

Compression

Compression is taking original **uncompressed data** and **compressing** it so that the result requires less data.

There are two types of compression. **Lossy compression** loses some of the information. The format of compression tries to lose information that cannot be noticed by the user. This type of compression is used for images, videos and music. The table below shows the result of JPEG compression on some text stored as an image. Small blocks of gray have been introduced – these are known as **digital artifacts**. You can sometimes also see these on video compression.

Original text	Compressed using lossy compression (JPEG)	Compressed using lossless compression (PNG)
2556 bytes	636 bytes	305 bytes

Compression that doesn't lose any of the original quality is known as **lossless compression**. PNG, GIF and ZIP files both use lossless compression. This is why GIF or PNG is good for compressing files of scanned text. ZIP files need to use lossless compression as they hold files such as text or programs that must **decompress (unzip)** to exactly the same as the original.

Internet File Formats

Files are compressed either because it reduces the storage space, for example on a DVD or Blu-ray disk, or because the data needs to go through a network connection, such as the Internet where there is limited bandwidth. For instance, HD (high definition) video uncompressed would require a data rate of 140MB/s. Compressed HD video using MPEG4 requires only 1MB/s.

File format and file extension	Type	Usage
JPEG – Joint Picture Electronics Group .jpeg / .jpg	Lossy	Used to compress images. Works well with photographs but not well on images with sudden changes in colour such as black text on white
MPEG – Motion Picture Electronics Group .mp4 / .mpeg / mpg	Lossy	Used to compress video and audio. Various standards have been made. Used for digital TV, DVD (MPEG2) and most online video (MPEG4)
MP3 – MPEG audio layer 3 .mp3	Lossy	Used to compress audio. It is the audio part of the MPEG2 video standard. The newer AAC (Advanced Audio Coding) has better sound quality
GIF – Graphics Interchange Format .gif	Lossless	Used for compression of images and also allows for short animation
PDF – Portable Document Format .pdf	Lossless	Now an open standard for text, fonts, graphics and layout. Used for books and documents
PNG – Portable Network Graphics .png	Lossless	Created as a replacement for GIF. Works well with photos and illustrations. Allows for transparency