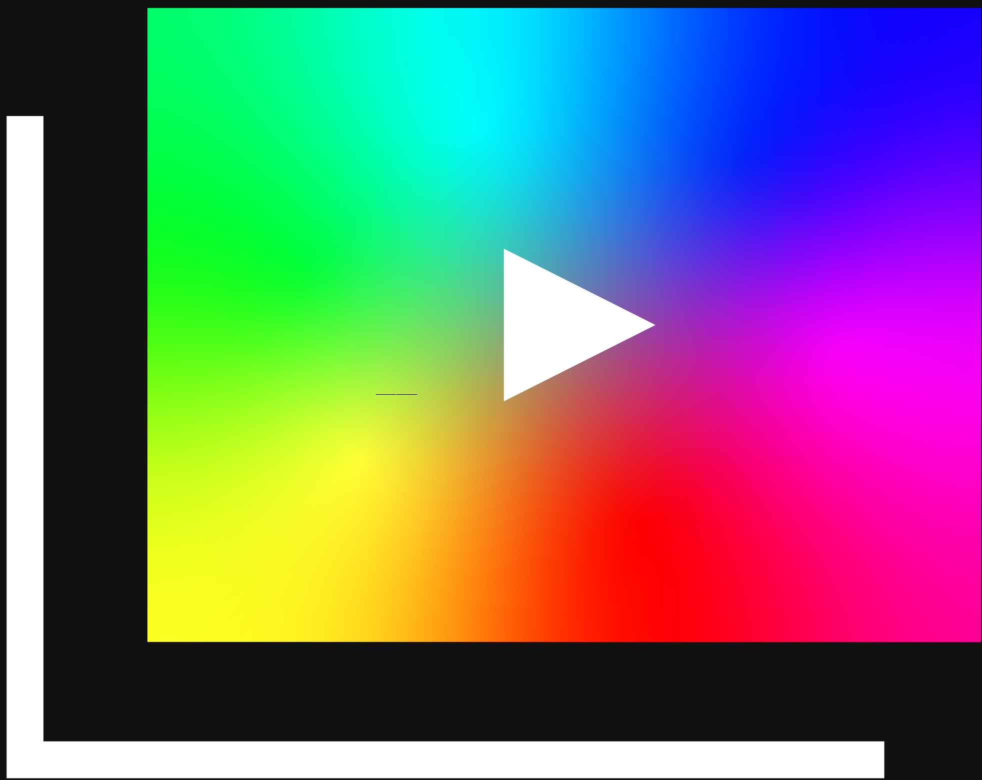
BITMOVIN



HANDBOOK

The practical guide to HDR by Bitmovin

Table of Contents

Introduction

Since its introduction, High Dynamic Range (HDR) video has quickly grown from a novel innovation to an expected feature when shopping for a new television. Brighter highlights, more detailed shadows and wider color palettes are all possible with HDR, so content creators and streaming platforms should be taking advantage of these improvements to create the best possible viewing experiences. Unfortunately, that's easier said than done, with multiple HDR formats and fragmented device support adding more complications to an already complex ecosystem.

To help demystify and highlight the essential considerations on the topic of HDR, we've created this practical guide to provide the background knowledge around formats, conversions and playback that will empower readers to make informed decisions about their own HDR workflows and support. This is not intended to be exhaustive or comprehensive, but we have added useful links for anyone interested in more detail.

HDR vs SDR

There are 3 main factors that allow HDR formats to provide better visual quality than SDR:

Higher Dynamic Range - There is more contrast available between the light and dark areas of HDR video. The darks are darker, the bright areas are brighter, and there is more detail in both the shadows and the highlights.

Wider Color Gamut - HDR formats utilize an expanded range of colors that allow them to more closely match the capabilities of the human eye.

Higher Bit Depth - HDR content is encoded with higher bit depths than SDR, allowing for finer distinctions, smoother gradients, and more accurate representations of colors, tones and brightness.

Transfer Functions

For video purposes, transfer functions are used to map the human eye's brightness and contrast detection capabilities to the electronic devices being used to capture and display video. Humans evolved to have better contrast perception in darker regions than in brighter ones and transfer functions account for that discrepancy, enabling higher dynamic range by ensuring data is allocated to areas where it can be appreciated. HDR formats utilize and apply more advanced transfer functions that bring us closer to the limits of human perception, taking advantage of the enhanced brightness and contrast capabilities of more modern displays and projectors.

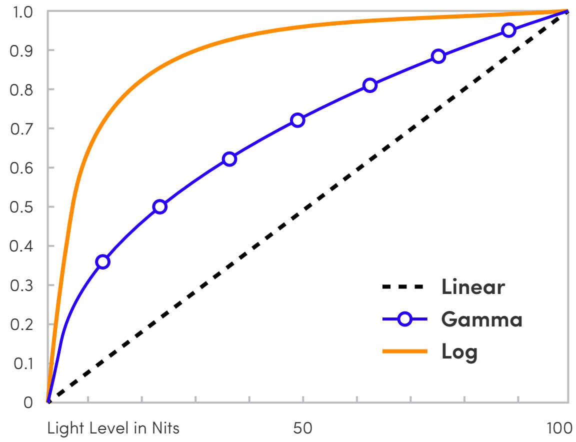
You may see mention of Opto-Electrical Transfer Functions(OETF), used by cameras to convert visible light into an electric signal and their inverse, counterpart Electro-Optical Transfer Functions (EOTF) used by displays to convert the electric signal into visible light patterns that we recognize as video. Sometimes they are combined and referred to as a combined end-to-end Optical-Optical Transfer Function (OOTF).

Common Photo and Video Transfer Functions

Linear – For applications where transmission bandwidth is not a constraint, linear transfer functions can still be employed to preserve the maximum amount of source information, albeit with extremely high data rates. This is common when shooting in RAW formats intended for post-production mastering with professional monitors and color-grading software.

Gamma – Originally developed to compensate for the non-linear response of the cathode ray tube, gamma correction functions or gamma curves are still employed by modern displays for SDR content. ITU-R recommendations Rec. 601, Rec. 709 and Rec. 2020 describe the reference OETFs for SD-TV, HD-TV and UHD-TV respectively, and are all based on the gamma curve.

Logarithmic – Several camera manufacturers have developed their own logarithmic OETFs in order to increase the captured dynamic range, including Sony's S-Log, Canon's Canon Log and ARRI's Log C.

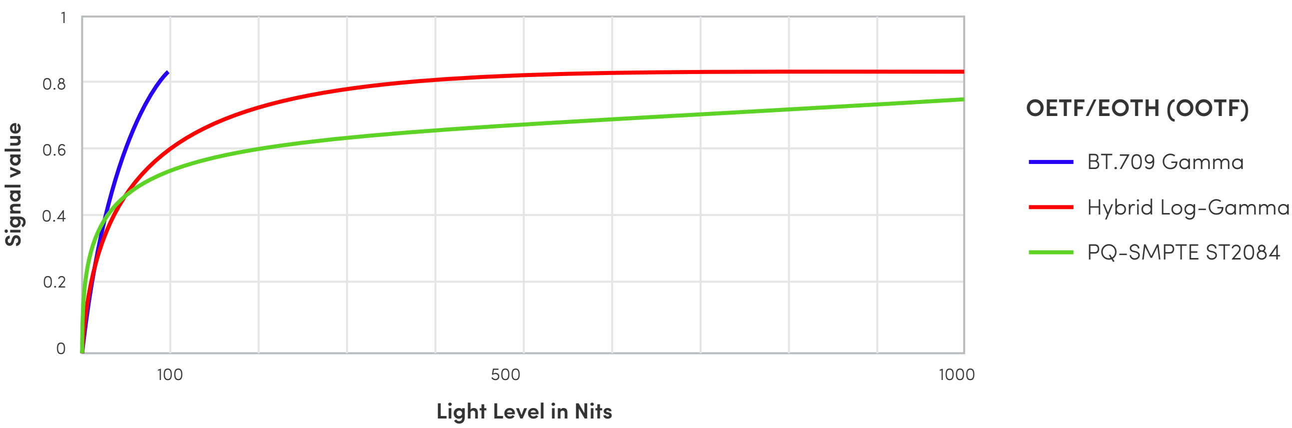


HDR Transfer Functions

ITU-R BT.2100 defines 2 types of EOTFs for displaying HDR content:

PQ – Perceptual Quantizer (also standardized in SMPTE ST 2084) was developed by Dolby to replace the gamma curve for HDR content and is utilized by multiple formats including Dolby Vision, HDR10 and HDR10+.

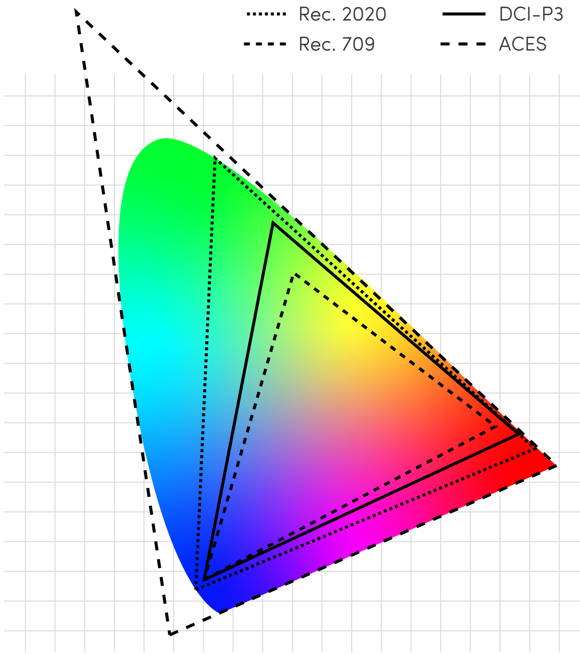
HLG – Hybrid-Log Gamma was jointly developed by the BBC and NHK for broadcast television HDR applications. As the name suggests, HLG uses a hybrid approach, combining the SDR gamma curve with a logarithmic function for the HDR range, allowing HLG content to be viewed as intended on both SDR and HDR capable monitors



Color Spaces

Color Spaces are defined collections of colors composed of a color model (system for representing individual colors) and a color gamut (the range of colors that can be represented). Most video display color spaces are based on RGB color models, where different levels of red, green, and blue light are combined to produce each hue.

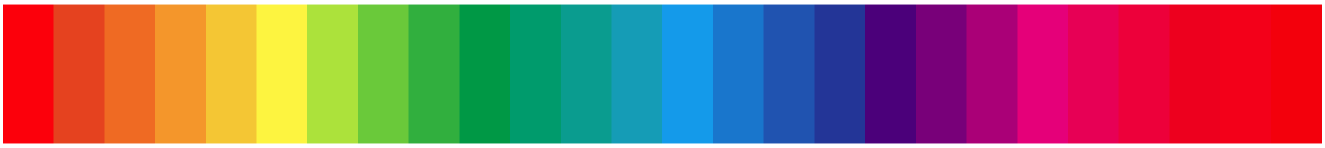
The table below shows commonly used RGB color spaces along with their purpose, year of standardization and the percentage of the human visible spectrum of colors they can represent.



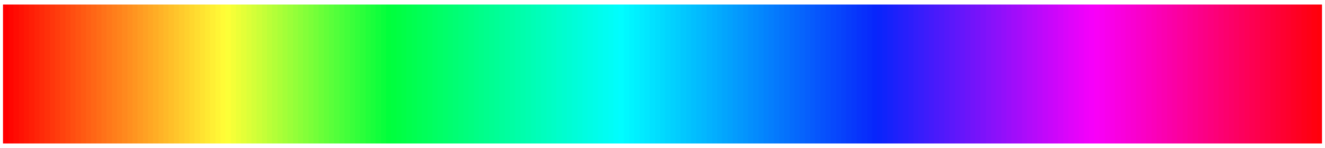
COLOR SPACE	YEAR	PURPOSE	% HUMAN VISUAL SPECTRUM
CIE 1931	1931	Developed in 1931 by International Commission on Illumination(CIE), the CIE 1931 chromaticity diagram maps out all human visible colors	100%
Rec. 709	1990	ITU recommendation for HDTV displays, used by most Standard Dynamic Range content	~ 33%
DCI-P3	2007	SMPTE standard further expanded color gamut for digital cinema projection	~ 42%
ACES	2007	Developed by Academy of Motion Picture Arts & Sciences for movie production, contains 100% of human visible spectrum, but no displays are capable of reproducing it	100%
Rec. 2020	2012	ITU recommendation for UHDTV displays, relevant for HDR formats	~ 57%

Pixel Depth

Pixel depth or bit depth refers to how many bits are used to represent the color of each individual pixel. HDR video is almost always created with a higher bit depth of 10 or 12-bits, compared to 8-bit SDR. 10-bit color allows for over 1 billion different hues compared to 16.7 million possibilities for 8-bit. Having so many more options allows for finer distinctions, smoother gradients, and more accurate representations of the scene being captured. 8-bit video can suffer from "banding" where there are noticeable borders and jumps in color in areas where there should be a smooth transition.



8-bit



10-bit

HDR Formats

Dolby Vision

The original High Dynamic Range format, Dolby Vision is a dynamic HDR format with accompanying metadata that is used to optimize contrast and light levels for every scene or even frame of a video in order to take full advantage of the display's capabilities and most accurately reproduce the content creator's intent.

- Developed by Dolby Laboratories, released in 2014
- Proprietary standard with licensing fees; Requires use of Dolby certified tools.
- EOTF: PQ
- Color gamut: Rec. 2020
- Pixel depth: 10 or 12-bit
- Max luminance: At least 1,000 nits; can also be mastered to a maximum of 4,000 nits. Dolby Vision standard technically supports up to 10,000 nits, but currently available consumer displays do not.
- Metadata: Dynamic, can adjust on a per-scene or frame basis.
- Compatibility: Dolby Vision compatibility profiles can be used to provide Dolby Vision metadata enhancements, while retaining backward compatibility with static HDR and SDR formats.

HLG

Hybrid Log-Gamma (HLG) was jointly developed by the BBC and NHK as a broadcast television HDR format that has a wider dynamic range while remaining backward-compatible with existing SDR transmission standards. As such, it can be displayed correctly on both SDR and HDR monitors. The Ultra HD Forum defined the term HLG10 to refer to an HLG system that is 10-bit, with wide gamut colorimetry.

- Release: 2015 by BBC and NHK
- Royalty-free open standard
- EOTF: HLG
- Color gamut: Rec. 2020
- Pixel depth: 10-bit
- Max luminance: 1,000 nits
- Metadata: None
- Compatibility: Backward compatible for viewing on SDR displays

HDR10

HDR10 is an open standard and the most widely used and supported HDR format. It has static metadata that is used to map the content to maximize the display's brightness and contrast capabilities, applied uniformly to the entire file. PQ10, a less used and discussed format, is essentially HDR10 without metadata.

- Release: 2015 by the Consumer Technology Association
- Royalty-free open standard
- EOTF: PQ
- Color gamut: Rec. 2020
- Pixel depth: 10 bit
- Max luminance of at least 1,000 nits; can also be mastered to maximum of 4,000 nits. Standard technically supports up to 10,000 nits, but currently available consumer displays do not
- Metadata: Static:
  - SMPTE ST 2086 (mastering display color volume)
  - MaxFALL (maximum frame-average light level, denotes the average light level of the brightest frame in the video)
  - MaxCLL (maximum content light level, denotes the light level of the brightest single pixel in the video)

HDR10+

Announced in 2017 by Samsung and Amazon Video, HDR10+ is an enhancement to HDR10 with dynamic metadata for scene by scene optimizations. HDR10+ metadata can co-exist with HDR10 static metadata, making HDR10+ content backward compatible with TVs that only support HDR10.

- Release: 2017 by Samsung and Amazon
- Royalty-free for content creators, annual license for display/device makers
- EOTF: PQ
- Pixel depth: 10, 12 or 16-bit
- Color gamut: Rec. 2020

- Max luminance of at least 1,000 nits; can also be mastered to maximum of 4,000 nits. Standard technically supports up to 10,000 nits, but currently available consumer displays do not
- Metadata: Dynamic - can be updated on per-scene or frame basis, co-exists with static HDR10 metadata
- Compatibility: Backward compatible for viewing on HDR10 displays

HDR Conversions

Methods

Dolby Vision Content Mapping – The Dolby Vision ecosystem is tightly controlled and only Dolby certified devices and software (like the Bitmovin VOD Encoder) can be used for creating Dolby Vision OTT outputs or converting Dolby Vision to static HDR and SDR formats.

Tone-mapping – Technique used in image and video processing to approximate the appearance of HDR within SDR formats by mapping one color space and dynamic range combination to another. You may see the term color volume used to describe that combination of color space and dynamic range, as well as up-mapping or down-mapping, which respectively refer to increasing or reducing color volume.

Gamut-mapping – Technique involving use of color adjustment algorithms applied to enable the original colors to 'fit' inside differently shaped color gamuts, authentically transferring images across a range of media (ex. converting an RGB image for printing on a CMYK printer).

LUTs - In the broadcast and film industries, Lookup tables or LUTs are used to map one color space to another, commonly used to calculate and preview how an image will be reproduced on another type of display. 3D-LUTs offer more color adjustment points and improved blending performance compared to traditional LUTs.

HDR10 ↔ HLG

HDR10 → HLG

Since HDR10 can contain peaks as high as 10,000 nits, BT.2390 suggests that HDR10 content be pre-normalized to a nominal peak luminance of 1000 nits prior to conversion. Once normalized to 1000 nits, it is possible to perform the required conversion to HLG with a 3D LUT or with tone-mapping.

HLG → HDR10

This conversion can be accomplished with a 3D LUT prior to encoding. HLG content is not mastered to a target peak luminance like HDR10, so a target needs to be established as part of the LUT conversion function. ITU-R Report BT.2390 recommends targeting a peak luminance of 1,000 nits, with the peak value of the HLG signal (100% HLG) mapping to the HDR10 value of 1,000 nits.

SDR ↔ HDR

HDR → SDR

HDR to SDR down-conversion can be accomplished through 3D LUT mechanisms or invertible down-mapping, in which HDR/WCG content is down-mapped in real time to SDR/BT.709 at or prior to the emission encoder.

SDR → HDR

Direct Mapping: SDR/BT.709 content is decoded and repackaged as PQ10 (HDR10 without metadata) or HLG10 containers, but while changing the system colorimetry, remapping does not change the color gamut or the dynamic range of the content; the content is simply mapped across to the equivalent color and brightness values.

Up-Mapping: SDR/BT.709 is decoded and then enhanced/modified to emulate PQ10/HLG10 and repackaged as above. Unmonitored up-mapping can lead to undesirable results depending on the technology so care should be taken when converting SDR to HDR. Conversion algorithms or equipment should be carefully evaluated prior to their use.

OTT Support

In general, there is no difference between HDR and SDR for OTT playback via DASH and HLS, as the difference resides within the delivered video codec bitstream and HDR support will depend on the capabilities of the playback device. There are however recommendations for the preparation and presentation of HDR content provided within the HLS and DASH specifications.

HLS

Apple provides a specification for HLS (HDR Metadata). Highlights include:

- 1.7. High Dynamic Range (HDR) HEVC video MUST be HDR10, HLG, or DolbyVision.
- 1.23. If multiple video streams are provided (H.264, HEVC, HDR), each stream SHOULD provide all anticipated bandwidths. Clients SHOULD NOT be required to switch codecs.
- 9.22. If you supply HDR content, you SHOULD provide both Dolby Vision and HDR10.

MPEG-DASH

The DASH Industry Forum provided guidelines for provisioning HDR content with MPEG-DASH. The related chapters are:

- 10.3. DASH-IF IOP HEVC HDR PQ10
- 10.4. DASH-IF IOP UHD Dual-Stream (Dolby Vision)
- 11.3.3. DASH-IF VP9-HDR

Device Support

Televisions

Note: The tables below only note confirmed support. Omission does not necessarily indicate lack of support.

BRAND	DOLBY VISION	HLG	HDR10	HDR10+	DETAILS
Sony	✔	✔	✔	✘	2017+ HD/UHD
LG	✔	✔	✔	✘	2017+ HD/UHD
Vizio	✔	✔	✔	✘	Some 2016, All 2017+ UHD
Samsung	✘	✔	✔	✔	Mid-2016 and above VIZIO SmartCast UHD HDR TVs, (P, M, E series) support HLG
Roku TV (Hisense, TCL)	✔	✘	✔	✘	Only the latest 2018/2019 TCL 4K TVs support HDR

Set-Top Boxes and Consoles

DEVICE	DOLBY VISION	HLG	HDR10	HDR10+
Fire TV Edition 4K	✔	✔	✔	✔
Fire TV Cube	✔	✘	✔	✘
FireStick 4K	✔	✘	✔	✘
Android TV STB	✘	✔	✔	✘
AppleTV 4K	✔	✔	✔	✘
Roku 4K	✔	✔	✔	✘
Chromecast Ultra	✔	✘	✔	✘
Nvidia TV Shield (Pro)	✔	✘	✔	✘
Xbox One S/X	✔	✘	✔	✘
PS4 Pro / PSD5	✘	✘	✔	✘

Bitmovin HDR Support

Bitmovin was the first Dolby Pro Partner to integrate the Dolby Encoding Engine to support both Dolby Vision and Dolby Atmos encoding in the cloud and can convert and transcode OTT compatible Dolby Vision, HDR10, and SDR outputs all in a single workflow. We've also provided dynamic range format presets for easy conversion between HLG, HDR10 and SDR content.

Conversions		Outputs			
		DOLBY VISION	HLG	HDR10	SDR
Inputs	**Dolby Vision***	✓	✗	✓	✓
	HLG	✗	✓	✓	✓
	HDR10	✗	✓	✓	✓
	SDR	✗	✓	✓	✓

* Dolby Vision mastered MXF files in ProRes or JPEG2K format

Bitmovin resources

Code Samples

https://github.com/bitmovin/bitmovin-api-sdk-examples#hdr-sdr-conversion

Blogs

Streamlining HDR Workflows with Bitmovin
Upconverting SDR to HDR
Why HDR Should Be Part of Your Video Workflow
Cloud-agnostic encoding solutions for Dolby Vision

Webinars

5 Myths About Dolby Vision and HDR Debunked
How to Enable and Optimize HDR Encoding
Deploying HDR for Streaming Services: Challenges and Opportunities

Links

Video: How does high dynamic range video work?
Glossary: Quick Reference HDR Glossary by Technicolor
Reference: Understanding Color Spaces
Guide: Dolby Vision for OTT Services
Whitepaper: Understanding the HDR10 Ecosystem

Specification docs:
- ITU Recommendations
- Apple HLS Authoring Specification
- High Dynamic Range Metadata for Apple Devices
- DASH-IF Interoperability Guidelines
- Dolby Vision Streams - HLS Format
- Dolby Vision Streams - MPEG-DASH Format

About Bitmovin

Bitmovin is an Emmy Award-winning leading provider of video software and cloud infrastructure for online media companies and enterprises globally. Bitmovin technology innovations focus on video encoding, playback, and analytics around user experiences. Innovations include the co-authoring of the MPEG-DASH streaming protocol and the first commercial HTML5 MPEG-DASH player, as well as massively parallel cloud-native API-driven encoding, featuring the first commercial AV1 next-generation codec. Another feature is industry-leading transcode speed reaching 100 times real-time.

The Bitmovin Player runs on the widest array of compelling consumer devices, ranging from mobile handheld devices to large screen televisions fed by dongle devices or with native smart TV capabilities - providing a rich feature set with consistent UIs and APIs.

Contact us at

US Headquarters
41 Drumm Street
San Francisco, CA 94111

European Headquarters
Schleppe Platz 7,
Klagenfurt, Austria 9020

Email us at
sales@bitmovin.com