5.1 Components in a 2.0 Broadcast Chain: A Brief Overview

Robert Reams, Neural Audio

The Big Picture

It is obvious that the interspersion of legacy 2.0 and 5.1 content is a reality. Unless this is dealt with on a system basis, the result will be less than transition proof. In fact, inability to successfully integrate legacy content with "modern" content in such a way that meets the consumers expectation is unfortunately naive and slows the adoption of 5.1.

Matrix methodologies do not satisfy the distributor's and consumer's expectation of what is now called "discrete" 5.1.

Neural's spatial compression and rendering methodology allows the distributor or broadcaster the ability to capture original source 5.1 content and "downmix" it to a 2.0 channel format while allowing the consumer to render the content in any spatial format they choose; mono, stereo, 3 channel stereo, quad, 5.1 or whatever.

In the absence of spatial rendering (which will be the case during transition toward 2.0/5.1 interoperability) the content is received and perceived as "stereo" on existing receivers.

Capture and Encoding

5.1 content is downmixed into two fine structures. The 5.1 image envelope is transformed to a two dimensional version of the original 5.1 image envelope. The azimuth of the original image envelope is represented by inter-channel intensity differences. The front-to-back, or "depth", of the original image envelope is represented by average inter-channel coherence (normalized cross-correlation).

The 2-D image envelope of the 5.1 content is imbedded in the two downmixed audio channels in the form of watermarking (patent pending). Intensity/coherence watermarking is an excellent choice because of its similarity to the image construct of naturally occurring 2-D stereo and compatibility with already prevalent Lt/Rt content.

Decoding and Rendering

Upon decoding, the image envelope of the original 5.1 content is re-synthesized based on the

intensity/coherence information contained in the watermark of the two fine structures. Using this methodology, an impression of the original source 5.1 content is rendered from the two downmixed audio channels with a high degree of merit.

The 5.1 rendering is accomplished by a (patent pending) programmable, transform based spatial rendering system. SEE (Spatial Environment Engine) can render any two dimensional audio source (both 5.1 and stereo are 2-D) into as few as 2 to as many as 256 outputs with a high degree of perceived separation.

The spatial elements of 2.0 stereo (or Lt/Rt) are segregated based on the 2-D image envelope naturally residing in the content, nothing is either created or destroyed. "Re-downmixing" of the 5.1 rendering of stereo back to 2.0 stereo (which happens quite often in the HDTV industry) results in "near perfect reconstruction" of the original stereo content with the stereo image completely intact.

Production/Editing

All downmixing and rendering duties of the broadcaster are handled by the Harris/Neural 5225 surround production appliance.

Existing stereo editing systems will work just fine with 5225 downmixed 5.1 because it's not a bitstream (unlike certain other methodologies).

Neural spatial compression survives editing and even conversion to analog. After 5.1 content elements are downmixed to 2.0 they may be stored in the stereo archive (server) where they may be imported into the stereo editor and treated like stereo elements. Cutting, pasting and mixing of the downmixed 5.1 elements with stereo 2.0 elements is allowed.

Downmixed 5.1 elements may be loaded, stored or recorded to any cart, sampler or rec/rep.

The 5225 works well with 5.1 editors. The editor 5.1 outputs may be routed directly into the 5225's downmix inputs. The 2.0 downmix output is routed to the 5225 spatial rendering input as well as the 2.0 track on the editor. The session may be monitored in real time through the 5225. The session 2.0 file may be dumped onto the server afterwards.

Monitoring

Encoded content may be broadcast, stored, loaded, sampled, recorded or distributed through existing 2.0 infrastructures. It may be rendered to 5.1 at any point in the broadcast chain for production QC or broadcast "confidence monitoring".

5.1/2.0 mixes may be tested for 5.1 compatibility at any time using the 5225's 2.0 to 5.1 rendering facilities.

Content may be tested anywhere from post capture to post receiver using the 5225 rendering

facilities.

Session work may be monitored in realtime through the 5225 while tracking.

Storage, Archiving and the Server

After capture, downmixed 5.1 content may be stored in the existing server (compressed or not) and treated as any other stereo content.

Broadcast Processing

Mono, stereo or 5.1 content of unknown pedigree can cause trouble with broadcast codecs if not properly treated. Image packing and noise reduction are tantamount to maximizing the performance of today's perceptual coders.

Level and spectral (loudness and signature sound) processing are performed commonly on both 2.0 and downmixed 5.1 content as both are represented in 2-D with two channels.

The stereo and interspersed 5.1 content are 100% compatible with the Harris/Neural Neustar and UltraLink codec conditioners. Placing the Neustar between the content "play-out" and the IBOC exciter results in HDC compatible content that is controlled, consistent and "renderable" to a 2.0 or 5.1 spatial environment.

Benefits

The methodology demonstrates a high degree of immunity to lossy compression, D/A/D and A/D/A conversion. It may be stored or broadcast in either digital or analog.

To broadcasters, this level of backward compatibility could be a godsend as most existing infrastructures, including production and storage, are 2.0. This greatly simplifies the inevitable integration of 2.0 and 5.1 content on both the broadcaster and consumer sides.

The use of intensity/coherence watermarking allows the spatial codec to dedicate significantly higher bits/sec to the fine structures. At ultra low data rates (48kB/s) this would increase the number of bits available for the fine structure by 33% as there is no side data to transport. There is no argument that the AAC codec and its derivatives "sound better" at 64kB/s vs. the 48kB/s or 48kB/s vs. the 32kB/s that would be allowed if the (16kB/s of) side information approach is used. In the realm of lossy compression for broadcast 16kB/s is a lot of bits to use for data other than the fine structure of the audio.

Automotive adoption

Automotive engineers are adopting a consumer version of the Neural's 2-D downmix/SEE rendering technology as the automotive audio infrastructure is also 2.0. SEE drastically reduces

the cost of implementing 5.1, thus driving (pun intended) the ubiquity of the automotive 5.1 system. Any 5225/Neustar treated content that is introduced to the SEE based automotive audio infrastructure will be rendered in “5.1". Original sourced 5.1 will be rendered as the original source 5.1 and stereo will be rendered in a 5.1 spatial environment.

Conclusion

Wide spread acceptance of 5.1 content isn’t as far away as once thought. 2.0 and 5.1 content interoperability doesn’t have to be scary, expensive or dangerous. Modern digital signal processing can modernize “work horse” infrastructures every bit as effectively as replace them.

The key is planned, staged transitioning from where the broadcaster is now to where the broadcaster wants to be at a rate of adoption that makes tactical and fiscal sense. The rate of transition should parallel the efforts of the content providers (record labels and artists) as well as manufacturers of home and automotive audio systems.