

# IBS Meeting

*JOHN WESTBURY MIBS reports on a presentation about Advanced Audio Coding.*

Steve Church, President of Telos Systems, gave a presentation to the IBS entitled "On Beer and Audio coding." The title came from the thought that more decisions about MPEG audio codecs were made over fine pilsners in Bavarian Biergartens than in the offices of the Fraunhofer Institut!

When MPEG 4 was being defined in the mid 1990s, Fraunhofer - along with several other companies - suggested a new coding scheme based on the principles of MPEG Layer 3 but using a 'modular' approach to avoid it becoming outdated. This new scheme was initially referred to as 'NBC' or Non-Backward Compatible MPEG coding, but subsequently went under the banner of 'Advanced Audio Coding' (AAC).

## Advanced Audio Coding

AAC is far more complex than Layer 3, with four times the number of filter bands (2048) and several new processes. These include 'Temporal Noise Shaping' (benefiting voice signals in particular), and 'Perceptual Noise Shaping' to enable finer control of quantisation resolution, thus using bits more efficiently. A new predictive model also improves coding efficiency.

**The delay characteristics for the XStream codec.**

Coding	Bitrate (kb/s)	Sampling Rate	Delay (ms)
G.722	64	48kHz	10
AAC-LD mono	64	48kHz	50
AAC-LD stereo	128	48kHz	60
AAC stereo	128	48kHz	172
Layer 2 stereo	128	48kHz	224
Layer 3 stereo	128	48kHz	326

In tests carried out by the BBC and NHK, and corroborated by the Communications Research Centre (CRC) in Canada, AAC was judged to "provide performance superior to any known codec at bit rates greater than 64kb/s." These tests also showed that AAC at 96kb/s provides quality comparable to Layer 2 at 192kb/s, and Layer 3 at 128kb/s.

The CRC also rated codecs in order of audio merit, awarding AAC first place, followed by PAC (from AT&T/Lucent), Layer 3, Dolby AC-3, and Layer 2. The highest quality was obtained from AAC operating at 128kb/s, and from AC-3 operating at 192kb/s, and the CRC concluded that AAC achieves the ITU requirement of "indistinguishable quality" at 128kb/s.

## Delays

In the broadcast industry, quality is not all we have to consider - coding delay is often a major problem. Research suggests delays of about 50ms start to concern listeners, so the designers of AAC have produced a low delay variant called AAC-LD. However, trade-offs must be made to achieve lower coding delays: spectral filters can have either low delay or good frequency resolution,

but not both; and although longer frames are more bit-efficient they can cause delay and pre-echoes too. AAC-LD overcomes these issues by changing its coding methods dynamically. For example, the 'window shape' of the spectral filter is adaptive, and the frame length changes according to the programme characteristics.

AAC-LD retains the same coding power as Layer 3 - so high quality 15kHz mono audio can be sent over one 64kb/s ISDN bearer - but has a third the delay of standard AAC. A series of tests at 64kb/s demonstrated that AAC-LD had superior performance to Layer 3 with about 50% of the test material, and was never worse, although the glockenspiel proved a tough test.

Currently only two ISDN codecs are known to offer AAC coding - the Mayah Centauri and the Telos Zephyr Xstream - but are compatible in the AAC mode. Auditioning the XStream running AAC-LD demonstrated that it certainly achieved its goals in terms of delay and quality, and could even provide "quite acceptable" stereo over one 64kb/s bearer channel! Where both quality and low delay is important - for example in a two way live interview - the XStream also supports asymmetric coding. Thus the contribution feed can employ AAC-LD over 128kb/s, while G.722 could be used for the cue feed, minimising the round loop delay.

## Spectral Band Replication

A new version of AAC called CT-aacPlus and developed by Coding Technologies in Germany, uses the modular nature of AAC combined with a new process called Spectral Band Replication (SBR). Research carried out at the Fraunhofer Institut showed that the high frequencies of typical audio signals are closely correlated to the low frequencies. Thus further coding savings can be achieved by synthesising high frequency content during decoding based on the LF information, using state-of-the-art transposition methods. In some ways this approach is similar to the 'object audio' technology of MPEG 7 - that is, sending information about the audio, rather than the audio itself! BBC tests revealed over 30% better coding efficiency in AAC+SBR (at 24kb/s), compared with standard AAC.

This SBR technology has also been applied to a derivative of Layer 3 known as MP3Pro, and due to its impressive performance at extremely low bit-rates (down to 16 kb/s), CT-aacPlus has been adopted by Digital Radio Mondial to provide near-FM performance on short-wave digital broadcasts. Comparing the performance of commercial codecs, the EBU found that CT-aacPlus came top, followed by MP3Pro, AAC, RealAudio8 (v8.5), WMA 8.5 and Layer 3, in that order.

Special thanks to Serge Auckland of PRECO for arranging for Steve Church's visit, and allowing us to enjoy his fascinating insight into the latest high-tech codecs.

**ibs**