Verification Test of the Low Complexity Enhancement Video Coding (LCEVC) Standard

Stefano Battista⬢, Guido Meardi, Simone Ferrara⬢, Lorenzo Ciccarelli⬢, Florian Maurer⬢,
Massimo Conti⬢, *Member, IEEE*, and Simone Orcioni⬢, *Senior Member, IEEE*

Abstract—This paper reports the methodology and results of the test campaign performed in the context of the development of the Low Complexity Enhancement Video Coding (LCEVC) ISO standard. LCEVC is a video coding technology that can be applied in conjunction to any other single-layer video coding technology to re-encode the residuals between the original video and its compressed representation. The first part of the paper describes the methodology for the verification tests: Requirements, Test Conditions, Test Sequences, Test Methods (Metrics); the second one reports the results of the objective and subjective tests carried out applying such methodology. The bitrate savings achieved with LCEVC when used in conjunction with four video codecs (AVC, HEVC, EVC, VVC) are reported.

Index Terms—LCEVC, MPEG, video coding, low-complexity, enhancement, objective test, subjective test.

I. INTRODUCTION

THE Low Complexity Enhancement Video Coding (LCEVC) specification is an MPEG Video Coding standard developed between April 2018 and November 2021 by the ISO Working Group known as ISO/IEC JTC1 SC29/WG11 (MPEG) until June 2020 and currently as SC29/WG04 (MPEG Video Coding). The specification consists of ISO/IEC 23094-2 [3] (MPEG-5 Part 2, "Low complexity enhancement video coding") and ISO/IEC 23094-3 [4] (MPEG-5 Part 3, "Conformance and reference software for low complexity enhancement video coding").

LCEVC is designed to be an "enhancement," in the sense that it is a toolset for encoding the residual differences between the original video and a compressed video, encoded with any existing video codec, like Advanced Video Coding (AVC) / H.264 [5], [6], High Efficiency Video Coding (HEVC) / H.265 [7], [8], Essential Video Coding (EVC) [9], [10], and Versatile Video Coding (VVC) / H.266 [11], [12] or future

video codecs. In other words, LCEVC encodes the coding errors between an original video sequence and its compressed representation by any video codec, working on the differences in the pixel domain, so without any dependency from the base codec.

LCEVC is also designed to be "low complexity," in the sense that the toolset is composed of typical video coding processing blocks (temporal prediction, transformation, quantization and entropy coding), all of them intended for implementation with low complexity in terms of processing power and memory requirements.

Consequently, LCEVC is intended to be efficiently and effectively implemented in software via existing processing blocks in existing devices, such as Single Instruction Multiple Data (SIMD) processors and Graphics Processing Unit (GPU) processors. LCEVC achieves a balance between low complexity and high rate-distortion performance.

The goal of any video compression scheme is the optimization of the Rate Distortion (RD) function, i.e. minimization of the bitrate required to achieve a given video quality, or, symmetrically, the maximization of the video quality achieved for a given bitrate. As with other standards developed by ISO/MPEG (Moving Picture Experts Group) and ITU-T/VCEG (Video Coding Experts Group), the final phase of the standardization process is the quantitative verification of the compression performance of the codec, by means of the execution of tests to measure the objective and subjective video quality, using metrics unanimously adopted by the video coding research community [1], [2].

LCEVC has been compared to four single-layer video codecs developed by MPEG and VCEG, specifically AVC/H.264, HEVC/H.265, EVC and VVC/H.266, that represent a set of single-layer codecs of increasing compression efficiency and at the same time increasing computational complexity.

In literature, attention in evaluating LCEVC performance is recently spreading. An overview of the LCEVC specification and a preliminary comparison of coding efficiency for LCEVC used with AVC, HEVC, and VVC, in terms of PSNR, VMAF, and MOS are presented in [13]. The paper also provides test results on computational complexity and a study of the correlation of bitrate savings for LCEVC with the temporal and spatial complexity of the video sequences, reporting low correlation with temporal activity and high correlation with spatial activity. In [14] the authors report a comparison of LCEVC to AVC (in its implementation x264) and HEVC (in its implementation x265), when applied to High Dynamic

Range (HDR) video sequences. The paper shows that LCEVC is capable of producing an HDR quality video starting from an 8-bit x264 base layer, thus providing HDR video in streaming applications. Enhancing 10-bit base layers (e.g, x265), the authors state a significant improvement of LCEVC when compared with the equivalent native implementation of the base layer. A comparison of LCEVC enhancing AVC (in its implementation x264) and HEVC (in its implementation x265), in the context of Live Gaming Video Streaming applications is reported in [15]. A bitrate saving of about 40% with respect the base codecs at full resolution for the VMAF metric, and a gain when using AVC and a loss when using HEVC for the PSNR metric are achieved. About the MOS metric, the authors report a discussion of LCEVC bitrate savings as a function of bitrate, suggesting the need for a methodology for estimating a metric, averaged over the video sequences, but calculated as a function of bitrate. Such a methodology is also proposed and applied in this work.

In [16] the authors provide a description of the LCEVC specification and an analysis of the individual coding tools comprised in LCEVC, highlighting the low complexity and the innovative aspects of such tools. If [16] is an overview of LCEVC, the current paper provides a synthesis of the methodologies used and the results obtained in the LCEVC verification tests. As noted above, the verification test is the final phase of the development of an MPEG Video coding standard, since it is crucial to confirm the performances of the new video coding technology with respect to the objectives set forth at the beginning of the standard development activity. The verification tests are performed within the scope of the MPEG standardization group, and cross checked by the other subject matter experts. Although the methodology adopted to perform the verification tests is a consolidated practice established in MPEG and VCEG, the original aspect of this work is that it reports the first in depth campaign of tests for the new coding technology specified by LCEVC [1], [2]. Besides, the verification involves a comparison with the RD performance of well established (AVC and HEVC) and emerging (EVC and VVC) video standards. Finally, a new methodology to model the RD performance of a generic video codec as a mean polynomial model has been applied, to have an insight on the compression efficiency over a range of bitrates rather than scalar values which represent the integral of the differences of RD curves.

The paper is organized as follows. Section II summarizes the requirements set at the beginning of the LCEVC development and used at the end to verify its performances. Sections III to V describe the rationale under the test conditions (which codecs to test and in what configuration), and the choice of the video test sequences, and finally the objective and subjective methods adopted for the verification tests. Section VI reports the results of LCEVC in terms of Rate Distortion (RD) performance, with a subset of PSNR and VMAF results [1], plus the complete set of MOS results [2]. Section VI also reports an alternative methodology for the analysis of the same data to study the RD performance as a function of Bit Rate (BR) [17]. Finally, Section VII summarizes the conclusions of the paper.

II. REQUIREMENTS

Following the 30 years of experience and consolidated practices in MPEG, the goals of the developments of LCEVC were set at the beginning of the project, in April 2018.

The basic requirements to be met by LCEVC were the following:

1) Compression: When used as enhancement to a base codec, e.g. AVC, the compression efficiency of the aggregate base plus LCEVC bitstream shall be significantly higher than the same base codec used at full resolution.

2) Complexity: The encoding and decoding complexity for the aggregate base plus LCEVC video shall be comparable to the encoding and decoding complexity of the base video at full resolution.

Such requirements, set at the beginning of the standardization process, were subject to verification after the specification reached the stage of Final Draft International Standard, with a campaign of objective and subjective testing performed by the experts of SC29/WG04 (MPEG Video Coding Working Group, chaired by Prof. Lu Yu) and SC29/AG05 (MPEG Visual Quality Assessment Advisory Group, chaired by Dr. Mathias Wien), and finalized in April 2021.

The first requirement, on the performance of the enhancement when comparing with a single-layer of the same codec, is intended to verify the performance of LCEVC in terms of "enhancement." The second requirement, on the encoding and decoding complexity, is intended to verify the requirement of "low complexity."

These are the two dimensions of innovation of LCEVC, since it has been designed not to replace existing single-layer codecs, but rather to perform as a Low Complexity Enhancement to any single-layer codec, maintaining the characteristic of being agnostic of the underlying base codec.

III. TEST CONDITIONS

For the verification of the requirements set for the LCEVC specification, four existing video coding standards specified by MPEG and VCEG were selected. Two codecs widely deployed in broadcasting and streaming, i.e. AVC/H.264, defined in 2003, and HEVC/H.265, defined in 2013. Two codecs standardized by MPEG in the same time frame of LCEVC, i.e. EVC, finalized in 2020, and VVC/H.266, finalized in 2021.

The verification test included two types of tests:

1) Requirements test. These tests were designed to verify the satisfaction of the LCEVC requirements, by comparing full resolution LCEVC-enhanced encoded sequences versus full resolution anchors encoded with the four selected native codecs.

2) Resolution Enhancement test. These tests were designed to verify that LCEVC has a better RD performance than unguided upsampling, by comparing full resolution LCEVC-enhanced encoded sequences versus quarter resolution (i.e. half width and half height) anchors encoded with the four selected native codecs and then upsampled using a fixed Lanczos filter.

The rationale for the second type of tests is that the simplest alternative to enhancing a codec with LCEVC is to just use the native codec at a lower resolution and let the end user
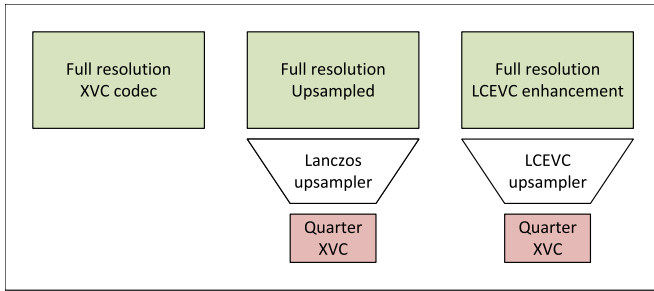
Fig. 1. Verification test conditions.

TABLE I
VERIFICATION TESTS VIDEO SEQUENCES

Codec	Class	Sequence	Resolution	Picture Rate
AVC	SDR-A	CatRobot	3840×2160	60
AVC	SDR-A	LupoPuppet	3840×2160	50
AVC	SDR-A	**DrivingLogo**	3840×2160	60
AVC	SDR-A	**BoxeLogo**	3840×2160	60
AVC	SDR-B	**TrafficLogo**	1920×1080	60
AVC	SDR-B	**Starcraft**	1920×1080	60
HEVC	SDR-A	CatRobot	3840×2160	60
HEVC	SDR-A	LupoPuppet	3840×2160	50
HEVC	SDR-A	**DrivingLogo**	3840×2160	60
HEVC	SDR-A	**BoxeLogo**	3840×2160	60
HEVC	SDR-B	**TrafficLogo**	1920×1080	60
HEVC	SDR-B	**Starcraft**	1920×1080	60
EVC	SDR-A	CatRobot	3840×2160	60
EVC	SDR-A	BarScene	3840×2160	60
EVC	SDR-A	**DrivingLogo**	3840×2160	60
EVC	SDR-A	**BoxeLogo**	3840×2160	60
EVC	SDR-B	**TrafficLogo**	1920×1080	60
EVC	SDR-B	**Starcraft**	1920×1080	60
VVC	SDR-A	Marathon	3840×2160	30
VVC	SDR-A	MountainBay	3840×2160	30
VVC	SDR-A	**DrivingLogo**	3840×2160	60
VVC	SDR-A	**BoxeLogo**	3840×2160	60
VVC	SDR-B	**TrafficLogo**	1920×1080	60
VVC	SDR-B	**Starcraft**	1920×1080	60

device upsample the quarter resolution encoded sequence to full resolution by means of unguided upsampling.

It is worth noting that the second requirement on Resolution Enhancement was not part of the original set of requirements, but was added during the development of the LCEVC specification, in response to investigations on the preliminary results presented in the MPEG Video Coding working group.

Thus, the comparison was performed among three test conditions:
- the single-layer encoding at full resolution, denoted by the label "Full,"
- the single-layer encoding at quarter resolution, followed by a fixed Lanczos upscaling, denoted as "Upsampled,"
- the multi-layer encoding at full resolution, using the same single-layer codec for the quarter resolution, and LCEVC for the enhancement, denoted as "LCEVC."

The three test conditions are graphically depicted in Fig. 1, showing the three conditions from left to right.

IV. TEST SEQUENCES

Since the verification tests were designed to include comparisons between full resolution encoded sequences and upsampled quarter resolution encoded sequences, the test set had to include sequences where the full resolution version has significant difference to the quarter resolution upsampled version, in terms of subjective quality, as formally measured using MOS as defined in Recommendation ITU-R BT.500 [18]. In particular, it was decided to include both relatively "smooth" sequences (where the expectation is that upsampled quarter resolution encodes have similar quality to full resolution encodes, independently of the codec compression efficiency) and "sharp" sequences with some high contrast details (where the expectation is that the difference between full resolution and upsampled quarter resolution encoded sequences can be perceived by a non-expert viewer).

The main characteristic of the "smooth" sequences is that they contain a lower amount of energy in the high spatial frequency, and conversely the "sharp" sequences contain a higher amount of energy at high frequency.

Taking into account such considerations, the following mix of sequences was selected for each tested codec:
- two Ultra High Definition (UHD) at 3840×2160 sequences without sharp details (smooth);
- two Ultra High Definition (UHD) at 3840×2160 sequences with some sharp details (medium);
- two High Definition (HD) at 1920×1080 sequences with many sharp details (sharp).

The "smooth" sequences are different for the four base codecs, since for EVC and VVC sequences from the respective test set were used. The "medium" and "sharp" sequences are the same for all base codecs, AVC, HEVC, EVC, and VVC, and they are highlighted in bold in Tab. I. An image from each of the test sequences is presented in Fig. 2.

Specifically, the two "medium" sequences (DrivingLogo and BoxeLogo) consist of natural video with an overlay of graphical content on about 25% of the picture area. while the two "sharp" sequences (TrafficLogo, Starcraft) consist for TrafficLogo of natural video with an overlay of 25% graphical content and for Starcraft completely of graphical content from a video game.

Tab. I summarizes the choice of video sequences selected for the LCEVC verification tests. All original video sequences are represented with 10 bits per pixel, except for the tests with AVC, in which the representation is with 8 bits per pixel.

All test sequences have a duration of 10 seconds, that is 300 pictures for 30 fps, 500 pictures for 50 fps, and 600 pictures for 60 fps.

V. TEST METHODS

Concerning the objective and subjective metrics adopted for the LCEVC Verification Test, the most widely used objective metric is the Peak Signal to Noise Ratio (PSNR), and the most widely used subjective metric is the Mean Opinion Score (MOS), as defined in ITU-R BT.500 DSIS MOS [18]. A third metric, Video Multi-method Assessment Fusion (VMAF) [19], recently developed from studies on objective evaluation of subjective quality, was also included in the test campaign, to have an intermediate tool for evaluation. The three metrics are briefly described in the following subsections.

After performing the objective and subjective tests, the relative performance of the different methods under test,

Fig. 2.   Verification test sequences (smooth, medium, sharp).

in terms of bitrate saving for the same quality range, can be computed applying the Bjontegaard Delta (BD) rate methodology, described in Sec. V-D.

A. Peak Signal to Noise Ratio

PSNR is a purely statistical metric, and computes an average of the Mean Square Error between the original sample values and the respective encoded values. PSNR of the overall YUV video sequence is a linear combination of PSNR for a single component (Luminance or Chrominance [17]:

$$PSNR_{YUV} = (6\,PSNR_Y + PSNR_U + PSNR_V)\,/8. \qquad (1)$$

with PSNR defined as

$$PSNR = 10 \log \frac{\left(2^B - 1\right)^2}{MSE}, \qquad (2)$$

$$MSE = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} (\text{original}(i,j) - \text{coded}(i,j))^2, \quad (3)$$

where
- M is the number of pixel rows (height);
- N is the number of pixel columns (width);
- $\text{original}(i,j)$ are the Y, U and V values of the pixel of the original pictures at position (i,j);
- $\text{coded}(i,j)$ are the Y, U and V values of the pixel of the coded pictures at position (i,j);
- B is the number of bits per pixel of Luminance (Y) and Chrominance (U, V).

B. Video Multi-Method Assessment Fusion

VMAF [19], on the contrary of PSNR, uses a combination of objective metrics to better model the results of subjective evaluation of video quality. The three elementry metrics combined in VMAF are:
- **Visual Information Fidelity (VIF)**. VIF is based on the concept that quality is complementary to the measure of

information fidelity loss. The VIF score is measured as a loss of fidelity combining four scales. VMAF uses a modified version of VIF, where the loss of fidelity in each of the four scales is included as an elementary metric.
- **Detail Loss Metric (DLM)**. DLM is based on the concept of separately measuring the loss of details which affects the content visibility, and the redundant impairment which distracts the viewer attention. The original metric described in [20] combines both DLM and Additive Impairment Measure to compute a final score. while VMAF uses only DLM as an elementary metric.
- **Motion**. Motion is a simple measure of the temporal difference between adjacent pictures, obtained calculating the average absolute pixel difference for the luminance component.

C. Mean Opinion Score

The subjective metric MOS is specified in Recommendation ITU-R BT.500 [18]. The method applies a Double-Stimulus Impairment Scale (DSIS), which consists in presenting a test (compressed) video sequence following the corresponding reference (uncompressed) video sequence to a non-expert viewer. The subject evaluates the deterioration level of the test video with respect to the reference video using a five grades scale (5, imperceptible; 4, perceptible, but not annoying; 3, slightly annoying; 2, annoying; and 1, very annoying). To collect a significant sample of evaluations, the number of subjects should be at least 15.

MOS is the most significant method for testing the performance of a video coding algorithm, since it is based on the real viewing experience of users of the video coding technology. Consequently, all official results of video quality assessments in MPEG and VCEG are based on the MOS methodology.

D. Bjontegaard Delta Rate

This method aims at calculating the average difference of BR between two curves, over a range of PSNR. It was

originally proposed in [21], where the interpolation of BR in dB with a third degree polynomial was recommended. The method, originally used for the interpolation of BR expressed as a function of PSNR, can be extended to a generic metric X,

$$\mathrm{BR_{dB}} = b_0 + b_1 X + b_2 X^2 + b_3 X^3 \qquad (4)$$

where $\mathrm{BR_{dB}} = 10 \log \mathrm{BR}$, and $\boldsymbol{b} = [b_0, \ldots, b_3]$ is a vector of coefficients to be estimated from data. Then the difference between polynomials can be averaged by integrating it in the full range of the metric X and dividing by the same range of X. Because of the polynomial interpolation, this can be also easily accomplished analytically as

$$\Delta \mathrm{BR_{dB,\,int}} = \delta b_0 X + \frac{\delta b_1}{2} X^2 + \frac{\delta b_2}{3} X^3 + \frac{\delta b_3}{4} X^4 \qquad (5)$$

where δb_i, $i = 0 \ldots 3$, are the differences between the interpolating coefficients of the two codecs under test. Then the integral is evaluated at the extremes of the range and divided by the range itself to obtain the Bjontegaard Delta Rate (BD rate),

$$\mathrm{BDR_{dB}} = \frac{\Delta \mathrm{BR_{dB,\,int}}(X_2) - \Delta \mathrm{BR_{dB,\,int}}(X_1)}{X_2 - X_1} \qquad (6)$$

The result can then be easily converted from dB to percentage as

$$\mathrm{BDR_\%} = 100(10^{\frac{\mathrm{BDR_{dB}}}{10}} - 1) \qquad (7)$$

VI. TEST RESULTS

This section presents a selection of the results related to [1]. In particular, medium and sharp video sequences, DrivingLogo, BoxeLogo, TrafficLogo, Stracraft, highlighted in bold in Tab. I, have been selected to compare the performance of the codecs, because these sequences are common to all four codecs. All test results are reported in supplementary material.

Fig. 3 and Fig. 4 show the result for PSNR, on the left column of panels and VMAF metric, on the right column, for the AVC and VVC, respectively. The comparison between the graphs for PSNR and those for VMAF shows the different behaviour of the three test conditions with respect to the two metrics.

Fig. 3 shows that with respect to a purely statistical metric like PSNR, LCEVC performs better on average (panels a,c,e versus g) than the full resolution AVC, while for the most advanced full resolution VVC it is true the contrary. But wit respect to the objective metric VMAF, Fig. 3 shows that the LCEVC outperforms the full resolution AVC codec and it gives comparable results (although in two panels, slightly lower) to the full resolution VVC. Since VMAF models the artefacts introduced by video coding and their effects on the viewer, thus providing a more accurate estimate of the subjective quality of the video sequences under test, this results show that LCEVC performs better with metrics that takes into account the perceived effect of quantization than with purely statistical metric like PSNR. This will be confirmed by the comparison of LCEVC and the other four codecs performance with respect a subjective metric like MOS.
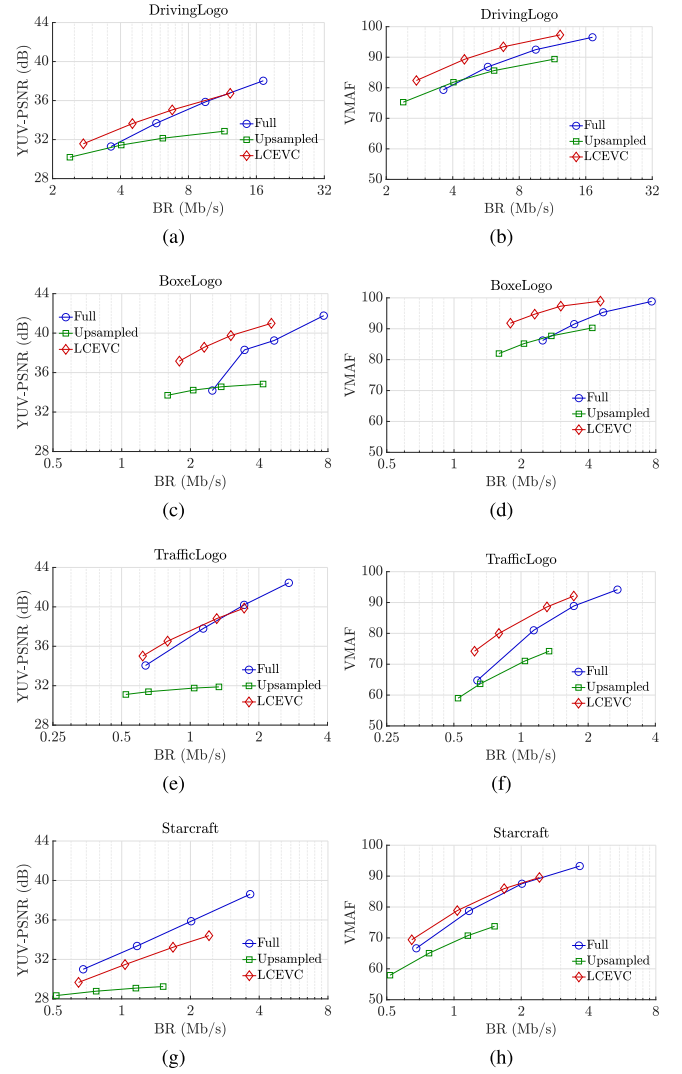

Fig. 3. LCEVC with AVC test results. Left panels: PSNR. Right panels: VMAF.

As for any other video coding standard specified by ISO MPEG and ITU-T VCEG, the verification of the requirements was also evaluated by the subjective tests, performed by real users following the procedures specified by ITU-R BT.500 DSIS MOS [18]. The viewing set up at the test laboratory was as follows:

- 65" TV set, with OLED screen set with "standard" viewing options and HDMI 2.1 input interface, capable of accepting and displaying 10-bit content.
- Suitable video player system, able to play out YUV UHD content up to 60 fps and 420 progressive colour scheme, in a fluid way (i.e., at full frame rate and without frame skipping) and without impairments.
- Protected viewing area (that is, no external video or audio pollution), with low illumination behind the screen (around 30 nits) not visible to the viewing subject, and without any other ambient light.
- Two seats for each testing room, at a distance of 1 meter from each other.
- Viewing distance 2H (two times the height of the screen).
- Two separate waiting areas for viewing subjects, while waiting and resting.
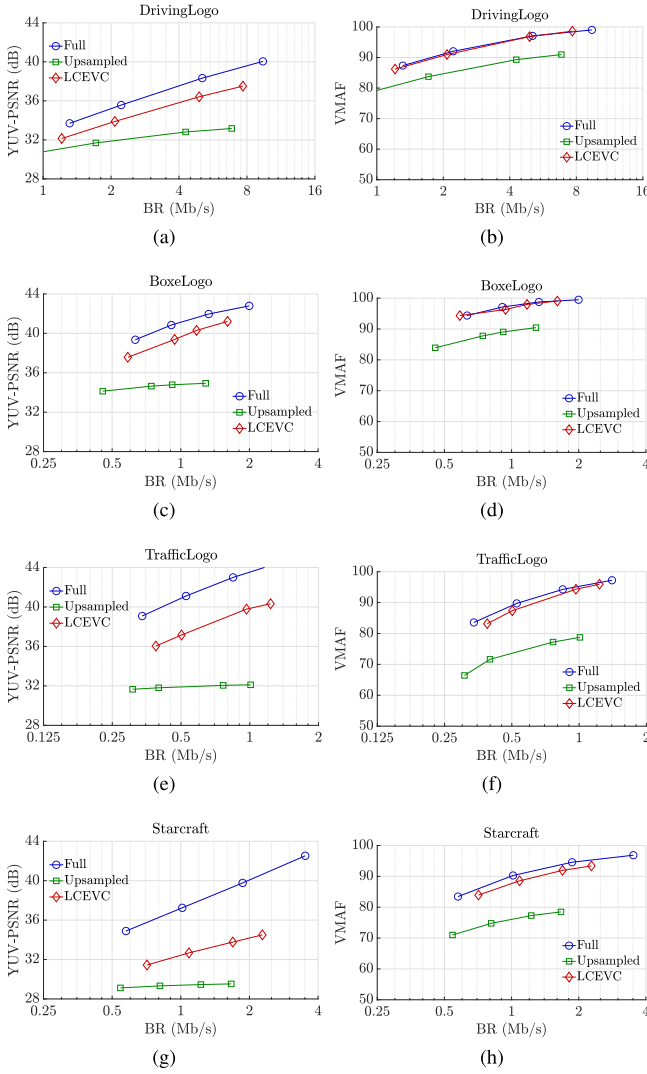
Fig. 4. LCEVC with VVC test results. Left panels: PSNR. Right panels: VMAF.

Fig. 5. LCEVC with AVC (left) and HEVC (right) test results. MOS.

The results of the subjective tests performed for LCEVC are reported in Fig. 5 for AVC and HEVC, and in Fig. 6 for EVC and VVC, respectively. All subjective test results, object of [2], are reported in the supplementary material. Fig. 5 show a good performance margin for LCEVC with respect both AVC and HEVC. The advantage is confirmed by Fig. 6 for newest codecs although with a reduced margin. Details about numerical amount of the BR saving will be given in the next subsection.

A. Numerical Summary of the Codec Performance

For each sequence, the bitrate saving in terms of BD rate for the common MOS range is computed. Then, the BD rates are averaged among the sequences of the set, to obtain a single numerical estimate of the bitrate saving comparing the Base codec at quarter resolution in conjunction with LCEVC, to the Base codec at full resolution.

Using LCEVC (in its implementation LTM, LCEVC Test Model) with the base codec AVC (implementation JM, AVC Joint Test Model [22]), the BD rates are those reported in Tab. II, and the bitrate savings for the same quality range are
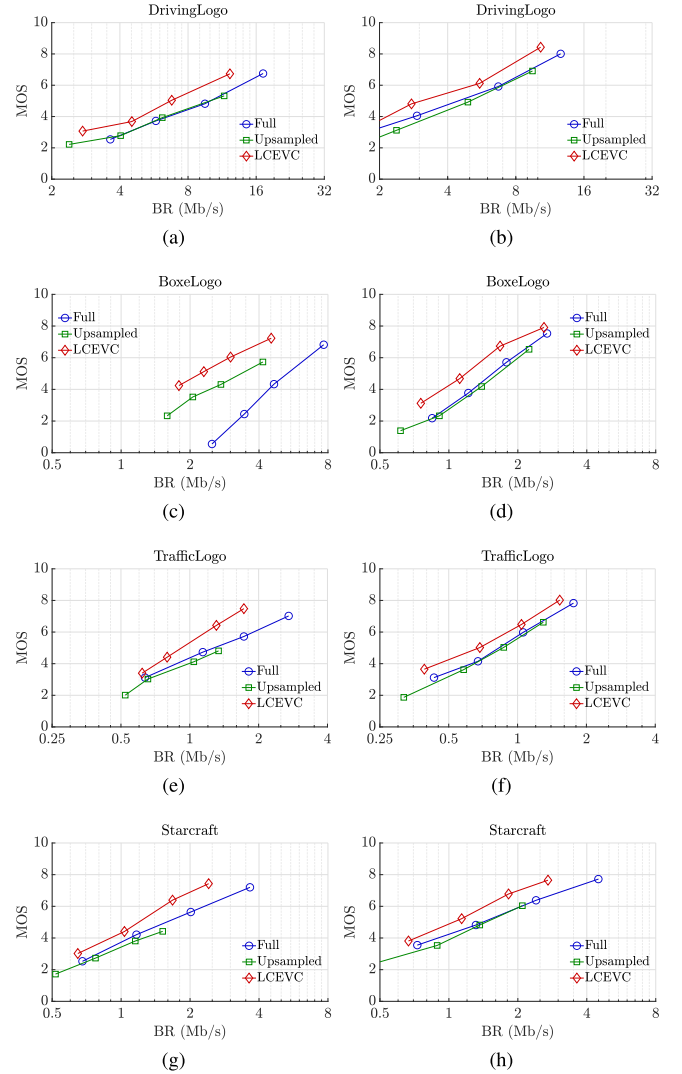
TABLE II
BD RATES FOR MOS FOR LCEVC WITH AVC

LCEVC/AVC	BD rate
CatRobot	-43.83%
LupoPuppet	-53.98%
DrivingLogo	-30.19%
BoxeLogo	-55.61%
Average UHD	**-45.90%**
TrafficLogo	-30.18%
Starcraft	-26.75%
Average HD	**-28.47%**

46% for the UHD sequences, 28% for the HD sequences, and 40% is the average over the 6 sequences.

Using LCEVC (in its implementation LTM) with the base codec HEVC (implementation HM, HEVC Test Model [23]), the BD rates are those reported in Tab. III, and the bitrate savings for the same quality range are 31% for the UHD sequences, 24% for the HD sequences, and 29% is the average over the 6 sequences.

Using LCEVC (in its implementation LTM) with the base codec EVC (implementation ETM, EVC Test Model), the BD
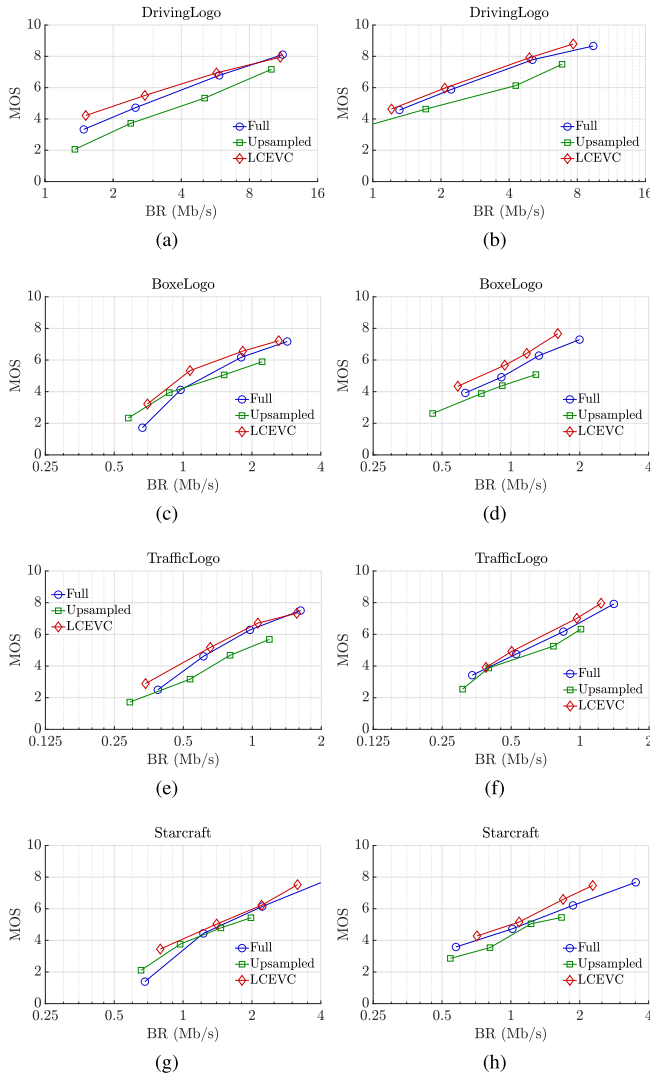
Fig. 6.   LCEVC with EVC (left) and VVC (right) test results. MOS.

TABLE III
BD Rates for MOS for LCEVC With HEVC

LCEVC/HEVC	BD Rate
CatRobot	-41.88%
LupoPuppet	-31.14%
DrivingLogo	-26.01%
BoxeLogo	-24.44%
Average UHD	**-30.87%**
TrafficLogo	-18.52%
Starcraft	-29.76%
Average HD	**-24.14%**

rates are those reported in Tab. IV, and the bitrate savings for the same quality range are 18% for the UHD sequences, 9% for the HD sequences, and 15% is the average over the 6 sequences.

Using LCEVC (in its implementation LTM) with the base codec VVC (implementation VTM, VVC Test Model [24]), the BD rates are those reported in Tab. V, and the bitrate savings for the same quality range are 16% for the UHD sequences, 14% for the HD sequences, and 15% is the average

TABLE IV
BD Rates for MOS for LCEVC With EVC

LCEVC/EVC	BD rate
CatRobot	-18.54%
BarScene	-20.50%
DrivingLogo	-14.83%
BoxeLogo	-17.49%
Average UHD	**-17.84%**
TrafficLogo	-9.27%
Starcraft	-7.83%
Average HD	**-8.55%**

TABLE V
BD Rates for MOS for LCEVC With VVC

LCEVC/VVC	BD rate
Marathon	-16.86%
MountainBay	-16.90%
DrivingLogo	-10.33%
BoxeLogo	-18.55%
Average UHD	**-15.66%**
TrafficLogo	-10.33%
Starcraft	-17.96%
Average HD	**-14.15%**

TABLE VI
BD Rates for MOS for LCEVC

codec	BD rate
AVC	40.1%
HEVC	28.6%
EVC	14.7%
VVC	15.2%

over the 6 sequences. The BD rate saving, averaged for all six test sequences, are summarized in Table VI.

B. Graphical Summary of the Codec Performance

Although the BD rate methodology described in the previous Sec. VI-A is the best practice implemented by MPEG and VCEG, it is also possible to average the RD models for a set of sequences, and compute an average model to describe the RD behaviour of a specific video codec as a function of BR as discussed in [17].

An advantage of this alternative methodology is the possibility to describe the RD characteristic of a codec over a set of sequences as a function of BR (rather than a scalar as with the BD rate). Such model of the codec under test will be more accurate increasing the number of test sequences.

The performance of a codec, calculated over a set of sequences, can be modeled by a cubic polynomial representing the logarithm of the bitrate as a function of the metric of interest. If the metric of interest is the MOS, as in our case, it is possible to write a model analog to (4)

$$\text{BR}_{\text{dB}} = b_0 + b_1\text{MOS} + b_2\text{MOS}^2 + b_3\text{MOS}^3. \qquad (8)$$

The vector $\boldsymbol{b}$ of coefficients is estimated for each sequence, using a Least Square estimator and, unlike the BD rate method in V-D, it is averaged over the complete set of
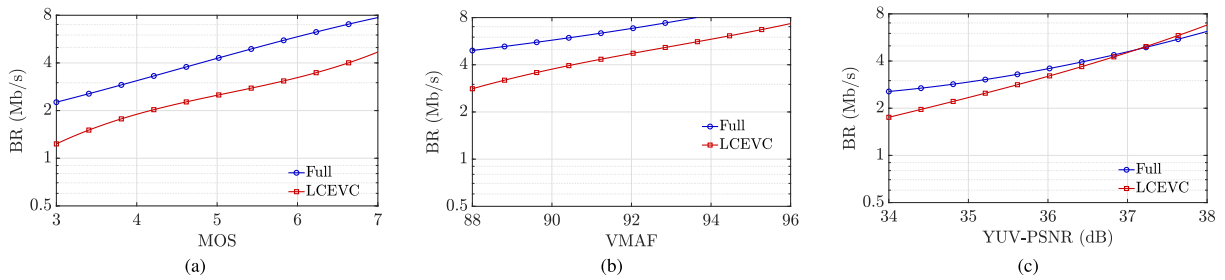
Fig. 7.   LCEVC with AVC. (a) MOS; (b) VMAF; (c) PSNR.
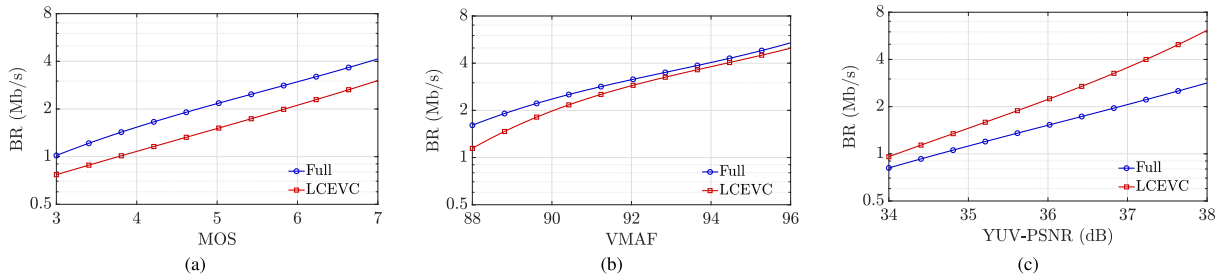

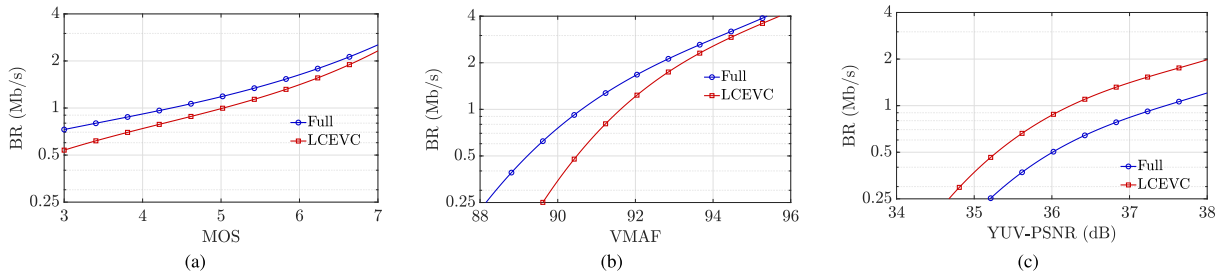Fig. 8.   LCEVC with HEVC. (a) MOS; (b) VMAF; (c) PSNR.

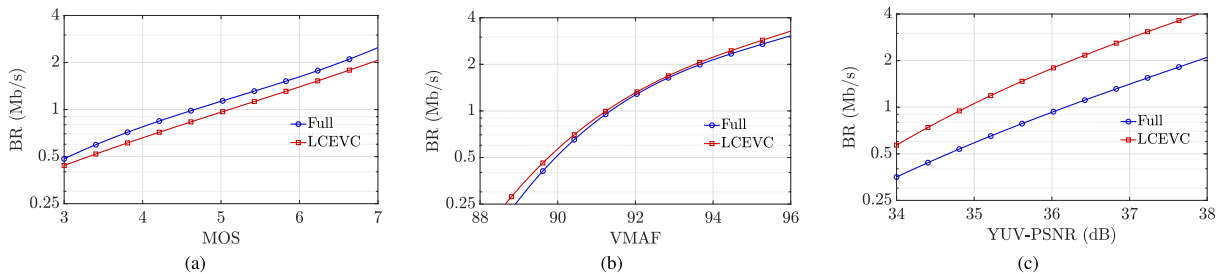Fig. 9.   LCEVC with EVC. (a) MOS; (b) VMAF; (c) PSNR.

Fig. 10.   LCEVC with VVC. (a) MOS; (b) VMAF; (c) PSNR.

sequences, so obtaining an average polynomial model, valid in the operational ranges chosen.

Figs. 7–10 show these polynomial models for all the four codecs and the three metrics adopted for the verification test. With respect to the BD rate method, this model provides the difference as a function of the metric, and it can also be used to obtain an average bitrate saving, equivalent to that obtained with (6) and (7).

Comparing the BD rates of Table VI, with those obtained with this average model, shown in Table VII, we can notice some differences. The deviations between the two estimates are $+1.8\%$ of estimated bitrate saving for the AVC codec, $+0.3\%$ for the HEVC codec, $+2.0\%$ for the EVC codec, and finally -0.9% for the VVC codec.

These slight differences are due to the the small sample size and they are bound to decrease as the number of sequences

TABLE VII
BITRATE SAVINGS FOR MOS FOR LCEVC

codec	Bitrate saving
AVC	41.9%
HEVC	28.9%
EVC	16.7%
VVC	14.3%

increases. Nevertheless this methodology has the advantage of obtaining a graphical representation of the difference in RD performance between two codecs, shown by the Figs. 7–10.

For example, from Figs. 9–10, observing the the full resolution encoding (blue line) and the LCEVC encoding (red line) it is possible to estimate which of the two settings provides better results at lower and higher bitrates. With AVC

and HEVC the bitrate saving is almost constant with respect to the bitrate, as shown in Fig 7(a) and Fig 8(a). With EVC, the bitrate saving for LCEVC is larger at lower bitrates (Fig. 9(a)). With VVC, vice versa, the bitrate saving for LCEVC is slightly larger at higher bitrates (Fig. 10(a)).

VII. Conclusion

This paper describes and presents the results of the verification tests [1], [2] performed in the final phase of the development of the ISO standard LCEVC (Low Complexity Enhancement Video Coding) [3], finalized by the MPEG Video Coding Working Group (ISO/IEC JTC1 SC29 WG04) in November 2021.

In particular, three visual quality metrics have been used: PSNR, VMAF, MOS. The results of the verification test confirm also that the objective metric VMAF is an adequate estimate of the subjective quality computed by the subjective metric MOS, based on the assessment by real viewers.

As an alternative to the best practice BD rate methodology, for comparing different codecs, a new methodology based on the average of the polynomial model of RD curves for different codecs was also applied. The results show that it is possible to use this methodology to obtain an estimate of the RD performance as a function of BR.

In summary, the bitrate savings, computed using the BD rate methodology, when comparing the four single-layer codecs used at full resolution, with the same codec used at quarter resolution in conjunction with LCEVC at full resolution are:
- around 40% with AVC (JM);
- around 30% with HEVC (HM);
- around 15% with EVC (ETM);
- around 15% with VVC (VTM).

Acknowledgment

References

[1] F. Maurer, L. Ciccarelli, and S. Ferrara. (2021). *LCEVC Verification Test—Overview of Bitrate, PSNR, VMAF, and MD5 Checksums*. [Online]. Available: https://dms.mpeg.expert/doc_end_user/documents/134_OnLine/wg11/m56330-v%6-m56330-Verification]_Test_PSNR_VMAF_MD5-v6.zip

[2] (2021). ISO/IEC WG04 N0076. *Verification Test Report on the Compression Performance of Low Complexity Enhancement Video Coding*. [Online]. Available: https://dms.mpeg.expert/doc_end_user/documents/134_OnLine/wg11/MDS20173% _WG04_N00076.zip

[3] *Information Technology—General Video Coding—Part 2: Low Complexity Enhancement Video Coding*, Standard ISO/IEC 23094–2, 2021.

[4] *Information Technology—General Video Coding—Part 3: Conformance and Reference Software for Low Complexity Enhancement Video Coding*, Standard ISO/IEC 23094–3, 2021.

[5] *Information Technology—Coding of Audio-Visual Objects—Part 10: Advanced Video Coding*, Standard ISO/IEC 14496-10, 2003.

[6] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 560–576, Jul. 2003.

[7] *Information Technology—High Efficiency Coding and Media Delivery in Heterogeneous Environments—Part 2: High Efficiency Video Coding*, Standard ISO/IEC 23008–2, 2013.

[8] G. J. Sullivan, J.-R. Ohm, W.-J. Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1649–1668, Dec. 2012.

[9] *Information Technology—General video coding—Part 1: Essential Video Coding*, Standard ISO/IEC 23094–1, 2020.

[10] K. Choi, J. Chen, D. Rusanovskyy, K.-P. Choi, and E.-S. Jang, "An overview of the MPEG-5 essential video coding standard," *IEEE Signal Process. Mag.*, vol. 37, no. 3, pp. 160–167, May 2020.

[11] *Information Technology—Coded Representation of Immersive Media—Part 3: Versatile Video Coding*, Standard ISO/IEC 23090–3, 2021.

[12] B. Bross et al., "Overview of the versatile video coding (VVC) standard and its applications," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 31, no. 10, pp. 3736–3764, Oct. 2021.

[13] S. Ferrara et al., "MPEG-5 part 2: Low complexity enhancement video coding (LCEVC): Overview and performance evaluation," *Proc. SPIE*, vol. 11510, p. 44, Aug. 2020.

[14] A. Jiménez-Moreno, L. Ciccarelli, R. Clucas, and S. Ferrara, "HDR video coding with MPEG-5 LCEVC," in *Proc. 1st Mile-High Video Conf.*, New York, NY, USA, Mar. 2022, pp. 25–31.

[15] N. Barman, S. Schmidt, S. Zadtootaghaj, and M. G. Martini, "Evaluation of MPEG-5 part 2 (LCEVC) for live gaming video streaming applications," in *Proc. 1st Mile-High Video Conf.*, New York, NY, USA, Mar. 2022, pp. 108–109.

[16] S. Battista et al., "Overview of the low complexity enhancement video coding (LCEVC) standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 32, no. 11, pp. 7983–7995, Nov. 2022.

[17] S. Battista, M. Conti, and S. Orcioni, "Methodology for modeling and comparing video codecs: HEVC, EVC, and VVC," *Electronics*, vol. 9, no. 10, p. 1579, Sep. 2020.

[18] *Methodologies for the Subjective Assessment of the Quality of Television Images*, document ITU-R BT.500:2019, 2019.

[19] Z. Li, A. Aaron, I. Katsavounidis, A. Moorthy, and M. Manohara. (2016). *Toward A Practical Perceptual Video Quality Metric*. [Online]. Available: https://netflixtechblog.com/toward-a-practical-perceptual-video-quality-%metric-653f208b9652

[20] S. Li, F. Zhang, M. Lin, and K. N. Ngan, "Image quality assessment by separately evaluating detail losses and additive impairments," *IEEE Trans. Multimedia*, vol. 13, no. 5, pp. 935–949, Oct. 2011.

[21] G. Bjontegaard. (2001). *Calculation of Average PSNR Differences Between RD-Curves*. [Online]. Available: https://www.itu.int/wftp3/av-arch/video-site/0104_Aus/VCEG-M33.doc

[22] MPEG/VCEG. (2022). *AVC Reference Software*. [Online]. Available: https://avc.hhi.fraunhofer.de/

[23] MPEG/VCEG. (2022). *HEVC Reference Software*. [Online]. Available: https://hevc.hhi.fraunhofer.de/

[24] MPEG/VCEG. (2022). *VVC Reference Software*. [Online]. Available: https://vvc.hhi.fraunhofer.de/

Stefano Battista received the Laurea degree in electronic engineering from Università Politecnica delle Marche, Italy, in 1990. From 1991 to 1997, he was a Researcher with the Telecom Italia Laboratory, Video and Multimedia Group, Torino, Italy. From 1997 to 1999, he was a System Engineer with the Advanced Systems Technology Group, STMicroelectronics, Agrate, Milano, Italy. His research interests include video coding, video analysis, 3D modeling, and multimedia systems, with a focus on standardization for multimedia communications. His main activities were in the development of new architectures for multimedia consumer products, specifically on DVB/DVD platforms, and their evolution towards interactive TV. From 1999 to 2018, he was a Researcher and a System Engineer with bSoft, Macerata, Italy, where he has been active in research and development. He has published several research papers on video coding and multimedia systems, and contributed to books related to the fields of communications and multimedia. He actively participated in the standardization effort of MPEG, in particular to MPEG-4 reference software development, co-authoring part 5 of the standard. For this contribution to the standardization activity, he received an ISO Certificate of Appreciation. He contributed to co-operative research projects in the third (RACE), fourth (ACTS), fifth (IST), sixth, and seventh (ICT) Framework Programs of the European Commission. He served as a consultant for several companies in Italy and abroad, and an External Evaluator and a Reviewer of the European Commission.

Guido Meardi received the M.B.A. degree from MIT Sloan and the dual M.Sc. degree in computer engineering from the Politecnico di Milano and The University of Texas at Austin. He is currently the Co-Founder and the CEO of V-Nova, leading company in data compression and artificial intelligence. He is also a Keen Innovator, an Entrepreneur, and an Investor, with relevant business building experience and half a dozen exits. Having retained in-depth scientific and engineering expertise, he contributed to the foundational development work at V-Nova Intellectual Property portfolio. He is a Siebel Scholar with MIT Sloan. He was an Intel Scholar with the Politecnico di Milano. He is an Inventor or a Joint Inventor of several essential aspects of the coding standards MPEG-5 Part 2 Low Complexity Enhancement Video Coding (LCEVC) and SMPTE VC-6 ST-2117, with over 200 patents filed. He is the Former Senior Partner at McKinsey, where he was the Head of the Organization and Operations Practices of the Mediterranean Complex. He also has a breadth of business experience and access to senior executives in a variety of industries and geographies, with well-established experience in telecoms, technology, healthcare, insurance, banking, automotive, aerospace, and defense. He led transformational projects in all continents and was instrumental in setting up some of McKinsey's own innovation-related business building activities.

Florian Maurer received the master's degree in electrical engineering, information technology and computer engineering from RWTH Aachen University, Germany, in 2020. He has focused his studies on signal processing with a particular interest in video coding and gained industry experience while working for a leading company in video compression. During his work at the RWTH Institute of Communications Engineering chaired by Prof. Jens-Rainer Ohm, he acquired comprehensive knowledge of different video coding standards. He was involved in the standardization process of MPEG-5 Part 2—Low Complexity Enhancement Video Coding (LCEVC). His contributions helped to move LCEVC forward to the International Standard (IS) stage and he has taken part in the development of a reference implementation of LCEVC. Recently, he has also shown interest in graph signal processing (GSP) and is researching motion estimation in dynamic 3D point cloud sequences.

Simone Ferrara received the Graduate degree in telecommunications engineering from the Politecnico di Milano and the M.Sc. degree in electrical engineering from Washington University in St. Louis, USA. He is currently the Senior Vice President at V-Nova, where he is responsible for the technology, IP, and standardization strategy, including driving its execution across the company. He has been leading the development of MPEG-5 Part 2 Low Complexity Enhancement Video Coding and has contributed to many of the coding tools included in the standard. Prior to V-Nova, he worked in the telecommunications industry first as a Researcher and then as an IP Expert.

Massimo Conti (Member, IEEE) received the Graduate degree in electronics engineering from the University of Ancona, Italy, in 1987. He is currently an Associate Professor at the Dipartimento di Ingegneria dell'Informazione (DII), Università Politecnica delle Marche (UNIVPM), Ancona, Italy. His research activity in the field of microelectronics is mainly devoted to system level design of low power integrated circuits, design of energy harvesting systems, battery management systems, and vehicle to grid connection. He is the coauthor of more than 190 papers on international books, journals, or conferences. He has 175 Scopus publications, 1190 citations, and H-index of 17. He is the coordinator of European and national research projects. He is an editor of seven international books, the guest editor of three special issues, and the general chairperson of seven international conferences. For information visit the link (www.univpm.it/massimo.conti).

Lorenzo Ciccarelli received the Laurea degree in electronic engineering and telecommunications from Università Politecnica delle Marche, Italy, in 1998. His master's thesis on video coding and in particular working on algorithm optimization for H.263+ video codecs. From 1999 to 2006, he was worked on several aspects of video compression algorithm design and implementation participating and leading projects focused on developing codecs on different VLIW architectures mainly used for videoconference terminal and multiconference unit. In 2006, he moved to U.K. joining Ericsson SATTV (former Tandberg TV) Research and Development Department, where he had the opportunity to deepen his knowledge of the TV broadcasting side of the video coding gaining expertise in rate control and statmuxing while porting video compression algorithm on FPGAs. Between 2008 and 2014, he was involved in several projects to design software test models used to design, test, and improve different algorithms based on MPEG2, AVC, and HEVC for large broadcasting systems based on multiple FPGA, DSPs, and CPUs. Between 2014 and 2016, he was leading the design of one of the first hardware implementation of a HEVC full UHD encoder based on a hybrid x86 and FPGA architecture to then join BBC video coding research and development, where he spent two years in improving internal video encoding testing platform (Turing encoder) and working on different European funded projects. In 2018, he joined V-Nova Ltd., with the title of a Principal Research Engineer. During the last two years, he has been involved in the algorithm design and standardization process for MPEG-5 Part 2 LCEVC.

Simone Orcioni (Senior Member, IEEE) received the Laurea degree in electronics engineering from the Università degli Studi di Ancona, Italy, in 1992, and the Ph.D. degree in 1995. From 1997 to 1999, he was Post-Doctoral Fellow at the Università degli Studi di Ancona. In 2000, he became an Assistant Professor, teaching courses in analog and digital electronics, and publishing a text book. In 2017, he was an Adjunct Professor at the Ubiquitous Computing Laboratory (UC-Laboratory), Hochschule Konstanz University of Applied Sciences (HTWG), Konstanz, Germany, where he is currently the Guest Researcher. Since 2021, he has been an Associate Professor at the Department of Information Engineering, Università Politecnica delle Marche. He is a member of EURASIP. He was the Guest Editor of *EURASIP Journal on Embedded Systems*, *Frontiers in Energy Research*, and *Sensors* (MDPI). He was the program chair of three international conferences, the editor of four international books, and an inventor in two patents. He is currently the Guest Editor of *Mathemathics* (MDPI). He has been working in statistical device modeling and simulation, analog circuit design, cyber-physical system simulation, and linear and nonlinear system identification. His current research interests include non-linear digital signal processing and power electronics for renewable energies.