Gerald's Column
by Gerald Fitton

I've had quite a lot of feedback this month about some recent and some long ago articles; I wasn't sure whether to do a PipeDream/Fireworkz article, a domain name or email article, an SMS article or what. Finally I decided to discuss Media files. Why? One reason is because, like many of my correspondents, I believe that it is one of the fastest growing uses of computers. There are other reasons which I shall come to later.

In this article I shall discuss the lossy compression of bitmap (picture) files.

What are "Media" files?

Any definition of "Media Files" must include picture, audio and video files but I feel that I have to include two other formats in these discussions. These are html and pdf. Don't ask my why I want to include html and pdf, I just do! Consequently, like most totally rational human beings who trust their instincts, I shall be totally irrational and include them!

Perhaps, if you manage to follow my rambling discourse then you might be able to help me rationalise my reasons for linking html and pdf files with more conventional media files?

Picture files come in all sorts of formats from sprites to drawfiles and mixtures of the two. There are two main categories, bitmaps and vector files. Sprites are bitmaps; they are saved pixel by pixel. Drawfiles are vector files; they are saved as a series of commands for drawing lines and filling shapes.

By Audio files I include not only music (including MIDI formats) but also speech, the sound of a peewit and even a streamed (live) BBC radio program.

With the 'every home must have one' video camera era upon us it is time to think about how we might process our home movies on our home computers. Even if you don't yet want to make your own home movies I am sure that you will be considering downloading video files from the Internet.

Finally, with your permission, I would like to include text files as a wonderfully expressive form of media file. Whilst a picture might paint a thousand words, there are writers whose words create wonderful pictures in the minds of the reader. One of my lasting childhood memories is having read to me the sad tale of the "Lady of Shalott" - the pictures which this story imprinted on my mind are better than the famous classic by Waterhouse!

I have sent Paul a digitised copy of this beautiful painting. In it the Lady of Shalott has just embarked on her last journey down the river to Camelot, having been enraptured by her view in her mirror of the most handsome Sir Lancelot. My 'mind' pictures are better!

Presentation

When you have constructed your masterpiece you will want to present it in some way to others. Perhaps you will invite them to see a film show on your computer, or you will transfer it to a CD or DVD or possibly to a medium growing in popularity, the Internet.

John William Waterhouse: *The Lady of Shalott* (1888)

If you upload your files to the Internet then your friends and relatives from all around the world will be able to access them - not only that but the Internet is a great place for the secure storage of your new novel, your diary, holiday pictures, favourite music, the sound of a chiff-chaff or even a movie of the first steps of your new born!

File Sizes

In earlier articles I have dealt with file sizes at length. Text files are compact; pictures in vector format are small; bitmaps (eg sprites) are larger; audio files are yet larger; video files are the largest of all.

Text file sizes can generally be measured in KB; vector graphics files are a few tens of KB; a RAW format 6MP (Mega Pixel) snapshot will be about 18MB; a WAV (wave) format audio file lasting a few minutes (a musical track from a CD) will be about 40MB; an AVI video lasting a few minutes will set you back about 1GB!

Compression

The formats which I quoted above are not compressed. There are many common and some unusual standards for compression which go under the general heading of 'codecs'. As far as I can determine the 'coder - decoder' co-dec applies to an analogue to digital conversion (and back) whereas a 'compression - decompression' codec is all digital. I could be wrong.

You will all be familiar with zip files. This is a compressor which reduces the size of files but it has a very important characteristic; it is lossless. A file which has been compressed using a lossless codec can be recovered to the original file exactly as it was. In some circumstances, for example compressing a program, lossless compression is essential.

Lossy Compression

Most codecs used for media files are lossy.

One example of a codec for bitmap pictures is the jpeg codec. Pictures taken with a digital snapshot camera in RAW format and then compressed using the jpeg codec lose some information along the way and the lost detail contained in the original RAW format can never be recovered.

When you use any lossy codec on a picture, audio or video file, you will lose some information. Generally you can choose just how much compression you use and therefore you can choose just how much information you lose. There is always a trade off between quality and file size. With nearly every lossy codec I have seen there is an option which allows me to choose just how much loss my media file can stand.

I have been experimenting with pictures, sound and video files using different codecs and discovering what compromise between quality and file size suit my creations. In this and following articles I shall share my findings with you so that you too can choose the degree of compression which suits your picture, audio and video files.

JPEG Compression

The picture below is a detail from a digital copy of Waterhouse's painting "The Lady of Shalott (1888)". By the way, the "(1888)" is part of the title because Waterhouse painted another picture with the title "The Lady of Shalott (1894)".

Look back to the full picture and you will see that this is a detail taken from the 'car rug' (which the Lady wove on her loom) draped over the side of the boat.

In the pictures below I have extracted an even smaller detail which is 225 pixels square.

In these two pictures you will see the effect of compression on a 225x225 pixel detail. The picture on the left is a 34339 Byte jpeg; the picture on the right is a 2055 Byte jpeg. I hope that you will be able to see the seriously bad pixellation in the picture on the right.

Percentages

Often the degree of compression is stated as a percentage. I don't know what it is a percentage of! Indeed, I don't think that it is a percentage at all! With the software I normally use, the picture on the left results from using a value of 5% and the picture on the right corresponds to 95%.

To confuse you even further, with some of the packages I've tried the percentages work the other way around. By this I mean that a compression of 95% gives a larger file and a better picture than 5%!

What is certain is this. The original uncompressed file (a bitmap - such as a sprite file) is 148KB. The calculation is 225*225*3/1024=148.3KB. The (lossless) tiff file of this 225x225 pixel detail contains extra non picture information which brings it up to 151KB.

The reduction in file size for the picture on the left is from 148KB to 34KB. The picture on the right is reduced from 148KB to 2KB, a substantially greater reduction in file size.

With this and many other pictures I have found that a reduction of file size to about 10% of the original (10% of the bitmap file size) will be acceptable to most people for most subjects. With the software I use this corresponds to a compression of 25%.

Have a look at the picture below. It is about 15KB in size. Compare the pixellation with the 34KB file and I think you'll agree that there is not a lot of difference.

To be on the safe side I usually use 12% as my default value of the compression. When I apply this degree of compression to the detail from Waterhouse's oil colour this jpeg is a little under 23KB, about 15% of the original (148KB) file size. I believe that the increase from 10% (15KB) to 15% (23KB) of the original file size is a small price to pay for not having to scrutinise every picture in detail to be sure I've not lost too much of the quality.

Audio and Video

The size of uncompressed pictures are just about acceptable but for audio and video some compression is essential, particularly if you wish to stream the files live over the Internet.

I shall look at some popular audio codecs next month.

AVLine

I am working on producing a CD for distribution. The initial version is available now.

I have called it AVLine (cf PipeLine, FireLine and GoldLine). It contains media files and articles which relate to media files. The articles are in PDF format and are linked together using htm files containing links. You will be able to read this CD not only on a RISC OS machine but also on other platforms (eg Windows).

Current members of GoldLine receive a copy as part of their GoldLine subscription.

Return your GoldLine CD now to receive your AVLine CD.

Like GoldLine, the AVLine CD will be upgraded as new articles are added.

Communication

Write to me about Archive things to archive@abacusline.co.uk.
Write to me about GoldLine things to goldline@abacusline.co.uk.
Write to me about AVLine things to avline@abacusline.co.uk.

You might like to try http://archive.abacusline.co.uk with Username "archive" and Password "amusement" (without the inverted commas) if you wish to download this and other articles I have written for Archive.

At present most of the articles are in Fireworkz format - I am in the process of converting them to PDF so that anyone can read them using any platform.

The Fireworkz format files can be read with Fireworkz for RISC OS or with Fireworkz for Windows or with a Fireworkz Reader (there is one on the site which you can download).

The PDF files can be read with any PDF reader.