MANAGING HDR CONTENT PRODUCTION AND DISPLAY DEVICE CAPABILITIES

M. Zink; M. D. Smith

Warner Bros., USA; Wavelet Consulting LLC, USA

ABSTRACT

The introduction of next-generation video technologies, particularly high dynamic range (HDR), provides a compelling new palette for content creators. Similarly, consumer displays with brighter screens and darker black levels are very appealing to consumers. However, there are a number of challenges for content producers and device manufacturers that will be discussed in this paper. It will examine some commercial HDR content and consider the impact of creative choices on consumer electronics devices, particularly related to power management. This paper further explores opportunities to manage content and display device capabilities by analyzing content light levels during production, to help improve rendition of the content in display devices. This paper further discusses opportunities for display manufacturers to manage content that exceeds the device capabilities, to ensure a compelling user experience.

INTRODUCTION

Next-generation video technologies include a variety of features – 4K resolution, high frame rates, wide color gamut, and high dynamic range. One of the most novel features is high dynamic range (HDR), as it provides a user experience clearly distinguishable from existing HDTV programming. But it also poses a number of challenges – both for content producers and for consumer electronics manufacturers.

CONTENT PRODUCTION CHALLENGES

The process for mastering motion picture content has been fairly unchanged for a long time. Usually the process consists of creating a theatrical master, commonly in a theatre environment with a theatre projector, and then applying a so-called “trim pass” to create a home video master. The color grading of the home video master typically uses a professional mastering monitor to ensure the most accurate image reproduction. Since the legacy workflows, based on the HDTV video specification defined by BT.709 [1] and BT.1886 [2], have been well established for many decades, both professional mastering monitors and consumer devices meet those specifications, albeit with varying degrees of accuracy. In other words, the content is limited by those specifications in terms of color gamut, peak luminance and contrast ratios, even though modern display devices often exceed the capabilities of those specifications. While this legacy workflow allows a filmmaker to create the desired look on a professional mastering monitor and ensure a reasonable reproduction of this look in consumer homes, the introduction of next-generation video technologies, particularly HDR, requires changes to the established workflows.

Next-Generation Workflows

From a content producer perspective, HDR provides a new palette for creative filmmakers to enhance their storytelling and render their creative vision. Since these features are still very new, the workflows to produce such content are not yet fully established. In fact, many aspects of the content mastering workflow, including the availability of professional mastering monitors, remain under development. Additionally, one other challenge becomes very obvious.

The next-generation home video specifications, including Ultra HD Blu-ray™, based on BT.2020 [3] and SMPTE ST2084 [4], permit a very wide range of color and luminance to be represented in the video signal. Building an affordable consumer display that can accurately render the available large color volume that is representable using BT.2020 and ST2084 may initially be difficult until display designs evolve further. Consequently, some professional mastering monitors may exceed the initial next-generation consumer display device capabilities, making it possible to create content that exceeds what consumer displays are able to reproduce.

CONSUMER ELECTRONICS CHALLENGES

As stated above, the implementation of the full BT.2020 and SMPTE ST2084 specifications can be challenging for consumer display devices – both technically and economically. While brighter screens and darker black levels are very appealing to consumers, and provide a much improved visual experience, there are concerns that increased contrast ratio and peak luminance may also result in increased power consumption. Since power management is a very important consideration for consumer electronics manufacturers, HDR poses some challenges in that regard.

Power Consumption

In recent years, the brightness of consumer displays has increased, partially driven by the fact that brighter TVs apparently sell better. However, brighter displays also require more energy to generate the increased light output. Therefore, the balance between higher brightness and reasonable power consumption has been of paramount importance for display manufacturers.

A previous study [5] shows a correlation between power consumption and the average picture level (APL) of the content on LCD displays with individual backlight segment dimming, such that when the APL increases, the power consumption also increases. Another study [6] has shown that achieving increased luminance of 500 cd/m^2 in a portion of the frame (12% screen area) is possible while maintaining a similar power consumption (80W) as occurs when the whole frame (100% screen area) is displaying a lower luminance 300 cd/m^2 (90W). The relationship between power consumption, peak luminance and percentage of screen area is shown [6] to be dependent on the number of backlight segments, such that a 6x8 backlight segment arrangement achieves peak luminance with a 12% screen area box while a 12x12 backlight segment arrangement can achieve peak luminance even with a smaller, 4% screen area box. This means that a larger number of segments provide more granular image control, which generally results in better image quality. We expect that additional advancements have been made since this study [6] was published in 2009, and that HDR displays being introduced this year will be able to manage better performance tradeoffs.

This relationship between brightness and power consumption becomes even more important when HDR is introduced. Content may not only contain higher brightness, but also an increased simultaneous contrast ratio, where an HDR image may for example have image detail in dark shadow regions, while also having image detail in bright specular highlight regions. In that regard, remember that APL has been shown to drive the power consumption, not the peak luminance of a given pixel.

MANAGING CONTENT & DISPLAY CAPABILITIES

Since mastered content may exceed consumer display capabilities, it should be considered how this discrepancy could be managed in the market to ensure the best consumer experience. Recognizing that display capabilities will continue to improve every year, there may come a time when next-generation video content can be fully reproduced on consumer display devices. In the meantime, understanding the parameters of the mastered content is important, as is attempting a high quality reproduction on consumer display devices that matches the color and luminance values defined in the content as best as possible.

If the content contains color and luminance values that cannot be shown on the particular consumer display being used, then the consumer display must render different color and luminance values that can be shown. Recently this process has been described as "color volume mapping" to consider both color and luminance replacement simultaneously. More traditional approaches separate this problem into two parts, considering color gamut mapping separately from luminance tone mapping. Regardless of the techniques used by the consumer displays to replace color and luminance values that cannot be shown (due to consumer device limitations), the color and luminance replacement process can be informed by some simple metadata items about the mastering display and statistics about the light levels in the mastered content. While one mastering facility may use the same mastering display to master many titles, the light level statistics about the content mastered on the same mastering display will probably vary from title to title.

Mastering Metadata - Mastering Display Color Volume

The color volume of the mastering display can be described with chromaticity of the red, green and blue display primaries as well as white point in addition to the minimum and maximum luminance of the mastering display. These mastering display attributes can be represented by the metadata fields defined in the SMPTE ST2086 standard [9]. Traditional output-referred mastering workflows are commonly setup in a conservative manner in which the signal that is mastered is constrained to exist only within the color volume that can be shown on the mastering display that is used for the creative approval process (i.e. "what you see is what you get" = WYSIWYG). One caveat is that usually the encoding of black exceeds the color volume of the mastering display to avoid the visibility of shadow detail crushing on other displays that may have better black level performance. See SMPTE EG432-1 engineering guideline [10] (Annex I) for more details. This document also describes (in Section 8) how mastering display metadata can be used to inform a color gamut mapping strategy. While the document is oriented towards Digital Cinema applications, the general concepts described are also related to the next-generation home video specifications that permit representation of wide-color gamut and high dynamic range video.

Content Metadata - Maximum Content Light Level

With respect to describing the content itself, describing the maximum brightness, or Maximum Content Light Level (MaxCLL), of the content being mastered is important. The max(R,G,B) operator is applied to all pixels in all frames of the content to determine the maximum value (MaxCLL) for that particular content. If a simple tone mapping approach is used by the consumer display to replace pixel values in the content that are not representable on the consumer display with pixel values that are representable, then the MaxCLL could be used to define the upper bound of the value of the pixels that will be encountered in that particular content. More information about the techniques that can be used to compute the MaxCLL value appears in Annex A of the recently revised CEA 861.3 document [11] and also in a recent JCT-VC document [13].

Content Metadata – Maximum Frame-Average Light Level

Since MaxCLL only describes the brightest pixel, it does not provide any indication of the overall brightness of a specific frame, or the entire video stream. As such, considering an image of a sky at night, the stars in the sky may appear very bright, but also cover a very small area of the frame. Since the remainder of the frame is dark, the overall average luminance for this frame may be much lower than compared with a regular daylight scene. However, this increase in dynamic range (dark sky with accented bright stars) is exactly what makes HDR an interesting storytelling tool.

Further, considering the challenges within consumer electronics devices, the frame-average luminance is much more important from a power management perspective than the peak luminance within a given frame. Therefore, the Frame-Average Light Level of a given piece of content will be an important parameter for consumer display devices to understand.

The average of the max(R,G,B) operator is applied to all pixels in each frame to determine the frame-average maxRGB for each frame. The Maximum Frame-Average Light Level (MaxFALL) value is set to the maximum value of frame-average maxRGB of all frames in the content. The computation of the MaxFALL values only considers the active image areas of the frame, which is relevant when for example a 2.40:1 aspect ratio content is stored in a 16x9 frame with letterbox mattes for distribution to the home, so the MaxFALL value can remain valid if cropping or zooming is applied to the image. More information about the techniques that can be used to compute the MaxFALL value appears in Annex A of the recently revised CEA 861.3 document [11] and also in a recent JCT-VC document [13].

Carriage and Transmission of Mastering and Content metadata

The Mastering metadata and Content metadata described in the previous sections should be available to next-generation display devices via multimedia interfaces or file-based mechanisms. For example, the revised CEA 861.3 uncompressed video specification [11] has defined a mechanism to transfer the ST2086, MaxCLL and MaxFALL values from a source device to a sink device (like a display). Additionally, a SEI message to carry the ST2086 metadata within an HEVC bitstream was defined in the recently published revision to the HEVC specification [12] (in sections D.2.27 and D.3.27) and a SEI message to carry MaxCLL and MaxFALL metadata within an HEVC bitstream was recently proposed [13] and will likely be included in a future revision publication of the HEVC specification.

Analysis of Commercial Content

Having identified two parameters, MaxCLL and MaxFALL, that are relevant when describing HDR content, we analyzed some commercial HDR content using the above methods. Figure 1 depicts the Frame-Average Light Level and the Maximum Light Level for each frame of a 3 minute portion of Reel 1 of the "Lego Movie" HDR version that was graded on a monitor with P3 RGB primaries, D65 white point minimum luminance of 0.005 cd/m^2 and maximum luminance of 4000 cd/m^2. It clearly shows that the frame-maximum light level for many frames exceeds 1,000 cd/m^2, with the absolute maximum content light level (MaxCLL) for this clip being 2,153 cd/m^2. At the same time, the frame-average light level is much more moderate, staying below 100 cd/m^2 for the majority of frames. In fact, the average frame-average light level for the entire Reel 1 is 26 cd/m^2, yet the maximum frame-average light level (MaxFALL) for Reel 1 is 768 cd/m^2.

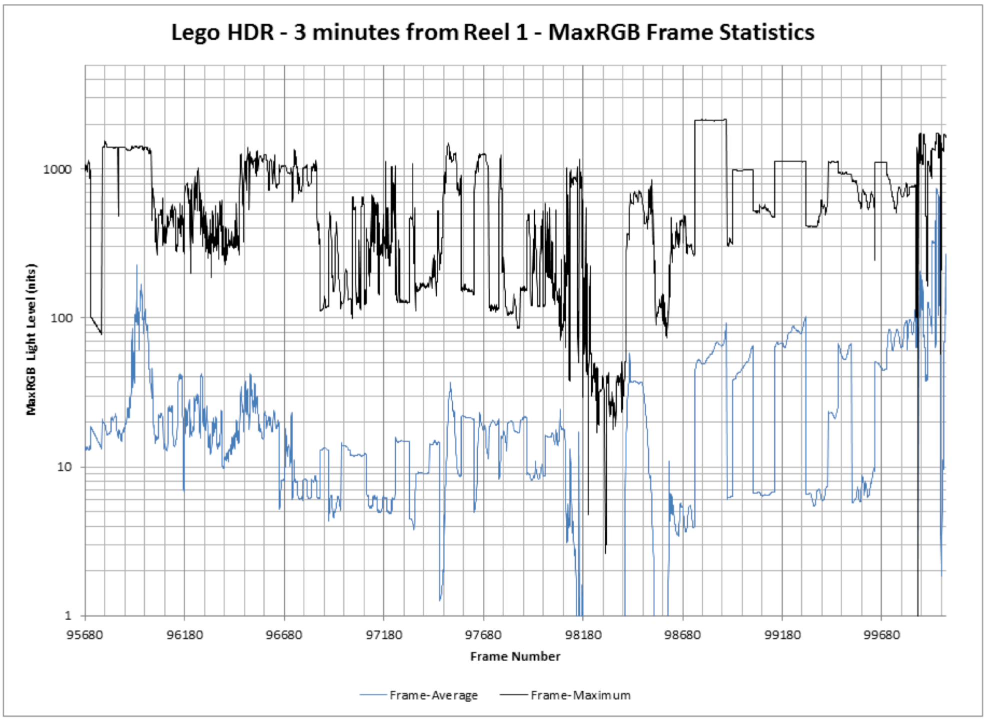

Figure 1 – Frame-Average Light Level and the Maximum Content Light Level for each frame of 3-minute portion of Reel 1 of the "Lego Movie".

Figure 2 explores the frame-average light level distribution across the entire “Lego Movie” – Reel 1 through 5. It shows a fairly consistent distribution, with the majority of frames averaging below 100 cd/m^2. Only 3.7% of the frames throughout the entire movie exhibit more than 100 cd/m^2 of frame-average light levels while only 0.02% of frames throughout the entire movie use more than 1,000 cd/m^2.

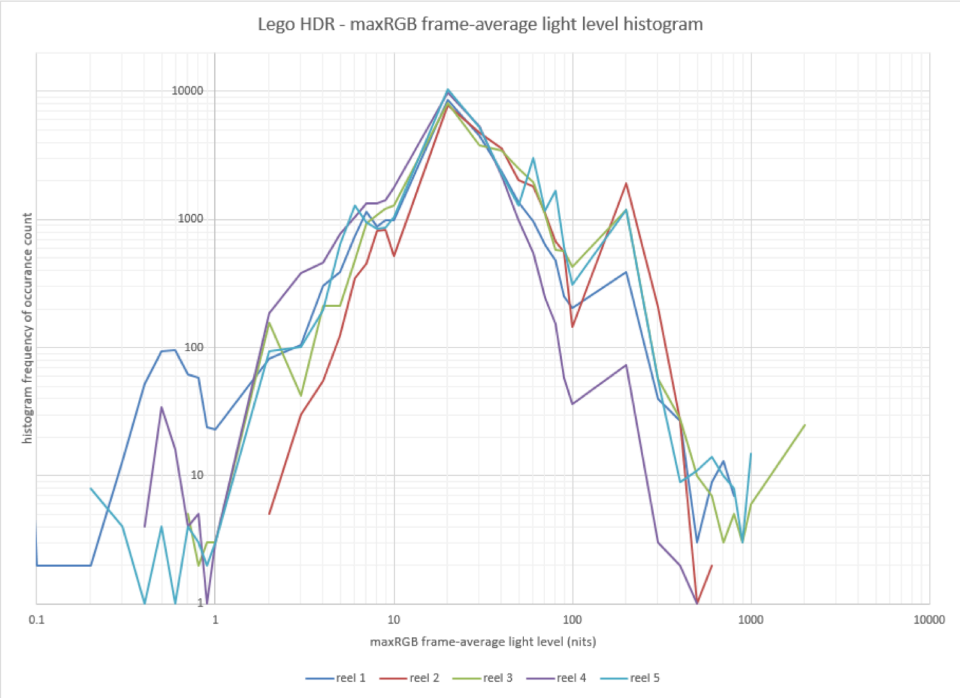

Figure 2 – Histogram of Frame-Average Light Levels for the entire “Lego Movie” (Reel 1 through 5).

Figure 3 further explores the frame-maximum light levels across the entire “Lego Movie” HDR version – Reel 1 through 5. It also shows a fairly consistent distribution, with 75.5% of the frames throughout the entire movie having a frame-maximum light level more than 500 cd/m^2 while 31.6% of frames having a frame-maximum light level more than 1,000 cd/m^2.

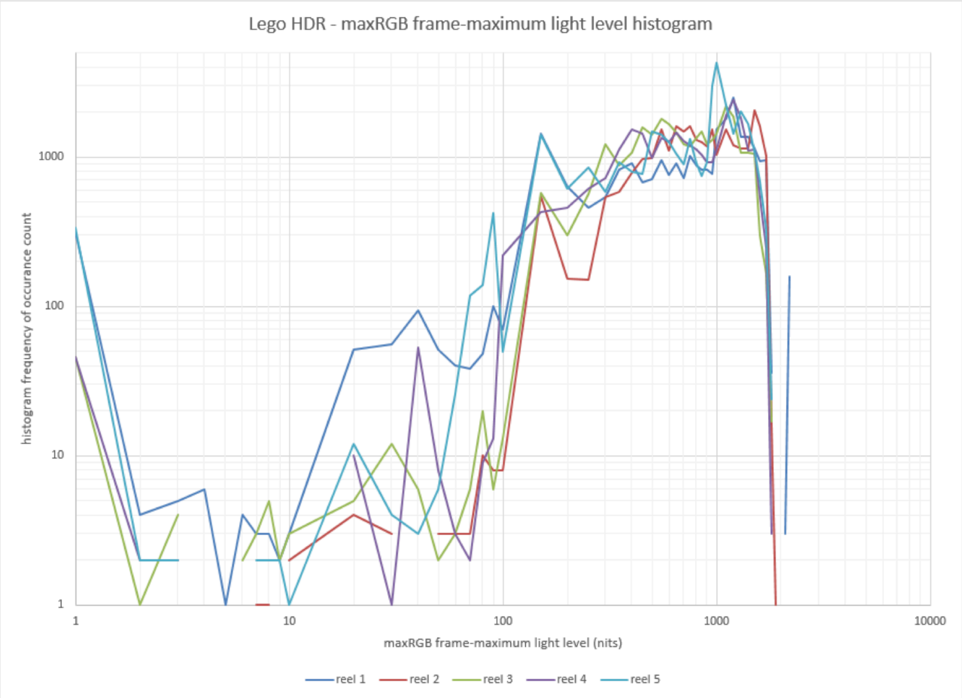

Figure 3 – Histogram of Frame-Maximum Light Levels for the entire "Lego Movie" (Reel 1 through 5).

Table 1 also lists the MaxCLL and MaxFALL values calculated for each reel of the "Lego Movie" HDR version – Reel 1 through 5.

	MaxCLL	MaxFALL
Reel 1	2,153.34	768.19
Reel 2	1,813.37	814.19
Reel 3	1,702.22	1,521.03
Reel 4	1,733.24	614.71
Reel 5	1,702.22	905.29
Across all Reels	2,153.34	1,521.03

Table 1 – MaxCLL and MaxFALL values for the "Lego Movie" (Reel 1 through 5)

In addition to the maximum and frame-average light levels, we further explored how the HDR version of the "Lego Movie" compared to the standard dynamic range (SDR) version that is currently being used for regular SD and HD home video distribution. For this purpose, we decided to compare the median values for each frame between both versions because the frame-median value would be less impacted by the higher peak luminance pixels that exist in the HDR version compared to the frame-average value. Figure 4 represents this comparison of the frame-median values of the first minute of Reel 3 of the "Lego Movie", and clearly shows that they are very similar. This indicates that the HDR version of the movie doesn't significantly change the mid tones, which usually contributes to the look and feel of a movie. Instead, the HDR version of the "Lego Movie" primarily takes advantage of expanding the dynamic range with more details in the dark areas, and accentuating highlights with increased peak brightness. In such a scenario, a reasonable expectation would be for a consumer display to roll off the brightness levels it cannot achieve without changing the overall look and feel of the movie.

Notably, this analysis of a single movie only provides one data point, and the distribution of the content luminance values is very content dependent. However, this analysis also suggests that information about the content light levels can be used to ensure that consumer display manufacturers can better manage the rendition of content in their respective devices.

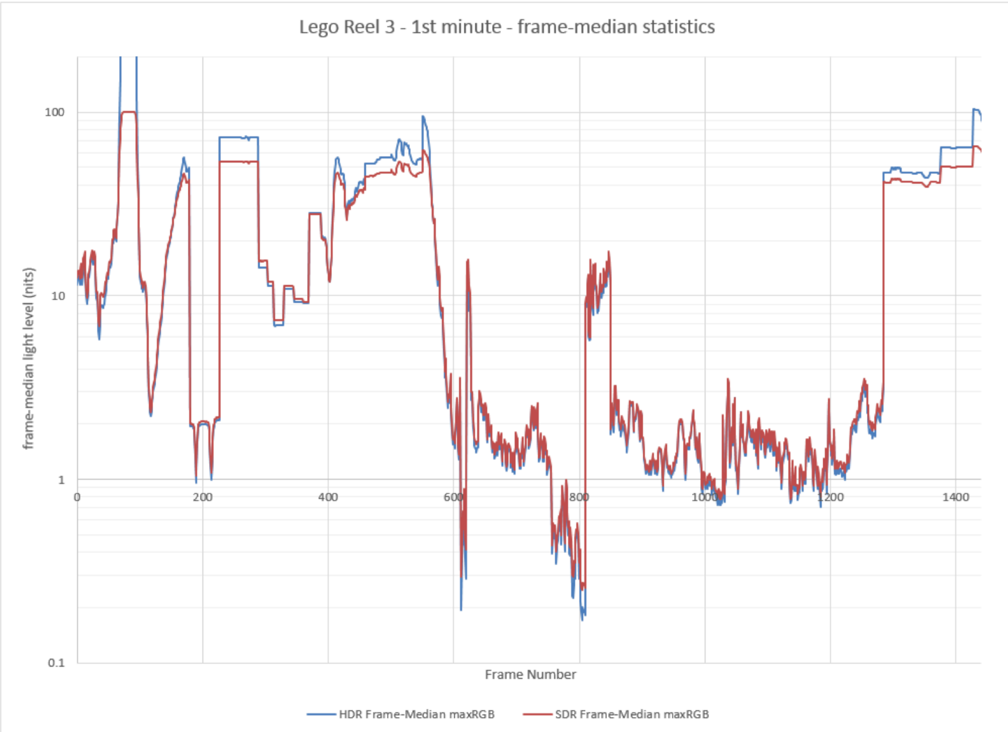

Figure 4 – Comparison of Frame-Median Light Level between SDR and HDR version of Reel 1 of the “Lego Movie”.

Considerations during Content Creation

HDR provides an exciting new tool for creative filmmakers to enhance their storytelling. But since the production tools and workflows have not been fully established yet, some consideration should be given to the content creation process.

As discussed above, understanding the parameters used by the content itself can be very helpful – both for content creators, and also for display devices. As such, content production tools, such as color grading systems, should perform relevant analysis of the content while it is being produced. Providing visual representations of the content parameters, with histograms or other means, would be very meaningful and helpful.

Since the content grading process is typically performed on a professional mastering monitor that may be capable of displaying much higher dynamic range compared to consumer displays, it will be important for the content producer to recognize that the experience may not be the same on a consumer display with lower dynamic range capabilities. In order to assist the content producer during the color grading process, there would be value in providing dialogues and warnings within production tools that signal when content exceeds the desired target values for frame-average light levels. This way, the content producer would be aware that the image seen on the professional mastering monitor might not be achievable on a consumer display, resulting in a slightly different experience. At the same time, a content producer would have the opportunity to make adjustments to the content, if desired, without necessarily changing the overall look and feel. Obviously, content can still be created exceeding the target values, but there are advantages for the content producer to understand possible implications related to the consumer device rendition.

The newly released ACES v1.0 source code [14] contains a set of Output Device Transforms (ODTs) targeting ST2084 HDR displays with various peak luminance values including 1,000 nits, 2,000 nits, and 4,000 nits. These ODTs roll off an HDR image's bright highlight detail to varying degrees while not changing the mid tone values. Using an ACES workflow incorporating these ODTs would allow the creative to quickly visualize the display of HDR content at various levels of display capability.

After all, the consumer display capabilities are more limited compared to professional mastering equipment, particularly with respect to power management. If content producers were to limit the average or median frame light levels of a given piece of content, this would ensure that consumer displays could maintain the overall look and feel of the content, and still provide a true HDR image by managing the peak luminance areas to their device capabilities, via roll off or other advanced algorithms.

Considerations in Consumer Displays

Achieving high brightness in consumer displays is important to deliver a true HDR experience for consumers [8]. At the same time, the impact to the power consumption has to be taken into account. There has been lots of research in the area of reducing power consumption while maintaining peak brightness through various implementations of local dimming and segmented backlighting [5-7]. Some segmented backlighting systems allow power savings over 50%, which can be used for boosting the peak luminance of the display. In fact, some segmented backlighting systems attempt to maintain a pre-defined power consumption level, and with some segments operating at low power levels, the remaining power can be delivered to other segments that need to provide a higher peak luminance. Obviously, there's a trade-off between the number of segments implemented in consumer displays, and the higher cost associated with it.

SDR content is often created with some pixels containing shadow detail that cannot be reproduced on all SDR displays. For example, the SDR EOTF BT.1886 [2] that is in common use in the industry avoids unpleasant clipping artifacts such as shadow detail crushing. The BT.1886 EOTF effectively describes a tone mapping of shadow detail that is dependent on the SDR display's black level (L_B) parameter. HDR content is often created with some pixels containing shadow detail beyond the capability of all HDR displays, and it is expected that high-quality HDR displays will perform tone mapping of shadow detail in a similar manner to high-quality SDR displays.

HDR content is often created with some pixels containing increased content peak brightness levels that cannot be reproduced on all HDR displays. High-quality HDR displays should avoid hard clipping at their peak luminance limit and instead roll off bright highlights that exceed the display capabilities, mapping the values that cannot be shown to values that can be shown on the display. Of course, more advanced tone mapping algorithms are desired, potentially even required, especially in scenarios where the frame-average luminance exceeds the display capabilities. In either case, consumer displays should be designed expecting to receive content with video parameters exceeding their own display capabilities, and the displays should be prepared to manage the signal processing without compromising the user experience.

CONCLUSIONS

The introduction of next-generation video technologies, particularly HDR, provides a very noticeable and improved user experience. While HDR is very compelling for content creators and consumer electronics manufacturers alike, HDR also provides some challenges. Content may be mastered with parameters exceeding today's consumer display device capabilities. At the same time, consumer display device limitations, particularly related to power consumption, should be taken into consideration. To strike a balance, analysis and understanding of the different content light levels during production ensures that consumer device manufacturers can better manage the rendition of content in their respective displays.

REFERENCES

[1] Recommendation ITU-R BT.709, "Parameter values for the HDTV standards for production and international programme exchange", ITU (2002)

[2] Recommendation ITU-R BT.1886, "Reference electro-optical transfer function for flat panel displays used in HDTV studio production", ITU (2011)

[3] Recommendation ITU-R BT.2020, "Parameter values for ultra-high definition television systems for production and international programme exchange", ITU (2014)

[4] SMPTE ST 2084:2014, "High Dynamic Range Electro-Optical Transfer Function of Mastering Reference Displays", SMPTE (2014)

[5] H. Chen *et at.*,"Evaluation of LCD local-dimming-backlight systems", SID Journal Vol. 18/1, pp. 57~65 (2010)

[6] G. Kim *et al.*, "Adaptive Luminance and Power Control (ALPC) for LED Backlight Units", SID Digest Vol. 48/1, pp. 723~726 (2009)

[7] E. Langendijk *et al.*, "Contrast gain and power savings using local dimming backlights", SID Journal, Vol. 16/12, pp. 1237~1242 (2008)

[8] P. Hanhart, P. Korshunov, T. Ebrahimi, Y. Thomas, H. Hoffmann, "Subjective Quality Evaluation of High Dynamic Range Video and Display for Future TV", IBC2014 Conference, 2014 page 2.2

[9] SMPTE ST 2086:2014, "Mastering Display Color Volume Metadata Supporting High Luminance and Wide Color Gamut Images", SMPTE (2014)

[10] SMPTE EG 432-1:2010 "Digital Source Processing - Color Processing for D-Cinema", SMPTE (2010)

[11] CEA Standard CEA-861.3 "HDR Static Metadata Extensions", January 2015

[12] Recommendation ITU-T H.265, "High efficiency video coding", ITU January 2015

[13] Chad Fogg, Jim Helman, Michael Smith, Michael Zink, "Content light level information SEI", JCT-VC-T0101, Joint Collaborative Team on Video Coding 20th Meeting: Geneva, CH, 10–18 Feb. 2015

[14] ACES v1.0 source code, available https://github.com/ampas/aces-dev/releases/tag/v1.0