IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol-6, Special Issue-4, May 2018 E-ISSN: 2347-2693

A Hybrid Lossless Encoding Method for Compressing Multispectral Images using LZW and Arithmetic Coding

S. Boopathiraja[1], P. Kalavathi[2*], and S. Chokkalingam[3]

[1,2,3]Department of Computer Science and Applications, The Gandhigram Rural Institute (Deemed to be University) Gandhigram, Tamil Nadu, India

[*]*Corresponding Author:* pkalavathi.gri@gmail.com*:*

Available online at: www.ijcseonline.org

Abstract—Most of the remote sensing images are multispectral image where these images are acquired in the form of several bands that constitute a spectral direction. As large amount of data is represented by multispectral image, a lot of memory space is needed for storage and transmission. Hence, there is big need for compression methods for multispectral images. The prime factor of any image compression method is the redundancy as well as correlation on an image. In this way, the multispectral images having high degree of correlation on spatial domain and redundancy on spectral domain. This leads to conception of several compression methods for these multispectral images. Moreover, every tiny information from multispectral image is very important for efficient processing and so the lossless encoding is always preferable. In this paper, we proposed a hybrid lossless method using Lempel-Ziv-Welch (LZW) and Arithmetic Coding for compressing the multispectral Images. The performance of our method is compared with existing lossless compression methods such as Huffman Coding, Run Length Coding (RLE), LZW and Arithmetic Coding.

Keywords— Lossless Compression, Multispectral Image, Huffman Coding, LZW, Run Length Coding, Arithmetic Coding

I. INTRODUCTION

Image storage and transmission is the biggest issue in this digital world. These issues can be avoided using the image compression techniques [1]. Image compression act as a vital role in digital image processing. Image compression is the technique of reducing unwanted or redundant information in the image. The ultimate aim of image compression is to reduce the size of an image so that it is possible to store more number of images in the less amount of memory as well as image transmission can be done very fast [2]. Compression is achieved by eliminating the coding redundancy, inter pixel redundancy and psycho-visual redundancy [3]. Coding redundancy refers to use of code which is more length instead of use logics. Inter pixel redundancy refers to pixel correlation of an image. Psycho visual redundancy is about human perspective that is some part of an image may not be required, those unwanted parts are eliminated using the compression techniques. Image compression technique are classified into two types such as lossless image compression and lossy image compression. As the name itself indicated there is some loss of information in the reconstructed image in the lossy compression methods [4]. Another one image compression technique is called lossless image compression. In this method, the reconstructed image is same as the original image and there is no loss of information.

Satellite images are taken by using the different sensors in the satellite. These images have a variety of uses that including cartography, military intelligence and meteorology. Satellite images can either be visible light images, water vapour images or infrared images [5]. The raw (unprocessed) satellite images are very huge in memory size so it needs more memory space to store huge amount of data. At the same time, those images had more time complexity while transmission. To avoid this type of problems, the satellite images need to compressed so that it reduces the space and transmission complexities [6]. Most of the satellite images can be Multispectral Image, Hyperspectral Image and Panchromatic Image. The remote sensors measure energy at many wavelengths that constitute a multispectral image which has seven bands. The lossless compression method is more compact for multispectral images because of every small amount of information on such images is important for above said applications which uses satellite images especially multispectral images [7]. In this paper, we proposed a hybrid lossless compression on multispectral images. The multispectral image is encoded using LZW and further the compressed details of LZW are encoded using Arithmetic Coding. Our proposed method is then compared against the existing lossless methods such as Huffman Coding, Run Length Coding, LZW and Arithmetic Coding.

Rest of the paper is organized as follows, Section I contain the introduction about our paper, section II reported the various lossless compression algorithms such as Huffman, RLE, Arithmetic Coding and LZW, section III contains the methodology of our proposed method and also various evaluation parameters are explained. In section IV the results of our proposed method and comparison of the results against various compression methods are discussed. In section V we concluded our paper.

II. LOSSLESS COMPRESSION METHODS

In this section various lossless compression methods existing are explained with its algorithmic steps.

A. Huffman algorithm

Huffman coding is classical data compression technique found by David Huffman. Huffman algorithm is a lossless compression technique [8]. Earlier, it is used for data compression and later it incorporated for images due to its coding efficiency. Huffman code procedure is based on the two observations [9].

- Symbol which is occurred more frequently will have smaller code words as well as symbol which occurred less frequently will have larger code words.
- The two least frequently occurred symbols have the same length.

The lowest probable symbols are merged and this merging process is repeated until only two probabilities of two compound symbols are left and thus a code tree is generated and Huffman codes are obtained from labelling of the code tree. After the code has been created decoding is performed from that we obtain the original data or image without loss of information. The algorithmic steps involved in Huffman Coding are given below.

Step 1: Calculate the probability of each and every symbol of the input.
Step 2: Then probability of symbols is arranged in decreasing order from
Step 3: The lowest probabilities are merged and this process is continued till last two probabilities are left.
Step 4: Based on the details obtained from step 3, the Huffman Tree is generated and bits are allocated accordingly.
Step 6: By traversing the Huffman tree the original symbols are decoded.

B. Run Length Coding:

Another type of the lossless image compression is Run Length Coding [10]. It is one of the easiest and simplest technique when compared with other techniques. Each matrix was scanned row wise for identifying repetitive pixels. Each group of such repetitions were then replaced by the pixel value and the frequency of occurrence [11]. This was done exhaustively throughout the image matrix. For single occurrence of a particular value frequency was not used because that would cause an overhead affecting the compression. The algorithmic steps involved in Run Length Coding are given below.

Step 1: Read the pixel continuously from first pixel of an image. If it is last pixel then exit.
Step 2: If the value of next pixel is same as previous then count value increment by one. Otherwise, store the pixel value in new array.
Step 3: Finally, the sequence of ordered pair is obtained with the intensity and its corresponding length.
Step 4: For reconstruction, the empty array with input size is created
Step 5: Construct the i^{th} row of compressed image with putting run length value in reconstruct array from compressed array.
Step 6: Then, construct $i+1^{th}$ row then next row and so on to get the reconstructed image
.

C. Arithmetic coding:

Arithmetic is a form of entropy encoding used in lossless data compression as well as in image compression [12]. In arithmetic coding, frequently used character will be stored with fewer bits and others are stored in more bits. Arithmetic coding represents a message in between 0 and 1.

Step1: Range of real numbers are taken as [0,1). This range divided into sub-ranges according to the symbols in the input. This sub-range denotes a real value equal to the source symbol probability.
Step2: Find the sub-range for each input symbol belong.
Step3: Again, this ranges are subdivided to get the next level sub-ranges.
Step4: Move to the next symbol and do it as same.
Step5: Step3 and Step4 is Repeated until all the symbols in the input are parsed.

D. LZW (Lembel-Ziv-Welch):

Another type of the lossless image compression technique is Lempel-Ziv-Welch (LZW). It is a dictionary based compression algorithm. It scans the entire file to find any sequence of data that occur more than one time [13]. These sequences are stored in dictionary that created. References are put where-ever repetitive data occurred. LZW compression replaces strings of characters with single codes. The code that the LZW algorithm outputs can be of any arbitrary length, but it must have more bits in it than a single character.

Step 1: The dictionary is initiated such that all input strings have length as one
Step 2: The longest string is identified for the current input symbol with the use of dictionary
Step 3: The dictionary index is emitted this string for output and removed from input
Step 4: Go to next symbol and do the same until all the input symbols are processed

III. PROPOSED METHOD

A. Methodology

The flowchart of our proposed method is given in Figure 1.

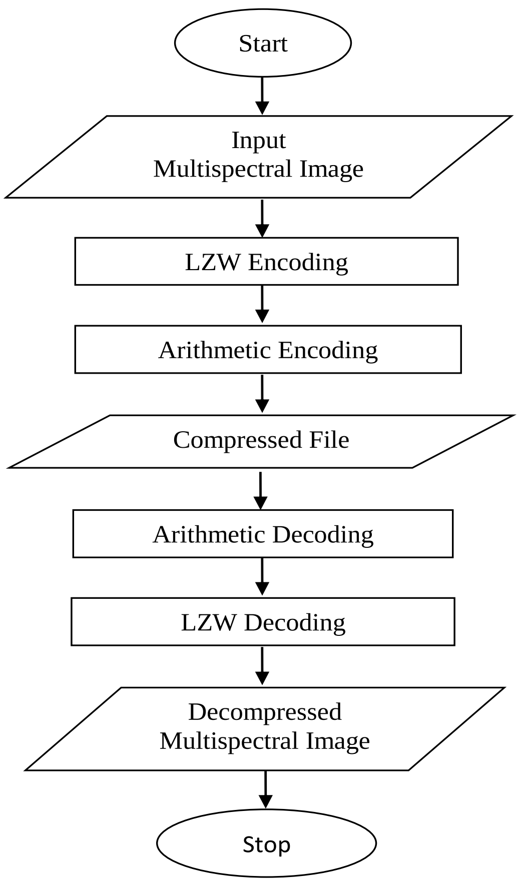

Figure 1: Flowchart of our proposed method

The input multispectral satellite image was taken to compress. In the first phase, the input image is encoded using the LZW encoding method. The compressed format of input image is the sent for the second encoding phase called Arithmetic encoding. In this phase, the encoding is done through the arithmetic encoding on the LZW compressed format of input image. Then, the inverse processes such as arithmetic decoding and LZW decoding are performed accordingly to get the decompressed image. The algorithmic steps of our proposed method are given below.

Step 1: Get the input Multispectral Image
Step 2: Applying LZW encoding on Input Image
Step 3: Applying Arithmetic Encoding on LZW Encoded Image
Step 4: Obtain the compressed bitstream
Step 5: Applying Arithmetical Decoding on compressed bitstream
Step 6: Applying LZW Decoding
Step 7: Obtain the Reconstructed image using the LZW decoding details

B. Evaluation Parameters

Two of the commonly used error metrics for qualitatively comparing various image compression techniques are Peak Signal to Noise Ratio (PSNR) and Structural Similarity Index (SSIM). Compression ratio (CR) and Bits per Pixel (BPP) are another metric for compare the original image and reconstructed image in terms of compression performance.

Compression Ratio (CR):
Compression ratio is defined as number of bit to represent the size of original image to the number of bit to represent the compressed image. Compression ratio can be determined using the following formula.

$$CR=n1/n2 \quad (1)$$

n1 - Number of bits in the original image
n2 - Number of bits in the compressed image

Bits per Pixel (BPP)
Number of bits are needed to encode the unit of pixel is known as BPP. The formula for BPP is given below

$$BPP=C/N \quad (2)$$

where, C- Compressed Image bits, N - Size of an Image.

Peak Signal to Noise Ratio:

The PSNR is the widely used measure to find the quality of the reconstructed image. The following formula is used to calculate the PSNR value.

$$PSNR=10.\log_{10} (MAX^2/MSE) \quad (3)$$

where, MAX is the maximum intensity pixel value of the image.

The MSE is the cumulative squared error between the compressed and original image. The following formula is used to find the MSE value

$$MSE = \frac{1}{MN}\sum_{y=1}^{M}\sum_{x=1}^{N}[Im(x,y) - Im'(x,y)]^2 \quad (4)$$

Structural Similarity Index (SSIM)

It is used to measure the similarity between the given original and compressed image. It is the perceptual metric to quantify the image quality degradation. The reference and the processed images are comparing through SSIM. Local image structure, luminance, and contrast are analyzed by SSIM metric.

$$SSIM(x,y) = [l(x,y)]^{\alpha}.[c(x,y)]^{\beta}.[s(x,y)]^{\gamma} \quad (5)$$

where, $$l(x,y) = \frac{2\mu_x\mu_y + c_1}{\mu_x^2 + \mu_y^2 + c_1} \quad (6)$$

$$c(x,y) = \frac{2\sigma_x\sigma_y + c_2}{\sigma_x^2 + \sigma_y^2 + c_2} \quad (7)$$

$$s(x,y) = \frac{2\sigma_{xy} + c_3}{\sigma_x\sigma_y + c_3} \quad (8)$$

where, μ_x and μ_y are the local means, σ_x and σ_y are the standard deviations and σ_{xy} is cross covariance for images x and y.

IV. RESULTS AND DISCUSSION

The input multispectral images are taken from the USGS (United States Geological Survey) website [14]. Our proposed method is compared against the various existing methods such as Huffman Coding, RLE, LZW and Arithmetic Coding. This comparison is depicted in the Table 1.

In Table 1, The Compression Ratio and the BPP values of all the methods including our proposed method are given. As all the methods are lossless in nature, the PSNR value is tends to infinity (i.e. the MSE values for all the methods is zero). Whereas the SSIM value of all method are one. Our proposed method gives increased amount of Compression Ratio as well as the reduced BPP. This indicates our proposed method yields better performance than other traditional lossless methods such as Huffman Coding, RLE, LZW and Arithmetic Coding. Also, for the quality measure PSNR and SSIM gave the lossless indication for our proposed method. The resultant image of proposed method is given in Figure. 2. The first band and sixth band of the Dataset 1 is given in Figure. 2.

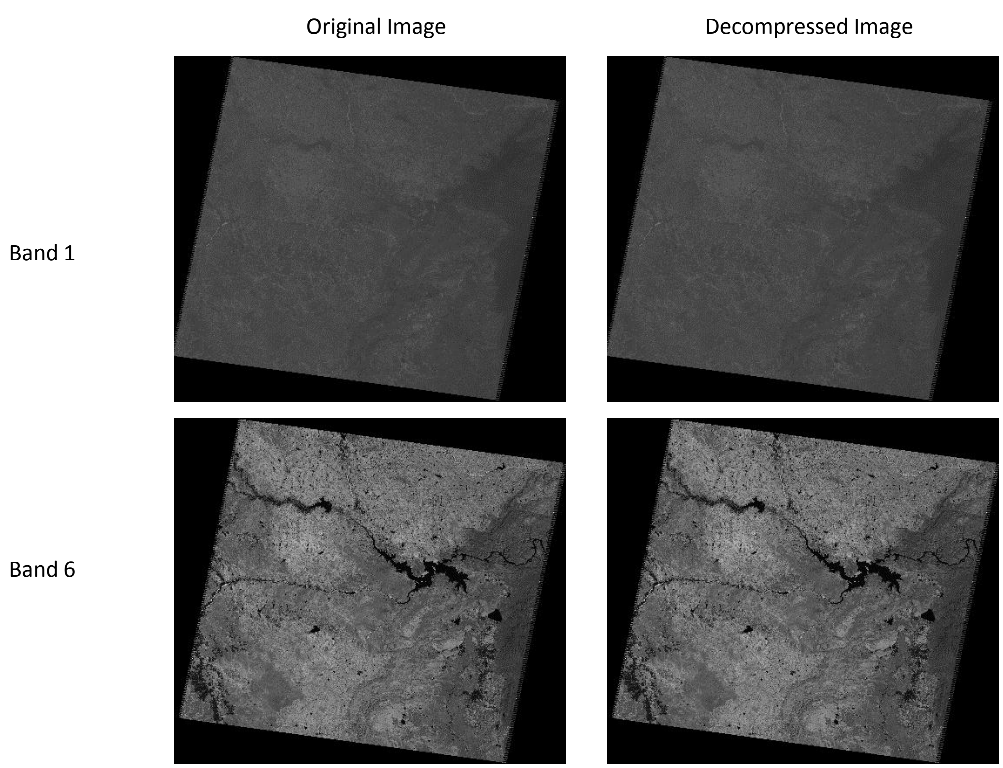

Figure 2. Resultant Images of Band 1 and Band 6

Table 1. Performance comparison of the proposed method

Dataset	Huffman		RLE		LZW		Arithmetic Coding (AC)		Proposed method (LZW+AC)		All the methods	
	CR	BPP	CR	BPP	CR	BPP	CR	BPP	CR	BPP	MSE	SSIM
1	1.34	5.98	0.77	10.37	0.80	10.02	1.34	5.96	1.62	4.94	0	1
2	1.34	6.00	0.79	10.14	0.76	10.52	1.34	5.93	1.62	4.94	0	1
3	1.35	5.94	0.77	10.33	0.83	9.60	1.35	5.91	1.64	4.89	0	1
4	1.42	5.64	0.82	9.77	0.81	9.84	1.42	5.61	1.74	4.58	0	1

V. CONCLUSIONS

As every information contained in the satellite multispectral images are important for some kind of applications, we need a lossless compression for multispectral images. Many of the lossless encoding of those images such as Huffman Coding, RLE, LZW and Arithmetic Coding are discussed elaborately in this paper and we proposed a hybrid lossless method that combining both the LZW and the Arithmetic Coding. With the results of work, we have concluded that our proposed method gives better performance than other existing lossless methods in both quality and compression factors.

REFERENCES

[1] M. Rabbani, and P.W Jones, *"Digital image compression techniques*", SPIE Press Vol. 7, 1991.

[2] P. Kalavathi and S. Boopathiraja "*A wavelet based image compression with RLC encoder*" National Conference on Computational Methods, Communication Techniques and Informatics, pp.289–292, 2017.

[3] J. Douglass, Ventana Medical Systems Inc," *Color image compression via spectral decorrelation and elimination of spatial redundancy"*. U.S. Patent 6,944,333

[4] A.M. Rufai, G. Anbarjafari, H. Demirel, "*Lossy image compression using singular value decomposition and wavelet difference reduction"*., Digital signal processing. Vol.1, Issue 24, pp. 117-124, 2014.

[5] The Sunitha Abburu, and Suresh BabuGolla, "*Satellite Image Classification Methods and Techniques: A Review*", International Journal of Computer Applications (0975 – 8887), Vol 119 – No.8, 2015.

[6] T. Markas and J. Reif, "*Multispectral image compression algorithms*", Data Compression Conference, pp. 391-400, 1993.

[7] ND. Memon, K. Sayood, SS. Magliveras, "Lossless *compression of multispectral image data",* IEEE Transactions on Geoscience and Remote Sensing, March 1994.

[8] D. Huffman, "*A Method for the Construction of Minimum-Redundancy Codes*", Proceedings of the IRE. Vol. 40, No. 9, pp.1098–1101, 1952.

[9] J.H. Pujar, LM. Kadlaskar*" A new lossless method of image compression and decompression using Huffman coding techniques*", Journal of Theoretical & Applied Information Technology,2010.

[10] EL. Hauck, inventor; Intelligent Storage Inc, assignee. "*Data compression using run length encoding and statistical encoding"*. United States patent US 4,626,829. 1986.

[11] P. Kalavathi, and S. Boopathiraja, "*A medical image compression technique using 2D-DWT* with run length encoding", Global Journal of Pure and Applied Mathematics (GJPAM), 2017.

[12] Witten, H. Ian, Neal, M. Radford, Cleary, and G. John, "*Arithmetic Coding for Data Compression"* (PDF). Communications of the ACM. Vol.6, issue.30, pp.520–540, 1987.

[13] J. Ziv, A. Lempel, "*Compression of individual sequences via variable-rate coding"*. IEEE transactions on Information Theory, Vol .5, issue.24 pp. 530-536,1978.

[14] www. glovis.usgs.gov

Authors Profile

.Mr. S. Boopathiraja is a Research Scholar in Department of Computer Science and Applications, The Gandhigram Rural Institute (Deemed to be University), obtained his Under-Graduate degree in Mathematics from GRI(DU); Post-Graduate Degree in Computer Applications (MCA) from Institute of Road and Transport Technology (IRTT), Erode; M. Phil in Computer Science from GRI(DU); He has two years of Research Experience and his research area focuses on Medical Image Anaysis and Digital Image Compression.

Dr. P. Kalavathi is an Associate Professor and Head in Computer Science and Applications, The Gandhigram Rural Institute (Deemed to be University), obtained her BCA degree from Mother Teresa Women's University, Kodaikanal; Post-Graduate Degree in Computer Applications (MCA) in GRI-DU; M. Phil in Computer Science from Bharathidasan University, Trichirappalli; and Doctoral Degree in Computer Science and Applications from GRI(DU), Gandhigram. She Qulified UGC-NET for Lectureship in 2000. She has 17 years of

Teaching and 10 years of Research Experience. The Author has more than 50 research articles in various journals and edited volumes. Her research area focuses on Digital Image Processing, Medical image Segmentation & Analysis and Medical Image Compression. She is also Serving as a reviewer of many international Conferences and various journals in IEEE, Springer, Elsevier etc. The author is Life Time Member of Indian Society for Technical Education (ISTE), New Delhi.

Mr S Chokkalingam pursed BCA in NPR Arts and Science College Dindigul, India in 2012 and Master of Computer Applications in PSNA College of Engg and Tech in Dindigul, India. He is currently pursuing M.Phil(C.S) in Gandhigram Rural Institute (Deemed to be University), Gandhigram, India.

