

International Telecommunication Union

ITU-R
Radiocommunication Sector of ITU

Report ITU-R BT.2408-1
(04/2018)

**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 Annex 1 of 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, 2018

© ITU 2018

All rights reserved. No part of this publication may be reproduced, by any means whatsoever, without written permission of ITU.

REPORT ITU-R BT.2408-1

Operational practices in HDR television production

(2017-2018)

Summary

These operational practices are intended to provide initial guidance 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. They should be read in conjunction with Report ITU-R BT.2390 which provides additional background information.

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.

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/m² 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.

Initial findings are given 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.

A practical guide is included for those transitioning from standard dynamic range (SDR) to high dynamic range (HDR) production. Key factors are also described for broadcasters to consider in facilitating the successful introduction of HDR-TV.

Experience of television programme production in HDR continues to grow, but is still in its relative infancy. These operational practices provide guidance based on knowledge gained so far.

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 of a 100% reflectance white card. 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 white¹. 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, 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/m² (nominal peak luminance) display, under controlled studio lighting². 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/m² 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.

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.

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.

TABLE 1
Nominal signal levels for PQ and HLG production

Reflectance Object or Reference (Luminance Factor, %) ³	Nominal Luminance, cd/m ² (PQ & 1000 cd/m ² HLG)	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 1 – The signal level of “HDR Reference White” is not 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 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.

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.

TABLE 2

Preliminary signal levels for common objects in PQ and HLG production

Reflectance Object	Nominal Luminance cd/m ² (PQ & 1 000 cd/m ² HLG)	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”. The narrow range representation is in widespread use and is considered the default. The full range representation was newly introduced into Recommendation ITU-R BT.2100 with the intention of being used only when all parties agree.

The use of narrow range signals is strongly preferred for HLG, to preserve the signal fidelity and to reduce the risk of mistaking full range for narrow range signals (and vice versa) in production. Common video processing techniques such as image re-sizing, filtering and compression create overshoots that extend above the nominal peak luminance into the “super-white” region (where the signal $E' > 1.0$), and create under-shoots that extend below black into the “sub-black” region (where the signal $E' < 0.0$). In order to maintain image fidelity, it is important that the over-shoots and under-shoots are not clipped, which would happen if full-range signals were used. Furthermore, the black level of an HLG display used in production should be adjusted using the Recommendation ITU-R BT.814 PLUGE signal, which is made easier if sub-blacks are present in the signal.

The full range representation is useful for PQ signals and provides an incremental advantage against visibility of banding/contouring and for processing. Because the range of PQ is so large, it is rare for content to contain pixel values near the extremes of the range. Therefore, over-shoots and under-shoots are unlikely to be clipped.

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 $Y' C'_B C'_R$ signal format and the constant intensity $IC_T C_P$ format.

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.

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 $Y'C'_B C'_R$ 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/m² 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/m²). 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) 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 legacy 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 BT.2100 as having a 5 cd/m² surround), manufacturers may provide modified brightness and display characteristics intended to compensate for the different viewing environment.

3.2 Display of HLG signals

Table 5 of Recommendation ITU-R BT.2100 specifies the HLG EOTF for reference displays. Note 5e 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.

Note 5f 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, as appropriate.

Firstly, the monitor gamma is adjusted, according to the formula in Note 5e 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/m² surround).

TABLE 3

HLG Display Gamma

Nominal Peak Luminance (cd/m²)	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/m²)	HDR Reference White (cd/m²)
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 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.

⁵ Measured perpendicular to the screen.

⁶ Assuming ~ 60% reflectance surround.

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.

4.1 Comfortable brightness of static images

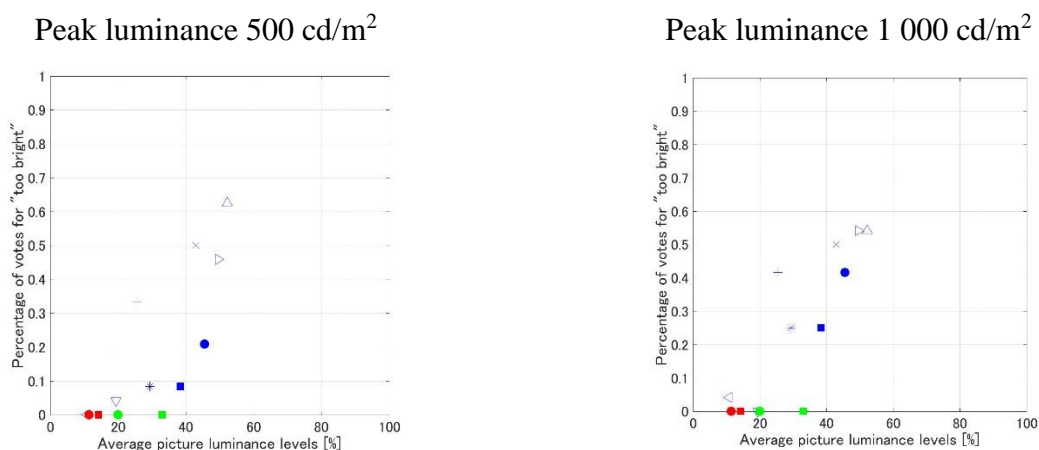
A study was performed by NHK to learn what range of luminances are judged comfortable by viewers. A number of SDR images that, on a 100 cd/m² reference monitor, varied in average luminance over a range of 10-50 cd/m², were used. The study was conducted using a relative display system that employed a 3 500 cd/m² 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/m² were simulated. Viewers were asked to judge whether images were “appropriate”, “too bright”, or “too dark”.

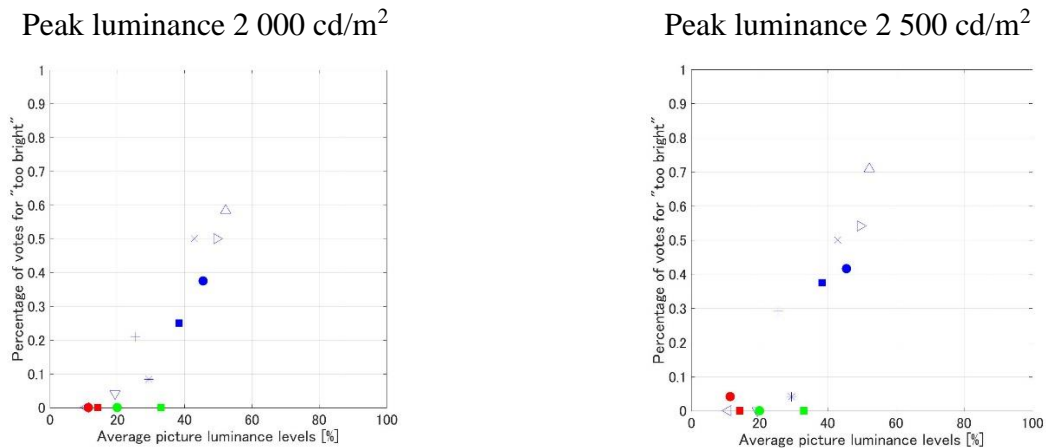
Figure 1 shows the results in the reference viewing environment (dim surround). For each simulated display peak luminance, images with average luminance less than 25% of the peak luminance being simulated were not judged as “too bright”. Images with average luminance greater than 25% of peak luminance 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/m² on a 1 000 cd/m² HLG monitor, would not be judged as too bright on an HLG monitor of any luminance up to at least 2 500 cd/m².

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/m² and 4 000 cd/m², 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/m² on a 1 000 cd/m² display. Similar comments were not made about the test scenes that had average luminances of 144 and 128 cd/m² on a 1 000 cd/m² display.

FIGURE 1

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





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 Brightness Shifts

Unexpected changes in image brightness might occur at programme junctions, with interstitials or channel changes. 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 (21). For the tests, the luminance behind the screen was 5 cd/m^2 , and the peak screen luminance was $1\,000 \text{ cd/m}^2$ (3). Subjects were asked to rate the change in overall brightness between two still HDR images.

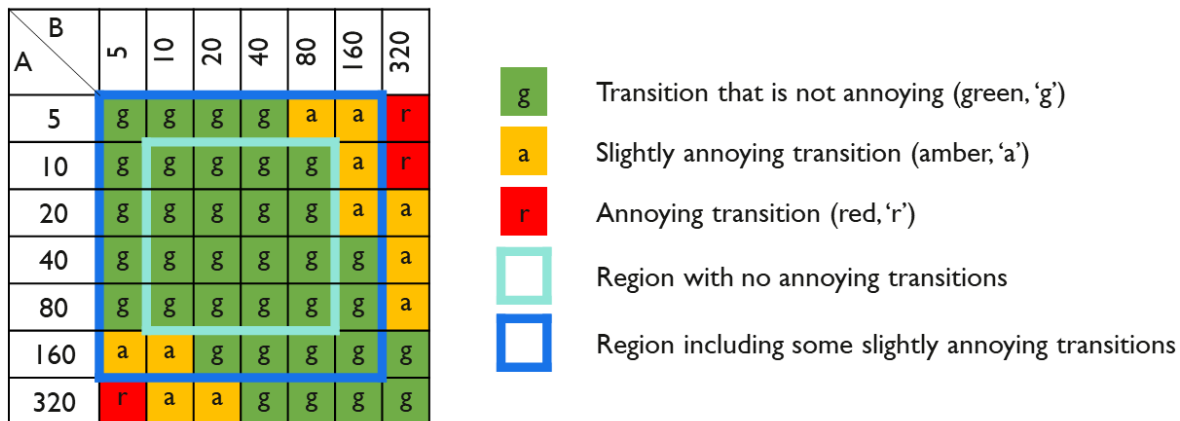
Figure 2 shows the overall results, with transitions from the first brightness A to the second brightness B categorised according to whether they are “not annoying”, “slightly annoying”, or “annoying”. Two regions are marked in the figure with thick blue lines. The inner region, with mean display luminance levels of 10 to 80 cd/m^2 , contains no annoying jumps, and so could be considered a suitable range for normal operation. The outer region, with mean display luminance levels of 5 to 160 cd/m^2 , includes some slightly annoying jumps, and so could be considered an extended range for creative effect. It is expected that night scenes will usually have an overall brightness at the lower end of the normal operating range, and sunny outdoor scenes will have an overall brightness at the upper end of the range. 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. The suggested ranges can be freely exceeded over a short timescale for special creative effect, but it is proposed that the average brightness over the length of a programme be within the normal operating range of 10 to 80 cd/m^2 . It should be noted that this range still allows for significant differences in the average brightness of whole programmes, so, for example, a “moody” or “bright” look can be achieved overall.

The results presented previously in Fig. 1 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. Further experiments reported by the BBC suggest that the ranges are also applicable for HLG displays up to a peak luminance of $4\,000 \text{ cd/m}^2$.

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 2

Transitions from brightness A (cd/m²) to brightness B (cd/m²) categorised by level of annoyance



5 Inclusion of standard dynamic range content

Definitions

Mapping – Placing SDR content in an HDR signal to preserve the “look” of the SDR content (note that the word “mapping” (when used alone) is not a shortened form of the term “tone mapping”, which is a different process, as defined below).

Inverse tone mapping – Placing SDR content in an HDR signal with expanded luminance range to emulate an HDR look (i.e. up-conversion).

Tone mapping – Converting HDR content to an SDR signal range (i.e. down-conversion).

Standard dynamic range (SDR) content may either be “mapped” or “up-converted” into an HDR format for inclusion in HDR programmes.

Methods for mapping, place 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 “look” of the SDR content when shown on an HDR display.

In contrast, up-conversion (sometimes known as inverse tone mapping) is intended to expand luminance values to use more of the available HDR luminance range, and thereby leverage the display capabilities. In this Report, the focus is on mapping of SDR content into an HDR container.

There are two possible approaches to SDR 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 BT.709 or BT.2020 display, when the content is shown on a BT.2100 HDR display; possibly at a slightly higher peak luminance to provide a value for diffuse white and skin tones that is more consistent with the brightness of native HDR content. An example of which is the inclusion of SDR graded content within an HDR programme.
- Scene-referred mapping is used where the source is a direct SDR camera output and the goal is to match the colours and lowlights and mid-tones of a BT.2100 HDR camera. An example of which is the inter-mixing of SDR and HDR cameras within a live television production.

More detailed technical descriptions of the above mapping process, including how to ensure comparable brightness of skin tones between HDR and mapped SDR content, are given in Report ITU-R BT.2390-3.

The nominal signal levels described in § 2.2 may be considered for format conversions, including when mapping SDR content into HDR. HDR-produced content is expected to vary around the nominal signal levels to allow artistic freedom, and to adapt to different scenes and objects.

5.1 Matching BT.709 SDR and BT.2100 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. 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. 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.

5.2 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-conversion rather than mapping is used to place the content into an HDR signal container. The up-conversion 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.

6 Conversion between PQ and HLG

Methods of converting between the PQ and HLG formats are described in section 7 of Report ITU-R BT.2390-2 (or later).

Notably, because of the difference in the way that PQ and HLG signals are rendered on displays of different peak luminance, a conversion rather than a simple transcode is required. By choosing a reference peak displayed luminance (L_w) of 1 000 cd/m² for the HLG signal, and requiring that the PQ signal be limited to the same peak luminance, consistent brightness is achieved in the converted signals.

In general, signals converted from HLG to PQ will retain the HLG “look”, while signals converted from PQ to HLG will retain the PQ “look”. So care should be taken when measuring test signals (e.g. colour bars, camera test charts) using a vector-scope or CIE colour chart after conversion.

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

The displayed “look” of SDR BT.709, SDR BT.2020, HLG and PQ signals are all different, because they all utilise different end-to-end OOTFs. In the reference viewing environment, SDR BT.709 and BT.2020 have an end-to-end gamma of 1.2 applied to red, green and blue components, but with different colour primaries. PQ applies a similar OOTF to the SDR BT.2020 format on red, green and blue, but extends the range with a modified OOTF for higher luminance levels. In contrast, by design, HLG applies its end-to-end gamma on the luminance component, which produces a different “look” described elsewhere as “natural”. So extreme care must be exercised when converting between production formats.

Displayed light format conversions will usually maintain the “look” of the original production format, following conversion to the new format. They should be used for “graded” content.

Scene light conversions will change the “look” of content, to match the new output format. They are particularly useful when matching signals from different types of cameras (see § 5).

The Figures below present examples of production architectures for HLG, but similar architectures may also be used for PQ.

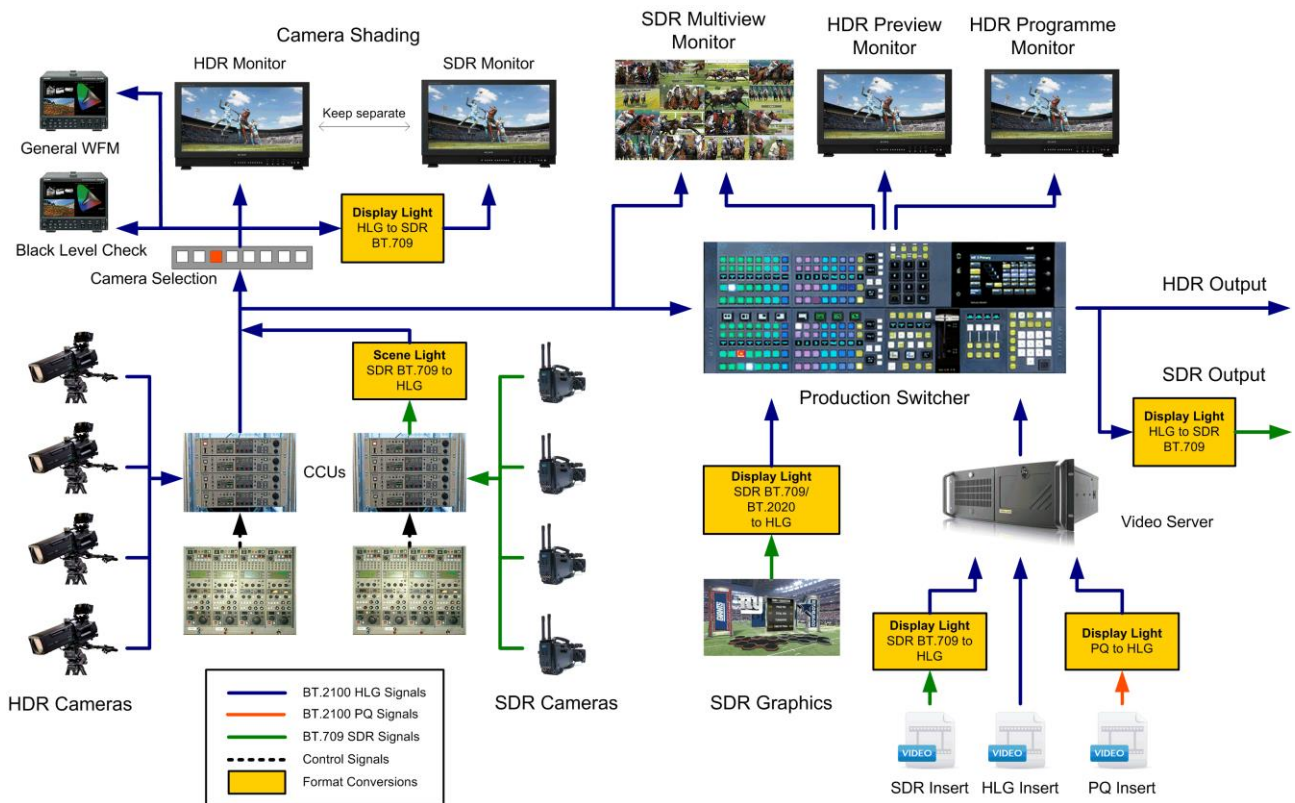
7.1 Live production

During the transition from SDR to HDR production, the majority of viewers will be watching in SDR. So it is essential that the SDR production is not compromised in any way 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.

A native HDR production architecture is illustrated in Fig. 3. This can be used to either maintain or enhance the quality of the SDR output, as described in §§ 7.1.1 and 7.1.2 respectively. When the utmost compatibility is required between SDR productions made, on different occasions, with or without HDR cameras, or when utmost consistency is required using cameras with different colour gamuts, the architecture illustrated in Fig. 4, and described in § 7.1.3, may be preferred. 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.

FIGURE 3

HDR production with SDR derived by down-conversion



In this example, input to the production switcher is HDR. This removes the need to process separate HDR and SDR feeds throughout the production chain. The primary output from the production

switcher is HDR, and the SDR BT.709 output is derived by a down-conversion of the HDR BT.2100 signal. The Figure illustrates HLG production, but PQ based production is similar.

SDR BT.709 cameras can be included in the production by using the “scene-referred” SDR mapping technique, with OOTF adjustment, as described in § 5 (additional technical details are given in Report BT.2390).

Graded SDR content and SDR graphics may be included in the HDR programme. To ensure that the SDR “look” is maintained, the “display-referred” SDR mapping described in § 5 should be used.

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.

Where SDR content is mapped into an HDR container and the HDR signal then down-converted to feed an SDR service (known as SDR-HDR-SDR “round-tripping”), the mapped SDR content may appear darker on the SDR service than if it had been broadcast directly. Section 7.2 discusses issues that can arise from “round-tripping”.

Differences in black level may be more visible in the down-converted SDR signal than in the HDR signal, as 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.

This architecture can be used to either maintain or enhance the quality of the SDR output, as described below.

7.1.1 HDR focussed production

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

7.1.2 SDR focussed 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-converter.

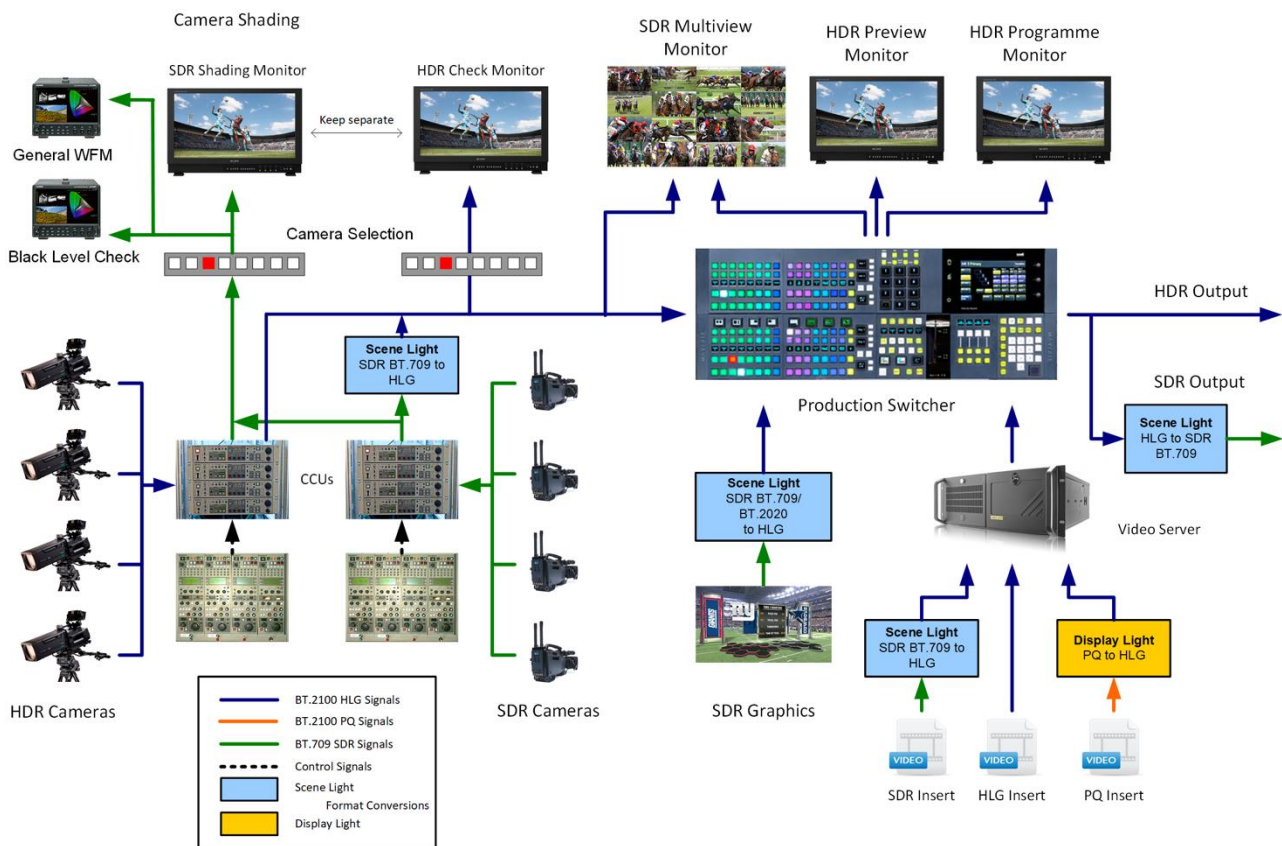
Note, however, that the eye may adapt to the brighter HDR screen, affecting the appearance of signals on the dimmer SDR screen. So the HDR and SDR screens should be separated for critical assessment of the down-converted SDR signal.

7.1.3 Production for maximum SDR compatibility

Although the arrangement of Fig. 3 can, potentially, produce a high quality for both HDR and SDR it sometimes most important to achieve the maximum consistency with the traditional SDR “look”. This can only be achieved by shading (or “racking”) the cameras in SDR and, also, by limiting the colour gamut to the smallest gamut available from any camera used. A suitable architecture is illustrated in Fig. 4 for HLG production, but a similar architecture is also appropriate for PQ.

FIGURE 4

HDR production for maximum SDR compatibility with SDR derived by down-conversion



This architecture is similar to that of Fig. 3. Here, however, shading or racking is performed directly on the SDR output available from the Camera Control Units (CCUs). The SDR output available from the CCU of an HDR camera is typically derived from the linear sensor output such that the desired white level from the camera achieves the correct signal level for both the SDR and HDR outputs. This is achieved by applying a 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 HDR signal. In the SDR signal, highlights greater than the super-white signal level (109%) are lost, but they are retained in the HDR signal.

Further considerations for HLG

A fundamental difference between Fig. 4 and Fig. 3 is that here, in Fig. 4, the SDR signal has the “traditional” look whilst the HLG HDR signal has the “natural” look. By design HLG signals have a “natural” look compared to the “traditional” look of SDR cameras (as described in Report ITU-R BT.2390). By contrast, in Fig. 3, both the SDR and the HDR signals have the same “natural” look⁷.

It is essential that the final SDR production corresponds precisely with the pictures seen by the shader or camera operator. In this architecture, in Fig. 4, the shader or camera operator views a conventional SDR monitor and hence is viewing a picture with the “traditional” look. Clearly this traditional look must be maintained in the final SDR production output. However the production, with the exception of shading/racking, is performed in the higher quality HDR format which, for HLG, has the natural

⁷ In Fig. 3 the “Display Light” format conversions may be changed to “Scene Light” format conversions. If this is done then the both the SDR shading monitor, and the final, downconverted, SDR output will have the “traditional” look, and Fig. 3 becomes equivalent to Fig. 4.

look. Consequently, to maintain the same look seen by the shader, the conversion of the final HDR signal to SDR, on the righthand side of Fig. 4, must be performed using a scene referred conversion. In Fig. 3 this final conversion is display referred, so this is a crucial difference between the two architectures.

7.2 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-conversion to HDR and down-conversion from HDR to SDR.

There are two main approaches to including SDR content in HDR programmes. The first is mapping the signal to HDR so that an SDR signal is “contained” within the HDR signal. The alternative approach, up-conversion, or inverse tone mapping, aims to increase the dynamic range.

Down-conversion of HDR to SDR is considered in Report BT.2390; for example, in the section on display mapping where the conversion can be regarded as a mapping of HDR onto an SDR display. Typically, HDR to SDR conversion might use 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-conversion and down-conversion, 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 highlights (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).

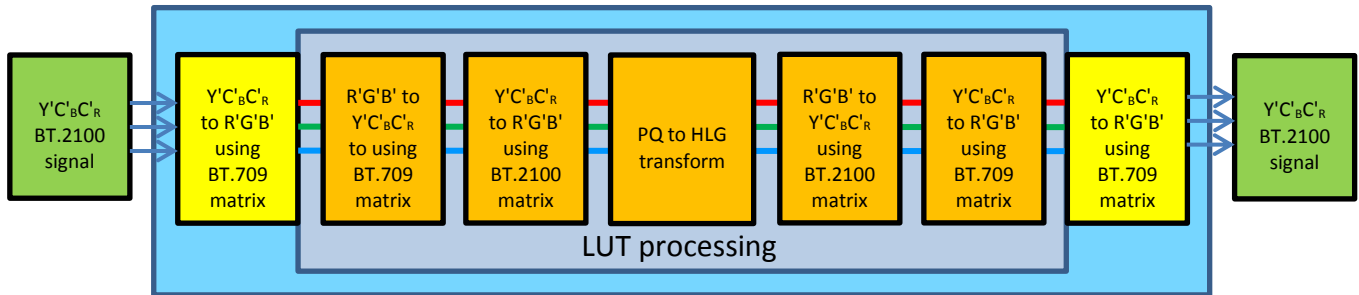
The optimum techniques for up-conversion followed by down-conversion are still under investigation.

7.3 Hardware colour matrices

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

Where it is not possible to configure a device for BT.2100 colour, 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. 5, within a look-up table performing a format conversion.

FIGURE 5
Example of colour matrix compensation within a LUT



Annex 1

Study to evaluate levels for PQ content

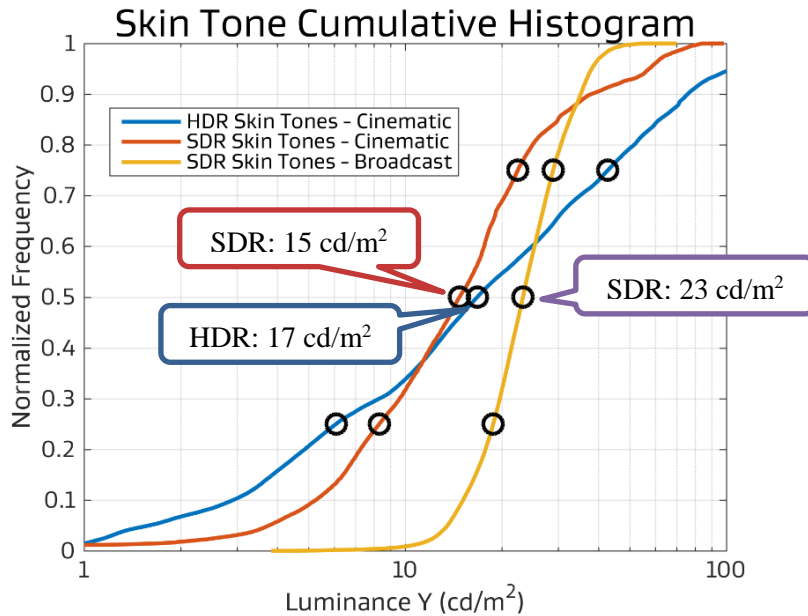
A study was performed to gain information that could be used to inform initial guidance on video levels for HDR production. The study used existing SDR materials from both broadcast content and home video content. The study also used PQ HDR materials, mostly from home video grades of movies that were done on a $4\,000\text{ cd/m}^2$ 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:

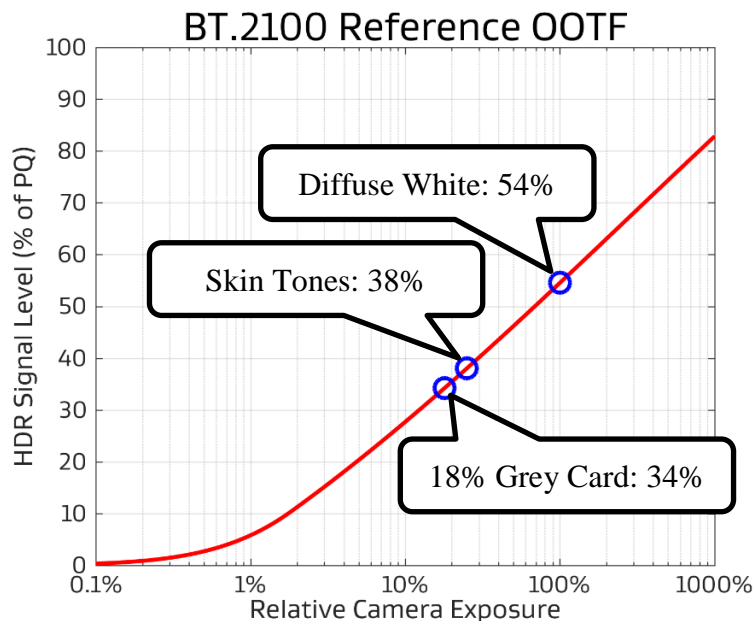


Due to the scarcity of HDR broadcast content currently available, in order to compare HDR and SDR content, the same analysis was completed utilizing HDR and SDR graded indoor scenes from cinematic content for home distribution. The cumulative histogram is given below.



For cinematic content for the home, HDR Caucasian skin tones are very similar to SDR skin tones (17 cd/m² compared to 15 cd/m²), but the standard deviation is larger. Extrapolating from this, it is hypothesized that indoor Caucasian skin tones in HDR broadcast may average 26 cd/m² with a larger deviation than SDR broadcast. The 26 cd/m² 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/m² 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/m² 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

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.

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/m² display device. This dataset is included in Report ITU-R BT.2245-4. In this dataset the arithmetic mean luminance is 65.47 cd/m² (standard deviation 83.99 cd/m²). The geometric mean luminance is 9.17 cd/m² (standard deviation 24.74 cd/m²).

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:

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

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 1000 cd/m² 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 cd/m ²
1%	321.89 (262.14)
5%	195.04 (206.15)
10%	145.03 (170.56)
20%	(145.01)

3 Diffuse white elements in live HLG encoded broadcast content

Diffuse white elements⁸ 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 cd/m^2 assuming a display peak luminance of $1000 \text{ cd}/\text{m}^2$, 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.

TABLE A2.3

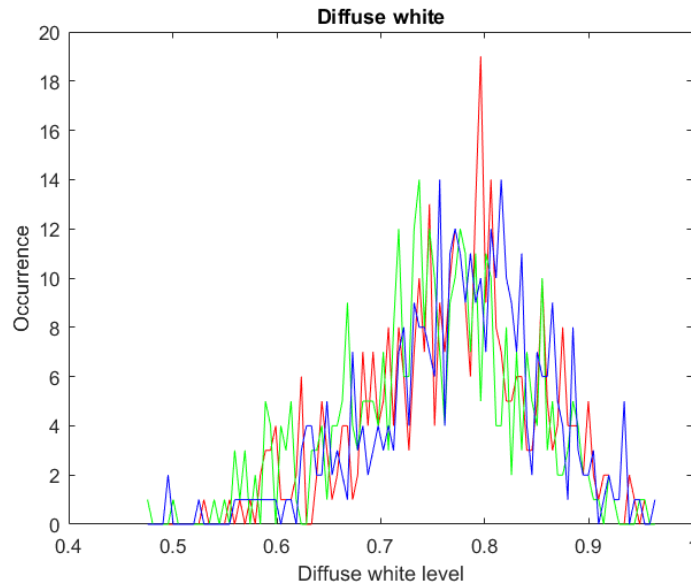
Analysis diffuse white of HLG encoded live broadcast content. The %HLG column was measured, while the remaining columns were derived from these measurements (152 frames, 378 points analysed)

Diffuse White	$\frac{\text{cd}}{\text{m}^2}$	%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)

⁸ For the purpose of this Annex, the level of diffuse white elements is referred to as “diffuse white”.

FIGURE A2.1

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



4 Diffuse white in an HDR dataset of 1000 cd/m² PQ encoded images

A dataset of 54 EXR images containing diffuse white patches was analysed using the same methodology as described in § 3. The dataset contains images that are graded for a 1000 cd/m² 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/m² 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	$\frac{\text{cd}}{\text{m}^2}$	%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, 1000.0, 946.5)	(98.2, 100, 99.1)	(74.1, 75.2, 74.6)

FIGURE A.2.2

**Distribution of Diffuse White patches in a test database of 1000 cd/m² images.
The values on the horizontal axis are in %PQ**

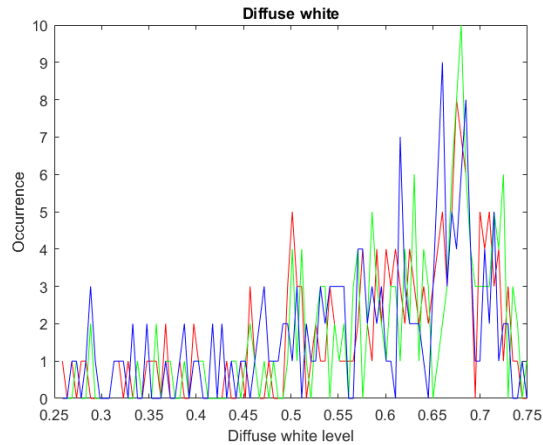


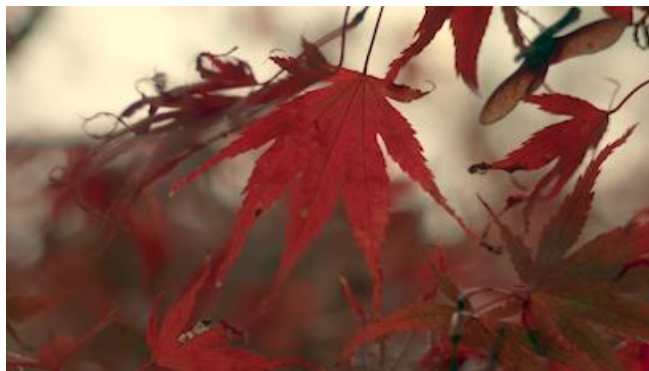
FIGURE A2.3

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)



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 cd/m² in a set of 107 HDR images that were graded at 1000 cd/m². Likewise, 95% of the same images have between 6.6% and 13% of their pixels larger than 203 cd/m².

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 cd/m².

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 cd/m² for reference level. However, the standard deviation was about 8.3% (measured in %HLG), which – for an assumed 1000 cd/m² signal - translates to a range between around 123 and 345 cd/m² (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 cd/m² 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.

6 Conclusions

The HDR Reference White level of 203 cd/m² 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.

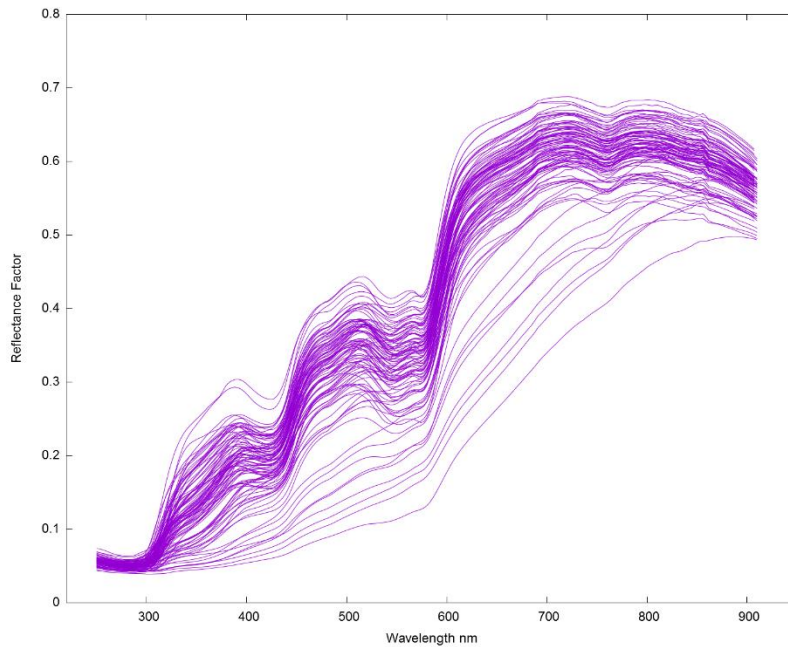
1 Study 1: using a skin tone database and an ideal model of a camera

A skin tone reflectance database from the US government National Institute of Standards and Technology (NIST) [4] 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 [5], 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.

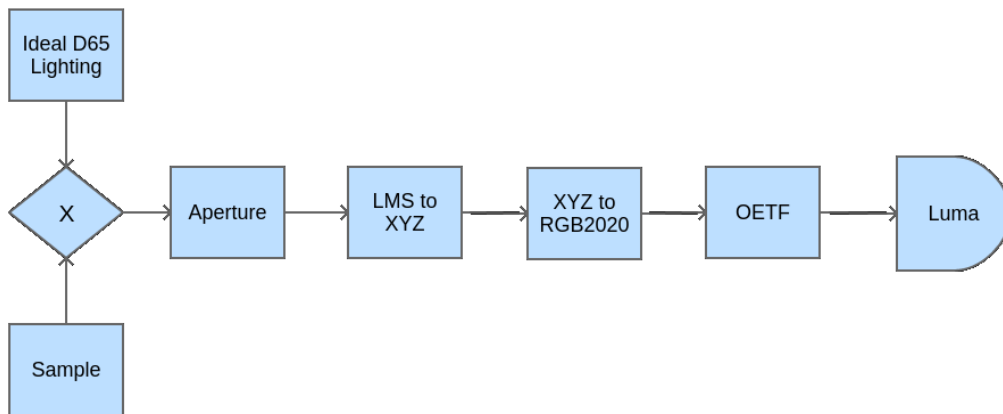
FIGURE A3.1
NIST Dataset
Each line corresponds to one skin sample



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 [6] fed through an aperture (a fixed scalar). A set of CIE 1931 2 degree observer LMS to XYZ curves [7] 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 signal 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 [6]. These regional labels have been added to the plot.

A further plot of skin tone reflectance against screen emittance for a 1000 cd/m² 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 [5]

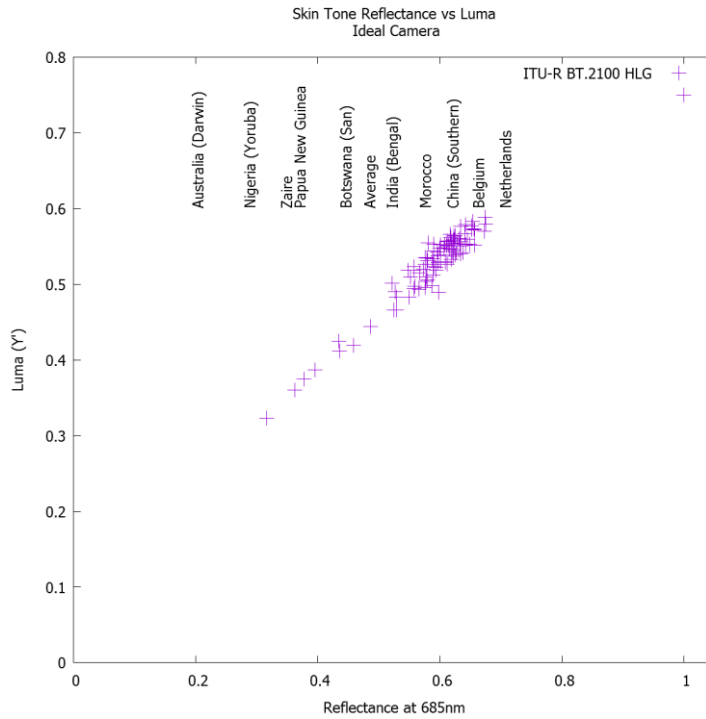
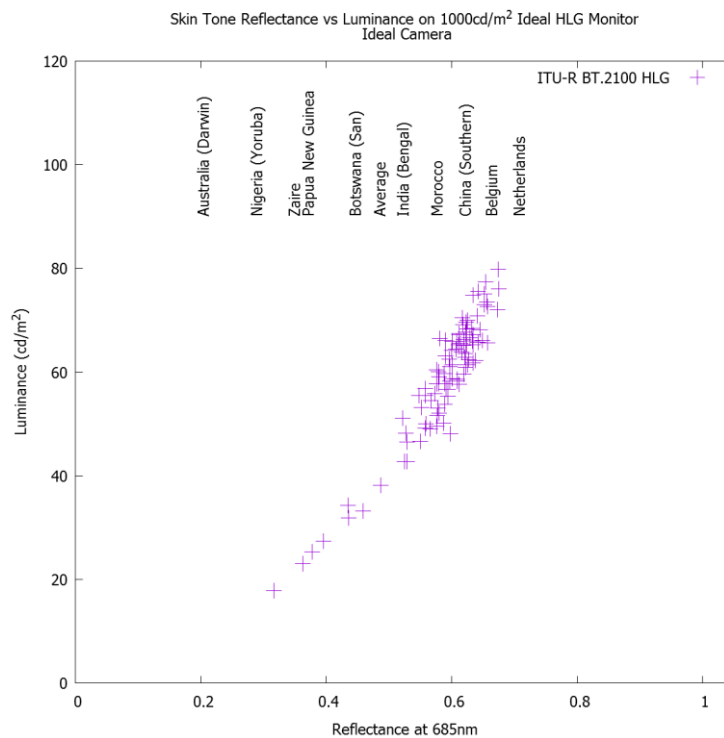


FIGURE A3.4

Skin tone reflectance at 685 nm against HLG luminance on a 1000 cd/m² display, for ideal camera, with regional labels from [5]



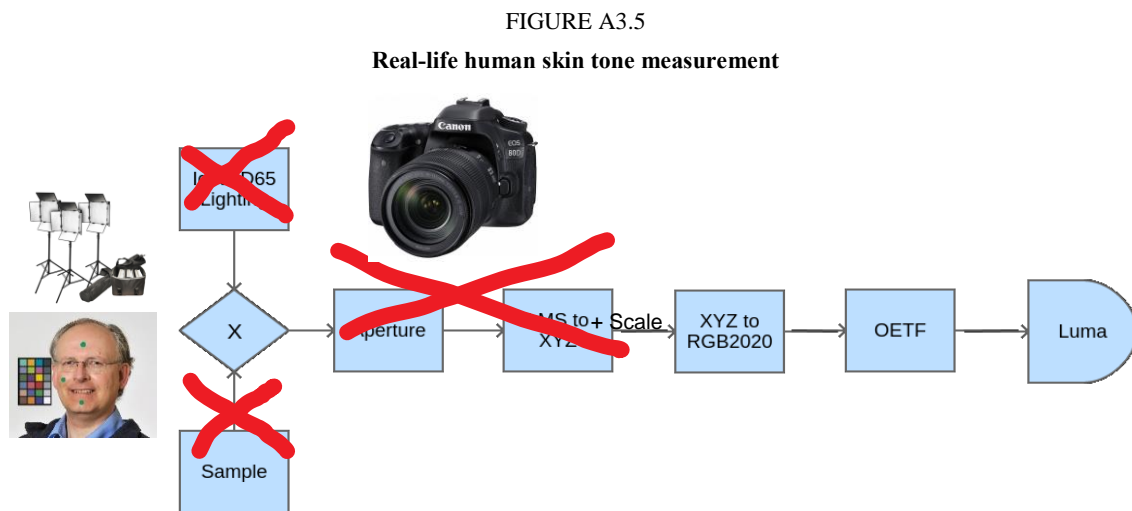
2 Study 2: using human subjects and a RAW recording camera

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 [8] 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 ITU-R BT.2100) and then calculate the Y' luma 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.



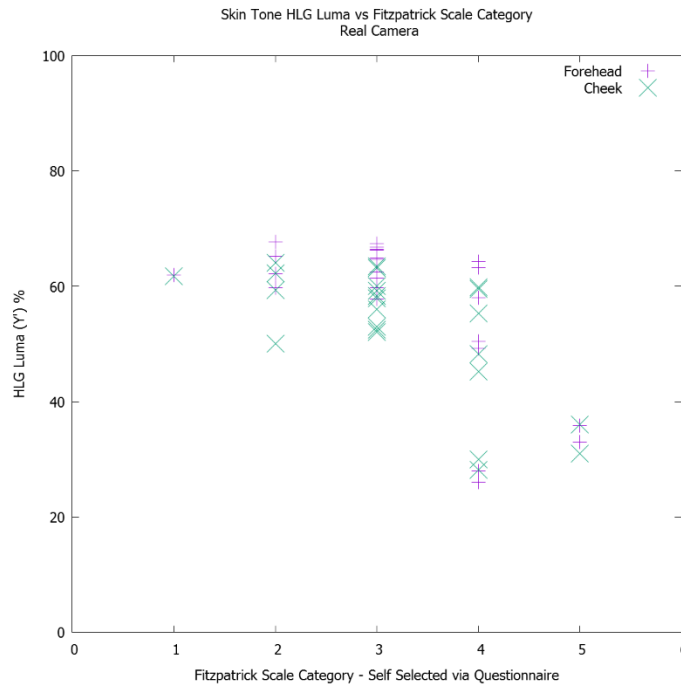
In order to match the test subject to the Fitzpatrick Scale classifications, a questionnaire from the Australian Government Radiation Protection and Nuclear Safety Agency was used [9].

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

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.

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

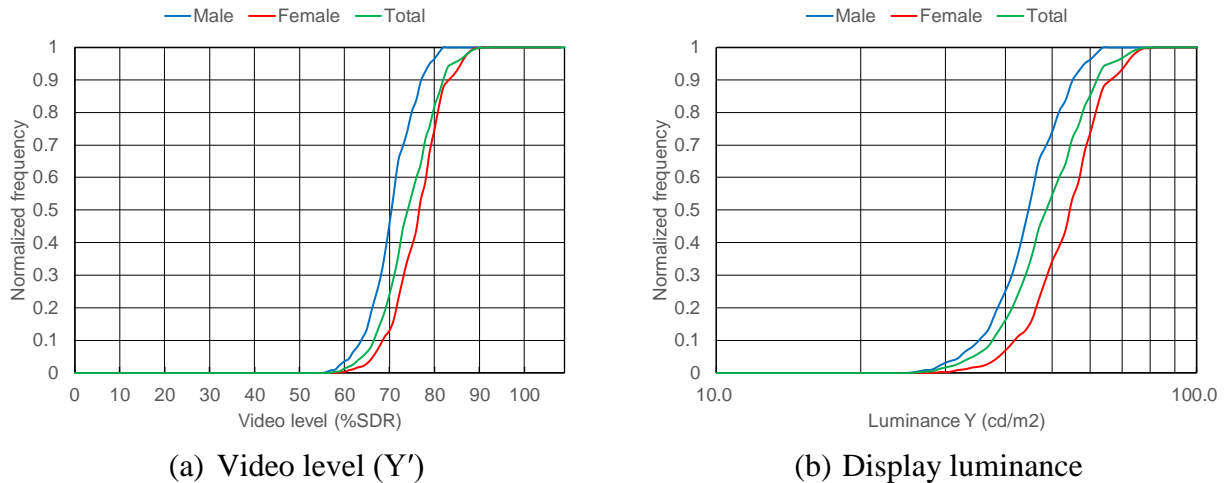
⁽¹⁾ Japan Broadcasting Corp., Asahi Broadcasting Corp, Nippon Television Network Corp., Tokyo Broadcasting System Television, Fuji Television Network, TV Asahi Corp., TV Tokyo Corp., and WOWOW.

⁽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 ($\sigma=5.2$), 77.6 ($\sigma=5.7$), and

74.6 ($\sigma=6.2$) %SDR, respectively. These video levels correspond to luminance of 45 cd/m², 55 cd/m², and 49 cd/m² on a display with the peak luminance of 100 cd/m². The luminance of facial skin tones is more than twice the 23 cd/m² reported in Annex 1 to Report ITU-R BT.2408 for SDR broadcast content.

FIGURE A4.1

Cumulative histogram of facial skin tones

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.

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

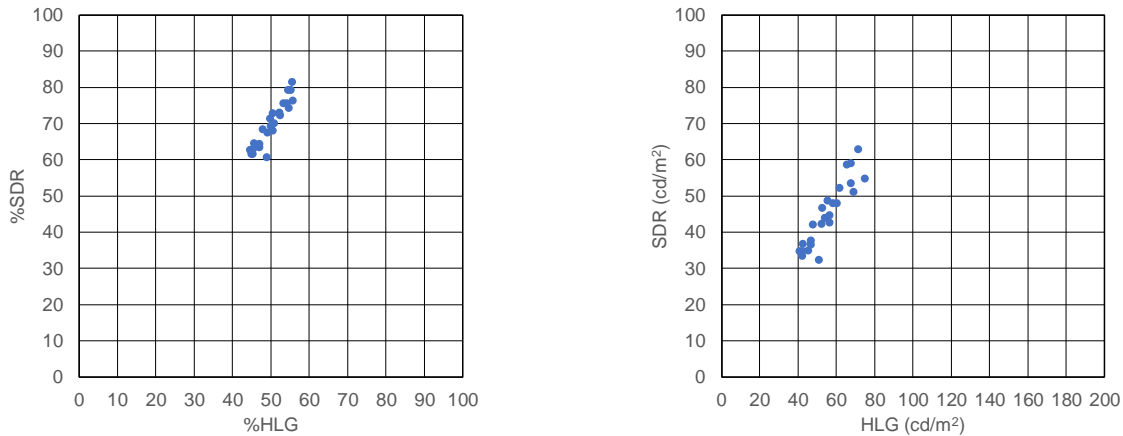
A preliminary study was conducted on skin tones in HLG HDR content in comparison with SDR content. Both HLG and SDR content were produced independently for the same NHK music programme in a concert where musicians performed on stage under special lighting and set. In the HDR production, video engineers 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/m² on HLG displays with a peak luminance of 1000 cd/m², and 61-82%SDR (70%SDR on average) and 32-63 cd/m² on SDR displays with a peak luminance of 100 cd/m².

The values for SDR correspond well to those for the SDR news and information programmes in studio described in § 1 of this Annex. 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



(a) Video level (Y')

(b) Display luminance

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 viceversa. 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

Programme genre		News and information in studio			Music programme in concert hall	
Format		SDR			SDR	HLG
Graphics white		100% SDR			100% SDR	75% HLG
Average skin tones	Signal level	Male	Female	Total	70% SDR	50% HLG
		72% SDR	78% SDR	75% SDR		
	Display luminance	45 cd/m ²	55 cd/m ²	49 cd/m ²	45 cd/m ²	55 cd/m ²
on a display of 100cd/m ² peak						on a display of 1000 cd/m ² peak

Annex 5

Factors facilitating successful HDR-TV

Broadcasters will benefit from a graceful, non-disruptive introduction of HDR-TV into television broadcasting and this will require the study of an end-to-end migration path from current standard dynamic range (SDR) broadcasting to HDR-TV broadcasting. With that in mind, following are some key factors to be considered when selecting which of the two BT.2100 HDR systems are selected for a given application.

The majority of television audiences will, for a period of several years, continue to watch programmes on consumer displays that were not designed to render HDR images. It will be important that the HDR-TV image systems will allow easy automatic conversion of an HDR-TV programme master to a standard dynamic range version. Also, mapping of SDR programmes into HDR will be key to inter-mixing old and new content in HDR programming.

The HDR-TV image system should provide, where appropriate, a degree of compatibility with existing workflows and broadcasters' legacy infrastructure⁹, including the possibility to use HDR-TV in live and non-live workflows and to easily intermix HDR-TV and SDR-TV programme material both in the temporal and in the spatial domain, including graphics and video overlays. The HDR-TV image system should allow easy image and waveform monitoring throughout the broadcast chain allowing for different viewing environments while providing consistent image reproduction at each point.

The HDR-TV image system and the creative practices in production should be arranged so they lead to no adverse effects such as visual fatigue or discomfort when viewed for a significant period of time. Additionally, care should be taken in HDR production and in SDR up-conversion with regard to the effect a greater image dynamic range may have on those viewers affected by visual disturbances such as photosensitive epilepsy. This will require some study with respect to the types of scene content that may trigger such adverse effects in such viewers.

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] Cooksey, C., Allen, D.W., et.al., "Reference Data Set of Human Skin Reflectance", *NIST Journal of Research*, NIST, USA, Jun 2017.
- [5] Jablonski N.G., Chaplin, G., "The evolution of human skin coloration", *Journal of Human Evolution*, Vol. 39, 2000
- [6] ISO/CIE 11664-5:2016(en), "Colorimetry — Part 2: CIE Standard Illuminants", International Organisation for Standardisation, Geneva 2016.

⁹ In this context, the term "infrastructure" includes all processing and connectivity (SDI, Bit Rate Reduction, Switchers, Routers, etc.).

- [7] Colour and Vision Research Laboratory, University College London, <http://cvrl.ioo.ucl.ac.uk/>.
- [8] DCraw <https://www.cybercom.net/~dcoffin/dcraw/>.
- [9] Australian Radiation Protection and Nuclear Safety Agency, “*Fitzpatrick Skin Type*”, questionnaire to determine Fitzpatrick Skin Type
<https://www.arpsa.gov.au/sites/g/files/net3086/f/legacy/pubs/RadiationProtection/FitzpatrickSkinType.pdf>.
-