



Ultra HD Forum Guidelines

Orange Book – Foundational Technologies for Ultra HD

Ultra HD Forum

8377 Fremont Blvd., Suite 117

Fremont, CA 94538

UNITED STATES

VERSION 3.0.0

Apr 7, 2023



1. Foreword

This new version v3 of the Ultra HD Forum Guidelines provides a holistic view of modern media systems, their mechanisms and workflows, and how those are impacted by the latest generation of improvements – the “Ultra HD” technologies, those that take media beyond the limits established at the start of this millennia, characterized in large part by the video resolutions and the dynamic range offered for media in “high definition”, i.e., ITU-R Rec. BT.709. The Forum considers Ultra HD to not only be any UHD media (i.e., 4K resolution, or higher), but also HD-resolution media with enhancements such as High Dynamic Range, Wide Color Gamut, etc. Ultra HD is a constellation of technologies that can provide significant improvements in media quality and audience experience.

This work represents over eight years of collaborative effort. These new books would not have been possible without the leadership of Jim DeFilippis, who represents Fraunhofer and chairs our Guidelines Work Group with invaluable support from the co-chair, Pete Sellar of Xperi as well as technical assistance from Ian Nock of Fairmile West Consulting.

Our gratitude to all the companies listed in the Acknowledgments that have participated in this effort over the years and specifically to Nabajeet Barman (Brightcove), Andrew Cotton (BBC), Jean Louis Diascorn (Harmonic), Richard Doherty (Dolby), Chris Johns (Sky UK), Katy Noland (BBC), Bill Redmann (InterDigital), Chris Seeger (Comcast/NBCUniversal), and Alessandro Travaglini (Fraunhofer).

This document, *Foundational Technologies for Ultra HD* (Orange Book), is one of a series of books, referred to as the Rainbow Books, that compose the Ultra HD Forum Guidelines. If any of these terms sound unfamiliar, follow the link below to the Black Book. If a particular standard is of interest, links such as the one above are available to take you to the White Book, where references are collected.

The Rainbow Books are, in their entirety:

White Book [Guidelines Index and References](#)

Red Book [Introduction to Ultra HD](#)



Orange Book	<u>Foundational Technologies for Ultra HD</u>
Yellow Book	<u>Beyond Foundational Technologies</u>
Green Book	<u>Ultra HD Distribution</u>
Blue Book	<u>Ultra HD Production and Post Production</u>
Indigo Book	<u>Ultra HD Technology Implementations</u>
Violet Book	<u>Real World Ultra HD</u>
Black Book	<u>Terms and Acronyms</u>

Updates in this new version of the Ultra HD Forum Guidelines are described on the following page.

I hope you will enjoy reading today.

If you want to know more about Ultra HD, and join our discussions on how it can be deployed, I invite you to join the Ultra HD Forum. You can start by visiting our website: www.ultrahdforum.org.

Nandhu Nandhakumar, President, Ultra HD Forum
April 2023



1.1 Changes from version 2.6 to 3.0

What's new in the Spring 2023 version of the UHDF Guidelines Orange Book, *Foundational Technologies for Ultra HD* (v3.0), edited by Nabajeet Barman.

Foundational Technologies for Ultra HD is the second of the series of Rainbow Books on the Guidelines for Ultra HD. The scope and purpose of this book is to describe the foundational Ultra HD technologies, including compression technologies (HEVC and AVS 2), High Dynamic Range (HDR), Wide Color as well as audio formats, including channel based immersive audio.

While most of the information in this edition is material from the previous version of the Guidelines (v2.6), the information and references have been updated. Each section includes references to more detailed information contained in the companion Rainbow Books.

We hope this new format will be helpful in understanding UHD technologies as well as planning for new or expanded Ultra HD services.

Jim DeFilippis and Pete Sellar,

Guidelines Working Group Co-Chairs, Ultra HD Forum, April 2023



2. Acknowledgements

We would like to provide the acknowledgement to all the member companies, past and present, of the Ultra HD Forum who have contributed in some small or large part to the body of knowledge that has been contributed to the Guidelines Color Books, including the specific subject of this book.

ARRIS	ATEME	ATT DIRECTV
British Broadcasting Corporation	BBright	Beamr
Brightcove Inc.	Broadcom	B<>COM
Comcast / NBC Universal LLC	Comunicare Digitale	Content Armor
CTOIC	Dolby	DTG
Endeavor Streaming	Eurofins Digital Testing	Fairmile West
Fraunhofer IIS	Harmonic	Huawei Technologies
InterDigital	LG Electronics	Mediakind
MovieLabs	NAB	Nagra, Kudelski Group
NGCodec	Sky UK	Sony Corporation
Xperi	Technicolor SA	Verimatrix Inc.
V-Silicon		



3. Notice

The Ultra HD Forum Guidelines are intended to serve the public interest by providing recommendations and procedures that promote uniformity of product, interchangeability and ultimately the long-term reliability of audio/video service transmission. This document shall not in any way preclude any member or nonmember of the Ultra HD Forum from manufacturing or selling products not conforming to such documents, nor shall the existence of such guidelines preclude their voluntary use by those other than Ultra HD Forum members, whether used domestically or internationally.

The Ultra HD Forum assumes no obligations or liability whatsoever to any party who may adopt the guidelines. Such an adopting party assumes all risks associated with adoption of these guidelines and accepts full responsibility for any damage and/or claims arising from the adoption of such guidelines.

Attention is called to the possibility that implementation of the recommendations and procedures described in these guidelines may require the use of subject matter covered by patent rights. By publication of these guidelines, no position is taken with respect to the existence or validity of any patent rights in connection therewith. Ultra HD Forum shall not be responsible for identifying patents for which a license may be required or for conducting inquiries into the legal validity or scope of those patents that are brought to its attention.

Patent holders who believe that they hold patents which are essential to the implementation of the recommendations and procedures described in these guidelines have been requested to provide information about those patents and any related licensing terms and conditions.

All Rights Reserved

© Ultra HD Forum 2023



4. Contents

1. Foreword	2
1.1 Changes from version 2.6 to 3.0	4
2. Acknowledgements	5
3. Notice	6
4. Contents	7
5. Figures	8
6. Tables	9
7. Foundational Technologies for Ultra HD	10
7.1. Video: HEVC, AVS 2	12
7.2. HDR: HLG, HDR10, PQ 10	12
7.2.1. Perceptual Quantization (PQ) and PQ10	13
7.2.2. Hybrid Log-Gamma (HLG) and HLG10	14
7.2.3. Recommendation ITU-R BT.2100	15
7.2.4. Static Metadata – SMPTE ST 2086, MaxFALL, MaxCLL	16
7.2.5. HDR10	17
7.2.6. Foundation Ultra HD HDR Technologies	17
7.2.7. HDR10 Metadata Generation	18
7.2.8. HDR10 Metadata Carriage	18
7.3. Wide Color: Rec 2020, P3	19
7.4. Audio: 5.1 Surround and Channel based immersive	24



5. Figures

List Of Images

- [Figure 1. Channel-based Immersive Audio Live Production](#)



6. Tables

List Of Tables

- [Table 1. Foundation Ultra HD Workflow Parameters](#)
- [Table 2. Foundation Ultra HD Content Parameters](#)
- [Table 3. File-Based Signaling for SDR/BT.709](#)
- [Table 4. File-Based Signaling for SDR/BT.2020](#)
- [Table 5. File-based Signaling for HDR/BT.2020](#)



7. Foundational Technologies for Ultra HD

Foundation Ultra HD content and services are often those that were commercially available as early as 2016, and thus have reached a level of market adoption and maturity. For the purposes of this document, Foundation Ultra HD comprises the following characteristics:

- Resolution – greater than or equal to 1080p and lower than or equal to 2160p, (progressive format; BT.2100 [\[5\]](#) does not include interlaced formats)
- Wide Color Gamut – color gamut wider than BT.709 [\[2\]](#)
- High Dynamic Range – Image dynamic range that provides a dynamic range larger than SDR, capturing and displaying increased highlight and shadow details. Use of tone curves as referenced in BT. 2100 (HLG and PQ).
- Bit depth – 10-bit
- Frame rates – up to 60fps (integer frame rates are preferred; note that cinematic content may opt to use lower frame rates, e.g., see DCI specification [\[12\]](#))
- Audio – 5.1 or higher channel surround sound or channel-based Immersive Audio
- Closed Captions/Subtitles – CTA 708/608, ETSI 300 743, ETSI 300 472, SCTE-27, IMSC1

The following terms are used in this document for HDR and HDR plus WCG:

- HLG: Hybrid Log-Gamma (HLG) OETF, EOTF, and OOTF transfer functions specified in BT.2100 [\[5\]](#)
- HLG10: HDR systems or content employing Hybrid Log-Gamma (HLG), the wide color gamut specified in BT.2100 [\[5\]](#) and 10-bit depth
- PQ: HDR systems or content employing Perceptual Quantization (PQ), the wide color gamut specified in BT.2100 [\[5\]](#), and 10-bit depth
- PQ10: HDR systems or content employing Perceptual Quantization (PQ), the wide color gamut specified in BT.2100 [\[5\]](#), and 10-bit depth
- HDR10: HDR systems or content employing PQ10 and further including or capable of providing SMPTE ST 2086 [\[10\]](#), MaxFALL, and MaxCLL metadata



Content Creation & Mastering	Defined and documented standard workflows for Live and Pre-recorded content
Service Type	Real-time Program Services; On-Demand content that was originally offered as Live content
Network Type	Unicast (including Adaptive Bit Rate) Broadcast, Multicast
Transport/Container/Format	MPEG TS, Multicast IP, DASH ISO BMFF
Interface to TVs (source format)	IP connected (for OTT content delivered via managed or unmanaged network) HDMI (for services delivered via a STB, e.g., OTT, MVPD)
Backward Compatibility	Native (HLG), simulcast (HDR10/PQ10), decoder based (optional)

Table 1. Foundation Ultra HD Workflow Parameters

Spatial Resolution	1080p ¹ or 2160p
System Colorimetry	BT.709 [2], BT.2020 [3]
Bit Depth	10-bit
Dynamic Range	SDR, PQ, HLG
Frame Rate ²	24(23.976), 25, 30(29.97), 50, 60(59.94)
Video Codec	HEVC, Main 10 Profile, Level 5 or 5.1 (single layer) ³
Audio Channels	5.1 or higher channel surround sound or channel-based Immersive Audio
Audio Codec	AC-3, E-AC-3, DTS-HD, E-AC-3 + JOC, HE-ACC, AAC-LC
Captions/Subtitles Coding	CTA 608/708, ETSI 300 743, ETSI 300 472, SCTE-27, IMSC1

Table 2. Foundation Ultra HD Content Parameters

Table 2 Notes:



1. 1080p, together with WCG and HDR, fulfills certain use cases for Foundation Ultra HD services and is therefore considered to be an Ultra HD format for the purpose of this document. 1080p without WCG or HDR is considered to be an HD format. The possibility of 1080i or 720p plus HDR and WCG is not considered here. HDR and WCG for multiscreen resolutions may be considered in the future.
2. Fractional frame rates for 24, 30, and 60 fps are included for compatibility with current plant video clock reference but are ultimately not preferred. Fractional frame rates will be necessary during migration from legacy video systems. The lower frame rates may be common for cinematic content.
3. For use in China, the AVS2 codec, Main10 profile, is used in addition to HEVC. See [Section 7.4.1 in the Yellow Book](#).

For the purpose of this document, including the above constraint on 1080p content, various combinations of these Foundation UltraHD parameters can be combined to produce Ultra HD content. Additional, non-Foundation technologies may also be employed.

7.1. Video: HEVC, AVS 2

The Foundation Ultra HD codec is HEVC for distribution to consumers, due to its support for HDR/WCG (10-bit) as well as improved coding efficiency, which makes 4K content more feasible. AVC specifications have been updated to include support for 10-bit depth with BT.2020/2100 system colorimetry and support of PQ/HLG High Dynamic Range formats. However, the Ultra HD Forum finds that deployed consumer decoders are mostly hardware-based with no feasible mechanism to upgrade to match these latest updates, and thus these new AVC capabilities are not generally usable for distribution purposes. For AVC, consumer decoders generally support 8-bit 4:2:0 formatted content normally with limitations with regard to the maximum bit rates that would normally preclude decoding 4K content at the compression rates that AVC can achieve.

Using AVC for production, contribution and mezzanine workflows is more viable because changing to encoders with support of the HDR/10-bit capabilities is something that is more easily accomplished.

7.2. HDR: HLG, HDR10, PQ 10

There are many terms in use in the field of HDR television. This section explains the existing terminology and how terms are used in these Guidelines.

Note that currently some UHD displays are capable of accepting BT.2020 [3] content, but as of this publication, no direct view display is available that is capable of rendering the full gamut of colors in the BT.2020 [3] system colorimetry. It is assumed that in these cases, the device employs “best effort” gamut mapping tailored to its particular display characteristics, and thus these devices are considered BT.2020 [3]-compatible.



7.2.1. Perceptual Quantization (PQ) and PQ10

One HDR transfer function set defined for use in television is the “Perceptual Quantization” (PQ) developed by Dolby. PQ is defined as a reference display transfer curve, or EOTF, that is employed on the mastering monitor used in the reference viewing environment. The transfer function is designed to minimize the visibility of banding on a display over the brightness range of 0 to 10,000 cd/m². SMPTE standardized the PQ EOTF in ST 2084 [9] and the ITU-R standardized the EOTF, OETF, and OOTF transfer functions in BT.2100 [5]. ST 2084 specifies the shape of the curve over a range of 0 to 1 but does not specify SDI code values for the 0 and 1 values. Table 9 in BT.2100 [5] describes the code mapping for both narrow range and full range code values¹.

The Ultra HD Forum has defined the term PQ10 to refer to an HDR system employing the PQ transfer functions and wide gamut colorimetry specified in BT.2100 [5] and 10-bit depth. PQ10 content or systems may or may not be metadata-capable. For example, HDR10 is a Foundation Ultra HD PQ10-based format that has the capability of including certain static metadata (see [Section 7.2.4](#)).

The PQ signal is “display-referred”, meaning that the pixel-encoded values represent specific values of luminance for displayed pixels. The intent is that only the luminance values near the minimum or maximum luminance capability of a display are necessarily adjusted to utilize the available dynamic range of the display. Some implementations may apply a “knee” at the compensation points in order to provide a smoother transition from the coded values to the display capabilities, e.g., to avoid “clipping”.

When default display settings are engaged, PQ enables pixel values in the mid-range, including skin tones, to be rendered on a display at the same (absolute) luminance level that was determined at production.

For example, if a scene was graded on a 1000 cd/m² grading monitor and then displayed on a 4000 cd/m² display, the skin tones can be rendered at the same luminance values on the 4000-nit display as on the 1000-nit monitor per the grader’s intent, while the speculars and darker tones can be smoothly extended to take full advantage of the 4000-nit display. [See also Section 7.5 of the Blue Book](#).

The Ultra HD Forum InterOp testing has shown that consumer displays vary significantly, especially when different viewing “modes” are selected (e.g., “sports mode” or “movie mode”). Research has also shown that ambient light in the viewing room can have an impact on humans’ perception of a displayed image. Displays used in non-reference viewing environments may employ adjustments to the PQ curve in order to provide some compensation for the difference between the actual viewing environment and the reference viewing environment.

¹CTA 861-G[31] offers additional information about code ranges of video signals.



PQ content requires a down-mapping step in order to provide acceptable SDR quality. See Section 0 below for a deeper discussion of backward compatibility, including the pros, cons, and open questions that apply to various possible methods.

7.2.2. Hybrid Log-Gamma (HLG) and HLG10

Another HDR transfer function defined for use in television is “Hybrid Log-Gamma” (HLG) developed by the BBC and NHK. This is defined as a camera capture transfer curve, or OETF. This curve was designed to provide HDR while maintaining a degree of backward compatibility with SDR/BT.2020 displays. The HLG curve was first specified in ARIB STD-B67 [15]², and later in ITU-R BT.2100 [5]. The two OETF specifications are identical. BT.2100 [5] additionally specifies the display EOTF and the OOTF, which are not defined by the ARIB standard.

The Ultra HD Forum has defined the term HLG10 to refer to an HDR system employing HLG, wide gamut colorimetry specified in BT.2100 and 10-bit depth with black at code 64, and nominal peak at code 940³.

HLG is a “scene-referenced” HDR technology that uses pixel-encoded values that are intended to represent picture luminance levels relative to each other in a given scene. The intent is that all the displayed pixel luminance values may be adjusted in a defined manner to compensate for specific display capabilities (e.g., peak luminance) and viewing environments in such a way that the values retain their perceptual appearance relative to one another.

As the eye’s sensitivity to light intensity is approximately logarithmic, a power (or “gamma”) law is applied to the HLG relative scene-light pixel values, to scale them to span the luminance range of the display; thereby approximating the same relative subjective brightness values that were determined at production. Since no one luminance range is “fixed”, the (absolute) displayed luminance of mid-range values, including skin tones, will increase or decrease to scale along with all other values in order to better preserve the appearance of the relative brightness values of the pixels in the scene.

For example, if a scene was graded on a 1000 cd/m² grading monitor and then displayed on a 2000 cd/m² display, the skin tones and other scene elements will be brighter on the 2000 cd/m² display than on the 1000 cd/m² monitor. However, given that all pixel values are adjusted through the “gamma law”, the overall image remains perceptually similar to the original “look” selected by the grader. See also Section 6.1.10⁴.

² The original v1.0 ARIB specification referred to the breakpoint between the logarithmic and gamma portions of the HLG curve as the “reference white level”, which some manufacturers incorrectly assumed to be the signal level for “HDR Reference White”. Later versions of the ARIB specification have removed the term “reference white level” to avoid any such confusion. The “HDR Reference White” level for both HLG and PQ production is defined in ITU-R BT. 2408 [8].

³ CTA 861-G [14] offers additional information about code ranges.

⁴ It is well known that changes in brightness of a display may affect the perception of image color (Hunt Effect) [16] as well as image contrast (Stevens) [17]. Research on the Color Appearance Model (CAM),



As mentioned above, Ultra HD Forum InterOp testing has shown that consumer displays vary significantly, especially when different viewing “modes” are selected (e.g., “sports mode” or “movie mode”). Research has also shown that ambient light in the viewing room can have an impact on humans’ perception of a displayed image⁵. Section 6.2 of ITU-R report BT.2390 [6] describes how the display gamma may be adjusted to provide some compensation.

Content produced using HLG can be displayed on SDR/WCG devices with a degree of compatibility that may be judged acceptable for programs and services according to Report ITU-R BT.2390 [6], and subjective tests performed by the EBU, RAI, IRT and Orange Labs [24]. Backward-compatible HLG is only intended to support SDR/BT.2020 (WCG) displays and not intended for displays which only support SDR/BT.709. See Section 0 below for a deeper discussion of backward compatibility, including the pros, cons and open questions that apply to various possible methods.

7.2.3. Recommendation ITU-R BT.2100

In July 2016, ITU-R Study Group 6 approved the publication of a Recommendation on HDR for use in production and international program exchange and is known as ITU-R BT.2100 [5]. This recommendation includes the following specifications:

- Spatial resolutions: 1080p, 2160p, 4320p
- Frame rates: 24/1.001, 24, 25, 30/1.001, 30, 50, 60/1.001, 60, 100, 120/1.001, 120
- System Colorimetry: Same as BT.2020 [3]
- Reference Viewing Environment: 5 cd/m² background and less than or equal to 5 cd/m² surround
- Reference non-linear transfer functions EOTF, OOTF, OETF: PQ and HLG
- Signal Format: Y’C_B’C_R’, IC_TC_P, and RGB.
- Color sub-sampling: same alignment as specified in BT.2020 [3]
- Bit depth value ranges:

where brightness exceeds 700 cd/m², agrees with these observations [17]; however as of 2020 the BBC has found no evidence of significant changes to the perception of color or contrast in their research using HLG with displays of different peak brightness [18], additional research is on-going on these effects with respect to HDR imagery.

⁵ Simon Thompson, Katy Noland and Manish Pindoria, “Evaluation of Required Adjustments for HDR Displays under Domestic Ambient Conditions”, Proc. ICCE-Berlin, Berlin, Germany, September 2017.

Hunt, R. W. G., “The Reproduction of Colour in Photography, Printing and Television,” Section 6.5 “Different Surround Conditions”, 6th ed., John Wiley & Sons Ltd, 2004.

Bartleson, C. J. and Breneman, E. J., “Brightness Perception in Complex Fields,” J. Optical Soc. America, vol. 57, no. 7, pp. 953-957, July 1967.



- 10 bit, Narrow (64-940) and Full (0-1023) ranges
- 12 bit, Narrow (256-3760) and Full (0-4095)
- Floating point signal representation: Linear RGB, 16-bit floating point

BT.2100 [5] is the primary reference document on HDR for use in production and international program exchange. Note that a full signal specification will need to include the following attributes: spatial resolution, frame rate, transfer function (PQ or HLG), color signal format, integer (10 or 12 bits, narrow or full range) or floating point.

Not all of the parameter values listed above have been deployed by Ultra HD service providers, but are included as informative details.

7.2.4. Static Metadata – SMPTE ST 2086, MaxFALL, MaxCLL

SMPTE has specified a set of static metadata in the ST 2086 [10] Mastering Display Color Volume Metadata Supporting High Luminance and Wide Color Gamut Images standard. Parameters included indicate the characteristics of the mastering display monitor. The mastering display metadata indicates that the creative intent was established on a monitor having the described characteristics. If provided, the implication is that the artistic intent of the content is within the subset of the overall container per the metadata values. The mastering display characteristics include the display primaries and white point as x,y chromaticity coordinates, and the maximum and minimum display luminance. For example, the metadata may indicate that the system colorimetry of the mastering display is the DCI-P3 gamut in the BT.2020 [3] container and the luminance range is a subset of the 0 to 10,000 cd/m² representable in PQ.

The Blu-ray Disc Association and DECE groups have defined carriage of two additional metadata items:

- MaxFALL – Maximum Frame Average Light Level; this is the largest average pixel light value of any video frame in the program
- MaxCLL – Maximum Content Light Level: this is the largest individual pixel light value of any video frame in the program

Static metadata may be used by displays to control color volume transform of the received program to better reproduce the creative intent as shown on the mastering display, given the capabilities of the display device. However, the use of MaxFALL and MaxCLL static metadata have limitations for use with Live broadcasts since it is difficult to determine a program's maximum pixel light values during a Live production. According to the UHD Alliance, today's mastering practices may generate outlier values unintentionally, causing the content's associated MaxCLL value to be higher than expected. As a response to that, some content providers use statistical analysis to calculate a MaxCLL value that is more representative of the statistically significant brightest pixels contained in the image sequence. The Ultra HD Forum further suggests that such statistical methodology may also apply to MaxFALL.



It is worth noting that code levels below minimum mastering display luminance and code levels above maximum mastering display luminance are likely clipped on a professional reference monitor and therefore any detail in pictures utilizing code levels outside this range are likely not seen in the content production process.

7.2.5. HDR10

HDR10 is a PQ10-based HDR format. The term “HDR10” is in widespread use and has been formally and consistently defined by several groups, including DECE and BDA, as:

- Transfer function: Rec. ITU-R BT.2100 [\[5\]](#) (PQ)
- System Colorimetry: Rec. ITU-R BT.2100 [\[5\]](#)
- Bit depth: 10 bits
- Metadata: SMPTE ST 2086 [\[10\]](#), MaxFALL, MaxCLL

Several delivery formats (e.g. Ultra HD Blu-ray™ and CTA’s HDR10 profile) have specified delivery using the above parameters with the metadata mandatory but are still considered to be using HDR10 for the purposes of this document. While ATSC A/341 does not use the term “HDR10,” when the PQ transfer function is used, the static metadata video parameters described therein are consistent with HDR10 as defined herein.

Note that some displays ignore some or all static metadata (i.e., ST 2086, MaxFALL, and MaxCLL); however, HDR10 distribution systems must deliver the static metadata, when present.

7.2.6. Foundation Ultra HD HDR Technologies

The following HDR systems are recommended for Foundation Ultra HD due to their conformance with the criteria listed in [Section 7.2.1](#) and [7.2.2](#) above:

- HLG10: HDR systems or content employing Hybrid Log-Gamma (HLG), the wide color gamut specified in BT.2100 [\[5\]](#) and 10-bit depth
- PQ10: HDR systems or content employing Perceptual Quantization (PQ), the wide color gamut specified in BT.2100 [\[5\]](#), and 10-bit depth
- HDR10: HDR systems or content employing PQ10 and further including or capable of providing SMPTE ST 2086 [\[10\]](#), MaxFALL, and MaxCLL metadata

It should be noted that Real-time Program Services are typically comprised of both Live and Pre-recorded content, and it is not recommended that service providers alternate between SDR and HDR signal formats or mix different HDR formats. See [Section 9.1 of the Blue Book](#) for details.



7.2.7. HDR10 Metadata Generation

HDR10 includes the static metadata described in [Section 7.2.4](#). MaxFALL and MaxCLL metadata could be generated by the color grading software or other video analysis software. In Live content production, MaxFALL or MaxCLL metadata is not generated during the production process. By definition, it is not possible to generate MaxFALL or MaxCLL for a Live program because these cannot be known until the entire program is produced, i.e., after the program is over. The special values of '0' (meaning, “unknown”) are allowed for MaxFALL and MaxCLL. It may be possible to set limits on the output and thus pre-determine MaxFALL and MaxCLL even for Live content production. For example, if MaxFALL and MaxCLL metadata values are provided, a video processor could be inserted in order to reduce brightness of any video frames that would exceed the indicated values (similar to the way audio processors are often inserted to enforce audio loudness and peak levels).

If it is desired to use HDR10 in Live content production, but the production facility does not support carriage of the metadata, then default values for ST 2086 [\[10\]](#), MaxFALL and MaxCLL should be determined and entered directly into the encoder via a UI or a file. SMPTE 2086 metadata could be set to values that represent the monitors used for grading during the production of the content.

7.2.8. HDR10 Metadata Carriage

When the Ultra HD Forum Guidelines were originally published, standards did not exist to carry HDR10 static metadata over an SDI interface. The metadata had to be either embedded within the content intended for encoding (for Pre-recorded content) or supplied to the encoder via a file or UI. However with the publication of SMPTE ST 2108-1 [\[21\]](#) and ST 2108-2 [\[19\]](#), there are now a variety of static and dynamic color transform metadata items that can be carried in the Vertical Ancillary Data Space (VANC) as Ancillary Data (ANC) Messages over an SDI interface or over an SMPTE ST 2110 IP connection (per SMPTE ST 2110-40 [\[20\]](#), See also [Section 7.5.1 of the Blue Book](#)).

In HEVC [\[23\]](#) and AVC [\[22\]](#) bitstreams, MaxFALL and MaxCLL may be carried in the Content Light Level (CLL) static SEI message (sections D.2.35 (Syntax) and D.3.35 (Semantics), [\[23\]](#)). SMPTE ST 2086 metadata is carried in the Mastering Display Color Volume (MDCV) SEI, (section D.2.28 (Syntax) and D.3.28 (Semantics), [\[23\]](#)). Both the CLL and MDCV SEI messages can be represented by their corresponding ANC messages as specified by SMPTE ST 2108-1 [\[21\]](#) and accompany each video frame to which they apply. An SMPTE ST 2108-1 ANC message is a simple, bit-accurate encapsulation of the corresponding SEI message, to ease handling by an encoder or decoder; as a result, SMPTE ST 2108-1 is also able to support certain dynamic metadata technologies (see Section 12).

Note that there may be multiple encoders in the end-to-end broadcast chain. In the event that HDR10 is used and the metadata does not reach the decoder/display, it is expected that the decoder/display will employ “best effort” to render the content accurately. Note that such an



end-to-end chain is operating in a manner indistinguishable from PQ10 content (i.e. PQ without metadata).

7.3. Wide Color: Rec 2020, P3

The system colorimetry, transfer function (SDR or HDR), and matrix coefficients must be known to downstream equipment ingesting or rendering content in order to preserve display intent. This is true for file transfers in file-based workflows and in linear content streams in linear workflows.

In file-based workflows, mezzanine file formats such as IMF/MXF (stored in picture essence descriptors) and QuickTime (stored in Video Information NCLC atoms) are often used.

In linear production workflows, these values are transmitted via VPID (Video Payload Identifier) within HD-SDI [99]. In linear compressed transmission workflows, these values are typically signaled in the VUI of an H.265/HEVC or H.264/MPEG-4 AVC bitstream. Details on SEI and VUI messaging are available in the HEVC specification [23], in particular, Appendix D (SEI) and Appendix E (VUI) of reference [23].

The tables below ([Table 3](#), [Table 4](#) and [Table 5](#)) summarize HEVC Main10 Profile bitstream SDR, PQ and HLG indicators. (In HEVC and AVC specifications, the bitstream elements are bolded and italicized to distinguish them from temporary variables and labels). As shown in Table 9 there are two methods of signaling the HLG transfer function.



	<u>System Identifier</u>	<u>BT. 709 YCC</u>	<u>BT. 709 RGB</u>	<u>Full-Range 709 RGB</u>	<u>BT. 601 525</u>	<u>BT. 601 625</u>
Color properties	Color primaries	BT.709	BT.709	BT.709	BT.601	BT.601
	Transfer Characteristics	BT.709	BT.709	BT.709	BT.709	BT.709
	Signal Format	Y'CbCr	R'G'B'	R'G'B'	Y'CbCr	Y'CbCr
Other	Full/narrow range	Narrow	Narrow	Full	Narrow	Narrow
	4:2:0 chroma sample location alignment	Interstitial	N/A	N/A	Interstitial	Interstitial
CICP parameters Rec. ITU-T H.273 ISO/IEC 23091-2 [113] (QuickTime/HE/AVC) ¹	ColorPrimaries	1	1	1	6	5
	TransferCharacteristics	1	1	1	6	6
	MatrixCoefficients	1	0	0	6	5
	VideoFullRangeFlag	0	0	1	0	0
SMPTE MXF parameters SMPTE ST 2067-212	Color primaries	06.0E.2B.34.04.01.01.06.04.01.01.01.03.03.00.00			06.0E.2B.34.04.01.01.06.04.01.01.01.03.03.00.00	06.0E.2B.34.04.01.01.06.04.01.01.01.03.02.00.00
	Transfer Characteristic	06.0E.2B.34.04.01.01.01.04.01.01.01.01.02.00.00				
	Coding Equations	06.0E.2B.34.04.01.01.01.04.01.01.01.01.00.00	N/R	N/R	06.0E.2B.34.04.01.01.01.04.01.01.01.02.01.00.00	
	Full/narrow level range	Inferred (indicated in black reference level, white reference level, Color range)				
	4:2:0 chroma sample location alignment	Inferred (ChromaLocType=0)	N/A	N/A	Inferred (ChromaLocType=0)	Inferred (ChromaLocType=0)

Table 3. File-Based Signaling for SDR/BT.709

Notes to Table 3:

1. QuickTime/MP4/HEVC/AVC: ITU-T Rec-H.Sup 19, Usage of Video Signal Type Code Points, April 2021, <https://www.itu.int/ITU-T/recommendations/rec.aspx?rec=14652&lang=en> 2. SMPTE UL's available: https://registry.smpte-ra.org/view/published/labels_view.htm |



	System Identifier	BT.2020 YCC	BT.2020 RGB
Color properties	Color Primaries	BT.2020	BT.2020
	Transfer Characteristics	BT.2020	BT.2020
	Signal Format	Y'CbCr	R'G'B'
Other	Full/narrow range	Narrow	Narrow
	4:2:0 chroma sample location alignment	Co-sited	N/A
CICP parameters Rec. ITU-T H.273 ISO/IEC 12091-2 [113] (QuickTime/ HEVC/AVC) ¹	Color Primaries	9	9
	TransferCharacteristics	14	14
	MatrixCoefficients	9	0
	VideoFullRangeFlag	0	0
SMPTE MXF parameters SMPTE ST 2067-21 ²	Color Primaries	06.0E.2B.34.04.01.01.0D.04.01.01.01.03.04.00.0	
	Transfer Characteristic	06.0E.2B.34.04.01.01.0E.04.01.01.01.01.09.00.00	
	Coding Equations	06.0E.2B.34.04.01.01.0D.04.01.01.01.02.06.00.00	N/R
	Full/narrow level range	Inferred (indicated in black reference level, white reference level, color range)	
	4:2:0 chroma sample location alignment	Inferred (ChromaLocType = 2)	N/A

Table 4. File-Based Signaling for SDR/BT.2020

Table 4 Notes:

1. QuickTime/MP4/HEVC/AVC: ITU-T Rec-H.Sup 19, Usage of Video Signal Type Code Points, April 2021. <https://www.itu.int/ITU-T/recommendations/rec.aspx?rec=14652&lang=en>
2. SMPTE UL's available: https://registry.smpte-ra.org/view/published/labels_view.html



	<u>System Identifier</u>	<u>BT2100</u> <u>PQ</u> <u>YCC</u>	<u>BT. 2100</u> <u>HLG</u> <u>YCC</u>	<u>BT.2100</u> <u>PQ</u> <u>IC_TC_P</u>	<u>BT.2100</u> <u>PQ</u> <u>RGB</u>	<u>BT.2100</u> <u>HLG</u> <u>RGB</u>
Color properties	Color primaries	BT.2020 / BT.2100	BT.2020 / BT.2100	BT.2100	BT.2020 / BT.2100	BT.2020 / BT.2100
	Transfer Characteristics	BT.2100 PQ	BT.2100 HLG	BT.2100 PQ	BT.2100 PQ	BT.2100 HLG
	Signal Format	Y'CbCr	Y'CbCr	IC _T C _P	R'G'B'	R'G'B'
Other	Full/narrow range	Narrow	Narrow	Narrow	Narrow	Narrow
	4:2:0 chroma sample location alignment	Co-sited	Co-sited	Co-sited	N/A	N/A
CICP parameters Rec. ITU-T H.273 ISO/IEC 23091-2 [113] (QuickTime/ HEVC/ AVC) ¹	Color Primaries	9	9	9	9	9
	Transfer Characteristics	16	18	16	16	18
	MatrixCoefficients	9	9	14	0	0
	Video Full-Range Flag	0	0	0	0	0
SMPTE MXF parameters SMPTE ST 2067-21 ²	Color Primaries	06.0E.2B.34.04.01.01.0D.04.01.01.01.03.04.00.00				
	Transfer Characteristic	06.0E.2B.34.04.01.01.0D.04.01.01.01.01.01.01.0A.00.00	06.0E.2B.34.04.01.01.0D.04.01.01.01.01.0B.00.00	06.0E.2B.34.04.01.01.0D.04.01.01.01.01.0A.00.00	06.0E.2B.34.04.01.01.0D.04.01.01.01.01.0A.00.00	06.0E.2B.34.04.01.01.0D.04.01.01.01.01.0B.00.00
	Coding Equations	06.0E.2B.34.04.01.01.0D.04.01.01.01.02.06.00.00		06.0E.2B.34.04.01.01.0D.04.01.01.01.02.07.00.00	N/R	N/R
	Full/narrow level range	Inferred (indicated in black reference level, white reference level, Color range)				
	4:2:0 chroma sample location alignment	Inferred (ChromaLocType = 2)	Inferred (ChromaLocType = 2)	unknown	N/A	N/A

Table 5. File-based Signaling for HDR/BT.2020



Table 5 Notes:

1. QuickTime/MP4/HEVC/AVC: ITU-T Rec-H.Sup 19, Usage of Video Signal Type Code Points, April 2021. <https://www.itu.int/ITU-T/recommendations/rec.aspx?rec=14652&lang=en>
2. SMPTE UL's available: https://registry.smpte-ra.org/view/published/labels_view.html

In one method, the SDR transfer function indicator is signaled in the VUI, and the HLG transfer function indicator is transmitted using an alternative transfer characteristics SEI message embedded in the bitstream. In this way, an “HLG aware” STB or decoder/display would recognize that the bitstream refers to content coded with HLG (since it is indicated by the preferred_transfer_characteristics syntax element of the SEI). If an “HLG aware” STB is connected to a TV that does not support HLG, the STB would transmit the SDR indicator over HDMI to the TV. If it is connected to a TV that supports HLG, the STB would copy the transfer function value in the SEI (to indicate HLG) and transmit this over HDMI to the TV.

In the other method, the HLG transfer function indicator is directly signaled in the VUI in the same way PQ or SDR would be signaled.

In theory, it is possible to achieve a lossless conversion between the two methods of signaling HLG by flipping the VUI transfer function characteristics indicator value and inserting or removing the alternative transfer characteristic SEI.

Using the first method (i.e., including the SDR transfer function indicator in the VUI and the HLG transfer function indicator in the SEI) enables backward compatibility with SDR/WCG displays. Service providers may also deem results to be acceptable on SDR/WCG displays using the second method (i.e., including the HLG transfer function indicator in the VUI). Service providers may wish to test both methods.

Additional use cases for production-level signaling and metadata in wrappers include color atoms in the Apple QuickTime™ Wrapper via NCLC (Non-Constant Luminance Coding) and MDCV (Mastering Display Color Volume), CLLI (For MaxFall/MaxCLL), using MPEG C1CP (Coding Independent Code Point) and MXF Image Essence Descriptors from ST.2067-21:2016 [13].

As described in [Section 9 of the Yellow Book](#), service providers should convert or remap all content into a single, consistent system colorimetry and transfer function. Setting the initial values in the encoder should be adequate, assuming the encoder is dedicated to a single format.

As of publication, while there are methods for signaling system colorimetry, transfer function, matrix coefficients, and HDR-related metadata through the end-to-end supply chain, however, due to devices with limited support for these capabilities, it is extremely difficult to successfully ensure that these signals and metadata survive through the entire broadcast linear production chain. Because gaps exist with legacy devices using SDI interfaces, as well as HDMI interfaces



and software that don't consistently support file-based signaling in mezzanine file wrappers, verification of proper signaling is recommended. The Ultra HD Forum observes that standards bodies have attempted to address these issues and provided documentation, much of which is referenced herein but some of which is still under development.

7.4. Audio: 5.1 Surround and Channel based immersive

For a live event, Immersive Audio can be created using existing mixing consoles and microphones used during the event. Using additional console busses, height-based ambience, and effects can be added to a traditional 5.1 or 7.1 channel mix which can then be encoded by an ETSI TS 103 420 compliant E-AC-3 encoder. A local confidence decoder can be employed to check typical downmixes, including the backwards compatible 5.1 channel render described in ETSI TS 103 420. During normal mixing, this confidence decoder can serve as a useful continuity check and display (i.e. to make sure the mix is still "on-air"), though due to normal latencies, it will likely be found to be impractical to be kept in the monitor path full time. A high-level diagram of a live mixing environment recently used at a major televised event using Dolby Atmos using E-AC-3+JOC [11] is shown in [Figure 1](#) below.

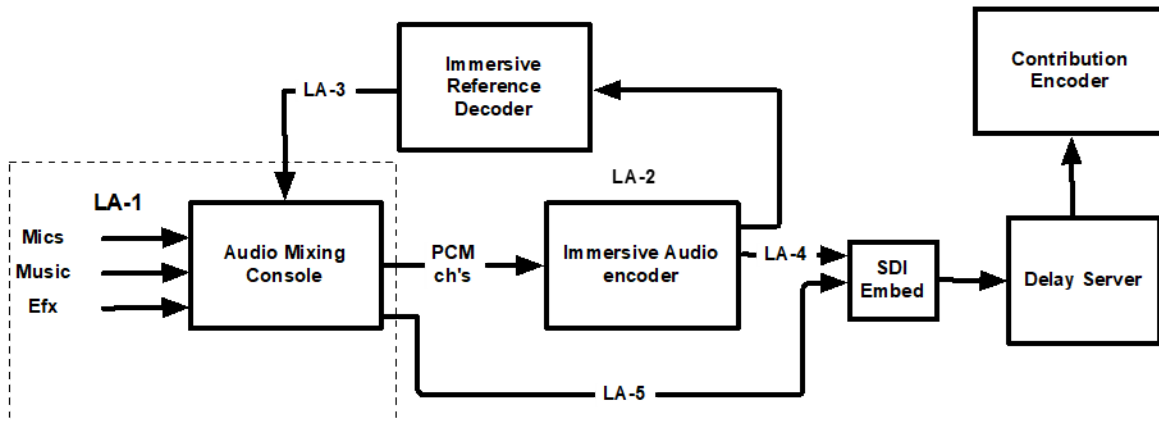


Figure 1. Channel-based Immersive Audio Live Production



Reference Point	Content Creation Functions	Reference Point Description
LA-1	Audio Capture and mixing	Microphones are placed throughout the live venue. Microphone feeds are brought into the audio mixing console in the same fashion as 5.1 production. Sound elements are mixed into a final immersive program.
LA-2	Immersive Audio Authoring	Audio groups are created in the mixing console representing the immersive mix as channel-based (e.g. 5.1.2, 5.1.4, 7.1.4) audio ⁶ .
LA-3	Immersive Audio Monitoring	The Next-Generation Audio Processor renders the audio to 5.1.2, 5.1.4 or 7.1.4 speaker feeds and sends these feeds back to the audio mixing console for monitoring in the live mix environment.
LA-4	Immersive Audio Encoding	Atmos immersive program, encoded as E-AC-3+JOC [11] is delivered as a 5.1 channel bitstream + parametric side-data (steering data) to the contribution encoder ⁷ transported over either a MADI or SDI link.
LA-5	Legacy Audio Delivery	Stereo or 5.1 complete mixes may be created at the audio mixing console and delivered via traditional means.

In this channel-based immersive audio example, a Dolby Atmos enabled E-AC-3+JOC [11] encoder generates a compressed bitstream containing 5.1 channels of audio that is backwards compatible with legacy E-AC-3 decoders. In parallel, the encoder generates an additional set of parameters (specified in ETSI TS 103 420 [11]) that are carried in the bitstream, along with the 5.1 audio, for use by a Dolby Atmos E-AC-3+JOC [11] decoder. The full Atmos decode process reconstructs the original channel-based immersive audio source from the 5.1 backwards compatible representation and the additional parameters. A typical channel-based immersive program can be encoded, for example, at 384-640 kbps total, thus fitting into existing emission scenarios.

(End of Orange Book)

⁶ The downstream Atmos Channel-Based Immersive emissions encoder, using E-AC-3 + JOC [11] will render a legacy 5.1 audio program. It is recommended to verify the rendered 5.1 audio program using a suitable E-AC-3 decoder in the monitor chain.

⁷ See [Section 11.5 of the Violet Book](#) regarding backward compatibility.