|  |
| --- |
| **INTERNATIONAL ORGANIZATION FOR STANDARDIZATION ORGANISATION INTERNATIONALE DE NORMALISATION ISO/IEC JTC 1/SC 29/WG 5** **MPEG JOINT VIDEO EXPERTS TEAM WITH ITU-T SG21** |
| **ISO/IEC JTC 1 / SC 29 / WG 5 N 388** |
| **Virtual, 14–23 January 2026** |
| |  |  | | --- | --- | | **Title:** | **Draft Joint Call for Proposals on video compression with capability  beyond VVC** | | **Source:** | **Convenor (Jens-Rainer Ohm)** | | **Type:** | **General** | | **Subtype:** | **N/A** | | **Status:** | **Approved** | | **Date:** | **2026-02-05** | | **Expected Action:** | **Info** | | **Action due date:** | **N/A** | | **No. of pages** | **20** (without this cover page) | | **Email of convenor:** | **ohm @ ient . rwth-aachen . de** | | **Committee URL:** | **https://sd.iso.org/documents/ui/#!/browse/iso/iso-iec-jtc-1/iso-iec-jtc-1-sc-29/iso-iec-jtc-1-sc-29-wg-5** | |

|  |  |
| --- | --- |
| **Joint Video Experts Team (JVET)**  **of ITU-T SG21 WP3/21 and ISO/IEC JTC 1/SC 29**  41st Meeting, by teleconference, 14–23 January 2026 | Document: JVET-AO2026-v1 |

****

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Draft Joint Call for Proposals on Video Compression with Capability beyond VVC** | | |
| *Status:* | Approved | | |
| *Purpose:* | Draft CfP | | |
| *Author(s) or Contact(s):* | Jens-Rainer Ohm  Mathias Wien  Frank Bossen | Email: | [ohm@ient.rwth-aachen.de](mailto:ohm@ient.rwth-aachen.de)  [wien@lfb.rwth.aachen.de](mailto:wien@lfb.rwth.aachen.de)  [frank@bossentech.com](mailto:frank@bossentech.com) |
| *Source:* | JVET | | |

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

# Introduction

This document is a draft Call for Proposals on video coding technology with capabilities that significantly exceed those of the Versatile Video Coding (VVC) standard (Rec. ITU-T H.266 | ISO/IEC 23090-3) and its current extensions. This draft Call for Proposals (CfP) has been issued by the Joint Video Experts Team (JVET) on behalf of its ITU-T and ISO/IEC parent bodies (ITU-T SG21 WP3/21 and ISO/IEC JTC 1/SC 29). It is planned to issue a final draft Call for Proposals shortly after the April 2026 meeting of JVET. The Call for Proposals will be issued at the July 2026 meeting. The evaluation of submissions is planned to be conducted during a meeting in January 2027.

# Purpose and procedure

A new generation of video compression technology that substantially exceeds the VVC standard in terms of compression, encoder/decoder implementability, applicability to a variety of content, and other features such as transmission latency/robustness, scalability, etc. is targeted.

Example sources include camera-captured content, computer-generated content, user-generated content, and high dynamic range content, while example applications include broadcast (both live or pre-authored content), real-time video conferencing, video chat, on-demand viewing, gaming, video upload, storage-based media replay, and surveillance with fixed or moving cameras. More background information, as well as information about applications and requirements, is given in [1].

The CfP requests submission of video compression technology that has compression performance or additional functionality beyond that of VVC, where the trade-off in terms of implementation cost is an important criterion as well. As it is recognized that in a growing number of applications practical fast encoding is required, information on this aspect is also separately requested in the context of the CfP. As it cannot be expected that encoding with lower run time comes without a penalty in compression, this will be tested separately.

Multiple test cases are defined: one test case for improved compression at large and three test cases for improved compression with runtime-constrained encoding. For each test case, test sequences are organized in test categories which represent a content type and an associated encoding configuration. Submitters are encouraged (but not required) to submit results for all test cases. However, submitters are required to provide complete results for all categories in a given test case[[1]](#footnote-1).

Responses to the CfP will be evaluated at the 45th JVET meeting in January 2027, as further described below.

Companies and organizations who have developed compression technology that they believe to be superior to the Main 10 Profile of the VVC standard regarding the aforementioned aspects are kindly invited to propose such technology to the JVET in response to this CfP. Additionally, contributions are also welcome regarding technology that better supports newly emerging application areas of video coding by additional functionality (see section 6 of this document).

To evaluate the proposed compression technologies, formal subjective tests will be conducted. Results of these tests will be made public (although no direct identification of the proponents will be made in the report of the results unless it is specifically requested or authorized by a proponent to be explicitly identified and a consensus is established to do so). Prior to having evaluated the results of the tests, no commitment to any course of action regarding the proposed technology can be made.

Descriptions of proposals shall be registered as JVET input documents to the proposal evaluation meeting of January 2027 (see the timeline in section 3). Proponents also need to attend this meeting to present their proposals. Further information about logistical steps to attend the meeting can be obtained from the listed contact persons (see section 11).

# Timeline

The timeline of the Call for Proposals is as follows (it is noted that dates for preparation milestones may change in final call):

2026-05-15: Final Draft Call for Proposals[[2]](#footnote-2)

2026-05-31: VTM anchors, runtime-constrained VTM encodings, additional VTM encodings with   
 RPR enabled, and VTM encoder configuration information available

2026-07-17: Call for Proposals

2026-08-01: Formal registration period opens.

2026-09-01: Formal registration period ends.

2026-09-07: Final fee (see section 9) is determined and a formal offer for conduction of the tests is sent by the test coordinator.

2026-10-26: Coded test material shall be available at the test site[[3]](#footnote-3). By this date, confirmation of the purchase order shall be received.

2026-11-02 Subjective assessment starts (planned to be finalized by 12/21).

2026-11-02: Set of additional sequences provided to the proponents.

2026-11-30: Start of cross-checking of bitstreams and binary decoders (participation mandatory for proponents, coordinated by the JVET chair and test coordinator).

2026-12-21: Submission of bitstreams and associated materials for the set of additional sequences.

2027-01-06: Registration and submission of documents describing the proposals[[4]](#footnote-4).

2027-01-06: Reporting of results about cross-checking of bitstreams and binary decoders to the JVET chair and test coordinator

2027-01-13: Subjective test results made available to proponents and within standardization body

2027-01-13/22: Evaluation of proposals at JVET meeting[[5]](#footnote-5)

Anticipated tentative timeline after CfP:

(referring to the first version of an anticipated new standard, which may be extended later in subsequent work):

2027-01 Test model selection process begins by starting to conduct Core Experiments

2027-10 Test model selection established

2029-10 Final standard completed

# Test case on improved compression

The following categories are considered. Each category name includes a characterization of the content type, the suggested encoder configuration and an indication of the resolution.

* **SDR RA UHD/4K**: Representing the use case of distribution of standard dynamic range UHD/4K video content e.g. in a streaming scenario, using a random-access configuration.
* **SDR RA HD**: Representing the use case of distribution of standard dynamic range HD video content e.g. in a streaming scenario, using a random-access configuration.
* **SDR LB HD**: Representing the use case of conversational and other low delay applications at HD resolution, correspondingly using a low-delay configuration.
* **HDR-PQ RA UHD**: Representing the use case of distribution of high dynamic range UHD/4K/8K video content using the PQ transfer function e.g. in a streaming scenario, using a random-access configuration. In order to reduce the encoding workload for assessment of this category and allow investigation on 4K displays, cropped regions of 3840×2160 resolution are used for 4K and 8K content.
* **HDR-HLG RA UHD**: Representing the use case of distribution of high dynamic range UHD/4K/8K video content using the HLG transfer function e.g. in a streaming scenario, using a random-access configuration. In order to reduce the encoding workload for assessment of this category and allow investigation on 4K displays, cropped regions of 3840×2160 resolution are used for 4K and 8K content.
* **Gaming LB HD**: Representing the use case of online gaming with a low-delay configuration.
* **UGC RA**: Representing the use case of user generated content at 1080×1920 or 1920x1080 resolution using a random-access configuration.

## General coding conditions

The following rules apply:

* Encoded bitstreams shall not exceed the target bit rates defined for each of the sequences.
* Sequences are expected to be encoded at full input resolution for all rate points. If coding at reduced resolution is part of the algorithmic concept, it shall be described[[6]](#footnote-6).
* Quantization settings should be kept static. When quantization parameters or more generally rate-distortion parameters (e.g., lambda) change from frame to frame (e.g., QP offset dependent on temporal ID) or within a frame, it shall be described.
* Notwithstanding the previous rule, a one-time change of the quantization settings to meet the target bit rate is allowed and must be documented.
* Optimizing encoding parameters using non-automatic means as well as optimizing encoding parameters on a per sequence basis within a category is discouraged. When any such optimization is employed, it shall be described.
* No part of test sequences shall be used as a training set or part of a training set for training entropy coding tables, VQ codebooks, transforms, predictors, filters, neural network models, etc. It is furthermore mandatory to report which material was used to train corresponding parts of the algorithm.
* Use of preprocessing and/or postprocessing shall be described.
* Use of perceptual optimization shall be described.
* Use of multi-pass encoding shall be described.

If any methods are applied in preprocessing, postprocessing or encoder optimization, it is encouraged to provide also results when applying equivalent methods to the anchor. Such information may also be requested to be provided at the evaluation meeting in January 2027, or during subsequent investigations.

For the test cases in the tables of sections 4.2-4.8, only rates 1-4 will be included in the subjective visual quality tests. The points at higher rates will only be used to investigate the gain based on objective metrics over a broader range of rate/quality.

A set of additional sequences and associated rate/quality ranges will be announced to the proponents after the submission of decoder binaries and bitstreams for the sequences listed in the tables of sections 4.2-4.8. The sequences in this set will not be included in the formal subjective evaluation conducted for the proposals. It is expected that the amount of content to be processed will be 50% or less compared to all the test sequences described in this section (similar test categories, i.e., similar resolutions, structural delay constraints, bitrates and content type). Bitstreams submitted for the additional sequences will be required to stay within 80-100% of the defined target bitrates. Bitstreams and results for this additional set shall be provided prior to the evaluation meeting in January 2027. Those bitstreams are expected to be decodable using the decoder binary previously submitted.

It is to be noted that test sequences and target bitrates currently listed in the test categories may be subject to change until the Final Draft Call is issued.

## SDR RA UHD/4K

### Sequence formats and target bit rates

All sequences have a picture size of 3840×2160, chroma format 4:2:0 YCbCr according to ITU-R BT.2020, and bit depth 10. Further information can be found in Table 1.

Table 1: Test sequences

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Sequence ID (SID)** | **Sequence name** | **Bit depth** | **Frame count** | **Frame rate** | **MD5Sum** |
| SRU1 | CrowdRun | 10 | 500 | 50 | 17282413c9a523a02e0c8463c097221f |
| SRU2 | DrivingPOV3 | 10 | 600 | 60 | e81b65724c4235128b2749ccb3b0fb4a |
| SRU3 | FireDance | 10 | 250 | 25 | bfcfc333924835aaf4ed89efa8428a36 |
| SRU4 | HallwayScene | 10 | 250 | 25 | b0c8718579998060e375fcdb77243013 |

Target bit rates for the encoding of the sequences are listed in Table 2.

Table 2: Target bit rates

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | | **Target bit rates [kbit/s]** | | | | |
| **SID** | **Sequences name** | **Rate 1** | **Rate 2** | **Rate 3** | **Rate 4** | **Rate 5** |
| SRU1 | CrowdRun | 700 | 1500 | 3200 | 7000 | 14000 |
| SRU2 | DrivingPOV3 | 300 | 600 | 1200 | 2400 | 4800 |
| SRU3 | FireDance | 300 | 700 | 1300 | 2500 | 5000 |
| SRU4 | HallwayScene | 100 | 250 | 500 | 1000 | 2000 |

### Coding conditions for VVC anchors

The general description in section 4.1 applies.

Specifically, in this test category, a random-access scenario (RA) is used for evaluation and follows the encoder settings in JVET common test conditions and software reference configurations [3]. The intra refresh period is dependent on the frame rate of the source: a value 32 shall be used for sequences with a frame rate equal to 24fps, 25fps or 30fps, and a value 64 for 50fps and 60fps.

### Coding conditions for submissions

Submissions to the Call for Proposals shall obey the following rules:

* The general rules as described in section 4.1.
* Allow for random access at intervals not larger than the intra refresh period of the respective anchor.

## SDR RA HD

### Sequence formats and frame rates

All sequences have a picture size of 1920×1080, chroma format 4:2:0 YCbCr according to ITU-R BT.709. Further information can be found in Table 3.

Table 3: Test sequences

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Sequence ID (SID)** | **Sequence name** | **Bit depth** | **Frame count** | **Frame rate** | **MD5Sum** |
| SRH1 | DucksTakeOff | 8 | 500 | 50 | 3504582275e36e310fae72ffa6a0f8aa |
| SRH2 | DrivingPOV4 | 10 | 600 | 60 | 4724c33b6b27669a406587bdb8d43dc8 |
| SRH3 | Seeking | 8 | 500 | 50 | 0f26ad79d085895299860530e145a8c9 |
| SRH4 | Umbrella | 8 | 500 | 50 | 87572379ef39bedf55f9c9d8152086e7 |

Target bit rates for the encoding of the sequences are listed in Table 4.

Table 4: Target bit rates

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | | **Target bit rates [kbit/s]** | | | | |
| **SID** | **Sequences name** | **Rate 1** | **Rate 2** | **Rate 3** | **Rate 4** | **Rate 5** |
| SRH1 | DucksTakeOff | 300 | 900 | 2400 | 4000 | 8000 |
| SRH2 | DrivingPOV4 | 100 | 250 | 550 | 1400 | 2800 |
| SRH3 | Seeking | 200 | 400 | 800 | 1600 | 3200 |
| SRH4 | Umbrella | 300 | 600 | 1400 | 3500 | 7000 |

### Coding conditions for VVC anchors

The description in Section 4.2.2 applies.

### Coding conditions for submissions

The description in Section 4.2.3 applies.

## SDR LB HD

### Sequence formats and frame rates

All sequences have a picture size of 1920×1080 (landscape orientation, ***L***) or 1080×1920 (portrait orientation ***P***), chroma format 4:2:0 YCbCr according to ITU-R BT.709. Further information can be found in Table 5.

Table 5: Test sequences

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Sequence ID (SID)** | **Sequence name** | **Bit depth** | **Frame count** | **Frame rate** | **MD5Sum** |
| SLH1 | Beatriz ***L*** | 8 | 500 | 50 | fe74cd5046fa033b4f743f42b29e69cd |
| SLH2 | GregoryCactus2 ***P*** | 10 | 300 | 30 | 4fb703c30a3fb06bacef461c69908b93 |
| SLH3 | GregoryScarf2 ***P*** | 10 | 300 | 30 | 7608f5bf81b4f3f11d0c7a489a5ff26c |
| SLH4 | OfficeWalkAtWall ***L*** | 8 | 300 | 30 | 529c15491ea8e1eb0320244a6ff902bb |

Target bit rates for the encoding of the sequences are listed in Table 6.

Table 6: Target bit rates

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | | **Target bit rates [kbit/s]** | | | | |
| **SID** | **Sequences name** | **Rate 1** | **Rate 2** | **Rate 3** | **Rate 4** | **Rate 5** |
| SLH1 | Beatriz | 50 | 100 | 200 | 400 | 800 |
| SLH2 | GregoryCactus2 | 200 | 600 | 1500 | 4000 | 8000 |
| SLH3 | GregoryScarf2 | 200 | 600 | 1800 | 5000 | 10000 |
| SLH4 | OfficeWalkAtWall | 50 | 150 | 350 | 900 | 1800 |

### Coding conditions for VVC anchors

The general description in section 4.1 applies.

Specifically, in this test category, a low-delay scenario with B pictures (LB) is used for evaluation and follows the encoder settings in JVET common test conditions and software reference configurations [3]. No picture reordering is applied between decoder processing and output.

### Coding conditions for submissions

Submissions to the Call for Proposals shall obey the following rules:

* The general rules as described in section 4.1.
* Overall structural delay shall not be larger than that of the respective anchor.

## HDR-PQ RA UHD

### Sequence formats and frame rates

UHD, 4K and 8K HDR PQ test sequences with chroma format 4:2:0 YCbCr, bit depth 10bit, according to ITU-R BT.2100, with the applicable transfer function indicated. Sequences of 4K or 8K resolution have been cropped to a picture size of 3840×2160. Further information can be found in Table 7.

Table 7: Test sequences

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Sequence ID (SID)** | **Sequence name** | **Transfer fct.** | **Frame count** | **Frame rate** | **MD5Sum** |
| HPQ1 | ChandelierCropBR | HDR10 PQ | 360 | 60 | bb42a5d627330c652f19aca4b184b47f |
| HPQ2 | FashionLadyCrop1 | HDR10 PQ | 380 | 60 | d8546edaa260468f027c0e1c28d422b0 |
| HPQ3 | MeridianHDR2 | P3 PQ 4000nits | 600 | 60 | 9e889c7d78b1b5a0fdebb88871a362dd |
| HPQ4 | NeptuneFountain3R1 | HDR10 PQ | 600 | 60 | 4d7aa4113ddbf6de8a3a9856168f6d5c |
| HPQ5 | SparksWelding | HDR10 PQ 1000nits | 600 | 60 | 7e0c6bd867c370dde9b11c76ed93409f |

Target bit rates for the encoding of the sequences are listed in Table 8.

Table 8: Target bit rates

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | | **Target bit rates [kbit/s]** | | | | |
| **SID** | **Sequences name** | **Rate 1** | **Rate 2** | **Rate 3** | **Rate 4** | **Rate 5** |
| HPQ1 | ChandelierCropBR | 300 | 650 | 1300 | 2800 | 5600 |
| HPQ2 | FashionLadyCrop1 | 200 | 500 | 1000 | 2000 | 4000 |
| HPQ3 | MeridianHDR2 | 150 | 300 | 600 | 1200 | 2400 |
| HPQ4 | NeptuneFountain3R1 | 350 | 850 | 2000 | 5000 | 10000 |
| HPQ5 | SparksWelding | 400 | 1000 | 2600 | 7000 | 14000 |

### Coding conditions for VVC anchors

In this test category, the random-access scenario is used for evaluation. The description in Section 4.2.2 applies, with the encoder settings in common test conditions for HDR/WCG content being used [4].

### Coding conditions for submissions

The description in Section 4.2.3 applies.

## HDR-HLG RA UHD

### Sequence formats and frame rates

UHD, 4K and 8K HDR HLG test sequences with chroma format 4:2:0 YCbCr, bit depth 10bit, according to ITU-R BT.2100, with the applicable transfer function indicated. Sequences of 4K or 8K resolution have been cropped to a picture size of 3840×2160. Further information can be found in Table 9.

Table 9: Test sequences

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Sequence ID (SID)** | **Sequence name** | **Transfer fct.** | **Frame count** | **Frame rate** | **MD5Sum** |
| HLG1 | AMS06 | HLG10 | 600 | 60 | d1f11c771febbb8a2bbb7faadc13cbbb |
| HLG2 | WaterfallForest | HLG10 | 500 | 50 | 54abb97e78a7255885fc283d02f9c964 |
| HLG3 | WomenFootball | HLG10 | 500 | 50 | 94bb484ab23ac900d1f417c52972535f |

Target bit rates for the encoding of the sequences are listed in Table 10.

Table 10: Target bit rates

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | | **Target bit rates [kbit/s]** | | | | |
| **SID** | **Sequences name** | **Rate 1** | **Rate 2** | **Rate 3** | **Rate 4** | **Rate 5** |
| HLG1 | AMS06 | 600 | 1300 | 3500 | 8000 | 16000 |
| HLG1 | WaterfallForest | 1000 | 2500 | 6000 | 14000 | 28000 |
| HLG2 | WomenFootball | 300 | 600 | 1100 | 2000 | 4000 |

### Coding conditions for VVC anchors

The description in Section 4.5.2 applies.

### Coding conditions for submissions

The description in Section 4.5.3 applies.

## Gaming LB HD

### Sequence formats and frame rates

All sequences have a picture size of 1920×1080, chroma format 4:2:0 YCbCr according to ITU-R BT.709. Further information can be found in Table 11.

Table 11: Test sequences

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Sequence ID (SID)** | **Sequence name** | **Bit depth** | **Frame count** | **Frame rate** | **MD5Sum** |
| GLH1 | DOTA2s360 | 8 | 550 | 60 | 999ae67b022892a7b70ab4dcc70f1043 |
| GLH2 | GTAVs090 | 8 | 600 | 60 | 43643d686a91a51758c63346fab7eca3 |
| GLH3 | Level1 | 10 | 600 | 60 | 87c7055771366d16b0d69cb56f6c66fb |
| GLH4 | Minecraft | 8 | 600 | 60 | 4f1055d558284e789fcae7686f3419bc |

Target bit rates for the encoding of the sequences are listed in Table 12.

Table 12: Target bit rates

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | | **Target bit rates [kbit/s]** | | | | |
| **SID** | **Sequences name** | **Rate 1** | **Rate 2** | **Rate 3** | **Rate 4** | **Rate 5** |
| GLH1 | DOTA2s360 | 180 | 300 | 550 | 1000 | 2000 |
| GLH2 | GTAVs090 | 400 | 900 | 2000 | 3600 | 7200 |
| GLH3 | Level1 | 400 | 1000 | 2000 | 4000 | 8000 |
| GLH4 | Minecraft | 300 | 600 | 1200 | 2400 | 4800 |

### Coding conditions for VVC

The general description in section 4.1 applies.

Specifically, in this test category, a low-delay scenario with B pictures (LB) is used for evaluation and follows the encoder settings in JVET common test conditions and software reference configurations [3]. No picture reordering is applied between decoder processing and output.

### Coding conditions for submissions

The description in Section 4.4.3 applies.

## UGC RA

### Sequence formats and frame rates

All sequences have a picture size of 1920×1080 (landscape orientation, ***L***) or 1080×1920 (portrait orientation ***P***), chroma format 4:2:0 YCbCr according to ITU-R BT.709. Further information can be found in Table 13.

Table 13: Test sequences

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Sequence ID (SID)** | **Sequence name** | **Bit depth** | **Frame count** | **Frame rate** | **MD5Sum** |
| URH1 | Camellia ***P*** | 8 | 600 | 60 | 6e5413666e42d3c24db9b81b5db641ff |
| URH2 | Hobby-w5xz-backpack ***P*** | 8 | 240 | 24 | b8ee8e98b5d34afcfc24a9d02d8f3a43 |
| URH3 | NightLandscapeRunning ***L*** | 8 | 300 | 30 | 8eed52f009ed4ca2d31abe9e61b30432 |
| URH4 | Sports-76a2-iceball ***L*** | 8 | 600 | 60 | 3030f3760f34ce97e67d7e1a02eedd3b |
| URH5 | VerticalVideo-3709-snow ***P*** | 8 | 300 | 30 | 3db32dc2c295dc25622ebc516420c63b |
| URH6 | VerticalVideo-3d96-walk ***P*** | 8 | 300 | 30 | 5eb532977244f8be48591bf06592579a |

Target bit rates for the encoding of the sequences are listed in Table 14.

Table 14: Target bit rates

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | | **Target bit rates [kbit/s]** | | | | |
| **SID** | **Sequences name** | **Rate 1** | **Rate 2** | **Rate 3** | **Rate 4** | **Rate 5** |
| URH1 | Camellia | 250 | 500 | 1000 | 2000 | 4000 |
| URH2 | Hobby-w5xz-backpack | 90 | 160 | 280 | 500 | 1000 |
| URH3 | NightLandscapeRunning | 100 | 250 | 600 | 1200 | 2400 |
| URH4 | Sports-76a2-iceball | 80 | 160 | 250 | 400 | 800 |
| URH5 | VerticalVideo-3709-snow | 80 | 160 | 300 | 500 | 1000 |
| URH6 | VerticalVideo-3d96-walk | 80 | 160 | 300 | 600 | 1200 |

### Coding conditions for VVC anchors

The description in Section 4.2.2 applies.

### Coding conditions for submissions

The description in Section 4.2.3 applies.

# Test cases on improved compression with runtime-constrained encoding

The same categories and sequences as defined in sections 4.2 – 4.8 are considered.

Submitters are strongly encouraged (but not required) to submit results for all three test cases. However, submitters are required to provide results for all sequences in a given test case.

## Sequence formats and frame rates

For each category, sequence formats and frame rates are identical to those defined in section 4.

## Coding conditions for anchors

For each category, the coding conditions for the anchor are identical to those defined in the corresponding part of section 4, except as noted below.

For VVC, bitstreams are generated for the following VTM configurations documented in [3][[7]](#footnote-7):

* Default (equivalent to the one used in section 4)
* High performance (encoding times are approximately 2 times those of the default)
* Reduced encoding time (3 variants, encoding times are approximately 0.2 to 0.75 times those of the default)

VTM with default configurations is the anchor with respect to which relative compression performance and run times are calculated.

## Coding conditions for submissions

Submissions to the Call for Proposals shall obey the following rules:

* For each category, the rules defined in the corresponding part of section 4 apply.
* Encoder run times must be in the range of targets defined below and should match the runtime targets as closely as possible.
* No target is defined for decoder run times. However, decoder run times are to be reported and may be considered in the evaluation of this Call.
* Code optimization beyond what is typically done for reference software such as VTM is discouraged. The level of code optimization shall be described.

Three encoder run time targets are defined as:

* 5 times (500%) the encoder run time of the VTM anchor
* Same (100%) as the encoder run time of the VTM anchor
* 0.2 times (20%) the encoder run time of the VTM anchor

For each bitstream, run time is defined to be the total processing time to generate the bitstream and includes processing time for any pre-processing operations. Where multithreading is used, the run time is the sum of the run times of each thread. Design parts affected by multithreading must be described.

The run time target is defined for the aggregate run time across all bitstreams. It is not required to meet to the target run time for each bitstream. It is discouraged to tune encoder configurations for each bitstream such as to meet the target run time for every bitstream.

For the purpose of matching run time targets, the aggregate run time across multiple sequences and rate points shall be computed as follows: 1) run times for all rate points are summed for each sequence; 2) the geometric mean is computed across all sequences; 3) the ratio to anchor runtime is computed. Alternate ways to aggregate run times may be additionally used when evaluating responses. For example, run times may be aggregated across a single rate point (e.g., the highest rate point).

Matching the run time targets exactly is not necessary. Run times should be sufficiently close to the target to be able to plot an encoder run time vs compression performance curve for each submission and have this curve cover a range similar to the one spanned by the three run time targets defined above.

It is strongly encouraged to include a set of bitstreams and results for each of the three run time targets in a response. Additional run time vs compression trade-offs may be included in a response such as to have a more densely sampled curve. For example, results for run times approximately 0.5 times or 2 times those of the VTM anchor may be included.

Depending on the number of submissions, it may be necessary to select a subset of representative video sequences for the subjective assessment if otherwise it would not be possible to complete such assessment within the timeline foreseen for the subjective tests. This subset would be selected by the JVET chair and the JVET test coordinator to properly represent as many test categories as possible. Assessment based on objective metrics will be done for all submitted sequences.

# Technology supporting other functionality

Companies and organizations who have developed compression technology that provides additional functionality with a better compression trade-off than existing standards are also invited to submit such information in the context of this Call for Proposals. Proposals on how the benefit of such functionality could be assessed during the development of a video compression standard would also be welcome. For examples of such additional functionality, you may refer to [1].

Submissions to this section will not be part of the formal subjective test. Potential expert viewing or informal demonstration could be conducted at a JVET meeting.]

# Test sites and delivery of test material

The proposals submission material will be evaluated by means of a formal subjective assessment process. The tests will be conducted by the Test Coordinator and one or more additional sites. The names of the sites will be provided in the Final Call for Proposals.

All proponents need to deliver, by the due date of 2026-10-26, an SSD to the address of the Test Coordinator (see section 11; for the content of this disk, see section 10A).

The correct reception of the material will be confirmed by the Test Coordinator. Any inconvenience caused by unexpected delivery delay or a failure of the disk will be under the complete responsibility of the proponents, but solutions will be negotiated to ensure that the data can still be included in the test if feasible, which means that correct and complete data need to be available before the beginning of the test at the latest.

All the bitstreams, the decoder executable, and the YUV files shall be accompanied by an MD5 check-sum file to verify their correct storing on the disk.

Further technical details on the delivery of the coded material are provided by the test coordinator.

# Testing fee

A testing fee will be charged for each response to this Call for Proposals. The purpose of the testing fee is to cover costs to conduct the subjective assessment which requires the hiring of subjects. The amount of the fee will be determined when the number of registered proposals is known. It is expected to be in the order of EUR 10-20k for one full set of categories with all sequences and rate points. A maximum fee for proposals submitting to all test cases will be determined in the Final Draft CfP document.

# Requirements for submissions

More information about file formats can be found in Annex A. Files of decoded sequences and bitstreams shall follow the naming conventions as specified in Annex A.

Proponents shall provide the following; incomplete proposals may not be considered:

1. Coded test material submission to be received by 2026-10-26[[8]](#footnote-8):
2. Bitstreams for all submitted test cases.
3. Binary decoder and encoder executables including relevant encoder configuration settings. The decoder must be capable of decoding the bitstreams and storing the decoded data in 10-bit YUV 4:2:0 uncompressed format as needed for the subjective evaluation. Executables should be built to run on Ubuntu 24.04 (x86-64)[[9]](#footnote-9). A single decoder executable shall be provided for all test cases. A decoder executable should require only two input parameters to reconstruct the YUV file from a bitstream file, for example: “<decoder\_exe> -b <bitstream\_filename> -o <yuv\_filename>”.
4. Decoded sequences (YUV files 4:2:0 uncompressed format) for all submitted test cases.
5. For evaluation by objective metrics, CSV files containing the values listed in annex D.
6. MD5 checksums for all submitted files. Reporting these checksums in one aggregated files is encouraged.
7. Document to be submitted before the evaluation meeting in 2027/01 shall contain[[10]](#footnote-10)
8. A technical description of the proposal sufficient for full conceptual understanding and generation of equivalent performance results by experts and for conveying the degree of optimization required to replicate the performance. This description should include all data processing paths and individual data processing components used to generate the bitstreams. It does not need to include complete bitstream format or implementation details, although as much detail as possible is desired.
9. csv files for all submitted test cases as specified in annex D.
10. The technical description shall also contain a statement about the programming language in which the software is written, e.g. C/C++, possible need for external libraries and platforms on which the binaries were compiled.
11. The technical description shall state how the proposed technology behaves in terms of random access to any picture within the sequence. For example, a description of the GOP structure and the maximum number of pictures that must be decoded to access any picture could be given.
12. The technical description shall specify the expected encoding and decoding delay characteristics of the technology, including structural delay e.g. due to the amount of picture reordering and DPB buffering, the degree of picture-level multi-pass decisions and the degree by which the delay can be minimized by parallel processing.
13. The technical description shall contain information suitable to assess the complexity of the implementation of the technology:

* Encoding and decoding run time relative to the VTM anchors, measured using the same simulation environment. Proponents shall provide a description of the platform and methodology used to determine the time. To help interpretation, a description of software optimisations undertaken, if any, is encouraged.
* Complexity characteristics of encoder and decoder by filling the complexity reporting template [8].
* Degree of capability for parallel processing.

1. Furthermore, the technical description should point out any specific properties of the proposal (e.g., additional functionality such as benefit for 4:4:4 coding, error resilience, scalability).
2. Optional information

Proponents are encouraged (but not required) to allow members of JVET to have access, on a temporary or permanent basis, to their encoded bitstreams and binary executables or source code.

Any additional information about properties and performance of the algorithm or specific tools is welcome.

# Subsequent provision of source code and IPR considerations

Proponents are advised that, upon acceptance for further evaluation, it will be required that certain parts of any technology proposed be made available in source code format to participants in the core experiments process and for potential inclusion in the prospective standard as reference software. When a particular technology is a candidate for further evaluation, commitment to provide such software is a condition of participation. The software shall produce identical results to those submitted to the test. Additionally, submission of improvements (bug fixes, etc.) is certainly encouraged.

In the above, “certain parts of any technology proposed” may include training scripts or equations used to derive parameters of a tool.

Furthermore, proponents are advised that this Call is being made subject to the common patent policy of ITU-T/ITU-R/ISO/IEC (see http://[www.itu.int/ITU-T/dbase/patent/patent-policy.html](http://www.itu.int/ITU-T/dbase/patent/patent-policy.html) or [ISO/IEC Directives Part 1](http://isotc.iso.org/livelink/livelink?func=ll&objId=4230455&objAction=browse&sort=subtype), Appendix I) and the other established policies of the standardization organizations.

# Contacts

Prospective contributors of responses to the Call for Proposals should contact the following people:

Jens-Rainer Ohm (JVET chair)

RWTH Aachen University, Institute of Communications Engineering

Melatener Str. 23, 52074 Aachen, Germany

Tel. +49-241-8027671, email [ohm@ient.rwth-aachen.de](mailto:ohm@ient.rwth-aachen.de)

Mathias Wien (JVET test coordinator)

RWTH Aachen University, Institute of Imaging and Computer Vision

Kopernikusstr. 16, 52074 Aachen, Germany

Tel. +49-241-8027867, email [wien@lfb.rwth-aachen.de](mailto:wien@lfb.rwth-aachen.de)

Registration of a response shall be made by submitting the registration form in Annex E to the persons above on or before September 1, 2026. Interested parties are kindly invited to express their intent as early as possible.

Details on how to format and submit documents, bitstreams, and other required data will be communicated directly to those who register a response. Additionally, the JVET chair will provide assistance to submitters from outside JVET in order for them to attend the JVET meeting.

Test sequences, anchors, and configuration information will also be made available by contacting the persons above (see section 3 for availability dates).

# References

1. “Draft of use cases and requirements for potential next-generation video coding standard beyond VVC capability”, ISO/IEC SC 29/WG 2 document [N 448](https://www.mpeg.org/wp-content/uploads/mpeg_meetings/150_OnLine/w25071.zip) and ITU-T Q6/21 document [VCEG-BX25](https://www.itu.int/wftp3/av-arch/video-site/2503_Tel/VCEG-BX25-v1-NGVC-Use-Reqs.docx), April 2025
2. “Algorithm description for Versatile Video Coding and Test Model 23 (VTM 23)”, Joint Video Experts Team (JVET), 41st Meeting, Virtual, January 2025, Doc. [JVET-AO2002](https://jvet-experts.org/doc_end_user/current_document.php?id=16691).
3. “VTM and HM common test conditions and software reference configurations for SDR 4:2:0 10 bit video”, Joint Video Experts Team (JVET), 38th Meeting, by teleconference, March/April 2025, Doc. [JVET-AL2010](https://jvet-experts.org/doc_end_user/current_document.php?id=15680).
4. “VTM and HM common test conditions and evaluation procedures for HDR/WCG video”, Joint Video Experts Team (JVET), 41st Meeting, by teleconference, January 2026, Doc. [JVET-AO2011](https://jvet-experts.org/doc_end_user/current_document.php?id=16692).
5. ITU-T, Working practices using objective metrics for evaluation of video coding efficiency experiments, Doc. [HSTP-VID-WPOM](https://www.itu.int/dms_pub/itu-t/opb/tut/T-TUT-ASC-2020-HSTP1-PDF-E.pdf), 2020.
6. International Telecommunication Union – Radiocommunication Sector, *Methodology for the subjective assessment of the quality of television pictures*, Recommendation ITU-R BT.500-15 (available at <https://www.itu.int/rec/R-REC-BT.500>).
7. International Telecommunication Union – Telecommunication Standardization Sector, Subjective video quality assessment methods for multimedia applications, Recommendation ITU-T P.910 (available at <https://www.itu.int/rec/t-rec-p.910>).
8. “Complexity reporting template for coding algorithms and tools”, 41st Meeting, by teleconference, January 2026, Doc. [JVET-AO2040](https://jvet-experts.org/doc_end_user/current_document.php?id=16698)
9. Distribution formats for test sequences, decoded results, and bitstreams

Distribution of original video material files containing test sequences is done in YUV files with extension “.yuv”. A description of the YUV file format is available at <http://www.fourcc.org/> web site, designated as “yuv420p10le iyuv”. [Note: .pyuv instead of .yuv files might be used, to be clarified in the Final Draft CfP]

VVC Anchor bitstreams are provided with extension “.bit”. Bitstream formats of proposals can be proprietary, but must contain all information necessary to decode the sequences at a given data rate (e.g. no additional parameter files). The file extension of a proposal bitstream shall be “.bit”.

Decoded sequences shall be provided in the same “.yuv” format as originals, with the exception that the colour depth shall be 10 bits per component for all sequences. The same format shall be output by the binary decoder executable.

Filenames must follow the following format: xxxx\_Pyy\_Rz\_Cw.eee where

* xxxx is the sequence ID, as defined in section 4
* yy is a proponent ID to be provided upon registration, the value 00 is reserved for the VVC anchor
* z is a rate target (1 to 5)
* w is a test case (0 = improved compression, 1 = runtime constraint 5x, 2 = runtime constraint 1x, 3 = runtime constraint 0.2x, 4.. = additional runtime constraints (optional))
* eee is an extension (yuv for decoded YUV files, bit for bitstream files)

1. Description of testing environment and methodology

The test method adopted for this evaluation is degradation category rating (DCR) [7].

* 1. Degradation Category Rating (DCR)

The degradation category rating (DCR) method presents stimuli in pairs. The first stimulus presented in each pair is always the reference. The second stimulus is that reference stimulus after processing by the systems under test. DCR is a double stimulus method. The DCR method is also known as the double stimulus impairment scale (DSIS) method defined in ITU-R BT.500 [6].

An impairment rating scale made of 11 levels is planned to be used, ranging from “0” (severely annoying) to “10”

The structure of the basic test cell (BTC) of the DCR method includes the presentation of the uncompressed reference of the video sequences followed by the Processed Video Sequence (PVS) under evaluation. Then a message displays for 5 seconds asking the viewers to vote. The presentation of the video clips is preceded by a mid-grey screen displaying e.g. “Source” for the original and “Test” for the coded version of the sequence under test for one second. During the voting time, the number of the vote is indicated on the screen.



**Figure B.1 – Structure of a DCR BTC**

* + 1. Training and stabilization phase

The outcome of a test is highly dependent on a proper training of the test subjects.

For this purpose, each subject has to be trained by means of a short practice (training) session demonstrating the range of qualities to be expected in the test.

The stabilization phase uses the test material of a test session, e.g. three BTCs containing one sample of best quality, one of the worst qualities and one of medium quality, are presented at the beginning of the test session. By this way, the test subjects have an immediate impression of the quality range they are expected to evaluate during that session.

The scores of the stabilization phase are discarded for the evaluation.

* 1. The laboratory setup

The laboratories for subjective assessments are supposed to comply with established recommendations [6][7]. Play-out of the video sequences must be done at native resolution.

When a video sequence is shown at a resolution lower than the native resolution of the display itself, the video has to be presented in the center of the display; the active part of the display (i.e. that is actually showing the video signal) must have a dimension equal in rows and columns to the raster of the video; the remaining part of the screen has to be set to a mid-grey level (e.g., 128 in 0-255 range). This constraint guarantees that no interpolation or distortion artefacts of the video images will be introduced.

The video play server, or the PC used to play video has to be able to support the display of both HD and UHD video formats, at 24, 25, 30, 50 and 60 frames per second, without any limitation, or without introducing any additional temporal or visual artefacts.

* + 1. Viewing environment

The viewing distance is set to 1.5H, where H is equal to the height of the active part of the screen, depending on the size of the active part of the screen and its native resolution.

The test rooms are expected to be protected from external visual or audio pollution. Internal general light must be low (just enough to allow the viewing subjects to fill out the scoring sheets) and no direct light must be visible to the viewing subjects seated in front of the screen; the light behind the monitor should be dimmed to an intensity and should have a colour tuned as close as possible to D65. No other light source, in particular any light source directed to the screen or creating reflections are permitted.

* 1. Statistical analysis and presentation of the results

The data collected from the score sheets, filled out by the viewing subjects, will be stored in a database for further processing.

For each coding condition, the Mean Opinion Score (MOS) and associated Confidence Interval (CI) values will be computed and reported.

The MOS and CI values will be used to draw graphs. The graphs will be drawn grouping the results for each video test sequence. No graph grouping results from different video sequences will be considered.

1. Using HDRTools to compute distortion metrics

This annex describes the usage of HDRTools to compute distortion metrics. The HDRTools software package can be downloaded from <https://gitlab.com/standards/HDRTools>. Building instructions are provided in the README.md file included in the HDRTools software package.

Version 0.26 of the HDRTools software should be used. Note that Version 0.26 of the HDRMetrics utility may produce different results on different platforms when non-integer frame rates are provided. To ensure consistent results of the HDRMetrics utility across different platforms, use of an integer frame rate input parameter is recommended. The value of the frame rate parameter should not affect results of HDRMetrics computation.

To compute the PSNR, wPSNR, and MS-SSIM metrics using the HDRMetrics utility, the following command is used from the root HDRTools directory:

build/bin/HDRMetrics \  
-f cfg/HDRMetricYUV.cfg \  
-p Input0File=<original.yuv> \  
-p Input1File=<decoded.yuv> \  
-p Input0Height=<height> \  
-p Input0Width=<width> \  
-p Input1Height=<height> \  
-p Input1Width=<width> \

-p Input0BitDepthCmp0=<bitdepth of original.yuv> \  
-p Input0BitDepthCmp1=<bitdepth of original.yuv> \  
-p Input0BitDepthCmp2=<bitdepth of original.yuv> \  
-p Input1BitDepthCmp0=10 \  
-p Input1BitDepthCmp1=10 \  
-p Input1BitDepthCmp2=10 \  
-p NumberOfFrames=<num\_frames> \  
-p EnableWTPSNR=1 \  
-p EnableJVETPSNR=1 \

-p EnableJVETMSSSIM=1 \  
-p WeightTableFile=cfg/hdrTable.txt \  
-p Input0Rate=60 \  
-p Input1Rate=60

-p MaxSampleValue=1023.0

The distortion values that are to be reported (see Annex B) are listed on the output line starting with “D\_Avg”.

1. Data submission format

Data should be submitted in CSV format. For each sequence and each rate point, the following information should be provided, as applicable. When submitting data for runtime-constrained encoding as per section 5, a separate CSV file shall be provided for each target run time. Additionally for each combination of sequence and rate point, a separate text file should be provided with a bit count for each frame in display order (one bit count per text line).

|  |  |
| --- | --- |
| Column name | Description |
| Sequence | Sequence name, e.g., SRU1 |
| Rate point | Rate point, e.g., R1 |
| Bitrate | Bitrate in kbit/s |
| PSNR Y | PSNR metric for Y component (See Annex A and [1]) |
| PSNR U | PSNR metric for U component (See Annex A and [1]) |
| PSNR V | PSNR metric for V component (See Annex A and [1]) |
| SSIM Y | MS-SSIM metric for Y component (See Annex A) |
| SSIM U | MS-SSIM metric for U component (See Annex A) |
| SSIM V | MS-SSIM metric for V component (See Annex A) |
| WPSNR Y | Weighted PSNR metric for Y component (See Annex A). To be reported for sequences in the HPQ category only. |
| WPSNR U | Weighted PSNR metric for U component (See Annex A). To be reported for sequences in the HPQ category only. |
| WPSNR V | Weighted PSNR metric for V component (See Annex A). To be reported for sequences in the HPQ category only. |
| EncT | Encoder run time in seconds as measured on CPU in single-threaded mode |
| DecT | Decoder run time in seconds as measured on CPU in single-threaded mode |
| relative EncT | Encoder run time ratio to VTM anchor (e.g., 1.0 if runtime is equal to anchor) |
| relative DecT | Decoder run time ratio to VTM anchor |
| peak Enc Memory | Peak memory use of encoder in KB (see getProcStatusValue("VmPeak:") in VTM software) |
| peak Dec Memory | Peak memory use of decoder in KB |
| MD5sum | MD5 checksum of the decoded YUV file |

The first line in the CSV file should contain the name of the columns. For example:

Sequence;Rate point;Bitrate;PSNR Y;PSNR U;PSNR V;SSIM Y;SSIM U;SSIM V;WPSNR Y;WPSNR U;WPSNR V;EncT;DecT;relative EncT;relative DecT;peak Enc Memory;peak Dec Memory;MD5sum

1. Registration Form

**Registration of response to the Joint Call for Proposals on video compression with capability beyond VVC**

|  |  |
| --- | --- |
| **Organization:** |  |
| **Contact person:** |  |
| **Contact email:** |  |

## Test case on improved compression: ☐

## Test cases on improved compression with runtime-constrained encoding: ☐

Intended runtime targets: ☐ 0.2x ☐ 1x ☐ 5x ☐ other \_\_\_ (not included in the formal subjective test)

☐ Submission includes a decoder executable runnable under Ubuntu 24.04 (x86-64) to decode sequences from bitstreams (if box not ticked, please contact the test coordinator about usage of a different platform).

☐ Submission includes additional functionality according to Section 6 which requires subjective evaluation (not included in the formal subjective tests).

Remarks:

1. If a proponent intends to submit only for a subset of categories, it is requested to inquire via the contact persons (see sec. 11). [↑](#footnote-ref-1)
2. Test categories and test cases are not expected to be modified after release of the Final Draft CfP. [↑](#footnote-ref-2)
3. People who formally registered will receive instructions regarding how to submit the coded materials. If material is received later, the proposal may be excluded from testing. [↑](#footnote-ref-3)
4. Contact persons will provide information about document submission process. Note that submitted documents will be made publicly available. Exceptions to public availability will be considered on a case-by-case basis upon request by the contributor. [↑](#footnote-ref-4)
5. Proponents are requested to attend this standardization meeting. [↑](#footnote-ref-5)
6. To allow a comparison against variable-resolution coding, rate-matched versions of VTM enabling its tool of GOP-adaptive reference picture resampling [2] will be included in the visual test in addition to VTM anchors. [↑](#footnote-ref-6)
7. Configurations for high performance and reduced encoding time are not described for HDR categories in [4], but can be configured by using equivalent settings as in [3]. [↑](#footnote-ref-7)
8. The same material shall also be brought to the meeting in January 2027. The bitstreams and binary executables will be provided to other parties for crosscheck purposes, without identifying the respective proposing party. [↑](#footnote-ref-8)
9. The test coordinator shall be informed well in advance if this is not possible. [↑](#footnote-ref-9)
10. A template providing more detail on the structure and content of the proposal description document will be provided along with the final CfP. [↑](#footnote-ref-10)