**** **ISO/IEC JTC 1/SC 29/WG 04 N0076**

**ISO/IEC JTC 1/SC 29/WG 04  
MPEG Video Coding   
Convenorship: CN**

**Document type:** Output Document

**Title:** Verification Test Report on the Compression Performance of Low Complexity Enhancement Video Coding

**Status:** Approved

**Date of document:** 2020-05-13

**Source:** ISO/IEC JTC 1/SC 29/WG 04

**Expected action:** None

**Action due date:** None

**No. of pages:** 24

**Email of Convenor:** yul@zju.edu.cn

**Committee URL:** <https://isotc.iso.org/livelink/livelink/open/jtc1sc29wg4>

**INTERNATIONAL ORGANIZATION FOR STANDARDIZATION**

**ORGANISATION INTERNATIONALE DE NORMALISATION**

**ISO/IEC JTC 1/SC 29/WG 04 MPEG VIDEO CODING**

**ISO/IEC JTC 1/SC 29/WG 04 N** **0076**

**April 2021, Online**

|  |  |
| --- | --- |
| **Title** | **Verification Test Report on the Compression Performance of Low Complexity Enhancement Video Coding** |
| **Source** | **WG 04, MPEG Video Coding** |
| **Status** | **Approved** |
| **Serial Number** | **20173** |

# Executive Summary

This document reports verification test results for MPEG-5 Part 2 Low Complexity Enhancement Video Coding (LCEVC) compression performance verification.

The purpose of the verification tests was to confirm that the coding efficiency objective for the LCEVC standard has been met achieving a bit-rate reduction at a similar level of subjective visual quality relative to a single-layer video codec. This document reports the results of the verification test to confirm that this goal was achieved and to estimate the magnitude of this achievement.

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 AVC were determined to be approximately 46% for UHD and 28% for HD. The average bit rate savings for LCEVC when enhancing 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 with these two codecs.

The second set of tests aimed to confirm that LCEVC provided a more efficient means of resolution enhancement of half resolution anchors than unguided up-sampling. For these tests, the test sequences were coded using AVC, HEVC, EVC, or VVC at half resolution in both horizontal and vertical direction. For anchor generation, the half resolution encoded sequences were upsampled with Lanczos filters to full resolution for visual assessment. The same half resolution encoded sequences were also used as base layers for LCEVC and hence not all curves may overlap as much as would be ideal when calculating a BD-rate. Comparing LCEVC full-resolution encoded sequences with the up-sampled half-resolution anchors, the average bit-rate savings when using LCEVC with AVC, HEVC, EVC, and VVC were calculated to be approximately 28%, 34%, 38%, and 33% respectively for UHD, and 27%, 26%, 21%, and 21% respectively for HD.

# Standard Dynamic Range Test Conditions

The verification test for LCEVC includes standard dynamic range (SDR) UHD test sequences encoded in random-access configuration and SDR HD test sequences encoded in random-access configuration. A short description of the tests performed is reported in Annex A.

All test materials are progressively scanned and use 4:2:0 colour sampling with 8/10 bits per sample bit depth. Table 1 through Table 4 summarize test material information.

Table 1. Test video sequences for AVC test

| **Class** | **Sequence name** | **Resolution** | **Frame count** | **Frame rate** | **Chroma format** | **Bit depth** |
| --- | --- | --- | --- | --- | --- | --- |
| SDR-A | CatRobot | 3840×2160 | 600 | 60 | 4:2:0 | 8 |
| LupoPuppet | 3840×2160 | 500 | 50 | 4:2:0 | 8 |
| DrivingPOVLogo | 3840×2160 | 600 | 60 | 4:2:0 | 8 |
| BoxeLogo | 3840×2160 | 596 | 60 | 4:2:0 | 8 |
| SDR-B | TrafficLogo | 1920×1080 | 600 | 60 | 4:2:0 | 8 |
| Starcraft | 1920×1080 | 600 | 60 | 4:2:0 | 8 |

Table 2. Test video sequences for HEVC test

| **Class** | **Sequence name** | **Resolution** | **Frame count** | **Frame rate** | **Chroma format** | **Bit depth** |
| --- | --- | --- | --- | --- | --- | --- |
| SDR-A | CatRobot | 3840×2160 | 600 | 60 | 4:2:0 | 10 |
| LupoPuppet | 3840×2160 | 500 | 50 | 4:2:0 | 10 |
| DrivingPOVLogo | 3840×2160 | 600 | 60 | 4:2:0 | 10 |
| BoxeLogo | 3840×2160 | 596 | 60 | 4:2:0 | 10 |
| SDR-B | TrafficLogo | 1920×1080 | 600 | 60 | 4:2:0 | 10 |
| Starcraft | 1920×1080 | 600 | 60 | 4:2:0 | 10 |

Table 3. Test video sequences for EVC test

| **Class** | **Sequence name** | **Resolution** | **Frame count** | **Frame rate** | **Chroma format** | **Bit depth** |
| --- | --- | --- | --- | --- | --- | --- |
| SDR-A | BarScene | 3840×2160 | 600 | 60 | 4:2:0 | 10 |
| CatRobot | 3840×2160 | 600 | 60 | 4:2:0 | 10 |
| DrivingPOVLogo | 3840×2160 | 600 | 60 | 4:2:0 | 10 |
| BoxeLogo | 3840×2160 | 596 | 60 | 4:2:0 | 10 |
| SDR-B | TrafficLogo | 1920×1080 | 600 | 60 | 4:2:0 | 10 |
| Starcraft | 1920×1080 | 600 | 60 | 4:2:0 | 10 |

Table 4. Test video sequences for VVC test

| **Class** | **Sequence name** | **Resolution** | **Frame count** | **Frame rate** | **Chroma format** | **Bit depth** |
| --- | --- | --- | --- | --- | --- | --- |
| SDR-A | Marathon | 3840×2160 | 300 | 30 | 4:2:0 | 10 |
| MountainBay2 | 3840×2160 | 300 | 30 | 4:2:0 | 10 |
| DrivingPOVLogo | 3840×2160 | 600 | 60 | 4:2:0 | 10 |
| BoxeLogo | 3840×2160 | 596 | 60 | 4:2:0 | 10 |
| SDR-B | TrafficLogo | 1920×1080 | 600 | 60 | 4:2:0 | 10 |
| Starcraft | 1920×1080 | 600 | 60 | 4:2:0 | 10 |

# Evaluation Procedure

For SDR content the DSIS test method [1] was used. The SDR subjective test was carried out at the VABTech and GBTech laboratories in Rome (Italy), acting in parallel on the same visual content.

Viewing set up at the test laboratories were as follows, considering the severe restrictions imposed by the COVID-19 pandemic:

* Daily deep clean-up of both viewing room and waiting room (using cleaning products specific to protect from COVID diffusion) to avoid any contamination and protect both viewers and test managers.
* Air purification one hour before the beginning of the testing activities using disinfectant spray cans to produce a cleaning tide for 45 minutes; then fresh air renewal for 15 minutes (opening all windows and doors) before test.
* All viewers and test managers wore high efficiency masks (e.g., U-Mask), washed hands and used disinfectant liquid.
* All viewers and test managers provided a COVID throat swab certificate showing a negative result to a COVID test performed not more than 2 days before.
* Each laboratory is equipped with a viewing and a resting area, each with one seat.
* Viewing distance: 2H
* 65” TV set, with OLED screen set with “standard” viewing option 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 420p colour scheme, in a fluid way (i.e., faithful frame rate and no frame jump) and no impairments.
* Viewing area protected by external audible and visible disturbance (no external video or audio pollutions)
* Low illumination behind the screen (around 30 nits) not visible to the viewing subject(s) and no other ambient light.
* A total of 50 viewers participated to the tests, spread across 10 days of testing activities.
* All viewers were screened for visual acuity (Snellen chart) and colour blindness (Ishihara tables).
* 20 female and 30 male viewers, all university students aging from 18 to 24 years.
* All scores were collected by means of paper scoring sheets.

# Test results and graphs

The anchors have been generated using AVC (JM 19.0), HEVC (HM 16.20), EVC (ETM 6 rc 1), and VVC (VTM 11). LCEVC encoded sequences were generated using the software model LTM 5.1.

Throughout this document, the following terminology is used to refer to different types of encoded sequences (xM indicates a generic test model):

|  |  |
| --- | --- |
| * full-xM: | anchor encoded at full resolution using xM as test model |
| * half-xM: | anchor encoded at half resolution in both horizontal and vertical direction, using xM as test model, and then upsampled with Lanczos filter [2] to full resolution using FFmpeg 4.3.1 |
| * LTM+xM: | LCEVC encoded at full resolution using LTM and enhancing xM used for encoding the base layer |

* 1. **SDR UHD resolution – LTM over JM vs. JM**

Table 5. Table of results for SDR UHD sequences

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **full-JM** | | | | **half-JM** | | | | **LTM+JM** | | | |
| *sequence* | *QP* | *bitrate* | *MOS* | *CI* | *QP* | *bitrate* | *MOS* | *CI* | *QP* | *bitrate* | *MOS* | *CI* |
| CatRobot | 27 | 15544.26 | 7.65 | 0.31 | 23 | 10549.13 | 6.62 | 0.24 | 23 | 10620.90 | 7.35 | 0.24 |
| 32 | 8207.94 | 5.12 | 0.44 | 28 | 5328.52 | 4.73 | 0.31 | 28 | 5359.38 | 5.77 | 0.24 |
| 37 | 4677.54 | 3.53 | 0.24 | 33 | 2836.47 | 3.38 | 0.26 | 33 | 2883.36 | 3.82 | 0.27 |
| 42 | 2696.16 | 2.08 | 0.32 | 37 | 1735.31 | 2.63 | 0.38 | 37 | 1834.08 | 3.01 | 0.27 |
| LupoPuppet | 30 | 11802.85 | 6.92 | 0.27 | 26 | 7253.05 | 6.63 | 0.30 | 26 | 7284.90 | 7.48 | 0.30 |
| 34 | 7243.20 | 5.03 | 0.37 | 29 | 4673.85 | 5.15 | 0.19 | 29 | 4694.90 | 6.12 | 0.46 |
| 38 | 4695.95 | 3.12 | 0.43 | 33 | 2691.10 | 3.52 | 0.28 | 33 | 2732.70 | 4.23 | 0.34 |
| 42 | 3090.75 | 2.02 | 0.33 | 37 | 1654.25 | 2.04 | 0.25 | 37 | 1722.80 | 2.95 | 0.39 |
| DrivingPOVLogo | 28 | 17165.46 | 6.75 | 0.27 | 24 | 11552.58 | 5.32 | 0.26 | 24 | 12248.78 | 6.73 | 0.42 |
| 32 | 9502.20 | 4.82 | 0.25 | 28 | 6158.82 | 3.93 | 0.28 | 28 | 6779.34 | 5.04 | 0.20 |
| 36 | 5767.26 | 3.72 | 0.25 | 31 | 4029.72 | 2.78 | 0.42 | 31 | 4517.04 | 3.68 | 0.24 |
| 42 | 3628.44 | 2.54 | 0.28 | 35 | 2386.80 | 2.22 | 0.40 | 35 | 2738.34 | 3.07 | 0.35 |
| BoxeLogo | 30 | 7664.70 | 6.82 | 0.22 | 24 | 4159.14 | 5.73 | 0.36 | 24 | 4522.26 | 7.23 | 0.26 |
| 34 | 4649.76 | 4.33 | 0.30 | 28 | 2726.88 | 4.31 | 0.34 | 28 | 3005.34 | 6.04 | 0.27 |
| 38 | 3449.92 | 2.44 | 0.39 | 31 | 2056.08 | 3.52 | 0.37 | 31 | 2296.86 | 5.12 | 0.34 |
| 42 | 2492.52 | 0.55 | 0.41 | 34 | 1588.80 | 2.33 | 0.42 | 34 | 1788.42 | 4.24 | 0.29 |

Fig. 1. Results for sequence CatRobot

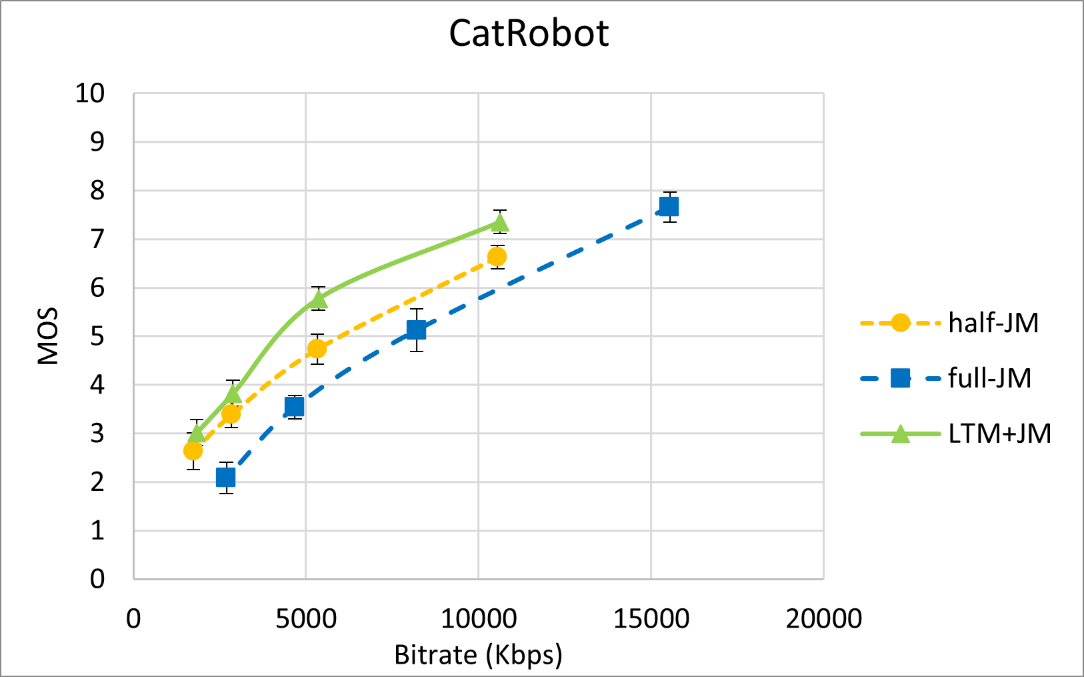


Fig. 2. Results for sequence LupoPuppet

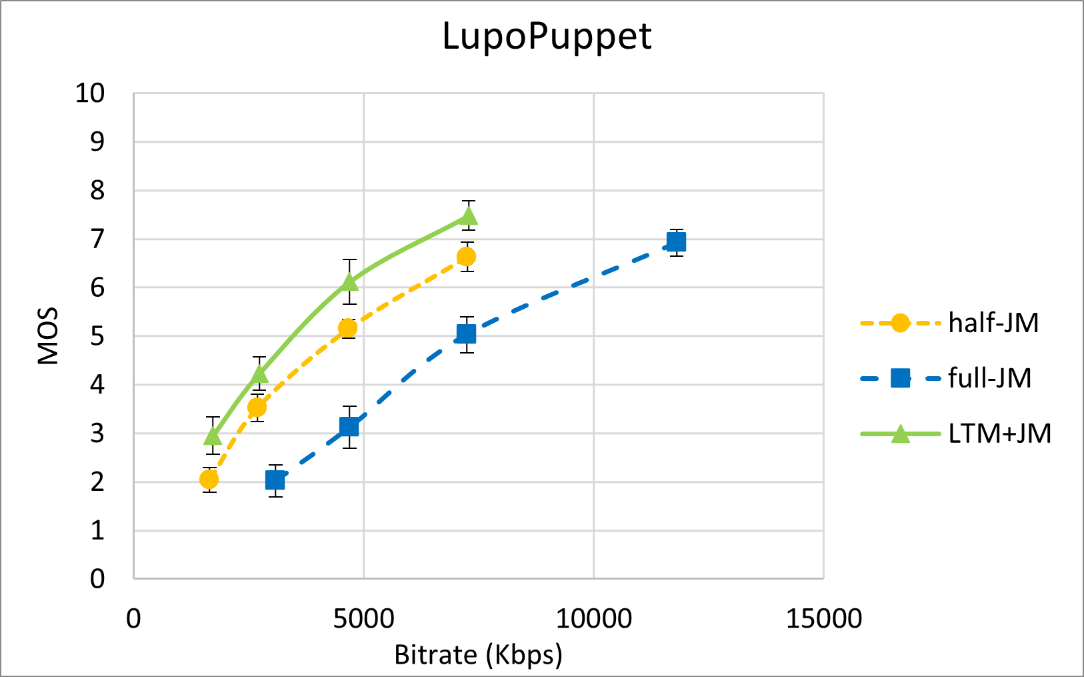


Fig. 3. Results for sequence DrivingPOVLogo

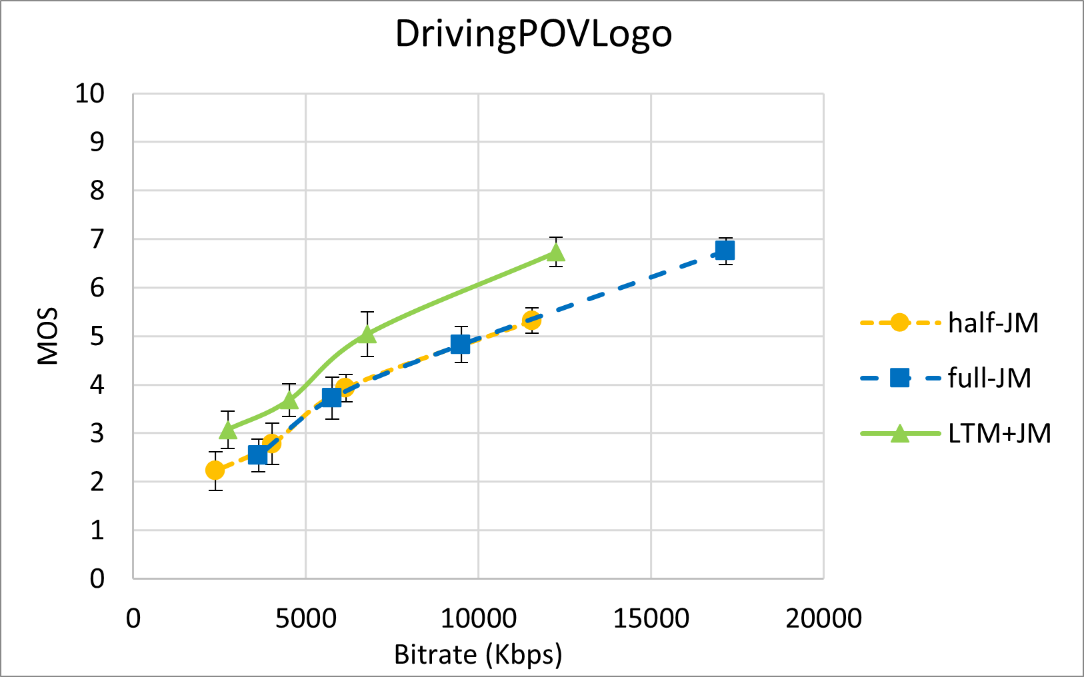
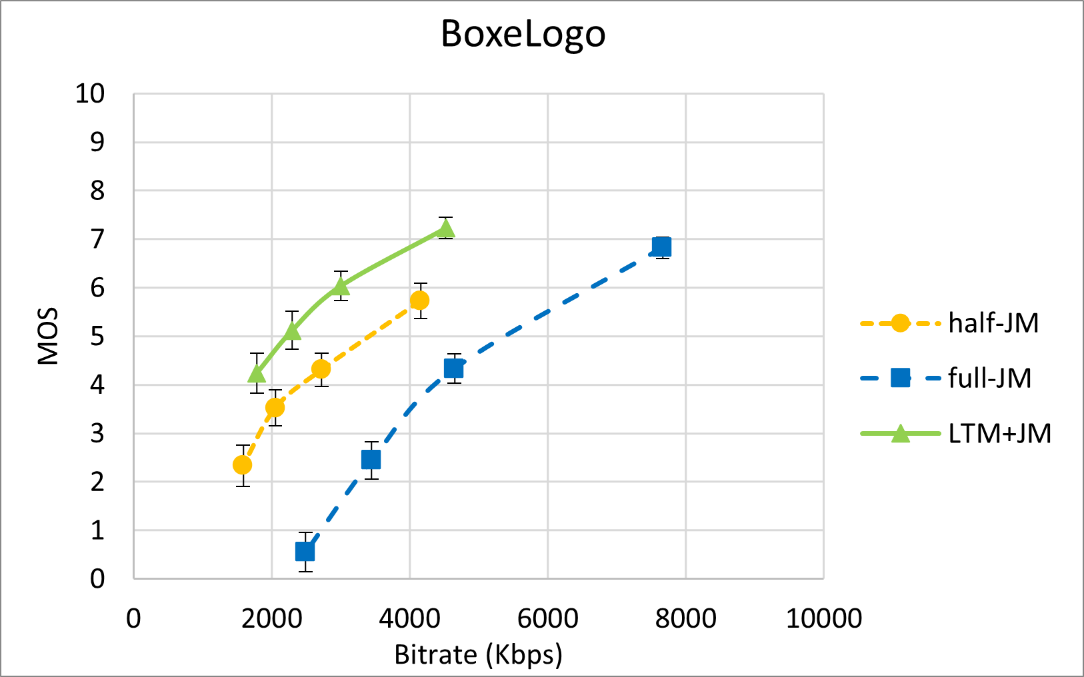


Fig. 4. Results for sequence BoxeLogo



* 1. **SDR** HD resolution – LTM over JM vs. JM

Table 6. Table of results for SDR HD sequences

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **full-JM** | | | | **half-JM** | | | | **LTM+JM** | | | |
| *sequence* | *QP* | *bitrate* | *MOS* | *CI* | *QP* | *bitrate* | *MOS* | *CI* | *QP* | *bitrate* | *MOS* | *CI* |
| TrafficLogo | 26 | 2705.64 | 7.02 | 0.48 | 24 | 1336.02 | 4.81 | 0.33 | 24 | 1724.46 | 7.48 | 0.30 |
| 30 | 1723.74 | 5.72 | 0.38 | 26 | 1041.36 | 4.12 | 0.30 | 26 | 1306.98 | 6.42 | 0.56 |
| 34 | 1139.64 | 4.73 | 0.52 | 30 | 655.68 | 3.03 | 0.44 | 30 | 796.20 | 4.41 | 0.62 |
| 40 | 636.96 | 3.12 | 0.38 | 32 | 521.64 | 2.01 | 0.42 | 32 | 618.90 | 3.40 | 0.51 |
| Starcraft | 28 | 3651.72 | 7.20 | 0.46 | 25 | 1518.87 | 4.42 | 0.30 | 25 | 2410.56 | 7.43 | 0.33 |
| 32 | 2014.02 | 5.64 | 0.61 | 27 | 1152.84 | 3.81 | 0.47 | 27 | 1677.30 | 6.39 | 0.35 |
| 36 | 1166.88 | 4.21 | 0.45 | 30 | 772.74 | 2.73 | 0.40 | 30 | 1034.40 | 4.41 | 0.30 |
| 40 | 678.06 | 2.53 | 0.55 | 33 | 517.02 | 1.72 | 0.31 | 33 | 647.10 | 3.02 | 0.57 |

Fig. 5. Results for sequence TrafficLogo

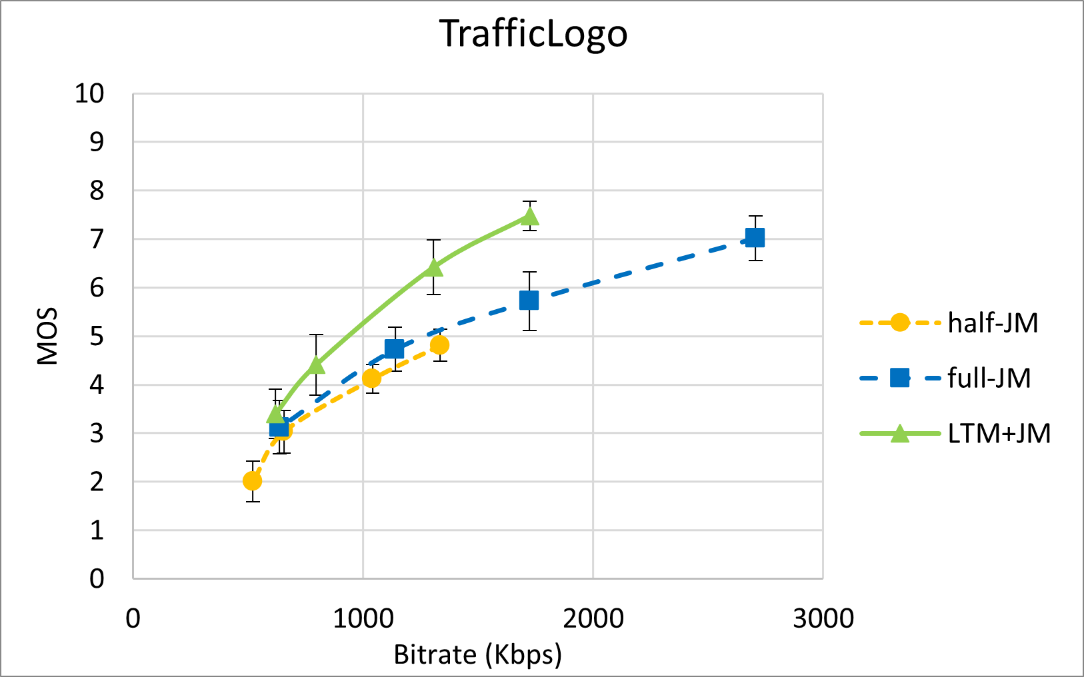
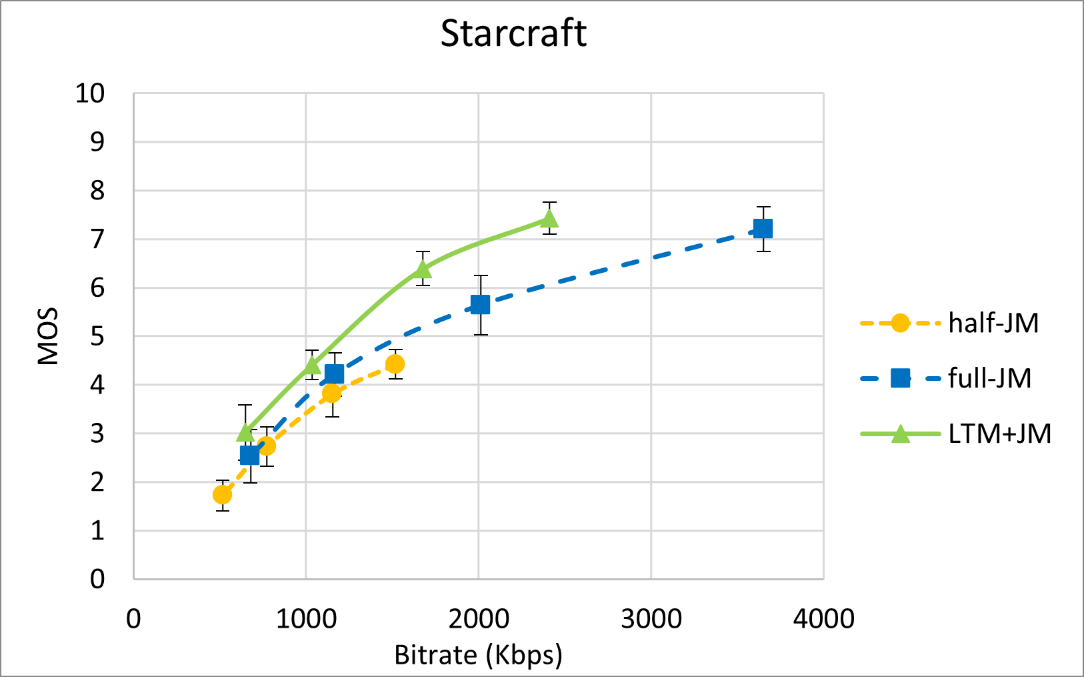


Fig. 6. Results for sequence Starcraft



* 1. **SDR** UHD resolution – LTM over HM vs. HM

Table 7. Table of results for SDR UHD sequences

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **full-HM** | | | | **half-HM** | | | | **LTM+HM** | | | |
| *sequence* | *QP* | *bitrate* | *MOS* | *CI* | *QP* | *bitrate* | *MOS* | *CI* | *QP* | *bitrate* | *MOS* | *CI* |
| CatRobot | 27 | 10803.96 | 7.65 | 0.29 | 23 | 8252.52 | 6.82 | 0.38 | 23 | 8346.96 | 7.82 | 0.27 |
| 32 | 5453.52 | 6.27 | 0.44 | 27 | 4675.38 | 5.63 | 0.44 | 27 | 4712.64 | 6.93 | 0.33 |
| 37 | 2960.58 | 4.65 | 0.41 | 32 | 2432.94 | 3.81 | 0.43 | 32 | 2483.34 | 5.82 | 0.44 |
| 42 | 1651.90 | 2.94 | 0.33 | 36 | 1455.66 | 2.38 | 0.35 | 36 | 1568.34 | 4.74 | 0.34 |
| LupoPuppet | 30 | 8552.70 | 7.55 | 0.44 | 25 | 6531.25 | 7.25 | 0.47 | 25 | 6589.80 | 8.31 | 0.31 |
| 34 | 4646.94 | 5.83 | 0.37 | 29 | 3376.65 | 5.28 | 0.29 | 29 | 3515.45 | 6.12 | 0.33 |
| 38 | 2610.12 | 4.02 | 0.49 | 32 | 2042.50 | 3.63 | 0.30 | 32 | 2242.45 | 4.53 | 0.30 |
| 42 | 1504.38 | 2.73 | 0.38 | 36 | 1104.55 | 2.63 | 0.33 | 36 | 1449.10 | 3.51 | 0.30 |
| DrivingPOVLogo | 28 | 12624.66 | 8.01 | 0.60 | 24 | 9477.84 | 6.91 | 0.46 | 24 | 10300.68 | 8.43 | 0.36 |
| 32 | 6697.32 | 5.92 | 0.34 | 28 | 4912.68 | 4.92 | 0.37 | 28 | 5537.52 | 6.12 | 0.40 |
| 38 | 2927.70 | 4.05 | 0.36 | 33 | 2377.20 | 3.12 | 0.46 | 33 | 2767.86 | 4.81 | 0.28 |
| 42 | 1769.34 | 3.02 | 0.45 | 36 | 1583.70 | 2.11 | 0.24 | 36 | 1812.36 | 3.43 | 0.29 |
| BoxeLogo | 30 | 2682.48 | 7.53 | 0.41 | 25 | 2239.38 | 6.53 | 0.19 | 25 | 2601.54 | 7.91 | 0.41 |
| 34 | 1783.86 | 5.71 | 0.52 | 29 | 1390.26 | 4.18 | 0.26 | 29 | 1672.74 | 6.73 | 0.47 |
| 38 | 1216.14 | 3.77 | 0.28 | 33 | 908.82 | 2.33 | 0.39 | 33 | 1115.34 | 4.68 | 0.40 |
| 42 | 844.98 | 2.18 | 0.48 | 37 | 615.30 | 1.39 | 0.21 | 37 | 753.54 | 3.12 | 0.34 |

Fig. 7. Results for sequence CatRobot

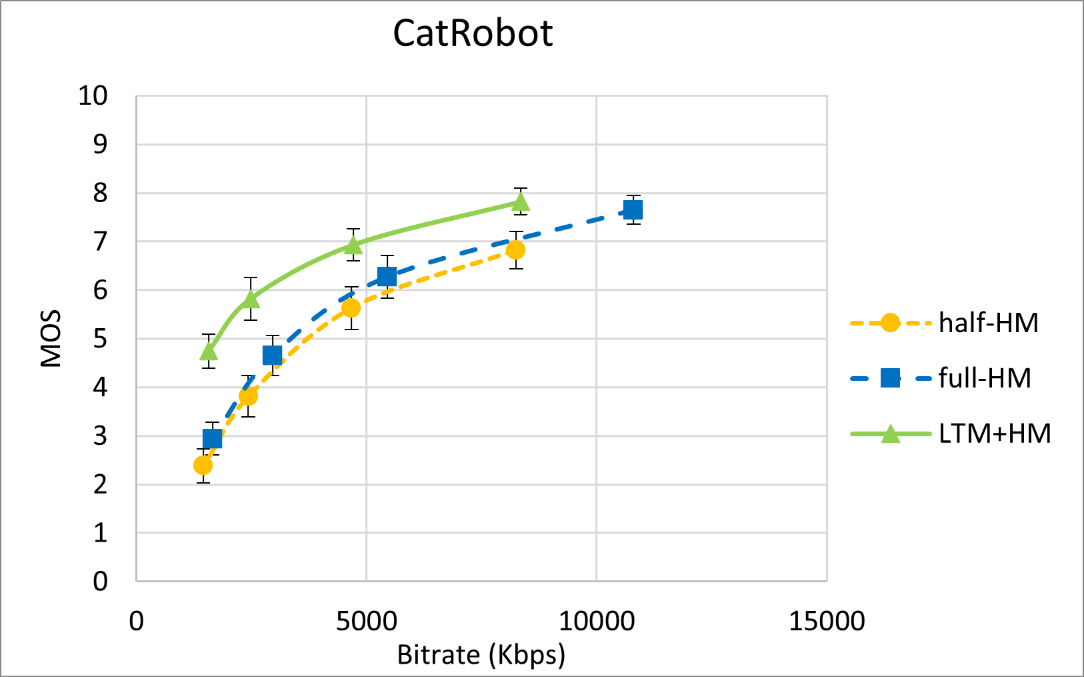


Fig. 8. Results for sequence LupoPuppet

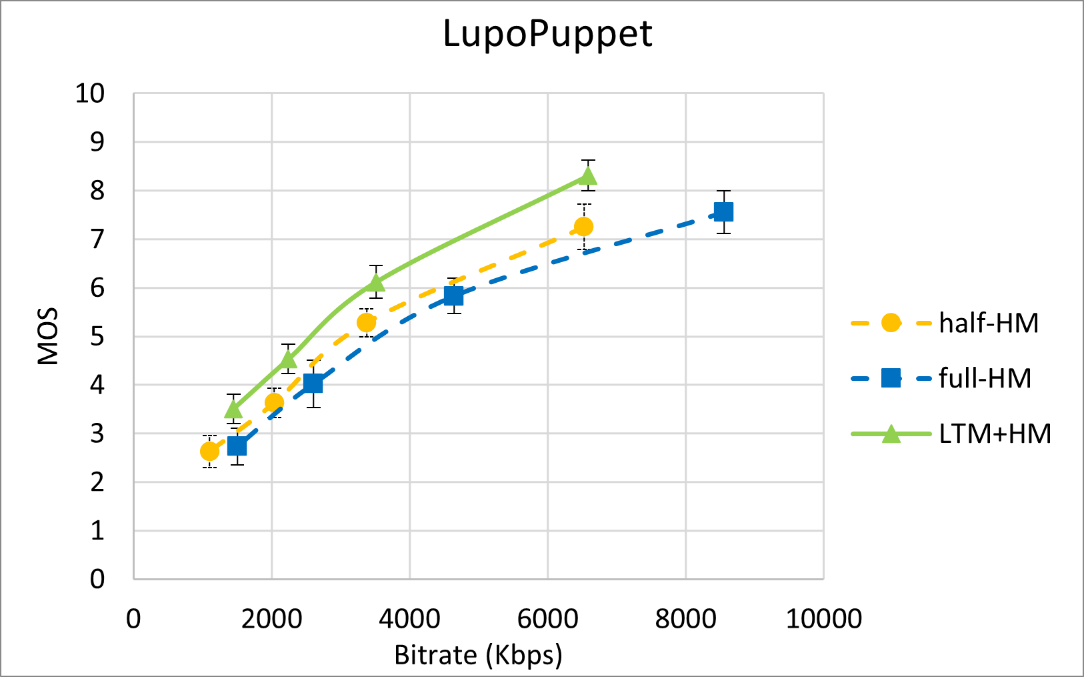


Fig. 9. Results for sequence DrivingPOVLogo

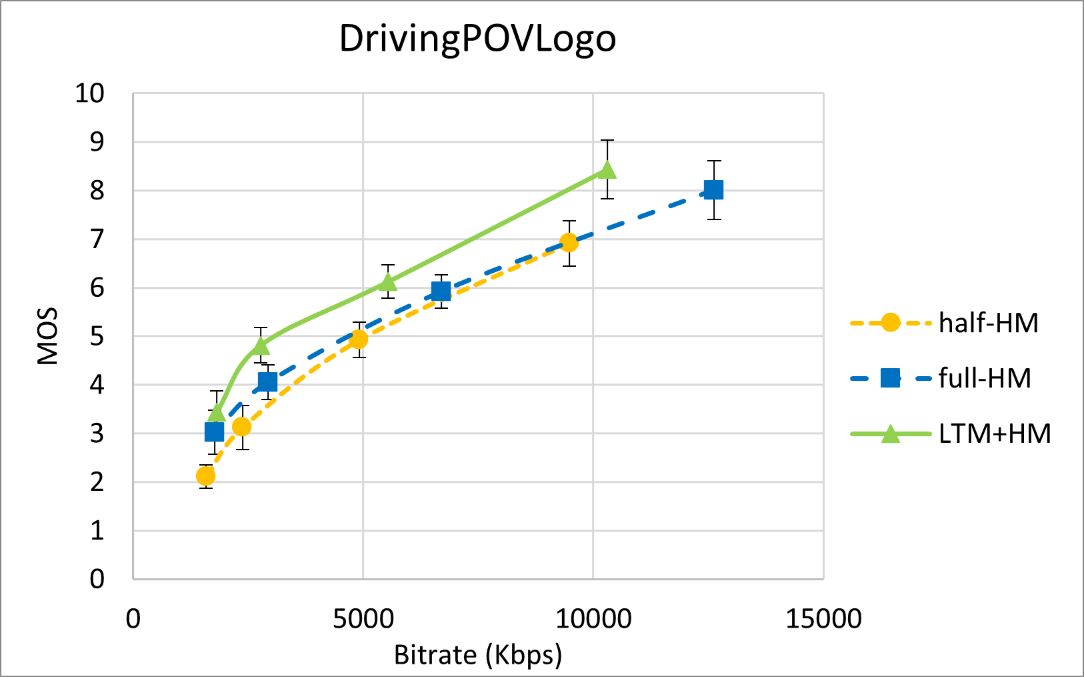
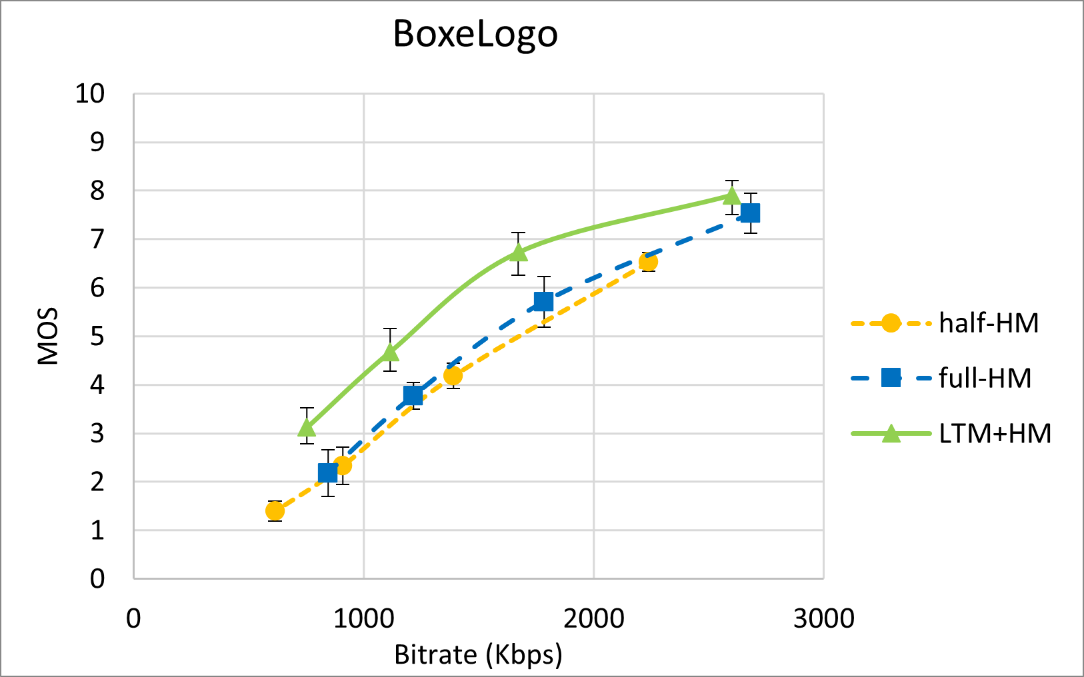


Fig. 10. Results for sequence BoxeLogo



* 1. **SDR** HD resolution – LTM over HM vs. HM

Table 8. Table of results for SDR HD sequences

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **full-HM** | | | | **half-HM** | | | | **LTM+HM** | | | |
| *sequence* | *QP* | *bitrate* | *MOS* | *CI* | *QP* | *bitrate* | *MOS* | *CI* | *QP* | *bitrate* | *MOS* | *CI* |
| TrafficLogo | 26 | 1755.66 | 7.83 | 0.19 | 22 | 1296.30 | 6.62 | 0.29 | 22 | 1529.10 | 8.02 | 0.39 |
| 30 | 1058.16 | 5.98 | 0.30 | 25 | 870.30 | 5.03 | 0.27 | 25 | 1040.28 | 6.48 | 0.41 |
| 34 | 671.34 | 4.15 | 0.26 | 28 | 581.22 | 3.62 | 0.32 | 28 | 684.06 | 5.02 | 0.28 |
| 38 | 431.22 | 3.12 | 0.31 | 33 | 317.76 | 1.87 | 0.39 | 33 | 391.08 | 3.65 | 0.29 |
| Starcraft | 26 | 4493.76 | 7.72 | 0.21 | 22 | 2093.46 | 6.04 | 0.38 | 22 | 2707.62 | 7.65 | 0.28 |
| 30 | 2400.90 | 6.38 | 0.25 | 25 | 1364.28 | 4.82 | 0.33 | 25 | 1821.60 | 6.78 | 0.28 |
| 34 | 1312.02 | 4.82 | 0.27 | 28 | 888.84 | 3.53 | 0.33 | 28 | 1136.52 | 5.22 | 0.47 |
| 38 | 728.46 | 3.55 | 0.30 | 32 | 458.94 | 2.34 | 0.31 | 32 | 665.94 | 3.81 | 0.33 |

Fig. 11. Results for sequence TrafficLogo

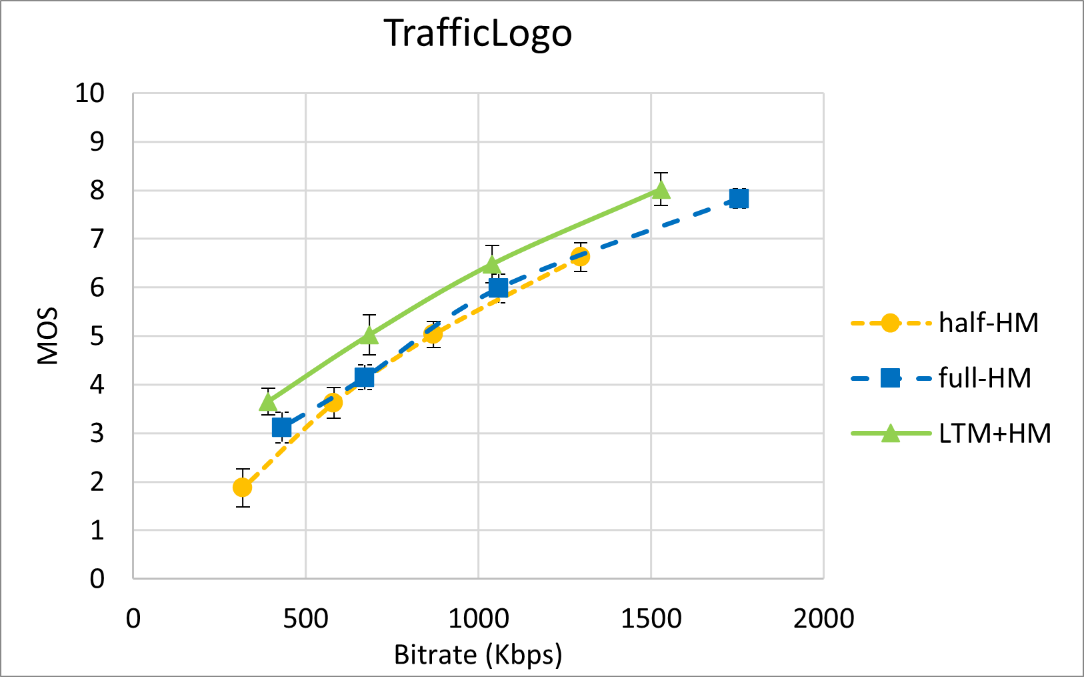
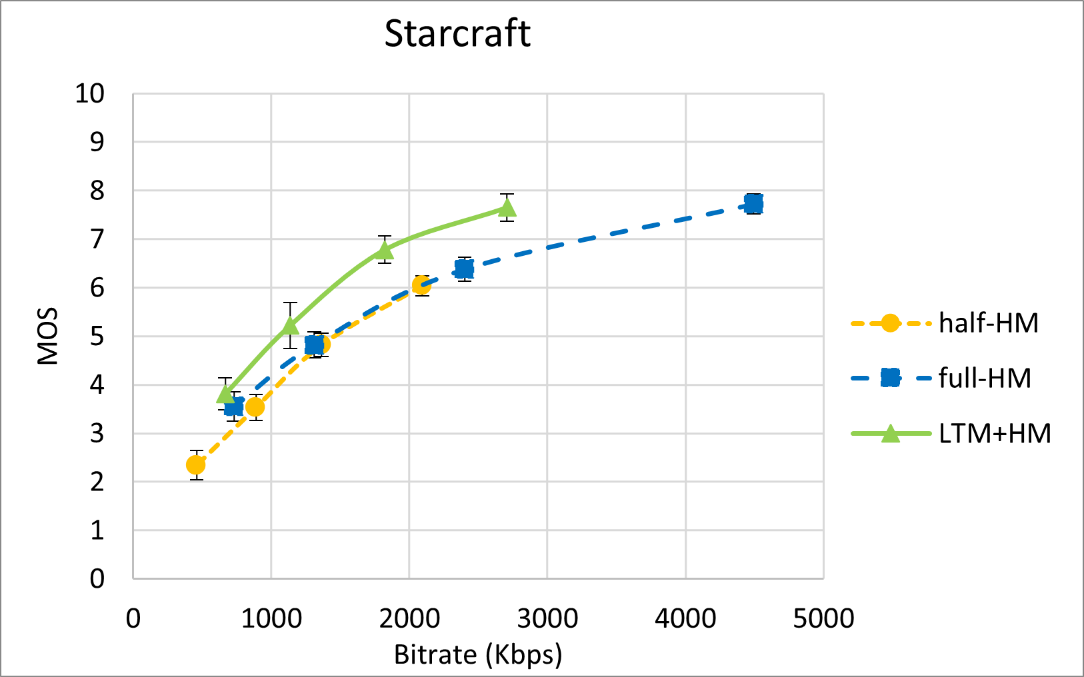


Fig. 12. Results for sequence Starcraft



* 1. **SDR** UHD resolution – LTM over ETM vs. ETM

Table 9. Table of results for SDR UHD sequences

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **full-ETM** | | | | **half-ETM** | | | | **LTM+ETM** | | | |
| *sequence* | *QP* | *bitrate* | *MOS* | *CI* | *QP* | *bitrate* | *MOS* | *CI* | *QP* | *bitrate* | *MOS* | *CI* |
| CatRobot | 28 | 7829.00 | 7.86 | 0.39 | 23 | 7203.00 | 6.94 | 0.46 | 23 | 7232.94 | 7.97 | 0.36 |
| 36 | 2627.00 | 6.03 | 0.38 | 30 | 2657.76 | 5.13 | 0.21 | 30 | 2670.78 | 6.18 | 0.34 |
| 40 | 1594.00 | 4.06 | 0.32 | 34 | 1555.98 | 3.63 | 0.17 | 34 | 1569.60 | 5.22 | 0.30 |
| 44 | 999.00 | 2.28 | 0.25 | 38 | 905.16 | 2.00 | 0.26 | 38 | 922.38 | 2.91 | 0.21 |
| DrivingPOVLogo | 28 | 11217.00 | 8.11 | 0.26 | 23 | 9996.70 | 7.17 | 0.27 | 23 | 10942.20 | 7.94 | 0.24 |
| 32 | 5880.00 | 6.78 | 0.41 | 27 | 5072.04 | 5.33 | 0.21 | 27 | 5720.34 | 6.94 | 0.32 |
| 38 | 2511.00 | 4.72 | 0.20 | 32 | 2392.44 | 3.72 | 0.33 | 32 | 2761.98 | 5.50 | 0.23 |
| 42 | 1483.00 | 3.33 | 0.54 | 36 | 1354.80 | 2.06 | 0.35 | 36 | 1513.74 | 4.22 | 0.28 |
| BarScene | 28 | 2614.00 | 8.11 | 0.34 | 23 | 1954.26 | 6.53 | 0.30 | 23 | 1979.70 | 7.71 | 0.37 |
| 36 | 595.00 | 6.18 | 0.36 | 31 | 516.60 | 5.06 | 0.25 | 31 | 529.56 | 6.56 | 0.37 |
| 40 | 368.00 | 4.22 | 0.28 | 34 | 355.38 | 3.94 | 0.22 | 34 | 368.40 | 5.33 | 0.35 |
| 44 | 239.00 | 1.89 | 0.36 | 38 | 218.22 | 2.33 | 0.33 | 38 | 231.42 | 3.22 | 0.30 |
| BoxeLogo | 28 | 2853.00 | 7.17 | 0.38 | 24 | 2216.16 | 5.89 | 0.33 | 24 | 2619.24 | 7.22 | 0.32 |
| 32 | 1796.00 | 6.17 | 0.46 | 27 | 1512.30 | 5.06 | 0.28 | 27 | 1823.52 | 6.56 | 0.34 |
| 38 | 975.00 | 4.11 | 0.33 | 32 | 870.18 | 3.94 | 0.41 | 32 | 1072.38 | 5.33 | 0.30 |
| 42 | 664.00 | 1.72 | 0.29 | 36 | 575.58 | 2.33 | 0.26 | 36 | 698.76 | 3.22 | 0.32 |

Fig. 13. Results for sequence DrivingPOVLogo

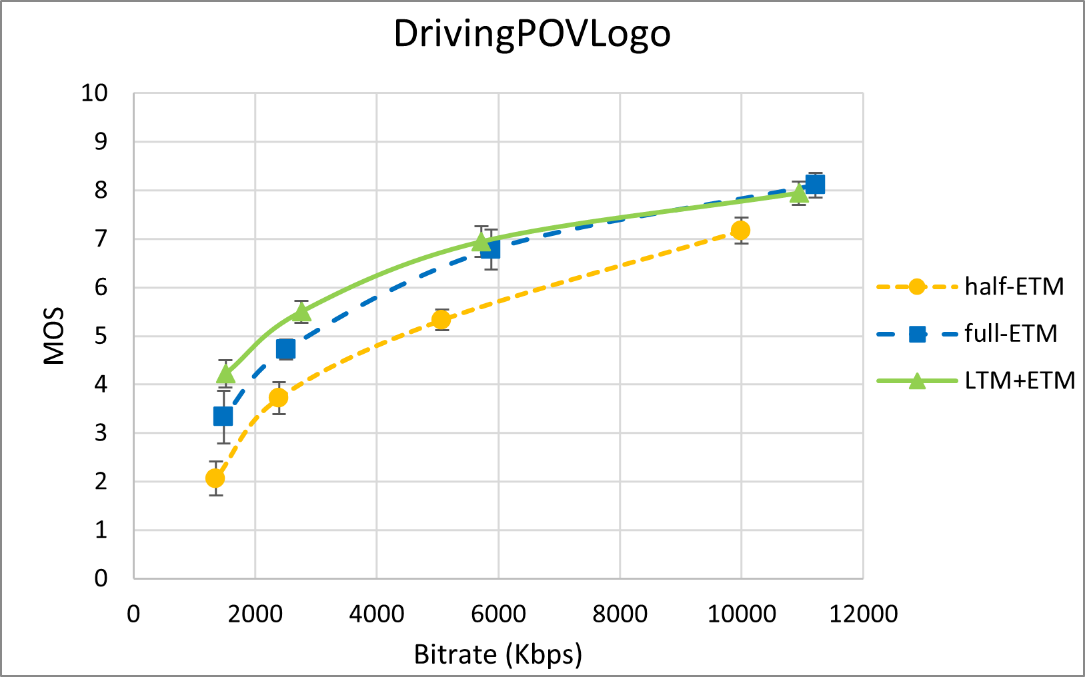


Fig. 14. Results for sequence CatRobot

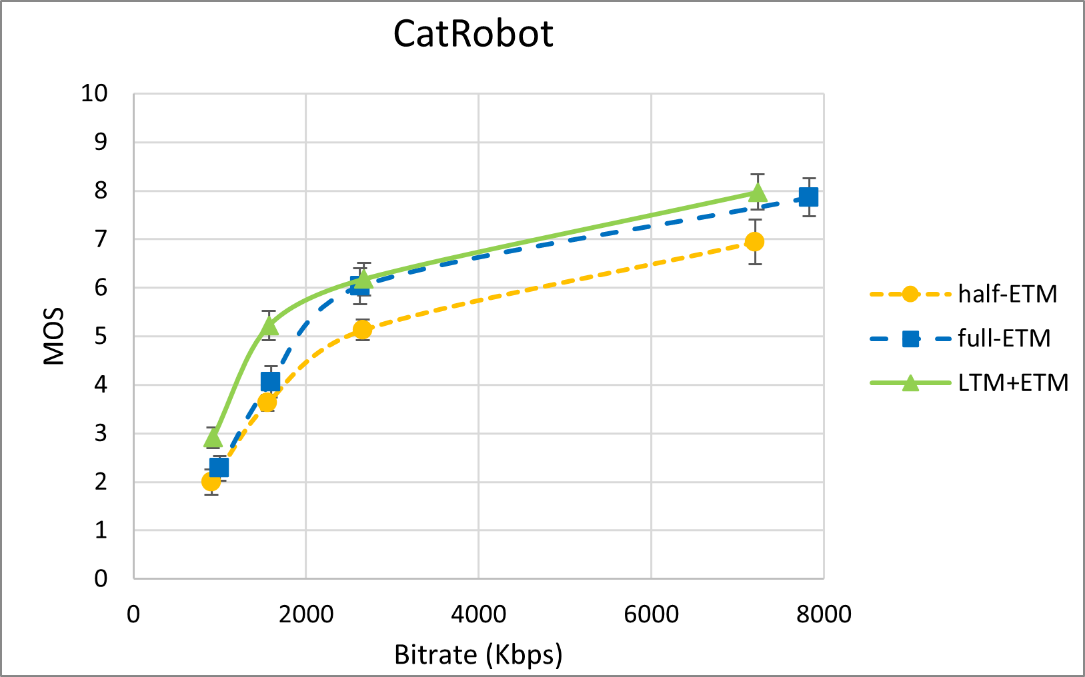


Fig. 15. Results for sequence BarScene

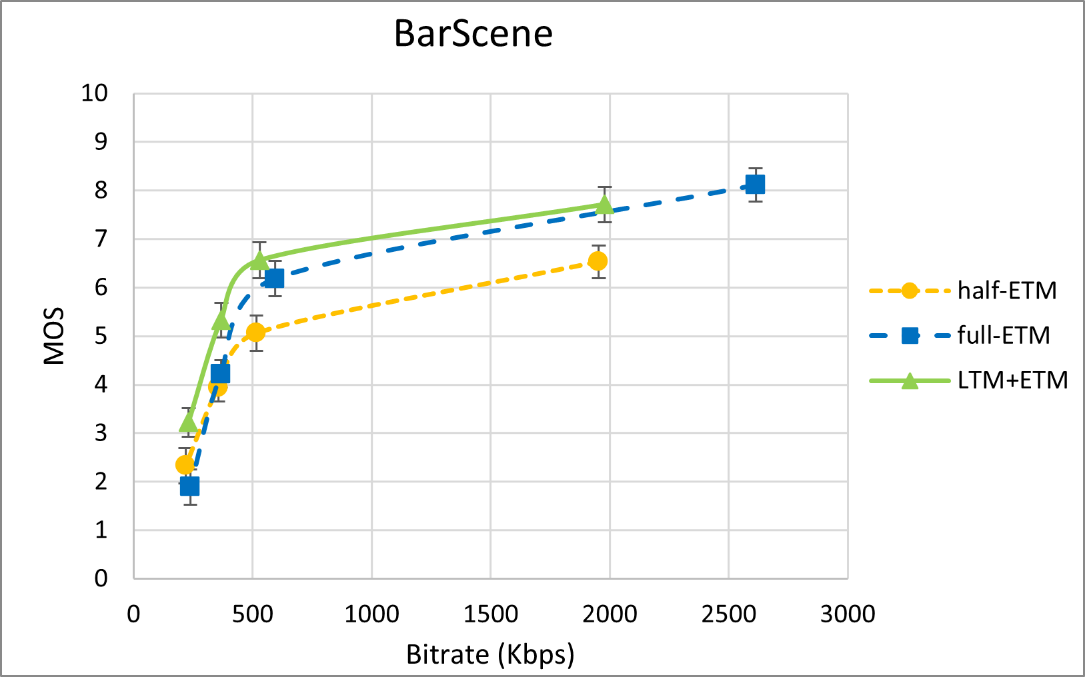
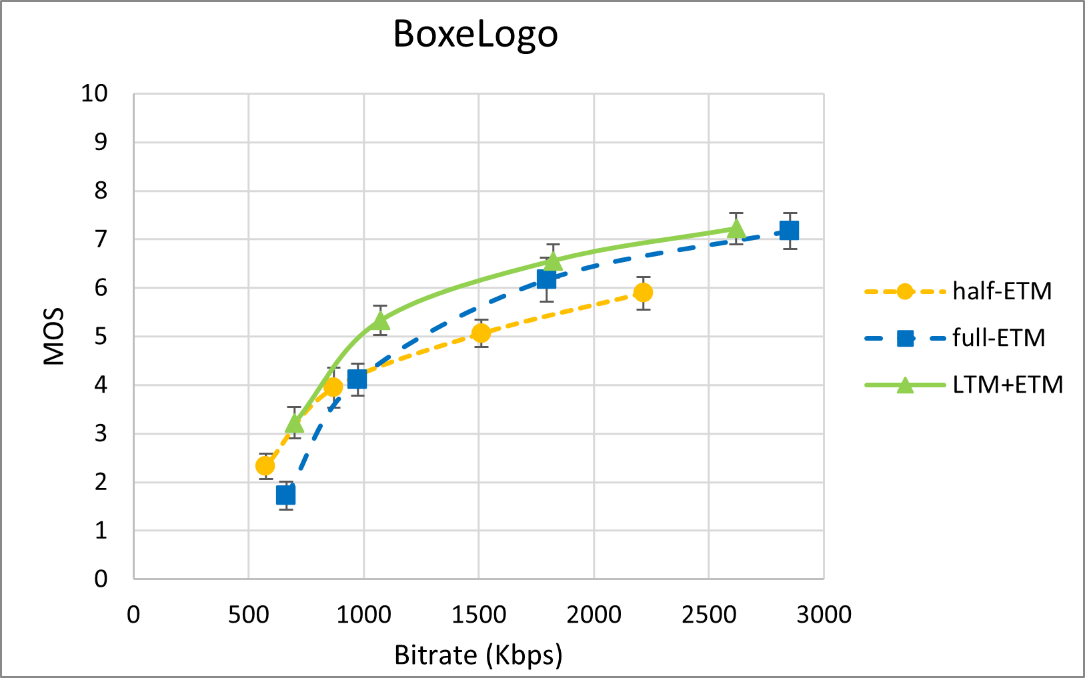


Fig. 16. Results for sequence BoxeLogo



* 1. **SDR** HD resolution – LTM over ETM vs. ETM

Table 10. Table of results for SDR HD sequences

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **full-ETM** | | | | **half-ETM** | | | | **LTM+ETM** | | | |
| *sequence* | *QP* | *bitrate* | *MOS* | *CI* | *QP* | *bitrate* | *MOS* | *CI* | *QP* | *bitrate* | *MOS* | *CI* |
| TrafficLogo | 26 | 1625.00 | 7.50 | 0.27 | 22 | 1186.80 | 5.68 | 0.33 | 22 | 1563.60 | 7.33 | 0.37 |
| 30 | 977.00 | 6.28 | 0.29 | 25 | 799.92 | 4.68 | 0.25 | 25 | 1058.22 | 6.71 | 0.38 |
| 34 | 612.00 | 4.61 | 0.31 | 28 | 534.78 | 3.17 | 0.31 | 28 | 654.66 | 5.17 | 0.31 |
| 38 | 387.00 | 2.50 | 0.34 | 33 | 292.31 | 1.72 | 0.42 | 33 | 342.48 | 2.89 | 0.30 |
| Starcraft | 26 | 4210.00 | 7.78 | 0.27 | 22 | 1980.30 | 5.43 | 0.48 | 22 | 3164.52 | 7.52 | 0.34 |
| 30 | 2218.00 | 6.14 | 0.27 | 24 | 1463.40 | 4.79 | 0.41 | 24 | 2205.48 | 6.22 | 0.41 |
| 34 | 1227.00 | 4.44 | 0.34 | 27 | 970.92 | 3.76 | 0.31 | 27 | 1402.38 | 5.03 | 0.43 |
| 38 | 681.00 | 1.39 | 0.37 | 30 | 655.26 | 2.11 | 0.33 | 30 | 795.00 | 3.45 | 0.34 |

Fig. 17. Results for sequence TrafficLogo

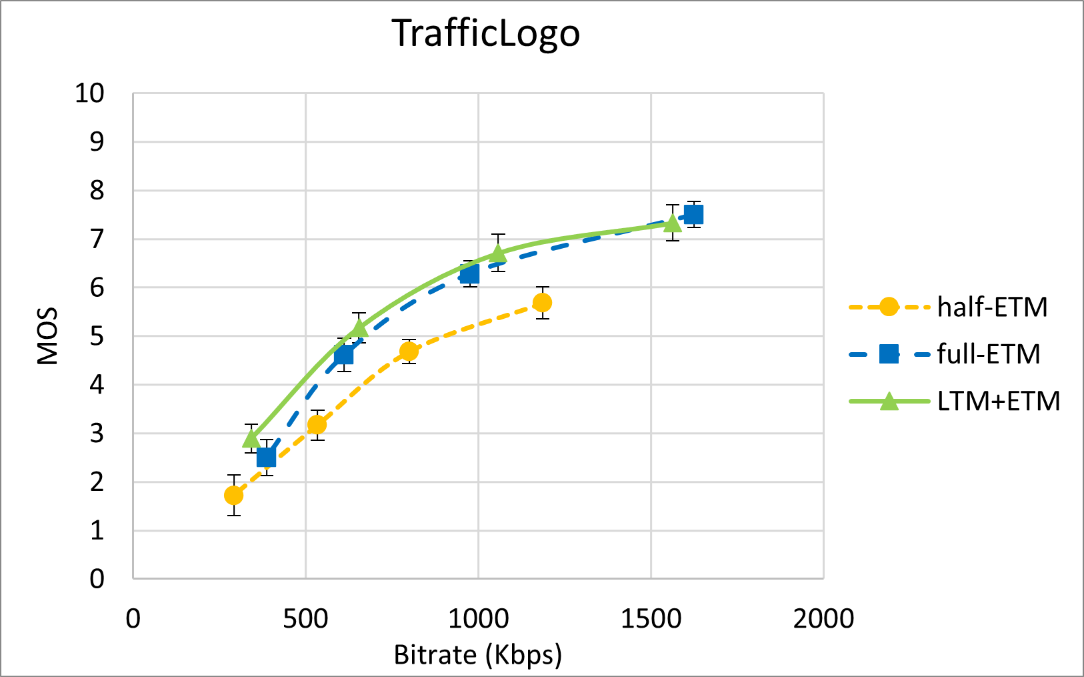
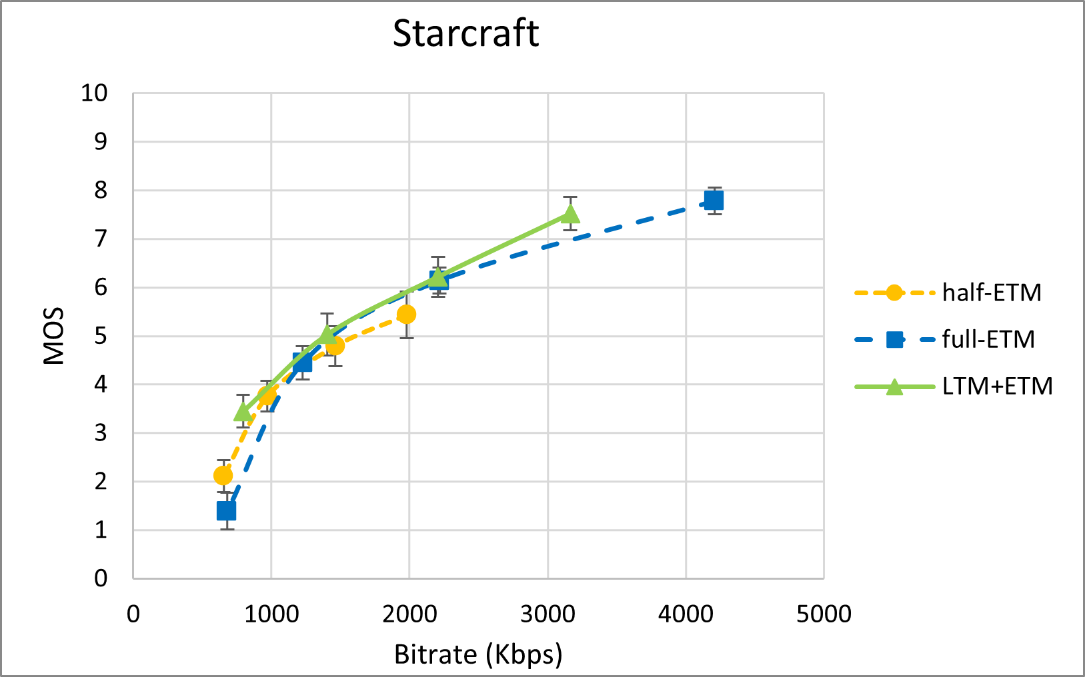


Fig. 18. Results for sequence Starcraft



* 1. **SDR** UHD resolution – LTM over VTM vs. VTM

Table 11. Table of results for SDR UHD sequences

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **full-VTM** | | | | **half-VTM** | | | | **LTM+VTM** | | | |
| *sequence* | *QP* | *bitrate* | *MOS* | *CI* | *QP* | *bitrate* | *MOS* | *CI* | *QP* | *bitrate* | *MOS* | *CI* |
| Marathon | 30 | 9256.35 | 7.97 | 0.25 | 26 | 7580.79 | 7.04 | 0.30 | 26 | 7599.06 | 7.97 | 0.27 |
| 34 | 5362.65 | 6.63 | 0.33 | 29 | 4762.11 | 5.58 | 0.22 | 29 | 4770.36 | 6.64 | 0.30 |
| 38 | 3100.83 | 4.83 | 0.36 | 32 | 2997.54 | 4.38 | 0.38 | 32 | 3005.76 | 5.42 | 0.28 |
| 42 | 1777.83 | 3.02 | 0.38 | 36 | 1673.64 | 2.25 | 0.30 | 36 | 1688.61 | 3.61 | 0.31 |
| MountainBay2 | 34 | 1672.14 | 6.63 | 0.26 | 29 | 1498.41 | 6.26 | 0.29 | 29 | 1505.88 | 6.79 | 0.42 |
| 38 | 860.58 | 5.61 | 0.35 | 33 | 777.09 | 5.05 | 0.27 | 33 | 786.03 | 5.75 | 0.27 |
| 42 | 442.47 | 3.71 | 0.34 | 37 | 408.30 | 3.49 | 0.29 | 37 | 419.37 | 4.21 | 0.24 |
| 46 | 226.91 | 1.79 | 0.37 | 40 | 248.70 | 2.03 | 0.22 | 40 | 260.34 | 2.50 | 0.26 |
| DrivingPOVLogo | 28 | 9413.40 | 8.66 | 0.30 | 24 | 6850.98 | 7.49 | 0.29 | 24 | 7675.50 | 8.79 | 0.26 |
| 32 | 5073.78 | 7.78 | 0.24 | 27 | 4279.14 | 6.13 | 0.29 | 27 | 4915.26 | 7.91 | 0.29 |
| 38 | 2218.44 | 5.88 | 0.28 | 33 | 1712.58 | 4.63 | 0.28 | 33 | 2078.94 | 5.97 | 0.21 |
| 42 | 1311.42 | 4.57 | 0.28 | 37 | 993.90 | 3.66 | 0.26 | 37 | 1209.36 | 4.63 | 0.31 |
| BoxeLogo | 30 | 1996.44 | 7.29 | 0.35 | 27 | 1286.34 | 5.08 | 0.40 | 27 | 1603.74 | 7.67 | 0.46 |
| 34 | 1326.42 | 6.28 | 0.34 | 30 | 917.16 | 4.38 | 0.29 | 30 | 1172.64 | 6.42 | 0.22 |
| 38 | 908.58 | 4.91 | 0.24 | 32 | 742.80 | 3.88 | 0.25 | 32 | 939.84 | 5.67 | 0.27 |
| 42 | 632.04 | 3.92 | 0.39 | 37 | 455.46 | 2.62 | 0.24 | 37 | 586.32 | 4.34 | 0.26 |

Fig. 19. Results for sequence Marathon

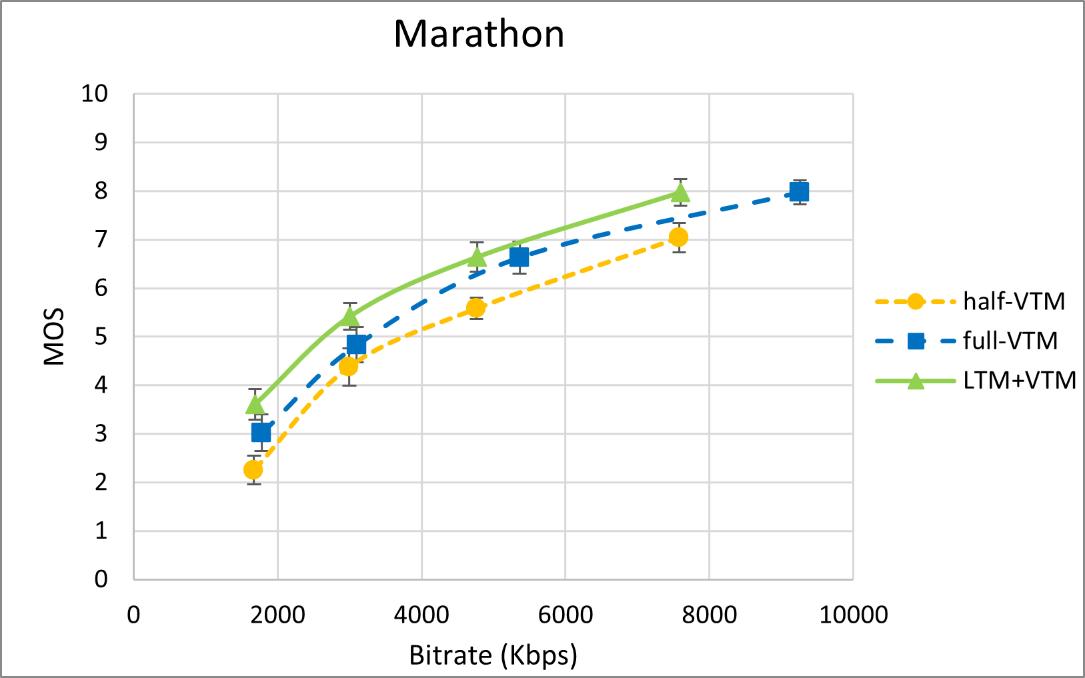


Fig. 20. Results for sequence MountainBay2

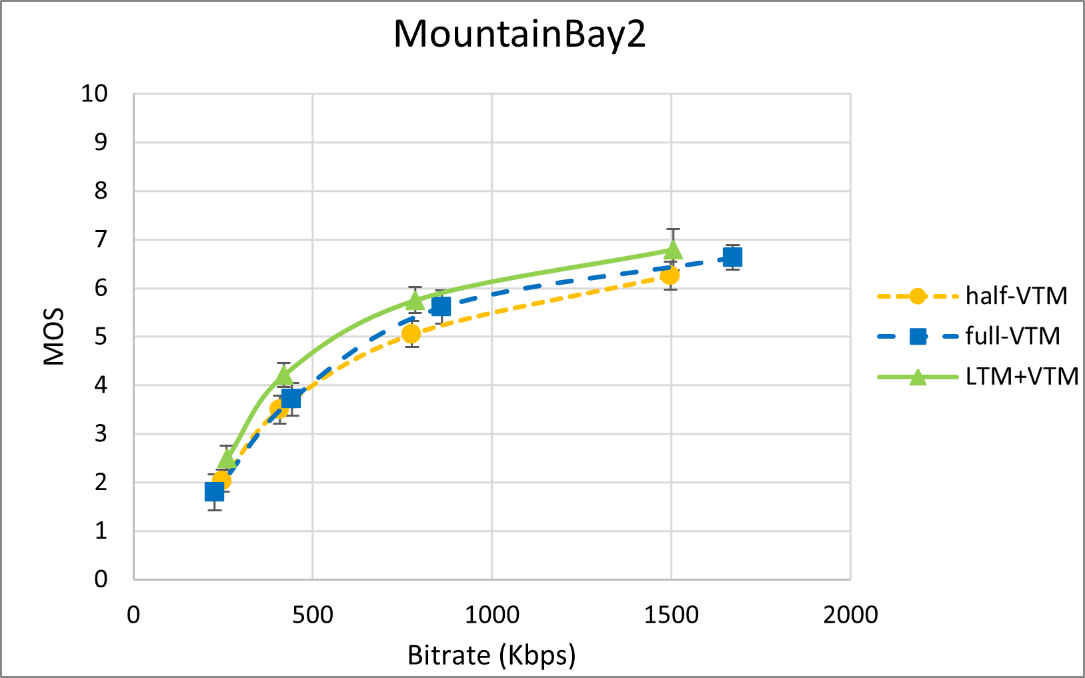


Fig. 21. Results for sequence DrivingPOVLogo

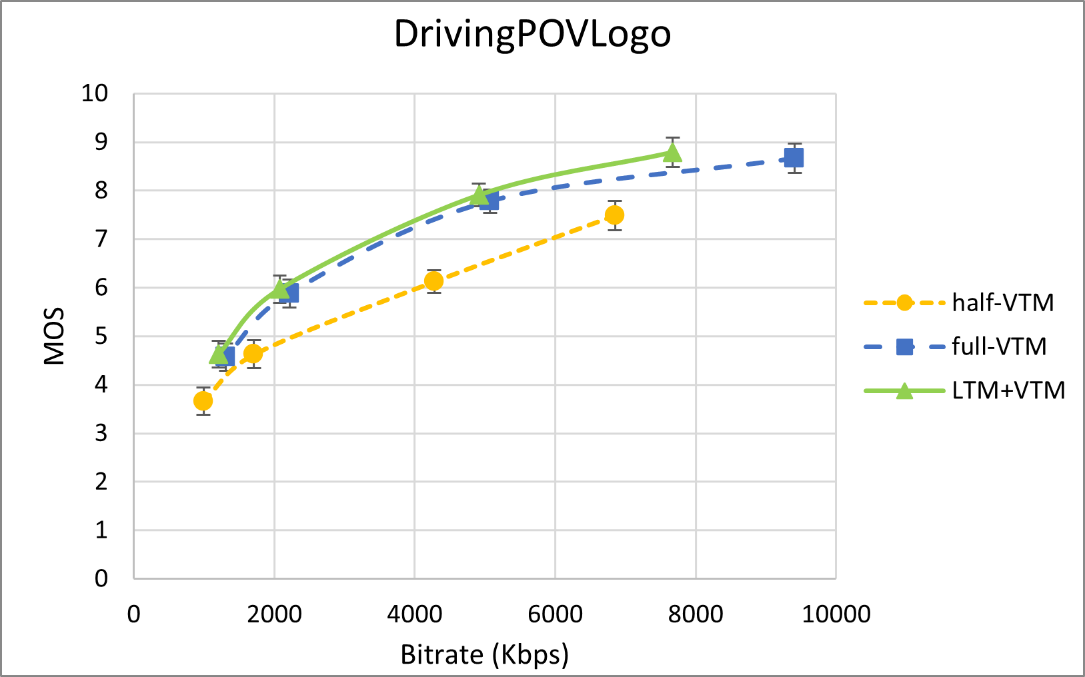
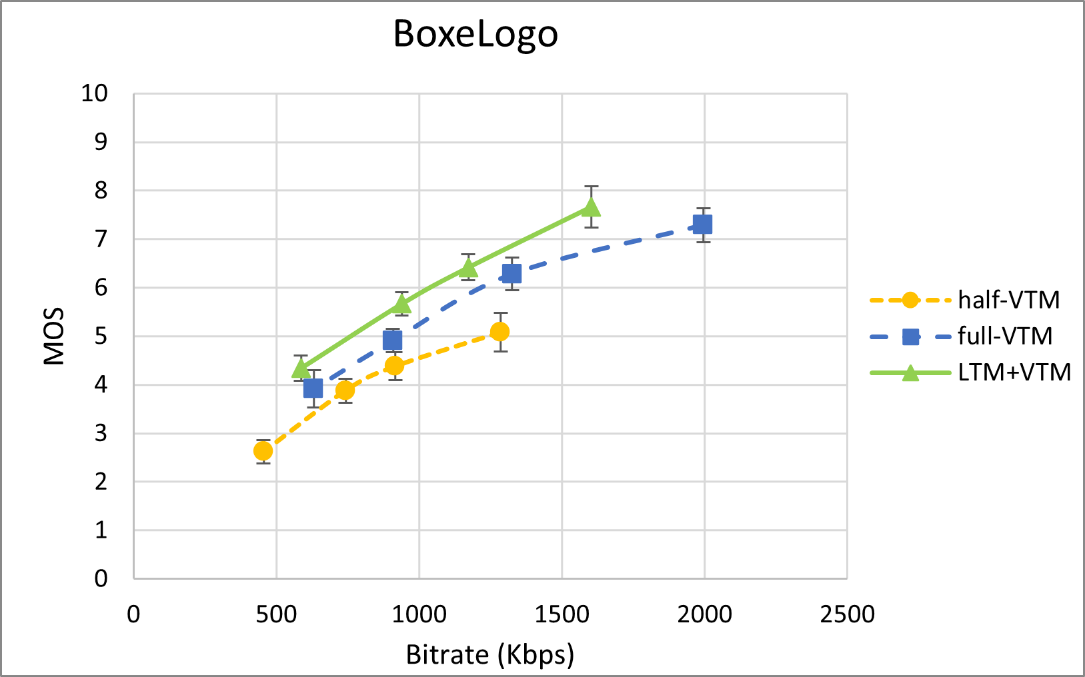


Fig. 22. Results for sequence BoxeLogo



* 1. **SDR** HD resolution – LTM over VTM vs. VTM

Table 12. Table of results for SDR HD sequences

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **full-VTM** | | | | **half-VTM** | | | | **LTM+VTM** | | | |
| *sequence* | *QP* | *bitrate* | *MOS* | *CI* | *QP* | *bitrate* | *MOS* | *CI* | *QP* | *bitrate* | *MOS* | *CI* |
| TrafficLogo | 26 | 1403.46 | 7.92 | 0.49 | 22 | 1008.48 | 6.33 | 0.38 | 22 | 1233.78 | 7.96 | 0.41 |
| 30 | 845.40 | 6.17 | 0.43 | 24 | 764.40 | 5.25 | 0.30 | 24 | 967.92 | 7.01 | 0.48 |
| 34 | 526.44 | 4.75 | 0.39 | 29 | 399.24 | 3.88 | 0.31 | 29 | 502.38 | 4.92 | 0.32 |
| 38 | 338.46 | 3.42 | 0.29 | 31 | 306.84 | 2.54 | 0.40 | 31 | 388.02 | 3.91 | 0.39 |
| Starcraft | 26 | 3513.72 | 7.67 | 0.28 | 22 | 1664.94 | 5.45 | 0.35 | 22 | 2279.64 | 7.47 | 0.42 |
| 30 | 1867.56 | 6.21 | 0.31 | 24 | 1224.78 | 5.05 | 0.23 | 24 | 1691.58 | 6.59 | 0.48 |
| 34 | 1015.08 | 4.73 | 0.35 | 27 | 810.30 | 3.54 | 0.16 | 27 | 1086.24 | 5.16 | 0.33 |
| 38 | 575.16 | 3.58 | 0.29 | 30 | 543.72 | 2.86 | 0.22 | 30 | 710.52 | 4.28 | 0.29 |

Fig. 23. Results for sequence TrafficLogo

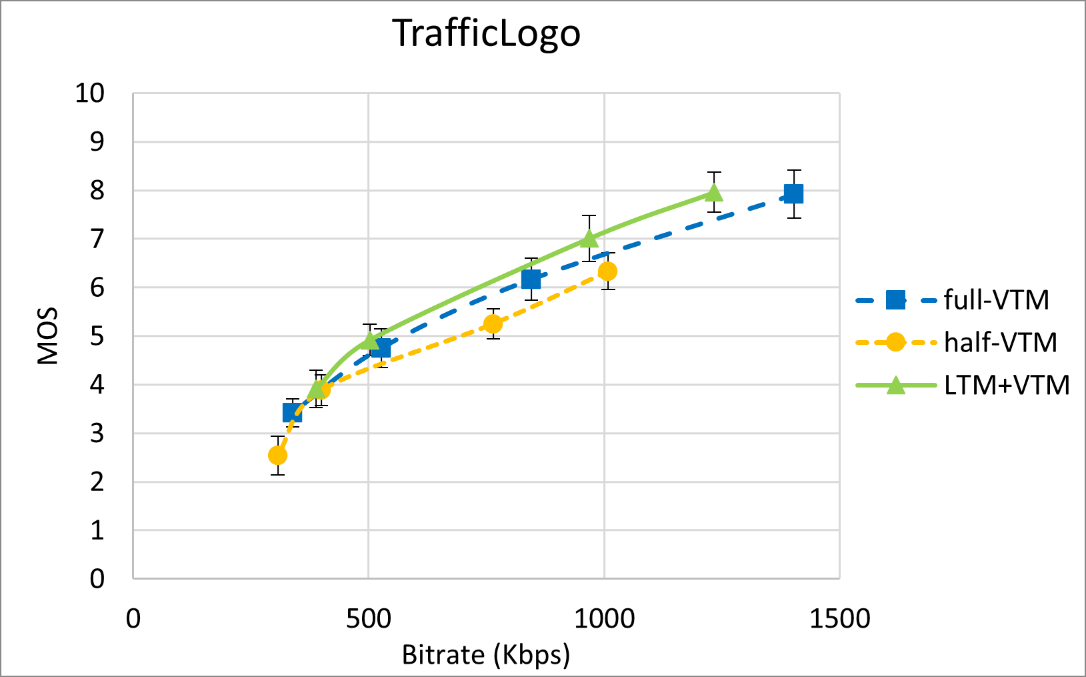
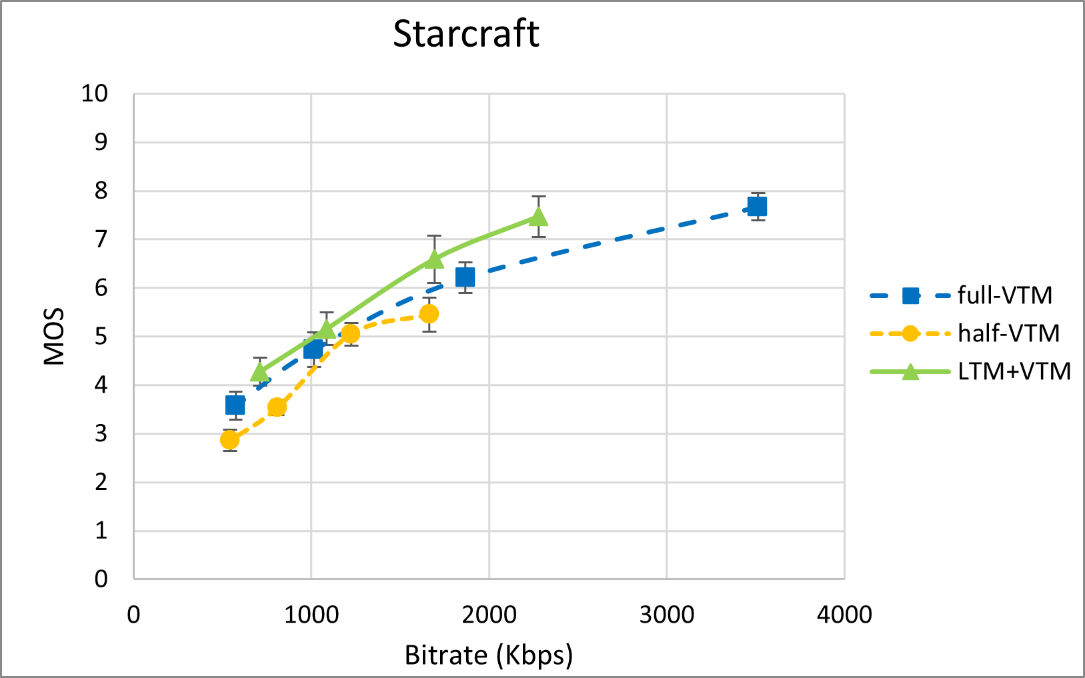


Fig. 24. Results for sequence Starcraft



# MOS BD-rate

In this section, the average bit rate savings of LCEVC profiles compared to references for each sequence were computed from the MOS vs. bit rate data in the same manner that was done in [3] to further quantify the bit rate savings achieved. The bit rate savings are averaged over the whole range where the same MOS scores for LCEVC, AVC (Full Resolution), AVC (Half Resolution), HEVC (Full Resolution), HEVC (Half Resolution), EVC (Half Resolution), and VVC (Half Resolution) can be interpolated from subjective test results shown in the plots in Section 3.

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, as such the MOS BD-rate savings are not reported. However, the test results tend to indicate an overall benefit when using LCEVC with these two codecs.

For the tests indicated as Half Resolution, the half-resolution anchors were used as the base layer for LCEVC and hence not all curves may overlap as much as would be ideal when calculating a BD-rate.

Tables 13 and 14 show the MOS BD-rate for the sequences in this test. The BD-rate measure described in [4] and [5] is used with MOS scores taking the place of PSNR. A piece wise cubic interpolation used in the LCEVC common conditions spreadsheet is used.

Table 13. MOS BD-rate – LTM vs JM and HM anchors at full resolution

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **LTM 5.1 vs JM 19.0** | | **BD-rate** |  | **LTM 5.1 vs HM 16.20** | | **BD-rate** |
| UHD | LupoPuppet | −53.98% |  | UHD | LupoPuppet | −31.14% |
| CatRobot | −43.83% |  | CatRobot | −41.88% |
| DrivingPOVLogo | −30.19% |  | DrivingPOVLogo | −26.01% |
| BoxeLogo | −55.61% |  | BoxeLogo | −24.44% |
| **Average** | | **−45.90%** |  | **Average** | | **−30.87%** |
|  |  |  |  |  |  |  |
| **LTM 5.1 vs JM 19.0** | | **BD-rate** |  | **LTM 5.1 vs HM 16.20** | | **BD-rate** |
| HD | TrafficLogo | −30.18% |  | HD | TrafficLogo | −18.52% |
| Starcraft | −26.75% |  | Starcraft | −29.76% |
| **Average** | | **−28.47%** |  | **Average** | | **−24.14%** |

Table 14. MOS BD-rate – LTM vs JM, HM, ETM, VTM upsampled half-resolution anchors

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **LTM 5.1 vs JM 19.0** | | **BD-rate** |  | **LTM 5.1 vs HM 16.20** | | **BD-rate** |
| UHD | LupoPuppet | −22.66% |  | UHD | LupoPuppet | −20.64% |
| CatRobot | −26.25% |  | CatRobot | −50.77% |
| DrivingPOVLogo | −26.95% |  | DrivingPOVLogo | −33.96% |
| BoxeLogo | −34.20% |  | BoxeLogo | −29.23% |
| **Average** | | **−27.52%** |  | **Average** | | **−33.65%** |
|  |  |  |  |  |  |  |
| **LTM 5.1 vs JM 19.0** | | **BD-rate** |  | **LTM 5.1 vs HM 16.20** | | **BD-rate** |
| HD | TrafficLogo | −27.80% |  | HD | TrafficLogo | −22.51% |
| Starcraft | −26.28% |  | Starcraft | −29.10% |
| **Average** | | **−27.04%** |  | **Average** | | **−25.80%** |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **LTM 5.1 vs ETM 6 rc1** | | **BD-rate** |  | **LTM 5.1 vs VTM 11** | | **BD-rate** |
| UHD | BarScene | −40.54% |  | UHD | Marathon | −30.44% |
| CatRobot | −39.64% |  | MountainBay2 | −22.52% |
| DrivingPOVLogo | −46.67% |  | DrivingPOVLogo | −43.42% |
| BoxeLogo | −24.58% |  | BoxeLogo | −37.07% |
| **Average** | | **−37.86%** |  | **Average** | | **−33.36%** |
|  |  |  |  |  |  |  |
| **LTM 5.1 vs ETM 6 rc1** | | **BD-rate** |  | **LTM 5.1 vs VTM 11** | | **BD-rate** |
| HD | TrafficLogo | −30.91% |  | HD | TrafficLogo | −21.11% |
| Starcraft | −11.28% |  | Starcraft | −19.96% |
| **Average** | | **−21.09%** |  | **Average** | | **−20.53%** |

# Conclusions

This document provides the report on LCEVC compression performance verification: Verification Testing of LCEVC for standard dynamic range (SDR) content.

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 AVC were determined to be approximately 46% for UHD and 28% for HD. The average bit rate savings for LCEVC when enhancing 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 with these two codecs.

The second set of tests aimed to confirm that LCEVC provided a more efficient means of resolution enhancement of half resolution anchors than unguided up-sampling. For these tests, the test sequences were coded using AVC, HEVC, EVC, or VVC at half resolution in both, horizontal and vertical direction. For anchor generation, the half resolution encoded sequences were upsampled with Lanczos filters to full resolution for visual assessment. The same half resolution encoded sequences were also used as base layers for LCEVC and hence not all curves may overlap as much as would be ideal when calculating a BD-rate. Comparing LCEVC full-resolution encoded sequences with the up-sampled half-resolution anchors, the average bit-rate savings when using LCEVC with AVC, HEVC, EVC, and VVC were calculated to be approximately 28%, 34%, 38%, and 33% respectively for UHD, and 27%, 26%, 21%, and 21% respectively for HD.

# References

1. International Telecommunication Union – Radio Communication Sector; Recommendation ITU‑R BT.500-14
2. “FFmpeg Filters Documentation”, The FFmpeg developers, Available: <https://ffmpeg.org/ffmpeg-filters.html#scale> (Retrieved: 2021/05/10)
3. “Report on Essential Video Coding Compression Performance Verification Testing for SDR Content”, ISO/IEC JTC 1/SC 29/WG 04 N0047 (w20000), January 2021
4. Gisle Bjøntegaard, “Calculation of Average PSNR Differences Between RD Curves”, ITU-T SG16/Q6, 13th VCEG Meeting, Austin, Texas, USA, April 2001, Doc. VCEG-M33
5. Gisle Bjøntegaard, “Improvements of the BD-PSNR model”, ITU-T SG16/Q6, 35th VCEG Meeting, Berlin, Germany, July, 2008, Doc.VCEG-AI11
6. “Requirements for Low-Complexity Video Coding Enhancements” ISO/IEC JTC 1/SC 29/WG11 N18098, Macau, October 2018

# Acknowledgements

The WG4 wishes to thank all the organizations and individuals who contributed to LCEVC verification tests:

* Dr. Vittorio Baroncini (MPEG VQA Advisory Group SDR Test Chair) for his guidance, design and coordination of the subjective tests;
* ISO/IEC JTC 1/SC 29/AG 5 for guidance on the verification test procedures;
* VabTech Lab and GBTech Lab for conducting the subjective assessments, despite the strong COVID restrictions imposed;
* Giacomo Baroncini for managing the subjective assessments conducted in GBTech Lab;
* V-Nova for generating the bitstreams, and Dolby, Tencent, Samsung, Università Politecnica delle Marche (IT) and Sun Yat-sen University (CN) for providing the cross-checks of the bitstreams;
* All experts who actively participated in discussions on the verification testing during meetings and provided valuable comments and suggestions.

1. – Short Description of the Verification Tests

## LCEVC Requirements

LCEVC performance requirements are detailed in N18098 [6]:

*“When enhancing an n-th generation MPEG codec (e.g., AVC), compression efficiency for the aggregate stream is appreciably higher than that of the n-th generation MPEG codec used at full resolution and as close as possible to that of the (n+1)-th generation MPEG codec (e.g., HEVC) used at full resolution, at bandwidths and operating conditions relevant to mass market distribution;*

*and*

*Encoding and decoding complexity for the aggregate full resolution video (i.e., base plus enhancement) shall be comparable with that of the base encoder or decoder, respectively, when used alone at full resolution.”*

## Tests Conducted

The verification tests have been conducted using AVC, HEVC, EVC, and VVC encoded sequences as anchors and base layers for LCEVC.

The verification tests included two types of tests:

1. **Requirements test**. These tests verify satisfaction of the LCEVC requirements, by comparing between full-resolution LCEVC-enhanced encoded sequences vs. full-resolution native anchors.
2. **Resolution Enhancement test**. These tests verify that LCEVC is a more efficient resolution enhancer than unguided upsampling, by comparing full-resolution LCEVC-enhanced encoded sequences vs. half-resolution native anchors upsampled with FFmpeg Lanczos*.*

*Rationale for this test: the simplest alternative to enhancing a codec with LCEVC – e.g., when there is a constraint of either bandwidth or processing (or both) – is to just use the native codec at a lower resolution and let the end user device upsample the lower-resolution encoded sequence to full resolution by means of unguided upsampling.*



The verification tests have been conducted by comparing the LCEVC-enhanced full resolution encoded sequences against the anchors encoded at full resolution and anchors encoded at half resolution (the latter upsampled to full resolution using the FFmpeg Lanczos upsampler).

The full-resolution anchors have been selected in order to enable assessment of the sequences over a relevant MOS range. For upsampled half-resolution anchors, the same encoded sequences produced to serve as base for LCEVC-enhanced encoding are used, in order to provide a meaningful comparison and avoid quality saturation issues with the half-resolution anchors.

## Video test sequences

Since the verification tests include comparisons between full resolution encoded sequences and upsampled half-resolution encoded sequences, the test set had to include sequences where the difference made by full resolution is appreciable in terms of formal ITU-R BT.500 DSIS MOS. In particular, it was decided to include both relatively soft sequences (where the expectation is that upsampled half-resolution encoded sequences may be judged similarly to full-resolution encoded sequences, for reasons independent of compression efficiency) and sequences with some visible sharp details (where the expectation is that the difference between full resolution and upsampled half-resolution encoded sequences can be perceived by a viewer, especially for higher rate points).

For sequences with some sharp details, the quality of half-resolution anchors tends to saturate, and full-resolution anchors quickly outperform even the maximum theoretical quality achievable by upsampled half-resolution (i.e., uncompressed half-resolution anchor). In practice, for sharper sequences resolution does matter more, and not even a perfect half-resolution anchor can match the quality of higher rate points of full-resolution encoding.

For the verification tests, the following mix of sequences was selected for each tested codec:

* two Class A (UHD) sequences without particularly sharp details
* two Class A (UHD) sequences with some sharp details
* two Class B (HD) sequences with many sharp details

For EVC and VVC, two of the tested sequences were taken from the test set of the respective verification test.

### Normal and/or relatively soft sequences (Class A, UHD)

|  |  |
| --- | --- |
| CatRobot  (JM, HM, ETM only) | CatRobot is a popular JVET sequence, used also for the VVC verification tests.  It is characterized by a good amount of detail (for both luma and chroma), although for the most part in the medium frequencies, rather than in the highest frequencies. Naïve viewers presented with a Lanczos-upsampled uncompressed half-resolution of the sequence may find it hard to distinguish it from the original full-resolution source.  The Lanczos-upsampled uncompressed half-resolution (i.e., the maximum theoretical quality achievable with half-resolution encoded sequences) scores 40.09 dB PSNR and 98.83 VMAF. |
| LupoPuppet (JM, HM only) | LupoPuppet is another popular JVET sequence. The sequence is very good to test the ability of codecs to properly deal with a mixture of gradients (prone to banding/blocking), odd movements, hair details, etc.  Also in this case, the energy in the highest frequencies, although present and noticeable to attentive expert viewers is low, so that many naïve viewers may find it hard to distinguish full resolution source from a Lanczos-upsampled uncompressed half-resolution of the sequence.  The Lanczos-upsampled uncompressed half-resolution (i.e., the maximum theoretical quality achievable with half-resolution encoded sequences) scores 39.61 dB PSNR and 95.34 VMAF. |
| BarScene (ETM only) | Popular sequence from the Netflix El Fuente test set, also used for the EVC verification tests. The scene includes talking heads and gradients with potential for banding.  Naïve viewers may find it hard at times to distinguish full resolution from half resolution. |
| Marathon (VTM only) | Often mentioned as a modern version of the all-time classic CrowdRun, Marathon is a sequence full of gradients, details and movement and, in general, complexity.  Despite the abundance of detail, and possibly due to the “busy nature” of the scene, naïve viewers may still find it hard at times to distinguish full resolution from half resolution.  In terms of objective metrics, the Lanczos-upsampled uncompressed half-resolution anchor scores 37.53 dB PSNR and 99.69 VMAF. |
| MountainBay2 (VTM only) | MountainBay2 is a popular scenery sequence, of the kind that is full of details, but in fact not particularly sharp (also due to the difficulty of distant panoramic elements to be all in perfect focus).  For this scene, the “asymptotic” Lanczos-upsampled uncompressed half-resolution anchor scores 43.19 dB PSNR and 99.62 VMAF. |

### Normal sequences with some sharp elements (Class A, UHD)

|  |  |
| --- | --- |
| DrivingPOVLogo | DrivingPOVLogo was obtained by combining the popular JVET sequence DrivingPOV with a graphic overlay with some sharp logos and text, generated by a third-party marketing agency to simulate the types of fonts and logos found in actual use cases such as sport overlays, sports stadium billboards, eSports, advertising, film credits, etc. The overlay thus includes sharp graphic elements and text of different sizes, from small to very large.  Aside from the graphics, DrivingPOV is a useful sequence to test the ability of codecs to properly encode road details and complex divergent movements.  The presence of sharp graphics makes it easier for viewers to appreciate true full-resolution UHD vs. upsampled half-resolution. From an objective metrics point of view, the uncompressed half-resolution anchor scores 34.13 dB PSNR and 94.97 VMAF, due to visible softening of the text and logos. |
| BoxeLogo | This Boxing sequence comes from the Netflix El Fuente test set, and sports a mix of face details and fast-motions at risk of banding, blocking and dragging impairments. Similarly to DrivingPOVLogo, some overlay graphics elements were added.  Also in this case, the presence of sharp details makes it easier for naïve viewers to appreciate true full-resolution UHD vs. upsampled half-resolution. From an objective metrics point of view, the uncompressed half-resolution anchor scores 35.42 dB PSNR and 93.33 VMAF. |

### Sharper sequences (Class B, HD)

|  |  |
| --- | --- |
| TrafficLogo | TrafficLogo is another sequence coming from the Netflix El Fuente test set. Originally shot in UHDp60, the sequence was downsampled to 1080p, which made it sharper than sequences natively shot in 1080p. Aside from the sharp road details, a graphic overlay was added.  TrafficLogo is a good sequence to stress test the behaviour of codecs in the context of dark low-contrast moving objects, sharp road details, and night light beams prone to banding.  The presence of sharp details sequence makes it easier for naïve viewers to discern full resolution from upsampled half-resolution, especially for high rate points. In terms of objective metrics, the uncompressed half-resolution anchor scores 32.30 dB PSNR and 82.51 VMAF. |
| Starcraft | Starcraft is a popular eSports sequence, and the sharpest sequence in the test set. Lots of details in both the background and the moving elements, with a combination of fast-motion panning and slower-paced action. Volumetric consistency of motion of background and moving objects – ill captured by frame-based metrics such as PSNR or VMAF, or even by split-screen expert viewing check – is relevant to MOS scores provided by naïve viewers to this sequence.  Naïve viewers are capable to appreciate true full-resolution 1080p vs. upsampled half-resolution, especially for high rate points. In terms of objective metrics, the uncompressed half-resolution anchor scores 29.69 dB PSNR and 81.21 VMAF. |

### Graphic overlays

In order to include in the tests also realistic sequences that would feature a mix of normal/softer areas and sharper elements, graphic overlays were added to two of the Class A sequences and one of the Class B sequences. The overlays were added by a marketing agency, briefed to include logos and text of different sizes, representative of full-resolution details found in real use cases such as news, sports, eSports, advertising, product placement, background details, film credits, etc.

Some examples of sizes of small text details found in real sequences (Tears of Steel, Starcraft) are reported below, compared to the smallest fonts used for the overlays added to some of the test sequences. It should be also noted that the large majority of text and logos used for the graphic overlays is much larger than what shown in these examples.

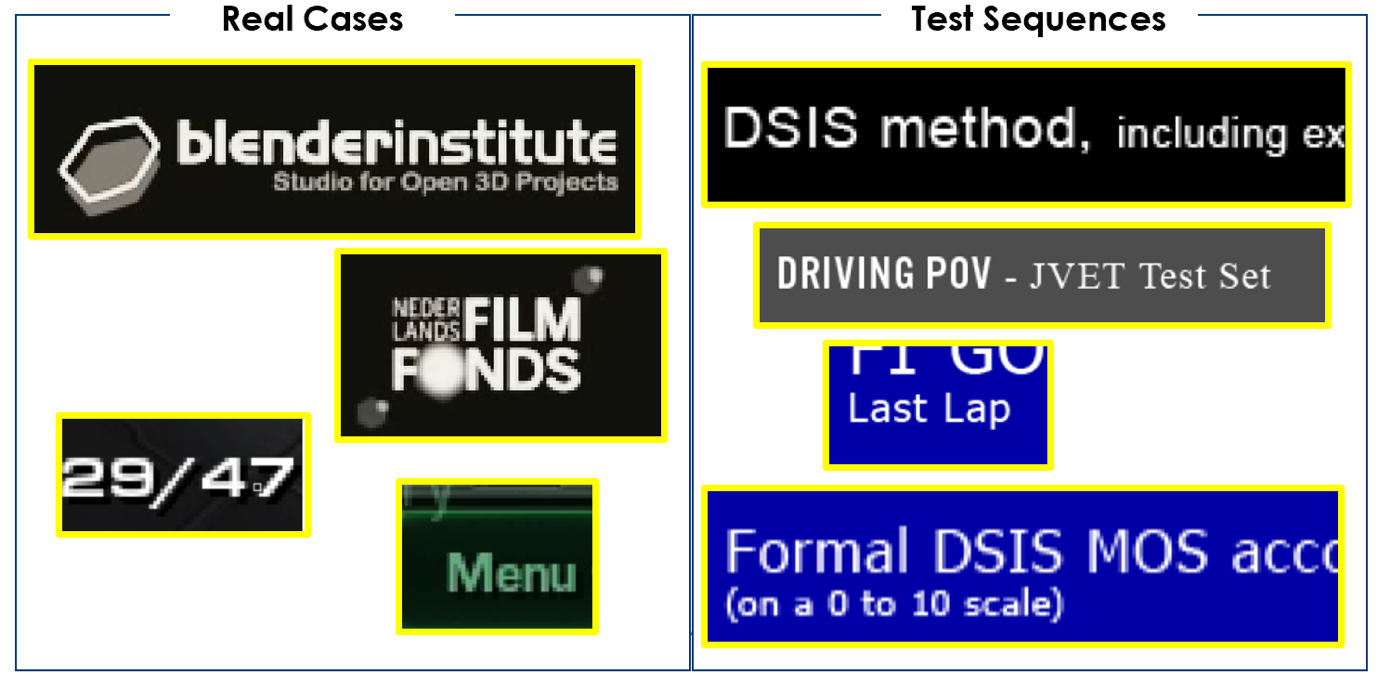


Fig. 25. Examples of small text details included in real sequences (left) vs. smallest text included in the graphic overlays.

The overlays successfully generated sequences with RD-curve profiles of full-resolution vs. upsampled half-resolution anchors “in between” those of relatively softer sequences and sharper sequences.

## Anchor encoding parameters

The full-resolution anchors have been selected in order to enable assessment of the sequences over a relevant MOS range. For upsampled half-resolution anchors, the same encoded sequences produced to serve as base layers for LCEVC-enhanced encoding were used. The same configuration files were used for both full-resolution anchors and LCEVC base layers / half-resolution anchors.

The initial selection of QPs for every anchor was based on targeting a suitable range of operating points spanning a range of expected MOS between 3 and 9. In response to feedback received during the collaborative phase (i.e., there was a doubt, then disproved by additional tests, that the efficiency benefits provided by LCEVC would be comparatively greater at medium qualities and would fade at higher visual qualities), the range of QPs focused on ensuring that also high qualities were included in verification tests, while still avoiding to reach MOS saturation.

For some of the sequences, it was noted that both PSNR and VMAF RD-curves showed extremely high values, but the actual corresponding qualities were quite low, and so the range of selected QPs was varied accordingly. An example of that is BoxeLogo, where the range of anchor QPs that was eventually chosen spans a VMAF range from 86 to 98.8 (with PSNR spanning from 34.1 dB to 41.8 dB). Although that may seem excessively high, in truth the sequence is extremely prone to visible banding, blocking and dragging artefacts, which are quite jarring even at VMAF scores higher than 95. It was also noticed, especially with VVC, that for sharp sequences objective metrics may tend to be very high even when actual subjective quality is lower than the score would suggest.

## LCEVC-enhanced encoding parameters

For LCEVC, LTM configurations specify the quantization stepwidth for sub-layer 2. All other LCEVC configuration parameters (e.g., QP for the base layer, stepwidth for sub-layer 1, upsampling method, transform type, quantization matrix, temporal stepwidth modifier, dequantization mode) are configured automatically by the LTM default configuration according to fixed formulas based on enhanced codec, resolution, and magnitude of stepwidth for sub-layer 2. These formulas are further specified in Section A.5.1 for resolutions above 1920×1080 and in Section A.5.2 for resolutions below or equal to 1920×1080.

The dithering tool was turned off for the verification tests.

### Relation between base layer QP and enhancement layer stepwidth for resolutions above 1920×1080

The following formulas are used in the LTM encoder depending on the codec used for the base layer:

* AVC base codec

If then = 3600 else = 32767

* HEVC base codec

If then = 3600 else = 32767

* EVC base codec

If then = 3600 else = 32767

* VVC base codec

If then = 3600 else = 32767

### Relation between base layer QP and enhancement layer stepwidth for resolutions below or equal to 1920×1080

The following formulas are used in the LTM encoder depending on the codec used for the base layer:

* AVC base codec

If then = 3600 else = 32767

* HEVC base codec

If then = 3600 else = 32767

* EVC base codec

If then = 3600 else = 32767

* VVC base codec

If then = 3600 else = 32767