5.1 Multi Channel Audio Contribution and Distribution - Matrix, Watermark or Discrete? APT Presents a Working Example for Each Requirement

Abstract

If radio is to maintain and improve its market share in a highly competitive and cluttered media sector, then 5.1 Multi Channel Audio needs to be embraced. Fundamental to this process will be contribution from live events (concert halls and sports stadiums) and the preservation of this content. This document outlays various solutions for broadcasters that are currently researching the use of 5.1 Multi Channel Audio.

Competition

In recent years, 5.1 Surround Sound has overtaken simple two-channel stereo as the accepted mode for audio in the Film, TV and Music industries. This increase in audio channels has been driven largely by consumer demand with most households now owning a DVD player complete with a 5.1 speaker configuration. Add to this HDTV audiences, cinema-goers and young "gamers" who grew up with the Surround Sound experience and it quickly becomes clear that that there is now a level of consumer expectation that has to be met by content providers.

Historical Review

This is an opportune time to highlight the fact that Surround Sound (in various guises) is actually a relatively old technology. The first attempt to get Surround Sound off the ground for broadcasters was a quadraphonic arrangement in the early 1970's. For the radio industry this died on the vine for a number of different reasons including the difficulty radio audiences faced in decoding content, consumer confusion, multiple formats and debates about standards (Matrix versus Discrete, which is a re-occurring theme). Arguably, and most importantly, its failure was down to a lack of content (either live or recorded) that was created primarily for radio. The failure could also have been due to the fact that Stereo was still in its infancy and the movement to multiple channel was just too much of a fundamental leap for the radio broadcast industry to deal with at that time.

The film industry took up the matrix version of quadraphonic, which became Surround Sound with Dolby Labs leading from the front. The Dolby multi-channel approach for cinema, in the Analogue domain, used Left, Centre, Right and Surround and became the standard approach, particularly after the success of Star Wars in the late 1970's. The Left, Centre and Right speakers were all placed in front of the

cinema audience and the Surround channel was reproduced by multiple speakers distributed around the rear and sides of the theatre. After that came the digital approach, which was 5.1 (Left, Centre, Right, Left Surround, Right Surround and the Low Frequency Effects channel (LFE)). Several other organisations in competition to Dolby Labs (for example DTS and Sony) offered different approaches to Surround Sound. Following the success of Surround Sound for cinema came the migration and subsequent explosion of the home theatre and audio content on DVD-A and SACD.

Why Radio Needs Surround

Now that the film industry has refined the concept, it is a simple fact that the Radio Broadcast industry needs to provide 5.1 Multi Channel Audio content. Radio broadcasters are in a highly pressurised market competing against information and entertainment mediums such as TV, Internet, DVD, MP3 players and iPods. An old two-channel programme is simply not going to meet the demands of their target audience. It may be argued that a stereo signal is abnormal. Although we have two ears, the brain has the ability to interpolate many signals simultaneously, which then creates a multi-dimensional image. In comparison, a two-channel signal will be both flat and unnatural.

Digital Radio has made the incremental step of improving quality by increasing audio bandwidth ("crystal clear CD quality") and adding ancillary services (scrolling text), but this is primarily still focussed on delivering two-channel content. To make a significant step, the next generation of digital radio services needs to offer the Surround Sound experience.

Several 5.1-for-Radio pilot projects have been completed and the war of standards is raging for a suitable transmission protocol. Although several parties are involved, the primary participants are Dolby, DTS and SRS who between them have many millions of decoders in the market place. However, the actual methodology of the transmission is an interesting topic in that it offers up the options of Discrete, Matrix and now a third alternative known as Watermarking.

Examining The Transmission Options

Discrete is self-explanatory in that a broadcaster transmits five separate channels plus the LFE – an approach that is well received amongst audio purists.

The Matrix solution has two parties - Fraunhofer and Coding Technologies - pushing this agenda. This approach advocates the use of a parametric steering channel as a side channel to an existing stereo platform and uses a perceptual coding technique to achieve the bit rate reduction. Multichannel decoders can reconstruct a surround mix from the Left/Right and Steering channel signal. This process assumes that content should be kept in discrete format up to the point of transmission.

The third alternative is watermarking, which generates a parametric steering channel upstream during the broadcast production phase. This signal encoded as a watermark and is perceptually hidden in an uncompressed

digital stereo audio signal. Neural Audio and Harris Broadcast are offering this solution to broadcasters, which should ensure that the watermark remains imperceptible after, broadcast processing, and can be correctly decoded by a Surround Sound system.

The manufacturers of digital receivers will probably take the well-trodden route of spinning-in a variety of solutions to decode Surround Sound. As such, in principle, listeners will be able to get 5.1 content into their homes (through Set Top Boxes). Hopefully, the various alternatives for transmission will be largely irrelevant to the listening audience as the Set Top Boxes will be sophisticated enough to decode content regardless of the transmission algorithms/protocols. However the home listener is only a small minority of the target audience. The segment that really gets the juices flowing for broadcasters and technology providers is the automotive industry - those people in their cars listening to the radio on their way to and from work and generally going about their daily business.

Now comes the bigger challenge: how does the broadcaster move live content from remote locations (concert halls and sporting venues), through to their studios and then out to their transmitter sites? This is one of the main issues for Surround Sound, especially in light of the fact that failing to provide surround sound content was the primary reason why Quadraphonic failed in the 1970's.

Compression, Codecs And Algorithms

One solution is to process all five audio channels and the LFE in linear PCM. For the record, a 24 bit, 48kHz sampled program requires a data rate of almost 6 Mbit/s for 5.1 channels. For most broadcasters, such an option is completely cost prohibitive and will kill any contribution and distribution projects at birth.

This brings us to our old friend digital audio data compression and the balancing act between low bit rate algorithms using perceptual coders and slightly higher bit rates that use ADPCM principles. Perceptual coders will certainly reduce the network costs but add substantial latency and run the risk of destroying the phase relationship between the individual channels. Given the loss of stereo separation caused by perceptual coders in stereo signals, this destruction is almost a certainty. It is also worth noting that the final transmission algorithm, whether DTS, Dolby or SRS, will be highly bit rate reduced and all efforts toward conserving content should be made prior to the final transmission. Many lessons should have been learnt following DAB and HD Radio, not least that the artefacts in the content was directly due to the number of perceptual coding passes in the broadcast chain.

An ADPCM-based algorithm will offer a much lower coding delay and will retain the phasing between the channels. Enhanced apt-X® from APT is generally considered by the Broadcast and Post Industry (and cinema audiences listening to films in the DTS format - apt-X® is licensed to DTS) to be relatively non-

destructive in nature. In addition Enhanced apt-X® offers an end-to-end system latency of fewer than five milliseconds making it a suitable solution for 5.1 contribution and distribution. Another option is Dolby E, which has been touted as a solution given its success in DTV. Dolby E was specifically designed to ease the transition for DTV broadcasters from two channels to Multi Channel Audio for distribution applications using existing AES3 infrastructures. However, radio broadcast does not suffer from the same constraints as DTV i.e. video frame rates. As such Dolby E has a few fundamental issues that would have to be overcome by radio broadcasters. These include a latency of over 60 milliseconds, bit polling between channels, an inability to process individual channels, a word resolution limited to approximately 22 bits and a set sampling frequency of 48 kHz. Also, and perhaps most fundamentally, Dolby E is based on perceptual coding techniques.

Heineken don't do 5.1 Codecs, but if they did:

If a broadcast Engineer could write a "wish-list" for a 5.1 codec suitable for live applications, the list may look like this:

1. Self contained unit i.e. all features available in a 19" Rackmount, preferably no bigger than 3U high.
2. 5.1 channels with 0º Phase Shift. This is an important consideration; in that there are several products on the market that claim to be Multi Channel i.e. has 6 channels. Unfortunately these channels are not phase related and depending on the power wake up mode, can have an arbitrary figure of phase relationship between 0º and 180º. This is merely Multiple Channels, rather than true Multi Channel.
3. Low delay i.e. under 5 milliseconds end-to-end.
4. Several transport options available i.e. IP (TCP & UDP), T1 / E1, V.35 / X.21, ISDN.
5. Return path for monitoring content
6. 4 Wire for Engineering talk back and possibly FXO / FXS.
7. Ancillary data i.e. RS232, Contact closure
8. 2 channel for traditional stereo applications.
9. Complete redundancy i.e. PSU (supporting both AC and DC) and Transport card.
10. Supporting various formats i.e. Compression and Linear PCM.
11. Variety of word depth (16 and 24 bit) and Sampling Frequencies (32, 44.1, 48 and 96 kHz)
12. Easy and intuitive Graphical User Interface for control

Real World Solutions

There have been a few pioneering broadcasters who have put to air 5.1 pilot projects using live source material. In most cases the live material has been either been Classical or Jazz content. Interestingly, these projects are happening in parallel on both sides of the Atlantic – at ORF in Austria and NPR in the USA being two examples. Unsurprisingly, the Europeans and the Americans have explored two

fundamentally different techniques, but key to both was the use of Enhanced apt-X®.

TCP/IP

ORF, under the guidance of Karl Petermichl (and closely observed by his fellow EBU members) chose Enhanced apt-X™ wrapped up in the WorldNet SkyLink. This unit is a codec with eight discrete channels (5.1 and a Stereo pair) and it uses an Ethernet port to present the compressed data to the outside world. The WorldNet SkyLink units were used for projects that included "Night of the Long Radio" and ORF's New Years Eve broadcast. The discrete channel approach enabled ORF to process individual channels and keep to a minimum the amount of hardware used in the broadcast chain. Data capacity was provided by Austria Telecom, which supplied a 2Mbit/s ADSL circuit.

E1 / T1

It is worth outlining that both WDR and BR in Germany are also using Enhanced apt-X® for 5.1 but are using an E1 interface (for these projects the broadcasters used the APT WorldNet Oslo). The actual transport medium, Synchronous or IP, would appear to be a decision based on what service the local Telco provides.

Required bit rates over a synchronous T1 or E1 circuit will depend on the word depth and Sampling Frequency. Assuming that the Sampling Frequency is 48kHz, the figures below related to word depths of 16 bit and 24 bit:

- Bit Rate for 5.1 @ 16 bit will require – 5 x 192kBit/s + 64kBit/s = 1024kBit/s.
- Bit Rate for 5.1 @ 24 bit will require – 5 x 256kBit/s + 64kBit/s = 1344kBit/s.

Note: the 64kBit/s bit rate is for the Low Frequency Effect channel. This will provide an audio bandwidth from 0Hz to 7Khz. The Enhanced apt-X algorithm uses 4 sub bands and allocates bits to each sub band. The lowest sub band i.e. 0Hz to 1750Hz, will have either 7 or 9 bits allocated (depending on the word depth i.e. 16 bit or 24 bit) as this is where the majority of energy will be focussed for a .1 channel.

Over an IP network, the required bit rates are approximately the same as synchronous networks but it is prudent to add an additional 10% for overhead.

ISDN

In the USA, NPR, directed by Mike Papas, has put to air a number of broadcasts including a Diane Reeves concert and the "Toast of the Nation" New Years Eve event in December 2004 and the broadcast from New Orleans in 2005.

Papas used the fundamentally different approach of Watermarking for moving Surround Sound content. In this case, at the concert site Mike and his Production team down mixed the 5.1 channels to 2 channels using the Neural Audio 5225 system. He then transported the two channels using ISDN codecs. The ISDN codec in question was the APT WorldNet Tokyo, which bonds together 4

x ISDN lines to create a 512kBit/s data pipe and uses Enhanced apt-X™ at 24 bit word resolution, 48kHz sampling frequency. As Enhanced apt-X® is a relatively non-destructive algorithm with a gentle compression ratio, this removed the incidence of any artefacts affecting the down mix (something that an MPEG algorithm could not achieve in an ISDN codec). At the receiving end, NPR reconstituted the two channels back up to a 5.1 signal.

The results of pilot projects undertaken by these pioneering broadcasters are likely to shape and influence the decisions taken when large-scale deployment of 5.1 multi-channel broadcasting begins in earnest. While ORF and NPR adopted two fundamentally alternative approaches using different equipment and transport mediums (and APT does not claim that one approach is superior to another), the constant in both scenarios was the choice of coding algorithm. All parties agreed that only the APT products and the Enhanced apt-X® algorithm, provided the right balance of network efficiency, latency and flawless audio quality. Regardless of the coding or production techniques used, the radio industry can now put to air live 5.1 material, which should offer a compelling reason for listeners to stay with radio as an entertainment and information medium.

Summary

With the volume of pre-recorded content currently being produced in 5.1, radio broadcasters have sufficient material to bring a Multi Channel 24/7 channel to air. However, to truly add the "WOW" factor, the live events need attention. This is now possible and a variety of options are available to transport content and importantly to preserve the acoustic integrity of the original material.

Acknowledgements

Mike Pappas, KUVO
Karl Petermichl, ORF