Fraunhofer’s Interactive 3D Audio System for Television Broadcasting – What’s New

Based on the MPEG-H Audio standard

Robert Bleidt – April 7, 2014 – NAB Convention

www.iis.fraunhofer.de/tvaudio
Our Interactive 3D TV Audio System – A Quick Review

- A new audio system for TV broadcasting – being considered for ATSC 3.0 and future DVB work
- Up to 128 channels or audio objects
- Includes a new, high efficiency audio codec based on AAC evolution + new technologies
- Flexible rendering to play any content on speakers from full 22.2 3D audio to earbuds
- Enhanced loudness and dynamic range control for multiple device types and environments
- Delivery over broadcast, internet, and mobile networks to first and second screen devices
- Audio from the new-media experts powering Android and iOS today, with 15 years of experience providing half the world’s surround sound for TV
- Enables: Interactive Audio and Immersive 3D Sound
- (See Fraunhofer presentation from yesterday for details or visit www.iis.fraunhofer.de/tvaudio)
Starting with the basics – what happens if I move to Fraunhofer’s system?
Today’s surround sound at half the bitrate

- Drop-In replacement for AC-3
- Carry 5.1 Surround Sound at 160 kb/s
  - Equivalent quality to AC-3 at 384/448 kb/s
- All AC-3 features and profiles supported
- No changes to operational practices
- New Fraunhofer audio encoder built into either:
  - HEVC video encoder – from usual suppliers such as Harmonic, Ericsson, Thomson, Harris, Cisco, etc
  - Or audio processor – from Linear Acoustic, etc.
- Replace your old MPEG-2 encoder with HEVC encoder, and upgrade your audio codec at the same time
- Proven performance of AAC codec family, powering half the world’s TV surround sound
What about Internet delivery?

- Since new media devices are mostly software-based, upgrading to new audio systems is easier.

- Does the AAC family offer surround sound on:
  - Android – Yes, Fraunhofer FDK library is built into Android OS.
  - iOS – Yes, Apple HE-AAC decoder built into iOS.
  - Windows – Yes, Microsoft HE-AAC MC decoder built into Windows.
  - Mac – Yes, Apple HE-AAC MC decoder built into MacOS Core Audio.
  - Linux – Yes, Fraunhofer HE-AAC MC FDK library available in ffmpeg.
  - Browsers – Yes, IE, Chrome, Safari, Android Browser, Firefox (on many OS).
  - Devices – Yes, Chromecast, Playstation, other game consoles.
  - Flash – Yes, stereo downmix output.
  - DSPs – Yes, TI, Analog Devices, Tensilica, ARM, MIPS, Tilera, Freescale, Broadcom, Qualcomm, MediaTek, etc.
Suppose we get on the air with the Fraunhofer system
What do we gain?

- Support for audio objects
  - Ability to offer personalized sound
  - Ability to efficiently carry other languages and VI

- Support for 3D sound
  - Offer immersive sound on par with cinema sound systems
  - (with objects) Offer personalized 3D sound

- Best listening experience on the wide range of viewers’ devices and environments:
  - Full home theater to phone earbuds
Audio Objects: What are they and what will they enable?
Objects allow a different approach to transmission

- An audio object is basically a channel that can be positioned or panned during playback
- Think of transmitting mixer channels or stems instead of final outputs
  - And sending pan-pot settings as well
- Objects can be moved during playback to track action in the video,
- Rendering under the viewer’s control:
  - Listener can adjust volume of objects, or turn on/off, or perhaps adjust position
- Rendering to match the listening environment:
  - Objects can be rendered to best match the speaker configuration available
- Objects can be combined with traditional channels
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
Early Concept of Object-based Interactive Audio
Field Test with objects – Winter Sports Event
Try it at the Fraunhofer booth SU 6117

- Uses audio and video from network live winter sports broadcast in January
  - Pick your announcer and language
  - Want more sound effects? More audience?
- Also 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
- See and hear it at our booth SU 6117
How we used objects in our winter sports field test mix

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 kbps.
Summary: Interactive Audio and Audio Objects

- New interactive features:
  - Viewers can adjust the dialogue level to their liking
  - Hard-of-hearing or non-native viewers can boost dialogue
  - Second or third language or VI support just requires additional objects – 20-40 Kb/s each
  - Viewers can personalize the audio mix to “hear what they want to hear”
  - Existing technical infrastructure can be used for static (non-moving) objects
Improving The Listening Experience With 3D Sound And Flexible Rendering
Surround sound is great, and 3D audio is even better

- 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

From Silzle et al, Investigation on the Quality of 3D Sound Reproduction, ICSA 2011
Playback rendering decouples channels and objects from speakers

- Rendering adapts the transmitted channels or audio objects to the speakers in the listener's environment.
- A simple example today is a downmix from 5.1 surround to stereo speakers.
- 3D audio systems typically have more flexible rendering, to allow for speakers in arbitrary positions, and to render objects.
- Rendering may include, with varying results, psychoacoustic processing to render the impression of height without height speakers, for legacy 5.1 or 2.0 consumers.
- HDMI 2.0 will allow communication of consumer's installed speaker locations (though there will likely be opportunities for CE manufacturers to add value here).
“Broadcasters will never produce live sports in 3D”

- “...Just mixing surround in a live broadcast is too complex”
- Well, we did.
- (sort of)
- We took signals from the host broadcaster at a sports event in January, and mixed to 3D in post-production
- It was not live since:
  - No time for pre-production or rehearsal
  - We covered three overlapping events from two production teams
  - Console tools for panning still need some work
- As far as we know, the first U.S. sports broadcast mixed in 3D audio
- Hear it at the Fraunhofer IIS booth SU 6117
“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?
Fraunhofer demonstrates future 3D listening for mainstream consumers with prototype sound frame

- 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
- Hear it at the Fraunhofer IIS booth SU 6117
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
- Flexible rendering will play the received audio in best quality on the speakers the viewer has
- “Sound Frame” technology could be built into TVs or soundbars to enable a dramatic improvement in sound quality and immersive sound for the mainstream consumer.
Implementing Our System in TV Plants and Networks
Inputs and Outputs of Fraunhofer Audio Codec

**Emission / Consumer Delivery**

- **Video Encoder**
  - Channels (PCM)
  - Objects (PCM)
  - Object Positioning
  - Loudness Settings

- **Fraunhofer Audio Encoder**
  - MPEG-H Audio Bitstream

- **Video Decoder**
  - Speakers (PCM)

**Contribution / Distribution**

- **Video Encoder**
  - Channels (PCM)
  - Objects (PCM)
  - Dynamic Object Metadata

- **Fraunhofer Audio Encoder**
  - MPEG Systems Bitstream
  - MPEG-H Audio Bitstream (Contribution or Distribution Quality)

- **Video Decoder**
  - Channels (PCM)
  - Dynamic Object Metadata
Proposal for Implementation in Stages

- Stage 1 – Drop-In replacement for AC-3
  - Benefit: Lower audio bitrate, better loudness control

- Stage 2 – Add some static objects to the 5.1 or 2.0 channel bed:
  - Second or third language dialogue
  - Audio Description / Visually Impaired
  - Primary language dialogue (for Dialogue Enhancement)
  - Sound Effects
  - Benefit: Multiple Languages, VI description at 20-40 kb/s each, User-adjustable dialog level, Personalized audio mix

- Stage 3: move to practical 3D audio – perhaps 5.1 or 7.1 + 4 height channels
  - Benefit: Immersive, 3D audio – parity with cinema, Blu-ray

- Stage 4: Add dynamic (moving) objects:
  - Benefit: ability to track screen action with mono track, carry Atmos or Auro 3D theatrical mix
What else could we talk about?

- Improved loudness control across devices and in different listening environments
- Adaptive Internet delivery and second-screen applications
- Efficient delivery to mobile networks, including object priority and bifurcated streaming
- 3D sound on mobile devices with Fraunhofer Cingo
- Based on an open international standard from ISO/MPEG
- Source code from Fraunhofer for equipment manufacturers
- Application support from Fraunhofer for broadcasters
AAC’s proven track record in multichannel TV sound

- AAC provides one-half to two-thirds of world’s surround TV
  - Part of DVB toolbox
  - Audio codec of ISDB (Japan, South America)
- Fifteen years of surround audio for Japan
  - AAC encoding hardware since 1998
  - Multichannel decoding supported in all TVs, AVRs, soundbars
  - AAC is codec of new 22.2 audio in NHK SHV broadcasting
- On the air for BBC’s HD service since 2009
  - Sole audio codec used for all UK Freeview TV
- Used in television broadcasting around the world:
  - United Kingdom, Norway, Denmark, Ireland, Israel, Slovenia, Australia, New Zealand, Hungary, Vietnam, South Africa, Tunisia, Turkey, Ghana, Kenya, Malawi, Nigeria, Uganda, Zambia
  - Japan, Brazil, Peru, Argentina, Chile, Venezuela, Ecuador, Costa Rica, Paraguay, Philippines, Bolivia, Nicaragua and Uruguay
  - Others planned but not on air or in trials yet, including: Russia, Singapore, Malaysia, Thailand, ...
Case Study - Google

- Google wanted multichannel audio and the ability to deliver cinema-quality sound to Android users

- Fraunhofer:
  - Provided open-source FDK software to Android
  - Worked with Android engineers on new metadata and audio processing
    - Additional metadata target levels and presentation modes
    - Movie loudness matched to music loudness
    - Dynamic range optimized for home theater, living room, tablet, earphone playback
    - Foolproof, automatic surround sound with HDMI output to AVR and TV
  - Fraunhofer Cingo incorporated into new Nexus devices for earphone and tablet speaker surround
  - On the air today with Google Play Movies
Hear our results at NAB booth SU 6117

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