

Different Compression Techniques Used for Different Applications

Alka Chauhan*

Assistant Professor, Department of Computer Science, D. M. College, Moga

Abstract – Data compression is a technique use to compress data. It is a process which reduce the size of data by reducing redundant data from it. Hence choosing the best machine learning algorithm is really important. In addition to different compression technologies and methodologies, selection of a good data compression tool is most important. There is a complete range of different data compression techniques available such that it becomes really difficult to choose which technique serves the best. Here comes the necessity of choosing the right method for text compression purposes and hence an algorithm that can reveal the best tool among the given ones. In this paper we represent different algorithms for different applications to compress the text data.

-----X-----

1. INTRODUCTION

Data compression is the process of modifying, encoding or converting the bits structure of data in such a way that it consumes less space on disk. It is a process by which a file can be compressed, such that the original file may be fully recovered without any loss of actual information especially over the internet since they can upload and download much faster. Data compression is a method of encoding rules that allows substantial reduction in the total number of bits to store or transmit a file. Data Compression is used because most of the real world data is very redundant. Data Compression is basically defined as a technique that reduces the size of data by applying different methods that can either be Lossy or Lossless (Rastogi & Segar, 2014). The algorithm which loss some part of data is called Lossy data compression. And the algorithm that achieve the same what we compressed after decompression is called lossless data compression.

2. TEXONOMY OF LOSSLESS DATA COMPRESSION

In this compression technique, no data is lost. The compressed file and decompressed file is almost same without any loss of data. In this type of compression generally the encrypted file is used for storing or transmitting data. For general purpose use we need to decrypt the file. These types of compression are also known as noiseless as they never add noise to signal or image. It is also termed as the entropy coding as it uses the techniques of decomposition/statistics to remove/reduce the redundancy. It is also used only for the some specific applications along with the rigid needs like a medical-

imaging. Below mentioned techniques consists in the lossless compression:

1. Huffman encoding
2. Run length encoding
3. Arithmetic coding
4. Dictionary Techniques.
 - a) LZ77
 - b) LZ78
 - c) LZW
5. Bit Plane coding

Here we are going to represent different applications where we can use these algorithms (Bhattacharjee, et. al., 2013).

- File Systems

OpenZFS (LZ4)

- Web

all major web browsers, Apache HTTP Server (all Brotli)

- Databases/Data Warehouses

MySQL, Apache HBase (both LZ4), Amazon Redshift (Zstandard)

- Big Data Services

Apache Hadoop (LZ4, Zstandard), Apache Spark (LZ4), Presto (Zstandard)

- Processing Pipelines

Facebook, "The Guardian" publication pipeline (both Zstandard), Dropbox static assets (Brotli)

3. TEXNOMY OF LOSSY DATA COMPRESSION

Lossy Compression is generally used for image, audio, video. In this compression technique, the compression process ignores some less important data and the exact replica of the original file can't be retrieved from the compressed file. To decompress the compressed data we can get a closer approximation of the original file (Bhattacharjee, et. al., 2013). In these methods few loss of the information is acceptable. Lossy compression techniques are used for pictures and music files that can be trimmed at the edges.

The examples of frequent use of Lossy data compression are on the Internet and especially in the streaming media and telephony applications. Some examples of lossy data compression algorithms are JPEG, MPEG, MP3. Most of the lossy data compression techniques suffer from generation loss which means decreasing the quality of text because of repeatedly compressing and decompressing the file. Lossy image compression can be used in digital cameras to increase storage capacities with minimal degradation of picture quality

A. Vector and Scalar Quantization

B. Video Compression: MPEG Compression Algorithm (Paper, 2008)

4. TODAY'S DATA COMPRESSION:

AREAS

Tailoring the different compression methods to the many rapidly changing environments in which compression is of potential value has led to applications of data compression to such diverse areas as:

- * satellite imagery
- * mini discs
- * MP3 technology
- * fax
- * digital cameras

- * DVD technology

- * modems

- * wireless telephony

- * database design

- * storage and transmission of CT and MRI scans

- * mammography

- * digital images, high definition television (HDTV), and video games.

CONCLUSION

Today, with growing amount of data storage and information transmission, data compression techniques have a significant role. Even with the advances in bandwidth and storage capabilities, if data were not compressed, many applications would be too costly and the users could not use them. In this research survey, I attempted to introduce two types of compression, lossless and lossy compression, algorithms and discussed their different applications. Here I discussed some major everyday applications regarding data compression; JPEG as an example for image compression and MPEG as an example of video compression in our everyday life. At the end of this survey I discussed major areas leveraging data compression algorithms.

REFERENCES

1. K. Rastogi, K. Segar (2014). "Analysis and Performance Comparison of Lossless Compression Techniques for Text Data" International Journal of Engineering Technology and Computer Research (IJETCR) 2 (1), pp. 16-19
2. https://www.researchgate.net/publication/318020143_Lossless_Data_Compression_-_Modern_Scope_and_Applications
3. A. K. Bhattacharjee, T. Bej and S. Agarwal (2013). "Comparison Study of Lossless Data Compression Algorithms for Text Data", IOSR Journal of Computer Engineering (IOSR-JCE), Volume 11, Issue 6 (May. - Jun. 2013), PP 15-19, e-ISSN: 2278-0661, p- ISSN: 2278-8727, www.iosrjournals.org
4. A. C. W. Paper (2008). "An explanation of video compression techniques."

5. <http://www.ams.org/publicoutreach/feature-column/fcarc-compression6>

Corresponding Author

Alka Chauhan*

Assistant Professor, Department of Computer
Science, D. M. College, Moga