

HDR

An introduction to color management for motion-picture, television and games production, including visual effects, animation. It expands on the original's focus on motion-picture visual effects and animation as a reflection of the expanded scope of digital imaging in the film and television production process, from capture to post-production to grading and final display, and as a reflection of the consistency in basic approaches used by visual effects, animation and increasingly games projects.

Color impacts many areas of the movie and the game-making process. From on-set capture to the many phases of the computer graphics pipeline: texture painting, look development, lighting, rendering, and compositing, to grading and finishing for the theater and home. Handling color is a tricky problem. We will be further exploring about the color science, measuring and encoding color and core concepts like gamuts, transfer functions, and scene-referred and output-referred colorimetry. These concepts are being used in modern motion-picture and games production, also including an introduction to efforts on digital color standardization in the motion-picture industry, ACES and ASC CDL, and how to experiment with all of these concepts for free using open-source software like OpenColorIO and Color Science for Python.

Topics Covered

Includes Color Science and Workflow, focussing on specific production workflows or challenges referring to Color Science with examples and reference images.

The Goal

Professionals in the motion picture, visual effects, animation, and games industries encounter color management challenges which are not covered in either traditional color science textbooks or online resources, leaving digital artists and computer graphics developers to fend for themselves.

Best practices are too often maintained through tribal knowledge: passed along by word of mouth, user forums, or scripts copied between facilities. This document's goal is to provide a better mechanism for communicating this information. It also outlines the color pipeline challenges in modern feature film, visual effects, animation and games production, from on-set capture to the computer graphics pipeline to digital intermediate (DI) grading and final display. It presents techniques and best practices currently in use at major production facilities and notes the areas that need further development and standardization. One note to begin with: Color pipelines are not static nor is there a single best approach. They grow and evolve with the constraints of each project. This paper intends to present color measurement, processing, and pipelines in enough generality that the reader will be able to address the specific needs of their productions, even if they are not described explicitly in the text. It is unlikely that this paper will address the exact requirements of any given production in whole.

Journey towards HDR

Each image is generated using a roughly consistent set of stages and processing steps. The succession of image states is as follows: 1) Camera image, in camera-native color space 2) Ungraded camera image or renderer output in an ungraded working color space 3) Image with generic grade 4) Final graded image for primary output with composited background and foreground elements and 5) graded image for an alternate output like PQ or DCDM.

There has been a convergence of approaches across cinematography, visual effects, animation, games, finishing, and grading. Shared challenges in the evolution of display and capture technology: a shift from photochemical to digital image capture, widely available wide-gamut high-dynamic range (HDR) displays, advances in rendering and image generation research and the increased integration of different production departments have driven this harmonization. Each production will define a slightly different overall color pipeline, but most discussion, debate, and variation boils down to choices made around the transforms and formats used. Image and color data with specific meaning, referring to scene or display intensities, and transforms, with specific expectations for input and output format, are concepts that are used remarkably consistently across on-set capture, visual effects, games, finishing, grading and software and hardware color processing pipelines.

History

In February and April 1990, Georges Cornu  jols introduced the first real-time HDR camera combining two successively or simultaneously - captured images.

In 1991, the first commercial video camera using consumer-grade sensors and cameras was introduced that performed real-time capturing of multiple images with different exposures, and producing an HDR video image, by Hymatom, licensee of Cornu  jols.

Also in 1991, Cornu  jols introduced the principle of non linear image accumulation HDR+ to increase the camera sensitivity in low-light environments, several successive images are accumulated, increasing the signal-to-noise ratio.

Later, in the early 2000s, several scholarly research efforts used consumer-grade sensors and cameras. A few companies such as RED and Arri have been developing digital sensors capable of a higher dynamic range. RED EPIC-X can capture time-sequential HDRx images with a user-selectable 1–3 stops of additional highlight latitude in the "x" channel. The "x" channel can be merged with the normal channel in post production software. The Arri Alexa camera uses a dual gain architecture to generate an HDR image from two exposures captured at the same time.

Modern movies have often been filmed with cameras featuring a higher dynamic range, and legacy movies can be converted even if manual intervention would be needed for some frames (as when black-and-white films are converted to color). Also, special effects, especially those that mix real and synthetic footage, require both HDR shooting and rendering. HDR video is

also needed in applications that demand high accuracy for capturing temporal aspects of changes in the scene. This is important in monitoring of some industrial processes such as welding, in predictive driver assistance systems in automotive industry, in surveillance video systems, and other applications. HDR video can be also considered to speed image acquisition in applications that need a large number of static HDR images are, for example in image-based methods in computer graphics.

OpenEXR was created in 1999 by Industrial Light & Magic (ILM) and released in 2003 as an open source software library. OpenEXR is used for film and television production.

What is HDR? Why does it matter? Will it last? And how should we handle it?

We've been hearing about 'HDR' since 2014, and it's clearly an important and much debated development in the industry. Rather confusingly, HDR has come to mean three quite different things:

1. The use of multiple exposures merged together to capture a wider dynamic range than is usually possible. This is the meaning used in photography.
2. Working in a scene linear environment that allows accurate modelling of physical light. This is the meaning used in Visual Effects.
3. The use of a display with significantly brighter whites and, usually, deeper blacks to display a higher dynamic range than the current standards (rec 709 and DCI P3).

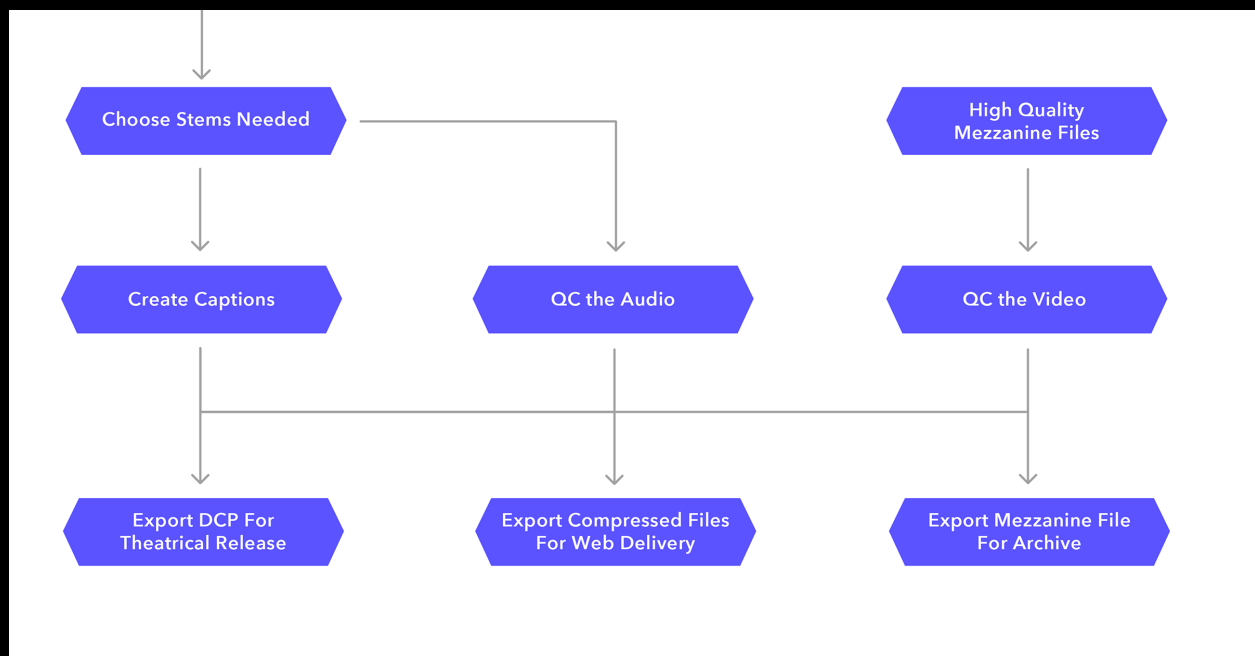
Clearly, some clarity is needed. The first definition – HDR capture – is also known as Photo HDR or HDR imaging (Hdri); this article won't discuss that. Instead, we are going to focus on HDR displays – in particular, why they are here, how they are different and why they are better.

Terms like HDR are relative and often change as technology advances. For example, "HD" has frequently been used to describe new standards. When the old black and white 405 line broadcasts were replaced by 625 line color, the new technology was described as "high definition", but it has of course been used for many more recent shifts.

HDR, or High Dynamic Range, is the newest technology to get the best image on screen, regardless of whether you are working in HD or 4K/UltraHD. It is an increase in dynamic range and luminance values and wide color gamut or WCG, meaning richer colors across the board. WCG is defined by the P3 or BT.2020 standards for color reach. SDR, or Standard Dynamic Range, can have 6 stops of dynamic range, however HDR offers up to 12 stops more of dynamic range for a total of 16/18 stops. That added dynamic range offers a wider reach into shadows and highlights in terms of luminance detail and a real pop to the color range with the ability to show true greens and magnificent golds.

But before discovering exactly what HDR is, there are some other persistent myths to dispense with.

- HDR isn't a fad. It is a technology shift and it is here to stay
- HDR is not painful or tiring to watch. This is a subjective statement, like saying that 5.1 audio is too loud. It might be true in individual cases, but there will always be good and bad examples.
- There is no single standard for HDR and probably never will be. HD doesn't have a single standard either so there's nothing new in that.
- Increased brightness on its own doesn't equate to HDR. For example, a BT.709 image displayed at 1000 nits would not be HDR, and would look terrible. HDR is an entirely different format, not just a display technology.
- HDR isn't dependent on 4k resolution, wider color gamut or any other improvements that are bundled with it. The requirements for HDR are images with a minimum 10 bits, and a display with enough dynamic range that can handle HDR formats.



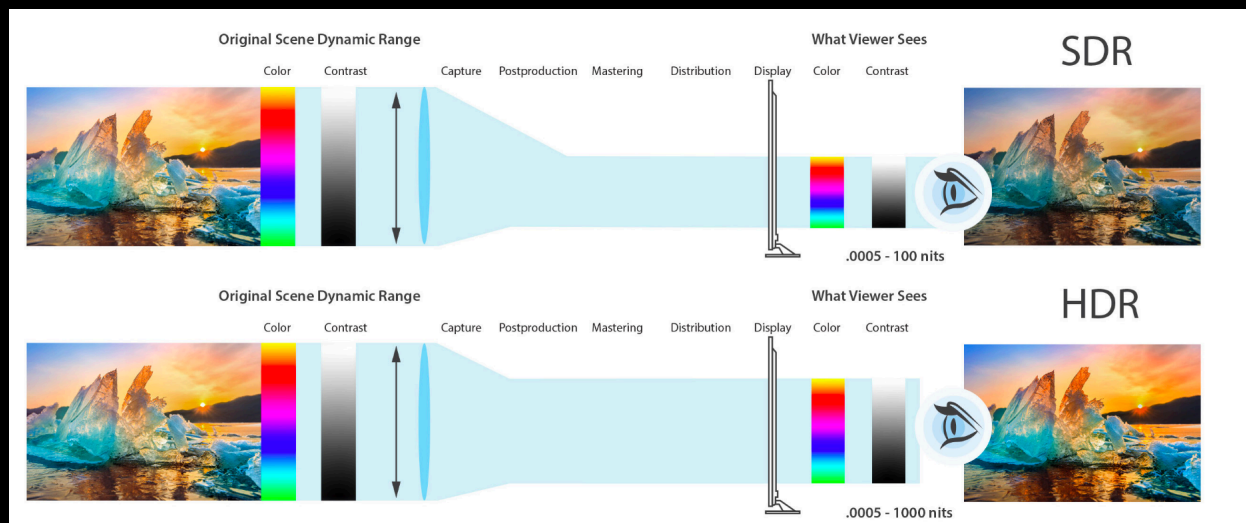
Why HDR is appealing

HDR is a technology that allows us to get closer to the Human Visual System (HVS). Our eyes do a much better job at perceiving the dynamic range in both bright and dark lighting situations, with greater detail in the shadows and highlights than video can. HDR is an effort to increase the intensity and range of that detail as well as provide a richer range of shades for colors displayed by a screen. That increase in both leads to a better sense of texture.

In other words, the closer a display's dynamic range and color gamut can get to the human visual system, the better a pixel will be perceived to be. That's why the addition of high dynamic range (HDR) and wide color gamut (WCG) to 4K UHD production and displays is regarded as two of the most important enhancements to television since HDTV. And as many are finding, it's an easily perceivable quality jump for HD itself.

With HDR, previously undetectable picture details in a shadowy area of an image are revealed, and a bright noonday sun doesn't bloom. For photographers and videographers, that's the difference between six stops in standard dynamic range images and 17.6 stops in their HDR counterparts.

Add to that a wide color gamut that skyrockets the number of colors that can be reproduced from mere millions to a billion or more, and the result is a video image that approaches the higher end of what the human visual system can perceive with pixels that are clearly better. While it is still early days for HDR and WCG, it's not too soon for film makers to begin understanding more about the technologies that make them possible and to start planning for how they will introduce them into their existing production workflows.



The number of colors and the dynamic range of the images that viewers see at home on an HDR display more closely approximate those of the original scene than those reproduced on an SDR monitor.

HDR is here to stay

HDR did not come about because the film and television industry demanded it. There is a wider market that drove its adoption and now it's unlikely that suppliers will want to continue to manufacture both SDR and HDR displays. There are plenty of similar historic circumstances where one technology won out.

- 1990s - CRTs were replaced with LEDs, plasmas and technology that wasn't as good for television. Why? Because CRTs are bulky and won't fit in a laptop. We have persevered to maintain standards and today's screens are excellent, but that was not always the case.
- 2012 - 4k screens were introduced, despite few in our industry advocating them as superior. There is justification for 4k capture – but much less for the deliverable. Even the consumer press acknowledges that there is little benefit for domestic use. However, the technology is needed by the computer and games industries to put bigger GUIs and more windows on the screen. So 4k is here to stay and again we have adapted to it.
- 2016 - Screens are getting brighter and brighter, to the point that ordinary BT.709 content, which is mastered for 100 nits inherited from the characteristics of the CRT, doesn't look right. But the phone and tablet industries need brighter screens to make their displays readable in daylight and in areas with high ambient lighting. In addition, research suggests that HDR displays are preferred by all audiences - a preference not shared for stereoscopic and 4k. Manufacturers will eventually stick with the brighter screens and not offer anything as low as 100 nits (most TVs are brighter than that these days anyway)

The adoption of HD required upgrades throughout post-production, but in many ways the workflow is the same as that of SD. The HD specifications were mostly based on SD. There was an opportunity to adopt a single standard in the transition, but the reality was that we gained many more standards.

The shift to color on the other hand was less about the technology and more about aesthetics. B/W content could remain B/W and be appreciated for that, but audiences globally preferred color and as soon as it was technologically possible that became the norm. The shift caused dramatic changes to costumes, makeup and of course color-grading amongst other things. The change from B/W to color was more creative than technical and so it is with HDR. Both are made possible by technology, but focusing on the hardware misses the point.

In fact, the biggest problem is trying to demonstrably explain the concept of HDR. No words can do it, nor can diagrams or images on a display that is not HDR. It is like trying to explain why a color photograph is better than the black and white version. Sometimes the black and white might look nicer, but usually people will say that the color image is more life-like and has more detail. Even showing HDR and SDR on the same display, or on side-by-side displays can be confusing, because the human eye can't adjust for both at the same time. The best way to understand HDR is to look at a perfectly good SDR image on an SDR display and then see a properly mastered HDR version on an HDR display in the same viewing environment.

If you are a colorist with access to the appropriate monitoring, you can take this experiment one step further. Grade the SDR version as you normally would, then start again and grade it on the HDR display at a minimum of 1000 nits.

The problem with comparisons is that the SDR display on its own can look great....until you put an HDR display next to it. Then it appears dull and flat because the human eye adjusts to the increased dynamic range.

Color Processing

Digital color management follows a clear general path tightly coupled with digital imaging from reflectance, image formation, storage and processing to display.

While there are many imaging considerations along the full imaging pipeline, such as spatial and temporal considerations, this discussion restricted itself to the propagation and transformation of color in the imaging pipeline.

The simplest example of the digital color pipeline for cinema is exemplified by a digital camera, covering the full imaging pipeline from lens to sensor to the display on the back of the camera. Even this simplest case involves processing and transforms that are a microcosm of more complex pipelines.

From this example, it is possible to expand to cover pipelines and workflows that have processing steps in-camera, on-set, in a scanning house, with editors, in a renderer, visual effects, animation and games teams, colorists, distribution processing, cinemas, streaming services and everywhere in between. In this section, we will present simple models for thinking about color processing and understanding which aspects are important in each phase of production.

In the simple case mentioned above, an image captured by a digital camera and displayed on the screen on the back of the camera, there are a series of states and transforms to consider: Light reflected off a surface, into camera's lens and onto an image sensor. Sensor photosite values encoded and stored in memory. A transform to combine adjacent Red, Green and Blue photosites into single pixel values. A "working space" image in memory. A transform to remap the intensities of the combined pixels for display. A "displayable" image ready to be sent to the display hardware showing the image.

The aspects of the example above that will show up repeatedly throughout this document are:

Input, Working, Output: The example includes the stages that are common to most modern color pipelines, namely an Input, Capture or Generated image state, a Working image state used for editing and production and an Output image state used for display. There may be other steps along the way, but virtually all modern color pipeline designs contain these three distinct states. The Input image for a visual effects or animated feature or game may be the product of a renderer.

Transforms - Input, Output, Look: Going between each image state is a transformation that has a specific, meaningful effect. The most widely used transforms will typically include a remapping of intensities and some degree of remapping of the color gamut: the underlying

meaning of the red, green and blue of each pixel. There are many more exotic transforms, and often a number of these simpler transforms are concatenated. Transforms going from Input to the Working space are usually strictly defined based on the capabilities of the capture device and the attributes of the working space. Transforms from the Working space to the Output space usually combine an element of artistry, commonly referred to as "the Look", with an explicit mapping based on the capabilities of the output device. Later sections of the document contain many examples of transforms used in practice.

Staged Processing: The process to go from light hitting the sensor to a display emitting light is defined by a specific, staged processing pipeline. Transforms take color from one image state to the next, constrained by the processing needs of each successive stage and the limitations and capabilities of the capture, processing and display hardware. Images may be written to disk or passed along an IO line representing the state of color at any of the points in the processing chain, as different applications or departments within production may be the source or the destination for the data.

Meaning and Placement - Understanding how to work with an image or color data set is more natural when you can be clear about what the data is supposed to represent and where it lives in the production or product's processing chain. These two elements, meaning and placement, provide the context needed to apply transformations and make aesthetic judgements in production.

Introduction to color spaces

It seems like the term color space is thrown around more than ever these days, but still many of us (even seasoned post professionals) are a bit fuzzy on what it specifically refers to, and the role it plays in the acquisition and manipulation of video.

But why has "color space" become so prevalent on set and in our post pipelines, and why is the knowledge gap among professionals so wide?

This is a byproduct of the explosive advancements of digital cameras, displays, and workflows we've seen in the last 10-15 years. Virtually every professional-grade camera out there now lets users choose between multiple capture color spaces, which makes it an upfront concern for production teams.

On the post-production side, intermediate color spaces have to be selected to make sure the exchange of footage, VFX, and graphics assets between facilities and teams runs smoothly. Additionally, with HDR and other next-gen display formats on the rise, we're faced with even more choices when it comes time for delivery.

To add even more fuel to the fire, it's becoming increasingly necessary to choose the color space during pre-production, at the same time as the codecs, resolutions, and other technical image specifications are decided, as it's a key detail necessary to plotting out a project's

workflow. That means more people than ever before are concerned with color spaces, and need to understand what they are and how they should be used. So how do we choose between color spaces? What makes one better or worse for our content and/or workflow? And what are the hidden opportunities and hazards hiding behind each of these technical choices?

So, we'll explore these questions and definitively explain what a color space is. Along the way, we'll discover why your choice of color space may be the single most important factor in capturing and delivering beautiful motion images.

What is color space?

A color space describes a specific, measurable, and fixed range of possible colors and luminance values. Its most basic practical function is to describe the capabilities of a capture or display device to reproduce color information.

For example, if I have a camera that shoots in an arbitrary "Color Space A," anything I capture which has color and/or luminance values beyond what that space can define will not be accurately captured. These "out of bounds" values can be dealt with in different ways (some more visually pleasing than others), but they can not be accurately captured as they appear in the real world.

The same is true of a display device designed to reproduce "Color Space B" — any color or luminance value stored in a digital image that falls outside "Color Space B" cannot be accurately shown on that particular screen. In both cases, larger color spaces mean a greater range of colors can be accurately captured and/or reproduced.

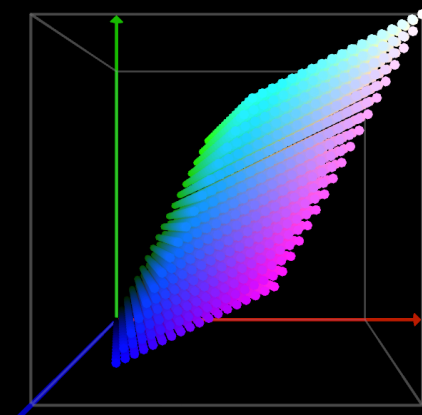
Furthermore, the only good way to ensure accurate reproduction of a scene captured in Color Space A and rendered in Color Space B is to be aware of the discrepancy between capture and delivery space, and to mathematically transform the signal from one color space to the other. Providing a means of ensuring color accuracy from capture to display is the other fundamental function of explicitly defined color spaces.

Some color spaces whose names you may already be familiar with include

Rec. 709, BT.1886, sRGB, Rec. 2020, ST 2084 / PQ Hybrid Log Gamma (HLG), P3 DCI, P3 D65, X'Y'Z', Arri ALEXA LogC, Blackmagic Film, Canon CLog, GoPro ProTune, Panasonic V-Log, RED Log3G10 or RedWideGamutRGB, Sony SLog, ACES2065-1, ACEScg, ACEScc, ACEScct, Cineon

When viewed as 3D plots, the concept of color "space" begins to make more intuitive sense.

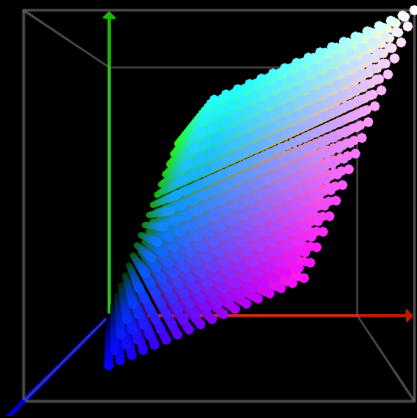
Ok, so we now know that a color space is an explicitly defined range of colors and luminances. But we also know that no two humans see things the same way—so how do you explicitly define something as slippery as



REC-709-COLOR-SPACE-3D-PLOT

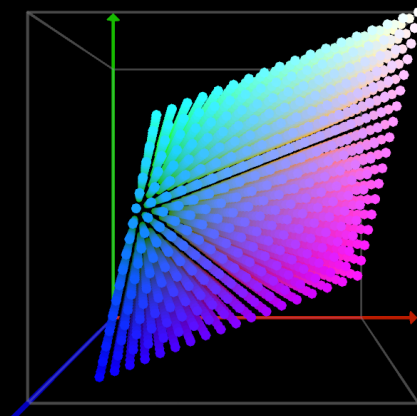
color? The maddening answer is that a color space can only be defined in reference to another color space.

So where does the madness end? Better yet, where does it begin? Is there a foundational color space against which all others can be defined? Thankfully, there is.



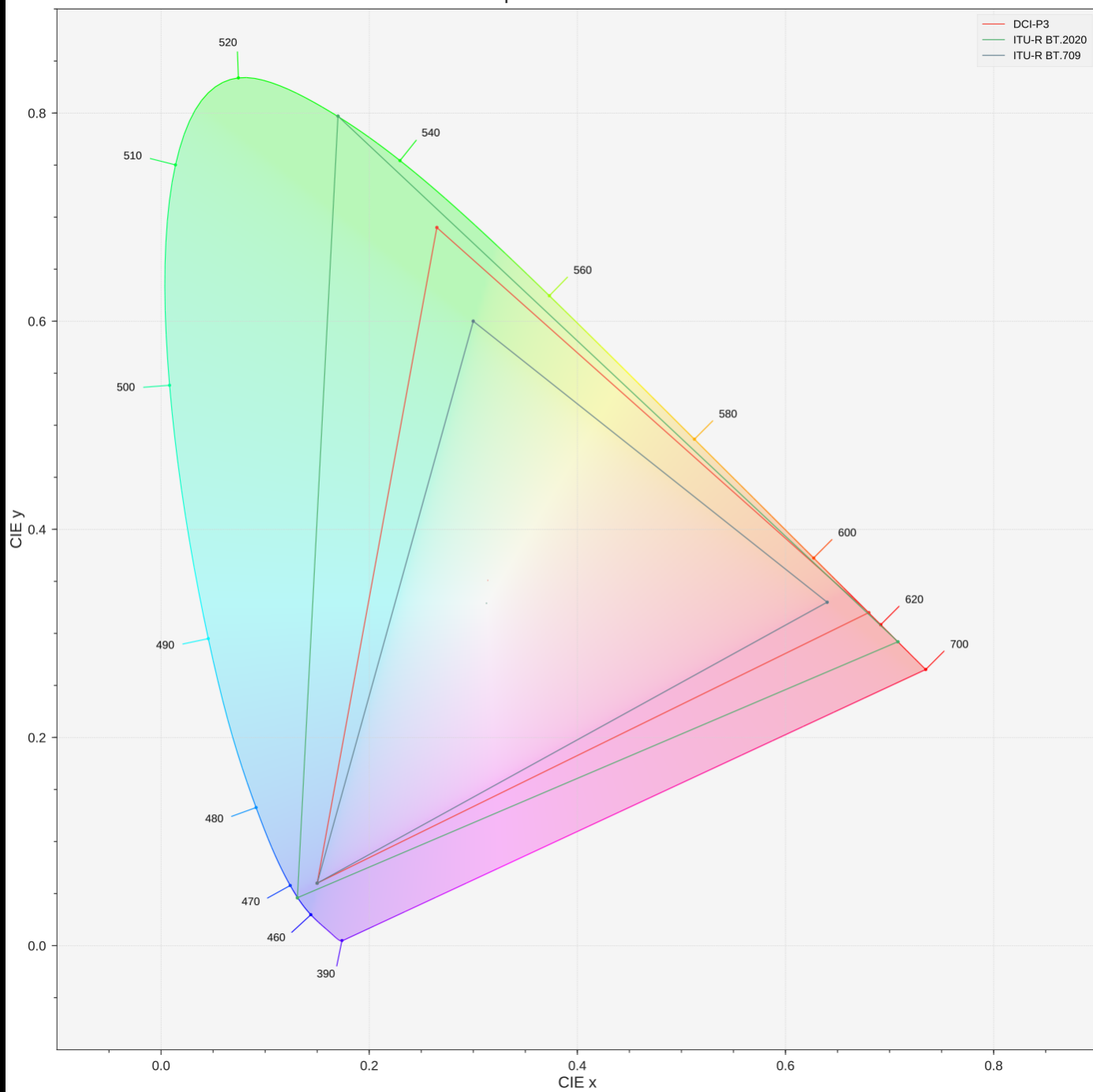
P3-COLOR-SPACE-3D-PLOT

In 1931, the International Commission on Illumination, or the CIE (from the French “Commission internationale de l’éclairage”) defined an all-encompassing color space based on human perception, using averaged data from experiments conducted with a small set of test subjects. Nearly a century later, this space, CIE 1931, remains the standard reference used to describe all other color spaces.

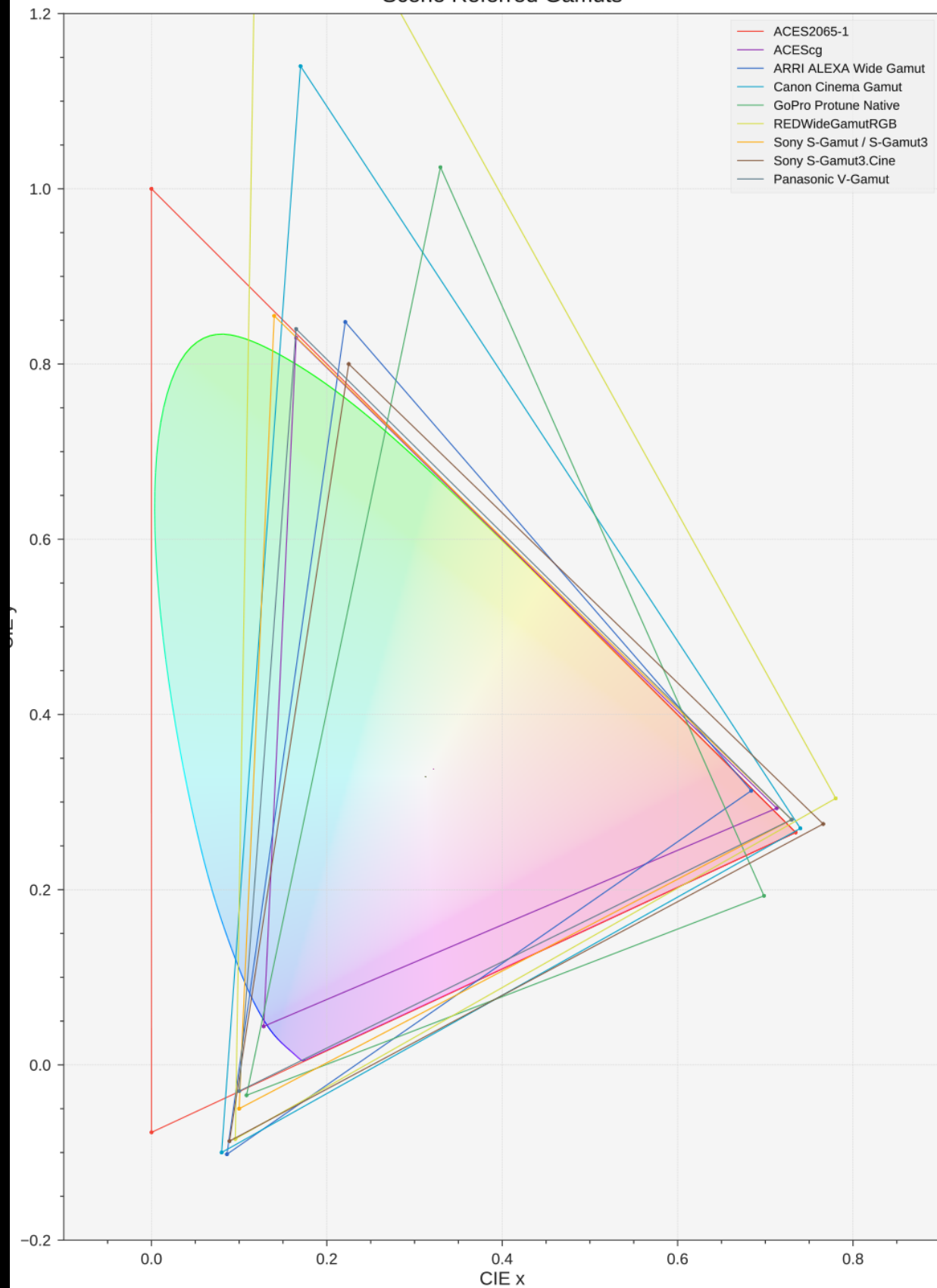


REC-2020-COLOR-SPACE-3D-PLOT

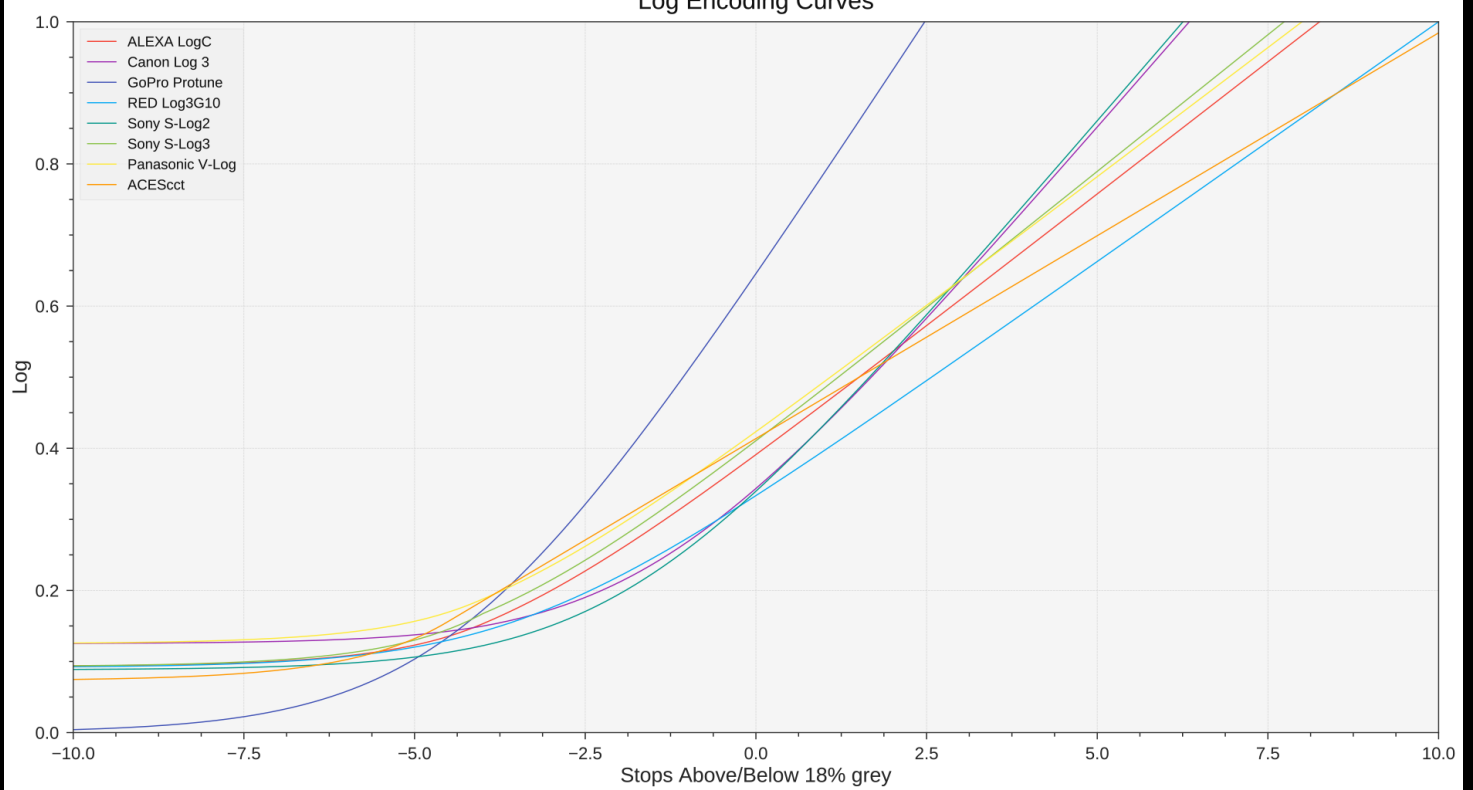
Output Referred Gamuts

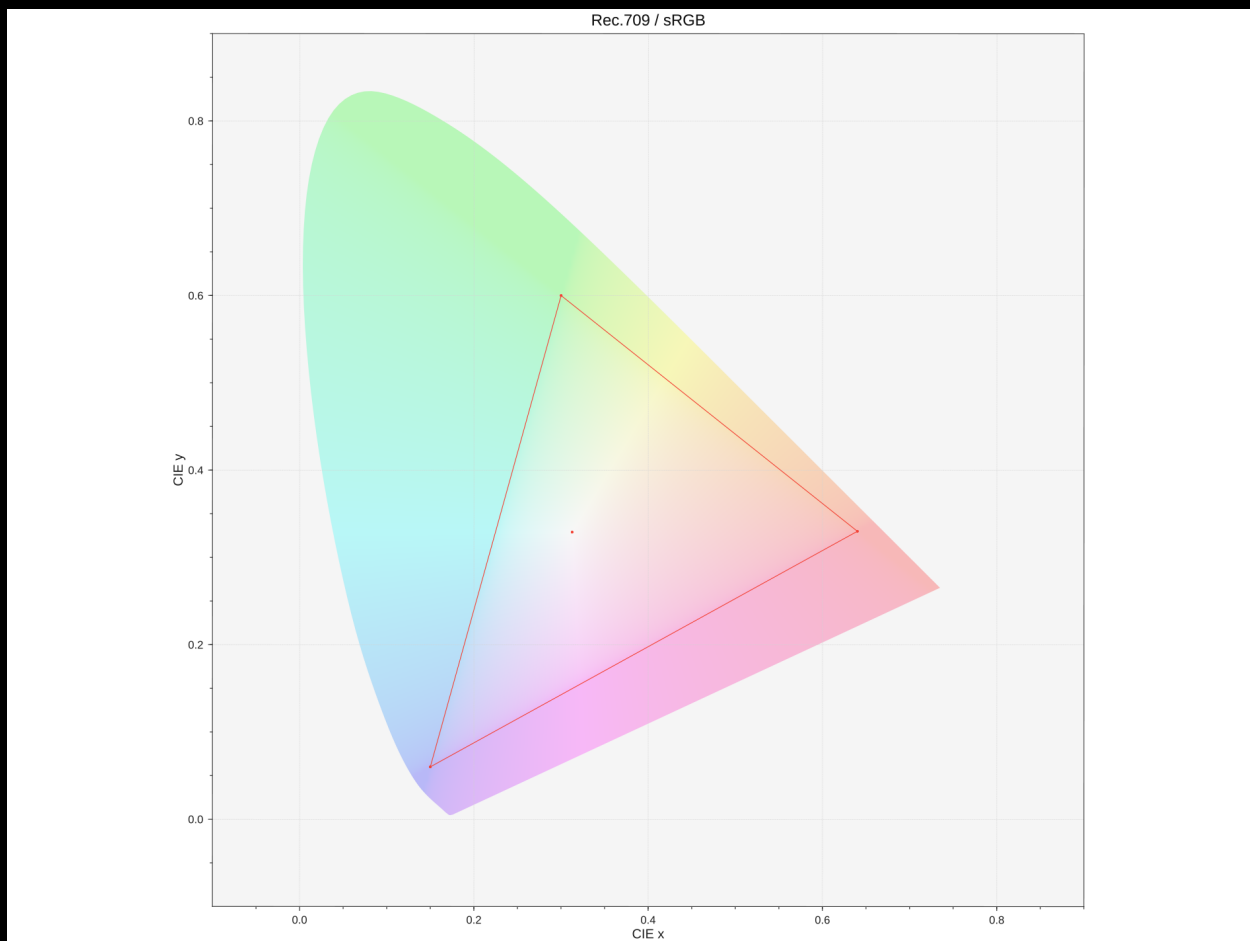


Scene Referred Gamuts



Log Encoding Curves





The ACES Still Life reference image encoded with the Rec. 709 (BT.1886) Output Transform Copyright © A.M.P.A.S

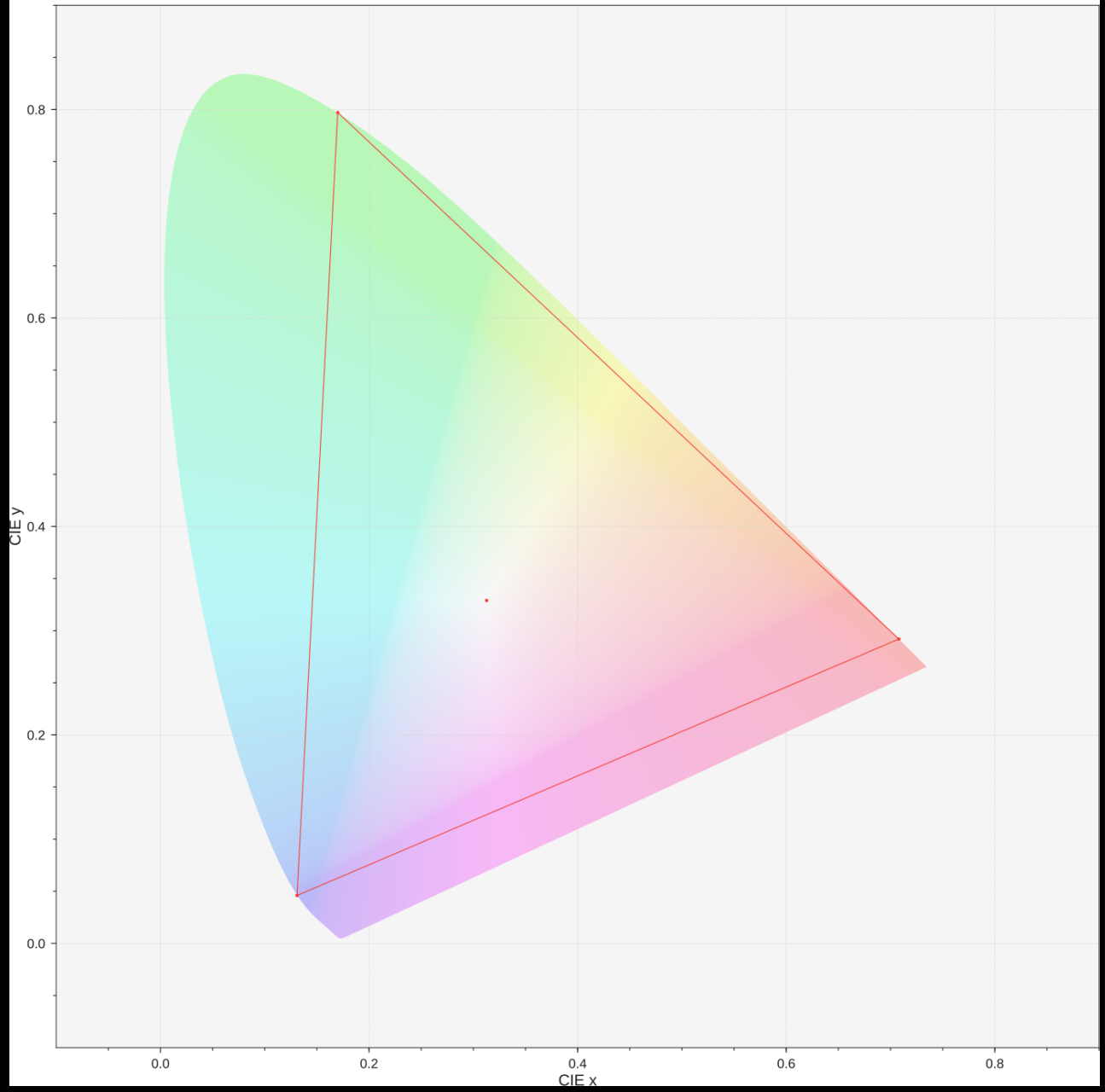




The ACES Still Life reference image encoded with the sRGB Output Transform.
Copyright © A.M.P.A.S



The ACES Still Life reference image encoded with the Rec. 2020 100 nits Output Transform.
Copyright © A.M.P.A.S

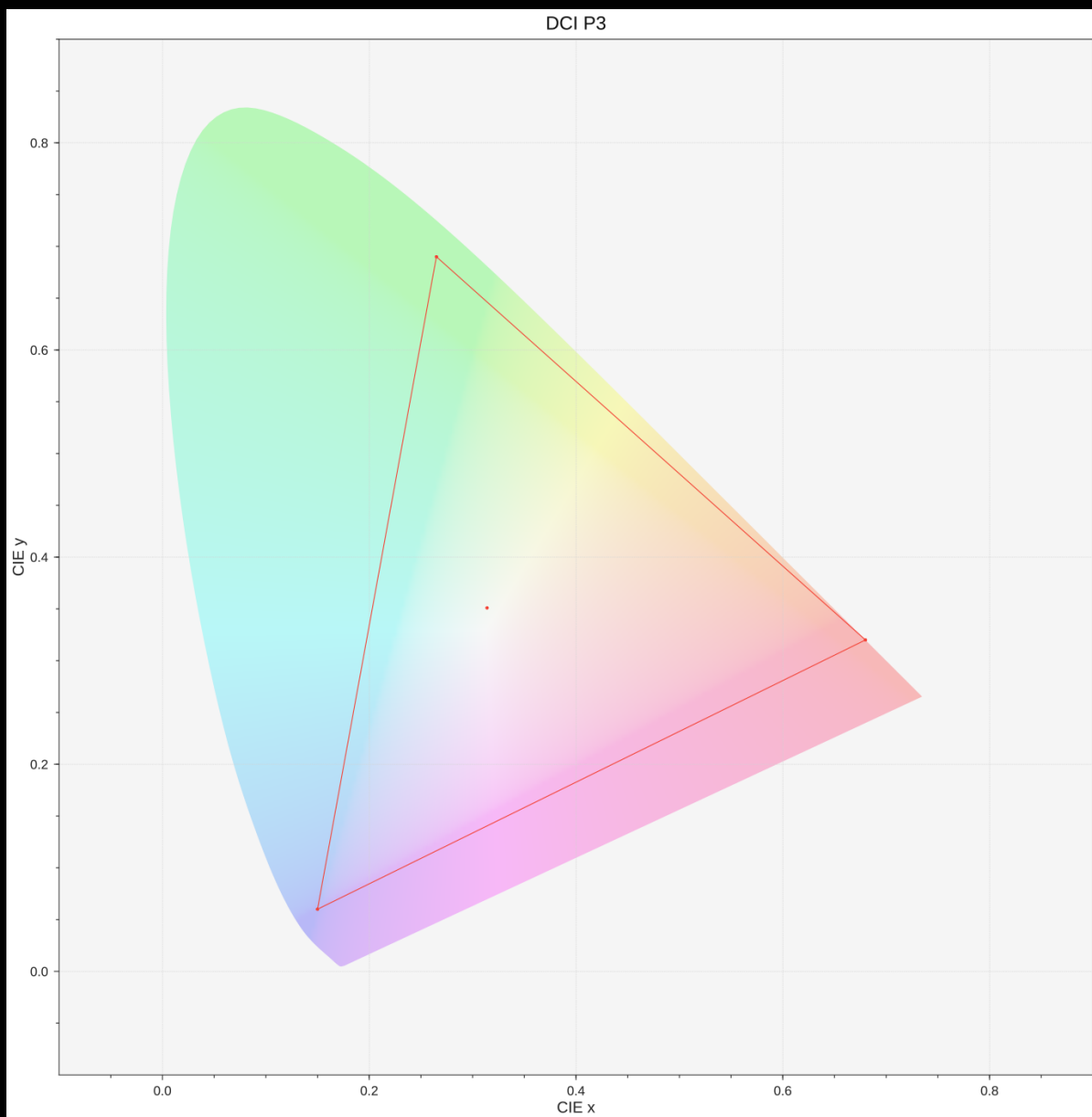




The ACES Still Life reference image encoded with the HLG Output Transform. Notice that the image looks “normal”, if somewhat muted on an SDR display. Copyright © A.M.P.A.S



The ACES Still Life reference image encoded with the P3 DCI 48 nits Output Transform. Note that this output transform includes a “D60 sim”, giving the image a slight magenta tint when viewed on an sRGB (D65) display. On the intended P3 DCI display neutrals would appear with ACES (-D60) white Copyright © A.M.P.A.S





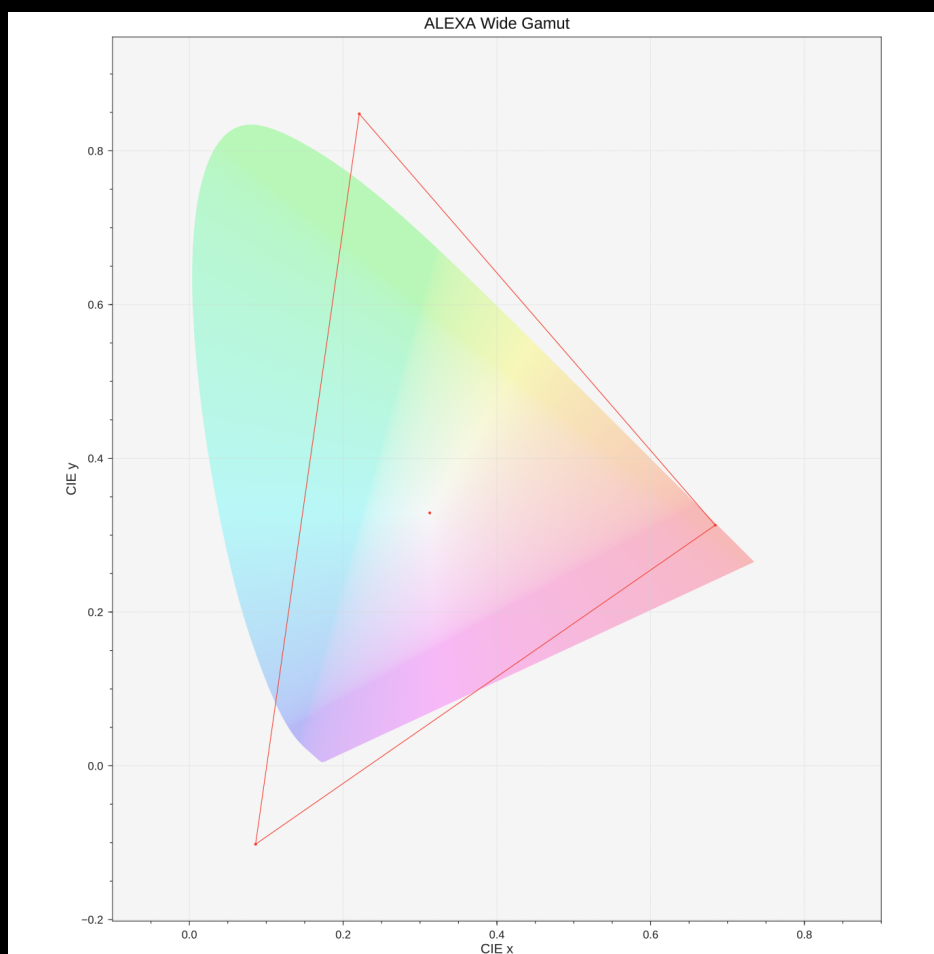
The ACES Still Life reference image encoded with the P3 D65 48 nits Output Transform.
Copyright © A.M.P.A.S



The ACES Still Life reference image encoded with the DCDM Output Transform.
On an sRGB display it is low contrast (due to the 2.6 gamma) and the colors appear incorrect, since $X'Y'Z'$ values are being interpreted as if they were $R'G'B'$ Copyright © A.M.P.A.S

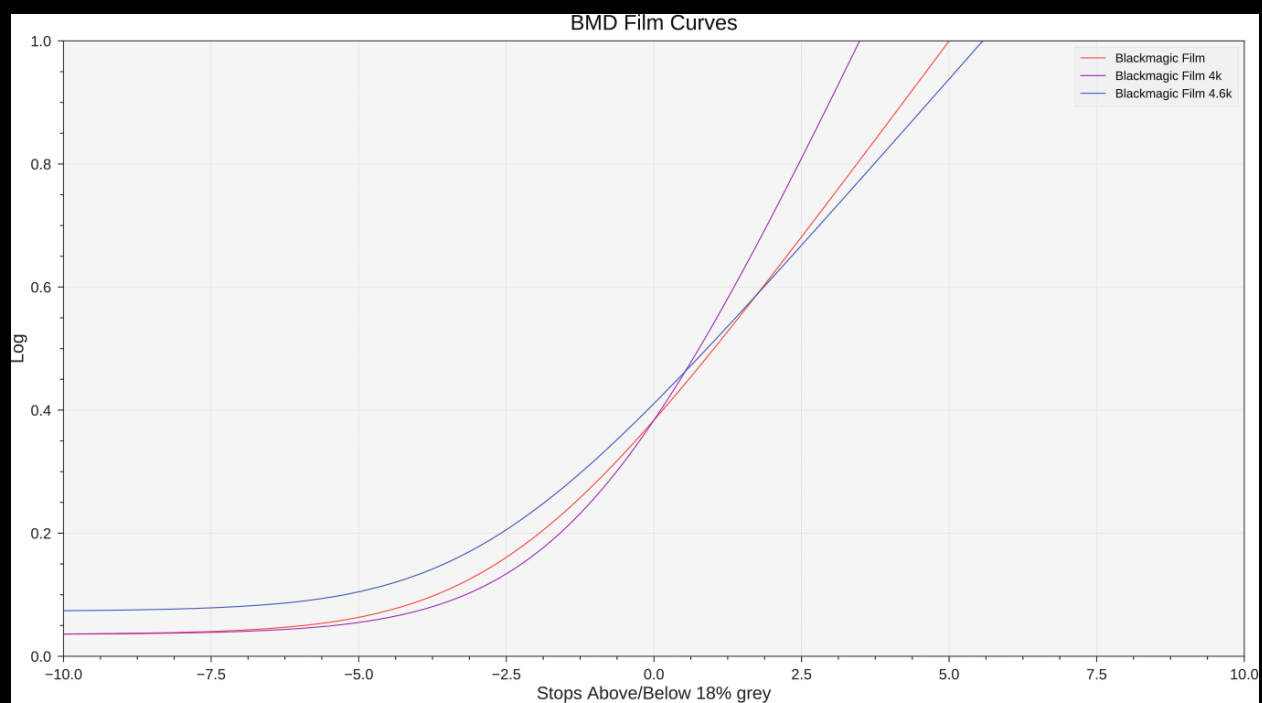


The ACES Still Life reference image converted to LogC / ALEXA Wide Gamut.
Copyright © A.M.P.A.S





The ACES Still Life reference image converted to Blackmagic Film 4.6K.
Copyright © A.M.P.A.S





The ACES Still Life reference image converted to Canon Log3 / Cinema Gamut.
Copyright © A.M.P.A.S





The ACES Still Life reference image converted to GoProTune Native.
Copyright © A.M.P.A.S



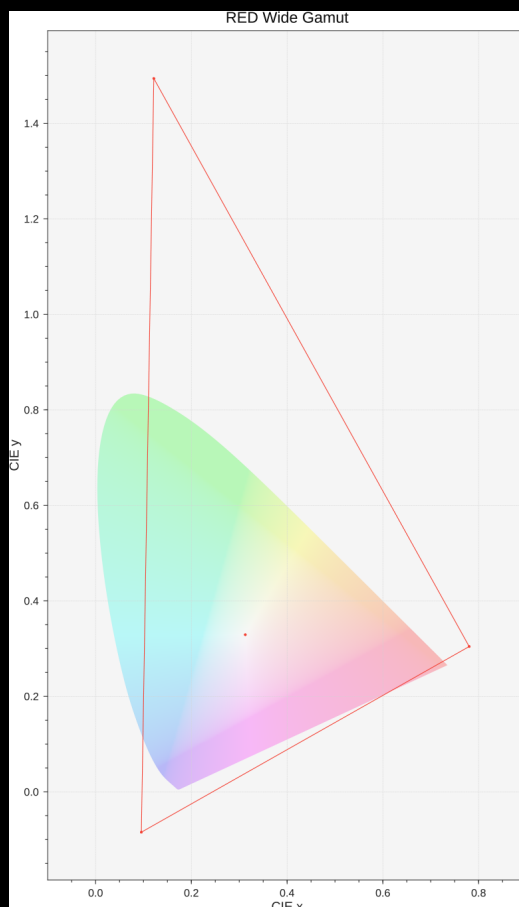


The ACES Still Life reference image converted to V-Log / V-Gamut.
Copyright © A.M.P.A.S



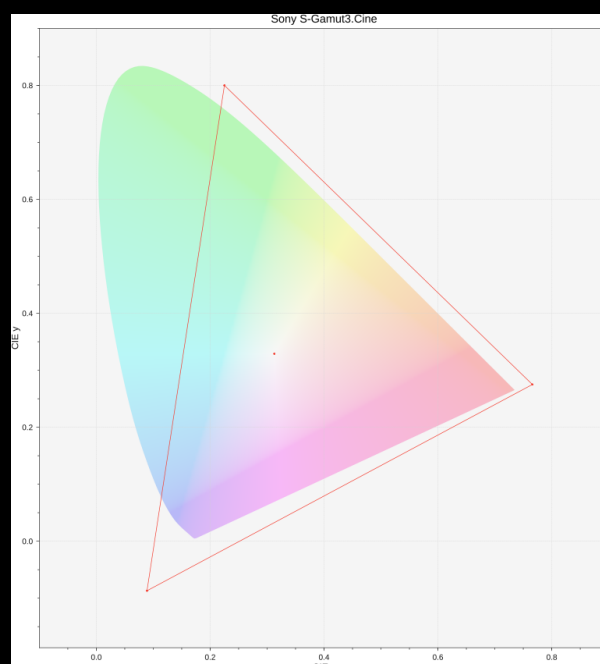


The ACES Still Life reference image converted to Log3G10 / RED Wide Gamut.
Copyright © A.M.P.A.S



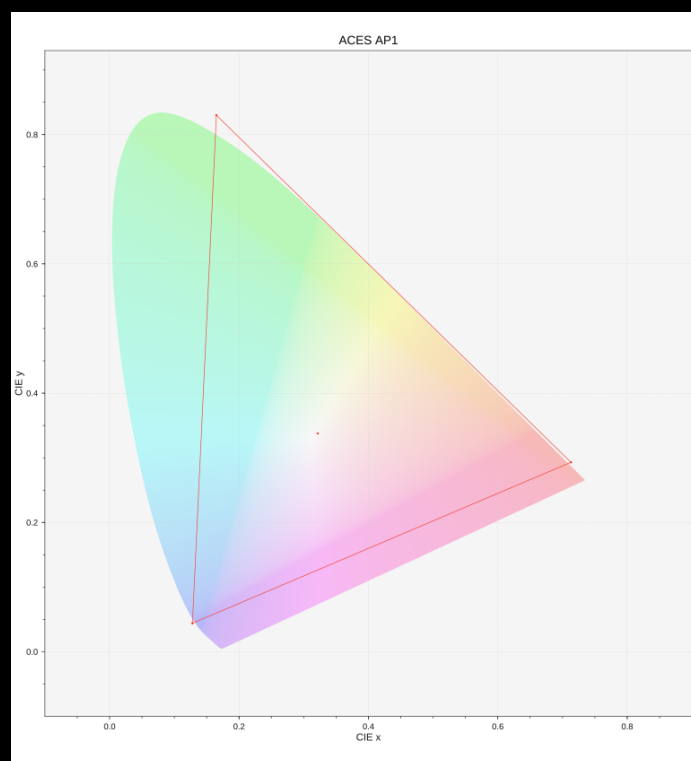


The ACES Still Life reference image converted to S-Log3 / S-Gamut3.
Copyright © A.M.P.A.S





The ACES Still Life reference image as ACES2065-1.
Pixel values >1 are present in the original, but clipped in this rendering.
Copyright © A.M.P.A.S





The ACES Still Life reference image converted to ACEScg.
Pixel values >1 are present in the original, but clipped in this rendering.
Copyright © A.M.P.A.S



The ACES Still Life reference image converted to ACEScc.
Copyright © A.M.P.A.S



The ACES Still Life reference image converted to ACEScct.
Copyright © A.M.P.A.S



The ACES Still Life reference image converted to ADX10.
Copyright © A.M.P.A.S

Why should I care about this?

Great question! The truth is that, for most of the history of motion imaging, none of this was important to anyone but the image scientists who designed film stocks, and later the engineers responsible for standardizing video capture and broadcast. The rest of us got to stand on their shoulders, operating within preconceived pipelines which ran from capture all the way through delivery. There was virtually no choice or control involved on the part of the content creator.

Today, those fixed pipelines and workflows are a thing of the past. Source material can come from one or several of the dozens of available capture formats — iPhone, GoPro, Canon, Sony, RED, Alexa, 35mm, etc.—many of which themselves offer multiple choices of color space. On the delivery side, a given piece of content may need to be played in theaters, as well as on SDR and/or HDR TVs, not to mention the ever-growing, ever-changing list of mobile devices and VR sets.

This is a complicated state of affairs, but it points to a simple truth: The color space(s) we capture to, the way we guide footage to our delivery color space(s), and where and how in that process we choose to do our grading is up to us, and can have as much impact on our images as the grade itself.

This is why it's so important to understand color space—because to be ignorant of it is to risk creating inferior images.

Defining a color space

As we learned earlier, we have the ability to define a color space against a standardized reference, thanks to the CIE.

But in practical terms, how do we clearly and simply invoke these definitions to describe a capture device or display? The most common way is to specify a gamut, gamma, and white point.

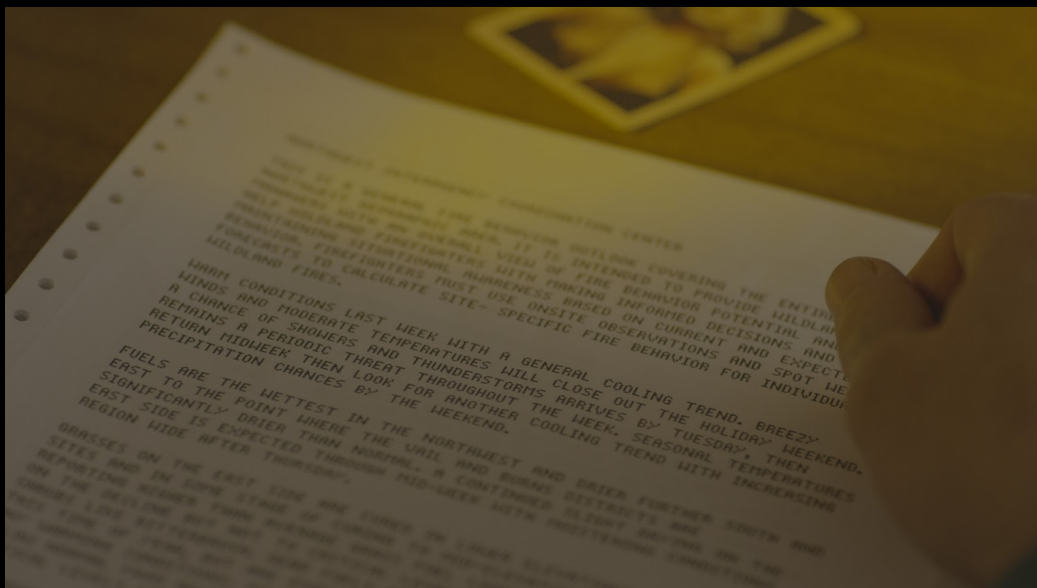
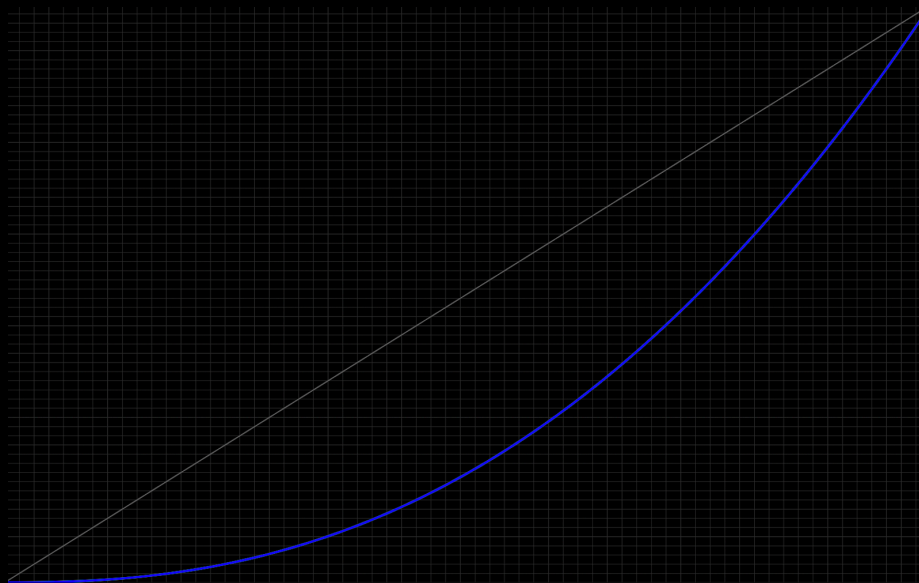
Gamut

A color gamut is the defining a range of chromaticities—essentially a set of possible hues and their respective maximum saturations.

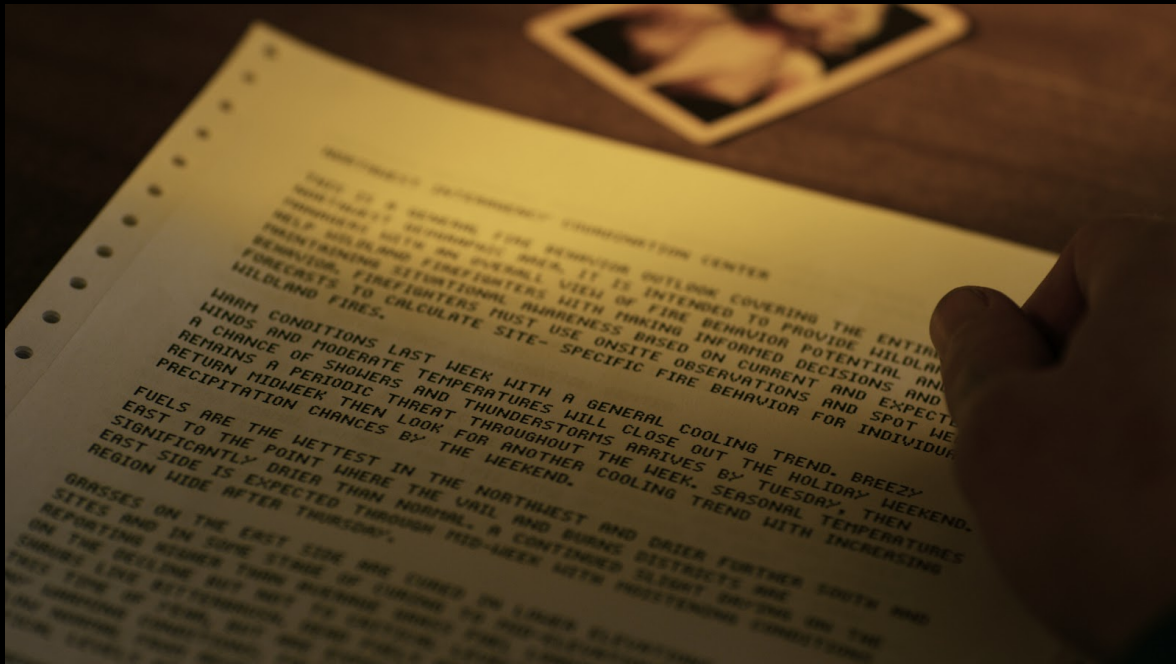
Gamma/Tone Mapping

A gamma curve or tone mapping curve is concerned with defining a specific non-linear distribution of luminance values. Different curves are designed for different purposes.

2D plot of a Gamma 2.4 tone mapping curve. This shows the non-linear distribution of luminance values as we move from pure black (bottom left) to pure white (upper right). A linear distribution would move in a straight line from bottom left to upper right.



A log image prior to gamma correction.



The same log image with a 2.4 gamma curve applied.

White point

In our day-to-day lives, we have no problem recognizing white. If I hand you a sheet piece of paper inside a fluorescent-lit office, then later show you the same piece of paper outside on a sunny day, you'll identify it as white in both scenarios, despite the fact that any digital color measurement device — including a camera — would get vastly different readings in these respective lighting conditions.

This is because our eyes are constantly adapting to our environment, using context and visual cues to define white. Digital sensors and displays are (generally) not designed to do this, so we have to provide this information for them.

This also means we need this information in order to precisely define a color space. In the case of a camera, we basically need a numerical color value to explain what the camera saw as white at the time of capture, since a display may have a different target white point.

It can be tricky to think of white as a non-fixed variable, but that's the reality when it comes to imaging and displays. White point is often expressed as a color temperature, such as 3200K or 5600K, or as one of a list of standard illuminants defined by the CIE.

To make things slightly more confusing, white point is often implied, rather than explicitly denoted, when describing color spaces. And as we'll see below, in the case of cameras, it is variable depending on the original scene.

Now that we've discussed these three parameters, here are some practical examples:

- An Arri Alexa records media in Arri Wide Color Gamut, with an Arri Log C tone mapping curve, and a white point ranging from 2,000K to 11,000K.

- A RED Dragon captures media in RedWideGamutRGB gamut, with a Log3G10 tone mapping curve, and a white point ranging from 1,700K to 10,000K (other gamut and gamma choices are available).
- A cinema projector has a DCI-P3 gamut, a Gamma 2.6 tone mapping curve, and a standard illuminant D63 white point.
- An SDR TV has a Rec 709 gamut, a Gamma 2.4 tone mapping curve, and a standard illuminant D65 white point.



Why Every Editor, Colorist, and VFX Artist Needs to Understand ACES

ACES aims to standardize the color science used across a huge breadth of software and hardware tools, and enable projects of all types to preserve the highest level of image quality throughout the entire workflow, including production, post-production, distribution/presentation, and archiving.

The History Of ACES

In 2004, when the ACES project started, the Academy and many industry experts realized a threat was looming over the future of the cinematic arts.

Decades of innovation in digital workflows had led to dozens of different camera systems, encoding options, display devices, and other hardware and software tools with very little standardization. While this rapid technological advancement brought tremendous benefits to the industry, the Academy realized the lack of commonality might potentially result in the irreparable loss of thousands of digital films.

Thus, the ACES project began. The goal was to build a standardized system capable of managing color in a precise, yet straightforward way, no matter the camera or display being used. The standard also needed to be open, allowing easy adoption in all parts of the industry, yet also robust and adaptable enough to handle the complexity of modern digital production.

“The Academy Color Encoding System is a set of components that facilitates a wide range of motion picture workflows while eliminating the ambiguity of today’s file formats.”

ACES Today

After over 15 years of far-reaching collaboration between the industry and the Academy, ACES is slowly but steadily gaining momentum as the standard color management system across all types of projects.

Since **version 1.0** was released in 2014, ACES has been used on dozens of Hollywood blockbusters, and is making its way into TV series and independent films. ACES pipelines are also being adopted in animation and VFX workflows, where the improved image quality and interoperability is critical to producing accurate results.

ACES is also being adopted by independent manufacturers and developers who aren’t founding partners with AMPAS on the standard. ACES compatibility is already available in many high-end software applications, including DaVinci Resolve, FilmLight’s Baselight, and Nucoda, with new programs adding it regularly.

Conceptually, I love ACES. In practice, I don’t have a terrible amount of hands-on experience with it. What I do know is from my own research and extensive testing in Resolve using Resolve color management (RCM) also uses ACES 1.1 It’s quite brilliantly conceived but it seems to be catching on rather slowly despite established standards and practices and a good track record. Moving forward, I think as the production community meanders farther and farther from HDTV, ACES will emerge as the most appropriate workflow.

The principal developer of **ACES** is the **Academy of Motion Picture Arts and Sciences (AMPAS)**.

Lately I've been taking a hard look at the Academy Color Encoding System aka ACES and trying to wrap my head around it. There are a handful of decent white papers on the topic but they tend to be overly technical. Through pulling tidbits from multiple sources, one can come to a decent understanding of the how's and why's of ACES but I was hoping to find some kind of an overview; something that presented the "need to knows" in a logical and concise way and wouldn't require a big time commitment to understand. I did not find such a resource so I'm writing it myself.

This description is vague but it seems to me that the Academy's hope for ACES is to resolve two major problems.

1. To create a theoretically limitless container for seamless interchange of high resolution, high dynamic range, wide color gamut images regardless of source. ACES utilizes a file format that can encode the entire visible spectrum in 30 possible stops of dynamic range. Once transformed, all source material is described in this system in the exact same way.

2. To create a future-proof "Digital Source Master" format in which the archive is as good as the source. ACES utilizes a portable programming language specifically for color transforms called CTL (Color Transform Language). The idea is that any project mastered in ACES would never need to be re-mastered. As future distribution specifications emerge, a new color transform is written and applied to the ACES data to create the new deliverable.

How ACES Works

It's important to understand that ACES is not just a program or plugin you can download. Nor is it just a firmware update for your camera or monitor. ACES is also not a particular workflow, a creative look for your footage, or a file format for your project.

Rather, ACES is a collection of rules for encoding and transforming data, along with metadata definitions for that data. ACES also encompasses a set of developer tools for integrating those data specifications into software and hardware. Finally, all of ACES' rules, guidelines, and tools are formulated based on a set of standards laid down by the Society of Motion Picture & Television Engineers (SMPTE) and the International Organization for Standards (ISO).

To really come to grips with what ACES is, let's take a look at what it does inside your workflow.

Understanding Capture & Scene-Referred Data

One of the most important things to understand about ACES is that its processing pipeline utilizes capture-referred data—i.e. the color science (and sometimes secret sauce) that each camera system uses and bakes into the signal.

ACES reverse-engineers that data (through what's known as an ACES Input Transform or IDT) back into the pure linear light information that was in the actual scene in front of the camera. Theoretically, without any camera bias.

This is why ACES is often discussed as being scene-referred or, with the more technical phrase, scene-linear. Abstracting away the camera's bias allows us to get closer to the actual real-world scene that the camera was pointing at.

Additionally, ACES color uses a color space so large it actually encompasses the entire visual locus (every color humans can see) and even colors we can't see! Even the smaller working color spaces of ACES (which we'll cover later) are much larger than Rec. 709 and Rec. 2020.

The key to understanding ACES is to acknowledge the difference between **“scene referred”** and **“display referred”** images in simple terms.

A **scene referred image** is one whose light values are recorded as they existed at the camera focal plane before any kind of in-camera processing. These linear light values are directly proportional to the objective, physical light of the scene exposure. By extension, if an image is scene referred then the camera that captured it is little more than a photon measuring device.

A **display referred image** is one defined by how it will be displayed. Rec.709 for example is a display referred color space meaning, the contrast range of Rec.709 images is mapped to the contrast range of the display device, a HD television.

ACES is a unifying standard, which allows you to transform footage captured from many

Unlike traditional DI and video workflows, ACES maintains maximum image quality until conversion for the intended viewing formats. Image © Academy of Motion Picture Arts and Sciences.

different cameras, into a scene-referred color science with a common linear starting point.

ACES is also broad enough to encompass most every color space, without the limitations of smaller gamuts. That means “future proof” formats for client handoffs and archival processes are possible, since the ACES color space is so large.

But the scene-linear approach of ACES is only part of the overall pipeline.

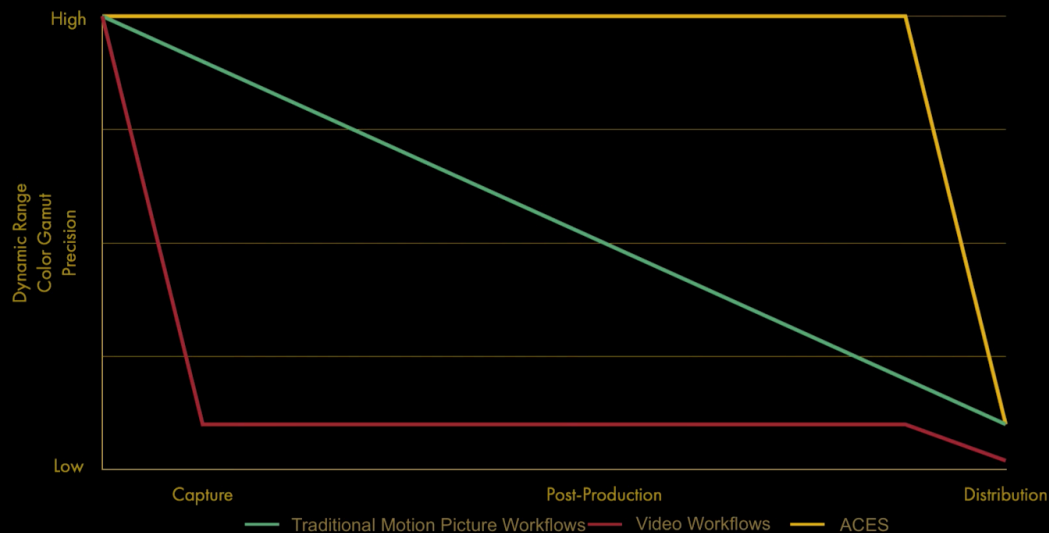
Because the human eye doesn't work in a linear fashion, projects will ultimately have to be viewed on TVs, displays, and projectors that assign gamma or EOTF (Electro Optical Transfer Function) curves. These curves aren't linear, and our current display devices are much more limited in the color spaces they can accurately reproduce (compared to ACES).

ACES already takes this into account and incorporates display-referred color management into the overall pipeline. Linear ACES data is parsed through different transforms for different color spaces and display devices, so that an image appears accurately in a variety of viewing contexts.

That means you can capture and work with footage in the highest quality all the way until final delivery, and then ACES makes it easy to adapt footage for almost any screen.

When it comes to precision, ACES uses Open EXR 16-bit half-float processing which results in 30+ stops of scene-referred exposure.

Image Quality Throughout the Workflow



Keep in mind that even though EXR is used, often this processing is just internal to the app you're using and no EXR files are created for you to manage—except in renders.

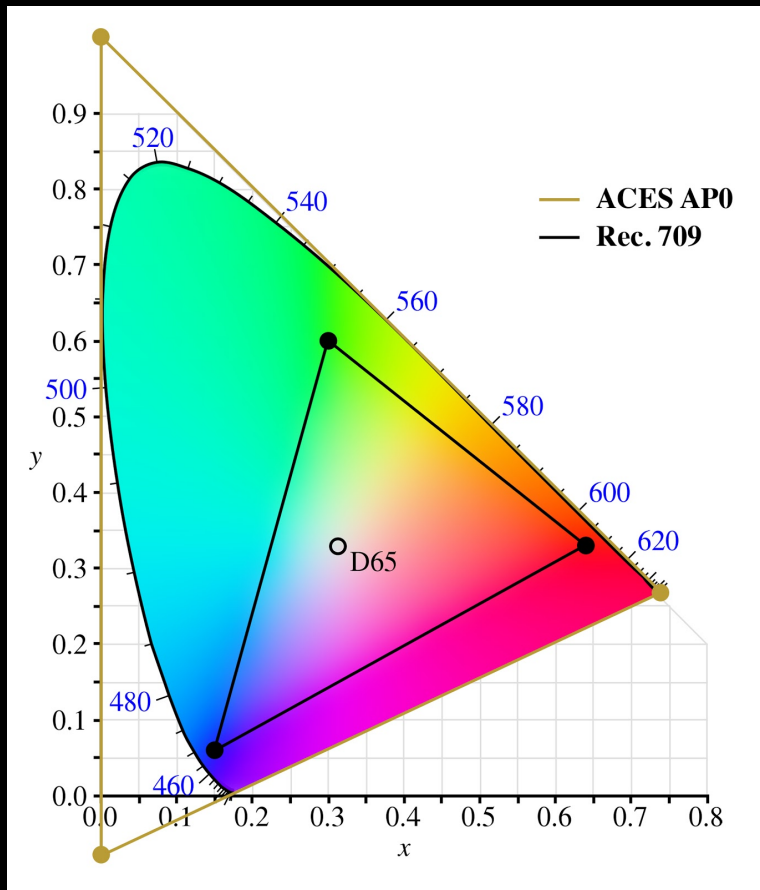
In sum, these are the benefits of ACES pipelines:

- Camera System Color Science Unification – Because of the capture-referred, scene-linear transform at the start of an ACES processing pipeline (no matter the camera(s) that were used on a project) there is a common color science starting point that all ACES capable applications can use and understand.
- No Guessing For VFX/CGI Workflows – One reason that ACES has been embraced by VFX/CGI heavy films is that they're compositing and working in linear anyway! And that linear data can then be rendered back to whatever is appropriate to the project, or kept linear and given back to a colorist to simply reapply their existing grade with no look shifting.
- Ready For Wide Gamut/High Dynamic Range – Now and in the future. Because ACES is capable of retaining 30+ stops of image data and the gamuts are so large, it's a great match for HDR, wide gamut projects, however those concepts may develop in the future.
- Evergreen Digital Masters – One major thing that the Academy pushes about ACES is that it allows for a true evergreen digital master because of the ultrawide/high dynamic range nature of the system.

The 'Parts' Of ACES

Even though ACES and its various transforms are quite mathematically complex, you can understand ACES better by understanding what each part or transform in the pipeline does.

Here's the terminology for each of these transforms:



ACES AP0 color space encompasses all the colors humans can see, dwarfing the range of colors Rec. 709 is capable of encoding.

• **ACES Input Transform (aka: IDT or Input Device Transform)** – The Input Transform takes the capture-referred data of a camera and transforms it into scene linear, ACES color space. Camera manufacturers are responsible for developing IDTs for their cameras but the Academy tests and verifies the IDTs. In future versions of ACES, the Academy may take on more control in the development of IDTs. IDTs, like all ACES transforms, are written using the CTL (Color Transform Language) programming language. It's also possible to utilize different IDTs for a camera system to compensate for different camera settings that might have been used.

• **ACES Look Transform (aka: LMT or Look Modification Transform)** – The first part of what's known as the ACES Viewing Transform (the Viewing Transform is a combination of LMT, RRT, & ODT transforms). LMTs provide a way to apply a look to a shot in a similar way to using a Look Up Table (LUT). It's important to note that the LMT happens after color grading of

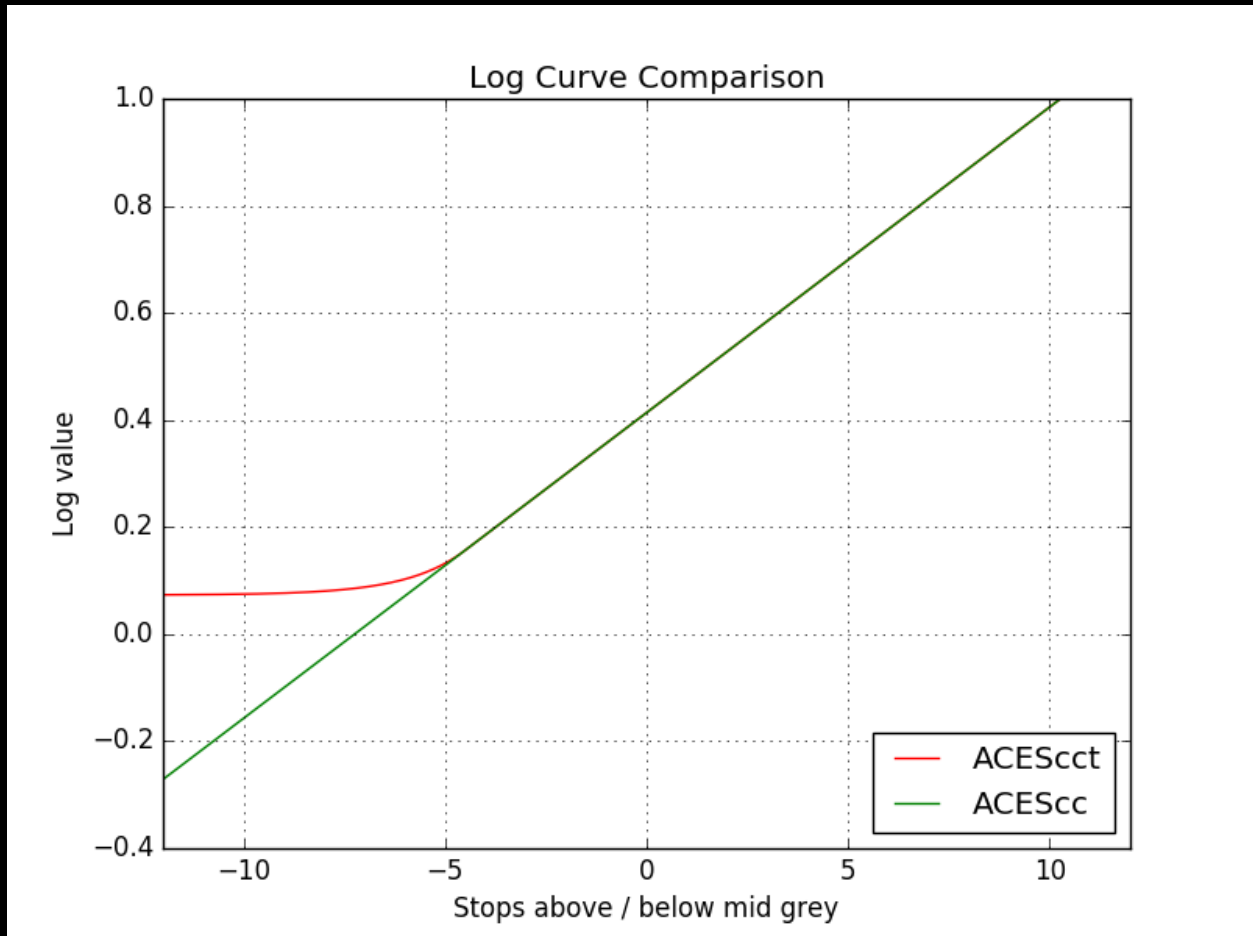
ACES data and not every tool supports the use of LMTs.

- **RRT (Reference Rendering Transform)** – Think of the RRT as the render engine component of ACES. The RRT converts scene referred linear data to an ultrawide display-referred data set. The RRT works in combo with the ODT to create viewable data for displays and projectors. While the Academy publishes the standard RRT, some applications have the ability to use customized RRTs (written with CTL), but many color correction systems do not provide direct access to the RRT.
- **ACES Output Transform (also known as the ODT or Output Device Transform)** – The final step in the ACES processing pipeline is the ODT. This takes the ultrawide, high dynamic range data from the RRT and transforms it for different devices and color spaces like P3 or Rec 709, 2020, etc. ODTs like IDTs and RRTs are written with CTL.

There are also three main subsets of ACES used for finishing workflows called ACEScc, ACEScct and ACEScg:

- **ACEScc** uses logarithmic color encoding and has the advantage of making color grading tools feel much more like they do when working in a log space that many colorists prefer.

- **ACEScct** is just like ACEScc, but adds a ‘toe’ to the encoding so that when using lift operations the response feels more similar to traditional log film scans. This quasi-logarithmic behavior is described as being more “milky,” or “foggier.” ACEScct was added with the ACES 1.03 specification and is meant as an alternative to ACEScc based on the feedback of many colorists.
- **ACEScg** utilizes linear color encoding and is designed for VFX/CGI artists so their tools behave more traditionally.



ACEScct provides colorists a more familiar feel for grading compared to ACEScc.

While ACEScc, ACEScct and ACEScg and the transforms they use are what you'll most often see when it comes to ACES, there are some additional terms you may encounter:

- **APD (Academy Printing Density)** – AMPAS supplied reference density for calibrating film scanners.

- **ADX (Academy Density Exchange)** – Used for scanning film and getting those scans into ACES – similar to the Cineon system for scanning.
- **ACESproxy** – uses logarithmic color encoding and is an integer, range limited version of ACEScc. ACESproxy is meant to be used on set with compatible equipment over SDI.

The ACES Pipeline

Now that we've defined the transforms used for ACES, understanding how the various transforms combine to form an **ACES processing pipeline** is pretty straightforward:

Camera Data -> Input Transform -> Color Grading -> Look Transform (optional) -> Reference Rendering Transform -> Output Transform

As mentioned, ACES is a hybrid color management system of scene referred/scene linear and display referred data.

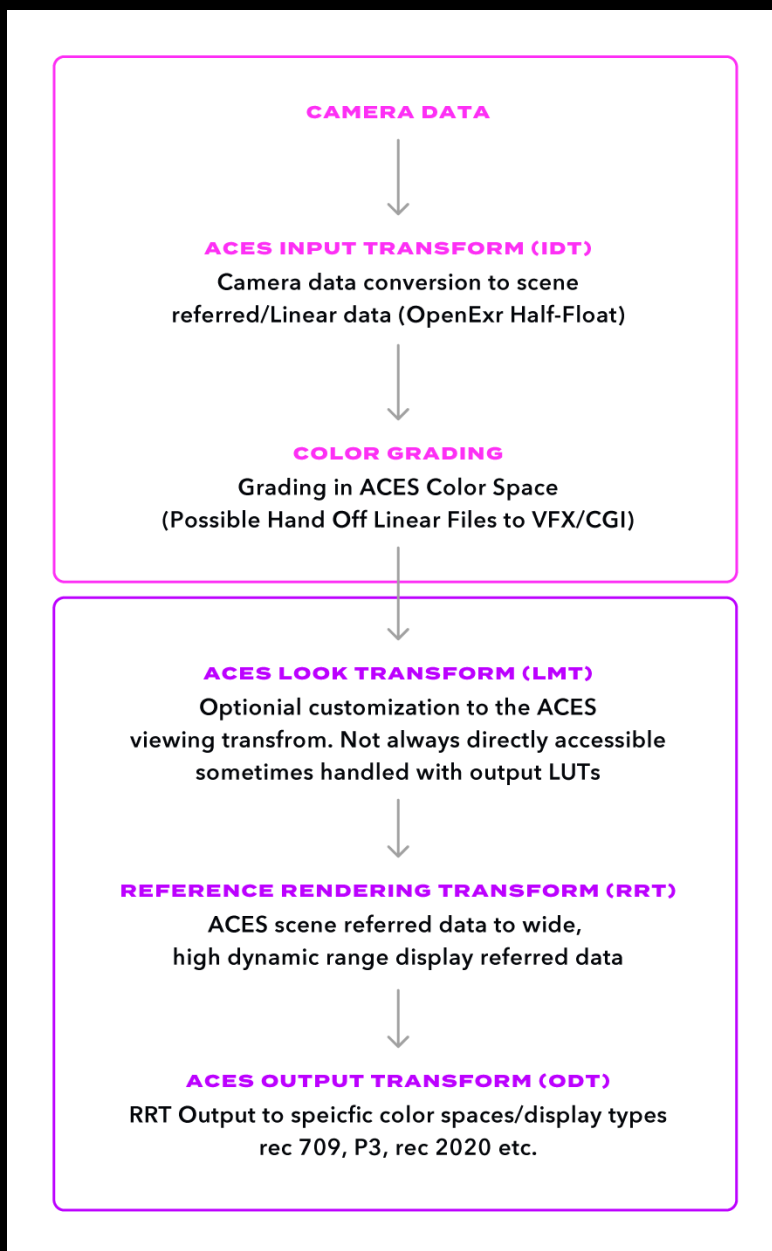
*In the graphic, the various ACES transforms fit into the **scene referred (top section)** and **display referred (bottom section)** part of the pipeline:*

ACES Color Spaces

When ACES is discussed, you'll often hear terms like 2065-1, AP0, AP1, and Rec. 2020+ thrown around.

What do these terms mean?

SMPTE ST 2065-1 is the SMPTE standardization of ACES.



While this standard has many parts, in daily usage ACES 2065 has come to mean the full linear version of ACES that has a larger gamut than the visual locus set of RGB primaries. ACES 2065 uses a set of primaries known as AP0 (ACES Primaries 0).

While ACES AP0 encompasses a much larger range of colors than most all other gamuts, ACES AP1 is only slightly larger than Rec. 2020, lending it the name Rec. 2020+

2065-1/AP0 is mainly meant for archival and file exchange. For grading, vfx, and editing AP1 is more likely to be used currently. ACEScc, ACEScct, and ACEScg all utilize AP1 primaries.

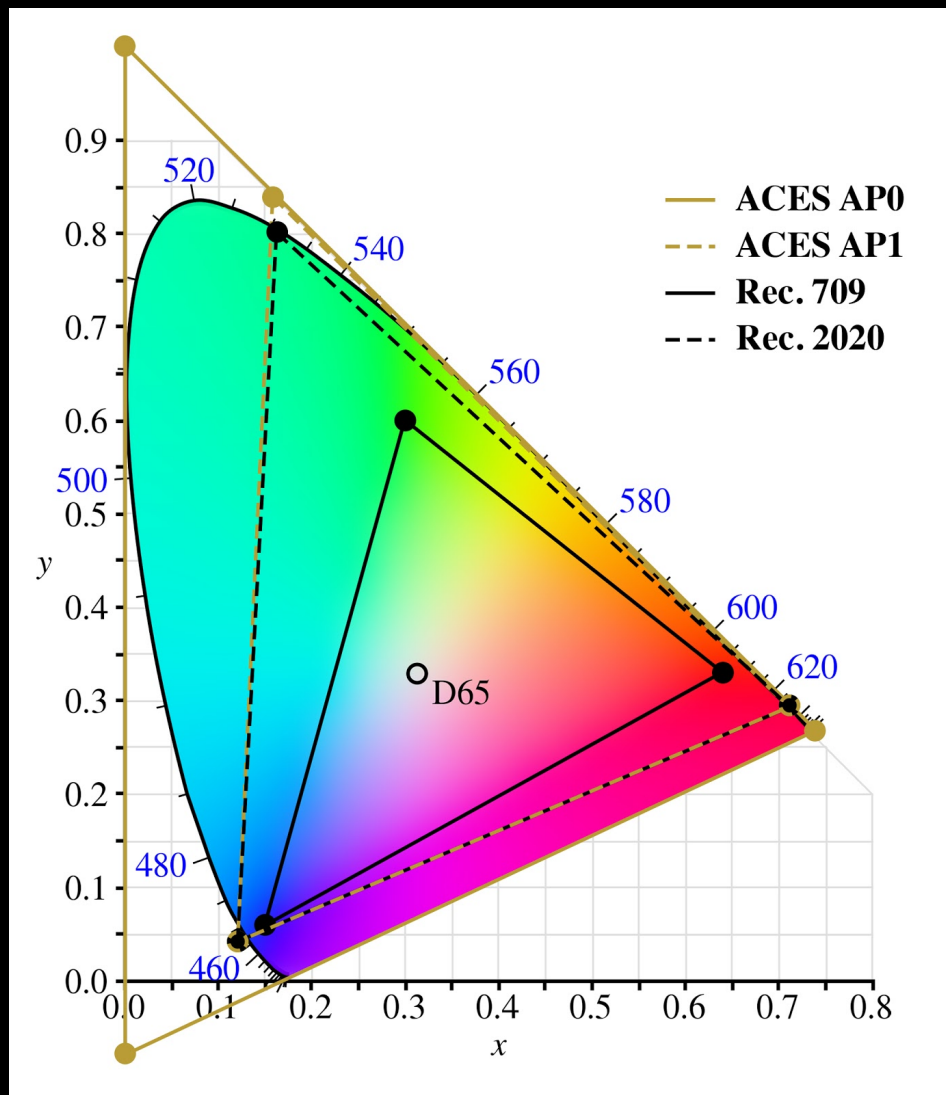
Using ACES

As you can see, ACES represents a huge leap over the color management systems you might be used to in your everyday workflow. At this point, you might be thinking “Yes, I want that now. How do I get it?”

As mentioned earlier, several post-production applications are already offering ACES compatibility, so if you use any of those programs regularly, you’re only a few steps away from diving into ACES.

That said, as we’ve discussed, ACES is a pipeline. For it to be valuable, every step of your workflow needs to follow the ACES rules.

So, how you can truly adopt ACES is very much dependent on what you are working on and what you need ACES to do.\



For example, VFX artists can already viably use ACES for improved compositing accuracy and handing off work between teams. The increased color accuracy can add incredible levels of realism to your scenes, and solves a myriad of technical issues artists have been battling for years.

But for many creatives, who deal with live-action footage on a daily basis, complete adoption of ACES will depend on production teams embracing the new standard. Of course, this is largely influenced by what transforms are available for which camera manufacturers. That's the necessary link for incorporating camera formats into the ACES system.

Obviously, it will take some time for hardware manufacturers to catch up for more workflows to be able to adopt ACES in earnest. But that doesn't mean everyone has to wait around to get their hands dirty with ACES.

Open Source Development

Because ACES is an open source project, anyone can help bring the standard to life in new ways.

While AMPAS actively manages and curates the development of the ACES system, the development of the system is held in a GitHub repository for anyone to contribute. ACES transforms are written in Color Transform Language (CTL), so if you're mathematically inclined and have experience in color science, you can contribute there.

If you're less technically-inclined but still want to contribute, you can visit ACES Central, an AMPAS run community. There you can connect with other film and video professionals who are exploring the future of ACES, and learn how other end users and ACES certified partners are adopting the technology.

Versions

The good news about ACES' continuous development is that new versions are always intended to be backwards compatible. Data encoded using older versions of ACES can be opened using a newer version of ACES that an application uses, though this may cause some differences in image representation compared to the original project. Many applications offer the ability to switch the version of ACES being used, so that you can match the appropriate version to the original project.

The most up-to-date official version of ACES is 1.1. Many applications do not yet support it, so version 1.03 remains the most widely supported version currently.

Unless there are major future changes and/or needs for end-users as determined by AMPAS, we will probably only see the major release versions appear in software, while the minor .0x releases may just be implemented behind the scenes as performance and bug fixes.

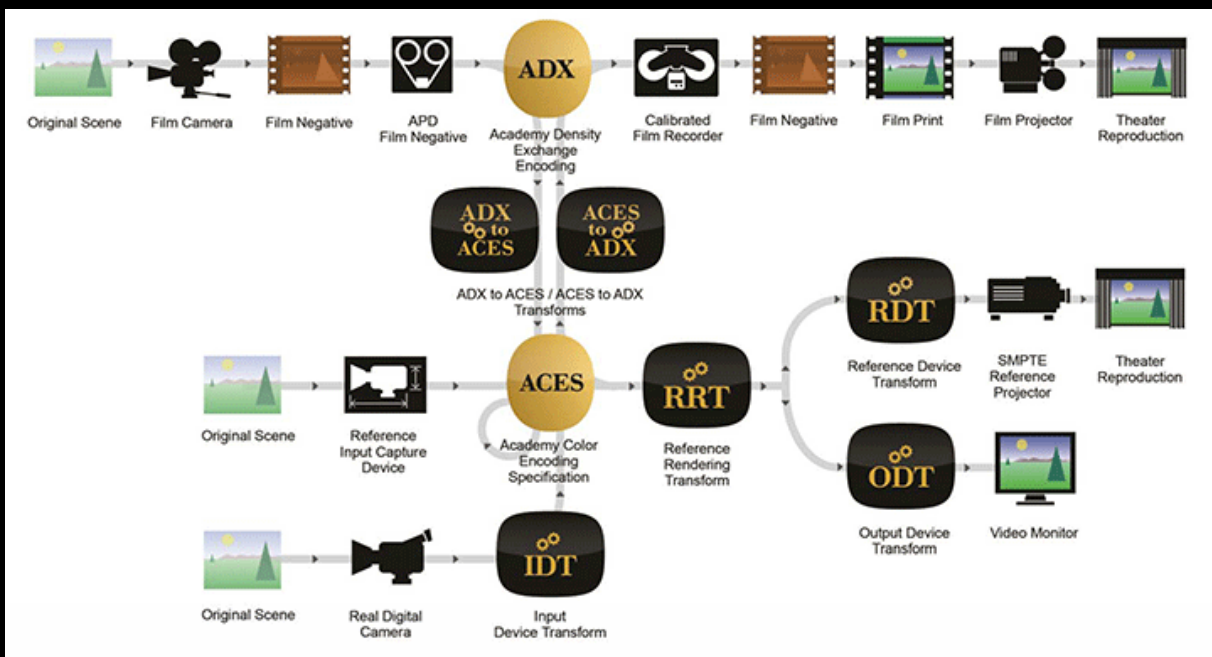
ACES till now and what the future holds

We know this is a lot to take in, but it is worth your time to get a basic understanding of ACES now so that you can adopt it as it becomes viable for your workflow in the future.

ACES isn't a magic bullet that will fix every technical issue or creative challenge of filmmaking. No technology will ever do that.

That said, ACES does offer some incredible advances for the future of video and cinema. This is a future we could only dream about in the not too distant past, so it's exciting to see it just over the horizon.

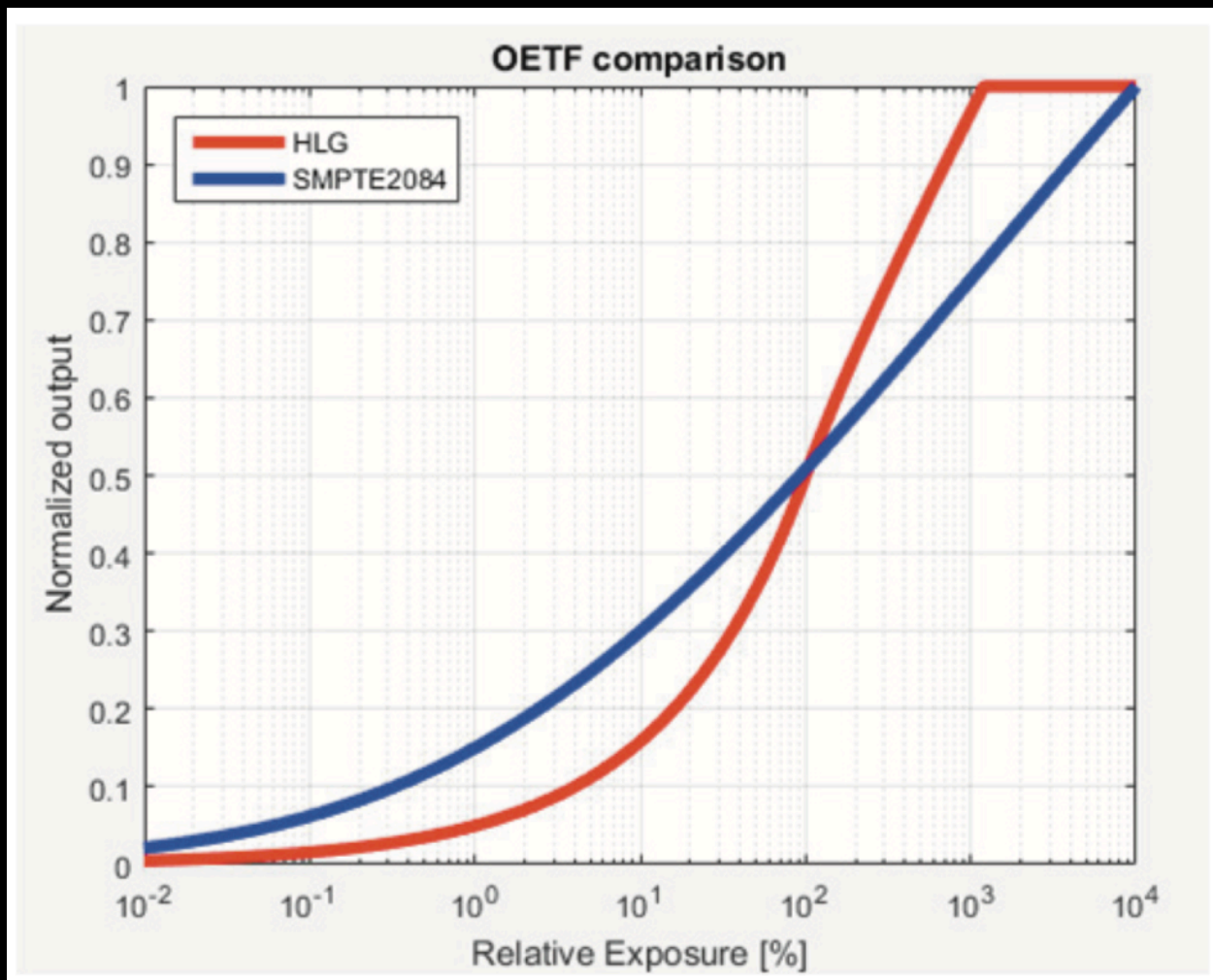
Time will tell how fast it gets here.



ITU-R BT.2100 recommends: "...that for program production and international exchange of HDR-TV, the perceptual quantization (PQ) or Hybrid Log-Gamma (HLG) specifications should be used...high dynamic range television production and display should make consistent use of the transfer functions of one system or the other and not inter-mix them."

HDR and Its Impact on Workflow

Now that we have discussed in detail of the color process, color spaces and ACES we further go onto discuss about HDR Standards, HDS/SDR compatible parallel workflows, conversions static vs Dynamic and how to integrate the benefits of High Dynamic Range in to existing video production and distribution.



HDR is winning over consumers today because of its ability to deliver a visibly wider range of highlights and shadows, as well as more realistic color and detail. Thanks to its contrast ratio, which is much closer to the conditions found in real life, HDR allows image reproduction that is much closer to reality. Furthermore, HDR also allows more dependable results under difficult shooting conditions — such as irregular lighting or partial shade — found at many outside broadcast productions. An additional advantage of HDR is that it is fully format-independent, and does not need any specialized viewing conditions to show its advantages.

For broadcasters and content creators, embracing an HDR workflow process raises a number of questions and poses some challenges. Most notable is the need for a parallel SDR/HDR production workflow, one where the signal can be adapted with up/down mapping as required to mix and match incoming content formats and output signals without sacrificing any quality.

For example, operators may need to integrate existing SDR content into new HDR productions, or may need to send HDR content to SDR multi viewing screens in the studio. In either case, the creator's intent for each individual signal must be maintained.

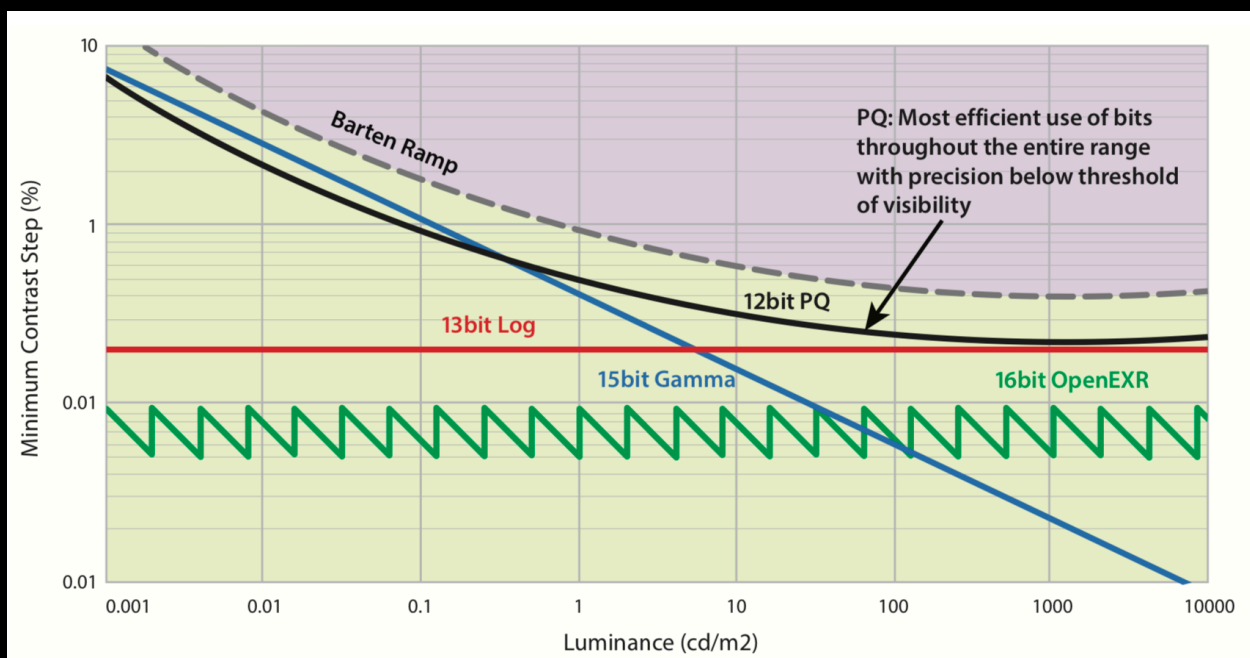
Let's now dive into the challenges and the best possible solutions to produce native 10-bit HDR in either HD or 4K UHD, and deliver that content in HDR and SDR simultaneously, as necessary, thanks to the high-quality conversion that can be done with a down mapping

process. These solutions provide native support for both of today's worldwide standard: Hybrid Log-Gamma (HLG) and Perceptual Quantization (PQ or SMPTE ST 2084).

The Flavors of HDR

Unlike the offerings of a certain ice cream shop, there aren't 31 flavors of high dynamic range—but there might as well be. Because just as it would be impossible to cram all of those different flavors of ice cream into one cone, **it's not possible to lump all of the types of HDR into a single signal. And with the exception of one approach, it's not even possible to mash HDR and standard dynamic range (SDR) into the same video signal.**

That exception is **Hybrid Log Gamma (HLG)**. Developed cooperatively by the BBC and the national Japanese broadcaster NHK, HLG specifically addresses the complexity that broadcasters otherwise would face in their production and distribution chains if they were required to deliver separate SDR and HDR versions of the same program. After all, broadcasters can't abandon the hundreds of millions of homes with legacy SDR TV viewers simply to deliver HDR content to a relatively nascent, though quickly growing, audience segment.



HLG combines SDR and HDR into one signal so that conventional SDR TVs can decode the dynamic range information they need and ignore the HDR part of the signal in a fashion that's reminiscent of how black-and-white TVs once ignored newly added chroma signals in the early days of color TV. With HLG, a typical gamma curve that legacy HDTVs can recognize is used for the lower part of the signal, and a logarithmic curve is used for the upper portion. Combined, they create HDR images in HLG-enabled televisions.

SDR imagery may be slightly affected by the wider **ITU BT.2020** color gamut in bright areas and highly saturated colors, and **HLG LUTs** continue to evolve to address this. Attention may need to be paid during transforms to ensure that these aberrations are minimized.

Another distinguishing characteristic of HLG is that it is scene-referred. In other words, the HLG signal represents and encodes the light detected by the camera. No additional control over dynamic range is provided or required. That's by design, because HLG is intended to deliver HDR signals from the camera to a home display without complication, which makes it well-suited for live productions such as sports.

The other major flavor of **HDR is PQ**, or **Perceptual Quantizer**. Like the chocolate in the freezer case at the ice cream shop, there isn't just chocolate—there's also chocolate with almonds and chocolate with marshmallows.

Unlike HLG, **SMPTE ST 2084 PQ** does not use a typical gamma curve in combination with a logarithmic curve. Rather, PQ represents the true light values of a scene and conveys them via a digital signal to be output as light by the display, from black up to 10,000 nits, which currently is beyond the upper brightness capacity of TVs. (A nit, a measure of brightness, is defined as candela per square meter.)

To represent true light values in a scene requires 16-bit code words. However experiments conducted by Dolby Labs revealed that there's a more efficient way to use code words to reduce the number of bits required. Starting with the Barten Ramp, named for Peter Barten, author of Contrast Sensitivity of the Human Eye and Its Effects on Image Quality, Dolby set out to find the point at which viewers would spot that banding in images as the number of bits making up a code word was reduced. They defined this point as the Just Noticeable Difference (JND). Dolby also defined a special curve that follows where humans detect the JND, known as the PQ curve. The researchers found this curve required 11-bit code words, which was rounded up to 12 bits.

It's possible to apply this curve using fewer bits and to create an image that looks good to viewers, which the Consumer Technology Association did in defining **HDR10**, which uses the **PQ curve** and 10 bits. **PQ** also differs from **HLG** by giving creatives, such as a colorist in postproduction, a means to make adjustments to express their vision and then use metadata to control viewer TVs so that they reproduce images as intended. Effectively, concurrent metadata is sent with the video signal that communicates with the display over HDMI as to how to map both light and color most effectively for the display's capabilities.

Three flavors of PQ are available today. **HDR10** is an open-source **HDR standard** based on **SMPTE ST 2084 PQ EOTF** and **Rec. 2020** with a bit depth of 10 bits. This approach to HDR leverages the **SMPTE ST 2086** metadata standard to convey static data describing the color calibration of the mastering display to televisions and other displays. That calibration data includes the maximum frame average light level, known as **MaxFALL**, and maximum content light level, called **MaxCLL**. TVs then reproduce the images displayed based on this static HDR metadata.

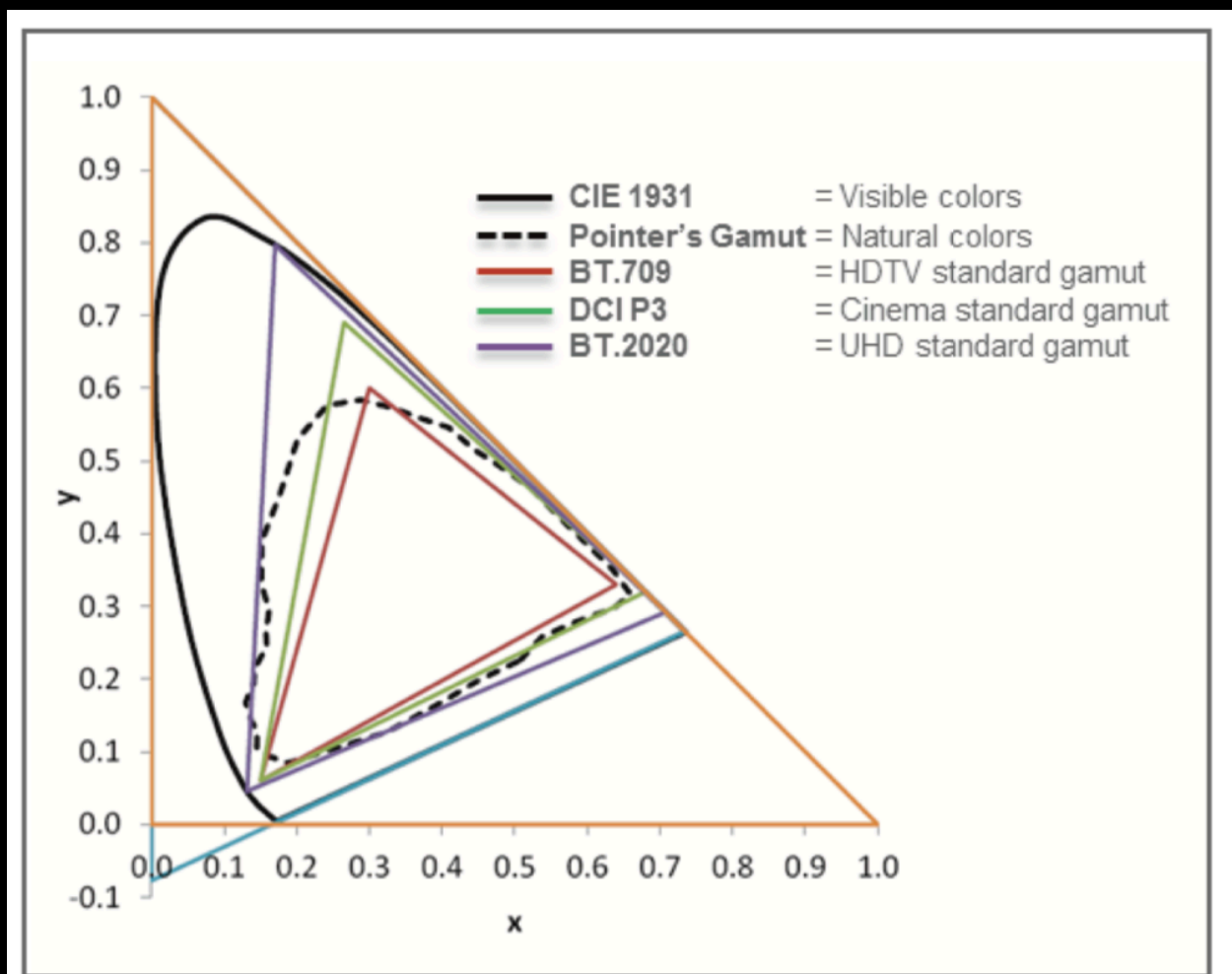
HDR10+, supported by companies like **Samsung**, **Amazon**, **20th Century Fox**, **Panasonic**, and **Warner Brothers**, builds on this approach with the **addition of dynamic metadata**—thus, the plus. Dynamic metadata allows creatives to adjust brightness and color volume on a scene-by-scene, or even frame-by-frame, basis. **HDR10+** relies on Samsung's implementation of the **SMPTE ST 2094-40** standard for this dynamic metadata.

Dolby Vision is Dolby Laboratories' proprietary implementation of PQ. While it implements **SMPTE ST 2084 PQ EOTF** and **Rec. 2020** color space, one of its key differentiators is support

for 12-bit color. Another is its support for both static and dynamic metadata, **SMPTE ST 2086** and **SMPTE ST 2094-10**, respectively. Because it's Dolby's technology, its use by production equipment vendors and consumer electronics makers **requires a license**.

All three PQ approaches are **display-referred systems**. In other words, when a certain code value is assigned in postproduction, a certain number of nits will be reproduced on the television. Thus, if the TV receives the code word for 303 nits, it must put out 303 nits of brightness.

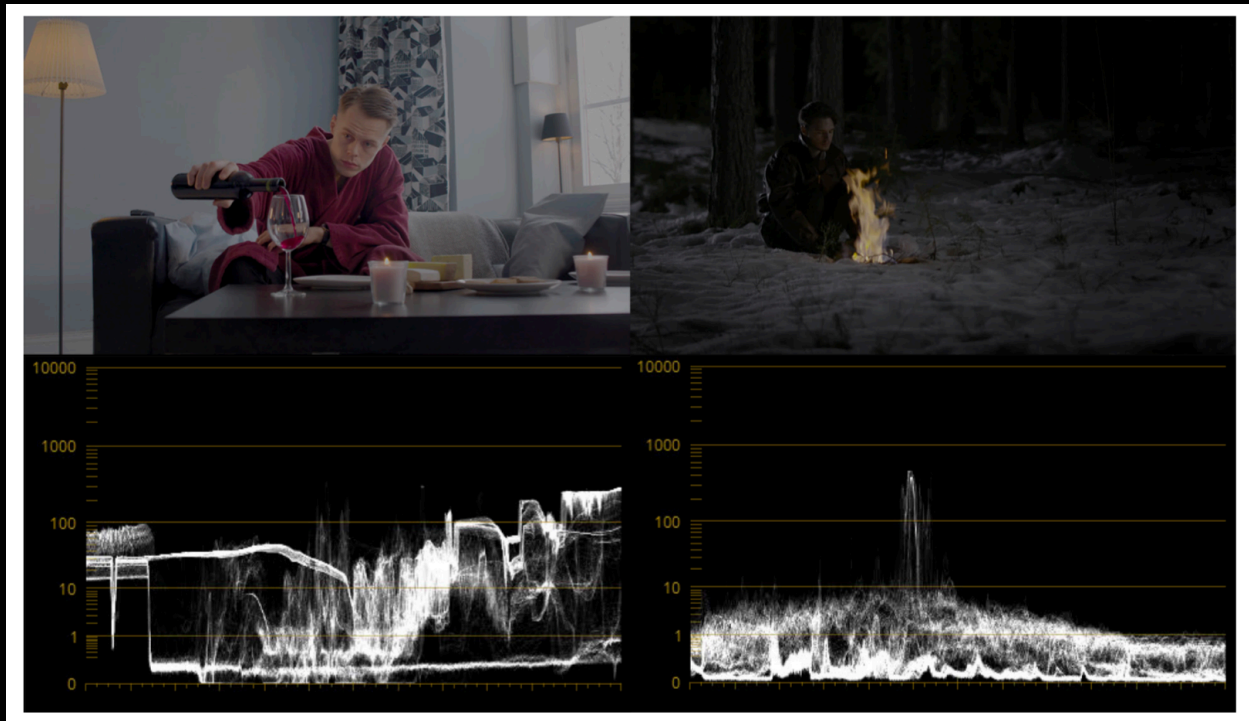
Like a hand in a glove, **HDR fits with wide color gamut, or WCG**. They're nearly inseparable because each HDR implementation specifies a color space recommendation and bit depth. Thus, **HDR10** and **HDR10+** use **Rec. 2020 10-bit color**, a far wider color palette than is available to an **HDTV display** that uses **Rec. 709 8-bit color**. Specifically, **the former can reproduce 1,024 color tones** per R, G, and B sub pixel, or a **total of 1.07 billion colors**, while the latter is capable of **just 256 tones** for R, G, and B, or a **total of 16.8 million colors**. (That's $1,024 \times 1,024 \times 1,024$ vs. $256 \times 256 \times 256$.) A color depth of 12-bit takes the total to 6.8 billion. **Together, HDR and WCG produce what truly can be described as a better pixel.**



HDR and the Colorist

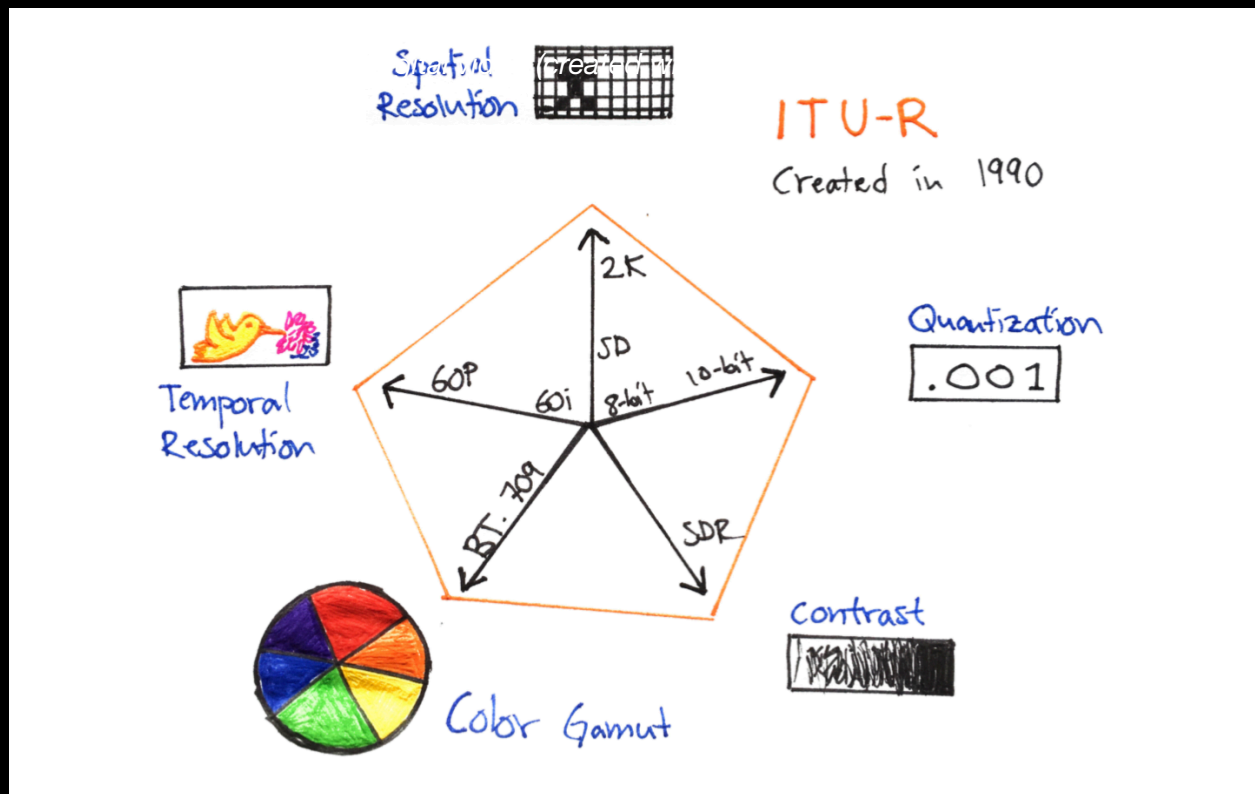
How can color grading in HDR be approached?

The objective is to understand how color grading in HDR can be approached; to find out how it can differ from color grading in SDR (Standard Dynamic Range), with the motivation to better understand how this technology can affect colorists working in the motion picture industry.



What effect can HDR have on visual modality?

On the Canon Pro Vimeo channel the video Canon Color Gamut3 defines the “five key imaging parameters” of digital image technology as; **temporal resolution**, **spatial resolution**, **quantization**, **contrast** and **color gamut**. These five parameters can be divided into two categories of either contributing to the amount of pixels (**temporal and spatial resolution**), or the quality of the pixels (**quantization, contrast, color gamut**). While both the quantitative and qualitative categories are important in their own right the main focus lies in the qualitative part; especially regarding contrast. Although, it is essential to have an understanding of all five of these parameters, therefore, each one will be briefly presented below.



Quantitative properties

The quantitative properties are quite straight forward and easy to describe. Spatial resolution in a digital image describes the amount of pixels and is presented by the amount of horizontal and vertical rows of pixels, for example; 1920x1080 (Full-HD) or 3840x2160 (UHD) which are both HDTV standards. Temporal resolution refers to the frequency or refresh rate of the video signal, often referred to as FPS (frames per second). As mentioned in the introduction; 24fps is common in cinema, although we have seen examples of HFR (high frame rate) for example in the 3D version of The Hobbit. Raising the spatial and temporal resolutions to higher values will simply put result in more pixels more often. Finally, when you see attributes such as 1080p30, then it refers strictly to resolution, both spatial and temporal.

Qualitative properties

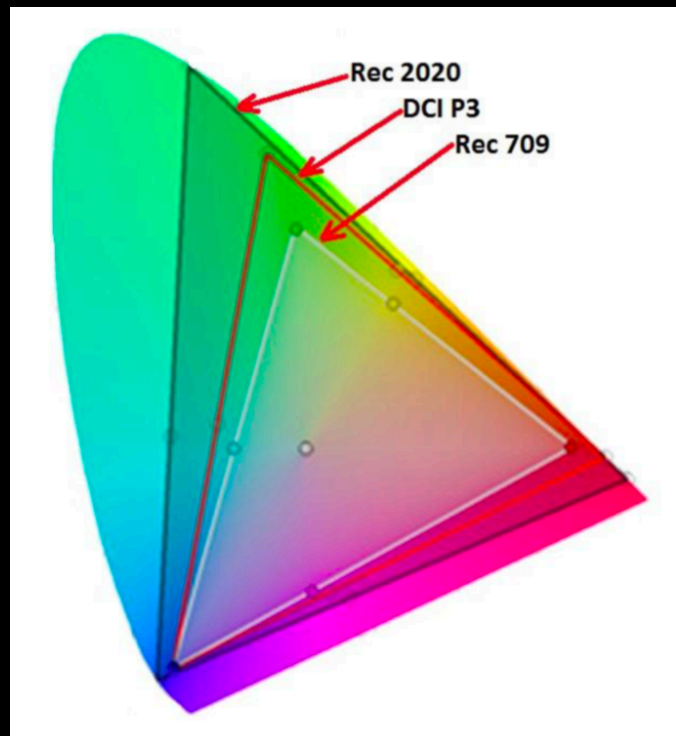
If higher quantitative values are described by more pixels more often, then higher qualitative values can be described as a larger volume of possible pixel brightness and color values with finer numerical gradation, let's find out why. The finer numerical gradation part refers to quantization. According to Stump "quantization is the process of converting continuously

varying analog voltages into a series of numerical values called samples”. “An 8-bit sampling resolution means that the continuous values of the input signal will be quantized to 2 to the 8th power, or 256 code values —in other words, 256 shades of red, 256 shades of green, and 256 shades of blue. When the red, green, and blue color palettes are multiplied to define the entire color palette, $256 \times 256 \times 256$, the result is 16,777,216— defined as 8-bit color in the digital realm.” Even if 16,7 million colors might intuitively seem like a lot it’s actually quite limiting, and can often times result in banding artifacts over areas with fine gradation, such as a sky. Since 8-bits results in 256 code values per channel there’s also only 256 code values between black and white; where all three channels are used in equal amount to generate the different shades of grey. To put this into perspective; 256 shades of grey in 8-bit can be compared with 1024 in 10-bit or 4096 in 12-bit. According to Stump “practical testing of human sensitivity to ‘just noticeable differences’ in contrast conducted recently determined the minimum quantizing level for the threshold of quantizing visibility in digital cinema to be 12 bits as a minimum”. Quantizing is also relevant to discuss in relation to the contrast of an image; since the actual brightness difference between the code values will vary depending of the overall tonal range. This master thesis will use the terms contrast, tonal range and dynamic range interchangeably. “The dynamic range of a TV refers to its luminance, the maximum and minimum amount of light the TV is capable of producing.”

While luminance is measured in candela per square meter (cd/m^2), commonly referred to as nits, dynamic range is presented as a ratio (for example 1500:1) between the brightest and darkest values a display can produce, where SDR- standards support luminance values in the range of 0.0002 to 100 cd/m^2 .¹⁰ The last of these five parameters is color gamut (or color space), which defines the reach or richness of hues. Different gamut’s encompass various amounts of the CIE 1931 chromaticity diagram (also called color horseshoe), which was “specifically designed to encompass all colors the average human can see”. It’s worth noting that there are other versions of the CIE 1931 such as the 1960 and 1976 versions, though the differences won’t be discussed here. Schulte and Barsotti explains that the legacy color space called BT.709 (also known as Rec.709) leaves “a large set of visible colors that cannot be rendered” and that “larger color spaces, such as DCI-P3 and BT.2020 can represent a much larger set of visible colors”. The BT.709 gamut is what we are used to see from our televisions, while DCI-P3 is a standard seen in the cinema, and BT.2020 as of today can only be displayed in full by certain laser projectors.

An introduction to the colorist

“Twenty-five years ago, the term colorist did not exist. We lived in a world where what you shot was pretty much what you got.” Hullfish writes that the first colorists were engineers in the film chain; where a camera pointed into a projector turned the motion picture content recorded on



film into a signal that could be broadcasted or recorded on to videotape. In this process, the engineers could adjust a very limited set of values on the camera to compensate for issues with the film. The film chain later developed into telecine, where instead of projection; the film was scanned using a gas electron beam. Combined with a much more gentle transport of the film this meant that instead of a print, the original film negative now safely could be used. As the available tools developed so did the people associated with the task of correcting color; “who began to experiment with the controls, developing different looks or styles with film”. Today, the telecine has moved into the digital domain where a datacine creates digital images (similar to frames from a digital still camera) which is called a digital intermediate (DI). Nowadays though, it is more common to capture motion picture digitally, rather than scanning and digitalizing images captured on film stock.

Color correction and color grading

To get a better understanding of the colorist’s work; present terms, methods and goals of color grading featured in the literature will now be presented. The historical overview above was presented in the context of motion picture film, which was clearly separate from video. With the development of digital technologies these two mediums have more or less merged together, as well as their respective terminologies. In the past, color correction was associated with working on video while grading related to film. Although, Van Hurkman argues that color correction refers to a more technical process to fix problems with the image and making it fairly neutral, while grading refers to a more intensive process of creating an appropriate look for the image; which he defines as “visible stylizations of an image with the intent to create a clear mood of reference”. The two terms of color correction and color grading will be used interchangeably, which Peters means that they also are, even if they aren’t always used that way. It’s relevant to point out that color correction can refer to both an individual correction and also a collection of adjustments, while grading implies several corrections which together realizes the look of the image. This is something Van Hurkman brings to attention with a quote from colorist Joe Owens: “correction is a sword fight while grading is the war”.

Primary and secondary adjustments

Hullfish distinguish two processes from color correction; primary and secondary color correction. He writes that these two processes probably always will be referred to as two distinct processes, even though “technology itself is starting to change the perception of how and why these two processes are used and when the colorist moves from one process to another”. Hullfish describes primary color correction as general adjustments which are applied to the entire image, while secondary adjustments are applied to a specific portion of the image. Inhofer confirms that primary color correction typically affects the whole image, and also explains that keys and shapes can be utilized to isolate specific parts of an image; to go from a broad primary adjustment to a more specific secondary adjustment.

Van Hurkman refers to secondary color correction as an essential part of any colorist’s toolkit; which, as the name suggest, “are generally made after the primary color correction”. Peters argues that this division is something that’s inherited from older technology, which is still present today because of its implementation in modern color correction software.

Goals and methods of the colorist

In *The Art of Short Form Content Cook* identifies three general steps of color correction. The first one is to balance the image so it looks “correct”, while the second step is to manipulate the image so that the viewer knows where to look. The third and final step according to Cook is to use color to evoke an emotional response amongst the audience, which he refers to as being the more mystical and subjective part of color correction; where colorists who excel at this often has somewhat of a celebrity status in the world of postproduction. Inhofer (who applies a more tool oriented perspective) defines the work of the colorist by the use of either primary or secondary adjustments, where each contains three functions. Normalizing the image, which he calls a base grade, followed by shot-matching and the creation of a look makes use of the primary toolkit. Fissoun differentiates between the terms normalizing and balancing, and points out that the former is about adjusting luminance and the latter color. Inhofer then lists fixing specific problems, achieving creative goals and controlling the viewers eyes as functions related to secondary adjustments. Van Hurkman also lists six labors of the colorist. The first one is to correct errors in exposure and color. This is followed by making key elements like people or products look right and balancing shots in a scene so they match each other, which is commonly referred to as shot-matching and done to enhance the continuity of a scene. Lastly, Van Hurkman lists creating style, depth, and adhering to quality control standards. The reference manual for the color correction software DaVinci Resolve 15 also lists six different goals of color correction. The first goal is to maximize the look of the image by adjusting exposure, contrast, tone and saturation, which is done to realize the intent of the director and cinematographer. Next is to guide the viewer’s eye by emphasizing what’s important in the image. To play with the expectation of the audience by isolating and adjusting specific parts of the image is also a part of the colorist’s toolbox. This can be done by changing a subject’s skin tone, the hue of foliage or the sky, or another specific object. Finally, the reference manual also lists shot-matching, adding style and adhering to quality control standards. Table 1 provides an overview of the above mentioned goals and methods, and is organized into four categories. Even if the sources occasionally use different words, they refer to similar functions. In the category Secondary tools the goals are more diverse than those of the other three categories. What brings these goals together under the same group is the use of keys and/or shapes.

	Cook	Inhofer	Van Hurkman	DaVinci Resolve
Base grade	Balancing the image	Normalizing	Correcting errors in color and exposure	Maximizing the look
Shot-matching		Shot-matching	Balancing shots in a scene to match	Balancing scenes
Secondary tools <i>(using keys and/or shapes)</i>	Guiding the eye	Fixing specific problems, Achieving creative goals, Controlling the eye	Making key elements look right, Creating depth	Emphasizing what’s important, Playing with expectations
Creative grade	Evoking emotion	Creating a look	Creating style	Adding style
			Adhering to quality control standards	

Adhering to quality control standards didn’t fit into any of the four categories and was therefore placed on a separate row. Harris and Walker⁴⁰ differentiate between color grading and color

management; where, in the latter group, adhering to quality control standards could potentially fit in. In their presentation they also mention how color management becomes more important today due to multiple source formats and deliverables, which they discuss in the context of HDR.

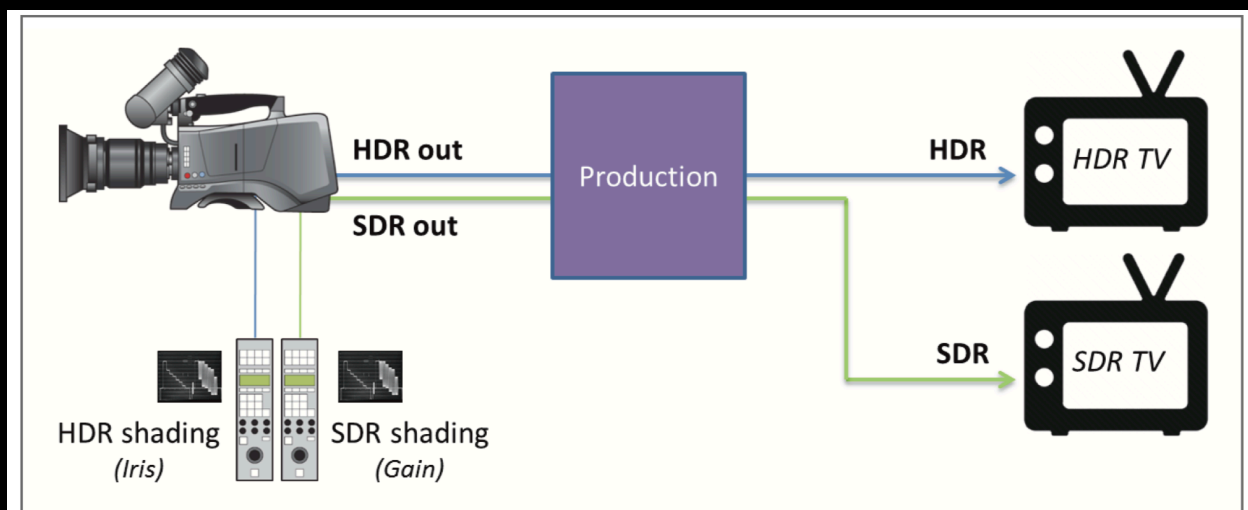
HDR/SDR Compatible Workflows

A full parallel HDR and SDR workflow is the easiest way to produce both HDR and SDR at the same time, from a single camera system through the full production chain.

In a full parallel scenario, the camera delivers two simultaneous signals, one HDR and one SDR. In this simultaneous HDR and SDR production workflow, the camera lens iris is set to HDR output and the SDR gain is used to control the SDR output to the required level. As long as the scene lighting does not change significantly, the lens iris will not need to be changed and a moderate variation of the lighting condition will be well inside the headroom of the HDR output.

In comparison, SDR does not offer this additional headroom, so a much more precise adaptation of the sensitivity will be required. Challenges in this workflow include the simultaneous shading of HDR (which requires less work due to the higher dynamic range) and SDR signals (which requires more work due to the limited dynamic range) as well as the handling of both signals separately through the full production chain. This translates into a more complex and expensive workflow which might be accepted for certain applications and/or for an intermediate period of time, but might not be acceptable in the long term. It's akin to the early days of HD, when remote sports productions would use separate SD and HD trucks and crews.

An alternative to the full parallel HDR/SDR workflow is to use only native HDR signals from the camera and then perform an HDR-to-SDR conversion somewhere during the production. While this simplifies the workflow and reduces the amount of resources required, in contrast to the parallel HDR/SDR workflow, there is no separate control for the SDR and HDR outputs, and the SDR gain cannot be independently controlled from the HDR signal. As a result, the success depends on the quality of the HDR-to-SDR conversion under all types of lighting conditions.



Static Conversions vs. Dynamic Conversion

Broadcasters and content creators planning to add HDR to their workflows have two types of HDR-to-SDR conversions to consider: static conversions with a fixed LUT or a selectable LUT (the “look up table” that is the template for the conversion), and dynamic conversions that analyze the picture content and apply content-based settings.

Static conversion offers always predictable results and allows controlling the offset between the SDR and HDR depending on the users' requirements. On the other hand there will be limitations especially during very challenging lighting conditions to get the best SDR and HDR outputs at the same time.

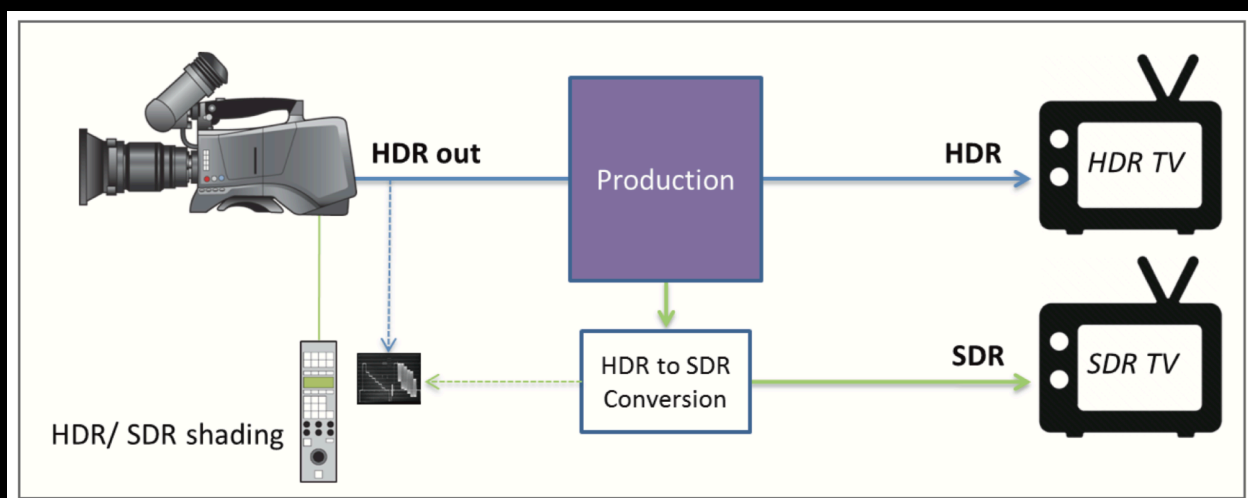
Dynamic conversion appears to offer a larger capability of processing inside the production chain, although automatically adopting the look of the image might not be acceptable for all users or in all cases.

Versions of both types of converters are available today and initial testing in live environments continues to deliver very promising results. No doubt a variety of solutions are available today and even more will become available in the near future covering a wider range of typical live applications.

Operational aspects for HDR

Before preparing an HDR production, a few things need to be cleared up front:

- Which version of HDR is requested, HLG or PQ?– If PQ will be used, what is the clip level, 500 – 10,000 Nits?
- What luminance and color difference signal representation is re- requested?
 - » Typically YCbCr is used.
- What code value range is requested, narrow or full?
 - » For most live applications, it's narrow.



- What kind of SDR downmapping is requested, static or dynamic?
 - » Typically it's set to 10,000 Nits.
 - If HLG will be used, what is the SDR set point, 50% or 75%?
 - » After latest changes to the recommendations, it's typically 75%.
- What color gamut is requested, BT.709 or BT.2020?
 - » Typically it's BT.2020 for HDR productions.
- » Depending on the downmapping, shading needs to be done on HDR or SDR

In most cases, at least for the time being, not all sources will be able to deliver the signal formats as required for the production. Since HDR and SDR or PQ and HLG or BT.709 and BT.2020 color gamut are not compatible, all signals have to be up/downmapped to the selected production format layer

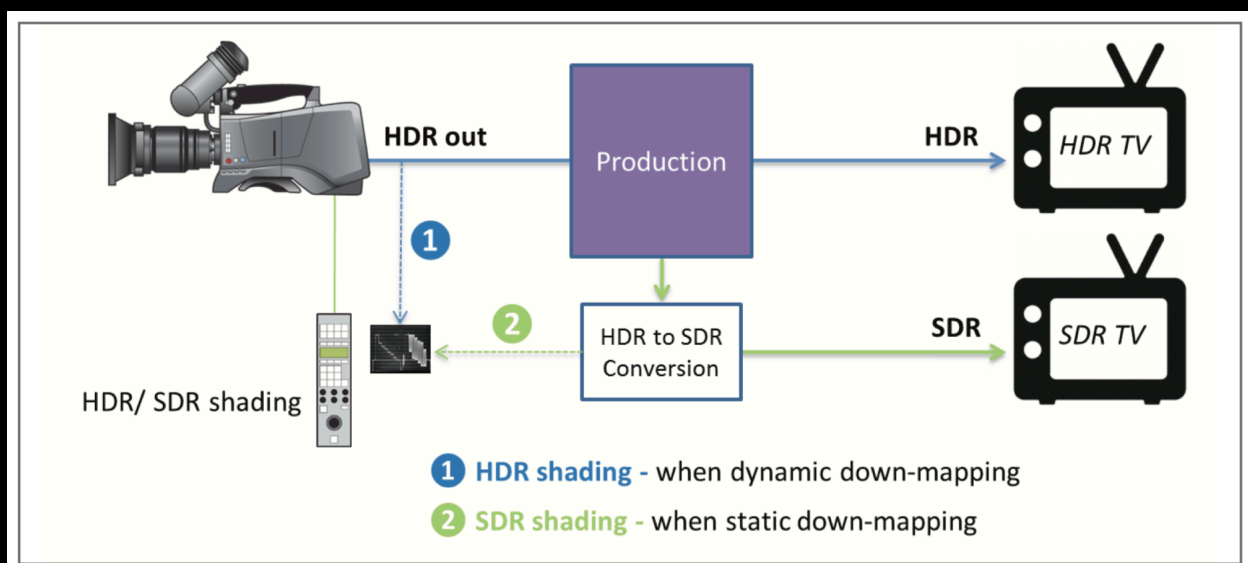
Quite a large number of up/downmappers might be required and the requirements must be carefully calculated upfront for an HDR production to avoid any shortage. In addition, it is highly recommend to always keep a few channels of up/downmappers on spare for unexpected requirements.

On the distribution side of the equation, HLG offers broadcasters a means to deliver HDR content to viewers watching new televisions equipped with HLG decoders while continuing to serve viewers using legacy SDR TVs with the same signal. Another approach is to simulcast content in HDR and SDR on separate channels or as on-demand content. For instance, Netflix is streaming HDR content in HDR10 and Dolby Vision while maintaining its SDR catalog for on-demand streaming. Similarly, Samsung and Amazon began streaming about a hundred TV shows using HDR10+ in December 2017 while continuing to offer SDR versions to customers.

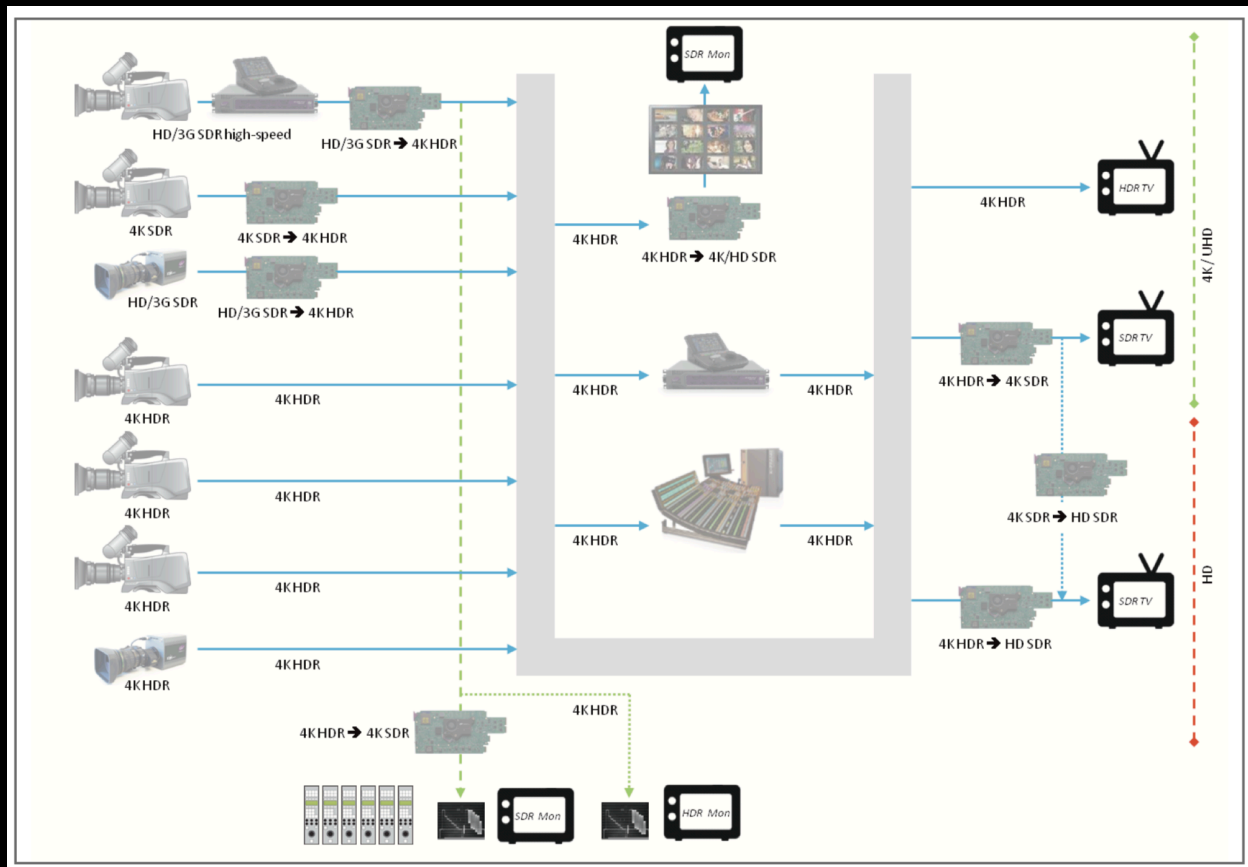
Mastering display

Having discussed the unpredictable accuracy of the average end-viewer display, we can start to understand just how important it is to have an accurate mastering display.

The reason is simple: any inaccuracies in our mastering display will compound with those of the end viewer display, creating a far larger maximum error than either display has individually.



This demand for accuracy in professional mastering displays is the key reason why they're significantly more expensive than consumer models—they have lower tolerances and are harder to manufacture. Even with a professional display, it's necessary to regularly calibrate to account for "drift" as the display ages. Since this is the only display we have direct control over, we need it to be as dead-on as possible. A proper grading monitor made for mastering HDR content is very expensive; where the cheaper Eizo Prominence CG3145 has a suggested retail price of \$30,995 US¹²⁵, and the Flanders Scientific XM310K is priced at \$45,000 US



Encoding and metadata

Video and image files carry not only picture and audio data, but metadata, which is exactly what it sounds like: data about data. Metadata for a given file can contain hundreds of entries, anything from the program title to the name of the camera assistant who loaded the magazine.

It can also contain information about the color gamut and/or gamma curve of the content, which some devices and software will detect and use to make image compensations. A simple example would be an image mastered in Rec. 709 unintentionally encoded with metadata identifying it as a Rec. 2020 master. Certain software and/or hardware will read this Rec. 2020 metadata and make a compensation intended to best reproduce a Rec. 2020 master gamut within a Rec. 709 display gamut.

Strategies for more consistent images

Articulate your deliverables

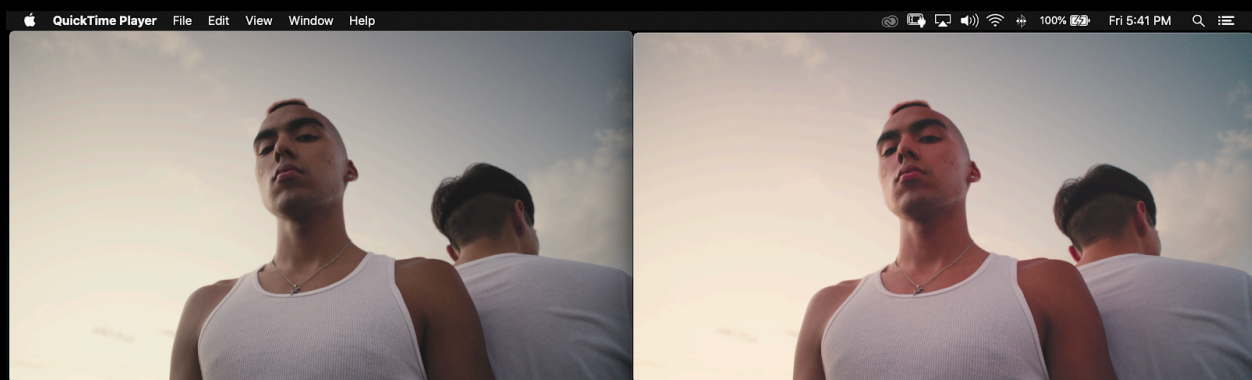
Take an inventory of the outlets for which you're delivering. If your content is destined for screens of more than one gamut/gamma standard, you need more than one gamut/gamma encoded deliverable.

For example, if I need a Quicktime from which to strike a DCP, as well as a Quicktime for Vimeo upload, these will need to be separate assets: one targeting P3 Gamma 2.6 (for DCP), and one targeting Rec. 709 Gamma 2.2 (for Vimeo). If you try to deliver the same file to both outlets, you'll end up getting at least one bad rendering.

Use a color-managed workflow

Using a color-managed workflow, you can target alternate gamut/gamma pairs for your renders without needing to make creative adjustments or guesses. It also means you can target display spaces without needing to monitor for them.

For example, if I do a color-managed grade on a film targeting my Rec. 709 display, I can later switch that target to P3 Gamma 2.6, and even though I'm not able to monitor that color space, I can be confident it will look the same on a P3 Gamma 2.6 projector as it did on my Rec. 709 display.



These images are identical in every way except for their metadata tags — the image on the left is tagged Rec. 2020, which causes Quicktime Player to saturate it in an attempt to give the most accurate reproduction possible. (Image courtesy of Director Leo Aguirre, BL&S Films)

Targeted web deliverables

Perhaps the trickiest outlet for ensuring accurate renderings is the web, simply because it encompasses so many platforms and devices whose display standards can significantly vary from one another.

Unfortunately, there's no cut and dried answer for the perfect gamut/gamma target for web; but in my experience, you're better off targeting sRGB's gamma of 2.2, rather than the TV standard of gamma 2.4 or BT.1886. This will get you a closer match to your master on nearly any screen that's not a TV — phone, laptop, tablet, etc.

Be mindful of encoded metadata

As discussed above, incorrect or incomplete metadata can cause a wildly inaccurate reproduction of your images. At a minimum, you should use a metadata inspector/editor to ensure that there are no incorrect entries, such as a flag for Rec. 2020 when you're mastering gamut was Rec. 709.

Additionally, you can ensure that there are no absent entries pertaining to mastering gamut or gamma. Digital Rebellion's Pro Media Tools is a great product for reviewing and editing metadata as needed.

A difference between SDR and HDR EOTF's are their relative vs absolute nature. The actual output brightness of a bit value in SDR will vary depending on the display and its settings, while in HDR (PQ) a bit value refers to a specific brightness level output. The problem with the absolute approach is the different specifications between the consumer TVs and professional grading monitors regarding color gamut coverage and especially peak brightness, this is where the metadata comes in Metadata "conveys the color volume of the mastering display and the luminance of the content [...] described by the chromaticity of the red, green, and blue display primaries and white point of the mastering display, plus its black level and peak



The web is challenging for color grading, as different devices and platforms can render colors very differently, even when watching the exact same video.

luminance level” The metadata also includes information about the Maximum Content Light Level (MaxCLL); which defines the luminance of the brightest pixel, and the Maximum Frame-Average Light Level (MaxFALL); which is the highest frame average brightness of all frames in the content. Static metadata uses the same values for the entire duration of a given program, while dynamic metadata changes as often as needed.

There are several systems that have emerged to address the needs of HDR.

Dolby Vision

For a vivid image, which HDR promises upon delivery to the display, the challenge has been to provide enough useful steps within the color and luminance range to provide smooth gradations and satisfy the viewer. Dolby Vision is a new way of defining how a picture looks on a TV. It is a "display referred" system where each code value defines a specific brightness (i.e. there is no user brightness control). Dolby Vision has a method to send 12-bit data down a HDMI cable utilizing 4096 shades for smooth gradations. For contemporary production pipelines, the challenge is that most pipelines are 10-bit with 1024 shades or steps available. HDR 10, Dolby Vision and HLG have been designed to solve this dilemma by efficiently choosing to utilize curves across the luminance and color ranges that minimize perceptual banding and achieve results that are closer to the human vision system, enabling incredible results.

- Proprietary format
- SMPTE ST-2084 Perceptual Quantizer (PQ)
- Metadata: Dynamic per ST 2094*
- Electro-optical transfer function, up to 4K resolution
- Playback bit depth: 12-bit
- Up to 10,000-nit maximum brightness (mastered to 4,000-nit in practice)
- Wide Color Gamut (WCG) color space (ITU-R Rec. 2020)

HDR10

HDR10 is a Media Profile defined by the Consumer Technologies Association (CTA).

HDR10 uses SMPTE ST-2086 "Mastering Display Color Volume" static metadata to send color calibration data of the mastering display, as well as MaxFALL (Maximum Frame Average Light Level) and MaxCLL (Maximum Content Light Level) static values. HDR10 is an open standard supported by a wide variety of companies.

- Open format
- SMPTE ST-2084 Perceptual Quantizer (PQ)
- Minimum Signal Interface: HDMI 2.0a
- Playback color representation BT.2020 (same as Dolby Vision)
- Playback transfer function: ST 2084 (same as Dolby Vision)
- Metadata: Static per ST 2086
- Playback bit depth: 10-bit

HDR10+

HDR10+ builds from HDR10*, adding dynamic metadata and more, allowing for scene by scene or even shot by shot variation. The standard allows for recognizing a display's color space and more. Samsung is the primary driver for this evolution of the original standard.

HDR10+ has recently been adopted by a range of companies and resulted in the creation of the HDR10+ Alliance with members comprising Amazon, Samsung, Panasonic, 20th Century Fox, and Warner Bros and more.

- SMPTE 2094-40 Dynamic Metadata for Color Volume Transforms (DMCVT)
- Color transforms optimized for each scene, and each display
- Metadata: Dynamic per ST 2094
- Metadata tracks carry supplementary color grading information
- Minimum Signal Interface: HDMI 2.0a
- Playback color representation BT.2020 or BT.709 dependent on display
- Playback transfer function: ST 2084
- Playback bit depth: 10-bit

HLG

HLG (Hybrid Log Gamma) is an HDR approach that was jointly developed by the BBC and NHK. HLG uses relative luminance values and builds on the BT.2020 standard for a wider color gamut and adds PQ and HLG dynamic range for a standard known as BT.2100. The HLG standard is royalty free and is compatible with SDR displays as well as HDR displays. HLG is supported by HDMI 2.0b*, HEVC, and VP9. The key appeal of HLG is that HDR content can be reasonably viewed on SDR displays.

- Hybrid between Gamma and Log
- Metadata: None
- Relative Luminance
- Very similar to SDR curve
- Playback bit depth: 10-bit
- Linear light image data from a camera is mapped to the HLG OETF Curve and to the WCG color space BT.2020

*Since HLG has no metadata, HLG works with any HDMI spec and HDMI 2.0b is a clarification of HDMI 2.0a

Production and Distribution

On the distribution side of the equation, HLG offers broadcasters a means to deliver HDR content to viewers watching new televisions equipped with HLG decoders while continuing to serve viewers using legacy SDR TVs with the same signal. Another approach is to simulcast content in HDR and SDR on separate channels or as on-demand content. For instance, Netflix is streaming HDR content in HDR10 and Dolby Vision while maintaining its SDR catalog for on-demand streaming. Similarly, Samsung and Amazon began streaming about a hundred TV shows using HDR10+ in December 2017 while continuing to offer SDR versions to customers.

When it comes to television production, producers will find themselves working in a mixed SDR-HDR ecosystem for the foreseeable future, as economics and depreciation schedules dictate that they wring as much use out of their legacy

When it comes to television production, producers will find themselves working in a mixed SDR-HDR ecosystem for the foreseeable future, as economics and depreciation schedules dictate that they wring as much use out of their legacy cameras, switchers, and other devices as possible. As a result, they will be required to convert SDR to HDR, one flavor of HDR to another, and HDR to SDR depending upon the unique circumstances of their production.

There are many other production hypotheticals that will also require signal conversion. Understanding the specifics of each at this time, however, is probably less important than realizing that the demand for HDR content is on an upward trajectory while the technology used to produce that content will be a mix of SDR- and HDR-native equipment for quite some time.

Keeping in mind the details we discussed, the OTT platforms' requirements are stated in their support page. They have several recommendation from the cameras used to the final acceptable deliverable format.

Create your Dolby Vision grade, with your tools

Dolby Vision being a Proprietary format and backward compatible with other HDR formats we look into the process of creating one.

The color-grading workflow for home

Creating the Dolby Vision master grade begins with creating the look for your master HDR grade first, using a compatible professional color-grading application with the working color space and EOTF of the HDR mastering monitor set to SMPTE ST. 2084 (or PQ) for HDR. In order to see the high brightness levels, deep blacks and wide color gamut, it's important to rely on an HDR reference grade mastering monitor that meets the Dolby Vision minimum specifications.

When the master HDR grade is done, you simply run the automated Dolby Vision analysis to generate (L1) metadata that describes the Minimum, Mean and Maximum pixel value for each shot.

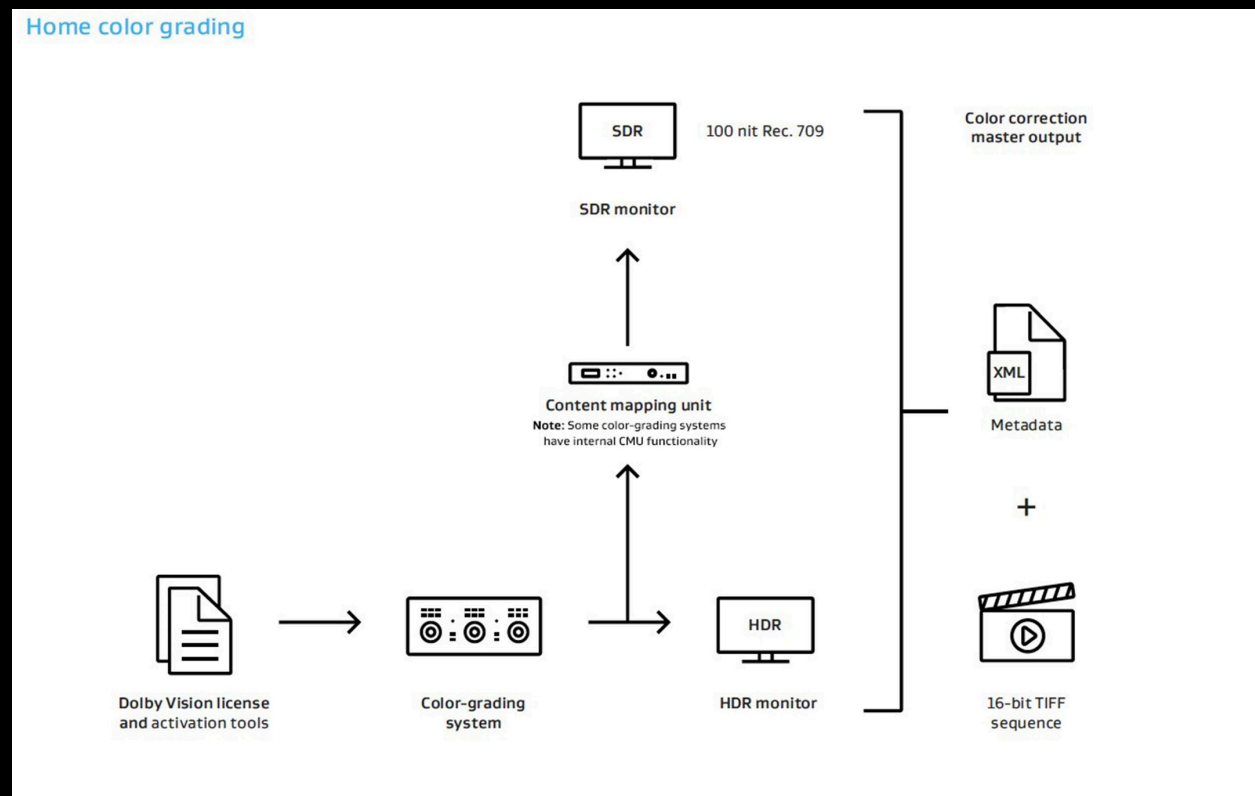
This function is integrated into many Professional video software tools that support Dolby Vision. The Dolby Vision metadata maps the content in real time, to an SDR Rec. 709, at a standard brightness of 100 nits (as well as other HDR displays). Some content owners may want to make adjustments to the mapping on a shot by shot basis. Using a Colorgrading system that supports Dolby Vision, you can check the mapping and if desired, make adjustments or 'artistic trims' (L2 metadata) to the mapped image, to best deliver the creative intent of the HDR master across a wide range of devices. Analysis Metadata creation is included in many professional video tools. Artistic Trims (L2) are available in colorgrading systems and requires a license from Dolby.

Dolby Vision Compatible Color-grading Systems

- Blackmagic Design DaVinci Resolve
- FilmLight Baselight
- Digital Vision Nucoda
- GVG Quantel Rio
- Autodesk Lustre
- SGO Mistika

The New TV Standard: Dolby Vision

With the expanded luminance of Dolby Vision, each scene is created with a range of colors much closer to what the eye actually sees.



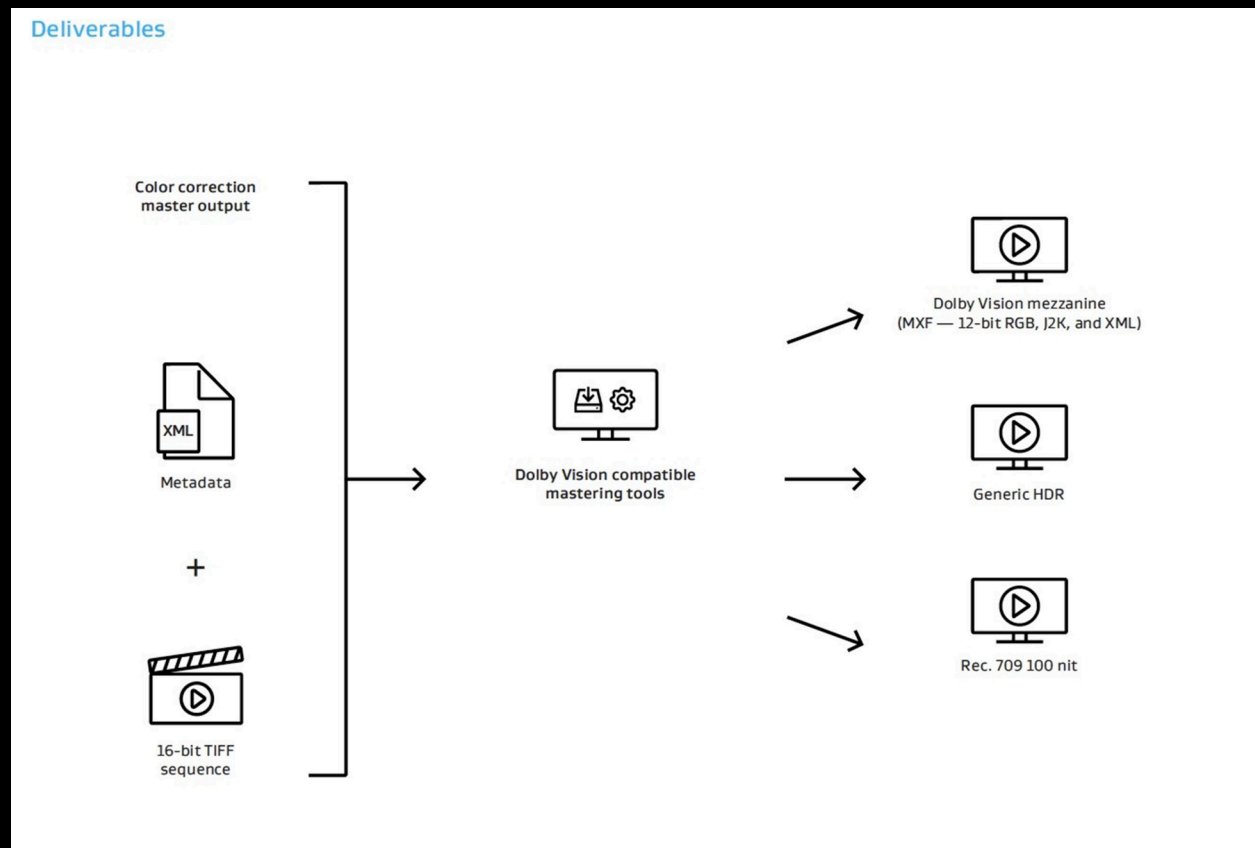
The mastering and QC workflow for home

Create all your high-quality deliverables using the Dolby Vision master grade. Once the color grading is complete, you export the final grade in the highest image quality, typically a 16-bit TIFF or other uncompressed image container and the accompanying Dolby Vision dynamic metadata as an XML file.

These archive masters are then imported into a Dolby Vision mastering tool to create the mezzanine file (or Service Master). Mezzanines are typically stored in a JPEG2000 or ProRes 4444, 4444XQ image container. For this process, you can use the Dolby® Professional Tools, or supported 3rd party mastering solution.

Professional Video tools for Dolby Vision

- Colorfront Transcoder
- MTI Film Cortex
- Marquise Technology MIST
- Rohde & Schwarz Clipster



That's a Wrap

Television has begun a grand journey towards delivering better pixels, something viewers increasingly will expect. In less than five years, annual shipments of HDR-enabled televisions to consumers will climb to 245 million, according to ABI Research.

Programmers, producers, and commercial clients who ignore the accompanying demand for HDR content do so at their own peril. They must ask themselves whether it's more desirable to stand out from the rest today because their pixels are better or to stand out in the not-too-distant future because their inferior pixels are surrounded by an ocean of HDR/WCG content. Either way, viewers will notice.

High dynamic range and wide color gamut are the twin foundational pillars to creating better pixels. Together, they produce images that are more pleasing by more closely aligning with the upper reaches of the human visual system.

Producers choosing to walk the path to better pixels—even in these early days—can do so with confidence. They can be assured that taking on HDR and WCG today will not require the complexity and expense of two separate production workflows. Equally important, they can be certain that doing so will put their content in sync with where consumers and their televisions are headed. Better pixels are at hand. The time to begin planning how best to leverage them in production is upon us.

Glossary

ACES (Academy Color Encoding System) Color management system standardized by A.M.P.A.S. which defines a wide gamut scene-referred linear working space, and the transforms needed to map to that from other color encodings. ACES also defines the transforms to display ACES image data on various standard displays, as well as metadata to communicate creative intent and other information. The acronym ACES may also stand for Academy Color Encoding Specification, the ACES2065-1 color encoding.

ASC CDL American Society of Cinematographers Color Decision List. See Appendix 4.5 for more detail. A list of color parameters that can be shared as XML or within EDL files and applied in most color grading systems.

Block Compression A form of lossy texture compression, usually hardware accelerated, which operates on fixed resolution blocks. Typically these blocks are 4x4 with compression ratios of 4:1 or 8:1, but this is not always the case. Follow the links in the "Block compression" part of Appendix 4.14.4 for more details.

Bokeh The appearance of the out of focus part of an image, characterised by discs with the same shape as the iris of the lens, which are most apparent around bright sources.

BT.2100 ITU-R BT.2100, referred to as BT.2100 or Rec. 2100 is the main ITU standards document for HDR. It defines the OETF and EOTF for both PQ and HLG, as well as specifying an OOTF for each. It also gives the spatial and temporal characteristics (pixel dimensions and frame rate) for HDR, and details of Y'CbCr encodings and SDI transmission.

CSI Colorist Society International (CSI) strives to advance the craft, education and awareness of the art and science of color grading and color correction. ColoristSociety.com

Code Values Discrete values for 8, 10, 12 or 16 bit integer encodings. For example, a 10 bit image has 1024 code values, numbering from 0 to 1023.

Color Management Techniques used to ensure that the intended appearance of an image is reproduced in the same way everywhere it is viewed. This may include Color Appearance Modeling to account for differences in viewing environment and display brightness, or may simply ensure identical absolute display colorimetry on all displays (within the limitations of a particular display).

Digital Intermediate (DI) The process of manipulating color and other characteristics of digital images prior to theatrical release. DI is currently used for both film and video markets and uses software with a data workflow rather than hardware and a tape-based pipeline. The term was originally used by Kodak to define a workflow from film negatives to digital manipulation and back to film. (from <http://www.finalcolor.com/glossary-1/>). It is commonly used in VFX to mean color grading and finishing, which is a narrower definition than the term originally intended but is the preferred meaning in this document.

Defocus The optical effect caused by light rays converging in front of or behind the plane of the sensor in a camera. Simulated in post by convolving the image with a kernel with the same shape as the lens iris, and size proportional to the desired defocus.

Display-referred see output referred image state

Dmin Minimum density on the characteristic curve of a film stock's response.

Display Transform See Output Transform

D65 The CIE standard daylight illuminant (approximately 6504K, although not actually on the Planckian locus) used as the white point for televisions, HD video monitors and most computer displays. The CIE defines the chromaticity coordinates of D65 as (0.31272, 0.32903) but Rec. 709 and many other color encoding specifications use a form rounded to four decimal places.

EDL (Edit Decision List) Text file used to convey the shots used in a cut, and their positions in the timeline, between different editing and grading systems. Historically EDLs were used to transfer edits from non-linear offline editing systems (and before that tape based offlines) to the high quality multi-machine tape editing systems. Various forms exist, but the most commonly used are based on the CMX3600 format.

EOTF (Electro-Optical Transfer Function) The mathematical mapping between the output-referred pixel values and the luminance on a particular display.

Encoding gamut The gamut chosen by a camera manufacturer for use in mapping the specific response from their sensor and filter's spectral response curves to chromaticities.

Gamut The range of colors able to be represented by positive values in a particular color encoding. Gamuts are often shown as triangles on a CIE xy or u^*v^* plot, but these, in fact, are only a projection of the volume of the color gamut, as the luminance component is not shown.

HDR (High Dynamic Range) HDR has multiple meanings: It is a set of techniques that allow a greater dynamic range of exposures or values (i.e. a wider range of values between light and dark areas) to be handled than normal digital imaging techniques. HDR image capture encodes a wider than normal dynamic range from the scene, without applying any tone mapping or other processing to make the result suitable for direct display. This may be done either by combining multiple bracketed exposures from a lower dynamic range device, or by using a camera which has an inherently high dynamic range. HDR image processing is the most common usage in this document and within the VFX community. It is used to mean working with wide dynamic range scene referred data in a manner which preserves the dynamic range, either by using unclamped floating point linear processing, or a logarithmic representation of the linear image data. HDR display is a set of new technologies which use screen brightness significantly greater than conventional displays. These displays use either a Perceptual Quantizer (PQ/ ST.2084) or Hybrid Log Gamma (HLG) EOTF to display images. Colorists, consumers, and distributors are likely to mean HDR displays when they use the term HDR. Note: The term HDR is also used to refer to images which use local tone mapping to prepare a high dynamic range image for a standard dynamic range display, producing the unnatural looking images often seen online. This document does not ever use the term in that sense.

HLG (Hybrid Log Gamma) HDR encoding developed by the BBC and NHK, and specified in ITU-R BT.2100, which defines an OETF and EOTF, producing a system gamma OOTF. Unlike ST.2084, HLG is a relative standard, with output brightness dependent on the black and white levels of the display.

IDT see Input Transform

Input Transform Transform provided by a camera manufacturer, or specified by A.M.P.A.S, to map a source encoding to ACES2065-1.

IRE While strictly a unit of measurement of the voltage in composite video signals, the term IRE is commonly used to refer to the level of any signal between black and white as a percentage, with 0 IRE being black and 100 IRE white. It is often used to abstract the encoded signal from the code values used to represent it, so in "Legal Range" 0 IRE is 10 bit 64 and 100 IRE is 940. In "Full Range" 0 IRE is 0 and 100 IRE is 1023.

Look Management Techniques used to communicate the intended appearance of an image, including creative as well as technical transforms, so that the same Look can be seen by all involved.

LMT See Look Transform

Look Transform In an ACES pipeline a transform applied downstream of grading operations to impart an overall look to a show or scene.

LUT Box Hardware image processing device capable of applying transforms implemented as LUTs to SDI or HDMI signals.

Munsell Renotation System System developed by Professor Albert H. Munsell in the early 20th century which describes colors in terms of perceptually uniform hue, value (lightness) and chroma.

NTSC (National Television System Committee) Organization in the US founded to standardize the format for television broadcasts. Commonly used to refer to the color television standard defined by that committee. NTSC is a standard definition interlaced format with 29.97 frames (59.94 fields) per second, but the term is often loosely used to refer to any 30 or 29.97 fps video format, to distinguish it from formats based on the European PAL 25 frames (50 fields) per second system. The precise frame rate is in fact $30 \times 1000/1001$.

ODT (Output Device Transform) See Output Transform

OECF (Opto-Electronic Conversion Function) See OETF

OETF (Opto-Electronic Transfer Function) The mathematical mapping between scene light and scene-referred pixel values.

OOTF (Optical to Optical Transfer Function) The mathematical mapping used to convert between scene-referred and display-referred image data for the purposes of picture rendering.

Output-referred image state Image state associated with image data that represents the colorimetry of the elements of an image that has undergone color rendering appropriate for a specified real or virtual output device and viewing conditions.

Output Transform The ACES term for what is also called a Display Transform, a mathematical mapping which transforms scene-referred image data into display-referred, targeting a particular class of display. This transform is not purely colorimetric, but includes tone mapping and picture rendering.

Picture rendering A modification of the image contrast to create a perceptual match between scene and display.

Protect To choose operations, formats and processing pipelines that do not clip intensities, gamuts or image content, even where not noticeable in the current form, in order to preserve maximum flexibility downstream.

PQ (Perceptual Quantizer aka Perceptual Quantisation) EOTF standardized by SMPTE ST.2084 for HDR displays. PQ is an output referred absolute encoding of HDR display luminance, covering the range from 0 to 10,000 cd/m². Originally designed by Dolby under the name PQ for Perceptual Quantizer, because each PQ code value increment represents an equal increase

in perceived brightness throughout the range. Quantizer and Quantization are used interchangeably in discussions about PQ. Quantizer was used originally by Dolby, but Quantization become more commonly used later on, as in the ITU-R BT.2100 document.

Rushes Original unmodified files recorded by a camera. Historically the term referred to the exposed negative or recorded video tape, but the term is now used to refer either to the files on the camera media, or a copy of those files made on set or in a data lab.

Scene-referred image state Image state associated with image data that represents estimates of the colorimetry of the elements of a scene. Scene-referred image data typically represents relative scene colorimetry estimates. Absolute scene colorimetry estimates may be calculated using a scaling factor.

SDR (Standard Dynamic Range) Term used to describe display systems not capable of high dynamic range; normally used in contrast to HDR displays. Also used to describe output-referred image data restricted to the 0-1 range (or integer equivalent) suitable for display on SDR displays.

Small Floats These are floating-point representations with fewer bits than the usual 16-bit implementations. Overall range is similar to 16-bit since the 5-bit exponent is preserved, but a reduced size mantissa reduces precision and the sign bit is dropped. Links with more information be found in the "Small float formats" part of Appendix 4.14.4.

ST.2084 See PQ

Trims Grading adjustments made after grading of the primary master is complete, in order to make the image suitable for a secondary deliverable.

View Transform See OOTF

Working Space The color space in which a particular set of operations is performed. The choice of working space has a significant effect on the way such operations act. Some spaces may be more appropriate than others for certain operations, and some may be entirely unsuitable. A color pipeline is likely to use different working spaces at different points. They are often transient color spaces, existing only within a grading or compositing system, with transforms to and from the working space applied either side of the operations.

Bibliography

This entire Paper is created with information collected and compiled from multiple sources of webpages and videos. Following are the links for the same, if you may need to explore further. All the information and copyright belongs to the respective owners.

https://en.wikipedia.org/wiki/High-dynamic-range_video
<https://www.finalcolor.com/colorblog/hdr-an-introduction>
<https://nick-shaw.github.io/cinematiccolor/description.html>
<https://workflow.frame.io/guide/#ch=delivery&s=end-point-determines-the-format>
https://blog.frame.io/2020/02/03/color-spaces-101/#What_is_color_space
<http://shootdatapost.com/blog/2014/5/16/aces-in-10-minutes>
<https://blog.frame.io/2019/09/09/aces/>
https://www.grassvalley.com/docs/WhitePapers/broadcast/technology/hdr/GVB-1-0676A-EN-WP_HDR_Live_Guide.pdf
https://www.aja.com/pdf/2018/aja_ebook_hdr.pdf
<https://blog.frame.io/2019/10/14/grading-mixed-delivery/>
<http://www.diva-portal.se/smash/get/diva2:1330365/FULLTEXT01.pdf>
https://www.mapyourshow.com/mys_shared/nab20/handouts/KosukeH30042020.pdf
<https://partnerhelp.netflixstudios.com/hc/en-us/articles/229150847-What-is-Dolby-Vision->

About Me

Arun Em DFT, Alumuni of FTIT 2001 Batch, Film editor, Blackmagic Davinci Resolve Official Trainer and Colorist since 2016.
email : editor.emarun@gmail.com
mobile : 9445001499

Each new medium is justified because it fills a lack or repairs a fault in its predecessor, because it fulfills the unkept promise of an older medium. (Typically, of course, users did not realize that the older medium had failed in its promise until the new one appeared.)¹⁹⁹
-Bolter & Grusin