Lecture Six
Multimedia Systems

Image Compression

Part One

By: Eng. Haydara Mahmoud.

Damascus University – Faculty of Informatics - 2019

Data Compression Concept:

- Most of the real world data is very redundant. Data Compression is basically defined as a technique that decreases the size of data.
- It also reduces the storage space and hence storage cost.
- And reduces time to retrieve and transmit data.
- Compression Ratio = Uncompressed Size / Compressed Size.
- Data-compression techniques can be divided into two major families:
    - Lossy Data Compression.
    - Lossless. Data Compression.

Lossless Compression:

1. The ***Lossless compression*** technique means when the data is decompressed, the result is a bit-for-bit perfect match with the original one.
2. The name of lossless means no data is lost, the data is only saved more efficiently in its compressed state, but nothing of it is removed.

Original Image

Lossless Compression

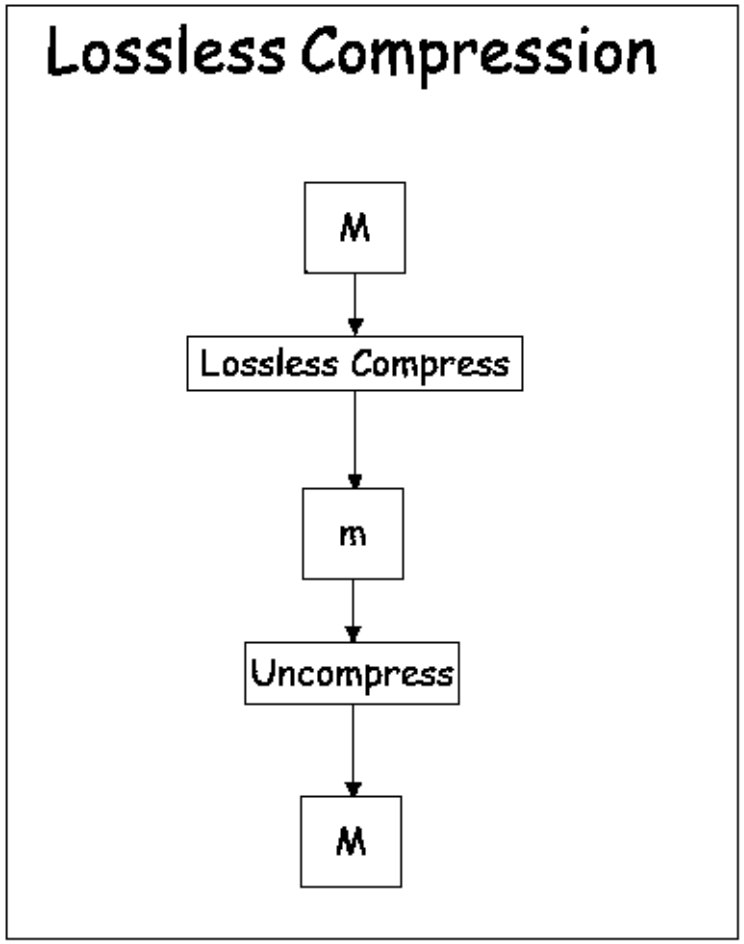
Lossy Compression:

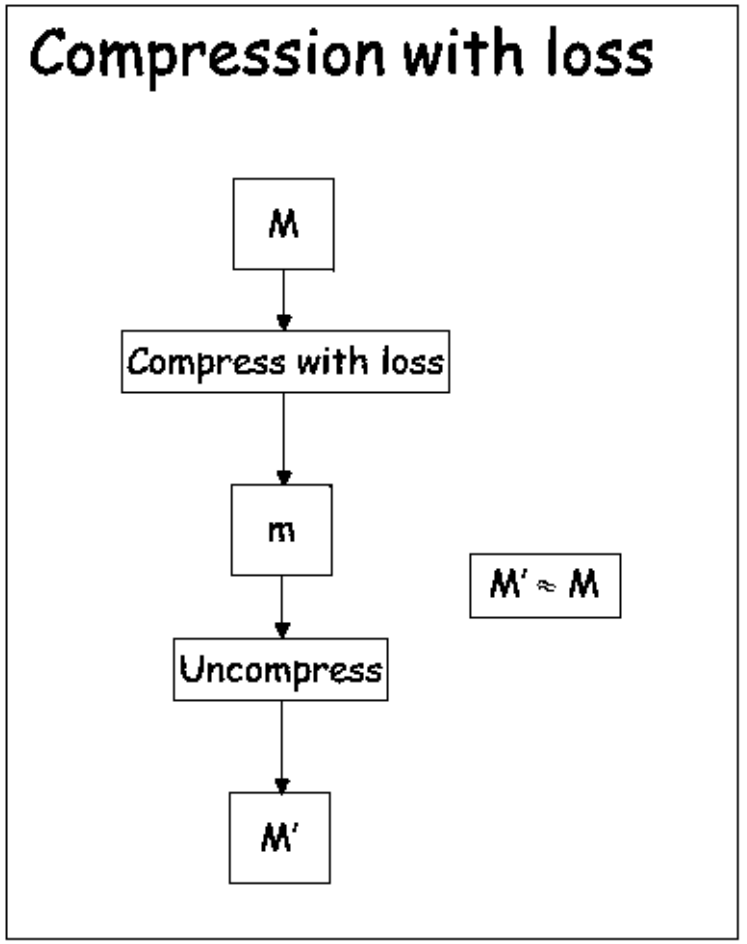
1. The **Lossy compression** technique means that some of the data is lost when it is decompressed.
2. Compression bases on the assumption that the recent data files save more information than human beings can "perceive". Thus, the irrelevant data can be removed.
3. Lossy image compression can be used in digital cameras, ***to increase storage capacities with minimal degradation of picture quality***.

Original Lena Image (12KB size)

Lena Image, Compressed (85% less information, 1.8KB)

Lena Image, Highly Compressed (96% less information, 0.56KB)


Lossless Compression
M
Lossless Compress
m
Uncompress
M


Compression with loss
M
Compress with loss
m
Uncompress
M'
M' ≈ M

Run Length Encoding (RLE):

- RLE is a simple compression algorithm used to compress sequences containing subsequent repetitions of the same character.

Input sequence:
0,0,-3,5,1,0,-2,0,0,0,0,2,-4,3,-2,0,0,0,1,0,0,-2
Run-length sequence (1): (15 samples)
(0,2)(-3,1)(5,1)(1,1)(0,1)(2,1)(0,4)(2,1)(-4,1)(3,1)(-2,1)(0,3)(1,1)(0,2)(-2,1)
Run-length sequence (2): (10 samples)
(# zeros to skip, next non-zero value)
(2,-3)(0,5)(0,1)(1,-2)(4,2)(0,-4)(0,3)(0,-2)(3,1)(2,-2)
• Method 2 reduces the number of samples to code.
• It is effective when data stream contains long runs of zeros
Implementation is simple.

10	10	10	10	10	10	10	10
10	10	10	10	10	12	12	12
10	10	10	10	10	12	12	12
0	0	0	10	10	10	0	0
5	5	5	0	0	0	0	0
5	5	5	10	10	9	9	10
5	5	5	4	4	4	0	0
0	0	0	0	0	0	0	0

First row :10,8

Second row :10,5,12,3

Third row:10 5 12 3

Fourth row : 0,3,10,3,0,2

Fifth row :5 3 0 5

Sixth row :5,3,10,2,9,2,10,1

Seventh row :5,3,4,3,0,2

Eighth row :0,8

Try Column – by – Column ?
Which better ??

Exercise:

Write a MATLAB code to:

1. Function to encode an image in RLE compression algorithm?
    - It will give you the original size and compressed image.
2. Function to decode a compressed image by RLE compression image?
    - You must enter the original size.

Entropy:

The definition of entropy is aimed at identifying often-occurring symbols as short *codeword*

For example, E occurs frequently in English, so we would give it a shorter code than Q.

$$\eta = H(s) = \sum_{i=0}^{n} p_i \log_2 \frac{1}{p_i} = -\sum_{i=0}^{n} p_i \log p_i$$

p_i is the probability that symbol Si in S will occur.

$\log_2(\frac{1}{p_i})$ indicates the amount of information contained in characters.

Mathematical Exercise:

Suppose you have the 4*4 gray-scale image block:

39	**39**	**126**	**126**
39	**39**	**126**	**126**
39	**39**	**126**	**126**
39	**39**	**126**	**126**

1. Apply (RLE) – *mathematically* – on the block?
2. What is the compression Ratio?
3. Calculate the Entropy?