International Telecommunication Union

ITU-R

Radiocommunication Sector of ITU

Report ITU-R BT.2446-0

(04/2019)

Methods for conversion of high dynamic range content to standard dynamic range content and vice-versa

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, 2019

REPORT ITU-R BT.2446-0

Methods for conversion of high dynamic range content to standard dynamic range content and vice-versa

(2019)

Summary

High dynamic range (HDR) content may be converted to a significantly lower dynamic range so that it may be displayed on a standard dynamic range (SDR) display device. As the dynamic range mismatch between content and display capabilities can be significant, the manner in which the dynamic range is reduced impacts the visual quality that can be achieved. Likewise, SDR content, such as for instance legacy content, may be converted to HDR content to enable display on a HDR display device, or to bring such content into a HDR video workflow. This Report provides guidance toward the design of methods that achieve such conversions, which are referred to as tone mapping and inverse tone mapping, and it presents example methods along with their evaluation.

Keywords

Standard dynamic range, high dynamic range, tone mapping, inverse tone mapping

Abbreviations/Glossary

HDR	High dynamic range
HDR-TV	High dynamic range television
HDTV	High definition television
ITMO	Inverse tone mapping operator
SDR	Standard dynamic range
TMO	Tone mapping operator
UHDTV	Ultra-high definition television

Related ITU Recommendations, Reports

Recommendation ITU-R BT.1886 – Reference electro-optical transfer function for flat panel displays used in HDTV studio production

Recommendation ITU-R BT.2020 – Parameter values for ultra-high definition television systems for production and international programme exchange

Recommendation ITU-R BT.2087 – Colour conversion from Recommendation ITU-R BT.709 to Recommendation ITU-R BT.2020

Recommendation ITU-R BT.2100 – Image parameter values for high dynamic range television for use in production and international programme exchange

Report ITU-R BT.2245 – HDTV and UHDTV including HDR-TV test materials for assessment of picture quality

Report ITU-R BT.2390 – High dynamic range television for production and international programme exchange

Report ITU-R BT.2408 – Operational practices in HDR television production

1 Introduction

This Report describes methods for conversion of standard dynamic range (SDR) content to high dynamic range (HDR) content and vice-versa. In this context, SDR content may be defined as a signal that is nominally produced to be viewed at a peak luminance of 100 cd/m^2 (nits) in the reference viewing environment, a black level close to zero and a bit-depth of 8 or 10 bits. As defined in Recommendation ITU-R BT.2100, High Dynamic Range Television (HDR-TV) provides viewers with an enhanced visual experience by providing images that have been produced to look correct on brighter displays, that provide much brighter highlights, and that provide improved detail in dark areas.

An important aspect of an HDR eco-system involves the meaningful conversion between SDR content and HDR, for example to enable existing content to be brought into an HDR workflow, or to check SDR content on an HDR display device. Likewise, HDR content may have to be converted to SDR for monitoring purposes, or for distribution. Conversions whereby SDR content is placed into an HDR container without changing the dynamic range is generally termed 'mapping', and this is discussed in Report ITU-R BT.2390.

Instead, the focus of this Report is on conversion methods whereby the dynamic range is changed, either from SDR levels to HDR, or vice-versa. The goal of such methods is to reproduce content at a different dynamic range without altering the visual experience to the extent possible. One goal of tone mapping HDR content is to allow the output to be intercut with natively produced SDR content, whereas a goal of inverse tone mapping of SDR content is to allow the output to be intercut with natively produced HDR content[1].

In this Report reducing the dynamic range of content will be termed "tone mapping". Similarly, increasing the dynamic range of content will be termed "inverse tone mapping". This is consistent with the academic literature on these topics [12].

The ability to perform round-tripping, i.e. to convert SDR content to HDR and then back to SDR again, whilst minimising the change to the visual experience, is an important requirement. Likewise, round-tripping from HDR to SDR and then back to HDR may also be useful. While such ability is very useful, it is advised to limit the number of such repeated conversions as much as possible, as some signal degradation is likely.

Tone mapping and inverse tone mapping are inherently methods that require trade-offs between, for example, computational complexity, handling of visual artefacts such as over-exposed areas, as well as the general mapping of luminance values, which may be different between live and graded content. The selection of an appropriate method should therefore be made bearing in mind any specific use case, or the nature of the content. To help such a selection to be made, this Report contains three methods that each make different trade-offs, as discussed in the introductory paragraphs of the sections in which they are described. The methods have in common, however, that they are static, offering a fixed processing independent of content or viewing environment. When evaluating the subjective performance of these methods it becomes clear that, depending on the particular video content, a different method may be preferred. It should be noted that adaptive, dynamic and spatially varying methods also exist, and it is possible that a suitable dynamic method could provide better performance across a wider range of video content.

The structure of this Report is as follows. Section 2 describes general guidelines for HDR to SDR conversion, SDR to HDR conversion, and round-trip performance. Section 3 discusses a pair of tone mapping and inverse tone mapping methods that map between 100 and 1 000 cd/m^2 under the assumption that the input data is of sufficiently high quality. When the quality of the input data is lower, for example in the presence of large over-exposed areas, less strong mappings may be favourable. Section 4 describes such methods for tone mapping and inverse tone mapping. Section 5 presents a method that is designed to preserve skin tones, and to adhere to reference levels (see Report

---

1 The term 'intercut' broadly refers to intermixing natively produced and converted content temporally, but also refers to spatial montages such as for example picture-in-picture applications.

ITU-R BT.2408). Section 6 offers a feature comparison between the three methods presented in this Report. Finally, Annex 1 presents an evaluation of the round-trip performance of the methods presented in § 3.

2 Guidelines for conversion between SDR and HDR

2.1 Guidelines for SDR to HDR conversion

Recommendation ITU-R BT.2100 specifies methods for encoding HDR content, enabling effective programme interchange. The perceptual quality (PQ) and hybrid log-gamma (HLG) systems documented in that Recommendation require an HDR linear light signal to be present as input. The origins of such signals are not described. Section 5 of Report ITU-R BT.2408-0 describes methods for placing SDR content in PQ and HLG containers without changing the actual dynamic range of the content.

SDR content comes from a variety of sources, with different levels of quality, for example as function of the technologies available at the time of production, or the presence of compression. As inverse tone mapping is designed to expand the dynamic range of content, such expansion may amplify artefacts, dependent on their presence and severity. As a consequence, some content may best be mapped without expanding the dynamic range, as described in Report ITU-R BT.2408.

Generally, however, a non-linear expansion may be preferred, thereby offering a different trade-off between management of artefacts and visual appearance. Such inverse tone mapping methods may take inspiration from knowledge about non-linearities in human visual perception, and thereby produce results that better match human vision.

The desire in such cases would be to increase the dynamic range of the content to effectively enhance its visual appearance. This requirement may be translated into several objectives that any SDR to HDR conversion process should adhere to:

1) maintain details in the shadows;
2) ensure that mid-tones are not unduly expanded;
3) expand highlights up to the peak display luminance, insofar the quality of the content allows (see discussion below);
4) ensure chromatic content is adjusted appropriately;
5) maintain temporal stability.

The motivations for these requirements are as follows. First, much of the look-and-feel of the content is encoded in the shadows and the mid-tones. By applying a non-linear function to the content may change the contrast between pixels. As a result, care should be taken that changing pixel relationships between pixels in the darks and mid-tones do not unacceptably alter the appearance of the image.

On the other hand, darks and mid-tones cannot be processed in exactly the same way, as shadow details would become relatively less visible, especially when viewed adjacent to much brighter highlights as a result of veiling glare. As such, dark tones would have to be expanded slightly more than the mid-tones. The mid-tones cannot be expanded by a large amount, because this would raise the mean luminance level too much, leading to the potential for significant viewer discomfort. Furthermore, account should be taken of the changes in perception of relative tones as the eye adapts to brighter inverse tone-mapped images.

The brightest parts of the image can receive the greatest amount of expansion, striking a balance between creating the increased impact associated with HDR imaging and maintaining the director’s intent. Note that expanding non-graded content, for instance live broadcast content captured with an SDR camera would benefit from the same considerations.

The design of an SDR to HDR conversion process should, however, take into consideration that very strong expansions may reveal artefacts. Even if a well-designed inverse tone mapping method does not introduce artefacts by itself, it may amplify artefacts that were already present in the input material. For instance, while aggressive compression schemes may introduce quantization that at SDR display levels may be sub-threshold, at HDR display levels such artefacts may end up supra-threshold. Such content should not be scaled to high luminance levels. However, compression in production and interchange environments is often light or absent, leaving much content which could be expanded with little worry for the visible appearance of quantization artefacts.

Further, over-exposed areas in the SDR content should be carefully considered, as under certain circumstances expanding those to high luminance levels may lead to visually objectionable results. SDR content exhibiting large over-exposed areas may be processed in one of several ways. First, it may be possible to apply image restoration techniques to such areas. In particular, de-clipping operators may restore some of the lost detail [1]. Second, over-exposed SDR content may be placed in an HDR container without increasing the luminance range, to at least enable such content to be incorporated into an HDR workflow.

Chromatic content needs to be tuned, as human visual perception of chromatic content is dependent on luminance levels [5] [9]. Premier among such luminance induced appearance effects is the Hunt effect, which states that as luminance levels increase, the associated colourfulness also increases [6]. In addition, chromatic content should be scaled in accordance with the amount of luminance expansion, as saturation would otherwise be off.

Finally, no method should introduce temporal artefacts.

With these requirements in mind, current proposals may be assessed. For example, to prepare SDR content for HDR displays, reverse or inverse tone mapping operators (ITMO) can be employed. Such algorithms process the luminance information in the image with the aim of recovering or recreating the appearance of the original scene. Typically, ITMOs take an SDR image as input, expand the luminance range of the image in a global manner, and subsequently process highlights or bright regions locally to enhance the HDR appearance of the result. This is at least in part consistent with the above requirements. Separate consideration of dark tones, however, is rare.

Although several ITMO solutions exist within the academic literature [2] [3] [4] [7] [8] [10], they focus on perceptually reproducing the appearance of the original scene and rely on strict assumptions about the content. Many such techniques propose spatially varying processing, which, if used with care, can indeed produce superior results. However, in the context of programme production and interchange, limits on processing time would prohibit the introduction of such computationally demanding techniques. Additionally, existing solutions are rarely created with video in mind, and therefore have the potential for being temporally unstable.

An obvious way in which inverse tone reproduction may be achieved is to begin by taking standard content which may, for instance, follow Recommendation ITU-R BT.709 (BT.709) or Recommendation ITU-R BT.2020 (BT.2020). In such example case, it would be possible to linearize the signal by applying an appropriate electro-optical transfer function (EOTF). This would produce a normalised relative signal, as information on absolute luminance levels at recording time is not available. An HDR signal may be derived by scaling this content in some way, for instance by linear scaling. However, there are a few concerns with such a simple approach.

First, experience has shown that linear scaling does not produce the best possible visual experience, as this would increase the mean luminance too much. In essence, the mid-tones become too high for comfortable viewing. Such a linear scaling would only work if the illumination of the viewing environment were also increased by the same amount. As this is not the case, linear scaling is inappropriate for luminance expansion. Second, human visual perception does not treat luminance and chrominance independently [5] [6] [9], which means that linear scaling in any colour space would create issues with chroma/saturation.

As a consequence, a better approach would be to apply a non-linear transform. This could take the form of a simple power function, where the exponent would be a constant (i.e. a gamma curve). This would exclude processing in RGB space, as this might introduce hue shifts. Once again, experience has shown that such solution does not yield the best possible visual experience, especially not for graded content where processing of dark tones requires extra care and precision.

In sum, converting SDR to HDR involves requirements that are not met by the simplest approaches such as linear scaling or gamma-like transfer functions. Spatially varying methods, on the other hand, tend to be too computationally demanding for a production environment. A well-designed middle ground between these two extremes would be desirable.

Bearing in mind these guidelines, as well as the need for methods that are highly predictable, the following two conversions are proposed.

2.2 HDR to SDR conversion

Tone mapping, i.e. the non-linear mapping between HDR to SDR content is a well-researched topic (see for example [12] and its references). There is a general notion that a reduction of the dynamic range needs to be governed by at least one secondary goal, i.e. the visual quality needs to be preserved in some way. It has not been clear, however, what aspect(s) of visual quality need to be preserved. Different approaches to tone mapping have aimed to preserve for example brightness, local contrast or visual appearance, each leading to imagery with a different look and feel. An example of how different tone mapping methods can produce different results is shown in Fig. 1.

FIGURE 1

Different tone mapping algorithms tend to produce images with a different look and feel

Report BT.2446-01

Despite several decades of research and hundreds of published papers on tone reproduction, an important observation remains that most available methods for tone mapping inherently incorporate some personal preference of the designers. The methods described in this Report, however, aim to limit this by offering designs that produce SDR content that, to the extent possible, is visually matched to its HDR origins under comparable viewing conditions.

2.3 Round-trip conversion

Currently, SDR and HDR productions occur side-by-side, a situation that will last for the foreseeable future. As a consequence, content may have to undergo conversions back and forth between SDR and HDR. While in general it is advisable that the number of such conversions is limited as much as possible, there are cases where they cannot be avoided. In those cases, it is highly desirable that methods for tone mapping and methods for inverse tone mapping are chosen that are matched/complementary, as this brings the possibility of limiting the loss of visual quality due to round-trip conversions. Examples of matched tone mapping/inverse tone mapping pairs are presented in §§ 3, 4 and 5. Method A, presented in § 3, is evaluated in terms of round-trip conversion quality in Annex 1.

3 Conversion Method A

Content produced for broadcast may originate from a variety of sources including episodic content, movies, ads, as well as live content. Especially movies and episodic content are produced to a high visual quality, as full control can be exerted over the scene, its lighting, as well as the camera settings. Moreover, in a post-production studio, content is typically colour graded to produce the highest possible visual quality. In such use cases, conversion between SDR and HDR content (and vice-versa) can be successfully achieved with methods that do not explicitly anticipate issues related to the quality of the input content. The tone mapping and inverse tone mapping techniques described in this section map content between 100 cd/m^2 and 1 000 cd/m^2 and vice-versa. These levels were chosen because most SDR content is nominally produced at 100 cd/m^2, while HDR content may require further conversions between the PQ and HLG systems, which is most straightforward to accomplish for signals that do not exceed 1 000 cd/m^2, as discussed in Report ITU-R BT.2390, § 7.

For both tone mapping and inverse tone mapping, methods suitable for this use case are presented in the following sub-sections. The tone mapping and inverse tone mapping methods (§§ 3.1 and 3.2, respectively) are intended to be complementary, allowing a conversion back and forth with only a minimum of quality loss, as evaluated and discussed in Annex 1.

3.1 HDR to SDR conversion

The purpose of this method is to convert HDR content, which is captured or produced at 1 000 cd/m^2, to content suitable for display on an SDR display device with a peak luminance of 100 cd/m^2. The choice of using a 1 000 cd/m^2 input signal is motivated by the bridge point used in the conversion between PQ and HLG signals (and vice-versa), as discussed in Report ITU-R BT.2390-2, § 7.2. The method allows HDR content to be used in SDR BT.2020 productions, or to be distributed in SDR. In addition, the method works in conjunction with the inverse tone mapping method described in § 3.2, allowing round-tripping, the performance of which is documented in § 3.3.

A normalized full-range linear display-light HDR signal RGB using colorimetry specified in Table 2 of Recommendation ITU-R BT.2100, which is assumed to be produced at 1 000 cd/m^2, may be tone mapped to an SDR luminance signal Y_{TMO} according to the specification presented in Table 1.

As human colour perception interacts with brightness, changing the overall level of an image as displayed will cause the perception of colour to be altered. In particular, saturation is affected, for example as described by the Hunt effect. As a consequence, a correction needs to be applied so that colours before and after tone mapping are perceived similarly. Here, colour correction is applied using as input the HDR $R'G'B'$ signal, as well as the luma signals Y' and Y'_{SDR}. As this correction and corresponding reconstruction of a colour signal is carried out on an SDR signal, the output is an SDR signal specified in $Y'C_b'C_r'$ space. It may be displayed directly, or it may be converted to RGB using the information contained in Table 4 of Recommendation ITU-R BT.2020. The colour correction reference is given in Table 2.

TABLE 1

HDR to SDR conversion reference functions[2]

Parameter	Values
Non-linear transfer function	$R' = R^{\frac{1}{2.4}}$ $G' = G^{\frac{1}{2.4}}$ $B' = B^{\frac{1}{2.4}}$
Luma	$Y' = 0.2627R' + 0.6780G' + 0.0593B'$
Tone mapping step 1	$Y'_p = \dfrac{\log(1 + (\rho_{\text{HDR}} - 1)\, Y')}{\log(\rho_{\text{HDR}})}$ $\rho_{\text{HDR}} = 1 + 32\left(\dfrac{L_{\text{HDR}}}{10000}\right)^{\frac{1}{2.4}}$
Tone mapping step 2	$Y'_c = \begin{cases} 1.0770Y'_p & 0 \le Y'_p \le 0.7399 \\ -1.1510{Y'_p}^2 + 2.7811Y'_p - 0.6302 & 0.7399 < Y'_p < 0.9909 \\ 0.5000Y'_p + 0.5000 & 0.9909 \le Y'_p \le 1 \end{cases}$
Tone mapping step 3	$Y'_{\text{SDR}} = \dfrac{\rho_{\text{SDR}}^{Y'_c} - 1}{\rho_{\text{SDR}} - 1}$ $\rho_{\text{SDR}} = 1 + 32\left(\dfrac{L_{\text{SDR}}}{10000}\right)^{\frac{1}{2.4}}$
Peak mastering display luminance	L_{HDR} is the assumed peak mastering display luminance. $L_{\text{HDR}} = 1000\ cd/m^2$
Peak target display luminance	$L_{\text{SDR}} = 100\ cd/m^2$ is the assumed peak target display luminance.

2 Note that the processing described in this table may be implemented using one 1D look-up table.

TABLE 2

Colour correction reference functions

Parameter	Values
Colour difference signals	$C'_{b,\mathrm{TMO}} = f(Y'_{\mathrm{SDR}})\,\dfrac{B'-Y'}{1.8814}$ $C'_{r,\mathrm{TMO}} = f(Y'_{\mathrm{SDR}})\,\dfrac{R'-Y'}{1.4746}$
Colour scaling function	$f(Y'_{\mathrm{SDR}}) = \dfrac{Y'_{\mathrm{SDR}}}{1.1\,Y'}$
Adjusted luma component	$Y'_{\mathrm{TMO}} = Y'_{\mathrm{SDR}} - \max(0.1\,C'_{r,\mathrm{TMO}}, 0)$
Colour space conversion	$Y'_{\mathrm{TMO}}\,C'_{b,\mathrm{TMO}}\,C'_{r,\mathrm{TMO}}$ values may be converted to $R'_{\mathrm{TMO}}\,G'_{\mathrm{TMO}}\,B'_{\mathrm{TMO}}$ according to Recommendation ITU-R BT.2020.

3.2 SDR to HDR conversion

The purpose of the method described in this section is to provide a baseline conversion taking as input SDR signals following BT.2020 colorimetry that are intended to be displayed in a reference viewing environment at 100 cd/m^2. Signals using BT.709 colorimetry may also serve as input, albeit that they should first be mapped or converted to use BT.2020 colorimetry. These signals are then converted to HDR signals suitable for use according to Recommendation ITU-R BT.2100. This method aims to provide visual similarity between input and output content.

A target luminance of 1 000 cd/m^2 is chosen to provide compatibility with the ‘bridge point’ of 1 000 cd/m^2 proposed to guide conversions between PQ and HLG systems (and vice-versa), see § 6 in Report ITU-R BT.2408 as well as § 7.2 of Report ITU-R BT.2390-2 (or later). This enables SDR content to be brought into an HDR workflow, without the need to commit to a specific workflow.

The method is intended for SDR content that does not contain overly large areas near maximum signal level. This method is compatible with the tone mapping method described in § 3.1, in that the two methods together facilitate round-trip performance. In this regard, the round-trip performance of the two methods is assessed in § 3.3.

Starting with an SDR RGB signal encoded in the Recommendation ITU-R BT.2020 colour space, conversion to HDR proceeds by first converting this signal to $Y'C'_bC'_r$, as specified in Table 4 of Recommendation ITU-R BT.2020. Then, the (previously normalised) Y' channel is expanded, and the C'_b and C'_r channels are adjusted to match the visual perception of chromatic content at different luminance levels. The peak luminance $L_{\max}$ is specified to be 1 000 cd/m^2. Note that both input and output signals are assumed to be full range. If required, legal or limited range input signals should be converted to full range. The inverse tone mapping process is described in Table 3.

TABLE 3

SDR to HDR conversion reference functions

Parameter	Values	
Range of input signal	$Y' \in [0,1]$	
Range adjustment	$Y'' = 255.0\,Y'$	
SDR to HDR luma mapping	$Y'_{\mathrm{HDR}} = Y''^{E}$	
Exponent	$E = \begin{cases} a_1 Y''^2 + b_1 Y'' + c_1, & Y'' \le T \\ a_2 Y''^2 + b_2 Y'' + c_2, & Y'' > T \end{cases}$ $T = 70$	
Constants	$a_1 = 1.8712e-5$ $b_1 = -2.7334e-3$ $c_1 = 1.3141$	$a_2 = 2.8305e-6$ $b_2 = -7.4622e-4$ $c_2 = 1.2528$
Chroma mapping	$C'_{b,\mathrm{HDR}} = C'_b\,S_C$ $C'_{r,\mathrm{HDR}} = C'_r\,S_C$	
Chroma scaling factor	$S_C = \begin{cases} 1.075\,\dfrac{Y'_{\mathrm{HDR}}}{Y'} & \text{if } Y' > 0 \\ 1 & \text{if } Y' = 0 \end{cases}$	
Maximum display luminance	$L_{\max} = 1000\ cd/m^2$	
Absolute HDR[3]	$(R'_{\mathrm{abs}}, G'_{\mathrm{abs}}, B'_{\mathrm{abs}}) = f^{-1}(Y'_{\mathrm{HDR}}, C'_{b,\mathrm{HDR}}, C'_{r,\mathrm{HDR}})\,L_{\max}$	
$f^{-1}(Y'_{\mathrm{HDR}}, C'_{b,\mathrm{HDR}}, C'_{r,\mathrm{HDR}})$	$\begin{cases} \left(\dfrac{\mathrm{clamp}(Y'_{\mathrm{HDR}} + 1.4746\,C'_{r,\mathrm{HDR}}, 0, 1000)}{1000}\right)^{2.4} \\ \left(\dfrac{\mathrm{clamp}(Y'_{\mathrm{HDR}} - 0.16455\,C'_{b,\mathrm{HDR}} - 0.57135\,C'_{r,\mathrm{HDR}}, 0, 1000)}{1000}\right)^{2.4} \\ \left(\dfrac{\mathrm{clamp}(Y'_{\mathrm{HDR}} + 1.8814\,C'_{b,\mathrm{HDR}}, 0, 1000)}{1000}\right)^{2.4} \end{cases}$ $\mathrm{clamp}(a,b,c) = \begin{cases} b & \text{if } a < b \\ c & \text{if } a > c \\ a & \text{otherwise} \end{cases}$	

4 Conversion Method B

Notably in live broadcast, but also in other use cases, the quality of captured SDR content may exhibit certain flaws that require mapping to HDR to be tailored to minimise amplification of such flaws. In particular large over-exposed areas, which are common in certain SDR content, may become overly bright under inverse tone mapping, unless the method is tailored to reduce the visual impact of such areas. This section presents a tone mapping / inverse tone mapping pair of methods, whereby the inverse tone mapping is designed to avoid producing objectionably bright large areas. This is achieved with a global curve that limits the peak luminance, relative to the method presented in § 3.

The inverse tone mapping method is presented in § 4.1, whereas the corresponding tone mapping method is discussed in § 4.2.

---

3 The signal produced here ranges between 0 and $L_{\max}$, and can be used for encoding into the PQ system, as defined in Recommendation ITU-R BT.2100. It may subsequently be encoded into the HLG system following the procedure outlined in § 7.2 of Report ITU-R BT.2390.

4.1 SDR to HDR conversion

Where a single non-adaptive method of SDR to HDR conversion is required to work well across a wide range of SDR content, including archive and live material which may contain large areas near maximum signal level, a smaller amount of highlight expansion may be more appropriate.

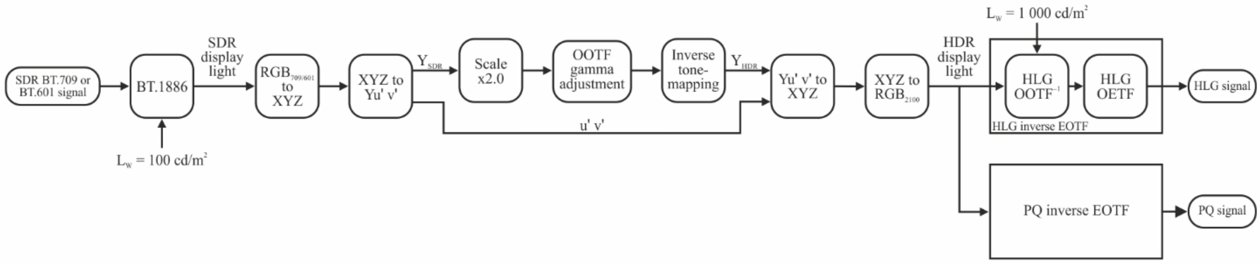
The method described below attempts to reverse the highlight compression commonly applied in SDR cameras using a "knee" function, or that is implicit in SDR graded material. The initial parameters for the inverse tone mapping operator were derived through analysing the distribution of displayed luminance levels across seventy-five SDR programmes, spanning a wide range of genres. By assuming an even distribution of luminance levels within the scene, the typical "knee" breakpoint and compression factor (or equivalent) were deduced by looking for an increase in the distribution of code values. Those initial estimates were then optimised through a series of expert subjective assessments, using a wide range of content that included 8-bit archive material, compressed and interlaced content. Each of these limits the degree of highlight expansion that can be applied.

As the goal of the inverse tone mapping process is to maintain the "look" of the SDR material, a display light conversion is described. A similar scene-light conversion may be used where the goal is to match SDR and HDR cameras. See Report ITU-R BT.2390, § 10.

An overview of the generalised inverse tone mapping process for both PQ and HLG output is illustrated in Fig. 2. The steps of the method are described in the sub-sections that follow.

FIGURE 2

Processing steps for inverse tone mapping method B



Report BT.2446-02

4.1.1 Linearise SDR input to RGB display light

The incoming SDR BT.709 or BT.601 signal is first converted to SDR display light using the Recommendation ITU-R BT.1886 (BT.1886) display EOTF at 100 cd/m^2 peak luminance. The SDR display light is then converted to CIE 1931 XYZ, and subsequently to CIE 1976 $Yu'v'$. The transformation allows the luminance component (Y) to be modified to increase the signal's contrast ratio, without affecting the image chrominance (u' and v').

4.1.2 Gamma compensation and scaling

A scaling factor is applied to the luminance signal to double the SDR displayed light, to be consistent with the HDR Reference White of 203 cd/m^2 (see Report ITU-R BT.2408) for a 1 000 cd/m^2 HLG display. In order to maintain the level of detail in the shadows and the appearance of mid-tones, an OOTF gamma adjustment must also be applied to compensate for the increase in displayed luminance.

The gamma adjustment for SDR content has been found to be quite close to that for HDR content, so the extended range gamma formula in footnote 2 of Recommendation BT.2100-2 can also be adapted to calculate the change in gamma for the scaled SDR. For a doubling of displayed luminance, from 100 to 200 cd/m^2, the formula becomes,

$$\text{Gamma adjustment} = k \wedge (\log_2 (200/100)), \text{ where } k = 1.111$$

Thus, from the formula, a good estimation of the gamma adjustment necessary is approximately 1.1. Slightly different values might be used in practice, to achieve the desired artistic effect.

4.1.3 Inverse tone mapping

In its most basic form, all pixel values above a specified breakpoint are increased by a multiplier, in effect expanding the highlights to more natural levels in relation to the shadows and mid-tones. In order to meet the criteria of being simple and robust, a fixed breakpoint is used, rather than a variable one that is dependent on the content. A fixed breakpoint also makes the process easier to reverse (for round-tripping) as no metadata is required. Subjective tests found that the exact breakpoint is less critical than the scaling multiplier. Therefore, a value of ~80% SDR signal (equivalent to ~60 cd/m^2 SDR displayed light on a 100 cd/m^2 BT.1886 display) is recommended as that is very close to the average breakpoint found in the analysis of programme content.

Camera knees can compress the highlights significantly, so a simple reversal may not be appropriate. In choosing the multiplier one must be careful to ensure that the image is natural in appearance, without adding any banding in the highlights introduced during the SDR production process. Initial testing showed that a value of 2.3 (display light) gave a subtle increase in highlights without causing banding. It also ensured that large over-exposed and clipped areas of the SDR image were not reproduced at too high a luminance level on HDR screens. This value was found to work well for both 8-bit and 10-bit content as even 8-bit video has fine enough quantisation in the upper signal ranges to allow some expansion of highlights.

In practice, to avoid artefacts at the intersection of the unity scaler and highlight expansion function, a Bezier curve or similar can be used to blend between the two.

4.1.4 Convert to RGB and apply the inverse EOTF

The modified display luminance and *u'v'* chrominance signals are then converted back to *XYZ*, to allow conversion to *RGB* displayed light with Recommendation ITU-R BT.2100 colour primaries. The HLG signal is obtained by applying the inverse HLG EOTF at 1 000 cd/m^2 peak luminance. A PQ signal is obtained by applying the PQ inverse EOTF.

4.1.5 Simplification for HLG output

Where only an HLG output signal is required, the inverse tone mapping process can be simplified by exploiting the ability of the HLG EOTF to operate at different peak luminance levels. Through careful choice of L_W, the HLG EOTF nominal peak luminance, a 1:1 mapping of SDR luminance to HDR luminance can be achieved over a large portion of the signal range, thereby avoiding the need for any luminance scaling and compensating OOTF adjustment. The simplified processing steps are shown in Fig. 3.

For graded SDR content, for which display light conversions are most often used, "diffuse white" (equivalent to the "HDR Reference White" defined in Report ITU-R BT.2408) is usually set at around 90% signal level. So the value of L_W used for the HLG inverse EOTF is chosen such that a 90% SDR achromatic signal (with no inverse tone mapping) would deliver the "HDR Reference White" of 75% signal for HLG.

Using this formula, the 1:1 mapping of SDR to HDR luminance is achieved by setting the SDR BT.1886 nominal peak luminance L_W, to 100 cd/m^2, the HLG nominal peak luminance, L_W to 291 cd/m^2 with a HLG gamma value of 0.99[4]. Once inverse tone mapping has been applied, 90% SDR signal is mapped to 77% HLG and the image brightness slightly increased. So marginally better

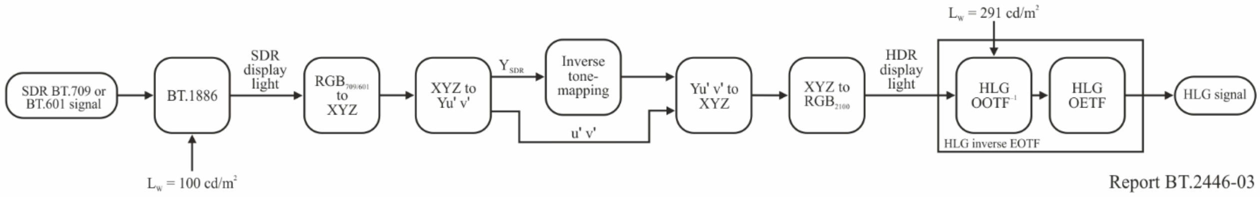
[4] Calculated using the extended range gamma formula in footnote 2 of Recommendation ITU-R BT.2100.

results are achieved with this inverse tone mapper by using a slightly higher gamma value of 1.03, thereby compensating for the increase in average luminance as a result of the inverse tone mapping process.

The simplified inverse tone mapping process also makes it easier to calculate a complementary tone mapper, to minimise round-tripping losses.

FIGURE 3

Simplified inverse tone mapping for HLG output



Report BT.2446-03

A suitable inverse tone mapping curve is illustrated in Fig. 4. The function provides a 1:1 mapping of SDR display light to HDR display light up to the breakpoint, ensuring that shadow details and mid-tones are faithfully reproduced in the HDR image. In this example, the breakpoint is set to 78% SDR signal (55 cd/m^2) as it closely matches the average breakpoint of SDR cameras and has been found to offer a good balance between image brightness, and the portrayal of highlights in the complementary tone mapper described in § 4.2.

Beyond the breakpoint an exponential function, rather than the basic linear function and Bezier curve (described in § 4.1.3) is used to expand the highlights.

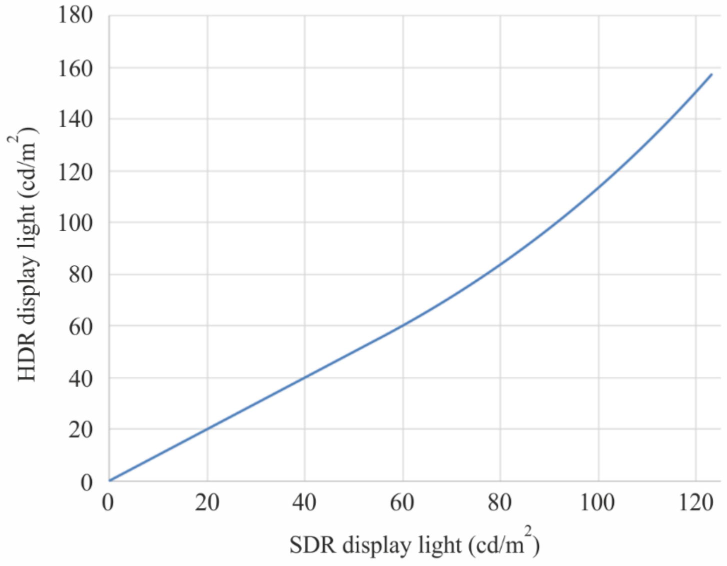
The curve is extended into the SDR "super-white" region (up to the maximum of 123 cd/m^2) to exploit any SDR highlights which may be present in the signal.

The gradient of the exponential function is set to unity at the breakpoint to ensure a smooth transition between the linear and exponential portions. Other parameters are adjusted to ensure that a 100% achromatic SDR signal delivers ~83% HLG signal (114 cd/m^2 for HLG L_W = 291 cd/m^2 with the increased gamma of 1.03).

Other functions are possible, provided a smooth transition is ensured at the breakpoint and the amount of highlight expansion is limited to similar levels.

FIGURE 4

Example inverse luminance tone mapping curve



Report BT.2446-04

4.2 HDR to SDR conversion

An alternative method of tone mapping is described below, which has been designed to complement the inverse tone mapping algorithm described in § 4.1, whilst also delivering a high quality SDR BT.709 image from natively produced BT.2100 HDR content. To achieve these goals, the method:

– is display-referred, to ensure the HDR and SDR images have a similar "look";

– uses a tone-mapping curve that, over a large part of the signal range, is the complement of the inverse tone-mapping curve described in § 4.1.5, thereby reducing round-tripping losses;

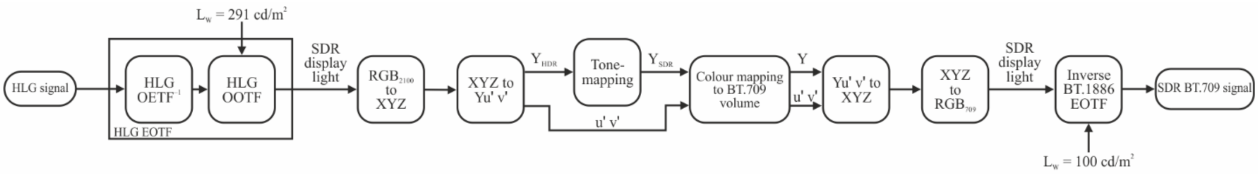
– maps the "HDR Reference White" level (75% HLG or 58% PQ) to approximately 90% of the SDR signal, to ensure consistent image brightness between tone-mapped and natively produced SDR content, whilst also allowing some headroom in the SDR signal for compressed highlights;

– provides a hard clip of out-of-gamut colours to BT.709 colour volume to minimise "round-tripping" chrominance losses.

The processing blocks for HLG signals, are illustrated in Fig. 5. As the method exploits the ability to configure the HLG EOTF to different peak luminance levels, PQ signals should first be converted to HLG using one of the methods specified in Report ITU-R BT.2390, § 7.

FIGURE 5

HDR to SDR tone-mapping method B



Report BT.2446-02

The algorithm is described in the following sub-sections.

4.2.1 Conversion to HDR linear RGB display light

The HLG display EOTF (comprising an inverse HLG Opto-Electronic Transfer Function ($OETF^{-1}$) and the Opto-Optical Transfer Function (OOTF)) is applied to the input signal to create linear RGB display light signals.

In graded SDR content, diffuse white is often set at around 90% of the SDR signal, which allows a small amount of "headroom" for highlights. So, to ensure similar brightness between tone-mapped HDR and native SDR content, the tone-mapping algorithm should also map "HDR Reference White" (75% HLG) to around 90% of the SDR signal.

Rather than using the usual bridge condition of 1 000 cd/m^2 for format conversions, the HLG display EOTF is conveniently set to a nominal peak luminance, L_W, of 291 cd/m^2, as that renders "HDR Reference White" at 78 cd/m^2, which is equivalent to 90% of the SDR signal on a BT.1886 reference display with nominal peak luminance of 100 cd/m^2. By doing so, the brightness of the tone mapped images are comparable to SDR graded content, and a direct 1:1 mapping between HDR displayed light and SDR displayed light can be used for the conversion, over most of the signal range.

The exact HLG gamma display gamma may be adjusted slightly from the value specified for HDR images in Recommendation ITU-R BT.2100, to achieve the desired artistic effect. But a gamma value of 1.03 for an EOTF L_W of 291 cd/m^2 has been found to work well and complements the inverse tone mapper described in § 4.1.5.

4.2.2 Convert to *Yu'v'*

To ensure that the hue and saturation of the signals are unaffected by the highlight compression, the tone mapping is only applied to the luminance component (*Y*), and the chrominance components (*u'* and *v'*) bypass the process and remain unaltered. The conversion comprises two steps:

1) RGB with BT.2100 primaries to CIE 1931 XYZ
2) CIE 1931 XYZ to CIE 1976 *Yu'v'*

4.2.3 Tone Mapping

A 1:1 mapping between HDR and SDR display light is used over most of the signal range, thereby ensuring the consistent portrayal of shadow detail and mid-tones between the HDR and SDR images. A compression function is then applied to the higher luminance level signals.

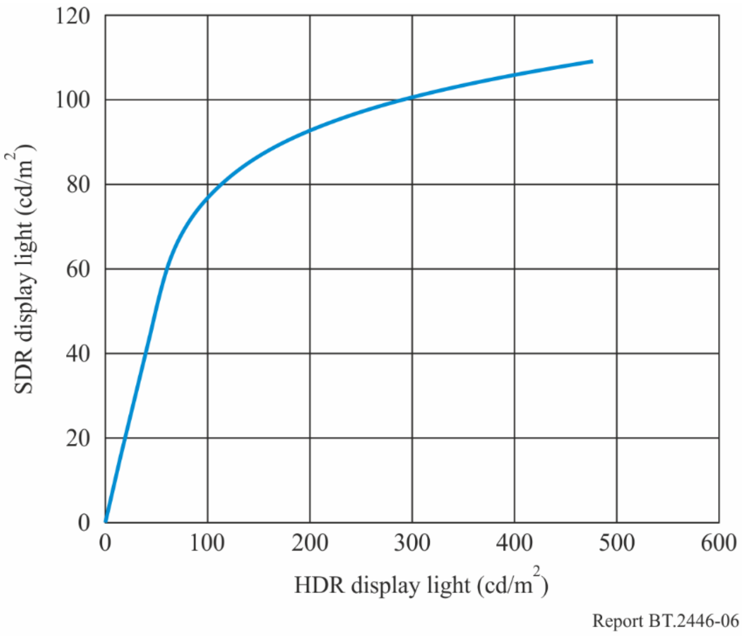
To both complement the inverse tone mapping described in § 4.1, and provide a good balance between the brightness of the SDR signal and the portrayal of highlights, the breakpoint between the linear 1:1 mapping and the compression function is set to 78% of the SDR signal level (~55 cd/m^2). The gradient of the compression portion of the curve should be unity at the breakpoint, in order to match the gradient of the lower linear portion and minimise banding artefacts.

In this example, luminance highlights are compressed using a logarithmic curve, similar to the inverse of the exponential curve used for the inverse tone mapper described in § 4.1.5. Following tone mapping, HDR Reference White (75% HLG) is mapped to 86% SDR signal.

The example luminance tone mapping curve is shown in Fig. 6.

FIGURE 6

Example luminance tone-mapping curve



4.2.4 Colour volume reduction

Following the highlight compression, some colours in the video signal will be outside the target BT.709 colour volume. These colours must be "reduced" so that they lie on or within the colour volume boundary. Careful hard clipping of out-of-gamut colours to the target colour volume surface improves the round-tripping performance when used in conjunction with the inverse tone-mapping algorithm described in § 4.1. It was also found to work well when developing the more sophisticated conversion techniques reported in Report ITU-R BT.2407, Colour gamut conversion from Recommendation ITU-R BT.2020 to Recommendation ITU-R BT.709.

The hard clipping is achieved by desaturating out-of-gamut colours so that they lie on the BT.709 colour volume surface. Care should be taken when choosing the working colour space, to avoid hue shifts.

Tests have shown that using CIE 1976 *Yu'v'* causes saturated reddish-orange colours, to desaturate towards pink. Similarly, saturated yellowish-green changes hue towards green. A small hue shift also occurred in blue.

Further tests performed using CIE *L*a*b** found that there were significant hue shifts from blue to purple in saturated water and sky scenes. A small hue shift also occurred in the reddish-orange colours of saturated sunset scenes.

No hue shifts were seen when desaturating colours using the IC_TC_P colour representation, standardised in Recommendation ITU-R BT.2100.

4.2.5 Convert to RGB and apply the inverse BT.1886 EOTF

An SDR non-linear signal is created by first converting CIE 1976 *Yu'v'* to *XYZ*, then converting *XYZ* to *RGB* with BT.709 colour and applying the inverse SDR EOTF described in Recommendation ITU-R BT.1886. The resulting output should be a perceptually close match (without the high contrast and brightness perceived in HDR) to the original HDR video without hue shifts.

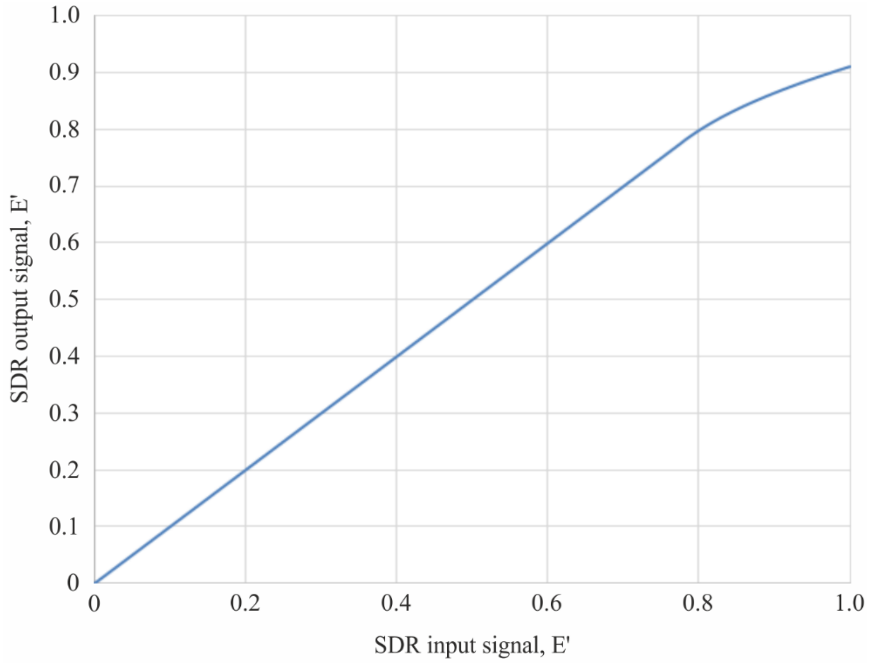
4.3 Round-tripping

The tone and inverse tone mappers described in §§ 4.1 and 4.2 are not the exact inverse of one another. When tone mapping, SDR images benefit from exploiting the full HDR signal range, to avoid hard clipping in the converted signal. The inverse tone mapper, however, only extends to around 83% of the HLG signal (for SDR nominal peak white signals) to avoid large overly bright clipped areas in the HDR image. Through careful design of the tone and inverse tone mappers it is, however, still possible to reduce round-tripping losses to acceptable levels.

Figure 7 shows the round-tripping signal losses for the cascaded SDR to HDR and HDR to SDR conversions described in § 4, for achromatic signals. Figure 8 shows the same signal losses, as they would appear if shown on a 100 cd/m^2 reference display. As the tone mapping and inverse tone mapping curves follow one another precisely up to the breakpoint (78% SDR signal), excellent round-tripping performance is achieved for the majority of the signal range, which includes the ranges commonly used for SDR skin tones. The highlight expansion and compression functions then slowly begin to diverge above the breakpoint, and some small losses can be seen thereafter. The peak luminance of signals following the double conversion is slightly reduced.

FIGURE 7

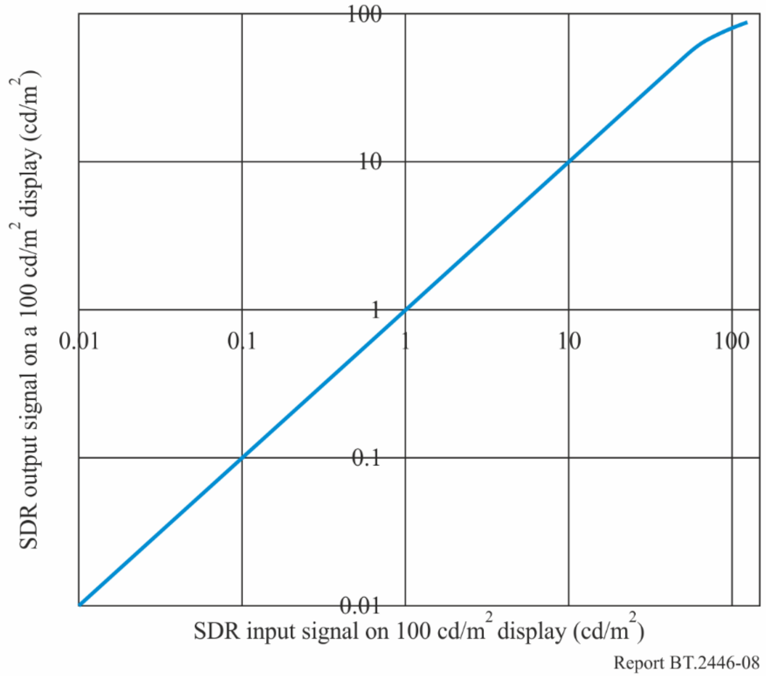
SDR-HDR-SDR round-tripping signal losses



Report BT.2446-07

FIGURE 8

SDR-HDR-SDR round-tripping signal losses, shown on a 100 cd/m² reference display



5 Conversion Method C

This conversion method is designed as a parametric representation to ensure that the optimum conversion can be adjusted depending on the content. In designing the conversion, the following considerations were made:

– A tone-mapping function is composed of a linear mapping function and a log function. Below highlight parts, HDR signals are linearly mapped to SDR signals to preserve the look of mid-tone. The gain of the linear mapping is determined from the relationship between the skin tones in the HDR and SDR content. The highlight parts of the HDR signals are compressed using the log function to fit within the SDR range while retaining as much gradation as possible.

– HDR Reference White (75%HLG) should be mapped to approximately 95% SDR signal level to ensure consistency with SDR original content, assuming that the peak luminance of HDR and SDR displays are 1 000 cd/m^2 and 100 cd/m^2, respectively.

– The hues of highlights in the original HDR content should be maximally retained after tone-mapping. To avoid hue shifts caused by the clipping of compressed highlight parts, pre- and post-processing are applied, in which a crosstalk matrix reduces saturation before tone-mapping and an inverse crosstalk matrix recovers the saturation after tone-mapping.

– In SDR production, highlight parts are sometimes intentionally expressed as achromatic. Thus, an optional processing step is included to change highlight parts to achromatic.

– Display-referred mapping is adopted so that the converted SDR content looks similar (except in dynamic range) to the original HDR content.

– Inverse conversion can be used for SDR to HDR conversion.

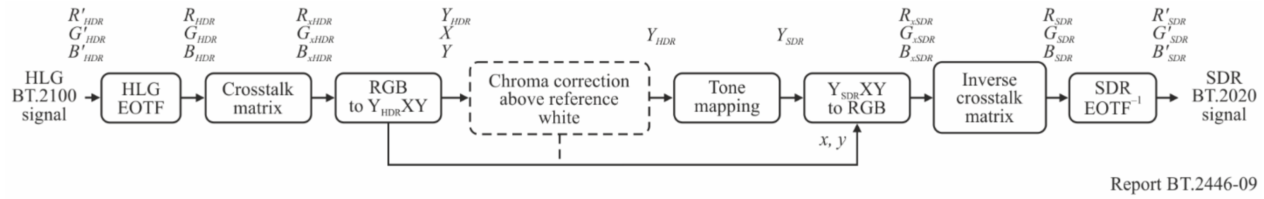
5.1 HDR to SDR conversion

A block diagram for HDR to SDR conversion is illustrated in Fig. 9. The conversion method assumes HLG input signals, however, similar conversion might be possible for PQ input signals. When colour conversion from BT.2020 to BT.709 is required, one of the methods described in Report ITU-R BT.2407 may be used after HDR to SDR conversion.

The details of the algorithms used are described in the following sub-sections. The processing may be implemented using one 3D look-up table.

FIGURE 9

Block diagram of HDR to SDR conversion



5.1.1 Conversion to linear display light signals

The input HLG video signals are converted to linear display light signals by applying the HLG EOTF as specified in Table 5 of Recommendation ITU-R BT.2100, where the system gamma of 1.2 may be used.

$$\begin{bmatrix} R_{HDR} \\ G_{HDR} \\ B_{HDR} \end{bmatrix} = \mathrm{EOTF_{HLG}} \begin{bmatrix} R'_{HDR} \\ G'_{HDR} \\ B'_{HDR} \end{bmatrix} \tag{1}$$

5.1.2 Crosstalk matrix

The crosstalk matrix is applied such that saturations of linear signals are reduced to achromatic to avoid hue changes caused by clipping of compressed highlight parts.

$$\begin{bmatrix} R_{xHDR} \\ G_{xHDR} \\ B_{xHDR} \end{bmatrix} = \begin{bmatrix} 1-2\alpha & \alpha & \alpha \\ \alpha & 1-2\alpha & \alpha \\ \alpha & \alpha & 1-2\alpha \end{bmatrix} \cdot \begin{bmatrix} R_{HDR} \\ G_{HDR} \\ B_{HDR} \end{bmatrix}, \tag{2}$$

where α determines the degree of de-saturation and should be set to a value in the range of $0 \leq \alpha \leq 0.33$.

5.1.3 Conversion to Yxy

The linear RGB signals after the crosstalk matrix are converted to CIE 1931 XYZ signals, and subsequently to Yxy signals. This conversion makes the luminance signal separated from the colour components, and the successive tone mapping is applied only to the luminance signal.

$$\begin{bmatrix} X_{HDR} \\ Y_{HDR} \\ Z_{HDR} \end{bmatrix} = \begin{bmatrix} 0.6370 & 0.1446 & 0.1689 \\ 0.2627 & 0.6780 & 0.0593 \\ 0.0000 & 0.0281 & 1.0610 \end{bmatrix} \cdot \begin{bmatrix} R_{xHDR} \\ G_{xHDR} \\ B_{xHDR} \end{bmatrix} \tag{3}$$

$$x = \frac{X_{HDR}}{X_{HDR}+Y_{HDR}+Z_{HDR}}, \quad y = \frac{Y_{HDR}}{X_{HDR}+Y_{HDR}+Z_{HDR}} \tag{4}$$

5.1.4 Tone mapping

The tone-mapping function is composed of a linear mapping function below highlight parts and a log function for the highlight parts:

$$Y_{SDR} = \begin{cases} k_1 \cdot Y_{HDR} & Y_{HDR} < Y_{HDR,ip} \\ k_2 \cdot \ln\left(\frac{Y_{HDR}}{Y_{HDR,ip}} - k_3\right) + k_4 & Y_{HDR} \geq Y_{HDR,ip} \end{cases} \tag{5}$$

where parameters k_1 to k_4 determine the tone mapping characteristics and $Y_{HDR,ip}$ is the inflection point corresponding to the knee point of SDR. The tone-mapping function can be uniquely specified by determining the gain of the linear mapping, the inflection point at which the function is changed from the linear mapping function to the log function, and the SDR white level corresponding to the HDR Reference White.

The parameter k_1 is determined by the gain of linear mapping below the inflection point, for example, from the relationship between the skin tones in SDR and HDR content.

The inflection point $Y_{HDR,ip}$ corresponds to the knee point of SDR. Because the highlight part above the inflection point is compressed using the log function, the inflection point should be higher than the skin tone to maintain fresh tones. The inflection point may be set to 80% SDR; that is, equivalent to 58.5 cd/m^2 on a 100 cd/m^2 SDR display, which has been found to work well. The inflection point $Y_{HDR,ip}$ is calculated as follows:

$$Y_{HDR,ip} = 58.5/k_1 \tag{6}$$

The parameters k_2 and k_4 are determined by the continuous condition at the inflection point, which is the continuous condition of the value and the first derivation, using the following equations:

$$k_2 = k_1(1 - k_3) \cdot Y_{HDR,ip} \tag{7}$$

$$k_4 = k_1 \cdot Y_{HDR,ip} - k_2 \cdot \ln(1 - k_3) \tag{8}$$

The parameter k_3 is determined by the condition that the increment of SDR linear signal from the inflection point to the white point corresponds to that of the log function.

$$Y_{SDR,wp} - k_1 \cdot Y_{HDR,ip} = \left\{k_2 \cdot \ln\left(\frac{Y_{HLG,Ref}}{Y_{HDR,ip}} - k_3\right) + k_4\right\} - \left\{k_2 \cdot \ln\left(\frac{Y_{HDR,ip}}{Y_{HDR,ip}} - k_3\right) + k_4\right\} \tag{9}$$

where $Y_{HLG,Ref}$ indicates the HDR linear signal corresponding to the HDR Reference White (75%HLG) and $Y_{SDR,wp}$ indicates the SDR linear signal for the white point corresponding to the HDR Reference White.

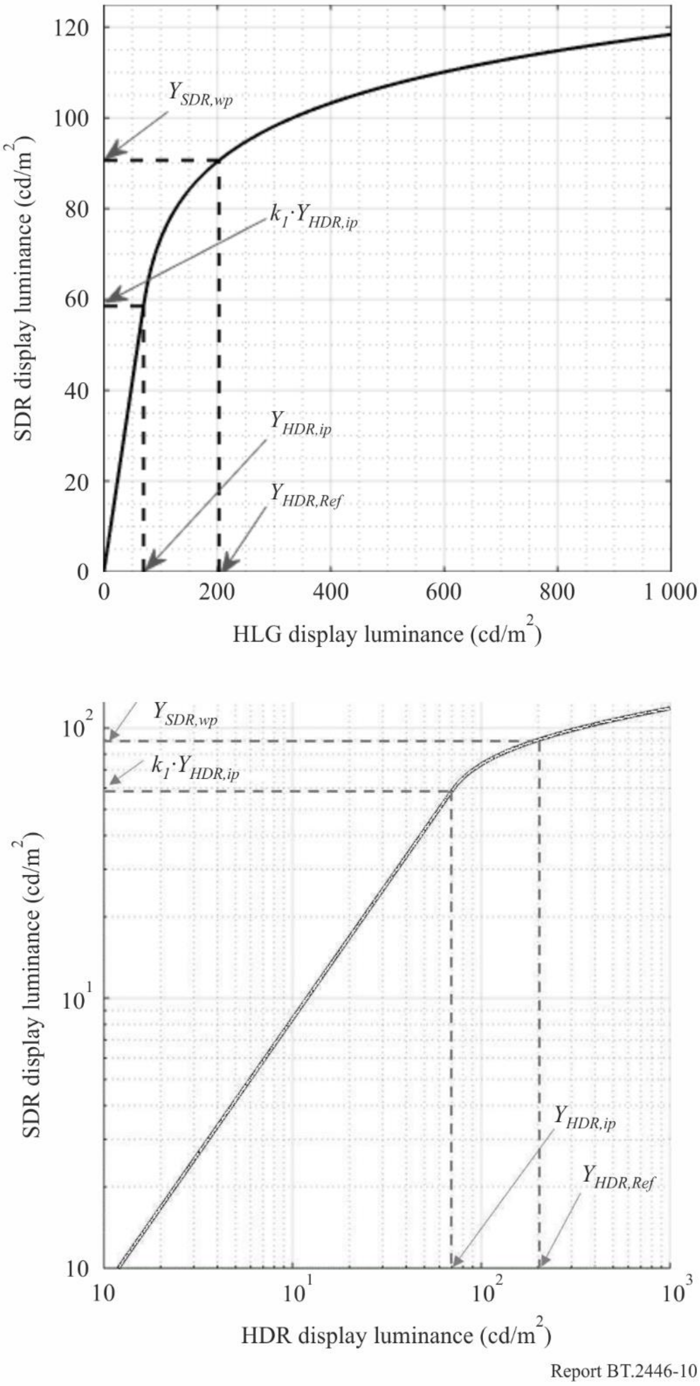
In the condition that the gain of linear mapping is determined by the relationship of HDR and SDR skin tones described in Annex 4 of Report ITU-R BT.2408, which is 50% HLG and 70% SDR, the inflection point is set to 80% SDR, and the SDR white level corresponding to the HDR Reference White is set to 96% SDR, as shown in Fig. 10, the following parameter values can be derived:

$$k_1 = 0.83802, k_2 = 15.09968, k_3 = 0.74204, k_4 = 78.99439 \tag{10}$$

By setting a different condition on the relationship of HDR and SDR levels according to different production intent, a different set of parameter values for k_1 to k_4 can be derived.

FIGURE 10

Tone-mapping function in luminance domain



5.1.5 Conversion to RGB linear signal

The tone-mapped SDR luminance signal Y_{SDR} is converted to RGB linear signals using the xy values calculated in § 5.1.3.

$$X_{SDR} = \left(\frac{x}{y}\right) \cdot Y_{SDR} \tag{11}$$

$$Z_{SDR} = \left(\frac{1-x-y}{y}\right) \cdot Y_{SDR} \tag{12}$$

$$\begin{bmatrix} R_{xSDR} \\ G_{xSDR} \\ B_{xSDR} \end{bmatrix} = \begin{bmatrix} 1.7167 & -0.3557 & -0.2534 \\ -0.6667 & 1.6165 & 0.0158 \\ 0.0176 & -0.0428 & 0.9421 \end{bmatrix} \cdot \begin{bmatrix} X_{SDR} \\ Y_{SDR} \\ Z_{SDR} \end{bmatrix} \tag{13}$$

5.1.6 Inverse crosstalk matrix

The inverse crosstalk matrix is applied to ensure that the original hues of input HDR images are recovered.

$$\begin{bmatrix} R_{SDR} \\ G_{SDR} \\ B_{SDR} \end{bmatrix} = \begin{bmatrix} 1-2\alpha & \alpha & \alpha \\ \alpha & 1-2\alpha & \alpha \\ \alpha & \alpha & 1-2\alpha \end{bmatrix}^{-1} \cdot \begin{bmatrix} R_{xSDR} \\ G_{xSDR} \\ B_{xSDR} \end{bmatrix}$$

$$= \frac{1}{1-3\alpha} \begin{bmatrix} 1-\alpha & -\alpha & -\alpha \\ -\alpha & 1-\alpha & -\alpha \\ -\alpha & -\alpha & 1-\alpha \end{bmatrix} \cdot \begin{bmatrix} R_{xSDR} \\ G_{xSDR} \\ B_{xSDR} \end{bmatrix} \tag{14}$$

5.1.7 Inverse SDR EOTF

The output SDR video signals are obtained by applying the inverse SDR EOTF, as described in Recommendation ITU-R BT.1886.

$$\begin{bmatrix} R'_{SDR} \\ G'_{SDR} \\ B'_{SDR} \end{bmatrix} = EOTF_{BT.1886}{}^{-1} \begin{bmatrix} R_{SDR} \\ G_{SDR} \\ B_{SDR} \end{bmatrix} \tag{15}$$

5.1.8 Optional processing of chroma correction above HDR Reference White

In SDR production, highlight parts are sometimes intentionally expressed as white. The processing described in this section is optionally used to shift chroma above HDR Reference White to achromatic when the converted SDR content requires a degree of consistency for SDR production content. This processing is applied as needed before the tone-mapping processing.

(1) Conversion of Yxy to CIE *L*a*b**

The linear Yxy signals derived from the HDR input are converted into CIE $L^*a^*b^*$ signals. In this conversion, the XYZ tristimulus values of the reference white point are defined by the HDR Reference White, that is $[X_n, Y_n, Z_n] = [192.93, 203, 221.05]$.

$$L^* = 116 \cdot f(Y_{HDR}/Y_n) - 16 \tag{16}$$

$$a^* = 500 \cdot [f(X_{HDR}/X_n) - f(Y_{HDR}/Y_n)] \tag{17}$$

$$b^* = 200 \cdot [f(Y_{HDR}/Y_n) - f(Z_{HDR}/Z_n)] \tag{18}$$

where:

$$X_{HDR} = \left(\frac{x}{y}\right) \cdot Y_{HDR} \tag{19}$$

$$Z_{HDR} = \left(\frac{1-x-y}{y}\right) \cdot Y_{HDR} \tag{20}$$

$$f(t) = \begin{cases} t^{\frac{1}{3}} & t > \delta^3 \\ \left(\frac{1}{\delta}\right)^3 \cdot t & otherwise \end{cases} \tag{21}$$

$$\delta = 6/29 \tag{22}$$

(2) Chroma correction above HDR Reference White

Chroma correction is applied only on the chroma above HDR Reference White. The chroma C^*_{ab} and hue h_{ab} are derived from the values of $a*$ and $b*$, and the correction value of chroma $C^*_{ab,cor}$ is calculated using the correction factor f_{cor}.

$$C^*_{ab} = \sqrt{a^{*2} + b^{*2}} \tag{23}$$

$$h_{ab} = \tan^{-1}(b^*/a^*) \tag{24}$$

$$C^*_{ab,cor} = f_{cor} \cdot C^*_{ab} \tag{25}$$

$$f_{cor} = \begin{cases} 1 & L^* \leq L^*_{Ref} \\ \left(1 - \sigma \cdot \frac{L^* - L^*_{Ref}}{L^*_{max} - L^*_{Ref}}\right) & L^* > L^*_{Ref} \end{cases} \tag{26}$$

where L^*_{Ref} is the lightness corresponding to the HDR Reference White, L^*_{max} is the lightness corresponding to the HDR peak level, and σ is a user parameter that determines the degree of white shift. When $\sigma = 1$, the corrected chroma value corresponding to L^*_{max} is to be 0, which means that the colour of the HDR peak level has completely shifted to white. When $f_{cor} < 0$, f_{cor} is set to 0.

Using the above corrected value of $C^*_{ab,cor}$, the correction values of a^*_{cor} and b^*_{cor} are calculated as follows:

$$\begin{pmatrix} a^*_{cor} \\ b^*_{cor} \end{pmatrix} = C^*_{ab,cor} \cdot \begin{pmatrix} \cos h_{ab} \\ \sin h_{ab} \end{pmatrix} \tag{27}$$

(3) Conversion of CIE *L*a*b to Yxy**

The correction values of $L^*a^*b^*$ are converted back to linear Yxy signals. When converting to RGB linear signals as described in § 5.1.5, the values of x and y are replaced with the values calculated in this section.

$$Y_{HDR} = \begin{cases} Y_n \cdot f_y^3 & f_y > \delta \\ \left(f_y - \frac{16}{116}\right) \cdot 3\delta^2 Y_n & otherwise \end{cases} \tag{28}$$

$$X_{HDR} = \begin{cases} X_n \cdot f_x^3 & f_x > \delta \\ \left(f_x - \frac{16}{116}\right) \cdot 3\delta^2 X_n & otherwise \end{cases} \tag{29}$$

$$Z_{HDR} = \begin{cases} Z_n \cdot f_z^3 & f_z > \delta \\ \left(f_z - \frac{16}{116}\right) \cdot 3\delta^2 Z_n & otherwise \end{cases} \tag{30}$$

where:

$$f_y = \frac{L^* + 16}{116} \tag{31}$$

$$f_x = f_y + \frac{a^*{}_{cor}}{500} \tag{32}$$

$$f_z = f_y - \frac{b^*{}_{cor}}{200} \tag{33}$$

x and *y* are then derived as follows:

$$x = \frac{X_{HDR}}{X_{HDR} + Y_{HDR} + Z_{HDR}} \tag{34}$$

$$y = \frac{Y_{HDR}}{X_{HDR} + Y_{HDR} + Z_{HDR}} \tag{35}$$

5.2 SDR to HDR conversion

The conversion from SDR to HDR may be implemented as the inverse conversion described in § 5.1. Under the assumption that the optional processing described in § 5.1.8 is not included in the HDR to SDR conversion, the SDR to HDR conversion will be the mathematical inverse and therefore the round-trip conversion will be exact.

6 Feature-based comparison

Methods A, B and C for tone mapping and inverse tone mapping, as discussed in §§ 3 to 5, exhibit different features that would each given them a different performance, which may benefit different use cases. To facilitate the comparison between them, Table 4 provides a brief comparison to highlight the main features of each method.

TABLE 4

Features of the methods presented in this Report

TMO/ITMO Feature	**A**	**B**	**C**
Intended primarily for	Graded content	Live broadcast; SDR content with possibly clipped areas	Live broadcast
Curve	Fixed	Fixed	Fixed (per-session adjustable parameters)
Round-trip performance	Good (see Annex A1)	Good	Mathematical inverse
Colorimetry	BT.2020/2100	BT.2020/2100 / BT.709	BT.2020/2100
TMO Feature	**A**	**B**	**C**
Conversion peak luminance	1 000 cd/m^2	291 cd/m^2	1 000 cd/m^2
Output peak luminance	100 cd/m^2	100 cd/m^2	120 cd/m^2
SDR signal range	0 – 100%	0 – 100%	0 – 109%
Perceptual colour management	Yes, in YC_bC_r	No	Yes, in Yxy and optionally in CIE $L^*a^*b^*$
Colour volume management	No	Yes (IC_TC_P)	No
Key features	Produce visual match with HDR input	Produce a visual match with graded SDR content	Produce a visual match with graded SDR content
ITMO Feature	**A**	**B**	**C**
Input peak luminance	100 cd/m^2	120 cd/m^2	120 cd/m^2
Conversion peak luminance	1 000 cd/m^2	291 cd/m^2	1 000 cd/m^2

TABLE 4 (*end*)

TMO/ITMO Feature	A	B	C
Input signal range	0-100%	−7 – 109%	0 – 109%
Related to TMO	Approximate inverse of TMO	Approximate inverse of TMO	Mathematical inverse of TMO
Perceptual colour management	Yes, in YC_bC_r	No	Yes, in *Yxy*
Colour volume management	No	No	No
Key features	Black level management, mid-tone preservation, visual match with SDR input	Robust to SDR content with large clipped areas	Robust to round-tripping

Annex 1

Evaluation of Round-Trip Performance

A1 Round-trip performance of Method A

This section reports on two psychophysical experiments that were designed to understand if observers would be capable to observe degradation of the images under tone mapping followed by inverse tone mapping, and under inverse tone mapping followed by tone mapping using the methods described in §§ 3.1 and 3.2[5].

A1.1 Database of HDR images

A set of 115 exposure stacks sized 1920 by 1080 were obtained with a variety of cameras (Minolta, Nikon D2h, Nikon D7100), and combined into HDRs with standard multiple exposure techniques, as described below. Some exposure stacks were captured manually using a tripod, whereas most exposure stacks were captured using auto-bracketing on tripod or hand-held. Auto-bracketing for the images taken with the Nikon D2h involved 9 exposures taken one f/stop apart. A small selection of images taken with the Nikon D7100 were captured with 3 exposures spaced 2 f/stops apart. The Minolta camera is older than the other two cameras and did not include viable auto-bracketing technology. Thus, each scene captured with this camera involved a variable number of exposures. The exposures were saved as JPG images. Some more recent captures were saved in RAW format.

The exposures were captured in the Adobe RGB colour space or converted to this colour space in post-processing. The exposures were merged into HDR images using either Greg Ward's Photosphere

[5] Note that the inverse tone mapping algorithm presented in § 3.2 is a slightly updated version from the one evaluated in this section. The parameter values used in this section are: $T = 155$, $a_1 = 6.3170e - 6$, $b_1 = -1.6080e - 3$, $c_1 = 1.2996$, $a_2 = -6.8603e - 7$, $b_2 = 6.6047e - 4$ and $c_2 = 1.1162$.

Software or Adobe Photoshop. Merging a stack of exposures into a single HDR image is achieved by the following sequence of steps:

1) linearizing each exposure to compensate for the camera response function (for JPG captures);
2) weighting each pixel such that over- or under-exposed pixels are given a lower weight;
3) dividing each pixel in each exposure by its associated exposure time;
4) summing corresponding pixels from each exposure;
5) saving the resulting image in an HDR image format.

The HDR image format used is commonly known as the ".hdr" format, and is documented in the Radiance lighting simulation package [11].

Acquisition took place in a variety of different countries (USA, Bahamas, UK, France, Spain, Germany, the Netherlands, Greece, Cyprus, Northern Cyprus) over a period of 15 years.

The images were regraded at 1 000 cd/m^2 using a Sony BVM-X300 professional mastering display. They were also stored in a BT.2020 container using the methodology described in Recommendation ITU-R BT.2087, albeit that an appropriate matrix M for converting between Adobe RGB and XYZ colour spaces was used, namely:

$$M = \begin{bmatrix} 0.5767309 & 0.1855540 & 0.1881852 \\ 0.2973769 & 0.6273491 & 0.0752741 \\ 0.0270343 & 0.0706872 & 0.9911085 \end{bmatrix}$$

The images were then stored in EXR format, forming the basis for all subsequent experimentation. The dataset is further described in Report ITU-R BT.2245-4.

A1.2 Stimuli

To produce stimuli, all images I_{HDR} in the HDR database described above were first tone mapped from 1 000 cd/m^2 to 100 cd/m^2 using the tone reproduction method described in § 3.1. The resulting images are referred to as I_{TM} The images I_{TM} were then inverse tone mapped using the method described in § 3.2, producing image set I_{TM-ITM}. Finally, a second round of tone mapping was applied to I_{TM-ITM}, yielding $I_{TM-ITM-TM}$.

These four image sets allow various comparisons to be made. First, the image sets I_{HDR} and I_{TM-ITM} have the same high dynamic range, and they could therefore be directly compared for degradation of the pathway between HDR to SDR and back to HDR. In the following, it will be referred to as the HDR comparison.

Second, the image sets I_{TM} and $I_{TM-ITM-TM}$ can be compared to understand whether the pathway from SDR to HDR and back to SDR produces a visually degraded result. In the following, it will be referred to as the SDR comparison.

From each image only the middle 1920/2 by 1080 pixels were used, so that two images can be shown side-by-side on the same monitor. Using only one monitor rules out any possible biases that may be introduced due to calibration mismatches if multi-display set-ups were used. Thus, stimuli were produced by placing side-by side the following images:

1 HDR Comparison: Left: I_{HDR}, Right: I_{TM-ITM}.
2 HDR Comparison: Left: I_{TM-ITM}, Right: I_{HDR}.
3 SDR Comparison: Left: I_{TM}, Right: $I_{TM-ITM-TM}$.
4 SDR Comparison: Left: $I_{TM-ITM-TM}$, Right: I_{TM}.

To each stimulus a three-pixel wide vertical bar at 0 cd/m^2 was added, separating the left and right halves to avoid comparisons being made on the border between the images. For the same reason, the images were not butterflied to ensure that comparisons were made over the entire surface of each image.

A1.3 Experimental set-up

Participants were seated 3.5 screen heights away from a Sony BVM-X300 grading monitor, and directly in front of a laptop (Dell Precision 7710) which was used to display instructions as well as to record keypresses. The walls of the room as well as most furniture and other objects in the room were painted black. During the experiment, the lights in the room were switched off. The laptop displays instructions when key-presses are required, but otherwise displays a black screen. The light leakage inherent in LCD displays was the only source of illumination in the room, other than the light coming from the images displayed on the Sony BVM-X300. The keyboard is also backlit, so that participants were able to find the correct keys to press during the experiment.

Each participant was presented with a short questionnaire to collect basic statistics on the population of participants. Further, all participants had to give consent for their anonymous results to be used in publications.

After completing the questionnaire, the laptop display would show written instructions. The instructions were written to ensure that each participant would carry out the same task. The instructions read as follows:

> ------ Instructions ------
>
> You will be shown pairs of images in the following screens
where one of the images is the input, and the other has been
processed in terms of luminance and chrominance, possibly degrading the image.
Please indicate with the left and right cursor keys
which image you think has been processed. You must select one
of the images, even if you do not know. You can only select
after the screen has gone blank, i.e. after 5 seconds of observation.

The instructions were designed to expressly ask participants to search for degradations relative to the ground-truth image. The experiment was a two-alternative forced choice (2AFC) experiment, meaning that for every image the left or right presentation had to be selected, even if the participant did not know the answer. Such designs are common for evaluation tasks, and they admit a relatively straightforward analysis.

Where possible stimuli presentations were randomized:

– The HDR and SDR comparison experiments were carried out separately, one after the other. HDR and SDR stimuli were not mixed, to avoid requiring participants to continuously adapt to different illumination levels. The order in which participants did these two experiments was randomized.
– For each stimulus in the HDR comparison, I_{HDR} was shown either left or right, with I_{TM-ITM} taking the remaining location.
– Stimuli in the SDR comparison were randomized similarly.
– The order of presentation of each image was randomized.

Such randomization is necessary to avoid learning effects, which might otherwise pollute the results.

Stimuli were shown for five seconds, as since approximately 2010 this is currently the average shot length in movies. This duration was chosen as the use case for the technologies under investigation ultimately lies in video processing. Participants were not allowed to select left or right presentations before they had observed each stimulus for five seconds, to avoid participants rushing the experiment. In practice, participants quickly settled into a rhythm whereby responses were recorded within a second after the end of each stimulus presentation. As a result, each of the HDR and SDR comparisons lasted for around 12 to 13 minutes, and the total time spend on both experiments together was around 25 minutes.

A1.4 Participants

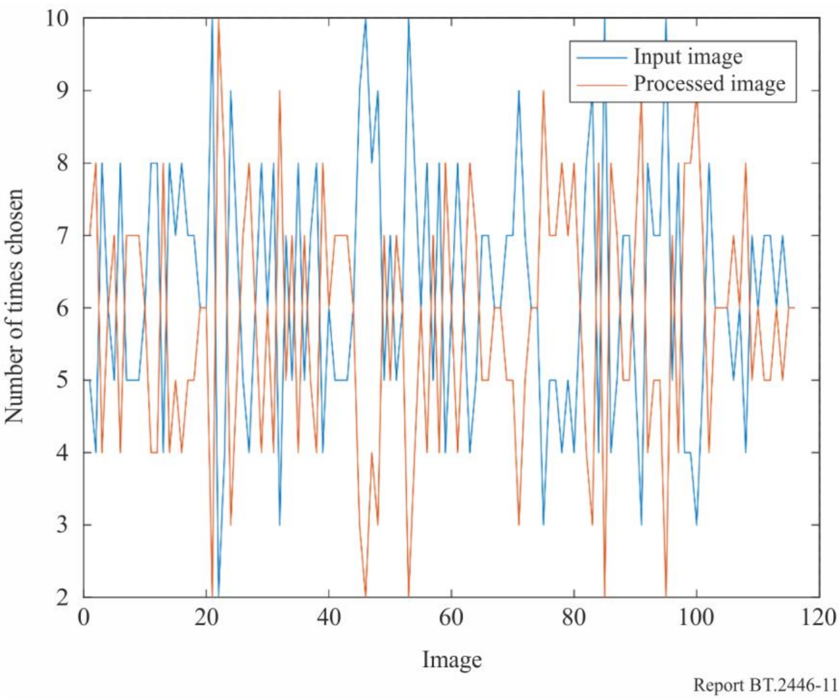
Participants were drawn from the population of Technicolor employees, Ph.D. students, post-docs as well as interns. Six males and six females participated in the experiments (all participants did both HDR and SDR comparisons). Their ages ranged between 22 and 49 years (mean: 33.4 years). One male described himself as colour anomalous. All participants had normal or corrected-to-normal vision. Ten participants were naïve as to the purpose of the experiment. The pool of participants consisted of a mixture of experts and non-experts in the area of high dynamic range imaging. All participants have given their consent for their anonymised results to be used in this publication.

A1.5 HDR comparison results

The HDR comparisons test whether a round trip from HDR to SDR and then back to HDR leads to a loss of visual quality. The responses from all participants for each image are plotted in Fig. 11. To determine if the processed image is chosen more often than the input image, the data in this plot was subjected to an unpaired *t*-test, which tests if the means of these two distributions are the same. The null-hypothesis that the means are the same cannot be rejected ($p = 0.167$), so that statistically, input images was chosen equally often as the tone mapped and subsequently inverse tone mapped images[6].

FIGURE 11

HDR comparison results, counting (over 12 participants), for each image in the database the number of times the HDR input image was chosen, and conversely, how many times the tone mapped and inverse tone mapped image was identified as processed

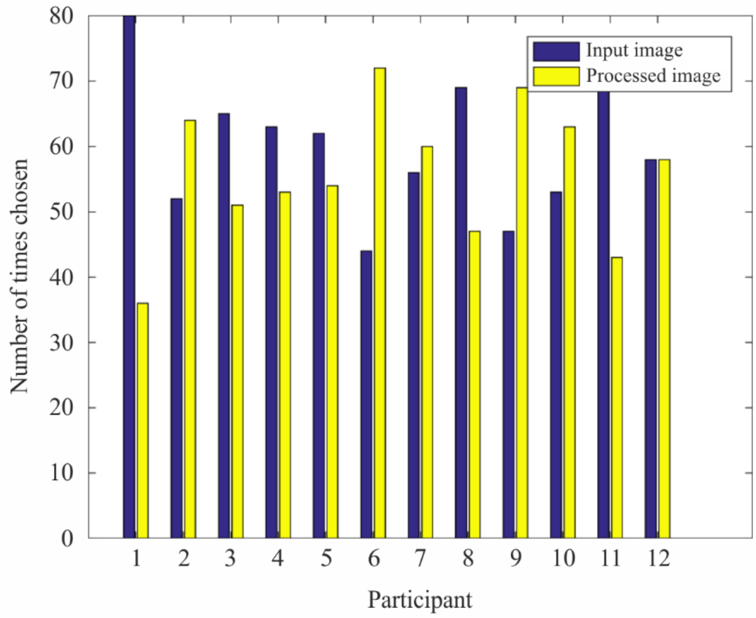


[6] Note that we can equally test whether the number of times the processed images were chosen is equal to half the number of participants. Such *t*-tests, however, produce the same results.

Figure 12 shows the distribution of responses per participant. A *t*-test on this data shows that the means of the two distributions shown in this plot are statistically the same ($p = 0.495$), confirming the results presented above.

FIGURE 12

HDR comparison, choices shown per participant



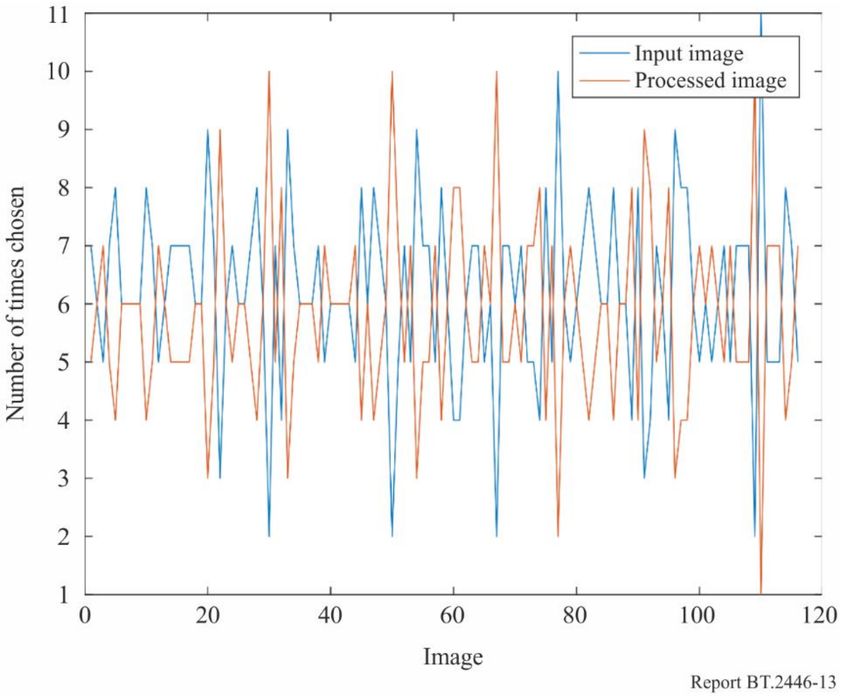
Report BT.2446-12

A1.6 SDR comparison results

In the SDR comparison results, the tone mapped images were subsequently inverse tone mapped, and then tone mapped for a second time. The tone mapped and twice tone mapped images were compared. Counts for each image are shown in Fig. 13. Here, a *t*-test revealed that once again the nullhypothesis- is rejected ($p = 0.196$), so that the means of the two distributions plotted in Fig. 14 are statistically the same.

FIGURE 13

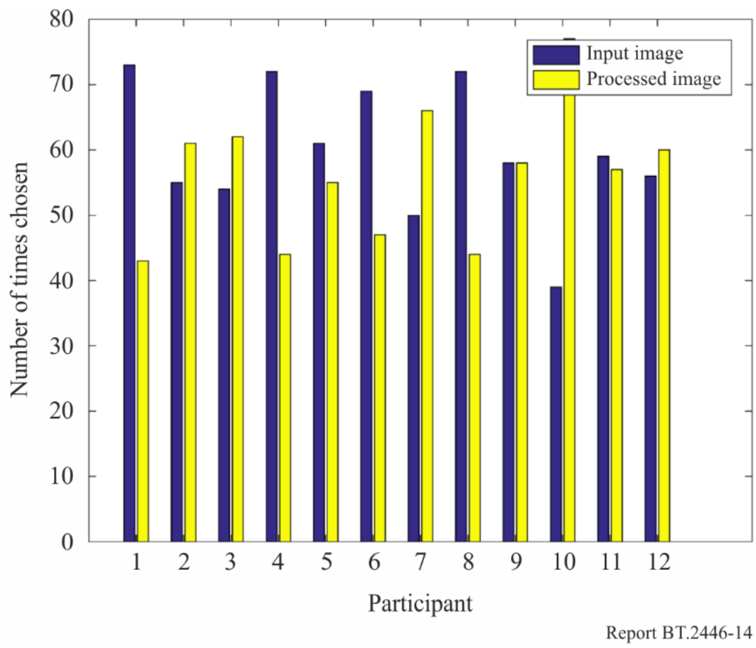
SDR comparison results, counting (over 12 participants), the number of time the input SDR image was chosen over the inverse tone mapped and subsequently tone mapped image



Confirming these results, a t-test on the per-participant choices (shown in Fig. 14) shows that participants chose the input and processed images equally often ($p = 0.549$).

FIGURE 14

SDR comparison results per participant



A1.7 Conclusions

The two experiments presented in this section reveal that the tone reproduction operator described in § 3.1 is well matched to the inverse tone reproduction operator presented in § 3.2. The instructions to the participants were designed to let participants look for flaws in the tone mapped and inverse tone mapped images. The analysis has shown that participants are not able to indicate which images have been processed, leading to the conclusion that tone mapping followed by inverse tone mapping leads to imagery of equal quality. The same conclusion is drawn for SDR images that were inverse tone mapped and subsequently tone mapped.

In addition, the SDR comparison is using SDR inputs that were created by first tone mapping HDR content. These images were then compared to images that were tone mapped twice and inverse tone mapped once. Despite this longer chain of processing, degradation of content could not be identified as such: the twice tone mapped images and the once tone mapped images remained of equal quality.

These experiments were performed under controlled conditions, designed to maximally reveal flaws in the tone reproduction and inverse tone reproduction operators (use of a dark room and a high-end grading monitor, as well as the formulation of the instructions to the participants). Under such conditions, participants were not able to distinguish processed images from their unprocessed inputs. There is confidence that in practice round-trip performance is sufficiently high that the proposed technologies can be used to bring SDR content into an HDR workflow, even if the content is subsequently tone mapped for display on an SDR monitor. In addition, the opposite workflow (HDR content passed through an SDR workflow), is equally achievable.

Bibliography

[1] Mekides Abebe, Tania Pouli, Jonathan Kervec and Chaker Larabi, *Color Clipping and Over-Exposure Correction*, In *Eurographics Symposium on Rendering*, 2015.

[2] Francesco Banterle, Patrick Ledda, Kurt Debattista and Alan Chalmers, *Inverse tone mapping*. In *GRAPHITE '06*, pages 349-356, 2006.

[3] Francesco Banterle, Patrick Ledda, Kurt Debattista and Alan Chalmers, *Expanding low dynamic range videos for high dynamic range applications*. In *SCCG '08: Proceedings of the 4th Spring Conference on Computer Graphics*, pages 349-356, New York, NY, USA, 2008. ACM.

[4] Piotr Didyk, Rafa Mantiuk, Matthias Hein and Hans-Peter Seidel, *Enhancement of bright video features for HDR displays*. In *Proceeding of Eurographics Symposium on Rendering 2008*, Computer Graphics Forum. Eurographics, 2008.

[5] Fairchild, Mark D., *Color appearance models*. John Wiley & Sons, 2013.

[6] Hunt, Robert W. G., *Light and dark adaptation and the perception of color*, Journal of the Optical Society of America, 42(3):190-199, 1952.

[7] Belen Masia, Sandra Agustin, Roland W. Fleming, Olga Sorkine and Diego Gutierrez. *Evaluation of reverse tone mapping through varying exposure conditions*. *ACM Transactions on Graphics*, 28(5), 2009.

[8] Laurence Meylan, Scott Daly and Sabine Süsstrunk. *Tone mapping for high dynamic range displays*. In *Electronic Imaging*, volume 6492, 2007.

[9] Reinhard, Erik, Erum Arif Khan, Ahmet Oguz Akyuz and Garrett Johnson. *Color imaging: fundamentals and applications*. CRC Press, 2008.

[10] Rempel, Allan G., *et al.*. *LDR2HDR: on-the-fly reverse tone mapping of legacy video and photographs*. *ACM Transactions on Graphics*, 26(3), 2007.

[11] Larson, Greg Ward, and Rob Shakespeare. Rendering with Radiance: the art and science of lighting visualization. Booksurge Llc, 2004.

[12] Reinhard, Erik, Greg Ward, Paul Debevec, Sumanta Pattanaik, Wolfgang Heidrich and Karol Myszkowski. *High dynamic range imaging: acquisition, display, and image-based lighting*. 2nd edition, Morgan Kaufmann, 2010.