|  |  |
| --- | --- |
| **Joint Collaborative Team on Video Coding (JCT-VC)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  17th Meeting: Valencia, ES, 27 March – 4 April 2014 | Document: JCTVC-Q0190  M33077 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Evaluation of distortion metrics on HDR video content** | | |
| *Status:* | Input Document to JCT-VC | | |
| *Purpose:* | Information | | |
| *Author(s) or Contact(s):* | Edouard François, Patrick Lopez, F. Le Léannec, S. Lasserre Technicolor 975 avenue des champs blancs CS 17616-35576 Cesson-Sévigné Cedex, France | Tel: Email: | +33 (0)2 99 27 37 21 [edouard.francois@technicolor.com](mailto:edouard.francois@technicolor.com) |
| *Source:* | Technicolor | | |

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

# Abstract

This contribution reports results of measuring different distortion metrics on HDR video content. The metrics are evaluated for different HDR video codecs at different operational points. One goal is to provide input data to identify possible distortion metrics helping to evaluate HDR video coding solutions.

# Introduction

In the context of the JCT Adhoc 18 on high bitdepth investigations and of the MPEG Adhoc on HDR/WGC, the problem of the objective evaluation of HDR video coding solutions has been raised. The topic of perceptual quality metric for video is highly complex and has been leading to substantial studies. When moving to HDR and WCG content, the problem is increasingly more complex.

Usually, when dealing with non-HDR Y’CbCr video content (e.g. BT.709 content), it is very common to use PSNR measured on each of the component as an objective distortion metric to evaluate or compare video coding solutions. As such metrics have been used for decades, people are used and able to interpret them. But when dealing with HDR content, such metrics are much more questionable and alternate metrics should also be considered.

This contribution aims at providing input data that could potentially enable identifying a set of distortion metrics to help in evaluating and comparing different HDR video coding solutions. Nevertheless the intent is not to investigate the field of perceptual quality metric or to propose simple distortion metrics as possible perceptual quality metrics. The evaluated metrics and test conditions are given in the next sections.

# Evaluated distortion metrics

The considered HDR content correspond to the 5 sequences described in document JCTVC-P0228. These sequences are represented in the linear-light half-float format, BT.709 RGB, 1080p.

Two types of metrics are considered: metrics computed in the OETF/EOTF domain (after having applied the PQ-EOTF), and metrics computed in the Linear or Lab color space (without OETF/EOTF).

We have considered two types of test chains:

* The first case, illustrated in Figure 1, involves before the encoding/decoding steps a mapping of the input half-float linear-light HDR content into a quantized non-linear space using the PQ-OETF/PQ-EOTF and the conversion to the Ydzdx space. The following configurations are considered:
  + PQ-EOTF+Y’dz’dx’ 10bits 4:2:0
  + PQ-EOTF+Y’dz’dx’ 12bits 4:4:4
  + PQ-EOTF+ Y’dz’dx’ 12bits 4:4:4 + chroma qp offset (-9)



Figure 1: encoding/decoding working with non-linear input/ouput data.

* The second case, illustrated in Figure 2, directly operates the encoding/decoding with the linear-light HDR data. The tested codec (TCH codec) is based on the design previously described in contribution m32076/JCTVC-P0159, and consists in decomposing the HDR signal into two signals, one LDR signal of full resolution and one low frequency signal of significantly reduced resolution. In the version tested in this contribution, the constraint of a viewable LDR signal on a standard LDR display has been removed, and the LDR signal is encoded using the HM\_RExt 4:4:4 10bits.



Figure 2: encoding/decoding working with linear input/ouput data.

The following metrics have been measured:

**EOTF domain (after EOTF and Y’dz’dx’ or Y’U’V’ conversion)**

* PSNRY\_EOTF / PSNRU\_EOTF / PSNRV\_EOTF
  + computed on the Y, U, V or Y, dz, dx components as  PSNR = 10.log10( 2N / MSE ) where
    - MSE is the mean square error of the component, N the bit-depth of the signal
* PSNRYUV\_EOTF = 10.log10( 2N / MSE\_YUV )
  + MSE\_YUV= (6.MSE\_Y+MSE\_U+MSE\_V) / 8
* SNRY\_EOTF / SNRU\_EOTF / SNRV\_EOTF

*Note :*

* *for PQ-EOTF+Y’dz’dx’ scheme, the measures are performed on the Ydzdx components ;*
* *for TCH scheme, the measures are performed on the YUV components.*

**Linear domain**

* SNR\_R / SNR\_G / SNR\_B
* MSE\_R / MSE\_G / MSE\_B

**Lab domain**

* MSE\_DE where
  + MSE\_DE is the mean of square E2000 averaged over the successive frames
* PSNR\_DE = 10.log10(65504 / MeanDE)
  + 65504 is the maximum value of the half-float representation
  + MeanDE is the mean of E2000 averaged over the successive frames
* var\_DE : variance of DE averaged over the successive frames
* PSNR\_L1 / MAD\_L1
  + After conversion to Lab color space, the mean of absolute value MAD\_L1 of the L component error is computed.
  + PSNR\_L1 = 10.log10(65504 / MAD\_L1)
* PSNR\_L2 / MSE\_L2
  + After conversion to Lab color space, the mean square error MSE\_L2 of the L component error is computed. Then
  + PSNR\_L2 = 10.log10(655042 / MSE\_L2)

*Note : DE stands for E2000 [1]*

Some other metrics have been evaluated but finally not considered because they turn out to be difficult to interpret:

* min and max values of E2000 per frame.
* the limit value of the 1st decile, averaged over the successive frames.
* the limit value of the 9th decile, averaged over the successive frames.

# Some comments

The attached excel file provides the full data for the different tested configurations and metrics. Figures 3 to 6 show 4 examples of metrics (PSNRY’, PSNRY’U’V’, PSNR\_L2 and PSNR\_DE) for 4 sequences. For each sequence, the 2 left graphs correspond to metrics in the EOTF domain and the 2 right graphs to metrics in the Lab domain.

It can be observed that PSNRY’ and PSNR\_L2 show very similar trends, which is logical since they both correspond to a metric applied to a non-linear luminance signal (derived in two different ways). These metrics are of course not able to capture chroma issues. In particular, they do not reflect that the PQ-EOTF+Ydzdx 444 version with a negative chroma QP offset results in a better perceptual quality than the PQ-EOTF+Ydzdx 444 version without chroma QP offset.

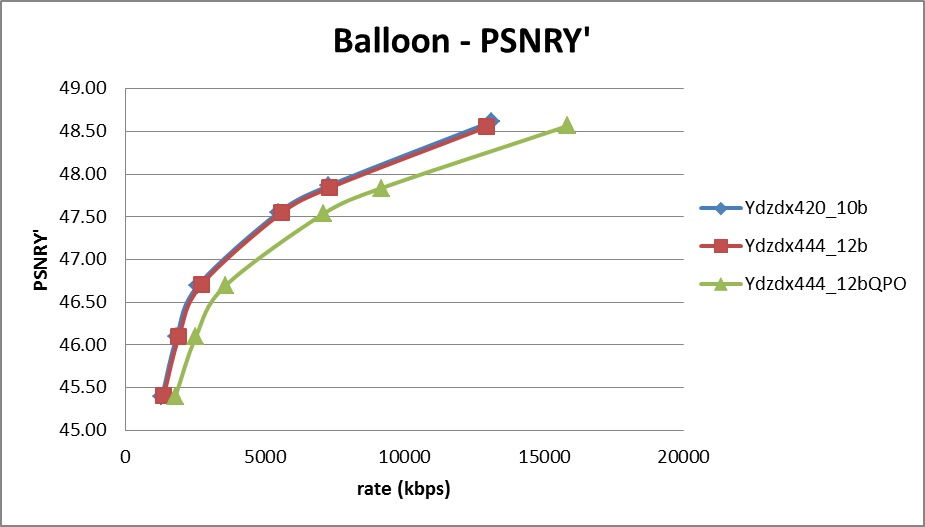
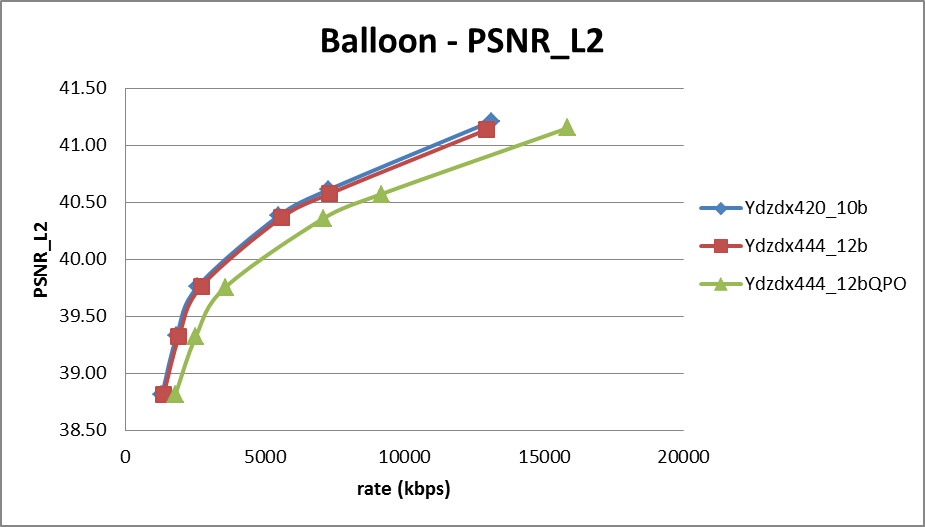
The PSNRY’U’V’ in the EOTF domain partly corrects this issue, but as observed for instance on the Balloon, Market3 or Seine sequences, it is not possible with this metric to clearly identify the perceptual advantage of the use of a negative chroma QP offset in the PQ-EOTF approach.

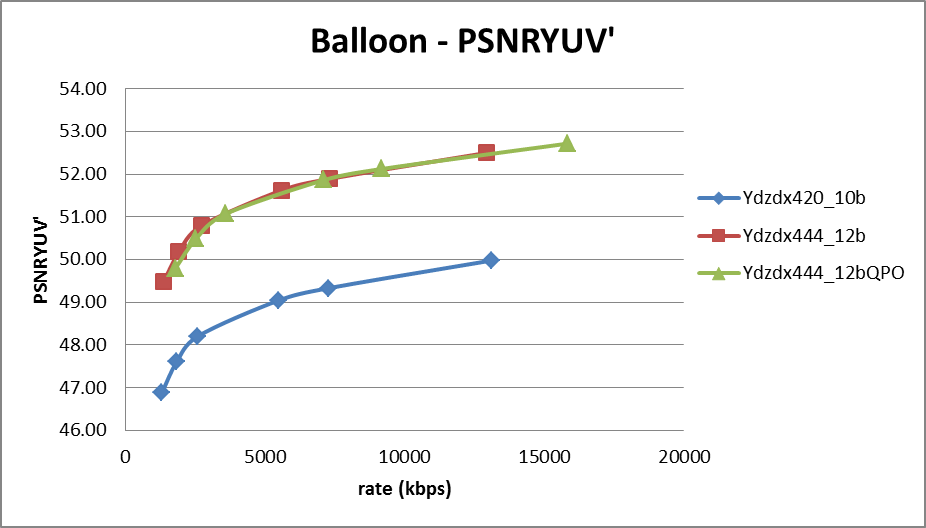
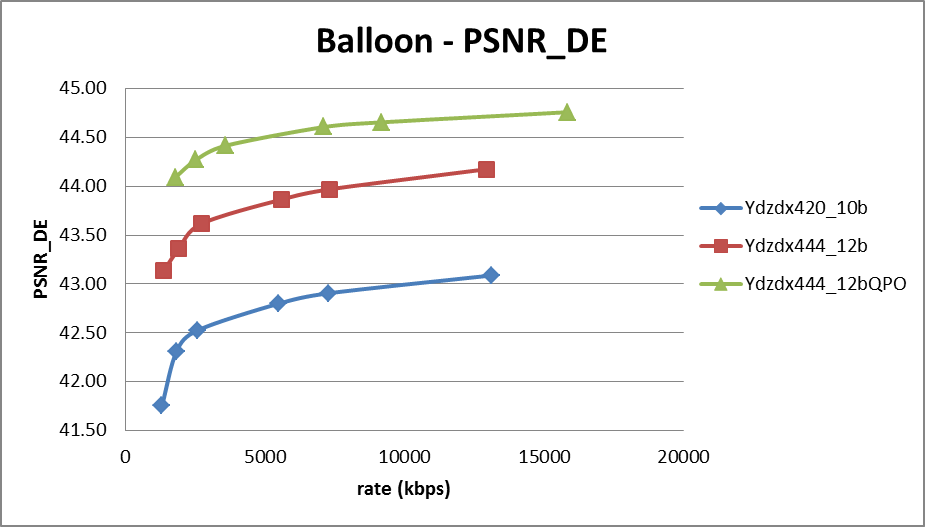
The PSNR\_DE metric turns out to generate more differentiated results. For instance, for Balloon and Market3, PSNR\_DE curve of PQ-EOTF+Ydzdx 444 version with a negative chroma QP offset is clearly above the PSNR\_DE curve of PQ-EOTF+Ydzdx 444 version without chroma QP offset.

Based on internal visual testing, we also notice that the PSNR\_DE is very well correlated with the visual perception and is able to rather well classify the performance of different codecs or settings.

*Notes :*

* *we have observed that the results for PQ-EOTF-Ydzdx using the HM12.0\_RExt5.0 4:2:0 12b are slighly worse than using the HM13.0\_RExt6.0 4:2:0 10b. This seems to come from the encoding process of HM\_RExt.*
* *Missing results for TCH codec will be added in a revised version.*

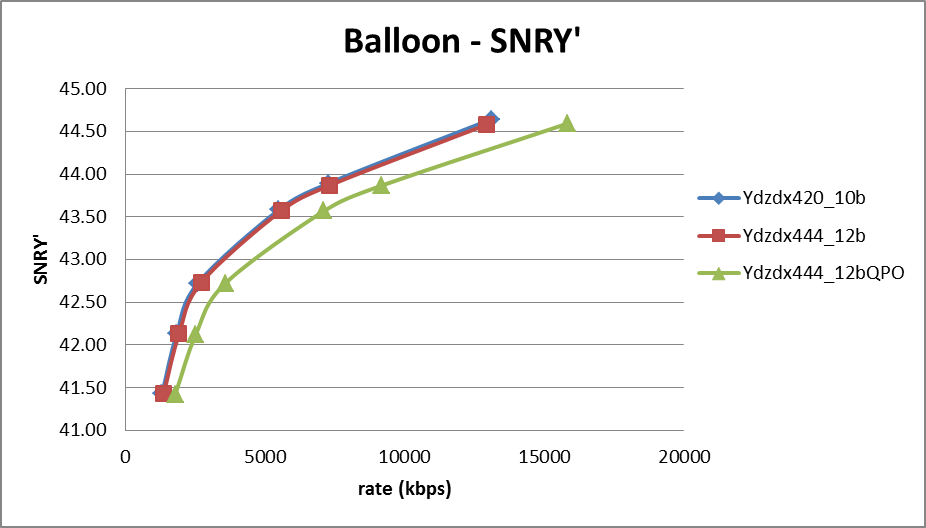
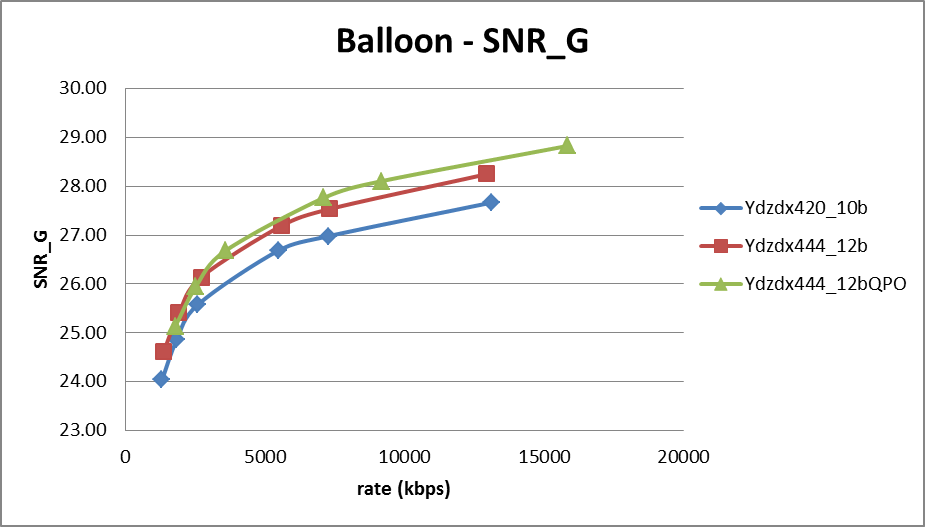
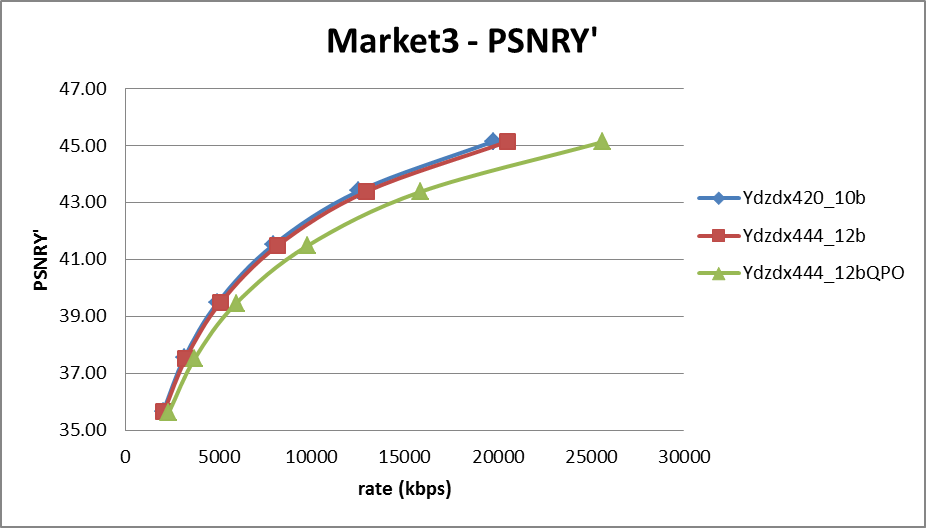
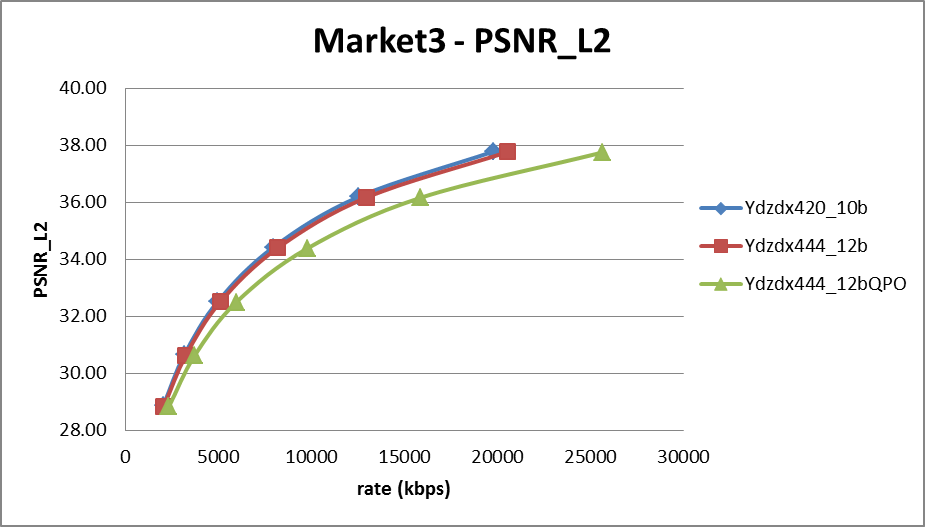
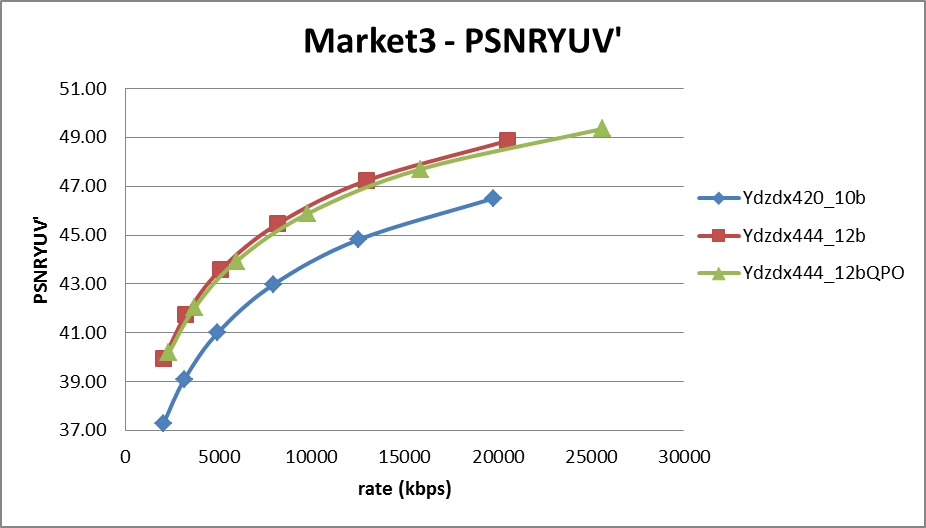
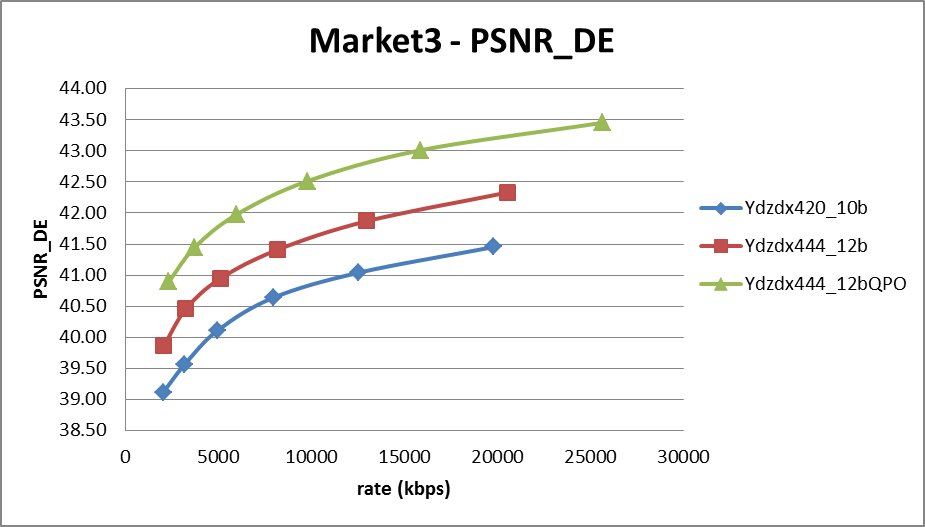
 

Figure 3: distortion metris on Balloon – left side: EOTF domain, right side: Lab domain.

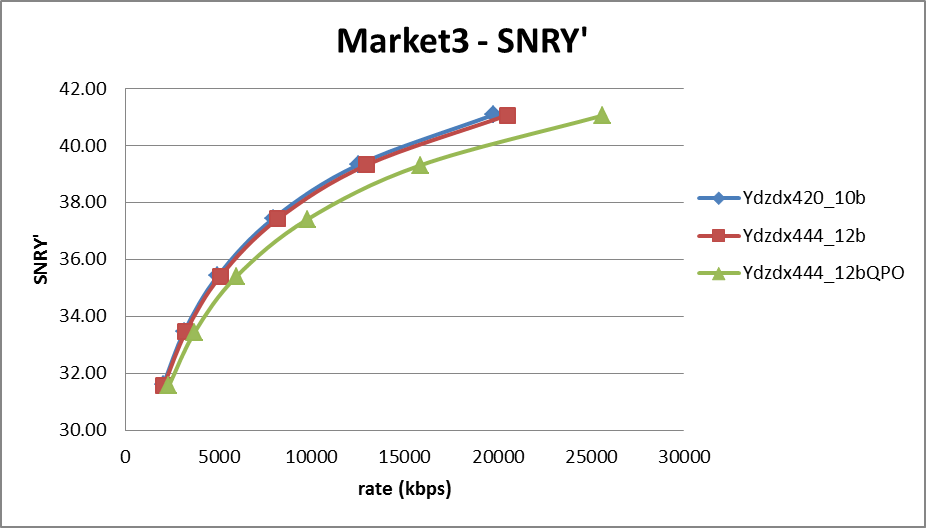
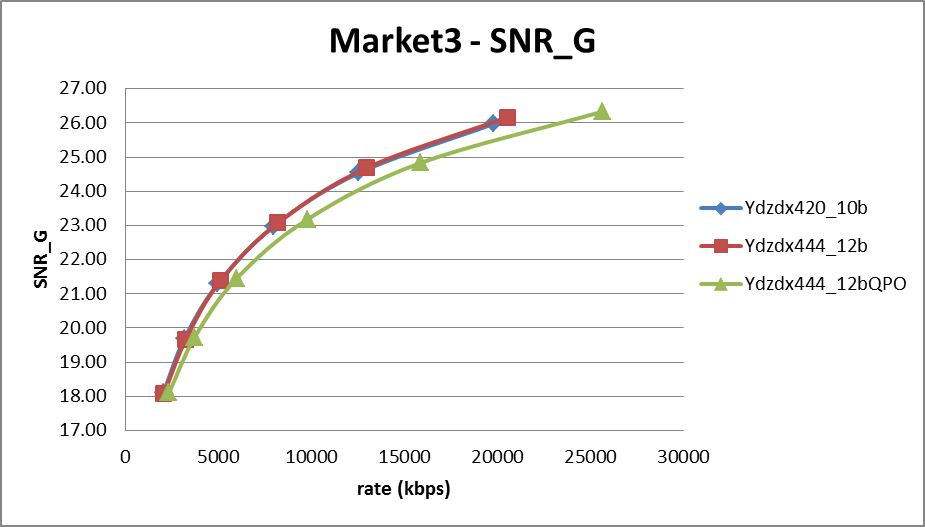
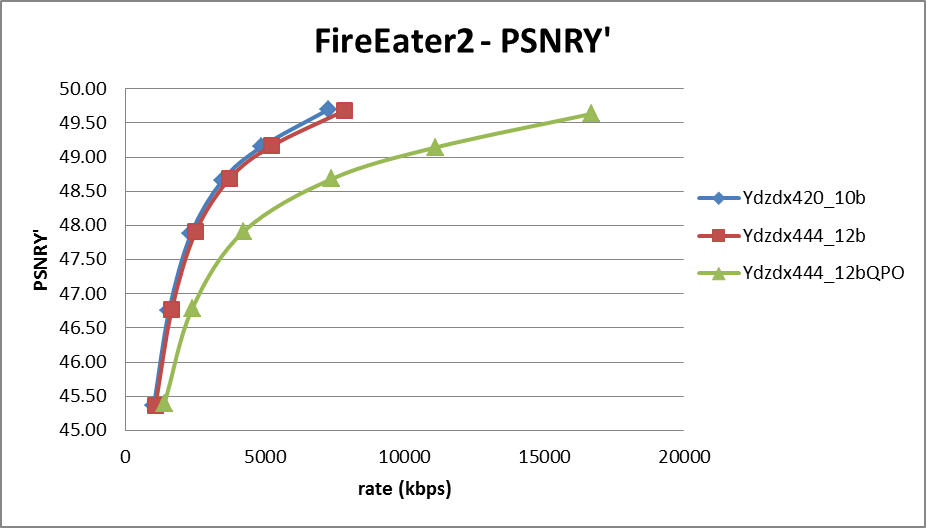
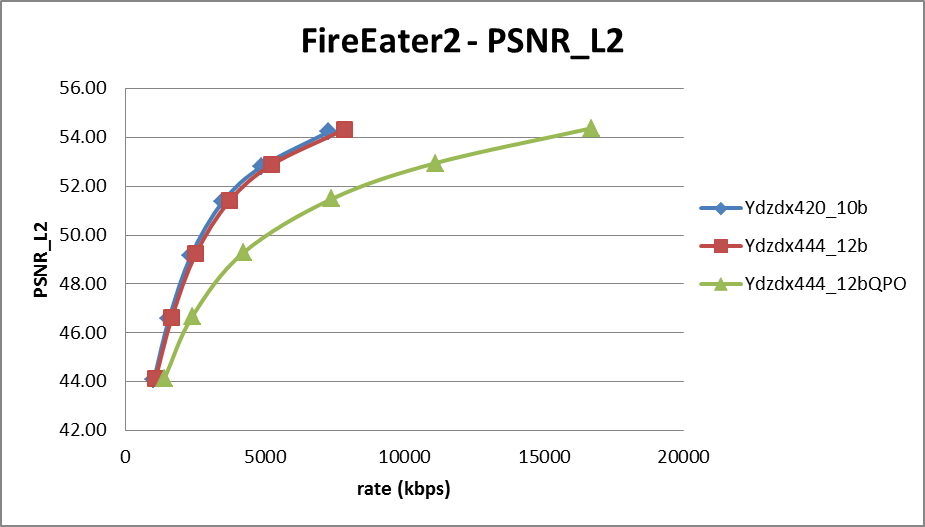
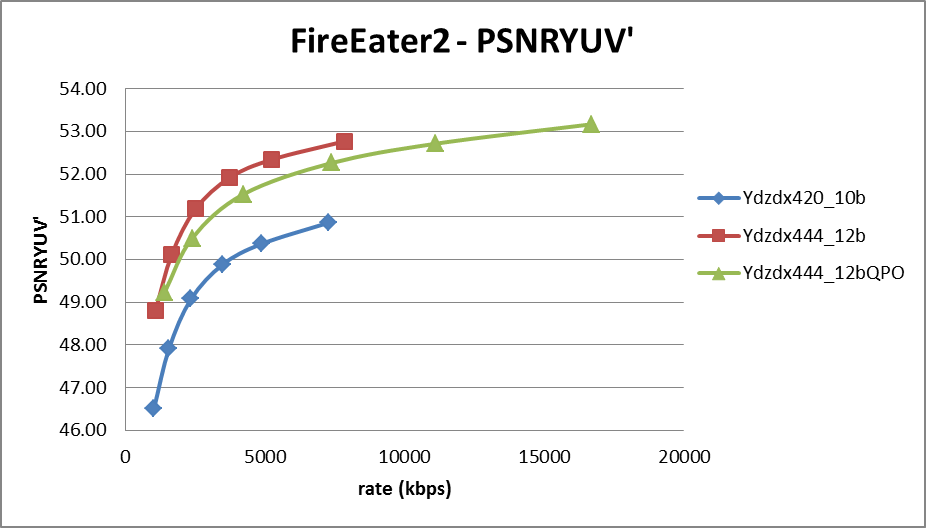
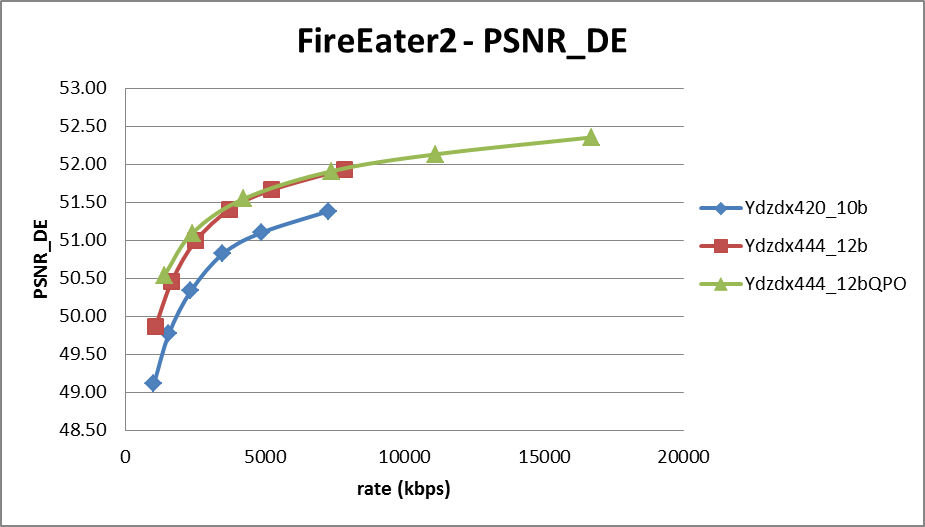
 

Figure 4: distortion metris on Market3 – left side: EOTF domain, right side: Lab domain.

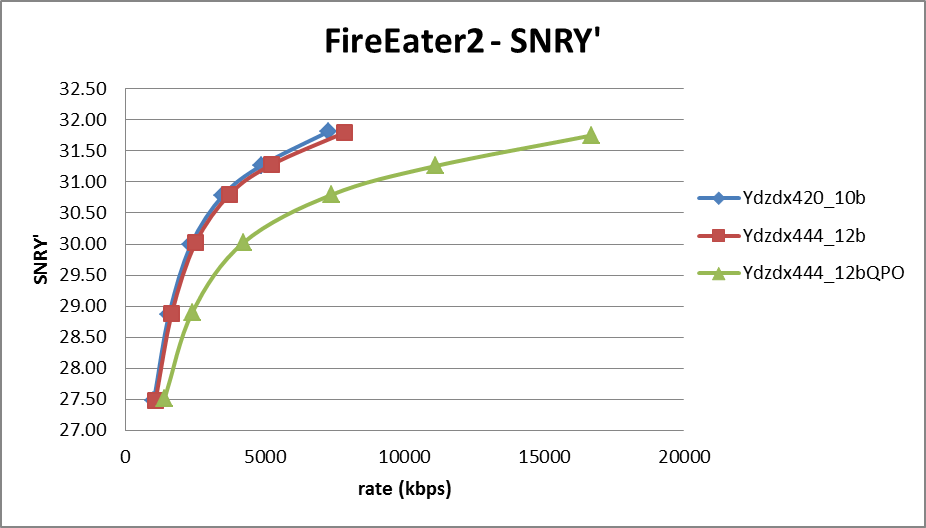
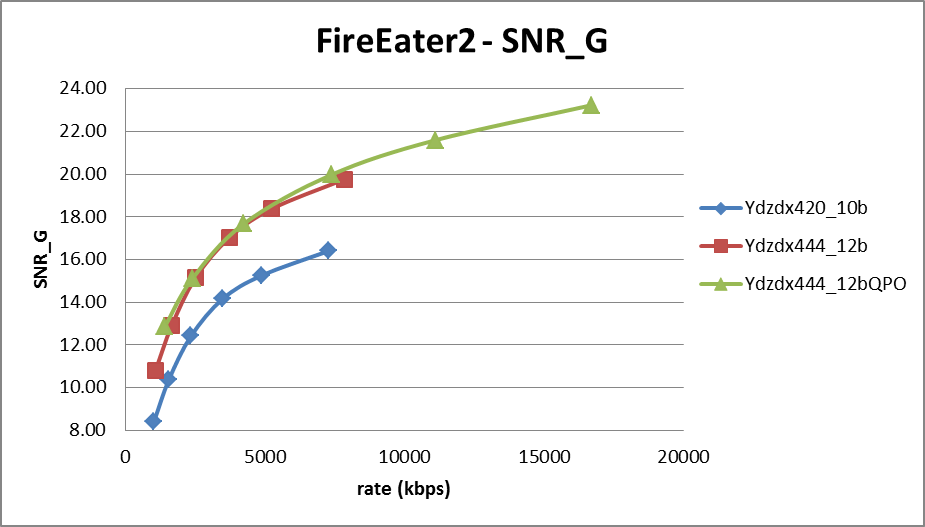
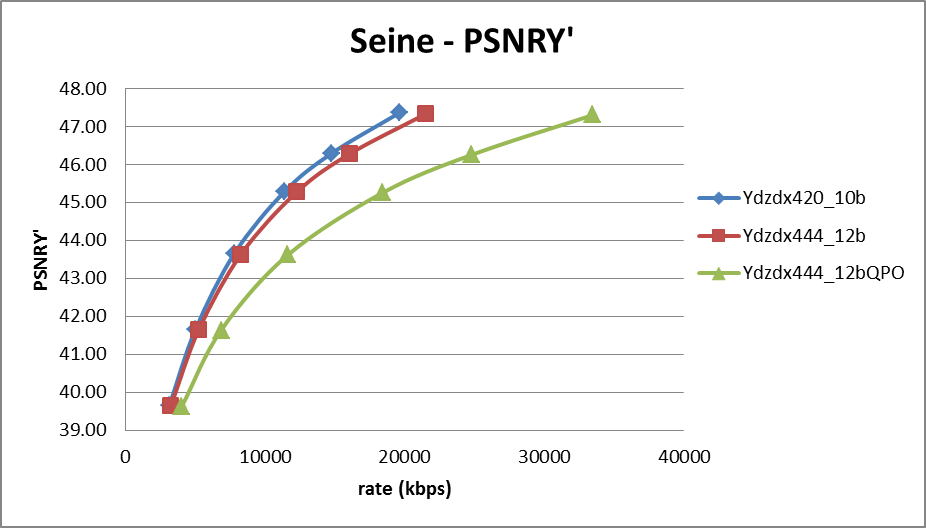
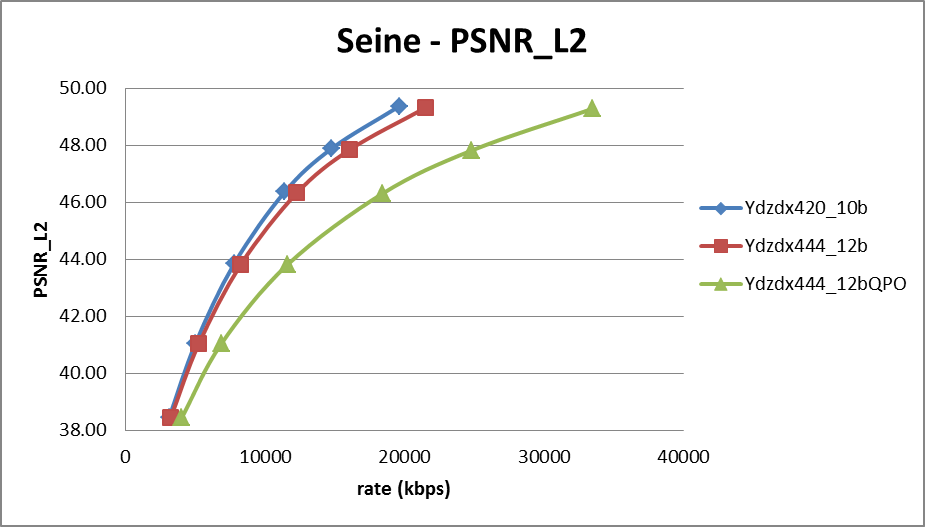
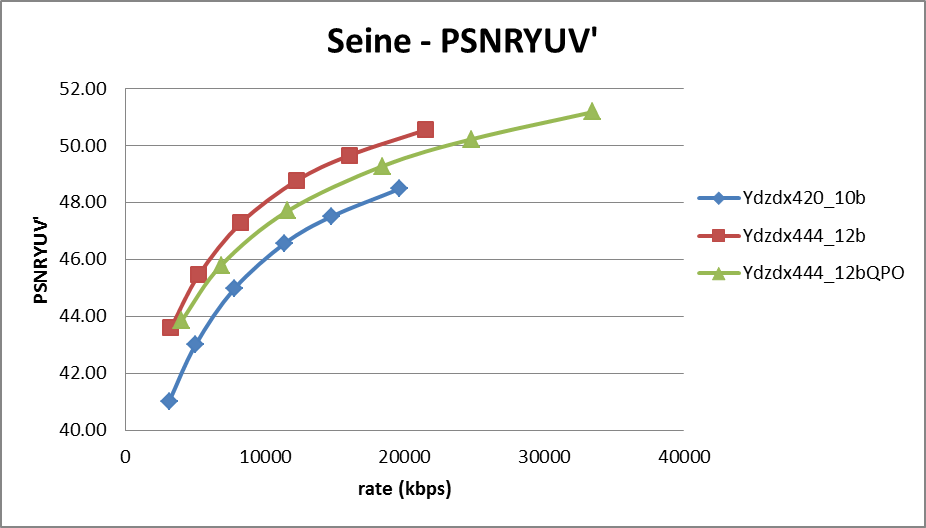
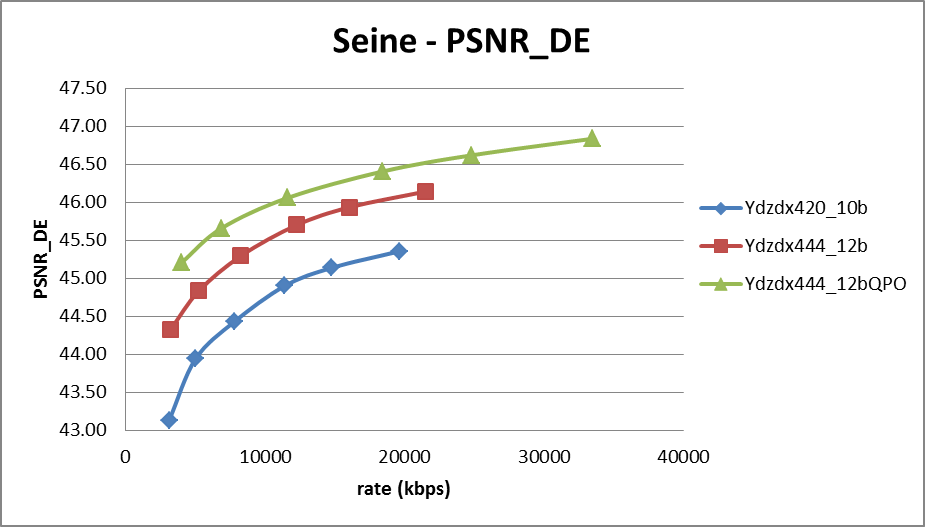
 

Figure 5: distortion metris on FireEater2 – left side: EOTF domain, right side: Lab domain.

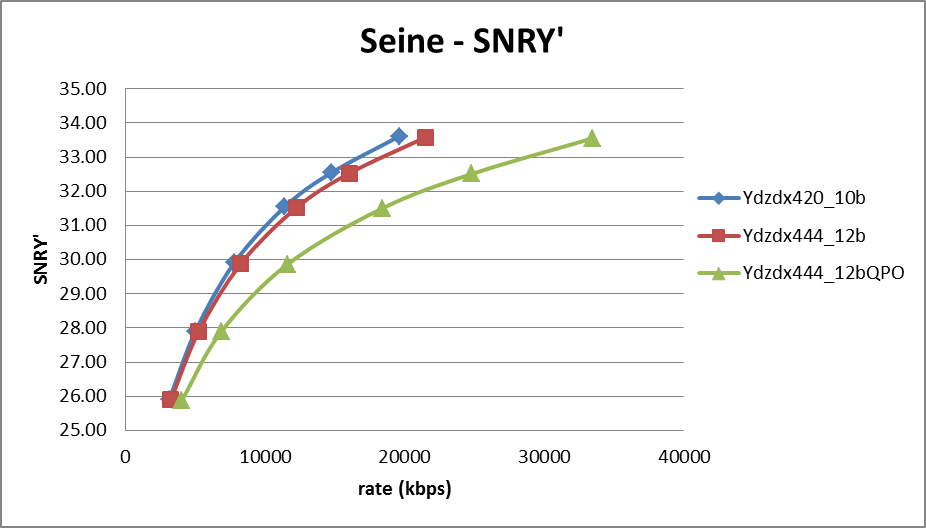
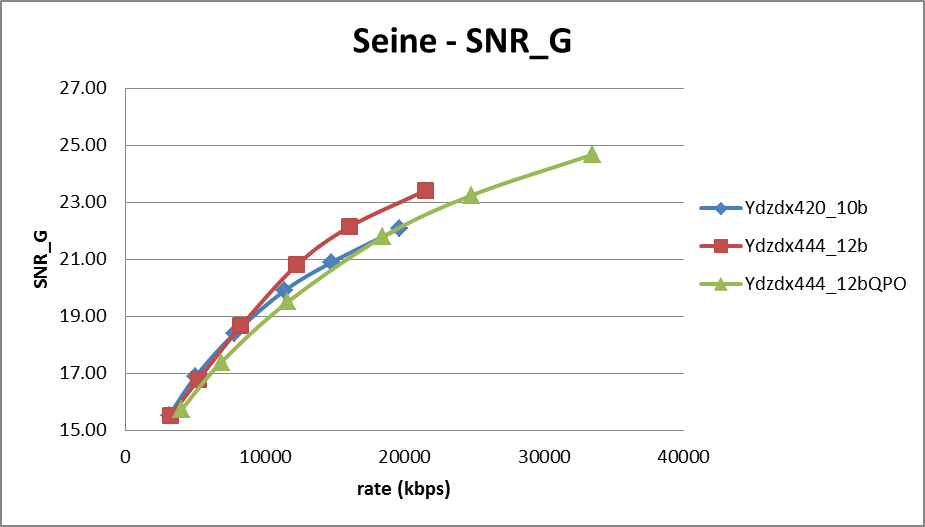
 

Figure 6: distortion metris on Seine – left side: EOTF domain, right side: Lab domain.

# Patent rights declaration

**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).**