CHARACTERIZATION OF PROCESSING ARTIFACTS IN HIGH DYNAMIC RANGE, WIDE COLOR GAMUT VIDEO

O. Baumann, A. Okell, J. Ström

Ericsson

ABSTRACT

A new, more immersive, television experience is here. With higher resolution, wider color gamut and extended dynamic range, the new Ultra High Definition (UHD) TV standards define a container which allows content creators to offer the consumer a much more immersive visual experience.

Although very little content exploiting the full range of the container is yet available, some artifacts associated with the compression of high dynamic range (HDR) content have already been identified and reported in the literature. Specifically, the chroma subsampling process has been shown to cause disturbing artifacts for image regions of certain color and luminance.

This paper quantifies the distortion and identifies regions of the extended color volume where artifacts associated with standard image processing techniques are more likely to occur. In doing so, it highlights that the problems will become greater as more content exploiting the full UHD container becomes available, requiring additional care and processing in content production and delivery. Finally, the paper references ways of overcoming these issues.

INTRODUCTION

The DVB UHD1 Phase 2 (1) and ATSC 3.0 (2) standards are becoming well established and allow significant benefits over High Definition television standards beyond simply more pixels. The technological enhancements of High Dynamic Range (HDR), Wide Color Gamut (WCG) and High Frame Rate (HFR) all contribute to more life-like images and hence a more immersive viewing experience.

ITU-R Recommendation BT.2100 (Rec. 2100) (3) defines parameters and formats for HDR television, using a minimum $Y'CbCr$ component bit-depth of 10-bits, color primaries from ITU-R Recommendation BT.2020 (Rec. 2020), and two alternative Electro-Optical Transfer Functions ($EOTF$s): Perceptual Quantizer (PQ, standardized as SMPTE ST 2084) and Hybrid Log-Gamma (HLG, standardized as ARIB STD-B67).

This paper examines a system using the PQ transfer function, but many of the principles will also apply to HLG. The PQ transfer function is specifically designed to exploit the relative insensitivity of the human visual system to absolute differences in luminance when the luminance is high. It achieves this by allocating more code words to lower values of the

red, green and blue component signals, than to higher values (5), and as such is much more non-linear than traditional SDR transfer functions.

A typical signal processing chain to convert linear light RGB camera data to Rec. 2100 4:2:0 $Y'CbCr$ format using the PQ transfer function is shown in Figure 1. Conversion from $Y'CbCr$ 4:2:0 format to linear light RGB for display is the reverse processing chain, with chroma upsampling, and $EOTF$s applied to R', G' and B'.

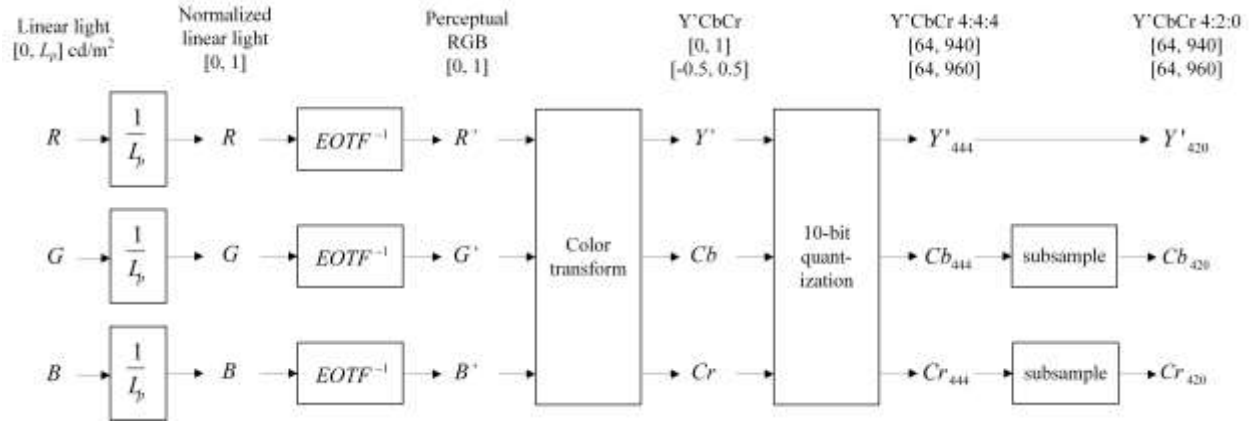


Figure 1 – Typical processing chain to convert linear light RGB camera data to Rec. 2100 Y'CbCr 4:2:0 format

The non-linearity associated with the PQ transfer function can result in significant chroma leakage for some colors, meaning that the luma component Y', does not completely represent the desired luminance output from the display, but part of that signal is instead carried in the Cb and Cr components. This in turn means that errors introduced on the chroma components by, for example, chroma subsampling, can result in visible distortion in the output luminance (6). It is the study of this distortion, where it occurs, and how it can be avoided, which forms the basis of this paper.

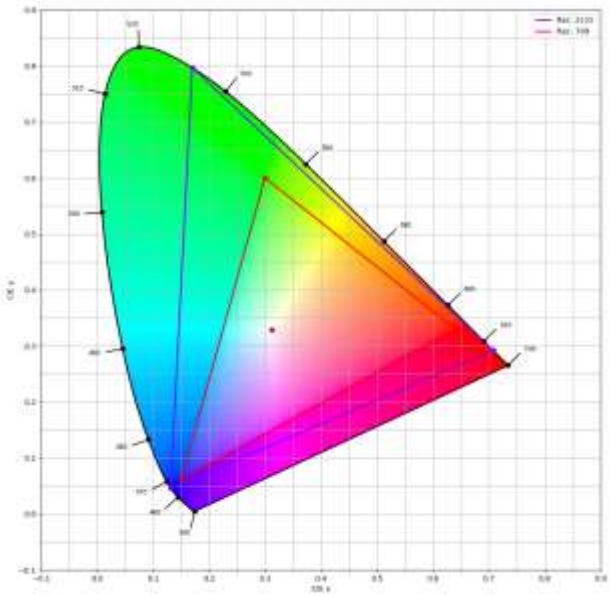


Figure 2 – Plots of the extent of the Rec. 2100 and Rec. 709 color spaces in CIE xy

A Note on Color Spaces

The Rec. 2100 and ITU-R Recommendation BT.709 (Rec. 709) color spaces are shown in the CIE 1931 xy coordinates along with the familiar horseshoe of the visible spectrum in Figure 2. It is worth noting that the chromaticity points, and therefore the color space of Rec. 2100 are identical to those of Rec. 2020 (3). Clearly, Rec. 2100 allows a far greater proportion of visible colors to be represented than Rec. 709. In

this paper, we use the xyY color space for analyzing the properties of visible light. The x and y components define the hue and purity, and the Y component the luminance in candelas-per-meter-squared (cd/m^2) (4).

CHROMA SUBSAMPLING AND LUMINANCE ERRORS

It is well understood that the human visual system has a lower acuity to chrominance than to luminance and as such the Cb and Cr components of the $Y'CbCr$ container are often subsampled either vertically as in 4:2:2, or both horizontally and vertically as in 4:2:0 (4). To avoid aliasing in the reconstructed chroma components, the full resolution chroma components are low pass filtered prior to decimation.

This filtering process is a weighted average of neighboring samples. Let's take an example, previously reported in (6), of two neighboring pixels in linear light RGB. Both pixels are very similar color and luminance:

$$RGB_1 = (1000, 0, 100)$$

$$RGB_2 = (1000, 4, 100)$$

The two colors are shown in the top two squares of Figure 3. It should be noted that all four squares in the figure have been identically scaled in order to see them in a non-HDR workflow.

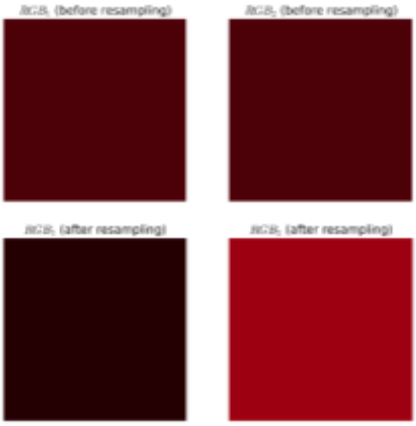


Figure 3 - Squares of plain color scaled from the RGB values of a worked chroma resampling example

Converting these neighboring pixels to $Y'CbCr$ and back to RGB as per the workflow presented in Figure 1 results in the following RGB values:

$$RGB_1 = (484, 0.03, 45)$$

$$RGB_2 = (2061, 2.2, 216)$$

Clearly these are quite different, not only from one another but each is also different from the original color. This can be seen from the lower two squares of Figure 3. We can look at the difference in luminance in the CIE color space by converting to xyY:

$$xyY_1 = (0.6405183, 0.26326786, \mathbf{129.9302})$$

$$xyY_2 = (0.63216754, 0.26031566, \mathbf{555.785})$$

Whilst the hue and purity have remained close to one another, the luminance is significantly different. In real images, this effect manifests itself as adding noise artifacts in chroma subsampled images. It has been previously observed in MPEG by Lopez et al. (7) and François (8). Examples from the *Market* sequence are shown in Figure 4. Since the original capture was in Rec. 709 color space, a Rec. 709 container has been used for this sequence in order to simulate Rec. 2100 content that fills up the Rec. 2100 color space.

It has been reported that the distortion induced is greatest for colors at the edges of the color space (8), i.e., for more saturated colors, and (9) provided some examples, but to the authors' knowledge no systematic analysis of where the distortion occurs has yet been undertaken.

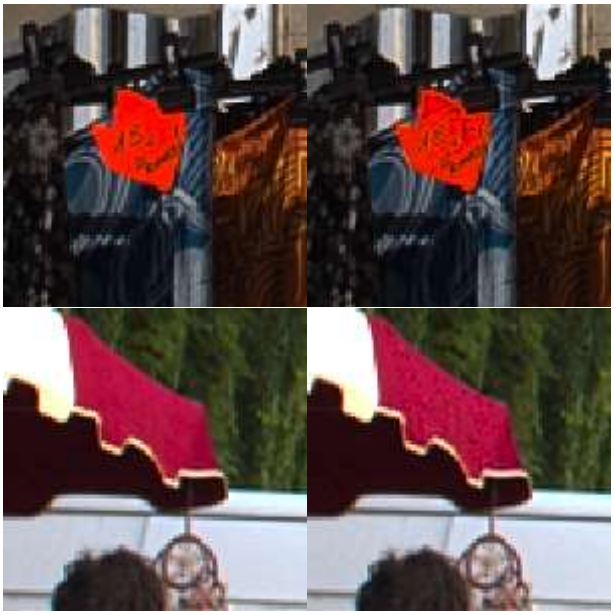


Figure 4 – Left column: Original 4:4:4. Right column: After conversion to 4:2:0 and back. Image sequence courtesy of Technicolor and the NevEx project.

IDENTIFYING WHERE ERRORS OCCUR

Ideally, we seek to identify an error surface in the xy plane of CIE 1931, showing how the distortion varies throughout the color space. This has two complicating factors:

1. The distortion is not simply a function of the color of a single pixel, but the combination of neighboring pixels.
2. The region of the color space with greatest distortion may change as a function of luminance.

An exhaustive search of every combination of neighbors would be not only time consuming, but very difficult to visualize.

Instead, we create a synthetic image, or "swatch", of 16x16 pixels, consisting of a single color of known linear light RGB value (and therefore known xyY value) and add zero-mean Gaussian noise. No attempt has been made to simulate the characteristics of camera or film noise. The resulting swatch is referred to as the reference swatch. We apply the Rec. 2100 processing chain to obtain 4:2:0 $Y'CbCr$ data. Converting back to linear light RGB results in what we refer to as the reconstructed swatch. We are then able to compare the reference and reconstructed swatches either by viewing or by measuring the error objectively. Most of the processing in this analysis is performed in Python using the open source *Colour Science* library (10).

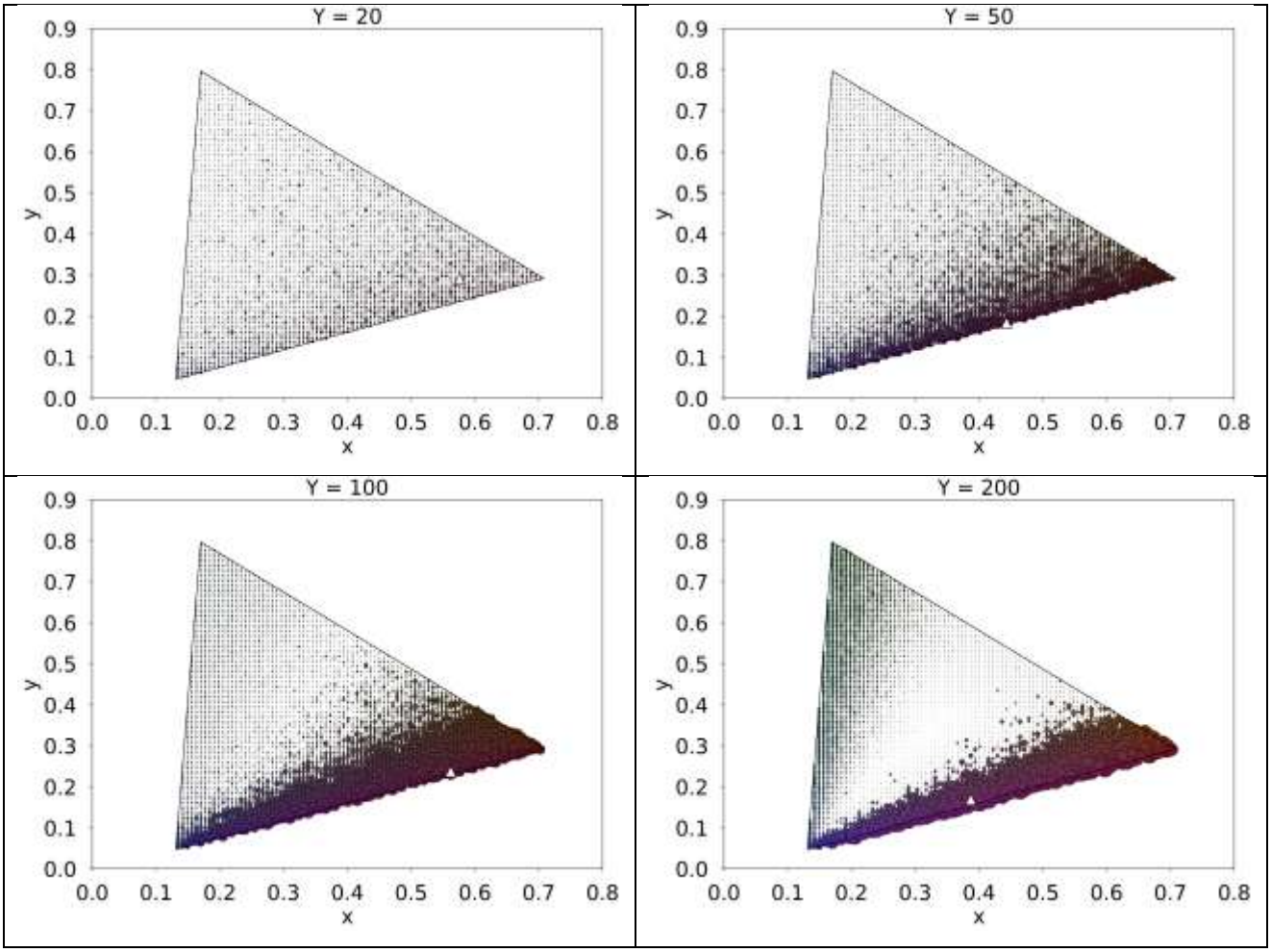
Objective Analysis

We have analyzed values of $Y \in \{20, 50, 100, 200, 500, 1000, 2500, 5000, 7500, 9000\}$, and for each value, we define a linearly spaced grid of 100×100 points in xy. For each point on this grid that results in a “legal” linear light RGB value, we create a swatch and perform the processing described above. By “legal” we mean that the xyY triple corresponds to a Rec. 2100 normalized linear light RGB triple with each component in the interval $[0,1]$. This allows us to visualize how the error changes through an xy plane for several “slices” of Y.

For the objective analysis of the error, we convert both swatches from linear light RGB to xyY, and calculate the sum of the squared difference (SSD) between the luminance samples of the reconstructed swatch, $\hat{Y}_i$, and the luminance samples of the reference swatch, Y_i:

$$SSD = \sum_{i=1}^{256} (\hat{Y}_i - Y_i)^2$$

This provides a single value of the expectation of the error for every color in the color space which we refer to as the reconstruction error. Figure 5 shows xy planes for the set of luminance values, Y, described above. The reconstruction error for a swatch with each color defined on the 100×100 grid is represented as a circular marker. The color of the marker is a representation of the color (subject to the constraints already discussed) and the radius of the marker is linearly proportional to the magnitude of the reconstruction error.



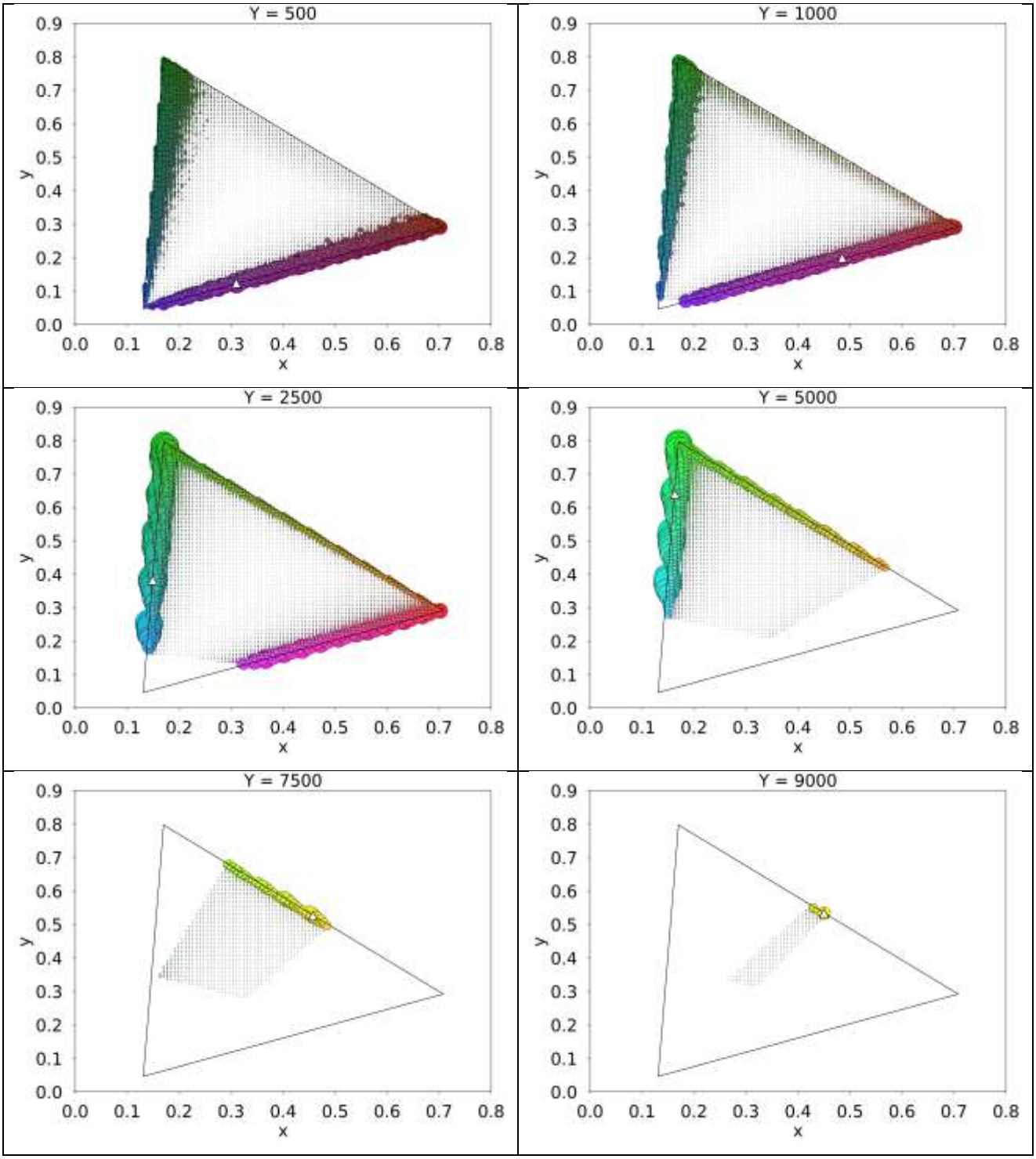


Figure 5 – Plots of the reconstruction error planes for values of Y between 20 and 9000. The point of maximum reconstruction error is marked with a white triangular marker in each case

We see that for all Y values, the errors are largest at the edges of the color gamut. For example, for $Y = 100$, the largest errors occur at the lower edge of the color space in what might be referred to as the saturated blues, purples and reds.

Furthermore, we see that for different values of Y, the area of the color space where larger reconstruction errors occur changes. Specifically, at a value of $Y = 200$, we start to see increasing reconstruction errors in the blue-green axis. At a value of $Y = 1000$ we see that errors begin to increase in the yellows. It should be noted that the total addressable area of the xy plane decreases with increasing Y since the range of "legal" linear light RGB values decreases.

Subjective Assessment

The reconstruction error above represents one simple metric of the error introduced by chroma subsampling in a Rec. 2100, PQ container. No account has been made for how the noise level might be assessed subjectively. For example, an absolute error in the luminance will be easier to see when the luminance of the image is lower than when it is very bright. With this in mind, we viewed swatches for the xy values which resulted in the maximum and minimum objective reconstruction error for several values of luminance, Y, on the SIM2 HDR47 monitor.

We observed that for very low luminance values (for example $Y = 10$) the noise on the 4:4:4 version was very significant, so large in fact that it masked any errors introduced by the subsampling process in the 4:2:0 version. For $Y = 100$ however, the swatch resulting in the maximum reconstruction error had noise that was barely perceptible on the 4:4:4 version but very clear on the 4:2:0 version. The swatch with a minimum error, occurring at $xy = (0.33, 0.34)$, showed barely perceptible noise on both the 4:4:4 and the 4:2:0 versions.

For a luminance of $Y = 2500$, the swatch having the greatest reconstruction error can be described as a light cyan color. The noise is not visible in 4:4:4 but can still be clearly seen in 4:2:0. Increasing the luminance further to $Y = 7500$, the color of the swatch with largest error is yellow. The noise is now not visible on either the 4:4:4 or the 4:2:0 versions, despite the reconstruction error being 9.7×10^5 which is twice that of the error for $Y = 100$.

In conclusion, the subjective assessment of the swatches revealed that for a fixed noise level, the perception of noise is higher when the swatch luminance is lower as one might expect. Furthermore, the perception of noise increases with increasing reconstruction error for a fixed luminance level. The precise relationship, however has not been investigated.

ARTIFACT AVOIDANCE

The artifacts analyzed in the preceding can be avoided in two main ways. The first involves ensuring that the content does not have pixel values at the edge of the color gamut. This will be the case where, for example, the camera used to capture the content cannot address the whole color volume. This restricts the benefit of the increased color volume and is therefore sub-optimal. The second involves modifying the process of converting RGB to $Y'CbCr$. One could use a constant luminance workflow. This decouples the luminance from the chroma components and in so doing eradicates the artifact. It is accepted, however, that this would require a complete overhaul of the broadcast chain.

Finally, we can modify the subsampling process. Two classes of technique to avoid luminance artifacts for non-constant luminance $Y'CbCr$ processing are presented in the following.

Luma Adjustment

The first class of techniques is collectively called luma adjustment, and here the idea is to compensate for the error in the luminance by changing the luma value Y' in each pixel. Since its introduction by Ström et al. (6), it has been the subject of several implementation optimizations (cf. Norkin (14), Ström et al. (11)(13), Rosewarne and Kolesnikov (12)). It has also been included as the enhanced processing mode for the technical report on conversion and coding practices for HDR/WCG (15). An example of artifact reduction using luma adjustment can be seen in Figure 6.

Chroma Adjustment

The second type of technique for artifact avoidance is known as chroma adjustment, and was introduced by Ström and Wennersten (16). Instead of changing only the Y' of the $Y'CbCr$ representation, the original color of each pixel is changed slightly in all three components so that it is more similar to the color of its neighbors. The change is small enough so that the difference should not be visible, yet it has the effect of almost completely removing the subsampling artifacts. After that, the luminance is reintroduced to avoid any luminance smoothing, and finally a step of luma adjustment is used.

In performing chroma adjustment on the artificially generated swatches as part of the subsampling process, it was observed that the reconstruction error is several orders of magnitude lower than that observed using standard subsampling techniques.

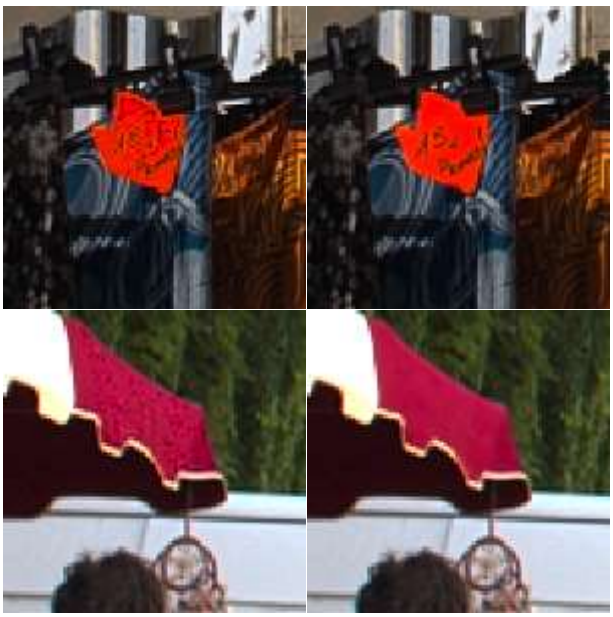


Figure 6 –Left column: after subsampling to 4:2:0 using traditional processing. Right column: after subsampling to 4:2:0 using luma adjustment. Image sequence courtesy of Technicolor and the NevEx project

CONCLUSIONS

This paper has presented an analysis of the error induced by chroma subsampling in an HDR, Rec. 2100, non-constant luminance workflow. By creating synthetic images of plain color contaminated by noise we have been able to objectively assess the magnitude of these errors as a function of both color and luminance. The artifacts have been demonstrated to be worse for colors of high purity, i.e., occupying space at the edge of the color space as previously observed. Additionally, this contribution has demonstrated that the colors for which the error is worse change as a function of the luminance.

Visible artifacts are unlikely to occur when the captured video data does not fill the Rec. 2100 color space. As more WCG content becomes available, however, the industry is likely to see more artifacts associated with chroma subsampling and will therefore need to take steps to avoid it. These fall in to two main areas:

- Prior to conversion to $Y'CbCr$: e.g. avoiding regions of the color space known to cause the issue
- Conversion to $Y'CbCr$:
    - Using a constant luminance workflow – it is acknowledged by the authors that this may not be practical where existing broadcast equipment is to be used
    - Using chroma / luma adjustment algorithms in the subsampling process

The method used in this paper could be significantly enhanced by deriving a noise model more commonly observed in camera data, and using a modified error metric which reflects the subjective assessment of the noise artifacts.

REFERENCES

1. DVB, 2016, Specification for the use of Video and Audio Coding in Broadcasting Applications based on the MPEG-2 Transport Stream, DVB Document A157.

2. ATSC, 2017, ATSC Proposed Standard: Video – HEVC (A/341).

3. ITU-R, 2016, Image parameter values for high dynamic range television for use in production and international programme exchange, Recommendation ITU-R BT.2100-0.

4. Poynton, C., 2003, Digital Video and HDTV Algorithms and Interfaces, Morgan Kaufmann.

5. SMPTE, 2014, High Dynamic Range Electro-Optical Transfer Function, ST 2084.

6. Ström, J., Samuelsson, J., Dovstam, K., 2016, Luma Adjustment for High Dynamic Range Video, Proceedings of the IEEE Data Compression Conference (DCC), Snowbird.

7. Lopez, P., François, E., Yin, P., Lu, T., Luthra, A., Minoo, K., Lee, S., Chen, J., 2014, Not public: Generation of anchors for the explorations for HDR /WCG Content Distribution and Storage, m34077, 109th MPEG meeting in Sapporo, Japan.

8. Francois, E., 2014, Not public: MPEG HDR AhG: about using a BT.2020 container for BT.709 content, 110th MPEG meeting in Strasbourg, France.

9 Ström J., 2015, Not public: Investigation of HDR Color Subsampling, m35841, 111th MPEG meeting in Geneva, Switzerland.

10. Mansencal, T., Mauderer, M., Parsons, M., Canavan, L., Cooper, S., Shaw, N., Wheatley, K., 2017, Colour 0.3.9, Zenodo.

11. Ström, J., 2016, AhG-13 related: Multi-LUT Luma Adjustment Implementation, JCT-VC, JCTVC-Y0030, Chengdu.

12. Rosewarne, C., and Kolesnikov, V., 2016, AHG13: Further results for luma sample adjustment, JCT-VC, JCTVC-Y0034, Chengdu.

13. Ström, J., Andersson, K., Pettersson, M., Hermansson, P., Samuelsson, J., Segall, A., Zhao, J., Kim, S-H., Misra, K., Tourapis, A. M., Su, Y., Singer, D., 2016, High Quality HDR Video Compression using HEVC Main 10 Profile, Proceedings of the IEEE Picture Coding Symposium (PCS), Nuremberg.

14. Norkin, A., 2016, Fast algorithm for HDR video pre-processing, Proceedings of the IEEE Picture Coding Symposium (PCS), Nuremberg.

15. ISO/IEC TR 23008-14, Conversion and Coding Practices for HDR/WCG Y′CbCr 4:2:0 Video with PQ Transfer Characteristics.

16. Ström, J., Wennersten, P., 2017, Chroma Adjustment for HDR Video, in 2017 IEEE Int. Conf. Image Process.