International Telecommunication Union

ITU-R

Radiocommunication Sector of ITU

Report ITU-R BT.2408-4

(03/2021)

Guidance for operational practices in HDR television production

BT Series

Broadcasting service (television)

International Telecommunication Union

Foreword

The role of the Radiocommunication Sector is to ensure the rational, equitable, efficient and economical use of the radio-frequency spectrum by all radiocommunication services, including satellite services, and carry out studies without limit of frequency range on the basis of which Recommendations are adopted.

The regulatory and policy functions of the Radiocommunication Sector are performed by World and Regional Radiocommunication Conferences and Radiocommunication Assemblies supported by Study Groups.

Policy on Intellectual Property Right (IPR)

ITU-R policy on IPR is described in the Common Patent Policy for ITU-T/ITU-R/ISO/IEC referenced in Resolution ITU-R 1. Forms to be used for the submission of patent statements and licensing declarations by patent holders are available from http://www.itu.int/ITU-R/go/patents/en where the Guidelines for Implementation of the Common Patent Policy for ITU-T/ITU-R/ISO/IEC and the ITU-R patent information database can also be found.

Series of ITU-R Reports

(Also available online at http://www.itu.int/publ/R-REP/en)

Series	Title
BO	Satellite delivery
BR	Recording for production, archival and play-out; film for television
BS	Broadcasting service (sound)
BT	**Broadcasting service (television)**
F	Fixed service
M	Mobile, radiodetermination, amateur and related satellite services
P	Radiowave propagation
RA	Radio astronomy
RS	Remote sensing systems
S	Fixed-satellite service
SA	Space applications and meteorology
SF	Frequency sharing and coordination between fixed-satellite and fixed service systems
SM	Spectrum management

***Note**: This ITU-R Report was approved in English by the Study Group under the procedure detailed in Resolution ITU-R 1.*

Electronic Publication
Geneva, 2021

REPORT ITU-R BT.2408-4

Guidance for operational practices in HDR television production

(2017-2018-04/2019-07/2019-2021)

Summary

This Guidance for operational practices is intended to help ensure optimum and consistent use of high dynamic range in television production using the Perceptual Quantization (PQ) and Hybrid Log-Gamma (HLG) methods specified in Recommendation ITU-R BT.2100. Additional background information on HDR is available in Report ITU-R BT.2390, while Report ITU-R BT.2446 provides guidance towards the design of methods of conversion between HDR and SDR content.

TABLE OF CONTENTS

Page

1 Introduction

Recommendation ITU-R BT.2100 (BT.2100) specifies HDR-TV image parameters for use in production and international programme exchange using the Perceptual Quantization (PQ) and Hybrid Log-Gamma (HLG) methods. Since its first publication in 2016, television programme production in high dynamic range (HDR) continues to grow and is attracting increasing interest from content creators and broadcasters wishing to benefit from the improved viewing experience that HDR offers. At the same time, standard dynamic range (SDR) and high dynamic range will need to coexist for many years to come. These operational practices propose guidance to programme makers and broadcasters based on knowledge and practical experience gained so far. A glossary of terms is included at the end of this Report.

Production in PQ is similar to standard dynamic range production. During capture, the scene may be exposed to produce the desired appearance on a reference monitor, ideally operating in the reference environment. Exposure setting may be assisted for example by setting a grey or diffuse white card to the desired signal level. It is possible for the PQ system to capture and encode information that is beyond the capabilities of a specific monitor, if that monitor cannot reach both the ideal peak luminance of 10 000 cd/m^2 and the full extent of the BT.2100 wide colour gamut. If the PQ signal is not actively constrained to the capability of the reference monitor in use, more information may be revealed on a subsequent display with higher peak luminance or colour gamut.

HLG has been designed to enable a straightforward migration towards HDR television production, with few changes to SDR production working practices. The compatible nature of the HLG signal allows standard dynamic range monitors to be used in non-critical monitoring areas. HDR monitors are necessary only for critical monitoring, such as when colour grading, camera shading (or racking) and monitoring programme and preview outputs in a production gallery.

Just as line-up levels are useful for audio production, nominal signal levels for standard test charts are also useful for HDR television production. Nominal signal levels are given in order to facilitate camera line-up to help ensure consistency both within and between programmes, together with advice on monitoring and displaying HDR content.

Initial findings are presented on viewer tolerances to variations in image brightness, aimed in particular at avoiding viewer discomfort at junctions between programmes and other items of content, as well as when switching between programme channels.

Techniques are described for including SDR content in HDR productions, as are the principles of transcoding between PQ and HLG. Experience gained from trials with live production is documented, offering a practical guide for transitioning from SDR to HDR.

Annexes provide further technical details and background information. Annexes 1, 2, 3 and 4 present the results of studies analysing skin tones and other existing content which have been used to help inform guidance on video levels in HDR production.

Annex 5 compares various approaches that can be used to map PQ signals to displays with a lower dynamic range than contained in the signal; such processes may also be required during conversion from PQ to HLG.

Annex 6 compares the native displayed "look" of each SDR and HDR production format.

Annex 7 gives technical details on how to calculate the normalized primary matrix (NPM) needed for conversion to and from the CIE XYZ colour space and the BT.2100 colour space.

2 Reference levels and signal format

During set-up, camera controls such as gain and shutter and others may be pre-adjusted to make best use of camera sensor capabilities, i.e. a balance between signal to noise ratio (SNR) and achieved sensor peak capability, and to establish a creative intent. During capture, the exposure may then be adjusted taking consideration of the reference levels listed below as well as the creative intent.

2.1 HDR Reference White

The reference level, HDR Reference White, is defined in this Report as the nominal signal level obtained from an HDR camera and a 100% reflectance white card resulting in a nominal luminance of 203 cd/m^2 on a PQ display or on an HLG display that has a nominal peak luminance capability of 1 000 cd/m^2. That is the signal level that would result from a 100% Lambertian reflector placed at the centre of interest within a scene under controlled lighting, commonly referred to as diffuse white[1]. There may be brighter whites captured by the camera that are not at the centre of interest, and may therefore be brighter than the HDR Reference White.

Graphics White is defined within the scope of this Report as the equivalent in the graphics domain of a 100% reflectance white card: the signal level of a flat, white element without any specular highlights within a graphic element. It therefore has the same signal level as HDR Reference White, and graphics should be inserted based on this level.

The nominal signal level corresponding to HDR Reference White, diffuse white and Graphics White is shown in Table 1.

Signal levels for common test charts and reflectance cards with different reflectances are calculated using scene-light (the light falling on a camera sensor), from HDR Reference White. Details are given in § 2.2.

2.2 Signal levels for line-up in production

Signal levels in these operational practices are specified in terms of %PQ and %HLG. These percentages represent signal values that lie between the minimum and maximum non-linear values normalized to the range 0 to 1.

The values in Table 1 are presented as nominal recommendations for test charts and graphics for PQ production and for HLG production on a 1 000 cd/m^2 (nominal peak luminance) display, under controlled studio lighting[2]. For PQ, the nominal luminance values are consistent on PQ reference displays. For HLG, the nominal luminance values will differ from those in Table 1 when the display's peak luminance is lower or higher than 1 000 cd/m^2. The nominal signal levels in Table 1 do not change. There is a practical benefit to the use of common levels for both PQ and HLG and Table 1 reflects guidance to use common levels. However, as PQ and HLG have different capabilities, and as HLG levels are influenced by a desire to maintain a degree of compatibility with SDR displays and PQ levels are not, as experience is developed in the use of both PQ and HLG this guidance to use common levels may need to be adjusted. Annex 1 describes a study of early HDR movies graded on a 4 000 cd/m^2 PQ monitor. According to that study, luminance levels for indoor scenes were found to be typically about two thirds of the values indicated in Table 1, however those for outdoor scenes were found to be brighter. As producers of PQ content gain more experience, it is possible that levels in PQ indoor content may increase.

---

[1] Diffuse white is the white provided by a card that approximates to a perfect reflecting diffuser by being spectrally grey, not just colorimetrically grey, by minimizing specular highlights and minimizing spectral power absorptance. A "perfect reflecting diffuser" is defined as an "ideal isotropic, nonfluorescent diffuser with a spectral radiance factor equal to unity at each wavelength of interest".

[2] The test chart should be illuminated by forward lights and the camera should shoot the chart from a non-specular direction.

It is important to know the reflectance of greyscale charts and white cards, to ensure that cameras are aligned to deliver the appropriate signal level and consistency in production.

An 18% grey card is commonly used for camera set-up in non-live workflows as it is the closest standard reflectance card to skin tones. It may also be useful when trying to match SDR and HDR cameras as the 18% grey should not be affected by any SDR camera “knee”.

A 75%HLG or 58%PQ marker on a waveform monitor, representing the reference level, will help the camera shader ensure that objects placed at the centre of interest within a scene are placed within the appropriate signal range, and that sufficient headroom is reserved for specular highlights.

TABLE 1

Nominal levels for PQ and HLG production

Reflectance object or reference (luminance factor, %)[3]	**Nominal luminance, cd/m^2 (for a PQ reference display, or a 1 000 cd/m^2 HLG display)**	**Nominal signal level**	
		%PQ	**%HLG**
Grey Card (18%)	26	38	38
Greyscale Chart Max (83%)	162	56	71
Greyscale Chart Max (90%)	179	57	73
Reference Level: HDR Reference White (100%) also diffuse white and Graphics White	203	58	75

NOTE – The signal level of “HDR Reference White” is not directly related to the signal level of SDR “peak white”.

In an experiment described in full in Annex 2, the levels of white objects in different types of HDR content were assessed, including an early live shoot of a baseball game, as well as a collection of HDR still photographs. In both cases, the mean white level is consistent with the HDR Reference White level as given in Table 1. However, for both types of content the spread around this mean value is significant, indicating that in practice the measured white levels can be expected to vary significantly around this target value.

When test charts are either not available or impractical, other objects such as skin tones or grass are often used to set signal levels. Approximate signal levels are given in Table 2.

The Fitzpatrick Skin Tone Scale [1] is used to classify skin types, which will vary by region. It was originally developed as a way to estimate the response of different types of skin to ultraviolet light. It may be used to provide a convenient classification method for the range of skin tones seen in television production.

Annex 3 describes how both experimental data, and a theoretical model of an ideal HDR television camera, have been used to determine the expected signal ranges for the Fitzpatrick skin types illustrated in Table 2. These ranges assume that content has been produced using the HDR Reference White signal levels specified in Table 1.

Annexes 1 and 4 report on skin tones in broadcast SDR content produced in studios in different regions. The skin tones in SDR content were found to be much different by regions. This may be

---

[3] “Luminance factor” is the ratio of the luminance of the surface element in the given direction to the luminance of a perfect reflecting or transmitting diffuser identically illuminated.

mainly due to a difference in long-standing production practice for SDR rather than a difference in skin reflectance. Annex 4 also reports on a study on skin tones in HLG HDR content with camera shading compliant to the reference level of 75%HLG in comparison with SDR content, both produced independently for the same programme. The facial skin tones in the HLG content correspond to the Type 3-4 (medium skin tone) in Table 2.

Variations in these signal levels can be expected. The value for grass, for example, will depend on the type of grass planted for a given sport. Creatives making programme content may choose to encode content at differing levels, i.e. a dark indoor drama may put a grey card (and thus skin tones) at a lower level than shown in Table 1. Also, some productions may employ higher/brighter levels for outdoor scenes or for dramatic effect.

As with the values for HDR Reference White, the nominal luminance values for PQ are the same on a PQ reference display, whereas the nominal luminance values vary for HLG depending on the display's peak luminance. Table 2 gives values for an HLG display with 1 000 cd/m^2 nominal peak luminance. The nominal signal levels do not change.

TABLE 2

Preliminary levels for common objects in PQ and HLG production

Reflectance object	**Nominal Luminance, cd/m^2 (for a PQ reference display, or a 1 000 cd/m^2 HLG display)**	**Signal level**	
		%PQ	**%HLG**
Skin Tones (Fitzpatrick Scale)			
Type 1-2 Light skin tone[4]	65-110	45-55	55-65
Type 3-4 Medium skin tone	40-85	40-50	45-60
Type 5-6 Dark skin tone[4]	10-40	30-40	25-45
Grass	30-65	40-45	40-55

2.3 Bit depth

High quality HDR programmes can be produced using conventional 10-bit infrastructure and 10-bit production codecs, with similar bitrates used for standard dynamic range production.

The use of 12-bit production systems will, however, give greater headroom for downstream signal processing for both PQ and HLG.

2.4 Signal range

Recommendation ITU-R BT.2100 specifies two different signal representations, "narrow" and "full". Narrow range signal representations are traditionally used for television programme production. They provide headroom above the code value of the nominal peak (where the signal $E' > 1.0$) and below zero light (where the signal $E' < 0.0$) to accommodate signal overshoots and undershoots. Signals above the nominal peak are often termed "super-whites" and those below zero light termed "sub-blacks", although they need not be achromatic signals. Full range signal representations are more common in cinematic workflows. The movie industry has traditionally followed the computer

[4] Experimental data for Type 1, Type 5 and Type 6 skin types is limited. So there is less certainty on the signal ranges for these skin types.

graphics industry and placed zero light at digital code value "0", and the code value of the nominal peak at the maximum code value for the given bit-depth. Full range signals do not, therefore, provide any headroom for signal overshoots or undershoots.

Signal overshoots and undershoots are produced by video processing techniques such as image re-sizing, filtering and compression, that are common in television production workflows. Overshoots and undershoots may also be present in the SDR signal after HDR to SDR down-mapping, particularly if the SDR super-white signal range is used to accommodate some of the highlights from the HDR source (see § 7.3.4). In order to maintain image fidelity, it is important that such overshoots and undershoots are not clipped. Any signal clipping introduces harmonic distortion, which makes the task of subsequent video compression or filtering even harder. Full range signals, which cannot accommodate signal overshoots and signal undershoots, are thus generally avoided in broadcasting systems. Furthermore, the black level of a display to represent an HLG signal should be adjusted using the Recommendation ITU-R BT.814 PLUGE signal, which is only possible if sub-blacks are present in the signal. Where HLG is used for programme production and exchange the full range signal representation should not be used.

The full range representation for PQ signals may, however, be useful as it provides an incremental advantage against visibility of banding/contouring and for processing. Furthermore, because the range of PQ is so large, it is rare for content to contain pixel values near the extremes of the range. Signal overshoots are therefore less likely to exist. It should be noted that full range signals may not be supported by broadcast distribution systems. For broadcast contribution, programme exchange or distribution, the full range signal representation of PQ should be used only when all parties agree. In the absence of such agreement, any PQ full range signals should be mapped to narrow range.

2.5 Colour representation

Recommendation ITU-R BT.2100 describes two luminance and colour difference signal representations, suitable for colour sub-sampling and/or source coding: the non-constant luminance $Y'C'_BC'_R$ signal format and the constant intensity IC_TC_P format.

As the IC_TC_P signal format is not compatible with conventional SDR monitors, and any benefits of the IC_TC_P colour representation are anticipated to be less for HLG than for PQ, so the non-constant luminance $Y'C'_BC'_R$ signal format is preferred for HLG.

For PQ, the IC_TC_P format has been shown to be advantageous in a number of respects (see Report ITU-R BT.2390), but compatibility with signal handling equipment must be considered before choosing to employ this format.

3 Monitoring

Ideally, critical monitoring, such as the production switcher's "programme" and "preview" outputs, should take place using a display that supports the full colour gamut and dynamic range of the signals. Monitors that support the BT.2100 colour space should include means to manage colours outside of their native display gamut.

3.1 Display of PQ signals

The content represented by PQ signals may be limited to the expected capabilities of the displays on which they are intended to be viewed, or they may be unlimited and represent the full level of highlights captured by the camera. In practice, monitors may not reach the full extent of the BT.2100 gamut or the 10 000 cd/m^2 limit of the PQ signal, resulting in the possibility that some encoded colours may not be displayable on some monitors.

Monitors that support PQ may or may not include tone-mapping to bring very high brightness signals down to the capability of that monitor. Some monitors may clip at their peak output capability (e.g. 2 000 cd/m^2). Some monitors may contain tone mapping that provides a soft-clip.

For production use, monitors should generally perform a hard clip to the display capabilities, and should provide a means to identify pixels that are outside the display’s capability (either in brightness or colour). If a soft-clip is desired, a Look-up-table (LUT) such as that described in § 3.1.1 can be applied to the signal to provide any desired tone mapping. Care should be taken for any content that is allowed to go outside the reference monitor colour gamut or dynamic range as that would not have been accurately presented to the operator and cannot be trusted as part of the approved or intended appearance. Reference monitors could provide a selectable overall brightness-attenuation in order to temporarily bring high brightness signals down to be within the display capability in order to provide a check on any content encoded brighter than the capability of the reference display.

If the BT.2100 PQ signal is presented to a monitor that expects a Recommendation ITU-R BT.709 (BT.709) input, the image will appear dim and washed out; colours will be desaturated and there will be some hue shifts. An external 3D LUT can provide the down-mapping function necessary to bring both colour and brightness into the BT.709 colour volume, thus allowing satisfactory display on the BT.709 monitor. Some monitors may provide this function by means of an internally provided 3D LUT. While this allows viewing on the BT.709 monitor, the resulting images should not be used to make critical judgements of the HDR production.

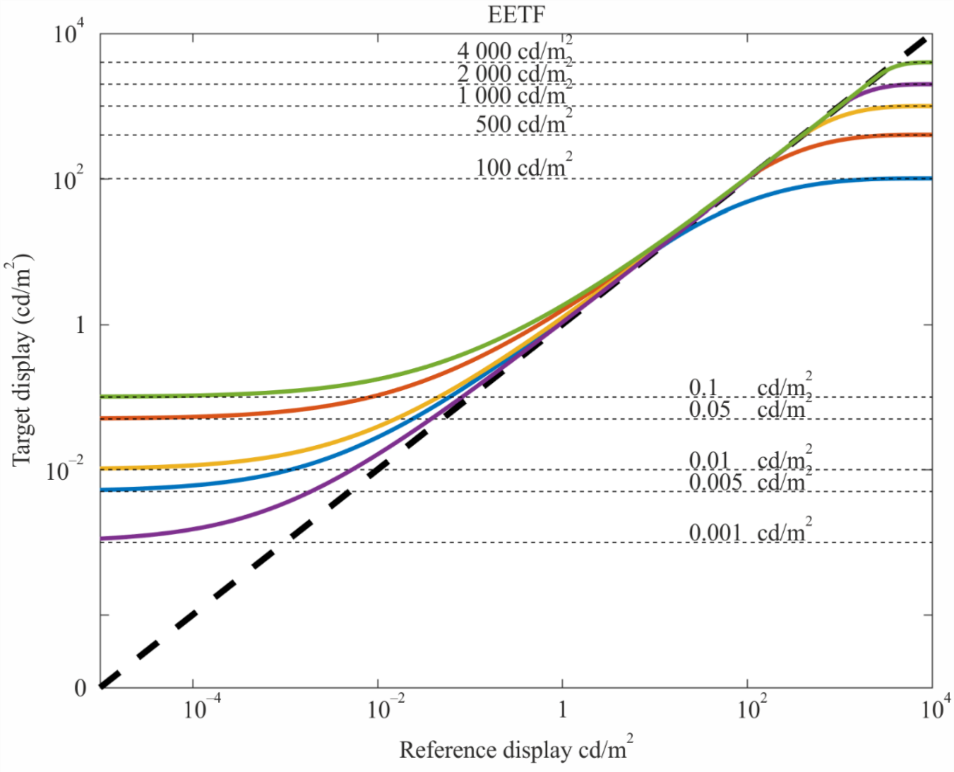
If PQ signals must be monitored in an environment brighter than the reference environment (specified in Recommendation ITU-R BT.2100 as having a 5 cd/m^2 surround), manufacturers may provide modified brightness and display characteristics intended to compensate for the different viewing environment.

3.1.1 Mapping to displays with limited luminance range

To view the entire range of HDR content on displays with a lower dynamic range, display mapping should be performed. This can take the form of an EETF (electrical-electrical transfer function) in the display. This function provides a toe and knee to gracefully roll off the highlights and shadows providing a balance between preserving the artistic intent and maintaining details. Figure 1 is an example EETF mapping from the full 0 – 10 000 cd/m^2 dynamic range to various target displays.

FIGURE 1

Example EETF from 0 – 10 000 cd/m² to various target displays

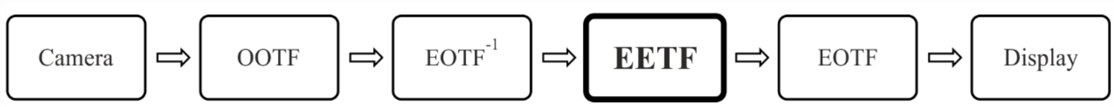


Report BT.2408-01

Annex 5 gives the specific mathematical steps to implement this tone mapping function for displays of various black and white luminance levels. Figure 2 shows the block diagram of where the EETF should be applied.

FIGURE 2

Block diagram of signal chain showing location of EETF application



Report BT.2408-02

3.2 Display of HLG signals

Table 5 of Recommendation ITU-R BT.2100 specifies the HLG EOTF (electro-optical transfer function) for reference displays. Note 5f specifies how the display's gamma is adjusted to compensate for changes in the response of the human visual system as the eye adapts, when using HLG displays of different peak luminance. The gamma adjustment allows consistent signals to be produced from a range of displays with different peak luminance. Details can be found in § 6.2 of Report ITU-R BT.2390.

The luminance on a production monitor corresponding to nominal peak, 100%, signal level, should be adjusted to a comfortable level for the viewing environment. Nominal peak signal level does not have to be set to the peak luminance of the monitor, which may be too bright for comfortable viewing. The nominal peak luminance of 1 000 cd/m², identified in Recommendation ITU-R BT.2100, has been found to work well in typical production environments.

Note 5g of Recommendation ITU-R BT.2100 recognises that the display's gamma should further be adjusted to compensate for the adaptation state of the eye in non-reference production environments. A formula specifying the gamma adjustment is also given in § 6.2 of Report ITU-R BT.2390.

Contrast, brightness and display system gamma (α, β and γ in Table 5 of Recommendation ITU-R BT.2100) are adjusted according to the viewing environment and nominal peak luminance of the display, as appropriate.

Firstly, the monitor gamma is adjusted, according to the formula in Note 5f of Recommendation ITU-R BT.2100, to the appropriate value for the target nominal peak luminance of the display. The target nominal peak luminance may depend on the viewing environment.

Table 3 shows the gamma values for a range of typical production monitors in the reference viewing environment (5 cd/m^2 surround).

TABLE 3

HLG display gamma

Nominal peak luminance (cd/m^2)	Display Gamma
400	1.03
600	1.11
800	1.16
1 000	1.20
1 500	1.27
2 000	1.33

The display's nominal peak luminance is then adjusted using the user gain control (legacy "contrast" control) and a photometer, with an HDR reference white (75%HLG) window test patch (typically 1% screen area). Table 4 shows the luminance levels for a range of typical production monitors.

TABLE 4

Test patch luminance levels for different nominal peak displays

Nominal peak luminance (cd/m^2)	HDR reference white (cd/m^2)
400	101
600	138
800	172
1 000	203
1 500	276
2 000	343

In non-reference viewing environments, a further adjustment should be made to the display's system gamma to compensate for the adaptation state of the eye. Table 5 illustrates the recommended gamma adjustments for a range of common production environments, assuming a surround reflectance of approximately 60%, typical of light coloured walls. However, for the greatest signal consistency, the reference conditions specified in Recommendation ITU-R BT.2100 should be used.

TABLE 5

Typical production environments with different surround conditions

Typical environment	Typical Illumination[5] (Lux)	Typical luminance[6] (cd/m^2)	Typical gamma adjustment
Office based production sunny day	130	25	−0.05
Office based production cloudy day	75	15	−0.04
Edit Suite	50	10	−0.02
Grading Suite	25	5	0.00
Production gallery/ Dark grading suite	3	0.5	+0.08

As a guide, a gamma adjustment of 0.03 is just visible to the expert viewer when viewed side-by-side. Thus, no additional gamma adjustment is necessary across the majority of critical television production environments.

However, a gamma adjustment is suggested for bright environments such as those sometimes used for news production, or where a colourist prefers to work in a very dark environment.

Lastly, the display black level is adjusted using the black level lift control (legacy "brightness" control) and the Recommendation ITU-R BT.814 PLUGE signal, such that the negative stripes on the test pattern disappear, whilst the positive stripes remain visible.

3.2.1 Display of HLG signals on SDR screens

For best results when displaying HLG signals on SDR screens, the SDR monitor should support the Recommendation ITU-R BT.2020 (BT.2020) colour gamut. However, for simple confirmation of the presence or absence of a signal, BT.709 colour monitoring may be sufficient. However, BT.709 colour monitors will show a de-saturated image with visible hue shifts.

Non-critical production monitors, such as multi-view production monitors, may be SDR BT.709 displays. A three-dimensional look-up table (3D-LUT) may be included in the monitoring chain to down-convert from BT.2100 HDR signals to BT.709 SDR, minimising colour distortions on such displays. Suitable look-up tables are often included within the display monitors themselves.

4 Image brightness

Work has commenced on developing automatic objective measures for brightness, akin to those in common use for audio loudness today. Experimental results [2] show that a simple mean of displayed pixel luminances provides a good correlation with subjective brightness at 3.2 picture heights from the screen. The effectiveness of this simple objective metric suggests that real-time brightness monitoring in production is a realistic goal. This would give guidance to content producers, enabling comfortable viewing in the home, whilst allowing a range for artistic freedom. The metric could be used further to characterise long-term and short-term average brightness.

---

5 Measured perpendicular to the screen.

6 Assuming ~ 60% reflectance surround.

4.1 Comfortable brightness of static images

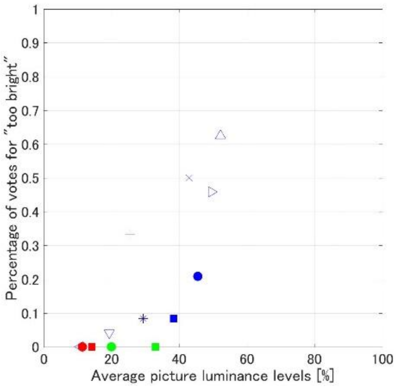
A study was performed by NHK to learn what range of luminances are judged comfortable by viewers. A number of SDR images that, on a 100 cd/m^2 reference monitor, varied in average luminance over a range of 10-50 cd/m^2, were used. The study was conducted using a relative display system that employed a 3 500 cd/m^2 display that was adjusted to simulate a range of display luminance levels, thus the results are relevant to the HLG system that also employs displays with relative luminance. Peak luminances of 500, 1 000, 2 000, and 2 500 cd/m^2 were simulated. Viewers were asked to judge whether images were “appropriate”, “too bright”, or “too dark”.

Figure 3 shows the results in the reference viewing environment (dim surround). For each simulated display peak luminance, images with average luminance less than 25% of the peak luminance being simulated were not judged as “too bright”. Images with average luminance greater than 25% of peak luminance began to be judged as “too bright” by many viewers. The judgements were essentially independent of the peak luminance being simulated on the display; this indicates that viewers’ eyes were adapting to the different display luminances. The implication of these results is that HLG images with average luminance of less than 250 cd/m^2 on a 1 000 cd/m^2 HLG monitor, would not be judged as too bright on an HLG monitor of any luminance up to at least 2 500 cd/m^2.

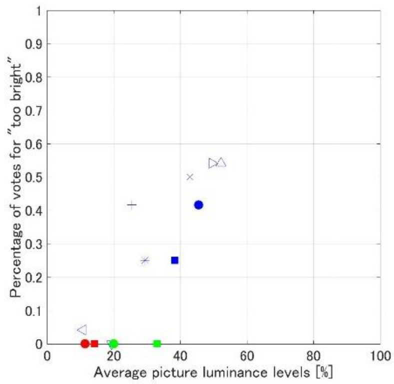
FIGURE 3

Percentage of votes for “too bright” in the reference environment (dim surround)

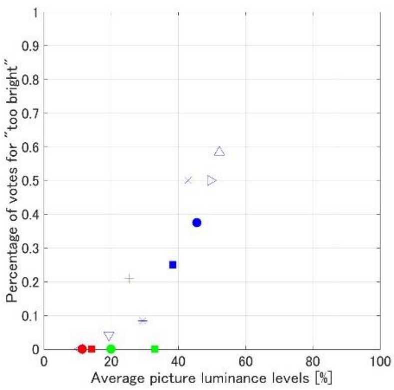
Peak luminance 500 cd/m^2



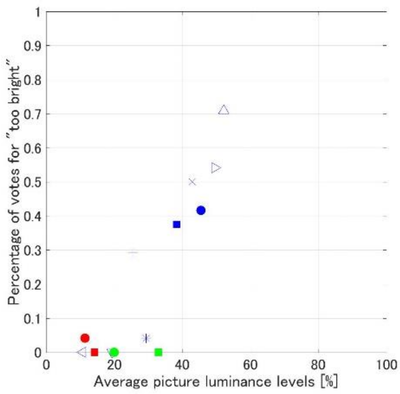
Peak luminance 1 000 cd/m^2



Peak luminance 2 000 cd/m^2



Peak luminance 2 500 cd/m^2



This is consistent with informal comments from subjects in separate tests performed by the BBC, which were targeted at measuring tolerance to brightness jumps (see § 4.2). Having seen HDR video sequences on HLG displays with peak luminance levels of 1 000 cd/m^2 and 4 000 cd/m^2, 25% of subjects commented informally that the brightest scenes were uncomfortably bright regardless of any jumps. These scenes had average luminance levels of 268 and 363 cd/m^2 on a 1 000 cd/m^2 display. Similar comments were not made about the test scenes that had average luminances of 144 and 128 cd/m^2 on a 1 000 cd/m^2 display.

Even when the static levels would be acceptable, sudden changes in brightness can be uncomfortable even when the static levels would be acceptable, so different requirements are needed to ensure viewer comfort when brightness jumps can occur.

4.2 Tolerance to programme brightness shifts

Unexpected changes in image brightness might occur between programmes, for example with interstitials. It is important to ensure that the brightness variations within HDR programmes are constrained to avoid viewer discomfort.

Subjective tests reported by the BBC investigated viewer tolerance to sudden changes in overall brightness for HDR television, using the mean pixel display luminance as a measure of brightness as described in [2]. This measure has been shown to correlate well with subjective ratings of the overall brightness, but there may occasionally be a scene with a non homogeneous spatial luminance distribution where the measure does not fully correspond to subjective brightness. For the tests, the luminance behind the screen was 5 cd/m^2, and the peak screen luminance was 1 000 cd/m^2 [3]. Subjects were asked to rate the change in overall brightness between two still HDR images.

Figure 4 shows the overall results, with transitions from the first mean luminance A to the second mean luminance B categorised according to whether they are “not annoying”, “slightly annoying”, or “annoying”. Two regions are marked in the figure with thick blue lines. The inner region, with mean display luminance levels of 5 to 80 cd/m^2, contains only one possible “slightly annoying” jump, and so could be considered a suitable range for operation that will not cause viewer discomfort. The outer region, with mean display luminance levels up to 160 cd/m^2, includes several slightly annoying jumps, and so could be considered an extended range for creative effect. Further experiments reported by the BBC show that this outer region can be extended down to 2.5 cd/m^2, and production trials with a prototype meter suggest that this extended range is appropriate.

Specific delivery requirements for luminance ranges are left to individual service providers, depending on their requirements. An example of requirement could be that the suggested ranges can be freely exceeded over a short timescale, but the mean luminance over the length of a programme is kept within an operating range of 5 to 80 cd/m^2. It should be noted that this range still allows for significant differences in brightness between programmes, so, for example, a “moody” or “bright” look can be achieved overall.

The results presented previously in Fig. 3 provide evidence that the eye adapts to a particular luminance level. Hence the scene-light levels corresponding to specified brightness shift tolerances are likely to be broadly applicable for HLG displays over a range of different peak luminances. This is supported by experiments reported by the BBC, which suggest that the ranges are applicable for HLG displays up to a peak luminance of 4 000 cd/m^2.

It should be noted that shadow detail may be lost after a transition from a bright scene to a very dark scene, even if the transition is not uncomfortable, because it takes time for the eyes to adapt. Also, a comfortable overall brightness does not ensure that the content makes good use of the available dynamic range. Further guidance may be useful to characterise best use of the dynamic range for common scene types.

FIGURE 4

Transitions from mean luminance A (cd/m^2) to mean luminance B (cd/m^2) categorised by level of annoyance

A \ B	5	10	20	40	80	160	320
5	g	g	g	g	a	a	r
10	g	g	g	g	g	a	r
20	g	g	g	g	g	a	a
40	g	g	g	g	g	g	a
80	g	g	g	g	g	g	a
160	a	a	g	g	g	g	g
320	r	a	a	g	g	g	g

5 Inclusion of standard dynamic range content

Definitions

Tone Mapping (TM) – Compression of the image dynamic range of content. It may be used to "down-map" (down-convert) HDR content to SDR content.

Inverse Tone Mapping (ITM) – Expansion of the image dynamic range of content. It may be used to "up-map" (up-convert) SDR content to emulate the appearance of HDR content. Also referred to as "up-mapping".

Direct Mapping – In the context of converting SDR content to HDR content, Direct Mapping preserves the appearance of the SDR content so that the HDR version displayed on a reference HDR monitor will look identical to the original SDR version displayed on a reference SDR monitor. A luminance gain (e.g. 2x) and other processing will provide a better match to the luminance of a native HDR image while maintaining the SDR appearance.

Hard Clipping – When converting from HDR to SDR there are some circumstances when hard clipping rather than tone mapping (akin to soft clipping) may be more appropriate. With hard clipping all signals above a threshold are clipped to that threshold. Hard clipping is useful when the signal from an HDR camera is required to look similar to the signal delivered by an SDR camera operated without a "knee".

Artistic Intent – A creative choice that the programme maker would like to preserve, primarily conveyed through the use of colour and tone.

Look – A characteristic of the displayed image. The native appearance of colours and tones of a particular system (for example, PQ, HLG, BT.709) as seen by the viewer.

SDR content may either be directly mapped or inverse tone mapped (up-mapped) into an HDR format for inclusion in HDR programmes. Direct mapping places SDR content into an HDR container, analogously to how content specified using BT.709 colorimetry may be placed in a BT.2020 container. This approach is intended to preserve the appearance of the SDR content when shown on an HDR display. In contrast, inverse tone mapping (up-mapping) is intended to expand the content to use more of the available HDR luminance range, and thereby leverage more of the display capabilities. Up-mapping is intended to make content captured in SDR look more as if it had been captured in HDR.

There are two possible approaches to both SDR direct mapping and up-mapping depending on the application:

– Display-referred mapping is used when the goal is to preserve the colours and relative tones seen on an SDR display, when the content is shown on an HDR display; an example of which is the inclusion of SDR graded content within an HDR programme. Display-referred mappings are derived by scaling the light reproduced by a reference display. These are known as "display-light" conversions.

– Scene-referred mapping is used when the goal is to match the colours and relative tones of an HDR and SDR camera; an example of which is the inter-mixing of SDR and HDR cameras within a live television production. Scene-referred mappings are based on the light falling on the camera sensor, but they include any camera characteristics, white balance, and any artistic camera adjustments. These are known as "scene-light" conversions.

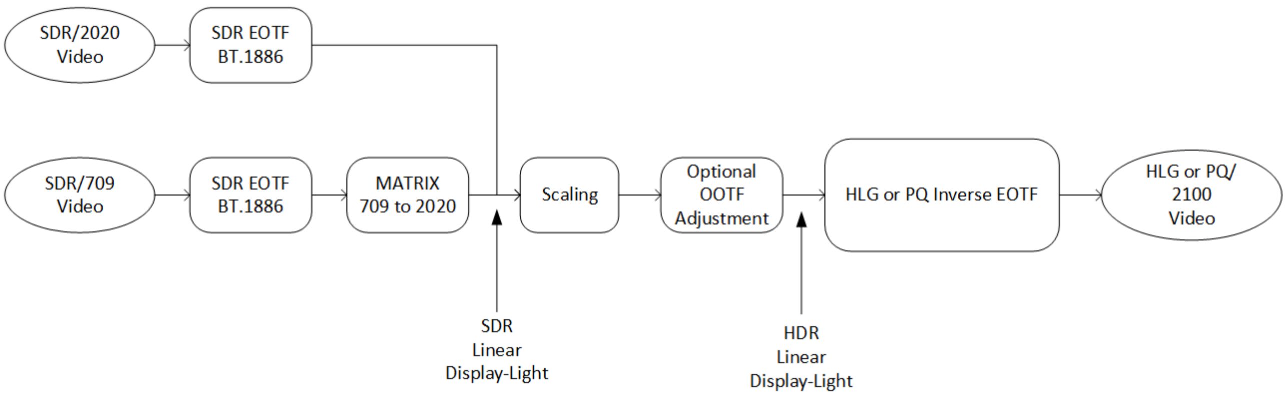
The nominal signal levels described in § 2.2 may be helpful to guide mid-tone levels during mapping.

5.1 Display referred mapping

Figure 5 illustrates the display-referred mapping of SDR signals into either HLG or PQ.

FIGURE 5

Generic method for "display-referred" mapping of SDR into HLG or PQ



The SDR signal is first passed through the BT.1886 reference EOTF to derive SDR linear display light. An approximation of the electro-optical transfer function (EOTF) from Recommendation ITU-R BT.1886 may be used:

$$E = (E')^{2.40} \qquad , \qquad 0 \leq E' \leq 1$$

where:

E' is the non-linear signal (R', G', B') in the range [0:1]

E is the normalised linear display light in the range [0:1].

A colour space conversion from BT.709 primaries to BT.2020/BT.2100 colour primaries is performed if necessary, details of which can be found in Recommendation ITU-R BT.2087.

The linear SDR display light may then be scaled to ensure that SDR and native HDR content have a similar level for HDR reference white. Where scaling is performed, a small optional adjustment to the OOTF (opto-optical transfer function) may then be applied to compensate for the subjective change in appearance of the SDR signal arising from a simple linear scaling; thereby ensuring that the visibility of detail in the shadows is maintained and that the level of skin tones in HDR and mapped SDR content are similar.

Having scaled and adjusted the SDR display light, the resulting signal is passed through an HLG or PQ inverse EOTF to provide either an HLG or PQ signal.

5.1.1 Display referred mapping of SDR into PQ

The following procedure may be followed to achieve consistent mid-tone luminance levels when mapping standard dynamic range content into PQ.

Standard dynamic range BT.2020 content should be mapped to PQ by applying the BT.1886 display EOTF and then applying the PQ $EOTF^{-1}$.

$$E' = EOTF_{PQ}^{-1}\left[scaling \times EOTF_{BT.1886}[V, L_W, L_B]\right]$$

V: Input SDR video signal level (normalized, black at $V = 0$, to white at $V = 1$)

L_W: SDR screen luminance for white = 100 cd/m^2

L_B: Screen luminance for black = 0 cd/m^2

E': Output PQ video signal level (normalized [0:1])

Scaling: $EOTF_{PQ}(E'_{V=1})$ / 100 cd/m^2

Example: for scaling = 2.0, $E'_{V=1} = 0.58$ and $EOTF_{PQ}(E'_{V=1}) = 200$ cd/m^2

For unity mapping the peak signal of standard dynamic range content would be set to 100 cd/m^2 or 51%PQ.

Unity mapping does not change the display of the SDR content (it will display on the PQ HDR reference monitor the same as it displayed on the reference SDR monitor). Thus, no OOTF adjustment of the SDR display light signal is necessary.

If the SDR content is being inserted into HDR programming, and there is desire to more closely match the brightness of the HDR content, and that brightness is known, scaling can be done to bring up the brightness of the mapped SDR content. Scaling should be performed with care lest scaled SDR content, in particular skin tones, becomes brighter than in the HDR content.

A scaling factor of 2.0 is consistent with the HDR level guidance of § 2.2, as that will map the 100 cd/m^2 peak white level of SDR to approximately the 203 cd/m^2 level suggested for HDR or 58%PQ. Also noteworthy is that MovieLabs has recommended a scaling factor of 2.0 when converting for consumer displays, as MovieLabs has found this to provide a good match to the way such displays show SDR content in their “home cinema” viewing modes [4].

For standard dynamic range BT.709 content the same process may be used, with the BT.709 to BT.2020 conversion matrix applied before the scaling as shown in Fig. 5.

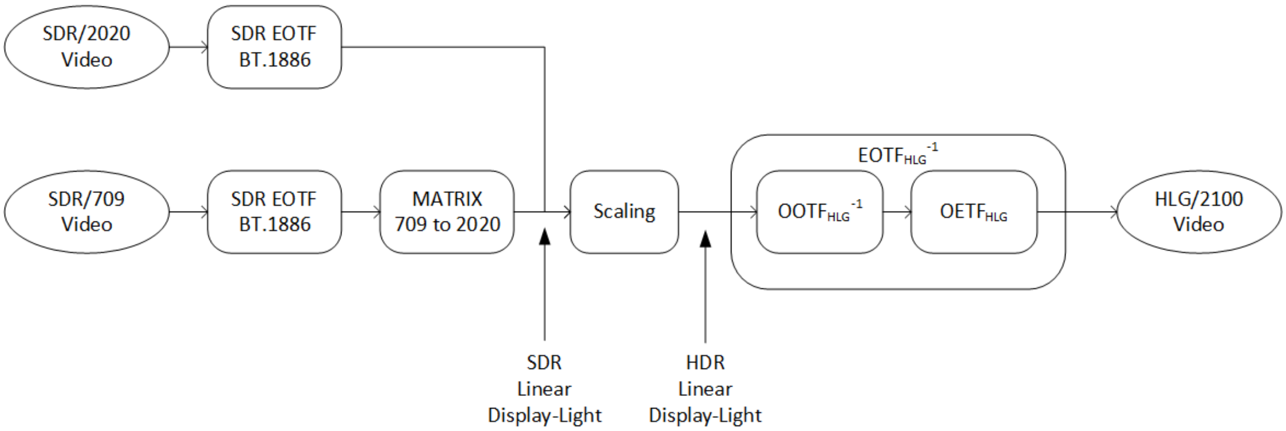
5.1.2 Display referred mapping of SDR into HLG

5.1.2.1 Mapping without gamma adjustment

The ‘display-referred’ method of mapping SDR content into a Hybrid Log-Gamma (HLG) container is illustrated below in Fig. 6.

FIGURE 6

SDR to HLG mapping without gamma adjustment (display-referred)

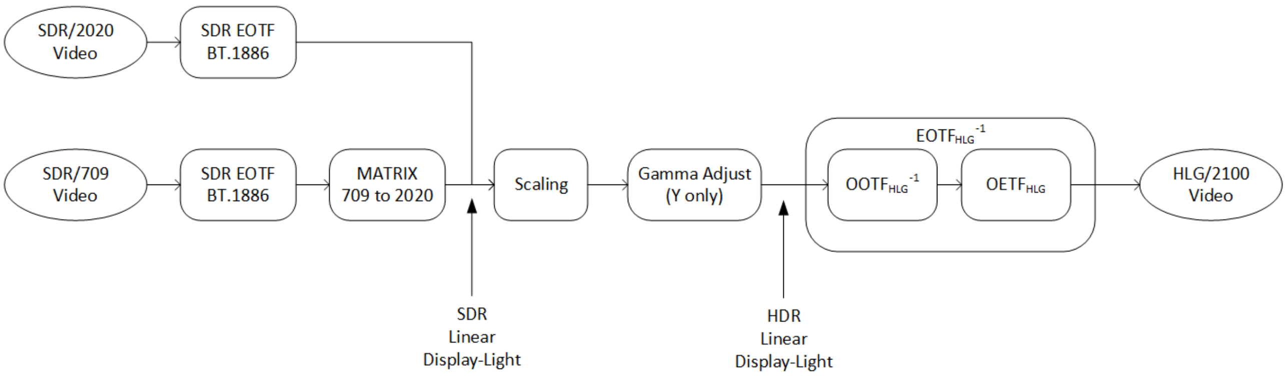


5.1.2.2 Mapping with gamma adjustment

For the case when gamma adjustment is made to the scaled SDR display light, the process is shown in Fig. 7.

FIGURE 7

Model for "display-referred" mapping with gamma adjustment of SDR into HLG



The linear SDR display light is scaled to ensure that 100% of the SDR signal is mapped to the HLG reference level 75%HLG. A small gamma adjustment may then optionally be applied to the luminance component, to compensate for the subjective change in appearance of the SDR signal arising from a simple linear scaling of the SDR display light signal.

Having scaled and adjusted the SDR display light, the resulting signal is passed through an HLG inverse EOTF to provide the HLG signal.

5.1.2.3 Scaling

When (100X)%SDR signal is mapped to (100Y)%HLG signal, a scaling gain is calculated by the following equation:

$$\text{Gain} = \frac{\text{EOTF}_{\text{HLG}}(\text{Y})}{\text{EOTF}_{\text{SDR}}(\text{X})}$$

For example, when 100%SDR signal is mapped to 75%HLG (203 cd/m^2 on a 1 000 cd/m^2 display), the scaling gain is calculated as follows:

$$\text{Gain} = \frac{\text{EOTF}_{\text{HLG}}(0.75)}{\text{EOTF}_{\text{SDR}}(1.0)} = \frac{\text{OOTF}_{\text{HLG}}\left(\text{OETF}^{-1}_{\text{HLG}}(0.75)\right)}{\text{EOTF}_{\text{SDR}}(1.0)} = \frac{0.265^{1.2}}{1.0^{2.4}} = 0.203$$

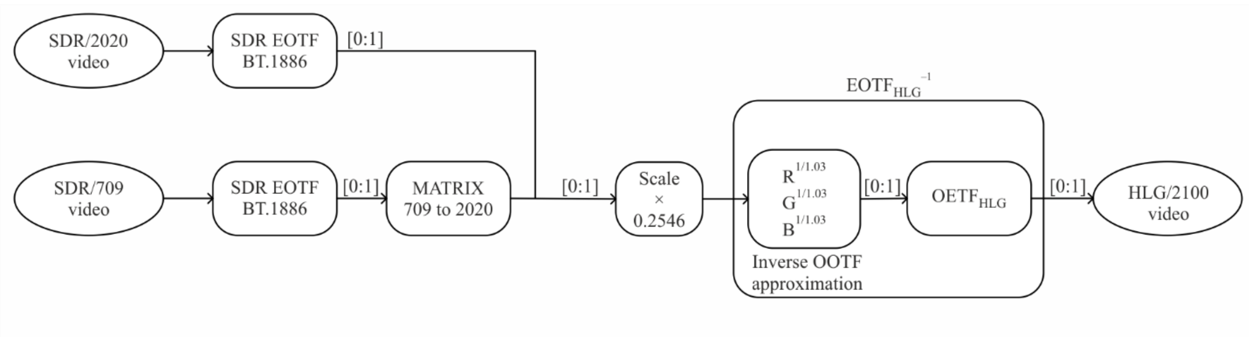
5.1.2.4 Simplification of the HLG mapping process

Through careful choice of the HLG inverse EOTF parameters, it is possible to avoid the need to scale and adjust the gamma of the SDR linear display light signal. By configuring the HLG inverse EOTF with a nominal peak luminance, L_W, of 392 cd/m^2, an input of 100 cd/m^2 from the SDR EOTF will directly deliver an HLG signal of 75%, satisfying the requirement to map 100%SDR signal to 75%HLG signal, without further scaling and gamma adjustment.

Figure 8 illustrates how, for all but the most critical applications, it is possible to simplify the conversion yet further. When applying the HLG inverse EOTF with L_W set to 392 cd/m^2, Note 5e of Recommendation ITU-R BT.2100 requires a gamma value of 1.03. As this is close to unity, in most applications there is no need to apply the inverse OOTF gamma to the luminance component, it can instead be applied independently to R, G and B components; greatly simplifying the mapping process. Colour distortions that usually arise through applying gamma to red, green and blue, rather than luminance, are barely visible for such low values of gamma.

FIGURE 8

Simplified (display-referred) SDR to mapping into HLG



Report BT.2408-08

As normalised signals are used throughout, a different scaling is required to match the signal ranges of the SDR EOTF and HDR inverse EOTF, thereby ensuring that 100%SDR signal maps to 75% of the HLG HDR signal. Note that as the normalised signals are dimensionless, the scaler is not adjusting the peak luminance of the SDR display light, so no additional gamma compensation for the signal scaling is required. Allowing for the inverse OOTF gamma of 1.03, the correct scale factor is 0.2546.

5.2 Scene referred mapping

It is particularly important that the scene-referred mapping is used for matching signals from BT.709 and BT.2020 SDR cameras with signals from HLG cameras. This is because, direct from the camera (and prior to subjective adjustment), both signals represent light from the scene captured by the camera.

If the display-referred mapping were used, which maintains the appearance of SDR images on an HLG display, the signals from SDR cameras and HLG cameras would not match. This is because the displayed 'look' of SDR and HLG images, from cameras that implement the reference OETFs (opto-electronic transfer functions), is different (see § 7.3.3 and Annex 6).

Scene-referred mapping will also work for mapping SDR to PQ. However, because the 'look' of PQ and BT.2020 SDR signals is very similar, for BT.2020 SDR signals the display-referred mapping will generally work well. To best match the PQ ‘look’, BT.709 SDR camera signals could be converted to BT.2020 SDR camera signals (using an OETF-based conversion similar to that specified in Recommendation ITU-R BT.2087) before display-referred mapping is applied.

The schematic diagram of the scene-referred mapping is illustrated in Fig. 9 for both PQ and HLG. It includes an optional artistic OOTF adjustment, for example to match the ‘traditional colour reproduction’ described in § 6.5 of Report ITU-R BT.2390.

FIGURE 9

SDR to HDR mapping (scene-referred)

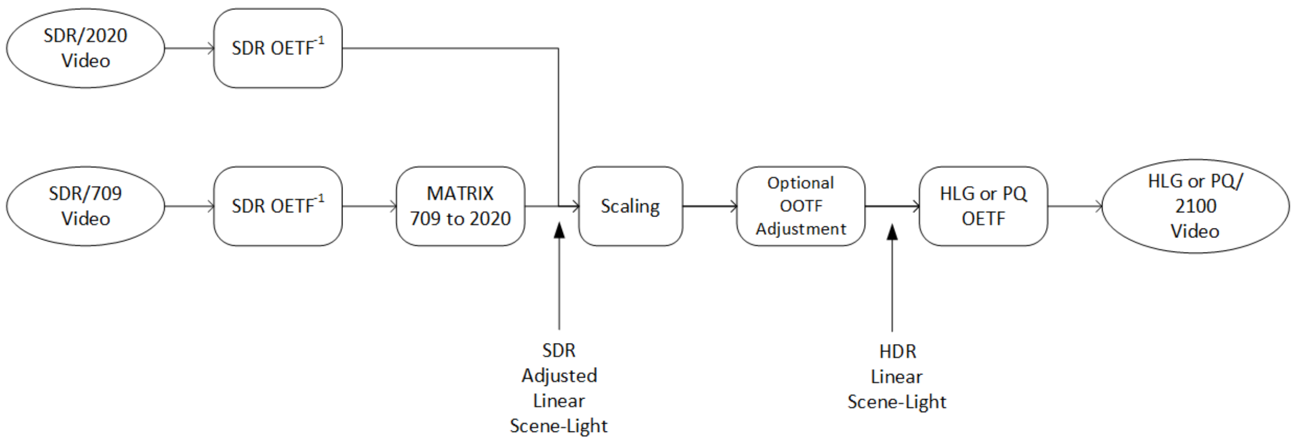


Figure 9 above shows how the non-linear SDR BT.709 or BT.2020 video signal is converted to linear ‘scene light’ by applying the approximate inverse of SDR OETF, $E=(E')^2$, as described in BT.2087. When the SDR source is with the BT.709 colorimetry, the conversion is followed by the colour conversion matrix as described in Recommendation ITU-R BT.2087.

The scene light signal is then scaled so that the non-linear signal, after applying the reference PQ or HLG OETF, is at the appropriate signal level for HDR reference white: 58%PQ or 75%HLG respectively. Following any OOTF adjustment, the HLG or PQ OETFs are applied to derive the non-linear signals.

Section 5.2.1 describes how to calculate the scale factor for HLG, as well as how to adjust the OOTF to preserve a traditional SDR look.

5.2.1 Scene referred mapping of SDR into HLG

When (100X)%SDR signal is mapped to (100Y)%HLG signal, a scaling gain is calculated by the following equation:

$$\text{Gain} = \frac{\text{OETF}_{\text{HLG}}^{-1}(\text{Y})}{\text{OETF}_{\text{SDR}}^{-1}(\text{X})}$$

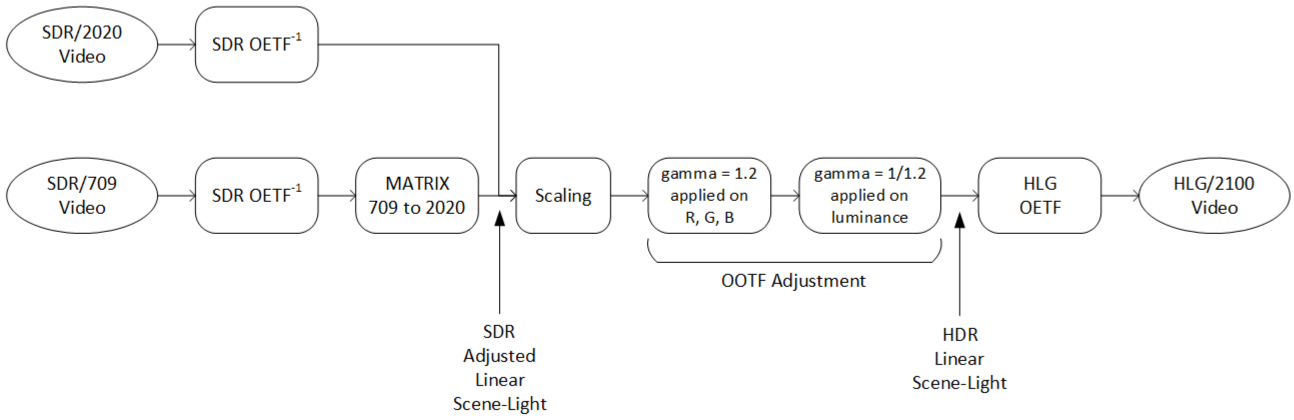
For example, when 100%SDR signal is mapped to 75%HLG signal, the scaling gain is calculated as follows:

$$\text{Gain} = \frac{\text{OETF}_{\text{HLG}}^{-1}(0.75)}{\text{OETF}_{\text{SDR}}^{-1}(1.0)} = \frac{0.265}{1.0^{2.0}} = 0.265$$

Where the SDR “look” is maintained during the conversion from SDR to HDR or the HLG camera is designed to deliver a traditional ‘look’ (see § 6.5 of Report ITU-R BT.2390), a small optional adjustment to the OOTF may then be applied to compensate for the subjective change in appearance of the SDR signal arising from a difference between HLG and SDR OOTFs. For the case when gamma adjustment is made to the scaled SDR scene light, the process is illustrated in Fig. 10.

FIGURE 10

SDR to HLG mapping with gamma adjustment (scene-referred)



5.3 Handling negative values in format conversion

It is common practice for camera OETFs and display EOTFs implemented within format converters to be extended to handle negative signals by reflecting the transfer functions around the zero light and zero signal axes. Extending the transfer functions in this way can be useful for increasing the colour gamut carried by a “narrow” range signal (see § 5.4) and for processing test signals such as PLUGE.

In format conversion, however, this could lead to an increase in “round-trip” errors. So the best approach will depend on the application.

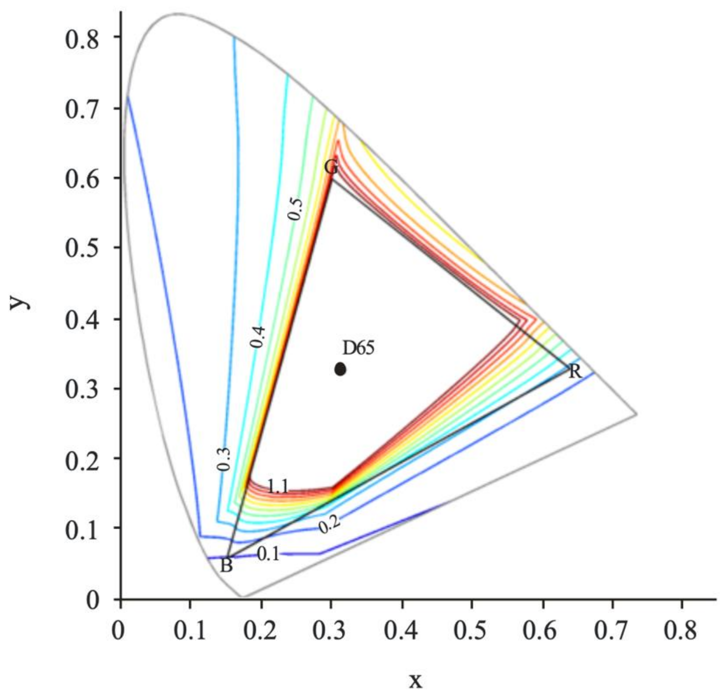
5.4 Adjustments to BT.709 cameras

It may be beneficial to include signals below black (sub-blacks) and above the SDR nominal peak white (super-whites) in the conversion process from SDR BT.709 to HDR. Such signals, which are often present in live SDR television production, effectively increase the colour gamut captured by the camera beyond the BT.709 colour primaries. More details are provided in Report ITU-R BT.2250.

The permitted SDR signal ranges vary between geographical regions. By way of an example, EBU R103 [5] allows SDR signals to span −5% to +105%. Figure 11 illustrates the maximum transmissible $Y'C'_BC'_R$ colour gamut. The contours are drawn for each normalized Y at an interval of 0.1 on the CIE 1931 xy chromaticity diagram. Negative values of R′, G′ and B′ widen the effective colour primaries. The gamut is increased in the red and the blue, and a smaller increase is also made in the green. Allowing the *R′G′B′* signals to extend above 100% increases the colour volume by allowing more saturated colours at higher luminance.

FIGURE 11

Extending the BT.709 camera colour gamut



The technique can be used to ensure a closer match between BT.709 and BT.2100 cameras for colours that are close to the BT.709 colour volume boundary.

Where the SDR BT.709 camera output is only used for shading and as the input to an SDR to HDR format converter, the signal clippers can be fully relaxed to maximise the captured colour volume. Not all format converters and production infrastructure are capable of passing the sub-black and super-white signals.

5.5 Use of 8-bit content

Although a minimum of 10-bits should be used for HDR production, there may be occasions when it might not be possible to avoid including 8-bit SDR content within an HDR programme. In such cases, care should be taken if up-mapping rather than direct mapping is used to place the content into an HDR signal container. The up-mapping process typically expands the SDR highlights. The 8-bit resolution, compounded by any 8-bit video compression, will limit the amount of highlight expansion that can be applied before banding and other artefacts become visible.

5.6 Mapping of SDR graphics

SDR graphics should be directly mapped into the HDR signal at the “Graphics White” signal level specified in Table 1 (75%HLG or 58%PQ) to avoid them appearing too bright, and thus making the underlying video appear dull in comparison. Where the desire is to maintain the colour branding of the SDR graphics, a display-light mapping should be used. Where the desire is to match signage within the captured scene (in-vision signage; e.g. a score board at a sporting event), a scene-light mapping is usually preferred.

6 Conversion between PQ and HLG

6.1 Transcoding concepts

Transcoding aims to produce identical display light when the transcoded signal is reproduced on a display of the same peak luminance as the original signal. This section describes how a PQ signal may be transcoded to an HLG signal and vice versa, although cascaded conversions are to be discouraged to avoid risking loss of quality.

Figure 12 illustrates the concept behind transcoding from the PQ signal to the HLG signal. The PQ signal is decoded by the PQ EOTF to yield a signal that represents linear display light. This signal is then encoded by the HLG inverse EOTF to produce an equivalent HLG signal. When this HLG signal is subsequently decoded by the HLG EOTF in the display, the result will be the same display light that would be produced by decoding the original PQ signal with the PQ EOTF. The HLG inverse EOTF is the HLG inverse OOTF followed by the HLG OETF.

FIGURE 12

Concept of transcoding from PQ to HLG

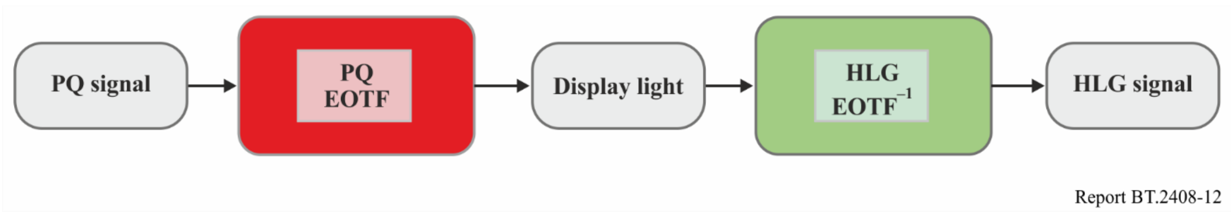
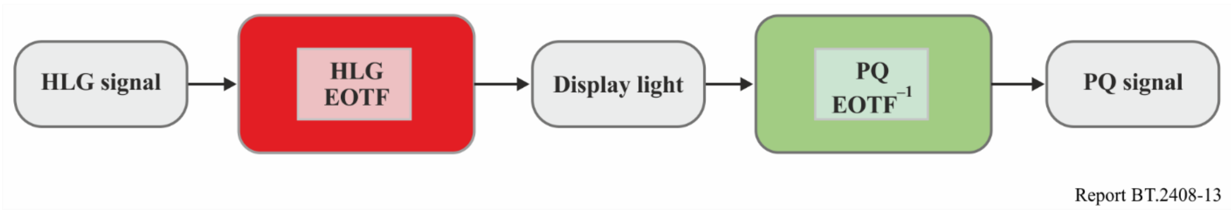


Figure 13 illustrates the concept behind the transcoding from the HLG signal to the PQ signal. The HLG signal is decoded by the HLG EOTF to yield a signal that represents linear display light. This signal is then encoded by the PQ inverse EOTF to produce an equivalent PQ signal. When this PQ signal is subsequently decoded by the PQ EOTF in the display, the result will be the same display light that would be produced by decoding the original HLG signal with the HLG EOTF.

FIGURE 13

Concept of transcoding from HLG to PQ



6.2 Conversion concepts using a reference condition at 1 000 cd/m^2

The transcoding concepts in the previous section produce the same displayed light for both PQ and HLG signals only when they are viewed on displays with the same peak luminance.

However, the difference in the way that PQ and HLG signals are rendered on displays of different peak luminance complicates the conversions between PQ and HLG signals. If, for example, PQ signals, representing different peak luminances, are simply transcoded to HLG, the signal level for diffuse white will vary. Similarly, when HLG content is transcoded to PQ the brightness of diffuse white will vary depending on the assumed peak luminance of the HLG display.

To avoid such brightness changes, it is needed to convert, rather than simply transcode, the signals. Consistent brightness in the converted signals may be achieved by choosing a reference peak displayed luminance (L_W) for the HLG signal, and requiring that PQ signal be limited to the same peak luminance. With these constraints consistent brightness is achieved in the converted signals.

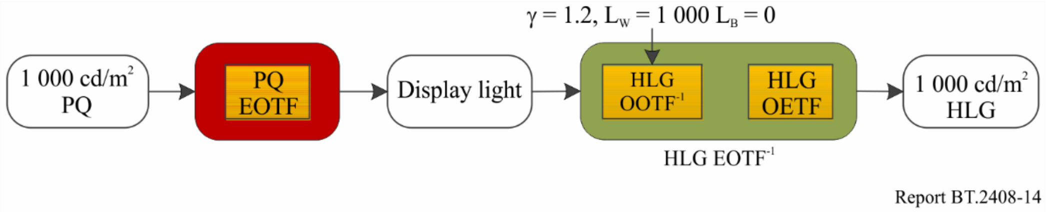
Therefore it is desirable that conversion between PQ and HLG should take place using the same reference peak displayed luminance for the signals used in the conversion. There is currently an industry consensus that this common peak luminance should be 1 000 cd/m^2.

For both transcoding and conversion a black level for the HLG EOTF also needs to be specified. The HLG black level, L_B, should be set to zero for transcoding and conversion.

With the choice of 1 000 cd/m^2 as the common peak luminance, the conversion outlined above is completely specified for any HLG signal to PQ and, for PQ signals not exceeding 1 000 cd/m^2, from PQ to HLG. Figure 14 illustrates the conversion from PQ to HLG.

FIGURE 14

Conversion from PQ to HLG at a common peak luminance of 1 000 cd/m^2



The following is an elaboration of Fig. 14 in terms of the three most fundamental transformations:

(1) The PQ EOTF and its inverse

(2) The HLG OETF and its inverse

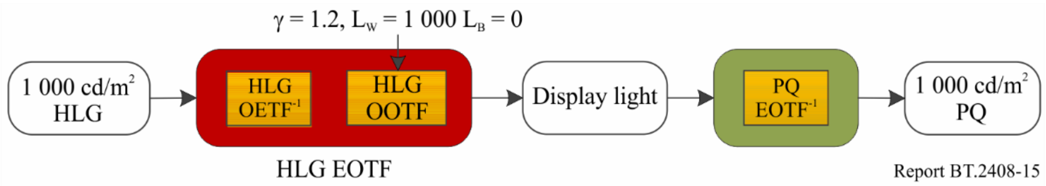
(3) The HLG OOTF and its inverse.

The HLG EOTF is derived from (2) and (3). The Figure also includes the parameters for HLG OOTF^{-1}. The resulting HLG signal will produce images identical to the original PQ images for all content that is within the colour volume of the 1 000 cd/m^2 HLG reference display.

Analogously, the conversion from HLG to PQ at 1 000 cd/m^2 is the inverse of the above as illustrated in Fig. 15.

FIGURE 15

Conversion from HLG to PQ at a common peak luminance of 1 000 cd/m^2



This conversion always produces a PQ image identical to HLG.

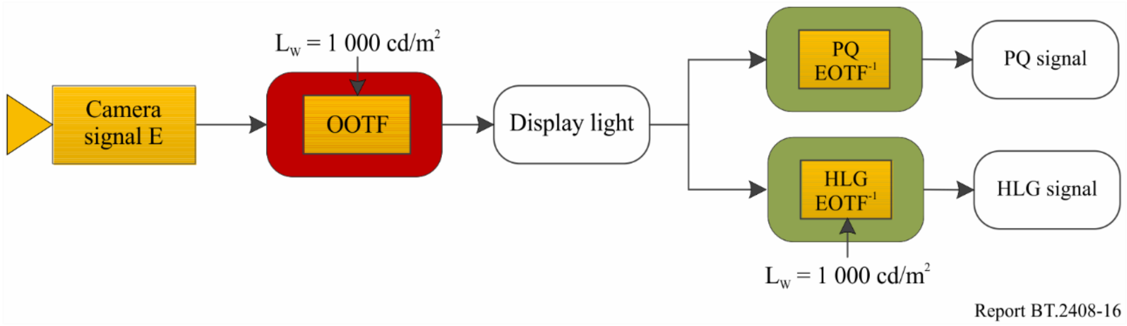
6.3 Cameras using a common OOTF at a reference peak luminance of 1 000 cd/m^2

Cameras could apply a common OOTF to produce PQ and HLG signals with identical displayed images at a reference peak luminance of Lw = 1 000 cd/m^2.

This OOTF could be the PQ OOTF, or the HLG OOTF, and might include additional modifications applied in the camera, as illustrated in Fig. 16. PQ and HLG signals are obtained using their respective inverse EOTFs.

FIGURE 16

Use of a common OOTF to provide both PQ and HLG at a common peak luminance of 1 000 cd/m²



The appearance of the displayed images will be the same on displays with a peak luminance capability of 1 000 cd/m^2, for both the PQ and HLG signals. The appearance of the image is determined by the OOTF.

6.4 Handling PQ signals with greater than 1 000 cd/m^2 peak luminance

PQ signals can represent a peak luminance of up to 10 000 cd/m^2. In order to enable the reference conversion described above, PQ content must be limited to have a peak luminance that does not exceed 1 000 cd/m^2. There are, in general, three approaches to achieving this:

(1) Clip to 1 000 cd/m^2

(2) Static mapping to 1 000 cd/m^2 (e.g. using an EETF curve like those described in § 3.1.1)

(3) Dynamic mapping to 1 000 cd/m^2

The first method, clipping to 1 000 cd/m^2, is simple to implement. While multiple round trip conversions between PQ and HLG are to be discouraged, with this method content undergoes no additional limiting/clipping in the event of multiple round-trip conversions (i.e. PQ->HLG->PQ->HLG) beyond the initial clipping.

The second method, static mapping to 1 000 cd/m^2 can be implemented by a LUT containing an EETF such as that described in § 3.1.1. While this avoids hard clipping of detail in the highlights, it is not invariant under blind multiple round-trip conversions.

The third method, dynamic mapping to 1 000 cd/m^2, utilizes adaptive processing, for example on a frame-by-frame, or scene-by-scene basis. An adaptive algorithm could vary the EETF described in § 3.1.1 based on statistics of the image content (scene maximum for example). For non-live content, dynamic mappings could be generated offline by the content producer (either manually or using algorithmic processing). Except for the initial stage of limiting the PQ signal to 1 000 cd/m^2, this approach could survive multiple round-trip conversions, because subsequent dynamic processing should be inactive given that the signal would already have been limited to 1 000 cd/m^2.

6.5 Possible colour differences when converting from PQ to HLG

In principle, the conversion of PQ images to HLG could give rise to hue shifts or desaturation on bright highly saturated areas of the picture, although such effects are believed to be rare in practice.

Mathematically, this arises because the OOTF applied in the display for HLG is a function of overall luminance rather than identical functions of R, G, and B. Consider the equations for luminance in both the display and scene domains along with the EOTF for HLG:

$$Y_D = 0.2627R_D + 0.6780G_D + 0.0593B_D$$

$$Y_S = 0.2627R_S + 0.6780G_S + 0.0593B_S$$

$$R_D = \alpha Y_S^{\gamma-1} R_S$$

$$G_D = \alpha Y_S^{\gamma-1} G_S$$

$$B_D = \alpha Y_S^{\gamma-1} B_S$$

Table 6 below summarizes the peak values that can be displayed for pure white, and for the red, green and blue primaries, for a 1 000 cd/m^2 PQ monitor, and for a 1 000 cd/m^2 nominal peak HLG monitor. The Table shows values of 'x' such that when the non-linear signal values $R' = G' = B' = x$ the resulting white is 1 000 cd/m^2. For PQ, this occurs when x is approximately 0.75; for a 1 000 cd/m^2 HLG display, this occurs when $x = 1.0$. For a 1 000 cd/m^2 PQ display, the maximum luminance of each of these colours is calculated using Y_D and is shown in the middle column of the Table. For HLG we can simplify the EOTF by normalizing scene colours within [0:1]. Thus:

$$R_D = 1000 Y_S^{\gamma-1} R_S, \text{ etc.}$$

When $x = 1$, so is the normalized scene colour and its non-linear representation, E', for the given component. This determines $\{R_D, G_D, B_D\}$ and the resulting luminance is calculated using Y_D.

However, in production, HLG signals usually adopt the narrow range quantization levels specified in Recommendation ITU-R BT.2100. As noted in § 6.1 of Report ITU-R BT.2390, the conventional "narrow range" digital signal can support signal levels of up to 109% of nominal full scale when carried as 10-bit or 12-bit signals ($E' = 1.090$). So, the extended signal range between $E' = 1.0$ and $E' = 1.090$, sometimes referred to as "super-whites", may be used to increase the HLG colour volume. The rightmost column in the Table shows the HLG maximum signal level required to match the displayed luminance of a 1 000 cd/m^2 PQ display. Thus, they also represent the HLG signal ranges necessary to precisely convert a 1 000 cd/m^2 PQ signal to HLG without clipping, thereby eliminating any risk of hue shifts or desaturation.

TABLE 6

Signal ranges and achievable colour volume for PQ and HLG on a 1 000 cd/m^2 nominal peak luminance display

Colour	BT.2100 PQ Y cd/m^2 for 1 000 cd/m^2 peak white	BT.2100 HLG Y cd/m^2	
	x = 0.75	x = 1.0	Max non-linear signal, E', to match PQ luminance
{x,x,x} // Peak white	1 000.0	1 000.0	$R' = G' = B'$=1.000
{x,0,0} // Peak red	262.7	201.1	R'=1.041
{0,x,0} // Peak green	678.0	627.3	G'=1.012
{0,0,x} // Peak blue	59.3	33.7	B'=1.086

NOTE – In practice some displays might not achieve a luminance output higher than their nominal peak value.

By exploiting the quantization levels above E' = 1.0, HLG is able to deliver the same 1 000 cd/m^2 colour volume as PQ, without clipping. This could be particularly useful when converting graded PQ content to HLG. Furthermore, the peak luminance for white on an HLG reference display is increased from 1 000 cd/m^2 to 1 811 cd/m^2. Note: the HLG OOTF system gamma is still calculated using the display's nominal peak luminance at x = 1.0.

7 Transitioning from SDR BT.709 to HDR BT.2100 production

During the transition from SDR to HDR production, the majority of viewers will be watching in SDR, so it is important that the SDR production is not significantly compromised by the introduction of HDR. It is, however, unlikely to be economic or practical to cover live programmes and events with totally independent HDR and SDR production facilities. As well as the cost of the two productions, there may simply be insufficient camera positions available for both HDR and SDR cameras.

Native HDR production architectures highlighting either HDR or SDR focussed production are illustrated in Figs 17 and 18. Over time, as audiences adopt HDR television displays designed for BT.2100 signals, production architectures may be expected to shift from focussing on delivering primarily for SDR, to delivering primarily for HDR.

Note that in both production architectures the eye may adapt to the brighter HDR monitor, affecting the appearance of signals on the dimmer SDR screen. So, the HDR and SDR screens should be physically separated for critical assessment of the SDR signal.

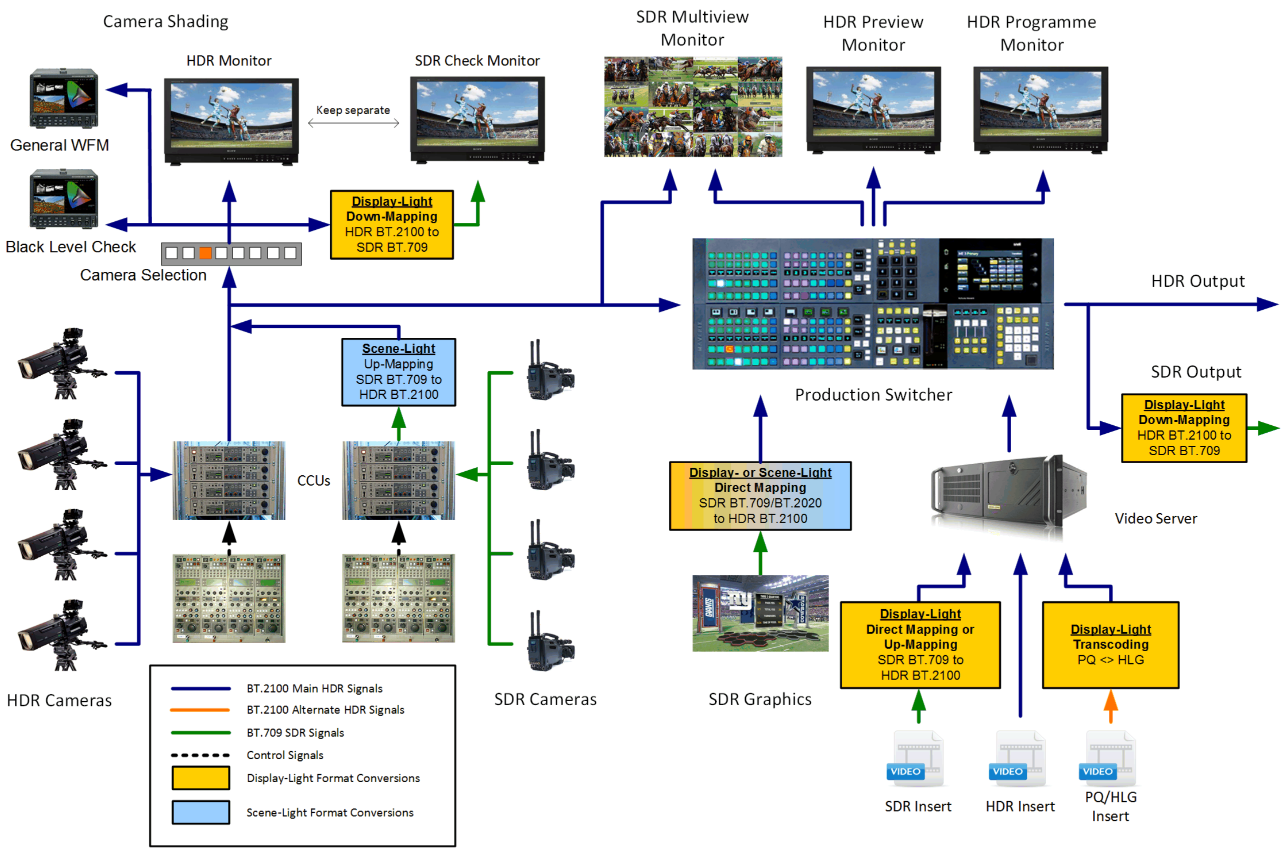
7.1 HDR focused production

For optimum quality HDR pictures, both HDR and SDR cameras should be shaded using an HDR monitor, as illustrated in Fig. 17. Nominal signal levels for shading are given in § 2.2.

As the exposure latitude of HDR images is far greater than SDR, a dynamic HDR to SDR converter may be required to deliver a satisfactory SDR output. A dynamic converter is designed to optimise the HDR to SDR tone mapping curve for any scene, thereby accommodating a wider range of exposures than might be possible with a fixed (or static) tone mapping curve.

FIGURE 17

HDR production with SDR derived by down-mapping



In this HDR focused production, BT.709 cameras may be included in the production by using the "scene-referred" SDR direct mapping technique, as described in § 5. To ensure a closer match between HDR and SDR cameras, up-mapping (which expands highlights in the SDR signal) is preferred. As highlights are often heavily clipped by SDR cameras, only a small amount of highlight expansion may be possible. Further colour match improvements can be made by relaxing the SDR signal clippers, as described in § 5.4.

In Fig. 17, all inputs to the production switcher are HDR. This removes the need to process separate HDR and SDR feeds throughout the production chain. Graphics may be inserted as per § 5.6. Work is currently underway to determine the best practice for HDR key signals. In the interim, using an SDR key signal directly has been found to deliver satisfactory results.

The primary output from the production switcher is HDR. The SDR output is derived via display referred (display-light) down-mapping. A display-light conversion ensures that both the SDR and HDR signals have the same look. A dynamic down-mapper may sometimes provide a more satisfactory SDR output than a static down-mapper, but attention should be paid to graphics which may need to be inserted after dynamic down-mapping, to ensure a fixed signal level. A scene-light HDR to SDR conversion may also be included (not shown in Fig. 17) where it is important to colour match the converted PQ or HLG output to downstream SDR BT.709 cameras. However, consideration should be given to potential changes in colour saturation of graded content (see § 7.3.3). Ultimately, the choice of HDR to SDR down-mapping depends on the application.

Differences in black level may be more visible in the down-converted SDR signal than in the HDR signal, as glare from bright highlights in the HDR image can mask detail in the shadows. To help ensure a consistent black level in the HDR and down-converted SDR signals, a dedicated waveform monitor displaying the lower portion of the signal range is recommended.

7.2 SDR focused production

If the SDR production must not be compromised, both HDR and SDR cameras should be shaded using an SDR monitor fed via a down-mapper. Whilst the HDR signals may not always exploit the full potential of the HDR production formats, the HDR pictures can still show significant improvement over SDR.

7.2.1 PQ production

SDR focused PQ production uses the same workflow as shown in Fig. 17 except:

- the SDR check monitor is now the shading monitor;
- the HDR shading monitor is now the check monitor.

An additional scene-light PQ to SDR BT.709 conversion may also be included for colour matching with downstream SDR BT.709 cameras.

7.2.2 HLG production

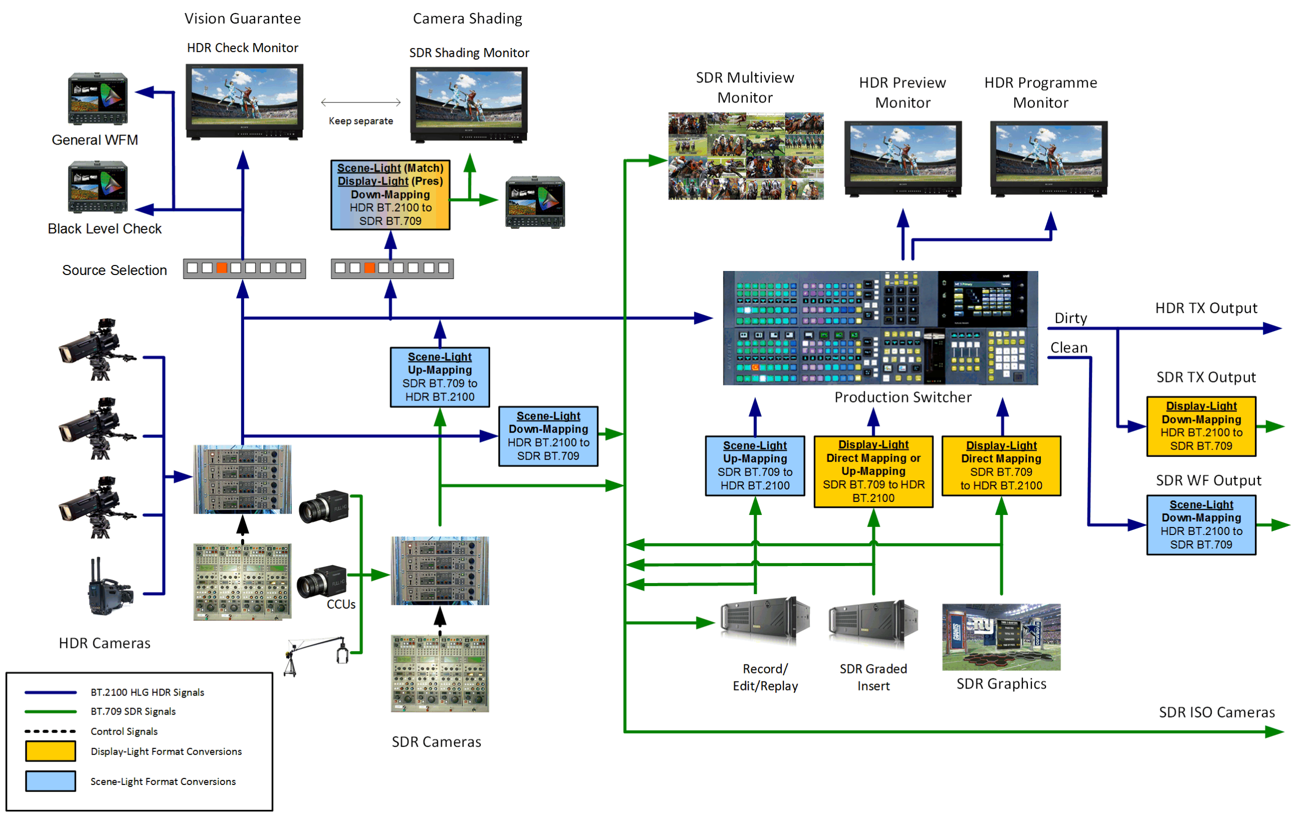
SDR focused HLG production can use the same workflow as shown in Fig. 17 except:

- the SDR check monitor is now the shading monitor;
- the HDR shading monitor is now the check monitor.

A trial carried out by the BBC is illustrated in simplified form in Fig. 18.

FIGURE 18

HDR production with camera shading in SDR



In a large sports production, such as that shown in Fig. 18, it is common for the host broadcaster to provide a "clean" SDR programme output (i.e. without graphics) to other broadcasters, sometimes called a "World Feed". This may be derived from the HDR signal using a scene-light conversion, to match other broadcasters' SDR cameras that may also be present at the venue. For any SDR output containing graphics (e.g. for the broadcaster's own SDR service) a display-light conversion is recommended, as that should ensure the same hue and saturation of graphics in both HDR and SDR outputs (see § 7.3.3). Down-mapping (tone-mapping) when converting to SDR, rather than hard clipping, will allow the SDR output to benefit from the high dynamic range production by preserving some detail in the image highlights.

As with Fig. 17, graded content should be inserted into the programme using display-light direct mapping or up-mapping, to preserve its original "look" and the artistic intent; SDR graphics should be directly mapped into the HDR format (as per § 5.6). Note that after "round-tripping" any SDR material included in the production using a display-light conversion (e.g. graded inserts or graphics) will appear more saturated in the SDR scene-light output, than in the original SDR version. So, in general, a display-light conversion to SDR on the final programme output is preferred.

During the transition to full HDR production, not only will it be common to include SDR BT.709 cameras within a production, but locally recorded action replays and programme inserts may also be limited to SDR BT.709. Additionally, a host broadcaster may be required to provide SDR BT.709 ISO (isolated/independent) camera feeds to other broadcasters. In such circumstances, the SDR camera outputs can be recorded and output directly (illustrated in Fig. 18), but HDR cameras should be converted to SDR BT.709 using a scene-light conversion to match the native SDR cameras. Complementary scene-light down-mapping and scene-light up-mapping can be used on the input and output of the replay servers, to minimise the "round-trip" losses.

To ensure the highest quality SDR output, cameras are shaded using an SDR monitor fed via identical HDR to SDR converters to those used on the main programme output. Where the "World Feed" is considered the main output, it may be via a scene-light converter. Where the broadcaster's SDR services is considered the main output, it should be via a display-light converter. In practice, the differences between the two may be small, and within the usual range of artistic tolerances for SDR production. In the case of the BBC trial illustrated in Fig. 18, a scene-light conversion was used for the cameras covering the main football match ("Match") and a display-light conversion was used for the cameras covering the presentation ("Pres") studio.

Changes in exposure of the image may be more visible in the HDR output than the SDR output. So rapid adjustments in exposure whilst shading in SDR should be avoided.

Under controlled studio lighting, a possible option may be to shade the cameras using the HLG backwards compatible SDR picture, rather than via a dedicated HDR to SDR converter. In this case, the SDR shading monitor should be set to a display gamma of 2.2 with BT.2020 colour, to resemble a typical display-light conversion from HLG to SDR as shown on a BT.1886 (gamma 2.4) production monitor. However, under variable lighting conditions or in territories where SDR skin tones are set brighter, a dedicated HDR to SDR converter is preferred.

The SDR monitors used for camera shading should be separated from the HDR check monitor that is used to ensure that high quality HDR output is being maintained (indicated as "Vision Guarantee" in Fig. 18). In the BBC live production trial occasional checks of the HDR output by a vision supervisor were found to be sufficient, with operators concentrating on the SDR monitors used for camera shading.

The BBC also found that in some situations, for example within the confined space of an outside broadcast truck, it is not practicable to achieve complete separation between the SDR and HDR monitors in the control room. As critical monitoring is in SDR, to avoid camera shader operators being affected by glare from an HDR check monitor, the nominal peak luminance of the HLG HDR monitor can be reduced, for example to 600 cd/m^2 (with an appropriate gamma adjustment, see § 3.2) to reduce the disturbance.

A fundamental difference between Figs 18 and 17 is that here, in Fig. 18, an additional scene-light SDR output signal is provided with the "traditional" BT.709 look, whilst the HLG HDR signal and display-light SDR signal have the HLG look. By design, if no further artistic adjustments are made, HLG signals preserve the chromaticity of the scene as imaged by the camera, when compared with the "traditional" looks of SDR BT.709 and BT.2020 cameras (as described in Report ITU-R BT.2390).

7.2.2.1 Further details for HLG production

Some HDR cameras conveniently provide parallel HDR and SDR signal outputs. Where that is the case, the cameras can be shaded using their SDR output, and the HDR output allowed to follow with a fixed "gain offset" (equivalent to an exposure offset) relative to the SDR. This approach relies on the SDR and HDR camera outputs precisely tracking one another, which may not always be the case. Operational staff may also have concerns about shading the cameras using a signal that is not exactly the same as that being used to feed their main SDR output.

Shading the cameras using an SDR output allows the HDR signals to be created in such a way that they closely follow the reference levels specified in Table 1, and are therefore well-conditioned for conversion to SDR. This is achieved by applying the fixed gain to the linear HDR signal such that a 90% reflectance object is portrayed with a 100% signal level in the SDR signal, and a 73% signal level in the HLG HDR signal. In the SDR signal used for shading, highlights greater than the super-white signal level (109%) are lost, but they are retained in the HDR signal.

When this approach is used for the main cameras, those cameras that only provide an HDR output should be shaded using a scene-light conversion to SDR. The scene-light conversion will provide images that more closely resemble those from traditional SDR cameras (and any HDR camera with SDR outputs) that may also be included in the production. To ensure that the HDR signals comply with the levels specified in Table 1, and to better match those of SDR cameras (in situations when a "knee" is not used), a hard clip to SDR rather than tone mapping is preferred.

7.3 SDR-HDR and HDR-SDR format conversion

This section consists of a summary of the format conversions suggested in §§ 7.1 and 7.2. The final choice of conversion will, however, be dependent on the producer's intent.

7.3.1 PQ conversion

Table 7 illustrates the suggested format conversions for PQ production.

TABLE 7

Suggested format conversions for PQ live production

Signal		Conversion Type		SDR to PQ		PQ to SDR		HLG to PQ
		Scene-light	Display-light	Direct mapping	Up-mapping	Hard clip	Down-mapping	Trans-coding
Graded content	SDR graded inserts		✓	✓ (1)	✓ (2)			
	HLG graded inserts		✓					✓
Cameras	SDR camera (relaxed clippers for BT.709)	✓ (4)			✓			
	HLG camera	✓						✓
Graphics	SDR matching colour branding		✓	✓				
	SDR matching in-vision signage	✓		✓				
SDR output(3)	SDR complete programme		✓				✓	
	SDR for downstream mixing with SDR cameras	✓					✓	

(1) Direct mapping faithfully maintains the original SDR look.

(2) Up-mapping adjusts the distribution of highlights of the original SDR look.

(3) SDR Output refers to conversion from HDR to both the final programme output as well as the SDR shading/check monitor.

(4) In PQ based production, the difference between display-light and scene-light conversion of BT.2020 signals is relatively minor and current practice is to use display-light conversion. Conversion from BT.709 to BT.2020 is defined in Recommendation ITU-R BT.2087.

7.3.2 HLG conversion

Table 8 illustrates the suggested format conversions for HLG production.

TABLE 8

Suggested format conversions for HLG live production

Signal			Conversion Type		SDR to HLG		HLG to SDR		PQ to HLG
			Scene-light	Display-light	Direct mapping	Up-mapping	Hard clip	Down-mapping	Trans-coding
Graded Content		SDR graded inserts		✓	✓ (1)	✓ (2)			
Graded Content		PQ graded inserts		✓					✓
Cameras	To switcher	SDR camera (relaxed clippers for BT.709)	✓			✓			
Cameras	To switcher	PQ camera	✓						✓
Cameras	To shading	HDR camera with SDR shading	✓				✓		
Cameras	To shading	SDR camera with HDR shading	✓			✓			
Graphics		SDR matching colour branding		✓	✓				
Graphics		SDR matching in-vision signage	✓		✓				
SDR Output(3)		SDR complete programme		✓				✓	
SDR Output(3)		SDR for downstream mixing with SDR cameras	✓					✓	

(1) Direct mapping faithfully maintains the original SDR look.

(2) Up-mapping adjusts the distribution of highlights of the original SDR look.

(3) SDR Output refers to conversion from HDR to both the final programme output as well as the SDR shading/check monitor.

7.3.3 The displayed "look" of content following format conversion

The different SDR and HDR production formats have different looks, as discussed in detail in Annex 6.

Thus, SDR to HDR and HDR to SDR format conversion may change the displayed look of content. Tables 9 and 10 summarise the look of content for HLG and PQ live production, after the format conversions specified in Tables 7 and 8.

One notable consideration is the possible change of look occurring when the input and output conversion types do not match. Scene-light HDR to SDR format conversion, necessary for downstream mixing with SDR BT.709 cameras, may cause some SDR graded content (inserted via display-light conversion) to appear more saturated than intended for HLG HDR production, or slightly less saturated than intended for PQ HDR production. Scene-light conversion to SDR should therefore be used with care, and multiple such conversions should be avoided.

Graded content does not carry a specific SDR or HDR look, but instead has an artistic look imposed upon it by the colourist.

TABLE 9

Display look of content after format conversion for HLG Production

Signal			Input conversion type		SDR output conversion following HLG production		
			Scene-light	Display-light	Scene-light(1)		Display-light(2)
					To BT.709	To BT.2020	BT.709 and BT.2020
Graded Content		SDR graded inserts		✓	Over saturated	Over saturated	Maintaining artistic intent(4)
Graded Content		PQ graded inserts		✓	Over saturated	Over saturated	Maintaining artistic intent(4)
Cameras	To switcher	SDR BT.709 camera	✓		SDR BT.709 look	SDR BT.2020 look	HLG look(3)
Cameras	To switcher	SDR BT.2020 camera	✓		SDR BT.709 look	SDR BT.2020 look	HLG look(3)
Cameras	To shading	HDR camera with SDR shading	✓		SDR BT.709 look	SDR BT.2020 look	HLG look(3)
Cameras	To shading	SDR camera with HDR shading	✓		SDR BT.709 look	SDR BT.2020 look	HLG look(3)
Graphics		SDR matching colour branding		✓	Over saturated	Over saturated	Maintaining artistic intent(4)
Graphics		SDR matching in-vision signage	✓		SDR BT.709 look	SDR BT.2020 look	HLG look(3)

(1) Scene-light conversion is used to match downstream SDR cameras but is not the preferred method for SDR output conversion.

(2) Display-light conversion is generally the preferred SDR output method and will preserve the look of graded content and graphics that originated in SDR or PQ.

(3) HLG, SDR BT.2020 and SDR BT.709 have different looks, as discussed in Annex 6.

(4) Graded Content and Graphics content do not necessarily have the native SDR or HLG look. The "Artistic Intent" may have been to make them more saturated, have different contrast, etc.

TABLE 10

Display look of content after format conversion for PQ Production

Signal		Input conversion type		SDR output conversion following PQ production		
		Scene -light	Display -light	Scene-light(1)		Display-light(2)
				To BT.709	To BT.2020	BT.709 and BT.2020
Graded Content	SDR graded inserts		✓	Slightly under saturated	Similar to artistic intent	Maintaining artistic intent(4)
	HLG graded inserts		✓	Slightly under saturated	Similar to artistic intent	Maintaining artistic intent(4)
Cameras	SDR camera	✓		SDR BT.709 look	SDR BT.2020 look	PQ look(3)
	HDR camera	✓		SDR BT.709 look	SDR BT.2020 look	PQ look(3)
Graphics	SDR matching colour branding		✓	Slightly under saturated	Similar to artistic intent	Maintaining artistic intent(4)
	SDR matching in-vision signage	✓		SDR BT.709 look	SDR BT.2020 look	PQ look(3)

(1) Scene-light output conversion may be appropriate for an SDR Output that needs to match with secondary production cameras.

(2) Display-light conversion is generally the preferred SDR output method and will preserve the look of graded content and graphics that also originated in SDR, or HLG.

(3) PQ and SDR BT.2020 have a similar look, as discussed in Annex 6 on native looks.

(4) Graded content and graphics content do not necessarily have a native SDR or PQ look. The “Artistic Intent” may have been to make them more saturated, have different contrast, etc.

7.3.4 Signal range considerations for HDR to SDR conversion

When converting signals from HDR to SDR, one approach is to hard clip the HDR signal so that signals below a given threshold (e.g. HDR Reference White) are mapped into the SDR signal range, and signals above the threshold are lost (see § 5). This approach works well when the HDR signal is tightly controlled (for example by using the production workflow described in § 7.2.2.1) to ensure that critically important image detail lies below the clipping threshold. However, to allow the SDR signal to benefit from the HDR production workflow, down-mapping (tone-mapping) is preferred.

With down-mapping, HDR highlights (for example signals above HDR Reference White) are compressed to lie within the upper portion of the SDR signal range. Signals at and below the HDR Reference White level will occupy the remaining SDR signal range. The level at which HDR Reference White is mapped to the SDR signal range is chosen to balance the overall brightness of the SDR image (including graphics) and the amount of detail that is preserved in the image highlights.

The SDR “super-white” code value range (i.e. signals above nominal peak white) is intended to accommodate signal transients and ringing which help to preserve signal fidelity after cascaded processing (e.g. filtering, video compression). In situations where it is known that these signals will not be clipped, they may also be exploited to preserve additional highlights after HDR to SDR down-mapping [5]. However, in other situations (e.g. use of some existing equipment), “super-white” and/or "sub-blacks" could be clipped. In such situations, detail that is critical to the artistic rendition of an image should not be placed in the SDR super-white region after conversion.

7.4 SDR-HDR-SDR "Round-Tripping"

As described in § 7.1, SDR signals will be converted to HDR during production and back again to SDR for distribution. This is the process known as "round-tripping".

Ideally, the process of round-tripping would be transparent. However, in practice, this is difficult to achieve and is the subject of on-going investigation. To understand the difficulties that can arise it is helpful to consider the individual processes of up-mapping to HDR and down-mapping to SDR.

There are two main approaches to including SDR content in HDR programmes: direct mapping and up-mapping (more information in § 5).

Conversion from HDR to SDR is considered in § 5; for example, there is a section on display mapping where the conversion can be regarded as a direct mapping of HDR onto an SDR display. Typically, HDR to SDR conversion uses a non-linearity, similar (and analogous) to the "knee" function found in cameras. This non-linear mapping reduces the dynamic range of highlights but does not completely remove them.

In both up-mapping and down-mapping, careful attention should be paid to those "diffuse" parts of the scene that can be supported in both SDR and HDR formats. However, this is made difficult by variation of the scene luminance factor corresponding to reference white (100%SDR signal) in SDR productions. SDR signals provide little "headroom" for highlights. Some SDR signals are simply clipped of most of the highlight information (e.g. live sport), but in other cases include more highlights through the use of a camera "knee" (e.g. drama or sport "beauty" shots).

The optimum techniques for up-mapping followed by down-mapping are still under investigation.

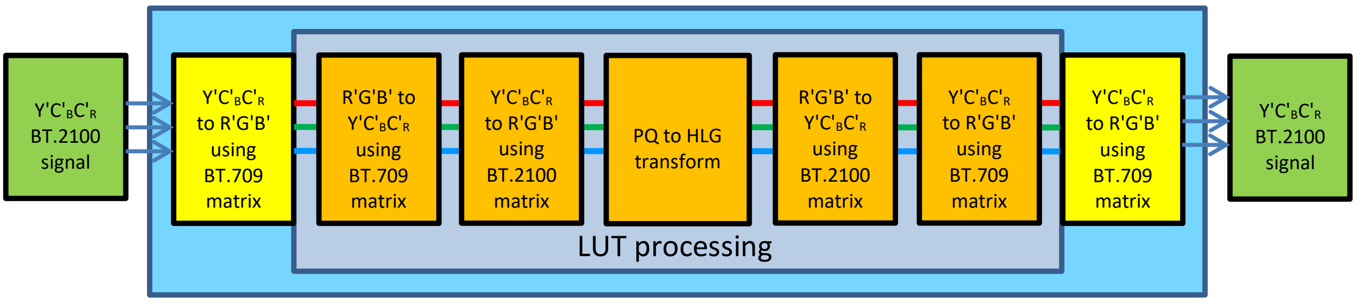
7.5 Hardware colour matrix compensation

Many of the existing hardware devices assume BT.709 colorimetry when converting between $R'G'B'$ and $Y'C'_BC'_R$ signal formats.

Where it is not possible to configure a device for BT.2100 colorimetry, a correction needs to be applied elsewhere. This might be in the conversion matrix on the complementary interface at the other end of the link (e.g. within a display) or, as illustrated in Fig. 19, within a look-up table performing a format conversion.

FIGURE 19

Example of colour matrix compensation within a LUT



7.6 Signal line-up

Prior to any live transmission, it is common practice for broadcasters to check the end-to-end integrity of the production and contribution signal chain. Typically a signal generator producing colour bars and a lipsync test is fed into the production switcher or matrix. The video waveform and lipsync is then checked for accuracy at various points along the chain, including the broadcaster's MCR (Master Control Room).

If BT.2111 Colour Bars are used as the signal source, after any HDR to SDR conversion (e.g. to feed an SDR contribution circuit) the wide colour gamut bars within the test pattern should not be expected to land on the colour bar targets of a standard BT.709 vectorscope; as the SDR BT.709 and HDR BT.2100 colour primaries are different, the true displayed colours of the respective primary (red, green, blue) and secondary (yellow, cyan, magenta) colour bar signals are also different. The BT.709 (scene light) colour bars within the BT.2111 test pattern may also not land on the colour bar targets after scene light down-mapping, as their luminance could be affected by any tone-mapping from HDR to SDR. They should not be expected to land on the colour bar targets after display light down-mapping.

Work is currently underway to design test patterns for signal line-up that should provide a predictable output after display-light and scene-light HDR to SDR conversion.

8 Conversion practices for camera and display RGB colorimetry

Several camera and display systems, for both professional and consumer applications, use their own colour primaries, a practice that may give them certain advantages during capture or display respectively. However, content captured or displayed on such devices would still have to be transformed to or from a BT.2100 workflow, respectively. It should be noted that the transformations in this document only apply under the following conditions:

– The source and target white points are the same and should be equal to D65.

– The source and target white point brightness is the same. For scenarios where brightness is different, refer to Report ITU-R BT.2446.

Furthermore, these transformations are not applicable for camera RAW signals.
Camera and display systems are commonly defined by their normalized primary matrix, NPM, which is specified as follows:

$$\mathrm{NPM} = \begin{bmatrix} \mathrm{X_R} & \mathrm{X_G} & \mathrm{X_B} \\ \mathrm{Y_R} & \mathrm{Y_G} & \mathrm{Y_B} \\ \mathrm{Z_R} & \mathrm{Z_G} & \mathrm{Z_B} \end{bmatrix}, \tag{1}$$

where the elements of the matrix depend on the chromaticity coordinates, (x_R, y_R), (x_G, y_G), (x_B, y_B), and (x_W, y_W) for red, green, blue, and white, respectively, that characterize each system.

The NPM is needed for the conversion process to and from the CIE XYZ colour space and the BT.2100 colour space. The specifics of the computation may be found in Annex 7.

Annex 1

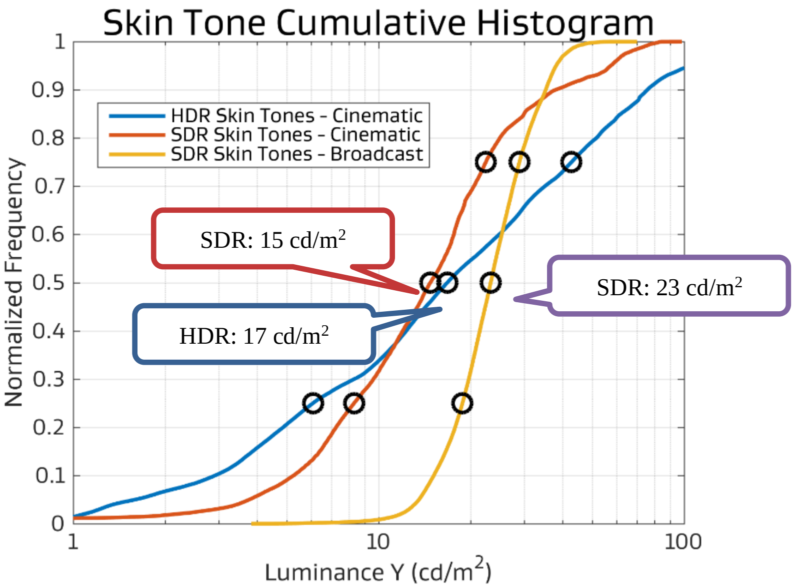
Study to evaluate levels for PQ content

A study was performed to gain information that could be used to inform initial guidance on video levels for HDR production. The study used existing SDR materials from both broadcast content and home video content. The study also used PQ HDR materials, mostly from home video grades of movies that were done on a 4 000 cd/m^2 PQ monitor. From this study, some data on levels is shown. While much of the study employed (for convenience) Caucasian skin levels, existing data on the reflectance of the Caucasian skin was employed to change the reference from skin levels (which of course are not consistent) to use of the conventional 18% grey card.

Details

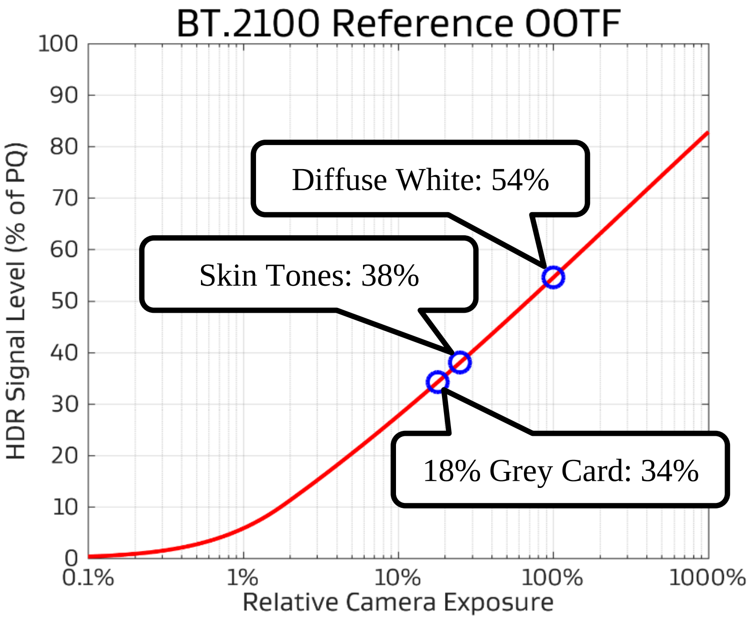
Skin tones from both broadcast content and home cinema release content were analysed. The indoor SDR broadcast content was manually segmented for well-exposed (Caucasian) skin tones and was analysed assuming a BT.1886 reference monitor with 100 cd/m^2 reference white and BT.709 colour primaries. A sampling of the images analysed (courtesy of SVT and FOX) is shown below:

Due to the scarcity of HDR broadcast content currently available, in order to compare HDR and SDR content, the same analysis was completed utilizing HDR and SDR graded indoor scenes from cinematic content for home distribution. The cumulative histogram is given below.



For cinematic content for the home, HDR Caucasian skin tones are very similar to SDR skin tones (17 cd/m^2 compared to 15 cd/m^2), but the standard deviation is larger. Extrapolating from this, it is hypothesized that indoor Caucasian skin tones in HDR broadcast may average 26 cd/m^2 with a larger deviation than SDR broadcast. The 26 cd/m^2 value maps to 38% of full scale in PQ space (or 38%PQ).

Utilizing skin tones as a reference level is, of course, not satisfactory because they vary widely across ethnicities and environments. To achieve consistency, an 18% grey card may be used instead to calibrate camera exposure. To convert from Caucasian skin tone brightness and its 38%PQ level to find the %PQ level of an 18% grey card, a database of 340 measured samples of skin tones (Sun, Fairchild) was used to determine skin tone reflectance levels. This database shows that Caucasian skin tones have a reflectivity of 25% of that of a diffuse white object (white card: 100% Lambertian reflector).



Using the BT.2100 reference PQ OOTF, 26 cd/m^2 may be related to relative scene exposure. Then the 25% and 18% reflectivity relationship may be used to solve for the appropriate 18% grey card level: 17 cd/m^2 on a PQ reference display or 34% on the PQ scale. This is the expected luminance for a grey card anchor in HDR broadcast content for indoor scenes, for content consistent with existing practice. A diffuse white would be expected to yield 54%PQ.

By segmenting HDR indoor and outdoor scenes, it was found that outdoor skin tones were an average of 1.7 stops brighter than indoor skin tones. Assuming a 1.7 stop increase in brightness from an indoor to outdoor scene, the exposure for an 18% grey card outdoors would be set to 45%PQ.

The Table below summarizes Dolby's findings for current content; these values could be considered tentative recommendations on settings of an 18% grey card and diffuse white objects in terms of both %PQ value and reference display brightness.

	Indoor		**Outdoor**	
	cd/m^2	**%PQ**	**cd/m^2**	**%PQ**
18% Grey Card	17	34	57	45
Diffuse White	140	54	425	66

The levels shown in this study are representative of some early HDR PQ content. More experience with HDR in broadcast is needed to settle on final values to be recommended. A major finding is that early HDR production has employed skin levels similar to those used in SDR content. The SDR skin levels are of necessity limited in order to leave room for full diffuse whites, and some trace of highlights. HDR signals have enough range that skin levels do not need such limitations. Given that in HDR production there is no need to limit the skin levels to those used in SDR production, it is possible that these may increase in brightness in subsequent productions. Thus, the values in the Table above might be considered the lower end of future operating levels.

Annex 2

Analysis of reference levels

A2.1 Introduction

The reference levels of Tables 1 and 2 of this Report are intended to provide guidance for the production of HDR content. This Annex presents a Technicolor analysis of existing content relative to several reference levels. The content chosen included frames from an HLG-based live broadcast, as well as a set of test images that were converted to PQ. The purpose of this Annex is to document how the defined reference levels relate to currently produced content, and to assess the variability in luma/luminance levels seen in current content.

A2.2 Analysis of reference levels

Several reference levels are analysed in the context of a database of 107 linear EXR images, graded for a 1 000 cd/m^2 display device. This dataset is included in Report ITU-R BT.2245. In this dataset the arithmetic mean luminance is 65.47 cd/m^2 (standard deviation 83.99 cd/m^2). The geometric mean luminance if 9.17 cd/m^2 (standard deviation 24.74 cd/m^2).

To understand how a given recommended reference level relates to the content presented in this database, the percentage of pixels that have values larger than the reference level is calculated. For each image, this percentage will be different, giving rise to a distribution of percentages. Then, a range of percentages was calculated that represents the 95% confidence interval. This means that this range of percentages represents 95% of the images in the database. To determine a confidence interval, the following equation was used:

$$\mathrm{CI} = \bar{x} \pm z^{*} \frac{\sigma}{\sqrt{n}}$$

where:

$n = 107$: number of images analysed

x : mean number of pixels above the selected reference level

σ : associated standard deviation.

The value of z^{*} is 1.96 for a 95% confidence interval. Likewise, the 99% confidence interval is computed, using a value of z^{*} of 2.58. The results are shown in Table A2.1.

TABLE A2.1

On a 1 000 cd/m^2 image dataset, the 95% and 99% confidence intervals are shown indicating the percentage of pixels that are larger than the reference luminance level

Description	Luminance (range)	95% Confidence interval	99% Confidence interval
Grey card (18%)	26	33.21% - 45.87%	32.21% - 47.88%
Greyscale chart max (83%)	162	8.89% - 16.23%	7.73% - 17.39%
Greyscale chart max (90%)	179	7.82% - 14.77%	6.72% - 15.87%
HDR Reference white	203	6.65% - 13.10%	5.62% - 14.13%
Grass	30-65	19.82% - 43.41%	18.16% - 45.37%
Ice rink	155	9.37% - 16.90%	8.18% - 19.09%
White Objects	140-425	1.79% - 18.59%	1.25% - 19.85%

Further, the same set of images were analysed to understand which luminance level marks the threshold so that 1% of the pixels lies above this level. This calculation was repeated for 5%, 10% and 20% of the pixels. The results are shown in Table A2.2.

TABLE A2.2

Luminance level marking the top *N*% pixels in a set of 107 HDR images

Percentile	Mean (Std) in $\mathrm{cd/m^2}$
1%	321.89 (262.14)
5%	195.04 (206.15)
10%	145.03 (170.56)
20%	(145.01)

A2.3 Diffuse white elements in live HLG encoded broadcast content

Diffuse white elements[7] in HLG encoded live broadcast content (“Dodgers Game”) were analysed by taking one frame every five seconds, and manually clicking in each frame on patches that appeared to represent diffuse white elements which were directly illuminated, without being over-exposed. The total number of analysed frames was 152, and the number of diffuse white points identified in this manner is 378. The content was a baseball game, interspersed with commercials, and containing scenes from a game played in daylight and a game played at night under artificial illumination.

The pixels identified in the manner described above represent values as %HLG. Statistics (mean, standard deviation, minimum and maximum RGB values) are given in the %HLG column of Table A2.3. These numbers were subsequently converted to $\mathrm{cd/m^2}$ assuming a display peak luminance of 1 000 $\mathrm{cd/m^2}$, and to %PQ. These values are also reported in Table A2.3. Finally, Fig. A2-1 shows a histogram of the distribution of diffuse white levels for each of the red, green and blue channels, with the horizontal axis indicating values in %HLG.

7 For the purpose of this Annex, the level of diffuse white elements is referred to as “diffuse white”.

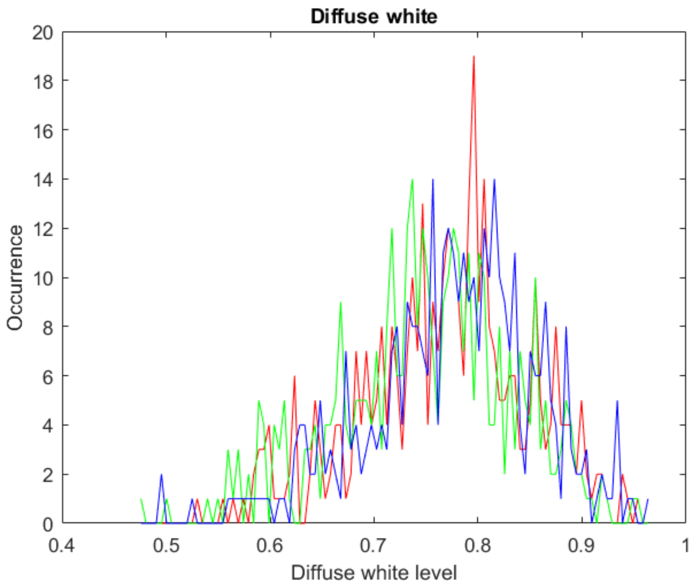
TABLE A2.3

Analysis diffuse white of HLG encoded live broadcast content. The %HLG column was measured, while the remaining columns were derived from these measurements (152 frames, 378 points analysed)

Diffuse White	cd/m^2	%HLG	%PQ
Mean	(222.1, 204.3, 231.3)	(76.6, 75.0, 77.4)	(59.0, 58.1, 59.4)
Std	(134.7 – 373.5 123.6 – 345.4 141.0 – 386.7)	(8.3, 8.4, 8.2)	–
Min	(44.6, 44.5, 48.9)	(47.4, 47.3, 49.6)	(42.9, 42.9, 44.0)
Max	(747.1, 735.3, 789.9)	(95.6, 95.3, 96.6)	(72.0, 71.8, 72.6)

FIGURE A2-1

Distribution of diffuse white patches in HLG live broadcast content. The values on the horizontal axis are in %HLG



A2.4 Diffuse white in an HDR dataset of 1 000 cd/m^2 PQ encoded images

A dataset of 54 EXR images containing diffuse white patches was analysed using the same methodology as described in § A2.3. The dataset contains images that are graded for a 1 000 cd/m^2 display device. The linear EXR images were first PQ encoded. A total of 169 white patches were identified, producing the distribution shown in Fig. A2-2 and the derived statistics shown in Table A2.4. In this Table, the %PQ column was measured from the pixels that were selected, whereas the columns indicated with cd/m^2 and %HLG were calculated from the %PQ column.

TABLE A2.4

Analysis diffuse white of PQ encoded content. The %PQ column was measured, while the remaining columns were derived from these measurements (54 frames, 169 points analysed)

Diffuse White	cd/m^2	%HLG	%PQ
Mean	(231.8, 244.2, 193.3)	(77.1, 78.1, 73.5)	(59.5, 60.0, 57.6)
Std	(76.6 – 665.6 80.6 – 703.4 58.9 – 594.1)	–	(11.3, 11.3, 12.0)
Min	(5.6, 7.0, 6.7)	(19.7 22.0, 21.5)	(25.6, 27.2, 27.0)
Max	(903.0, 1 000.0, 946.5)	(98.2, 100, 99.1)	(74.1, 75.2, 74.6)

FIGURE A2-2

Distribution of Diffuse White patches in a test database of 1 000 cd/m^2 images. The values on the horizontal axis are in %PQ

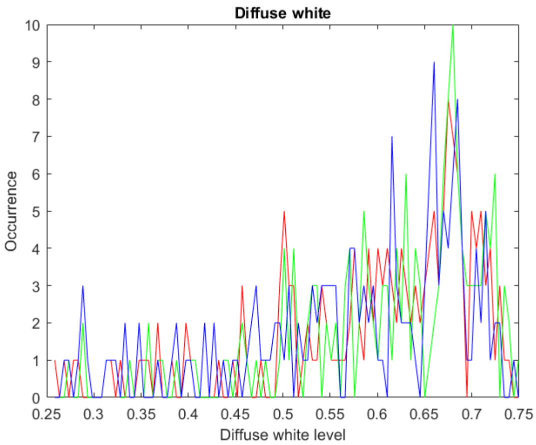


FIGURE A2-3

An image with 1.4% of its pixels above diffuse white (prior to tone mapping for display)

FIGURE A2-4

An image with 17% of its pixels with values above diffuse white (prior to tone mapping for display)

A2.5 Discussion

Two types of analyses were performed to help understand the relationship with pre-defined reference levels and content. In the first analysis, the number of pixels that have values higher than a given reference level was computed. A 95% and a 99% confidence interval was calculated, indicating the percentage of pixels that may be expected to be above the reference level.

For HDR Reference White, for example, it was determined that 99% of the images have between 5.6% and 14% of their pixels result in levels greater than 203 $\mathrm{cd/m^2}$ in a set of 107 HDR images that were graded at 1 000 $\mathrm{cd/m^2}$. Likewise, 95% of the same images have between 6.6% and 13% of their pixels larger than 203 $\mathrm{cd/m^2}$.

To illustrate, compare the images shown in Figs A2-3 and A2-4, which have 1.4% and 17% of their pixels above HDR reference white, respectively. Figure A2-3 shows a clear case of an image where the extra headroom afforded by HDR technologies is spent on the highlights. Figure A2-4, on the other hand, has a significant part of the sky in the background at values above 203 $\mathrm{cd/m^2}$.

In a second analysis, diffuse white was measured by manually identifying pixels in a set of frames/images. Over-exposed pixels were avoided, while diffuse white surfaces not receiving direct illumination were also excluded. The signal levels of white pixels were analysed. For the HLG-based live broadcast content, the mean diffuse white level was 75%HLG, which is the same as the recommended reference level in Table 1 – even if the content was produced without specifically using a target 203 $c\mathrm{d/m^2}$ for reference level. However, the standard deviation was about 8.3% (measured in %HLG), which – for an assumed 1 000 $\mathrm{cd/m^2}$ signal - translates to a range between around 123 and 345 $\mathrm{cd/m^2}$ (i.e. mean $\pm$ one standard deviation). This suggests that the diffuse white level as measured in live broadcast content varies significantly.

These results are broadly replicated with the test set of 107 HDR images which are PQ encoded. Here, the mean diffuse white level was determined to be around 60%PQ, which is close to 58%PQ as recommended in Table 1. The standard deviation was 11% (in %PQ), however, which translates to a range between around 80 and 700 $\mathrm{cd/m^2}$ for mean $\pm$ 1 standard deviation. The variability of diffuse white in this dataset is therefore significant, and it is larger than measured in the HLG-produced live broadcast content.

A2.6 Conclusions

The HDR Reference White level of 203 $\mathrm{cd/m^2}$ in Table 1 of this Report is consistent with the mean diffuse white as measured in the content analysed in this Annex. However, the standard deviation of diffuse white in two different sources of content are large, indicating a significant spread of diffuse white around the mean.

Annex 3

Two studies of skin tones, using a reflectance database and using real subjects

This Annex reports two studies of skin tones, one that uses an existing database of skin reflectances and a model of an ideal camera, and one that uses real subjects and RAW camera recording. Luma values are proposed for different skin tones in HLG high dynamic range video.

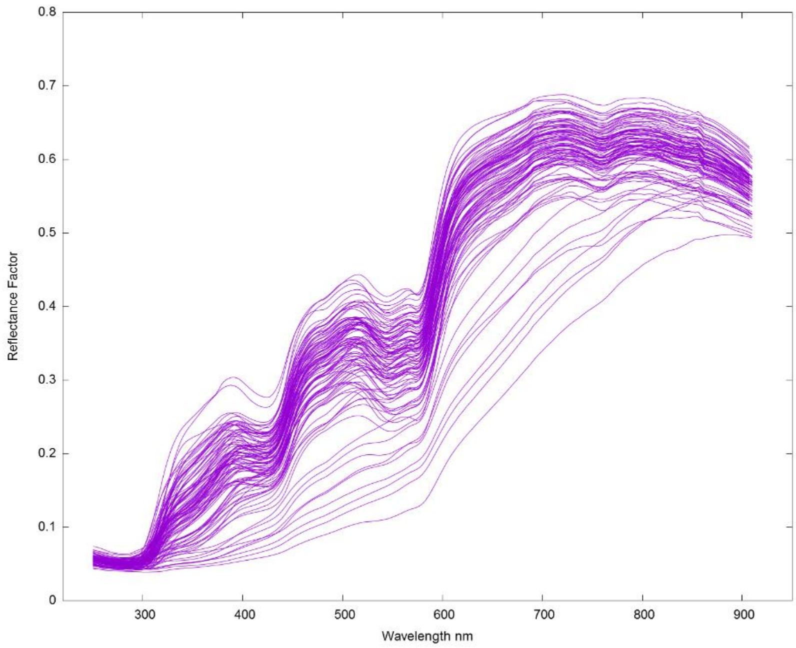
A3.1 Study 1: using a skin tone database and an ideal model of a camera

A skin tone reflectance database from the US government National Institute of Standards and Technology (NIST) [6] was used for this study. The database covers a wide range of skin tones, however when comparing the 685 nm reflectances with those given elsewhere [7], it can be seen that it does not cover the full range of expected global reflectances.

The NIST database contains measures of skin reflectance of the inner forearm at a number of wavelengths. These tend to be slightly higher than the face. This dataset is shown in Fig. A3-1.

FIGURE A3-1

NIST Dataset
Each line corresponds to one skin sample

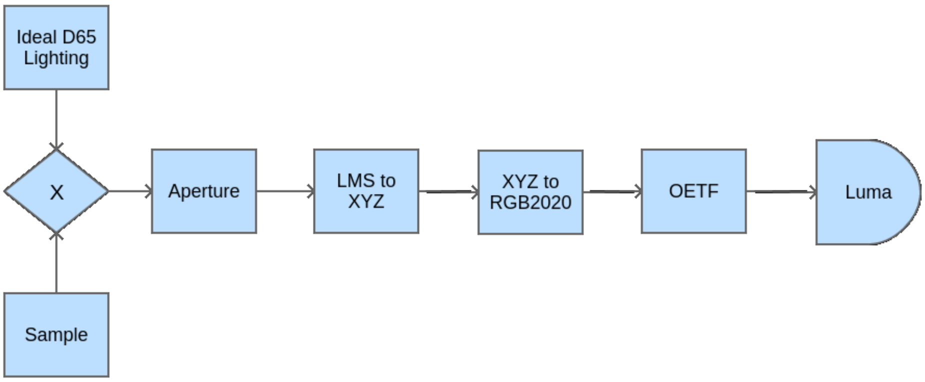


A software model of an ideal camera and lighting scenario was used (illustrated in Fig. A3-2) to generate values for Hybrid Log-Gamma (HLG) luma.

The model consists of a sample multiplied by the spectral curve of an ideal D65 illuminant [8] fed through an aperture (a fixed scalar). A set of CIE 1931 2 degree observer LMS to XYZ curves [9] are then used to convert to a known imaging format. These XYZ values are then converted to Recommendation ITU-R BT.2020/BT.2100 linear RGB values and the HLG Opto-Electronic Transfer Function (OETF) is applied. Finally, the luma value is calculated for the HLG $R'G'B'$ values.

FIGURE A3-2

Block diagram of ideal camera model



The NIST data set, ideal D65 illuminant curves and LMS to XYZ curves all used different wavelength step sizes in presenting the data, so, where data points did not align, a linear interpolation was used.

The first step in using the model was to calculate the required input aperture. By setting the input sample to a fixed value of 1.0 at all wavelengths to represent diffuse white, the aperture (a scalar) was adjusted such that the HLG luma value was equal to 0.75, the HLG signal level for HDR Reference White. This value of aperture was then used for all further samples.

The second step is to apply the model for each skin reflectance curve given in the NIST dataset. The results of this are shown in Fig. A3-3. Luma values are plotted against the skin reflectance at 685 nm to allow comparison with regional labelling from [7]. These regional labels have been added to the plot.

A further plot of skin tone reflectance against screen emittance for a 1 000 cd/m^2 HLG display is given in Fig. A3-4.

FIGURE A3-3

Skin tone reflectance at 685 nm against HLG Luma for ideal camera, with regional labels from [7]

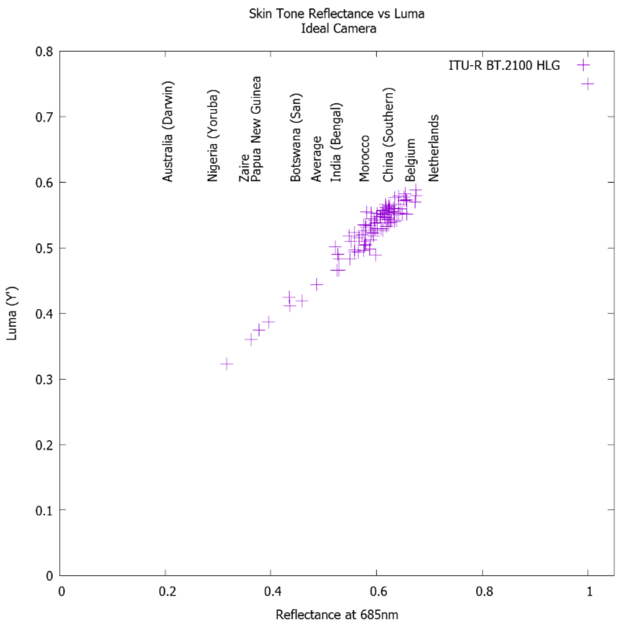
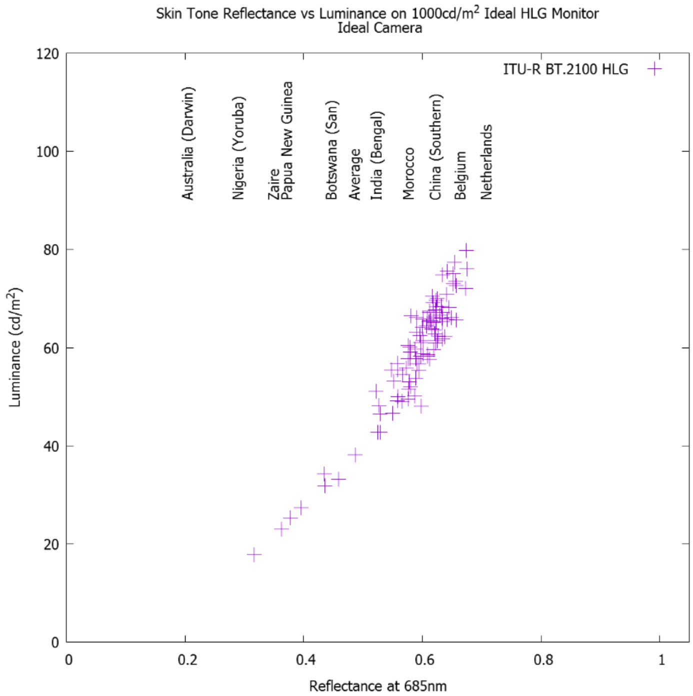


FIGURE A3-4

Skin tone reflectance at 685 nm against HLG luminance on a 1 000 cd/m^2 display, for ideal camera, with regional labels from [7]



A3.2 Study 2: using human subjects and a RAW recording camera

In conjunction with the European Broadcasting Union, a second experiment was conducted using real people and a DSLR RAW-recording camera. To categorise the subjects, the Fitzpatrick Skin Tone Scale [1] was used.

The first stage of the experiment was to calculate the reflectance of a small test chart that could be used in shot when photographing test subjects under practical D65 LED lighting. Using a Konika-Minolta CS2000 photospectrometer, the reflectances of the test chart white and black patches, a magnesium carbonate reference (97.5% reflectance) and a Gregory hole reference (black velvet lined box – 0% reflectance) were measured. The test chart white patch reflected 81.2% of light, the black patch 3.9%.

The processing chain for the images was designed to closely replicate the ideal camera workflow shown in Fig. A3-2. This is shown in Fig. A3-5. To convert the camera RAW file to linear XYZ, the open source package DCRaw [10] was used. This file was then processed to:

1 Convert the XYZ values to ITU-R BT.2020 linear RGB values and then to CIE Yu′v′;

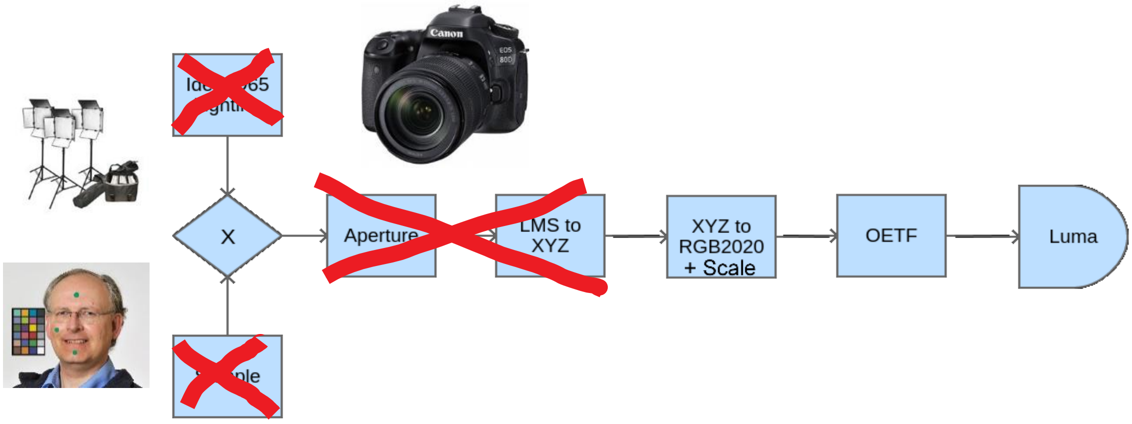
2 Scale Y such that the average black patch pixel value equalled 3.9% and the average white patch pixel value equalled 81.2%, then convert back to ITU-R BT.2020 linear RGB values;

3 Apply the HLG OETF to the R, G and B channels (using the equations found in Recommendation ITU-R BT.2100) and then calculate the Y′ channel;

4 Crop two 50 pixel by 50 pixel areas of skin tone (forehead and cheek) and calculate the average luma value. Care was taken to ensure that the chosen areas are co-planar with the physical luminance ramp test chart.

FIGURE A3-5

Real-life human skin tone measurement



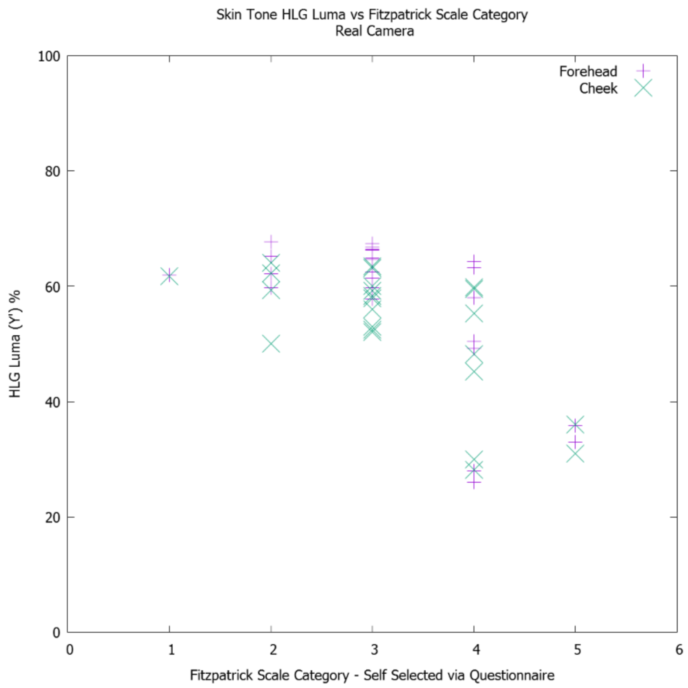
In order to match the test subject to the Fitzpatrick Scale classifications, a questionnaire from the Australian Government Radiation Protection and Nuclear Safety Agency was used [11].

The results of these photographic tests are shown in Fig. A3-6. Skin tone measurements range from approximately 26%HLG to 67%HLG dependant on skin tone. It can also be seen that there is an issue with two peoples' replies to the questionnaire. Both individuals are deeply pigmented and should either be type V or VI but have self-identified as type IV. Following discussions with these individuals, and others identifying as type IV, V or VI, it appears that there is an issue with the questions relating to tanning: people either reported that they were permanently tanned or that they never tanned, which led to changes in the result. Finally, it can be seen that there is a small difference across the face, with the forehead being more reflective than the cheek for persons with skin types II to IV.

It should be noted that the event at which measurements were taken occurred in the Northern Hemisphere during winter (so few people were currently tanned) and the attendee demographic was skewed towards categories II, III and IV.

FIGURE A3-6

HLG signal levels measured from human subjects



Based on these experimental results, Table A3.1 shows suggested HLG luma ranges for each skin type. In formulating the values, the two people discussed previously in this section have been re-categorised as category VI, which gives values consistent with those presented in Fig. A3-3. To accurately represent the majority of the exposed skin which does not exhibit issues with perspiration shine, the ranges are chosen to cover the majority of the cheek skin tone measurements for each category, ignoring obvious outliers. A small amount of leeway is allowed at the bottom end of the ranges for categories I-IV to allow for summer tanning. Camera zebras should be set 2 to 3% above these ranges to take account of perspiration shine. Values are chosen to be easily used by productions using waveform monitors only.

TABLE A3.1

Suggested HLG signal ranges for different skin types

Fitzpatrick skin type	HLG signal level (%HLG)
I and II	55-65
III and IV	45-60
V and VI	25-45

A3.3 Conclusions

1 HLG luma levels measured with the DSLR camera (Study 2) are similar to those calculated with the computer camera model (Study 1).

2 Results are valid when the sample is from areas of skin co-planar with the physical test chart. Due to using a single light source, there is a marked drop off in reflectance when moving away from areas of the face that are co-planar. In one instance, the side of the face reflects less light than the black test colour on the chart.

3 There is an issue with "forehead shine" caused by both perspiration under the studio lights and a matching of the angle of incidence and reflection such that light is reflected directly towards the camera.

4 The Australian Government questionnaire is designed to suggest levels of skin protection required in the southern hemisphere tropics and, therefore, is most suited to Fitzpatrick Skin Types I-IV. There is a possible mis-categorisation of two test subjects.

Annex 4

Study of facial skin tones in broadcast content

This Annex reports on studies of facial skin tones in broadcast content in Japan.

A4.1 Facial skin tones in SDR news and information programmes in studio

Eight Japanese broadcasters contributed SDR broadcast content produced under controlled lighting in studios to this study. Table A4.1 shows the overview of the content. Target areas within a face, i.e. forehead and cheeks, were clipped out from the images and their average signal levels were measured.

TABLE A4.1

SDR broadcast content produced in studio used for study

Content holders	8 Japanese broadcasters(1)
Programme genre	News, information, and talk shows produced in studio
Types of framing	Single shot, two-shot, and group shot
People in scenes	Male and female Japanese/Mongoloid
Target area for analysis	Forehead and cheeks(2)
Number of sample images	387 in total
Number of faces analysed	Male: 365, female: 348, and total: 713

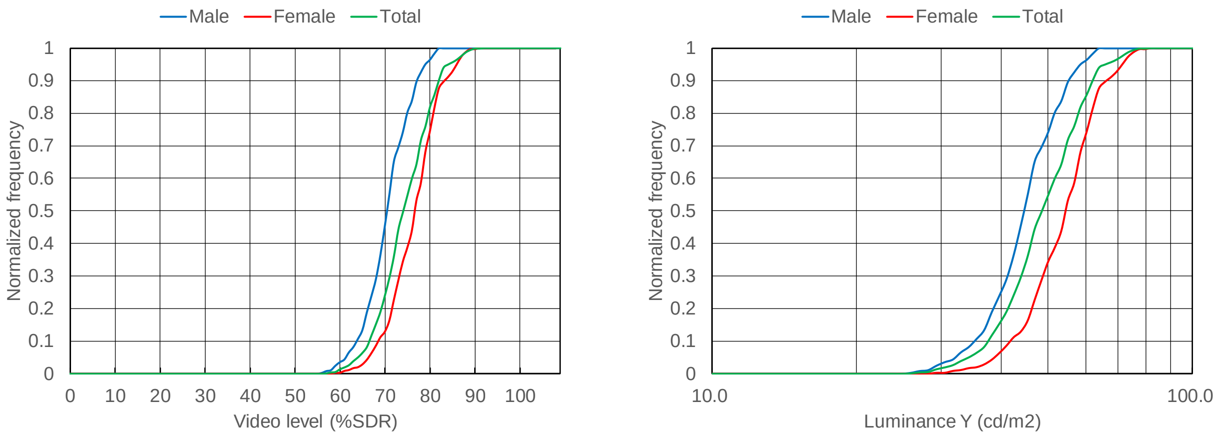
(1) Japan Broadcasting Corp., Asahi Broadcasting Corp, Nippon Television Network Corp., Tokyo Broadcasting System Television, Fuji Television Network, TV Asahi Corp., TV Tokyo Corp., and WOWOW.

(2) Areas that exhibit the highest signal level within a face except for specular reflection and shine. A single person was charged with analysing the skin tones for consistent analysis.

Figure A4-1 shows the cumulative histogram of the facial skin tones. The average video levels (Y′) and standard deviations for male, female, and total are 71.8 (σ=5.2), 77.6 (σ=5.7), and 74.6 (σ=6.2) %SDR, respectively. These video levels correspond to luminance of 45 cd/m^2, 55 cd/m^2, and 49 cd/m^2 on a display with the peak luminance of 100 cd/m^2. The luminance of facial skin tones is more than twice the 23 cd/m^2 reported in Annex 1 for SDR broadcast content.

FIGURE A4-1

Cumulative histogram of facial skin tones



- Video level (Y′)
- Display luminance

Facial skin reflectance was estimated in one of the SDR programmes, in which video level of facial skin was 81%SDR (Y′), by placing the 11-step grey scale chart at the caster's position under the same lighting and exposure conditions in the studio. From the measurement, the reflectance of the facial skin was estimated to be 31% for luminance.

A4.2 Comparison of facial skin tones in HLG HDR and SDR content in a music programme

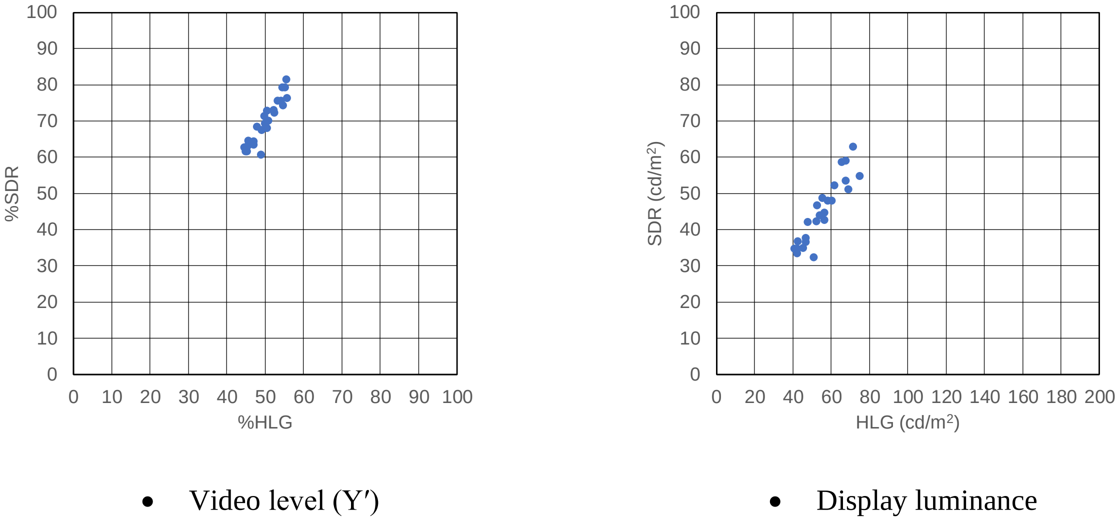
A preliminary study was conducted on skin tones in HLG HDR content in comparison with SDR content. Both HLG and SDR content were produced independently for the same NHK music programme in a concert where musicians performed on stage under special lighting and set. In the HDR production, video engineers paid attention to the reference level of 75%HLG. 75%HLG was also used for graphics white in captions.

The facial skin tones of 11 people (musicians and hosts) from 24 scenes in each of the HLG and SDR programmes were analysed. Figure A4-2 plots average levels of each face. The facial skin tones were found to be 45-56%HLG (50%HLG on average) and 41-71 cd/m^2 on HLG displays with a peak luminance of 1 000 cd/m^2, and 61-82%SDR (70%SDR on average) and 32-63 cd/m^2 on SDR displays with a peak luminance of 100 cd/m^2.

The values for SDR correspond well to those for the SDR news and information programmes in studio described in § A4.1. Since the HLG signal level for 30% reflectance is 49%HLG when 75%HLG corresponds to 100% reflectance, the skin tones in the HLG programme well match the HDR reference level.

FIGURE A4-2

Comparison of facial skin tones in HLG and SDR content in music programme



A4.3 Conclusion

Facial skin tones in SDR content in news and information programmes in studios and those in HDR and SDR content in a music programme in a concert hall were studied. The results are summarized in Table A4.2. The facial skin tones in Japanese SDR programmes were found to be much higher than those reported for European and American programmes in Annex 1. This may be mainly due to a difference in long-standing production practice for SDR rather than a difference in skin reflectance. It should also be noted that makeup also affects skin tones significantly.

The relationship in facial skin tones between HDR and SDR should provide a foundation for establishing guidelines for converting HDR content into SDR and vice versa. Although HDR production is anticipated to universally follow the HDR reference levels described in this Report, different conversion characteristics may be needed for the conversion from HDR to SDR to obtain SDR pictures with familiar facial look in different regions or countries, yet more research is desirable.

TABLE A4.2

Summary of facial skin tones in Japanese content

<table>
<tr><th colspan="2">Programme genre</th><th colspan="3">News and information in studio</th><th colspan="2">Music programme in concert hall</th></tr>
<tr><td colspan="2">Format</td><td colspan="3">SDR</td><td>SDR</td><td>HLG</td></tr>
<tr><td colspan="2">Graphics white</td><td colspan="3">100%SDR</td><td>100%SDR</td><td>75%HLG</td></tr>
<tr><td rowspan="4">Average skin tones</td><td rowspan="2">Signal level</td><td>Male</td><td>Female</td><td>Total</td><td rowspan="2">70%SDR</td><td rowspan="2">50%HLG</td></tr>
<tr><td>72%SDR</td><td>78%SDR</td><td>75%SDR</td></tr>
<tr><td rowspan="2">Display luminance</td><td>45 cd/m^2</td><td>55 cd/m^2</td><td>49 cd/m^2</td><td>45 cd/m^2</td><td>55 cd/m^2</td></tr>
<tr><td colspan="4">on a display of 100 cd/m^2 peak</td><td>on a display of 1 000 cd/m^2 peak</td></tr>
</table>

Annex 5

Displaying PQ - calculating the EETF

This Annex describes approaches to mapping HDR signals to displays with a lower dynamic range, i.e. how to calculate the necessary EETF (electrical-electrical transfer function) in order to adapt to the display, see § 3.1.1. Such mapping may also be required during conversion from PQ to HLG.

The central region of the tone mapping curve is defined as a 1:1 mapping. A 'knee' roll off may be calculated using a hermite spline to create a mapping that will reduce the luminance range to the capability of the display. The black level lift is controlled by an offset, b, which may be determined by a PLUGE adjustment as specified in Recommendation ITU-R BT.814. The difference between this proposal and the black level adjustment per Recommendation ITU-R BT.1886 is the addition of a tapering factor $(1 - E_2)^4$. Without such a tapering factor, a constant offset throughout the entire signal range has the effect of increasing the brightness at the high end. With Recommendation ITU-R BT.1886 this effect was limited and not problematic due to the large number of code values at the high end of the gamma curve. The perceptual uniformity of the PQ EOTF causes this effect to be unacceptable. The tapering function allows fine-tuning the lift without a significant impact on mid-tones or highlights.

In the case where the mastering display minimum black and peak white luminances are known or reasonably can be assumed, the first step in applying the EETF is to normalize the PQ values based on the mastering display black and white luminances, L_B and L_W:

Step 1: $E_1 = (E' - EOTF_{PQ}^{-1}[L_B])/(EOTF_{PQ}^{-1}[L_W] - EOTF_{PQ}^{-1}[L_B])$

where E' is the I, Y′ or R′, G′, or B′ PQ component and E_1 is the corresponding mastering display black and white normalized PQ component.

In the case where the mastering display minimum black and peak white luminances are not known and reasonably cannot be assumed, a value of 0 can be used for L_B and a value of 10 000 can be used for L_W, corresponding to the entire PQ encoding luminance range.

The next step is to calculate the mastering display black and white normalized PQ values, *minLum* and *maxLum*, corresponding to the target display minimum (L_{min}) and maximum (L_{max}) luminances, including ambient, as follows:

$$minLum = (EOTF_{PQ}^{-1}[L_{min}] - EOTF_{PQ}^{-1}[L_B])/(EOTF_{PQ}^{-1}[L_W] - EOTF_{PQ}^{-1}[L_B])$$

$$maxLum = (EOTF_{PQ}^{-1}[L_{max}] - EOTF_{PQ}^{-1}[L_B])/(EOTF_{PQ}^{-1}[L_W] - EOTF_{PQ}^{-1}[L_B])$$

The next step is to calculate the 1:1 mapping and knee (E_2). The turning point (KneeStart or KS) for the spline is the point where the roll off will begin, as follows:

Step 2: $KS = 1.5\ maxLum - 0.5$

$b = minLum$

The next step is to solve for the EETF (E_3) with given end points.

Step 3: $E_2 = E_1$ for $E_1 < KS$

$E_2 = P[E_1]$ for $KS \le E_1 \le 1$

$E_3 = E_2 + b(1 - E_2)^4$ for $0 \le E_2 \le 1$

Hermite spline equations:

Step 4: $P[B] = (2T[B]^3 - 3T[B]^2 + 1)KS + (T[B]^3 - 2T[B]^2 + T[B])(1 - KS) + (-2T[B]^3 + 3T[B]^2)maxLum$

T[A] = (A – KS)/(1 – KS)

The final step is to invert the normalization of the PQ values based on the mastering display black and white luminances, L_B and L_W, to obtain the target display PQ values.

Step 5: $E_4 = E_3\,(EOTF_{PQ}^{-1}[L_W] - EOTF_{PQ}^{-1}[L_B]\,) + EOTF_{PQ}^{-1}[L_B]$

The EETF may be applied in many colour representations [12]. Here are some options:

1) IC_TC_P

$$I_2 = EETF(I_1)$$

$$C_{T2}, C_{P2} = min\left(\frac{I_1}{I_2}, \frac{I_2}{I_1}\right) \times (C_{T1}, C_{P1})$$

2) $Y'C'_BC'_R$

$$Y'_2 = EETF(Y'_1)$$

$$C'_{B2}, C'_{R2} = min\left(\frac{Y'_1}{Y'_2}, \frac{Y'_2}{Y'_1}\right) \times (C'_{B1}, C'_{R1})$$

3) *YRGB*

$$Y_1 = 0.2627R_1 + 0.6780G_1 + 0.0593B_1$$

$$Y_2 = EOTF_{PQ}(EETF(EOTF_{PQ}^{-1}(Y_1)))$$

$$(R_2, G_2, B_2) = \frac{Y_2}{Y_1} \times (R_1, G_1, B_1)$$

4) $R'G'B'$

$$(R'_2, G'_2, B'_2) = EETF(R'_1, G'_1, B'_1)$$

5) *maxRGB*

$$M_1 = max(R_1, G_1, B_1)$$

$$M_2 = EOTF_{PQ}\left(EETF\left(EOTF_{PQ}^{-1}(M_1)\right)\right)$$

$$(R_2, G_2, B_2) = (M_2/M_1) \times (R_1, G_1, B_1)$$

As summarized in Table A5.1, since the IC_TC_P $Y'C'_BC'_R$, and *YRGB* methods can produce colours significantly outside the destination gamut, the degree to which these methods preserve the creative intent can be dependent on the gamut mapping used.

TABLE A5.1

Advantages of EETF application space

	IC_TC_P	**$Y'C'_BC'_R$**	**YRGB**	**R'G'B'**	**maxRGB**
Does not produce colours significantly outside the Target Colour Volume	×	×	×	✓	✓

It is possible to blend results from multiple methods. For example, with highlight compression, desaturation and hue changes can be controlled to some extent without requiring gamut mapping by using a blend of *R'G'B'* with *maxRGB*.

The degree to which the creative intent is maintained also depends on the amount of tone and colour compression applied, with more compression producing more significant differences. See Report ITU-R BT.2446 for methods for converting between HDR and SDR.

The following is a short list expanding on the characteristics of each mapping space:

IC_TC_P

1 Has the potential to produce colours outside the destination gamut, which then require gamut mapping.

2 Since it is a perceptual colour difference space, it is a good space for gamut mapping. No need to convert to a different colour space if gamut boundary information and appropriately configurable gamut mapping algorithms are available.

3 Includes a desaturation function to produce a "natural" looking desaturation where the source image lightness is changed by the EETF. Natural refers to the desaturation that results from the roll-offs in the human visual system response with colours that are extremely darker or extremely lighter than the adapted luminance. A common example is walking out of a dark theatre into the sun - initially the outdoor colours will look very bright and "washed out."

4 Preserves hue in IC_TC_P space, which should be close to preserving perceptual hue due to the design of the IC_TC_P space.

$Y'C'_BC'_R$

1 Has the potential to produce colours outside the destination gamut, which then require gamut mapping.

2 The colour space can be used for gamut mapping. No need to convert to a different colour space if gamut boundary information and appropriately configurable gamut mapping algorithms are available.

3 Includes a desaturation function to produce a "natural" looking desaturation where the source image lightness is changed by the EETF.

4 Preserves hue in the nonlinear $Y'C'_BC'_R$ space, which departs in some areas from perceptual hue but can still produce acceptable results.

YRGB

1 Has the potential to produce colours outside the destination gamut, which then require gamut mapping.

2 Preserves chromaticity except for where gamut mapping is applied. Does not produce a "natural" looking desaturation of tonally compressed colours.

3 Problems can be avoided by using in combination with a variable desaturation and gamut mapping algorithm, although such algorithms generally perform best in hue, saturation and lightness colour spaces (requiring a colour space change).

R'G'B'

1 Does not produce colours outside the destination gamut.

2 Generally tends to produce "natural" looking colours, although saturation of extreme colours is reduced substantially and some colours may be changed in hue.

3 Problematic when it is desired to retain bright saturated colours, such as coloured lights at night.

4 Depending on the amount of compression, the saturation decrease may be excessive, and occasionally hue changes can be objectionable.

5 Results depend on RGB primaries used. It has been found that primaries close to the BT.2020 primaries tend to work well.

maxRGB

1 Does not produce colours outside the destination gamut when used to compress highlights.

2 Preserves chromaticity at the expense of lightness.

3 Can produce un-natural colours that look like artifacts, due to lightness differences being compressed without associated saturation changes. Typical examples include very bright skin tones and sunsets where luminance differences are obscured. In these cases, other methods will produce better looking results.

4 Does a good job maintaining bright coloured lights.

5 Problems can be reduced by avoiding problematic colours in the original, for example exposing so skin tones are never near the top of the range, and mastering using IC_TC_P or $R'G'B'$ so there is some built-in saturation difference as a function of lightness for very bright colours.

Annex 6

Comparison of the native looks of HDR and SDR production

As mentioned in § 7.3.3, when colour matching cameras in live production, it is important to note that the native displayed "look" of each SDR and HDR production format is different as, by design, they all have different OOTFs. Even though cameras usually provide "painting controls' that adjust the OOTFs to deliver the desired artistic "look", they are often insufficient to exactly match the displayed "look" of cameras using the different formats. For that reason, when converting signals from different format cameras into a common format for live production, scene-light rather than display-light conversions are preferred, as they are agnostic to the OOTF differences.

To quantify the differences in the displayed look of the different formats, the different television production formats BT.709, BT.2020, PQ, HLG, and HLG with traditional colour reproduction (as described in § 6.5 of Report ITU-R BT.2390) are compared based on their different renderings (different displayed light) of the same scene data, i.e. their different "native looks", as determined by their different OOTFs. For each format, the display light is obtained by passing selected reference colour pattern data (a "colour chart") through its OOTF, where the OOTF of each format is determined by the concatenation of its respective OETF (camera side) and EOTF (display side). The reference colour data is the television colour reference pattern of [13], which describes a colour chart containing three lines of six coloured swatches as well as one line of six neutral swatches (with different reflectances). To enable an objective comparison between the different formats, the luminance of the displayed white swatch (the neutral swatch with the highest reflectance) is normalized to approximately 200 cd/m^2 for each format. This normalization is performed by linearly scaling the scene linear reference colour data (equivalent to adjusting the camera iris).

The differences between the display light colour charts for the different formats can be characterized by the differences in chromaticity as well as luminance between the displayed colour swatches.

The CIE 1976 uniform chromaticity scale plot [14] of the display light chromaticity values for each format shows that the chromaticity differences between the different formats can be substantially characterized as saturation differences (the differences in hue between the formats are small).

The HLG format, by design, has the lowest saturation of all formats because it preserves the chromaticity of the scene as imaged by the camera; all other formats increase saturation compared to the scene as imaged by the camera. Earlier studies have shown that colorimetrically accurate reproduction of natural scenes does not necessarily ensure the highest perceived image quality and a reliable enhancement of perceived image quality can be produced by selectively increasing saturation values [15]. Most cameras offer a saturation adjustment in the CCU to deliver artistically pleasing images.

When ranking the formats from low to high saturation for the non-neutral colour swatches in the chart, the ranking is as follows for average and median colour swatch saturation:

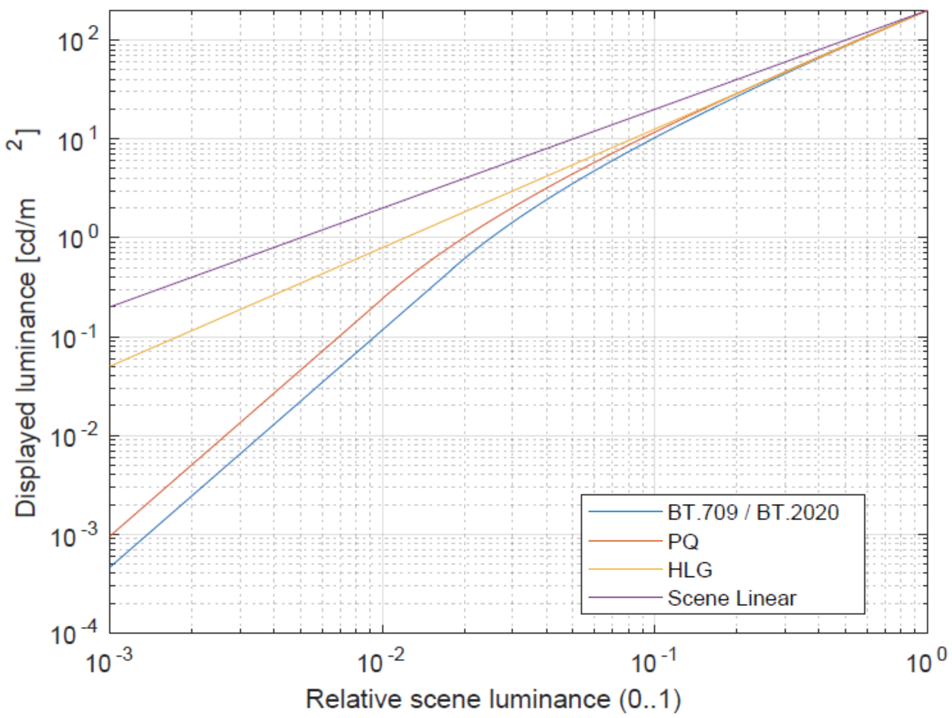
Average: HLG < HLG traditional colour < PQ < BT.709 < BT.2020

Median: HLG < BT.709 < HLG traditional colour < PQ < BT.2020

In addition to the differences in saturation, there are also differences in luminance between the formats. These differences are more pronounced for relatively (or absolutely) lower luminances and can be explained by the differences in the respective OOTFs, as shown in Fig. A6-1 (where the scene luminance has been normalized on the white swatch luminance). The HLG format OOTF has a gamma 1.2 across the luminance range and, for the lower luminances, is closest to the scene light (which has a linear OOTF or a gamma of 1). The BT.709, BT.2020, and PQ formats all have a gamma 2.4 near black. It can be observed that the HDR formats (particularly HLG) preserve a higher luminance near black than the SDR formats, so the HDR formats show/preserve more detail in the dark. The SDR formats, on the other hand, produce images with a higher perceived contrast.

FIGURE A6-1

OOTFs for the SDR and HDR formats compared to scene light



The following sub-sections provide details of the analysis of the differences in chromaticity/saturation and luminance of the different SDR and HDR formats.

A6.1 Differences in chromaticity and saturation

Colours may be characterized by their chromaticity, which is the property of colour that is independent of luminance. To visualize chromaticity and chromaticity differences, the CIE 1976 uniform chromaticity scale diagram with chromaticity coordinates *u*' and *v*' may be used [14]. Each point in the diagram can be described either directly by its coordinates, or indirectly by its hue (or hue angle) and saturation.

The hue (angle) is defined as $h_{uv} = \arctan[(v' - v'_n)/(u' - u'_n)]$ and the saturation as $s_{uv} = 13\,[(u' - u'_n)^2 + (v' - v'_n)^2]^{1/2}$ where u'_n and v'_n are the coordinates of the white point, which for television is CIE D65 with coordinates (0.1978, 0.4683). Thus, the saturation corresponds to the Euclidian distance from the white point.

The uniform chromaticity diagram in Fig. A6-2 shows the display light colour swatch chromaticity for the different production formats. It can be observed that the colour swatch chromaticity for the different production formats falls approximately on a line of constant hue. Therefore, the chromaticity differences can substantially be characterized as saturation differences.

It can be observed that BT.2020 generally provides the highest saturation while, by design, HLG provides the lowest saturation. This can also be observed from the saturation values shown in Fig. A6-3. The saturation differences with BT.709 are shown in Fig. A6-4 and the saturation differences with BT.2020 are shown in Fig. A6-5.

FIGURE A6-2

Display light colour swatch chromaticity for the different production formats

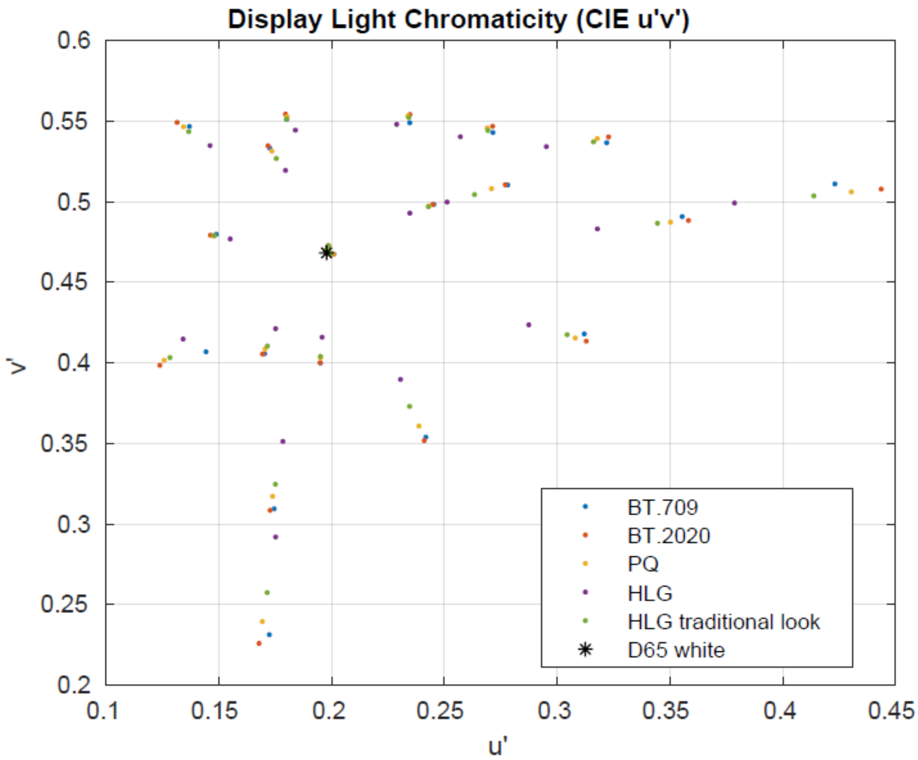


FIGURE A6-3

Display light colour swatch saturation

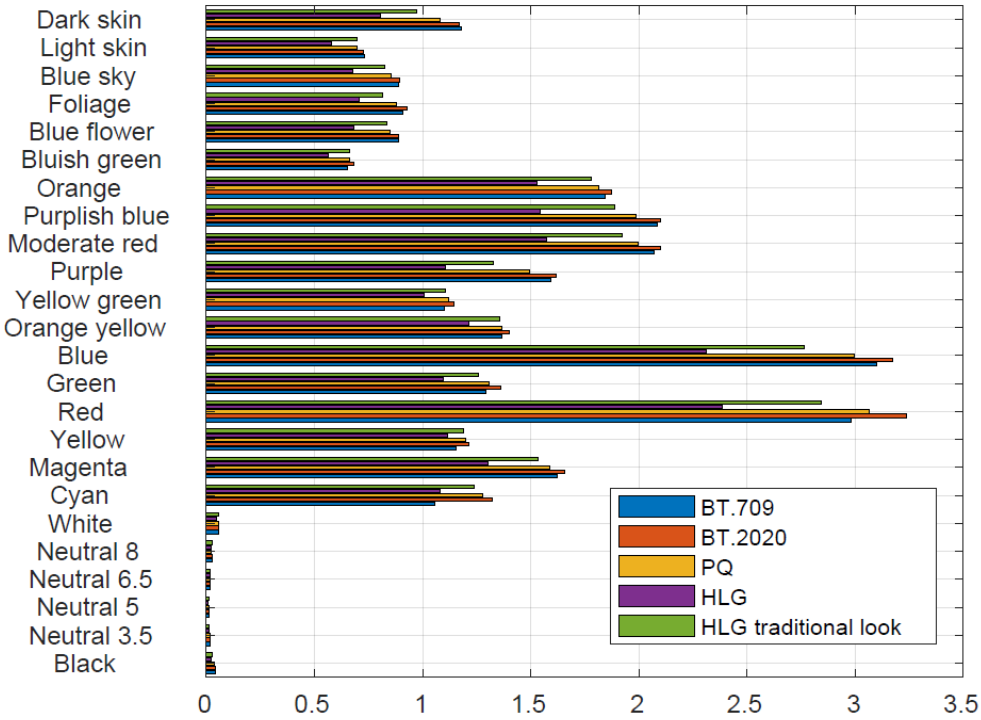


FIGURE A6-4

Display light colour swatch saturation differences with BT.709

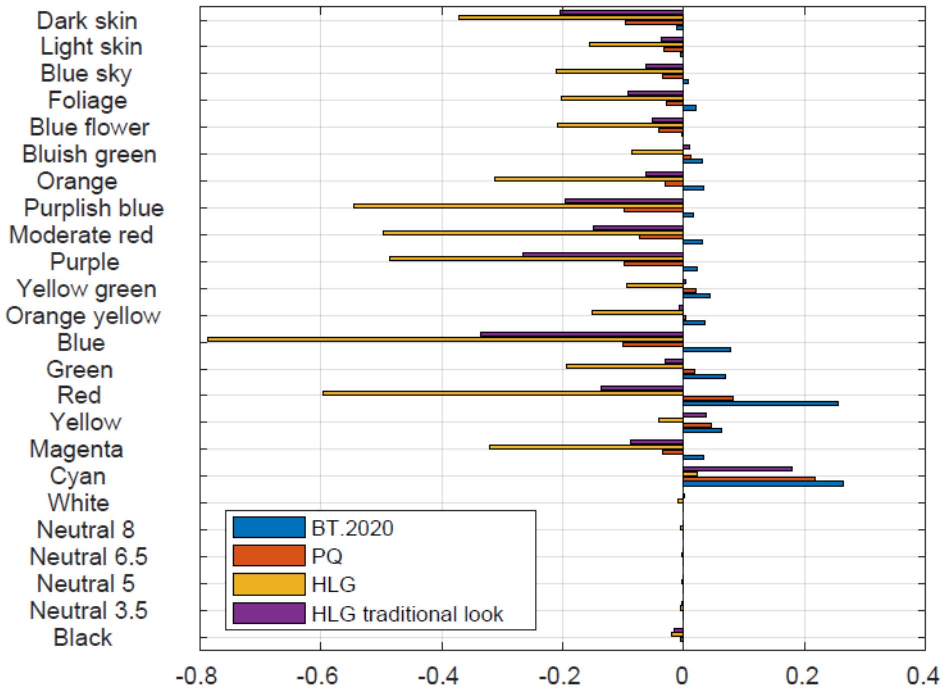
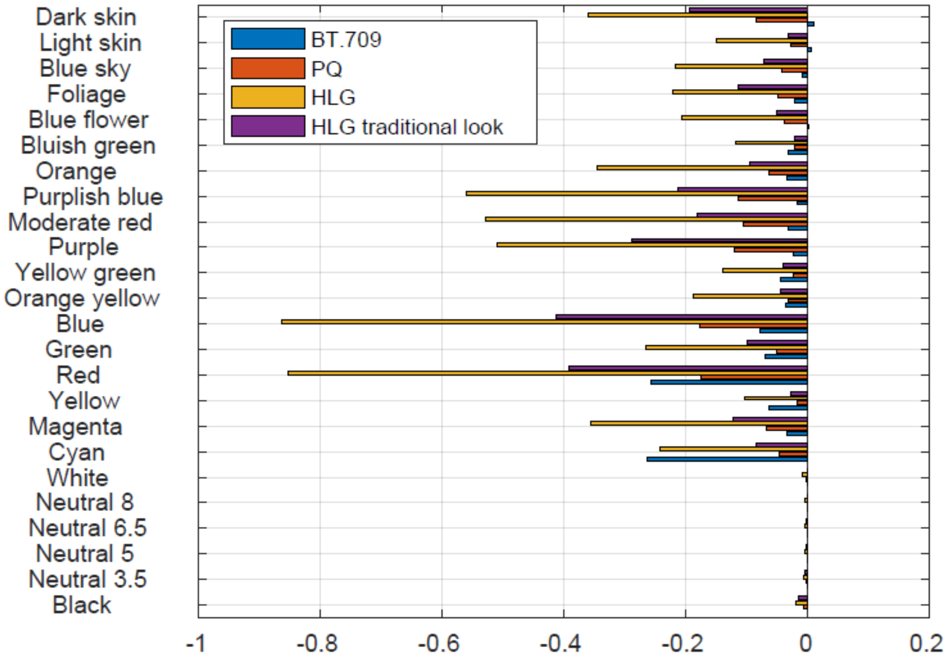


FIGURE A6-5

Display light colour swatch saturation differences with BT.2020



A6.2 Quantifying the total colour differences

While the differences in chromaticity and saturation of the colour swatches were shown in § A6.1, those differences do not take into account the luminance differences between the swatches and therefore do not represent the total colour differences. To quantify the total differences, a metric should be applied that takes into account the chromaticity differences as well as the luminance differences, such as e.g. the Delta E_{2000} metric defined by the CIE [16], or the new Delta E_{ITP} metric defined in Recommendation ITU-R BT.2124. The latter metric is applied in the following, using either the BT.709 or BT.2020 display light colour swatches as a reference.

The Delta E_{ITP} differences with BT.709 are shown in Fig. A6-6 and those with BT.2020 are shown in Fig. A6-7. It can be observed, e.g. that the differences for the highly saturated colours (such as Red and Blue) are larger than the differences with BT.709.

Note also the differences for the White/Neutral/Black colour swatches, which are luminance differences caused by the differences between the SDR and HDR OOTFs (as shown in Fig. A6-1). The relatively darker/lower scene luminances are displayed brighter in the HDR formats than in the SDR formats (so the HDR formats show/preserve additional detail in the dark).

FIGURE A6-6

Display light colour swatch Delta E_{ITP} with BT.709

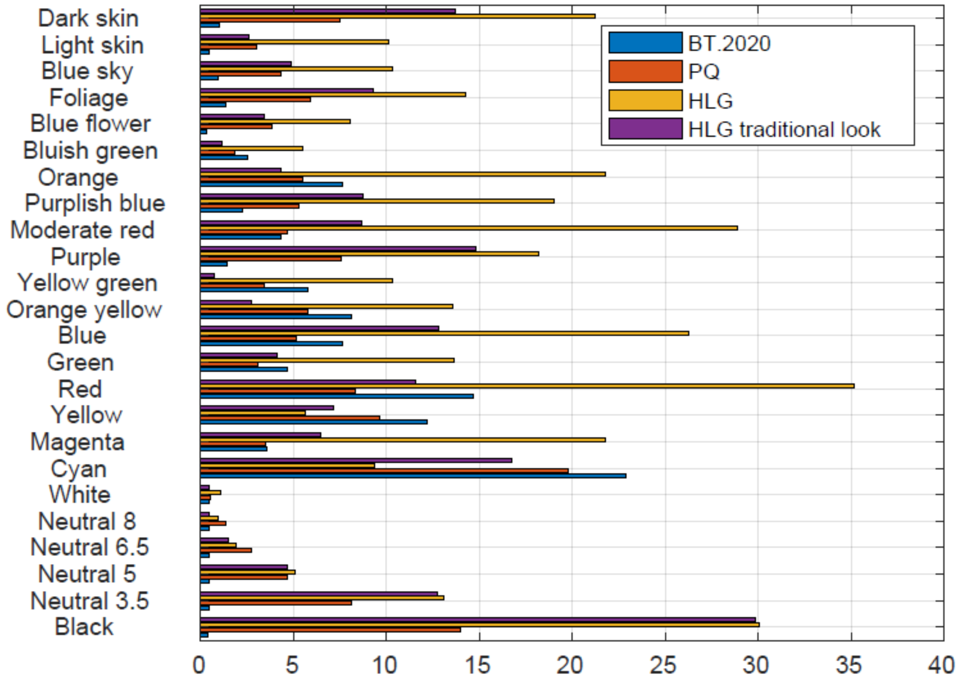
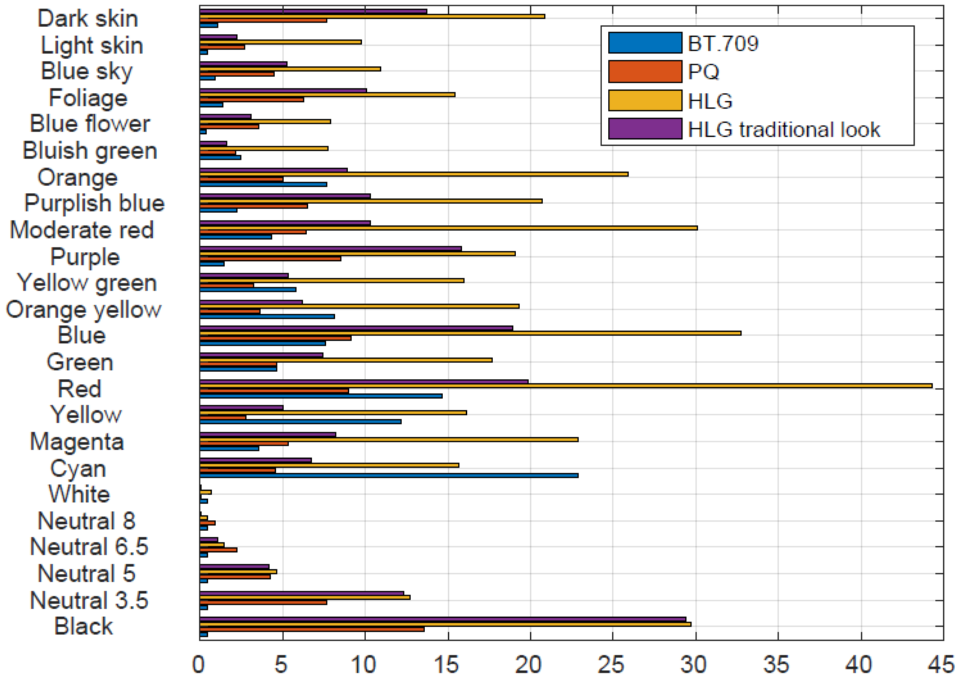


FIGURE A6-7

Display light colour swatch Delta E_{ITP} with BT.2020



A6.3 Comparison with the reference colour pattern data

Instead of using one of the production formats as a reference for comparison, the original colour pattern reference data [13] can be used as well. To do so, the reference data is linearly scaled (i.e. a linear OOTF is applied) such that the white swatch luminance is approximately 200 cd/m^2 as for the other formats.

The saturation differences with the scaled colour pattern reference data are shown in Fig. A6-8. For the HLG format, the saturation differences are all 0, because the HLG format preserves the chromaticity (and therefore the saturation) of the scene as imaged by the camera. Note that this does not imply that the HLG format preserves the chromaticity of the original scene, since camera image adjustments, such as white balancing, will change the chromaticity.

The Delta E_{ITP} differences with the scaled reference data are shown in Fig. A6-9. For the HLG format, the differences are purely luminance differences, caused by the difference between the relative scene luminances (linear OOTF) and the displayed luminances for the HLG format (OOTF gamma 1.2).

FIGURE A6-8

Display light colour swatch saturation differences with the reference colour pattern

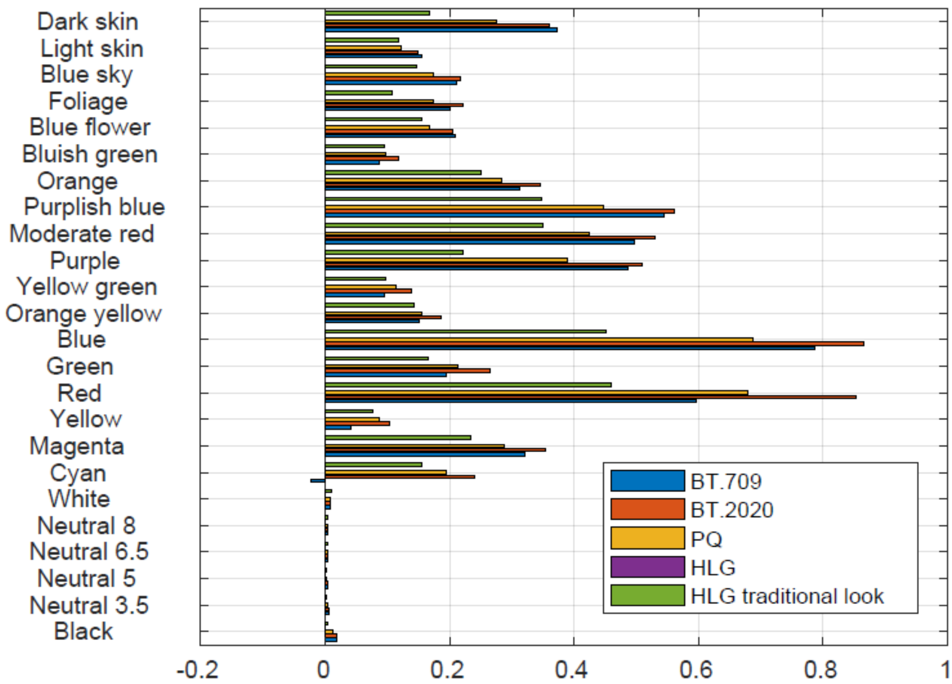
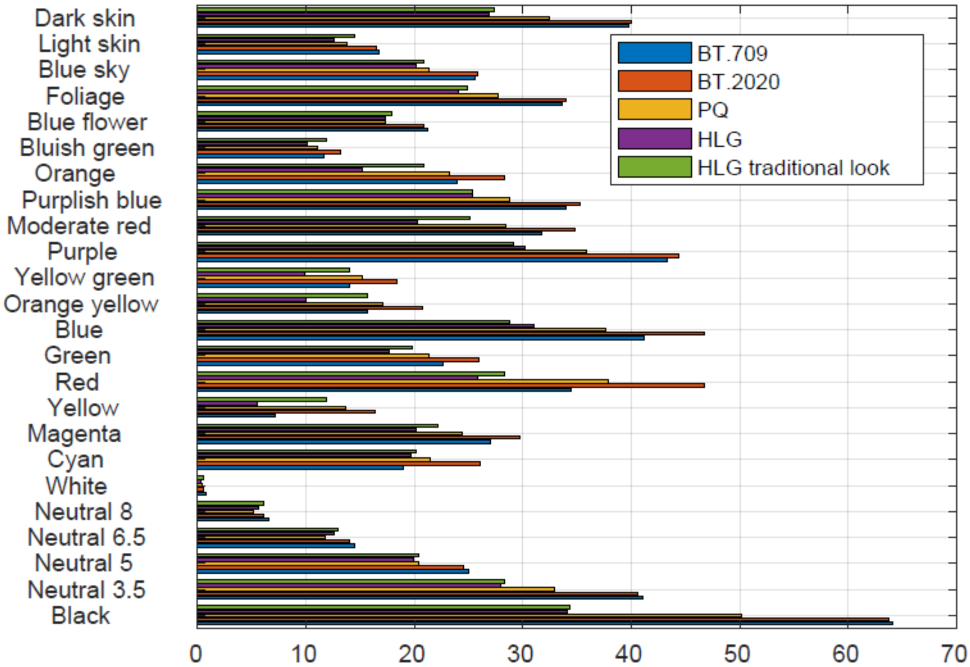


FIGURE A6-9

Display light colour swatch Delta E_{ITP} with the scaled reference colour pattern



Annex 7

Calculating the normalized primary matrix

The normalized primary matrix is needed for the conversion process to and from the CIE XYZ colour space and the BT.2100 colour space, as described in § 8.

Camera and display systems are commonly defined by their normalized primary matrix, NPM, which is specified as follows:

$$\mathrm{NPM} = \begin{bmatrix} X_R & X_G & X_B \\ Y_R & Y_G & Y_B \\ Z_R & Z_G & Z_B \end{bmatrix}, \quad (1)$$

where the elements of the matrix depend on the chromaticity coordinates, (x_R, y_R), (x_G, y_G), (x_B, y_B), and (x_W, y_W) for red, green, blue, and white, respectively, that characterize each system.

The NPM is needed for the conversion process to and from the CIE XYZ colour space and the BT.2100 colour space. Its elements could be computed as follows:

First, compute the z coordinates for all colour primaries as follows:

$$z_R = 1 - (x_R + y_R) \quad (2)$$

$$z_G = 1 - (x_G + y_G) \quad (3)$$

$$z_B = 1 - (x_B + y_B) \quad (4)$$

$$z_W = 1 - (x_W + y_W) \quad (5)$$

Then the matrix elements of NPM are derived as follows:

$$X_R = \frac{((y_G*z_B-y_B*z_G)*x_W+(x_B*z_G-x_G*z_B)*y_W+(x_G*y_B-x_B*y_G)*z_W)*x_R}{(x_R*(y_G*z_B-y_B*z_G)-x_G*(y_R*z_B-y_B*z_R)+x_B*(y_R*z_G-y_G*z_R))*y_W} \quad (6)$$

$$X_G = \frac{((y_B*z_R-y_R*z_B)*x_W+(x_R*z_B-x_B*z_R)*y_W+(x_B*y_R-x_R*y_B)*z_W)*x_G}{(x_R*(y_G*z_B-y_B*z_G)-x_G*(y_R*z_B-y_B*z_R)+x_B*(y_R*z_G-y_G*z_R))*y_W} \quad (7)$$

$$X_B = \frac{((y_R*z_G-y_G*z_R)*x_W+(x_G*z_R-x_R*z_G)*y_W+(x_R*y_G-x_G*y_R)*z_W)*x_B}{(x_R*(y_G*z_B-y_B*z_G)-x_G*(y_R*z_B-y_B*z_R)+x_B*(y_R*z_G-y_G*z_R))*y_W} \quad (8)$$

$$Y_R = \frac{((y_G*z_B-y_B*z_G)*x_W+(x_B*z_G-x_G*z_B)*y_W+(x_G*y_B-x_B*y_G)*z_W)*y_R}{(x_R*(y_G*z_B-y_B*z_G)-x_G*(y_R*z_B-y_B*z_R)+x_B*(y_R*z_G-y_G*z_R))*y_W} \quad (9)$$

$$Y_G = \frac{((y_B*z_R-y_R*z_B)*x_W+(x_R*z_B-x_B*z_R)*y_W+(x_B*y_R-x_R*y_B)*z_W)*y_G}{(x_R*(y_G*z_B-y_B*z_G)-x_G*(y_R*z_B-y_B*z_R)+x_B*(y_R*z_G-y_G*z_R))*y_W} \quad (10)$$

$$Y_B = \frac{((y_R*z_G-y_G*z_R)*x_W+(x_G*z_R-x_R*z_G)*y_W+(x_R*y_G-x_G*y_R)*z_W)*y_B}{(x_R*(y_G*z_B-y_B*z_G)-x_G*(y_R*z_B-y_B*z_R)+x_B*(y_R*z_G-y_G*z_R))*y_W} \quad (11)$$

$$Z_R = \frac{((y_G*z_B-y_B*z_G)*x_W+(x_B*z_G-x_G*z_B)*y_W+(x_G*y_B-x_B*y_G)*z_W)*z_R}{(x_R*(y_G*z_B-y_B*z_G)-x_G*(y_R*z_B-y_B*z_R)+x_B*(y_R*z_G-y_G*z_R))*y_W} \quad (12)$$

$$Z_G = \frac{((y_B*z_R-y_R*z_B)*x_W+(x_R*z_B-x_B*z_R)*y_W+(x_B*y_R-x_R*y_B)*z_W)*z_G}{(x_R*(y_G*z_B-y_B*z_G)-x_G*(y_R*z_B-y_B*z_R)+x_B*(y_R*z_G-y_G*z_R))*y_W} \quad (13)$$

$$Z_B = \frac{((y_R*z_G-y_G*z_R)*x_W+(x_G*z_R-x_R*z_G)*y_W+(x_R*y_G-x_G*y_R)*z_W)*z_B}{(x_R*(y_G*z_B-y_B*z_G)-x_G*(y_R*z_B-y_B*z_R)+x_B*(y_R*z_G-y_G*z_R))*y_W} \quad (14)$$

All the chromaticity values for R, G, B, and White are defined in three or four decimal digits in ITU-R texts, from which the transformation matrices or NPMs are derived. All values shown in the matrices below were calculated with high precision and then rounded to four decimal digits. It is suggested that the matrix calculations should be performed using high precision coefficient values without rounding.

A7.1 Conversion of normalized linear colour signals to Recommendation ITU-R BT.2100

In the case for conversion to the BT.2100 colour space, where the source colour space is linear, normalized within the [0:1] range, and defined by a particular NPM, conversion can be done as follows:

$$\begin{bmatrix} E_R \\ E_G \\ E_B \end{bmatrix}_{BT.2100} = \begin{bmatrix} 1.7167 & -0.3557 & -0.2534 \\ -0.6667 & 1.6165 & 0.0158 \\ 0.0176 & -0.0428 & 0.9421 \end{bmatrix} * \mathrm{NPM}_{Source} * \begin{bmatrix} E_R \\ E_G \\ E_B \end{bmatrix}_{Source}$$

and:

$$\begin{bmatrix} E_R \\ E_G \\ E_B \end{bmatrix}_{BT.2100} = \begin{bmatrix} 1.7167 & -0.3557 & -0.2534 \\ -0.6667 & 1.6165 & 0.0158 \\ 0.0176 & -0.0428 & 0.9421 \end{bmatrix} * \begin{bmatrix} X_R & X_G & X_B \\ Y_R & Y_G & Y_B \\ Z_R & Z_G & Z_B \end{bmatrix}_{Source} * \begin{bmatrix} E_R \\ E_G \\ E_B \end{bmatrix}_{Source} \quad (15)$$

Finally, since not all colours in the source representation may be within the BT.2100 representation, an additional clipping process may be performed. The negative values may be clipped to zero. The positive values may also be clipped to the capabilities of the interface. Although both soft or hard clipping could be performed (see Report ITU-R BT.2407, in many applications hard clipping is preferred. In the scenario that hard clipping of only the negative values is performed the process would be as follows:

$$E_R = \mathrm{Max}(0, E_R) \quad (16)$$

$$E_G = \mathrm{Max}(0, E_G) \quad (17)$$

$$E_B = \mathrm{Max}(0, E_B) \tag{18}$$

The above transformations could be applied in both display and scene referred workflows.

The conversion process, assuming a display referred camera workflow, as well as the final conversion to a BT.2100 representation, is shown in Fig. A7-1. For conversion to HLG, a bridge point of 1 000 cd/m^2 is assumed, and can therefore use the reference OOTF (see § 6.2 of Report ITU-R BT.2390).

FIGURE A7-1

Conversion of arbitrary display referred linear light signals to BT.2100 signals using a display referred workflow

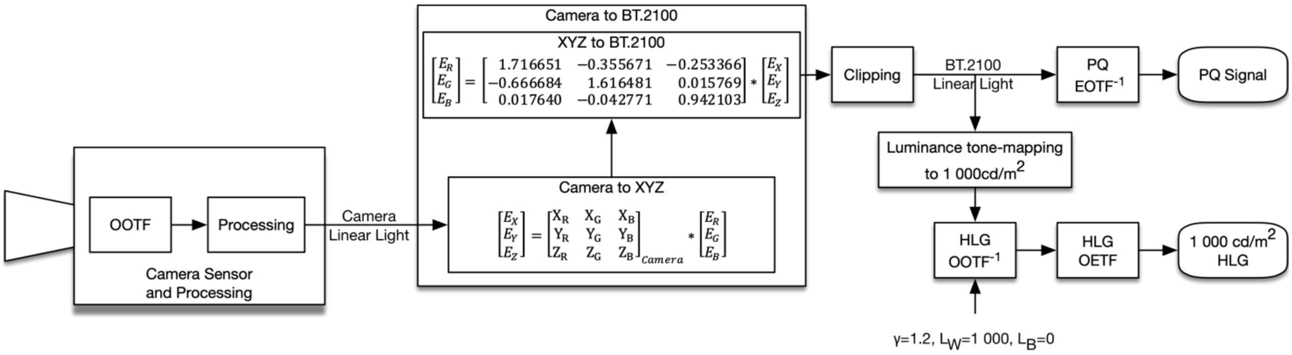


Figure A7-2 depicts the conversion process when applied on a scene referred workflow with the HLG BT.2100 signal as its output. Figure A7-3 depicts the same conversion process when applied on a scene referred workflow with the PQ BT.2100 signal as its output.

FIGURE A7-2

Conversion of arbitrary scene referred light signals to a BT.2100 HLG signal using a scene referred workflow

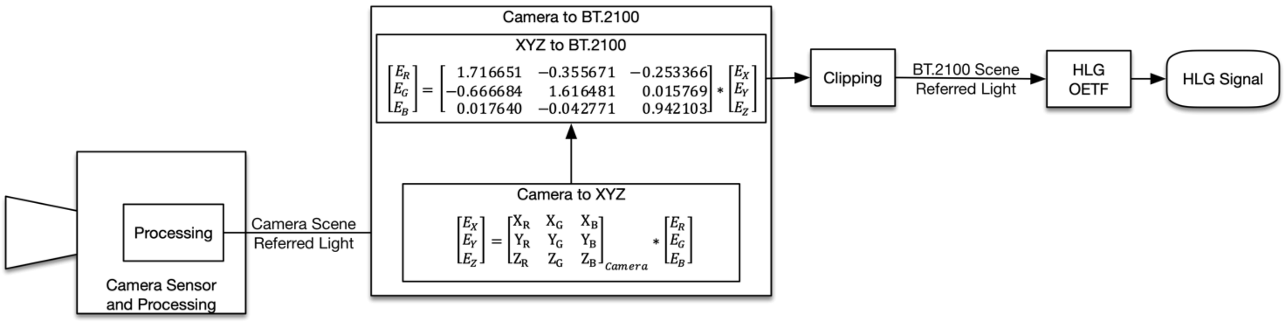
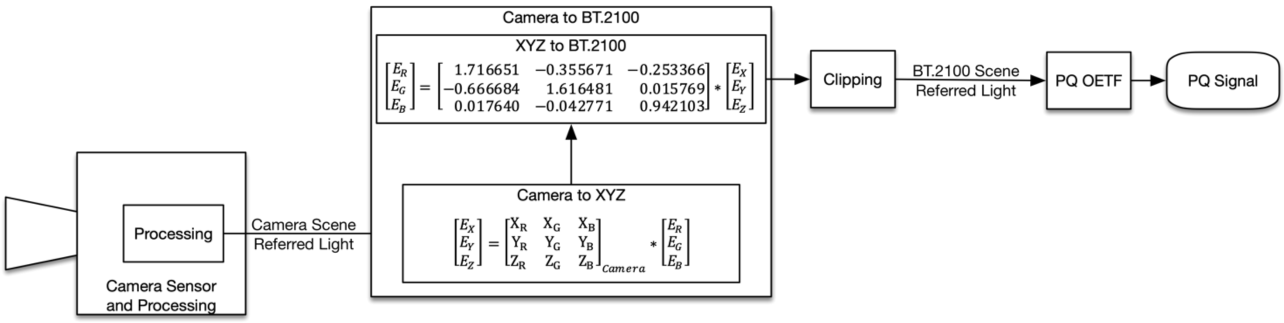


FIGURE A7-3

Conversion of arbitrary scene referred light signals to a BT.2100 PQ signal using a scene referred workflow



A7.2 Conversion of BT.2100 to arbitrary linear colour signals for display systems

Similarly, conversion from linear and normalized BT.2100 RGB primaries to the RGB primaries of an arbitrary display system can be performed as follows:

$$\begin{bmatrix} E_R \\ E_G \\ E_B \end{bmatrix}_{Display} = \mathrm{NPM}_{Display}^{-1} * \begin{bmatrix} 0.6370 & 0.1446 & 0.1689 \\ 0.2627 & 0.6780 & 0.0593 \\ 0.0000 & 0.0281 & 1.0610 \end{bmatrix} * \begin{bmatrix} E_R \\ E_G \\ E_B \end{bmatrix}_{BT.2100}$$

and:

$$\begin{bmatrix} E_R \\ E_G \\ E_B \end{bmatrix}_{Display} = \begin{bmatrix} \mathrm{X_R} & \mathrm{X_G} & \mathrm{X_B} \\ \mathrm{Y_R} & \mathrm{Y_G} & \mathrm{Y_B} \\ \mathrm{Z_R} & \mathrm{Z_G} & \mathrm{Z_B} \end{bmatrix}_{Display}^{-1} * \begin{bmatrix} 0.6370 & 0.1446 & 0.1689 \\ 0.2627 & 0.6780 & 0.0593 \\ 0.0000 & 0.0281 & 1.0610 \end{bmatrix} * \begin{bmatrix} E_R \\ E_G \\ E_B \end{bmatrix}_{BT.2100} \tag{19}$$

Not all colours in the original representation may be within the target representation.

The negative values may be clipped to zero. The positive values may also be clipped to the capabilities of the display. Although both soft or hard clipping could be performed, in many applications, such as when using a reference display, hard clipping is preferred. In the scenario that hard clipping of only the negative values is performed the process would be as follows:

$$E_R = \mathrm{Max}(0, E_R) \tag{20}$$

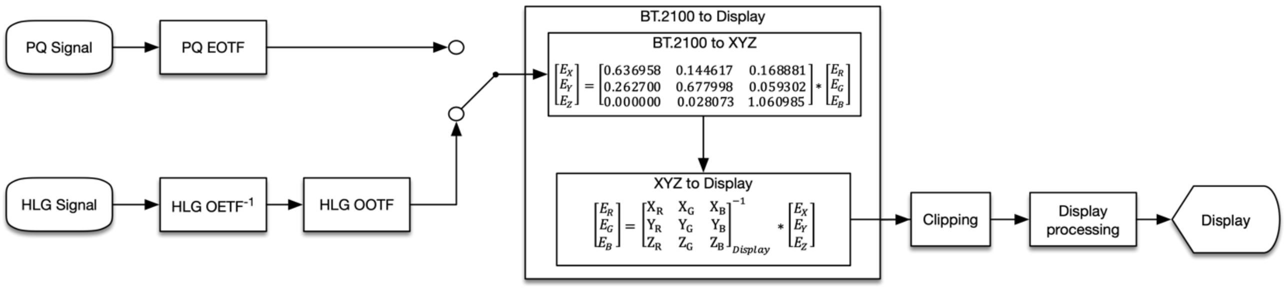
$$E_G = \mathrm{Max}(0, E_G) \tag{21}$$

$$E_B = \mathrm{Max}(0, E_B) \tag{22}$$

Figure A7-4 depicts this conversion process assuming a display referred workflow for both PQ and HLG. For conversion from HLG, the nominal peak luminance of the target display (and the appropriate system gamma) is used for the HLG OOTF.

FIGURE A7-4

Conversion of BT.2100 signals to an arbitrary display using a display referred workflow



References

[1] Fitzpatrick, T.B. (1988), "The validity and practicality of sun reactive skin types I through VI". Arch Dermatol 124(6), pp. 869-871.

[2] K.C. Noland, M. Pindoria and A. Cotton, "Modelling brightness perception for high dynamic range television", *Ninth International Conference on Quality of Multimedia Experience (QoMEX)*, Erfurt, 2017.

[3] K.C. Noland and M. Pindoria, "A Brightness Measure for High Dynamic Range Television", IBC Conference, September 2017.

[4] MovieLabs, "MovieLabs Best Practices for Mapping BT.709 Content to HDR10 for Consumer Distribution v1.0", https://movielabs.com/download/938/.

[5] EBU Recommendation R103 version 3.0, "Video Signal Tolerance in Digital Television Systems", May 2020 (https://tech.ebu.ch/docs/r/r103.pdf).

[6] Cooksey, C., Allen, D.W., et.al., "Reference Data Set of Human Skin Reflectance", NIST Journal of Research, NIST, USA, Jun 2017.

[7] Jablonski N.G., Chaplin, G., "*The evolution of human skin coloration*", Journal of Human Evolution, Vol. 39, 2000.

[8] ISO/CIE 11664-5:2016(en), *"Colorimetry – Part 2: CIE Standard Illuminants",* International Organisation for Standardisation, Geneva 2016.

[9] Colour and Vision Research Laboratory, University College London, http://cvrl.ioo.ucl.ac.uk/.

[10] DCraw https://www.cybercom.net/~dcoffin/dcraw/.

[11] Australian Radiation Protection and Nuclear Safety Agency, "*Fitzpatrick Skin Type*", questionnaire to determine Fitzpatrick Skin Type https://www.arpansa.gov.au/sites/g/files/net3086/f/legacy/pubs/RadiationProtection/FitzpatrickSkinType.pdf.

[12] Burk, A., Smith, M., Zink, M., "Color Volume and Hue-Preservation in HDR Tone Mapping", SMPTE Motion Imaging Journal. May 2020.

[13] SMPTE ST 303:2002, "Color reference pattern".

[14] ISO/CIE 11664-5:2016, "Colorimetry — Part 5: CIE 1976 L*u*v* colour space and u', v' uniform chromaticity scale diagram".

[15] Elena A. Fedorovskaya and Huib de Ridder "Subjective matters: from image quality to image psychology", Proc. SPIE 8651, Human Vision and Electronic Imaging XVIII, 86510O (23 March 2013); https://doi.org/10.1117/12.2008681

[16] ISO/CIE 11664-6:2014, "Colorimetry — Part 6: CIEDE2000 Colour-difference formula".

Glossary

Following is a list of terms within Report ITU-R BT.2408 which may not have been encountered by the reader in the context of High Dynamic Range.

Camera RAW signals - image data produced by, or internal to, a digital camera that has not been processed, except for A/D conversion and the following optional steps: linearization, dark current/frame subtraction, shading and sensitivity (flat field) correction, flare removal, white balancing (e.g. so the adopted white produces equal RGB values or no chrominance), missing colour pixel reconstruction (without colour transformations).

Colour branding - the intentionally applied look of graphics added to a programme designed to create or maintain a consistent and familiar appearance.

Display-light conversion - conversion of image data performed by converting the image data to display light using the reference EOTF of the first encoding, and then applying the reference inverse EOTF for the second encoding, typically used for preserving the appearance of graded content.

Gregory hole reference - an object used as reference black with 0% reflectance, typically a box lined with black material designed to absorb light and prevent reflections of light which is viewed through a small opening.

HDR reference white - the nominal signal level obtained from an HDR camera and a 100% reflectance white card resulting in a nominal luminance of 203 cd/m^2 on a PQ display or on an HLG display that has a nominal peak luminance capability of 1 000 cd/m^2.

Lambertian reflector - a reflecting surface which reflects incident light in all directions, giving the same apparent brightness regardless of the angle of view of an observer.

Luma - a term specifying that a signal represents the monochrome information related to non-linear colour signals. The symbol for luma information is denoted as Y'.

NOTE – The term luma is used rather than the term luminance in order to signify the use of *non-linear* light transfer characteristics as opposed to the linear characteristics in the term luminance. However, in many of the ITU-R Recommendations on television systems, the term “luminance signal” is used rather than “luma” for Y' together with C'_B and C'_R.

Luminance - the photometrically weighted flow of light per unit area travelling in a given direction. It describes the amount of light that passes through, is emitted from, or is reflected from a particular area, and falls within a given solid angle. It is expressed in candelas per square meter (cd/m^2).

NOTE – The relative luminance of a pixel can be approximated by a weighted sum of the *linear* colour components; the weights depend on the colour primaries and the white point.

Scene-light conversion - conversion of image data performed by converting the image data back to scene light using the inverse of the first encoding reference OETF, and then applying the reference OETF for the second encoding, typically used for matching cameras.

NOTE – The image data can be from a graded source (for example, from a camera that has been adjusted to give a particular desired appearance on a reference monitor), but scene light conversions might not preserve the colour appearance produced on the reference monitor.

Overshoots - signal excursions above nominal peak level.

Sub-blacks - in a narrow range signal, a video signal of lower than 0% black level extending down to approximately 6.8% below black level. In the case of 10-bit digital coding this range lies below value 64 (black level) extending down to value 4, while in 12-bit digital coding this range lies below value 256 extending down to value 17.

Super-white - in a narrow range signal, a video signal of greater than 100% nominal peak level extending up to 109% of nominal peak level. In the case of 10-bit digital coding this range lies above value 940 (nominal peak) extending to value 1 019, while in 12-bit digital coding this range lies above value 3 760 extending to value 4 079.

Undershoots - signal excursions below black level.