audioXpress
Tube, Solid State, Loudspeaker Technology

AES Multi-Tracks in New York 2007, Pt. 1

By David J. Weinberg

Here are some highlights from last fall's AES Convention in New York City.

Almost 21,000 professionals invaded New York's Jacob Javits Convention Center for the 123rd Audio Engineering Society convention. Attendees had to choose from a diversity of topics in more than 150 sessions including paper presentations, panels, tutorials, workshops, master classes, and special events on broadcasting, education, pro-audio history, live sound, architectural acoustics, and hearing. Coincident was the Broadcast Conference—co-presented by the AES and Dolby—which covered the entire broadcast chain from facility design to the receiver. Almost 500 companies had exhibits and demos. There was far too much for me to experience it all.

LIVE SOUND SYMPOSIUM: SURROUND LIVE 5

Fred Ampel (Technology Visions; www.TechnologyVisions.com) produced and chaired the fifth Surround Live event, with major support from the Sports Video Group and sponsorship from Sennheiser, Neural Audio, Harris, Digico, Beyer, and Klein & Hummel (which provided the loudspeaker systems). Strictly speaking, Surround Live is not part of the formal AES convention, but is an official symposium event that is always held in conjunction with the convention. Surround Live 5 (SL5) drew well in excess of 100 attendees. A surround sound system was set up around the seating area for presenters' demonstrations. Ampel used Audio Control's Iasys for initial calibration, which shortened the process.

Due to scheduling conflicts, the event was held in the Broad Street Ballroom, within a block of the New York Stock Exchange, instead of its usual location within a short walking distance from the convention center. The ballroom is bare marble and concrete with acoustics like the world's largest bathroom! Ampel had heavy drapes hung around the seating area, which helped a great deal and enabled most of the demonstrated effects to be heard.

FEATURED SPEAKERS

Kurt Graffy (Arup Acoustics; www.Arup.com), SL5's keynote speaker, talked about how you determine a sound's location—the inter-aural level difference (>1.5kHz) and inter-aural time difference (<0.5kHz; a 1.5ms time difference moves a sound from centered to one side), the head-related transfer function (including the pinnae effect), the ISO-226 equal loudness curves (which look like the Stephens curves, slightly different from the Fletcher-Munson curves), masking, cochlear-cortical filters, and other physical and psychoacoustic characteristics. Human hearing is nonlinear with frequency, time, level, and angle (vertical and horizontal). Factors that you perceive and affect what you hear include depth, width, distance, direction, balance, timbre, dynamics, and direct-versus-reverberant sound. There are substantial environmental distortions to dimensionality including room reflections, noise, reverberation time, room modes, speaker location(s), and listener(s) location(s). His overall message was that hearing is complex, and you make many pro-audio decisions based on simplified models.

Jim Hilson (Dolby Labs; www.Dolby.com) addressed the pressing issue of levels and Dolby Digital metadata. DialNorm is a parameter to normalize the dialog level to -31dBFS. Dolby Digital playback (decoding) will drop the level from the number set in metadata to -31dBFS; thus if DialNorm is set to –27, program level will be lowered 4dB. He mentioned three measurement periods (a 10-second average, a 10-minute average, and a full program average), emphasizing that the program should be of normal dialog, not shouting or whispering, and should not include other loud effects or music, which lead to an incorrect value. DialNorm decoding is mandatory and always active by default in all Dolby Digital decoders (some decoders offer a menu choice to turn it off). For the full dynamic range experience it is critical that only one DialNorm value be set for the entire program, as is the case for almost all movies.

Tom Sahara (senior director of remote operations, TNT Sports, Turner Networks; www.Turner.com) spoke about the difficulties in getting good live-event sound for broadcast. Problems include mike selection and placement (for example, miking the track at NASCAR races involves consideration of levels, dirt, and possible rain), a contract sound-mixer monitoring in a remote truck he's not familiar with, while keeping the mix balanced for surround, stereo, mono, and playback systems from movie/home theaters to hand-held devices.

His fundamental premise is to keep it simple. He picks the viewer's sound field perspective, plus how fast and often it changes—should the viewer's sound field match the image for the in-car camera as well as various close-up and far-field shots; should the sound transitions be hard-cut or cross-fades, and so on. The surround channels' sounds must be carefully chosen to not distract the viewer from the screen and to ensure stereo and mono mixdown compatibility, particularly because most viewers use stereo or mono playback. Mike Pappas (KUVO-FM) emphasized that point, saying that at least 30% of his audience listens on small radios in mono and that he is careful to ensure they get good sound.

Sahara pointed out that at each step in production-through-playback there are many potential causes of sound problems. He strongly recommends—and practices—staying away from the "this is how we've always done it" approach and enter each situation by gaining a thorough understanding of how the considered technology works, what it can and can't do, and how it can most effectively be used to produce the sonic reality sought. Program creators are challenged to deliver sound that is compatible with mono, stereo, and poorly set-up surround systems, yet produce a sound that blows the minds of those with great, properly calibrated playback systems.

Intelligibility is the most important characteristic in design and delivery of broadcast sound. Sahara has determined that the announcer should have a maximum average level around +12dBA (slow-averaged) over the effects, and the surround levels should be –6dBA compared with the effects. If an LFE is used, its level should be even with the effects' level. He ended his presentation with an impressive video excerpt from a NASCAR race, with the image of the cars, and their sound, coming toward the camera and receding off to the right-rear.

Michael Nunan (CTV Television; www.CTV.ca) has long said that in a surround system the front channels are not merely stereo-plus-centered-dialog, but should be used to produce a sound field continuum across the screen. His production manager complained that while watching hockey, football, or even ice skating, the announcers are center-front even when they're not on the screen—their disembodied voices interfering with the front sound field. Nunan produced a demo of sporting-event sound fields with two announcers, one in the left-surround and the other in the right-surround. With the right levels, and properly set-up surround playback, there was no problem understanding the announcers while more fully experiencing the front sound field. The announcers being behind was not distracting from the event; in fact, it allowed better integration of the image and sound field. Obviously there must be consideration for badly set-up surround systems, plus stereo and mono mixdowns, but I found the demonstration thought-provoking.

Other presenters were Fred Aldous (Fox Sports), Randy Conrod (Harris), and Mike Pappas (KUVO radio). The panel discussion "Surround Operational Issues: What Happens When It Leaves the Truck?" was moderated by Ken Kerschbaumer (Sports Video Group), with Ron Scalise (ESPN Remote Operations Audio Project Manager), Jim Starzynski (principal audio architect for NBC/Universal Advanced Engineering Growth Center), Robert P. Seidel (VP Engineering & Advanced Technology, CBS Television Network), Bruce Goldfeder (Director of Engineering, CBS Sports), and Randy Conrod.

Beyerdynamic (www.northern-america.beyerdynamic.com) demonstrated its Headzone® earphones that take a digital audio input and reproduce a stereo or surround sound source with "out-of-head localization of the center channel" and offer limited room size and acoustics adjustment, plus head tracking to keep the image at the "stage." These are fine if you don't need a speaker-reference, but the Smyth Research system should be a better fit for location mixers who must work in an unfamiliar environment, such as a remote location truck at an NFL game.

Surround for live events is challenging, but when done well, involves the viewer much more than traditional sound approaches and adds lots of fun and a sense of reality to the program. Surround Live 6 will coincide with the 125th AES convention in San Francisco next fall.

SAMPLING THE CONVENTION

Perception

"Room Reflections Misunderstood?" (AES preprint 7162) Siegfried Linkwitz (Linkwitz Lab; www.LinkwitzLab.com) evaluated a dipolar versus a monopolar speaker for "perceived differences in their reproduction of acoustic events" in a domestic living room. His criteria were timbre, phantom image placement, and sound stage width. He claims they sound similar even though their frequency responses and sound fields are significantly different, plus the effect is "further enhanced by extending the dipole behavior to frequencies above 1.4kHz." The monopole speaker is omnidirectional below about 3kHz.

Linkwitz noted six "impediments to creating a realistic impression of an acoustic event:" non-uniform polar response, inadequate dynamic range (not able to reproduce realistic levels), speakers too close together and asymmetrically placed, room treatment that changes the spectral balance of reflected sound, room equalization above low frequencies, and recordings with too many microphones in separate sonic spaces. He seems recently to have realized that off-axis frequency response needs to closely match the on-axis response—a fundamental truth that is decades old to many.

Linkwitz initially tunes his systems with test equipment, but performs fine tuning by ear, with source material he knows well.

SIGNAL PROCESSING FOR ROOM CORRECTION

"A Low-Complexity Perceptually Tuned Room Correction System." (AES preprint 7263) The instigation behind this effort by James D. Johnston and Serge Smirnov (Microsoft; www.Microsoft.com) was that in many rooms speakers are not identical, locations are not ideal (asymmetrical and at various distances), and room acoustics are antagonistic. This results in non-optimal reproduction. Johnston explained their "room-correction algorithm that restores imaging characteristics, equalizes the first-attack frequency response of the loudspeakers [in the critical bandwidth, the first 6-8ms is the most important for room EQ, even at higher frequencies], and substantially improves the listeners' experience by using relatively simple render-side DSP in combination with a sophisticated room analysis engine that is expressly designed to capture room characteristics that are important for stereo imaging and timbre correction."

BROADCAST SESSIONS

"Audio for HDTV: DialNorm." This panel discussion involved chair Andy Butler, Mike Babbitt (Dolby Labs), Tim Carroll (Linear Acoustic; www.LinearAcoustic.com), and Robert Seidel (CBS). Seidel said that an acceptable Leq(A) needs to stay within about a 7.5dB range to not be bothersome to listeners. CBS receives many program files with widely varying levels and incorrectly set DialNorm, and has had many instances of a quiet program segment just before a loud commercial, causing an unacceptable jump in measured and perceived level (sometimes this is unavoidable because the program's level, while OK overall, happens to have a quiet segment just before a commercial break). Because of these problems, CBS sets the DialNorm metadata to –31, effectively turning off the volume leveling function in all receivers.

Turning off the DialNorm function results in the widest variation in playback levels possible. Steve Lyman (Dolby Labs) told me that the DialNorm and DynRng metadata parameters must be processed by all Dolby Digital decoders, and that the default in consumer equipment is almost universally ON, attempting to level the inter-program volume and compress the dynamic range of the soundtrack. Because of the nonlinear transfer function in the playback processing, turning the DialNorm function off by setting it to –31 exacerbates the level-change problem.

The author continues his look at the extensive offerings at the 123rd AES Convention in next month's issue. *aX*