Digital Audio, Digital Video and Compression

- Digital Data Rates
- Digital Audio
- Digital Video
- Digital Compression

Digital Data Rates

- CD Audio (44.1 kHz, 16 bit) - 1.4112 Mbits / sec
- Uncompressed digital video - 167 Mbits / sec
- Video CD (MPEG1, constant bit rate) - 1.2 Mbits / sec
- DVD Video (MPEG2, variable bit rate) - 3.5 Mbits / sec average, 11 Mbits / sec peak

Digital Audio

When a recording is made, vibration in the air - sound waves - are converted by a microphone into an electrical signal. This signal is a complex, constantly changing voltage whose many subtleties describe the sound captured. It must be converted intoa digital signal, a series of '1's and '0's or bits, before it can be recorded onto a CD or any other digital format.

As with any conversion from analogue to digital of anything but the most simple waveforms, a decision has to be made about how much digital data should be used. The more bits available to the analogue to digital (A to D) converter, the more accurate the conversion will be and the closer the decoded version will sound to the original signal.

If you imagine a graph with an analogue wave drawn on it, the wave can be converted to digital information by measuring the height of the wave above the x axis at regular intervals along it. The resulting values can be used to reconstruct the wave by plotting them on another graph at the same regular intervals along the x axis. The more accurately the height can be measured, and the smaller the regular intervals the more accurately the wave will be reproduced.

In digital audio terms, the accuracy to which the height above the x axis can be measured is described in bits per sample, and the size of the intervals is the sampling frequency. On a CD, each of the two audio channels (left and right) is sampled using 16 bit accuracy - that is to say there are 65536 discrete levels available to describe the height of the wave. Each channel is sampled 44,100 times a second, a sampling frequency of 44.1 kHz. The amount of data used per second can be calculated as follows -

bits per sample x sampling frequency x no. of channels = data per second

16 x 44100 x 2 = 1411200 bits per second

While this is the data rate used for CD, there are other formats being developed which offer more bits per sample, up to 24 bits, faster sampling frequencies, 96 kHz and even 192 kHz, and more channels. The data required for each of these combinations can still be calculated using the equation above.

Digital Video

Video is also an analogue signal, and is converted to a digital one in much the same way as for audio. However, video is rather more complex than audio, so it requires significantly more bits to do it accurately. Audio professionals have been taking advantage of 24 bit recording for some time, using data rates of more than 2 million bits per second for their master tapes which are then converted down to 16 bits for transfer onto CD. The equivalent digital video recorder needs a data rate of around 167 million bits per second to record video material with enough accuracy to satisfy the professional market.

Video recorders (analogue or digital) simulate moving pictures by showing a series of still pictures in quick succession. If the picture changes frequently enough, our brains are fooled into seeing smooth motion on the screen. There are two major systems for displaying video pictures in wide use.

NTSC is the US standard, also in use in Japan and around Asia and elsewhere. It uses 30 interlaced frames (or pictures) per second. Each frame has a vertical resolution of 480 pixels and a horizontal resolution of 720 pixels.

PAL is the European standard, also used in China, Australia and in some parts of Asia. It uses 25 interlaced frames per second. Each frame has a vertical resolution of 576 pixels and a horizontal resolution of 720 pixels.

Interlacing is a system which divides each frame into two fields. Each field has half the vertical resolution of the full frame, and is displayed on alternate lines of the television screen. Fields are updated alternately at twice the normal frame rate. So for an NTSC picture, fields are updated every 60th of a second.

When video is converted from analogue to digital, the signal undergoes much the same process as it does for audio. Each pixel has colour and brightness information, which is described digitally using 16 bits. The calculation to work out the total data rate for digital video is therefore -

bits per pixel x vertical resolution x horizontal resolution x frame rate = data per second

For NTSC - 16 x 480 x 720 x 30 = 165888000 bits per second

For PAL - 16 x 576 x 720 x 25 = 165888000 bits per second

Digital Compression

There are circumstances where it is desirable for audio and video material to occupy less digital space. This could be for audio delivery over the Internet, or storage of a movie on a DVD for example. In these cases it would be possible to reduce the accuracy with which the analogue signal is converted by simply reducing the number of bits per sample or bits per pixel, but the result would be seriously compromised in terms of its accuracy - it would sound or look horrible. A much better result can be achieved by converting to digital using the same accuracy and then using a computer processor to analyse the information and squeeze it into a smaller space. This is what is meant by digital compression.

DVD relies on a compression system called MPEG2. It was developed by a body called the Motion Picture Experts Group as part of an extended project to develop a series of compression systems for video and audio material.

The first consumer application for compressed digital video, Video CD, used a system called MPEG1. It introduced us to many of the basic building blocks used in MPEG2. Its aim was to reduce the data required for digital video to the CD Audio data rate, to squeeze 167 Mbits per second into about 1.4 Mbits per second, a reduction of over 99%, while still achieving the equivalent of VHS picture quality.

MPEG1 uses several stages to reduce the amount of data used. Firstly the picture size is reduced, halving the horizontal resoution to 352 pixels and throwing away one of the interlaced fields, effectively halving the vertical resolution. The decoder reproduces the horizontal resolution by interpolation and the vertical by playing the same field twice for each frame. Next the picture is divided into 16 x 16 pixel squares and encoded using a more efficient coding scheme called DCT (discrete cosine transform).

Most of the data reduction is achieved by the last stage - the frames are analysed as a sequence and any parts of the frame that do not change from picture to picture are not stored, but are simply repeated from the previous picture. Each 16 x 16 pixel block is also compared with areas around it on successive frames to discover if the picture is the same but has moved to another place on the screen. This is called motion estimation and can result in significant data savings.

MPEG2 uses the same principals as MPEG1, but without the resolution reduction of MPEG1. MPEG2 is widely used in digital satellite, cable and terrestrial broadcasting as well as in DVD Video. For its DVD application, it achieves the most efficient compression by varying the data used to store video material depending on how difficult it is to encode. Due to the way in which MPEG compares successive frames, slow moving video material tends to compress more successfully than fast moving sequences, or sequences when the entire frame changes frequently. When compressing for DVD, the encoder can change the number of bits per second used depending on the material, and consequently use the disc space most efficiently.

As with any digital systems used to describe analogue signals, like audio or video, there has to be a compromise between quality and data rate. However, digital compression will continue to improve in quality as smarter systems are developed and as it becomes possible to put more and more processing power into a single inexpensive chip. The next generation of optical discs, High Density DVD, is likely to be developed for High Definition Television, which has a much higher resolution picture and therefore needs more bits per second. Another possible application for a higher capacity disc will be home digital video recording - it may be cheaper to develop a disc with more recording space and to compress the video less than to develop a domestic MPEG2 encoder capable of achieving the relatively low data rates recordable onto current DVD Recordable formats.