A “DISPLAY INDEPENDENT”
HIGH DYNAMIC RANGE TELEVISION SYSTEM

T. Borer and A. Cotton

BBC R&D, 56 Wood Lane, London, W12 7SB, UK

ABSTRACT

High Dynamic Range (HDR) television has captured the imagination of the broadcast and movie industries. This paper presents an overview of the BBC’s “Hybrid Log-Gamma” solution, designed to meet the requirements of high dynamic range television. The signal is “display independent” and requires no complex “mastering metadata”. In addition to providing high quality high dynamic range (HDR) pictures it also delivers a high quality “compatible” image to legacy standard dynamic range (SDR) screens and can be mixed, re-sized and compressed using standard tools and equipment. The technical requirements for a high quality HDR television system are presented. Quantisation effects (or “banding”) are analysed theoretically and confirmed experimentally. It is shown that quantisation effects are comparable or below competing HDR solutions. The psychovisual effects of presenting images on emissive displays in dim environments, “system gamma”, is shown experimentally and the analysis informs the design of this HDR system.

INTRODUCTION

With improvements in technology, television with greater impact, more “presence”, deeper “immersion”, a “wow factor”, or, in short, better pictures, is now possible. Ultra high definition, UHD, is not just about more pixels, it has the potential to deliver wider colour gamut (WCG), higher frame rates (HFR), and higher dynamic range (HDR). Of these perhaps high dynamic range offers the greatest improvement, and the costs of upgrading to HDR can be relatively low for both production and distribution. High dynamic range offers unmistakeably better pictures, across the living room, even on a smaller displays, and even to those with less than perfect vision. Potentially HDR may be produced using mostly installed, legacy, standard dynamic range (SDR) infrastructure, and distributed over largely unchanged distribution networks. No wonder that, even before the standards have been finalised, some movie studios are already talking about creating movies in HDR, Ultra HD for home viewing.

This paper describes the signal processing technology required for high dynamic range in the television production and distribution chains. It describes how one solution, the “hybrid log-gamma” approach, provides a “display independent” signal that can produce high quality images, which maintain the director’s artistic intent on a wide range of displays in diverse viewing environments. So, for example, precisely the same signal may be viewed in a controlled production suite, a home cinema, an ordinary living room, or on a laptop or mobile device. Furthermore the signal may be displayed on a conventional standard dynamic range display to provide a high quality “compatible” image. The log-gamma HDR signal may be mixed, re-sized, compressed, and generally “produced”, using conventional tools and equipment. The only specifically high dynamic range equipment needed is cameras and displays for quality monitoring (signal monitoring may continue to use SDR displays). No complex mastering metadata is required. Conventional end user distribution techniques may be used (although a 10 bit signal is required). No layered or multichannel codecs are required. Only a single signal is required for both SDR and HDR displays and expensive multiple “grades” (for both HDR and SDR) are not necessary.

The paper continues by discussing the meaning of high dynamic range. To understand HDR production and display we need to look at the television signal chain, which is discussed next. This then allows us to consider the design of the camera transfer characteristic (the opto-electronic transfer function (OETF)). Next we discuss an important psychovisual aspect of HDR TV, the “system gamma”. Whilst this effect has long been known in television, movies, and the academic literature, it assumes an enhanced importance for HDR. Based on an understanding of system gamma we discuss the design of the electro-optic transfer function (EOTF) in the display, and how this can allow the display of high quality pictures on a diverse range of displays. Once the EOTF is defined we can analyse the likely effect of quantisation and the performance, in terms of dynamic range, that may be expected from the system. This is compared to alternative HDR proposals and the theoretical analysis is compared to experimental results. The paper ends with some concluding remarks.

DYNAMIC RANGE

Dynamic range is the ratio between the whitest whites and blackest blacks in an image. For example printed images have a dynamic range of less than 100:1 (because it is difficult to make a black ink that reflects less than 1% of incident light). Dynamic range is often measured in “stops”, which is the logarithm (base 2) of the ratio. So printed images have less than 7 stops of dynamic range. Standard dynamic range consumer television (8 bit video, e.g. DVD, SD and HD DVB) only supports about 6 stops of dynamic range, as discussed below. Professional SDR video (10 bits) supports about 10 stops. But the human eye can see up to about 14 stops (1) of dynamic range in a single image. Higher dynamic range results in an experience closer to reality, and hence of greater impact or “immersion”. Furthermore higher dynamic range also increases the subjective sharpness of images and so provides a double benefit.

Some debate has confused high dynamic range with high brightness. The two are not the same. You can have high dynamic range in a dark movie environment, with pictures of only 48cd/m2. Alternatively you can have standard dynamic range on very bright screens of hundreds, or even thousands, of cd/m2. What high brightness does allow is to see high dynamic range without needing a very dark viewing environment.

It might be thought that higher dynamic range could be achieved by simply making displays brighter. But this is analogous to suggesting that you can increase the dynamic range of audio by turning up the volume. With audio, turning up the volume merely emphasises the noise. The same is true for video. With video the “noise” is quantisation noise, where the steps between quantisation levels become clearly visible. This is manifest as “banding”, “contouring, or “posterisation”. An extreme example of banding is shown below.

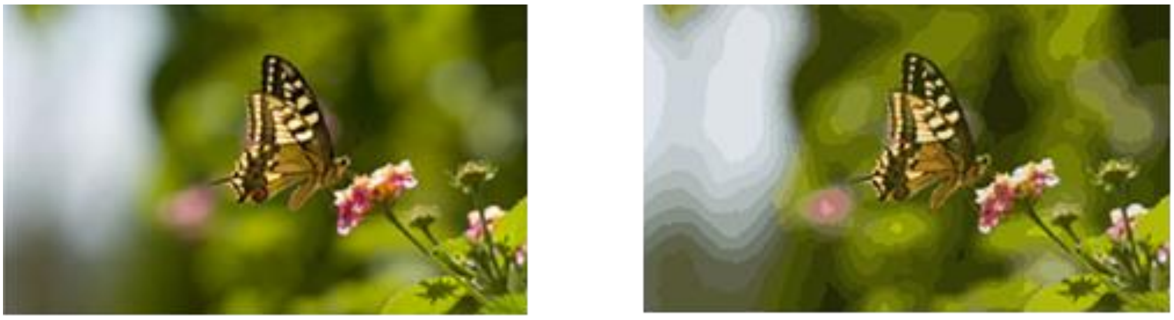

Figure 1:Image Quantisation, left original, right extreme banding

The useful dynamic range of video is determined by the ratio between adjacent quantisation levels. If the adjacent quantisation levels differ in luminance by less than 2% the difference is probably imperceptible in the image. This threshold of visibility increases at low luminance. The Schreiber curve[1] (shown later) is an approximation to the threshold of visibility.

In television systems the luminance represented by the signal is a non-linear function of the signal value. Conventionally this is a gamma curve, exemplified by ITU-R Rec 1886. For example the displayed luminance L may be the signal V raised to the power gamma, $L=V^{\gamma}$. The ratio between adjacent quantisation levels, also known as the Weber fraction, is given mathematically by:

$$\text{Weber fraction} = \frac{1}{\mathrm{N}\cdot\mathrm{V}}\frac{dL}{dV} = \frac{\gamma}{N}\frac{1}{L^{1/\gamma}}$$

Where N is the number of quantisation levels (220 for 8 bit video, 876 for 10 bit), and the algebraic form is given for a conventional gamma curve. Using this equation we may find the luminance corresponding to the threshold at which banding is visible, which then gives the usable dynamic range. For conventional, 8 bit SDR video, with gamma 2.4 and a 5% threshold (allowing for a dim display), this yields a dynamic range of only 5.27 stops. This is not a high dynamic range, a bit less than a photographic print, although it can be extended by using techniques such as dither. Overall the conventional display gamma curve is not adequate for HDR reproduction and a different non-linearity is required.

TELEVISION SIGNAL CHAIN

The television signal chain, shown below, intentionally includes signal non-linearities in both cameras and displays. The camera non-linearity is known as the opto-electronic transfer function (OETF) and the display non-linearity is known as the electro-optic transfer function (EOTF). Colloquially and confusingly in conventional television both transfer functions are known as “gamma curves”. As is well known the EOTF is not the inverse of the OETF, so overall the signal chain has a non-linear (or “system”) opto-optic transfer function (OOTF). The OOTF compensates for the psycho-visual effects viewing pictures on an emissive display in dim or dark surroundings, and is sometimes known as “rendering intent” (2). This is discussed in more detail below.

---

[1] Schreiber measured the threshold of contrast visibility experimentally (3). Schreiber’s results are broadly consistent with experiments on video quantization reported by Moore (4), and also with the DICOM model (5), which itself is derived from the Barten model (6)

Originally the OETF, in combination with the CRT display EOTF, was designed to make the effects of camera noise more uniform at different brightnesses. In digital systems the non-linearities help to minimise the visibility of quantisation (or "banding"). But, by modifying these non-linearities, it is possible to further reduce the effects of banding and so increase dynamic range.

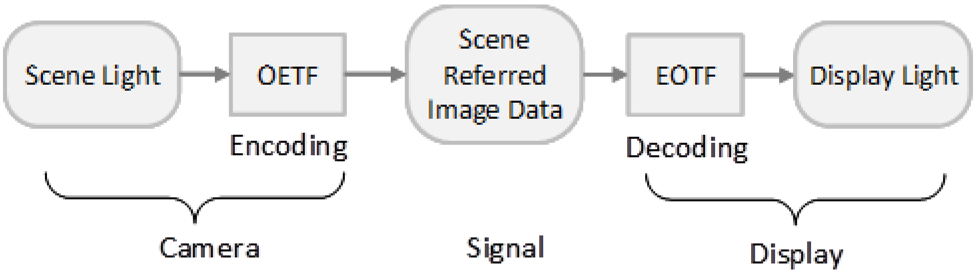

Figure 2: Television signal chain

In this conventional model of the television chain, the relative luminance in the scene is captured by the camera and encoded in the signal. The light from the scene defines the signal. The EOTF then renders the light from scene so that, *subjectively*, it appears the same as reality. Historically, with CRT displays, display brightness was fairly consistent because CRTs simply could not be made very bright. This allowed a single EOTF to be used to render the signal for all CRT displays. With the availability of a plethora of bright display technologies different EOTFs are needed to ensure that pictures look *subjectively* the same on displays of different brightness. The approach described in this paper allows the signal to be rendered on any display (OLED, LCD, local backlight dimming, or quantum dot), preserving the director's artistic intent, without the need for metadata and without needing to re-grade for different displays. That is, the hybrid log-gamma approach defines a signal that is independent of the display.

THE HYBRID LOG-GAMMA OPTO-ELECTRONIC TRANSFER FUNCTION (OETF)

In the brighter parts and highlights of an image the threshold for perceiving quantisation is approximately constant (know as Weber's law). This implies a logarithmic OETF would provide the maximum dynamic range for a given bit depth. Proprietary logarithmic OETFs, such as S-Log, Panalog and Log C are, indeed, widely used. But in the low lights it becomes increasingly difficult to perceive banding. That is, the threshold of visibility for banding becomes higher as the image gets darker. This is known as the De Vries-Rose law. The conventional gamma OETF comes close to matching the De Vries-Rose law, which is perhaps not coincidental since gamma curves were designed for dim CRT displays. So an ideal OETF would, perhaps, be logarithmic in the high tones and a gamma law in the low lights, which is essentially the form of the hybrid log-gamma OETF.

The dynamic range of modern video cameras is considerably greater than can be conveyed by a video signal using a conventional gamma curve (i.e. ITU-Rec 709). In order to exploit their full dynamic range conventional video cameras use a "knee" characteristic to extend the dynamic range of the signal. The knee characteristic compresses the image highlights to prevent the signal from clipping or being "blown out" (overexposed). A similar effect is also a characteristic of analogue film used in traditional movie cameras. When a hybrid gamma HDR video signal is displayed on a conventional SDR display the effect is similar to the use of a digital camera with a knee or using film. It is not surprising therefore, that the hybrid gamma video signal is highly compatible with conventional SDR displays, because what you see is very similar to the signal from an SDR camera. Indeed the knee characteristic of the hybrid gamma characteristic, defined below, is conservative, providing only 300% overload.

A hybrid gamma signal is defined as:

OETF:

$$E' = \begin{cases} r\sqrt{E} & 0 \le E \le 1 \\ a\ln(E-b)+c & 1 < E \end{cases}$$

where *E* is proportional to the light intensity detected in a camera colour channel (R, G, or B), normalized by the reference white level. *E'* is the non-linear, or “gamma corrected”, signal, where the non-linearity is applied separately to each colour channel. The reference value of E’ is 0.5, denoted “r”, and corresponds to reference white level. Constants, *a*= 0.17883277, *b*= 0.28466892, *c*= 0.55991073, are defined so that the signal value is unity for a (relative) luminance of 12.0.

The hybrid log-gamma OETF is shown below alongside the conventional SDR gamma curve and a knee characteristic. Note that the horizontal axis for the hybrid log-gamma curve, as defined above, has been scaled to emphasise compatibility with the conventional SDR gamma curve. Furthermore, because the hybrid log-gamma signal only describes the light representing the scene, it is independent of the display. Consequently, with a suitable EOTF, it may be used with any display.

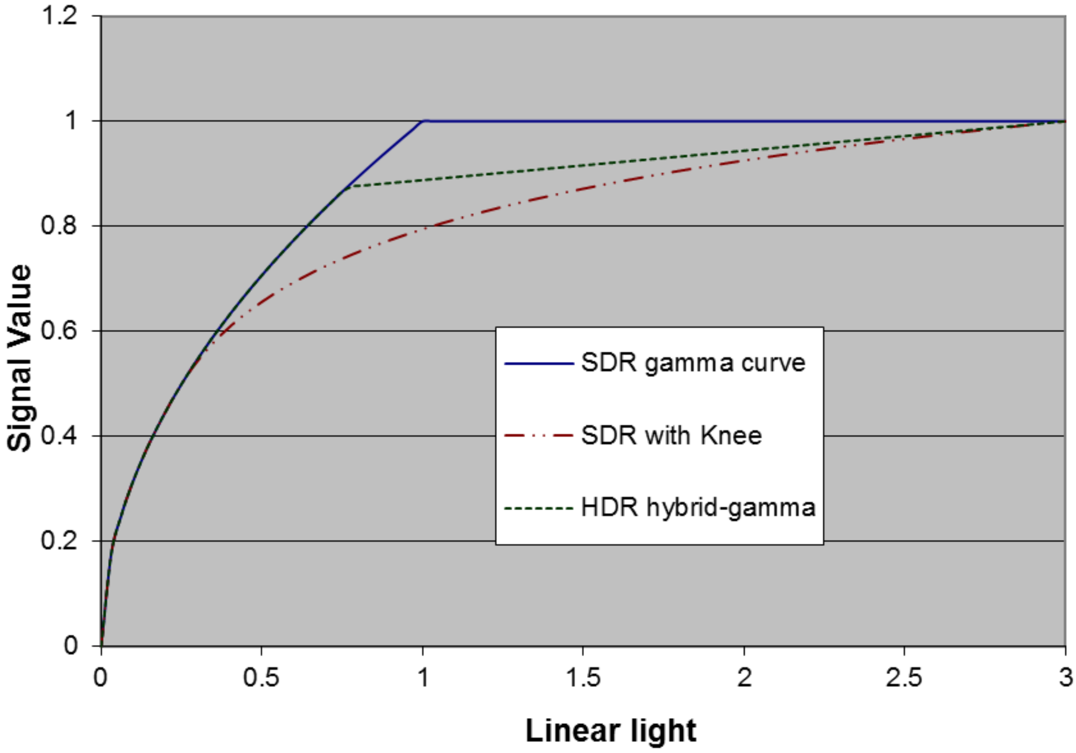

Figure 4: Hybrid log-gamma and SDR OETFs

SYSTEM GAMMA AND THE OPTO-OPTIC TRANSFER FUNCTION (OOTF)

As is well known, and noted above, the light out of a television display is not proportional to the light detected by the camera. The overall system non-linearity, or “rendering intent” (2) is defined by the opto-optic transfer function, or OOTF. Rendering intent is needed to compensate for the psychovisual effects of watching an emissive screen in a dark or dim environment, which affects the adaptation state (and hence the sensitivity) of the eye. Without rendering intent pictures would look too bright or “washed out”. Traditionally movies were, and often still are, shot on negative film with a gamma of about 0.6. They were then displayed from a print with a gamma of between 2.6 and 3.0: This gives movies a system gamma of between 1.6 and 1.8, which is needed because of the dark viewing environment. Conventional SDR television has an OOTF which is also a gamma curve with a system gamma of 1.2. But, for HDR, the brightness of displays and backgrounds will vary widely, and the system gamma will need to vary accordingly.

In order to determine the necessary system gamma we conducted experiments viewing images with different gammas at different luminances (and with a fixed background luminance). The pictures were derived from HDR linear light images selected from Mark Fairchild’s HDR Photographic Survey. A reference display (Dolby PRM4220) and a test display (SIM2) were placed about 1 metre apart, in a controlled viewing environment (room illumination 10Lux, D65). The reference image, shown at 600cd/m^2 on the reference display, was chosen by participants from 9 images with different gammas (1.0 to 2.4). This allowed them to choose the artistic effect they preferred. The participants then choose a picture on the test display, from 9 different gammas (1.0 to 2.4) that best matched the reference image. The test images were shown at different luminances on the test display. The results, illustrated below, provide an estimate of the preferred system gamma, (excluding artistic preferences), at a range of display brightnesses from 68 to 5200cd/m^2. Whilst only a small number of participants were involved, and further experimental results would be most welcome, the results are quite consistent and provide a good guide to the necessary system gamma for different display brightness relative to background luminance.

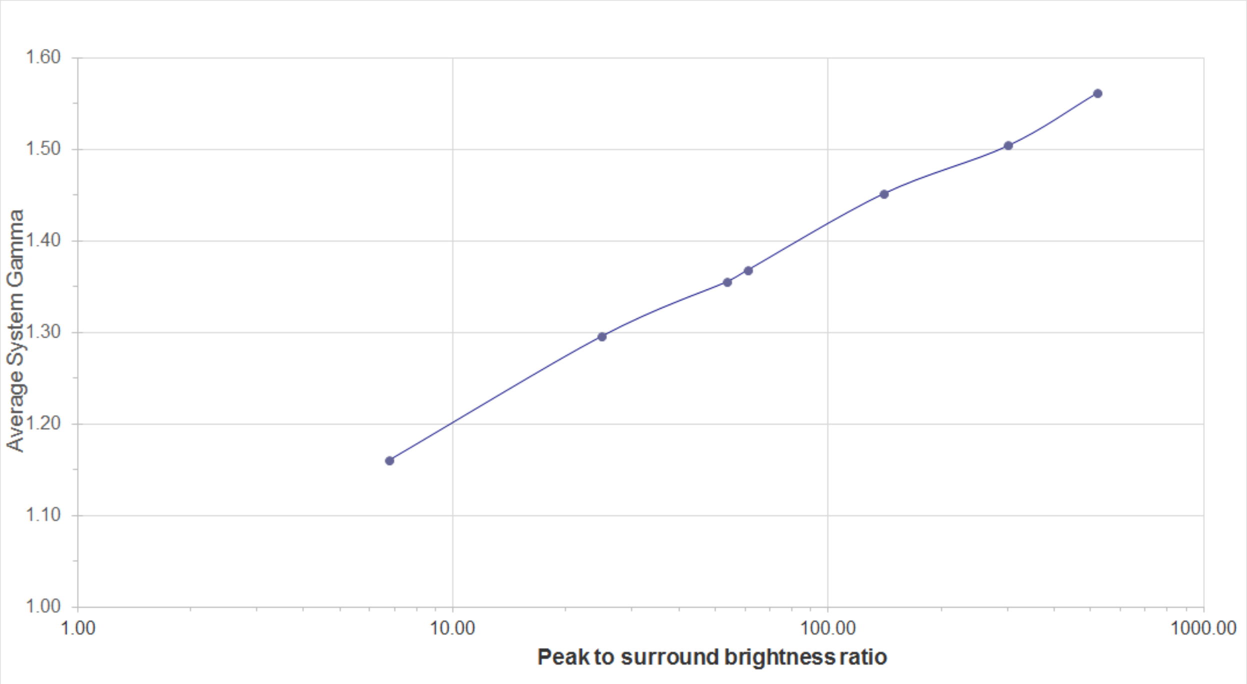

Figure 5: Preferred system gamma versus normalised display brightness

These empirical results may be approximated by the following formula for system gamma, where Y represents luminance.

$$\gamma = 1 + \frac{1}{5}\log_{10}\left(\frac{Y_{peak}}{Y_{surround}}\right)$$

The results clearly show that the end-to-end system gamma of the HDR TV system has to be adjusted to accommodate displays of differing peak luminance. They suggest that, with a background luminance of 10cd/m^2, an OLED display of around 1000 cd/m2 would require a system gamma of around 1.4, whilst a brighter LCD of a few thousand cd/m2 would require a system gamma closer 1.5. Whilst these variations in system gamma appear small, they have a significant impact on the subjective appearance of an image.

THE HYBRID LOG-GAMMA ELECTRO-OPTIC TRANSFER FUNCTION (EOTF)

In order to specify the complete television system we need an EOTF as well as the OETF defined above. This maps the relative light representing the scene to the light emitted from the display. The EOTF should perform this mapping 1) whilst preserving the artistic intent of the programme maker (and providing a suitable rendering intent), 2) allowing for the dynamic range of the display from black level to peak white, and 3) minimising quantisation artefacts. The EOTF defined below is similar to the conventional display gamma curve, thereby maximising backward compatibility, whilst also meeting the three preceding requirements;

EOTF:

$$Y_d = \alpha Y_s^{\gamma} + \beta$$

where Y_d is the luminance of a pixel presented on the display, Y_s is the relative luminance representing the scene for that pixel, and γ is the system gamma discussed above. Parameters α, and β correspond to similar parameters in ITU-R Rec 1886, which are the traditional “contrast” and “brightness” controls respectively. They determine the peak displayed luminance and the minimum luminance, i.e. the black level[2].

The EOTF maps the linear scene luminance, Y_s, to the linear display luminance, Y_d. This differs from current practice for SDR, which applies the EOTF to each colour component independently. But applying the EOTF to each component changes the saturation, and to a lesser extent hue, of the picture. Since the EOTF needs to change with the display it must be applied to luminance to avoid inconsistent colours.

Scene luminance Ys may be recovered from the signal by first applying the inverse of the OETF to each colour component R', G', and B' to yield the linear colour components R, G and B. With the same nomenclature as the OETF;

Inverse OETF:

$$E = \begin{cases} (E'/r)^2 & 0 \le E' \le r \\ \exp((E'-c)/a) + b & \mathrm{r} < E' \end{cases}$$

From the linear colour components the scene luminance may be derived as follows (assuming ITU-R Rec 2020 colorimetry);

Scene Luminance:

$$Y_s = \frac{0.2627R + 0.6780G + 0.0593B}{12}$$

Note that the factor of 12 in the denominator is because the signal normalisation for the OETF yields a maximum value of 12 for each linear colour component, rather than the

---

[2]
$$\alpha = L_P - L_B \quad \beta = L_B$$
where L_P is the displayed luminance for peak white (Y_s = 1.0), and L_B is the displayed luminance for black (Y_s = 0.0).

more conventional value of 1.

Having determined the linear scene luminance the displayed luminance may be derived from the EOTF, where parameters α, β, and γ depend on the display and the viewing environment. Given the displayed luminance we still need to determine the individual R_d, G_d, and B_d values that should be displayed for each pixel. We obtain these simply by scaling the linear scene colour components as follows:

Displayed Colour Components:

$$R_d = R \times \left((Y_d - \beta) / 12Y_s \right) + \beta$$
$$G_d = G \times \left((Y_d - \beta) / 12Y_s \right) + \beta$$
$$B_d = B \times \left((Y_d - \beta) / 12Y_s \right) + \beta$$

where R_d, G_d, B_d, are the luminances presented on the display.

Minimising the visibility of quantisation, or “banding” is an important aspect of the EOTF. We can estimate its visibility by calculating the weber fraction and comparing it to the Schreiber limit, as discussed above. Doing so and plotting the results yields the flowing graph for Weber fraction versus displayed luminance. The EOTF may be used for displays with different peak luminance and black levels. This example assumes a 10 bit signal, with peak white of 2000 cd/m^2, a black level of 0.01 cd/m^2, and a system gamma of 1.5. This corresponds to the use of bright displays for programme monitoring and grading, viewed in a “dark” environment relative to the brightness of the display. It represents a dynamic range of 200,000:1 or 17.6 stops, which is more than the dynamic range the eye can perceive in a single image.

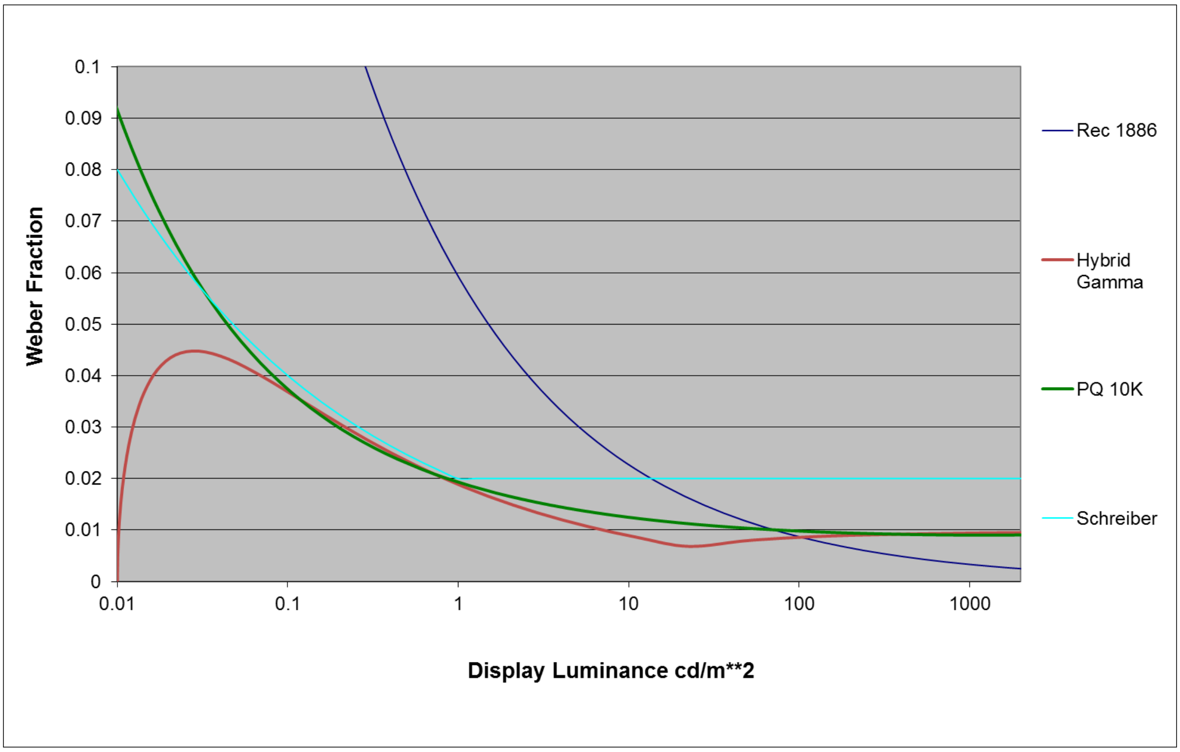

Figure 6: Weber fractions versus display luminance

For comparison this graph also includes the Schreiber limit, the conventional SDR gamma curve (Rec 1886), and an alternative HDR, a perceptual quantisation curve (PQ 10K) defined in SMPTE ST 2084. Banding is likely to be visible when the Weber fraction for an

EOTF is above the Schreiber limit. With a 10 bit signal this indicates that for the hybrid gamma EOTF banding will, at worst, be at the threshold of visibility across the whole luminance range and similar to or below that of the PQ curve. It also shows that the conventional gamma curve is not adequate, with banding expected to be visible below 20 cd/m^2. Note that this analysis is for a 10 bit signal. With a 12 bit signal, which has been proposed for HDR production, Weber fractions would be much lower and banding would be significantly below the threshold of detectability across the whole luminance range.

To confirm the theoretical analysis we performed experiments on the comparative visibility of banding. Highly critical 10 bit, horizontal, “shallow ramps”, with adjacent patches varying by 1 quantisation level, were compared to a continuous reference[3] using the ITU-R Rec 500 double stimulus impairment scale method. 33 subjects were tested. The test images were displayed on a Dolby PRM 4220 monitor configured (using a custom internal LUT) to emulate the low tones of a display with 2000cd/m^2 peak luminance and black level of 0.01cd/m^2. Each horizontal grey scale ramp occupied 25% of screen height, and included 20 adjacent grey levels each spanning 1/24th picture width. The experimental results are shown below. In these results -20 is one grade of impairment. The error bars indicate the 95% confidence intervals.

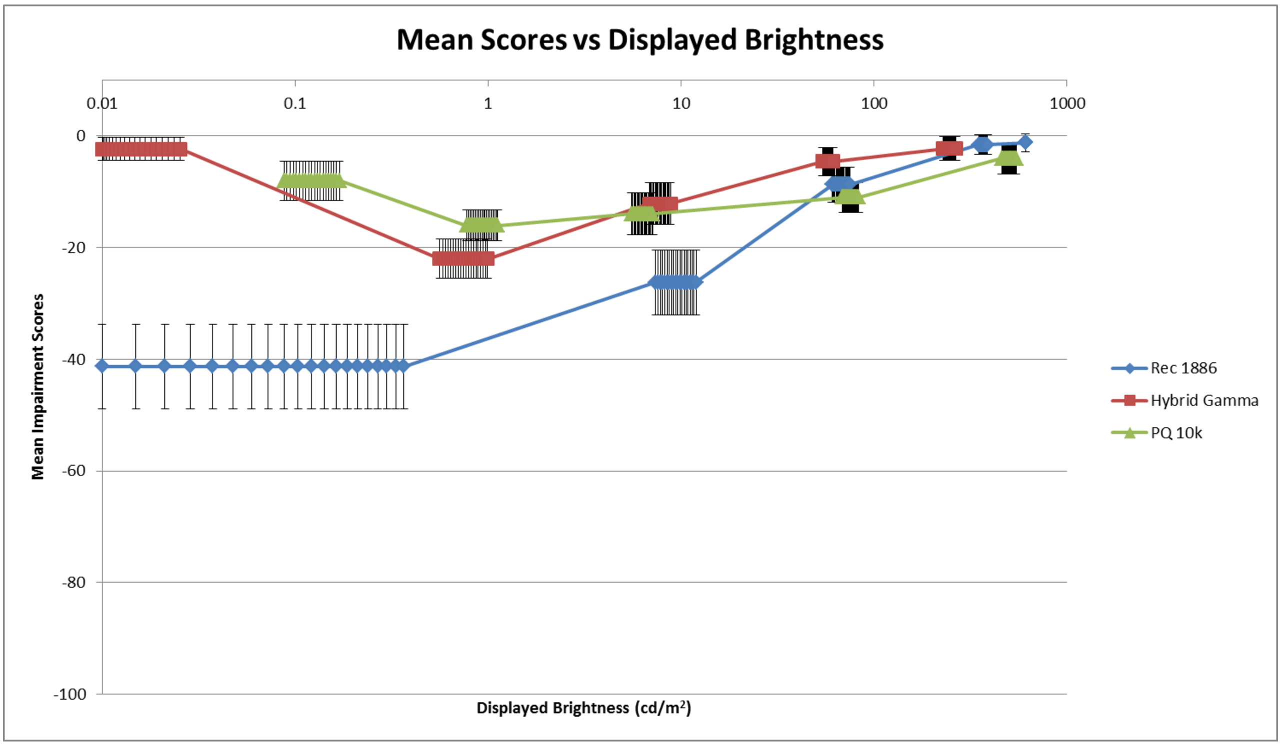

Figure 7: Banding impairment versus displayed brightness

These results appear to closely corroborate the theoretical analysis. Both the hybrid gamma and ST 2084 are less than or about 1 grade of impairment, which is described as “imperceptible” or just “perceptible but not annoying” at their very worst. Furthermore the hybrid gamma EOTF shows marginally more banding than ST 2084 in the region of 1cd/m^2, and marginally less banding elsewhere, which is in line with the theoretical analysis. The impairment for conventional SDR gamma (ITU-R Rec 1886) rises from

---

[3] The reference was a carefully dithered 10 bit signal in which banding was undetectable.

“imperceptible” to “perceptible but not annoying” below 20cd/m^2, up to “slightly annoying” in the region below 1cd/m^2, also in line with the theoretical analysis. Overall the hybrid gamma curve provides acceptable banding performance at 10 bits for highly critical material, equivalent to that of ST 2084, for a display with 17.6 stops of dynamic range. In practice it is highly unlikely that any banding will be visible on naturally occurring scenes.

CONCLUSIONS

This paper has presented the rationale and design for a HDR television system. The hybrid gamma approach can support a range of displays of different brightness, without metadata, and so is display independent. A 10 bit signal is substantially compatible with conventional SDR signals. Shown unprocessed on an SDR display the picture is of high quality and so may be used for signal monitoring. This also means production can use existing SDR infrastructure, tools and equipment. Only quality monitoring requires an HDR display. No metadata is required, thereby simplifying the production chain. These features mean SDR production may be upgraded to HDR at relatively modest cost. Only a single signal is required for both SDR and HDR displays and expensive multiple “grades” (for both HDR and SDR) are not necessary. Consequently layered, or multichannel, coding for end users is unnecessary, thereby simplifying distribution and minimising cost. For a 2000cd/m2 HDR display, with a black level of 0.01cd/m2, i.e. 17.6 stops dynamic range, it is shown, both theoretically and experimentally, that quantisation artefacts (“banding”) will not be visible on real pictures and that “banding” is comparable, or less, than the competing, more complex, ST 2084 HDR system. Finally note that HDR Rec 2020 signals may be formatted to look like conventional SDR Rec 709 signals, raising the possibility of conventional SDR media carrying HDR signals in a completely compatible way. This will be the subject of a future paper.

REFERENCES

1. Kunkel, T. and Reinhard, E., 2010. A reassessment of the simultaneous dynamic range of the human visual system. Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization. ISBN 978-1-4503-0248-7 , pp. 17–24. July 2010.

2. Poynton, C., 2012. Digital Video and HD.Morgan Kaufmann; 2nd edition edition (2 Dec 2012). ISBN-13: 978-0123919267

3. Schreiber, W. F., 1992. Fundamentals of Electronic Imaging Systems, *Third Edition*: Springer-Verlag, 1992, ISBN: 978-3-540-56018-0

4. Moore, T.A. 1974. Digital Video: The number of bits per sample required for reference coding of luminance and colour-difference signals. BBC Research Department Report. BBC RD 1974/42.

5. DICOM, 2008. DICOM.Digital Imaging and Communications in Medicine (DICOM), Part14: Grayscale Standard Display Function. National Electrical Manufacturers Association, PS3.14, 2008

6. Barten,P.G. J., 2004. Formula for the Contrast Sensitivity of the Human Eye. Proceedings SPIE-IS&T, 5294:231–238, Jan.2004.

ACKNOWLEDGEMENT

We would like to thank Manish Pindoria of BBC R&D for his work on this project, without which this paper would not have been possible.