1.3 Data storage and file compression

1.3.1 Measurement of data storage

A **bit** is the basic unit of all computing memory storage terms and is either 1 or 0. The word comes from **b**inary dig**it**. The byte is the smallest unit of memory in a computer. 1 byte is 8 bits. A 4-bit number is called a nibble – half a byte.

1 byte of memory wouldn't allow you to store very much information so memory size is measured in the multiples shown in Table 1.4:

▼ **Table 1.4** Memory size using denary values

Name of memory size	Equivalent denary value
1 kilobyte (1 KB)	1 000 bytes
1 megabyte (1 MB)	1 000 000 bytes
1 gigabyte (1 GB)	1 000 000 000 bytes
1 terabyte (1 TB)	1 000 000 000 000 bytes
1 petabyte (1 PB)	1 000 000 000 000 000 bytes
1 exabyte (1 EB)	1 000 000 000 000 000 000 bytes

The above system of numbering now only refers to some storage devices but is technically inaccurate. It is based on the SI (base 10) system of units where 1 kilo is equal to 1000.

A 1 TB hard disk drive would allow the storage of 1×10^{12} bytes according to this system.

However, since memory size is actually measured in terms of powers of 2, another system has been adopted by the IEC (International Electrotechnical Commission) that is based on the binary system (Table 1.5):

▼ **Table 1.5** IEC memory size system

Name of memory size	Number of bytes	Equivalent denary value
1 kibibyte (1 KiB)	2^{10}	1 024 bytes
1 mebibyte (1 MiB)	2^{20}	1 048 576 bytes
1 gibibyte (1 GiB)	2^{30}	1 073 741 824 bytes
1 tebibyte (1 TiB)	2^{40}	1 099 511 627 776 bytes
1 pebibyte (1 PiB)	2^{50}	1 125 899 906 842 624 bytes
1 exbibyte (1 EiB)	2^{60}	1 152 921 504 606 846 976 bytes

Advice
Only the IEC system is covered in the syllabus.

This system is more accurate. Internal memories (such as RAM and ROM) should be measured using the IEC system. A 64 GiB RAM could, therefore, store 64×2^{30} bytes of data (68 719 476 736 bytes).

1.3.2 Calculation of file size

In this section we will look at the calculation of the file size required to hold a bitmap image and a sound sample.

The file size of an image is calculated as:

image resolution (in pixels) × colour depth (in bits)

The size of a mono sound file is calculated as:

sample rate (in Hz) × sample resolution (in bits) × length of sample (in seconds)

For a stereo sound file, you would then multiply the result by two.

Example 1

A photograph is 1024 × 1080 pixels and uses a colour depth of 32 bits. How many photographs of this size would fit onto a memory stick of 64 GiB?

1. Multiply number of pixels in vertical and horizontal directions to find total number of pixels = (1024 × 1080) = 1 105 920 pixels
2. Now multiply number of pixels by colour depth then divide by 8 to give the number of bytes = 1 105 920 × 32 = 35 389 440/8 bytes = 4 423 680 bytes
3. 64 GiB = 64 × 1024 × 1024 × 1024 = 68 719 476 736 bytes
4. Finally divide the memory stick size by the files size = 68 719 476 736/4 423 680 = 15 534 photos.

Example 2

A camera detector has an array of 2048 by 2048 pixels and uses a colour depth of 16. Find the size of an image taken by this camera in MiB.

1. Multiply number of pixels in vertical and horizontal directions to find total number of pixels = (2048 × 2048) = 4 194 304 pixels
2. Now multiply number of pixels by colour depth = 4 194 304 × 16 = 67 108 864 bits
3. Now divide number of bits by 8 to find the number of bytes in the file = (67 108 864)/8 = 8 388 608 bytes
4. Now divide by 1024 × 1024 to convert to MiB = (8 388 608)/(1 048 576) = 8 MiB.

Example 3

An audio CD has a sample rate of 44 100 and a sample resolution of 16 bits. The music being sampled uses two channels to allow for stereo recording. Calculate the file size for a 60-minute recording.

1. Size of file =

sample rate (in Hz) × sample resolution (in bits) × length of sample (in seconds)

2. Size of sample = (44 100 × 16 × (60 × 60)) = 2 540 160 000 bits
3. Multiply by 2 since there are two channels being used = 5 080 320 000 bits
4. Divide by 8 to find number of bytes = (5 080 320 000)/8 = 635 040 000
5. Divide by 1024 × 1024 to convert to MiB = 635 040 000 / 1 048 576 = 605 MiB.