CONVERSION AND HEVC COMPRESSION OF HIGH DYNAMIC RANGE (HDR) VIDEO

Jonatan Samuelsson and Jacob Ström

Ericsson, Sweden

ABSTRACT

High Dynamic Range (HDR) video constitutes a new type of video with brighter brights and darker darks compared to conventional video. Recent developments in display technology have made it possible to deliver a more immersive viewing experience through being able to reproduce HDR video. This new video type has caused experts to investigate whether existing compression tools can operate efficiently or whether new tools need to be introduced. In MPEG and VCEG the current state is somewhere in-between: existing tools work well with HDR but adjusting their settings to specifically optimize for HDR video, makes it possible to reduce the bit-rate and improve visual quality. This paper will present background information around compression of HDR video and the work on HDR that has been, and still is being, performed in MPEG and VCEG.

INTRODUCTION

Digital video compression has transformed the way in which television has been consumed and delivered to consumers for more than 25 years. Video coding standards such as MPEG-2 (Part 2) and AVC (MPEG-4 Part 10 / H.264) have played a fundamental role in enabling worldwide interoperability of compressed video. The most recent video coding standard is the High Efficiency Video Coding (HEVC) standard (also known as MPEG-H Part 2 and H.265), which was published in its first version in 2013. The development of HEVC has been performed jointly by the Moving Picture Experts Group (MPEG), formally known as ISO/IEC JTC1/SC29/WG11, and the Video Coding Experts Group (VCEG), formally known as ITU-T Q.6/SG16 in the Joint Collaborative Team on Video Coding (JCT-VC). In recent years, several extensions have been added to HEVC including for example SHVC for scalable video coding and MV-HEVC for multi-view video coding. Common to all of these video coding standards and extensions is that a vast amount of technical work is being conducted by a large number of experts in the field in order to firmly investigate the properties of the technology being included in the standards. Typically, the work of a new standard or extension is preceded by a Call for Evidence (CfE) and/or a Call for Proposals (CfP). These calls encourage submissions relevant to a specific area and further development work is commonly initiated with such inputs as a starting point.

In 2015, MPEG issued a CfE on High Dynamic Range (HDR) and Wide Colour Gamut (WCG) video following the several exploration activities on HDR that had been performed by MPEG starting in 2013. One of the key questions to determine was if technology could be identified which would motivate the creation of a new extension of HEVC specifically for

HDR, but in early 2016 it was determined that no such technology had been identified. However, during the course of the investigations, several findings and differences were identified between how to convert and compress HDR video compared to SDR (Standard Dynamic Range) video. These findings are being documented in a technical report that is expected to be completed in October 2016 and approved in January 2017 as a new part of ISO/IEC 23008 called MPEG-H Part 14. This paper will provide more details of this work and the methods described in MPEG-H Part 14.

BACKGROUND OF HIGH DYNAMIC RANGE VIDEO

Conventional video signals can be referred to as Standard Dynamic Range (SDR) video in order to emphasize the difference from High Dynamic Range (HDR) video. SDR video is created with the target of being displayed on screens ranging from roughly 0.01 cd/m^2 to 100 cd/m^2 corresponding to a 10000:1 relationship between the brightest and the darkest pixels. The reason for this limitation is very simple: up until quite recently consumer display technology had not been able to support a wider dynamic range. However, recently the peak luminance level of consumer displays has been increasing at the same time as the black level has remained constant or even been reduced. There are already TVs in the market supporting peak luminance levels of up to 1000 cd/m^2 and professional monitors reaches up to 10,000 cd/m^2. These displays are based on LCD technology with LED backlight, a method that makes it possible to substantially increase the amount of light being emitted by the display while still preserving good black levels. OLED displays have also been able to reach higher luminance levels (although not as high as LCD with LED backlight) and have the advantage of being able to reach black levels lower than 0.0005 cd/m^2.

ITU-R Recommendations for SDR video

A large number of the fundamental pieces used for representing SDR video are described in Recommendations provided by ITU-R. ITU-R Recommendation BT.709 "Parameter values for the HDTV standards for production and international programme exchange" includes a description of how to convert a linear light RGB signal into a non-linear R'G'B' signal suitable for being represented with a limited bit-depth, typically 8 bits per component. This conversion is commonly called the OETF (opto-electronic transfer function). BT.709 also includes definitions of colour primaries (expressed in CIE1931 chromaticity coordinates) and a description of how to convert from red, green and blue (R'G'B') to luma and chroma (Y'CbCr), which is called a colour space conversion. The human eye is much more sensitive to differences in brightness compared with differences in colour and therefore this colour space conversion is performed in order to enable sub-sampling of colour information (by reducing the spatial resolution of the chroma components) while keeping full luma resolution. It should however be noted that the luma component does not correspond directly to the actual luminance and that some of the luminance information is carried in the chroma components. This has the effect that modifications in the chroma components (such as sub-sampling) can lead to luminance changes. This is sometimes called *chroma leakage* and as it turns out, chroma leakage is much more severe for HDR video than for SDR video. This is explained in more detail in the section on luma adjustment below. ITU-R Recommendation BT.1886 "Reference electro-optical transfer function for flat panel displays used in HDTV studio production" describes the conversion from a non-linear R'G'B' signal to a linear RGB signal. This conversion is performed before the video is displayed and is called the electro-optical transfer function (EOTF). It should be noted that the EOTF described in BT.1886 is not the inverse of the OETF described in BT.709. The difference between the EOTF and the inverse of the OETF can be represented by a *system gamma* or an OOTF (opto-optical transfer function) but will not be described further in this paper. In short it can be said that without system gamma most images would look unnaturally pale and dull and the system gamma is a straightforward way of making images look better and more realistic. Finally, ITU-R Recommendation BT.2035 "A reference viewing environment for evaluation of HDTV programme material or completed programmes" provides reference settings for viewing environments specifying, among other things, the luminance range from 0.01 cd/m^2 to 100 cd/m^2.

ITU-R Recommendation BT.2020 "Parameter values for ultra-high definition television systems for production and international programme exchange" introduces 4K resolution (3840x2160) and Wide Colour Gamut (WCG). WCG means that the colour primaries are spread further apart compared with conventional video making it possible to represent more saturated colours such as deep-blue, magenta and neon green. It should be noted that deployment of WCG without HDR is, and is expected to be, quite limited. Similarly, there is no or very little interest in deploying HDR without WCG. HDR and WCG go naturally together and both of these dimensions are targeting improved reconstruction of each pixel, providing more lifelike video. In the remainder of this paper the term HDR will be used with an inherent assumption that WCG is also included. The effect of HDR and WCG is more or less completely independent of spatial resolution and may be used for example with 4K resolution as well as with 1080p resolution. BT.2020 includes the same colour space conversion as BT.709 called *non-constant luminance* but also a different one that is applied directly on the RGB signal (instead of on the R'G'B' signal) called *constant luminance*. In constant luminance the OETF is applied after the colour space conversion, but this approach has not become as widely used as non-constant luminance.

ITU-R Recommendation BT.2100

For HDR video the recently published ITU-R Recommendation BT.2100 "Image parameter values for high dynamic range television for use in production and international programme exchange" includes descriptions of two different systems for HDR: the Hybrid Log-Gamma (HLG) system and the perceptual quantization (PQ) system. The OETF of the HLG system in BT.2100 is technically identical to the OETF in ARIB STD-B67, and has been constructed to offer a degree of compatibility with legacy SDR displays that support WCG (i.e. BT.2020 colour primaries). It can be noted that HLG alone does not offer compatibility with displays only supporting BT.709 colour primaries. The EOTF of the PQ system in BT.2100 is technically identical to the EOTF in SMPTE ST.2084 and is constructed to have luminance levels distributed in alignment with the human visual system. SMPTE ST.2084 was published in 2013 and has become adopted and deployed by the industry and by different standardization bodies. Support for signalling ST.2084 EOTF is present in HDMI since version 1.4a. The work in MPEG and JCT-VC has been using the PQ EOTF as the basis for experiments and as a reference for investigations. The remainder of this paper will focus on the PQ EOTF but the methods described here are also applicable to other transfer functions such as the HLG. The colour primaries and the colour space conversion of BT.2100 are identical to the ones in BT.2020.

HDR10

The Blu-ray Disc Association (BDA) was one of the first groups to add support for HDR video. Several different technologies for HDR are included by the BDA but the only one which is mandatory is the one commonly called *HDR10* consisting of: the PQ EOTF in combination with BT.2020 colour primaries; 10-bit, non-constant luminance Y'CbCr; 4:2:0 chroma sub-sampling; and, when compressed, the HEVC Main10 Profile is used. In addition to the actual video, HDR10 may also include optional metadata such as Mastering Display Colour Information and Content Light Level Information.

THE WORK IN MPEG AND VCEG ON HDR

When MPEG first started to investigate compression of HDR video it was primarily in the form of exploratory experiments, breaking out individual pieces of the video compression and processing chain. Transfer functions were investigated in isolation, different colour representation formats where evaluated and various sub-sampling methods were tested. At the same time, methods for subjective video quality testing and metrics for objective quality measurements were explored. It became apparent that conventional error metrics such as Peak Signal-to-Noise Ratio (PSNR) did not provide much useful information when applied to high precision HDR video. Several quality metrics specifically created for HDR video were tested but as of today there is still no agreement around any metric being able to provide good a correlation between objective and subjective scores.

The MPEG CfE on HDR and WCG

In early 2015, MPEG decided to issue a Call for Evidence on HDR and WCG video (1). The purpose of the CfE was to identify if there existed indications of novel technology being able to provide improved compression of HDR video. In order to identify such technologies, a reference configuration was defined which all responses to the CfE would be compared against. This configuration later became called *Anchor 1.0* and consisted of pieces fully compatible with HDR10. In particular, Anchor 1.0 uses the PQ EOTF and non-constant luminance Y'CbCr compressed using HEVC Main10 Profile. Anchor 1.0 was created using an unmodified version of the HEVC reference software HM together with a software package called HDRTools to perform the necessary conversion steps. HDRTools was, and still is, used by MPEG and VCEG to implement the different methods under evaluation. The software package is free and publically available (2).

A total of 9 different responses were received from Apple, Arris, BBC, Dolby, Ericsson, FastVDO, InterDigital, MovieLabs, NGCodec, Philips, Qualcomm, Technicolor and University of Warwick/goHDR. All of the responses were evaluated in subjective tests conducted by two test labs comparing each response to Anchor 1.0. The results from the subjective tests (3) indicated that for several sequences a statistically significant difference was seen between some of the responses and Anchor 1.0. This result, in combination with input from the industry on timelines for deployment of HDR, led to initiation of a "fast-track" standardization process. The idea was to make additions to the HEVC specification, for example a new profile, which could improve coding of HDR video without modifying the low-level processing of HEVC Main10 Profile decoders. The work was conducted in a Core Experiment (CE2) in MPEG and an Exploratory Test Model (ETM) was developed introducing a "reshaper box" effectively approximating an adaptive transfer function to be applied directly on pictures output by the HEVC decoder.

Improvements of the Anchor

In parallel to the "fast-track" activity, a different Core Experiment was created (CE1) with the target of improving the performance of the Anchor by using technology already present in the HEVC specification. It had been identified that some of the sequences in the CfE, Anchor 1.0 demonstrated visual compression artefacts that could be alleviated by shifting bits from the luma component to the chroma components using an HEVC feature called chroma QP offset (4). Later input (5) also highlighted that in Anchor 1.0, too many bits were spent in the dark parts of the images and too few bits were spent on the bright parts. Modifications to address these two deficiencies together with a method called *luma adjustment* (described in more detail below) formed a new Anchor called Anchor 3.2, which clearly provided improved subjective quality relative to Anchor 1.0. It should be noted that the changes in Anchor 3.2 compared to Anchor 1.0 are all at or before the HEVC encoding and no changes are done at the decoder side, which means that Anchor 3.2 is also fully aligned with HDR10. An example of the improvements of Anchor 3.2 relative to Anchor 1.0 is shown in Figure 1 (tone mapped in order to be presented in SDR).

Figure 1 – Example of Anchor 1.0 (left) and Anchor 3.2 (right) compressed at equal bit-rate. Sequence courtesy of Technicolor and the NevEx project.

Conclusion of no new HEVC Profile

At the MPEG meeting in February 2016 the work of CE2 (the ETM) was compared with the work of CE1 (Anchor 3.2) and it was concluded that there was no technology identified that motivated the creation of a new HEVC profile or other normative specification. The focus would now instead be turned towards development of a technical report providing a reference for the methods for efficient conversion and coding of HDR using HEVC Main 10 Profile i.e. the methods included in Anchor 3.2. The work on creating such a technical report had started already in 2015 and the first official draft was approved in February 2016. It was decided that this would become a new part of MPEG-H which would be numbered Part 14.

MPEG-H PART 14

During 2015 it became apparent that early adopters of HDR video would support the PQ transfer function in combination with 4:2:0 Y’CbCr and compressed using HEVC Main10 Profile. The work in MPEG and VCEG had identified certain HDR-specific properties related to conversion and compression of HDR video and so it was decided to create a technical report documenting these properties and providing guidance on how to deal with HDR video. The first draft (6) was called “Conversion and coding practices for HDR/WCG video” but later it was decided to use a more specific title and so the most recent draft (7) is called “Conversion and Coding Practices for HDR/WCG Y′CbCr 4:2:0 Video with PQ Transfer Characteristics”.

The scope of the technical report is illustrated in Figure 2, which is borrowed from (7). The technical report includes descriptions of all conversions steps necessary for going from a linear light RGB representation with BT.2020 colour primaries to a 10-bit, narrow range, PQ, 4:2:0, non-constant luminance Y′CbCr representation. This includes the application of the EOTF, the colour space conversion, quantization and chroma sub-sampling. Corresponding processes are described for the conversion steps performed after video decoding in order to convert back to linear light RGB. These building blocks are relatively straightforward, but one method, the luma adjustment, is rather unconventional and will be presented in more detail below. The encoding part does not include a full description of video encoding; it merely highlights two specific methods that are useful for compression of HDR video: QP selection based on average luma and chroma QP offset based on content colour gamut. Both of these methods are briefly presented below.

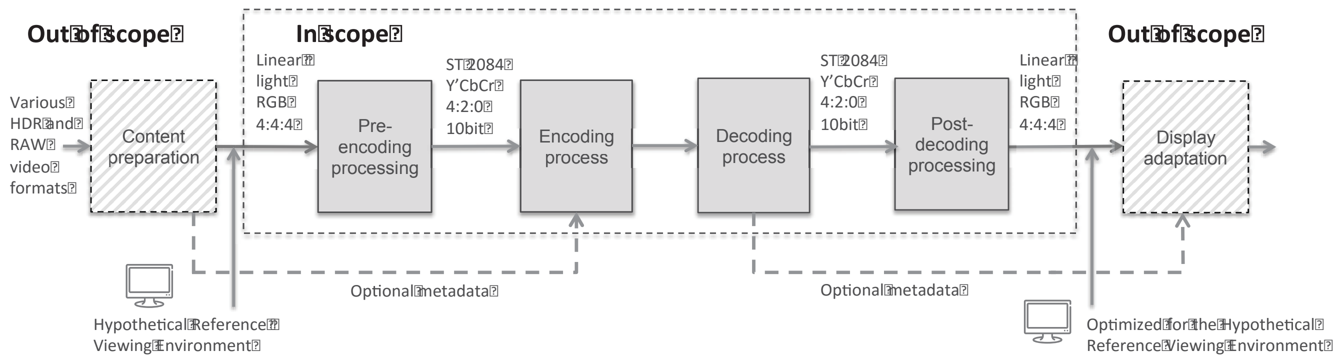

Figure 2 – Scope of MPEG-H Part 14

Luma adjustment

In the background section of this paper it is mentioned that in a Y’CbCr representation, some of the luminance information is carried in the chroma components, which can cause chroma leakage when the chroma components are subsampled. For HDR video this effect can be quite large due to the highly non-linear nature of the PQ EOTF and the fact that luma can differ significantly between neighbouring pixels. Figure 3 shows an example of chroma leakage and how the problem can be solved using the luma adjustment method. It should be noted that the example is processed in a form which makes it use all of the available code values for chroma i.e. graded for BT.709 display and represented in a BT.709 container. This is because no test material has been available in MPEG with full

BT.2020 colour gamut, but similar effects are expected for BT.2020 material in a BT.2020 container. The example in Figure 3 is tone mapped in order to be presented in SDR.

Figure 3 – Chroma leakage example. Original (left), Conventional 4:2:0 (middle) and 4:2:0 with luma adjustment (right). Sequence courtesy of Technicolor and the NevEx project.

In short, the luma adjustment method (8) modifies the chroma sub-sampling so that not only Cb and Cr is affected by the sub-sampling step; luma is adjusted to ensure that the correct luminance is retained when chroma is up-sampled. The scheme can be understood by reference to Figure 4 borrowed from the technical report (7).

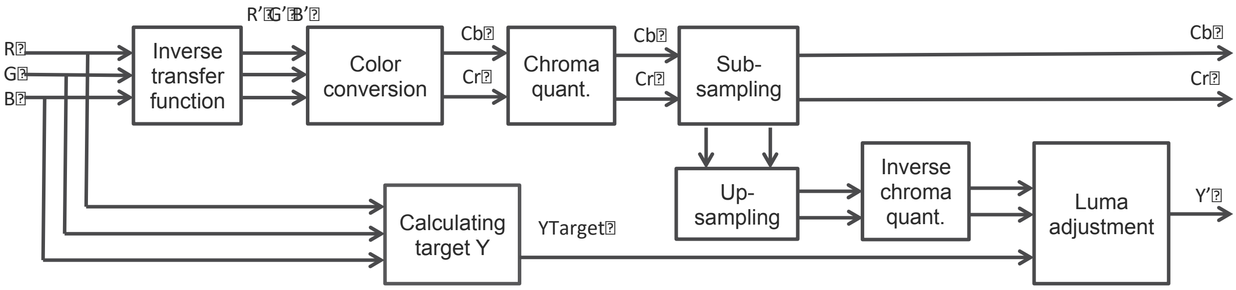

Figure 4 – Schematic overview of the luma adjustment method

QP selection based on average luma

In HEVC it is possible to set the QP for each Coding Unit individually. This is a feature that is fundamental in order to distribute the bits spent on a compressed image in alignment with the human visual system. It is very common to use a lower QP for a flat area (low variance) and a higher QP for a highly textured area (high variance). For HDR video, represented with the PQ EOTF, the subjective quality of the images can be improved by using a higher QP for blocks with low average luma value (dark area) and a lower QP for blocks with a high average luma value (bright area).

Chroma QP offset based on content colour gamut

For HDR video, the combination of the PQ EOTF and the BT.2020 colour primaries causes the Cb and Cr values to become more closely centred around zero compared with SDR video. When the HDR video is compressed, this results in the chroma values being more often lost during quantization, especially at high QP values. HEVC contains a feature called chroma QP offset, which controls the relation between the QP used for luma and the QP used for chroma. By setting the chroma QP offset individually for Cb and Cr it is possible to ensure that when high luma QP values are used, the chroma QP can still be at a level which does not cause the chroma information to be lost during quantization. If the content of the video does not exercise the full colour gamut and this is known during the encoding process (through analysis of the video or through metadata accompanying the video), then the chroma QP offset can be further optimized to provide an even better result.

VERIFICATION TEST OF HDR/WCG VIDEO CODING USING HEVC MAIN 10 PROFILE

When it had been decided that no new profile would be created in HEVC for HDR, and when it had furthermore been decided to create a technical report for coding and conversion of HDR, it was also decided to conduct a formal verification test of HDR compression using the HEVC Main10 Profile. The test would be using the same configuration as was used in the Anchors and in the technical report. Since no legacy method for compressing HDR video exists, it was agreed that the test would include two different variants of compressing HDR using the HEVC Main10 Profile, PQ, 4:2:0 and non-constant luminance Y'CbCr: one corresponding to Anchor 1.0 and one corresponding to Anchor 3.2. That way the test would not only be able to verify that HEVC is capable of efficiently compressing HDR video, but it would also be able to verify that the methods described in the technical report (corresponding to Anchor 3.2) give a subjective quality improvement over conventional methods (corresponding to Anchor 1.0).

The test was conducted at two test sites and concluded that HEVC is capable of effectively coding HDR video content (9). It was further concluded that for some sequences a benefit was demonstrated when using the methods described in the technical report. On average approximately 27% bit-rate reduction (calculated using MOS BD-Rate) could be achieved by using the methods from Anchor 3.2 compared to the methods in Anchor 1.0

REFERENCES

1. A. Luthra, E. Francois, W. Husak, "Call for Evidence (CfE) for HDR and WCG Video Coding", MPEG2014/N15083, 110th MPEG Meeting, Geneva, 2015.

2. A. Tourapis, HDRTools software package: https://gitlab.com/standards/HDRTools/

3. V. Baroncini, T. Ebrahimi, "Test Results of Call for Evidence (CfE) for HDR and WCG Video Coding", N15350, June 2015.

4. J. Samuelsson, M. Pettersson, J. Ström, K. Andersson, Not Public: "Using chroma QP offset on HDR sequences", MPEG m36581, June 2015

5. J. Zhao, S.-H. Kim, A. Segall, K. Misra, Not Public: "Performance investigation of high dynamic range and wide color gamut video coding techniques", MPEG m37439, Oct 2015.

6. J. Samuelsson, "Conversion and coding practices for high-dynamic-range and wide-colour-gamut video", Draft 1, MPEG N16063, February 2016.

7. J. Samuelsson, C. Fogg, A. Norkin, A. Segall, J. Ström, G. Sullivan, A. Tourapis, "Conversion and Coding Practices for HDR/WCG Y′CbCr 4:2:0 Video with PQ Transfer Characteristics", Draft 2, JCTVC-X1017, June 2016.

8. J. Strom, J. Samuelsson and K. Dovstam, "Luma Adjustment for High Dynamic Range Video", Proceedings of the IEEE Data Compression Conference (DCC), March 2016.

9. V. Baroncini, K. Andersson, A. K. Ramasubramonian, G. J. Sullivan, "Verification Test Report for HDR/WCG Video Coding Using HEVC Main 10 Profile", N16273, June 2016