THE ART OF BETTER PIXELS

D. G. Brooks

Dolby Laboratories, Inc., USA

ABSTRACT

The quality of colour television has been dominated by the physics of the Cathode Ray Tube (CRT) in determining the spatial resolution, the image dynamic range, the colour gamut and even the frame rate of the images displayed.

UltraHD (UHD) as specified in Recommendation ITU-R BT.2020 was the first standard to partially break free from these constraints, however it retained the 100 candela/m^2 CRT reference brightness limit and the corresponding Electro-Optical Transfer Function (EOTF) based on the gamma characteristic of the CRT.

In April 2012, the United States made a submission to ITU-R Working Party 6C which proposed a new EOTF in order to enable the carriage of High Dynamic Range (HDR) content through the 10 and 12 bit baseband interfaces specified in the then draft version of ITU-R BT.2020. The new perceptual EOTF proposed was based not upon the gamma function of a CRT, but directly on the contrast sensitivity ratio of the human eye as measured by Barten and referenced in Report ITU-R BT.2246-2. The submission further proposed that, based on viewer preference testing, an HDR system should be capable of handling signals with a brightness range from 0 to 10,000 candela/m^2.

INTRODUCTION

This paper gives an overview of the key experiments used to determine the parameters required for baseband video combining both HDR and WCG or EDR (extended dynamic range) for short.

Furthermore, it provides an insight into the EDR content creation process and proposes how the creative intent can be propagated throughout the distribution chain all the way to the TV receiver, to provide the viewer with the best possible experience on a wide range of EDR capable TV receivers.

It’s hard to exaggerate the impact that the cathode ray tube (CRT) had and still continues to have on the design of television standards. Our CRT heritage affects our HDTV digital standards in a number of key ways

- The colour gamut (as defined by the rare earth phosphors used)
- The brightness was limited to ~100 candela/m^2 (often referred to as “nits”) not only to control large area flicker, but also to prevent the electron beam from spreading and reducing spatial resolution.

LOOKING TO THE FUTURE

There are 3 fundamental ways to improve moving image quality

- Increase spatial resolution – more pixels
- Increase the frame rate – faster pixels
- Increase the capability of each pixel to represent more colours and increased dynamic range – better pixels.

The ITU-R UHDTV Recommendation BT.2020 (1) published in August 2012 offered both more and faster pixels with an increased colour gamut, but did not consider the need for pixels with an increased dynamic range and brightness.

This vestige of the CRT lives on despite the increasing interest in High Dynamic range systems in still photography.

What dynamic range is required for Entertainment content?

Dolby developed the experimental display shown in Figure 1 to answer this question. This display consists of a digital cinema projector pointed at a 23" monochrome LCD panel. The digital projector image and the LCD panel are dual modulated to create a display capable of a black level of 0.004nits and a peak white level of 20,000nits.

Figure 1 - Experimental HDR Display

Tests were performed to find out the preferred viewer experience for

- Black level
- Diffuse White level
- Highlight level

Content was carefully chosen to ensure that the image contrast ratio did not change while testing with average luminance level and largely achromatic images were used to avoid viewer preferences for more contrasty or colourful images.

The full details of this experiment can be found in the papers published by Daly et al. (2) Daly et al. (3) and Daly, Kunkel and Farrell (4)

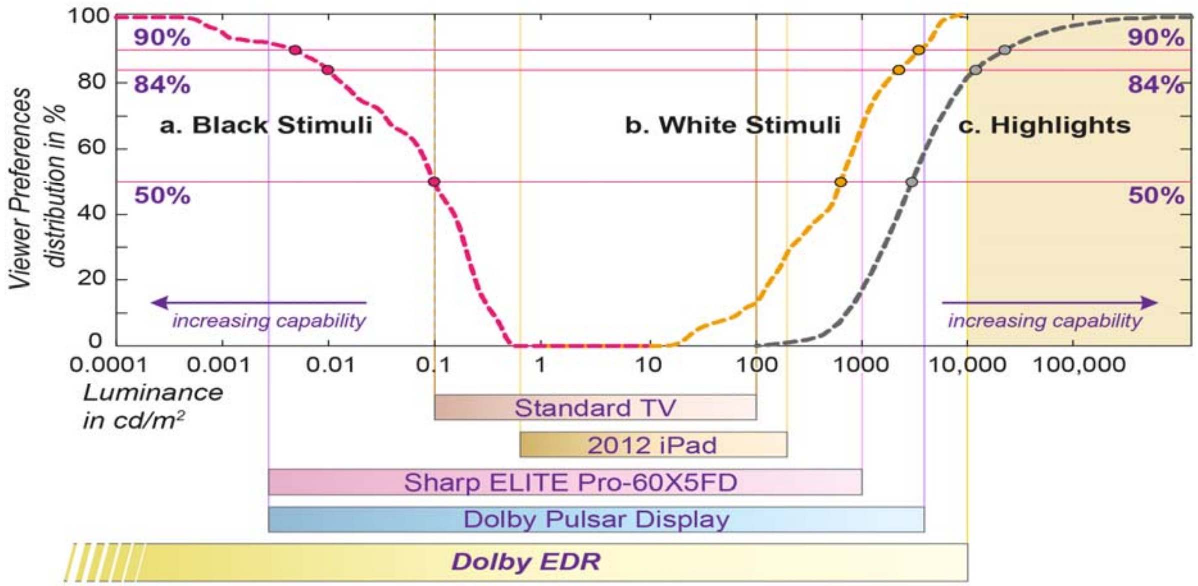

Figure 2 - Summary of Dynamic Range Subjective Tests

From Figure 2, it is evident that the current TV standards are inadequate as regards both dynamic range and brightness, with consumer preferences being orders of magnitude greater than today's television systems can provide.

Since the above results were first published, there have been questions as to whether these diffuse white and highlight levels are equally applicable to larger displays. An additional experiment was performed, with the same stimuli, using a 4 metre screen and a digital cinema projector. In both experiments, the viewer was positioned 3 picture heights from the screen, the results of the later experiment is compared with small screen results in Figure 3.

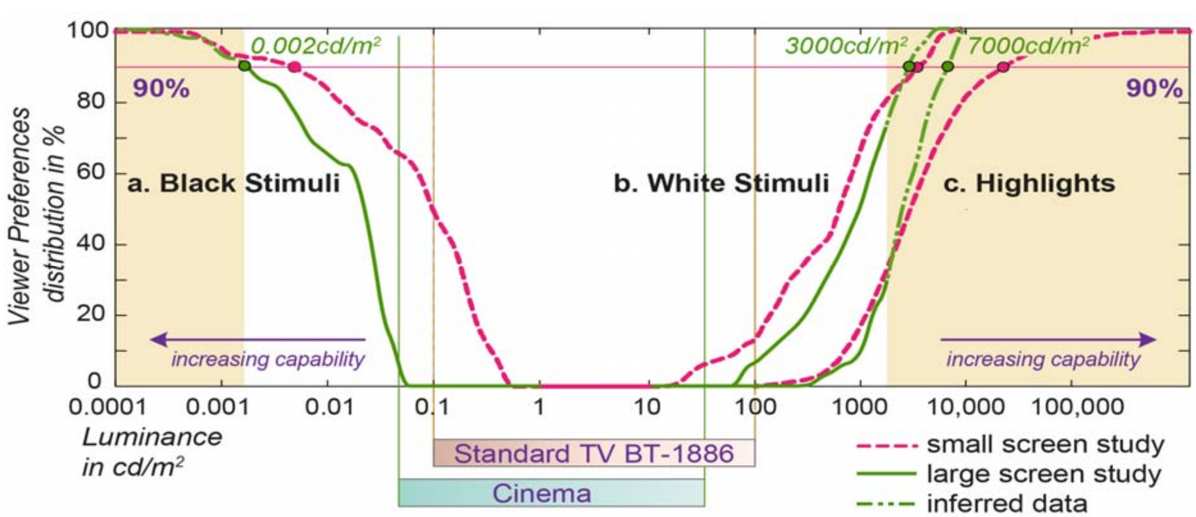

Figure 3 - Summary of Large Screen Dynamic Range Subjective Tests

	Luminance level to satisfy 90% of Viewers	
	Large Screen	Small Screen
Black	0.002 nits	0.005 nits
	21 F-stops	20 F-stops
Diffuse White	3,000 nits	4,000 nits
	1 F-stop	2 F-stops
Highlights	7,000 nits	20,000 nits

Table 1 - Summary of Dynamic Range Experiments

Table 1 summarizes these experiments and shows that a dynamic range of 22 F-stops is needed to meet 90% of viewer preferences for both small and large screen applications. One can further conclude that a television system which is capable of delivering content with a dynamic range from 0.001 nits to 10,000nits would satisfy the vast majority of viewers on a range of consumer devices from tablets to very large LCD displays.

COMBINING HIGH DYNAMIC RANGE WITH WIDER COLOUR GAMUT

Recommendation ITU-R BT.2020 (BT.2020) provides next generation standard dynamic range TV systems with a much wider colour gamut than current Recommendation ITU-R BT.709 (BT.709) (5), consequently any new television system must be capable of delivering both wider colour gamut and high dynamic range content.

Colour Volume

Traditionally we have used the colour "horseshoe" diagram to represent the colour gamut of a television signal such as shown in Figure 4, however for each colour shown there is a corresponding maximum luminance. White is the brightest colour, as television uses an additive colour system.

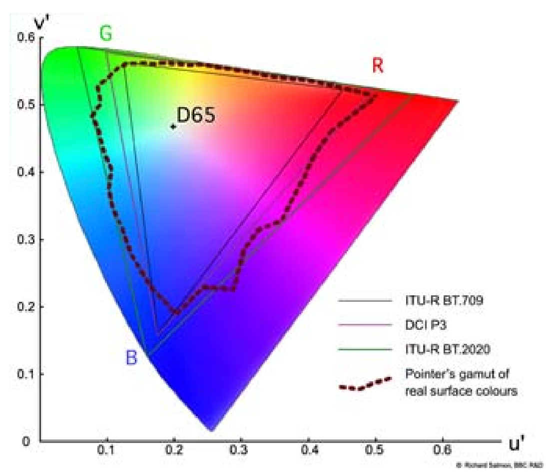

Figure 4 - The Colour "Horseshoe"

An alternative representation, which allows both the colour gamut and the dynamic range to be represented, is shown in Figure 5. Figure 5 also shows both today's BT.709 colour space and the new EDR colour space encompassing both the high dynamic range requirement from 0 to 10,000nits and the colour gamut as defined in BT.2020.

This representation has the advantage of showing all possible colours at all available luminance levels.

Having now defined the required colour volume needed to satisfy the vast majority of viewers, how can this be represented in a practical television system?

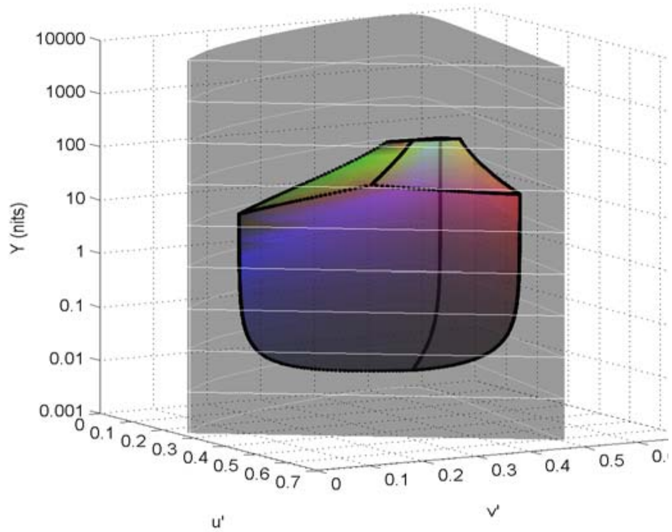

Figure 5 – Colour Volume EDR v. BT709

EDR Colour Volume Quantization

What bit depth is required to accurately represent this EDR colour volume shown in Figure 5? To answer this question requires an investigation into both luminance and chrominance elements.

Luminance Quantization

Report ITU-R BT.2246-2 (6) makes use of Barten's model for the contrast sensitivity function of the human visual system to ensure that the contouring artefacts in the BT.2020 UHDTV standard are acceptable.

Using the existing Electro-Optical Transfer Function (EOTF) of Gamma 2.4, as defined in Recommendation ITU-R BT.1886 (7), would require a 15bit representation to match the contrast sensitivity or contouring performance of the human visual system as shown in Figure 6.

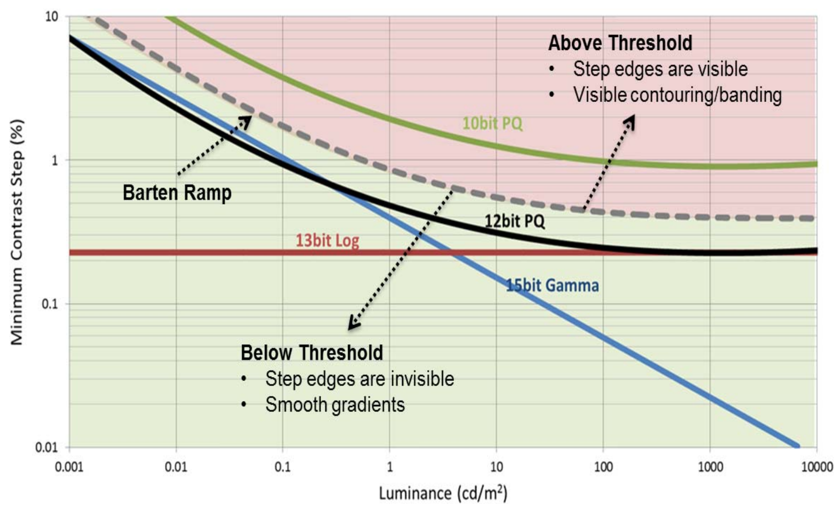

Figure 6 - Barten's Model over 10,000nit range

However, for a luminance range of 10,000 nits, Gamma 2.4 does not accurately match the contrast sensitivity function of the human eye. Gamma wastes bits coding bright areas which are well below the Barten threshold. Alternatively, using a 13bit Log representation wastes bits in the darker regions.

To resolve these issues a new Perceptual Quantizer (PQ) (8) was developed which follows the Barten curve to ensure the optimum use of the bits available. From Figure 6, it can be seen that 12bit PQ is below the visible threshold for contouring artefacts whereas 10bit PQ is above.

Colour Quantization

Whilst using PQ as the EOTF can provide the solution for monochrome images, an additional metric is required to ensure that colour images are also artefact free. The JND Cross test, Figure 7, was developed for this purpose.

Each square in Figure 7 represents a block of pixels which have been perturbed by 1 code word away from the grey background. All possible variations of this in RGB result in 26 different colour patches.

If the image looks like a flat grey image, then there is no visible quantization. If some patches are visible, then some contouring or banding is likely.

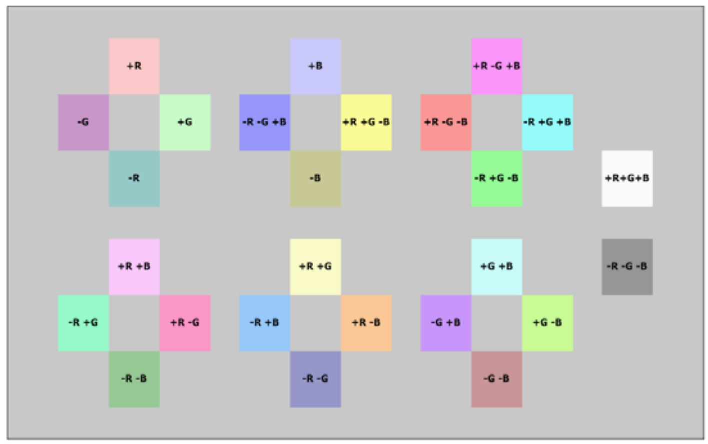

Figure 7 - The JND Cross (colours exaggerated)

This test is then performed for multiple different grey levels. The industry accepted metric to judge the visibility of these patches is CIEDE2000 (9) colour difference formula.

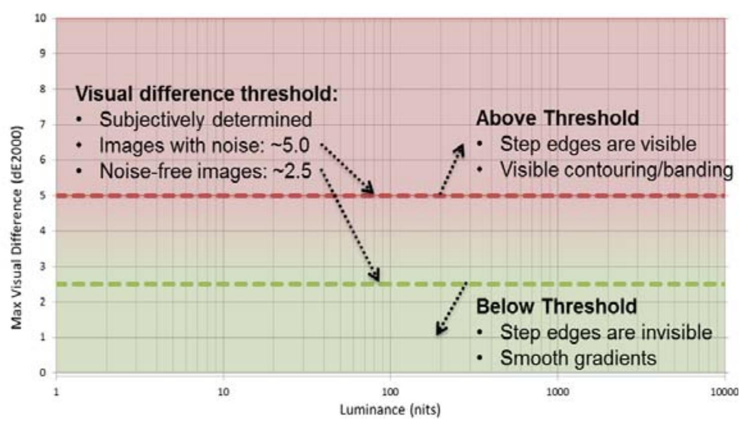

Figure 8 - CIEDE2000 overview

Subjective tests by Nezamabadi et al. (10) indicate that for noise free images such as animation, CGI and graphics a DE2000 threshold of ~2.5 is needed to avoid contouring artefacts. When noise is present, this threshold increases to ~5.

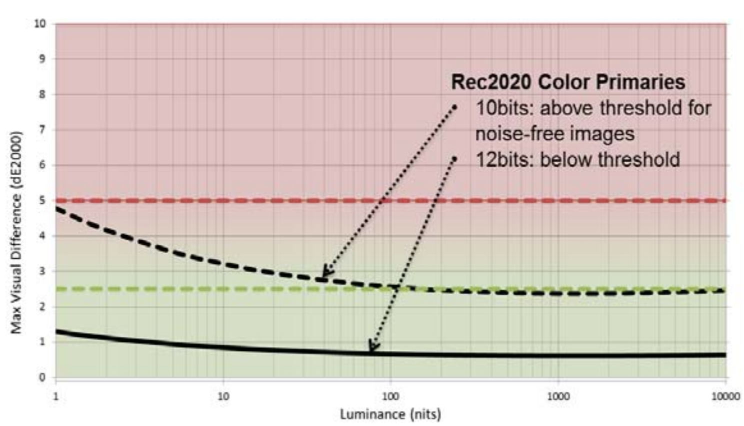

Figure 9 – EDR evaluation using DE2000

It should also be noted that the DE2000 metric starts to loose accuracy below about 3nits; this corresponds to the point where the human visual system begins to switch between photopic to scotopic vision.

By combining the use of this metric with the JND colour patches allows the quantification of possible colour quantization artefacts. The results are shown in Figure 9.

EDR Colour Volume Conclusions

For noise free EDR images such as animation, CGI or graphics, using 12bit PQ ensures that both monochromatic and colour banding artefacts are below the visible threshold. This matches the MovieLabs specification for next generation content (11).

For captured images that contain some noise, using a 10bit PQ representation is adequate, but care needs to be taken especially with graphics. However, this is an area well understood by broadcasters who daily deal with these issues for today's 8bit BT.709 transmissions.

Based on the above studies, a detailed submission was made in July 2012 by the United States to ITU-R Working Party 6C proposing the PQ EOTF for High Dynamic Range / Wider Colour Gamut images (12). A Rapporteur Group was subsequently set up to examine the options and work in this group is on-going. In parallel with this effort, the PQ EOTF is in the process of being standardized by the Society of Motion Picture and Television Engineers (SMPTE) (8)

GETTING EDR CONTENT TO THE HOME

Having established the definition for EDR content, how best can this content be transmitted to the home and presented to the viewer?

The following criteria were defined to meet this requirement:

- Single production workflow required for both SDR and EDR content
- Compatibility with existing off-line and real-time infrastructures
- The highest quality images, matching the creative intent, should always be available to the home viewer
- Image quality will increase, in both professional and consumer applications, as display technology evolves, the transmission system must not be the bottleneck
- The transmission system should allow backwards compatibility with today's BT.709 or BT.2020 systems
- Should be independent of any spatial resolution, colour gamut or frame rate
- The EDR transmission system should use industry standard codecs
- A bitrate efficient transmission method is required for EDR
- New EDR TVs must be capable of mapping the transmitted EDR images to the TV display's native colour volume (defined by the display's black level, peak white level, colour temperature and colour primaries) thereby remaining as true as possible to the original creative intent within the confines of the display's capabilities

An end-to-end system overview

Figure 10 provides an end-to-end architecture / workflow for off-line content creation and distribution which meets the requirements described above.

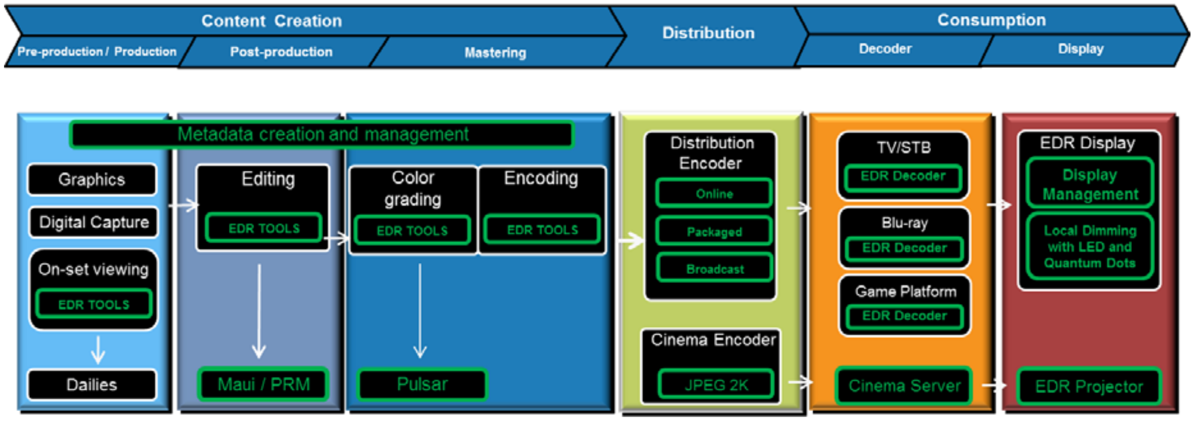

Figure 10 - Off-line EDR System Architecture / Workflow

The ability to create EDR content has been limited, not by today's cameras, but by existing 100nit broadcast reference monitors. The development of the Pulsar monitor, which has a peak brightness of 4000nits with a DCI-P3 colour gamut, plus a set of EDR tools as plug-ins to commonly used off-line production suites meet the needs of the creative.

In parallel with this process, image metadata is automatically generated and stored along with any creative input. This metadata is used downstream to guide the display management block, based on both automatic and creative input captured during production, in the TV receiver. The definition of this metadata has begun and is now in the process of becoming a SMPTE standard (13).

To meet the practical requirements of backwards compatibility with today's BT.709 and future BT.2020 transmission systems, a dual layer codec architecture has been developed which allows the existing standards to be transmitted as the base layer. An enhancement layer contains the information necessary to recreate the EDR signal. Figure 11 shows a block diagram of this architecture. It is worth noting that this dual layer codec works outside of the coding loop (unlike MVC or SVC, for example) so that no changes are required to either encoders or decoders.

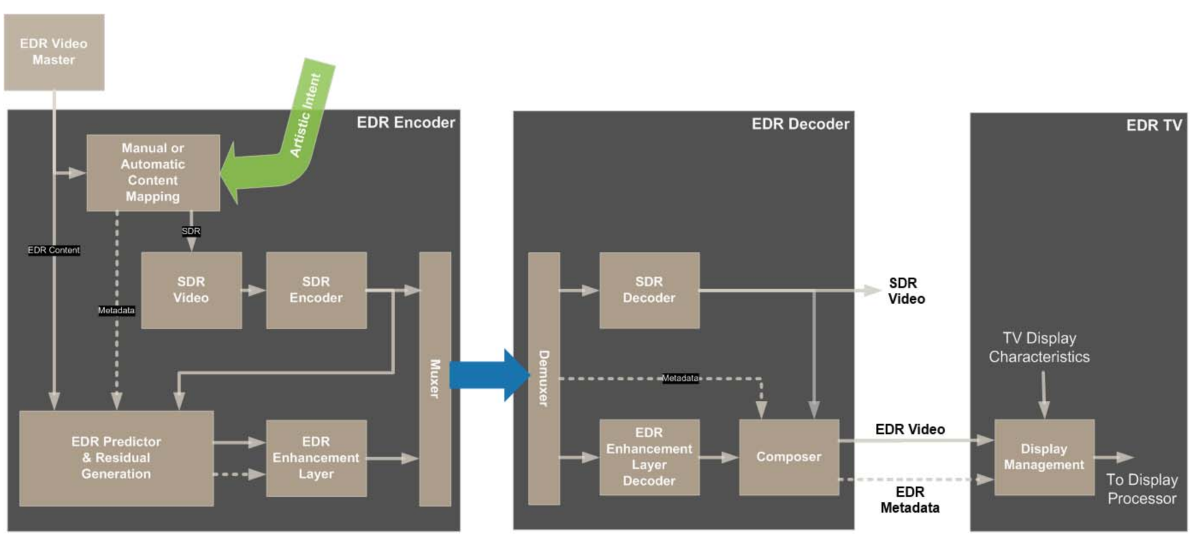

Figure 11 – The EDR Dual Layer Codec

Moreover the dual layer codec as shown in Figure 11 allows for either an 8bit AVC or an 8/10bit HEVC base layer for compatibility with existing services. Similarly the EDR input signal can be either 10bit PQ or 12bit PQ and either will be faithfully reconstructed at the dual decoder output.

Using this dual layer codec also allows existing TVs to receive the base layer standard dynamic range signal, whilst new EDR televisions can receive the EDR signal and adapt the content to match the capabilities of the display. The EDR enhancement layer only increases the bitrate by ~25% compared with the backwards compatible base layer.

The capabilities of the EDR TV will evolve over time as new technologies such as quantum dots allow wider colour primaries and brighter displays. By the use of a combination of television's inherent display characteristics plus the EDR metadata transmitted, the display management function inside the TV is able to optimally adapt the incoming EDR images to match each display's characteristics whilst maintaining the creative intent.

CONCLUSION

This paper outlines the experiments made to determine the image dynamic range required for television and cinematographic entertainment purposes.

Using this information, a series of tests and experiments were performed to determine the baseband bit depth required to ensure that no contouring artefacts would be visible.

In making these tests, the traditional TV gamma 2.4 non-linear curve was proven to be inefficient for high dynamic range images and a new Perceptual Quantizer was designed to match the characteristics of the human visual system.

Both 10 and 12 bit PQ EDR signals can be encoded by the dual layer codec proposed. The SDR base layer can be coded as either 8 or 10 bit depending on the backwards compatibility requirements.

Existing SDR TV receivers will ignore the enhancement layer and display the SDR content as today.

New EDR receivers will decode both base and enhancement layers and combine these together to faithfully reproduce EDR images at either 10 or 12bit PQ depending on the source. The TV receiver will also incorporate a new display management block which maps the content to the evolving characteristics of each display by using the metadata present in the EDR stream.

REFERENCES

1. Recommendation ITU-R BT.2020 "Parameter values for ultra-high definition television systems for production and international programme exchange", August 2012

2. S. Daly, T. Kunkel, X. Sun, S. Farrell, and P. Crum (2013) "Preference limits of the visual dynamic range for ultra high quality and aesthetic conveyance" SPIE Electronic Imaging, HVEI, Burlingame, CA, Feb 2013

3. Scott Daly, Timo Kunkel, Xing Sun, Suzanne Farrell, and Poppy Crum (2013) "Viewer Preferences for Shadow, Diffuse, Specular, and Emissive Luminance Limits of High Dynamic Range Displays" , SID Display Week, paper 41.1 , Vancouver, Canada.

4. Scott Daly, Timo Kunkel, and Suzanne Farrell (2013) "Using Image Dimension Dissection to Study Viewer Preferences of Luminance Range in Displayed Imagery",

Invited talk, IDW, International Workshop on Displays , SID, Sapporo , Japan , Dec. 6, 2013

5. Recommendation ITU-R BT.709-5 “Parameter values for the HDTV standards for production and international programme exchange”, April 2002
6. Report ITU-R BT.2246-2 “The present state of ultra-high definition television”, November 2012
7. Recommendation ITU-R BT.1886 “Reference electro-optical transfer function for flat panel displays used in HDTV studio production”, March 2011
8. Draft SMPTE Standard 2084 “High Dynamic Range Electro-Optical Transfer Function of Mastering Reference Displays”
9. CIE Publication 142:2001 “Improvements to industrial colour difference evaluation”, Commission Internationale de l’Eclairage, Vienna, Austria (2001)
10. M. Nezamabadi, S. Miller, S. Daly, and R. Atkins “Color signal encoding for high dynamic range and wide color gamut based on human perception.” Color Imaging at SPIE’ s Electronic Imaging Conference (2014).
11. MovieLabs “Specification for Next Generation Video” – Version 1.0, 2013
12. Input document USA002 to ITU-R Working Party 6C, “Parameter values for UHDTV systems for production and international programme exchange” April 2012
13. Draft SMPTE Standard 2086 “Mastering Display Color Volume Metadata Supporting High Luminance and Wide Color Gamut Images”

ACKNOWLEDGEMENTS

The author would like to thank his colleagues for all their work in developing EDR (Dolby Vision) from Image Research & Development in Burbank, Sunnyvale, San Francisco and Yardley. I would also like to thank my wife, Naomi, for encouraging me to write this paper over a beautiful long Spring weekend.

Figure 4 is courtesy of Richard Salmon at the BBC.