S34-2:AHG on Audio for ATSC 3.0

Jim Starzynski,
NBCUniversal
S34-2 Chairman
From Evolution to Revolution

• **ATSC S34-2** is the ad-hoc group currently working on the next generation of DTV audio. The group was formed in August of 2013 and meets roughly 2 times per month in a combination of teleconferences and face to face meetings.

• The group members are representatives of organizations that are involved in the technical aspect of TV Audio.

![Group Members and Logos]...and many more...
The Future of TV Sound
Audio for ATSC 3.0
Overview

• We can consider ATSC’s requirement of “immersive” audio to consist of two components:

  1.) Personalization

  2.) Enhanced Surround Sound
## The Requirements:

<table>
<thead>
<tr>
<th>Req. #</th>
<th>Requirement Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>65</td>
<td>The system shall support Immersive Audio with enhanced performance when compared with existing 5.1 or 7.1 channel-based systems.</td>
</tr>
<tr>
<td>66</td>
<td>The system shall support delivery of Content from mono, stereo, 5.1 channel and 7.1 channel audio sources, as well as sources supporting Immersive Audio.</td>
</tr>
<tr>
<td>70</td>
<td>The system shall enable Immersive Audio on a wide range of loudspeaker configurations.</td>
</tr>
<tr>
<td>71</td>
<td>The system shall enable Immersive Audio on loudspeaker configurations with suboptimum loudspeaker locations.</td>
</tr>
<tr>
<td>72</td>
<td>The system shall enable Immersive Audio for headphones</td>
</tr>
<tr>
<td>73</td>
<td>The system shall enable audio reproduction on loudspeaker configurations not designed for Immersive Audio such as 7.1 channel, 5.1 channel, two channel and single channel loudspeaker configurations.</td>
</tr>
<tr>
<td>74</td>
<td>The system should enable user control of certain aspects of the sound scene that is rendered from the encoded representation, (e.g., relative level of dialog, music, effects, or other elements important to the user).</td>
</tr>
<tr>
<td>75</td>
<td>The system shall enable user-selectable alternative audio tracks to be delivered via Terrestrial Broadcast or via Broadband and in Real Time or Non-real Time. Such audio tracks could be intended to replace the primary audio track or be mixed with the primary audio track and shall be delivered for synchronous presentation with the corresponding video content.</td>
</tr>
<tr>
<td>76</td>
<td>The system shall enable mixing in the receiver of alternative audio tracks (e.g., assistive audio services, other language dialog, special commentary, music and effects) with the main audio track or with each other, with relative levels and position in the sound field and with receiver adjustments suitable to the user.</td>
</tr>
<tr>
<td>77</td>
<td>The system shall enable broadcasters to provide users with the option of varying the loudness of a TV Program’s dialog relative to other elements of the audio mix, for purposes of increased intelligibility.</td>
</tr>
<tr>
<td>78</td>
<td>The system shall enable adapting the loudness and dynamic range of audio Content as appropriate for the receiving device and environment of the Content presentation.</td>
</tr>
<tr>
<td>79</td>
<td>The system shall enable audio reproduction for mono, 2, 5.1, and 7.1 channel and Immersive audio sources, of a quality equal to or better than that specified in Recommendation ITU-R BS.1548-4, 2013 (User requirements for audio coding systems for digital broadcasting), Appendix2, Section 2.1.1.1, as determined via the methodology described in Recommendation ITU-R BS.1116-1 extended to cover the 7.1 channel and Immersive Audio cases.</td>
</tr>
<tr>
<td>80</td>
<td>The system shall support information and functionality to normalize and control the loudness of reproduced audio Content.</td>
</tr>
<tr>
<td>81</td>
<td>The system shall support alternative audio tracks to be delivered via Terrestrial Broadcast or via Broadband and in Real Time or Non-real Time. Such audio tracks could be intended to replace the primary audio track or be mixed with the primary audio track and shall be delivered for synchronous presentation with the corresponding video content.</td>
</tr>
<tr>
<td>82</td>
<td>The system shall enable mixing in the receiver of alternative audio tracks (e.g., assistive audio services, other language dialog, special commentary, music and effects) with the main audio track or with each other, with relative levels and position in the sound field and with receiver adjustments suitable to the user.</td>
</tr>
<tr>
<td>83</td>
<td>The system shall enable broadcasters to provide users with the option of varying the loudness of a TV Program’s dialog relative to other elements of the audio mix, for purposes of increased intelligibility.</td>
</tr>
<tr>
<td>84</td>
<td>The system shall enable audio reproduction for mono, 2, 5.1, and 7.1 channel and Immersive audio sources, of a quality equal to or better than that specified in Recommendation ITU-R BS.1548-4, 2013 (User requirements for audio coding systems for digital broadcasting), Appendix2, Section 2.1.1.1, as determined via the methodology described in Recommendation ITU-R BS.1116-1 extended to cover the 7.1 channel and Immersive Audio cases.</td>
</tr>
</tbody>
</table>
Audio: Personalization

- Ability to *efficiently* customize an audio experience with channels/objects:
  - Choose language
  - Choose commentary
  - Address impairments with description and improved intelligibility
  - Normalize loudness of all content
  - Contour dynamic range to the unique user, device and environment
The Future of TV Sound: Personalized Features
Personalization: Dialog Enhancement
Personalization: Language
Personalization: Description
Personalization: MY WAY!

Streaming

MLB NETWORK

ON FOX

ATSC

NORMAL
Loudness Normalization

Commercial arrives as videotape or digital file

-18 LKFS (LOUD)

**Commercial Aggregator**

-6 LKFS SHIFT

ALL -24 LKFS LOUDNESS

**SCALER**

C.A.L.M. IDENTIFIER INSERTED

ON AIR VIDEO SERVER

ALL AUDIO
-24 LKFS AVERAGE

ATSC 3.0 Encoder
Passes Audio Unaltered

COMM. ARRIVES AS VIDETAPRE OR DIGITAL FILE

-18 (VARIOUS) LKFS LOUDNESS

**Commercial Aggregator**

ON AIR VIDEO SERVER

NO C.A.L.M. IDENTIFIER

ALL AUDIO
-24 LKFS AVERAGE

ATSC 3.0 Encoder
Normalizes Loudness in Real Time

COMM. ARRIVES AS VIDETAPRE OR DIGITAL FILE
Dynamic Range Control  Soft To LOUD
Creative Intent: Full Range

Consumer’s Capabilities

- MOBILE STANDARD WIDE NORMAL
- NARROW STANDARD WIDE NORMAL
- NARROW STANDARD WIDE NORMAL
The Future of TV Sound: Efficiency
Transmitting Audio: ATSC-1

- 1 Complete Main (Stereo or 5.1) **384 kb/s**
- 1 Ancillary Soundtrack (Alternate Language or VDS) ~**192 kb/s** (usually mono)

**576 kb/s**
Transmitting Audio: ATSC-3

192 kbps
Object-based Personalized
  Music & Effects

48 kbps
English Dialogue
Spanish Dialogue
English DVS
Spanish DVS
Metadata

English Complete Main
English Main + DVS
Spanish Complete Main
Spanish Main + DVS

4 Complete Presentations at ~384kbps

384 kb/s!
The Future of TV Sound: Enhanced Surround Sound
Immersive, Enhanced Surround Sound

• Improved spatial resolution in sound source localization
  – Sound with improved azimuth, elevation and distance perspective
  – Use of channels and objects or “elements” and metadata (similar to fader automation)
  – Metadata allows rendering at the decoder, customized to the user’s sound system
  – The decoder places the sound in the most accurate position the user’s sound system supports

ATSC 3.0 Performance Supports Multiple Platforms!
Objects:

“An audio object is basically a channel that can be positioned or panned during playback”

- **ATSC 1.0**
  - Up To 16 channels of routed audio
  - Up to 6 rigid (bound) channels
  - To 1 speaker
  - Up to 5.1 channels

- **ATSC 3.0**
  - Up To 40+ elements/ch routed audio & metadata
  - Up to 40+ unbound Elements and Channels
  - Up to ~11.1 speakers
Rendering:
Creative Intent: Multiple Elements

Decoder’s Renderer

Consumer’s Capabilities

Creative Intent: Multiple Elements

Decoder’s Renderer

Consumer’s Capabilities
Rendering for Headphones

Headphone Reproduction Will Simulate Height & Depth—Sounding more like speakers and the feeling of listening in a real room
Rendering Home Theater

AVR Learns Home Sound Environment:
• User Enters Specifics During Set-up
• AVR Performs Auto Analysis w/mic.
Speakers for Immersive

7.1+4

5.1+4

Immersive Soundbar

Conventional or Reflective Approach
S34-2: Process
Call for Proposal Roadmaps the Project

• CFP main sections:
  • Audio System Requirements
  • Proponent Registration
  • Precertification Listening Tests
  • Phase 1: Listening Tests
    • Formal Feature Evaluation
    • Proponent Demonstrations
  • Phase 2: Decision Process
    • Test Coordinator Review
    • Information Analysis
    • Member/Group Decision on Technology

• Candidate Standard
CFP Specifies Test Methods
esp. small impairments & sound quality...


- **ITU-R BS.1534-2**, “Method for the subjective assessment of intermediate quality level of audio systems”.

- And **30-channel** loudspeaker position table (incl. azimuth, elevation) to be used by all testing labs.
Audio Test Content:
Acquisition and Licensing

- >15 content owners provided total of 115 audio test clips in 7 genres for 6 audio formats. (each clip 10 ~ 30 sec)

<table>
<thead>
<tr>
<th>CATEGORY</th>
<th>Total Clips</th>
<th>Stereo</th>
<th>5.1</th>
<th>7.1+4</th>
<th>22.2</th>
<th>HOA</th>
<th>Objects</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>by Format →</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>by Category ↓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DRAMA</td>
<td>11</td>
<td>0</td>
<td>1</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>MDA</td>
</tr>
<tr>
<td>EFFECTS</td>
<td>29</td>
<td>1</td>
<td>2</td>
<td>7</td>
<td>3</td>
<td>14</td>
<td>2</td>
<td>(38, /31)</td>
</tr>
<tr>
<td>F/M VOICE</td>
<td>7</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MUSIC-Classical</td>
<td>39</td>
<td>4</td>
<td>10</td>
<td>4</td>
<td>2</td>
<td>19</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>MUSIC-Jazz</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>MUSIC-Pop/Rock</td>
<td>16</td>
<td>1</td>
<td>6</td>
<td>1</td>
<td>3</td>
<td>5</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>SPORTS</td>
<td>8</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>No Category</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>- hdpdBRIR</td>
</tr>
<tr>
<td>Total</td>
<td>115</td>
<td>12</td>
<td>23</td>
<td>17</td>
<td>12</td>
<td>46</td>
<td>4</td>
<td>1</td>
</tr>
</tbody>
</table>

- And every clip must be licensed by its owner to each Proponent; and to ATSC including Test Coordinators.
CFP Speaker Configurations

30-Channel Speaker Map

Example Immersive Layout (7.1 + 4)
Audio Test Content:
Content Selection Meeting

- December: Qualcomm’s BS.1116 compliant listening room in San Diego
- The sweet-spot is right about here
- 60 Clips were identified
Test Method - Precertification

- BS.1116 recommends:
  - Listener selection/validation
  - Selection of test content
  - Double-blind test using 5-grade scoring system
  - Detailed listening room specifications
  - Reveals Bit-Rate Required to Score 4 or better
Precertification Results:
2 Proponents are certified in the 3.0 Audio Process
Test Method – Phase-1

- Content Preparation
  - Performed in April at Studio 22, McGill University in Montreal
  - Test Coordinators Validate Usability of Coded Content Streams Prior to Release
  - Additional Expert Confirms Content and the Process
• BS.1534 recommends:
  – Listener selection/validation
  – Selection of test content incl. low and mid-anchors (3.5, 7kHz LPFs)
  – Double-blind test using a 0 ~ 100 scale and anchors*
  – Listening room same requirements as BS.1116.
  – More rigorous statistical analysis of listener scores compared to BS.1116.

* This methodology is referred to as MUSHRA (MUlti Stimulus test with Hidden Reference and Anchor)
Phase 1: Feature Evaluation and Test Site Visit:

- Weeks of June 1 and June 8 - Test Coordinators (and other interested ATSC members) visit Proponent facilities during Phase 1 testing

- Allows Test Coordinators to:
  - Review any Phase 1 testing in-process
  - Judge features on Pass/Fail basis if proponent has met required elements of CFP (Annex E)

- Permits proponent to demonstrate features which may be in addition to and exceed those required by the CFP
Proponent Demonstrations
Conducted during July Face to Face Meetings

- July 13-16 Hilton Marietta—Atlanta, Georgia
- Separate from the CFP Test and Evaluation process
- Open to all ATSC members
- Provides an opportunity for any member to acquire the most current information, and for proponents to demo their system:
  - Members may schedule private sessions addressing their concerns and needs affording them an opportunity to ask any and all questions
  - May use “more interesting, specific material” (e.g. movie trailers, etc.)
  - Show and discuss features that exceed or are above and beyond those required by the CFP
- Gather valuable feedback from interested ATSC members
# Audio CFP Schedule

<table>
<thead>
<tr>
<th>Test Element</th>
<th>Result/Deliverable</th>
<th>CFP</th>
<th>Deadlines</th>
</tr>
</thead>
<tbody>
<tr>
<td>Registration</td>
<td>CFP response from prospective proponent</td>
<td>6.1</td>
<td>12-Jan-2015</td>
</tr>
<tr>
<td>Acceptance</td>
<td>Notice to proponent from ATSC</td>
<td>6.1</td>
<td>16-Jan-2015</td>
</tr>
<tr>
<td></td>
<td>Precertification content made available</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Precertification</td>
<td>Proponent self-test report due</td>
<td>6.2</td>
<td>09-Mar-2015</td>
</tr>
<tr>
<td></td>
<td>Detailed system descriptions due</td>
<td></td>
<td></td>
</tr>
<tr>
<td>System Presentations</td>
<td>Proponent F2F presentations to ATSC</td>
<td>6.2</td>
<td>11-Mar-2015</td>
</tr>
<tr>
<td>Certification</td>
<td>Notice to proponent of ATSC acceptance</td>
<td>6.2</td>
<td>20-Mar-2015</td>
</tr>
<tr>
<td>Phase 1 Original Content Available</td>
<td>Original test content to proponenets</td>
<td>6.3</td>
<td>23-Mar-2015</td>
</tr>
<tr>
<td>Demo Date/Location Set</td>
<td>Proponents advised of Demo Event details</td>
<td>6.4</td>
<td>23-Mar-2015</td>
</tr>
<tr>
<td>Coded Audio Uploaded</td>
<td>Proponents deliver coded content to ATSC</td>
<td>6.3</td>
<td>10-Apr-2015</td>
</tr>
<tr>
<td>Phase 1 Test Content Available</td>
<td>All Phase 1 test material to proponenets</td>
<td>6.3</td>
<td>01-May-2015</td>
</tr>
<tr>
<td>Phase 1 Tests</td>
<td>Listening test scores (per BS.1534)</td>
<td>6.3</td>
<td>26-Jun-2015</td>
</tr>
<tr>
<td></td>
<td>Required features (Pass/Fail)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Non-required features (Narrative)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Demo Event</td>
<td>At Proponent’s option</td>
<td>6.4</td>
<td>13/14-Jul-2015</td>
</tr>
<tr>
<td>Phase 2</td>
<td>Delivery of test results to S34-2</td>
<td>6.5</td>
<td>15-Jul-2015</td>
</tr>
<tr>
<td>Recommendation</td>
<td>Audio System decision delivered to S34</td>
<td>6.5</td>
<td>14-Aug-2015</td>
</tr>
</tbody>
</table>

Table 3: Summary of ATSC 3.0 Audio CFP Process Schedule
**Summary:**

- ATSC 1.0 audio “evolution-ized” Television sound by introducing **digital technology**, quality driven **loudness control** and **5.1 discrete surround sound** to the audience
- ATSC 3.0 will **revolutionize** Television sound by expanding digital technology in an **extremely efficient** manner, bringing a host of new features to the user, enabling **personalization** and **enhanced, immersive surround sound** with almost unlimited boundaries

**Important to our Members:**

- A successful Television Audio **user experience** is the result of a well planned marriage of broadcast and consumer technology!
- In ATSC 1, we learned that clear understanding and design of user accessibility of features is key to creating a technology that meets the goals of the system designers

**Moving Forward:**

- Leverage the Test and Feature Evaluation Visits and the July Demos to gather critical information for making the best decision for your company’s unique operation and needs!
Thank You!
ATSC 3.0
UltraHD. Delivered.

Tune In to the Future
Thanks to our Sponsors

LG  zenith  DOLBY

BTS  CEA  COX MEDIA

Fraunhofer  NAB LABS  NERCO DTV

ONEMedia  pearl  SAMSUNG  SONY

triveni DigiITAL.  VERANCE  PBS  bim