A New Technique for Compression and Storage of Images

Dr. S. Selvam
Head & Asst.Prof., Dept. of Computer Applications,
NMSSVN College, Madurai.

Abstract:- **During the past decades we have been observing a permanent increase in image data, leading to huge repositories. [1]. This paper focuses on image compression and storage. We define methods to improve the efficiency and effectiveness of image compression .Image compression refers to represent an image with as few bits as possible while preserving the level of quality and intelligibility required for a particular application. The present work aims to develop an efficient algorithm for compression and storage of two-tone image.**

In this paper, an efficient coding technique is proposed and termed as line- skipping coding for two-tone image. The technique exploits 2-D correlation present in the image. This new algorithm is devised to reduce fifty to seventy five percentage of memory storage.

Keywords:- 2-D correlation, 1LSC, 2LSC, 3LSC, DCT, RLC.

I. INTRODUCTION

The need for compression and electronic storage of two-tone image such as line diagrams, weather maps and printed documents has been increasing rapidly. It has endless applications ranging from preservation of old manuscripts, Paperless office and electronic library etc. For text like English and Arabic, good quality optical character readers are available and provide good compression. But they accept only limited fonts and style of characters. Also for line diagram and text whose OCR'S are not readily available, the materials has to be considered as an image. Electronic storage of such image required a very large amount of memory. To reduce the memory requirements and hence the cost of storage, efficient coding techniques are used. A large number of coding techniques have been proposed and studied by different researchers. These techniques are broadly classified into two categories: Loss less and loss.[4].

Lossless techniques do not introduce any distortion. From the coded bit stream, we can reconstruct the digitized original image extract. Lossy techniques introduce some distortion to the reconstructed image while achieving high compression and retaining image usability.

A scan line of a two-tone image consists of runs black pixels separated by runs of white pixels. The spatially close pixel are significantly correlated, source coding technique exploit this correlation either along a single scan line or along many scan lines. Simplest and most commonly used techniques is run length coding, which exploit the correlation along with a single scan line to code the runs of black or white pixels. More complex techniques exploit the correlation along many scan lines to give better compression. However this is achieved at the cost of increased system complexity.[2].

In this paper, a very simple and efficient coding technique is proposed and termed as skip-line coding for two-tone image. The technique exploits 2-D correlation present in the image. The technique is based on the assumption that if there exist very high degree of correlation between successive scan lines, then there is no need to code each of them, only one of them need be coded and other may be skipped. While decoding, skipped lines are taken to be identical to the pervious line. This reduced the storage requirement significantly. The performance of this technique is compared with the run length coding. This paper also concentrates on Discrete Cosine Transform based compression so as to have comparative study of its performance.[3].

II. EXISTING SYSTEM

Image Data Compression is one of the major areas of research in image processing. There were several algorithms designed for compression of images. Here two methods have been explained for compression of image data, which is considered to be an existing system.

i) Block Truncation coding
ii) Run Length Coding

i) Block Truncation coding

In Block Truncation coding method, an image which is to be compressed will be divided into many disjoint blocks. In each block, mean and standard deviation are computed and each pixel will be coded as either zero or one depending on its level with mean. The recovered image is compared with original image and the error rate/pixel is calculated.[15].

ii) Run Length Coding:

Run length coding (RLC) is the simplest and most popular technique to code the runs of black and white pixels. It explicit the correlations along a single scan line. For storage, coding is done from top to bottom to achieve higher compression. The runs can be coded by fixed length code words. Since coding is done from top-to-bottom, maximum run length will be equal to the file size, in case the whole document is white or black, we divide runs into two categories: (i) Runs having length less than 255. (ii) Runs having length equal to or greater than 255.The runs between 0 and 254 are coded as single byte, when the runs are greater than 254.It is coded as three bytes. The first byte, FF indicates that the coded runs is greater than or equal to 255.The second byte gives the factor by which the run being coded is greater than 255. The third represents the remainder. Thought this way we can represent a maximum length of 65535. This coding scheme is not restricted to A4 size documents, but can

ISSN: 2278-0181
ICRADL - 2021 Conference Proceedings

be used for other standard type documents. Since black and white runs are coded by same codes there is no way to find tone of the pixels. Therefore one bit of information is provided to synchronize the color for runs of the pixels here synchronizing bit is provided only at the beginning and every alternate run will be of the same color.[6].

The decoded image is obtained by the following reverse process of coding. Since no information is lost in the coding process, reconstructed image is replica of the original image.

III. PROPOSED SYSTEM

In an image, very high degree of 2D correlation is presented. Efficient coding technique will exploit this correlation to give much higher compression. A number of such techniques are given. They provide maximum compression only if the correlations between the successive lines are perfect (i.e.) successive lines are similar. However similar lines will have similar codes. So they will skip the other similar lines. The proposed technique s based on this assumption. The algorithm is termed as **Line Skipping Coding Method**.

3.1. Line Skipping Coding Method

In this coding scheme, it is assumed that vertical correlations between successive scan lines are perfect. So they will have same code. First it is assumed that the two successive lines are similar, and then coding is performed only on one of them and skip the next line. This is termed as One Line Skipping Coding (1LSC). Here only half of the total scanned lines need to be coded. The runs of the black and white pixel are coded by using run length coding technique. [1].

Next, it is assumed that vertical correlation among three successive lines have similar codes, so it will perform on only one of them and skip the next two consecutive lines. This is termed as Two Line Skipping Coding (2LSC). Here only one third of the total scanned line need to be coded.

Finally coding is performed over one line and skipping three consecutive lines and this technique is termed as Skip Three Line Skipping Coding (3LSC). Here only one fourth of the total scan line is run length coded. It is clear that compression achieved through the proposed technique will be much higher compared to Run-Length Coding (RLC) but introduces distortion in the reproduced document. If line-to-line correlation is poor, the distortion will not be visible for documents having high line-to-line correlation. This process is shown in the figure1.

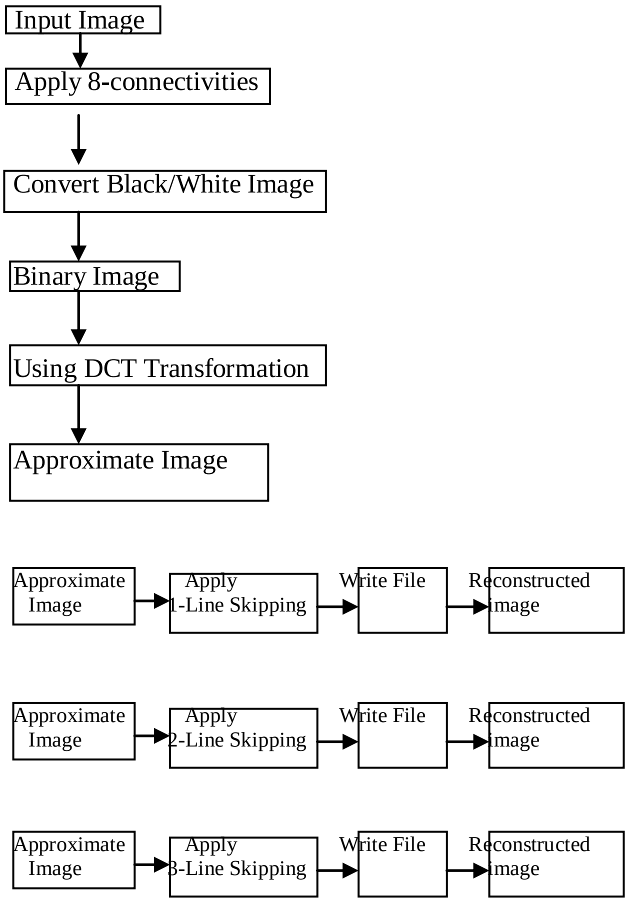

Figure:1 Process of the Line Skipping Coding Method

3.2. Algorithm for proposed method

1. Input to the image (gray or color image)
2. Apply m-connectivity relationship in the image.
3. Convert the image into Binary image.
4. Binary image is converted into Approximated image using DCT Transformation.
5. To get the color of first pixel.
6. Compute the Run length.
7. Compute One Line Skipping Coding (1LSC), Two Line Skipping Coding (2LSC) and Three Line Skipping Coding (3LSC).
8. Open and Display the compressed file for reconstruction.
9. Produce the reconstructed image matrix.
10. Display the reconstructed image.

3.3.DCT-Baseline Mode Coding

The algorithm for compression is used in the baseline sequential mode. First partition the original image into non-overlapping blocks of 8 X 8 pixels as shown below figure 2.

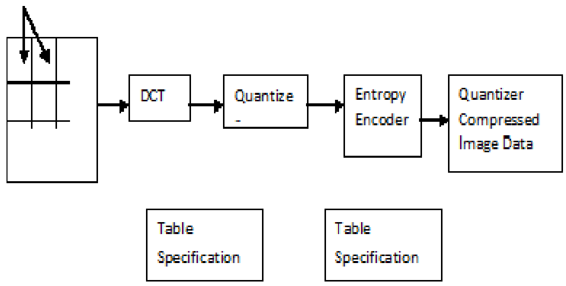

Figure: 2: Block diagram of JPEG Encoder.

Each block of 8 X 8 pixels , then transformed by using DCT into an array of 8 X 8 coefficients as illustrated by the equation-(1) given below.

$$F(u,v) = [1/4]\, C(u)\, C(v) \sum_{i=0}^{7} \sum_{j=0}^{7} f(i,j) x \cos[(2i + 1)u\, \pi/16] \cos[(2i + 1)v\, \pi/16]$$

Where $C(x) = \frac{1}{\sqrt{2}}$ for $x= 0$

$\quad 1$ Otherwise

f(i,j) = level shifted value of pixel(i,j) in image.

F(U,V) = DCT Coefficient with frequency indices (U,V); U,V= 0,1, … 7.

The first coefficient (0,0) of every block is called coefficient, while the rest of the coefficient are called AC coefficients. DCT (low frequency) AC coefficients as shown figure 3.

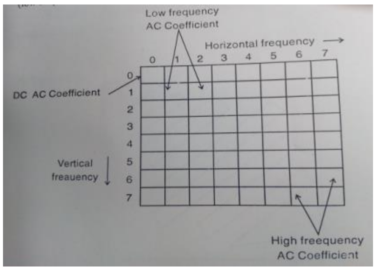

Figure 3: Block diagram for AC & DC coefficient

In other word, the higher frequency coefficients contain relatively less crucial details. All the coefficients are quintile using a uniform misstep quintile and rounded to the nearest integer as expressed in the equation()

$C(u,v) = F(u,v) \pm (Q(u,v))/2 \,/\, Q(u,v)$

Where, Q(u,v) = quantization step size for coefficient (u,v);

C(u,v) = rounded value of the quantized coefficient;

Each application can have its own quantization tables, which is usually designed to provide the best possible reconstructed image quality. In order to obtain a good subjective image quality, the DC and the low frequency AC coefficient are quantized by using large step sizes. Hence there is a tradeoff between the quantization steps size (i.e. image quality) and the compressive achieved: Smaller the steps size, better the image quality and smaller the compression ratio. The correlation between the DCT Coefficients of adjacent blocks and exploited by using DCPM to achieve further compression.

The DCT coefficient of each block is reordered in 1-D sequence using Zigzag scan as shown in the figure 4.

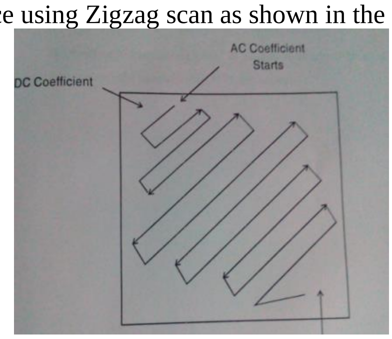

Figure 4: 1-D sequence using Zigzag scan

The scheme generates long runs or zero value coefficients (corresponding to the high frequency AC coefficient) in most of the image. The Zigzag ordered coefficient and run length and Huffman coded (i.e.) a code stored / transmitted for each DC coefficient and non zero AC coefficient indicating the magnitude and position in the Zigzag order.

Finally, the image blocks are rested scanned to generate the image bit stream.

The image is reconstructed by performing the decompression operations in the reverse order. Each block of 8 X 8 pixels is hence transformed back to spatial domain using the inverse discrete cosine transformation (IDCT) in equation(1).

$$F(u,v) = [1/4]\, C(u)\, C(v) \sum_{i=0}^{7} \sum_{j=0}^{7} f(i,j) x \cos[(2i + 1)u\, \pi/16] \cos[(2i + 1)v\, \pi/16] \quad \text{---(1)}$$

Where, f(i,j) = (i,j)th pixel in the reconstructed image block.

The baseline sequential algorithm is used to reconstruct the image in its original size at a specific image quality (SNR resolution).

IV PERFORMANCE EVALUATION

We select to test Line Skipping Coding Method, to compare the new system results with Run length coding method. The quality of the image after decompression has been found quantitatively by using signal to noise ratio.

The proposed method gives maximum compression than run length code and Discrete Cosine Transform coding. It gives an option of maximum compression of 98% to 99% at the cost of relative degradation in the output.

The following table:1 gives the performance of all the compression algorithms in terms of compression ratio and signal-to-noise ratio.

Table: 1 Comparison of Line Skipping Coding Method with other Methods.

Specimen Name	Original image size	Compressed file size in Bytes				
		RLC	1LSC	2LSC	3LSC	DCT Based
MOON	14,400	450	225	145	120	9.12 kb
SATURN	14,200	428	214	134	109	9.5 kb
HALF MOON	13,500	380	192	188	166	8.6 kb

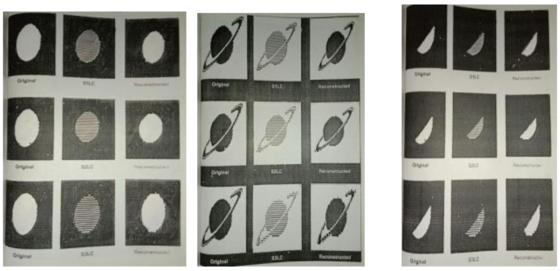

Figure 5: Sample Output of the Moon, Saturn and Half Moon Images.

RESULT:

Results are presented for these images. The images have been taken in its gray level form and have been converted to two-tone image if they are not in the two-tone format in its original form. The algorithm for run length coding(RLC), One Line Skipping Coding algorithm along with RLC, Two Skipping algorithm along with RLC,

Three Line Skipping algorithms along with RLC, and Discrete Cosine Transformation based compression have been coded in Mat Lab software and show result obtained in implementing them.

CONCLUSION:

In this work a novel method of compression image has been presented. The method presented exploits the usefulness of Run Length Coding technique. Also presented the variations of Run Length Coding Technique by introducing the concept of Line Skipping Coding Method. The technique based on DCT has also been carried out and presented. It has been proved that the proposed work shows superiority over the Standardized Discrete Cosine based method and simple run length coding method. Further work would involve in compressing image with bit planes after skipping lines. It would also involve in color and multispectral image coding.

REFERENCES

[1] O.R Mitchell and E.J. Delp, "Multilevel graphics representation using block truncation coding", proceedings of the IEEE, vol.68,no.7, pp.868-873,July 1980.

[2] J. Polec and J. Pavlovicova,; , "A new version of region Based BTC," EUROCON'2001, Trends in Communications,International Conference on. , vol.1, no., pp.88-0 vol.1, 4-7 July 2001.

[3] C.K. Yang, and W.H. Tsai, Improving block truncation coding by line and edge information and adaptive bit plane selection for gray-scale image compression, Pattern recognition letters,volume.16,number1,pages=67-75,1995.

[4] T.M. Amarunnishad, V. K. Govindan and Abraham T. Mathew, Improved BTC image compression using a fuzzy complement edge operator, signal processing operator, signal Processing, vol- 88, issue 12, (2008)2989- 2997, Elsevier 2008.

[5] T.M. Amarunnishad, V. K. Govindan and Abraham T. Mathew,Use of Fuzzy Edge Image in Block Truncation Coding for Image compression, International Journal of signal Processing,Vol 4, N0 3, pp 215-221, 2008.

[6] S.Selvam and Dr.S.Thabasu Kannan, "titled "Image retrival optimization with genitic algorithm", published IJAER ,International Journal of Applied Engineering Research as Special issue, Volume No 10, Issue No 55(2015) and IJAER is indexed by **SCOPUS,** EBSCOhost, Google Scholar, JournalSeek, J-Gate etc. And also listed in Anna University Chennai Annexure II-2014(Sl.No.8565).

[7] A. Aggoun and A. El-Mabrouk; , "Image compression algorithm using local edgedetection," Wireless Image/Video Communications, 1996., First International Workshop on, vol., no., pp.68-73, 4-5 Sep 1996.

[8] S.Selvam and Dr.S.Thabasu Kannan, "titled " A New Technique for Color-Based Image Retrieval System Using Histrogram", published IJAER ,International Journal of Applied Engineering Research as Special issue, Volume No 10, Issue No 82(2015), ISSN 0973-4562 and IJAER is indexed by **SCOPUS,** EBSCOhost, Google Scholar, JournalSeek, J-Gate etc. And also listed in Anna University Chennai Annexure II-2014(Sl.No.8565).

[9] U.Y. Desai, M. M. Mizuki and I. Masakiand Horn; B.K.P., Edge and mean based image compression,1996.

[10] R. Redondo and G. Cristobal; "Lossless chain coder for gray edge images," Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on , vol.2, no., pp. II- 201-4 vol.3, 14-17 Sept. 2003.

[11] D. E. Tamir , K. Phillip and Abdul-Karim, , "Efficient chain-code encoding for segmentation-based imagecompression," Data Compression Conference, 1996. DCC '96. Proceedings , vol., no., pp.455, Mar/Apr 1996.

[12] H. Sung and W.Y. Kuo. A skip-line with threshold algorithm for binary image compression. In Image and Signal Processing (CISP), 2010 3rd International Congress on, volume 2, pages 515-523. IEEE, 2010.

[13] E.J. Delp, M. Saenz and Salma, article BLOCK TRUNCATION CODING (BTC),2010.

[14] S.Selvam and Dr.S.Thabasu Kannan, "titled "An Empirical Review on Enhancing the Robustness of Multi resolution Water Marking", published IJAER ,International Journal of Applied Engineering Research as Special issue, Volume No 10, Issue No 82(2015), ISSN 0973-4562 and IJAER is indexed by **SCOPUS,** EBSCOhost, Google Scholar, JournalSeek, J-Gate etc. And also listed in Anna University Chennai Annexure II-2014(Sl.No.8565).

[15] Rafel C. Gonzalez and Richard E. Woods, "Digital Image Processing", Second Edition, Pearson Education Asia, 2005.