IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 12, December 2015

Lossless Image Compression using Hybridization of Entropy Encoding and Data Folding

Malwinder Kaur[1], Navdeep Kaur[2]

Student, Department of CSE, (Rayat) Bahra Group of Institutes, Patiala, Punjab, India [1]

Assistant Professor, Department of CSE, (Rayat) Bahra Group of Institutes, Patiala, Punjab, India [2]

Abstract: The compression models play an important role in saving the storage space on the local or online storage or for the smooth transfers over the give network architecture. The robust and lossless compression is always remained as the point of interest for the researchers. The higher compression ratios heavily impact the data quality, which has turned the research over compression in various regions of the world. There are several entropy encoding models, arithmetic encoding like encoding techniques, etc has already been used for the purpose of data compression. The wavelet models have also been proposed in various compression models under various research projects. In this paper, the author worked over the robust, stronger and lossless compression model for various model of data using the amalgamation of Huffman encoding, Run-length encoding, LZW encoding along with Discrete wavelet transform (DWT) based wavelet compression model in order to provide the highly robust and lossless compression. The discrete wavelet transform (DWT) has already been proved as the best image compression algorithm. The DWT technique decomposes the image matrix into various sub-matrices to create a compressed image. The new compression technique will be developed by combining the most effective and fast wavelets of DWT technique for image compression. The quality of the new image compression technique will be evaluated using the peak signal to noise ratio (PSNR), mean squared error (MSE), compression ratio (CR) and elapsed time (ET). Also, the new techniques would be compared with the existing image compression techniques on the basis of the latter mentioned parameters. The experimental results have proved the effectiveness of the proposed model in the terms of MSE, CR and ET.

Keywords: Huffman encoding, run-length encoding, LZW encoding, DWT.

I. INTRODUCTION

Compression is the process of reducing the size of a file or of a media such as high-tech graphical images etc, by encoding its data information more efficiently. By doing this, there is a reduction in the number of bits and bytes used to store the information. Therefore, a smaller file or image size is generated in order to achieve a faster transmission of electronic files or digital images and a smaller space required for downloading. Compression basically employs four types of redundancy in the data:

- **Temporal**: This is present in 1D data, 1D signal, Audio etc.
- **Spatial**: It occurs due to correlation between neighbouring pixels or data items.
- **Spectral**: This is present due to correlation between color or luminescence components. This uses the frequency domain to exploit relationships between frequencies of change in data.
- **Psycho-visual**: This redundancy exploits perceptual properties of the human visual system.

1.1 General Compression Model

Compression is achieved by applying linear transform, quantizing the resulting transform coefficients and entropy coding the quantized values.

Image Compression Model as shown in figure 1.1 consists of three components namely Source Encoder, Quantize and Entropy Encoder.

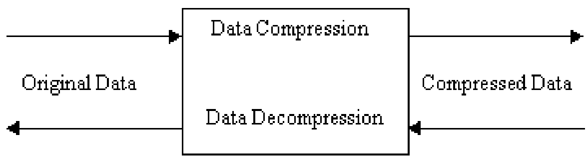


Figure 1.1: Image Compression Model

(i) Source Encoder: In the first step, signal is processed with a reversible linear mathematical transform to map the pixel values onto a set of coefficients which are then quantized and encoded. This step is intended to de correlate the input signal by transforming its representation in which the set of data values is sparser which compact the information content of the signal into smaller number of coefficients. The choice of transform to be used depends on a number of factors such as computational complexity, coding gain etc. For compression purpose, higher the capabilities of compressing information in fewer coefficients better the transform. Most widely used transform coding techniques are DCT (Discrete Cosine Transform), DWT (Discrete Wavelet Transform), etc.

(ii) Quantizier: It is an irreversible step. It represents the lossy stage in the process. A good quantizer tries to assign more bits for coefficients with more information content and fewer bits for coefficients with less information content, based on the given fixed bit rate. The choice of the quantizer depends on the transform that is selected.

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 12, December 2015

Some quantization methods perform better with particular transform methods. Quantization can be performed on each individual coefficient is called Scalar Quantization (SQ). Quantization can also be applied on a group of coefficients together known as Vector Quantization (VQ).

(iii) Entropy Encoder: It removes redundancy from the output of the quantizer. This redundancy is in the form of repeated bit patterns in the output of the quantizer. The frequently occurring symbols are replaced with shorter bit patterns while infrequently occurring symbols are replaced with longer bit patterns, resulting in a smaller bit stream overall. It uses a model to perfectly determine the probabilities for each quantized value and produces an appropriate code based on these probabilities so that the resultant output code stream is smaller than the input stream. Most commonly used entropy coding techniques are RLE (Run Length Encoding), Huffman Coding, Arithmetic Coding, Lempel-Ziv (LZ) algorithms, etc. The properly designed quantizer and entropy encoder are absolutely necessary along with optimum signal transformation to get the best possible compression.
Data compression has several advantages such as:

- Extending the life of the source node. Unless the source has data to transmit, and consumes less energy at the radio transceiver. This is true as the complexity of the compression algorithm adopted is reliable enough to be profitable. The compression process should not cost more in terms of energy gain that it brings about communication, if the presence of a compression process could reduce the lifetime of the node.
- Extending the life of the intermediate nodes. For the same reasons, reducing the amount of data at the source will be necessarily beneficial to the nodes responsible for relaying packets between the source and the destination node. They receive fewer packets of data, so they have fewer packets and acknowledgments to transmit.
- Reduced risk of congestion. A decrease in the amount of data transmitted over the network will lead to reduced risk of network congestion, thus a decrease in packet loss and transmission delays.
- Tolerance to losses. Some reinforcement on tolerance to packet loss can be achieved by applying some processing mechanisms at the source, such as mixing or tattoo image.

II .PROPOSED METHODOLOGY

1. **Input data:** The data input has been performed in double type procedure. The first procedure has been programmed for the pre-defined input data in the form of large sized matrix, small sized matrix and image data (testing). Also the data input module has been programmed to receive the input in the form of custom text data or image data. The input data matrix can be defined as the following:

$$f(x) = \begin{cases} 1, & \text{when input is pre} - \text{saved data} \\ 0, & \text{when input is custom data} \end{cases}$$

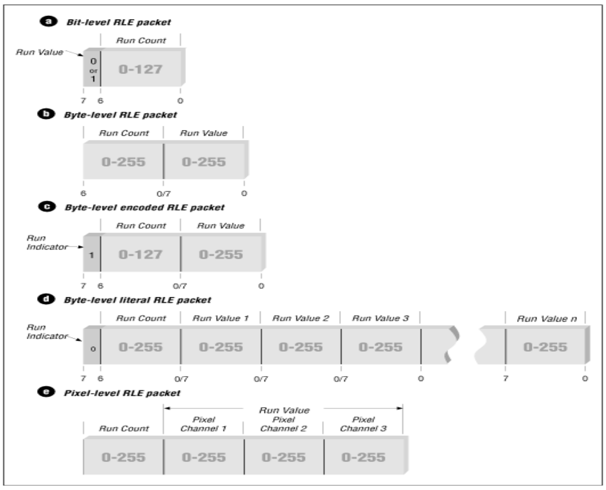
For the pre-saved data, when f(x) is set to 1:

$$f(x) = \begin{cases} 0, & \text{For large file} \\ 1, & \text{For small file} \\ 2, & \text{For image data} \end{cases}$$

For the custom data, when f(x) is set to 1:

$$f(x) = \begin{cases} 1, & \text{For text input} \\ 0, & \text{For image input} \end{cases}$$

2. Apply Run Length Encoding: RLE is the simplest approach for the lossless data compression in the continuous data, which is arranged in the particular length sequences such as an n by m matrix. The RLE requires all of the values to be stored as the single data and count of the data should be appropriate to contain the successful runs of the data on different iterations for the proper validation. RLE may increase the data size in the files don't have runs or continuous patterns. The RLE encoding can be defined with the following sequence:



3. Apply Huffman Encoding: Huffman can be classified as the prefix code model based encoding scheme for the data compression. The Huffman coding was developed by David A. Huffman for the purpose of constructing the minimum redundancy codes. The Huffman encoding scheme returns the variable-length code with the table to store the source symbol data similar to the characters in any language. The algorithm constructs the symbol table on the basis of estimated probability and the frequency of the patterns, which describes the weight for each possible value for the given source symbol.

Input:
Alphabet A={a_1, a_2, a_3 …. a_N}, for the symbols or alphabet list of size n
Set W = {w_1, w_2, w_3 …. w_N}, is the set of positive weights assigned to each symbol (usually termed as proportional to probabilities), i.e. wi= weight (ai), $1 \le i \le n$

Output: Code C(A,W) = {c_1, c_2, c_3 …. c_N}, which is the tuple of binary codeword's or codebook, where each c_i codeword in the give codebook for symbol a_i defined using the condition, $1 \le i \le n$

Goal: Let $L(C)=\sum_{i=1}^{n} w(i) * \text{length}\{c(i)\}$ become the weighted path length of the codebook C. Condition $L(C) \le L(T)$ for any code T(A,W).

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 12, December 2015

4. Apply LZW Encoding: The Lempel-Ziv-Welch (LZW) encoding is the universal lossless data compression model to obtain the high throughput. The algorithm is widely used in the UNIX applications of the data compression. The dictionary of the strings is initialized with the length 1, and the then the longest strings are marked in the input string. The dictionary index is emitted for each string coefficient W, which is added to output and removed from input. The iterative method is applied on all possible W coefficients in given string to return the compressed data.

5. Apply Discrete wavelet compression: This section illustrates the wavelet compression technique with pruning proposal based on discrete wavelet transform (DWT). The wavelet based compression technique first decomposes an image into coefficients called sub-bands and then the resulting coefficients are compared with a threshold. Coefficients below the threshold are set to zero. Finally, the coefficients above the threshold value are encoded with a loss less compression technique. The main equation is applied over the given data as described in the below equation:

$$y[n] = (x * g)[n] = \sum_{k=-\infty}^{\infty} x[k] g[n-k]$$

The high and low frequency bands are extracted from the input data using the following equations:

$$y_{\mathrm{low}}[n] = \sum_{k=-\infty}^{\infty} x[k] h[2n-k]$$

$$y_{\mathrm{high}}[n] = \sum_{k=-\infty}^{\infty} x[k] g[2n-k]$$

III.RESULTS AND DISCUSSIONS

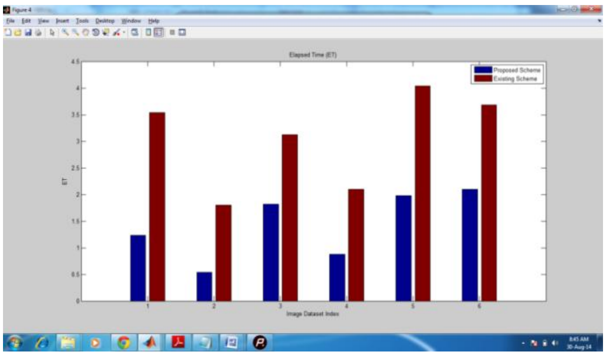


Figure I: Elapsed time between proposed and existing system

Elapsed time is the total time taken by system to execute its operations for compression mechanism on the selected data. The above graph has clearly shown that proposed algorithm has done way better than the existing algorithm. The elapsed time of the proposed algorithm is lower for all image categories in the dataset.

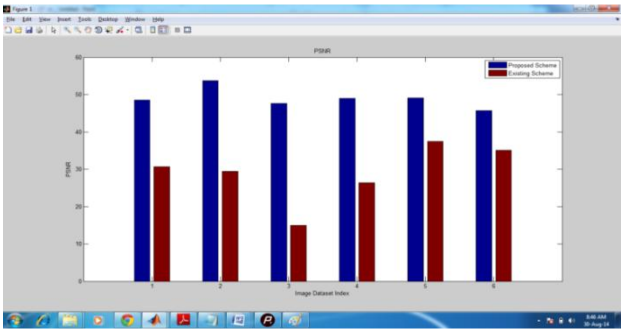


Figure II: PSNR comparison between proposed and existing system

PSNR represents the quality of the image by comparing images of before and after processing on the selected image data. The above graph has clearly shown that proposed algorithm has done way better than the existing algorithm in the terms of PSNR. The PSNR value is higher in the case of proposed algorithm than the existing algorithm for all image categories in the dataset, which shows that proposed algorithm creates clearer image at the end of the processing.

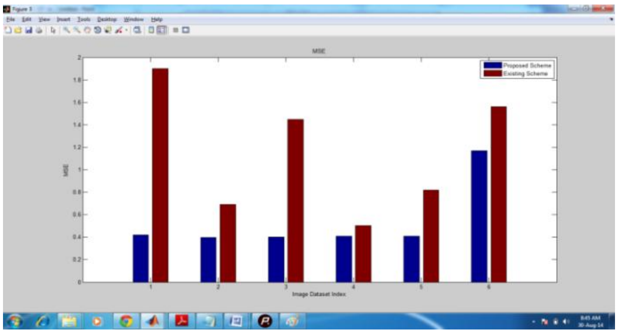


Figure III: MSE comparison between proposed and existing system.

Mean squared error is calculated by calculating the error bits over all bits, which represents the total error in the received data when it is compared to the data sent at the other end or data before and after processing. MSE value should be less to represent the less damage to the quality of the image. In the above graph, the MSE value for proposed system is lower as compared to the existing system on different image categories in the image dataset.

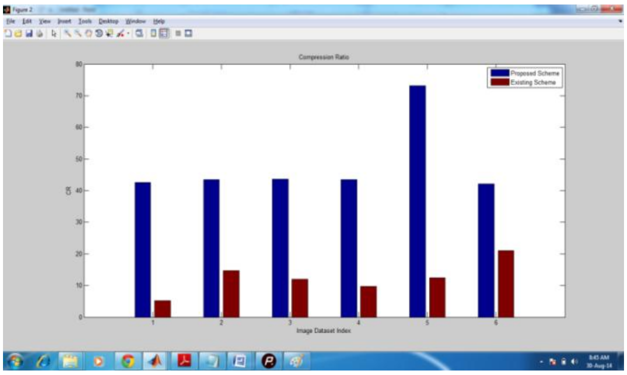


Figure IV: Compression Ratio based comparison between proposed and existing system.

Compression Ratio represents the reduction in the size of the image after the compression process. Higher is the compression ratio; lower is the transmission effort and disk space consumption. In the case of proposed model, the compression is recorded way higher than the existing model.

IV. CONCLUSION

The aim of the proposed project is to design and implement a new compression algorithm using wavelet filters in unique and effective combination. The proposed compression technique has been decompose the image matrix and has been reproduce the compressed image. It has already been proved in number of researches that the wavelet transform techniques are very effective for image compression. The proposed technique has been compared with the existing ones on the basis of various existing techniques on the basis of peak signal to noise ratio (PSNR), mean square error (MSE), elapsed time (ET) and Compression Ratio (CR).

The new technique has been the unique combination of the various compression techniques with the DWT in MATLAB simulator. The proposed model has been design and implement to provide more robust compression using the entropy encoding techniques (such as LZW, Run-Length and Huffman) along with DWT compression. The proposed model has been implemented using the MATLAB. The algorithm has been designed and the design to build stronger, effective and lossless compression technique by keeping an eye over the image matrix quality.

Additionally, the DWT based compression model has been evaluated for their performance with the existing compression models. The performance evaluation tells us about the best and worst performing coefficients among the all being surveyed under this research project. The highest compression ratio of approximately 70% has been achieved through the proposed model test bed simulation.

V. FUTURE WORK

The proposed model has been proved itself as the best coefficient or combination of the coefficients among the other existing coefficients. It means the new coefficients can be carried forward for the future developments. The proposed model can be taken as a base for the future developments on the compression coefficients based on DWT. In the future, a new series of the coefficients may be also proposed based on the various mathematical or statistical computations as per the existing.

ACKNOWLEDGMENT

Thanks to my family, my guide **Er. Navdeep Kaur** (AP, CSE Department), **Er. Harpreet Kaur** (AP & HOD, CSE Department), Bahra Group of Institute, Bhedpura, Patiala for having faith in me and allowing me to work with all terms and conditions and advising me time to time about this survey paper.

REFERENCES

1. Sandeep Kaur, Gaganpreet Kaur and Dr. Dheerendra Singh (2013), "Comparative Analysis of Haar And Coiflet Wavelets Using Discrete Wavelet Transform In Digital Image Compression "International Journal of Engineering Research and Applications (IJERA) Vol. 3, Issue 3, pp.669-673 669
2. Noor HudaJa'afar et.al (2013) "Distributed Arithmetic Architecture of Discrete Wavelet Transform (DWT) with Hybrid Method"
3. Oussama GHORBEL and Walid AYEDI (2012)," Images Compression in WSN: Performance Analysis"IEEE.
4. Aggarwal, Preeti. and Rani, Babita (2010), "Performance Comparison of Image Compression Using Wavelets", International Journal of Computer Science & Communication, Vol No. 1, Issue No. 2, pp. 97-100.
5. Al-lahan, Mohammed. And El Emary, Ibrahim M. M. (2007), "Comparative Study between Various Algorithms of Data Compression Techniques", Proceedings of the World Congress on Engineering & Computer Science.
6. Anish Kumar, M. S., Roy, Rajesh Cherian. And Gopika kumari, R. (2006), "A New Image Compression and Decompression technique based on 8x8 MRT", GVIP Journal, Volume 6, Issue 1, July 2006.
7. Ashraf, Robina. And Akbar, Muhammad. (2005), "Diagnostically Lossless Compression of Medical Images-2", IEEE-2005, pp.227-230.
8. Averbuch, A. and Zheludev, VA. (2004), "A new family of spline based biorthogonal wavelet transforms and their application to image compression", IEEE Transactions on Image Processing, 2004, 13, (7), pp. 993–1007.
9. Baligar, V. P., Patnaik, L. M. and Nagabhushana, G. R. (2006), "Low complexity and high fidelity image compression using fixed threshold method" , International Journal of Information Sciences, 176 (2006), pp. 664-675.
10. Galabov, Miroslav. (2003), "Fractal Image Compression", International Conference on Computer Science & Technologies-CompSys Tech 2003.
11. Ms. Sonam Malik and Mr. Vikram Verma (2012), "Comparative analysis of DCT, Haar and Daubechies Wavelet for Image Compression", International Journal of Applied Engineering Research, ISSN 0973-4562 Vol.7 No.11.
12. B. B. S. Kumar and P. S. Satyanarayana (2013), "Compression and Denoising - Comparative Analysis from still Images using Wavelet Techniques", ITSI Transactions on Electrical and Electronics Engineering (ITSI-TEEE) ISSN (PRINT): 2320 – 8945, Volume -1, Issue -6.
13. Amrut N. Patel, Ravindra M. Patel, D. J. Shah (2013), "Performance Analysis of Wavelet Families for Image Compression", International Journal of Engineering and Innovative Technology (IJEIT) Volume 2, Issue 12.