Comparative Analysis of Huffman and Arithmetic Coding Algorithms for Image Compression

Swetha Vura[1,2], Yamuna Devi C.R[3]

[1]*Assistant Professor, Dept. of ECE, School of Engineering and Technology, CMR University, Bengaluru, India*
[2]*Research Scholar, VTU, Belagavi, India*
[3]*Associate Professor & HOD, Dept. of ETE, Dr. Ambedkar Institute of Technology, Bengaluru, India*

Abstract - *Images are one of the most important visual representations used almost in every field. They require memory for their storage which necessitates a lot of space. Image compression plays a pivotal role in reducing the size of an image so that more images can be stored and thereby increasing the transmission speed. Many coding algorithms are written to compress images and reduce redundancy. This paper analyses and compares Huffman and Arithmetic coding algorithms with reference to image compression.*

Key Words: Image compression, Huffman coding, Arithmetic coding, compression ratio, compression time

1. INTRODUCTION

Image is represented as a 2D matrix in terms of rows and columns. The intersection of a particular row and column is called a pixel. To store these images, lot of memory and storage space is required. This gives rise to the need for image compression where the image size is reduced to accommodate more images. As the resources are finite, the techniques used in image compression have to effectively employ the storage space and the bandwidth [1]. Images find their application in almost every possible field like satellites [2-5], medical [6,7], artificial intelligence (AI), machine learning, IOT, robotics, computer vision, pattern recognition etc. The real challenge is to decrease the size without deteriorating the quality of the image. Joint Photographic Experts Group (JPEG) is a standard for compression [8] which comprises of steps including transformation, quantization and coding processes. The major classification of image compression is lossy and lossless compression techniques [9]. In lossy technique, we have data loss but in lossless, the decompressed and original image are identical [10,11]. Some of the lossless compression methods [12,13] are Shannon Fano Coding, Run-length Encoding, Huffman Coding, Arithmetic Coding, Lempel Ziv Coding etc. The coding algorithms are assessed [14] on various aspects like compression ratio, compression time, Peak signal-to-noise Ratio (PSNR) etc.

2.COMPARISON BETWEEN HUFFMAN AND ARITHMETIC CODING ALGORITHMS

Huffman coding is a type of entropy encoding algorithm which finds the improved method of encoding strings of data depending on their respective frequency. On the basis of the assigned frequency, the minutest variable length code is designated to each character based on its frequency of occurrence in the data. This coding incorporates a distinct method for selecting the representation for each symbol, which results in a prefix-free code [15] where the more recurrently used characters are allotted smaller codes and less recurrently used characters are allotted larger codes. Huffman algorithm builds an expanded binary tree of minimal weighted path length from a series of weights. This list consists of probabilities of symbol occurrences. This Huffman tree is used to define variable sized representations.

Arithmetic coding circumvents the thought of substituting an input symbol with a distinct code. The less frequently used symbols are represented with more bits and more frequently used symbols are represented with less bits [16]. The channel of input symbols is substituted with a single point floating number which lies between 0 and 1. This unique number can be decoded to construct the exact channel of symbols that went into its design (Fig. 1).

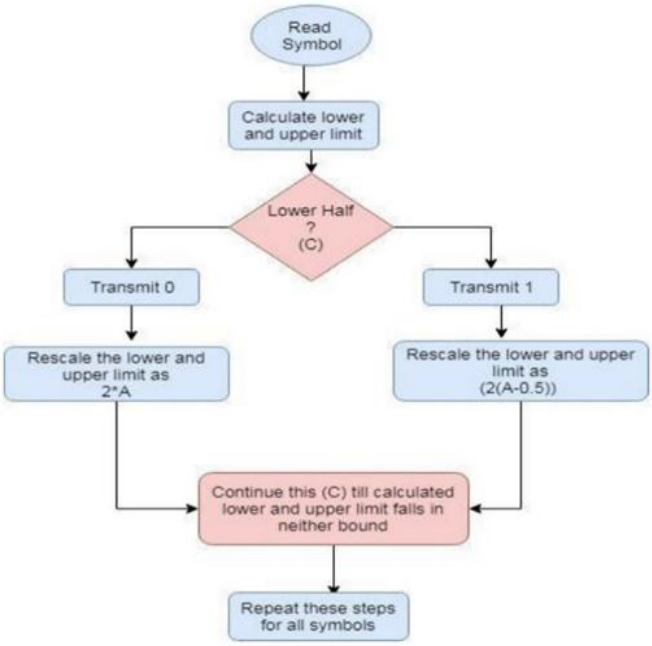

Fig -1: Block Diagram of Arithmetic Coding Algorithm [16]

3. LITERATURE SURVEY

The efficiency of Adaptive Huffman and Arithmetic coding is compared for different sets of images in [17]. An application named CMedia Compressor is proposed where analysis is performed on a quantitative basis. The sample images which were selected for the process of compression are of nature and architecture and taken from the public domain. The proposed application provides the option to select either the Adaptive Huffman Coding or Arithmetic Coding. The parameters calculated are compression time and space-savings to determine the efficiency of the coding algorithms. The simulation results conclude that Arithmetic coding has less compression time for all the sample images and high space-savings parameter for most of the cases. For text files, Adaptive Huffman coding performed better compared to Arithmetic coding.

A hybrid combination of DWT and DCT is used to analyse the performance of Huffman and Arithmetic Coding in [18]. The color images of Lena and Baboon are used for the process of compression as depicted in Fig. 2. The original image is Level 3 decomposed by Discrete Wavelet Transform and the quantized sub bands are encoded with Huffman and Arithmetic coding techniques. Then the 2D DCT is applied to get the compressed image. This compressed image has to undergo inverse transformation using IDCT and decoded along with inverse DWT to obtain the decompressed image. On calculating PSNR and Compression Ratio (CR), it is inferred that Huffman coding has higher PSNR in comparison to Arithmetic coding but has less CR.

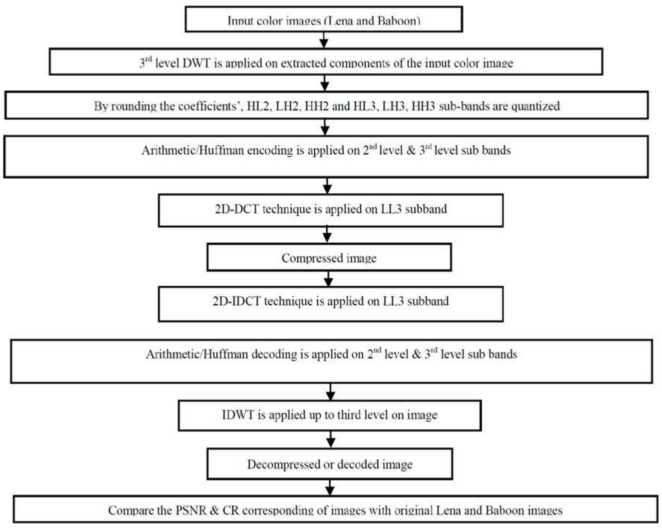

Fig -2: Flowchart of the process of compression and decompression [18]

Stereoscopic images taken in 3D have more depth and need more storage space in contrast to normal images. Compression of such images becomes a herculean task. The proposed method in [19] compares the results of stereoscopic compression using Arithmetic coding (SCAC) and stereoscopic compression using Huffman coding (SCHC). In the SCAC method, the left and right images are initially read to extract the RGB components. After quantization, Arithmetic coding is applied to both the images. The CR is first calculated. The PSNR value is obtained when the decoding and retrieving process is completed. The same procedure is similarly carried out for SCAC. Both the coding approaches are analysed for lossy as well as lossless techniques. The time taken for execution of SCHC is much more than SCAC. It can also be noticed that Lossy SCAC gives more CR and Lossless SCAC produces more PSNR.

The images captured by the satellite in the raw format occupy enormous amount of memory and are complex in nature. The images used for remote sensing are generally multispectral images which are large in size and hence require a lot of space for storage. In [20], a hybrid methodology is proposed which consists of Lempel-Ziv-Welch (LZW) Coding and Arithmetic Coding to compress such images. The simulation results for the proposed hybrid method are analysed in comparison with other lossless methods like Run Length Coding, Huffman coding etc. The multispectral image is first applied with LZW encoding. On the resultant encoded image, Arithmetic encoding is employed. Once the bitstream is compressed, it is decoded with Arithmetic and LZW methods. The deconstructed images (Fig. 3) are evaluated with parameters like CR, PSNR, Bits per Pixel (BPP) and Structural Similarity Index (SSIM). The results indicate that the adopted hybrid method performs better compared to traditional compression methods. Higher compression ratio and less BPP are achieved in the proposed hybrid methodology.

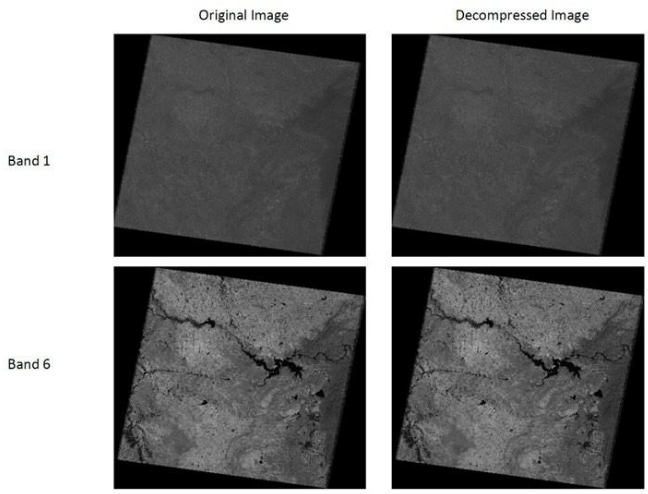

Fig -3: Resultant multispectral images of hybrid methodology [20]

Hospitals generate enormous data daily and therefore necessitate a lot of storage space. These medical images have to be compressed so that more number of images can be stored. Lossless image compression algorithms like Run-length encoding, LZW, Shannon-Fano, Huffman and

Arithmetic coding are analysed in [21]. RLE performs well only if the image has long runs of similar data. LZW is a lossless compression algorithm but is complex and consumes a lot of time. In Shannon-Fano coding, the generated output code may vary each time. There is a huge loss of data which is not acceptable. Huffman coding adopts a bottom-up approach where an optimal prefix-free code is generated (Fig. 4). The major disadvantage is that even if there is a small change in the code, the entire message is lost. A fixed number of bits are used in Arithmetic coding which gives better compression ratio but increases the compression time. It is concluded that Huffman coding surpasses other algorithms for real time applications.

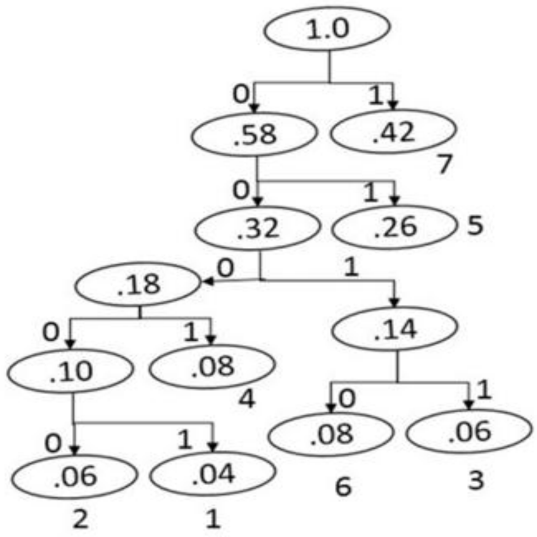

Fig -4: Encoding of Huffman tree [21]

Digital Image Processing has a plethora of applications which include remote sensing, medical, robotics, machine learning, artificial intelligence, pattern recognition etc. The algorithms and techniques required for image compression are analysed in [22]. JPEG compression inputs an image to subdivide it as n x n images. These sub-images are transformed using algorithms like DCT, DWT etc. They are then quantized where there is a particular loss of data. The quantized output is encoded using techniques like Run-length, Huffman, Arithmetic coding etc. It is observed that Huffman coding [23] has a faster compression rate while Arithmetic coding gives higher compression ratio [24, 25]. Huffman coding and Arithmetic coding are extensively used in the field of image compression. As both the coding techniques have their own limitations, an effective and alternate algorithm is proposed in [26] which is based on Huffman coding. Initially Huffman coding algorithm is employed to the data (Fig. 5) and then the resultant data is read whether it is long or short. The short bits are ignored and if the data has more number of bits, further compression takes place. Depending on whether the current and the next character are in the same leaf, the flag bits “1” and “0” are added. The addition of the flag bits ensure that the bits can either be truncated or written directly, avoiding repetition of data. An increased compression ratio of 15.30% was attained when compared to Huffman coding algorithm. This proposed algorithm was less complex and produced greater compression ratios than Huffman coding with the same compression time.

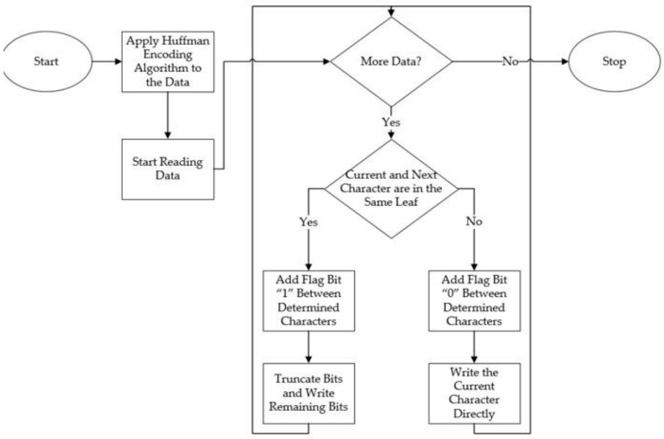

Fig -5: Flow diagram of proposed algorithm [26]

4. CONCLUSION

With the growing data and information in images in every field, image compression is inevitable. Huffman coding and Arithmetic coding have been applied and tested on a diverse set of images from satellite, multispectral, medical, architectural, stereoscopy etc. It has been observed that Arithmetic coding has higher compression ratio but increased compression time in comparison to Huffman coding. In terms of performance, Huffman coding surpasses Arithmetic coding with higher PSNR. Some of the hybrid methodologies have also been discussed in this paper which show better performance than traditional methods.

REFERENCES

[1] Uthayakumar, J., T. Vengattaraman, and P. Dhavachelvan. "A survey on data compression techniques: From the perspective of data quality, coding schemes, data type and applications." Journal of King Saud University-Computer and Information Sciences (2018).

[2] Gunasheela K.S., Prasantha H.S. (2018) Satellite Image Compression-Detailed Survey of the Algorithms. In: Guru D., Vasudev T., Chethan H., Kumar Y. (eds) Proceedings of International Conference on Cognition and Recognition. Lecture Notes in Networks and Systems, vol 14. Springer, Singapore. https://doi.org/10.1007/978-981-10-5146-3_18 .

[3] Shihab, H.S., Shafie, S., Ramli, A.R. et al. Enhancement of Satellite Image Compression Using a Hybrid (DWT–DCT)

Algorithm. Sens Imaging 18, 30 (2017). https://doi.org/10.1007/s11220-017-0183-6.

[4] Telles J., Kemper G. (2021) A Multispectral Image Compression Algorithm for Small Satellites Based on Wavelet Subband Coding. In: Iano Y., Arthur R., Saotome O., Kemper G., Padilha França R. (eds) Proceedings of the 5th Brazilian Technology Symposium. BTSym 2019. Smart Innovation, Systems and Technologies, vol 201. Springer, Cham. https://doi.org/10.1007/978-3-030-57548-9_17.

[5] M. Olaru and M. Craus, "Lossless multispectral and hyperspectral image compression on multicore systems," 2017 21st International Conference on System Theory, Control and Computing (ICSTCC), 2017, pp. 175-179, doi: 10.1109/ICSTCC.2017.8107030.

[6] Liu, F.; Hernandez-Cabronero, M.; Sanchez, V.; Marcellin, M.W.; Bilgin, A. The Current Role of Image Compression Standards in Medical Imaging. Information 2017, 8, 131. https://doi.org/10.3390/info8040131.

[7] Pardeep Kumar, Ashish Parmar, Versatile Approaches for Medical Image Compression: A Review, Procedia Computer Science, Volume 167, 2020, Pages 1380-1389, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2020.03.349.

[8] Shuyun Yuan, Jianbo Hu, Research on image compression technology based on Huffman coding, Journal of Visual Communication and Image Representation, Volume 59, 2019, Pages 33-38, ISSN 1047-3203, https://doi.org/10.1016/j.jvcir.2018.12.043.

[9] Khandwani, Fouzia I., and P. E. Ajmire. "A survey of lossless image compression techniques." International Journal of Electrical Electronics & Computer Science Engineering 5.1 (2018): 39-42.

[10] A.J. Hussain, Ali Al-Fayadh, Naeem Radi , Image Compression Techniques: A Survey in Lossless and Lossy algorithms., Neurocomputing (2018), doi: 10.1016/j.neucom.2018.02.094.

[11] Arora, Sunny, and Gaurav Kumar. "Review of Image Compression Techniques." International Journal of Recent Research Aspects 5 (2018): 185-188.

[12] Mander, Kuldeep, and Himanshu Jindal. "An improved image compression-decompression technique using block truncation and wavelets." International journal of image, graphics and signal processing 9.8 (2017): 17.

[13] Raghavendra, C., Sivasubramanian, S. & Kumaravel, A. Improved image compression using effective lossless compression technique. Cluster Comput 22, 3911–3916 (2019). https://doi.org/10.1007/s10586-018-2508-1.

[14] Wahba, Walaa Z., and Ashraf YA Maghari. "Lossless image compression techniques comparative study." International Research Journal of Engineering and Technology (IRJET), e-ISSN 2395-0056 3.2 (2016).

[15] Alistair Moffat. 2019. Huffman Coding. ACM Comput. Surv. 52, 4, Article 85 (July 2020), 35 pages. DOI:https://doi.org/10.1145/3342555.

[16] Harika Devi Kotha et al 2019 J. Phys.: Conf. Ser. 1228 012007

[17] P. Mbewe and S. D. Asare, "Analysis and comparison of adaptive huffman coding and arithmetic coding algorithms," 2017 13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), 2017, pp. 178-185, doi: 10.1109/FSKD.2017.8393036.

[18] Kumar, Gaurav & Kumar, Rajeev. (2021). Analysis of Arithmetic and Huffman Compression Techniques by Using DWT-DCT. International Journal of Image, Graphics and Signal Processing. 13. 63-70. 10.5815/ijigsp.2021.04.05.

[19] T. K. Poolakkachalil and S. Chandran, "Analysis of Stereoscopic Image Compression Using Arithmetic Coding and Huffman Coding," 2018 International Conference on Inventive Research in Computing Applications (ICIRCA), 2018, pp. 214-220, doi: 10.1109/ICIRCA.2018.8597216.

[20] S. Boopathiraja & Palanisamy, Kalavathi & Chokkalingam, S. (2018). A Hybrid Lossless Encoding Method for Compressing Multispectral Images using LZW and Arithmetic Coding. INTERNATIONAL JOURNAL OF COMPUTER SCIENCES AND ENGINEERING. 06.

[21] Rahman, M.A.; Hamada, M. Lossless Image Compression Techniques: A State-of-the-Art Survey. Symmetry 2019, 11, 1274. https://doi.org/10.3390/sym11101274

[22] P. T. Chiou, Y. Sun and G. S. Young, "A complexity analysis of the JPEG image compression algorithm," 2017 9th Computer Science and Electronic Engineering (CEEC), 2017, pp. 65-70, doi: 10.1109/CEEC.2017.8101601.

[23] Liu, X., An, P., Chen, Y. et al. An improved lossless image compression algorithm based on Huffman coding. Multimed Tools Appl (2021). https://doi.org/10.1007/s11042-021-11017-5.

[24] Salman, Nassir H. "Compare Arithmetic Coding to The Wavelet Approaches for Medical Image Compression" Journal of Engineering Science and Technology 16.1 (2021): 737-749.

[25] T. Koya, S. Chandran and K. Vijayalakshmi, "Analysis of application of arithmetic coding on dct and dct-dwt hybrid transforms of images for compression," 2017 International Conference on Networks & Advances in Computational Technologies (NetACT), 2017, pp. 288-293, doi: 10.1109/NETACT.2017.8076782.

[26] Erdal, E.; Ergüzen, A. An Efficient Encoding Algorithm Using Local Path on Huffman Encoding Algorithm for Compression. Appl. Sci. 2019, 9, 782. https://doi.org/10.3390/app9040782