|  |  |
| --- | --- |
| **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-Z0024 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **AHG14: Evaluation of SDR quality from backward compatible HDR video technologies** | | |
| *Status:* | Input Document to JCT-VC | | |
| *Purpose:* | Information | | |
| *Author(s) or Contact(s):* | Edouard François Franck Hiron Christophe Chevance (Technicolor) | Email: | [edouard.francois@technicolor.com](mailto:edouard.francois@technicolor.com) [franck.hiron@technicolor.com](mailto:franck.hiron@technicolor.com) [christophe.chevance@technicolor.com](mailto:christophe.chevance@technicolor.com) |
| *Source:* | JCT-VC | | |

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

# Abstract

This document is a report of evaluation tests performed by Technicolor on the SDR quality from various bitstream SDR backward compatible technologies described in the draft Technical Report on Signalling, Backward Compatibility and Display Adaptation for HDR/WCG Video Coding (JCTVC document JCTVC-Y1012). The tests have consisted in assessing the conformity of the SDR rendering, in terms of color and texture, comparatively to the HDR rendering.

# Introduction

In the draft Technical Report (TR) on Signalling, Backward Compatibility and Display Adaptation for HDR/WCG Video Coding (JCTVC document JCTVC-Y1012 [1]), a number of technologies enabling the bitstream SDR backward compatibility are reviewed. In HDR/WCG distribution systems that support bitstream SDR backward compatibility, the decoded video signal from a standard-compliant decoder can be directly displayed on an SDR-capable display without adaptation. Different technologies are mentioned in the TR. The first solution is the HLG, a fixed transfer function applying to an input scene-referred signal (when the input HDR signal is display-referred, a pre-process depending on the mastering display peak luminance is applied). The other set of solutions is based on the usage of content-dependent dynamic metadata, conveyed using the CRI SEI message, the TMI SEI message, or a user-data registered SEI message in ETSI TS 103 433 [2] (a.k.a. SL-HDR1).

This document reports results of visual evaluation tests made by Technicolor to evaluate the quality of the SDR video generated from the HDR content generated from these different SDR backward compatible solutions. The tests have been done in uncompressed domain and only aimed at checking the SDR quality resulting from the conversion of input HDR content represented in 4:4:4 RGB linear-light to a 4:2:0 Y’CbCr 10-bits signal. The goal is to assess how much conform are the SDR versions generated by the different tested SDR backward compatible solutions, to the HDR version. The assessment is primarily focused on the color and texture rendering, not on the brightness that, obviously, may significantly differ between the HDR and the SDR versions.

# Test content

Different sequences with various peak luminance, from 1000 to 5000 cd/m2, have been used. The content color gamut is either BT.709 or P3D65. All sequences are represented in BT.2020 color primaries container, and are represented in EXR RGB 4:4:4 linear-light half-float format.

The sequences are considered as display-referred, as they do not result directly from the scene capture, but from a grading process on a grading monitor.

The conversion and inverse conversion processes, for all tested solutions, are performed in BT.2020 color primaries container.

The test sequences are listed in Table 1, and illustrative snapshots are depicted in Figure 1 - Figure 5.

Table 1. Test sequences.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Sequence** | **Peak**  **Luminance (cd/m2)** | **Content**  **gamut** | **Container**  **gamut** | **Native resolution** |
| Tch\_Balloon | 1000 | 709 | 2020 | UHD |
| Tch\_Zombie1 | 1000 | P3D65 | 2020 | UHD |
| Tch\_Kayak | 1250 | P3D65 | 2020 | UHD |
| Hurdles | 3000 | 709 | 2020 | HD |
| Starting | 3000 | 709 | 2020 | HD |
| Market | 4000 | 709 | 2020 | HD |
| Netflix Chimera2 | 4000 | P3D65 | 2020 | UHD |
| BalloonFestival | 5000 | 709 | 2020 | HD |

Figure 1. Snapshot of Tch\_Balloon (left) and Tch\_Zombie1 (right) - 1000 cd/m2.



Figure 2. Snapshot of Tch\_Kayak - 1250 cd/m2.

Figure 3. Snapshot of EBU Hurdles and EBU Starting - 3000 cd/m2.

Figure 4. Snapshot of Market and Netflix Chimera2 - 4000 cd/m2.



Figure 5. Snapshot of BalloonFestival - 5000 cd/m2.

# Evaluated technologies

## HLG

The first considered technology is the HLG, as described in section 7.1 of JCTVC-Y1012 [1]. The conversion and inverse conversion processes are depicted in Figure 6. As the input HDR content used for the tests is display-referred, the conversion from linear-light to 4:2:0 Y′CbCr 10-bits includes the initial inverse system gamma step, as described in section 7.1.2.2 of JCTVC-Y1012 [1]. The value of the gamma correction, g, is derived from the peak luminance of the content mastering display, as:

g = 1.2 + 0.42 \* Log10( P ÷ 1000 )



Figure 6. Conversion (top) and inverse conversion (bottom) using HLG.

## SL-HDR1

The second considered solution, mentioned in section 7.2 of JCTVC-Y1012 [1], is the technology specified in recommendation ETSI TS 103 433 [2], also known as SL-HDR1. The solution is based on the usage of dynamic metadata, conveyed in a user-data registered SEI message. The metadata should be used in the decoder to perform the reconstruction of the HDR video starting from the SDR video generated at the encoder from the input HDR video. The reconstructed HDR output from the decoder is close to the input HDR of the encoder. The metadata mostly consists of a luma mapping function applicable to the Y component, and of luma-dependent chroma scaling functions. The scheme, illustrated in Figure 7, is close to what was described in contribution JCTVC-U0085 [3].

Automatic grading is used to derive the dynamic metadata. An example of such an algorithm can be found in JCTVC-W0031 [4], section 3.2.



Figure 7. Conversion (top) and inverse conversion (bottom) with SL-HDR1 (ETSI TS 103 433).

## CRI applied in Y′CbCr 4:2:0 domain

The third considered solution is described in section 7.2.1 of JCTVC-Y1012 [1]. It is also based on the usage of dynamic metadata to perform at decoder side the inverse conversion from the decoded SDR signal to an HDR version close to the input HDR. The conversion and inverse conversion processes, shown in Figure 8, are directly applied in the Y′CbCr 4:2:0 10-bits domain. They involve a dynamic range adaptation (DRA) at encoder to convert the HDR PQ BT.2020 video into an SDR BT.2020 version. The inverse DRA at decoder uses metadata conveyed in CRI SEI messages. The three Pre-LUTs of CRI are activated. The three-by-three matric and the three Psot-LUTs are not used.

In the experiments, the three CRI Pre-LUTs are automatically generated using the SL-HDR1 SDRs as reference SDRs. The dynamic range adaptation functions, that map the input HDR signal into an SDR version, aim at mapping the Y′, Cb and Cr HDR componennts as close as possible to the reference Y′, Cb and Cr SDR components. These mapping functions must be invertible. The 3 Pre-LUTs conveyed in the CRI are the inverse of these mapping functions. One CRI SEI per sequence is generated, based on the first HDR PQ and SDR SL-HDR1 picture of the sequence.



Figure 8. Conversion (top) and inverse conversion (bottom) with CRI applied in Y′CbCr 4:2:0 domain.

## TMI applied in R′G′B′ 4:4:4 domain

The fourth considered solution is described in section 7.2.3 of JCTVC-Y1012 [1] and in JCTVC-Y0042 [5]. It is also based on the usage of dynamic metadata to perform at decoder side the inverse conversion from the decoded SDR signal to an HDR version close to the input HDR. The DRA and inverse DRA processes, shown in Figure 9, are applied in the R′G′B′ 4:4:4 domain. An invertible mapping function is used in the pre-processing to map the luminance signal to SDR. A scaling factor is then derived and applied to the RGB components. The TMI is used to embed the inverse mapping function.



Figure 9. Conversion (top) and inverse conversion (bottom) with TMI applied in RGB 4:4:4 domain.

For this case, the results from contribution JCTVC-Y0042 [5] were used. Only the SDR versions from JCTVC test content were available. Therefore only Hurdles, Starting, Market and BalloonFestival have been used for this case.

# Visual tests description

## Test environment and methodology

The goal of the tests is to verify the conformity of colors and texture of the SDR generated by the various tested technologies to the HDR content. For each solution, the viewers have to to check the **conformity of the colors and the preservation of the texture**. It was indicated to the viewers that the brightness should, as much as possible, not be taken into account in the conformity assessment.

The methodology used for the tests is inspired from the methodology described in MPEG document m37407. The main difference is that, instead of using two different displays as in m37407 [6], the test set-up uses one Sim2 for viewing both the HDR and SDR content. This guarantees that the colors are rendered in the same way for all test cases. It was also judged that the Sim2 is relevant to check content with high mastering peak luminance (up to 5000 cd/m2).

For each test case, the test was achieved as follows:

* HDR video (8 to 10 sec)
* Grey picture (0,5 sec)
* SDR video (8 to 10 sec.)
* message “Vote N” (5 sec.)
* Grey picture (0,5 sec)

For each sequence, the different test cases (resulting from the different tested technologies) were presented, according the sequence explained above. The order of presentation of the various test cases has been randomized. One test session involved two viewers during around 15 min. The tests were performed with 26 viewers, among which 1 was considered as outlier (based on detection test derived from Rec. ITU-R BT.500-13 [7]).

The eleven-grade mean-opinion score (MOS) scale ranging from 0 to 10 was used. They may be interpreted as shown in Table 2.

Table 2. Interpretation of MOS score values.

|  |  |
| --- | --- |
| **MOS** | **Explanation** |
| 10 | denotes a quality of reproduction that is perfectly faithful to the original.  No further improvement is possible. |
| 9 | Excellent |
| 8 |  |
| 7 | Good |
| 6 |  |
| 5 | Fair |
| 4 |  |
| 3 | Poor |
| 2 |  |
| 1 | Bad |
| 0 | denotes a quality of reproduction that has no similarity to the original.  A worse quality cannot be imagined. |

## Conversion process to Sim2

The conversion process for Sim2 rendering is based on HDRTools v0.13, ffmpeg, and Sim2Convert v0.8.1.

In case of native UHD HDR content, a preliminary downsampling process is applied to HD. The new EXR HD version version is generated in three steps: conversion from linear-light RGB 4:4:4 EXR to RGB 4:4:4 PQ (typically 16 bits), downsampling of the RGB 4:4:4 PQ content to HD resolution, inverse conversion from RGB 4:4:4 PQ to linear-light RGB 4:4:4 EXR.



Figure 10. Downsampling chain of UHD linear-light RGB 4:4:4 HDR content.

The conversion of HDR linear-light RGB 4:4:4 EXR content to logLuv Sim2 is made using Sim2Convert, v0.8.1. Clipping to BT.709 color gamut is applied for P3D65 content.

The conversion chain of SDR YCbCr 4:2:0 10-bits content is made in two steps. First the SDR YCbCr 4:2:0 10-bits signal is concerted to RGB 4:4:4 EXR SDR, with a peak luminance set to 100 cd/m2. This is indicated using the parameter SourceNormalizationScale of HDRTools (v0.13), that is set equal to 100. The resulting signal is then converted to Sim2 using Sim2Convert, v0.8.1. Clipping to BT.709 color gamut is applied for P3D65 content.

Content with frame rate above 30 fps are actually displayed at 30 fps.

# Tests results

Solutions HLG, SL-HDR1 and CRI were evaluated on the eight test sequences. For solution TMI, only the results for the four JCTVC test sequences, Hurdles, Starting, Market and BalloonFestival were used, as the other ones were not available from contribution JCTVC-Y0042 [5].

The results are summarized in Table 3, Figure 11 and Figure 12. Average scores and confidence intervals are derived as described in Rec. ITU-R BT.500-13 [7]. The results show that, in average, HLG solution has been assessed around a MOS of 5. CRI and SL-HDR1 have higher average MOS scores, 5.6 and 6.8 respectively. The difference is larger for content with high mastering peak luminance (above 1250 cd/m2), with an average MOS difference of up to 2.2 MOS points between SL-HDR1 and HLG. The difference is noticeably less significant for content with mastering peak luminance below 1250 cd/m2, where an average MOS difference of 1.2 points between SL-HDR1 and HLG. CRI-based solution is in between these two solutions, which can be explained by the fact that the reference SDR used for the CRI derivation was the ones resulting from SL-HDR1. Again, the difference is lower for content below 1250 cd/m2 (–0.7 average MOS difference compared to SL-HDR1) than for content above 1250 cd/m2 (–1.2 average MOS difference compared to SL-HDR1). TMI-based solution was assessed as equivalent to HLG, with a score around 4.1, but the assessment was only based on four sequences with mastering peak luminance equal to or higher than 3000 cd/m2.

Plots for each individual sequence are provided in Annex A.

Table 3. Average MOS score values.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | HLG | SL-HDR1 | CRI | TMI |
| all seqs | 4.9 | 6.8 | 5.6 |  |
| <= 1250 cd/m2 | 5.3 | 6.5 | 5.8 |  |
| > 1250 cd/m2 | 4.7 | 6.9 | 5.5 |  |
| TMI seqs | 4.2 | 6.9 | 5.6 | 4.1 |

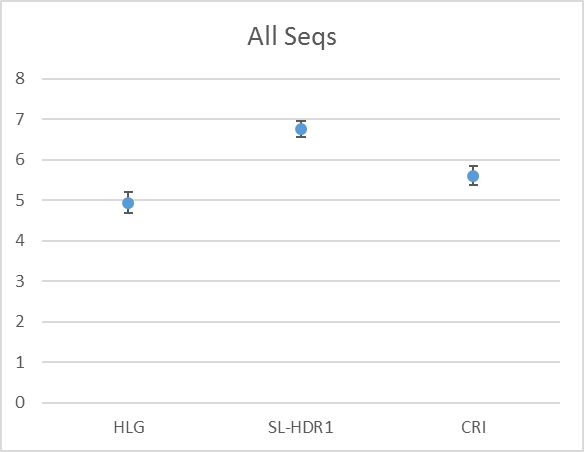
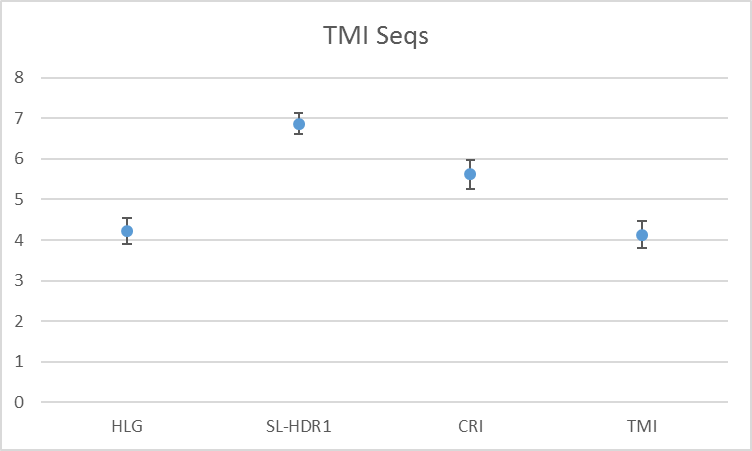
 

Figure 11. Average MOS score over all sequences, and for sequences used in TMI solution.

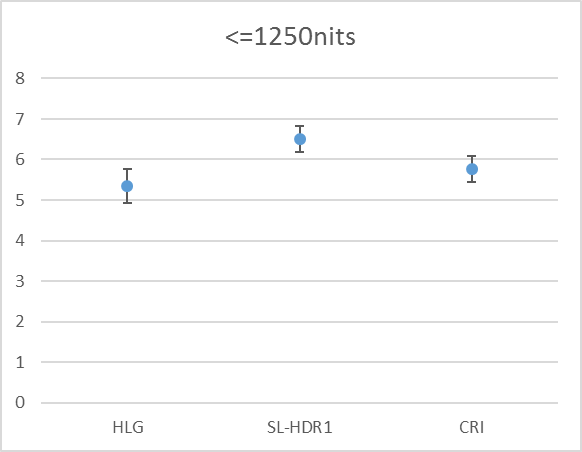
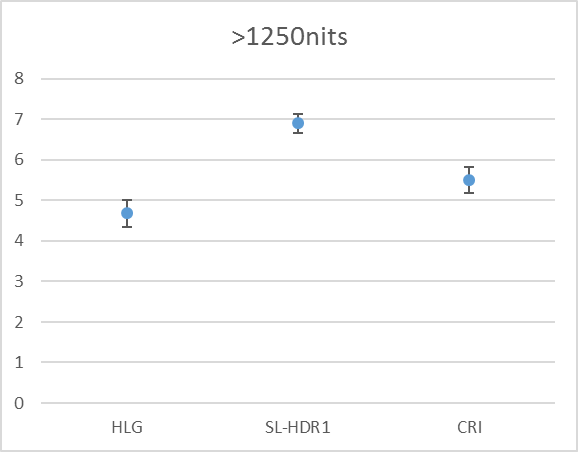
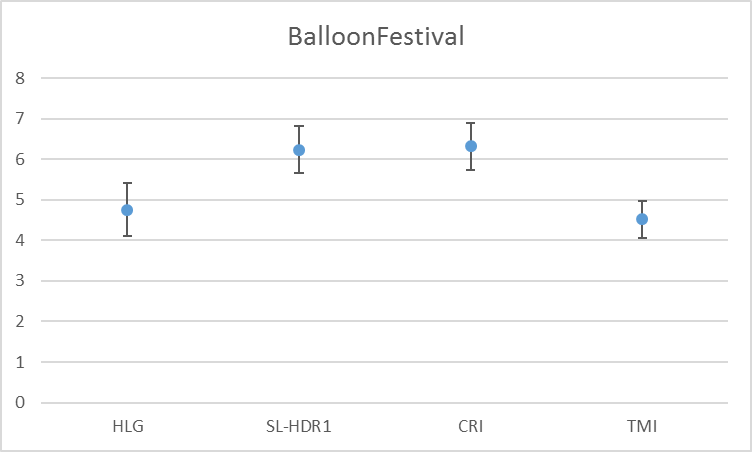
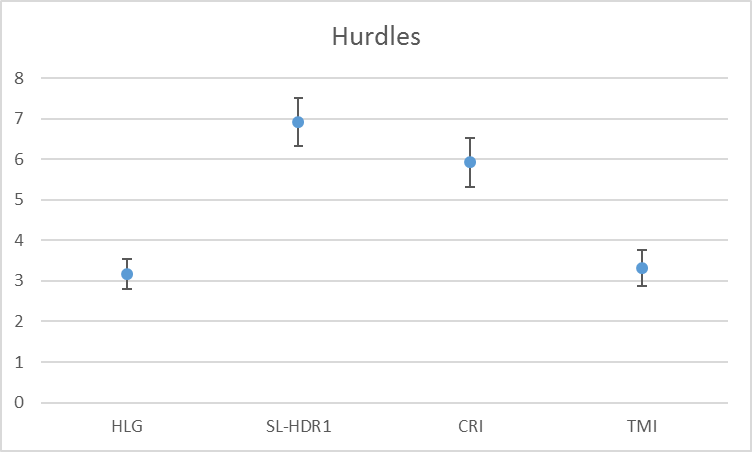
 

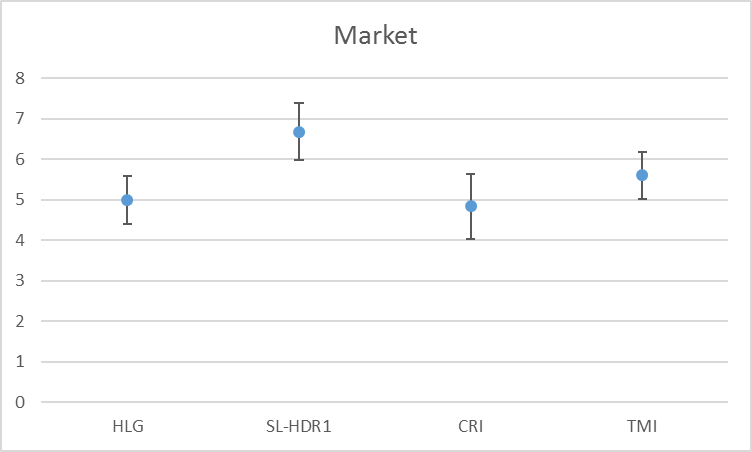
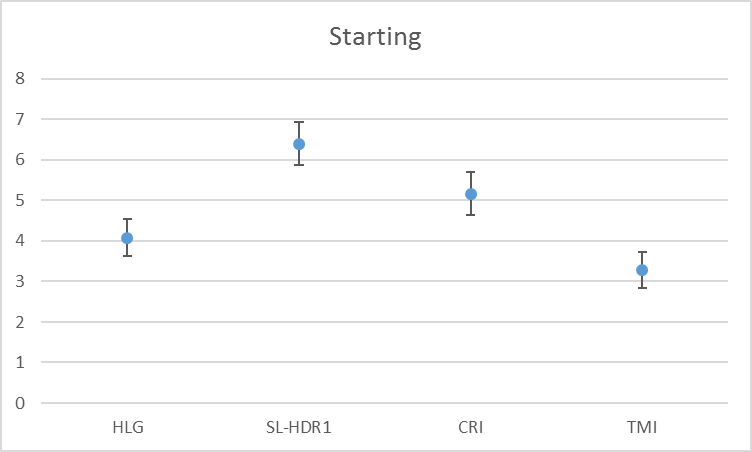
Figure 12. Average MOS score over sequences below and above 1250 cd/m2.

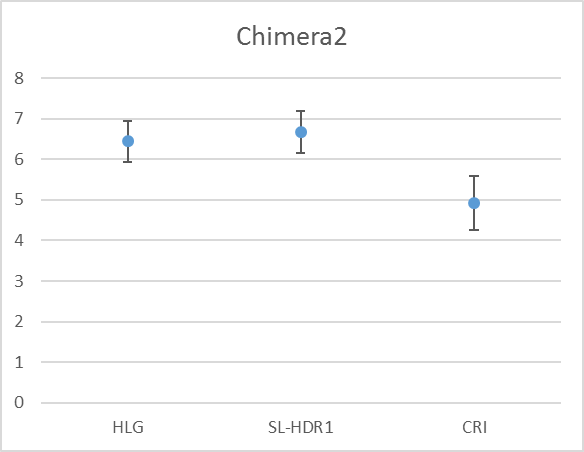
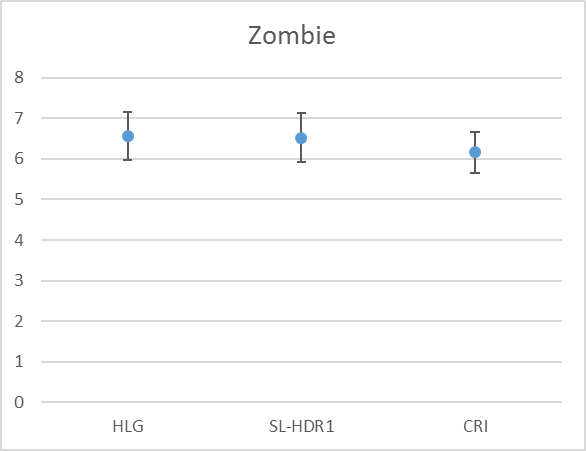
# References

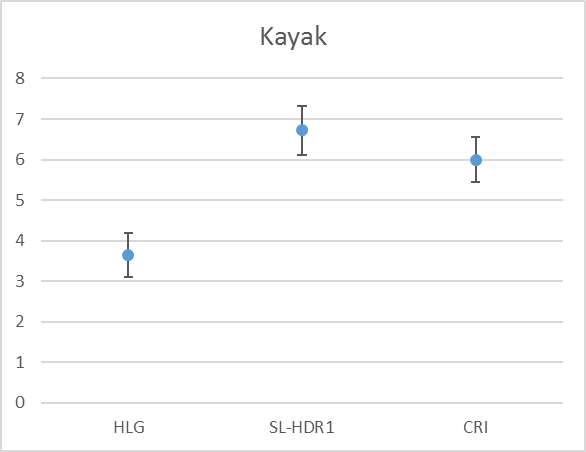
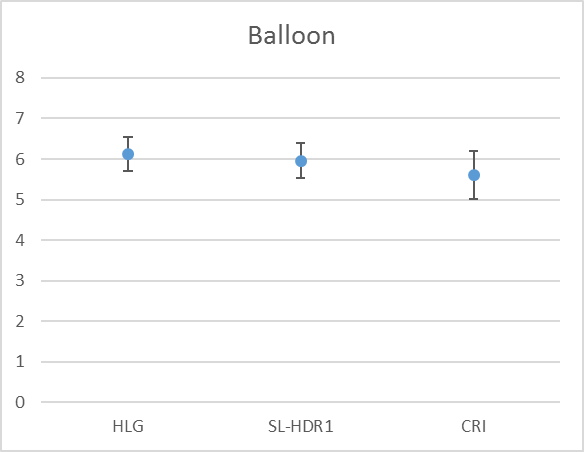
1. E. François, D. Rusanovskyy, G. J. Sullivan, P. Topiwala, P. Yin, M. Naccari, “Signalling, Backward Compatibility, and Display Adaptation for HDR/WCG Video Draft 1”, JCTVC-Y1012, 25th JCT-VC Meeting, Chengdu, CN, Oct. 2016.
2. ETSI TS 103 433 (2016), “High-Performance Single Layer Directly Standard Dynamic Range (SDR) Compatible High Dynamic Range (HDR) System for use in Consumer Electronics devices (SL-HDR1)”.
3. S. Lasserre, F. Le Leannec, E. François, T. Poirier, “Modulation-based HDR video coding with SDR backward compatibility”, JCTVC-U0085, 21th JCT-VC Meeting, Warsaw, PL, Jun. 2015.
4. K. Minoo, T. Lu, P. Yin, L. Kerofsky, D. Rusanovskyy, E. Francois, Description of the reshaper parameters derivation process in ETM reference software”, JCTVC-W0031, 23th JCT-VC Meeting, San Diego, USA, Feb. 2016.
5. P. Topiwala, W. Dai, M. Krishnan, “AHG14: Tone Mapping and Related SEIs for HDR Coding”, JCTVC-Y0042, 25th JCT-VC Meeting, Chengdu, CN, Oct. 2016.
6. V. Baroncini, Proposed test methodology of SDR quality evaluation with HDR reference, MPEG document ISO/IEC JTC1/SC29/WG11 MPEG2015/M37407, Geneva, CH, October 2015.
7. Recommendation ITU-R BT.500-13 (01/2012), “Methodology for the subjective assessment of the quality of television pictures”.

# Annex A – MOS scores per sequence

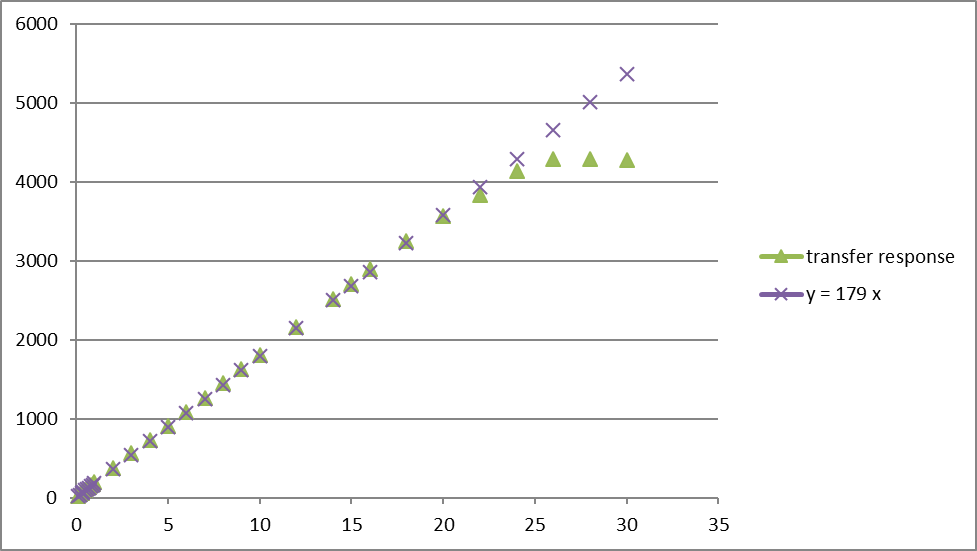
# Annex B – Characteristics of the display using for the tests

The display used for viewing the content is a SIM2 display, with main characteristics provided in the table below.

|  |  |
| --- | --- |
| Technology | HDR LCD Display with individually controlled LED backlight modulation |
| Resolution | 1920 x 1080 pixels |
| Display size | 47” |
| Panel aspect ratio | 16:9 |
| Number of real colours | 16 bits per channel |
| Number of LED | 2202 |
| Brightness | 4000 cd/m2 |
| ANSI Contrast | >20.000:1 |
| FULL ON/OFF Contrast | virtually infinite (>1.000.000:1) |
| White point | 6500K (totally adjustable 5000k–9000k) |

The transfer response and the color primaries of the unit used to grade the content have been characterized using the transfer characteristics of the EasyHDRPlayer v2.0.0 tool provided by Sim2 using default shader version 3.0 equations provided in the tool.

The transfer response is plotted below. It is observed that the response of the display is linear up to 4000 cd/m2.



Each color primary has been characterized using a different test file with the same pixel values for all pixels of the file:

* For Red primary : R = 1.0f, G = 0.0f, B = 0.0f
* For Green primary : R = 0.0f, G = 1.0f, B = 0.0f
* For Blue primary : R = 0.0f, G = 0.0f, B = 1.0f

The results are:

* For Red primary : x = 0.6215, y = 0.3388
* For Green primary : x = 0.2864, y = 0.6319
* For Blue primary : x = 0.1593, y = 0.0677