Upmixing—Stereo to Surround

New Demands and Requirements for a Modern Upmix Process

As NUGEN Audio prepares to launch the NUGEN Audio Halo Upmix Plug-in, a new upmixer for the post-production market at IBC 2015, the company's Creative Director and Co-Founder Jon Schorah reflects on the changing landscape and increasingly complex demands of the modern audio production environment.

By
Jon Schorah
(United Kingdom)

The notion of upmixing audio from stereo to surround formats is not a new one. For many years, cinematic audio has been routinely delivered as a 7.1 audio mix. Where necessary, upmixing has been used to generate audio from original stereo material in a suitable format for the surround environment.

More recently, the demand for original TV programming in surround formats has led to a widespread use of upmixing in post-production. Even more recently, we've seen the emergence of object audio (e.g., Dolby Atmos), which makes use of 9.1 audio beds and promises to draw new attention to the subject of upmixing.

A movie theater is a relatively predicable environment, acoustically treated and relying on a playback system properly calibrated for 7.1 playback. However, the home listening environment has no such constraints; additionally, the delivered 5.1 surround mix may be automatically down-mixed into stereo by the consumer set-top box (STB). Object audio is also subject to these challenges. Although object audio mixes are largely destined for the movie theater, there is no reason that these same mixes cannot be re-interpreted for home listeners. Since object audio is theoretically infinitely scalable, we can join the movie theater and the home TV experience together via a single translatable mix.

Good Upmix Source

For an upmix process to be able to perform well, beyond the traditional 7.1 cinematic presentation, several new requirements come into play.

First and foremost, the process must be able to provide a good upmix from the original source material. Of course, "good" is somewhat subjective, but two seemingly conflicting criteria often present themselves depending upon the purpose of the upmixed audio. One is a requirement for the upmix to respect the original source material, providing

a naturally extended panorama that "unfolds" the original into a surround context. This is important when the original source is a broadcast mix that is simply being translated for broadcast in surround. In this situation, creative decisions have been made and these preferences need to be respected and represented in the upmix produced. The other situation involves upmixing of original stereo material as part of a more complex arrangement being produced in surround. In this case, getting the upmix to fit and contribute to the overall sound of the surround mix is as important, if not more, than maintaining a direct relationship with the original source audio.

Flexible Process

The upmix process also needs to be flexible. Traditionally, the majority of cinematic releases have used a 7.1 configuration, with a 5.1 configuration for TV. For film audio engineers, a 7.1 upmix solution is perfectly adequate. However, many freelance engineers routinely offer services to both the film and television industries and, therefore, need a certain level of flexibility. For those with an eye on the future, the advent of object audio offers an enticing 9.1 bed track, including additional information for overhead loudspeaker arrays. According to one manufacturer's technical white paper—in theory, an object-based mix, including 9.1 beds, can be automatically translated to 7.1, 5.1, and stereo formats allowing the audio professional to "author once, optimize everywhere." This has the potential of introducing a universally translatable mix, but it remains to be seen how and if object audio will be established in practice.

Dialog Isolation

In the realm of restoration and re-imagining of old recordings and upmixing for television broadcast, there is an additional concern: dialog isolation. Most delivery specifications require a high degree of dialog isolation into the center channel for aesthetic purposes and delivery consistency. Recent increases in computing power have led to great improvements in processes that facilitate this isolation. Techniques that go far beyond simple mid-side and frequency isolation are now possible, with minimal over-isolation and transparent transmission of directional content passed to the sides.

Downmix Compatible

In addition, an upmix must, in most cases, be highly downmix-compatible. This means that a downmixed version of the upmixed audio (the upmix collapsed back down to stereo) must still sound very close to the original. This is important if the upmixed audio is ever to be repurposed in a different format (e.g., a feature film repurposed for 5.1 and/or stereo TV in which downmixes are commonly used to produce the broadcast audio). There are many upmix processes that can be used to generate a full-sounding upmix. However, some of these can cause serious downmix compatibility issues. For original TV production, the 5.1 downmix needs to correlate closely to the original stereo source as it is highly likely that the downmix will be what the consumer eventually hears. Even in the context of a new full-surround mix, where the downmix will not be directly related to a single upmixed audio source, it is still an important requirement to deliver a down-mixed track free from unnatural sounding delay and phase artefacts.

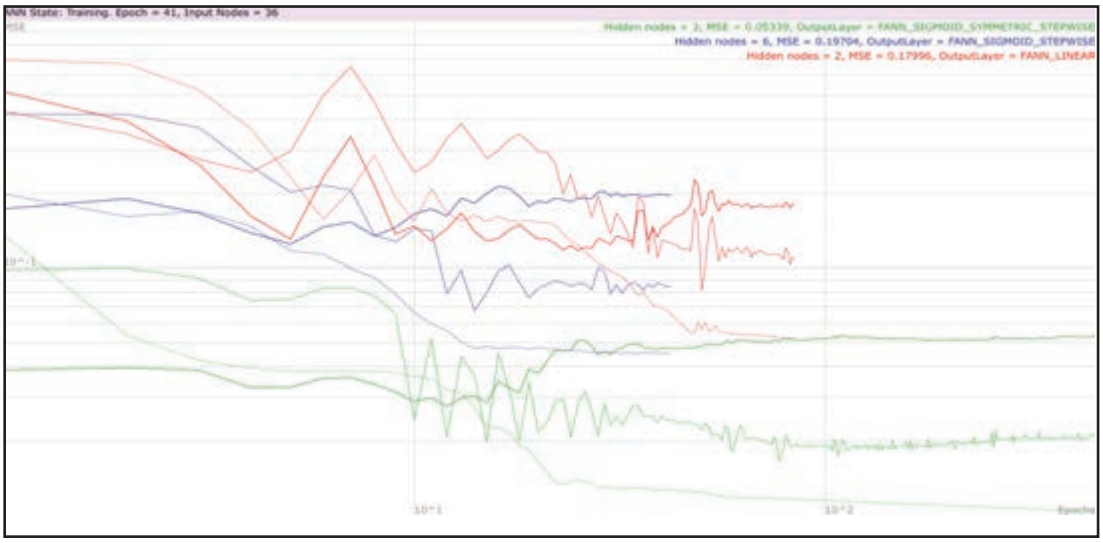
Loudness

Finally, there is the consideration of loudness. International standards now regulate the loudness of audio delivery for television across the world, and cinema sound is moving toward similar recommended practices. This raises a potential

About the Author

Jon Schorah is the creative director and co-founder of NUGEN Audio, one of the world's leading manufacturers of loudness products. Jon has a background in mastering and engineering and has considerable experience in wider aspects of the industry. A 1992 Leeds University (UK) graduate, in recent years Jon has focused on product design with a particular interest in the usability and workflow aspects of audio software.

NUGEN Audio is preparing to release its newest plug-in the NUGEN Audio Halo Upmix plug-in.



Halo Upmix uses a neural network artificial intelligence algorithm to extract the dialog portion from a piece of audio to allow isolation in the center channel, or in left-center-right (LCR). The neural network used was produced by running a genetic algorithm on trained neural networks. The graph shows some competing neural networks' test scores as they are trained using backward propagation.

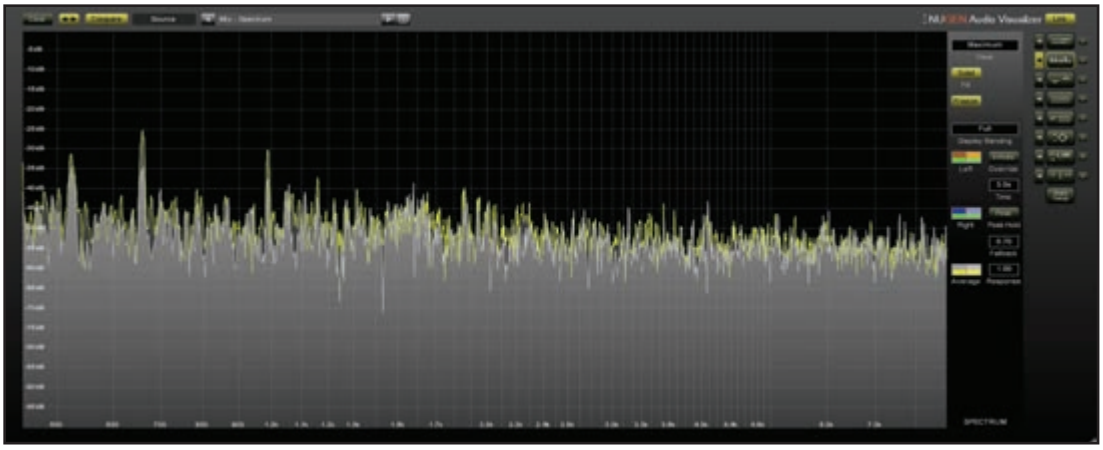
problem, since a 7.1 original mix, a 5.1 TV reversion, and a stereo TV downmix of the same audio may not all have the same loudness. If each mix could be discretely delivered this would not be a major issue because each mix could be normalized to loudness compliance with minor global offsets. However, this is not standard practice.

For the majority of home listeners, as we've noted, the 5.1 mix is delivered to the consumer's STB where it is automatically downmixed into stereo. To make matters worse, the difference in loudness between the surround and downmix versions is not a constant value from one program to another. Nor is the difference in loudness always in the same direction, it can be louder or quieter for a number of different reasons. Plus, the high degree of center channel dialog isolation often demanded for television broadcast can further exacerbate this loudness disparity.

With a possible difference of several loudness units (LU) in either direction, in some situations we might see a re-introduction of loudness jumps from one program to the next simply due to differing loudness in the consumer downmix. Gaining an understanding of this difference in loudness during the upmix process is of critical importance if delivery specifications are to be met, yet little research or recommended practice is currently available on the matter.

Upmix Seen in Context

In summary, as much as any other aspect of audio production, upmix requirements now need to be seen in the context of the "digital content everywhere" consumer revolution. Gone are the days of isolated production for a single playout scenario. Cinematic releases are routinely repurposed for television and premium television content broadcast in 5.1 is routinely downmixed into stereo, then further repurposed for mobile listening. An understanding of these different processes and how the original mix will translate is an important part of the upmix decision if creative intent and consumer satisfaction are to be preserved throughout the distribution chain.



Here a stereo downmix is compared with the original stereo source.