Philips HDR technology – white paper

 Version 0.1 2015-08-21

1. Introduction

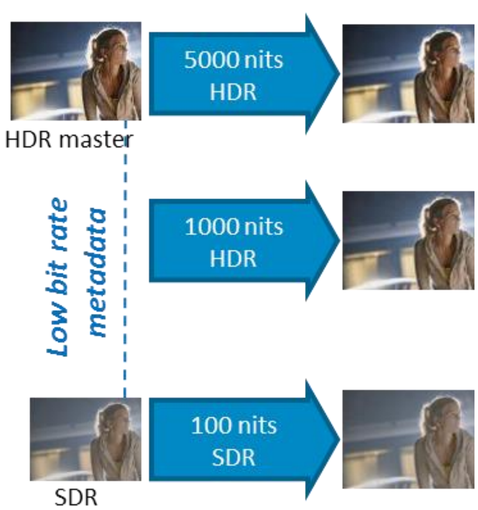
Current video formats are still based upon the display capabilities of CRT monitors. They cannot capture the detail of what professional cameras can record, high-end TVs can display, and what the human eye can see.

Philips has been working on higher brightness and High Dynamic Range (HDR) television since 2003. Initially, our research focused on the displays, resulting in a prototype/demonstrator LCD display with a peak luminance of 3 000 cd/m2 and a high contrast and wide color gamut using a local-dimming RGB LED backlight [1]. This display was shown to selected customers during the 2009 IFA in Berlin in a private area of the Philips booth and received favorable response. The difference with normal TVs is obvious even from a distance. With the right content, HDR offers a truly immersive experience.

Since then we have extended our studies to cover the entire HDR television chain, focusing on achieving high quality HDR using a single layer 10 bit video stream, while at the same time ensuring low decoder implementation cost.

2. Concept

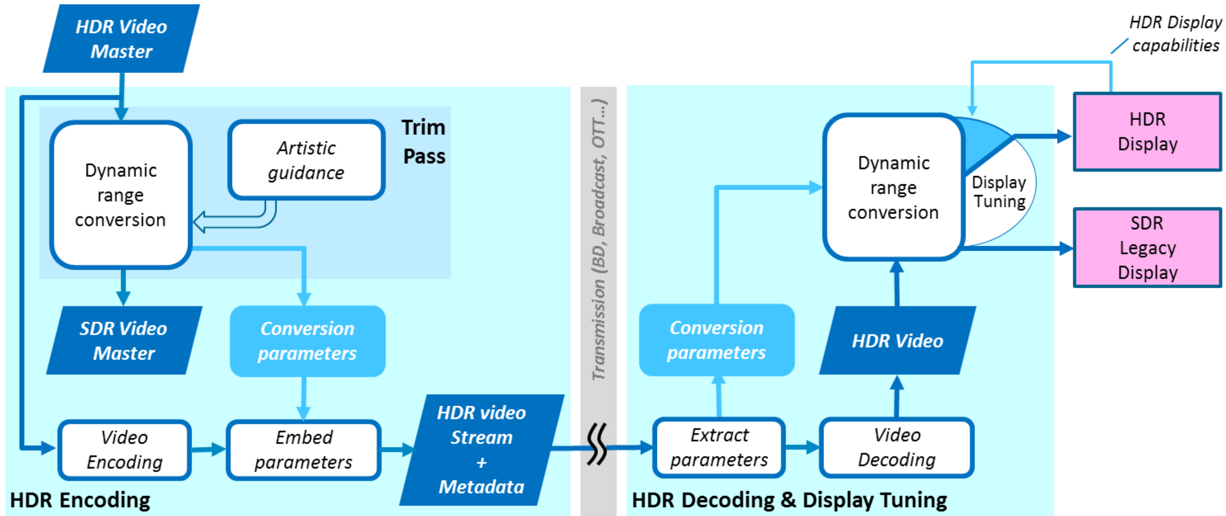
A new HDR video standard must address a wide variety of displays: from high-end to mainstream TVs, but also tablets and mobile phones. Ideally, all these different displays receive a video signal that has been graded for their own specific capabilities. In practice the optimum result for a wide variety of displays can be derived from only two reference grades: a high level HDR grade and a conventional SDR grade. These two grades are however related; it turns out that the difference can be captured with a set of conversion parameters. By adding this set of parameters to the video stream, both the HDR grade and SDR grade can be encoded in a single video bit stream, with only a negligible extra bit rate for the conversion parameters.



Using ‘smart interpolation’, the received video signal can now be tone mapped to the specific capabilities of the display, not only the high level HDR an the conventional SDR, but also anything in between.

[1] M. Hammer, “High-dynamic-range displays : contributions to signal processing and backlight control”, TU/e 2014. This Ph.D. thesis contains a description of the prototype HDR display developed at Philips Research (Chapter 2)

A functional block diagram of an example HDR video transmission system is shown below.



The process starts from an HDR video master. This HDR master must be graded on a high-dynamic-range display with high peak luminance and low black level. In the case of live production, such display should be used to monitor the HDR video master. The HDR video master is encoded. Typically, an HEVC video encoder will be used, with SMPTE ST 2084 as EOTF.

An SDR video master is derived in a semi-automatic way from the HDR master. First, an initial tone mapping is proposed. This may be based upon an average tone mapping curve, or may be created by an automatic analysis of the input image. A colorist then adds corrections to this initial to optimize the SDR master. This is typically done on a scene-by-scene basis, but may be done on a frame-by-frame basis if desired. This 'artistic guidance' process is typically done in a trim-pass.

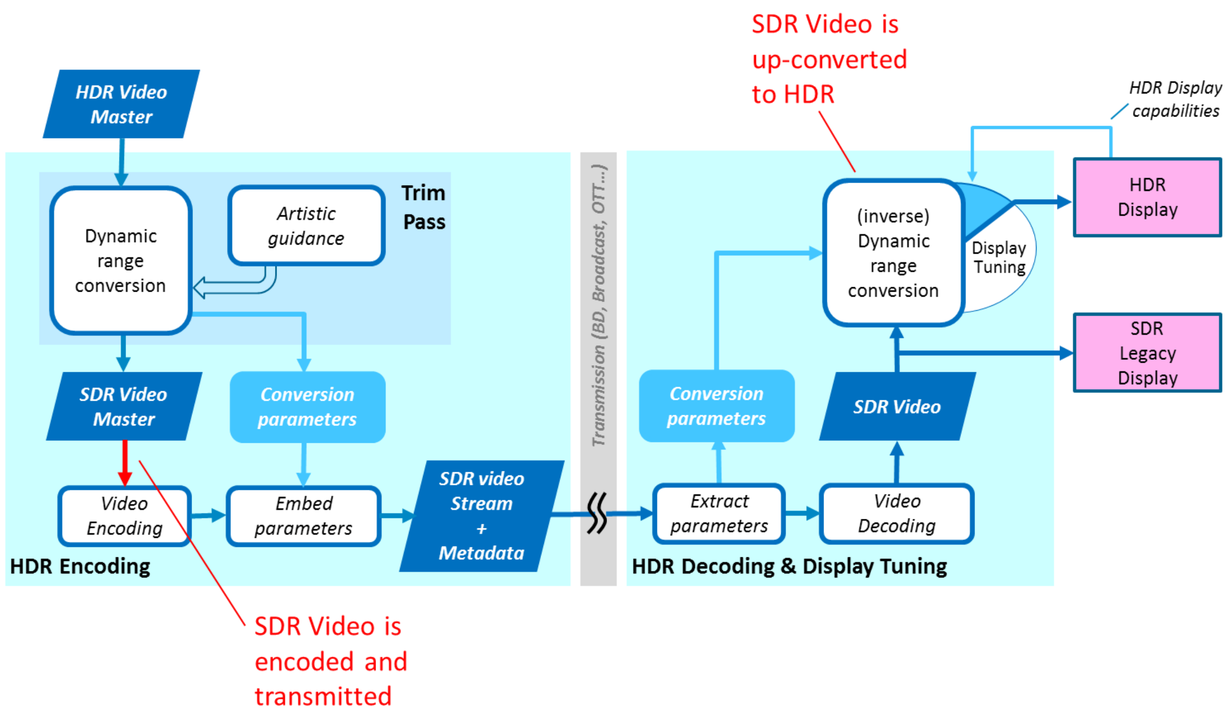
The corrections by the colorist are captured in conversion parameters, which are transmitted as content-dependent metadata along with the encoded HDR video using SEI messages embedded in the video stream. The content-dependent metadata is defined in SMPTE standard ST 2094-20. The characteristics of the display used for grading or monitoring, such as peak luminance and black level, are added as SMPTE ST 2086 metadata to the video stream.

At the receiving side, the video stream is decoded, producing HDR video. The conversion parameters are extracted from the SEI messages. If an SDR output signal is desired, exactly the same dynamic range conversion process as set by the colorist on the encoder side is repeated in the decoder, producing an SDR signal that is – apart from artifacts introduced by the video encoding-decoding – equal to the SDR master on the encoder side. If an HDR display is connected, the conversion parameters are used to produce an HDR output signal optimized for the specific display capabilities of the display. In case the decoder is built in a STB or BD player, the information in display capabilities can be sent to the HDR decoder using the HDMI 2.0a protocol.

The conversion parameters and HDR processing are agnostic to the applied video codec or format. They can be used together with 4K video or with FHD video.

3. SDR-compatible mode

For applications that require an SDR compatible transmission signal, a compatible mode is available, in which not the HDR master but the SDR master video is (HEVC) encoded and transmitted along with the conversion parameters. The functional block diagram is shown below.



The decoder receives the SDR signal that can go directly to an SDR display. An 'HDR-aware' decoder retrieves the conversion parameters from the video stream. Now, the *inverse* dynamic range conversion has to be applied to the video to obtain the HDR video.

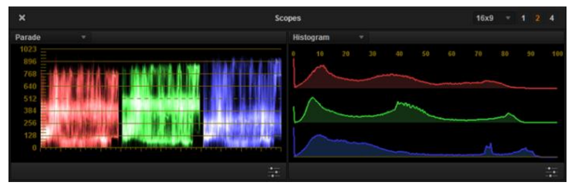
In this compatible mode, the conversion needs to be invertible. This puts some constraints on the HDR to SDR conversion, e.g. hard clipping is not allowed. In a recent MPEG test, it was shown that the Philips HDR system using this SDR-compatible mode of operation actually provides clearly better video quality than straightforward encoding of the HDR video. See the following link:

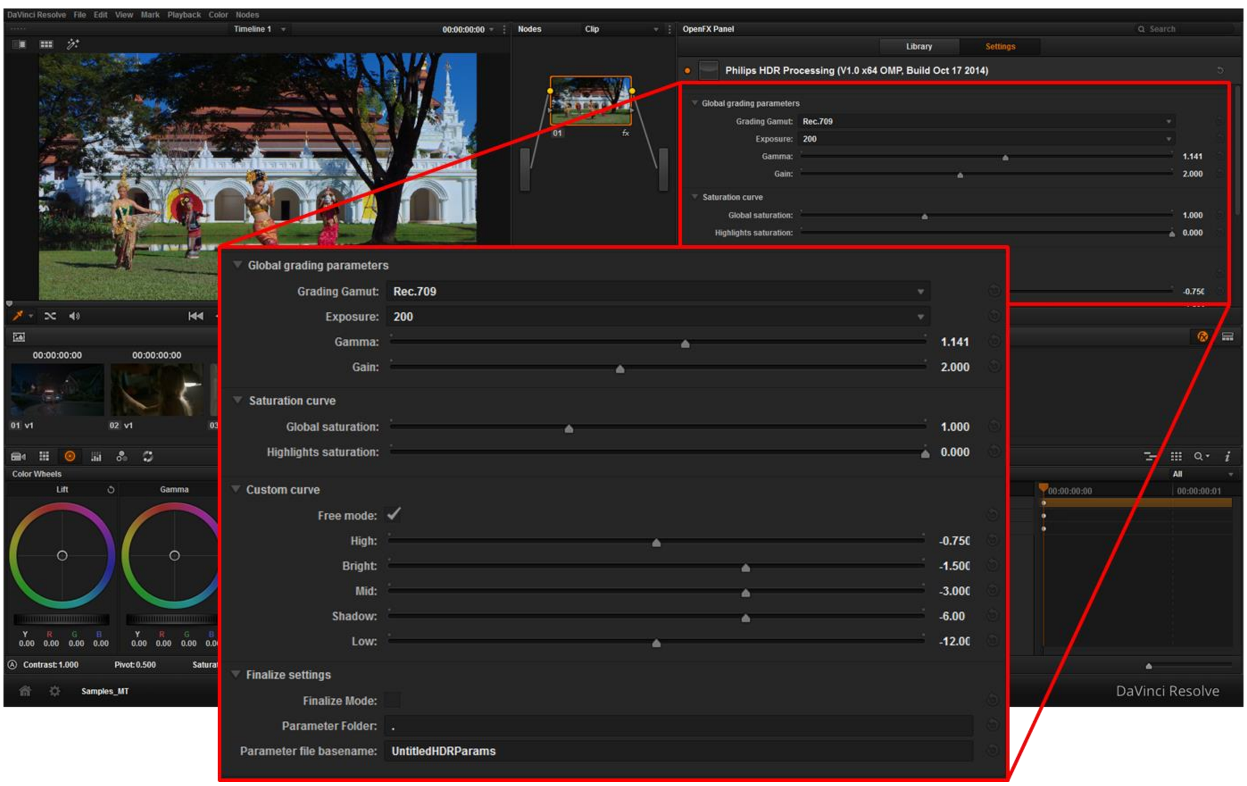
http://mpeg.chiariglione.org/sites/default/files/files/standards/parts/docs/w15350%20v2.docx

The test results are made anonymous in this report; the different systems tested are denoted with an identifier only. The Philips SDR-compatible HDR system is denoted P22. Please note that only one system comes close to Philips HDR, but that system is not SDR compatible. All systems were tested against anchors that correspond to straightforward encoding of the HDR video using the ST 2084 EOTF.

4. Artistic Guidance in a non-real time workflow

In a non-real time workflow, the SDR master video is created Starting from the HDR master video in a trim pass under control of a colorist. The colorist can apply an HDR to SDR tone mapping curve that is defined by three parameters. For further fine tuning, piece-wise linear adjustments to the tone mapping curve can be applied, as well as a luminance dependent saturation control. For those cases where the global adjustments above are not sufficient, local processing windows can be used. Typically one set of parameters is created per scene. Philips has developed an OpenFX plugin that can be used with all major grading suites. See an example below of the GUI of the tool when used in DaVinci Resolve.





5. Artistic Guidance in a real time workflow

In a real time workflow, as used in live broadcast, typically only one (shadow gain control) or maybe two (plus highlight gain control) parameters would be adjusted on the fly. For all other parameters, default values would be used.

If metadata cannot be used at all, the decoder will do a typical SDR to HDR conversion for average broadcast material. The HDR system will still function well, while allowing optimal SDR and HDR video quality during non-real time programming, when the use of metadata is less problematic.

6. Standardization status

BDA. Philips HDR is one of the HDR technologies for UltraHD Blu-ray.

SMPTE. The Philips dynamic metadata set is being standardized as ST 2094-20.

HDMI/CEA. HDMI 2.0a covers HDR EOTF signaling and static metadata. Dynamic metadata is to be covered in HDMI 2.1.

MPEG. The Philips HDR technology (SDR-compatible mode) has been submitted to the Call for Evidence: we obtained HDR results better than the anchor (the anchor corresponds to the BD base layer), better than any other SDR compatible submission, and even on par with the best non-SDR compatible submission.

7. Value chain

STUDIOS	OPERATORS	TV/STB manufacturers	CONSUMER
✓ Bring Artistic Intent (“director mode”) to the home ✓ High level of automation to generate SDR from HDR ✓ SDR/HDR grading plug-in for major grading tools	✓ Single Layer / very low bandwidth metadata: ▪ no simulcast needed ▪ Transmit HDR within existing frequency plans ✓ 10 bits (or 12 bits)	✓ Offer “director mode” viewing option ✓ Optimal adaptation to actual display luminance capability ✓ Can add own “taste” in further tuning ✓ Low-cost decoder	✓ Get content providers’ artistic intent ✓ Keep flexibility to further tune the content in line with the user’s own preference ✓ Low cost

8. Conclusions

Philips HDR offers

- Single layer video decoding
- Excellent HDR quality in 10 bits video
- Very low additional bit rate requirements (keeping standard bandwidth)
- Cost optimized HDR processing (optimized HW costs)

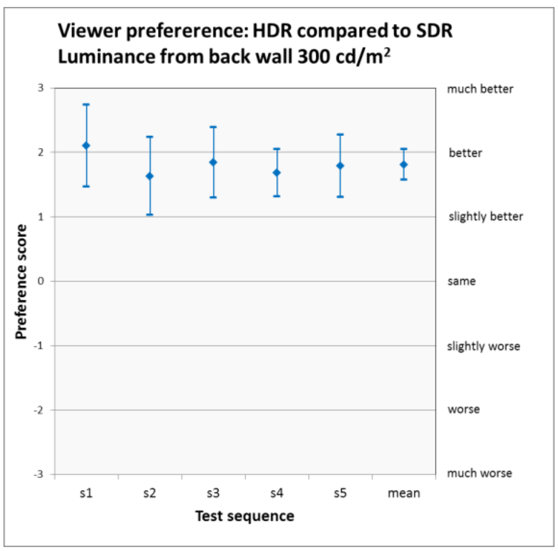
Annexes

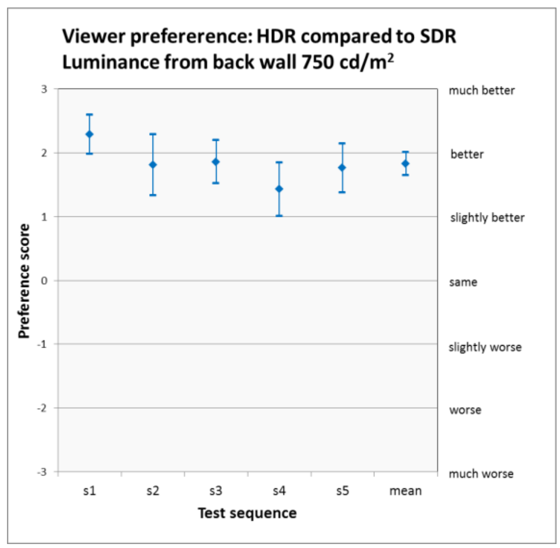
A.1. Peak luminance

An HDR signal format needs to support both high peak luminance as well as low black levels. Philips executed tests to assess the visual preference of high peak luminance High Dynamic Range (HDR) versus Standard Dynamic Range (SDR) images in a daytime television viewing environment. For this test we used 20 non-expert viewers. The test was run with two different levels of environmental luminance:

- 300 cd/m^2, corresponding to the indoor luminance level on a heavily overcast day, and
- 750 cd/m^2, corresponding to the indoor luminance level on a partially clouded day.

The HDR signal was displayed at a peak luminance of 5 000 cd/m^2, the SDR at a peak luminance of 400 cd/m^2. The figures below show the mean and 95% confidence interval for the 5 test sequences, plus the overall mean and confidence interval.





Both tests show a clear preference for HDR video compared to SDR video.

In the same test, we asked the viewers about their preference for the level of the peak luminance. The majority of viewers preferred a maximum peak luminance of around or slightly below 5 000 cd/m^2.

A.2. A future EOTF

The Electro-Optical transfer Function (EOTF) of current video formats is the well-known gamma curve. It is based upon the physical characteristics of CRTs, but turned out to work well for SDR signals. For the transmission of HDR signals, such a gamma curve is however not suitable: the bit depth needed to provide sufficient accuracy for the signals over the whole luminance range would be excessively high.

In order to most efficiently use the available bit depth or, in other words, to minimize the required bit depth for an HDR television system, an appropriately designed EOTF should be applied. Philips has extensively studied display EOTFs with the goal to minimize the visibility of quantization or compression errors and we have found (and demonstrated) that such quantization and compression errors can be made invisible to a human observer for an appropriately designed HDR television systems under practical viewing conditions. Philips proposes the following reference EOTF for high dynamic range displays:

$$L = L_m \left(\frac{\rho^{v}-1}{\rho-1}\right)^{\gamma} \quad \text{eq.(1)}$$

where L is the luminance in cd/m^2, v is the electrical value normalized to the 0..1 range, and L_m is the peak luminance value of the reference display in cd/m^2. The proposed values of the constants are, for a 5 000 cd/m^2 peak luminance system:

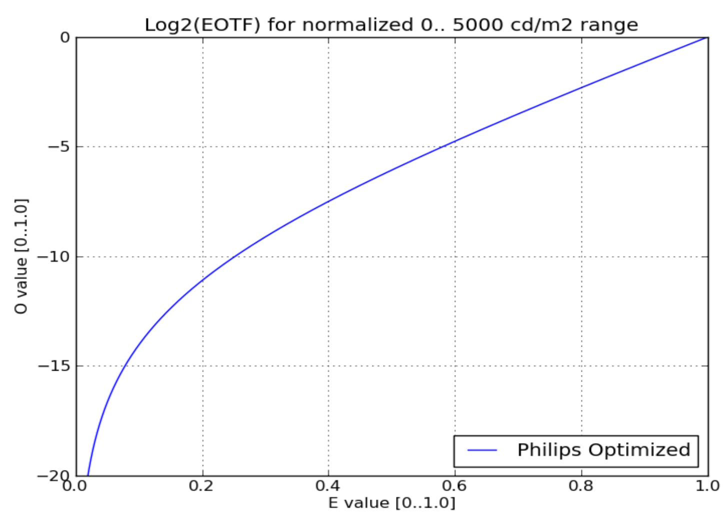
$$\rho = 25, \gamma = 2.4, L_m = 5\,000 \text{ cd/m}^2 \quad \text{eq.(2)}$$

For a 10 000 cd/m^2 peak luminance system, the proposed values of the constants are:

$$\rho = 33, \gamma = 2.4, L_m = 10\,000 \text{ cd/m}^2 \quad \text{eq.(3)}$$

This EOTF behaves like a gamma curve at low luminance levels, gradually turning into an exponential curve at high luminance levels.

We have optimized the EOTF parameters for minimizing the visibility of quantization errors, using a variety of images ranging from very dark (night scene) to very bright (sunny day scene), displayed on a calibrated high-dynamic-range display with a peak luminance of 5 000 cd/m^2. The figure below shows the EOTF in a graphical form. The horizontal axis shows the electrical value *E*, corresponding to *v* in equation (1), the vertical axis shows the optical value *O*, corresponding to a normalized *L* in equation (1).



Note that this optimized display EOTF is close to the predictions of the model obtained by retired Philips employee P.G.J. Barten [2].

[2] P. G.J. Barten, "Contrast Sensitivity Of The Human Eye And Its Effects On Image Quality", SPIE Press, 2000.

A.3. The corresponding OETF

Current television systems have an end-to-end (optical to optical) non-linear transfer characteristic. This transfer characteristic provides the correct rendering intent for the typical dim surround television viewing environment [3].

Philips has investigated the end-to-end television system transfer characteristic for future high dynamic range television systems with high peak luminance displays (specifically a display with a peak luminance of 5 000 cd/m^2 was employed in Philips' experiments) and has found that the current end-to-end transfer characteristic is also applicable to these future systems. The explanation for this observation is that the transfer characteristic is determined by the television viewing environment, which for high dynamic range television will be the same as it is for current television.

The end-to-end transfer characteristic for current television systems is determined by the concatenation of the recommended OETF (Rec. ITU-R BT.709 and Rec. ITU-R BT.2020) and EOTF (Rec. ITU-R BT.1886). Philips also proposes an Opto-Electrical Transfer Function (OETF) that is appropriate for the proposed EOTF and fully preserves the current rendering intent. The proposed OETF is given by:

$$E' = \begin{cases} \dfrac{\log(35.445\, E \cdot (\rho\text{-}1)+1)}{\log(\rho)}, & 0 \le E < \beta \\[2ex] \dfrac{\log((\alpha E^{0.508} - (\alpha - 1)) \cdot (\rho\text{-}1)+1)}{\log(\rho)}, & \beta \le E \le 1 \end{cases}$$

where *E* is the voltage normalized by the reference white level and proportional to the implicit light intensity that would be detected with a reference camera color channel *R, G, B; E'* is the resulting non-linear signal, and:

$$\rho = 25,\ \alpha = 1.00622 \text{ and } \beta = 0.0001812 \qquad \text{eq.(3)}$$

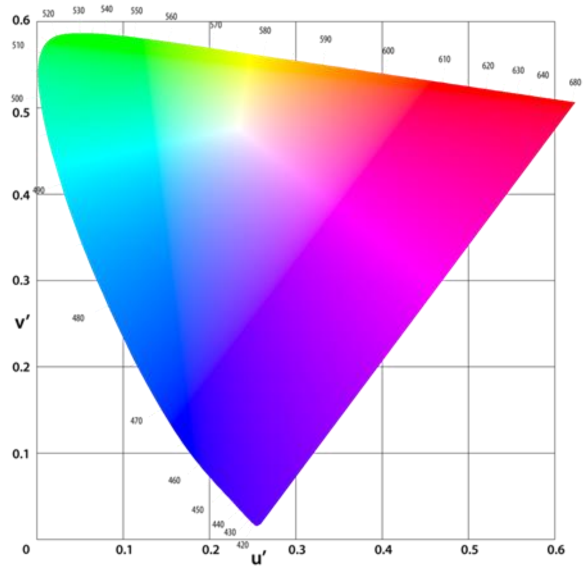
[3] See e.g. sections 11.9, 19.13, and 23.14, of "The Reproduction of Colour" by R.W.G. Hunt (Sixth ed., Wiley, 2006).

A.4. A future color representation

Philips HDR works with standard YCbCr color representation. This makes it possible to re-use existing professional equipment, IP blocks, and maybe most importantly, knowledge. And for an SDR-compatible system, YCbCr is of course the only option.

Philips is in its research however also looking into future technologies, even though these are not expected to reach the market in the short term future. Philips developed amongst others the Y”u”v” color representation to optimize the quality and the bandwidth of an HDR encoded signal. Y”u”v” has major advantages:

- It is a perceptually uniform color space
- Has true constant luminance
- Wide color gamut can be supported at low cost in terms of required bit depth
- 10 bits Y”u”v” is equivalent to 12 bits XYZ or YDzDx
- The u”, v” components can be linearly quantized
- Color subsampling behavior is much better than with YCbCr



See further the publications on this topic:

[1] Charles Poynton, Jeroen Stessen, Rutger Nijland “Deploying Wide Colour Gamut and High Dynamic Range in HD and UHD”, IBC 2014

[2] Charles Poynton, Jeroen Stessen, Rutger Nijland “Deploying Wide Colour Gamut and High Dynamic Range in HD and UHD”, SMPTE Motion Imaging journal April 2015