



MPEG-5 LCEVC for 3.0 Next Generation Digital TV in Brazil

Lorenzo Ciccarelli*, Simone Ferrara and Florian Maurer

V-Nova Limited, R&D, London, United Kingdom

TV 3.0 is the next generation digital broadcasting system developed in Brazil by the SBTVD Forum. The ambition of TV 3.0 is significantly higher than that of previous generations as it targets the delivery of IP based signals for applications, such as 8K, HDR, virtual and augmented reality, video enhancement and scalability. To deliver such services, more advanced and flexible compression technologies are required. MPEG-5 Part 2 Low Complexity Enhancement Video Coding (LCEVC) is a new video coding standard which works in combination with a separate video standard (e.g., H.264/AVC [H.264/AVC], H.265/HEVC [H.265/HEVC], H.266/VVC [H.266/VVC], AV1 [AV1]) to enhance the quality of a video. In the typical scenario, the enhanced quality is provided in terms of a higher resolution video obtained by adding details coded through an enhancement layer to a lower resolution version of the same video coded through a base layer. The LCEVC format also provides the ability to signal the bit-depth of the base layer independently from that of the enhancement layer and allowing up to 14-bit depth HDR. MPEG-5 LCEVC has been selected by the SBTVD committee as part of the TV 3.0 in December 2021. In this paper we describe the proposal submitted for LCEVC in response to the SBTVD Call for Proposals (CfP) for TV 3.0.

Keywords: video coding HDR, MPEG-5 LCEVC, 10-bit video coding, enhancement layer, DTT, OTT, HFR, DRE

OPEN ACCESS

Edited by:

Matteo Naccari,
Audinate, United Kingdom

Reviewed by:

Xiem Hoang,
Vietnam National University, Hanoi,
Vietnam
Fan Zhang,
University of Bristol, United Kingdom
Giuseppe Valenzise,
UMR8506 Laboratoire des Signaux et
Systèmes (L2S), France

*Correspondence:

Lorenzo Ciccarelli
lorenzo.ciccarelli@v-nova.com

Specialty section:

This article was submitted to
Image Processing,
a section of the journal
Frontiers in Signal Processing

Received: 26 February 2022

Accepted: 20 June 2022

Published: 08 September 2022

Citation:

Ciccarelli L, Ferrara S and Maurer F
(2022) MPEG-5 LCEVC for 3.0 Next
Generation Digital TV in Brazil.
Front. Sig. Proc. 2:884254.
doi: 10.3389/frsip.2022.884254

1 INTRODUCTION

The TV 3.0 project aims at a next generation television system using two complementary delivery methods: Over-The-Air (OTA) and Over-The-Top (OTT). The project is led by the SBTVD Forum (Fórum Sistema Brasileiro TV Digital Terrestre), a Brazilian organization responsible for development of digital television in Brazil. In July 2020, SBTVD issued a CfP aiming at extending the current digital television system to new use cases and applications. Among these applications, video format such as 8K, HDR, AR/VR shall be supported and delivered over broadcast, or hybrid broadband/broad networks, based on IP-centric protocols. In the video coding part, the SBTVD committee has introduced the concept of scalability and the concept of video codec enhancement. In response to the CfP, MPEG-5 Low Complexity Enhancement Video Coding (LCEVC) was proposed as a candidate technology for the *Enhancement Video Coding* use case which envisages a scalable video coding solution used to deliver content across users with different bandwidth and decoding capabilities. LCEVC enables deployments where a service provider intends to enhance existing codecs (e.g., H.264/AVC-H.265/HEVC-based services which can be upgraded to LCEVC-over- H.264/AVC-H.265/HEVC based services) or allows the adoption of new standard by enhancing new codec (e.g., LCEVC-over-AV1- or H.266/VVC based services). LCEVC allows to deliver higher resolutions, frame rate, maintaining or improving the quality of the existing service while using lower bitrate

and reducing the complexity. The ability of LCEVC to enhance any codec has been demonstrated in several use cases ((Jiménez-Moreno et al., 2022; Martini) and (Lcevc, 2020a) but in particular during the MPEG Verification Tests (Mpeg) where the coding efficiency LCEVC has been appraised when enhancing H.264/AVC, H.265/HEVC, (EVC) and H.266/VVC. The purpose of those verification tests was to confirm that the main goal set by MPEG for LCEVC, i.e. reducing the bit rate at the same level of visual quality of a single-layer video codec, was successfully accomplished. The document in (V-Nova, 2021) reports the results as follows. The first set of tests compared full-resolution LCEVC-enhanced encoded sequences with full-resolution single-layer anchors. The average bit rate savings for LCEVC when enhancing H.264/AVC were determined to be approximately 46% for UHD (2160p) and 28% for HD (1080p) content. The average bit rate savings for LCEVC when enhancing H.265/HEVC were determined to be approximately 31% for UHD and 24% for HD. Numerical analysis of the average benefit of LCEVC and its statistical significance compared to the corresponding full resolution EVC or VVC codec was more difficult to interpret, due to several test points having overlapping confidence intervals. However, the test results tend to indicate an overall benefit when using LCEVC on top of these two codecs. The second set of tests aimed to confirm that LCEVC provides a more efficient solution for video content upsampling. In particular, the test clips were downsampled by a factor of two along the horizontal and vertical directions and then coded using the H.264/AVC, H.265/HEVC, EVC, or H.266/VVC standards as base layers. Anchors were generated by upsampling with Lanczos filters to full resolution (for visual assessment). Comparing LCEVC full-resolution encoded sequences with the up-sampled anchors, the average bit rate savings when using LCEVC with H.264/AVC, H.265/HEVC, EVC, and H.266/VVC were calculated to be approximately 28%, 34%, 38%, and 33% respectively for UHD, and 27%, 26%, 21%, and 21% respectively for HD.

In this context, the main goal and contribution of this paper is to overview the MPEG-5 LCEVC standard, most notably its architecture, coding toolset and main benefits offered during the deployment of streaming and broadcasting services. Moreover, the paper presents the coding efficiency offered by LCEVC over the test conditions specified by the TV 3.0 project. Accordingly, the remainder of the paper is organized as follows. **Section 2** provides a brief description of the TV 3.0 project main targets, **Section 3** introduces the MPEG-5 LCEVC standard along with its main technology components and key benefits for the deployment. **Section 4** presents the results of the experiments used for the assessment and the selection of MPEG-5 LCEVC in the multiple use cases envisaged in the TV 3.0 project. The results demonstrate the LCEVC's coding efficiency along with its flexibility to serve multiple application scenarios, confirming its readiness for commercial deployment in the context of TV 3.0 in Brazil. Finally, conclusions are drawn in **Section 5**.

2 THE TV 3.0 PROJECT

The TV 3.0 project aims at a next generation television system using two complementary delivery methods: OTA and OTT.

The project is led by the SBTVD Forum (Fórum Sistema Brasileiro TV Digital Terrestre), a Brazilian organization responsible for development of digital television in Brazil. The organization gathers more than eighty members, both from Academia and Industry, representing all parties involved in the complete ecosystem, including broadcasters, manufacturers, research institutions, and Universities. A CfP (Call for Proposal Phase 2, 2021) was issued in July 2020 to solicit novel technologies able to meet the requirements described in the call for proposals document. Such requirements related to the different aspects involved in the video content delivery chain (e.g. channel modulation techniques, audio compression, etc.) As far as the video compression technology is concerned, the main requirements may be briefly summarized as follows:

- Spatial resolutions spanning from 1280×720 (aka 720p) up to 7680×4320 with 16:9 aspect ratio and progressive only scanning format with 4:2:0 chroma subsampling.
- Colorimetry with ITU-R BT.2020 color primaries and bit depths up to 10 bits per pixel. Transfer characteristics functions (i.e. Opto-Electronic Transfer Function (OETF) and Electro-Optical Transfer Function (EOTF)) from ITU-R BT.2100, that is either Hybrid Log-Gamma (HLG) or Perceptual Quantizer (PQ).
- Temporal resolutions spanning from 25 to 120 frames per second (fps).
- Both single layer as well as scalable

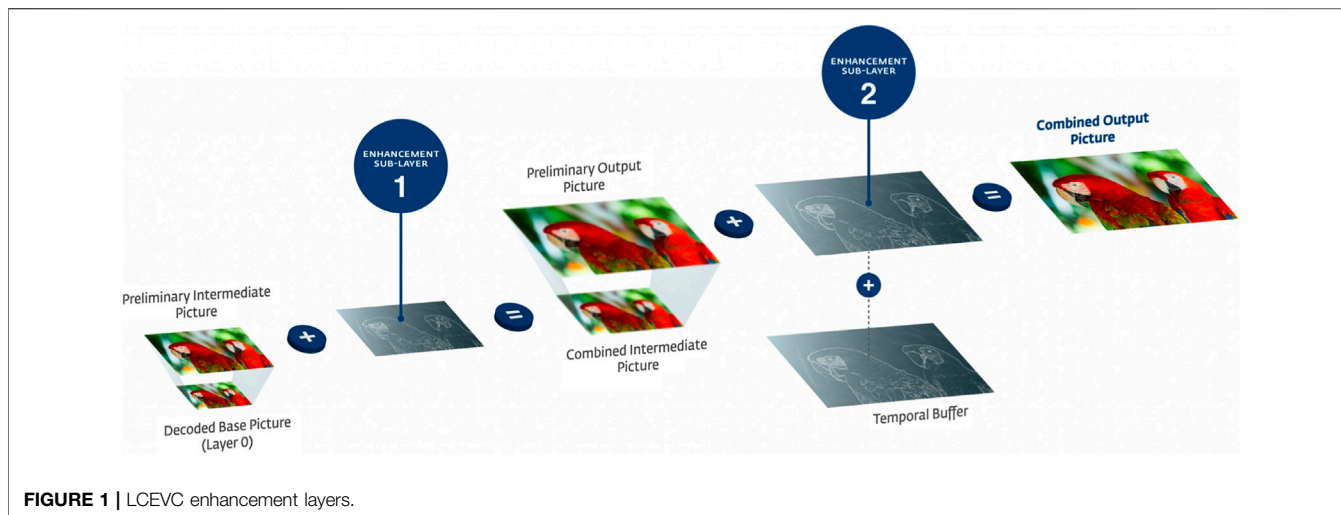
Operating coding rates are not explicitly specified in the CfP document but rather constraints on the quality are set, for example the bitrate should guarantee a Mean Opinion Score (MOS) value of 4 and above. Different combinations of spatial, temporal and pixel dynamic range resolutions are grouped in the so-called use cases, identified by "VCx" whereby "x" denotes the identification number of the use case.

MPEG-5 LCEVC was proposed as a candidate technology for the Enhancement Video Coding use case which addresses application scenarios where scalable video coding technology may be used to serve end users, each having different bandwidth limitations and decoding capabilities **Figure 1** and **Figure 2**.

In January 2022, the final decision was published by SBTVD Forum—for further details please see https://forumsbtvd.org.br/tv3_0/. The decision, which has been ratified by the Brazilian Ministry of Communications, was reached unanimously by the SBTVD Forum considering the results of the testing and evaluation, as well as market and intellectual property aspects of the candidate technologies. **Figure 3** shows a summary of the decision.

Considering the video coding components of TV 3.0, the following decisions were made and next steps planned accordingly.

H.266/VVC has been adopted as main codec for the video base layer (both OTA and OTT) with H.264/AVC and H.265/HEVC



continued to be supported for distribution of alternative content over the Internet; and LCEVC has been adopted for the video enhancement (OTA and OTT, for both legacy codecs and H.266/VVC) in combination with Dynamic Resolution Encoding (DRE).

Phase 3 of TV 3.0 is being planned and expected to last about 2 years (2022–2023), contemplating, among other activities, complementary tests for the selection of the physical layer technology, development of the necessary adaptations and extensions to the transport layer specification, subjective assessment of the video coding quality (determination of the necessary bitrate), development of adaptations and extensions to DTV Play for TV 3.0 Application Coding, elaboration of ABNT technical standards for TV 3.0, development of interoperability tests, demonstrations, etc.

The TV 3.0 is expected to launch in 2024.

3 MPEG-5 LOW COMPLEXITY ENHANCEMENT VIDEO CODING

This section overviews the LCEVC standard by providing the user with a description of the main codec architecture as well as the main coding tools used to compress the enhancement layers. A discussion on the main benefits associated with LCEVC and different with respect to other scalable video coding technologies is also provided.

3.1 Architecture and Coding Toolset Overview

The MPEG-5 Low Complexity Enhancement Video Coding (LCEVC) standard specifies a two-level pyramidal coding scheme where the base layer is compressed with any format

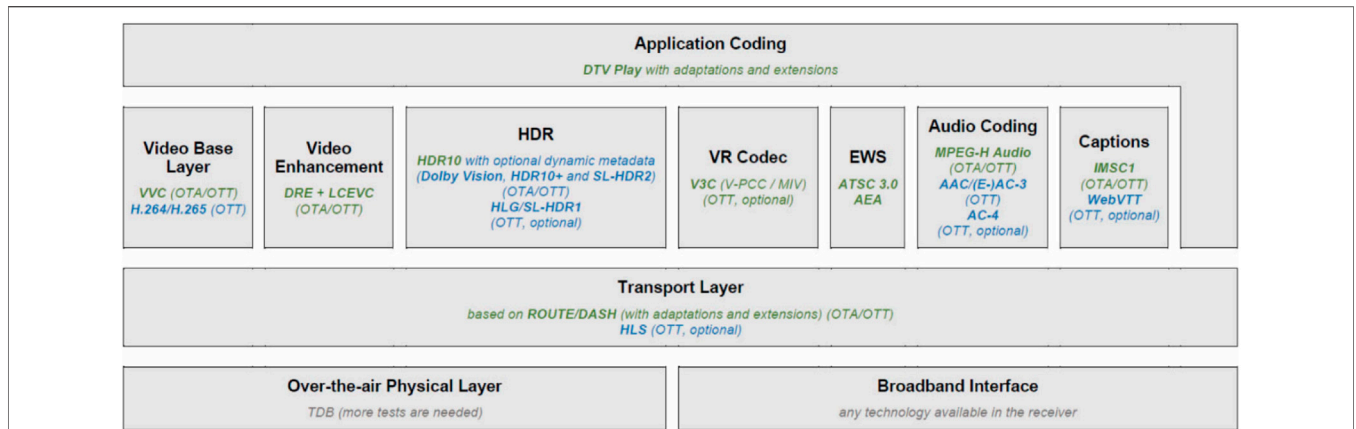


FIGURE 3 | Summary of the TV 3.0 adopted technologies

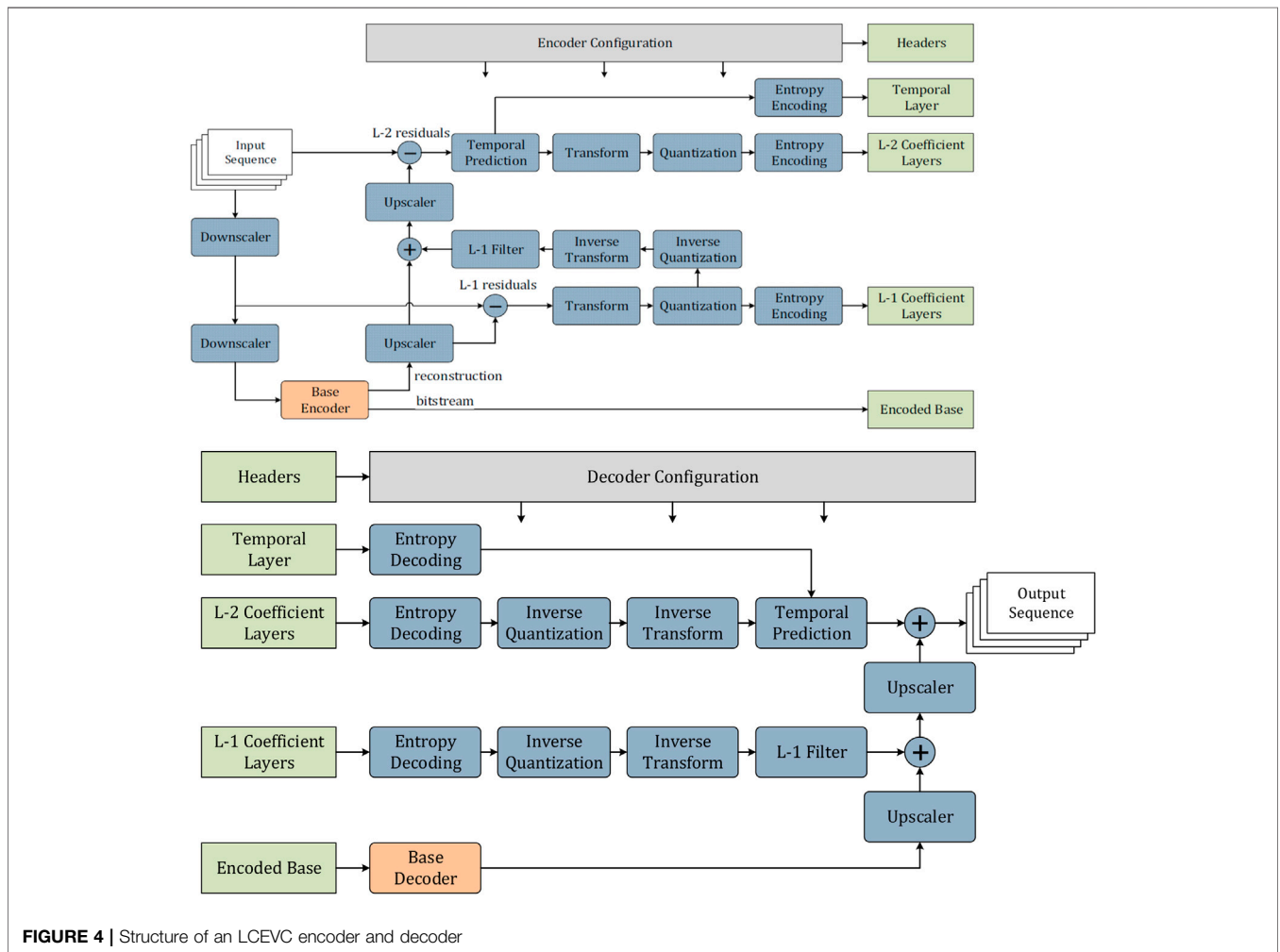


FIGURE 4 | Structure of an LCEVC encoder and decoder

selected by the user whilst the enhancement layers are compressed using the coding tools from the standard’s toolset. The structure of an LCEVC encoder is depicted in Figure 4. The high level encoding process can be divided into three main steps:

Base layer compression, Sub-layer 1 (L-1) and Sub-layer 2 (L-2) compression. Each of these steps then involves some low level processing carried out with the coding tools specified by the standard. In the following subsections, the three steps associated

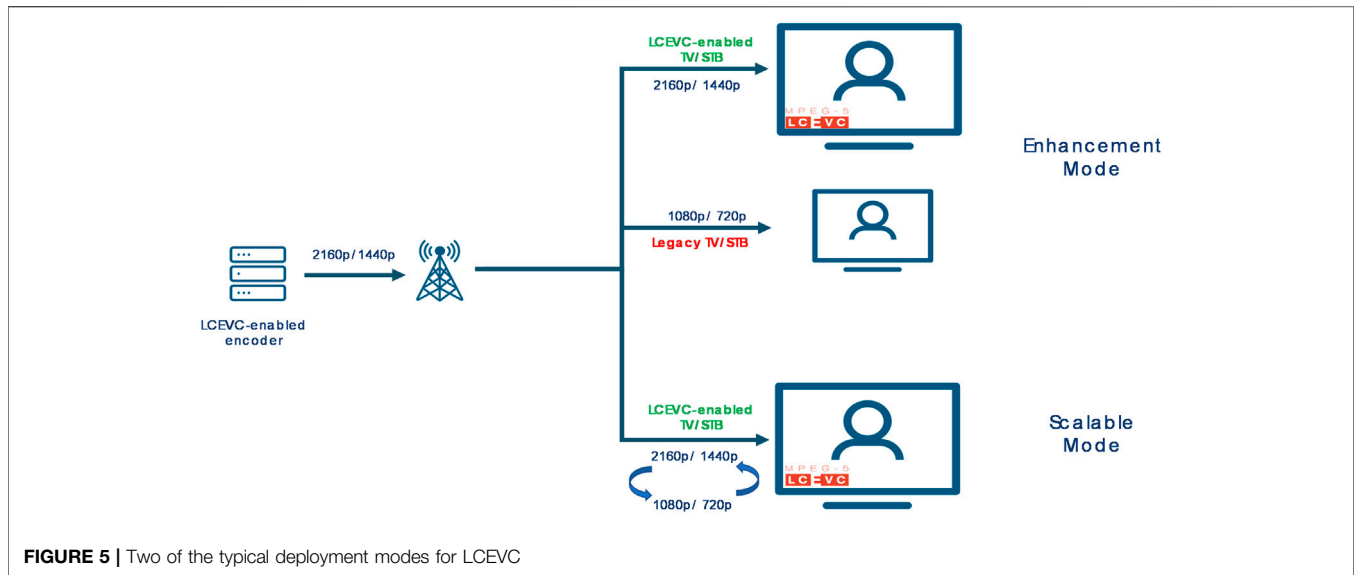


FIGURE 5 | Two of the typical deployment modes for LCEVC

with the high level encoding process are introduced first, followed by a description of each coding tool. The picture below shows how the enhancement is applied at every sublayers.

3.1.1 Base-Layer Compression

Firstly, the input sequence is downscaled using a non-normative down-scaler. Depending on the chosen encoder configuration, the downscaling can be applied up to two consecutive times. The video, now at a lower resolution than the input sequence, is fed into the base encoder, which may be compliant with the H.264/AVC, H.265/HEVC, H.266/VVC standards or even a proprietary solution: LCEVC is agnostic to the base layer compression technology employed. Accordingly, the base layer compression process is not further specified in LCEVC: any encoder that produces a decodable bitstream can be used.

3.1.2 Sub-Layer 1 (L-1) Compression

The normative up-scaler specified in the standard (ISO/IEC DIS, 2022), is employed to produce an upscaled version of the base-layer which is then used as input to this encoding stage. The enhancement sub-layer 1 (L-1) residuals are created by subtracting, for each color plane, the downscaled input sequence and the reconstructed base-layer. The residuals obtained are typically sparse, that is containing samples with value close to zero, except where image edges or textured details are present in the original content. Residuals associated with each color plane are then partitioned over a non overlapping grid of *coding blocks* with size either 4×4 or 2×2 , depending on the selected encoder's configuration. Over each coding block, spatial transformation, quantization and entropy encoding are applied. Details of the different tools used for transform, quantization and entropy encoding will be provided after the three main encoding stages are described. After the sub-layer 1 residuals are encoded, the inverse quantization and inverse transform stage are applied. Additionally, a filter can be applied to the sub-layer 1 reconstructed residuals which functions as a simple

deblocking filter. Worth mentioning here that the LCEVC standard supports the signaling of different values for the kernels used to upscale the base-layer to the sub-layer L-1, moreover a nonlinear correction denoted as *Predicted Residuals (PR)* may also be used to zero out and/or adjust the sample values with the ultimate goal of improving the coding efficiency of the different coefficient groups. The PR may be seen as a pre-processing step which increases the probability of having a sparse residual signal with long runs of zero values, so that the associated entropy is further reduced. The transform used in LCEVC has a simple structure and uses a small kernel of size 2×2 or 4×4 . This allows to both efficiently code sparse information and parallelize the transforms, since individual blocks are not dependent on other blocks within a frame. A linear quantizer, which may include an adaptive dead-zone, is used to further process the transform coefficients. The entropy encoder, which consists of a run-length encoder (RLE) and an optional prefix encoder (Huffman encoder), processes the quantized transform coefficients and creates the coefficient groups for sub-layer 1.

3.1.3 Sub-Layer 2 (L-2) Compression

The sub-layer 1 reconstructed residuals are upscaled to full resolution and subtracted from the original input sequence pixels. The resulting enhancement sub-layer 2 (L-2) residuals are fed into the temporal prediction module depicted in **Figure 4**. LCEVC uses a zero-motion vector temporal scheme which operates on a coding block basis. The residuals from the previous frame are stored in a temporal buffer and are added to the L-2 residuals in case the temporal prediction is activated. To reduce the signaling overhead in the case of a fast-moving sequence, where the zero-motion compensation scheme would likely not be beneficial, temporal prediction may be disabled by simply transmitting a binary flag for a group of pixels of size 32×32 or for the entire frame. As for the case of sub-layer 1 residuals, also the sub-layer 2 residuals are transformed, quantized and entropy encoded using the same tools from sub-layer 1.

TABLE 1 | MOS scores and CI for PQ sequences.

Sequence	LCEVC (config A) vs. HM	LCEVC (config B) vs. HM
	BD-rate	BD-rate
vc10-lcevc-1	-45.93%	-71.52%
vc10-lcevc-2	-56.86%	-37.13%
vc10-avs3	-27.51%	10.24%
vc10-dre	-79.54%	-37.37%
vc10-globo	-78.40%	-42.24%
vc10-vcv	-27.23%	-41.73%
Average	-52.58%	-36.63%

TABLE 2 | MOS scores and CI for HLG sequences.

Sequence	LCEVC (config A) vs. HM	LCEVC (config B) vs. HM
	BD-rate	BD-rate
vc04-lcevc-1	-54.24%	-82.51%
vc04-lcevc-2	-62.22%	-42.46%
vc04-avs3	-37.46%	-17.90%
vc04-dre	-57.66%	-38.15%
vc04-globo	-72.57%	-54.99%
vc04-vcv	-43.46%	-59.01%
Average	-54.60%	-49.17%

Now that the high level workflow specified by the LCEVC standard has been described, the details associated with the coding tools for transform, quantization and entropy coding can be provided.

3.1.4 Spatial Transformation

LCEVC uses a linear block-based transformation which is similar to 2×2 or 4×4 Hadamard kernel as shown by the following two matrix formulations where R_{xy} are the residuals and d_{xy} are the transform coefficients.

$$\{R_{00}\}\{R_{01}\}\{R_{10}\}\{R_{11}\} = \begin{matrix} \{1, 1, 1, 1\} & \{d_{00}\} \\ \{1, -1, 1, -1\} & \{d_{01}\} \\ \{1, 1, -1, -1\} & \{d_{10}\} \\ \{1, -1, -1, 1\} & \{d_{11}\} \end{matrix}$$

$$\begin{matrix} \{R_{00}\} & \{1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1\} & \{d_{00}\} \\ \{R_{01}\} & \{1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1\} & \{d_{01}\} \\ \{R_{02}\} & \{1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1\} & \{d_{02}\} \\ \{R_{03}\} & \{1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 & -1 & -1\} & \{d_{03}\} \\ \{R_{10}\} & \{1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1\} & \{d_{10}\} \\ \{R_{11}\} & \{1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1\} & \{d_{11}\} \\ \{R_{12}\} & \{1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1\} & \{d_{12}\} \\ \{R_{13}\} & \{1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & 1 & -1 & 1 & -1\} & \{d_{13}\} \\ \{R_{20}\} & \{1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & 1 & 1 & -1 & -1 & 1 & 1\} & \{d_{20}\} \\ \{R_{21}\} & \{1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 & 1 & 1 & -1 & -1 & -1 & 1\} & \{d_{21}\} \\ \{R_{22}\} & \{1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1\} & \{d_{22}\} \\ \{R_{23}\} & \{1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 & -1 & 1 & -1 & 1 & 1 & 1 & -1 & -1\} & \{d_{23}\} \\ \{R_{30}\} & \{1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & 1\} & \{d_{30}\} \\ \{R_{31}\} & \{1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & 1 & 1 & -1 & -1\} & \{d_{31}\} \\ \{R_{32}\} & \{1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & -1 & 1 & 1 & -1 & 1 & 1 & -1\} & \{d_{32}\} \\ \{R_{33}\} & \{1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & -1 & 1 & -1 & -1 & 1\} & \{d_{33}\} \end{matrix}$$

The use of the Hadamard transform is due to its better energy compaction of sparse input data, which is indeed the case for the enhancement layers prediction residuals. Moreover, the block size

TABLE 3 | MOS scores and CI for PQ sequences.

Sequence	Rates	Configuration A		Configuration B		Anchor (HM)	
		MOS	CI	MOS	CI	MOS	CI
vc10-lcevc-1	R4	7.81	0.26	8.19	0.31	7.50	0.25
vc10-lcevc-1	R3	6.63	0.42	7.50	0.55	5.12	0.30
vc10-lcevc-1	R2	5.38	0.46	6.56	0.49	3.99	0.35
vc10-lcevc-1	R1	2.94	0.41	4.69	0.38	2.25	0.37
vc10-lcevc-2	R4	7.75	0.44	7.75	0.41	7.10	0.24
vc10-lcevc-2	R3	6.25	0.54	6.00	0.35	5.16	0.42
vc10-lcevc-2	R2	5.31	0.33	4.50	0.39	3.76	0.35
vc10-lcevc-2	R1	3.81	0.26	2.81	0.36	2.44	0.30
vc10-avs3	R4	8.13	0.29	7.50	0.49	7.19	0.21
vc10-avs3	R3	6.94	0.41	6.25	0.37	6.19	0.40
vc10-avs3	R2	5.63	0.42	4.81	0.31	5.13	0.24
vc10-avs3	R1	3.75	0.21	2.38	0.29	3.56	0.46
vc10-dre	R4	8.56	0.30	8.06	0.27	6.54	0.30
vc10-dre	R3	7.38	0.29	6.00	0.35	5.34	0.34
vc10-dre	R2	6.75	0.56	4.94	0.37	4.05	0.49
vc10-dre	R1	5.00	0.81	3.25	0.48	2.31	0.33
vc10-globo	R4	8.06	0.32	7.94	0.37	7.25	0.57
vc10-globo	R3	7.69	0.29	6.44	0.30	6.03	0.34
vc10-globo	R2	6.81	0.26	5.56	0.30	4.44	0.35
vc10-globo	R1	5.50	0.30	3.63	0.34	2.13	0.29
vc10-vcv	R4	7.88	0.38	8.06	0.37	7.69	0.38
vc10-vcv	R3	5.69	0.59	6.13	0.38	5.42	0.55
vc10-vcv	R2	4.63	0.42	5.13	0.24	3.84	0.24
vc10-vcv	R1	3.44	0.39	3.75	0.27	2.13	0.41

TABLE 4 | MOS scores and CI for HLG sequences.

Sequence	Rates	Configuration A		Configuration B		Anchor (HM)	
		MOS	CI	MOS	CI	MOS	CI
vc04-lcevc-1	R4	8.00	0.35	8.75	0.37	7.17	0.25
vc04-lcevc-1	R3	7.13	0.49	8.31	0.29	5.12	0.30
vc04-lcevc-1	R2	5.63	0.24	7.44	0.42	3.99	0.35
vc04-lcevc-1	R1	3.75	0.21	6.44	0.57	2.25	0.37
vc04-lcevc-2	R4	7.75	0.44	7.75	0.41	6.31	0.33
vc04-lcevc-2	R3	6.25	0.54	6.00	0.35	4.81	0.36
vc04-lcevc-2	R2	5.31	0.33	4.50	0.39	3.56	0.35
vc04-lcevc-2	R1	3.81	0.26	2.81	0.36	1.94	0.32
vc04-avs3	R4	8.50	0.30	8.25	0.37	8.00	0.43
vc04-avs3	R3	7.94	0.44	7.56	0.24	7.25	0.27
vc04-avs3	R2	6.88	0.52	6.25	0.51	5.69	0.33
vc04-avs3	R1	5.13	0.49	4.00	0.25	3.63	0.29
vc04-dre	R4	8.31	0.29	7.88	0.46	6.81	0.36
vc04-dre	R3	7.31	0.23	7.00	0.30	6.00	0.51
vc04-dre	R2	5.94	0.41	5.19	0.58	4.81	0.36
vc04-dre	R1	5.50	0.30	4.50	0.30	3.50	0.30
vc04-globo	R4	8.88	0.34	8.31	0.33	7.69	0.23
vc04-globo	R3	7.88	0.38	7.00	0.46	5.56	0.39
vc04-globo	R2	6.50	0.55	5.81	0.31	4.63	0.29
vc04-globo	R1	4.63	0.38	3.88	0.29	3.19	0.36
vc04-vcv	R4	7.88	0.38	8.06	0.37	5.88	0.52
vc04-vcv	R3	5.69	0.59	6.13	0.38	5.13	0.30
vc04-vcv	R2	4.63	0.42	5.13	0.24	3.88	0.33
vc04-vcv	R1	3.44	0.39	3.75	0.27	2.25	0.32

is limited to 4×4 to allow for parallel encoding implementations, whereby each block can be transformed independently on a multi threaded CPU or GPU architecture. The size used by the encoder is transmitted in the bitstream as coding metadata.

3.1.5 Quantization

LCEVC specifies a uniform scalar quantizer with dead-zone adjustment. The quantizer can use different quantization steps (denoted as quantization parameters in the standard document) for the two enhancement sub-layers allowing to balance the bit budget between the two sub-layers and decide where to add more details. Reconstruction of residuals using the inverse quantization process can be performed using asymmetric dequantization whereby the dequantized coefficient values can be offset to correct the edge of the bins depending on the magnitude of the quantization parameter used.

3.1.6 Entropy Coding

The quantized transform coefficients are grouped together with respect to their associated frequency band. Over each band, the quantized coefficients of each coding block are encoded using a simple run-length encoding. More precisely, (r, v) pairs are formed from the coefficients with r being the run length of a consecutive group of zero valued quantized coefficients and v is the value which interrupts the run. The values for r and v are entropy encoded using prefix coding, whereas a different table is used for r and v and transmitted in the bitstream.

3.1.7 Temporal Prediction

As already mentioned, LCEVC uses a zero motion compensation temporal prediction scheme which operates

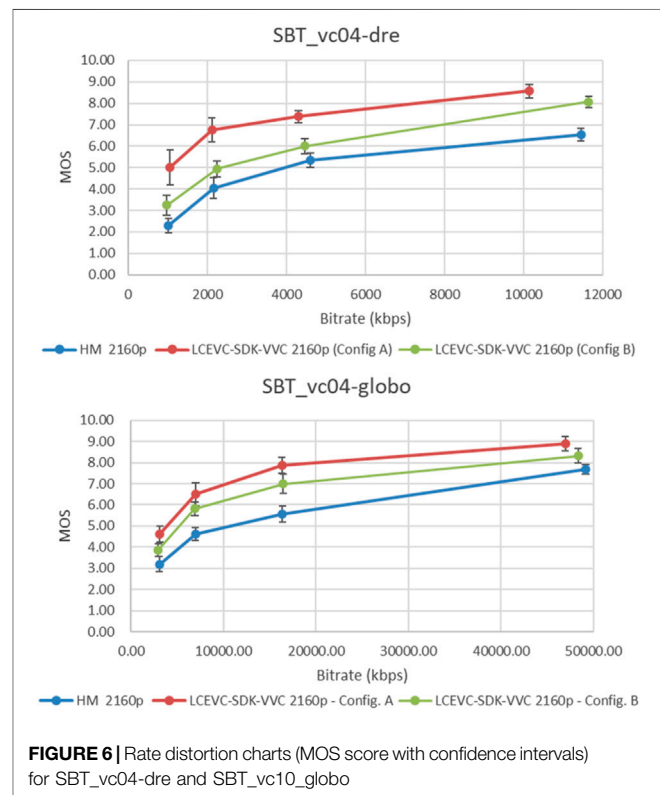


FIGURE 6 | Rate distortion charts (MOS score with confidence intervals) for SBT_vc04-dre and SBT_vc10_globo

over the prediction residuals. The standard's syntax supports the possibility to signal the use of inter prediction on a coding block basis via a binary flag. The ensemble of these

TABLE 5 | BD-rate savings for PQ sequence– Configuration A.

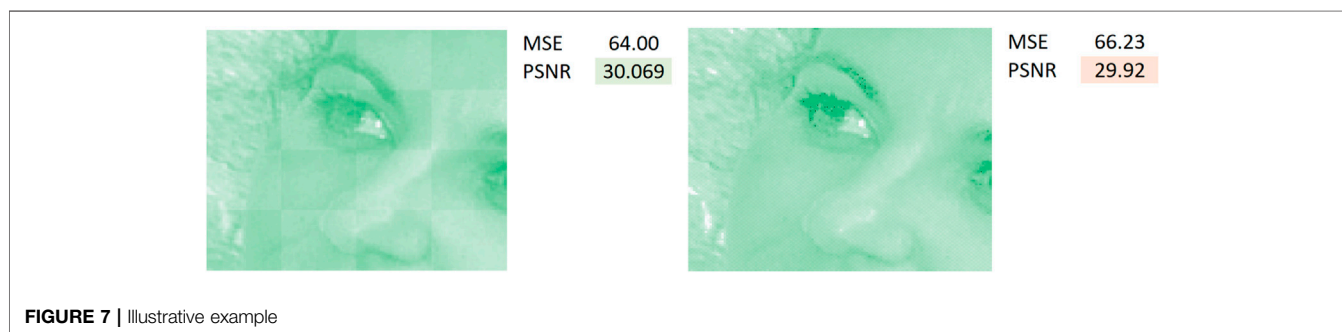
Sequence	VMAF	MS-SSIM-Y	MS-SSIM-U	MS-SSIM-V	PSNR-Y	PSNR-U	PSNR-V
vc10-lcevc-1	-29.32%	-12.49%	-69.30%	-63.63%	14.42%	-31.56%	-34.26%
vc10-lcevc-2	-49.75%	-17.57%	-58.60%	-73.19%	-25.90%	-64.55%	-78.31%
vc10-avs3	-17.48%	-10.11%	-9.14%	-40.89%	23.73%	-9.14%	-27.42%
vc10-dre	-27.12%	-33.43%	-26.95%	-38.71%	-5.76%	-54.54%	-69.47%
vc10-globo	-27.13%	-5.58%	-60.30%	-58.78%	5.20%	-0.19%	-41.61%
vc10-HEVC	-57.90%	-30.88%	-66.82%	-74.07%	-30.76%	-60.97%	-73.90%
Average	-34.78%	-20.01%	-48.52%	-58.21%	-3.18%	-36.83%	-54.16%

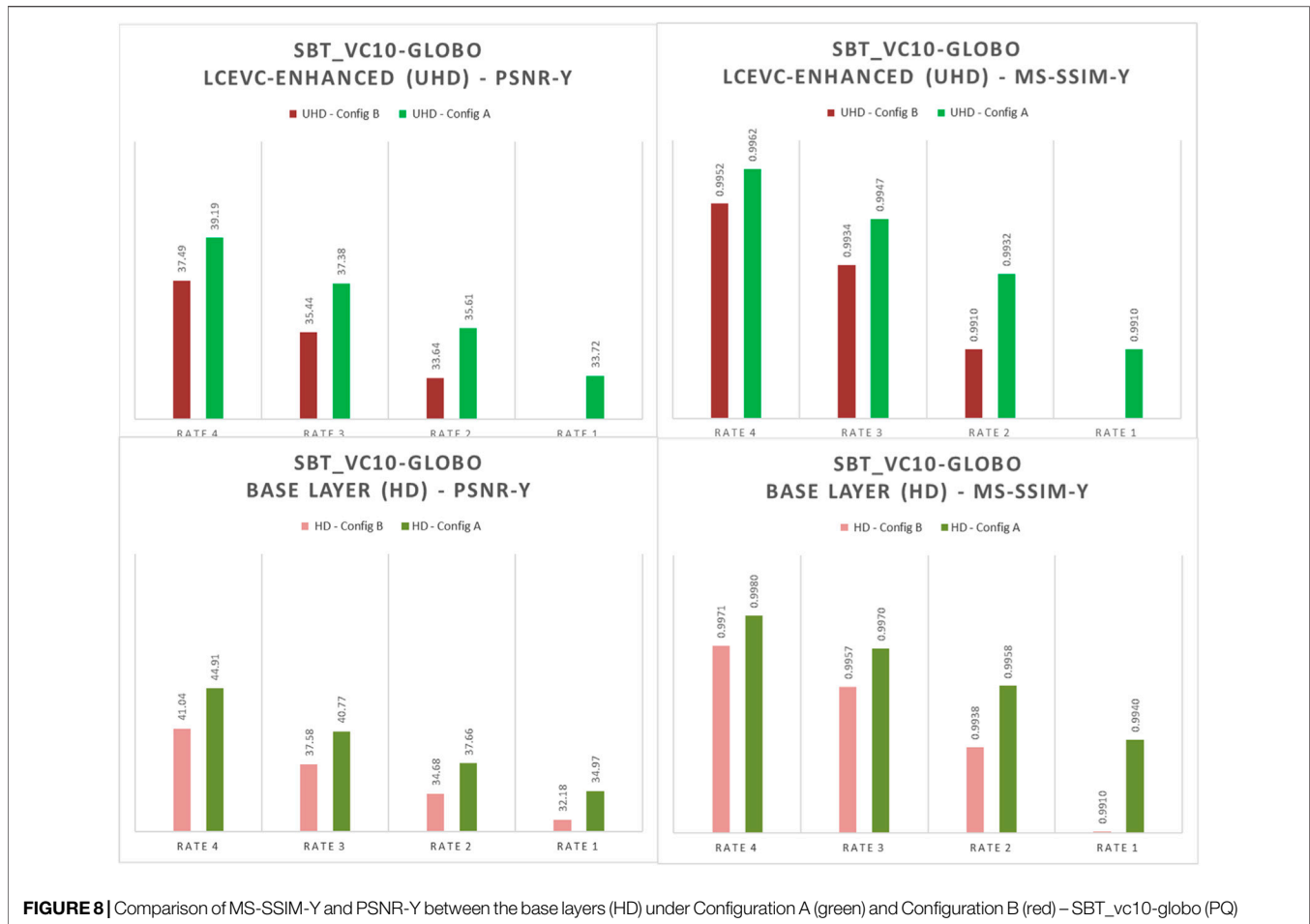
Sequence	H.265/HEVC (HM)			LCEVC-enhancing H.266/VVC		
	Bitrate	MOS	CI	Bitrate	MOS	CI
vc10-lcevc-1	5798	6.35	0.24	5087	7.29	0.42
vc10-lcevc-2	5892	6.38	0.34	5239	8.38	0.34
vc10-avs3	4444	4.70	0.49	5140	6.59	0.35
vc10-dre	5569	6.01	0.29	5073	7.49	0.37
vc10-globo	5366	3.23	0.46	5120	3.5	0.49
vc10-vc	5574	6.00	0.32	5048	7.18	0.51
Average	5440	5.45	-	5152	6.74	-

TABLE 6 | BD-rate savings for HLG sequence – Configuration A.

Sequence	VMAF	MS-SSIM-Y	MS-SSIM-U	MS-SSIM-V	PSNR-Y (%)	PSNR-U	PSNR-V
vc04-lcevc-1	-39.86%	4.5%	-66.80%	-70.47%	55.90	-15.25%	-40.00%
vc04-lcevc-2	-36.86%	-15.98%	-33.60%	-33.60%	3.98	-19.21%	-23.47%
vc04-avs3	19.74%	7.26%	8.46%	8.29%	47.41	10.50%	-1.42%
vc04-dre	-26.96%	-33.12%	-34.69%	-36.62%	3.69	-36.09%	-57.99%
vc04-globo	-31.57%	-21.2%	-92.85%	-55.50%	20.92	-18.20%	-36.09%
vc04-vc	-8.42%	-13.56%	-29.28%	-71.03%	80.58	198.52%	0.64%
Average	-20.61%	-12.01%	-41.46%	-46.90%	35.41	20.05%	-26.39%

Sequence	H.265/HEVC (HM)			LCEVC-enhancing H.266/VVC		
	Bitrate	MOS	CI	Bitrate	MOS	CI
vc04-lcevc-1	4472	4.57	0.61	5051	5.79	0.69
vc04-lcevc-2	5274	4.95	0.55	5085	7.98	0.44
vc04-avs3	4175	7.81	0.47	5215	8.19	0.36
vc04-dre	5179	5.81	0.50	5210	6.27	0.32
vc04-globo	5054	4.11	0.57	5188	5.83	0.58
vc04-vc	4705	4.49	0.83	5390	5.81	0.63
Average	4810	5.29	-	5240	6.65	-





binary flags is entropy encoded to improve the whole coding efficiency.

3.2 Key Benefits of Deploying MPEG-5 LCEVC

Due to its multi-layer structure, LCEVC can *also* be used as a scalable codec, although it has not been primarily built as a scalable codec.

Because of this unique feature, LCEVC can typically be used in two modes (see **Figure 5**):

- **Enhancement Mode:** to provide an enhanced full resolution video - albeit starting from a lower resolution video which is separately and independently decodable; or
- **Scalable Mode:** to provide a scalable video, where a same receiver can switch between a lower and a higher resolution.

Indeed, current commercial deployments use LCEVC to provide an enhanced full resolution video on top of a base codec widely available in the ecosystem (e.g., H.264/AVC, H.265/HEVC or AV1), knowing that devices which, for one reason or another cannot decode the LCEVC component, can

still decode the lower resolution base layer and display a video at a quality which is typically similar or higher than that they would otherwise receive.

This aspect allows for a phased deployment of LCEVC without affecting current services. For example, LCEVC enables deployments where a service provider intends to enhance existing codecs (e.g., H.264/AVC- H.265/HEVC -based services which can be upgraded to LCEVC-over- H.264/AVC- H.265/HEVC based services).

However, LCEVC is also very important for deployments where a service provider intends to enhance new codecs (e.g., LCEVC-over-AV1- H.266/VVC based services). In this case, besides the inherent advantages of using LCEVC as described above, LCEVC enables a more efficient and sustainable use of the new codec, particularly for live scenarios, as well as helps accelerating the deployment and ease transitioning from existing system to the new ones.

The low-complexity nature of LCEVC has other deployment advantages. Naturally, like any other coding technology, LCEVC can be deployed in silicon/ASIC.

In contrast to other coding technologies though, efficient LCEVC implementations can also be obtained without need for bespoke hardware, with decoding done using general



FIGURE 9 | Comparison of MS-SSIM-Y and PSNR-Y between the base layers (HD) under Configuration A (green) and Configuration B (red) – SBT_vc10-lcevc-2 (PQ)

processing units (CPU, GPU, DSP). This may enable quicker deployment by adding the LCEVC functionality to the decoding stack and, where possible, by software upgrading the existing devices.

Finally, LCEVC can reduce the processing cost of these newer compute-intensive codecs by up to 70%. Initial test results show that LCEVC reduces Video-on-Demand transcoding time (i.e., energy consumption) of the full Adaptive Bit Rate (ABR) ladder by 70% - or speeds up transcoding by 3.2x - while at the same time generating a superior video quality (Lcevc, 2020b).

3.3 Differences Between LCEVC and other Scalable Video Coding Technologies

As the described in Section 3.2.1 and 3.2.2, MPEG-5 LCEVC has some similarities with a scalable codec (i.e. spatial scalability thanks to the upsampler) but it is also substantially different for the following reasons:

- Generally, in scalable codec, the base layer is encoded with the same standard of the enhancement layer. As specified in the description LCEVC is codec agnostic. The base layer used in LCEVC can be any codec. This particular feature

allows LCEVC to be used with any standard, from the older one (i.e. H.264/AVC) to the latest one (H.266/VVC) and also with any others that have not been developed by MPEG (i.e. AV1, VP8, VP9 etc.) (Verification Test Report on the Compression Performance of Low Complexity Enhancement Video Coding, 2021) (Jiménez-Moreno et al., 2022) (Martini). It is worth mentioning here that even the scalable extension of the H.265/HEVC standard (Scalable High efficiency Video Coding, SHVC) allows the base layer to be compliant with any standard or proprietary format. However, SHVC will then use coding tools such as the Context Adaptive Binary Arithmetic Coding (CABAC) which may not be supported on legacy hardware (think about an H.264/AVC receiver compliant with the baseline profile) hence its deployment in application scenarios targeted by LCEVC may not be possible.

- The MPEG-5 LCEVC structure is using simple tools specifically designed for the sparse nature of the residual data which allow to keep the complexity low and limit the overhead associated with the enhancement layers, a common problem of scalable codecs. This makes possible to have a software version of the MPEG-5 LCEVC that can run on existing hardware and on top of existing base codec

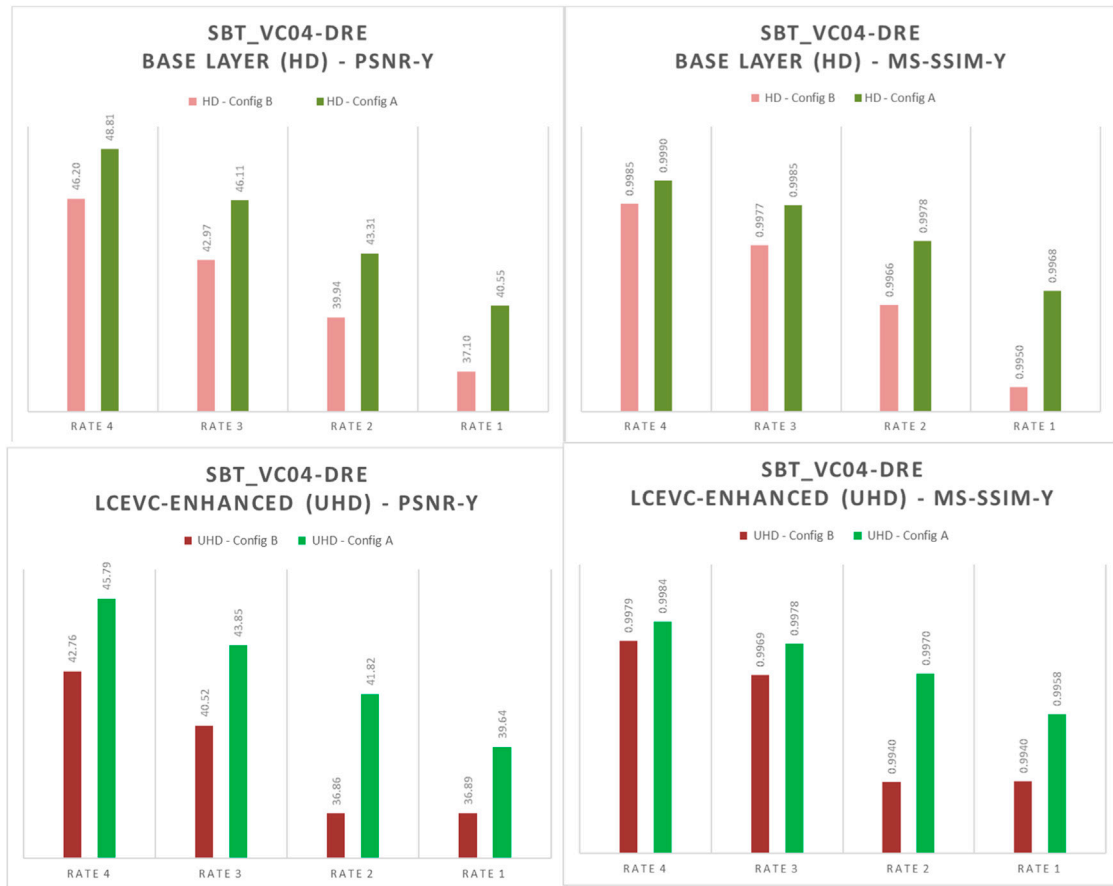


FIGURE 10 | Comparison of MS-SSIM-Y and PSNR-Y between the base layers (HD) under Configuration A (green) and Configuration B (red) – SBT_vc04-dre (HLG)

TABLE 7 | Bitrates, MOS scores and CI for PQ sequences. LCEVC and base layer in two separate 6MHz/5Mbps channels.

Sequence	H.265/HEVC (HM)			LCEVC-enhancing H.266/VVC		
	Bitrate	MOS	CI	Bitrate	MOS	CI
vc10-lcevc-1	10470	6.89	0.57	8631	7.56	0.35
vc10-lcevc-2	7995	7.01	0.32	8599	7.81	0.19
vc10-avs3	9026	6.24	0.30	8132	6.13	0.38
vc10-dre	9203	6.29	0.46	7836	7.50	0.46
vc10-globo	9284	4.02	0.42	8549	3.81	0.36
vc10-vc	9387	6.80	0.42	8927	7.50	0.49
Average	9302	6.21	-	8446	6.72	-

TABLE 8 | Bitrates, MOS scores and CI for HLG sequences. LCEVC and base layer in two separate 6MHz/5Mbps channels.

Sequence	H.265/HEVC (HM)			LCEVC-enhancing H.266/VVC		
	Bitrate	MOS	CI	Bitrate	MOS	CI
vc04-lcevc-1	9885	6.20	0.57	7796	6.13	0.49
vc04-lcevc-2	9916	5.98	0.35	8505	8.26	0.41
vc04-avs3	7926	8.13	0.29	8440	8.56	0.24
vc04-dre	9701	6.44	0.46	7556	6.75	0.59
vc04-globo	11661	5.02	0.65	8333	5.75	0.56
vc04-vc	10111	5.25	0.79	8957	6.54	0.46
Average	9867	6.17	-	8264	7.00	-

with no need to develop a specific hardware for it. As a consequence, the base codec can work more efficiently and faster given the ability of LCEVC to work with a base codec running at a quarter of the resolution. As a demonstration of this the results of MPEG verification test are described in³.

- Differently to most of scalable codecs, MPEG-5 LCEVC provides two levels of enhancement that can be applied at different stages or resolutions. Each level has its own independent quantization module and sublayer of the bitstream that can easily decoupled from the other. This also allows bitrate allocation flexibility to cope with different type of content.
- From **Figure 4** and Section 3.2.1, it may be noted that MPEG-5 LCEVC offers up to two cascade scaling processes in order to

³<https://www.v-nova.com/lcevc-enhanced-video/lcevc-video-sdk/>

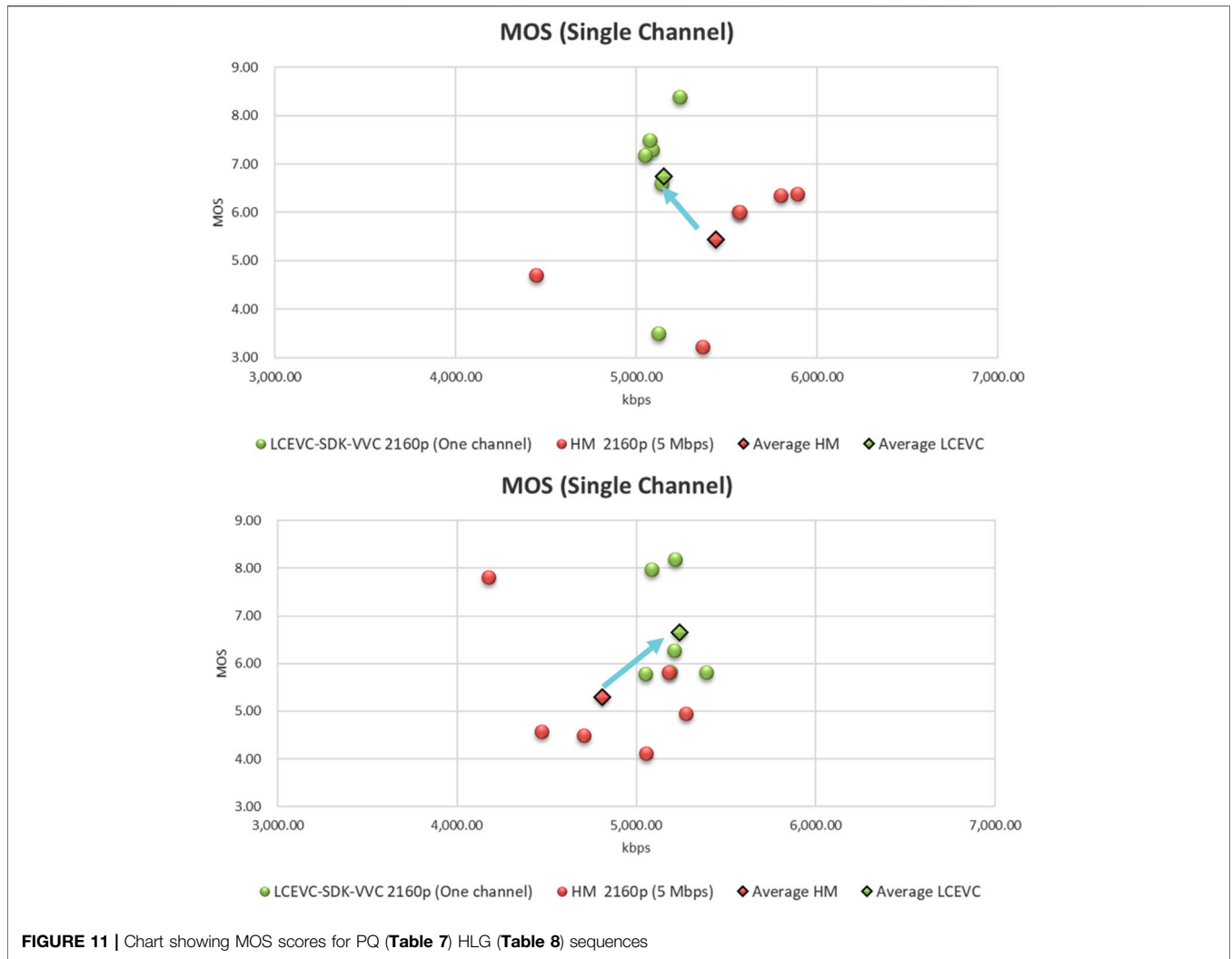


FIGURE 11 | Chart showing MOS scores for PQ (Table 7) HLG (Table 8) sequences

TABLE 9 | BD-rate savings for LCEVC-enhanced x265 over x265 (PQ).

Sequence	VMAF	MS-SSIM-Y	PSNR
vc10-lcevc-1	-27.12%	-55.53%	-20.07%
vc10-lcevc-2	-28.06%	-57.80%	-58.29%
vc10-avs3	-4.33%	-2.80%	3.02%
vc10-dre	-17.36%	-33.03%	-17.73%
vc10-globo	-4.52%	-21.04%	-24.96%
vc10-vc	-26.55%	-56.29%	-57.24%
Average	-17.99%	-37.75%	-29.21%

TABLE 10 | BD-rate savings for LCEVC-enhanced x265 over x265 (HLG).

Sequence	VMAF	MS-SSIM-Y	PSNR
vc04-lcevc-1	-14.90%	-18.29%	14.18%
vc04-lcevc-2	-11.59%	-13.95%	-0.59%
vc04-avs3	-19.20%	-14.34%	-6.67%
vc04-dre	-15.72%	-26.37%	-7.56%
vc04-globo	-1.43%	-15.24%	-6.02%
vc04-vc	-13.61%	-38.28%	-2.81%
Average	-12.80%	-21.08%	-1.58%

further improve the efficiency of the base layer. Each scaler can be user defined, along the following degrees of freedom: kernel size, type of upscaling (i.e. which sub-layer, L-1 or L-2) and kernel values. MPEG-5 LCEVC offers 4 normative upsamplers and one 4 taps user defined kernel. Scalable codecs are generally offering only one fixed scaling engine and it is not programmable.

- MPEG-5 LCEVC can handle different bit depths up to 14 bits per pixel in the main profile. The standard allows

the base layer to work on a different bit depth compared the input signal one. This operation can effectively enhance a base layer working at a lower bit depth to a higher one contributing to maintain the fidelity of the input signal. An example of this application is delivering HDR with technologies that cannot deliver more than 8bit like AVC⁴.

⁴<https://www.xilinx.com/products/acceleration-solutions/v-nova-lcevc.html>

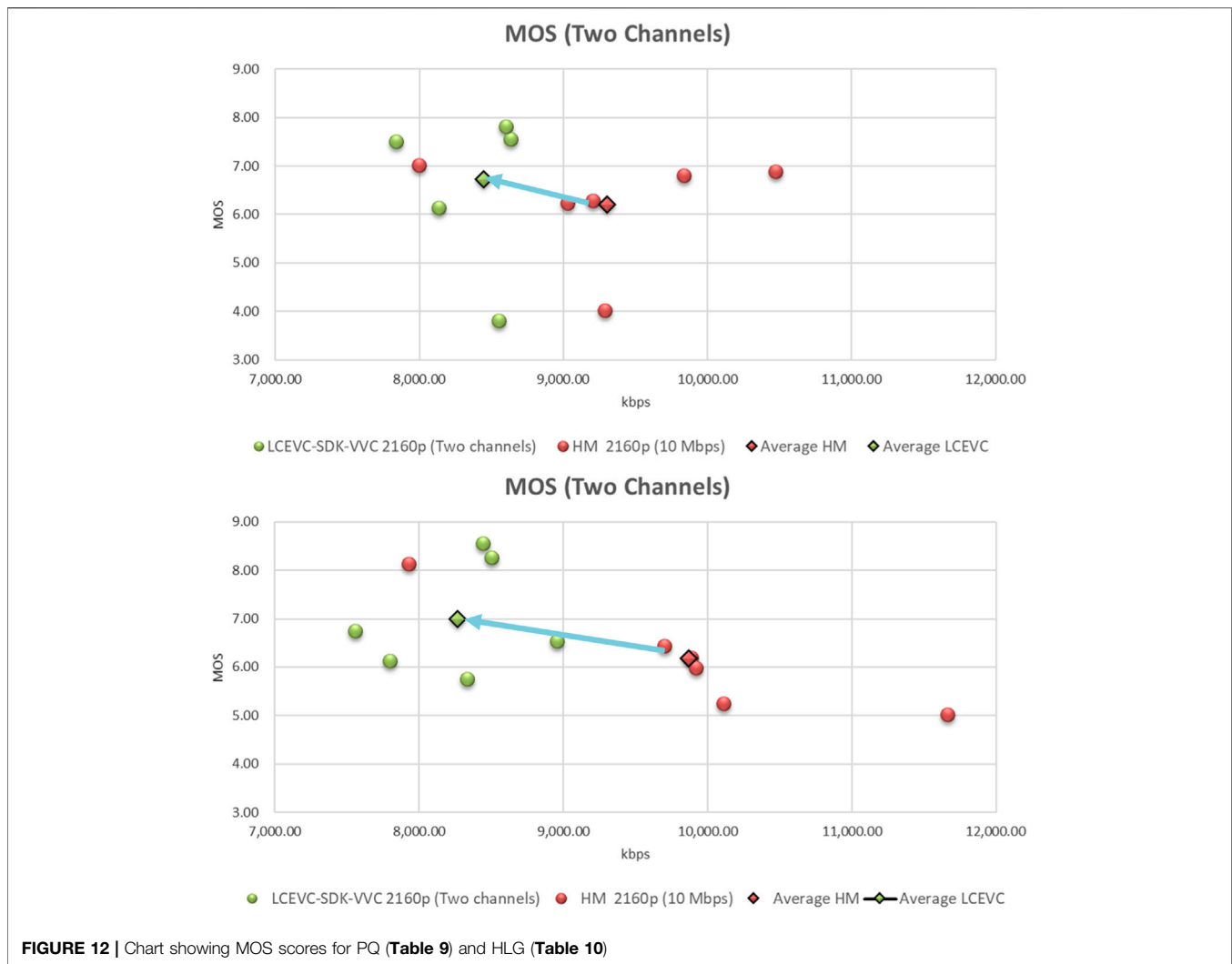


FIGURE 12 | Chart showing MOS scores for PQ (Table 9) and HLG (Table 10)

In light of the differences described above, LCEVC has the potential to deliver on the application scenarios envisaged in the TV 3.0 project, most notably when it comes to mitigate the encoding complexity required by standards such as H.266/VVC and the need for a quality of the video received comparable with the non-scalable case.

4 EVALUATION OF LCEVC ON THE TV 3.0 APPLICATION SCENARIOS

The CfP of the TV 3.0 project aims to test how new video coding technologies can improve the state-of-the-art ones currently used in broadcasting and in the application scenarios envisaged in the project. In the context of the TV 3.0 project, new video coding technologies were assumed to be the H.266/VVC and the LCEVC standard whilst legacy compression systems use the H.265/HEVC standard. Accordingly, the TV 3.0 project assessed the improvement of H.266/VVC over H.265/HEVC for single layer coding scenarios and, additionally, the coding efficiency

improvement in a scalable scenario when LCEVC is used to improve H.266/VVC/. With this in mind, the selected anchors for TV 3.0 CfP used always H.265/HEVC and in the scalable scenario the base line codec was H.266/VVC. LCEVC has been then tested using H.266/VVC as base layer video codec.

In the following sections we report a series of results and analyses for LCEVC tested under different conditions. Besides the results following the conditions set out in the CfP Phase 2 (Section 4.1.1), we also provide herein several additional results, specifically testing LCEVC in capacity-constrained channels (Section 4.1.2), in conjunction with Harmonic Dynamic Resolution Encoding (Section 4.1.3) and finally using x265 as a base codec (Section 4.1.4).

The tests below should enable evaluation of LCEVC and its unique features and benefits in a more complete fashion and under multiple testing conditions.

4.1 Test Set 1: LCEVC Under CfP Conditions

In this test we assessed LCEVC enhancing H.266/VVC using H.265/HEVC as anchor. In particular, for H.265/HEVC, the HM

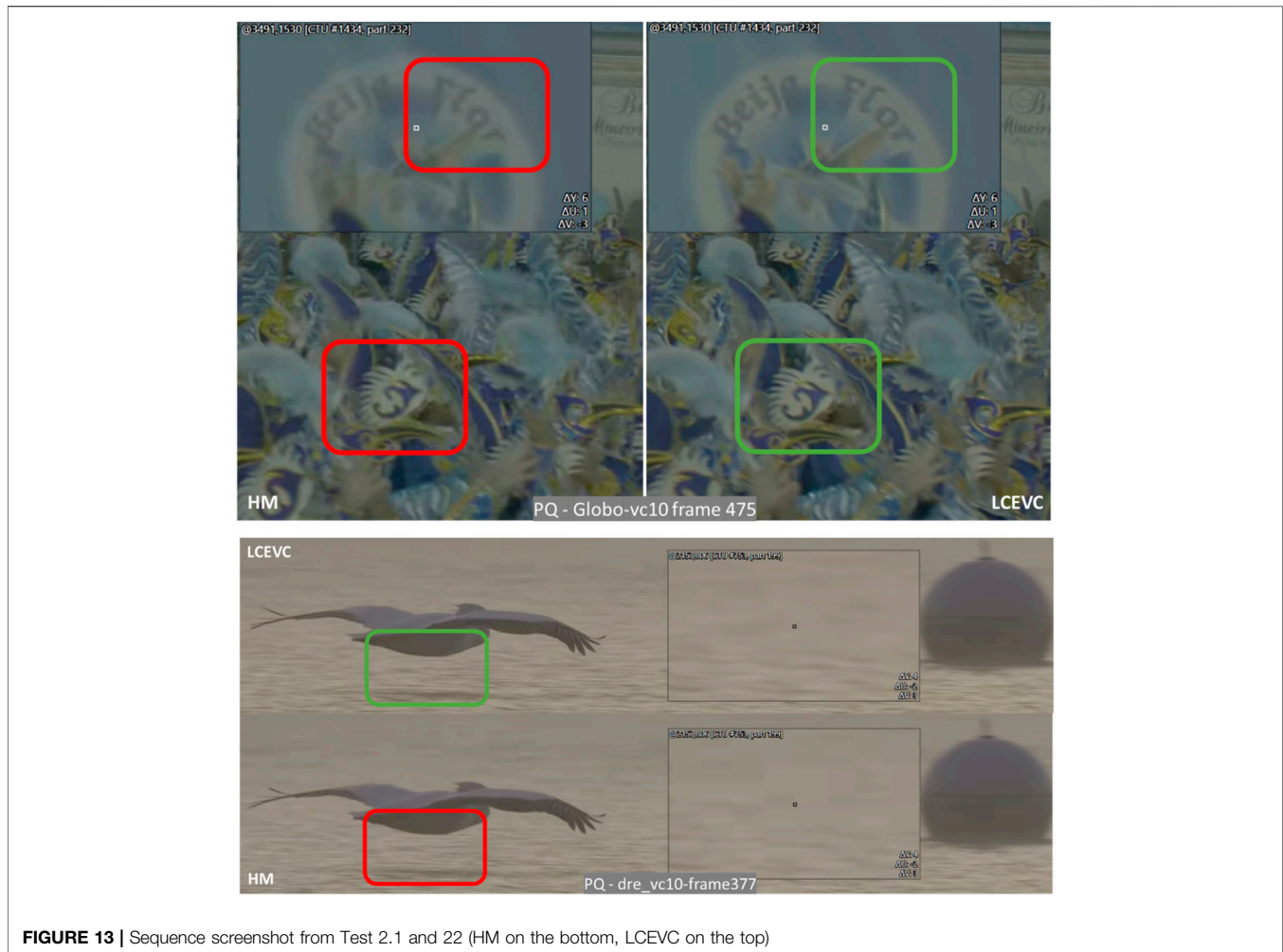


FIGURE 13 | Sequence screenshot from Test 2.1 and 22 (HM on the bottom, LCEVC on the top)

16.22 software was used using the configuration specified in the Annex A of the CfP document (Call for Proposal Phase 2, 2021). For LCEVC, we used the V-Nova LCEVC SDK and the VVenC by HHI² as the base layer.

Note that, for the purpose of this experiment, the V-Nova LCEVC SDK has been used in test-model mode, using fixed quantization and de-activating any pre-processing or post-processing optimizations. Using rate control, pre/post-processing and other tools is expected to provide even better results than those reported here.

We performed two different tests and then provided a comparative analysis between these two tests to highlight some of the key characteristics of LCEVC.

The set of sequences used in the TV 3.0 project tests are all HDR, with both PQ and HLG transfer characteristics. The test content is a mix of high motion, sharp details and highly textured areas. The figure below depicts the screenshots of the sequences used.

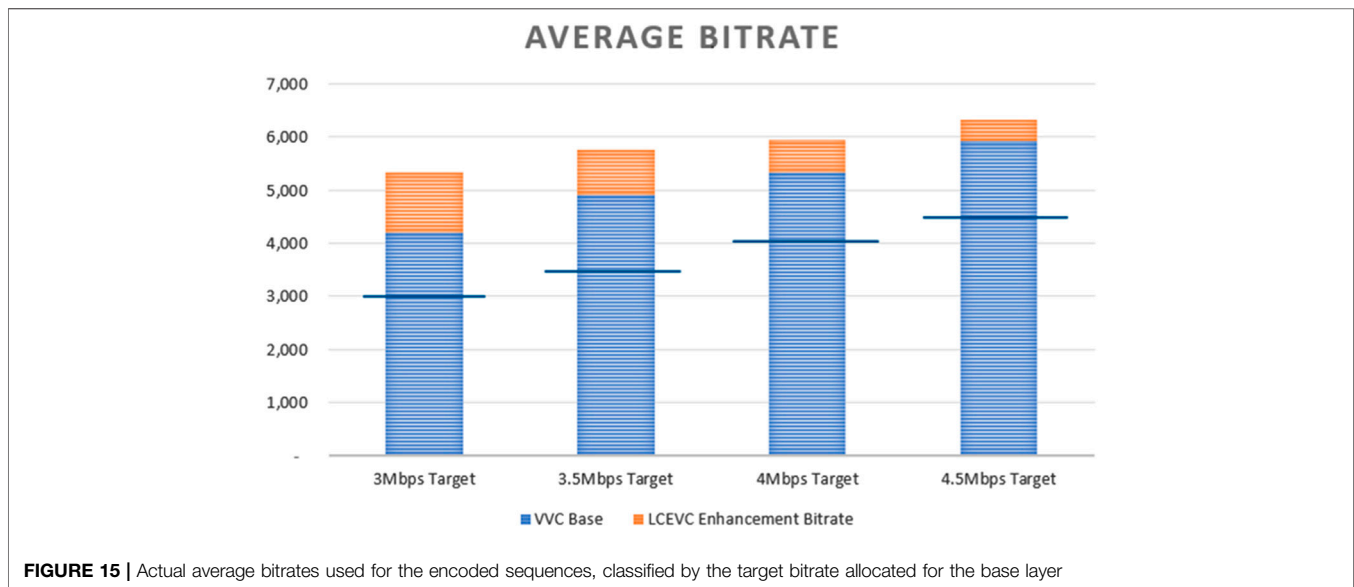
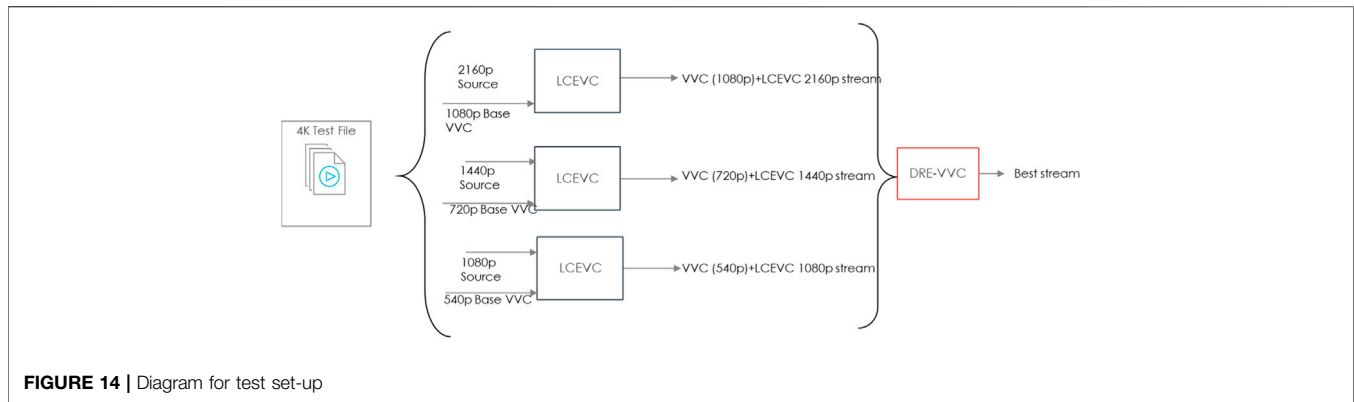
4.1.1 Configurations

4.1.2 LCEVC Enhancing H.266/VVC—Configuration A

For the target bitrates, the following Quantization Parameter (QP) values have been selected:

- For the full resolution anchors, we have used the overall target bitrate as specified in (Call for Proposal Phase 2, 2021) (i.e., TC1.1.4, TC1.2.4, TC11.4 and TC11.5 from the CfP), namely:
 - Rate 4 = 2160p, HM 16.22, QP = 22;
 - Rate 3 = 2160p, HM 16.22, QP = 27;
 - Rate 2 = 2160p, HM 16.22, QP = 32;
 - Rate 1 = 2160p, HM 16.22, QP = 37;
- for the Base Layer (BL), the target bitrate is a fixed percentage of the full-resolution encoded sequences, namely:
 - BL-R4 = ca. 80% of Rate 4;
 - BL-R3 = ca. 85% of Rate 3;
 - BL-R2 = ca. 90% of Rate 2;
 - BL-R1 = ca. 95% of Rate 1.

²<https://github.com/fraunhoferhhi/vvenc>



Although not necessarily optimal, the proportions specified above for the Base Layer coding rate are more aligned with those used in the MPEG Verification Test (Verification Test Report on the Compression Performance of Low Complexity Enhancement Video Coding, 2021), and are closer to the appropriate proportions used by LCEVC when encoding the enhancement layers. Because of the enhancement nature of the layers above the one, such a proportions results in good visual quality (see **Section 4.4** for further details). It is worth noting that the spread of the total bitrate among the base and enhancement layers is central for rate control purposes in practical implementation of LCEVC. However, given that the TV 3.0 project target the assessment of the shear coding advantage brought by the LCEVC technology, a more naive and fixed rate allocation across layers was used.

4.1.3 LCEVC Enhancing H.266/VVC—Configuration B

For the target bitrates, the following QP values have been selected:

- For the full resolution anchors, we have used the overall target bitrate as specified in (Call for Proposal Phase 2, 2021)

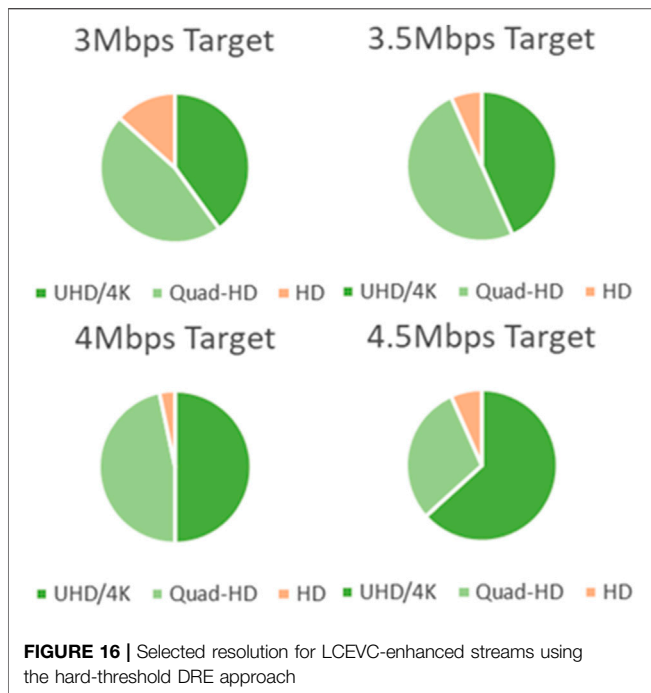
(TC1.1.4, TC1.2.4, TC11.4 and TC11.5 from the CfP) namely:

- Rate 4 = 2160p, HM 16.22, QP = 22;
- Rate 3 = 2160p, HM 16.22, QP = 27;
- Rate 2 = 2160p, HM 16.22, QP = 32;
- Rate 1 = 2160p, HM 16.22, QP = 37;
- for the Base Layer (BL), we have used the target bitrate as specified in (Call for Proposal Phase 2, 2021), namely:
 - BL-R4 = 1080p, HM 16.22, QP = 22;
 - BL-R3 = 1080p, HM 16.22, QP = 27;
 - BL-R2 = 1080p, HM 16.22, QP = 32;
 - BL-R1 = 1080p, HM 16.22, QP = 37

4.1.4 Subjective Evaluation Results

Formal subjective evaluations have been conducted for the tests specified in **Section 4.1**. The tests were conducted by GBTech laboratories and VABTech laboratories under the following conditions:

- Test Method: DSIS (11 grades impairment scale) according to ITU-R BT.500



- Naïve viewers: 30 (14 female, 16 males) all screened for visual acuity (Snellen chart) and colour blindness (Ishihara tables)
- Display: LG 65" CX6LA (3840 × 2160) @ 2H viewing distance
- Post-test viewers screening (Pearson correlation)
- 12 test sessions, each 12 min long

- Each session included a stabilization phase and the reference vs. reference test for all source sequences
- Careful distribution of quality levels across each test session
- Results of the assessment provided with MOS and CI values

In **Tables 1–4** we report the BD-rates computed using the MOS scores for both PQ And HLG material including the MOS scores and the confidence intervals (CI)

As it may be seen, LCEVC enhancing H.266/VVC provides an average bitrate saving of between 52% and 54% for both PQ and HLG sequences when Configuration A is used, and an average bitrate saving of almost 50% for HLG sequences and 36% for PQ sequences when Configuration B is used instead.

In **Figure 6** we include a couple of exemplary charts - one for PQ (clip SBT_vc04-dre) and one for HLG (clip SBT_vc10_globo) - to illustrate the typical rate distortion curves for the subjective formal assessments.

4.1.5 Objective Metrics

In **Tables 5, 6**, we report the Bjøntegaard Delta (BD) on rate (BD-rate) savings for Configuration A computed based on a selection of the objective quality metrics used in (Call for Proposal Phase 2, 2021) (i.e. VMAF, MS-SSIM and PSNR).

It is important to note that objective metrics based on MSE, such as PSNR, are known to be an unreliable predictor of visual quality (see for example (Wang and Bovik, 2009), (Lin et al., 2003) and (On learning based video quality metrics, 2021)), and this is particularly true for multi-resolution schemes such as LCEVC (see for example (Lcevc, 2020b)). Therefore, it is expected that BD-rates based on PSNR are not favorable for



TABLE 11 | BD-rate savings for LCEVC-enhanced x265 over x265 (HLG).

Sequence	VMAF	MS-SSIM-Y	PSNR
vc04-lcevc-1	-14.90%	-18.29%	14.18%
vc04-lcevc-2	-11.59%	-13.95%	-0.59%
vc04-avs3	-19.20%	-14.34%	-6.67%
vc04-dre	-15.72%	-26.37%	-7.56%
vc04-globo	-1.43%	-15.24%	-6.02%
vc04-vcv	-13.61%	-38.28%	-2.81%
Average	-12.80%	-21.08%	-1.58%

LCEVC, especially in the range of non visually lossless coding¹.

Figure 7 shows an illustrative example of the phenomena at play when small quantization errors in mid-low frequencies (quite visible due to block correlation) are substituted with more accurate mid-low frequencies and larger quantization errors in fine details. Due to the quadratic nature of Mean Squared Error, in lossy range the left picture is scored better by PSNR despite obviously worse visual quality and a higher number of errors than the right picture.

Although perceptual-based objective metrics such MS-SSIM and VMAF are better predictors of visual quality - with VMAF being probably the best to date - they are still not fully correlated with visual quality, which for LCEVC tends to outperform even those metrics.

Accordingly, we provide the results of the formal subjective assessments to reliably evaluate the comparative quality of the encoded sequences and use the objective metrics purely for cross-checking the sequences.

4.1.6 Comparison Between Configuration A and Configuration B

In this section we have compared the results of Configuration A and Configuration B.

In both tests the overall target bitrate (i.e., the bitrate for the base layer *plus* the bitrate for the enhancement layer) is substantially the same. However, the percentage of the overall target bitrate allocated to the base layer (and, consequently, to the enhancement layer) is different.

Specifically, in Configuration A more bitrate is allocated to the base layer than to the enhancement layer (on average, between 82% and 85% to the base layer and between 15% and 18% to the enhancement), whilst in Configuration B more bitrate is allocated to the enhancement layer than to the base layer (on average, between 38% and 41% to the base layer and between 59% and 62% to the enhancement—more akin to what would be used for a traditional scalable scheme).

As it may be seen from the results, Configuration A provides a much better performance as it is better suited for how LCEVC works.

Contrary to typical scalable encoding schemes, LCEVC is very efficient at predicting from the base layer and at encoding the

residuals in the enhancement layer, therefore requiring only a small percentage of an overall bitrate. The larger remaining portion of the overall bitrate can be used to encode a base layer at a much higher quality.

The overall effect is that with LCEVC used with appropriate bitrate allocation not only the quality of the full resolution is higher, but also the quality of the base layer is much higher than if a scalable-like bitrate allocation were used. This means that—using the same overall bitrate—it is possible to provide both a higher-quality higher resolution service (e.g., 4k/UHD) and a higher-quality lower resolution service (e.g., 1080p).

In Figures 8–10 we show a selection of graphs comparing Configuration A (LCEVC-like proportion—green tones) and Configuration B (scalable-like proportion—red tones) for the base layer (HD) and the LCEVC-enhanced full resolution (UHD).

As it may be seen, both base layers and LCEVC-enhanced full resolution in Configuration A are significantly better for all metrics and at all rate points.

4.2 Test Set 2: LCEVC Under Capacity-Constrained Channels

In these tests we have assessed LCEVC when enhancing H.266/VVC using H.265/HEVC as an anchor. In particular, for H.265/HEVC, the HM 16.22 software was used. For LCEVC, we used the V-Nova LCEVC SDK using the VVenC by HHI² as the base layer.

Note that in all these tests, VVenC was set in “fast” pre-set to simulate as close as possible real-time transmission. Using VVenC in a slower pre-set, which would take advantage by the processing savings generated by LCEVC, may further improve the coding efficiency^{3,4}.

These tests are aimed at providing data relative to two commercially relevant scenarios. In the first scenario (Section 4.2.1), both base layer and enhancement layer are transmitted within a single 6 MHz/5 Mbps channel. In the second scenario (Section 4.2.2), the base layer is transmitted in a first 6 MHz/5 Mbps channel whilst the enhancement layer is transmitted in a second 6 MHz/5 Mbps channel.

4.2.1 Test 2.1: LCEVC and Base Layer Within a Single 6 MHz/5 Mbps Channel

In this scenario, we compared HM as an anchor and LCEVC-enhanced H.266/VVC (i.e., base layer and enhancement layer), both transmitted over a single 6MHz/5 Mbps channel. As such, the target bitrate for each encoded stream is 5 Mbps. The actual bitrate achieved is typically within 10% of the target bitrate, mainly due to the HM quantization and VVenC rate control.

As done with Test 1, in Tables 7, 8 are reporting the results of formal subjective assessments performed on these sequences and conducted at the same time and with the same settings as reported in Section 4.1.5 for Test 1.

¹Note that—aside from being poorly correlated with subjective quality—relative MSE performance of LCEVC-enhanced coding vs. native coding may vary significantly across lossy vs. near-lossless coding ranges, often with LCEVC achieving better PSNR for very lossy bitrate ranges (where errors in low frequencies are large enough to matter to MSE), then worse for lossy ranges, then better again for near-lossless ranges

²<https://github.com/fraunhoferhhi/vvenc>

³<https://www.v-nova.com/lcevc-enhanced-video/lcevc-video-sdk/>

⁴<https://www.xilinx.com/products/acceleration-solutions/v-nova-lcevc.html>

In **Figure 11** we also represent graphically the same data reported in the above tables.

As it can be seen LCEVC-enhancing H.266/VVC provides a significant improvement in visual quality over the HM anchor—over 1.3 MOS points of improvement on average—whilst staying very close to the 5 Mbps target bitrate.

4.2.2 Test 2.2: LCEVC and Base Layer in Two Separate 6 MHz/5 Mbps Channels

In this scenario, we again compared HM as an anchor and LCEVC-enhanced H.266/VVC.

However, in this scenario H.266/VVC as a base layer is transmitted over one 6MHz/5 Mbps channel, and the LCEVC enhancement layer is transmitted over another, separate 6 MHz/5 Mbps channel. As for the HM anchor the target bitrate is 10 Mbps. Of course, this is a purely “notional” anchor, as in the case of two separate 6MHz/5 Mbps channels it would not be possible to transmit a 10 Mbps HM-encoded sequence.

The actual bitrate achieved for HM is typically within 5%–10% of the target bitrate, mainly due to the HM quantization. For the LCEVC-enhanced sequences, the actual bitrate is typically lower than the 10 Mbps, and on average 15%–18% less than the target bitrate. The reason for this is that the LCEVC-enhanced sequences, the base layer component is encoded with VVenC targeting 5 Mbps, whilst the enhancement layer component is encoded with LCEVC using significantly less than the target 5 Mbps due to the ability of LCEVC to efficiently compress the enhancement component without needing too much bitrate, as discussed and demonstrated in **Section 4.1.6**.

As done with Test 1, in **Tables 9, 10** we are reporting the results of formal subjective assessments performed on these sequences and conducted at the same time and with the same settings as reported in **Section 4.1.5** for Test 1.

As it may be seen from **Figure 12**, LCEVC-enhancing H.266/VVC provides a significant improvement in visual quality over the “notional” HM anchor—between 0.5 and almost 0.9 MOS points of improvement on average—whilst also requiring much less than the 10 Mbps target bitrate and, crucially, being able to leverage two separate 6MHz/5 Mbps channels.

4.2.3 Exemplary Screenshots

In this section, we report some exemplary screenshots, mainly from the Test 2.2 as reported in **Section 4.2.2**, to provide some evidence of the improvements associated with the LCEVC-enhanced sequences when compared to their HM counterpart.

In general, it is possible to notice that LCEVC provides a generally improved sharpness and accuracy in reproducing details as demonstrated in **Figure 13**.

4.3 Tests Set 3: LCEVC in Combination With Dynamic Resolution Encoding

In this test we have assessed LCEVC-enhancing H.266/VVC with the Dynamic Resolution Encoding (DRE) technology. In particular, the test was set-up as illustrated in **Figure 14**.

Starting from the same source material at 2160p (“Regatta” sequence from Harmonic), we have created three different resolutions by downsampling the original source to 1080p, 720p and 540p resolution, encoded the downsampled sequences with VVenC in “fast” preset to generate encoded base layers, and then used LCEVC to enhance each encoded base layer to three different higher resolutions 2160p, 1440p and 1080p, respectively.

The overall target bitrate for each resolution is 5 Mbps - in order to simulate a scenario similar to that in Test 2.1 (see **Section 4.2.1**). The base layer was encoded using VVenC with different target bitrates for the base—namely, 3, 3.5, 4 and 4.5 Mbps—with the LCEVC enhancement using the remaining bitrate. In **Figure 15** we report the actual average bitrate used for each encoded set—as it can be seen, the H.266/VVC base layer typically overshoots compared to the base target bitrate, with LCEVC then adding the typical 10%–20% bitrate on top.

The final LCEVC-enhanced sequences were then split into 2-s chunks, and for each chunk the three resolutions (2160p, 1440p and 1080p) were compared using VMAF to select the best resolution.

In particular, two different algorithms were used to select the best resolution:

- 1) Hard-threshold: In this scenario, the selected resolution for each 2-s chunk was the one with the highest VMAF score for that chunk.
- 2) Soft-threshold: In this scenario, the selected resolution for each 2-s chunk was either the one with the highest VMAF score for that chunk or, in case there was a higher resolution with a VMAF score within 0.2 points from the highest VMAF score, this higher resolution.

In **Figures 16, 17** we show the resolution selected according to the two different algorithms provided below, classified by the target bitrate allocated to the base layer.

As it may be seen from the results, the selected LCEVC-enhanced 2-s chunk is in most cases the highest possible resolution available, namely 2160p, with the second highest resolution (i.e., 1440p) used in certain cases. Only in a handful of chunks the lowest resolution (1080p) is used.

This test shows that the combination of LCEVC-enhanced streams and DRE technology can be beneficial as it allows to select the highest possible resolution in most cases and, when useful to have a better quality, fallback to the second highest resolution.

4.4 Tests Set 4: LCEVC With x265

In this section we report a further test performed using x265 as an anchor and LCEVC-enhanced x265 as a target. This test is done primarily to demonstrate that, although Test 1, test 2 and Test 3 have been run using H.266/VVC as the coding scheme for the base layer, LCEVC is agnostic to the specific coding scheme used for the base layer and can provide improvement enhancing other codecs.

In this test, 8 different target bitrates were tested for each sequence (namely, 750, 1500, 3000, 6000, 10,000, 15,000, 20,000 and 30,000 kbps) and various objective metrics were measured. For these tests, no formal subjective assessment was made. The BD-rates savings for LCEVC-enhanced x265 over x265 are reported in **Table 11**.

As it can be seen, LCEVC can provide a significant improvement when enhancing H.265/HEVC. These results are consistent with the results provided in multiple other tests on LCEVC enhancing H.265/HEVC, for example the official MPEG verification tests reported in (Verification Test Report on the Compression Performance of Low Complexity Enhancement Video Coding, 2021), typically showing -30–40% BD-rate MOS gains and slightly lower gains on objective metrics. Since these were the first tests executed on HDR sequences, results may also improve with further calibrations aimed at HDR.

5 CONCLUSION

In this paper, the LCEVC proposal to the TV-3.0 CfP has been described. The flexibility and the efficiency of LCEVC has been discussed showing that it addresses all the TV-3.0 use-cases and requirements in an efficient manner. An overview of the toolset for the LCEVC standard has been provided along with the benefits this innovative scheme can provide as an enhancement codec. The combination with multiple technology and base codec has been described showing the LCEVC readiness for commercial deployment in the context of TV-3.0 in Brazil.

REFERENCES

- Av1, AV1. Available at: <https://aomediacodec.github.io/av1-spec/av1-spec.pdf>
 Call for proposal phase 2 (2021). Testing and evaluation: TV 3.0 project. Available at: <https://forumsbvtv.org.br/wp-content/uploads/2020/07/SBTVDTV-3-0-CfP.pdf>.
 Evc. ISO/IEC 23094-1, information technology — general video coding — Part 1: Essential video coding.
 H.264/Avc. ISO/IEC 14496-5/AMD 6, information technology — coding of audiovisual objects — Part 5: Reference software — amendment 6: Advanced video coding (AVC) and high efficiency advanced audio coding (HE AAC) reference software
 H.265/Hevc. ISO/IEC 23008-5, information technology — high efficiency coding and media delivery in heterogeneous environments — Part 5: Reference software for high efficiency video coding
 H.266/Vvc. ISO/IEC 23090-16, Information technology – coded representation of immersive media – Part 16: Reference software for versatile video coding
 Iso/Iec Dis, (2022), 23094-3 – joint technical committee ISO/IEC JTC 1, information technology – general video coding – Part 3: Conformance and reference software for low complexity enhancement video coding,
 Jiménez-Moreno, A., Ciccarelli, L., Clucas, R., and Ferrara, S. (2022). “HDR video coding with MPEG-5 LCEVC,” in MHV '22: Proceedings of the 1st Conference on Mile-High Video, 17 March 2022. Available at: <https://www.lcevc.org/wp-content/uploads/Evaluation-of-MPEG-5-Part-2-LCEVC-for-Live-Gaming-Video.pdf>.
 Lcevc (2020a). Avc – incredible 28% gain at 3x speed – OTTVerse – OTTVerse. Available at: <https://ottverse.com/lcevc-vs-avc-using-ffmpeg/>.
 Lcevc (2020b) Experimental Results of LCEVC versus conventional coding methods” ISO/IEC JTC1/SC 29/WG4 m53806,
 Lin, W., Li, D., and Xue, P. (2003). “Discriminative analysis of pixel difference towards picture quality prediction,” in Proceedings 2003 International Conference on Image Processing (Cat. No. 03CH37429, Barcelona, Spain, 14-17 September 2003, 193.
 Martini, M. G.. *Evaluation of MPEG-5 part 2 (LCEVC) for live gaming video streaming applications nabajeet barman, stephen shmidt, saman zadtootaghaj* in MHV '22: Proceedings of the 1st Conference on

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

ACKNOWLEDGMENTS

The authors of this report would like to thank Xavier Ducloux and Thierry Fautier of Harmonic for their support and help in conducting Test 3, GBTech Laboratories and VABTech Laboratories for conducting the formal subjective assessments reported in Test 1 and Test 2, and all the people in V-Nova who have helped in preparing test material and software, and designing, generating and analysing the results reported in this document.

Mile-High Video, Kingston University, London, UK. Available at: https://www.lcevc.org/wp-content/uploads/MHV22_paper_HDR_video_coding_with_MPEG5_LCEVC_camera_ready.pdf.

Mpeg. Verification test report on the compression performance of LCEVC. Available at: <https://www.lcevc.org/wp-content/uploads/MPEG-Verification-Test-Report-on-the-Compression-Performance-of-LCEVC-Meeting-MPEG-134-May-2021.pdf>.

On learning based video quality metrics (2021), ISO/IEC JTC 1/SC 29/AG 5 m56636.

V-Nova (2021). LCEVC licensing terms announced for entertainment video services. Available at: <https://www.v-nova.com/press-releases/v-nova-lcevc-licensing-terms-announced-for-entertainment-video-services/>.

Verification test report on the compression performance of low complexity enhancement video coding (2021). ISO/IEC jtc 1/SC 29/WG 04 N0076 (20173). Available at: https://www.mpegstandards.org/wp-content/uploads/mpeg_meetings/134_OnLine/w20173.zip.

Wang, Z., and Bovik, A. (2009). Mean squared error: Love it or leave it? *IEEE Signal Process. Mag.* 26, 98–117. doi:10.1109/MSP.2008.930649

Conflict of Interest: Authors LC, FM and SF are employed by V-Nova.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Ciccarelli, Ferrara and Maurer. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.