USING A COLOUR APPEARANCE MODEL ACCOUNTING FOR THE HUMAN VISUAL SYSTEM FOR DOWN MAPPING HYBRID LOG-GAMMA (HLG) TO STANDARD DYNAMIC RANGE (SDR)

Daniel-Dee Lofthouse-Smith

BBC Research and Development, United Kingdom

ABSTRACT

Most of today's High Dynamic Range (HDR) Hybrid Log-Gamma (HLG) to Standard Dynamic Range (SDR) conversions rely on a mathematical approach for colour volume reduction from a BT.2100 container into a BT.709 one, typically through "tone-mapping" applied to a luminance signal. Examples can be found in ITU-R report BT.2446. Such "down-mapping" conversions attempt to maintain colour fidelity between the two visual representations and are usually implemented in a three-dimensional Look Up Table (LUT). Such approaches, however, cannot take full account of psychovisual factors and may struggle to differentiate between colours such as yellow and brown, that are similar in hue, but vary in luminance. An alternative down-mapping approach can be carried out through a more sophisticated Colour Appearance Model (CAM) that is combined with a Chromatic Adaptation Transform (CAT). The CAT transforms the image into a uniform colour space and the CAM calculates values of perceived quantities of the viewed media, such as lightness and colourfulness given specific display parameters.

A CAM has been extended in this work in order to compute HLG to SDR down-mapping. Perceptual quantities in the model are for the most part conserved to compute the down-mapping. This method has shown great promise in several natural images and colours, whilst having limitations with engineering test signals. SDR footage derived from an HDR capture using a CAM LUT results in natural pictures that are more visually accurate than those from conventional conversion technologies or SDR capture, since many colours in nature exist outside of a BT.709 container. This approach may not deliver mathematically ideal output values for artificial test patterns so the paper will also discuss what changes may need to be made to TV production workflows, if the improvements offered by CAM based down-mapping are to be fully realised.

INTRODUCTION

British Broadcasting Corporation (BBC) trials for High Dynamic Range (HDR) production have focused on "establishing a single production workflow with Standard Dynamic Range (SDR) output derived from Hybrid Log-Gamma (HLG) HDR programme output". To this end, a single HLG HDR production workflow will require a robust method for HLG to SDR conversion. The current BBC implemented method (which is largely explained in BT.2446 4.2 ITU-R, (1) "Method B"), employs a tone mapping method with additional proprietary modifications which down maps HLG HDR signal in a BT.709 colour volume.

Typically, tone mapping looks to map a luminance component in one colour space to another to achieve the HDR to SDR down mapping. Tone mapping is defined in BT.2446 as "the non-linear mapping between HDR to SDR content" (1) with BT.2446 presenting three different conversion methods: A B and C. More detail for each method can be found in BT.2446 Sections 3.1, 4.2 and 5.1 respectively (1). After down mapping these methods also require colour space conversion to go from a BT.2100 colour volume to a BT.709 colour volume where the typical methods for this are described in ITU-R BT.2407 ITU-R, (3).

Colour space conversion in BT.2407 can be carried out through a simple conversion from BT.2020 to BT.709 based on a linear matrix conversion (3), Section 2 or through more complex methods that include a colour gamut mapping step which can be carried out by way of a uniform colour space. In BT.2446 Section 3, conversion method A (1), care is made to account for the human colour perception brightness interaction. If it is desirable to produce a visual match between the HLG and SDR content, it is thought that the use of a colour appearance model for the mapping of HLG to SDR can be used to best ensure this visual matching.

A Colour Appearance Model (CAM) is a mathematical model that attempts to represent the perceptual aspects of a viewing stimuli by the human visual system. A CAM can be used to obtain the basic appearance parameters of this stimuli for a given viewing environment. These parameters are brightness (perceived strength of illumination), lightness (perceived strength of illumination with respect to the reference white), hue (ratio of the achromatic components of the colour in question), chroma (colourfulness judged with respect to a similarly illuminated white), colourfulness (perception of an area to be more or less chromatic) and saturation (ratio between brightness and colourfulness). In this paper we derive a Colour Appearance Model (CAM) using a similar approach to that shown in Kim et al, (2) but include serveral modifications and apply to HDR to SDR conversion to preserve the HLG BT.2100 input appearance in an SDR BT.709 volume with both the luminance tone mapping and the chromatic adaption being computed by the model.

CURRENT HDR TO SDR REDUCTION METHOD

As mentioned previously, the existing BBC implemented method is largely explained in (1), 4.2 where some key features of the method will be discussed here.

In this model, the nominal peak luminance (L_W) of the orignal HLG signal for conversion is set to approximately 400 cd m^{-2}. The selection of this L_W means that the reference white of the HDR content (75% HLG) is mapped to the peak white in a standard SDR display (100 cd m^{-2} as defined in BT.1886 (9) and BT.2035 (10)).

The following steps are implemented in this method in order compute the HLG to SDR down mapping:

- Compute the HLG EOTF on the HLG signal with L_W ~ 400 cd m^{-2} to get SDR Display Light
- Compute RGB_{2100} to XYZ
- Transform this XYZ signal into a luminance and colour opponents signal
- Tone map the luminance component in order to transform it to SDR
- Colour map the signal into a BT.709 volume
- Return the luminance and colour opponent values into XYZ
- Compute XYZ to RGB_{709} to get SDR BT.709 display light

- Compute inverse BT.1886 EOTF with L_W = 100 cd m^{-2} to give the final SDR BT.709 signal

These steps are implemented to ensure that the HLG and the SDR output have the same "look".

This conversion also maps the 75% BT.709 "equivalent" display light (DL) colour bars embedded within the EBU Tech 3373 EBU, (4) colour bar pattern to a 75% SDR output signal. The EBU Tech 3373 colour bars test pattern is used as an engineering test signal for HDR broadcast and ilustrated below. Monitoring the output of the respective colour bars after DL or scene light (SL) conversion, allows video engineers to confirm the down mapping is working correctly.

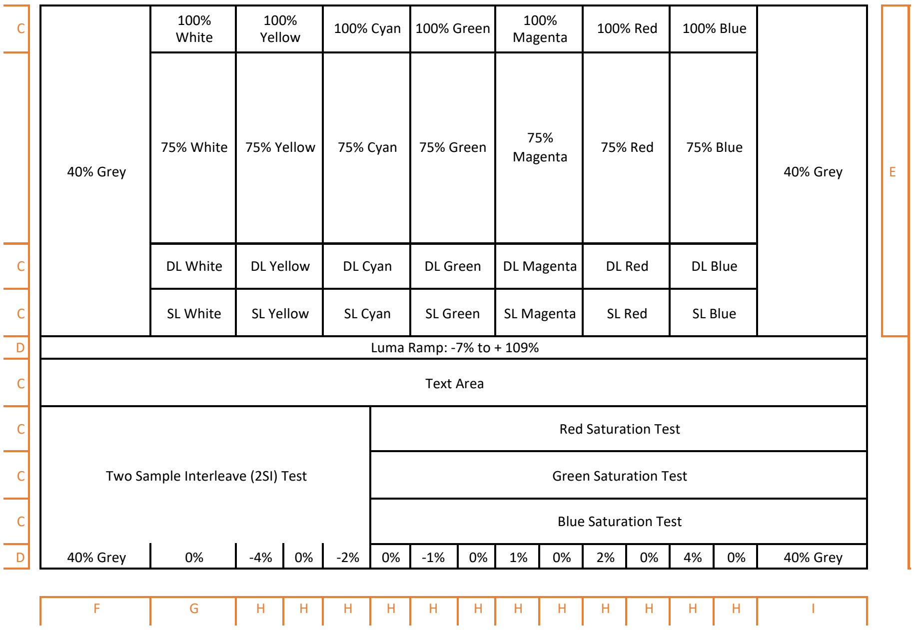

Figure 1: EBU Tech 3373 Colour Bars (4)

Difficulties with existing implementations

Using the current implementation, and other traditional methods that rely on a conversion based on tone mapping the luminance, has a few difficulties. One such difficulty arises when attempting to down map colours that have similar chromatic components but vary only in their luminance. In this circumstance, approaches based on tone mapping, which compresses the luminance, will reduce the perceived difference between the colours. They will, therefore, struggle to faithfully represent these colours in a BT.709 colour volume. It is thought that using a visual perception model (i.e., a CAM) will mean that these colours that are visually different but are essentially mathematically identical in their chrominance can be better produced in a BT.709 colour volume.

THE NEW COLOUR APPEARANCE MODEL HLG TO SDR DOWN MAPPING

The CAM described in (2) deconstructs an HDR image into different visual perception components. Some of these include lightness (J), Colourfulness (M) and Hue (h). In the creation of an HLG to SDR down mapping pipeline, in order to match the output SDR to the input HLG image we aim to conserve components such as the colourfulness and hue, whilst adjusting the lightness to ensure that the output fits within the BT.709 colour volume.

Some modifications were also made to the model shown in (2). The first of these is a change in the CAT through altering the Uniform Colour Space (UCS) that was used. The model states the use of the CIECAT02 colour space, however works by Changjun et al, (5) have shown that this colour space has a negative value problem. Their suggestion for solving this issue is by adopting a new UCS known as CAT16 and that has been implemented in this HLG to SDR down mapping. Further refinements to this CAT16 UCS have been noted in Schlomer,(6) which attempts to simplify the CAT whilst also solving a problem with inputs of zero generating non zero values, so again these changes have been implemented here.

In addition to this, to reduce the mathematical complexity of this model, the lightness function from (2) has been removed and replaced with a simplified estimation for the lightness for the down mapping. This new implementation uses the ratio of the achromatic signal component to the achromatic reference white of the display (either HDR or SDR). A scale factor reduction is also applied to this variable when converting from HLG to SDR.

Lastly, an addition has been made in the final step of the HLG to SDR down mapping. Throughout the model, colourfulness is conserved between the HLG input and the SDR output. However, there are instances when the RGB output is still outside of the BT.709 colour volume and so a desaturation of these colours is required. In the model the colourfulness is reduced until this colour is within gamut. This is seen in Figure 2 , where in CIE 1976 u' v' colour space, a u' v' combination (shown in orange) that existed outside of the BT.709 gamut (green triangle) is brought within the gamut throught the process of colourfulness reduction. A flowchart for the colourfulness reduction is shown in Figure 4.

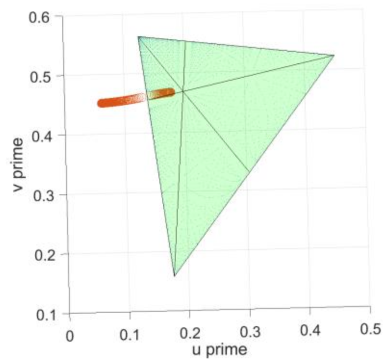

Figure 2: Colourfulness reduction of an RGB triplet shown in CIE 1976 which brings the u' v' components of this triplet within the BT.709 gamut (shown in green)

Following the implementation of these changes the processing blocks for HLG signals to SDR is shown in Figure 3. In the first stage of this process the chromaic adaptation (L_c) is computed using a simplified version of the CAM 16 UCS. "Cone" responses (i.e., LMS) are calculated by the multiplication of the $M16_{D65}$ matrix to the CIE 1931 XYZ components of the input display light. Colour adaption is computed by Equation 1, where D_R is a surround parameter with L being the cone response. Equation 1 can be interchanged with either M or S.

$$\boldsymbol{L_c} = D_R \cdot L \qquad \text{Equation 1}$$

The absolute cone response is then derived by the following formula taken from (2) and is shown in Equation 2.

$$V' = \frac{V^{n_c}}{V^{n_c} + L_a^{n_c}} \qquad \text{Equation 2}$$

Here, V is the response of either the L,M or S cone, L_a is the absolute level of adaptaion (taken as 20% of the reference white luminance from the "grey world" assumption) and n_c is a constant where n_c=0.57.

The derivation of the colourfulness (M), chroma (C) and the hue angle (h) and the constants used in their equations are taken from from (2) are used in this down mapping along with an adjusted lightness parameter used for this down mapping. The equations for the calculation of each of these parmeters is shown below,

$$C = \alpha_k \cdot (\sqrt{a^2 + b^2})^{n_k}$$ Equation 3

$$M = C \cdot (\alpha_m \log_{10} L_w + \beta_m)$$ Equation 4

$$h = \frac{180}{\pi} \tan^{-1}(b/a)$$ Equation 5

$$x = {}^{A}/_{A_w}$$ Equation 6

Where α_k is 456.5, n_k is 0.62, α_m is 0.11 and β_m is 0.61. L_w is the reference white luminance, A_w the achromatic component of this reference white, and A and a,b being the achromatic and chromatic components respetively derived from the absolute cone responses of the human visual system. It should be noted that PQ signals can be used with this down mapping process but must be first converted to HLG via the transcoding method described in BT.2390 ITU-R, (7) i.e. computing a PQ EOTF followed by an HLG inverse EOTF.

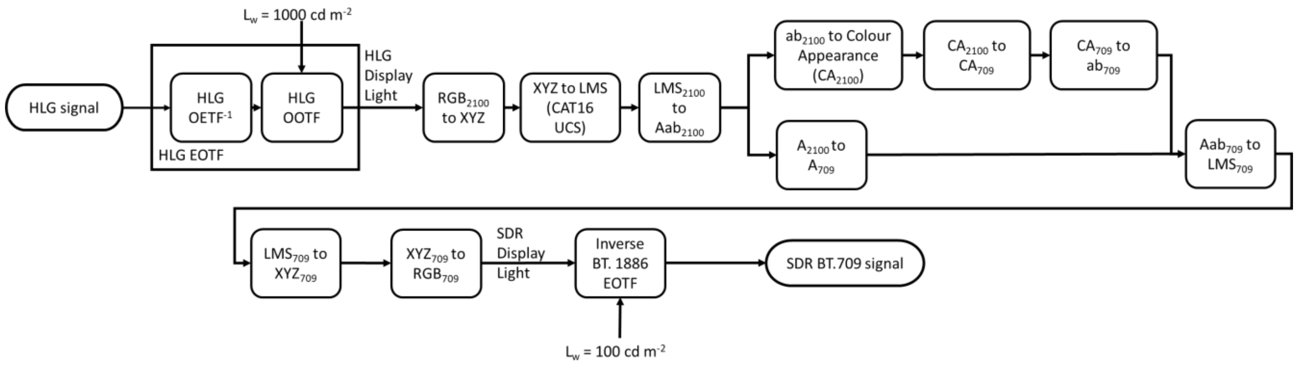

Figure 3: HLG to SDR down mapping process

The processing block for desaturation of the RGB_{709} SDR signal back into BT.709 gamut is shown in Figure 4. This is an iterative process and is used to ensure that all colours down mapped exist inside the BT.709 colour volume.

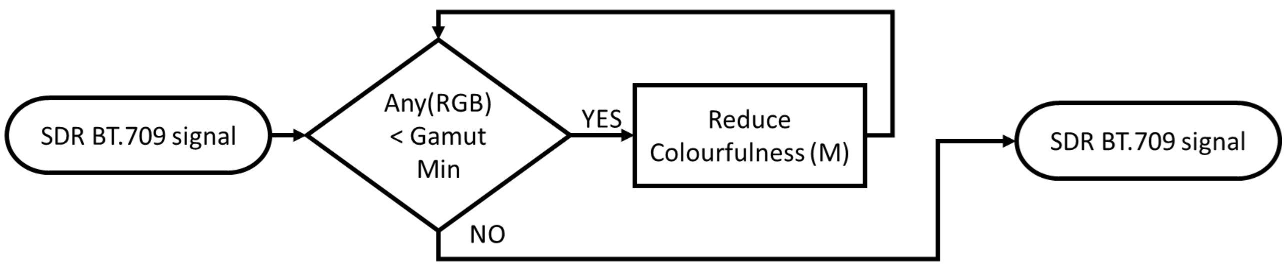

Figure 4: Process for desaturation of colours into the BT.709 colour volume

RESULTS

HLG images can be broken down into their constituent colour appearance components during the down mapping process. Using the CAM down mapping, the colourfulness of the original image can be calculated and compared to the colourfulness of the down mapped image. This can be used to identify how well the colour fidelity of the image is maintained, and therefore maintain the same "look" as the original HLG. Colourfulness can also be obtained for down mapped images (through the CAM model) using the BBC's "traditional" process and the ITU-R BT.2446, method A in order to compare the colourfulness conversation of the varied down mappers.

The colourfulness of the down-mapped images is derived independently of their down mapping process. Colourfulness is obtained from SDR display light given a display with a peak luminance (L_W) of 100 cd m^{-2}. Therefore the colourfulness of the SDR content can be thought of as being independent of the down mapping process, but based on the visual perception one might have viewing any down mapped content on an SDR display given the colour appearance model parameters.

Image selection favours images that have many pixels that have colours that are very similar in chromatic coordinates but differ largely only in brightness. Moreover, selection also favours images with a significant portion of the image existing outside of the BT.709 gamut. This will better test the limitations of the CAM down mapping implementation. Several images have been chosen and are shown in Figure 5. These have been chosen due to their inherent difficulty to visually match the original HDR footage following down mapping as a large proportion of the image exists outside of the BT.709 colour volume and is shown in Figure 5 as the black regions of the respective images. Note, these HLG images are for reference only and the colours have not been adjusted for print media.

The images chosen are:

- Dogs at sunset
- Sun Flare
- Stadium Sky
- Red Sail

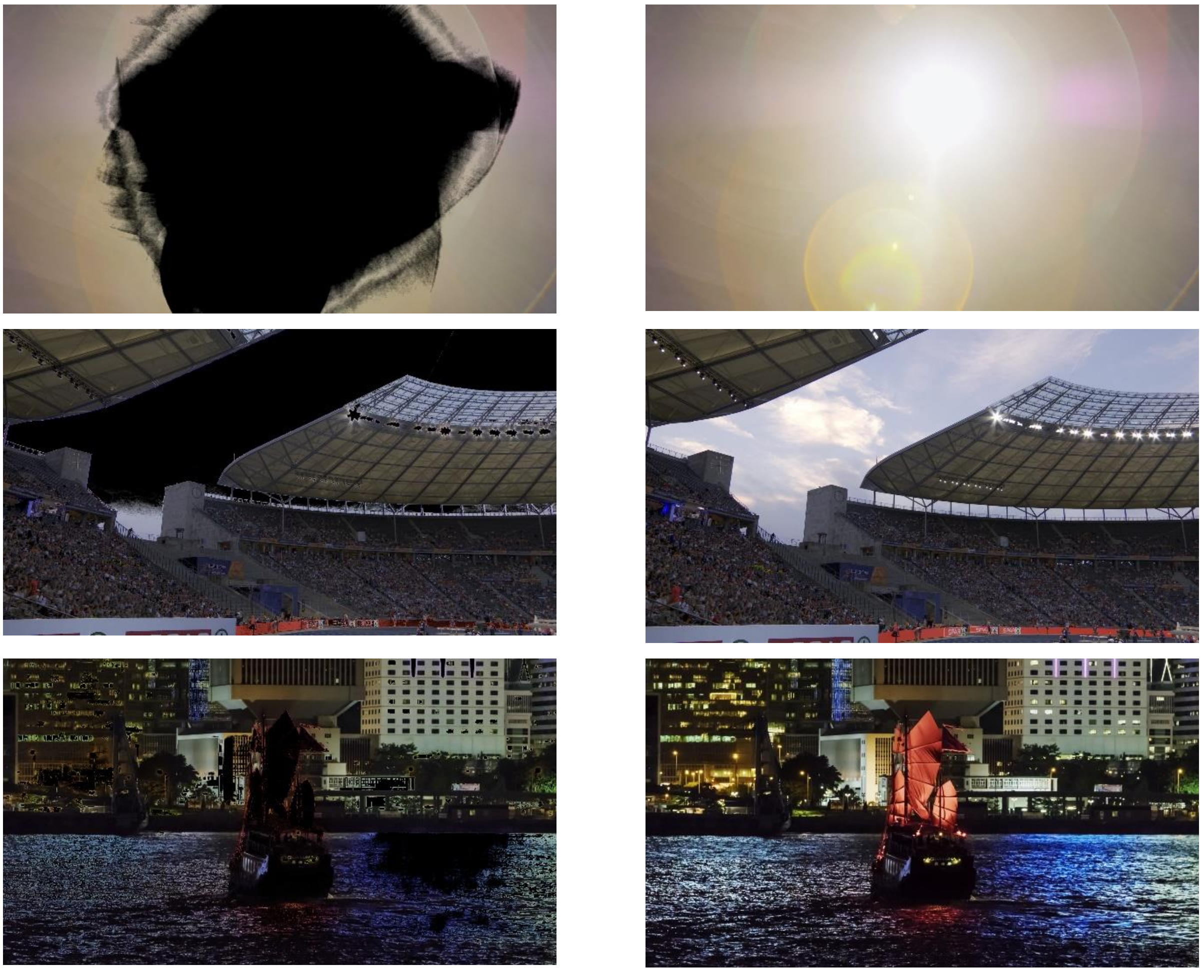

Figure 5: The HLG image original, and the image, left, with regions that are outside of the BT.709 triangle shown in black.

In order to speed up computational time in the analysis, small regions of interest can be selected from these images (favouring the parts of the image that are outside of gamut). From the results presented in Figure 6 and Figure 7 it is clear that the change in colourfulness has been significantly less in the CAM down mapping than in the current BBC implementation. This suggests that the colour appearance of the down mapped footage using the CAM implementation will better match the original HLG footage than the most up to date BBC conversion techniques. This can also be compared to the change in colourfulness of the ITU-R BT.2446 Method A tone mapping down mapping (1). The figure shows that although the BBC implemented down mapping does not maintain the colourfulness as well as the CAM model, it significantly outperforms the ITU-R BT.2446 Method A down mapping implementation.

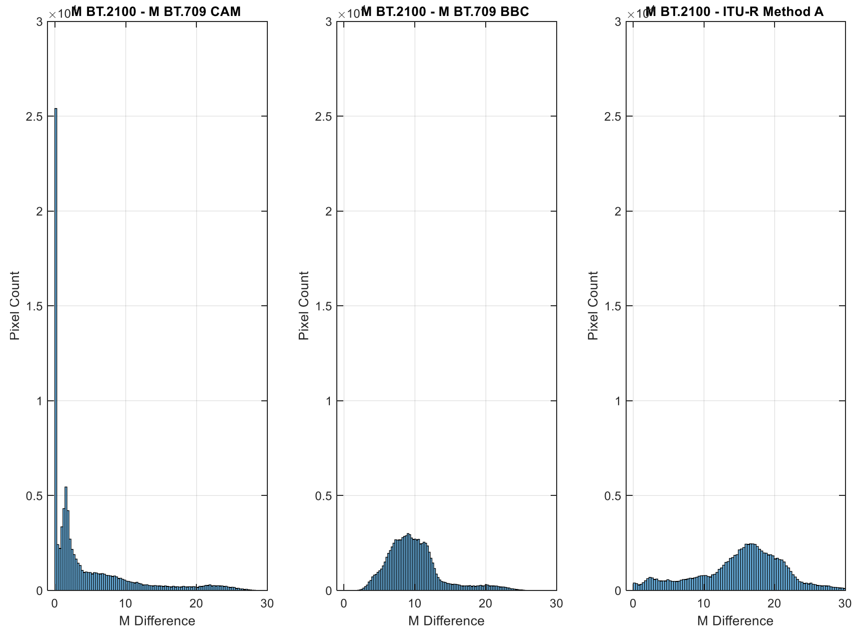

Figure 6: Histogram of the difference between HLG colourfulness and down mapped colourfulness for the CAM down mapping, left, the BBC down mapping, middle and the ITU-R BT.2446 Method A tone mapping, right for the ‘Dogs at Sunset’ file.

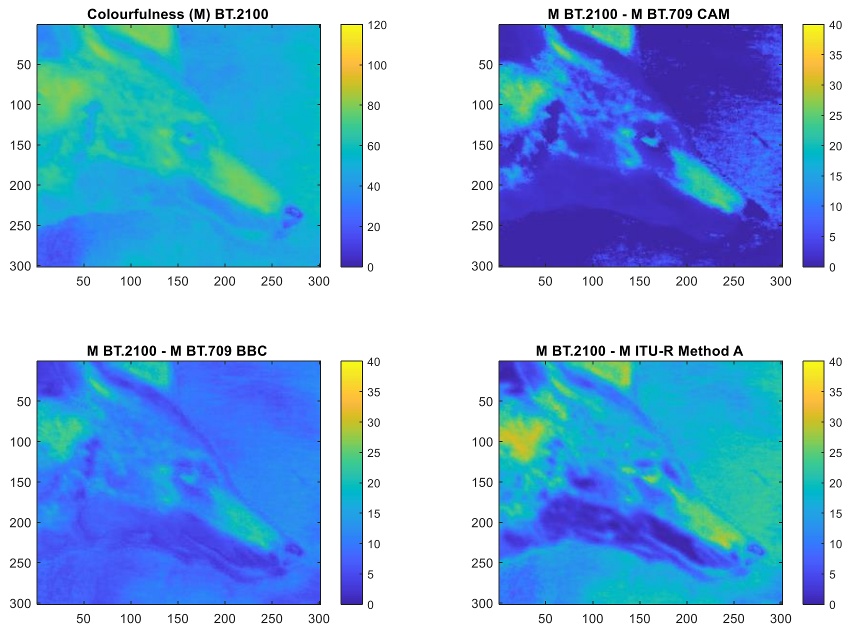

Figure 7: Original HLG colourfulness for ‘Dogs at Sunset’, top left, and the difference between this and down mapped colourfulness for the CAM down mapping, top right, the BBC down mapping, bottom left and the ITU-R BT.2446 Method A, bottom right.

As shown in Figure 5, several images are used to compare the colourfulness changes following the three down mapping processes. The difference in their colourfulness along with the original colourfulness in HDR is shown in the following figures. Like the results presented in Figure 6, the CAM down mapping shows the lowest difference in colourfulness when compared to the original HLG still, with the ITU BT.2446 Method A implementation performing the worst. This hypothesis is further supported by Figure 8, Figure 9 and Figure 10, where the CAM colourfulness difference is the lowest, and the colourfulness difference is greatest in the ITU-R Method A down mapping.

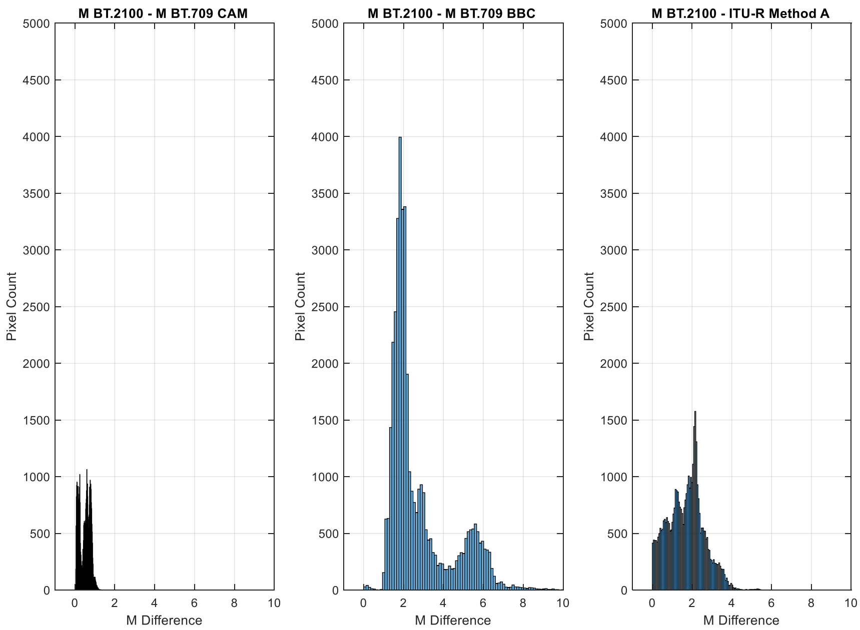

Figure 8: Histogram of the difference in colourfulness from the original HLG for the ‘Sun Flare’ file for the CAM down mapping, left, the BBC down mapping, middle and the ITU-R BT.2446 Method A tone mapping, right.

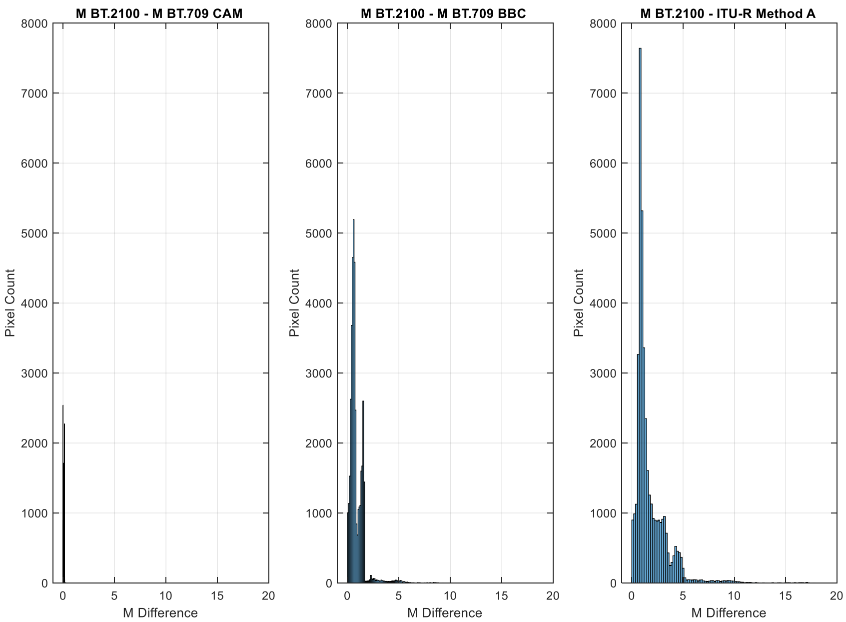

Figure 9: Histogram of the difference in colourfulness from the original HLG for the ‘Stadium Sky’ file derived from the CAM, left, the BBC down mapping, middle and the ITU-R BT.2446 Method A tone mapping, right.

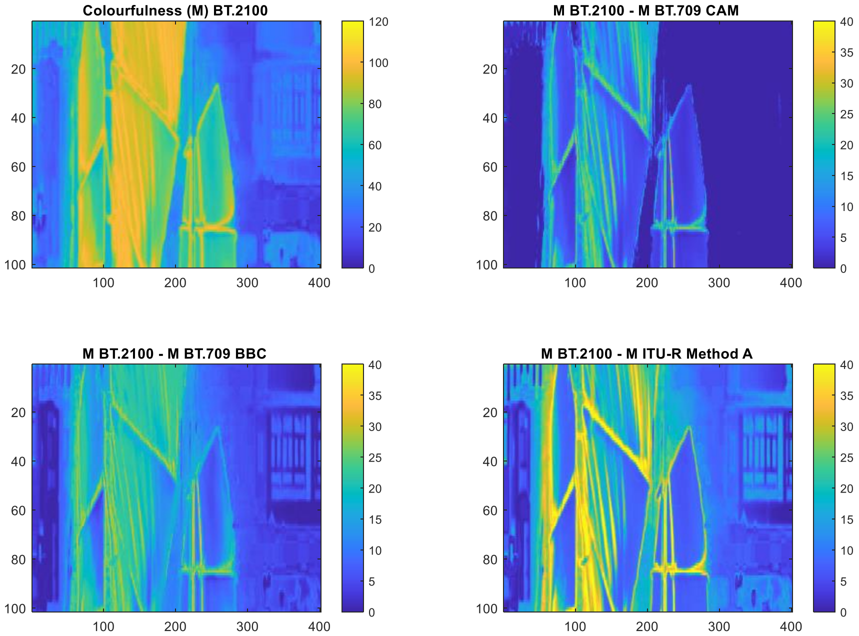

Figure 10: Original HLG colourfulness for ‘Red Sail’, top left, and the difference between this and down mapped colourfulness for the CAM down mapping, top right, the BBC down mapping, bottom left and the ITU-R BT.2446 Method A, bottom right

Down mapping is typically implemented using a 3D-LUT and this can then be compared with down mapped HLG content with 3D-LUTs created using the most up to date BBC down mapping processes available. The results of these comparisons are much less computationally expensive than a pixel-by-pixel implementation, and in real broadcast content, a 3D-LUT (typically a 33 cube LUT) is favoured here. This 3-D LUT implemented transform is shown in Figure 11 comparing the BBC LUT, the CAM LUT and a down mapping created from ITU-R BT.2446 method A.

CAM LUT	BBC LUT	ITU-R BT.2446 Method A
		
		

Figure 11: A selection of down mapped images using the CAM implementation (left) the most up to date BBC LUT (middle) and the ITU-R BT.2446 Method A (right).

These images highlight the benefit of implementing a colourfulness conservation, since all the different down mapping methods appear to have a different colour appearance across the whole image. Maintenance of the colour appearance between HLG and SDR is therefore best using the CAM approach.

CAM BASED CONVERSION WITH COLOUR BARS

As discussed in the current HDR reduction methods section, the down mapping used for television production relies on the preservation of the EBU Tech 3373 DL or SL bars. However, the CAM down mapping is only able to conserve the achromatic DL bars at 75% grey. Other bars are not output at 75% SDR signal. The CAM down mapping suggests that these EBU test bars may not be fully representative of the visual perception of these colour bars, or that the CAM needs to account for these DL bars to make use of these engineering test signals. A better approach might be to include a third set of BT.709 equivalent bars alongside the existing DL and SL bars, embedded within the signal using a CAM based method. These might more reasonably be expected to survive CAM based conversion, to deliver accurate BT.709 colour bars after conversion.

CONCLUSIONS

This work shows promise in the maintaining of colour appearance between HLG and SDR footage. It consistently outperforms the most up to date BBC tone mapping implementation in colourfulness preservation in test images that have a large proportion of the image outside of gamut. Although this is the case, there are some difficulties that arise from this implementation. Further work on the EBU test pattern alongside the CAM is therefore required to better understand this phenomenon. This mathematical analysis on SDR down mapped content would profit from subjective testing to confirm which model most faithfully represents the original HLG footage amongst a range of test subjects. However, due to the current Covid-19 restrictions in place in the United Kingdom, it has not been possible to conduct any testing. If subjective testing proves this new down mapping method provides the best representation of HLG in SDR, the current engineering test pattern would need reviewing. This subjective testing would therefore be an area of interest for the further development of this work.

BIBLIOGRAPHY

1. **ITU-R.** *Report ITU-R BT.2446-0 Methods for conversion of high dynamic range content to standard dynamic range content and vice-versa.* Geneva : ITU, 2019.
2. *Modelling human colour perception under extended luminance levels .* **Min H Kim, Tim Weyrich, Jan Kautz.** 2009, ACM SIGGRAPH, pp. 1-9.
3. **ITU-R.** *Report ITU-R BT.2407-0 Colour gamut conversion from Recommendation ITU-R BT.2020 to Recommendation ITU-R BT.709.* Geneva : ITU, 2017.
4. **EBU.** *Tech 3373 Colour bars for use in the production of Hybrid Log Gamma (HDR) UHDTV.* Geneva : EBU, 2020.
5. *A Revision of CIECAM02 and its CAT and UCS.* **Changjun Li, Zhiqiang Li, Zhifeng Wang, Yang Xu, Ming Ronnier Luo, Guihua Cui, Manuel Melgosa Michael Pointer.** San Diego : Society for Imaging Science and Technology, 2016. Vol. 1, pp. 208-212.

6. **Schlomer, Nico.** *Algorithmic improvements for the CIECAM02 and CAM16 color appearance models.* s.l. : arXiv:1802.06067, 2018.
7. **ITU-R.** *Report ITU-R BT.2390-8 High dynamic range television for production and international programme exchange.* Geneva : ITU, 2020.
8. **Francesco Banterle, Artusi Alessandro, Debattista Kurt, Chalmers Alan.** *Advanced High Dynamic Range Imaging.* 2nd. Natick, MA, USA : AK Peters (CRC Press), 2017.
9. **ITU-R.** *Recommendation ITU-R BT.1886 Reference electro-optical transfer function for flat panel displays used in HDTV studio production.* Geneva : ITU, 2011.
10. **ITU-R**. *Recommendation ITU-R BT.2035 A reference viewing enviroment for evaluation of HDTV program material or completed programmes.* Geneva : ITU, 2013.

ACKNOWLEDGEMENTS

The author gratefully acknowledges, and thanks, other members of the team, particularly Simon Thompson and Andrew Cotton, for their contribution to this work.