Survey of Lossy Data Compression Algorithms

Author: Maulik Kumbhani, Ashish Gotekar,

Co-Author: Prof. Kaushal Gor

Department of MCA, Parul University, Vadodara, India

Abstract: Data compression is an area which are covered maximum and almost attention on digital data like text, image, video and etc. data compression was introduce by shanon, fano and Huffman with aim of reduce redundancy in stored and communication or transference of data on network. Also, they represent different technologies and methodologies and applied to the algorithms includes types lossless and lossy compression with Huffman coding, arithmetic coding, dictionary techniques and many more techniques to compress text and multi-media files

Keyword: Data Compression, Lossy Data Compression.

I. INTRODUCTION

Lossy compression is most generally used to pack sight and sound data (sound, video, and images), particularly in applications, for example, spilling media and web communication. On the other hand, lossless compression is ordinarily required for content and data documents, for example, bank records and content articles. It very well may be worthwhile to make an ace lossless document which would then be able to be utilized to deliver extra duplicates from. This permits one to abstain from basing new packed duplicates off of a lossy source record, which would yield extra antiquities and further superfluous data misfortune.

All around planned lossy compression innovation regularly diminishes record estimates fundamentally before debasement is seen by the end-client. In any event, when perceptible by the client, further data decrease might be alluring (e.g., for ongoing correspondence, to lessen transmission times, or to diminish capacity needs).

Lossy image compression is one of the mainstream strategies which are utilized for image compression because of the sudden increment of lumps of data, which is being utilized for Internet and different applications. We need to ensure that it is vital just to guarantee that the compression part isn't debasing the nature of the image in a Lossy Compression. The working standard of a Lossy Compression calculation is assessed, undoubtedly. PSNR values are utilized to decide Compression proportion and image quality.

Lossy compression is a method of data compression in which the size of the file is truncated by eliminating data in the file. In doing so, quality of image is reduced to decrease the size of image. Any data that the compression algorithm deems expendable is abstracted from the image, thereby truncating its size. Albeit the file doesn't have the same data as it did afore compression was utilized, often this won't be salient, albeit the resolution of the image has suffered.

Where lossy compression becomes noticeable when a file is modified and made copy multiple times. In comparing the two images below Figure, you will see that the image on the left is of higher quality. As the image on the left has been saved multiple times. Each time it was saved, lossy compression was used, and clarity of the image size has been reduced. Although the single time copy of an image is saved won't result in noticeable differences, the below Figure shows that multiple saves of an image using lossy compression will eventually degrade the image so that the decrease in quality becomes *recognizable*.
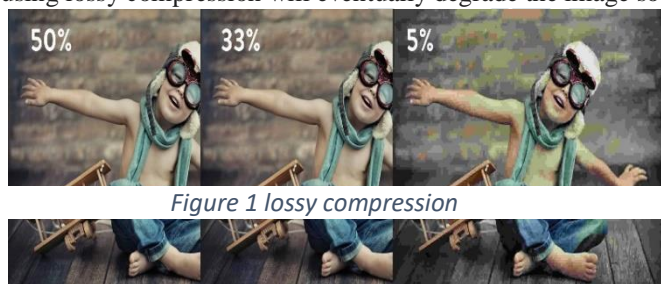

Figure 1 lossy compression

Saying that lossy compression sacrifices data doesn't mean it removes data arbitrarily. After all, it would be Addressable if something in a JPEG photo was suddenly missing an eye. Rather, lossy compression removes data from an image using one or more techniques:

- Chroma sub sampling
- Colour reduction
- Fractal compression
- Transform coding

II. APPLICATION AREAS

2.1 Satellite imagery: Satellite Data Compression covers recent progress in compression techniques for multispectral, hyperspectral and ultra-spectral data.

2.2 Mini discs: The difference is that a Minidisc uses a digital compression technique called ATRAC (Adaptive Transform Acoustic Coding) when storing music

2.3 MP3 technology: The MPEG audio standard is a high-complexity, high-compression, and high audio quality algorithm.

2.4 Modems: Data compression is the ability of the modem to take data in from the computer, reduce it in volume, and then send it out via the modem.

2.5 Digital image: Image compression is useful because it decreases the amount of memory required to store images digitally or communicate these images over a network such as the internet.

III. METHODOLOGY

Lossy compression methods have bigger compression proportions when contrasted with the lossless compression procedures. Lossy methods are utilized for most applications.

Lossy compression includes following methods:

1. Block truncation coding
2. Code Vector quantization
3. Fractal coding
4. Transform coding
5. Sub-band coding

IV. TECHNIQUES OF LOSSY COMPRESSION

Image compression might be lossy or lossless. Lossless compression is favoured for authentic purposes and regularly for clinical imaging, specialized drawings, cut craftsmanship, or funnies.

Lossy compression techniques

4.1 Transform coding:

Transform coding is a method of data compression for data like audio signals or photographic images. The transform is typically lossy, resulting in a lower quality duplicate of the original input. Knowledge of the application is used to choose data to discard, thereby lowering its bandwidth. The rest of the data can be compressed via different types of methods, the result may not be identical to the original input, but is predicted to be close enough for the purpose of the application when the output is decoded.

• Input $N \times N$ image is subdivided into sub images of size $n \times n$.

• $n \times n$ subimages are converted into transform arrays. This tends to decor relate pixel values and pack as much information as possible in the smallest number of coefficients. • Quantizer by selection eliminates or coarsely quantizes the coefficients with the least info.
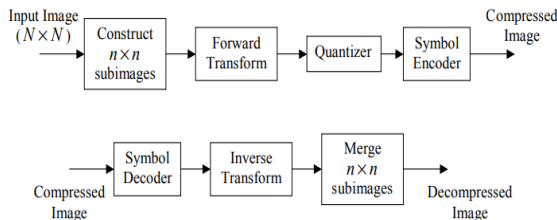
•Variable-length code to encode the quantized coefficients is used by Symbol encoder.

• Any of the above steps can be adapted to each sub-image, based on local image information, or fixed for all sub-images.

4.2 Discrete Cosine Transform:

A discrete expresses a finite sequence of data points. Discrete Cosine Transform is a lossy compression technique that is widely used in the area of image and audio compression. DCTs are used to change data in the addition of series of cosine waves oscillating at different frequencies.

4.3 Discrete Wavelet Transform (DWT):

The Discrete Wavelet Transform (DWT) is an implementation of the wavelet transform using a discrete set of wavelet scales and translations obeying some defined rules. In other words, this change decomposes the signal into a mutually orthogonal set of wavelets, which is the main difference from the continuous wavelet transform, or its implementation for the discrete-time sometimes called Discrete-time continuous wavelet Transform (DT-CWT).

Difference between some of the Lossy Compression Techniques

Factors	Lossy Compression Techniques		
	Transform coding	DCT	DWT
Advantages	It produces awesome image quality. Expanded encoding with versatile piece task (ATC), an encoding intricacy tantamount to that of completely versatile prescient coding (APC).	It is really valued. Better energy compaction. Coefficients are nearly correlated. Experimentally observed to work well.	It offers a concurrent area of time and frequencies. It very well may be utilized to break down a sign into part wavelet. Extremely little wavelet can be utilized to segregate fine subtleties in a sign, while enormous wavelet can distinguish coarse subtleties.
Computation	It is computationally intensive.	It is a fast-computational approach.	It is computationally very fast.
Application	JPEG, MPEG	JPEG, MPEG	JPEG, MPEG
Drawback	Transform matrix cannot be factored into sparse matrix. High computational complexity.	Truncation of higher otherworldly coefficients brings about obscuring of the images, particularly any place the subtleties are high. Coarse quantization of a portion of the low otherworldly coefficient presents graininess in the smooth bits of the images.	DWT is move touchy on the grounds that info signal movements produce erratic changes in DWT coefficients. It experiences poor directionality in light of the fact that DWT coefficients uncover just three spatial directions.

V. Tools & Technology

- Winzip
- 7-zip
- WinRAR
- PeaZip
- Zipware
- Hamster Zip Archiver
- Express Zip File Compression

VI. LITERATURE REVIEW

Paper 1: A Survey on Lossless and Lossy Data Compression Methods
Paper 2: An Improved Data Compression Method for General Data.
Paper 3: Image Compression Techniques: Lossy and Lossless
Paper 4: Image Compression Using VQ for Lossy Compression.
Paper 5: A Review of Lossless and Lossy Based Image Compression Technique