|  |  |
| --- | --- |
| **Joint Collaborative Team on Video Coding (JCT-VC)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11** | Document: JCTVC-AA0037 |

27th Meeting: Hobart, AU, 31 March – 7 April 2017

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **User Data Unregistered SEI for Backward Compatible HDR Coding** | | |
| *Status:* | Input Document to JCT-VC | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | Pankaj Topiwala Wei Dai Madhu Krishnan  3097 Cortona Dr., Melbourne, FL 32940 | Tel: Email: | 321-355-7376 {pankajtva @ gmail.com} |
| *Source:* | FastVDO LLC | | |

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

# Abstract

This proposal describes the usage of User Data Unregistered (UDU) SEI for transmitting metadata related to backward compatible solutions to coding HDR video. The bandwidth for UDU SEI is three times lower than the bandwidth consumed by the Tone Mapping Information (TMI) SEI.

# Introduction

As part of the investigations done to leverage existing video standards for efficiently encoding HDR/WCG videos, FastVDO developed several effective approaches for HDR video coding [1]. All our methods can be implemented entirely as pre/post processing steps, and do not necessitate any normative changes to the core Main10 encode/decode processes. Moreover, the output of a Main10 decoder is an optimized SDR rendering. In input document JCTVC-Y0042 [2] a variety of methods were described for signalling the associated metadata: a) TMI SEI message and b) Color Remapping Information (CRI) SEI message for content mapping. These were developed in both linear light and nonlinear light. Both types of methods were cross-checked by Technicolor. The method using nonlinear light has been reported in a JCTVC Technical Report [8]. In this proposal the additional method using linear light, signaled by the user data unregistered (UDU) SEI [4] is described which performs well, and could reduce the payload size of the metadata compared to TMI.

# System Overview

There are a variety of related, but somewhat different methods for achieving our goals as described in JCTVC-Y0042 [2]. In this document, the pre/post processing step applied to linear light in 4:4:4 domain, along with usage of UDU SEI [4] for signalling the content mapping is explained.

The HDR-to-SDR conversion approach is shown in Figure 1. In this method, the dynamic range adaptation (DRA) is performed on linear light input signal. Figure 2 describes some details of the DRA scheme as pre and post processing.

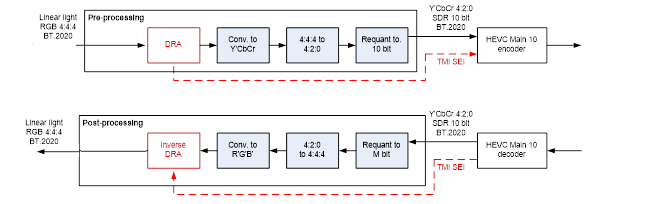
****

Figure 1. Generic signal processing scheme with DRA, as pre- and post processing.

Grade HDR to SDR in 4:4:4

RGBHDR RGBSDR

Generate Data Dependent HDR to SDR Grading Function

Metadata

Figure 2. DRA scheme, which regrades HDR to SDR.

*Pre-processing:* This step involves the computation of a data dependent HDR-to-SDR grading function/process. Since the input linear light data is presented in floating point, one example DRA works in floating point as well. A base transfer function, which can be either static or adaptive, is modified depending on the statistical analysis of the linear light HDR RGB input signal. Example base transfer functions that can be used include ST-2084 (Dolby PQ) [5], Philips TF [3], and HLG (ARIB STD-B67) [6]. The computed grading function/process is represented as a 1D LUT. The generated metadata consists of the color volume information and the 1D LUT represented using a piecewise linear function with 10 pivot points. The grading function/process is then used to grade the HDR RGB input signal to an SDR RGB signal. The SDR RGB signal is then color transformed to Y’CbCr, downsampled from 4:4:4 to 4:2:0 and encoded using HEVC reference codec (Main10 profile). The metadata is signaled using mastering display color volume (MDCV) SEI [4] and the UDU SEI.

***1) Data dependent HDR-to-SDR grading function/process:*** Figure 3 describes the technique involved in generating the grading function/process.

Piecewise

Linear Tone

Mapping

Generate

Grading

Factor

Apply adaptive OETF

Apply adaptive EOTF

Coding TF

Generate

Luminance

RGBHDR

Figure 3. Data dependent grading

***2) Generate Luminance:*** A luminance component (Y) is generated using the common conversion to the Y’CbCr color space. This part is linear. All other components are nonlinear, and they can be all combined into a single function for purposes of implementation.

***3) Apply data adaptive OETF:*** The linear luminance (Y) input signal is converted to a perceptually uniform signal, but using a FastVDO-modified version of ST-2084 [5], herein “FVPQ”. The FVPQ function is described as:

Here, and is new parameter, indicative of the peak brightness of the data. When the parameter equal 10,000, this reduces to the usual PQ function.

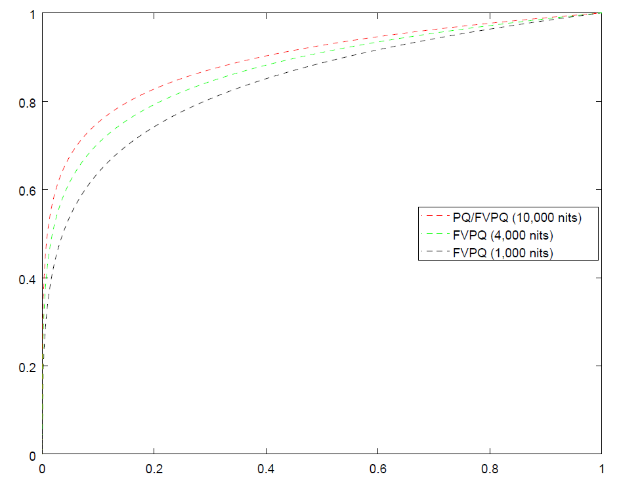
 This represents a first level of adaptivity, and is an important step in creating SDR intermediates that better capture the artistic intent of the HDR content. This is signalled using the MDCV SEI.

Figure 4. FVPQ OETF, for values of *InputPeakBrightness =* 10000, 4000, and 1000 nits (at 10K nits it coincides with the ST-2084 PQ function), as pictured from top to bottom. Note that when the input data is below 10000 nits, our TF affords more codewords at the brighter end than the standard PQ transfer function.

***4) Piecewise Linear Tone Mapping:*** A great variety of tone mapping schemes can be applied for HDR coding, and implemented in a variety of ways as well. In this paper, for convenience, we will only mention a piecewise linear mapping scheme. To add a modest amount of flexibility, we consider an L-piece linear model, where L >=1. For each piece the model can be represented as:

where, *k* represents the *kth* piece or bin. These equations are applied to the to generate a new signal .Let *MAX\_INPUT\_LUMA* be the maximum luminance of the input signal and *MAX\_DISPLAY\_LUMA* be the maximum luminance of the target SDR display. The coefficients *a(k), b(k)* are sequence dependent and are derived based on the distribution characteristics of the HDR input signal. As a specific example, they can be based on a histogram analysis, as follows. The input luma signal is segmented to *L* bins, and a count is kept for the number of pixels in each bin. Under an optimization process, the number of bins and their respective sizes are chosen such that the lower and mid tone levels have an improved representation compared to the unadjusted transfer function representation. Thus the parameters *a, b* are chosen for each piece such that:

where *MAX\_LUMA\_BINk* is the maximum luma value in bin *k*, and *binsizek* is the number of pixels in bin *k*.

|  |  |
| --- | --- |
|  |  |

Figure 5. Example histograms of first frame of (a) BalloonFestival, and (b) Market sequence, as example statistical basis for tone maps.

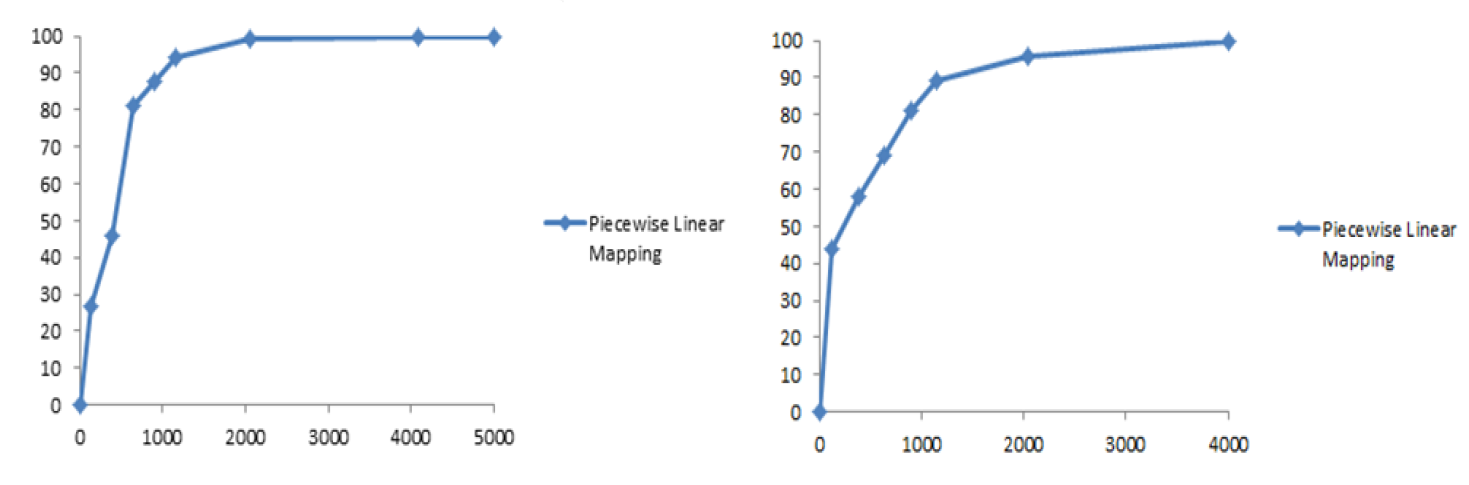


Figure 6. Example piecewise linear mappings corresponding to the statistics (histogram here) for (a) BalloonFestival, and (b) Market.

***5) Apply data adaptive EOTF:*** The perceptual domain signal is converted back to linear light (Y’) using the matching inverse FastVDO-modified ST-2084 OETF (FVInvPQ). The function is described as:

Here *dMax(a,b)* returns *a* if *a > b,* else *b*.

***6) Generate Grading Factor:*** The grading factor (actually, a function of x, y, and t) is calculated as:

This is then used to generate an intermediate signal (RGBSDR)’ = RGBHDR \* λ.

***7) Coding TF:*** Finally, a backward compatible SDR signal, RGBSDR, is generated by further applying a function like the ITU-R Rec. BT.2020 power-law gamma curve [7] or the ST-2084 (PQ TF).

The entire processing chain in Figure 2 can be visualized as a single tone mapping scheme.

*Post-processing:* The post-processing step in the proposed HDR coding chain is presented in Figure 7.

Regrade SDR to HDR in 4:4:4

RGBSDR RGBHDR

Metadata

Generate Data Dependent SDR to HDR Regrading Function

Figure 7. Post-Processing, which recovers HDR from SDR.

The HEVC reference codec is used to decode the bitstream and generate the reconstructed Y’CbCr 4:2:0 signal and metadata. The Y’CbCr 4:2:0 signal is then upsampled from 4:2:0 to 4:4:4 and color transformed to SDR RGB. The regrading function/process is computed using the metadata. The process is very similar to that described in Figure 2, except that the blocks are in reverse order. Finally, regrading is performed on the SDR RGB input to obtain the linear light reconstructed HDR RGB signal.

**Video Usability Information (VUI)**

The following VUI settings are used for rendering the SDR content:

* colour\_description\_present\_flag 1
* colour\_primaries 9 (BT.2020)
* transfer\_characteristics 1, 6, 14 or 15 (SDR transfer function)
* matrix\_coeffs 9 (BT.2020 NCL)
* video\_full\_range\_flag 0 (narrow range)

**User Data Unregistered (UDU) SEI**

The UDU SEI syntax as described in H.265/HEVC Specification [4] contains unregistered user data identified by a universally unique identifier (UUID). The message consists of two fields as defined below:

* uuid\_iso\_iec\_11578 is a unique 16 Byte field that shall be set to the value:

Bytes 1-8: 0x46, 0x61, 0x73, 0x74, 0x56, 0x44, 0x4f, 0x55

Bytes 9-16: 0x44, 0x55, 0x53, 0x65, 0x69, 0x4d, 0x65, 0x73

This represents the ASCII string: “FastVDOUDUSeiMes”

* user\_data\_payload\_byte is a variable length field depending on the signalled data. In the current application, 20 bytes are used to represent the content mapping information.

**Mastering display color volume (MDCV) SEI**

The MDCV SEI syntax describes the color volume of the mastering display used for viewing while authoring the video content.

An example of the settings used in the SEI:

SEIMasteringDisplayColourVolume: 1

SEIMasteringDisplayMinLuminance: 47

SEIMasteringDisplayMaxLuminance: 40000000

SEIMasteringDisplayPrimaries: 13250,34500,7500,3000,34000,16000

SEIMasteringDisplayWhitePoint: 15635,16450

**Results**

The objective results obtained by the proposed systems are very similar to that presented in JCTVC-Y0042 [2]. The only difference is in the size of the metadata. Table 1 compares the total bandwidth of UDU and TMI SEI for the different sequences.

|  |  |  |
| --- | --- | --- |
| **Sequences** | **TMI SEI Bandwidth (bytes)** | **UDU SEI Bandwidth (bytes)** |
| BalloonFestival | 134 | 36 |
| BikeSparklers cut 1 | 134 | 36 |
| BikeSparklers cut 2 | 134 | 36 |
| FireEaterClip4000r1 | 134 | 36 |
| GarageExit | 134 | 36 |
| EBU\_04\_Hurdles | 134 | 36 |
| StEM\_MagicHour cut 1 | 134 | 36 |
| StEM\_MagicHour cut 2 | 134 | 36 |
| StEM\_MagicHour cut 3 | 134 | 36 |
| Market3Clip4000r2 | 134 | 36 |
| ShowGirl2Teaser | 268 | 72 |
| EBU\_06\_Start | 134 | 36 |
| SunRise | 134 | 36 |
| StEM\_WarmNight cut 1 | 134 | 36 |
| StEM\_WarmNight cut 2 | 134 | 36 |

Table 1. Bandwidth Comparison

**Conclusion**

A technique is presented for the transmission of metadata effectively using the user data unregistered (UDU) SEI. The metadata size is more than three times lower than the case using TMI SEI.

**References**

[1] Pankaj Topiwala, Wei Dai and Madhu Krishnan, “HDR CE5: Report of Experiment 5.3.2”, JCTVC-W0055, San Diego, Feb., 2016.

[2] Pankaj Topiwala, Wei Dai and Madhu Krishnan, “Tone Mapping Information and Related SEIs for HDR Coding”, JCTVC-Y0042, Chengdu, Oct., 2016.

[3] R. Goris, R. Brondijk, and R. van der Vleuten, “Philips Response to CFE on HDR and WCG,” ISO/IEC MPEG doc. m36266, Warsaw, PL, July, 2015.

[4] ITU-T Rec. H.265 (2016/12), High Efficiency Video Coding.

[5] Society of Motion Picture and Television Engineers ST 2084 (2014), High Luminance EOTF.

[6] Recommendation ITU-R BT.2100 (2016), “Image parameter values for high dynamic range television for use in production and international programme exchange”.

[7] ITU-R BT.2020 (2012), “Parameter Values for ultra-high definition television systems for production and internal programme exchange.”

[8] E. Francois et al, “Signalling, Backward Compatibility and Display Adaptation for HDR/WCG Video Coding (Draft 2),” JCTVC-Z1012, Geneva, CH, 12-20 January, 2017.

# Patent rights declaration(s)

**FastVDO LLC 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).**