Luma HDRv: an open source high dynamic range video codec optimized by large-scale testing

Gabriel Eilertsen[1] Rafał K. Mantiuk[2] Jonas Unger[1]
[1]Linköping University, Sweden [2]Computer Laboratory, University of Cambridge, UK

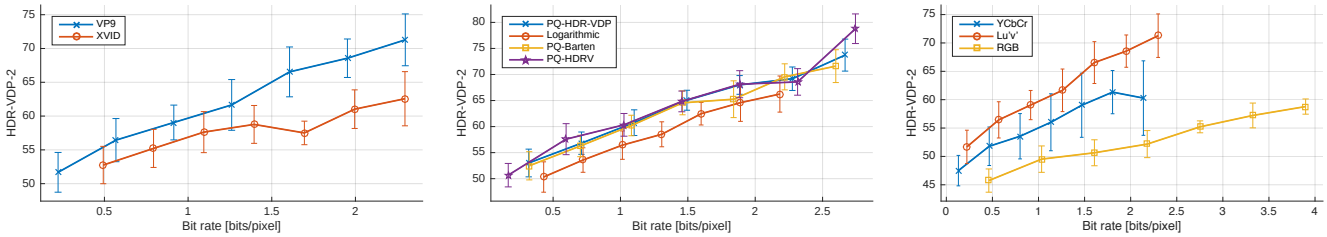

Figure 1: *Bit rate plots, comparing 2 codecs* ***(left)****, 4 luminance transfer functions* ***(middle)****, and 3 colorspaces* ***(right)****. The plots are the result of comparing 33 HDR video sequences in terms of HDR-VDP-2 quality predictions. The errorbars represent standard errors.*

Abstract

We present *Luma HDRv* – an open source solution for encoding of high dynamic range (HDR) video. The software makes use of techniques for adapting the HDR video for compression with a standard video codec. In the design of the encoder we perform a large-scale test, using 33 HDR video sequences in order to compare 2 video codecs, 4 luminance encoding techniques (transfer functions) and 3 color encoding methods. This serves both as an evaluation of existing techniques for encoding of HDR luminances and colors, as well as to optimize the performance of *Luma HDRv*.

Keywords: high dynamic range (HDR) video, HDR video coding, open source software, perceptual image metrics

Concepts: •**Information systems** → **Open source software;** •**Computing methodologies** → *Image processing;*

1 Introduction

High dynamic range (HDR) video is one of the key components in the development of next generation imaging technologies. An important challenge within HDR is the development of efficient solutions for storage and distribution of video content. Inter-frame encoding of HDR video is typically done using existing high-bit depth video compression algorithms, where the HDR input is mapped to a format suited for the encoder. Although techniques for this purpose are under active development, they lack in comprehensive evaluation. Also, while there are a few ongoing initiatives towards HDR video compression, and the first steps are being taken in standardization, most existing solutions are proprietary, and HDR video encoding software is currently not available on open source terms.

We address these problems with our open source HDR video codec, *Luma HDRv* (http://lumahdrv.org). *Luma HDRv* is based on a large-scale test of existing methods for encoding of HDR luminances and colors. Using a number of objective metrics, and comparing on a large range of different input video sequences, we design *Luma HDRv* to have the best possible compression performance. The result is an efficient codec for HDR video with a user friendly API.

SIGGRAPH 2016 Talks, July 24-28, 2016, Anaheim, CA
ISBN: 978-1-4503-ABCD-E/16/07
DOI: http://doi.acm.org/10.1145/9999997.9999999

2 HDR video encoding

In order to utilize the efficiency of existing video codecs to encode HDR, the linear luminances of the HDR data need to be transformed and quantized before the encoding. With a perceptually motivated luminance transformation, this can be achieved in such manner to make quantization errors invisible across luminances, as described in the single stream HDR video encoding method by Mantiuk *et al.* [Mantiuk et al. 2004]. For this method 11 bits were determined to be enough to allow for visually lossless encoding. An alternate approach is to store the HDR video in two separate 8-bit streams; one tone-mapped video, and one containing a residual required to reconstruct the HDR frames [Mantiuk et al. 2006].

For *Luma HDRv* we consider the single stream approach from [Mantiuk et al. 2004], which allows for better compression performance as compared to using multiple streams [Azimi et al. 2015]. For this method, there are three main steps affecting the performance of the final encoding: 1) the codec used for compressing the final bit stream, 2) the transformation from linear luminances to values for encoding, and 3) the space used for representing colors and luminances. In each of theses steps we compare the original techniques to standard techniques, and consider possible new candidates. Then, optimizing for the best performance in a large-scale evaluation, we design an HDR video encoding configuration from the top performing algorithms in each step.

3 Evaluation

To extensively evaluate the performance of the HDR video encoding, we make use of 33 HDR video sequences at 1080p resolution, containing approximately 100 frames per sequence. In total we use 9 conditions at 15 quality settings, which makes for 442 395 pairs of frames to compare. To make this possible, we employed a large-scale computer cluster, where all calculations could be performed within a few days.

The sequences are taken from the comprehensive collection of HDR videos by Froehlich *et al.* [Froehlich et al. 2014]. To provide a reliable estimate of the perceptual differences between the methods tested, we use objective metrics such as the visual difference predictor for HDR images (HDR-VDP-2, v2.2) [Mantiuk et al. 2011]. HDR-VDP-2 is a full-reference metric using a model of the HVS to assess visual differences, and it has been demonstrated to correlate well with subjective studies [Artusi et al. 2015].

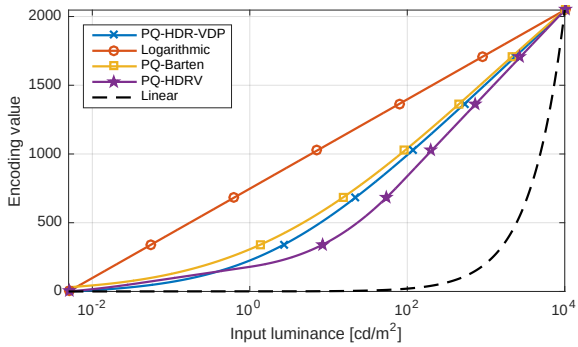

Figure 2: *Luminance transfer functions, mapping physical units to integer values for encoding. Here, luminances in the range* $[0.005, 10^4]$ *cd/m*^{2} *are mapped to 11 bits,* $[0, 2047]$.

Video codec: A number of existing video codecs provide the precision needed for visually lossless compression of HDR video, with high bit depth profiles for encoding at up to 12 bits. Since our goal was to release *Luma HDRv* under open source terms, we selected Google's VP9 codec as its license permits such use. The performance of the codec is on par with the widely used H.264 standard [Řeřábek and Ebrahimi 2014].

To confirm that VP9 is a suitable option, we evaluated its performance against the XVID MPEG-4 Part 2 compression scheme used in the original single stream HDR video encoding, [Mantiuk et al. 2004]. Figure 1 (left) shows a comparison of the codecs using HDR-VDP-2 (v2.2) quality predictor "Q". It is evident that VP9 clearly outperforms XVID using only about half the bit rate for the same quality.

Luminance encoding: Similarly as a "gamma" transfer function is required for standard dynamic range content, linear HDR pixel values need to be compressed with an appropriate transfer function before encoding. The simplest choice, justified by the approximately logarithmic response of the eye (according to the Weber–Fechner law), is to encode values in the logarithmic domain. However, more sophisticated methods have been derived using perceptual measurements. These are referred to as perceptual transfer functions (PTFs) or electro-optical transfer functions (EOTFs), and their purpose is to translate linear floating point luminances to the screen-referred integer representation of an encoding system, with quantization errors that are perceptually equally distributed across all luminances.

In our comparison, we consider the four most commonly used transfer functions. One is a logarithmic function, while the three others are perceptually based: PQ-HDRV [Mantiuk et al. 2004], PQ-HDR-VDP [Mantiuk et al. 2011], and PQ-Barten [Miller et al. 2013], see Figure 2. Their performances are compared in Figure 1 (middle), and – as expected – the perceptual transfer functions show a substantial improvement over the logarithmic encoding. Differentiating between these, however, is difficult considering the variance of the measurements. Even though there are no clear evidence in favor of any of the perceptual encodings, we employ PQ-Barten as default for our open source codec. This is based on observations that PQ-Barten distributes distortions more uniformly across luminances as compared to the other methods [Boitard et al. 2015].

Color encoding: Figure 1 (right) compares the performance of encoding the HDR video sequences in RGB, YC_bC_r and $Lu'v'$. For YC_bC_r and $Lu'v'$, the chroma channels have been sub-sampled to half the image width and height (4:2:2 sampling), while RGB uses the full-sized channels. To encode luminances, the PQ-HDRV transformation has been used on the luminance channels, and on RGB channels separately. As expected, RGB is clearly inefficient. Also, although variance is high, comparing average performance $Lu'v'$ shows a great improvement over YC_bC_r, with about half the bit rate for the same quality.

4 Open source codec

Guided by the results of the tests in the previous section, we selected the best combination of codec, transfer function and color coding for the *Luma HDRv* open source video compression software. The software has been released under the BSD license. *Luma HDRv* provides libraries for including the HDR video encoding and decoding in software development, as well as applications to perform encoding and decoding with a number of different settings. Furthermore, we also provide an HDR video player for real-time decoding, playback and tone-mapping of encoded HDR videos.

Luma HDRv uses VP9 for encoding, and the default settings are according to the results discussed above, with PQ-Barten for luminance quantization at 11 bits, and encoding in the $Lu'v'$ colorspace. The encoded HDR videos are stored using the Matroska container (http://www.matroska.org), for flexibility and easy integration into existing software.

References

ARTUSI, A., MANTIUK, R. K., RICHTER, T., HANHART, P., KORSHUNOV, P., AGOSTINELLI, M., TEN, A., AND EBRAHIMI, T. 2015. Overview and evaluation of the JPEG XT HDR image compression standard. *Journal of Real-Time Image Processing*.

AZIMI, M., BOITARD, R., OZTAS, B., PLOUMIS, S., TOHIDYPOUR, H. R., POURAZAD, M. T., AND NASIOPOULOS, P. 2015. Compression efficiency of HDR/LDR content. In *IEEE Quality of Multimedia Experience (QoMEX)*.

BOITARD, R., MANTIUK, R. K., AND POULI, T. 2015. Evaluation of color encodings for high dynamic range pixels. In *Proc. SPIE, Human Vision and Electronic Imaging XX*.

FROEHLICH, J., GRANDINETTI, S., EBERHARDT, B., WALTER, S., SCHILLING, A., AND BRENDEL, H. 2014. Creating cinematic wide gamut HDR-video for the evaluation of tone mapping operators and HDR-displays. In *Proc. SPIE, Digital Photography X*.

MANTIUK, R. K., KRAWCZYK, G., MYSZKOWSKI, K., AND SEIDEL, H.-P. 2004. Perception-motivated high dynamic range video encoding. *ACM Trans. on Graphics 23*, 3.

MANTIUK, R. K., EFREMOV, A., MYSZKOWSKI, K., AND SEIDEL, H.-P. 2006. Backward compatible high dynamic range MPEG video compression. *ACM Trans. on Graphics 25*, 3 (July).

MANTIUK, R. K., KIM, K. J., REMPEL, A. G., AND HEIDRICH, W. 2011. HDR-VDP-2: A calibrated visual metric for visibility and quality predictions in all luminance conditions. *ACM Trans. Graphics 30*, 4.

MILLER, S., NEZAMABADI, M., AND DALY, S. 2013. Perceptual signal coding for more efficient usage of bit codes. *SMPTE Motion Imaging Journal 122*, 4.

ŘEŘÁBEK, M., AND EBRAHIMI, T. 2014. Comparison of compression efficiency between HEVC/H. 265 and VP9 based on subjective assessments. In *SPIE Optical Engineering + Applications*.