|  |  |
| --- | --- |
| **Joint Collaborative Team on Video Coding (JCT-VC)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  26th Meeting: Geneva, CH, 12–20 January 2017 | Document: JCTVC-Z0043 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **AHG 7: Content colour volume SEI message - Observations and findings** | | |
| *Status:* | Input Document to JCT-VC | | |
| *Purpose:* | Information | | |
| *Author(s) or Contact(s):* | M. Meyer, A.M. Tourapis, D. Singer, Y. Su  1 Infinite Loop Cupertino, CA 95014 USA | Tel: Email: | +1-408-228-7983 atourapis@apple.com |
| *Source:* | Apple Inc | | |

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

# Abstract

This contribution presents initial observations and findings when using the content colour volume SEI message, that was adopted at the previous meeting, as well as some of its possible extensions.

# Introduction

A new content colour volume SEI message [3] was adopted in the HEVC specification at the previous meeting since it was understood that such an SEI message could potentially be beneficial for the processing and display of video content. It was suggested that this might especially be true in the context of HDR applications. This SEI message essentially specifies a nominal description of the expected colour volume that the content actually covers, which may be quite different, and substantially smaller than the colour volume container used for the content representation. As an example, the content may be currently represented using a BT.2100 colour volume container, however, the actual content might only contain colours within the BT.709 colour gamut. Peak brightness may also be limited. Knowledge of such information could potentially be beneficial for a variety of applications, including display mapping, filtering, interpolation, among others.

Although there was a desire [2] to initially have a much more very flexible and comprehensive description of the signal in the originally proposed SEI message, a simplified form was adopted in [3] until the advantages of this SEI as well as its possible extensions are better understood.

In this contribution, we provide information on some of the work that our team performed in the context of this work. It should be noted that given the limited time that we had from the previous meeting until now, as well as certain problems that we discovered with some of the content that we originally intended to use for our experiments, our experimentation and findings are not as extensive as we would have liked them to be. Our work is further described in the next sections.

# Colour Volume Processing and Analysis Framework

Given the intent of these experiments, and to have the ability to more flexibly analyze and process video content we implemented the entire processing chain that is currently performed in HDRTools [1], except the luma adjustment method, also in Matlab[[1]](#footnote-1). Essentially, we can load any image/video format also supported in HDRTools in Matlab, and then by applying appropriate color space conversion processes, different transfer functions, as well as downsampling and upsampling mechanisms, appropriately convert this image to a desired format. As an example, we can convert an SDR video signal that uses the BT.709 RGB or YCbCr limited range representation, to a 4:2:0 Non-Constant Luminance YCbCr BT.2100 [5] representation using the PQ transfer function and the BT.2100 colour primaries. Converting this new signal back to its original representation is also possible. Such conversions would essentially result in a video signal that contains information in a considerably narrower “spectrum” than what the actual signal container can represent. Figure 1 and Figure 2 show the forward and inverse conversion steps supported by this Matlab software implementation.

Figure 1. Conversion of an SDR signal to the BT.2100 representation in our Matlab Framework



Figure 2. Conversion of a BT.2100 representation signal back to SDR



We have selected to use Matlab instead of HDRTools for these conversion steps since that allowed us to more easily analyze and visualize the video data at every step of the conversion process, something that is not as easy to perform currently with HDRTools.

# Content Colour Volume SEI Specific Processing

The currently adopted SEI in [3] is limited by the fact that it only permits signaling of the content colour volume information using RGB primaries, as well as by providing only limits for the minimum and maximum luminance. However, in [2], the color volume could be alternatively described using multiple other colour volumes, using regions, as well as using minimum and maximum values in other colour representations such as RGB, YCbCr, XYZ, HSV, and LMS among others. To understand the effect of these additional descriptions of the content colour volume, we introduced into our processing additional modules that allow us to perform analysis and processing of the data, such as clipping, in these spaces as well. In particular, we have considered the use of adaptive clipping of the video content based on the known content colour gamut by cascading multiple clipping stages in these different descriptions. Given the limited time between the previous meeting and the current, unfortunately we could not test other types of processing. There were several other things that unfortunately we could not test, including:

* Clipping using the content colour gamut information is performed only at the last conversion stage and not during any intermediate processing/conversion stages. Such method could potentially help with reducing error propagation.
* Pixel-wise clipping that accounts for the YCbCr color space limits as per JCTVC-U0042, was not considered.
* When content colour gamut information was used with partitioning partitions were processed in isolation. Ideally, pixels near edges or corresponding to objects spanning multiple partitions should be handled jointly to potentially avoid the introduction of partitioning artifacts.
* Clipping in different domains was done serially, using a fixed order, and without any adaptation process. Intuitively, clipping may benefit if the order was adaptive, e.g. based on the number of values required to be clipped for each colour representation, could be performed pixel-wise instead of framewise, account for clipping errors and so on.
* The use of the HSV colour space.

# Simulations

We have used several SDR MPEG content to evaluate the adaptive clipping mechanism. The content were basically converted to the YCbCr BT.2100 representation, format converted to 4:2:0, and then converted back again to a BT.709 representation. It should be highlighted here that unfortunately considerably amount of time was spent initially on a different set of sequences that seemed to have misclassified as being in limited range, but apparently were in full range, and which misdirected completely our analysis and processing work. Given the limited time remaining, the new sequences were selected without much consideration of their characteristics and might not be the most appropriate content for this conversion. Other material may exhibit better properties than what we have observed using this content. All of this material were in the RGB domain initially.

Table 1. Experimental Content

|  |  |
| --- | --- |
| **Sequence** | **Bit-depth** |
| DucksAndLegs | 10 |
| Kimono | 10 |
| BigBuck | 8 |
| EBULupoCandlelight | 10 |
| EBURainFruits | 10 |
| VenueVue | 8 |
| KimonoError2 | 8 |
| Sc\_map | 8 |
| Sc\_SlideShow | 8 |
| MissionControlClip3 | 8 |
| Sc\_cg\_twist\_tunnel | 8 |
| Sc\_robot | 8 |

Table 1 presents the impact of performing the different clipping processes on these sequences for all frames in the sequence. The first column presents the results without clipping, the second the results with only RGB limits considered, the third uses different RGB limits using 4x4 (16 regions) partitioning, the 4th column accounts for the performance impact using the concatenated approach and a single region, while the last column presents the results for the concatenated method and 16 regions.

It can be observed that although small (0.0252dB) improvement can be seen for the BigBuck sequence using the concatenated approach, the benefits do appear to be consistent and improving using the different methods.

Table 2. Average PSNR performance using different clipping mechanisms

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **unclipped** | **RGB Frm** | **RGB Rgn** | **Conc. Frm** | **Conc. Rgn** |
| DucksAndLegs | 33.4325 | 33.4337 | 33.4340 |  |  |
| Kimono | 41.3561 | 41.3566 | 41.3573 |  |  |
| BigBuck | 41.6158 | 41.6158 | 41.6264 | 41.6269 | 41.6410 |
| EBULupoCandlelight | 49.7765 | 49.7767 | 49.7777 |  |  |
| EBURainFruits | 46.6554 | 46.6554 | 46.6566 |  |  |
| VenueVue | 46.1202 | 46.1205 | 46.1241 |  |  |
| KimonoError2 | 22.8438 | 22.8438 | 22.8438 |  |  |
| Sc\_map | 36.6525 | 36.6525 | 36.6583 | 36.6534 | 36.6664 |
| Sc\_SlideShow | 45.6732 | 45.6732 | 45.6738 |  |  |
| MissionControlClip3 | 29.6490 | 29.6490 | 29.6490 |  |  |
| Sc\_cg\_twist\_tunnel | 36.0406 | 36.0406 | 36.0466 |  |  |
| Sc\_robot | 38.6684 | 38.6843 | 38.6904 | 38.6843 | 38.6907 |

More significant benefits are apparent when examining per frame benefits (Table 3). A gain of 0.089dB can be observed for the BigBuck sequence using the regionwise method and multiple color representations. Since our current scheme is quite crude, we do expect that improvements in performance could still be achieved. Using also more appropriate content, may yield better results as well.

Table 3. Best Frame PSNR improvements compared to the unclipped process

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **unclipped** | **RGB Frm** | **RGB Rgn** | **Conc. Frm** | **Conc. Rgn** |
| DucksAndLegs | 0.0000 | 0.0026 | 0.0027 |  |  |
| Kimono | 0.0000 | 0.0029 | 0.0045 |  |  |
| BigBuck | 0.0000 | 0.0000 | 0.0630 | 0.0133 | 0.0893 |
| EBULupoCandlelight | 0.0000 | 0.0003 | 0.0021 |  |  |
| EBURainFruits | 0.0000 | 0.0001 | 0.0022 |  |  |
| VenueVue | 0.0000 | 0.0098 | 0.0168 |  |  |
| KimonoError2 | 0.0000 | 0.0000 | 0.0000 |  |  |
| Sc\_map | 0.0000 | 0.0000 | 0.0089 | 0.0015 | 0.0249 |
| Sc\_SlideShow | 0.0000 | 0.0000 | 0.0006 |  |  |
| MissionControlClip3 | 0.0000 | 0.0000 | 0.0001 |  |  |
| Sc\_cg\_twist\_tunnel | 0.0000 | 0.0000 | 0.0059 |  |  |
| Sc\_robot | 0.0000 | 0.0192 | 0.0260 | 0.0193 | 0.0265 |

# Conclusion

In this contribution, we have presented some preliminary results on how the Content Colour Volume SEI could be used and how improvements in performance could be achieved by allowing for an enhanced representation to be signaled using this SEI. Further refinements of the processing methods used should likely result in further performance improvements. It is expected that such improved methods will be presented at the next meeting in Hobart.

# References

1. HDRTools Master branch: <https://gitlab.com/standards/HDRTools/>
2. A. M. Tourapis, Y. Su, D. Singer, H. M. Oh, J. W. Choi, J.-Y. Suh, A. K. Ramasubramonian, D. B. Sansli, J. Sole, D. Rusanovskyy, M. Karczewicz, “Content colour volume SEI message,”, JCTVC contribution, JCTVC-Y0040, October 2016, Chengdu, China
3. A. Tourapis, E. François, H. M. Oh, A. K. Ramasubramonian, P. Yin, “Content Colour Volume SEI Message Draft 1”, JCTVC contribution, JCTVC-Y1005, October 2016, Chengdu, China
4. A.M. Tourapis, M. Meyer, and D. Singer, “A new tool for Color Gamut Analysis of MPEG video content,” JCTVC contribution, JCTVC-Z0040, January 2017, Geneva, Switzerland
5. Recommendation ITU-R BT.2100 (2016), *Image parameter values for high dynamic range television for use in production and international programme exchange*

# Patent rights declaration(s)

**Apple Inc may have current or pending patent rights relating to the technology described in this contribution and, conditioned on reciprocity, is prepared to grant licenses under reasonable and non-discriminatory terms as necessary for implementation of the resulting ITU-T Recommendation | ISO/IEC International Standard (per box 2 of the ITU-T/ITU-R/ISO/IEC patent statement and licensing declaration form).**

1. This code could be made available to anyone that may be interested in it. [↑](#footnote-ref-1)