Introduction

File handling is one of the primary functions of a computer system. Based on the type of data that needs to be stored, several types of file formats are available. Each file format occupies certain storage space.

An image file with good quality occupies around 1 MB and a video file needs to store 25 frames per second occupying large storage space. Thus, compression methods are used to reduce the size of the files. Compression is also helpful in reducing the download time of image, audio and video from the internet. In this chapter, different types of files formats and compression methods are explained.

Compressing Image files

An image file is compressed to reduce the amount of memory required to store the file. Compressing an image also changes its attributes such as file type, resolution, dimensions and bit depth.

There are two types of compression:

a) Lossless compression:

When the file is compressed, the quality of the image remains the same and the image can be reconstructed to its original form. In this case, information is very important and cannot be lost. Let us consider a text file with the following sentence:

“See a pin and pick it up, all the day you'll have good luck; see a pin and let it lie, bad luck you'll have all day”

This text file can be compressed by making a table for this information.

Index	word
1	see
2	a
3	pin
4	and
5	pick
6	it
7	up
8	all
9	the
10	day
11	you'll
12	have
13	good
14	luck
15	let
16	lie
17	bad

The sentence can be coded in the form of numbers in the above table and stored in the computer:
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 1 2 3 4 15 6 16 17 14 11 12 8 9 10
This saves memory but using codes for words that are repeated. This is a type of lossless compression. With the code and the index table, the complete sentence can be recreated.

b) Lossy compression:

When a file is compressed, the unnecessary bits of information are removed permanently. This information is less likely to be noticed by humans. This type of compression is used for photographs where the information to be compressed cannot be predicted.

TIFF (.tif) and BMP (.bmp) refers to raw bitmap, that is, uncompressed image files. As these file formats are uncompressed, they represent images with highest image quality.

The different types of compressed images formats are:

a) Loseless

PNG: PNG is a lossless compression type that is mainly used for transferring images over the internet.

GIF (.gif): GIF (Graphics interchange format) compresses images to a maximum of 8-bit depth (256 colours). It is thus not used for high-quality images. Mainly used for images with large areas of solid colour. A sequence of gif images is used to store animated graphics. This a lossless compression type and is used for small images such as logos, icons, etc.

PDF: PDF (Portable document format) is an open standard to encode text and graphics. Many applications are capable of decoding PDF files.

b) Lossy

JPEG or JPG (Joint Photographic Experts Group): JPEG is a lossy compression type that offers a high compression rate compared to PNG. This method is used in digital cameras. User is allowed to choose the compression ratio. Images to be used on a web page can be compressed with a large compression ratio.

JPEG offers a colour depth of 24 bits and can represent 16.7 million different colours.

Musical instrument digital interface (MIDI)

MIDI is a communication protocol that enables electronic musical instruments to interact with each other using information and control signals.

This protocol uses a asynchronous serial 8-bit transmission for communication. A start and stop bit is required in asynchronous transmission.

MIDI file instructs a device on how to produce a musical note using several commands. A command is identified by its specific sequence of bytes. There are two types of MIDI bytes such as status and data. Status bytes indicate instruments about the type of action to be performed. MIDI files operate on 16 channels and the action for all these channels are present in status byte. Some of the actions instructed are: notes on/off, polyphonic key pressure, pitch bend, etc. A note on/off specifies whether the key is pressed or released in the instrument. The key pressure indicates how hard a key is pressed and hence, denotes the loudness of music note.

The data bytes gives information to status byte. Pitch byte denotes the note to be played. Velocity byte denotes the loudness of the note.

MIDI files are developed when sound from an instrument is recorded. The extension for MIDI files is .mid. Using a MIDI file, a music identical to original can be played every time. This file also enables us to play the same music through different instruments. A sequence software is required so that the commands in MIDI files are recognised.

As this protocols allows the usage of 16 channels, 16 different instruments can be instructed at the same time. Hence, an electronic orchestra can be created by giving input to each instrument through a sequence software. As a result, MIDI files are widely used in recording studios, orchestras and musical scores in films.

Compared to a MP3 file, the size of MIDI file is small. This is because MIDI file does not contain any audio track. MIDI files are used in web pages. Due to its less memory requirement, MIDI files are also used in mobile phones to store ring tones.

Videos

Digital video is created by playing a series of images at high speed. A typical HD video has a frame rate of 60 fps. Advanced video standards support up to 300 fps. The sampling rate of a video is given in frames per second. This is also measured in Hertz. Video files also have a bit rate that defines the quality of audio and image.

Compressing a video file reduces the resolution, dimensions and bit rate. Compressing a video file may also lead to poor quality and random coloured blocks on the screen. These blocks are called artefacts. MP4 and MOV are few examples of lossy video file formats. The following image shows the quality of uncompressed and compressed video.

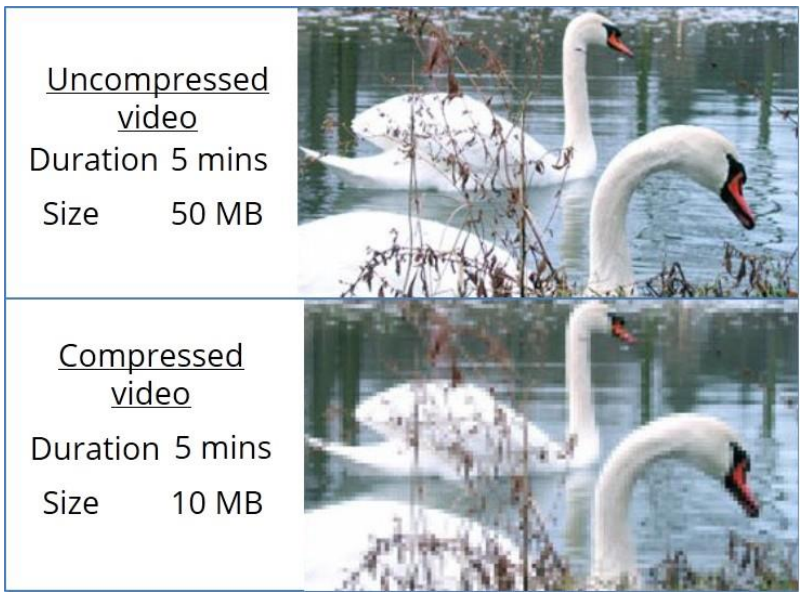


Figure 5: Video compression

Streaming audio and video files

Compression is very helpful in streaming and downloading audio and video files.

MP3 file format is used for audio compression. This ensures up to 90% of compression. MP3 files are also used for storing files in computers, MP3 players, mobile phones, etc. The CD files are converted to MP3 file format using file compression software. Even though, the quality of MP3 file cannot match the original CD file, it is still satisfactory for various purposes. This is a lossy file compression technique. User does not feel any difference while listening to the MP3 file because of perpetual music shaping in file compression algorithms that only removes details that humans cannot hear. Therefore, only few parts of the sound file are removed without compromising much on quality.

The bit rate affects the quality of MP3 file. It ranges from 80 kbps to 320 kbps. An original CD file has a bit rate of 200 kbps and above.

MP4 files is similar to MP3 files, but also allow storage of other multimedia files such as video, animation, photos, etc. Video files are also compressed to stream online in MP4 format. This file is used for transmission over digital channels, cables and satellites. DVD movies are available in this format. A high-definition 720p (HD 720) video has a bit rate of 1200-4000 kbps. A 4K video has a bit rate of 8000-14000 kbps. The stream rate for standard definition video is 1 Mbps, for HD video is 3.5 Mbps and 4K ultra HD video is 15 Mbps.

FLAC and ALAC are examples of open source lossless compression formats. File sizes can be reduced up to 50% without losing quality. A 1,411 kbps WAV file can be compressed to a 64 kbps MP3 (Moving pictures experts group audio layer 3)) file by losing some data permanently. The bit depth is also reduced to remove data. Hence, in lossy compression, the bit depth is a variable. MP3 and AAC are patent codecs and Ogg Vorbis is an example of an open-source lossy compression technique.

Codecs and Compression Algorithms

Codecs are programs that encode or decode an audio, image or video file. Compression codecs are aimed at reducing the size of a file without affecting its quality. Algorithms decide the amount of data that can be removed to reduce the file size.

Run length encoding (RLE) is an example of a compression algorithm that converts the consecutive similar values into a code. This code consists of the identical value and the number of times this value is repeated. This is a lossless type of compression.
Let us consider a black-and-white image carrying 8 bytes of data as shown in the figure below. The computer stores binary value 1 for white and binary value 0 for black for each row of the image.

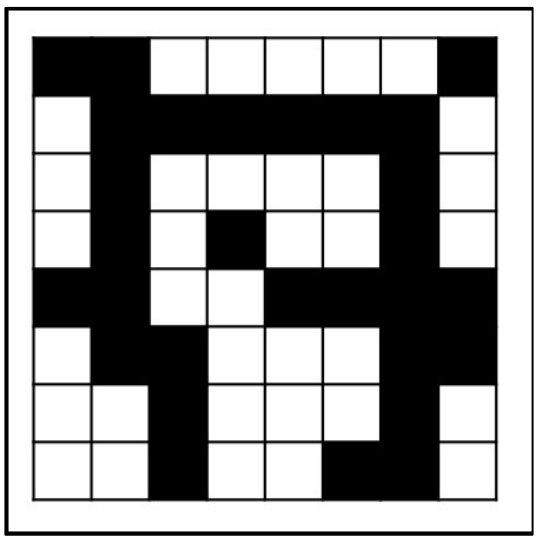


Figure 6: A black and white image

The first row in the image can be represented as 2 0 5 1 1 0. This code represents 2 black pixels, 5 white pixels and 1 black pixels. Similarly, the second row in the image is represented as 1 1 6 0 1 1.
This type of coding is not efficient if the file does not have many runs. In such cases, the file size may increase instead of getting compressed. RLE is thus used only in simple images with a large area of same colour.

RLE are also used to compress video files.

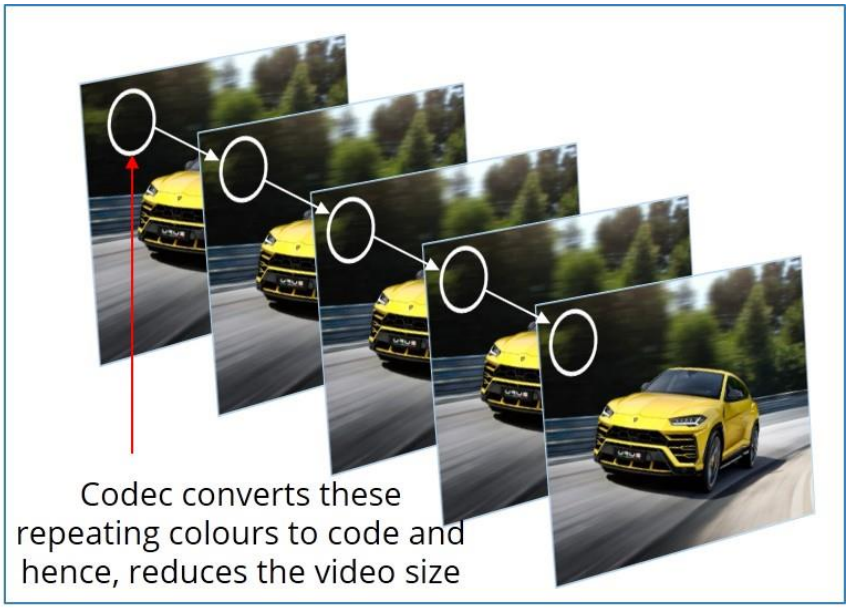


Figure 7: Compression algorithm

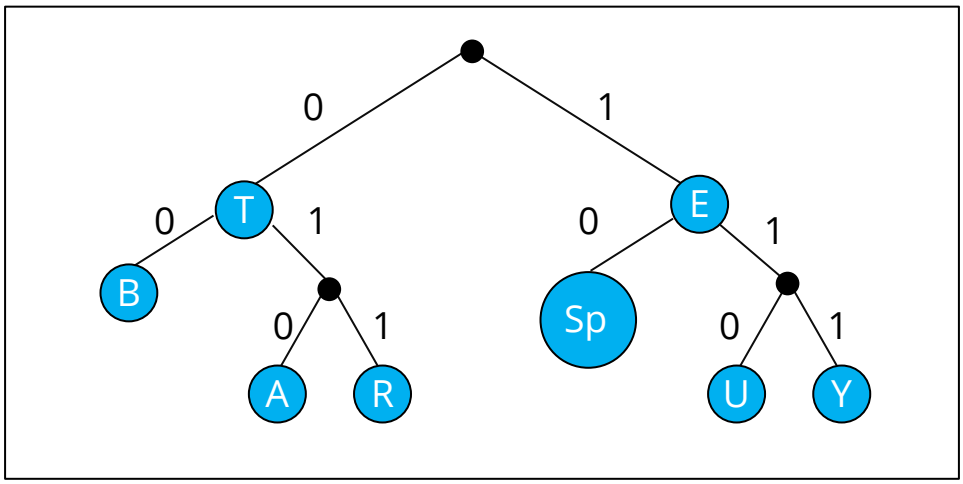
Huffman coding

Huffman coding is a compression technique used to reduce the number of bits that represents each letter. In ASCII coding, each letter is represented using 7 bits. In Huffman coding, each letter is represented with a different number of bits. The most frequently appearing letters are represented with a lesser number of bits and, hence, the number of bits required to store information is reduced.

Consider the sentence: Betty ate butter. The frequency of characters in this sentence is shown in the table below. There are 17 characters in total (including spaces). Therefore, the total number of bits used to represent their ASCII codes are: 17×7= 119 bits.

Letter	A	B	E	R	T	U	Y	Space
Frequency	1	2	3	1	5	1	1	3

A simple Huffman binary tree is shown in the figure below. A binary tree is a data structure made of nodes and is constructed based on hierarchy. A parent node in a binary tree has up to two child nodes.



Each letter is now assigned a binary value:

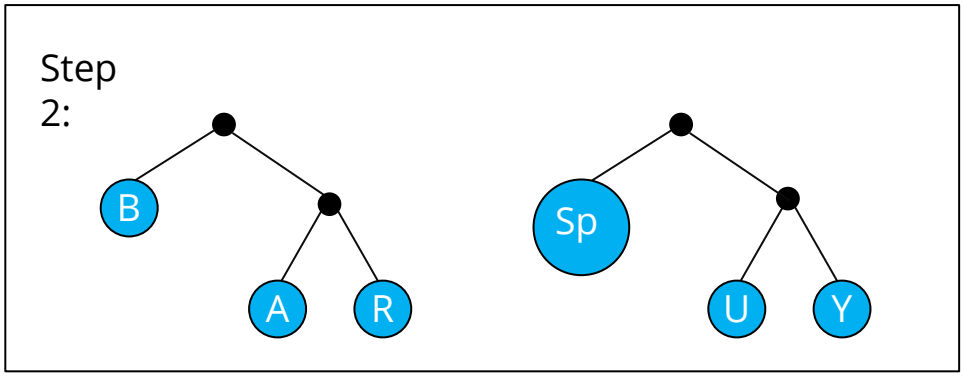
Letter	A	B	E	R	T	U	Y	Space
Frequency	1	2	3	1	5	1	1	2
Binary value	010	00	1	011	0	110	111	10

It can be noticed that letters with higher frequency are assigned with a smaller number of bits. Substituting these values in the sentence and calculating the total number of bits: 3 + 4 + 3 + 3 + 5 + 3 + 3 +4 = 28 bits. Using Huffman coding, we have saved 119 – 28 = 91 bits.

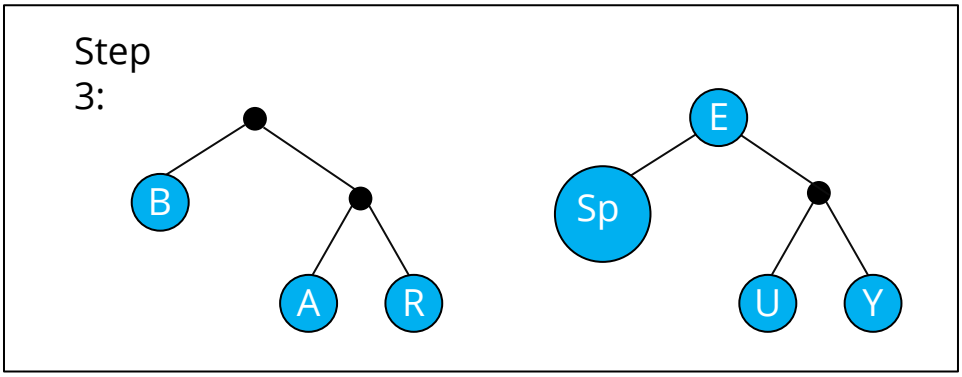
Let us now learn the method to build this tree. We need the lowest number of bits for letters with higher frequency. Hence, we must award slightly a greater number of bits for letters with lesser frequency. We shall start with the letters with lower frequency. Letters A, R, U and Y have the lowest frequency of 1.

Step 1:

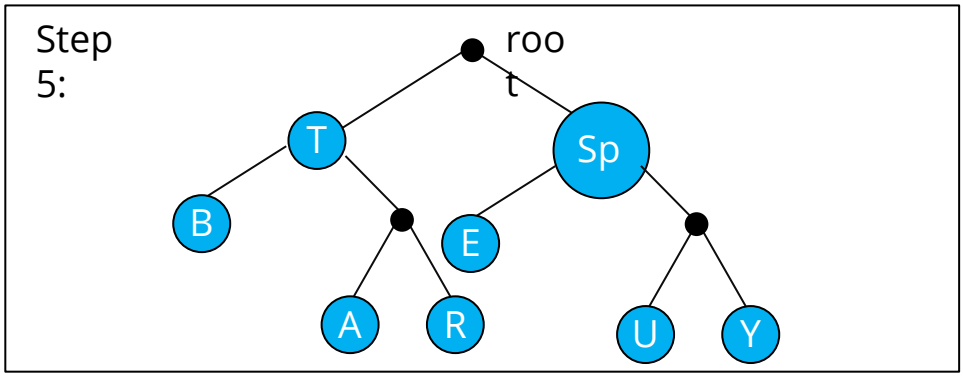
Next, let us consider characters B and space with a frequency of 2 each.



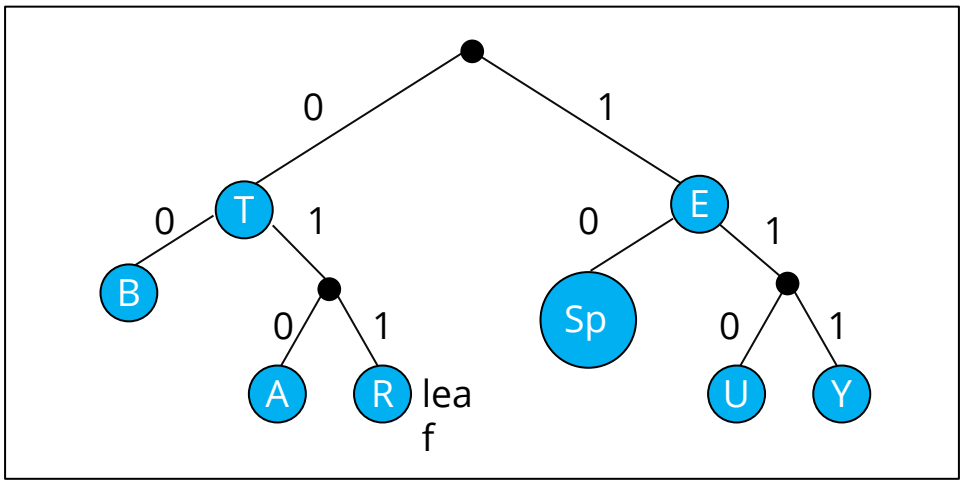
Next, let us consider letter E with a frequency of 3.



The letter T has the highest frequency of 5. Let us include T and connect the nodes to the root.



The final step of forming a Huffman tree is giving binary values for each connection. Left branches are assigned a value of 0 and right branches are assigned a value of 1. Each path terminates to a leaf.



Using the tree, each letter is assigned a binary value, starting from the root to the leaf.

Letter	A	B	E	R	T	U	Y	Space
Binary value	010	00	1	011	0	110	111	10

This is an example to use Huffman coding. There shall be different character coding for a letter.