SOUND AND FURY:
BRINGING DOLBY ATMOS TO THE NHRA

Michael Babbitt
Dolby Laboratories
San Francisco, CA, USA
mgb@dolby.com

ABSTRACT

The National Hot Rod Association (NHRA) approached Dolby Laboratories in early 2017 to explore what value immersive audio might bring to the NHRA's television broadcasts. After site visits, audio capture, video production analysis and production tests, a hybrid production strategy leveraging conventional live and cinematic post production methodologies using both the static placement of objects and the dynamic panning of audio objects in the immersive space and in real time was employed. This approach took advantage of the relatively predictable nature of each NHRA drag race to enable the use of live, dynamically panned audio objects combined with statically panned objects like ambience and crowd paired with substantial use of audio scene changes. This hybrid live/post approach to the creation of live Dolby Atmos immersive audio experiences could be used as a model for other sports, and exposes opportunities for a measure of automated audio production control through the capture and use of telemetry data from spider-cams, jib-cams and the participants themselves.

Understanding the Challenges

The Importance of Audio to the NHRA

The island launch pad during an NHRA top fuel Funny Car drag race

The NHRA produces 24 live events each season held at race tracks around the United States during a season that lasts roughly 10 months. Each three-day live event (Friday through Sunday) involves several rounds of qualifying and elimination races and attracts thousands of fans over the course of the weekend, who come to enjoy what some call a "full body experience." One of the main highlights for these race fans are the "Top Fuel" vehicle classification that burn a concoction known as nitromethane, and are separated into two categories: "Funny Cars" and "Top Fuel Dragsters." These supercharged vehicles output more than 10,000 horsepower each, which translates into a sound pressure level of more than 160dB. In fact, in a 2007 experiment with a single 7,000 horsepower vehicle (state of the art at that time), the launch of the vehicle registered 2.3 on the Richter Scale[1].

[1] http://www.espn.com/espn/news/story?id=5759488

NHRA fans at these events crowd around the retaining fence at the vehicle staging and launch area, about 50 feet from the starting line, where the cars perform a routine which includes a tire burn-out and last minute mechanical adjustments prior to the line-up and launch. When the cars launch down the track, the fans experience a physical shock strong enough to knock them back and even distort their vision. Fans find this sensation to be addictive.

Each NHRA event is broadcast either live or same-day time-shifted in the US mainly on FOX Sports 1 (FS1), a channel delivered via subscription services like cable, satellite or IP, but also occasionally on the main FOX flagship broadcast channel. In 2015, the NHRA live production team transitioned from stereo program production to 5.1. Still, the televised product was no match for the live experience to which drag race fans had come to expect, and the NHRA looked to Dolby to help them create a more representative televised experience that better matched the live experience, with the expressed intention of growing their fan base.

Immersive Audio Opportunities for the NHRA

The full body experience of the NHRA's live events drove NHRA leadership to consider the relatively new concept of immersive audio as a possible solution to their intention of bringing the live experience to the home, and approached Dolby to discuss an approach leveraging immersive audio using Dolby Atmos.

Immersive audio approaches in general enhance typical multichannel, 5.1 surround sound by adding more speaker channels, some placed overhead or at different heights, thereby "immersing" the viewer in audio and creating the opportunity to relay a more lifelike presentation in the home. Dolby Atmos' immersive approach for live television production is to add up to four new overhead channels to the main 5.1 mix or "5.1 bed," creating a configuration known as "5.1.4."

In most stadium sports, the overhead register will typically contain general stadium sounds like crowd ambience, public address, referee mics, etc. This will generally be treated as ambience, and the audio elements are intended to convey the excitement of the crowd, reflecting the action on the field. This approach brings a sense of being present at the event and in the stands, but is predominantly static and contains no motion or panning other than that captured by individual microphones statically panned to specific channels. This approach is currently quite typical for live immersive audio events due to the unpredictability of ball sports and the speed of most motor sports, resulting in few if any opportunities to manually add motion to the audio program. Whether this approach would work for the NHRA was initially unclear, so arrangements were made to capture the live, isolated microphone splits at an NHRA event and bring those isolated element captures back to Dolby Laboratories for analysis and remixing on a post-production dub stage to determine the best approach and the one that would be the most realistically achievable for these uniquely challenging events.

It quickly became apparent that the typical ambience approach used for stadium sports mentioned above and often captured using multi-capsule or steerable microphones would not work for the NHRA. The cars were incredibly loud, which required the mic gains to be set quite low, making them relatively useless for the capture of crowd reactions or other generic elements that would be found in an ambience bed. In an experiment during the initial test capture effort, an RF mic was placed roughly 0.5 miles from the launch point and still required the mic preamp gain to be set exceedingly low to prevent clipping.

During post-production of the NHRA event capture at Dolby Laboratories, what became clear was that NHRA drag races are quite different from typical ball sports and other motor sports. Those differences include:

- Races involve two vehicles at a time (called “2-Wide”) or at some specialty track venues, four vehicles (known as “4-Wide”).
- Except for emergency situations where the vehicles explode in a fireball (which happens quite often, by the way), the races themselves are highly predictable and involve six main components in a sort of “recipe” for each race:

A Funny Car explodes (courtesy NHRA)

  1. An engine starting procedure known as “lighting up”
  2. A “burn-out,” where the drivers smoke the tires to heat them up in an effort to increase traction (this process also results in a significant amount of toxic vapor and hot, flying debris)
  3. The roll-back of the car to behind the starting line after the burn-out
  4. Last minute adjustments to the vehicle’s engine
  5. The roll up and staging of the vehicle at the starting line
  6. The races themselves, which generally lasts less than 4 seconds and take the same straight line path from A to B

Due to the brief, roughly 4 second duration of the actual races themselves where the vehicles frequently exceed 300mph over the course of only 1000 feet, the staging and launch of the vehicles, which can take about two minutes or so, provided the best opportunity to leverage an immersive experience and to relay the sense of anticipation experienced by the fans at the live events. The staging and launch of the vehicles is where the team concentrated their efforts at replicating a live immersive experience.

During the staging sequence, interesting opportunities for a more unique live immersive experience leveraging motion within the audio playback environment were identified during post-production. One of the featured cameras at the starting line is a boom or “jib” camera. This camera is able to swoop over the cars while they are within the launch area, during the first five steps listed above. The combination of a relatively predictable sequence of events prior to the launch of the vehicles and the motion of the jib-cam presented an opportunity to bring a post-produced-like, cinematic approach to creating the Dolby Atmos experience to the televised product where the live audio mixer could, with close cooperation from the program director and the video team overall, have an opportunity to manually pan an audio element across the 5.1 audio bed, upwards into the overhead height speakers and back, thereby enhancing the visible motion of the jib-cam.

The post-production effort also seemed to suggest that while we were able to closely match the movements of the jib-cam in post on the dub stage, this didn’t appear to be strictly necessary as the combination of video and audio movement seemed to be resolved within the viewer’s brain and didn’t result in any clearly objectionable results. The correlation of audio and video movement could be an interesting area for continued research using a statistically significant sample size, however we moved forward with the intention of pursuing best effort audio panning correlation with the jib-cam’s movement, in large part due to the

difficulty that a live production mixer would have in manually achieving tight correlation with the jib-cam's movement.

Another opportunity involved two cameras mounted on the track, just past the starting light cluster, aka the "Christmas Tree." These cameras conveyed an up-track and down-track perspective of the vehicles during the burn-outs and the roll back to the starting line afterwards. Additionally, the camera facing up-track (towards the starting line) is often used as a launch camera. Since the cameras were placed directly on the track surface, the immersive opportunity was such that the sound from the vehicles could be made to literally pass over the viewer either front to back or back to front, depending on the camera perspective. This effect could reasonably be achieved through static signal panning and not require manual manipulation of a pan control. Even so, the perspective change from the car moving front-to-back and back-to-front when the director called one or both of the cameras in sequence would require an audio scene change.

The NHRA demo footage which was captured live at the event and posted on a Dolby sound stage, was presented to the NHRA leadership, who found the result to be extremely compelling, differentiating and overall the closest match to the spectral density and impact of the live experience, with the additional, unique and unexpected bonus of motion in the audio space which is realistically nearly impossible to achieve in any other sport.

Dolby Atmos Sound Design for the NHRA

The Existing Approach

The NHRA broadcast team approach their events with two Game Creek Video broadcast trucks, aptly named Nitro A and Nitro B. The local live productions comprise three distinct efforts that are largely split between the two trucks and one production trailer:

1. The main TV broadcast video and audio production, which consumes the entire Nitro A truck and the audio suite in the Nitro B truck, and
2. The on-site program production for in-venue replays and general entertainment of the fans called "All Access/Sunoco Vision," which consumes most of the resources in the Nitro B truck, and
3. An online program called "NHRA All Access" which produces an online product focused at the "Sportsman" and other classes of races.

The audio effects submixer in the Nitro B audio suite produces a submix of all the action on the race track for the broadcast product. That race track submix is relayed to the A1 in the Nitro A unit, where the program for air is mixed to include music, announcers, pre-recorded elements, etc.

The effects submixer in Nitro B had long been creating different submixes for different camera cuts and creating an audio-follow-video track mix. This existing discipline, refined over several seasons of both stereo and 5.1 program production, enhanced and jump-started the team's ability to create a new immersive experience.

Immersive Audio Considerations

In order to produce an immersive program for a live broadcast, the capabilities of the existing production needed to be assessed and in some areas, expanded. Notably:

- The EVS system needed to be updated to capture 16 channels of audio. This enabled the capture of both the existing 5.1 product and the new 10 channel (5.1.4) Atmos product.
- The mixing desk in Nitro B was originally a Calrec Brio and did not have the more extensive resources required to mix the planned immersive experience. Ultimately, the desk was swapped out with a Calrec Summa.
- The audio-follow-video approach to video scene changes which worked well for stereo and 5.1 programs needed to be revisited to ensure a stable immersive sound field when switching between different camera perspectives on the track.
- The existing mic plan, which relied heavily upon specific camera mics, needed to be revised for immersive audio capture. This required identifying the launch point as the area on the track where the most value for the immersive audio product could be realized.
- While adding new microphones to the effort was not completely out of the question, the effort focused on using the existing inventory of mics available in the truck which allowed the team to quickly test different mic placements and combinations. The harsh environment was also a consideration since if a mic failed, it could quickly be replaced from existing stock.
- Adding more mics specifically for the capture of audience reactions was considered but ultimately discarded due to the variance in infrastructures needed to support this approach in the different race venues across all 24 events. Sweetening the live crowd reactions using canned applause or cheers was deemed not to be an acceptable alternative either. Additionally, it became clear that the value-add of an immersive experience for the NHRA would not realistically embody live ambience capture due to the freakishly massive SPL generated by the vehicles themselves, and which is the main draw to the events for NHRA race fans.
- To maintain an immersive experience throughout the program and to avoid a collapse of the immersive sound field, a generalized ambient signal needed to be identified and captured to feed the overhead channels during non-race activity like driver interviews, taped segments from the team pits, the pre-recorded sponsored segments, etc.
- Monitoring capabilities within the Nitro trucks needed to be expanded to include overhead speakers and immersive volume control.
- The existing 5.1 product delivery needed to be maintained in the near term, so the immersive program would be an additional backhaul signal to the broadcast partner (FS1).

The technical areas were relatively easy to address (the production budget permitting), as those required only upgrades to firmware, the purchase of enhanced licenses for the products or, in the case of the mixing desk, replacement. Retrofitting the existing audio suites to accommodate overhead speakers was a bit tricky but was ultimately quite effectively achieved by Game Creek Video using a mix of the existing Neumann speakers for the front LCR, new Genelec 2110 for the Ls/Rs and high-end Klipsch in-wall speakers for the new overheads.

Opportunities for the capture of natural sound for spatial audio filler elements for the production of non-race segments are still being evaluated. Interesting audio elements from the race car pits are one option, with the capture of ambient sound from the rebuilding of the

cars' engines which happens after every run down the track. This ambient sound, called "thrash" in the vernacular of the NHRA, is a common audio element heard by the crowds walking through the pits at an NHRA event.

For the dual-program delivery of both 5.1 and 10 channel (5.1.4) immersive audio, the NHRA audio team decided to go with a bold but simple approach and begin mixing the live show solely as an immersive product and using the Calrec Artemis in Nitro A to fold down the immersive mix by lowering the overhead channels 9dB prior to folding them down into their respective L/R and Ls/Rs channels in the 5.1 "bed" to create the 5.1 product automatically. Since the overhead immersive channels would be carrying unique audio signals during periods of manual panning when the jib-cam was called, throwing away the overhead channels to create the 5.1 product, as has been done on other recent immersive audio events, wasn't an option. A distinct benefit to this approach was the fact that the 5.1 product would contain all of the panning motion present in the immersive product, only without the overhead channels, thereby improving the viewer's quality of experience when watching the 5.1 product.

Backhaul of the existing 5.1 program had been via up to 8 channels of PCM audio within SDI video. Adding immersive audio would expand that backhaul signal to a full 16 audio channels where 5.1 would continue to occupy channels 1-6 to serve the existing signal paths within the broadcast partner's infrastructure, and the immersive program in channels 7-16 feeding a new signal path. Within the FOX Network Operations Center in Los Angeles, in the near-term the two programs would be split and go through separate master control rooms (MCR) where commercials would be inserted. All commercials are played out in 5.1, and so on the Dolby Atmos channel, during periods of commercial breaks, the overhead channels would go silent. This would be a temporary approach for commercial insertion until an immersive upmixer was tested and put in place within the MCR. A mid-term goal is to backhaul only the 10-channel immersive product and render to 5.1 at the FOX NOC to feed legacy systems.

A trickier challenge for the audio team was the overall immersive mic plan for the events.

The Immersive Mic Plan

The physical experience of being in the presence of a pair of idling 10,000 horsepower race cars is awesome. Experiencing the launch of two, or at some events four of these vehicles simultaneously is indescribable. The result of all this power unleashed from a broadcast microphone standpoint is the clipping of input channels if not the utter destruction of the microphone itself. In addition to the threat of massive SPL levels, the burn-outs and launches of the vehicles generate a large amount smoke and debris, consisting of small, hot bits of rubber which can easily infiltrate the capsule or melt through a foam microphone windscreen. Additionally, the exhaust from a nitromethane-burning top fuel dragster contains nitric acid, which is highly corrosive. Selection and placement of microphones around the starting line was a critical factor in the success of this project.

Placing the microphones around the staging and launch area required a few considerations. Among them were:

- It is not possible at most race tracks to place any mics in the far lane due to access issues. No wires can be laid on the track and many venues lack wire access underneath the track. RF mics are possible if there is available bandwidth not already taken up by cameras and other devices.

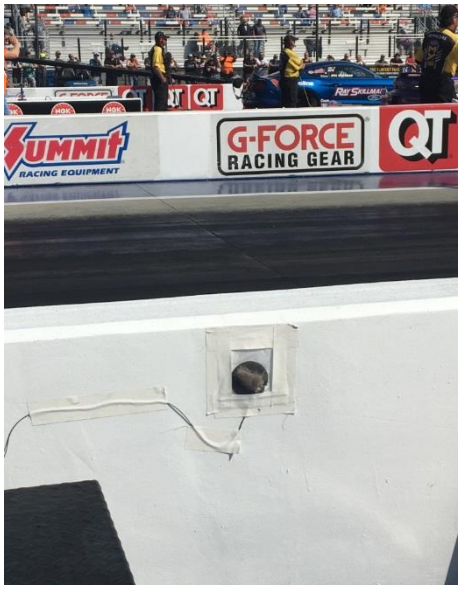
- At about 30 feet down track from the starting line, the vehicles achieve full power and their highest SPL, making this a particularly dangerous place for microphones.
- Placing mics within the island at the starting line requires attention to safety concerns as this is an area of high foot traffic and contains critical safety and timing systems.
- Access to the track and the launch island during Top Fuel races is limited and can be physically dangerous.

A DPA-4007A mic placed at the tree, bagged and ready for action

Testing using several different microphone combinations resulted in a few damaged capsules and discoveries. While experimenting with mic placement, a microphone "death zone" was found about 30'-50' from the launch point, as the vehicles reached maximum power there and all mics clipped, even when placed on the ground behind the concrete safety wall.

When changing audio scenes and during pan moves, it was observed that the low frequency component of the signal often seemed to go missing. This was caused by an abundance of directional microphones in the launch area and their physical proximity to the vehicles themselves as they moved back and forth in the launch area. In order to maintain a stable spectral density within the sound field where the low end remained robust and powerful, several directional mics in the launch area were eliminated from the mic plan and four omnidirectional mics were added, the physical placement of which was intended to be replicable at most, if not all race tracks where NHRA events are scheduled.

After extensive trial and error testing, ultimately a total of 10 microphones were strategically placed in the island launch area. Please refer to figure 1:

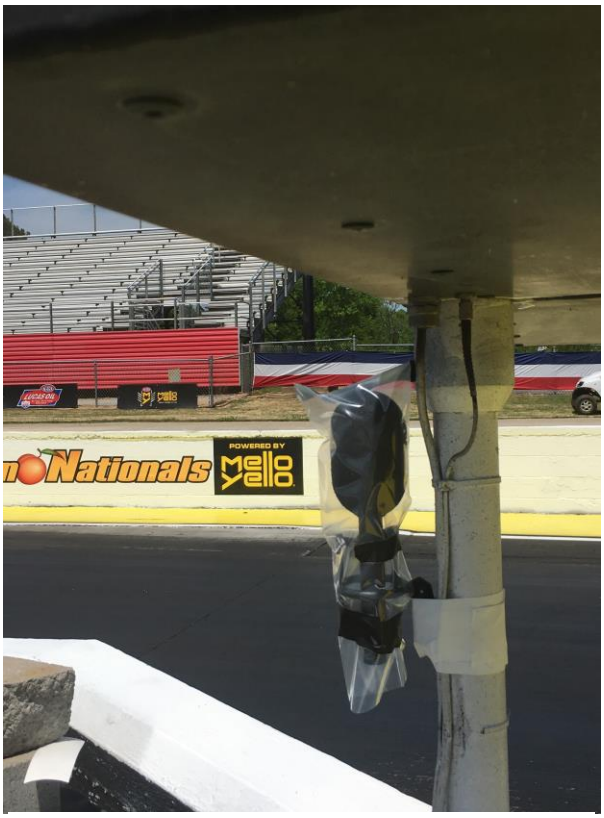


A Sony ECM-77B mic, bagged and taped to the interior of the island retaining wall.

- Sony ECM-77B omnidirectional lavalier mics taped to the interior wall of the island in locations 1 and 2. These were chosen due to their surprising performance and availability. While they (and most other mics) clipped when placed anywhere in the direct path of the vehicles, taping them to the retaining wall behind the launch area provided excellent coverage of the opposite lane.
- Two DPA-4007 omnidirectional mics at locations 3 and 4. These mics have a maximum dynamic range of 159dB and have a very wide capture coverage, which allowed us to use only two mics at either end of the launch area. Their frequency response is quite flat and provides a superb and stable low end response irrespective of their proximity to the vehicles, which was critical to the success of the project.

- A pair of Sennheiser 416 shotgun mics at locations 5 and 6. These mics were retained from the original stereo and 5.1 mic plan and are mainly used in a specific audio scene paired with a low POV, down-track facing camera to capture the sound from the near down track area after the burnouts and during the vehicles' roll back to the starting line.
- Four Earthworks M30 high SPL instrumentation mics in locations 7 through 10. These mics have a very small aperture which allow them to capture extremely high SPL sources. A pair of the M30s set at the "tree" were retained from the original mic plan and are mainly used in an audio-follow-video scene complementary to the Sennheiser scene mentioned above and paired with a low POV, up-track facing camera to capture the roll back after the burnouts and in some cases the initial launch of the vehicles. A second pair of the M30s are pointed at each lane towards the rear of the island to capture the idling vehicle.
- All of the mics are bagged and sleeved to help prevent damage from debris flung by the vehicles during the burn-out and launch.

Figure 1



A DPA-4007A mic, bagged and placed at the rear of the island in the launch area.

The Sennheiser pair (mics 5/6) and Earthworks M30 pair (mics 7/8) at the top of the island are used for a specific audio-follow-video scene change using two small RF cameras mounted near the surface of the track. After the burn outs, the down-track facing camera captures the vehicles rolling back to the launch pad and the Sennheiser mics capture the sound of the vehicles moving front to back. When the director calls the companion up-track facing camera, the visual perspective switches as does the audio perspective, and now the vehicles are heard moving back to front. Since the camera's point of view is so low, this simple camera cut sequence also leverages the front and rear height

channels available in Dolby Atmos and requires no manual panning, unlike the jib camera cut. When the up-track camera is used for the launch of the vehicles, the sensation is as if the vehicles are driving over the viewer, and is quite compelling.

The island mic plan is intended to be the stable bed for the immersive mix within which the individual mono, and stereo camera mics are added to leverage the production team's existing audio-follow-video scene approach. In addition, the jib-cam's mic and any appropriate hand-held camera mics would be brought in and panned in real-time through the island bed audio to create audio motion across both the lower and upper planes of the immersive environment.

When the cars launch, the down-track audio capture of the race itself is done via existing and race-proven conventional methods, using a combination of statically placed mics and camera mics that are preassigned a place in the audio scene, generally assigning them in a left-to-right sound field. As the races themselves are only around 4 seconds in duration, this is the only practical approach to audio panning and capture and works quite well overall.

At the 1000-foot finish line, if all goes as planned, the parachutes pop and the engines shut off. The pop and fluttering of the parachutes are captured by additional mics down track. If the 'chutes or brakes malfunction, there are additional mics in a gravel pit at the end of the track to capture the sound of the car entering this area, and in some unfortunate cases, the car being shredded by a steel capture net.

For the immersive audio plan, only the staging area was addressed and the existing mic plan in place for down-track capture of the races was left in place as it was already quite good and effective for both conventional and immersive production.

Executing the Immersive Audio Strategy

Bringing the immersive audio plan to reality was a year of effort which included testing of mic placements and temporary installation of monitoring in the Nitro audio suites, prior to the permanent installation of the immersive production kit. Additionally, production meetings began to include audio coverage strategies for the on-track experience and tight coordination between the video and audio teams.

Determining the proper approach to the immersive mix was another area of hard work and trial and error. Ultimately, the team landed on a recipe of a stable bed using the statically placed mics as described previously combined with highlight mics from the cameras as called by the director, with the jib-cam including manual panning along with statically panned motion, as with the track-level cameras covering the burn outs and roll backs.

Panning Audio in Real-Time

Mixing for immersive audio (5.1.4) using a 5.1 mixing desk requires the combination of two 5.1 busses. The Calrec Summa is a 5.1 desk containing 8 5.1 busses, which meant that for immersive audio, there would be 4 immersive busses. Two of those 5.1 busses are consumed by the manual panning activity (actually, 5.0 busses since the LFE channel is not panned).

To pan in an immersive environment in real-time, the effects submixer assigns one fader to a 5.0 buss and another fader to another 5.0 buss. The first 5.0 buss feeds the L/C/R/Ls/Rs mix output while the second 5.0 buss feeds the four overhead channels. Assigning the same signal to the two faders and crossfading with one hand while manipulating the panning control with the other achieves the desired result of manual panning in X-Y-Z planes.

Clearly, such an approach requires close coordination between the audio and video teams and a fair bit of practice. The director needs to announce the camera cut early before taking it in order for the submixer to prepare the desk for the move. When the cut is made, the camera operator needs to keep a pace to their movements such that the submixer isn't moving the audio too quickly during a manual pan, which could result in a jarring and suboptimal experience, nor so slowly that the audio motion would be less perceptible and unengaging.

The result is quite effective, but like any complicated manual process that requires a significant amount of cooperation, it is also quite difficult to execute. It requires significant practice, the opportunities for which often come only during live broadcasts, and like any live broadcast, actions on the field of play will affect the director's decisions. During an NHRA broadcast, not all individual races will be an opportunity for manual panning, however the immersive audio effect is always present due to the combination of audio elements statically placed within the sound field and the audio scene changes that follow the director's camera cuts.

Looking to the Future

The NHRA's events provide a unique opportunity to create a unique hybrid approach to both conventional and immersive live television audio, leveraging aspects of conventional live audio production with cinematic, post-production-like techniques. For most sports, this approach would be unrealistic since typical ball sports and even oval and road track motor sports are significantly less predictable, and otherwise don't offer the repeatability found in the staging of an NHRA drag race.

Even so, with a view towards adding motion in the audio space to a live sports production, opportunities can be found and leveraged, like the common use of cable and drone cameras and once these opportunities are found, the obstacle then becomes the execution of the audio pan.

An automated approach to live audio panning would potentially remove this obstacle. Leveraging the existing telemetry used for the insertion of on-screen virtual graphics and advertisements could provide enough steering information to an external device or a mixing desk which could drive this panning capability, thereby removing the manual panning responsibility from the live mixer, which could more accurately match the camera move itself and add production value to the program in both conventional 5.1 and immersive audio productions.

References

[1] McGee, Ryan "Drag Racing: The Loudest Sport," November, 2010
http://www.espn.com/espn/news/story?id=5759488

Acknowledgements

Mike Rokosa, NHRA Technology Executive
Russel Roark, Audio Mixer
Josh Daniels, Audio Mixer
Robb Scheetz, Audio Mixer