Dolby

The Sound Makes the Picture:
Making Audio Unforgettable with NEXTGEN TV and Dolby AC-4

—

Tim Carroll

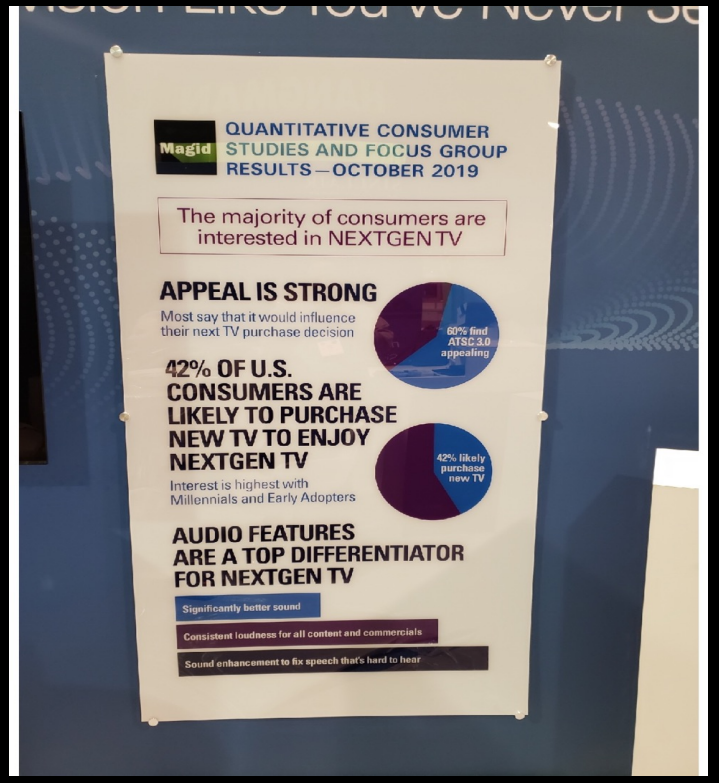
tcarr@dolby.com

ATSC 3.0

- A/300 – ATSC 3.0 System; Section 5.1.15 Audio
  - "All ATSC 3.0 terrestrial and hybrid television services emitted within a given region shall use one audio system selected for that region from those defined in A/342 Parts 2 and higher."
  - "…broadcast organizations in North America have selected the audio system defined in A/342, Part 2 as the audio system for use in Mexico, Canada and the U.S."
  - Crucial language for CE manufacturers to include what is necessary in a given region. Other formats might result in audio but by chance not design.
- A/342 Part 2 "AC-4 System"
  - Enhanced feature set
  - Improves upon the capabilities of past ATSC audio systems

Consumer Study Reinforces NEXTGEN TV Audio Importance for Consumers

- Pearl TV presented the results of a second consumer study that reinforced the findings of their initial study: Audio is at the top of the list of features consumers desire
- Conducted by consumer research firm Magid, the study highlighted that ATSC 3.0 NEXTGEN TV "Audio Features are a Top Differentiator for Consumers," specifically:
    - Consistent loudness for all content and commercials
    - Immersive, theatre-like sound
    - Fixes speech that's hard to hear: Voice + (dialog enhancement), exclusive to AC-4



DOLBY AC-4 ENABLES DELIVERING WHAT CONSUMERS WANT

ACCESSIBLE

- Dialog audibility
- Audio description

ADAPTABLE

- Multi-device Single stream
- Optimized output

IMMERSIVE

- Three-dimensional sound
- Premium audio

PERSONALIZED

- Customized experiences
- Multi-language

Improvements since ATSC 1.0 and AC-3 (1995)

We've come a long way baby!

- Processing power
  - 1995: DSP farm required for encode; barely enough cycles to just decode 5.1
  - Today: The heaviest lift still done by the encoder, BUT consumer side is very powerful: supports decoding & post processing (DRC, virtualization, speaker EQ, etc…)
- New modes:
  - Channels and objects (which are channels with positional data) supporting mono, stereo, 5.1, 5.1.4, and "compositional" configurations with separate M&E and D(s)
  - Dialog Enhancement (Voice +)
  - Dolby Atmos

So, what the heck is Dolby Atmos?

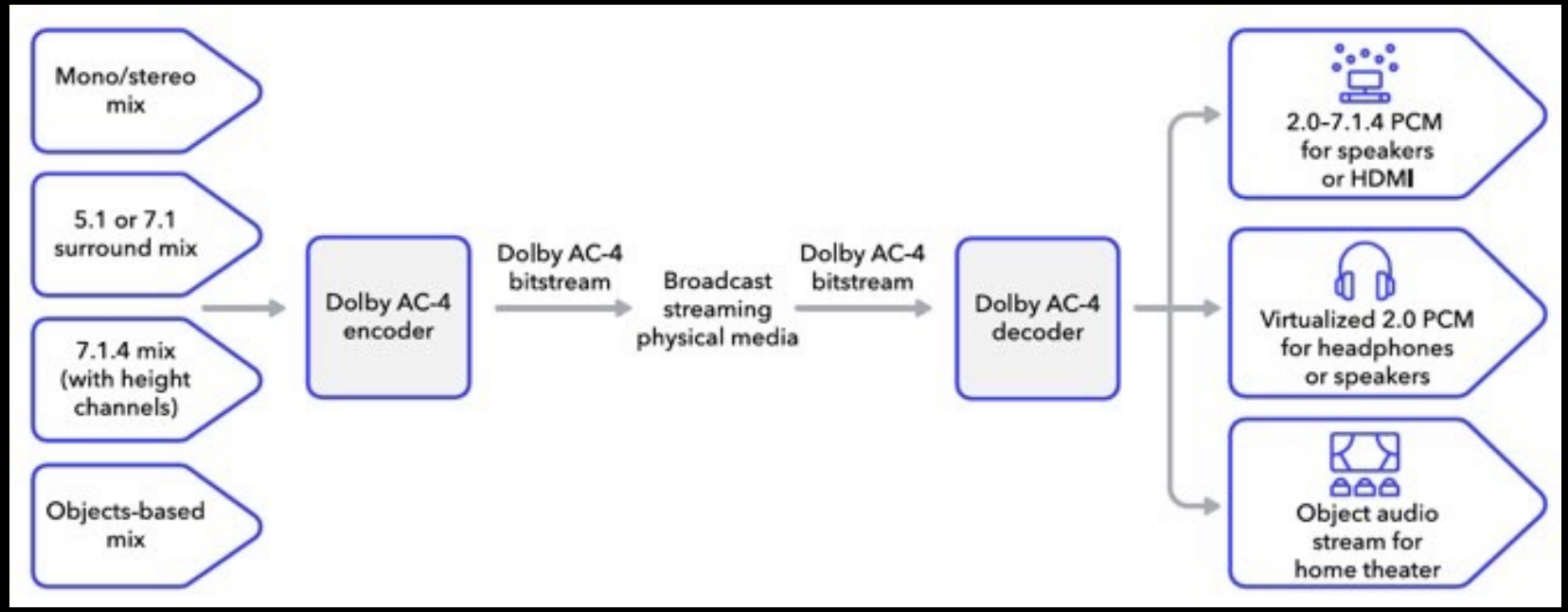
- It’s not just height!
- Encoding height channels is not useful if they fall on the floor
- Height speakers are not useful if sources limited to 5.1 channels
- Atmos is a *translational format*
- It enables decoupling the output and input channels

Channels in ATSC 1.0 and AC-3

Output follows input. Sources can be downmixed to match speakers, surround decoding can sometimes be invoked as a post-process

Channels in ATSC 3.0 with AC-4

Outputs and inputs are decoupled: any source can be translated to any output configuration in a deliberate and predictable manner



Broadcaster Summary

- Metadata very automatic – power on and go for today's formats
- Loudness compliance is built-in and automatic
- New mixing and monitoring tools exist from well-known 3rd parties
- System is ready for production at the network level to catch up, though questions around how to deliver via traditional infrastructure
  - Creativity with SDI enables carrying legacy and some new capabilities
  - IP infrastructure removes limitations and includes metadata

Get Up and Running Today

- Video encoders have or will have full AC-4 audio encoding functionality built-in:

Pro

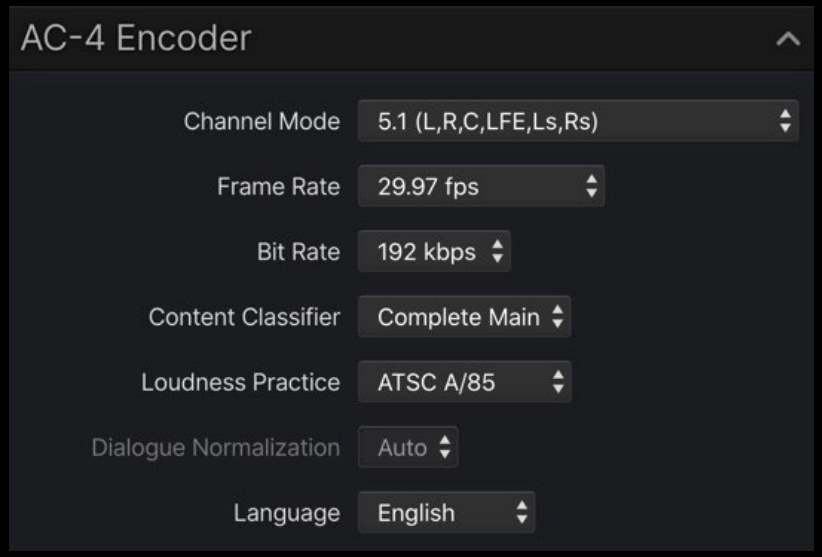
- Simplifies A/V sync
- Packaging more consistent

Con

- May not support additional audio functions like watermarking, upmixing
- Video experts might not be audio experts

- Video encoders are required to support passthrough of externally encoded audio
- In either case, AC-4 is AC-4 and the consumer audio experience is better right out of the box

Get Up and Running – Suggested Settings



- Framerate – Audio frames should match video frames
- Bit Rate - Minimum recommended for 5.1 channel audio is 192kbps (preserves watermarking, headroom for downstream processes)
- Content Classifier – CM (complete main; other choices are compositional)
- Loudness Practice – ATSC A/85 (sets dialnorm, adjusts audio if necessary)

Production

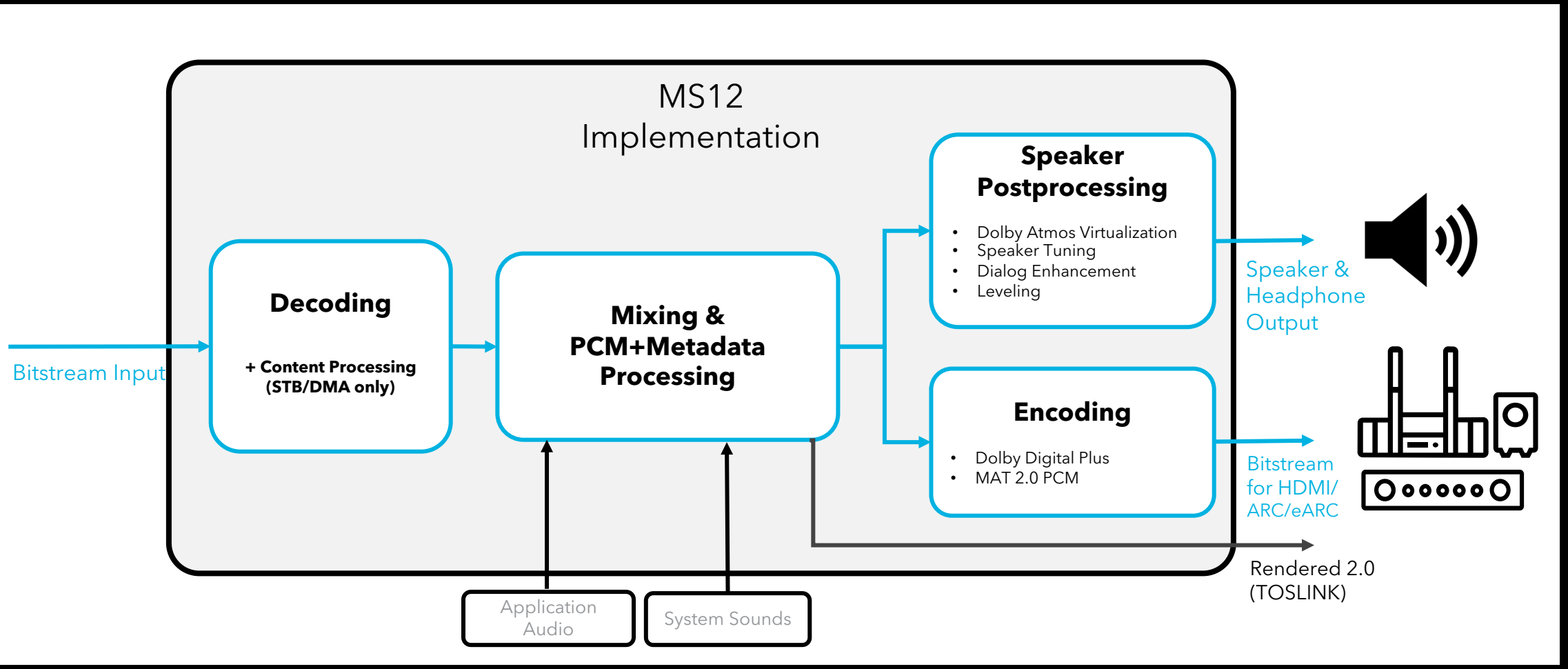
- Upmixing is (still) an easy way to improve things locally
  - Works with all existing production techniques though stereo is a minimum
  - Broadcaster controls their sound, “lights the light” as a differentiator
  - Keeps dialog centered and can improve results of virtualization
  - Professional bests experimental – has to work in many modes
  - Easily accomplished with external devices such as:
    - Linear Acoustic UPMAX ISC – stand alone unit
    - Linear Acoustic LA-5300 – integrated with AC-4

Consumer

- Solid progress delivering next gen experiences:
  - Dolby Atmos available via all major streaming services
  - Many CE devices support Dolby Atmos
  - Soundbars are a good thing (and no more need for ladders, saws or excuses)
- AC-4 now brings automatic consistent loudness, Voice + and of course Dolby Atmos via NEXTGEN TV
- Relies on fielded methods for widest compatibility:
  - Dolby MS12
  - Dolby MAT 2.0

MS12

Dolby's one-stop-shop audio solution for consumer playback devices



Dolby MAT 2.0

MAT 2.0 is a solution to transmit more audio

- via HDMI 1.4 and higher
- via eARC in HDMI v2.1 and higher

MAT 2.0 carries up to 31 objects + LFE + Metadata and transmits Dolby Atmos between source and sink devices.

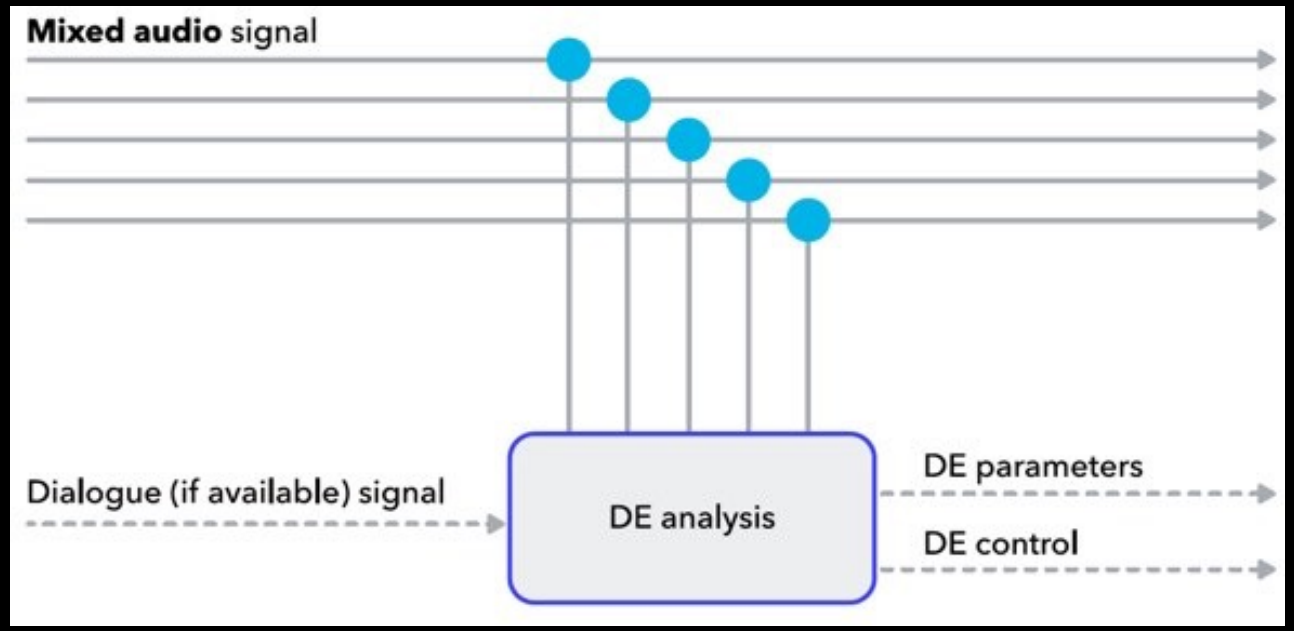
Present in:

- 100% of Dolby Atmos TVs, soundbars, and AVRs shipping today
- XBOX, AppleTV, Roku 4k Ultra 2020, etc…
- Enables ATSC 3.0 audio compatibility

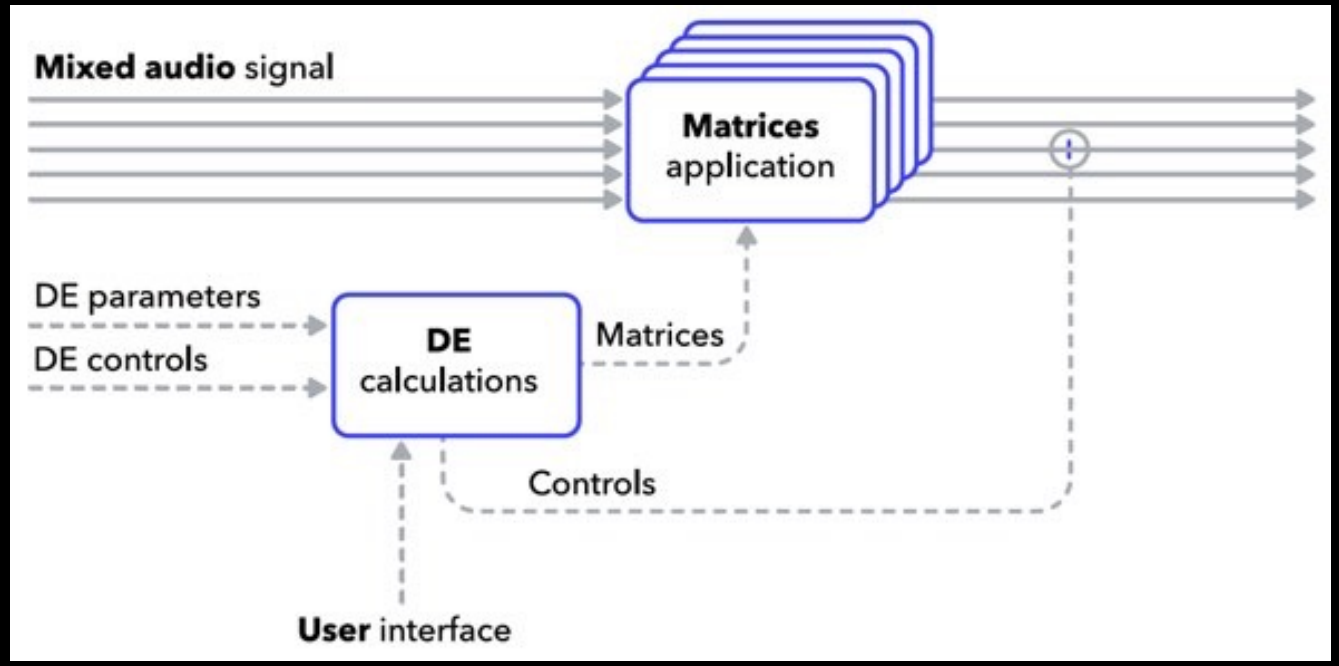
Voice + is an Essential Feature for Consumers

- Pearl, the CTA and Dolby collaborated to create the consumer-facing “Voice +” moniker to describe the dialog enhancement function uniquely provided by AC-4
- As popular with consumers as immersive audio and other audio features, Voice + provides immediate benefits for existing content and more for new content
- It is cost-free, operates automatically in the encoder, present in all decoders
- Can extract dialog from legacy 5.1, stereo, and even mono audio
- Further improved results possible in the future if dialog separated, e.g. M&E+D

Dialog Enhancement (Broadcast Side)



Dialog Enhancement (Consumer Side)



Accessibility: Better than Basic

- Why shouldn’t all services be compelling to listen to?
- Voice + is an accessibility feature – it can improve all audio and it is automatic
- Mono AD is a legacy constraint - stereo or 5.1 often takes little additional effort, networks exploring, also good recent experiences with live describers
- 5.1 channel M&E+D is possible with existing consoles, enables the following:
  - 5.1 M&E + English D + Spanish D + Barstool D + AD
  - Consumer selects desired presentation
  - Easier with live production

Resources

https://professional.dolby.com – Interesting view across several areas including Dolby Atmos and Dolby Vision, including helpful tutorials

www.atsc.org – Lots of standards and helpful recommended practices

Dolby ATSC 3.0 Audio Handbook – Overview of integration, recommended operating points (data rates, etc…)

Thank You!