n ;k++)

IF (k(n) is dissimilar to k+1)

Put k into the array of dissimilarity

Else

K = similar pixel

**Step 4:** Remove dissimilar pixels from the pixel

**Step 5:** Resize the original image to generate the compressed image.

Input: Compressed image for decompression

Output: Decompressed image

**Step 1:** Select the compressed image for the decompression

**Step 2:** Traverse the image to analyzed removed pixels from the image

**Step 3:** Check the dissimilar pixels for the image

**Step 3.1:** Execute for loop, for all pixels in the image

If (a(K,m)!=a(k+1,m+1))

Select a (K,m)

Else

Select (a (k+1,m+1))

End

**Step 4:** Insert the selected pixels in the input image

**Step 5:** Repeat step 3 and 3.1 until all pixels removed

**Step 6:** End

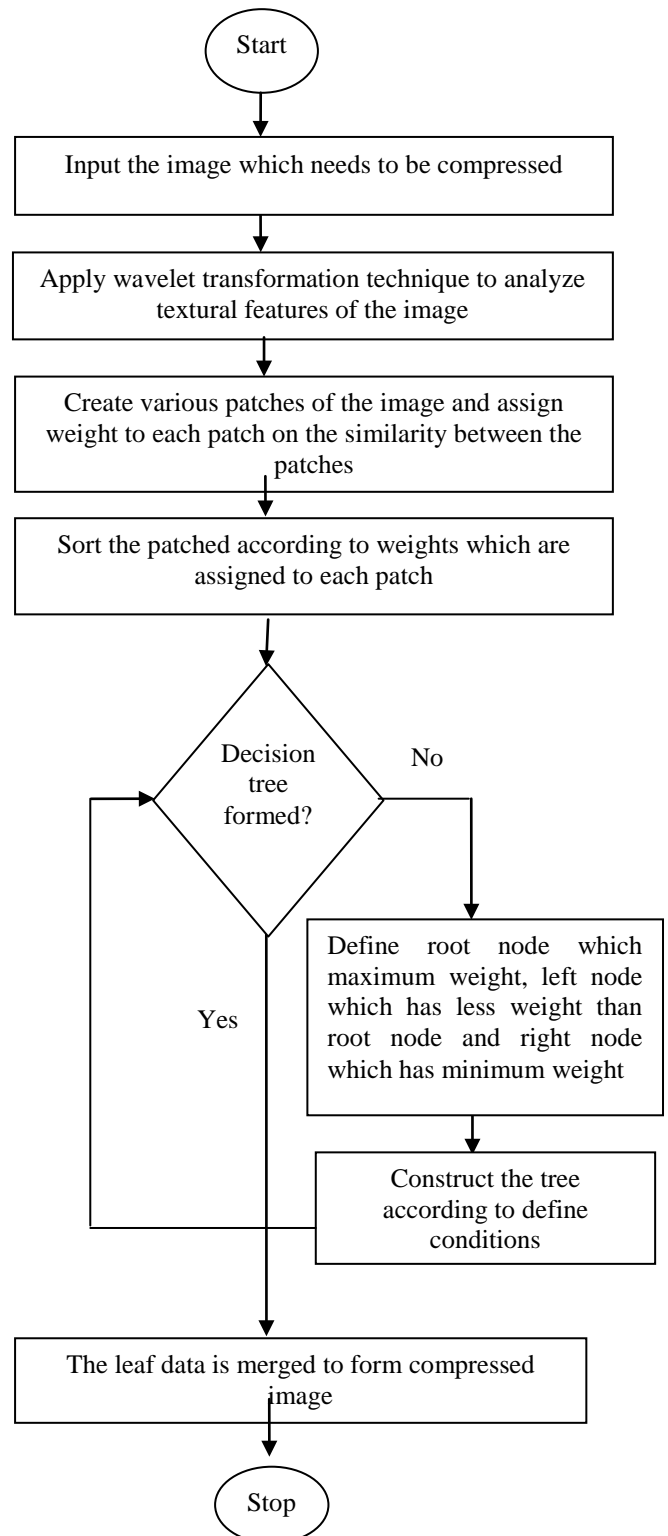


Figure 4. Flowchart of the proposed method

## IV.RESULTS AND DISCUSSIONS

In this section results of the proposed method is presented and discussed. In the proposed work decision tree algorithm has been used for compression of an image. The proposed technique has been implemented in MATLAB 2012a. Four grey scale images like Chess, Lena, Cameraman and Roots

images are used to test the proposed method. The result has been shown only on chess image. Not all the images result has been shown because of the shortage of space in research paper. The performance of the proposed method is measured using image quality measures (IQM) like Compression Ratio, MSE, Elapsed time, PSNR [10, 11]

- (a) **PSNR:** It is the peak error between the compressed image and input image given by the equation (1). The range of grey scale images lie between 0 to 255.

$$\text{PSNR} = 10 \log_{10} \left( \frac{255^2}{\text{MSE}} \right) \quad (1)$$

- (b) **MSE:** It is the difference between the compressed image and input image given by the equation (2).

$$\text{MSE} = \frac{1}{HW} \sum_{k=1}^H \sum_{l=1}^W [Y(k, l) - Z(k, l)]^2 \quad (2)$$

$Y(k, l)$  is input image and  $Z(k, l)$  is output image and  $HW$  are dimensions of an image.

- (c) **Compression Ratio:** The ratio between the total bits in an input image to total bits in the encoded image is known as Compression Ratio. The Compression Ratio given by the equation (3).

$$\text{CR} = \frac{b_1}{b_2} \quad (3)$$

$b_1$  represents the input image data rate and  $b_2$  represents the bit stream that is encoded.

- (d) **Elapsed time:** It is duration from when the processing started until the time it terminated. It is a measure from start to end including time that passes due to process delays or waiting for value to become available. The equations (4,5,6,7,8) are given as:

$$\text{Elapsed time} = \text{toc} \quad (4)$$

$$\text{toc}(\text{time val}) = \text{Display the time Elapsed} \quad (5)$$

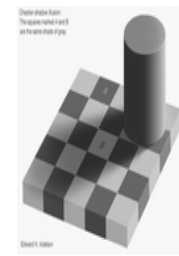
$$\text{Elapsed Time} = \text{toc}(\text{time val}) \text{ return the elapsed time} \quad (6)$$

$$\text{time val} = \text{value of the internal timer tic command} \quad (7)$$

$$\text{Elapsed Time} = \text{Finish Time} - \text{Start Time} \quad (8)$$

The performance of the proposed method is compared with Huffman method and Wavelet Difference Reduction technique. Different outputs of Chess image with different techniques are shown in Figure 5. (a) Shows the original image (b) Compressed image using Huffman method (c) Compressed image using Wavelet Difference Reduction method (d) Shows Compressed image using proposed method (e) Decompressed image.

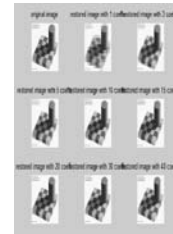
From the experimental results, it has been concluded that the proposed method is better than Wavelet Difference Reduction method and Huffman method.



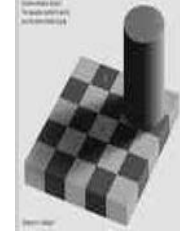
(a)



(b)



(c)



(d)



(e)

Figure 5. Image compression results for chess image (a) Original image (b) Compressed image using Huffman Method (c) Compressed image using Wavelet Difference Reduction Method. (d) Compressed image using Proposed Method (e) Decompressed image.

Comparison of the graphical representation of the proposed method with Huffman method and Wavelet Difference Reduction method in terms of shown. PSNR is shown in Figure 6, In terms of MSE is shown in Figure 7 and in terms of CR is shown in Figure 8.

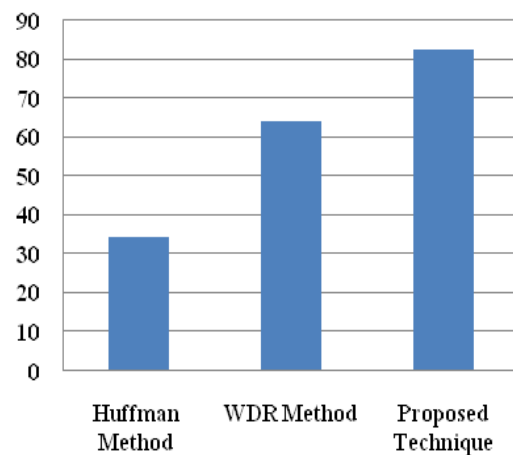


Figure 6. Comparison of the proposed method with Huffman method and Wavelet Difference Reduction method in terms of PSNR

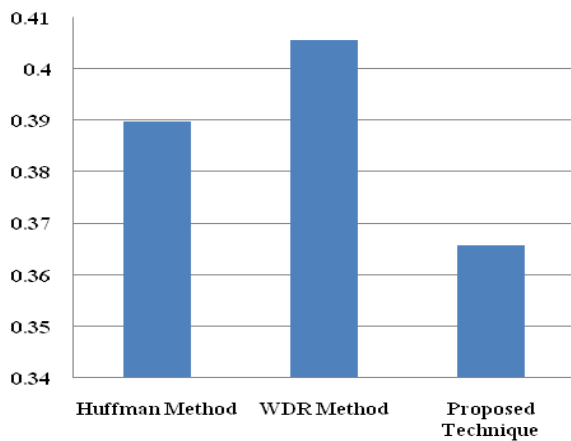


Figure 7. Comparison of the proposed method with Huffman method and Wavelet Difference Reduction method in terms of MSE

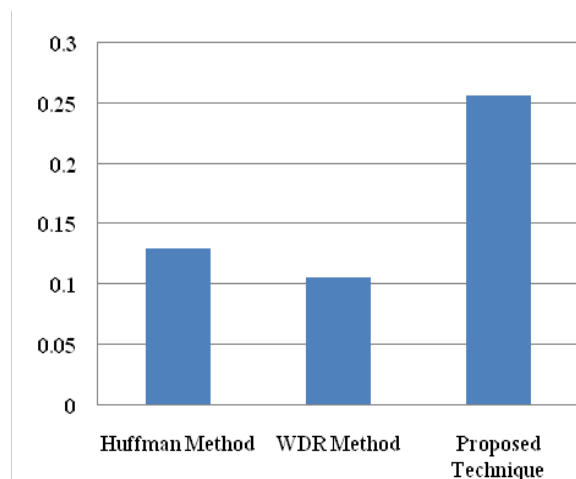


Figure 8. Comparison of the proposed method with Huffman method and Wavelet Difference Reduction method in terms of Compression Ratio

Table 1 Comparison between Huffman technique, Wavelet Difference Reduction technique and the proposed technique in terms of PSNR, Compression Ratio, MSE and Elapsed time for Chess image

Parameters	Compression Techniques		
	Huffman Technique	Wavelet Difference Reduction Technique	Proposed Technique
PSNR	34.52892	64.00525	82.37340
MSE	0.38974	0.40567	0.36586
CR	0.13030	0.10657	0.25625
Elapsed Time(sec)	5.597756	3.804984	0.054821

Comparison of the Huffman method, Wavelet Difference Reduction method and Proposed method in terms of PSNR, CR, MSE and Elapsed time.

## V.CONCLUSION

In this paper various compression methods are studied and reviewed for compression of grey scale image. In this method decision tree algorithm is used for image compression.

The performance of the proposed method is evaluated using Image Quality Measures (IQM) like Compression Ratio, MSE, Elapsed time and PSNR of an image. From experimental result it has been concluded that the proposed method increase the PSNR and CR (Compression Ratio) and as compared to Wavelet Difference Reduction technique and Huffman method.

## REFERENCES

- [1] Kaur, M. and Kaur N. 2015. A Literature Survey on Lossless Image Compression. International Journal of Advanced Research in Computer and Communication Engineering, 4(3): pp. 491-493.
- [2] Sharma, M. 2010. Compression Using Huffman Coding. International journal of computer science and network security, 10(5): pp.133-141.
- [3] Bansal, V., Gupta, P. and Tomar S. 2014. The Implementation of Run Length Encoding for RGB Image Compression. International Journal of Advanced Research in Computer Engineering and Technology, 3 (12): pp. 4397-4401.
- [4] Pujar, J. and Kadlaskar. 2010. A New Lossless Method of Image Compression and Decompression Using Huffman Coding Techniques. Journal of Theoretical and Applied Information Technology, 15(1): pp.18-23.
- [5] Kaimal, A. and Manimurugan, S. 2013. Image Compression Techniques: A Survey International Journal of Engineering Inventions, 2 (4): pp. 26-28.
- [6] Sharma, S. August 2010. Digital Image Processing. Department of Electronics & Communication Engg.
- [7] Anitha, S. 2015. Lossless Image Compression And Decompression Using Huffman Coding. International Research Journal of Engineering and Technology, 3(1): pp. 240-247
- [8] Benchikh, S and Corinthios, M. A Hybrid Image Compression Technique Based on DWT Transforms. Proc. of the IEEE Transactions on industrial electronics, June, Canada.
- [9] Islam, R., Baki, A. and Palash, S. 2009. A New Image Compression Scheme Using Repeat Reduction and Arithmetic Coding. International Conference on Computer and Information Technology, pp. 21-23 dec., pp. 209-214.
- [10] Raja, S. and Suruliandi, A. 2011. Image Compression using WDR & ASWDR Techniques with different Wavelet Codecs. Proc. of Int. Conf. on Advances in Computer Engineering, 1(2): pp. 102-105.
- [11] Grgic, S. and Grgic, M. 2001. Performance Analysis of Image Compression Using Wavelets. Proc. Of the IEEE Transactions on industrial electronics, June, Zovko-Cihlar, 682-695.
- [12] Geng, X. and Ding, Q. 2012. Similar-short periodicity analysis and application in image compression encryption of digital chaos. International Workshop on Chaos-fractals Theories and Applications. 18-21 Oct. .pp.167-171.
- [13] Kau, L. and Lin , Y. 2007. Least-Squares-Based Switching Structure for Lossless Image Coding. Proc. of the IEEE transactions on circuits and system, 54(7): pp. 1529-1541.
- [14] Jasmi, R., Permual, B., and Rajasekaran, P. 2015. Comparison of image compression techniques using Huffman coding, DWT and Fractal algorithms. International

- Conference on Computer Communication and Informatics (ICCCI), 08 -10 Jan. Coimbatore.
- [15] Parmer, H. 2014. Comparison of DCT and Wavelet based Image Compression Techniques. International Journal Engineering Development and Research. 2(1):pp.664-669.
- [16] Kaur, D. and Kaur, K. 2013. Huffman Based LZW Lossless Image Compression Using Retinex Algorithm. International Journal of Advance Research in Computer and Communication Engineering. 2(8): pp. 3145-3151.
- [17] Kaur, M. and Kaur, G. 2013. A Survey of Lossless and Lossy Compression Techniques. International Journal of Advance Research in Computer Science and Software Engineering. 3(2): pp. 323-326.