LIMITATIONS OF AUDIO CODECS AND THE USE OF "DITHER" TO REDUCE SUBJECTIVE IMPAIRMENTS AND THEREFORE COST

Paul A. Conway

Dept. of Electrical Engineering, Leeds Polytechnic.

INTRODUCTION

The topics discussed in this paper are concerned with the development of techniques to ensure that subjective impairments are minimised for a given CODEC strategy thereby implying a minimum cost solution for a given level of subjective quality.

There are two main categories under which research towards further optimisation may be possible.

i) Improvements in CODEC hardware design.

ii) Pre-processing of signals prior to A-D conversion and post-processing after conversion.

It is the latter category in which I have a research interest.

THE NATURE OF SOUND

If we consider any signals suitable for "human consumption" as packets of information, then an understanding of the nature of these sounds is important to the signal processing engineer.

For example:-

i) It is well known that a typical passage of recorded music does not make full use of the baseband channel bandwidth as a function of time.

ii) We know that high frequency distortion is less noticeable subjectively than low frequency distortion.

iii) The human ear is not able to detect changes in phase (although there are those who would not agree with this.)

iv) Noise is less noticeable in high frequency than low frequency material.

v) High quality music generally only contains local components (in the time domain) with large dynamic range for a small percentage of the total time. i.e a large dynamic range is only usually required for a short time.

Furthermore, speech has additional parameters that we need to understand and may need to take account of. For example:-

i) Normal everyday speech requires a typical vocabulary of 10000 to 15000 words and these words consist of some 40 to 50 phonemes.

ii) Does a 3 dimensional plot of time, amplitude and frequency (i.e a spectrograph plot - which has been used by phoneticians for many years) give the total picture or is there other information present (or absent) which the spectrograph doesn't show?

If it is required to process audio signals and/or examine the limitations of a CODEC system, then it is vital that the signal processing engineer develops an understanding of the nature of sound such that the question "What are the important bits?" can be answered.

SOME SPECIFICATION PARAMETERS FOR AN ADC.

Other than the obvious requirement of good linearity, accuracy and resolution, the two cost related parameters of an ADC are:-

i) Conversion Time (i.e maximum sampling rate).

For a typical frequency response of (say) 15kHz, a minimum sampling frequency defined by the Niquist Sampling Theorum would be 30kHz. This however would require a "brick wall" anti-aliasing filter and therefore to allow the implementation of practical filters, a sampling frequency of around 45kHz would seem reasonable.

ii) Sample size (bits/ word).

There is a simple relationship between the theoretical dynamic range of a CODEC system and the number of bits per sample that can be shown to be equal to 6dB/bit.

A recording studio may have a SPL of 20dB and a symphony orchestra may peak at 120dB. i.e a dynamic range of 100dB would be required. Furthermore, rock music may exceed the threshold of pain (130dB) but at this level a little clipping may not be noticeable.

Taking more typical extremes, the average home may exhibit a background SPL of around 43dB and the upper level may be set by many at something less than 120dB SPL. Therefore a dynamic range of something less than 80dB would seem to be quite adequate for many applications.

There is of course another argument that a listener could subject himself to a greater dynamic range when listening with headphones in a quiet room.

LIMITATIONS OF AUDIO CODECS AND THE USE OF "DITHER"

A CODEC with a sampling period of around 22μs would seem to be required with a word size of around 13 to 16 bits.

It seems therefore that any attempt to reduce the cost of a CODEC would result in a reduction in sampling rate and/or a reduction in the size (in bits) of each sample. In fact the subjective outcome of varying these two parameters is well known.

Clearly, a reduction in sampling rate (with suitably modified anti-alias filter) will result in a reduced frequency response and in the absence of any known real time interpolator, the solution to this compromise seems extremely complex.

SPECTRA OF QUANTIZED SIGNALS

Quantization of an audio signal (time sampling) means that the original analog signal is replaced by a waveform constructed of quantized values selected on a minimum error basis from the discreet set available. Clearly, if we assign the quantum values with sufficiently close spacing we may make the quantized wave indistinguishable (subjectively) from the original.

The purpose of quantization of magnitudes is to suppress the effects of interference in the transmission medium.

In general, if we have n digit positions (ie we quantize to n bits) we can construct 2_n different numbers. If we need no more than 2_n different discreet magnitudes for sound transmission, complete information can be sent by a sequence of n ON or OFF pulses during each sampling interval.

By means of the use of binary levels to represent the signal, it can be regenerated to a "noise free" signal provided that the interference does not reach the threshold limit of the receiver. This of course represents a significant advantage over analog transmission techniques and means that theoretically (and to a large extent practically) the signal can be routed through an infinite number of signal processing stages without degradation.

To determine the number of quantized steps required to handle specific signals, we require a knowledge of the relationship between distortion and step size. This consideration can be regarded in two ways:-

i) Quantization of magnitude only.

ii) Combined quantization of magnitude and time.

The first part can be treated using the well known "staircase transducer". Signals impressed are sorted into voltage slices and all signals within ± half a step of the mid value of a slice are replaced at the output by the mid value.

The distortion error consists of the difference between the input and output signals. The maximum instantaneous value of distortion is half of one step (1/2 LSB) and the total pk.to pk. variation is from -1/2 LSB to +1/2 LSB. i.e 1 LSB pk.to pk. If there is a large number of small steps, the error signal resembles a series of straight lines with varying slopes. The limiting condition of closely spaced steps enables us to derive quite simply an approximate value for the mean square error. This approximation consists of calculating the mean square value of a straight line from minus half a step to plus half a step with arbitrary step.

The mean square error is therefore:-

$$MSE = p(e_n)e_n^2 \quad de_n$$

$$= Q^2/12 \tag{1}$$

Not all of the distortion falls within the signal band. The distortion may be considered to result from a modulation process consisting of the application of the component frequencies of the original signal to the non linear staircase characteristic. High order modulation products may have frequencies quite remote from those within the original signal and these can be excluded by a filter passing only the signal band.

It becomes important therefore to be able to calculate the spectrum of the error signal. This can be shown for a generalised signal using a method of correlation based on the fact that the power spectrum of a signal is the Fourier Cosine Transform of the correlation function.

QUANTIZATION EFFECTS

When an analog signal is quantized into discreet amplitudes, an error between the original signal and the quantized version is produced. This error is called quantization noise and in a digital audio system is experienced as Granular Noise.

A sufficient number of quantization levels must therefore be used such that granular effects are unoticeable to the human listener and depending on quality, somewhere between 256 and 65536 levels are required.

Quantization noise is of a wide-band nature but is amplitude dependent. Its fundamental frequency is that of the signal itself. However, during each cycle of the signal, (if the signal is periodic) the error signal undergoes 2N alternations (where N = No

of quantization levels) and thus the noise spectrum is much wider than the signal, extending well beyond the 2Nth harmonic.
With a sampled signal, the entire quantization noise spectrum is folded back into the baseband region and the total noise power is given in equation (1).

IMPAIRMENTS

The quantization stage and initial sampling considerations are quite separate. During sampling, the complex audio signal is chopped (or switched) (at a rate greater than the Niquist limit) into a series of individual samples. The number of "bits per sample" defining the resolution and therefore the potential dynamic range of the system. It can also be shown [1] that the signal to noise ratio is defined and related to the size of the sample word.
This in fact leads to an interesting fact in that the maximum output of any coder must be related to the power supply voltage. The minimum output must be related to the input noise under "no signal" conditions and therefore the maximum dynamic range of the coder is clearly defined (in the limit) by the ratio of these two parameters, regardless of the number of quantization levels.

Clearly therefore, the ratio of power supply voltage to input noise must always be greater than 2_n-1 for any coder system.

In the time domain, the original signal becomes a series of sampling points prior to quantization and provided that the input signal is band limited (by means of an anti-alias filter) to less than half the sampling rate, the original signal will always be recoverable in frequency terms. It should also be noted that because of the multiplication process that takes place between the sampling clock and the input samples, the recovered signal after D-A conversion is unlikely to look identical to the source signal but nevertheless can still appear to be (and can be) free of subjective impairments.
This is because of the fact that the Fourier Transform of both signals after multiplication indicates that the new signal contains frequency components that were not present in the source signal. It is therefore hardly surprising that the CODEC output looks different in the time domain. The low pass filter at the output of the DAC, from a spectral point of view, removes those frequencies that were created by the multiplicative action of the sampler. In fact, the unwanted frequency components are unlikely to be within the audio frequency range and therefore the only justification for an output filter is to remove (or at least significantly attenuate) the out of band components as a safeguard against interference or "beats" with other equipment.

It is interesting to note that, as described by Blesser [2] and others, time sampling can be a lossless process, whereas amplitude quantization always destroys information.

THE APPLICATION OF DITHER

The addition of random noise to an audio signal prior to quantization breaks up the contours resulting from quantization and trades granular noise for noise (ie non amplitude dependent).

Quantization distortion is more objectionable than additive random noise of the same mean squared value. To achieve minimum total error for a signal with added noise, a "noise" source identical to that added prior to A-D conversion may be subtracted from the signal after D-A conversion, with the two noise generators synchronised.

A further consideration is that with a much reduced signal level into a digital coder (ADC), what was maybe a perfectly satisfactory transfer characteristic, now resembles a course staircase. The addition of dither can therefore have a much greater effect on the quantization of low level signals. The course transfer characteristic being effectively smoothed out by noise averaging to a LINEAR AVERAGE CHARACTERISTIC.

The distortion is distributed over the whole of the channel bandwidth and the 3rd harmonic is reduced by around 25%.

This of course is extremely useful, but there is in fact another advantage in that the input signal, in the absence of dither, will never be coded when its amplitude falls below the threshold of 1 LSB. However, the application of dither to a small signal such as this, has the effect of sometimes coding the signal and sometimes not. The statistics of this situation cause signals that are below the threshold level of detection under "normal" conditions to be "lifted" above the noise floor and therefore coded.

Previous tests [3] have shown that a sinusoidal tone may be perceived whilst buried in wide-band noise where the signal is 16dB down on the noise.

It therefore follows that with a 16 bit CODEC with a theoretical dynamic range of 96dB, the dynamic range should be subjectively increased to around 112dB with the addition of dither and an associated subjective degradation of the signal to noise ratio by about 3dB.

Dither has the effect of allowing the AVERAGE value of a quantized signal to assume any value between two adjacent quantization levels, the statistical effect of the dither operating with the source signal being to vary the time that the ADC output spends at these two levels in a very similar way to a pulse width modulation system. The modulating signal in this case being:-

LIMITATIONS OF AUDIO CODECS AND THE USE OF "DITHER"

$$V_s + V_d \qquad (2)$$

-where Vs represents the time varying input signal and Vd represents the random dither signal.

BIT RATE REDUCTION

In 1962, L.G Roberts [4] published a paper in which he discussed a principle by which the number of bits required to encode a picture element could be reduced without producing the contouring effect normally associated with this action. He suggested that the application of noise to a picture signal prior to quantization would have the effect of smoothing out the steps between quantization levels such that the subjective effect called "contouring" could be traded for more noise on the picture. More recent research by th BBC has described methods for the removal of this resulting noise by making use of the frame to frame correlation that occurs between adjacent television "frames" such that when adjacent frames are added, the noise and picture signals do not add in the same way. This results in an improvement in the signal to noise ratio and in fact, with a non moving picture can result in the noise being completely eliminated.

Unfortunately, this "adjacent frame correlation" does not appear to be present with an acoustic signal. However, computer simulation has shown that the spectral choice of added noise can have a significant effect upon the subjective signal to noise ratio for a given number of bits per sample.

EXPERIMENTAL WORK AND RESULTS.

A 16 bit audio CODEC has been completed and a programmable digital PRNG is in its final stages of completion. With the addition of a simple "bit switch" subjective tests can be carried out in the proposed areas.

In parallel with the "real engineering" experiments, software is being developed to determine the effects of different spectra of dither on the output spectrum. It is also hoped to simulate (and therefore optimise) the adaptive dither generator discussed in the concluding section of this paper.

CONCLUSIONS

There is still much to be learned about the subjective effects of different noise sources when used as a dither signal in audio CODEC applications. There is some discussion as to the merits of subtracting the noise after digital to analog conversion and this is therefore also considered, the overall objective being to reduce the bit rate with less degradation than that which would normally be experienced with this action. Any time domain correlation that may be found to exist, or that can be implanted into the signal without becoming subjectively noticeable, would allow further noise reduction processing to be carried out. This would represent a major breakthrough in noise reduction techniques.

A clearer understanding of the statistical properties of audio signals is required and it is hoped that this may yield a further improvement by the use of an "adaptive, dynamic dither injection system". For example it is commonly accepted that there is a greater probability of low amplitude sound signals rather than large ones and low frequencies rather than high ones. The proposed adaptive system will therefore "adapt" to the signal as viewed in a three dimensional plane such that the dither signal will change dynamically and should always therefore be optimum.

REFERENCES

[1] B.P Lathi, "Modern Digital Communication Systems" Holt, Rinehart & Winston, 184-187, (1983)
[2] B.A Blesser, "Digitization of Audio" J.A.E.S, Vol 26, Vol. 26, no.10, 739-771, (1978).
[3] Hi-Fi News, "Review of Sony PCM-F1", October (1982).
[4] L.G Roberts, "Picture Coding Using PRN", IRE Trans. Info. Th. 8, 145-154, Feb (1962).