KEYWORDS AMBIENT VIEWING ENVIRONMENT // ENVIRONMENT-DRIVEN RENDERING // VIEWING EXPERIENCE // HDR

Preserve HDR Creative Intent with Ambient Viewing Environment Metadata

By Qing Song, Taoran Lu, Giggs Xu, Chris Chang, and Xin Wang

Abstract
Consumer mobile captures and professional movies are graded in different nominal ambient viewing environments, but many of them lack the corresponding Ambient Viewing Environment (AMVE) metadata. This causes problems in video rendering and degrades visual quality. This paper presents a recommended practice for handling AMVE metadata to benefit the industry and address practical issues. A theoretical model and a closed-form solution are provided, along with application examples and usage scenarios. We encourage the industry to explore this area further, align videos from different sources, and optimize the viewing experience.

It is well known that the viewing environment, including ambient, viewing distance, and viewing angle, significantly influences the video viewing experience. HDR videos are particularly affected by ambient viewing conditions because of the abundance of information in both shadows and highlights. To optimize visual quality for users, environment-driven video rendering has been explored. Numerous methods have been proposed to compensate for detail or contrast loss caused by ambient conditions. Most methods increase the overall brightness of the video for a bright ambient viewing environment, implicitly assuming that videos are originally mastered in a dim ambient viewing environment. This is true for professional studio-graded movies and broadcasting programs, and it complies with ITU-R BT. 2100,[1] which recommends 5 nits surround (16 lux ambient). However, when we investigated

1545-0279/26©2026SMPTE

Consumer mobile HDR captures are often graded distinctly from professional content mastered in dim studios to provide satisfactory looks for various everyday viewing conditions. By promoting awareness and utilization of Ambient Viewing Environment metadata, the industry can better align videos from different sources, reduce manual content tuning, and optimize viewing experiences while addressing rendering inconsistencies.

consumer HDR mobile captures, it was found that this is not always the case.

Table 1 elucidates the fundamental differences between the anticipated and prevalent practices of the HDR video production workflow in the present era. Consumer mobile HDR videos and professional HDR videos differ in aspects of video capture, grading, and consumption environment. Before Apple developed consumer HDR video capture on mobile phones, HDR video production was exclusively feasible with high-end, professional devices (camcorders, DSLRs) and by individuals with expertise and a comprehensive understanding of the domain.

Grading is performed in the dim grading suite, and the nominal environment in which users are expected to consume the content is a cinema or a living room. In contrast, consumer mobile HDR captures aim to provide a satisfactory look suitable for various viewing conditions in consumers' everyday life. To achieve that, the nominal ambient viewing illumination may differ from the 16-lux (5 cd/m^2) dim studio environment.

The disparities between the two categories are reflected in the distribution of absolute signal levels. **Figure 1** shows the statistics for professional movies and mobile-captured videos. Six mobile phones capable of HDR video recording were used to capture three typical scenes: night, sunny outdoor, and well-lit indoor. Similar scenes in professional 1000-nit mastered HDR10 movies were selected for comparison. The average frame brightness shows that most mobile captures are graded at distinctly higher brightness levels than professional movies. That implies a higher nominal ambient environment, but what is the value?

iPhone captured HDR videos[2] gives us a hint. The iPhone captures Ambient Viewing Environment (AMVE) metadata[3] indicating that ambient mastering is assumed to be 314 lux (100 cd/m^2). AMVE metadata was introduced in the AVC,[4] HEVC,[5] and V-SEI[6] specifications as a means to preserve creative intent on displays with color volumes and ambient viewing illumination that may differ from those of the described mastering display and its corresponding ambient viewing environment. The metadata was also later added to file format specifications[7,8] to make the signaling codec agnostic and to avoid parsing low-level codec signaling to find and utilize such information.

314-lux implies a bright office environment. When such a video is displayed in dim ambient conditions or when it is to be viewed or composited with HDR content mastered in accordance with the BT.2100 standard, some adaptation should be applied to adjust the video brightness and contrast **(Fig. 2)**. Otherwise, the video may appear uncomfortably bright to users. This is contrary to the common practice of addressing the viewing ambient, which almost always adapts from dim

TABLE 1. Differences Between Consumer HDR Video and Professional HDR Video.

	Consumer HDR Mobile Capture	Broadcast / HDTV Production	Movie / Hollywood
Requirement for the skill of person doing the capture	No specialized skills are required, provided that the individual possesses the ability to operate a smartphone effectively.	Expertise required	Expertise required
Environment for capture	A wide range of lighting conditions, including complete darkness and full sunlight, without the use of assistant lighting or controlled lighting in general.	A diverse range of lighting conditions, with potential assistance from additional lighting sources, or in a controlled studio environment.	A diverse range of lighting conditions, with potential assistance from additional lighting sources, or in a controlled studio environment.
Expectation for having a reasonable "look" or grading in capture environment?	Yes, expectation is a grading that closely approximates reality.	No in general, usually in RAW/LOG format, which requires post-production, except for "live broadcasting" case.	No, usually in RAW/LOG format and expect to go through professional post-production.
Grading and post-production	Post-production is unnecessary. The grading is determined and optimized on the fly for immediate file generation and playback within the capture environment.	The grading is done in the dim grading suite under reference ambient conditions.	The grading is done in the dim grading suite under reference ambient conditions.
Typical "nominal ambient environment" we expect user to consume the content.	Home, office, restaurants, public transportation, outdoor shaded area, etc.	Living room	Cinema, living room
Assumed ambient viewing illumination	314 lux (100 cd/m^2), D65 white point, as indicated by iPhone[2]	16 lux (5 cd/m^2)	16 lux (5 cd/m^2)

1545-0279/26©2026SMPTE

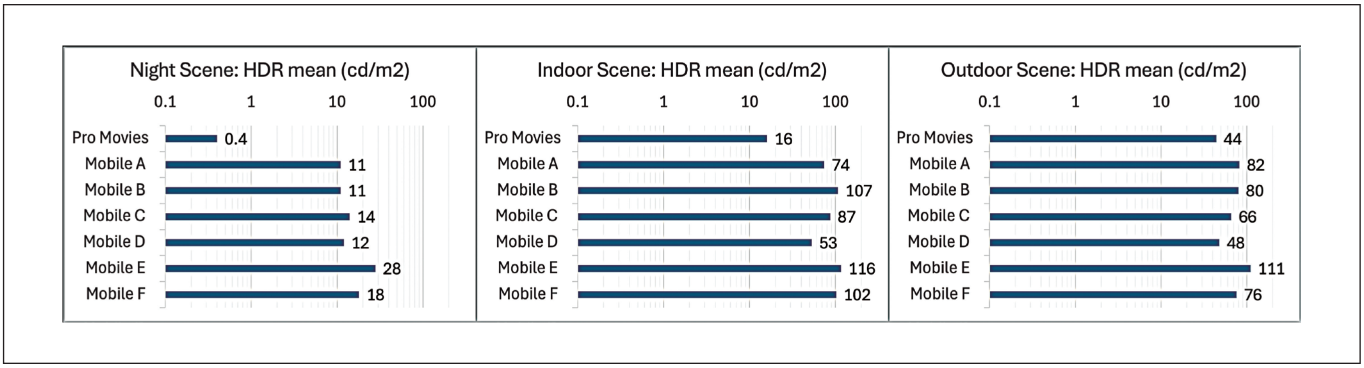


FIGURE 1. Mean luminance of professional HDR movies and consumer mobile captured HDR video.

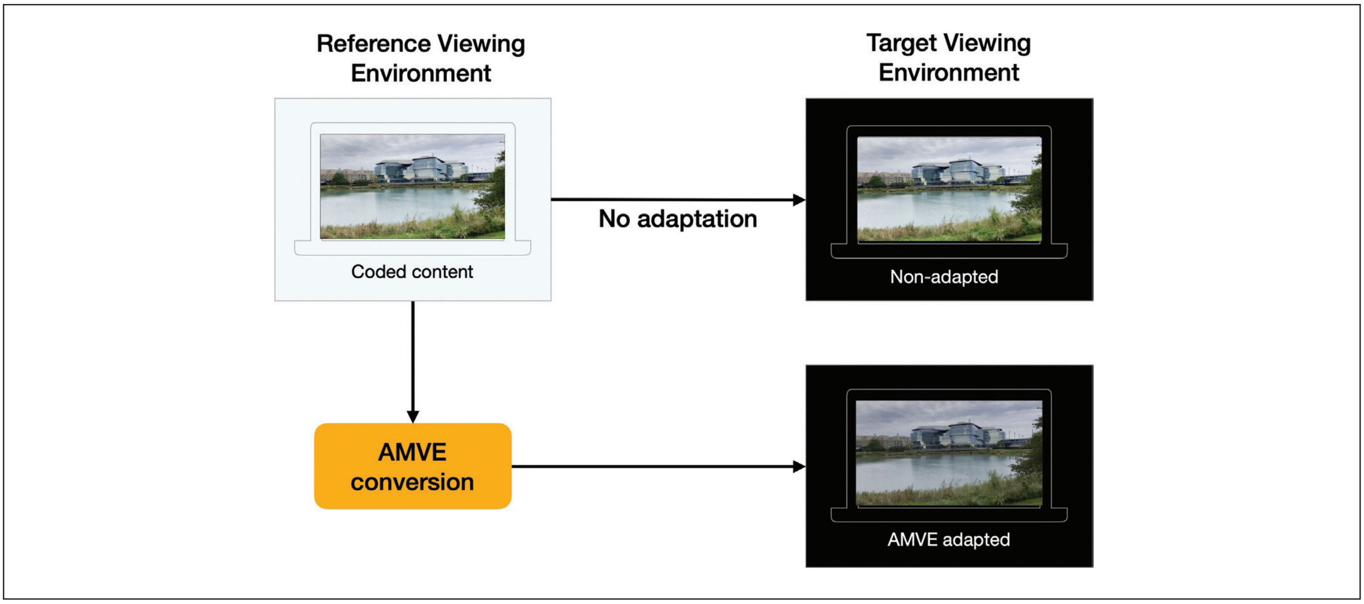


FIGURE 2. Example of content adaptation with AMVE metadata. The content is mastered under a single nominal ambient illumination (e.g., 314 lux). When displayed under a target viewing ambient illumination (e.g., 16 lux), which differs from the reference viewing illumination, AMVE adaptation can be performed to preserve the creative intent better. Without adaptation, content mastered in a bright ambient environment may appear too bright when displayed under dim viewing conditions.

ambient to bright ambient, as explained at the beginning of the paper.

The current situation in the industry, however, is that the AMVE metadata is not widely utilized or respected, resulting in rendering inconsistencies and suboptimal experiences. All the mobile brands in our investigation, with the exception of iPhone, do not tag any AMVE metadata, despite the grading being similarly bright to iPhone's. Since playback cannot distinguish the mastering environment without the AMVE metadata, mobile captures and professional movies are displayed using the same mechanism. As a result, users experience differing brightness across the two video categories. It is very common for videos from various sources to be fed to users in a sequence or even simultaneously. Finding it hard to adapt eyes to different brightness levels in a short period, users often complain about eye fatigue.

Even if the AMVE metadata is tagged to the video, it is frequently lost in the existing video distribution pipelines. To our best knowledge, Instagram preserves the AMVE metadata upon video uploading and distribution. Most other social media do not, unfortunately. Some might not be aware of the existence of the AMVE, while others might lack knowledge of how to handle it, leading them to discard it.

The primary intent of this paper is to raise awareness of and promote the use of the AMVE metadata. Once the AMVE metadata is preserved and carried, the next question is, "What can I or should I do with that?" With this missing piece filled in, we hope the metadata can be better adopted and respected in the entire industry. Note that, although our study primarily focuses on HDR videos, a similar practice can be applied to SDR videos. The rest of the paper is organized as follows: the next section describes a model that can adapt videos from the nominal ambient viewing environment to a target ambient viewing environment; then, examples are provided to illustrate the application of the conversion function; the use cases are discussed later; the last section concludes the paper.

The AMVE Conversion Model

Generally, decoder/receiver implementors can design an algorithm tailored to the specific receiving system, as long as it correctly handles different mastered ambient videos. For the convenience of industry practitioners, we provide

1545-0279/26©2026SMPTE

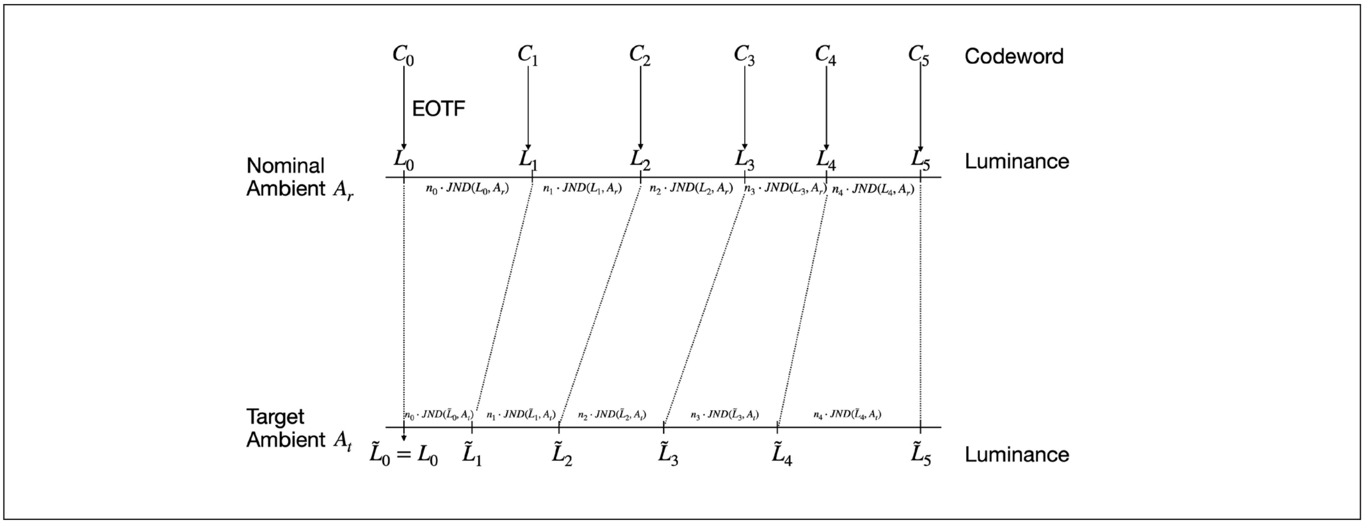
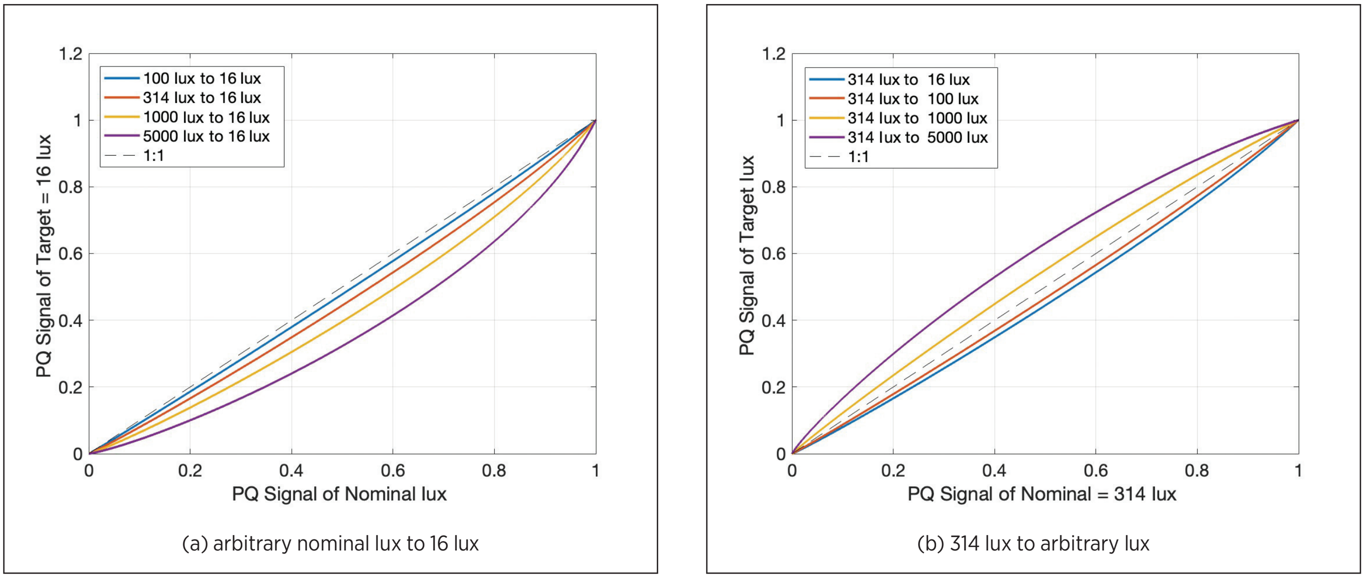


FIGURE 3. Construction of the AMVE conversion function. C_i is the codeword, L_i is the corresponding display luminance, which can be computed using EOTF, $JND(L_i, A_r)$ is the JND of the stimulating luminance L_i under the adapted luminance A_r, n_i is the number of JND between L_i and L_{i+1}, $\tilde{L}_{i+1}$ is the luminance that can preserve the same number of JND away from $\tilde{L}_i$ under the adapted luminance A_t. $L_i \rightarrow \tilde{L}_i$ constructs the mapping function.



(a) arbitrary nominal lux to 16 lux

(b) 314 lux to arbitrary lux

FIGURE 4. Illustration of AMVE conversion functions from a nominal AMVE to a target AMVE, derived by the theoretical model.

a model for converting videos from the nominal viewing environment to the target viewing environment. As dynamic metadata that indicate pixel distributions, such as the content color volume SEI,[6] may not always be available in the video bitstream, we aim for a general, content-independent model that can be easily applied to any video. The conversion model is designed to consider the following points:

- A static function for setting a quality baseline, for the ease of implementation and adoption.
- Support for the full luminance range (0 ~ 10000 cd/m^2) defined by the PQ transfer function, which is the most extensive range supported in the HDR field.
- Support for round-trip conversion with minimal round-trip error.
- The capability to handle an arbitrary pair of AMVE values from one input value to an output value.

Model Construction

To understand why grading is impacted by the ambient viewing environment, the crispening effect[9] is examined. The human visual system's sensitivity is maximized at the luminance level to which it is adapted. That is why a car headlight looks fairly dim during the day: our eyes are adapted to the bright environment. However, at night, the headlight seems quite bright, even blinding. We probably cannot see the details in the headlights because our eyes are adapted to the dark environment. Similarly, one would grade the midtone of a video that looks comfortable in their grading ambient and distribute pixels to shadows and highlights to reveal good details in that viewing environment.

A conversion model is proposed[10] based on Barten's Contrast Sensitivity Function (CSF). The CSF defines the sensitivity of the human eyes, and the reciprocal 1/CSF defines the minimum detectable modulation. Barten's CSF model is a

1545-0279/26©2026SMPTE

function of frequency, stimulating luminance, and adapted luminance.[11,12] Just-Noticeable Difference (JND) can be derived from the minimum detectable modulation. Since digital videos are represented by a limited number of codewords, each codeword corresponds to a display luminance, and the number of JND between two neighboring codewords is computed under the nominal ambient viewing environment. Then, under the target ambient viewing environment, maintain the same number of JND and derive the next display luminance level. Hence, we can obtain the luminance pairs in the two environments and construct a mapping function. **Figure 3** illustrates the concept.

The luminance for the target ambient can be derived in the following step-by-step fashion. If it is desired to map the luminance of the last codeword to the peak brightness of the display, multiply a constant α to $n_i \cdot JND(\tilde{L}_i, A_t)$. Refer to the reference work[10] to find the optimal α.

$$n_i = \frac{L_{i+1} - L_i}{JND(L_i, A_r)}$$

$$\tilde{L}_{i+1} = \tilde{L}_i + \alpha \cdot n_i \cdot JND(\tilde{L}_i, A_t)$$

Encoding luminance pairs using PQ[13] yields the PQ conversion function. **Figure 4** plots the conversion functions from arbitrary nominal ambient to 16 lux and the conversion functions from 314 lux to arbitrary target ambient. When a high nominal ambient is converted to a low target ambient, the conversion function reduces mid-tone brightness, resulting in a dimmer video appearance. Conversely, when the nominal ambient exceeds the target, the mid-tone is boosted, and the contrast of the shadows is enhanced, thereby improving the video's perceivability in bright ambient conditions.

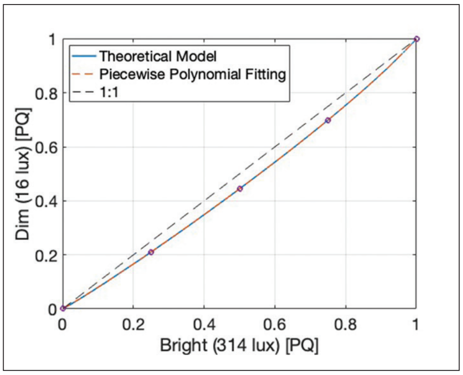
Why is screen reflection not considered here? One reason is that our target is to develop a display-agnostic model. Any display-related compensation should be done at the display side. The other reason is that screens nowadays, especially high-end devices, typically have low reflection rates. It is reasonable to assume that the screen reflection of the mastering display used for grading can be neglected. Ambient white point is not considered in this paper, but can be explored further in the future.

Closed-Form Math Representation

The model described above is constructed in a step-by-step fashion. Once constructed, it can be implemented as a look-up table (LUT). Alternatively, the function can be approximated by a piecewise polynomial, which could be easier to distribute and implement.

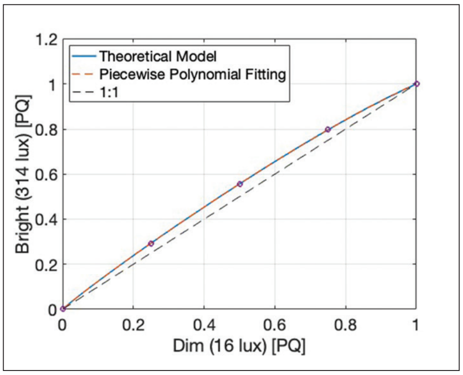
$$f(x) = c_{0,k} + c_{1,k} \cdot x + c_{2,k} \cdot x^2 + c_{3,k} \cdot x^3 \text{ , if } p_k \le x < p_{k+1}$$

- $c_{i,k}$ is the coefficient of the i^{th} order of the k^{th} piece
- p_k is the starting point of the k^{th} piece



$c_{i,k}$	$0 \le x < 0.25$	$0.25 \le x < 0.5$	$0.5 \le x < 0.75$	$0.75 \le x \le 1$
$c_{0,k}$	0	-0.0006	-0.0462	-0.5320
$c_{1,k}$	0.7378	0.7914	1.0542	2.8973
$c_{2,k}$	0.6291	0.2297	-0.2746	-2.5983
$c_{3,k}$	-0.8424	-0.0629	0.2593	1.2326

(a) 314 lux to 16 lux



$c_{i,k}$	$0 \le x < 0.25$	$0.25 \le x < 0.5$	$0.5 \le x < 0.75$	$0.75 \le x \le 1$
$c_{0,k}$	0	-0.0026	0.0164	0.1302
$c_{1,k}$	1.3208	1.2738	1.1236	0.6611
$c_{2,k}$	-0.9166	-0.4137	-0.0413	0.5848
$c_{3,k}$	1.2938	0.2028	-0.0935	-0.3760

(b) 16 lux to 314 lux

FIGURE 5. Illustration of the piecewise polynomial fitting and its coefficients for the conversion between 314 lux and 16 lux in the PQ domain.

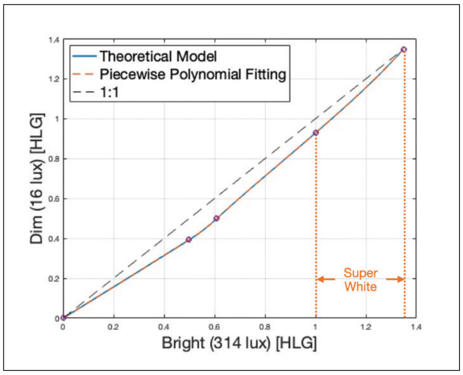
Let us define the following:

Static conversion: $y = f_{nom->tar}(x)$, where both x and y are floating-point PQ values in the range of 0 to 1, inclusive.

Round-trip error: $err = x - f_{A->B}(f_{B->A}(x))$.

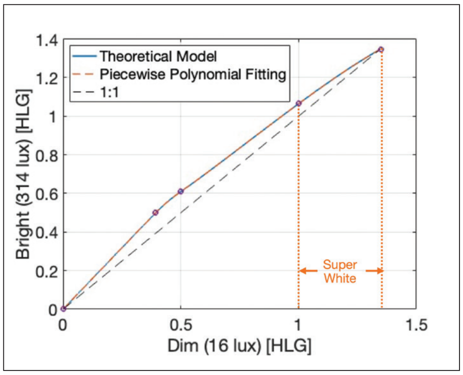
The round-trip error between 314 lux and 16 lux can be calculated as: $err = x - f_{16->314}(f_{314->16}(x))$.

- Max round-trip error in floating point: 6.21e-4
- Max round-trip error in 10-bit: 9.78e-4 => at most one codeword in 10-bit



$c_{i,k}$	$0 \le x < 0.5$	$0.5 \le x < 0.6096$	$0.6096 \le x < 1$	$1 \le x \le 1.3468$
$c_{0,k}$	0	0.2164	-0.4582	-0.9737
$c_{1,k}$	0.7827	0.2192	2.1581	3.5081
$c_{2,k}$	-0.0002	-0.3438	-1.2597	-2.4129
$c_{3,k}$	0.0155	1.2257	0.4885	0.8072

(a) 314 lux to 16 lux



$c_{i,k}$	$0 \le x < 0.3938$	$0.3938 \le x < 0.5$	$0.5 \le x < 1$	$1 \le x \le 1.3468$
$c_{0,k}$	0	-0.7402	0.2738	0.3782
$c_{1,k}$	1.2727	5.8726	0.4126	0.1610
$c_{2,k}$	0.0198	-9.0162	0.6549	0.8448
$c_{3,k}$	-0.0676	5.3375	-0.2768	-0.3195

(b) 16 lux to 314 lux

FIGURE 6. Closed-form math model of round-trip conversion between AMVE 314 lux and 16 lux in the HLG domain.

By default, the AMVE conversion function is defined in the PQ domain. It can be easily represented in the linear light domain by applying PQ-to-linear light conversion, as in the SMPTE 2084 standard. If one wants to use the AMVE conversion directly in the HLG code domain, the equivalent piecewise polynomial model in the HLG domain is also provided. Note that the PQ to HLG conversion follows the ITU-R BT.2100[1] standard, leading to the PQ value of 1.0 (10,000 cd/m^2) being converted to 1.3468 in the HLG domain, where 1.0 corresponds to the peak luminance of 1000 cd/m^2 defined in BT.2100 and the values in the range [1.0, 1.3468] corresponds to HLG "super-white" values between 1000 cd/m^2 and 10,000 cd/m^2.

- Max round-trip error in floating point: 5.29e-4
- Max round-trip error in 10-bit: 9.78e-4 => at most one codeword in 10-bit

Evaluation

To evaluate the conversion model, we invited a professional, award-winning colorist to grade 2819 scenes captured by an iPhone. The colorist graded in the 16-lux studio environment, while the iPhone carries the 314-lux AMVE. The conversion function is computed from each pair of the two gradings. Then the average of the conversion functions is calculated. **Figure 7** shows a comparison of the conversion functions constructed from manual grading and the proposed model. The two curves exhibit similarities, particularly in the mid-to-high region, suggesting that the proposed AMVE conversion aligns with human perception to some extent. The lower part has a slight discrepancy. Upon consultation with the colorist, we found that he sometimes did local tuning to enhance back-lit objects, resulting in boosted shadows. Apart from this, the proposed AMVE conversion model aligns well with his manual grading. Note that iPhone captures have a maximum pixel brightness of 1000 nits, and the colorist graded on a 1000-nit display; hence, the function derived from the manual grading spans only from 0 to 1000 nits (represented as 0.75 in PQ).

One could question how the proposed approach is different from the gamma adjustment in BT.2390[14] for HLG adaptation to bright ambient. We applied the gamma adjustment to ambient luminance of 16 lux (5 cd/m^2) and 314 lux (100 cd/m^2), respectively, and computed the display luminance corresponding to each codeword. This forms mapping pairs between the two ambiance. The display luminance is then converted to PQ for illustration. **Figure 8** shows the conversion curve of the BT.2390 approach and the proposed approach. The two curves are similar in the lower part. The BT.2390 curve then grows more rapidly and reaches the 1:1 curve at 1000-nit (PQ 0.7518), as designed. This causes the HLG-based curve to deviate from the manually graded curve and produce brighter output at mid-tone and bright regions.

Example Application of Conversion Functions

Several example usages of the AMVE conversion block in an HDR conversion workflow are illustrated in this section. These examples are not exclusive, and implementors can decide where and how to use the AMVE conversion block to fit their own needs. More methods of application in other color spaces can be found in ITU-R BT. 2446.[15]

Example #1:

Apply the AMVE conversion on PQ or HLG encoded R', G', and B' components individually. This is the simplest method for use in a non-linear (PQ/HLG) domain directly and without introducing any out-of-gamut colors. Users need to configure the AMVE conversion model (PQ or HLG) appropriately to match the real signal **(Fig. 9).**

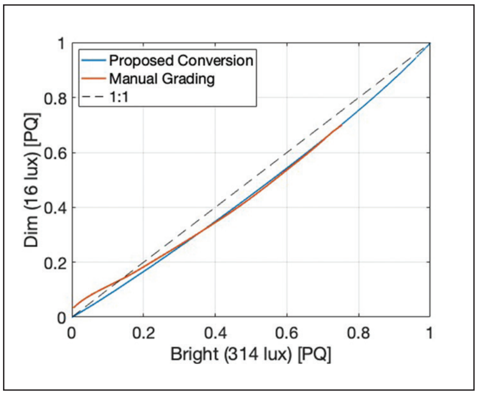


FIGURE 7. Comparing the proposed AMVE conversion function and the conversion function built from the manual grading.

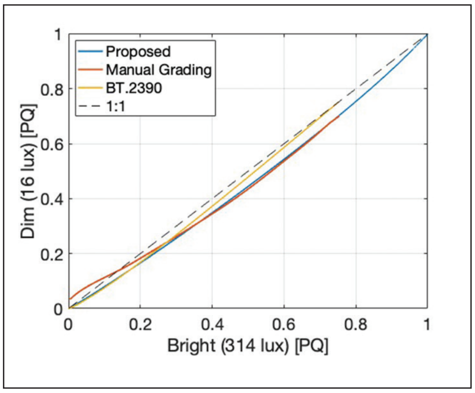


FIGURE 8. Comparing the proposed AMVE conversion function and the conversion function built from BT.2390.

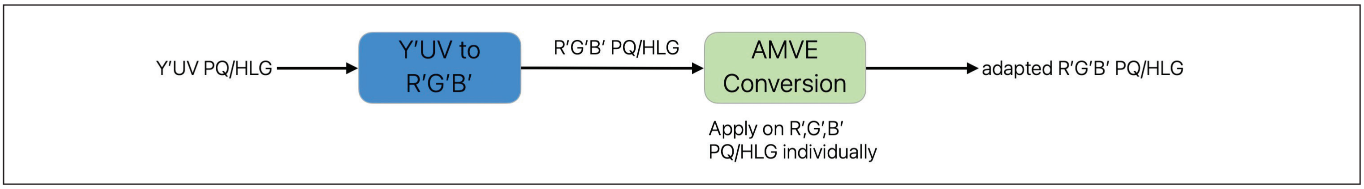


FIGURE 9. Example #1 of the application of the AMVE conversion.

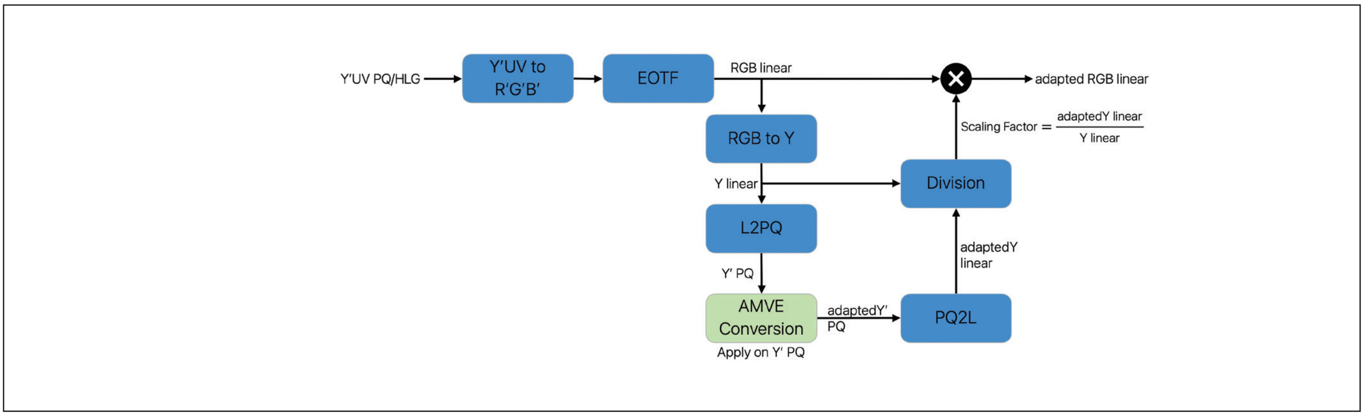


FIGURE 10. Example #2 of the application of the AMVE conversion.

Example #2:
Perform the AMVE conversion on constant-luminance Y and apply the same scaling factor to all three R, G, and B components in the linear light domain. The AMVE conversion model in the PQ domain should be used **(Fig. 10).**

Example #3:
Apply the AMVE conversion on non-constant-luminance Y' and apply the same scaling factor to all three R, G, and B components in the linear light domain. Users need to configure the AMVE conversion model (PQ or HLG) appropriately to match the real signal **(Fig. 11).**

Example #4:
Apply the AMVE conversion directly on the luma or luminance component of a color differencing signal space (e.g. Y'UV, Yu'v', ICtCp, IPT). Users need to configure the AMVE conversion model (PQ or HLG) appropriately to match the real signal. Note that adapting in such a space may generate out-of-gamut colors and may require additional gamut mapping after the AMVE conversion. Such mapping is outside the scope of this document **(Fig. 12).**

The output of the example applications is shown in **Fig. 13**. All the outputs have a similar brightness to the manually graded look. We also compute the average PQ luminance of the four converted outputs for all videos in our dataset and compare it with that of the manually graded look. The average absolute difference between the four converted outputs and the manual grading is 0.0196, 0.0195, 0.0196, and 0.0199, respectively.

Usage Scenarios

This section identifies the usage scenarios for the AMVE conversion. These are not exclusive cases.

1. Playback: if the system can detect the user's viewing environment, the AMVE conversion can be applied to convert the video from its indicated nominal ambient

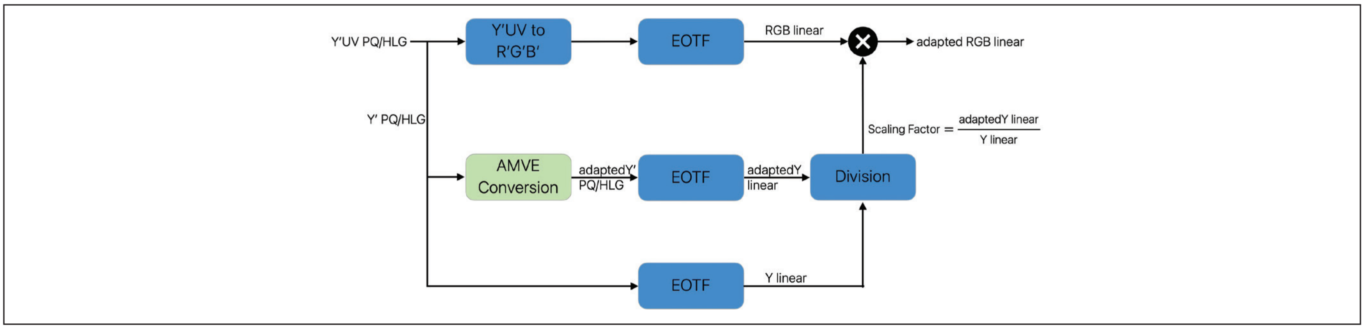


FIGURE 11. Example #3 of the application of the AMVE conversion.

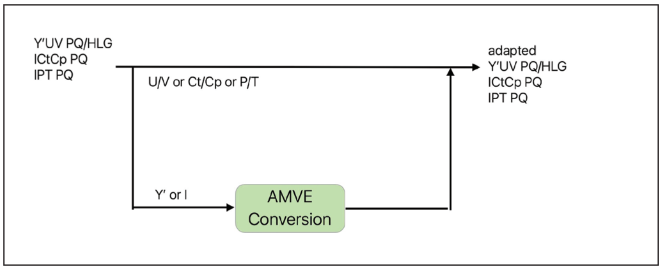


FIGURE 12. Example #4 of the application of the AMVE conversion.

viewing environment to the actual ambient that the user is experiencing.

2. Video editing: when a user attempts to composite videos with different AMVE metadata (such as concatenating multiple videos in a "reel," or placing two videos side by side in the same frame), they can convert all the videos to one target ambient viewing environment using the proposed conversion before editing, and export the final video with the AMVE metadata indicating the target ambient. This can help reduce the manual tuning of the content. The AMVE metadata is expected to be persistent throughout the entire bitstream and cannot change on a per-frame basis. It is essential to maintain consistent grading across all frames under the same ambient viewing environment to prevent significant variations in video brightness.
3. Video creation: Some mobile HDR cameras may need to determine the appropriate AMVE value for their captures. The captured videos by different mobile brands in Figure 1 exhibit different mean brightness. iPhone has AMVE of 314 lux. Other brands might use values that best suit their grading. They could experiment with various nominal ambient values and apply the conversion model to convert their captures to a 16-lux target ambient viewing environment. Then they can compare the outputs' brightness with studio-graded movies to determine the nominal ambient that yields the most similar brightness. This value can be used as the AMVE metadata for their original captures.

Conclusion

In this paper, we identified the problem of the absence of the AMVE metadata in HDR video production and distribution, as well as the negative impact on playback. A conversion model is provided to transfer a video from the nominal ambient viewing environment to a target ambient. The source code will be publicly available for generating a conversion transform between an input ambient lux and an output ambient lux. Example applications and usage scenarios are elucidated. While the industry is not required to follow the proposed model or the example applications exactly, the recommended practice for handling AMVE is provided to benefit the entire industry and to address practical issues in HDR video rendering.

To best preserve the creative intent and improve the viewing experience, it is recommended as a best practice to:

1. Carry and respect any AMVE metadata present in a bitstream whenever possible through the HDR video distribution pipeline to ensure proper interpretation of a video signal by receivers/displays. Such information should be treated similarly to how the mastering display color volume information and the content color volume information metadata that may also be associated with each bitstream. It should be preserved and acknowledged during distribution.
2. Assume 5 cd/m^2 (16 lux) to be the ambient viewing illumination when AMVE metadata is absent. This provides backward compatibility with most broadcast and studio HDR content following BT.2100 practice.
3. Adapt an HDR signal graded to the ambient illumination indicated in the AMVE metadata to a target viewing ambient illumination using the static AMVE conversion model provided in this paper, if the implementor finds it fits their needs.

We encourage the industry to explore this area further, including:

1. Extending to a white point adaptation, which is the other component of the AMVE metadata.
2. Identifying an appropriate AMVE value for each mobile HDR camera that most accurately reflects their grading intent.

(a) Source with 314-lux AMVE. Average luminance PQ = 0.5213 (114.02 nits).

(b) Manually graded in 16 lux by colorist. Average luminance PQ = 0.4406 (50.17 nits).

(c) Converted to 16-lux by #1. Average luminance PQ = 0.4625 (63.02 nits).

(d) Converted to 16-lux by #2. Average luminance PQ = 0.4626 (63.08 nits).

(e) Converted to 16-lux by #3. Average luminance PQ = 0.4626 (63.08 nits).

(f) Converted to 16-lux by #4. Average luminance PQ = 0.4626 (63.08 nits).

FIGURE 13. Conversion outputs (one scene example).

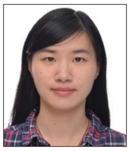
References

1. International Telecommunication Union-Radiocommunication (ITU-R), Recommendation BT.2100-3: Image parameter values for high dynamic range television for use in production and international programme exchange, Feb. 2025.
2. Apple Inc., "HDR video metadata," *Apple Technotes*, TN3145. Accessed: Jun. 25, 2025. [Online]. Available: https://developer.apple.com/documentation/technotes/tn3145-hdr-video-metadata#Incorporating-HDR-Video-in-your-apps
3. G. J. Sullivan, "Ambient viewing environment SEI message," JCT-VC Doc. JCT-VC-U0112, Jun. 2015.
4. International Telecommunication Union-Telecommunication (ITU-T), H.264 | ISO/IEC 14496-10: Information technology—Coding of audio-visual objects—Part 10: Advanced Video Coding, Jul. 2025.
5. International Telecommunication Union-Telecommunication (ITU-T), H.265 | ISO/IEC 23008-2: Information technology—High efficiency coding and media delivery in heterogeneous environments—Part 2: High Efficiency Video Coding, Mar. 2025.
6. International Telecommunication Union-Telecommunication (ITU-T), H.274 | ISO//IEC 23002-7: Information technology—MPEG video technologies—Part 7: Versatile supplemental enhancement information messages for coded video bitstreams, Oct. 2024.
7. International Organization for Standardization/International Electrotechnical Commission (ISO/IEC), ISO/IEC 14496-12: Information technology—Coding of audio-visual objects—Part 12: ISO base media file format, Jan. 2022.
8. International Organization for Standardization/International Electrotechnical Commission (ISO/IEC), ISO/IEC 23008-12: Information technology—High efficiency coding and media delivery in heterogeneous environments—Part 12: Image File Format, Jul. 2025.
9. P. Whittle, "Brightness, discriminability, and the crispening effect," *Vision Research*, 32 (8): 1493-1507, Aug. 1992.
10. Q. Song and P. C. Cosman, "Luminance enhancement and detail preservation of images and videos adapted to ambient illumination," *IEEE Trans. on Image Process.*, 27 (10): 4901-4915, Oct. 2018.

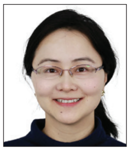
1545-0279/26©2026SMPTE

11. P. G. J. Barten, “Contrast sensitivity of the human eye and its effects on image quality,” *Proc. SPIE*, vol. PM72, p. 232, Dec. 1999.
12. P. G. J. Barten, “Formula for the contrast sensitivity of the human eye,” *Proc. SPIE*, vol. 5294, pp. 231–238, Dec. 2003.
13. S. Miller, M. Nezamabadi, and S. Daly, “Perceptual signal coding for more efficient usage of bit codes,” presented at *the SMPTE Annu. Tech. Conf. & Exhib.*, Hollywood, CA, Oct. 2012.
14. International Telecommunication Union-Radiocommunication (ITU-R), Recommendation BT.2390: High dynamic range television for production and international programme exchange, 2018.
15. International Telecommunication Union-Radiocommunication (ITU-R), Recommendation BT.2446: Methods for conversion of high dynamic range content to standard dynamic range content and vice versa, 2019.

About the Authors

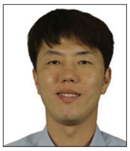


Qing Song is a video engineer in Apple, Inc., focusing on HDR video processing. She received a PhD degree from University of California, San Diego.



Taoran Lu is a system software engineer with Apple Inc., Her responsibilities include applied R&D in algorithms and software for video quality enhancement, tone mapping, display management and metadata handling.

Giggs Xu, a video processing engineer at Apple Inc., specializes in developing algorithms and frameworks for HDR video tone-mapping, transcoding, and conversion.



Chris Chang is a senior software engineer at Apple focusing on HDR video processing, tone mapping, and display architecture for improved video quality.

Xin Wang earned his PhD degree from Purdue University. He is currently engaged in HDR video development at Apple Corporation.

Presented at the SMPTE 2025 Media Technology Summit, Pasadena, CA, 16-19 Oct. 2025. Copyright © 2026 by SMPTE.

DOI: 10.5594/JMI.2026/RRHS3606
Date of publication: 1 April 2026