HDR video coding with MPEG-5 LCEVC

Amaya Jiménez-Moreno
V-Nova Limited
London, UK
amaya.moreno@v-nova.com

Lorenzo Ciccarelli
V-Nova Limited
London, UK
lorenzo.ciccarelli@v-nova.com

Rick Clucas
V-Nova Limited
London, UK
rick.clucas@v-nova.com

Simone Ferrara
V-Nova Limited
London, UK
simone.ferrara@v-nova.com

ABSTRACT

High Dynamic Range (HDR) video content is continuing to gain market relevance for both streaming and brodacasting services, providing video with improved contrast and colour depth. However, the predominance of 8-bit based codecs and the wide availability of Standard Dynamic Range (SDR) devices still poses challenges regarding effective deployment of HDR content.

MPEG-5 LCEVC is a new video coding standard that works in combination with a separate video standard (e.g., AVC, HEVC, VVC, AV1) to enhance the quality of a video. The enhanced quality is provided by adding details coded through an enhancement layer to a lower resolution version of the same video coded through a base layer. These enhancement layers can be used to add an HDR enhancement layer to any underlying codec, even to 8-bit based codecs, which helps to achieve a more efficient encoding and solves backward-compatibility issues.

In this paper, we describe how LCEVC enables the encoding of HDR video, explaining some of the main tools to provide higher efficiency for this content. Moreover, we provide a series of test results and comparisons of encoding HDR using LCEVC to enhance different video codecs.

CCS CONCEPTS

• **Applied computing** → **Engineering**; **Telecommunications**.

KEYWORDS

enhancement layer, HDR, MPEG-5 LCEVC, video coding, 10-bit video coding, 12-bit video coding

ACM Reference Format:
Amaya Jiménez-Moreno, Lorenzo Ciccarelli, Rick Clucas, and Simone Ferrara. 2022. HDR video coding with MPEG-5 LCEVC. In *Mile-High Video Conference (MHV '22), March 1–3, 2022, Denver, CO, USA.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3510450.3517307

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
MHV '22, March 1–3, 2022, Denver, CO, USA
© 2022 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-9222-8/22/03.
https://doi.org/10.1145/3510450.3517307

1 INTRODUCTION

High Dynamic Range (HDR) technology uses higher bit-depth to represent luminance and chrominance planes in video content, offering the customers a high quality image with levels of brightness (and darkness) and a range of colors closer to those perceived by the human visual system. However, the increase in bit-depth requires a very large storage space as well as higher transmission bandwidths than a corresponding Standard Dynamic Range (SDR) video. For this reason, the development of efficient HDR coding systems is extremely important to enable effective deployment of HDR technology. Moreover, the majority of existing digital display devices can still only support SDR (8-bit) video content, so it is also necessary to provide backward-compatibility.

We can find some approaches in the state-of-the-art for managing HDR content and facing the backward-compatibility issue, e.g., [4], [6], and [5]. [4] was focused on the problem of finding an optimal tone-curve mapping algorithm for a backward-compatible encoding scheme using an 8-bit SDR signal encoded with H.264/AVC. In [6] and [5] the authors proposed backward-compatible HDR video compression algorithms based on MPEG video coding standards. The first one made an adaptation of MPEG-4 to HDR data. In the second one, a tone-mapped SDR version of the input HDR frames was compressed using a MPEG-2 encoder to produce a backward-compatible SDR stream.

The characteristics of HDR data are different compared to those of SDR data. Thus, if an encoder primarily devised for SDR treats an error in an HDR sequence the same way as it does for an SDR sequence, such encoder could incur in errors which are very large in some areas of the sequence [8]. This may result in more compression artifacts for HDR sequences [9].

For this reason, several methods have been proposed to improve the quality of the HDR encoded sequences, all of them related to some kind of Quantization Parameter (QP) adjustment for luminance and chrominance, e.g., the proposals done by Zhang et al. in [10] and [11], or [3]. This QP adjustment has also been included in the recommended guidelines for the operation of the AVC or HEVC video coding system for compressing HDR video [8].

MPEG-5 LCEVC is a new video coding standard that enhances any other codec (e.g. AVC, VP9, HEVC, AV1, EVC, or VVC) to produce a higher compression efficiency than that achievable by the enhanced codec used alone, as demonstrated during the MPEG verification tests report [2]. It uses any conventional video codec as a base codec at a lower resolution and adds details to the low-resolution version of the video, through an enhancement layer. In

doing so, it also reduces the overall encoding complexity by operating the enhanced codec at a fraction of the encoding complexity required for encoding the video at full resolution, with LCEVC adding a very small encoding complexity on top.

The LCEVC format also provides the ability to signal the bit-depth of the base layer independently from that of the enhancement layer. This is particularly relevant for encoding HDR content. In addition to transmitting the base layer and enhancement layer at the same bit-depth (e.g., 10-bit), the format allows transmission of a lower-bit depth base layer (e.g., 8-bit) and a higher bit-depth enhancement (e.g., 10-bit). This latter feature allows further quality enhancement capabilities, for example, transmission of an SDR base layer with an HDR enhancement layer or, in the future, a 10-bit depth HDR base layer with a 12-bit depth HDR enhancement layer.

In this paper we describe and analyze the way HDR content can be managed with MPEG-5 LCEVC. There are two modes of operation: full HDR and extended HDR. In the full HDR mode, the bit-depth of the base layer and that of the enhancement layer are the same - today this is typically specified as 10-bit to reflect the commercial HDR formats available today, but the bit-depth can extend to up to 14 bit. In the extended HDR mode, the bit-depth of the base layer is lower than that of the enhancement layer, meaning for example that we can encode an HDR source using for the base layer an AVC/H.264 or HEVC/H.265 8-bit encoding and then adding on top an LCEVC 10-bit enhancement layer to bring the full-resolution video to HDR quality, allowing HDR support to devices and workflows which are not designed to deliver HDR quality video. In this way, both devices which are HDR capable and devices that are non-HDR capable can receive the same bitstream and display HDR quality.

In particular, the following analysis shows how the enhancement layer is in charge of adding the HDR details - particularly for the extended HDR mode - improving the quality of the obtained encoded sequence and avoiding the appearance of artifacts. The use of lower resolution sequences in the base layer significantly reduces the generated bitrate to store or transmit the content, which can be very helpful when managing HDR content. Note that as LCEVC is agnostic of the base codec, there are no modifications required for the base codec in order to deliver HDR content.

In this paper, we describe how LCEVC enables the encoding of HDR content in terms of format, syntax, and tools. We further provide a series of preliminary test results and comparisons of encoding HDR content using LCEVC to enhance different video codecs, using different bit-depth configurations for base layer and enhancement layer. Finally, we discuss the results and provide examples of real-life implementations where LCEVC can provide benefits for HDR content.

2 MPEG-5 LCEVC DESCRIPTION

In Figure 1 we can see a block diagram of the MPEG-5 LCEVC encoder [7]. Depending on the chosen configuration, the downscaling can be applied up to two consecutive times. The encoding process can be divided into three main steps. Firstly, the input sequence is downscaled. The video, at a lower resolution, is fed into the base encoder that can be any video coding standard (e.g. AVC, VP9, HEVC, AV1, EVC, or VVC) producing a corresponding decodable bitstream.

The video reconstructed from the encoded base bitstream is used as the input for the enhancement layers, following an optional upscaling process if the downscaling process was applied twice. At the enhancement layer 1 (L-1), the residuals are created by subtracting the downscaled input sequence and the base video reconstruction. These residuals are transformed, quantized, and encoded. The transform used in LCEVC has a small Hadamard-type kernel of size 2×2 or 4×4. A uniform quantizer with variable deadzone is used on the transform coefficients. The entropy encoder, which consists of a run-length encoder (RLE) and an optional prefix encoder (Huffman encoder), receives the quantized transform coefficients and generates the encoded bitstream. The inverse tools of the quantization and transform are applied, creating the L-1 reconstruction. Additionally, an L-1 filter can be added which functions as a simple deblocking filter.

Finally, the L-1 reconstruction is upscaled to full resolution and subtracted from the original input sequence. The resulting residual is presented to the enhancement layer 2 (L-2) and fed into a zero-motion vector temporal prediction module. If the temporal module decides to use the temporal prediction, the residual will be the difference between the input residual to L-2 and the predicted one. The residuals are transformed, quantized, and encoded using the same tools used for L-1.

2.1 MPEG-5 LCEVC features for HDR content

LCEVC already has some specific tools for managing HDR content and gives the possibility of adding new ones to better fit the HDR features.

2.1.1 Bitdepths up to 14-bit. Currently, LCEVC allows up to 14 bits in its main profile data representations, which covers the main HDR formats, i.e. HDR10, HLG, a big part of HDR10+ and Dolby Vision. So, LCEVC can cover the encoding necessities in the great majority of HDR architectures at this time. Moreover, the great flexibility of the LCEVC encoder allows us the combination of different bit-depths for base layer and enhancement layers. For example, in a full HDR mode a bitdepth of 10 or 12 bits can be used for both base and enhancement layers. In extended HDR mode, an 8-bit base layer with 10-bit enhancement layer or a 10-bit base layer with 12-bit enhancement layer can be used.

2.1.2 Dithering. One of the drawbacks when using a reduced bit-depth for the base encoder is the appearance of visible artifacts that can degrade the visual quality of the output picture. One of the most common artifacts is the banding that appears when a sequence contains a big area with a soft color transition, as can be seen in Figure 2a, this artifact also can be easily obtained when using a hard quantization. To solve this, LCEVC allows the use of a post-processing dithering module in the decoding process. The MPEG-5 LCEVC standard is able to indicate the strength of the dithering to be applied, but the type is not mandated. In the present work, we used the V-Nova SDK implementation of LCEVC [1], and in this implementation the dithering module adds random noise to the luminance and chrominance decoded planes. This helps to visibly reduce the banding artifact, as can be seen in Figure 2b.

However, although the perceived quality is improved using the post-processing dithering module, it is not possible to completely

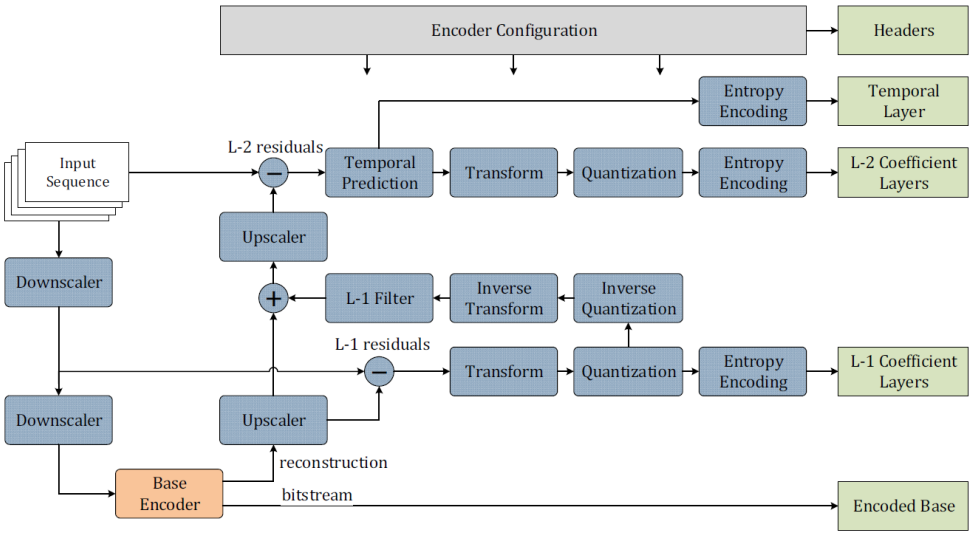


Figure 1: Structure of an LCEVC encoder

remove the banding when the artifact is very strong. Thus, to reduce the impact of reducing the bit-depth for the base encoder further, we propose the use of a new dithering module on the encoder side, before the input to the base encoder (pre-processing dithering module). First, the input sequence is downsampled and, during the conversion from 10 to 8-bit, a new dithering technique is applied. This new dithering module consists of adding 1 to bit 0 of the 8-bit resulting word, when bit 1 of the 10-bit source word is 1, depending on the x and y position, so a checkerboard pattern is followed. The effect on the input sequence to the base encoder can be seen in Figure 2c. The strength of this new dithering module could also be controlled, e.g., adding a value greater than 1 to the resulting 8-bit word if the base QP is going to be high, keeping the dithering pattern in the base decoded sequence to reduce the appearance of the artifacts. It must be noted that this pre-processing dithering module is not part of the MPEG-5 LCEVC standard and it is proposed here as an improvement.

Combining the two dithering methods, the perceived quality can be enhanced greatly, removing the majority of the artifacts resulting from the normal encoding process and, more specifically, from the bit-depth reduction.

2.1.3 Separated quantization for the chroma. We have previously seen how the QP adjustment is a common way to improve the quality of the HDR encoded sequences, reducing the artifacts generated when the HDR content is managed by a video encoder devised to deal with SDR sequences. Moreover, the QP adjustment is also recommended in the guidelines for the operation of the AVC or HEVC video coding systems for compressing HDR video [8].

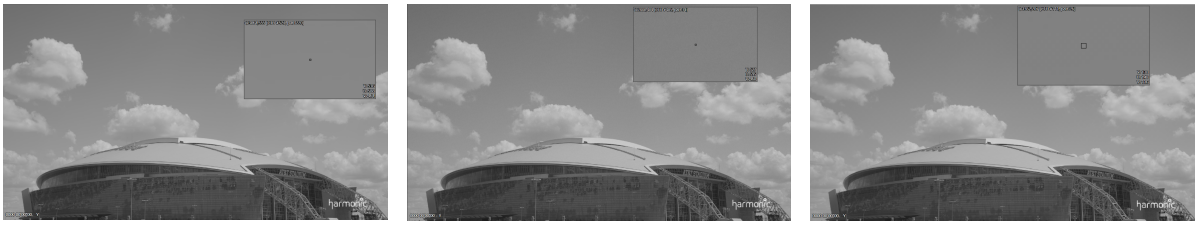
LCEVC allows for the use of different QP values for the luminance and chrominance planes and the different transformed coefficients. This flexibility when selecting the QP, can be used to make a proper adjustment of the QP values for the HDR content. In the current version of the LCEVC reference software, the QPs are adjusted for every type of coefficient after the transformation. Also, LCEVC signals a modifier for the QP of the chrominance planes. There are 8 bits available to specify a value used to scale the QP of L-2 for the chrominance.

2.1.4 User data as additional information. LCEVC gives the possibility of including additional information in the regular bitstream. To do so, at transform level (so, for every 2×2 or 4×4 data block, depending on the encoder configuration) we can use up to 6 bits of the first transform coefficient to signal extra information. These bits will be extracted in the decoder side and used to adjust the decoding process as the user had specified. This information can be used to make an adaptive adjustment of the QP values for luma or chrominance, depending on the brightness or color values, which allows extra flexibility to avoid quantization artifacts for HDR content without adding extra bits in the bitstream.

This signalling is part of the LCEVC standard, but the specific use is not normative. So, a LCEVC compliant decoder must, at least, be able to read and store this extra information. The specific use of it, will depend on the decoder implementation.

2.1.5 Benefits of using MPEG-5 LCEVC with HDR content. We have seen in the previous sections the possibilities that LCEVC has to manage HDR content. It is able to manage different scenarios, encoding HDR content up to 14 bits and allowing backward-compatibility frameworks using lower bit-depths for the base encoder. The encoding artifacts can be reduced using the new pre-processing and post-processing dithering modules and adjusting the QP values for the different transformed coefficients and between luminance and chrominance planes, so the encoder is adapted to the specific features and range of values of the HDR content.

Moreover, LCEVC is agnostic to PQ or HLG transfer functions as the input format when a backward-compatible stream is desired. The input sequence can be converted from the original format to



(a) Sequence converted to 8-bit presenting banding artifacts **(b) Reconstructed sequence after using the post-processing dithering module** **(c) Input sequence to the base encoder after using pre-processing dithering module**

Figure 2: Dithering modules analysis with the sequence FB_AmStadiumOutside_2160p60Hz_v210

the lower bit-depth format and then use this lower bit-depth sequence as input to the base encoder. Then, the enhancement layers will work with the original bit-depth converting back the decoded sequence of the base to the original format. In the implementation presented in this paper, we have simply implemented a bit-shifting operation. This is not sufficient to preserve a good quality for the 8-bit version of the sequence, especially for PQ signals. A current line of work is dedicated to improve this implementation by adding a tone-mapping algorithm to perform the bit-depth conversions.

3 EXPERIMENTAL EVALUATION

3.1 Rate-Distortion performance evaluation

We evaluated the performance of MPEG-5 LCEVC with HDR content using several encoding configurations and various types of sequences. The V-Nova LCEVC SDK implementation [1] was used for these experiments. The LCEVC encoder was configured without tone-mapping (conversion from 10-bit to 8-bit is done by simply removing the last two bits of the pixel representation) or scaling of the chrominance QP. As discussed above, pre-processing and post-processing dithering modules were used. In the experiments, only one downscaling and upscaling operation was used. A comprehensive set of HD and UHD sequences was used, covering different temporal and spatial complexities. Specifically, we tested 21 PQ sequences (6 HD and 15 UHD) and 18 HLG sequences (3 HD and 15 UHD), all in 420 progressive and 10-bit format (in Table 1 the complete list of sequences is presented). We are going to test two configurations, the first one using the same preset for all the tested encoders, and the second one to compare the encoders with a more similar computational complexity [1].

In Table 2 we report the results of a comparison between (i) a 10-bit LCEVC encode enhancing an 8-bit x264 base layer and (ii) a native 10-bit x265. We computed the Rate-Distortion (RD) curves using the following target bitrates: 500kb, 1Mb, 2Mb, 3Mb, and 5Mb for HD sequences, and 2Mb, 5Mb, 8Mb, 12Mb, and 16Mb for UHD. We then measured the Bjontegaard Delta (BD) bitrate and quality changes when using LCEVC in comparison to x265. We considered VMAF, PSNR, and MS-SSIM metrics, these two last calculated using the HDR Tools. For the comparisons, we considered two sets of configurations. A first set includes native x265 with preset *veryslow* and LCEVC enhancing x264 base encoder with preset *veryslow* - note that with these configurations the encoding time for LCEVC is significantly lower than that of the native configuration. A second set includes native x265 *medium* and LCEVC enhancing x264 base encoder with preset *veryslow* - with these configurations the encoding times for native and LCEVC are more comparable.

[1]The computational complexities were evaluated with an Intel(R) Core(TM) i9-9900K CPU at 3.60GHz

Table 1: List of sequences used in the performance evaluation

Transfer function	Resolution	Sequence name (fps, number of frames)
PQ	HD	BalloonFest (24fps, 240), ShowGirl2Clip (25fps, 339), Starting (50fps, 500), Hurdles (50fps, 500), A1_Tango_bl (60fps, 600), HDR_meridian (60fps, 600)
PQ	UHD	A1_Tango (60fps, 294), A1_Tango_bl (60fps, 600), A1_Tango_or (60fps, 600), A2_DaylightRd (60fps, 300), A2_RollerCstr (60fps, 300), CampfireParty (30fps, 300), HDR_meridian1 60fps, 382), HDR_meridian2 (60fps, 392), HDR_meridian3 (60fps, 482), HDR_meridian4 (60fps, 378), HDR_meridian5 (60fps, 425), HDR_meridian6 (60fps, 327), HDR_meridian7 (60fps, 493), HDR_meridian (60fps, 600), HDR_nocturne (60fps, 600)
HLG	HD	DayStreet_bl (60fps, 600), PeopleShopCn (60fps, 600), SunsetBeach (60fps, 601)
HLG	UHD	H3_AMS01 (60fps, 600), DayStreet_bl (60fps, 600), FlyingBirds2 (60fps, 300), LG_BeachPan (60fps, 154), LG_BeachContrast (60fps, 123), DayStreet (60fps, 600), FlyingBirds (60fps, 600), DayStreet_or (60fps, 600), PeopleShopCn (60fps, 600), H3_AMS03 (60fps, 600), SunsetBeach (60fps, 601), SunBeach_LgoOr (60fps, 601), H3_AMS09 (60fps, 600), LG_BeachSand (60fps, 182), SunBeach_LgoBl (60fps, 601)

Table 2 reports the average results. Comparing with native 10-bit x265 *veryslow*, the BD-Rate(VMAF) increments around 10% for UHD sequences and between 8 and 24% for HD. However, when comparing with native 10-bit x265 *medium*, the results show an average reduction in BD-Rate(VMAF) between 6% and 10% for UHD sequences and up to 14% for HD sequences. In terms of MS-SSIM, we measure increments in the BD-Rate, except when comparing with native x265 *medium* for HD HLG sequences. When looking at PSNR, the BD-Rate results are less favourable to LCEVC - this is expected as PSNR has been shown to penalise LCEVC when compared to native implementations, and with the use of dithering this penalisation is further accentuated.

In Table 3 we report the results of a comparison between (i) a 10-bit LCEVC encode enhancing a 10-bit x265 base layer and (ii) a native 10-bit x265. For the comparisons, we considered two sets of configurations. A first set includes native x265 with preset *veryslow* and LCEVC enhancing x265 base encoder with preset *veryslow* - note that with these configurations the encoding time for LCEVC is significantly lower than that of the native configuration. A second

Table 2: Average BD-VMAF, BD-Rate(VMAF), BD-PSNR, BD-Rate(PSNR), BD-MS-SSIM, and BD-Rate(MS-SSIM) average results for LCEVC(10-bit) + x264(8-bit) in comparison to x265(10-bit)

	x265(10-bit) (anchor) vs LCEVC(10-bit)+x264(8-bit) (target)							
	Conf 1) Preset native x265 *veryslow*, LCEVC x264 *veryslow*				Conf 2) Preset native x265 *medium*, LCEVC x264 *veryslow*			
	PQ		HLG		PQ		HLG	
	BD-VMAF	BD-Rate(VMAF)(%)	BD-VMAF	BD-Rate(VMAF)(%)	BD-VMAF	BD-Rate(VMAF)(%)	BD-VMAF	BD-Rate(VMAF)(%)
UHD	-0.67	9.69%	-1.46	12.16%	0.53	-6.39%	0.99	-10.85%
HD	-2.39	24.66%	-1.30	8.06%	0.10	0.49%	2.76	-14.70%
	BD-PSNR	BD-Rate(PSNR)(%)	BD-PSNR	BD-Rate(PSNR)(%)	BD-PSNR	BD-Rate(PSNR)(%)	BD-PSNR	BD-Rate(PSNR)(%)
UHD	-0.35	15.26%	-1.63	73.85%	-0.50	50.96%	-1.25	49.85%
HD	-2.52	137.30%	-2.11	78.07%	-2.08	81.57%	-1.62	48.90%
	BD-MS-SSIM	BD-Rate(MS-SSIM)(%)	BD-MS-SSIM	BD-Rate(MS-SSIM)(%)	BD-MS-SSIM	BD-Rate(MS-SSIM)(%)	BD-MS-SSIM	BD-Rate(MS-SSIM)(%)
UHD	0	1.69%	0	21.10%	0	18.76%	0	7.25%
HD	0	43.94%	0	8.52%	0	23.52%	0	-5.86%

Table 3: Average BD-VMAF, BD-Rate(VMAF), BD-PSNR, BD-Rate(PSNR), BD-MS-SSIM, and BD-Rate(MS-SSIM) average results for LCEVC(10-bit)+x265(10-bit) in comparison to x265(10-bit)

	x265(10-bit) (anchor) vs LCEVC(10-bit)+x265(10-bit) (target)							
	Conf 1) Preset native x265 *veryslow*, LCEVC x265 *veryslow*				Conf 2) Preset native x265 *medium*, LCEVC x265 *slow*			
	PQ		HLG		PQ		HLG	
	BD-VMAF	BD-Rate(VMAF)(%)	BD-VMAF	BD-Rate(VMAF)(%)	BD-VMAF	BD-Rate(VMAF)(%)	BD-VMAF	BD-Rate(VMAF)(%)
UHD	2.87	-48.49%	2.73	-28.82%	3.40	-52.10%	4.27	-37.25%
HD	2.73	-22.31%	3.46	-21.30%	3.99	-28.49%	5.94	-30.08%
	BD-PSNR	BD-Rate(PSNR)(%)	BD-PSNR	BD-Rate(PSNR)(%)	BD-PSNR	BD-Rate(PSNR)(%)	BD-PSNR	BD-Rate(PSNR)(%)
UHD	0.38	-33.98%	-0.87	34.75%	0.29	-20.64%	-0.44	15.89%
HD	-1.80	87.19%	-1.68	44.56%	-1.53	49.15%	-1.32	29.64%
	BD-MS-SSIM	BD-Rate(MS-SSIM)(%)	BD-MS-SSIM	BD-Rate(MS-SSIM)(%)	BD-MS-SSIM	BD-Rate(MS-SSIM)(%)	BD-MS-SSIM	BD-Rate(MS-SSIM)(%)
UHD	0	-42.17%	0	-3.57%	0	-35.36%	0	-15.90%
HD	0	5.98%	0	-5.4%	0	-2.50%	0	-12.70%

set includes native x265 *medium* and LCEVC enhancing x265 base encoder with preset *slow* - with these configurations the encoding times for native and LCEVC are more comparable. Table 3 reports the average results which have been computed using the same target bitrates and metrics reported for the previous experiment. These results show that LCEVC enhancing a 10-bit x265 base is able to obtain an average bitrate reduction in terms of VMAF up to 48% and 52% for UHD sequences and up to 22% and 30% for HD, for the first and second set of configurations, respectively. When looking at MS-SSIM, in the first set of configurations LCEVC reduces the BD-Rate up to 42% and 5% for UHD and HD sequences, respectively, whereas with the second set of configurations the BD-Rate reduction is up to 35% and 12% for UHD and HD. When looking at PSNR, the BD-Rate results are less favourable to LCEVC - again this is expected for the same reasons explained above.

Finally, we also performed a comparison between a 10-bit LCEVC encode enhancing an 8-bit x265 base layer and a native 10-bit x265 one. We used the same two configurations, target bitrates, and metrics reported for the tests reported in Table 3. The results showed that LCEVC enhancing an 8-bit x265 base compared with 10-bit x265 obtained, in terms of VMAF, an average bitrate reduction up to 48% and 50% for UHD sequences and up to 24% and 36% for HD sequences, for the first and second set of configurations, respectively. In this experiment, we also observed BD-Rate reductions when looking at MS-SSIM for UHD sequences and HD HLG sequences.

3.2 Subjective analysis

In this section, we present a subjective analysis performed by visually inspecting two of the tested PQ sequences. We collected the reconstructed yuv files of the two sequences with the different coding options evaluated with the presets set to *veryslow* for both encoders. Then, we configured an HDR monitor Asus ProArt PA27UCX to work with PQ transfer function and used the Vooya player to reproduce the yuv files.

In Figure 3 we show some screenshots extracted from the sequence HDR_meridian1_2160p60_10b (UHD, PQ) at 12Mb. This is a sequence that gave a good performance with LCEVC. The screenshots reproduce a zoomed area of reconstructed frame 112. As we can see, the reconstruction using LCEVC 10-bit + x264 8-bit is already giving a comparable visual quality when compared with native 10-bit x265. The use of a 10-bit x265 further improve the visual quality. Specifically, the details around the eye are always clearer using LCEVC and x265 base. It must be noted that looking at the whole frame and sequence, the subjective quality obtained with LCEVC was comparable or better to that of native x265 (consistent with the results of the VMAF metrics).

In Figure 4 we show some screenshots extracted from the sequence A1_Tango_2160p60_10b_PQ at 5Mb. For this sequence, all metrics showed an improvement using LCEVC, except for VMAF when enhancing an x264 base. The screenshots in Figure 4 reproduce a zoomed area of reconstructed frame 124. In this example, the most unpleasant artifacts that can be seen with the native x265 reconstruction (lines in the middle of the necklace), disappear when LCEVC is used. In general, LCEVC with x265 10-bit improves the

(a) x265 10-bit **(b) LCEVC 10-bit + x264 8-bit** **(c) LCEVC 10-bit + x265 10-bit**

Figure 3: Zoom on a particular region of reconstructed frame 112 of HDR_meridian1_2160p60_10b 12 Mb (UHD, PQ)

(a) x265 10-bit **(b) LCEVC 10-bit + x264 8-bit** **(c) LCEVC 10-bit + x265 10-bit**

Figure 4: Zoom on a particular region of reconstructed frame 124 of A1_Tango_2160p60_10b_PQ 5Mb (UHD, PQ)

perceived quality, especially in the shadow in the right top corner of the image. In this example, looking at the whole frame and sequence, the sharpness in the details obtained with LCEVC improve the perceived quality.

It must be noted that we did not observe any perceivable difference in terms of brightness and color between the native HDR encoders (using 10-bit) and LCEVC enhancing an 8-bit base layer. This implies that LCEVC is able to produce an HDR quality reconstructed sequence starting from an SDR version used for the base encoder and adding the 10-bit LCEVC enhancement layer.

4 CONCLUSIONS

In this paper, we have discussed how MPEG-5 LCEVC can be used to manage HDR content. With MPEG-5 LCEVC, we can either encode directly at the HDR native bit-depth (e.g., 10-bit) or we can start from a lower-bit depth for the base layer (e.g., 8-bit) and achieve a higher bit-depth for the enhancement layer (e.g., 10-bit), thus extending the capabilities of existing 8-bit based systems. Moreover, we have described several tools available in LCEVC that can help the processing of HDR content, analyzing all their potential benefits.

We have evaluated the performance of LCEVC enhancing either an 8-bit base format or a 10-bit base format and comparing it with 10-bit native x265. The tests show that LCEVC is capable of producing an HDR quality video starting from an 8-bit x264 base layer, thus providing HDR video also in those scenarios where only x264 can be supported in end-user devices (e.g., streaming applications). Further, when enhancing 10-bit base layers (e.g, x265), LCEVC produces significant improvement when compared with the equivalent native implementation of the base layer. A preliminary subjective evaluation was carried out to validate those results. Moreover, these promising results could be additionally improved with the use of the other described HDR features. Future work is planned to explore how additional tools (such as tone-mapping) can further improve the performance.

REFERENCES

[1] [n. d.]. MPEG-5 LCEVC Low Complexity Enhancement Video Coding. http://www.lcevc.org/. Accessed: 2021-12-09.

[2] Apr. 2021. Verification Test Report on the Compression Performance of Low Complexity Enhancement Video Coding. ISO/IEC JTC1/SC29/WG04 N0076. Online.

[3] Jong-Hyeok Lee, Young-Woon Lee, Dongsan Jun, and Byung-Gyu Kim. 2020. Efficient Color Artifact Removal Algorithm Based on High-Efficiency Video Coding (HEVC) for High-Dynamic Range Video Sequences. *IEEE Access* 8 (2020), 64099–64111. https://doi.org/10.1109/ACCESS.2020.2984012

[4] Zicong Mai, Hassan Mansour, Rafal Mantiuk, Panos Nasiopoulos, Rabab Ward, and Wolfgang Heidrich. 2011. Optimizing a Tone Curve for Backward-Compatible High Dynamic Range Image and Video Compression. *IEEE Transactions on Image Processing* 20, 6 (2011), 1558–1571. https://doi.org/10.1109/TIP.2010.2095866

[5] Rafal Mantiuk, Alexander Efremov, Karol Myszkowski, and Hans-Peter Seidel. 2006. Backward Compatible High Dynamic Range MPEG Video Compression. In *ACM SIGGRAPH 2006 Papers* (Boston, Massachusetts) *(SIGGRAPH '06)*. Association for Computing Machinery, New York, NY, USA, 713–723. https://doi.org/10.1145/1179352.1141946

[6] Rafal Mantiuk, Grzegorz Krawczyk, Karol Myszkowski, and Hans-Peter Seidel. 2004. Perception-Motivated High Dynamic Range Video Encoding. *ACM Trans. Graph.* 23, 3 (Aug 2004), 733–741. https://doi.org/10.1145/1015706.1015794

[7] Guido Meardi, Simone Ferrara, Lorenzo Ciccarelli, Guendalina Cobianchi, Stergios Poularakis, Florian Maurer, Stefano Battista, and Ahmad Byagowi. 2020. MPEG-5 part 2: Low Complexity Enhancement Video Coding (LCEVC): Overview and performance evaluation. In *Applications of Digital Image Processing XLIII*, Andrew G. Tescher and Touradj Ebrahimi (Eds.), Vol. 11510. International Society for Optics and Photonics, SPIE, 238 – 257. https://doi.org/10.1117/12.2569246

[8] J. Samuelsson, C. Fogg, A. Norkin, A. Segall, J. Straom, G. Sullivan, P. Topiwala, and A. Tourapis. 2017. Conversion and Coding Practices for HDR/WCG YCbCr 4:2:0 Video with PQ Transfer Characteristics (Draft 4). In *JCTVC-Z1017, 26th Meeting, Geneva, Switzerland*. 1–4. https://doi.org/standards/mpeg-h/conversion-and-coding-practices-high-dynamic-range-and-wide-colour-gamut-video-1

[9] Yang Zhang, Dimitris Agrafiotis, Matteo Naccari, Marta Mrak, and David R. Bull. 2013. Visual masking phenomena with high dynamic range content. In *2013 IEEE International Conference on Image Processing*. 2284–2288. https://doi.org/10.1109/ICIP.2013.6738471

[10] Yang Zhang, Matteo Naccari, Dimitris Agrafiotis, Marta Mrak, and David R. Bull. 2013. High dynamic range video compression by intensity dependent spatial quantization in HEVC. In *2013 Picture Coding Symposium (PCS)*. 353–356. https://doi.org/10.1109/PCS.2013.6737756

[11] Yang Zhang, Matteo Naccari, Marta Mrak, Dimitris Agrafiotis, and David Bull. 2016. High dynamic range video compression exploiting luminance masking. *IEEE Transactions on Circuits and Systems for Video Technology* 26, 5 (1 May 2016), 950–964. https://doi.org/10.1109/TCSVT.2015.2426552