|  |  |
| --- | --- |
| **Joint Collaborative Team on Video Coding (JCT-VC)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  23rd Meeting: San Diego, USA, 19–26 February 2016 | Document: JCTVC-W0089 / m37729 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **HDR CE2-related: some experiments on ETM with dual grading input** | | |
| *Status:* | Input Document to JCT-VC | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | E. François, Y. Olivier, C. Chevance | Tel: Email: | [Edouard.francois@technicolor.com](mailto:Edouard.francois@technicolor.com) [Yannick.olivier@technicolor.com](mailto:Yannick.olivier@technicolor.com) |
| *Source:* | Technicolor | | |

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

# Abstract

This document reports preliminary experiments on ETM with dual SDR/HDR grading. In this configuration, a graded SDR version as well as the usual graded HDR version are given as inputs to the reshaper. The reshaping process uses the input SDR as reference and generates reshaping data to match as much as possible the reshaped input HDR to the input SDR. In the experiments, reshape setting 1 configuration is used (cross-plane chroma reshaping). It is reported that these preliminary experiments tend to show that ETM models are able to properly reproduce an SDR intent.

# Introduction

This documents reports preliminary experiments using ETM for dual-grading HDR distribution scenario. We consider in this scenario that both SDR and HDR versions are given as input of the HDR distribution system. The goal of this initial work is to adapt the ETM algorithm in order to generate a reshaped version of the HDR content that matches as much as possible the input SDR.

The document reports the achieved preliminary work based on the reshape setting 1 configuration (PWP 8 pieces for luma, cross-channel PWL 8 or 32 pieces for chroma). In next section, we shortly remind the ETM reshape setting 1 configuration process. Next section describes the adaptation made to ETM to work with an input SDR. Next sections reports the preliminary experiments and results.

# ETM\_RC\_r1 process overview for reshapeSetting 1 configuration

Reshaper setting 1 configuration applies cross-plane mode for chroma reshaping, using two piece-wise-linear (PWL) modelled look-up-tables (LUTs) to control the chroma components. The luma component is reshaped using an intra-plane piece-wise-polynomial (PWP) modelled LUT. In the current ETM settings, 8 pieces are used both for luma PWP and for chroma PWL models, [1].

The reshaping process [2] works in four main steps, as shown in Figure 1:

1. Conversion of input HDR signal into Y’CbCr 10b 4:2:0 with inverse ST.2084 (PQ) EOTF;
2. Reshaping of 10b luma PQ to get an SDR-like 10b luma;
3. Automatic generation of an intermediate SDR 4:2:0 Y’CbCr picture;
4. Derivation of cross-plane chroma PWL models parameters to fit as much as possible the chroma components of the intermediate 4:2:0 Y’CbCr SDR picture.

In the following sub-sections, short description of these processes is provided. More details on the reshaping algorithms in ETM\_RC\_r1 are provided in document JCTVC-W0031.



Figure 1: simplified block diagram of ETM\_RC\_r1 process in reshapeSetting 1 configuration.

## Y reshaping

Y reshaping is achieved using the following function

F(L) = log( 1. + (L / Ba)g ) / log( 1. + (P / Ba)g )

controlled by two adaptive tuning parameters Ba, g, and by the peak luminance P. In ETM\_RC\_r1, g is set to 2.8, and Ba varies depending on the content. The resulting function is modeled by a PWP 8 pieces model. Y is then scaled to and quantized in Full or Standard Range signal.

## Intermediate SDR generation

To generate the intermediate SDR (YintermUintermVinterm), the following steps apply:

1. Generate HDR luminance Lhdr from input linear-light HDR R,G,B samples
2. Convert Yreshape to linear-light Lreshape as follows

Lreshape = ( (Yreshape – off) / scale ) (1 / 0.45)

where off and scale are the parameters for Full or Standard Range normalization of Yreshape

1. Scale linear-light HDR R,G,B samples by ratio w = Lreshape / Lhdr
2. Convert scaled R,G,B to SDR R’,G’,B’ as follows

R’ = 1023 \* (Rscaled)0.45 G’ = 1023 \* (Gscaled)0.45 B’ = 1023 \* (Bscaled)0.45

1. Convert SDR R’,G’,B’ to 10b Y’CbCr then downsample chroma to get 10b 4:2:0 signal YintermUintermVinterm

## Chroma LUTs fitting

The chroma LUTs fitting computes the two LUTs LUTcb and LUTcr in order to fit as much as possible the chroma components of the intermediate SDR picture, using the cross-channel chroma mode. The LUTs are derived by least mean squares optimization, with minimization of the following functions:

with OC equal to 512 (29) for 10 bits content.

More details on the cross-plane LUT optimization for reshape setting 1 are provided in document JCTVC-W0031 [2].

# ETM adaptations for input SDR

## Change compared to ETM

The main change to ETM for dealing with input SDR is shown in shown in Figure 2. Instead of generating automatically an intermediate SDR, we use the input SDR. The input SDR is used as reference to find a LUTY enabling to match the input YPQ signal to the input YinSdr signal. Then the input SDR is used as reference to find LUTcb and LUTcr enabling to match the input UPQ and VPQ signal to the input UinSdr and VinSrd signal using the cross-channel mode.



Figure 2: simplified block diagram of modified reshaping algorithm.

## Derivation of a LUTY to match the input SDR luma

The problem of luma matching is to find a LUTY so that LUTY[YPQ(i,j)] is close to YinSdr(i,j) for any pixel i,j. In general, and especially when different local grading apply for the SDR and HDR versions, there is not one unique matching between a value of YPQ and its corresponding value of YinSdr. This is illustrated in Figure 3 that depicts for two different content (first picture of BalloonFestival and WarmNight sequences) the minimum, maximum and average YinSdr values for eachYPQ value. It can be observed that, especially for BalloonFestival, the YinSdr values can noticeably vary for a given YPQ value. This means that a simple global LUT matching cannot in general perfectly match the input SDR from the input HDR PQ signal. The goal is however to obtain an SDR version that globally renders as far as possible as the input SDR.

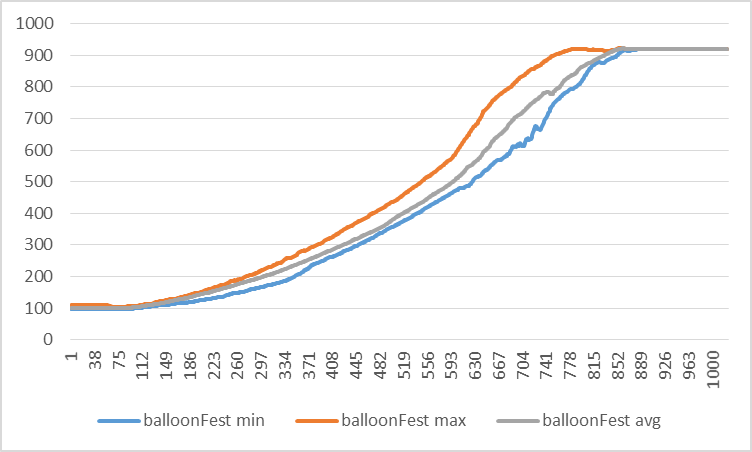
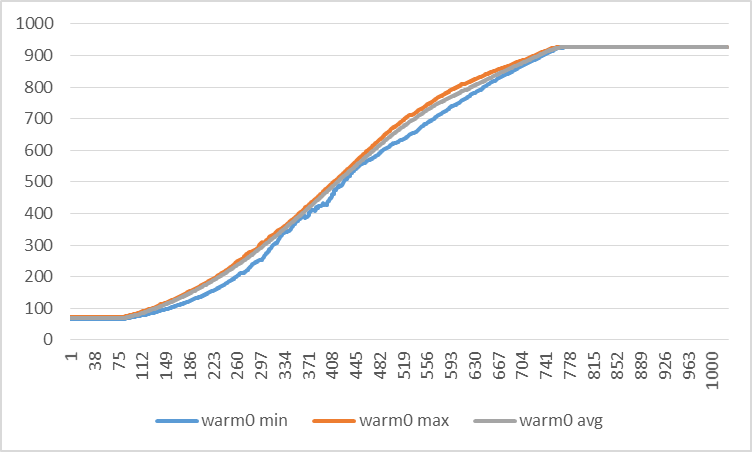
 

Figure 3. Min, max and average LUTs for balloonFestival and WarmNight.

In this preliminary work, two different mapping solutions are tested for luma.

### First luma mapping solution

In the first solution, the mapping is achieved using the following equation:

with a weighting function. In practice, has been used, which actually corresponds to non-weighted Least Mean Squares estimation.

### Second luma mapping solution

In the second solution, a robust estimation is applied. For each YPQ value, a histogram of the YinSdr values in built. The mapped value LUTY[YPQ] is estimated as the value Y such that N% of the YinSdr samples mapping to YPQ are above this value Y. A typical value for N is 50 (corresponding to median estimation). If the goal is to generate a mapped version with more dense colors, lower values of Y are preferable, and the ratio N shall be reduced. In the illustrations below, a value N=20 was used.

Figure 4 shows the LUTs of these two different algorithms for different tested content (first picture considered). Slight differences in the LUT shape can be observed between these two algorithms.

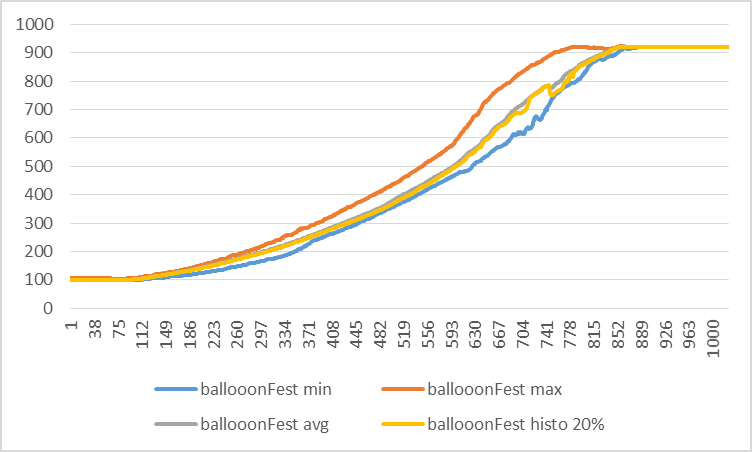
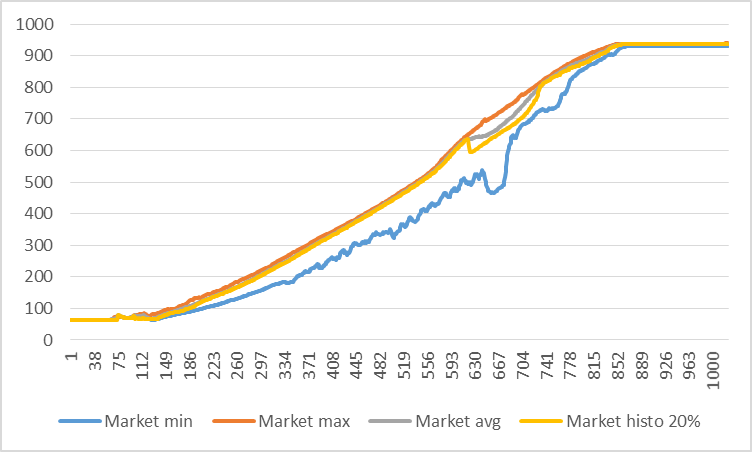
 

Figure . Min, max, average and histogram-based (N=20%) LUTs for balloonFestival and Market.

## Derivation of the chroma LUTs to match the input SDR luma

The same algorithm as used in the default ETM (cf section 2.3) is used to derive the cross-plane chroma LUTs.

# Experiments illustrations and comments

The following content, for which a graded version is available to the MPEG members, have been tested: BalloonFestival, FireEater, Market, StEM MagicHours, StEM WarmNight. The content is natively in BT.709 container. It is converted to BT.2020 container before being processed. For final rendering, the mapped SDR in BT.2020 container is converted to BT.709 SDR.

It can be observed that the mapped versions obtained from the two tested algorithms are in general close. The histogram-based algorithm generates slightly darker pictures. In both cases, the mapped picture is quite similar to the graded SDR picture. Some limited hue changes can be observed, in particular in BalloonFestival, and at a lesser level, in Market. The flame of fireEater is slightly less bright in the mapped version than in the graded SDR. The matching for StEM content is particularly efficient and very few differences can be observed between the graded version and the mapped version. Overall, the EMT can map with a good conformity the input SDR, even for critical cases where non-global grading was used.

## Comparisons of average LUT vs histogram-based matching algorithm

The following pages include snapshots of input graded SDR, mapped SDR using the average luma algorithm, and mapped SDR using the histogram-based (N=20%) algorithm. For chroma LUT modeling, a PWP model with 32 pieces was used. Next sub-section shows some results with 8 pieces.







Figure . BalloonFestival – top: grade SDR, middle: average LUT, bottom: histogram-based LUT.







Figure . Market – top: grade SDR, middle: average LUT, bottom: histogram-based LUT.







Figure . Fire image 0 – top: grade SDR, middle: average LUT, bottom: histogram-based LUT.







Figure . Fire image 51 – top: grade SDR, middle: average LUT, bottom: histogram-based LUT.







Figure . MagicHours cut1 – top: grade SDR, middle: average LUT, bottom: histogram-based LUT.







Figure . MagicHours cut2 – top: grade SDR, middle: average LUT, bottom: histogram-based LUT.







Figure . MagicHours cut3 – top: grade SDR, middle: average LUT, bottom: histogram-based LUT.







Figure . WarmNight cut1 – top: grade SDR, middle: average LUT, bottom: histogram-based LUT.







Figure . WarmNight cut2 – top: grade SDR, middle: average LUT, bottom: histogram-based LUT.

## Using 8, 16 or 32 pieces for chroma LUT modeling

Some illustrations showing the difference of using 8, 16 or 32 pieces for the chroma LUTs modeling for fireEater are shown below (average algorithm is used for luma fitting). For this sequence, using 16 or 32 pieces provides better matched colors and sharpness in the color pictures. There is no big difference between using 16 or 32 pieces. For most sequences, this observation is confirmed. Difference between 16 and 32 pieces is small, but 8 pieces leads to more noticeable changes in the colors.

Figure 14. FireEater 1st picture: top left: graded SDR, top right: PWL32, bottom left: PWL16, bottom right: PWL8.

# References

1. “Proposed initial Test Model for HEVC HDR extension,” R. Brondijk, R. Goris, R. van der Vleuten, L. van de Kerkhof, D. Rusanovskyy, A. Ramasubramonian, D. Bugdayci, S. Lee, J. Sole, M. Karczewicz, F. Galpin, S. Lasserre, F. Le Leannec, T. Poirier, E. François, MPEG document M37285, Geneva, Switzerland, Oct. 2015.
2. “Description of the reshaper parameters derivation process in ETM reference software,” K. Minoo, T. Lu, P. Yin, L. Kerofsky, D. Rusanovskyy, E. François, JCTVC-W0031/MPEG M37536, San Diego, USA, Feb. 2016.

# Patent rights declaration(s)

**Technicolor 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).**