Audiovisual Updates

Latest Additions to the DVB Toolbox

Ken McCann, Chair, TM-AVC
Virginie Drugeon, Chair, TM-AVC Video Drafting Sub-Group
David Daniels, Co-Chair, TM-AVC NGA Sub-Group

Since the earliest days of DVB, a single unified specification for video and audio coding has been maintained for all applications based on the MPEG-2 Transport Stream, whether delivered by terrestrial broadcasting, satellite, cable or IPTV. This specification was originally made available in 1995 as DVB BlueBook 001 and published by ETSI as TS 101 154. It has subsequently been updated several times to include further options for video and audio coding and it's being revised again in 2016 to add support for two important new features: Ultra High Definition Video Phase 2 (UHD-1 Phase 2) and Next Generation Audio.

What are the technical challenges for UHD-1 Phase 2?

DVB approved new Commercial Requirements for UHD-1 Phase 2 in November 2015. The new DVB UHD-1 Phase 2 format will be an extension to the DVB UHD-1 Phase 1 format that was finalized in 2014. DVB UHD-1 Phase 2 adds two main technical features to Phase 1: High Dynamic Range (HDR) and High Frame Rates (HFR). Both features aim to increase the sense of reality in pictures by representing a greater range of luminance and color in the case of HDR, and improving motion representation in the case of HFR. Both features are technically independent from each other, so that DVB expects that there will be bitstreams that use HDR only, bitstreams that use HFR only, and bitstreams that use both.

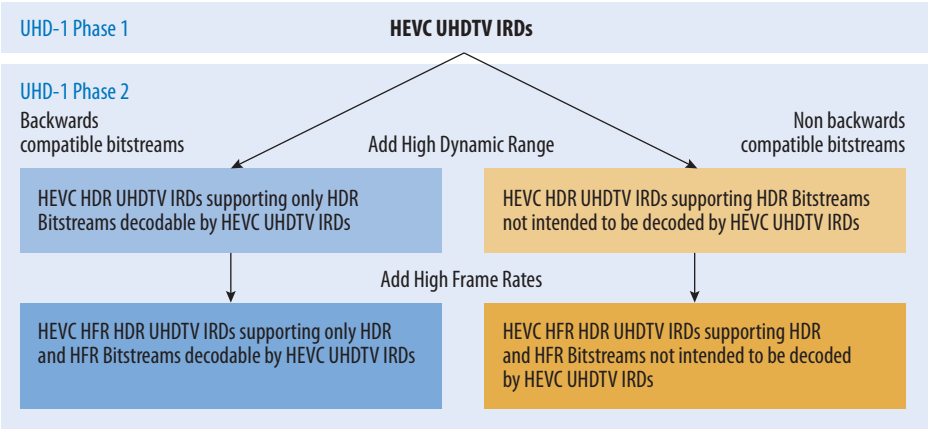
While the technical details of the HDR and HFR formats are still under discussion in the DVB TM-AVC group (the Sub-Group of the Technical Module responsible for audio and video coding), some aspects of the new UHD-1 Phase 2 standard have already been identified as working assumptions for the current discussions:

- HEVC is the codec used for all UHDTV applications;
- Progressive video formats only;
- Square pixel resolutions only;
- 10 bits only;
- BT. 2020 color primaries only;
- 4:2:0 chroma subsampling.

Furthermore, the Commercial Requirements state that at least one backwards compatible profile will need to be specified. Backwards compatibility is understood as the capability for new UHD-1 Phase 2 bitstreams to be decodable by already specified and existing UHD-1 Phase 1 receivers (called HEVC UHDTV IRDs in the currently published version of TS 101 154).

DVB TM-AVC has identified four receiver conformance points to be specified in the next revision of TS 101 154 (see Figure 1) after analysis of the DVB Commercial Requirements for UHD-1 Phase 2. The two receiver conformance points in the upper row of Figure 1 extend the existing UHD-1 Phase 1 receiver conformance point with HDR, while the two in the lower row further add HFR. Therefore, although bitstreams will be able to use HDR or HFR independently, it is expected that all HFR capable receivers will also be HDR capable. Furthermore, the first and second columns of Figure 1 differentiate between receivers that can support only backwards compatible bitstreams and receivers that can support only non-backwards compatible bitstreams. Indeed, the Commercial Requirements identified two distinct market segments for UHD-1 Phase 2 with different requirements. Those markets that do not require backwards compatible bitstreams are expected to benefit, in terms of coding efficiency, from using a solution that is not backwards compatible.

Figure 1. Four receiver conformance points identified for UHD-1 Phase 2 specification



There are currently many potential technical solutions for the definition and transmission of HDR signals under discussion in various standardization bodies (such as ITU-R, SMPTE and MPEG). DVB TM-AVC has therefore started an analysis of potential HDR solutions based on a common framework. Both the framework for analysis and the analysis itself are still work-in-progress, as is the exact number of technical HDR solutions to consider.

Figure 2 attempts to broadly summarize the different possibilities that have been brought to the attention of TM-AVC so far by grouping solutions in three different classes: single layer solutions, dual layer solutions, and solutions currently under discussion in the MPEG HDR fast track effort that may result in an HDR specific extension to the currently published HEVC specification. The criteria listed in the table are relevant for the TM-AVC effort to draft a specification. Identification of dependencies on other standardization bodies is essential to make sure that the DVB solution will be available in the required timeframe. Other relevant aspects include specification of the bitstream (such as HEVC profile and level, single or dual layer, specification of enhancement layer for dual layer solutions) and signaling of the characteristics of the HDR signal at the video elementary stream level (such as signaling of the transfer function or metadata). The choice and signaling of the transfer function used to grade the HDR content and therefore increase the range of available luminance values is one major discussion point. Two possibilities have been identified: the Perceptual Quantizer Electro-Optical Transfer Function standardized by SMPTE, and the Hybrid Log-Gamma Opto-Electrical Transfer Function standardized by ARIB. Both transfer functions have pros and cons in terms of backwards compatibility and representation of artistic intent.

The situation is easier for HFR since the technical challenges for HFR extension had already been considered during the specification of UHD-1 Phase 1, so that the currently published version of TS 101 154 already covers the basic tools for encoding HFR signals. The main challenges are the specification of the backwards compatible profile that uses temporal scalability and the transmission of the low frame rate base layer and the high frame rate enhancement layer in separate PIDs of the MPEG-2 Transport Stream.

What can we expect from Next Generation Audio?

While the video experts are working on

bringing us closer to creating realistic pictures, some might argue that in the audio world we are already there. But there are still some limitations to be overcome, and the work on Next Generation Audio (NGA) aims to address these. So, what is NGA? Well, NGA is a complete rethink of how audio is produced and delivered, increasing the functionality that can be delivered to the listener, while simplifying the process of production. NGA systems deliver audio as Scene or Object based streams, making it possible to deliver audio and a description of how that audio relates to the soundstage.

In the home, both Scene (also known as Higher Order Ambisonics) and Object based audio systems are loudspeaker agnostic, adapting the rendering to the local speaker and the room's geography and ambience. In the case of a Scene based encoder, the soundfield is captured as a representation of the actual sound pressure map. In the case of Object based audio, a series of audio objects are captured. For example, for an orchestra these objects may be the soloist, the strings section, the percussion section, the hall and audience, etc. For each object, its exact placement in the sound field is also captured.

This allows new use cases to be supported, such as providing improved accessibility by increasing the dialogue object above the ambience. It also supports personalization of the audio within limits set by the broadcaster, giving an immersive experience by allowing a broader soundstage to be rendered in the home and the possibility to better adapt to the device on which the audio is played. Of course, as the audio is rendered by the listening device according to what it knows about the setup in the home, best use can be made of a user's speaker placement and traditional stereo and surround can easily be replicated without the need for multiple mixes to be produced in the studio.

The work on including NGA in TS 101 154 is still in its early stages, with a particular focus on one specific commercial requirement, namely Audio Codec Performance Evaluation. This requirement is the first time that the DVB has tried to crystallize such an evaluation, and as such it is challenging the keenest minds in the business with questions such as what is better: a codec that perfectly reproduces the tone of the sound but puts it above your head when it should be to the right of the soundstage or one that puts it in exactly the right place but sounds a bit fuzzy?

When do we plan to be finished?

The goal is to complete the revised version of TS 101 154, including support for both UHD-1 Phase 2 and Next Generation Audio, in time for approval at the October Technical Module meeting. This is a tough timescale, requiring face-to-face TM-AVC meetings every two months as well as approximately two conference calls per week to progress the details of the video and the audio work. But if the level of dedication and collaboration that has been demonstrated so far can be maintained, we can do it!

Ken McCann is a director and founder of Zetacast, an independent technology consultancy company specializing in digital TV and related areas. Ken has chaired the DVB technical group responsible for audio-visual coding specifications (TM-AVC) since its inception over 20 years ago.

Virginie Drugeon is an engineer at Panasonic AVC Langen Development Center near Frankfurt, Germany. Her main areas of work include video coding, digital TV and related areas. She leads the Sub-Group responsible for drafting the text of the DVB Ultra High Definition Video standard.

David Daniels is a Senior Technologist at Sky plc. in London, and has been an active member of the DVB for many years. In that time he has chaired a number of groups and currently Co-Chairs the TM-AVC NGA Sub-Group.

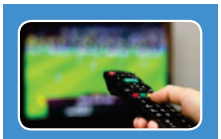
Figure 2. Summary of HDR solutions under consideration in TM-AVC

	Single Layer solutions (class of solutions using current HEVC Main 10 Profile with content graded using a new transfer function)	Dual Layer solutions (class of solutions that use scalable coding technologies)	HDR extensions to HEVC (class of solutions discussed in the MPEG HDR fast track for improved HDR coding efficiency)
Number of solutions in that class that are currently under consideration	At least 2	At least 2	At least 5
Dependencies on other standardization bodies	SMPTE, ARIB, MPEG/ITU-T (either already published or technically stable references available)	SMPTE, MPEG/ITU-T (already published references) ETSI (under development, only relevant for one of the solutions under discussion)	MPEG/ITU-T (under development)
Transfer function	Two are under consideration: HLG OETF or PQ EOTF	Typically based on PQ EOTF	Unclear for now
Coding technology	HEVC Main 10 (either edition already published or edition to be published shortly depending on the solution)	Either SHVC Scalable Main 10 or another profile of HEVC (unclear for now) depending on the solution	Potentially another profile of HEVC (unclear for now)
Metadata	Optional SEI messages such as Mastering Display Color Volume, Color Remapping Information, Content Light Level	Unclear for now	Unclear for now
Backwards compatibility to DVB UHD-1 Phase 1 receivers	Yes if HLG OETF is used No if PQ EOTF is used	Yes	Both are under consideration

Figure 3. Next Generation Audio

Accessible

Descriptive Audio, Dialog Enhancement, Multiple Language



Personalized

Modify the presentation to the listener's preference

Immersive

Put the consumer in the action and the venue

Lifelike experience

Adaptable

Optimal playback on every device