Introducing the MPEG-H Audio Alliance's New Interactive and Immersive TV Audio System

Presentation at AES 137th Convention, Los Angeles

Robert Bleidt – October, 2014

Based on the new MPEG-H Audio Open International Standard ISO/IEC23008-3

Fraunhofer

IIS

MPEG-H Audio improves the listening experience

- The newest audio codec from MPEG expands the features of today's codecs:

- Personalize the sound to what you want to hear

- Make surround sound more realistic

- Tailor reproduction to your listening environment and hardware

- And the technical stuff:
  - One stream to every device
  - Combine multiple streams for hybrid delivery
  - Audio I-Frames for easy DASH segmentation
  - Unsurpassed coding efficiency

Fraunhofer
IIS

MPEG-H Audio improves the listening experience

Personalize the sound to what you want to hear

Fraunhofer

IIS

MPEG-H Audio improves the listening experience
Personalize the sound to what you want to hear

- In addition to traditional channel-based audio, MPEG-H includes audio objects, sound elements that are transmitted separately to the decoder

- This allows the listener to adjust the elements to his liking

- Examples:
  - Simple case: Adjust the dialogue or commentary level in relation to the other audio elements such as ambient sound, music, or sound effects
    - Precursor 2011 Wimbledon "Grunt Controller" trials with BBC showed most listeners preferred either a higher or lower dialog level
    - Helpful for listeners with hearing loss or non-primary language or dialect
    - Additional language support becomes a simple matter of an additional 20-40 kb/s object

Fraunhofer
IIS

First practical use of objects: Dialogue Enhancement
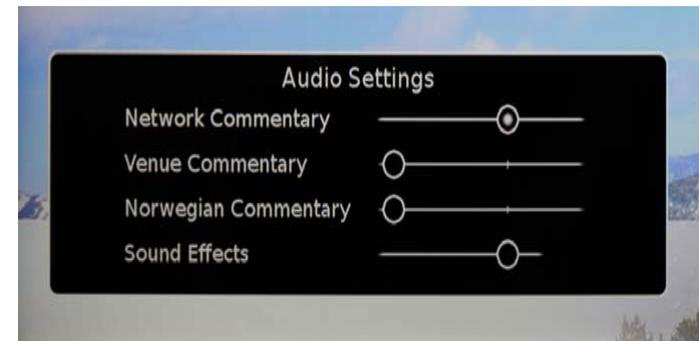## Allows viewers to set their own Dialogue levels

- User benefit
  - Dialogue/Announcer can be adjusted up or down
  - Helps hearing impaired understand Dialogue
  - Helps for listening in noisy environment
- Broadcaster benefit
  - One signal broadcast to all viewers
  - Default mix played by existing receivers

- Public test by BBC during the Tennis Grand Slam Championships 2011 in Wimbledon

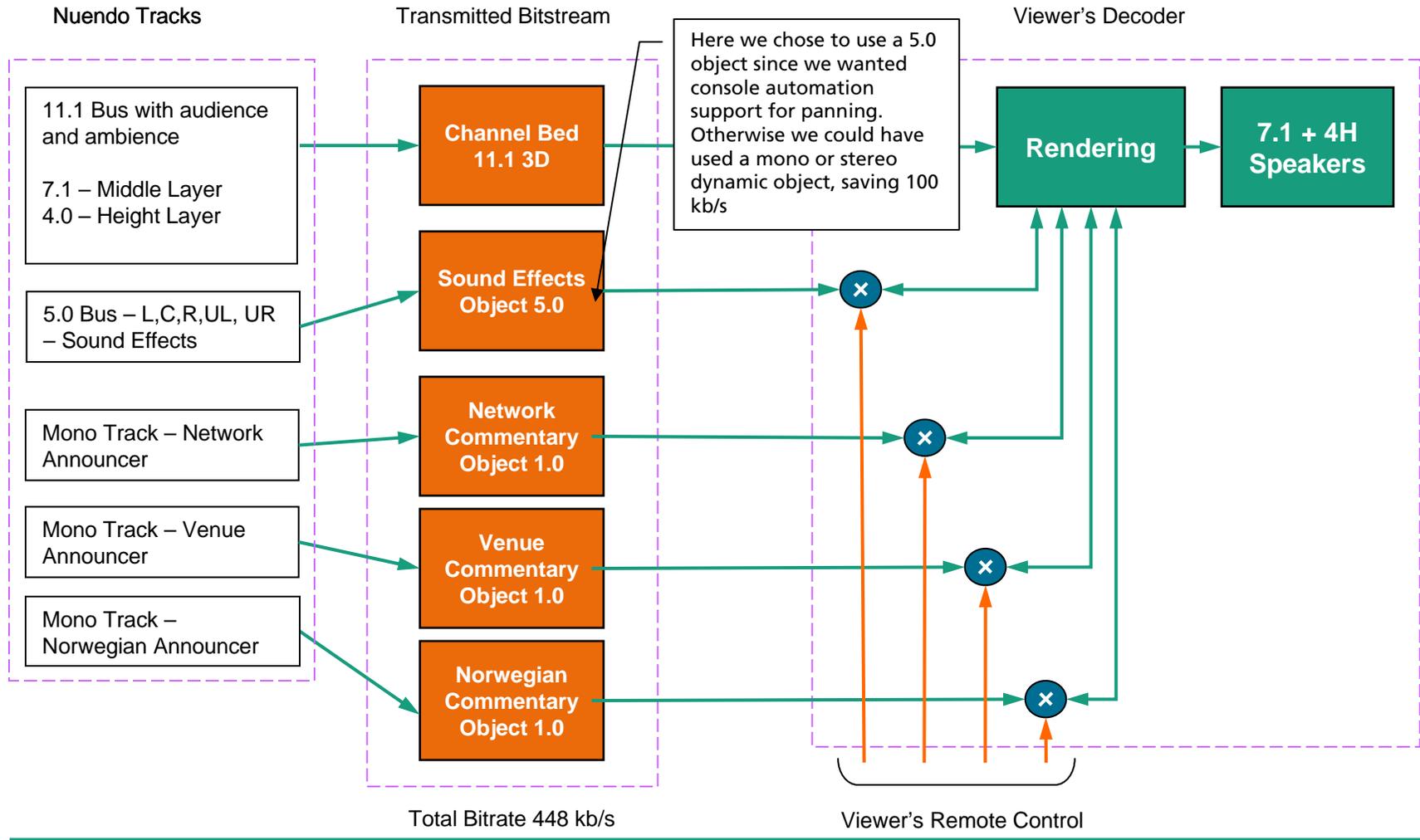- Now undergoing standardization in DVB

MPEG-H Audio improves the listening experience
Personalize the sound to what you want to hear

- More complex examples:
  - 2014 Field Test with leading U.S. sports network at winter sports event
    - Pick your announcer and language
    - Want to hear more sound effects?
      More audience sounds?
  - Demo content from NASCAR race
    - Pick your announcer and language
    - Mix in driver's radio
    - Hear car sounds/transducers
- Real-time decoding and rendering
- User selects audio mix from on-screen display sliders or **preset buttons**

≡ Fraunhofer
IIS

How we Used Objects in our Field Test Mix

Nuendo Tracks

Transmitted Bitstream

Viewer's Decoder

11.1 Bus with audience and ambience

7.1 – Middle Layer
4.0 – Height Layer

5.0 Bus – L,C,R,UL, UR – Sound Effects

Mono Track – Network Announcer

Mono Track – Venue Announcer

Mono Track – Norwegian Announcer

Channel Bed 11.1 3D

Sound Effects Object 5.0

Network Commentary Object 1.0

Venue Commentary Object 1.0

Norwegian Commentary Object 1.0

Here we chose to use a 5.0 object since we wanted console automation support for panning. Otherwise we could have used a mono or stereo dynamic object, saving 100 kb/s

Rendering

7.1 + 4H Speakers

Total Bitrate 448 kb/s

Viewer's Remote Control

Fraunhofer IIS

MPEG-H Audio improves the listening experience
Personalize the sound to what you want to hear

- Summary: Personalization

- Audio objects allow listeners to:

  - Adjust elements in the sound mix to their liking

  - Boost dialogue for hearing-impaired or second language listeners

  - Listen to alternate audio content such as:

    - Home vs away team commentary

    - Driver to crew race audio

    - Alternate languages, simultaneous translation, or audio description

    - Many other future creative and revenue opportunities

≡ Fraunhofer
IIS

MPEG-H Audio improves the listening experience

Making surround sound more realistic

Fraunhofer

IIS

MPEG-H Audio improves the listening experience
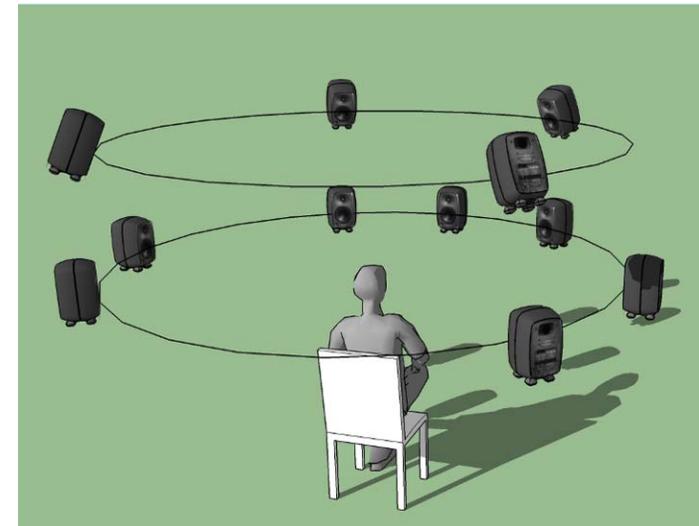## Making surround sound more realistic

- 3D audio systems add height information (more channels or audio objects) to the sound field to improve realism

- It becomes difficult for the listener to perceive he is hearing a recording

- Our studies show adding **four height speakers** provides a substantial improvement without the 22 speakers of classical 3D systems
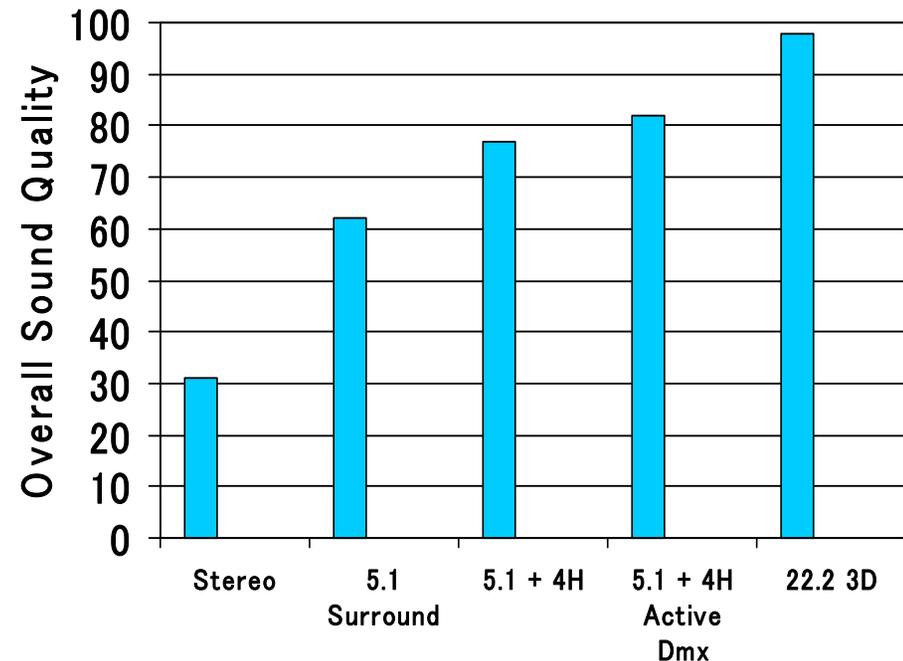
Fraunhofer
IIS

MPEG-H Audio improves the listening experience
## Making surround sound more realistic

- MPEG-H currently has defined speaker configurations from 1.0 (mono) to 22.2

  - Bitstream support for up to 128 channels and 128 objects

- **Our recommendation for 3D sound:**

  - **7.1 Blu-ray surround configuration**

  - **4 Height Speakers above corner speakers as shown**

- Other configurations are possible, for example:

  - 5.1 + 4H

  - Using front wide 7.1 configuration or 9.1

≡ Fraunhofer
IIS

MPEG-H Audio improves the listening experience
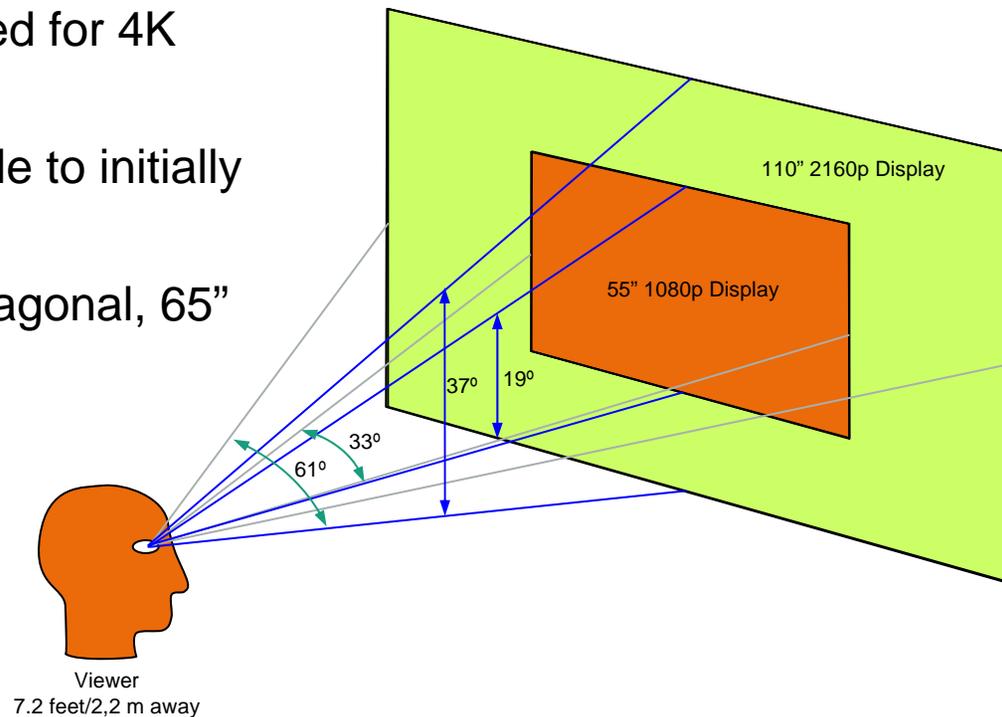## Making surround sound more realistic

- MPEG-H includes sound field capture technology for alternative means of capturing and reproducing 3D sound

- Based on Higher-Order Ambisonics techniques

- Transmits predominant sounds and ambience in separate tracks, independent of speaker format

- Allows one-microphone capture of sound field or ambience

- As with channel-based system, may be combined with audio objects for interactivity or personalization

Fraunhofer
IIS
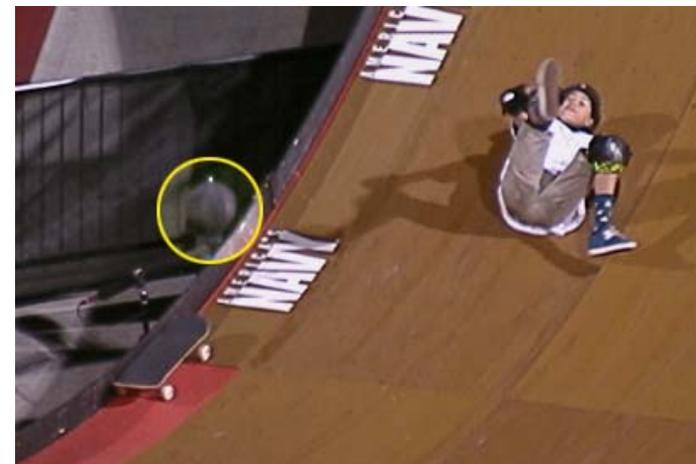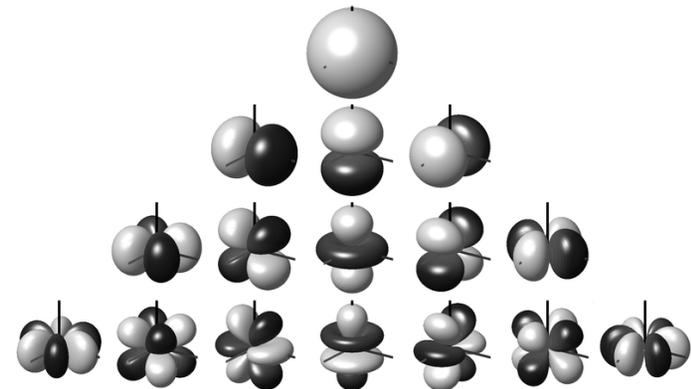
Recommended Vertical Viewing Angles Increase with 4K

System	1920 x 1080	3840 x 2160
Viewing distance (relative to picture height)	3.2	1.6
Viewing angle - horizontal (degrees)	31	58
Viewing angle - vertical (degrees)	18	**35**

- 1.6 PH viewing is recommended for 4K (BT.2022)

- (though consumers may be able to initially afford/accept only 2 PH)

- Setup note: 2 PH ~= screen diagonal, 65" away from 65" TV

110" 2160p Display

55" 1080p Display

37º 19º
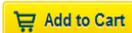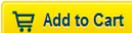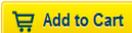
33º

61º

Viewer
7.2 feet/2,2 m away

Fraunhofer
IIS

"The average consumer will never hear 3D, they listen to stereo today"

- It is true today's consumers listen mainly on stereo in-TV speakers or with simple 2.0 or 2.1 soundbars

- 5.1 or 7.1 AVR + speaker systems aren't seen as the logical match for that new sleek, slim flat-panel TV

- So why do 3D audio if most of the audience won't hear it?

  - The same reason we do surround today – the high-value viewers will have 3D in their home theaters

  - Right?

Recommended For You

Samsung - 2.1-Channel Soundbar with Wireless Subwoofer
$299.99
Add to Cart

Samsung - AudioBar 2.1-Channel Soundbar System with Wired
$179.99
Add to Cart

Samsung - 2.1-Channel Soundbar System with Wireless Active Subwoofer
$1,299.99
Add to Cart

Fraunhofer

IIS

Fraunhofer demonstrates future 3D listening for mainstream consumers with prototype "3D Sound Bar"

- Concept demonstration of our vision of mass consumer delivery of 3D audio

- Décor-friendly – no external wires or speakers

- No confusing AVR setup menus – un-box, hang on wall, enjoy

- Could be integrated into TV or offered as a separate frame

- Dramatic improvement over today's stereo soundbars

- Home theater enthusiasts will still want 11.1 speakers (or more) for ultimate sound quality

- Shown at the Fraunhofer IIS booth at NAB 2014, IBC 2014

FRAUNHOFER
INTERACTIVE 3D AUDIO

Winter Sports Event
January 2014
Field Test

Fraunhofer
IIS

Fraunhofer
IIS

MPEG-H Audio improves the listening experience
Making surround sound more realistic

- Summary: 3D Audio

- 3D provides a substantial increase in sound quality beyond surround sound

- Four height speakers provide a good tradeoff between cost and sound quality

- "3D sound bar" technology could be built into TVs or soundbars to enable a dramatic improvement in sound quality and immersive sound for the mainstream consumer

- Higher Order Ambisonics offers an alternative method of capturing and transmitting 3D sound

Fraunhofer
IIS

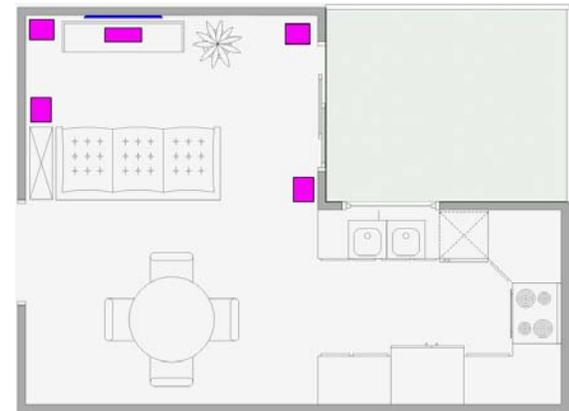# MPEG-H Audio improves the listening experience
Tailor reproduction to your listening environment and hardware

- Multi-platform loudness control

  - Loudness measurement built into the encoder

    - Or use external values

  - Uses existing Dolby metadata in legacy mode

  - Adjustment of dynamic range to match listening environment and device capabilities

    - Examples:

      - Home listening on good speakers

      - Listening on earbuds at airport on mobile phone

      - Listening at home on tablet

≡ Fraunhofer
IIS

MPEG-H Audio improves the listening experience
Tailor reproduction to your listening environment and hardware

- Speaker-foolproof rendering for legacy speakers

  - May be able to correct for misplaced speakers in consumer's home

    - HDMI 2.0 will allow communication of consumer's installed speaker locations

  - May offer a limited impression of height on existing 5.1 or 2.0 consumer speakers

- Binaural rendering for headphone reproduction

  - Listener perceives sound stage in front of him, instead of between his ears

  - May be extended to tablet speakers with Fraunhofer Cingo technology

Fraunhofer
IIS

MPEG-H Audio improves the listening experience

The Technical Stuff

A Brief introduction to some selected
MPEG-H concepts and features

Fraunhofer

IIS

MPEG-H offers the best coding efficiency available today

■ Includes a highly efficient new audio codec to enable true 3D Audio transmission in broadcasting and streaming:

Bitrates in kb/s for:	Good	Recommended	Transparent
22.2 Channels	256	512	1200
7.1 + 4 Height Channels + 4 Objects	200	**384**	800
5.1 Channels	96	160	256
2.0 Channels	32	56	160

■ Includes the unified Fraunhofer/VoiceAge improved speech codec and new spatial coding techniques

■ Better speech quality at low bitrates

■ Improved stereo and multichannel imaging

■ Improved coding efficiency

Fraunhofer

IIS

Combining streams in the decoder

- In MPEG-H, objects may be delivered over disparate networks and supplied to the decoder separately

- Enables broadcast TV delivery of main content, with alternative or auxiliary content delivery over the Internet, for example

- Effectively, late binding in the MPEG-H decoder

Fraunhofer

IIS

Implementing MPEG-H in TV Broadcasting

Easy four-stage process as individual broadcasters desire

Fraunhofer

IIS

Developing a TV Audio System Based on MPEG-H

- The MPEG-H Audio Alliance is developing a practical TV audio system based on the MPEG-H standard

- Includes not just audio coding, but a complete end-to-end system for content production, transmission and consumer delivery

- We propose a four-stage process for implementation by broadcasters:

    - From the start, consumer decoders will support all features of our system

    - Broadcasters can choose what features to use and when to implement them

Fraunhofer
IIS
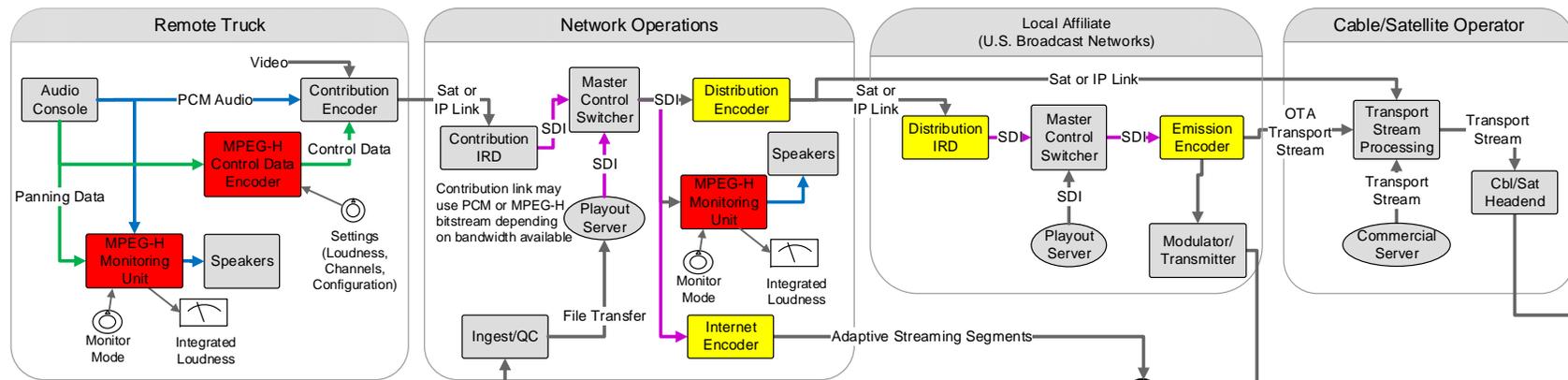
Stage 1 – today's surround with current practices

- Replace AC-3 encoders with MPEG-H encoders

- (likely simultaneous with replacing MPEG-2 video encoders with HEVC or SHVC

- Benefits:

  - Same audio quality with half the bitrate of AC-3

  - Similar operational practices using traditional loudness dialnorm and compressor profiles of AC-3

  - Automatic loudness profiles generated for mobile or tablet delivery

Fraunhofer
IIS

Stage 2 –
Add objects for consumer choice of audio elements

- Primary language dialogue (for Dialogue Enhancement – consumer adjustment of dialogue level)

- Second or third language dialogue

- Audio Description / Visually Impaired

- Home vs. Away team announcer

- Driver – Pit Crew radio

- Athlete's playlist?

- Sound Effects

- Mono objects require at 20-40 kb/s each

- Objects can be programmed individually or in presets or groups

- Requires additional channels in TV plant for objects

Fraunhofer
IIS

Stage 3 – Adding Immersive Sound

- Add practical immersive audio – perhaps 5.1 or 7.1 + 4 height channels

- Increasing the realism of the audio image to create the "you're in the scene or at the event" experience.

- May use Channels-based or HOA-based immersive sound bed + audio objects

- Requires four additional audio channels in TV plant for height channels

- System supports audio up to 128 channels, configurations to 22.2 defined in standard.

Fraunhofer
IIS

Stage 4 – Adding Dynamic Objects

- Dynamic Objects – change position over time as specified by control data in bitstream

- Interesting not only for feature film or episodic drama, but for effects in sports broadcasting and other genres

- Requires transmission of control data through broadcast plant to final emission encoder

  - We'll have a future paper on that…

Fraunhofer

IIS

Progress with MPEG-H continues

- MPEG Standard at DIS stage

  - Final International Standard expected in January

- Being considered for upcoming ATSC 3.0 and DVB TV standards

- Pioneering experiments in sports broadcasting and other media to test production with MPEG-H's new features

- Visit the Fraunhofer's AES listening room on the show floor to learn more: Booth 1509

- Or go to www.iis.fraunhofer.de/tvaudio or www.mpeghaa.com

Fraunhofer
IIS

USA: codecs@dmt.fraunhofer.org   EU: amm-info@iis.fraunhofer.de

Fraunhofer
IIS