IJCSE International Journal of Computer Sciences and Engineering Open Access
Survey Paper Vol.-7, Special Issue-8, Apr 2019 E-ISSN: 2347-2693

# Survey on Image Compression Techniques

**C. Jenifer Kamalin[1*], G. Muthu Lakshmi[2]**

[1,2] Dept. of Computer Science and Engineering, ,Manonmaniam Sundaranar University, Abishekapatti, Tirunelveli-12, Tamilnadu, India.

*Corresponding Author: manisha.jeni@gmail.com*

DOI: https://doi.org/10.26438/ijcse/v7si8.129132 | Available online at: www.ijcseonline.org

***Abstract***— The goal of this paper is to survey about various image compression techniques. Image compression is a very important to implement in digital image processing. It is the case of converting an image file in such a way that it consumes smaller space than the original file. Compression technique that reduces the size of an image file without disturbing or degrading its quality to a greater extent. The main intention of the compression is to minimize the amount or unnecessary data.The goal behind is to save the amount of memory required to save the images or to make practical network bandwidth in efficient manner.

***Keywords***— Image Compression, Run length encoding, Entropy encoding, LZW.

## I. Introduction

### *A Image Compression*

Image compression is minimizing the size in bytes of a graphics file without corrupting the quality of the image to an unacceptable level[1]. The reduction in file size grants more images to be stored in a given amount of disk or memory space. It also decreases the time required for images to be sent over the Internet or downloaded from Web pages. There are lot of different ways in which image files can be compressed. For Internet use, the two suggestive compressed graphic image formats are the JPEG format and the GIF format. The JPEG method is regularly used for photographs, while the GIF method is commonly used for line art and other images in which geometric shapes are relatively simple.These methods have not extended widespread acceptance for use on the Internet as of this writing. However, both methods offer agreement because they offer higher compression ratios than the JPEG or GIF methods for some types of images. Another new method that may in time swap the GIF format is the PNG format.

A text file can be compressed without the errors, but only up to a some extent. This is called lossless compression. Beyond this point, errors are introduced. In text and program files, it is vital that compression be lossless because a single error can actively damage the meaning of a text file, or cause a program not to run. In image compression, a small loss in quality is usually not notable.

### *B Lossless and Lossy Compression*

Lossless and lossy compression are terms that describe whether or not, in the compression of a file, all original data can be rescued when the file is uncompressed[1]. With lossless compression, each and every bit of data that was originally in the file remains after the file is uncompressed. All the information are completely built up. The Graphics Interchange File (GIF) is an image format used on the Web that supports lossless compression. Another side, lossy compression reduces a file by permanently erasing certain information, especially redundant information. When the file is uncompressed, only a part of the original information is deathly there. Lossy compression is generally used for video and sound, where a certain amount of information loss will not be discovered by most users.

### *C Lossy Compression*

Lossy compression depends on patterns being present within the information[1]. The algorithm looks at the data, tries to identify the patterns and decides how much it can throw away without noticeably affecting the quality of the data.Different file types can be compressed to different amounts. Some information can be safely discarded, while other information may be essential to retain.For example a document needs to store every character. Imagine the word *fury* and *furry* in a sentence, they mean something completely different - yet they are only one character different - you cannot afford to lose that single character.On the other hand a music file will have sounds at frequencies that the human ear cannot hear. Those frequencies can be safely discarded without the listener ever noticing their absence.

**Lossy Advantages and Disadvantages**

*Advantages* - more tools, plugins, and small file sizes.

*Disadvantages*-Quality disgrace with higher ratio of compression.

*D Lossless Compression*

Lossless compression refers to which the **image is reduced without any quality loss.** compression technique that decompresses data back to its original without any quality loss[1]. The decompressed file and the original are same. Usually this is done by removing unnecessary data from JPEG and PNG files.

Lossless compression is used in cases where it is important that the original and the decompressed data be identical, or where deviations from the original data could be deleterious[4]. Typical examples are executable programs, text documents, and source code. Some image file formats, like PNG or GIF, use only lossless compression, while others like TIFF and MNG may use either lossless or lossy methods

**Lossless Advantages and Disadvantages**

*Advantages* - No quality loss, little decreases in image file sizes.

*Disadvantages* **-** More subtle.

## II . METHODS FOR LOSSLESS IMAGE COMPRESSION

- *Run-length encoding*
- *DPCM and Predictive Coding*
- *Entropy encoding.*
- *Adaptive dictionary algorithms such as LZW – used in GIF and TIFF.*

*A Run Length Encoding*

Usually a data files will be containing the same characters as repeatedly in a row. For example, text files use spaces to separate lines,paragraphs,tables&charts etc.,Digital signals can also have runs of the same value,that means the signal is not changing.for instance Sky at the nighttime would contain fully black .Run length encoding is a simplest method of compressing these types of files.

A good example of a run-length scheme is PackBits, created for Macintosh users. Every byte from the input file is recovered by nine bits in the compressed file. The added ninth bit is clarified as the sign of the number. That is, each character will be read from the input file that is between 0 to 255, while each character written to the encoded file that is between -255 and 255. To understand how this is used,

Consider, the input file: 1, 2, 3,4,2,2,2,2,2,4,

Compressed file generated by the PackBits algorithm: 1, 2, 3, 4, 2,-3,4.

The compression program simply assign each number from the input file to the compressed file, with the exception of the run: 2,2,2,2,2. This can be written in the compressed file by the two numbers: 2,-3. The first number ("2") shows what character the run consists of. The second number ("-3") indicates the number of time of characters in the run, found by taking the complete value and adding one.

Run length encoding on an image, transmission of digital line scan is replaced by transmission of a quantity count of each of successive run of block or white scanned picture element.

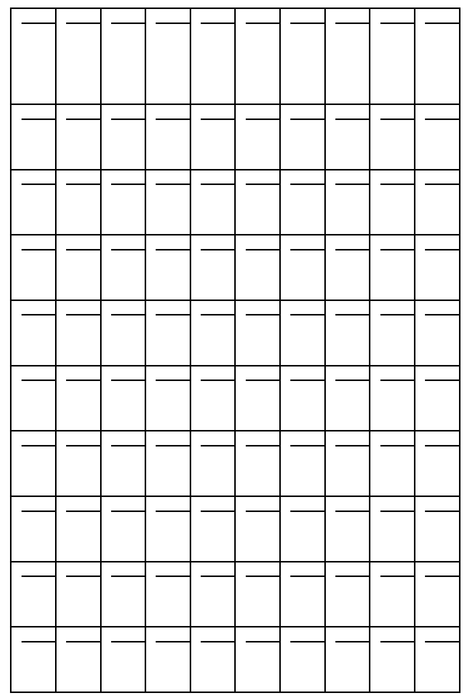
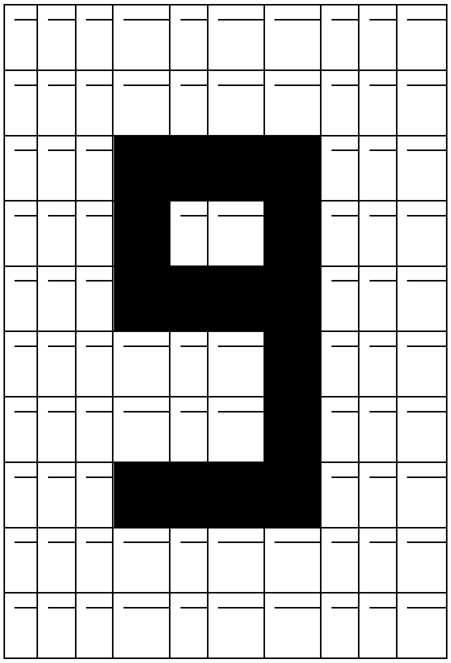


**Fig1. Run length code for a 100 pixel image**



As in Fig 1. 23W 4B 6W 1B 2W 1B 6W 4B 9W 1B 9W 1B 6W 4B 23W

Or

23 4 6 1 2 1 6 4 9 1 9 1 6 4 23

Length: 15 Characters=120 bits

*B DPCM (Differential pulse code modulation):*

Differential pulse code modulation is a technique that converts analog to digital signal. This technique appraises the analog signal and then quantizes the difference between the sampled value and its predicted value, then encodes the signal to form a digital value. Coding a difference is based on the fact that most source signals show significant correlation between successive samples so encoding uses redundancy in sample values which implies lower bit rate.

The samples of a signal are highly equated with each other. The signal's value from the present sample to next sample does not differ much. The neighboring samples of the signal carry the same information with a little difference. When these samples are encoded by the standard PCM system, the resulting encoded signal contains some repeated information bits.

*C Entropy Encoding*

One of the main types of entropy coding builds and assigns a unique prefix-free code to each unique symbol that occurs in the input. These entropy encoders then compress data by changing each fixed-length input symbol with the corresponding variable-length prefix-free output codeword[6]. The length of each codeword is almost proportional to the negative logarithm of the probability. That means, the most common symbols use the shortest codes. As attested by Shannon's source coding theorem, the optimal code length for a symbol is $-\log_b P$, where $b$ is the number of symbols used to make output codes and $P$ is the probability of the input symbol. Entropy is a lower bound on the average number of bits required to represent the symbols (the data compression limit).

The typical entropy encoding techniques are Huffman coding and Arithmetic coding. If the approximate entropy characteristics of a data stream are known in advance it especially for signal compression,a simpler static code may be useful. These static codes embrase universal codes such as Elias gamma coding or Fibonacci coding and Golomb codes such as unary coding or Rice coding.

*D LZW*

Lzw Coding And Decoding Lempel, Ziv and Welch (LZW) compression name is originate from the scientists Abraham Lempel, Jakob Ziv and Terry Welch. LZW compression algorithm is lossless and dictionary based compression algorithm[2]. Dictionary based algorithms scan a file and search the sequences of data or string that appear more than once in a file. LZW compression works by substitution strings of characters with single codes without doing any study of the incoming text data. It adds every new found characters of string in the dictionary and data compression present on the single code. These strings are then stored in a dictionary and the compressed file with references are put wherever redundant data occurred. The replaced code can be of any arbitrary length, but it must have more bits in it than a single character [3]. The first 256 codes when using eight bit characters are initially assigned to the standard character set and the resting codes are assigned to strings as the algorithm proceeds. The sample program runs as shown with 12 bit codes. LZW is an adaptive technique. As the compression algorithm runs, a changing dictionary of the strings that have came in the text so far is maintained. Because the dictionary is pre- loaded with the 256 different codes that may appear in a byte, it is guaranteed that the whole input source may be converted into a series of dictionary indexes. If "A" and "B" are two strings that are held in the dictionary, the character sequence "AB" is changed into the index of "A" followed by the index of "B". "A" greedy string matching algorithm is used for scanning the input, so if the first character of "B" is "x", then "Ax" cannot be an element of the dictionary. This means codes 0-255 refer to individual bytes, while codes 256-4095 refers to substrings. Thus, there are Advantages and disadvantages of LZW compression; the size of files usually increases to a great extent when it includes lots of repetitive data or monochrome images[2]. LZW compression is the better technique for reducing the size of files containing more repetitive data.LZW compression is fast and simple to apply.Since this is a lossless compression technique none of the contents in the file are lost during or after compression.The decompression algorithm always follows the compression algorithm[3]. LZW algorithm is efficient because it don't necessary to pass the string table to the decompression code. The table can be recreated as it was during compression, using the input stream as data. This avoids insertion of large string translation table with compression data.

*LZW Encoding*

LZW encoding is working based on the occurrence multiplicity of bit sequences in the pixel to be encoded. Its principle consists in substituting patterns with an input image, by progressively building a dictionary.

1. Initial table with initial character strings
2. P=first input character
3. WHILE not end of input stream
4. C=next input character
5. IF P+C is in the string table
6. P=P+C
7. ELSE
8. output the code for P
9. add P+C to the string table
10. P=C
11. END WHILE
12. output code for P

*LZW Decoding*

In decoding process, the algorithm rebuilds the dictionary in the opposite direction; it thus does not need to be stored.

1. Initialize table with single character strings
2. OLD = first input code
3. output translation of OLD
4. WHILE not end of input stream
5. NEW = next input code
6. IF NEW is not in the string table
7. S = translation of OLD
8. S = S+C
9. ELSE
10. S = translation of NEW
11. output S
12. C = first character of S
13. OLD + C to the string table
14. OLD = NEW
15. END WHILE

## III. CONCLUSION

In this paper lossless image compression techniques Run-length encoding ,DPCM ,Entropy encoding, LZW are discussed. Image compression is categorized by lossy compression and Lossless compression. RLE is best for compressing any type of data regardless of its information content, but the content of the data will disturb the compression ratio achieved by RLE. Although most RLE algorithms cannot reach the high compression ratios of the more advanced compression methods. The DPCM conducted on signals with the interaction between successive samples leads to good compression ratios. The LZW algorithm could reduce the compression ratio and boost the compression efficiency.

## REFERENCES


[1]. Yogesh Chandra, Shikha Mishra "Review Paper of Image Compression", International Journal of Engineering and Applied Sciences (IJEAS) ISSN: 2394-3661, Volume-2, Issue-2, February 2015.

[2]. Dalvir Kaur, Kamaljit Kaur, "Huffman Based LZW Lossless Image Compression Using Retinex Algorithm", International Journal of Advanced Research in Computer and Communication Engineering, Vol. 2, Issue 8, August 2013.

[3]. Mrs.Bhumika Gupta, "Study Of Various Lossless Image Compression Technique", IJETTCS, volume 2, issue 4, July- August 2013.

[4]. David Jeff Jackson & Sidney Joel Hannah, "Comparative Analysis of image CompressionTechniques," System Theory 1993, Proceedings SSST '93, 25th Southeastern Symposium,pp 513-517, 7 –9March 1993.

[5]. Malwinder Kaur, Navdeep Kaur, "A Litreature Survey on Lossless Image Compression", International Journal of Advanced Research in Computer and Communication Engineering, Vol. 4, Issue 3, March 2015.

[6]. Mridul Kumar Mathur, Seema Loonker, Dr. Dheeraj Saxena, "Lossless Huffman Coding Technique For Image Compression And Reconstruction Using Binary Trees", IJCTA, Vol 3 , Jan-Feb 2012.

[7]. Jagadeesh B, Ankitha Rao, "An approach for Image Compression Using Adaptive Huffman Coding", International Journal of Engineering Research & Technology (IJERT), Vol. 2 Issue 12, December – 2013.

[8]. Xiaoqiang Ma ,Chengdong Xu ,Pengfei Zhang ,Chunsheng Hu, "The Application of the Improved LZW Algorithm in the Data Processing of GNSS Simulation", Fourth International Conference on Computational and Information Sciences 2012.

[9]. JayavrindaVrindavanam ,SaravananChandran, Gautam K. Mahanti, "A Survey of Image Compression Methods" International Journal of Computer Applications® (IJCA) 2012.


## AUTHOR PROFILE

C.Jenifer kamalin is currently pursuing her research in the Department of Computer Science and Engineering, Manonmaniam Sundaranar University, Tirunelveli. She completed her M.Tech in Information Technology from kalasalingam University and B.Tech in Information Technologyfrom Scad College of Engineering and Technology. Her Research Interests include Image compression.

Dr.G.Muthu Lakshmi is currently working as a Assistant Professor in the Department of Computer Science andEngineering, Manonmaniam Sundaranar University, Tirunelveli. She received her B.E degree in Department of Computer Science and Engineering, PSR Engineering College, Viruthunagar and M.E degree in Computer Science and Engineering from Manonmaniam Sundaranar University and She received her Doctorate in Computer Science and Engineering from Manonmaniam Sundaranar University. She has published papers in many National and International level Journals and Conferences. Her Research Interests are in the field of Digital Image Processing, Pattern Recognition.