 ISO/IEC JTC 1/SC 29/AG 3 N0046

**ISO/IEC JTC 1/SC 29/AG 3**

**MPEG Systems   
Convenorship: KATS (Korea, Republic of)**

**Document type:** Output Document

**Title:** White paper on Essential Video Coding (EVC)

**Status:** Approved

**Date of document:** 2021-11-12

**Source:** ISO/IEC JTC 1/SC 29/AG 3

# Expected action: None

# Action due date: None

**No. of pages:** 16 (without cover page)

**Email of Convenor:** kyuheonkim@khu.ac.kr

**Committee URL:** <https://isotc.iso.org/livelink/livelink/open/jtc1sc29ag3>

**INTERNATIONAL ORGANISATION FOR STANDARDISATION**

**ORGANISATION INTERNATIONALE DE NORMALISATION**

**ISO/IEC JTC 1/SC 29/AG 3**

**CODING OF MOVING PICTURES AND AUDIO**

**ISO/IEC JTC 1/SC 29/AG 3 N** **00046**

**Online – October 2021**

|  |  |
| --- | --- |
| **Source:** | **Convenor of ISO/IEC JTC 1/SC 29/AG 03** |
| **Status:** | **Approved by AG 03** |
| **Subject:** | **White paper on Essential video coding (EVC)** |
| **Date:** | **November 12, 2021** |
| **Serial Number:** | **21036** |

**Abstract**

This document provides a draft text for EVC white paper.

1. **Introduction**

The Essential Video Coding (EVC) standard has been developed as Part 1 of the MPEG-5 project to meet the requirements of significantly improving compression efficiency over existing video coding standards with a timely publication of licensing terms. The normative standard was finalized by the ISO/IEC Moving Picture Experts Group (MPEG) in April 2020 [1]. This white paper reviews the architecture and building blocks of EVC, which were carefully selected to consider compression capability, complexity and timely publication of licensing terms, which are the key goals of the EVC project. This paper also reports results of MPEG verification testing that have shown that the compression capability objective of the EVC standard development was achieved.

1. **Background**

There is a constant demand for more efficient video coding technologies, however coding efficiency is not the only factor which determines the industry choice of video coding technology for products and services. Video coding technologies should address the needs of existing and emerging real-world use cases and should also be easy to adopt from both technological and business perspectives.

As motivated by business perspectives but still existing demands on a high performance license-friendly video codec, MPEG issued a requirement to develop a new video coding standard that addresses both technical and business requirements that are not adequately satisfied by previous video coding standards [2].

To achieve these technical and business requirements, the EVC standard was developed with a royalty-friendly Baseline profile and a royalty bearing Main profile which includes a small number of improved coding tools on top of the Baseline profile.

The Baseline profile contains technologies assessed to be over 20 years to maximally approach to a royalty free codec. The Baseline profile builds a video codec using only conventional coding technologies which consist of traditional methods from the early 1980s to the end of the 1990s. On top of the Baseline profile, a small number of tools were added in order to improve compression performance for the Main profile. Each of these tools in the Main profile provide a significant improvement in terms of compression efficiency and are capable of being activated or deactivated on an individual basis. The Main profile follows the Type 2 declarations in ITU-R/ITU-T/ISO/IEC Common Patent Policy [3].

The purpose of such a licensing-friendly video codec is to reduce business uncertainty by minimizing any risk associated with the often lengthy delay in understanding the full licensing cost of launching products or services based on the standard. After finalization of the EVC standard, the major contributors to EVC have jointly reaffirmed their commitments to offer the EVC business requirement [4].

1. **Key technology and feature**

EVC employ a hybrid block-based video coding approach. Each colour component of the input picture is partitioning into a group of non-overlapped blocks. Largest allowed block in each colour component is called Coding Tree Block (CTB), combination of luma CTB and associated chroma component’s CTBs comprise Coding Tree Unit (CTU). A group of CTUs covering a rectangular region of the coded picture can be further grouped into a slice and/or a tile. CTUs are coded in a raster scan order utilizing a set of defined coding tools.

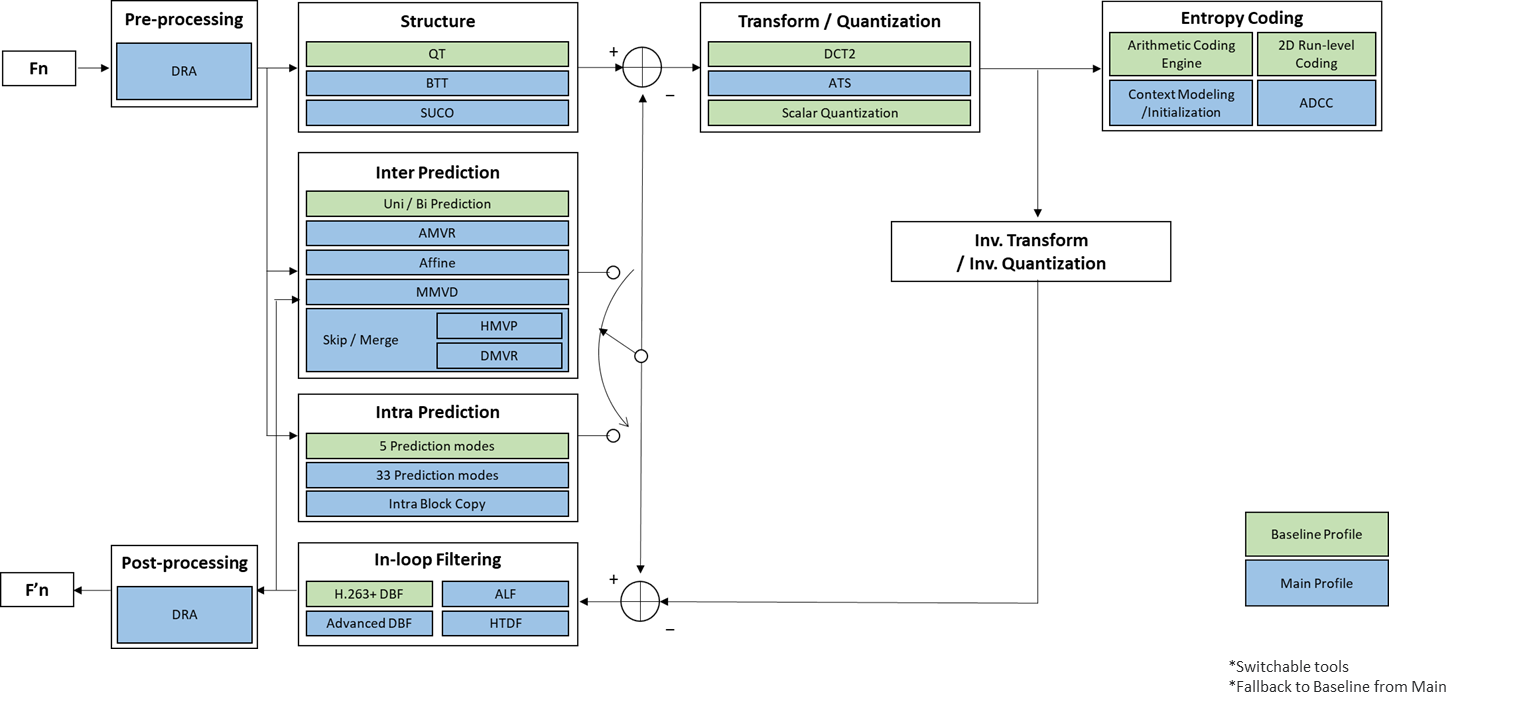


Figure 1. Block diagram of EVC coding tools

Figure 1 shows a tool-level summary of EVC, which is illustrated in a block diagram of tools associated with its profile. All of tools in the Baseline profile are based on technologies assessed as available for over 20 years. The tool set of the Main profile, which is added on top of the basic set of the Baseline profile, were carefully chosen by considering an interaction between the tools to enhance a performance with a limited number of coding tools.

* 1. **Tools of Baseline profile**
     1. **Coding structure**

Baseline profile of EVC coding input picture with a partitioning to CTUs with a maximal size of 64x64 luma samples. Group of CTU can form a slice each of which can be coded independently.

Coding structure would be the most efficient tool to dramatically improve compression performance. Historically, new video codecs have always been proposed with new coding structures. A 16x16 samples macroblock structure was employed until the AVC standard [8]. Since a coding structure based on 16x16 sample blocks gave poor coding performance in the case of large resolutions such as 1080p or 4K, the HEVC standard and subsequent video coding standards employed larger block structures such as 64x64 with a flexible splitting mechanism to divide a large coding block into small blocks. The coding structure in the Baseline profile employed a quad-tree based approach which can use blocks up to 64x64 samples. The quad-tree based coding algorithm was introduced in the early 1990s and an efficient splitting method was disclosed in 1994 [5].

* + 1. **Intra prediction**

Intra prediction is a technology to exploit spatial correlation between neighbouring pixels. While prediction is performed in the transform domain in H.263 and MPEG-4 visual, AVC introduced prediction in the pixel domain by referencing the neighbouring pixels of the previously coded blocks [6][7]. Higher compression performance was achieved by increasing a number of prediction directions in the HEVC standard. In EVC Baseline profile, 5 prediction modes are employed: a so-called DC mode, where predicted samples are calculated as the arithmetic mean of the reference samples, and 4 directional prediction modes ­­– horizontal (H), vertical (V), diagonal left (DL), diagonal right (DR). A codeword for prediction mode of the current block is adaptively assigned by using a mapping table between symbol and prediction mode, which is selected based on the prediction modes of the neighbouring upper and left blocks.

* + 1. **Inter prediction**

For temporal prediction, the Baseline profile exploits three neighbouring motion vectors and a motion vector of temporally co-located blocks. After choosing one of the candidate motion vectors as a predictor, the index of the predictor is coded to the bitstream. The difference between the motion vector for the current block and the predictor can be coded or not. If encoder decides that the predictor is good, the motion vector difference and a block residue are not coded, this mode is called the skip mode. Otherwise, the motion vector difference and the block residue are coded and signalled in the bitstream. The EVC Baseline profile also allows a bi-directional prediction, which is a linear combination of two motion compensated blocks that involve two motion vectors.

* + 1. **Transform and Quantization**

As has been used in numerous codecs for video compression, the EVC Baseline profile also adopts a discrete cosine transform (DCT), which is well-known as having good energy compaction efficiency. The transforms are applied to a residual block between an original block and the corresponding prediction block, as a conventional hybrid video codec does. Since transforms are applied to coding blocks, the transform size is equal to the coding block size, i.e. anywhere from 2x2 to 64x64 samples. After the transform is conducted, scalar quantization is applied to the transformed coefficients. The quantization parameter (QP) is taken from a range of 0..51 and a scaling factor (SF) corresponding to each QP is defined by a look-up table.

* + 1. **Loop filter**

Most block-based video coding schemes introduce blocking artifacts due to the per‑block analysis process and these are very noticeable especially at low bitrates. To remove such blocking artifacts, post filtering and loop filtering have been used. The loop filters, unlike the post filters, operate inside of coding loop and improve the quality of reference frames. Since refining reference frames results in improvement in visual quality and compression efficiency, loop filtering plays an important role in video coding. In the EVC Baseline profile, a loop filter based on H.263 Annex J [6] was employed to increase objective and subjective image quality. The algorithm is relatively simple enough to allow its operating in the coding loop, at the same time it is able to successfully remove blocking artifacts.

* + 1. **Entropy coding**

For entropy coding in video and image compressions, Huffman and arithmetic coding have been employed. The state-of-the-art entropy coding scheme is context-based adaptive binary arithmetic coding (CABAC), introduced in the AVC standard. Instead of CABAC, the binary arithmetic coding scheme from JPEG Annex D is applied as the entropy coding engine of the EVC Baseline profile [9]. After a binarization process of the given symbol, an arithmetic coding engine encodes each binary value with the corresponding context that stores the occurrence probability of a given value. After each binary value is encoded, the probability is updated by using a look-up table and the binary value of symbol is stored in the corresponding context. To code the transformed and quantized coefficient values, run-level symbols are generated after scanning with a zig-zag pattern. Each run or level symbol is binarized by unary coding and the binary value is coded with the corresponding run or level contexts. The sign value and the last symbol indicate whether the level is the last one in the block should be followed to each level.

* + 1. **High level syntax**

An EVC bitstream consists of network abstract layer (NAL) units with a 2-byte NAL unit header which contains properties of a NAL unit such as the type of data and temporal ID. A sequence parameter set (SPS) contains parameters that apply to the whole coded video sequence (CVS), a picture parameter set (PPS) contains parameters that apply to one or more pictures of a CVS, and an adaptation parameter set (APS) contains parameters of a coding tool which apply to one or more pictures of a CVS.

Reference picture management in the EVC Baseline profile is implemented by well-known implicit sliding window concept that is used in AVC standard [8]. This concept marks a reference picture as "unused for reference" when the number of reference frames is equal to a given maximum number. The reference pictures are stored in a first-in, first-out manner so that the most recently decoded pictures are kept in the Decoded Picture Buffer (DPB). It should be noted that in contrast to the AVC standard, the EVC Baseline profile does not support the explicit memory management control operation (MMCO) process.

* 1. **Tools of Main profile**
     1. **Coding structure**

Main profile of EVC coding input picture with a partitioning into CTUs with a maximal size of 128x128 luma samples. Group of CTU can form a tile each of which can be coded independently. A group of tiles can be further grouped into a slice.

The EVC Main profile supports a flexible block partitioning structure, which is based on the binary and ternary trees mixture scheme (BTT) with the Split Unit Coding Order (SUCO) method, for efficient and flexible representation of video content with various resolutions. BTT has coding unit (CU) shapes described by the ratio between the width and height of a block as shown in Figure 2. For instance, if the width and height of a block are the same it can be represented as a 1:1 ratio CU or a square CU, and if the width is equal to 64 samples and the height is equal to 16 samples it can be represented as a 1:4 ratio CU. CU partitioning is conducted based on the allowed CU shapes and their allowed maximum and minimum sizes.

**텍스트, 낱말맞추기게임, 클립아트, 실루엣이(가) 표시된 사진

자동 생성된 설명**

Figure 2. Binary/Ternary split modes

As shown in Figure 3, the SUCO method enables a more flexible coding order, such as left to right (L2R) and right to left (R2L) orders, to allow intra prediction from right reference pixels and inter prediction with right motion vector predictors. If a split unit (SU) is partitioned vertically (vertical splitting), a flag is signalled to indicate L2R or R2L coding order of partitioned units. If an SU is partitioned by a quad tree structure, a flag is shared for the two above units and two bottom units. If no flag is signalled for coding order of an SU, the coding order follows its parent’s SU coding order.



Figure 3. Allowed coding order in Quad-tree/Binary/Ternary split modes

* + 1. **Intra prediction**

To exploit spatial correlation efficiently based on flexible coding structure, a total of 33 intra prediction modes for luma component and 5 modes for chroma component are applied. DC, Bi-linear, Plane and Direct Mode (DM) modes are similar to that of the AVS2 standard [10], with a straightforward extension for a flexible block size while Plane and angular prediction mode are different from that in AVS2.

* + 1. **Merge with Motion Vector Difference**

Merge with motion vector difference (MMVD) approach provides a new motion vector (MV) expression method with simplified signalling. Similar to the skip and merge modes in HEVC, MMVD makes a candidate list from neighbouring motion information but MMVD can cover more extended motions that are not limited to the neighbouring motions. In order to construct more precise candidates, a starting point, a motion magnitude and a motion direction are utilized.

* + 1. **Adaptive motion vector resolution**

Adaptive motion vector resolution (AMVR) supports multiple motion vector resolutions. Supported motion vector resolutions in the EVC Main profile range from 1/4–pel to 4–pel (1/4–pel, 1/2–pel, 1–pel, 2–pel and 4–pel). Information about the motion vector resolution is signalled at the CU level. Depending on the resolution of CU, both motion vector (MV) and motion vector predictor (MVP) of the CU are adjusted accordingly.

* + 1. **Affine Mode**

Affine prediction of the EVC Main profile allows the use three different affine motion modes, namely four and six parameters model modes and an affine merge mode. In EVC, the affine motion field for a CU is described by MVs of two control points located at the top-left and top-right corners (four parameter model) or MVs of three control points located at the top-left, top-right and bottom-left corners (six parameter model). In four and six parameters model modes, the control point MVs for the current CU are signalled in the bitstream. For affine merge mode, the control point MVs of the current CU are derived based on the motion information of the neighbouring CUs. When merge or skip mode is applied, and both width and height for the CU are larger than or equal to 8 samples, a CU level affine flag is signalled in the bitstream in order to indicate whether affine merge mode is used. In this mode, the merge candidate index with a maximum value of 5 is signalled for specifying which motion information candidate in the affine merge candidate list is used for the CU.

* + 1. **Decoder-side motion vector refinement**

The Decoder-side motion vector refinement (DMVR) method operates with the two motion vectors of the bi-prediction which are further refined by a bilateral matching process. In bi-prediction mode, refined MVs are searched around the initial MVs in the reference picture list 0 and the reference picture list 1. The DMVR searching process consists of an integer sample MV offset search and fractional sample MV refinement. The integer sample MV searching calculates the distortion between the two candidate reference blocks in the reference picture list L0 and list L1. The Sum of Absolute Differences (SAD) metric between the reference blocks based on each MV candidate around the initial MV is calculated. The MV candidate with the lowest SAD becomes the refined MV and used to generate the bi-predicted signal. The search range is 2 integer luma samples from the initial MVs. The fractional sample refinement is conditionally invoked based on the output of the integer sample search stage. To reduce the computational complexity, the fractional sample refinement is derived by using a parametric error surface equation, instead of an additional search with SAD comparison. In parametric error surface based sub-pixel offsets estimation, the center position cost and the costs at four neighbouring positions from the center are used to fit a 2-D parabolic error surface equation.

* + 1. **History-based Motion Vector Prediction**

The History-based motion vector prediction (HMVP) method is an inter coding tool which can be applied to both merge candidate list and motion vector prediction process. In HMVP, a table of HMVP candidates is maintained and updated on-the-fly. Whenever a non-affine inter coded block is decoded, the decoded motion information is used to update the HMVP table in the last position following a first-in-first-out (FIFO) rule to remove and add entry. The HMVP table size is set to be 23 and subsampling fetching process is employed to avoid pruning over the HMVP table and thus reduce complexity. For each of the HMVP entry, a single MV and a reference index for uni-prediction or two MVs and two reference indices for bi-prediction are stored.

* + 1. **Advanced Deblocking Filter**

A deblocking filter based on the design utilized in ISO/IEC 14496-10 [8] is employed to increase objective and subjective image quality for the EVC Main profile. Utilized filter is applied at the coded blocks boundaries and adapts its strength depending on several factors, including utilized QP, coding modes for current and neighbouring blocks, motion information, as well as signal gradient at the block boundaries. Strength of the filter is also can be adjusted by user-defined control parameters.

* + 1. **Adaptive Loop Filter**

To suppress compression artifacts and improve the visual and objective quality of the decoded and reference pictures, decoded samples are filtered with an adaptive loop filter (ALF). Parameters of the ALF are signalled in independent NAL units, called Adaptation Parameter Set (APS), which once made available at decoder side can be referred from different coded pictures. . For luma filtering, two types of diamond filter patterns (5x5 and 7x7) are defined while for chroma samples only the 5x5 pattern is used. Each ALF structure may consist of up to 25 different luma filters and the utilized luma filter is selected through a classification process for each 4x4 block. For chroma filtering, a single filter per APS can be signalled and used in a slice or tiles group.

To reduce bit overhead, several filter coefficient prediction techniques are employed. To benefit from symmetrical properties of filters, the utilized ALF employs a filter coefficient transformation process, thus reducing the number of independent filter coefficients to 13. Following this, a set of static filter coefficients is provided to the decoder as side information and used as predictors for signalled ALF coefficients.

And finally, to benefit from temporal correlation, the employed ALF design utilizes re-usage of the ALF coefficients signalled earlier. Each signalled APS consisting of an ALF is identified by the unique adaptation parameter set ID which is used for referencing the relevant ALF information from other syntax elements, e.g. from slice or tiles group. All signalled APS with unique set id value are stored in APS buffer size up to 32 entries. To enable the random access (RA) coding configuration, the encoder’s choice of the APS ID usage is constrained. For example, to keep the temporal scalability, only APS from the same or lower temporal layers can be referenced.

* + 1. **Hadamard transform domain filter**

In addition to deblocking filter and ALF, Hadamard transform domain filter aims to reduce ringing artifacts caused by quantization of residual coefficients is introduced in EVC. Hadamard transform domain filter (HTDF) is applied to luma reconstructed blocks when the quantization parameter (QP) is larger than 17. The transform core is 2x2 samples Hadamard transform, which results in that HTDF is a 3x3 low pass smoothing filter, taking into account 2x2 blocks overlapping. The filter parameters are derived from the coded information, including transform block size and QP. The HTDF filter is applied within Rate-Distortion Optimization (RDO) loop prior to the deblocking filter and ALF.

* + 1. **Adaptive transform selection**

Adaptive transform selection (ATS) is exploited in the EVC Main profile. In addition to traditional for video compression DCT-II transform core, EVC additionally introduces DST-VII and DCT-VIII transform cores that can be applied for intra and inter predicted residual. For intra coded block, a flag is used to signal to the decoder whether ATS is applied or not. Basically, the encoder determines the use of ATS based on the RDO process at CU level. When the encoder selects usage of ATS in a CU as a core transform, two more flags are signalled to the decoder in order to indicate which type is used, for the horizontal and vertical directions respectively. For inter coded block that contains residuals, a flag is signalled to indicate whether the whole residual block or a sub-part of the residual block should be decoded. When only a sub-part of the residual block is coded, the part of the residual block is coded with inferred transform type (DST-VII or DCT-VIII) and the other part of the residual block is zeroed out.

* + 1. **Advanced coefficient coding**

Transform coefficients of the coded block (residual data) after quantization are scanned in a predefined scan pattern and entropy coded. To employ statistical properties of transform coefficients, advanced coefficient coding (ADCC) utilizes the bit-plane like coding approach. In particular, transformed coefficients are scanned and coordinates (X and Y) of the last non-zero transform coefficient in the scan order are signalled. Following this, coefficients are parsed in inverse zig-zag scan order and processed in chunks of the maximal size of 16. Coefficients within each processing chunk are signalled as a sequence of significance (i.e., sigMapFlag) and levels flags (i.e., flagLevelA and flagLevelB), sign flag and remaining levels (i.e., levelRem). Among these symbols, the bins of sigMapFlag, flagLevelA and flagLevelB are encoded with adaptive context models; signFlag and binarized bins of levelRem are encoded through by-pass mode. To decrease the number of context-coded bins, the explicit flagLevelA and flagLevelB are adaptively switched into levelRem coding, which is binarized with Exponential-Golomb code and encoded with bypass mode with equal probability.

* + 1. **Dynamic Range Adjustment (DRA)**

To allow efficient coding of various types of video content, e.g. video in HDR/WCG formats, EVC utilizes Dynamic Range Adjustment (DRA) at the normative post-processing stage. The forward DRA is performed on the input video samples prior to coding and the inverse process is performed on the samples after decoding. The DRA modifies luma and chroma samples in the picture based on the value of the luma or chroma samples and on the value of other component samples located at the same or closest relative location. Parameters of the inverse DRA signalled in the bitstream as Adaptation Parameter Set (APS) and used at the decoder side to configure the inverse DRA which is applied to the decoded picture at the picture output process.

* + 1. **High level syntax**

Reference picture management in the EVC Main profile is implemented by a modern RPL (reference picture lists) concept. It assumes explicit signalling of two reference picture lists, list 0 and list 1 that determine the reference picture management process for a specific picture in the Group of pictures (GOP) structure. Reference picture marking is directly based on reference picture lists 0 and 1, utilizing both active and inactive entries in the reference picture lists, while only active entries may be used as reference indices in inter prediction of CTUs. Predefined reference picture lists can be signalled in the SPS, for use by referencing in the slice header and additional reference picture lists can be signalled for a picture in the Slice Header.

1. **Profiles and levels**

EVC specifies four profiles including “Baseline profile”, “Main profile”, “Baseline Still Picture profile”, and “Main Still Picture profile”.

Similar to previous video compression standards, EVC defines conformance points in terms of profiles (combinations of decoding tools and associated bitstream syntax that is expected to be interpreted by a decoder), and levels (typically maximum sizes of pictures and frame rates, maximum bit rate, buffer capacity etc.). Currently, a total of 13 levels, ranging in support for typical picture sizes as small as VGA at the low end, up to as large as 8K, 4K ultra HD at the high end, are specified.

1. **Performance**

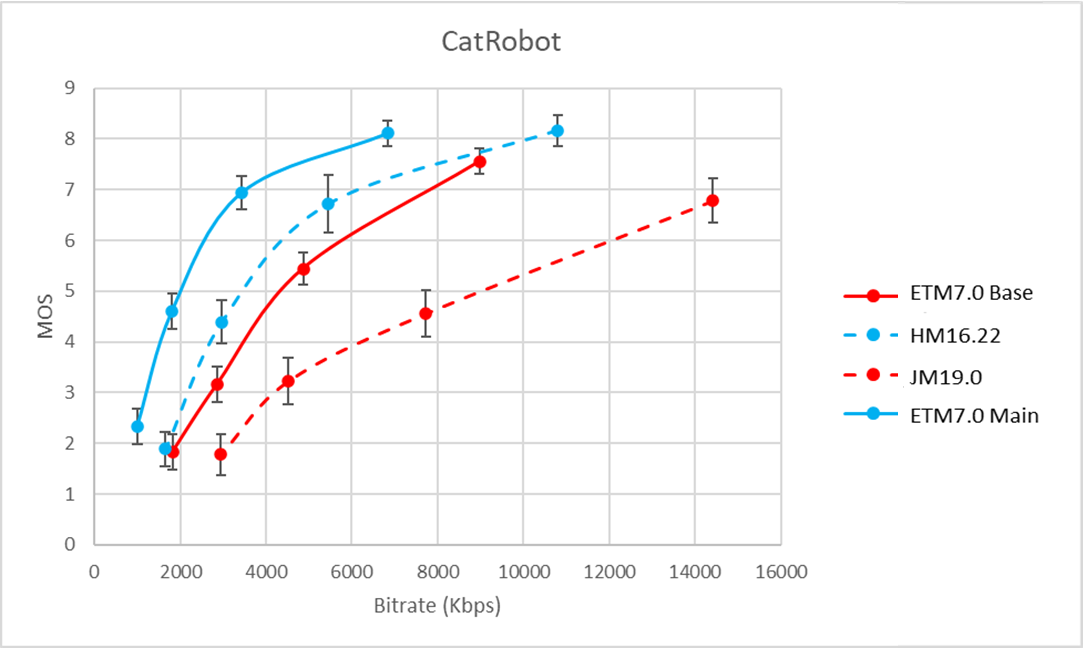
MPEG carried out a formal subjective verification test of EVC Main and Baseline profiles for standard dynamic range (SDR) content and high dynamic range (HDR) and wide color gamut (WCG) content, respectively [11][12]. The purpose of the verification test was to confirm that the compression efficiency objective for the EVC standard has been met: achieving a substantial bit-rate reduction at the same level of subjective visual quality relative to the HEVC Main10 profile for the EVC Main profile and the AVC High10 profile for the EVC Baseline profile, respectively.

* 1. **Results on SDR sequence**

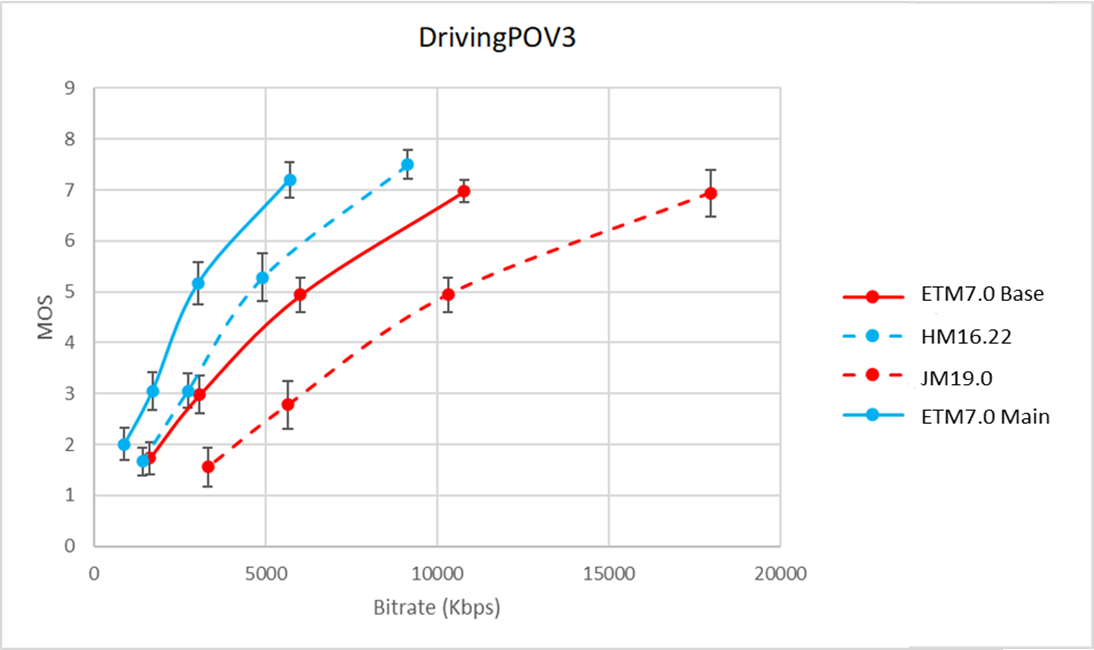
The assessment included SDR ultra high definition (UHD, a.k.a. 4K, 3840×2160) test sequences encoded in random access (RA) configuration and SDR full high definition (HD, a.k.a. 2K, 1920×1080) test sequences encoded in low delay (LD) configuration.

The average bit rate saving of the EVC profiles relative to the anchors were computed from the MOS vs. bit rate data for each sequence, in the same manner that was done in [13], to further quantify the bit rate savings achieved.

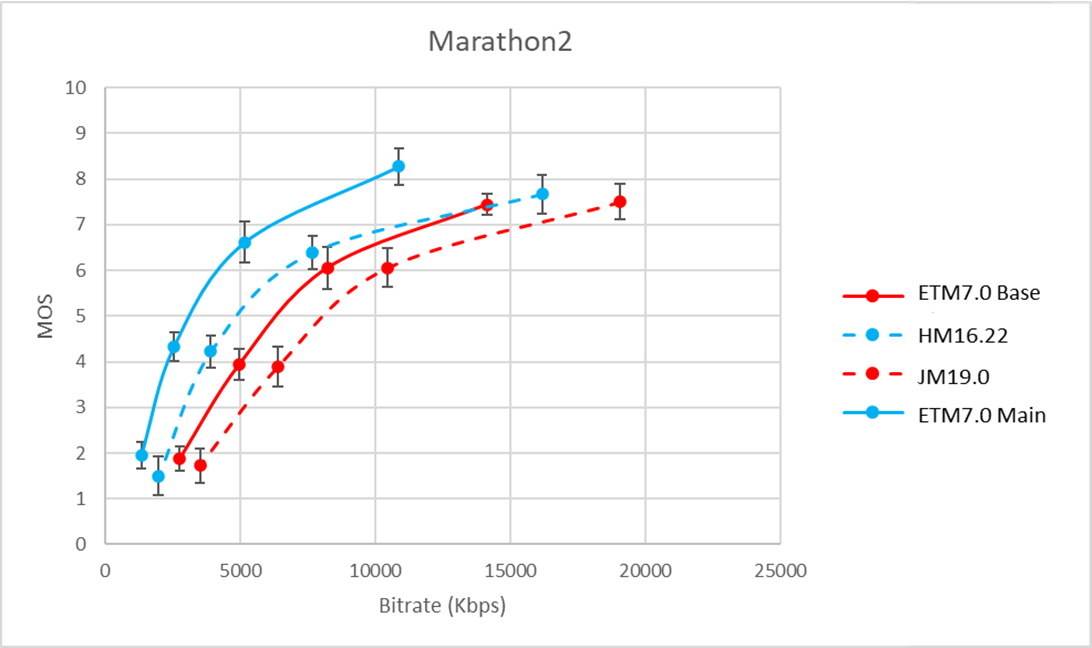
The bit rate savings were averaged over the whole range where the same MOS scores for the EVC Main profile and the EVC Baseline profile, HEVC and AVC could be interpolated from subjective test results shown in Figure 4 and 5.



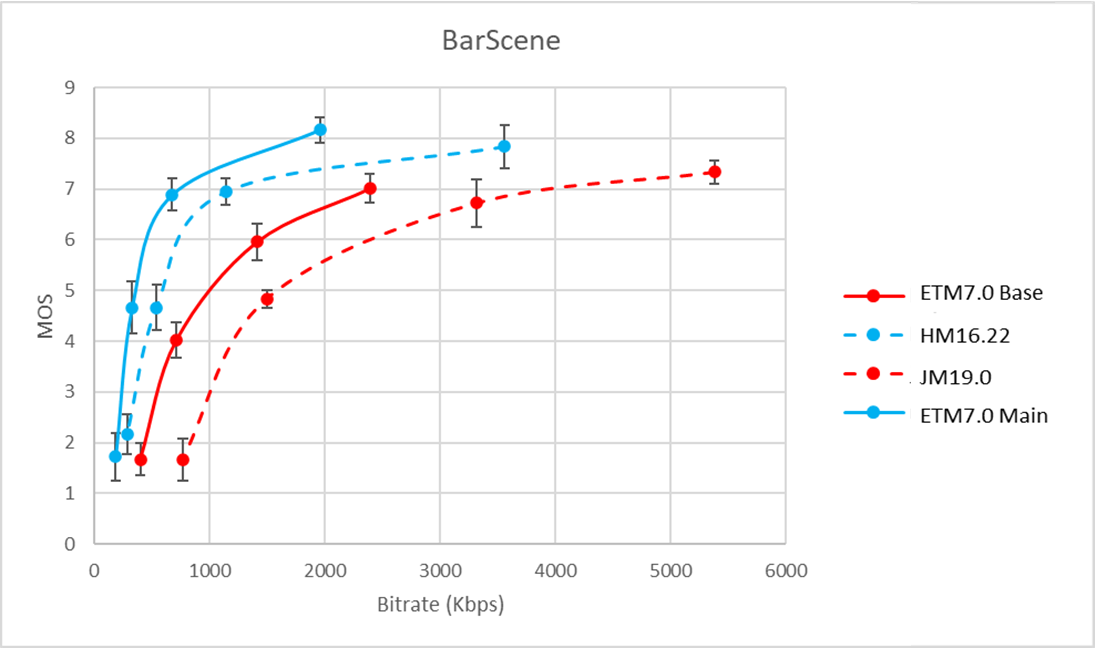
1. CatRobot test sequence



1. DrivingPOV3 test sequence

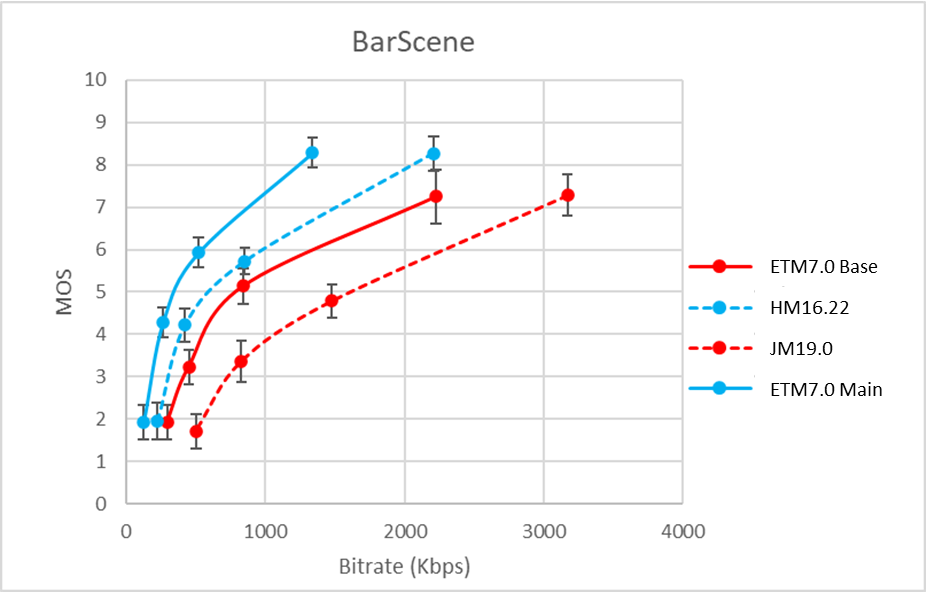


1. Marathon2 test sequence

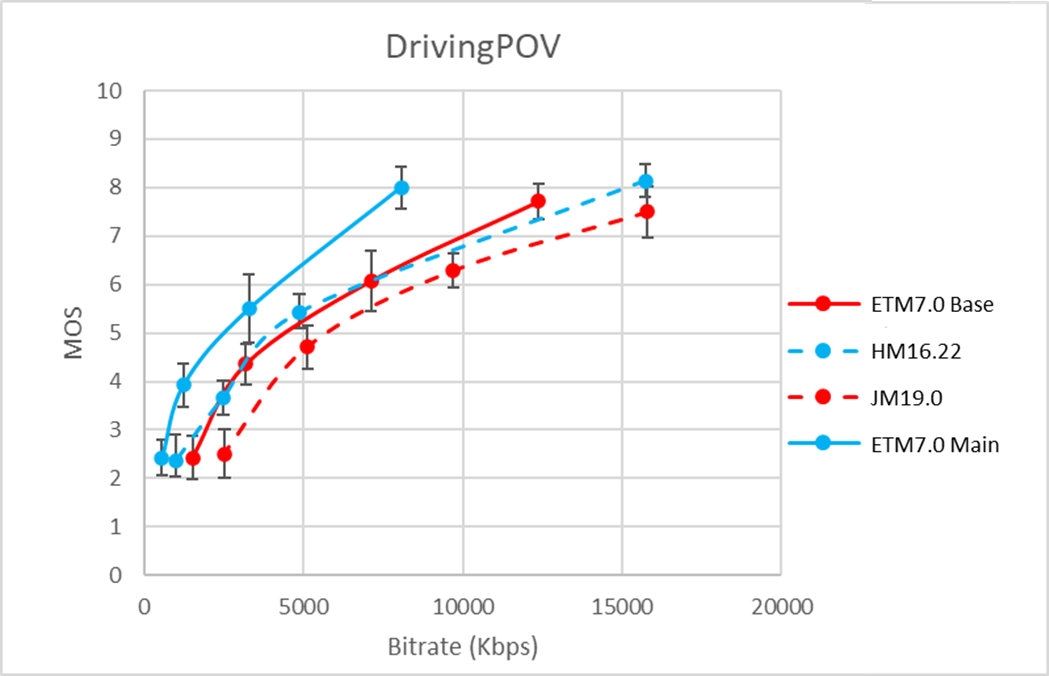


1. BarScene test sequence

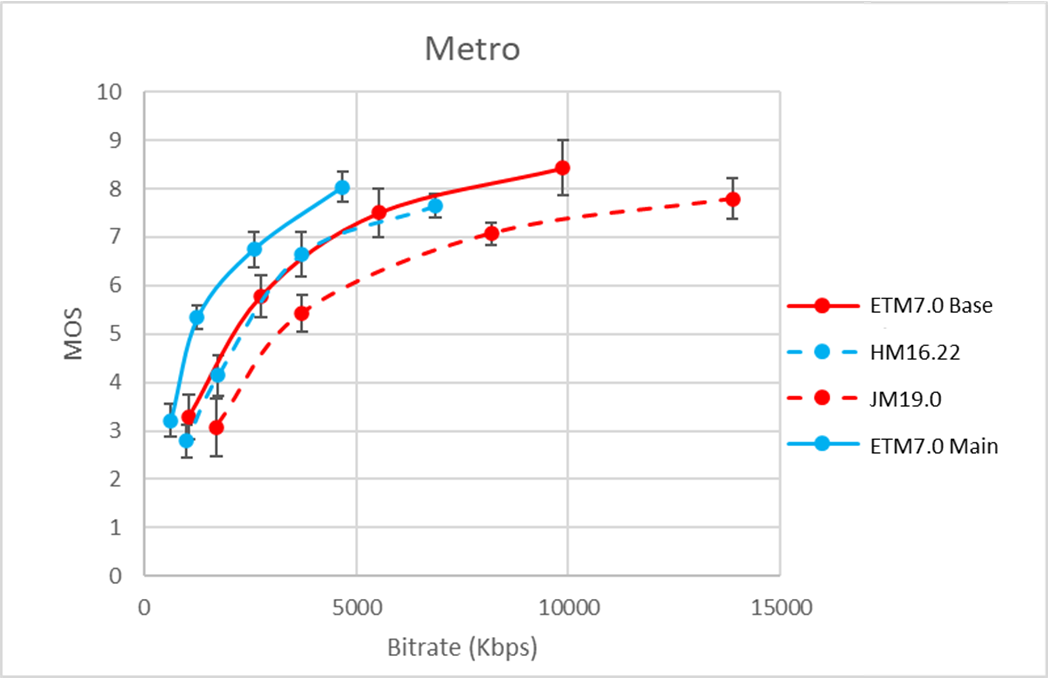
Figure 4. Results on SDR UHD resolution with RA configuration



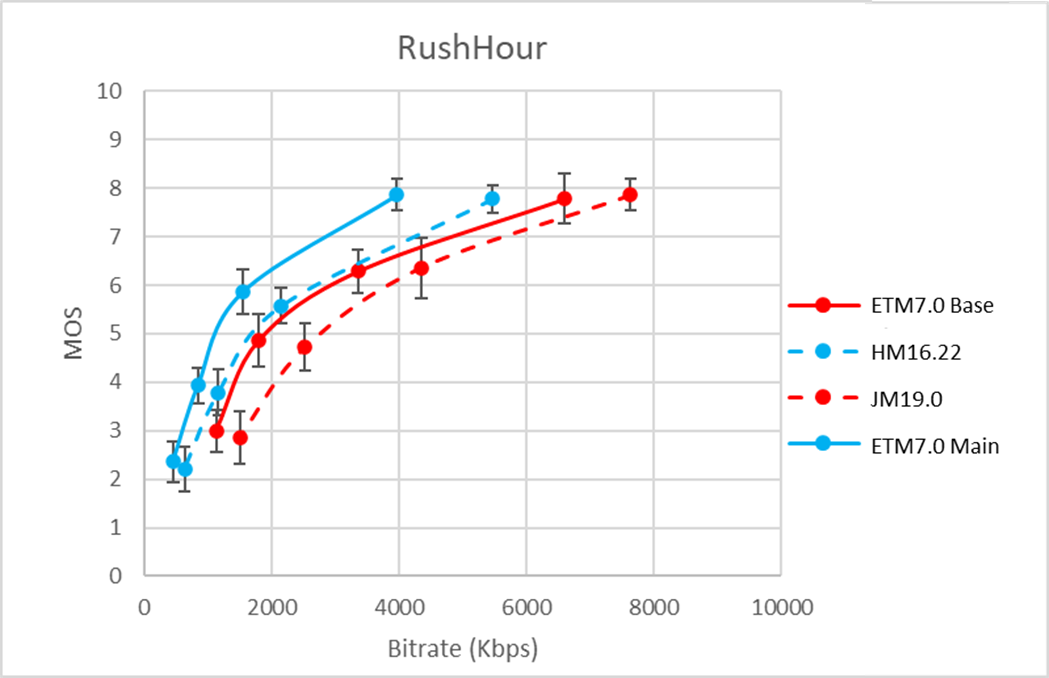
1. BarScene test sequence



1. DrivingPOV test sequence



1. Metro test sequence



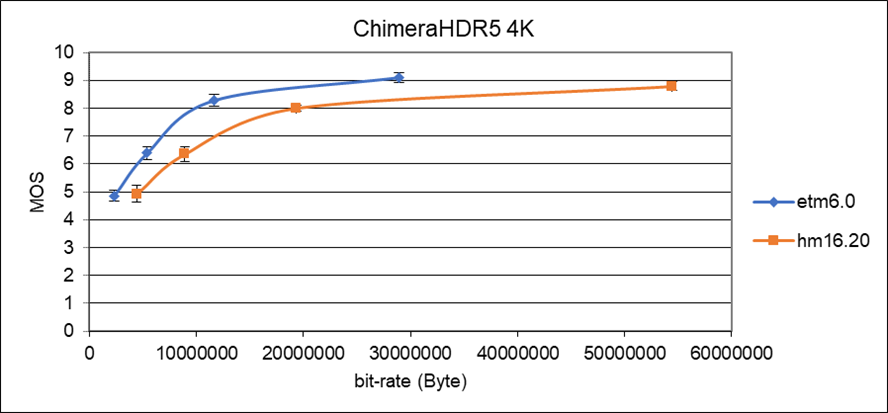
1. RushHour test sequence

Figure 5. Results on HD resolution with LD configuration

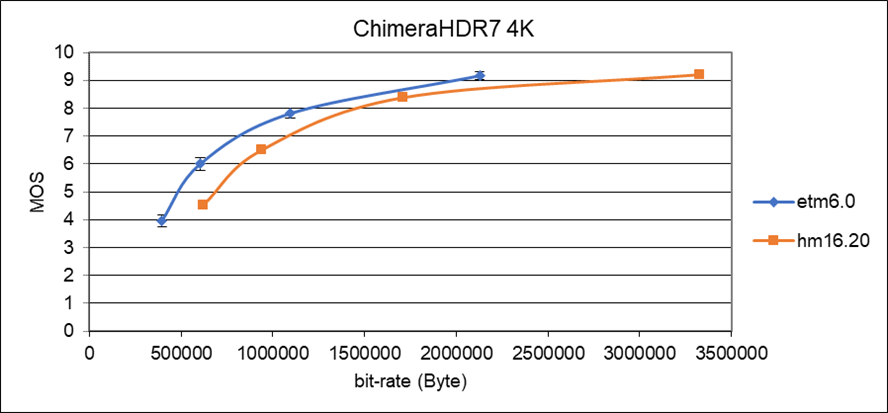
The MOS BD-rate calculation [14] shows that the average bit rate saving for the EVC Main profile compared to the HEVC Main10 profile was approximately 39% using UHD SDR content encoded in random access configuration, and approximately 41% using HD SDR content encoded in low delay configuration. The average bit rate saving for the EVC Baseline profile compared to the AVC High10 profile was approximately 39% using UHD SDR content encoded in random access configuration, and approximately 34% using HD SDR content encoded in low delay configuration.

* 1. **Results on HDR/WCG sequence**

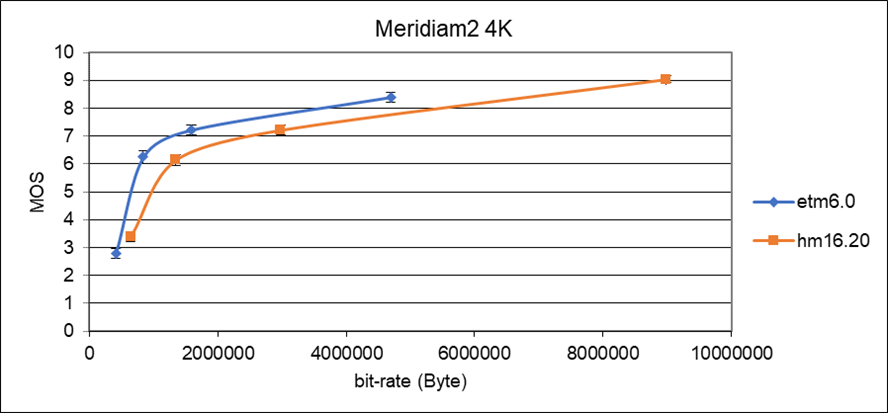
Subjective testing in HDR category was conducted through a subjective evaluation, comparing coding performance of the EVC Main profile to the HEVC Main 10 profile for video sequences in BT.2100/PQ representation [15]. Figure 7 shows the visual testing results in HDR/WCG category. The average bit rate savings for the EVC Main profile (etm6.0) with HDR-UHD and HDR-HD are estimated at approximately 36% and 35% respectively compared to the HEVC Main10 profile (hm16.20).



1. Chimera5 test sequence

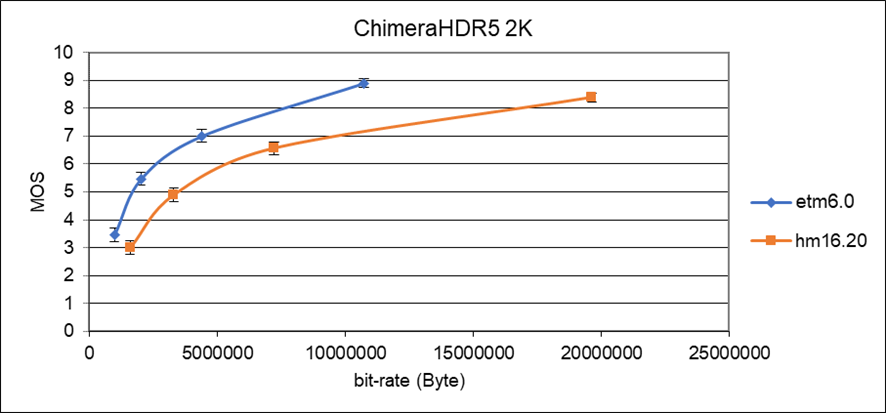


1. Chimera7 test sequence

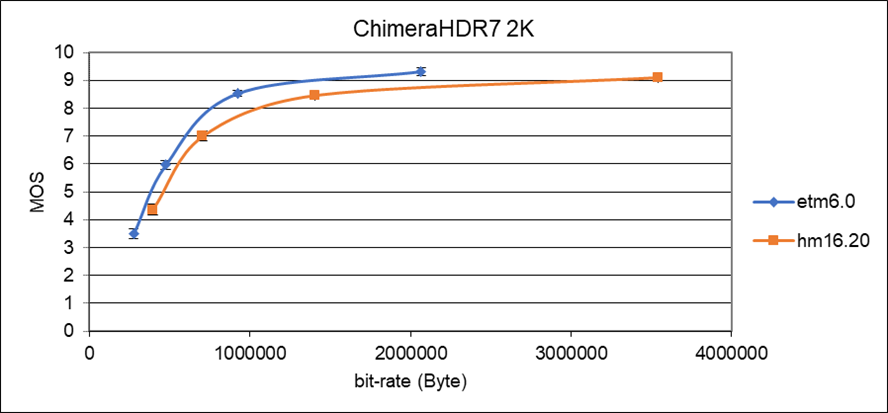


1. Meridian2 test sequence

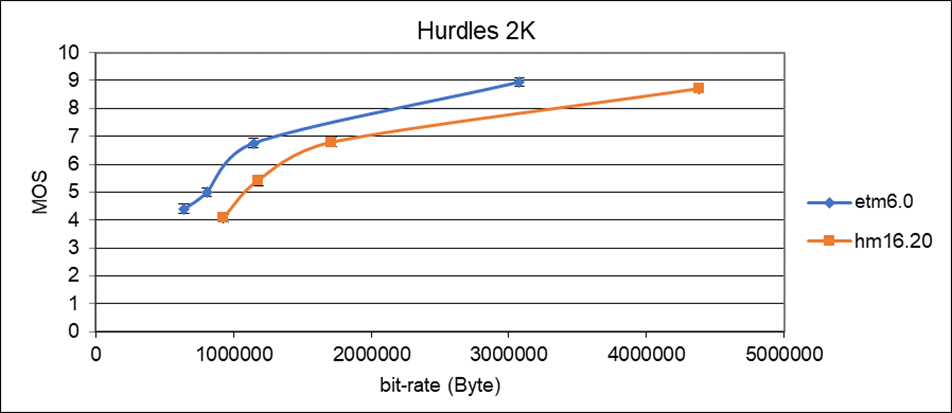
Figure 6. Results on UHD resolution with RA configuration



1. Chimera5 test sequence



1. Chimera7 test sequence



1. Hurdles test sequence

Figure 7. Results on HD resolution with RA configuration

1. **Resources**

EVC test model (ETM) description and Common test conditions (CTC) document are available at [16][17], and the ETM reference software and the EVC conformance bitstreams are available at [18][19] respectively. The open-source encoder and decoder of EVC, which are XEVE [20] and XEVD [21], are available in a public code repository.

1. **Reference**

[1] ISO/IEC 23094-1:2020 Information technology — General video coding — Part 1: Essential video coding.

[2] “Requirements for a New Video Coding Standard”, ISO/IEC JTC1/SC29/WG11 N17928, Oct. 2018.

[3] “Common Patent Policy for ITU-T/ITU-R/ISO/IEC”, [Online]. Available: https://www.itu.int/en/ITU-T/ipr/Pages/policy.aspx

[4] “Samsung Electronics, Huawei and Qualcomm Welcome the Release of a New Video Coding Standard”, [Online]. Available: https://research.samsung.com/news/Samsung-Electronics-Huawei-and-Qualcomm-Welcome-the-Release-of-a-New-Video-Coding-Standard

[5] G.J. Sullivan and R.L. Baker, “Efficient quadtree coding of images and video”, IEEE Trans-actions on Image Processing, 327-331, May 1994.

[6] Video Coding for Low bit rate Communication, ITU-T Recommendation H.263, version 1, Nov. 1995; version 2, Jan. 1998; version 3, Nov. 2000.

[7] Coding of audio-visual objects—Part 2: Visual”, ISO/IEC 14496-2 MPEG-4 visual, Apr. 1999.

[8] “Advanced Video Coding (AVC)”, ITU-T Recommendation H.264 and ISO/IEC 14496—10, May 2003.

[9] Information technology-digital compression and coding of continuous-tone still images-requirements and guidelines. International Telecommunication Union. CCITT recommendation, 81, p.09.

[10] He, Zhichu, Lu Yu, Xiaozhen Zheng, Siwei Ma, and Yun He. “Framework of AVS2-video coding”, IEEE International Conference on Image Processing, 1515-1519, 2013.

[11] “Report on Essential Video Coding Compression Performance Verification Testing for HDR/WCG Content”, ISO/IEC JTC 1/SC 29/WG 04 N0030, Oct. 2020.

[12] “Report on Essential Video Coding compression performance verification testing for SDR Content”, ISO/IEC JTC 1/SC 29/WG 04 N0047, Jan. 2021.

[13] T .Tan, M. Mrak, V. Baronicini, N. Ramzan, “Report on HEVC compression performance verification testing”, JCTVC-Q1011, April 2014.

[14] Gisle Bjøntegaard, “Improvements of the BD-PSNR model”, ITU-T SG16/Q6, 35th VCEG Meeting, Berlin, Germany, July, 2008, Doc.VCEG-AI11.

[15] International Telecommunication Union – Radio Communication Sector, Recommendation ITU-R BT.2100-2: Image parameter values for high dynamic range television for use in production and international programme exchange.

[16] “Test Model Description of Essential Video Coding”, ISO/IEC JTC 1/SC 29/WG 04 N0127, July 2021. [Online]. Available: https://www.mpegstandards.org/wp-content/uploads/mpeg\_meetings/135\_OnLine/w20741.zip

[17] “Common Test Conditions for Essential Video Coding”, ISO/IEC JTC 1/SC 29/WG 04 N0126, July 2021. [Online]. Available: https://www.mpegstandards.org/wp-content/uploads/mpeg\_meetings/135\_OnLine/w20739.zip

[18] Essential Video Coding Test Model. [Online]. Available: https://github.com/MPEGGroup/MPEG-EVC-ETM

[19] Conformance bitstreams for Essential Video Coding. [Online]. Available: https://standards.iso.org/iso-iec/23094/-4/ed-1/en/

[20] eXtra-fast Essential Video Encoder (XEVE). [Online]. Available: https://github.com/mpeg5/xeve

[21] eXtra-fast Essential Video Decoder (XEVD), [Online]. Available: https://github.com/mpeg5/xevd