Live Single-Stream HDR Production for HDR and SDR Cable Distribution

What Happens When "Post Production" Is Real-Time

prepared for SCTE•ISBE by

Chris Seeger
Director, Advanced Content Production
NBC Universal, LLC
900 Sylvan Avenue
Englewood Cliffs, NJ 07632
chris_seeger@nbcuni.com

Table of Contents

Title **Page Number**

1. INTRODUCTION 3
2. CONTENT CONVERSION AND PERCEPTUAL MEASUREMENT 3
   2.1. How the Human Visual System Sees Light and Color 5
   2.2. SINGLE-STREAM HDR AND SDR PRODUCTION 9
   2.3. Hybrid Log-Gamma 11
   2.4. HLG from camera to display 12
   2.5. Color Conversion and objective color accuracy measurements 12
3. SINGLE-STREAM HDR-SDR PRODUCTION 15
4. HDR/SDR SINGLE-STREAM DISTRIBUTION 18
5. CONCLUSION 20

ABBREVIATIONS 21

BIBLIOGRAPHY & REFERENCES 21

List of Figures

Title **Page Number**

Figure 1 - How Do We See Light and Color? (© Universal Pictures from PacRim) 4
Figure 2 - The Human Visual System and Color 6
Figure 3 - Static versus Dynamic Adjustments to Light Levels 6
Figure 4 - The Human Visual System is Relative 7
Figure 5 - Color Appearance and HDR Production 8
Figure 6 - Nothing in This Image is Actually Moving 9
Figure 7 - Single-Stream HDR and SDR Production (ITU-R Report BT.2408) 10
Figure 8 - Objective Color Accuracy Measurement 13
Figure 9 - LUT Conversion (devices & software): Sarnoff(SRI) “Yellow Brick Road” Pattern 14
Figure 10 - LUT Conversion: Delta-E ITP (SDR-to-HDR) Using Sarnoff(SRI) Color Checker 2014 15
Figure 11 -: Examples of Shading: No Highlights 17
Figure 12 - Examples of Shading: Highlights 18
Figure 13 - This conversion produces a PQ image which is visually identical to HLG 19
Figure 14 - A simplied view of the NBCU/Comcast production through distribution in PQ 19

1. Introduction

Developing video content for the extended dynamic range and color depth in high dynamic range (HDR) and wide color gamut (WCG) now common in consumer displays isn't necessarily new: the first functional 4K/HDR televisions came out in 2016, and popular over-the-top (OTT) video platforms have routinely produced content in HDR since then. However; it is far, far more difficult to shoot and deliver HDR content for live linear workflows than it is for most file-based post produced approaches. That's because in live HDR production environments, handling graphical placements, color correction, and luminance normalization across what is usually a mix of SDR (standard dynamic range, which is essentially contemporary HDTV) and HDR cameras must be done on-the-fly and in real-time.

Also, differences in system colorimetry exist within the primary HDR systems, specifically Hybrid Log-Gamma (HLG) and Perceptual Quantizers (PQ), and took us some time to understand. A functional workflow for live delivery requires multiple levels of conversion for various signals throughout the infrastructure - like graphics, video playback, SDR and native HDR cameras. This three-part paper will explore 1) conversion and perceptual measurement for HDR and SDR; 2) how HDR production workflows are orchestrated for live content, and 3) how those live linear workflows are distributed – a topic likely of most interest to the cable television industry.

In addition, the paper will illuminate the research and collaborative processes necessary to build a functional, single-stream HDR-SDR workflow, at each stage of the pipeline, without compromising the artistic intent or quality of the distribution paths for either the HDR or the primary (and revenue-generating) SDR. In-depth research of color conversions will be discussed and explored, such as the specialized techniques described in the International Telecommunications Union – Radiocommunications Sector's (ITU-R) Recommendation BT.2124. Production layouts will be explored and described, so that readers gain a deeper understanding of decisions made in the baseband (serial digital interface(SDI) and IP – Internet protocol), file-based conversions/transcodes, and orchestration layers.

2. Content Conversion and Perceptual Measurement

Content conversion and perceptual measurement matters because of the vast differences in a scene you see with your own eyes, and that same scene captured by a camera and displayed on a screen. In real life, we see these bright and colorful scenes – but when we capture and move the image through to legacy TV's, some of the highlights and light levels get muted. That's because a traditional TV can was designed to display 100 nits (a "nit" is a measure of brightness, roughly equivalent to the light of one candle per square meter). Figure 1 shows a simulated scene represented in both HDR (lower image) and SDR.

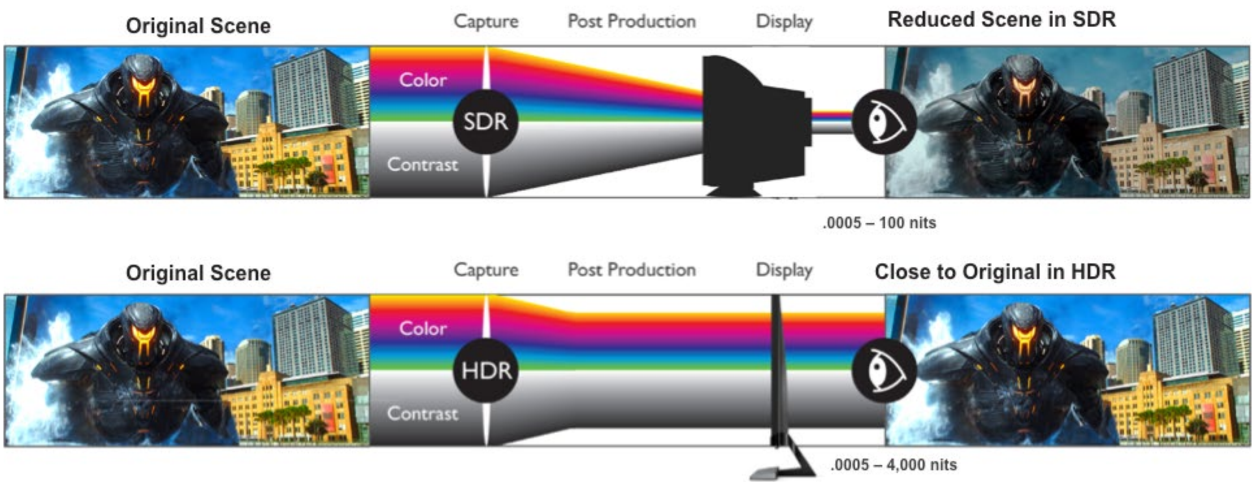


Figure 1 - How Do We See Light and Color? **(© Universal Pictures from PacRim)**

The interplay between cameras, display screens and brightness dates back to the early days of analog cathode ray tune (CRT), which could achieve only a relatively marginal level of brightness, measured at approximately 100 nits. Compare this against today's HDR TVs (and some smartphones) – which can be as bright as 4,000 nits; and SDR TVs, which is to say "traditional HDTVs," that can scale the peak white to a maximum of around 400 nits (note that a recommendation that is in development would make it around 200nits so it matches reference/graphics white in HDR). As time went on, and digital techniques replaced analog, cameras advanced to the point of being able to record a much wider dynamic range – meaning more detail to describe a larger range of luminance, from dark to bright, and a far richer span of colors that is closer to what the human visual system is capable of. HDR(High Dynamic Range) gives you substantially better highlights (the twinkly stuff), and WCG(Wide Color Gamut) gives you redder reds, bluer blues, blacker blacks.

HDR got a better foothold, commercially, within the medical imaging community, which required a method to see in a highly detailed and reliable way. As it turned out, the HDR images were so vivid, and so much closer to the workings of the human visual system (which we'll get to next), that it began to be developed for consumers. In today's consumer electronics marketplace, it's getting harder to find a television that doesn't include HDR, than one that does.

Dolby Laboratories (Dolby), in particular, focused on a system of quantization that could standardize the way HDR content is displayed, known in the industry as "PQ," for perceptual quantization. In essence, PQ adds more bits to areas where the human visual system can see more detail and fewer bits where we see less detail. The thought is, if we can't see the detail, why store more detail. Additionally, PQ decided to use an absolute mapping which associates a specific code value with a luminance level.

If history is our guide, a proposed video standard isn't really a technology story unless a competing standard is vying for consideration. Which, in the case of HDR video, is HLG, co-developed by the British Broadcasting Corporation (BBC) and Nippon Hoso Kyokai (NHK – also known as the Japan

Broadcasting Corporation). Unlike PQ, HLG deliberately has some aspects of backwards compatibility to SDR(SDR, which is to say "HDTV" that uses traditional gamma mapped signals) when used with wide color gamut or ITU-R Recommendation BT.2020. Backward compatibility does not exist with older legacy displays that use the narrower color space in HDTV (ITU-R Recommendation BT.709). It works by defining a nonlinear transfer function, known as an electro-optical transfer function (EOTF), in which the lower half of the signal values use a gamma curve, and the upper half use a logarithmic curve.

SDR and HLG are relative systems that scale the entire image from a lower luminance level to a higher luminance value (the dynamic range doesn't change only the luminance level). SDR was designed as a 100 nit system and we scale it to 200-400 nits for consumer delivery.

HLG was originally designed as a 1000 nit system and was adapted with a newer algorithm (ITU-R Recommendation BT.2100) in order to scale to other luminance levels without hue shifts. Scaling shifts shadows, midtones and highlights which can be an important consideration when distributing a signal to consumer displays that are getting brighter and brighter. Scaling does not increase the dynamic range (it's like turning the volume up in audio).

The differences between the absolute (PQ) and relative (HLG) HDR systems are important to remember when we get to the distribution section of this document.

Our work producing and distributing HDR content began after a test run with the 2015 Independence Day fireworks in New York city, and then in earnest with the opening ceremonies for the 2016 Olympic Games, in Rio de Janeiro, Brazil. Since then, Comcast/NBCUniversal produced and distributed HDR video for the 2018 PyeongChang Winter Olympics; the 2018 and 2019 Notre Dame football season; the 2018 men's Federation Internationale de Football (FIFA) games; and, earlier this year, Chicago Blackhawks hockey games. We learn something new every time we venture into an HDR/WCG event, and in particular these learnings tend to focus on ways in which to optimize content conversions, and avoid unnecessary duplications. The aim is a live production that "produces once, outputs twice:" Once in native HDR; once in converted SDR. This is no small feat. The rest of this paper will explain why.

2.1. How the Human Visual System Sees Light and Color

If the ultimate goal of video production is to faithfully reproduce what the eye sees, then a rudimentary understanding of the human visual system (HVS) is useful. The basics: the human eye takes in light, which is processed primarily by cone cells, which are sensitive to certain wavelengths – long, medium and short. Our rod cells are sensitive to dark scenes, and thus aren't as involved in the processes of, say, watching television. Rod cells are more numerous than cone cells, but are less important, perceptually, during normal lighting conditions.

Figure 2 illustrates how the HVS processes information. Incoming light is captured by the three photo receptors (cones) in the eye, which have peak sensitivities in the Long, Medium and Short wavelengths, designated "L," "M" and "S." Linear light is converted into a non-linear signal response, to mimic the adaptive cone response of the HVS. Notice the different ratio of L/M to S wavelengths.

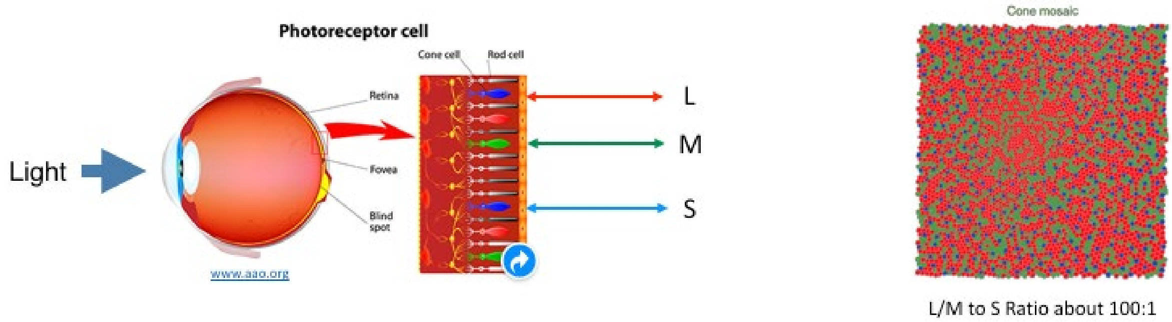


Figure 2 - The Human Visual System and Color

Not only that, but the HVS is adaptive, so, the responses can change, and those changes must be taken into account, especially when considering a much wider dynamic range.

Figure 3 illustrates static versus dynamic adjustments to light levels. Adapting to darker environments – like attempting to see objects in the night sky, after being in a lit room – can take as long a 30 minutes; in brighter environments, it can take five to 10 minutes for the eyes to adapt. In general, the HVS can see about 12 simultaneous stops; “stops” are a scaled measure of dynamic range. But! The HVS can operate within a 24-stop range, by doing a number of things, like scanning up and down in different brightness ranges, then adapting(perceptual) while adjusting the human pupil (physical), which when dilated, allows us to see darker detail, but loses brighter detail. Pupil contraction allows us to see detail in brighter objects, but sacrifices darker detail.

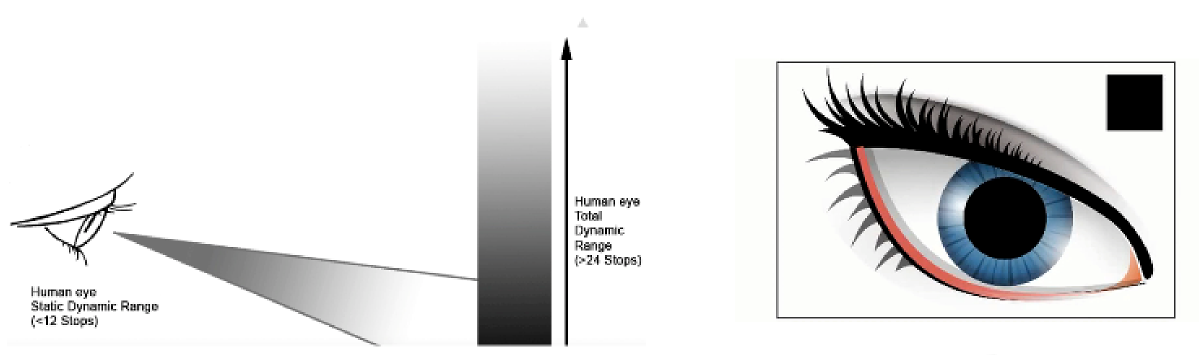


Figure 3 - Static versus Dynamic Adjustments to Light Levels

Because of adaptive factors (some of which aren’t mentioned), the HVS is relative. It plays tricks on you!

Allow me to prove it to you:

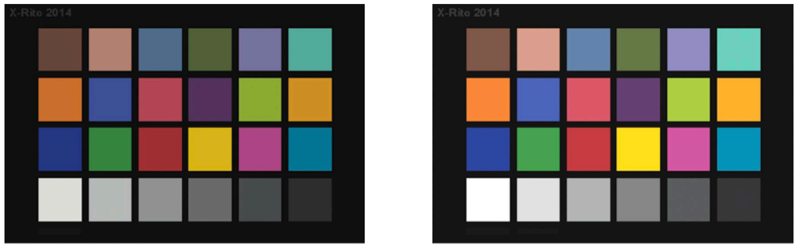
In Figure 4, the stripe in the middle of the image appears brighter on the left side than the right – but it's not. As you scan from left to right on Figure 4, the images will appear to change – but they are not. It's merely your brain, adapting.At any moment in time, the HVS is juggling a complex mixture of perception, adaptation, sensitivity, acuity, and "day vs. night" variables. As content producers, it's important to pay attention to these variables, so as to provide consistent content to our viewers.

Figure 4 - The Human Visual System is Relative

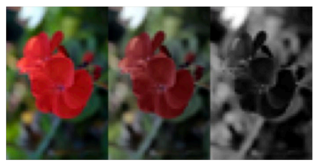
Then there's the matter of how the Human Visual System interprets color and contrast, and how many different ways the brain can see things.

For instance:

- The "Hunt Effect," simulated in the upper portion of Figure 5, illustrates how color saturation appears to increase – but the only thing that actually increased was the luminance/brightness.
- The "Stevens Effect," does something similar with contrast, in that it appears to increase with brighter light.
- The "Bartleson-Breneman Effect," investigates how display image contrast increases with the luminance of surround lighting.
- The "Helmholtz-Kohlrausch," shows that where there is a saturated and a neutral color of the same luminance, the saturated color may appear brighter. This is another example of how saturation plays a role in perceived brightness, which will change with the viewing environment.
- The Perkinje Effect (illustrated in the lower image in Figure 5) describes the tendency for the peak luminance sensitivity of the eye to shift toward the blue end of the color spectrum at low illumination levels as part of dark adaptation. In consequence, reds will appear darker relative to other colors as light levels decrease



Hunt Effect



Perkinje Effect

Figure 5 - Color Appearance and HDR Production

Because of the vast increases in luminance that come with HDR technologies, a significant issue to resolve is overall viewing comfort. Simply put, with HDR, as objects of certain colors get larger, they can become uncomfortably bright to view. The ITU-R produced standards for all levels – but they don't include object size, color and temporal elements. As a result, small, twinkly images can be really beautiful, but large, bright objects can quickly become offensively bright.

This requires broadcasters and video content producers to purposefully shade material by observing additional elements mentioned earlier (object size, color and duration on-screen). Eventually, my hope is that "color loudness" rules will be established, just as there are for audio. Although out of scope for this paper, we recently filed a preliminary patent with some concepts which we believe will help to develop a "comfort metric" for HDR measurement tools.

Figure 6 is a bit of "eye-candy" to expose other strange perceptual elements. In the image, nothing is moving, but the blue dots appear to be moving. It illustrates the impact of drastically different contrast levels, which essentially make our brains freak out.

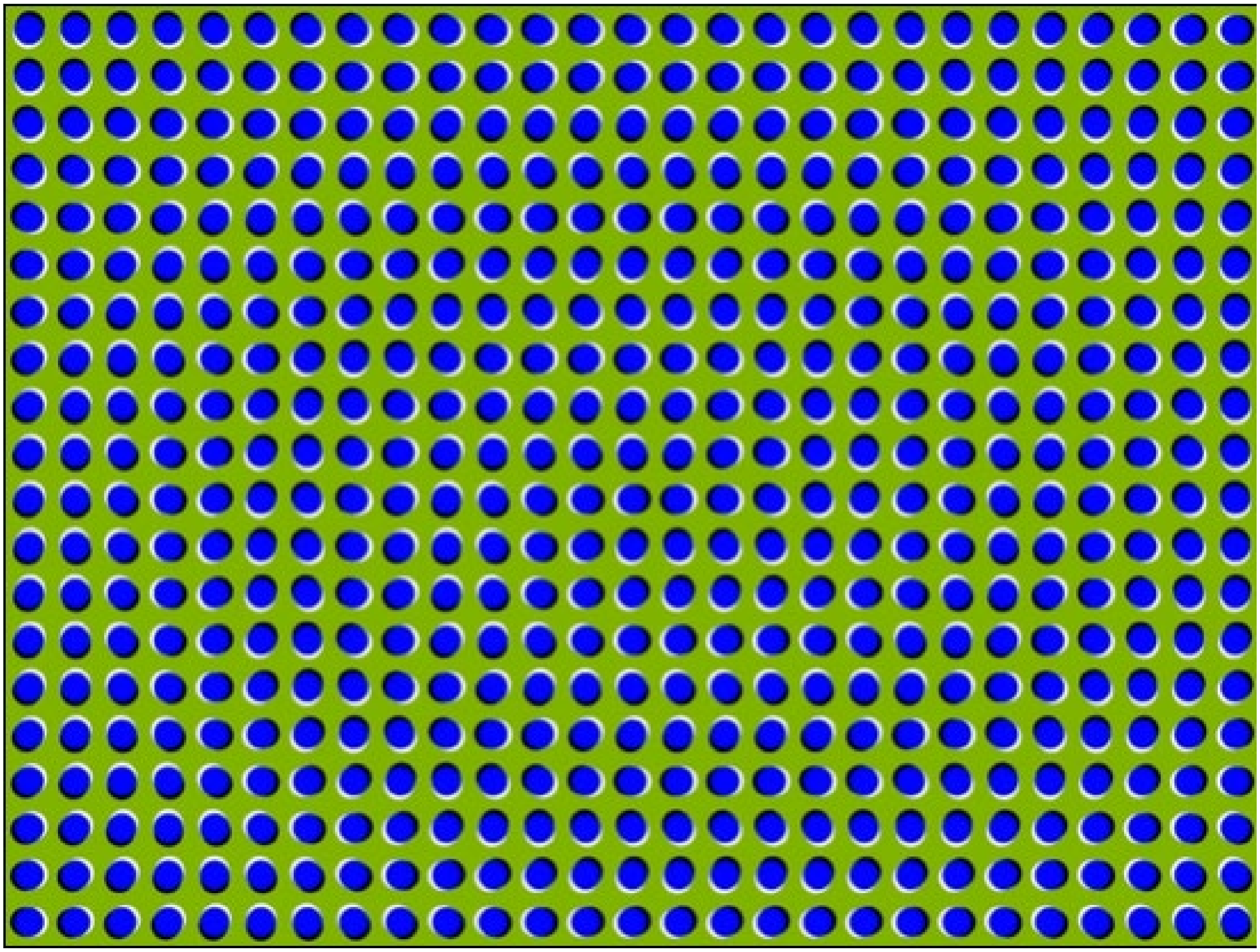


Figure 6 - Nothing in This Image is Actually Moving

2.2. SINGLE-STREAM HDR AND SDR PRODUCTION

Figure 7 takes us directly into the deep waters for content producers. It depicts the inner workings of a video control facility, as well as a control room, based on ITU-R Report BT.2408. In essence, we mapped out all of the pieces of the live HDR puzzle, necessarily including cameras, displays, the video switcher, graphics, distribution outputs, and how a typical HDR production would be laid out. Few (if any) broadcasters have budgets to support a full complement of HDR cameras; generally speaking, the HDR cams are used for "important capture," while the extant suite of SDR cameras are put to use in, say, helicopter shots or those less vital to the overall imagery. Over time, all the camera's will be capable of HDR.

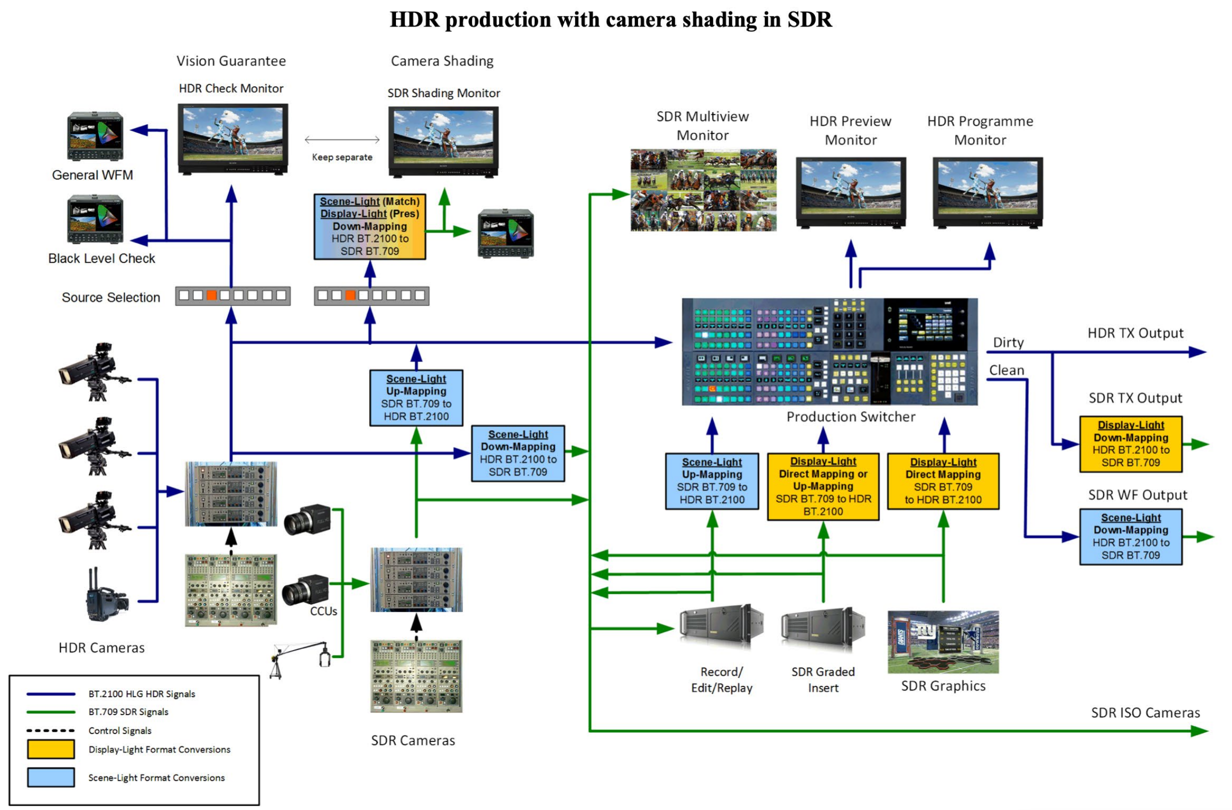


Figure 7 - Single-Stream HDR and SDR Production (ITU-R Report BT.2408)

2.2.1. Video Control

The video control area includes both HDR and SDR cameras, managed by a "shader" – essentially a "golden eye" for color and light – who matches the outputs of both camera types before they go on air. All sources need to be converted to the same HDR system before they enter the production, therefore, SDR sources are converted by "direct-mapping" or "up-mapping," to HDR, using specific look-up tables (LUT). Specific LUTs are used and are chosen dependent on each production. If, for instance, the production uses S-log3 versus HLG there will be a different LUT set. Also, some productions prefer a different mapping of light levels requiring a different set of LUTs. The primary goal for the design of any LUT-set is to preserve artistic intent on the HDR and SDR sides of the output.

In HLG productions, two types of conversion must be applied. One is called "Scene Light Conversion," which gets applied to older SDR camera's in order to match them with native HLG's "natural look", and the other is "Display Light Conversion." which get's applied to all other inputs and outputs (HLG to SDR).

Scene Light Conversion applies the source video's inverse-OETF (optical-to-electrical transfer function) to generate scene light prior to conversion, using the output OETF (to HLG, SDR or PQ) as the video signal. In essence, it converts the SDR signal, with its color balance, and rebalances the colors in order to better match with a native HLG camera. HLG-BT.2100 cameras produce what has been labeled as the "Natural Look". Some markets set their camera's to output what has been labeled the "Traditional Look" which restores a similar color balance and saturation compared to SDR or PQ. This could complicate media exchange in the future.

SDR-to-HLG Display light conversion (DLC) applies a the SDR EOTF prior to the conversion to HLG, in order which converts the source to displayed light which properly preserves the artistic intent (color balance) of the source. SDR-to-HLG scene-light conversion will desaturate images while HLG-to-SDR scene-light conversion would oversaturate images because the variable luma-gamma in HLG will not have been applied.

Predictive LUTs are also an important video control room component, so that HDR to SDR conversion can be previewed, such that we know exactly what the SDR output feed to legacy distribution will look like, before it goes to air.

2.2.1. Control Room

In the control room, depicted in the right box of Figure 7, a multi-viewer screen is used by the director and technical director, to examine all camera outputs and select which shot to use. A LUT can be used here which will output the HDR feed into the SDR multi-viewer compilation. The control room is also where slow motion replays, SDR graphics, and HLG replays are composited in the video switcher, all while preserving the color balance of the source. Those elements are direct-mapped using Display Light Conversion (DLC). Native HDR playback and HLG replays do not require upconversion, and feed directly into the video switcher.

The switcher outputs directly to a native HDR distribution feed, which then goes to the transmission system; the signal will also be transmitted in SDR to our legacy broadcast network. Legacy SDR output from the HDR production requires a DLC (HDR-to-SDR), which reduces the dynamic range but maintains the color balance and therefore as much of the original artistic intent in the HDR production as possible. We have found that mapping HLG-75% to SDR-95% and then constructing a subjective knee which preserves additional highlights from HLG between SDR 95% and 109%. An HLG "pre-compression" knee applies compression natively in HLG from 90%-109% and then applies an additional knee which "post-compresses" HLG 75%-109% to SDR 95%-109%. This process creates substantially more compression in upper end of the HLG curve beyond SDR "legal range" and preserves additional highlights within SDR "legal range."

2.3. Hybrid Log-Gamma

HLG works by capturing scene light, then sending it through an OETF. The resul is a curve, representing the scene's light level, that determines the quantization of the light levels into certain bits or code values. The scene light signal is what gets carried to the display, via SDI, IP or file for video distribution. Inside the receiving display, assuming it is HLG-equipped, several things happen: it applies an inverse-OETF, to reverse what was done in the camera, then an OETF is applied using gamma-to-luma only (not R,G,B - red, green, blue), so that it can scale the luminance of the images depending on the display's peak brightness capabilities. A 1,000 nit display will use one overall system gamma value, while a 2,000 nit version will use another, and that gamma value will scale only the light level, separate from the color. Perceptual effects like the Hunt Effect are important to remember when varying luminance and still require more study to understand what their impact might be. Variables like the perceptual effects are part of why we convert to PQ for distribution.

2.4. HLG from camera to display

HLG signals, from camera to display, apply a variable luma gamma in the display, which is determined by the display's peak brightness capability. By contrast, SDR and PQ apply gamma to R, G & B. This difference in gamma application for HLG produces a non-linear difference in chromaticity, relative to SDR or PQ.

The offshoot of this rather difficult process is that the signal looks desaturated compared to what we're used to in SDR. In some countries, that desaturated look is considered "less colorful," even though more accurate to the scene. The desaturated version is commonly known as a "natural look," whereas a "traditional look" re-saturates the images. As a direct result, HDR cameras carry two HLG settings – one for the natural look, the other for the traditional look. Deciding and/or converting between the two "looks" adds an additional level of complexity in broadcast workflows. A "Mild" setting is provided in some camera settings. It achieves a setting in between natural and traditional.

2.5. Color Conversion and objective color accuracy measurements

At this point, we know that we have to convert between SDR, HLG, or PQ. For the live-linear realtime conversions we commonly use 3D LUTs with a minimum size of 33 points and recommend a high quality interpolation method. 3D LUTS allow our file-based conversion workflows to match our baseband hardware conversions (live and post production conversions will match). Newer conversion methods are becoming available that allow for purely mathematical transforms. Purely mathematical transforms perform much less interpolation than LUTs and may become preferred over time.

In the early days of our HDR journey, we made subjective measurements for every LUT and every piece of new conversion hardware – we'd put the video in (test patterns and actual video), then examine the conversion vs the original side-by-side on a displays, and on a scope. It was a fairly inaccurate way of looking at very complex things like light and color. We found ourselves needing an objective measure that could make critical measurements. The ITU adopted metric created by Dolby (BT.2124) for objective color volume measurement. ITU-R BT.2124 uses a "color difference metric, called Delta-E ITP which measures the color volume difference (Delta) of one normalized sampled color volume (HLG, PQ and SDR) and another. ITP is a derivative of ICtCp which applys a small conversion to Ct (Ct/2). I and Cp remain the same.

ICtCp was designed to use key aspects of the HVS, in three steps, illustrated in Figure 8.

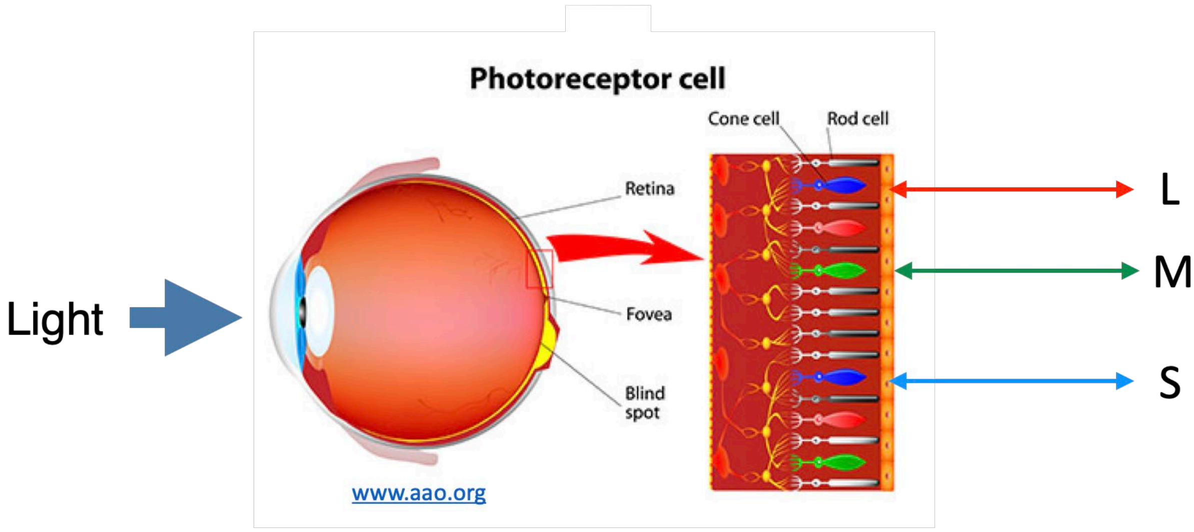


Figure 8 - Objective Color Accuracy Measurement

First, incoming light is captured by the three photo receptors (cones) in the eye, that have the previously discussed peak sensitivities in the Long, Medium and Short wavelengths. Next, the linear light is converted into a non-linear signal response, to mimic the adaptive cone response of the HVS. It is ICtCP's ability to algorithmically mimic the HVS that allows us to measure the perceptual appearance. It's not perfect, but it goes a long way to mimic how our eyes and brains translate color and light.

It's worth noting that this was important enough for us to build a plug-in to perform this measurement, using mathematical formula's contributed by Dolby and other recommendations from Philips. Normalizing all signals into ICtCp-PQ-BT.2020 (one large color volume container) enables us to make a single plot for comparisons of conversion accuracy between all current SDR-HDR-WCG systems.

First, after normalizing the signals, the T/P (color) components of ITP can be plotted on an X-Y axis to check for hue shifts throughout the LUTs conversions points.

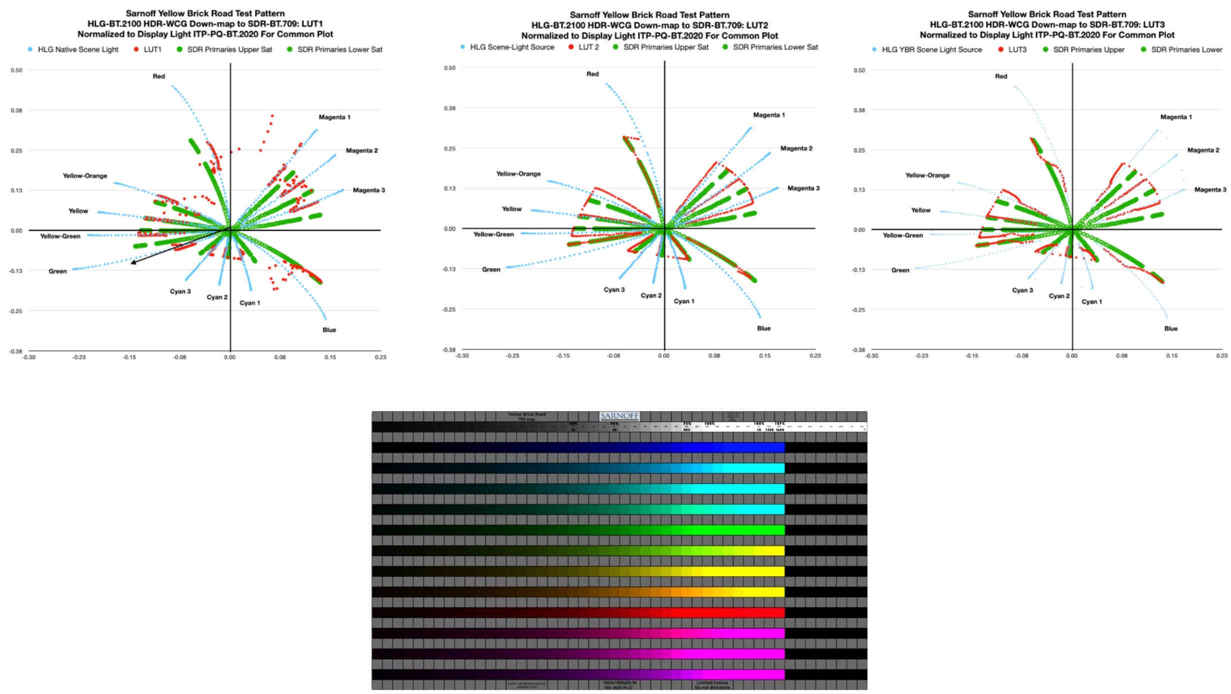


Figure 9 - LUT Conversion (devices & software): Sarnoff(SRI) “Yellow Brick Road” Pattern

Figure 9 compares three LUTs, with 576 sample points, to test up/down mapping and round-trips for color accuracy using the T/P components of ITP. All three samples are measuring the same pattern – a compilation of 576 color chips, to ascertain the ramp up in primary colors, and also the accuracy of compound colors (magenta, yellow, cyan, etc). In each sample, the blue dots represent the original(following BT.2020 primaries), the red dots represent the conversion, and. The green lines show the measurement of the SDR color primaries (BT.709) , because we want to know if the conversion is tracking the BT.2020 or BT.709 primaries so we can understand the conversion strategies.

LUT 1, on the far left, shows a conversion following the ITU-R Report BT.709 primaries for RGB for the majority of the mappings, but is extremely inaccurate with shadows and has less color detail. LUT 2, in the middle, shows excellent linear tracking of a conversion from ITU-R Recommendation BT.2020 to follow the BT.709 primaries. LUT 3, on the right, is a conversion that reproduces the BT.2020 colors within the BT.709 color triangle until the saturation/luma gets too high, then it slews toward the BT.709 primaries.

Of the three, the left-most conversion (LUT 1) is the worst of the bunch; the middle (LUT 2) is very good, and the far right is somewhat more subjective, if only because of the inevitable light and color tradeoffs that occur when converting from a larger to a smaller container (from HDR to SDR; BT.2020 to BT.709).

Figure 10 is an example of Delta-E ITP measurements for a conversion from SDR to HLG. Delta-E ITP provides a simple metric representing the difference of intensity and color between one image and another. An major benefit to the Delta-E ITP algorithm is that any unit measured above a "1" is considered a "Just Noticeable Difference," or JND. This means that humans are able to discern a difference between two object's light or color, if it measures above a 1. When calculated one against the other with the Delta-E ITP algorithm, it gives us a reliable indicator of noticeable differences, indicating accuracy issues with color and light conversions between HLG, PQ and SDR.

ITP Difference Original vs LUT Converters	
Colorchecker CSM Suite HLG 709 colors in 2020	Display Referred SDR->HLG
	ΔE-ITP
dark skin	0.6107
light skin	0.4276
blue sky	0.5727
foliage	0.4997
blue flower	0.4987
bluish green	0.3500
orange	0.9263
purplish blue	0.4047
moderate red	0.8893
purple	0.3227
yellow green	0.5079
orange yellow	0.5393
blue	0.4771
green	0.5339
red	0.7252
yellow	0.3631
magenta	0.7047
cyan	0.3159
white	0.2364
Gray 2	0.2208
Gray 3	0.2589
Gray 4	0.2547
Gray 5	0.2193
Gray 6	0.1750

Figure 10 - LUT Conversion: Delta-E ITP (SDR-to-HDR) Using Sarnoff(SRI) Color Checker 2014

Since implementing the original color metrics workflow, we've also added the ability to plot our SDR, PQ, HLG sources normalized into CIE1976 u'v' in order to examine absolute chromaticity based on recommendations from Philips in ITU.

3. Single-Stream HDR-SDR Production

The bulk of our experience in producing live content in HDR is, perhaps predictably, sports. As mentioned previously, our HDR/WCG adventures began in earnest with the 2016 Olympic Games, in Rio de Janeiro, Brazil. Subsequent HDR productions included the 2018 PyeongChang Winter Olympics; the 2018 and 2019 Notre Dame football season; the 2018 men's FIFA games; and, earlier this year, Chicago Blackhawks hockey games. Producing live content in HDR, while preserving and/or converting live feeds for SDR broadcasts, involves a blend of baseband and file-based workflows and overall orchestration.

Specifically, three different tone maps are used to manage overall brightness, for scene light, HDR display light, and SDR display light. Live HDR cameras feed into the switcher are output in native HDR. Legacy SDR camera's are up-mapped using Scene-Light so that the SDR color balance matches HLG. Other legacy pre-produced content and graphics from SDR are direct-mapped using Display-Light on the fly to preserve artistic intent.

File movement is challenging, from an orchestration perspective. Files need to be moved efficiently, converted, and transcoded and then blended with the live, native feeds.

Consistently shading HDR content is another significant component of live productions, to set levels, color balance and obviate uncomfortable light levels. Through the ITU-R, broadcasters have established recommendations for a reference white / graphics white level, at 203 nits (75% HLG; 58% PQ in narrow range.) Reference white acts as an "HDR anchor point" for all other focal points like skin tones, grass, and other images, to prevent HDR brightness levels from going too high -- while reserving enough space for highlights up to 1,000 nits for some HLG material (1,000 nits is an established normalization point for HLG to PQ conversions). Many non-HDR consumer TVs scale SDR to approximately the same peak white brightness level as the HDR anchor point (reference white at 203 nits).

For right now, most HDR productions use SDR graphics which mean they don't have to change their workflow. The SDR-to-HDR conversion LUT handles the mapping of graphics into HDR. By establishing an anchor point, HDR graphics producers could build elements and know where to map them – when looking at a scope, they can reference those known white levels. This reference white level is extremely important to making the production consistent, especially for live material. Again, the over-riding goal is to ensure that the colors and light levels delivered are consistent and aren't so uncomfortably bright as to be garish.

Figures 10 and 11 illustrate the concept of HDR shading, with no highlights (Figure 10) and highlights (Figure 11.) Highlights (like the sky in Figure 11) are typically a small part of the content. In Figure 10, an average scene from a game correlates on the scope image to roughly a 75% level, which indicates that the levels are set well. In Figure 11, a segment of the sky is in the shot, which immediately skews the corresponding scope level above 75%. Some of the scene reaches 100%, which is equivalent to 1,000 nits. The CIE1931 (an earlier version of the chromaticity measure for color gamut devised by the Commission Internationale de L'Eclairage) in Figure 10+11 maps the visible chromaticity range and shows how the current image occupies a specific color space. It is the colorful plot in the bottom middle of the scope screen capture. The plot for BT.2020 WCG imagery shows the picture information well inside the BT.709 triangle. The circle on both figures is a vector scope, which helps identify standard color targets which color bars should fall into.

HLG images are normalized to 1,000 nits as per ITU-R Recommendations during conversion to PQ so that when the signal is distributed, displays of all brightness levels will match the original intent and not be uncomfortable. Using this normalization level means that PQ's absolute mapping will preserve brightness levels at their originally-created levels seen by the production shader.

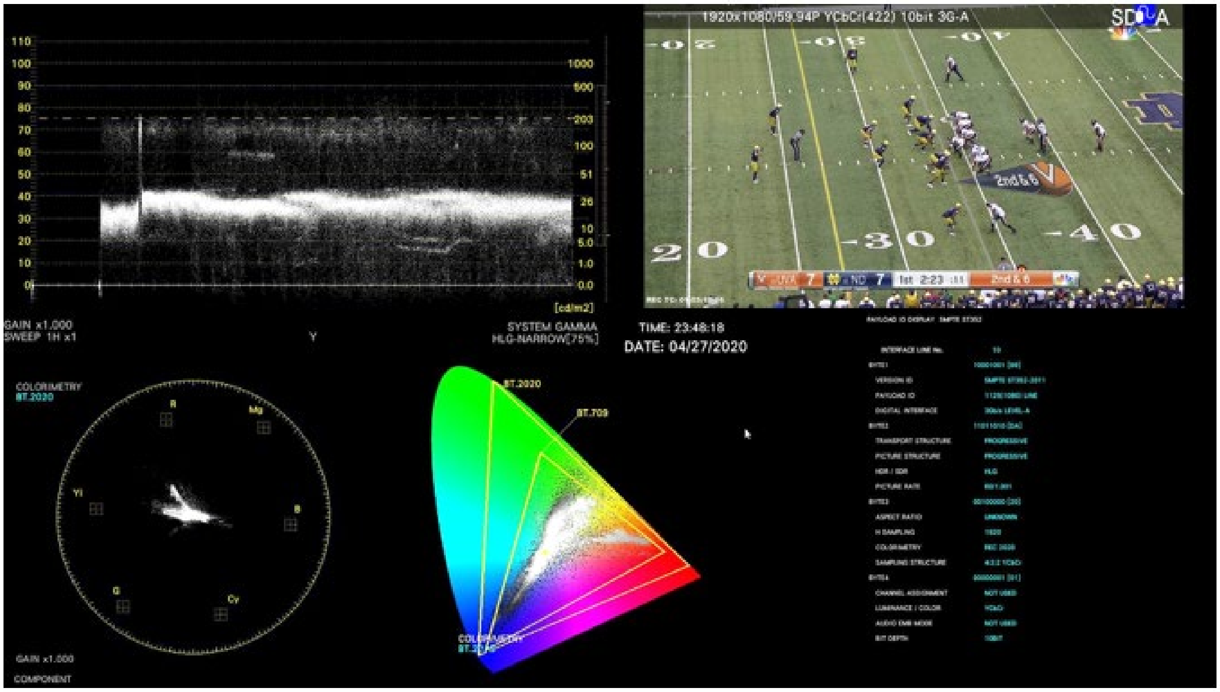


Figure 11 -: Examples of Shading: No Highlights

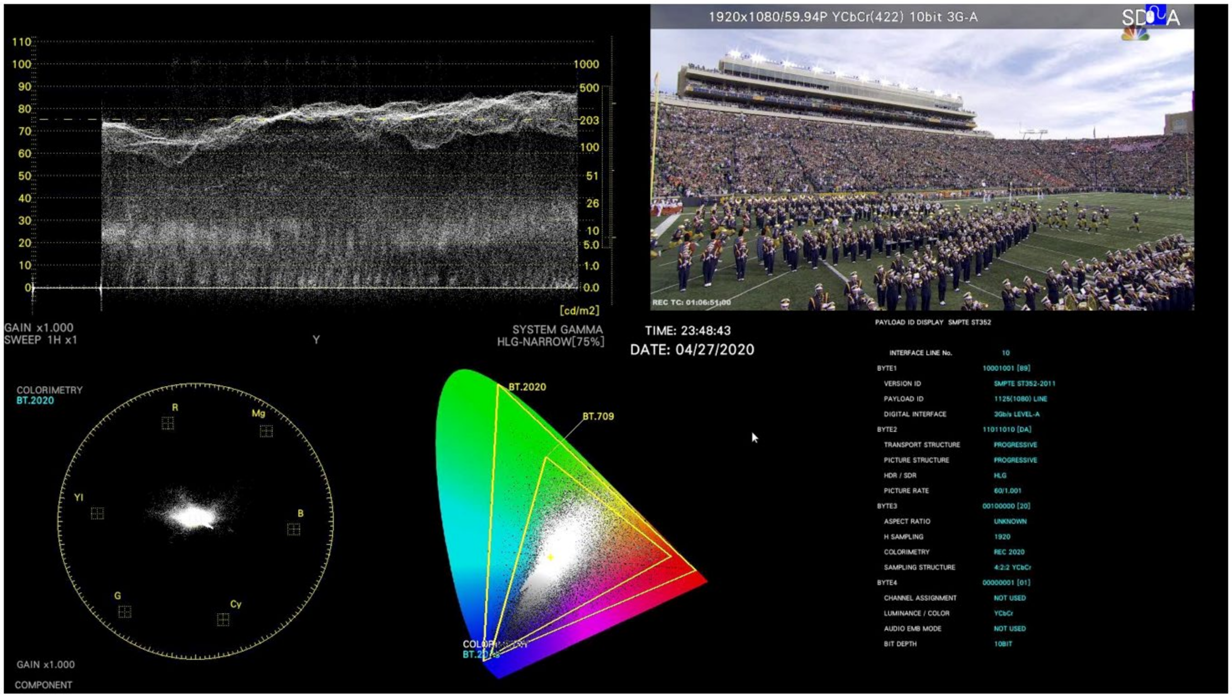


Figure 12 - Examples of Shading: Highlights

4. HDR/SDR Single-Stream Distribution

Figure 14 illustrates a simplified, single-stream HDR workflow for distribution. At the network origination point, all the production sources – SDR sources are de-interlaced, tone-mapped, scaled and color space converted. Since NBCU is agnostic to the HDR production format, native HDR formats (HLG or camera logs like S-log3) are converted to the final distribution format, PQ and resolution-scaled if necessary. Any third-party, native HDR content is also readied for transmission.

The process of cross-conversion from HLG to PQ has been defined in ITU-R Report BT.2390 (Figure 12) and produces an mathematically transparent conversion. It specifies a normalized peak white at 1,000 nit using an HLG system gamma of 1.2. Since HLG has a smaller color volume compared to PQ, we typically allow overshoots up to 109% for HLG (9% above peak-white) during conversion.

Conversion from HLG to PQ at a common peak luminance of 1 000 cd/m²

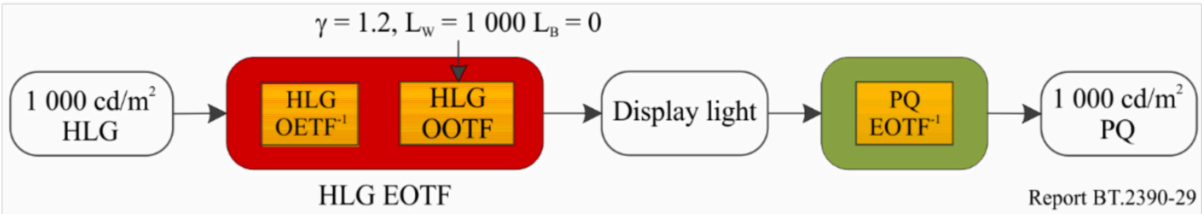


Figure 13 - This conversion produces a PQ image which is visually identical to HLG

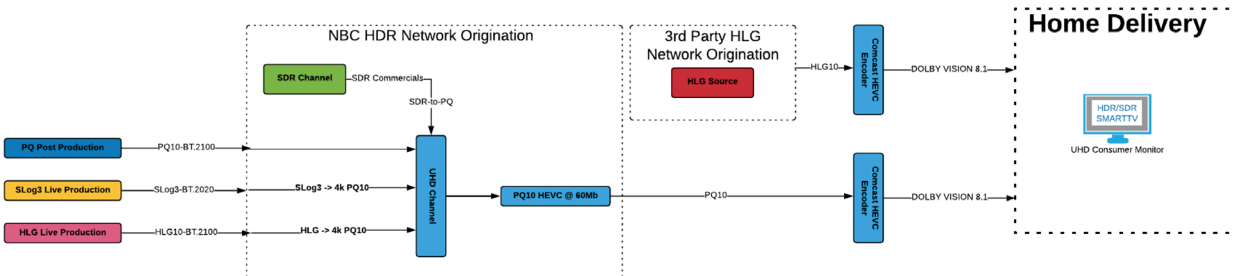


Figure 14 - A simplied view of the NBCU/Comcast production through distribution in PQ

The final distribution includes:

- Event production video: Converted from HLG/Slog3 to PQ
- Commercials: Converted from SDR to PQ
- Audio is encoded externally as Dolby ATMOS immersive audio and then embedded in the video prior to the HEVC encoder.
- Video is encoded using high efficiency video coding (HEVC) at 60-80 Mbps
- Mezzanine encode is transmitted to distribution partners (cable, satellite, and IP/OTT.).

NBCUniversal and Comcast use PQ for linear distribution because, for starters, it is the most common HDR system in-market: it's in every HDR display TV, and supported in OTT streamers like AppleTV and Roku; and, OTT providers like Netflix use it. It's used in PQ10, HDR10, HDR10+ (in each case, the "10" represents 10 bits, not the 8-bit transmissions of SDR). It can use static or dynamic color volume mapping (as mentioned earlier), which improves tone-mapping to displays with different brightness capabilities. Its conversions from other HDR formats are transparent, and normalized conversion from SDR to PQ preserves the original artistic intent.

Comcast is able to take NBCU's higher contribution rate HEVC video and re-encode it to a lower distribution rate while also adding Dolby Vision dynamic metadata to the HEVC stream. The Dolby Vision metadata provides scene or frame-based dynamic tone-mapping which provides benefits when the consumer display has different capabilities compared to the source. As televisions peak brightness capabilities improve, the content producers may choose to improve the content as well but initially not all of the more affordable consumer displays will support the superior capabilities. Using dynamic metadata can improve tone-mapping from content to display which much of the consumer ecosystem can benefit from.

An additional benefit of mapping an HLG production to PQ for distribution is that we can take advantage of PQ's support of static or dynamic metadata for improved tone-mapping on less expensive displays that don't support luminance at or above 1,000nits. Finally, many HLG consumer displays clip above nominal video levels (100 IRE), but we can preserve levels above HLG-100IRE during the conversion because PQ supports up to 10,000nits(10x that of HLG). The IRE unit is used in the measurement of composite video signals.

In live-linear broadcast distribution, PQ, with proper shading, means that as displays get brighter, the images will remain at the original brightness unless the consumer manually changes settings. The original artistic intent is preserved and does not suffer from some other perceptual effects referenced in section 2.1 (Hunt, Stevens, etc.). Since some cinema or pre-produced content are currently shaded at levels up to 4,000 nits, relative systems like HLG, could be restrictive as displays improve. In short, for live-linear distribution, PQ's absolute mapping, combined with the productions white level anchor point, make beautiful imagery and preserves artistic intent all the way from the production to the consumers HDR display.

5. Conclusion

Producing and delivering HDR content for live linear workflows, like sports, is considerably more complex than it is for most file-based, VOD-styled approaches, which have post production resources that simply don't exist for live, on-the-fly material. Live productions must handle graphics, color correction, and luminance normalization in real-time, across what is usually a mix of SDR (which is essentially contemporary HDTV) and HDR cameras.

This paper details differences in system colorimetry which exist within the primary HDR systems, and specifically HLG) and PQ. It describes functional workflows for live delivery, including what are multiple levels of conversion for various signals throughout the infrastructure -- like graphics, video playback, SDR and native HDR cameras.

It has covered how format conversion and perceptual measurement work for HDR and SDR productions, how HDR production workflows are orchestrated for live content, and how those live linear workflows are distributed in PQ. Most contemporary broadcasters and cable providers are following a similar, if not identical, path: Producing in HLG or other camera-logs and distributing in PQ.

This paper also explained what it takes to build a functional, single-stream HDR-SDR workflow, at each stage of the pipeline, without compromising the artistic intent or quality of the distribution paths for either the HDR or the primary SDR feeds. In-depth research of color conversions were discussed and explored, as well as production layouts.

Finally, distribution requirements, delivery infrastructure and technical specifications were described and explored. In particular, we described why production environments tend to prefer to use the HLG transfer function, while distribution functions favor PQ.

Abbreviations

BBC	British Broadcasting Corporation
CIE	Commission Internationale de L'Eclairage
CRT	cathode ray tube
DLC	display light conversion
EOTF	electrical-optical transfer function
FIFA	Federation Internationale de Football
HDR	high dynamic range
HEVC	high efficiency video coding
HLG	Hybrid-Log Gamma
HVS	human visual system
IP	Internet protocol
IRE	Institute of Radio Engineers
JND	just noticeable difference
LUT	look-up table
NHK	Nippon Hoso Kyokai (Japanese Broadcasting Association)
OETF	optical-electrical transfer function
OTT	over-the-top
PQ	perceptual quantization
RGB	red, blue, green
SDI	serial digital interface
SDR	Standard dynamic range
WCG	wide color gamut

Bibliography & References

ITU: ITU-R Recommendation BT.709-6 (2015), "Parameter values for the HDTV standards for production and international programme exchange ," International Telecommunications Union, Geneva

ITU: ITU-R Recommendation BT.2020-2 (2015), "Parameter values for ultra-high definition television systems for production and international programme exchange," International Telecommunications Union, Geneva

ITU: ITU-R Recommendation BT.2100-2 (2018), "Image parameter values for high dynamic range television for use in production and international programme exchange," International Telecommunications Union, Geneva

ITU: ITU-R Recommendation BT.2124-0 (2019), “Objective metric for the assessment of the potential visibility of colour differences in television,” International Telecommunications Union, Geneva

ITU: ITU-R Report BT.2390-8 (2020), “High dynamic range television for production and international programme exchange,” International Telecommunications Union, Geneva

ITU: ITU-R Report BT.2408-3 (2019), “Operational practices in HDR television production,” International Telecommunications Union, Geneva

ITU: H-series Recommendations–Supplement19, “Usage of video signal type code points,” International Telecommunications Union, Geneva.

Min, H., Kim, T., & Weyrich, JK. (2009). Modeling human color perception under extended luminance levels. *Association for Computing Machinery Transactions on Graphics, Vol. 28, No. 3*. http://reality.cs.ucl.ac.uk/projects/xlrcam/kim09xlrcam.pdf

vooya video sequence player and YccIccSee color plugin. https://www.offminor.de/

