ISSN No. 0976-5697

Volume 8, No. 5, May-June 2017

International Journal of Advanced Research in Computer Science

REVIEW ARTICLE

Available Online at www.ijarcs.info

A critical study on the applications of run length encoding techniques in combined encoding schemes

Sanjoy Mitra
Dept. of Computer Science and Engineering
Tripura Institute of Technology, Tripura, India

Debaprasad Das
Dept. of Electronics and Communication Engineering
TSSOT, Assam University, Silchar, Assam, India

Abstract : Recent technological breakthrough in high speed processing units and communication devices have enabled the development of high data compression schemes Run Length Encoding (RLE) is one of the most significant entropy encoding compression techniques for compressing any type of data. Run length encoding algorithm performs compression of input data based on sequences of identical values. In this paper, we have studied and analyzed the applications of run length encoding technique in various combined encoding schemes and also pointed out their key features. Merits and shortcomings of respective encoding schemes are also pointed out.

Keywords: RLE, compression, decompression, compression ratio

1. INTRODUCTION

Data compression facilitates lowering down of storage space requirements, bandwidth requirements, reduced time on image viewing and loading, speedy transmission of ordinary data and multimedia data files. Lossless compression generates exactly alike the original data and mostly applied for compressing word processing files or spreadsheets, database records where exact reproduction of the original data is necessary. On the other hand, lossy compression permanently eliminates data and is particularly suitable for compressing digitized voice and graphics images where data loss beyond visual or real perception may be tolerated. In this paper we focus our study primarily on lossless RLE technique as this algorithm is easy to implement and easy to decode, memory space economy and faster execution. RLE is combined with different encoding schemes in order to improve the compression efficiency. Run-length encoding compresses runs of the identical value in a column to a close singular representation.

Here in this paper we also discuss various applied fields of RLE. For example, RLE is used to reduce power consumption during transmission in wireless sensor networks. Researcher shows that for single bit wireless transmission thousand times more energy required than a single 32-bit transmission [6]. Normal compression technique cannot save energy data, so the compression algorithms having lower memory space and less computational effort capability are well suited in wireless sensor networks.

Performance comparison of run length encoding with an application of Huffman coding for binary image compression is also discussed here in this paper.

Efficient selection of decoding aware data combined with run length coding of repetitive word pattern which is used for compressing bit stream of configuration files of FPGA in order to improve compression ratio, communication bandwidth and to reduce memory space, bit stream size and decrease decompression ratio is also discussed here in this paper.

2. REVIEW OF PRIOR WORKS

A. Combination of RLE and Huffman Coding to compress Fluoroscopy Medical Images [2]

CT, MRI and Fluoroscopy diagnostic images of a patient consists of a succession of images. Bunch of storage space is required besides the cost and time sustained during transmission. Lossless compression of medical data have a preference over the larger feasibility of lossy compression in order to preserve the accuracy. The method in [2] states about a lossless compression approach using correlation and combination of run Length and Huffman coding on the difference pairs of images classified by correlation and applied to pharynx and esophagus fluoroscopy images,. Their experimental results claims improved compression ratio of 11.41 as compared to 1.31 for the ordinary images

Their method consists of three fundamental steps:-

1. Correlation

They classified medical images into groups by using correlation coefficients (CC). The CCs were obtained between the first image and the second image for the first step, followed by the CC between the second image and the third image, and so on.

2. Preprocessing step

This step is accomplished by:-

a. Black area elimination from the fluoroscopy images.

b. Subtraction of white area from the fluoroscopy images. It does not contain any relevant data. ROI is extracted through steps a-b.

c. Image difference is computed by subtracting the test image from the reference image, as most images captured from the identical view are typically similar.

3. Coding step

Features like repetition and estimated probability of occurrence for each possible value of the source symbol motivates the combination of RLE and Huffman coding

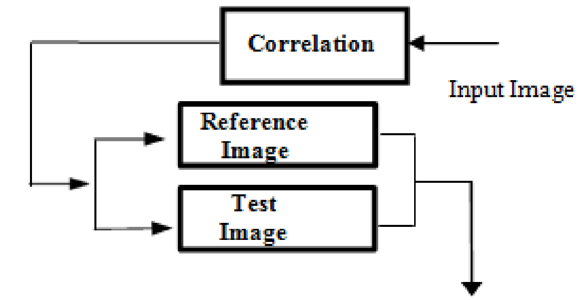

Fig.1 preprocessing steps [2]

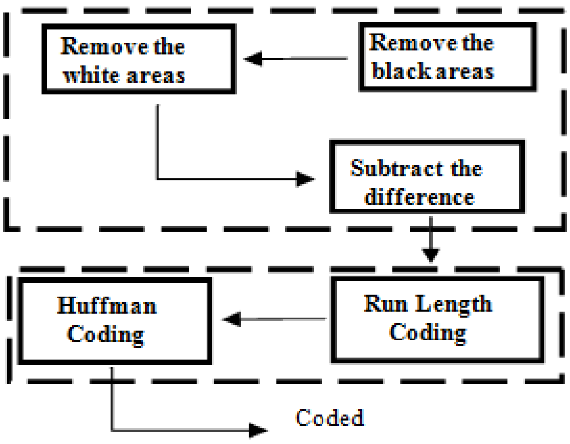

Fig 2. The framework of the approach [2]

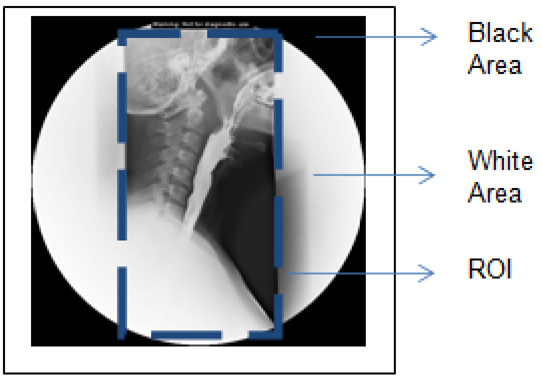

Fig 3. Sample of the Fluoroscopy image approach [2]

B. ***RLE for JPEG Image Compression in Space Research Program***

RLE scheme encodes image data in case of JPEG compression as a pair of (RUN, LEVEL) where RUNS shows the number of consecutive LEVELS [3]. Consecutive zeros (RUN) are only counted by this method and the non-zero coefficient (LEVEL) are appended.

But error rate is increased in this original RLE encoding technique. To overcome this problem optimized RLE technique was put forwarded by Muhammad Bilal Akhtar, Adil Masoud Qureshi and Qamar-ul-Islam [4] in 2011 where single zero between non-zero characters is represented by a pair of (RUN, LEVELS) if and only if a pattern of consecutive zeros occur at the input of the encoder. The non-zero digits are encoded as their respective values in LEVELS parameter. They have applied the run-length encoding (RLE) data compression algorithm to the 8×8 Image block and following results they have got:

[(0,-7) (0,5) (0,-1) (0,4) (0,-3) (0,-4) (0,2) (0,-6) (0,-2) (0,4) (1,1) (2,3) (4,-5) (3,2) (8,5) (0,0)].

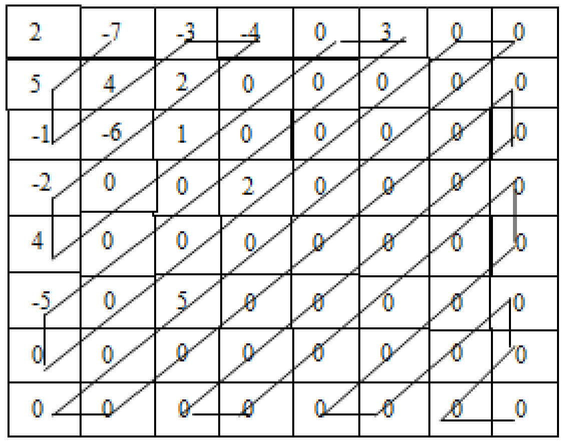

Fig.4: 8×8 Image block [4]

This RLE coding technique increases the error rate. To improve the conventional Run Length Coding In 2011 [4] Optimized RLE has been developed. In this method a FLAG bit is used to find the repeating zeros which appears in the input data.

at the input of the encoder and the non-zero digits are encoded as their respective values in LEVELS parameter. So we output of this method would be:

[-7 5 -1 4 -3 -4 2 -6 -2 4 (1, 0) 1 (2, 0) 3 (4, 0) -5 (3, 0) 2 (8, 0) 5 (0, 0)]

In 2013 Amritpal Singh and V.P. Singh [3] improved this RLE scheme by eliminating unwanted RUN, LEVEL means if we get (1, 0) pair in output then single '0' is encoded. (1, 0) pair in output then single zero present between the non-zero characters. Therefore we get the following output:

[-7 5 -1 4 -3 -4 2 -6 -2 4 0 1 (2, 0) 3 (4, 0) -5 (3, 0) 2 (8, 0) 5 (0, 0)]

C. Modified RLE in Wireless Sensor Networks

A sensor network is formed by a big number of sensor nodes where reduction of transmitted bit plays a significant role in wireless transmission method. Data compression scheme in particular to wireless transmission may be primarily distinguished as distributed data compression and local data compression [5]. High spatial correlation data in dense networks are used is distributed data compression. In this scheme, some information gets damaged at source during energy conservation. Temporal correlation between sensor data and compressed data based on some compression techniques such as RLE, Lossless Entropy compression, Huffman compression scheme is accomplished in local data compression. Modified bit-level RLE algorithm in wireless sensor networks is described in [5].

This algorithm can compressed each measure m_i obtained from a sensor node with a small dictionary whose size is determined by the Analog-to-Digital converter (ADC) to a binary representation r_i on *R* bits, where *R* is the resolution of the ADC [6]. For example- since this model is applied in environmental monitoring system, we can take 7 days humidity rate in binary. For each new measure m_i the compression algorithm computes the difference between resolutions ($\mathbf{d_i = r_i - r_{i-1}}$) where i > 0 and perform following conditions:

a. If i=0, then $h_i = r_0$ means first record.
b. If i>0 and $r_i > r_{i-1}$ then $h_i = 2 \times \mathbf{d_i}$.
c. Otherwise, $h_i = -2 \times \mathbf{d_i} + 1$

But this method is not suited because of short size of data on an average. To overcome this problem and to improve compression ratio rearrange the bit stream [6]. In bit stream rearrangement h_0 will not be altered because the value of d_0 and r_0 same so it will be kept reserved and remaining bits are encoded by the RLE compression technique as a pair of (value, run). To get the result of bit stream rearrangement transposes the figure: matrix. After getting compressed value subtracts 1 from every run length because the minimum run-length in original RLE is 1 and 0 is unused.

D. ***Compression and Decompression of FPGA Bit Stream***

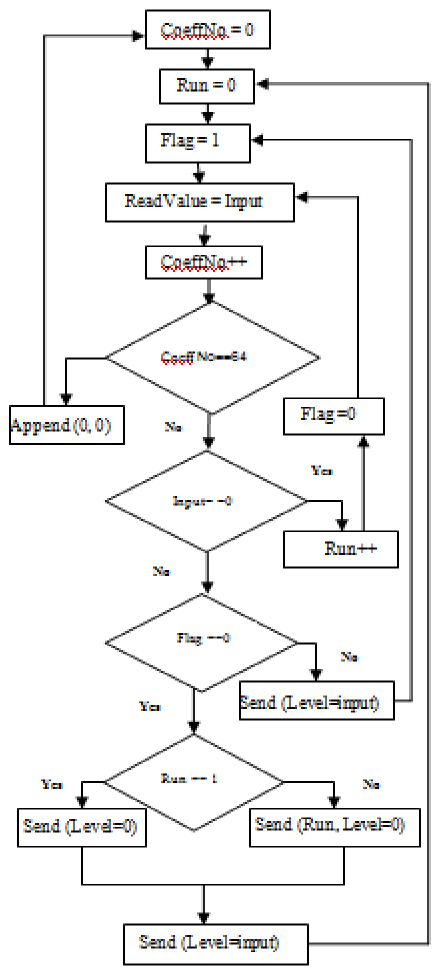

Fig.5 Flow diagram of RLE technique in [3]

Bit stream compression is performed by fixed length coding that contains consecutively repeated. In the fraternity of encoding methods, RLE based compression and bitmask based compression are more appropriate for bit stream compression because of their ability to encode such pattern repeated of bits. The technique depicted in [10] consists of four steps: bitmask selection, dictionary selection, RLE compression and decode aware placement. Fig: 6 shows decode aware bit stream.

Decode-aware placement algorithm is placed the compressed bit-stream in the memory for efficient decompression. During run-time, the compressed bit-stream is transmitted from the memory to the decompression engine, and the original configuration bit-stream is produced by decompression.

Bitmask selection requires three properties [9, 10]: compression using dictionary, compression using bitmask and without compression. Fixed bitmask is applied on fixed position symbol in bitmask based compression and sliding bitmask can be applied at any position. But the problem is fixed bitmasks encoded comparatively less than the sliding bitmask because fixed position and sliding bitmask consume more bits to encode. So both fixed and sliding bitmask are used for bitmask selection.

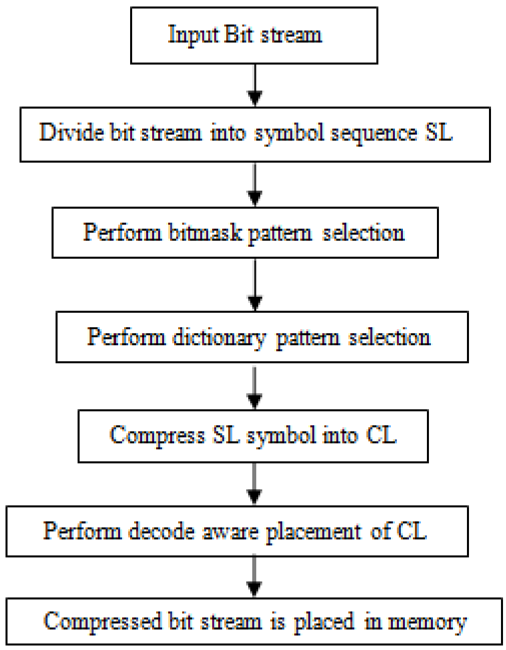

Fig.6 Flow diagram of decode aware bit stream compression [10]

Size of dictionary may be reduced through designing dictionary selection. In [9] a node adjacent to the most profitable node is eliminated if the profit is less than the certain threshold. Improper selection of threshold may result in incorrect removal of high frequency symbols. Since in bit stream the dictionary size is negligible, hence it is not suitable to reduce the dictionary size at the cost of the compression ratio.

The RLE compression method in[9] may yield a better compression ratio. Extra bits are not needed for this scheme and bitmask '0'is never used because it indicates an extra match and would have encoded using zero bitmasks. Repetitions of bits can be encoded using the mask pattern '00' a special marker of a RLE compression [8].For example: if the input contains word "00000000" which is repeated six times. The first occurrence "00000000" is encoded normally and the remaining five repeated symbols are encoded by RLE compression. The times of repetitions is encoded by the combination of bitmask address "0000" and the dictionary index "0100".

The decompression engine consists of a hardware component which is used to decode the compressed configuration bit-stream. A decompression engine is having two units: a buffering circuitry and circuitry for aligning codes fetched from the memory, while decoders perform decompression operation to generate original bit stream.

E.Binary image compression using run length encoding and multiple scanning techniques [16]

Binary image compression may be done using RLE compression technique. The corresponding integer values are stored in a 1D array or 2D array after compression [11]. Horizontal correlation between pixels on the same scan line is coded through1D run length coding. Every scan line denotes a sequence of alternating independent black and white runs. By considering the representation of first run value in each scan line as white color and set the white color value equal to zero if the first color is black. At the end each scan line has a special codeword.

But 1D RLC is used only for horizontal pixels. Horizontal and vertical correlation between pixels are achieved by applying 2D RLC and higher coding efficiency is derived This tag on three factors to encode an image- two lines (first line is reference line which has been coded and another line is the coding line which is being coded), five changing group (a_0, a_1, a_2, b_1, b_2) and three coding mode (pass mode horizontal mode and vertical mode).

The five changing group's are-

1. Assumed reference pixel a_0 is the white pixel which can moves both horizontal and vertical direction.
2. Coding line a_1 is the black pixel which can moves right to a_0.
3. a_2 contains same color as a_0
4. b_1 is the reference line contains same color as
5. b_2 is the changing pixel which has same color as

If the is b_2 at left of a_1 then pass mode is used means run in reference line starts from

If the distance between a_1 and b_1 is larger than three pixels then horizontal coding technique is used else vertical coding is used.

3. PERFORMANCE ANALYSIS OF DIFFERENT TECHNIQUES

In this section a performance analysis of RLE compression techniques presented in different scheme is done. The efficiency of RLE compression in hybrid scheme is measured using compression ratio. It is defined as the ratio between the compressed image or data size and the original data or image size. A smaller compression ratio implies a better compression technique. Graphs are plotted to test compression ratio for multiple scheme as shown below

JPEG Image Compression

The result shows an efficient compression using the new Run Length Coding scheme in JPEG compression which can applied on all type of images

Table1. Result of JPEG image compression using RLE

Algorithm	No. of bits before compression	No. of bits after compression	Compression Ratio
Simple RLE	64	32	50.00%
Optimized RLE	64	27	42.19%
RLE in [3]	64	26	40.63%

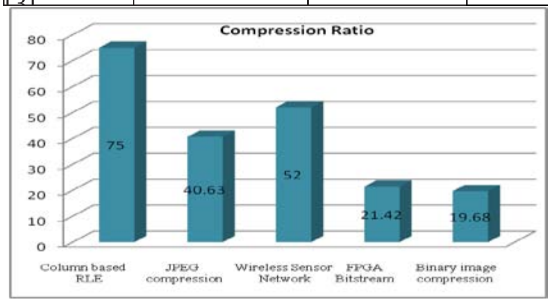

Fig 7: Compression ratio of different RLE algorithm in hybrid scheme

4. CONCLUSION

In this paper, we have studied various RLE compression algorithm particularly applied to the reduce storage and computational resources in combined encoding schemes to achieve system performance, efficiency, compression ratio. A snapshot of the different RLE based recent combined encoding schemes along with their merits and shortcomings are shown in Table 2. The study reveals that combination of the different RLE compression algorithms has made the compression more secure and the also made it competent to be employed in any of the transmission schemes. This study may render a significant role keeping in view the ever increasing need for low bit rate compression methods and efficacy of lossless compression methods; scope exists for new RLE methods as well as evolving more efficient algorithms in the existing methods.

Table 2

Method / Author	**Key features**	**Advantage**	**Shortcomings**
D. Sharma et. Al[11]	• Multi code compression scheme • It is based on different coding techniques like Golomb, Frequency-Directed Run-length, alternating Run Length, Extended Frequency-Directed run length ,IFDR and Huffman coding	Significant improvement in compression ratio by applying RL- Huffman coding	Increased Computational cost

J.G Abhraham et al[12]	• RLE coding deals with the compression of repeated sample values. • The modified MTF (Move to front) coding facilitates the RLE coding of sample values that differ by a constant difference. • The pattern detection algorithm detects the number patterns, that repeats consecutively in the input sample set.	When RLE, MTF and Pattern RLE is applied jointly, compression ratio becomes 4.15 and 2.75 in case of without noise and with noise respectively	Compression efficiency needs improvement in case of noisy channel
R. Sahoo et. Al.[13]	• Method 1 uses DCT, Quantization, Entropy Coding(Zigzag Scan and Conventional Run Length Coding) • Decompression is done using IDCT (Inverse Discrete Cosine Transform) • Method 2 uses Haar wavelet, Hard Thresholding, Entropy Coding (Zigzag Scan and Conventional Run Length Coding)	• Robustness against transmission error • Low bit rate transmission • Compression ratio improved significantly in method 1	Method 1 cannot improve the quality of image due to low PSNR
S. Mirthulla et. Al[14]	• Three multistage compression techniques are introduced to reduce the test data volume • The three encoding schemes namely equal run-length coding (ERLC), extended frequency directed run length (EFDR) coding, alternating variable run length (AVR) is used for computing the data. • These encoding scheme together with nine coded (9C) technique enhance the test compression ratio**.**	• By using the nine coded compression (9C) coding test power is reduced • 9C-AVR provides better compression ratio than the other compression coding techniques.	Increased computational time due to processing in multiple stages
Hend. A. Elsayed et. Al [15]	• Lossless audio coding is performed by applying Burrows-Wheeler Transform (BWT) and a combination of a Move-To-Front coding (MTF) and Run Length Encoding (RLE). • The BWT is applied to these floating point values of audio signals to get the transformed coefficients and then these resulting coefficients are converted using the MTF coding. • Further compression is performed by using a combination of the run length encoding and entropy coding	BWT and the combined MTF coding and RLE method performs better than the other lossless audio coding using BWT method, the combined BWT and MTF coding method, and combined BWT and RLE coding method	No any comparison was shown with respect to conventional compression techniques

REFERENCES

1. Mingyuan An, "Column-Based RLE in Row Oriented Database" in 2009.
2. Arif Sameh Arif, Sarina Mansor, Hezrul Abdul Karim, Rajasvaran Logeswaran "Lossless Compression of Fluoroscopy Medical Images using Correlation and the Combination of Run-length and Huffman Coding" In Proc. of IEEE EMBS International Conference on Biomedical Engineering and Sciences | Langkawi | 17th - 19th December 2012.
3. Amritpal Singh and V.P. Singh "An Enhanced Run Length Coding for JPEG Image Compression" International Journal of Computer Applications, Volume 72– No.20, June 2013.
4. Muhammad Bilal Akhtar, Adil Masoud Qureshi2and Qamar-ul-Islam "Optimized Run Length Coding for JPEG Image Compression Used in Space Research Program of IST". In proc. of International Conference on Computer Networks and Information Technology (ICCNIT), 2011
5. S. Kalaivani and C. Tharini, "Efficient data compression technique using modified adaptive Rice Golumb coding for wireless sensor networks" In ARPN Journal of Engineering and Applied Sciences Vol. 10, No. 12, July 2015 pp 5395-5401
6. Shengchun Long and Pengyuan Xiang "Lossless Data Compression for Wireless Sensor Networks Based on Modified Bit-level RLE", In Proc. of 8th International Conference on Wireless Communications, Networking and Mobile Computing, 2012, pp1-4
7. Saumya.Sadanandan, V. K. Govindan, "Modified Skip Line Encoding for Binary Image Compression" In Proc. of Int. Conf. on Advances in Information Technology and Mobile Communication 2013, pp 44-47.
8. P.M. Sandeep and C.S Manikandababu, "FPGA Bit-stream compression Using Run-length Encoding", International Journal of Electronics Communication and Computer Technology (IJECCT) Volume 3 Issue 2 (March 2013) pp 386-389.
9. D. Koch, C. Beckhoff, and J. Teich, "Bitstream decompression for high speed FPGA configuration from slow memories," in Proc. of International Conference on Field-Programmable Technology, ICFPT 2007 pp. 161–168.
10 P .Hemanth and V. Prabhu, "Compression of FPGA bit streams using improved RLE algorithm", In proc. of International Conference on Information Communication and Embedded Systems (ICICES), 2013 .

11. Deepika Sharma, Debbrat Ghosh, Harpreet Vohra, Test Data Volume Minimization using Double Hamming Distance Reordering with Mixed RL Huffman based compression scheme for System-on-chip, In proc. of Nirma University International Conference on Engineering, IEEE NUiCONE-2012, December, 2012
12. Jijo George Abraham, Rahul Mishra and Deepa.J, A Lossless Compression Algorithm for Vibration Data of Space Systems, in proc. of International Conference on Next Generation Intelligent Systems (ICNGIS) 2016
13. Rashmita Sahoo, Sangita Roy, Sheli Sinha Chaudhuri, Haar Wavelet Transform Image Compression Using Run Length Encoding, in proc of International Conference on Communication and Signal Processing, April 3-5, 2014, India
14. S.Mirthulla and A. Arulmurugan, Improvement of Test Data Compression using Combined Encoding, in proc. of IEEE sponsored 2nd International Conference on Electronics and communication, ICECS 2015
15. Hend. A. Elsayed, Burrows-Wheeler Transform and Combination of Move-to-Front Coding and Run Length Encoding for Lossless Audio Coding, In proc. of 9th International Conference on Computer Engineering & Systems (ICCES), 2014
16. Frank J. Merkl, “Binary image compression using run length encoding and multiple scanning techniques”, in 1988

