FRAUNHOFER EXPLAINS: MPEG-H IMMERSIVE SOUND FOR BROADCAST, STREAMING, AND MUSIC

AES Los Angeles
Section Webinar

Robert Bleidt
April 2020
codecs@dmt.fraunhofer.org

360 Reality Audio Information: sony-360ra-license-office@sony.com
https://www.sony.net/Products/360RA/licensing/
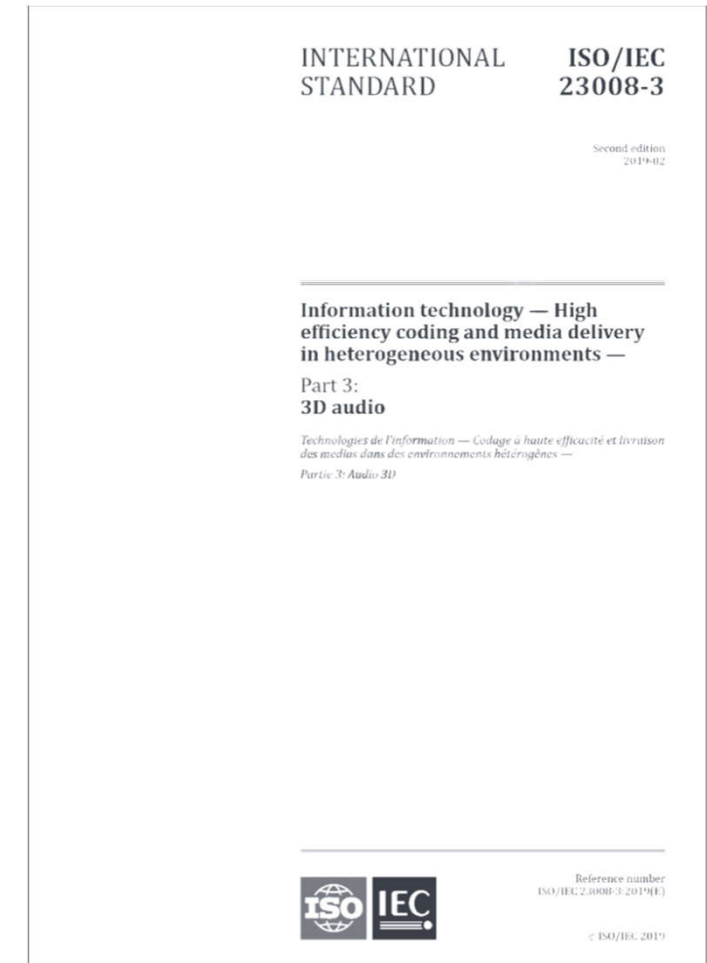
MPEG-H AUDIO

Fraunhofer

IIS

Agenda

- Introducing MPEG-H

- From mono to immersive sound – getting to "you are there"

- Not just immersion – solving two other big problems

- Developing the MPEG-H Audio System

- MPEG-H Tests and Adoption

- MPEG-H Playback Devices

- Mixing and Mastering in MPEG-H

- Content interchange standards, archiving, conversion tools

- Demonstrations

     www.mpeg-h.com   www.iis.fraunhofer.de/audio
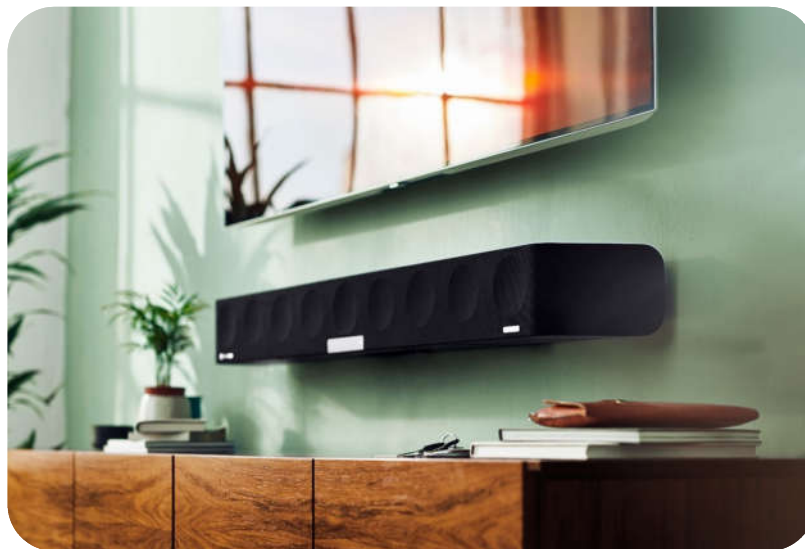
MPEG-H AUDIO

Fraunhofer

IIS

MPEG-H is:

- An international standard from MPEG, the organization behind MP3, AAC, MPEG-2, AVC, HEVC and other audio and video standards

    - Immersive "3D" Sound

    - Consumer Personalization or Interactivity

    - Universal Delivery to Playback Devices

- A complete consumer audio system developed around the standard by Fraunhofer and its partners

    - Software implementations and accessory products

    - Production and archiving tools

    - Decoder and product testing

- The basis of the 360 Reality Audio music format developed by Sony

INTERNATIONAL STANDARD

ISO/IEC 23008-3

Second edition 2019-02

Information technology — High efficiency coding and media delivery in heterogeneous environments —

Part 3: 3D audio

Technologies de l'information — Codage à haute efficacité et livraison des médias dans des environnements hétérogènes —

Partie 3: Audio 3D

Reference number ISO/IEC 23008-3:2019(E)
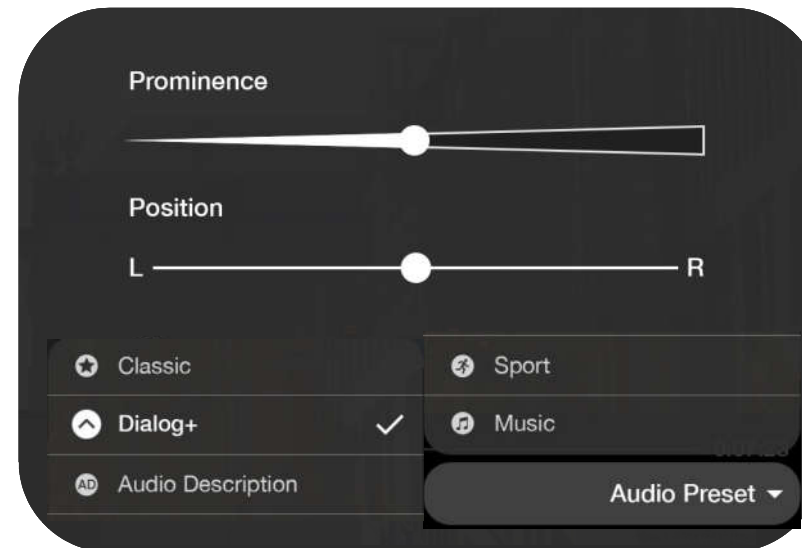
© ISO/IEC 2019

MPEG-H AUDIO

Fraunhofer

IIS

MPEG-H Audio – Three Main Feature Sets

Immersive Sound

- A viewer becomes part of the audience
- Delivered to mainstream consumers, not just enthusiast viewers

Interactivity

- Hear your home team
- Turn the announcer or dialogue up or down
- Hear your pit crew

Universal Delivery

- Home Theater
- Headphones
- Tablet Speakers
- Earbuds on airplane

MPEG-H AUDIO

Fraunhofer
IIS

GETTING TO "YOU ARE THERE"

The development of immersive audio

Rock in Rio 2019, MPEG-H test by Globo TV
on ISDB-Tb, 5G wireless test channel, HLS streaming

   www.mpeg-h.com   www.iis.fraunhofer.de/audio

MPEG-H AUDIO

Fraunhofer

IIS

Getting to "you are there"

- How can a listener tell he is not at an event or performance?
    - Basic sound quality is not realistic
        - Frequency response, distortion, transient response, SPL, …
        - Even modest consumer systems today can do pretty well on basic sound quality.
    - Sounds or ambience do not appear to come from realistic directions
        - This is what immersive audio can improve
    - Sound sources don't seem in the same room – you can't walk around them.
        - MPEG-I, game audio, wavefield synthesis are partial solutions
- Sounding better than "you are there" through production
- Visual/Audio perceptual fusion can help the sound image, but not for audio-only content.
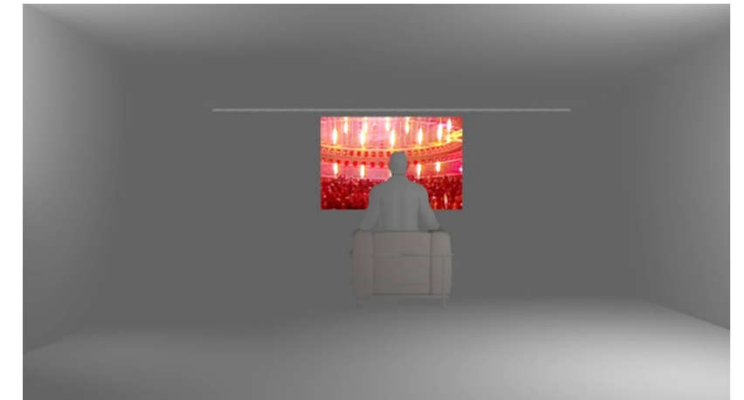

Stereo

Surround

Immersive

MPEG-H AUDIO

Fraunhofer

IIS
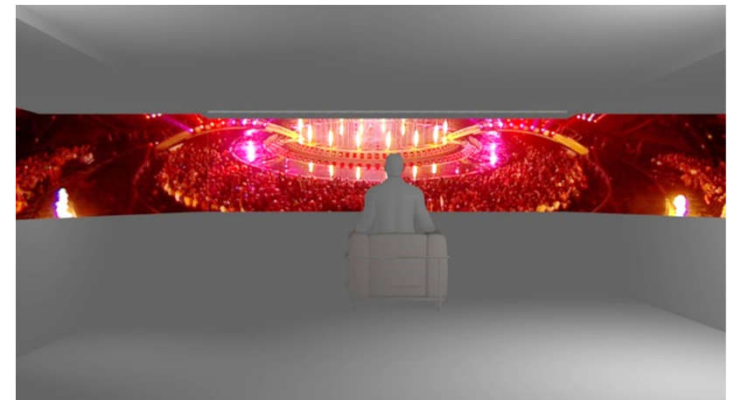
# Improving spatial resolution in channel-based audio – more speakers

- Monophonic reproduction: Audio appears to come from a single speaker.

- **Stereo**: relies on "phantom images" produced by panning a signal between speakers.

  - This is a psychoacoustic effect, sound waves from the two speakers are different than what would come from a sound source at the panned position

  - It works pretty well, so people don't usually think about how it works

- **Surround**: extends the stereo concept to more speakers horizontally

  - Panning is between two speakers except for divergence/spread effect

  - Typical layout: 5.1 or 7.1

- **Immersive**: Adds speakers above and perhaps below the horizontal plane, extending sound image to three dimensions

  - Panning is typically between three speakers (VBAP technique)

  - Typical layout: 5.1+4H or 7.1+4H, 22.2 for true envelopment

7.1+4H speaker layout

MPEG-H AUDIO

Fraunhofer

IIS
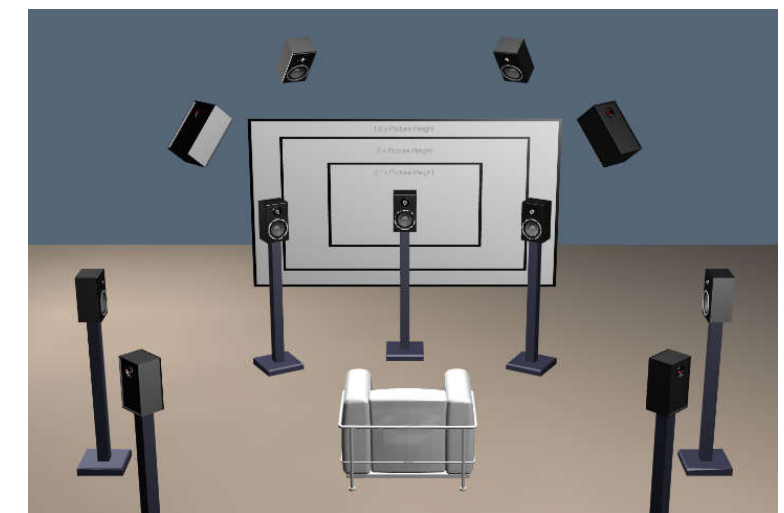
# Audio Objects - Moving panning and faders from the studio to the home

- Instead of panning tracks or stems to channels in the console or DAW, we send them separately to the home and pan (render) them there using positions and gains we send in object metadata

- Metadata can change every few milliseconds to move and fade objects dynamically - it's like an automation track.

- Objects allow interactivity and consumer adjustment (more later on that)

- Objects are decoupled from production or playback channel layout. In theory, this allows infinite spatial resolution, but this is limited by the playback speakers. Spatial resolution improvement with objects is primarily for cinema, not home playback

- Objects generally have less coding efficiency than channels and practical bandwidth to the home limits the number we can send. (Typical use cases are 16 or 24 objects or channels in total)
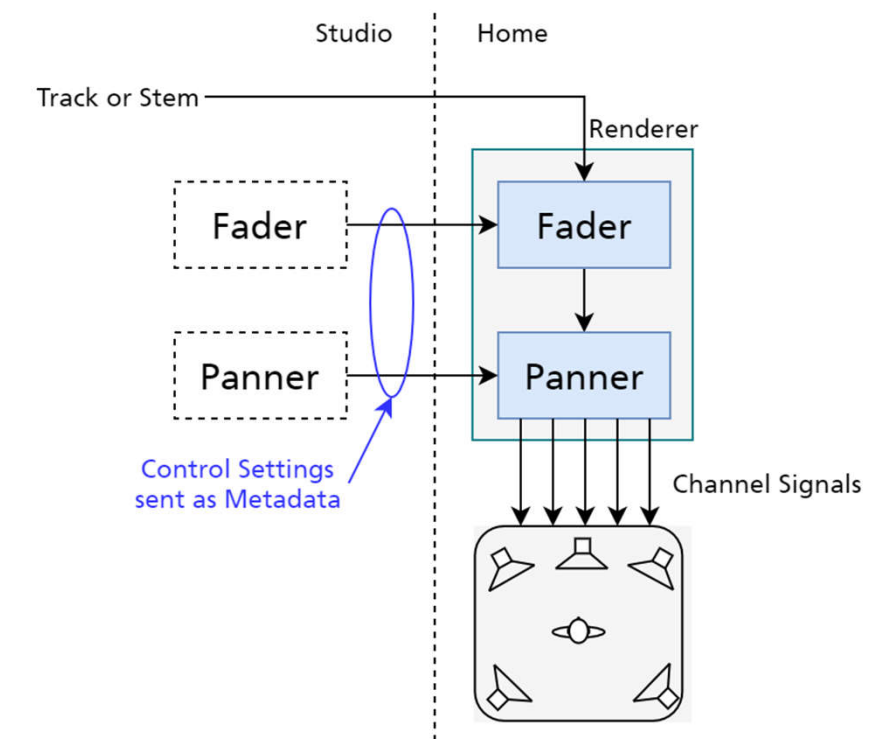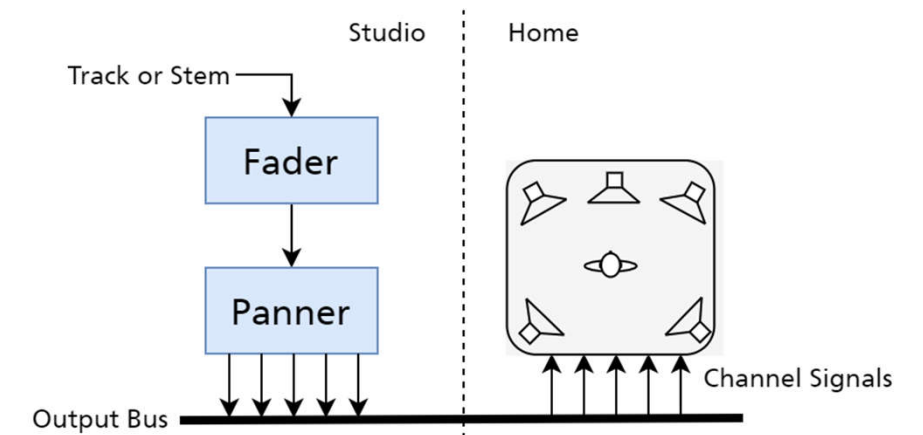
Ambisonics – Alternate technique that is theoretically interesting

- Instead of practical and psychoacoustic techniques, high-order Ambisonics attempts to recreate the sound field at a point theoretically exactly through decomposition of the acoustic wave equation using spherical harmonic basis functions

- Similar to Fourier or wavelet series but in the spatial domain, as basis functions are organized in a hierarchy of increasing spatial resolution

- Appealing in virtual reality use case for earphone playback since sound image can be easily rotated using public techniques

- Small (typically one foot or less) sweet spot on loudspeaker playback

- Requires new mixing techniques due to large number of signals (16, 25, 36, or 49) for each track for practical resolutions (basis orders)

- Thus useful more for VR "reality capture" than produced sound

- Ambisonics is only available in the older MPEG-H LC profile
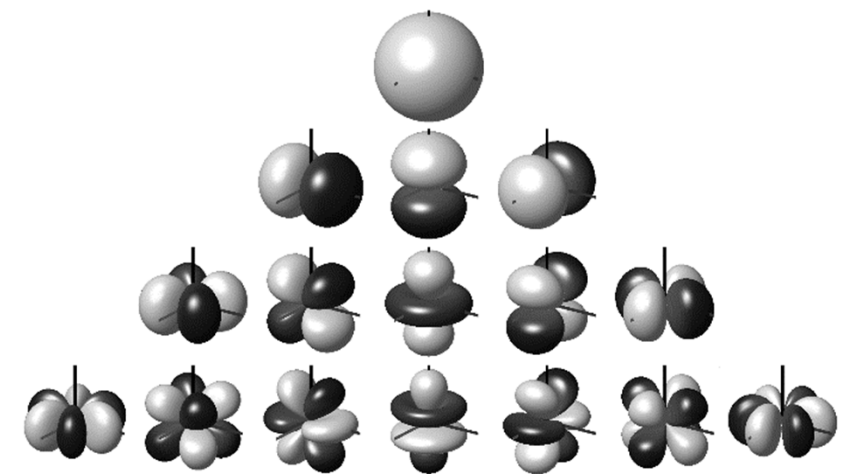
Ambisonic basis functions, order 0 to 3

MPEG-H AUDIO

Fraunhofer

IIS

Speaker "Virtualization" in Advanced Soundbars and Smart Speakers

- Consumers strongly favor convenient listening today as opposed to precise imaging

  - They want a one-box, one-minute install, not a 7.1+4 speaker upgrade project

- Acoustic and psychoacoustic techniques can be used to make a single soundbar or smart speaker create a sound image similar to an immersive speaker installation – termed "virtualization"

  - Image extends mainly to the sides of the listener. Rear imaging difficult without additional speakers.

  - Sound direction is not as precise as with traditional immersive speaker playback

  - "Upfiring" speakers that bounce sound off the ceiling are a simple example

  - Sophisticated implementations can provide a realistic and satisfying sound image

   www.mpeg-h.com   www.iis.fraunhofer.de/audio

MPEG-H AUDIO

Fraunhofer

IIS

Making immersive sound on earphones through binaural rendering

- Concept is to create the sound at each ear that would be heard from loudspeakers playing in a room by:
    - Playing back the sound through simulated speakers at channel or object positions
    - Adding room reflections of a simulated room to sound from each speaker
        - Rooms can be typical idealized room or measured ones
    - Accounting for the delay and attenuation of the head and ear in hearing these signals (HRTF)
        - Unfortunately, this varies between people due to different head and ear shapes
        - Can be measured (tedious lab or studio procedure) or estimated from 3D CAD model of head (Anthropometric HRTF)
        - 3D head model can be estimated from 2D photos
    - Accounting for changes in sound as head is turned (Head Tracking)
        - Needed to resolve front/back ambiguity

MPEG-H AUDIO

Fraunhofer
IIS
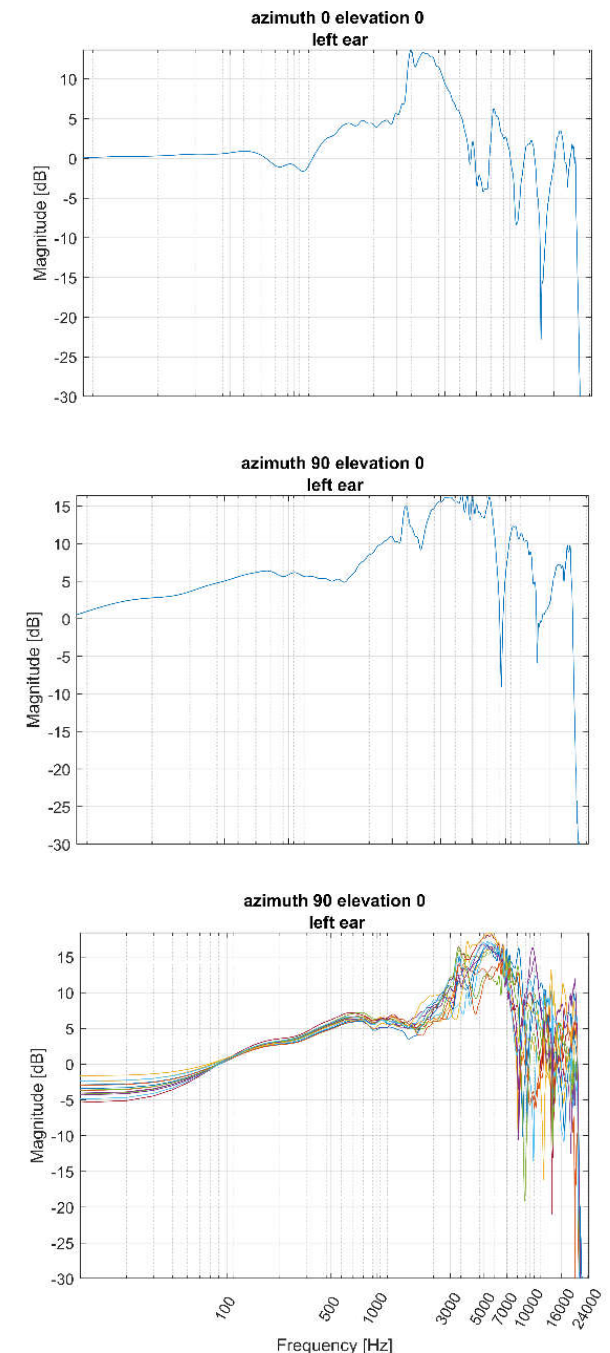
Upmixing

- Upmixing uses an algorithm to extract and distribute ambience information present in a stereo or surround recording into immersive channels

- Good upmixers product a nice effect that can add to the listening experience when true immersive content is not available.

- Upmixing can be done before encoding in the studio or during playback after decoding
  - High-quality upmixers are available for studio use (i.e. Illusonic) or for consumer products (i.e. Fraunhofer Symphoria)

- Upmixing does not know artistic intent or tell a story – An upmix will not purposely put a backing vocal behind you or make a tambourine fly over your head.

- Upmixing is not part of the MPEG-H standard and is supplied separately by Fraunhofer

CES 2013: Fraunhofer Symphoria upmix and rendering announced for Audi cars

MPEG-H AUDIO

Fraunhofer

IIS

# Not just immersion – solving two other big problems

- **Interactivity**
  - For TV, objects are used for different dialogue languages or biased commentary.
  - For music, objects can be used to change your perceived location – on stage with the conductor or row G at the symphony
  - Preset mixes of objects can be selected by the user or he can be given limited control over the mix.
    - **Turning the announcer up or down** is a highly-rated consumer feature for sports, for example.

- **Universal Delivery**
  - Allows sending the same bitstream to multiple devices – phone earbuds, tablets, or living room speakers.
  - Loudness and dynamic range adjusted to suit playback device
  - Energy preserving downmix and advanced downmix gain matrix improve downmix quality
  - Binaural rendering allows for headphone playback

www.mpeg-h.com   www.iis.fraunhofer.de/audio

MPEG-H AUDIO

Fraunhofer

IIS

Interactivity Examples (these options set by MPEG-H authoring)
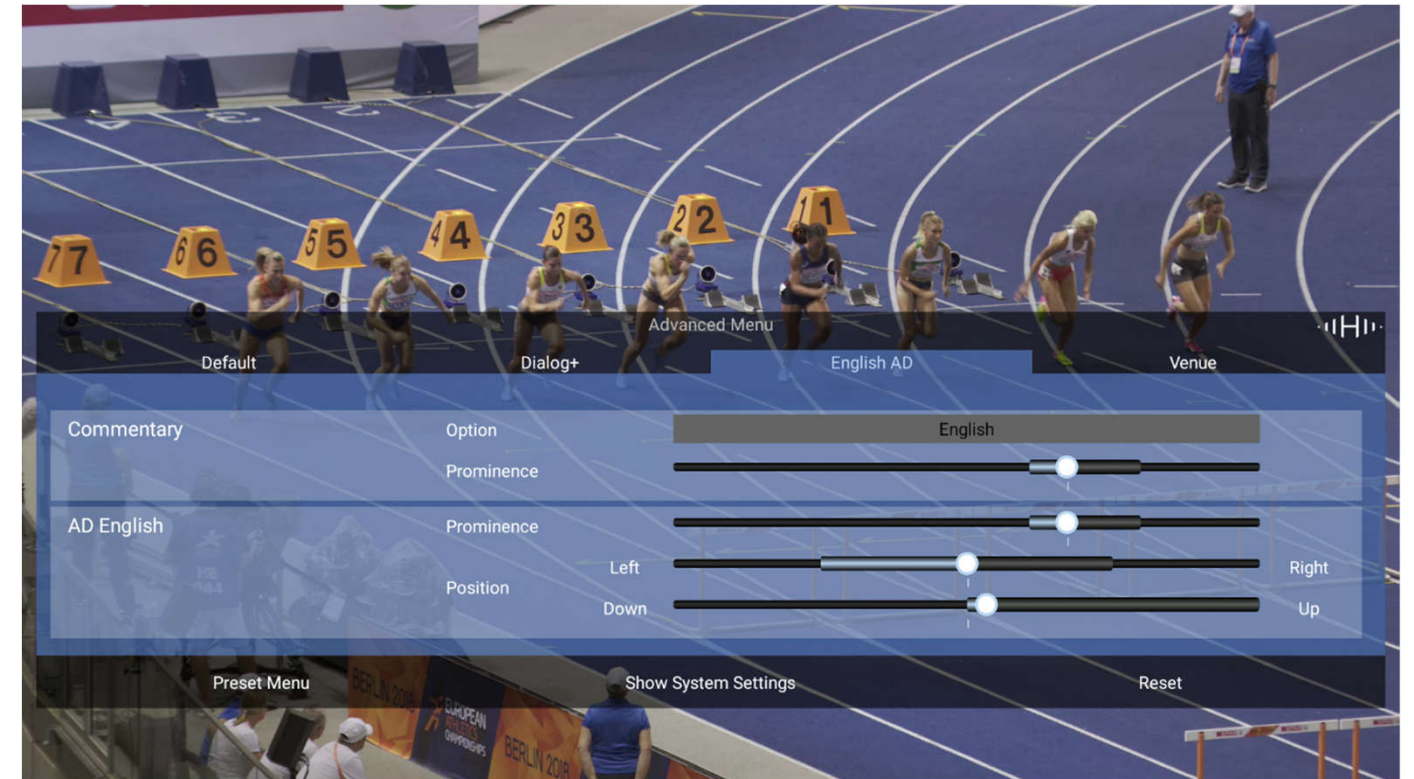
- 2019 LG TV with MPEG-H User Interface built-in

- French Tennis Open broadcast by France TV in MPEG-H

- User can select normal mix, boosted commentary, or just stadium ambience and PA

- Advanced user interface on Fraunhofer Android TV app

- From 2018 European Athletic Championships

- User can select desired language, level of announcer, or pan audio description narration

www.mpeg-h.com   www.iis.fraunhofer.de/audio

MPEG-H AUDIO

Fraunhofer

IIS

MPEG-H Universal Delivery

- Goal: Play the same MPEG-H bitstream on any device while delivering the best possible sound experience in each situation

- Issue: Differences in ambient noise level

  - Living room 30-40 dB SPL

  - Airliner 80-85 dB SPL

- Issue: Difference in device capabilities:

  - 12-15 mm phone or tablet speaker with 0.5mm x_{max}

  - Premium soundbar with 100 dB SPL

  - Enthusiast AVR system with 105-110 SPL speakers and amps

  - In-ear earphones 105-120 dB SPL / mW, 30 dB isolation

- Solution: Adaptation to the listening situation

  - Improved loudness control

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

MPEG-H AUDIO

Fraunhofer
IIS

Example – Listening on Airline Flight

- Average broadband noise level: 80-85 dBA SPL

  - Does not include PA or passenger conversation

- Maximum peak SPL available from earbuds for listening: 100 dBA legal limit (EN 50332)

- Simple earbuds – no acoustic isolation

  - Active headphones could provide ~20 dB improvement, sealed earbuds ~35 dB improvement

- Decoder target level: -16 dBFS

- Average loudness: 84 dB SPL (assuming peaks are not clipped or limited)

- Resulting signal to noise ratio: -1 to +4 dB

  - Extremely challenging use case

- -> Advanced Dynamic Range Control and post-processing required

MPEG-H AUDIO

Fraunhofer

IIS

Industry Dichotomy Collides In Converged Mobile Devices

Industry	Music, Radio	Film, TV
Traditional Loudness Strategy	"Pre-Normalize" (Server-side)	Loudness Metadata (Playback-side)
Misguided Goal	"Make it Louder"	Preserve Cinema Dynamic Range
Exceptions	Sound Check, Replay Gain	Fixed Metadata TV Plant
Typical loudness	-15 to -7 LKFS	-31, -24 LKFS
Developments	Streaming Services begin normalization to -14 and below	AES71, CTA-2075

   www.mpeg-h.com   www.iis.fraunhofer.de/audio

MPEG-H AUDIO

Fraunhofer
IIS

MPEG-D DRC Standard Enables Universal Delivery

- Joint development of Apple and Fraunhofer in MPEG audio subgroup

- Comprehensive metadata scheme integrated in xHE-AAC and MPEG-H

- Key features:

  - Metadata for track normalization and album normalization

  - Metadata for dynamic range control at decoder side

  - Mandatory peak limiter at decoder side

  - Encoded audio content stays untouched

- Flexible decoder configuration dependent on device type and listening condition

  - Loudness request (e.g. -31, -24, -16 LKFS)

  - Track normalization or album normalization

 - Selectable DRC profiles for playback optimization

   www.mpeg-h.com   www.iis.fraunhofer.de/audio

MPEG-H AUDIO

Fraunhofer

IIS

Concept of Loudness Normalization

■ Goal: Assure consistent loudness across programs and channels.

playback loudness

target loudness at decoder

No normalization

commercial

film

sports

With normalization

commercial

film

sports

true peak

loudness range

loudness [LKFS]

© 2020 Fraunhofer IIS   www.mpeg-h.com   www.iis.fraunhofer.de/audio

MPEG-H AUDIO

Fraunhofer

IIS
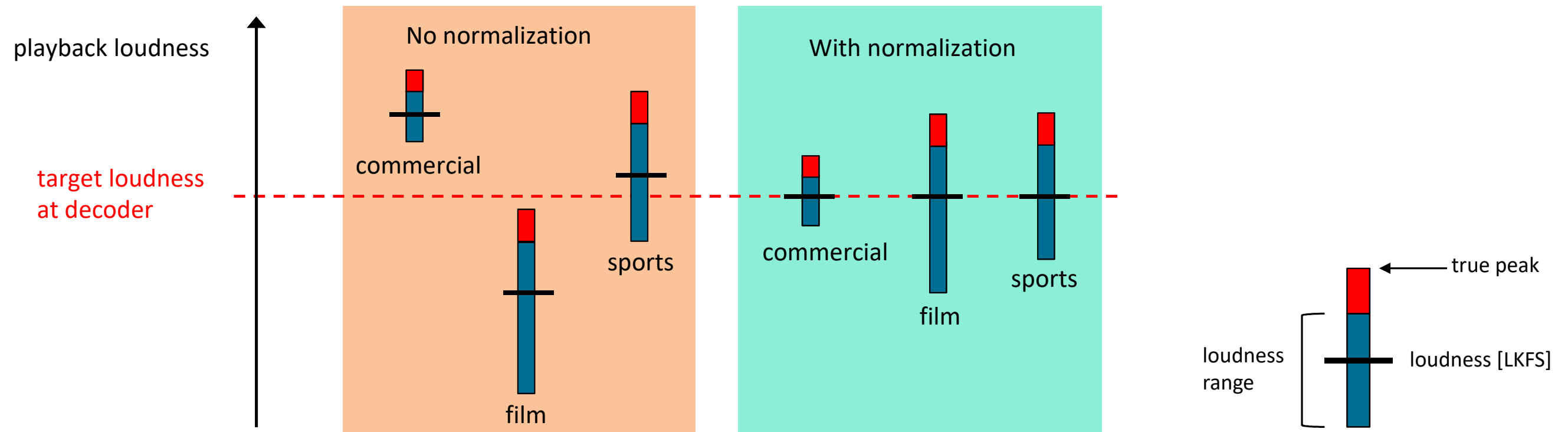
Dynamic Range Control

- **Goal**: Adjust the dynamics of the content as appropriate for the given listening situation.
- Typical relation of playback level and dynamic range for
  - different receiver types
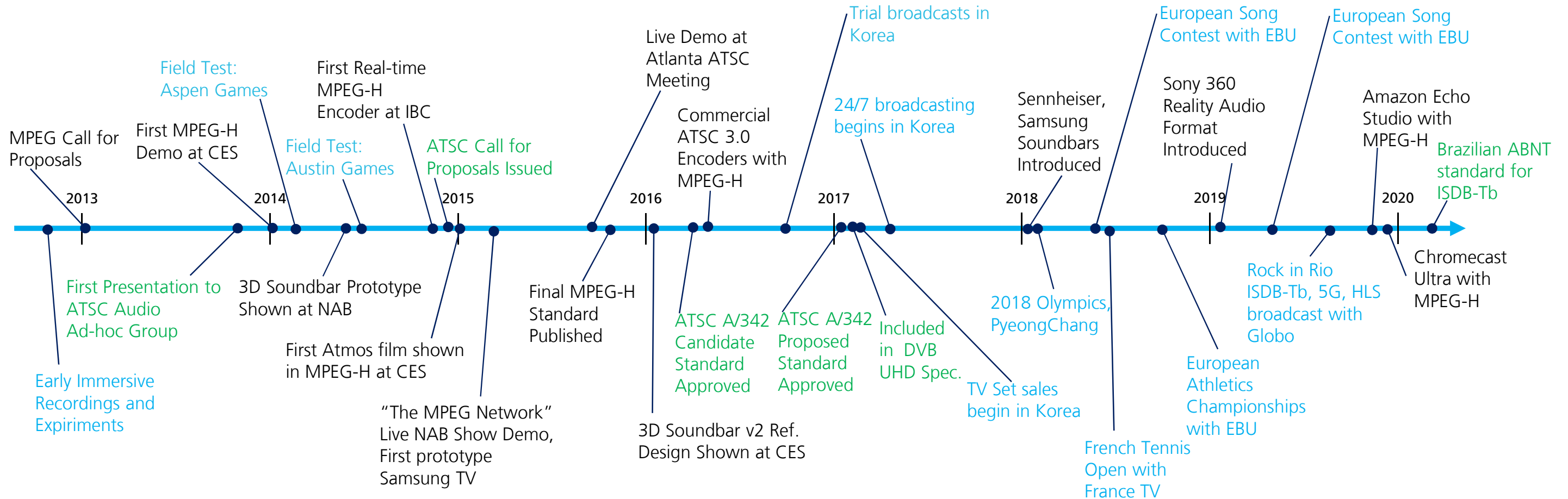 - listening conditions

   www.mpeg-h.com   www.iis.fraunhofer.de/audio
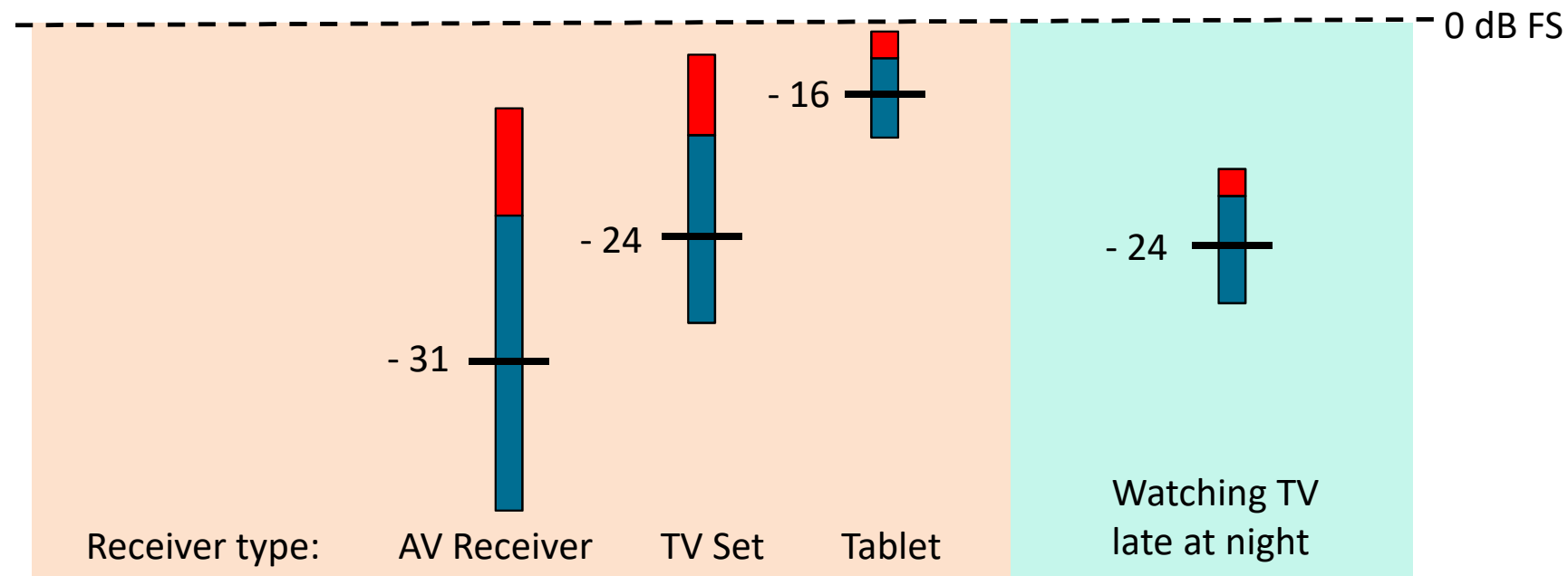
DEVELOPING THE MPEG-H AUDIO SYSTEM

	Format	Presentations	Language/Dialog		Program Log at each demo location		
				Remote Truck	MPEG Network	WMPG	
				Aspen	New York	Birmingham	
Intro	2.0	Broadcast	MOS		Opening Title		
	7.1 + 4H	Broadcast	ENG		Introducing Demonstrations		
	2.0	Broadcast	MOS		Show Title		
	5.1 + 4H + 2dynO	Broadcast	Music Only		Network ID Long		
Network Show	2.0 + 2statO	Broadcast, Dialog+	ENG		SportsTech Show - opening segment		
	5.1	Broadcast	ENG		PB: Big Air - Host Mix		
	2.0 + 2statO	Broadcast, Dialog+	ENG		SportsTech Show - setup of H mix		
	HOA + 1statO	Broadcast, Dialog+, Live	ENG		PB: Big Air - MPEG-H Version		
	2.0 + 2statO	Broadcast, Dialog+	ENG		SportsTech Show - half-pipe setup		
	5.1	Broadcast	ENG		PB: Half-pipe - Host Mix		
	2.0 + 2statO	Broadcast, Dialog+	ENG		SportsTech Show - setup of half-pipe live H mix		
	5.1+4H + 3statO + 1dynO	Broadcast, Dialog+, Live	ENG(Network), ENG(Venue), NOR	Half-pipe (live)	Cut to Aspen - live mix of half-pipe		
	2.0 + 2statO	Broadcast, Dialog+	ENG		SportsTech Show - throw to commerical		
Local Break	5.1	Broadcast	ENG		National Spot - AAA		
	5.1+4H	Broadcast	ENG		Network Cover: Technicolor Promo	WMPG ID - WeatherCenter 84	
	3 x 2.0statO	Broadcast, Dialog+	ENG, SPA, CHI			Local spot #1 - Crown Nissan	
	3 x 2.0statO	Broadcast, Dialog+	ENG, SPA, CHI			Local spot #2 - airbag lawyer	
	5.1+4H + 2dynO	Broadcast	Music Only		Network ID Short		
Network Show	2.0 + 2statO	Broadcast, Dialog+	ENG		SportsTech Show - setup NASCAR		
	5.1 + 5.0statO + 4 x 1.0statO	Broadcast	ENG, ITA		PB: Nascar		
	2.0 + 2statO	Broadcast, Dialog+	ENG		SportsTech Show - close		
Break	HOA + 2statO	Broadcast	ENG, CHI		National Spot - Qualcomm: SnapDragon		

2015: Program Log of Demonstration Network, 13 formats

2011: Audio Objects used to adjust commentary level

MPEG-H AUDIO

© 2020 Fraunhofer IIS   www.mpeg-h.com   www.iis.fraunhofer.de/audio

Fraunhofer
IIS

Timeline of MPEG-H Development

Standards: MP3: 1992, AAC: 1997, AAC-LD: 1999, HE-AAC: 2003, HD-AAC: 2006, HE-AACv2: 2006, MPEG Surround: 2007, AAC-ELD: 2009, xHE-AAC: 2012, EVS: 2014, MPEG-H: 2015

System Issues we had to solve to deploy MPEG-H

- No way to experience immersive audio without a 10 or 12 speaker AVR setup
  - Development of practical high-performance 3D soundbar
  - Extension of Fraunhofer Cingo binaural rendering to immersive
- Production and post systems had no way to carry metadata with the audio, as needed for dynamic objects
  - MPEG-H Production Format: PCM audio plus a time code-like "control track"
- No commercial TV consoles could mix immersive
  - Development of authoring and monitoring unit to adapt existing consoles for MPEG-H production
- How to enable listeners to select mix presets or adjust objects
  - Distributed MPEG-H User Interface with control packets sent over HDMI or S/PDIF

MPEG-H AUDIO

Fraunhofer

IIS

The "3D Soundbar" makes true immersive sound possible as a "one minute install"

- 2014: First concept prototype
- Loudspeakers in a frame surrounding the TV
- Hundreds of speakers, complex DSP (not practical to manufacture)

- 2016: Enclosure similar to traditional soundbar
- 14 Speakers

- 2019: Launch of Sennheiser AMBEO soundbar with MPEG-H playback

Carrying metadata in the time code-like Control Track

- Design approach: "metadata modem" makes an analog signal than can be carried in spare audio channel similar to time code.

- No need to configure and maintain data mode settings on audio channel (as required for carrying compressed audio in AES or SDI)

- No compensating video frame delays needed

- Survives sample rate or time base conversion and gain changes

- Can be edited as a normal audio track in video editors such as Adobe Premiere

    www.mpeg-h.com   www.iis.fraunhofer.de/audio
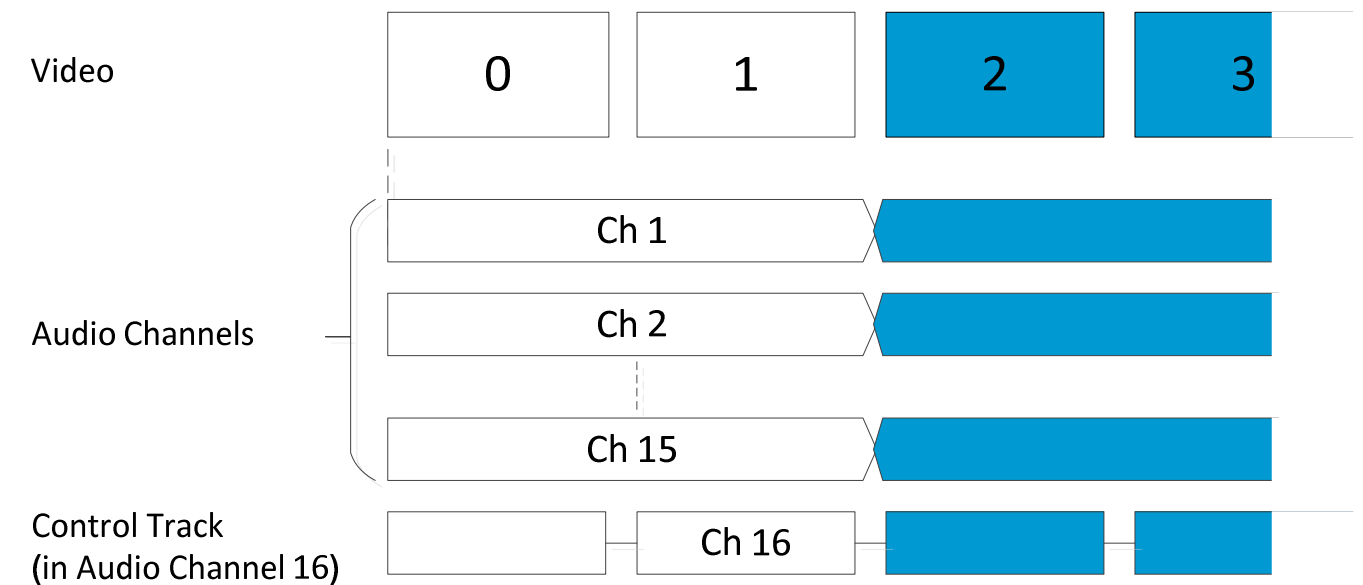
Mapping audio objects to channels

- As programs become more complex, industry conventions on channel assignment will break down

- No longer a question of "is the center channel on SDI channel 4 or channel 6?"

- Channel 14 may be "away team commentary" on one show and "Spanish Dialogue" on another

- This problem envisioned in 2014 when we designed the system and explained in our 2015 Facilities Paper

- MPEG-H Control Track automatically maps SDI channels to MPEG-H channels or objects and provides text labels

1. **Track allocations**

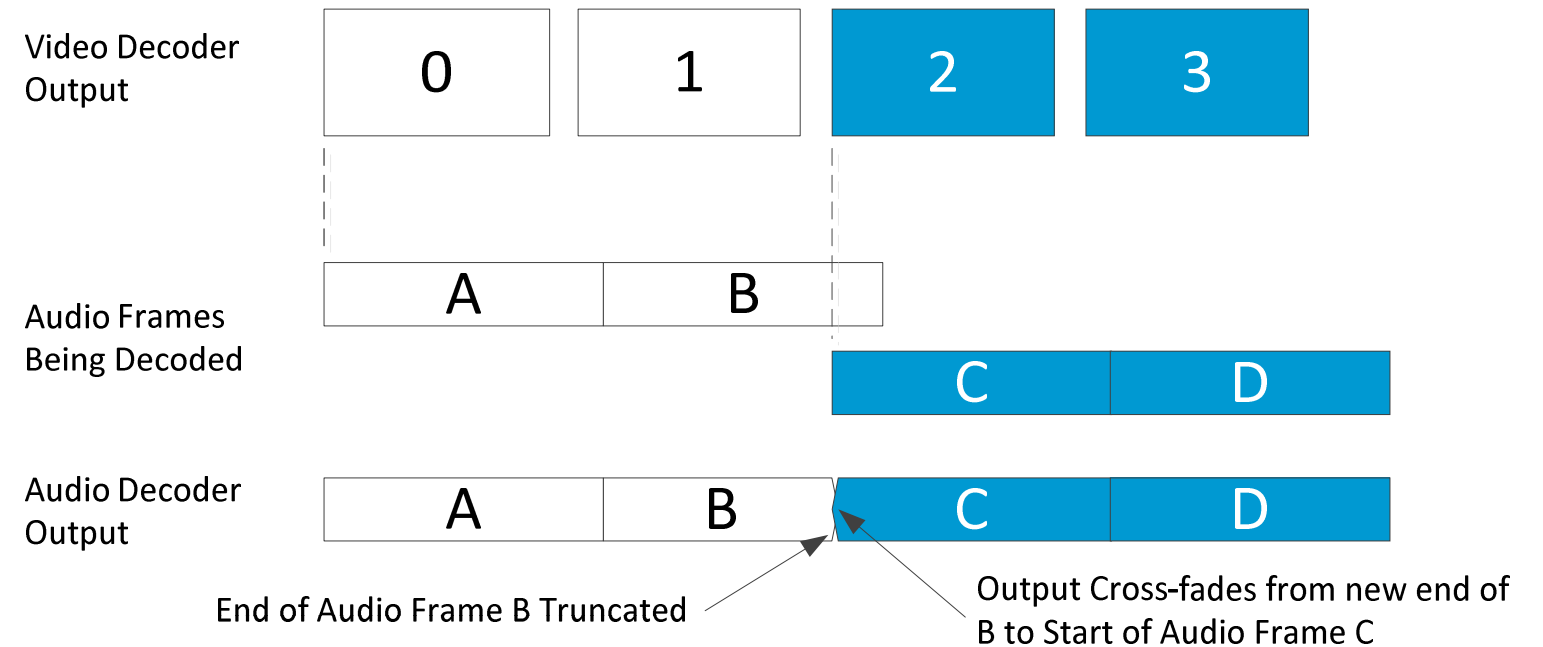Audio tracks in multichannel formats sl
recorded on all non-available tracks.

1	2	3	4
L	R	C	LFE (if available)

Where:

Symbol	Audio Signal
L	Left
R	Right
C	Centre
LFE	Low frequency effects
LS	Left surround
RS	Right surround
MS	Monophonic surround
A	Left (two channel stereo)
B	Right (two channel stereo)

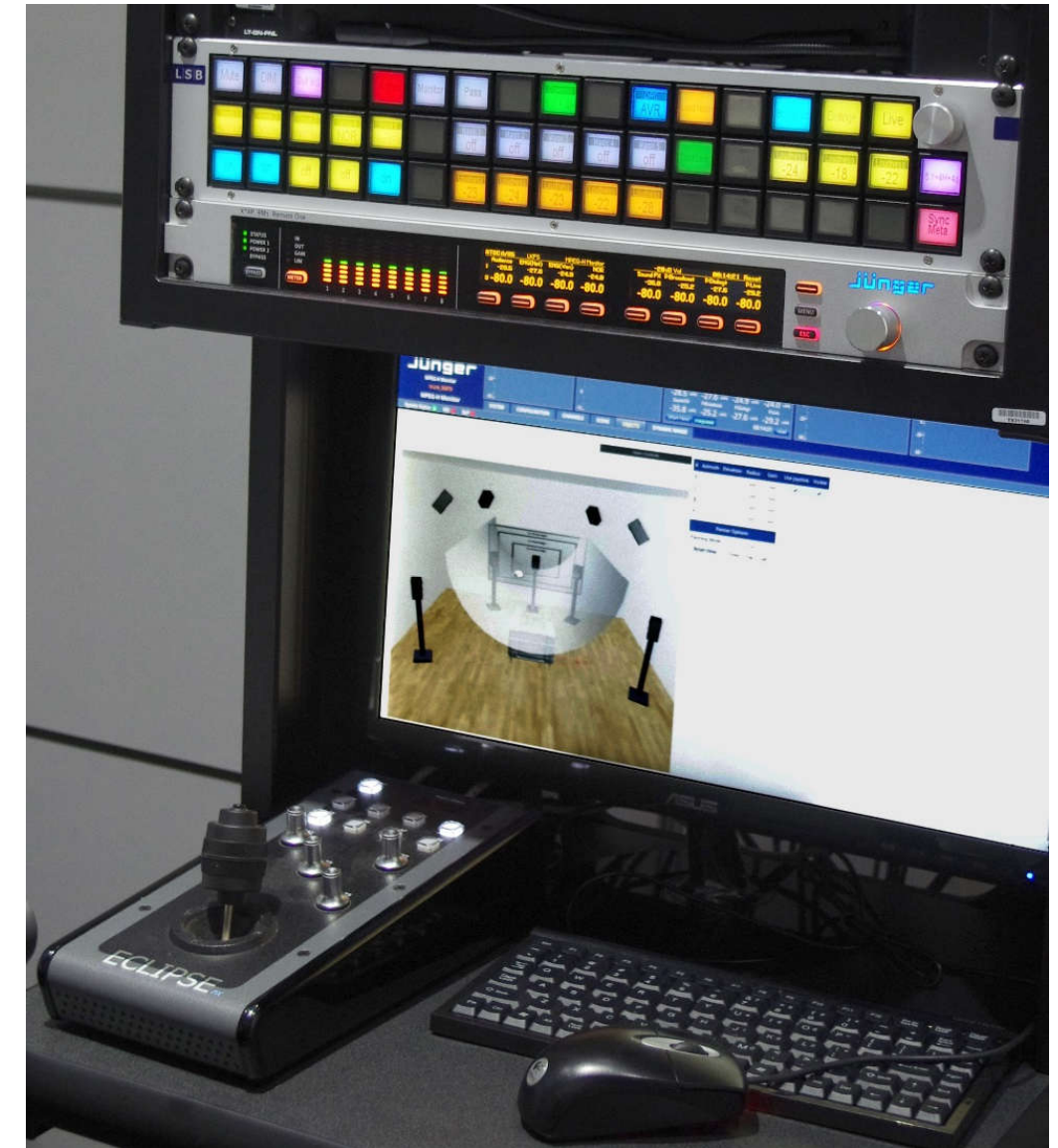   www.mpeg-h.com   www.iis.fraunhofer.de/audio

Splicing of MPEG-H Audio Streams at Video Frame Boundaries

- Control track and PCM audio may be cut at any frame

- MPEG-H *Encoded* audio partitioned into audio frames containing one audio scene or channel configuration

- Audio and Video frames align once every few hours

- Solution: Send additional audio frame at video cut and cross-fade

- Eliminates loss of coding efficiency from locking audio frame rate to video frame rate

Video Decoder Output

| 0 | 1 | 2 | 3 |

Audio Frames Being Decoded

| A | B | | C | D |

Audio Decoder Output

| A | B | C | D |

End of Audio Frame B Truncated

Output Cross-fades from new end of B to Start of Audio Frame C

© 2020 Fraunhofer IIS   www.mpeg-h.com   www.iis.fraunhofer.de/audio

MPEG-H AUDIO

Fraunhofer

IIS

Adapting Live TV Consoles With the AMAU

- Situation:
    - Broadcast consoles limited to 5.1 mix busses
    - Plug-ins are available only on an outboard PC
    - Monitor Control limited to 5.1
    - Loudness Monitoring limited to 5.1
- Solution: make an accessory box that adds these features to an existing console
    - MPEG-H **A**udio **M**onitoring and **A**uthoring **U**nit
    - Developed in collaboration with Junger Audio

© 2020 Fraunhofer IIS   www.mpeg-h.com   www.iis.fraunhofer.de/audio

MPEG-H AUDIO

Fraunhofer

IIS

MPEG-H audio playback may be distributed over multiple devices

- A common scenario is the display of the MPEG-H user interface on a source device such as a TV or STB, while the audio decoding is done on a Soundbar or AVR

- User interaction data is sent in the MPEG-H bitstream over the HDMI interface to the Soundbar or AVR for processing by the MPEG-H decoder

- Bitstream is carried in MHAS, the native transport format of MPEG-H

- Transport specified in HDMI, IEC and CTA standards

© 2020 Fraunhofer IIS   www.mpeg-h.com   www.iis.fraunhofer.de/audio
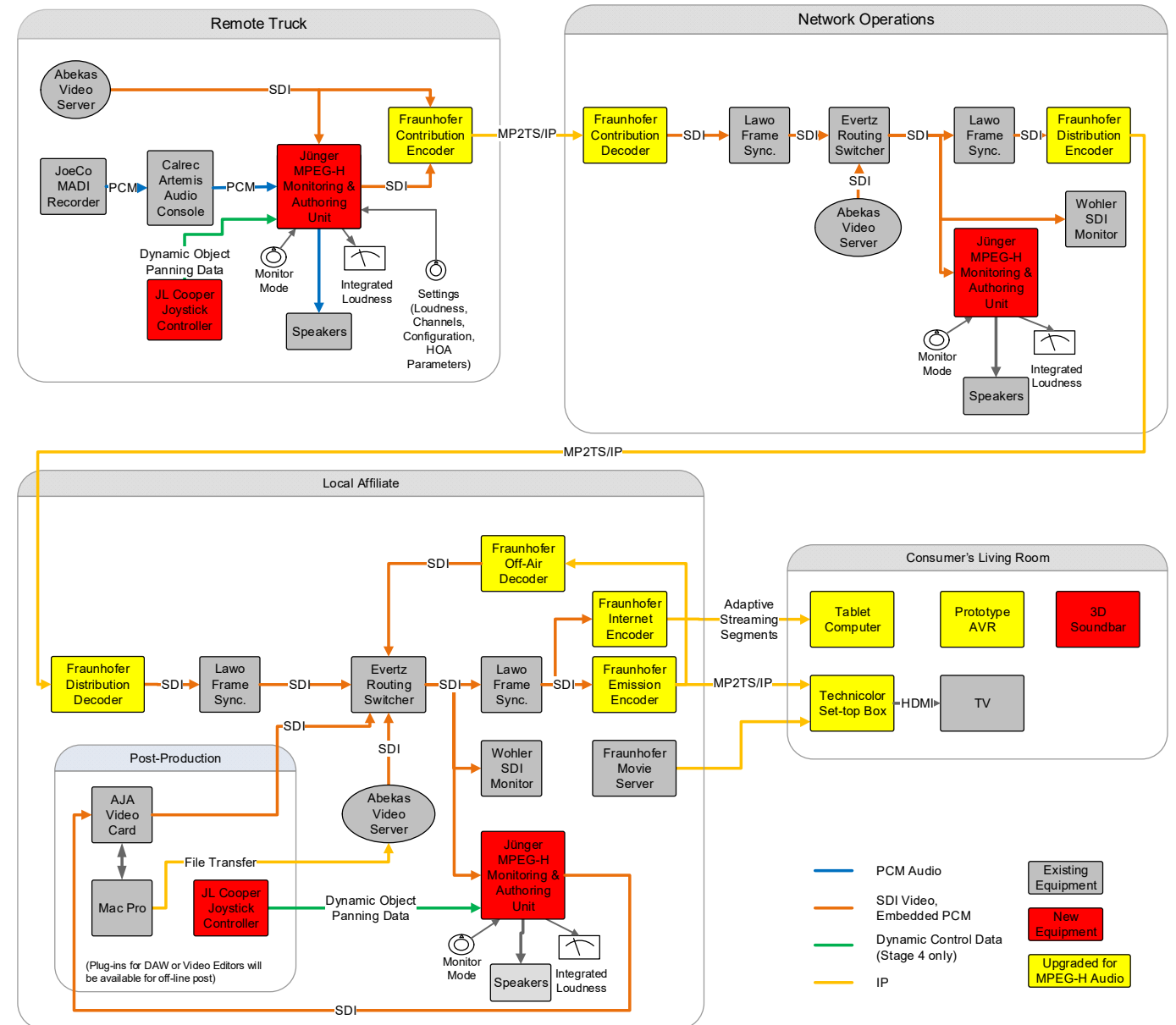
MPEG-H AUDIO

Fraunhofer

IIS
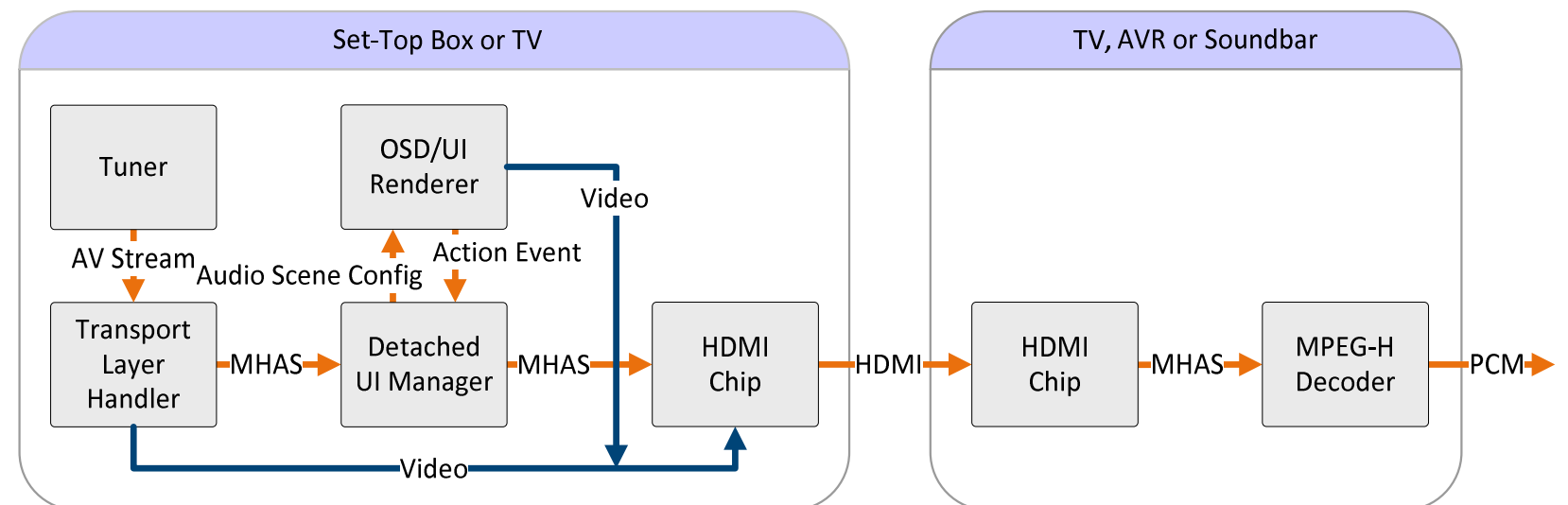
MPEG-H Delivery to the Home - features

- MPEG-H works over HDMI 1.4 HBR mode for forward and ARC connections,

- No eARC needed

- No transcoding needed

- Distributed User Interface concept allows use of source remote (STB/DMA), not sink device (Soundbar/AVR) remote

- MPEG-H is fully specified in CTA-861 G, IEC 61937-13, HDMI. Bits reserved for other flavors

- Lip sync managed through certification to +10/-20 ms

© 2020 Fraunhofer IIS   www.mpeg-h.com   www.iis.fraunhofer.de/audio

MPEG-H AUDIO

Fraunhofer

IIS

2015: Building a testbed to prove the system

- The "World's Most Complex TV Network" from an audio standpoint – 13 formats

- Constructed in four rooms:
  - Remote Truck
    - Live mixing of pre-recorded microphone signals
  - Network Operations Center
    - Playout of 13 formats from automation playlist
    - Insert of live feed from truck
  - Local Affiliate
    - Insertion of local commercials from automation playlist
    - Editing of local spots and sports highlights in Premiere
  - Consumer Living Room
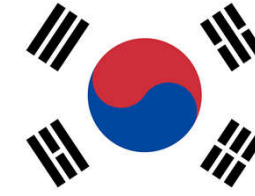   - Playback on Technicolor STB and Fraunhofer 3D Soundbar

MPEG-H TESTS AND ADOPTION

© 2020 Fraunhofer IIS   www.mpeg-h.com   www.iis.fraunhofer.de/audio

MPEG-H AUDIO

Fraunhofer

IIS

First TV market using MPEG-H
Terrestrial UHDTV Service in South Korea

- First and currently only regular terrestrial UHDTV service worldwide using a Next Generation Audio Codec

  - Regular service started in May 2017, nationwide service in 2020

- MPEG-H Audio is the only audio codec specified for the broadcast services

  - TV sets and STBs as well as encoders support the full feature set of the MPEG-H Audio:

    - up to 32 elements for the transmission and simultaneous decoding of 16 elements

    - Advanced accessibility and personalization options

MPEG-H Audio adoption in Brazil
Selected for ISDB-T broadcast

- SBTVD Forum has selected MPEG-H Audio for enhancing the terrestrial broadcast over ISDB-Tb in Brazil with immersive and personalized sound.

- MPEG-H Audio - the Next Generation Audio system with the most advanced personalization and accessibility features

- Availability of production and broadcast equipment from 3rd party companies essential for fast adoption

- Broadcasters can now use MPEG-H Audio in simulcast with existing AAC system

- First live production with MPEG-H Audio conducted by TV Globo during Rio de Janeiro Carnival 2019

Rock in Rio 2019
First broadcast in MPEG-H Audio over ISDB-Tb

- Globo, the largest media group in Brazil successfully tested MPEG-H Audio during Rock in Rio over:
  - ISDB-Tb terrestrial broadcast
  - 5G broadcast (experimental UHF channel)
  - HLS streaming

- Globosat sound engineers have produced the immersive mix in 5.1+4H.

- Additional mono and stereo stems for ambience, instruments or vocals were used to enable personalization features.

- Visitors at Globoplay booth had the option to experience the immersive sound and interact with the content.

https://www.audioblog.iis.fraunhofer.com/globo-rockinrio-mpegh-isdbtb-5g

MPEG-H AUDIO

Fraunhofer
IIS

Eurovision Song Contest 2019
MPEG-H Live production and broadcast

- Parallel MPEG-H production of the event:

  - Live mixing of more then 100 mic feeds

  - Additional microphones placed on the ceiling of the arena for better ambience capturing

  - Immersive mix using 5.1+4H together with 5 additional objects for 5 languages

  - Broadcasted live via the Eurovision FINE network to Geneva and Madrid

- MPEG-H Audio partners:

  - **ATEME, Jünger Audio, Sennheiser, Solid State Logic and TELOS ALLIANCE.**

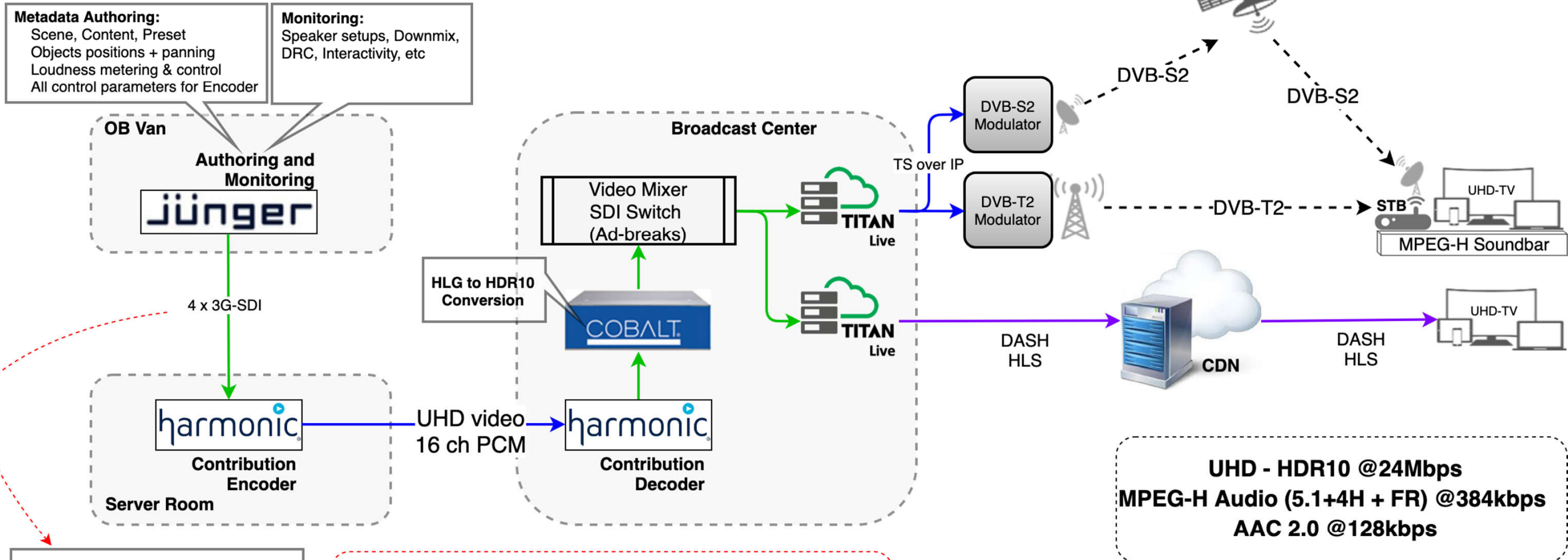*https://tech.ebu.ch/news/2019/05/immersive-and-personalized-audio-at-the-eurovision-song-contest*

© 2020 Fraunhofer IIS   www.mpeg-h.com   www.iis.fraunhofer.de/audio

MPEG-H Audio during the French Tennis Open 2019
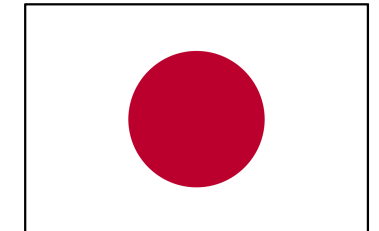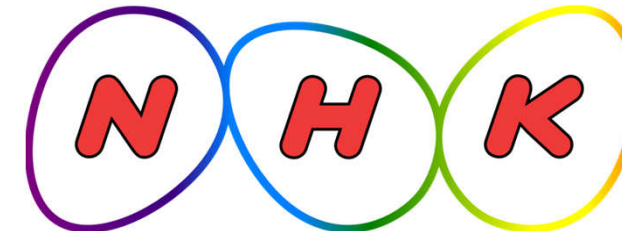## Successful terrestrial and satellite reception

MPEG-H Audio
China and Japan

- China is in the final stage to standardize the China 3D Audio transmission codec for UHDTV services based on MPEG-H

- Fraunhofer IIS and its partners Jetsen/Auro, Jünger, Kuvision, Hisilicon and Skyworth Digital have already been put to the test with a trial at CCTV during the 2018 soccer World Cup

- NHK is testing MPEG-H Audio for their next generation digital terrestrial broadcasting and future services in Japan.

Object-based sound production and broadcast systems
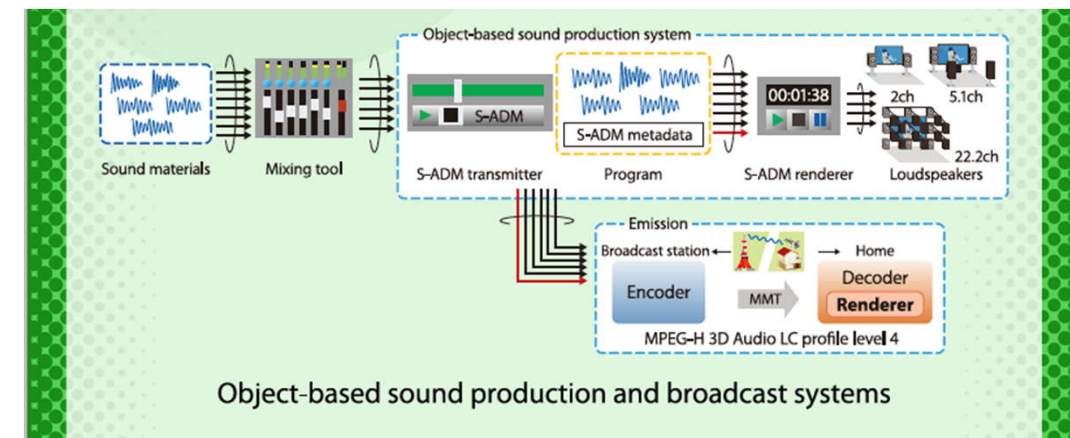
https://www.nhk.or.jp/strl/nab2019/05_NAB2019.pdf

© 2020 Fraunhofer IIS   www.mpeg-h.com   www.iis.fraunhofer.de/audio

Official Launch Event (NYC) – Oct 15, 2019

- Guests : 200+ media and influencers

- Key messages:

  - Partnering with the entire music industry to deliver a new music experience

  - (4) service partners started 360RA service from Oct 28th, 2019

  - Works across both Headphones and Speakers

Press event

Panel Discussion

Live performance

[LIST OF PARTNERS]
Streaming Services
Amazon Music HD
Deezer
nugs.net
TIDAL

Music Labels
Sony Music Entertainment
Universal Music
Warner Music

Platforms
Amazon Alexa
Google Chromecast

Chipset
Qualcomm Technologies International, Ltd.
NXP Semiconductors N.V.
Media Tek Inc.

Additional Partners
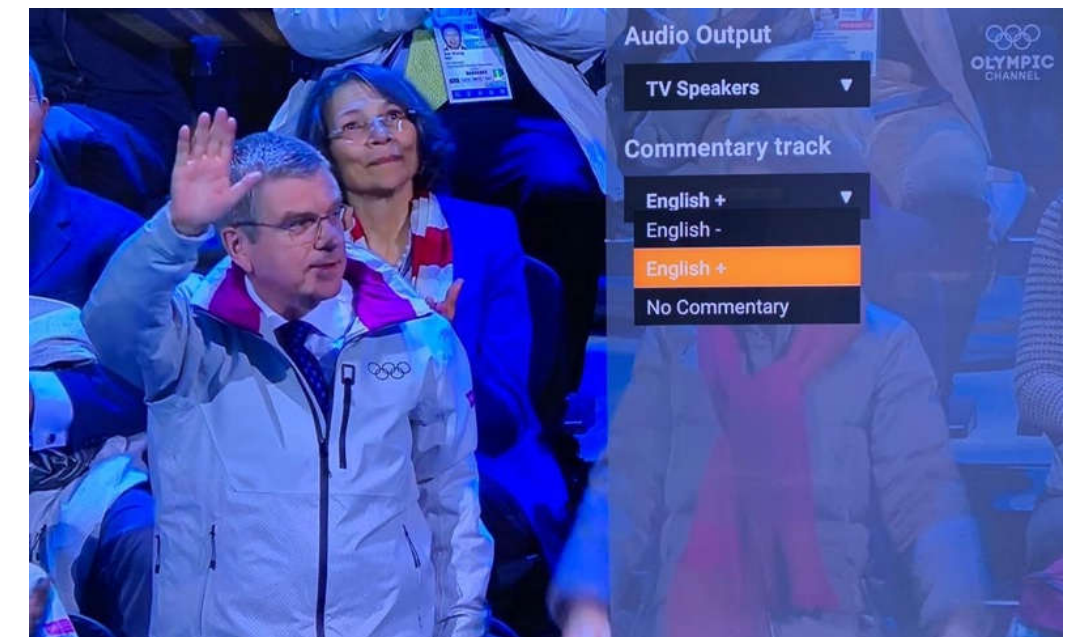Live Nation
Fraunhofer
Napster

MPEG-H Audio live streaming
Youth Games January 2020

- The Opening Ceremony for the Youth Olympic Games (Lausanne, Switzerland) was streamed live using MPEG-H Audio on the Olympic Channel Apps for Android TV and Swisscom Android Set-top boxes

- OBS has prepared the interactive and immersive audio for live and VOD:

  - Personalization – Dialogue Enhancement and Venue Presets

  - Immersive Audio – Passthrough to Soundbar

https://tech.ebu.ch/contents/publications/next-generation-audio-nga-at-the-olympics
https://play.google.com/store/apps/details?id=com.olympicchannel.olympics&hl=en

www.mpeg-h.com   www.iis.fraunhofer.de/audio

MPEG-H AUDIO

Fraunhofer

IIS

MPEG-H PLAYBACK DEVICES

www.mpeg-h.com   www.iis.fraunhofer.de/audio

MPEG-H AUDIO

Fraunhofer

IIS

MPEG-H Audio Deployment
Support in TV sets

■ LG and Samsung TVs support MPEG-H Audio
since 2017

■ LG enabled native support for MPEG-H User
Interface since 2019 models

© 2020 Fraunhofer IIS   www.mpeg-h.com   www.iis.fraunhofer.de/audio

MPEG-H AUDIO

Fraunhofer
IIS

MPEG-H Audio Deployment
Support in Soundbars

- Sennheiser AMBEO Soundbar (*Best of Show: CES 2018 and 2019*)

  - https://de-de.sennheiser.com/ambeo-soundbar

  - *"Using the latest virtualization technology jointly developed with Fraunhofer, the AMBEO Soundbar captures knowledge of your room size and its reflective surfaces, adapting the acoustics to fit your individual environment.*

© 2020 Fraunhofer IIS   www.mpeg-h.com   www.iis.fraunhofer.de/audio

MPEG-H AUDIO

Fraunhofer

IIS

MPEG-H Audio Deployment
Support in Soundbars

- At CES 2018 Samsung released two Soundbar models with an MPEG-H Audio decoder integrated:

  - 7.1.4 Ch Soundbar HW-N950

  - 5.1.2 Ch Soundbar HW-N850

- With MPEG-H bitstream input over HDMI, all audio channels are available in the Soundbar for reproducing a truly immersive experience.

Source: https://www.samsung.com/nz/audio-video/hw-n950/

© 2020 Fraunhofer IIS   www.mpeg-h.com   www.iis.fraunhofer.de/audio

MPEG-H AUDIO

Fraunhofer

IIS

MPEG-H Audio
360 Reality Audio: Amazon Echo Studio

- In November 2019, Amazon launched a new immersive smart speaker, the Echo Studio, which plays music from Amazon Music HD in the 360 Reality Audio format based on MPEG-H

© 2020 Fraunhofer IIS   www.mpeg-h.com   www.iis.fraunhofer.de/audio

MPEG-H Audio – Chromecast MPEG-H Pass-through

- Google Casting with MPEG-H pass-through support is available today, Cast built-in to follow soon
- https://developers.google.com/cast/docs/media

 www.mpeg-h.com   www.iis.fraunhofer.de/audio

MPEG-H AUDIO

Fraunhofer

IIS

MPEG-H Audio
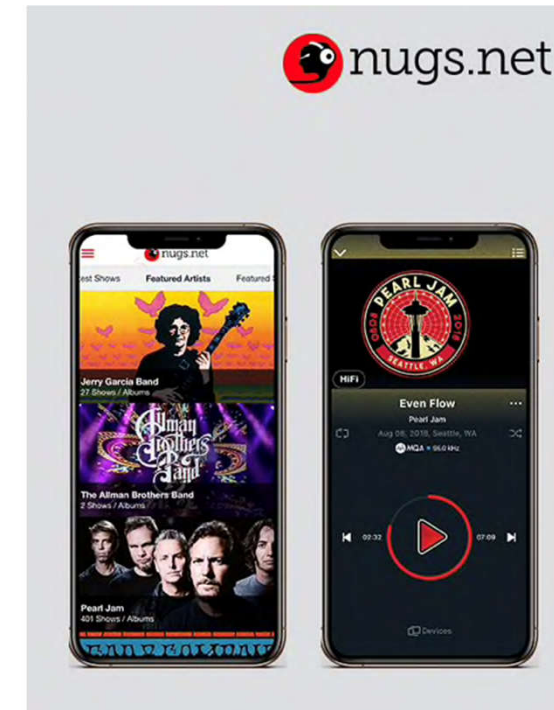## 360 Reality Audio Mobile Apps

360 Reality Audio music can be enjoyed by consumers using mobile apps from :
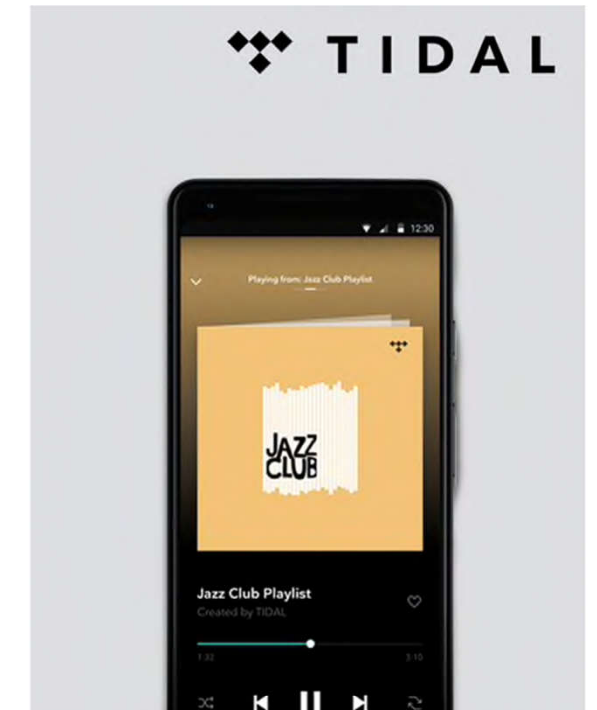
- Tidal
- Deezer
- Nugs.net

360 REALITY AUDIO
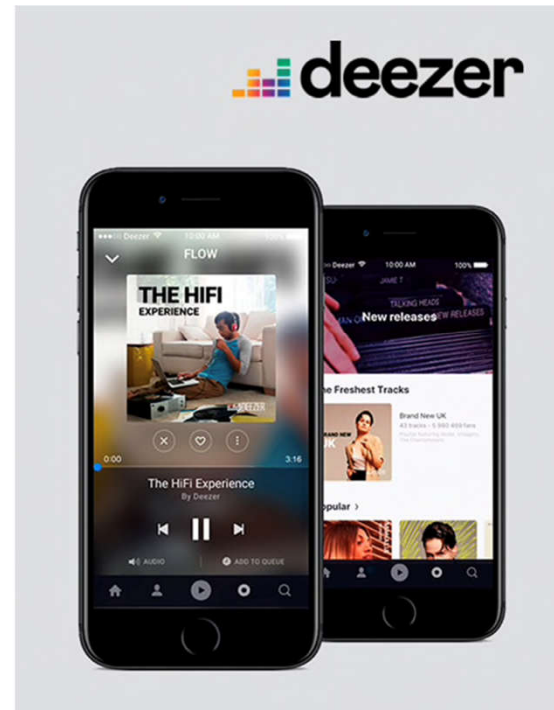
deezer

Hear your music as it's meant be heard. 30 days free trial available. [4]

nugs.net

Over 10,000 professionally recorded live concerts streamed or offline. 30 days free trial available.

TIDAL

All your favorite music. Best sound quality available. Start your 30 day trial now. [5]

https://www.sony.com/electronics/360-reality-audio

MPEG-H AUDIO

Fraunhofer

IIS

MIXING AND MASTERING IN MPEG-H

Fraunhofer Main Listening Room "Mozart"

Fraunhofer Project Studio "Bach"

© 2020 Fraunhofer IIS   www.mpeg-h.com   www.iis.fraunhofer.de/audio

The MPEG-H TV Audio system logo is a trademark of Fraunhofer IIS and is registered in Germany and other countries

MPEG-H AUDIO

Microphone Techniques

- **Spot microphones for "multitrack"-style recording or dialogue work as they always have**
  - Usual cautions against bleed when dynamically panned, just as for stereo or surround

- **Ambience Capture Options**
  - Discrete (ordinary) microphones widely separated – as in an arena
  - Discrete microphones arranged in a tree configuration, typically 0.5 to 2 meter extent
  - Purpose-built immersive microphones – compact and sometimes costly
    - Usually contained in a blimp or in the same mic head
    - Schoeps ORTF 3D
    - Eigenmic
    - Ambeo Mic

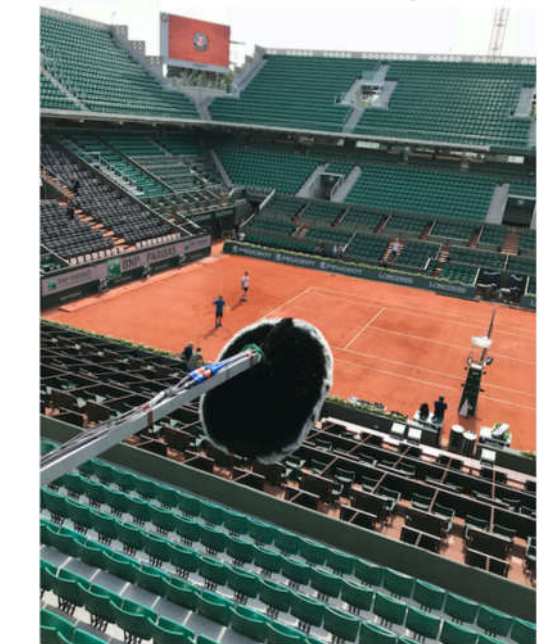- **Mic technique could be another whole webinar…**

Hamasaki square for ambience capture at Eurovision Song Contest 2019

2L-Cube mic tree for in-the round location recording of classical music – Lindberg, 2012

Schoeps ORTF 3D mic array, in windscreen at French Tennis Open

© 2020 Fraunhofer IIS www.mpeg-h.com www.iis.fraunhofer.de/audio

MPEG-H Production offers two workflows: #1 - Live

- Live Production using TV consoles and real-time TV equipment:

- MPEG-H Audio Monitoring and Authoring Units from Linear Acoustic and Junger Audio:

  - Authoring of metadata

  - Loudness metering

  - Panning of audio objects to track action

  - Interfacing to traditional audio consoles

- MPEG-H encoding built into broadcast video encoders from Ateme, Ericsson, DS broadcast, Kai Media, others

- MPEG-H Production Format stores metadata in time-code-like signal on spare SDI audio channel, allows carriage through technical plant and video editing without any changes
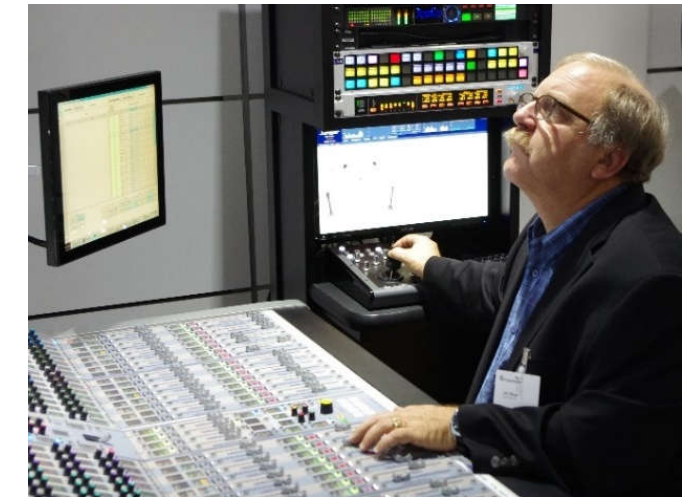
Panning an object live at NAB 2015 with Junger authoring and monitoring unit and Calrec console

Authoring Units from Linear Acoustic and Junger

Mixing live on location at 2019 European Song Contest broadcast using SSL console in container

Live Production Workflow

© 2020 Fraunhofer IIS   www.mpeg-h.com   www.iis.fraunhofer.de/audio
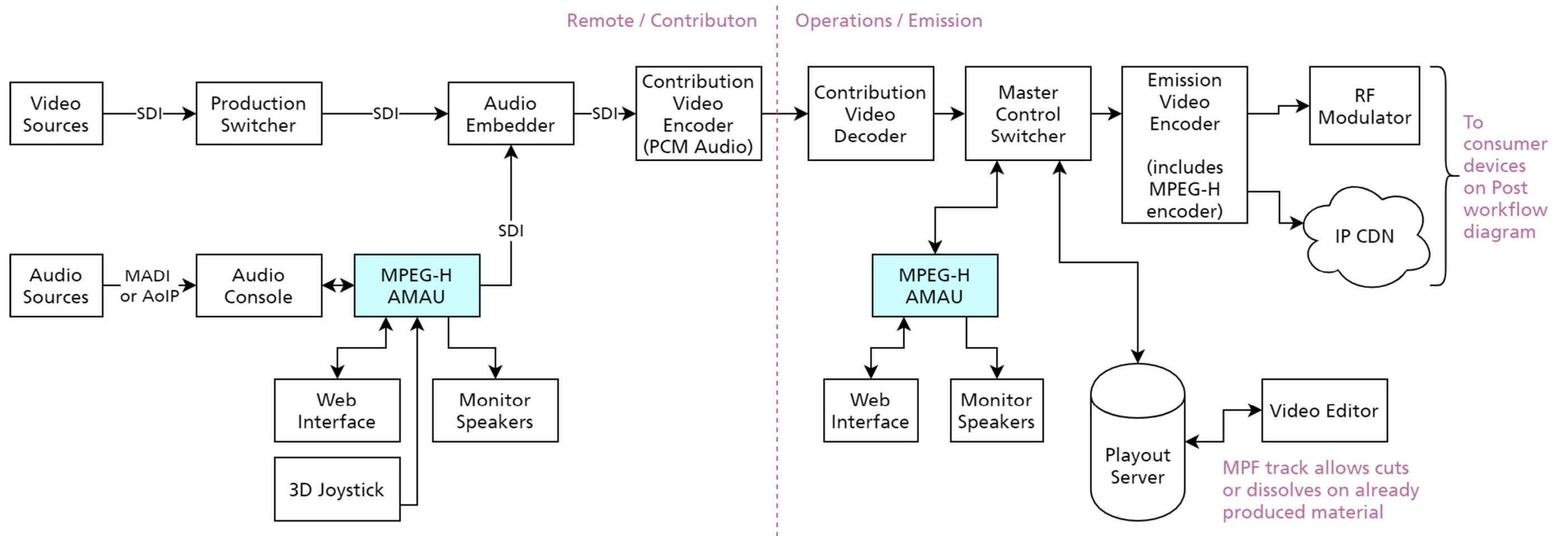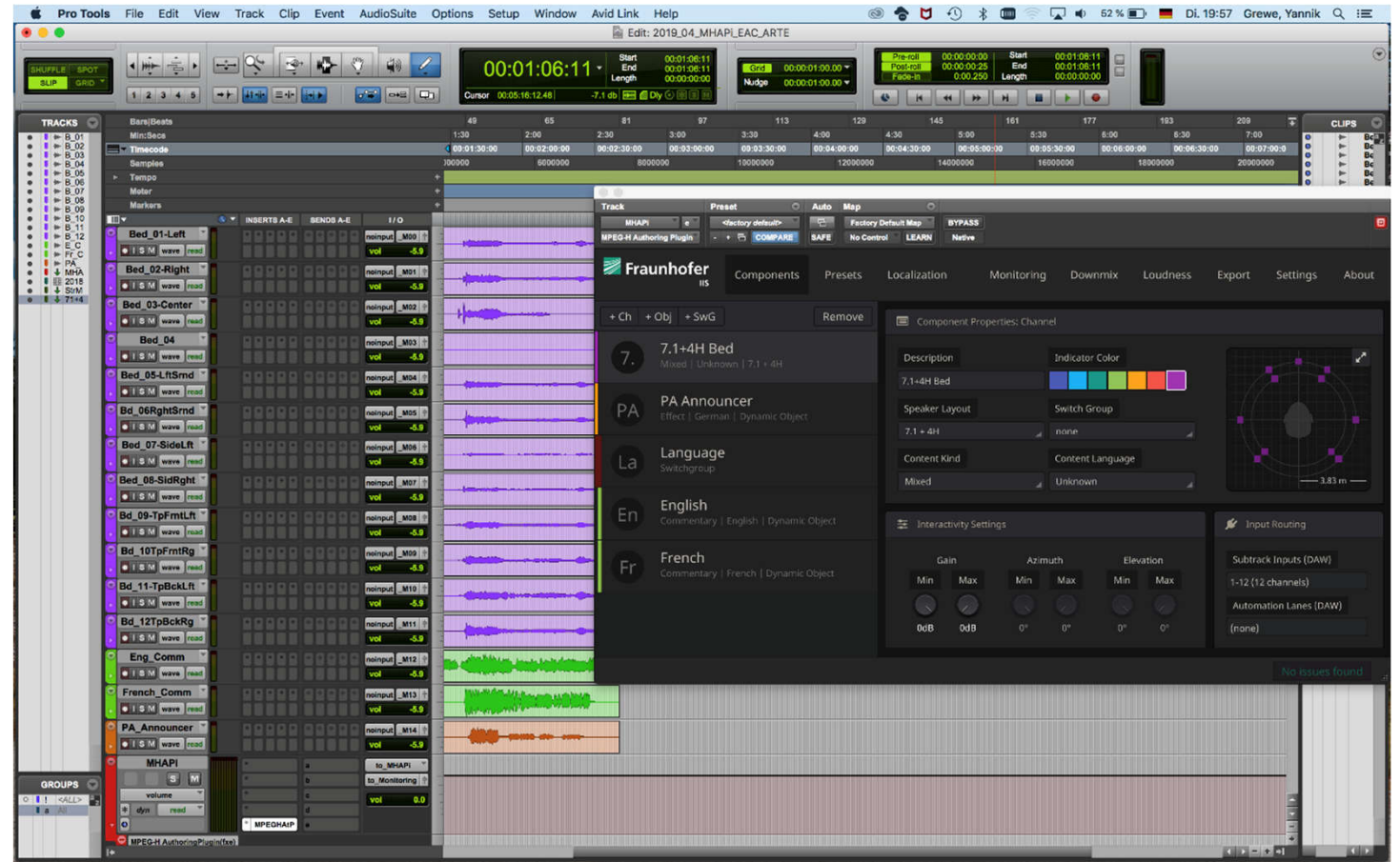
MPEG-H Production offers two workflows: #2 - Post

- Post-Production using VST or AAX plug-ins for DAWs:
    - Fraunhofer MPEG-H Authoring Plug-In
    - Fraunhofer 3D Reverb
- Blackmagic Design DaVinci Resolve Fairlight
    - integrated MPEG-H authoring and panning
- Fraunhofer EncMux tool
    - Encodes MPEG-H audio and combines with video into a mp4 file
- Fraunhofer Atmos ADM Converter
    - Converts Atmos BWF-ADM to MPEG-H Production Format or MPEG-H BWF-ADM

© 2020 Fraunhofer IIS   www.mpeg-h.com   www.iis.fraunhofer.de/audio

Post-Production Workflow

© 2020 Fraunhofer IIS   www.mpeg-h.com   www.iis.fraunhofer.de/audio

MPEG-H AUDIO
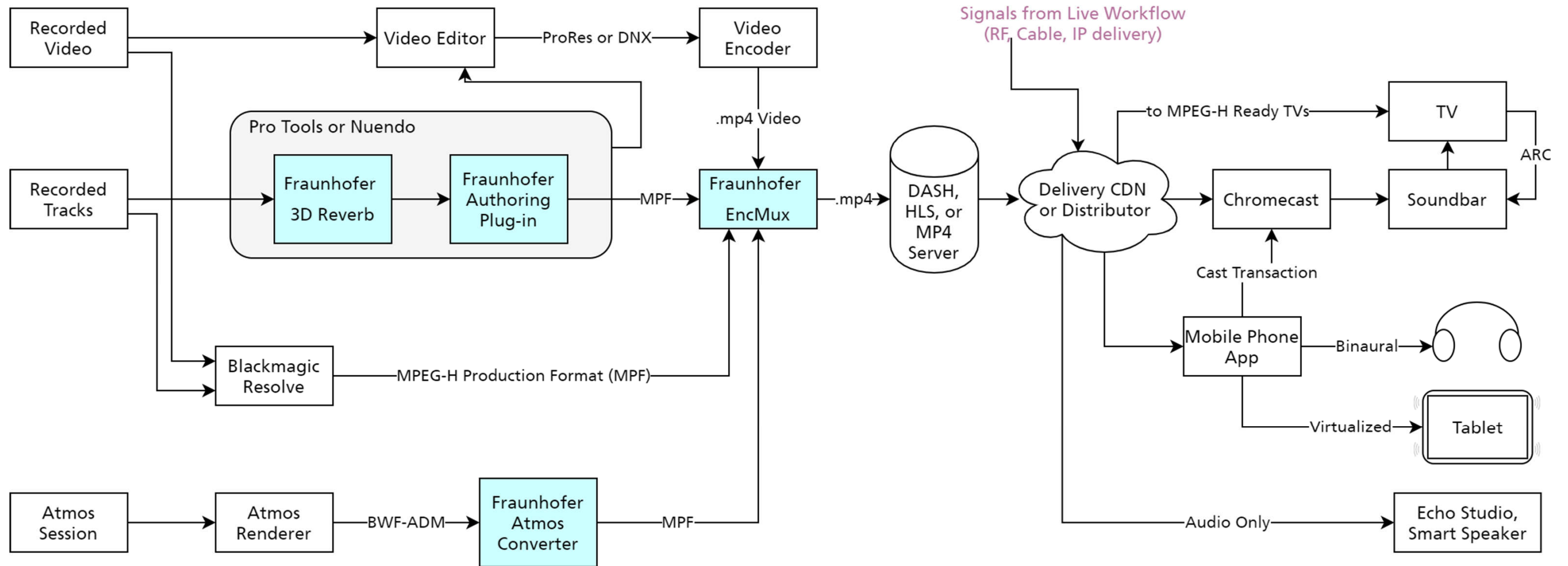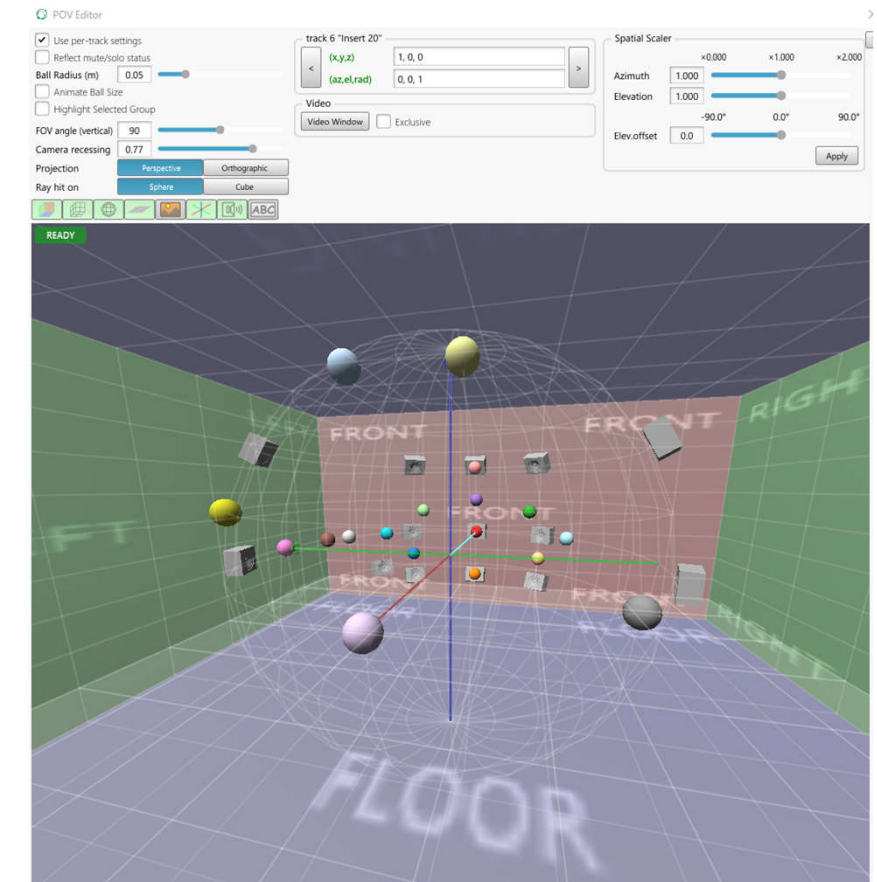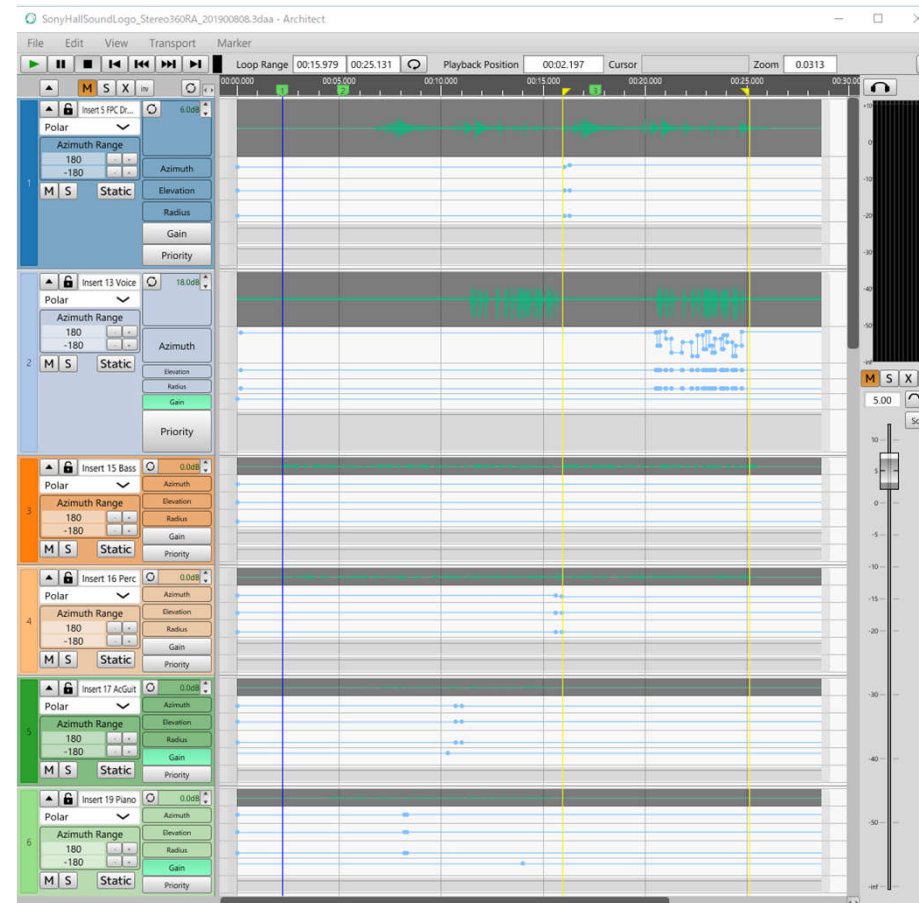
Sony 360 Reality Audio Production

- Mixing of tracks or stems in DAW such as Pro Tools or Neundo
    - Output as wav files to:
- Sony Architect Mixing Tool
    - Panning and automation of objects with loudspeaker or headphone monitoring
- Sony Encoder

© 2020 Fraunhofer IIS   www.mpeg-h.com   www.iis.fraunhofer.de/audio

MPEG-H AUDIO

Fraunhofer

IIS

Content creation – from recoding to delivery

Music Studio / Music Label	Music Service	Listener

Recording	Editing	Encoding	Distribution	Playback

New Recording
- Studio
- Live

From Archive
- Stem files
- Multi tracks

Mixing by 360RA editing tool

Encoding

Music Search & Delivery

Note:
Utilizes the same stem files used with stereo. With no special recording requirements.

Editing Tool

Encoding Tool
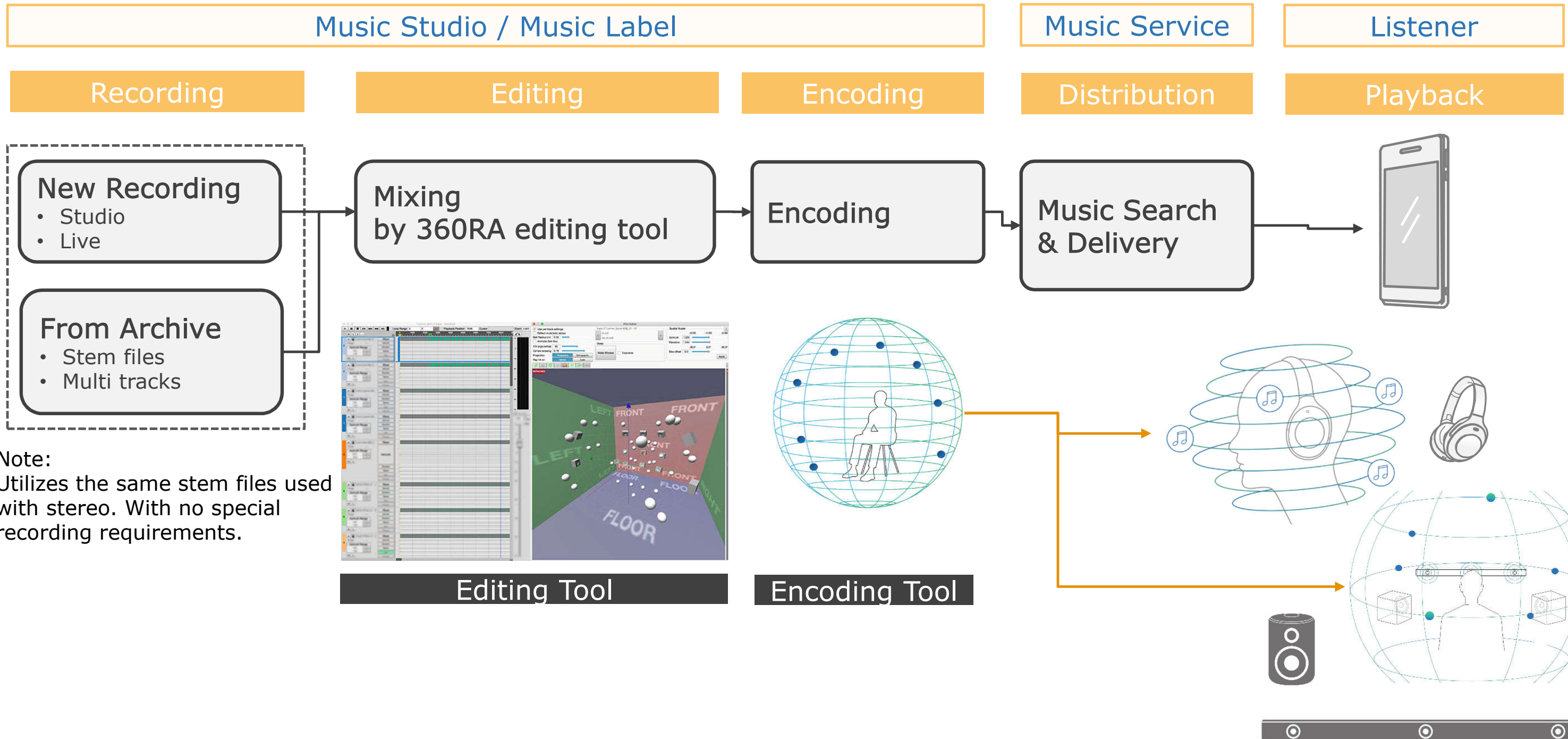
Bus limitations in most DAW software require additional plug-ins

Popular DAW Software	Avid Pro Tools	Steinberg Nuendo	BMD DaVinci Resolve (Fairlight)
Maximum Bus Width	16 signals	22.2	26
3-D panning to channels	7.1.2 only, use Fraunhofer 3D Reverb	22.2, Atmos, or use Fraunhofer 3D Reverb	Yes
3-D panning to objects	Yes, through a renderer box or plug-in	Yes, through a renderer box or plug-in	Yes
Room reverb	Fraunhofer 3D Reverb	Fraunhofer 3D Reverb	Fraunhofer 3D Reverb
Authoring for interactivity	Fraunhofer Authoring Plugin	Fraunhofer Authoring Plugin	Native

www.mpeg-h.com   www.iis.fraunhofer.de/audio

MPEG-H AUDIO

Fraunhofer
IIS

Upgrading a control room to immersive

- Instead of two or five speakers, you need 10 or 12
    - With this number of speakers, self-powered speakers are very convenient
    - Space and cost concerns may indicate use of smaller speakers with 80 Hz crossover to a bass-managed subwoofer

- The control or listening room would ideally have a 10 or 12 foot ceiling to allow sufficient room for upper speakers. In a remote truck, in-wall consumer speakers may need to be considered for upper speakers.

- Bass management should include the height channels as well, if you plan on mixing a helicopter flyover.

- Control room design for immersive is still evolving. The older design styles that were expressly optimized for stereo, such as LEDE, DELE, soffit-mounted monitors, etc. probably won't work well for immersive. A good existing surround room is usually well suited for immersive.

- The basics – good control of bass modes and first order reflections, appropriate reverberation time, low ambient noise, appropriately distributed absorbtion and diffusion, etc. work for immersive just as they have for older formats.

     www.mpeg-h.com   www.iis.fraunhofer.de/audio

MPEG-H AUDIO

Fraunhofer
IIS

Replacing the Auratones or NS-10s

- Ordinary consumers unlikely to be listening on a 7.1+4H AVR system at home
- How do we evaluate their listening experience in the mastering room or listening room?
  - Sennheiser Soundbar ($2500)
    - Highest quality soundbar available today, supports all popular immersive formats
    - Drive with Android TV and Fraunhofer app or Chromecast
  - Echo Studio ($199)
    - Great sound for the price, very accessible to consumers due to distribution
    - Amazon offers professional playback options – see me for details
- Both products partially rely on room reflections, need normal consumer walls and ceilings for good playback
  - Office with "acoustic" grid drop ceiling or dead control room with foot-thick fiberglass on walls won't work well.

© 2020 Fraunhofer IIS   www.mpeg-h.com   www.iis.fraunhofer.de/audio

ARCHIVING IMMERSIVE AUDIO PRODUCTIONS

How do you safely store this stuff?

Current EBay Listing , buy it now for $1,199.00:
"WORKED PERFECTLY UNTIL IT WAS TAKEN OUT
OF ROTATION 3-4 YEARS AGO.
- NOT FULLY TESTED. **SOLD AS IS.**"

Archiving Immersive Audio Productions

- What is a suitable format to store a production for future use?

- For playout, distribution, or short term storage:

  - MPEG-H Production Format (Control Track)

    - Editable, storable, transmittable using all legacy software and hardware. Inherently works with AES 67, SMPTE 2110, other audio over IP standards since it is an audio signal.

    - Unique feature of MPEG-H System

- For archives or content interchange:

  - ITU BWF/ADM

  - SMPTE IAB and IMF

  - Import to MPEG-H using Fraunhofer tools

© 2020 Fraunhofer IIS   www.mpeg-h.com   www.iis.fraunhofer.de/audio

MPEG-H AUDIO

Fraunhofer

IIS

ITU BWF/ADM

- Broadcast Wave File with additional chunks to represent program and object metadata

- Standardized in ITU with participation from Fraunhofer, Dolby, Xperi/DTS, BBC, IRT, NHK, EBU, …

- ADM Profiles define interoperability points:
    - Dolby Atmos ADM Profile – 128 channels/objects, no interactivity, similar to cinema master
    - MPEG-H ADM Profile – 16 simultaneous channels/objects with interactivity/personalization for broadcast and streaming use cases

- Conversion of **MPEG-H ADM** to **MPF** and vice versa

- Specification available on Fraunhofer Website

- Conformance check with Fraunhofer ADM Info Tool

- **Most likely the archive format of choice for sports, news, reality, and other broadcast content.**

ADM ⟷ MPF

S-ADM ⟷ MPF

MPEG-H AUDIO

Fraunhofer

IIS
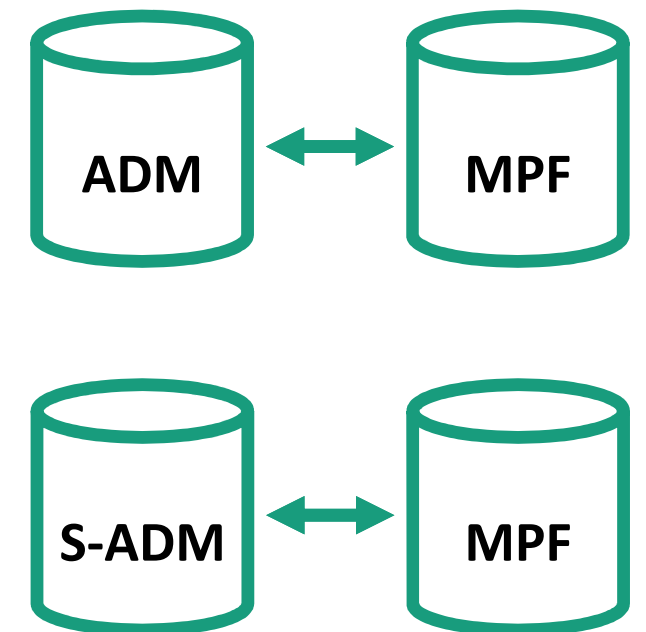
SMPTE IAB (Immersive Audio Bitstream)

- Superset of Atmos theatrical bitstream format, with industry improvements

- 128 channels/objects, no interactivity, similar to cinema master

- Standardized in SMPTE with participation from Fraunhofer, Dolby, Xperi/DTS, Deluxe, Technicolor, Fox, Netflix, …

- Netflix is a major supporter of the standard

- Transport in MXF inside DCP or Interoperable Master Format (IMF)

- **Most likely will be the archive format of choice for film and episodic content**

IAB → MPF

   www.mpeg-h.com   www.iis.fraunhofer.de/audio

MPEG-H AUDIO

Fraunhofer
IIS
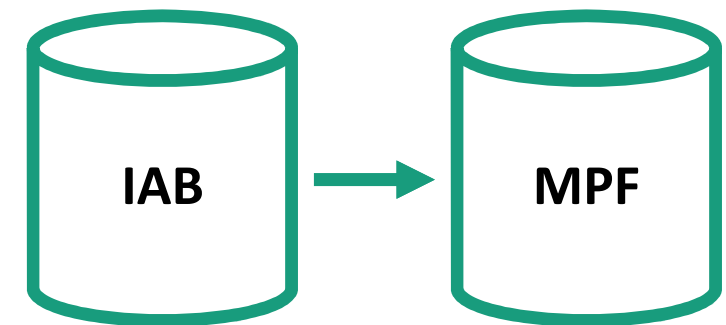
Why consider immersive sound and MPEG-H?

- MPEG-H is one of the primary immersive sound systems implemented today:
    - In all Korean TV sets and on the air with all Korean commercial networks since 2017
    - Used for the new 360 Reality Audio format from Sony
        - Music from Universal, Warner, Sony, and Amazon Music
    - In Amazon Echo Studio, Sennheiser Soundbar, Google Chromecast, and other consumer devices
    - Adopted in Brazil for ISDB-Tb
    - Considered for next generation of TV broadcasting in Japan
    - MPEG-H offers field-proven technical excellence, with no legacy baggage, and uniquely supports true interactivity today
- Consumer-friendly smart speakers, soundbars, and binaural playback make for an easier and less expensive consumer entry point than with DVD-audio, SACD, and surround sound in general.
    - Hopefully this leads to a market of sufficient size for immersive content to grow
- Immersive production is routine in film industry, but not in music or TV – production capabilities will need to be improved

© 2020 Fraunhofer IIS   www.mpeg-h.com   www.iis.fraunhofer.de/audio

MPEG-H AUDIO

Fraunhofer

IIS

PRODUCTION SOFTWARE DEMONSTRATION

Fraunhofer 3D Reverb Plugin, Fraunhofer MPEG-H Authoring Plug-in, Blackmagic Resolve

MPEG-H AUDIO

Fraunhofer

IIS

© 2020 Fraunhofer IIS   www.mpeg-h.com   www.iis.fraunhofer.de/audio