

# Image Compression Based on Wavelet and Quantization with Optimal Huffman Code

Ashanta Ranjan Routray  
Fakir Mohan University  
Balasore, India

Munesh Chandra Adhikary  
Fakir Mohan University  
Balasore, India

## ABSTRACT

This paper proposes a method for the compression of color images by Haar transform, quantization and construction of minimum redundancy code using conventional Huffman coding. Wavelet compression is accomplished by decomposing the row and column of the image matrix using the Harr transform. And the reconstruction of the image is feasible just from  $1/4^{\text{th}}$  of the decomposed image and even  $1/16^{\text{th}}$  of the decomposed image is enough for re-construction and the quality relies on the nature the image. A fast and effective histogram-based quantization is applied to the decomposed image. The weighted minmax quantization incorporates activity weighting, whereby obtaining high quality quantized image with significantly less visual distortion. Partition based Huffman coding divides symbols based on sorted probabilities of symbols into two equal halves and generates codes for each portioned symbols. Analytical and experimental results suggest that the optimum code can be generated for images with balanced binary partition Huffman coding, which is not only decodable but also offers the possibility of realizing an average code-word length that can be made arbitrarily close to the source entropy.

## General Terms

Image Compression, Decoding, and Prefix code

## Keywords

Entropy, Haar Transform, Quantization, Huffman code

## 1. INTRODUCTION

A common characteristic of image data is that, they contain a significant amount of information that is redundant. . The amount of data associated with visual information is so large and its storage requires enormous storage capacity. The transmission of this redundant data is wasteful of primary communication resources. For efficient data transmission, the redundant information should be removed from the signal prior to transmission [1]. The image coding may be lossless or lossy based on application. Huffman coding has been widely used in lossless image compression. The famous Huffman code is an instantaneous uniquely decodable block code which assigns shorter code words to frequent source symbols, and longer code words to rare source symbols of the image to be encoded for transmission and storage. The goal of this paper is to provide a complete coding system which includes wavelet transform and quantization plus balanced binary partition Huffman coding, and propose a novel context-based compression approach to provide high compression performance with reduced memory utilization.

## 2. HAAR WAVELET

As we are going to deal with compression of images, it is obvious that manipulation in raw image is quite impossible and if the image is represented in some mathematical form then the manipulation will be simpler and easier. Hence, the raw image needs to be transformed. To what extent a particular transform will support data compression depends on both the transform and the nature of images being compressed. The practicality of an image coding scheme depends on the computational workload of the encoding and decoding steps, as well as the degree of compression obtained. The availability of a fast implementation algorithm can greatly enhance the appeal of the particular transform. Some of the transformations are Sine transform, Cosine transform, Haar transform, Slant transform etc [2].

The Haar transform is a very fast transform. It is a compact, dyadic, orthonormal wavelet transform. The basic wavelet is progressively narrowed (reduced in scale) by power of two. Each smaller wavelet is then translated by increments equal to its width, so that the complete set of wavelets at any scale completely covers the interval (As the basis wavelet is scaled down by power of two, its amplitude is scaled up by powers of  $\sqrt{2}$ , to maintain orthonormality)[3].

The procedure goes like this. A low pass filter and a high pass filter are chosen, such that they exactly halve the frequency range between themselves. This filter pair is called the Analysis Filter pair. First, the low pass filter is applied for each row of data, thereby getting the low frequency components of the row. But since the low pass filter is a half band filter, the output data contains frequencies only in the first half of the original frequency range. So, by Shannon's Sampling Theorem, they can be subsampled by two, so that the output data now contains only half the original number of samples [4]. Now, the high pass filter is applied for the same row of data, and similarly the high pass components are separated, and placed by the side of the low pass components. This procedure is done for all rows. Next, the filtering is done for each column of the intermediate data. The resulting two-dimensional array of coefficients contains four bands of data, each labeled as LL (low-low), HL (high-low), LH (low-high) and HH (high-high). The LL band can be decomposed once again in the same manner, thereby producing even more subbands. This can be done up to any level, thereby resulting in a pyradecomposition [5]. Just as a forward transform to used to separate the image data into various classes of importance, a reverse transform is used to reassemble the various classes of data into a reconstructed image. A pair of high pass and low pass filters is used here also. This filter pair is called the Synthesis Filter pair. The filtering procedure is just the opposite - we start from the topmost level, apply the filters column wise first and

then row wise, and proceed to the next level, till we reach the first level [6, 7].

### 3. QUANTIZATION

In color quantization a true color image is transformed into a color-mapped image consisting of K carefully selected representative colors. The goal of this quantization is to discard information, which is not visually significant. Quantization is a many-to-one mapping, and therefore is fundamentally lossy. A fast and effective (improves image quality) method for color quantization, which uses a histogram to weight, each color proportion to its frequency, is presented here. This is a modified version of simple minmax algorithm proposed by Gonzalez. This quantization is applied to transformed image. The quantized image is then subjected to entropy coding and the coded image is stored in a storage medium.

Modification to simple minmax as follows

1. The centroid of each of the K clusters becomes the representative colors for each of K clusters.
2. The distance function is scaled based on there relative contribution to luminance to partially compensate for the non-uniform nature of the RGB color space:

$$d(x,y) = \sqrt{((0.25 \cdot (x-R-y \cdot R))^2 + (x \cdot G - y \cdot G))^2 + (0.5 \cdot (x \cdot B - y \cdot B))^2}$$

Above modification enhances the picture quality as compared to existing one. The final version of this quantization uses acting weighted histogram to incorporate activity weighting into the Minmax algorithm, whereby favoring colors with higher activity weight.

The activity level  $a(x,y)$  of a pixel at row  $x$ , column  $y$  of a true color image  $I$  is measured by the magnitude of the luminance gradient of the pixel given by

$$a(x,y) = |Y(I(x,y)) - Y(I(x-1,y))| + |Y(I(x,y)) - Y(I(x,y+1))| \quad (1)$$

Where  $Y$  is the luminance of a pixel

$$\text{and } x > 0 \text{ \& } y < H-1 \text{ (assumption)}$$

The activity weighting  $w(x,y)$  of a pixel at row  $x$ , and column  $y$  is calculated from its activity level using

$$\begin{aligned} w(x,y) = & 1/4 & \text{if } a(x,y) = 0 \\ & 1/3 & \text{if } a(x,y) = 1 \\ & 1/a(x,y) & 1 < a(x,y) < 6 \\ & 1/a(x,y) \cdot 1.25 & 6 < a(x,y) < 12 \\ & 1/161.25 & 12 < a(x,y) < 16 \\ & 1/231.4 & a(x,y) \geq 16 \end{aligned} \quad (2)$$

The quantization gives cluster based activity weighting which is lossy and can be used effectively for partition based Huffman coding.

#### Pseudocode Description Of Simple Minmax Quantization

Input: The set of N unique RGB colors in image and  
the numbers of colors to be selected.

Assumption:  $N > K$

Output: Selected representative RGB colors for image.

The basic steps of original minmax algorithm are as follows.

1. Initially all of the colors in the image are in same cluster, and its representative color is arbitrarily chosen.

2. Color that is farthest away from representative color (i.e max color) is taken as representative color of new cluster.

3. Colors that are closest to new cluster than to the existing one is moved to newly formed cluster.

Above two steps (2 & 3) are repeated until the number of clusters is equal to the number of selected colors. The representative colors of each cluster are the selected colors. These selected colors are used to map the original image to quantized image. When the activity level is extremely low (0 or 1) the pixel is considered to be part of flat region; therefore it is considered less important than when its activity level is low.

When the activity level is greater than 16, the pixel is considered to be part of an edge. Thus to prevent blurring of edges, the activity weighting of a pixel is never less than 1/161.25

The weighted histogram  $H_w$  of a true color image  $I$  is defined as a total function:

$$H_w : RGB \rightarrow R$$

Where  $H_w$  is the total activity weighting of the pixel.

By simply substituting the weighted histogram, any histogram-based algorithm becomes activity weighted. Activity weighting has been shown to be a fast and effective way to enhance the quality of images produced by histogram-based color image quantization techniques. This weighted histogram is used for calculating distance function. Weighted minmax quantization is selected for its better performance. For 1/4th of transformed image this quantization is applied before entropy coding.

### 4. ENTROPY CODING OF IMAGES

For a set of pixel values  $S$  with probability  $p(s)$ ,  $s \in S$ , the self information of  $s$  is:

$$i(s) = \log \frac{1}{p(s)} = -\log p(s)$$

Measured in bits if the log is base 2.

The lower the probability, the higher the information. Entropy is the weighted average of self information.

$$H(S) = \sum_{s \in S} p(s) \log \frac{1}{p(s)}$$

For a code  $C$  with associated probabilities  $p(c)$  the average length is defined as

$$l_a(C) = \sum_{c \in C} p(c) l(c)$$

We say that a prefix code  $C$  is optimal if for all prefix codes  $C'$ ,  $l_a(C) \leq l_a(C')$

A code  $C(X)$  is uniquely decodable [2] if, under the extended code  $C^+$ , no two distinct strings have the same encoding, i.e.

$$\forall x, y \in A_x^+, x \neq y \Rightarrow c^+(x) \neq c^+(y)$$

The Huffman coding algorithm generates an optimal prefix code [8].

#### 4.1 THE HUFFMAN CODING ALGORITHM

After the data has been quantized into a finite set of values, it can be encoded using an Entropy Coder to give additional compression. By entropy, we mean the amount of information

present in the data, and an entropy coder encodes the given set of symbols with the minimum number of bits required to represent them. The most popular technique for removing coding redundancy is Huffman coding. A binary decoder tree generally does decoding a stream of Huffman code. Huffman codes are built from the bottom up, starting with the leaves of the tree and working progressively closer to the root.

The objective of image compression is to represent the gray levels of an image with a minimum amount of data that is sufficient to describe completely an image without loss of information [9].

If the gray levels of an image are equiprobable, then entropy corresponding to maximum uncertainty. Since the source generates gray levels randomly, we require the knowledge of the statistics of the source. If some gray levels are known to be more probable than others, then we may compress data by assigning shorter code words to frequent source symbols, and longer code words to rare source symbols. The resultant code is variable length code[3].

Generally, there are two steps in constructing Huffman codes.

1. The source symbols are listed in order of decreasing probability. Take the two least probable symbols. Assign longest codewords of equal length, and differ only in the last digit.
2. Combine these two symbols into a single symbol, and repeat.

#### 4.2 BALANCED BINARY PARTITION HUFFMAN CODING

Least probable symbols are assigned with longest codewords. The number of bits in longest codewords can be reduced by dividing sorted source symbols based on probability. The code for two set of symbols are constructed independently by Huffman code tree. Finally two sets are merged to get codewords for each symbol by prefixing a 0 to codewords of first set which will have higher frequent symbols and a 1 to codewords of second set which will have lower frequent symbols. To decode partitioned Huffman code the minimum code length is 2 as it satisfies the condition of uniquely decodable.

**Table 1. Example of Optimal Huffman code and partition Huffman code**

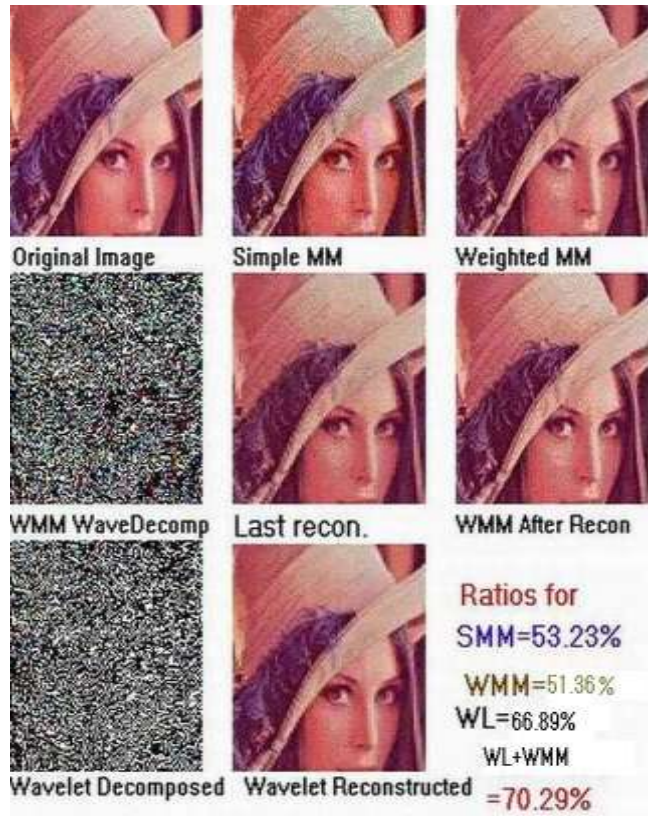
Symbols	Probability	HuffCode	PartitionHuff
A	1/2	0	00
B	1/4	10	010
C	1/8	110	0110
D	1/16	1110	0111
E	1/32	11110	10
F	1/64	111110	110
G	1/128	1111110	1110
H	1/128	1111111	1111

To decode the string ADG using Huffman coding we require 12 bits (0 1110 1111110) whereas for partitioned Huffman code it requires 10 bits (00 0111 1110). This technique is applied to

quantized images to reduce number of bits required to represent a digital image.

### 5. EXPERIMENTAL RESULT

In compressed file, the index table that contains the unique symbols and its code and the maximum and minimum length of the Huffman code are written [10,11]. The image can be reconstructed from this compressed file. The reconstructed images for complete coding system with simple minmax(MM), weighted minmax(MM), wavelet and minmax(WMM) and WMM Partitioned Huffman coding are shown in figure. The quality and compression ratio of the image codec is comparable with different image compression standards.



**Figure 1**

The table 1 below shows the peak signal to noise ratio (PSNR) of three test images by complete image codec after decompressing the color image.

**Table 1**

Images	Peak Signal to Noise Ratio
Figure 1 (Lena)	37.6542
Figure 2 (Waterfall)	41.9830
Figure 3 (Palace)	66.2156





Figure 2

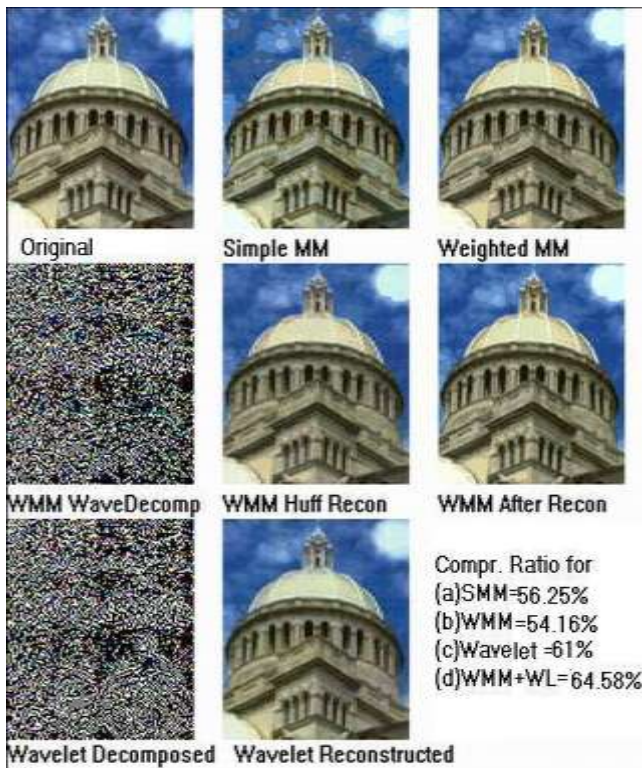


Figure 3

## 5. CONCLUSION

The compression ratio is good for Wavelet transform method and even by storing 1/16th of the image, the reconstruction of original image is found to be good. A comprehensive comparison between MinMax algorithm and weighted MinMax algorithm is presented. The quantized image has considerable contouring because it treats all colors equally. Weighted Minmax is a high quality quantization technique which uses a histogram to weight each color proportional to its frequency i.e, colors that occurs more often in the image are given greater precedence.

The running time of the Weighted MinMax algorithm is slower than the MinMax algorithm but the quality of quantized output produced was shown to be superior. The generalized activity weighting method was shown to be a fast and effective way to enhance the quality quantized images. After adding weighted minmax quantization between the wavelet transform and the partitioned Huffman coding, the compression ratio and quality of reconstructed image is found to be better than wavelet based image compression. The partitioned Huffman code gives optimum code for weighted minmax quantized pixel values by reducing number of bits for least frequent symbols.

## 6. REFERENCES

- [1] Communication Systems, 4th edition, Simon Haykin, John Wiley and Sons, INC
- [2] Anil K. Jain, "Fundamentals of Digital Image Processing", Prentice-Hall of India, India, 2001.
- [3] Robi Polikar, "Fundamental Concepts and An overview of the wavelet theory" – 2nd edition.
- [4] Cristos Chrysafis and Antonio Ortega, Member,IEEE "Line-Based,Reduced Memory, Wavelet Image Compression", IEEE Trans. VOL 9, NO 3, MARCH 2000.
- [5] Francois G. Meyer, Member, IEEE, Amir Z. Averbuch, and Jan-Olov Stromberg. "Fast Adoptive Wavelet Packet Image Compression",IEEE Trans. VOL 9, NO.5, MAY 2000.
- [6] Image Compression with Adaptive Arithmetic Coding, 2010 International Journal of Computer Applications (0975 - 8887) Volume 1 – No. 18
- [7] Sergio D. Servetto, Member, IEEE, Kannan Ramachandran, Member,IEEE,Vinay A. Vaishampayan, and Klara Nahrstedt, Member, IEEE, "Multiple Description Wavelet Based Image Coding" IEEE Trans. VOL. 9,NO. 5,MAY 2000.
- [8] Information Theory, Inference, and Learning Algorithms David J.C. MacKay Cambridge University Press 2003
- [9] D. A. Huffman, "A method for the construction of minimum redundancy codes," Proc. IRE, vol. 40, pp. 1098–1101, Sept. 1952.
- [10] Digital Image Processing Using MATLAB, Gonzalez, Woods, Eddins, Pearson Education
- [11] Lawson S. and Zhu J Image Compression Using Waveletsand JPEG2000", Comm, Electronics and Commn. Engg. Journal, June 2002.