Split the difference!

I've recently been experimenting with creating sound files in 'mp3' format (MPEG1 Layer III). This made me wonder about how much detail is being 'lost', and how does that vary with the choice of bitrate? Rather than just try to reach a subjective judgement I decided to see if I could obtain some objective measure of the loss.

I started with a test soundfile in LPCM Wave format. For the sake of example I chose a 2 minute section of a work by Ibert which I'd recorded at 48ksample/second from a BBC Prom, stored in a file, Bert/wav. I used Lame to generate a 320kbps stereo mp3 version. Having obtained the mp3 version I then used Lame again to 'decode' the result and create a fresh LPCM Wave version I called Bert320/wav. This meant I now had two LPCM versions, one the original, the other altered by having been converted into mp3 and back.

For the tests I used the Linux version of Lame, and wrote a simple ROX app interface to make conversions 'drag and drop'. The ROX app should be available with this article in the usual way. If you drop an LPCM Wave file onto its icon it will generate an mp3 version. And if you drop the mp3 onto the icon it will decode it into an LPCM Wave.

I also wrote a new RISC OS application I've called '!WAV_Difference'. If you give this the names of two LPCM Wave files it will check to make sure they share the same sample rate, duration, etc. If they are comparable it then reads though the two input files and takes the difference between the pairs of sample values. The results are then written into a new LPCM Wave file. Hence if the two input files were identical the resulting output 'difference' file would contain a series of zeros and be totally silent if played.

Having done the above I could use some of my previous RO apps to analyse the results. I could also listen to the 'diff' file generated by the above. Alas this made it obvious that the difference was far from silent, and that the difference between the initial soundfile and the one that had been cycled though being a 320kbps mp3 were eaily loud enough to hear when played in isolation!

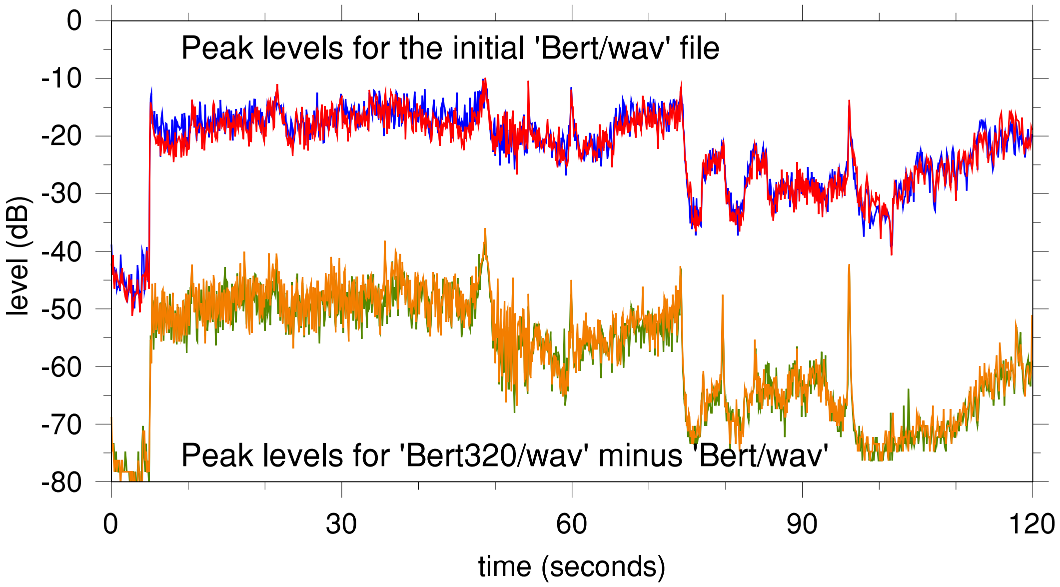

Figure 1

Figure 1 shows the results of using !WAV_stats on the original file and the difference. The plots are of the peak level found during each 100 millisecond chunk of the files. The top pair of lines are the left and right channel for the original soundfile. The lower pair of lines are for the difference made by being turned into mp3 and back to LPCM. For the sake of the test I'd chosen 320kbps to aim at 'high quality'. I'd also told Lame to use true stereo and the highest 'q' value (q 0). However I know it is common for people to use lower bitrates so I repeated the process but this time converting into a 128kbps mp3 and back.

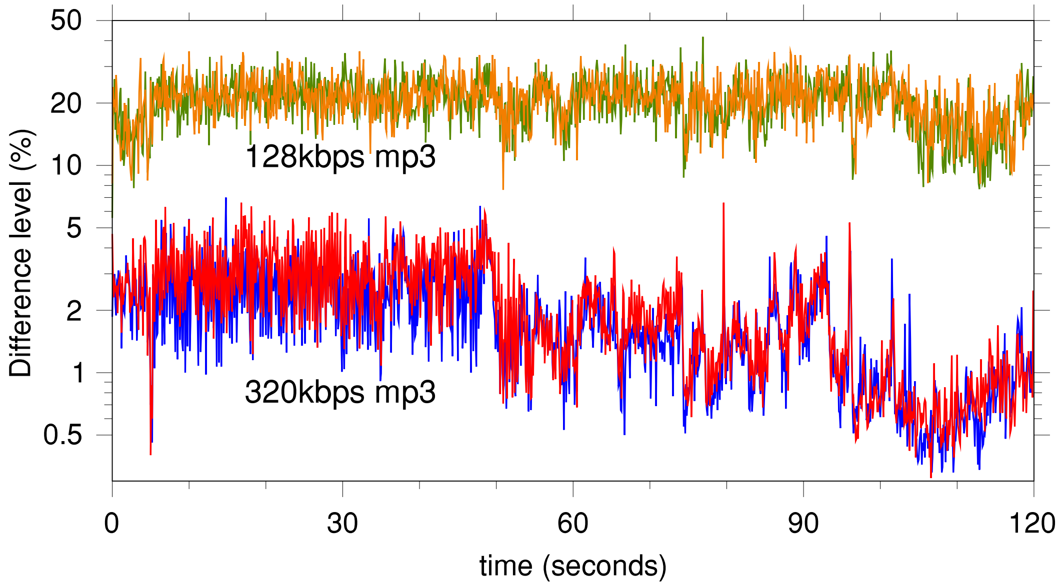

Figure 2

Figure 2 compares how much each of the mp3 conversions tends to alter the result. Here I've plotted the difference between the orginal and the output versions as a percentage of the amplitude of the original. The higher the percentage values, the more the mp3 conversion differs from the original. Looking at the results we can see that for the 320kbps mp3 the departures from the original occasionally peak at over 5%, but are typically in the region of a few percent. Whereas the 128kbps version tends to differ from the original by around 20%!

I must admit that this result did surprise me as it seems so poor. If the 128kbps result were conventional harmonic distortion it would be totally unacceptable for almost any musical purpose. Even the 320kbps result would be worryingly high from a 'hifi' viewpoint. However – as with image files like the JPEG – the departures from the original are being created by a process that tries to apply a set of judgement rules to ensure the result is perceived as being as similar as possible to the original. So the audible degradation may be quite small despite the alarmingly high percentage values.

I did check to ensure that the difference wasn't simply due to the mp3 version being slightly louder or softer in volume. I won't show the results of that examination here, but it showed that overall, the mean levels of the original and conversions differed by much less than 0·1dB. That is essentially inaudible as a change in loudness and would be too small to cause the above differences. So it is the detailed shapes of the waveforms that are being altered by the mp3 encoding process, not the overall level of loudness.

Out of curiosity I worked out the mean spectra of the original wave file and the the difference between the original and 128kbps mp3 version. The results are shown in Figure 3.

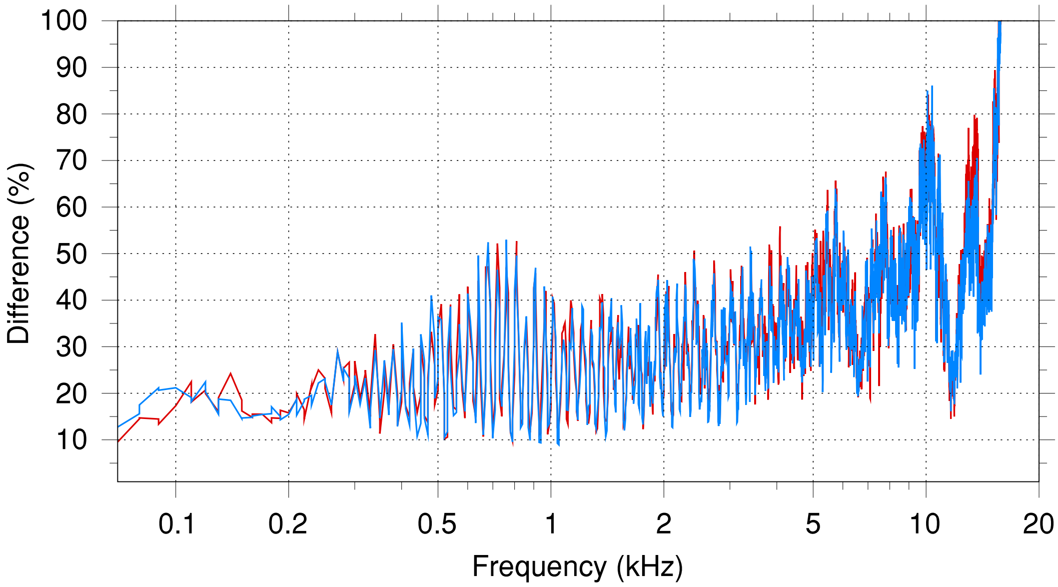

Figure 3

Again I did find these results slightly suprising. I'd expected that the mp3 encoding process would tend to preferentially omit high frequency components. This is because the decision process assumes high frequencie components are less likely to be missed. Overall, the differenced did seem to be in terms of alterations at high frequencies. But there were still makred alterations at lower frequencies.

All being well, the applications I wrote to do the examination will be available from Archive in the usual way. I'd be interested to see how others get on. So try the applications and form your own conclusions about how well the Lame mp3 encoder does at producing results that sound like the original LPCM.

Jim Lesurf
780 words
24th Nov 2009